Three-Dimensional Solution Structure of *Escherichia coli* Periplasmic Cyclophilin[†,‡]

Robert T. Clubb,[§] Stephen B. Ferguson,[∥] Christopher T. Walsh, and Gerhard Wagner*

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115

Received October 8, 1993; Revised Manuscript Received December 16, 1993[⊗]

ABSTRACT: The solution structure of the periplasmic cyclophilin type cis–trans peptidyl–prolyl isomerase from *Escherichia coli* (167 residues, MW > 18.200) has been determined using multidimensional heteronuclear NMR spectroscopy and distance geometry calculations. The structure determination is based on a total of 1720 NMR-derived restraints (1566 distance and 101 ϕ and 53 χ_1 torsion angle restraints). Twelve distance geometry structures were calculated, and the average root-mean-square (rms) deviation about the mean backbone coordinate positions is 0.84 ± 0.18 Å for the backbone atoms of residues 5–165 of the ensemble. The three-dimensional structure of *E. coli* cyclophilin consists of an eight-stranded antiparallel β-sheet barrel capped by α-helices. The average coordinates of the backbone atoms of the core residues of *E. coli* cyclophilin have an rms deviation of 1.44 Å, with conserved regions in the crystal structure of unligated human T cell cyclophilin [Ke, H. (1992) *J. Mol. Biol. 228*, 539–550]. Four regions proximal to the active site differ substantially and may determine protein substrate specificity, sensitivity to cyclosporin A, and the composite drug:protein surface required to inhibit calcineurin. A residue essential for isomerase activity in human T cell cyclophilin (His126) is replaced by Tyr122 in *E. coli* cyclophilin without affecting enzymatic activity.

Folding and sorting of proteins *in vivo* are believed to require a set of enzymes and chaperones that accelerate the rates of folding and increase the yield of native protein structures. Included among these enzymes is *Escherichia coli* cyclophilin (eCyP),[1] a peptidyl–prolyl isomerase (PPIase), which catalyzes the interconversion of cis–trans Xaa–Pro bonds in peptides and proteins. Two structurally distinct branches of PPIases have been identified: the cyclophilins and the FK506-binding proteins (FKBP). Both are distributed ubiquitously in nature and in subcellular compartments where proteins fold and refold (Schreiber, 1991; Walsh et al., 1992). *In vivo*, cyclophilins are required for the proper folding of Rh 1 rhodopsin in *Drosophila* eye tissue (Ondek et al., 1992), the folding and trafficking of transferrin in hepatocyte G2 cells (Lodish & Kong, 1991), increases in the rates of collagen triple-helix formation (Steinmann et al., 1991), and interact with the gag protein from human immunodeficiency virus type 1 (Luban et al., 1993). *In vitro*, cyclophilins increase the folding rates of ribonuclease T1 (Kiefhaber et al., 1990), chymotrypsin inhibitor (CI_2) (Jackson & Fersht, 1991), carbonic anhydrase II (Freskgard et al., 1992), IgG light chain and Fab fragment (Lang et al., 1987), and the rate of interconversion between the cis and trans forms of the hormone calcitonin (Kern et al., 1993).

Cyclophilin and FKBP are of interest for a separate pharmacological reason: both are targets for the blockade of graft rejection and T cell activation by the immunosuppressant drugs cyclosporin A (CsA) (Handshumacher et al., 1984; Harding et al., 1986) and FK506 (Harding et al., 1989; Siekierka et al., 1989), respectively. Immunosuppression is thought to occur in the T cell when the complex of CyP:CsA or FKBP:FK506 binds to and inhibits the calcium- and calmodulin-dependent protein serine phosphatase, calcineurin (CN) (Liu et al., 1991b). Inhibition may prevent the dephosphorylation and subsequent nuclear translocation of the transcription factor NFAT, a protein required for T cell activation (Emmel et al., 1989; Flanagan et al., 1991). Cyclosporin A and FK506 also inhibit the PPIase activity of their respective target proteins, but PPIase and immunosuppressive activity appear to be unrelated (Zydowsky et al., 1992).

The dual function of cyclophilins in both immunosuppression and protein folding has resulted in the intense structural study of human T cell cyclophilin isoform A (hCyP A). Crystal structures have been reported for unligated hCyP A (Ke et al., 1991; Ke, 1992), for a hCyP A:CsA complex (Pflügl et al., 1993), and for complexes of hCyP A with tetrapeptide (Kallen et al., 1991; Kallen & Walkinshaw, 1992) and dipeptide substrates (Ke et al., 1993). NMR investigations have identified hCyP A:CsA interactions (Neri et al., 1991; Spitzfaden et al., 1992; Fesik et al., 1992) and the conformation of CsA bound to hCyP A in solution (Fesik et al., 1991; Weber et al., 1991; Hsu & Armitage, 1992). The secondary structure

[†] This research was supported by NIH Grant GM 47467 and by NIH NSRA Grant 14252 to S.B.F.

[‡] Protein coordinates have been deposited in the Brookhaven Protein Data Bank under file name 1CLH.

* Address correspondence to this author.

[§] Current address: NIDDK, National Institutes of Health, Bldg. 5, Bethesda, MD 20892.

[∥] Current address: Hoffmann-La Roche, Nutley, NJ 07110-1199.

[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1994.

[1] Abbreviations: AC, *N*-acetyl; AMC, amidomethylcoumarin; CN, calcineurin; CsA, cyclosporin A; eCyP, *Escherichia coli* cyclophilin; hCyP A, human T cell cyclophilin isoform A; DQF-COSY, two-dimensional double quantum filtered correlation spectroscopy; HMQC, heteronuclear multiple quantum coherence; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; PPIase, cis–trans peptidyl–prolyl isomerase; TOCSY, total correlation spectroscopy; rms, root-mean-square; <rmsd>, average root-mean-square deviation; ppm, parts per million.

0006-2960/94/0433-2761$04.50/0

of hCyP A (Neri et al., 1991; Wüthrich et al., 1991) and the structure of the hCyP A:CsA complex (Thériault et al., 1992) in solution have been determined as well.

In contrast to the numerous structural studies of hCyP A, the three-dimensional structure of any other member of the cyclophilin family remains unknown. There are at least 27 cyclophilins in both prokaryotic and eukaryotic organisms with varying affinities for CsA and potentially different substrate preferences. In general, eukaryotic cyclophilins share greater than 60% sequence identity with hCyP A, while the three bacterial cyclophilins identified thus far are more distantly related (Trandinh et al., 1992). Cytoplasmic and periplasmic cyclophilins have been cloned, sequenced, and purified from *E. coli* (Liu & Walsh., 1990; Hayano et al., 1991; Compton et al., 1992) and are approximately 50% identical in their amino acid sequences (Hayano et al., 1991). However, cytoplasmic and periplasmic *E. coli* cyclophilins share only 36 and 34% identity with hCyP A, respectively (Trandinh et al., 1992). Variations between the primary sequences of eCyP and hCyP A are manifested in eCyP's significantly reduced affinity for the drug CsA relative to hCyP A (hCyP A: IC_{50} = 6 nM; periplasmic eCyP: IC_{50} = 3000 nM) (Handschumacher et al., 1984; Harding et al., 1986; Liu & Walsh, 1990). A single-point mutation in periplasmic eCyP (Phe112Trp), however, results in a 35-fold increase in affinity for CsA; the periplasmic Phe112Trp eCyP:CsA complex can also bind and inhibit calcineurin (CN) (Fejzo et al., 1993). Therefore, even the most distantly related cyclophilins, when complexed with CsA, possess all of the structural and dynamic determinants necessary to inhibit calcineurin.

Multidimensional heteronuclear magnetic resonance spectroscopy has been utilized to determine the solution structure of periplasmic *E. coli* cyclophilin. The structure elucidates the degree of conservation in the cyclophilin family and provides insights into the mechanism of peptidyl–prolyl isomerization and the molecular details of cyclophilin-mediated immunosuppression. We find that eCyP and hCyP A share a common global fold but differ substantially in their respective active sites. Major variations occur in four surface loops near the active site. The effects of these variations on CsA binding, CN inhibition, and PPIase substrate specificity are discussed.

MATERIALS AND METHODS

Sample Preparation. Uniformly ^{15}N and $^{13}C,^{15}N$ isotopically labeled protein was expressed and purified as described previously (Clubb et al., 1993). Samples for NMR studies contained 1–1.5 mM protein in 50 mM phosphate buffer (pH 6.2) dissolved in either 95% H_2O/5% ^2H_2O or 99.996% ^2H_2O. pH meter readings in ^2H_2O solutions were not corrected for isotope effects.

NMR Spectroscopy. All spectra were recorded at 298 K with Bruker AMX500 and 600-MHz spectrometers. The ^1H–^{15}N TOCSY-HMQC (Marion et al., 1989), ^1H–^{15}N NOESY-HMQC (Fesik & Zuiderweg, 1988; Zuiderweg & Fesik, 1989; Marion et al., 1989), DQF-COSY (Piantini et al., 1984; Shaka & Freeman, 1983; Rance et al., 1984), HCCH-COSY (Kay et al., 1990b; Bax et al., 1990a), and ^1H–^{13}C heteronuclear single-quantum coherence (HSQC) spectra (Bodenhausen & Ruben, 1980) were acquired as described previously (Clubb et al., 1993). The HNHB spectrum (Archer et al., 1991) was acquired with 123, 112, and 512 real data points in $t_1(^{15}N)$, $t_2(^1H)$, and $t_3(^1H)$, respectively. The final zero-filled data matrix contained 128, 256, and 256 points in ω_1, ω_2, and ω_3, respectively. The HNHA spectrum (Vuister & Bax, 1993) was acquired without gradient pulses with 69, 96, and 512 real data points in $t_1(^{15}N)$, $t_2(^1H)$, and $t_3(^1H)$, respectively. The final zero-filled data matrix contained 128, 128, and 256 points in ω_1, ω_2, and ω_3, respectively. Totals of 16 and 32 transients were recorded per increment in the HNHB and HNHA experiments, respectively. Both the HNHA and HNHB experiments were performed at 600 MHz with continuous wave presaturation during the 1-s recycle delay. The HNHB experiment used sweep widths of 2016, 7862, and 9615 Hz in ω_1, ω_2, and ω_3, respectively. The HNHA experiment used sweep widths of 2016, 6410, and 9615 Hz in ω_1, ω_2, and ω_3, respectively. Two 3D ^1H–^{13}C NOESY-HMQC spectra (Fesik & Zuiderweg, 1988; Zuiderweg & Fesik, 1989; Marion et al., 1989a) with mixing times of 60 and 100 ms and a single ^1H–^{15}N NOESY-HMQC spectrum with a mixing time of 60 ms were acquired as well.

Side-Chain Assignments. In order to obtain a sufficient number of restraints to define the tertiary structure of eCyP, it was necessary to assign the proton and carbon resonances of the side chains (R. T. Clubb, P. Zhang, J. W. Peng, J. Fejzo, S. B. Ferguson, C. T. Walsh, and G. Wagner, manuscript in preparation). Aliphatic nuclei were assigned primarily through analysis of an HCCH-COSY spectrum (Kay et al., 1990b; Bax et al., 1990a) of eCyP. ^1H–^{15}N and ^1H–^{13}C NOESY-HMQC spectra (Fesik & Zuiderweg, 1988; Zuiderweg & Fesik, 1989; Marion et al., 1989a) were also utilized to confirm and obtain additional assignments, correlating the previously assigned $^1H^N$ and $^1H^\alpha$ protons with their respective intraresidue side-chain nuclei ($^1H^\alpha$, $^1H^\beta$, $^1H^\gamma$, etc.). An HCCH-TOCSY spectrum (Fesik et al., 1990; Bax et al., 1990b) of eCyP was not extensively utilized, due to its low sensitivity.

Assignment of aromatic resonances involved the analysis of homonuclear DQF-COSY, TOCSY, and NOESY and heteronuclear ^1H–^{13}C correlation spectra of eCyP dissolved in ^2H_2O. Aromatic spin systems were identified, and each system was then located in the primary sequence by observation of NOEs between the $^1H^\delta$ proton of the aromatic ring and the previously assigned $^1H^\beta$ and $^1H^\alpha$ nuclei. A ^1H–^{15}N NOESY-HMQC spectrum was also used to confirm the aromatic assignments, by observation of an NOE between backbone $^1H^N$ and the $^1H^\delta$ nuclei of the attached ring system.

ϕ Torsion Angle Restraints. The ϕ torsion angles of eCyP were determined by measuring the $^3J_{HN\alpha}$ coupling constant from an HNHA spectrum of eCyP (Vuister & Bax, 1993). This quantitative approach allowed 76 $^3J_{HN\alpha}$ coupling constants to be measured. Diagonal and cross-peaks were integrated in the ^1H–^1H slice of maximum intensity. Measured $^3J_{HN\alpha}$ coupling constants were increased by 21% (empirically determined) to compensate for the effects of $^1H^\alpha$ relaxation. Residues known to participate in sheets or helices on the basis of amide exchange rates and NOE patterns (Clubb et al., 1993) were used to determine the deviation between experimental and actual $^3J_{HN\alpha}$ coupling constants. Interestingly, this 21% correction factor is greater than the 11% increase suggested for proteins of similar molecular mass. The 11% increase was derived theoretically assuming selective spin-lattice relaxation times (T_{1sel}) equal to 100 ms for the H^α proton (Vuister & Bax, 1993). In order to accommodate any potential errors in our adjusted $^3J_{HN\alpha}$ values, all coupling constants were loosely converted into ϕ angle restraints. Corrected $^3J_{HN\alpha}$ values were converted to the following ϕ angle restraints: >9 Hz, −80° to −160°; >8 Hz, −70° to −170°; 7 to 8 Hz, −180° to −60°; and <5.5 Hz, −20° to −90°. In addition, the ϕ torsion angle was restrained to negative values (−20° to −180°) for 25 residues whose amide protons gave a

more intense NOE to the preceding residue's $^1H^\alpha$ proton compared with the NOE to its own $^1H^\alpha$ proton.

χ_1 Torsion Angle Restraints and Stereospecific Assignments. Stereospecific assignments of $^1H^\beta$ protons and χ_1 torsion angle restraints were determined by qualitative estimates of the $^3J_{\alpha\beta}$ and $^3J_{N\beta}$ coupling constants and the relative intensities of intraresidue NOEs (Wagner et al., 1987; Hyberts et al., 1987; Montelione et al., 1989; Clore et al., 1991).

$^3J_{N\beta}$ coupling constants were estimated for 24 residues from an HNHB spectrum (Archer et al., 1991). $^3J_{\alpha\beta}$ coupling constant estimates were obtained for 25 residues through analysis of a DQF-COSY spectrum for resolvable H^α–H^β cross-peaks. The H^N–H^β cross-peak intensity in an ^{15}N-edited TOCSY experiment (30-ms mixing time) was also utilized to estimate eight additional $^3J_{\alpha\beta}$ coupling constants (Clore et al., 1991). Coupling constant information was then complemented with short mixing time (60 ms) NOESY spectra, utilized to determine the relative intensities of intraresidue cross-peaks. The methodology was successful in restricting the χ_1 angles of 49 of 93 non-proline residues with β-methylene protons in eCyP. If one considers that only 68 of 93 residues in eCyP have nondegenerate and/or two assigned β-methylene protons, 72% (49 of 68) of the χ_1 angles of these residues could be restricted. For a total of 15 residues, stereospecific assignments of β-methylene protons were obtained, and χ_1 was restricted to a single rotamer (–60° ± 30°, 60° ± 30°, 180° ± 30°), while 34 residues were restricted to either 150° to –30° or –90° to 90°.

The χ_1 angles and stereospecific assignments of the γ-methyls of four valines were determined from intraresidue NOEs and qualitative estimates of $^3J_{\alpha\beta}$ (Zuiderweg et al., 1985).

Hydrogen-Bond Restraints. Hydrogen-bond restraints were imposed for 54 backbone $^1H^N$ protons that exchanged slowly with the solvent and exhibited characteristic medium- and long-range NOEs (Wüthrich, 1986). Slowly exchanging H^N protons were identified from a series of 2D HSQC spectra (Bodenhausen & Ruben, 1980) recorded after dissolving a sample of lyophilized eCyP in ^2H_2O. Hydrogen-bond restraints (1.5–2.3 Å $^1H^N$–CO and 2.4–3.3 Å N–CO for each hydrogen bond) were introduced *after* the initial set of structure calculations.

Assignment and Evaluation of NOE Cross-Peaks. NOEs were assigned from 3D heteronuclear edited NOESY spectra and a homonuclear 2D NOESY spectrum, each recorded with mixing times of 60 and 100 ms. When NOEs were derived from 3D spectra they were quantified in the ^1H–^1H plane of maximum intensity. Cross-peak intensities, as determined by the number of contour levels, were converted to loose distance restraints to avoid a bias due to the effects of spin diffusion. NOEs were classified as strong, medium, weak, and very weak, corresponding to interproton distance restraints of 1.8–3.0, 1.8–4.0, 1.8–4.5, and 1.8–5.0 Å, respectively. Upper distance restraints involving non-stereospecifically assigned methylene, aromatic, and methyl protons were replaced by the appropriate pseudoatoms (Wüthrich et al., 1983).

Initial data interpretation was facilitated by a model of the conformation of eCyP, constructed from the results of a preliminary analysis of the NMR data (Clubb et al., 1993) and sequence and secondary structure homology to human T cell cyclophilin (hCyP A) (Havel & Snow, 1991). The model accurately predicted NOEs between the main-chain proton atoms of eCyP, but failed to correctly predict side-chain–side-chain NOEs. Reference to the model structure allowed ~300 distance restraints to be extracted, and a structure of eCyP consistent with this data was calculated. A series of calculations with additional restraints incorporated at each successive stage was then conducted *without* reference to the human structure. The list of constraints used is provided in the supplementary material.

Structure Calculations. Structures were computed using distance geometry methods (Havel, 1991) in the NMR refine module of the INSIGHTII package (BIOSYM Technologies, San Diego). This procedure includes bound smoothing in which triangular inequality distance limits are determined. Subsequently, random distance matrices that satisfy the inequality limits are generated by random metrization, and coordinates with appropriate distances are obtained by angular embedding and majorization procedures in 4D. Structures are then optimized in 3D by simulated annealing and conjugate gradient minimization procedures described by Havel (1991). Converged structures are then subjected to 100 iterations of steepest descent and 1000 iterations of conjugate gradient minimization to reduce repulsive nonbonded contacts. Minimization was performed using the INSIGHTII package with the CVFF force field (BIOSYM) and included potential energy terms for the distance and dihedral angle restraints. In the minimization all force constants were scaled to 1.0, and force constants of 10 and 30 kcal Å^{-2} mol^{-1} were used to enforce distance and dihedral angle restraints, respectively.

Calculations were performed on a Silicon Graphics 4D/480 computer. The amino acid residues are numbered as described previously (Clubb et al., 1993). The N-terminal alanine, designated Ala1′, was not restrained by the NMR data and was therefore omitted from the structure calculations. Excluding the bound smoothing routine, which required ca. 22 h for the ensemble, approximately 20 h of CPU time was required to generate a single structure. An iterative approach to the structure calculations was used by carrying out a series of calculations with additional restraints incorporated at each successive stage.

The total of 1720 experimental restraints was used in the final structure calculations. These comprised 1458 interproton distance, 108 hydrogen-bonding, and 101 ϕ and 53 χ_1 torsion angle restraints. The interproton distance restraints consisted of 252 intraresidue, 483 sequential, 149 medium-range (less than five residues apart), and 574 long-range restraints (five or more residues apart). The NOE spectra have not yet been analyzed exhaustively, and assignment of a significantly larger number of NOEs should lead to an increased precision of the structure.

RESULTS

Converged Structures. A total of 34 structures was calculated, 15 of these possessed no residual distance violation greater than 0.35 Å, and they were energy-minimized using the program DISCOVER. After minimization, 12 structures possessed no residual distance violation greater than 0.5 Å; restraint and structural statistics of this ensemble are presented in Table 1. These 12 structures were submitted to the Brookhaven Protein Data Bank. The histogram of the rms deviations of the backbone nuclei (C, C^α, N) to the average structure of the ensemble, shown in Figure 1A, indicates that the NMR data define the structure of eCyP reasonably well. For the majority of residues, the average rms deviations from the average structure are below 0.8 Å (Figure 1A). Six regions of nonregular secondary structure show deviations from the average structure of >1.5 Å; these are the N- and C-termini (residues Ala1′-Asp4 and Pro166) and residues Gly80, Thr93–

Table 1: Restraint Statistics and Atomic rms Deviations[a]

A. Restraint Statistics

total	distance	rms deviations from experimental distance restraints (Å)
1566	all	0.0264 ± 0.0051
1458	NOE	0.0274 ± 0.0053
483	interresidue ($\|i-j\| = 1$)	0.0227 ± 0.0107
149	interresidue ($1 < \|i-j\| \leq 4$)	0.0260 ± 0.0108
574	interresidue ($\|i-j\| > 4$)	0.0350 ± 0.0082
252	intraresidue	0.0100 ± 0.0062
108	H-bonds	0.0055 ± 0.0043

total	torsion	rms deviations from experimental dihedral restraints (deg)
154	all	0.5499 ± 0.2305
101	ϕ angle	0.5918 ± 0.2918
53	χ angle	0.4086 ± 0.5166

B. Structural Statistics

Cartesian coordinate rms deviations (Å)	backbone (N, C^α, C')
⟨DG⟩ vs ⟨$\overline{DG}$⟩ all[b]	0.84 (± 0.18)
⟨DG⟩ vs ⟨$\overline{DG}$⟩ selected[c]	0.65 (± 0.15)
⟨$\overline{DG}$⟩ vs hCyP X-ray[d]	1.44

[a] ⟨DG⟩ represents the 12 NMR structures and ⟨$\overline{DG}$⟩ represents the mean atomic structure obtained by averaging the least-squares fitted atomic coordinates of the 12 structures. [b] rms deviations of the 12 structures and th emean structure for backbone heavy atoms (N, C^α, C') of residues 5–165. [c] rms deviations of the 12 structures and the mean structure of the well-defined protein core (residues 5–77, 83–92, 99–110, 121–144, 151–165. [d] rms deviations between the underlined residues of Figure 5 in the mean atomic solution structure and the X-ray structure of unligated hCyP A (Ke, 1992).

Asp95, Gln117–Asp119, and Pro148. A histogram of NOE distance restraints vs sequence number is presented in Figure 1B and clearly shows an inverse correlation with the rms deviations of Figure 1A.

Another measure of the precision of the ensemble of structures is the angular order parameter, S^{angle} (Hyberts et al., 1992). Figure 1C,D shows plots of the angular order parameters of the ϕ and ψ dihedral angles, respectively. Precisely determined dihedral angles are characterized by S^{angle} values near unity. In general, the residues with higher order parameters exhibit lower rms deviations from the average structure and are defined by a larger number of NOE restraints. The quality of the structures can be best visualized by inspection of Figure 2, which displays two views of the best fit superposition of the backbone atoms of residues Pro5–Leu165. The average rms deviation of these residues about their mean backbone coordinate positions is 0.84 ± 0.18 Å.

In general, all structures exhibit good covalent geometry and nonbonded contacts. The Ramachandran plot (Ramachandran & Sasisekharan, 1968) of ϕ,ψ angles for residues Pro5–Leu165 for the 12 structures in the ensemble is shown in Figure 3A. The majority of ϕ,ψ values are within energetically favorable regions. If only those residues with angular order parameters larger than 0.9 are plotted, the number of ϕ,ψ values located in energetically unfavorable regions dramatically decreases (Figure 3B). Therefore, all residues whose ϕ,ψ angles have been determined with high precision are energetically reasonable. None of the residues with well-defined ϕ and ψ angles ($S^{angle}(\phi,\psi) > 0.9$) have positive ϕ angles; however, Gly86 consistently exhibits a ϕ angle near 150° ($S^{angle}(\phi) > 0.87$).

Description of Cyclophilin Structure. The structure of eCyP consists of an eight-stranded antiparallel β-sheet and three helices. The schematic diagram of the polypeptide fold presented in Figure 4A reveals that the strands of the β-sheet form a barrel, with helices resting at the top and bottom. The secondary structural elements of eCyP in solution are listed in Table 2. These elements were deduced from characteristic NOE connectivities (Wüthrich, 1986), hydrogen-bonding patterns, and the average value of the backbone dihedral angles. Several loops, extended strands, and turns connect elements of regular secondary structure (Table 2); they also play key roles in defining the active site of eCyP and are discussed in detail.

The first two strands of the β-sheet, B1 and B2 (B1, Pro5–Thr10; B2, Asn15–Leu20), are connected by a tight turn comprising residues Thr11–Gly14 (B1–B2 turn) and are located at the rear of the barrel, as viewed in Figure 4A. The polypeptide chain exits B2 at the top of the barrel, adopting a helix-like conformation (Hx1, Lys22–Ala25), before bending toward the core of the protein, closing the barrel in a long α-helix (Hα1, Pro26–Asn37). Residues Ser38–Tyr41 (H1–B3 connector) then form a short turnlike structure, which may be a continuation of Hα1 (Figure 2A). The chain then descends on the right face of the barrel to form strand B3 (B3, Thr44–Val49), which is drawn with two arrows to account for the presence of a β-bulge. After a short turn (Ile50–Phe53), the strand reverses its direction to begin strand B4 (B4, Met54–Gly57), which interacts with strands B3 and B6. The polypeptide chain exits at the top of the barrel and utilizes residues Gly58–Arg85 (B4–B5 connector) to cross the barrel along the protein surface in a long strand of irregular secondary structure. The beginning of the B4–B5 connector forms a short hairpin loop or even a short antiparallel β-sheet, while the second half of the connector may possess rudimentary helical structure (Figure 2B). Assignment of regular secondary structure in the B4–B5 connector is tentative and therefore is not drawn in Figure 4. Strand B5 (Gly86–Met90) pairs with B6 and B7 before forming the long B5–B6 loop (Leu91–Gln102) . After strand B6 (Phe103–Val107), the polypeptide chain exits at the bottom of the barrel, ascends in the B6–H2 connector (Ala108–Ala111) reaching a short helix (Hx2, Phe112–Asp114) and forms the long H2–B7 loop (His115–Ala123). Strand B7 (Val124–Lys130) pairs with B5 and B2; a third helix (Hα3, Met132–Ile138) then follows B7, closing the barrel at the bottom. The long H3–B8 loop (Ser139–Lys155) then overlooks strand B3 before finally leading into strand B8 (Pro156–Leu165) and the C-terminus of the protein. At least three β-bulges occur in the structure of eCyP: two bulges in strand B7 are formed by residues Phe125–Gly126 and Val129–Lys130, while residues Leu160–Ser161 form a bulge in strand B8. A fourth β-bulge is seen in strand B3 at residues His47–Arg48 in >75% of the structures. Residues His47 and Arg48 are not well-defined by the NMR data. In particular, the amide resonances of these residues have weak intensities and exhibit few NOEs. It may be that they rapidly exchange with solvent or undergo conformational averaging.

The core of the barrel is hydrophobic and is predominantly constructed from the side chains of residues that participate in regular secondary structure (Figure 4A). These residues include Val7 and Leu9 of B1; Ile16, Leu18, and Leu20 of B2; Phe32 and Tyr35 of Hα1; Phe46 and Val49 of B3; Ile55 of B4; Ile88 and Met90 of B5; Phe103 and Ile105 of B6; Phe125 and Val128 of B7; Ala135 and Ile138 of Hα3; and Val157, Ile159, and Ala162 of strand B8. The side chains of Thr11, Phe40, and Tyr41 contribute to the core as well.

The surface of eCyP is quite varied: the rear face of the protein (Figure 4A) formed by strands B1, B2, and B8 is

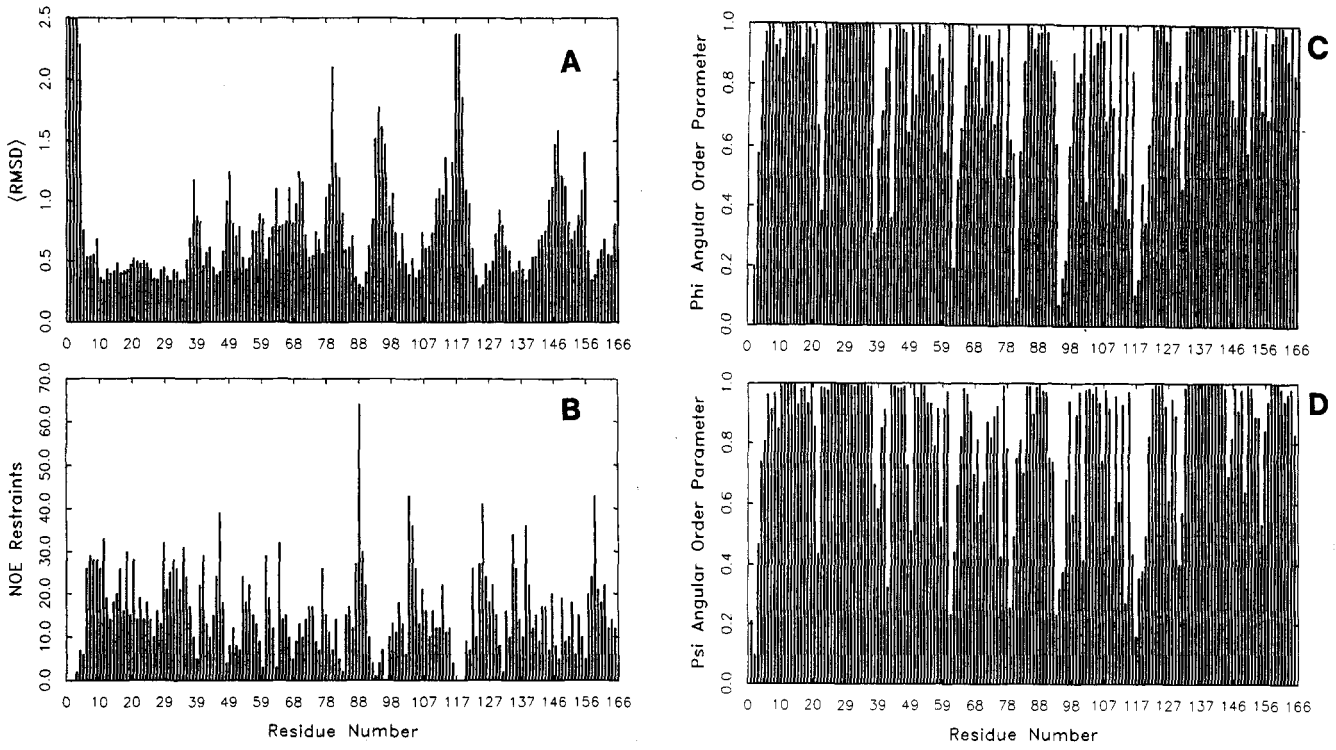

FIGURE 1: (A) Histogram of <rmsd> values of the backbone nuclei (C, C^{α}, and N) from the average structure vs sequence number. (B) Histogram of NOE restraints vs sequence number. The average structure was calculated by averaging the positions of the ensemble after superposition of the backbone nuclei (C, C^{α}, and N) of residues Pro5–Leu165. The <rmsd> values reported were calculated by averaging the individual rms deviations of the three nuclei. (C and D) Histograms of the angular order parameters, S^{angle}, of the ϕ and ψ angles of the ensemble vs sequence number.

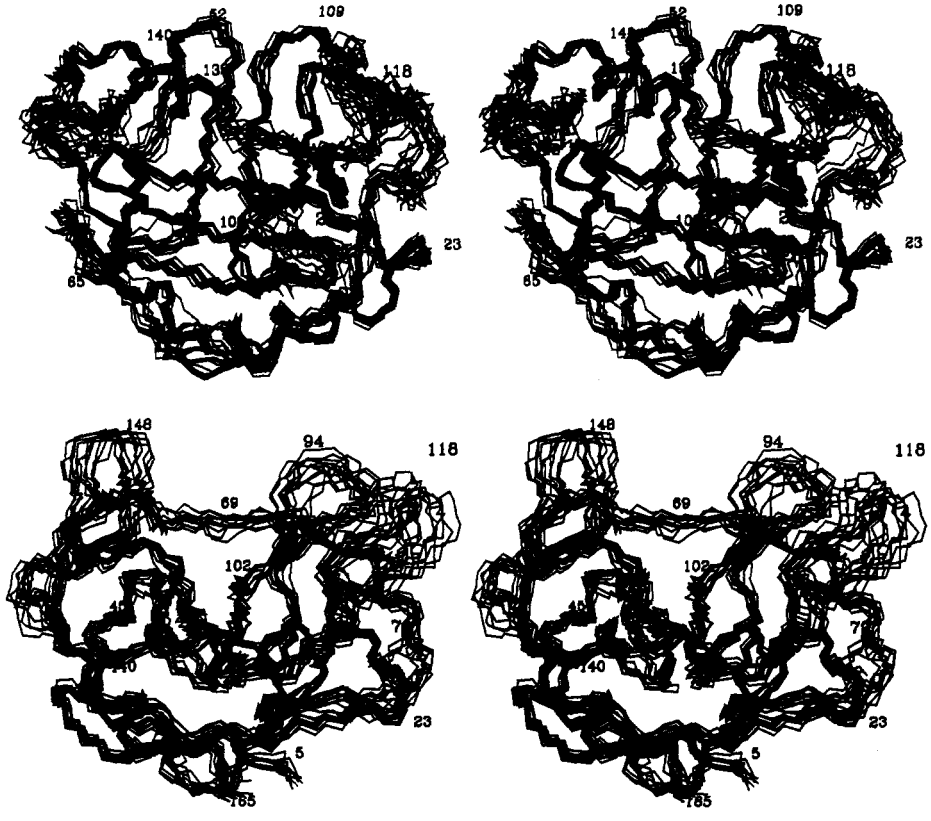

FIGURE 2: (A, top, and B, bottom) Two different orientations of the stereo representation of the 12 accepted structures of *E. coli* cyclophilin (Pro5–Leu165). Structures were aligned by superimposing the C, C^{α}, and N atoms of residues 5–165.

smooth with only side chains protruding. In contrast, the front face of the protein forms a large crevice with strands B3, B4, B6, and B5 lining the floor (Figures 2B and 4A). The left wall of the crevice is formed by the B5–B6 surface loop (Ala91–Ser101), a short helix (Hx2, Phe112–Asp114), and the H2–B7 loop (His115–Gly121). The right side of the crevice is formed by the long H3-B8 surface loop consisting of residues His144–Asn151, with the backbone atoms of residues Pro148–Asn151 nearest the pocket. The pocket is the putative PPIase active site and is also homologous to the

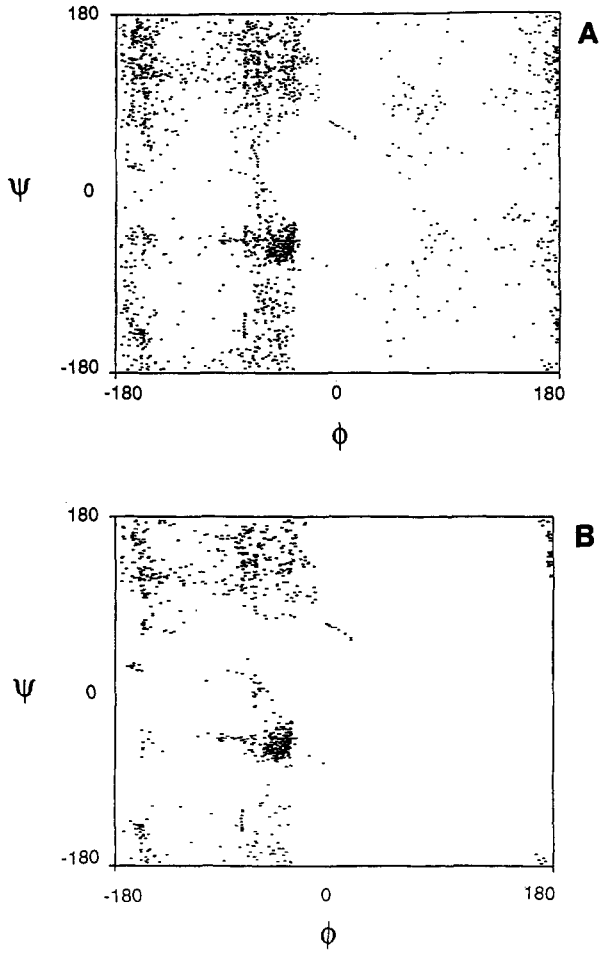

FIGURE 3: Ramachandran plots for (A) all ϕ,ψ angles for the 12 structures and (B) only residues with angular order parameters $S^{angle} > 0.9$.

CsA binding site of hCyP A. Mechanistic implications of its structure as well its relationship to the structure of hCyP A will be addressed in the Discussion section.

The structure of eCyP is well-defined by the NMR data (rms deviation of 0.84 ± 0.18 Å about the mean backbone), especially if one considers that eCyP (>18.2 kDa) is one of the largest protein structures solved by NMR. Several regions in the structure are disordered, however, beginning with the N-terminus (residues Ala1′-Asp4), which has rms deviations between its backbone atoms (C^{α}, N, and C) and the average structure of between 2.3 and 7.2 Å. The backbone atoms of residues Pro5–Asn37 and Gly39–His47 are more precisely determined and have rms deviations of <0.9 Å to the average structure (Figure 1A). This region corresponds to strands B1 and B2 and helices Hx1 and Hα1. The structure becomes less well-defined (rms deviations >0.9 Å) at residues Arg48 and Val49, which are located at the end of strand B3. Further refinement of the structure is necessary, since few distance restraints characterize this region (Figure 1B). Refinement will also help to define residues in strand B3, which could not be unambiguously identified as participants in the β-sheet; only Ile50 and Thr44 were restrained by hydrogen bonds.

Following strand B3, residues Ile50–Glu62 (the B3–B4 turn, B4 and the beginning of the long B4–B5 connector) all possess rms deviations of less than 0.9 Å. In the remaining parts of the long extended chain connecting strands B4 and B5 (B4–B5 connector, Gly58–Arg85), residues Gln63, Lys67, Pro69–Pro71 and Asp78–Arg82 are not well-defined by the NMR data (rms deviations >0.9 Å). However, the overall direction and placement of the chain on the protein surface are consistent among the conformers. Amino acid segments including residues Thr93–Ser98, Ala111–Gly121, and Asp145–Gln150 exhibit rms deviations of >0.9 Å to the average structure. These residues are in or near surface loops (His115–Gly121, Ala91–Ser101, and His144–Asn151) that overlook the active site of eCyP. Five other residues exhibit rms deviations of >0.9 Å to the average structure: Ser38, Lys130, Lys155–Pro156, and Pro166. Additional analysis of the NMR data is required to determine the source of conformational disorder in the structures of eCyP, including heteronuclear relaxation rate measurements to probe the dynamics of the backbone nuclei.

Several features in the NMR spectrum of eCyP hindered the extraction of conformational restraints, including the rapid exchange of the $^1H^N$ protons of some residues with solvent and chemical heterogeneity in the nitrogen-15-labeled sample of the eCyP. Heterogeneity was manifested in the doubling of $^1H^N$ and $^{15}N^H$ resonances for residues His6-Leu9, Ile16–Leu20, and Val129–Asp136. The doubled resonances appear in a 60:40 ratio, with maximum chemical shift differences of 0.65 ($^{15}N^H$) and 0.08 ($^1H^N$) ppm at Lys130. These residues reside on the rear face of the protein in strands B1 and B2 and near the beginning of helix Hα3. The source of this heterogeneity is unknown, but may result from the deamidation of Asn15. However, additional chromatography or mass spectrometry analysis would have to be carried out to clarify this question. Since the heterogeneity is distant from the active site, it is not expected to perturb the structure of this region. Several residues exhibit rapid $^1H^N$ proton exchange with solvent (Clubb et al., 1993), including the N-terminus (Ala1′–Asp4), residues in the extended chain that connects strands B4 and B5 (Gly58–Gly59, Glu62–Gln63, Glu76, and Asn83), and residues in the two active-site surface loops (Gln117–Phe120 and Arg92–Asp97). The $^1H^N$ proton of Gly131 also rapidly exchanges with the solvent and is located in a turn connecting strand B7 and helix Hα3.

DISCUSSION

Comparison of Human and E. coli Structures. The cyclophilin family of proteins is ubiquitously expressed in a variety of subcellular compartments, in both prokaryotic and eukaryotic organisms. At least 27 cyclophilins have been cloned and their amino acid sequences determined (Trandinh et al., 1992). Eukaryotic cyclophilins typically possess greater than 60% sequence identity with human T cell cyclophilin, the only cyclophilin for which a three-dimensional structure is known. The three bacterial cyclophilins are distantly related, possessing only 34–36% sequence identity with human T cell cyclophilin (Trandinh et al., 1992). The three-dimensional structure of periplasmic *E. coli* cyclophilin (34% sequence identity with hCyP A), therefore, provides the first assessment of the structural diversity of this important family of proteins.

Schematic representations of the unligated average solution structure of eCyP and the X-ray structure of unligated hCyP A (Ke, 1992) are presented in Figure 4A,B, respectively. Clearly, the two structures are very similar; structural variations predominantly occur in regions of irregular secondary structure. Figure 5 shows an amino acid sequence alignment of hCyP A and eCyP based on their tertiary structures, which differs from the previous alignment based on secondary structure (Clubb et al., 1993). In particular, eCyP residues Thr84–Thr93, Gly116–Ala123, and Met132–Gly147 have been realigned. Structurally conserved regions are underlined, and the backbone atoms (C, C^{α}, and N) of

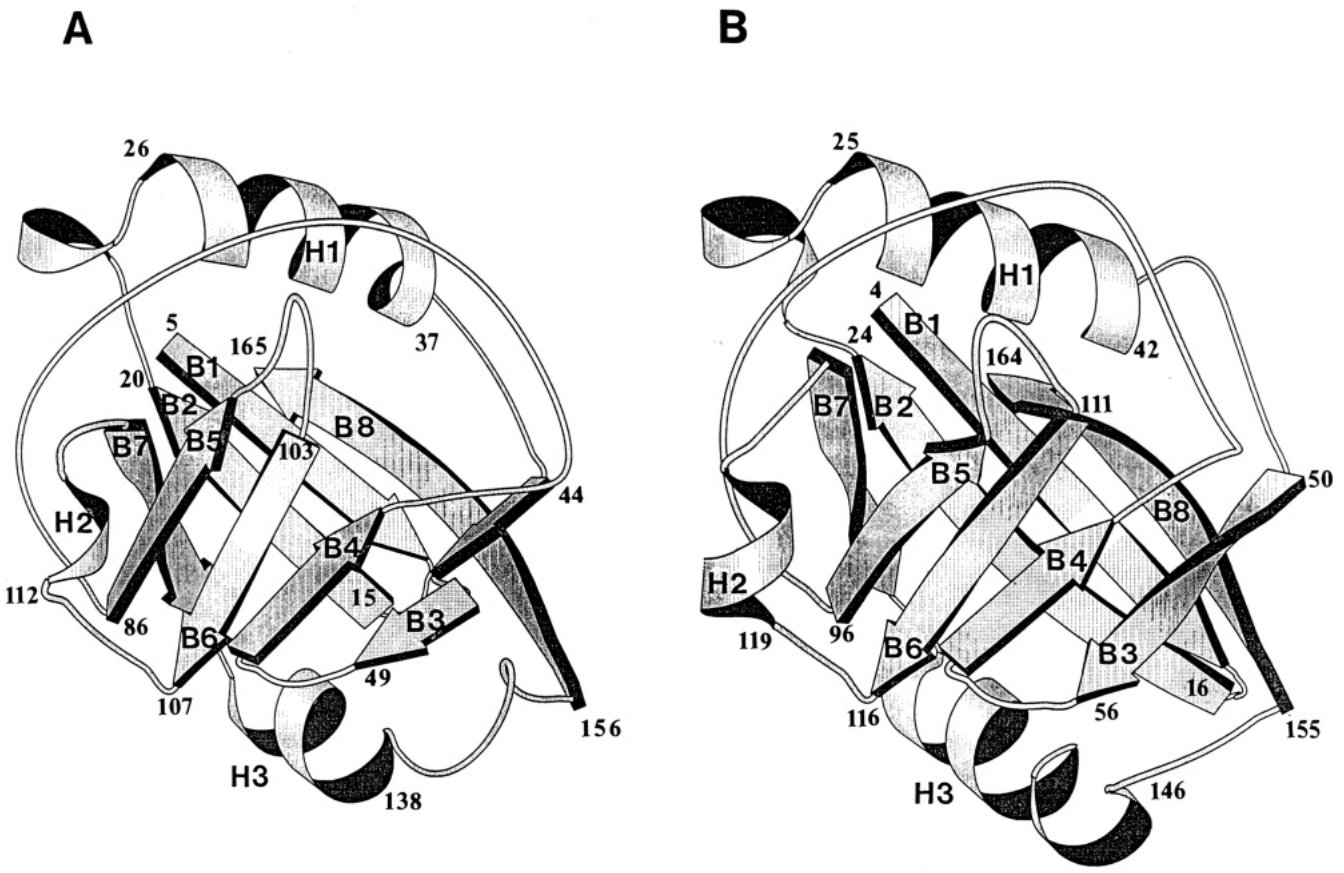

FIGURE 4: Schematic representation of (A) the average solution structure of eCyP and (B) the X-ray structure of unligated hCyP A (Ke et al., 1991; Ke, 1992). β-Sheet strands and helices are labeled B and H, respectively. Secondary structure designations in hCyP A are identical to those defined by Ke (1992) (Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, accession number 2CPL). Diagrams were generated with the program MOLSCRIPT (Kraulis, 1991).

Table 2: Summary of Structural Elements in the Solution Structure of eCyP[a]

structural elements	residue number	structural elements	residue number
B1	5–10	B5	86–90
B1–B2 turn	11–14	B5–B6 loop	91–102
B2	15–20	B6	103–107
Hx1	22–25	B6–H2 connector	108–111
Hα1	26–37	Hx2	112–114
H1–B3 connector	38–43	H2–B7 loop	115–123
B3[b]	44–49	B7	124–130
B3–B4 turn	50–53	Hα3	132–138
B4	54–57	H3–B8 loop	139–155
B4–B5 connector	58–85	B8	156–165

[a] Key: Hα, α-helix; B, β-sheet; Hx, 3_{10}- or α-helix. [b] All residues in B3 could not be unambiguously identified as participants in the β-sheet.

these regions in the average solution structure and X-ray structure (Ke, 1992) have an rms deviation of 1.44 Å.

Two variations are distal to the active site. In eCyP, strands B1 and B2 are shorter than in hCyP A. As shown in Figure 5, hCyP A has an insertion consisting of residues Ala11h–Glu15h (h identifies the hCyP A sequence). A β-bulge in hCyP at residues Leu17h–Gly18h is replaced by a reverse turn in eCyP (Thr11–Gly14). The human structure also has an insertion of three residues corresponding to Glu43h–Gly45h, which forms a small surface loop after the first α-helix. This loop in hCyP A appears to be more mobile than the rest of the protein backbone, exhibiting higher *B* factor values (Ke, 1992). Interestingly, this insertion is conserved in all eukaryotic proteins, but is absent from the three known prokaryotic cyclophilins. Both insertions in the hCyP structure are located on the rear face of the protein (Figure 4) distal to the active site and are unlikely to affect CsA binding or PPIase activity.

Four structural variations between eCyP and hCyP A occur in or near their active sites and may therefore affect the functional properties of eCyP. These variations are illustrated in Figure 6, which displays a stereo representation of the superimposed tertiary structures of eCyP (red, Thr45–Ser154) and hCyP A (purple, Cys52–Ser152). Three surface loops (B5–B6, H2–B7, and H3–B8) overlook the active site in eCyP and each differs structurally from hCyP A. A five-residue insertion occurs in the H3–B8 surface loop of eCyP and is reminiscent of two strands of an antiparallel sheet, with Pro148 and Gly147 forming the reverse turn between strands. Hydrophobic contacts with the Gly59–Lys68 loop position the H3–B8 loop toward the bottom of the pocket (lower left of Figure 6), with residues Tyr149 and Val146 anchoring the loop to Phe60 and Met64. The nonconserved H2–B7 loop in eCyP, characteristic of the prokaryotic cyclophilins, protrudes from the active site and results from a five-residue insertion (Gly116–Phe120) in the eCyP sequence. A third active-site loop (B5–B6) formed by residues Ala91–Ser101 in eCyP also contains a one-residue insertion relative to hCyP A. The homologous loop in hCyP A forms essential contacts with CsA and tetrapeptide substrates.

The structure of the long connecting sequence between strands B4 and B5 also differs between eCyP and hCyP A. In particular, residues Gly59–Lys68 in eCyP form a loop that points away from the active site, while the comparable loop in hCyP A faces toward the active site. Not surprisingly, residues in this region possess low sequence homology (eCyP, G59–K68; hCyP A, Asp66h–Gly75h). The remainder of the

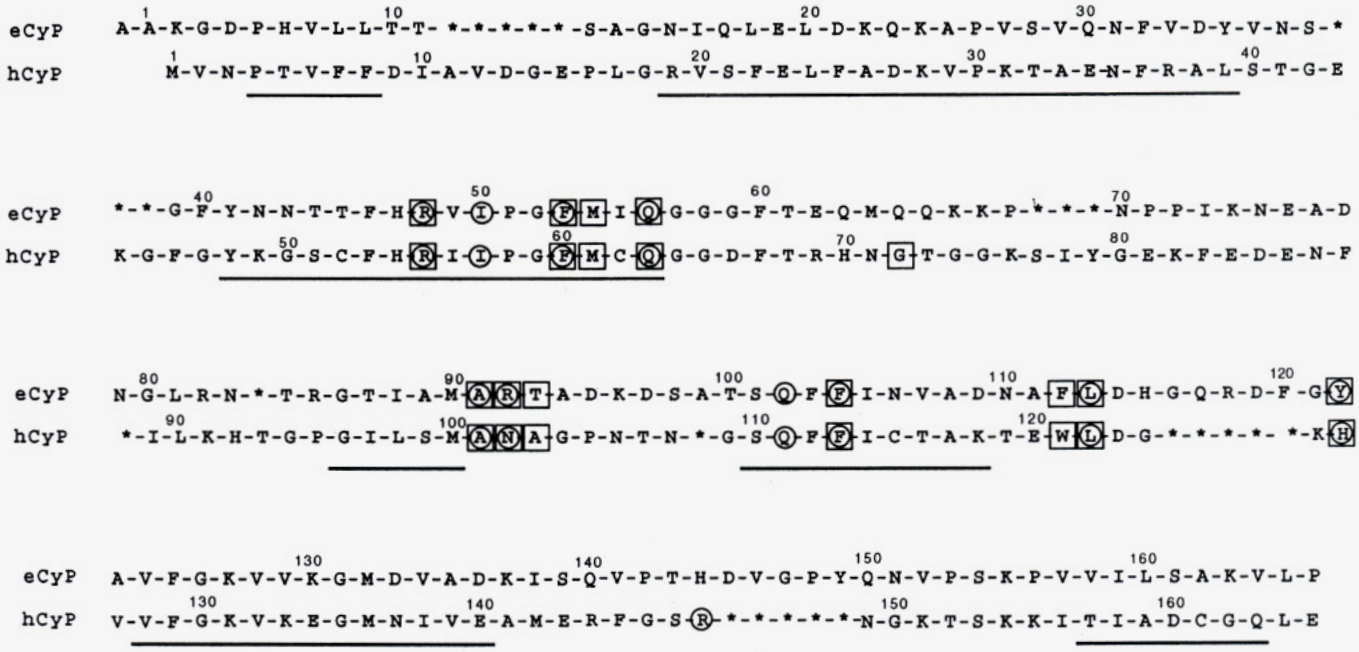

FIGURE 5: Alignment of the primary sequences of *E. coli* and human T cell cyclophilins based on their tertiary structures. Underlined residues are structurally homologous. Residues in hCyP enclosed in boxes have heavy atoms within 4.0 Å of CsA (Pflügl et al., 1992; Thériault et al., 1992); homologous residues in eCyP are also marked with boxes. Residues in hCyP A enclosed in circles have heavy atoms within 3.5 Å of the tetrapeptide substrate AC-Ala-Ala-Pro-Ala-AMC (Kallen et al., 1991; Kallen & Walkinshaw, 1992); homologous residues in eCyP are also enclosed by circles.

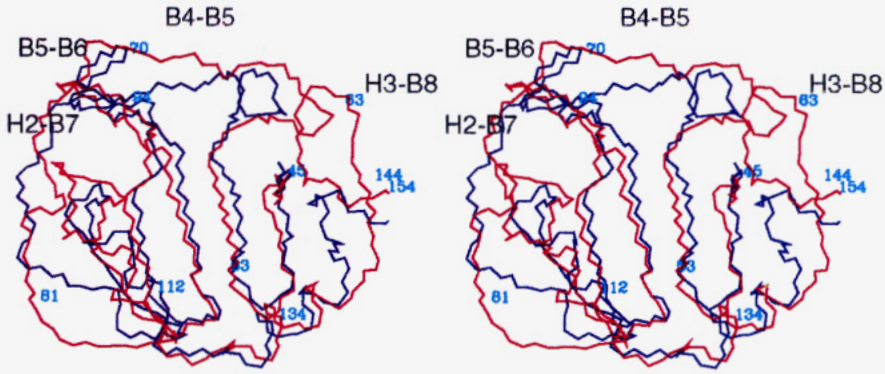

FIGURE 6: Stereo plot of the superimposed tertiary structures of hCyP A (purple) and eCyP (red). The backbone atoms of the underlined residues of Figure 5 were superimposed. Backbone atoms of residues Cys52–Ser153 of hCyP A and Thr45–Ser154 of eCyP are displayed. The loops discussed in the text are annotated. Sequence numbers refer to the eCyP structure.

surface connector (eCyP, Pro70–Arg85) also differs substantially; however, both chains transverse a similar path and differ only in the locations of kinks and bends in the chain.

Peptidyl–Prolyl Isomerization. *E. coli* and human cyclophilins catalyze the interconversion of Xaa–Pro bonds in tetrapeptides with equal catalytic efficiency (hCyP A, K_{cat}/K_m = 16 μM^{-1} s^{-1}; eCyP, K_{cat}/K_m = 20 μM^{-1} s^{-1}) and differ only slightly in their preference for the amino acid (Xaa) preceding the proline (Liu et al., 1991a). *E. coli* and human cyclophilins differ substantially, however, in their sensitivity to CsA. The catalytic activity of eCyP is insensitive to CsA (K_i > 8.8 μM), while hCyP A activity is strongly inhibited (K_i = 17 nM). The proteins can also be distinguished by their ability to catalyze the slow folding reactions of larger partially folded substrates (Schonbrunner et al., 1991). In the folding of ribonuclease T1, an intermediate step whose rate is determined by the trans to cis conversion of its Ser54–Pro55 peptide bond is accelerated equally well by a variety of cyclophilins. The terminal phase of folding, however, which is limited by the trans to cis conversion of its Tyr38–Pro39 peptide bond, is accelerated only by eCyP but not by hCyP A (Schonbrunner et al., 1991). Interconversion of the Tyr38–Pro39 peptide bond in ribonuclease T1 occurs in a largely folded intermediate; eCyP therefore differs from its eukaryotic counterparts in its steric requirements for isomerization and may, in general, be more efficient at isomerizing sterically restricted polypeptide chains (Schonbrunner et al., 1991).

Mechanistic studies of cyclophilin from calf thymus led to the proposal that catalysis occurs by distortion of the amide bond (Harrison & Stein, 1990). In this cyclophilin (calf thymus), catalysis is independent of pH between 5.5 and 9.0 and exhibits no primary or secondary isotope effects, suggesting that the enzyme uses favorable noncovalent interactions with the substrate to stabilize a transition state in which the C–N amide bond is partially rotated (Harrison & Stein, 1990). Recently, Ke et al. have proposed a related mechanism of catalysis, on the basis of crystallographic analysis of an hCyP A:Ala-Pro complex (Ke et al., 1993). In this model, the peptidyl–prolyl bond is desolvated in the ground state by binding to the active site of hCyP A, and a twisted intermediate state is then stabilized by a hydrogen bond between the carbonyl oxygen of the amide bond and a bound water molecule. Unfortunately, X-ray structures of hCyP A in complex with di- and tetrapeptide substrates indicate that the

ground state (cis form) of the peptide is bound, failing to distinguish the two mechanisms (Kallen & Walkinshaw, 1992; Ke et al., 1993). Site-directed mutagenesis studies of hCyP A have also provided mechanistic insights, identifying three residues essential for PPIase activity: Arg55h (B3), Phe60h (B3–B4 loop), and His126h (H2–B7 loop) (Zydowsky et al., 1992).

The solution structure of eCyP reveals that the PPIase active sites of eCyP and hCyP A are highly conserved and likely bind peptide substrates in a similar manner (Figure 6). There are 11 residues in hCyP A, with heavy atoms within 3.8 Å of the tetrapeptide substrate AC-Ala-Ala-Pro-Ala-AMC (enclosed in circles in Figure 5) (Kallen & Walkinshaw, 1992). Eight of these eleven residues are conserved in eCyP, while Asn102h, His126h, and Arg148h are not conserved. In eCyP, Asn102h is changed to Arg92; this change is probably insignificant, since only the backbone carbonyl oxygen and amide proton of Asn102h interact with the substrate in the hCyP A:tetrapeptide complex. In eCyP there is no obvious homologue to Arg148h of hCyP A, which is located in a short surface loop near the coumarin group of the tetrapeptide (Kallen & Walkinshaw, 1992). The analogous loop is extended in eCyP and positioned toward the N-terminus of a potential substrate. Since Arg148h does not make any specific interactions with the tetrapeptide, its importance in catalysis is probably minimal.

A significant change occurs when the essential residue His126h of hCyP A (in the H2–B7 loop) is changed to Tyr122 in eCyP. If tetrapeptide substrates bind to eCyP the same as in hCyP A, and no substantial structural rearrangements occur in eCyP upon substrate binding, the larger size of the tyrosyl ring would cause a steric clash with the prolyl ring of the substrate. The hydroxyl group of Tyr122 may therefore require the substrate to adopt a different position in the pocket. If so, the hydroxyl group of Tyr122 may form a hydrogen bond or protonate the tetrahedral nitrogen lone pair of the substrate as the amide bond is rotated. This would imply that the mechanism of catalysis in eCyP is different from that in the human protein, which possesses no serine, threonine, or tyrosine residues near the peptide amide bond. It is also equally plausible, however, that Tyr122 in eCyP and His126h in hCyP are required as hydrophobic residues only and are essential in buttressing the prolyl ring of the substrate. While mutation of His126h to glutamic acid abolishes enzymatic activity, our results suggest that a His126Tyr mutant of hCyP A would retain enzymatic function.

The structure of eCyP may also provide an explanation for the variable accelerations seen in the folding of ribonuclease T1. In hCyP A, the B5–B6 loop (Ala101h–Ser110h) and the hairpin-like structure at the beginning of the long B4–B5 connector (Arg69h–Gly74h) may form the walls of a "pass", through which the N-terminal portion of the polypeptide substrate may pass (Kallen & Walkinshaw, 1992). The substrate pass in eCyP is wider, since the loop analogous to Arg69h–Gly74h of hCyP A (Glu62–Lys67 in eCyP) points away from the active site (Figure 6). The recognition of protein substrates appears to be a very complex process since the structures of eCyP and hCyP A differ substantially around their active sites. Conservation of the majority of core residues near the peptide bond of the substrate, however, ensures that cyclophilins from different species can catalyze short peptide substrates with nearly equal efficiency. Structural analysis of eCyP does not distinguish between the two aforementioned PPIase mechanisms: (1) stabilization of a twisted amide transition state or (2) hydrogen-bond formation between a bound water molecule and the peptide carbonyl oxygen. It is interesting to note, however, that the protein side chain utilized to bind the relevant water molecule in unligated and ligated structures of hCyP A (Gln63h) (Ke et al., 1993) is conserved in eCyP (Gln56) and all other known cyclophilins (Trandinh et al., 1993).

Cyclosporin A Sensitivity. Cyclosporin A binds to and inhibits the PPIase activity of hCyP A at nanomolar concentrations (IC_{50} = 6 nM); in contrast, eCyP is insensitive to inhibition (IC_{50} = 3000 nM) (Handschumacher et al., 1984; Harding et al., 1986; Liu et al., 1990). A single-point mutation in eCyP, Phe112Trp (located at the end of the B6–H2 connector), however, dramatically increases the sensitivity of PPIase activity to inhibition (IC_{50} = 130 nM) (Liu et al., 1991). Sensitivity to CsA cannot be completely attributed to the presence of a single tryptophan residue, since the mutant Phe112Trp eCyP is still 10-fold less sensitive than hCyP A. NMR studies of CsA bound to the Phe112Trp mutant of eCyP indicate that CsA adopts a conformation similar to its conformation when bound to hCyP A and that the structure of mutant Phe112Trp eCyP is similar to wild-type unligated eCyP (Fejzo et al., 1993). Therefore, the three-dimensional structures of wild-type unligated eCyP and the hCyP A:CsA complex provide an opportunity to explore the structural basis of CsA binding and PPIase inhibition.

In the hCyP A:CsA complex, the drug:protein interface is formed by the side chains of CsA (MeBmt1, Abu2, Sar3, MeLeu9, MeLeu10, and MeVal11) and 13 residues of hCyP A (enclosed in boxes in Figure 5) (Pflügl et al., 1992; Thériault et al., 1992). The solvent-accessible surface of the hCyP A interface residues form a groove that also serves as the PPIase active site; these residues are shown in Figure 7A colored yellow. In eCyP, eight of these thirteen interface residues are conserved; Figure 7B shows the solvent-accessible surface of the conserved residues colored yellow, and the five nonconserved residues are colored red. The five variations correspond to the conversion of hCyP A residue Asn102h to Arg92, Ala103h to Thr93 (B5–B6 loop), Trp121h to Phe112 (B6–H2 connector), His126h to Tyr122 (H2–B7 loop), and likely Gly72h to Phe60 (B4–B5 connector).

Four intermolecular hydrogen bonds join CsA to hCyP A in the X-ray and NMR structures of the hCyP A:CsA complex (Pflügl et al., 1992; Thériault et al., 1992). Structural homology between eCyP and hCyP A indicates that three of the four hydrogen bonds should be conserved in an eCyP:CsA complex; the hydrogen bond between Trp121h (ϵNH) and MetLeu9 (CO) cannot be formed by Phe112 of eCyP. The remaining four amino acid differences between hCyP A and eCyP cause a 10-fold reduction in affinity for CsA; therefore, the surface formed by these residues in eCyP may be sterically and electrostatically less attractive to CsA. When the structure of CsA is manually docked to the active site of eCyP in an orientation similar to that observed in the hCyP A:CsA complex, the side chain of Tyr122 overlaps with the *N*-methyl group of MetVal11. This steric hindrance is relieved by slightly moving CsA, and rotation of the tyrosyl ring then leaves the hydroxyl group poised to form a hydrogen bond with the carbonyl oxygen of MeVal11. Asn101h and Ala102h of hCyP A are not conserved and they correspond to Arg91 and Thr92 in eCyP. These changes are probably insignificant since only the backbone carbonyl oxygen of Asn102h interacts with CsA, and Ala102h and Thr92 share similar hydrophobic methyl groups. A more detailed analysis is not possible since Arg91 and Thr92 are not well-determined in the ensemble of eCyP structures (Figure 1A).

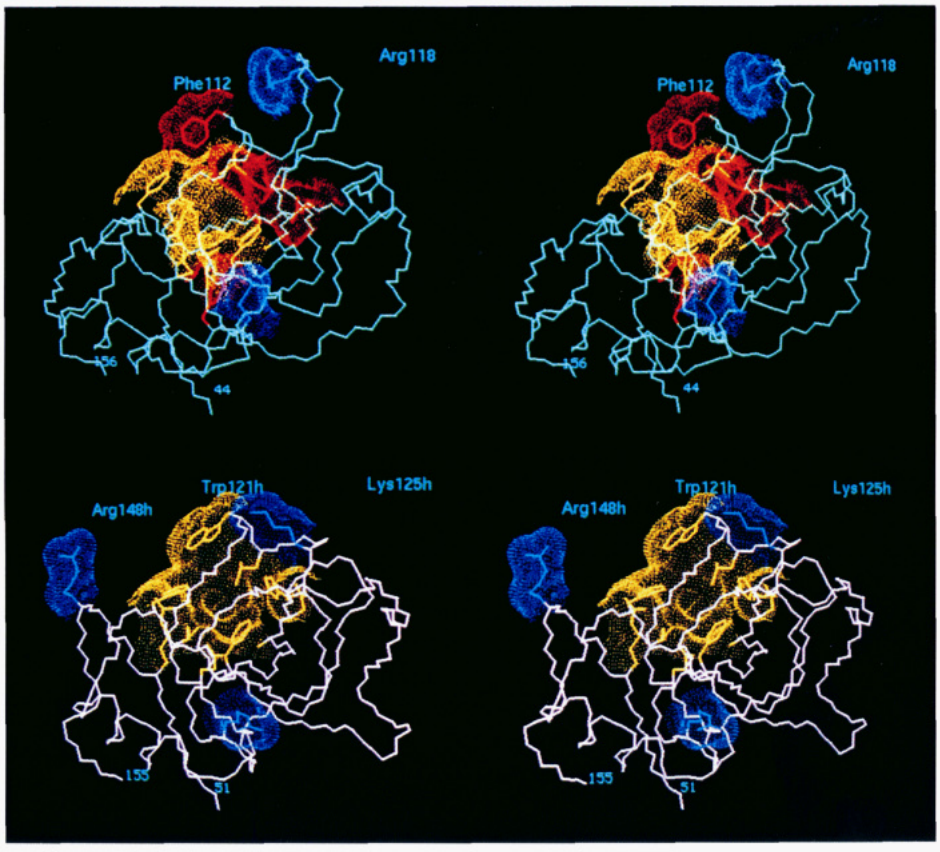

FIGURE 7: Solvent-accessible surfaces of the CsA binding interface of (A, top) hCyP A and (B, bottom) the homologous region of eCyP. Side-chain surfaces colored yellow are residues in hCyP A that contact CsA (A) and those residues in eCyP that are completely conserved (B). Interface residues not conserved in eCyP are colored red (B). Blue side-chain surfaces correspond to residues in hCyP A that affect CN inhibition in the hCyP A:CsA complex (A) and their potential homologues in eCyP (B).

The protein interfaces of eCyP and hCyP A differ substantially near Gly72h of hCyP A (B4–B5 connector). In the homologous region of eCyP, two loops, Glu62–Lys67 and His144–N151, provide a unique protein interface near residues Abu2–Val5 of CsA. The distinctive loop conformation of Glu62–Lys67 in eCyP positions the side chain of Phe60 near Sar3 and MeLeu4 of CsA. The side chain of Tyr149 in the extended His144–N151 loop contacts Phe60 and may also interact with CsA. The structure of CsA bound to mutant Phe112Trp eCyP has been solved (Fejzo et al., 1993); further analysis of the complex will elucidate the structure of the protein:drug interface.

In summary, the lower affinity of wild-type eCyP relative to hCyP A for the immunosuppressant drug CsA can be attributed to the absence of a single hydrogen bond to MeVal9 of CsA. The presence of a tyrosyl ring near MetVal11 and the dramatically different eCyP conformation near the putative interface with residues Abu2–Val5 of CsA may account for the remaining 10-fold discrepancy in CsA affinity.

Calcineurin Inhibition. CsA bound to hCyP A blocks T cell activation and graft rejection by inhibiting the protein serine phosphatase activity of calcineurin (CN) (Liu et al., 1991b). Cyclophilin acts as a scaffold, presenting the correct conformation of CsA to CN. Recent mutagenesis studies of hCyP A, however, indicate that a composite surface formed by CsA and hCyP A interacts with CN (Zydowsky et al., 1992; Etzkorn et al., 1993). Mutant Phe112Trp eCyP and hCyP A bind CsA in similar conformations, and the mutant Phe112Trp eCyP:CsA complex also binds and inhibits CN (Fejzo et al., 1993). The three-dimensional structures of unligated wild-type eCyP and mutant Phe112Trp eCyP in complex with CsA are also similar (Fejzo et al., 1993). Therefore, the results of mutagenesis studies of hCyP A and a comparison of its structure with unligated wild-type eCyP can be utilized to determine many of the essential components of the protein:drug interface responsible for CN inhibition.

Biochemical studies of Etzkorn et al. (1993) indicate that three cationic residues, Arg69h (B4–B5 connector), Arg148h (H3–B8 loop), and Lys125h (H2–B7 loop), modulate the inhibitory properties of the hCyP A:CsA complex without significantly affecting PPIase activity or CsA affinity. In addition, the indole ring of Trp121h also forms part of the recognition surface, since the hCyP A Trp121Ala mutant binds CsA but is a weak inhibitor of CN (Zydowsky et al., 1992). In hCyP A, conversion of Arg69h to glutamate (Arg69Glu) results in a 13-fold decrease in CN inhibition, while conversion to a neutral side chain (Arg69Leu) results in only a 3-fold decrease. Surprisingly, mutation of Arg148h (H3–B8 loop) to glutamate (Arg148Glu) increases the inhibitory effectiveness of the hCyP A:CsA complex, while leucine at this position (Arg148Leu) has little effect (Etzkorn et al., 1993). A questionable role in CN inhibition is played by Lys125h (H2–B7 loop), since its mutagenesis to glutamine impaired CN inhibition 6.3-fold, while conversion to glutamic acid had no effect (Etzkorn et al., 1993).

Superposition of the structures of eCyP and hCyP A indicates that Gln66 of eCyP replaces Arg69h of hCyP A. The side chains of Gln66 and Arg69h are colored blue in Figure 7A,B, respectively. In contrast, Arg148h of hCyP A is not conserved in eCyP; instead different loop positions result in a cleft in the eCyP structure where the side chain of Arg148h is located (colored blue in Figure 7A). In eCyP, Lys125h may be mimicked by a surface loop that contains Arg118 at its tip. Mutagenesis studies therefore predict at least a 3-fold

decrease in the inhibitory effectiveness of the Phe112Trp eCyP mutant relative to hCyP A (3-fold decrease for the neutral Gln66 instead of Arg69h, no change for the missing positive charge of Arg148h, and possibly no change for substitution of Arg118 for Lys125h). This prediction is consistent with the lower inhibitory effectiveness of Phe112Trp eCyP:CsA (K_i = 1070 nM) compared to the hCyP A:CsA complex (K_i = 336 nM) (Etzkorn et al., 1993). Therefore, the lack of a positive charge on the Phe112Trp eCyP:CsA interface at position Gln66 may explain its decreased inhibitory effectiveness.

Large variations in the tertiary structures of eCyP and hCyP A in regions proximal to the CsA binding site indicate that the CN specificity of the complex is primarily determined by the conformation of the bound drug. Interestingly, one of the three hCyP A residues discovered by mutagenesis to affect CN inhibition is structurally conserved in eCyP, while another may be effectively mimicked. The physiological relevance of this apparent structural conservation remains unknown. It is worthwhile to note that all eukaryotic cyclophilins, with the exception of *Mus musculus*, retain cationic or neutral residues at positions homologous to Arg69h and Lys125h in hCyP A. Since these residues are not essential for PPIase activity or CsA binding, they may be conserved to allow cyclophilins to interact with CN or interact/fold some other as yet unidentified protein. The importance of cationic residues in CN recognition is emphasized further by mutagenesis studies of FKBP12. In this immunophilin:drug complex, Arg42 is important for CN inhibition (Aldape et al., 1992; Yang et al., 1993), while overlay of the FKBP12:FK506 and CyP:CsA complexes indicates that Arg42 may play a role similar to that of Arg69h of hCyP A (Etzkorn et al., 1993).

CONCLUSION

The solution structure of the *E. coli* cyclophilin-type peptidyl–prolyl isomerase (eCyP) has been determined by NMR spectroscopy. The protein consists of an eight-stranded antiparallel β-barrel with helices at the ends and is one of the largest structures determined by NMR. The global fold of eCyP is similar to that of human T cell cyclophilin, but differs substantially in the conformation of its active site. In eCyP, three surface loops overlook the active site and are extended relative to hCyP A. The majority of residues that form the PPIase active site are conserved between eCyP and hCyP A, consistent with their nearly equal PPIase catalytic efficiencies toward peptide substrates. Interestingly, the PPIase-active eCyP substitutes Tyr122 for a histidine residue known to be essential for hCyP A enzymatic activity. The lower affinity of wild-type eCyP for CsA compared to hCyP A can be attributed substantially to the absence of a single hydrogen bond to MeVal9 of CsA. The presence of the a tyrosyl ring near MetVal111 and the dramatically different eCyP conformation near the putative interface with residues Abu2–Val5 of CsA may account for the remaining 10-fold discrepancy in CsA affinity.

ACKNOWLEDGMENT

We are indebted to Dr. Sekhar Talluri for writing the software used in the analysis of the NOESY data and Dr. Sven Hyberts for providing the software to generate Figures 1C and 3. We also thank Drs. Timothy Havel, Andrzej Krezel, and Jane Withka for assistance with the structure calculations and Dr. Jasna Fejzo and Mrs. Pei-li Zhang for useful discussions.

SUPPLEMENTARY MATERIAL AVAILABLE

Table of experimental NMR restraints used to determine the three-dimensional solution structure of periplasmic cyclophilin from *E. coli* (32 pages). Ordering information is given on any current masthead page.

REFERENCES

Aldape, R. A., Futer, O., DeCenzo, M. T., Jarrett, B. P., Murcko, M. A., & Livingstone, D. J. (1992) *J. Biol. Chem. 267*, 16029–16032.
Archer, S. J., Ikura, M., Sporn, M. B., Torchia, D. A., & Bax, A. (1991) *J. Magn. Reson. 95*, 636–641.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., & Kay, L. E. (1990a) *J. Magn. Reson. 87*, 620–627.
Bax, A., Clore, G. M., & Gronenborn, A. M. (1990b) *J. Magn. Reson. 88*, 425–431.
Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 320–346.
Bodenhausen, G., & Ruben, D. L. (1980) *Chem. Phys. Lett. 69*, 185–188.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Clore, G. M., Bax, A., & Gronenborn, A. M. (1991) *J. Biomol. NMR 1*, 13–22.
Clubb, R. T., Thanabal, V., Fejzo, J., Ferguson, S. B., Zydowsky, L., Baker, C. H., Walsh, C. T., & Wagner, G. (1993) *Biochemistry 32*, 6391–6401.
Compton, L. A., Davis, J. M., MacDonald, J. R., & Bachinger, H. P. (1992) *Eur. J. Biochem. 206*, 927–934.
Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T., & Gronenborn, A. M. (1990) *Biochemistry 29*, 3542–3556.
Emmel, E. A., Verweij, C. L., Durand, D. B., Higgins, K. M., Lacy, E., & Crabtree, G. (1989) *Science 246*, 1617–1620.
Etzkorn, F. A., Chang, Z., Stolz, L. A., & Walsh, C. T. (1993) *Biochemistry* (submitted for publication).
Fejzo, J., Etzkorn, F. A., Clubb, R. T., Stolz, L. A., Shi, Y., Wagner, G., & Walsh, C. T. (1994) *Biochemistry* (submitted for publication).
Fesik, S., & Zuiderweg, E. R. P. (1988) *J. Magn. Reson. 78*, 588–593.
Fesik, S. W., Eaton, H. L., Olejniczak, E. T., Zuiderweg, E. R. P., McIntosh, L. P., & Dahlquist, F. W. (1990) *J. Am. Chem. Soc. 112*, 886–888.
Fesik, S. W., Gampe, R. T., Jr., Eaton, H. L., Gemmecker, G., Olejniczak, E. T., Neri, P., Holzman, T. F., Egan, D. A., Edalji, R., Simmer, R., Helfrich, R., Hochlowski, J., & Jackson, M. (1991) *Biochemistry 30*, 6574–6583.
Fesik, S. W., Neri, P., Meadows, R., Olejniczak, E. T., & Gemmecker, G. (1992) *J. Am. Chem. Soc. 114*, 3165–3166.
Fischer, G., Wittmann-Liebold, B., Lang, K., Kiefhaber, T., & Schmid, F. X. (1989) *Nature 337*, 476–478.
Flanagan, W. M., Corthesy, B., Bram, R. J., & Crabtree, G. (1991) *Nature 352*, 803–807.
Freskgard, P., Bergenhem, N., Jonsson, B., Svensson, M., & Carsson, U. (1992) *Science 258*, 466–469.
Handschumacher, R. E., Harding, M. W., Rice, J., Drugge, R. J., & Speicher, D. W. (1984) *Science 226*, 544–547.
Harding, M. W., Handschumacher, R. E., and Speicher, D. W. (1986) *J. Biol. Chem. 261*, 8547–8555.
Harding, M. W., Galat, A., Uehling, D. E., & Schreiber, S. L. (1989) *Nature 341*, 758–760.
Harrison, R. K., & Stein, R. L. (1990) *Biochemistry 29*, 1684–1689.
Havel, T. F. (1991) *Prog. Biophys. Mol. Biol. 56*, 43–78.
Havel, T., & Snow, M. (1991) *J. Mol. Biol. 217*, 1–7.
Hayano, T., Takahashi, N., Kato, S., Maki, N., & Suzuki, M. (1991) *Biochemistry 30*, 3041–3048.
Hsu, V. L., & Armitage, I. A. (1992) *Biochemistry 31*, 12778–12784.

Hyberts, S. G., Marki, W., & Wagner, G.(1987) *Eur. J. Biochem 164*, 625–635.
Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G.-(1992) *Protein Sci. 1*, 736–751.
Ikura, M., Kay, L. E., Tschudin, R., & Bax, A. (1990) *J. Magn. Reson. 86*, 204–209.
Jackson, S., & Fersht, A. (1991) *Biochemistry 30*, 10428–10433.
Kallen, J., & Walkinshaw, M. D. (1992) *FEBS Lett. 300*, 286–290.
Kallen, J., Spitzfaden, C., Zurini, M. G. M., Wider, G., Widmer, H., Wüthrich, K., & Walkinshaw, M. D. (1991) *Nature 353*, 276–279.
Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (1990a) *J. Magn. Reson. 89*, 496–514.
Kay, L. E., Ikura, M., & Bax, A. (1990b) *J. Am. Chem. Soc. 112*, 888–889.
Ke, H. (1992) *J. Mol. Biol. 228*, 539–550.
Ke, H., Zydowsky, L. D., Liu, J., & Walsh, C. T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9483–9487.
Ke, H., Mayrose, D., & Cao, W. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3324–3328.
Kern, D., Drakenberg, T., Wikstrom, M., Forsén, S., Bang, H., & Fischer, G. (1993) *FEBS 323*, 198–202.
Kiefhaber, T., Quass, R., Hahn, U., & Schmid, F. X. (1990) *Biochemistry 29*, 3061–3070.
Kraulis, P. J. (1991) *J. Appl. Crystallogr 24*, 946–950.
Kumar, R. M., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.
Lang, K., Schmid, F. X., & Fischer, G. (1987) *Nature 329*, 268–270.
Liu, J., & Walsh, C. T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4028–4032.
Liu, J., Chen, C. M., & Walsh, C. T. (1991a) *Biochemistry 30*, 2306–2310.
Liu, J., Farmer, J. D., Lane, W. S., Friedman, J., Weissman, I., & Schreiber, S. L. (1991b) *Cell 66*, 807–815.
Lodish, H. F., & Kong, N. (1991) *J. Biol. Chem. 266*, 14835–14838.
Luben, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V., & Goff, S. P. (1993) *Cell 73*, 1067–1078.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A., & Clore, G. M. (1989a) *Biochemistry 28*, 6150–6156.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989b) *J. Magn. Reson. 85*, 393–399.
Montelione, G. T., Winkler, M. E., Rauenbuehler, P., & Wagner, G. (1989) *J. Magn. Reson. 82*, 198–204.
Neri, P., Meadows, R., Gemmecker, G., Olejniczak, E., Nettesheim, D., Logan, T., Simmer, R., Helfrich, R., Holzman, T., Severin, J., & Fesik, S. (1991) *FEBS 294*, 81–88.
Ondek, B., Hardy, R. W., Baker, E. K., Stamnes, M. A., Shieh, B.-H., & Zuker, C. S. (1992) *J. Biol. Chem. 267*, 16460–16466.
Pflügl, G., Kallen, J., Schirmer, T., Jansonius, J. N., Zurini, M. G. M., & Walkinshaw, M. D. (1993) *Nature 361*, 91–94.
Piantini, U., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.
Ramachandran, G. N., & Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283–437.
Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1984) *Biophys. Biochem. Res. Commun. 117*, 479–485.
Schonbrunner, E. R., Mayer, S., Tropschug, M., Fischer, G., Takahashi, N., & Schmid, F. X. (1991) *J. Biol. Chem. 266*, 3630–3635.
Schreiber, S. L. (1991) *Science 251*, 283–287.
Shaka, A. J., & Freeman, R. (1983) *J. Magn. Reson. 51*, 169–173.
Siekierka, J. L., Hung, S. H. Y., Poe, M., Lin, C. S., & Sigal, N. H. (1989) *Nature 341*, 755–757.
Spitzfaden, C., Weber, H., Braun, W., Kallen, J., Wider, G., Widmer, H., Walkinshaw, M. D., & Wüthrich, K., (1992) *FEBS 300*, 291–300.
Steinmann, B., Bruckner, P., & Superti-Furga, A. (1991) *J. Biol. Chem. 266*, 1299–1303.
Takahashi, N., Hayano, T., & Suzuki, M. (1989) *Nature* **337**, 473–475.
Thériault, Y., Logan, T. M., Meadows, R., Liping, Y., Olejniczak, E. T., Holzman, T. F., Simmer, R. L., & Fesik, S. W. (1992) *Nature 361*, 88–91.
Trandinh, C. C., Pao, G. M., & Saier, M. H., Jr. (1992) *FASEB J. 6*, 3410–3420.
Vuister, G. W., & Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772–7777.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.
Walsh, C. T., Zydowsky, L. D., & McKeon, F. D. (1992) *J. Biol. Chem. 267*, 13115–13118.
Weber, C., Wider, G., von Freyberg, B., Traber, R., Braun, W., Widmer, H., & Wüthrich, K. (1991) *Biochemistry 30*, 6563–6574.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
Wüthrich, K., Billiter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.
Wüthrich, K., Spitzfaden, C., Memmert, K., Widmer, H., & Wider, G. (1991) *FEBS 285*, 237–247.
Xu, R. X., Olejniczak, E. T., & Fesik, S. W. (1992) *FEBS 2*, 137–143.
Yang, D., Rosen, M. K., & Schreiber, S. L. (1993) *J. Am. Chem. Soc 115*, 819–120.
Zuiderweg, E. R. P., & Fesik, S. W. (1989) *Biochemistry 28*, 2387–2391.
Zuiderweg, E. R. P., Boelens, R., & Kaptein, R. (1985) *Biopolymers 24*, 601–611.
Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., & Fesik, S. W. (1989) *J. Magn. Reson. 86*, 210–216.
Zydowsky, L. D., Etzkorn, F. A., Chang, H. Y., Ferguson, S. B., Stoltz, L. A., Ho, S. I., & Walsh, C. T. (1992) *Protein Sci. 1*, 1092–1099.