Stryer, L., Hurley, J. B., & Fung, B. K.-K. (1981) *Curr. Top. Membr. Transp. 15*, 93-108.
Tolkovsky, A. M., & Levitzki, A. (1978) *Biochemistry 17*, 3795-3810.
Tolkovsky, A. M., Braun, S., & Levitzki, A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 213-217.
Wessling-Resnick, M., & Johnson, G. L. (1987a) *J. Biol. Chem. 262*, 3697-3705.
Wessling-Resnick, M., & Johnson, G. L. (1987b) *J. Biol. Chem. 262*, 12444-12447.
Yatani, A., Codina, J., Brown, A. M., & Birnbaumer, L. (1987) *Science 235*, 207-211.

Structure, Multiple Site Binding, and Segmental Accommodation in Thymidylate Synthase on Binding dUMP and an Anti-Folate[†,‡]

William R. Montfort, Kathy M. Perry, Eric B. Fauman, Janet S. Finer-Moore, Gladys F. Maley,[§] Larry Hardy,[∥] Frank Maley,[§] and Robert M. Stroud*

Department of Biochemistry and Biophysics, S-960, University of California in San Francisco, San Francisco, California 94143-0448

Received November 20, 1989; Revised Manuscript Received March 22, 1990

ABSTRACT: The structure of *Escherichia coli* thymidylate synthase (TS) complexed with the substrate dUMP and an analogue of the cofactor methylenetetrahydrofolate was solved by multiple isomorphous replacement and refined at 1.97-Å resolution to a residual of 18% for all data (16% for data $>2\sigma$) for a highly constrained structure. All residues in the structure are clearly resolved and give a very high confidence in total correctness of the structure. The ternary complex directly suggests how methylation of dUMP takes place. C-6 of dUMP is covalently bound to γS of Cys-198(146) during catalysis, and the reactants are surrounded by specific hydrogen bonds and hydrophobic interactions from conserved residues. Comparison with the independently solved structure of unliganded TS reveals a large conformation change in the enzyme, which closes down to sequester the reactants and several highly ordered water molecules within a cavernous active center, away from bulk solvent. A second binding site for the quinazoline ring of the cofactor analogue was discovered by withholding addition of reducing agent during crystal storage. The chemical change in the protein is slight, and from difference density maps modification of sulfhydryls is not directly responsible for blockade of the primary site. The site, only partially overlapping with the primary site, is also surrounded by conserved residues and thus may play a functional role. The ligand-induced conformational change is not a domain shift but involves the segmental accommodation of several helices, β-strands, and loops that move as units against the β-sheet interface between monomers.

Thymidylate synthase (TS)[1] (EC 2.1.1.45) catalyzes the reductive methylation of deoxyuridine monophosphate (dUMP) by 5,10-methylenetetrahydrofolate (CH_2-H_4folate) to form thymidine monophosphate (dTMP) and dihydrofolate (H_2folate) (Figure 1) [reviewed by Pogolotti and Santi (1977)]. Conformation changes are associated with initial binding of dUMP (Galivan et al., 1975; Leary et al., 1975), with even larger changes accompanying binding of cofactor, evaluated both by spectroscopic (Sharma & Kisliuk, 1973; Danenberg et al., 1974; Santi et al., 1974; Donato et al., 1976) and by hydrodynamic methods. Gel filtration and sedimentation analyses show a large apparent decrease of 3.5% in Stokes radius of *Lactobacillus casei* and human protein in response to ternary complex formation (Lockshin & Danenberg, 1980). The structure of unliganded[2] TS from *L. casei* was solved first (Hardy et al., 1987). However, to describe these large structural changes and to help delineate their role in catalysis, we sought to determine the structure of a ternary complex of a close analogue of the normal TS·dUMP·CH_2-H_4folate complex that could be compared directly with the structure of the unliganded protein of the same species. We first crystallized such a pair of structures for this comparison with TS from *Escherichia coli*. The structure of the ternary complex reported here was solved and refined independently from the structure of unliganded *E. coli* TS (Perry et al., 1990), thus eliminating any bias in interpretation of conformational differences.

As TS is the terminal step in the sole de novo synthetic pathway to dTMP, it has, since the discovery of 5-fluorouracil

[†] Supported by National Institutes of Health Grants RO1-CA-41323 to J.S.F.-M. and R.M.S., GM24485 to R.M.S., and CA44355 to F.M., by NSF Grant DM386-16273 to G.F.M., and by postdoctoral fellowships from the American Chemical Society (S-49-87 to W.R.M.) and from NIH (GM-11846 to K.M.P.). E.B.F. is a Howard Hughes Medical Institute Predoctoral Fellow.

[‡] Crystallographic coordinates have been submitted to the Brookhaven Protein Data Bank under 1TSC,1TMS.

* Author to whom correspondence should be addressed.

[§] Present address: Wadsworth Center for Laboratories and Research, New York State Department of Health, Empire State Plaza, Albany, NY 12201-0509.

[∥] Present address: Department of Pharmacology, University of Massachusetts Medical School, Worcester, MA 01655.

[1] Abbreviations: TS, thymidylate synthase; dUMP, 2′-deoxyuridine 5′-monophosphate; CH_2-H_4folate, 5,10-methylenetetrahydrofolate; dTMP, thymidine 5′-monophosphate; H_2folate, dihydrofolate; CB3717, 10-propargyl-5,8-dideazafolate; MIR, multiple isomorphic replacement; R_{crys}, crystallographic R factor.

[2] Unliganded structures refer to structures with inorganic phosphate bound at the site for phosphate in dUMP.

0006-2960/90/0429-6964$02.50/0

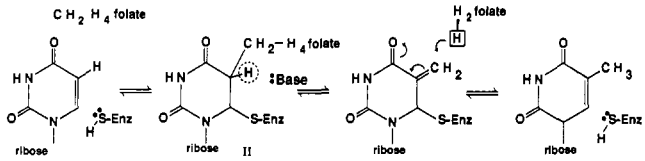

FIGURE 1: One possible mechanism for TS. Substrates for the reaction catalyzed by TS are dUMP and CH_2-H_4folate. The products are dTMP and H_2folate. dUMP binds first in the ordered reaction. An intermediate is formed which contains a covalent bond between the SH of cysteine-198(146) and C-6 of dUMP and between C-5 of dUMP and the methylene group attached at N-5 of the cofactor.

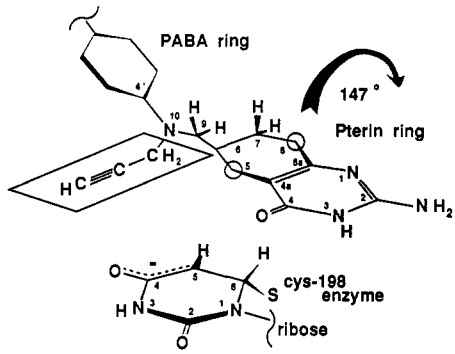

FIGURE 2: Chemical drawing and numbering scheme for CB3717. For reference, the structure of dUMP as observed in the crystal structure of TS·dUMP·CB3717 is also shown. The conformation of CB3717 is the conformation seen in the crystal structure of the fully reduced complex. In crystals for which reducing agent was not added continuously before data collection, the quinazoline ring of CB3717 is rotated 147°, as shown by the arrow.

(Heidelberger, 1957), been targeted for design of inhibitors as potential anticancer drugs. The inhibition characteristics of many dUMP analogues have been determined [summarized by Lewis and Dunlap (1981)]. The structural diversity available in the analogues of folate is greater for the larger molecule [summarized by Santi and Danenberg (1984)].

Quinazoline-based analogues of CH_2-H_4folate have carbon substituted for nitrogen at the 5- and 8-positions and are much stronger inhibitors of TS than the folate analogues. Further substituents placed at the 5-position reduce binding, while substituents at the 10-position greatly enhance binding (Jones et al., 1981). The 10-propargyl derivative 10-propargyl-5,8-dideazafolate (CB3717) (Figure 2) is the most potent of the 10-substituted derivatives, having biological effects in the 1 nM range (Jackson et al., 1983). It is 30-fold more effective than the 10-propyl derivative ($-CH_2CH_2CH_3$). Thus, the mechanism by which the propargyl group ($-CH_2C{\equiv}CH$) acts is responsible for its effectiveness, though it does not attach to the enzyme covalently (Pogolotti et al., 1986). Clinical interest in CB3717 pertains to its effectiveness in treating mice with L-1210 tumors (Jones et al., 1981). It has also shown some effectiveness in treatment of ovarian and breast cancer (Cantwell et al., 1988; Sessa et al., 1988).

CB3717 was developed as an anti-folate to inhibit thymidylate synthase alone, and it is competitive with CH_2-H_4folate (Jones et al., 1981; Jackson et al., 1983) with an inhibition constant of 10 nM (Pogolotti et al., 1986). After incubation of the enzyme with dUMP and CB3717, the inhibition slowly becomes noncompetitive, and a ternary complex is formed which is 30-fold more stable than the TS·FdUMP·CH_2-H_4-folate complex and which dissociates very slowly (Pogolotti et al., 1986). Spectroscopic data and tritium isotope effects suggest that the stable complex forms a covalent bond between the reactive SH of Cys-198(146)[3] and C-6 of dUMP (as is formed in the enzymatic reaction), with protonation of C-5. We have solved the crystal structure of this complex, TS·dUMP·CB3717. Goals of our structural analysis are to understand the reaction mechanism for TS and to relate the liganded enzyme structures with the altered binding properties found for the substrate analogues.

Table I: MIR Statistics for Heavy-Atom Derivatives

derivative	R_{merge}[a]	no. of sites	res (Å)[b]	no. of ref[c]	$\langle F_h/E \rangle$[d]	m[e]
MeHg[f]	0.20	8	3.2	8194	1.9	
$Pt(CN)_4$	0.10	9	3.2	8176	0.94	
PCMB	0.16	8	3.2	8191	1.6	0.58

[a] Residual for derivative data merged to native data. [b] Resolution in Å for the derivative data used. [c] Number of independent reflections used. Redundancy of data collected was >36. [d] Phasing power as calculated heavy-atom scattering/rms lack of closure error. [e] Mean figure of merit for derivative set. [f] Derivatives are methylmercuric chloride, platinum cyanide, *p*-(chloromercuri)benzoate.

Experimental Procedures

Protein Isolation and Crystal Growth. Thymidylate synthase from *E. coli* was isolated from a cloned source as described by Maley and Maley (1988) and Belfort et al. (1983) and stored as an ammonium sulfate precipitate at −80 °C. For crystal growth, the enzyme was dialyzed against 20 mM potassium phosphate, pH 7.0, 0.1 mM disodium ethylenediaminetetraacetic acid, and 20 mM β-mercaptoethanol. Crystals were grown by vapor diffusion from a solution that contained protein (5–9 mg/mL), 20 mM potassium phosphate, 2.3 M ammonium sulfate, 100 mM dUMP, and 4.78 mM CB3717, pH 8.0. Crystals grew over a 3-week period, and β-mercaptoethanol was added to the well buffer every 4 days. An approximate density for the crystals was measured by flotation in bromobenzene and 2-chloroethanol.

Enzymatic activity was assayed by measuring formation of the product, dihydrofolate, from absorbance at 340 nm (Wahba & Friedkin, 1961, 1962; Blakley & McDougall, 1962). Activity of the crystallized protein was determined by rinsing crystals in buffer, dissolving them in inhibitor-free buffer, and assaying enzyme activity. Protein concentration was determined colorimetrically (Bradford 1976).

Data Collection. X-ray intensities were collected on a Nicolet IPC area detector on a three-circle goniometer. Cu Kα X-rays were from a 200-μm focal spot source on a Rigaku rotating target tube with a graphite monochromator or from a 100-μm focal spot size with Franks focusing optics (Franks, 1955). A modified version of the data collection software of Blum et al. (1987) was used to process data frames. Oscillation widths were generally 6 min in ω (for the fine-focus system) or 15 min (for the monochromator-based system). Counting times were as high as 60 min per degree of rotation for the high-resolution data with the focused source. Intensities were reduced with the programs of Howard et al. (1985) and Blum et al. (1987). The integration program SCAN of Blum et al. coupled with the single-parameter scale option in Howard's SCALEI produced the most consistently reliable intensities.

Structure Solution. Initial structure solution was attempted unsuccessfully by molecular replacement, using our initial 3.0-Å structure of the unliganded thymidylate synthase from

[3] The amino acid sequence numbering used here is that of the *L. casei* enzyme, so as to be consistent with our previous publications. The *E. coli* sequence number is listed in parentheses. To convert from *L. casei* to *E. coli* numbering, subtract 2 for numbers ≤89 and subtract 52 for residues >89. A complete alignment of the 17 known TS sequences can be found in Perry et al. (1990). In addition, residues from the "second" monomer which enter into the discussion of the "first" monomer are indicated with a prime, e.g., R179′.

L. casei. The *L. casei* structure has subsequently been refined to an *R* factor of 21% at 2.3 Å from a different crystal form. The rotation function did not result in a clear solution; therefore, heavy-atom derivatives were prepared for multiple isomorphous replacement (MIR). Numerous heavy-atom-containing compounds were soaked into crystals, resulting in three useful derivatives (Table I). Phases were refined by the method of Dickerson et al. (1961). Two mercurial derivatives had nearly identical heavy-atom sites, but with different occupancies. The platinum derivative was poor but useful in confirming the positions and relative occupancies of the two mercurial derivatives. Molecular boundaries and some polypeptide chains were visible in the resulting 3.2 Å resolution MIR electron density map. Phases were further refined through the density modification algorithm of Wang (1985), modified so as to use MIR rather than SIR phases.

Building of *E. coli* TS into the MIR map began by manual fitting of the *L. casei* TS model to the electron density with Jones' model building program FRODO (Jones, 1985). Because of the large conformational change found between unliganded and liganded structures, the two monomers of the dimer were treated independently. Those side chains that differed between the two species were changed to alanine, and two inserts (residues 1–2 and 90–141) unique to *L. casei* removed. This model was subjected to rigid body refinement with the program CORELS (Sussman, 1985), initially on the entire molecule and later on secondary structural segments, separated so as to allow independent rigid body refinement. The remaining *E. coli* side chains were built into an electron density map calculated with the terms $2F_o - F_c$, and the resulting model was subjected to restrained least-squares refinement (Hendrickson & Konnert, 1981) vectorized to run on an FPS264 by Furey (1984).

Extensive rebuilding and refinement followed. The two monomers were built and refined independently of one another to preserve any conformational differences that might exist. The most interpretable electron density maps for much of the rebuilding were calculated from combined calculated (ϕ_{calc}) and MIR phases and $2F_o - F_c$ amplitudes. The phases were combined by a modified weighting scheme (Sim, 1959) similar to the one outlined by Rice (1981), using Hendrickson and Lattman coefficients (1970). Atoms to be rebuilt were removed from the model before phases were calculated for the phase combination. As the model continued to improve, the reliance on MIR phases was no longer necessary, and rebuilding continued into $2F_o - F_c$, ϕ_{calc} electron density maps, where residues to be rebuilt were omitted from phasing. The data used were extended to higher resolution in steps as the refinement progressed. Individual atomic temperature factors were incorporated into the refinement after the *R* factor was reduced below 30%. Molecular dynamics coupled with an X-ray minimization term was run for 9 ps, with the program GROMOS (Fujinaga et al., 1989), and was followed by further rebuilding, placement of water molecules, and restrained least-squares refinement. All residues in the structure are clearly resolved in the final 1.97-Å density map; $R = 18\%$ for all data (16% for $>2\sigma$).[4]

Comparison with Other TS Structures. To compare the complexed structure with the unliganded structure that we independently determined and refined (Perry et al., 1990), a common core was identified and used for least-squares alignment. The common core was defined, by use of difference distance matrices, as the largest group of α-carbon atoms in the complete dimer whose distances from each other did not change by more than ±0.5 Å between the two structures and where each atom was within 10.0 Å of at least one other atom in the core. The 10.0-Å limit preserved the ability to identify independent domains and to search for domain movement. We describe the core selection algorithm elsewhere (Perry et al., 1990). A total of 155 α-carbons (out of 528) were so identified and used to overlap the two structures by least-squares minimization for comparison of the difference in conformation.

After alignment, it was apparent that segments of the protein had different positions, consistent with the ligands having induced coordinated movements of segments of the structure. For each of these segments individual vector shifts of Cα atoms were summed together to give a total shift vector, which was compared with the total expected from summation of the same vector magnitudes in random orientation. It is well established that the summation of a large number of vectors $\mathbf{v}_i$ of uniform length, in three dimensions, leads to a probability for a sum of a given magnitude x given by a Maxwellian distribution (Maxwell, 1860; Rayleigh, 1919):

$$P(x) = Ax^2 \exp[-3x^2/(2\sigma^2)]$$

where σ, the root mean square value is given by

$$\sigma = \sqrt{(\sum|\mathbf{v}_i|^2)} = \sqrt{(3/2)}x'$$

where x' is the most probable length. The normalizing term A is $(1/\sigma^3)(54/\pi)^{1/2}$. Simulations to test whether addition of small numbers of shift vectors of *nonuniform* length, oriented randomly, would follow a similar distribution were carried out. A random number generator was used to pick vectors, and 50 000 simulated additions were used to obtain an expected distribution of lengths. For additions involving more than four vectors, the distribution was approximately Maxwellian. For a Maxwellian distribution, 95% of the lengths in the distribution are less than 1.6σ. For three and four vector distributions the 95% level is at about 1.5σ.

Shift vectors of adjacent α-carbons between two structures are necessarily not randomly oriented since they are coupled through covalent bonds. To test whether the addition of shift vectors from random errors in a protein structure approximates a Maxwellian distribution, eight segments of at least six residues long were picked from a pair of independently determined bovine trypsin structures (Chambers & Stroud, 1979) from current coordinate sets, and their shift vectors were summed. Their magnitudes follow that predicted by the Maxwellian distribution. Thus, summed shift vectors whose total is greater than that at the 95% confidence level are considered to move significantly in the direction of the final vector.

Results and Discussion

Crystal Characterization. Hexagonal rod-shaped crystals appear in 2–3 weeks and grow to 0.3 × 1 mm in size. They diffract to approximately 1.9 Å, though the data beyond 2.3 Å are weak. The space group is $P6_3$. Unit-cell dimensions are $a = 127.1$ Å and $c = 67.9$ Å. If the reducing agent β-mercaptoethanol was not added repeatedly, the cell dimensions increased almost imperceptibly to $a = 127.3$ Å and $c = 68.2$ Å, and intensities in the diffraction pattern changed by a few percent. They have a density of ~ 1.25 g/cm^3 and contain one dimer (60 800 daltons) per asymmetric unit.

[4] The structure of an *E. coli* ternary complex with dUMP and CB3717 crystallized in a different space group from that reported here has been determined by D. Matthews and his colleagues at Agouron Pharmaceuticals. We jointly decided not to compare our results except in general terms prior to publication. Results from the two studies were submitted simultaneously to separate journals with no prior exchange of coordinates or manuscripts.

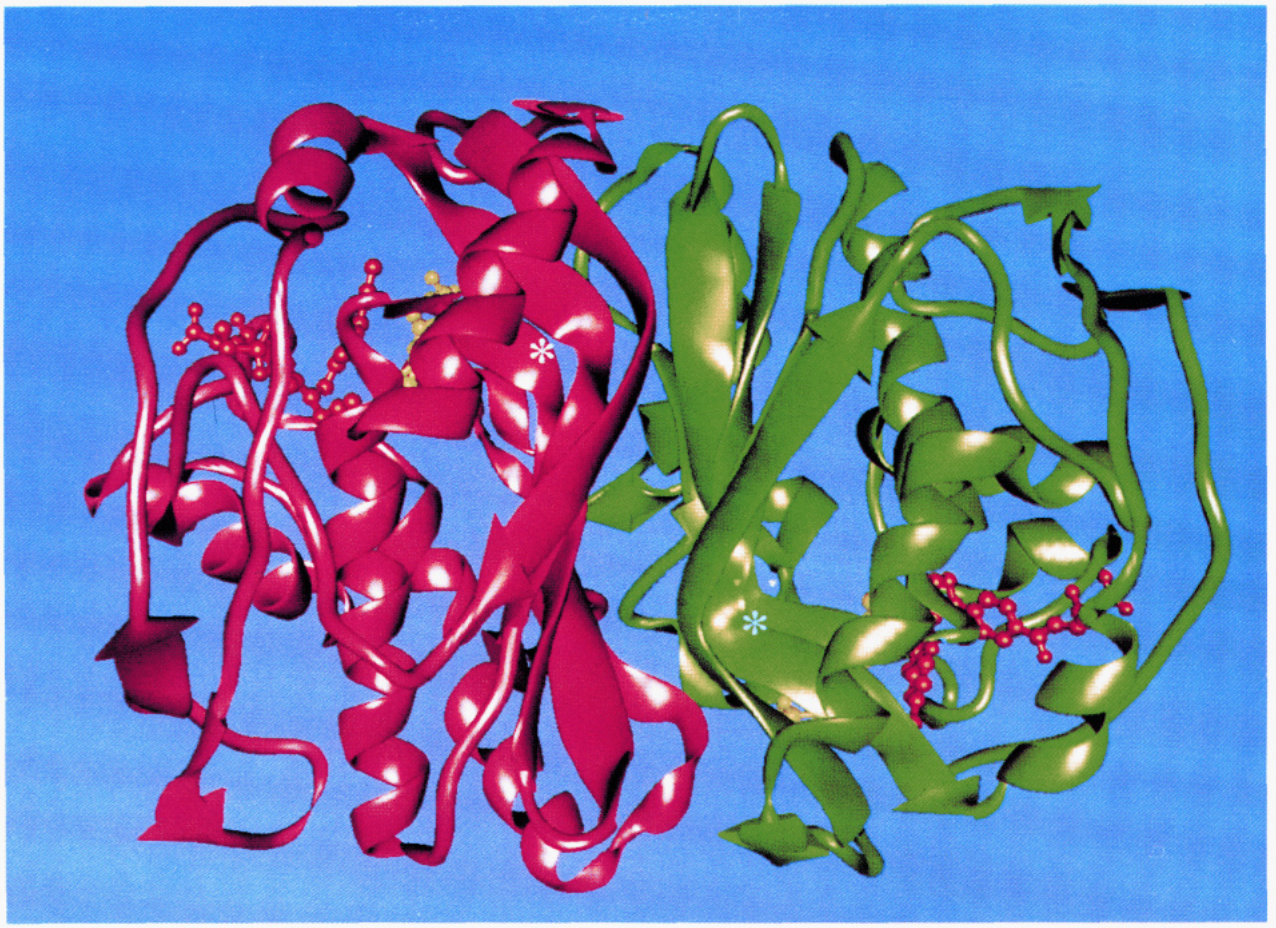

FIGURE 3: Ribbon drawing of folding in the TS dimer. There is an extensive six-stranded β-sheet interface between monomers that supports a β-kink in three strands of the β-sheet, i–iii. The kink is indicated with an asterisk.

The integrity of the inhibited complex was measured by rinsing and dissolving crystals. The specific activity of the complex at a concentration of 0.24 mg/mL was 4% of that of fully active unliganded enzyme (0.09 unit/mL). Dilution by a factor of 10 resulted in an increase in activity to 30% that of free enzyme (0.61 unit/mL), as expected for the reversible dissociation of CB3717. Thus, most of the enzyme is inhibited in the crystallized ternary complex, and after dissociation of the inhibitor the enzyme is fully active. Crystals of unliganded TS from *E. coli* are of a different space group (Perry et al., 1990) and are stable in well buffer but dissolve immediately upon addition of a well buffer containing CB3717 and dUMP.

Structure Solution. Model building began in an MIR electron density map subjected to density modification by Wang's solvent flattening approach (Wang, 1985). Secondary structural features from the *L. casei* TS structure were apparent in this map, including the six-stranded β-sheet at the dimer interface and several helices. However, there were numerous breaks in the electron density representing the polypeptide backbone, and many cycles of model rebuilding followed by crystallographic refinement were required to solve the structure. For much of this analysis, $2F_o - F_c$ vs ϕ_{calc} electron density maps were uninterpretable, and interpretation proceeded in maps calculated with combined MIR and model phases. The current model is based on data to 1.97-Å resolution with a crystallographic R factor (R_{crys}) of 18% for all non-zero data between 15.0- and 1.97-Å resolution. The RMS deviations of bond and angle distances from ideal values are 0.022 and 0.059 Å, respectively. Shown in Figure 3 is a representation of the ternary complex model. The conformations of all 528 residues are clear in the final electron density map, although some surface side-chain conformations are difficult to interpret. Two dUMP molecules, two CB3717, and 255 water molecules have been located.

Lack of Errors in Interpretation. To address the issue of what errors may be expected in the structures we report here, we show the worst electron density in the map of the ternary complex, for three consecutive residues that have the highest temperature factors (Figure 4). The fit is unambiguous. Therefore, it is unlikely that any significant errors are to be found in the current structures we discuss.

Overall Structure. The overall structure of TS·dUMP·CB3717 is similar to our 2.3-Å phosphate-inhibited TS structure from *L. casei* or *E. coli* (Perry et al., 1990). Segments of the secondary structure lie in slightly different positions, and coordinated shifts of residues toward the active center are apparent in the complex. Crystals of the ternary complex contain one dimer in the asymmetric unit; however, conformational differences between monomers are small (Figures 3 and 5).

(A) β-Sheet Interface and β-Kink. The dimer interface is quite unique in having a right-handed twist between β-sheets (Hardy et al., 1987; Mathews et al., 1989). It is comprised of a back to back apposition of two five-stranded sheets with hydrophobic and hydrophilic contacts between them, one residue from a sixth strand ($-i$), and several residues from other regions of the protein. Defining enclosure in terms of a change in the solvent-accessible surface area of a residue from more than 11% accessible in the structure of the monomer to less than 11% in the dimer, this interface encloses 14 hydrophobic, 11 polar, and 5 charged side chains. Six water molecules make hydrogen bonds to both monomers, with three of these six completely buried (0% solvent accessible) in the interface. One of these buried water molecules lies on the

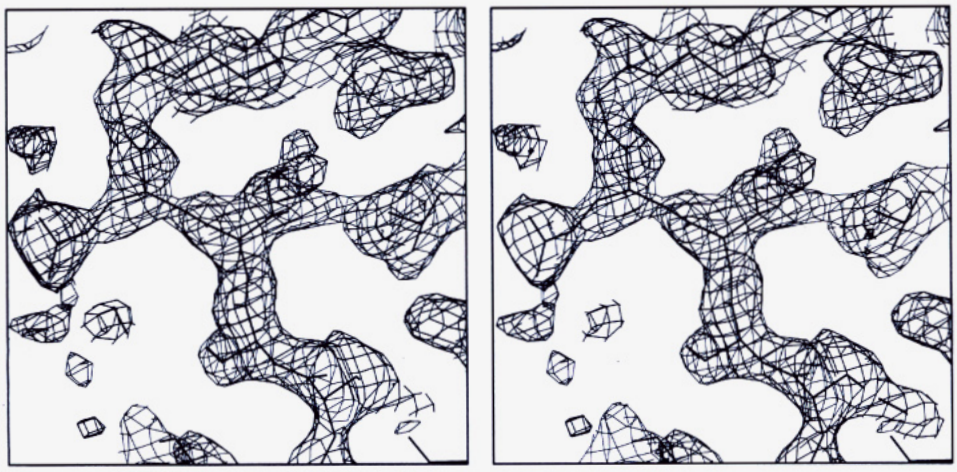

FIGURE 4: To address the question of reliability in crystal structure determinations, we show here the electron density seen in the worst region of the $2F_o - F_c$ map, that surrounding the three consecutive residues with the highest temperature factors in the structure. The electron density for this region is nevertheless clearly interpretable. Therefore, it is very unlikely that there are any significant errors in our TS structures.

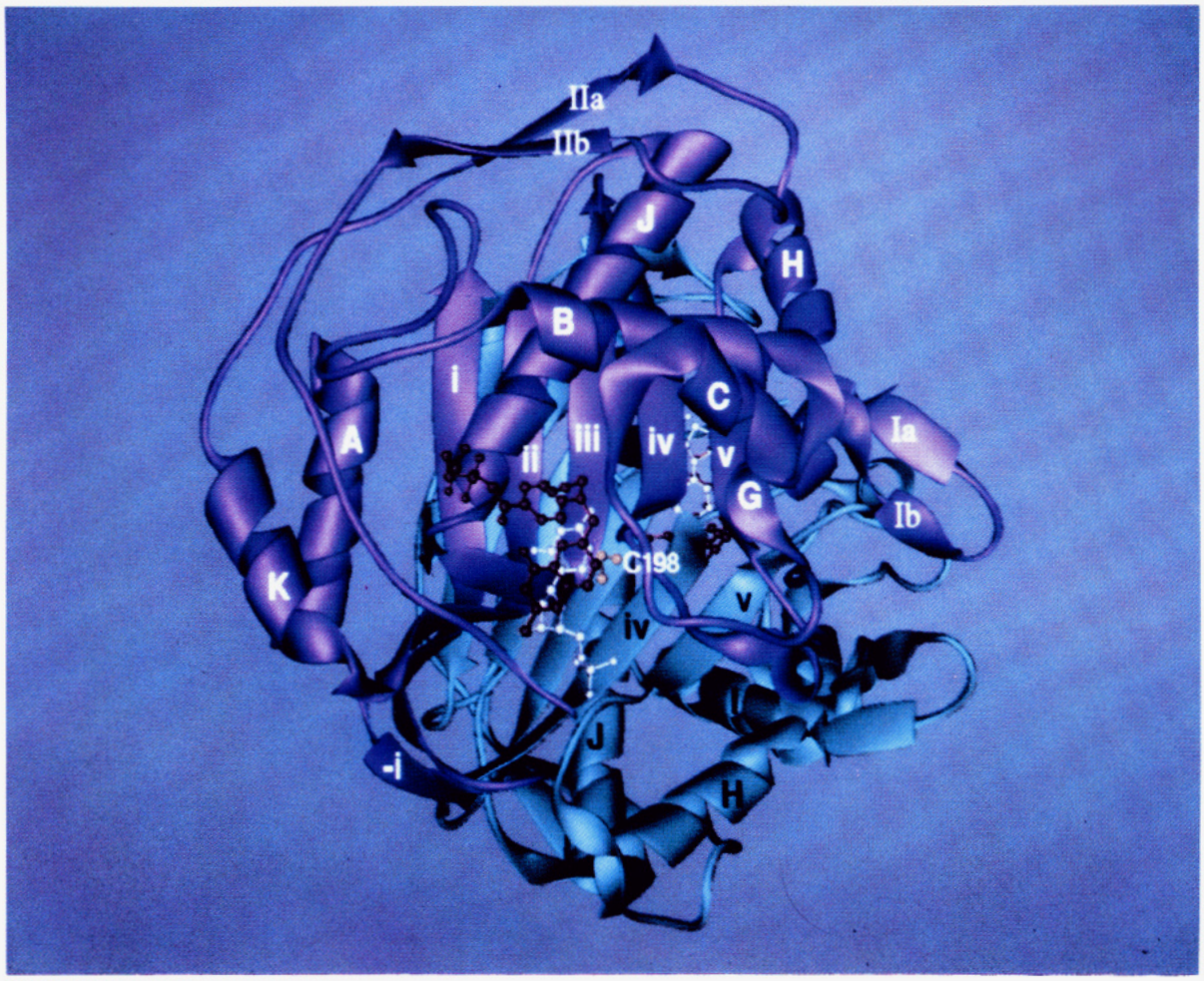

FIGURE 5: Ribbon representation of the ternary complex in the TS dimer showing ligands in the monomer I active site. The two ligands contact all sides of the active site. The amino terminus is in the upper left-hand corner, the carboxy terminus in the active site. The secondary structure is labeled as in our original publication of the *L. casei* structure (Hardy et al., 1987) for consistency. In *E. coli* TS there are fewer residues due to the deletions of residues 1–2 and 90–139. The deletion includes helices E and F, and helix D is no longer truly α-helical in *E. coli*. The definition of segments is alphabetic for helices in order within the sequence as follows: helix A, residues 1–16(14); B, 54(52)–67(65); C, 71(69)–78(76); G, 145(93)–152(100); H, 162(110)–174(122); I, 189(137)–193(141); J, 225(173)–245(193); K, 264(212)–271(219). For β-sheets, numbering in Roman numerals is from i at the left of the β-sheet interface to v at the right of the sheet. β-Strand –i was not found in the *L. casei* structure. Otherwise, the β-sheet strands are as follows: –i, 18(16)–20(18); i, 27(25)–38(36); ii, 250(198)–262(210); iii, 210(158)–222(170); iv, 198(146)–207(155); v, 180(128)–184(132). β-Ribbon I, 153(101)–155(103) and 159(107)–161(109);; β-ribbon II, 282(230)–286(234) and 298(246)–302(250).

molecular pseudo-2-fold symmetry axis. Of the four side chains hydrogen bonded to the buried waters in the interface, two are absolutely conserved [Trp-185(133) and Gln-163-(111)], and two have conserved hydrogen bond donor character [Lys-20(18) and Ser-183(131) near the dimer 2-fold axis]. In general, the water molecules outside the molecule are not

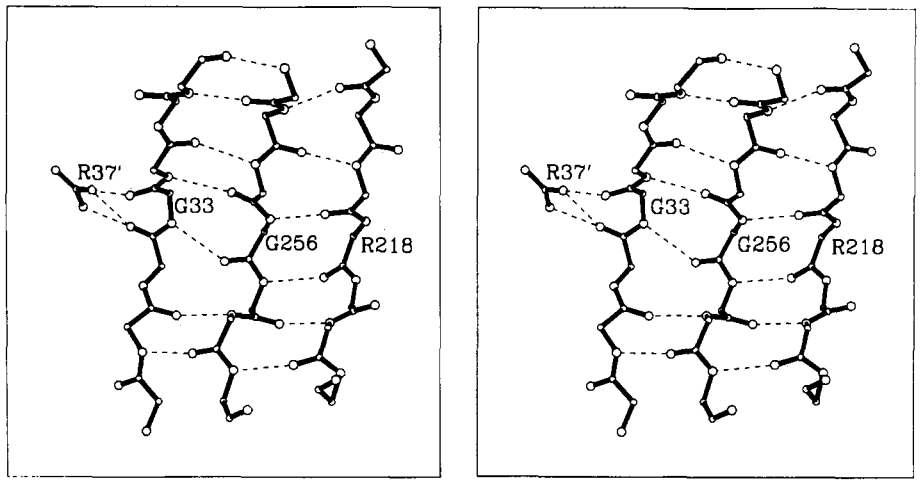

FIGURE 6: Hydrogen bonding among the residues associated with the β-kink. The main-chain atoms of β-strands i–iii are shown. The β-sheet interface has paired regions of polar and hydrophobic patches. These may force the unique right-handed twist between strands of the interface in TS, and so generate the three-strand β-kinks that curl around key residues of the active center. Normal β-sheet hydrogen bonding is maintained through the β-kink, although it is somewhat strained. the ϕ,ψ angles for the kink in strand ii are those of a left-handed helix; those of strands iii and iv are those of an α-helix. The single-letter amino acid code and the numbering scheme from the *L. casei* enzyme are used to identify residues at the kink and Arg-37′(35′). Hydrogen bonds are denoted by dashed lines and oxygens and nitrogens by larger circles. This and all other stereo plots are "crossed-eye" (convergent) stereoviews.

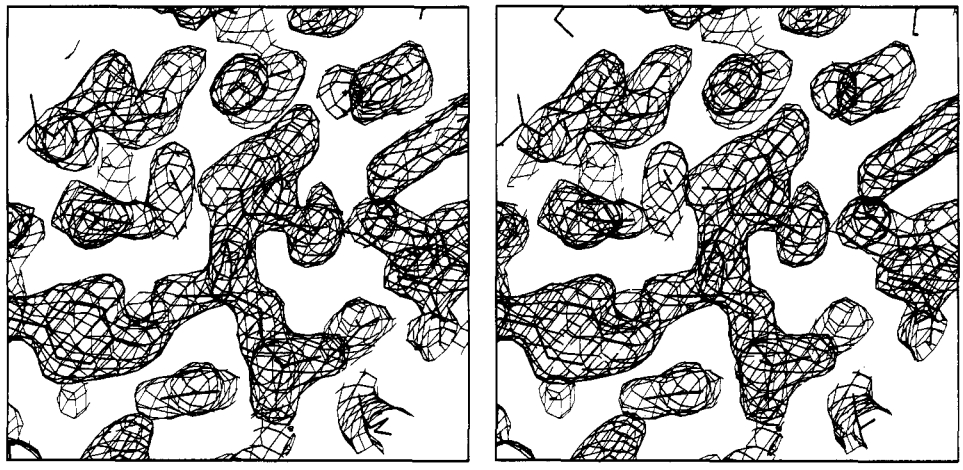

FIGURE 7: Electron density for the substrate dUMP. Water molecule H_2O 1 is located 3.1 Å from O-4 of dUMP, which becomes partially negatively charged during the reaction. H_2O 1 is hydrogen bonded to conserved His-199(147) and Glu-60(58). This and other well-ordered water molecules in the active center are probably of key importance to the enzymatic reaction.

two-fold related. The spatial arrangement of patches of hydrophobic and hydrophilic side chains may determine how the two monomers associate, resulting in the only known case of right-handed β-sheet packing. In the lower left-hand corner of the model as displayed in Figure 5, three strands of the β-sheet bend sharply away from the remainder of the sheet in what we term a β-kink. This kink forms a part of the active site pocket and constrains several key substrate binding residues. The main-chain hydrogen bonding in this region is quite unusual, as illustrated in Figure 6. The peptide planes in all three strands are rotated around such that two side chains in a row are on the same side of the β-sheet, although all main-chain hydrogen bonding is satisfied.

(*B*) *Identity of Monomers.* Thymidylate synthase is a symmetric dimer of identical monomers with two active sites formed in the cleft between subunits on either side of the molecule. Binding studies using equilibrium dialysis and spectroscopy (Santi & Danenberg, 1984) indicate that the two active sites may not react equivalently. In the most extreme form, it has been suggested that the enzyme uses only one active site at a time (Danenberg & Danenberg, 1979), even though both active sites can be filled through covalent attachment of the inhibitor 5-fluoro-dUMP and CH_2-H_4folate. The difficulty in performing and interpreting these experiments has led to conflicting results and considerable debate in the literature.

The asymmetric (unique) portion of the ternary complex crystal form is a dimer, whereas in all of the unliganded TS crystal structures we have solved to date (human, T4, *E. coli*, *L. casei*) the asymmetric unit is a monomer. Therefore, structural differences between monomers can only be observed in the ternary complex crystal form. The monomers in the ternary complex structures are similar to one another but differ substantially from the unliganded structures. However, the average temperature factor (*B*) for one monomer is considerably less than for the second monomer (11.6 compared with 18.6 $Å^2$ for main-chain atoms), even though the distribution of *B* values was similar within each monomer. The *B* factors for the corresponding monomers when β-mercaptoethanol was withheld from the ternary complex crystals were 15.7 and 23.7 $Å^2$, respectively. It is possible that the difference in mobility between the monomers in the complex structure reflects a mechanistic nonequivalence of active sites in TS. The difference may also be the result of crystal packing forces. At present we cannot distinguish these possibilities.

Ligand Binding. (*A*) *The Substrate dUMP.* dUMP binds

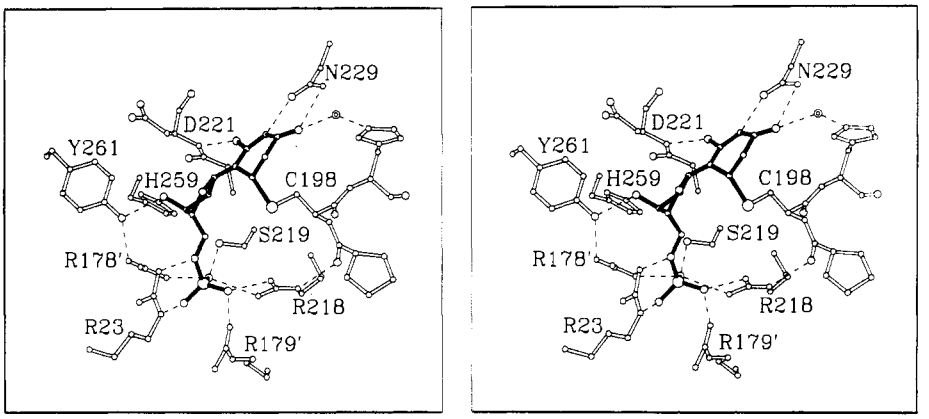

FIGURE 8: dUMP binding site shows the covalent bond between C-6 of dUMP and the SH of Cys-198(146) and the donor–acceptor hydrogen-bonding pair from Asn-229(177) to donor N-3 and acceptor O-4 of dUMP. Key water molecule H_2O I is the donor to O-4 also and in turn is bonded to conserved His-199(147) and Glu-60(58).

Table II: Hydrogen Bonds to dUMP Phosphate

atom	residue	atom	distance (Å)
γOH	Ser-219(167)	O-1	2.7
εNH	Arg-178′(126′)	O-1	2.7
η_1NH	Arg-218(166)	O-1	2.9
η_1NH	Arg-218(166)	O-2	3.2
η_2NH	Arg-218(166)	O-2	2.5
ηNH	Arg-179(127′)	O-2	2.9
εNH	Arg-23(21)	O-3	2.7
ηNH	Arg-23(21)	O-4	3.1

first to TS, before folate, in an ordered binding sequence. The electron density seen in a $2F_o - F_c$ map around the dUMP site is shown in Figure 7. The phosphate ion is bound between four arginine side chains, two from each monomer, and one serine, 219(167). Seven hydrogen bonds, each with good geometry to be donors, and one close contact with poor geometry are made with the oxygen atoms of the phosphate ion as shown in Figure 8 and listed in Table II. These residues are absolutely conserved in all 17 known species of TS [aligned in Perry et al. (1990)], with the exception of Ser-219 → Cys in *Crithida fasiculata* (Hughes et al., 1989) and Arg-23 → Gly in *Bacillus subtilis* phage φ3t (Kenny et al., 1985).

Hydrogen bonds and other contacts to the uridine ring may encode base specificity and define possible binding modes for other nucleotide analogues (Figure 8). All residues discussed are absolutely conserved except as noted. Asn-229(177) makes the most specific interaction; the δNH_2 is a hydrogen-bond donor to O-4, and the δO is a hydrogen-bond acceptor from N-3H on dUMP. These hydrogen bonds would be compromised in any nucleotide modified at the 4-position. In support of this, the K_m for 4-thio-dUMP is increased to 7×10^{-5} M (Kalman et al., 1973) from the wild-type K_m of 1.2×10^{-6} M. This corresponds to a difference in binding energy of about 2.5 kcal/mol. Further discrimination at O-4 of dUMP may be provided by a key well-ordered water molecule (H_2O 1) which may also have a role in catalysis. This water molecule is a hydrogen-bond donor to O-4 (3.1 Å away), as well as to the εO of Glu-60(58) (2.7 Å away), and is a hydrogen-bond acceptor from the εNH of His-199(147) (2.7 Å away). The orientation of the imidazole of His-199(147) is stabilized by a hydrogen bond between the δN and ηO of Tyr-233(171), and the environment is highly conserved, with the exception that His-199(147) is substituted for a nonpolar valine in *B. subtilis* phage φ3t (Kenney et al., 1985). Lack of conservation of His-199(147) in the most distantly related species along with insensitivity of the enzyme to mutations of His-199(147) to small, uncharged residues (Dev et al., 1989) indicates it has no indispensable role in catalysis.

In addition to specific interactions at N-3H, O-4, and C-6 of the dUMP pyrimidine ring, O-2 is hydrogen bonded to the main-chain amide NH of Asp-221(169) (2.9 Å). This amide group is further constrained by a hydrogen bond from the side-chain δO to CB3717 and from Ser-219(167) to the phosphate of dUMP.

The ribose ring is bonded by several absolute conserved residues also. These include hydrogen bonds between ribose O-3′ and His-259(207) (2.7 Å) and between O-3′ and Tyr-261(209) (2.6 Å), both of which have ideal geometry for the role of hydrogen donor or acceptor. TS is very sensitive to changes at the 3′-position. 3′-Deoxy- or dideoxy-UMP does not bind appreciably (Holy & Votruba, 1974). Presumably, this conserved hydrogen bonding at the O-3′ position is a major determinant of specificity. Replacement of Tyr-261(209) by anything other than Met yields an enzyme that is compromised by a factor of 10^{-4} or greater, since it will not complement a strain that lacks TS (referred to as thy$^-$) (Climie et al., submitted for publication).

Binding of uridine monophosphate (with oxygen at position 2′) to TS is 40 times worse than that for dUMP (Reyes & Heidelberger, 1965). Analogues with anything larger at the ribose 2′-position do not shown any inhibitory activity at all (Wohlrab et al., 1978). The carbon atom C-2′ lies at the back of the TS binding pocket and is 3.5 Å from the NH of conserved Asp-221(169) and 3.3 Å from γO of Ser-219(167). Thus, binding of nucleotides modified at 2′ is sterically hindered by the protein.

(B) CB3717 Binding. The folate analogue CB3717 binds in a substrate-like conformation with the quinazoline ring parallel to and in van der Waals contact with the uridine ring of dUMP. The electron density map for this region is shown in Figure 9. The structure of CB3717 and its environment are shown in Figure 10. Like the cofactor CH_2-H_4folate, it is composed of three parts, with somewhat independent degrees of freedom. These include (a) the quinazoline ring which is the dideaza analogue of the pterin ring of CH_2-H_4folate, (b) the *p*-aminobenzoic acid (PABA) ring, and (c) the glutamate residue(s) attached to the PABA ring. CB3717 contains only one glutamate, though the physiological form of the cofactor contains as many as seven. The connecting region between the pterin and PABA rings in CH_2-H_4folate would contain the methylene group that is to be transferred to the C-5 of dUMP.

The quinazoline and PABA rings are surrounded by highly conserved residues which close down onto the reactants. On the quinazoline ring, NH_2-2 is the donor to the carbonyl of the penultimate residue 315. A water molecule (H_2O 2)

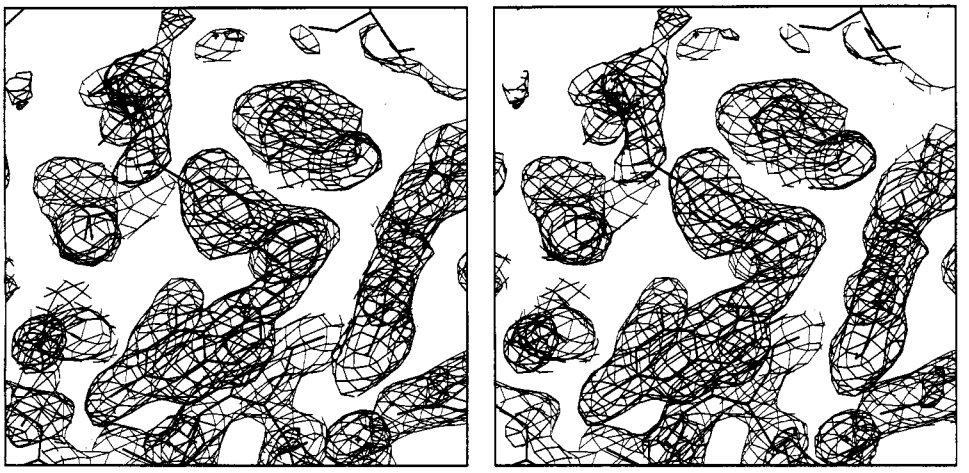

FIGURE 9: Electron density seen in a $2F_o - F_c$ map for CB3717 as seen in the major conformation, probably very close to that of the bound substrate in the catalyzed reaction.

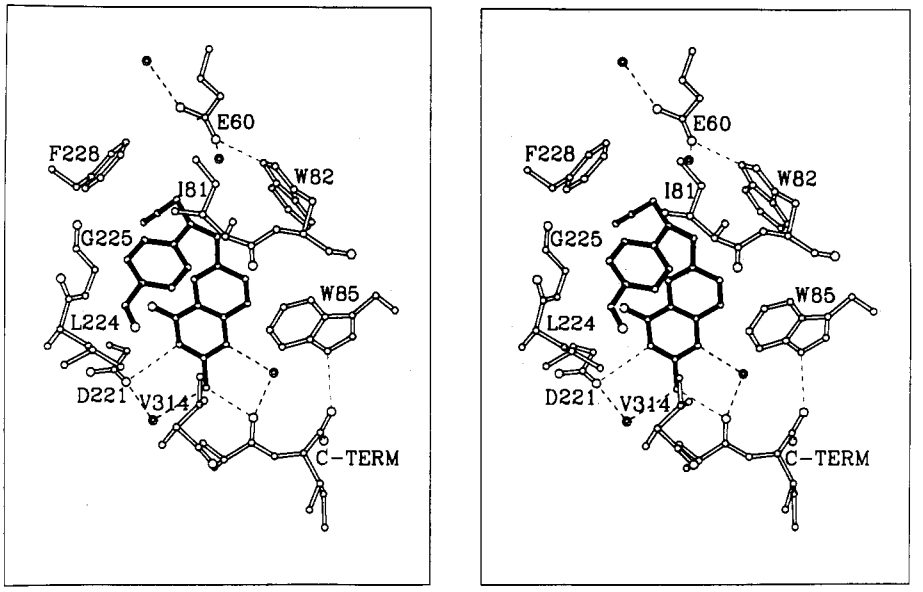

FIGURE 10: Surrounding structure of the major site for the CB3717 binding. The propargyl group is in contact with conserved Phe-228(176), and the quinazoline ring lies on top of the substrate uridine ring.

cross-links this C=O to N-1 of the ring system. NH-3 is the donor to Asp-221(169), and O-4 is 3.26 Å from the NH of Gly-225(173) though the geometry is poor for a hydrogen bond. A third ordered water (H_2O 3) is a hydrogen-bond acceptor and donor to NH_2-2 on the quinazoline ring and Asp-221(169), respectively. There are no obvious receptors for the NH at position 8, present only in the cofactor but not in the quinazoline ring. Trp-85(83), Leu-195(143), and conserved Trp-82(80) make hydrophobic contacts with the quinazoline ring, and Trp-82(80) is hydrogen bonded to conserved Glu-60(58), in the network around the reactive site. The most important contact for the quinazoline ring is with dUMP. Nine of the twelve atoms in the quinazoline are less than 3.8 Å from some part of dUMP. The PABA ring is surrounded by several conserved hydrophobic residues [Leu-224(172), Ile-81(79), and Phe-228(176)].

The quinazoline ring of CB3717 binds in a manner close to that required of the pterin ring in the TS·dUMP·CH_2-H_4folate ternary complex, suggesting a mechanism for how the normal catalytic reaction may occur. In particular, C-5 of the quinazoline ring is 3.6 Å from C-5 of dUMP. In the TS·dUMP·CH_2-H_4folate ternary complex, pterin N-5 and dUMP C-5 are connected through a methylene bridge. The details and mechanistic implications of this arrangement are explored by Finer-Moore et al. (1990). dUMP forms a large part of one surface of the binding site for the quinazoline (or pterin) ring, which explains the ordered binding of reactants (Danenberg & Danenberg, 1978).

The propargyl group, responsible for the highest specificity of the series of 10-substituted quinazoline-based compounds for TS, lies against conserved Phe-228(176) and enters the "alternate site" for the quinazoline ring (see below). The triple-bonded terminal carbon of the propargyl group lies close to Asn-229(177) and N-3 of dUMP.

(C) The Alternate Site for the Quinazoline Ring of CB3717. Remarkable is the finding of a second site for binding of the quinazoline, a large two-ring system that is a close analogue of the pterin ring, in a second crystal structure of a chemically altered TS. Electron density in a $2F_o - F_c$ map for this crystal structure is shown in Figure 11. The change that led to occupation of the site is produced by withholding the readdition of reducing agent, which is initially present in crystallizing media. All crystals are prepared in the presence of 20 mM β-mercaptoethanol (BME). However, if allowed to remain for some time without reintroduction of reducing agent, they change unit-cell dimensions slightly, and a second binding site for the quinazoline ring is seen which is rotated approximately 147° away from the main, substrate-like site (Figures 12 and

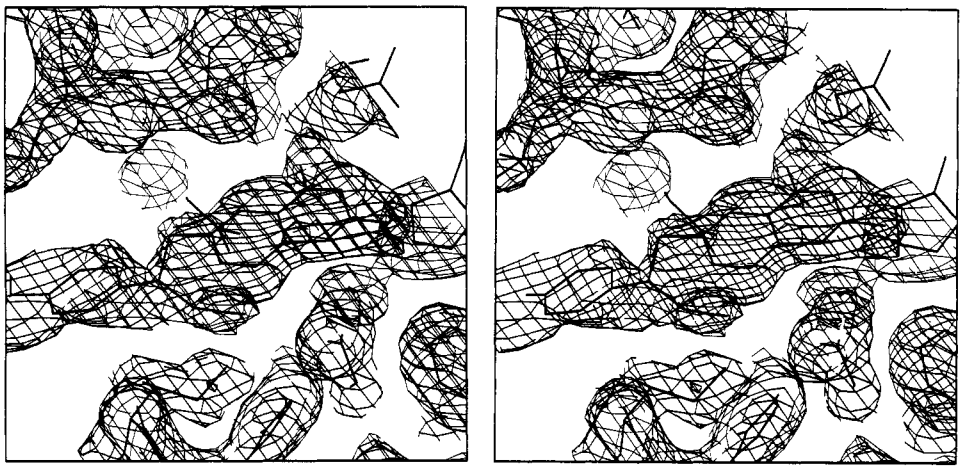

FIGURE 11: Electron density seen in a $2F_o - F_c$ map, for CB3717 in the alternate site. This site is surrounded by many conserved residues to which it is hydrogen bonded. This site is vacant, filled with ordered water in the structure of the major, cofactor-like binding site. However, its preservation suggests a clearly functional role for this alternate site in the TS reaction.

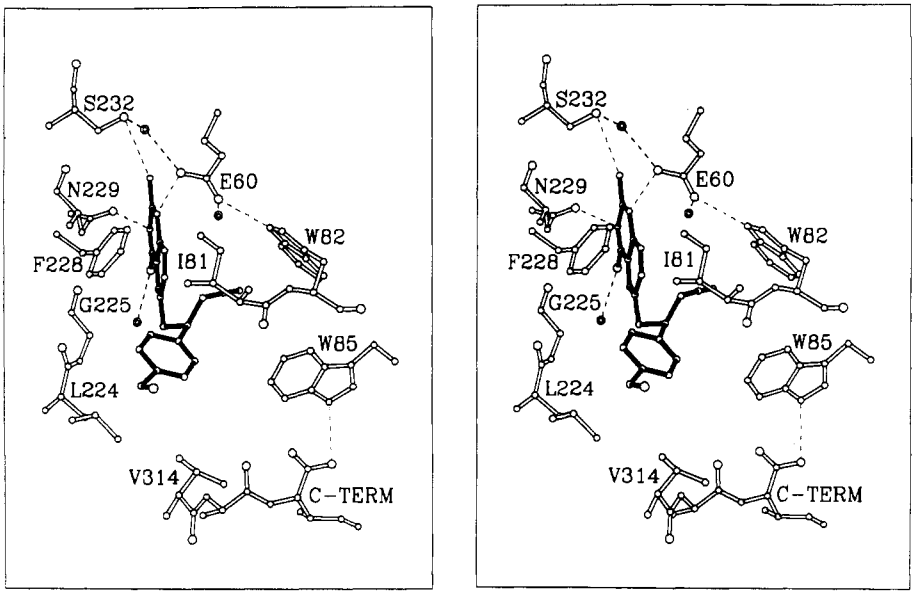

FIGURE 12: Minor site for binding of CB3717 is surrounded by several conserved residues and could only be occupied by folates in which the imidazolidine ring is opened.

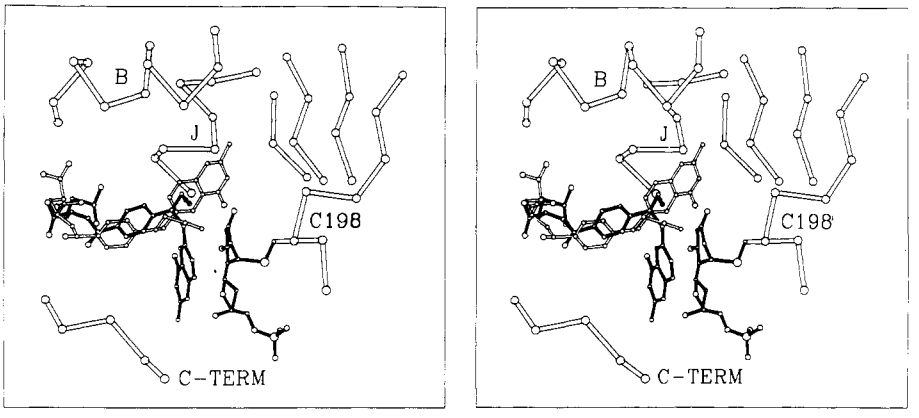

FIGURE 13: Shown are the major (in solid line) and minor (in hollow line) sites for bound CB3717, as solved in two different crystal structures. In its major site, the only one that could be occupied by the initial form of the cofactor, the quinazoline ring of the folate analogue packs against the uridine ring of dUMP (solid line). C-6 of the uridine ring is bonded to Cys-198(146), and C-5 is placed close to the position where a methylene attached between N-5 and N-10 of the ring would lie in CH_2-H_4folate. The *p*-aminobenzoic acid ring (PABA ring) differs by 1 Å but is still surrounded by conserved hydrophobic residues. The two positions for the quinazoline ring are related by a 147° rotation.

13). The alternate binding is reversible by addition of BME to the crystal. This second site could not occur in binding of the initially reduced cofactor, since the cofactor is further constrained by the methylene-bridge-containing imidazolidine ring. However once the imidazolidine ring is opened, as occurs as a consequence of transfer of the methylene group to C-5 of dUMP, the pterin ring of H_2folate could occupy this site.

Difference maps show that there are only small differences

in several of the cysteines, which are modified. This reaction in proteins is poorly characterized. There are ten cysteines (-SH) in *E. coli* TS of which several are unmodified in the structure. Those that are modified seem to have as little as one atom covalently bound.

In the alternate site the quinazoline ring is hydrogen bonded to even more conserved residues than in the primary site. N-1 is hydrogen bonded to the ϵO of Glu-60(58). NH_2 at position 2 is hydrogen bonded to the OH of Ser-232(180). NH-3 is the donor to the δO of Asn-229(177), which is displaced from its position in the major conformation. The C=O at position 4 is a receptor for a hydrogen bond from a nearby ordered water molecule, H_2O 4. Most surprising is that this large site is vacant in the primary CB3717 structure. The site is surrounded by highly conserved residues found in all 17 known sequences for species of TS, suggesting a role in the reaction mechanism. There are many hydrogen bonds and close contacts characteristic of a functional role for this site in the mechanism. The remaining parts of the quinazoline derivative, the PABA ring and glutamate, lie in about the same position in both binding modes. (Figure 13). This second site also offers possible new avenues for design of more specific inhibitors of TS.

The propargyl group lies in contact with the pyrimidine ring and with Trp-82(80). Thus, the electrons of the triple bond are in contact with the electron-poor region at the center of the indole ring of Trp-82(80).

Each monomer is different in the alternate-site structure, due to structural feedback between active sites for which there is some evidence in the normal enzymatic reaction or due to the different stability of each monomer in the crystal toward partial oxidation. In the incompletely reduced form the main differences seen between monomers are in the active site. One active site contains both CB3717 in the alternate site and dUMP. In the second site the cofactor analogue is shifted up to 1 Å from that of the first active site, and only a phosphate ion is bound in place of dUMP.

(D) CH_2-H_4folate Binding. The binding site for the quinazoline ring immediately suggests the very similar manner in which the pterin ring of CH_2-H_4folate binds after dUMP to initiate the reaction. Hydrogen bonds with the protein are made by N-1 with H_2O 2 that in turn bonds to the C=O of 315(263), NH_2-2 and the C=O of 315(263), and NH-3 and the carboxylate of Asp-221(169). From the point of view of drug design it is immediately noticeable that there is room for modification at NH_2-2, within a polar environment on the enzyme. Modeling of the chemical reaction, beginning with the TS·dUMP·CH_2-H_4folate complex, is described by Finer-Moore et al. (1990).

Conformational Change. TS from both human and *L. casei* undergo a reduction in Stokes radius of about 3.5% upon ternary complex formation (Lockshin & Danenberg, 1980), suggesting that a large conformational change takes place. In a comparison of the structures from *E. coli*, the conformational changes that do occur generally serve to close down the active site and sequester the reactants, along with a large number of ordered water molecules, away from the bulk solvent. The radius of gyration calculated for the dimer complex with ligands present was 23.10 Å compared with 23.20 Å for the unliganded dimer from *E. coli*. This 0.4% difference is not inconsistent with the 3.5% difference in Stokes radius observed upon ligation, since hydrodynamic methods are sensitive to outer radius, shape, and flexibility, which change dramatically.

The change in conformation is not a shift of domains as seen in other proteins [such as yeast hexokinase (Anderson et al., 1978; Bennett & Steitz, 1980) and citrate synthase (Remington et al., 1982; Huber & Bennett, 1983; Lesk & Chothia, 1984)]. To quantitate conformational change, a core of α-carbons of residues whose mutual disposition remains unchanged ±0.5 Å between liganded and unliganded *E. coli* TS was selected by use of difference distance matrices, and is shown in Figure 14. To find the core, it is necessary to use the full dimer. Searching with each monomer inevitably gives the erroneous impression that the dimer interface twists around the line joining subunit centers. This occurs because monomer cores consisted of a different subset of residues whose relative interatomic distances remain the same. As these lie more to one side of the monomer, the general closure around reactants would appear falsely as an angular displacement.

The two structures were overlapped by least-squares minimization of the difference between the cores. The rms deviation between core atoms (155 atoms) was 0.24 Å, between all α-carbons was 0.90 Å, and between all main-chain atoms was 0.99 Å. (A similar overlap performed on the two monomers in the ternary complex yields an rms deviation between α-carbons of 0.22 Å for the core atoms, 0.37 Å for all.)

To accurately describe shifts between two protein structures, it is essential to compare differences in position with those that would be expected to arise from random errors, as would occur between multiple cycles of refinement after convergence. The random errors are a function of thermal factor B, resolution, and R factor (quality of converged refinement). Expected random errors in position of each structure were estimated according to the quadratic relationship between error and temperature factor B [Chambers and Stroud (1979) and eq 5 of Perry et al. (1990)]:

$$\sigma_{xyz,\mathrm{error}}(B) = (3/4)rR(aB^2 + bB + c)$$

where $a = 0.00146$, $b = -0.0203$, and $c = 0.359$. r is the resolution in angstroms. R is the residual (>2σ data) as a fraction, and B is the thermal factor in Å^2. This relationship provides an empirically determined estimate of the standard deviation of a normal error distribution expected for an atom of given thermal factor B. After liganded and unliganded structures were overlapped, the rms deviation in position (0.99 Å for main-chain atoms) was much larger than the overall standard deviation expected for commuted errors in both structures (0.27 Å) on the basis of average thermal factors (Figures 14 and 15). Many regions of low temperature factor in both structures, indicative of well-determined structure, differed by 0.5–1.25 Å. With the exception of the C-terminus, which differs by over 4.0 Å, other shifts are less than 1.5 Å.

The structure of the ternary complex differs from the unliganded structure by a large ligand-induced conformational change. In contrast to the movements of large domains seen in other proteins upon ligand binding (Huber & Bennett, 1983), the change in TS involves segmental accommodation, in which segments of β-sheet, α-helix, and connecting loops are moved in concert toward the active center. The two different binding modes for CB3717 differ by a smaller shift in these residues. There is little evidence yet for cooperativity between the subunits in TS; thus, the large conformational change that involves many highly conserved residues is most probably linked to correct binding and orientation of components that must themselves undergo some dramatic change in stereochemistry and structure, during the multistep reaction mechanism.

Segmental Accommodation. Shown in the upper half of Figure 15 are vectors relating the α-carbons of unliganded TS to the ternary complex. The active sites close to surround the

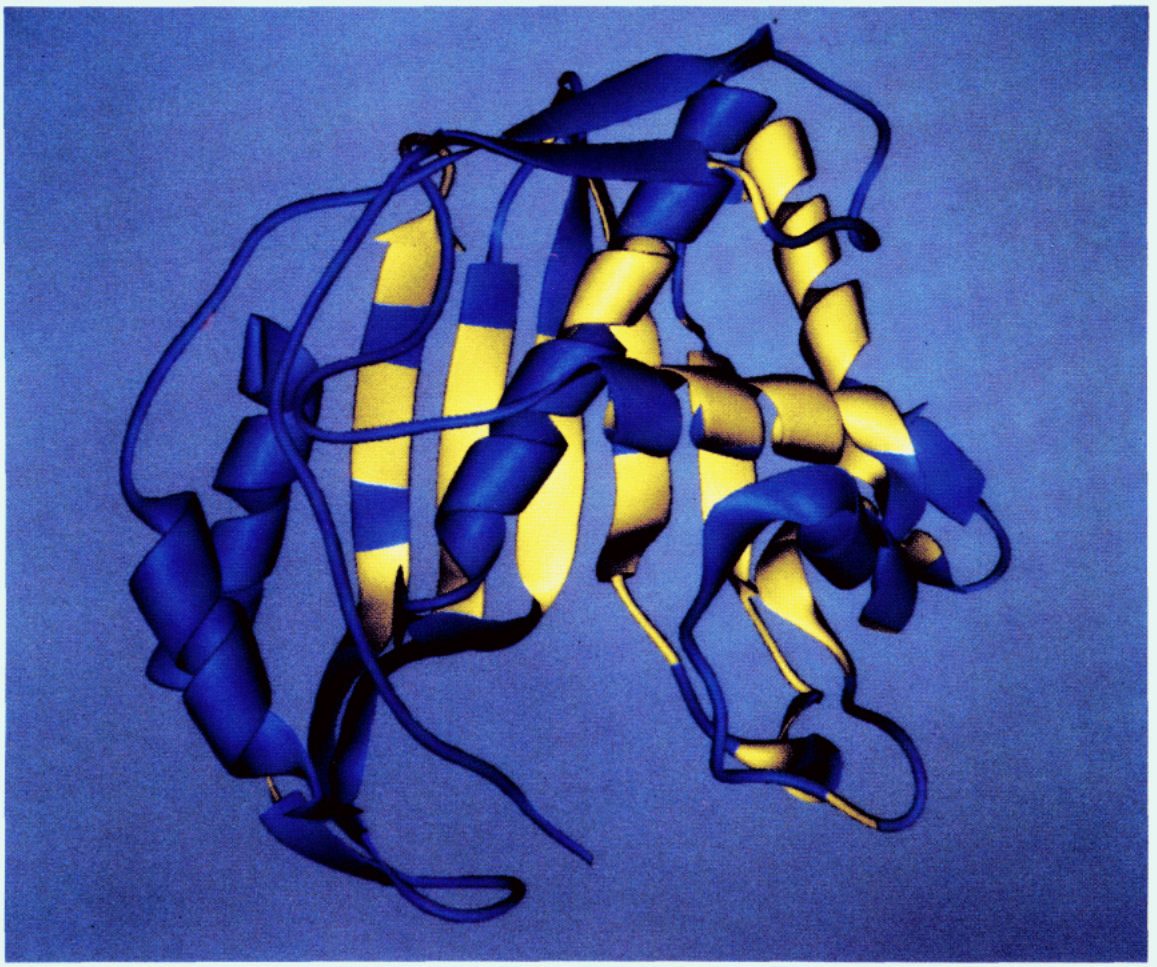

FIGURE 14: Residues contained in the constant core identified for the comparison of unliganded and liganded structures of *E. coli* TS. It was important that the difference distance matrices were applied to the entire dimer. The main segmental accommodation is seen in the coordinate movement of segments on the left toward the active center upon ligation.

ligands in TS·dUMP·CB3717. Many of the shift vectors point in the same direction, toward the active site. To illustrate this more clearly, vectors from 19 stretches of residues were vectorially summed and plotted on the α-carbon tracing in the lower half of Figure 15. The magnitudes of these vectors, listed in Table III, are substantially larger than would be obtained from random addition of the shift vectors. Regions that shift are not domains in the usual sense but segments of secondary structure which move as units.

The 21-residue J helix, which runs through the center of the protein, is displaced. This helix contains a slight bend, lines one side of the active site cleft, and contacts the inhibitor at several points. The interaction with inhibitor results in a slight rotation of the helix and an increase in the helix bend, moving the amino (bottom in Figure 15) end about 0.5 Å toward the active site. Moving along with the J helix are the three β-strands below and left of the β-kink as seen in Figure 15.

The β-kink acts as one of several "hinge" points in the structure, about which other units in the protein rotate; the further from the hinge, the greater the displacement. Displacements range from about 0.3 Å near the hinge to 1.1 Å at the farthest residue from the active site, residue 25(23), in the loop that contains phosphate-binding Arg-23(21). Residues attached to or packed with this portion of the β-sheet move along with it. For instance, the K helix, on the outside of the protein (lower left in Figure 15), is displaced about 1.0 Å upon ligand binding.

Table III: Segmental Shifts of Secondary Structural Elements

	residues[a]	vector sum (Å)[b]	95% confidence[c]	shift/atom (Å)
1	1–12	5.3	3.1	0.44
2	18–20	2.9	3.0	0.97
3	21–28	8.6	9.0	1.07
4	42–52	8.2	4.5	0.74
5	69–80	6.0	3.4	0.50
6	81–88	3.7	2.2	0.46
7	100–110	7.0	3.8	0.63
8	143–148	2.0	1.5	0.33
9	152–161	3.9	3.1	0.39
10	167–179	8.7	4.2	0.67
11	180–185	2.0	1.4	0.34
12	186–194	3.2	2.4	0.36
13	197–200	1.3	1.1	0.32
14	204–209	4.0	2.9	0.67
15	210–219	13.4	7.2	1.34
16	220–229	8.2	4.4	0.82
17	230–234	2.1	2.0	0.42
18	242–248	3.2	2.4	0.46
19	250–255	4.0	2.9	0.67
20	256–261	6.7	5.2	1.12
21	262–264	15.1	13.5	5.03

[a] Residues used for summation of shift vectors. [b] Magnitude of the sum of the vectors relating α-carbons of TS–P_i to TS–tern after overlapping, for the residues indicated, as shown in Figure 15. Values for monomer I and monomer II are averaged. [c] 95% of all additions of the same vectors in random orientation are below this value.

On the right side of the molecule (Figure 15), the displacements are smaller. The β-bulge containing Cys-198(146),

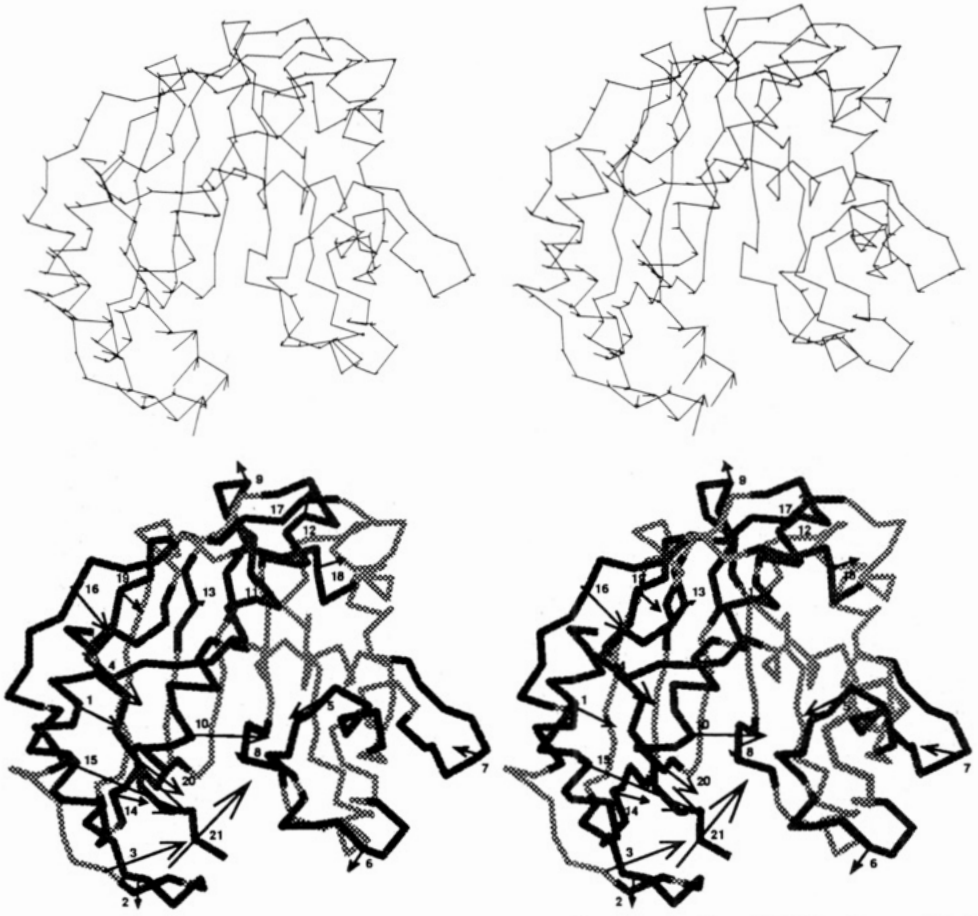

FIGURE 15: (Top) Stereoscopic view of the individual shift vectors seen upon ligation of unliganded TS. (Bottom) Indicated are the sums of vector shifts within segments of several residues within coherent regions of secondary structure elements.

which initiates catalysis by nucleophilic attack of dUMP, is moved further into the active site. The angle between the B and C helices is altered slightly and the loop connecting helix C to helix G (lower right) has adjusted. This loop contains Trp-85(83), which hydrogen bonds to the C-terminus in TS·dUMP·CB3717 but not in unliganded TS. The G and H helices lead into the other active site and are displaced somewhat in that direction. The loop connecting the H helix and β-strand vi contains Arg-178′(126′) and -179′(127′) which interact with the phosphate of dUMP. These arginines are located slightly further into the active site.

The largest shift between the two proteins involves the six residues at the carboxy terminus. The terminal region moves 4.0 Å further into and partly covers the active site. The carboxyl forms a hydrogen bond with Trp-85(83) and with Arg-23(21) that in turn provides two hydrogen bonds to phosphate oxygens. The C-terminal valine has been implicated in TS function since its removal by carboxypeptidase leads to complete inactivation (Aull et al., 1974). The structure suggests that it is the position of the charged carboxyl group rather than the nature of the terminal side group that is important. The side chain lies in a hydrophobic region. Mutagenesis of Val-316 in *L. casei* TS tends to confirm this since it can be replaced by small hydrophobic or neutral side chains, while mutations to glycine or charged side chains generally inactivate the protein (Climie et al., unpublished results).

It is possible that some of the displacements discussed above were due to crystal packing. We have discussed only those movements found in both monomers to avoid this bias. The structures of these proteins in different crystal forms would further address this source of error. We have also solved structures of unliganded TS from three other species, the best refined of which is for *L. casei*. (R = 21% at 2.3-Å resolution.) It contains a 50-residue insert between the C and G helices and has only 60% sequence agreement in the remainder of the protein. These differences make a comparison between it and the ternary complex structures more difficult, since the ligand-induced displacements are of the same magnitude as those from sequence divergence (Perry et al., 1990). Nevertheless, in comparing the TS ternary complex and *L. casei* structures, we found the general result of active site closure to hold, with many of the same residues involved.

We describe a ligand-induced segmental accommodation that takes place in thymidylate synthase. Although this is the first time to our knowledge that a conformational change of this sort has been seen, or described in detail, it probably represents a general mechanism for proteins to adjust upon ligand binding or covalent modification. Smaller local changes are seen in most proteins, as described for example in the comparison of diisopropyl phosphoryl inhibited trypsin and benzamidine trypsin (Krieger et al., 1974), or seen upon activation of trypsinogen to trypsin (Kossiakoff et al., 1977). Lesk and Chothia (1984) have demonstrated that the domain closure seen in citrate synthase and other enzymes is the cumulative effect of small shifts and rotations of secondary structure similar to the adjustments which occur in TS on ligand binding.

Segmental accommodation in TS may be related to apparent half of the sites phenomena. Also, Aull et al. (1974) showed that treatment of TS with carboxypeptidase inactivates the entire dimeric enzyme after cleavage of only one chain, implying that the effect of one cleavage is relayed to the other active site 30 Å away. No functional advantage to segmental accommodation is yet apparent. However, TS is among the

most conserved enzymes of any (Perry et al., 1990), and roles for many of the conserved regions are yet to emerge. Interactions with other proteins in the "replitase" complex and the finding that TS in all parasites for which there are sequence data (17 at this date) are bifunctional enzymes linked to dihydrofolate reductase in the same gene suggest that interactions with DHFR and other proteins may take advantage of the segmental motion that accompanies binding of ligands to TS and may also help to account for the unusually high conservation of sequences in TS.

ACKNOWLEDGMENTS

We are grateful to Jo Davisson and Dan Santi for useful discussions and Julie Newdoll for preparing Figures 3 and 5.

REFERENCES

Anderson, C. M., Stenkamp, R. E., McDonald, R. C., & Steitz, T. A. (1978) *J. Mol. Biol. 123*, 207–219.

Aull, J. L., Loeble, R. B., & Dunlap, R. B. (1974) *J. Biol. Chem. 249*, 1167–1172.

Belfort, M., Maley, G., & Maley, F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4914–4918.

Bennett, W. S., Jr., & Steitz, T. A. (1980) *J. Mol. Biol. 140*, 211–230.

Blakley, R. L., & McDougall, B. M. (1962) *J. Biol. Chem. 237*, 812–818.

Blum, M., Metcalf, P., Harrison, S. C., & Wiley, D. C. (1987) *J. Appl. Crystallogr. 20*, 235–242.

Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.

Cantwell, B. M. J., Macaulay, V., Harri, A. L., Kaye, S. B., Smith, I. E., Milsted, R. A. V., & Calvert, A. H. (1988) *Eur. J. Cancer Clin. Oncol. 24*, 733–736.

Chambers, J. L., & Stroud, R. M. (1979) *Acta Crystallogr. B35*, 1861–1874.

Cleland, W. W. (1967) *Adv. Enzymol. Relat. Areas Mol. Biol. 29*, 1–32.

Danenberg, P. V., & Danenberg, K. D. (1978) *Biochemistry 17*, 4018–4024.

Danenberg, K. D., & Danenberg, P. V. (1979) *J. Biol. Chem. 254*, 4345–4348.

Danenberg, P. V., Langenbach, R. J., & Heidelberger, C. (1974) *Biochemistry 13*, 926–933.

Dev, I. K., Yates, B. B., Atashi, J., & Dallas, W. S. (1989) *J. Biol. Chem. 264*, 19132–19137.

Dickerson, R. E., Kendrew, J. C., & Strandberg, B. E. (1961) *Acta Crystallogr. 14*, 1188–1195.

Donato, H., Jr., Aull, J. L., Lyon, J. A., Reinsch, J. W., & Dunlap, R. B. (1976) *J. Biol. Chem. 251*, 1303–1310.

Finer-Moore, J. S., Montfort, W. M., & Stroud, R. M. (1990) *Biochemistry* (following paper in this issue).

Franks, A. (1955) *Proc. Phys. Soc., London, Sect. B 68*, 1054.

Fujinaga, M., Gros, P., & van Gunsteren, W. F. (1989) *J. Appl. Crystallogr. 22*, 1–8.

Furey, W. J. (1984) in *Methods and Applications in Crystallographic Computing* (Hall, S. R., & Ashinda, T., Eds.) pp 353–371, Clarendon, Oxford, U.K.

Galivan, J. H., Maley, G. F., & Maley, F. (1975) *Biochemistry 14*, 3338–3344.

Hardy, L. W., Finer-Moore, J. S., Montfort, W. R., Jones, M. O., Santi, D. V., & Stroud, R. M. (1987) *Science 235*, 448–455.

Heidelberger, C., Chaudhari, N., Danenberg, P., Mooren, D., & Griesback, L. (1957) *Nature 179*, 663–666.

Hendrickson, W. A., & Lattman, E. E. (1970) *Acta Crystallogr. B26*, 136–143.

Hendrickson, W. A., & Konnert, J. H. (1981) in *Biomolecular Structure, Conformation, Function and Evolution* (Srinivasan, R., Subramanian, E., & Yathindra, N., Eds.) Vol. 1, pp 43–57, Pergamon, Oxford.

Holy, A., & Votruba, I. (1974) *Collect. Czech. Chem. Commun. 39*, 1646–1661.

Howard, A. J., Nielsen, C., & Xuong, Ng. H. (1985) *Methods Enzymol. 114*, 452–472.

Huber, R., & Bennett, W. S., Jr. (1983) *Biopolymers 22*, 261–279.

Hughes, D. E., Shonekan, O. A., & Simpson, L. (1989) *Mol. Biochem. Parisitol. 34*, 155–166.

Jackson, R. C., Jackman, A. L., & Calvert, A. H. (1983) *Biochem. Pharmacol. 32*, 3783–3790.

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

Jones, T. R., Calvert, A. H., Jackman, A. L., Brown, S. J., Jones, M., & Harrap, K. R. (1981) *Eur. J. Cancer 17*, 11–19.

Kalman, T. I., Bloch, A., Szekeres, G. L., & Bardos, T. J. (1973) *Biochem. Biophys. Res. Commun. 55*, 210–217.

Kenny, E., Atkinson, T., & Hartley, B. S. (1985) *Gene 34*, 335–342.

Kossiakoff, A. A., Chambers, J. L., Kay, L. M., & Stroud, R. M. (1977) *Biochemistry 16*, 654–664.

Krieger, M., Kay, L. M., & Stroud, R. M. (1974) *J. Mol. Biol. 83*, 209–230.

Leary, R. P., Beaudette, N., & Kisliuk, R. L. (1975) *J. Biol. Chem. 250*, 4864–4868.

Lesk, A. M., & Chothia, C. (1984) *J. Mol. Biol. 174*, 175–191.

Lewis, C. A., & Dunlap, R. B. (1981) *Top. Mol. Pharmacol. 1*, 169–217.

Lockshin, A., & Danenberg, P. V. (1980) *Biochemistry 19*, 4244–4251.

Maley, G., & Maley, F. (1988) *J. Biol. Chem. 263*, 7620–7627.

Matthews, D. A., Appelt, K., & Oatley, S. J. (1989) *J. Mol. Biol. 205*, 449–454.

Maxwell, J. C. (1860) *Philos. Mag. 19*, 19–31.

Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Montfort, W. R., Maley, G. F., Maley, F., & Stroud, R. M. (1990) *Proteins* (in press).

Pogolotti, A. L., Jr., & Santi, D. V. (1977) *Bioorg. Chem. 1*, 277–311.

Pogolotti, A. L., Jr., Danenberg, P. V., & Santi, D. V. (1986) *J. Med. Chem. 29*, 478–482.

Rayleigh, O. M. (1919) *Philos. Mag. 37*, 321–347.

Remington, S., Wiegand, G., & Huber, R. (1982) *J. Mol. Biol. 158*, 111–152.

Reyes, P., & Heidelberger, C. (1965) *Mol. Pharmacol. 1*, 14–30.

Rice, D. W. (1981) *Acta Crystallogr. A37*, 491–500.

Santi, D. V., & Danenberg, P. V. (1984) in *Folates and Pterins*, Vol. 1, Chemistry and Biochemistry of Folates (Blakley, R. L., & Benkovic, S. J., Eds.) pp 345–398, Wiley, New York.

Santi, D. V., McHenry, C. S., & Sommer, H. (1974) *Biochemistry 13*, 471–481.

Sessa, C., Zucchetti, M., Ginier, M., Willems, Y., D'Incalci, M., & Cavalli, F. (1988) *Eur. J. Cancer Clin. Oncol. 24*, 769–775.

Sharma, R. K., & Kisliuk, R. L. (1973) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 32*, 591 (Abstract).

Sim, G. A. (1959) *Acta Crystallogr. 12*, 813–815.

Sussman, J. L. (1985) *Methods Enzymol. 115*, 271–303.
Wahba, A. J., & Friedkin, M. (1961) *J. Biol. Chem. 236*, PC11–PC12.
Wahba, A. J., & Friedkin, M. (1962) *J. Biol. Chem. 237*, 3794–3801.
Wang, B.-C. (1985) *Methods Enzymol. 115*, 90–112.
Wohlrab, F., Haertle, T., Trichtinger, T., & Guschlbauer, W. (1978) *Nucleic Acids Res. 5*, 4753–4759.

Pairwise Specificity and Sequential Binding in Enzyme Catalysis: Thymidylate Synthase†

Janet S. Finer-Moore, William R. Montfort, and Robert M. Stroud*
Department of Biochemistry and Biophysics, S-960, University of California in San Francisco, San Francisco, California 94143-0448

Received November 20, 1989; Revised Manuscript Received March 22, 1990

ABSTRACT: The structures of thymidylate synthase (TS) from *Escherichia coli*, in ternary complexes with substrate and an analogue of the cofactor, are the basis of a stereochemical model for a key reaction intermediate in the catalyzed reaction. This model is used to compare the reaction chemistry and chirality of the transferred methyl group with structures of the components, to identify those residues that participate, and to propose a stereochemical mechanism for catalysis by TS. Effects of chemical modification of specific amino acid residues and site-directed mutations of residues are correlated with structure and effects on enzyme mechanism. The ordered binding sequence of substrate deoxyuridine monophosphate and methylenetetrahydrofolate can be understood from the structure, where each forms a large part of the binding site for the other. The catalytic site serves to orient the reactants, which are sequestered along with many water molecules within a cavernous active center. Conformational changes during the reaction could involve nearby residues in ways that are not obvious in this complex.

Thymidylate synthase (TS)[1] catalyzes methyl transfer from the cofactor 5,10-methylenetetrahydrofolate (CH_2-H_4folate) to 2′-deoxyuridine 5′-monophosphate (dUMP) to give 2′-deoxythymidine 5′-monophosphate (dTMP) and dihydrofolate (H_2folate). The cofactor CH_2-H4folate is normally polyglutamylated by up to seven glutamyl residues. Here we relate what is known of the mechanism of this multistep reaction to the high-resolution structures of TS that we solved first for the *Lactobacillus casei* enzyme (Hardy et al., 1987) and later for complexes formed between the *Escherichia coli* enzyme, substrate, and cofactor analogues (Perry et al., 1990; Montfort et al., 1990). These are three of six highly refined TS structures solved in our laboratory. In the *E. coli* pair there are no disordered residues; thus, we have very high confidence in the total correctness of the structures.

A proposed chemical mechanism, consistent with the protein structure for TS, is shown in Figure 1 (Pogolotti & Santi, 1977; Heidelberger et al., 1983). The cofactor CH_2-H_4folate plays a dual role. First, it is a one-carbon donor, and then it is a reductant of the transferred methylene at different steps in the reaction (Humphries & Greenberg, 1958; Friedkin, 1959). It is regenerated stereospecifically by dihydrofolate reductase and serine hydroxymethyltransferase in the thymidylate synthase cycle. Because of the increased needs for thymine in dividing cells, TS, with its essential role in DNA synthesis, is a target for anticancer drug design.

A major factor in understanding catalysis by TS has been the discovery of analogues of reaction intermediates. The stable, reversible ternary complex of TS, CH_2-H_4folate, and the 5-fluoro derivative of the substrate dUMP (FdUMP) has been particularly useful (Santi & McHenry, 1972). This complex is an analogue of the steady-state intermediate II of Figure 1, differing only by substitution of a fluorine for a hydrogen atom at C-5 of dUMP (Santi & Danenberg, 1984; Moore et al., 1986). C-6 of FdUMP is covalently bound to Cys-198 (Pogolotti et al., 1976; Bellisario et al., 1979). Cys-198 lies in a wide cavity on the protein surface, surrounded by conserved hydrogen-bonding residues. Reaction of this thiol with C-6 of dUMP activates C-5 for condensation with cofactor.

The stable TS·FdUMP·CH_2-H_4folate complex provides a means of determining kinetics of the first half of the reaction in detail since the mechanism of formation of this analogue complex is thought to be very similar to the mechanism of formation of II (Santi et al., 1987). FdUMP and monoglutamated folate bind in an ordered, sequential manner (Lorenson et al., 1967; Danenberg & Danenberg, 1978). TS is an obligate dimer, with components of each monomer contributing to each of two active sites. Some evidence suggests that the two active sites in TS, which are identical with one another in all four structures we determined for unliganded TS species, also react sequentially (Danenberg & Danenberg, 1979). This result implies an asymmetry during the binding and during reaction of substrates with enzyme.

The stereochemistry of the stable ternary TS·FdUMP·CH_2-H_4folate complex has been characterized in great detail. The stability of the covalently bound cofactor (Pogolotti et al., 1976) and the sensitivity of the C-9–N-10 bond of the cofactor to cleavage by specific chemical reagents (Pellino &

†Supported by National Institutes of Health Grants RO1-CA-41323 to J.S.F.-M. and R.M.S. and GM24485 to R.M.S. and by a postdoctoral fellowship from the American Chemical Society (S-49-87 to W.R.M.).
* Author to whom correspondence should be addressed.

[1] Abbreviations: TS, thymidylate synthase; dUMP, 2′-deoxyuridine 5′-monophosphate; CH_2-H_4folate, 5,10-methylenetetrahydrofolate; dTMP, 2′-deoxythymidine 5′-monophosphate; H_2folate, dihydrofolate; CB3717, 10-propargyl-5,8-dideazafolate; FdUMP, 5-fluoro-2′-deoxyuridine 5′-monophosphate.

0006-2960/90/0429-6977$02.50/0