Biochemistry

© Copyright 1998 by the American Chemical Society Volume 37, Number 37 September 15, 1998

Accelerated Publications

Identification of Protein Folding Patterns Using Site-Directed Spin Labeling. Structural Characterization of a β-Sheet and Putative Substrate Binding Regions in the Conserved Domain of αA-Crystallin[†]

Hanane A. Koteiche, Anderee R. Berengian, and Hassane S. Mchaourab*

National Biomedical ESR Center, Biophysics Research Institute, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, Wisconsin 53226

Received June 15, 1998; Revised Manuscript Received July 15, 1998

ABSTRACT: The folding pattern of the segment of αA-crystallin encoded by exon 2 and containing putative substrate binding sites was explored using site-directed spin labeling (SDSL). For this purpose, a nitroxide scan was carried out between residues 60 and 108. At each site, structural constraints describing the local environment and topography were obtained from analysis of the nitroxide mobility and its solvent accessibility. Periodic patterns in the sequence-specific variation of these parameters were used to assign the secondary structure along the sequence. Geometric constraints describing the packing of secondary structure were deduced from patterns of proximities in 20 nitroxide pairs, specifically designed to differentiate between supersecondary structural motifs. Our data, in conjunction with those of Berengian et al. [Berengian, A. R., Bova, M. P., and Mchaourab, H. S. (1997) *Biochemistry 36*, 9951−9957], reveal that the fold of the segment between residues 84 and 120 consists of an antiparallel β-sheet of three strands arranged in consecutive β-hairpins. The boundaries of the sheet are defined at one end by a surface of isologous association and on the other end by an unstructured, charged interdomain segment. One of the putative substrate binding segments overlaps a buried loop, suggesting that the structural origin of the thermal activation of binding is the transient exposure of this site. This paper describes and implements a general strategy for experimental fold recognition using SDSL. The results of its application to αA-crystallin provide the first experimental insight into the folding pattern of the subunit and establish the structural context necessary to understand molecular recognition and substrate binding.

The α-crystallin domain is a protein module used in the construction of the small heat-shock protein family (sHSP)[1] including lens α-crystallins (*1*). It consists of a stretch of 80−100 amino acids located in the C-terminal region of the sequence and flanked by an N-terminal domain characterized by extensive sequence and length variability. Among heat-shock proteins, the sHSP family is the most divergent in sequence, molecular mass, and abundance within species. In fact, the overall sequence similarity is marginal when distant sHSP are compared (*2*). Nevertheless, the presence of conserved residues in the C-terminal domain allows these proteins to be grouped together. Members of the sHSP family from diverse organisms also share common structural

---

[†] This work was supported by Grants EY12018, GM22923, and RR01008 from the National Institutes of Health.

* To whom correspondence should be addressed.

[1] Abbreviations: CD, circular dichroism; EPR, electron paramagnetic resonance; NiEDDA, nickel(II) ethylenediaminediacetate; SDSL, site-directed spin labeling; sHSP, small heat-shock proteins; DTT, dithiothreitol; TCEP, tris(2-carboxyethyl)phosphine; WT, wild-type; WT*, cysteine-less wild-type; EDTA, ethylenediaminetetraacetic acid.

S0006-2960(98)01407-X CCC: $15.00

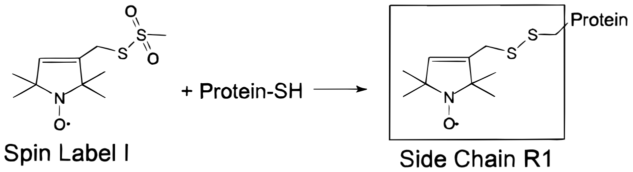
and functional features. They form oligomeric structures of 9−40 subunits and can suppress the aggregation of denaturing proteins (*2*). In the vertebrate lens, αA- and αB-crystallins form hetero-oligomers of variable size and quaternary structure that bind and sequester damaged proteins, preventing the formation of particulates that scatter light (*3*).

Dissecting the molecular mechanism mediating the chaperone-like function of α-crystallins and sHSP has proven challenging. One of the major obstacles is the lack of structural information concerning the oligomeric and subunit structures of the protein scaffold. It is likely that the common properties of these proteins are associated with a conserved structural framework encoded in the α-crystallin domain. In mammalian sHSP, this domain forms an independent folding unit in a two-domain structure (*4*). Recombinant C-terminal domains of αA- and αB-crystallins and mouse hsp 25 are folded in isolation and form multimeric structures (*5*). In the case of αA-crystallin, the multimers consist of a mixture of dimers and tetramers. On the basis of these observations, Wistow proposed that the α-crystallin domain in sHSP mediates the formation of a basic tetrameric unit (*6*). Multiple tetramers then assemble into the native oligomeric structure via interactions in the N-terminal domain. In support of Wistow's model, we have demonstrated that conserved sequences in the α-crystallin domain of both αA-crystallin and hsp 27 are involved in subunit interactions (*7*, *8*).

Numerous studies have been undertaken to determine the functional elements involved in the chaperone-like function of sHSP. Sharma et al. and Lee et al. have identified a stretch of residues, corresponding to the sequence between residues 71 and 88 of αA-crystallin, as potential substrate binding sites (*9*−*11*). While sHSP bind denaturing proteins at room temperature, the efficiency of the chaperone-like function is increased following exposure to higher temperature (*9*, *12*). For αA-crystallin, thermal activation of substrate binding correlates with an increase in the rate of subunit exchange between the oligomers (*13*), suggesting an important role for the dynamics of the oligomeric structure in the exposure of otherwise inaccessible functional sequences (*14*). However, little is known concerning the structure and topography along these sequences, and along the α-crystallin domain in general, hampering further insight into the physical basis of molecular recognition and the structural and dynamic features that modulate substrate binding.

With the goal of understanding sequence−function and sequence−structure relationships in sHSP, we are using cysteine-scanning mutagenesis and SDSL (*15*) to define the common structural framework in sHSP and to explore the impact of sequence divergence on the structural and functional features. In previous work, a limited sequence-correlated structural comparison of αA-crystallin and hsp 27 was reported (*7*, *8*). This paper extends the SDSL investigation of αA-crystallin to include the sequence between residues 60 and 108. An approach for experimental fold recognition, based on SDSL, is used to determine the local folding pattern of the sequence (*16*). In this approach, the sequence-specific secondary structure, assigned from nitroxide scanning (*15*), is used to identify a set of most likely supersecondary structural motifs. The presence of these motifs, each characterized by a distinct geometric arrangement of the secondary structure, is evaluated from patterns of proximities in nitroxide pairs. The results reveal the presence of an antiparallel β-sheet along the 84−120 segment, define the structural features of the putative substrate binding sites, and suggest a structural origin for the temperature activation of αA-crystallin chaperone-like function.

Scheme 1



EXPERIMENTAL PROCEDURES

Materials. Spin label I was a generous gift from Professor Kalman Hideg, University of Pécs, Hungary. The Superose 6 column was obtained from Amersham-Pharmacia Biotech. Tris(2-carboxyethyl)phosphine (TCEP) was obtained from Molecular Probes.

Protein Expression, Purification, and Spin Labeling. Site-directed mutagenesis, protein expression, and purification were carried out as described by Berengian et al. (*7*). The purified mutants were then reacted with a 5-fold excess of spin label I for 4 h at room temperature to yield the side-chain R1 as shown in Scheme 1 (*17*). The 1:1 labeling of each cysteine in the double mutants was determined by reduction of the disulfide bond linking the nitroxide to the protein backbone using TCEP, followed by quantitation of the released label (*8*).

Circular Dichroism. CD analysis was performed on a Jasco 710 spectropolarimeter. Protein samples were prepared in 20 mM phosphate, pH 7.1, at a concentration of 0.15 mg/mL.

Chaperone Assays. Chaperone assays of the mutants were performed in a microtiter plate using a BioTek EL311 plate reader at room temperature. Aggregation of the insulin B chain was monitored by measuring the increase in absorption due to scattering at 490 nm. Insulin, dissolved in 50 mM phosphate, pH 7.0, was reduced with 20 mM DTT in the presence of varying amounts of WT* or mutant αA-crystallin. Protein concentrations were determined by UV absorption using $\epsilon_{280} = 16\,500$ M^{-1} cm^{-1}, and the relative concentration was verified using the Bradford assay.

Size Exclusion Chromatography. R1-labeled mutants were injected from 100 μL volumes on a Superose 6 column calibrated according to the manufacturer's instructions. The samples were chromatographed at a flow rate of 0.5 mL/min.

EPR Measurements. Samples for EPR spectroscopy were prepared in 20 mM MOPS, 50 mM NaCl, and 0.1 mM EDTA, pH 7.2. EPR measurements were carried out using a Varian E102 spectrometer fitted with a two-loop one-gap resonator. The microwave power was 2 mW incident, corresponding to a microwave field of 0.20 G at the sample. A 1.6 G Zeeman modulation amplitude was used. Power saturation measurements were carried out under nitrogen and in the presence of 3 mM NiEDDA, and the data were analyzed to obtain the parameter $P_{1/2}$ (*18*). The EPR

Table 1: Chaperone Activity of αA-Crystallin Cysteine Mutants

mutant	% activity	mutant	% activity	mutant	% activity
WT	100	75	120	92	100
WT*	100	76	300	93	140
60	100	77	100	94	140
61	100	78	140	95	200
62	100	79	500	96	180
63	100	80	800	97	100
64	100	81	100	98	140
65	100	82	100	99	180
66	100	83	120	100	240
67	100	84	140	101	100
68	100	85	200	102	120
69	120	86	100	103	100
70	100	87	120	104	100
71	140	88	140	105	240
72	120	89	100	106	240
73	100	90	100	107	100
74	120	91	100	108	100

accessibility parameter, Π, was calculated as previously described (*18*).

RESULTS

Effects of Cysteine Substitution on Chaperone-like Function. To explore the sequence determinants of the chaperone-like function in αA-crystallin, we evaluated the efficiency of each mutant in suppressing the aggregation of the insulin B chain at room temperature (*19*). The chaperone assays were run in the successive rows of a microtiter plate. Each plate contained two WT* control rows consisting of six different mass ratios of insulin to αA: 1:0, 1:1, 1:1.5, 1:2, 1:3, and 1:5. The chaperone activity of each mutant was measured as the relative inhibition with respect to WT* at three different mass ratios (1:1; 1:2; 1:5), and the average deviations are reported in Table 1 in increments of 20%. All mutants are at least as efficient as the WT* in suppressing the aggregation of the insulin B chain. In fact, the deviations observed for many mutants tend to be in the direction of increased chaperone efficiency. Notable examples are residues 76, 79, and 80, where the increase in efficiency varied between 3- and 8-fold. For these residues, assays were run in triplicate and included 1:0.25 and 1:0.5 insulin to αA mass ratios.

To determine the statistical significance of these deviations, we compared the activity of the WT* obtained from assays performed on different days normalized to their respective insulin control. We found that a 20% deviation in suppression can be expected, and the variations were randomly distributed around the mean. Thus, the increased efficiency obtained as a result of cysteine substitution at sites such as 79 and 80 is not the result of experimental errors.

Structural Consequences of the R1 Substitution. The structural integrity of the R1-labeled mutants was evaluated using size exclusion chromatography and far-UV circular dichroism. Size exclusion assesses the ability of the labeled mutants to assemble into the native oligomer with an average molecular mass of 640 kDa. Table 2 lists the molecular mass of each R1-labeled mutant. Most mutants form native-like oligomers with molecular masses in the range of that of the WT* and do not show any evidence of dissociation into smaller species. Significant deviations in the average molecular mass were observed at residues 60, 75, 80, 82, 100, 103, 106, 107, and 108. Except for P82 and R103, all of these residues are in a relatively buried environment (see below). Thus, in general, the increased molecular mass can be attributed to changes in the hydrodynamic radius of the subunit needed to accommodate the molar volume of R1 (*20*). However, the larger deviations in molecular mass, such as those observed at residues 75 and 80, might also reflect underlying changes in the aggregation state. Residues 83R1 and 106R1 also displayed a wider distribution of molecular masses than that of the WT* (data not shown). For instance, the width at half-height of 83R1 was about 7 min compared to 3.5 min for the WT*. Although no separate peaks could be resolved in the chromatogram, the width of the distribution for this mutant indicates a significant population of larger and smaller oligomers.

Table 2: Molecular Mass of R1-Labeled αA-Crystallin Mutants

mutant	molecular mass (MDa)	mutant	molecular mass (MDa)	mutant	molecular mass (MDa)
WT*	0.64	76	0.78	93	0.61
60	0.84	77	0.60	94	0.59
61	0.70	78	0.76	95	0.79
62	0.79	79	0.72	96	0.64
63	0.75	80	0.92	97	0.66
64	0.63	81	0.64	98	0.68
65	0.59	82	0.84	99	0.62
66	0.62	83	0.74	100	0.91
67	0.67	84	0.72	101	0.74
68	0.53	85	0.70	102	0.79
69	0.78	86	0.63	103	0.88
70	0.57	87	0.62	104	0.74
71	0.57	88	0.61	105	0.75
72	0.61	89	0.60	106	0.86
73	0.63	90	0.66	107	0.83
74	0.67	91	0.64	108	0.93
75	1.13	92	0.79		

The effects of R1 at sites 83 and 106 on the apparent molecular mass are accompanied by changes in the far-UV CD spectrum. For both mutants, a distinct increase in the 205 nm signal was observed, suggesting some degree of unfolding. The remaining mutants segregated into two categories. Both groups have CD spectra very similar to that of WT* in the 208−220 nm region, indicating an intact overall secondary structure (data not shown). However, changes in the 195−200 nm region were more pronounced in one of the categories, and the pattern was similar to that reported in refs *7* and *8* in that residues in this category tend to have a buried environment. The absence of drastic changes in the global secondary structure and oligomeric structure at most sites is consistent with the conclusion that R1 substitution results in relatively small perturbation at the level of the backbone-fold (*20*).

Secondary Structure and Solvent Accessibility in the 60−108 Sequence. To assign the sequence-specific secondary structure and solvent accessibility, the EPR spectral line shapes of the R1-labeled mutants were analyzed in terms of the side-chain mobilities and the accessibilities to NiEDDA determined from analysis of R1 relaxation properties. Figure 1 shows the EPR spectra of R1 in the 60−108 sequence. The line shape in these spectra reflects the rotational mobility of the nitroxide around the bonds tethering it to the protein backbone and has contributions from backbone fluctuations particularly in solvent-exposed loops (*20*). The term mobility is used here in a general sense to include the effects of both molecular ordering and motional frequency of these dynamic

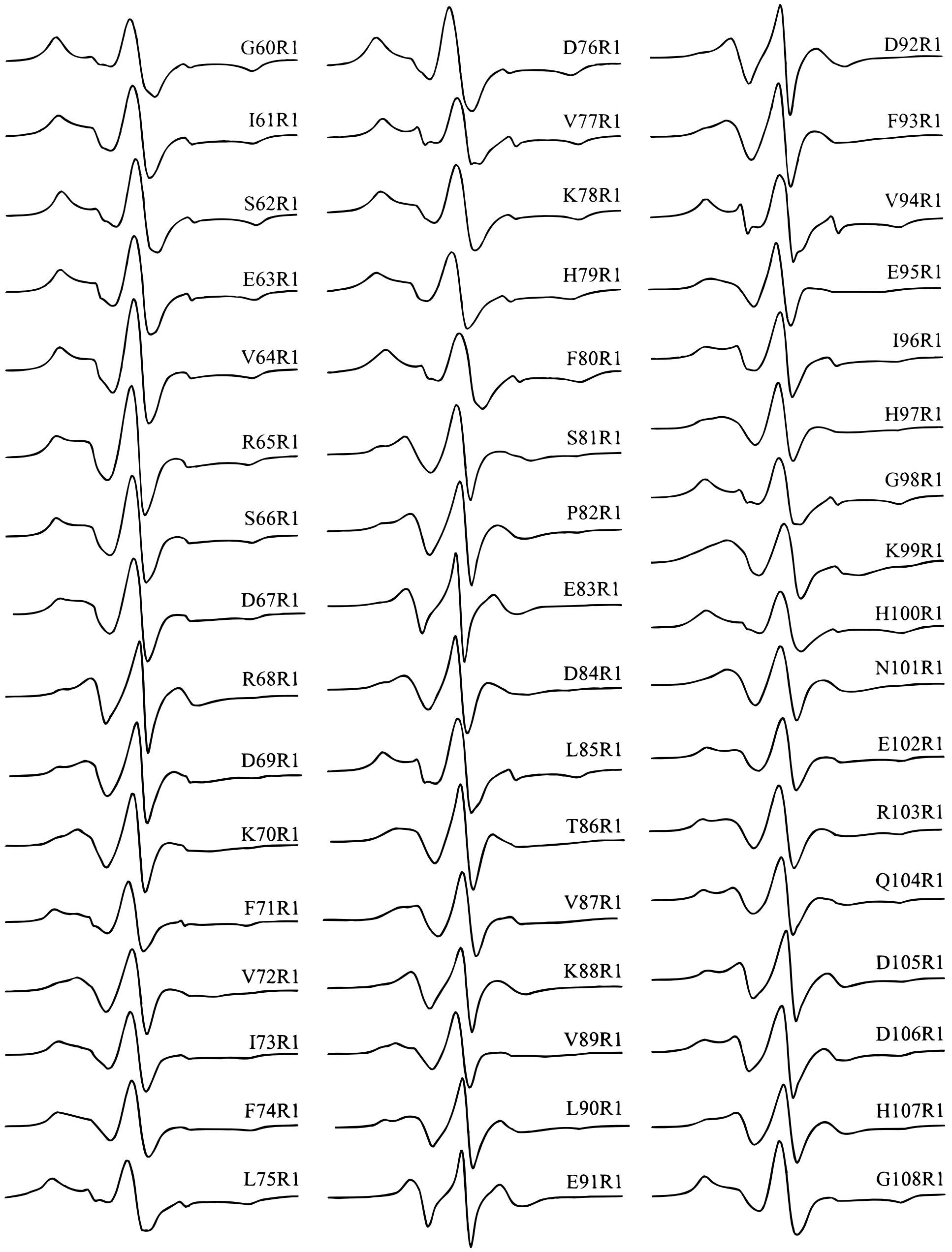


FIGURE 1: Room temperature EPR spectra of the R1-labeled αA-crystallin mutants. The spectra were normalized to represent the same number of spins and scaled for convenience of representation. All spectra were recorded with a 100 G scan width.

modes. In the absence of detailed line shape analysis, the width of the central resonance line is a convenient measure of the relative variation in mobility (*20*).

The EPR accessibility parameter, Π, is a measure of the solvent exposure of R1 at a particular site. It is determined from the change in the parameter $P_{1/2}$ in the presence of NiEDDA, a reagent exclusively soluble in the aqueous phase. In monomeric proteins, both Π and $(\Delta H_0)^{-1}$ are primarily determined by the level of tertiary contact interactions (*20, 21*). In oligomeric proteins, steric contact of the nitroxide with neighboring subunits, i.e., quaternary interactions, is another determinant of mobility and accessibility. In general, mobility and accessibility allow the classification of sites as exposed, buried, or in steric contact (*20*). The periodic sampling of these environments when R1 is scanned across a segment of regular secondary structure results in periodic patterns in Π and $(\Delta H_0)^{-1}$ that reflect the type of the secondary structure and its orientation relative to the protein fold (*15*).

Figure 2 shows the sequence-specific variation in Π and $(\Delta H_0)^{-1}$ of R1 along the sequence between residues 60 and 108. Also shown are the data in the 109−120 stretch reported in ref *7*. Periodic patterns in both parameters are observed in the sequences 84−90, 93−101, and 110−120. In all three regions, Π and $(\Delta H_0)^{-1}$ vary with a period of two, consistent with the presence of three β-strands, hereafter referred to as C1, C2, and C3, respectively. In general, R1 scanning through a β-strand results in the sampling of two environments with distinct solvation and structural order as is the case for C3 (*15*). However, superimposed on the rapid variation in Π and $(\Delta H_0)^{-1}$ along C1 and C2 is a low-frequency component with a global maximum at residues 91 and 92. The bell-like variation results in a measurable accessibility at the more sequestered surface of strand C1,

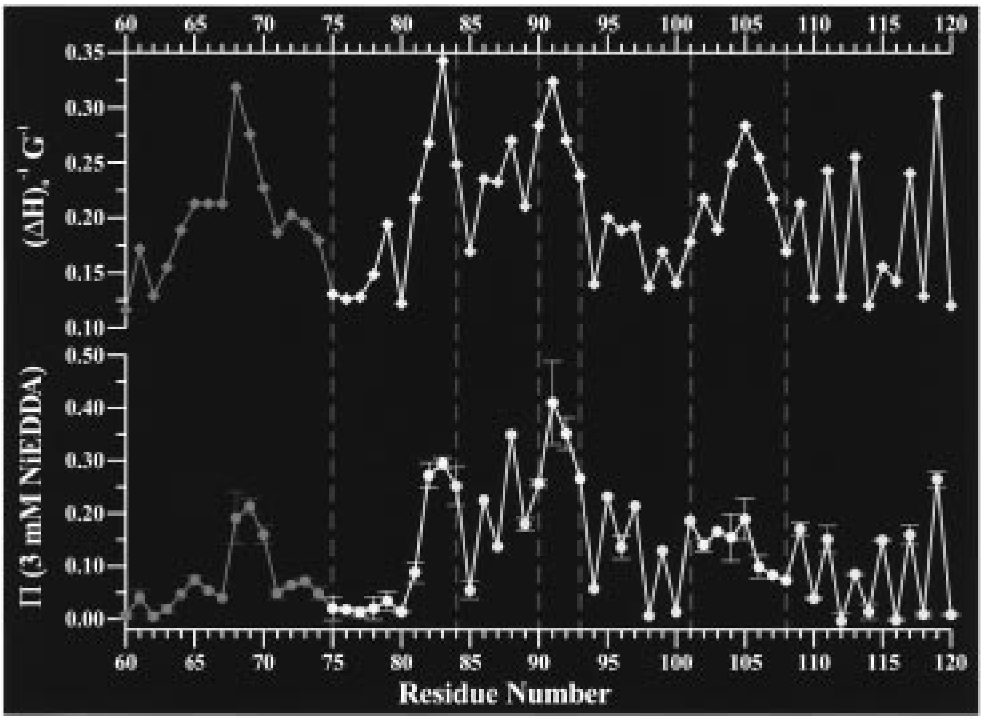


FIGURE 2: Π (NiEDDA) (●, bottom) and $(\Delta H_0)^{-1}$ (◆, top) versus residue number. Error bars are the standard deviation of two to six measurements performed on different days using different protein preparations. The experimental error of individual Π measurements is within the size of the symbol. The data points are color coded to reflect the secondary structure assignment: β-strands are shown in yellow and unstructured regions are shown in white. The dotted green lines indicate possible boundaries between adjacent secondary structures. The interdomain segment is coded in red.

and the line shapes at these sites (e.g., V87R1) are more mobile than at sites buried in the protein hydrophobic core (*20*). At residue 92, the spectral line shape of R1 is indicative of large amplitude motion, and the absolute values of Π at 91 and 92 approach the range expected at exposed sites in monomeric proteins, indicating that this region is located at the solvent−oligomer interface.

Between residues 63 and 75, the putative interdomain segment, Π varies with a period ≥4. The interpretation of this periodicity is problematic since amphipathic helical structures are expected to have a period of 3.6 (*15*, *21*). However, helices with extensive steric contacts can display a slightly higher periodicity particularly when the steric contacts occur along both faces. That the interdomain segment is in strong steric contact along both the first and last turns, i.e., residues 63−67 and 70−75, respectively, is indicated by the immobilized EPR spectra and the marginal acessibility contrast. Along this segment, residues 68 and 69 are local maxima both in mobility and in accessibility.

There are two relatively exposed unstructured regions as deduced from the absence of periodic patterns. The first, consisting of residues 90−93, is a β-turn linking the two β-strands into a possible β-hairpin motif. The second is a loop between residues 102 and 108 linking strand C2 to strand C3. Since the strand spanning residues 110−120 is located at a subunit interface, it qualifies as an end strand.

The three β-strands are connected to the interdomain segment via a loop spanning residues 76−83. The assignment is based on the absence of any periodic variation in either Π or $(\Delta H_0)^{-1}$ even when 10 mM NiEDDA or molecular oxygen was used (data not shown). The residues along this sequence segregate into two groups. One, comprised of residues 76−80, has a highly structured environment with virtually no accessibility to the solvent; the other, from residue 81 to 83, is more exposed with less steric constraints on the dynamics of R1.

Packing of Secondary Structures. To determine the local folding pattern, the β-strands should be arranged relative to each other and their packing pattern determined. A direct approach to this end is through the determination of internitroxide distances for selected pairs introduced on the more exposed surfaces of the β-strands. Because of the limited number of possible packing arrangements between neighboring β-strands in the primary sequence (*16*), the pairs can be designed to provide a distinctive pattern of distances that are unique to one particular motif (see below). To simplify the interpretation, members of the pairs can also be selected to have either a separation of less than 10 Å, resulting in strong spin−spin interactions, or a separation larger than 18 Å with marginal interaction. A qualitative indicator of in-close (≤10 Å) proximity is the broadening of the line shape, drop in the overall normalized spectral intensity relative to that of the sum of the EPR spectra of the single mutants, and the appearance of intensities separated by more than 90 G that dominate the spectral line shape (*22*, *23*).

Two sets of double mutants were constructed. The first, shown in Figure 3a, consists of pairs with one member in strand C1 and the other in strand C2. The second, shown in Figure 3b, consists of pairs with one member in strand C2 while the second member is in strand C3. All double mutants formed native-like oligomers and were at least as efficient as the WT* in suppressing the aggregation of the insulin B chain. The choice of the pairs was based on a β-hairpin arrangement of the strands, suggested by the accessibility profile, and the pattern of distance separation is entirely consistent with this interpretation. Residue D84

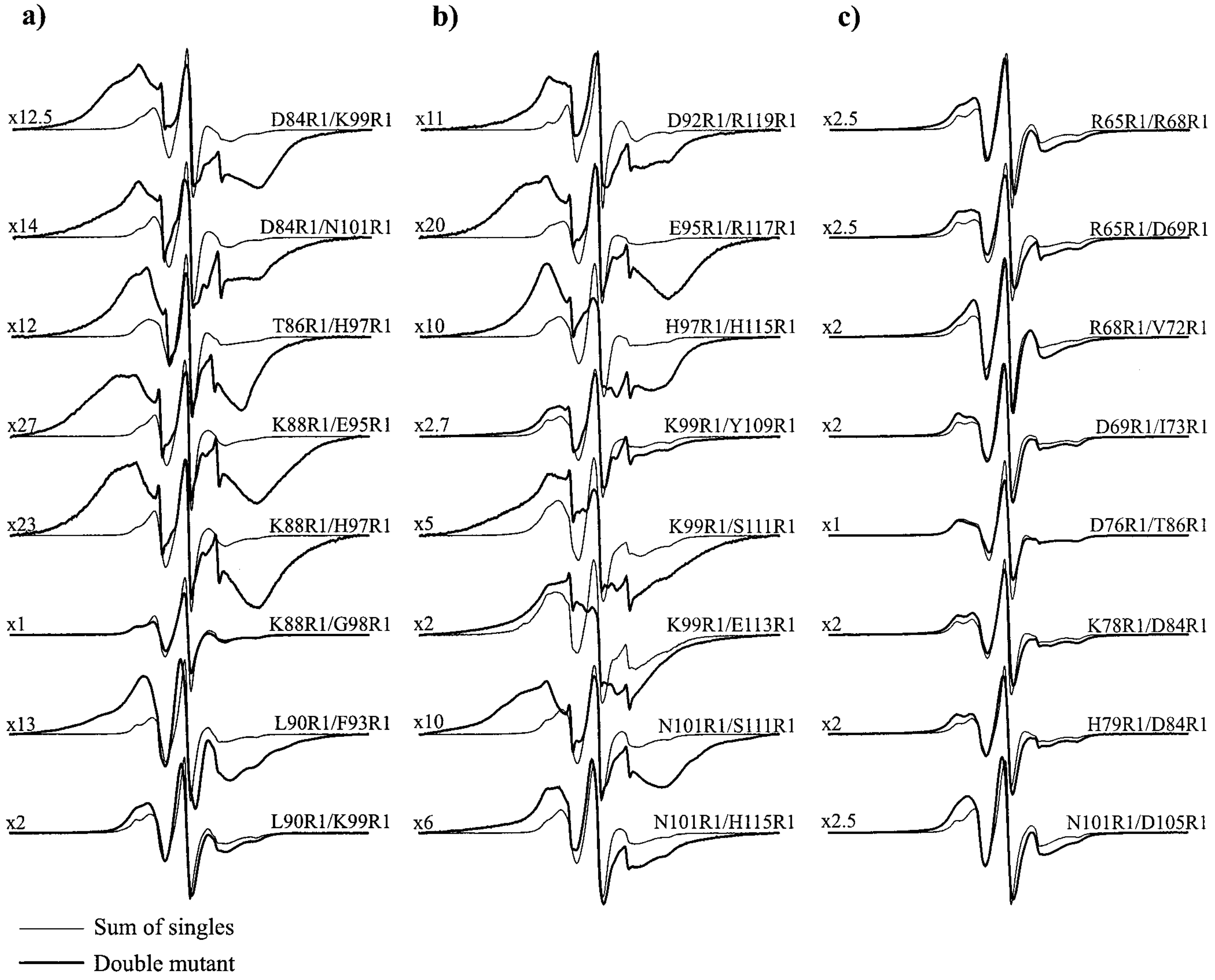


FIGURE 3: Room temperature EPR spectra of the indicated double mutants (heavy trace) and the sum of the corresponding single mutants (light trace). The spectra were normalized to represent the same number of spins. The spectrum of each double mutant was scaled with respect to the corresponding sum of singles to show details of the line shape. The scaling factor is shown at the left of each spectrum. All spectra were recorded with a 200 G scan width.

at the N-terminus of strand C1 is in close contact with residues K99 and N101 at the C-terminus of strand C2. Similarly, R117 at the C-terminus of strand C3 is in close contact with residue E95 near the N-terminus of strand C2. Residues in between have the pattern of proximities expected based on an antiparallel arrangement. In all cases the spin−spin interaction results in broadening of the hyperfine features similar to that observed for spin pairs in T4 lysozyme separated by about 10 Å (*23*). The primary mechanisms of interactions are static dipolar coupling and through-space spin exchange (*24*). The pairs 88/98, 90/99, and 101/115 were designed such that weak spin−spin interactions are expected on the basis of an antiparallel arrangement of the strands. The pair 88/98 is particularly instructive since the 88/97 pair shows extensive broadening. In a β-hairpin, the nitroxide groups at residues 88 and 98 are separated by at least 15−20 Å due to the projection of the side chain in opposite directions. Except for residues 111 and 113, dipolar couplings are not observed in the spectra of the single mutants (Figure 3, thin traces), indicating that the broadening is not the result of intermonomer interactions. The broadening in the 99/113 pair is the result of interactions between three spins as R1 at residue 113 interacts across a subunit interface (*7*). Similar patterns of spin−spin interactions were obtained when the double mutants were refolded in the presence of equimolar amounts of WT* (data not shown).

We also attempted to determine the location of the 75−80 segment relative to strand C1 and simultaneously determine whether this region forms a β-strand, the periodic pattern of which is masked by the lack of anisotropic solvation. The pairs 76/86, 78/84, and 79/84, shown in Figure 3c, were constructed to specifically test the possibility of antiparallel arrangement relative to strand C1. In all cases no significant spin−spin interaction was detected, in agreement with the interpretation above that this region is a continuation of the 81−83 loop, buried due to quaternary interactions between subunits.

Spin−spin interaction was also used as a probe for secondary structure in the interdomain region. The rationale is that if the segment is perfectly helical, residues at i and $i + 4$ are expected to be within 10 Å and extensive broadening is expected (*23*). Attempts to detect such interactions were unsuccessful as shown in Figure 3c for the pairs 65/68, 65/69, 68/72, and 69/73.

DISCUSSION

Despite the length of the investigated region and the variability in the secondary structure and topographical locations, none of the cysteine substitution resulted in a decrease of αA-crystallin efficiency to suppress the aggregation of the insulin B chain. A simple interpretation of this result is that αA binds insulin over an extended surface area, and therefore individual residue substitutions are unlikely to result in significant changes in the chaperone-like efficiency. However, this interpretation cannot explain the increase in efficiency observed at a number of residues. It is generally expected that mutations in buried regions of proteins will affect the activity via long-range effects. For instance, substitutions at buried sites in T4 lysozyme result in a dramatic drop in activity even when the mutations are located away from the catalytic site (*25*). For αA-crystallin, the more buried mutants result in an *increase* in the suppression of aggregation efficiency.

The advantage of the room temperature insulin assay is that it avoids the use of thermal unfolding of the substrate, which can have a differential effect on the structure of the mutants and therefore complicate the interpretation. For α-crystallins, an increase in temperature has the added complication of increasing the efficiency of the chaperone-like function of the WT (*12*, *26*). At 40 °C, Smulders et al. (*26*) report that a 1:1 mass ratio of αA to insulin almost completely suppresses the aggregation of the B chain compared to a ratio of 5:1 at room temperature (*19*). Furthermore, Das and Surewisz report that the room temperature chaperone function is increased, although to a lesser extent, following preincubation of α-crystallin at elevated temperature (*13*). This effect has been interpreted as the result of temperature activation of kinetic processes that expose otherwise inaccessible binding surfaces.

Likewise, if the cysteine substitutions result in an increase in these kinetic rates at room temperature, they will also result in an increase in the apparent stoichiometry. The word “apparent” is used here to emphasize that the aggregation assays, including the insulin assay, are conducted under nonequilibrium conditions. Thus they will reflect, in addition to changes in binding stoichiometry, changes in the kinetics of binding (*19*).

Consistent with this interpretation, the residues most affected by the cysteine substitutions are located within or near the putative binding regions identified in refs *9*−*11*. Structurally, residues 76−80 are in a buried environment although the segment appears to be part of a loop. It is noted, however, that the absence of periodic contrast in Figure 2 might be due to the absence of anisotropic solvation. Therefore, the possibility of the presence of a regular secondary structure cannot be completely excluded. Presumably, this region is buried by quaternary interactions between subunits. Thus access must involve a transient exposure through rearrangement of the quaternary structure. This can be mediated by subunit exchange between oligomers (*13*). The binding region, identified in refs *10* and *11*, also overlaps strand C1. Remarkably, the hydrophobic face of this strand is relatively solvent accessible, consistent with the proposal that substrate binding is mediated by exposed hydrophobic surfaces. Cysteine substitution also resulted in a measurable increase in the chaperone activity at sites along the 98−108 region. Although this stretch was proposed to mediate substrate binding in αB-crystallin rather than αA (*11*), its highly conserved sequence makes it also a possible substrate binding domain in αA.

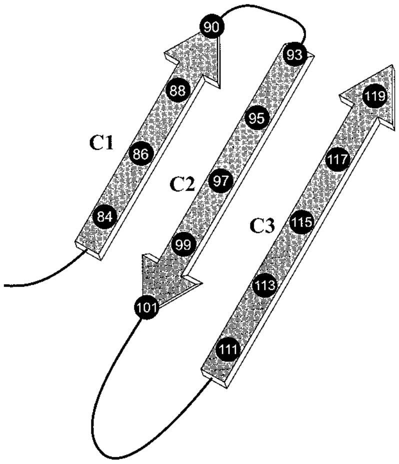


FIGURE 4: Topology diagram of αA-crystallin between residues 83 and 120. β-strands, represented by arrows, are sequentially named C1, C2, and C3. Residues along the exposed surfaces of the strands are pointing out of the plane of the figure. Sites in β-strands used to derive geometric constraints are indicated by the solid circles. The lines representing the turns and loops are not drawn to scale.

An overall picture of the structural framework that mediates the chaperone function of αA-crystallin is starting to emerge. Figure 4 is a schematic representation of the structural data reported in this paper. The data define the main structural characteristics of a β-sheet: the number of strands, their relative direction, and their connectivity along the polypeptide chain. Although these characteristics were used to generate an atomic model using the Insight II software (Molecular Simulations), the absence of a packed core involving the hydrophobic surface of the sheet precludes the use of energy minimization at this stage. Thus the results are best represented using the topology diagram (*27*) of Figure 4.

The secondary structure assignment is based on the periodic patterns in the accessibility and mobility and has an inherent uncertainty in defining the ends of the strands. Nevertheless, the biphasic periodicity clearly indicates the presence of three β-strands in the region between residues 60 and 120. Strands C1 and C2 are connected via a β-turn that defines a solvent−oligomer interface, while strands C2 and C3 are connected by a highly conserved polar loop. The absence of intervening helices generally indicates an antiparallel arrangement. However, the data of Figure 2 do not distinguish between β-hairpin motifs, with side to side packing, and greek key motifs that connect antiparallel strands on opposite sides of a barrel or sandwich.

That the three strands form a β-sheet with pairwise β-hairpin packing is established from patterns of proximities between pairs of nitroxide side chains introduced on the more exposed surfaces of neighboring strands in the primary sequence. The use of the pattern approach is based on the fact that β-strands tend to pack in a limited number of motifs, each characterized by a unique set of proximities (*16*). The expected differences in the proximity patterns are larger than the uncertainty in the orientation of the side chain. For instance, a β-hairpin will result in R1 separations of less than

10 Å at sites with hydrogen-bonded backbones. The pattern of such interactions, where residues at the N-terminal end of one strand are in close contact with residues at the C-terminal end of a neighboring strand in the primary sequence, is unique to the β-hairpin motif. A distinct pattern is expected from a greek key connectivity where equivalent pairs are in different sheets and point in opposite directions.

The advantage of the pattern approach, used previously to investigate conformational changes in T4 lysozyme and rhodopsin (*23*, *28*), is that accurate distances are not required. Conclusions are based on the coarse agreement of a statistically significant set of pairs selected to distinguish between models. Therefore, the impact of the inherent lack of accuracy in the measured distance, due to uncertainty in the position of the nitroxide relative to the α-carbon, on the structural interpretation is minimal. However, many important structural details, such as the right-handed twist of the β-strands and the exact alignment of neighboring β-strands, are lost. These details can be obtained in later stages by quantitative analysis of distances (*22*, *29*) and relative orientation (*22*) between the nitroxides and will transform topology diagrams to relevant atomic models.

The β-sheet is connected to the N-terminal domain by the interdomain segment and is demarcated at one end by interaction with another subunit. The absence of the interdomain segment from Figure 4 reflects the uncertainty in its location relative to the β-sheet. Attempts to obtain inter-nitroxide interactions between R1 pairs at residues i and $i + 4$ along this segment were not successful. Thus, if the segment has a helical conformation, it is likely that the helix is either irregular or consists of less than two full turns. The absence of regular secondary structure in this region is consistent with the expected flexibility, which allows the insertion of a 13 amino acid peptide in a variant of rat αA-crystallin (*30*).

Finally, the work presented in this paper provides the structural basis to guide further biochemical and mutagenic studies addressing recognition and binding in sHSP. In addition, our results suggest that temperature activation of the chaperone-like function will have defined structural consequences such as a measurable increase in solvent accessibility in the 76−80 loop. The pattern recognition approach can be extended to identify the overall fold of a subunit. The geometric aspect of known all-β folds can be reduced to distinctive sets of spatial proximities. The compatibility of the structure with these folds can be evaluated using SDSL.

REFERENCES

1. Caspers, G., Leunissen, J. A. M., and de Jong, W. W. (1995) *J. Mol. Evol. 40*, 238−248.
2. Lee, G. J., Pokala, N., and Vierling, E. (1995) *J. Biol. Chem. 270*, 10432−10438.
3. Horwitz, J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10449−10453.
4. Merck, K. B., De Haard-Hoekman, W. A., Oude Essink, B. B., Bloemendal, H., and De Jong, W. W. (1992) *Biochim. Biophys. Acta 1130*, 267−276.
5. Merck, H. B., Horwitz, J., Kersten, M., Overkamp, P., Gaestel, M., Bloemendal, H., and De Jong, W. W. (1993) *Mol. Biol. Rep. 18*, 209−215.
6. Wistow, G. (1993) *Exp. Eye Res. 56*, 729−732.
7. Berengian, A. R., Bova, M. P., and Mchaourab, H. S. (1997) *Biochemistry 36*, 9951−9957.
8. Mchaourab, H. S., Berengian, A. R., and Koteiche, H. A. (1997) *Biochemistry 36*, 14627−14634.
9. Lee, G. J., Roseman, A. M., Saibil, H. R., and Vierling E. (1997) *EMBO J. 16*, 659−671.
10. Sharma, K. K., Kaur, H., and Kester, K. (1997) *Biochem. Biophys. Res. Commun. 239*, 217−222.
11. Sharma, K. K., Kaur, H., and Kester, K. (1998) *Invest. Ophthalmol. Visual Sci. 39*, S870.
12. Das, K. P., and Surewicz, W. K. (1995) *FEBS Lett. 369*, 321−325.
13. Bova, M. P., Ding, L., Horwitz, J., and Fung, B. K. (1997) *J. Biol. Chem. 272*, 29511−29517.
14. Haley, D. A., Horwitz, J., and Stewart, P. L. (1998) *J. Mol. Biol. 277*, 27−35.
15. Hubbell, W. L., Mchaourab, H. S., Altenbach, C., and Lietzow, M. A. (1996) *Structure 4*, 779−783.
16. Chothia, C., and Finkelstein, A. V. (1990) *Annu. Rev. Biochem. 59*, 1007−1039.
17. Berliner, L. J., Hankovsky, H. O., and Hideg, K. (1982) *Anal. Biochem. 119*, 450−453.
18. Farahbakhsh, Z. T., Altenbach, C., and Hubbell, W. L. (1992) *Photochem. Photobiol. 56*, 1019−1033.
19. Farahbakhsh, Z. T., Huang, Q., Ding, L., Altenbach, C., Steinhoff, H., Horwitz, J., and Hubbell, W. L. (1995) *Biochemistry 34*, 509−516.
20. Mchaourab, H. S., Lietzow, M. A., Hideg, K., and Hubbell, W. L. (1996) *Biochemistry 35*, 7692−7704.
21. Hubbell, W. L., and Altenbach, C. (1994) *Curr. Opin. Struct. Biol. 4*, 566−573.
22. Hustedt, E. J., Smirnov, A. I., Laub, C., Cobb, C. E., and Beth, A. H. (1997) *Biophys. J. 72*, 1861−1877.
23. Mchaourab, H. S., Oh, K. J., Fang, C. J., and Hubbell, W. L. (1997) *Biochemistry 36*, 307−316.
24. Eaton, G., and Eaton, S. (1989) in *Biological Magnetic Resonance* (Berliner, L. J., and Reuben, J., Eds.) Vol. 8, pp 340−397, Plenum, New York.
25. Matthews, B. W. (1995) *Adv. Protein Chem. 46*, 249−278.
26. Smulders, R. H. P. H., Carvers, J. A., Lindner, R. A., Van Boekel, M. A. M., Bloemendal, H., and De Jong, W. W. (1996) *J. Biol. Chem. 271*, 29060−29066.
27. Levitt, M., and Chothia, C. (1976) *Nature 261*, 552−558.
28. Farrens, D. L., Altenbach, C., Hubbell, W. L., and Khorana, H. G. (1996) *Science 274*, 768−770.
29. Rabenstein, M. D., and Shin, Y. K. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8239−8243.
30. Smulders, R. H. P. H., Van Geel, I. G., Gerards, W. L. H., Bloemendal, H., and De Jong, W. W. (1995) *J. Biol. Chem. 270*, 13916−13924.

BI9814078