

Supporting Information for:

Aspects of Weak Interactions between Folate and Glycine Betaine

Purva P. Bhojane[‡], Michael R. Duff, Jr. [‡], Khushboo Bafna[¥], Gabriella Rimmer[‡],

Pratul Agarwal^{‡,¥,‡} and Elizabeth E. Howell^{‡,¥*}

[‡]Department of Biochemistry, Cellular, and Molecular Biology, University of Tennessee, Knoxville, TN 37996-0840, [¥]Genome Science and Technology Program, University of Tennessee, Knoxville, TN 37996-0840, [‡]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

Materials and Methods

Materials Betaine, folic acid, indole acetate, *m*-aminobenzoate, *o*-aminobenzoate, *p*-toluic acid, *p*-aminobenzoate-glutamate, pyrrole-2-carboxylate, adenosine 5'-monophosphate, guanosine 5'-monophosphate, cytidine 2'-monophosphate and thymidine 5'-monophosphate were purchased from Sigma-Aldrich. Nicotinic acid, nicotinamide, pyrimidone, pyridoxine-HCl were from Acros Organics, *p*-aminobenzoate was from MP Biomedicals, and phenylalanine-HCl was from Fisher Scientific. Uridine 3'-monophosphate was from Chem-Impex International Inc. Pteroyltetra- γ -L-glutamate (PG4) was from Schircks Laboratories.

Protein Purification R67 DHFR was expressed and purified as described previously.¹ Briefly, ammonium sulfate precipitation and ion-exchange column chromatography were used to purify the protein to homogeneity. EcDHFR was expressed and purified as published earlier.² His-tagged protein was purified using two affinity chromatography columns - a nickel-NTA column followed by a methotrexate (MTX) affinity column. Elution of EcDHFR from the MTX affinity column required addition of folate, which was subsequently removed with a DEAE column. Purified samples were dialyzed against distilled, deionized water and then lyophilized. Protein concentrations were determined using a bicinchoninic acid (BCA) (Pierce) assay.

Folate dimerization at pH 10 by NMR spectroscopy A 1D proton NMR experiment was performed as described by Duff et al.³ to study dimerization of deprotonated folate. Stock solutions of folate were prepared in 10 mM deuterated Tris (pD 10) with and without 20% deuterated betaine. An NMR sample with 300 mM folate was prepared at pH 10 and the spectrum was recorded. The sample was diluted with buffer and the same procedure was repeated until a folate concentration of 0.5 mM was reached. Spectra were recorded on a Varian

500 MHz NMR spectrometer with a pulse length of 3.7 μ s using 16 scans from 14 to -0.5 parts per million (ppm) per spectrum. Data analysis used MestreNova version 10.0 (Mestrelab Research, Compostela, Spain).⁴ The spectra were phase and baseline corrected, and the peaks were referenced to the water peak (chemical shift for water, 4.80 ppm). The proton chemical shifts were fit to a dimerization equation as described previously.³ Similar NMR experiments and analyses were done in the presence of 20% deuterated betaine.

Simulation of folate in betaine Computer simulations of folate surrounded by betaine and water were performed using the AMBER simulations package.⁵ A folate molecule was placed in the center of a periodic box surrounded by betaine and water (SPC/E water model). Betaine, or trimethyl-glycine, was modeled using AMBER's parm 14SB force-field; the aliphatic carbons, hydrogens and nitrogen were parameterized using the lipid related parameters while the remaining atoms were parameterized based on glycine.⁶ The charges for betaine were calculated using a procedure similar to that used for the folate molecule. The ratio of folate to betaine molecules was 1:76, corresponding to a 1.35 M concentration of betaine in a periodic box of 46.74 Å \times 49.64 Å \times 50.12 Å. The initial placement of betaine around folate was performed using PackMol software⁷, followed by filling the remaining space with SPC/E water using AMBER's xleap module. The prepared systems were slowly equilibrated using a procedure developed in our group, and described previously.⁸ The production run of 200 nanoseconds was performed for each system at 300 K in an NVE ensemble using a 2 femtosecond time-step. A total of 200 conformations (every 1 ns) were used for analysis.

Isothermal Titration Calorimetry (ITC) Binding affinities, stoichiometries and enthalpies were determined using either a Nano-ITC (TA Instruments) or a VP-ITC (Microcal). For studies with R67 DHFR, binding of folate or pteroyltetra- γ -L-glutamate (PG4 from Schircks Laboratories) to

a 1:1 R67 DHFR- NADPH complex was monitored. Titrations were performed in duplicate at 13 °C and pH 8.0 to minimize catalysis. The R67 DHFR concentration was 100- 150 μM and the buffer was MTA (100 mM MES, 50 mM Tris and 50 mM acetic acid) pH 8.0. The ligand concentration ranged from 1.2 – 1.4 mM for the experiments with no betaine and from 1.8 mM- 1.95 mM for experiments with 5 % and 10 % betaine in the MTA buffer. The time between injections was 240-300 seconds, allowing for baseline equilibration. The software supplied by the manufacturer was initially used for analysis. The data were then exported into SEDPHAT; this program allows global fitting of replicate data sets.⁹ A single sites model ($A + B \leftrightarrow AB$) was used for the fitting process. Similar experiments were performed with binding of folate and PG4 to EcDHFR in MTA buffer pH 7.0 at 25 °C. EcDHFR concentrations ranged from 10-15 μM. Folate and PG4 concentrations ranged from 350 – 550 μM for titrations in the absence of betaine. The folate concentration for binding to EcDHFR in presence of 10 % and 20 % betaine was in the range of 600-850 μM. The “c value” ($= [P_{total}] / K_d$) ranged from 1-10, within the suggested values of 1-1000.¹⁰

Comparison of $\Delta\mu_{23}/RT$ predictions of ligand binding with ITC Data To determine how accurately the predicted μ_{23}/RT values reflect experimental data, $\Delta\mu_{23}/RT$ values were calculated for ligands binding to the two DHFR enzyme types. Using data previously obtained by ITC, the $\Delta\mu_{23}/RT$ for binding can be calculated from the slopes of $\ln(K_a)$ versus molality using Eq (1) from Guinn *et al.*¹¹

$$-\frac{\ln K_a}{m_3} = \frac{\Delta\mu_{23}}{RT} \quad \text{Eq (S1)}$$

where K_a is the association constant and m_3 is molality of betaine. Predicted μ_{23}/RT values were calculated for the apo-proteins or protein-ligand complexes using Eq (4) in the main text with a

Python script. Waters were removed from the PDB file and the surface areas of each of the atom types were calculated using SurfaceRacer.¹² The product of the atomic surface areas and the corresponding atom-type α value were summed to obtain the predicted μ_{23}/RT . The μ_{23}/RT values for the ligands in their bound conformation were calculated in a similar manner. To calculate the $\Delta\mu_{23}/RT$ for the binary protein-ligand complexes, the sum of the μ_{23}/RT s for the apo-protein and the unbound ligand was subtracted from the μ_{23}/RT of the complex. To obtain the $\Delta\mu_{23}/RT$ for ternary complexes, the μ_{23}/RT of the ternary complex was subtracted from the μ_{23}/RT of the binary protein-ligand complex plus the unbound ligand.

Results

Folate dimerization at pH 10 NMR experiments at pH 10 noted the change in chemical shifts for the pteridine (C7H), C9H, benzoyl ring protons (C2'H/C6'H, C3'H/C5'H) with increasing folate concentration while the glutamate proton shifts were unchanged (see Figure S1 for numbering of atoms). The data and fits for each of the proton chemical shifts are shown in Figure S2. Fitting the sum of the C7, C9, C3'/C5', and C2'/C6 proton chemical shifts with no betaine to a dimerization equation yielded a K_d of 960 ± 140 mM for folate at pH 10 (Table S2). As the K_d was higher than the highest folate concentration used for the experiment, it suggests that folate dimerization at pH 10 is very weak and its K_d cannot be accurately determined. Higher concentrations of folate cannot be achieved because of limited solubility. Although the K_d s obtained were much higher than the folate concentrations used for the experiments, we can qualitatively see a trend for a lower K_d in the presence of betaine.

Betaine-imidazole interaction by VPO Imidazole, a small molecule, has a pK_a of 6.5;¹³ it also

dimerizes with a K_d of 1 mM for the protonated form and a K_d of 33 μM for the deprotonated form.¹⁴ VPO studies of imidazole at pH 4 showed a slightly negative μ_{23}/RT value, whereas at pH 10, its μ_{23}/RT was near zero and had large errors. Supplemental Figure S3 shows the data for 250-270 mm of imidazole. As the change in μ_{23}/RT values for protonated and deprotonated imidazole was not large, scatter was observed in the data, and dimerization was a concern, we did not analyze these data.

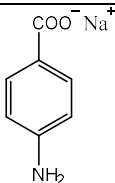
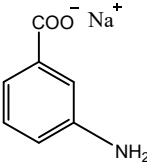
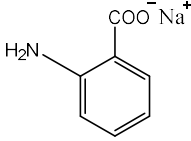
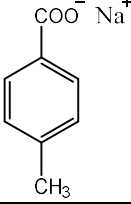
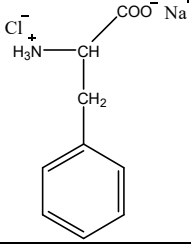
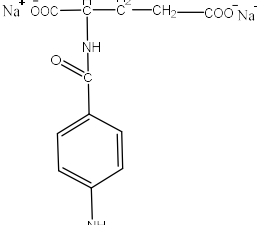
Folate MD simulation in betaine The simulation of folate in betaine and water yielded similar results as the simulation of folate in water alone. A range of conformations were obtained and a μ_{23}/RT value was calculated for each conformation. The average μ_{23}/RT value for folate in betaine plus water was $-0.019 \pm 0.045 m^{-1}$, which is in the same range for that of folate with water only ($-0.003 \pm 0.05 m^{-1}$). The overlay plot of predicted μ_{23}/RT values obtained for folate conformers sampled during the course of each simulation shown in Supplemental Figure S4B suggests betaine has no significant effect on the conformational sampling of folate.

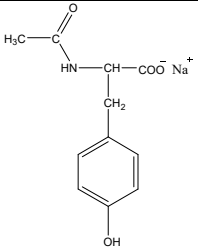
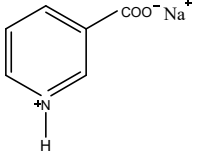
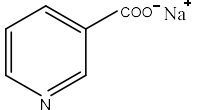
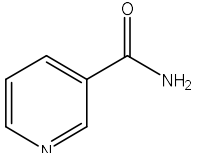
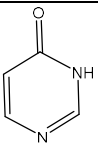
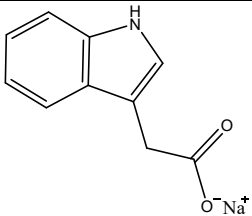
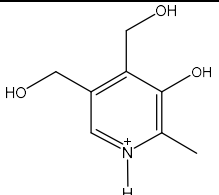
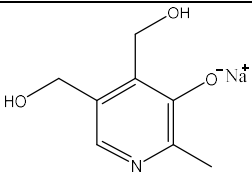

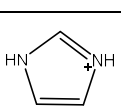
Binding of folate and PG4 to R67 DHFR and EcDHFR α values can be used to predict ligand-osmolyte interactions. Can this information be used to predict effects on ligand binding to proteins? The caveat is whether all surfaces of the ligand are used in the binding interaction. For example, we consider the case of folate polyglutamylation. As glutamate excludes betaine, addition of extra glutamates to folate (extended conformer) increases the predicted μ_{23}/RT value from $-0.09 \pm 0.04 m^{-1}$ to $1.22 \pm 0.04 m^{-1}$ for pteroyltetra- γ -L-glutamate (PG4). If the polyglutamate tail is involved in binding to DHFR, this increase in μ_{23}/RT predicts lesser osmotic stress effects. However, our ITC experiments found betaine addition weakens binding of folate or PG4 to R67 DHFR (see Figure S6A and Table S5). Thus use of a calculated μ_{23}/RT

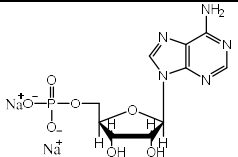
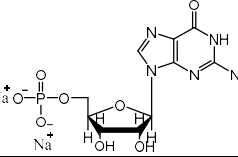
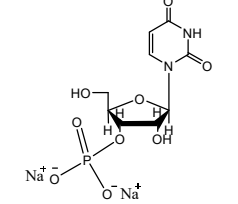
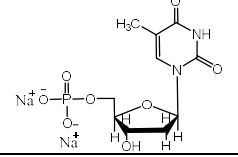
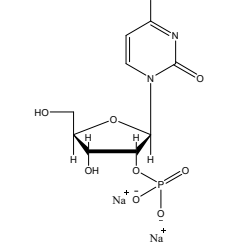
value for a ligand is not sufficient to predict effects of betaine on binding.

While many folate pathway enzymes show tighter binding to polyglutamylated folate redox states, we did not find any information addressing this issue in EcDHFR. Thus we measured the affinity for folate to EcDHFR and found it also decreased linearly with increasing betaine concentration. The affinity for PG4 binding to EcDHFR was similar to that of folate (see Figure S6B). These results predict that the additional glutamates will not contribute to binding to EcDHFR. Thus an important parameter in predicting betaine effects on binding is whether all the ligand atoms are used in the interaction. It should be noted that osmolytes could also potentially interact with the protein. If osmolytes were to interact with the protein, especially the ligand binding surface, then additional complications would arise in predicting osmolyte effects on ligand binding.

Table S1. A list of compounds tested for preferential interactions with betaine by the VPO method. Structures with correct protonation states are shown along with the source from which each of the structures were obtained.

Compound	Structure	Source
<i>p</i> -Amino-benzoate		MOE ^a
<i>m</i> -Amino-benzoate		MOE ^a
<i>o</i> -Amino-benzoate		MOE ^a
<i>p</i> -Toluic acid		MOE ^a
Phenylalanine-HCl		BMRB ^b
<i>p</i> -Amino-benzoyl-Glutamate		PDB 3DL6 ^c

N-acetyl-tyrosine		MOE ^a
Nicotinic acid (Protonated)		BMRB ^b
Nicotinic acid (Deprotonated)		BMRB ^b
Nicotinamide		BMRB ^b
Pyrimidone		MOE ^a
Indole Acetate monosodium salt		BMRB ^b
Pyridoxine- HCl (protonated)		BMRB ^b
Pyridoxine (deprotonated)		BMRB ^b
Imidazole (deprotonated)		MOE ^a
Imidazole (protonated)		MOE ^a

5'AMP monosodium salt		PDB 12AS ^c
5'GMP disodium salt		PDB 5C46 ^c
3'UMP disodium salt		PDB 4J7L ^c
5'dTMP disodium salt		BMRB ^b
2'CMP disodium salt		PDB 1ROB ^c

^a built using MOE (versions 2012.10 and 2015.1001, Chemical Computing Group, Montreal, QC). ^b obtained from the Biological Magnetic Resonance Bank, (<http://www.bmrwisc.edu/>). ^c obtained from the Protein Data Bank (PDB, <http://www.rcsb.org>)

Table S2. Dimerization constants obtained from fitting the concentration dependence of chemical shifts obtained by NMR as described in Duff et al.³ Data and fits shown in Supplementary Figure S2.

Chemical shifts	No betaine	20 % betaine
	K_d (mM)	K_d (mM)
Sum of all proton chemical shifts	960 ± 140	710 ± 60
C7H	2200 ± 260	980 ± 60
C2'H	1800 ± 370	830 ± 90
C6'H	1500 ± 440	900 ± 90
C3'H	860 ± 80	510 ± 60
C5'H	880 ± 90	590 ± 80
C9H	1100 ± 140	700 ± 60

Table S3. A list of experimental and predicted μ_{23}/RT values for test compounds in addition to compounds listed in Table 1 in the main text. These compounds were not used in our fits to obtain α values.

Compound	Experimental μ_{23}/RT (m^{-1})	Predicted μ_{23}/RT (m^{-1})	pH	Reason for not including in α value fit
<i>p</i> -Toluic acid	-0.46 ± 0.04	-0.21 ± 0.03	7	Outlier on plot of experimental vs. predicted μ_{23}/RT values
<i>p</i> -Amino-benzoyl-glutamate	0.50 ± 0.02	-0.12 ± 0.02	7	Changed R^2 of fit from 0.93 to 0.86
N-acetyl-tyrosine	-0.66 ± 0.03	-0.05 ± 0.02	4	Low solubility \rightarrow low concentration
Quinolinic acid	0.86 ± 0.04	0.20 ± 0.01	7	Low solubility \rightarrow low concentration
Imidazole	-0.07 ± 0.01	-	4	Dimerization
Imidazole	0.01 ± 0.01	-	10	Dimerization

Table S4. Calculation of K_p partition coefficient values for the different atom types.

Atom Type	K_p
Aliphatic C	1.08 ± 0.03^a
Hydroxyl O	0.81 ± 0.03^a
Amide O	0.11 ± 0.05^b
Amide N	1.89 ± 0.05^a
Carboxylate O	0.24 ± 0.03^a
Cationic N	1.37 ± 0.03^a
Aromatic C	1.83 ± 0.03^a
Phosphate O	0.13 ± 0.04^c
Amine N off Aromatic rings	2.43 ± 0.08^a
Aromatic N	0.27 ± 0.08^a

^a b_i is set to 0.18, which is approximately 2 layers of water.¹¹ ^b b_i is set to 0.27, which gives a K_p value greater than 0. $b_i = 0.27$ is equivalent to three layers of water surrounding the amide oxygen atoms ^c b_i is set to 0.27, or three layers of water surrounding the phosphate oxygens, in accordance with Capp et. al.¹⁵

Table S5. Thermodynamic parameters obtained for binding of folate and PG4 to the R67 DHFR-NADPH binary complex and EcDHFR by ITC. The osmolalities of the buffer with and without betaine are also listed.

Complex	Buffer and/or osmolyte addition	K_d (μM)	ΔG (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)	n	Osmolality (Osm)
Folate binding to R67DHFR-NADPH	MTA pH 8	23 ± 5	-6.1 ± 0.1	-8.1 ± 0.8	-1.9	0.60 ± 0.1	0.25
	MTA pH 8 + 5 % betaine	46 ± 16	-5.7 ± 0.2	-5.6 ± 1.2	0.12	0.58 ± 0.1	0.84
	MTA + 10 % betaine	52 ± 27	-5.6 ± 0.4	-2.0 ± 0.6	3.6	0.68 ± 0.1	1.25
PG4 binding to R67DHFR-NADPH	MTA pH 8	16 ± 6	-6.3 ± 0.3	-6.8 ± 1	-0.52	0.61 ± 0.1	0.25
	MTA pH 8 + 5 % betaine	29 ± 10	-5.9 ± 0.2	-4.4 ± 0.8	1.6	0.58 ± 0.1	0.84
	MTA pH 8 + 10 % betaine	37 ± 13	-5.8 ± 0.3	-1.3 ± 0.2	4.5	0.55 ± 0.1	1.25
Folate binding to EcDHFR	MTA pH 7	2.9 ± 1.7	-7.6 ± 0.3	-9.0 ± 2.3	-1.5	0.80 ± 0.1	0.19
	MTA pH 7 + 10 % betaine	7.1 ± 1.8	-7.1 ± 0.2	-9.6 ± 1.3	-2.5	0.93 ± 0.1	1.15
	MTA pH 7 + 20% betaine	13 ± 4	-6.7 ± 0.2	-9.7 ± 2.5	-2.9	0.83 ± 0.1	2.05
PG4 binding to EcDHFR	MTA pH 7	2.6 ± 0.7	-7.6 ± 0.2	-8.9 ± 0.9	-1.4	0.70 ± 0.1	0.19

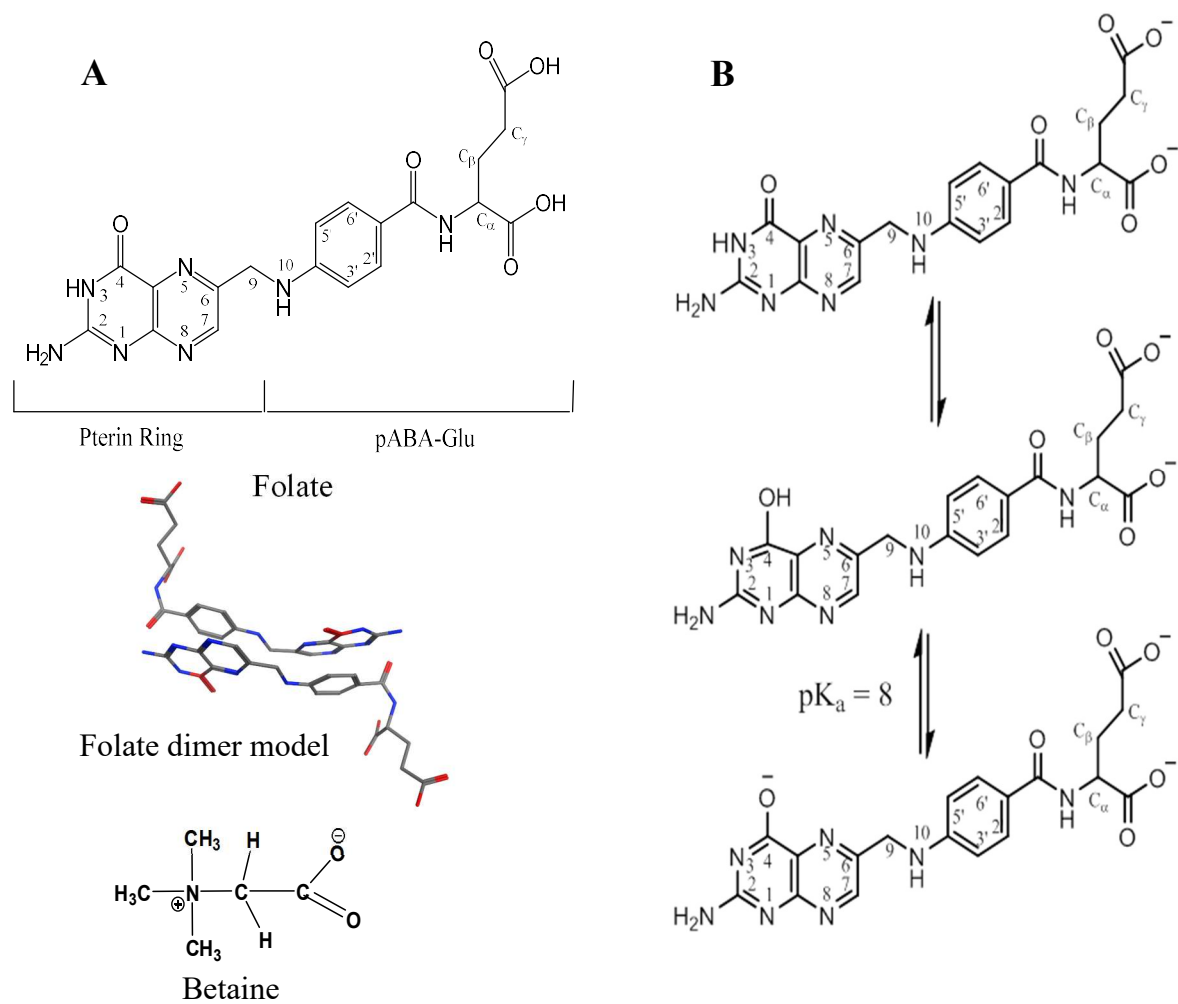


Figure S1. Panel A shows the structures of folate with atom numbers, a model for the folate dimer and the betaine structure. Panel B shows the keto-enol tautomerization and deprotonation of the N3-O4 group of the pterin ring of folate.

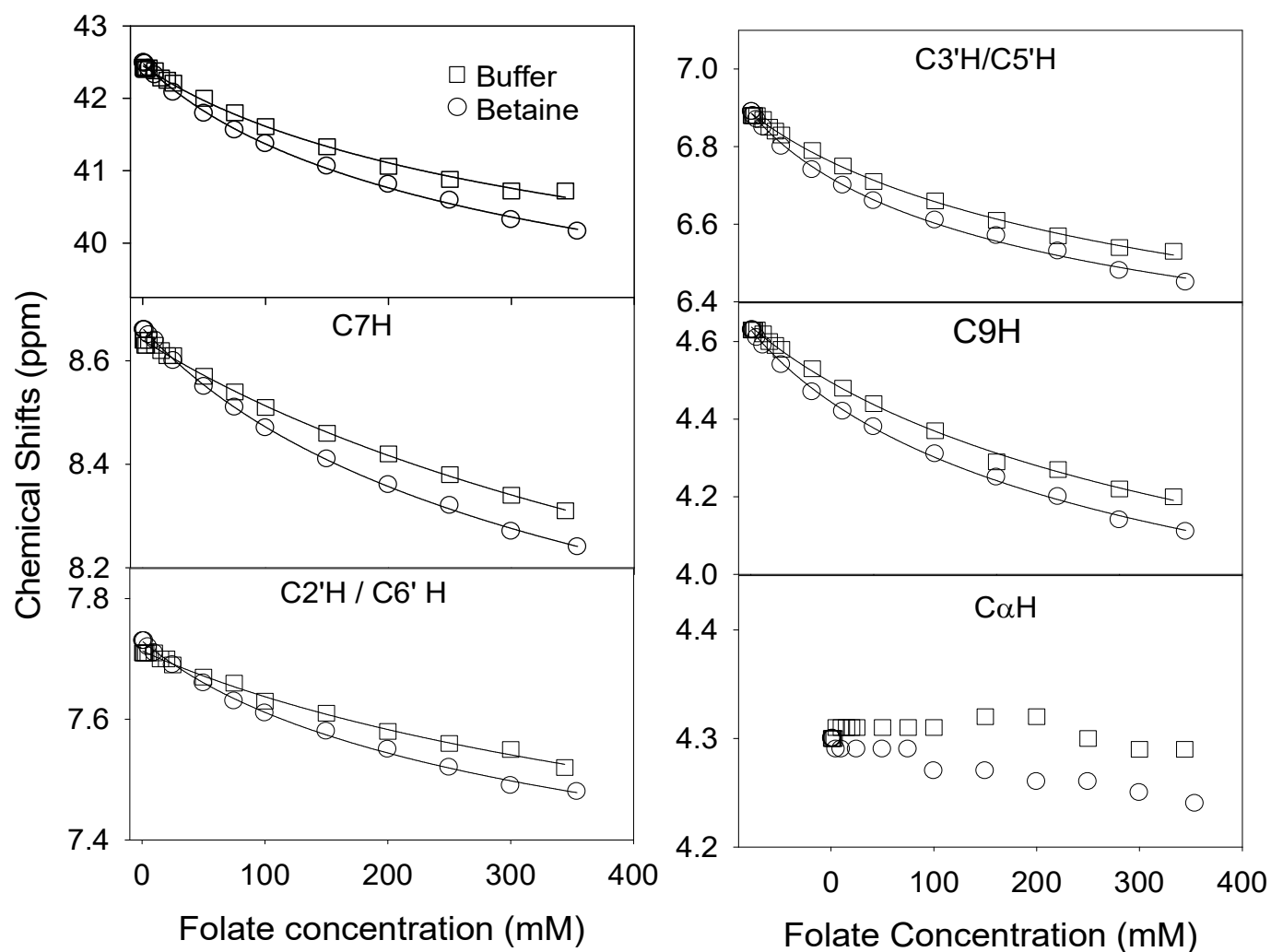


Figure S2. Folate dimerization at pH 10. The panels show chemical shifts noted for folate protons (numbered as in Figure S1) in 10 mM deuterated Tris (\square), 10 mM deuterated Tris with 20 % deuterated betaine (\circ). The lines are the fits to the dimerization equation in Duff et al.³

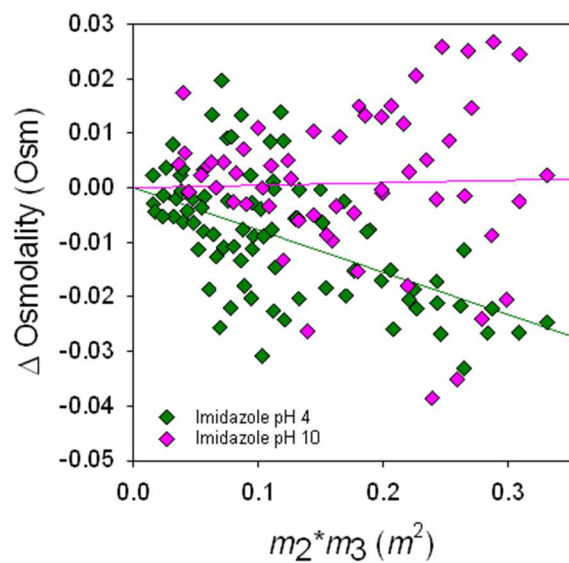


Figure S3. Vapor pressure osmometry studies of imidazole. Data for imidazole at pH 4 (green diamonds) and pH 10 (magenta diamonds). The concentration of imidazole used in the experiment was 250 *mm*.

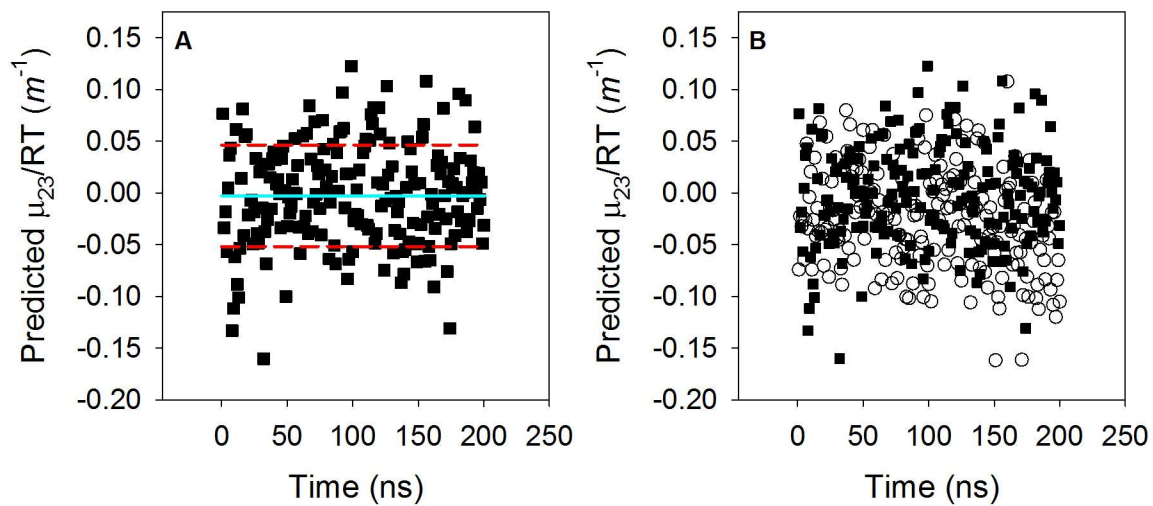


Figure S4. Predicted μ_{23}/RT values for folate associated with the MD simulations of folate in water with betaine (■). The average of the μ_{23}/RT values is shown by a cyan line. The red dashed lines show one standard deviation from the average value. Panel B shows predicted μ_{23}/RT values from the frames of simulations of folate in just water (○) and in water with 1.35 M betaine (■).

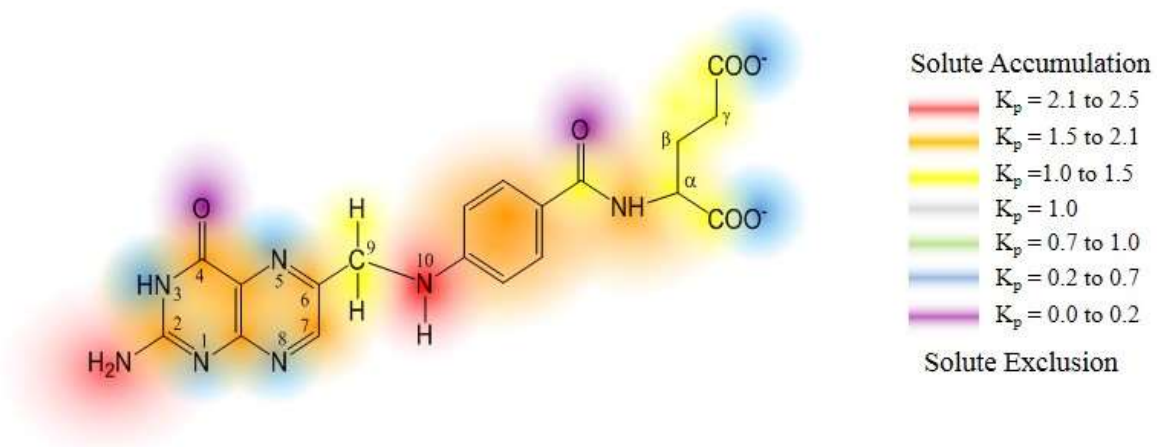


Figure S5. A representation of K_p values for each atom type in folate is shown. Aromatic carbons, amide nitrogen, amine nitrogens off aromatic rings and aliphatic carbons accumulate betaine whereas aromatic nitrogens, carboxylate oxygens and amide oxygen exclude betaine.

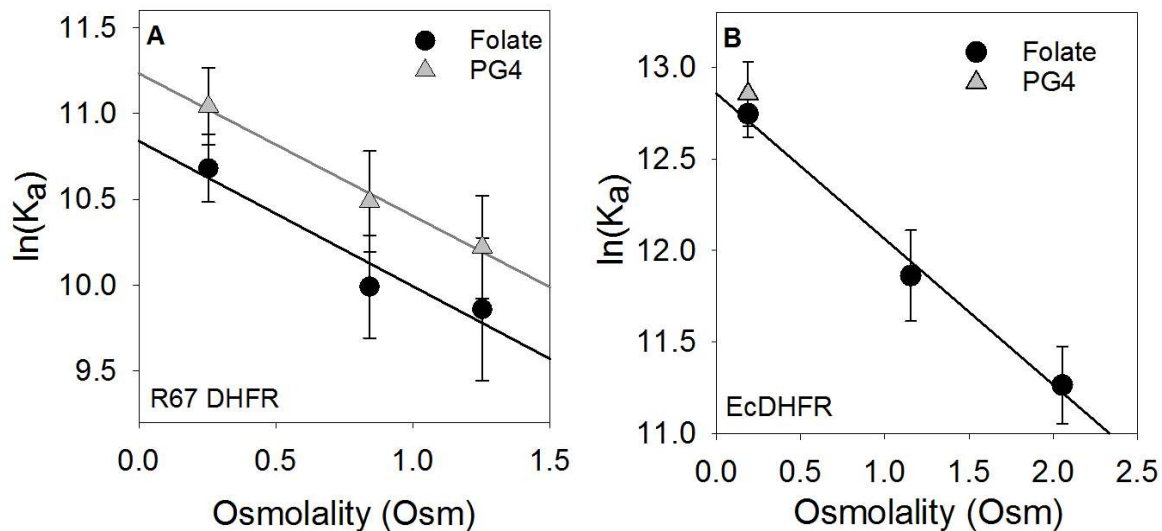


Figure S6. The effect of osmolality on the binding affinities of folate (●) and PG4 (▲) to DHFRs. Panel A plots $\ln K_a$ vs. osmolality for folate and PG4 binding to R67 DHFR-NADPH. The negative slopes indicate weaker binding of both folate and PG4 in the presence of betaine. Panel B shows similar results obtained for EcDHFR. No significant difference can be noted in the binding affinities of folate and PG4 to EcDHFR, predicting no effect of polyglutamylation on binding.

References

1. Reece, L. J., Nichols, R., Ogden, R. C., and Howell, E. E. (1991) Construction of a synthetic gene for an R-plasmid-encoded dihydrofolate reductase and studies on the role of the N-terminus in the protein, *Biochemistry* 30, 10895-10904.
2. Grubbs, J., Rahmanian, S., DeLuca, A., Padmashali, C., Jackson, M., Duff, M. R., Jr., and Howell, E. E. (2011) Thermodynamics and solvent effects on substrate and cofactor binding in *Escherichia coli* chromosomal dihydrofolate reductase, *Biochemistry* 50, 3673-3685.
3. Duff, M. R., Jr., Grubbs, J., Serpersu, E., and Howell, E. E. (2012) Weak interactions between folate and osmolytes in solution, *Biochemistry* 51, 2309-2318.
4. Willcott, M. R. (2009) MestRe Nova, *J. Am. Chem. Soc.* 131, 13180-13180.
5. Case, D. A., Babin, V., Berryman, J. T., Betz, R. M., Cai, Q., Cerutti, D. S., T.E. Cheatham, I., Darden, T. A., R.E.Duke, Gohlke, H., Goetz, A. W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolosváry, I., Kovalenko, A., Lee, T. S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K. M., Paesani, F., Roe, D. R., Roitberg, A., Sagui, C., Salomon-Ferrer, R., Seabra, G., Simmerling, C. L., W. Smith, J. S., Walker, R. C., Wang, J., Wolf, R. M., X.Wu, and Kollman, P. A. (2014) AMBER14, University of California, San Francisco.
6. Ma, L., Pegram, L., Record Jr, M. T., and Cui, Q. (2010) Preferential interactions between small solutes and the protein backbone: a computational analysis, *Biochemistry* 49, 1954-1962.
7. Martinez, L., Andrade, R., Birgin, E. G., and Martinez, J. M. (2009) PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations, *J. Comput. Chem.* 30, 2157-2164.
8. Ramanathan, A., Savol, A. J., Langmead, C. J., Agarwal, P. K., and Chennubhotla, C. S. (2011) Discovering conformational sub-states relevant to protein function, *PLoS One* 6, e15827.
9. Houtman, J. C., Brown, P. H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L. E., and Schuck, P. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling, *Protein Sci.* 16, 30-42.
10. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter, *Anal. Biochem.* 179, 131-137.
11. Guinn, E. J., Pegram, L. M., Capp, M. W., Pollock, M. N., and Record, M. T., Jr. (2011) Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer, *Proc. Natl. Acad. Sci. U. S. A.* 108, 16932-16937.
12. Tsodikov, O. V., Record, M. T., Jr., and Sergeev, Y. V. (2002) Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature, *J. Comput. Chem.* 23, 600-609.
13. Pace, C. N., Grimsley, G. R., and Scholtz, J. M. (2009) Protein ionizable groups: pK values and their contribution to protein stability and solubility, *J. Biol. Chem.* 284, 13285-13289.
14. Peral, F., and Gallego, E. (1997) Self-association of imidazole and its methyl derivatives in aqueous solution. A study by ultraviolet spectroscopy, *J. Mol. Structure* 415, 187-196.

15. Capp, M. W., Pegram, L. M., Saecker, R. M., Kratz, M., Riccardi, D., Wendorff, T., Cannon, J. G., and Record, M. T., Jr. (2009) Interactions of the osmolyte glycine betaine with molecular surfaces in water: thermodynamics, structural interpretation, and prediction of m-values, *Biochemistry* 48, 10372-10379.