Journal of Advanced Bioinformatics Applications and Research
ISSN 0976-2604.Online ISSN 2278–6007
Vol 6, Issue2, 2015, pp45-51
http://www.bipublication.com

Prediction of Aggregation Sites in Human Eye Lens γ-Crystallins: Implications in Cataract

Vandna Sharma, Priyanka Chauhan and Kalyan Sundar Ghosh*

Department of Chemistry,
National Institute of Technology Hamirpur, Himachal Pradesh 177005, India
*Corresponding author: e-mail: kalyan@nith.ac.in
Tel: +91-1972-254104; Fax: +91-1972-223834;

Received: 11/04/2015 Accepted: 17/05/2015 Published: 29/05/2015

ABSTRACT:
Cataract, a leading cause of blindness, occurs due to aggregation of crystallin proteins in eye lens. In this paper, the aggregation site in three major γ-crystallins of human eye lens has been identified unambiguously by using their primary sequences with the help of *PASTA*, *Tango* and *AGGRESCAN*. We have found that some hydrophobic residues in C-terminal domain of γ-crystallins play putative role in initiation of aggregation of these proteins. Suppression of aggregation of γD-crystallin by molecular chaperone in lens like αB-crystallin is also documented mechanistically through docking studies.

Keywords: *Cataract, γ-crystallins, aggregation site, hydrophobic interaction, αB-crystallin, molecular docking*

INTRODUCTION

Aggregation of protein molecules has been linked with several neurodegenerative disorders like Alzheimer's, Parkinson's disease etc [1-3]. Cataract, a common disease of old age occurs due to increased light scattering by the aggregated crystallin proteins in eye lens. Three types of crystallins (α-, β- and γ-) contribute to 90% of total protein content of the fiber cells of lens tissue. Optimum homologous and heterologous interactions among crystallins maintain the transparency and refractive index of lens [4, 5].

But with increase in age of an individual, β- and γ-crystallins start to form their partially unfolded species. Alpha-crystallins as molecular chaperone prevent the aggregation of β- and γ-crystallins by binding with their partially unfolded forms to maintain the transparency of the lens [6, 7]. Different environmental stresses like exposure to UV light, oxidizing agents etc. can damage the native conformation of γ-crystallins and generate partially unfolded species for further aberrant protein-protein interactions leading to cataract [8-10]. As there is no turnover of crystallins in aged lens and due to the lack of any suitable way-out for removing aggregated proteins from lens, there is almost no other alternate except eye surgery for treating cataract. Therefore, specific knowledge about aggregation sites in β/γ-crystallins would help to combat against age-onset cataract and based on that, suitable inhibitor could be designed.

In this work, we have explored the sites of aggregation in different γ-crystallins. We have shown that the protein aggregation is triggered by short aggregation prone regions, which act as nucleating sites for aggregation. The buried aggregation prone regions in native conformation become exposed on partial unfolding of proteins and hence the process of aggregation will be initiated. In two earlier works on some other aggregating proteins such as polyQ and ployA,

our group had identified aggregating regions in those proteins and our results were able to explain the experimental findings too [11, 12]. In this paper, the residues having inherent propensity towards aggregation in three major human eye lens γ-crystallins have been identified. The predicted residues seem to be responsible for aggregation of γ-crystallin and this will provide a clue for possible way out for inhibiting aggregation of γ-crystallins.

METHODS

To find out aggregation sites in three human γ-crystallins (γB, γC and γD, further denoted as HGB, HGC and HGD respectively), aggregation propensity calculators like *PASTA, Tango* and *AGGRESCAN* were used. *PASTA2.0* (prediction of amyloid structure aggregation) [13] calculates the aggregation propensity of the residues of a given amino acid sequence and as an output it generates an aggregation profile. *Tango* [14] uses statistical mechanics algorithm for scanning a given primary sequence and evaluates the probability for each residue to adopt the β-aggregate structure. The algorithm also takes into account of other factors like protein stability and some physico-chemical parameters such as pH, concentration, ionic strength etc. While *AGGRESCAN* [15] quantitatively calculates the aggregation-propensity value for each residue in a sequence and from its output aggregation profile, putative aggregation sites (hot spots) can be identified.

For structural studies on these proteins, we have used the three dimensional structures of HGD and HGB downloaded from Protein Data Bank (PDB) [PDB entries 1HK0-chain A; 2JDF-chain A respectively] (Fig. 1 A and B). Due to unavailability of any three dimensional structure of HGC, a *BLAST* search of its amino acid sequence was performed and 84% identity in sequence was found with an available 3D structure in PDB (2V2U-chain A). Therefore, based on significant identity in amino acid sequences and e-value as obtained from *BLAST* search, homology model structure of HGC (Fig. 1C) had been made using 2V2U-chain A as the template. *SWISS-MODEL* work space [16] was used for making the homology model structure of HGC (Fig. 1C). After generating homology model structure, it is essential to assess its structural quality. Initially the model structure of HGC was verified by Ramachandran plot using *Rampage* [17]. The amino acid residues of HGC model structure are falling in two regions: favored (95.3%) and allowed (4.7%) with none of the residues in outlier. A high percentage of residues in favored region for the model structure suggest a highly reliable quality of the built model.

The model structure of HGC was also tested in *QMEAN* server for quality estimation [18]. *QMEAN* generated two scoring functions namely *QMEAN* score and *QMEAN* Z-score. The *QMEAN* score for HGC model structure was 0.694, and *QMEAN* Z-score was -0.73. *QMEAN* score of the model structure was in between 0 and 1, which suggests the model structure is reliable and of good quality.

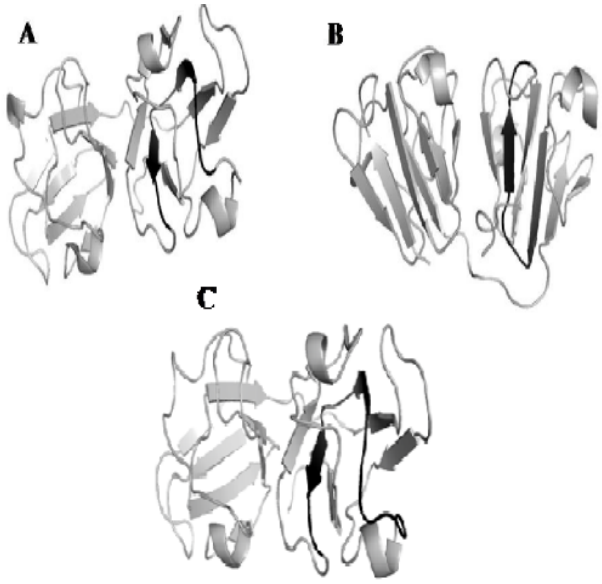


Fig: 1. Structure of human γD –crystallin (A), γB-crystallin (B), γC- crystallin (C).

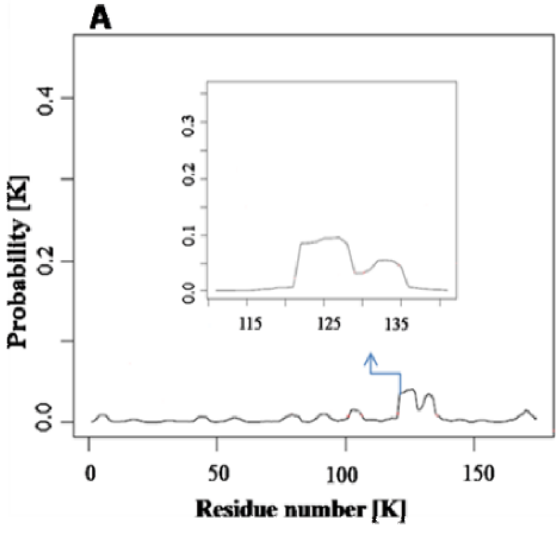
Solvent accessible surface area (ASA) of the amino acid residues was calculated through *NACCESS* [19]. ASA of the residues was calculated using either the structures directly downloaded from PDB (for HGD and HGB) or the homology model structure (for HGC). The identified aggregating region of HGD was docked with human αB-crystallin (HAB). Autodock tools were used to remove all water

molecules from the PDB file of HAB, to add polar hydrogen atoms and to assign Gasteiger charges prior to docking. Docking was carried out using *Autodock 4.2* Lamarckian Genetic Algorithm. All other docking parameters were kept as default. *PyMol* was used for visualization and measurement of distances between interacting residues.

RESULTS AND DISCUSSION

PASTA, Tango and *AGGRECSCAN* were used to predict the residues with high propensity toward aggregation in HGB, HGC and HGD. Aggregation profiles obtained from *PASTA, Tango* and *AGGRECSCAN* for one representative γ-crystallin (HGD) is shown in Fig. 2. We have used three different programs to identify the aggregation sites unambiguously. The residues having high aggregation propensity as obtained from *PASTA, Tango* and *AGGRECSCAN* mutually match significantly well for all three γ-crystallins. We have considered only that region having high aggregation propensity as well as mutual consensus among the outputs of three programs. The identified region has been predicted as aggregation site. The identified residues in the three γ-crystallins are listed in Table 1.

Gamma crystallins contain two homologous domains namely N-terminal domain (N-td) and C-terminal domain (C-td) [20]. The folded structure of HGD contain four Greek key motifs — motif 1 (residues 1-40), motif 2 (42-83), motif 3 (88-128) and motif 4 (131-174). Exceptional



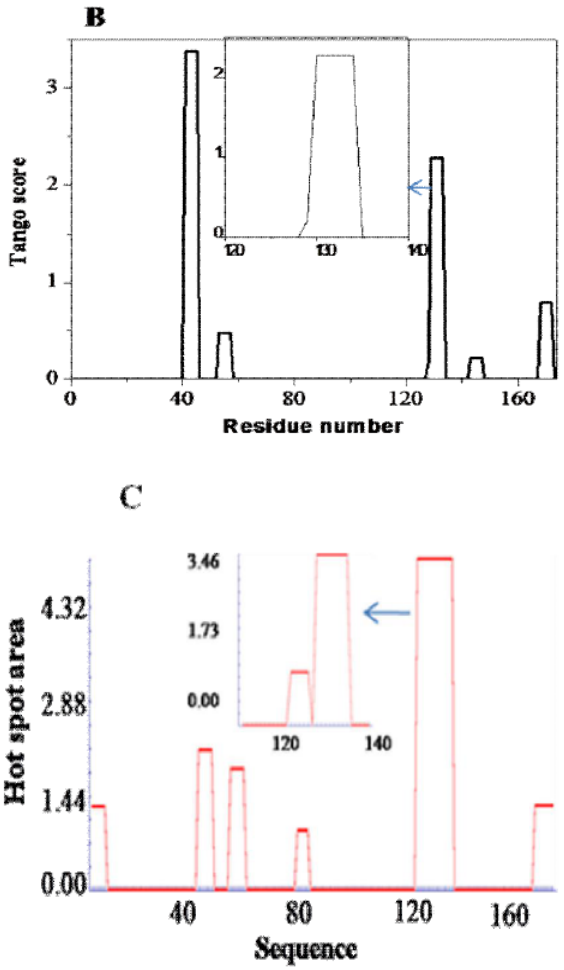


Fig: 2. (A) Aggregation profile obtained from *PASTA* 2.0 for HGD. The aggregation probability [k] is plotted as a function of residue number. The peak having highest aggregation propensity is shown in inset. (B) Predicted aggregation propensity in terms of *Tango* score for the residues of HGD. (c) *AGGRESCAN* output for HGD.

stability of the native state of HGD is generally accounted for the significant stabilizing effects due to intra- and inter-domain hydrophobic interactions [21, 22]. Hydrophobic residues in Ntd like Met 43, Phe 56 and Ile 181 form a cluster with C-td residues Val 132, Leu 145 and Val 170 through hydrophobic contact [21]. The stability of motif 2 in N-td is dependent on interdomain hydrophobic interface [23] and this interface at C-td provides the template for folding of N-td [21, 23, 24].

The aggregation site in HGD indentified by us also includes hydrophobic residues like Ile 121, Leu 124, Val 126, Leu 127, Trp 131, Val 132, Leu 133, Tyr 134, Leu 136 in the C-td (Table 1).

These residues are found to be unaccessible in the native state of HGD (Table 1). As N-td unfolds prior to the C-td [21], therefore partial unfolding of N-td exposes the above mentioned hydrophobic residues in the identified region. This will lead to an initiation of aggregation of HGD molecules through further hydrophobic interactions. Most interestingly, our indentified region encompasses the last part of motif 3, the initial residues of motif 4 and the residues in between these two motifs. The residues (Gly 129 and Ser 130) intervening motif 3 and motif 4 are on a flexible loop structure. This loop structure will be more flexible upon partial unfolding of Ntd. From earlier experimental observation, the residues 80-163 are found at the core of amyloid fibril as identified by MALDI mass, 2D IR spectroscopy and 13C isotope labelling [25, 26].

γ-crystallins	Aggregation predictor			Predicted aggregating region	Secondary structure of the aggregating region	Hydrophobic residues present in that region	ASA of the Hydrophobic residues (Å2)
	PASTA	*Tango*	*AGGRESCAN*				
HGB (containing 175 residues)	122-136 (Highest peak)	41-46 (Highest peak) 130-135 (Second Highest peak)	128-137 (Highest peak)	122-137	β-sheet and loop	Leu 124	0.00
						Val 126	2.08
						Leu 127	71.68
						Trp 131	0.00
						Ile 132	0.00
						Leu 133	0.11
						Tyr 134	1.43
HGC (containing 174 residues)	116-128	41-46 (Highest peak) 129-134 (Second Highest peak)	121-136 (Highest peak)	116-136	β-sheet and loop	Phe 116	2.57
						Leu 118	51.07
						Ile 121	0.00
						Leu 124	0.00
						Val 126	2.69
						Leu 127	42.46
						Trp 131	0.00
						Val 132	0.00
						Leu 133	0.07
						Tyr 134	5.17
						Leu 136	74.44
HGD (containing 174 residues)	119-135 (Highest peak)	41-45 (Highest peak) 129-134 (Second highest peak)	127-136 (Highest peak)	119-136	β-sheet and loop	Ile 121	0.00
						Leu 124	0.00
						Val 126	2.66
						Leu 127	68.72
						Trp 131	0.00
						Val 132	0.34
						Leu 133	0.00
						Tyr 134	4.38
						Leu 136	68.92

Table 1: Predicted residues having aggregation propensity in three γ-crystallins

These residues are mostly from C-td and the linker between two domains. Presence of these residues at the core of the amyloid invariably suggests that aggregation initiates from this region, which again validates our predicted aggregation site.

Aromatic interactions between Tyr residue pair (Y134 and Y139) also contribute significantly to the stability of C-td and N-td as well as that of the whole protein [27]. Y134 is also a part of our predicted aggregation site and any change in this region due to partial unfolding will therefore perturb the stabilizing effects of aromatic interactions. At the same time the "sandwichlike" Y134-R168-Y139 cluster (cation–π interaction) [28] may also be disturbed. In addition to that another residue E135 is also predicted by us which contributes to the stability by forming salt bridge with R142 and this interaction will also be disrupted [22]. All these findings strongly support our observation that the aggregation process of HGD starts from the identified region. One earlier result mechanistically indicates that due to unfolding of N-td, hydrophobic residues

in the region 135-164 of motif 4 in C-td become highly exposed along with partial unstructuring of C-td [29]. Molecular dynamics simulations probed that the residues 135-164 of C-td of one partially unfolded HGD molecule can make frequent contact with the N-td of another HGD molecule to continue the process of aggregation [29]. Therefore, very crucial hydrophobic contact at domain–domain interface in HGD will be significantly compromised. So in case of age-onset cataract, HGD molecules aggregate through successive intermolecular domain swapping and some residues of motif 4 of C-td play significant roles. This is also supporting our predictions on aggregation site mechanistically. Another study on only HGD also predicts that the residue S130 acts as a hot spot for forming amyloid-like structure by computational as well as mutational studies [30].

A similar phenomenon is also expected for HGB and HGC based on their homologous structures with HGD. Our results are able to provide a molecular insight on age-onset cataract due to aggregation of γ-crystallins and it will be a basis to develop strategies to prevent and/or treat cataract. Strong correlation between our predicted aggregating region with earlier experimental findings corroborates our results. Finally molecular docking was performed to visualize the suppression of aggregation of HGD by HAB. The HGD fragment having residues 121 to 134 was docked with HAB. It was observed that the hydrophobic residues in the identified region of HGD are interacting with a major hydrophobic patch on the surface of HAB (composed of Val 91, Val 93, Leu 131 and Ile 133) (Fig. 3). The hydrophobic residues of HGD such as Val 132 and Tyr 134 show prominent interaction with HAB. The distances between the interacting residues of HGD and HAB are given in Table 2.

We have also calculated the ASA of the residues of HAB before and after interaction with HGD. It is found that the ASA of the interacting hydrophobic residues of HAB are diminished on complexation with HGD as shown in Table 3. This suggests prominent hydrophobic interaction between HGD and HAB can inhibit the aggregation of HGD.

Table 2: Distances between interacting residues of HGD and HAB

Interacting residues of HGD	**Interacting residue of HAB (Distance in Å)**
δ1C Val 132	Val 91 δ1C (3.25) Ile 133 γ2C (3.87)
δ2C Val 132	Ile 133 βC (3.80)
δ1C Tyr 134	Val 93 βC (3.29) Val 93 δ1C (3.88)
ε1C Tyr 134	Val 93 βC (3.55) Val 93 δ1C (3.59)
ε2C Tyr 134	Leu 131 δ1C (3.43)
ζC Tyr 134	Leu 131 δ1C (3.45)

Distances between interacting residues of HGD and HAB crystalline (up to 4A0)

Table 3: ASA of HAB residues before and after interaction with HGD

Interacting residues of HAB	**ASA of residues before interaction with HGD (Å2)**	**ASA of residues after interaction with HGD (Å2)**
Val 91	53.79	1.66
Val 93	53.53	6.394
Leu 131	147.65	99.077
Ile 133	37.117	17.238
Leu 137	62.93	44.87

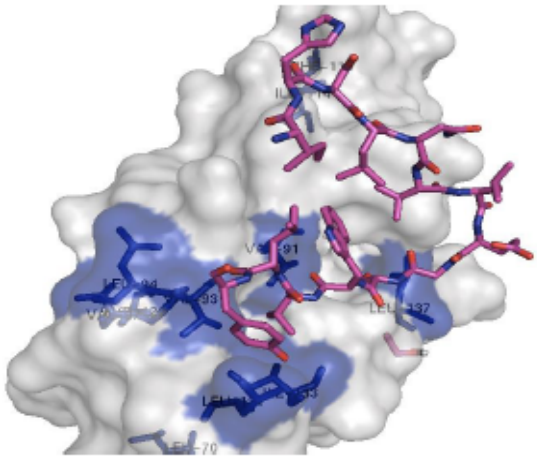


Fig: 3. Docked conformation of identified region of HGD with HAB. HGD is shown in pink color and HAB is represented as surface HAB. Hydrophobic patch on the surface of HAB is shown in blue color and interacting residues of HAB are labeled by their respective residue number.

ACKNOWLEDGEMENT

KSG is grateful to Science and Engineering Research Board (SERB) under Department of Science and Technology, Govt. of India for funding through its sanctioned project (No. SB/FT/LS-277/2012). VS is thankful to the same agency for her project assistantship.

REFERENCES

1. Aguzzi, A and O'Connor, T, (2010), Protein aggregation diseases: pathogenicity and therapeutic perspectives, Nature Reviews Drug Discovery, vol-**9**, issue 3, pg 237-248.
2. Jucker, M, and Walker, L. C, (2013), Self propagation of pathogenic protein aggregates in neurodegenerative diseases, Nature, vol-501, issue 7465, pg 45–51.
3. Pedersen, J. T and Heegaard, N. H, (2013)**,** Analysis of protein aggregation in neurodegenerative disease, Anal. Chem, vol-85, issue 9, pg 4215–4227.
4. Benedek, G. B, (1971), Theory of transparency of the eye, Applied Optics vol. 10, issue 3, pg 459-
   473.
5. Tardieu, A, (1988), Eye lens proteins and transparency: from light transmission theory to solution X-Ray structural analysis, Annual Review of Biophysics and Biophysical Chemistry,vol-17, pg 47-70.
6. Horwitz, J, (1992), α-Crystallin can function as a molecular chaperone, Proc. Natl. Acad. Sci. USA
   vol-89, issue 21, pg 10449-10453.
7. Sampson, L. A and King, J, (2010), Partially folded aggregation intermediates of human γD-,γC-, and γS-crystallin are recognized and bound by human αB-crystallin chaperone, J. Mol. Biol. vol-401, issue 1, pg 134–152.
8. McCarty, C. A., Taylor, H. R., 2002, A review of the epidemiologic evidence linking ultraviolet radiation and cataracts. Dev. Ophthalmol. 35, 21-31.
9. Spector, A, (1984), Oxidation and cataract, In:human cataract formation, Ciba Foundation Symposium 106, Pitman, London, pg 48-64.
10. Harding, J, (1991), Cataract: Biochemistry, epidemiology and pharmacology, Chapman & Hall, London.
11. Ghosh, K. S. (2013), Inherent aggregation propensity of flanking residues attached to polyglutamines: implication to aggregation inhibition, *Prot. Pept. Lett. 20*, 569-572.
12. Sharma, V., Ghosh, K. S. (2014), Involvement of non-polyalanine (polyA) residues in aggregation of polyA proteins: Clue for inhibition of aggregation *Comp. Biol. Chem. 53*, 318–323.
13. Walsh, I, et al., (2014), PASTA2: An improved server for protein aggregation prediction, Nucleic Acids Res. 42(web server issue) W301-W307.
14. Fernandez-Escamilla, A. M, (2004), Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins, Nature Biotechnology, vol- 22,issue 10, pg 1302-1306.
15. Solé, C et al., (2007), AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides, BMC Bioinf., vol-8, pg 65.
16. Arnold, K, (2006), The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling, Bioinformatics, vol-22, issue 2, pg 195-201.
17. Lovell, S.C, et al., (2002), Structure validation by Calpha geometry: phi,psi and Cbeta deviation, *Proteins*, vol-50, pg 437-450.
18. Benkert, P, et al., (2009), QMEAN server for protein model quality estimation, Nucleic Acids Res, 37(web server issue) W510-4
19. Hubbard, S. J., Thornton, J. M., (1993), 'NACCESS' Computer Program, Department of Biochemistry and Molecular Biology, University College London.
20. Basak A, et al. (2003) High-resolution X-ray crystal structures of human gammaD crystalline (1.25 angstrom) and the R58H mutant (1.15 angstrom) associated with aculeiform cataract. J Mol Biol 328:1137–1147
21. Flaugh, S. L, et al., (2005), Contributions of hydrophobic domain interface interactions to

the folding and stability of human γD-crystallin, Prot. Sci. vol- 14, issue 3, pg 569–581.

22. Das, P, et al., (2010), β-strand interactions at the domain interface critical for the stability of human lens γD-crystallin, Prot. Sci. vol-19, issue 1, pg 131-140.
23. Flaugh S. L, Kosinski-Collins MS, King J (2005) Interdomain side-chain interactions in human gamma D crystallin influencing folding and stability. Protein Sci 14:2030–2043.
24. Flaugh SL, Mills IA, King J (2006) Glutamine deamidation destabilizes human gammaDcrystallin and lowers the kinetic barrier to unfolding. J Biol Chem 281:30782–30793.
25. Moran, S. D, et al., (2012), Structural and Sequence Analysis of the Human γD-Crystallin Amyloid Fibril Core Using 2D IR Spectroscopy, Segmental 13C Labeling, and Mass Spectrometry, J. Am. Chem. Soc., vol-134, issue 44, pg 18410−18416.
26. Moran, S. D, et al., (2012), Two-dimensional IR spectroscopy and segmental 13C labeling reveals the domain structure of human γD-crystallin amyloid fibrils, Proc. Natl. Acad. Sci. USA vol-109, issue 9, pg 3329–3334.
27. Kong, F, and King, J, (2011), Contributions of aromatic pairs to the folding and stability of longlived human γD-crystallin, Prot. Sci., vol-20, issue 3, pg 513-528.
28. Yang, Z, et al., (2014), Dissecting the contributions of β-hairpin tyrosine pairs to the folding and stability of long-lived human γDcrystallins, Nanoscale, vol-6, issue 3, pg 1797-807.
29. P. Das, J. A. King, and R. Zhou, Aggregation of γ-crystallins associated with human cataracts via domain swapping at the C-terminal β-strands, Proc. Natl. Acad. Sci. USA (201) 10514–10519 vol. 108 no. 26.
30. Sahin, E, et al., (2011) Computational design and biophysical characterization of aggregation resistant point mutations for γD crystalline illustrate a balance of conformational stability and intrinsic aggregation propensity, Biochemistry, vol-50, pg 628-39.