Structure of the Hexapeptide Xenobiotic Acetyltransferase from *Pseudomonas aeruginosa*[†,‡]

Todd W. Beaman,[§] Michele Sugantino,[§] and Steven L. Roderick*

Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received January 14, 1998; Revised Manuscript Received March 16, 1998

ABSTRACT: The crystal structure of the xenobiotic acetyltransferase from *Pseudomonas aeruginosa* PA103 (PaXAT) has been determined, as well as that of its complex with the substrate chloramphenicol and the cofactor analogue desulfo-coenzyme A. PaXAT is a member of the large hexapeptide acyltransferase family of enzymes that display tandem repeated copies of a six-residue hexapeptide repeat sequence motif encoding a left-handed parallel β helix (LβH) structural domain. The xenobiotic acetyltransferase class of hexapeptide acyltransferases is composed of microbial enzymes that utilize acetyl-CoA to acylate a variety of hydroxyl-bearing acceptors. The active site of trimeric PaXAT is a short tunnel into which chloramphenicol and the cofactor analogue desulfo-CoA project from opposite ends. This tunnel is formed by the flat parallel β sheets of two separate LβH domains and an extended 39-residue loop. His 79 of the extended loop forms hydrogen bonds from its imidazole NE2 atom to the 3-hydroxyl group of chloramphenicol and from its ND1 group to the peptide oxygen of Thr 86. The interactions of this histidine residue are similar to those found in the structurally unrelated type III chloramphenicol acetyltransferase and suggest that His 79 of PaXAT may be similarly positioned and tautomerically stabilized to serve as a general base catalyst.

Inspection of the amino acid sequence database reveals the existence of a family of proteins that are composed of imperfect tandem repeated copies of a hexapeptide sequence, generally described as [LIV]-[GAED]-X_2-[STAV]-X (*1*). This *hexapeptide repeat* motif (*2*, *3*) is characterized by the frequent occurrence of an aliphatic residue (Leu, Ile, or Val) at position *i*, a small residue (Ser, Thr, Ala, or Val) at position $i + 4$ and Gly at position $i + 1$. Proteins that contain tandem repeated copies of this six-residue repeat are called *hexapeptide proteins*. Hexapeptide proteins often display catalytic function as acyltransferases and participate in a variety of enzymatic processes, including cell wall biosynthesis, amino acid metabolism, and detoxification (*4*−*8*). They are found predominantly in microorganisms and higher plants. The hexapeptide acyltransferases utilize the phosphopantothenyl moiety of either acyl-coenzyme A (*6*) or the acylated acyl carrier protein (*4*, *5*), to which the transferred acyl group is attached in thioester linkage. The acyl groups transferred are either acetate (*6*), succinate (*7*), or *R*-3-hydroxy fatty acids (*4*, *5*). The acceptor substrates are amino acids, sugars, metabolic intermediates, or natural products bearing free hydroxyl or amine groups.

The crystal structures of three hexapeptide proteins have been determined. These include the *Escherichia coli* UDP-GlcNAc *O*-acyltransferase (*9*), a thermophilic carbonic anhydrase from *Methanosarcina thermophila* (*10*), and a bacterial tetrahydrodipicolinate *N*-succinyltransferase (*11*). The carbonic anhydrase represents an unusual example of a hexapeptide enzyme that does not function as an acyltransferase. The crystal structure of LpxA[1] revealed that the hexapeptide repeat sequence directs folding of a highly unusual structural domain termed a left-handed parallel β helix that is unique to hexapeptide proteins (*9*). The LβH domain is intimately involved in the subunit interface of these trimeric enzymes and has been suggested to function catalytically (*10*, *11*), although the three-dimensional structure of any hexapeptide acyltransferase in complex with substrates has yet to be reported.

It has become apparent in the last several years that there exists a homologous class of hexapeptide enzymes that utilize acetyl-CoA to *O*-acetylate a variety of natural products. These enzymes have been termed xenobiotic acetyltransferases, a class that can be further subdivided into two groups on the basis of shared amino acid sequence similarity and acceptor substrate specificity (*8*). Among the enzymes of the first group are the acetyltransferases from *Agrobacterium tumefaciens* C58 (*12*), the *E. coli* multiresistance transposon Tn2424 (*13*), and *Pseudomonas aeruginosa* PA103 (*14*). These enzymes are all able to acetylate chloramphenicol in

[†] This work was supported by Grants AI38328 and AI42154.

[‡] The coordinates corresponding to the *Pseudomonas aeruginosa* xenobiotic acetyltransferase structures have been deposited in the Brookhaven Protein Data Bank (identification codes 1XAT and 2XAT).

* Address correspondence to this author: Phone (718) 430-2784; Fax (718) 430-8565; E-mail roderick@aecom.yu.edu.

[§] Authors who contributed equally to this work.

[1] Abbreviations: PaXAT, *Pseudomonas aeruginosa* hexapeptide xenobiotic acetyltransferase; LβH, left-handed parallel β helix; LpxA, *Escherichia coli* UDP-*N*-acetylglucosamine acyltransferase; Cam, *Methanosarcina thermophila* carbonic anhydrase; DapD, tetrahydrodipicolinate *N*-succinyltransferase; Cm, chloramphenicol; CoA, coenzyme A; MIR, multiple isomorphous replacement; RMS, root-mean-square; σ, root-mean-square units; PHMPS, *p*-hydroxymercuribenzenesulfonic acid; CAT_{III}, type III chloramphenicol acetyltransferase; SA, simulated annealing.

S0006-2960(98)00106-8 CCC: $15.00
Published on Web 04/23/1998

Table 1: Data Measurement Statistics[a]

data set	unique reflections	observed reflections	completeness (%)	redundancy (%)	$\langle I/\sigma(I)\rangle$	R_{merge}	R_{iso}
native I	10 949	231 793	1.00	21.2	3.3	0.217	
cis-$Pt(NH_3)_2Cl_2$	5745	123 688	0.99	21.5	2.3	0.341	0.126
PHMPS	10 939	125 703	0.99	11.5	3.1	0.267	0.096
native II	10 448	34 784	0.95	3.2	13.6	0.078	0.146
Cm/desulfo-CoA	10 328	42 067	0.94	3.8	10.9	0.094	0.236

[a] X-ray data were measured using either highly redundant fast scans or less redundant slow scans (Experimental Procedures). The data sets are the fast scan data for native I and two heavy atom derivatives, cis-$Pt(NH_3)Cl_2$ (2 mM; 2 h soak) and PHMPS (0.2 mM; 90 min soak), and slow scan data for native II and Cm/desulfo-CoA (2 mM Cm, 50 mM desulfo-CoA). The resolution of the data is 3.2 Å for all of the data sets except for the cis-$Pt(NH_3)Cl_2$ data set (4.0 Å). The number of unique reflections, total number of observed reflections, fractional completeness, redundancy, ratio of individual intensities to their standard deviations, and the merging and isomorphous residuals are given. $R_{merge} = \sum|I_i - \langle I\rangle|/\sum|I_i|$ within one data set and $R_{iso} = \sum|F_{nat} - F_{der}|/\sum|F_{nat}|$ between native and derivative data sets. For R_{iso} calculations, the heavy atom derivative data sets are compared to native I and the Cm/desulfo-CoA data are compared to native II. The value of R_{iso} for native II is a comparison to native I.

Table 2: Heavy Atom Parameter Refinement and Phase Calculation Statistics[a]

derivative	occupancy	*X*	*Y*	*Z*	*B* (Å^2)	f_c/LOC	R_c
cis-$Pt(NH_3)_2Cl_2$	0.58	0.6001	0.0199	0.0060	29.4	1.46	0.60
PHMPS	0.83	0.5307	0.8887	0.1028	10.7	1.21	0.71

[a] The refined relative occupancies, fractional coordinates, and thermal factors (Å^2) resulting from the 3.2 Å heavy atom parameter refinement and isomorphous replacement phase calculation are shown. MIR phase calculations were conducted using cis-$Pt(NH_3)_2Cl_2$ data to 4.0 Å and PHMPS data to 3.2 Å resolution. f_c/LOC = RMS heavy atom structure factor amplitude/RMS lack-of-closure. $R_c = \sum||f_{H,obs}| - |f_{H,calc}||/\sum||f_{H,obs}||$ for the 1254 and 1958 centric reflections measured from the two heavy atom derivative crystals. Overall figure-of-merit for 10 674 reflections (0.97 complete) to 3.2 Å = 0.43.

vitro but are distinct from the well-known type III chloramphenicol acetyltransferase (*15*−*17*) and may acetylate other molecules in vivo. The second group of hexapeptide xenobiotic acetyltransferases has been associated with resistance to the virginiamycin class of antibiotics and they generally do not acetylate Cm (*18*−*24*).

We have previously reported the overexpression, purification, and crystallization of the xenobiotic acetyltransferase from *Pseudomonas aeruginosa* PA103, an enzyme composed of three identical subunits of 212 residues (23.4 kDa), and have demonstrated that the purified enzyme acetylates Cm (*14*). We report here the 3.2 Å resolution X-ray crystallographic structure of PaXAT and its complex with Cm and the cofactor analogue desulfo-CoA. These results define the active-site geometry and conformation of substrates bound to a hexapeptide acyltransferase and reveal active-site similarities to the nonhomologous CAT_{III} and to other members of the hexapeptide acetyltransferase family of enzymes.

EXPERIMENTAL PROCEDURES

Purification and Crystallization. The *Pseudomonas aeruginosa* xenobiotic acetyltransferase was overexpressed and purified as previously described and crystallized by the hanging drop method from solutions containing 10−20% poly(ethylene glycol) monomethyl ether 2000, 100 mM Tris, pH 8.5, and 10 mM $NiCl_2$. The crystals belong to the cubic space group $P4_132$ (a = 154.8 Å) (*14*). The largest crystals attain dimensions of 0.6 × 0.6 × 0.8 mm but diffract rather weakly, yielding X-ray reflections to 3.2 Å resolution using a laboratory rotating-anode X-ray source. Isomorphous crystals of PaXAT in complex with chloramphenicol and desulfo-CoA were prepared by soaking intact apoenzyme crystals for 3 days at ambient temperature in 20% poly(ethylene glycol) monomethyl ether 2000, 100 mM Tris, pH 8.5, 10 mM $NiCl_2$, 2 mM Cm, and 50 mM desulfo-CoA. The pH at which apoenzyme and complex crystals were prepared and X-ray data were measured is identical to the pH used to assay the enzyme (*14*).

Data Measurement and Phasing. X-ray diffraction data were measured using a Siemens X1000 area detector mounted on a Rigaku RU200 rotating-anode generator operating with fine focus at 50 kV and 80 mA. X-ray data were measured from single crystals at ambient temperature for a 36−48 h period over which the crystals were stable. These data sets were measured with either fast-scan (2−3 min) or slow-scan (15 min) exposure times for each 0.25° crystal rotation frame width and were reduced with either the XDS (*25*) or XENGEN (*26*) program packages (Table 1). The relatively high values of the merging residuals for fast-scan data are a consequence of the weak diffraction obtained from these crystals (unit cell volume 3.7 × 10^6 Å^3; solvent content 79%) and reflect the statistical uncertainty of individual observations. However, due to the high symmetry of these cubic crystals, the observed data were very redundant and produced acceptable merged data after averaging, as indicated by the typical values of the isomorphous residuals between merged native and derivative data sets as well as the lower merging residuals obtained from the slow-scan data for which fewer, but more accurate, individual observations were made.

Primary phases were determined by the method of isomorphous replacement using two heavy atom derivatives. Analysis of heavy atom difference Patterson functions immediately revealed the presence of one cis-$Pt(NH_3)_2Cl_2$ site (average vector density 2.1σ for 16 distinct self-vector positions) and also one distinct PHMPS site (3.0σ) for crystals soaked in solutions of these heavy atom compounds. The PHASES program package (*27*) was used to conduct heavy atom parameter refinement and multiple isomorphous replacement (MIR) phase calculations using cis-$Pt(NH_3)_2Cl_2$ data to 4.0 Å and PHMPS data to 3.2 Å resolution (Table 2). The mean overall figure-of-merit for the MIR protein phase determination to 3.2 Å resolution was 0.43 using

Table 3: Structure Refinement Statistics[a]

	apoenzyme	Cm/desulfo-CoA
resolution (Å)	8.0−3.2	8.0−3.2
data completeness (%)	95.1 (87.3)	93.8. (87.9)
protein atoms	1573	1565
substrate atoms		67
RMS deviation from ideality		
bond lengths (Å)	0.009	0.011
bond angles (deg)	1.4	1.4
temperature factors (Å^2)		
mean (all atoms)	21.7	18.0
mean (substrate atoms)		34.3
Wilson	24.8	22.7
R-free (%)	25.9 (27.7)	25.5 (25.3)
R-factor (%)	22.1 (26.4)	21.3 (25.2)

[a] Refinement statistics for PaXAT apoenzyme and substrate complexes are shown. The X-ray data used in these refinements were the slow-scan data sets (Experimental Procedures). *R*-free (%) = $\sum|F_o - F_c|/\sum|F_o|(100)$ for a 5% subset of X-ray diffraction data omitted from refinement calculations. *R*-factor (%) = $\sum|F_o - F_c|/\sum|F_o|(100)$ for all available data, including data reserved for the calculation of *R*-free. Values in parentheses refer to the corresponding statistic calculated for data in the highest resolution bin (3.39−3.20 Å).

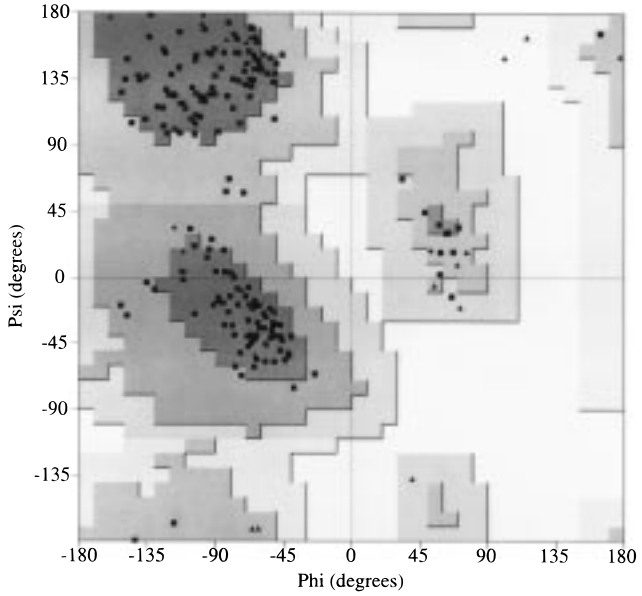

FIGURE 1: Ramachandran diagram (*35*) of the main-chain torsion angles ϕ and ψ for the refined apoenzyme model of PaXAT. The shading indicates levels of conformational energy for non-glycine residues, with white signifying energetically forbidden regions. Glycine residues are depicted as filled triangles. The figure is based on the output from PROCHECK (*34*).

10 674 reflections (0.97 complete).

Although the MIR electron density map calculated on the basis of the two heavy atom derivatives showed evidence of the expected and distinct LβH structural domain, it was clear that this map was not of good quality. Solvent flattening procedures were therefore carried out using either PHASES (*27*) or the program DM of the CCP4 suite (*28*). Since the composition of the asymmetric unit was not known, and rotation function analyses did not clearly indicate the presence of molecular symmetry, assumed solvent volumes corresponding to one, two, or three subunits per asymmetric unit were used in these trial calculations. Inspection of these solvent-flattened maps revealed clearly that the asymmetric unit was composed of just one subunit (79% solvent volume) and that the program DM, employing solvent flattening and histogram matching, worked best and improved the quality of electron density dramatically.

Model Building and Atomic Parameter Refinement of PaXAT. The solvent-flattened 3.2 Å resolution electron density map supported a polypeptide chain trace using the graphics program O (*29*). The initial model building efforts were aided by the known three-dimensional structure of the coiled LβH domain and the rather large fraction of aromatic residues in PaXAT (13%) and resulted in a model that included 195 of 211 residues of a single subunit. Approximately 5% of the 3.2 Å resolution diffraction data were assigned to a test set and omitted from refinement procedures in order to calculate the refinement working set residual R_{work} as well as the cross-validation statistic R_{free} (*30*, *31*). R_{work} for the model prior to the first parameter refinement was 40.7% (R_{free} 40.6%) for 1299 atoms (0.79 complete).

Simulated annealing atomic model parameter refinements were conducted with the X-PLOR program package (*32*) using 3.2 Å data and a single overall thermal factor. The first SA refinement reduced R_{work} to 28.3% (R_{free} 34.0%). This model was rebuilt and a second SA refinement reduced R_{work} to 23.7% (R_{free} 29.3%) for 1529 atoms (0.93 complete). At this stage, a series of 11 sequential omit maps were calculated, each preceded by a 4000 K SA refinement of nonomitted atoms in order to reduce model bias. These maps served to confirm the polypeptide chain trace and to correct side-chain positioning errors. The conjugate direction procedures encoded in the TNT suite of computer programs (*33*) and Powell refinement of X-PLOR were then used in the final several cycles of positional refinement against the apoenzyme slow-scan data (native II; Table 1), retaining the same subsets of reflections for calculation of the working set and cross-validation residuals. A final grouped refinement of thermal factors (one value per residue) reduced R_{work} to 22.1% (R_{free} 25.9%; Table 3).

To determine the active-site location and conformation of bound substrates, intact apoenzyme crystals were soaked in a crystal stabilization solution containing Cm and desulfo-CoA. Previous experiments using either acetyl-CoA or CoA did not clearly reveal the conformation of the pantetheine arm, perhaps due to hydrolysis and oxidation of the free thiol group by the Ni^{2+} present in the crystal soaking solution. The location and conformation of bound substrates was revealed most clearly in an unbiased $F_o - F_o$ map calculated using the apoenzyme coordinates as a phase source and the observed substrate complex and apoenzyme amplitudes. Least-squares refinement of the substrate complex model was carried out using the Powell refinement procedure of X-PLOR while maintaining the same set of reflections for the calculation of R_{work} and R_{free} as for the apoenzyme structure refinement. Cyclic model building and refinement reduced R_{work} to 21.3% (R_{free} 25.5%; Table 3).

RESULTS AND DISCUSSION

Atomic Model of PaXAT. Structure refinement statistics for the apoenzyme and substrate complex models of the *Pseudomonas aeruginosa* xenobiotic acetyltransferase are given in Table 3. The stereochemical quality of the model is good, with RMS deviations of bond lengths and bond angles from ideality of 0.009 Å and 1.4° for the apoenzyme model. The program ProCheck (*34*) was used to monitor

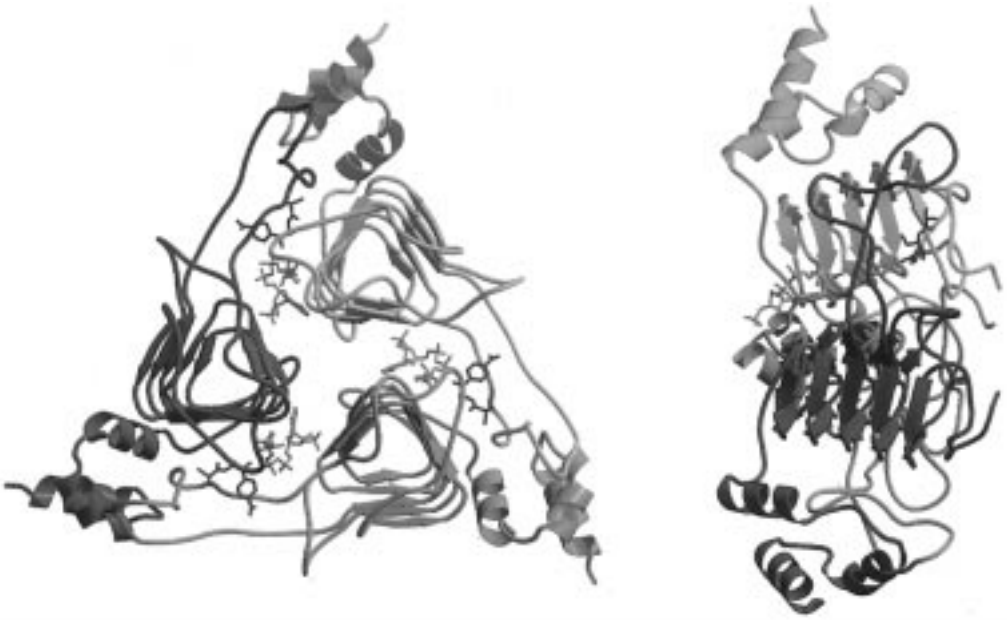

FIGURE 2: PaXAT in complex with chloramphenicol (blue) and desulfo-CoA (gray). (Left) Viewed parallel to the crystallographic 3-fold axis that relates subunits of the trimer. (Right) Viewed perpendicular to the 3-fold axis. Drawn by MOLSCRIPT (*41*) and Raster3D (*42*, *43*).

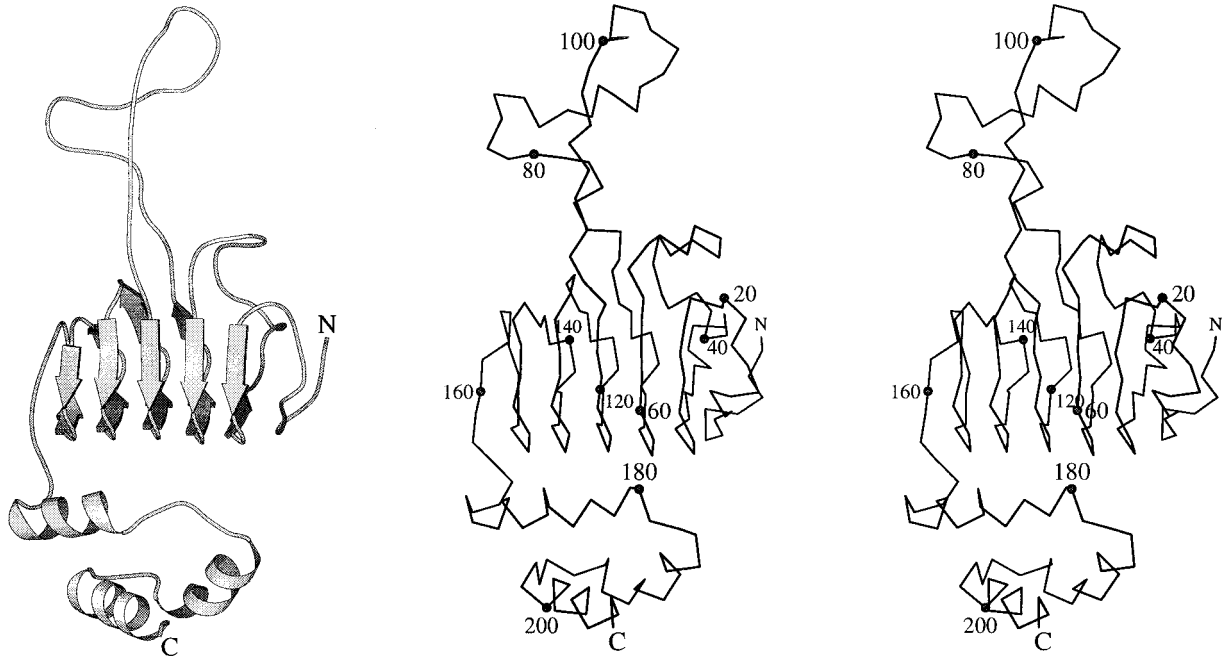

FIGURE 3: Polypeptide chain conformation of a PaXAT subunit. (Left) Ribbon diagram. (Right) Divergent stereoview α-carbon diagram.

the quality of the model, which was found to be better in all geometric categories for both main-chain and side-chain groups compared to other structures determined to similar resolution. A Ramachandran plot (*35*) (Figure 1) of the apoenzyme main-chain torsion angle pairs ϕ and ψ shows that 7.1% of all nonglycine residues are placed in the outlier region of the plot, a better fraction than for 20 of 26 3.2 Å resolution structural models in the Brookhaven Protein Data Bank (*36*). The final refined model is also sterically and chemically consistent with the positions of heavy metals used in MIR phasing. A difference Fourier map calculated with phases from the refined model of PaXAT identified Cys 43 as the residue to which PHMPS is bound (18σ $F_o - F_o$ map density) and Met 126 as the cis-$Pt(NH_3)_2Cl_2$ binding site (22σ). An additional lower occupancy platinum site was also located at this stage bound to Met 48 (9σ).

Oligomeric Structure and Crystal Packing. The oligomeric structure of most enzymes containing the LβH domain fold is trimeric (*9*−*11*) and the elution volume for PaXAT from gel-filtration columns is consistent with this structural assignment (data not shown). The three-dimensional structure of PaXAT determined here is also based on a trimeric arrangement with each subunit situated near a crystallographic 3-fold rotation axis (Figure 2). The overall dimensions of the trimer are 85 × 85 × 32 Å with the shortest dimension parallel to the 3-fold axis. Contacts between subunits that produce the cubic crystals studied here are of three types. The first is the association of the three

subunits surrounding a crystallographic 3-fold axis to form the trimeric enzyme. The second is the association of two distinct trimers in a "head-to-head" arrangement to form a hexamer with 32-point symmetry. The centroid of this hexamer is located at the crystallographic coordinate ($^3/_8$, $^3/_8$, $^3/_8$); a location in the crystal at the intersection of a crystallographic 3-fold rotation axis and three 2-fold axes. The third type of contact is made between the C-terminal α-helical domain and the extended loop domain of a single subunit (see below) with similar residues from a distinct hexamer.

Overall Polypeptide Chain Fold. A single subunit of the enzyme is formed by four domains (Figure 3). An NH_2-terminal domain corresponding to residues 3−22 is involved in formation of the crystalline hexamer and caps an LβH domain of five coils (23−153). The coiled fold of the LβH structural domain is interrupted by an extended loop domain formed by residues 72−110 and is terminated at its C-terminal coil by a β-hairpin loop (154−160). A C-terminal domain is composed predominantly of three α-helices in the residue ranges 167−176, 183−193, and 198−210. The overall polypeptide chain fold of PaXAT is unrelated to that of the type III chloramphenicol acetyltransferase (*16*).

Left-Handed Parallel β Helix Domain. The largest structural domain of PaXAT is the LβH structural domain (residues 23−153). This domain corresponds to the imperfect tandem repeated hexapeptide sequence and folds as a large coiled prism as if wound in a left-handed spiral around the surface of an equilateral prism (Figures 3 and 4). The faces of the LβH are formed by three extremely flat parallel β-sheets, each of which is composed of either three or five short β-strands. The first coil of the LβH domain is irregular and includes a residue range (32−56) that does not form interstrand hydrogen bonds with residues from the succeeding coil. The crossover connections between the parallel β-strands of this domain are all left-handed, an extremely rare occurrence in proteins not possessing the LβH structural domain (*37*). Direct connections between these strands are made by tight turns that are similar in conformation to classical type II tight turns (*9*, *38*), placing main-chain amide and carbonyl groups of the central peptide units in position to hydrogen bond to similar groups in the turns of adjacent coils.

The pleat of the three β-sheets that comprise the LβH structural domain causes side chains to either project into the lumen of this domain or outward and arranges amino acid side chains in stacks. A structure-based amino acid sequence alignment depicts the arrangement of these structurally equivalent residues in adjacent coils of the LβH domain and defines a nomenclature for the parallel β-strands (PB1, PB2, and PB3) and tight turns (T1, T2, and T3) of each coil (Figure 5). The conserved hydrophobic residues at position *i* (Leu, Ile, and Val) of the hexapeptide repeat motif ([LIV]-[GAED]-X_2-[STAV]-X) project into the interior of the LβH domain (Figure 4). Valine and isoleucine residues at this position frequently adopt staggered side-chain conformations with respect to the main chain (side-chain torsion angles $\chi_1 = 180°$ and $-60°$, respectively). The structure of PaXAT is unusual in that two aromatic residues (Phe 71 and Phe 125) are present at the *i* position of the hexapeptide repeat sequence, which is ordinarily occupied by aliphatic residues (Leu, Ile, or Val). These substitutions

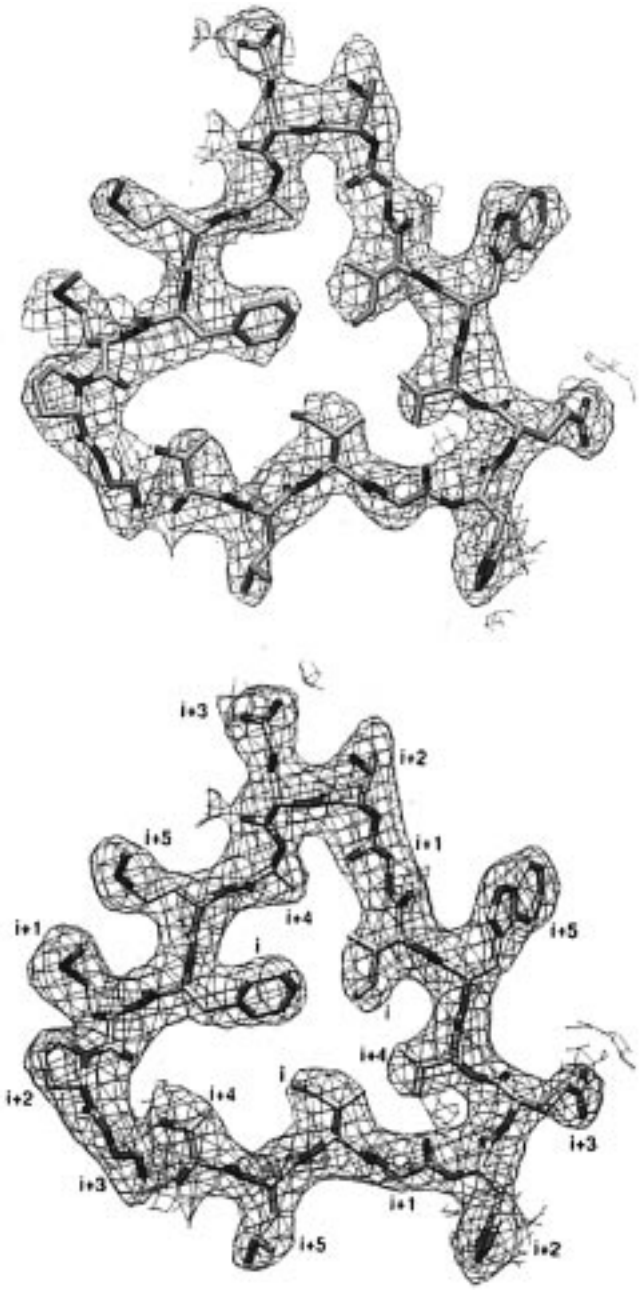

FIGURE 4: Left-handed parallel β helix (LβH) domain of PaXAT. Residues 111−128 (coil C3) of the LβH from the final refined model of the apoenzyme are depicted. (A, top) Electron density (red) from the solvent-flattened MIR map used for initial fitting. (B, bottom) Density from the final $2F_o - F_c$ map (blue). The hexapeptide repeat residue types (*i*, $i + 1$, ...) are also indicated (see text and legend to Figure 5). This figure and Figure 6 were drawn by SETOR (*44*).

are accommodated by a widening of the LβH near the T3 turn position of these coils, which relieves otherwise close contacts to preceding residues at the *i* and $i + 4$ positions and also allows position $i + 4$ to accommodate an unusual leucine residue type at position 57. Residues at position $i + 4$ are usually small (Ser, Thr, Ala, or Val) and project into the LβH at the triangular corners of each coil. Polar residues at this position form hydrogen bonds to inward-pointing main chain peptide groups in the tight turn of the succeeding coil, while nonpolar residues allow ordinary hydrogen-bond formation between main-chain peptide groups across these turns. No charged residues are present in the

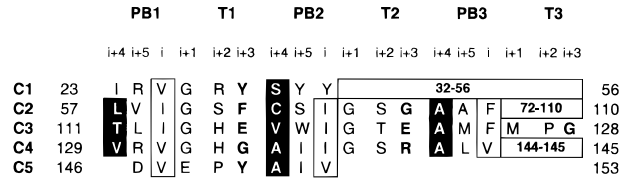

FIGURE 5: Structure-based amino acid sequence alignment of the LβH domain of PaXAT (residues 23−153) identifying structurally equivalent residues in each of the five coils. Each horizontal line represents one complete or partial coil. The parallel β strands that form the planar faces of the LβH are denoted PB1, PB2, and PB3 and the turns separating these strands are denoted T1, T2, and T3. The residues types within each hexapeptide unit are indicated (*i*, *i* + 1, ...). The conserved hydrophobic residues at position *i* are boxed. Residues in left-handed conformation (main-chain torsion angle $\varphi > 0$) at position $i + 3$ are depicted in boldface type. The small residues at the corners of each coil at position $i + 4$ are reverse shaded. Residues not involved in the coiled β-helical structure are enclosed in boxes and include the extended loop domain formed by residues 72−110.

interior (*i*, $i + 4$) positions of the LβH domain of PaXAT, as has been found for all other known hexapeptide protein structures. An additional structural determinant of the LβH domain occurs at position $i + 3$. Residues at this position project outward from the LβH and adopt a left-handed conformation (main-chain torsion angle $\varphi > 0°$) for 9 of 9 residues, as has been observed in other proteins containing the LβH domain (9−11).

The axis of each LβH domain makes an angle of 21° with the 3-fold axis of the trimeric enzyme. This degree of inclination is unusual for hexapeptide proteins, all of which previously had been found to adopt a very nearly parallel arrangement of their LβH domain axes with the trimeric 3-fold axis (e.g., 1−2° for LpxA and Cam, 3−4° for DapD). This may be related to the shorter length of the LβH domain of PaXAT (five coils) compared to LpxA (10), Cam (seven), and DapD (seven), which would allow for comparatively fewer inter-LβH interactions across the 3-fold axis as a result. Subunit−subunit interactions distant from the 3-fold axis involve the extended loop (residues 72−110) of one subunit with the LβH and C-terminal domains of an adjacent subunit. The amino acid sequence of this extended 39-residue loop does not obey the hexapeptide repeat sequence rule and its conformation is not accommodated in the coiled LβH fold, instead exiting from and reentering the LβH at the same tight turn position of the second coil (Figures 3 and 5). Such loops are a common structural feature of hexapeptide proteins, although the extended 39-residue loop of PaXAT is an unusually long example. These loops may represent an important means by which hexapeptide proteins have attained structural and functional diversity in the context of the largely invariant structure of the coiled LβH domain.

Active Site of PaXAT. The refined structure of PaXAT in complex with Cm and desulfo-CoA defines the location of the active site and illustrates its use of residues from two subunits (termed A and B). The ligands bind to the opposite ends of a short tunnel formed by the extended loop from subunit B (72B−110B) and residues from the LβH domains of both the A and B subunits. This tunnel is given accessible volume by the presence of glycine residues at the $i + 1$ position of the β-sheet PB2 of the LβH in three successive coils (residues Gly 66A, Gly 120A, and Gly 138A; Figure 5) and the first residue of the C-terminal domain, Gly 154A.

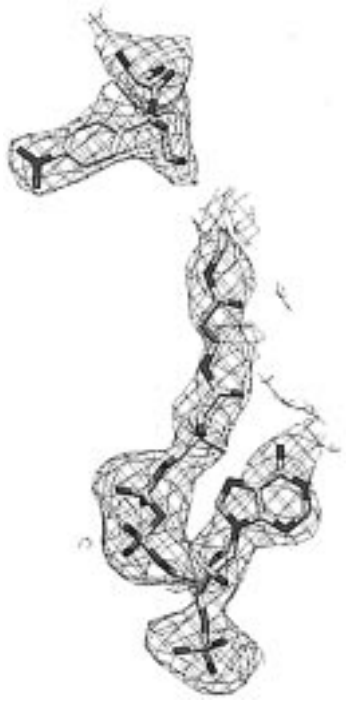

FIGURE 6: Refined coordinates of Cm and desulfo-CoA superimposed with the unbiased electron density (2.0σ) calculated using the apoenzyme model as a phase source and $F_o - F_o$ coefficients.

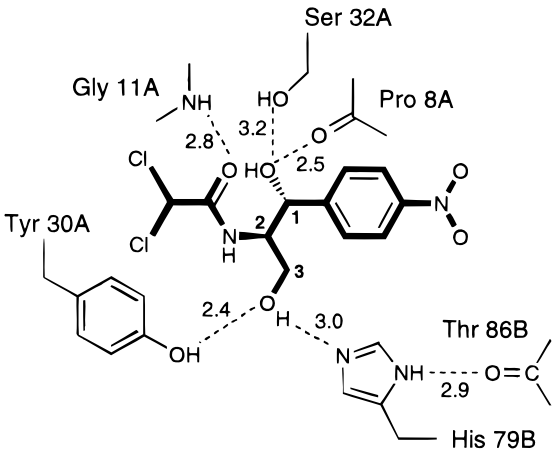

FIGURE 7: Schematic diagram of the chloramphenicol binding site. Hydrophilic contacts are indicated by dotted lines with associated distances (in angstroms).

The main-chain conformations of the apoenzyme and substrate complex structures are nearly identical, with an RMS discrepancy between paired α-carbon atoms of 0.24 Å for residues 3−210, indicating that little conformational rearrangement of the main chain is required to bind substrates. Only slight rearrangements of side-chain positions accompany ligand binding.

The binding site for the cofactor analogue desulfo-CoA is located at the C-terminal end of the active-site tunnel and includes residues from both the A and B subunits. The cofactor conformation is bent at the pyrophosphate group and adopts the conformation of a fishhook (Figure 6). The 3′-phosphate ADP moiety binds to residues in the C-terminal domain and to the C-terminal end of each LβH domain. The extended pantetheine arm is oriented parallel to the 3-fold axis of the enzyme and points directly toward the Cm binding site. The ribose 2′-hydroxyl is within hydrogen-bonding

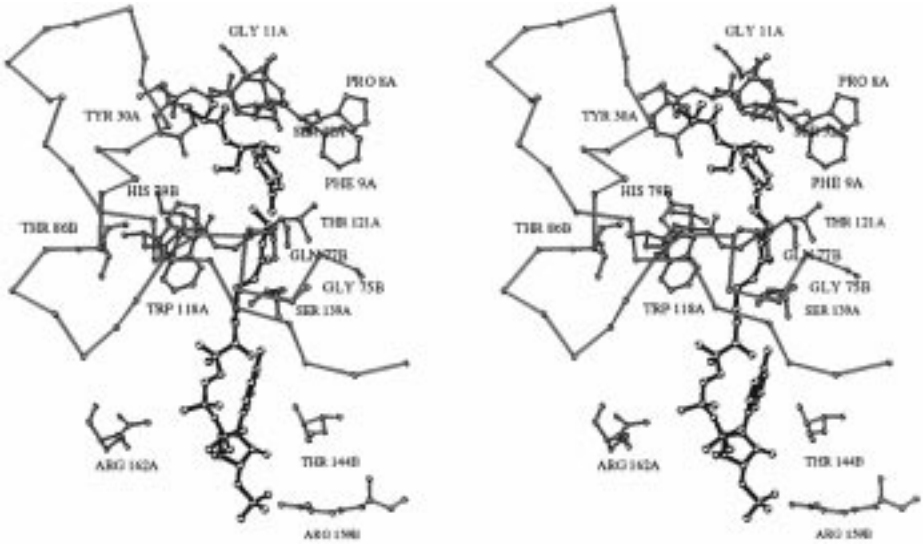

FIGURE 8: Divergent stereoview of the active site of PaXAT. Residues from the A subunit (blue) and B subunit (green) are depicted, as are the substrate Cm and cofactor analogue desulfo-CoA (black).

distance of Thr 144B and the 3′-phosphate contacts Arg 159B. The pyrophosphate group interacts with Arg 162A. The amide carbonyl oxygen atoms of the pantetheine β-alanyl units are within hydrogen-bonding distance of the main-chain nitrogen of Thr 121A and the hydroxyl and main-chain nitrogen groups of Ser 139A. The amide nitrogen of the terminal β-alanine residue is within hydrogen-bonding distance of the side-chain oxygen of Gln 77B. The adenine ring is directed toward the protein and its exocyclic amino group is within hydrogen-bonding distance of the peptide oxygen of Ser 139A. The overall bent conformation of the cofactor, the extended pantetheine arm, the solvent-exposed 3′-phosphate group, and the direction of the adenine group toward the protein in anti glycosidic linkage are all frequently observed features of protein-bound CoA conformations (*39*).

The Cm binding site is located at the NH_2-terminal end of the active-site tunnel and consists chiefly of residues from the NH_2-terminal 32 residues of the A subunit. The solvent accessibility of Cm is reduced by 95% on binding to PaXAT, largely due to its binding position under a loop formed by the NH_2-terminal 12 residues of the A subunit and its insertion into one end of the active-site tunnel. The observed conformation of Cm bound to PaXAT cannot be superimposed on that of Cm bound to CAT_{III}, differing by 120° torsion angle rotations about the C1−C2 and C2−C3 bonds. The C-1 hydroxyl group of Cm bound to PaXAT is within hydrogen-bonding distance of the hydroxyl of Ser 32A and the peptide oxygen of Pro 8A (Figures 7 and 8). The Cm carbonyl oxygen accepts a hydrogen bond from the peptide nitrogen of Gly 11A and its *p*-nitrophenyl ring participates in a perpendicular ring interaction with Phe 9A. The primary hydroxyl group of Cm at the C-3 position is within hydrogen-bonding distance of the phenolic hydroxyl of Tyr 30A (2.4 Å) and the imidazole NE2 atom of His 79B (3.0 Å), an extended loop residue of the B subunit. His 79B also donates a hydrogen bond from its imidazole ND1 to the peptide oxygen of Thr 86B and participates in a parallel ring stacking interaction with the indole of Trp 118A belonging to the LβH domain of an adjacent subunit. The distance between the C-3 hydroxyl of Cm and the terminal methyl group of the desulfo-CoA is 4.8 Å and a cavity is located between these binding sites. Modeling experiments indicate that this cavity is of sufficient volume to accommodate the acyl group of acetyl-CoA and that the carbonyl carbon of its thioester group can be modeled in several different ways to bring it in close proximity to the primary C-3 hydroxyl group of Cm.

The importance of His 79 is indicated by its conservation in the sequence of all xenobiotic acetyltransferases and the inactivation of *Agrobacterium tumefaciens* xenobiotic acetyltransferase by the site-directed replacement of this residue to alanine (*8*). The indole ring of Trp 118A stacks against His 79B in the structure of PaXAT and this tryptophan is also conserved among xenobiotic acetyltransferases as well as two hexapeptide acetyltransferases that are not members of the xenobiotic class—*E. coli* galactoside acetyltransferase (*40*) and *Rhizobium leguminosarum* lipooligosaccharide acetyltransferase (*6*). The galactoside acetyltransferase homologue of PaXAT Trp 118 (Trp 139) has been proposed to lie near the active site of this enzyme on the basis of its site-directed replacement to alanine, which abolishes the intrinsic fluorescence quench observed on acetyl CoA binding (*40*). These results suggest that some nonxenobiotic hexapeptide acetyltransferases may share active-site structural features in common with PaXAT and other members of the hexapeptide xenobiotic acetyltransferase class of enzymes.

Although the precise role of PaXAT His 79B is unknown, its function may be revealed by a comparison of the active sites of PaXAT and CAT_{III} (*16*, *17*). The overall polypeptide chain folds of these enzymes are completely unrelated but both enzymes are trimers and bind similar substrates to opposites ends of an active-site tunnel located at the boundary of adjacent subunits. PaXAT His 79B projects into the active-site tunnel from an extended loop domain and appears to form interactions with its side chain that are identical to those made by CAT_{III} His 195, a residue that has been proposed to function as a general base catalyst (*15*). Both residues appear to donate a hydrogen bond from their imidazole ND1 to a main-chain oxygen atom and to receive a hydrogen bond from the C-3 hydroxyl of Cm with their imidazole NE2 atoms. Although attempts to superimpose additional active-site features do not reveal further structural

correspondence, the environmental similarities of these histidine residues could indicate that His 79B of PaXAT plays a similar role to that proposed for His 195 of CAT_{III}. In such a case, the hydrogen bond donated by the imidazole ND1 of PaXAT His 79B to the peptide oxygen of Thr 86B would stabilize the proper tautomeric form of the imidazole lacking a proton on NE2 and could promote the role of this residue as a general base catalyst capable of deprotonating the C-3 hydroxyl group of the Cm acceptor.

ACKNOWLEDGMENT

We gratefully acknowledge the technical assistance of Yu Tian and helpful discussions with Drs. John S. Blanchard, Laurence R. Olsen, and Giovanna Scapin and Mr. Yoram Puius and Renjian Zheng.

REFERENCES

1. Bairoch, A. (1993) *Nucleic Acids Res. 21*, 3097−3103.
2. Vaara, M. (1992) *FEMS Microbiol. Lett. 97*, 249−254.
3. Dicker, I. B., and Seetharam, S. (1992) *Mol. Microbiol. 6*, 817−823.
4. Anderson, M. S., and Raetz, C. R. H. (1987) *J. Biol. Chem. 262*, 5159−5169.
5. Kelly, T. M., Stachula, S. A., Raetz, C. R. H., and Anderson, M. S. (1993) *J. Biol. Chem. 268*, 19866−19874.
6. Downie, J. A. (1989) *Mol. Microbiol. 3*, 1649−1651.
7. Gilvarg, C., and Weinberger, S. (1970) *J. Bacteriol. 101*, 323−324.
8. Murray, I. A., and Shaw, W. V. (1997) *Antimicrob. Agents Chemother. 41*, 1−6.
9. Raetz, C. R. H., and Roderick, S. L. (1995) *Science 270*, 997−1000.
10. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. C. (1996) *EMBO J. 15*, 2323−2330.
11. Beaman, T. W., Binder, D. A., Blanchard, J. S., and Roderick, S. L. (1997) *Biochemistry 36*, 489−494.
12. Tennigkeit, J., and Matzura, H. (1996) *Gene 98*, 113−116.
13. Parent, R., and Roy, P. H. (1992) *J. Bacteriol. 174*, 2891−2897.
14. Tian, Y., Beaman, T. W., and Roderick, S. L. (1997) *Proteins: Struct., Funct., Genet. 28*, 298−300.
15. Kleanthous, C., Cullis, P. M., and Shaw, W. V. (1985) *Biochemistry 24*, 5307−5313.
16. Leslie, A. G. W., Moody, P. C. E., and Shaw, W. V. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4133−4137.
17. Leslie, A. G. W. (1990) *J. Mol. Biol. 213*, 167−186.
18. Monod, M., Mohan, S., and Dubnau, D. (1987) *J. Bacteriol. 169*, 340−350.
19. Alignet, J., Loncle, V., Mazodier, P., and El Solh, N. (1988) *Plasmid 20*, 271−275.
20. Toriya, M., Sakakibara, M., Matsushita, K., and Morohoshi, T. (1992) *Chem. Pharm. Bull. 40*, 2473−2477.
21. Alignet, J., Loncle, V., Simenel, C., Delepierre, M., and El Solh, N. (1993) *Gene 130*, 91−98.
22. Rende-Fouriner, R., Leclercq, R., Galimand, M., Duval, J., and Courvalin, P. (1993) *Antimicrob. Agents Chemother. 37*, 2119−2125.
23. Alignet, J., and El Solh, N. (1995) *Antimicrob. Agents Chemother. 39*, 2027−2036.
24. Bunny, K. L., Hall, R. M., and Stokes, H. W. (1995) *Antimicrob. Agents Chemother. 39*, 686−693.
25. Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.
26. Howard, A. J. (1986) *A guide to data reduction for the Nicolet Imaging Proportional Counter: the XENGEN system*, Protein Engineering Department, Genex Corp., 16020 Industrial Drive, Gaithersburg, MD 20877.
27. Furey, W., and Swaminathan, S. (1990) American Crystallographic Association Annual Meeting Program Abstract 18, 73 (abstr).
28. Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsl. Protein Crystallogr. 31*, 34−38.
29. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
30. Brunger, A. T. (1992) *Nature 355*, 472−475.
31. Kleywegt, G. J., and Brunger, A. T. (1996) *Structure 4*, 897−904.
32. Brunger, A. T., Krukowski, A., and Erickson, J. W. (1990) *Acta Crystallogr. A46*, 585−593.
33. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
34. Laskowski, R. A., MacArthur, M. W., Moss, S. D., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
35. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol. 7*, 95−99.
36. Kleywegt, G. J., and Jones, T. A. (1996) *Structure 4*, 1395−1400.
37. Richardson, J. S. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2619−2623.
38. Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167−339.
39. Engel, C., and Wierenga, R. (1996) *Curr. Opin. Struct. Biol. 6*, 790−797.
40. Lewendon, A., Ellis, J., and Shaw, W. V. (1995) *J. Biol. Chem. 270*, 26326−26331.
41. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
42. Bacon, D. J., and Anderson, W. F. (1996) *J. Mol. Graphics 6*, 219−220.
43. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.
44. Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.

BI980106V