REFERENCES
Alcock, N. W., Benton, D. J., & Moore, P. (1970) *Trans. Faraday Soc. 66*, 2210–2213.
Bray, R. C., & Watts, D. C. (1966) *Biochem. J. 98*, 142–148.
Bray, R. C., Chisholm, A. J., Hart, L. I., Meriwether, L. S., & Watts, D. C. (1966) in *Flavins and Flavoproteins* (Slater, E. C., Ed.) p 117, Elsevier, Amsterdam.
Chalmers, R. E., Parker, R., Simmonds, H. A., Snedden, W., & Watts, R. W. E. (1969) *Biochem. J. 112*, 527–532.
Eberlein, G., & Bruice, T. C. (1982) *J. Am. Chem. Soc. 104*, 1449–1452.
Ghisla, S., Hartmann, U., Hemmerich, P., & Müller, F. (1973) *Justus Liebigs Ann. Chem.*, 1388–1414.
Hemmerich, P., Massey, V., Michel, H., & Schug, C. (1981) *Struct. Bonding (Berlin) 48*, 93–121.
Hille, R., & Massey, V. (1981) *Pharmac. Ther. 14*, 249–263.
Ingold, C. K. (1969) *Structure and Mechanism in Organic Chemistry*, pp 467–468, Cornell University Press, Ithaca, New York.
Komai, H., & Massey, V. (1969) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 28*, 868.
Komai, H., & Massey, V. (1971) in *Proceedings of the Third International Symposium on Flavins and Flavoproteins* (Kamin, H., Ed.) pp 399–423, University Park Press, Baltimore, MD.
Komai, H., Massey, V., & Palmer, G. (1969) *J. Biol. Chem. 244*, 1692–1700.
Krenitsky, T. A., Elion, G. B., Strelitz, R. A., & Hitchings, G. H. (1967) *J. Biol. Chem. 242*, 2675–2682.
Leonard, N. J. (1982) *Acc. Chem. Res. 15*, 128–135.
Massey, V., Brumby, P. E., Komai, H., & Palmer, G. (1969) *J. Biol. Chem. 244*, 1682–1691.
McCollister, R. J., Gilbert, W. R., Jr., Ashton, D. M., & Wyngaarden, J. B. (1964) *J. Biol. Chem. 239*, 1560–1568.
McGartoll, M. A., & Bray, R. C. (1969) *Biochem. J. 114*, 443–444.
Moore, H. W. (1977) *Science (Washington, D.C.) 197*, 527–532.
Moore, H. W., & Czerniak, P. (1981) *Med. Res. Rev. 1*, 249–280.
Segel, I. H. (1975) *Enzyme Kinetics*, p 128, Wiley, New York.
Skibo, E. B. (1986) *J. Org. Chem. 51*, 522–527.

Crystal Structure of a Novel Trimethoprim-Resistant Dihydrofolate Reductase Specified in *Escherichia coli* by R-Plasmid R67†

David A. Matthews,*,‡ S. L. Smith,§ D. P. Baccanari,§ J. J. Burchall,§ S. J. Oatley,‡ and J. Kraut‡

Department of Chemistry, University of California—San Diego, La Jolla, California 92093, and Burroughs Wellcome Company, Research Triangle Park, North Carolina 27709

Received January 3, 1986; Revised Manuscript Received March 21, 1986

ABSTRACT: Crystalline R67 dihydrofolate reductase (DHFR) is a dimeric molecule with two identical 78 amino acid subunits, each folded into a β-barrel conformation. The outer surfaces of the three longest β strands in each protomer together form a third β barrel having six strands at the subunit interface. A unique feature of the enzyme structure is that while the intersubunit β barrel is quite regular over most of its surface, an 8-Å "gap" runs the full length of the barrel, disrupting potential hydrogen bonds between β-strand D in subunit I and the adjacent corresponding strand of subunit II. It is proposed that this deep groove is the NADPH binding site and that the association between protein and cofactor is modulated by hydrogen-bonding interactions along one face of this antiparallel β-barrel structure. A hypothetical model is proposed for the R67 DHFR–NADPH–folate ternary complex that is consistent with both the known reaction stereoselectivity and the weak binding of 2,4-diamino inhibitors to the plasmid-specified reductase. Geometrical comparison of this model with an experimentally determined structure for chicken DHFR suggests that chromosomal and type II R-plasmid specified enzymes may have independently evolved similar catalytic machinery for substrate reduction.

Dihydrofolate reductase (DHFR; EC 1.5.1.3)[1] is a very widely occurring enzyme that catalyzes the NADPH-dependent reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate. Trimethoprim (TMP) selectively inhibits bacterial DHFR and for the past 15 years has been used clinically in combination with sulfonamides as a broad spectrum antibiotic. In 1972 Fleming et al. reported that certain clinical strains of *Escherichia coli* and *Klebsiella aerogenes* carry R factors that confer high levels of resistance to TMP. In accord with already known mechanisms by which R factors confer resistance to other drugs, it was anticipated that in this case as well resistance would most likely occur because of drug inactivation or impermeability (Amyes & Smith, 1974; Benveniste & Davies, 1973). However, it was subsequently shown by Amyes and Smith (1974) and Skold and Widh (1974) that plasmids in these resistant organisms coded for a novel DHFR

† Supported by U.S. Public Health Service, National Institutes of Health, Research Grant CA17374.

* Address correspondence to this author at the Agouron Institute, 505 Coast Boulevard South, La Jolla, CA 92037.
‡ University of California—San Diego.
§ Burroughs Wellcome Co.

[1] Abbreviations: DHFR, dihydrofolate reductase; NADPH, nicotinamide adenine dinucleotide phosphate; MTX, methotrexate; DTT, dithiothreitol; MPD, 2-methyl-2,4-pentanediol; NMN, nicotinamide mononucleotide; AMN, adenosine 2′,5′-phosphate; TMP, trimethoprim.

0006-2960/86/0425-4194$01.50/0

Table I: Diffraction Data for R67 Dihydrofolate Reductase Structure Determination

derivative	resolution (Å)	Bijvoet data	unique reflections including Bijvoet	scaling R factor	no. of crystals	cell parameters
parent	$\infty \rightarrow 2.8$	no	4369	0.061	2	40.80, 46.13, 86.64
$Na_2Hg(SCN)_4$ (I)	$\infty \rightarrow 3.8$	yes	3105	0.014	1	40.75, 46.17, 86.54
$Na_2Hg(SCN)_4$ (II)	$\infty \rightarrow 2.8$	3.2 → 3.8 Å only	5344	0.036	3	40.58, 46.39, 86.83
$Pt(NH_3)_2(NO_2)_2$	$\infty \rightarrow 2.8$	no	4347	0.033	2	40.76, 45.95, 86.82
$K_3UO_2F_5$ (I)	$\infty \rightarrow 3.3$	no	2284	0.025	1	41.02, 45.63, 87.12
$K_3UO_2F_5$ (II)	$3.5 \rightarrow 2.8$	no	2718	0.053	2	40.95, 45.53, 87.03

with greatly reduced sensitivity to TMP so that microbial replication continued despite drug concentrations in excess of those required to completely inhibit the chromosomal enzyme.

Later studies by Pattishall et al. (1977) showed that R-factor-specified DHFRs fall into at least two distinct groups designated type I and type II while more recently Joyner et al. (1984) have characterized a so-called type III enzyme. Burchall (1984) has classified the R-factor-containing bacteria into two groups, those that are insensitive to TMP and those that are completely resistant. Type I and type III DHFRs confer insensitivity to TMP but probably not for identical reasons. The type I enzymes are dimeric molecules of M_r 35 000 having TMP K_i values of about 10 μM (Pattishall et al., 1977) compared to 0.4 nM for *E. coli* chromosomal DHFR. Type III DHFRs are monomeric proteins with molecular weights of around 18 000, TMP K_i's of 19 nM, and K_m's for dihydrofolate almost 10-fold lower than those observed for *E. coli* chromosomal DHFR (Joyner et al., 1984). Hence, whereas type I enzymes are insensitive to TMP primarily because of reduced affinity for the inhibitor, type III DHFR insensitivity results from weaker binding of TMP (but a higher affinity than that for the type I enzyme) augmented by an enhanced affinity for substrate. Key amino acids important for substrate and cofactor binding in chromosomal DHFRs are conserved in both type I and type III plasmid induced DHFRs, suggesting that these TMP-insensitive R-factor enzymes are related genetically to normal cellular DHFR (Kraut & Matthews, 1986 and references therein).

The type II enzyme is 1000-fold less sensitive to TMP than even the type I DHFR, and bacteria harboring plasmids specifying this reductase are resistant to therapeutic levels of the drug. Furthermore, unlike other known DHFRs, the enzyme is only very weakly inhibited by methotrexate (MTX; $K_i = 0.9 \times 10^{-4}$ M) despite having a K_m for dihydrofolate close to that for chromosomal *E. coli* DHFR (Pattishall et al., 1977). The type II plasmid encoded enzyme is reported to have a molecular weight of around 34 000 with four identical monomeric subunits M_r 8500 each (Smith et al., 1979). The protomer is a single polypeptide consisting of 78 amino acids having no apparent sequence homology with any known chromosomal DHFR (Stone & Smith, 1979). In order to further compare this novel enzyme with its chromosomal counterpart, we have crystallized the type II plasmid specified DHFR from *E. coli* (R67) and determined its structure using X-ray diffraction methods.

Preparation and Characterization of Crystals

Enzyme was isolated and purified as described earlier (Smith et al., 1979). A 20-mg sample was dialyzed against 0.001 M dithiothreitol (DTT) followed by pressure dialysis to achieve a final protein concentration of 15 mg/mL. Solid DTT was then added to the protein solution to raise its molarity to 0.02 M. Fifteen-microliter droplets of the enzyme solution were placed on previously silicone-coated glass slides to which were added 5 μL of a 50% solution of glass-distilled 2-methyl-2,4-pentanediol (MPD-Kodak). The slides were placed over reservoirs containing unbuffered solutions of 0.02 M DTT in 25% MPD. Reserviors were sealed with vacuum grease and kept at 4 °C. Crystals in the shape of large rhomboids (1.2 × 0.6 × 0.6 mm) formed within 24 h.

Examination of diffraction patterns revealed that the crystals have an orthorhombic unit cell with $a = 46.26$ Å, $b = 87.06$ Å, and $c = 40.86$ Å and strongly suggested that the space group was $P2_12_12_1$ on the basis of the absence of measurable intensity for axial reflections with an odd Miller index. If R67 DHFR is a tetrameric protein of M_r 33 766 as reported previously (Smith et al., 1979), then V_M is 1.22 Å^3/dalton assuming four tetramers per unit cell in accord with the putative space group assignment. This is well below the lowest V_M reported by Matthews (1974) for any known protein crystal, suggesting that either the initial space group assignment was incorrect or the molecular weight of the crystalline R67 DHFR molecule was much less than originally reported. It is shown below that the original choice of space group is correct but that the crystallographically repeating unit is a dimer rather than a tetramer. The true V_M is 2.44 Å^3/dalton and lies near the middle of the range for known protein crystals (Matthews, 1974).

Crystals of R67 DHFR diffract well to a minimum Bragg spacing of at least 2.0 Å and when freshly grown can be exposed to Cu Kα radiation for 150 h with a resultant intensity loss of less than 20%.

Heavy Atom Derivatives and X-ray Data Collection

Crystals of R67 DHFR grown from unbuffered solutions of aqueous MPD could be transferred successfully to 40% MPD but often dissolved following exposure to low (1 mM) concentrations of buffers or heavy atom salts. In most instances even before potentially useful heavy atom soaks could be screened on a precession camera, it was necessary to experiment with the composition of the artificial mother liquor on a case by case basis in order to prevent crystal degradation and/or large swings in pH upon addition of heavy atom compounds. The search for usable heavy atom derivatives was further complicated by three additional factors that were found in certain instances to influence the outcome of a particular heavy atom soaking experiment: (1) the age of the crystal, (2) its prior exposure to DTT, and (3) its external morphology, although in no case could we demonstrate any variation in cell parameters or relative intensities for crystals of differing external shape. Irreproducibility was a problem with both the mercuric thiocyanate and uranyl fluoride heavy atom derivatives so that in each case we were forced to work with multiple X-ray diffraction data sets that apparently did not conform to a single unique pattern of heavy atom substitution. Information on heavy atom derivatives used to phase the R67 DHFR crystal structure is presented in Tables I and II.

X-ray diffraction data were collected by using an Enraf-Nonius CAD4 single crystal diffractometer at an ambient temperature of 18 °C. Data collection procedures were similar to those reported by Volz et al. (1982). Crystal decomposition

Table II: Soaking Conditions for Heavy Atom Derivatives

heavy atom reagent	conditions	soak time
$Na_2Hg(SCN)_4$ (I)	4 mM in 50% MPD	22 h
$Na_2Hg(SCN)_4$ (II)	1 mM in 45% MPD	20 h
$Pt(NH_3)_2(NO_2)_2$	3 mM in 33% MPD	1 month
$K_3UO_2F_5$ (I)	1.5 mM in 40% MPD	1 day
$K_3UO_2F_5$ (II)	2 mM in 32% MPD	2 weeks

was followed by monitoring the intensities of four standard reflections, and except as noted below, data collection on each individual crystal was terminated when an intensity decrease of 20% was observed for any standard reflection. Crystals soaked in uranyl fluoride were especially sensitive to X-ray damage, and for each crystal, intensity decreases at the termination of data collection were 35%. All diffraction data were reduced with the XTAL 80 programs (Hall et al., 1980) on a VAX 11/780 computer.

Heavy atom difference Patterson syntheses showed two nearly equally occupied sites for each mercury and platinum derivative while the uranium derivative showed three sites with significantly different occupancies. In each case the Patterson functions could only be interpreted by assuming the true space group to be $P2_12_12_1$. Difference Patterson syntheses for each of the mercuric thiocyanate partial data sets (see Table III) showed the same two sites of mercury substitution. However, careful scaling between data sets established that two quantitatively different substitution patterns could be distinguished. Therefore, the mercuric thiocyanate heavy atom diffraction data were not merged but instead treated as distinct separate heavy atom derivative data for the purpose of multiple isomorphous replacement phase refinement. Two distinct substitution patterns were also found for crystals soaked in uranyl fluoride. There are three uranium sites common to both $K_3UO_2F_{5(I)}$ and $K_3UO_2F_{5(II)}$ but the relative occupancies of these sites are distinctly different in each (see Table III).

Phase refinement with all five heavy atom derivatives was performed by using the multiple isomorphous replacement phase refinement program included in the XTAL 80 package. The refinement procedure was the customary minimization of weighted closure errors, incorporating anomalous scattering data when available except that the residuals were computed in terms of F^2 as proposed by Hendrickson and Lattman (1979) rather than in terms of F. Statistics from the last cycle of heavy atom parameter refinement are given in Table III.

Structure Determination

A preliminary electron density map based on centroid phases was calculated on a scale of 4 Å/cm and displayed on plexiglass plates. The overall protein envelope was easily distinguishable from solvent regions, and long stretches of continuous electron density were readily identifiable. The asymmetric unit consists of two monomeric subunits related by a dyad axis perpendicular to the crystallographic b axis. The mercury heavy atom sites are positioned equivalently in the two subunits, each one two residues removed from a flat electron density rich side chain that we tentatively identified early on as Trp-45 on the basis of its proximity to the anticipated mercury substitution site, namely, Cys-47. This interpretation was confirmed by locating density for a second large flat side chain nine residues from the putative Cys-47, where we expected to find a second tryptophan (Trp-38). At this point detailed model building was begun by using a map displayed on an Evans and Sutherland Picture System II.

Backbone density was strong and continuous for both subunits from residue 22 to the carboxy terminus except for a break in the density of both subunits at Gly-44. Proceeding backward toward the amino terminus from residue 22, the density weakened for the next five or six residues in each subunit and then became disjointed, diffuse, and difficult to follow. With the exception of Val-8, the first 10 amino acids are hydrophilic, and it is likely that residues 1–10 are projecting out into solution and are probably completely disordered.

A starting set of coordinates consisting of all atoms in residues 22–78 and 23–78 in subunits I and II, respectively, was used in conjunction with the 2.8-Å diffraction data to initiate the parameter optimization procedure of Hendrickson and Konnert (1980). Our hope was that by appropriate combination of multiple isomorphous replacement and calculated phases on the basis of a refined partial model, a new electron density map could be computed into which some of the remaining amino-terminal residues missing from our initial model could be fit. The refinement process converged after 10 cycles to a crystallographic R of 0.30. A new set of centroid phases was then derived by combining multiple isomorphous replacement and model-phase probabilities (Hendrickson & Lattman, 1970). Combined phases were used for all reflections within the resolution range 6–3.5 Å while at low resolution (<6 Å) and at high resolution (>3.5 Å) the multiple iso-

Table III: Statistics from the Last Cycle of Heavy Atom Parameter Refinement[a]

derivative	fractional coordinates x	y	z	B	PP	PPA	K	C	$\langle E\rangle/\langle F_H\rangle$	R_C	R_K
$Na_2Hg(SCN)_4$ (I)							1.69	9.99	0.39	0.49	0.12
site 1	0.1390	0.2652	0.1551	54.08	1.28	0.73					
site 2	0.1113	−0.0647	−0.1062	47.73	1.38	0.89					
$Na_2Hg(SCN)_4$ (II)							1.51	−1.73	0.45	0.52	0.21
site 1	0.1449	0.2627	0.1534	30.40	1.65	0.73					
site 2	−0.1138	−0.0606	−0.1066	35.42	1.66	1.12					
$Pt(NH_3)_2(NO_2)_2$							2.69	2.00	0.68	0.62	0.18
site 1	0.1118	0.3069	0.1895	43.22	1.23						
site 2	0.0963	0.4068	0.5810	55.16	1.42						
$K_3UO_2F_5$ (I)							3.94	11.51	0.47	0.48	0.19
site 1	0.1900	0.0936	0.1598	29.74	0.92						
site 2	0.3401	0.1296	0.8949	50.34	0.84						
site 3	0.1471	0.3387	0.7424	17.58	1.09						
$K_3UO_2F_5$ (II)							3.00	2.00	0.43	0.51	0.17
site 1	0.1908	0.0961	0.1609	37.17	1.02						
site 2	0.3377	0.1244	0.8980	53.16	1.67						
site 3	0.1489	0.3394	0.7434	25.35	1.15						

[a] Mean figure of merit 0.68; B, isotropic temperature factor; PP, relative occupancy; PPA, relative anomalous electron occupancy; K, linear scale factor; C, exponential scale factor; E, closure error, F_H, calculated heavy atom contribution; $R_C = \sum|F_{PH} \pm F_P| - |F_{H,calcd}|/\sum|F_{PH} - F_P|$; $R_K = \sum|F_{PH,obsd} - F_{PH,calcd}|/\sum F_{PH,obsd}$.

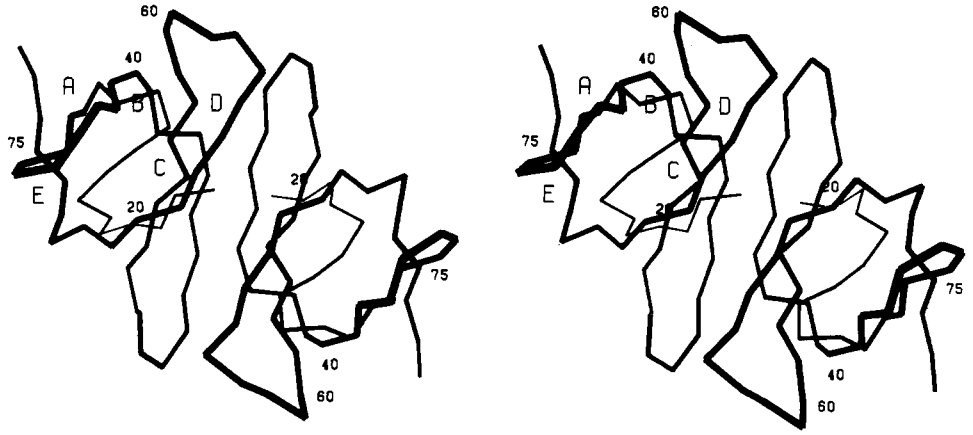

FIGURE 1: α-Carbon backbone conformation for the R67 dihydrofolate reductase dimer viewed approximately along the molecular twofold axis. The amino-terminal tail in both subunits is apparently disordered (see text).

morphous replacement and calculated phases were chosen, respectively (Remington et al., 1982). Electron density computed in this way was significantly improved compared to the original density map calculated by using only the multiple isomorphous replacement phases. In particular, it was now possible to unambiguously assign all the electron density extending back toward the amino terminus in both subunits as far as Phe-18 at which point the density again grew weak and ill-defined. Following this round of model building the new expanded coordinate set was used as input to a second round of Hendrickson–Konnert refinement. The model reported here is the result of 10 additional least-squares cycles, has a calculated R of 0.28 for all data between 8 and 2.8 Å, and includes all atoms in residues 18–78 of both subunits.

Polypeptide Chain Folding

The polypeptide chain conformation for residues 18–78 of R67 DHFR is shown in Figure 1. Each monomeric subunit is best described as approximating a pure up and down six stranded antiparallel β barrel except that in this case the fifth strand is absent. Strand D is connected directly to the carboxy-terminal strand of the β barrel by three consecutive residues in 3_{10} conformation forming two interlocked tight turns that redirect the last strand away from strand D and into a position where it completes the β barrel by forming classical antiparallel hydrogen bonds with the beginning amino strand. The second and third β strands (B and C) are considerably longer than the other three, forming a strongly twisted two-stranded ribbon running the full width of the molecule. The outer edge of this ribbon is aligned in an antiparallel fashion with the corresponding strand in the other protomer to form a sheet with intersubunit antiparallel hydrogen bonds. As a consequence of this geometry, the outer surfaces of strands B–D from one barrel associate with corresponding strands from the second subunit such that together they form a third β barrel having six strands at the protomer–protomer interface.

A qualitatively similar arrangement of β barrels has been observed in the immunoglobin V_L dimer (Epp et al., 1974), although the size of the barrel (eight strands) and the connectivity between individual strands is distinctly different from that reported here for R67 DHFR. The eight-stranded intersubunit β barrel observed for the V_L immunoglobin dimer is regular in that each strand is hydrogen-bonded to adjacent strands on both sides producing a continuously hydrogen-bonded cylindrical sheet. A unique feature of the R67 DHFR structure is that while the β barrel at the subunit interface is quite regular over most of its surface, an extended gap, 8 Å in width, running the full length of the barrel disrupts hydrogen bonding between β-strand D in subunit I and the adjacent corresponding strand of subunit II. This slit in the intersubunit B barrel opens into a pronounced cleft which traverses one whole face of the dimer. Although as yet we have no electron density maps for NADPH binding to R67 DHFR, we postulate that this deep groove that separates the subunits and slices through the intersubunit β barrel at the protomer–protomer interface is, in fact, the cofactor binding site. Details of this proposal and a model for NADPH and folate bound to R67 DHFR are presented below. If subsequently confirmed by diffraction experiments, this represents a novel mode of protein–cofactor interaction in which unsatisfied hydrogen bonding along one edge of an antiparallel β-barrel structure provides the locus for dinucleotide binding.

There are compelling reasons for believing that the amino-terminal tail of R67 DHFR has little if any role in ligand binding or substrate turnover and probably is unimportant for correct protein folding as well. At least three lines of evidence support this conclusion. (1) The putative cofactor and substrate binding sites on the one hand and the protomer amino termini on the other are located on opposite sides of the intersubunit β barrel. (2) Most of the amino acid sequence differences between the two type II R plasmid DHFRs of known primary structure (R67 and R388) cluster within the 21 residues at the amino terminus where there are 11 variations in amino acid type while for the succeeding 65 residues there are but four differences (Zolg & Hanggi, 1981). (3) Finally, and most definitive is the fact that R67 DHFR cut at Phe-16 by chymotrypsin retains full enzymic activity (E. Howell, private communication).

Structural Resemblance between R67 DHFR and the R Subunit of Papain

DHFR from bacteria and vertebrates is an M_r 18 000–22 000 monomeric protein that folds into an eight-stranded mixed β sheet surrounded by four helices and extensive loop regions connecting these secondary structural elements [for a recent review, see Kraut & Matthews (1986)]. As evidenced by its dissimilar primary, secondary, tertiary, and quaternary structure, we can state with considerable confidence that the R67 enzyme must be unrelated genetically to the ubiquitous chromosomally encoded DHFRs. Virtually nothing is known concerning the possible antecedent function and genesis of the type II plasmid encoded DHFR. However, numerous other proteins of known three-dimensional structure consist in part of highly twisted antiparallel β sheets of varying size, shape, and topology so we were encouraged to compare the R67 DHFR structure with β-barrel type folds that occur in proteins

Table IV: Topologically Equivalent α Carbons Used in Comparing β Barrels in R67 Dihydrofolate Reductase and Papain

conformation	R67 DHFR	papain	R67 DHFR	papain
strand A	27–32	129–134	27–32	129–134
strand B	40–47	161–168	40–47	161–168
strand C	55–60	171–176	55–60	171–176
strand D	65–69	185–189	65–69	185–189
strand E	75–77	205–207		
residues equivalenced	28		25	
rms difference (Å)	1.76		1.25	

with coordinates on deposit at the Protein Data Bank. Our search has turned up a intriguing structural resemblance between R67 DHFR and the R domain of the thiol protease papain.

The carboxy-terminal domain of the papain molecule folds into an antiparallel β barrel with simple +1 connections, meaning that adjacent β strands are sequentially contiguous and thus there are no connecting segments across intervening strands of sheet. The β-barrel connectivity for papain is therefore topologically identical with that of R67 DHFR. What is even more intriguing however is the close geometrical similarity between the two structures.

In order to establish a mapping of R67 DHFR α-carbon positions onto structurally analogous carbon atoms in papain (Kamphuis et al., 1984), we attempted to superpose the β-barrel portions of the two structures using an Evans and Sutherland PS II computer graphics terminal. The four major β strands in each molecule have nearly identical twists so that α-carbon equivalences were readily determined by visual alignment of corresponding strands. As is the case for R67 DHFR, the R-domain β barrel of papain has "gaps" in hydrogen bonding owing to the absence of an intervening β strand between strand D and the carboxy-terminal strand. In papain the region between these two strands consists of a stretch of 14 amino acids folded into a series of tight interlocking turns whereas in R67 DHFR the corresponding link is a much shorter stretch of only four residues. As a result the terminal β strands in these two molecules twist across the respective beginning β strands at somewhat different angles. One consequence of this structural variation is that least-squares refinement of our initial visual alignment of R67 DHFR and papain β barrels, to be discussed next, gives significantly higher residuals for α carbons in the carboxy-terminal strands than for α carbons in strands A–D.

Table IV lists the α-carbon pairs that were equivalenced for each comparison and gives the final root-mean-square (rms) separation distances obtained by using the Rao and Rossmann (1973) algorithm for refining independent sets of coordinates to an optimal least-squares fit. A comparison in which α-carbon coordinates 18–78 for one subunit of R67 DHFR were fit to corresponding atoms in the second crystallographically independent subunit gave an rms value of 0.67 Å.

These results reveal a striking geometrical similarity between the β-barrel portions of the R67 DHFR protomer and the R domain of papain. This overall topographical resemblance exists despite some differences that might have been expected to weaken the observed three-dimensional similarity including (1) the intimate association of the three longest strands in one R67 DHFR protomer with corresponding strands in the second protomer to form the intersubunit β barrel, (2) the longer connecting loops between individual β strands in papain, and (3) the absence in R67 DHFR of a disulfide bridge analogous to that between Cys-153 and Cys-200 in papain.

The rms deviation for 28 equivalenced pairs of atoms in R67 DHFR and papain (1.76 Å) is somewhat lower than that observed by Richardson et al. (1976) in their structural comparison of the immunoglobin domain and the copper, zinc superoxide dismutase subunit but quite similar to values reported by Rossmann et al. (1975) for geometrical alignments between β-sheet portions of dinucleotide binding domains in three apparently evolutionarily related dehydrogenases. Note that when the three α carbons in strand E are eliminated from the calculation, the rms difference is reduced from 1.76 to 1.25 Å in accord with our previous observation concerning geometrical differences between the terminal β strands in the two structures.

In the past when close structural similarities between proteins of unrelated function have been noted, it invariably raises questions concerning possible evolutionary relationships. The only definitive proof that such a relationship exists would be to demonstrate clear amino acid sequence homology between the two proteins. No such homology is found when amino acids for the α-carbon pairs listed in Table IV are compared. In fact out of 28 such pairs there are only two instances in which the equivalenced α carbons correspond to identical amino acids. This negative result does not disprove an evolutionary relationship between R67 DHFR and the R domain of papain, but it does allow for an alternative interpretation of the observed structural resemblances, namely, that constraints of protein stability and folding requirements may permit only a limited repertoire of independent three-dimensional structural domains and therefore, significant geometrical resemblances between unrelated proteins are likely to occur more frequently than would be expected from "random" probability alone (Richardson et al., 1976).

Model for the R67 DHFR–NADPH–Folate Ternary Complex

The past 10 years has seen a dramatic increase in the number of proteins whose three-dimensional structures have been "solved" with an accuracy at least sufficient to permit tracing the polypeptide backbone folding. This wealth of geometrical information has encouraged scientists to search for common structural motifs among such proteins, both as a method for simply grouping and cataloging them and as a means for attempting to correlate protein structure with function. One significant observation resulting from these types of studies is that proteins that bind pyridine dinucleotides, and in many cases mononucleotide binding proteins as well, often utilize geometrically similar super secondary structural domains consisting of parallel β sheets flanked by adjoining α helices (Rossmann et al., 1975). Hol et al. (1978) have pointed out a further similarity among proteins that bind pyridine dinucleotides: the presence of α-helical dipoles in close proximity to pyrophosphate binding sites where they presumably stabilize the coenzyme's negatively charged phosphate groups. Exceptions to this generalization occur in the case of phosphorylase *b* where no such α-helix–phosphate interaction is found for NAD bound at the allosteric site (Stura et al. 1983) and also for both 6-phosphogluconate dehydrogenase (Adams et al., 1983) and beef liver catalase (Fita & Rossman, 1985) which bind NADP near the carboxy ends of α helices. A key finding of the present study is that R67 DHFR has neither parallel β sheet nor α-helical secondary structure, and therefore, it follows that NADPH binding to the plasmid-encoded reductase must be significantly different from that observed for any other pyridine dinucleotide dependent enzyme of known structure.

Unfortunately we have been unable to bind NADPH to apoenzyme crystals of R67 DHFR by simply exposing pregrown crystals to free cofactor in solution, nor have we been

successful in obtaining isomorphous crystals of the binary complex by the alternative strategy of first forming the complex in solution and then crystallizing it. Had either of these approaches worked, we could have examined dinucleotide binding to the plasmid-encoded reductase by straightforward crystallographic methods. Nevertheless, using molecular modeling techniques, we have identified putative binding sites for both NADPH and dihydrofolate. In what follows we discuss the methods used in constructing a detailed model for the enzyme–NADPH–folate ternary complex and go on to show how the proposed model can account in a self-consistent way for known properties of the enzyme. In retrospect we now realize that out inability to obtain isomorphous crystalline complexes with NADPH or folate is a result of the way adjacent molecules pack in the crystallographic unit cell, causing the enzyme's active sites to be blocked by carboxy-terminal residues of symmetry-related molecules. We have however been able to make some use of this unfortunate twist of fate in helping to guide our hypothetical model-building experiments. In particular, the backbone NH, COO^-, and side-chain carboxamide of the C-terminal residue Asn-78 are all involved in hydrogen-bond interactions at the putative active site of an adjacent molecule and can therefore serve to identify protein backbone and side-chain groups within the active site that may be important in binding cofactor and substrate.

Initially there were two pieces of evidence that suggested to us that NADPH probably binds along the lengthwise "slit" in the intersubunit β barrel composed of strands B–D from both protomers. In the first place earlier studies in solution demonstrated that less than one molecule of NADPH is bound per 78 amino acid protomer (Smith & Burchall, 1980) and therefore the cofactor binding site is most probably shared between subunits. Furthermore, careful examination of the R67 DHFR structure uncovered no alternative site that could possibly accommodate a molecule the size of NADPH without major alterations in the enzyme's three-dimensional architecture. We even considered the possibility that perhaps only the NMN portion of NADPH bound specifically to the enzyme, thus reducing by half the required size of the cofactor binding site. This possibility was rejected because NMN and other analogues of NADPH in which the adenine mononucleotide is altered are extremely poor cofactors (Smith & Burchall, 1983). We did however note two other clefts, considerably smaller in size than the putative NADPH binding site, located at opposite ends of the intersubunit β barrel, which may be the binding sites for substrate. Smith and Burchall (1980) have shown that, for dihydrofolate, the stoichiometry of binding is one substrate molecule per R67 DHFR protomer.

Our modeling studies indicate that in order for NADPH to fit snugly into the lengthwise opening in the intersubunit β barrel while simultaneously positioning its nicotinamide and adenine rings in close proximity to the putative substrate binding sites at opposite ends of the β barrel, it is necessary that the cofactor have an open, extended conformation similar to that observed in other protein–nucleotide complexes. Detailed model building was guided by attempts to maximize favorable protein–NADPH hydrogen bonding and van der Waals interactions subject to conformational energy constraints that restrict free rotation about various NADPH single bonds and consistent with the requirement that the nicotinamide ring be positioned for hydride transfer from its C4 carbon to the C6 *si* face of dihydrofolate (Smith & Burchall, 1983 and references therein). For chromosomal DHFRs, the A side hydrogen of nicotinamide is transferred to substrate (Pastore & Friedkin, 1962) while for R67 DHFR the stereoselectivity of hydride transfer is unknown. However, Benner's suggestion that the stereoselectivity of nicotinamide dinucleotide dependent dehydrogenases correlates with the thermodynamic stability of their substrates (Benner, 1982) predicts A-side transfer for the type II plasmid specified reductase as well. In agreement with this prediction attempts to construct a model on the basis of hydrogen transfer from the B side of nicotinamide were unsuccessful since the proposed binding site will not accommodate cofactor in the syn conformation.

Another important consideration that further restricts possible NADPH torsion angles is the presence of a molecular twofold axis that relates one subunit to the other, bisects the intersubunit β barrel, and passes throught the center of the putative cofactor binding site perpendicular to the barrel's surface. If the NADPH molecule binds to R67 DHFR in a groove having twofold symmetry, then the molecular dyad and the dinucleotide's pseudo twofold axis must coincide. This observation further reduces the conformational space that must be searched in order to optimize a fit between NADPH and the enzyme because it at least requires that $\alpha_n \equiv \alpha_a$ and $\beta_n \equiv \beta_a$ [see Saenger (1984) for atomic numbering and torsion angle definitions].

We began modeling NADPH binding to R67 DHFR by requiring that the molecular dyad pass through the cofactor's bridging pyrophosphate oxygen and then proceeded to systematically explore possible translations of the NADPH molecule along this twofold axis, global rotations about the axis, and internal torsional rotations around specific NADPH single bonds subject to the constraints and considerations discussed earlier. In the case of mono- and dinucleotides, rotational freedom about single bonds is usually restricted to a small subset of sterically allowed conformations. The extensive literature on this subject has been tabulated and reviewed by Saenger (1984) and is not considered further here. Although, as mentioned above, we required that α and β angles for both mononucleotides be equal, torsion angles χ_a and χ_n were allowed to vary independently for each mononucleotide unit since orientation about the glycosyl bond will be influenced both by the nature of the bases themselves and by differences in how dissimilar bases are accommodated in a symmetrical cofactor binding site. Similar arguments suggest that γ_a and γ_n should also be unteathered. For nucleotides, the preferred sugar puckering modes are either C2′-endo or C3′-endo. Both possibilities were explored independently for each ribose ring of the cofactor. Finally, we assumed that the enzyme itself would not undergo significant conformational rearrangement concomitant with binding NADPH.

With regard to how substrate might bind to the reductase, we have already called attention to clefts at each end of the intersubunit β barrel. In modeling folate binding at this site we attempted to maximize polar interactions between the pteridine ring and specific hydrogen-bonding groups of the enzyme while at the same time keeping in mind that the C6 carbon should be positioned close to C4 of nicotinamide and oriented to agree with the known stereoselectivity of the hydride-transfer reaction (Pastore & Friedkin, 1962; Charlton et al., 1979). The substrate's *p*-aminobenzoyl glutamate must then project out of the cavity and interact with surface residues adjacent to the active site. Because there are few a priori constraints on how this long flexible side chain might be accommodated at the surface of the R67 DHFR molecule, no attempt was made to predict its conformation or binding site.

The above considerations together provide a stringent set of conditions that led directly to the proposed model of the R67 DHFR–NADPH–folate ternary complex shown in Fig-

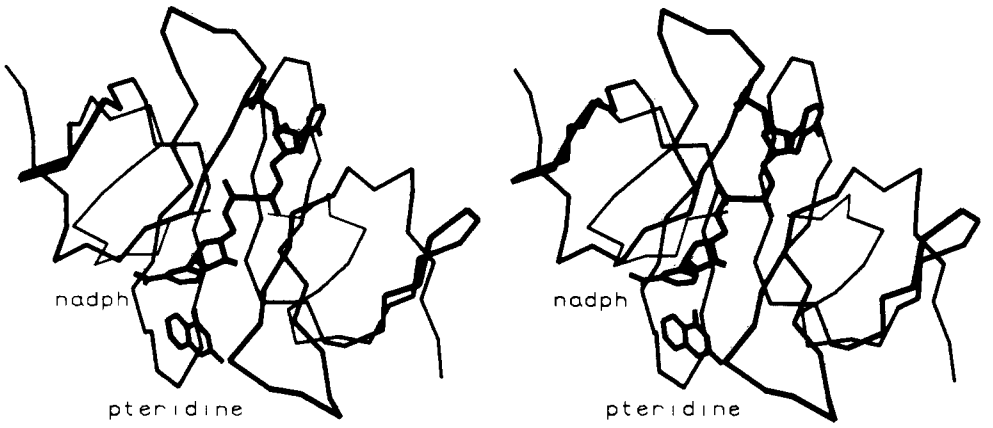

FIGURE 2: Hypothetical model for NADPH and 2-amino-4-oxopteridine binding to an α-carbon backbone representation of R67 dihydrofolate reductase.

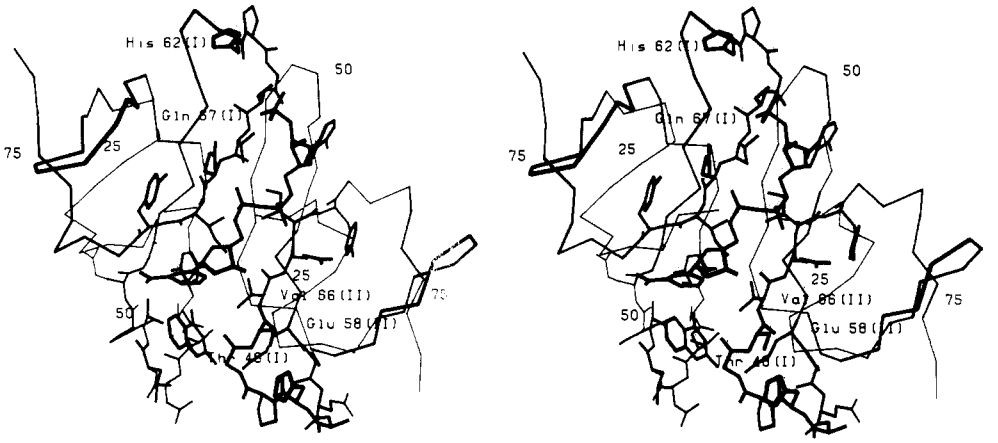

FIGURE 3: View of the hypothetical R67 dihydrofolate reductase–NADPH–pteridine ternary complex similar to that in Figure 2 but showing, in addition, main-chain and side-chain atoms that may be important in binding cofactor and substrate.

ures 2 and 3. The most significant features of the model can be described as follows. Conformations for the two individual mononucleotide portions of NADPH are very similar. When Klyne–Prelog nomenclature is used, orientations about the C4′–C5′ bond γ, the C5′–O5′ ester bond β, and the O5′–P ester bond α are respectively antiperiplanar, antiperiplanar, and –synclinal for both ribosyl 5′-phosphate moieties. Torsional angles about the P–O bridge bonds together with α rotations specify the geometry of the pyrophosphate bridge and disposition of the two ribose moieties around the P–P axis. Viewed down the P–P axis the model predicts that the six attached oxygens (excluding the pyrophosphate bridge oxygen) are arranged in a staggered conformation with the O5′ ester oxygens –anticlinal to one another. Suggested sugar puckerings are 2′-endo and 3′-endo for the NMN and AMN ribose rings, respectively, while both bases are in the anti conformational range with torsion angles about the glycosyl bonds of approximately 120°. Its interesting to note that the hypothetical model for NADPH bound to R67 DHFR predicts the same overall extension for the cofactor, 16.7 Å between C6 (adenine) and C2 (nicotinamide), as was found experimentally in our earlier high-resolution structural studies of NADPH binding to *L. casei* and chicken DHFR (Filman et al., 1982; Matthews et al., 1985a). The corresponding distance for NAD^+ bound to lactate and malate dehydrogenases is some 2.5 Å shorter (Rossmann et al., 1975).

Our model predicts that a key structural feature of R67 DHFR important for cofactor recognition is the presence of two β chains, strand D from each protomer, running antiparallel to one another separated by a distance of 8 Å and twisted so as to expose main-chain amido and carbonyl groups that can provide a complementary docking site for the cofactor's pyrophosphate backbone. Torsional movement around the C4′–C5′ bond of NADPH changes the position of the ribose ring relative to the attached 5′-phosphate groups and for γ angles near 180° (antiperiplanar) opens up the nucleotide units exposing the sugar and phosphate groups for mutual recognition and binding contacts with the protein backbone to ensure correctness of fit. Model-building experiments suggest that the 5′-phosphate group of NMN hydrogen bonds directly with the main-chain amido of Ile-68(I) and with backbone carbonyls of Val-66(II) and Ile-68(I) via intervening water molecules. The side chain of Gln-67(I) may provide a second hydrogen-bonding interaction with the phosphate O2. By symmetry, the AMN portion of NADPH makes identical interactions in the upper half of the cofactor binding site as pictured in Figure 3.

The important point to note is that, in contrast to other proteins of known structure which stabilize anion binding by salt links or favorable electrostatic interactions with macrodipoles from amino ends of nearby α helices (Wierenga et al., 1985 and references therein), in the case of the plasmid-encoded DHFR fixed hydrogen-bond dipoles alone may be sufficient to perform an analogous function. In this regard it is interesting to note that recently a bacterial sulfate binding protein has been shown to sequester sulfate anions in a water-inaccessible cleft also without benefit of salt linkages to protein side chains (Pflugrath & Quiocho, 1985). However, in this case the anion binding site is situated at a juncture formed by the amino termini of three α helices that, according

to the arguments of Hol et al. (1978), should provide a strong dipole field especially suited for stabilizing anion binding.

Our analysis suggests that the O5′ ester oxygen and the NMN ribose 2′-hydroxyl probably make water-mediated hydrogen bonds with the side chains of Tyr-69(I) and Ser-65(II), respectively. Similar interactions are predicted to occur between protein and corresponding portions of AMN except that the presence of a phosphate group attached to the 2′ carbon of the AMN ribose provides additional possibilities for anchoring this portion of the cofactor into a shallow cleft at the protein surface. The model suggests that in addition to Ser-65(I), the side chains of His-62(I) and/or Gln-67(I) could also hydrogen bond with the 2′-phosphate oxygens. The side chain of His-62(I) in the apoenzyme is too distant from the hypothetical phosphate binding site for direct interaction, but rotations about Cα–Cβ and Cβ–Cγ concomitant with ligand binding could easily reposition the imidazole ring for optimal hydrogen bonding, or alternatively a hydrogen-bonded link between the two could be mediated by an intervening solvent molecule. Water-mediated hydrogen bonds apparently provide a useful mechanism for linking dinucleotides and presumably other ligands as well to their protein binding sites. For example, in an earlier high-resolution crystallographic study of NADPH binding to *L. casei* DHFR, we found seven instances in which hydrogen-bonding contacts between protein and cofactor were mediated by fixed water molecules (Filman et al., 1982).

The nicotinamide and adenine bases are sandwiched between protomers at opposite ends of the gap or slit in the intersubunit β barrel. Side-chain hydroxyl groups from Tyr-46(I) and Tyr-46(II) are positioned to hydrogen bond with the 6-amino group of adenine and the carboxamide of nicotinamide, respectively, while a water-mediated hydrogen-bonding link may occur between the adenine's N7 and the backbone carbonyl of Ile-68(II).

It is worth reiterating that because of the molecular dyad relating the two R67 DHFR subunits, the nicotinamide and adenine binding pockets are equivalent, meaning that either site can bind either base. What is even more remarkable is that the enzyme functions about equally well with α-NADPH or the more well-known β epimer (Smith & Burchall, 1983) and therefore can presumably accommodate α-NMN, β-NMN or AMN in the same binding pocket!

The major groove separating the two R67 DHFR protomers, into which we have fit a molecule of NADPH, makes an abrupt turn in the neighborhood of the nicotinamide binding site and terminates in a pronounced cavity at the end of the intersubunit β barrel. The molecular architecture of this cleft defines a receptor site into which the pteridine portion of a folate molecule can be inserted with remarkable precision. There is but a single binding mode that provides convincing geometrical and chemical complementarity between substrate and enzyme, and this geometry is consistent with the known stereoselectivity of hydride transfer to the C6 *si* face of substrate. As discussed below, the proposed model also provides a simple explanation for the intriguing observation that MTX, TMP, and other 2,4-diamino heterocycles are very poor inhibitors of type II plasmid encoded DHFRs.

In an attempt to define the principal interactions between the pteridine ring of folate and the enzyme, we have modeled the fit of the 4-keto tautomer to the putative binding site. The 4-keto tautomer of dihydrofolate is known to be the stable form in solution (Pfleiderer et al., 1960; Brown & Jacobsen, 1961) and when bound to *L. casei* DHFR (Hood & Roberts, 1978). The substrate's pteridine ring is sequestered from solvent on one side by the cofactor's nicotinamide ring, the side chain of Tyr-46(I), and the loop connecting strands B and C in subunit I (Figure 3). The opposite edge of the cleft is formed by polypeptide backbone and side-chain atoms from an extended loop that connects strands C and D in subunit II as well as by amino acid side chains from Glu-58(II) and Val-66(II). Thus, residues from both subunits are important in binding both substrate and cofactor. The model suggests that the 4-keto oxygen of dihydrofolate can accept a hydrogen bond from the side-chain hydroxyl of Ser-51(I) and possibly from the hydroxyl of Tyr-46(I) as well. Thr-48(I) is positioned to make a hydrogen bond with N3 while the carboxylate group of Glu-58(II) apparently accepts a hydrogen bond from the substrates 2-amino group. As has been observed for substrate binding to chicken DHFR, the pyrazine portion of the pteridine ring is exposed to solvent on one edge and does not participate in specific hydrogen bonding with the protein.

Active Site Similarities among Chromosomal and R67 DHFRs: Implications for Catalysis

One reason for our initial interest in studying the plasmid encoded type II DHFR was the likelihood that this enzyme had evolved completely independently of its chromosomal counterpart. Since both enzymes catalyze the same reaction, we anticipated that a detailed comparison of the two three-dimensional structures might direct attention to certain specific chemical and geometrical features that the active sites have in common and thus provide additional insight into the catalytic mechanism. This assumes that reduction of dihydrofolate proceeds by a similar mechanism in both enzymes, which is unproven but at least consistent with the known reaction stereoselectivity at C6 of dihydrofolate which is the same for both chromosomal and type II plasmid specific DHFRs (Smith & Burchall, 1983). We describe below comparisons of active site geometries between our hypothetical model for the R67 DHFR–NADPH–folate ternary complex and an experimentally determined structure for the chicken DHFR–NADP–biopterin complex. An important discovery of the current study is that two conserved active site amino acids in chromosomal DHFRs have structural counterparts in the postulated active site of the R67 plasmid enzyme but their locations relative to the substrate's bound pteridine ring are interchanged.

Analysis of refined high-resolution X-ray structures for DHFRs from *E. coli*, *L. casei*, and chicken as well as sequence comparisons with a total of 10 chromosomal DHFRs of known primary structure reveals that only two polar residues that interact with the substrate's pteridine ring are conserved across all enzyme species—an active site carboxylic acid (glutamic acid in vertebrate DHFRs and aspartic acid in bacterial enzymes) and a threonine residue that hydrogen bonds with both the conserved acidic side chain and an invariant water molecule linked directly to the substrate's 2-amino group (Kraut & Matthews, 1986). Recent site-directed mutagenesis experiments on *E. coli* DHFR suggest the following roles for these two residues in the chromosomal DHFRs. Thr-113,ec[2] apparently plays a purely structural role in providing hydrogen bonds required to bind substrate with a low K_m (Chen et al., 1986). Mutagenesis studies on the effect of replacing the active site carboxyl in *E. coli* DHFR, Asp-27, with asparagine or serine suggest that the function of this residue must be to catalyze protonation of substrate prior to hydride transfer from NADPH (Howell et al., 1986), in agreement with an earlier

[2] The symbols "ec" and "cl" are appended to residue numbers to indicate the *E. coli* and chicken enzymes, respectively.

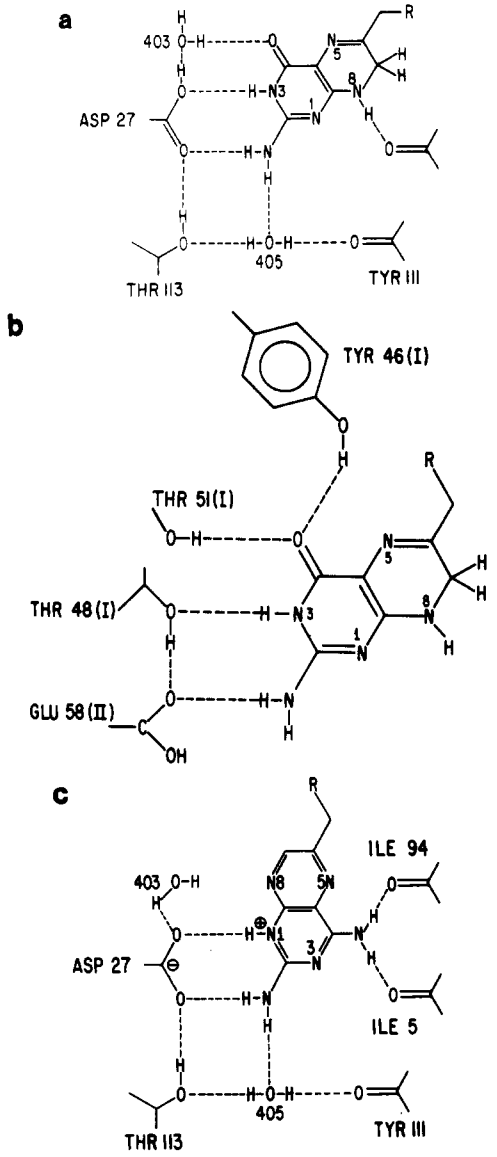

FIGURE 4: (a) Probable hydrogen bonding interactions between *E. coli* dihydrofolate reductase and 2-amino-4-oxopteridine. (b) Inferred hydrogen bondings between R67 dihydrofolate reductase and 2-amino-4-oxopteridine on the basis of hypothetical model-building experiments described in the text. (c) Hydrogen bonds between *E. coli* dihydrofolate reductase and the 2,4-diaminopteridine portion of MTX.

suggestion made originally by Matthews et al. (1978). In the course of our studies on ligand binding to chicken DHFR, we have determined the 2.8-Å structure of an abortive ternary complex containing enzyme, NADP, and an inferior substrate, biopterin [L-erythro-6-(1,2-dihydroxypropyl)pterin]. These experimental results essentially confirm the model for substrate binding proposed earlier by Bolin et al. (1982) (see Figure 4a). Armed with an experimentally determined model for cofactor and 2-amino-4-oxopteridine binding to a chromosomal DHFR and a knowledge of which specific active site amino acids are functionally important for substrate reduction across 10 species of enzyme from bacteria to human, we can now search the putative R67 DHFR active site for any key structural features that it might have in common with its chromosomal counterpart. Our approach was simply to superpose C4 of NADP and the substrate's pteridine ring in the chicken DHFR structure with the corresponding moieties in our hypothetical model for the R67 DHFR ternary complex and then to look for functionally analogous residues that might be similarly positioned in the two active sites.

It is obvious from a geometrical comparison of these two unrelated protein structures that the overall construction of the putative pteridine binding pocket in R67 DHFR is completely different from that for the known substrate binding site in chromosomal DHFR. Despite these differences, the structural superposition described earlier revealed a remarkable geometrical correspondence between the conserved active site carboxyl and threonine residues in chromosomal DHFRs (Glu-30 and Thr-136 in the chicken enzyme) and Glu-58(II) and Thr-48(I) in the plasmid encoded reductase, except that the side-chain positions relative to bound pteridine and C6 of NADP are reversed in the two structures. Thus, Glu-30 and Thr-136,cl coincide geometrically with Thr-48(I) and Glu-58(II) in R67 DHFR, respectively (Figure 4a,b).

Although these observations do not provide conclusive evidence, they at least suggest the possibility that if our model for substrate binding to R67 DHFR is substantially correct, chromosomal and type II R-plasmid encoded DHFRs may have independently evolved similar catalytic machinery for reducing dihydrofolate. Protonation at N5 of dihydrofolate is generally believed to be a necessary preliminary to hydride transfer from NADPH (Huennekens & Scrimgeour, 1964; Scrimgeour, 1976). As mentioned earlier there is convincing evidence that the conserved active site carboxyl in chromosomal DHFR has an elevated pK_a and is involved in protonation of the substrate [see Kraut & Matthews (1986) and references therein]. Moreover, it is equally clear from the crystallographic studies that the geometrical positioning of this side chain relative to bound substrate precludes direct proton transfer to N5. Likewise, for R67 DHFR our model predicts that if proton transfer occurs from Glu-58 to the substrate's N5, then it also must be indirect. However, it is noteworthy that the overall proton transfer pathway in the two cases must be different since, viewed from the substrate position, the spatial locations of Glu-30 and Thr-136 in chicken DHFR are geometrically permuted compared to the orientations of Glu-58(II) and Thr-48(I) in the R67 enzyme. Perhaps a proton-donating group in the neighborhood of bound substrate is essential for ring protonation proceeding hydride transfer, but the exact location of this side chain with respect to the dihydrofolate molecule may be less crucial. The hypothesis could be tested by swapping amino acid types at sequence positions corresponding to the active site carboxyl and threonine residues in, for example, *E. coli* DHFR by using the methods of site-directed mutagenesis. Relative to bound pteridine the active site carboxyl and threonine side chains in the mutated *E. coli* enzyme would then be geometrically homologous with chemically identical residues in the putative active site of R67 DHFR leading to the prediction that the reengineered chromosomal enzyme should be catalytically active.

Inhibitor Binding to R67 DHFR

A novel property of R67 DHFR, and indeed an important factor accounting for its emergence among clinical isolates of plasmid-containing TMP-resistant bacteria, is the enzyme's extremely poor binding of antifolates. In what follows we attempt to account for this observation by considering how 2,4-diamino heterocycles might interact with the putative active site of the plasmid-encoded reductase, but first we briefly review what is known about MTX binding to chromosomal DHFR in order to identify the stereochemical factors responsible for potent inhibition of these reductases.

Virtually all potent inhibitors of chromosomal DHFR contain 2,4-diamino heterocycles, and it is very cler from

crystallographic and other studies of inhibitor binding to these enzymes that polar interactions between these inhibitors and specific groups lining the protein active site are of primary importance in accounting for this tight binding (Matthews et al., 1985a,b).

It is now well established that the pteridine ring of MTX when bound at the active site of chromosomal DHFR is turned over compared to that of folate or dihydrofolate [see Bolin et al. (1982)] and protonated at N1 (Blakely, 1984 and references therein). Replacing the substrate's 4-oxo group by a 4-amino group promotes tight binding in at least two ways. (1) A 4-amino substituent increases the basicity of the heterocycle's N1, facilitating protonation and permitting a very favorable interaction with the active site carboxylate via a doubly hydrogen-bonded salt linkage. (2) In contrast to an oxo group, a 4-amino substituent can donate hydrogen bonds to appropriately positioned acceptor atoms. Remarkably, it turns out that two backbone carbonyl oxygens are correctly disposed to accept hydrogen bonds from the inhibitor's 4-amino group only if the ring flips into the inverted binding mode, the same configuration that simultaneously optimizes interaction between both N1 and the 2-amino group of the inhibitor and key conserved active site side chains (Bolin et al., 1982).

In this view then, potent inhibition of chromosomal DHFR by MTX (and certain other 2,4-diamino heterocycles) occurs not because of the inhibitor's structural resemblance to the enzymic transition state as has been suggested by others (Williams et al., 1979), but rather because of the fortuitous occurrence at the substrate binding site of a molecular architecture that provides a complementary docking site for 2,4-diamino heterocycles bound in the "inverted" configuration. Therefore, unless the entire chromosomal DHFR pteridine binding site, in all its geometrical and chemical complexity, turned up intact in R67 DHFR, potent MTX inhibition of the type II DHFR would be paradoxical. In fact, as discussed earlier, the active sites of the plasmid-encoded reductase and of chromosomal DHFR are not structurally homologous. Thus, MTX and TMP are poor inhibitors of R67 DHFR owing both to the absence of hydrogen bond acceptor groups at the pteridine binding site positioned to interact with a 4-amino substituent and to the geometrical "switch" of glutamate and threonine in R67 DHFR compared to the active site orientations of the chemically identical residues in chromosomal reductases.

These structural arguments can be appreciated by referring to Figure 4 in which we compare our model for folate binding at the putative active site of R67 DHFR with substrate and MTX binding to the *E. coli* enzyme. If MTX bound to R67 DHFR in a manner similar to that shown in Figure 4b for substrate, the 4-amino group could not hydrogen bond with the side-chain hydroxyls of either Tyr-46 or Ser-51 since both these groups are positioned well out of the pteridine plane. Moreover, the carboxyl group of Glu-58 could not stabilize protonation at the inhibitor's N1 via hydrogen bonding. The "flipped" pteridine binding mode corresponding to that shown in Figure 4c for MTX bound to *E. coli* DHFR is also unfavorable. Again there would be no protein atoms positioned to accept hydrogen bonds from the inhibitor's 4-amino substituent, and N1 would reside close to the side chain of Thr-48 rather than a negatively charged acidic side chain analogous to that stabilizing binding of protonated MTX to chromosomal DHFR. Thus, our model suggests that weak binding of MTX and TMP to R67 DHFR is due, at least in part, to the location of key active site carboxylate and threonine side chains that, relative to bound substrate, are geometrically permuted compared to locations for two chemically identical side chains common to all chromosomal DHFRs of known structure.

Registry No. DHFR, 9002-03-3; NADPH, 53-57-6; 2-amino-4-oxopteridine, 2236-60-4.

REFERENCES

Adams, M. J., Archibald, I. G., Bugg, C. E., Carne, A., Gover, S., Helliwell, J. R., Pickersgill, R. W., & White, S. W. (1983) *EMBO J. 2*, 1009–1014.

Amyes, S. G. B., & Smith, J. T. (1974) *Biochem. Biophys. Res. Commun. 58*, 412–418.

Benner, S. A. (1982) *Experientia 38*, 633–637.

Benveniste, R., & Davies, J. (1973) *Annu. Rev. Biochem. 42*, 471–506.

Blakley, R. L. (1984) in *Folates and Pterins: Chemistry and Biochemistry of Folates* (Blakley, R. L., & Benkovic, S. J., Eds.) Vol. 1, Chapter 5, pp 191–253, Wiley, New York.

Bolin, J. T., Filman, D. J., Matthews, D. A., Hamlin, R. C., & Kraut, J. (1982) *J. Biol. Chem. 257*, 13650–13662.

Brown, D. J., & Jacobsen, N. W. (1961) *J. Chem. Soc.*, 4413–4420.

Burchall, J. J. (1984) in *Folate Antagonists as Therapeutic Agents* (Sirotnak, F. M., Burchall, J. J., Ensminger, W. B., & Montgomery, J. A., Eds.) Vol. 1, pp 130–150, Academic, Orlando.

Charlton, P. A., Young, D. W., Birdsall, B., Feeney, J., & Roberts, G. C. K. (1979) *J. Chem. Soc., Chem. Commun. 20*, 922–924.

Chen, J. T., Mayer, R. J., Fienke, C. A., & Benkovic, S. J. (1986) *J. Cell. Biochem. 29*, 73–82.

Epp, O., Colman, P., Fehlhammer, H., Bode, W., Schiffer, M., Huber, R., & Palm, W. (1974) *Eur. J. Biochem. 45*, 513–524.

Filman, D. J., Bolin, J. T., Matthews, D. A., & Kraut, J. (1982) *J. Biol. Chem. 257*, 13663–13672.

Fita, I., & Rossmann, M. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1604–1608.

Fleming, M. P., Datta, N., & Grüñeberg, R. N. (1972) *Br. Med. J. 1*, 726–728.

Hall, S. R., Stewart, J., & Munn, R. J. (1980) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr. Theor. Gen. Crystallogr. A36*, 979–989.

Hendrickson, W. A., & Lattman, E. E. (1970) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B26*, 136–143.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., Ed.) Vol. I, pp 43–57, Pergamon, Oxford, England.

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

Hood, K., & Roberts, G. C. K. (1978) *Biochem. J. 171*, 357–366.

Howell, E. E., Villafranca, J. E., Warren, M. S., Oatley, S. J., & Kraut, J. (1986) *Science (Washington, D.C.) 231*, 1123–1128.

Huennekens, F. M., & Scrimgeour, K. G. (1964) in *Pteridine Chemistry* (Pfleiderer, W., & Taylor, E. C., Eds.) pp 335–379, Pergamon Press, Oxford.

Joyner, S. S., Fling, M. E., Stone, D., & Baccanari, D. P. (1984) *J. Biol. Chem. 259*, 5851–5856.

Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A., & Drenth, J. (1984) *J. Mol. Biol. 179*, 233–257.

Kraut, J., & Matthews, D. A. (1986) in *Biological Macromolecules and Assemblies* (Jurnak, F., & McPherson, A., Eds.) Wiley, New York Vol. III (in press).

Matthews, B. W. (1974) *J. Mol. Biol. 82*, 513-526.
Matthews, D. A., Alden, R. A., Bolin, J. T., Filman, D. J., Freer, S. T., Hamlin, R., Hol, W. G. J., Kisliuk, R. L., Pastore, E. J., Plante, L. T., Xuong, N., & Kraut, J. (1978) *J. Biol. Chem. 253*, 6946-6954.
Matthews, D. A., Bolin, J. T., Burridge, J. M., Filman, D. J., Volz, K. W., Kaufman, B. T., Beddell, C. R., Champness, J. N., Stammers, D. K., & Kraut, J. (1985a) *J. Biol. Chem. 260*, 381-391.
Matthews, D. A., Bolin, J. T., Burridge, J. M., Filman, D. J., Volz, K. W., & Kraut, J. (1985b) *J. Biol. Chem. 260*, 392-399.
Pastore, E. J., & Friedkin, M. (1962) *J. Biol. Chem. 237*, 3802-3810.
Pattishall, K. H., Acar, J., Burchall, J. J., Goldstein, F. W., & Harvey, R. J. (1977) *J. Biol. Chem. 252*, 2319-2323.
Pfleiderer, W., Liedek, E., Lohrmann, R., & Rukwied, M. (1960) *Chem. Ber. 93*, 2015-2024.
Pflugrath, J. W., & Quiocho, F. A. (1985) *Nature (London) 314*, 257-260.
Rao, S. T., & Rossmann, M. G. (1973) *J. Mol. Biol. 76*, 241-256.
Remington, S., Wiegand, G., & Huber, R. (1982) *J. Mol. Biol. 158*, 111-152.
Richardson, J. S., Richardson, D. C., & Thomas, K. A. (1976) *J. Mol. Biol. 102*, 221-235.
Rossmann, M. G., Liljas, A., Brändén, C.-I., & Banaszak, L. J. (1975) *Enzymes (3rd Ed.) 11*, 62-102.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Scrimgeour, K. G. (1976) in *Chemistry and Biology of Pteridines* (Pfleiderer, W., Ed.) pp 731-751, de Gruyter, Berlin.
Sköld, O., & Widh, A. (1974) *J. Biol. Chem. 249*, 4324-4325.
Smith, S. L., & Burchall, J. J. (1980) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 39*, 1771.
Smith, S. L., & Burchall, J. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4619-4623.
Smith, S. L., Stone, D., Novak, P., Baccanari, D. P., & Burchall, J. J. (1979) *J. Biol. Chem. 254*, 6222-6225.
Stone, D., & Smith, S. (1979) *J. Biol. Chem. 254*, 10857-10861.
Stura, E. A., Zanotti, G., Babu, Y. S., Sansom, M. S. P., Stuart, D. I., Wilson, K. S., Johnson, L. N., & Van de Werve, G. (1983) *J. Mol. Biol. 170*, 529-565.
Volz, K. W., Matthews, D. A., Alden, R. A., Freer, S. T., Hansch, C., Kaufman, B. T., & Kraut, J. (1982) *J. Biol. Chem. 257*, 2528-2536.
Wierenga, R. K., De Maeyer, M. C. H., & Hol, W. G. J. (1985) *Biochemistry 24*, 1346-1357.
Williams, J. W., Morrison, J. F., & Duggleby, R. G. (1979) *Biochemistry 18*, 2567-2573.
Zolg, J. W., & Hänggi, U. J. (1981) *Nucleic Acids Res. 9*, 697-710.

Characterization of Phenylalanine Hydroxylase†

L. M. Bloom and S. J. Benkovic*
Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802
Betty Jean Gaffney
Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218
Received December 20, 1985; Revised Manuscript Received March 20, 1986

ABSTRACT: Iron can be bound to phenylalanine hydroxylase (PAH) in two environments. The assignment of the electron paramagnetic resonance spectrum of PAH to two, overlapping high-spin ferric signals is confirmed by computer simulation. Both environments are shown to be populated in the crude enzyme. Reconstitution of the apoenzyme demonstrated that the two iron environments are not interconvertible. Oxygen consumption during PAH reduction by tetrahydropterin in the absence of phenylalanine but not in its presence explains the different reduction stoichiometries (tetrahydropterin:enzyme) that have been observed.

Phenylalanine hydroxylase (EC 1.14.16.1) catalyzes the hydroxylation of phenylalanine to tyrosine. Although iron is required for the hydroxylation of phenylalanine (Fisher et al., 1972; Gottschall et al., 1982), the role played by iron in the reaction has only begun to be elucidated. We previously demonstrated (Wallick et al., 1984) that iron is reduced during an obligatory reduction by the pterin cofactor in order to activate the enzyme (Marota & Shiman, 1984). Our work on the EPR[1] of PAH revealed that whereas there is one iron per subunit there are at least two environments for iron in the enzyme; the assignments of the EPR signals to two species with different ligand environments were made as described in Wallick et al. (1984). The magnitude of the signals at g_y and g_x of 6.7 and 5.4 correlates with activity; the magnitude of the broad signal centered around the $g_{eff} = 4.3$ region is inversely

† This work was supported by National Science Foundation Grants DMB-831 6425 A02 (S.J.B.) and PCM-8303948 (B.J.G.) and National Institutes of Health Grant GM-28070 (B.J.G.). This work is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy by L.M.B.

[1] Abbreviations: PAH, phenylalanine hydroxylase; SDS, sodium dodecyl sulfate; $6MPH_4$, 6-methyltetrahydropterin; $7,8PH_2$, 7,8-dihydropterin; EPR, electron paramagnetic resonance; Phe, phenylalanine; Tyr, tyrosine; 5-thiapterin, 6-phenyltetrahydro-5-thiapterin; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol.

0006-2960/86/0425-4204$01.50/0