7384 *Biochemistry* **2010,** *49,* 7384–7392
DOI: 10.1021/bi1005943

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Radical Redesign of a Tandem Array of Four R67 Dihydrofolate Reductase Genes Yields a Functional, Folded Protein Possessing 45 Substitutions[†]

Jian Feng, Jordan Grubbs, Ashita Dave, Sumit Goswami, Caroline Glyn Horner, and Elizabeth E. Howell*

Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee 37996-0840

Received April 19, 2010; Revised Manuscript Received July 26, 2010

ABSTRACT: R67 dihydrofolate reductase (DHFR) is a plasmid-encoded, type II enzyme. Four monomers (78 amino acids long) assemble into a homotetramer possessing 222 symmetry. In previous studies, a tandem array of four R67 DHFR gene copies was fused in frame to generate a functional monomer named Quad1. This protein possessed the essential tertiary structure of the R67 "parent". To facilitate mutagenesis reactions, restriction enzyme sites were introduced in the tandem gene array. S59A and H362L mutations were also added to minimize possible folding topologies; this protein product, named Quad3, possesses 10 substitutions and is functional. Since R67 DHFR possesses a stable scaffold, a large jump in sequence space was performed by the further addition of 45 amino acid substitutions. The mutational design utilized alternate sequences from other type II DHFRs. In addition, most of the mutations were positioned on the surface of the protein as well as in the disordered N-terminal sequence, which serves as the linker between the fused domains. The resulting Quad4 protein is quite functional; however, it is less stable than Quad1, suffering a $\Delta\Delta G$ loss of 5 kcal/mol at pH 5. One unexpected result was formation of Quad4 dimers and higher order oligomers at pH 8. R67 DHFR, and its derivative Quad proteins, possesses a robust scaffold, capable of withstanding introduction of ≥ 55 substitutions.

Dihydrofolate reductase (EC 1.5.1.3) catalyzes the reduction of dihydrofolate (DHF)[1] to tetrahydrofolate (THF) using the cofactor, NADPH. Two different DHFRs have been identified; the first is chromosomal DHFR, which (in Gram-negative and -positive bacteria) is inhibited by the antibacterial drug, trimethoprim (TMP). The second is a type II DHFR, encoded by R-plasmids, which provides resistance to high concentrations of TMP (*1*). R67 DHFR, a well-characterized example of a type II DHFR, is unrelated genetically and structurally to chromosomal DHFRs (*1*−*3*).

The crystal structure of R67 DHFR describes a homotetramer with a single active site pore that is solvent-accessible, except perhaps at the hourglass center when ligands are bound (*3*−*5*). The overall structure possesses 222 symmetry. This symmetry results in overlapping binding sites for DHF and NADPH, which can be seen experimentally as R67 DHFR binds a total of two ligands. The combinations are either two NADPH molecules or two folate/DHF molecules or one NADPH plus one folate/DHF molecule (*6*). The first two complexes are dead-end complexes, while the third yields the reaction products.

Due to the 222 symmetry and the overlapping binding sites, one mutation per gene produces four mutations per active site pore, which often result in large cumulative effects (*7*, *8*). Also, since each monomer is only 78 amino acids long, each residue likely participates in a number of roles, including contributions to folding, stability, and oligomerization state as well as binding and catalysis.

As a first step in breaking the 222 symmetry in this protein, we made a tandem array of four identical gene copies by removing the stop codons. In this construct, the C-terminus of monomer 1 is fused in-frame with the N-terminus of monomer 2 and so on (*9*). The linker between gene copies corresponds to the natural N-terminus. Recombination of the four identical gene copies down to fewer gene copies occurred in all *Escherichia coli* strains utilized, except for STBL2 (*F- endA1 glnV44 thi-1 recA1 gyrA96 relA1* Δ(*lac-proAB*) *mcrA* Δ(*mcrBC-hsdRMS-mrr*) λ^- (*10*)). Provided the cell stock is frozen in glycerol and the strain is grown at 30 °C and not subjected to prolonged growth under TMP pressure, the four gene copy construct is stable in this strain. The protein resulting from this construct is named Quad1, and SDS−PAGE shows Quad1 possesses four times the mass of the R67 monomer. Quad1 is almost fully functional as k_{cat}/K_m is decreased ~2-fold compared to homotetrameric R67 DHFR. These observations indicate the feasibility of linking the four monomers from R67 DHFR into a single, functional protein, which now possesses four domains.

Mutagenesis of the Quad1 construct did not allow control of either the number of mutations or their placement since each of the gene copies was identical. Thus restriction enzyme sites were designed between the gene copies, allowing each member of the tandem array to be maintained separately and mutagenized as desired (*11*). Recloning then allowed introduction of asymmetric mutations into the full-length construct. From examination of the homotetramer structure, it appeared feasible for the protein to fold into either an ABCD or an ABDC topology or a mixture of the two species in the quadruplicated gene product. To minimize the possible topologies, a S59A mutation was added to gene copy 1 and a H62L mutation was placed in gene copy 4. (Individually, the S59A and H62L mutations destabilize

[†]This work was supported by NSF Grant MCB-0817827.
*Corresponding author: phone, 865-974-4507; fax, 865-974-6306; e-mail, lzh@utk.edu.
[1]Abbreviations: R67 DHFR, R67 dihydrofolate reductase; DHF, dihydrofolate; NADP(+/H), nicotinamide adenine dinucleotide phosphate (oxidized/reduced); wt, wild type; TMP, trimethoprim; MTA buffer, 100 mM Tris, 50 mM Mes, 50 mM acetic acid polybuffer.

 Published on Web 07/27/2010

homotetrameric R67 DHFR, forming inactive dimers. However, 1:1 mixtures restore almost full activity, indicating they complement each other and form heterotetramers (*12*).) The resulting tandem array was named Quad3. Again, the resulting protein was almost fully functional.

Asymmetric mutations (Q67H, K32M, or Y69F) have been added to Quad3, and the resulting proteins indicate that addition of one mutation is well tolerated as single mutants have near-wild-type k_{cat} and K_m values (*11*, *13*, *14*). Conversely, to produce a substantial effect and begin to mimic the mutation in the R67 homotetrameric context, three mutations are usually required. Mutations at the center of the pore (Q67H) produce different effects than mutations further out on the pore surface (K32M and Y69F). However, the effect of mixing mutations has been difficult to predict as parts of the mechanism remain unresolved. Thus a genetic approach could provide an alternate route to generation of mutants with increased catalytic efficiency.

As gene duplication is proposed as a way for gene products to diverge and evolve (*15*−*18*), a class of interesting variants of Quad3 could allow evolution of the binding site so that half the active site pore preferentially binds DHF and the other half of the pore, NADPH. However, recombination between gene copies limits use of a genetic approach as prolonged growth under TMP selection typically results in shorter gene copies. To break the symmetry of the protein and minimize recombination, we have designed a four gene copy sequence that utilizes the redundancy of the genetic code as well as the natural variation in type II DHFR sequences. The resulting construct, Quad4, remains stable in numerous strains of *E. coli*.

MATERIALS AND METHODS

Gene Synthesis and Mutation. Gene copies 2−4 of Quad4 were synthesized by Bio S&T Inc., Montreal, Canada (http://www.biost.com/) and cloned into the *Eco*RV−*Eco*RI sites of Quad3 (*11*). The promoter for this construct is the constitutive, up promoter for *E. coli* chromosomal DHFR (*19*−*21*). DNA sequencing confirmed the desired construct. K32M mutations were introduced into gene copies 1 and 3 as previously described (*14*). The resulting construct is named K32M:1 + 3 in Quad4.

Protein Expression and Purification. *E. coli* cells carrying the Quad4 gene were lysed by sonication in the presence of a protease inhibitor cocktail (Sigma-Aldrich). Separation of Quad4 from the cell lysate used the protocol provided by Qiagen for purification with a Ni-NTA affinity column. The resulting protein had several contaminants, which were removed by a DEAE-fractogel column equilibrated in 50 mM potassium phosphate buffer, pH 8, plus 1 mM EDTA (*19*). All solutions contained 0.1 g/L PEG 3350 to minimize aggregation as previously described (*13*).

Steady-State Kinetics. Steady-state kinetic data were obtained at 30 °C in MTA polybuffer at pH 7.0 using a Perkin-Elmer λ35 spectrophotometer as described elsewhere (*8*). MTA buffer contains 50 mM MES, 100 mM Tris, and 50 mM acetic acid and maintains a constant ionic strength of 0.1 from pH 4.5 to pH 9.5 (*22*). Protein concentration ranges were 13−65 nM for Quad4 and 16−66 nM for the K32M:1 + 3 mutant in Quad4. A nonlinear, global fit of all the data using SAS (*11*, *23*) resulted in best fit values for k_{cat} and both K_m values. The SAS macro (NLINEK) is available at http://animalscience.ag.utk.edu/faculty/saxton/software.htm. DHF was prepared by reduction of folate as described by Blakley (*24*). NADPH was obtained from Alexis Biochemicals. Concentrations of DHF and NADPH were measured using their respective extinction coefficients at 340 nm, 7.75×10^3 M^{-1} cm^{-1} and 6.23×10^3 M^{-1} cm^{-1} (*25*) at 340 nm. The extinction coefficient for the DHFR reaction is 12.3×10^3 M^{-1} cm^{-1} (*26*). For mutants with high K_m values, measurements at 360 nm were performed; values for these extinction coefficients are 4.02×10^3 M^{-1} cm^{-1} for NADPH, 2.63×10^3 M^{-1} cm^{-1} for DHF, and 5.02×10^3 M^{-1} cm^{-1} for the reaction (*14*).

pH Dependence of Enzyme Conformation. A pH-dependent equilibrium in R67 DHFR occurs which describes dissociation of active tetramer into protonated, inactive dimers due to protonation of symmetry-related H62 residues located at the dimer−dimer interfaces (*27*). This equilibrium can be monitored by the environment of W38 residues also located at the dimer−dimer interfaces. Since the Quad1−Quad3 proteins are monomers, they do not undergo dissociation; however, a conformational shift from compact monomer to an extended conformation occurs as the pH is decreased. To monitor these titrations, fluorescence measurements were made using a Perkin-Elmer LS50 spectrometer. Tryptophan residues were excited at 295 nm, and emission was monitored from 300 to 450 nm using 2 μM monomer in MTA buffer. Each sample was titrated with small aliquots of 1 M HCl, and the pH was measured from pH 8 to pH 4. The intensity-averaged emission wavelength, $\langle\lambda\rangle$, for each emission spectrum was calculated using the equation:

$$\langle\lambda\rangle = \sum(I_i\lambda_i)/\sum(I_i) \quad (1)$$

where I is intensity and λ is the wavelength (*28*). The fluorescence data were fit to a standard titration equation. The number of protons, n, was allowed to vary; best fits occurred with $n = 3$.

Single Turnover Experiments. Stopped-flow experiments were performed on an SX20 model Applied PhotoPhysics system, monitoring absorbance at 340 nm. The resulting data were fit using Pro-Data SX software. A total of 5−10 traces from single turnover experiments with final concentrations of 10 μM protein, 8 μM DHF, and 50 μM NADPH (saturating) were averaged and fit to several possible functions, with best fits observed for a single exponential process.

Sedimentation Velocity. Sedimentation velocity experiments were conducted using a Beckman Optima XL-I ultracentrifuge and absorbance optics. Protein samples were dialyzed into 50 mM phosphate buffer, pH 5 or 8, plus 1 mM EDTA, and the dialysis solution was used as the optical reference. Protein was loaded (390 μL loading volume of 2.9−23 μM monomer) into double-sector charcoal-filled Epon centerpieces, and sedimentation velocity analysis was carried out at 50000 rpm and 25 °C using an An50 Ti eight-hole rotor. Sedimentation velocity analysis was performed by direct boundary modeling using the Lamm equation and the Sedfit program (ref *29*; see www.analyticalultracentrifugation.com). Partial specific volume (=0.7269), buffer density, and viscosity values were determined using the SEDNTERP program (ref *30*; see www.jphilo.mailway.com/download.htm). Global fitting of several sedimentation velocity data sets was performed using the Sedphat program (ref *31*; see http://www.analyticalultracentrifugation.com/sedphat/sedphat.htm).

Equilibrium Unfolding. Equilibrium unfolding curves at pH 5.0 were monitored as a function of guanidine hydrochloride (GdnHCl) concentration (*9*, *32*). The fluorescence emission spectra for the DHFRs (excitation, 295 nm) were measured from 310 to 450 nm at each denaturant concentration (between 0 and 5 M GdnHCl). The intensity-averaged emission wavelength for

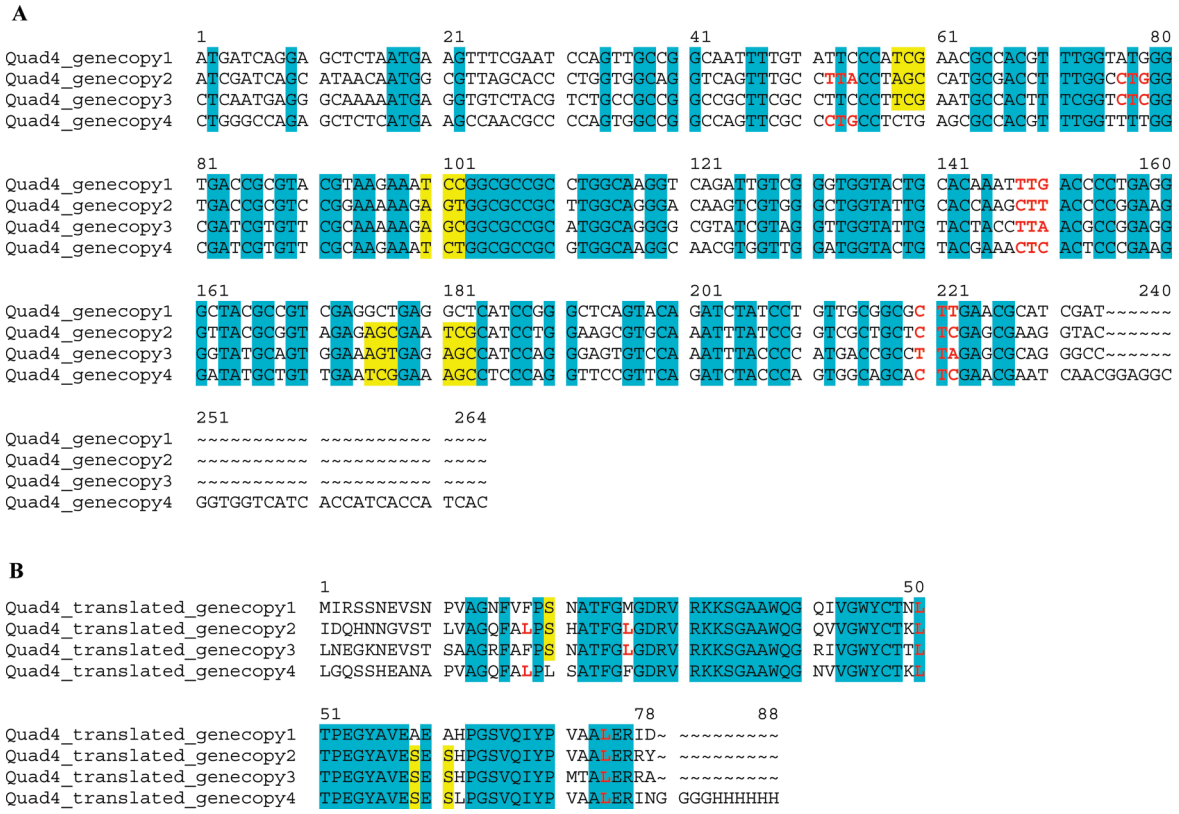

FIGURE 1: DNA and protein sequences used for the Quad4 construct. Panel A: The DNA sequences for the four gene copies are aligned. Several examples where the redundancy of the genetic code was used to promote divergence of the fused gene copies are shown in bold, red letters for leucine codons and are highlighted in yellow for serine codons. Other changes were used to maximize divergence of the gene copies and are listed in Supporting Information Table S1. Panel B: The protein sequences deriving from the four gene copies are shown. Identical sequences in all four copies are highlighted in cyan. The leucines and serines described in panel A are typed in red and highlighted in yellow, respectively.

each emission spectrum was calculated and normalized as described above. The Gibbs free energy change, ΔG, between folded and unfolded states can then be calculated by nonlinear regression of the equation:

$$Y_{obs} = \{(Y^{o}_{N} + M_{N}[\text{GdnHCl}]) + (Y^{o}_{U} + M_{U}[\text{GdnHCl}])\exp(-1(\Delta G_{H_2O} + M_{G}[\text{GdnHCl}])/RT)\} / \{1 + \exp(-1(\Delta G_{H_2O} + M_{G}[\text{GdnHCl}])/RT)\} \quad (2)$$

where R is the gas constant, T is the temperature, ΔG_{H_2O} is the free energy change between native and unfolded protein in the absence of GdnHCl, M_G is the slope describing the dependence of ΔG on denaturant concentration, Y^{o}_{N} and Y^{o}_{U} are the concentration-independent optical values, and M_N and M_U are the slopes of the pre- and posttransitional slopes, respectively (*6*). To allow ready comparison, data were normalized using the equation:

$$F_{app} = (Y_{obs} - Y_N)/(Y_U - Y_N) \quad (3)$$

where F_{app} is a fractional value between 0 and 1 and Y_{obs}, Y_N, and Y_U are the optical values associated at the observed [GdnHCl] and with the native and unfolded forms, respectively (*19*, *33*).

RESULTS

Sequence Design. Numerous steps and considerations were used in the design of alternate sequences for gene copies 2−4 of Quad4. First, even though a low level of divergence (< 3%) has been shown to significantly decrease recombination in *E. coli* (*34*), we aimed for a higher level of divergence to further decrease recombination rates (*35*, *36*). Second, the redundancy of the codon table was used to engineer alternate DNA sequences of conserved amino acid residues. Third, a BLAST search using the R67 DHFR protein sequence finds numerous homologues, where the main difference resides in the N-terminal 15−20 residues. This result agrees with the observation that the N-terminus of dimeric R67 DHFR is disordered since these residues do not appear in its X-ray structure (*37*). Additionally, 16 N-terminal residues can be removed by chymotrypsin treatment, yet the truncated protein remains fully functional (*19*). However, gene sequences with large truncations do not confer resistance to TMP (*19*, *38*, *39*). These observations indicate that while the β-barrel core of R67 DHFR carries the information for catalytic function and variations in the N-terminus are readily tolerated, removal of the N-terminus does not result in expression of functional protein *in vivo*.

Another observation concerning the use of different sequences in gene copies 2−4 arises from analyzing the various type II DHFR protein sequences in algorithms that predict disordered sequences. Intrinsically unstructured proteins or regions of proteins have been recently described and proposed to be important as bristles and linkers as well as other functions (*40*, *41*). When the type II DHFR sequences are submitted to various disorder predictors (DisEMBL (*42*), IUPred (*43*), GlobPlot (*42*), and Pondr (*44*)), all sites indicate a high propensity for the N-terminal sequences of R67 DHFR to be intrinsically disordered. Observation of this pattern suggests that in a protein derived from a tandem gene array the N-terminal sequences may act as spacers between the domains rather than folding onto the surface of the protein. Thus the linker design probably does not need to consider unique contacts between the linker and the surface of the protein.

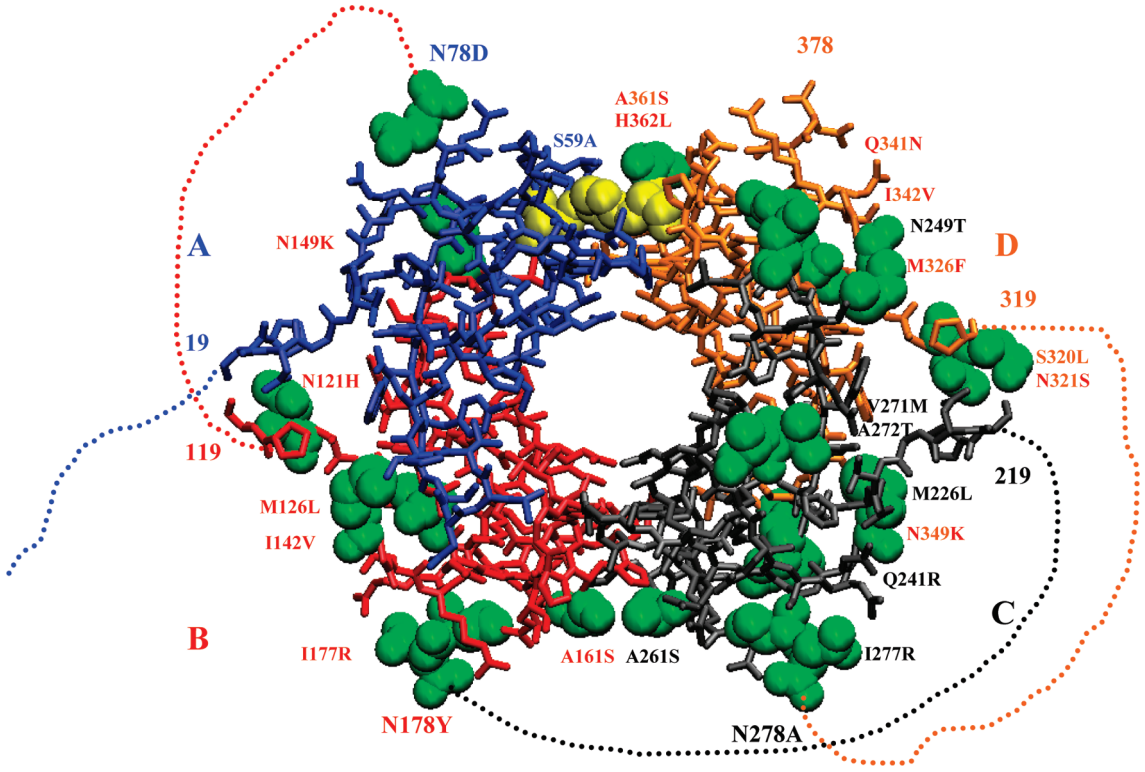

FIGURE 2: The structure of homotetrameric R67 DHFR (1VIE from the Protein Data Bank (*3*)). Each monomer is colored differently. Residues are not numbered consecutively; rather the amino acids in the first monomer (A) are labeled 1−78; those in the second monomer (B), 101−178; those in the third monomer (C), 201−278; and those in the fourth monomer (D), 301−378. The active site pore occurs in the middle of the structure. Since residues 1−17 were disordered in the dimer structure (*37*), residues 1−16 were truncated by chymotrypsin treatment (*19*), and this species crystallized as a tetramer (*3*). The dimer−dimer interfaces occur at the top and bottom of the figure. The monomer arrangement going counterclockwise in the crystal structure 1VIE is ABDC. To minimize confusion in the quadruplicated gene construct, we have relabeled the monomers ABCD going counterclockwise. In Quad4, the N-terminal 18 residues of monomer A are modeled as a blue dotted line. The C-terminus of monomer A (N78) would be connected to the N-terminus of monomer B, the C-terminus of monomer B would be connected to the N-terminus of monomer C, and the C-terminus of monomer C would be connected to the N-terminus of monomer D. The expected connections are drawn with dotted lines. In Quad4, the monomers would become domains. The residues in Quad4 that differ from the R67 DHFR sequence are shown in green with CPK surfaces. The identity and approximate location of the substitutions are shown. Additional changes occur in the linker sequences (residues 1−19), which contain the most variability. The S59A and H62L mutations, designed to prevent formation of alternate topologies, are shown in yellow.

A minor change in Quad4 design was addition of a His tag sequence to the C-terminus to facilitate protein purification. Maintenance of unique restriction sites for cloning or screening of PCR mutagenesis reactions was also desired.

Using these various considerations, sequences for gene copies 2, 3, and 4 were designed. Gene copy 1 remained our synthetic gene for R67 DHFR with numerous unique restriction sites (*19*, *45*). The sequence for gene copy 2 came from *dfrB3*, GenBank accession number AY123252, also known as R751 DHFR (*46*, *47*). The gene copy 3 sequence originates from *dfrB4*, GenBank accession number AJ429132 (*48*, *49*), while gene copy 4 comes from *dfrB2*, GenBank accession number J01773, also known as R388 DHFR (*50*). Figure 1A shows an alignment of the four designed gene sequences. The highest degree of identity between gene copies is 71%, while the lowest degree is 62%. Figure 1B shows the alignment of the corresponding protein sequences. Because we used the redundancy of the codon table, the protein sequences are more closely related, ranging from 68% to 78% identity. The four gene copy fusion and its resulting protein are both named Quad4.

The amino acids in Quad4 that have been changed are mapped onto the structure of R67 DHFR as shown in Figure 2. Since the crystal structure is of the β-barrel core where the N-terminus is cleaved off by chymotrypsin, all the changes that occur in the N-terminus (residues 1−18) and the linkers (residues 101−118, 201−218, 301−318) do not appear. The changes shown in Figure 2 are on the surface of the protein, not in the active site pore, suggesting that the protein should be functional. Overall, 10 amino acid substitutions were added to Quad1 to engineer Quad3. An additional 45 substitutions were then added to Quad3 to obtain Quad4. This results in a total of 55 substitutions incorporated into a 312 amino acid protein (not counting the C-terminal GGGG linker and His tag sequence), resulting in an ~18% change in sequence.

Analysis. The Quad4 gene was synthesized and cloned as described above. It was initially transformed into the Stbl2 strain of *E. coli*, previously found to maintain the tandem array with minimal recombination (*10*). Quad3 and Quad4 have subsequently been transformed into several other *E. coli* strains, including DH5α (*F' endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15* Δ(*lacZYA-argF*)*U169*, *hsdR17*(r_K^- m_K^+), λ^-) as well as LH18 (Δ(*lac pro*) *supE hsdR5 F' lac* (I^q-*ZΔM15*)$^+$ *pro*$^+$ *thyA* Δ*fol::kan* (*51*)) and MH829 (*thyA* (*ts*) *argE3 rna* Δ*folaA::kan3 λ* (*52*)). *Bam*HI digestion of the plasmid carrying both Quad3 and Quad4 DNA from these strains shows a band corresponding to the quadruplicated gene construct. Previous analysis of Quad1 in JM107 and Sure *E. coli* (Stratagene) strains showed fragment ladders, indicating recombination of the

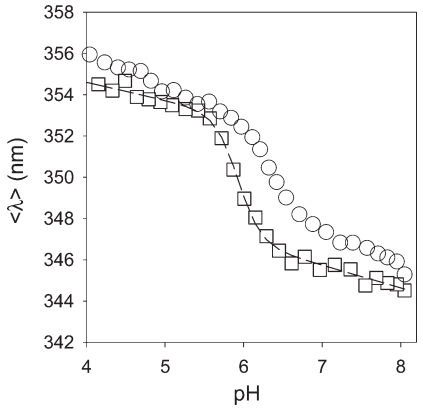

FIGURE 3: The pH dependence of Quad3 and Quad4 fluorescence. The *y*-axis describes the intensity-averaged emission wavelength, $\langle\lambda\rangle$, calculated by eq 1. The titration curve monitoring the equilibrium between the "closed" and "open" forms of Quad3 is shown by a dashed line and □ points. A best fit value for the pK_a for Quad3 is 6.0. The response of Quad4 to pH is shown by ○ points.

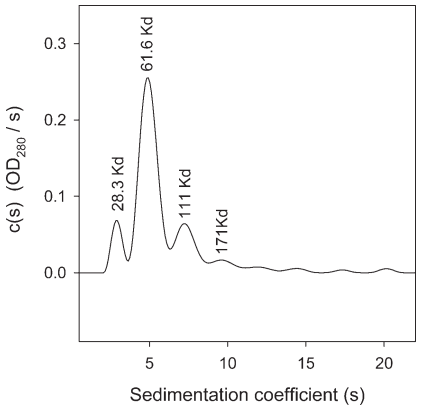

FIGURE 4: Sedimentation coefficient distributions, or *c*(*s*), calculated from analysis of sedimentation velocity data. The sedimentation pattern of 9.3 μM Quad4 at pH 8 was monitored by absorbance at 280 nm. The resulting data were analyzed using the *c*(*s*) model in Sedfit. The contribution of each component to the overall absorbance is shown. The goodness of the fit is shown by a rmsd of 0.004377.

four gene copies down to fewer copies (*10*). Thus the later Quad3 and Quad4 designs appear stable in several different *E. coli* strains. Even after daily subculturing for a week in the presence of trimethoprim, recombination was not observed.

pH Dependence of Enzyme Conformation. The Quad4 protein was purified and characterized by a variety of techniques. To monitor the pH dependence of the Quad4 conformation, changes in fluorescence were monitored. This approach parallels that used in R67 DHFR, as at pH 8, W38 residues (at the dimer−dimer interfaces) are buried in the native homotetramer, whereas at pH 5, W38s are exposed to solvent in the inactive dimer. Since W38 and its symmetry-related partners remain in the quadruplicated gene products, we monitor their fluorescence as a function of pH. For Quad1 and Quad3, the protein changes conformation as the domains splay apart upon H62 protonation. Figure 3 shows the titration of Quad3, with a K_a of $1.08 \times 10^{-6} \pm 3.7 \times 10^{-8}$, which corresponds to a pK_a of 6.0. Figure 3 also shows the titration of Quad4 at an equivalent protein concentration. Since this pH titration describes a more complex process (see Sedimentation Velocity studies below), we simply show the data and note that the titration for Quad4 in Figure 3 is shifted to the right compared to Quad3, consistent with a lower stability in acidic conditions. For comparison, Quad1 is even more pH stable, displaying a pK_a of 5.5 (*9*).

Sedimentation Velocity. At pH 5, ultracentrifugation experiments find Quad4 to be predominately monomeric (85−90% of the population for three different runs at protein concentrations of 9−12 μM). The concentration of discrete species (or *c*(*s*)) plot for one of the three data sets is shown in Supporting Information Figure S1. Global fitting of the three data sets using Sedphat (*31*) yields a monomer with an *s* value of 3.17, which corresponds to a mass of 33800 Da. This mass compares well to a calculated value of 34560 Da.

At pH 8, sedimentation velocity experiments found Quad4 to be mostly dimeric, with some monomer present as well as some higher order oligomers. Figure 4 shows a *c*(*s*) plot describing the various species observed. In contrast to these results, previous ultracentrifugation studies have found Quad1 and Quad3 to be monomeric at both pH 5 and pH 8 (*9*, *14*). To investigate the concentration dependence of the Quad4 dimers, sedimentation velocity experiments were performed with three different protein concentrations (2.9, 9.3, and 23 μM monomer). The data were globally analyzed with Sedphat, and the resulting fit is given in Supporting Information Figure S2. As the protein concentration increases, so does the average oligomeric size. We were able to fit the data using a hybrid local continuous and global discrete species model; the *c*(*s*) values for monomer, dimer, and tetramer are 3.068, 5.153, and 7.650, respectively, corresponding to masses of 28300, 61600, and 111000 Da.

Since the peak positions do not change as a function of protein concentration, we were unable to fit the data using a monomer to *n*-mer self-association model, most likely as interconversion between species is slow on the sedimentation time scale. However, the integrated peak areas can be used to estimate K_d values for the various equilibria. For the 9.3 and 23 μM monomer data sets, the monomer to dimer K_d value was calculated as 0.6 or 1.5 μM. While the estimated K_d values are ~2-fold different, the data collectively indicate dimer predominates between 2 and 10 μM.

How Active Is Quad4? Steady-state kinetic analysis yielded a k_{cat} value of 1.8 ± 0.1 s^{-1} (assuming a monomer mass) at pH 7.0, 30 °C, and K_m values of 5.6 ± 0.3 μM and 2.6 ± 0.2 μM for DHF and NADPH, respectively. (Since our steady-state assays utilized 13−66 nM enzyme, it is likely that k_{cat} reports a value for the monomeric species.) The kinetic values for Quad3, Quad1, and the wt homotetramer are also listed in Table 1. Comparing all of these values shows 2-fold effects at the most, indicating minimal effects on enzyme activity due to covalent linking of the monomers and the redesign of the protein sequence to decrease the symmetry within the structure.

Single Turnover Experiments. To address the question of whether dimeric Quad4 is active, stopped-flow analysis was performed with a limiting DHF concentration. Quad4 (10 μM) was used, as dimer predominates at this concentration. If monomeric Quad4 is more active than dimeric Quad4, we might expect to observe a better fit to more than one exponential. If only monomeric Quad4 is active, we might expect a linear phase, consistent with multiple turnovers (*53*). For Quad4 mixed with 8 μM DHF and 50 μM NADPH, the data were clearly fit by a single exponential with a rate of 0.69 ± 0.0004 s^{-1} (see Supporting Information Figure S3). This result is consistent with both monomer and dimer being active and possessing reasonably similar rate constants. R67 DHFR (10 μM) was used as a control; the data were fit to a single exponential with a rate of 0.92 ± 0.0002 s^{-1}.

Table 1: Steady-State Kinetic Values at pH 7.0, 30 °C

protein species	k_{cat} (s^{-1})	$K_{m(NADPH)}$ (μM)	$K_{m(DHF)}$ (μM)
Quad4	1.8 ± 0.1[a]	2.6 ± 0.2	5.6 ± 0.3
Quad3[b]	0.81 ± 0.02	4.4 ± 0.4	6.7 ± 0.4
Quad1[c]	0.75 ± 0.01	4.5 ± 0.2	8.0 ± 0.3
R67 DHFR[d]	1.3 ± 0.1	3.0 ± 0.1	5.8 ± 0.1
K32M:1 + 3 in Quad4	1.4 ± 0.1[a]	81 ± 7	165 ± 14
K32M:1 + 3 in Quad3[e]	≥3.7	≥145	≥400

[a]Calculated using the monomer mass. [b]From ref *11*. The construct describes four identical R67 DHFR gene copies fused in frame with S59A and H362L mutations in gene copies 1 and 4, respectively, and unique restriction enzyme sites designed between gene copies. [c]From ref *9*. The construct describes four identical R67 DHFR gene copies fused in frame. [d]From ref *19*. [e]From ref *14*.

A Dimer Model. From our pH titration data, the oligomeric state of high concentrations of Quad4 varies as a function of pH; thus dimer formation may involve intermolecular interactions at the loop regions containing H62, H162, and H262 (*27*). If so, the presence of the S59A and H362L mutations in gene copies 1 and 4 should constrain the possible modes of assembly. In other words, the S59A and H362L mutations in domains A and D must interact with each other to maintain DHFR activity and confer TMP resistance. A simple model, as shown in Figure 5, suggests two monomers might be able to associate in an antiparallel fashion. This model allows the S59A mutation in domain A of the first Quad4 molecule to interact with the H362L mutation in domain D of the second Quad4 molecule (and vice versa). In this scenario, Quad4 would have an elongated conformation (like beads on a string) as a monomer at pH 5, and then two of these elongated species could interact to form a dimer at pH 8. Nonspecific association may also occur in Quad4, but as long as the building block(s) involved in oligomer formation retain(s) their predicted, unique active site configuration, the kinetic behavior should be predictable (see below).

Mutational Analysis? Our model for the Quad4 dimer shown in Figure 5 predicts that mutations can be added asymmetrically and that the mutations will remain in the same relative positions as in the compact Quad3 monomer. In other words, there should not be a mechanism by which mutations can be oriented such that some configuration yields a wild-type site. To test this hypothesis, K32M mutations were added asymmetrically to gene copies 1 and 3; this construct is named K32M:1 + 3 in Quad4. This construction places both mutations on one side of the protein pore (see Figure 1B in Hicks et al. (*14*)), which strongly affects ligand binding.

The K32M:1 + 3 double mutant in the Quad4 context was constructed and purified. Analysis of the steady-state kinetic data yields a k_{cat} of 1.4 ± 0.1 s^{-1}, a $K_{m(NADPH)}$ of 81 ± 7 μM, and a $K_{m(DHF)}$ of 165 ± 14 μM at pH 7 (see Supporting Information Figure S4). For comparison, we previously reported a k_{cat} of ≥3.7 s^{-1}, a $K_{m(NADPH)}$ of ≥145 μM, and a $K_{m(DHF)}$ of ≥400 μM for this mutant in Quad3 (also see Table 1). Since ligand binding is greatly weakened in these mutants, it is difficult to saturate the enzyme and obtain accurate kinetic values. As all of the values for the K32M:1 + 3 in Quad4 mutant are approximately 2-fold lower than those for the K32M:1 + 3 mutant in Quad3, it may be that these differences are really due to the different protein contexts. Alternatively, since addition of the His tag greatly aided purification of the Quad4 mutant, it could be that the difficulty and time

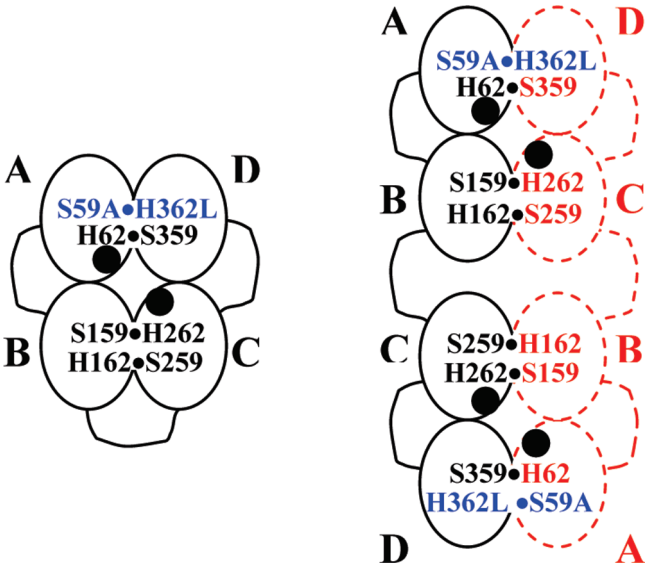

FIGURE 5: Cartoon representations of Quad4. At left is a compact monomer depicting the structure shown in Figure 2. Each domain is depicted as a circle, and the active site pore is shown by a central diamond shape. The covalent linkages between domains are shown by the connecting lines. Since the S59A and H362L mutations must interact to allow a functional protein, they constrain the possible topologies to either an ABCD or ACBD arrangement. We prefer the simpler ABCD arrangement as the protein product of a duplicated gene is monomeric at pH 5 (*32*). If oligomers form due to specific interactions, the right drawing shows a simple model for dimeric Quad4. Monomer 1 is shown in black with solid lines while monomer 2 is shown in red with dashed lines. The S59A and H362L mutations (blue type) again constrain the possible packing arrangements. A probe of this model may be addition of asymmetric mutations, for example, the K32M:1 + 3 construct with mutations in gene copies 1 and 3, which places mutations in domains A and C (shown by ●). In the compact monomer model, the mutations occur diagonally across one-half of the active site pore. This topological relationship is conserved in the dimer model. We also note another possible dimer/oligomer model could use interlocking ring structures (not shown).

taken in purifying the K32M:1 + 3 mutant in Quad3 could have negatively impacted the efficiency of that variant. It also could be that the wider [DHF] and [NADPH] ranges used in this study resulted in better bracketing of the K_m values and thus more accurate values. In any case, it is clear that the k_{cat} and K_m values of this mutant are quite different from those of Quad4, indicating addition of asymmetric mutations alters function in a (mostly) predictable manner. Finally, the *R* squared value for the fit in Supporting Information Figure S4 is 97%, indicating a good fit of the data to the bisubstrate Michaelis−Menten equation. This result is consistent with the active sites in the various oligomeric states being equivalent and does not support alternate packing events such that some assemblies possess wild-type active sites. The latter should be readily apparent due to the different k_{cat}/K_m values for the wild-type vs mutant sites (*54*).

A *c*(*s*) plot from a sedimentation velocity experiment with this mutant is shown in Supporting Information Figure S5. The plot indicates monomer is the predominant species (64%) with some dimer (27%) and fewer higher order oligomers. The *s* coefficient for monomer is 3.14, corresponding to a mass of 30500 Da. Addition of the K32M mutations appears to shift the oligomer equilibrium toward monomer. Thus, while analysis of this mutant supports using Quad4 as a scaffold for investigating mutational effects, the fact that it is mostly monomeric limits its usefulness for analysis of the dimer model.

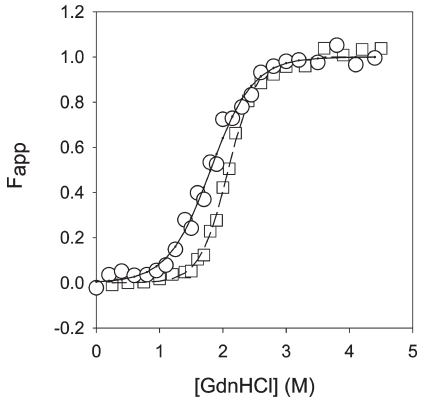

FIGURE 6: Equilibrium unfolding/refolding curves for Quad3 and Quad4 at pH 5 monitored by fluorescence. The stabilities of extended conformations of Quad3 (dashed line and □ points) and Quad4 (solid line and ○ points) were monitored as a function of guanidine hydrochloride concentration. The intensity-averaged wavelength, $\langle\lambda\rangle$, was calculated, and the data were fit to eq 2. To aid comparison, the data were converted to a fraction using eq 3. Best fit values are given in Table 2.

Equilibrium Unfolding Curves. Next, the stability of Quad4 was assessed by equilibrium unfolding/refolding in guanidine hydrochloride. Since previous studies with Quad1 showed full reversibility at pH 5, but only partial reversibility at pH 8, equilibrium unfolding was studied at pH 5 where the Quad variants are monomeric and possess an extended configuration. Figure 6 compares the unfolding curves for Quad3 and Quad4. The ΔG for Quad4 unfolding is -3.2 ± 0.2 kcal/mol with a M_G (slope) of -1.8 ± 0.2 kcal mol^{-1} M^{-1}. In comparison, the ΔG for Quad3 is -5.6 ± 0.4 kcal/mol with a M_G of -2.7 ± 0.2. These values are compiled in Table 2. Quad1 is 2.6 kcal/mol more stable than Quad3, which, in turn, is 2.4 kcal/mol more stable than Quad4. The cumulative addition of substitutions to the Quad proteins decreases their stability. Since the 10 and 45 amino acid replacements were added *en bloc*, it is not clear whether specific substitutions and/or loss of symmetry is involved in these effects.

DISCUSSION

Effects of Substitutions on Function and Stability. Previously, 10 substitutions were added to Quad1 to generate Quad3 (*11*). In this study, addition of 45 more substitutions generated Quad4. No loss of activity (Table 1) was observed, probably as amino acid substitutions were not added to the active site. An unexpected result was formation of dimers and higher order oligomers at pH 7−8. A first reason considered for this change was that dimerization might have occurred due to introduction of the A161S and A261S mutations at one of the domain−domain interfaces (corresponding to a dimer−dimer interface in R67 DHFR; see Figure 2). However, when these substitutions were removed from Quad4, the protein still remained dimeric (data not shown). Since our previous modeling had suggested the N-terminal linker sequence was barely long enough to link the 2 domains together (*9*), perhaps changes in the new sequence between gene copies 2 and 3 are responsible for the change in oligomeric state.

Design Aspects. Several R67 DHFR features likely contribute to our ability to engineer Quad4. A first aspect is that most of the substitutions occur in the disordered linker sequences. Indeed, randomization of disordered sequences has sometimes been found to still allow function (*55*). Second, all four gene sequences used were based on natural sequences and their ability

Table 2: Best Fit Values of Equilibrium Folding Curves to eq[2] As Monitored by Fluorescence at pH 5.0[a]

species	ΔG (kcal mol^{-1})	M_G (slope) (kcal mol^{-1} M^{-1})
Quad4	-3.2 ± 0.3	-1.8 ± 0.2
Quad3	-5.6 ± 0.4	-2.7 ± 0.2
Quad1[b]	-8.2 ± 0.3	-2.8 ± 0.1

[a]Figure 6 shows the experimental data normalized to the fraction unfolded (F_{app}), a value between 0 and 1. [b]From ref *9*.

to confer TMP resistance, indicating they were individually viable (*46*−*50*). As R67 DHFR is a homotetramer, we can apparently treat the individual monomers as modules or building blocks and swap out one functional monomer (or domain in Quad4) for another. Third, R67 DHFR does not possess a typical hydrophobic core as the structure is doughnut shaped. Thus core repacking effects might be expected to be minimal (*56*). Fourth, R67 DHFR is quite stable with a $\Delta G_{H_2O}°$ for the dimer of −13.9 kcal/mol at pH 5 (*19*) and a $\Delta G_{H_2O}°$ for the tetramer of −34.3 kcal/mol at pH 8.0 (*57*). Fifth, other mutagenesis experiments have found R67 DHFR to be a robust enzyme, capable of accumulating numerous mutations, yet retaining function. For example, an RNA hypermutagenesis experiment using R67 DHFR generated three sequential libraries of TMP-resistant clones (*39*). Sequencing indicated up to 18 mutations per gene (a 22% change), with hypermutagenesis in the N-terminus and only 18 residues (mostly mapping to amino acids 43−51 and 65−70) remaining mutation free.

While 55 substitutions were involved in generation of Quad4, a subset of the replacements were symmetry related; for example, M101I, M201L, and M301L all introduce substitutions at the beginning of the second, third, and fourth gene copies. In all, 28 different positions were mutated. Symmetry usually provides a simple strategy for building oligomers, leading to protein stability (*58*−*60*). Thus the ability to add 1−3 asymmetric mutations per tandem array might have a lesser effect than addition of 4 symmetric mutations per homotetramer. This has been previously observed in Quad3 as addition of one asymmetric mutation typically does not impact function (*11*, *13*, *14*). This may be important, for if residues in R67 DHFR play multiple roles in protein function, folding, and/or oligomerization, then alteration of <4 residues should have a smaller impact than changing all symmetry-related residues. The remaining 1−3 residues could continue to serve their other function(s).

There Are Several Views of Mutational Effects. One view posits that most mutations are neutral and do not affect function (*61*−*63*). Another view proposes that most mutations have deleterious effects; however, these effects can be masked as proteins have a fitness threshold, allowing accumulation of mutations with minimal consequences (*64*−*66*). Past this threshold, the toll of mutations on fitness increases and protein function suffers. The threshold correlates mainly with thermodynamic stability. For example, introduction of increasing numbers of mutations into TEM-1 β-lactamase was correlated with loss of fitness, and addition of 16 mutations was poorly tolerated (*64*).

Past studies by the Tawfik and Serrano groups have found an average destabilizing effect of ~0.6 kcal/mol per surface substitution (*66*). However, for the Quad variants, addition of 10 substitutions going from Quad1 to Quad3 results in a $\Delta\Delta G$ of 2.6 kcal/mol for the monomer at pH 5, corresponding to a $\Delta\Delta G$ of 0.26 kcal/mol per mutation. The further addition of 45 substitutions

results in an additional $\Delta\Delta G$ of 2.4 kcal/mol. This mutational jump corresponds to a $\Delta\Delta G$ of 0.05 kcal/mol per mutation. These values are quite small, indicating the ability of R67 DHFR to carry a large mutational load. Further, these results support proposals that link robustness and evolvablility with thermodynamic stability (*67*−*70*). Here, the "extra" thermodynamic stability of R67 DHFR makes its structure and function more robust to the addition of 45−55 mutations.

ACKNOWLEDGMENT

We thank Cynthia Peterson for critical reading of the manuscript and Larry Thompson and Dan Roberts for help with the stopped flow.

SUPPORTING INFORMATION AVAILABLE

The results of sedimentation coefficient distributions, or $c(s)$ analysis, calculated from analysis of sedimentation velocity data of Quad4 at pH 5, Sephat analysis of sedimentation patterns at pH 8 for three different Quad4 concentrations, the results from a single turnover experiment using 10 μM Quad4 monitored with a limiting DHF concentration, and a steady-state kinetic analysis of the K32M:1 + 3 mutant in Quad4 as well as Sedfit analysis of sedimentation velocity data. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. White, P. A., and Rawlinson, W. D. (2001) Current status of the *aadA* and *dfr* gene cassette families. *J. Antimicrob. Chemother. 47*, 495–496.
2. Howell, E. E. (2005) Searching sequence space: two different approaches to dihydrofolate reductase catalysis. *ChemBioChem 6*, 590–600.
3. Narayana, N., Matthews, D. A., Howell, E. E., and Nguyen-huu, X. (1995) A plasmid-encoded dihydrofolate reductase from trimethoprim-resistant bacteria has a novel D2-symmetric active site. *Nat. Struct. Biol. 2*, 1018–1025.
4. Krahn, J., Jackson, M., DeRose, E. F., Howell, E. E., and London, R. E. (2007) Structure of a type II dihydrofolate reductase ternary complex: use of identical binding sites for unrelated ligands. *Biochemistry 46*, 14878–14888.
5. Narayana, N. (2006) High-resolution structure of a plasmid-encoded dihydrofolate reductase: pentagonal network of water molecules in the D2-symmetric active site. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 62*, 695–706.
6. Bradrick, T. D., Beechem, J. M., and Howell, E. E. (1996) Unusual binding stoichiometries and cooperativity are observed during binary and ternary complex formation in the single active pore of R67 dihydrofolate reductase, a D_2 symmetric protein. *Biochemistry 35*, 11414–11424.
7. Strader, M. B., Chopra, S., Jackson, M., Smiley, R. D., Stinnett, L., Wu, J., and Howell, E. E. (2004) Defining the binding site of homotetrameric R67 dihydrofolate reductase and correlating binding enthalpy with catalysis. *Biochemistry 43*, 7403–7412.
8. Strader, M. B., Smiley, R. D., Stinnett, L. G., VerBerkmoes, N. C., and Howell, E. E. (2001) Role of S65, Q67, I68, and Y69 residues in homotetrameric R67 dihydrofolate reductase. *Biochemistry 40*, 11344–11352.
9. Bradrick, T. D., Shattuck, C., Strader, M. B., Wicker, C., Eisenstein, E., and Howell, E. E. (1996) Redesigning the quaternary structure of R67 dihydrofolate reductase. Creation of an active monomer from a tetrameric protein by quadruplication of the gene. *J. Biol. Chem. 271*, 28031–28037.
10. Strader, M. B., and Howell, E. E. (1997) Stable maintenance of a tandem array of four R67 dihydrofolate reductase genes. *Gibco-BRL Focus 19*, 24–25.
11. Smiley, R. D., Stinnett, L. G., Saxton, A. M., and Howell, E. E. (2002) Breaking symmetry: mutations engineered into R67 dihydrofolate reductase, a D2 symmetric homotetramer possessing a single active site pore. *Biochemistry 41*, 15664–15675.
12. Dam, J., Rose, T., Goldberg, M. E., and Blondel, A. (2000) Complementation between dimeric mutants as a probe of dimer-dimer interactions in tetrameric dihydrofolate reductase encoded by R67 plasmid of *E. coli*. *J. Mol. Biol. 302*, 235–250.
13. Stinnett, L. G., Smiley, R. D., Hicks, S. N., and Howell, E. E. (2004) "Catch 222," the effects of symmetry on ligand binding and catalysis in R67 dihydrofolate reductase as determined by mutations at Tyr-69. *J. Biol. Chem. 279*, 47003–47009.
14. Hicks, S. N., Smiley, R. D., Stinnett, L. G., Minor, K. H., and Howell, E. E. (2004) Role of Lys-32 residues in R67 dihydrofolate reductase probed by asymmetric mutations. *J. Biol. Chem. 279*, 46995–47002.
15. Vogel, C., and Morea, V. (2006) Duplication, divergence and formation of novel protein topologies. *BioEssays 28*, 973–978.
16. Ohno, S. (1988) On periodicities governing the construction of genes and proteins. *Anim. Genet. 19*, 305–316.
17. Ohno, S. (1970) Evolution by Gene Duplication, Springer-Verlag, Berlin.
18. Li, W.-H., and Graur, D. (1991) Molecular Evolution, Sinauer Associates, Sunderland, MA.
19. 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.
20. Smith, D. R., and Calvo, J. M. (1980) Nucleotide sequence of the *E. coli* gene coding for dihydrofolate reductase. *Nucleic Acids Res. 8*, 2255–2274.
21. Smith, D. R., and Calvo, J. M. (1982) Nucleotide sequence of dihydrofolate reductase genes from trimethoprim-resistant mutants of *Escherichia coli*. Evidence that dihydrofolate reductase interacts with another essential gene product. *Mol. Gen. Genet. 187*, 72–78.
22. Ellis, K. J., and Morrison, J. F. (1982) Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol. 87*, 405–426.
23. Smiley, R. D., Saxton, A. M., Jackson, M. J., Hicks, S. N., Stinnett, L. G., and Howell, E. E. (2004) Nonlinear fitting of bisubstrate enzyme kinetic models using SAS computer software: application to R67 dihydrofolate reductase. *Anal. Biochem. 334*, 204–206.
24. Blakley, R. L. (1960) Crystalline dihydropteroylglutamic acid. *Nature 188*, 231–232.
25. Horecker, B. L., and Kornberg, A. (1948) The extinction coefficients of the reduced band of pyridine nucleotides. *J. Biol. Chem. 175*, 385–390.
26. Baccanari, D., Phillips, A., Smith, S., Sinski, D., and Burchall, J. (1975) Purification and properties of *Escherichia coli* dihydrofolate reductase. *Biochemistry 14*, 5267–5273.
27. Nichols, R., Weaver, C. D., Eisenstein, E., Blakley, R. L., Appleman, J., Huang, T. H., Huang, F. Y., and Howell, E. E. (1993) Titration of histidine 62 in R67 dihydrofolate reductase is linked to a tetramer to two-dimers equilibrium. *Biochemistry 32*, 1695–1706.
28. Royer, C. A., Mann, C. J., and Matthews, C. R. (1993) Resolution of the fluorescence equilibrium unfolding profile of trp aporepressor using single tryptophan mutants. *Protein Sci. 2*, 1844–1852.
29. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J. 78*, 1606–1619.
30. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Analytical Ultracentrifugation in Biochemistry and Polymer Science (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge.
31. Schuck, P. (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal. Biochem. 320*, 104–124.
32. Zhuang, P., Yin, M., Holland, J. C., Peterson, C. B., and Howell, E. E. (1993) Artificial duplication of the R67 dihydrofolate reductase gene to create protein asymmetry: Effects on protein activity and folding. *J. Biol. Chem. 268*, 22672–22679.
33. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1990) Protein Structure: A Practical Approach, Chapter 13, IRL Press, Oxford.
34. Worth, L., Jr., Clark, S., Radman, M., and Modrich, P. (1994) Mismatch repair proteins MutS and MutL inhibit RecA-catalyzed strand transfer between diverged DNAs. *Proc. Natl. Acad. Sci. U.S.A. 91*, 3238–3241.
35. Stambuk, S., and Radman, M. (1998) Mechanism and control of interspecies recombination in *Escherichia coli*. I. Mismatch repair, methylation, recombination and replication functions. *Genetics 150*, 533–542.
36. Shen, P., and Huang, H. V. (1986) Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics 112*, 441–457.
37. Matthews, D. A., Smith, S. L., Baccanari, D. P., Burchall, J. J., Oatley, S. J., and Kraut, J. (1986) Crystal structure of a novel

trimethoprim-resistant dihydrofolate reductase specified in *Escherichia coli* by R-plasmid R67. *Biochemistry 25*, 4194–4204.
38. Vermersch, P. S., and Bennett, G. N. (1988) Synthesis and expression of a gene for a mini type II dihydrofolate reductase. *DNA 7*, 243–251.
39. Martinez, M. A., Pezo, V., Marliere, P., and Wain-Hobson, S. (1996) Exploring the functional robustness of an enzyme by in vitro evolution. *EMBO J. 15*, 1203–1210.
40. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein. *J. Mol. Graphics Model. 19*, 26–59.
41. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol. 293*, 321–331.
42. Linding, R., Russell, R. B., Neduva, V., and Gibson, T. J. (2003) GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic Acids Res. 31*, 3701–3708.
43. Dosztányi, Z., Csizmók, V., Tompa., P., and Simon, I. (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics 21*, 3433–3434.
44. Li, X., Rani, M., Romero, P., Obradovic, Z., and Dunker, A. K. (1999) Predicting protein disorder for N-, C-, and internal regions. *Genome Inf. 10*, 30–40.
45. Brisson, N., and Hohn, T. (1984) Nucleotide sequence of the dihydrofolate-reductase gene borne by the plasmid R67 and conferring methotrexate resistance. *Gene 28*, 271–274.
46. Flensburg, J., and Steen, R. (1986) Nucleotide sequence analysis of the trimethoprim resistant dihydrofolate reductase encoded by R plasmid R751. *Nucleic Acids Res. 14*, 5933.
47. Radstrom, P., Skold, O., Swedberg, G., Flensburg, J., Roy, P. H., and Sundstrom, L. (1994) Transposon Tn5090 of plasmid R751, which carries an integron, is related to Tn7, Mu, and the retroelements. *J. Bacteriol. 176*, 3257–3268.
48. Grape, M., Farra, A., Kronvall, G., and Sundstrom, L. (2005) Integrons and gene cassettes in clinical isolates of co-trimoxazole-resistant Gram-negative bacteria. *Clin. Microbiol. Infect. 11*, 185–192.
49. Grape, M., Sundstrom, L., and Kronvall, G. (2003) New dfr2 gene as a single-gene cassette in a class 1 integron from a trimethoprim-resistant *Escherichia coli* isolate. *Microb. Drug Resist. 9*, 317–322.
50. Swift, G., McCarthy, B. J., and Heffron, F. (1981) DNA sequence of a plasmid-encoded dihydrofolate reductase. *Mol. Gen. Genet. 181*, 441–447.
51. Howell, E. E., Foster, P. G., and Foster, L. M. (1988) Construction of a dihydrofolate reductase-deficient mutant of *Escherichia coli* by gene replacement. *J. Bacteriol. 170*, 3040–3045.
52. Herrington, M. B., and Chirwa, N. T. (1999) Growth properties of a folA null mutant of *Escherichia coli* K12. *Can. J. Microbiol. 45*, 191–200.
53. Heredia, V. V., and Penning, T. M. (2004) Dissection of the physiological interconversion of 5alpha-DHT and 3alpha-diol by rat 3alpha-HSD via transient kinetics shows that the chemical step is rate-determining: effect of mutating cofactor and substrate-binding pocket residues on catalysis. *Biochemistry 43*, 12028–12037.
54. Spears, G., Sneyd, J. G., and Loten, E. G. (1971) A method for deriving kinetic constants for two enzymes acting on the same substrate. *Biochem. J. 125*, 1149–1151.
55. Hansen, J. C., Lu, X., Ross, E. D., and Woody, R. W. (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem. 281*, 1853–1856.
56. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science 255*, 178–183.
57. Zhuang, P., Eisenstein, E., and Howell, E. E. (1994) Equilibrium folding studies of tetrameric R67 dihydrofolate reductase. *Biochemistry 33*, 4237–4244.
58. Blundell, T. L., and Srinivasan, N. (1996) Symmetry, stability, and dynamics of multidomain and multicomponent protein systems. *Proc. Natl. Acad. Sci. U.S.A. 93*, 14243–14248.
59. Goodsell, D. S., and Olson, A. J. (2000) Structural symmetry and protein function. *Annu. Rev. Biophys. Biomol. Struct. 29*, 105–153.
60. Neet, K. E., and Timm, D. E. (1994) Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation. *Protein Sci. 3*, 2167–2174.
61. Kimura, M. (1991) The neutral theory of molecular evolution: a review of recent evidence. *Jpn. J. Genet. 66*, 367–386.
62. Miller, J. H., Coulondre, C., Hofer, M., Schmeissner, U., Sommer, H., Schmitz, A., and Lu, P. (1979) Genetic studies of the lac repressor. IX. Generation of altered proteins by the suppression of nonsence mutations. *J. Mol. Biol. 131*, 191–222.
63. Bowie, J. U., Reidhaar-Olson, J. F., Lim, W. A., and Sauer, R. T. (1990) Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science 247*, 1306–1310.
64. Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N., and Tawfik, D. S. (2006) Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. *Nature 444*, 929–932.
65. Bershtein, S., and Tawfik, D. S. (2008) Advances in laboratory evolution of enzymes. *Curr. Opin. Chem. Biol. 12*, 151–158.
66. Tokuriki, N., Stricher, F., Schymkowitz, J., Serrano, L., and Tawfik, D. S. (2007) The stability effects of protein mutations appear to be universally distributed. *J. Mol. Biol. 369*, 1318–1332.
67. Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006) Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U.S.A. 103*, 5869–5874.
68. O'Loughlin, T. L., Patrick, W. M., and Matsumura, I. (2006) Natural history as a predictor of protein evolvability. *Protein Eng. Des. Sel. 19*, 439–442.
69. Wagner, A. (2008) Robustness and evolvability: a paradox resolved. *Proc. Biol. Sci. 275*, 91–100.
70. Bloom, J. D., Wilke, C. O., Arnold, F. H., and Adami, C. (2004) Stability and the evolvability of function in a model protein. *Biophys. J. 86*, 2758–2764.