Non-3D domain swapped crystal structure of truncated zebrafish alphaA crystallin

A. Laganowsky\textsuperscript{1,2,3} and D. Eisenberg\textsuperscript{1,2,3}\textsuperscript{*}

\textsuperscript{1}UCLA-DOE, Institute for Genomics and Proteomics, Los Angeles, California
\textsuperscript{2}Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, California
\textsuperscript{3}Howard Hughes Medical Institute

Received 21 May 2010; Revised 5 July 2010; Accepted 6 July 2010
DOI: 10.1002/pro.471
Published online 28 July 2010 proteinscience.org

Abstract: In previous work on truncated alpha crystallins (Laganowsky et al., Protein Sci 2010; 19:1031–1043), we determined crystal structures of the alpha crystallin core, a seven beta-stranded immunoglobulin-like domain, with its conserved C-terminal extension. These extensions swap into neighboring cores forming oligomeric assemblies. The extension is palindromic in sequence, binding in either of two directions. Here, we report the crystal structure of a truncated alphaA crystallin (AAC) from zebrafish (\textit{Danio rerio}) revealing C-terminal extensions in a non-three-dimensional (3D) domain swapped, “closed” state. The extension is quasi-palindromic, bound within its own zebrafish core domain, lying in the opposite direction to that of bovine AAC, which is bound within an adjacent core domain (Laganowsky et al., Protein Sci 2010; 19:1031–1043). Our findings establish that the C-terminal extension of alpha crystallin proteins can be either 3D domain swapped or non-3D domain swapped. This duality provides another molecular mechanism for alpha crystallin proteins to maintain the polydispersity that is crucial for eye lens transparency.

Keywords: X-ray diffraction; small heat shock protein; protein chaperone; cataract; eye lens transparency; alpha crystallin

Introduction

Alpha crystallins are eye lens proteins functioning in light refraction and in maintaining lens transparency. In zebrafish (\textit{Danio rerio}), the alpha crystallins comprise 7–22\% of eye lens proteins depending on the age of the fish\textsuperscript{1,3}, a lower percentage than the up to 50\% reached in some mammalian lenses.\textsuperscript{4} Although the percentage of alpha crystallin is lower in zebrafish, the ratio of alphaA crystallin (AAC) to alphaB crystallin (ABC) is roughly 3:1 for zebrafish, similar to human.\textsuperscript{2,5} In function, zebrafish AAC (zfAAC) expression has been shown to rescue gamma crystallin protein insolubility in the zebrafish \textit{cloche} mutant, preventing cataract formation.\textsuperscript{6} This zebrafish \textit{cloche} mutant highlights the importance of AAC in eye lens development.
Recently, both we and Bagneris et al.\textsuperscript{7} discovered that truncated constructs of alpha crystallins form crystals. This discovery led to two crystal structures of truncated human ABC: one of the alpha crystallin core domain, residues 67–157,\textsuperscript{7} and the other, residues 68–162 (bABC\textsubscript{68–162}),\textsuperscript{1} containing the conserved C-terminal extension. Also, there are two crystal structures for a truncated bovine AAC, residues 59–163 (hABC\textsubscript{59–163}),\textsuperscript{1} containing the C-terminal extension.

From our previous work,\textsuperscript{1} we found that alpha crystallins encode polydispersity that supports eye lens transparency and chaperone function. A dimeric interface is formed by an antiparallel beta sheet,\textsuperscript{1,7–11} called the AP interface.\textsuperscript{1} This AP interface appears dynamic with three distinct registration states. The three registration states have been defined according to their degree of surface area buried as AP\textsubscript{I}, AP\textsubscript{II}, and AP\textsubscript{III},\textsuperscript{7} with AP\textsubscript{I} registration state having the greatest surface area buried. The C-terminal extension can structurally be divided into two sections; hinge loop and tail.\textsuperscript{1} The hinge loop forms a flexible loop that allows binding to neighboring molecules. For example, in one of the hABC\textsubscript{59–163} structures, the hinge loop is disordered, and in the second, the hinge loop is ordered, positioning a key histidine residue involved in a Zinc-binding motif (hABC\textsubscript{59–163}–Zn).\textsuperscript{1} The C-terminal tail contains a conserved palindromic sequence motif that allows for bidirectional binding.\textsuperscript{1} Taken together, all of these features enforce polydisperse and short-range interactions within the alpha crystallins in support of eye lens function.\textsuperscript{1,12–15}

Three-dimensional (3D) domain swapping has been observed for alpha crystallin and related proteins. In 3D domain swapping, a protein domain breaks its noncovalent bonds with the core domain and swaps into the same environment in an identical molecule, producing an oligomeric assembly (for review see Refs. 16–20). In crystal structures of truncated alpha crystallins (bAAC\textsubscript{59–163} and hABC\textsubscript{58–162}), the C-terminal extension is 3D domain swapped as in homologous structures (Wheat Hsp16.9\textsuperscript{21} and Methanococcus janaschii Hsp16.5\textsuperscript{22}). In these structures, the C-terminal extension is swapped into adjacent molecules forming run-away oligomeric assemblies. In the case of the homologous structures, Wheat Hsp16.9 and \textit{M. janaschii} Hsp 16.5, the C-terminal extensions 3D domain swap to form oligomeric assemblies.

Other eye lens crystallins also 3D domain swap. Gamma and beta crystallins contain two domains, with each domain containing two Greek key structural motifs, which are joined by a linker (for review see Ref. 23). For example, in the structure of \gamma-crystallin II monomer,\textsuperscript{24} the linker between domains supports non-3D domain swapping to form a closed monomer.\textsuperscript{19} For beta crystallin, the amino acid sequence and linker support both 3D domain swapping and non-3D domain swapped closed structures. For example, a 3D domain swapped dimer of \beta B2-crystallin, with an extended linker, assembles into a tetrameric oligomer.\textsuperscript{25,26} \beta B1-crystallin, with sequence identity of 58% to \beta B2-crystallin,\textsuperscript{26} forms a dimer without 3D domain swapping.\textsuperscript{25,26} Furthermore, a homologous protein of the beta–gamma crystallin family, nitrollin from \textit{Nitrosospira multiformis}, contains a single domain that forms a dimer through 3D domain swapping.\textsuperscript{28}

Figure 1. Amino acid sequence alignment of truncated alpha crystallins from zebrafish (\textit{Danio rerio}), antartic cod (\textit{Dissostichus mawsoni}), human (\textit{Homo sapiens}), and bovine (\textit{Bos taurus}). Residue numbers are listed below amino acid sequences. Beta-strand assignments are shown at top of alignment. Identical residues are shaded in gray and similar residues are shaded in light gray.
Here, we present the first structure of a truncated alpha crystallin with C-terminal extension in a non-3D domain swapped, closed state, and present evidence that the closed monomer can transform into a 3D domain swapped dimer.

**Results**

Our truncated form of zfAAC contains residues 60–166 (zfAAC60–166) including the C-terminal quasi-palindromic motif, with amino acid sequence of DRTIPVTDK. The truncated amino acid sequence of zfAAC is highly similar to AAC of antartic cod (Dissostichus mawsoni) Fig. 1. Both bovine and human AAC show major differences to zfAAC in the hinge loop region Fig. 1.

The structure of zfAAC60–166 reveals the first alpha crystallin structure having non-3D domain swapping of the C-terminal extension (Fig. 2 and Table 1). There are two chains located in the asymmetric unit cell of the immunoglobulin-like alpha crystallin core domain of seven beta strands that form a tetrameric assembly. Within this tetrameric assembly, the C-terminal extensions extend from the core domain wrapping around to cap their own domain by binding of the quasi-palindromic C-terminal tail. This non-3D domain swapping of the C-terminal extension is in contrast to the 3D domain swapped

![Figure 2](image2.png)

Figure 2. Crystal structure of truncated zebrafish alphaA crystallin (zfAAC60–166). Four molecules form a tetramer mediated through a previously unobserved dimeric interface (AB). The asymmetric unit contains two chains that form a dimeric interaction through an interface (AP) common to all vertebrate small heat shock proteins of known structure, with a registration state of AP1. Shown are four molecules generated by the twofold symmetry operation on the dimer of the asymmetric unit. N- and C-termini labels are color coded for each chain. A: A total of four chains form the closed tetramer were each C-terminal extension binds to its own domain (DS). The inset is an enlarged view of the dimeric interface, AB, created by beta strand i2 and the loop, located in between strands i5 and i6. Residue numbers are shown. Hydrogen bonds are shown by dashed yellow lines. B: A 90° rotation of the closed tetramer about a horizontal axis.

![Figure 3](image3.png)

Figure 3. Molecular interactions of the non-3D domain swapped C-terminal extension. The C-terminal extension, residues 146–165, backbone carbon atoms are colored in rainbow with carbonyl oxygen atoms red and nitrogen atoms blue. Remaining residues are shown as a gray ribbon. Hydrogen bonds are shown by dashed yellow lines and residues are labeled. A side view (panel A) and top view (panel B) are shown. Extending from the crystallin domain, the hinge loop forms a beta hairpin structure. R154 and the carbonyl of G153 of the structured hinge loop form hydrogen bonds with the crystallin core domain and position the C-terminal tail. The C-terminal tail binds in the orientation opposite to that in baAAC59–163, which is 3D domain swapped. Three backbone hydrogen bonds are made each by residues 157–159 and residues 162–164, located on both sides for the C-terminal tail. Additional hydrogen bonds are made by the side chain of D156 to the amide of D92, the side chain of R154 to the side chain of D93 and carboxyl of S123, and the side chain of R157 to carboxyl of I130 and Q127. Hydrophobic interactions are formed by the burial of I159 and V161 into a hydrophobic groove, created by strands i4 and i8, located on top of the alpha crystallin domain.
Statistics for Truncated Zebrafish AlphaA (zfAAC60–166) Crystallin

<table>
<thead>
<tr>
<th>Parameter</th>
<th>zfAAC60–166</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crystal parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P4$_3$2$_1$2</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td>$a = b = 33.29$, $c = 354.33$</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>$\alpha = \beta = \gamma = 90$</td>
</tr>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
</tr>
<tr>
<td>Synchrotron (beam line)</td>
<td>APS (24-ID-E)</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>27.82–1.75</td>
</tr>
<tr>
<td>$R_{free}$ (%)</td>
<td>6.3 (34.4)</td>
</tr>
<tr>
<td>I/σ</td>
<td>23.7 (4.3)</td>
</tr>
<tr>
<td>Reflections observed/unique</td>
<td>148,485/21,407</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98 (99)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
</tr>
<tr>
<td>$R_{work}$ (%)</td>
<td>20.5</td>
</tr>
<tr>
<td>$R_{free}$ (%)</td>
<td>24.8$^b$</td>
</tr>
<tr>
<td>Number of protein atoms</td>
<td>1705</td>
</tr>
<tr>
<td>Number of water molecules</td>
<td>172</td>
</tr>
<tr>
<td>RMSD bond/angle (Å/°)</td>
<td>0.007/1.11</td>
</tr>
<tr>
<td>Protein atoms average</td>
<td>24</td>
</tr>
<tr>
<td>B-factor (Å$^2$)</td>
<td>95/4.6</td>
</tr>
<tr>
<td>Ramachandran map, favored/allowed (%)</td>
<td>95/4.6</td>
</tr>
<tr>
<td>PDB accession code</td>
<td>3N3E</td>
</tr>
</tbody>
</table>

Values in parentheses correspond to the highest resolution shell.

$^a$ $R_{free} = \Sigma |I-I'|/\Sigma I$.

$^b$ $R_{free}$ calculated using 5% of the data.

C-terminal extensions of hAAC59–163 and hABC68–162, which are bound to adjacent molecules.

**Dimeric interface**

A tetrameric assembly is created by two dimeric interfaces. One dimeric interface, termed AP, is created by two identical beta strands on the neighboring molecules (Fig. 2). The registration of the two strands is of the AP$_{12}^3$ type we observed in hAAC59–163. A previously unobserved type of dimeric interface is created through an interaction of edge beta strands, $\beta 2$ of the alpha crystallin domain, to form the tetrameric assembly. This dimeric interface is further stabilized by residues located in the loop region between $\beta 5$ and $\beta 6 \pm 7$. These loops bend inward toward the $\beta 2$ beta strands forming hydrogen bonds through residues aspartate 107 and histidine 108 to serine 67, asparagine 124, and lysine 146 [Fig. 2(A), inset].

**Non-3D domain swapped C-terminal extension**

The C-terminal extension of zfAAC60–166 is unique among the alpha crystallins of known structure in that it is non-3D domain swapped. Extending outward the C-terminal extension wraps around to bind to the top of the core domain of the same molecule, which is defined as the closed state$^{29}$ (Fig. 3). In contrast, the swapped state refers to the C-terminal extension when 3D domain swapped, for example, in hAAC59–163 and hABC68–162$^1$. Thus, the non-3D domain swapped C-terminal extensions of zfAAC60–166 create a tetrameric oligomer.

The structure displays two conformations of the hinge loop. In one conformation, the hinge loop contains two beta hairpin structural motifs that are roughly perpendicular to the beta strands of the alpha crystallin core domain (Fig. 3) with clear electron density (Supporting Information Fig. S1). This beta hairpin motif is reminiscent of the "w"-like structure composed of three beta hairpin structural motifs in hABC59–163$^1$ (Supporting Information Fig. S2). In a second conformation, the hinge loop forms a large loop. Similar molecular interactions position the hinge loop of zfAAC60–166 in these two conformations. In both conformations, the carbonyl hydrogen of hinge loop glycine 153 bonds to the amide of serine 128. The side chain of arginine 154 hydrogen bonds to both the carbonyl of serine 123 and the side chain of aspartate 93 in one of its two conformations and to the side chain of aspartate 93 in its second conformation.

Several molecular interactions stabilize the bound C-terminal quasi-palindromic tail. Residues 156–166 in both chains of zfAAC60–166 are similar, with differences for residues 163–166. Two short beta strands, $\beta 10$ and $\beta 11$ Fig. 1, are formed on each side of the C-terminal tail (Figs. 3 and 4). For both chains, C-terminal tail side chain of aspartate 156 hydrogen bonds to amide of aspartate 92, and side chain of arginine 157 forms hydrogen bonds to both carbonyl of serine 128 and carbonyl of isoleucine 159. In short, these molecular interactions are similar to other truncated alpha crystallins.

The quasi-palindromic C-terminal tail supports bidirectional binding with pseudo-twofold symmetry. The C-terminal tail is bound to the core domain in the opposite direction of the tail of hAAC59–163$^1$ and in the same direction as the tails of hABC68–162$^1$, Wheat Hsp16.9$^{21}$ and $M.\ janaschii$ Hsp16.5$^{22}$. Structural comparison of the oppositely bound C-terminal tails of zfAAC60–166 and hAAC59–163 reveals a pseudo twofold symmetry axis centered on proline 160 (Fig. 4). The flanking palindromic arginines display differing rotamer conformations, while other residues nearly overlay.

**3D and Non-3D swapping of the C-terminal extension**

The energetic barrier to swapping of the C-terminal extension of AAC is evidently small as shown by the shift in equilibrium toward the 3D domain swapped state from the non-3D domain swapped state that is achieved by heating and subsequent cooling of a solution of zfAAC60–166. Native polyacrylamide gel electrophoresis resolves a native dimer, while on heating and subsequent cooling, an additional band of larger molecular weight is observed (Supporting Information Fig. S3). This larger molecular weight
complex (Supporting Information Fig. S3, lane 2) presumably arises from molecular interactions of the C-terminal extension with adjacent zfAAC60–166 chains. This observation shows that a small percentage of the C-terminal extensions have entered the 3D domain swapped state by heating AAC at a concentration of 1 mg/mL to 37°C for 1 h followed by rapidly cooling on ice. This result supports a modest energetic barrier to 3D domain swapping. A modest energy barrier would support a dynamic polydisperse oligomeric assembly with the ability of subunits to readily exchange. This is in direct agreement with other experimental work.30,31

Discussion

Alpha crystallins are highly dynamic proteins exhibiting structural plasticity.1,7,14,20–34 Regardless of whether the C-terminal extension in a swapped or closed state, the C-terminal hinge loop maintains structural rigidity while offering enough structural
flexibility to enable various oligomeric interactions. Our observation of both swapped and closed states of the C-terminal extension establishes another molecular mechanism by which the alpha crystallins maintain polydisperse oligomeric assemblies. 3D domain swapping, added to bidirectional binding, various AP interface registrations, and a flexible hinge loop of our previous work, show that at least four molecular mechanisms have evolved for the eye lens alpha crystallins to frustrate crystallization and to maintain eye lens function.

Material and Methods

Protein expression and purification
Plasmid construction, protein expression, and purification were performed as previously described. Briefly, zebrafish (D. rerio) AAC residues 60–166 was cloned into pET28b (Novagen, Gibbstown, NJ) with tobacco etch virus (TEV) protease cleavable N-terminal His-tag. Proteins were purified by affinity chromatography, followed by TEV protease cleavage of N-terminal His-tag. Proteins were purified by affinity chromatography, followed by TEV protease cleavage of N-terminal His-tag, and final purification by gel filtration chromatography, followed by TEV protease cleavage of N-terminal His-tag, and final purification by gel filtration chromatography, followed by TEV protease cleavage of N-terminal His-tag.

Protein crystallization
Concentrated zfAAC60–166 was diluted to 8 mg/mL in GF buffer, and crystals were grown in hanging drop VDX plates (Hampton Research, Aliso Viejo, CA) in condition Qiagen PACT B3 (0.1 M MIB (2:3:3—sodium malonate:imidazole:boric acid) buffer pH 6.0, 25% PEG 1500, 20% glycercol) prior to use. Protein concentration was determined by absorbance at 280 nm with the calculated extinction coefficient of 0.47.

Data collection and structure determination
Flash frozen cryoprotected crystals of zfAAC60–166 were used for data collection at APS beamline 24-ID-E. Data collection sets were collected at a set distance of 300 mm with a detector 20 of 0° and 18°. Data were processed using DENZO and SCALEPACK. Phases were obtained by molecular replacement with pdb 3L1E using Phaser and followed by 20 cycles of automated model building in ARP/wARP. Further manual model building was done using COOT. Model refinement was done using PHENIX.

Native polyacrylamide gel electrophoresis
zfAAC60–166 stored at 4°C was diluted to 1 mg/mL in GF buffer. The sample was then incubated at 37°C for 1 h, followed by rapidly cooling on ice. The control sample was stored on ice during the 1-h incubation period. The two samples were run on an 8–25% native gel using the pharmacia phast system (GE Healthcare, Piscataway, NJ) according to manufacturer protocol. The resulting native gel was stained with coomassie.

Acknowledgments
The authors thank Duilio Cascio, Joseph Horwitz, Meytal Landau, and Mike Sawaya for useful discussion, Mason Posner for zebrafish alpha crystallin clones, UCLA-DOE X-ray Crystallography Core Facility, Jason Navarro at the UCLA Crystalization Facility, M. Capel, K. Rajashankar, F. Murphy, J. Schuermann, and I. Kourinov at NE-CAT beamline 24-ID-E at APS, and NIH, NSF, and HHMI for support.

References
10. Berengian AR, Parfenova M, McHaourab HS (1999) Site-directed spin labeling study of subunit interactions in...


