



Crystal Structure of R120G Disease Mutant of Human α B-Crystallin Domain Dimer Shows Closure of a Groove

A. R. Clark, C. E. Naylor, C. Bagn eris, N. H. Keep and C. Slingsby*

Department of Biological Sciences, Crystallography, Institute of Structural and Molecular Biology, Birkbeck College, Malet Street, London WC1E 7HX, UK

Received 15 December 2010;
received in revised form
7 February 2011;
accepted 9 February 2011
Available online
15 February 2011

Edited by I. Wilson

Keywords:

chaperone;
disease biomarker;
myopathy;
proteostasis;
stress response

Small heat shock proteins form large cytosolic assemblies from an “ α -crystallin domain” (ACD) flanked by sequence extensions. Mutation of a conserved arginine in the ACD of several human small heat shock protein family members causes many common inherited diseases of the lens and neuromuscular system. The mutation R120G in α B-crystallin causes myopathy, cardiomyopathy and cataract. We have solved the X-ray structure of the excised ACD dimer of human α B R120G close to physiological pH and compared it with several recently determined wild-type vertebrate ACD dimer structures. Wild-type excised ACD dimers have a deep groove at the interface floored by a flat extended “bottom sheet.” Solid-state NMR studies of large assemblies of full-length α B-crystallin have shown that the groove is blocked in the ACD dimer by curvature of the bottom sheet. The crystal structure of R120G ACD dimer also reveals a closed groove, but here the bottom sheet is flat. Loss of Arg120 results in rearrangement of an extensive array of charged interactions across this interface. His83 and Asp80 on movable arches on either side of the interface close the groove by forming two new salt bridges. The residues involved in this extended set of ionic interactions are conserved in Hsp27, Hsp20, α A- and α B-crystallin sequences. They are not conserved in Hsp22, where mutation of the equivalent of Arg120 causes neuropathy. We speculate that the α B R120G mutation disturbs oligomer dynamics, causing the growth of large soluble oligomers that are toxic to cells by blocking essential processes.

  2011 Elsevier Ltd. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

Introduction

All cells can protect themselves from hostile environmental changes by the stress response system, which includes increasing the levels of heat shock proteins (Hsps).^{1,2} Hsps contribute to proteostasis³ by binding nonnative proteins, and they are integrated into cellular networks that

protect the cell from accumulating damaged proteins.⁴ Small Hsps (sHsps) form large oligomeric assemblies with a high capacity for holding onto nonnative proteins⁵ that are handed over to the ATP-driven chaperone machines for refolding.⁶ sHsps are ubiquitous among the kingdoms of life and within mammalian genomes occur in large families with members differentially expressed in many tissues including muscle, brain and the eye lens.^{7–9} Members of the family of 10 human sHsps (HSPB1–B10) are mainly cytoplasmic.¹⁰ α B-Crystallin (HSPB5), Hsp27 (HSPB1) and Hsp22 (HSPB8) are cytoprotective, with roles in apoptosis and autophagy.^{11,12} A cytoprotective function is a hindrance to health when exploited by cancer cells.¹³ α B-Crystallin behaves as an oncoprotein when expressed in basal breast cancer¹⁴ and blocks

*Corresponding author. E-mail address:

C.Slingsby@mail.cryst.bbk.ac.uk.

Abbreviations used: ACD, α -crystallin domain; AP, antiparallel; Hsp, heat shock protein; sHsp, small heat shock protein; MPD, methyl pentane diol; rTEV, recombinant tobacco etch virus; His, histidine.

apoptosis in highly migratory glioma cells,¹⁵ making it a potential drug target.

α B-Crystallin is constitutively and highly expressed in the eye lens¹⁶ and to a lesser extent in skeletal and heart muscle.⁷ Cells in these tissues are long-lived and highly elongated, making them vulnerable to protein aggregation and deficiencies in waste-disposal processes.¹⁷ Either event will likely cause upregulation of stress proteins, making sHsps, in particular, biomarkers of disease in these tissues.^{11,18,19} In the brain, α B-crystallin is upregulated in neuron-supporting glial cells during stress, for example, in astrocytes subject to blockade of the proteasome.²⁰ In the very early stages of multiple sclerosis, α B-crystallin is upregulated in oligodendrocytes that are adjacent to activated HLA-DR-expressing microglia.²¹ Severe challenges to proteostasis can occur on expression of stress proteins bearing dominant missense mutations as they signal further stress, leading to more dysfunctional protein. Unsurprisingly, sHsp stress proteins are associated with swathes of inherited human diseases, including cataract,²² skeletal myopathies,²³ cardiomyopathies²⁴ and neuropathies.^{25,26}

Small Hsp sequences have in common an “ α -crystallin domain” (ACD) of around 90 amino acids, with most having a short conserved C-terminal extension bearing an I-X-I motif and a longer variable N-terminal extension.²⁷ Kingdom-wide sequence comparisons show subgroups of sHsp sequences between metazoans and nonmetazoans.^{28,29} Attention has focussed on an arginine within the ACD that is conserved across all phyla, as its mutation in different sHSP family members causes several human diseases. The mutation R120G in α B-crystallin causes desmin-related myopathy.³⁰ The equivalent arginine mutation (R116C) in α A-crystallin (HSPB4) causes zonular central nuclear cataracts.³¹ In Hsp27 and Hsp22, the equivalent substitution (R127W and K141N/E, respectively) results in neuromuscular disorders, axonal Charcot-Marie-Tooth disease type 2F and 2L, and distal hereditary motor neuropathy.^{26,32} The sequence conservation of this arginine and the associated pathogenic mutations in the different sHsp sequences highlight that it is performing a general role common to the ACD in many sHsps.

A striking feature of sHsps is their association into large, soluble, globular assemblies with a quaternary organisation that allows a variable number of protomers.⁵ *In vitro*, recombinant α B R120G mutant protein forms soluble oligomers that are larger and more polydisperse than wild type³³ and that are prone to aggregation and precipitation on handling.³⁴ Although solution polydispersity in α -crystallins is advantageous for maintaining eye lens transparency, it is a hindrance for three-dimensional (3D) atomic characterisation.³⁵ X-ray crystallography has revealed high-resolution structures of two well-defined, and hence atypical,

assemblies: a 24-mer from an archaeal thermophile³⁶ and a 12-mer from wheat.³⁷ The ACD is an immunoglobulin-like beta sandwich comprising strands β 2– β 9, and the assembly unit is a dimer. The I-X-I motifs contribute to building dimers into an assembly by docking into pockets on one side of a neighbouring ACD beta sandwich. The structural role of the N-terminal extensions, which emerge from the other side of the beta sandwich, is less clearly defined. Single-particle electron microscopy studies of shell-like 24-mers from archaeal HSP16.5 indicate a buried location,³⁸ whereas in 24-mers from yeast HSP26, 12 N-terminal domains are buried inside the assembly and 12 are on the surface.³⁹ Only 6 N-terminal extensions were seen buried within the double-disc 12-mer in the crystal structure of wheat HSP16.9,³⁷ where they contributed to holding the assembly together.

Recent progress in crystallographic structure determination of vertebrate sHsps has relied on truncation of the flanking sequence extensions. The ACD dimers show a mode of dimerisation that is different from that of nonmetazoans (Fig 1):^{40–42} the fewer residues between strands β 5 and β 7 in metazoan sequences result in the loss of a distinct β 6 strand, and instead, β 7 strand is extended and referred to as β 6+7 strand. This extended strand interacts about the dimer interface with the equivalent strand in the other chain, whereas in nonmetazoans the β 6 strand exchanges between partner chains at the dimer interface. In effect, an interface region on one side of the ACD sandwich has switched from one sheet to the other (Fig. 1). Sequence differences support these distinct modes of dimerisation, in particular the presence of a PG motif in nonmetazoans compared with a KH motif in metazoans (Table 1). The conserved arginine in human sHsps is on the β 6+7 strand at the dimer interface, not far from the dyad axis.

The different dimeric arrangements mean that the ACD side pockets, which are docking stations for I-X-I C-terminal motifs, are in different relative orientations in nonmetazoan and metazoan sHsps. When the I-X-I motif is included in the human α B-crystallin construct, it binds into ACD pockets, forming runaway assemblies in the crystal lattice.⁴¹ Recent solid-state NMR studies on full-length human α B-crystallin show the ACD dimer structure within the context of the assembled oligomer and also show the binding of the I-X-I motif into ACD pockets.⁴⁴ The crystal structure of bovine α A-crystallin shows that the I-X-I C-terminal motif can bind in the reverse direction by exploiting its palindromic sequence.⁴¹ In zebrafish α A-crystallin, the motif can bind into ACD pockets formed from the same chain.⁴² These pockets are candidate-binding sites for nonnative protein substrates.^{45,46} It has been proposed that a similar motif sequence in Bag3, a co-chaperone that plays a role in

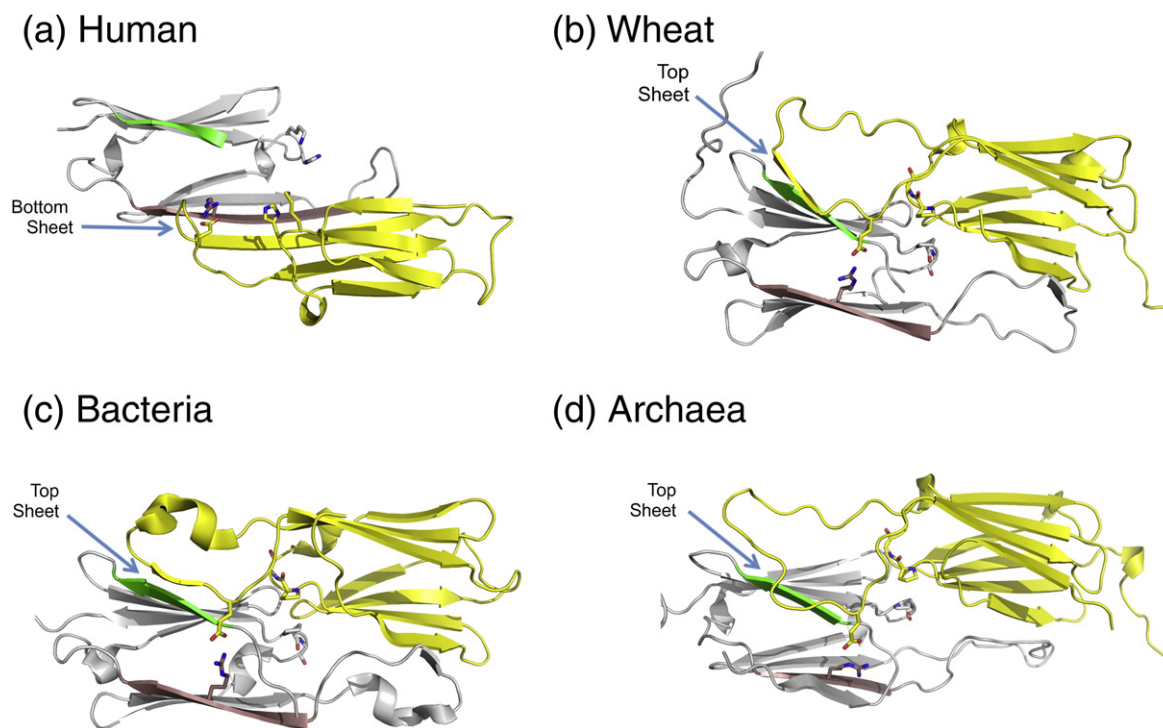


Fig. 1. The conserved arginine forms an interchain ion pair in the ACD dimers. (a) The mammalian ACD dimer from human (2WJ7). (b) The plant ACD dimer from wheat, *Triticum aestivum* (1GME). (c) The bacterial ACD dimer, from *Xanthomonas axonopodis* (3GLA). (d) The archaeal dimer, from *Methanococcus jannaschii* (1SHS). In all structures in one chain the β 7 strand (with the conserved arginine) is coloured brown, the N-terminal β 2 strand is green, and the second molecule in the dimer is yellow. Also, in the β 3/4 arch, the ProGly (of nonmetazoans) and LysHis in the human dimer are shown as sticks. The β -sheet ACD dimer interface that is formed by either the top or bottom β -sheet is indicated by a blue arrow.

macroautophagy,⁴⁷ binds into ACD pockets of Hsp22.¹²

Although the conserved arginine occurs in two kinds of structural environment, due to the different ways in which ACDs dimerise among the king-

doms of life (Fig. 1), it does form an interchain ion pair with a topologically equivalent aspartate (mammals) or glutamate (Table 1) in sHsps of all other phyla for which high-resolution structural data are currently available. Within vertebrates,

Table 1. A structural and sequence alignment of the sHsp ACD

Metazoans	β 2	β 3	β 4	β 5	β 6+7	β 8	β 9		
Human_alphaB	<u>EMRLEKDRFSVNL</u> <u>LDV</u> <u>KHFS</u> -	<u>EELKVKVLG</u> -DV--IEVHGK--HEERQ	-----E-HGFTS	REFH	KYRIPAD---VDPLTT	TSSLS	SDGVLTVNG		
Human_alphaA	<u>EVRSDRDKFVIFLDV</u> <u>KHFS</u> -	<u>EDLTVKQD</u> -DF--VEIHGK--HNERQ	-----D-HGYIS	REFH	RRYRLPSN---VDQSALS	CSLS	SADGMLTFCG		
Human_hsp27	<u>EIRHTADRWVSLD</u> <u>VNHFA</u> -	<u>EELTVKTKD</u> -GV--VEITGK--HEERQ	-----E-HGYIS	RCFTR	KYTLPPG---VDPTQV	SSLS	PEGTLTVEA		
Human_hsp22	<u>PPFPGEPPKVCVNVH</u> <u>SFK</u> -	<u>EELMVKTKD</u> -GY--VEVSGK--HEEKQ	-----E-GGIVS	KNFTK	KIQLPAE---VDPVTV	FASLS	PEGLLIEA		
Human_hsp20	<u>QVPTDPGPHSVLLD</u> <u>VKHFS</u> -	<u>EETIAVKVVG</u> -EH--VEVHAR--HEERPD	-----E-HGFVA	REFH	RRYRLPPG---VDPAAVTS	SALS	PEGVLSIQA		
Rat_hsp20	<u>QVPTDPGYFVLLD</u> <u>VKHFS</u> -	<u>EETISVKVVG</u> -DH--VEVHAR--HEERP	-----E-HGFTA	REFH	RRYRLPPG---VDPAAVTS	SALS	PEGVLSIQA		
Non-Metazoans	β 2	β 3	β 4	β 5	β 6	β 7	β 8	β 9	
Wheat_1GME	<u>DWKETPEAHVFKADL</u> <u>PGVK</u> -KEEVKVEVEDGNV--LVVSGE--RTKEKEDKNDK <u>WHRV</u> -	<u>ERSSGK</u> <u>FVRRFR</u> <u>LLED</u> ---	AKVEEV	KAGLE	-NGVLT	VTV			
Yeast_2H50	<u>DILDHDNNYELKVVV</u> <u>PGVKS</u> <u>KDIDI</u> <u>BYHQ</u> -NKNQILVSGEIPSTLNEESK-DKV-KVKESSGK	<u>F</u> -	<u>KRVITL</u> <u>PDYP</u> <u>PGVD</u> <u>ADNIK</u> <u>ADYA</u> -	<u>NGVLT</u> <u>TLV</u>					
Bacteria_3GLA	<u>DIKEEVNHFVLYADL</u> <u>PCID</u> -PSQIEVQMDK-GI--LSIRGE--RKSESSTETERFSRI-	<u>BRRYGS</u> <u>FHRR</u> <u>FALPDS</u> ---	ADADGI	TAAGR	-NGVLE	I			
Bacteria_2BYU	<u>DWKETPEAHVFKADL</u> <u>PGVK</u> -KEEVKVEVEDGNV--LVVSGE--RTKEKEDKNDK <u>WHRV</u> -	<u>ERSSGK</u> <u>FVRRFR</u> <u>LLED</u> ---	AKVEEV	KAGLE	-NGVLT	VTV			
Archaea_1SHS	<u>SIIEGDQHIKVI</u> <u>AWL</u> <u>PGVN</u> -KEDITLNAVQ-DT--LEIRAK--RSPLMITESERI	<u>IYS</u>	<u>IP</u>	<u>EEEEIY</u>	<u>RTIK</u>	<u>LPAT</u>	---	VKENASAKFE	-NGVLSVIL

The β strands for the wild-type α B-crystallin (2WJ7) and wheat HSP16.9 (1GME) are underlined and labeled β 2– β 9. In nonmetazoans, there is a conserved ProGly (red) in the connection between strands β 3 and β 4 that packs close to the dimer 2-fold axis, leaving little room for bulky residues at these positions. In metazoans, this loop is in a different environment and the ProGly sequence is not conserved, with the equivalent residues in human α B-crystallin being Lys82His83 (yellow), with a conserved proline now located four residues later in the sequence (green). Also shown is the conserved arginine (cyan) in which mutations are associated with a number of diseases. Acidic residues that form an interchain ion pair with the conserved arginine are shown in grey. There are two structurally conserved phenylalanines that form a hinge-like interaction in metazoans along the edge of the groove (purple). The T-Coffee server was used to align the above sequences and their respective structures (PDB codes 2KLR, 2WJ7, 3L1E, 2WJ5, 1GME, 3GLA, 1SHS).⁴³ The ion pairs are not indicated for 2H50 and 2BYU, as they are electron microscopy structures.

there is a source of oligomer heterogeneity that impacts on the environment of the conserved arginine and that is the register shift at the dimer interface. All the structures of the ACDs from vertebrates solved to date form an antiparallel (AP) dimer about the β 6+7 strand (Fig. 1) that has been termed the AP interface.⁴¹ The register of the strands can switch, with three different registers observed (see Supplementary Fig. S1). In both AP2 and AP3 registers, the conserved interchain ion pair (Arg120–Asp109 in α B) involving the conserved arginine is maintained.

Here we have solved the X-ray structure of the ACD dimer of human α B R120G close to physiological pH, as well as higher-resolution α B ACD wild-type controls. We compared the mutant dimer structure with several recently determined wild-type vertebrate ACD dimer crystal structures and with the solid-state dimer structure resolved in the context of the full-length sequence in a large assembly in order to assign the conformational changes attributable to the R120G mutation.

Results

Structure of the mutant α B-crystallin ACD dimer

Two mutations improved the diffraction resolution of crystals of human α B ACD: L137M substitution and truncation of the β 2-strand. The highest-resolution structure was solved from a β 2-strand truncated construct, α B ACD L137SeM (residues 71–157). The structure of the human α B ACD L137SeM (residues 67–157, the same residues as in 2WJ7) was also solved. The mutation L137M is located at the surface of a β 4/ β 8 side pocket, and in human α B ACD L137SeM (residues 67–157) the crystal packing was different from that of 2WJ7, even though crystallisation conditions were similar (Fig. S2). Both the higher-resolution L137SeM α B ACD structures are very similar to the wild-type α B ACD (2WJ7), with the RMSD of each construct (when residues 73–147 are overlaid) equal to 0.68 Å². This shows the tolerance for mutation L137M, a substitution that changes the crystal packing, but not the core ACD structure. Therefore, the R120G mutation site was engineered into the longer L137M construct, but for simplicity, from now on the R120G/L137M structure will be referred to as R120G. The RMSD between the high-resolution 71–157 α B ACD (2 Å resolution) and the R120G ACD (2.5 Å resolution) is equal to 1.589 Å².

The X-ray structure of the ACD from α B-crystallin R120G is very similar to that of the wild type, being composed of strands β 2– β 9.^{40,41,44} It is shown annotated with features common to ACDs in Fig. 2. There are highly ordered methyl pentane diol (MPD)

molecules bound near the β 6+7 loops and in the side pockets, where they mimic the binding of an isoleucine in the I-X-I C-terminal motif.⁴¹

There are now coordinates for several vertebrate sHsp ACD dimers in the Protein Data Bank (PDB),^{40–42,44} which show structural differences. When two wild-type human α B ACDs (2KLR and 2WJ7) are compared, there are significant differences between them: these structures were determined at different pH, with one being an excised ACD dimer solved by X-ray crystallography and the other an NMR structure solved in the context of a large oligomeric assembly. We will now compare all the vertebrate sHsp structures to help elucidate those changes caused by the R120G mutation. We begin by analysing discrete parts of the structure that show change and then attempt to build a more complete picture of how the ACD dimer performs its role.

Variability of the β 2 strand

The variable sequence location and conformation of the β 2 strand impacts on the accessibility of the β 3 strand. In the R120G and the zebrafish α A (3N3E) crystal structures, the β 2 strand is ordered and forms a curved β -sheet at a lattice interface, keeping the adjacent β 3 strand covered (Fig. S3). In human wild-type α B-crystallin ACD (2WJ7), the β 2 strand can be ordered or disordered (Fig. 3a). In the oligomeric structure of human α B solved by solid-state NMR (2KLR), β 3 strand residues Asp78 and Val76 are paired with upstream residues Ser59, Trp60 and Phe61 (termed β 2a) from the N-terminal extension of a separate dimer. The residues equivalent to the β 2 strand form a connection between β 2a and β 3⁴⁴ (Fig. 3c), referred to as heterogeneous region 1 (HR1), due to the many structural conformers that this sequence adopts.^{44,48}

In the rat Hsp20 (HSPB6) structure (2WJ5), the sequence corresponding to the β 2 strand is not hydrogen bonding with β 3, instead it extends across to a partner domain in the lattice and inserts a hydrophobic side chain, topologically equivalent to Met68 in the β 2 strand of human α B-crystallin, into a β 4/ β 8 side pocket. The absence of both β 2 strands from the two top sheets of rat Hsp20 ACD dimer creates space for the binding of the (C-terminal) peptide into the shared groove (Fig. 3b).

Taken together, these interchain interactions in the structures highlight the role in oligomerisation that the N-terminal and β 2 strand sequences might play and the opportunity for β 3 strand to bind variable partners.

Variability of the bottom β -sheet

The length of the β 6+7 strand is determined (in part) by the number of H-bonds to the β 5 strand and

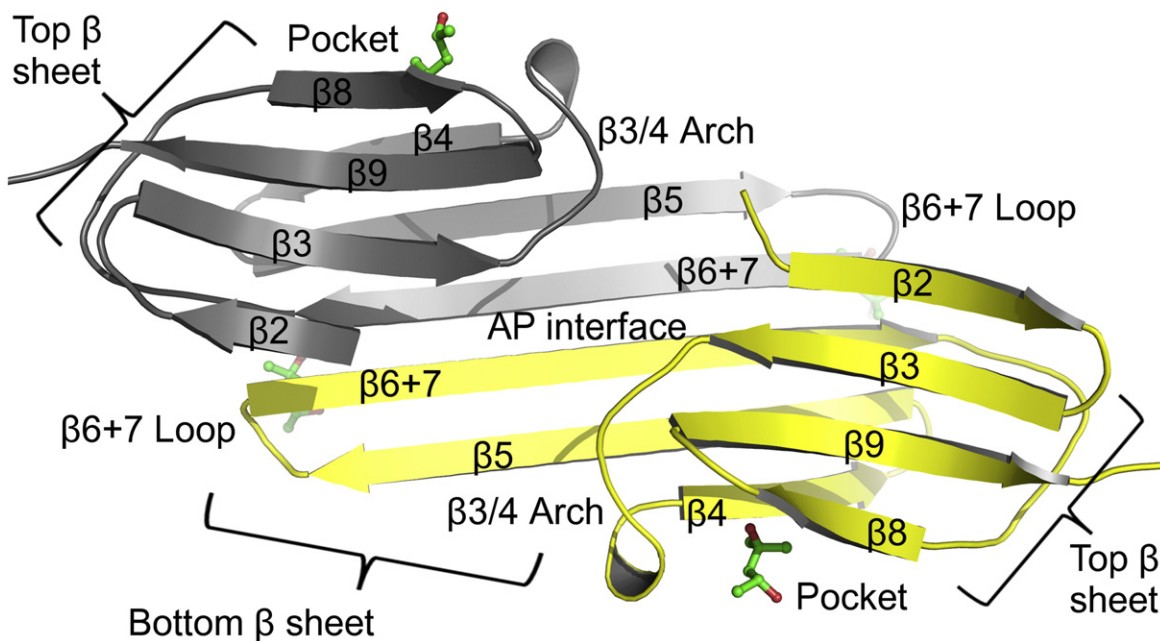


Fig. 2. The structure of the α -crystallin domain of α B-crystallin R120G/L137M. The structure is annotated with the structural features of ACDs. The β -sheets are labeled and the hairpin loop connecting the $\beta 6+7$ and $\beta 5$ strands is indicated. One of the loops that link the two sheets in each monomer (strand $\beta 3$ to $\beta 4$) is shown as the $\beta 3/4$ arch. The side pockets at the edges of $\beta 4$ and $\beta 8$ are labeled, and the AP dimer interface is labeled. The MPD molecules in the $\beta 4/\beta 8$ pockets and below the loops are depicted as green sticks.

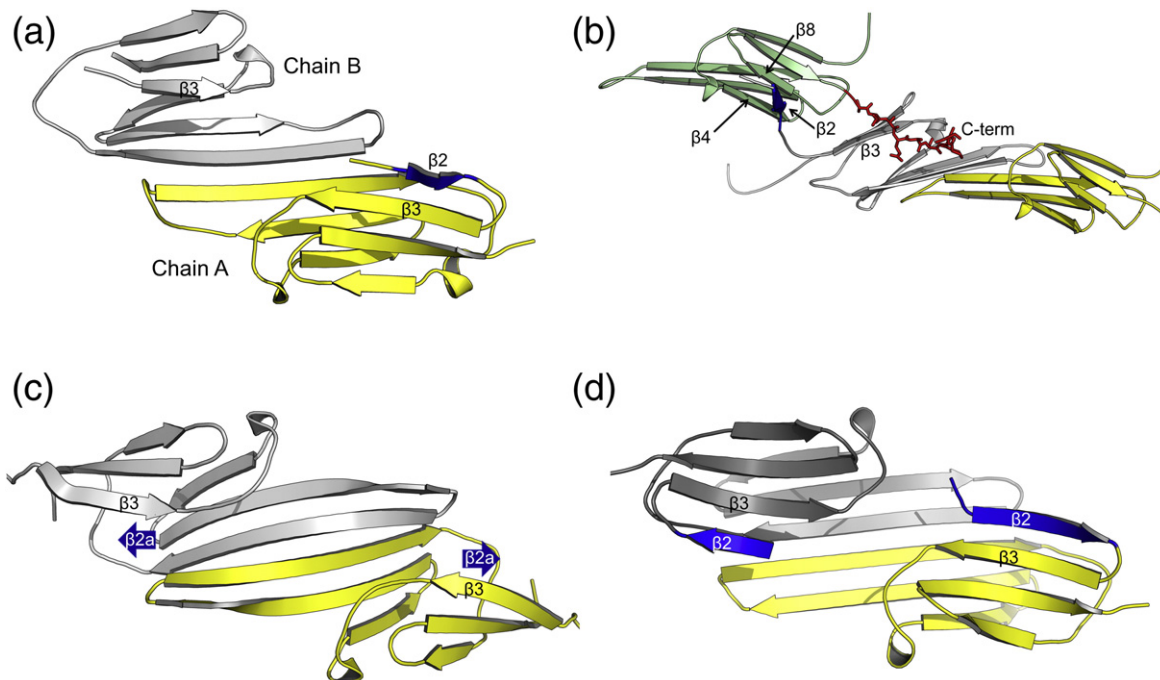


Fig. 3. Variability and strand swapping at the $\beta 3:\beta 2$ position. (a) In human α B (2WJ7) the amino acids adjacent to $\beta 3$ form the $\beta 2$ strand in one chain of the dimer. (b) In rat Hsp20 (2WJ5) the $\beta 2$ strand is not part of a β -sheet and (together with the construct N-terminus) is extended in both chains, allowing a hydrophobic side chain to bind in the $\beta 4/\beta 8$ pockets of an adjacent dimer. The C-terminal extension of a symmetry-related chain now binds to the $\beta 3$ strand (shown as red sticks). (c) In human α B (2KLR) the intrachain $\beta 2$ strands are not present; however, interactions from upstream N-terminal residues from another dimer are present. (d) The R120G structure showing intrachain $\beta 2$ strands for each chain.

is correlated with the length of the intervening β 6+7 loop. In wild-type α B ACD (2WJ7), at pH 9 the loops sit up, at almost 90° to the β -sheet (Fig. 4a and b), whereas in the wild-type α B structure solved by solid-state NMR (2KLR) and in R120G, the shorter loops sit below the plane of the β -sheet. The changes in loop conformation change the surface structure of the domain around the AP interface.

The α B-crystallin ACD R120G dimer structure is in the same AP register as the pH 9 wild-type α B-crystallin ACD (2WJ7) and the pH 7.5 full-length α B-crystallin solid-state NMR structure (2KLR) and is close to physiological at pH 7.4. The AP interface is formed by β -sheet strands β 6+7, β 5 and β 4 extending with the equivalent strands in the other chain into a larger β -sheet, referred to here as the bottom sheet. There is considerable flexibility in the structure of this extended sheet (Fig. 4a). It may be

flat (3L1G, 3L1E and 2WJ7). It may be curved due to the register of the individual sheets changing, such as rat Hsp20 (2WJ5), where although the strands remain fairly flat (Fig. 4c), the two monomers have moved further away from the midpoint (into register AP3), creating a concave and convex surface (Fig. 4A). Then, finally, the β -sheets may have an intrinsic twist, as seen in the solid-state NMR structure (2KLR) and, to a lesser extent, the two α A structures (3N3E and 3L1F) (Fig. 4a and c).

It has been suggested that the curvature of the β -sheet formed by β 4, β 5 and β 6+7 is pH dependent⁴⁴ based on the difference in curvature between the NMR structure at pH 7.5 (2KLR) and the crystal structure at pH 9 (2WJ7). However, the wild-type α B ACD crystal structures (that are now solved) are both flat at pH 9 and pH 4.6 (3L1G), and the α A structures at physiological pH are nearly flat (3L1E

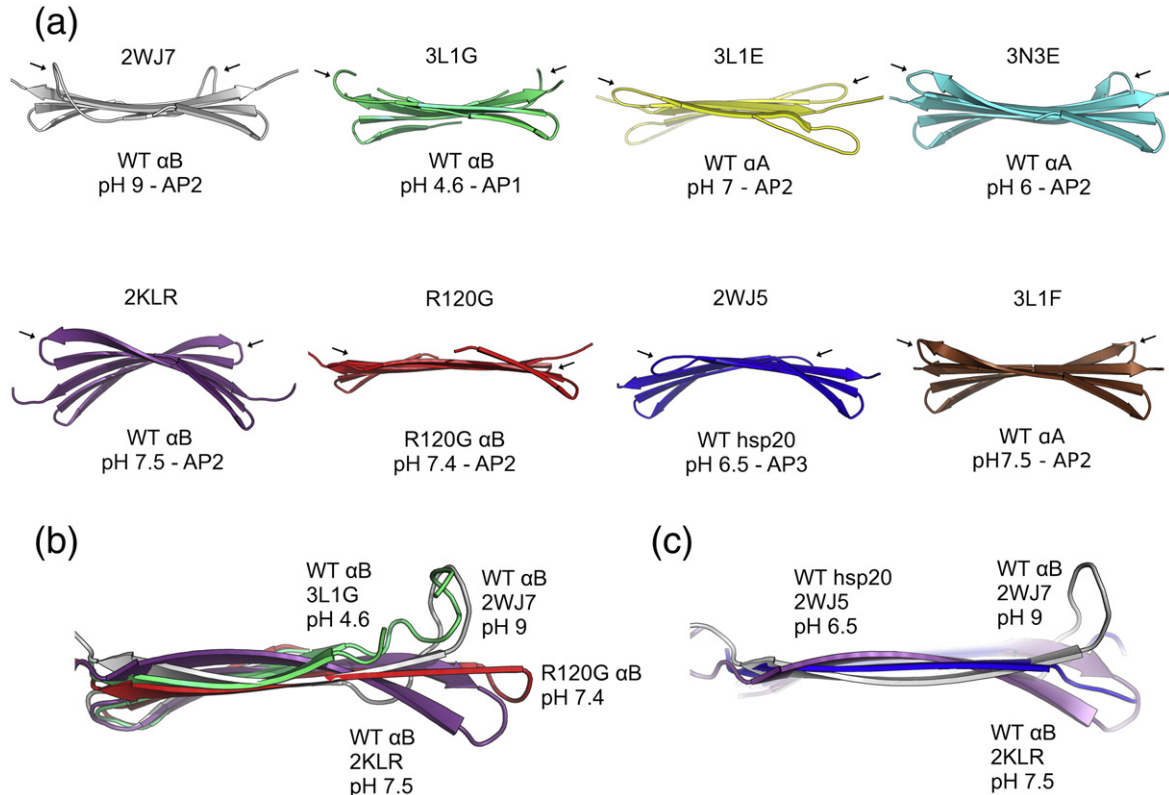


Fig. 4. The bottom sheet, strands β 6+7, β 5 and β 4, is variable. (a) The wild-type human α B ACD, 2WJ7 (grey); wild-type human α B ACD and C-terminal motif, 3L1G (green); wild-type bovine α A ACD and C-terminal motif plus zinc, 3L1E (yellow); wild-type zebrafish α A ACD and C-terminal motif, 3N3E (cyan); the solid-state NMR ACD structure from human α B full-length oligomers, 2KLR (purple); human α B ACD, R120G (red); wild-type rat Hsp20 ACD, 2WJ5 (blue); wild-type bovine α A ACD and C-terminal motif, 3L1F (brown). Note that some structures have a flat sheet, while other structures have varying degrees of curvature. The register is listed below the structures, as AP1, AP2 or AP3. The pH of the crystal is also noted below the sheet. The loop position is indicated with an arrow. (b) An overlay of the β 6+7 and β 5 strands from the α B structures, looking from the AP dimer interface showing the β 6+7 loop position and the length of the β 6+7 strands. The loop between the β 6+7 sheet and the β 5 strands sits up at 90° to the sheet in the wild type at pH 9 (grey) and pH 4.6 (green). The R120G (red) has longer β -sheets flanking the loop, and the loop does not sit up like those of the wild type. (c) The β 6+7 strands from 2WJ7, 2WJ5 and 2KLR are overlaid. The 2KLR shows the marked curve in the strand, whereas the 2WJ7 and 2WJ5 have essentially flat strands. The curvature observed in rat Hsp20 [in (a)] is due to the register shift, and not intrinsic curvature of the β -sheet.

and 3L1F), making it unlikely that pH is solely responsible for the variability seen in the curvature of the bottom sheet (Fig. 4).

The dynamic AP interface

The AP dimer formed about the β 6+7 strands has been observed in three different registers in the vertebrate sHsp structures⁴¹ (Fig. S1), with the AP1 register having the greatest overlap and AP3 the least. In addition to the variation in the area of the AP interface due to the different registers, there is also the impact of the length of the β 6+7 strand (Fig. 3b). The length of this strand is correlated with the number of H-bonds across the AP interface and therefore the dimer interface area. The longer β 6+7 strands of the pH 7.5 structure (2KLR) and the R120G structure form a more extensive interface than the pH 9 structure (2WJ7). The surface areas of the various ACD dimer AP interfaces were analysed by PISA.⁴⁹ In the wild-type α B-crystallin (67–157) structure (2WJ7), there are five chains in the asymmetric unit of the crystal C2 cell, forming three AP2 interfaces (two noncrystallographic and one crystallographic) with an average area of

633 Å². In the solid-state NMR structure of the full-length α B-crystallin structure (2KLR), the ACD dimer interface (also in AP2) has an area of 895 Å². With an interface area of 982 Å², the R120G has a greater interface than any of the wild-type α B structures, even when the register and β 6+7 sheet length are considered. This increased surface area is due to increased intermolecular interactions, described below. This increase in the dimer interface is also consistent with a tighter dimer interface for the R120G ACD dimer, shown by gel filtration at acidic pH.⁴⁰

The AP dimer interface is likely sensitive to the environment, especially pH, as it is rich in physiologically titratable histidines. All the α -crystallin dimer structures (α A and α B) solved at physiological pH (7–7.5) are in AP2, as well as α B ACD at pH 9, whereas the more distantly related rat Hsp20 ACD at pH 6.5 (2WJ5) is in register AP3. Lowering the pH to 4.6 (3L1G), in the case of α B-crystallin, may cause the register to shift from AP2 to AP1, which results in disruption of the entire ion pair network, including the ion pairs between Arg120 and Asp109 (Fig. 5). In 2WJ7 and 2WJ5 these ion pairs are bidentate, whereas in 2KLR they are

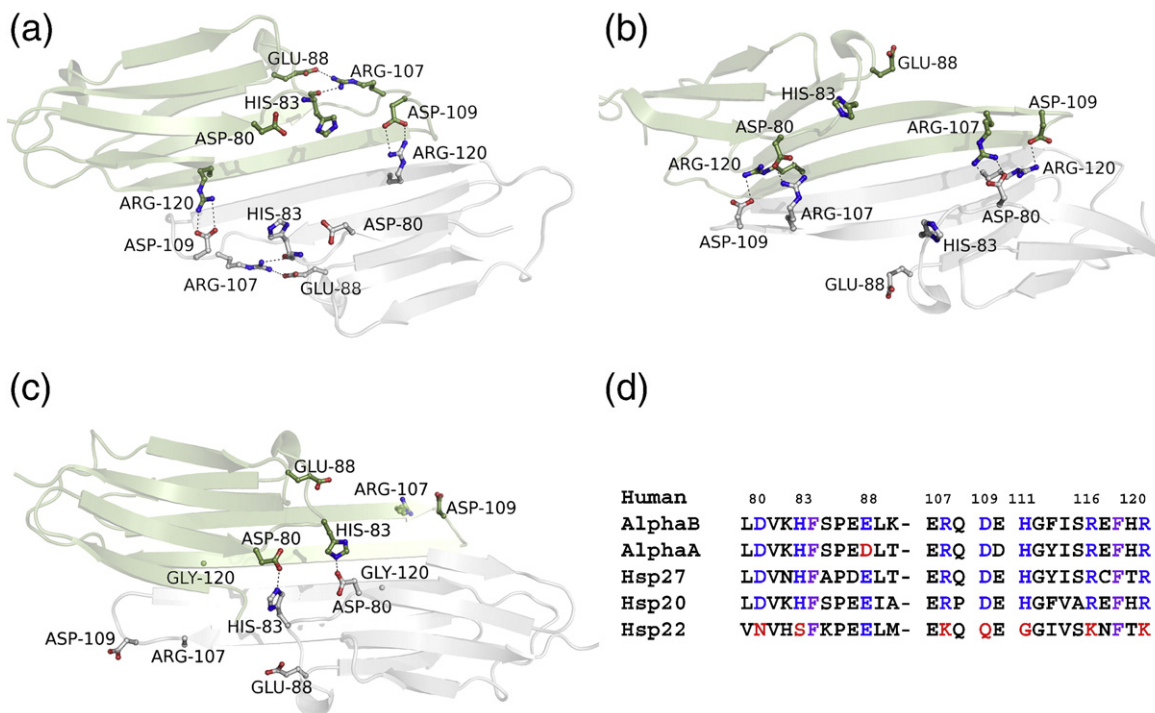


Fig. 5. Conservation and variation of charged residues at the AP interface. (a) The wild-type α B-crystallin ACD (2WJ7) (pH 9). (b) The ACD dimer from full-length wild-type α B-crystallin (2KLR) (pH 7.5), determined by solid-state NMR. (c) The α B-crystallin ACD R120G structure (pH 7.4). The two chains are coloured light green and grey, and the residues forming key interactions at and around the interface are shown as sticks. Interactions are shown with dashed lines. (d) The truncated sequence alignment shows that in α A, α B, Hsp27 and Hsp20 the residues involved in the ion pairs at the AP interface are conserved, but vary in Hsp22. Human α B-crystallin numbering is shown above; residues coloured blue are those indicated on structures in (a)–(c), and residues that are not identical to the α B sequence are coloured red. The two conserved phenylalanine residues (84 and 118) are also shown in purple.

monodentate. The presence of bidentate ion pairs between Arg107 and Asp80 across the AP2 interface in 2KLR is a major structural difference between the two wild-type α B structures, the curved 2KLR and the flat 2WJ7, and may be responsible for the extensive curvature of the β 6+7, β 5 and β 4 sheet of 2KLR (Fig. 5).⁴⁴

While clearly this is a structurally dynamic interface, the two ion pairs between Arg120 and Asp109 across the dimer interface are conserved in the wild-type α B ACD at pH 9 (2WJ7) and the full-length solid-state NMR α B at pH 7.5 (2KLR) and rat Hsp20 (2WJ5) (Fig 1 and Table 1). This interaction is even maintained when there is a register shift, from AP2 to AP3 (rat Hsp20) and when there is an increase in pH. It is this pair of salt bridges that is disrupted in the pathogenic mutation R120G (Fig. 5).

Arg120 is part of an extensive ionic interaction network that crosses the dimer interface (Fig 5). Sequence alignments show a set of residues (Asp 80, His83, Arg107, Asp109, His111, Arg116, Arg120) that is conserved in α B, α A, Hsp27 and Hsp20, yet they are all different in Hsp22 (inset in Fig 5). This strengthens the case that these charged side chains are functionally connected and that Hsp22 has diverged. In muscle α B, Hsp27 and Hsp20 can form co-oligomers,⁵⁰ and α A, α B and Hsp27 have been shown to co-oligomerise and exchange subunits *in vitro* in a temperature-dependent manner⁵¹ as have Hsp27 and Hsp20.⁵² Although co-assembly of Hsp22 with Hsp27 is less clear, there is cell biological evidence that mutant Hsp22 interacts more avidly with Hsp27 than wild-type Hsp22.³² If sHsps use this interface to bind unrelated protein partners (for insertion into networks) or nonnative protein substrates (as a chaperone), then the distinctive Hsp22 AP interface may facilitate binding of different partners and substrates from the other sHsps.

Changes to the charged groove at the mutant AP interface

The sHsp wild-type ACD X-ray structures all have a deep groove that runs down the AP dimer interface. The groove is maintained through sequence changes, register shifts and across a wide pH range (Fig. 6). The base of the groove is lined with four arginines, Arg120 and Arg116 on each of the β 6 +7 strands. The change in Arg120 to Gly120 results in the loss of two positive charges along the interface. This in turn results in loss of two interface ion pairs between Arg120 and Asp109 and gain of two new interface ion pairs between His83 and Asp80 (Figs. 5 and 6). These ion pairs between Asp80 and His83 are not observed in any of the currently available wild-type α B ACD structures. Both residues are on an arch between strands β 3 and β 4 that links the top sheet to the bottom sheet in the

ACD on each side of the groove (Fig. 2). Movement of these arches toward each other to form two pairs of interchain salt bridges blocks the groove (Fig. 5). In the wild-type α B 2KLR structure the cleft is also absent, but in this case the bottom sheet curves up into the area between the top sheets (β 3, β 9 and β 8) (Fig. 7).

The ability to open and close the groove may be facilitated by two phenylalanines that are conserved in all metazoan ACDs (purple in the alignment in Table 1 and the inset in Fig. 5). In the 3D structures, they are on the edge of the shared groove (α B residues 84 and 118) and interact by edge-to-face stacking of the side chains (Fig. S4). In the NMR structure of α B (2KLR), Phe118 also interacts with Arg116 across the AP interface *via* a pi-cation interaction.⁴⁴ Arg116 makes an intrachain cation–aromatic interaction with His83 in α B (2WJ7) at pH 9, α A (3L1E) at pH 7 and α A (3L1F) at pH 7.5 (yellow in Table 1). This set of amino–aromatic interactions is embedded under the network of ionic interactions. As His83 is in a central position, it may respond to pH changes. In 2KLR, the interaction between His83 and Arg116 is broken, as it is in the crystal structures of rat Hsp20 (2WJ5) at pH 6.5, α B (3L1G) at pH 4.6 and α A (3N3E) at pH 6. In 2KLR, the altered position of Phe118 correlates with the bottom sheet being twisted (Fig. 7).

It is possible that the ACD dimer is able to fluctuate between open- and closed-groove conformations, with a closed state being favoured in the oligomer and the open state in the dissociated dimer. The structure of the R120G indicates that in the mutant this dynamism would be dampened.

Discussion

Biophysical methods indicate that in solution α B-crystallin is polydisperse in size and in rapid dynamic equilibrium with subassembly species.^{5,16,35} We now propose, from analysing the 3D structures of ACD dimers currently available in the PDB, that the dimer itself exhibits conformational flexibility. Using N-terminally truncated sHsp sequences, crystallography has captured snapshots of dimers bearing a deep, shared groove at the dimer interface that is conserved in the structures of vertebrate ACDs of α B, α A and Hsp20.^{40–42} In the solid-state NMR structure, the α B ACD dimer has conformational differences from those solved by crystallography.⁴⁴ The extended β -sheet, formed by strands β 4, β 5 and β 6+7, is essentially flat in the crystal structures; however, in the NMR structure it is twisted, more like the shape of a hyperbolic paraboloid. The twisted bottom β -sheet curves up and fills the groove space. This form of α B-crystallin ACD dimer with the groove in a closed state either does not crystallise or exists only in the full-length highly

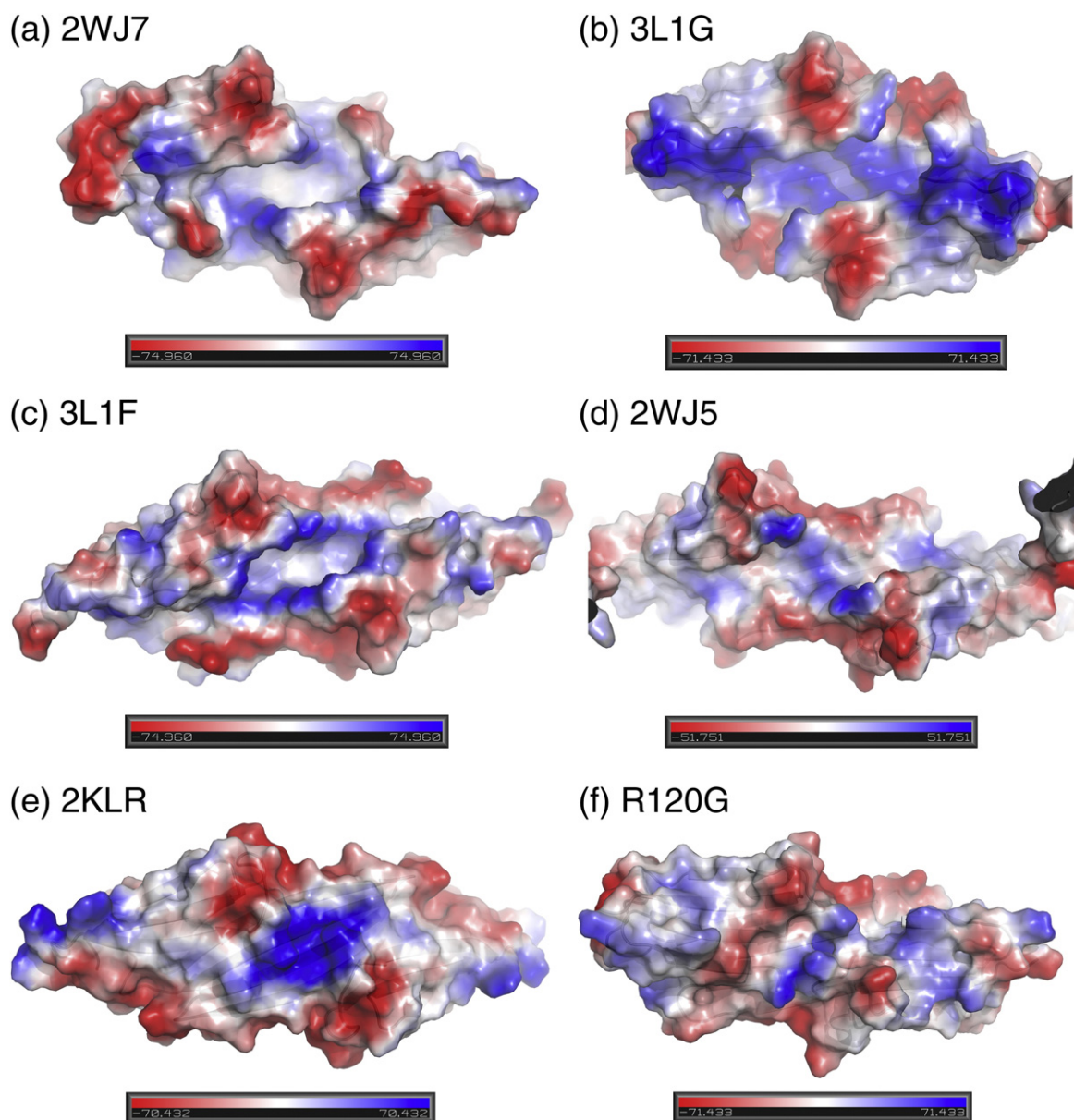


Fig. 6. The conserved positively charged groove along the AP interface is lost in R120G. This groove is conserved in wild-type ACD structures of human α B [2WJ7 (a) and 3L1G (b)], bovine α A (3L1F) (c) and rat Hsp20 (2WJ5) (d). In the R120G, the groove is absent, due to ion pairs between Asp80 and His83, on the arches connecting top (β 3, β 9 and β 8) and bottom sheets, β -sheets having moved closer to the centre in R120G. (e) The groove is closed in the ACD dimer from full-length wild-type α B-crystallin (2KLR).

oligomeric state. The crystal structure of R120G ACD dimer also reveals a closed groove, but in the mutant closure is caused by a pair of salt bridges between His83 and Asp80 close to the dimer 2-fold axis. The His83–Asp80 ion pair is not observed in any of the wild-type ACD dimer structures currently available. The two structures with closed grooves were solved at pH 7.4/5.

In the structures of Hsp20, α A and α B, there are molecules (peptide, MPD and sulfate) that have co-crystallised in the AP groove (Fig. 8). This highlights

the groove as a potential binding site for polypeptides and small molecules. The function of this dynamic groove down the AP interface may be to bind a phosphate-containing small molecule, which could allow regulation of these sHsps. In the wild-type α B structure (3L1G), a sulfate is bound at the 2-fold interface in the dimer (Fig. 8). Early experiments showed that ATP modulated the limited digestion of α B-crystallin with trypsin or chymotrypsin, which indicated that it could exert a conformational change of the ACD.⁵³ When the residues with a nucleotide-

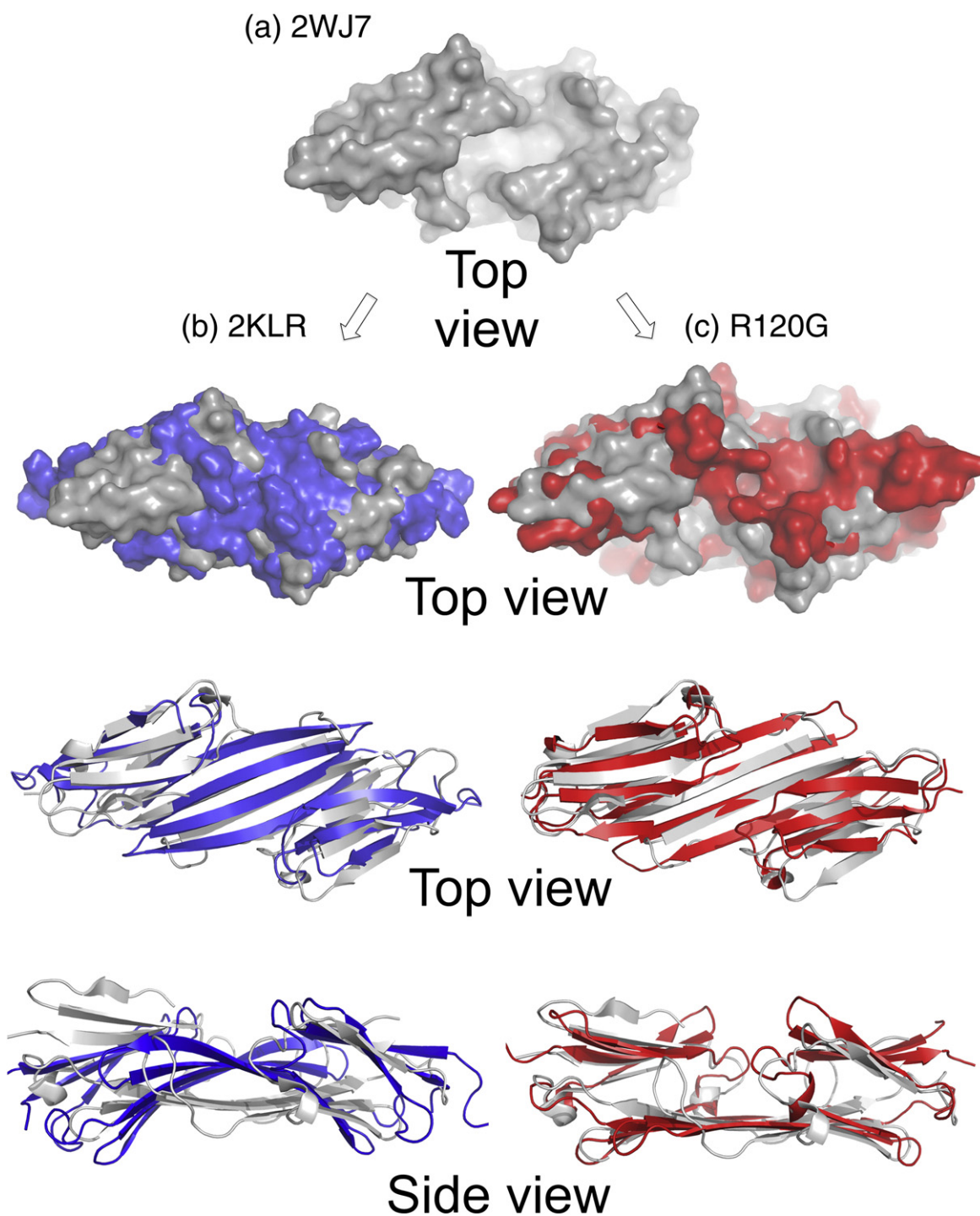


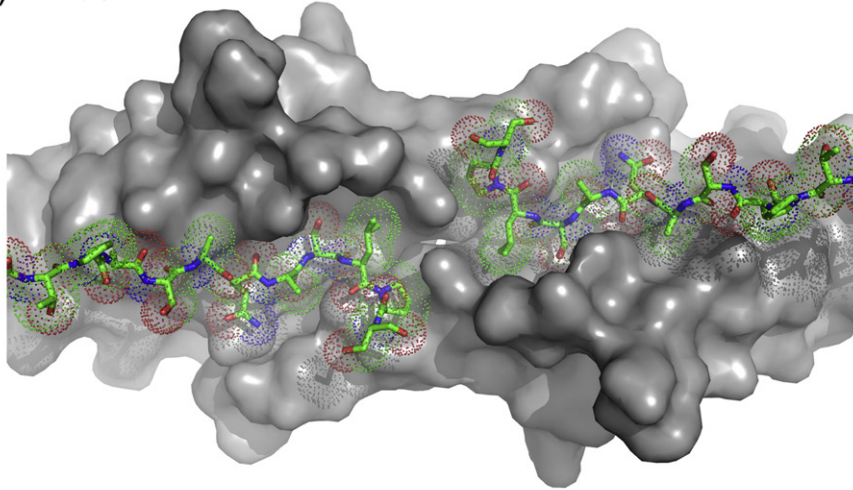
Fig. 7. Closing the groove in two different ways. (a) The wild-type α B-crystallin ACD (grey) showing the deep groove. (b) In the full-length wild-type α B-crystallin (blue) the lower β -sheet has moved up into the space where the groove was in structure (a). (c) In the R120G α B-crystallin ACD (red) the β 3– β 4 arches have moved into the groove. In (b) and (c) the 2WJ7 structure is shown in grey, overlaid with the second structure shown in either blue (2KLR) or red (R120G).

dependent protease susceptibility are mapped onto the crystallin domain, they are located around the site of the sulfate molecule, lining the groove at the AP interface (Supplementary Fig. S5). It is possible that

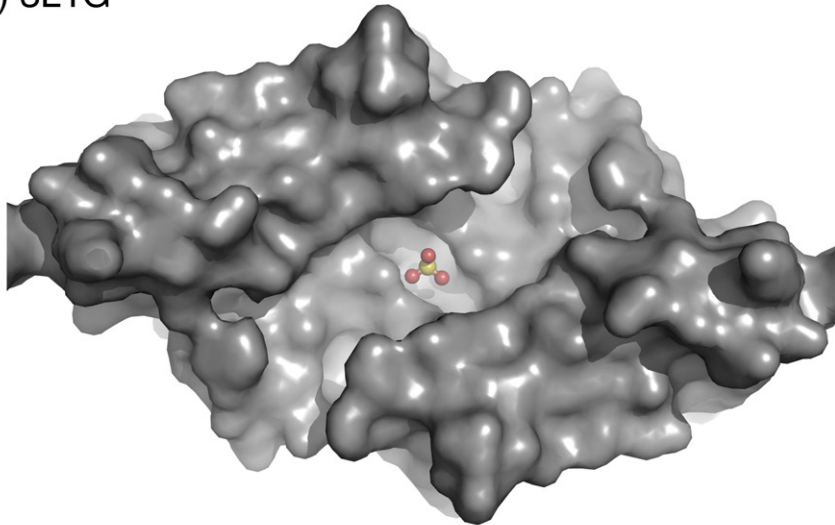
the sulfate in 3L1G may be bound in an ATP (or similar compound) binding site.

Recombinant full-length α B R120G mutant protein forms a soluble oligomer that is both larger and

(a) 2WJ5



(b) 3L1G



(c) 3L1F

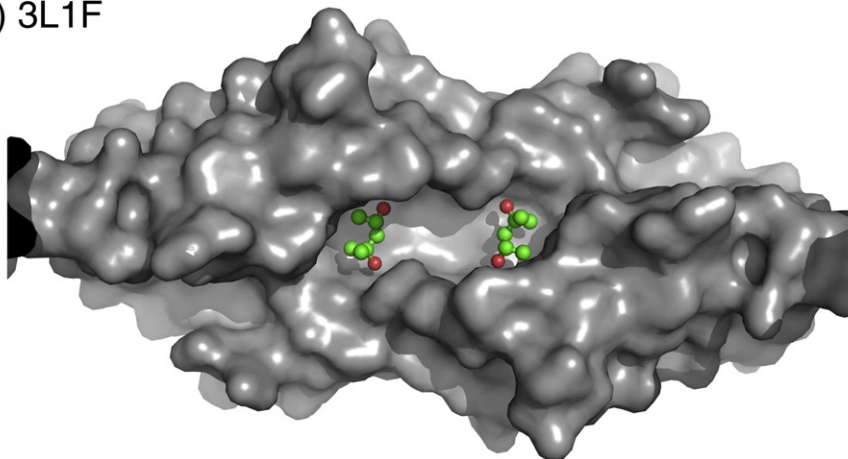


Fig. 8. Groove as a binding site for other molecules. (a) The rat Hsp20 ACD has the C-terminal of a chain from an adjacent dimer bound in the groove. (b) The human wild-type α B ACD (3L1G) has a sulfate bound at the crystallographic 2-fold interface between the molecules. (c) The bovine wild-type α A has two MPD molecules bound in the groove.

more polydisperse than wild type³³ and is prone to aggregation and precipitation on handling.³⁴ The mutation thus disturbs higher assembly of the oligomer, and the crystal structure of R120G ACD shows a dimer with a blocked groove. The N-terminal extension of α A has been shown to be necessary for its higher assembly.⁵¹ The solid-state NMR study provides evidence for the N-terminal extension playing a role in the oligomerisation of α B-crystallin by showing that the region around Trp60 (in what is termed the β 2a strand) interacts with the β 3 strand from a different dimer.⁴⁴ It is possible that the N-terminal extension is involved in regulation of assembly. We speculate that these extensions may bind in the open groove in the full-length subassembly state, as the rat Hsp20 C-terminus binds in the Hsp20 groove in the crystal structure (in 2WJ5). In this scenario, the burial of the largely hydrophobic N-terminal extension may drive the dissociation of the dimer out of the higher oligomer.

How might these structural observations contribute to understanding how mammalian sHsps function? α B-Crystallin, in common with other more regular nonmetazoan sHsp assemblies, is in dynamic equilibrium in solution with subassembly species.^{51,54,55} The R120G with its new set of ion pairs would favour the ACD dimer in the closed form and disturb the subunit-exchange-driven assembly dynamics. Recently, an investigation of the quaternary dynamics of pea HSP18.1 has shown how heat can cause the regular assembly to increase in size and polydispersity, along with increased dissociation into subassembly species, in a reversible manner.⁵⁶ These authors show evidence that this thermal plasticity correlates with chaperone activity. Stress in mammalian systems can be relayed by kinases, with serines in human sHsps often being phosphorylated. In Hsp27, there is clear evidence for their phosphorylation causing dissociation into smaller oligomers.^{57,58} In keeping with our speculative model of the released N-terminal extension binding the open groove of a dimer, the four arginines in the groove may bind with higher affinity to the phosphorylated N-terminus, favouring the dissociated state. The groove may also be a site of regulation, as reflected by a number of small molecules that have co-crystallised in it and from the evidence that ATP might bind here accompanied by a structural rearrangement.⁵³ The depth of the cleft and its possible role in the function of wild-type α B-crystallin make this region a possible candidate site for drug design.

The R120G mutant is a defective chaperone and coaggregates with other destabilised proteins in chaperone assays.³³ This may be due to the mutation disturbing the subunit exchange dynamics, although for α B-crystallin, there is evidence that the efficiency of exchange is only slightly decreased

by the R120G mutation⁵⁹ or is increased.³⁴ The groove may also be a client substrate-binding site (chemical probes also label β 3 strand sequence⁴⁵) whose access is regulated through the dynamics of the assembly in ways that involve the dynamics of the N-terminal extension. In a well-characterised nonmetazoan sHsp system for detecting substrate-binding sites, both the N-terminal extension and β 7 strand bound substrates.⁶⁰ However, the solution state of the dissociated full-length α B-crystallin dimer is not yet accessible to standard structural methods, and the geometric locations of the groove and ACD pockets are not defined in the ambient or in the activated forms of α B-crystallin. Our provisional model for how sHsps might bind substrates involves unfolded polypeptide chains wrapping around ACD domains by docking into substrate-binding regions formed by the shared groove and β 4/ β 8 pockets. Support for this model comes from the X-ray structure of an sHsp from a parasitic worm, where completely resolved N-terminal extensions wrap around a set of four ACDs.⁶¹

Forced expression of α B R120G is highly toxic to cells.⁶² It causes activation of the aggregosome pathway,⁶³ with formation of massively aggregated bodies in the cell, similar to the effects caused by blockage of the proteasome.²⁰ Animal models that overexpress α B R120G in heart muscle cells have increased levels of amyloid and develop cardiomyopathy.⁶⁴ The ability of R120G to form oligomers larger than wild type yet with some solubility may be the basis of its toxicity causing it to aggregate over time with crucial cellular components, possibly blocking or overwhelming the waste-disposal systems in muscle cells. Similar damage has been suggested to cause motor-neuron-specific neurite degeneration when the sequence equivalent mutation, Hsp22 K141N/E, is expressed in elongated cultured primary neuronal and glial cells.⁶⁵ In the eye lens, where most cellular organelles have faded away, R120G likely causes cataract by binding to other crystallins, causing aggregation that then scatters visible light.

Methods

Protein expression and purification

Three protein constructs of human α B-crystallin that encoded the residues of the ACD were used: α B L137M (residues 71–157), L137M (residues 67–157) and R120G/L137M (residues 67–157). The mutations R120G and L137M were generated by QuikChange® PCR mutagenesis, both in the longer construct (residues 67–157). After purification and recombinant tobacco etch virus (rTEV) cleavage of the histidine (His) tag from this construct, all the recombinant proteins contained three extra residues at the N-terminus, GAM.

The L137SeM (71–157) and L137SeM (67–157) constructs in the pProExTM Ht(b) vector were expressed in *Escherichia coli* grown in SelenoMetTM medium base (Athena Enzyme Systems) with a hexahistidine affinity tag and purified by Ni chromatography, rTEV cleavage, and then gel filtration, as previously described for the wild-type α B-crystallin domain.⁴⁰

To purify the R120G/L137M protein, a slightly different purification was used. Cells from a 6-L culture were lysed in buffer A (0.5 M NaCl, 50 mM potassium phosphate, pH 8) containing 20 mM imidazole and an EDTA (ethylenediaminetetraacetic acid)-free protease inhibitor cocktail tablet (Roche) by a cell disruptor (EmulsiFlex-C5, AVESTIN), and then briefly sonicated to shear the DNA. The cellular debris and any remaining whole cells were removed by centrifugation (20,000g for 30 min) and filtered through a 0.45 μ m filter. The protein solution was then loaded onto a preequilibrated 5 -mL HisTrapTM HP column (GE Healthcare). The unbound proteins were removed with 10 column volumes of buffer A containing 20 mM imidazole. The proteins were then eluted in a stepwise imidazole gradient of buffer A containing 100, 150 and 250 mM imidazole, with the recombinant protein eluting in the 250 mM fractions. These fractions were pooled and dialysed into buffer [25 mM Tris-HCl (pH 8), 100 mM NaCl, 1 mM DTT] at 4 °C. Then rTEV protease was added (100:1 ratio of substrate/protease) and incubated at 4 °C for 2 days to cleave the His tag. The protein was then dialysed into buffer A (also containing 20 mM imidazole) and the flow-through from a second preequilibrated 5 -mL HisTrapTM HP column was collected. This step removes the His tag, rTEV and any uncleaved recombinant protein that binds to the Ni column. The eluant was then further purified by gel filtration on a Superose75 HR10/60 column, in buffer (25 mM NaCl, 25 mM Tris-HCl, pH 8). Fractions containing the recombinant protein were pooled, concentrated to 16.5 mg/mL (as determined by A_{280} using an extinction coefficient of 1446 M⁻¹ cm⁻¹) and stored at -80 °C.

Crystallisation

Crystals were all produced using the vapour-diffusion technique; the conditions used for each construct are provided below. Following crystallisation, crystals were mounted in a loop in the drop solution (all conditions contained MPD concentrations sufficient for cryoprotection) and flash-frozen by immersion in liquid N₂.

L137SeM α B (71–157) pH 9

One microlitre of protein at 20 mg/mL in buffer [25 mM Tris-HCl (pH 8.5), 200 mM NaCl] was mixed with 1 μ L of mother liquor [40 mM bicine (pH 9.0), 40% MPD] and equilibrated against wells containing 50 μ L of mother liquor and incubated at 4 °C.

L137SeM α B (67–157) pH 9

One microlitre of protein at 20 mg/mL in buffer [25 mM Tris-HCl (pH 8.5), 200 mM NaCl] was mixed with 110 mM bicine (pH 9.0) and 55% MPD and equilibrated against wells containing 50 μ L of mother liquor at 16 °C.

R120G/L137M α B (67–157) pH 7.4

One microlitre of protein at 16.5 mg/mL in buffer [25 mM Tris-HCl (pH 7.5), 25 mM NaCl] was mixed with 1 μ L of mother liquor (36% MPD, 0.18 M potassium chloride), giving a final pH of 7.4 in the drop solution. Sitting drops were then equilibrated against wells containing 50 μ L of mother liquor at 16 °C.

Structure determination

L137SeM α B (71–157)

Data were collected at the Diamond Light Source, beamline I02, on an ADSC CCD detector. Eight hundred images with an oscillation angle of 0.45° were collected at a wavelength of 0.9794 Å. Initial indexing indicated point group 422. Following integration with Mosflm⁶⁶ and scaling with Scala,⁶⁷ systematic absences showed the space group to be $P4_{1/3}22$ with one molecule in the asymmetric unit (giving a Matthews coefficient of 2.0 and a solvent content of 38%). Data processing statistics are shown in Table 2.

The structure was solved by molecular replacement using Phaser⁶⁸ and the wild-type α B-crystallin (PDB code 2WJ7) lacking the β 2-strand as a model. The solution was in space group $P4_322$ and had a Z score of 19.6 and an initial R-factor of 45.6%. Rounds of refinement with Phenix,⁶⁹ alternating with manual rebuilding with Coot,⁷⁰ resulted in a final R and R_{free} of 20.4% and 22.1%, respectively. Refinement utilised automatic water picking, TLS refinement (with two domains corresponding to residues 10–58 and 59–87, or each of the two β -sheets) and refinement of the anomalous signal from the single ordered selenomethionine residue. An ordered MPD molecule was observed in the density proximal to Ser135 in the β 4/ β 8 ACD binding pocket.

L137SeM α B (67–157)

Data were collected at the European Synchrotron Radiation Facility (ESRF), beamline 14-1, where diffraction for this crystal extended to 3.7 Å. Three hundred and sixty images with an oscillation angle of 1° were collected for each of three wavelengths (peak at 0.978 Å). Initial indexing showed the data to be in point group 222. Following integration with Mosflm and scaling with Scala, systematic absences indicated the most likely space groups to be $P2_122$ or $P2_12_12$. Six α -crystallin domains in the asymmetric unit of this cell result in a Matthew's coefficient of 2.59 and an estimated solvent content of 52%. Data collection statistics are provided in Table 2.

The structure was solved by molecular replacement using Phaser and the higher-resolution L137SeM (71–157) coordinates as a model. Initially, four chains could be placed with good density, with a Z score of 25.7 and an initial R-factor of 48.5% in space group $P2_12_12$. However, these four chains did not form a lattice, and inspection of the initial maps revealed areas of positive density between the layers of molecules placed by Phaser. An additional α B ACD dimer was placed in this density by inspection, thus completing the lattice. The L137SeM low-resolution data set, originally collected as part of a multiwavelength anomalous dispersion experiment, was intended for use in structure solution. A significant anomalous signal is

Table 2. X-ray data collection and refinement statistics

	L137M residues 71–157	L137M residues 67–157	R120G/L137M residues 67–157
<i>Crystal parameters</i>			
Space group	$P4_322$	$P2_12_12$	$P3_221$
Cell dimensions (Å)	$a=36.03, b=36.03, c=148.32$	$a=67.12, b=78.42, c=130.67$	$a=60.73, b=60.73, c=100.15$
Angles (°)	$\alpha=\beta=\gamma=90$	$\alpha=\beta=\gamma=90$	$\alpha=\beta=90^\circ, \gamma=120^\circ$
<i>Data collection</i>			
Wavelength (Å)	0.9794	0.978	0.8726
Resolution limit (Å)	2.0–35.01 (2.11–2.0)	3.7–67.24 (3.90–3.70)	2.50–46.56 (2.64–2.50)
Mosaicity	0.4	0.6	1.3
R_{merge}	0.097 (0.322)	0.236 (0.66)	0.099 (0.536)
R_{anom}	0.039 (0.066)	0.058 (0.169)	
Total no. of observations	184850 (27285)	100933 (14795)	69470 (9859)
Total no. unique	7272 (1018)	7861 (1109)	7582 (1087)
Mean $I/\sigma I$	25.6 (11.6)	8.7 (4.0)	19.4 (4.3)
Completeness	100.0 (100.0)	100.0 (100.0)	98.0 (97.8)
Anomalous completeness	100.0 (100.0)	100.0 (100.0)	
Multiplicity	25.4 (26.8)	12.8 (13.3)	9.2 (9.1)
Anomalous multiplicity	14.6 (14.7)	7.0 (7.1)	
<i>Refinement</i>			
Reflections	12582	7846	6777
Protein atoms in model	624	3342	1405
Solvent atoms in model	58	0	14
R_{working}	0.204	0.213	0.173
R_{free}	0.221 ^a	0.277 ^b	0.241 ^c
RMSD from ideal geometry ^d			
Bond lengths (Å)	0.003	0.012	0.006
Bond angles (°)	0.67	0.28	0.98
Wilson B -factor (Å ²)	21.4	56.1	43.7
B -factor of protein atoms (Å ²)	33.6	88.8	31.2
B -factor of solvent atoms (Å ²)	44.2		17.2
Ramachandran plot ^d			
Most favoured (%)	97.4	95.8	97.0
Outlier (%)	0.0	0.44 ^e	0.6 ^f
PDB code	2Y1Y	2Y22	2Y1Z

Overall values are listed, with the outer-shell values in parentheses.

$$R_{\text{merge}} = \sum |I_i - bI_i N| / \sum I_i$$

$$R_{\text{working}} = \sum |F_o - F_c| / \sum F_o$$

R_{free} is the R -factor calculated for the cross-validated test set of reflections.

^a R_{free} is 4.67% of reflections.

^b R_{free} is 4.78% of reflections.

^c R_{free} is 10.4% of reflections.

^d As defined by MolProbity.

^e Outliers are Glu67 in chain A and Glu106 in chain C, in poor density.

^f Outlier is Asp109 in chain A, in poor density.

present in these data that has been used as an independent source of experimental phases to confirm the location of the poorly ordered chains. There are small but significant (3.5σ) electron density peaks corresponding to SeMet 137 for both poorly ordered chains and also for SeMet 68 from molecule E in anomalous difference maps calculated using peak wavelength data and phases from the final refined model with the two weak chains omitted. In addition, there is weak but significant density to support the location of the poorly ordered chains in maps calculated using phases optimised with Phaser from data collected at the peak wavelength and the partial molecular replacement solution consisting of the four well-ordered chains. This density is strongest where the two additional chains make crystal contacts with the rest of the structure, as might be expected.

Refinement was carried out with Refmac 5.6⁷¹ utilising local NCS restraints and the new Jelly Body refinement

protocol and was alternated with manual building using Coot. The final R and R_{free} were 21.3% and 27.7%, respectively. Density for possible MPD molecules was seen in the same location as for the high-resolution structure, close to Ser135 for the four ordered chains initially placed by Phaser. Additional density that could correspond to MPD molecules was seen on the other side of this serine for three of these molecules.

R120G/L137M (67–157)

Data were collected at ESRF, beamline 14.4, as 77 images, each with a 2° oscillation angle, at a wavelength of 0.87 Å. Data from 2.5 to 46.3 Å were integrated and indexed by Mosflm and then scaled with Scala (Table 2). The space group and point group were uncertain due to a combination of probable twinning and missing axial reflections. The apparent point group was $P6$, and the

likely space groups were $P3_n$ or $P3_121$, where the nature of the screw axis was not known. Solutions were determined by molecular replacement using the L137SeM structure (residues 71–157) in Phaser in a number of space groups. The solution in $P3_121$ refined to sensible R -factors. Building by hand in Coot and intensity-based twinning refinement in Refmac (final twin fraction of 47%) resulted in final R and R_{free} of 17.3% and 24.0%, respectively, with good geometry as determined by MolProbity.⁷² All structures were processed with the CCP4 program suite,⁷³ and images were generated with PyMOL.

PDB accession numbers

Coordinates and structure factors have been deposited in the PDB. For human α B R120G/L137M (67–157), the PDB code is 2Y1Z for the coordinate entry, and the code is R2Y1ZSF for the structure factors. For human α B L137SeM (67–157), the PDB code is 2Y22 for the coordinate entry, and the code is R2Y22SF for the structure factors. For human α B L137SeM (71–157), the PDB code is 2Y1Y for the coordinate entry, and the code is R2Y1YSF for the structure factors.

Supplementary materials related to this article can be found online at [doi:10.1016/j.jmb.2011.02.020](https://doi.org/10.1016/j.jmb.2011.02.020)

Acknowledgements

We thank Dr. Nora Cronin for help with data collection and the synchrotron beamline staff at the European Synchrotron Radiation Facility (Grenoble, France) and at the Diamond Light Source (Oxfordshire, UK). We thank Professor David Moss for helpful discussions on symmetry. We are grateful for the financial support of the Medical Research Council, UK (grant G0801846).

References

- Lindquist, S. (1986). The heat-shock response. *Annu. Rev. Biochem.* **55**, 1151–1191.
- Wickner, S., Maurizi, M. R. & Gottesman, S. (1999). Posttranslational quality control: Folding, refolding, and degrading proteins. *Science*, **286**, 1888–1893.
- Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. (2008). Adapting proteostasis for disease intervention. *Science*, **319**, 916–919.
- Hartl, F. U. & Hayer-Hartl, M. (2009). Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 574–581.
- Mchaourab, H. S., Godar, J. A. & Stewart, P. L. (2009). Structure and mechanism of protein stability sensors: chaperone activity of small heat shock proteins. *Biochemistry*, **48**, 3828–3837.
- Bukau, B., Weissman, J. & Horwich, A. (2006). Molecular chaperones and protein quality control. *Cell*, **125**, 443–451.
- Taylor, R. P. & Benjamin, I. J. (2005). Small heat shock proteins: a new classification scheme in mammals. *J. Mol. Cell. Cardiol.* **38**, 433–444.
- Quraishe, S., Asuni, A., Boelens, W. C., O'Connor, 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.
- Tebbenkamp, A. T. N. & Borchelt, D. R. (2010). Analysis of chaperone mRNA expression in the adult mouse brain by meta analysis of the Allen Brain Atlas. *PLoS One*, **5**, e13675.
- Kappe, G., Franck, E., Verschuure, P., Boelens, W. C., Leunissen, J. A. M. & de Jong, W. W. (2003). The human genome encodes 10 α -crystallin-related small heat shock proteins: HspB1–10. *Cell Stress Chaperones*, **8**, 53–61.
- Arrigo, A. P., Simon, S., Gibert, B., Kretz-Remy, C., Nivon, M., Czekalla, A. *et al.* (2007). Hsp27 (HspB1) and α B-crystallin (HspB5) as therapeutic targets. *FEBS Lett.* **581**, 3665–3674.
- Fuchs, M., Poirier, D. J., Seguin, S. J., Lambert, H., Carra, S., Charette, S. J. & Landry, J. (2010). Identification of the key structural motifs involved in HspB8/HspB6–Bag3 interaction. *Biochem. J.* **425**, 245–255.
- Powers, M. V., Clarke, P. A. & Workman, P. (2009). Death by chaperone HSP90, HSP70 or both? *Cell Cycle*, **8**, 518–526.
- Moyano, J. V., Evans, J. R., Chen, F., Lu, M. L., Werner, M. E., Yehiely, F. *et al.* (2006). α B-Crystallin is a novel oncoprotein that predicts poor clinical outcome in breast cancer. *J. Clin. Invest.* **116**, 261–270.
- Goplen, D., Bougnaud, S., Rajcevic, U., Boe, S. O., Skaftnesmo, K. O., Voges, J. *et al.* (2010). α B-Crystallin is elevated in highly infiltrative apoptosis-resistant glioblastoma cells. *Am. J. Pathol.* **177**, 1618–1628.
- Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N. H., Slingsby, C. & Tardieu, A. (2004). Ageing and vision: structure, stability and function of lens crystallins. *Prog. Biophys. Mol. Biol.* **86**, 407–485.
- De Vos, K. J., Grierson, A. J., Ackerley, S. & Miller, C. C. J. (2008). Role of axonal transport in neurodegenerative diseases. *Annu. Rev. Neurosci.* **31**, 151–173.
- Clark, J. I. & Muchowski, P. J. (2000). Small heat-shock proteins and their potential role in human disease. *Curr. Opin. Struct. Biol.* **10**, 52–59.
- Sun, Y. & MacRae, T. H. (2005). The small heat shock proteins and their role in human disease. *FEBS J.* **272**, 2613–2627.
- Goldbaum, O., Riedel, M., Stahnke, T. & Richter-Landsberg, C. (2009). The small heat shock protein HSP25 protects astrocytes against stress induced by proteasomal inhibition. *Glia*, **57**, 1566–1577.
- van Noort, J. M., Bsibsi, M., Gerritsen, W. H., van der Valk, P., Bajramovic, J. J., Steinman, L. & Amor, S. (2010). α B-Crystallin is a target for adaptive immune responses and a trigger of innate responses in preactive multiple sclerosis lesions. *J. Neuropathol. Exp. Neurol.* **69**, 694–703.
- Huang, B. & He, W. (2010). Molecular characteristics of inherited congenital cataracts. *Eur. J. Med. Genet.* **6**, 347–357.
- Selcen, D. (2010). Myofibrillar myopathies. *Curr. Opin. Neurol.* **23**, 477–481.

24. Goldfarb, L. G. & Dalakas, M. C. (2009). Tragedy in a heartbeat: malfunctioning desmin causes skeletal and cardiac muscle disease. *J. Clin. Invest.* **119**, 1806–1813.
25. Dierick, I., Irobi, J., De Jonghe, P. & Timmerman, V. (2005). Small heat shock proteins in inherited peripheral neuropathies. *Annals Med.* **37**, 413–422.
26. 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.
27. Haslbeck, M., Franzmann, T., Weinfurter, D. & Buchner, J. (2005). Some like it hot: the structure and function of small heat-shock proteins. *Nat. Struct. Mol. Biol.* **12**, 842–846.
28. de Jong, W. W., Caspers, G. J. & Leunissen, J. A. M. (1998). Genealogy of the α -crystallin–small heat-shock protein superfamily. *Int. J. Biol. Macromol.* **22**, 151–162.
29. Poulain, P., Gelly, J. C. & Flatters, D. (2010). Detection and architecture of small heat shock protein monomers. *PLoS One*, **5**, e9990.
30. Vicart, P., Caron, A., Guicheney, P., Li, Z., Prevost, M. C., Faure, A. *et al.* (1998). A missense mutation in the α B-crystallin chaperone gene causes a desmin-related myopathy. *Nat. Genet.* **20**, 92–95.
31. 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 α crystallin gene *CRYAA*. *Hum. Mol. Genet.* **7**, 471–474.
32. Irobi, J., van Impe, K., Seeman, P., Jordanova, A., Dierick, I., Verpoorten, N. *et al.* (2004). Hot-spot residue in small heat-shock protein 22 causes distal motor neuropathy. *Nat. Genet.* **36**, 597–601.
33. Bova, M. P., Yaron, O., Huang, Q. L., Ding, L. L., Haley, D. A., Stewart, P. L. & Horwitz, J. (1999). Mutation R120G in α B-crystallin, which is linked to a desmin-related myopathy, results in an irregular structure and defective chaperone-like function. *Proc. Natl Acad. Sci. USA*, **96**, 6137–6142.
34. Michiel, M., Skouri-Panet, F., Duprat, E., Simon, S., Ferard, C., Tardieu, A. & Finet, S. (2009). Abnormal assemblies and subunit exchange of α B-crystallin R120 mutants could be associated with destabilization of the dimeric substructure. *Biochemistry*, **48**, 442–453.
35. Horwitz, J. (2009). Alpha crystallin: The quest for a homogeneous quaternary structure. *Exp. Eye Res.* **88**, 190–194.
36. Kim, K. K., Kim, R. & Kim, S. H. (1998). Crystal structure of a small heat-shock protein. *Nature*, **394**, 595–599.
37. 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.
38. Haley, D. A., Bova, M. P., Huang, Q. L., Mchaourab, H. S. & Stewart, P. L. (2000). Small heat-shock protein structures reveal a continuum from symmetric to variable assemblies. *J. Mol. Biol.* **298**, 261–272.
39. White, H. E., Orlova, E. V., Chen, S. X., Wang, L. C., Ignatiou, A., Gowen, B. *et al.* (2006). Multiple distinct assemblies reveal conformational flexibility in the small heat shock protein Hsp26. *Structure*, **14**, 1197–1204.
40. Bagn eris, C., Bateman, O. A., Naylor, C. E., Cronin, N., Boelens, W. C., Keep, N. H. & Slingsby, C. (2009). Crystal structures of α -crystallin domain dimers of α B-crystallin and Hsp20. *J. Mol. Biol.* **392**, 1242–1252.
41. Laganowsky, A., Benesch, J. L. P., Landau, M., Ding, L. L., Sawaya, M. R., Cascio, D. *et al.* (2010). Crystal structures of truncated α A and α B crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci.* **19**, 1031–1043.
42. Laganowsky, A. & Eisenberg, D. (2010). Non-3D domain swapped crystal structure of truncated zebrafish α A crystallin. *Protein Sci.* **19**, 1978–1984.
43. Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217.
44. Jehle, S., Rajagopal, P., Bardiaux, B., Markovic, S., Kuhne, R., Stout, J. R. *et al.* (2010). Solid-state NMR and SAXS studies provide a structural basis for the activation of α B-crystallin oligomers. *Nat. Struct. Mol. Biol.* **17**, 1037–1042.
45. Sharma, K. K., Kumar, G. S., Murphy, A. S. & Kester, K. (1998). Identification of 1,1'-bi(4-anilino)naphthalene-5,5'-disulfonic acid binding sequences in α -crystallin. *J. Biol. Chem.* **273**, 15474–15478.
46. Ghosh, J. G., Estrada, M. R. & Clark, J. I. (2006). Structure-based analysis of the β 8 interactive sequence of human α B crystallin. *Biochemistry*, **45**, 9878–9886.
47. Gamerdinger, M., Hajieva, P., Kaya, A. M., Wolfrum, U., Hartl, F. U. & Behl, C. (2009). Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3. *EMBO J.* **28**, 889–901.
48. Jehle, S., van Rossum, B., Stout, J. R., Noguchi, S. M., Falber, K., Rehbein, K. *et al.* (2009). α B-Crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. *J. Mol. Biol.* **385**, 1481–1497.
49. Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797.
50. Sugiyama, Y., Suzuki, A., Kishikawa, M., Akutsu, R., Hirose, T., Wayne, M. M. Y. *et al.* (2000). Muscle develops a specific form of small heat shock protein complex composed of MKBP/HSPB2 and HSPB3 during myogenic differentiation. *J. Biol. Chem.* **275**, 1095–1104.
51. Bova, M. P., Mchaourab, H. S., Han, Y. & Fung, B. K. K. (2000). Subunit exchange of small heat shock proteins—analysis of oligomer formation of α A-crystallin and Hsp27 by fluorescence resonance energy transfer and site-directed truncations. *J. Biol. Chem.* **275**, 1035–1042.
52. Bukach, O. V., Glukhova, Seit-Nebi, A. S. & Gusev, N. B. (2009). Heterooligomeric complexes formed by human small heat shock proteins HspB1 (Hsp27) and HspB6 (Hsp20). *Biochim. Biophys. Acta*, **1794**, 486–495.
53. Muchowski, P. J., Hays, L. G., Yates, J. R. & Clark, J. I. (1999). ATP and the core “ α -crystallin” domain of the small heat-shock protein α B-crystallin. *J. Biol. Chem.* **274**, 30190–30195.
54. Sobott, F., Benesch, J. L. P., Vierling, E. & Robinson, C. V. (2002). Subunit exchange of multimeric protein complexes—real-time monitoring of subunit exchange between small heat shock proteins by using electrospray mass spectrometry. *J. Biol. Chem.* **277**, 38921–38929.

55. 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.
56. Stengel, F., Baldwin, A. J., Painter, A. J., Jaya, N., Basha, E., Kay, L. E. *et al.* (2010). Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proc. Natl Acad. Sci. USA*, **107**, 2007–2012.
57. 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 α by phosphorylation. *J. Biol. Chem.* **274**, 18947–18956.
58. Hayes, D., Napoli, V., Mazurkie, A., Stafford, W. F. & Graceffa, P. (2009). Phosphorylation dependence of Hsp27 multimeric size and molecular chaperone function. *J. Biol. Chem.* **284**, 18801–18807.
59. Liang, J. J. & Liu, B. F. (2006). Fluorescence resonance energy transfer study of subunit exchange in human lens crystallins and congenital cataract crystallin mutants. *Protein Sci.* **15**, 1619–1627.
60. Jaya, N., Garcia, V. & Vierling, E. (2009). Substrate binding site flexibility of the small heat shock protein molecular chaperones. *Proc. Natl Acad. Sci. USA*, **106**, 15604–15609.
61. Stamler, R., Kappe, G., Boelens, W. & Slingsby, C. (2005). Wrapping the α -crystallin domain fold in a chaperone assembly. *J. Mol. Biol.* **353**, 68–79.
62. 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 α B-crystallin. *Biochemistry*, **46**, 9605–9614.
63. Kopito, R. R. (2000). Aggregosomes, inclusion bodies and protein aggregation. *Trends Cell Biol.* **10**, 524–530.
64. Maloyan, A., Gulick, J., Glabe, C. G., Kaye, R. & Robbins, J. (2007). Exercise reverses preamyloid oligomer and prolongs survival in α B-crystallin-based desmin-related cardiomyopathy. *Proc. Natl Acad. Sci. USA*, **104**, 5995–6000.
65. Irobi, J., Almeida-Souza, L., Asselbergh, B., De Winter, V., Goethals, S., Dierick, I. *et al.* (2010). Mutant HSPB8 causes motor neuron-specific neurite degeneration. *Hum. Mol. Gen.* **19**, 3254–3265.
66. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 ESF-EAMCB Newslett. Protein Crystallogr.* No. 26.
67. Evans, P. R. (2005). Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 72–82.
68. 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.
69. 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.
70. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2126–2132.
71. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255.
72. 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.
73. Collaborative Computational Project, No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.