High Resolution Solution Structure of the 1.3S Subunit of Transcarboxylase from *Propionibacterium shermanii*†

D. Venkat Reddy,‡ Bhami C. Shenoy,§,|| Paul R. Carey,§ and Frank D. Sönnichsen*,‡

Department of Physiology & Biophysics, and the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Received November 3, 1999

ABSTRACT: Transcarboxylase (TC) from *Propionibacterium shermanii*, a biotin-dependent enzyme, catalyzes the transfer of a carboxyl group from methylmalonyl-CoA to pyruvate to form propionyl-CoA and oxalacetate. Within the multi-subunit enzyme complex, the 1.3S subunit functions as the carboxyl group carrier and also binds the other two subunits to assist in the overall assembly of the enzyme. The 1.3S subunit is a 123 amino acid polypeptide (12.6 kDa) to which biotin is covalently attached at Lys 89. The three-dimensional solution structure of the full-length holo-1.3S subunit of TC has been solved by multidimensional heteronuclear NMR spectroscopy. The C-terminal half of the protein (51−123) is folded into a compact all-β-domain comprising of two four-stranded antiparallel β-sheets connected by short loops and turns. The fold exhibits a high 2-fold internal symmetry and is similar to that of the biotin carboxyl carrier protein (BCCP) of acetyl-CoA carboxylase, but lacks an extension that has been termed "protruding thumb" in BCCP. The first 50 residues, which have been shown to be involved in intersubunit interactions in the intact enzyme, appear to be disordered in the isolated 1.3S subunit. The molecular surface of the folded domain has two distinct surfaces: one side is highly charged, while the other comprises mainly hydrophobic, highly conserved residues.

Transcarboxylase (TC)[1] from *Propionibacterium shermanii* occupies a unique position among biotin-containing enzymes, since it catalyzes the transfer of a carboxyl group from methylmalonyl-CoA to pyruvate to form propionyl-CoA and oxalacetate without the involvement of free CO_2, HCO_3-, or ATP (*1*). TC is a complex multi-subunit enzyme (1200 kDa), composed of 30 polypeptide chains of three different subunits. In the intact form of TC, there are a total of six 12S monomers (60 kDa) in a cylindrical central unit, twelve 5S monomers (60 kDa) in six dimeric outer subunits, and twelve 1.3S subunits (12.6 kDa) (*1*−*3*). The 1.3S subunit of TC is the biotin carboxyl carrier unit and consists of 123 amino acids. Biotin is covalently attached to the Lys 89 ϵ-amino group of 1.3S as a posttranslational modification by biotin holo-carboxylase synthetase (BHS) (*4*).

In the functional cycle of transcarboxylase (Scheme 1), the 1.3S participates in three heterologous protein−protein interactions. First, it serves as the substrate in the biotin ligation reaction catalyzed by biotin holo-carboxylase synthetase (*4*). Second, in its biotinylated form, holo-1.3S (1.3S-biotin) interacts with the biotin carboxylase subunit (12S) in the abstraction of CO_2 from the substrate methylmalonyl-CoA and its attachment to the ureido group of the biotin ring system. Finally, the carboxylated carboxyl carrier protein (1.3S-biotin-CO_2) interacts with carboxyl transferase subunit (5S) in which the CO_2 moiety is transferred to pyruvate to yield oxalacetate (Scheme 1).

The sequences of C-terminal regions among different biotin carboxyl carrier proteins are well-conserved, implying the existence of a common protein fold for these domains (*5*, *6*). Further, biotinyl domains have been shown to bear sequence and structural homology to the lipoyl domains of lipoic acid-dependent enzymes (*6*−*8*). Three-dimensional structures have been reported for six lipoyl domains (*9*−*14*), and one biotin carboxyl carrier protein (BCCP, from acetyl-CoA carboxylase) (*6*, *15*, *16*). The protein folds were shown to be quite similar, consisting of two four-stranded β-sheets.

Here, we report the structural characterization of the full-length holo-1.3S subunit of transcarboxylase (TC) from *Propionibacterium shermanii*, determined by multidimensional heteronuclear nuclear magnetic resonance (NMR) spectroscopy. While the N-terminal region of the protein, which has been shown to be involved in intersubunit interactions in the intact enzyme (*17*, *18*), appears to be disordered in the isolated 1.3S subunit, the C-terminal domain of 1.3S is folded into a compact β-sandwich structure. The domain exhibits a high 2-fold internal symmetry, and a fold similar to the homologous *Escherichia Coli* acetyl-CoA

|| Present address: Altus Biologics Inc., 625 Putnam Avenue, Cambridge, Massachusetts 02139-4807.

* To whom correspondence should be addressed: Frank D. Sönnichsen, Department of Physiology & Biophysics, 10900 Euclid Avenue, Case Western Reserve University, Cleveland, Ohio 44106-4970. Telephone: (216) 368-5405. Fax: (216) 368-1693. E-mail: frank@herring.phol.cwru.edu.

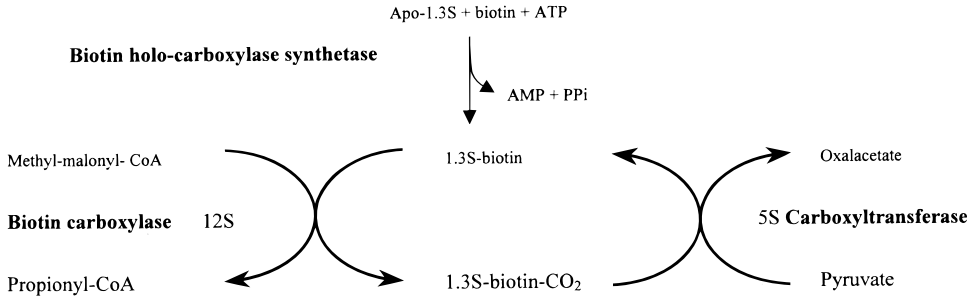
‡ Department of Physiology & Biophysics.

§ Department of Biochemistry.

[1] Abbreviations: 3D, three-dimensional; 4D, four-dimensional; BCCP, biotin carboxyl carrier protein; BHS, biotin holo-carboxylase synthetase; HSQC, heteronuclear single-quantum coherence; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser enhancement spectroscopy; rms, root-mean-square deviation; TOCSY, total correlation spectroscopy; TC, transcarboxylase.

10.1021/bi9925367 CCC: $19.00
Published on Web 02/17/2000

Scheme 1: The schematic representation of the three heterologous protein−protein interactions in which the 1.3S subunit participates in the functional cycle of transcarboxylase (TC).



carboxylase BCCP domain (*6*, *15*, *16*), but lacks an extension termed the "protruding thumb" found in BCCP. For 1.3S, we have demonstrated (*19*) that there is no evidence for an interaction between biotin and the protein. In BCCP, the thumb is in contact with the biotin, and the absence of a corresponding loop in 1.3S may explain our earlier NMR results (*19*).

MATERIALS AND METHODS

Sample Preparation. The singly ^{15}N-labeled 1.3S protein was obtained by the procedure described previously (*20*). The ^{15}N/^{13}C uniformly double-labeled 1.3S protein was produced by growing the cells at 37 °C in minimal medium, but containing $^{15}NH_4Cl$ and ^{13}C-glucose (Isotech, OH) as the sole source of nitrogen and carbon, respectively. Commercially purchased ^{13}C-labeled proline was added to the expression system to ensure high label incorporation for this residue type in protein. Unlabeled *d*-biotin was ligated to apo-1.3S by holo-carboxylase synthetase after prepurification (*4*), and the double-labeled holo-1.3S protein was isolated and purified as described previously for unlabeled samples (*21*). NMR samples contained ∼1.5−2 mM protein in 90% H_2O/10% D_2O (v/v) with the pH adjusted to 6.5.

NMR Spectroscopy. NMR spectra were recorded either at 20 or at 25 °C on a Varian Inova 600 MHz spectrometer equipped with pulsed field gradient units and triple-resonance probes with actively shielded z-gradients. The following experiments were performed on ^{15}N/^{13}C uniformly double-labeled 1.3S protein: 3D-HCCH-TOCSY (*22*), 3D-HC(CO)-NH-TOCSY (*23*), ^{15}N/^{13}C-edited 3D-NOESY (*24*), and ^{13}C/^{13}C-edited 4D-NOESY exeriments (150 ms mixing time) (*25*). The NMR data were processed and analyzed on IRIS Indigo2 Silicon graphics computers, using the nmrPipe software system (*26*) and analyzed with the program PIPP (*27*). Backbone ^{15}N, ^{1}H, and side chain ^{1}H assignments were already described elsewhere (*20*). ^{13}C assignments for Cα and side chain carbon atoms were obtained from the combination of HC(CO)NH-TOCSY and HCCH-TOCSY experiments. Previous ^{1}H assignments from 47 to 123 residues were confirmed and missing assignments completed from the analysis of the HCCH-TOCSY experiment. Residues 1−46 were not assigned due to weak amide proton intensity or absence of cross-peaks (*20*).

Structure Calculations. nOe-derived distance restraints were obtained from the ^{15}N-edited 3D-NOESY-HSQC, ^{15}N/^{13}C-edited 3D-NOESY, and ^{13}C/^{13}C-edited 4D-NOESY experimental data and were calibrated as described by Slupsky and Sykes (*28*). ϕ and ψ angle restraints were obtained from an analysis of the $d_{N\alpha}/d_{\alpha N}$ ratio according to Gagné et al. (*29*). $^{3}J_{HNH\alpha}$ coupling constants were calculated from a 3D ^{15}N-HNHA (*30*) experiment as described previously (*20*). χ_1 restraints and stereospecific assignments were determined with the help of a ^{15}N-edited short mixing 3D TOCSY-HSQC (35 ms), 3D HNHB (*31*), 3D NOESY-HSQC experiments, and a ^{15}N/^{13}C-edited 3D-NOESY experiment (*24*). Structure calculations for the 1.3S protein were performed using the standard DGSA protocol (*32*) in X-PLOR 3.8 software (*33*). A set of 1508 proton−proton distance restraints (552 long range, 119 medium range, 344 sequential, and 493 intra), 94 dihedral restraints (37 ϕ, 25 ψ, 32 χ1), along with the 47 $^{3}J_{HNH\alpha}$ coupling constants and 36 stereospecific assignments (for prochiral atoms/groups) were employed in the final structure calculations. From the final structure calculations, 32 structures were selected, which showed the lowest energies and no violations greater than 0.1 Å, 3°, or 2 Hz, and the overall quality of the structure was assessed using the program PROCHECK software (*34*). Biotin was not included in structure calculations, since no interactions were found between biotin and protein (*19*, *20*).

RESULTS AND DISCUSSION

The structure of 1.3S can be described as a protein having two domains with different structural features. The C-terminus (residues 51−123) of the 1.3S protein is well-structured and folded into a compact sandwich β-sheet structure, while the N-terminal of the protein appears to be disordered and has no detectable structure. This was first recognized in our previous relaxation studies where, in the ^{1}H−^{15}N heteronuclear steady state nOe experiments, the residues preceding the structured domain (residues from Gly 50 to the N-terminal) showed negative nOe ratios (*20*). However, the 1−50 N-terminal sequence of 1.3S is essential for the protein's function (*17*, *18*). Residues 32−55 are rich in Gly, Ala, and Pro residues and are predicted to form a flexible loop (or linker) in the intact enzyme. This region before the folded domain of 1.3S has some similarity to the linker regions, which follow the lipoyl domains in the sequences of E2-chains of dehydrogenases (*8*, *35*). The N-terminal 26 resides have been shown to be involved in binding with the other two subunits in the intact enzyme complex. In particular, residues 2−14 have been shown to bind to the central 12S subunit, and residues 15−26 have been shown to bind to the outer 5S subunit (*17*).

Structure Determination. The solution structure of the 1.3S subunit (residues 47−123 were used in the structure calcula-

tions) of TC was solved by multidimensional heteronuclear NMR spectroscopy, making use of uniformly labeled ^{15}N- and ^{15}N/^{13}C double-labeled proteins. A detailed account of the ^{1}H and ^{15}N resonance assignment procedure for the 1.3S using the uniformly ^{15}N singly labeled protein has already been presented elsewhere (*20*). Essentially complete ^{13}C resonance assignments for the Cα and side chain carbon atoms were obtained by running a set of standard heteronuclear NMR experiments using ^{15}N/^{13}C double-labeled 1.3S protein as described in the methods section. A total of 1508 inter proton distance restraints derived from 3D- and 4D-NOESY data were employed for structure calculations, including 552 long range, 119 medium range, 344 sequential, and 493 intraresidue nOes. A large number of nOes were observed for the residues that are found in the hydrophobic core (residues Ile 57, Val 65, Ile 68, Val 70, Val 82, Leu 83, Leu 85, Ile 94, Leu 119, Ile 120, and Ile 122) as explained below. A smaller number of nOes were observed for the turn regions (Figure 1a).

In the structure calculations (residues 47−123), a total of 94 experimental dihedral restraints including 37 ϕ, 25 ψ, and 32 χ_1 were used along with the 47 $^{3}J_{\mathrm{HNH}\alpha}$ coupling constants, and 36 stereospecific assignments of prochiral atoms/groups (Table 1). Biotin was not included in structure calculations, as no interactions were found between biotin and protein (*19*, *20*). Structure calculations were performed using the standard distance geometry simulated annealing (DGSA) (*32*) protocol in the X-PLOR programm (*33*). From the final 50 structures calculated, 32 structures were selected to represent the solution structure of the 1.3S, based on energetic criteria. These structures exhibit no violation greater than 0.1 Å, 3°, or 2 Hz, and their overall quality was assessed using the program PROCHECK (*34*). A Ramachandran plot (Figure 1d) for the averaged and energy minimized coordinates of 1.3S for the best defined region (55−123) shows 95% of the residues lying in the most favored regions of the ϕ, ψ conformational space. No residues were found in the disallowed regions, indicating the high quality of the structure.

A summary of the structural statistics is provided in Table 1, and a stereoview best-fit superposition of the final ensemble of 32 simulated annealing structures is shown in Figure 2a. A super-position of residues 55−123 of the ensemble to their mean coordinates yielded an average rms values of 0.33 and 0.79 Å for backbone and all heavy-atoms, respectively. Excluding the biocytin (Lys 89−biotin) turn that is slightly less well defined, the superposition of residues 56−86 and 93−123 shows a rms deviation of 0.28 Å for backbone and 0.74 Å for heavy atoms (Figure 1, b and c).

Description of the Structure. Overall Folding. The structure of the C-terminal of 1.3S is globular in shape and is formed by two very similar four-stranded antiparallel β-sheets, connected with small loops and turns (Figures 2b, 3a, and 4a), which are packed around a core of hydrophobic residues in a sandwich-like manner. These eight β-strands are designated sequentially as β1−β8 for the segments comprising residues Glu 56−Pro 58, Gly 63−Leu 69, Thr 75−Lys 77, Val 82−Glu 86, Thr 91−Asn 95, Gly 100−Leu 106, Ala 112−Gln 114, and Leu 119−Gly 123, respectively. The first β-sheet is comprised of strands β1, β3, β6, and β8, and the second β-sheet is formed by strands β2, β4, β5, and β7; with all strands arranged in an antiparallel fashion (Figures

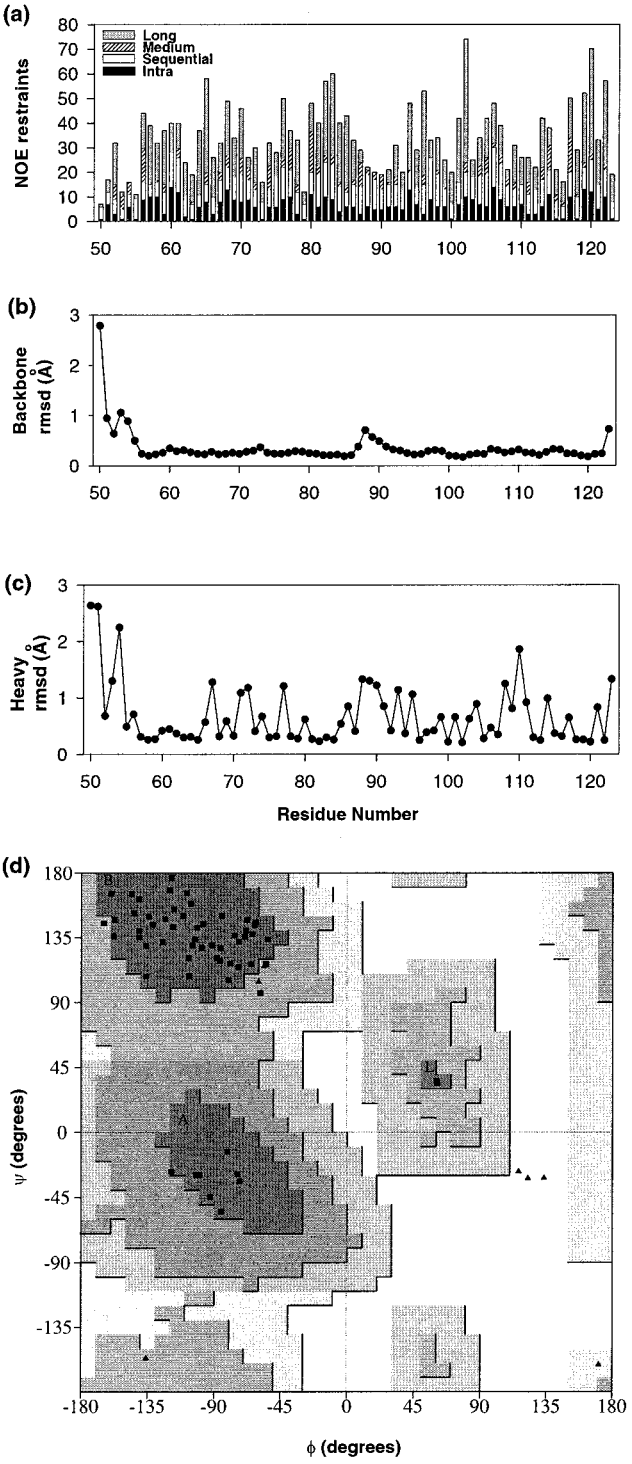


FIGURE 1: Structural data for the final 32 structures for 1.3S. (a) Distribution of intra, sequential, medium-range (2−5 residues), and long-range (>5 residues) nOe restraints per residue. Each nOe was counted twice, once for each of the residues involved in the interaction. The heights of the bars reflect the total number of nOe restraints for each residue. (b) Average root-mean-square deviation (rms) of backbone atoms and, (c) all heavy atoms for the 32 structures from the average structure, when the backbone of residues 55−123 are fitted. (d) Ramachandran plot of the ϕ, ψ angles observed in the averaged and energy-minimized coordinates of 1.3S (residues 55−123). Glycines are plotted as triangles. The shaded regions correspond to the most favored and allowed regions as defined by PROCHECK (*34*). More than 95% of the residues are found in the most favored regions, 5% in favored regions, and none in the disallowed regions, indicating the high quality of the structure.

Table 1: Structural Statistics of NMR Structures[a]

XPLOR potential energies (kcal mole^{-1})		
E_{total}	71.31 ± 4.04	68.05
E_{bond}	3.03 ± 0.41	2.51
E_{angle}	52.76 ± 2.00	51.58
$E_{improper}$	5.23 ± 0.26	5.81
E_{repel}	9.02 ± 1.40	7.55
E_{nOe}	1.09 ± 0.87	0.26
E_{cdih}	0.01 ± 0.01	0.00
E_{coup}	0.17 ± 0.41	0.33
rms deviations from experimental distance restraints (Å) (1508)		
intra (493)	0.0026 ± 0.0004	0.000
sequential (344)	0.0071 ± 0.0015	0.000
medium-range (119)	0.0120 ± 0.0019	0.004
long-range (552)	0.0059 ± 0.0007	0.002
rms deviations from experimental		
dihedral restraints (deg) 37 ϕ, 25 ψ, 28 χ	0.0168 ± 0.0044	0.000
rms deviations from $^3J_{HNH\alpha}$ (Hz) (*47*)	0.6581 ± 0.0057	0.695
rms deviations from idealized covalent geometry		
bonds (Å)	0.0016 ± 0.0001	0.0015
angles (deg)	0.4070 ± 0.0077	0.4031
impropers (deg)	0.2572 ± 0.0063	0.2711
measures of structure quality[b]		
Ramachandran plot (residues 55−123)		
residues in most favored regions(%)	86.8	94.7
residues in additional allowed regions(%)	13.1	5.3
residues in generously allowed regions (%)	0.1	0.0
residues in disallowed regions (%)	0.0	0.0
coordinate precision (Å)	backbone	heavy
residues 50−123	0.49 ± 0.08	0.90 ± 0.07
residues 55−123	0.30 ± 0.07	0.77 ± 0.06

[a] The notation of the NMR structures is as follows: ⟨*SA*⟩ are the final 32 simulated annealing structures; ⟨*SA*⟩ $_r$ is the energy-minimized structure of the mean structure obtained by averaging the coordinates of individual SA structures. [b] The overall quality of the structure was assessed using the program PROCHECK (*34*).

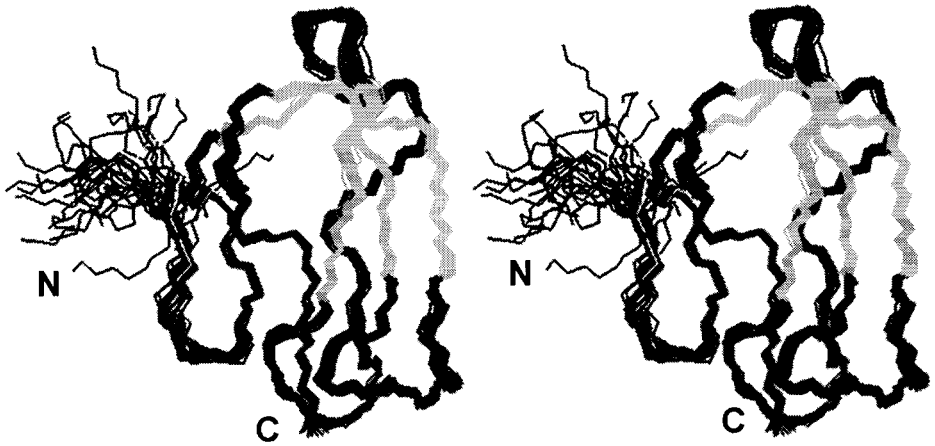


FIGURE 2: Stereoview of the backbone of final 32 refined structures of 1.3S (residues 47−123). The view was obtained by best-fit superposition using backbone atoms of residues 55−123 in the ensemble to the average structure (generated using the program InsightII/MSI, San Diego, CA).

2b, 3a, and 4a). The β-strands run alternatively from one sheet to another throughout the entire structure (except for β4−β5, which are part of the same β-sheet).

The observed NMR data indicate the presence of β-bulges in strands β2 (at Ser 66−Lys 67), β4 (at Val 82−Leu 83), β6 (at Glu 103−Lys 104) and β8 (at Leu119−Ile 120) (*20*). The presence of these β-bulges makes the respective strands form a curve/kink-like structure. The curvature of β2 makes it possible for it to pair up with β4 at one end, which in turn, pairs with β5 to form the flat region of the β-sheet. The β2 strand pairs with the β7 at the other end, which is almost perpendicular to the flat segment of the sheet and thus forms a cap of the β-sheet. The counterpart of β2, the β6, pairs with β8 at one end, which further pairs with β1. The β6 strand pairs with β3 at the other end and forms the second cap of the β-sheet (Figure 2b, 3a, and 4a). The core, formed by the two sheets, is tightly packed and composed principally of hydrophobic residues Ile 57, Val 102, Val 107, Leu 119, Ile 120, and Ile 122 from the first β-sheet, and residues Val 65, Ile 68, Val 70, Val 82, Leu 83, Leu 85, and Ile 94 from the second β-sheet. Most of these hydrophobic residues are very well-defined and presumably play an important role in stabilizing the β-sandwich structure. A large number of long-range nOes were observed across the two sheets.

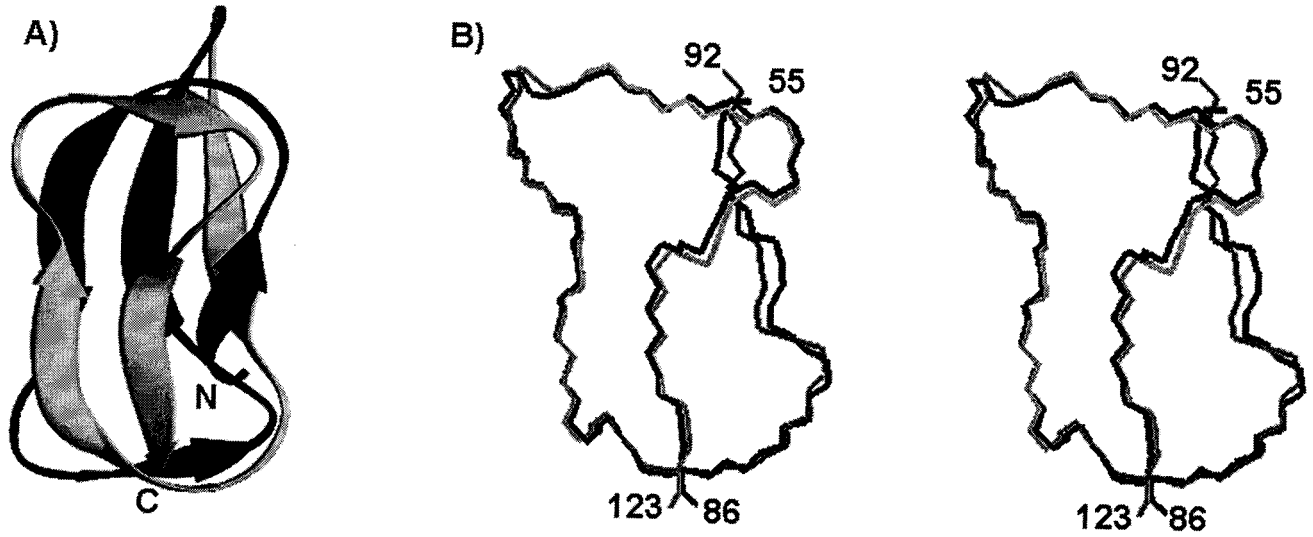


FIGURE 3: Figures showing the internal symmetry found in 1.3S: (a) ribbons presentation of the two sequential halves colored black and grey, (b) stereoview backbone superposition of the two halves (55−86 superimposed onto 92−123) as colored in (a).

The regular secondary structure elements were connected by four well-defined type II β-turns, one hairpin type I′ β-turn, and two short loops (Figures 2b, 3a, and 4a). The first loop consists of residues Ala 59−Ala 62, connecting the strands β1 and β2, and the second loop is constituted by residues Ala 96−Asp 99, connecting the strands β5 and β6, making a bridge between the two β-sheets at both ends. Residues Lys 71−Asp 74 and residues Lys 77−Gln 80 form two type II β-turns, creating a "turn-strand-turn like" motif around strand β3. Similarly, between strands β6 and β8, residues Glu 108−Asp 111, and residues Gln 114−Gln 117 form two type II β-turns completing another "turn-strand-turn like" motif along with strand β7 in the middle. These β-turns are stabilized by additional hydrogen bonds (*20*) (which, however, were not included in the form of constraints in the calculations).

The active site biocytin (biotinylated Lys-89) is located in a solvent exposed type I′ hairpin β-turn involving residues Ala 87−Met 90 between strands β4 and β5 (Figures 3a and 4a), connecting the N-terminal and C-terminal halves of the structured domain of 1.3S. Relatively high rms values (Figure 1, b, and c*)* were observed in the structure for the residues in this turn compared to the other well defined parts of the structure, indicating that this β-turn is relatively mobile. This supports our previous dynamic studies where the $^{1}H-^{15}N$ heteronuclear steady-state nOe experiments showed that this turn is somewhat flexible. Similarly, the Lys 89 side chain is not defined in the structure (except for a χ_1 preference for the −60 rotamer), which is in agreement with a negative $^{1}H-^{15}N$ heteronuclear steady-state nOe for the biotin-linked Lys 89 ϵ-NH, and the absence of nOes between the protein and biotin seen in this and a previous study (*19*).

Internal Symmetry. The structured domain of 1.3S is remarkably symmetric with three levels of symmetry. The first level is that related to the amino acid sequence. For the structured domain, the active site biocytin residue divides the sequence into N-terminal and C-terminal halves, which are very similar. Also, there are several conserved glycine and valine residues at identical positions in relation to the biocytin site. It has been speculated that these residues might be involved in apo-protein recognition by the holocarboxylase synthetase in the biotin ligation reaction (*36*, *37*). Moreover, these two halves centered around the biotinyl residue 89, exhibit nearly identical NMR properties (*20*). Consequently, the secondary structure elements derived from the NMR-data are nearly identical for these two regions. Backbone superposition of the two sequential halves (Figure 3 a and b) (residues 55−86 onto residues 92−123) yielded a rms value of 0.8 Å.

The second level of internal symmetry is that observed from a purely structural perspective. Two sheets of the β-sandwich make the two symmetrically related structural halves of the protein. Each β-sheet contains four antiparallel β-strands and two type II β-turns, which are similar in length (except strand β1) and are located in the similar positions of the other β-sheet (Figure 2, a and b). The third level of symmetry is the internal symmetry of the hammerhead-like structure. The structured domain of 1.3S contains two structural elements that each resemble a hammer, and which are again symmetric. This novel structural motif was also observed in BCCP and was named "hammerhead structure" (*6*) (Figure 3, a and b). Strands β6 and β8 in the first hammer and β2 and β4 in the second can be thought of as the hammer "handles", with the "hammerhead" formed by a "turn-strand-turn" motif. Each hammerhead (residues 107−118 and 70−81 in the first and second hammers, respectively) contains two type II β-turns (Lys 108−Asp 111 and Gln 114−Gln 117 in the first and Lys 71−Asp 74 and Lys 77−Gln 80 in the second) that are separated by the respective strands β7 and β3. The two "hammerhead" like units are remarkably similar and each "hammerhead" has its own internal symmetry (Figure 3a, b).

Structural Similarities to other Proteins. (a) BCCP. Athappilly and Hendrickson (*6*) determined the crystal structure of a C-terminal fragment (corresponding to residues 77−156 in the intact protein) of the biotin carboxyl carrier protein (BCCP), the biotinyl subunit of *E. coli* acetyl-CoA carboxylase. The fold of BCCP is very similar to the structure of 1.3S C-terminal folded domain (Figure 4, a and b). All strands and most turns are comparable in lengths and location, albeit some strands differ by one or two residues.

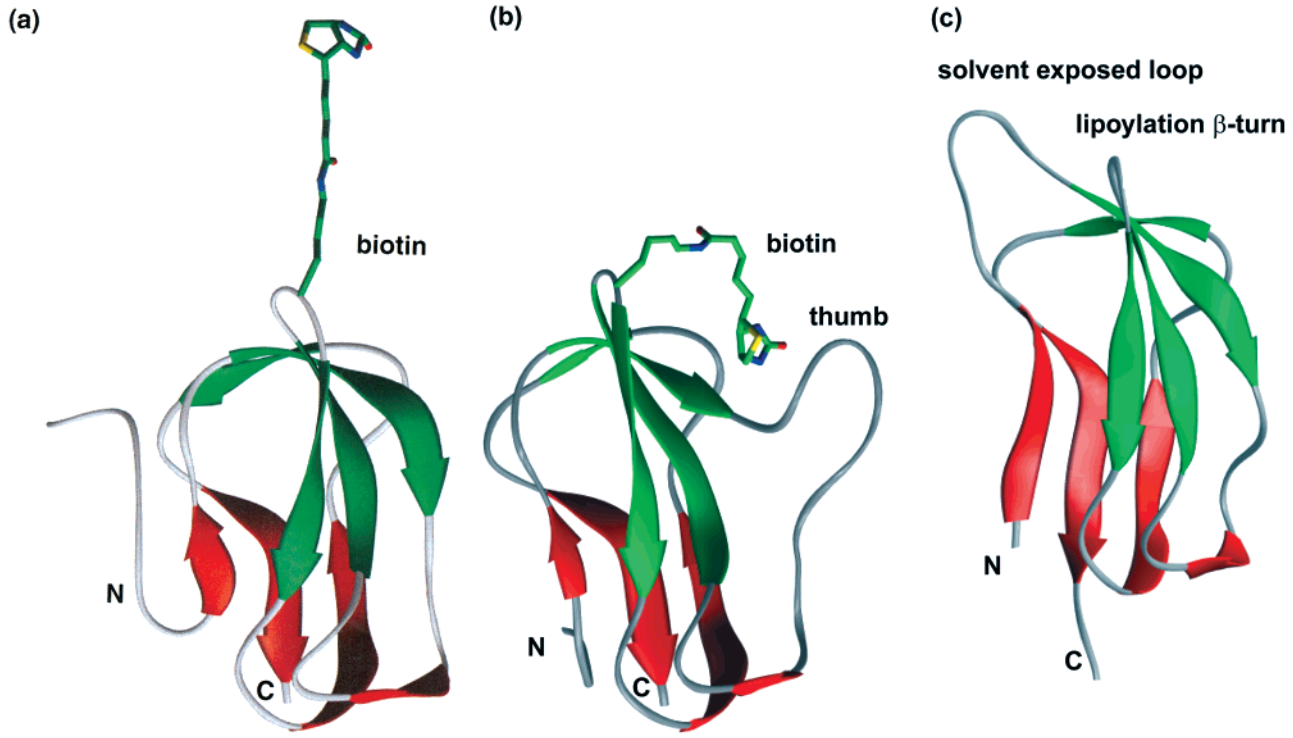


FIGURE 4: Structural comparison of 1.3S with BCCP and lipoyl domain proteins. (a) Ribbon diagram showing the structure of 1.3S protein. The biotin prosthetic group is attached arbitrarily, since no interactions were found between the biotin moiety and protein. The high internal 2-fold symmetry is seen by the two symmetry related strands in red and green colors. (b) Ribbon structure of the biotin carboxyl carrier protein (BCCP) from *E. coli* acetyl-CoA carboxylase (*6*), showing that the biotin is interacting with protein. (c) Ribbon structure of the apo-lipoyl domain of OGDHC from *A.vinelandi* (*12*).

In BCCP, the biotinylated Lys 122 (biocytin) is located at a hairpin β-turn between strands β4 and β5, connecting the N-terminal and the C-terminal halves. One striking feature of BCCP structure is that the biotin group is partially buried on the surface of the protein, with the ureido ring positioned between the protein core and a seven residue long "thumb-like" protrusion (with a type I β-turn in the center) formed between residues 94 and 101. In the present context, the "thumb" and its association with biotin is the most intriguing feature of this structure. The structure exhibits very specific interactions between the biotin and the protein moiety in this region, in particular two hydrogen bonds are formed between the side chain −OH and main-chain oxygen atom of Thr 94 with the ureido carbonyl (O2′) and N1′−H atoms of biotin, respectively. This indicates that in acetyl-CoA carboxylase the biotinyl group of BCCP is not free to "swing" between the active sites, at least in its noncarboxylated form, but is partially buried below the surface of the domain of BCCP.

More recently, solution structures of the C-terminal fragment of both apo- and holo- BCCP (residues 70−156) have been solved by NMR spectroscopy (*15*, *16*). It was found that the structure of the apo-domain is essentially identical to that of the holo-form, and the latter is in good agreement with that of the crystal structure of holo-BCCP solved by Athappilly and Hendrickson. Several nOe cross-peaks were observed between the biotin group and the protein, indicating that biotin−protein interactions occur in solution also, and that the biotinyl−lysine (K122) side chain is not freely moving in solution (*16*, *38*).

The structural comparison (Figure 4, a and b) between the holo-1.3S and BCCP clearly shows that there is no exposed loop/thumb present in 1.3S. In contrast to holo-BCCP, we found that in holo-1.3S the biotin is not interacting with the protein moiety (*19*). The absence of biotin−protein interactions in 1.3S can be explained by the absence of the "thumb", which is responsible for the majority of protein−biotin interactions in BCCP (*6*, *16*, *38*). Since there is no detectable interaction between biotin and protein in 1.3S, the biotin is apparently free to move in solution.

Thus, it is clear that the biotin prosthetic group is in different environment in 1.3S TC and in BCCP *E. coli* acetyl-CoA carboxylase. The significance of the biotin−protein interactions in BCCP and the relevance of the structural differences between BCCP and 1.3S are not known presently. However, it is worth remembering that the mechanisms of biotin carboxylation are different for the two holo proteins (*39*).

(b) Lipoyl Domains. The C-terminal sequences of biotin carboxyl carrier proteins show significant sequence similarity to the lipoyl domains from the dehydrogenase multienzyme complexes (*5*). Lipoyl domains comprise of approximately 80 residues (*8*) and contain a covalently bound lipoyl group linked to the Nϵ of a specific conserved lysine residue. The lipoyl domains function as an acetyl group carrier (from one component to another in the multi enzyme complex) similar to the biotin carboxyl carrier proteins. Recently, several three-dimensional structures of lipoyl domains have been solved by NMR spectroscopy (*9*−*14*). The structures are very similar showing two four-stranded antiparallel β-sheets similar to the biotin carboxyl carrier proteins. The ribbon structure in Figure 4c, of the apo-lipoyl domain of the 2-oxoglutarate dehydrogenase complex (OGDHC) from *Azobacter vinelandii* (*11*), is taken as an example for comparison.

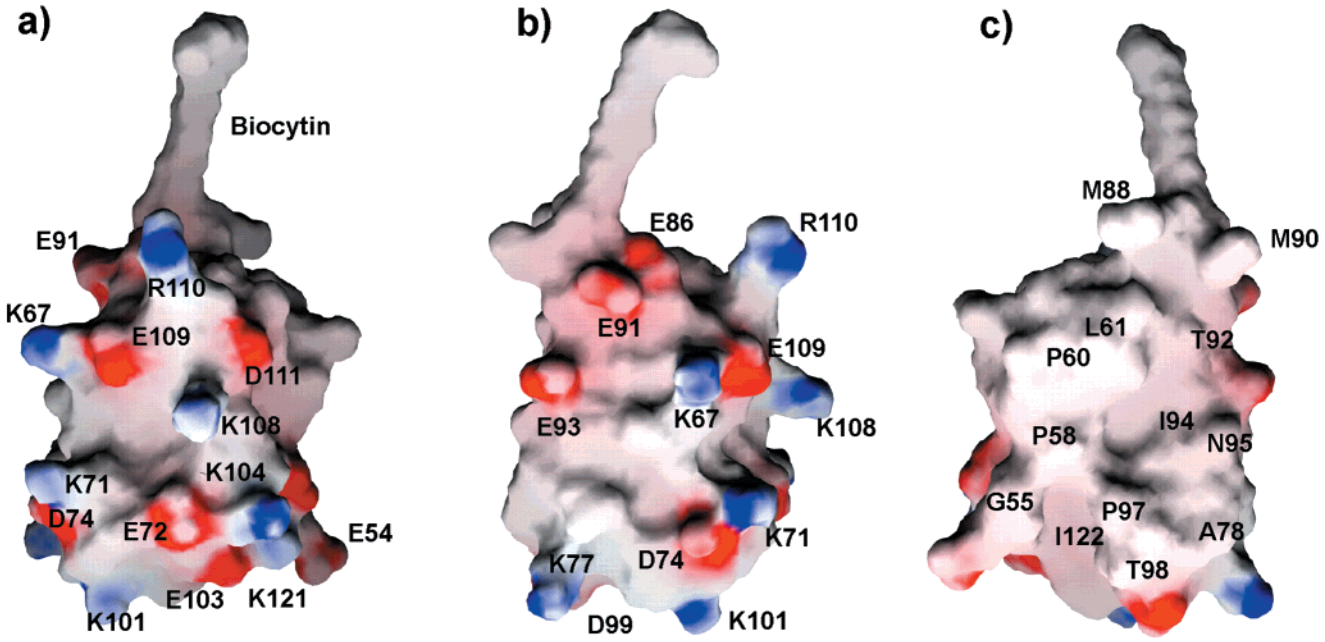


FIGURE 5: Distribution of the electrostatic potential on the molecular surface of 1.3S: (a) the hydrophilic face, (b) side view, and (c) the hydrophobic face of the protein. The surface map is generated using the program GRASP (*41*).

A comparison of 1.3S with that of the structure of lipoyl domain (OGDHC) reveals at least two interesting features (Figure 4, a and c). The overall fold of both proteins exists as a compact β-sandwich structure. However, the active site (biocytin) in 1.3S is located in a type I′ β-turn, whereas, in lipoyl domains, the corresponding β-turns at the lipoylation sites have a type I β-turn conformation. The second difference is that the lipoyl domains have an exposed loop, which lies close in space to the lipoylation site (Figure 4c).

Surface Characteristics of 1.3S and Functional Implications. The molecular surface of 1.3S exhibits (Figure 5, a, b, and c) two characteristic faces with clearly distinct properties. One face, formed by the edge of the two β- sheets (primarily residues 53−61 (β1) and 90−98 (β5)), and two turns of the "hammerhead" motifs (Ala 78/Gly 79 and Gly 115/Gly 116), is completely devoid of charges (Figure 5c). This region is composed of six glycines, three prolines, two methionines, two threonines, one alanine, one leucine, and one aspargine residue. Further, with the exception of aspargine and threonines, they all are nonpolar. Several of these glycine and proline residues are conserved among different biotin domains (*36*, *37*). The surface is relatively flat, as judged by the side view (Figure 5b), containing only a small groove near Ile 57 and Ile 122 (Figure 5c). This surface could dock onto a complementary surface on a 5S or 12S subunit, or biotin holocarboxylase syntetase (BHS), without loss of entropy, since the surface is devoid of mobile amino acid side chains. Such a docking could also give rise to 1.3S-subunit motion due to slippage of the surfaces, one over the other, since charge−charge interactions are absent.

The likelyhood that the "flat face" participates in protein−protein recognition is supported by the mutation/truncation studies of Samols et al. (*36*). The authors proposed that the recognition site for biotin holocarboxylase synthetase (BHS) in the biotinylation reaction involves Ile 122. The structure now shows that this residue is located in the apolar face (Figure 5c, at the bottom), with its side chain buried in the protein core.

The other face of the protein is polar. A number of charged residues are exposed, including putative salt bridges formed by Lys 121 and Glu 56, Lys67 and Glu 109, Lys 71 and Asp 74, Lys104 and Glu 72, Lys 108 and Asp 111, and Lys 77 and Asp 99 (Figure 5, a and b). Overall, an excess of negative charge distribution (five negative residues vs one free positive charge) is observed. In particular, a cluster of negative charges (Glu 86, Glu 90, Glu 93) are located in a row near the biotin binding site (Figure 5b). It is worth mentioning here that in the recent studies on the homologous protein BCCP from acetyl-CoA carboxylase, it was found that the substitution of the Glu with Lys (E119K) near the biotin binding site inactivates the protein as a substrate for biotin ligase (*40*). From this result and several other observations, it was suggested that charge maintenance might be particularly significant in the immediate vicinity of the biotin binding site. Glu 119 in BCCP is equivalent to Glu 86 in 1.3S, one of the three Glu residues comprising the cluster mentioned above (Figure 5b). Thus, the present structure adds credence to the notion that a negatively charged residue near the biotinylation site is needed for function.

Accession Numbers. The structures were deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) data bank under the accession codes 1DCZ (minimized average structure) and 1DD2 (ensemble of 32 structures).

ACKNOWLEDGMENT

We would like to thank Dr. Sven Rothemund, Dr. David Samols, and Rosa Rivera for many helpful discussions. The authors further thank Dr. Lewis Kay (University of Toronto), Dr. Stephane Gagné (University of Alberta), respectively, for making NMR pulse sequences and data processing scripts available to us, and F. Delaglio and D. Garrett for their software packages NMRpipe and PIPP. The present study

was supported by NIH Grant DK5305302 (P.R.C. and F.D.S.), NIH Grant GM 55362 (F.D.S.), and a Research Initiation Grant of the Ohio Board of Regents (F.D.S.).

REFERENCES

1. Wood, H. G. (1979) *Crit. Rev. Biochem. 7*, 143−160.
2. Maloy, W. L., Bowien, B. U., Zwolinski, G. K., Kumar, G. K., Wood, H. G., Ericsson, L. H., and Walsh, K. A. (1979) *J. Biol. Chem. 254*, 11615−11622.
3. Wood, H. G., and Kumar, G. K. (1985) *Ann. N.Y. Acad. Sci. 447*, 1−22.
4. Shenoy, B. C., and Wood, H. G. (1988) *FASEB J. 2*, 2396−2401.
5. Toh, H., Kondo, H., and Tanabe, T. (1993) *Eur. J. Biochem. 215*, 687−696.
6. Athappilly, F. K., and Hendrickson, W. A. (1995) *Structure 3*, 1407−1419.
7. Brocklehurst, S. M., and Perham, R. N. (1993) *Protein Sci. 2*, 626−639.
8. Berg, A., and de Kok, A. (1997) *Biol. Chem. 378*, 617−634.
9. Dardel, F., Davis, A. L., Laue, E. D., and Perham, R. N. (1993) *J. Mol. Biol. 229*, 1037−1048.
10. Green, J. D. F., Laue, E. D., Perham, R. N., Ali, S. T., and Guest, J. R. (1995) *J. Mol. Biol. 248*, 328−343.
11. Berg, A., Vervoort, J., and de Kok, A. (1996) *J. Mol. Biol. 261*, 432−442.
12. Berg, A., Vervoort, J., and de Kok, A. (1997) *Eur. J. Biochem. 244*, 352−360.
13. Ricaud, P. M., Howard, M. J., Roberts, E. L., Broadhurst, R. W., and Perham, R. N. (1996) *J. Mol. Biol. 264*, 179−190.
14. Howard, M. J., Fuller, C., Broadhurst, R. W., Perham, R. N., Tang, J. G., Quinn, J., Diamond, A. G., and Yeaman, S. J. (1998) *Gastroenterology 115*, 139−146.
15. Yao, X., Wei, D., Soden, C. Jr, Summers, M. F., and Beckett, D. (1997) *Biochemistry 36*, 15089−15100.
16. Roberts, E. L., Shu, N., Howard, M. J., Broadhurst, R. W., Chapman-Smith, A., Wallace, J. C., Morris, T., Cronan, J. E. Jr, and Perham, R. N. (1999) *Biochemistry 38*, 5045−53.
17. Kumar, G. K., Bahler, C. R., Wood, H. G., and Merrifield, R. B. (1982) *J. Biol. Chem. 257*, 13828−13834.
18. Shenoy, B. C., Kumar, G. K., and Samols, D. (1993) *J. Biol. Chem. 268*, 2232−2238.
19. Reddy, D. V., Shenoy, B. C., Carey, P. R., and Sönnichsen, F. D. (1997) *Biochemistry 36*, 14676−14682.
20. Reddy, D. V., Rothemund, S., Shenoy, B. C., Carey, P. R., and Sönnichsen, F. D. (1998) *Protein Sci. 7*, 2156−63.
21. Shenoy, B. C., Magner, W. J., Kumar, G. K., Phillips, N. F., Haase, F. C., and Samols, D. (1993) *Protein Expr. Purif. 4*, 85−94.
22. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Forman-Kay, J. D. (1993) *J. Magn. Res. B 101*, 333−337.
23. Logan, T. M., Olejniczak, E. T., Xu, R. X., and Fesik, S. W. (1993) *J. Biomol. NMR 3*, 225−231.
24. Pascal, S. M., Muhandiram, D. R., Yamazaki, T., and Forman-Kay, J. D. Kay L. E. (1994) *J. Magn. Res. B 103*, 197−201.
25. Kay, L. E. (1995) *Prog. Biophys. Mol. Biol. 63*, 277−299.
26. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
27. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) *J. Magn. Res. 95*, 214−220.
28. Slupsky, C. M., and Sykes, B. D. (1995) *Biochemistry 34*, 15953−64.
29. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) *Protein Sci. 3*, 1961−74.
30. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
31. Archer, S. J., Ikura, M., Torchia, D. A., and Bax, A. (1991) *J. Magn. Res. 95*, 636−641.
32. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
33. Brünger, A. T. (1992) X-PLOR. A System for X-ray crystallography and NMR. Yale University Press, New Haven, CT.
34. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
35. Perham, R. N. (1991) *Biochemistry 30*, 8501−8512.
36. Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C., and Wood, H. G. (1988) *J. Biol. Chem. 263*, 6461−6464.
37. Leon-Del-Rio, A., and Gravel, R. A. (1994) *J. Biol. Chem. 269*, 22964−22968.
38. Yao, X., Soden, C. Jr, Summers, M. F., and Beckett, D. (1999) *Protein Sci. 8*, 307−317.
39. Knowles, J. R. (1989) *Ann. Rev. Biochem. 58*, 195−221.
40. Chapman-Smith, A., Morris, T. W., Wallace, J. C., and Cronan, J. E. Jr (1999) *J. Biol. Chem. 274*, 1449−1457.
41. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins 11*, 281−296.

BI9925367