Structure, Volume *19*

Supplemental Information

The Polydispersity of αB-Crystallin Is Rationalized by an Interconverting Polyhedral Architecture

Andrew J. Baldwin, Hadi Lioe, Gillian R. Hilton, Lindsay A. Baker, John L. Rubinstein, Lewis E. Kay, and Justin L.P. Benesch

Inventory of Supplemental Information

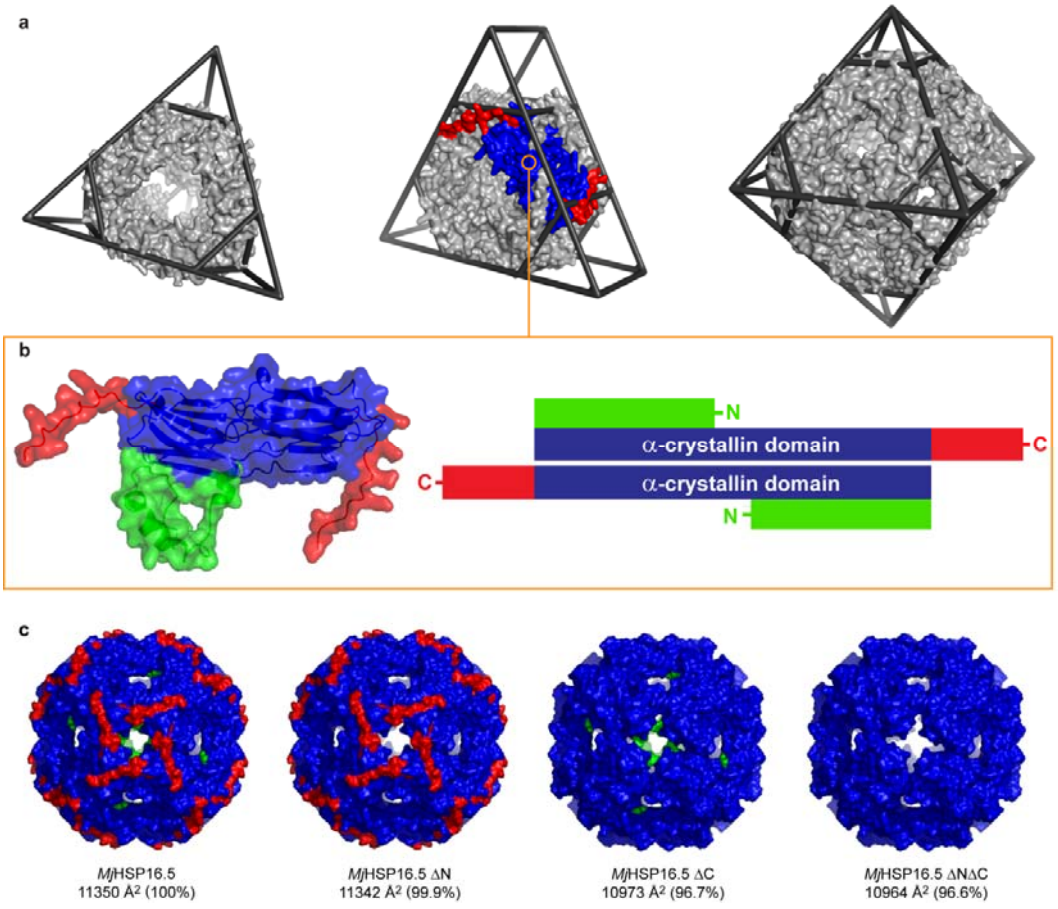
Supplementary Figure 1: related to Figure 1

Supplementary Figure 2: related to Figure 2

Supplementary Figure 3: related to Figure 4

Supplementary Tables 1-3: related to Experimental Procedures

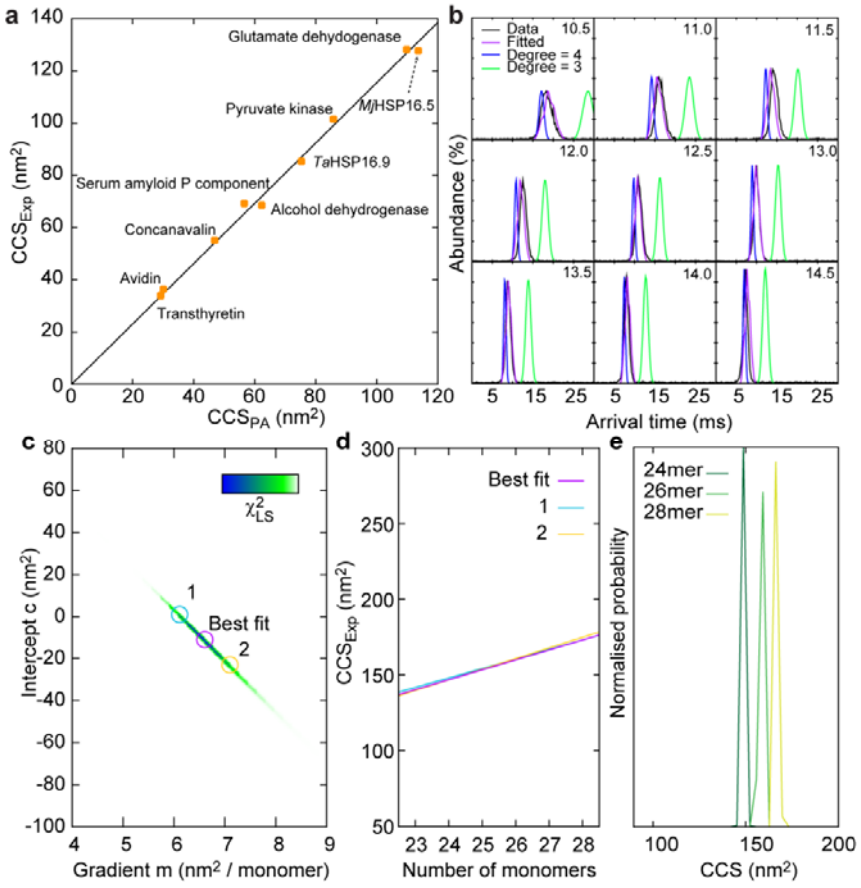
Supplementary Figure 1 (related to Figure 1)



Oligomeric structures of sHSPs display polyhedral and ring-like topology. Acr1 from *Mycobacterium tuberculosis* (Kennaway et al., 2005) (left), HSP16.9 from *Triticum aestivum* (van Montfort et al., 2001) (middle), and HSP16.5 from *Methanocaldococcus jannaschii* (Kim et al., 1998) (right) can be represented as a tetrahedron, double triangular ring, and octahedron, respectively (a). Note that a double triangular ring can also be thought of as a triangular prism which is missing the edges connecting the triangular faces. In each of these shapes the long edges are defined by a dimeric building block (b), composed of an α-crystallin core flanked by C-terminal arms which

link dimers together. Though relatively little is known about the N-terminal region of the sHSPs, the crystal structure of HSP16.9 (a, middle; b) and measurements on HSP16.5 (Koteiche et al., 2005) reveal it to be sequestered on the inside of the oligomers, as a recent study has similarly suggested for αB-crystallin (Jehle et al., 2011). These polyhedral and ring-like geometries therefore represent attractive scaffolds for constructing oligomeric models for αB-crystallin, starting with high-resolution structures of the truncated dimer as a building block. To assess whether the missing sequence would affect the CCS, we calculated CCS_{PA} values from the pdb file of HSP16.5 from *Methanocaldococcus jannaschii* (Kim et al., 1998) (left), or from pdb files with the flanking N- (centre left), C- (centre right), or both N- and C- terminal (right) sequences removed (c). The *in silico* truncation of the terminal regions causes only a small difference in CCS_{PA}, similar to the error of our measurement. Therefore calculated CCS values obtained from oligomeric models for αB-crystallin constructed from the dimeric αB-crystallin core (Bagneris et al., 2009) can be directly compared with the corresponding CCSs obtained from experiments recorded on the full-length protein.

Supplementary Figure 2 (related to Figure 2)



Comparison of experimental and in silico CCSs and ATDs. There is an excellent correlation between the CCSs calculated from high-resolution structures of protein complexes according to the projection approximation and those measured experimentally (a) (Benesch and Ruotolo, 2011). These measured CCSs are averaged over all charge states, as the difference in CCS for different charge states is small for large, highly charged protein assemblies (Bush et al., 2010). This allows for the direct comparison of our modelled structures with experimental measurements (b), performed at a range of wave heights, and using the equations in Experimental Procedures. Degree

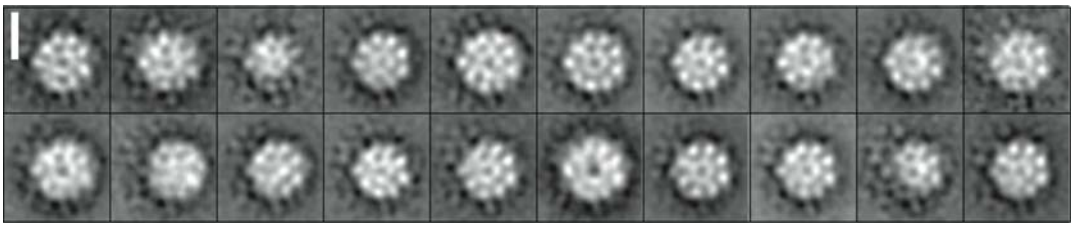
4 polyhedra (blue) fit the experimental data (black) much better than degree 3 (green), and the match between free fitting (purple) of CCS with oligomer size is excellent. This fitting is performed assuming that, to first approximation, CCS varies linearly with the number of monomers in the oligomer, *N*, such that $CSS_{(N)}=mN+c$ where *m* and *c*, the gradient and the intercept were two of the three free parameters in the fit of ATD versus subunit number (the third was the resolving power, κ, see Experimental Procedures). No statistical improvement to the fit was obtained by using a more complex model, such as $CSS_{(N)}=aN^2+mN+c$, for the trend in CCS with the oligomer stoichiometry. A plot of the Bayesian probability* of a given *m* and *c* being the 'true' solution reveals the strong correlation between gradient and an intercept expected when performing a linear regression (c). The most probable values of *m* and *c* were determined to be 6.6 $nm^2 monomer^{-1}$ and -11 nm^2, respectively. A plot of CCS versus number of monomers are shown for the three regions indicated in (c) on a scale comparable to Fig. 2b (d). The CCS values in the region of interest, *N*=22 to 29, are in very close agreement to those expected with a probability of ca. 50 % that of the maximum, denoted by the locations 1 and 2 in (c). The probability of a given CCS being the 'true' solution for the 24, 26 and 28mer (e). The standard deviations of these histograms indicate that the precision of individual CCS measurements are within 3% of the mean values.

* **Method.** The best fitting values of *m* and *c* were determined by minimising the least squared residual $\chi^2_{\mathrm{LS}} = \sum^{\mathrm{n}} \left(\mathrm{y}_{\mathrm{exp}} - \mathrm{y}_{\mathrm{calc}} \right)^2$ where *y* denotes a measured intensity of the ATD at a given drift time and wave height, and the sum is over all *n* data points (200 drift times per wave height, and 9 wave heights, Supplementary Fig. 2b). An estimate in the uncertainty in each individual intensity, σ, can be obtained by assuming that the

reduced χ^2 of the overall fit is unity, leading to $\sigma^2 \cong \frac{\chi^2_{LS}}{n-p}$ where *n* is the number of data points and *p* is the number of free parameters in the fit. By this method, each measured ATD value is predicted to have an uncertainty of ca. 3% of the maximum of each given ATD.

The Bayesian probability, *P*, of a given set of *m* and *c* fitting parameters being the 'true' solution is then proportional to $\exp\left(-\chi^2_{LS}/\left(2\sigma^2\right)\right)$, shown as a function of *m* and *c* in (a). The probability of a given *CCS* being the 'true' value for a given oligomer was estimated using the following method. For an oligomer size of interest, the *CCS* at each *m* and *c* position on the surface shown in (a) is evaluated, and the relevant bin of a corresponding *CCS* histogram is incremented by the probability associated with that combination of *m* and *c* values. The resulting *CCS* histograms for the 24, 26 and 28mers are shown. The standard deviation of the *CCS* histograms, an indication of the precision of the individual *CCS* measurements, was found to be within 3% of the mean values. The experimentally determined CCSs of the 24mer, 26mer and 28mer were 148 ± 3, 160 ± 4 and 174 ± 5 nm^2 respectively, where the uncertainties scale with the relative abundance of the oligomer in the peak undergoing analysis (Fig. 2f).

Supplementary Figure 3 (related to Figure 4)



Effect of bin size in classification of αB-crystallin particles. Negative stain EM was performed on αB-crystallin, with 3422 particle images selected for further analysis. The class averages obtained when the particles are initially band-pass filtered to 12 Å (Jehle et al., 2011; Peschek et al., 2009) and separated into 20 classes show nearly identical features, and an apparently narrow size distribution. Scale bar is 10 nm. This is in clear contrast to the heterogeneity observed in the individual particles (Fig. 4a). Relative to when the separation is performed into 200 classes (Fig. 4b), though the average radius remains the same, the standard deviation is reduced by a factor of two. This demonstrates the importance of sorting the particles into a sufficiently large number of classes such that the heterogeneity in the sample is properly accounted for. Here, approximately an order of magnitude larger than the number of oligomeric states populated in the sample, which in the case of αB-crystallin is in excess of 20, was required to clearly see the heterogeneity.

Supplementary Table 1 (related to Experimental Procedures)

Wave height (V)	***a***	***b***	***c***
10.5	3180.121	81.381	0.371
11.0	3011.959	88.199	0.373
11.5	2912.105	93.652	0.374
12.0	2876.519	100.088	0.366
12.5	2951.010	102.939	0.368
13.0	2925.148	112.015	0.346
13.5	3052.081	113.581	0.350
14.0	3243.924	113.544	0.356
14.5	3398.545	116.570	0.348

Calibration parameters for converting IM-MS arrival times to CCS_{Exp} values, as described in Experimental Procedures, for a range of wave heights.

Supplementary Table 2 (related to Experimental Procedures)

Class	Name	# Dimers	CCS_{PA} (A^2)
Mixed degree			
Polyhedron	AugmentedHexagonalPrism	22	25929
Polyhedron	AugmentedPentagonalPrism	19	18751
Polyhedron	AugmentedTriangularPrism	13	13596
Polyhedron	AugmentedTridiminishedIcosahedron	18	17949
Polyhedron	BiaugmentedPentagonalPrism	23	19873
Polyhedron	BiaugmentedTriangularPrism	17	15708
Polyhedron	Dipyramid(3)	9	8541
Polyhedron	Dipyramid(5)	15	17323
Polyhedron	ElongatedPentagonalDipyramid	25	21310
Polyhedron	ElongatedPentagonalPyramid	20	18547
Polyhedron	ElongatedSquarePyramid	16	15354
Polyhedron	ElongatedTriangularDipyramid	15	14598
Polyhedron	ElongatedTriangularPyramid	12	13468
Polyhedron	Gyrobifastigium	14	15015
Polyhedron	GyroelongatedPentagonalPyramid	25	23695
Polyhedron	GyroelongatedSquareDipyramid	24	23970
Polyhedron	GyroelongatedSquarePyramid	20	20069
Polyhedron	MetabidiminishedIcosahedron	20	20545
Polyhedron	PentagonalCupola	25	34560
Polyhedron	Pyramid(4)	8	9181
Polyhedron	Pyramid(5)	10	11488
Polyhedron	RhombicDodecahedron	24	23761
Polyhedron	SnubDisphenoid	18	14969
Polyhedron	Sphenocorona	22	18059
Polyhedron	SquareCupola	20	22189
Polyhedron	TriakisTetrahedron	18	16234
Polyhedron	TriangularCupola	15	16260
Polyhedron	TriaugmentedTriangularPrism	21	19848
Polyhedron	TridiminishedIcosahedron	15	16025
Degree = 4			
Polyhedron	Antiprism(4)	16	15742
Polyhedron	Antiprism(5)	20	20615
Polyhedron	Antiprism(6)	24	24454
Polyhedron	Cuboctahedron	24	23348
Polyhedron	ElongatedSquareDipyramid	20	18552
Polyhedron	Octahedron	12	12443
Polyhedron	TriangularOrthobicupola	24	21200
Degree = 3			
Polyhedron	Prism(3)	9	10805

Polyhedron	Cube	12	14860
Polyhedron	Prism(5)	15	17662
Polyhedron	Prism(6)	18	21693
Polyhedron	Prism(7)	21	24405
Polyhedron	Prism(8)	24	27946
Polyhedron	Tetrahedron	6	7310
Polyhedron	TruncatedTetrahedron	18	21304
Rings			
Single ring	SingleTriangle	3	4112
Single ring	SingleSquare	4	5241
Single ring	SinglePentagon	5	6972
Single ring	SingleHexagon	6	8295
Single ring	SingleHeptagon	7	9694
Single ring	SingleOctagon	8	11483
Single ring	SingleNonagon	9	12688
Single ring	SingleDecagon	10	14338
Single ring	SingleHendecadon	11	15968
Single ring	SingleDodecagon	12	17284
Single ring	SingleTridecagon	13	18653
Single ring	SingleTetradecagon	14	20523
Single ring	SinglePentadecagon	15	21518
Single ring	SingleHexadecagon	16	23590
Double ring	DoubleTriangle	6	5927
Double ring	DoubleSquare	8	7497
Double ring	DoublePentagon	10	9565
Double ring	DoubleHexagon	12	11408
Double ring	DoubleHeptagon	14	13612
Double ring	DoubleOctagon	16	16087
Double ring	DoubleNonagon	18	17765
Triple ring	TripleTriangle	9	7176
Triple ring	TripleSquare	12	9272
Triple ring	TriplePentagon	15	12804
Quadruple ring	QuadrupleTriangle	12	8827
Quadruple ring	QuadrupleSquare	16	11216

Database of candidate topologies for modelling αB-crystallin oligomers, the number of dimers required and the corresponding CCS calculated from the pdb file. This CCS_{PA} is subsequently scaled as described in the methods to give figure 2b. In all cases 2WJ7 (Bagneris et al., 2009) was used, as illustrated in figure 2a. For the pyramids, the dipyramids, the prisms and the antiprisms, the number of sides in the main face is indicated in brackets. The polyhedra with degree exactly equal to 3 and 4 are indicated.

Supplementary Table 3 (related to Experimental Procedures)

Dimer	**CCS_{PA} (A^2)**	
PDB ID	**0°**	**90°**
2WJ7	12443.21	13573.45
3L1G	11788.29	12363.98
2KLR	10843.86	12159.17

CCS values for obtained for an octahedral structure built using three alternative high resolution structures of the dimer. These structures differ slightly in terms of register at the intra-dimer interface, curvature, and loop position (Bagneris et al., 2009; Jehle et al., 2010; Laganowsky et al., 2010). Structures are built as depicted in figure 2a with the large face of the cuboid on the outside of the oligomer (column 0°), or with the dimer rotated 90° about the edge. In the manuscript the dimer and orientation in grey is used, but all other possibilities lie within 12% of this area.

Supplementary References

Bagneris, C., Bateman, O.A., Naylor, C.E., Cronin, N., Boelens, W.C., Keep, N.H., and Slingsby, C. (2009). Crystal structures of alpha-crystallin domain dimers of alphaB-crystallin and Hsp20. Journal of molecular biology *392*, 1242-1252.

Benesch, J.L.P., and Ruotolo, B.T. (2011). Mass Spectrometry: an Approach Come-of-Age for Structural and Dynamical Biology. Current opinion in structural biology *in press*.

Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C.V., and Ruotolo, B.T. (2010). Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. Anal Chem *82*, 9557-9565.

Jehle, S., Rajagopal, P., Bardiaux, B., Markovic, S., Kuhne, R., Stout, J.R., Higman, V.A., Klevit, R.E., van Rossum, B.J., and Oschkinat, H. (2010). Solid-state NMR and SAXS studies provide a structural basis for the activation of alphaB-crystallin oligomers. Nat Struct Mol Biol *17*, 1037-1042.

Jehle, S., Vollmar, B.S., Bardiaux, B., Dove, K.K., Rajagopal, P., Gonen, T., Oschkinat, H., and Klevit, R.E. (2011). N-terminal domain of alphaB-crystallin provides a conformational switch for multimerization and structural heterogeneity. Proc Natl Acad Sci U S A *108*, 6409-6414.

Kennaway, C.K., Benesch, J.L., Gohlke, U., Wang, L., Robinson, C.V., Orlova, E.V., Saibil, H.R., and Keep, N.H. (2005). Dodecameric structure of the small heat shock protein Acr1 from Mycobacterium tuberculosis. J. Biol. Chem. *280*, 33419-33425.

Kim, K.K., Kim, R., and Kim, S.H. (1998). Crystal structure of a small heat-shock protein. Nature *394*, 595-599.

Koteiche, H.A., Chiu, S., Majdoch, R.L., Stewart, P.L., and McHaourab, H.S. (2005). Atomic models by cryo-EM and site-directed spin labeling: application to the N-terminal region of Hsp16.5. Structure *13*, 1165-1171.

Laganowsky, A., Benesch, J.L.P., Landau, M., Ding, L.L., Sawaya, M., Cascio, D., Huang, Q., Robinson, C.V., Horwitz, J., and Eisenberg, D. (2010). Crystal Structures of Truncated AlphaA and AlphaB Crystallins Reveal Structural Mechanisms of Polydispersity Important for Eye Lens Function. Protein Science *19*, 1031-1043.

Peschek, J., Braun, N., Franzmann, T.M., Georgalis, Y., Haslbeck, M., Weinkauf, S., and Buchner, J. (2009). The eye lens chaperone alpha-crystallin forms defined globular assemblies. Proc Natl Acad Sci U S A *106*, 13272-13277.

van Montfort, R.L., Basha, E., Friedrich, K.L., Slingsby, C., and Vierling, E. (2001). Crystal structure and assembly of a eukaryotic small heat shock protein. Nat Struct Biol *8*, 1025-1030.