Caught in Action: X-ray Structure of Thymidylate Synthase with Noncovalent Intermediate Analog

Svetlana A. Kholodar,[§,†] Janet S. Finer-Moore,[‡] Katarzyna Świderek,[#] Kemel Arafet,[#] Vicent Moliner,[#*] Robert M. Stroud,[‡] and Amnon Kohen.[§]*

[§]Department of Chemistry, The University of Iowa, Iowa City, Iowa 52242, United States. ‡Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, San Francisco 94158, United States. #Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain.

ABSTRACT Methylation of 2-deoxyuridine-5′-monophosphate (dUMP) at the C5 position by the obligate dimeric thymidylate synthase (TSase) in the sole de novo biosynthetic pathway to dTMP proceeds by forming a covalent ternary complex with dUMP and co-substrate 5,10-methylenetetrahydrofolate. The crystal structure of an analog of this intermediate gives important mechanistic insights but does not explain the half-of-the sites activity of the enzyme. Recent experiments showed that the C5 proton and the catalytic Cys are eliminated in a concerted manner from the covalent ternary complex to produce a noncovalent bisubstrate intermediate. Here we report the crystal structure of TSase with a close synthetic analog of this intermediate in which it has partially reacted with the enzyme, but in only one protomer, consistent with the half-the-sites

activity of this enzyme. Quantum mechanics/molecular mechanics simulations confirmed that the analog could undergo catalysis. The crystal structure shows a new water 2.9 Å from the critical C5 of the dUMP moiety, which, in conjunction with other residues in the network, may be the elusive general base that abstracts the C5 proton of dUMP during the reaction.

Thymidylate synthase (TSase) methylates 2-deoxyuridine-5′-monophosphate (dUMP) to produce thymidylate (dTMP), a building block of DNA. Since TSase is the sole *de novo* source of dTMP it is essential in dividing cells and hence is an anti-cancer target.[1] A minimal mechanism has been proposed for the reaction in which Michael addition of the catalytic Cys146 to C6 of the dUMP pyrimidine ring activates C5 for addition to the N5-methylene of the co-substrate 5,10-methylenetetrahydrofolate (CH_2H_4F) (Figure 1A, step 1), producing a covalent ternary complex between protein, dUMP and CH_2H_4F (this intermediate we term **Int**). In this proposed mechanism, the Cys is eliminated during the final (hydride transfer) step of the reaction (pathway A).[2] Much of the current understanding of TSase mechanism derives from its covalent ternary complex with CH_2H_4folate and the mechanism-based inhibitor 5-fluoro-dUMP (FdUMP) (Figure 1B and 1C).[2] This structure is a close analog of **Int**: the replacement of hydrogen at the critical C5 methylation site by fluorine stalls the reaction before resolution of the complex into tetrahydrofolate (H_4F) and the covalent exocyclic methylene intermediate. Despite the high similarity of the FdUMPCH_2H_4F complex to the naturally occurring **Int**, the atomic details of this structure might not represent those of the analogous complex formed during the reaction.[3] Specifically, *Ec*TSase is a half-sites active homodimer,[4] but, seemingly inconsistent with this half-the-sites activity (or simply reflecting random substitutions of half-the-sites activated dimers), the structure of the intermediate analog structure is apparently two-fold symmetric. Secondly, no specific general base (BH:) has been identified in the reaction. Rather, the whole

network of residues consisting of Tyr94, Glu58, His147 and Asn177 and water molecules connected by H-bonds was suggested to serve as a general base.[5] No crystal structure, including that of the FdUMPCH_2H_4F complex, features a continuous H-bond network between all the residues involved in the proton transfer site.

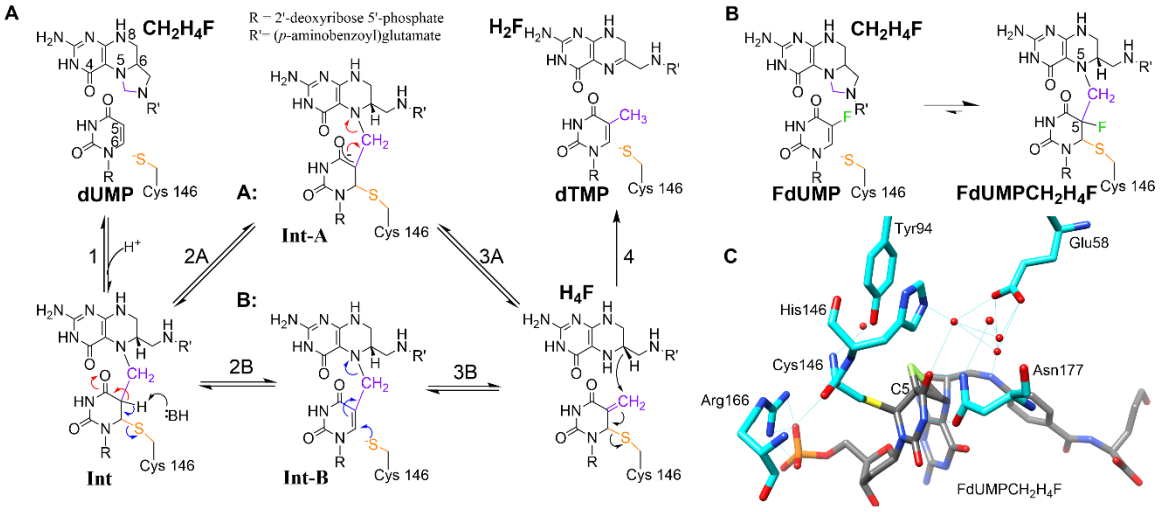


Figure 1. A. Mechanism of TSase catalyzed reaction. **B**. Inhibition of TSase by FdUMP. **C.** Structure of *Ec*TSase (PDB 1TLS) with covalently bound FdUMPCH_2H_4F (a mimic of **Int**) showing several important active site residues and water molecules. H-Bonds are drawn in cyan.

Recently, based on quantum mechanics/molecular mechanics (QM/MM) calculations, a modification of the TSase′s mechanism was proposed in which the proton at C5 of dUMP and the catalytic Cys are eliminated in a concerted manner from **Int** (Figure 1, step 2B), leaving a noncovalent bisubstrate intermediate **Int-B**.[3, 6] This proposal was supported when **Int-B** was synthesized and shown to be a chemically and kinetically competent intermediate in the TSase reaction[7]. Subsequently, the finding was reinforced by isolation of the **Int-B** formed in course of the *Ec*TSase catalyzed reaction.[8]

Previous studies reported synthesis and biological activity evaluation of the stable analogue of **Int-B**, 8-deaza-5,6,7,8-tetrahydrofolate,[9-10] a potent nanomolar competitive inhibitor of human TSase. Despite this inhibitor being the closest analog of the naturally occurring intermediate of the reaction, no study focused on the structural characterization of its complex with TSase. Such structural information would for the first time provide an atomic-level view of the reaction caught after formation of the covalent ternary complex **Int**. In order to explore this possibility, we synthesized a potential bisubstrate inhibitor of TSase: 4-amino-8-deaza-Int-B. The inhibitor was synthesized as a mixture of diastereomers (6*S*) and (6*R*), where C6 is on the folate analog moiety. Only the (6*S*)-diastereomer of CH_2H_4F (and hence of **Int-B**) is active; the 4-amino-8-deaza-Int-B corresponding to active **Int-B** is the (6*S*)-diastereomer (referred to further in the text as **4A8DZ-Int-B**). The diastereomers were separated by C18 HPLC and each was found to inhibit *Ec*TSase with IC50 of 4.7 and 17.3 μM for (6*S*)- and (6*R*)- diastereomer, respectively (Figure S1).

The two diastereomers, were each co-crystallized with *Ec*TSase, their crystal structures were refined at 1.8 Å and 1.55 Å resolution. Further discussion will focus primarily on the structure of the (6*S*) complex.

The **4A8DZ-Int-B** is clearly identifiable in difference density maps only in monomer B (Figure 2A). It closely resembles the *Ec*TSase·FdUMPCH_2H_4F structure but apparently is without a covalent bond between C6 of dUMP and the catalytic cysteine. Hydrogen bond interactions with *Ec*TSase are conserved. A unique crystallographic water molecule (W2) forming the continuous H-bond network between Tyr94, Glu58, His147 and Asn177, is observed only 2.9 Å away from the C5 of the inhibitor, making it a good candidate for the elusive general base in the TSase reaction.

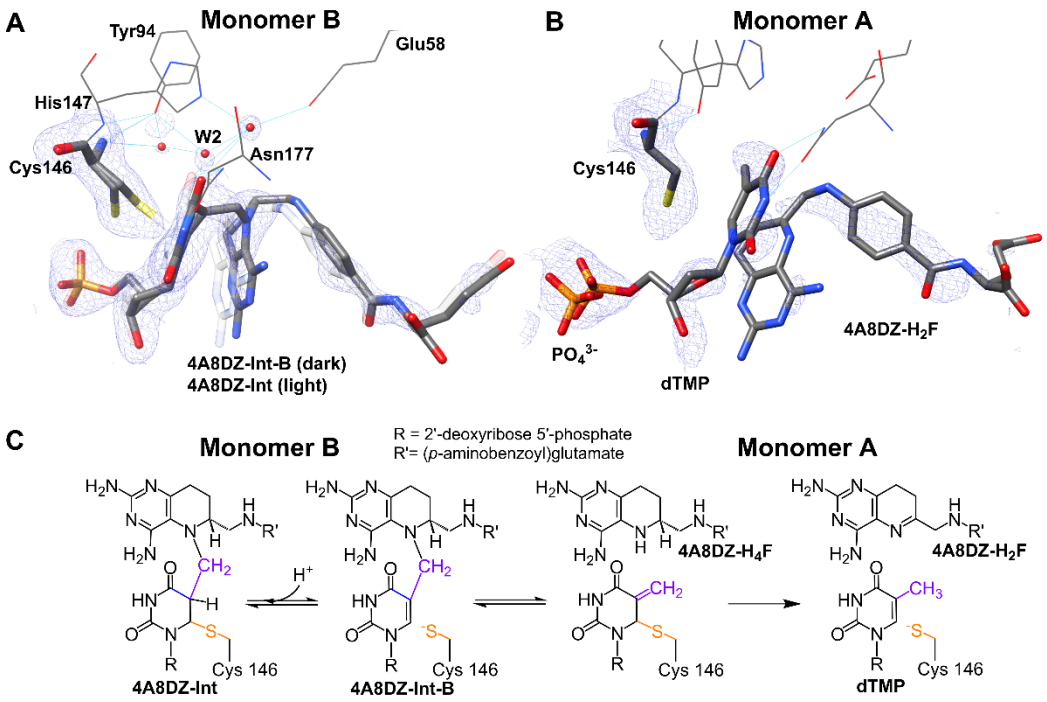


Figure 2. A. Monomer B active site with the inhibitor built as a statistical mixture of covalently (**4A8DZ-Int**, light) and non-covalently (**4A8DZ-Int-B**, dark) bound forms. The continuous H-bond network is outlined between residues Tyr94, Glu58, His147, Asn177, and several crystallographic water molecules, including the suggested general base in the reaction – W2. **B.** Monomer A active site showing the inhibitor resolved into products, 4-amino-8-deaza-H_2F and dTMP. dTMP is partially occupied; PO_4^{-2} partially occupies the dTMP density. Zoned electron density map (2mFo-DFc) is represented as an isomesh at 0.7 sigma. **C**. Crystal structure-suggested partitioning of the **4A8DZ-Int-B** in TSase active sites: equilibrium between **4A8DZ-Int-B** and its covalent C5-protonated form **4A8DZ-Int** in the Monomer B, and complete conversion of the **4A8DZ-Int-B** into products in the Monomer A.

Bridging density for the methylene group is clearly seen between the N5-methylene group of the pterin analog moiety and C6 of dUMP, but whereas dUMP C5 in **Int-B** is sp^2 hybridized, the

Int-B analog could be best fit to the density when C5 and its substituents were not coplanar, suggesting C5 was sp^3 hybridized.

In order to test the feasibility of the bisubstrate intermediate being protonated at C5, hybrid hybrid QM/MM free energy surfaces (FES) were generated with the QM region described semiempirical and DFT methods, while the protein and the solvent water molecules were described with classical force fields (see SI for details). Our selection of QM methods is in agreement with the previous findings of Kaiyawet *et al*., [11] who also pointed out the possible relevance of the QM-MM partitioning, and with our previous QM/MM benchmarking on this system.[12] The QM/MM free energy surfaces generated to explore the C6-Cys146 breaking bond without a concomitant proton transfer from the C5 carbon atom (Figure S6A) render structures unstable where other unexpected interactions are established. These involve the interaction between the active site sulfur atom of Cys146 and either N1 or C2 atoms of dUMP (Fig. S6B). In addition, the process involves activation energy barriers (ca. 35 kcal·mol^{-1}, as shown in Fig. S6A) too large to provide any competent intermediates. The exploration was carried out with **Int-B** but also with **4A8DZ-Int-B**, obtaining the same results. Subsequently, the species observed in the monomer B are consistent with a mixture of the intact **Int-B** analog and its covalent, C5-protonated form - an analog of **Int** (Figure 2A). This statistical disorder agrees with weak density for C5 and C6 of the pyrimidine, poorly resolved density for O4 and with the presence of weak density between the catalytic sulfhydryl and the pyrimidine of the inhibitor. We refined the inhibitor as a statistical mixture of **Int-B** and **Int** analogs. The C5 of dUMP in the **Int-B** analog is somewhat distorted from sp^2 geometry after refinement, which we attribute to strain imposed by the closed conformation of the active site in the crystal.

4A8DZ-Int-B in monomer A is more disordered. Bridging density between the N5-methylene group of the 4A8DZ-H_4F and C5 of dUMP is not visible in density maps, rather, there is a methyl group attached to C5 of dUMP. Density for the folate moiety of the inhibitor fits best to 4-amino-8-deaza-H_2F (4A8DZ-H_2F, Figure 2B). It therefore appears that in monomer A, **4A8DZ-Int-B** behaves as a substrate and most has been converted to dTMP and 4A8DZ-H_2F (Figure 2C). Density around some atoms in the ligands, particularly C6 of dTMP is weak or absent. The structure that best accounts for the density in the active site assumes dTMP was released and replaced by water and a phosphate ion a portion of the time. This interpretation is consistent with the disorder of the phosphate-binding loop, which shows the loop alternately in its conformation in apo-*Ec*TSase (PDB 1TJS) and in its conformation in the ternary complex.

Complete turnover of the **4A8DZ-Int-B** in the monomer A indicates that chemical alterations introduced to halt the reactivity of the **Int-B** analog were insufficient to fully prevent its conversion to the products in at least one active site under the crystallization conditions. Choice of the 4-amino and 8-deaza substitutions in the folate part of the molecule was largely dictated by an increased stability to nonenzymatic autoxidation provided by these alterations.[13] In order to determine if the enzymatic oxidation of the 4-amino-8-deaza-moiety of the **Int-B** analog is also hampered in the TSase active site we calculated the QM/MM free energy surface for the hydride transfer from C6 of the **4A8DZ-H_2F** to the exocyclic methylene intermediate (analogous to the step 4 in the Figure 1A) and found that the energy barrier for this reaction is essentially the same as for the natural reaction, within the uncertainty associated to the method (Figure S7). Since QM/MM calculations did not support that the chemical step is perturbed for the **Int-B** analog, its incomplete turnover (only in the monomer A) is more likely a result of the protein allosteric changes stalling it in the half-the-sites reactive state as discussed below.

Both the natural (6*S*) and unnatural (6*R*) diastereomers of the **Int-B** analogs were able to undergo elimination of the H_4F analog to give an exocyclic methylene intermediate (see details on (6*R*)-isomer structure in the Supporting Information). The exocyclic methylene was subsequently either reduced to yield products (in the case of the natural (6*S*) diastereomer) or undergo addition of water to give 5-hydroxymethyl-dUMP (in the case of the (6*R*) diastereomer). The products of these chemical steps were seen only in monomer A, which in both cases has the same conformation as in the covalent ternary complex analog *Ec*TSase·FdUMPCH_2H_4F.

Both of the **Int-B** analog complexes are highly symmetrical dimers, but differences in the dynamics of the two protomers may be one basis for activity of only protomer A. In both complexes the average B-factors in protomer B are approximately 9 $Å^2$ higher than in protomer A; in particular, the B-factor of Trp80, which interfaces with the inhibitor and sequesters the active site from bulk solvent,[14] is 14 $Å^2$ higher in protomer B than in protomer A. In addition to the B-factor differences, the phosphate-binding loop (residues 19-25) in protomer B is shifted ~ 1 Å out of the active site compared to the active protomer A. Thus, the closed conformation of the active site, which is important for catalysis, may be less stable in protomer B.[15]

There are also differences between the two protomers in the dynamics of the catalytic Cys146 and adjacent residues. The NMR titration experiments with the FdUMPCH_2H_4F di-ligand showed that allosteric changes accompanying di-ligand binding that may lead to half-the-sites reactivity are mostly transmitted through the β-sheet interface and involve changes to the dynamics of the two active sites.[16] Consistent with the NMR titration experiments, there are subtle differences in the strand containing active site nucleophile, Cys146 in the inactive monomer B. In monomer B of both **Int-B** analog structures, the side chain of Phe149 in this strand is in two alternate conformations, and in the **(6*R*)-Int-B** analog, Cys146 and His147 both

adopt alternate conformations (Figure S4). Disorder of these residues has also been seen in the asymmetric dimer structure of a ligand-bound *Ec*TSase mutant,[17].

The (6*S*)-diastereomer of the **Int-B** analog differs from **Int-B** only by replacement of O4 and N8 in the pterin ring by NH_2 and CH_2, respectively. Neither O4 nor N8 contact the protein in *Ec*TSase·FdUMPCH_2H_4F, so the analog would be expected to bind in the same way as **Int-B**. Thus, the co-crystal structure of *Ec*TSase with the (6*S*) inhibitor for the first time gives a glimpse of a TSase reaction intermediate formed after **Int**. The asymmetry of the structure validates the half-the-sites mechanism of *Ec*TSase and shows that half-site activity occurs before the hydride transfer step.

In summary we structurally characterized for the first time an intermediate in the TSase reaction that follows formation of the covalent ternary complex intermediate. The structure is an asymmetric dimer in which the bisubstrate analog has partially reacted with the enzyme in only one active site. The two protomers have very similar structures. Minor differences at the dimer interface and in the active site phosphate-binding loop are the main clues to the asymmetric activity. A new water 2.9 Å from C5 of the dUMP moiety in the inactive protomer is part of a network of conserved waters and residues that may jointly function to abstract the C5 proton prior to methyl transfer.

ASSOCIATED CONTENT

Supporting Information. The following file is available free of charge.
Procedures of the chemical synthesis, enzymatic assays, protein crystallography and QM/MM calculations can be found in the Supporting Information (PDF).

Accession Codes. *E.coli* TSase UniProtKB Accession ID, P0A884. Coordinates and structure factors for the EcTSase complexes with (6S)- 4A8DZ-Int-B and (6R)- 4A8DZ-Int-B are deposited in the Protein Data Bank with accession codes 7JX1 and 7JXF, respectively.

AUTHOR INFORMATION

Corresponding Author

finer@msg.ucsf.edu moliner@uji.es

Present Addresses

† Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, San Francisco 94158, United States.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by National Institutes of Health (NIH) Grant R01 GM024485 to R.M.S and NIH Grant R01 GM65368 to A.K, the Spanish Ministerio de Ciencia, Innovación y Universidades (Grant PGC2018-094852-B-C21 and PID2019-107098RJ-I00), Generalitat Valenciana (Grant AICO/2019/195, SEJI/2020/007 and APOSTD/2020/015) and Universitat Jaume I (UJI-A2019-04 and UJI-B2020-03).

ACKNOWLEDGMENT

We thank Ananda Ghosh for expression and purification of *E.coli* TSase used for the inhibition studies, and the computational resources of the Servei d'Informàtica of Universitat Jaume I.

ABBREVIATIONS

5A4DZ-H_4F, 5-amino-4-deaza H_4F; A4DZ-Int-B, 5-amino-4-deaza Int-B; CH_2H_4F, N5, N10-methylene-5,6,7,8- tetrahydrofolate; dTMP, thymidine 5′ -monophosphate ; dUMP, 2′ -deoxyuridine 5′ -monophosphate; FdUMP, 5-fluoro-2′ -deoxyuridine 5′ -monophosphate; H_4F, 5,6,7,8-tetrahydrofolate; Int, N5-thymidinyl 5′ -monophosphate enzyme-bound derivative of H_4F; Int-B, N5-thymidinyl 5′ -monophosphate derivative of H_4F; QM/MM, quantum mechanics/molecular mechanics; TSase, thymidylate synthase;

REFERENCES

1. Wilson, P. M.; Danenberg, P. V.; Johnston, P. G.; Lenz, H. J.; Ladner, R. D., Standing the test of time: targeting thymidylate biosynthesis in cancer therapy. *Nat Rev Clin Oncol* **2014,** *11* (5), 282-98.

2. Finer-Moore, J. S.; Santi, D. V.; Stroud, R. M., Lessons and conclusions from dissecting the mechanism of a bisubstrate enzyme: thymidylate synthase mutagenesis, function, and structure. *Biochemistry* **2003,** *42* (2), 248-56.

3. Hyatt, D. C.; Maley, F.; Montfort, W. R., Use of strain in a stereospecific catalytic mechanism: crystal structures of Escherichia coli thymidylate synthase bound to FdUMP and methylenetetrahydrofolate. *Biochemistry* **1997,** *36* (15), 4585-94.

4. Maley, F.; Pedersen-Lane, J.; Changchien, L., Complete restoration of activity to inactive mutants of Escherichia coli thymidylate synthase: evidence that E. coli thymidylate synthase is a half-the-sites activity enzyme. *Biochemistry* **1995,** *34* (5), 1469-74.

5. Ghosh, A. K.; Islam, Z.; Krueger, J.; Abeysinghe, T.; Kohen, A., The general base in the thymidylate synthase catalyzed proton abstraction. *Phys Chem Chem Phys* **2015,** *17* (46), 30867-75.

6. Wang, Z.; Ferrer, S.; Moliner, V.; Kohen, A., QM/MM calculations suggest a novel intermediate following the proton abstraction catalyzed by thymidylate synthase. *Biochemistry* **2013,** *52* (13), 2348-58.

7. Kholodar, S. A.; Kohen, A., Noncovalent Intermediate of Thymidylate Synthase: Fact or Fiction? *J Am Chem Soc* **2016,** *138* (26), 8056-9.

8. Kholodar, S. A.; Ghosh, A. K.; Swiderek, K.; Moliner, V.; Kohen, A., Parallel reaction pathways and noncovalent intermediates in thymidylate synthase revealed by experimental and computational tools. *Proc Natl Acad Sci U S A* **2018,** *115* (41), 10311-10314.

9. Singh, P.; Islam, Z.; Kohen, A., Examinations of the Chemical Step in Enzyme Catalysis. *Methods Enzymol* **2016,** *577*, 287-318.

10. Bayomi, S. M. M.; Brixner, D. I.; Eisa, H.; Broom, A. D.; Ueda, T.; Cheng, Y. C., Probing the thymidylate synthase active site with bisubstrte analog inhibitors. *Nucleosides and Nucleotides* **1988,** *7*, 103-115.

11. Kaiyawet, N.; Lonsdale, R.; Rungrotmongkol, T.; Mulholland, A. J.; Hannongbua, S., High-level QM/MM calculations support the concerted mechanism for Michael addition and covalent complex formation in thymidylate synthase. *J. Chem. Theory Comput*. **2015,** *13*, 1375-1388.

12. Swiderek, K.; Arafet, K.; Kohen, A.; Moliner, V., Benchmarking Quantum Mechanics/Molecular Mechanics (QM/MM) Methods on the Thymidylate Synthase-Catalyzed Hydride Transfer. *J Chem Theory Comput* **2017,** *13* (3), 1375-1388.

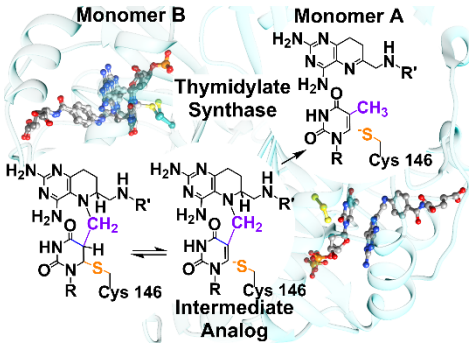
13. Blair, J., A,; Pearson, A. J., Kinetics and mechanism of the autoxidation of the 2-amino-4-hydroxy-5,6,7,8-tetrahydropteridines. *J. Chem. Soc., Perkin Trans. 2* **1974**, 80-88.

14. Fritz, T. A.; Liu, L.; Finer-Moore, J. S.; Stroud, R. M., Tryptophan 80 and leucine 143 are critical for the hydride transfer step of thymidylate synthase by controlling active site access. *Biochemistry* **2002,** *41* (22), 7021-9.

15. Newby, Z.; Lee, T. T.; Morse, R. J.; Liu, Y.; Liu, L.; Venkatraman, P.; Santi, D. V.; Finer-Moore, J. S.; Stroud, R. M., The role of protein dynamics in thymidylate synthase catalysis: variants of conserved 2'-deoxyuridine 5'-monophosphate (dUMP)-binding Tyr-261. *Biochemistry* **2006,** *45* (24), 7415-28.

16. Sapienza, P. J.; Popov, K. I.; Mowrey, D. D.; Falk, B. T.; Dokholyan, N. V.; Lee, A. L., Inter-Active Site Communication Mediated by the Dimer Interface beta-Sheet in the Half-the-Sites Enzyme, Thymidylate Synthase. *Biochemistry* **2019,** *58* (30), 3302-3313.

17. Finer-Moore, J. S.; Lee, T. T.; Stroud, R. M., A Single Mutation Traps a Half-Sites Reactive Enzyme in Midstream, Explaining Asymmetry in Hydride Transfer. *Biochemistry* **2018,** *57* (19), 2786-2795.


Monomer B
Monomer A
Thymidylate
Synthase
Cys 146
Cys 146
Cys 146
Intermediate
Analog