Three-Dimensional Structure of *Escherichia coli* Dihydrodipicolinate Reductase[†,‡]

Giovanna Scapin, John S. Blanchard, and James C. Sacchettini*

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

Received August 24, 1994; Revised Manuscript Received December 15, 1994[⊗]

ABSTRACT: Dihydrodipicolinate reductase is an enzyme found in bacteria and higher plants involved in the biosynthesis of diaminopimelic acid and lysine. Because these pathways are unique to bacteria and plants, they may represent attractive targets for new antimicrobial or herbicidal compounds. The three-dimensional structure of *Escherichia coli* dihydrodipicolinate reductase, complexed with NADPH, has been determined and refined to a crystallographic *R*-factor of 18.6% with diffraction data to 2.2 Å resolution. The refined model contains the complete protein chain, the cofactor NADPH, and 55 water molecules. The enzyme is composed of two domains. The dinucleotide binding domain has a central seven-stranded parallel β-sheet surrounded by four α-helices, with the cofactor binding site located at the carboxy-terminal edge of the sheet. The second domain contains four β-strands and two α-helices that form an open mixed β-sandwich. A possible binding site for dihydrodipicolinate has been identified in this second domain, about 12 Å away from the dinucleotide binding site. This would imply that the protein must undergo some conformational change in order to perform catalysis. In the crystal, the native enzyme is a homotetramer generated by a 222 crystallographic axis. Implications of the tetrameric structure for the enzyme function are presented. Dihydrodipicolinate reductase uses both NADH and NADPH as cofactors, and analysis of its cofactor binding site allows for a molecular understanding of the enzyme's dual specificity.

Dihydrodipicolinate reductase (DHPR)[1] catalyzes the NAD(P)H-linked reduction of dihydrodipicolinic acid (DHP) to tetrahydrodipicolinic acid in the biosynthetic pathway leading to diaminopimelic acid and lysine in bacteria. Diaminopimelic acid (DAP) is an essential component of the peptidoglycan layer of bacterial cell walls, and disruption of its biosynthesis results in cell death, probably due to instability of the peptidoglycan (Cirillo et al., 1994). Lysine is produced by the enzymatic decarboxylation of DAP and is essential for protein synthesis. The DAP pathway for lysine synthesis is also present in higher plants and most algae while a different pathway, the *α-aminoadipic acid* pathway, is used by euglenoid algae and higher fungi.

In bacteria, three routes are known for biosynthesis of DAP and lysine (Figure 1): one in which succinylated intermediates are used (succinylase pathway; Kindler & Gilvarg, 1960; Berges et al., 1986); a second one in which acetylated intermediates are used (acetylase pathway; Sundharadas & Gilvarg, 1967); and the dehydrogenase pathway in which the intermediate tetrahydrodipicolinate, the precursor of all three pathways (see Figure 1), is converted in a single step to DAP (White, 1986). The presence of multiple biosynthetic pathways for the synthesis of DAP and lysine may indicate their importance for bacterial survival. In DAP biosynthesis, dihydrodipicolinate reductase, encoded by the *dap*B gene, represents the enzyme which synthesizes the precursor used in all three pathways (Figure 1). To date, *Escherichia coli* (Bouvier et al., 1984), *Bacillus lactofermentum* (Pisabarro et al., 1993), and *Mycobacterium bovis* BCG (Cirillo et al., 1994) *dap*B genes have been sequenced. Given the fact that the DAP/lysine biosynthetic pathway is unique to plants and bacteria, DHPR may represent an attractive target for new herbicidal or antimicrobial compounds. For this reason, a more complete understanding of the structural basis for its mechanism and kinetic properties is highly desirable. The *dap*B genes encode for enzymes of 248−273 residues, which show a reasonable degree of homology along their entire length (24−33% identity; Cirillo et al., 1994). Sequence comparisons have revealed that dihydrodipicolinate reductase contains an N-terminal region which is similar to the dinucleotide binding domain of several dehydrogenases (Pisabarro et al., 1993), including the aspartic semialdehyde dehydrogenases of *E. coli*, *Saccharomyces cerevisiae*, and *Corynebacterium glutamicum*, the glyceraldehyde-3-phosphate dehydrogenase of *C. glutamicum*, white mustard, and *Mesembryanthemum crystallinum*, and malate dehydrogenases of pig mitochondria and *E. coli*. By comparison of these domains a consensus sequence, (V/I)(A/G)(V/I)-XGXXGXXG, was observed. This sequence is highly conserved in different NAD(P)H-dependent dehydrogenases, including DHPR (Pisabarro et al., 1993), and is part of the "fingerprint" used to identify the fold responsible for dinuceotide binding in proteins (Wierenga, 1986).

The *E. coli dap*B gene encodes a 273 amino acid polypeptide (M_r = 28 758), which has been expressed and purified to homogeneity (Reddy et al., 1995). The enzyme

[†] This work was supported by NIH Grants AI33696 and GM45859.

[‡] Coordinates for the three-dimensional structure of *E. coli* dihydrodipicolinate reductase have been deposited with the Brookhaven Protein Data Bank under the accession code 1DIH.

* Author to whom correspondence should be addressed [telephone (718) 430-2741; Fax (718) 597-5692].

[⊗] Abstract published in *Advance ACS Abstracts,* March 1, 1995.

[1] Abbreviations: NADPH, β-nicotinamide adenine dinucleotide phosphate, reduced form; NADH, β-nicotinamide adenine dinucleotide, reduced form; DHPR, dihydrodipicolinate reductase; DHP, dihydrodipicolinic acid; DAP, diaminopimelic acid; HEPES, *N*-(2-hydroxyethyl)-piperazine-*N*′-2-ethanesulfonic acid; EMP, ethylmercury phosphate; MA, mercury acetate; MIR, multiple isomorphous replacement.

0006-2960/95/0434-3502$09.00/0

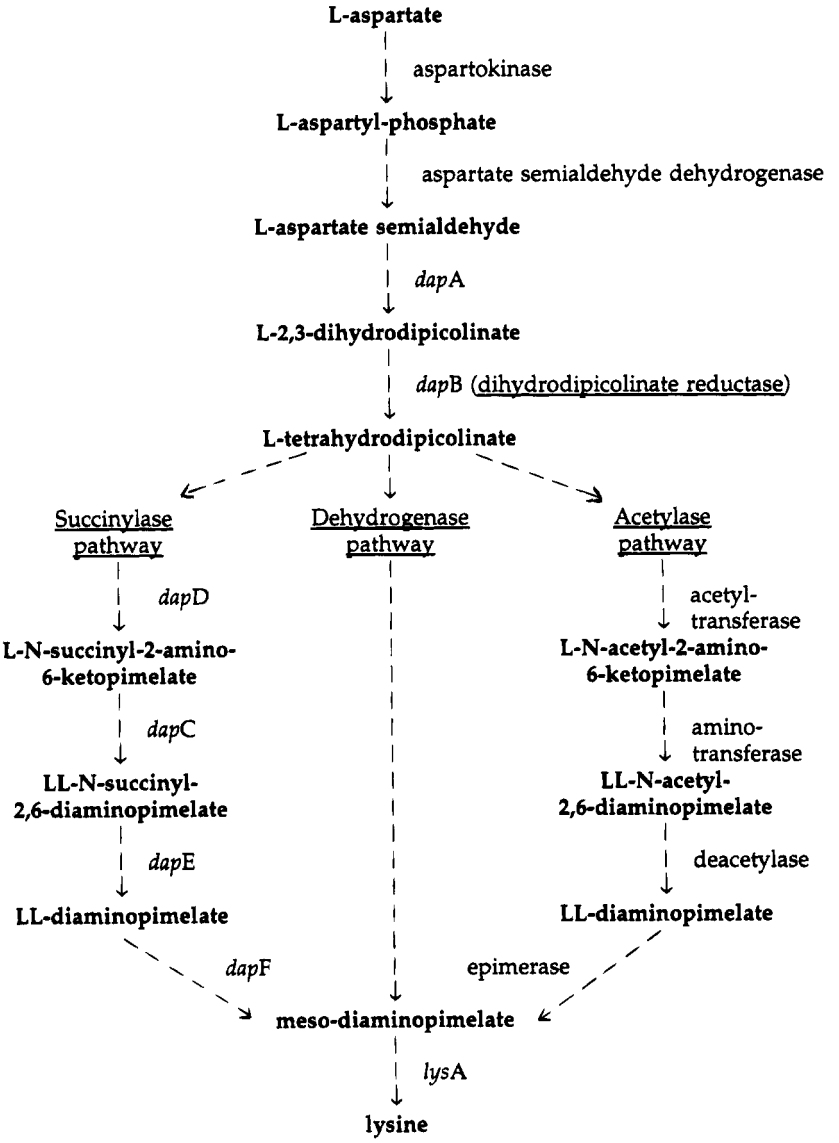


FIGURE 1: Pathways for biosynthesis of *meso*-diaminopimelic acid and lysine in bacteria.

shows unusual pyridine nucleotide substrate specificity, exhibiting only modest selectivity for its dinucleotide (Tamir & Gilvarg, 1974). In the preceding paper DHPR was shown to use NADH as a slightly better substrate than NADPH (the relative *V/K* value for NADH is approximaterly 2-fold higher). A 6-fold higher affinity of NADH over NADPH was also determined by titration calorimetry studies. This is quite unusual, since pyridine nucleotide dependent enzymes usually show a strong preference for one of the two cofactors and those exhibiting a dual-cofactor specificity preferentially utilize NADPH over NADH (Male & Storey, 1982; Lee & Levy, 1992; John et al., 1994). A number of studies have attempted to explain the molecular basis for this specificity, focusing on the fingerprint region of the nucleotide binding fold or on the presence of conserved basic residues at the C-terminal end of the second β-strand of the $\beta\alpha\beta$-fold (Scrutton et al., 1990; Bystroff et al., 1990; Baker et al., 1992b; Sem & Kasper, 1993). We report here the three-dimensional structure of *E. coli* DHPR complexed with NADPH, solved and refined to 2.2 Å resolution. *E. coli* DHPR represents the second case of a dual specificity enzyme (the other is glucose dehydrogenase; John et al., 1994) for which the three-dimensional structure is available. Analysis of the interactions between DHPR and NADPH may permit a molecular description of the enzyme's dual specificity.

EXPERIMENTAL PROCEDURES

Crystallization, Heavy Atom Preparation, and Data Collection. Recombinant *E. coli* DHPR was expressed and purified as described in the preceding paper (Reddy et al., 1995). Crystals of the apoprotein (i.e., without cofactor) and of the binary DHPR/NADPH complex were obtained using the hanging drop vapor diffusion method, with 2.2 M ammonium sulfate as precipitant, in 100 mM HEPES, pH = 7.5. Crystals were grown at room temperature, from a 6-μL droplet, which contained 3 μL of protein solution (18 mg/mL) and 3 μL of precipitant solution. NADPH (1.5–2.0 molar excess, when present) was incubated with the protein for about 30 min on ice before crystallization. Crystals appeared between 1 and 2 weeks and reached their maximum dimensions over a period of a month. Both apo- and DHPR/NADPH crystals typically measure 0.5 × 0.5 × 1.5 mm and diffract to better than 2.0 Å resolution. The crystals are orthorhombic, with unit cell parameters a = 75.7 Å, b = 81.2 Å, c = 94.2 Å, and $\alpha = \beta = \gamma = 90.0°$, and

Table 1: Data Collection and Statistics

data set	max resolution (Å)	no. of observations	no. of reflections	% completeness	$I/\sigma I$	R_{sym} (I)	R_{merge} (F)
native 1	2.5	32 982 (1877)[a]	9 569 (983)[a]	92.4 (61.4)[a]	14.6 (1.2)[a]	9.4	
native 2[b]	2.1	76 793 (5854)[a]	15 200 (2346)[a]	93.4 (74.5)[a]	16.0 (1.7)[a]	7.7	5.5
EMP	2.4	38 379	10 717	87.3	20.8	7.0	12.0
MA	2.4	52 097	9 026	76.0	15.1	5.4	9.5

[a] For the two native data sets, the number in parentheses represents measures in the last shell of resolution (2.6−2.5 Å for native 1 and 2.2−2.1 Å for native 2). [b] This data set was collected after the crystal was soaked in 2.5 mM NADPH, as described under Experimental Procedures, and using a San Diego MultiWire System (SDMS) area detector, equipped with a Rigaku RU 200 rotating anode X-ray source, operating at 50 kV and 150 mA.

Table 2: Statistics for Phase Calculation

derivative	x	y	z	OCC	B	no. of reflections	phasing power	R_{Cullis}	R_{Kraut}	resolution (Å)
EMP (iso)						6380	1.22	0.594	0.161	2.7
HG1	0.17	0.08	0.12	0.70	20.0					
HG2	0.22	0.28	0.06	0.88	20.0					
MA (iso)						6251	2.59	0.486	0.082	2.7
HG3	0.21	0.24	0.09	0.85	25.0					
MA (ano)						5438	3.17		0.211	2.7
HG4	0.21	0.24	0.09	1.0	25.0					
resolution (Å)	9.7	6.0	5.0	4.5	4.0	3.7 3.5	3.2 3.0	2.7	overall	
mean figure of merit	0.87	0.86	0.82	0.82	0.78	0.77 0.73	0.70 0.66	0.61	0.75	

contain one molecule per asymmetric unit. The space group was determined by pseudo-precession photograph (PHASES; Furey & Swaminathan, 1990) to be either $I222$ or $I2_12_12_1$. This ambiguity was easily resolved by preparing and collecting X-ray diffraction data on one heavy atom derivative (as described below) and analyzing the heavy atom Harker vectors in the correspondent difference (derivative minus native) Patterson map. This procedure defined the crystals as belonging to the space group $I222$. Only the NADPH-containing crystals were used for structure determination purposes.

X-ray diffraction data for native and heavy atom derivative crystals were collected on a Siemens multiwire area detector coupled to a Rigaku RU 200 rotating anode X-ray source operating at 55 kV and 80 mA. Data were processed and reduced using the XENGEN software (Howard, 1986). Table 1 reports statistics on the data collected.

Structure Determination. Two heavy atom derivatives, ethylmercury phosphate (EMP) and mercury acetate (MA), were used to calculate the initial phases for the *E. coli* DHPR/NADPH binary complex diffraction data. The EMP derivative was obtained by soaking a native protein crystal overnight in a solution containing 1 mM EMP in 2.2 M ammonium sulfate and 100 mM HEPES, pH = 7.5. For the MA derivative, a native crystal was soaked overnight in 0.25 mM MA in the same mother liquor. Analysis of difference Patterson maps for the EMP derivative showed the presence of two heavy atom sites. Interestingly, the MA derivative had one single site that was different from either of the two EMP binding sites. The heavy atom positions, as calculated from difference Patterson maps, were refined using the package PHASES (Furey & Swaminathan, 1990). Isomorphous data from both derivatives, together with anomalous data from the MA derivative, were used to phase the diffraction data. The resulting mean figure of merit for 6965 phased reflections to 2.7 Å, with $F > 2\sigma$, was 0.75. Table 2 summarizes the statistics for the MIR phase calculation. Solvent flattening (Wang, 1985) procedures, as implemented in PHASES, were used to further improve the MIR phases. The resulting electron density map permitted the tracing of nearly the entire polypeptide chain (excluding the first four residues, Met1 through Ala4, for which no electron density could be found) and the assignment of approximately 70% of the side chain. TOM (a derivative of FRODO; Jones, 1987), displayed on a Silicon Graphics Iris workstation, was used exclusively for model building. "Combined" electron density maps (i.e., maps calculated by combining MIR and model-based phases) were then used (PHASES; Furey & Swaminathan, 1990) to incorporate the complete amino acid sequence.

Table 3: Course of Refinement for the *E. coli* DHPR Binary Complex

	cycle			
	1	2	3	4
program	X-PLOR	TNT	TNT	TNT
resolution range (Å)	10.0−2.7	20.0−2.5	20.0−2.3	20.0−2.2
no. of reflections	7160	8151	12225	13459
protein atoms	1964	1964	1964	1979
substrate			NADPH	NADPH
solvent atoms		39	40	55
crystallographic R-factor (%)	25.2	19.0	20.4	18.6
rms deviation				
bond distance (Å)	0.025	0.025	0.022	0.020
bond angle (deg)	4.7	3.2	2.9	2.8

Structure Refinement. The protein coordinates determined from the MIR electron density were used as a starting point for refinement, as summarized in Table 3. In the first cycle the simulated annealing procedure (slow cool), as implemented in the X-PLOR manual (Brunger et al., 1987; Brunger, 1992), reduced the crystallographic R-factor from 43.2% to 25.2% for 7160 reflections between 10.0 and 2.7 Å. The resulting model had a reasonable geometry (rms bond lengths and bond angles were 0.025 Å and 4.7°, respectively). Difference Fourier ($(|F_o| - |F_c|)\phi_C$) and omit electron density maps (i.e., maps calculated omitting from the structure factor calculation regions of the protein) allowed for the repositioning of atoms in the loops Gly50−Val60, Asp71−Phe79, and Ala181−Gly204. Least squares positional and temperature factor refinement (TNT; Tronrud et

al., 1988) was subsequently included as part of the refinement strategy, initially (cycle 2 in Table 3) with data between 20.0 and 2.5 Å and removing all data with $I < 2\sigma I$ (8151 reflections, 79% of the possible). The weight for the temperature factor correlation (the restraint that causes the thermal factors for atoms bonded to each other to be similar) was set to 0.4 throughout the entire refinement. After 30 iterations, the crystallographic *R*-factor dropped to 19.0%, and the structure had rms deviations for bond lengths and bond angles of 0.025 Å and 3.2°, respectively. Inspection of a difference $((|F_o| - |F_c|)\phi_C)$ Fourier electron density map revealed the presence of electron density that, based on its location, size, and shape, appeared to be the bound NADPH. Because the electron density for the adenine ring was poor, a second diffraction data set was collected (to 2.1 Å resolution, see Table 1) after a binary complex crystal was soaked for 4 h in a freshly made 2.5 mM solution of NADPH in mother liquor. Difference Fourier electron density maps $((|F_o| - |F_c|)\phi_C)$ calculated with the new data set to 2.5 Å were of better quality than the previous, and the bound cofactor could be unambiguously fit to the electron density. Thirty-nine water molecules were also added to the structure at this stage, to account for positive peaks of density in the difference $((|F_o| - |F_c|)\phi_C)$ Fourier map that were greater or equal to three times the standard deviation of the map. Forty-five additional cycles of refinement using TNT were then run, incorporating the bound NADPH molecule into the structure and extending the resolution to 2.2 Å. The final model of *E. coli* DHPR has a crystallographic *R*-factor of 18.6% for 13459 reflections with $I > 2\sigma I$ (89% of the possible data) and rms deviations for bond lengths and bond angles of 0.02 Å and 2.8°, respectively.

The crystallographic *R*-factor for all 14 174 reflections (94% complete) is 21.2%.

RESULTS AND DISCUSSION

Figure 2a shows a ribbon diagram of the DHPR three-dimensional structure, and Figure 2b is a stereo diagram of the Cα trace. The structure was initially solved to 2.7 Å using MIR procedures with data from two heavy atom derivatives, as described under Experimental Procedures. Several cycles of refinement were then used to increase the resolution and to improve both the crystallographic *R*-factor and the overall geometry of the model. Omit maps (i.e., maps calculated without including a portion of the molecule in the structure factor calculation) were extensively used throughout the course of refinement, in order to rebuild loop regions (residues Glu40−Ser65, His180−Ala190, Ser255−Glu258, and the last 9 residues, Arg265−Leu273) or to properly assign side-chain positions (i.e., Arg16, Arg39, Glu40, Lys187).

Quality of the Structure. The current model contains 1979 protein atoms, 55 solvent molecules, and the 48 non-hydrogen atoms of bound NADPH. Electron density for the backbone atoms of residues 1 through 4 was visible only at the latter stages of refinement, after inclusion of 2.2 Å data, and these four residues were inserted as alanines (i.e., without side chains, for which no density was visible). The final model has an *R*-factor of 18.6% for 13459 reflections with $F > 2\sigma$ to 2.2 Å (89% of possible data) and rms deviations for bond lengths and bond angles of 0.020 Å and 2.8°, respectively. The crystallographic *R*-factor for all 14 174 reflections (94% complete) is 21.2%. Figure 3 is a Ramachandran plot (as calculated with PROCHECK; Laskowski et al., 1993) of the ϕ, ψ angles for the refined protein backbone. All residues fall within the allowed region of low conformational energy. Of the five residues located in generously allowed regions of low energy, two (Asp48 and Ala55) are located in a long surface loop (Glu40 through Gly59) that connects β-strands B2 and B3 in the N-terminal region of the protein. The electron density for this loop was weak, and the loop is probably flexible in the crystal. Of the remaining three residues, residues Ala3 and Asn5 appear to be relatively disordered in the N-terminal portion of the refined structure, and Thr80 is located in a loop that, as will be discussed later, is involved in NADPH binding. The electron density for this side chain is well-defined, and the unusual geometry may be related to the binding of the cofactor.

Description of the Structure. E. coli DHPR is a two-domain protein, with an elongated shape, measuring approximately 70 Å × 35 Å × 40 Å (Figure 2). Scheme 1 is a cartoon of the secondary structure elements in the *E. coli* DHPR three-dimensional structure. The first domain, which we refer to as the nucleotide binding domain, consists of four α-helices and seven β-strands. These secondary structure elements are arranged to form an α/β structure (typically referred to as a dinucleotide binding or Rossmann fold) similar to that of the dinucleotide binding domains of many dehydrogenases [Rossmann et al., 1975; for a review, see Branden and Tooze (1992)]. This structure can be divided into two similar halves: the first half contains β-strands B1, B2, and B3 and α-helix A1. β-strand B2 is connected to β-strand B3 by a long loop (Glu40 through Gly59) rather than by the helix usually found in the dinucleotide binding fold. Residues Ser65−Asp74 connect the first half of the Rossmann fold to the second, which contains β-strands B4, B5, and B6 and α-helices A2 and A3. The nucleotide binding domain contains two more secondary structural elements, α-helix A6 and β-strand B11, that are actually from the C-terminal sequence. Helix A6 is parallel to A3 and β-strand B11 is parallel to β-strands B1−B6 of the central β-sheet, respectively. The consensus sequence conserved among NAD(P)H binding proteins (Pisabarro et al., 1993), (V/I)(A/G)(V/I)XGXXGXXG, is found at the C-terminal end of β-strand B1, into the loop that connects it to α-helix A1 (^{8}VAIAGAGGRMG^{18}). This was expected from comparisons with other NAD(P)H-dependent proteins for which the three-dimensional structures have been reported (i.e., horse liver alcohol dehydrogenase, Eklund et al., 1976; glutathione reductase, Karplus & Schulz, 1987; malate dehydrogenase, Birktoft et al., 1989). The identification of this conserved sequence and the presence of an acidic residue (Glu38) 20 residues downstream from this Gly-rich region allow us to define the B1−A1−B2 region in DHPR as the conserved ADP-binding βαβ-fold (Wierenga et al., 1986) and the whole N-terminal domain as the dinucleotide binding domain.

A sharp, four-residue loop (Ala127−Ser130) connects B6 to the first helix (A4) of the second domain of *E. coli* DHPR, which we believe is the substrate binding domain (Figure 2). This domain contains two α-helices (A4 and A5) and four long β-strands (B7 through B10) that form an open mixed β-sandwich (Richardson, 1981), since the β-sheet contains both parallel (B7 and B8) and antiparallel (B7, B9,

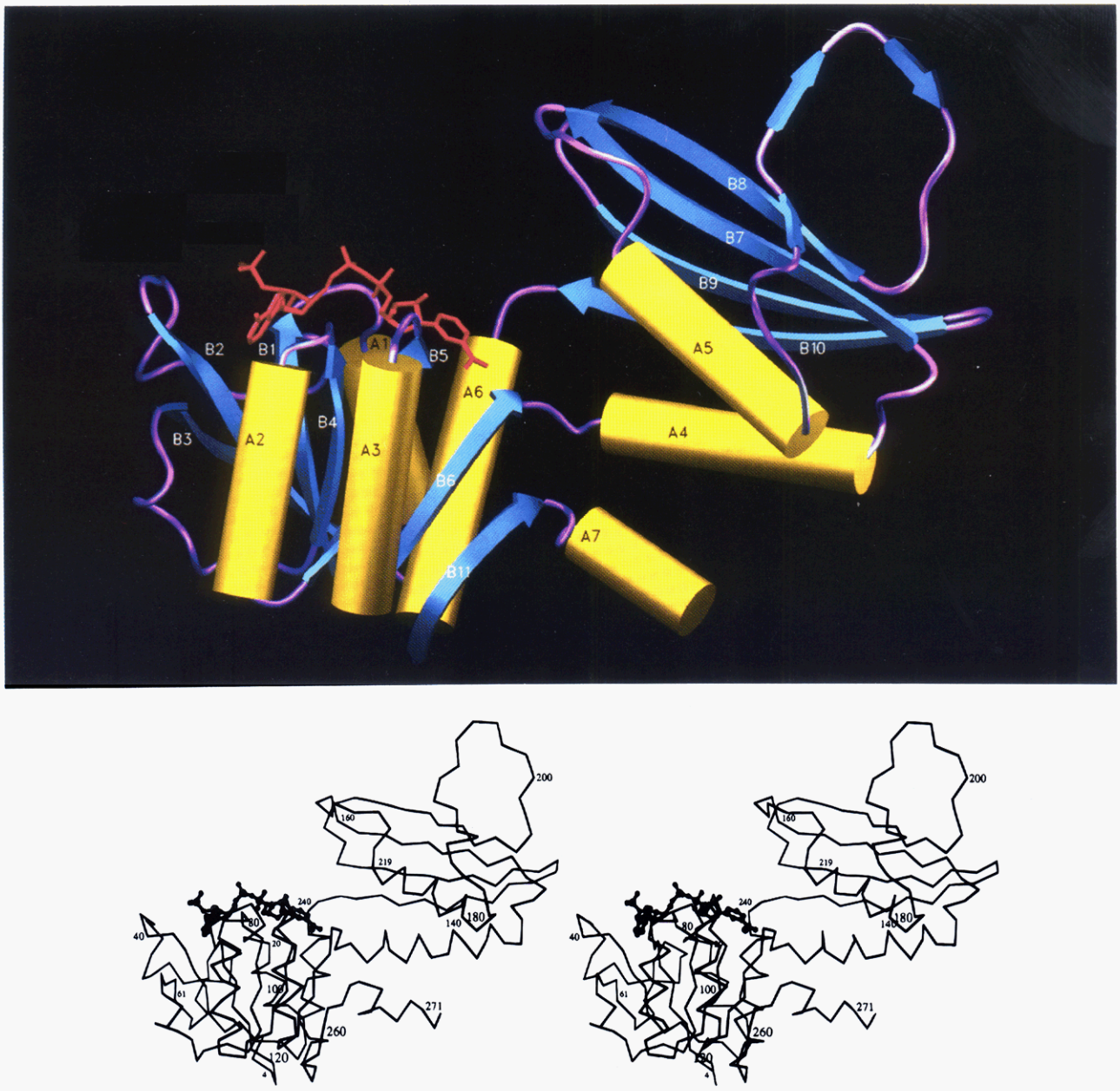


FIGURE 2: (a, top) Ribbon diagram of the three-dimensional structure of *E. coli* DHPR. Helices A1 through A3 and A6 and β-strands B1 through B6 and B11 define the nucleotide binding domain while helices A4 and A5 and β-strands B7 through B10 make up the substrate binding domain. The bound NADPH (red) is also shown. This figure (and Figure 4a) was prepared with the program SETOR (Evans, 1993). (b, bottom) Stereo diagram of the Cα trace of the DHPR monomer. The bound NADPH is shown as a ball-and-stick model. This figure (and Figure 5b) was generated using MOLSCRIPT (Kraulis, 1991).

and B10) β-strands. The C-terminal domain contains also a long loop (Leu182 through Gly204) that has three short β-strands, Cys189−Val191, Ser193−Glu195, and His197−Gly199, and extends from the body of the C-terminal domain forming an angle of about 60° with the β-sheet. This arrangement provides for two short connections (Ala127−Ser130 and Ser239−Arg240) between the two domains. These two loops are relatively unconstrained by packing, a common configuration for proteins in which a large domain motion is a consequence of substrate binding and is required for catalysis (Gerstein et al., 1994). In such linking hinge regions, large torsion angle changes involving a few residues are sufficient to produce the domain motion. The rest of the protein is envisioned to rotate as a rigid body with the rotation axis passing through the hinge region.

In solution, *E. coli* DHPR has been reported to be a tetramer of identical subunits (Tamir & Gilvarg, 1974; Reddy et al., 1995). Although DHPR crystals have one molecule per asymmetric unit, analysis of the crystal packing in the *I*222 space group shows the presence of a tetramer which is generated by 222 crystallographic symmetry. Figure 4 shows two ribbon diagrams of the DHPR tetramer. The interactions between the four monomers (yellow, orange, green, and blue) occur exclusively between residues of the substrate binding, carboxy-terminal domains.

The yellow and orange subunits, as well as the green and the blue, interact by pairing the four β-strands of their substrate binding domains to form an eight-stranded mixed β-sheet (B8−B7−B9−B10−B10′−B9′−B7′−B8′ in Figure 4a). Four helices, two from each one of the C-terminal domains (A4, A5, A4′, and A5′), are located on the exterior side of the β-sheet. β-strands B10 (residues 229−238) and its mate B10′ form the interface between the two monomers. These two β-strands mantain a typical hydrogen bond pattern

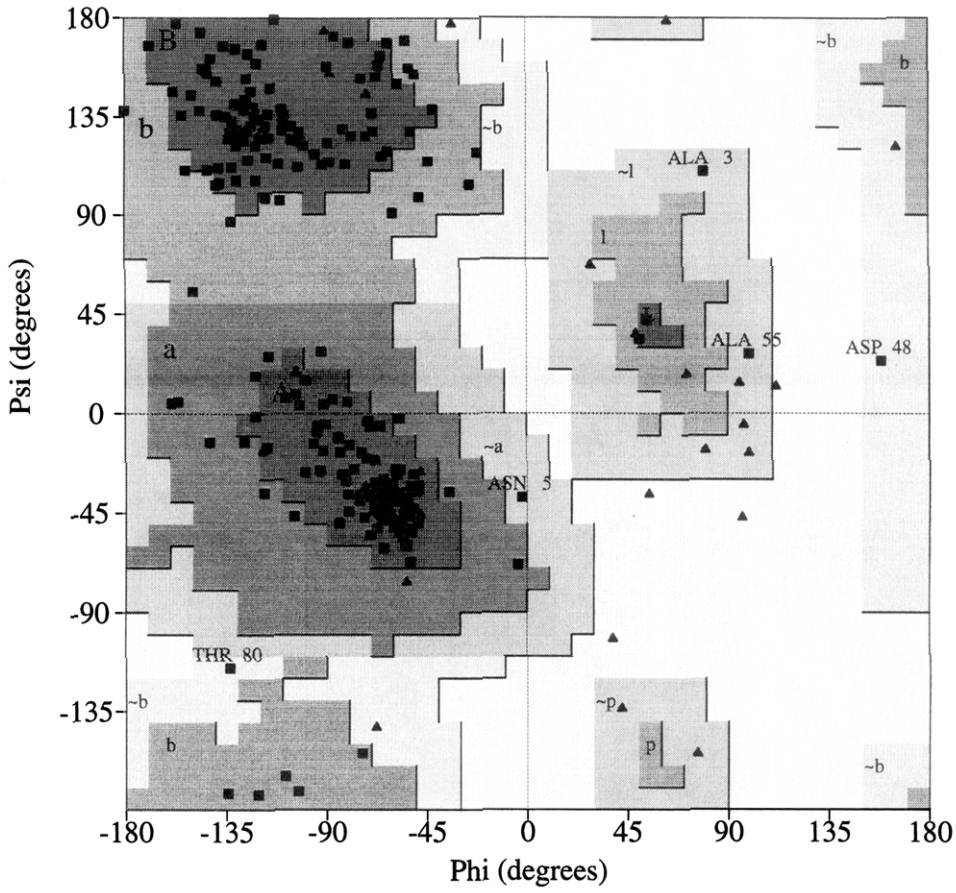


FIGURE 3: Ramachandran plot of ϕ, ψ angles for the refined structure of *E. coli* DHPR produced by the PROCHEK software (Laskowski et al., 1993). The "most favored" regions are shaded darkest (▲, glycines; ■, all other residues).

of two antiparallel β-strands. Helices A4 and A4′ are also arranged in an antiparallel fashion with an angle of about 60° between their axes, which is typical of helix−helix interactions (Richardson, 1981). Van der Waals interactions between three clusters of residues (Val135, Val146′, and Met147′ at one end of the helices, their symmetry-related mates Val146, Met147, and Val135′ at the other end, and Leu139/Leu139′ in the middle of the helices) may help stabilize the subunit−subunit association. Additional interactions are contributed by hydrogen bonds between the side-chain oxygen OE1 of Glu229 (of β-strand B10) and the main-chain nitrogen of Gly132 that belongs to α-helix A4.

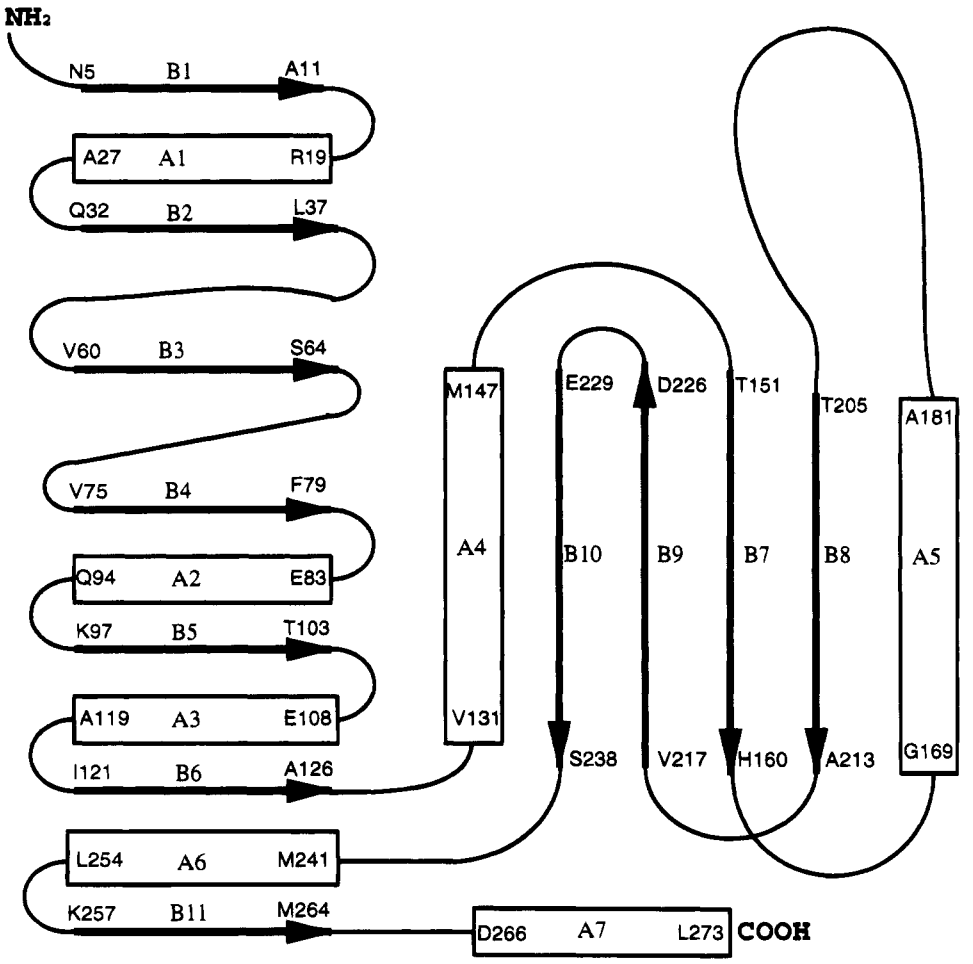
The yellow and the green subunits in Figure 4 and the orange and the blue subunits interact through β-strand B8 and its antiparallel symmetry-related strand B8′ (Figure 4a) and by residues Glu195−Pro203 (L1 in Figure 4a) and residues Val164−Pro167 (L2) and their symmetry-related loops L1′ and L2′. Residues Glu195−Pro203 are part of the long loop (Leu182−Ser204) that extends, as seen in Figure 2, from the core of the protein. This loop appears to play a role in the tetramer formation, wrapping around the mixed β-sheet of the neighboring monomer (Figure 4b) and providing favorable interactions.

The final result is that the two eight-stranded β-sheets are paired face to face forming a flattened 16-stranded mixed β-barrel anchored by the outermost β-strands and the four loops. A similar arrangement has been described for proteins of the lectin family (Hardman & Ainsworth, 1972; Dessen et al., 1994). The interior portion of the barrel may be divided in three parts: the upper and lower regions (as seen in Figure 4) which are made of mostly polar and charged residues. Residues of these two regions form an extensive network of hydrogen bonds that involves both main-chain and side-chain atoms. In the core of the barrel van der Waals interactions appear to contribute to the association of the tetramer.

The four dinucleotide binding domains of the tetramer extend from the core region as seen in Figure 4b. The mean temperature factors for the main-chain atoms of the N-terminal and C-terminal domains are quite different: for the nucleotide binding domain (including A7 and B11), it is 39.5 Å^2, while for the substrate binding domain, it is 24.5 Å^2. This rather large difference may be due to lattice contacts between the substrate binding domains which generate the tetramer.

Dinucleotide Binding Site. Crystals of *E. coli* DHPR used in the three-dimensional structure determination were obtained in the presence of NADPH. Analysis of $((|F_o| - |F_c|)\phi_C)$ difference Fourier electron density maps showed the presence of positive density across the C-terminal portion of the β-sheet of the first domain (Figure 5a). On the basis of its location and its shape this electron density was interpreted as being bound NADPH. Figure 5b is a stereo-view of the NADPH binding site. The dinucleotide is bound in an extended conformation, with the pyrophosphate moiety located in the crevice formed at the "switch point" of the β-sheet (Branden, 1981). Both ribose rings are 2′ *endo* and the adenine base is in an *anti* conformation. Atoms of the nicotinamide mononucleotide, especially the exocyclic amido group, are part of an extensive network of hydrogen bonds with protein atoms, while the adenine portion (AMP) is rather loosely associated as evidenced by the weaker density, fewer

Scheme 1: Schematic Representation of the Secondary Structure Element in the Three-Dimensional Structure of *E. coli* DHPR



hydrogen bonds, and higher temperature factors found for this portion of NADPH (the mean temperature factor for the AMP atoms is 49.4 Å^2, versus the 31.5 Å^2 of atoms of the nicotinamide ribose portion).

Four hydrogen bonds are present between the amido group of nicotinamide and the protein. The oxygen of the amido group hydrogen bonds to the main-chain nitrogen of Phe129 (2.9 Å), and the nitrogen of the amido group hydrogen bonds to the main-chain oxygens of Gly102, Ala126, and Ala127 (2.9, 3.2, and 3.1 Å, respectively). As a result, the nicotinamide ring is rigidly bound within a pocket (Branden & Tooze, 1992) formed by protein atoms (Figure 6). In this conformation, the nicotinamide C4S hydrogen points toward the protein, while the C4R hydrogen faces the carboxy-terminal domain. DHPR can be classified as a type A reductase (transfer of the *pro-R* hydride), in accord with the determined stereochemistry of hydride transfer catalyzed by DHPR (Reddy et al., 1995).

The nicotinamide ribose also forms an extensive set of hydrogen-bonding interactions with the protein through its O2′ and O3′ hydroxyls. O2′ is within hydrogen bond distance of the side-chain oxygen of Thr104 (2.5 Å), and O3′ forms hydrogen bonds with the main-chain carbonyl oxygen of Thr80 (2.8 Å), the main-chain nitrogen of Thr104 (2.7 Å), and the side-chain oxygen of Thr104 (3.1 Å).

The pyrophosphate moiety is near the Gly-rich loop (Gly12−Gly18) in the conserved dinucleotide binding motif and interacts with the main-chain nitrogens of Arg16 (two hydrogen bonds, at 2.9 and 3.4 Å, respectively) and Met17 (2.7 Å), as well as atoms of the guanidinium group of Arg16 (NE, 3.2 Å, and NH2, 3.3 Å), Arg81 (NH2, 3.1 Å), and Arg240 (NH2, 3.2 Å). Arg240 is located at the beginning of α-helix A6, and, together with β-strand B11, represents an extension of the typical Rossmann fold. The pyrophosphate is also stabilized by favorable interactions with the dipole of the "dinucleotide binding helix" (Hol et al., 1978); that in DHPR is A1.

The adenosyl ribose ring forms hydrogen bonds with the main-chain nitrogen of Gly15 (3.4 Å), the main-chain oxygen of Gly12 (2.85 Å), and the side-chain oxygen of Glu38 (3.6 Å). The adenine ring is within van der Waals contact distance of the aliphatic atoms (Cβ, Cγ, and Cδ) of the side chain of Arg39 and atoms of residues Thr80−Gly84 (Thr80C, Thr80Cγ, Arg81Cβ, Gly84Cα: these distances range from 3.3 to 4.6 Å). The N3 and N6 of the adenine ring also hydrogen bond to the main-chain nitrogen of Gly84 (3.3 Å), the side-chain oxygen of Thr80 (3.2 Å), and the side-chain nitrogen of Arg39 (NH2, 3.2 Å). The 2′-phosphate is located on a wide surface cavity bordered by Ala13−Gly15 and Asp38−Leu44. It appears to be rather loosely bound to the protein as there is a single hydrogen bond between one of its oxygens and the NE of Arg39 (2.5 Å). This is different from the extensive 2′-phosphate/protein interactions reported for other strictly NADPH-dependent enzymes (Baker et al., 1992a), in which as many as seven hydrogen bonds occur between the phosphate monoester and the protein. For example, in the structure of glutathione reductase (Karplus & Schulz, 1989) there are seven hydrogen

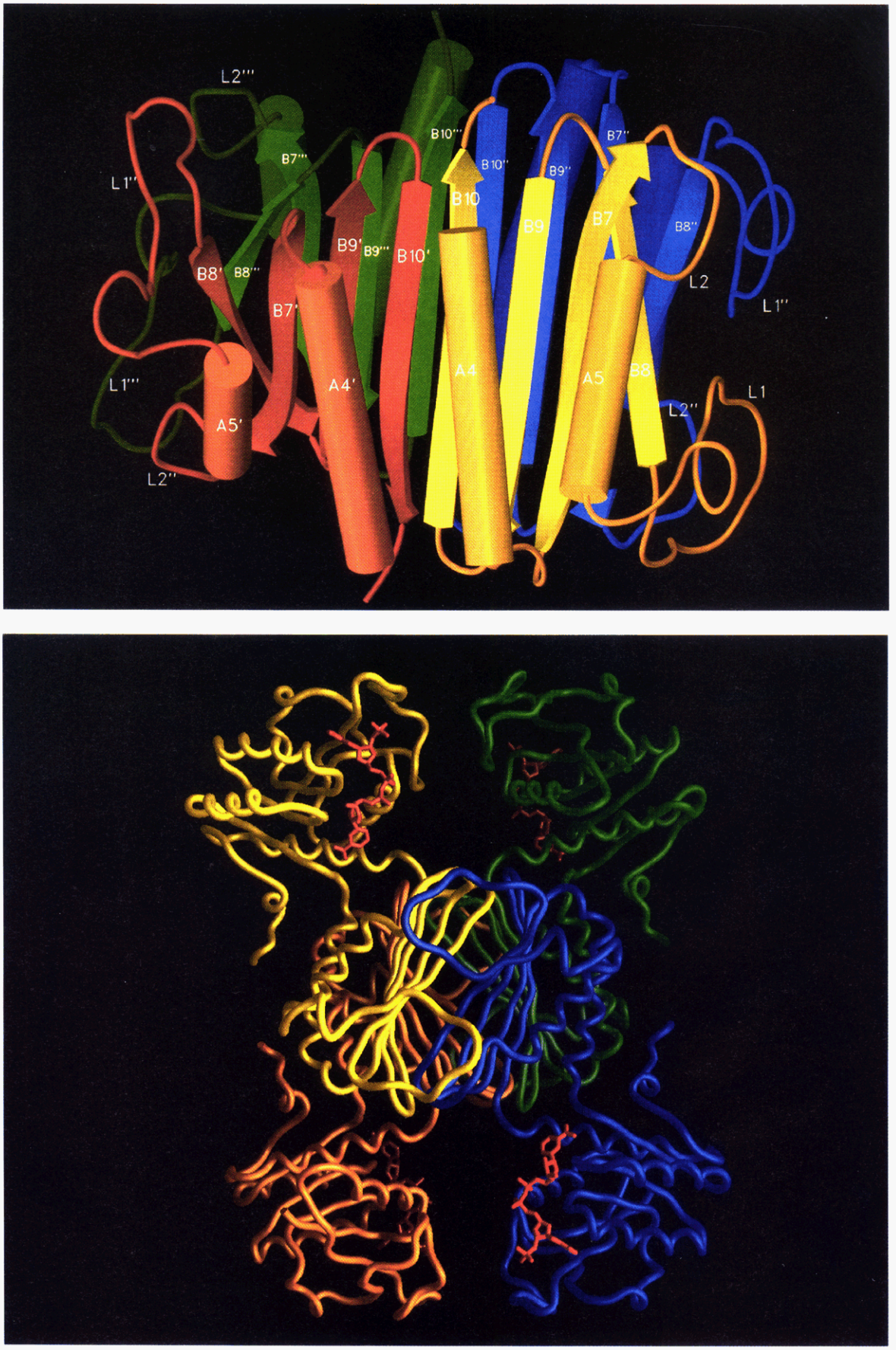


FIGURE 4: Ribbon diagrams of the DHPR tetramer. The four monomers are represented with different colors: yellow, orange, green, and blue. (a, top) Close-up view of the central core of the tetramer. See text for a detailed description. (b, bottom) Representation of the entire DHPR tetramer. The nucleotide binding domains and the bound NADPH (one molecule per monomer, red) are also displayed. Panel b and Figure 6 were made with the program GRASP (Nicholls et al., 1991).

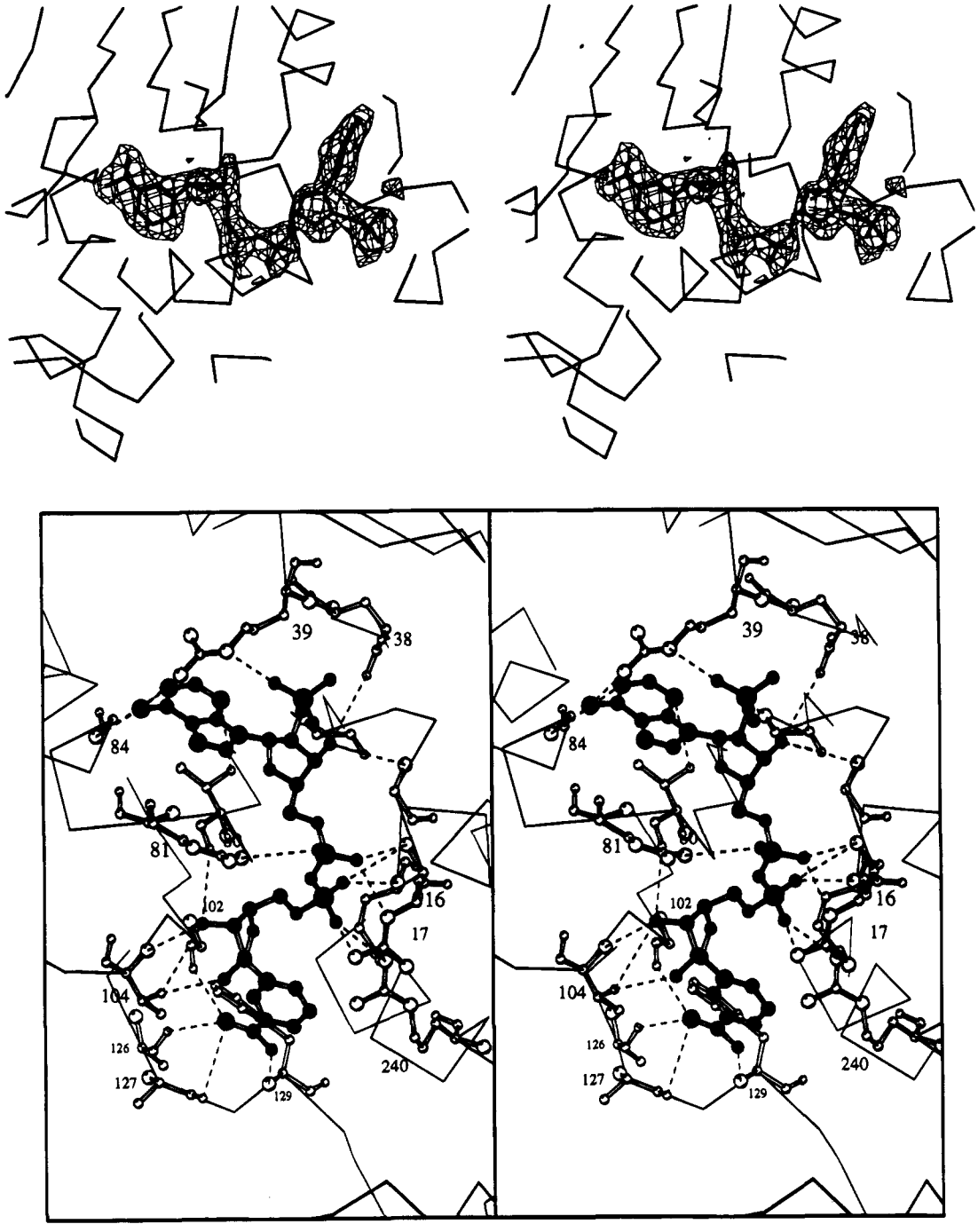


FIGURE 5: (a, top) Stereo diagram of a difference Fourier electron density map for the DHPR/NADPH complex showing the position of the bound cofactor. The map was calculated without including NADPH and solvent molecules in the model. (b, bottom) Stereo diagram of the NADPH binding site in DHPR. Residues involved in the nucleotide binding are displayed. Dotted lines represent hydrogen bond interactions.

bonds between the three terminal oxygens of the 2′-phosphate and protein atoms. In both dihydrofolate reductase (Bystroff et al., 1990) and ferrodoxin−$NADP^+$ reductase (Karplus et al., 1991), four hydrogen bonds between the 2′-phosphate and protein side- or main-chain atoms are reported.

E. coli DHPR does not significantly discriminate between NADPH (K_d = 1.8 μM) and NADH (K_d = 0.26 μM; Reddy et al., 1995). This is quite unusual, since nucleotide-dependent enzymes typically show a strong preference for one cofactor. Moreover, in other dual specificity enzymes studied, the steady-state K_m values of NADH are at least severalfold higher than the equivalent values for NADPH (Male & Storey, 1982; Lee & Levy, 1992; John et al., 1994). Additionally, the cofactor specificity may vary within related enzymes, from a strict preference for NADH to a strict preference for NADPH. For example, in the case of glutamate dehydrogenases (GDH; Baker et al., 1992b), the *Clostridium symbiosum* GDH only uses NADH, while the *E. coli* enzyme is strictly NADPH dependent. Enzymes that to date have been characterized as dual specificity include bovine liver GDH (Male & Storey, 1982), glyceraldehyde-3-phosphate dehydrogenase from maize chloroplast (Quigley et al., 1989), and glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* (DeMoss et al., 1953; Lee & Levy, 1992). A large number of comparative studies, based either on primary structures or, when available, on tertiary

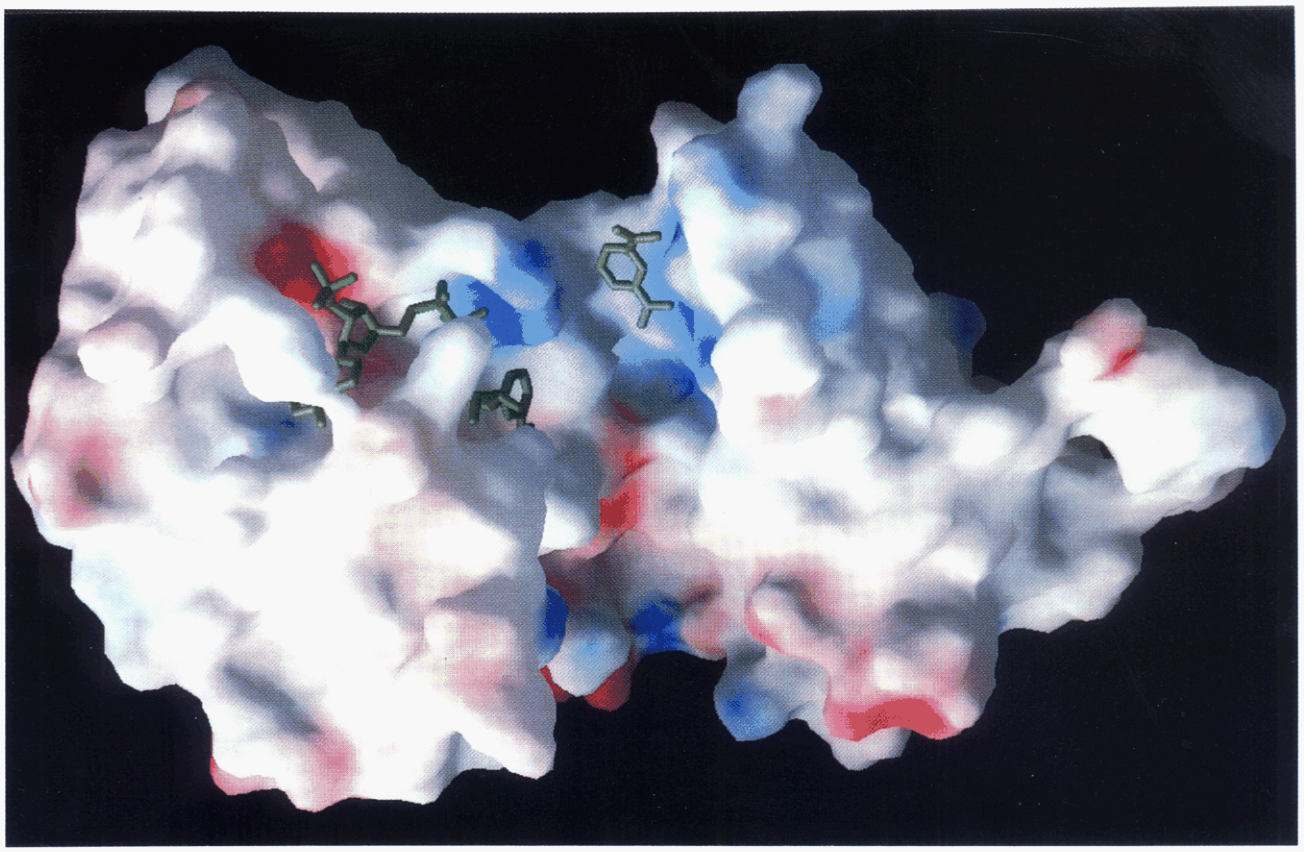


FIGURE 6: Surface of DHPR, colored according to potential. Blue and red colorations represent regions of positive and negative potential, respectively. NADPH is bound on the surface of the N-terminal domain, with its C4 facing the substrate binding domain. The blue coloration in this side of the substrate binding domain represents the stretch His159−Lys163 thought to be important in the substrate binding. A molecule of DHP is included in the model as described in the text.

structures, have attempted to explain the molecular basis for the specificity for one cofactor over the other. Several studies have focused on the fingerprint region of the nucleotide binding fold (the $\beta\alpha\beta$-fold, containing the conserved GXGXXG or GXXGXXG sequence and an acidic residue 20−30 residues downstream of the Gly-rich region) in order to identify a pattern that predicts the selectivity for either NADH or NADPH. These studies have failed to produce a general rule (Scrutton et al., 1990; Baker et al., 1992b). Baker et al. have suggested that the coenzyme specificity in a number of NAD-linked dehydrogenases might be based on hydrogen bond interactions between the main-chain nitrogen of the second residue of the consensus sequence and the "conserved acidic residue" (see above), which in turn makes hydrogen bonds to the O3′ and O2′ hydroxyls of the adenosyl ribose ring. However, the conserved acidic residue is often present in dual specifity and NADPH-dependent enzymes. In a second paper Baker et al. (1992a) propose that a discrimination in favor of NADH may depend on the size of the O2′ binding cavity. A number of studies have correlated the NADPH specificity to the presence of conserved basic residues, particularly arginine residues, near or in the place of the acidic residue at the C-terminal end of the second β-strand of the $\beta\alpha\beta$-fold. These residues are capable of providing stabilizing interactions for the negative charge(s) carried by the 2′-phosphate (Bystroff et al., 1990; Scrutton et al., 1990; Sem & Kasper, 1993). Karplus and Schulz (1989), in their structure of glutathione reductase, suggest that the total number and the quality of hydrogen bonds made in a NADH/inorganic phosphate complex or in the NADPH complex play a role in defining the overall energy of the binding and the specificity for NADPH.

Apart from the *Thermoplasma acidophilum* glucose dehydrogenase (John et al., 1994), which shows a 300-fold higher affinity for NADPH than for NADH and whose three-dimensional structure does not contain the cofactor, *E. coli* DHPR represents the only other case of a dual specificity enzyme for which the three-dimensional structure has been solved. Analysis of the interactions between the enzyme and NADPH allows for a molecular understanding of its dual specificity. DHPR has in the O2′ binding pocket characteristics that have been postulated to be responsible for specificity in both NADH- and NADPH-dependent enzymes. For example, the hydrogen bond between the side chain of Glu38 and the ribose O3′ is part of a network of hydrogen bonds similar to that of the "indirect recognition of adenine ribose" (Baker et al., 1992b) found in the NADH-dependent enzymes GAPDH, LDH, and ADH. On the other hand, a basic residue (Arg39) is also present in the binding pocket as found in many other NADPH-dependent enzymes (Bystroff et al., 1990; Scrutton et al., 1990; Sem & Kasper, 1993) and is thought to be important in interacting with the negatively charged 2′-phosphate moiety. Interactions between the 2′-phosphate and the protein's side chains, which are assumed to be the most important contribution to the cofactor specificity, are few in DHPR. Only one hydrogen bond is formed between a terminal oxygen of the 2′-phosphate and NE of Arg39; NH1 and NH2 of Arg39 are involved in loose hydrogen bond interactions with the adenine ring. The lack of extensive interactions with the 2′-phosphate may provide one basis for the dual specificity. In our current model, the

distance between NE of Arg39 and the ribose O2′ is about 4.6 Å. It is reasonable to imagine that, in the absence of a 2′-phosphate, a relatively small rearrangement of the Arg39 side chain could make its NE available to hydrogen bond the ribose O2′ while preserving the NH1 and NH2 interactions with the adenine. In both cases the number of hydogen bonds formed would be the same, and no distinction between the two cofactors could be made on this basis.

Dihydrodipicolinic Acid (DHP) Binding Site. Although the three-dimensional structure of *E. coli* DHPR we report here is that of the enzyme/NADPH complex, the structure of the binary complex provides clues which allow us to speculate on the substrate binding site on the enzyme. The NADPH is located on the surface of the N-terminal domain with the nicotinamide C4R hydrogen pointed toward the carboxy-terminal domain. It is reasonable to propose that the substrate binding site is located in this region. Since DHP at physiological pH is negatively charged, one can predict that the DHP binding site must contain positively charged residues. A cluster of five basic residues (His159, His160, Arg161, His162, and Lys163 identified by the blue coloration in Figure 6) is located in the loop connecting β-strand B7 to α-helix A5 (Figure 2) of the putative substrate binding domain. A DHP molecule was placed in this area (Figure 6) using Insight (Biosym Technologies, San Diego CA), such that the interactions between the negative charges carried by the DHP and the basic residues were optimized. The distance between the C4 position of DHP in this final position (C4 is the acceptor of the hydride transferred from NADPH; Reddy et al., 1995) and the C4 of the nicotinamide ring of NADPH is about 8 Å, too long for hydride transfer. A hinge movement (Gerstein et al., 1994) around the two short loops connecting the NADPH binding domain and the substrate binding domain (Ala127–Ser130 and Ser239–Arg240) could accomplish a reorientation to allow for catalysis. Because the substrate binding domain appears to be relatively rigid due to the extensive interactions in the tetramer, the hinge movement would require a rotation of the nucleotide binding domain of approximately 30°. On the basis of our model building studies, this motion would in fact position the cofactor with its nicotinamide C4 within a distance to the C4 of DHP that would allow hydride transfer. Crystallographic studies on a ternary complex [DHPR with bound NADPH and 2,6-pyridinedicarboxylate (a DHPR inhibitor); Reddy et al., 1995] are underway and will be necessary in order to confirm this hypothesis.

REFERENCES

Baker, P. J., Britton, K. L., Engel, P. C., Farrants, G. W., Lilley, K. S., Rice, D. W., & Stillman, T. J. (1992a) *Proteins 12*, 75–86.

Baker, P. J., Britton, K. L., Rice, D. W., Rob, A., & Stillman, T. J. (1992b) *J. Mol. Biol. 228*, 662–671.

Bartlett, A. T. M., & White, P. J. (1985) *J. Gen. Microbiol. 131*, 2145–2152.

Berges, D. A., DeWolf, W. E., Dunn, G. L., Newman, S. J., Schmidt, J., Taggart, J., & Gilvarg, G. (1986) *J. Biol. Chem. 261*, 6160–6167.

Birktoff, J. J., Rhodes, G., & Banaszak, L. J. (1989) *Biochemistry 28*, 6065–6081.

Bouvier, J., Richaud, C., Richaud, F., Patte, J.-C., & Stragier, P. (1984) *J. Biol. Chem. 259*, 14829–14834.

Branden, C.-I. (1981) *Q. Rev. Biophys. 13*, 317–338.

Brunger, A. T. (1992) *X-PLOR version 3 manual: a system for crystallography & NMR*, Yale University, New Haven, CT.

Brunger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Bukhari, A. I., & Taylor, A. L. (1971) *J. Bacteriol. 173*, 7772–7780.

Bystroff, C., Oatley, S. J., & Kraut, J. (1990) *Biochemistry 29*, 3263–3277.

DeMoss, R. D., Gunsalus, I. C., & Bard, R. C. (1953) *J. Bacteriol. 66*, 10–16.

Dessen, A., Gupta, D., Sabesan, S., Brewer, C. F., & Sacchettini, J. C. (1994) *Biochemistry* (submitted for publication).

Cirillo, J. D., Weisbrod, T. R., Banerjee, A., Bloom, B. R., & Jacobs, W. R., Jr. (1994) *J. Bacteriol. 176*, 4424–4429.

Eklund, H., Nordstrom, B., Zeppezauner, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B.-O., Tapia, O., & Branden, C.-I. (1976) *J. Mol. Biol. 102*, 27–59.

Evans, S. V. (1993) *J. Mol. Graphics 11*, 134–138.

Furey, W., & Swaminathan, S. (1990) *Am. Crystallogr. Assoc. Am. Meet. Program Abstr. 18*, 73.

Gerstein, M., Lesk, A. M., & Chothia, C. (1994) *Biochemistry 33*, 6739–6749.

Hardman, K. D., & Ainsworth, C. F. (1972) *Biochemistry 11*, 4910–4919.

Hol, W. G. J., Van Duijen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443–446.

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.

John, J., Crennel, S. J., Hough, D. W., Danson, M. J., & Taylor, G. L. (1994) *Structure 2*, 385–393.

Jones, T. A. (1985) *Methods Enzymol. 115*, 157–171.

Karplus, P. A., & Schulz, G. E. (1989) *J. Mol. Biol. 210*, 163–180.

Karplus, P. A., Daniels, M. J., & Herriot, J. R. (1991) *Science 251*, 60–66.

Kindler, S. H., & Gilvarg, C. (1960) *J. Biol. Chem. 235*, 3532–3535.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Laskowski, R. A., MacArthur, M. W., Moss, S. D., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283–291.

Lee, W. T., & Lewy, H. R. (1992) *Protein Sci. 1*, 329–334.

Male, K. B., & Storey, K. B. (1982) *Int. J. Biochem. 14*, 1083–1089.

Nicholls, A., Sharp, K., & Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281.

Pisabarro, A., Malumbres, M., Mateos, L. M., Oguiza, J. A., & Martin, J. F. (1993) *J. Bacteriol. 175*, 2743–2749.

Quigley, F., Brinkmann, H., Martin, W. F., & Cerff, R. (1989) *J. Mol. Evol. 29*, 412–421.

Reddy, S. G., Sacchettini, J. C., & Blanchard, J. S. (1995) *Biochemistry 34*, 3492–3501.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Rossmann, M. G., Liljas, A., Branden, C.-I., & Banaszak, L. J. (1975) *Enzymes (3rd Ed.) 11A*, 61–102.

Schrumpf, B., Schwarzer, A., Kalinowski, J., Pühler, A., Eggenling, L., & Sahm, H. (1991) *J. Bacteriol. 173*, 4510–4516.

Scrutton, N. S., Berry, A., & Perham, R. N. (1990) *Nature 343*, 38–43.

Sem, D. S., & Kasper, C. B. (1993) *Biochemistry 32*, 11548–11558.

Sundharadas, G., & Gilvarg, C. (1967) *J. Biol. Chem. 242*, 3983–3988.

Tamir, H., & Gilvarg, C. (1974) *J. Biol. Chem. 249*, 3034–3040.

Tronrud, D. E., Ten Eyck, L. F., & Mathews, B. W. (1988) *Acta Crystallogr., Sect. A 43*, 489–501.

Wang, Bi-C. (1985) *Methods Enzymol. 115*, 90–112.

Weinberg, S., & Gilvarg, C. (1970) *J. Bacteriol. 101*, 323–324.

White, P. J. (1983) *J. Gen. Microbiol. 129*, 739–749.

Wierenga, R. K., Terpstra, P., & Hol, W. G. J. (1986) *J. Mol. Biol. 187*, 101–107.

BI941966O