Protein Science (1999), *8*:1342–1349. Cambridge University Press. Printed in the USA.

The crystal structure of a bacterial, bifunctional 5,10 methylene-tetrahydrofolate dehydrogenase/cyclohydrolase

BETTY W. SHEN,[1] DAVID H. DYER,[1] JIE-YU HUANG,[1] LINDA D'ARI,[2] JESSE RABINOWITZ,[2] AND BARRY L. STODDARD[1]

[1]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, A3-023, 1100 Fairview Avenue North, Seattle, Washington 98109
[2]Department of Molecular and Cellular Biology, Barker Hall, University of California, Berkeley, California 94720

(RECEIVED December 30, 1998; ACCEPTED March 4, 1999)

Abstract

The structure of a bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase/cyclohydrolase from *Escherichia coli* has been determined at 2.5 Å resolution in the absence of bound substrates and compared to the NADP-bound structure of the homologous enzyme domains from a trifunctional human synthetase enzyme. Superposition of these structures allows the identification of a highly conserved cluster of basic residues that are appropriately positioned to serve as a binding site for the poly-γ-glutamyl tail of the tetrahydrofolate substrate. Modeling studies and molecular dynamic simulations of bound methylene-tetrahydrofolate and NADP shows that this binding site would allow interaction of the nicotinamide and pterin rings in the dehydrogenase active site. Comparison of these enzymes also indicates differences between their active sites that might allow the development of inhibitors specific to the bacterial target.

Keywords: bifunctional; channeling; cyclohydrolase; dehydrogenase; folate

Single carbon units are carried in a variety of oxidation states by the coenzyme folic acid (Blakeley, 1969; Benkovic & Bullard, 1973; MacKenzie, 1984). A variety of human pathologies, including conditions such as spina bifida and hyperhomocysteinemia, are related to metabolic or dietary folate deficiencies (Bendich & Butterworth, 1996; Jacobsen, 1998). Single carbon units in the oxidation states of formate and formyl-, methenyl-, and methylene-tetrahydrofolate (THF) are interconverted by a series of three enzyme functionalities: 10-formyl THF synthetase, 5,10-methenyl-THF cyclohydrolase, and 5,10-methylene-THF dehydrogenase (Fig. 1). The dehydrogenase and cyclohydrolase activities reversibly catalyze the oxidation of N^5,N^{10}-methylene-THF to N^5,N^{10}-methenyl-THF and the hydrolysis of N^5,N^{10}-methenyl-THF to N^{10}-formyl-THF. Many of these folate-single carbon adducts can be used by the cell as the source of biosynthetic carbon. N^5,N^{10}-methylene-THF is a substrate for thymidylate synthetase, while N^{10}-formyl-THF is a formyl group donor in purine and formylmethionine-tRNA syntheses.

Reprint requests to: Barry L. Stoddard, Division of Basic Sciences, Fred Hutchinson Cancer Research Center, A3-023, 1100 Fairview Avenue, North Seattle, Washington 98109; e-mail: bstoddar@fred.fhcrc.org.

Abbreviations: deh/cyc, 5,10-methylene tetrahydrofolate dehydrogenase/cyclohydrolase; hDC301, human 5,10-methylene tetrahydrofolate dehydrogenase/cyclohydrolase domains; NADP, nicotinamide adenine dinucleotide phosphate; PGF, poly-γ-glutamyl-tetrahydrofolate; syn, 10-formyl tetrahydrofolate synthetase; THF, tetrahydrofolate.

N^5,N^{10}-methylene-THF dehydrogenase and cyclohydrolase activities are generally found on a single, bifunctional enzyme (deh/cyc) in bacteria and in most mitochondria. Many eukaryotes also encode a longer, cytosolic polypeptide chain containing an N-terminal deh/cyc domain (with molecular mass similar to their bacterial and mitochondrial homologues), and a C-terminal domain of mass 60 to 70 kDa exhibiting 10-formyl-THF synthetase (syn) activity (MacKenzie, 1984). The bacterial enzymes and cytosolic enzymes of eukaryotes are predominantly NADP dependent, whereas most mitochondrial enzymes and monofunctional dehydrogenases use NAD. These enzymes bind a single folate substrate (Pelletier & MacKenzie, 1995) and a single NADP coenzyme (Allaire et al., 1998) per subunit.

A number of studies, interpreted together, suggest an efficient reaction mechanism for deh/cyc in which the enzyme binds the folate substrate, interacting with several residues of its polyglutamyl tail, and then allows the pterin ring system sufficient flexibility to move directly between two separate active sites. The methenyl-THF product of the dehydrogenase reaction displays significant kinetic "channeling" to the cyclohydrolase active site, where it is converted to formyl-THF (Cohen & MacKenzie, 1978; Wasserman et al., 1983; Hum & MacKenzie, 1991; Pelletier & MacKenzie, 1995). Several chemical studies indicate that the deh and cyc active sites are in close proximity and might overlap (Schirch, 1978; Drummond et al., 1983; Appling & Rabinowitz, 1985). Kinetic studies using porcine deh/cyc indicate that while the enzyme does not exhibit a marked specificity for polyglutamylated THF

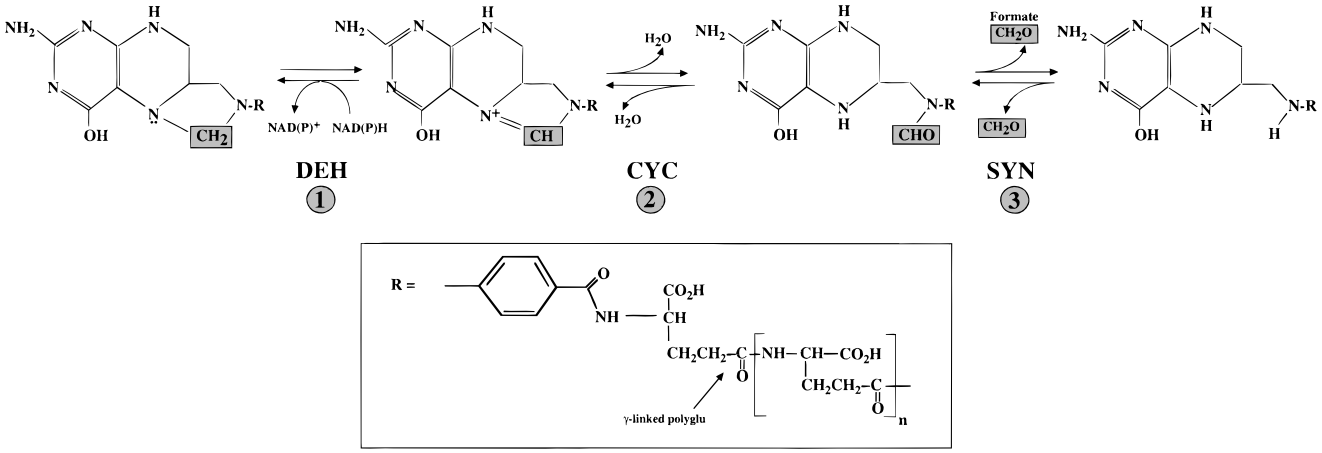

Fig. 1. Conversion of 5,10-methylene-THF, 5,10-methenyl-THF, 10-formyl-THF, and free formate (1) dehydrogenase, (2) cyclohydrolase, and (3) synthetase activities, respectively.

(as estimated by the effect of polyglutamylation on folate K_m), the enzyme velocity increases as the length of the folate polyglutamyl tail is increased, peaking at CH_2-H_4PteGlu$_6$ (Green et al., 1988).

The deh/cyc enzyme from *Escherichia coli* is a dimer composed of two identical 31 kDa subunits of 288 residues each (D'Ari & Rabinowitz, 1991). The structure of the homologous human deh/cyc domains (hDC301) from the trifunctional synthase enzyme was recently determined in complex with NADP (Allaire et al., 1998). On the basis of this structure, the NADP-binding domain of the enzyme was identified, and a group of conserved residues was implicated in dehydrogenase activity. The cyclohydrolase active site was not unambiguously assigned. In the work reported here, the structure of the bifunctional, bacterial deh/cyc apo-enzyme is reported. This structure allows, by direct examination of structural homology, further characterization of residues that might be important for enzyme activity. Analysis of the surface of the enzyme provides a possible model for the binding and dynamics of poly-γ-glutamyl-THF. Comparison of the active site structures for human and bacterial deh/cyc indicates local structural differences that may be useful for future development of microbe-specific enzyme inhibitors.

Results

Topology and overall structure

The bifunctional *E. coli* 5,10-methylene-tetrahydrofolate dehydrogenase/cyclohydrolase (deh/cyc) contains 288 residues and is active as homodimer. The enzyme is globular and has overall dimensions $35 \times 40 \times 55$ Å (Fig. 2). Each monomer can be divided into two distinct domains with typical α/β-folds, connected by two long helices and a loop-helix-loop motif. The smaller N-terminal domain consists of a three-stranded β-sheet flanked on either side by four helices and resembles the first domain of the D-ribose-binding protein (Protein Data Bank (PDB) code 2DRI) as previously described for the human enzyme. The second and larger domain consists of a six-stranded β-sheet and forms a typical dinucleotide-binding domain fold. The three-dimensional structure of the bac-

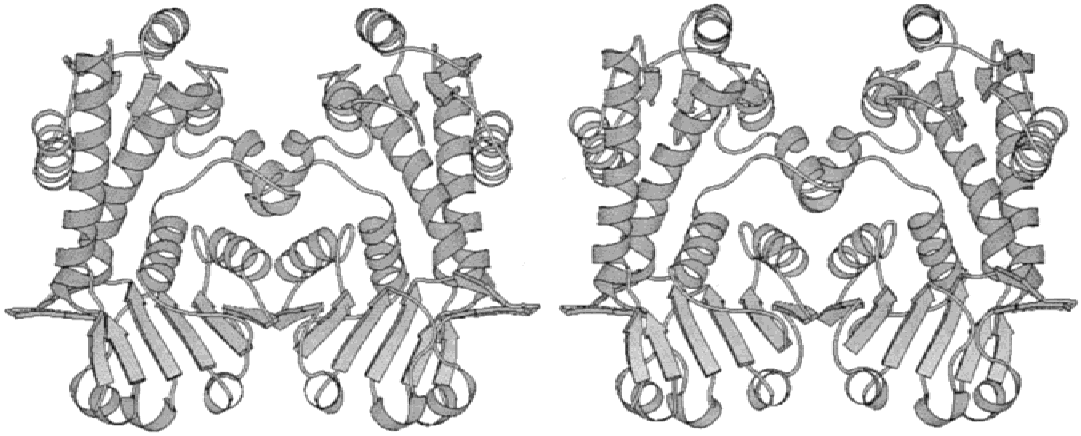

Fig. 2. Stereoview of the bacterial D/C enzyme dimer. The NADP binding site is located in the cleft between enzyme domains as described in the text. Figures 2–5 prepared with MOLSCRIPT (Kraulis, 1991).

terial apo-enzyme is very similar to that of the 301-residue human deh/cyc domain (hDC301), from the trifunctional human THF synthetase (Allaire et al., 1998), except that the extended β-hairpin (residues 239–253 in hDC301) is shortened by six residues. The angle and separation between enzyme domains display intermediate values compared to the two independent subunits of hDC301, as described in the next section.

Similar to many known folate-utilizing enzymes, deh/cyc from *E. coli* forms an active dimer. The current crystal form contains one monomer per asymmetric unit, and the two identical subunits of the enzyme are related to each other by a crystallographic dyad axis. Approximately 1,500 Å^2 accessible surface per monomer is buried upon dimerization, accounting for ~11% of the total surface of a monomer. The majority of the dimer interface contacts (~960 Å^2) occur between helix 7 and β-strand *f* in the large NADP-binding domain of the two subunits. Upon dimerization, the central β-sheets in the large domain of the two enzyme subunits associate through the formation of an antiparallel β-strand pair, thereby forming a continuous 12-strand β-sheet. Cys132 of helix 5 is positioned for the formation of a disulfide linkage with its symmetry mate, Cys132′ of helix 5′. The stability of the dimer is ensured by a number of salt bridges and hydrogen bonds, as well as the covalent disulfide linkage between Cys132 and 132′. Most notably, Glu112 of one subunit forms a H-bond with Tyr126′ and a salt bridge with Arg130′ of the other subunit. In the human enzyme, this disulfide bridge is not present, and a similar dimer interface is stabilized by a series of polar interactions between side chains.

Structural comparison with human deh/cyc enzyme domains

The topology of *E. coli* deh/cyc strongly resembles that of human DC301. This is not surprising, considering the high sequence identity between the two enzymes (>44% identity). However, it is interesting that the superposition between the Cα trace of the bacterial enzyme and either of the hDC301 subunits gives smaller root-mean-square deviation (RMSD) values than when the two noncrystallographic symmetry (NCS) related monomers of the hDC301 dimer are superimposed on one another (Table 1). For all cases, superposition of individual enzyme domains provides closer agreement than the superposition of entire enzyme subunits. This analysis clearly indicates that the peptide linkers between enzyme domains are hinge points that facilitate global dynamic flexibility. The large RMSD (1.403 Å) between the two NCS-related subunits of the NADP-bound hDC301 enzyme is attributed to differences in local crystal contacts and the visible flexibility of the enzyme even when complexed with NADP. In comparison, the two domains of the bacterial enzyme display a separation and intervening cleft dimensions midway between those observed for the A and B subunits of the hDC301 (Fig. 3). This is consistent with the assumption that the molecule is flexible and suggests that the active site cleft of the dehydrogenase and cyclohydrolase may exhibit similar thermal motion and flexibility in the presence and absence of bound NADP cofactor. A question that remains to be answered is whether the binding of folate, or formation of a ternary complex, will induce a more substantial domain closure as part of the enzyme's catalytic mechanism. Such a result has been previously described for thymidylate synthase (Matthews et al., 1990; Montfort et al., 1990; Kamb et al., 1992).

Table 1. *Structural superpositions and RMS differences between subunits and domains of E. coli deh/cyc and D/C bifunctional domain of human THF synthetase (D/C)*

Superposition	Number of equivalent Cα positions	RMSD (Å)
E. coli deh/cyc to hDC301-A	280 (98.2%)[a]	1.19
Small (N-terminal) domain	86 (100%)	0.86
Large (C-terminal) domain	105	0.75
E. coli deh/cyc to hDC301-B	278	1.27
Small (N-terminal) domain	86	0.92
Large (C-terminal) domain	101	0.71
hDC-A to hDC301-B	285 (97%)	1.40
Small (N-terminal) domain	90	0.28
Large (C-terminal) domain	110	0.271

[a]Numbers in parentheses denote percentage with respect to the number of Cα positions in *E. coli*. Deh/cyc, or in the case of superposition between the A and B subunits of hDC301, refers to the number in subunit A.

In the structure of hDC301-NADP complex, the NADP molecule was shown to bind to the C-terminal ledge of the central β-sheet in the large domain with the diphosphate linkage straddling the ridge between β-strands. Since the large domain of the *E. coli* deh/cyc apo-enzyme superimposes closely over that of the hDC301-NADP complex (RMSD < 1.0 Å), it is reasonable to assume that the bacterial deh/cyc would bind its NADP cofactor at the same location with minimal conformational rearrangement. Alignment of the sequence and structure of the two enzymes showed that 13 of the 20 residues within 4 Å of the bound NADP in the human DC301-NADP complex are conserved in the bacterial enzyme. To compare the structural differences in the vicinity of the NADP cofactor, the coordinates of NADP in the human hDC301 complex were transformed into the bacterial coordinates, using as a transformation matrix the rotation and translation operators derived from superposition of the large domains in the two structures. The residues of *E. coli* deh/cyc within 4 Å of the transformed NADP cofactor, before and after molecular dynamics minimization of the modeled NADP, were compared with corresponding residues from the structure of the human enzyme complex. Inspection of the environment around NADP in the two structures showed that there are close similarities between the binding sites for the NADP molecules in the two enzymes. All but two residues (Arg173 and Ser197) involved in the binding of NADP in the hDC301 complex are conserved both sequentially and structurally. Arg173 in hDC301, which makes three hydrogen bonds with the adenosine-2-phosphate of NADP, is changed to Ala167 in the bacterial enzyme, while Ser197, which also makes a hydrogen bond with the 2′-phosphate, is changed to Arg191. Because the Arg173 side chain of hDC301 and the Arg191 side chain of bacterial deh/cyc are positioned similarly, it is probable that they make similar contacts to the 2′-phosphate group in the NADP-enzyme complex. In the simulated bacterial deh/cyc-NADP complex, Asn233 is within 4 Å of the hydroxyl group of the nicotinamide ribose but not its counterpart (Asn239) in hDC301.

Putative methylene-tetrahydrofolate binding site

Kinetic studies suggest that interaction of the poly-γ-glutamyl tail of the methylene-tetrahydrofolate molecule with the enzyme provides binding energy and affinity for the substrate, while allowing

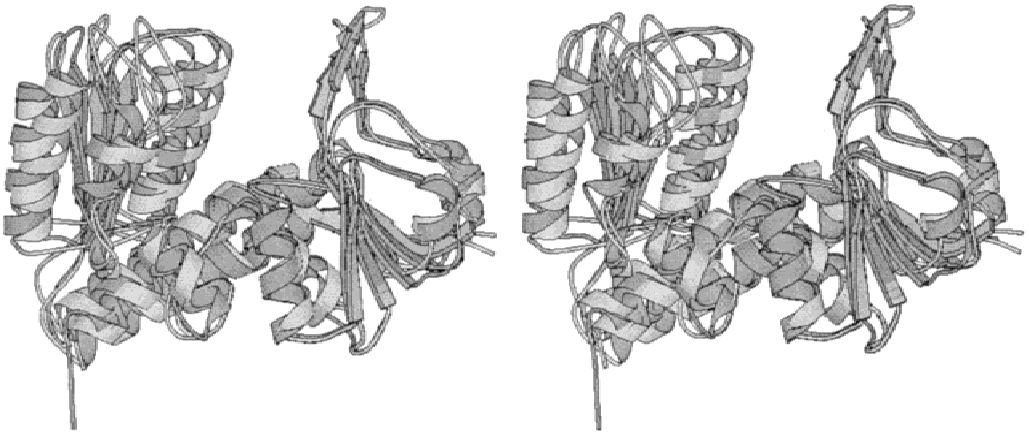

Fig. 3. Superposition of the two crystallographically independent subunits of the human D/C dimer (light subunits) and the single independent subunit of the bacterial enzyme (darker subunit). The enzyme monomers were aligned on their small domains. The point of maximum dynamic flexibility is localized to the interdomain cleft.

dynamic motion and kinetic channeling of the reactive methenylpterin ring system between dehydrogenase and cyclohydrolase active sites (Cohen & MacKenzie, 1978; Wasserman et al., 1983; Green et al., 1988; Hum & MacKenzie, 1991; Pelletier & MacKenzie, 1995). Inspection of the surface of the *E. coli* dph/cyc revealed that there are two significant regions of positive-charged residues, one at each end of the cleft between the two α/β domains of deh/cyc. To model the binding of poly-γ-glutamyl-methylene-tetrahydrofolate in the active site cleft, the coordinates of the *E. coli* dph/cyc with the modeled NADP was used as a receptor for the calculation of a molecular surface, using the DMS program of MIDAS (Ferrin et al., 1988). From that molecular surface, a cluster of spheres was generated and used for docking of the folate substrate into a box representing the active site cleft, using DOCK 4.0 (Meng et al., 1992; Oshiro & Kuntz, 1995).

The 10 most stable docked solutions for the poly-γ-glutamyl-5′-10′-methylene-tetrahydrofolate, based on contact score only, were generated and each inspected in O. As expected, the folate substrate is able to fit in the active site cleft in two orientations, with the poly-γ-Glu tail pointing toward or away from the monomer-monomer interface. A cluster of several solutions in this list all have the pterin ring pointing toward the dehydrogenase active site and the poly-γ-glutamyl tail pointed toward the exterior of the enzyme, where it interacts with a group of conserved basic residues. Molecular dynamics simulations of this docked ternary complex, using initially rigid-body minimization followed by dynamic simulation with temperature coupling, rapidly converged with minimal rearrangement and gave a distance between the methylene carbon and the C4 carbon of the nicotinamide of ~4 Å. Such a model appears attractive both on the basis of the conservation of residues proposed to bind the glutamyl substrate tail and on the basis of the modeled proximity between the nicotinamide and pterin rings of the cofactors. It appears that up to four glutamyl residues can be accommodated along this region of the enzyme surface, with reasonable electrostatic contacts to conserved basic residues. Figures 4 and 5 show details, respectively, of the proposed binding site of the THF poly-γ-glutamyl-tail and of the proposed ternary complex with PGF and NADP.

Discussion

The structure reported here allows direct comparison of a free-standing bifunctional enzyme (deh/cyc) from a prokaryotic source with the homologous enzyme domains liberated from a human trifunctional synthase enzyme. Previously published results have provided mixed indications as to whether the C-terminal 10-formyl synthetase (syn) domain is necessary for full or even partial activity of the deh/cyc domains in the trifunctional synthase enzymes. Studies of the deh/cyc domains from the trifunctional yeast synthase enzyme indicated that the presence of the syn domain was an absolute requirement for dehydrogenase activity, and that chimaeric fusions of deh/cyc with other C-terminal gene products displayed restored dehydrogenase activity (Song & Rabinowitz, 1995). A second study, however, compared the kinetic properties of the deh/cyc domains from the human trifunctional synthase with the same domains from the human mitochondrial NAD-dependent bifunctional enzyme and the NADP-dependent bifunctional enzyme from *Photobacterium phosphoreum.* These studies indicated that while the ratios of kcat(cyc)/kcat(deh) vary widely, the deh activity is measurable and significant for each enzyme and that all channel methenyl-THF to the cyc active site with equal efficiency (Pawelek & MacKenzie, 1998). There is very little difference between the structures of the *E. coli* deh/cyc and the expressed human domains from the trifunctional synthase enzyme, supporting the concept that the deh/cyc domains are fully separable and functionally independent from the synthetase domain. In this structure, the C-terminal end of the bacterial enzyme is located on the opposite side of the enzyme monomer from the dehydrogenase active site; however, it is considerably closer to the active site within the opposite subunit. It is straightforward to envision that a large enzyme domain (syn is over 500 residues in size) would interact with structural elements of the deh/cyc active site. Additionally, the residues of the final α-helix leading to the C-terminus of the protein are very well conserved across the known synthase and deh/cyc THF enzymes and may play an important role in protein folding and dynamics; truncation of the deh/cyc domains

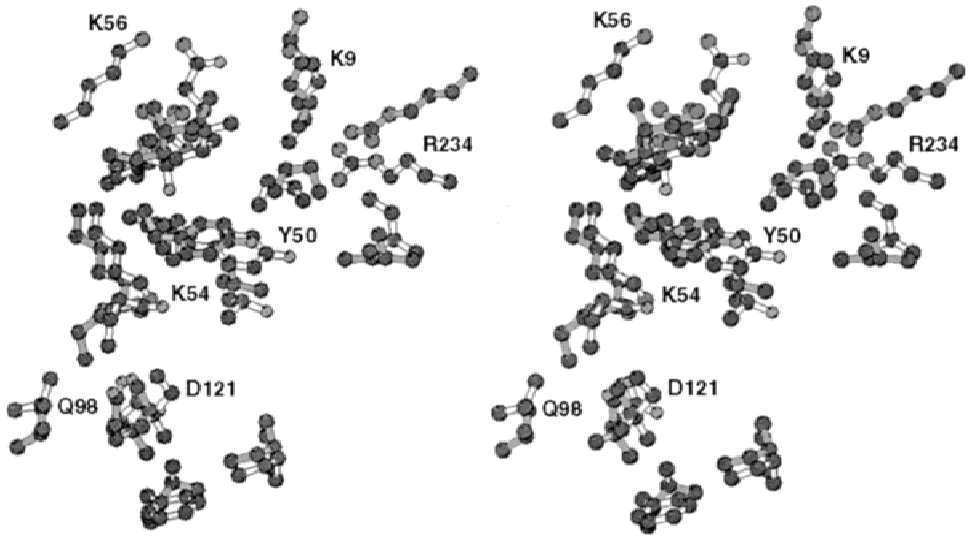

Fig. 4. Stereoview of the proposed conserved binding site for the tetrahydrofolate γ-glutamyl tail. The side chains of the bacterial basic residues are shown with light bonds and the human enzyme side chains with darker bonds.

from trifunctional synthase enzymes may dramatically affect the structure and function of the resulting protein construct.

The dinucleotides NAD and NADP are broadly distributed cofactors for oxidation-reduction processes. They serve as electron carriers through hydride transfer for a number of important metabolic pathways, including glycolysis, tricarboxylic acid cycle, fatty acid synthesis, and sterol synthesis. NADP and NAD mediated electron transfer also play important roles in the reversible interconversion of single carbon units when carried by folate cofactors. The structural basis of dinucleotide mediated hydride transfer has been most extensively studied in the NAD dependent dehydrogenases and aldolases, and in the NADP dependent dihydrofolate reductase. A clear picture emerges from the crystal structure of a large number of nucleotide dependent enzymes. It shows that in spite of low sequence homology among some of the studied enzymes, the structural features that are responsible for the binding of the nucleotide remain topologically the same; a six-stranded β-sheet in a typical α/β-fold for a dinucleotide with each mononucleotide binding to half of the extended α/β fold and the diphosphate linkage wedged between the boundary of the two three-stranded half domains (Ohlsson et al., 1974).

It has been shown in the structure of many dehydrogenase enzymes, and recently in hDC301, that the dinucleotide is anchored to the enzyme proper through multiple interactions with (1) the 2′-phosphate and 3′ hydroxyl oxygens of the adenosine ribose, (2) the oxygen of the diphosphate backbone, and (3) the amide group of the nicotine base. All residues in hDC301, except Arg173 and Ser174, involved in the binding of NADP are conserved. In hDC301, the 3′-phosphate group is anchored to the enzyme through multiple hydrogen bonds to the guanidinium nitrogens NE and NH2 of Arg173. In the bacterial enzyme, this residue is replaced by Ala167. However, a second amino acid substitution, Ser197 to Arg191, restores the *E. coli* enzyme's potential to bind the 2′-phosphate group. In the structure of hDC301 complexed with NADP, the

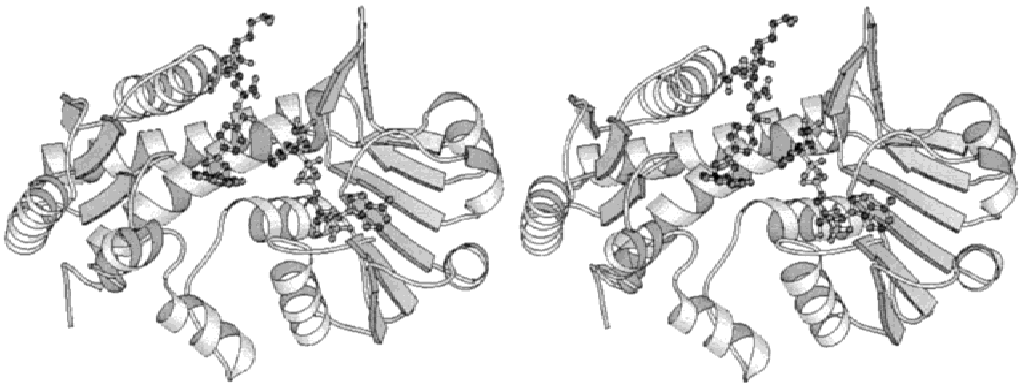

Fig. 5. Stereo view of the modeled ternary complex of D/C, NADP, and methylene-THF. The NADP cofactor is the ligand to the lower right, associated with the β-sheet of the Rossman fold domain, while the PGF substrate is the ligand to the upper right. The glutamyl tail, modeled here with four glutamate residues, points directly up and in the model described in the text makes contacts with the residues shown in Figure 4.

pro-S face of the nicotinamide ring is buried against the protein surface, while the pro-R face is exposed to the solvent, in position to accept a transferred hydride from the pterin-bound methylene carbon. This observation is consistent with the observed stereochemistry of the reaction (Green et al., 1986).

Using the modeled deh/cyc-NADP as receptor, we have docked and energy-minimized a model of poly-γ-glutamyl-THF in the active site cleft to form a putative enzymatic ternary complex. In this model, the poly-γ-glutamyl-THF (PGF) is bound to each enzyme subunit with its pterin head group close to the nicotinamide ring of NADP and oriented toward the monomer-monomer interface. The residues within 5 Å of the modeled PGF molecule are highly conserved (80% identity vs. 44% identity overall). In this model, the pterin ring rests against a large hydrophobic wall lined by three residues (Leu99, Phe123, and Ile170) and its 4-hydroxy group stabilized by D121 and T267. The π-electron cloud of the p-amino-benzoyl group contacts Tyr50 and Val268. Additionally, the negatively charged carboxylates of the poly-γ-glutamyl tail are within H-bond distance to five conserved basic residues (Lys9, Arg16, Lys 54, Lys56, and Arg234).

The general location of the dehydrogenase active site in this bifunctional enzyme appears to be well defined on the basis of the structures of bacterial and human deh/cyc, the bound position of NADP, and the conservation of residues near the nicotinamide ring. However, the position of the enzyme atoms that comprise the cyclohydrolase active site are less clear. Kinetic data indicate that the methenyl-THF pterin ring system is channeled between these active sites. Examination of the structure of hDC301 (Allaire et al., 1998) has led to the hypothesis that a pair of conserved peptide signatures ($Tyr_{52}X_3Lys_{56}$ and a Ser_{49}-Gln_{100}-Pro_{102} triplet) are involved in pterin-binding and the hydration reaction. In one possible reaction mechanism, the conserved tyrosine and/or lysine residues could participate in proton transfer to the N^5-nitrogen of the pterin ring during the reversible interconversion of methenyl-THF and formyl-THF. Assuming that all the residues of the γ-glutamyl tail are fixed in a relatively stable bound conformation, that the amide linkage to the PABA ring displays limited rotational freedom, and that the remaining rotatable groups of the THF cofactor allow an unhindered 180° rotation of the pterin ring, then the bound carbon substrate is allowed a range of motion of up to 10 Å from its location in the dehydrogenase active site. Within this distance lies a pair of conserved residues (Lys54 and Gln98 in the bacterial enzyme) that correspond to Lys56 and Gln100 of the hDC301 domains. These residues are complexed with a well-ordered water molecule, as observed in the structure of the human domains. The structure therefore appears to be in good agreement with the hypothesis that the YXXXK and S-Q-P signatures might constitute an important part of the cyclohydrolase active site. It is known that the yeast monofunctional dehydrogenase (West et al., 1993), which lacks the cyclohydrolase activity, has a threonine residue substituted for Lys54 or Lys56 in the bacterial and human enzymes, respectively.

To bring the methenyl carbon into proximity with these groups would require a rotation of approximately 90° about the N^{10}-C^{15} bond of the THF cofactor. Such a movement would appear to be possible in the bound model described here. Interestingly, modeling of the types of bond rotations that might position the methenyl carbon in alternate locations often appear to induce a steric clash with the bound NADP cofactor. This observation might be consistent with the observation that bound nicotinamide mononucleotide partially inhibits the cyclohydrolase activity of the enzyme. An additional possibility for the type of dynamic motion involved in substrate channeling would be a closure of the N- and C-terminal domains during turnover, which would bring the proposed cyclohydrolase catalytic residues closer to the dehydrogenase active site. The crystal structures of the bacterial deh/cyc and the human DC301 domains support such protein flexibility.

Materials and methods

The enzyme was purified and crystallized as previously described (Cheung et al., 1997). The crystals were transferred into 25% glycerol plus 13% PEG 10000, 100 mM Bis Tris (pH7.5) for flash-cooling in liquid nitrogen by sequential transfer in 2% increments, allowing the crystals to equilibrate for 3 min at each step. The crystals belong to a body centered orthorhombic space group, with unit cell dimensions: a = 64.52 Å, b = 84.95 Å, and c = 146.15 Å. Based on systematic absences in the resulting data set, the space group was assigned as I222. The related space group $I2_12_12_1$ was excluded by Patterson analysis of heavy metal derivatives and by examination of maps after initial multiple isomorphous replacement (MIR) phasing. The unit cell volume is 8.0×10^5 Å^3, and the crystal unit cell and diffraction is consistent with an asymmetric unit consisting of the enzyme monomer, and a specific volume of 3.2 Å^3/Da.

A native data set was collected to 2.5 Å resolution from a single crystal at beamline X12-C at the National Synchrotron Light Source at Brookhaven National Laboratory, using a MAR-research imaging plate detector and an incident X-ray wavelength at 1.1 Å. Data were reduced using the DENZO/SCALEPACK crystallographic data reduction package (Otwinowski & Minor, 1997). Data statistics are shown in Table 2. Three isomorphous heavy atom derivative data sets (one di-platinum and two gold) were used for MIR phase calculations. The crystals were soaked in 1 mM solutions of gold cyanide ($Au(CN)_2$) or PIP (di-μ-iodo-bis(ethylenediamine)-diplatinum-nitrate) for 1 h prior to data collection. The Pt and one of the Au derivative data sets were collected to 2.8 Å resolution at beamline X12-C as described for the native data set. A second gold data set was collected to 3 Å resolution on an RAXIS-II area detector using a rotating anode X-ray generator. The statistical heavy atom refinement and phasing program SHARP (LaFortelle et al., 1997) was used to calculate and improve the experimental phases to 3 Å resolution. The phases were extended to 2.5 Å resolution using solvent flattening and real-space density constraints in program DM (CCP4, 1994; Cowtan, 1994), Table 2 contains relevant phasing statistics. The map was of excellent quality showing good connectivity of all secondary structural features and directionality of the peptide bonds in the helices.

Crystallographic model building and refinement

The electron density maps were clear and unambiguous for most regions of the secondary structure, allowing an initial trace to be built for approximately 90% of the protein backbone using the BONES option of the model building program O (Jones et al., 1991; Jones & Kjeldgaard, 1998). The coordinates for the homologous bifunctional domain from the human trifunctional synthetase enzyme (Allaire et al., 1998) were superpositioned onto the protein backbone model of the bacterial enzyme, and the positions of easily recognizable residues [YITPVPGG] in a highly conserved peptide (255–269) were used to register the protein sequence onto the model. The side chains were built into the density,

Table 2. *Data, phasing, and refinement statistics*

Diffraction data	Native	Au1	Au2	PIP
Wavelength (Å)	1.100	1.542	0.903	1.542
Resolution	2.5	2.8	2.9	2.8
Completeness (%)	98.6 (99.8)	90.4 (59.9)	98.0 (99.0)	98.4 (91.7)
R_{merge}	3.3 (13.0)	5.7 (21.8)	7.0 (39.2)	4.1 (14.6)
Phasing statistics				
Number of sites		3	2	2
Phasing power		1.14	1.66	1.03
Overall figure of merit		0.40 (acentric)/0.43 (centric)		
Refinement				
R_{cryst}/R_{free}	23.2/29.9			
Protein/solvent atoms	2,759/35			
Ramachandran distribution (%)	90.2 core, 9.8 allowed, 0 generous, 0 disallowed			
RMS (bond lengths, Å)	0.006			
RMS (bond angles, °)	1.27			
Average B (Å^2) (protein)	38.8			

allowing space for weak and missing density of exposed glycine residues.

The model was subjected to iterations of conventional positional refinement using X-PLOR 3.8 (Brünger, 1992) and manual model rebuilding followed by torsion angle dynamics and *B*-factor refinements. The free *R*-factor (Brünger, 1993) was used to monitor all stages of the refinement. Restrained group *B*-factors were refined once the *R*-factors had dropped to reasonable values. For each residue, a single overall *B*-factor for its main-chain atoms and a second separate overall *B*-factor for its side-chain atoms were refined. The statistics for the refined model are given in Table 2. The working *R*-factor is 22.7% using data between 50 and 2.56 Å resolution. The free *R*-factor (10% of the data, 1,296 reflections) is 29.8% for the final model using reflections with $I/\sigma(I) > 2$. The stereochemical quality of the model was examined throughout the refinement using PROCHECK (Laskowski et al., 1993) and OOPS (Jones et al., 1991; Jones & Kjeldgaard, 1998). The quality of the final model was also confirmed by examining side-chain environments using a profile-fit analysis and found to be within acceptable ranges.

Acknowledgments

We thank William Scott, Roland Strong, and Kam Zhang for advice and help with X-ray crystallography, and Edwin Cheung for performing initial protein purifications. Robert MacKenzie and Miroslaw Cygler graciously provided protein coordinates for the homologous human enzyme domains that greatly accelerated our model building and refinement after initial phasing and chain trace. We also thank the beamline staff at X12-C (Robert Sweet, John Skinner, and Sal Sclafani) at the Brookhaven National Synchrotron Light Source. B.L.S. was funded for this project by the NIH (GM51224). Atomic coordinates have been deposited in the Brookhaven PDB.

References

Allaire M, Li Y, MacKenzie RE, Cygler M. 1998. The 3-D structure of a folate-dependent dehydrogenase/cyclohydrolase bifunctional enzyme at 1.5 Å resolution. *Structure 6*:173–182.

Appling DR, Rabinowitz JC. 1985. Evidence for overlapping active sites in a multifunctional enzyme: Immunochemical and chemical modification studies on C1-tetrahydrofolate synthase from *Saccharomyces cerevisiae*. *Biochemistry 24*:3540–3547.

Bendich A, Butterworth CEJ. 1996. Folic acid and the prevention of birth defects. *Ann Rev Nutr 16*:73–97.

Benkovic SJ, Bullard WP. 1973. On the mechanisms of action of folic acid cofactors. *Prog Bior Chem 2*:133–175.

Blakeley R. 1969. Biochemistry of folic acid and related pteridines. In: Bloch K, ed. *Accounts of chemical research*. New York: Wiley and Sons. pp 191–237.

Brünger A. 1992. *X-PLOR version 3.1: A system for X-ray crystallography and NMR*. New Haven Connecticut: Yale University Press.

Brünger A. 1993. Assessment of phase accuracy by cross validation: The free *R* value. Methods and applications. *Acta Cryst D49*:24–36.

Cheung E, D'Ari L, Rabinowitz JC, Dyer DH, Huang J-Y, Stoddard BL. 1997. Purification, crystallization, and preliminary X-ray studies of a bifunctional 5,10-methenyl/methylene tetrahydrofolate cyclohydrolase/dehydrogenase from *Escherichia coli*. *Proteins Struct Funct Genet 27*:322–324.

Cohen L, MacKenzie RE. 1978. Methylenetetrahydrofolate-dehydrogenase/methenyltetrahydro-folate cyclohydrolase/formyltetrahydrofolate synthetase from porcine liver. Interaction between the dehydrogenase and cyclohydrolase activities of the multifunctional enzyme. *Biochim Biophys Acta 522*:311–317.

Collaborative Computational Project 4. 1994. The CCP4 Suite: Programs for protein crystallography. *Acta Cryst D 50*:760–763.

Cowtan K. 1994. An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newslett Protein Crystallogr 31*:34–38.

D'Ari L, Rabinowitz J. 1991. Purification, characterization, clining, and amino acid sequence of the bifunctional enzyme 5,10-methylenetetrahydrofolate dehydrogenase/5,10-methenyltetrahydrofolate cyclohydrolase from *E. coli*. *J Biol Chem 266*:23953–23958.

Drummond D, Smith S, MacKenzie RE. 1983. Methenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase/formyltetrahydrofolate synthetase from porcine liver: Evidence to support a common dehydrogenase/cyclohydrolase site. *Can J Biochem Cell Biol 61*:1166–1171.

Ferrin TE, Huang CC, Jarvis LE, Langridge R. 1988. The MIDAS display system. *J Mol Graphics 6*:13–27.

Green JM, Mackenzie RE, Matthews RG. 1988. Substrate flux through methylenetetrahydrofolate dehydrogenase: Predicted effects of the concentration of methylenetetrahydrofolate on its partitioning into pathways leading to nucleotide biosynthesis or methionine regeneration. *Biochemistry 27*:8014–8022.

Green JM, Matthews RG, Mackenzie RE. 1986. Stereochemistry of hydride transfer to NADP+ by methylenetetrahydrofolate dehydrogenase from pig liver. In: Cooper BA, Whitehead VM, eds. *Chemistry and biology of pteridines*. Berlin: Walter de Gruyter and Co. pp 901–904.

Hum DW, MacKenzie RE. 1991. Expression of active domains of a human folate-dependent trifunctional enzyme in *E. coli*. *Protein Eng 4*:493–500.

Jacobsen DW. 1998. Homocysteine and vitamins in cardiovascular disease. *Clinical Chem 44*:1833–1843.

Jones TA, Kjeldgaard MO. 1998. Electron density map interpretation. *Methods Enzymol 277*:173–208.

Jones TA, Zou J-Y, Cowtan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst A 47*:110–119.

Kamb A, Finer-Moore JS, Stroud RM. 1992. Cofactor triggers the conformational change in thymidylate synthase: Implications for an ordered binding mechanism. *Biochemistry 31*:12876–12884.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst 24*:946–950.

LaFortelle E, Irwin JJ, Bricogne G. 1997. SHARP: A maximum-likelihood heavy-atom parameter refinement and phasing program for the MIR and MAD methods. In: Bourne P, Watenpaugh K, eds. *Crystallographic computing*. Oxford: Oxford Science Publications. pp 100–130.

Laskowski RJ, Macarthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:X83–X90.

MacKenzie RE. 1984. Biogenesis and interconversion of substituted tetrahydrofolates. In: Blakeley R, Benkovic S, eds. *Folates and pterins: Chemistry and biochemistry of folates*. New York: Wiley and Sons. pp 255–306.

Matthews DA, Villafranca JE, Janson CA, Smith WW, Welsh K, Freer S. 1990. Stereochemical mechanism of action for thymidylate synthase based on the X-ray structure of the covalent inhibitory ternary complex with 5-fluoro-2′-deoxyuridylate and 5,10-methylenetetrahydrofolate. *J Mol Biol 214*:937–948.

Meng EC, Shoichet B, Kuntz ID. 1992. Automated docking with grid-based energy evaluation. *J Comp Chem 13*:505–524.

Montfort WR, Perry KM, Fauman EB, Finer-Moore JS, Maley GF, Hardy L, Maley F, Stroud RM. 1990. Structure, multiple site binding, and segmental accommodation in thymidylate synthase on binding dUMP and an antifolate. *Biochemistry 29*:6964–6977.

Ohlsson I, Nordstrom B, Branden C-I. 1974. Structural and functional similarities within the coenzyme binding domains of dehydrogenases. *J Mol Biol 89*:339–354.

Oshiro CM, Kuntz ID. 1995. Flexible ligand docking using a genetic algorithm. *J Comp-Aided Mol Design 9*:113–130.

Otwinowski Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol 276*:307–326.

Pawelek PD, MacKenzie RE. 1998. Methenyltetrahydrofolate cyclohydrolase is rate limiting for the enzymatic conversion of 10-formyltetrahydrofolate to 5,10-methylenetetrahydrofolate in bifunctional dehydrogenase-cyclohydrolase enzymes. *Biochemistry 37*:1109–1115.

Pelletier JN, MacKenzie RE. 1995. Binding and interconversion of tetrahydrofolates at a single site in the bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase. *Biochemistry 34*:12673–12680.

Schirch L. 1978. Formyl-methenyl-methylenetetrahydrofolate synthetase from rabbit liver (combined). Evidence for a single site in the conversion of 5,10-methylenetetrahydrofolate to 10-formyltetrahydrofolate. *Arch Biochem Biophys 189*:283–290.

Song JM, Rabinowitz JC. 1995. The N-terminal, dehydrogenase/cyclohydrolase domain of yeast cytoplasmic trifunctional C1-synthase requires the C-terminal, synthetase domain for the catalytic activity in vitro. *FEBS Lett 376*:229–232.

Wasserman GF, Benkovic PA, Young M, Benkovic SJ. 1983. Kinetic relationships between the various activities of the formyl/methenyl/methylenetetrahydrofolate synthetase. *Biochemistry 22*:1005–1013.

West MG, Barlowe CK, Appling DR. 1993. Cloning and characterization of the *Saccharomyces cerevisiae* gene encoding NAD-dependent 5,10-methylenetetrahydrofolate dehydrogenase. *J Biol Chem 268*:153–160.