Lessons and Conclusions from Dissecting the Mechanism of a Bisubstrate Enzyme: Thymidylate Synthase Mutagenesis, Function, and Structure †

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 2003

Impact Factor: 3.02 · DOI: 10.1021/bi020599a · Source: PubMed

3 AUTHORS:

Janet Finer-Moore
University of California, San Francisco
102 PUBLICATIONS **4,765** CITATIONS
SEE PROFILE

Daniel Santi
prolynx
77 PUBLICATIONS **2,563** CITATIONS
SEE PROFILE

Robert Michael Stroud
University of California, San Francisco
337 PUBLICATIONS **18,709** CITATIONS
SEE PROFILE

Lessons and Conclusions from Dissecting the Mechanism of a Bisubstrate Enzyme: Thymidylate Synthase Mutagenesis, Function, and Structure[†]

Janet S. Finer-Moore, Daniel V. Santi, and Robert M. Stroud*

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

Received September 24, 2002; Revised Manuscript Received October 15, 2002

Thymidylate synthase (TS)[1] has been the subject of a concerted effort to relate catalytic function to dynamics and orientation using mutagenesis, chemistry, kinetics, and structure. It may be that few enzymes will ever be subjected to a more detailed dissection than has been achieved here. Therefore, we learn general lessons about the pathway to understanding this specific enzyme that are likely to be of broad interest.

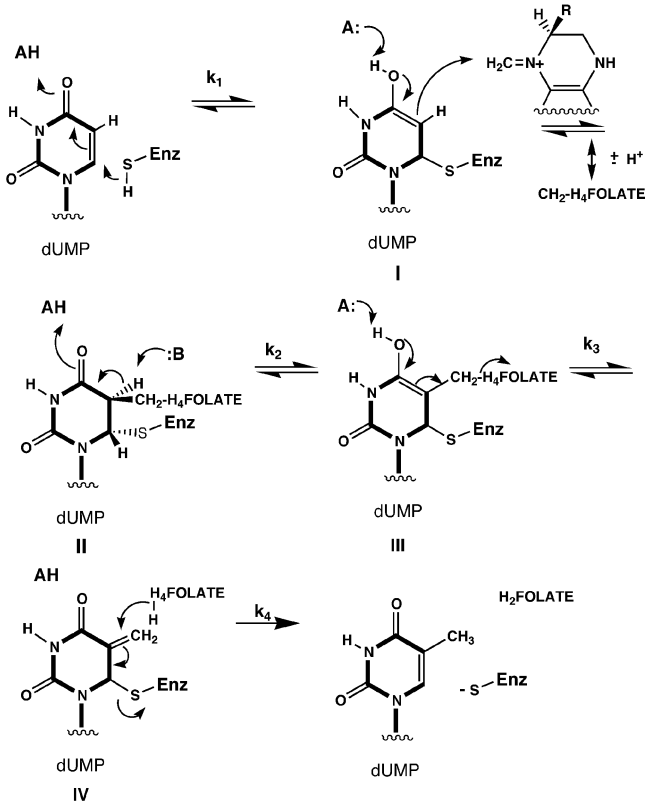
Thymidylate synthase is an essential enzyme for the *de novo* synthesis of dTMP, a nucleotide required for DNA synthesis, and therefore, it is an important drug target. It is a bisubstrate enzyme in which the cofactor, CH_2H_4folate, serves as both the one-carbon donor to the substrate dUMP and, subsequently, as the reductant. A one-carbon unit and hydride are transferred from different sites on the cofactor, and large contributions to the catalytic power of TS are achieved through the dynamic alignment and realignment of reactants at these chemical steps.

The correct alignment of the reactants is necessary but not sufficient for chemical catalysis in TS. It is clear from studies of chemical model systems that proton transfer between a general acid and dUMP aids methylene transfer to the dUMP base (steps 1−4, Scheme 1) (*1*), and a general base on the enzyme seems to be necessary for abstracting a proton from the covalent enzyme−substrate−cofactor intermediate (step 3, Scheme 1). A general acid has also been postulated to convert the cofactor into a reactive iminium ion by protonating C10 (*2*). A major challenge to understanding the function of conserved amino acids in TS is to identify additional roles for a subset of these residues in acid−base catalysis, or electrostatic or hydrogen bonding stabilization of the five presumed transition states in the reaction.

Mutagenesis Tests the Roles of Residues in Catalysis

The steps in the TS reaction, shown in Scheme 1, were deduced from chemical model systems, kinetic measurements, and spectroscopic data before the three-dimensional structure of the enzyme had been determined (*1*) The chemical steps have subsequently been validated by structures of reaction intermediate analogues with mechanism-based inhibitors (*2*−*6*) and by site-directed mutagenesis that isolates or traps specific reaction intermediates (*7*, *8*).

Scheme 1: Chemical Mechanism of TS



The X-ray structures dramatically altered the thinking about the roles of residues in TS (*2*, *9*, *10*). By the time the three-dimensional structure of thymidylate synthase was determined (*9*), sequences of TS from several species were known, providing a database of naturally occurring variants. Many of the conserved residues line the active site cavity, and several were potential substrate binding residues, or candidates for the general base required in step 3 of the reaction (Scheme 1). Each of these residues was mutated individually to all possible alternatives to probe its function. This was accomplished by construction of a synthetic *Lactobacillus casei* TS gene that allowed DNA cassettes containing mixtures of 32 codons (encoding all 20 amino acids and the AMBER stop codon) to be constructed at any particular amino acid position (*11*). With this synthetic gene,

[†] Supported by National Institutes of Health Grant CA-41323.

* To whom correspondence should be addressed.

[1] Abbreviations: TS, thymidylate synthase; dUMP, 2′-deoxyuridine 5′-monophosphate; CH_2H_4folate or mTHF, 5,10-methylene-5,6,7,8-tetrahydrofolate; dTMP, 2′-deoxythymidine 5′-monophosphate; H_2folate or DHF, 7,8-dihydrofolate; H_4folate or THF, (6*S*)-5,6,7,8-tetrahydrofolate; FdUMP, 5-fluoro-dUMP; CB3717, 10-propargyl-5,8-dideazafolate; PABA, *p*-aminobenzoic acid; LcTS, *L. casei* TS; EcTS, *E. coli* TS.

10.1021/bi020599a CCC: $25.00
Published on Web 12/17/2002

Table 1: Some Assays for Steps in the TS Reaction

reaction step	assays	most informative structure
dUMP binding	(1) k_{cat}/K_m^{dUMP}, at saturating mTHF (2) rate of fluorescence quenching by different concentrations of dUMP	TS-dUMP
mTHF binding	(1) pre-steady state kinetics of formation of the TS·FdUMP·mTHF complex (2) pre-steady state absorbance spectroscopy for TS plus substrates	noncovalent ternary complexes
Cys Sγ attack at C6 of dUMP	debromination of 5-bromo-dUMP in the absence of cofactor	TS−dUMP or TS−BrdUMP
abstraction of dUMP H-C5 from intermediate II	rate of tritium release from [2-^{14}C,5-^{3}H]dUMP during the enzyme reaction	TS·FdUMP·mTHF
hydride transfer	partitioning of intermediate II between dTMP and HETM-dUMP as a function of mTHF concentration	TS−dUMP−THF

"replacement sets" of up to 19 different variants for any given residue could be efficiently generated. High-level expression of these variants in a TS-deficient strain of *Escherichia coli* immediately defined substitutions that complemented the deficiency, thus allowing for rapid "ballpark" definition of activity of each mutant relative to ~1% of wild-type activity (*12*). The catalytic efficiencies of subsets of the several individual substitutions of each of the residues, selected on the basis of the complementation results, were then measured by steady state kinetics.

It must be stressed that mutation of a residue will reveal the upper bound of its effect on the wild-type enzyme reaction. Therefore, to evaluate the contribution of a functional group to reaction rates, the substitution that has the least impact on function may be the most informative. The rationale for saturation mutagenesis is that the functionally most conservative amino acid substitution at any particular site cannot be predicted from a wild-type structure, since misleading structural rearrangements, such as changes to solvent structure, often occur in response to any mutation.

High-Resolution Assays of Function

Understanding the effects of mutating a residue requires a more encompassing view of the reaction energetics than that provided by steady state kinetic analyses alone. Ideally, the view would include reaction rates for each chemical step and crystal structures of the corresponding reaction intermediates. In reality, effects of mutation on the rate constants for the steps in TS can only be estimated by analogy to partial reactions that are models for a particular step, and actual reaction intermediates in TS are not stable enough to be isolated in crystals. One or more specific assays are used to probe the energetic consequences of mutation on the individual reaction steps shown in Scheme 1, and the results are complemented by crystal structures of stable mimetics of reaction intermediates (Table 1). The approaches used to monitor discrete steps in the enzyme reaction are described below.

TS interacts with its ligands in an ordered fashion, with dUMP binding before CH_2H_4folate (*13*). In an ordered BiBi kinetic mechanism, k_{cat}/K_m^{dUMP}, measured under conditions of saturating cofactor, is equal to the on-rate for dUMP, k_a (*14*), which has also been measured by fluorescence quenching (*15*). TS binary complexes with dUMP readily crystallize for most mutants. Thus, steady state kinetics and crystallography provide an entrée to a description of the first step of the reaction, i.e., dUMP binding.

$K_m^{CH_2H_4folate}$, on the other hand, is an apparent binding constant that cannot be used to derive a rate for cofactor binding. Cofactor binding, in fact, involves several steps, including initial association with the binary complex, k_b, active site closure, opening of the imidazolidine ring of the cofactor, and covalent bond formation between the resulting iminium ion and C5 of dUMP (*15*, *16*). The rate for the transition from a loosely associated ternary complex to a tight covalent complex (k_1, Scheme 1) encompasses the latter three steps. The covalent TS·FdUMP·CH_2H_4folate complex is a close analogue of intermediate II in Scheme 1 and is likely to be formed by a very similar, if not identical, mechanism. This analogue provides a tool for defining the details of cofactor binding (*16*). Rate and equilibrium constants as well as $\Delta G^{\ddagger}$, $\Delta H^{\ddagger}$, and $\Delta S^{\ddagger}$ for formation of this complex are thought to parallel the rate and equilibrium constants for formation of intermediate II, and have been determined by a combination of steady state kinetics and pre-steady state kinetics at different temperatures (*16*). A similar analysis of the C-terminal deletion mutation V316Am has been used to measure effects of the mutation on cofactor binding and activation (*17*). Kinetic parameters for cofactor binding can also be calculated from pre-steady state absorbance spectroscopy (*15*).

Nucleophilic attack of Cys-198 on C6 of dUMP occurs simultaneously with tight binding of the cofactor, and a rate for Michael addition cannot be separated from the rate of transition to the covalent ternary complex. However, the Michael addition at C6 can be monitored by a partial reaction that does not require cofactor binding or protein conformational change. TS catalyzes debromination of 5-bromo-dUMP in the absence of cofactor. The mechanism for this reaction involves nucleophilic addition of the catalytic cysteine to C6 of 5-bromo-dUMP, analogous to the Michael addition step of the enzymatic reaction; therefore, k_{cat}/K_m for this reaction monitors the catalytic function of Cys-198 (*18*). The dehalogenation reaction is very sensitive to nucleotide misorientation and disorder, so is also a test for perturbations to the dUMP binding site (*19*).

Following Michael addition to form covalent intermediate II, the enzyme abstracts a proton from C5 of dUMP. The rate for this step in the reaction, (k_2, Scheme 1) can be measured by using [2-^{14}C,5-^{3}H]dUMP as the substrate and monitoring the decrease in the ratio of ^{3}H to ^{14}C (*17*). Although hydrogen exchange from dUMP C5 also occurs in TS−dUMP binary complexes, the rate of this off-pathway reaction is 45000-fold lower than that of dTMP formation (*20*), so tritium release is a good assay of k_2.

β-Elimination of the cofactor from intermediate III was detected as a separate step in the reaction using mutants of Glu-60, though the rate for this step (k_3, Scheme 1) is not

accessible with conventional kinetics. For these mutants, a covalent ternary complex intermediate accumulated on gels, showing that the rate-determining step came at or before breakdown of the complex (*7*). Since the rate of dTMP formation was slower than the rate of tritium release from [2-^{14}C,5-^{3}H]dUMP ($\sim k_2$), β-elimination of cofactor, not proton abstraction from intermediate II, was rate-limiting (*7*, *21*), and the covalent complex that accumulated on gels was presumably intermediate III in Scheme 1.

Finally, hydride transfer from the cofactor to the exocyclic methylene group at C5 of the pyrimidine in intermediate IV (Scheme 1) can be assayed by monitoring the formation of a side product, 5-(2-hydroxyethyl)thiomethyl-dUMP (HETM-dUMP). In mutants for which hydride transfer is much slower than in the wild-type enzyme, intermediate IV can be trapped with β-ME to give HETM-dUMP (*8*). The partitioning of intermediate IV between HETM-dUMP and dTMP as a function of cofactor concentration has been used to calculate rate constants for the transition from intermediate IV to dTMP (k_4, Scheme 1) and to HETM-dUMP (k_5) (*22*). These rate constants are a means of evaluating in a quantitative way the efficiency of hydride transfer in mutants.

The Five "Most Essential" Residues for Catalysis

Saturation mutagenesis of almost all of the 25 conserved residues surrounding the active site cavity, followed by complementation assays of *L. casei* TS (*12*) or *E. coli* TS (*23*), in TS$^-$ *E. coli* showed the initially surprising result that most of the residues were not essential for activity; that is, substitution with one or more alternate amino acids allowed the reaction to proceed. Conservation of amino acids across species implies conserved roles that may, and indeed in TS are found to, partition among the following general roles: covalent catalysis, stabilization of transition states, orientation of substrates (or reorientation at a later stage in the reaction), stabilization of catalytic residues or water, determining selectivity for substrate and against alternative ligands, and maintenance of the structure or dynamic trajectory of the protein. While any of these roles may be required for the catalytic event, several residues may contribute to each, and often any single residue can be substituted with other amino acids.

We assayed more than 500 LcTS active site mutants by complementation, measured steady state kinetic constants for 146 of these, and determined 52 crystal structures of apo or ligand-bound forms of EcTS or LcTS mutants specifically to address mechanism. Other investigators also used complementation or steady state kinetics to assay active site mutants in *E. coli* (115 mutants) (*23*−*25*) and human TS (26 mutants) (*26*, *27*). From these investigations, five residues proved to be critically important at different steps of the reaction. These residues are required for dUMP binding (Arg-218), cofactor binding (Asp-221), Michael addition to dUMP C6 (Cys-198), proton abstraction from intermediate II (Tyr-146), and cofactor elimination from the covalent ternary complex (Glu-60).

Arg-218 is one of four arginines that bind to the phosphate moiety of dUMP. Hydrogen bonds between the guanidinium group and backbone atoms of two loops at the active site are important scaffolding that maintains the structure of the active site. The crystal structure of LcTS R218K shows the active site is disordered and some of the residues may be rearranged (*28*). Arg-218 is adjacent to the active site cysteine, and we postulated that it activates the cysteine through electrostatic interactions (*9*). This role in catalysis has never been proven; however, the only variant of Arg-218 in *L. casei* TS with a k_{cat} close to that of the wild-type enzyme (~30-fold lower) is the lysine variant, a residue with electrostatic properties similar to those of arginine (*28*). R218K has a K_m^{dUMP} ~100-fold higher than for wild-type TS, and it does not complement TS deficiency. Nevertheless, the high k_{cat} of R218K compared with other variants is consistent with the proposal that Arg-218 plays a role in activating Cys-198.

Asp-221 is the only residue whose side chain hydrogen bonds directly to the pterin ring of the cofactor. D221 variants that cannot form this hydrogen bond either do not form a covalent ternary complex or, in the case of D221A, form the ternary complex with a 200-fold higher $K_d^{CH_2H_4folate}$ (*29*). Crystallography has shown that the hydrogen bond is critical for excluding nonproductive cofactor binding modes in the ternary complex. In the structure of D221N, for example, where the hydrogen bond acceptor, Asp, is replaced with the donor, Asn, the pterin ring of the cofactor binds in a flipped conformation and cannot condense with dUMP (*30*). Matthews et al. (*2*) and Chiericatti and Santi (*29*) have proposed that Asp-221 is a general acid that catalyzes protonation of N5 of the cofactor during breakdown of intermediate II.

Nucleophilic attack of Cys-198 Sγ at dUMP C6 catalyzes condensation of the cofactor methylene group to the C5 position of the uracil ring (Scheme 1) (*31*, *32*). Variants of Cys-198 are inactive or, in the case of EcTS C198(146)S,[2] 5000-fold less active (*33*). The crystal structure of the C198S·dUMP·CH_2H_4folate substrate ternary complex shows both dUMP and CH_2H_4folate bound at their wild-type binding sites, but with no covalent bonds between dUMP and Ser198 Oγ, or the N5 methylene of CH_2H_4folate, confirming that the Michael addition step is impaired by the mutation (unpublished).

Tyr-146 and Glu-60 are the two other residues in TS, besides Cys-198, that are required for chemical steps in the reaction. Both of these residues are conserved, but are not in direct contact with the substrate or cofactor. Crystal structures of the inactive variants Y146F and E60Q show that both can form steady state intermediates close in structure to that formed by wild-type TS (*34*, *35*). Mutations of these residues lead to greatly diminished k_{cat}s and accumulation of ternary complex reaction intermediates that point to the step that requires the residue (*7*, *36*). Tyr-146 appears to be essential for abstracting the proton from C5 of dUMP in intermediate II, either by itself or acting through water (*36*). Glu-60 is hydrogen bonded through water to O4 of dUMP and thus can act to transfer a proton to and from the pyrimidine ring in as many as four of the steps in the enzyme reaction, and is essential for β-elimination of H_4-folate from covalent intermediate III (Scheme 1) (*7*, *21*).

In addition to these five completely conserved residues, a sixth conserved residue, Asn-229, is not required for catalytic

[2] The numbering scheme of the *L. casei* enzyme is used. When a different species, such as *E. coli* TS, is discussed, its numbering is given in parentheses.

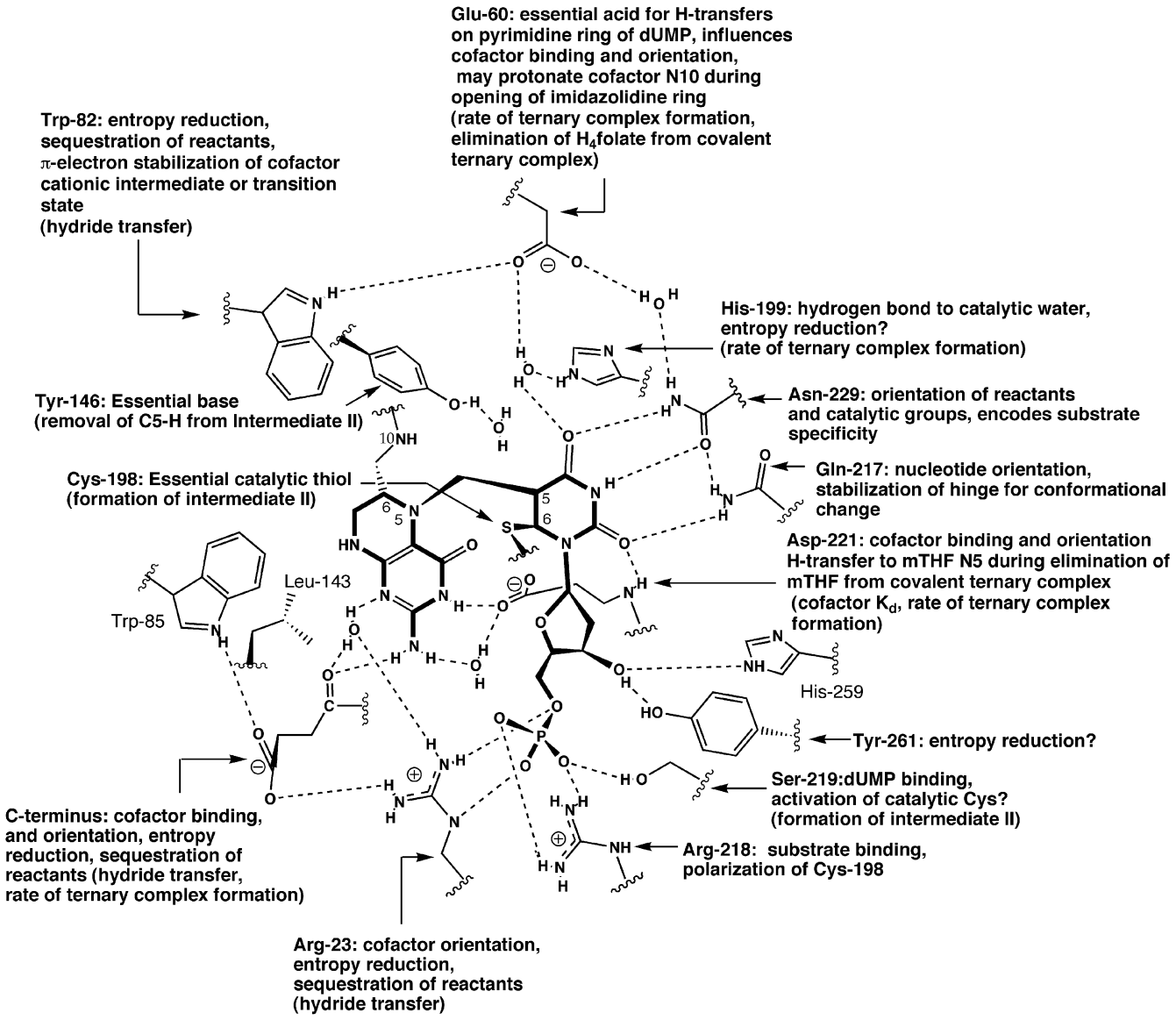


FIGURE 1: Diagram of the active site showing residues involved in substrate or cofactor binding. Hydrogen bonds are depicted with dashed lines. The two phosphate-binding arginines from the other protomer and residues that bind the PABA−Glu moiety of the cofactor are not shown. Five residues (E60, Y146, C198, R218, and D221) are critically important at different times during the reaction, as indicated in the annotations. Residues that are not essential, but are thought to contribute to catalysis through orientation of reactants, are also annotated. If mutants of the residue are known to impair a particular step in the reaction, that step is noted in parentheses.

chemistry, but has a critical role in substrate selectivity, particularly in excluding dCMP from the active site (*37*, *38*). In wild-type *L. casei* TS, K_d^{dCMP} is 400-fold higher than K_d^{dUMP} and the enzyme does not produce methylated dCMP. However, the mutation of Asn-229 to an Asp allows the enzyme to bind and methylate dCMP (*39*, *40*).

Since the first analysis of the structure of TS (*9*), the sequences of TS from many other species have been determined. Many of the residues that were totally conserved in the original set of available sequences are now seen to have conservative substitutions in one or more known sequences. This is consistent with the tolerance for mutation observed in our saturation mutagenesis experiments. When a very divergent TS from the archeon *Methanococcus jannaschii* (*41*) is included in the alignment of TS sequences, only six to eight residues (depending on the sequence alignment of *M. jannaschii* TS) are seen to be absolutely conserved, and these include those identified by saturation mutagenesis as having irreplaceable roles in catalysis, namely, Glu-60, Tyr-146, Cys-198, and Asp-221. Arg-218 is a Lys in the *M. jannashii* gene; thus, its role could still be assigned to activating the catalytic nucleophile by electrostatic interactions. Also conserved are Thr-48, which plays a role in maintaining the protein fold, and Arg-23, Arg-178, and Arg-179, the three other arginines besides Arg-218 that bind the dUMP phosphate.

Substrate Binding Residues and Their Contributions to Catalysis in TS

Most of the residues in TS that contribute to catalysis also have a role in substrate or cofactor binding (Figure 1); therefore, any mutation may potentially impair a particular reaction step by altering the substrate or cofactor orientation. For this reason, it is essential that we rationalize the effects of a mutation in the context of the most appropriate structural intermediates (Table 1).

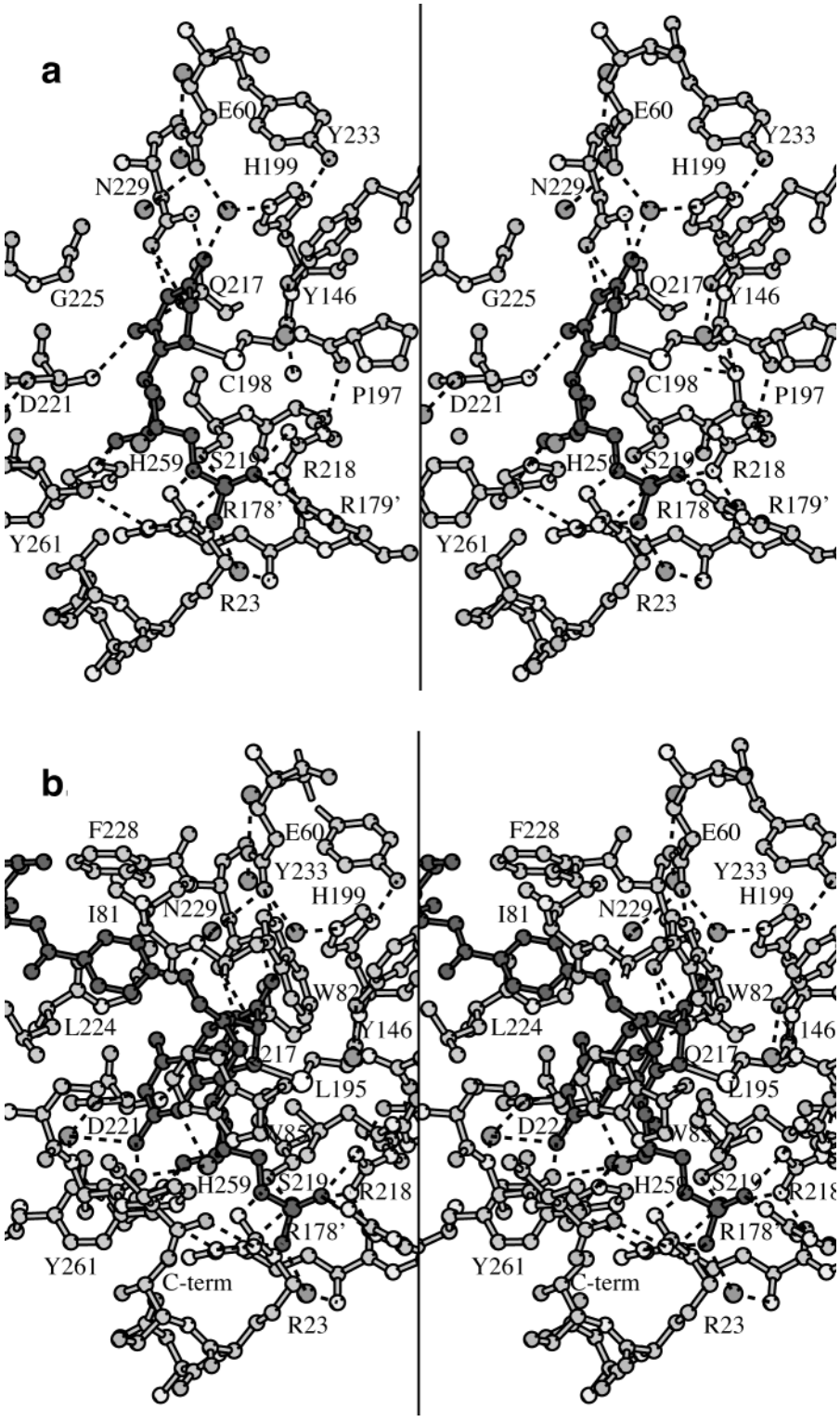


FIGURE 2: (a) Stereoplot of the dUMP binding site showing hydrogen bonds with dashed lines. dUMP is plotted in dark gray. Plot generated with MOLSCRIPT (*53*). (b) Same view as in panel a but showing CH_2H_4folate as well as dUMP in dark gray. Coordinates derived from PDB entry 1TSN.

dUMP binds to a preformed binding site on apo-TS (Figure 2a). Electrostatic interactions of the dUMP phosphate moiety with a conserved four-arginine binding site of TS provide much of the substrate binding affinity (*42*). Each protomer of the homodimer contributes two conserved arginines to the phosphate binding site. The arginines are located on noncontiguous regions in the sequence, and their guanidinium groups come together in an approximately tetrahedral arrangement around the phosphate moiety, where each donates two hydrogen bonds to the phosphate oxygens. In addition to these hydrogen bonds, a conserved serine residue is hydrogen bonded to one of the phosphate oxygens.

Charge Balance and Cooperativity in Phosphate Binding. The charged dUMP phosphate−arginine interactions are finely tuned to balance favorable electrostatic and hydrogen bonding energy against the energetic costs of (i) desolvation of both phosphate and arginines, (ii) the entropic cost of ordering mobile arginine side chains, and (iii) the cost of

imposing electrostatic repulsion between adjacent arginines. Partly as a result of the balance of these energetic forces, changes in the free energy of activation calculated from the change in the rate of dUMP binding ($k_a = k_{cat}/K_m^{dUMP}$ at saturating cofactor concentrations for an ordered BiBi kinetic mechanism) for single and double mutants of the four arginines are not additive (*43*). For example, the increases in the free energy of activation of dUMP binding compared to that of wild-type TS $\{\Delta\Delta G_{s\dagger} = -RT \ln[(k_{cat}/K_m)_{mut}/(k_{cat}/K_m)_{wt}]\}$ for the *L. casei* TS variants R179T, R23I, and R23I/R179T were 1, 3, and 7 kcal/mol, respectively, and there were no structural rearrangements in the mutants that would account for the nonadditive effects (*43*). Thus, the energetic cost from loss of one arginine is small, while the cost of losing a second is much greater, regardless of the order in which the arginines are removed. Cooperative binding of the arginines to the phosphate moiety, where dUMP hydrogen bonding to one arginine fixes the phosphate moiety in an optimal orientation for accepting hydrogen bonds from the others, may also contribute to these nonadditive effects (*43*).

Asn-229 Encodes Substrate Specificity. Interactions between the protein and the pyrimidine ring of dUMP are the primary determinants of the orientation of dUMP in TS. This was demonstrated using crystal structures of TS bound to chemical moieties that are "fragments" of dUMP. The binding sites for the pyrimidine rings of 2′-deoxyuridine, 2′,5′-dideoxyuridine, and uridine are the same as they are in the substrate, dUMP, showing that the phosphoribose moiety is not required for positioning the pyrimidine base for catalysis (*44*). Oδ1 and Nδ2 of the Asn-229 carboxamide group accept and donate, respectively, a hydrogen bond from N3 and to O4 of the pyrimidine ring. Asn-229 can be mutated to several other small, hydrophilic residues with negligible loss of dUMP binding affinity since water, or other side chains, can replace these hydrogen bonds (*37*).

The side chains of His-259 and Tyr-261 both form hydrogen bonds to the dUMP hydroxyl, O3′, and these are the only protein hydrogen bonds to the ribose ring. These hydrogen bonds are less important for orienting the substrate than the hydrogen bonds with Asn-229. They are not, for example, sufficient to position the ribose ring in the dUMP binding site in the absence of the pyrimidine ring (*44*). In isolation, the phosphoribose fragment of dUMP binds to *E. coli* TS in a disordered fashion, with the phosphate group anchored in the phosphate-binding site, but the ribose not tethered by hydrogen bonds to either His-259 or Tyr-261 (*44*). Tyr-261 can be mutated to residues as diverse as Trp or Ala with no change in the dUMP orientation.

The TS reaction rate is not very sensitive to poor ordering of dUMP in the binary TS−dUMP complex because subsequent binding of the cofactor is very effective in orienting and ordering dUMP, and the disorder in the active site of the binary dUMP complex is not propagated to the active sites of ternary complexes (*25*). Several mutants of key dUMP binding residues Asn-229 and Arg-178 have greatly elevated K_m^{dUMP}s, but k_{cat}s that are approximately equal to that of the wild-type enzyme (*28*, *45*). Crystal structures of dUMP binary complexes of these mutants show dUMP is disordered, with its equilibrium position displaced from the wild-type binding site (*19*, *43*). $K_m^{CH_2H_4folate}$ is also higher for these variants since the CH_2H_4folate binding site, formed in large part by the dUMP surface, is disordered. However, k_{cat} is not affected, suggesting that order and orientation are re-established upon formation of the ternary complex.

On the other hand, for some variants of Asn-229 that bind dUMP in a well-ordered but misoriented manner, misorientation persists after the cofactor binds, and side chains and ordered waters that may be involved in catalysis are also in different configurations. For these variants, K_ms for dUMP and CH_2H_4folate are similar to those of wild-type TS, but k_{cat} is diminished more than 100-fold (*45*, *46*).

In an experiment in which the sensitivity of k_{cat} to a "tuned" range of substrate misorientations was measured, the specificity of LcTS was changed from the normal substrate, dUMP, to dCMP with successive mutations that allowed dCMP to move progressively into the dUMP binding site and, at the same time, forced dUMP to bind slightly further from the site (*47*). First, Asn-229 was changed to Asp, with hydrogen bonding properties more compatible with dCMP than dUMP. dCMP made the anticipated hydrogen bonds to Asp-229, but collision of its 4-NH_2 substituent with His-199 caused it to bind with its pyrimidine ring tilted ∼20° away from the dUMP base in wild-type TS, and C6 ∼0.5 Å further from the catalytic sulfhydryl group. The k_{cat}/K_m for methylation of dCMP by N229D increased from 0 to 0.003 $s^{-1}\,\mu M^{-1}$. dUMP bound to N229D with its pyrimidine tilted such that O4 was ∼1 Å out of its binding site, thus eliminating hydrogen bonds to this carbonyl (*38*). Elimination of hydrogen bonds to O4, which are important for catalysis, resulted in a 10^{-4}-fold decrease in k_{cat}/K_m for methylation of dUMP.

Mutation of His-199 in N229D to Ala allowed dCMP to bind analogously to dUMP in the wild-type enzyme. The combination of an orientation adjustment to the pyrimidine ring and a 0.5 Å shift of dCMP C6 toward Cys-198 Sγ led to a 4-fold increase in k_{cat} and a 97-fold increase in k_{cat}/K_m^{dCMP} compared to those of N229D. The second mutation increased the disorder of dUMP binding, which reduced k_{cat}/K_m^{dUMP} by an additional 1 order of magnitude. Thus, the experiment in tuning the alignment of reacting groups suggested merely a 0.5 Å misalignment causes a 4-fold drop in k_{cat}.

Cofactor Binding Triggers a Protein Conformational Change that Creates Its Ultimate Binding Site. Cofactor binding triggers a shift of several protein segments toward the active site which surrounds the reactants and sequesters them from bulk solvent. In addition, several residues adopt different side chain rotamers to maximize hydrophobic contacts with the cofactor. Thus, the cofactor (or cofactor analogue) binding site in the covalent complex is fully formed only after extensive conformational change. The binding site consists of (i) a large hydrophobic surface area provided by dUMP, (ii) a number of hydrophobic amino acid side chains in van der Waals contact with the PABA or pterin rings, (iii) positively charged side chains that contribute electrostatic binding energy or water-mediated hydrogen bonds to the polyglutamyl moiety, and (iv) a hydrogen bond network to the pterin ring comprised of a main chain carbonyl, the C-terminal carboxyl, two ordered waters, and the side chain of Asp-221. A single hydrogen bond in the network, donated by CH_2H_4folate N3-H to Asp-221, is present when the enzyme is in an open conformation (*48*, *49*); the remaining hydrogen bonds are formed once the four

C-terminal residues and Arg-23 move into the active site cavity during closure of the enzyme (Figures 1 and 2b).

The k_{cat} decreases at least 10-fold when cofactor-binding residues are mutated (*28*, *29*, *50*), while it is less consistently affected by mutagenesis of the dUMP-binding residues (*28*, *45*). The hydride transfer step of the TS reaction, which is rate-determining in the wild-type enzyme (*15*), appears to be particularly sensitive to the orientation of reactants; mutants of three cofactor-binding residues (Arg-23, Trp-82, and Ile-316) reduce the hydride transfer rate (k_4, Scheme 1) (*22*). Since neither Arg-23 nor Ile-316 is close to the methyl or hydride transfer sites, it is most likely that their role in orienting the cofactor is important for hydride transfer.

These results suggested that variants of cofactor-binding residues could provide a more sensitive "ruler" of the proper orientation and proximity effects on catalysis than the binary complex structures of the variants of dUMP-binding residues. Ideally, structures of ternary complexes of TS mutants with dUMP and a cofactor analogue such as CB3717 would reveal different degrees of ligand misalignment that would correlate with the magnitudes of the decreases in reaction rates (k_4 or k_{cat}). With the aim of mapping the dependence of the catalytic rate on the orientation of reactants, we determined the structures of EcTS W82(80)G and V316(264)Am (in which the C-terminal residue was not encoded) in ternary complexes with dUMP and CB3717. In addition, we determined the structure of the dUMP−CB3717 ternary complex of EcTS D221(169)C. D221(169)C is the only catalytically active variant in either LcTS or EcTS of the important cofactor binding residue D221(169). Although it readily forms the covalent steady state intermediate, k_{cat} is significantly lower, suggesting that ligand alignment is not optimal in ternary complexes.

All three experiments gave similar results that did not fit the expected pattern. Each mutant crystallized with one or more dimers in the crystal asymmetric unit; thus, independent structures of at least two active sites were represented in each crystal structure. In each case, the dimers were asymmetric, with at least one active site being more open than in wild-type complexes. W82G ternary complex crystals had three dimers in the asymmetric unit, and the six protomers in the crystal structure adopted a spectrum of protein conformations (*51*). The protomers with more open conformations had more disordered, and in many cases misoriented, ligands. However, each crystal structure also had a protomer in the fully closed conformation with the ligands aligned as in the wild-type ternary complex (*51*).

Clearly, reactants could achieve the optimum transition state structure in all cases, and a loss of catalytic ability was less related to perturbation of the transition state structures than to increased disorder in the steady state intermediates. If the crystal structures are representative of the ternary complexes in solution, then compared to wild-type TS, the mutant complexes assume a broader range of protein conformations. These results imply that the closed form of the enzyme required for forming covalent ternary complexes, such as steady state intermediate II in Scheme 1, is less stable than in wild-type TS. This conclusion is corroborated by kinetics studies of the rates of formation and dissociation of the covalent TS·FdUMP·CH_2H_4folate complex (an analogue of intermediate II) in the *L. casei* TS C-terminal mutant, V316(264)Am (*17*). The equilibrium constant for the conversion of the reversible, loosely bound complex of V316(264)-Am with FdUMP and CH_2H_4folate to the covalent (closed) complex was 2.6×10^3-fold higher than in wild-type LcTS, corresponding to an ~4.5 kcal increase in the free energy of the TS·FdUMP·CH_2H_4folate complex. The advantage of the wild-type protein, therefore, is that it highly favors an ordered structure for the closed ternary complex structure in which the reactants are well positioned for the chemistry that needs to occur.

Mutants of cofactor-binding residues have revealed connections between cofactor binding, protein dynamics, and catalysis. TS relies on a conserved binding mechanism, in which binding interactions with the cofactor pay the energetic cost of closing the enzyme, for aligning the reactants, and sequestering them from bulk solvent. The protein may contribute more to catalysis by reducing the entropy in the ground state intermediates, and biasing the distribution of ternary complex conformations toward structures resembling the transition states, than by stabilization of the transition states.

Gln-217: Priming the Active Site for Cofactor Binding. Gln-217 is a highly conserved residue in van der Waals contact with the dUMP pyrimidine ring, which indirectly affects the nucleotide orientation through its key structural role in donating hydrogen bonds to Asn-229 Oδ1 and to the backbone carbonyl of Ser-219. Ser-219 lies in the β-kink that acts as a hinge during the conformational change between the open and closed enzyme (*5*). Thus, Gln-217 has two important roles: it orients the carboxamide of the primary pyrimidine-binding residue Asn-229, and it stabilizes the structure of the active site cavity.

The roles played by the conserved hydrogen bonds and hydrophobic contacts formed by Gln-217 are best evaluated by the conservative mutants Q217A, -G, -H, -N, and -S. These residues do not sterically interfere with the structure or introduce new charge, but cannot make all of the same favorable hydrogen bonding contacts with surrounding residues. These mutants in human TS have only slightly reduced k_{cat}s and approximately 10-fold higher K_m^{dUMP}s. More revealing is the 1.5−5-fold reductions in the rate of ternary complex formation exhibited by these mutants, which correlate inversely with the size of the substituted side chain (*27*). Presumably, the smaller side chains permit greater mobility of the surrounding residues; therefore, these rate reductions indicate the degree to which preordering and orientation of active site residues (and, indirectly, of dUMP) by Gln-217 contribute to an important step in ternary complex formation, cofactor orientation and closure of the enzyme.

Tyr-261 Contributes to k_{cat} by Stabilizing the Closed Enzyme. TS is exceedingly sensitive to mutation of Tyr-261. Only two variants of Tyr-261 complement TS$^-$ *E. coli*, and only the most conservative variant, Y261F, has the same k_{cat} as wild-type TS; k_{cat} for the second complementing variant, Y261M, is 17-fold lower than the wild-type value (*12*). The most obvious role for Tyr-261 is in dUMP binding since it forms one of only two hydrogen bonds to the ribose ring of dUMP. Yet mutation of this residue to amino acids as diverse as Ala and Trp has small effects on dUMP orientation and K_m^{dUMP} (3-fold increase for Y261M, 19-fold increase for Y261W) that do not explain the more than 100-fold decreases in k_{cat} for these variants. The crystal structures unexpectedly

revealed that even though this residue does not contact the cofactor directly, it might contribute to k_{cat} by preferentially stabilizing the closed conformation of the enzyme. Tyr-261 is buried in the closed but not the open enzyme conformation, and the degrees to which the substituted hydrophobic side chains filled the volume of the Tyr phenyl ring in the Y261 variants correlated with k_{cat}. Presumably, hydrophobic contacts made by Y261 only after the active site closes favor the closed conformation of the ternary complex, in which reactants are aligned for catalysis. Tyr-261 also positions two residues that are themselves directly involved in cofactor orientation. Substitutions of Y261 with smaller hydrophobic side chains alter the orientation of Asn-221 and increase the mobility of the Arg-23-containing loop. Both Asp-221 and Arg-23 contribute to the hydrogen bond network with the cofactor pterin ring in closed ternary complexes.

Conformational Change Closes the Glu-60−H_2O−dUMP Circuit Used in Catalysis. When TS undergoes the cofactor-induced conformational change, one or more ordered waters are expelled from a pocket above the pyrimidine ring and a water-mediated hydrogen bond forms between dUMP O4 and Glu-60 Oϵ1 (Figure 1). The conserved water in this link between Glu-60 and dUMP is proposed to function as a general acid in several steps in the enzyme reaction, including the Michael addition of dUMP to the catalytic cysteine, as shown in Scheme 1 (*7*, *35*). The hydrogen bonds connecting Glu-60 and dUMP are part of a hydrogen bond network in ternary complexes of TS that also involves His-199 and Asn-229. Although the side chains in this hydrogen bond network do not bind the cofactor, conservative mutations of these residues decrease k_1, the rate of formation of the covalent ternary complex, as well as k_{cat} (*24*, *45*, *52*). His-199 and Asn-229 may modulate the orientation and pK_a of the ordered water, thus fine-tuning its function as a general acid during nucleophilic attack of the catalytic cysteine at C6 of dUMP. The participation of these residues in the hydrogen bond network also may enhance the rate of ternary complex formation by stabilizing the closed form of the enzyme and fixing the pyrimidine ring of dUMP in the optimal orientation for alignment with the cofactor.

Summary

TS contains many highly conserved residues whose function we and others have investigated by site-directed mutagenesis, steady state kinetics, assays of chemical steps, and structure. These investigations identified three acid−base or covalent catalysts for different steps in the reaction. Cys-198 is the nucleophile required for the Michael addition step; Glu-60 catalyzes hydrogen transfers to and from the dUMP pyrimidine ring during many of the chemical steps, and Tyr-146 abstracts a proton from C5 of dUMP to catalyze the breakdown of the covalent ternary complex.

Most of the other conserved residues in the enzyme contribute to catalysis mainly through their role in driving a conformational change that closes the enzyme active site, thereby fixing the reactants in the optimal orientations for the chemical reactions and organizing the solvent structure at the active site. Thus, not one of this subset of residues is essential in the sense that it cannot be mutated without abrogating activity, yet all enhance the catalytic rate by limiting nonproductive conformations of ternary complex intermediates in the reaction. TS relies on this mechanism of entropy reduction to efficiently traverse a complicated reaction coordinate that has several steady state intermediates (*15*).

The thymidylate synthase mechanism emphasizes the strength of combining mutagenesis, kinetics, partial reactions and other probes of chemical steps, and multiple structures of key intermediates in defining the factors harnessed for enzyme catalysis. Several general lessons are learned with this approach. First, the structures of mutant complexes are essential to understanding the basis for the maximal factor contributed by the wild-type residue to k_{cat}/K_m. Second, 0.5 Å changes to the median orientations of reactants can cause a 4-fold reduction in the catalytic rate. Third, a residue's effects on catalysis can include direct and indirect effects in stabilizing different reaction intermediate states. Fourth, any individual mutagenesis experiment in an enzyme can be misleading and, at best, gives an incomplete understanding of the role of the residue in catalysis. Mutation may lead to a great rate reduction for any of many different reasons, including long-range effects on structure and electrostatics. Crystallography, while it can define atomic position, amplitudes of atomic vibration, and the trajectory boundaries of isolable states through the reaction, does not yet adequately provide the basis for describing protein dynamics, which is key to function. We can possibly compensate for these limitations by using molecular mechanics to assist in the interpretation of the results but also by the vast screen that is derived from the saturation mutagenesis approach. Thus, the results of many experiments, not one of which is definitive, combine to define residue roles and yield reliable insights into the origins of enzymatic catalysis.

From the perspective of evolution within the heterogeneous cellular milieu, it is remarkable that fewer than 20 amino acid side chains, along with amide bonds, an amino group, and a carboxyl group, in the context of a three-dimensional fold, can accomplish such selective catalysis. In general, enzymes use subtle contributions to catalysis to achieve reaction rates far higher than those that can be achieved for the same reaction in the test tube, in vitro, or even by artificial selection pressure applied to develop proteins that catalyze the same chemistry.

ACKNOWLEDGMENT

We thank Paola Costi for critically reading the manuscript.

REFERENCES

1. Pogolotti, A. L., Jr., and Santi, D. V. (1977) in *Bioorganic Chemistry* (van Tamelan, E. E., Ed.) pp 277, Academic, New York.
2. Matthews, D. A., Villafranca, J. E., Janson, C. A., Smith, W. W., Welsh, K., and Freer, S. (1990) *J. Mol. Biol. 214*, 937−948.
3. Danenberg, P. V., Langenbach, R. J., and Heidelberger, C. (1974) *Biochemistry 13*, 926−933.
4. Santi, D. V., McHenry, C. S., and Sommer, H. (1974) *Biochemistry 13*, 471−481.
5. Montfort, W. R., Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Maley, G. F., Hardy, L., Maley, F., and Stroud, R. M. (1990) *Biochemistry 29*, 6964−6977.
6. Matthews, D. A., Appelt, K., Oatley, S. J., and Xuong, N. H. (1990) *J. Mol. Biol. 214*, 923−936.
7. Huang, W., and Santi, D. V. (1994) *J. Biol. Chem. 269*, 31327−31329.
8. Barrett, J. E., Maltby, D. A., Santi, D. V., and Schultz, P. G. (1998) *J. Am. Chem. Soc. 120*, 449−450.

9. Hardy, L. W., Finer-Moore, J. S., Montfort, W. R., Jones, M. O., Santi, D. V., and Stroud, R. M. (1987) *Science 235*, 448−455.
10. Finer-Moore, J. S., Montfort, W. R., and Stroud, R. M. (1990) *Biochemistry 29*, 6977−6986.
11. Climie, S., and Santi, D. V. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 633−637.
12. Climie, S., Ruiz-Perez, L., Gonzalez-Pacanowska, D., Prapunwattana, P., Cho, S. W., Stroud, R., and Santi, D. V. (1990) *J. Biol. Chem. 265*, 18776−18779.
13. Santi, D. V., and Danenberg, P. V. (1984) in *Folates & Pterins: Vol. 1, Chemistry and Biology of Folates* (Blakley, R. L., and Benkovic, S. J., Eds.) pp 345−398, John Wiley and Sons, New York.
14. Santi, D. V., Pinter, K., Kealey, J., and Davisson, V. J. (1990) *J. Biol. Chem. 265*, 6770−6775.
15. Spencer, H. T., Villafranca, J. E., and Appleman, J. R. (1997) *Biochemistry 36*, 4212−4222.
16. Santi, D. V., McHenry, C. S., Raines, R. T., and Ivanetich, K. M. (1987) *Biochemistry 26*, 8606−8613.
17. Carreras, C. W., Climie, S. C., and Santi, D. V. (1992) *Biochemistry 31*, 6038−6044.
18. Garret, C., Wataya, Y., and Santi, D. V. (1979) *Biochemistry 18*, 2798−2804.
19. Finer-Moore, J., Liu, L., Birdsall, D. L., Brem, R., Apfeld, J., Santi, D. V., and Stroud, R. M. (1998) *J. Mol. Biol. 276*, 113−129.
20. Pogolotti, A. L. J., Weill, C., and Santi, D. V. (1979) *Biochemistry 18*, 2794−2798.
21. Hardy, L. W., Graves, K. L., and Nalivaika, E. (1995) *Biochemistry 34*, 8422−8432.
22. Variath, P., Liu, Y., Lee, T. T., Stroud, R. M., and Santi, D. V. (2000) *Biochemistry 39*, 2429−2435.
23. Michaels, M. L., Kim, C. W., Matthews, D. A., and Miller, J. H. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3957−3961.
24. Dev, I. K., Yates, B. B., Atashi, J., and Dallas, W. S. (1989) *J. Biol. Chem. 264*, 19132−19137.
25. Strop, P., Changchien, L., Maley, F., and Montfort, W. R. (1997) *Protein Sci. 6*, 2504−2511.
26. Williams, A. W., Dunlap, R. B., and Berger, S. H. (1998) *Biochemistry 37*, 7096−7102.
27. Steadman, D. J., Spencer, H. T., Dunlap, R. B., and Berger, S. H. (1999) *Biochemistry 38*, 5582−5587.
28. Kawase, S., Cho, S. W., Rozelle, J., Stroud, R. M., Finer-Moore, J., and Santi, D. V. (2000) *Protein Eng. 13*, 557−563.
29. Chiericatti, G., and Santi, D. V. (1998) *Biochemistry 37*, 9038−9042.
30. Sage, C. R., Michelitsch, M. D., Stout, T. J., Biermann, D., Nissen, R., Finer-Moore, J., and Stroud, R. M. (1998) *Biochemistry 37*, 13893−13901.
31. Pogolotti, A. L., Jr., Ivanetich, K. M., Sommer, H., and Santi, D. V. (1976) *Biochem. Biophys. Res. Commun. 70*, 972−978.
32. Bellisario, R. L., Maley, G. F., Galivan, J. H., and Maley, F. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1848−1852.
33. Dev, I. K., Yates, B. B., Leong, J., and Dallas, W. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1472−1476.
34. Hyatt, D. C., Maley, F., and Montfort, W. R. (1997) *Biochemistry 36*, 4585−4594.
35. Sage, C. R., Rutenber, E. E., Stout, T. J., and Stroud, R. M. (1996) *Biochemistry 35*, 16270−16281.
36. Liu, Y., Barrett, J. E., Schultz, P. G., and Santi, D. V. (1999) *Biochemistry 38*, 848−852.
37. Liu, L., and Santi, D. V. (1993) *Biochemistry 32*, 9263−9267.
38. Finer-Moore, J. S., Liu, L., Schafmeister, C. E., Birdsall, D. L., Mau, T., Santi, D. V., and Stroud, R. M. (1996) *Biochemistry 35*, 5125−5136.
39. Liu, L., and Santi, D. V. (1992) *Biochemistry 31*, 5100−5104.
40. Hardy, L. W., and Nalivaika, E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9725−9729.
41. Aurora, R., and Rose, G. D. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 2818−2823.
42. Finer-Moore, J. S., Fauman, E. B., Morse, R. J., Santi, D. V., and Stroud, R. M. (1996) *Protein Eng. 9*, 69−75.
43. Morse, R. J., Kawase, S., Santi, D. V., Finer-Moore, J., and Stroud, R. M. (2000) *Biochemistry 39*, 1011−1020.
44. Stout, T. J., Sage, C. R., and Stroud, R. M. (1998) *Structure 6*, 839−848.
45. Liu, L., and Santi, D. V. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8604−8608.
46. Reyes, C. L., Sage, C. R., Rutenber, E. E., Nissen, R., Finer-Moore, J. S., and Stroud, R. M. (1998) *J. Mol. Biol. 284*, 699−712.
47. Agarwalla, S., LaPorte, S., Liu, L., Finer-Moore, J., Stroud, R. M., and Santi, D. V. (1997) *Biochemistry 36*, 15909−15917.
48. Sotelo-Mundo, R. R., Ciesla, J., Dzik, J. M., Rode, W., Maley, F., Maley, G. F., Hardy, L. W., and Montfort, W. R. (1999) *Biochemistry 38*, 1087−1094.
49. Almog, R., Waddling, C. A., Maley, F., Maley, G. F., and Van Roey, P. (2001) *Protein Sci. 10*, 988−996.
50. Kealey, J. T., Eckstein, J., and Santi, D. V. (1995) *Chem. Biol. 2*, 609−614.
51. Fritz, T. A., Liu, L., Finer-Moore, J. S., and Stroud, R. M. (2002) *Biochemistry 41*, 7021−7029.
52. Zapf, J. W., Weir, M. S., Emerick, V., Villafranca, J. E., and Dunlap, R. B. (1993) *Biochemistry 32*, 9274−9281.
53. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI020599A