Supporting Information

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, California 94158, United States.

[†]Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain.

Corresponding Author

* E-mail: finer@msg.ucsf.edu, moliner@uji.es

Contents

Abbreviations

CH_2H_4F – N5,N10-methyelene-5,6,7,8- tetrahydrofolate; DCM – dichloromethane; DEA – diethylaniline; DMAD – dimethyl acetylenedicarboxylate; DMF – dimethylformamide; dTMP – thymidine 5'-monophosphate; t_R – retention time; DTT – D,L-dithiotreitol; dUMP – 2'-deoxyuridine 5'-monophosphate; EDTA – ethylenediaminetetraacetic acid; HRMS – high resolution mass spectrometry; pABAGlu – *para*-aminobenzoyl glutamate; RT – room temperature; TBAI – tetrabutylammonium iodide; TEAB – triethylammonium bicarbonate; THF – tetrahydrofuran; TLC – thin layer chromatography; TsCl – tosyl chloride; CB3717 – 10-propargyl-5,8-dideazafolate.

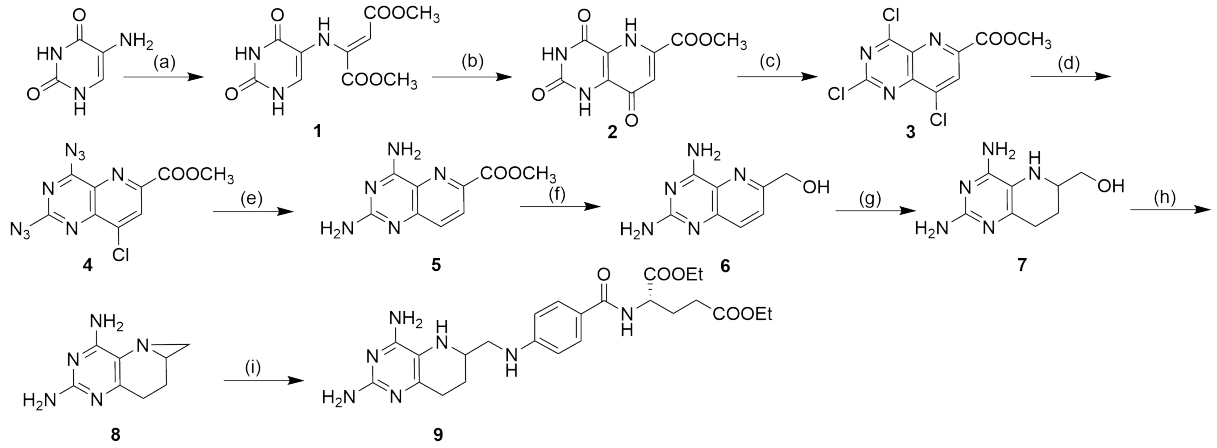
Materials and Methods

Unlabeled CH_2H_4F was a generous gift from EPROVA (Switzerland). *Ec*TSase was expressed and purified according to the established procedures,[1] and quantified spectrophotometrically at 280 nm using $\varepsilon_{280} = 0.591 \cdot 10^5$ $M^{-1}cm^{-1}$.[2] All other materials were synthesized as described below or purchased from Sigma or Fischer Scientific. Stocks of dUMP and CH_2H_4F were quantified using reported extinction coefficients: dUMP ε_{262} = 9.66 $mM^{-1}cm^{-1}$,[3] ; CH_2H_4F = 1.01 $mM^{-1}cm^{-1}$.[4] Semipreparative HPLC was performed using Agilent 1260 Infinity HPLC equipped with diode array detector and a C18 semipreparative column (5 µM particle size, 10 mm I.D. by 25 cm long; Supelco Discovery, Sigma-Aldrich). ^1H NMR spectra were acquired on a Bruker AVANCE III 600 or 400 MHz spectrometer equipped with a 1.7 mm ^1H,^{13}C,^{15}N probe using default pulse sequences. Mass spectrometry was performed on ESI-MS spectrometer Waters Acquity TQD operating in positive or negative ion mode.

Experimental Procedures

1. Organic Synthesis

Scheme S1. Synthesis of 4-amino-8-deazatetrahydrofolate diethyl ester (**9**)



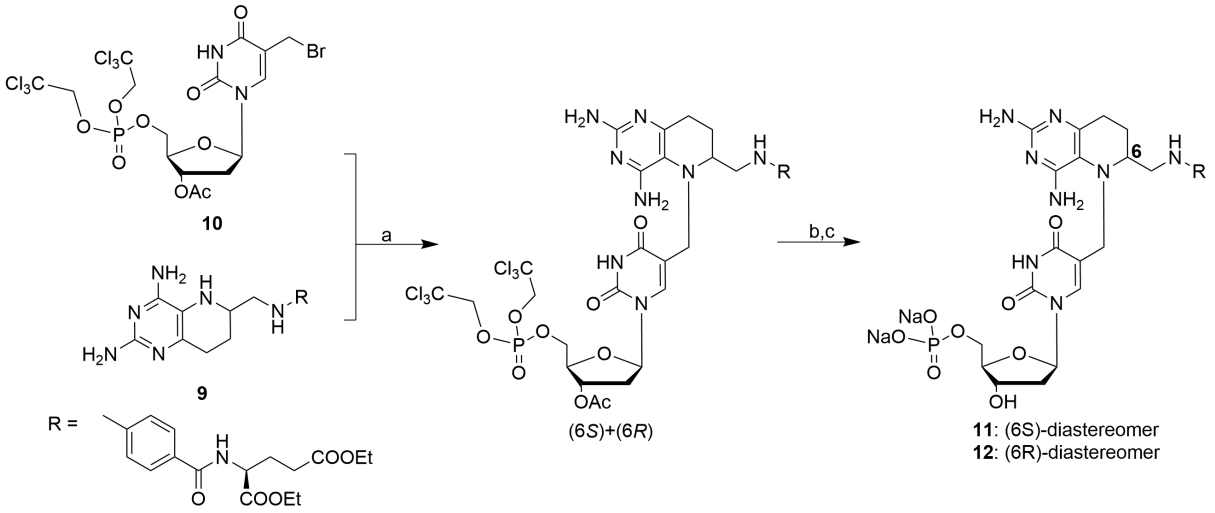
[a]Reaction conditions: (a) DMAD, MeOH, RT, 16 h; (b) Dowtherm A, ~257 °C (boiling point), 2 min; (c) $POCl_3$, *N*,*N*-DEA, 110 °C, 16 h (d) NaN_3, EtOH, RT, 2 h; (e) H_2, Pd/C, DMF, RT, 48 h; (f) $LiBH_4$, THF, 0 °C then RT, 24 h; (h) H_2, PtO_2, EtOH/HCl, RT, 24 h; (h) TsCl, TBAI, THF/NaOH, RT, 5 h; (i) pABAGlu, $Sn(OTf)_2$, DCM, RT, 48 h.

Dimethyl 2-(2,4-Dioxopyrimidin-5-yl)aminofumarate **1** was synthesized according to published procedures.[5] 6-(Carbomethoxy)-2,4,8-trioxopyrido[3,2-d]pyrimidine **2** and 6-(carbomethoxy)-2,4,8-trichloropyrido[3,2-d]pyrimidine **3** were synthesized as reported previously.[6] 6-(Carbomethoxy)-2,4-diazido-8-chloropyrido[3,2-d]pyrimidine **4**, 6-(carbomethoxy)-2,4-diaminopyrido[3,2-d]pyrimidine **5** and 6-(hydroxymethyl)-2,4-diaminopyrido[3,2-d]pyrimidine **6** were synthesized as reported previously.[7] (2,4-Diamino-5,6,7,8-tetrahydropyrido[3,2-d]pyrimidin-6-yl)methanol **7** and 2,4-Diamino-5,6-methylene-5,6,7,8-tetrahydropyrido[3,2-d]pyrimidine **8** were synthesized as reported previously.[8] ^{1}H-NMR spectra of the synthesized compounds were in agreement with those reported.

4-amino-8-deazatetrahydrofolate diethyl ester (9) (Scheme S1, step (i)). To a suspension of aziridine **8** (30 mg, 0.169 mmol) in anhydrous DCM was added pABAGlu diethyl ester diethyl ester (1.2 eq) and 3.5 mg $Sn(OTf)_2$ (5 mol%). Suspension was stirred at RT and followed by analytical HPLC. After 48 h complete consumption of **8** was

detected. Product was purified on a 1.5x20 cm silica column eluted with 10% MeOH in DCM (30% yield). ESI-MS (ESI+) calcd for **9** [M+H$^+$] 500.3, found 500.2.

Scheme S2. Synthesis of *N*5-thymidinyl derivatives of 4-amino-8-deazatetrahydrofolic acid (**11,12**).



[a]Reaction conditions: (a) $NaHCO_3$, DMF, RT, 16 h; (b) Zn/Cu, DMF, 55 °C, 5 h; (c) 1 M NaOH, RT.

***N*5-thymidinyl derivatives of 4-amino-8-deazatetrahydrofolic acid (11,12).** 106 mg (0.15 mmol) of 5-(bromomethyl)-3'-acetyl-2'-deoxyuridine 5'-(methyl 2,2,2-trichloroethyl phosphate) **10** prepared according to the published procedure[9] was dissolved in 2 ml of anhydrous DMF and added to dry residue of compound **9** (60 mg, 0.12 mmol), followed by 25 mg $NaHCO_3$. Resulting suspension was stirred for 36 h and partitioned between $CHCl_3$ and brine. Organic phase was washed 2-3 times with brine, dried over Na_2SO_4 and concentrated. Product was purified on a 1.5x40 cm silica gel column in 10% MeOH/$CHCl_3$. Fractions were analyzed by ESI MS and fractions containing product of expected mass were pooled and concentrated resulting in ~50 mg of product (63 % yield). (*R*)- and (*S*)- diastereomer were isolated by semipreparative C18 HPLC. Protected diastereomers peaks were detected at 19.7 min and 22.7 min upon elution with 1.25% MeOH/min in 10 mM TEAB pH 8.0 at 3.2 ml/min flow rate. Volatiles were removed by

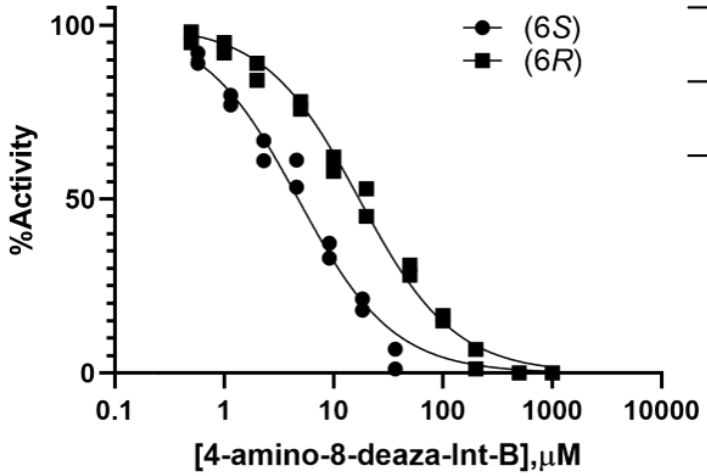
rotary evaporation and dry residues of diastereomers were deprotected as described below.

Fully protected precursor (5 mg, 4.5 μmole, single diastereomer) was dissolved in 1 mL anhydrous DMF containing 40 μL of acetylacetone and stirred with 30 mg of freshly prepared Zn/Cu couple (see preparation below) at 55 °C for 5 h. Resulting suspension was diluted with equal volume of 100 mM $Zn(OAc)_2$ and Zn-salt of the product was precipitated by centrifugation. Supernatant was removed and the pellet was washed with 1 mL of DMF, resuspended in 1 mL 1 M NaOH (aq.) and centrifuged. Supernatant adjusted to pH 8.0 and stirred with 3 mL of Chelex100 (Na^+ form) for 1 h. Suspension was filtered through 2 μM filter and adjusted to pH 7.0. The product was purified by semipreparative reverse phase chromatography using 25 cm X 10 mm Supelco Discovery C18 HPLC column. Peak of product was detected at 15.4 min (for 12) upon elution with a linear gradient of MeOH (1.25%/min) in 10 mM TEAB pH 8.0 at 3.2 ml/min. Collected product was lyophilized, dissolved in 50 μL of H_2O and added to 1 mL of ice-cold 100 mM NaI in dry acetone. After centrifugation, the pellet was washed with cold dry acetone and dried in vacuum at RT resulting in 2-3 mg of white solid (2.3-3.5 μmole, 51-78 % yield). ^{1}H NMR for **11** (600 MHz, D_2O): δ = 7.62 (s, 1H), δ = 7.47 (d, J = 8.7 Hz, 2H), δ = 6.46 (d, J = 8.8 Hz, 2H), δ = 5.58 (t, J = 6.6 Hz, 1H), δ = 4.32 (dt, J = 6.2, 3.3 Hz, 1H), δ = 4.21 (dd, J = 8.5, 4.7 Hz, 1H), δ = 4.09 (m, 1H), δ = 4.00-3.96 (m, 1H), δ = 3.91 (ddd, J = 11.5, 5.4, 2.9 Hz, 1H), δ = 3.67 (d, J = 14.3 Hz, 1H), δ = 3.34 (d, J = 11.1 Hz, 1H), δ = 3.11-2.98 (m, 2H), δ = 2.74 (dd, J = 15.0, 11.5 Hz, 1H), δ = 2.48 (dd, J = 18.5, 7.3 Hz, 1H), δ = 2.44-2.36 (m, 1H), δ = 2.23 (t, J = 8.0 Hz, 2H), δ = 2.18 (ddd, J = 13.9, 6.2, 3.4 Hz, 2H), δ = 2.04 (ddd, J = 12.5, 9.4, 5.5 Hz, 2H), δ = 1.9 (td, J = 14.1, 7.3 Hz, 1H), δ = 1.65 (dd, J = 13.8, 7.1 Hz, 1H). ESI-MS for **11** (ESI-) calcd for [M-H$^+$] 762.7, found 762.2. ^{1}H NMR for **12** (400 MHz, D_2O): δ = 7.55 (d, J = 8.8 Hz, 2H), δ = 7.5 (s, 1H), δ = 6.61 (d, J = 8.8 Hz, 2H), δ = 6.04 (dd, J = 8.0, 6.2 Hz, 1H), δ = 4.45-4.39 (m, 1H), δ = 4.26 (dd, J = 8.5, 4.6 Hz, 1H), δ = 4.07-4.01 (m, 1H), δ = 3.93-3.81 (m, 1H), δ = 3.31 (d, J = 11.0 Hz, 1H), δ = 3.18 (d, J = 14.5 Hz, 2H), δ = 2.66 (dd J = 14.9, 11.5 Hz, 1H), δ = 2.66-2.47 (m, 2H), δ = 2.3 (t, J = 8.0 Hz, 2H), δ = 2.18 (ddd, J = 14.0, 6.2, 2.4 Hz, 2H), δ = 2.14-2.04 (m, 2H), δ

= 2.02-1.91 (m, 1H), δ = 1.79-1.69 (m, 1H). ESI-MS for **12** (ESI-) calcd for [M-H$^+$] 762.7, found 762.5.

2. Inhibition Studies

Initial rates for *Ec*TSase-catalyzed reaction in the absence or presence of inhibitor were recorded upon monitoring of absorbance increase at 340 nm corresponding to conversion of CH_2H_4F to H_2F. Reaction mixtures containing 190 µL of 100 mM Tris·HCl pH 7.5, 50 mM $MgCl_2$, 1 mM EDTA, 7 mM HCHO, 25 mM DTT, 25 µM dUMP, 100 µM CH_2H_4F, and various concentrations of the inhibitor, were equilibrated at 25 °C for 2 minutes followed by starting the reaction upon addition of 10 µL of *Ec*TSase solution up to final concentration of 10 nM (dimer). Increase of the absorbance signal was monitored for 300 s using SpectraMax M5 plate reader (Molecular Devices, San Jose, CA).



	(6*S*)	(6*R*)
IC50, µM	4.7 ± 0.3	17.2 ± 0.8

Figure S1. Dependence of *Ec*TSase activity on concentration of (6*S*)- and (6*R*)-2-amino-8-deaza-Int-B at [dUMP] = 25 µM and [CH_2H_4F] = 100 µM. Curves are the result of nonlinear regression performed globally to two independent replicas per curve using equation %Activity = 100/(1+[4-amino-8-deaza-Int-B]/IC50).

3. X-ray Crystallography

Crystallization and data collection

*Ec*TSase- complexes were grown at 16°C by the hanging drop vapor diffusion method against 0.1 M HEPES, pH 7.1-7.5 and 1.3-1.35 M NaCitrate, One microliter protein solution [5.5 or 8 mg/ml *Ec*TSase for the (6*S*) and (6*R*)-4-amino-8-deaza-Int-B complexes, respectively, 3.3 mM inhibitor, 5 mM DTT] was mixed with an equal volume of well solution and suspended over 1000 μl reservoir solution. Hexagonal rods appeared after several days.

Data from crystals of the *Ec*TSase complexes were collected at the Lawrence Berkeley National Laboratory Advanced Light Source Beamline 8.3.1 at -170°C on a Dectris Pilatus3 S 6M detector using λ= 1.11587 Å. No cryoprotectant was required. Data were processed using XDS[10]. High resolution cutoffs for the data sets were determined based on $CC_{1/2}$ statistics[11]. Data collection statistics are in Table S1.

Structure solution and refinement

The structures of *Ec*TSase-(6*S*)-4-amino-8-deaza-IntB and *Ec*TSase-(6*R*)-4-amino-8-deaza-IntB were solved with Phaser from the PHENIX program suite[12] by molecular replacement using a dimer of *Ec*TSase-dUMP-CB3717 (PDB accession code 6NNR) with ligands removed as a starting model. Density fitting and assignment of alternate side chain conformations with Coot[13] was alternated with refinement of atomic positions, individual isotropic B-factors and occupancies against a maximum likelihood target in phenix.refine[12]. TLS refinement with a single TLS group was done as a final refinement step. Hydrogen atoms were included in riding positions during refinement. Refinement statistics are in Table S1.

Computer graphics

Figures 1,2 and S3 were made using UCSF Chimera.[14] Figures S2 and S4 were made using the PyMol Molecular Graphics System, Version 1.2r3pre, Schrodinger, LLC.

Table S1. Data collection and refinement statistics.

	*Ec*TSase-(6*S*)-4-amino-8-deaza-Int-B	*Ec*TSase-(6*R*)-4-amino-8-deaza-Int-B
Wavelength (Å)	1.116	1.116
Resolution range (Å)	62.99 - 1.821 (1.886 - 1.821)	108.6 - 1.5 (1.554 - 1.5)
Space group	P 63	P 63
Unit cell lengths (Å)	*a,b*=126.0 *c*=67.3	*a,b*=125.4 *c*=67.0
Total reflections	953580 (35955)	1397707 (25650)
Unique reflections	53737 (4581)	88795 (4716)
Multiplicity	17.7 (7.9)	15.7 (4.6)
Completeness (%)	97.54 (76.56)	91.77 (49.34)
Mean I/sigma(I)	17.35 (0.61)	26.24 (0.56)
Wilson B-factor (Å^2)	35.3	25.4
R-merge	0.1028 (3.267)	0.05009 (1.97)
R-meas	0.1058 (3.495)	0.05161 (2.215)
R-pim	0.02444 (1.191)	0.01216 (0.9855)
CC1/2	0.999 (0.147)	1 (0.212)
CC*	1 (0.507)	1 (0.592)
Reflections used in refinement	53308 (4167)	88135 (4716)
Reflections used for R-free	1984 (151)	1999 (105)
R-work	0.1507 (0.3717)	0.1513 (0.3320)
R-free	0.1901 (0.4415)	0.1746 (0.3291)
CC(work)	0.975 (0.493)	0.977 (0.533)
CC(free)	0.968 (0.545)	0.969 (0.513)
Number of non-hydrogen atoms	5020	5156
macromolecules	4593	4561
ligands	164	185
solvent	263	410
Protein residues	528	528
RMS(bonds) (Å)	0.012	0.010
RMS(angles) (°)	1.16	1.06
Ramachandran favored (%)	97.89	97.51
Ramachandran allowed (%)	1.72	2.11
Ramachandran outliers (%)	0.38	0.38
Rotamer outliers (%)	0.20	0.20
Clash score	6.05	4.88
Average B-factor (Å^2)	43.8	34.6
macromolecules	42.1	33.3
ligands	86.9	47.9
solvent	46.5	43.2
Number of TLS groups	1	1

Statistics for the highest-resolution shell are shown in parentheses.

*Ec*TSase-(6*R*)-4-amino-8-deaza-Int-B structure

(6*R*)-4-amino-8-deaza-Int-B was in two statistically disordered conformations in protomer B of the *Ec*TSase dimer (Figure S2A). In both conformations the dUMP moiety was in the substrate binding site for dUMP while the 4-amino-8-deaza-pterin moiety was in two novel positions. In one binding mode the 4-amino-8-deaza-pterin ring is positioned against the face of the dUMP moiety, flipped 180° with respect to the (6S)-4-amino-8-deaza-pterin moiety in the first structure, allowing the PABA-Glu moiety to bind in the PABA-Glu-binding groove of the enzyme (Figure S3A). In the second mode (Figure S3C, 2) the 4-amino-8-deaza-pterin moiety is bound to a pocket above the dUMP moiety, surrounded by conserved amino acids and donating two hydrogen bonds to the carboxyl of invariant residue Glu58. This same pocket housed the N10-propargyl group of 10-propargyl-5,8-dideazafolate (CB3717) in the ternary complex *Ec*TSase·dUMP·CB3717. The PABA-Glu moieties for both conformations lie in the PABA-Glu binding groove, but are in broken density, thus highly disordered. To accommodate inhibitor in the second site (above the dUMP) a helix-loop-helix subdomain containing the important cofactor-binding residues Trp80 and Trp83, has shifted outward by 1.0 -1.5 Å. The dUMP moiety C6 is not covalently attached to the catalytic Cys.

In monomer A the 4-amino-8-deaza-pterin N5 methylene has been transferred to C5 of dUMP (Figure S2B); there is very little bridging density between the dUMP and the 4-amino-8-deaza-pterin moiety. This implies that the reaction proceeded to yield the exocyclic methylene intermediate. The C6 hydrogen of the (6*R*)-4-amino-8-deaza-pterin moiety is not positioned for hydride transfer to the exocyclic methylene. Density suggests that instead a nucleophile such as water, added to the exocyclic methylene to give, for example, the 5-hydroxymethyl derivative of dUMP. There is no covalent bond between 5-hydroxymethyl-dUMP and the enzyme. The density for the C5-substituent was weaker than for the rest of the pyrimidine ring, suggesting the reaction occurred only partially. This finding agrees with the results of our previous study[15] where (6*R*)-diastereomer of the Int-B was slowly reacting in the presence of *Ec*TSase to form H_4F and dithiothreitol-

trapped dTMP (an adduct resulting from nucleophilic attack of dithiothreitol present in the reaction buffer on the methylene exocyclic intermediate).

As in monomer B, the (6*R*)-4-amino-8-deaza-pterin moiety is in two orientations, one parallel to the pyrimidine plane of dUMP and the second in the pocket above the dUMP (Figure S3D). The second site is shifted compared to the second site in monomer B; untethered to the dUMP moiety, the (6*R*)-4-amino-8-deaza-pterin can bind more deeply into the pocket, and the *Ec*TSase subdomain containing Trp80 and Trp83 is not shifted relative to this subdomain in the *Ec*TSase·FdUMPCH_2H_4F structure.

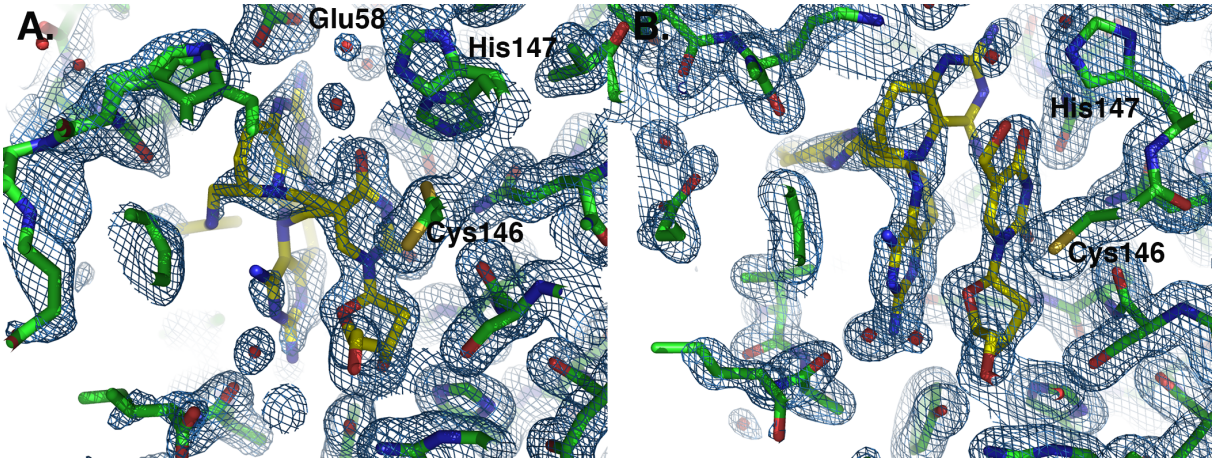


Figure S2. Refined structures of *Ec*TSase-(6R)-4-amino-8-deaza-Int-B plotted as sticks against the 2mFo-DFc density map calculated after refinement, shown as blue mesh. The protein is colored by atom type with green carbons, red oxygens, blue nitrogens and yellow sulfurs. The inhibitor has gold carbons. **A**. In monomer B, the 4-amino-8-deaza-pterin moiety of the inhibitor occupies either one of two partially occupied sites in the crystal structure. His147 and Cys146 at the dimer interface are in two conformations. **B.** In monomer A, the inhibitor has resolved into 5-hydroxymethyl-dUMP and 4-amino-8-deaza-H_2F. The latter is a statistical mixture of two conformations similar to the inhibitor in monomer B.

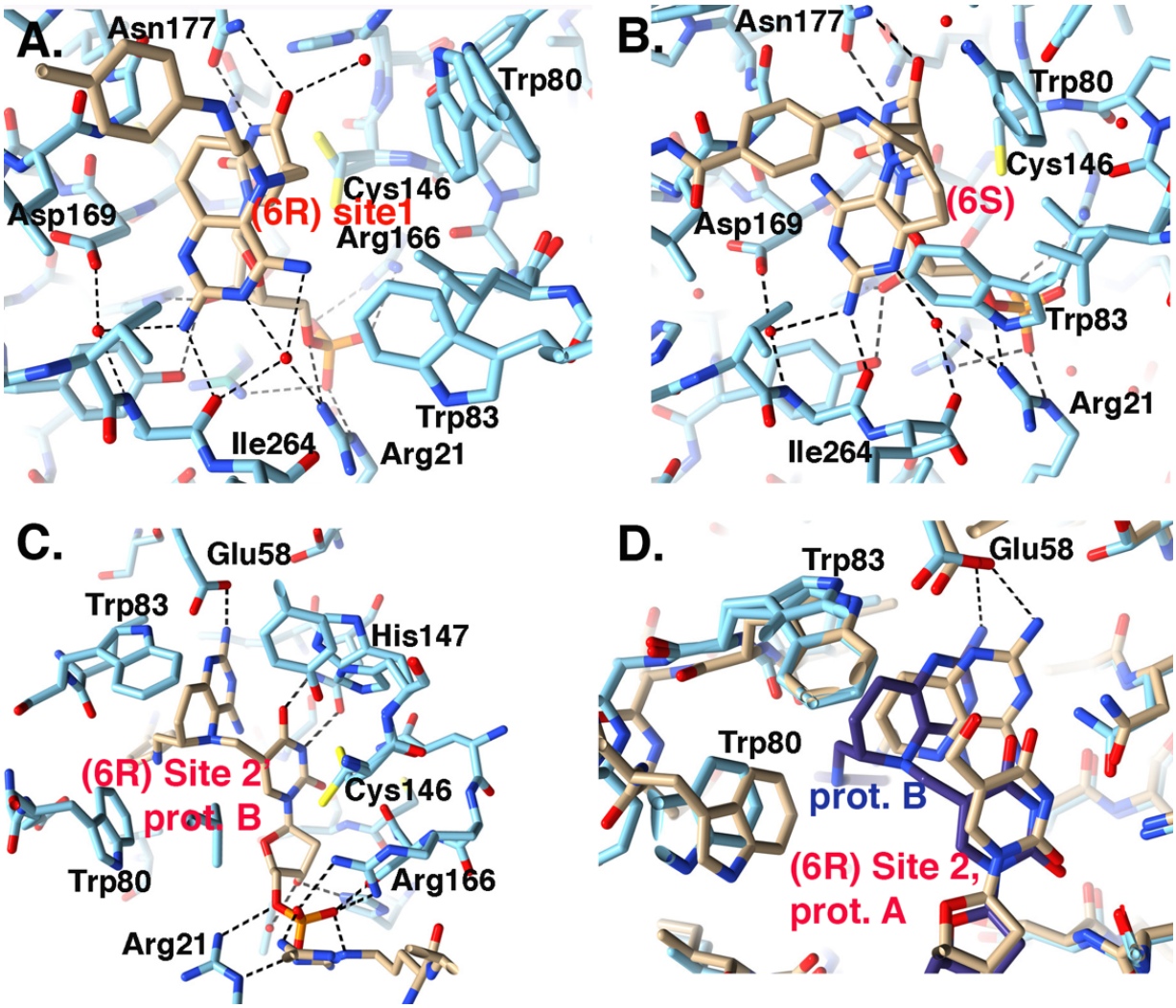


Figure S3. Two binding sites for the 4-amino-8-deaza-pterin moiety in *Ec*TSase- (6*R*)-4-amino-8-deaza-Int-B. **A.** In site 1 (shown here for monomer B) the 4-amino-8-deaza-pterin plane lies on top of and parallel to the plane of the dUMP moiety, similarly to, but flipped 180° from, the pterin moiety in the (6*S*)-inhibitor, shown in panel **B. C.** In site 2 (shown here for monomer B), the 4-amino-8-deaza pterin lies in a hydrophobic pocket above the dUMP moiety where it donates a hydrogen bond to Glu58. **D.** In monomer A, the 4-amino-8-deaza pterin is not attached to the 5-hydroxymethyl-dUMP so site 2 is more deeply buried in the pocket. The structure is drawn with sticks color coded as in panel **A** except with all carbons tan. For comparison, protomer B is superimposed in sticks with light and dark blue carbons for the protein and inhibitor, respectively. Structures in panels (A-C) are drawn with sticks colored by atom type: carbon, blue for protein and tan for inhibitor; oxygen, red; nitrogen, blue; sulfur, yellow. Dashed lines are hydrogen bonds.

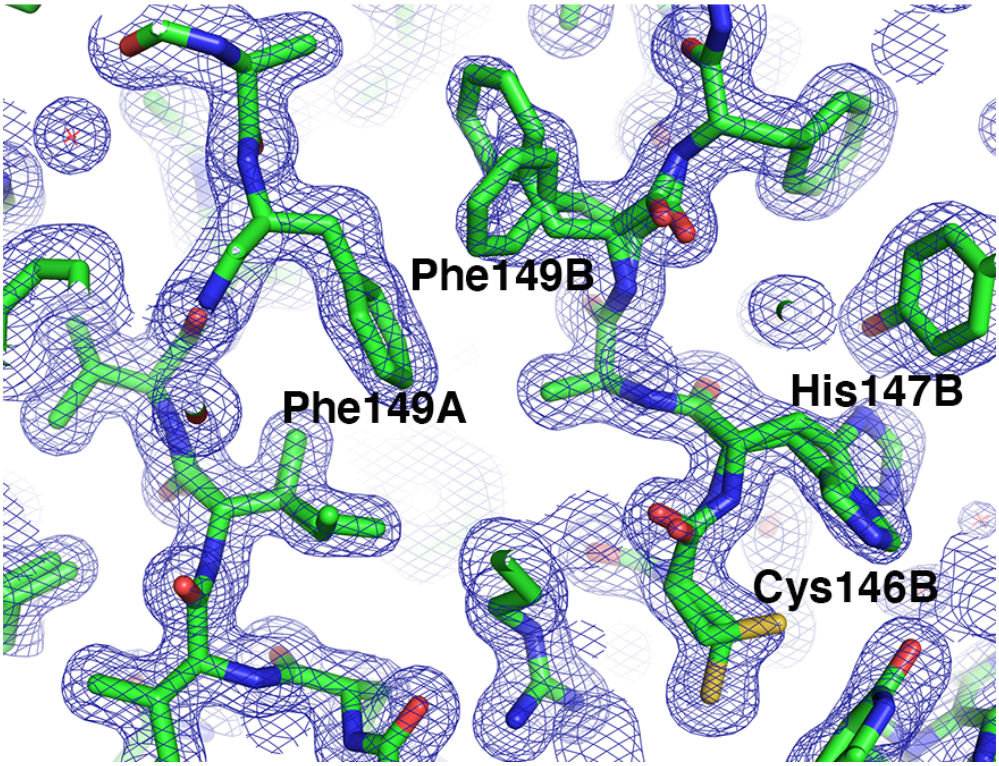


Figure S4. Beta-strands at the dimer interface that contains the catalytic Cys, shown for the *Ec*TSase-(6*R*)-4-amino-8-deaza-Int-B structure. Residues Cys146, His147 and Phe149 are statistically disordered only in protomer B, where each adopts two alternate conformations. Density from a 2mFo-DFc map is shown in blue mesh.

4. QM/MM Calculations

Setting up the molecular model. Simulations were performed based on two different starting geometries from Escherichia coli TSase: PDB code 1TLS[16]. The X-ray structure had both FdUMP and CH_2H_4folate bound at the active site of each monomer. The fluorine was replaced with a hydrogen atom at C5 of dUMP, and the CH_2H_4folate was replaced by the (6*S*)-4-amino-8-deaza H_4F. The coordinates of the hydrogen atoms were added to both structures using the fDYNAMO[17] library after computing the pK_a values of ionisable amino acids with the empirical PROPKA3 program[18-19], in order to determine their proper protonation state in the protein environment. A total of 22 counterions (Na^+) were placed in optimal electrostatic positions around the enzyme (further than 10.5 Å from any atom of the system and 5 Å from any other counterion, using a regular grid of 0.5 Å) because the total charge of the system was not neutral. Finally, the systems were solvated using boxes of water molecules of 100 × 80 × 80 $Å^3$, and the water molecules with an oxygen atom lying within 2.8 Å of any heavy atom were removed.

The whole system was divided into a QM part and a MM part to perform hybrid QM/MM calculations. The QM part considers 27 atoms of the (6S)-4-amino-8-deaza H_4F, 25 atoms of the dUMP and the side chain of Cys146, which gives a total of 56 QM atoms, and three hydrogen link atoms that were added at the boundary between QM and MM regions to satisfy the valence of the QM/MM frontier atoms (see figure S5)[20]. The calculations only modeled one active site in the QM region, leaving the second active site (ligands removed) in the MM region. The MM part comprises the rest of the folate and dUMP, the enzyme, the crystallization and solvation water molecules and the sodium counterions, which makes a total of 60826 atoms.

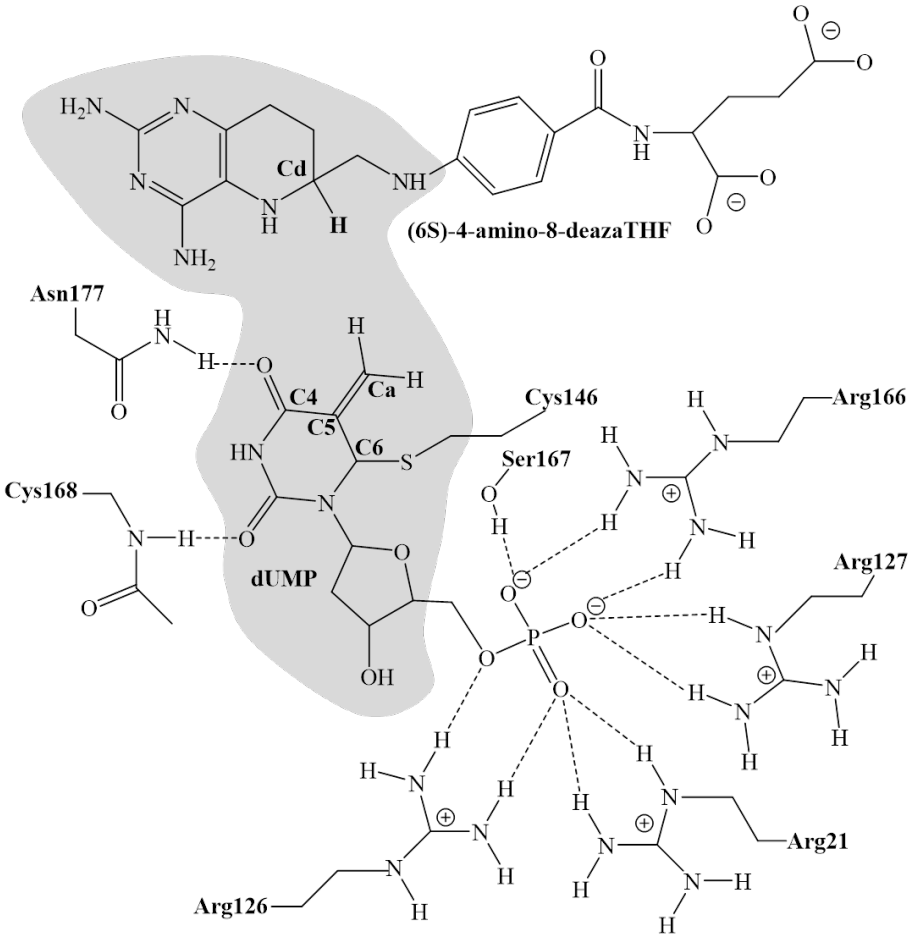


Figure S5. Schematic representation of the active site of the *E. coli* TSase prior to the hydride transfer step, with the (6S)-4-amino-8-deaza H_4F as substrate. Atoms in the grey region are treated by QM methods and QM-MM frontiers are depicted as black thick bars. The transferred hydride is highlighted, and its donor and acceptor atoms marked as Cd and Ca, respectively.

During the QM/MM energy optimizations, the atoms of the QM region were treated by the semiempirial Hamiltonians AM1[21] and RM1[22], and the B3LYP[23-24] and M06-2X[25] hybrid density functional theory (DFT) methods. The standard 6-31G+(d,p) basis set was used in the DFT calculations. The combination of B3LYP:AM1 (as explained below) was used to explore the C6-Cys146 breaking bond without a concomitant proton transfer from the C5 carbon (to perform direct comparisons with our last study on the exploration of the intermediate structures)[26], while the combination M06-2X:RM1 was used for the hydride transfer step (to perform direct comparisons with our last study on the exploration of this step).[27] Anyway, it is important to point out that the results obtained with the different

methods were equivalent. The rest of the system, protein and water molecules, are described using the OPLS-AA[28]and TIP3P[29] force fields, respectively, as implemented in the fDYNAMO library. Cut-offs for the non-bonding interactions are applied using a force switching scheme, within a range radius from 14.5 to 16 Å. After thermalization, QM/MM MD simulations of the system in the NVT ensemble (with the QM region treated at AM1 or RM1 level) were ran during 500 ps at a temperature of 300 K using the Langevin-Verlet algorithm using a time step of 1 fs. The molecular dynamics (MD) simulations were performed using an integration step size of 1 fs and the velocity Verlet algorithm.[30] The resulting structures had an energy fluctuation lower than 0.1%, a kinetic energy fluctuation lower than 1% and a change in temperature lower than 2.5 K over the last 2 ps of the MD. Moreover, according to the time-dependent evolution of the RMSD of those atoms belonging to the protein backbone, the systems can be considered equilibrated.

Potential Energy Surface, PES. After setting up the model, the full system was minimized using the Adopted Basis Newton Raphson (ABNR) method, keeping the backbone of the enzyme as frozen. Due to the huge dimensions of the system, all the residues further than 20 Å from the QM part were kept frozen. The structures were then used to generate a mono-dimensional (1D) QM/MM PES, to explore the C6-Cys146 breaking bond without a concomitant proton transfer from the C5 carbon starting from Int-B, and bi-dimensional (2D) QM/MM PESs for the r.d.s. for the hydride transfer step). The refined TSs were used as starting point to generate the 1D PMFs, while the geometries from the 2D PESs were used to generate the 2D PMFs.

Potential of Mean Force, PMF. Mono-dimensional PMFs, 1D PMF, were computed using the C6-SGCys146 distance as the distinguished reaction coordinate, as in the PES calculations. The weighted histogram analysis method (WHAM), combined with the umbrella sampling approach,[31-32] was employed to scan the reaction coordinate in a range from 1.7 to 4.1 Å, with a window width of 0.05 Å. Two-dimensional PMFs, 2D PMF, were obtained with the anti-symmetric combination of the distances describing the breaking and forming bonds on the hydride transfer step, d(Cd-H) – d(H-Ca), and the distance between the C6 of the dUMP and the sulfur atom of the Cys146 (dC6-S). A total

of 61 simulations were performed at different values of d(Cd-H) – d(H-Ca) with an umbrella force constant of 2500 kJ·mol^{-1}·A^{-1} for each particular value of the distance dC6-S (23 simulations with a force constant of 2500 kJ·mol^{-1}·A^{-1}, from 1.5 to 4.0 Å).

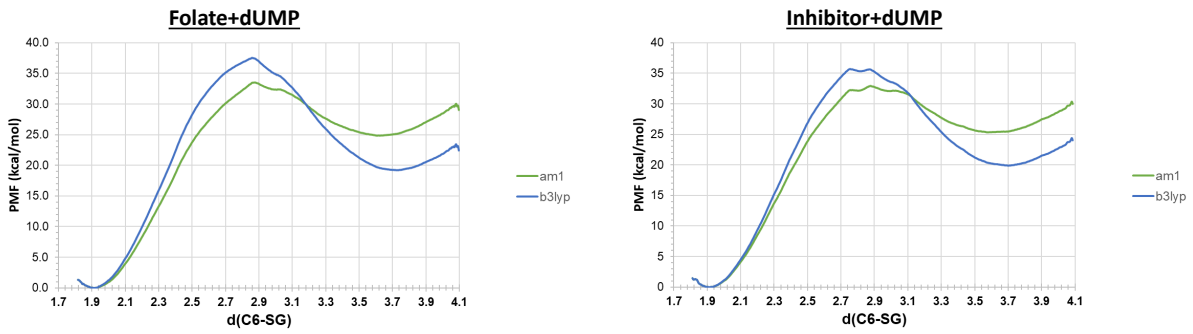
The values of the variables sampled during the simulations were then pieced together to construct a full distribution function from which the 1D-PMF and 2D-PMF were obtained. On each window, 5 ps of relaxation was followed by 10 ps of production with a time step of 0.5 ps due to the nature of the chemical step involving a hydrogen transfer. The velocity Verlet algorithm[30] was used to update the velocities in each window. All the PMFs were performed at 303 K, using as starting points of each window structures from the previous obtained 1D and 2D PESs.

Because of the large number of structures that must be evaluated during free energy calculations, QM/MM MD calculations were done with the AM1 semiempirical Hamiltonian. In order to improve the quality of our MD simulations, following the work of Truhlar et al.,[33-35] a spline under tension[36-37] was used to interpolate this correction term at any value of the reaction coordinates ξ_1 and ξ_2 selected to generate the 2D PMFs. In this way we obtain a continuous function to obtain corrected PMFs:[38-40]

$$E = E_{AM1/MM} + S\left[\Delta E_{LL}^{HL}(\xi_1, \xi_2)\right] \qquad (1)$$

where S denotes a two-dimensional cubic spline function, and its argument is a correction term evaluated from the single-point energy difference between the high-level (HL) and the low-level (LL) calculation of the QM subsystem. In particular, S was adjusted to a grid of points obtained as HL single energy calculation corrections. The semiempirical AM1 or RM1 Hamiltonians were used as LL method, while B3LYP or M06-2X hybrid functional, with the standard 6-31G+(d,p) basis set, was selected for the HL energy calculations. The functional and basis set were selected following the suggestions of Truhlar and co-workers,**Error! Bookmark not defined.** while the use of a semiempirical method such as AM1 (which gives reasonable results according to our own experience) is dictated due to the cost of the hybrid QM/MM MD simulations required to obtain PMFs. These calculations were carried out using the Gaussian09 program.[41]

A



B

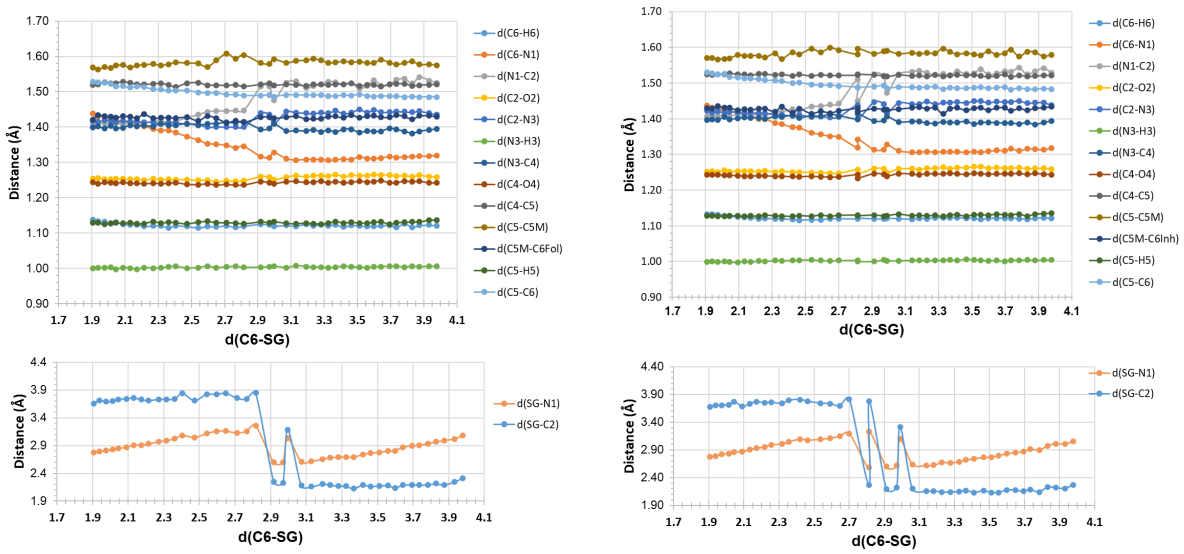


Figure S6. A) B3LYP:AM1/MM FESs for the exploration of the C6-Cys146 breaking bond without a concomitant proton transfer from the C5 carbon atom obtained with the AM1/MM (green line) or the B3LYP/MM method (blue line) to describe the QM subset of atoms. Results obtained with the folate (left panel) and with the inhibitor (right panel). B) Evolution of key inter-atomic distance along the reaction.

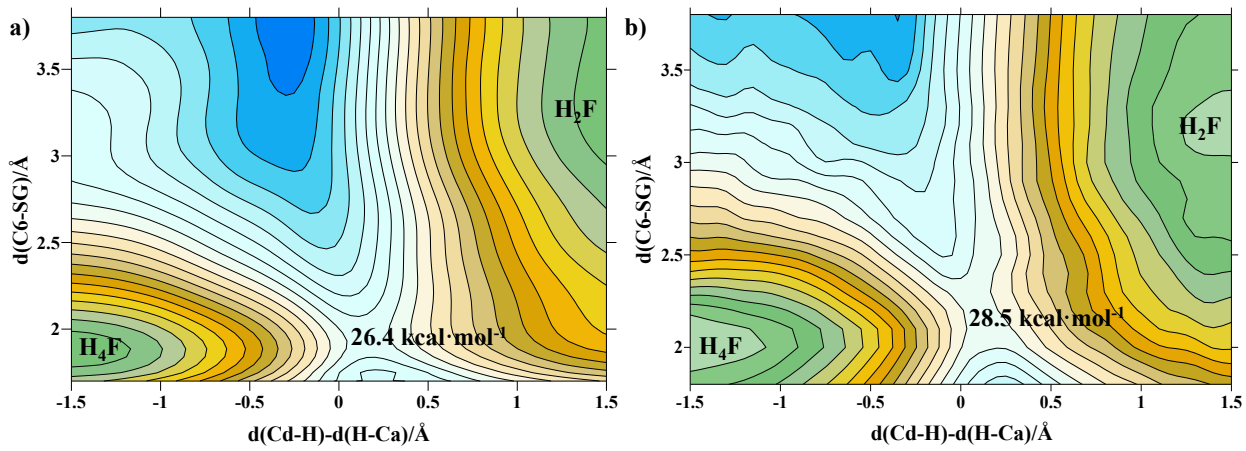


Figure S7. M06-2X:RM1/MM Free Energy Surfaces for the hydride transfer step catalyzed by TSase with (a) the natural substrate; and (b) the (6*S*)-4-amino-8-deaza H_4F. The surface on panel (b) was generated from the data published in ref. 26. Distances of axes are given in units of Å, while values on isoenergetic lines are given in units of kcal mol. Surfaces computed at 303 K.

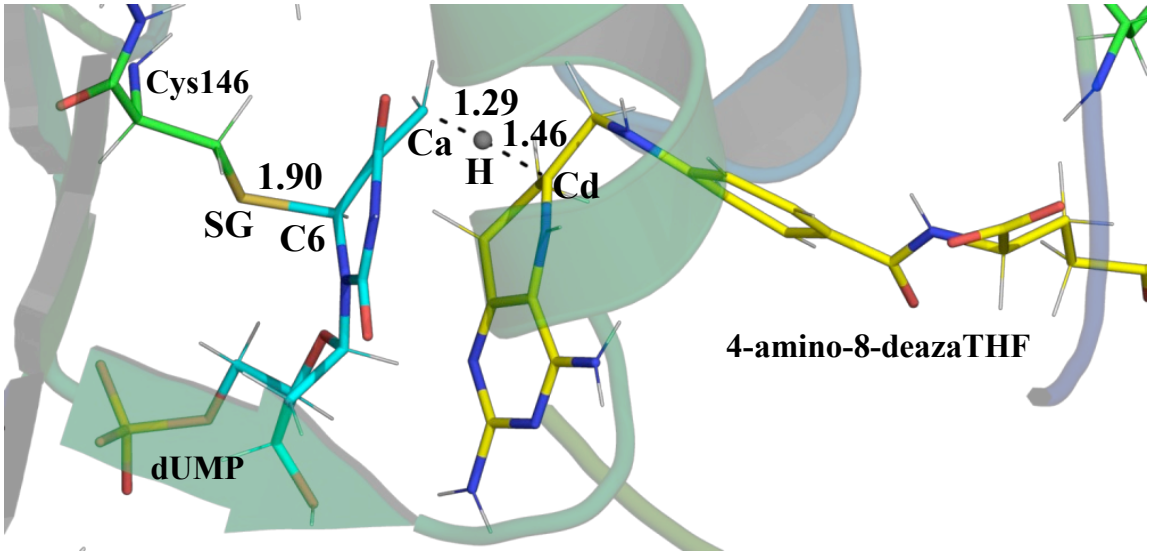


Figure S8. Representative snapshot of the transition state structure of the hydride transfer located on the M06-2X:RM1/MM Free Energy Surfaces with the (6*S*)-4-amino-8-deaza H_4F co-substrate. Distances in Å.

1. Changchien, L. M.; Garibian, A.; Frasca, V.; Lobo, A.; Maley, G. F.; Maley, F., High-level expression of Escherichia coli and Bacillus subtilis thymidylate synthases. *Protein. Expr. Purif.* **2000,** *19* (2), 265-70.

2. Saxl, R. L.; Maley, G. F.; Hauer, C. R.; Maccoll, R.; Changchien, L.; Maley, F., Significance of mutations on the structural perturbation of thymidylate synthase: implications for their involvement in subunit exchange. *Protein. Sci.* **2007,** *16* (7), 1439-48.

3. Cavaluzzi, M. J.; Borer, P. N., Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic. Acids. Res.* **2004,** *32* (1), e13.

4. Ramasastri, B. V.; Blakley, R. L., 5,10-METHYLENETETRAHYDROFOLIC DEHYDROGENASE FROM BAKERS' YEAST. II. USE IN ASSAY OF TETRAHYDROFOLIC ACID. *J. Biol. Chem.* **1964,** *239*, 106-11.

5. Srinivasan, A.; Broom, A. D., Pyridopyrimidines. 10. Nucleophilic substitutions in the pyrido [3, 2-d] pyrimidine series. *The Journal of Organic Chemistry* **1979,** *44* (3), 435-440.

6. Rosowsky, A.; Forsch, R. A.; Queener, S. F., 2, 4-Diaminopyrido [3, 2-d] pyrimidine inhibitors of dihydrofolate reductase from Pneumocystis carinii and Toxoplasma gondii. *Journal of medicinal chemistry* **1995,** *38* (14), 2615-2620.

7. Bayomi, S. M.; Brixner, D. I.; Eisa, H.; Broom, A. D.; Ueda, T.; Cheng, Y.-C., Probing the thymidylate synthase active site with bisubstrate analog inhibitors. *Nucleosides & Nucleotides* **1988,** *7* (1), 103-115.

8. Zhou, S.; Tian, C.; Li, C.; Guo, Y.; Wang, X.; Liu, J.; Zhang, Z., Novel Synthesis of 8-Deaza-5,6,7,8-tetrahydroaminopterin Analogues via an Aziridine Intermediate. *Molecules* **2012,** *17* (5), 5604-5614.

9. Kholodar, S. A.; Kohen, A., Noncovalent intermediate of thymidylate synthase: Fact or fiction? *Journal of the American Chemical Society* **2016,** *138* (26), 8056-8059.

10. Kabsch, W., Xds. *Acta crystallographica. Section D, Biological crystallography* **2010,** *66* (Pt 2), 125-32.

11. Karplus, P. A.; Diederichs, K., Linking crystallographic model and data quality. *Science* **2012,** *336* (6084), 1030-3.

12. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **2010,** *66* (Pt 2), 213-21.
13. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **2004,** *60* (Pt 12 Pt 1), 2126-32.
14. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **2004,** *25* (13), 1605-12.
15. Kholodar, S. A.; Kohen, A., Noncovalent Intermediate of Thymidylate Synthase: Fact or Fiction? *J Am Chem Soc* **2016,** *138* (26), 8056-9.
16. 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.
17. Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A., The dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *Journal Comput. Chem.* **2000,** *21*, 1088-1100.
18. Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **2005,** *61* (4), 704-21.
19. Bas, D. C.; Rogers, D. M.; Jensen, J. H., Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* **2008,** *73* (3), 765-83.
20. Field, M. J.; Bash, P. A.; Karplus, M., Combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990,** *11*, 700-733.
21. Dewar, M. J. S.; Zoebisch, E.; Healy, E. F.; Stewart, J. J. P., AM1 a New General Purpose Quantum Mechanical Molecular Model. *Journal of the American Chemical Society* **1985,** *107*, 3902-3909.

22. Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. J., RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *J. Comput. Chem.* **2006,** *27*, 1101-1111.
23. Becke, A. D., Becke's three parameter hybrid method using the LYP correlation functional. *J. Chem. Phys.* **1993,** *98*, 5648-5652
24. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H., Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989,** *157*, 200-206.
25. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, kinetics, nocovalent. *Theor. Chem. Acc.* **2008,** *120*, 215-241.
26. 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.
27. 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.
28. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J., Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996,** *118*, 11225-11236.
29. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.
30. Verlet, L., Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review* **1967,** *159* (1), 98-103.
31. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992,** *13*, 1011-1021.
32. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977,** *23*, 187-199.

33. Corchado, J. C.; Coitino, E. L.; Chuang, Y.; Fast, P. L.; Truhlar, D. G., Interpolated Variational Transition State Theory by Mapping. *J. Phys. Chem. A* **1998,** *102*, 2424-2438.
34. Nguyen, K. A.; Rossi, I.; Truhlar, D. G., A Dual-Level Shepard Interpolation Method for Generating Potential Energy Surfaces for Dynamics Calculations. *J. Chem. Phys.* **1995,** *103*, 5522-5530.
35. Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G., Mapped Interpolation Scheme for Single-Point Energy Corrections in Reaction Rate Calculations and a Critical Evaluation of Dual-Level Reaction-Path Dynamics Methods. *J. Phys. Chem. A* **1999,** *103*, 1140-1149.
36. Renka, R. J., Interpolatory tension splines with automatic selection of tension factors. *SIAM J. Stat. Comput.* **1987,** *8*, 393-415.
37. Renka, R. J., Algorithm 716: TSPACK:tension spline curve-fitting package. *ACM Trans. Math. Software* **1993,** *19*, 81-94.
38. Roca, M.; Moliner, V.; Ruiz-Pernia, J. J.; Silla, E.; Tunon, I., Activation free energy of catechol O-methyltransferase. Corrections to the potential of mean force. *J Phys Chem A* **2006,** *110* (2), 503-9.
39. Ruiz-Pernia, J. J.; Silla, E.; Tunon, I.; Marti, S., Hybrid quantum mechanics/molecular mechanics simulations with two-dimensional interpolated corrections: application to enzymatic processes. *J Phys Chem B* **2006,** *110* (35), 17663-70.
40. Ruiz-Pernia, J. J.; Silla, E.; Tunon, I.; Marti, S.; Moliner, V., Hybrid QM/MM Potentials of Mean Force with Interpolated Corrections. *J. Phys. Chem. B* **2004,** *108*, 8427-8433.
41. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, B.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Calicato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J.A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J.

C.; Iyengar, S. S.; Tomasi, J.; Cosssi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, A.; Daniels, D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*, Gaussian, INC: Wallingford, CT, 2009.