

Guanidine-HCl Dependent Structural Unfolding of M-Crystallin: Fluctuating Native State Like Topologies and Intermolecular Association

Ravi Pratap Barnwal^{*‡}, Geetika Agarwal, Kandala V. R. Chary*

Department of Chemical Sciences, Tata Institute of Fundamental Research, Colaba, Mumbai, India

Abstract

Numerous experimental techniques and computational studies, proposed in recent times, have revolutionized the understanding of protein-folding paradigm. The complete understanding of protein folding and intermediates are of medical relevance, as the aggregation of misfolding proteins underlies various diseases, including some neurodegenerative disorders. Here, we describe the unfolding of M-crystallin, a $\beta\gamma$ -crystallin homologue protein from archaea, from its native state to its denatured state using multidimensional NMR and other biophysical techniques. The protein, which was earlier characterized to be a predominantly β -sheet protein in its native state, shows different structural propensities (α and β), under different denaturing conditions. In 2 M GdmCl, the protein starts showing two distinct sets of peaks, with one arising from a partially unfolded state and the other from a completely folded state. The native secondary structural elements start disappearing as the denaturant concentration approaches 4 M. Subsequently, the protein is completely unfolded when the denaturant concentration is 6 M. The ^{15}N relaxation data (T_1/T_2), heteronuclear ^1H - ^{15}N Overhauser effects (nOes), NOESY data, and other biophysical data taken together indicate that the protein shows a consistent, gradual change in its structural and motional preferences with increasing GdmCl concentration.

Citation: Barnwal RP, Agarwal G, Chary KVR (2012) Guanidine-HCl Dependent Structural Unfolding of M-Crystallin: Fluctuating Native State Like Topologies and Intermolecular Association. PLoS ONE 7(12): e42948. doi:10.1371/journal.pone.0042948

Editor: Andreas Hofmann, Griffith University, Australia

Received: March 2, 2012; **Accepted:** July 16, 2012; **Published:** December 17, 2012

Copyright: © 2012 Barnwal et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The facilities provided by the National Facility for High Field NMR, and the support of the Department of Science and Technology, Department of Biotechnology, Council of Scientific and Industrial Research, and Tata Institute of Fundamental Research, Mumbai, are greatly acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: barnwal@u.washington.edu (RPB); chary@tifr.res.in (KVR)

‡ Current address: Department of Chemistry, University of Washington, Seattle, Washington, United States of America

Introduction

The phenomenon of protein folding-unfolding and its relevance to diseases [1,2,3,4,5,6] is yet to be understood. Protein folding always starts enthalpically/entropically from a state, which is non-native or denatured state and goes into a stable state i.e., native state. Deciphering the structural and dynamic basis of how a given protein sequence translates into a folded conformation requires understanding of the intermediate state(s) of the protein during the process of protein folding/unfolding. The information about the structure and dynamics of different protein states is useful to understand the enthalpically or entropically driven folding/unfolding processes in a given protein, which in turn help in the understanding of the design of protein function. In this endeavor, solution structures are ideally suited for studying such functional states. Earlier studies suggest that proteins exhibit different secondary structural elements under extreme conditions, and sometimes even possesses residual structures [2,7,8,9,10,11,12,13].

On the other hand, the eye-lens, its architecture and its evolution has been subjects of research for quite a long time [14,15]. The transparency of the lens is due to its complex architecture and due to the presence of high protein concentration. The function of the eye-lens is to focus the incoming light on the retina. The lens proteins, namely crystallins, accomplish this task of focusing the incoming light. Eye-lens crystallins are of three

kinds; α -, β - and γ - crystallins. The lens α -crystallins are known to act as molecular chaperones, while lens β - and γ -crystallins, which were thought to be structural proteins, were shown recently to have diverse roles in various cellular processes [16,17]. The β - and γ -crystallins are grouped under $\beta\gamma$ -crystallin superfamily, because of their structural similarity and these proteins comprise of two consecutive Greek key motifs, which together form a stable eight-stranded β -sheeted sandwich structure [18,19]. Some of the $\beta\gamma$ -crystallins that include several bacterial homologues have been shown to bind Ca^{2+} [20,21,22,23,24]. When the eye-lens becomes opaque, it results in cataract, a disorder of the eye. Causes of cataract include: mutation in one of the lens-crystallin proteins, ageing process, diabetes, and environmental factors (i.e., change in the pH, temperature and UV-exposure). During these processes, unfolding of any crystallin protein may occur and lead to protein aggregation or altered interaction/association between native crystallins, due to insolubility of proteins. Therefore, it is important to study the complete unfolding pathway of these proteins, which could aid in understanding cataract related problems.

In this backdrop, we have used M-crystallin, a putative $\beta\gamma$ -crystallin protein (accession NP_617429) from the genome of an archaea *Methanosarcina acetivorans*, as a model protein to study the folding/unfolding behavior of this oldest relative of $\beta\gamma$ -crystallins (Figure S1) [25]. We have earlier solved the 3D structure of M-

crystallin in its Ca^{2+} -bound form by NMR [25]. A temperature dependent oligomerization of this protein has been reported to have significance in the cataract [26]. The ^{15}N relaxation data (T_1/T_2), heteronuclear ^1H - ^{15}N Overhauser effects (nOes), NOESY data, and other biophysical data of M-crystallin taken together indicate that the protein shows a consistent, gradual change in its structural and motional preferences with increasing GdmCl concentration.

Results and Discussion

Resonance assignments of M-crystallin in 4 and 6 M GdmCl

The backbone ^1H , ^{13}C and ^{15}N resonance assignments of M-crystallin in 4 and 6 M GdmCl were carried out to an extent of 94 and 96%, respectively, using a suite of 3D experiments (HNCACB, HN(CO)CACB, HNCO and HN(CA)CO), as reported earlier [27,28]. Inherent spectral overlaps caused by the chemical shift degeneracies in ^1H and ^{13}C spins could be resolved using other experiments [29,30,31], namely (3, 2)D HNHA, (3, 2)D HNHB, (3, 2)D CB(CACO)NHN, (3, 2)D CT-HCCH-COSY (Figure 1 and Figures S2, S3A/B). All the chemical shifts, thus obtained, were deposited in BMRB under the accession numbers 15918 and 15934, respectively. The spectral signatures of the first two N-terminal residues (M1 and A2), were absent under both the denaturing conditions, as in the native state.

GdmCl induced unfolding as studied by CD and fluorescence

GdmCl induced unfolding of M-crystallin was characterized by CD and fluorescence emission spectra [32,33,34] as described in Materials and Methods. Figure 2 shows the denaturation curves of M-crystallin as obtained by far-UV CD and fluorescence spectroscopy, wherein the ellipticity at 218 nm (Figure 2A) and the Trp emission at 331 nm (Figure 2B) were monitored, respectively. The normalized fluorescence and CD curves (not shown here) show similar trend. Almost no changes were noticed

till the GdmCl concentration reached 1.1 M. Between 1.1 and 2.8 M, there was a sudden change, with no change thereafter (Figure 2B). As it is evident from the Figures 2A and 2B, both the curves depict protein unfolding as a two-state model. Transition mid-point and slope of the transition (m_1) thus determined from the fluorescence data were 1.9 M and $2.84 \text{ kCal mol}^{-1} \text{ M}^{-1}$, respectively.

GdmCl dependent [^{15}N - ^1H]-HSQC spectra

We used NMR to study the order-disorder transitions in M-crystallin under various denaturing conditions ranging from 0 to 6 M GdmCl. For this purpose, we recorded a set of [^{15}N - ^1H]-HSQC of M-crystallin taken in 0, 0.8, 1.7, 2.0, 3.1, 4.0, 5.2 and 6.0 M GdmCl (Figure S2). As evident from Figure 1 and Figure S2, the protein was in a completely folded state at 0 M GdmCl concentration, with a single set of peaks. However, we started observing two sets of peaks as the GdmCl concentration approached 2 M (Figure S2), one set of peaks corresponding to the folded state (well dispersed peaks; as seen at 0 M) and the other set corresponding to a partially unfolded state (with the $^1\text{H}^{\text{N}}$ chemical shifts in a narrow range of chemical shifts; between 7.8–8.8 ppm). This indicated that the protein was not in a completely unfolded state at the above-mentioned denaturant concentration. This observation of protein in its two states, correlates with the biophysical data described above. Further, as the denaturant concentration approached 3 M, the protein showed spectral signatures largely expected from the completely unfolded state of the protein. At this concentration and beyond, only one set of ^{15}N - $^1\text{H}^{\text{N}}$ peaks showed-up with the corresponding $^1\text{H}^{\text{N}}$ chemical shifts in a narrow range of 7.8–8.8 ppm. This observation is consistent with the fluorescence data discussed above, which indicates loss of the tertiary structure at higher GdmCl concentrations beyond 2.8 M. However, it is worth mentioning here that the protein showed some transient residual structures at denaturant concentrations greater than 3 M, which we probed and characterized, as discussed later.

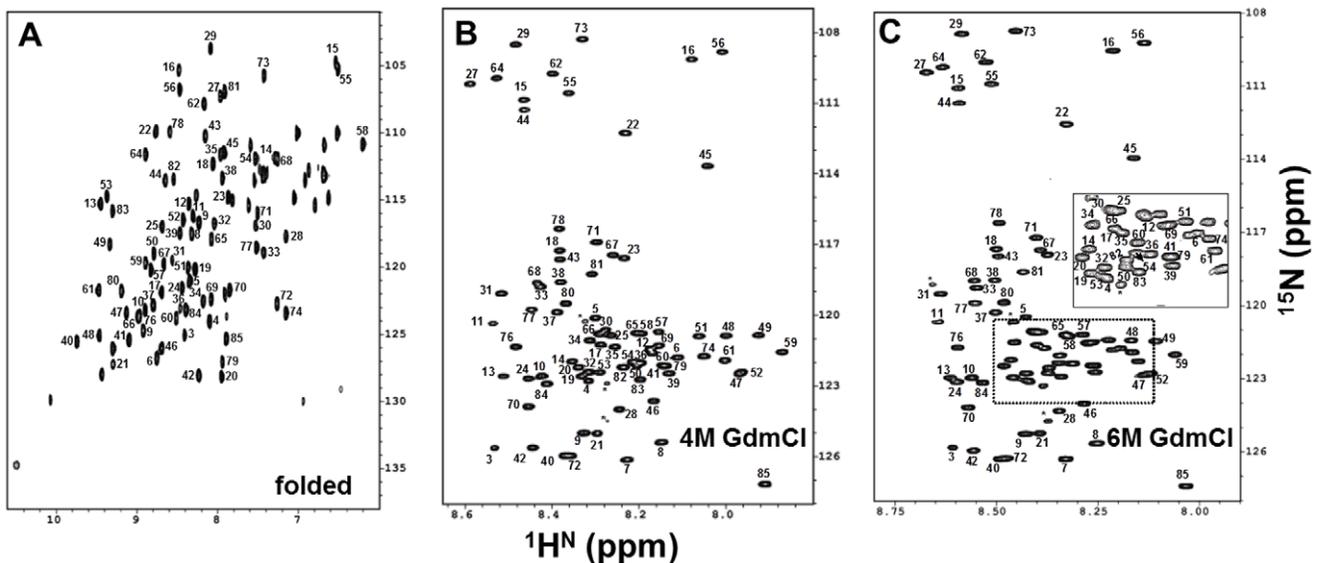


Figure 1. Sensitivity-enhanced 2D [^{15}N - ^1H]-HSQC of M-crystallin in the presence of (A) calcium at pH 7.5 and 298 K (folded), (B) 4 M GdmCl, and (C) 6 M GdmCl, at pH 5.5 and 298 K. These spectra were recorded on Bruker Avance 800 MHz spectrometer with 128 and 1024 points along t_1 and t_2 , dimensions, respectively. Individual peak assignments are shown by the corresponding single-letter code of the amino acid residue and its sequence number along the primary sequence. doi:10.1371/journal.pone.0042948.g001

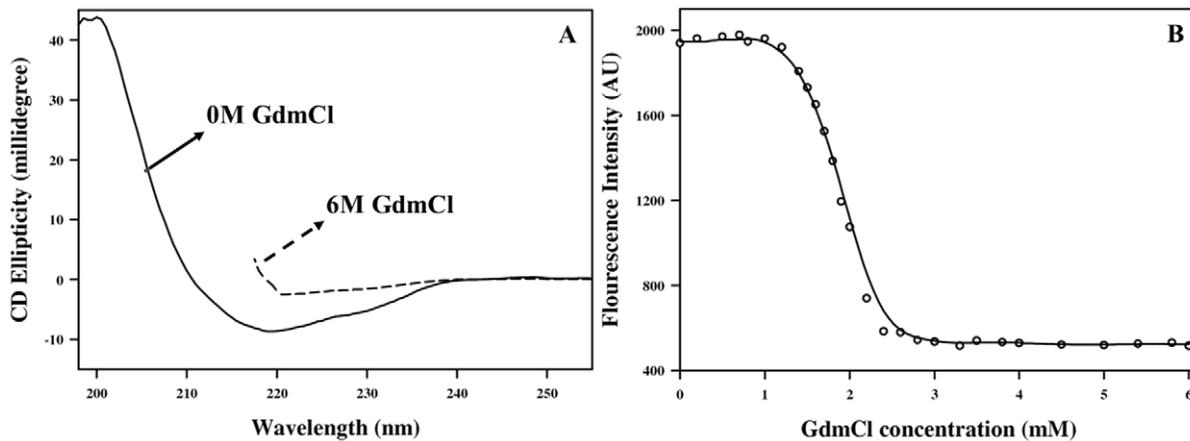


Figure 2. GdmCl dependent denaturation profile of M-crystallin. (A) Far-UV CD spectra of M-crystallin at pH 5.5 and 25°C, in the absence of GdmCl (0 M, solid line) and in presence of GdmCl (6 M, dotted line). (B) Fluorescence spectra of GdmCl-induced unfolding of M-crystallin. The line through the points is the fit to the two-state denaturation model. Protein was used at a concentration of 0.2 mg ml⁻¹ in 50 mM Tris-Cl (pH 5.5) buffer containing 50 mM NaCl.

doi:10.1371/journal.pone.0042948.g002

Structural preferences at different GdmCl concentrations and its comparison with folded protein

(a) Secondary structure chemical shift preferences (ΔC^{α} , ΔC^{β}) and $^3J(^1H^N-^1H\alpha)$. The secondary structural preferences of M-crystallin in 4 and 6 M GdmCl were estimated from the analysis of empirical relations of ¹³C chemical shifts (ΔC^{α} - ΔC^{β}) [35] (shown in Figure 3A) and $^3J(^1H^N-^1H\alpha)$ three-bond coupling constants (Figures S3 and S4). While, the folded protein has a rigid structure, which is predominantly β -sheet in structure, and made up of seven well defined β -strands (Val 6-Glu 10, Ser1 8-Ala 21, Ser 38-Val 41, Thr 45-Tyr 49, Trp 59-Gly 62, Gly 64-Tyr 66 and Ser 81-Gln 84), the protein under denatured condition shows similar secondary structural propensities in these regions. As discussed earlier by several researchers, the denatured state(s) of any given protein may either adopt a completely random coil conformation [36,37,38] or may have regions which adopt preferred conformations or secondary structure propensities for transient structure formations [39,40,41]. The regions, which have certain secondary structural propensities, are termed as folding cores. These folding cores indicate a possible initiation of the folding reaction upon dilution of the denaturant concentrations. Figure 3A suggests that M-crystallin retained its predominantly β -sheet conformation during its unfolding pathway. However, three polypeptide stretches (Asp 20-Ala 28, Asp 60-Tyr 66 and Ala 72-Asn 77) did show propensities for helical conformation for the M-crystallin in 6 M GdmCl.

As mentioned below in Materials and Methods, $^3J(^1H^N-^1H\alpha)$ were calculated from GFT (3, 2)D-HNHA spectra. We could measure $^3J(^1H^N-^1H\alpha)$ values for 77 out of 80 residues (>95%) in the native state of the protein, while the number of 3J values were 75 and 68 for the M-crystallin in 4 and 6 M GdmCl (Figure S4). These 3J couplings are highly sensitive to the backbone torsion angle ϕ and carry the information about the conformational preference of individual amino acid residues. The $^3J(^1H^N-^1H\alpha)$ for residues in α -helical segments (α -helix and PP_{||}) range between 4.0 and 5.5 Hz, while these values are >8 Hz for residues involved in β -strand structures and 5.5–8.0 Hz for random coil stretches. For the native state of M-crystallin (2k1w), while ~3% residues adopt an α -helical conformation, 42% were in β -sheet and 55% in random coil conformation. In the presence of GdmCl, most of these couplings range between 5.5 and 8.5 Hz, indicating their

involvement in an extended structure with some kind of conformational averaging. However, for the polypeptide stretches, Thr22-Gln25, Gly64-Ser68 and Glu70-Ile74, the measured coupling constants were less than 5 Hz indicating α -helical or PP_{||} preferences for these stretches.

(b) Temperature coefficients and residual structure formation. Amide proton (¹H^N) exchange rates and their temperature coefficients throw light on hydrogen bonding (if any), both in globular or denatured state of any protein [42]. If an ¹H^N is hydrogen bonded, it would have a temperature coefficient more positive than -4.5 ppb/K. Random coiled or non-hydrogen bonded ¹H^N tends to have more negative values than -4.5 ppb/K. In the present study, the ¹H^N temperature coefficients were measured by recording a set of HSQC spectra at different temperatures ranging from 15 to 36°C, at an interval of 3°C, for all the three states of the protein (Figure 3B). As seen from Figure 3B, the protein in its native state (0 M GdmCl) displayed the presence of several continuous polypeptide stretches involved in hydrogen bonding. On the other hand, the protein under denatured conditions (4 M and 6 M GdmCl) exhibits largely extended conformations, with very few signatures of intramolecular hydrogen bonding. The polypeptide stretches that showed signatures of their involvement in hydrogen bonding were Thr22-Asp24, Glu65-Tyr66, and Asn77-Ser78. These polypeptide stretches showed temperature coefficients more positive than -4.0 ppb/K values even in 4 M GdmCl, hinting at their possible involvement in some kind of residual structures even under such denatured conditions.

Protein dynamics in the denatured states

The ¹⁵N spin-lattice/spin-spin relaxation times (T_1/T_2), and heteronuclear ¹H-¹⁵N Overhauser effects (NOE) were measured at 25 °C and at two different external magnetic fields (600 and 800 MHz) for the folded protein as well as the protein in the denatured states to throw light on GdmCl induced motional perturbations. We carried out relaxation analysis of 69 and 65 residues of the protein in 4 and 6 M GdmCl, respectively. Relaxation data for the folded protein was taken from our earlier reported work [26]. Three prolines and two N-terminal residues, which do not show their spectral signatures could not be part of

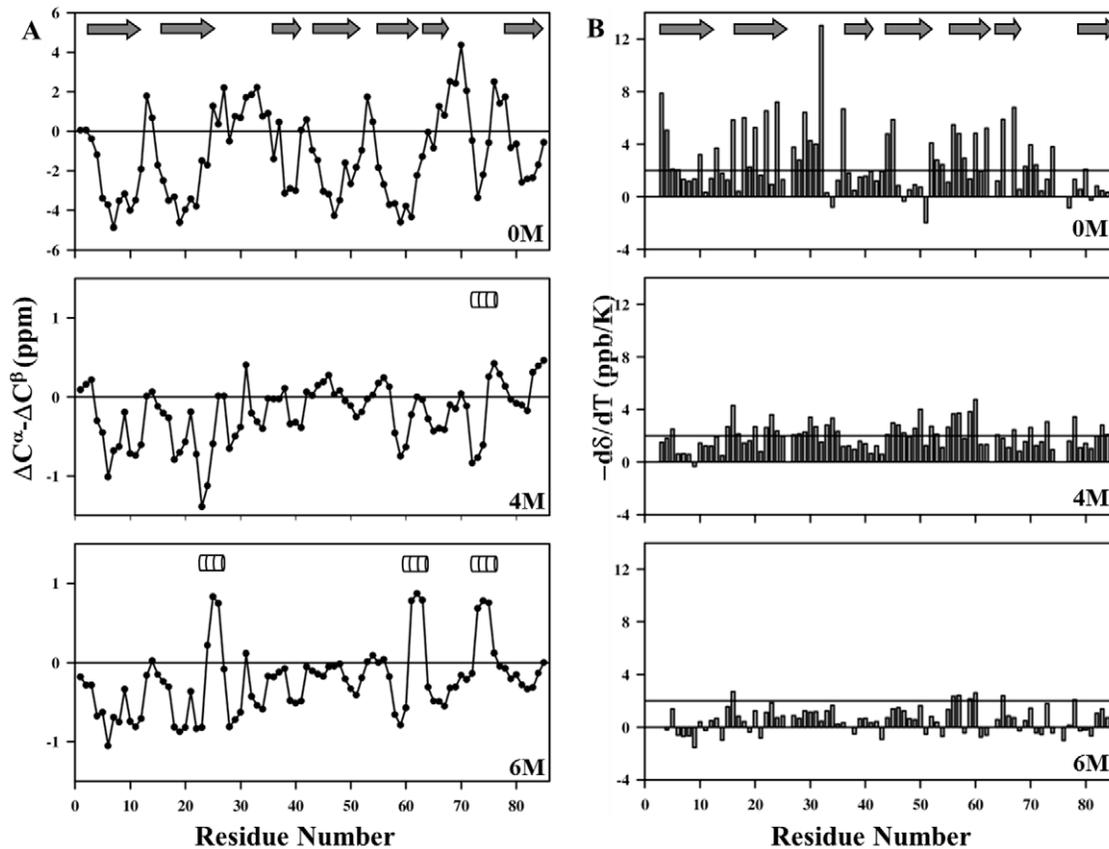


Figure 3. Secondary structure elements and temperature coefficients of M-crystallin. (A) Plot of secondary chemical shifts ($\Delta C^{\alpha}-\Delta C^{\beta}$) for M-crystallin with 4 and 6 M GdmCl concentrations, and their comparison with that of the folded monomer. New structural elements seen are shown with boxes. (B) Residue-wise temperature coefficients of M-crystallin taken in 0, 4 and 6 M GdmCl concentrations. The bar line was fixed at a value +2. Residues representing the secondary structure elements in folded case (0 M) are shown as arrows at top of the both the plot panels A and B. doi:10.1371/journal.pone.0042948.g003

this dynamic study either in the native or the denatured state of the protein.

(a) Relaxation parameters. The longitudinal relaxation (R_1) rates are highly field dependent and are sensitive to nanosecond-picosecond time scale motion. R_1 remained constant across the polypeptide stretch with average values of 1.9 ± 0.1 , 1.78 ± 0.06 and $2.65 \pm 0.1 \text{ s}^{-1}$ for the protein in 6, 4 and 0 M GdmCl, respectively (Figure 4A and Table 1). Both termini of the protein expectedly showed lower values of R_1 suggesting faster time scale motions at both ends. Significantly, large values of R_1 is seen for residues Asn 31, Lys 35, Ser 38, Tyr 51, Gly 55, Ser 68, Ser 71, and Ile 79 at 0 M; His 11, Glu 50, Asn 53, Asn 77, and Ser 78, at 4 M and Ser 23, Asp 24, Gly 27, Trp 59 and Ser 78 at 6 M GdmCl concentrations (Figure 4A and Figure S5A).

The transverse relaxation (R_2) rates are important to get information about the conformational transition in the millisecond to microsecond time-scale regime. They showed larger variation all along the sequence ranging from 2.66 to 6.58 s^{-1} , 1.56 to 8.54 s^{-1} and 5.01 to 14.18 s^{-1} for the protein in 6, 4 and 0 M GdmCl (Figure 4B). The polypeptide stretches Val 8-Phe 14, Asp 20-Gln 25, and Lys 42-Arg 57 showed higher R_2 values ($>6 \text{ s}^{-1}$) for the protein in 4 and 6 M GdmCl concentrations. Here, Ala 21 and Ala 28 show lower values of R_2 compared to their respective values seen in the native state of the protein [26]. These residues, belonging to a different Greek keys, are seen to have a higher value of exchange and R_2 value in case of native protein [26]. The N-terminal residues and Lys 17, Ala 28, Gly 29, Gly73, and Ile74

residues also show lower R_2 values due to their presence at the edge of structural propensities. This is also supported by the fact that these residues are in the neighborhood of residues Gly, Pro or Ala, which are supposed to create hinges with in a polypeptide stretches and/or act as secondary structure breakers.

The ^1H - ^{15}N steady state nOes have information about high frequency motions as is evident in Figure 4C, three N-terminal residues Asn 3, Ala 4 and Glu 5, and two C-terminal residues Gln 84 and Ile 85 show negative nOes, as a result of their higher order flexibility (Figure 4C and Figure S5C) in the presence of denaturant (6 and 4 M GdmCl), while the protein in its native state (at 0 M GdmCl) does not show such higher order flexibility at the C-terminal end. All the relevant R_1 , R_2 and nOe data are provided in Table 1.

(b) Spectral densities. ^{15}N -relaxation data acquired on 600 MHz NMR spectrometer were used to calculate the reduced spectral density functions (J values) at different frequencies (0, 60 and 540 MHz). These $J(\omega_H)$, $J(\omega_N)$ and $J(0)$ values were calculated based on Equations 2–5 (see Methods; Figure 4D–F). The $J(0)$ values showed the largest variation across the sequence and are mostly based on R_2 , whereas the $J(\omega_H)$ values, which are largely determined by nOe data, exhibited a lot of variation in the regions belonging to terminal residues, Asp 20-Ala 28, Gly 44-Ile 52, Asp 60-Tyr 66 and Ala 72-Asn 77. The $J(\omega_N)$ values, which are largely dependent on R_1 , were found to be constant for most of the residues, except for regions at the N and C-terminals, and the polypeptide stretches Thr 22-Ala 28, and Gly 73-Ser 80. A higher

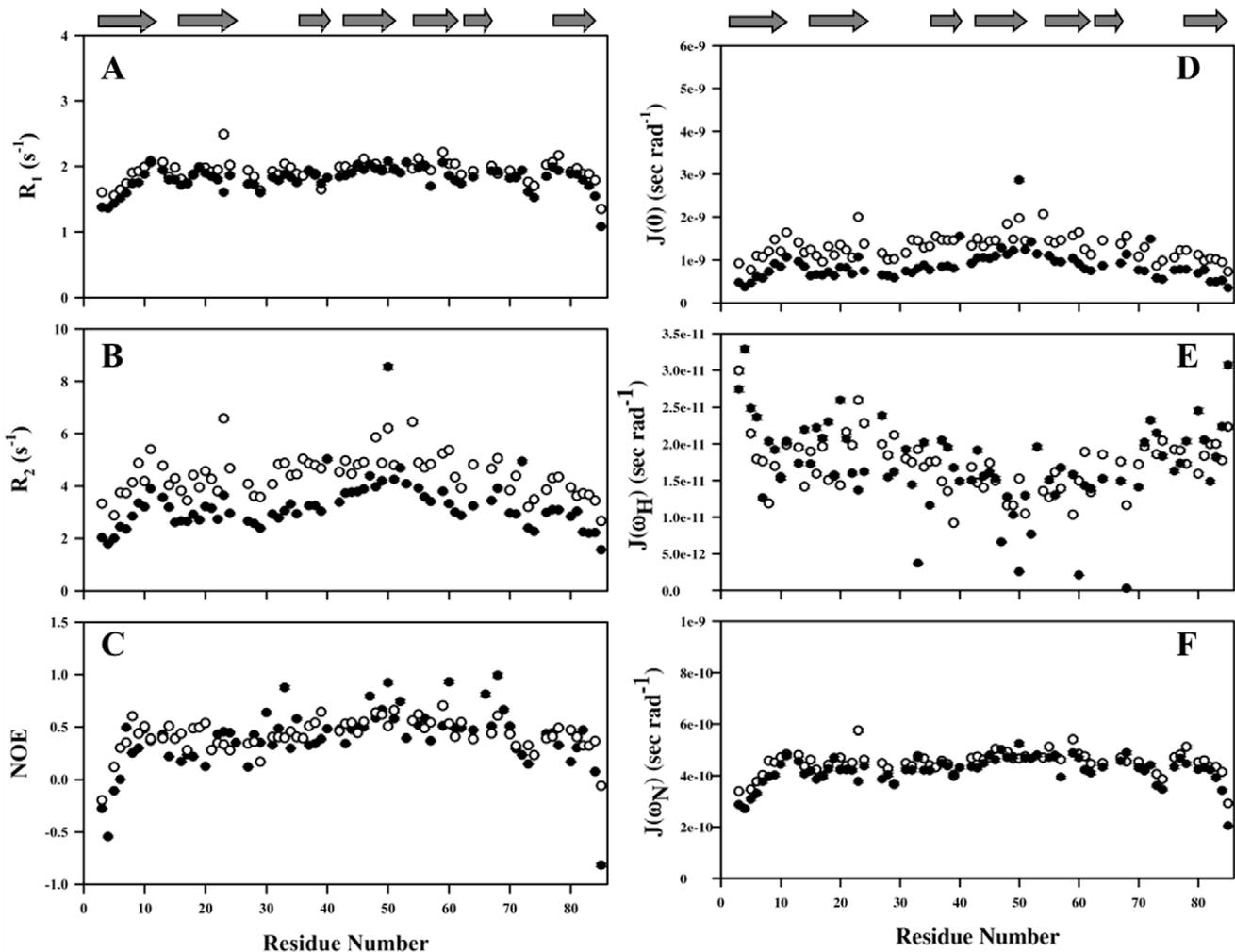


Figure 4. ^{15}N relaxation parameters and reduced spectral densities versus residue number. (A) Longitudinal relaxation times (T_1). (B) Transverse relaxation times (T_2). (C) ^{1}H - ^{15}N NOE enhancements are defined as $I_{\text{sat}}/I_{\text{eq}}$ where I_{sat} and I_{eq} are the intensities of peak in the 2D experiments with and without proton saturation described above, respectively. Error bar of the T_1 and T_2 data denote curve-fitting uncertainties; errors in the ^{1}H - ^{15}N NOEs are estimated from the signal/noise ratio of the spectra. Spectral densities are shown in (D) $J(0)$, (E) $J(\omega_H)$, and (F) $J(\omega_N)$, as a function of protein sequence number. Here, filled and open circles correspond to the 4 and 6 M GdmCl cases, respectively. Secondary structure elements seen in the folded protein (0 M) are depicted at the top of both the plots. doi:10.1371/journal.pone.0042948.g004

degree of variation was observed for the terminal residues at both 4 and 6 M concentrations of the denaturant. In the native state, large variations in $J(\omega_H)$ and $J(\omega_N)$ values were seen for polypeptide stretches Ala 28-Asn 33 and Asn 77-Arg 83. These residues showed higher $J(0)$ values and have chemical exchange contribution under native state [26]. The calculated average values of $J(0)$, $J(\omega_H)$ and $J(\omega_N)$ were 2.96 ns rad^{-1} , 4.35 ps rad^{-1} and $0.539 \text{ ns rad}^{-1}$, respectively at 0 M concentration, whereas they were $0.864 \text{ ns rad}^{-1}$, $17.08 \text{ ps rad}^{-1}$, $0.421 \text{ ns rad}^{-1}$ at 4 M and 1.29 ns rad^{-1} , $17.09 \text{ ps rad}^{-1}$, $0.449 \text{ ns rad}^{-1}$ at 6 M concentration, respectively (Figure 4 D-F. Figure S5 D-F).

(c) Correlation times. The correlation time provides details into different time-scale motions observed in a polypeptide chain. The linear correlation between $\mathcal{J}(\omega_{N,H})$ and corresponding $\mathcal{J}(0)$ values using the Equation $\mathcal{J}(\omega_{N,H}) = \alpha\mathcal{J}(0) + \beta$, and Equation (6) were used to calculate correlation times under different denaturing conditions (Figure 5 A-D). The calculated correlation times were 4.67, 1.39 and 1.68 ns for the native protein (0 M GdmCl); 3.54, 1.09, 1.28, 4.42, and 0.071 ns for the protein in 4 M GdmCl, and

3.97, 1.01, 1.27, 10.06 and 0.062 ns for the protein in 6 M GdmCl (Table 2). These values clearly depict different time-scale motions under different denaturant conditions. The correlation time 10.06 ns associated with the protein in 6 M GdmCl corresponds to global correlation and this value is double the value obtained for the natively folded protein. This is probably due to the following reasons: (i) complete unfolding of the protein, (ii) higher viscosity, as seen in the presence of 6 M GdmCl, and (c) intermolecular association under unfolding conditions [43,44]. The shorter correlation times 1.01, 1.27, and 0.062 ns reflect independently fluctuating segments of the protein in 6 M GdmCl whereas 3.97 ns corresponds to the overall tumbling time of M-crystallin in its native form (Table 2).

(d) Conformational exchanges (R_{ex}) contribution. The R_2 values are representative of slow time-scale motions including conformational exchanges. They provide evidence for motional restrictions and flexibilities in native as well as in denatured proteins. The R_2 values determined at different denaturing conditions are shown in Figure 6A. As evident in Figure 6A, R_2

Table 1. Comparison of Backbone Dynamics Parameter for greek-key1 and greek-key2 in M-crystallin.

	0 M GdmCl	4 M GdmCl	6 M GdmCl
Motional parameters for residues with secondary structures			
Greek key 1			
Average R1(s)	2.67 (± 0.09)	1.79 (± 0.09)	1.85 (± 0.10)
Average R2(s)	7.11 (± 0.37)	3.13 (± 0.21)	4.36 (± 0.17)
Average NOEs	0.71 (± 0.03)	0.35 (± 0.05)	0.45 (± 0.04)
Average S^2	0.83	ND	ND
Greek key 2			
Average R1(s)	2.68 (± 0.08)	1.81 (± 0.08)	1.91 (± 0.11)
Average R2(s)	6.69 (± 0.41)	3.36 (± 0.19)	4.33 (± 0.21)
Average NOEs	0.73 (± 0.03)	0.42 (± 0.06)	0.43 (± 0.05)
Average S^2	0.81	ND	ND
Average R_2 (and nOe) of β-strands and loop			
Greek key1 vs Greek key 2			
$\beta 1$ (6–9) vs. $\beta 5$ (44–50)	9.19(0.55)/8.45(0.61)	2.75(0.26)/4.64(0.63)	4.12(0.42)/5.19 (0.54)
$\beta 2$ (18–21) vs. $\beta 6$ (59–61)	7.28(0.58)/8.99(0.64)	3.00(0.28)/3.25 (0.60)	4.30 (0.45)/4.72 (0.54)
$\beta 4$ (38–41) vs. $\beta 8$ (80–85)	9.53(0.55)/7.56(0.59)	3.77(0.40)/2.35 (0.08)	4.73 (0.59)/3.51 (0.30)
$\beta 7$ (64–66) vs. $\beta 3^*$ (27–30)	7.09(0.66)/8.64(0.56)	3.24(0.64)/2.53 (0.38)	4.82 (0.38)/3.75 (0.29)
loop a (10–16) vs loop b (51–56)	7.34(0.88)/7.87(0.91)	3.19(0.34)/3.86 (0.49)	4.42 (0.43)/4.93 (0.55)
Calcium binding region1 (31–36)	10.89 (0.84)	3.00 (0.51)	4.61 (0.41)
Calcium binding region2 (76–80)	10.88 (0.88)	3.00 (0.34)	4.11 (0.44)

*-not a β -sheet in the NMR structure. ND- not determined. Calcium binding region 1 is annotated as loop-1 and calcium binding region 2 is annotated as loop-2.
doi:10.1371/journal.pone.0042948.t001

values showed finite variations all along the protein primary sequence indicating some degree of restricted motions even under denaturing conditions. Figure 6B shows the changes in the R_2 values for the protein in going from 0 to 4 M, 0 to 6 M and 4 to 6 M of GdmCl concentrations. The negative values in Figure 6B

represent increased conformational transitions, while positive deviations indicate decreased conformational transitions. These features were predominantly seen for residues Asp 24 and Glu 50, while going from 4 to 6 M GdmCl concentration. Here, Asp 24 shows decreased conformational transition whereas Glu 50

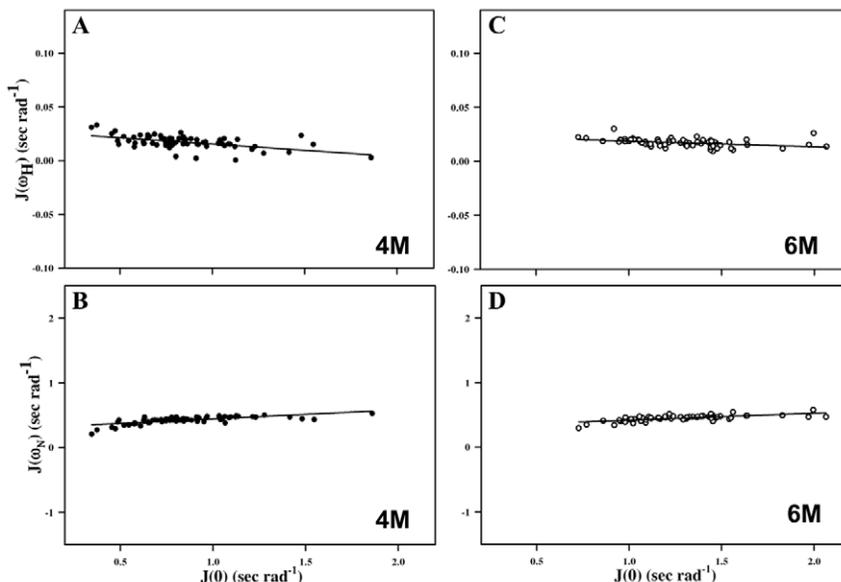


Figure 5. Correlation between spectral density functions. Plot of $J(\omega_H)$ versus $J(0)$ (A, C) and $J(\omega_N)$ versus $J(0)$ (B, D). The fits were obtained by linear regression, $J(\omega_{N,H}) = \alpha J(0) + \beta$. Here α is the slope and β is the intercept on ω_H/ω_N axis. Filled circled (A, B) plots denote M-crystallin taken in 4 M GdmCl concentration, whereas open circled (C, D) plots correspond to the protein taken in 6 M GdmCl concentration.
doi:10.1371/journal.pone.0042948.g005

Table 2. All tumbling times (in ns) at various GdmCl concentrations are given in the table.

GdmCl conc.	xN1	xN2	xN3	xH1	xH2	xH3
0 M	4.67	1.39	IM**	79.30*	1.68	0.042***
4 M	3.54	1.09	IM	1.28	4.42	0.071
6 M	3.97	1.01	IM	1.27	10.06	0.062

*sub- μ s motion,

**-imaginary value,

***-ps motion.

These measurements were done at 25°C. The equation

$$2\alpha\omega_{(H)}^2\tau_c^2 + 5\beta\omega_{(H)}^2\tau_c^2 + 2(\alpha - 1)\tau_c + 5\beta = 0$$
 was used for the calculation. xHs/xNs represent the respective roots for proton (using ω_H) and nitrogen (using ω_N).

doi:10.1371/journal.pone.0042948.t002

exhibited increased conformational transition (Figure 6B). Further, as is evident from Figure 6 A/B, Val 8, Thr 22, Ala 8 and Ser 78 were observed to have increased conformational exchange in its native state (0 M GdmCl) as compared to the protein with 6 M GdmCl. This could be attributed to the fact that, as the protein attains its native fold, the mechanism for conformational exchange is driven by several factors such as change of viscosity, presence of stable secondary structural elements etc [43,44].

Conformational exchange contributions (R_{ex}) to the R_2 were calculated based on equation $R_2 = R_1 + R_{ex}$ [45,46]. Higher values of R_2/R_1 ratio represent the presence of conformational exchange and shed light on motional fluctuations (Figure 7A/B). Conformational exchange is expected to increase with higher-field strength (not shown in this study). Significantly large conformational exchange contributions were observed for residues Tyr 9, Ala 21, Gln 25, Ala 28, Asn 31, Leu 32, Lys 40, Thr 45, Leu 61, Ser 71 and Ile 78 for the protein with 0 M GdmCl concentration; residues His 11, Glu 50, and Ala 72 at 4 M GdmCl, and residues

His 11, Ser 23, Tyr 54 and Ser 78 at 6 M GdmCl. These residues are present either at the edge of the secondary structure in M-crystallin or in the loop region of the protein.

Residual structural preferences in the unfolded states

The presence of hydrophobic clustering has been quite often reported to be the prime reason for the formation of residual structural element(s) in unfolded state of proteins [38,47,48,49]. The hydrophobic clustering in M-crystallin (if any) was quantified using the per-residue average area buried upon folding AABUF [50], which is defined as the “change in average area of each amino acid residue upon unfolding”. AABUF is well correlated in several unfolding studies [51,52,53] and is used as a measure for the identification of folding sites in proteins. The correlation of $J(0)$ values and AABUF with the protein primary sequence is shown in Figure 8A. As is evident from this figure, the AABUF values were relatively higher in the polypeptide stretches Val 6-Phe 14, Lys 42-Gly 62, and Ser 78-Phe 82. Interestingly, $J(0)$ and the AABUF showed a similar trend in these regions only under denaturing conditions of 4 and 6 M GdmCl. These regions with reduced flexibilities identified from $J(0)$ analysis were found to correlate well with increased AABUF values (increased hydrophobic content). This correlation suggests the presence of local hydrophobic clustering in these regions under denaturing conditions. A further analysis of HSQC peaks with lower intensities or absence of HSQC peaks in these regions led to our conclusion of occurrence of possible transient hydrophobic clustering under denaturing conditions mentioned above (Figure 8B). It is also interesting to note that region Asn 77-Arg 83 shows variation in $J(\omega_H)$ and $J(\omega_N)$ values in its native state (0 M GdmCl). Incidentally, the three polypeptide stretches Val 6-Phe14, Lys 42-Gly 62, and Ser 78-Phe 82 discussed above comprise four β -strands, namely β_1 , β_4 , β_5 and β_7 , in the native state protein structure (Figure 8A). Thus, it is obvious that local interactions in these regions resulted in restricted

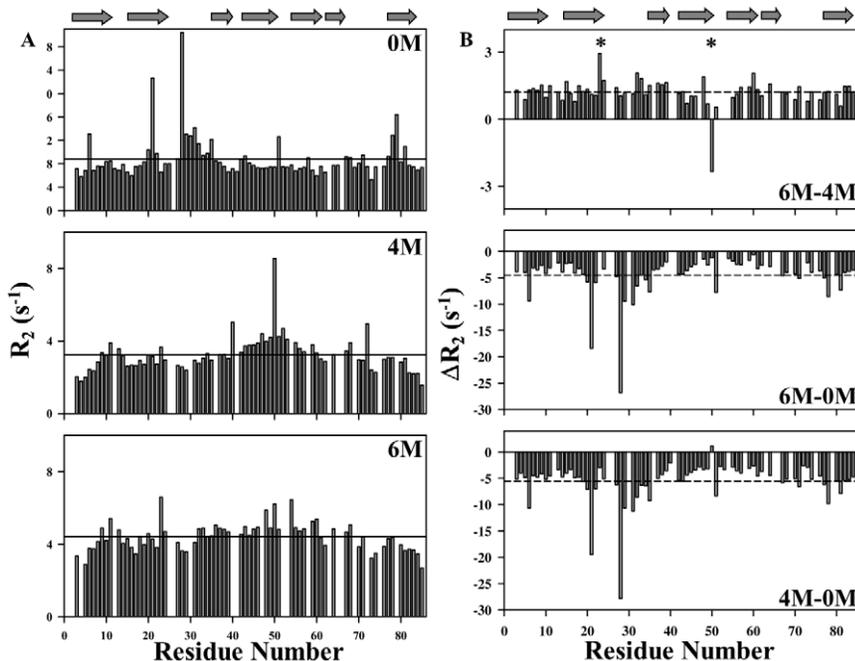


Figure 6. Transverse relaxation rates (R_2) in M-crystallin. Left panel (A) represents measured R_2 values at 6, 4 and 0 M GdmCl concentrations against protein sequence. Right panel (B) shows the deviations in R_2 values for 6, 4 and 0 M GdmCl. Horizontal lines in each box indicate average values. Residues Asp 24 and Glu 50 are marked with asterisks.

doi:10.1371/journal.pone.0042948.g006

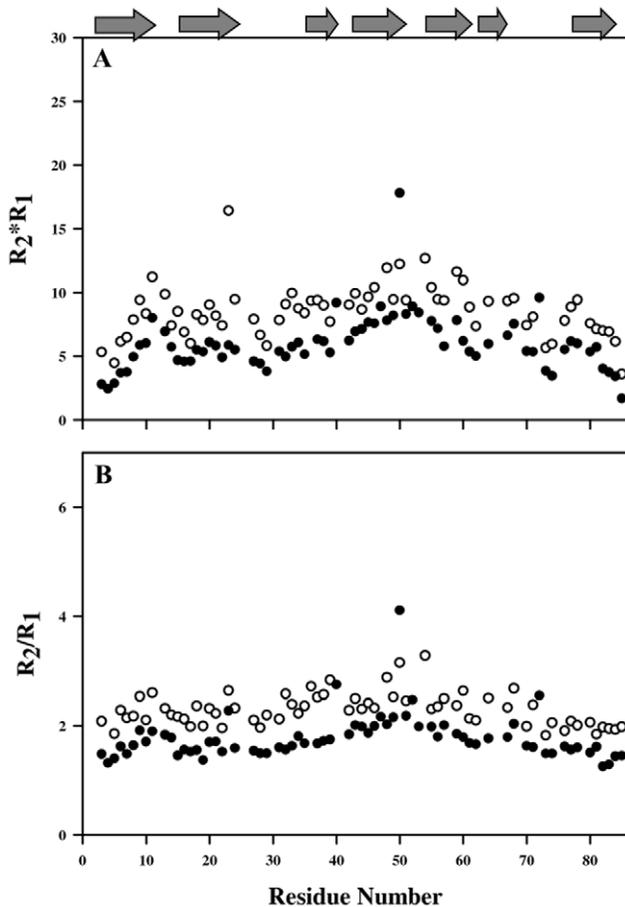


Figure 7. Conformational exchange in M-crystallin. Plots depicting the conformational exchange for the protein taken in 4 M (filled circles) and 6 M (open circles) GdmCl. (A) $R_2 \times R_1$ and (B) R_2/R_1 . doi:10.1371/journal.pone.0042948.g007

time scale motions, which can result in the formation of residual structures, especially in the regions Thr 22-Ala 28, Asp 60-Tyr 66 and Ala 72-Asn 77. This is in line with the α -helical propensities of these residues under 6 M GdmCl concentration.

Furthermore, we recorded ^{15}N - ^1H -HSQC-NOESY spectra with different mixing times (τ_m 100 and 140 ms) to look for medium- and long-range nOes (if any) in the NOESY spectra of the protein under different denaturant conditions. Unlike the native state protein, which showed several medium- and long-range nOe connectivities, the protein under different GdmCl concentrations did not show any long-range nOes, except few $d_{\text{NN}}(i, i+1)$ connectivities, which are commonly expected in α -helical segments, along the polypeptide stretches Asp 20-Gln 25, and Ser 68-Gly 73, of the protein in 6 M GdmCl (Figure S6). These ^1H - ^1H nOe connectivities throw light on the structural preference of these corresponding polypeptide stretches. These observations, taken together, support the propensity of both α -helical and β -strand structural elements of the M-crystallin under denaturing conditions.

Sequence-structure-function paradigm and folding mechanism

The primary sequence, structure and dynamics reflect the design of function of any protein. Structural Biology is guided by this motivation. Solution structures are ideally suited for unrav-

eling such functional states. It is widely accepted that residual structures under denaturing conditions are indicative of structural nucleation for proper folding of the protein [41]. For M-crystallin under different denaturing conditions, many segments followed the β -strand structural preferences, with few polypeptide stretches showing α -helical propensities. Based on the temperature coefficients, dynamics study and nOe information, we propose a model for the folding of M-crystallin. The protein starts from the natively folded state when the M-crystallin is in 0 M GdmCl [25]. Subsequently, protein undergoes transition and starts showing two distinct sets of peaks when it is in the 2 M GdmCl (Figure S2). Thereafter, when the protein is in 4 M GdmCl, it starts showing secondary structural elements and finally adopts an unfolded state under 6 M GdmCl.

Unfolding Study of M-crystallin and its biological relevance

The unfolding study of M-crystallin under different GdmCl concentrations provided residue level insights into the intrinsic conformational preferences of different polypeptide stretches in the protein. The backbone dynamics, NOESY data and biophysical data of the protein under different GdmCl concentrations taken together suggest that the protein adopts few α - and β -type structural preferences under different denaturing conditions. In conclusion, these fluctuating structural preferences and the possible intermolecular interaction/association, as predicted based on the spectral density data, might be a leading cause for aggregation of crystallins resulting in cataract.

Materials and Methods

Protein sample preparation

^{13}C - or/and ^{15}N -labelled M-crystallin were over-expressed and purified as mentioned elsewhere [16][54]. The purity of samples was checked using SDS-PAGE and mass spectroscopy (MALDI-TOF). The protein samples denatured in the presence of varying concentrations of *guanidine-HCl* (GdmCl; in 0–6 M range at an interval of 0.2–0.3 M) were prepared in milliQ H_2O solution (10 mM Tris, 50 mM KCl, 5–10 mM CaCl_2 , pH 5.5) for various purposes as discussed above.

For NMR measurements, the protein samples were concentrated to 1.2 mM and exchanged with the proper buffer containing GdmCl at concentrations 0, 1, 2, 4, and 6 M. NMR experiments were recorded after equilibrating the samples in GdmCl for 6–8 hrs. The data acquired with the protein samples in 4 and 6 M GdmCl concentrations is discussed in the results and discussion.

Denaturation studies in guanidine-HCl

The denaturation profile of M-crystallin was studied by optical (circular dichroism (CD) and fluorescence) and NMR spectroscopy. The GdmCl concentrations were determined using a refractometer. For optical measurements, the protein samples of 45 μM concentration with different concentrations of GdmCl were prepared and equilibrated for at least a period of 5–8 hrs.

Far-UV CD spectra were recorded at 25°C on a JASCO J-810 spectropolarimeter (JASCO, Europe) and corrected for buffer baseline. Each scan ranged from 200 to 260 nm with a scan speed of 20 nm/min with a quartz cell having a path length of 0.1 cm. The CD spectra below 200 nm were saturated due to high salt (particularly in the presence of 6 M GdmCl) and hence were not included in the data analysis. Each spectrum was an average of 4 scans (Figure 2A).

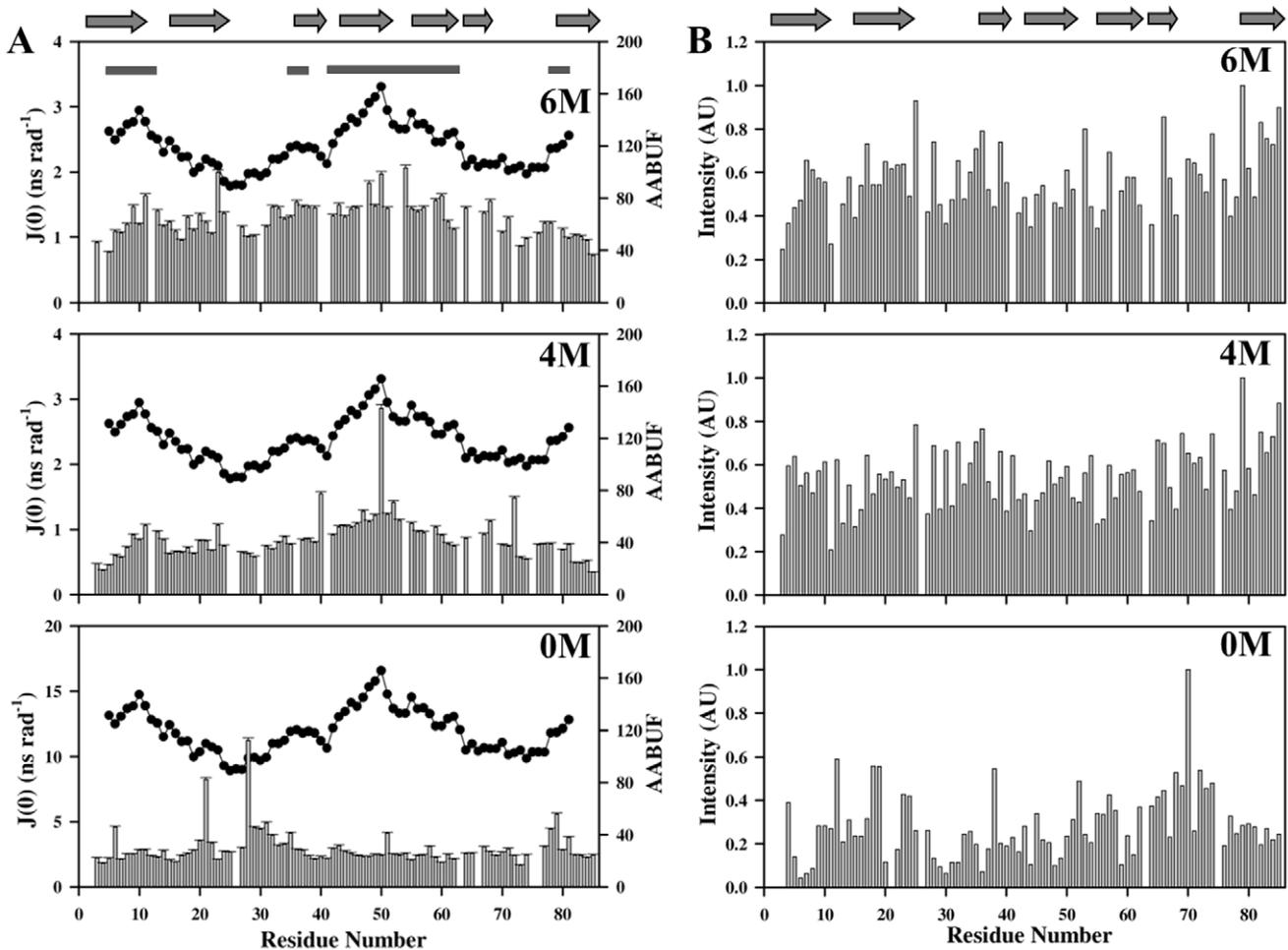


Figure 8. Spectral density, AABUF and normalized cross-peak intensity analysis. (A) Plots of $J(0)$ values (vertical bars with error) and AABUF (Average Area Buried upon Folding) (filled black circles). Horizontal bars on top of the plot indicate the regions with AABUF value more than the average. (B) Normalized cross-peak intensities derived from 2D $[^{15}\text{N}\text{-}^1\text{H}]$ -HSQC of M-crystallin at different denaturant concentrations. The folded secondary structural elements (0 M GdmCl) are shown on the top of panel. doi:10.1371/journal.pone.0042948.g008

Fluorescence spectroscopy experiments were performed on a Hitachi F-4500 spectrofluorimeter with a protein concentrations of 45 μM , as mentioned above (Figure 2B). The Trp emission spectra were recorded by exciting the protein samples at 295 nm. The intensity maxima, wavelength maxima, and intensity value at 331 nm were plotted as a function of GdmCl concentration to determine the fractions unfolded. The data were fitted to a two-state denaturation model, and parameters were determined by a non-linear curve fitting to the following equation [55]:

$$Y = \frac{\{s_n + s_d \exp(-(g_1 - m_1 D)/RT)\}}{\{1 + \exp(-(g_1 - m_1 D)/RT)\}} \quad (1)$$

where, Y is the observed spectroscopic signal; s_n and s_d represent the spectroscopic signals of folded and denatured proteins, respectively; g_1 and m_1 represent the free-energy change and slope of the transition, respectively; D is the denaturant concentration; T is the temperature (in K), and R is the universal gas constant (1.987 cal $\text{K}^{-1}\text{mol}^{-1}$).

NMR spectroscopy

NMR experiments were performed either on a Bruker Avance spectrometer with a ^1H frequency of 800.12 MHz, or on a Varian Inova spectrometer with ^1H frequency of 600.51 MHz at 25°C. Both the spectrometers were equipped with cryogenically cooled probes. Sensitivity-enhanced 2D $[^{15}\text{N}\text{-}^1\text{H}]$ -HSQC was recorded for each NMR sample, with 256 and 1024 complex data points along the ^{15}N and ^1H dimensions, respectively. The spectral widths along ^1H and ^{15}N dimensions were 12 and 30 ppm, respectively. No changes in HSQC spectra were noticed during the data acquisition, indicating a stable equilibrium under all the experimental conditions. The data were apodized using a sine-squared bell window functions along both the dimensions, zero filled to yield a final resolution of 0.9 and 2.5 Hz/pt along ω_1 and ω_2 dimensions, respectively, and followed by Fourier transformation. A suite of 3D experiments (CBCACONH, CBCANH, HNCO) and ^{15}N -edited TOCSY-HSQC with a mixing time of 80 ms were recorded for complete resonance assignments [27,28]. ^{15}N - and ^{13}C -edited NOESY-HSQC, with a mixing time of 100 and 140 ms, were recorded to get information about $^1\text{H}\text{-}^1\text{H}$ nOes. Number of scans were 8–24 scans for different experiments, with 32–64 and 48–80 complex increments along ^{15}N and ^{13}C

dimensions, respectively. 100–128 complex increments were used along the indirect ^1H dimension.

Relaxation data were recorded on both 600 and 800 MHz spectrometers at 25°C [56,57]. ^{15}N -spin lattice relaxation rates ($R_1 = 1/T_1$) were measured with inversion recovery delays of 10, 30*, 50, 70, 90*, 130, 240, 410*, 610, 860, and 1100 ms, where the delays marked with an asterisk were recorded twice. ^{15}N -spin-spin relaxation rates ($R_2 = 1/T_2$) were measured with Carr-Purcell-Meiboom-Gill (CPMG) delays of 10*, 30, 50*, 70, 90, 110*, 130, 150, 170, and 190 ms. Steady state [^{15}N - ^1H]-heteronuclear nOe measurements were carried out with the proton saturation time and relaxation delay of 3 s each. A relaxation delay of 6 s was used in the experiment without proton saturation. 32 scans were used during R_1 and R_2 measurements, while 48 scans were used for the heteronuclear nOe measurement.

Amide proton temperature coefficients were measured by recording a suite of sensitivity-enhanced 2D [^{15}N - ^1H]-HSQC at different temperatures ranging from 15 to 36°C, at an interval of 3°C, with 256 t_1 increments along the ^{15}N dimension. Number of scans was 16, and the spectral widths were 12 and 30 ppm along the ^1H and ^{15}N dimensions, respectively. The protein was stable over the entire temperature and pH (5.5 and 7.5) ranges. $^3J(\text{H}^{\text{N}}-\text{H}^{\alpha})$ coupling constants were measured from (3, 2) HNHA [31]. All the experiments were processed using Felix (Accelrys Software Inc., San Diego, CA), NMRPipe [58] and analyzed using Felix (<http://www.felixnmr.com>) and CARA (<http://cara.nmr.ch>). The (3, 2)D HNHA, (3, 2)D HNHB, (3, 2)D CB(CA-CO)NHN, (3, 2)D CT-HCCH-COSY spectral data were processed and used as mentioned elsewhere [31]. Sensitivity-enhanced 2D [^{15}N - ^1H]-HSQC were recorded at the beginning and end of all the 3D experiments and compared to make sure the stability of the protein. No observable changes were seen in the HSQC indicating high stability of the protein during the entire collection of the NMR Data. All the data were apodized with a sine-bell window function, shifted by 56° along both the dimensions of 2D data, and by 60° along all the three dimensions of 3D data, followed by zero-filling and Fourier transformation. The final processed data matrices had 2048*1024 and 1024*128*256 complex data points in all 2D and 3D spectra, respectively.

In the temperature coefficient measurement study, the chemical shift information derived from individual HSQC was tabulated and fitted to a straight line and the corresponding temperature coefficients ($d\delta/dT$) were determined from its slope. The digital resolutions in these HSQC spectra were 2.0 and 0.5 Hz/pt along ω_1 and ω_2 dimensions, respectively.

In relaxation data measurements, peak heights with their respective errors were measured. The peak-heights thus derived were fitted to a single exponential decay function,

$$I(t) = A + Be^{-R_1 t}$$

to derive the individual R_1 and R_2 values, where $I(t)$ is the intensity at delay t (ms) used in the measurement of R_1 and R_2 . $A+B$ is the intensity at initial time $t=0$, and A is the steady-state intensity at $t=\infty$.

The ^1H - ^{15}N heteronuclear nOe was calculated from the following equation,

$$\text{nOe} = I_{\text{sat}}/I_{\text{eq}}$$

where, I_{sat} and I_{eq} are the intensities of individual peaks in the spectra recorded with and without proton saturation. The errors in

the nOes were obtained using the root-mean-square value of the background noise as described by Farrow *et al.* [59,60].

Proton chemical shifts were referenced using DSS at 0.00 ppm in $\text{H}_2\text{O}/^2\text{H}_2\text{O}$ solution at 25°C (6 M GdmCl, pH 5), whereas ^{15}N and ^{13}C were referenced indirectly as described elsewhere [61].

Spectral density functions and correlation times

The spectral density functions $\mathcal{J}(0)$, $\mathcal{J}(\omega_N)$, and $\mathcal{J}(\omega_H)$ were calculated as described by Wagner *et al.* [59,60] and Lefevre *et al.* [62]. According to these approaches, $\mathcal{J}(0)$, $\mathcal{J}(\omega_N)$, and $\mathcal{J}(\omega_H)$ terms can be expressed in terms of ^{15}N spin-lattice (R_1) and spin-spin (R_2) relaxation rates and heteronuclear [^1H - ^{15}N] NOEs as follows:

$$J(0) = \frac{3}{2(3d^2 + c^2)} \left[-\frac{1}{2}R_1 + R_2 - \frac{3}{5}R_{\text{NOE}} \right] \quad (2)$$

$$J(\omega_N) = \frac{1}{3d^2 + c^2} \left[R_1 - \frac{7}{5}R_{\text{NOE}} \right] \quad (3)$$

$$J(\omega_H) = \frac{1}{5d^2} R_{\text{NOE}} \quad (4)$$

where,

$$R_{\text{NOE}} = [(\{^1\text{H}\} - ^{15}\text{N})\text{NOE} - 1]R_1 \frac{\gamma_N}{\gamma_H} \quad (5)$$

The constant d^2 is approximately equal to 1.35×10^9 (rad/s) 2 , whereas the constant c^2 is approximately 1.25×10^9 (rad/s) 2 and 2.25×10^9 (rad/s) 2 at 600 and 800 MHz, respectively. Errors in the individual spectral density functions were calculated from the error in the related parameters and by solving above equations [63]. After calculating $\mathcal{J}(0)$, $\mathcal{J}(\omega_N)$, and $\mathcal{J}(\omega_H)$, the linear correlation between $\mathcal{J}(0)$ and $\mathcal{J}(\omega_N)$, and $\mathcal{J}(0)$ and $\mathcal{J}(\omega_H)$ was examined (Figure 4).

Only few residues showed distinctly higher values of $\mathcal{J}(0)$, indicating that chemical exchange motions have significant contributions to the mobility of their NH vectors. A linear correlation between $\mathcal{J}(\omega_{\text{N,H}})$ and corresponding $\mathcal{J}(0)$ values using the equation $J(\omega_{\text{N,H}}) = \alpha J(0) + \beta$, has been proposed for the calculation of τ_m [60,62]. The values of α and β derived from the plot of $\mathcal{J}(\omega_N)$ versus $\mathcal{J}(0)$ were found to be 0.14 and 0.30 ns/rad, respectively, for 4 M GdmCl case, whereas 0.11 and 0.31 ns/rad, respectively, for 6 M case. These values were then used to calculate the overall correlation time (τ_m) using the following equation [60,62]:

$$2\alpha\omega_{(H,N)}^2\tau_m^3 + 5\beta\omega_{(H,N)}^2\tau_m^2 + 2(\alpha - 1)\tau_m + 5\beta = 0 \quad (6)$$

The solution of the above cubic equation yields three values of τ_m (Table 2). The values in milli-seconds to micro-seconds, sub-nano-seconds and pico-seconds represent the chemical exchange, overall rotational correlation time and internal motions, respectively, depending on their amplitudes.

Average area buried upon folding

The per-residue average area buried upon folding (AABUF) was calculated using the method of Rose *et al.* [50]. A nine-residue moving-average window was used in the AABUF calculations.

Supporting Information

Figure S1 Sequence alignment of M-crystallin with other known lens crystallins, β B2-crystallin, γ B-crystallin, ciona-crystallin and DdCAD, Trp corner and Tyr corner, special features of crystallins are shown with black arrow. Higher sequence similarity in the alignment chart is highlighted with red color.

(TIF)

Figure S2 Sensitivity-enhanced 2D [^{15}N - ^1H]-HSQC of [Ca^{2+}] $_2$ -M-crystallin in the presence of 0 (A), 2 (B), 4 (C) and 6 M (D) GdmCl (pH 5.5 and temperature 298 K). These spectra were recorded on a Bruker Avance 800 MHz spectrometer with 128 and 1024 points along t_1 and t_2 dimensions, respectively. Individual peak assignments are shown by the corresponding single-letter code of the amino acid residue and its sequence number along the primary sequence.

(TIF)

Figure S3 (3, 2) D HNHA spectra of M-crystallin taken in (A) 4 and (B) 6 M GdmCl.

(TIF)

Figure S4 $^3\text{J}(\text{H}^{\text{N}}-\text{H}\alpha)$ versus residue number along the primary sequence, as measured from (3, 2) D HNHA for different GdmCl concentrations. In (3, 2) D HNHA, couplings were measured using the formula mentioned in [31].

(TIF)

References

- Rezaei H, Choiset Y, Eghiaian F, Treguer E, Mentre P, et al. (2002) Amyloidogenic unfolding intermediates differentiate sheep prion protein variants. *J Mol Biol* 322: 799–814.
- Scheibel T, Buchner J (2006) Protein aggregation as a cause for disease. *Handb Exp Pharmacol*: 199–219.
- Scriven P, Brown NJ, Pockley AG, Wyld L (2007) The unfolded protein response and cancer: a brighter future unfolding? *J Mol Med* 85: 331–341.
- Kudo T, Katayama T, Imaizumi K, Yasuda Y, Yatera M, et al. (2002) The unfolded protein response is involved in the pathology of Alzheimer's disease. *Ann N Y Acad Sci* 977: 349–355.
- Forman MS, Lee VM, Trojanowski JQ (2003) 'Unfolding' pathways in neurodegenerative disease. *Trends Neurosci* 26: 407–410.
- Slavotteck AM, Biesecker LG (2001) Unfolding the role of chaperones and chaperonins in human disease. *Trends Genet* 17: 528–535.
- Dyson HJ, Wright PE (1998) Equilibrium NMR studies of unfolded and partially folded proteins. *Nat Struct Biol* 5 Suppl: 499–503.
- Dyson HJ, Wright PE (2002) Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 12: 54–60.
- Solomon B (2002) Anti-aggregating antibodies, a new approach towards treatment of conformational diseases. *Curr Med Chem* 9: 1737–1749.
- Wright PE, Dyson HJ (2009) Linking folding and binding. *Curr Opin Struct Biol* 19: 31–38.
- Silva JL, Cordeiro Y, Foguel D (2006) Protein folding and aggregation: two sides of the same coin in the condensation of proteins revealed by pressure studies. *Biochim Biophys Acta* 1764: 443–451.
- Santucci R, Sinibaldi F, Fiorucci L (2008) Protein folding, unfolding and misfolding: role played by intermediate States. *Mini Rev Med Chem* 8: 57–62.
- Almstedt K, Lundqvist M, Carlsson J, Karlsson M, Persson B, et al. (2004) Unfolding a folding disease: folding, misfolding and aggregation of the marble brain syndrome-associated mutant H107Y of human carbonic anhydrase II. *J Mol Biol* 342: 619–633.
- Bloemendal H, de JW, Jaenicke R, Lubsen NH, Slingsby C, et al. (2004) Ageing and vision: structure, stability and function of lens crystallins. *Prog Biophys Mol Biol* 86: 407–485.
- Benedek GB (1997) Cataract as a protein condensation disease: the Proctor Lecture. *Invest Ophthalmol Vis Sci* 38: 1911–1921.
- Ma B, Sen T, Asnaghi L, Valapala M, Yang F, et al. (2011) betaA3/A1-Crystallin controls anoikis-mediated cell death in astrocytes by modulating PI3K/AKT/mTOR and ERK survival pathways through the PKD/Bit1-signaling axis. *Cell death & disease* 2: e217.
- Sinha D, Valapala M, Bhutto I, Patek B, Zhang C, et al. (2012) betaA3/A1-crystallin is required for proper astrocyte template formation and vascular remodeling in the retina. *Transgenic research* 21: 1033–1042.
- Wistow GJ, Piatigorsky J (1988) Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. *Annu Rev Biochem* 57: 479–504.
- Wistow G (1990) Evolution of a protein superfamily: relationships between vertebrate lens crystallins and microorganism dormancy proteins. *Journal of molecular evolution* 30: 140–145.
- Aravind P, Mishra A, Suman SK, Jobby MK, Sankaranarayanan R, et al. (2009) The betagamma-crystallin superfamily contains a universal motif for binding calcium. *Biochemistry* 48: 12180–12190.
- Mishra A, Suman SK, Srivastava SS, Sankaranarayanan R, Sharma Y (2012) Decoding the molecular design principles underlying Ca(2+) binding to betagamma-crystallin motifs. *J Mol Biol* 415: 75–91.
- Wenk M, Mayr EM (1998) Myxococcus xanthus spore coat protein S, a stress-induced member of the betagamma-crystallin superfamily, gains stability from binding of calcium ions. *Eur J Biochem* 255: 604–610.
- Kretschmar M, Mayr EM, Jaenicke R (1999) Kinetic and thermodynamic stabilization of the betagamma-crystallin homolog spherulin 3a from Physarum polycephalum by calcium binding. *J Mol Biol* 289: 701–705.
- Suman SK, Mishra A, Ravindra D, Yeramala L, Sharma Y (2011) Evolutionary remodeling of betagamma-crystallins for domain stability at cost of Ca2+ binding. *J Biol Chem* 286: 43891–43901.
- Barnwal RP, Jobby MK, Devi KM, Sharma Y, Chary KV (2009) Solution structure and calcium-binding properties of M-crystallin, a primordial betagamma-crystallin from archaea. *J Mol Biol* 386: 675–689.
- Barnwal RP, Devi KM, Agarwal G, Sharma Y, Chary KV (2011) Temperature-dependent oligomerization in M-crystallin: lead or lag toward cataract, an NMR perspective. *Proteins* 79: 569–580.
- Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ, Rance M (2006) Protein NMR spectroscopy: Principles and practice: Academic Press.
- Chary KV, Govil G (2008) NMR in Biological Systems: From Molecules to Humans: Springer.
- Barnwal RP, Atreya HS, Chary KV (2008) Chemical shift based editing of CH3 groups in fractionally ^{13}C -labelled proteins using GFT (3, 2)D CT-HCCH-COSY: stereospecific assignments of CH3 groups of Val and Leu residues. *J Biomol NMR* 42: 149–154.
- Barnwal RP, Rout AK, Atreya HS, Chary KV (2008) Identification of C-terminal neighbours of amino acid residues without an aliphatic ^{13}C gamma as an aid to NMR assignments in proteins. *J Biomol NMR* 41: 191–197.
- Barnwal RP, Rout AK, Chary KV, Atreya HS (2007) Rapid measurement of $^3\text{J}(\text{H N-H alpha})$ and $^3\text{J}(\text{N-H beta})$ coupling constants in polypeptides. *J Biomol NMR* 39: 259–263.
- Patel S, Chaffotte AF, Goubard F, Pauthe E (2004) Urea-induced sequential unfolding of fibronectin: a fluorescence spectroscopy and circular dichroism study. *Biochemistry* 43: 1724–1735.
- Kumar A, Srivastava S, Kumar Mishra R, Mittal R, Hosur RV (2006) Residue-level NMR view of the urea-driven equilibrium folding transition of SUMO-1 (1–97): native preferences do not increase monotonously. *J Mol Biol* 361: 180–194.

Figure S5 ^{15}N relaxation parameters and reduced spectral densities versus residue number. (A) Longitudinal relaxation times (T_1). (B) Transverse relaxation times (T_2). (C) [^1H - ^{15}N] NOE enhancements derived from $I_{\text{sat}}/I_{\text{eq}}$ ratio, where I_{sat} and I_{eq} are the intensities of peaks in the 2D spectra recorded with and without proton saturation, respectively. Error bars in the T_1 and T_2 data denote curve-fitting uncertainties; errors in the [^1H - ^{15}N] NOEs are estimated from the signal/noise ratio of the spectra. Spectral densities are shown in (D) $J(0)$, (E) $J(\omega_{\text{H}})$, and (F) $J(\omega_{\text{N}})$, as a function of protein sequence number. Here, horizontal bars represent the data for the protein with 0 M GdmCl, filled and open circles correspond to the protein taken in 4 and 6 M GdmCl concentrations, respectively. Secondary structural elements of folded protein are depicted on top of the panels.

(TIF)

Figure S6 NOE connectivities obtained from ^{15}N -edited 3D NOESY-HSQC of M-crystallin taken in 6 M GdmCl (pH = 5.5; temperature = 25 °C). The native secondary structure elements are depicted on top of the panel with arrows. Folded protein has 7 stretches of β -strands.

(TIF)

Author Contributions

Conceived and designed the experiments: RPB KVRC. Performed the experiments: RPB. Analyzed the data: RPB GA. Wrote the paper: RPB GA KVRC.

34. Kumar A, Srivastava S, Mishra RK, Mittal R, Hosur RV (2006) Local structural preferences and dynamics restrictions in the urea-denatured state of SUMO-1: NMR characterization. *Biophysical journal* 90: 2498–2509.
35. Barnwal RP, Chary KV (2008) An efficient method for secondary structure determination in polypeptides by NMR. *Current Science* 94: 5.
36. Kazmirski SL, Daggett V (1998) Simulations of the structural and dynamical properties of denatured proteins: the “molten coil” state of bovine pancreatic trypsin inhibitor. *J Mol Biol* 277: 487–506.
37. Kazmirski SL, Daggett V (1998) Non-native interactions in protein folding intermediates: molecular dynamics simulations of hen lysozyme. *J Mol Biol* 284: 793–806.
38. Kazmirski SL, Wong KB, Freund SM, Tan YJ, Fersht AR, et al. (2001) Protein folding from a highly disordered denatured state: the folding pathway of chymotrypsin inhibitor 2 at atomic resolution. *Proc Natl Acad Sci U S A* 98: 4349–4354.
39. Meinhold DW, Wright PE (2011) Measurement of protein unfolding/refolding kinetics and structural characterization of hidden intermediates by NMR relaxation dispersion. *Proc Natl Acad Sci U S A* 108: 9078–9083.
40. Bezsonova I, Evanics F, Marsh JA, Forman-Kay JD, Prosser RS (2007) Oxygen as a paramagnetic probe of clustering and solvent exposure in folded and unfolded states of an SH3 domain. *Journal of the American Chemical Society* 129: 1826–1835.
41. Lundstrom P, Mulder FA, Akke M (2005) Correlated dynamics of consecutive residues reveal transient and cooperative unfolding of secondary structure in proteins. *Proc Natl Acad Sci U S A* 102: 16984–16989.
42. Baxter NJ, Williamson MP (1997) Temperature dependence of ¹H chemical shifts in proteins. *J Biomol NMR* 9: 359–369.
43. Shtilerman M, Lorimer GH, Englander SW (1999) Chaperonin function: folding by forced unfolding. *Science* 284: 822–825.
44. Booth DR, Sunde M, Bellotti V, Robinson CV, Hutchinson WL, et al. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis. *Nature* 385: 787–793.
45. Kneller JM, Lu M, Bracken C (2002) An effective method for the discrimination of motional anisotropy and chemical exchange. *Journal of the American Chemical Society* 124: 1852–1853.
46. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28: 8972–8979.
47. Eliezer D, Chung J, Dyson HJ, Wright PE (2000) Native and non-native secondary structure and dynamics in the pH 4 intermediate of apomyoglobin. *Biochemistry* 39: 2894–2901.
48. Eliezer D, Yao J, Dyson HJ, Wright PE (1998) Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nature structural biology* 5: 148–155.
49. Mok KH, Kuhn LT, Goez M, Day IJ, Lin JC, et al. (2007) A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature* 447: 106–109.
50. Rose GD, Geselowitz AR, Lesser GJ, Lee RH, Zehfus MH (1985) Hydrophobicity of amino acid residues in globular proteins. *Science* 229: 834–838.
51. Gerum C, Silvers R, Wirmer-Bartoschek J, Schwalbe H (2009) Unfolded-state structure and dynamics influence the fibril formation of human prion protein. *Angewandte Chemie* 48: 9452–9456.
52. Felitsky DJ, Lietzow MA, Dyson HJ, Wright PE (2008) Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. *Proc Natl Acad Sci U S A* 105: 6278–6283.
53. Gong H, Rose GD (2008) Assessing the solvent-dependent surface area of unfolded proteins using an ensemble model. *Proc Natl Acad Sci U S A* 105: 3321–3326.
54. Barnwal RP, Agarwal G, Sharma Y, Chary KVR (2009) Complete backbone assignment of a Ca²⁺-binding protein of the βγ-crystallin superfamily from *Methanosarcina acetivorans*, at two denaturant concentrations *Biomolecular NMR Assignments* 3: 107–110.
55. Hung HC, Chen YH, Liu GY, Lee HJ, Chang GG (2003) Equilibrium protein folding-unfolding process involving multiple intermediates. *Bulletin of mathematical biology* 65: 553–570.
56. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28: 8972–8979.
57. Kordel J, Skelton NJ, Akke M, Palmer AG III, Chazin WJ (1992) Backbone dynamics of calcium-loaded calbindin D9k studied by two-dimensional proton-detected ¹⁵N NMR spectroscopy. *Biochemistry* 31: 4856–4866.
58. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, et al. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6: 277–293.
59. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* 33: 5984–6003.
60. Farrow NA, Zhang O, Szabo A, Torchia DA, Kay LE (1995) Spectral density function mapping using ¹⁵N relaxation data exclusively. *J Biomol NMR* 6: 153–162.
61. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, et al. (1995) ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J Biomol NMR* 6: 135–140.
62. Lefevre JF, Dayie KT, Peng JW, Wagner G (1996) Internal mobility in the partially folded DNA binding and dimerization domains of GAL4: NMR analysis of the N-H spectral density functions. *Biochemistry* 35: 2674–2686.
63. Bevington PR, Robinson DK (2003) Data reduction and error analysis for the physical sciences. Boston: McGraw-Hill. xi, 320 p. p.