Supporting Information

Wan et al. 10.1073/pnas.1415856111

SI Materials and Methods

Proposals for Keto-Enol Tautomerization of Substrate and the DHFR Catalytic Mechanism. As noted in the main text, competing versions of this mechanism have been proposed previously (1–3). In one version, the Asp27 is ionized in the ground state. However, on substrate binding, water is released from the active site, lowering the local dielectric constant and raising the pK_a of Asp27, allowing its carboxylate to become protonated by the substrate N3 (1). This initiates tautomerization of DHF at the pterin N3-C4-O4 atoms, ultimately converting the O4 to a hydroxyl (enol). The enol at pterin O4 is the proton donor to N5. The Asp27 protonates N3, returning Asp27 to a carboxylate species and allowing the substrate to tautomerize to the keto form.

A related proposal is that Asp27 is protonated in the ground state and triggers a proton relay with the pterin O4 via a crystallographically observed water molecule (2), resulting in an enol at C4-O4 as described above. The enol tautomer is stabilized by the Gandour effect (4), and the O4 transfers its proton to a water molecule that has gained access to the active site due to Met20 loop fluctuations. This water, not the O4 enol, is the ultimate proton donor to N5 (3).

HDX Characteristics of the αA Helix. The αA helix, which undergoes complete HDX, is associated with substrate binding. Substrate binding expels water molecules from the active site and causes the αA helix to become more accessible, resulting in the main chain amides of the αA helix forming hydrogen bonds to 14 water molecules.

Purification, Crystallization, and Data Collection. ecDHFR was expressed, purified, and crystallized as described previously (5, 6). Folate and $NADP^+$ were added to the purified DHFR solution at a molar ratio of 3:3:1 and concentrated to 40 mg/mL. The ecDHFR-folate-$NADP^+$ complex was initially crystallized under the following conditions: 18% PEG 400, 100 mM $MnCl_2$, and 20 mM imidazole (pH 7.0). A single 3.6-mm^3 (4 × 1.3 × 0.7 mm) crystal was obtained by the microseeding technique, followed by the macroseeding technique. The 2.0-Å resolution neutron diffraction data were collected at 291 K using IMAGINE at Oak Ridge National Laboratory. The 1.6-Å resolution X-ray diffraction data were collected at 291 K with a smaller crystal using a Rigaku MicroMax-007HF X-ray generator, Osmic VariMax optics, and an R-AXIS IV++ image plate. Collection of the 291 K neutron and X-ray data sets has been described previously (6). The diffraction data were processed and scaled with the HKL3000 program (7).

For the 100 K and 277 K datasets, crystals were grown from recombinant protein overexpressed in the *E. coli* strain BL21 (DE3) (Novagen) using a pET22b construct (8). Cells were grown in LB medium with 100 μg/mL ampicillin and shaking at 37 °C until they reached an OD600 of 0.7, at which point they were induced with 0.4 mM IPTG. Induced cells were incubated at room temperature overnight and then harvested by centrifugation, frozen in liquid nitrogen, and stored at −80 °C.

Cell pellets were lysed on ice in 25 mM Hepes (pH 7.5), 1 mM EDTA, and 1 mM DTT (stabilization buffer) with 1 mg/mL lysozyme and brief sonication. The lysate was centrifuged at 10,000 × *g*, and streptomycin sulfate was used to precipitate nucleic acids. The sample was again centrifuged at 10,000 × *g* for 30 min, after which the supernatant was dialyzed overnight against stabilization buffer at 4 °C. DHFR was affinity-purified from the dialyzed clarified lysate using an MTX-agarose DHFR affinity resin (Sigma-Aldrich). Resin-bound DHFR was eluted with 200 mM sodium borate (pH 9.0), 3 mM folic acid, 0.5 M NaCl, 1 mM EDTA, and 1 mM DTT (elution buffer). Any endogenous copurifying nucleotide cofactors in the highly pure DHFR fraction were removed using a DEAE-sephacel anion-exchange column with a linear elution gradient from 0 to 0.5 M KCl in 10 mM Tris·HCl, 1 mM EDTA, and 1 mM DTT. DHFR-containing fractions were pooled, and their concentrations were determined using absorbance at 280 nm (ε280 = 33,690 M^{-1} cm^{-1}) and bound to folic acid by dialysis against 20 mM imidazole (pH 7.0) (crystallization buffer) with 1 mM folic acid at 4 °C. DHFR-folate was concentrated to 35 mg/mL and stored at 4 °C.

Freshly made $NADP^+$ was added to aliquots of the concentrated DHFR-folate complex to a final concentration of 6 mM. Crystals were grown at 4 °C using the hanging drop vapor diffusion method against a reservoir solution containing 17% PEG 400, 20 mM imidazole pH 7.0, and 125 mM $MnCl_2$. Crystals grew to their maximum dimensions after 1–2 wk. For the sample used to collect the 100 K X-ray dataset, one crystal was cryoprotected by transfer through increasing concentrations of PEG 400 to a final concentration of 30% in 20 mM imidazole (pH 7.0) and 125 mM $MnCl_2$. For the 277 K dataset, several crystals were placed into reservoir solution [17% PEG 400, 20 mM imidazole (pH 7.0), and 125 mM $MnCl_2$] and mounted into quartz capillary tubes (Hampton Research).

Structure Determination and Refinement. The joint X-ray neutron (XN) structure of ecDHFR-folate-$NADP^+$ ternary complex was refined against the 2.0-Å neutron and 1.6-Å X-ray diffraction data using the Phenix program suite (9) and manipulated manually using Coot (10). After removing all of the ligands and water from the initial model [PDB ID code 1RX2 (11)], the structure was solved using the molecule replacement program Phaser (12), incorporated in Phenix, and refined against the X-ray data. After two rounds of refinement and model manipulation, folate and $NADP^+$ were built in according to the clear electron density in the 2Fo − Fc and Fo − Fc omit maps. Hydrogen atoms and exchangeable deuterium atoms were added into the model and the model was refined against the X-ray and neutron diffraction data afterward. D_2O water molecules were finally added according to both the nuclear and electron density maps. The final R and R_{free} values of the refined model were 0.2310 and 0.2714, respectively (Table S1).

The atom depth, calculated as the distance between each atom to the closest solvent accessible neighboring atom, was carried out using the DPX server (13). The topology diagram depicting the hydrogen/deuterium exchange rate was generated using the TopDraw program (14) incorporated in the CCP4 program suite (15). The solvent accessible surface area was calculated using the Areaimol program (16) incorporated in the CCP4 program suite.

The 100 K and 277 K X-ray datasets were collected at beamline 11-1 at the Stanford Synchrotron Radiation Laboratory. For the 100 K dataset, the data were obtained from a single crystal using 13.776-KeV incident X-rays. Two orientations of this crystal were used to measure reflections that would otherwise have been lost in the blind region. Both high-resolution (2.0–0.85 Å) and low-resolution (75–1.5 Å) datasets were collected for both the 100 K and 277 K samples. For the 277 K dataset, three capillary-mounted crystals at 277 K were used to collect data in multiple wedges to minimize radiation damage. All diffraction data were indexed, integrated, and scaled in HKL2000, with data statistics provided in Table 1.

The models for the 100 K and 277 K X-ray data for the DHFR-folate-NADP$^+$ ternary complexes were refined with SHELX-97 (17) using 1RX2 as the initial model. Initial rigid body refinement was followed by stereochemically restrained refinement of coordinates and individual atomic isotropic displacement parameters against an intensity-based target. Restraints were used for all protein and cofactor atoms at each stage of the refinement. Individual anisotropic atomic displacement parameters were refined for all atoms. Subsequent cycles of refinement included the modeling of residues in alternate conformations and modeling of an ambiguously oxidized cysteine at residue 152 as a cysteine sulfinic acid in dual conformations. Riding hydrogen atoms were added to all atoms except histidine Nε1 and Nδ2 atoms, tyrosine O atoms, threonine Oγ2 atoms, and serine Oγ atoms, owing to the hydrogen positional ambiguity for these atoms.

In the 1.05-Å resolution 277 K model, the occupancy of the partially occupied water molecule near the N5 atom of folate and the two conformations of the Met20 residue were all fixed to 0.5. This approach is consistent with correlated disorder of the Met20 side chain and the water molecule, which is reasonable given the close contact distance (1.7 Å) between the water and Cε atom of the conflicting conformation of Met20. In the final 10 cycles of refinement, the test set and the working set of reflections were combined for both the 100 K and 277 K datasets.

1. Cummins PL, Gready JE (2001) Energetically most likely substrate and active-site protonation sites and pathways in the catalytic mechanism of dihydrofolate reductase. *J Am Chem Soc* 123(15):3418–3428.
2. Bystroff C, Oatley SJ, Kraut J (1990) Crystal structures of *Escherichia coli* dihydrofolate reductase: The NADP$^+$ holoenzyme and the folate/NADP$^+$ ternary complex. Substrate binding and a model for the transition state. *Biochemistry* 29(13):3263–3277.
3. Reyes VM, Sawaya MR, Brown KA, Kraut J (1995) Isomorphous crystal structures of *Escherichia coli* dihydrofolate reductase complexed with folate, 5-deazafolate, and 5,10-dideazatetrahydrofolate: Mechanistic implications. *Biochemistry* 34(8): 2710–2723.
4. Gandour RD (1981) On the importance of orientation in general base catalysis by carboxylate. *Bioorg Chem* 10:169–176.
5. Bennett BC, Xu H, Simmerman RF, Lee RE, Dealwis CG (2007) Crystal structure of the anthrax drug target, *Bacillus anthracis* dihydrofolate reductase. *J Med Chem* 50(18): 4374–4381.
6. Wan Q, et al. (2014) Preliminary joint X-ray and neutron protein crystallographic studies of ecDHFR complexed with folate and NADP$^+$. *Acta Crystallogr F Struct Biol Commun* 70(Pt 6):814–818.
7. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M (2006) HKL-3000: The integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* 62(Pt 8):859–866.
8. Cameron CE, Benkovic SJ (1997) Evidence for a functional role of the dynamics of glycine-121 of *Escherichia coli* dihydrofolate reductase obtained from kinetic analysis of a site-directed mutant. *Biochemistry* 36(50):15792–15800.
9. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
10. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
11. Sawaya MR, Kraut J (1997) Loop and subdomain movements in the mechanism of *Escherichia coli* dihydrofolate reductase: Crystallographic evidence. *Biochemistry* 36(3): 586–603.
12. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
13. Pintar A, Carugo O, Pongor S (2003) DPX: For the analysis of the protein core. *Bioinformatics* 19(2):313–314.
14. Bond CS (2003) TopDraw: A sketchpad for protein structure topology cartoons. *Bioinformatics* 19(2):311–312.
15. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–763.
16. Lee B, Richards FM (1971) The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol* 55(3):379–400.
17. Sheldrick GM, Schneider TR (1997) SHELXL: High-resolution refinement. *Methods Enzymol* 277:319–343.

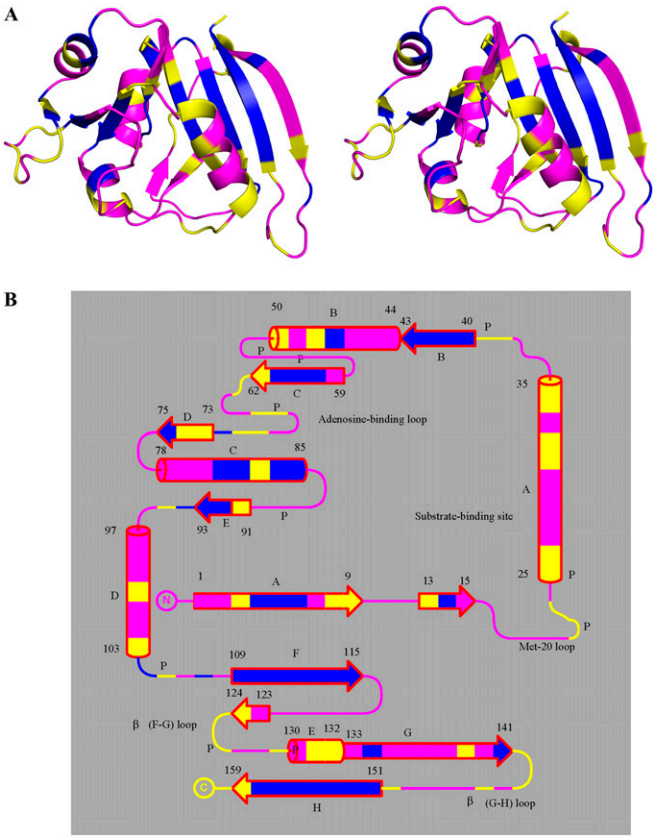

Fig. S1. HDX analysis of ecDHFR-folate-NADP$^+$. (*A*) Stereoview cartoon diagram of residues colored based on the occupancy of backbone amide HDX. (*B*) Topology diagram of backbone amide HDX pattern for the ternary complex. Nonexchanged (0–0.15) amides are represented in blue; partially exchanged (0.16–0.69) amindes, in yellow; completely exchanged (0.7–1.00) amides, in magenta.

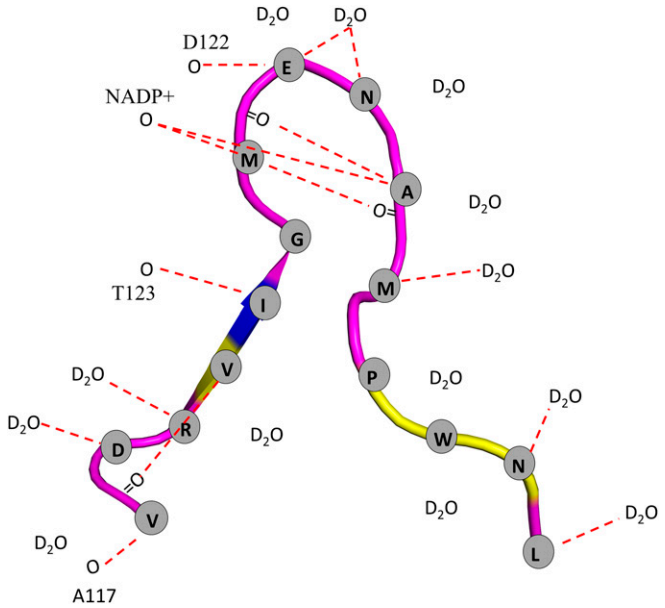

Fig. S2. HDX pattern of the backbone amides of the Met20 loop. The loop color is the same as that in Fig. S1. Waters within 4 Å but not forming H bonds with the backbone amides are also shown here without dashed lines.

Table S1. Data collection and structure refinement statistics

Variable	Neutron 291 K	X-ray 291 K	X-ray 100 K	X-ray 277 K
Data collection				
Ligands	Folate $NADP^+$	Folate $NADP^+$	Folate $NADP^+$	Folate $NADP^+$
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions				
a, b, c, Å	34.3, 45.7, 98.9	34.3, 45.6, 99.0	34.0, 44.8, 98.3	34.3, 45.5, 98.7
α, β, γ, °		$\alpha = \beta = \gamma = 90$		
Resolution, Å	2.0	1.6	0.85	1.05
Unique reflections	8,745	20,795	130,200	71,782
Redundancy	7.5 (5.8)	4.1 (4.0)	5.5 (4.0)	9.9 (9.1)
Completeness, %	79.3 (61.3)*	97.7 (94.0)*	98.2 (97.2)	98.3 (97.2)
R_{sym} †	0.188 (0.305)*	0.097 (0.446)*	0.049 (0.550)	0.045 (0.458)
$I/\sigma(I)$	5.3 (3.2)*	13.7 (2.97)*	36.3 (2.3)	40.8 (5.2)
Refinement				
R_{work} ‡/R_{free} §	0.2310/0.2714	0.1937/0.2178	0.128/0.112	0.134/0.108
No. of atoms				
Protein	2,799		1,779	1,716
$NADP^+$	80		48	48
Folate	54		32	32
Mn^{2+}	2		2	2
Water	357		356	230
B factors				
Protein	18.1		9.3	14.6
$NADP^+$	12.2		8.0	10.6
Folate	17.6		11.7	18.4
Mn^{2+}	37.1		7.5	13.0
Water	34.5		25.9	38.6
Deviation from ideality				
Bond length, Å	0.005		0.014	0.012
Bond angle, °	1.04		2.2	2.1
Ramachandran plot statistics, %				
Most favorable	99.36		98.6	99.5
Allowed	0.64		1.4	0.5
Disallowed	0		0	0
PDB ID code	4PDJ		4PSY	4RGC

*Numbers in parentheses represent values in the highest-resolution shell.

†$R_{sym} = \sum(|Ii - \langle I \rangle|)/\sum(I)$, where Ii is the measured intensity and $\langle I \rangle$ is the mean intensity of all measured observations equivalent to reflection Ii.

‡$R_{work} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$, where $|F_{obs}|$ is the observed diffraction amplitude and $|F_{calc}|$ is the corresponding calculated structure factor amplitude.

§R_{free} is defined by R_{work}, involving 5% of the measured reflections not used in refinement and set aside for cross-validation purposes.

Table S2. Refined deuterium occupancies of backbone amides for ecDHFR-folate-NADP$^+$

Residue	Occupancy	H/D exchange	Secondary structure
1	1.00	D**	B
2	0.97	D**	B
3	0.15	D*	B
4	0	H	B
5	0	H	B
6	0.04	H	B
7	0.97	D**	B
8	0.16	D*	B
9	0.19	D*	B
10	0.95	D**	H
11	0.74	D**	H
12	0.68	D**	H
13	0.50	D*	B
14	0	H	B
15	0.78	D**	B
16	1.00	D**	L
17	0.84	D**	L
18	1.00	D**	L
19	0.90	D**	L
20	0.92	D**	L
21	N/A	P	L
22	0.64	D*	L
23	0.69	D*	L
24	1.00	D**	L
25	N/A	P	H
26	0.65	D*	H
27	0.95	D**	H
28	0.88	D**	H
29	0.89	D**	H
30	0.74	D**	H
31	0.17	D*	H
32	0.59	D*	H
33	0.87	D**	H
34	0.28	D*	H
35	0.49	D*	H
36	0.84	D**	L
37	0.97	D**	L
38	0.30	D*	L
39	N/A	P	L
40	0.08	H	B
41	0.03	H	B
42	0	H	B
43	0.02	H	B
44	1.00	D**	H
45	1.00	D**	H
46	0.76	D**	H
47	0	H	H
48	0.41	D*	H
49	0.83	D**	H
50	0.36	D*	H
51	0.98	D**	L
52	0.92	D**	L
53	N/A	P	L
54	0.87	D**	L
55	N/A	P	L
56	0.91	D**	L
57	1.00	D**	L
58	0.98	D**	L
59	0.87	D**	B
60	0.04	H	B
61	0	H	B
62	0.24	D*	B

Table S2. Cont.

Residue	Occupancy	H/D exchange	Secondary structure
63	0.28	D*	B
64	0.64	D**	L
65	0.77	D**	L
66	N/A	P	L
67	0.64	D**	L
68	1.00	D**	L
69	0.70	D**	L
70	0.20	D*	L
71	0.33	D*	L
72	0	H	L
73	0.31	D*	B
74	0.74	D*	B
75	0	H	B
76	1.00	D**	L
77	0.87	D**	L
78	1.00	D**	H
79	0.92	D**	H
80	1.00	D**	H
81	0	H	H
82	0.11	H	H
83	0.50	D*	H
84	0	H	H
85	0	H	H
86	0.87	D**	L
87	0.73	D**	L
88	1.00	D**	L
89	N/A	P	L
90	1.00	D**	L
91	0.23	D*	B
92	0	H	B
93	0	H	B
94	0	H	L
95	0.37	D*	L
96	0.95	D**	L
97	0.91	D**	H
98	0.90	D**	H
99	0.79	D**	H
100	0.17	D*	H
101	1.00	D**	H
102	0.79	D**	H
103	0.25	D*	H
104	0.07	H	L
105	N/A	P	L
106	0.96	D**	L
107	0	H	L
108	1.00	D**	L
109	0.03	H	B
110	0	H	B
111	0.01	H	B
112	0	H	B
113	0	H	B
114	0.16	D*	B
115	0.08	H	B
116	0.74	D**	L
117	1.00	D**	L
118	0.87	D**	L
119	0.76	D**	L
120	0.80	D**	L
121	0.86	D**	L
122	0.83	D**	L
123	0.79	D**	B
124	0.42	D*	B

Table S2. Cont.

Residue	Occupancy	H/D exchange	Secondary structure
125	0.67	D*	L
126	N/A	P	L
127	0.82	D**	L
128	0.95	D**	L
129	0.46	D*	L
130	N/A	P	H
131	0.59	D*	H
132	0.70	D**	H
133	0.84	D**	B
134	0	H	B
135	0.90	D**	B
136	0.81	D**	B
137	1.00	D**	B
138	0.94	D**	B
139	0.60	D*	B
140	0.75	D**	B
141	0	H	B
142	0.76	D**	L
143	0.86	D**	L
144	1.00	D**	L
145	0.66	D*	L
146	1.00	D**	L
147	0.85	D**	L
148	0.97	D**	L
149	0.94	D**	L
150	0.63	D*	L
151	0	H	B
152	0.08	H	B
153	0.03	H	B
154	0	H	B
155	0	H	B
156	0	H	B
157	0	H	B
158	0.35	D*	B
159	0.55	D*	L

D*, refined deuterium occupancy 0.16–0.69 (partially exchanged); D**, refined deuterium occupancy $\geq$ 0.7 (fully exchanged); H, refined deuterium occupancy 0–0.15 (not exchanged); *n*, disordered main chain; N/A, deuterium occupancy value could not be refined; P, proline (no exchangeable atom on backbone amide nitrogen); SS, secondary structure type (B, β strand; H, α heli; L, loop).

Table S3. Analysis of DHFR dynamics based on solvent accessible area, atomic depth, and H/D exchange of backbone amides: Comparison with the DHFR-MTX neutron structure (2INQ)

	Mon A (2INQ); partially occluded	Mon B (2INQ); closed	4PDJ; closed
Solvent accessible area, Å^2			
Overall	6,304	5,849	8,655
Met20 loop	495	382	943
Substrate-binding region	348	280	599
Atomic depth, Å			
Overall	1,203	1,253	1,158
Met20 loop	67	82	55
Substrate-binding region	82	75	80
H/D exchange, %*			
Overall	95 (65.1)	81 (54.7)	112 (75.2)
Met20 loop	13 (86.7)	10 (66.7)	15 (100)
Substrate-binding region	5 (45.5)	5 (45.5)	11 (100)

*Numbers in parenthesis represent the percentage of fully or partially H/D-exchanged residues.