The Apo and Ternary Complex Structures of a Chemotherapeutic Target: Human Glycinamide Ribonucleotide Transformylase[†,‡]

Tanya E. S. Dahms,[§,||] Germaine Sainz,[§] Eugene L. Giroux,[⊥] Carol A. Caperelli,*[,⊥] and Janet L. Smith*[,§,#]

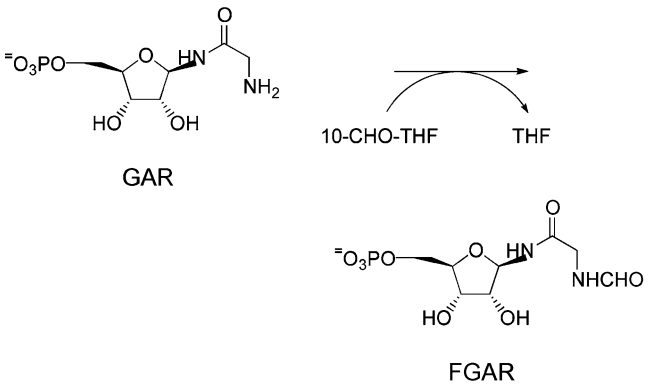
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, and College of Pharmacy, University of Cincinnati, Cincinnati, Ohio 45267

Received February 18, 2005; Revised Manuscript Received May 17, 2005

ABSTRACT: Glycinamide ribonucleotide transformylase (GART; 10-formyltetrahydrofolate:5′-phosphoribosylglycinamide formyltransferase, EC 2.1.2.2), an essential enzyme in de novo purine biosynthesis, has been a chemotherapeutic target for several decades. The three-dimensional structure of the GART domain from the human trifunctional enzyme has been solved by X-ray crystallography. Models of the apoenzyme, and a ternary complex with the 10-formyl-5,8-dideazafolate cosubstrate and a glycinamide ribonucleotide analogue, hydroxyacetamide ribonucleotide [α,*β*-*N*-(hydroxyacetyl)-D-ribofuranosylamine], are reported to 2.2 and 2.07 Å, respectively. The model of the apoenzyme represents the first structure of GART, from any source, with a completely unoccupied substrate and cosubstrate site, while the ternary complex is the first structure of the human GART domain that is bound at both the substrate and cosubstrate sites. A comparison of the two models therefore reveals subtle structural differences that reflect substrate and cosubstrate binding effects and implies roles for the invariant residues Gly 133, Gly 146, and His 137. Preactivation of the DDF formyl group appears to be key for catalysis, and structural flexibility of the active end of the substrate may facilitate nucleophilic attack. A change in pH, rather than folate binding, correlates with movement of the folate binding loop, whereas the phosphate binding loop position does not vary with pH. The electrostatic surface potentials of the human GART domain and *Escherichia coli* enzyme explain differences in the binding affinity of polyglutamylated folates, and these differences have implications to future chemotherapeutic agent design.

The discovery that glycinamide ribonucleotide transformylase (GART)[1] is the target of the potent antifolate 5,10-dideazatetrahydrofolate (DDATHF) (*1*) has promoted a resurgence of interest in both the enzyme and the purine biosynthetic pathway. In recent years, inhibitors of folate-dependent enzymes have been shown to demonstrate antitumor activity in vivo, and increasingly, inhibitors of these enzymes are being synthesized and tested for their chemotherapeutic activity (*2*).

In higher eukaryotes, the GART activity comprises the C-terminal domain of a trifunctional enzyme, which is also responsible for catalyzing the second (glycinamide ribonucleotide synthetase) and fifth (aminoimidazole ribonucleotide synthetase) reactions of de novo purine biosynthesis (*3*). GART catalyzes the third step of this biosynthetic pathway, namely, the conversion of glycinamide ribonucleotide (GAR) to *N*-formylglycinamide ribonucleotide (FGAR) accompanied by the conversion of the cosubstrate, 10-formyltetrahydrofolate (10-CHO-THF), to tetrahydrofolate (THF) (*4*, *5*).



Kinetic studies of the *E. coli* enzyme (*6*) and of the human (*7*) and murine (*8*) GART domains suggest a sequential mechanism in which the formyl group is transferred by a direct nucleophilic attack of the GAR amino group on the

[†] This work was supported in part by National Institutes of Health Grants GM-42663 to C.A.C. and DK-42303 to J.L.S., an Indiana Elk's Foundation grant to J.L.S., and a National Science and Engineering Research Council of Canada postdoctoral fellowship to T.E.S.D.

[‡] The atomic models are available in the Protein Data Bank with accession codes 1ZLX (apo) and 1ZLY (ternary complex).

* To whom correspondence should be addressed (JanetSmith@umich.edu; carol.caperelli@uc.edu).

[§] Department of Biological Sciences, Purdue University.

[||] Present address: Department of Chemistry and Biochemistry, Regina, SK, Canada S4S 0A2.

[⊥] College of Pharmacy, University of Cincinnati.

[#] Present address: Life Sciences Institute, University of Michigan, 210 Washtenaw Ave., Ann Arbor, MI 48109.

[1] Abbreviations: 10-CHO-THF, 10-formyl-5,6,7,8-tetrahydrofolate; 10-CF_3CO-DDACTHF, 10-trifluoroacetyl-5,10-didedideaza-acyclic-5,6,7,8-tetrahydrofolate; 10-CHO-DDACTHF, 10-formyl-5,10-didedideaza-acyclic-5,6,7,8-tetrahydrofolate; 10-CHO-DDF, 10-formyl-5,8-dideazafolate; DDF, 5,8-dideazafolate; FDH, 10-formyltetrahydrofolate dehydrogenase; FMT, 10-formyltetrahydrofolate:L-methionyl-tRNA *N*-formyltransferase; FGAR, *N*-formylglycinamide ribonucleotide; GAR, glycinamide ribonucleotide; GAR-OH, α,*β*-*N*-(hydroxyacetyl)-D-ribofuranosylamine; GART, glycinamide ribonucleotide transformylase or 10-formyltetrahydrofolate:5′-phosphoribosylglycinamide formyltransferase (EC 2.1.2.2).

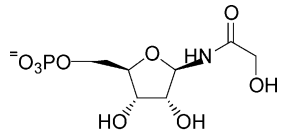
10.1021/bi050307g CCC: $30.25
Published on Web 06/28/2005

formyl carbon of the cosubstrate, leading to the formation of a tetrahedral intermediate. Formation of this intermediate and collapse to product require proton transfers. It has been proposed that a " fixed" water molecule, rather than the invariant amino acids in the active site (*9*), mediates the required proton transfer between substrate and cofactor, but this has not been verified.

Retention of the cosubstrate (or analogue) in the cell is affected by polyglutamylation of THF, DDATHF, and other potential folate analogue chemotherapeutic drugs (*10*). Previous studies (*8*) of murine GART have shown that the hexa-γ-glutamate of the 10-formyl-5,8-dideazafolate (10-CHO-DDF) cosubstrate binds almost 10-fold more tightly to the enzyme than the monoglutamate form and that the hexaglutamate of DDATHF was bound approximately 30-fold more tightly than the monoglutamate. To date, there is no structural information to explain the enhanced affinity resulting from polyglutamylation, but the high sequence identity between the murine and human GART domains (87%) merits a structural investigation of this phenomenon.

The structure of the *E. coli* GART enzyme has been solved at pH 3.5 (1.8 Å) (*11*), pH 6.75 (2.8 and 3.0 Å) (*12*, *13*), and pH 7.5 (1.9 Å) (*11*) and at neutral pH in complexes with GAR and 5-deazatetrahydrofolate (2.5 Å) (*12*), GAR and 10-formyl-5,8,10-trideazafolic acid (2.1 Å) (*14*), and two multisubstrate adduct inhibitors (1.96 and 1.6 Å) (*15*, *16*). The 38% sequence identity between the *E. coli* and human GART enzymes led to the premise that the *E. coli* GART structure could provide an appropriate model for its eukaryotic counterpart. However, recent kinetic (*17*) and structural (*18*) studies suggest that this is erroneous. The human GART domain is now readily available through cloning and overexpression (*7*, *19*); thus it is the most relevant protein for mechanistic studies and structure-based inhibitor design. Indeed, the structure of human GART has been reported at pH 4.2 (1.7 Å), at pH 8.5 (2 Å), at pH 8.5 in the binary complex with the substrate β-GAR (2.2 Å) (*18*) and at pH 7 in a binary complex with the cosubstrate analogue inhibitor 10-trifluoroacetyl-5,10-dideaza-acyclic-5,6,7,8-tetrahydrofolic acid (10-CF_3CO-DDACTHF) (2 Å) (*20*) and a series of folate inhibitors (*21*). Several questions remain regarding the structure-based mechanism of GART. First, does the pH of the protein environment affect the structure or dynamics of the enzyme, thereby regulating its activity? Wilson and co-workers have addressed this question by examining the *E. coli* (*11*) and human (*18*) GART structures at low and high pH values, representing the inactive and active forms of the enzyme, respectively. For both proteins, importance has been ascribed to structural differences as a function of pH. In the case of the *E. coli* enzyme, emphasis is placed on changes in the "flexible" (residues 111−131) and folate binding (residues 141−145) loops as a function of pH (*11*), while pH studies of the human GART enzyme highlight changes in the substrate (residues 8−14) binding site (*18*). The question remains as to the relevant structural changes associated with human GART activity solely as a function of pH. The pH (6.3) of the crystals used in this study is intermediate to the pH values of the two previously published models of human GART, affording structural comparisons as a function of pH.

To date, all crystal structures of GART, in the absence of substrate and/or cosubstrate, contain a phosphate, sulfate, or glycerol bound directly at, or just outside, the phosphate binding loop (residues 11−14) for the substrate, β-GAR. The human apoGART model reported herein represents a structure that is devoid of small molecules at either the β-GAR or 10-formyl-THF binding sites. The second model reported in this work corresponds to the first structure of human GART occupied at both the substrate and cosubstrate site. A structure has been determined for the ternary complex of an inhibitor/cosubstrate pair, hydroxyacetamide ribonucleotide [GAR-OH; $K_{is} \sim 1.3$ μM (*7*)], in which the side-chain



GAR-OH

terminal amino group of GAR is replaced by a hydroxy group, and 10-formyl-5,8-dideazafolic acid [10-CHO-DDF; $K_m \sim 1.5$ μM (*7*)] with the human GART domain. This unique pair of models offers the opportunity to investigate structural differences between the bound and unbound structures.

Herein we report the structures of the apo human GART domain and human GART ternary complex (GAR-OH/10-CHO-DDF) to 2.2 and 2.07 Å, respectively. We make a structural comparison of all human GART structures as a function of pH, examine the structural changes associated with substrate and cosubstrate binding, and propose a mechanism for the tight binding of polyglutamylated cosubstrate based on surface charge complementation.

EXPERIMENTAL PROCEDURES

Sequence Alignment. All known GART sequences (GART domains and monofunctional GART enzymes) were aligned with ClustalW (*22*).

Protein Preparation. The rhGART domain, which consists of residues 808−1010 of the GARS-AIRS-GART human trifunctional enzyme, was prepared and assayed according to published procedures (*7*). The protein was ultrapurified to >99% homogeneity by anion-exchange (DEAE-5PW, TSK, 7.5 mm × 7.5 cm, TosoHaas) high-performance liquid chromatography. A linear gradient (30 min from 0 to 0.2 M NaCl in 10 mM Tris, pH 7.5) was used for elution, and the peak fraction between the full width at half-maximum of the predominant peak (~23 min; data not shown) was collected. The HPLC-purified hGART was assayed (*7*) to verify formyltransferase activity. Selenomethionyl (SeMet) GART was produced in *E. coli* strain B834(DE3)pLysS (Met auxotroph) grown on defined media (*23*) containing 0.3 mM SeMet and 50 mg/L ampicillin. To minimize oxidation of the SeMet, 10 mM dithioerythritol (DTE) was included in all media and buffers.

Crystallization. All crystals were grown by vapor diffusion from a 1:1 mixture of protein solution (5 mg/mL GART, 10 mM Hepes, pH 7.5, 1 mM DTE) and well solution. Small crystals grew spontaneously with a well solution consisting of 0.1 M sodium citrate, pH 6.0, 20% 2-propanol (v/v), and 20% PEG-4000 (w/v). Larger hexagonal bipyramids (0.5 mm × 0.4 mm × 0.4 mm) of wild-type and SeMet GART were grown by seeding small crystals into fresh sitting drops using

Table 1: Crystallographic Summary

	GART	GART-DDF-GAR-OH
Diffraction Data		
space group	$P3_221$	$P3_221$
unit cell constants (Å)	75.7, 75.7, 101.7	75.9, 75.9, 101.0
X-ray source	CHESS F2	CHESS F2
data range (Å)	40.2−2.2	40.0−2.07
completeness (%)	97.9 (99.8)[c]	99.5 (97.4)
R_{sym}[a] (%)	8.9 (31)	7.3 (29.6)
$\langle I/\sigma_I \rangle$	8.3 (7.7)	7.7 (3.7)
unique reflections (no.)	20705	21025
average redundancy	6.9	3.9
Model Refinement		
data range (Å)	40.2−2.2	40.0−2.07
data cutoff	$F < 0$	$F < 0$
R_{work}[b] (%)	22.4	22.2
reflections (no.)	15686	18900
R_{free}[b] (%)	26.7	23.6
reflections (no.)	871	1015
residue range	1−200	1−200
solvent sites (no.)	248	335
average B (Å^2)	56.2	53.8
RMSD from target values		
bond lengths (Å)	0.007	0.006
bond angles (deg)	1.34	1.45
bonded B (Å^2)	4.47	2.94
estd coordinate error (Å)	0.24	0.28
overall anisotropic B-factor (Å^2)		
B_{11}, B_{22}	−7.51	−5.79
B_{12}	−3.13	−3.49
B_{33}	15.03	11.59

[a] $R_{sym} = \sum_j |I_h - I_j| / \sum_j I_j$, where I_h = average intensity for all observations of a reflection with unique indices h and $I_j = j$th observation of a reflection with average intensity I_h. [b] $R = \sum |F_o - F_c| / \sum |F_o|$. R_{free} is based on a 5% subset of the data. [c] Values in parentheses pertain to the outermost shell of data.

well solutions consisting of 0.1 M sodium citrate, pH 6.3, 10% 2-propanol, and 18% PEG-4000. Crystals were dissolved and tested for enzyme activity (*7*) to eliminate the possibility of protein inactivation during crystallization. Crystals of the ternary complex of GART with GAR-OH, a substrate analogue and competitive inhibitor, and 10-CHO-DDF, a cosubstrate, were produced by addition of 0.24 mM GAR-OH and 0.45 mM 10-CHO-DDF to a sitting drop containing one or more GART crystals. Immediately prior to crystal soaking, the integrity of 10-CHO-DDF was assessed spectroscopically, where loss of the formyl group alters the UV−vis spectrum. Crystals of the SeMet GART ternary complex were also grown by seeding into sitting drops with well solutions that contained 0.1 M sodium citrate, pH 6.3, 6−8% 2-propanol, and 16−18% PEG-4000. Crystals were cryoprotected by successive transfer into well solution (4 °C) augmented by increasing glycerol in 5% increments (2 min/increment) to a final concentration of 20%.

Data Collection. Preliminary 2.9 Å data were collected using Cu Kα radiation and a Rigaku R-axis IIc image plate detector for the wild-type free enzyme at room temperature and 2.85 Å data for the SeMet GART ternary complex at 2 °C. These data were used for initial phasing, model building, and refinement. Data used for the final refinements were collected on beamline F2 at CHESS using a Quantum 4 CCD detector (Area Detector Systems Corp.) from frozen crystals (−178 °C) of wild-type GART. Data were processed using Denzo and Scalepack (*24*) (Table 1), and the CCP4 suite of programs was used for further data manipulation (*25*). The space group was determined to be $P3_221$ or its enantiomorph ($P3_121$) with one (V_m = 3.7 Å^3/Da) or two monomers (V_m = 1.85 Å^3/Da) in the asymmetric unit.

Phase Determination. The structure of the apoenzyme was solved by molecular replacement from a truncated model of the 38% identical *E. coli* enzyme [2.0 Å; PDB entry 1GAR (*15*)] using the program AMoRe (*25*, *26*). The probe structure, for which nonidentical residues were truncated to alanine, contained approximately 70% of the total protein atoms. The *R*-factor (0.48%) and correlation coefficient (0.58) were determined for the correct solution, and the space group was found to be $P3_221$.

The molecular replacement solution contained only one molecule per asymmetric unit and a high solvent content (73%). Two Se sites from the SeMet-substituted GART (residues 89 and 104) were located in an $F_{SeMet} - F_{wild\ type}$ difference map, confirming the molecular replacement solution, the positions of the Met side chains, and the presence of only one GART domain per asymmetric unit.

Model Building and Structure Refinement. All model building was done using the program O (*27*), and refinement was conducted using X-PLOR 3.8 (*28*) and CNS (*29*). Rigid body refinement of the molecular replacement model was used initially on the entire molecule and then on individual secondary structural elements. The high solvent content of the GART crystals was exploited for phase refinement using SOLOMON (*25*, *30*).

Frozen GART crystals diffracted significantly further than unfrozen crystals but exhibited highly anisotropic diffraction (Table 1). The final models of free GART and the ternary complex contain 200 of 203 amino acids. No electron density was observed for the three residues at the C-terminus (201−203). The free GART model contains one glycerol and 248 water molecules, while the ternary complex contains GAR-OH up to and including the amide nitrogen, all of 10-CHO-DDF except the formyl group and α-carboxylate, and 335 water molecules. Structures were compared using a least-squares analysis (DEJAVU) (*25*, *31*).

RESULTS AND DISCUSSION

Structure Description. Common features of the hGART domain can be viewed in Figure 1a, and the active site is highlighted in Figure 1b. Figure 1a shows the relationship between the active site, the flexible region (residues 111−131) thought to be important for folate binding, the phosphate (residues 11−14) and folate binding (residues 141−145) loops, and the N- and C-termini. The overall structural details of human GART have recently been described (*18*), while the topology (*12*) and structural details (*12*, *13*) have been described elsewhere for the similar *E. coli* enzyme. The first residue (Figure 1a) in the human GART domain is numbered according to the *E. coli* enzyme to facilitate comparisons between the homologous enzymes, where Ala 1 of the GART domain corresponds to Ala 808 of the human trifunctional enzyme (*3*). Contrary to all previous "apo" GART structures, there are no small molecules associated with the phosphate loop of the substrate binding site of this hGART model, possibly as a result of the additional and more rigorous HPLC purification procedure used in these studies.

Structural Comparisons. Among the crystal structures of the human GART domain, in the presence and absence of

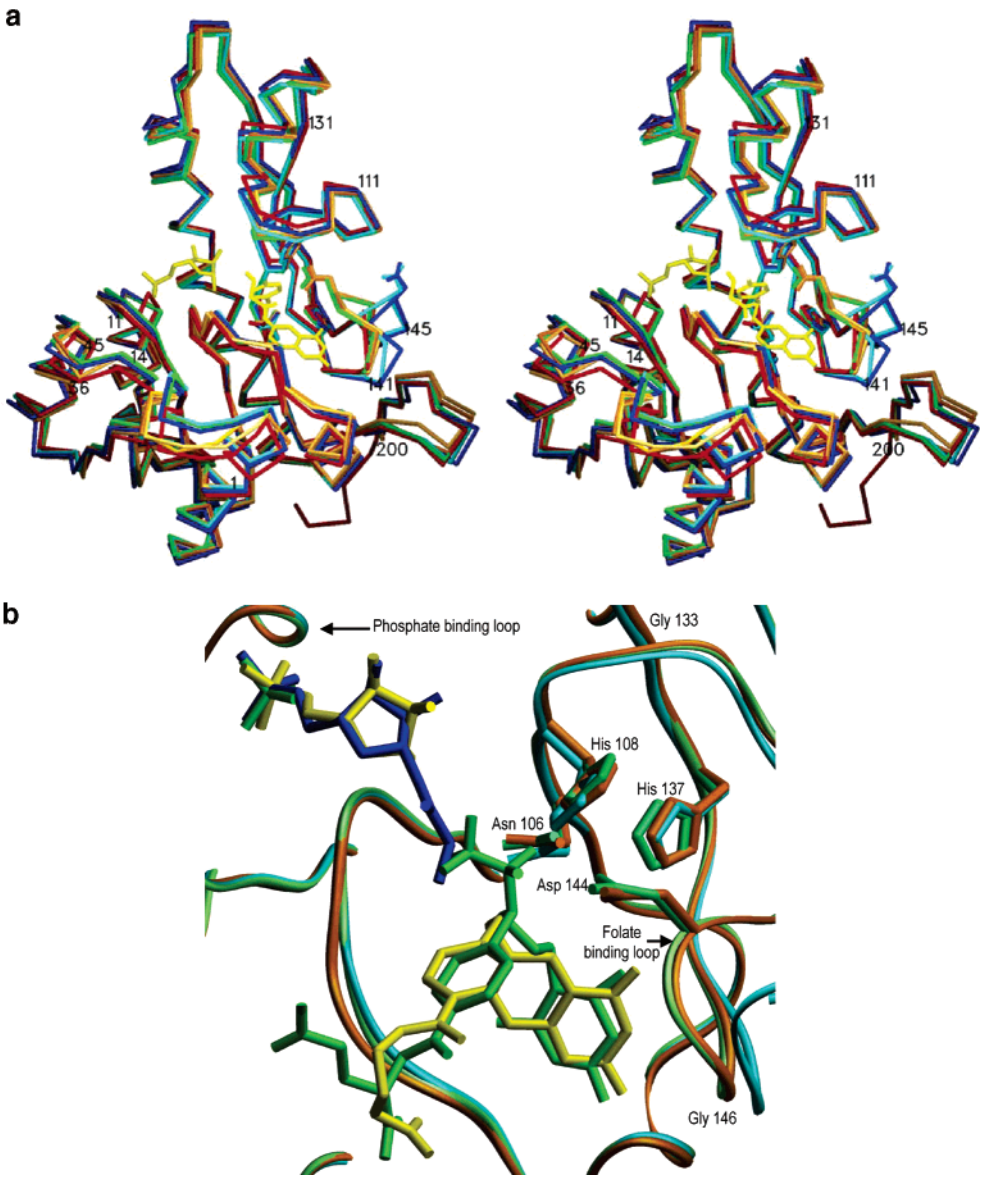


FIGURE 1: Overlay of the C^{α} coordinates for all high-resolution models of human GART: apo hGART (dark gold; this work), hGART/GAR-OH/CHO-DDF (light gold and yellow, respectively; this work), hGART, pH 4.2 (red; PDB 1MEO; *18*), hGART, pH 8.5 (blue; PDB 1MEJ; *18*), hGART/10-CF_3CO-DDACTHF (green; PDB 1NJS; *20*), hGART/GAR (cyan; PDB 1MEN; *18*). (a) Backbone traces of apo and complex structures. (b) Close-up view of the complexed structure backbones with bound substrates, cosubstrates, and inhibitors. Figures 1−4 were generated using the program SETOR (*36*).

substrates/inhibitors, the most significant structural variability occurs in the cosubstrate (folate) binding loop (residues 141−145) (Figure 1a). In addition, the outer loop region defined by residues 36−45, which is adjacent to the phosphate binding loop, shows slightly different positions for the various hGART structures.

The folate binding loop (residues 141−145), which displays significant variability in the hGART structures (see Figure 1a), is directly preceded by residue 140, which is involved in contacts with the cosubstrate pteridine ring in all of the complexed structures (human and *E. coli*). This folate binding loop displays a range of positions, 2−5 Å away from the pteridine ring, that do not appear to correlate with folate binding but rather changes in pH. The folate binding loop of hGART is largely disordered at pH 4.2, is ordered and closest to the pteridine ring at neutral pH (6.3, 7; apo and ternary), and is furthest from the pteridine ring at pH 8.5 (apo and GAR-bound) (Figure 1a).

Asp 144 in the folate binding loop was identified as being catalytically important in the *E. coli* enzyme (*32*). The Asp 144 side chain is positioned toward the folate site in all of the neutral pH structures (apo, DDF-bound, and ternary) but points away from the site at higher pH (8.5; apo and GAR-bound). In fact, the position of Asp 144 is nearly identical for bound and unbound forms of the enzyme at neutral pH, and therefore we conclude that folate binding is not a requirement for the correct positioning of the active site Asp 144, as has been suggested for the *E. coli* enzyme (*33*) and hGART (*18*). The high pK_a of 9.7, ascribed to His 108 in the *E. coli* enzyme, is attributed to the Asp144−His108 salt bridge (*33*). However, in both of the structures reported here, Asp 144 is near His 108 (3.4 Å), but the His imidazole and Asp144 O are not coplanar, ruling out a hydrogen bond under these conditions. In contrast to Asp 144, Asp 142 and Val 143 in the folate binding loop move in response to folate binding, while the Glu 141 side chain and its *E. coli*

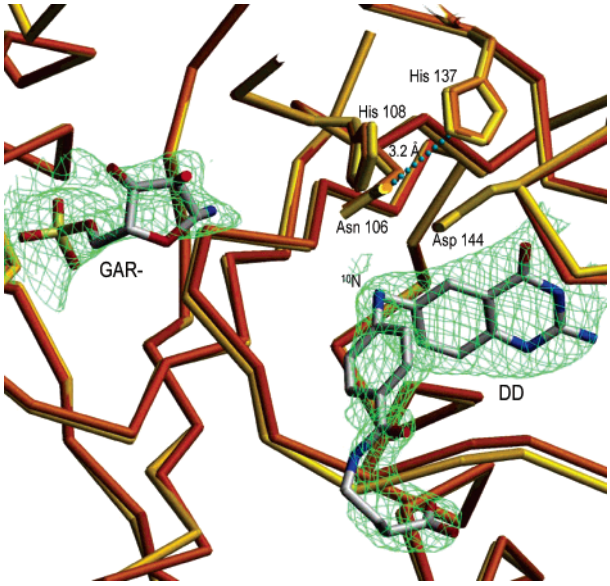


FIGURE 2: Superposition of apo and ternary hGART models to show the structural similarity of the active site and its invariant amino acids (Asn 106, His 108, Gly 133, His 137, Asp 144, and Gly 146) in the presence and absence of GAR-OH and CHO-DDF. The apo hGART is rendered in light gold, the ternary hGART complex in dark gold, and GAR-OH and CHO-DDF with atomic colors, and the $2F_o - F_c$ electron density map contoured at 0.8 RMS is shown in green for the substrate and cosubstrate.

counterpart, Asp 141, are positioned away from the pteridine binding site in all cases.

Substrate and Inhibitor Binding. The binding mode of the cosubstrate and inhibitor in the human GART ternary structure is similar to that observed for analogous ligands bound to the *E. coli* (*12*, *14*) ternary complexes and the human (*18*, *20*) binary complexes (Figure 1b). The electron density map (Figure 2) displays clear density for much of the GAR-OH inhibitor and DDF cosubstrate but not for the reactive groups: the GAR-OH side-chain hydroxyacetyl and the DDF formyl group. Although weak density appeared for the substrate hydroxyacetyl group, density for the entire formyl group was lacking whether ternary complex crystals were produced by soaking (Table 1) or by cocrystallization (data not shown). We conclude that the active site environment is particularly conducive to removal of the labile formyl group when a substrate or analogue is present. The ternary complex structure thus represents a mixed substrate−product complex, with GAR-OH as substrate analogue and deformylated 10-CHO-DDF (DDF) as product. Steric effects may have contributed to departure of the formyl group and to disorder of the GAR-OH hydroxyacetamide, which would be less than 1 Å from the formyl oxygen of 10-CHO-DDF based on the binary hGART complexes. There is little difference between the apo and ternary complex structures (Figure 2), especially at the phosphate binding loop. Unlike all other GART crystal structures, the apoenzyme structure lacks fortuitous binding of ions at the phosphate binding loop of the GAR binding site that would have to be replaced by its substrate.

The GAR-OH substrate analogue binds the enzyme with its phosphate in a loop (residues 11−14) at the N-terminus of a helix and is hydrogen bonded to backbone N−H groups of Ser 12 and Asn 13, to a water molecule anchored to the backbone at Leu 14, and to the Ser 12 side chain (Figure 3a). The ribose 3′-hydroxyl is hydrogen bonded to the Glu 173 side chain, and the GAR amide nitrogen is within hydrogen-bonding distance of the backbone carbonyl of Ile 107 (Figure 3a). These interactions are in general agreement with previous structures of GAR bound to the *E. coli* (*12*, *14*) and human (*18*) GART enzymes. The *B*-factors are high for the GAR-OH inhibitor, and the inhibitor is poorly ordered at the hydroxyl end (Figure 2).

The DDF product is well ordered at the bicyclic pteridine ring and becomes disordered toward the carboxylate end. The electron density is continuous to the γ-carboxylate, which clearly protrudes from the substrate binding cavity, but density is missing for the α-carboxylate (Figure 2). The bicyclic ring hydrogen bonds (Figure 3b) with the backbone NH groups of Leu 92 (N1) and Asp 144 (4-oxo) and the backbone carbonyl groups of Ala 140 (N3) and Glu 141

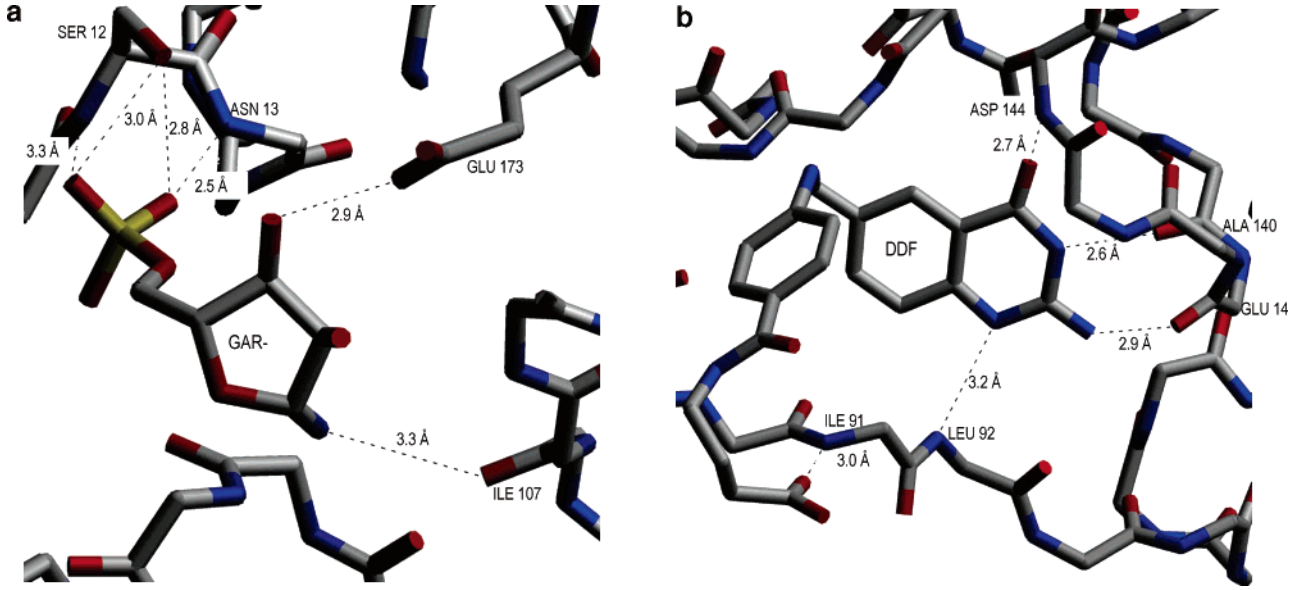
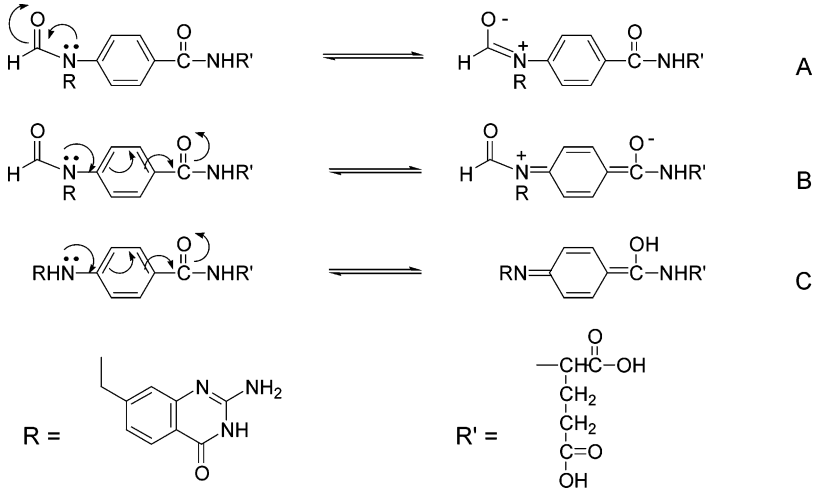


FIGURE 3: Close-up view of the binding sites of (a) GAR-OH and (b) CHO-DDF for the hGART ternary complex, including hydrogen-bonding interactions between substrate/cosubstrate and protein. The C^α coordinates of the ternary coordinates, along with GAR-OH and CHO-DDF, are shown in atomic colors.

Scheme 1



(NA2), in general agreement with the cosubstrate binding pattern for *E. coli* GART (*12*, *14*). One oxygen atom from the terminal carboxylate is hydrogen bonded to the amide NH of Ile 91.

GART was initially thought to follow an ordered sequential mechanism, but studies on the murine recombinant enzyme revealed that the binding mechanism was random sequential (*8*), leading to studies of the *E. coli* (*6*) enzyme that corrected the previous literature and suggested random addition. Thus, random sequential binding and direct nucleophilic attack have been proposed for the *E. coli* (*6*), murine (*8*), and human (*7*) GART enzymes. The structural flexibility of the substrate and cosubstrate and the preordering of the phosphate binding loop are consistent with the proposed reaction and binding mechanisms, respectively. The honeycomb packing of the crystals used in this study results in minimal crystal contacts and a high solvent content (>70%), providing a good model for GART in solution. The structural flexibility of the substrate (B_{av} = 79 Å^2) and cosubstrate (B_{av} = 71 Å^2) may facilitate the GART reaction, and the disordered portions of each substrate may represent flexibility that allows for optimal positioning of the reactive groups for a direct nucleophilic attack. We suggest, however, that binding of the substrate ribose ring along with the bicyclic ring of the cosubstrate may be important for ordering the entire active site.

It has been suggested (*33*) that the formyl group is activated by its twisting "out of the delocalized planar structure to relieve the resonance stabilization". Scheme 1 shows that there are two pathways for delocalization of the N^{10} lone pair. Path A would lead to deactivation of the formyl group toward nucleophilic attack, while path B would increase the positive charge on the formyl carbon, increasing its electrophilicity. In free DDF, one would expect that the lone pair would be delocalized into the benzene ring (C). The conformation of DDF bound to the enzyme (Figures 2 and 3B) suggests overlap of the N^{10}-nonbonding orbital with the π system of the benzene ring, favoring path B and deformylation of 10-CHO-DDF. If this conformation reflects that of enzyme-bound 10-CHO-DDF, this is the first direct structural evidence for conformational preactivation of the cosubstrate formyl group. We speculate that free 10-CHO-DDF would exist in a conformation with minimal overlap of the N^{10} lone pair with the benzene π system and that enzyme binding induces the conformational change.

This idea is also supported by the apparent lability of the DDF formyl group when bound to the enzyme in the presence of a GAR analogue inhibitor. The 10-CHO-DDF used in these experiments was intact when added to solutions used to soak protein crystals. We have found 10-CHO-DDF to be stable in aqueous solutions at neutral pH and, therefore, conclude that the enzyme accelerates loss of the formyl group.

A random-collision mechanism for catalysis is suggested by the preactivated cosubstrate and the flexibility at the reactive end of the substrate. If 10-CHO-DDF, with a preactivated formyl leaving group, and GAR, with a flexible reactive end, are bound in extremely high local concentration, then catalysis could proceed easily by random collision. In this circumstance, it seems unnecessary to preorient the reactive end of GAR. The enzyme need only provide an environment in which the GAR amino group is unprotonated and perhaps stabilize a transient negative charge on the formyl oxygen.

Role of Invariant Amino Acids. The sequence of the GART domain within the trifunctional enzyme is highly conserved. For example, 41% of residues are invariant among four eukaryotic organisms (human, mouse, chicken, fruit fly) (*3*), and the human GART domain is 38% identical to the monofunctional *E. coli* enzyme. Eighteen sites of conservative substitution are distributed throughout the molecule, and some of these residues are important in protein−substrate interactions. Residues Asn 106, His 108, Gly 133, His 137, Asp 144, and Gly 146 are invariant among the GART domains and enzymes. The invariant residues are in or adjacent to the GART active site. His 108, Gly133, Asp 144, and Gly146 are also invariant within the putative folate binding site of other 10-CHO-THF-utilizing enzymes (*34*). Although GART and 10-formyltetrahydrofolate:L-methionyl-tRNA *N*-formyltransferase (FMT) have a low sequence identity, they are structural homologues (*35*). Modeling studies of the 10-formyltetrahydrofolate dehydrogenase (FDH) structure indicate that it has the same fold as GART (*34*). FMT and FDH have putative catalytic residues analogous to His 108 and Asp 144 in GART, and FMT has the additional Asn 106 analogue that may be critical to formyltransferase activity (*34*), but neither related enzyme has an analogue of GART His 137.

The positions of Gly 133, His 137 (including the side chain), and Gly 146 are virtually identical for all GART

structures (human and *E. coli*) in the presence and absence of ligands. The role of His 137 has been largely ignored for both the human and *E. coli* enzyme. The fact that His 137 is completely conserved in GART but not in FMT and FDH indicates it may have a specific role in the GART mechanism. The position of the His 137 imidazole is fixed by a hydrogen bond to Asn 106 in both the apoGART and ternary complex (Figure 2), consistent with electrostatic calculations of the *E. coli* enzyme (*36*) that suggest His 137 is part of the proton relay system for removal of a proton from GAR and addition of a proton to the N^{10} of folate. The structure reported in this work shows evidence of four ordered water molecules in the active site, but none is found in a position suitable to act as the catalytic water proposed for the *E. coli* GART mechanism (*9*).

Gly 146 (Figure 1b) is located at the hinge point of the folate binding loop in all GART structures and may confer flexibility on the loop to facilitate folate binding, whereas conserved Gly 133 likely plays a pivotal role in the mobility of the flexible loop comprising residues 110−131 (Figure 1). On the basis of the loop (110−131) flexibility shown in the structural comparison (Figure 1) and the position of the conserved Gly 133 at its hinge point, this residue may be poised to play a pivotal role in the mobility of this region. Substitution of five residues (118−122) at the tip of the loop with Ser-Ser in the *E. coli* enzyme reduced its affinity for substrate (*9*), demonstrating the importance of the loop to substrate binding. The loop occupies a variety of positions and is disordered in some crystal structures. There is further support that Gly 133 and Gly 146 are pivotal to cosubstrate binding since they are strictly conserved not only in GART from different species but also in the related 10-CHO-THF-utilizing enzymes (*34*).

As previously stated, the position of Asp 144 is virtually identical for the apo and ternary complexes at pH 6.3. Therefore, these structures do not support the mechanism proposed on the basis of the structure of the *E. coli* enzyme, in which folate binding is a prerequisite to active site rearrangements allowing a salt bridge, "critical to catalysis", to form between His 108 and Asp 144 (*33*). In the apo and ternary hGART models, His 108 is not close enough to Asp 144 (3.5 Å apart), nor properly oriented to form a salt bridge (Figure 2).

Residues Asn 106 and His 108 make hydrogen bonds with the formyl group of the multisubstrate adduct complexed to the *E. coli* enzyme (*15*), implying that Asn 106 is directly involved in the catalytic mechanism. The formyl group is missing in the ternary complex structure reported herein, precluding comparison. Pairwise electrostatic interactions of conserved residues in the *E. coli* enzyme (*36*) have been used to examine the electrostatic environment of the GART active site and indicate that neutral His 137 relays an electrostatic interaction to GAR(NH_2) through Asn 106 and ultimately Asp 144 and His 108, thus contributing to the proton relay between GAR(NH_2) and His 108. The positions of Asn 106, His 108, His 137, and Asp 144 side chains are superimposable between the ternary hGART complex (this work) and the *E. coli* structure used for the computational studies (*12*). Given that His 137 is within hydrogen-bonding distance of Asn 106 (Figure 2), this residue may play a similar role in the hGART enzyme.

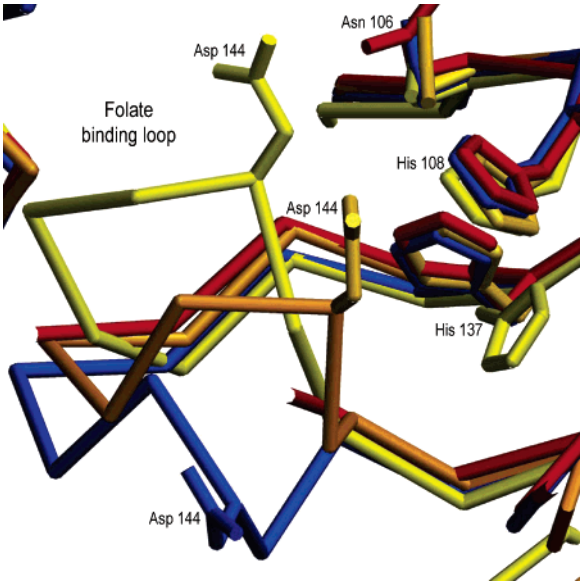


FIGURE 4: Comparison of the C^{α} coordinates of the folate binding loop, including invariant residues, for high-resolution models of hGART at pH 4.2, 6.3, and 8.5 and *E. coli* GART at pH 3.5 (PDB 2GAR; *11*). Backbone colors are the same as in Figure 1a, and the *E. coli* GART at pH 3.5 is rendered in yellow.

The hydrogen bond between the Asp 144 and His 108 side chains has been proposed as a critical determinant for the catalytic function of the *E. coli* enzyme by altering the pK_a of His 108, allowing it to act as a general acid catalyst (*33*), an idea that is supported by Poisson−Boltzmann-based electrostatic models (*36*). This interaction is not observed in the hGART enzyme, but it is nonetheless possible that this interplay between active site residues is involved in hGART enzyme activity.

Effects of pH on GART Structure. The human (*18*) and *E. coli* (*33*) enzymes show similar pH rate optima at approximately pH 8 (*36*). Crystal structures of human and *E. coli* GART at pH values ranging from 3.5 to 8.5 have various loops in different positions as a function of pH in the cosubstrate-free structures (Figure 4), but the loop positions are identical in the cosubstrate complexes at all pHs. In the absence of cosubstrate, the folate binding loop (residues 137−147) encroaches progressively deeper into the cosubstrate binding site with decreasing pH. However, the position of the loop at pH 6.3 is closest to its position in the presence of cosubstrates, which is virtually identical in all complexed structures. Therefore, a pH-dependent loop motion is observed for the free enzyme, but cosubstrate binding stabilizes the loop at the position observed in the structure derived from near-physiological pH (6.3). Further, the folate-dependent conformational changes, which were ascribed relevance in previous studies (*33*), do not appear to be relevant to the hGART structure because the folate loop is ordered both in the presence and in the absence of cosubstrate (Figure 2).

The substrate binding loop (residues 8−14), which cradles the GAR phosphate group, also exhibits different conformations in various crystal structures (Figure 1a). Zhang et al. highlighted an occluded substrate loop as the major source of hGART's inability to bind GAR and its inactivity at pH 4.2 (*18*). Others have suggested that the loss of activity for the *E. coli* enzyme at low pH results from titration of the

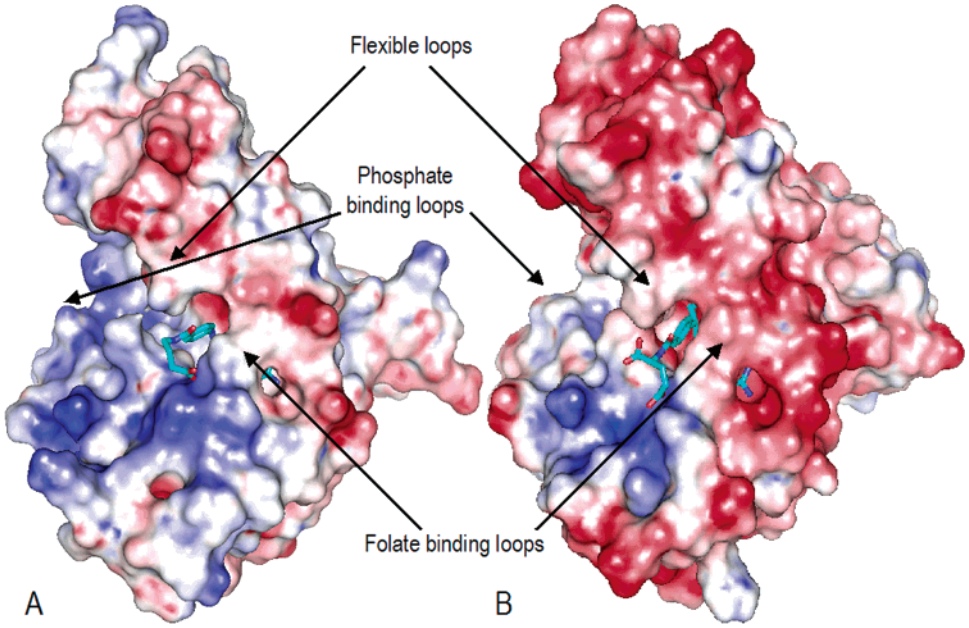


FIGURE 5: Surface charge representation of (a) the human GART structure in the presence of GAR-OH/CHO-DDF (this work) and (b) *E. coli* GART in complex with GAR and 10-formyltetrahydrofolate (PDB 1C3E; *14*) calculated using the programs GROMACS (*38*) and MEAD (*39*) and rendered by the program PyMol (40). The folate cofactor is shown as a stick representation using standard atomic colors (except cyan = carbon), and the surface potential color scheme indicates areas that are neutral (white, 0*kT*), positively (red, +6.7*kT*) and negatively (blue, −6.7*kT*) charged. The view is similar here and in Figure 1a.

GAR amino group (*33*, *36*). Direct comparison of the phosphate binding loops is not straightforward because only the pH 6.3 structure represents a truly apo structure. The difference in loop position (~2.3 Å, Figure 1a) observed at pH 4.2 is not observed either for the *E. coli* GART structure (*11*) at pH 3.5 (inactive) or for the human GART structure reported here at pH 6.3 (active). In fact, the phosphate binding loops in human GART structures at pH 8.5 (*18*) and 6.3, along with the *E. coli* GART structure at pH 3.5 (*11*), are virtually identical (Figure 1a; data not shown for *E. coli* structure). Therefore, previously observed phosphate loop movements (*18*) do not explain the pH profile for hGART and likely reflect extraneous ligand binding.

Implications of Surface Charge Pattern. A surface potential diagram has been constructed for the protein portion of the ternary complex of human GART (Figure 5a) and its *E. coli* counterpart (Figure 5b) (*14*). The electrostatic potential at the surface of hGART, extending away from the glutamyl region of the DDF and toward the outer side of the folate binding loop, is either neutral or positively charged and, therefore, would provide a suitable surface (neutral and positive) for interaction with a polyglutamate (Figure 5a) with alternate neutral and negatively charged regions. In contrast, the outer surface of the *E. coli* GART folate binding loop is predominantly negative, and the only surface complementary (positive/neutral) to a polyglutamate leads away from the folate binding loop (Figure 5b). Previous studies (*8*) examined the binding of folate substrates and inhibitors to recombinant mouse GART, finding that polyglutamylation of the tight-binding DDF ligands enhanced affinity of both substrate and inhibitor by an order of magnitude. Marsilje et al. (*17*) examined 10-CHO-DDACTHF and its pentaglutamate derivative as inhibitors of recombinant *E. coli* and human GART, revealing that the γ-pentaglutamylated inhibitor was more potent against the human enzyme (K_i ~ 14 nM) compared to the *E. coli* enzyme (K_i ~ 6 μM). This increased affinity of hGART for the polyglutamylated DDF is likely the result of charge complementarity of the glutamyl portion of the cofactor and the outer surface of the folate binding loop (141−145), giving rise to further stabilization of the substrate binding pocket. A structural role is certainly not unprecedented for the hexaglutamate moiety (*37*), and the comparison presented herein provides a plausible explanation for the role of polyglutamylation in the hGART reaction.

Conclusions. Comparison of the first apo and ternary hGART structures reveals that limited structural changes are induced with binding of both the substrate and cosubstrate, and their protein interactions are consistent with earlier work. We have elucidated roles for the invariant residues Gly 133, Gly 146, and His 137 and propose that substrate/cosubstrate flexibility and cosubstrate preactivation are critical for the catalytic mechanism of human GART. A structural comparison was conducted to identify flexible areas that are likely important to facilitate substrate and cosubstrate binding. The pH affects the position of the folate binding loop in the absence of cosubstrate, but at pH 6.3 folate binding does not appear to drive the position of the folate binding loop. Finally, the surface charge on the hGART exterior, leading from the γ-carboxylate of the monoglutamate to the folate binding loop, appears to be complementary to the charge pattern expected for a hexaglutamate and may explain the tight binding of the cosubstrate to the human GART domain conferred by glutamylation. This finding is relevant to future inhibitor design and to the in vivo hGART activity for which folates are known to be polyglutamylated.

ACKNOWLEDGMENT

We thank the staff of the MacCHESS facility for helpful support, and T.E.S.D. thanks Dr. J. Krahn for generous

advice on computation and data collection and Drs. S. Evans and P. Martel for software support and guidance.

REFERENCES

1. Beardsley, G. P., Moroson, B. A., Taylor, E. C., and Moran, R. G. (1989) A new folate antimetabolite, 5,10-dideaza-5,6,7,8-tetrahydrofolate is a potent inhibitor of de novo purine synthesis, *J. Biol. Chem. 264*, 328−333.
2. Costi, M. P., and Ferrari, S. (2001) Update on antifolate drugs targets, *Curr. Drug Targets 2*, 135−166.
3. Aimi, J., Qiu, H., Williams, J., Zalkin, H., and Dixon, J. E. (1990) De novo purine nucleotide biosynthesis: cloning of human and avian cDNAs encoding the trifunctional glycinamide ribonucleotide synthetase-aminoimidazole ribonucleotide synthetase-glycinamide ribonucleotide transformylase by functional complementation in *E. coli*, *Nucleic Acids Res. 18*, 6665−6672.
4. Dev, I. K., and Harvey, R. J. (1978) N10-Formyltetrahydrofolate is the formyl donor for glycinamide ribonucleotide transformylase in *Escherichia coli*, *J. Biol. Chem. 253*, 4242−4245.
5. Smith, G. K., Mueller, W. T., Benkovic, P. A., and Benkovic, S. J. (1981) On the cofactor specificity of glycinamide ribonucleotide and 5-aminoimidazole-4-carboxamide ribonucleotide transformylase from chicken liver, *Biochemistry 20*, 1241−1245.
6. Shim, J. H., and Benkovic, S. J. (1998) Evaluation of the kinetic mechanism of *Escherichia coli* glycinamide ribonucleotide transformylase, *Biochemistry 37*, 8776−8782.
7. Caperelli, C. A., and Giroux, E. (1997) The human glycinamide ribonucleotide transformylase domain: purification, characterization, and kinetic mechanism, *Arch. Biochem. Biophys. 341*, 98−103.
8. Sanghani, S. P., and Moran, R. G. (1997) Tight binding of folate substrates and inhibitors to recombinant mouse glycinamide ribonucleotide formyltransferase, *Biochemistry 36*, 10506−10516.
9. Warren, M. S., Marolewski, A. E., and Benkovic, S. J. (1996) A rapid screen of active site mutants in glycinamide ribonucleotide transformylase, *Biochemistry 35*, 8855−8862.
10. Moran, R. G. (1999) Roles of folylpoly-gamma-glutamate synthetase in therapeutics with tetrahydrofolate antimetabolites: an overview, *Semin. Oncol. 42*, 457−464.
11. Su, Y., Yamashita, M. M., Greasley, S. E., Mullen, C. A., Shim, J. H., Jennings, P. A., Benkovic, S. J., and Wilson, I. A. (1998) A pH-dependent stabilization of an active site loop observed from low and high pH crystal structures of mutant monomeric glycinamide ribonucleotide transformylase at 1.8 to 1.9 Å, *J. Mol. Biol. 281*, 485−499.
12. Almassy, R. J., Janson, C. A., Kan, C.-C., and Hostomska, K. (1992) Structures of apo and complexed *Escherichia coli* glycinamide ribonucleotide transformylase, *Proc. Natl. Acad. Sci. U.S.A. 89*, 6114−6118.
13. Chen, P., Schulze-Gahmen, U., Stura, E. A., Inglese, J., Johnson, D. L., Marolewski, A., Benkovic, S. J., and Wilson, I. A. (1992) Crystal structure of glycinamide ribonucleotide transformylase from *Escherichia coli* at 3.0 Å resolution. A target enzyme for chemotherapy, *J. Mol. Biol. 227*, 283−292.
14. Greasley, S. E., Yamashita, M. M., Cai, H., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (1999) New insights into inhibitor design from the crystal structure and NMR studies of *Escherichia coli* GAR transformylase in complex with beta-GAR and 10-formyl-5,8,10-trideazafolic acid, *Biochemistry 38*, 16783−16793.
15. Klein, C., Chen, P., Arevalo, J. H., Stura, E. A., Marolewski, A., Warren, M. S., Benkovic, S. J., and Wilson, I. A. (1995) Towards structure-based drug design: crystal structure of a multisubstrate adduct complex of glycinamide ribonucleotide transformylase at 1.96 Å resolution, *J. Mol. Biol. 249*, 153−175.
16. Greasley, S. E., Marsilje, T. H., Cai, H., Baker, S., Benkovic, S. J., Boger, D. L., and Wilson, I. A. (2001) Unexpected formation of an epoxide-derived multisubstrate adduct inhibitor on the active site of GAR transformylase, *Biochemistry 40*, 13538−13547.
17. Marsilje, T. H., Labroli, M. A., Hedrick, M. P., Jin, Q., Desharnais, J., Baker, S. J., Gooljarsingh, L. T., Ramcharan, J., Tavassoli, A., Zhang, Y., Wilson, I. A., Beardsley, G. P., Benkovic, S. J., and Boger, D. L. (2002) 10-Formyl-5,10-dideaza-acyclic-5,6,7,8-tetrahydrofolic acid (10-formyl-DDACTHF): a potent cytotoxic agent acting by selective inhibition of human GAR Tfase and the de novo purine biosynthetic pathway, *Bioorg. Med. Chem. 10*, 2739−2749.
18. Zhang, Y., Desharnais, J., Greasley, S. E., Beardsley, G. P., Boger, D. L., and Wilson, I. A. (2002) Crystal structures of human GAR Tfase at low and high pH and with substrate beta-GAR, *Biochemistry 41*, 14206−14215.
19. Kan, C.-C., Gehring, M. R., Nodes, B. R., Janson, C. A., Almassy, R. J., and Hostomska, Z. (1992) Heterologous expression and purification of active human phosphoribosyl-glycinamide formyltransferase as a single domain, *J. Protein Chem. 11*, 467−473.
20. Zhang, Y., Desharnais, J., Marsilje, T. H., Li, C., Hedrick, M. P., Gooljarsingh, L. T., Tavassoli, A., Benkovic, S. J., Olson, A. J., Boger, D. L., and Wilson, I. A. (2003) Rational design, synthesis, evaluation, and crystal structure of a potent inhibitor of human GAR Tfase: 10-(trifluoroacetyl)-5,10-didedideaza-acyclic-5,6,7,8-tetrahydrofolic acid, *Biochemistry 42*, 6043−6056.
21. Varney, M. D., Palmer, C. L., Romines, W. H., III., Boritzki, T., Margosiak, S. A., Almassy, R., Janson, C. A., Bartlett, C., Howland, E. J., and Ferre, R. (1997) Protein structure-based design, synthesis, and biological evaluation of 5-thia-2,6-diamino-4(3H)-oxopyrimidines: potent inhibitors of glycinamide ribonucleotide transformylase with potent cell growth inhibition, *J. Med. Chem. 40*, 2502−2524.
22. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res. 22*, 4673−4680.
23. Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J., and Huber, R. (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*, *Eur. J. Biochem. 230*, 788−796.
24. Otwinowski, Z. (1993) Data Collection and Processing, in *Proceedings of the CCP4 Study*, SERC Daresbury Laboratory, Warrington, U.K.
25. Bailey, S. (1994) The CCP4 Suite: Programs for Protein Crystallography, *Acta Crystallogr. D50*, 760−763.
26. Navaza, J. (1994) AMoRe: an Automated Package for Molecular Replacement, *Acta Crystallogr. A50*, 157−163.
27. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
28. Brünger, A. T. (1992) X-PLOR Version 3.8: A System for X-ray Crystallography and NMR, pp 1−382, Yale University Press, New Haven, CT.
29. Adams, P. D., Pannu, N. S., Read, R. J., and Brünger, A. T. (1999) Extending the limits of molecular replacement through combined simulated annealing and maximum-likelihood refinement, *Acta Crystallogr. D55*, 181−190.
30. Brünger, A. T., Krukowski, A., and Erickson, J. W. (1990) Slow-cooling protocols for crystallographic refinement by simulated annealing, *Acta Crystallogr. A46*, 585−593.
31. Kleywegt, G. J., and Jones, T. A. (1997) Detecting folding motifs and similarities in protein structures, *Methods Enzymol. 277*, 525−545.
32. Inglese, J., Smith, J. M., and Benkovic, S. J. (1990) Active-site mapping and site-specific mutagenesis of glycinamide ribonucleotide transformylase from *Escherichia coli*, *Biochemistry 29*, 6678−6687.
33. Shim, J. H., and Benkovic, S. J. (1999) Catalytic mechanism of *Escherichia coli* glycinamide ribonucleotide transformylase probed by site-directed mutagenesis and pH-dependent studies, *Biochemistry 38*, 10024−10031.
34. Krupenko, S. A., Wagner, C., and Cook, R. J. (1997) Domain structure of rat 10-formyltetrahydrofolate dehydrogenase. Resolution of the amino-terminal domain as 10-formyltetrahydrofolate hydrolase, *J. Biol. Chem. 272*, 10273−10278.
35. Schmitt, E., Blanquet, S., and Mechulam, Y. (1996) Structure of crystalline *Escherichia coli* methionyl-tRNA(f)Met formyltransferase: comparison with glycinamide ribonucleotide formyltransferase, *EMBO J. 15*, 4749−4758.
36. Morikis, D., Elcock, A. H., Jennings, P. A., and McCammon, J. A. (2001) Proton transfer dynamics of GART: the pH-dependent catalytic mechanism examined by electrostatic calculations, *Protein Sci. 10*, 2379−2392.

37. Nakamura, K., and Kozloff, L. M. (1978) Folate polyglutamates in T4D bacteriophage and T4D-infected *Escherichia coli*, *Biochim. Biophys. Acta 540*, 313−319.
38. Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis, *J. Mol. Model. 7*, 306−317.
39. Bashford, D., and Gerwert, K. (1992) Electrostatic calculations of the pK_a values of ionizable groups in bacteriorhodopsin, *J. Mol. Biol. 224*, 473−486.
40. DeLano, W. L. (2004) The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA (http://www.pymol.org).

BI050307G