

# SCIENTIFIC REPORTS



OPEN

## Structure of a Highly Active Cephalopod S-crystallin Mutant: New Molecular Evidence for Evolution from an Active Enzyme into Lens-Refractive Protein

Wei-Hung Tan\*, Shu-Chun Cheng\*, Yu-Tung Liu\*, Cheng-Guo Wu\*, Min-Han Lin, Chiao-Che Chen, Chao-Hsiung Lin & Chi-Yuan Chou

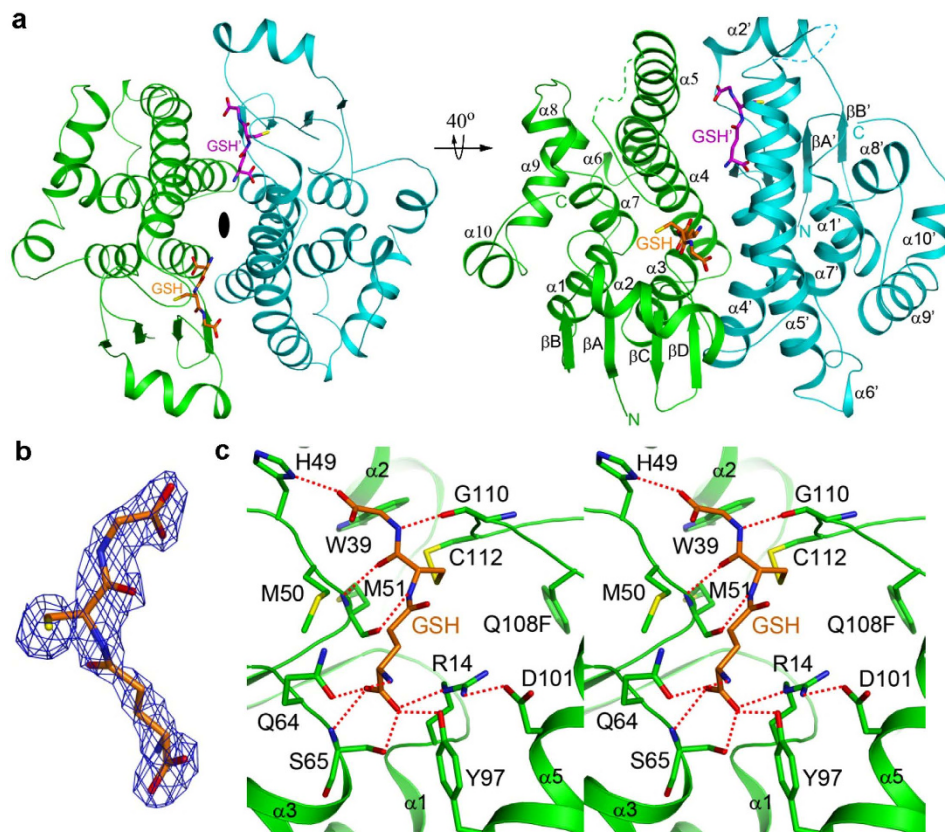
Received: 25 April 2016  
Accepted: 13 July 2016  
Published: 08 August 2016

Crystallins are found widely in animal lenses and have important functions due to their refractive properties. In the coleoid cephalopods, a lens with a graded refractive index provides good vision and is required for survival. Cephalopod S-crystallin is thought to have evolved from glutathione S-transferase (GST) with various homologs differentially expressed in the lens. However, there is no direct structural information that helps to delineate the mechanisms by which S-crystallin could have evolved. Here we report the structural and biochemical characterization of novel S-crystallin-glutathione complex. The 2.35-Å crystal structure of a S-crystallin mutant from *Octopus vulgaris* reveals an active-site architecture that is different from that of GST. S-crystallin has a preference for glutathione binding, although almost lost its GST enzymatic activity. We've also identified four historical mutations that are able to produce a "GST-like" S-crystallin that has regained activity. This protein recapitulates the evolution of S-crystallin from GST. Protein stability studies suggest that S-crystallin is stabilized by glutathione binding to prevent its aggregation; this contrasts with GST- $\sigma$ , which do not possess this protection. We suggest that a tradeoff between enzyme activity and the stability of the lens protein might have been one of the major driving force behind lens evolution.

Crystallins are soluble proteins in eye lenses that play an important role in the maintenance of lens transparency and optical clarity<sup>1</sup>. Their aggregation can result in the formation of cataracts, which are the most common cause of blindness worldwide<sup>2</sup>. Up to the present, the only treatment for this is the surgical removal of cataractous lens, albeit it has been reported very recently that lanosterol may reverse aggregation of crystallin<sup>3</sup>. Interestingly, some taxon-specific crystallins have been found to have structures that are related to cytosolic housekeeping enzymes<sup>4,5</sup>. How lens crystallins evolved from such housekeeping enzymes is a most important question and answers to this question will greatly help our understanding of protein function, folding, and stability<sup>6</sup>.

Among the crystallins, S-crystallin in the cephalopod lens has been shown to have an amino acid sequence similar to that of glutathione S-transferase (GST)<sup>6–8</sup>, which catalyzes the conjugation of glutathione (GSH) to various endogenous and xenobiotic electrophilic compounds and is a major detoxification mechanism of biological systems<sup>9–12</sup> (Extended data Fig. 1). Previous studies have led to a model that involves multiple gene duplications and exon shuffling to allow the evolution from GST to S-crystallin<sup>4,7,8</sup> and this process then creates a diverse and large family of S-crystallins. There are about ten, twenty four and four different S-crystallins in the lens of the squids, *Nototodarus sloanii* and *Loligo opalescens*, and *Octopus vulgaris* (the West Pacific octopus), respectively<sup>7,13,14</sup>. The cDNAs encoding these S-crystallins have nucleotide sequence identities ranging from 58% to 98%, while their proteins have identities ranging from 46% to 99% at the amino acid sequence level<sup>7</sup>. Among these proteins, "short-loop" S-crystallins (SL11, S4 and Cry9) are expressed less in the lens and are more GST-like, while "long-loop" S-crystallins

Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei 112, Taiwan.  
\*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to C.-Y.C. (email: cychou@ym.edu.tw)



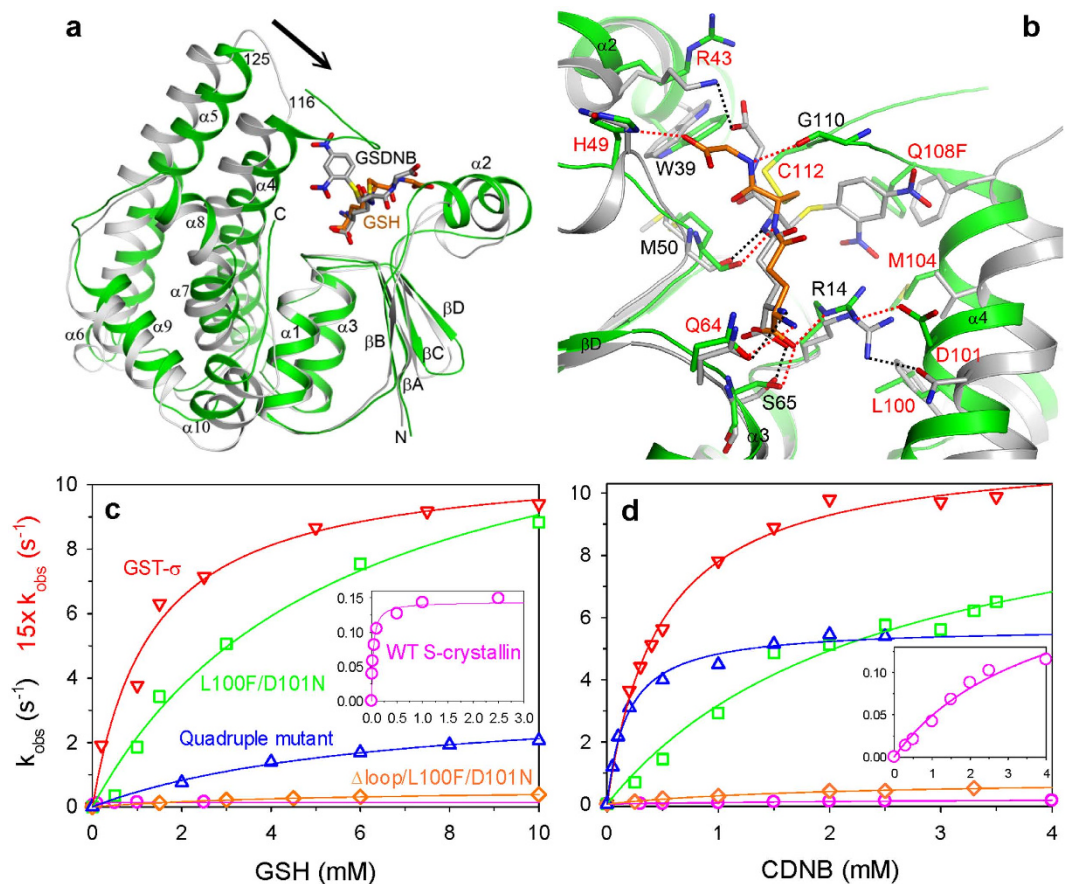
**Figure 1.** Crystal structure of the octopus S-crystallin Q108F mutant in complex with GSH. (a) A schematic drawing of the structure of the S-crystallin dimer (in green and cyan) in complex with one of two substrates, GSH (in orange and magenta), in two views. Residues (117 to 124), which are missing in the  $\alpha 4$ - $\alpha 5$  loop, are indicated by dashed lines. The 2-fold axis of the dimer is indicated by the black oval. (b) The  $F_o - F_c$  electron density at 2.35 Å resolution of GSH contoured at  $3.5\sigma$ . (c) A stereo pair showing the detailed interactions between the GSH and the active site of S-crystallin. Ion-pair and hydrogen-bonding interactions are indicated by red dashed lines. All structural figures were produced using PyMOL (<http://www.pymol.org>).

are dominantly expressed in the lens and are enzymatically inactive<sup>7,15</sup>. Up to the present there has been no direct structural information available that allows us to understand the reason for the differences.

## Results and Discussion

In the present study we overexpressed several long-loop S-crystallins of *Octopus vulgaris* in *Escherichia coli*. Only one of them, named OctS4, could be expressed and purified to homogeneity. However, efforts to crystallize the wild-type OctS4 were unsuccessful. Therefore, we modified the wild-type sequence by creating a number of mutants that are believed to be related to the binding of GSH or of electrophilic compounds (Extended data Fig. 1). After a significant number of trials, a highly active mutant, Q108F, was successfully expressed and crystallized in the presence of GSH. We determined the crystal structure of the S-crystallin Q108F mutant at 2.35 Å resolution (Fig. 1 and Extended data Table 1). A monomer of S-crystallin forms an asymmetric unit within the crystal, and a dimer of the protein that is similar to GST can be generated using a crystallographic 2-fold symmetry axis (Fig. 1a and Extended data Fig. 2a). Approximately 1600 Å<sup>2</sup> of the surface area of each monomer is buried in the interface of this dimer, comparing to 1300 Å<sup>2</sup> for the squid GST- $\sigma$ <sup>16</sup> with 38% sequence identity to S-crystallin. Besides, the dimerization of S-crystallin and several of its mutants was confirmed by analytical ultracentrifugation (Extended data Fig. 2).

The reported crystal structures of GST- $\sigma$  are in complex with either 1-(S-glutathionyl)-2,4-dinitrobenzene (GSDNB) or S-(3-iodobenzyl) glutathione (GSBzl)<sup>16</sup>. In contrast to the above, we observed the presence of a GSH molecule within the active site of S-crystallin, located between the N-domain ( $\beta A$  to  $\alpha 3$ ) and C-domain ( $\alpha 4$  to  $\alpha 10$ ) of the protein (Fig. 1b). There is a disulfide-bond between the thiol group of the GSH and residue Cys112 of S-crystallin. In addition, there is a polar interaction network (Fig. 1c) that consists of the glutamyl group of GSH interacting with residues Arg14, Gln64, Ser65 and Tyr97, the cysteinyl group of GSH interacting with residue Met51 and the glycyl group of GSH interacting with residues His49 and Gly110. The two domains of S-crystallin are closer compared to those of GST- $\sigma$  in complex with GSDNB; although the r.m.s distance between equivalent C $\alpha$  atoms of the two structures is 1.3 Å (Fig. 2a). When their N-domains are superposed, a difference of 3.5° can be seen in the orientation of their C-domains.



**Figure 2. Structural and biochemical characterizations of S-crystallin.** (a) An overlay of the structure of the S-crystallin Q108F mutant in complex with GSH (green) against GST- $\sigma$  in complex with GSDNB (grey)<sup>16</sup>. The arrow indicates the conformational differences in the  $\alpha 4$  and  $\alpha 5$  helices. (b) Overlay of the active sites of the two structures. Residues labeled in red were selected for mutagenesis. (c,d) The GST activity of S-crystallin. Plot of the rate constant as a function of the concentration of two substrates GSH (c) and CDNB (d) for the wild-type (magenta) and various S-crystallin mutants, the L100F/D101N double mutant (green), the  $\Delta$ loop/L100F/D101N double mutant (orange), and L100F/D101N/M104V/Q108F quadruple mutant (blue). The activity of GST- $\sigma$  from octopus liver tissues (red) was used as a positive control. Insets show plots of the wild-type S-crystallin at a suitable scale.

Compared with the binding of GSH to S-crystallin, GSDNB has fewer hydrophilic interactions with GST- $\sigma$  (Fig. 2b). GSDNB has no interaction with residue Arg13 of GST- $\sigma$ , which is equivalent to residue Arg14 of S-crystallin. In addition, due to the lack of a long loop between the  $\alpha 4$  and  $\alpha 5$  helices (residues 112 to 122), the glycyl group of GSDNB does not interact with the C-domain of GST- $\sigma$ , while its dinitrobenzyl group does have hydrophobic contact with Phe106. Similarly, the structure of GST- $\sigma$  in complex with GSBzI indicates that the iodobenzyl ring of GSBzI is located in a hydrophobic pocket that consists of residues Phe98, Val102 and Phe106 (Extended data Fig. 3). Like other GSTs<sup>17</sup>, the different binding modes of the GSDNB and the GSBzI to GST- $\sigma$  imply a more flexible active site that favors catalysis. In contrast, the equivalent residues in S-crystallin are Leu100, Met104 and Gln108, respectively (Fig. 2b). These differences result in a collapsed pocket that may disfavor the binding of aromatic compounds. We characterized the GST activity levels of S-crystallin and its various mutants in order to evaluate the role of the above residues. Octopus GST- $\sigma$  was purified from the digestive gland and then used as a standard for enzyme activity (Fig. 2c,d). The GST catalytic activity ( $k_{cat}$ ) of the wild-type S-crystallin is  $0.24 \text{ s}^{-1}$ , which is about the same to that of the S-crystallins purified from octopus lens<sup>18</sup> but only  $\sim 1/700$ , of that of GST- $\sigma$  (it will be  $\sim 1/6000$  if compared their catalytic efficiency by  $k_{cat}/K_{m,CDNB}$ ) (Table 1). Consistent with our structural observations, we found that the apparent binding affinity of 1-chloro-2,4-dinitrobenzene (CDNB) to S-crystallin is lower than that with GST- $\sigma$ , with an 8.3-fold increase in the  $K_{m,CDNB}$ . By way of contrast, the apparent binding affinity of GSH with S-crystallin is significantly tighter than that with GST- $\sigma$ , with a 43-fold decrease in  $K_{m,GSH}$ .

Kinetic analysis indicated that three single mutants, L100F, M104V and Q108F, have a higher  $k_{cat}$  than that of wild-type S-crystallin with 10- to 44-fold increases, respectively (Extended data Fig. 4c,d and Extended data Table 3). The  $K_{m,GSH}$  of the L100F and M104V mutants of S-crystallin have 37-fold and 83-fold increases and are similar to that of GST- $\sigma$ , while  $K_{m,CDNB}$  of the M104V mutant shows a significant decrease and is almost identical to that of GST- $\sigma$  (0.65 mM versus 0.47 mM). Interestingly, although the Q108F mutant of S-crystallin possesses

Protein <sup>a</sup>	$K_{m,GSH}$ (mM) <sup>c</sup>	$K_{m,CDNB}$ (mM) <sup>c</sup>	$k_{cat}$ (s <sup>-1</sup> ) <sup>c</sup>
Wild-type S-crystallin	0.03 ± 0.004	3.9 ± 1.2	0.24 ± 0.05
L100F/D101N	5.5 ± 1.0	2.5 ± 0.7	14.1 ± 1.0
L100F/D101N/M104V/Q108F	6.5 ± 0.7	0.18 ± 0.02	5.6 ± 0.1
$\Delta$ loop/L100F/D101N <sup>b</sup>	5.3 ± 0.4	2.0 ± 0.4	0.80 ± 0.07
Octopus GST- $\sigma$	1.3 ± 0.2	0.47 ± 0.04	173.6 ± 4.3
S-like GST	3.7 ± 0.4	unsaturated	1.45 ± 0.07

**Table 1. Steady-state kinetic parameters of S-crystallin, its various mutants and two GSTs.** <sup>a</sup>The concentration of the six proteins was 1.1, 0.04, 0.1, 0.4, 0.001, and 0.2  $\mu$ M, respectively. For the GSH titration, the concentration of CDNB was 3.5 mM. For the CDNB titration, the concentration of GSH was 10 mM. <sup>b</sup>The loop between  $\alpha$ 4 and  $\alpha$ 5 helices (residue 112–122) was deleted and is designated as  $\Delta$ loop. <sup>c</sup>Data were fitted to the Michaelis-Menten equation and the  $R_{sq}$  values were 0.980 to 0.999, respectively. All the assays were repeated at least twice to ensure reproducibility. The  $K_m$  was used as apparent binding affinity in the present studies.

the highest  $k_{cat}$  among the three mutants, its  $K_{m,GSH}$  is also very high and we were not able to estimate a reliable  $K_{m,CDNB}$  due to unsaturated kinetic assays. This indicates that the mutation of Gln108 to Phe is beneficial for catalysis but is detrimental to the binding of the two substrates. These results suggest that, in order to obtain a more “GST-like” S-crystallin, no single mutant is suitable, but rather multiple mutations are necessary (see below).

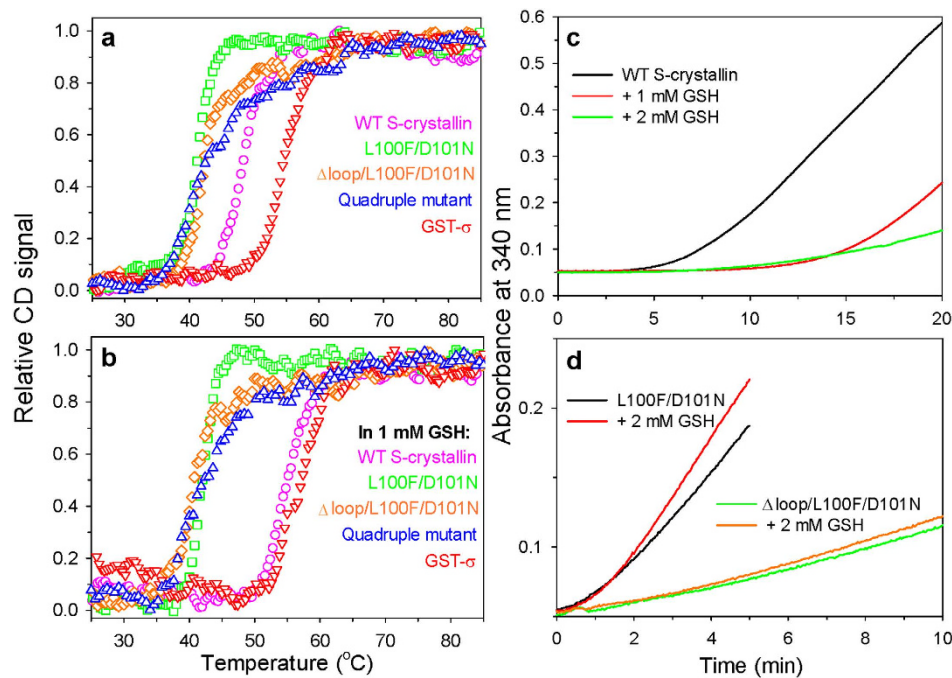
Another difference between S-crystallin and GST- $\sigma$  occurs with residue Asp101, which is at the active center (Asn99 in GST- $\sigma$ ) (Fig. 2b). Although both residues are located in the vicinity of an arginine residue, the “charge-charge” interaction between Asp101 and Arg14 in S-crystallin is likely to diminish the important function of arginine residue, which is to stabilize the negatively charged intermediate state (the Meisenheimer complex) during catalysis<sup>19</sup>. Indeed, the mutation of Asp101 to Ala or Asn resulted in a minor increase in  $k_{cat}$  in parallel with a decrease in the  $K_{m,CDNB}$  of S-crystallin; albeit a 13-fold to 87-fold increase in  $K_{m,GSH}$  (Extended data Fig. 4c,d and Extended data Table 3). Mutation of other residues involved in the binding of GSH, such as Arg43 (to Lys) and His49 (to Asn) (Fig. 2b), resulted in no significant differences in GST activity (Extended data Fig. 4a,b and Extended data Table 3). On the other hand, based on our structural observations and our kinetic assays, the highly conserved residue Gln64 would seem to be essential for the binding of GSH and for catalysis (Fig. 2b and Extended data Table 3).

Next we focused on four key residues of S-crystallin, Leu100, Asp101, Met104 and Gln108, mutation of which resulted in large changes in catalysis or binding. The GST activity of the various double, triple and quadruple mutants of S-crystallin was characterized in order to delineate their mutual effects (Fig. 2c,d, Extended data Fig. 4e,f and Extended data Table 3). The results suggest that, compared with those from the single mutants, all double and triple mutants of S-crystallin have comparable or even higher  $k_{cat}$  values, while there were no significant differences in the  $K_m$  with GSH or CDNB. Furthermore, the quadruple mutant of S-crystallin, L100F/D101N/M104V/Q108F, had the lowest  $K_{m,CDNB}$  (0.18 mM) and the highest  $K_{m,GSH}$ . Although there was only a 23-fold increase in activity for some unknown reason, the catalytic efficiency ( $k_{cat}/K_{m,CDNB}$ ) has a 518-fold increase (Table 1). Nevertheless, based on the switch in the two substrate-binding affinities (increased  $K_{m,GSH}$  and decreased  $K_{m,CDNB}$ ) and the significant recovery of GST activity and catalytic efficiency, the double mutant D101N/Q108F and the quadruple mutant of S-crystallin would seem to result in a GST-like S-crystallin that has recovered GST activity and probably a more flexible active site<sup>17</sup>. In addition, another double mutant of S-crystallin, D101A/Q108F, showed the highest enzymatic activity in these mutants, which is only a 7-fold lower than that of GST- $\sigma$  (Extended data Table 3).

Our structure demonstrates that the long loop between the  $\alpha$ 4 and  $\alpha$ 5 helices of S-crystallin is able to bind to GSH through disulfide-bonding interaction via residue Cys112 and via a hydrogen bonding interaction with residue Gly110 (Fig. 1c). Next we characterized the GST activities of the C112G and loop-deletion mutants of S-crystallin in order to evaluate the effect of these changes. Interestingly, the mutation of Cys112 to Gly did not influence the binding affinity of GSH to S-crystallin; although the GST activity of the mutant is 3-fold lower than that of the wild-type (Extended data Fig. 4a,b and Extended data Table 3). This indicates that the disulfide bonding between S-crystallin and GSH is not essential for binding. This also helps to explain the observation that the equivalent residue in most other S-crystallins is a glycine<sup>7,14</sup> (Extended data Fig. 1).

Not surprisingly, the deletion of the long loop (residues 112–122) resulted in an obvious decrease in the apparent binding affinity of GSH as with a 73-fold increase in  $K_{m,GSH}$  but a similar  $k_{cat}$  compared to the wild-type S-crystallin (Extended data Fig. 4g,h and Extended data Table 3). These features prompted us to make a mutant that consisted of the four key residues changed previously together with the loop deletion in order to generate a putative “ancestral” protein. However, other than the  $\Delta$ loop/Q108F mutant of S-crystallin, the other mutations together with the loop deletion resulted in a low level of recovery of activity and poor binding affinity with GSH (Fig. 2c,d, Extended data Fig. 4g,h and Extended data Table 3). Unlike the Q108F mutant alone form of the protein, the  $K_{m,CDNB}$  of the  $\Delta$ loop/Q108F mutant of S-crystallin can be precisely determined to be 1.9 mM. This suggests that existence of the long loop not only enhances the binding of GSH to S-crystallin but may interfere with the binding of electrophilic compounds.

To understand the mechanisms by which changes in gene sequence generate shifts in function and therefore phenotype<sup>20,21</sup>, it is necessary to analyze how the changes in protein structure mediate the effects that the mutations have on function<sup>22</sup>. Interestingly, a long-loop S-crystallin (OctS4) shows a significant increase in apparent GSH binding affinity compared to that of GST- $\sigma$ . To evaluate the possible biological significances of such strong



**Figure 3.** Protein stability of S-crystallin and its mutants by heat denaturation (a,b) and chemical denaturation (c,d). Plots of the relative CD signal from the ellipticity at 222 nm as a function of the temperature for the wild-type (magenta) and various S-crystallin mutants, the L100F/D101N double mutant (green), the  $\Delta$ loop/L100F/D101N double mutant (orange), and the L100F/D101N/M104V/Q108F quadruple mutant (blue) without (a) or with (b) 1 mM GSH. The thermal stability of GST- $\sigma$  (red) without or with 1 mM GSH was also measured. The protein concentration was 7.2  $\mu$ M. The results were fitted in order to calculate their  $T_m$  of each mutant, which are shown in Table 2. (c,d) Denaturant-induced light scattering of S-crystallin (c) and its mutants (d). The aggregation traces of S-crystallin in 3.5 M urea without or with GSH at 1–2 mM were observed by light-scattering at 340 nm. The protein concentration was 10  $\mu$ M. The assays were repeated twice to ensure reproducibility.

Protein <sup>a</sup>	$T_m$ (°C)	$T_m$ in GSH (°C)	$\Delta T_m$ 1 (mutant-WT)	$\Delta T_m$ 2 (in GSH-apoform)
Wild-type S-crystallin	48.2 $\pm$ 1.0	55.2 $\pm$ 1.2	—	7.0
L100F/D101N	41.5 $\pm$ 1.2	41.9 $\pm$ 1.6	-6.7	0.4
L100F/D101N/M104V/Q108F	39.2 $\pm$ 1.3	39.9 $\pm$ 1.4	-9.0	0.7
$\Delta$ loop/L100F/D101N	41.7 $\pm$ 2.0	40.3 $\pm$ 2.2	-6.5	-1.4
GST- $\sigma$	54.8 $\pm$ 0.9	56.5 $\pm$ 1.5	6.6	1.7
S-like GST	42.4 $\pm$ 1.9	41.7 $\pm$ 2.1	-5.8	-0.7

**Table 2.** Thermal stability of S-crystallin, its various mutants and two GSTs with or without 1 mM GSH. <sup>a</sup>The protein concentration was at 7.2  $\mu$ M. The ellipticity at 222 nm was monitored while varying the temperature ranging from 25 °C to 85 °C. The results were fitted to the two-state unfolding model in order to calculate  $T_m$  of the various proteins.

GSH binding, the thermal stability of S-crystallin without or with GSH was determined using circular dichroism spectrometry (Fig. 3a,b). In the presence of GSH, the melting temperature ( $T_m$ ) of S-crystallin was higher by 7 °C than that of the protein in the absence of GSH (Table 2). Again, the mutation of Cys112 to Gly did not influence the  $T_m$  enhancement in the presence of GSH (Extended Data Fig. 5a,d and Extended Data Table 4). This indicates that the disulfide bonding between S-crystallin and GSH is not essential for this protection. Furthermore, denaturant-induced aggregation of S-crystallin was also observed using light-scattering analysis (Fig. 3c). Our results suggested that S-crystallin aggregation is prevented by the presence of GSH in a dose-dependent manner. By way of contrast, GST- $\sigma$  showed only a 1.7 °C increase in  $T_m$  in the presence of GSH; although the  $T_m$  of the free GST- $\sigma$  is higher than that of the free S-crystallin. It indicated that GSH may not benefit to stabilize the GST- $\sigma$ . Differently, previous studies have suggested that GSH binding is required to stabilize GSTP1-1 and *Plasmodium falciparum* GST<sup>23–25</sup>.

The structural protection and aggregation prevention afforded to S-crystallin by GSH suggested that in the lens S-crystallin may bind to GSH and that this contributes to the overall stability of lens proteins. In addition, this explains why S-crystallin from the lens was unable to bind to a GSH affinity column in earlier studies<sup>19,26,27</sup>.

Interestingly, all animals have an abundance of reducing agents in the lens, including GSH (2–10 mM), in order to maintain the redox potential of the lens<sup>28</sup>. Traditionally, the high GSH content in lens is believed to protect free thiols, which allows the structural integrity and/or proper biological functioning of various enzymes to be maintained<sup>28</sup>. Aging or damaged lenses show a diminished size of GSH pool and accelerating cataract. Previous studies have suggested that glutathiolation of human gamma crystallin is able to enhance its degradation by a proteasome system<sup>29</sup>. Here our novel finding is that GSH can also be used as a ligand of S-crystallin in order to prevent its aggregation and possible cataract formation in the cephalopod lens.

If we considered the situation of S-crystallin in the lens, it is important for this protein to capture GSH for as long as possible and to minimize its catalytic activity, otherwise the GSH will be released as a product conjugate. This is likely to be a driving force that provided the impetus for the ancestral protein to evolve in a manner that weakens its binding affinity for electrophilic compounds by mutating residues 100, 104 and 108 and that interferes with the formation of the intermediate complex by mutating residue 101. In addition, inserting the long loop between  $\alpha 4$  and  $\alpha 5$  helices is also able to enhance the binding of GSH while at the same time interfering with the binding of electrophilic compounds. We believe that these mutations may play crucial roles in the evolutionary trajectory of S-crystallin. Similarly, it has been previously reported that the old human GST, GSTT2-2, shows a low affinity for GSH, compared to the more recently evolved human Alpha, Mu, and Pi GSTs<sup>17</sup>. It is possible to speculate that GST and S-crystallin may be evolved from a common GST ancestor under the same evolutionary pressure, i.e. in the direction of a decrease in  $K_m$  for GSH.

The thermal stabilities of the various mutations with or without the long  $\alpha 4$ - $\alpha 5$  loop were also characterized. Other than the Q108F mutant of S-crystallin, all the mutants that showed an increased  $K_m$  or a recovery in activity were found to have a lower  $T_m$  without GSH (Fig. 3a,b, Extended data Fig. 5 and Extended data Table 4). This suggests that the original residues and the insertion of the long loop of S-crystallin improve the stability of the protein. Furthermore, other than the L100F and D101N mutants of S-crystallin, all the other mutations were found to have lost the protective effect of GSH as there were no  $T_m$  increases when GSH was present. Based on these findings, we concluded that the existence of the long loop and the collapse of the binding pocket are both required for S-crystallin to be able to be protected in terms of stability when GSH is present. Furthermore, the mutation of the residue 101 from Asn to Asp is able to increase the stability of S-crystallin and this occurs because of a strengthening of the GSH protection effect.

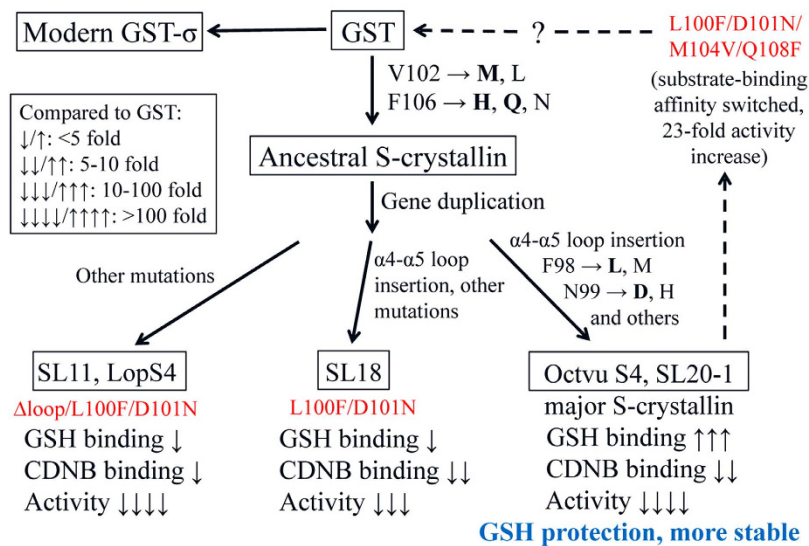
To recapitulate the evolutionary trajectory of S-crystallin from its putative ancestral protein, GST or a GST like protein<sup>8</sup>, it was possible to create various different proteins by mutation of “GST” towards a more “S-crystallin”-like protein and by mutation of “S-crystallin” towards a “GST”-like protein. We have successfully generated a GST-like S-crystallin that has a 100-fold increase in activity and shows a switch in substrate-binding affinity. As the next step, using the full gene synthesis, we created a “S-crystallin”-like GST, by a quadruple mutation, F98L/N99D/V102M/F106Q, and the insertion of the long loop into the  $\alpha 4$ - $\alpha 5$  helices of the GST- $\sigma$  sequence (Extended data Fig. 6). However, although the S-crystallin-like GST protein showed a 120-fold reduction in catalytic activity and an unsaturated pattern for the binding of CDNB (Table 1); this protein failed to gain the ability to be protected by GSH (Table 2). Mutations of other residues involved in the binding of GSH may be necessary to retain GSH in the protein.

Another concern when deciphering the evolution of S-crystallin is whether we can find a orthologous S-crystallin that is close to ancestral protein, in other words, a “molecular fossil”<sup>8</sup>. There are many different kind of S-crystallins found in the cephalopod lens as a result of gene duplication and/or the recruitment of novel genes into this morphogenetic pathways<sup>7,14,15</sup>. Previous studies have suggested dividing S-crystallins into two groups, the short-loop group, which is less expressed and the long-loop group, which are the dominant forms of S-crystallin. Based on sequence alignment, the short-loop crystallins are considered to be the earliest descendants from the ancestral gene<sup>7</sup>. In the present studies, we have identified one long-loop S-crystallin, SL-18, that is different to the others because it preserves residues 100 and 101 as Phe and Asn, respectively (Extended data Fig. 1). Actually, the short-loop S-crystallins (SL11, LopS4 and Cry9) from three species of squids also preserve the two residues<sup>7,13,15</sup>. We chose two mutants,  $\Delta$ loop/L100F/D101N and L100F/D101N, and used these to mimic the short-loop and SL18 S-crystallins, respectively (Figs 2c,d and 3). The long-loop deletion resulted in the loss of GSH protection and the ability to prevent aggregation (Table 2 and Fig. 3d); although the GST activity of short-loop mimic showed a 3.3-fold increase (Table 1). This slight increase in activity is compatible with the findings of previous studies wherein lysate containing expressed short-loop-crystallin showed minor GST activity<sup>7</sup>. On the other hand, the SL18 mimic protein showed significantly high GST activity (a 59-fold increase) as well as lower thermal stability and the inability to prevent aggregate (Tables 1 and 2 and Fig. 3). For the first time we have provided experimental evidence showing that a modern S-crystallin with high GST activity do exist and that this may be advantageous to help protect other proteins in cephalopod lens when GSH is present.

Here we have provided a schematic summary that demonstrates the structural/functional variations that affect duplicated S-crystallins (Fig. 4). To have a major dominant modern S-crystallins that has GSH protection, there are three groups of mutations needed: those affecting the binding pocket of the electrophilic compound (GST’s Phe98, Val102 and Phe106 residues), those affecting intermediate state stabilization (Asn99), and the presence of a long-loop insertion between the  $\alpha 4$  and  $\alpha 5$  helices. Sequence alignment suggests that all modern S-crystallin have mutations at Val102 and Phe106, which indicates that these events have happened quite early in the ancestral S-crystallin before gene duplication occurred. Later, Phe98 and Asn99 were mutated resulting in a further loss of enzyme activity with only the short-loop (SL11, LopS4, Cry9) and one long-loop S-crystallin (SL18) still retaining these two residues. In a final step, there is an insertion of the various long-loop, which produces the goal of GSH protection while making the protein more stable.

Overall, our findings have revealed a novel structure for the cephalopod S-crystallin and provided crucial evidence that greatly aids our understanding of the evolution of S-crystallin. Loss of GST enzyme activity is linked to gaining the enhanced protein stability of S-crystallin and this occurs via GSH binding. These changes can be

## Scheme of S-crystallin evolution



**Figure 4. Putative evolution of cephalopod S-crystallin.** The three different groups of S-crystallin are separated by the mutations at two residues F98 and N99 and by the insertion of a long loop between the  $\alpha 4$  and  $\alpha 5$  helices. Two other residues, V102 and F106, can be used to separate GST and the ancestral S-crystallin. With a longer  $\alpha 4$ - $\alpha 5$  loop and the four “key” mutations, major S-crystallins lose their GST activity but gain significantly increased GSH binding affinity. The binding of GSH to S-crystallin enhances its protein stability, which is beneficial for the life-span of the lens and the survival of cephalopods. According to previous studies<sup>29</sup>, there is between 2 mM and 10 mM GSH in animal lenses.

considered to be the major driving force during the evolution of S-crystallin. Moreover, the secret of the stability of crystallin may help us find a strategy that will allow us either to prevent or treat cataracts.

### Methods

**Protein expression and purification.** The west Pacific octopus (*Octopus vulgaris*) was purchased from local fish market and kept on ice during the transfer. The cDNA of the octopus was prepared by RNA extraction from lenses according to standard procedures<sup>14</sup>. One of the major S-crystallin cDNAs was then subcloned into pET-28a vector (Novagen). A fused protein with a N-terminal 6x His tag was then expressed by *E. coli* BL21 Rosetta (DE3) (EMD Millipore). Cultures were incubated overnight at 20 °C and induced with 0.4 mM isopropyl- $\beta$ -D-thiogalactopyranoside. The cell pellets were resuspended in lysis buffer (20 mM Tris, pH 8.5, 250 mM NaCl, 5% glycerol, 0.2% Triton X-100 and 2 mM  $\beta$ -mercaptoethanol), lysed by sonication and then centrifuged to remove any insoluble lysate. The remaining soluble proteins were then passed through nickel affinity beads (Qiagen) and then the bound expressed protein was released. This protein was subjected to further purification by gel-filtration (Sephacryl S-100; GE Healthcare) chromatography using a running buffer of 20 mM Tris, pH 8.5, 100 mM NaCl and 2 mM dithiothreitol. The purified protein was concentrated to 20 mg/ml, and the solution was supplemented with 5% (v/v) glycerol before being flash-frozen in liquid nitrogen and stored at -80 °C.

Octopus digestive gland GST was purified by a single GSH-Sepharose 4B affinity chromatography process that has been described previously<sup>18,26</sup>. Briefly, octopus digestive gland (20 g) from an octopus was homogenized with 10 mM phosphate buffer (pH 7.0) containing 2 mM mercaptoethanol and then subjected to ultracentrifugation at 4 °C and 80,000 g for 1 h. After filtration, the supernatant was further purified by GSH-Sepharose 4B affinity column. The fractions containing GST were pooled and supplemented with 5% (v/v) glycerol before being flash-frozen in liquid nitrogen and stored at -80 °C. All animal experiments conformed to the regulations drafted by Animal Protection Act in Taiwan and were approved by the laboratory animal center at National Yang-Ming University.

**Protein crystallization.** Crystals of the S-crystallin Q108F mutant in complex with GSH were obtained at 22 °C by the sitting drop, vapor diffusion method. Initial crystal screens were set up after incubating the protein at 6 mg/ml in 7 mM GSH and 0.35 mM CDNB for 1 h. Poorly diffracting crystals were grown in 0.5% PEG8000, 0.1 M HEPES, pH 6.8 and 1.65 M ammonium sulfate in 3 days and these were used for microseeding. Single crystals of a hexagonal shape and with dimensions of 0.15 mm were obtained in less than a week. All crystals were cryoprotected in reservoir solution supplemented with 25% (v/v) glycerol and flash-cooled in liquid nitrogen.

**Data collection and structure determination.** X-ray diffraction data were collected at 100 K on the SPXF beamline 13C1 at the National Synchrotron Radiation Research Center, Taiwan, R.O.C. using an ADSC Quantum-315r CCD detector (X-ray wavelength of 0.976 Å). The diffraction images were processed and scaled with the HKL-2000 package<sup>30</sup>. The crystals belong to space group P6<sub>4</sub>22, with cell dimensions of a = b = 114.3 Å, c = 63.9 Å,  $\alpha = \beta = 90^\circ$ , and  $c = 120^\circ$ . There is one S-crystallin subunit in the crystallographic asymmetric unit.

The structure of S-crystallin was solved by the molecular replacement method by the program Molrep<sup>31</sup> using a simulated model for octopus S-crystallin<sup>19</sup> as the model. The atomic model was built using the program Coot<sup>32</sup>. Structure refinement was carried out using REFMAC<sup>33</sup>. The crystallographic information is summarized in Extended Data Table 1.

**Analytical ultracentrifugation analysis.** The experiments were performed on an XL-A analytical ultracentrifuge using an An-60 Ti rotor (Beckman Coulter)<sup>34–36</sup>. The sedimentation-velocity experiments were performed using a double-sector epon charcoal-filled centerpiece at 20 °C with a rotor speed of 42,000 rpm. Protein solutions of S-crystallin and its mutants (0.025–1.0 mg/ml) and a reference buffer (50 mM phosphate, pH 6.5) were loaded into the centerpiece. The absorbance at 230 or 280 nm was monitored in continuous mode with a time interval of 300 s and a step size of 0.003 cm. Multiple scans at different time intervals were then fitted to a continuous  $c(M)$  distribution model using SEDFIT<sup>37,38</sup>.

**Mutagenesis and kinetic studies.** Site-specific mutants were designed based on the available sequence alignment and structural information. The mutations were introduced using the QuikChange kit (Stratagene) and verified by DNA sequencing. The primers used for site-specific mutations are summarized in Extended Data Table 2.

The GST activity of S-crystallin was determined by measuring the conjugation of GSH and CDNB<sup>26,39</sup>. The formation of the glutathione conjugate was monitored continuously at 340 nm. A molar absorption coefficient of  $9,600 \text{ M}^{-1} \cdot \text{cm}^{-1}$  for the conjugate was used in calculations. The initial velocity study for pseudo-first order kinetic assay was performed by varying the concentration of GSH from  $1 \mu\text{M}$  to 12 mM with a constant concentration of CDNB of 3.5 mM for the  $K_{m,GSH}$  calculation and by varying the concentration of CDNB from 0.1 to 3.8 mM with a constant concentration of GSH of 10 mM for  $K_{m,CDNB}$ . The  $K_m$  was used as apparent binding affinity in the present studies. The non-enzymatic conjugation of GSH with CDNB (Extended data Fig. 4I,j) was corrected for during each assay.

**Thermal stability analysis.** The unfolding of each protein was monitored by measuring the CD signal of 222 nm from 25 °C to 85 °C. The protein concentration was 0.2 mg/ml without or with 1 mM GSH. The data was analyzed using thermodynamic models by globally fitting the measurements to a two-state unfolding model in order to calculate the  $T_m$  of each protein<sup>40</sup>.

**Light-scattering measurement.** Polymerization and aggregation of S-crystallin and its various mutants was measured by denaturant-induced light-scattering<sup>41</sup>. The protein was incubated with 3.5 M urea in the absence or presence of GSH at 30 °C. The protein concentration was 0.3 mg/ml. The light-scattering at 340 nm was continuously monitored by UV/VIS spectrometry (Jasco).

## References

- De Jong, W. W., Hendriks, W., Mulders, J. W. & Bloemendal, H. Evolution of eye lens crystallins: the stress connection. *Trends Biochem. Sci.* **14**, 365–368 (1989).
- Pascolini, D. & Mariotti, S. P. Global estimates of visual impairment: 2010. *Br. J. Ophthalmol.* **96**, 614–618 (2012).
- Zhao, L. *et al.* Lanosterol reverses protein aggregation in cataracts. *Nature* **523**, 607–611 (2015).
- Piatigorsky, J. & Wistow, G. The recruitment of crystallins: new functions precede gene duplication. *Science* **252**, 1078–1079 (1991).
- Piatigorsky, J. Lens crystallins. Innovation associated with changes in gene regulation. *J. Biol. Chem.* **267**, 4277–4280 (1992).
- Sweeney, A. M., Des Marais, D. L., Ban, Y. E. & Johnsen, S. Evolution of graded refractive index in squid lenses. *J R Soc Interface* **4**, 685–698 (2007).
- Tomarev, S. I., Chung, S. & Piatigorsky, J. Glutathione S-transferase and S-crystallins of cephalopods: evolution from active enzyme to lens-refractive proteins. *J. Mol. Evol.* **41**, 1048–1056 (1995).
- Tomarev, S. I. & Piatigorsky, J. Lens crystallins of invertebrates—diversity and recruitment from detoxification enzymes and novel proteins. *Eur. J. Biochem.* **235**, 449–465 (1996).
- Mannervik, B. & Danielson, U. H. Glutathione transferases—structure and catalytic activity. *CRC Crit Rev Biochem* **23**, 283–337 (1988).
- Rushmore, T. H. & Pickett, C. B. Glutathione S-transferases, structure, regulation, and therapeutic implications. *J. Biol. Chem.* **268**, 11475–11478 (1993).
- Hayes, J. D. & Pulford, D. J. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit. Rev. Biochem. Mol. Biol.* **30**, 445–600 (1995).
- Armstrong, R. N. Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem. Res. Toxicol.* **10**, 2–18 (1997).
- Tomarev, S. I., Zinovieva, R. D. & Piatigorsky, J. Characterization of squid crystallin genes. Comparison with mammalian glutathione S-transferase genes. *J. Biol. Chem.* **267**, 8604–8612 (1992).
- Chiou, S. H. *et al.* Octopus S-crystallins with endogenous glutathione S-transferase (GST) activity: sequence comparison and evolutionary relationships with authentic GST enzymes. *Biochem. J.* **309**, 793–800 (1995).
- Yoshida, M. A. *et al.* Molecular Evidence for Convergence and Parallelism in Evolution of Complex Brains of Cephalopod Molluscs: Insights from Visual Systems. *Integr. Comp. Biol.* (2015).
- Ji, X. *et al.* Three-dimensional structure, catalytic properties, and evolution of a sigma class glutathione transferase from squid, a progenitor of the lens S-crystallins of cephalopods. *Biochemistry* **34**, 5317–5328 (1995).
- Caccuri, A. M. *et al.* Human glutathione transferase T2-2 discloses some evolutionary strategies for optimization of substrate binding to the active site of glutathione transferases. *J. Biol. Chem.* **276**, 5427–5431 (2001).
- Tang, S. S., Lin, C. C. & Chang, G. G. Isolation and characterization of octopus hepatopancreatic glutathione S-transferase. Comparison of digestive gland enzyme with lens S-crystallin. *J. Protein Chem.* **13**, 609–618 (1994).
- Chuang, C. C., Wu, S. H., Chiou, S. H. & Chang, G. G. Homology modeling of cephalopod lens S-crystallin: a natural mutant of sigma-class glutathione transferase with diminished endogenous activity. *Biophys. J.* **76**, 679–690 (1999).
- Golding, G. B. & Dean, A. M. The structural basis of molecular adaptation. *Mol. Biol. Evol.* **15**, 355–369 (1998).
- Stern, D. L. Evolutionary developmental biology and the problem of variation. *Evolution* **54**, 1079–1091 (2000).
- Ortlund, E. A., Bridgman, J. T., Redinbo, M. R. & Thornton, J. W. Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* **317**, 1544–1548 (2007).



23. Caccuri, A. M. *et al.* Temperature adaptation of glutathione S-transferase P1-1. A case for homotropic regulation of substrate binding. *J. Biol. Chem.* **274**, 19276–19280 (1999).
24. Ricci, G. *et al.* Glutathione transferase P1-1: self-preservation of an anti-cancer enzyme. *Biochem. J.* **376**, 71–76 (2003).
25. Tripathi, T., Rahlfs, S., Becker, K. & Bhakuni, V. Glutathione mediated regulation of oligomeric structure and functional activity of *Plasmodium falciparum* glutathione S-transferase. *BMC Struct. Biol.* **7**, 67 (2007).
26. Tang, S. S. & Chang, G. G. Steady-state kinetics and chemical mechanism of octopus hepatopancreatic glutathione transferase. *Biochem. J.* **309**, 347–353 (1995).
27. Tomarev, S. I., Chung, S. & Piatigorsky, J. Glutathione S-transferase and S-crystallins of cephalopods: evolution from active enzyme to lens-refractive proteins. *J. Mol. Evol.* **41**, 1048–1056 (1995).
28. Lou, M. F. Redox regulation in the lens. *Prog Retin Eye Res* **22**, 657–682 (2003).
29. Zetterberg, M. *et al.* Glutathiolation enhances the degradation of gammaC-crystallin in lens and reticulocyte lysates, partially via the ubiquitin-proteasome pathway. *Invest Ophthalmol Vis Sci* **47**, 3467–3473 (2006).
30. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
31. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25 (2010).
32. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst.* **D66**, 486–501 (2010).
33. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
34. Wu, C. G. *et al.* Mechanism for controlling the monomer-dimer conversion of SARS coronavirus main protease. *Acta Crystallogr D Biol Crystallogr* **69**, 747–755 (2013).
35. Lin, M. H. *et al.* Structural and functional characterization of MERS coronavirus papain-like protease. *J. Biomed. Sci.* **21**, 54 (2014).
36. Chou, C. Y. *et al.* Structural basis for catalysis and ubiquitin recognition by the severe acute respiratory syndrome coronavirus papain-like protease. *Acta Crystallogr D Biol Crystallogr* **70**, 572–581 (2014).
37. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
38. Chou, C. Y., Hsieh, Y. H. & Chang, G. G. Applications of analytical ultracentrifugation to protein size-and-shape distribution and structure-and-function analyses. *Methods* **54**, 76–82 (2011).
39. Tang, S. S. & Chang, G. G. Kinetic characterization of the endogenous glutathione transferase activity of octopus lens S-crystallin. *J Biochem* **119**, 1182–1188 (1996).
40. Pace, C. N. Measuring and increasing protein stability. *Trends Biotechnol.* **8**, 93–98 (1990).
41. Chang, H. C., Lin, T. L. & Chang, G. G. Molecular basis for the polymerization of octopus lens S-crystallin. *Biophys. J.* **78**, 2070–2080 (2000).

## Acknowledgements

We would like to thank Gu-Gang Chang and Liang Tong for helpful discussions. This research was supported by grants from Ministry of Science and Technology, Taiwan (101-2320-B-010-061, 103-2320-B-010-025 and 104-2320-B-010-034) and Ministry of Education, Taiwan (Aim for the Top University Plan) to CYC. We thank the technical services provided by the “Synchrotron Radiation Protein Crystallography Facility of the National Core Facility Program for Biotechnology, Ministry of Science and Technology” and the “National Synchrotron Radiation Research Center”, a national user facility supported by the Ministry of Science and Technology, Taiwan, ROC.

## Author Contributions

W.-H.T. and C.-G.W. performed protein expression, purification and crystallization experiments. S.-C.C. and Y.-T.L. performed the cloning, mutagenesis, kinetic assays and thermostability assays. M.-H.L. performed some kinetic assays, analytical ultracentrifugation and the dimerization analysis. C.-C.C. and C.-H.L. performed the mass spectrometry. C.-Y.C. performed the calculations and the light-scattering measurements. W.-H.T. and C.-Y.C. wrote the paper.

## Additional Information

**Accession codes:** Model coordinates and structure factors for the crystal structure of the octopus S-crystallin Q108F mutant was deposited in the Protein Data Bank (PDB code: 5B7C).

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Tan, W.-H. *et al.* Structure of a Highly Active Cephalopod S-crystallin Mutant: New Molecular Evidence for Evolution from an Active Enzyme into Lens-Refractive Protein. *Sci. Rep.* **6**, 31176; doi: 10.1038/srep31176 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016