Site-Directed Tryptophan Fluorescence Reveals the Solution Structure of Tear Lipocalin: Evidence for Features That Confer Promiscuity in Ligand Binding[†]

Oktay K. Gasymov, Adil R. Abduragimov, Taleh N. Yusifov, and Ben J. Glasgow*

Departments of Pathology and Ophthalmology, UCLA School of Medicine, Los Angeles, California 90095

Received May 21, 2001; Revised Manuscript Received October 18, 2001

ABSTRACT: The solution structure of human TL was deduced from the position of the emission peaks after site-directed tryptophan fluorescence (SDTF). The fluorescent amino acid tryptophan was sequentially substituted for each native amino acid in the sequence. Characteristic periodicities for eight β-strands that comprise the β-barrel and three α-helices were identified. The putative β-strand I was relatively exposed to solvent, suggesting it does not participate in the formation of the β-barrel. The β-strands A and F contain β-bulges. The average λ_{max} of emission maxima reveals that strand D is at the edge of the barrel and β-strand H interacts with the main α-helical domain. On the basis of the SDTF data, a 3D homology model was constructed for TL and compared to the known crystallographic structures of RBP and β-lactoglobulin. The small size and splayed open configuration of the E−F hairpin facilitate access of ligands into the cavity mouth of TL as compared to that of RBP with a long overhanging loop that restricts access. In the model of TL, four alanine residues are positioned in the binding site as compared to bulkier residues in the corresponding positions of β-lactoglobulin. Substitution of A51, A66, A86 to Trp results in a 3−4-fold decrease in binding affinity. The data suggest that the smaller side chains of Ala provide more capacity in the cavity of TL than the bulkier side chains (I56, I71, V92) in the cavity of β-lactoglobulin. The structural features provide an explanation for the promiscuous binding characteristics exhibited by TL. SDTF provides a general approach for determining the solution structure of many proteins and enhances homology modeling in the absence of high sequence identity.

TL[1] is the principal lipid binding protein in tears and is bound to fatty acids, glycolipids, phospholipids, and cholesterol in tears (*1*). The broad range of ligand specificity exhibited by TL is slightly greater than that of β-lactoglobulin but much more pronounced than other members of the lipocalin family. For example, RBP and bilin binding protein show highly selective ligand domains that are limited to very specific molecules. X-ray crystallography and computer-generated models of several lipocalins have revealed structural features in common (*2*−*5*). Eight strands are arranged in a β-barrel and are joined by loops between the β-strands (*2*). Three structurally conserved regions in β-strands A, F−G, and H have been identified for the lipocalins (*2*). Although the crystal structure of TL is not known, the secondary structure of β-strand G has been resolved by analysis of fluorescence after sequential substitution of native amino acids by tryptophan (SDTF) and later validated by site-directed spin-labeling (*6*, *7*). The spectral positions of the fluorescent emission maxima (λ_{max}) for tryptophan mutants reveal alternating periodicity of β-strand G residues in TL, accurately reflecting the β-strand pattern of solvent exposure. The relatively simple data obtained from λ_{max} correlate with detailed analysis of quenching experiments.

For other proteins, the analysis of λ_{max} for fluorescent chromophores placed along the peptide chains accurately reflects secondary structural elements. The emission peaks of fluorescent labeled cysteine mutants of T4 lysozyme represent solvent surface accessibility and correlate with the crystal structure (*8*). The λ_{max} of synthetic peptides with tryptophan substitutions provide resolution of α-helical properties with minimal perturbational effects (*9*). Similarly, loop positions are clearly differentiated from α-helical positions by analysis of λ_{max} at key residues for the Tet repressor (*10*).

Recent advancements in tryptophan fluorescence have yielded abundant local structural information regarding solvent accessibility, dynamics, relaxation, and distance between donor−acceptor pair, etc. (*11*). One of the applications, SDTF, has opened new possibilities for investigation of protein structure and dynamics. SDTF has been used in combination with nitroxide spin-labeled fatty acids to delineate ligand position and motion within the cavity of TL (*12*). Data obtained by SDTF have been in good agreement with other techniques (*7*, *8*). In some cases, SDTF has been informative about regions not resolved in crystal structures due to lack of electron density (*10*). These studies have focused on peptide fragments or parts of proteins. Here, we extend this work to the resolution of the secondary structure of an entire protein, TL. Despite the fact that TL has only

[†] Supported by USPHS NIH Grant EY 11224 and by an unrestricted grant from Research to Prevent Blindness to B.J.G.

* Correspondence should be addressed to this author at 100 Stein Plaza, Room B-279, Los Angeles, CA 90095. Fax: (310) 794-2144. Phone: (310) 825-6998. E-mail: bglasgow@mednet.ucla.edu.

[1] Abbreviations: 16-DSA, 16-doxylstearic acid; CD, circular dichroism; DAUDA, 11-(((5-(dimethylamino)-1-naphthalenyl)sulfonyl)amino)undecanoic acid; λ_{max}, fluorescent emission maxima; RBP, retinol binding protein; SDTF, site-directed tryptophan fluorescence; TL, tear lipocalin; UV, ultraviolet.

10.1021/bi0110342 CCC: $20.00
Published on Web 11/15/2001

23% sequence identity with β-lactoglobulin, the creation of a 3D computer model for TL was possible because secondary structure could be assigned experimentally by SDTF. The resulting model permits a detailed structural comparison of TL to β-lactoglobulin as well as RBP and reveals features in the lipocalin scaffold that may confer ligand specificity.

EXPERIMENTAL PROCEDURES

Site-Directed Mutagenesis and Plasmid Construction. The TL cDNA in PCR II (Invitrogen, San Diego, CA), previously synthesized (*13*), was used as a template to clone the TL gene spanning bases 115−592 of the previously published sequence (*14*) into pET 20b (Novagene). Flanking restriction sites for *Nde*I and *Bam*HI were added to produce the native protein sequence as found in tears (*15*). To avoid the use of subtraction spectra in analysis and to simplify the interpretation of quenching experiments in molecules with substituted tryptophans, it was first necessary to substitute the native tryptophan with an amino acid that did not induce significant structural perturbations and has similar binding characteristics (*6*, *16*). We prepared a TL mutant, W17Y, with oligonucleotides (Universal DNA Inc.) using the previously published method of introduction of a point mutation by sequential PCR steps (*17*). Using this mutant as a template, mutant cDNAs were constructed in which the corresponding amino acids, 1−93 and 106−159 (except C61 and C153, which form the disulfide bond), were additionally substituted sequentially with tryptophan. Tryptophan mutants were previously constructed for 94−105 (*6*). Amino acid 1 corresponds to His, bases 115−118 according to Redl (*14*). One additional mutant, M22W/K114C, was constructed to test the hypothesis that the positive charge of K114 shifts the λ_{max} of M22W.

Expression and Purification of Mutant Proteins. The mutant plasmids were transformed in *E. coli* BL21(DE3), and cells were cultured and protein was expressed according to the manufacturer's protocol (Novagene). Following cell lysis (*18*), the supernatant was treated with methanol (40% final concentration) at 4 °C for 2.5 h. Alternatively, mutant proteins expressed in inclusion bodies were dissociated in 8 M urea at room temperature for 2 h. In either case, the resulting suspension was centrifuged at 3000*g* for 30 min. The supernatant was dialyzed against 50 mM Tris-HCl, pH 8.4. The dialysate was treated with ammonium sulfate 45−75% saturation. The resulting precipitate was dissolved in 50 mM Tris-HCl, pH 8.4, and applied to a Sephadex G-100 column (2.5 × 100 cm) equilibrated with 50 mM Tris-HCl, 100 mM NaCl, pH 8.4. The fraction containing the mutant protein was dialyzed against 50 mM Tris-HCl, pH 8.4, and applied to a DEAE Sephadex A-25 column. Bound protein was eluted with a 0−0.8 M NaCl gradient. Eluted fractions containing mutant proteins were centrifugally concentrated (Amicon, Centricon-10). The purity of mutant proteins was verified by SDS−Tricine gel electrophoresis (*1*). The protein concentration was determined by the biuret method (*19*). For ligand binding experiments, TL-WT and mutants were delipidated with chloroform and methanol as described previously (*1*).

CD Spectral Measurements. Spectra were recorded (Jasco 600 spectropolarimeter, 0.2 mm path length for far-UV spectra) using protein concentrations of 1.2 mg/mL. Eight scans were performed from 190 to 260 nm. Results were recorded in millidegrees and converted to mean residue ellipticity in deg·cm^2·dmol^{-1}.

Fluorescence Spectroscopy. Fluorescence measurements were made on a Jobin Yvon-SPEX Fluorolog-3 spectrofluorometer; bandwidths for excitation and emission were 2 nm. The excitation λ of 295 nm was used to ensure that light was absorbed almost entirely by tryptophanyl groups. Protein solutions with less than 0.1 OD at 295 nm were analyzed. All spectra were obtained from samples in 10 mM sodium phosphate, pH 7.3, at room temperature. The fluorescence spectra were corrected for light scattering from buffer and the instrument response using the appropriate correction curve. The 467 nm Xenon lamp peak was used for calibration of the excitation monochromator. The emission monochromator was calibrated by the position of the water Raman band, to 397 nm with the excitation wavelength set to 350 nm.

Construction of a 3D Model of TL. For modeling, we utilized the INSIGHT II modeling package (MSI, Inc.) run on an Octane Silicon Graphics computer. The coordinates of β-lactoglobulin were retrieved from the Protein DataBank (1B0o). The experimentally determined secondary structure of TL was designated as structural conserved regions and aligned with that of β-lactoglobulin (Figure 2) using the HOMOLOGY module. Because β-lactoglobulin has a β-bulge in strand B, it was necessary to align this region of TL with that of RBP (PDB ID: 1RBP). The orientations of side chains within the β-strands (exposed or buried) and α-helix (sites with interaction to the β-barrel) were taken into account for this alignment. Based on the structurally conserved regions in the reference structure, residues from the β-strands of TL were assigned to the coordinates of β-lactoglobulin. This assignment yielded an identical match of 23% of the residues between TL and β-lactoglobulin. The structures of loops, as well as N- and C-terminal portions of TL, were modeled using randomly generated values for the loops optimized by energy minimization. In the final step, the entire protein structure was energy-minimized using 1000 steps of steepest descent, followed by 700 steps of conjugate gradient descent using the Discover module (*20*).

EPR Measurements. Electron paramagnetic resonance spectra were recorded at X-band with a Varian E-109 spectrometer fitted with two-loop one-gap resonator (*21*). Samples of 5 μL were loaded in 0.84 mm o.d. capillaries (Vitrocom Inc., Mountain Lakes, NJ) that were sealed at one end. All spectra were acquired at X-band using a 2mW incident microwave power and an 80 G scan width. Acquisition and manipulation of EPR spectra were carried out using computer programs written in LabVIEW (National Instruments, Austin, TX) by Dr. C. Altenbach. The modulation amplitude was 1 G and the time constant 0.3 s. Signal averaging was performed with 30 scans. For ligand binding, a stock solution of 16-DSA (Sigma Chemical Co., St. Louis, MO) was prepared in ethanol. A constant concentration of 16-DSA was titrated with progressively increasing concentrations of TL-WT or mutant proteins, and the EPR spectra were recorded. The final concentration of ethanol did not exceed 2%. The EPR spectra revealed the presence of both protein-bound and free spin-label. The ratio of bound to free spins in equilibrium was estimated by spectral titration (*1*). The free spectrum of the spin-label was subtracted from a composite spectrum of bound and free to give the pure bound

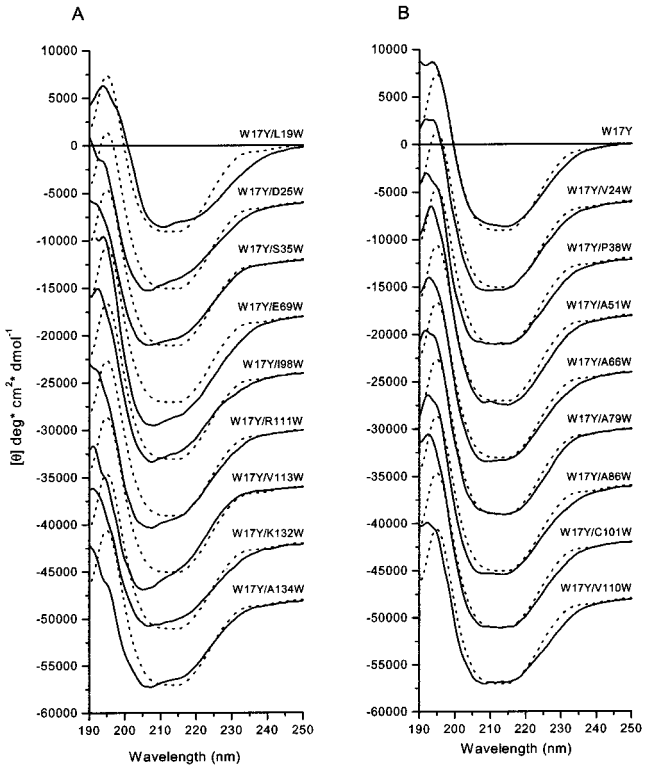

FIGURE 1: Far-UV CD spectra of WT-TL (dotted line) and its mutant proteins (solid lines). To facilitate comparison, the baseline of each successive spectrum has been shifted −6000 deg*cm^2*dmol^{-1}. (A) Mutants with distortion of secondary structure. (B) Mutant without tryptophan (template) and other representative mutants with minimal spectral changes.

spectrum. The ratio of double integrals of the pure bound to pure free gave the bound to free molar ratio. The concentrations of spin-label in solution, calculated by double integration of the EPR spectra, were compared with a standard solution of Tempol (*1*). The Scatchard plots were fitted by standard nonlinear regression techniques (using Microcal ORIGIN software, Northampton, MA) to a linear curve to give estimates of the dissociation constant (K_d).

RESULTS

Circular Dichroism. The circular dichroic spectra of expressed W17Y mutant TL alone and with sequentially substituted tryptophans for amino acids K94−L105 have been published previously (*6*). Of the 143 new mutants, 9 showed some distortion of secondary structure (L19W, D25W, S35W, E69W, I98W, R111W, V113W, K132W, A134W) (Figure 1A). The major perturbation was a shifted minimum at about 207 nm (Figure 1A) that may be attributed to changes in packing of secondary structural elements (*22*). The other mutants showed an expected loss of the positive contribution at 230 nm (Figure 1B) that is attributable to the absence of Trp 17 (*16*). There were no other significant differences in the far-UV CD spectra of these mutants compared with those of W17Y mutant or the wild-type TL (Figure 1B).

Fluorescence. Tryptophanyl residues that are exposed to a polar solvent and have high mobility reveal emission peaks that are red-shifted. Residues in an apolar environment (buried in the protein) reveal blue-shifted emission peaks (*11*, *23*, *24*).

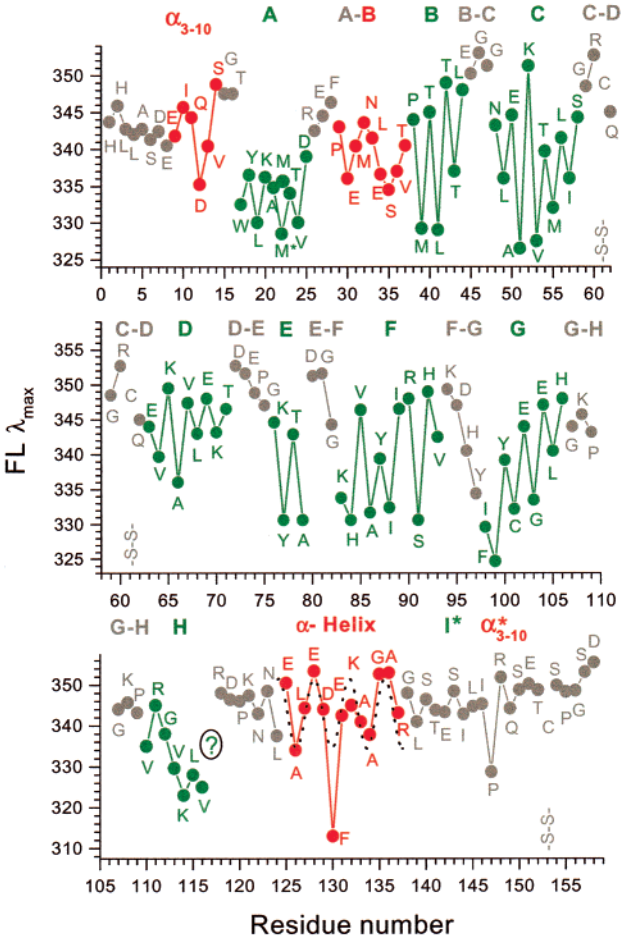

FIGURE 2: Site-directed tryptophan fluorescence data for TL. The maximum emission peaks of the corrected spectra of tryptophan-substituted mutants are shown in sequence. The β-strands are shown in green and the α-helices in red. Gray is used where no secondary structure is evident such as in some areas of loops. The color of the letters indicates secondary structure as above. Single letters refer to β-strands (e.g., A is the A β-strand). Loops are designated by the letters of the connecting strands (e.g., A−B is the loop between the A and B β-strands). The A−B loop is depicted by gray and red letters to indicate some α-helical content. Asterisks indicate regions of putative secondary structure that could not be unambiguously assigned. The dotted line is the fitting curve for α-helical content. The question mark is shown for position 117 because mutant G117W could not be expressed in a dimeric state. M* is λ_{max} for M22W/K114C.

Analysis of the patterns produced by the λ_{max} reveals the major structural motifs for TL including β-strands, β-bulges, α-helices, and loop regions (Figure 2). In assigned β-strand regions, the minimum shift of λ_{max} between adjacent residues was 3 nm, and the maximum shift was 25 nm. Alternating periodicity of the emission peaks in the regions encompassing residues 17−25 (A), 38−44 (B), 48−58 (C), 63−71 (D), 76−79 (E), 83−93 (F), I98−L106 (G), and V110−V116 (H) reflects the changing exposure to solvent as predicted for the β-strands that comprise the barrel (Figure 2). Even perturbations to alternating periodicity are informative. The alternating periodicity of strand H becomes disrupted at V113. V113 and L115 show less exposure than predicted for a β-strand in which one side faces the solvent. In the lipocalin family, a conserved α-helical segment interacts with strand H. Our data are consistent with the orientation of the side chains of V113 and L115 at the interface between the

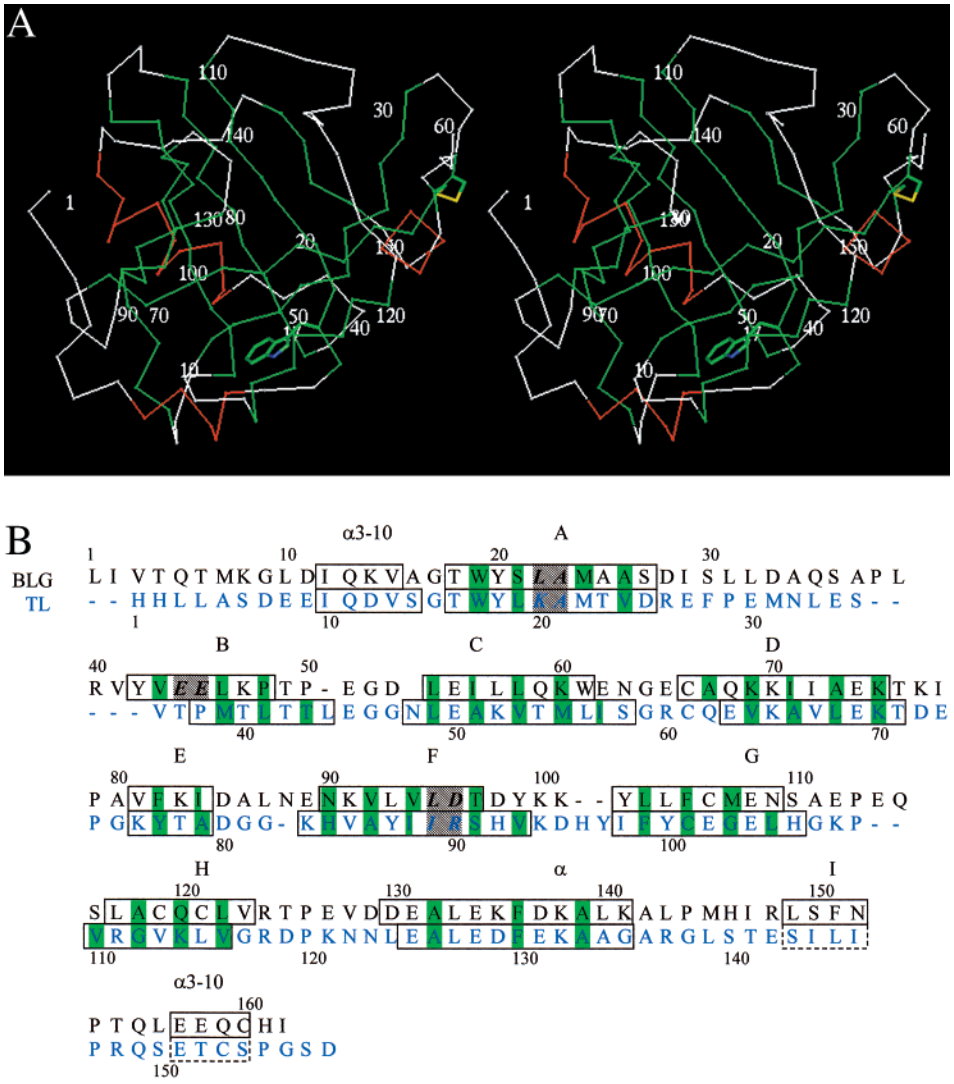

FIGURE 3: (A) Stereo diagram of the C_α trace for the TL model developed from SDTF data in combination with homology modeling. β-Strands (A−H) are shown in green, and α-helices are depicted in red. The side chain of Trp17 (nitrogen is shown in blue) and the disulfide bridge (sulfur is shown in yellow) connecting C61 to C153 are shown as the stick figure. (B) Sequence comparison of β-lactoglobulin and TL. Aligned buried residues of the β-strands are shaded in green. Residues of β-bulges are shaded with diagonal lines. Residues assigned to regions of designated secondary structure are boxed. Dashed boxes are regions of putative structure that could not be unambiguously assigned.

conserved α-helical segment and the β-barrel. In addition, close inspection of Figure 2 shows that alternating periodicity is interrupted for residues I89−R90 on strand F, indicating a β-bulge at this site. Similarly, residues near the end of strand A, K20−M22, may represent positions in another β-bulge. Our data fail to unambiguously identify the residues of strand I because only slight variation of λ_{max} was detected in alternating residues in the expected region.

Loops that join the β-strands of the TL calyx are apparent in SDTF as areas with patternless positions of emission peaks sometimes containing short segments of α-helix. The loops are variable in size. The longest loop in TL is loop A−B (12 residues, R26−T37). Crystallographic structures of other members of the lipocalin family show a consistent length of 10−14 residues comprising loop A−B (*2*−*5*, *25*−*27*). In contrast, loop E−F is short (4 residues) in TL, but similar to loop E−F in β-lactoglobulin (5 residues).

α-Helical structure is evident in the pattern of the emission peaks that reveal a periodicity of about 3.6 for residues P29−T37 and E125−R137 (Figure 2). The α-helix at P29−T37 in loop A−B of TL is analogous to the α-1 helix (D28−D33) of β-lactoglobulin (*25*). Residues E9−S14 may also represent limited α_{3-10} helical structure within the loop prior to β-strand A. This region corresponds to α_{3-10} helical structure in β-lactoglobulin (*25*).

Energy Minimization. Residues of TL were assigned to β-strand coordinates of β-lactoglobulin from our experimental data (Figures 2 and 3). Energy minimization results in the assignment of favorable conformations for the loops as well as the C- and N-terminal regions of TL. The resulting model produces several minor changes at the ends of some β-strands. For example, Y97 was initially assigned as a loop according to our SDTF data. However, in the model, Y97 achieves an energetically favorable orientation assigned to the end of strand G. Our data fit the model if Y97 is oriented in a more exposed position than the adjacent residue, I98. The adjustment is reasonable because in the model, Y97 appears to form the base of the cavity and would be partially exposed to solvent.

The results of energy minimization also provide a rational explanation for some ambiguous results. For example, on the basis of λ_{max} data, we assigned K20−A21 to a β-bulge in strand A. Given this assignment, M22 would be expected to exhibit a relatively blue-shifted emission peak. In fact, the λ_{max} of M22W is similar to A21W. The 3D model places M22 in close proximity [K114(NZ)−M22(CE) = 4.8 Å] to the positively charged residue K114. It is known that a positive charge near the benzene ring of the indole group creates a red shift of fluorescence (*28*). Hence, the red shift we observed with M22W may be produced by the close proximity of the side chain of K114. In keeping with this model, replacement of lysine at position 114 by cysteine (uncharged residue), mutant M22W/K114C, yields a blue shift of the λ_{max} from 335.7 to 328.5 nm (Figure 2).

Influence of Amino Acid Residue Side Chains in the Cavity on Ligand Binding. The results of ligand binding after substitution of the small amino acid residues (A51, A66, A79, A86) that putatively participate in the binding site of TL are shown in Figure 4. The composite EPR spectra demonstrate the signals of free and bound 16-DSA which can be easily monitored by high-field resonance peaks. Free 16-DSA gives rise to a narrow peak, while bound ligand is evident as a broad immobilized component (Figure 4, apo-TL). The presence of Trp at positions 51, 66, and 86 resulted in a diminution of the immobilized component and a concomitant increase in the narrow resonance peak, indicating decreased binding compared to apo-TL. However, at position 79, the substitution with Trp resulted in a similar free and bound 16-DSA component as compared to apo-TL, indicative of the close binding characteristics of these two proteins. The binding constants shown in parentheses were derived from the Scatchard plot of 16-DSA binding to each mutant (Figure 4B). It is evident that the presence of the larger residue at positions 51, 66, and 86 result in an approximate 3−4-fold decrease in binding. However, the K_d for the mutant with a substitution at position 79 is nearly the same as apo-TL.

DISCUSSION

SDTF has disclosed the solution structure of TL, a dimeric protein for which the size is not conducive to the current methods of NMR analysis and for which the crystal structure has not yet been obtained. To date, the predictions of TLs' structures were based on theoretical grounds. Nine regions of β-structure and one α-helix were predicted in TL by Redl et al. from the Chou−Fasman method (*14*). There are two major differences from our data. First, the region that was predicted as β-strand A, residues 9−12, is part of an α_{3-10} helix by SDTF (Figure 2). Second, the β-structure we observed in residues 76−79 (strand E) was not predicted from the hydropathy profile. These errors resulted in the misidentification of strands A−D. There are numerous other minor variations, but the general regions including that of α-helix are fairly close to the predicted values.

Our data indicate that TL has similar lengths and positions of the β-sheets to most other lipocalins. The backbone of lipocalin is a conserved scaffold; the specificity of ligand binding is conferred by differences in loop size and amino acid differences in key positions. Differences in length and conformation of loops C−D and E−F have been found to be an important determinant in the size of the cavity mouth

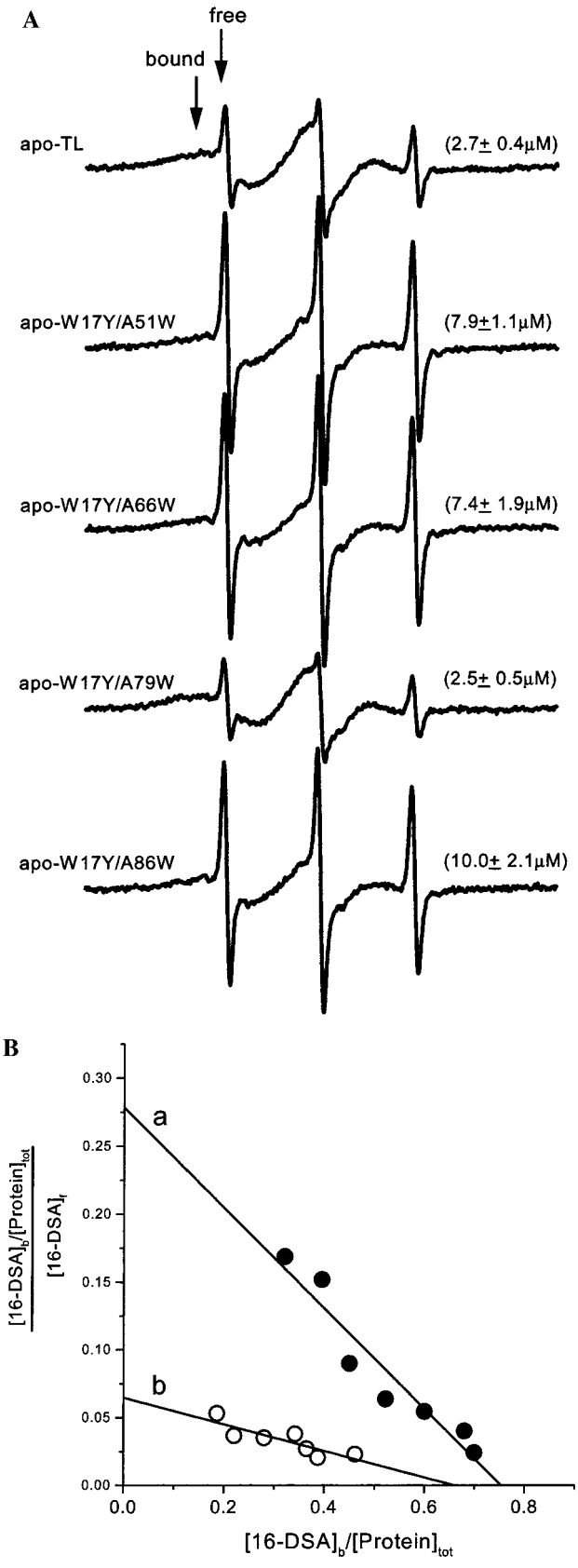

FIGURE 4: (A) X-band EPR of 16-DSA (50 μM) and protein complexes (100 μM). The dissociation constant for each protein is shown in parentheses. Scan width is 80 G. (B) Scatchard plot of 16-DSA binding to apo-TL (a) and apo-W17Y/A86W (b), where $[16\text{-DSA}]_{bd}$ = bound ligand, $[16\text{-DSA}]_f$ = free ligand, and $[\text{Protein}]_{tot}$ = total protein. The experiments were performed in 10 mM phosphate buffer, pH 7.3.

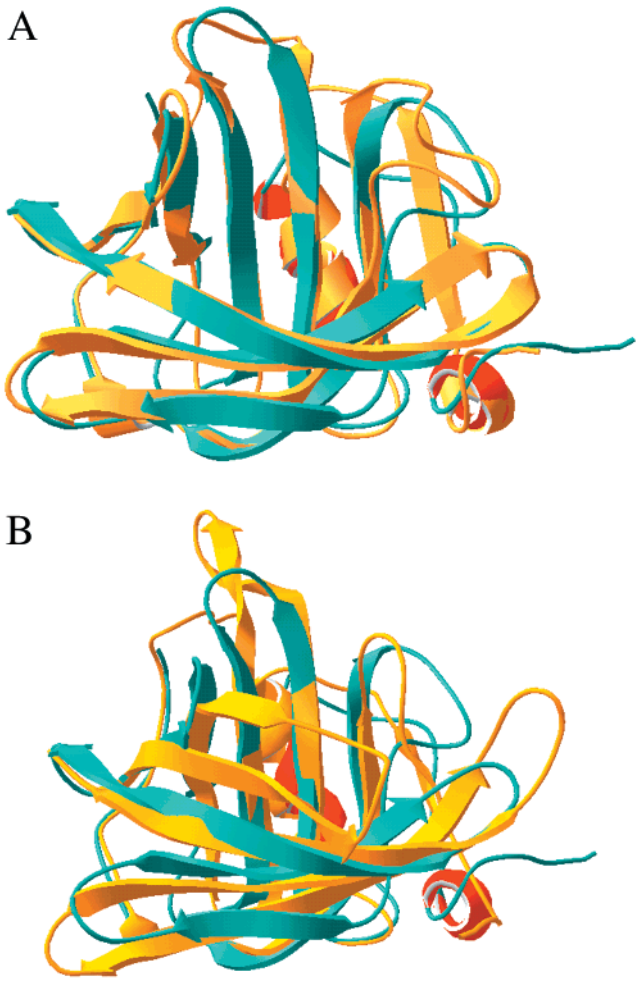

FIGURE 5: (A) Comparison of the homology-modeled structure of TL (green) with the crystallographic structure of β-lactoglobulin (yellow). β-Strands are shown as smoothly curving arrows, α-helices as red spiral ribbons, and loops as cords. (B) Comparison of the homology-modeled structure of TL (blue) with the crystallographic structure of RBP (yellow). β-Strands are depicted in a ribbon diagram.

for RBP and bilin binding protein (*25, 26*). In contrast to RBP and bilin binding protein, TL and β-lactoglobulin are promiscuous proteins, which bind a wide array of lipid ligands (*1*, *29*, *30*). The secondary structure of TL by SDTF is strikingly similar to the crystallographic structures of β-lactoglobulin (*25*). The spacing and lengths of the β-strands and α-helix correspond quite closely. For example, in both proteins, strand E is the shortest of the β-strands (four residues in TL and five residues in β-lactoglobulin). The hairpin formed from strands E and F and the adjoining loop overhang the mouth of the cavity in lipocalins. It is apparent from our data and our model that the E−F hairpin is much shorter in TL and β-lactoglobulin than in RBP (Figure 5). The binding domain of RBP is far more restrictive, and the E−F hairpin is much longer. The additional length in the E−F hairpin is derived from the greater length in both loop E−F and the corresponding strands (*27*). Hence, the short hairpin E−F in TL confers a wide cavity entrance and ligand promiscuity compared to RBP.

The average λ_{max} values represent the overall solvent exposure for a loop or β-strand and hence reflect their relative positions in the β-barrel. The residues assigned to the loops are generally red-shifted (Figure 6). The notable exception

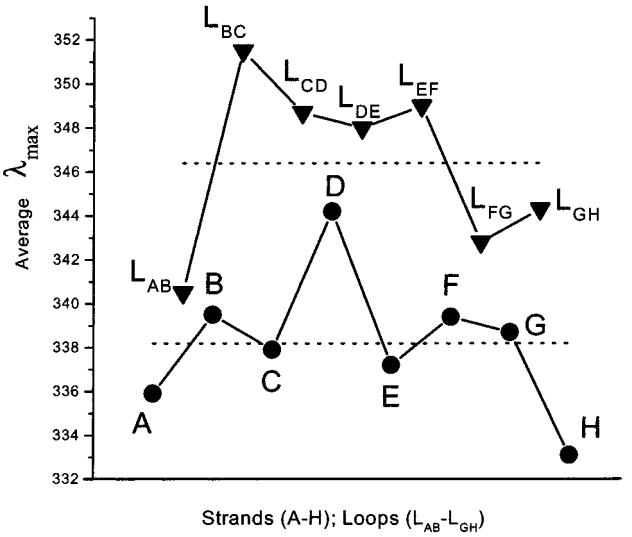

FIGURE 6: Average λ_{max} for loops (L) and β-strands (A−H) of TL. The subscripts for loops indicate adjoining strands. The dashed lines indicate the overall average of loops (upper dashed line) and β-strands (lower dashed line).

is loop A−B. A blue-shifted average λ_{max} suggests an interaction with other strands. Interestingly, in most lipocalins, loop A−B is folded back to join the two orthogonal β-sheets. Consistent with this feature, the computerized model of TL positions residues of loop A−B in a configuration conducive to interactions with residues of the barrel.

The most red-shifted and therefore the most exposed (flexible) loops are B−C and E−F, each of which has two Gly residues (Figure 6). The small size and less restricted dihedral angles of Gly residues are likely to confer flexibility and exposure to loops B−C and E−F. Loops C−D and D−E each have a single Gly residue. Loops A−B and F−G do not have Gly and on average are blue-shifted with respect to the other loops. However, loop F−G exhibits changes in relative exposure with ligand binding. We have shown previously that conformational transitions promote relative exposure of loop F−G in holo-TL relative to apo-TL (*12*). The functional implication of these dynamic alterations is that loop F−G in TL may be a receptor recognition site for the holo-protein. It has been hypothesized that loop F−G, at the closed end of the lipocalin fold, may be part of a receptor recognition site for lipocalins in general (*2*).

The average λ_{max} values for the β-strands are revealing about their relative positions in the TL barrel (Figure 6). Strand D residues have an average red-shifted λ_{max} indicative of a solvent-exposed position at the edge of the cavity. In contrast, strand H features a blue-shifted average λ_{max} that reflects a relatively buried environment. The most likely explanation for the relatively buried environment of strand H is an interaction with α-helix, a conserved relationship among members of the lipocalin superfamily. The homology model is consistent with such a position.

The emission peaks on the putative β-strand I are red-shifted and feature very slight differences from one residue to the next. Therefore, we cannot assign unambiguously strand I from our data. The lack of change in the alternating pattern suggests that these residues are relatively exposed to solvent and are not part of the β-barrel.

To detect α-helical structure by SDTF, at least one side of the helix must interact with the β-barrel. If the residues

Table 1: Residues of the β-Barrel That Are Found To Be Less than 4 Å from Palmitate Atoms in β-Lactoglobulin and Their Corresponding Residues in the TL Model

BLG (28,29)	L46	I56	K60	K69	I71	I84	V92	V94	L103	F105	M107
TL	L41	A51	M55	V64	A66	A79	A86	I88	F99	C101	G103

of the α-helix were totally exposed to solvent, the emission peaks would show a marked red shift without any discernible periodicity. The interaction of α-helix and β-barrel residues results in blue-shifted emission peaks that exhibit a periodicity close to 3.6 for the α-helical residues. For example, F130 shows an extremely blue-shifted emission peak indicating the side chain resides in a hydrophobic environment. There is a periodicity close to 3.6 for this region (Figure 2). The α-helix in β-lactoglobulin is known to have a similar hydrophobic interaction with the β-strand H. Our data in combination with energy minimization data produce a model in which the α-helix of TL interfaces strand H. This interpretation provides an explanation for the incomplete alternating periodicity in strand H by SDTF. Alternate strand H residues are surrounded by residues of the α-helix instead of solvent. In this model, residues V113 and L115 of the β-strand H interdigitate with residue F130 of the α-helical segment.

The combination of SDTF and energy minimization data indicates that K114 of strand H is oriented facing the interior of the β-barrel near the side chain of S35. The residue in β-lactoglobulin that corresponds to K114 in TL is Q120, a polar residue. The distortion of secondary packing observed with the tryptophan mutation at S35 may be explained by the absence of the predicted hydrogen bond joining the side chains of K114 and S35.

β-Bulges are produced in β-strands by additional residues that interrupt the alternating pattern of exposure to solvent. This results in the alteration of the hydrogen bonding pattern and a change in the direction of the projecting side chains in a localized area. In RBP, the β-bulges at M27−A28 and V107−D108 permit the sharing of β-strands A and F, respectively, between the two orthogonal sheets (*26*). Similarly, there are β-bulges in β-lactoglobulin at L22−A23 in β-strand A and L95−D96 in β-strand F (PDB ID: 1B0o). Our data indicate β-bulges in TL also occur in β-strands A (K20−A21) and F (I89−R90) and may allow sharing of the strands between the two orthogonal sheets while maintaining the hydrogen bonding pattern.

The position of the conserved disulfide in TL is consistent with the general motif published for other lipocalins (*31*). The disulfide bond links strand D of the β-barrel to the C terminus or tail of the lipocalin. A markedly decreased fluorescent yield may result if tryptophan approaches a disulfide bond (*32*). Mutants T37W and L56W of TL have unusually low fluorescent yields. Our TL model, which takes into account the conserved position of the disulfide, places both T37 (CG2) and L56 (CD2) 3.9 Å from the disulfide bond.

The introduction of perturbations in secondary structure by specific tryptophan mutants is informative about important local side chain interactions. For example, substitution-induced distortion is evident with V113W. In our model, V113 (CG2) on strand H lies only 4.0 Å from F130 (CE1) on the α-helix. The substitution of valine with the bulkier residue, tryptophan, apparently perturbs the interaction of strand H with the α-helix. The interaction of strand H with the α-helix is highly conserved in the lipocalins and may have a role in stabilization of structure.

In one case, G117W, tryptophan substitution affected the state of the expressed protein. We were unable to purify protein expressed in a dimeric state; perhaps it exists in a higher oligomeric state. G117 in TL corresponds to G121 in porcine odorant binding protein, which putatively has a role in restricting domain swapping via the main α-helix (*33*). It is intriguing to speculate that the tryptophan substitution in TL may free the α-helix from its interaction with the β-barrel and permit sharing with the domain of another TL molecule to form oligomers.

Our model provides a useful comparison of the position of ligands in TL versus β-lactoglobulin. It has been previously shown that fatty acids bind within the cavities of TL and β-lactoglobulin (*7, 12, 34, 35*). Table 1 shows the barrel residues in β-lactoglobulin that are assigned within 4 Å of palmitate atoms. Also shown are the corresponding residues derived from our model for TL. Previous work corroborates that C101 and G103 are in the binding site because tryptophans substituted at these positions were quenched by a nitroxide-labeled lauric acid analogue (*12*). Furthermore, it has been shown that the functional dimensions of the TL cavity are greater than those of β-lactoglobulin (*30*). While both proteins bind fatty acids, only TL binds the fluorescent-labeled fatty acid DAUDA (*16, 29*). Narayan and Berliner's data (*36*) with spin-labeled fatty acids in combination with crystallographic data (*34, 35*) suggest that the nitroxide side chain may be sterically restricted from entering the cavity of β-lactoglobulin. The differences in binding affinities of 5-doxyl stearic acid and 12-doxyl stearic acid for β-lactoglobulin, corresponding free energy difference of ∼3600 cal/mol, are accounted for by the contribution of six CH_2 [600 cal/mol each (*37*)]. These data in concert with the diminished binding of 16-doxyl stearic acid in β-lactoglobulin suggest that the penetration of the hydrocarbon tail of fatty acids is limited by the position of the relatively bulky nitroxide. In contrast, TL binds 5-, 12-, and 16-doxyl stearic acids.

Residues I56, I71, I84, and V92 in β-lactoglobulin (Table 1, Figure 7A) are part of the binding site (*34, 35*). The corresponding positions in TL are A51, A66, A79, and A86 (Table 1, Figure 7B). If the model is correct, then the smaller alanine side chains in TL could potentially result in a more capacious binding site compared to β-lactoglobulin. EPR-monitored ligand binding experiments with 16-DSA allowed us to test this hypothesis. Substitution with the bulkier residue tryptophan resulted in less binding to the proteins in all cases except position 79. Since the precise orientation of the ligand in the binding pocket is not known, it is possible that the ligand is positioned so that A79 does not restrict the entrance in the cavity (Figure 7).

The model of TL corroborates data regarding the contribution of the carboxyl group to ligand binding. Charged residues K60 and K69 in β-lactoglobulin are replaced by M55 and V64 in TL. An electrostatic interaction occurs

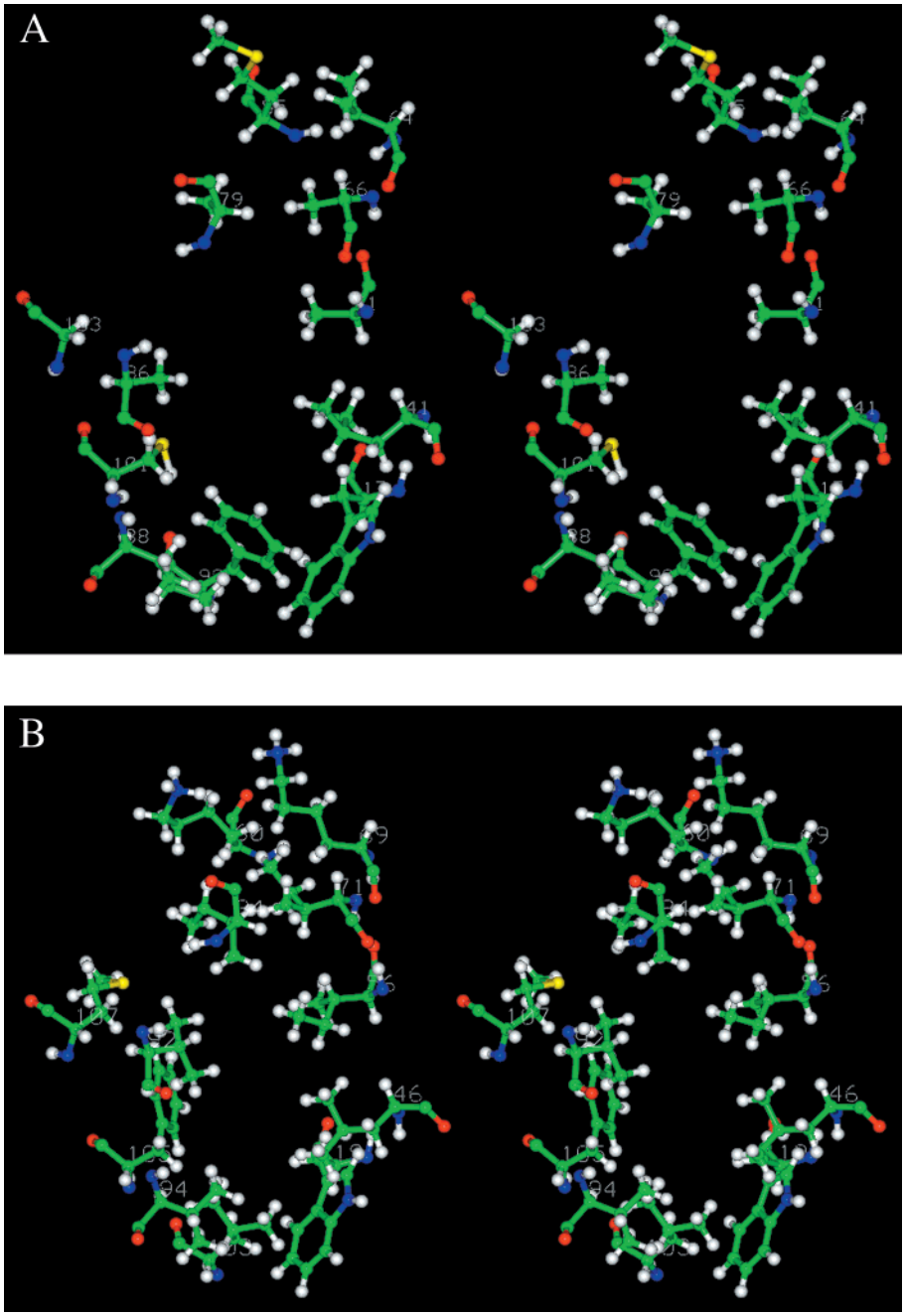

FIGURE 7: (A) Stereo diagram of the TL binding cavity in a stick and ball presentation. Colors represent atoms as follows: white, hydrogen; green, carbon; red, oxygen; blue, nitrogen; yellow, sulfur. (B) Stereo diagram of the β-lactoglobulin binding cavity with the same color scheme as shown in panel A.

between the carboxylic moiety of palmitic acid in β-lactoglobulin and two lysine residues, K60 and K69. Since M55 and V64 are not charged, such an electrostatic interaction would not be possible in TL. This interpretation of the model is supported by experimental data that suggest that the C1 groups of fatty acids are highly mobile in TL and do not interact with side chain residues (*7, 30*).

The predictions of our model were further tested with the mutant M22W/K114C. The red shift produced by the interaction of tryptophan in mutant M22W with K114 was eliminated by the cysteine substitution (Figure 2), confirming the close proximity of M22 to K114.

In summary, SDTF may be applied to the investigation of the solution structure of entire proteins. Loops, α-helices, β-strands, and β-bulges can be identified from the analysis of the maxima of the tryptophan emission spectra. The average λ_{max} of loops and β-strands reveal information about their relative positions and domain interactions. The information reflects the dynamic state of the protein. SDTF is ideally suited for proteins with a single or very few tryptophan residues. One must be cautious in the interpretation of function in a mutant in which an essential tryptophan has been altered. In the case of TL, the use of another aromatic residue (Tyr) restored functional qualities (*16*). Here, the structural features elucidated by SDTF provide a model for TL that is quite consistent with lipocalins in general, but also explain key functional differences between family members. The broad promiscuity of TL is conferred by the short E−F hairpin, flexible loop B−C, and a capacious binding site. SDTF is a useful tool to determine solution

structure for select proteins, especially in the absence of NMR or crystallographic data.

ACKNOWLEDGMENT

We thank Dr. Wayne Hubbell and Christian Altenbach for their guidance in operating the Octane Silicon Graphics Workstation for modeling of TL and access to the Varian 109 spectrometer.

SUPPORTING INFORMATION AVAILABLE

The variation of λ_{max} in β-strands of tear lipocalin and conformational changes in the G strand as a result of energy minimization (2 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Glasgow, B. J., Abduragimov, A. R., Farahbakash, Z., Faull, K. F., and Hubbell, W. L. (1995) *Curr. Eye Res. 14*, 363−372.
2. Flower, D. R. (1996) *Biochem. J. 318*, 1−14.
3. Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C., Findlay, J. B., Sivaprasadarao, R., Jones, T. A., Newcomer, M. E., and Kraulis, P. J. (1986) *Nature 324*, 383−385.
4. Molinari, H., Ragona, L., Varani, L., Musco, G., Consonni, R., Zetta, L., and Monaco, H. L. (1996) *FEBS Lett. 381*, 237−243.
5. Monaco, H. L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L., and Eliopoulos, E. E. (1987) *J. Mol. Biol. 197*, 695−706.
6. Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., and Glasgow, B. J. (1997) *Biochem. Biophys. Res. Commun. 239*, 191−196.
7. Glasgow, B. J., Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., Altenbach, C., and Hubbell, W. L. (1999) *Biochemistry 38*, 13707−13716.
8. Mansoor, S. E., Mchaourab, H. S., and Farrens, D. L. (1999) *Biochemistry 38*, 16383−16393.
9. O'Neil, K. T., Wolfe, H. R., Jr., Erickson-Viitanen, S., and Degrado, W. F. (1987) *Science 236*, 1454−1456.
10. Kintrup, M., Schubert, P., Kunz, M., Chabbert, M., Alberti, P., Bombarda, E., Schneider, S., and Hillen, W. (2000) *Eur. J. Biochem. 267*, 821−829.
11. Lakowicz, J. R. (1999) *Principles of Fluorescence Spectroscopy*, Second edition, Plenum Press, New York.
12. Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., and Glasgow B. J. (2000) *Protein Sci. 9*, 325−331.
13. Glasgow, B. J., Heinzmann, C., Kojis, T., Sparkes, R. S., Mohandas, T., and Bateman, J. B. (1993) *Curr. Eye Res. 11*, 1019−1023.
14. Redl, B., Holzfeind, P., and Lottspeich, F. (1992) *J. Biol. Chem. 267*, 20282−20287.
15. Glasgow, B. J. (1995) *Graefe's Arch. Clin. Exp. Ophthalmol. 233*, 513−522.
16. Gasymov, O. K., Abduragimov, A. R., Yusifov, T. N., and Glasgow, B. J. (1999) *Biochim. Biophys. Acta 1433*, 307−320.
17. Cormack, B. (1987) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Ed.) Suppl. 15, pp 8.5.1−8.5.9, Greene Publishing Associates and Wiley-Interscience, New York.
18. Marston, F. A. O. (1987) A Practical Approach. in *DNA Cloning* (Glover, D. M., Ed.) Vol. III, p 62, IRL Press, Oxford, England.
19. Bozimowski, D., Artiss, J. D., and Zak, B. (1985) *J. Clin. Chem. Clin. Biochem. 23*, 683−689.
20. Koteiche, H. A. and Mchaourab, H. S. (1999) *J. Mol. Biol. 294*, 561−577.
21. Hubbell, W. L., Froncisz, W., and Hyde J. S. (1987) *Rev. Sci. Instrum. 58*, 1879−1886.
22. Berengian, A. R., Bova, M. P., and Mchaourab, H. S. (1997) *Biochemistry 36*, 9951−9957.
23. Eftink, M. R., and Ghiron, C. A. (1976) *Biochemistry 15*, 672−680.
24. Burstein, E. A., Vedenkina, N. S., and Ivkova, M. N. (1973) *Photochem. Photobiol. 18*, 263−279.
25. Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., Baker, E. N., and Jameson, G. B. (1998) *Biochemistry 37*, 14014−14023.
26. Huber, R., Schneider, M., Mayr, I., Muller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H., and Kayser, H. (1987) *J. Mol. Biol. 198*, 499−513.
27. Cowan, S. W., Newcomer, M. E., and Jones, T. A. (1990) *Proteins: Struct., Funct., Genet. 8*, 44−61.
28. Callis, P. R., and Burgess, B. K. (1997) *J. Phys. Chem. 101*, 9429−9432.
29. Kennedy, M. W., Brass, A., McCruden, A. B., Price, N. C., Kelly, S. M., and Cooper, A. (1995) *Biochemistry 34*, 6700−6710.
30. Abduragimov, A. R., Gasymov, O. K., Yusifov, T. N., and Glasgow, B. J. (2000) *Curr. Eye Res. 4*, 824−832.
31. Glasgow, B. J., Abduragimov, A. R., Yusifov, T. N., Gasymov, O. K., Horwitz, J., Hubbell, W. L., and Faull, K. F. (1998) *Biochemistry 37*, 2215−2225.
32. Permyakov, E. A. (1993) *Luminescent spectroscopy of proteins*, p 48, CRC Press, Boca Raton, FL.
33. Spinelli, S., Ramoni, R., Grolli, S., Bonicel, J., Cambillau, C., and Tegoni, M. (1998) *Biochemistry 37*, 7913−7918.
34. Wu, S. Y., Perez, M. D., Puyol, P., and Sawyer, L. (1999) *J. Biol. Chem. 274*, 170−174.
35. Qin, B. Y., Creamer, L. K., Baker, E. N., and Jameson, B. B. (1998) *FEBS Lett. 438*, 272−278.
36. Narayan, M., and Berliner, L. J. (1997) *Biochemistry 36*, 1906−1911.
37. Tanford, C. (1980) *The Hydrophobic Effect*, p154, Wiley-Interscience, New York.

BI0110342