Subcloning, Characterization, and Affinity Labeling of *Escherichia coli* Glycinamide Ribonucleotide Transformylase†

James Inglese,‡ Dana L. Johnson,‡ Alan Shiau,§ John M. Smith,§ and Stephen J. Benkovic*,‡
Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16801, and Seattle Biomedical Research Institute, Seattle, Washington 98109

Received August 4, 1989; Revised Manuscript Received October 9, 1989

ABSTRACT: Glycinamide ribonucleotide transformylase (GAR TFase; EC 2.1.2.2) has been purified 70-fold to apparent homogeneity from *Escherichia coli* harboring an expression vector encoding the *pur*N gene product, GAR TFase. The protein is a monomer of M_r 23 241 and catalyzes a single reaction. Steady-state kinetic parameters for the enzyme have been obtained. The structural requirements for cofactor utilization have been investigated and found to parallel those of the multifunctional avian enzyme. The enzyme was inactivated with the affinity label N^{10}-(bromoacetyl)-5,8-dideazafolate in a stoichiometric and active-site-specific manner. The ionization state of the cofactor analogue in the enzyme–cofactor complex appears to require the dissociation of the proton at N3 of the pyrimidine within the complex.

Glycinamide ribonucleotide transformylase (GAR TFase;[1] EC 2.1.2.2) catalyzes the first reduced folate requiring step in de novo purine biosynthesis (Scheme I). This reaction was first discovered and reported by the Buchanan group (Buchanan & Hartman, 1959, and references cited therein; Buchanan, 1982) and recently has become of mechanistic and pharmacological interest. Earlier work has focused on this activity in chicken liver (Rowe, 1984, and references cited therein), which was discovered to be part of a multifunctional protein (Daubner et al., 1985; Henikoff, 1986) containing the additional, GAR synthetase and AIR synthetase, activities. The studies presented in this paper concern our investigations with GAR TFase from *Escherichia coli*.

Unlike its avian counterpart, the bacterial transformylase is encoded by a relatively small gene of 636 base pairs which has been cloned and sequenced by Smith and co-workers (Smith & Daum, 1987) as part of a program to investigate the control of the purine biosynthetic pathway at the transcriptional level. These studies suggest that the gene encodes a single enzymatic activity. The present account describes the overexpression and purification, in quantity, of this protein combined with the initial characterization of relevant physical and chemical properties. In addition, we describe the active-site labeling of the enzyme using a dideazafolate affinity label and observe an apparent cofactor ionization that occurs in the enzyme–N^{10}-formyl-DDF complex but not in the enzyme–N^{10}-formyl-H_4folate complex. The enzyme's lack of structural complexity relative to the avian protein identifies this protein as a more attractive candidate for in-depth mechanistic and structural study.

Experimental Procedures

Materials

The reagents used to prepare 5,8-dideazafolate (DDF) and analogues were purchased from Aldrich Chemical Co., Milwaukee, WI. NBS was recrystallized from water, and triethylamine was distilled from CaH (Perrin et al., 1980). Glycinamide ribonucleotide (GAR) was prepared by the procedure of Chettur and Benkovic (1977). 8-Deazafolate was prepared by the method of Srinivasan and Broom (1981). 5-Deazafolate was kindly provided by Dr. E. C. Taylor, Department of Chemistry, Princeton University. 2,4-Diamino-5,8-dideazafolate was the gift of Dr. J. B. Hynes, Department of Pharmaceutical Sciences, Medical University of South Carolina. Carbocyclic GAR was the gift of Dr. R. Vince, Department of Medicinal Chemistry, University of Minnesota. Tris, Hepes, PMSF, chicken egg white lysozyme, Sephadex G-100 (fine), QAE-Sephadex A-25 (40–120 m), molecular weight markers for gel electrophoresis (Dalton VII), and bovine serum albumin (BSA) were obtained from Sigma Chemical Co., St. Louis, MO. All other reagents were of the highest possible quality and used as obtained unless otherwise specified. Spectrophotometric assays were performed on either a Gilford Model 252 or a Cary 219 spectrophotometer. UV–visible spectra were recorded on a Perkin-Elmer Lambda Array 3840 UV–vis spectrophotometer. Densitometry was carried out on an Ultroscan XL enhanced laser densitometer (LKB, Piscataway, NJ).

Methods

Protein Determination. Concentrations of protein solutions were determined by the Pierce BCA protein assay (Pierce Chemical Co.). An extinction coefficient of 22.1 mM^{-1} cm^{-1} (280 nm) was calculated by using bovine serum albumin as a standard.

5,8-Dideazafolate and N^{10}-Substituted Analogues. 5,8-Dideazafolate (DDF) was prepared according to the method

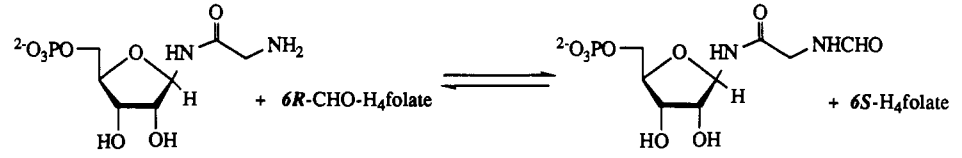
† This study was supported by grants from the National Institutes of Health of J.M.S. (AI20068) and to S.J.B. (GM24129). J.I. is the recipient of a Homer Braddock Research Fellowship. D.L.J. is the recipient of National Institutes of Health Postdoctoral Research Training Award GM12310-01.

* Author to whom correspondence should be addressed.
‡ The Pennsylvania State University.
§ Seattle Biomedical Research Institute.

[1] Abbreviations: GAR TFase, glycinamide ribonucleotide transformylase; GAR synthetase, glycinamide ribonucleotide synthetase; AIR synthetase, 5-aminoimidazole ribonucleotide synthethase; DHFR, dihydrofolate reductase; TS, thymidylate synthase; DDF, 5,8-dideazafolate; 5-DF, 5-deazafolate; 8-DF, 8-deazafolate; EDTA, ethylenediaminetetraacetic acid; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PMSF, phenylmethanesulfonyl fluoride; BSA, bovine serum albumin; FPLC, fast protein liquid chromatography; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; Mes, 4-morpholineethanesulfonate; BCA, bicinchonic acid; NBS, *N*-bromosuccinimide; carbocyclic β-GAR, carbocyclic glycinamide ribonucleotide; β-TGDDF, β-thioglycinamide dideazafolate.

0006-2960/90/0429-1436$02.50/0

Scheme I: Reaction Catalyzed by GAR TFase



of Achanya and Hynes (1975) with the modification of Calvert et al. (1980) at the 6-(bromomethyl)-4-hydroxy-2-(trimethylacetamido)quinazoline deacylation stage. Purification of DDF via anion-exchange chromatography employing DEAE-Sephadex was accomplished by using the procedure of Young et al. (1984). [*p*-(Benzylformylamino)benzoyl]-L-glutamate was prepared according to the method of Smith et al. (1981b). N^{10}-Formylation of analogues was carried out as described by Smith et al. (1981b).

N^{10}-Bromoacetylation of DDF, to generate the affinity label N^{10}-(bromoacetyl)-DDF, was accomplished by using the method of Daubner et al. (1986). The affinity label was purified by reversed-phase HPLC on a Perkin-Elmer analytical C-18 column. Elution conditions were 0–16% acetonitrile in a linear gradient over 70 min at a flow rate of 0.7 mL/min with both solvents containing 0.1% trifluoroacetic acid. The peak of interest had a retention time of approximately 55 min as determined at 310 nm. The sample was reduced to dryness on a Savant Speed-Vac. Approximately 1 μmol of material was applied to the column per injection (typically 100 μL of a 10 mM solution in 20 mM K_2HPO_4, pH 7.5, was injected). Solutions of the affinity label were prepared in either 20 mM K_2HPO_4, pH 7.5, or 20 mM Tris, pH 7.5. The concentration was determined by using an extinction coefficient of $\epsilon = 4.19$ mM^{-1} cm^{-1} at 310 nm (Daubner et al., 1986).

Electrophoresis. SDS–polyacrylamide gel electrophoresis was performed according to the method of Laemmli (1970) on slabs of 12% polyacrylamide. The standards for molecular weight calibration were bovine serum albumin, egg albumin, glyceraldehyde-3-phosphate dehydrogenase, carbonic anhydrase, trypsinogen, trypsin inhibitor, and α-lactalbumin. Samples were denatured by boiling for 10 min in a solution of 1% SDS, 1% β-mercaptoethanol, 50 mM Tris, pH 8.8, and 5% glycerol. Gels were stained for 15 min by immersion in a 1.2% Brilliant Coomassie Blue, 5:5:1 (methanol/water/glacial acetic acid).

Native Molecular Weight Determination. The native molecular weight of *E. coli* GAR TFase was determined by equilibrium sedimentation ultracentrifugation (Schachman, 1957; Chervenka, 1969). Centrifugation was performed on 3–200-μL enzyme samples (OD_{280nm} = 0.25–0.30) in the following buffers: 50 mM Tris, pH 7.5, 0.5 mM EDTA, and 50 mM K_2HPO_4, pH 7.5 (with separate determinations with and without 1 mM DTT). The samples were prepared by exhaustive dialysis against the respective buffer with 210 μL of the appropriate dialysis buffer used as a reference for each sample. Samples were centrifuged in a Model E Beckman analytical ultracentrifuge at 20000 rpm (An-F Ti rotor) for 72 h at 20 °C. After equilibrium conditions were obtained, the sample was centrifuged at 40000 rpm to clear protein from the meniscus to obtain a base-line value. Data from the equilibrium distribution curve were plotted as ln c vs r^2, where c is protein concentration and r is the distance from the center of the rotor, and the resulting slope was used to determine the molecular weight from

$$M_r = \frac{2RT}{(1 - \nu\rho)\omega^2} \frac{d(\ln c)}{d(r^2)} \quad (1)$$

where M_r is the molecular weight, R is the gas constant, T is the temperature of system (20 °C), ν is the partial specific volume [calculated from the amino acid sequence to be 0.7370 using the values of McMeekin and Marshall (1952)], ρ is the density of the medium, and ω^2 is the angular velocity of the rotor (4.387×10^6).

N-Terminal Sequence Analysis. Automated Edman degradation was carried out on an Applied Biosystems Model 477A pulsed-liquid sequenator. The resulting PTH derivatives were analyzed by using an on-line Applied Biosystems Model 120A microbore HPLC. Analysis was carried out on 1–2 nmol of protein.

Enzyme Assays. E. coli GAR TFase was assayed by the method of Young et al. (1984) utilizing N^{10}-formyl-DDF as the formyl donor. The assay was carried out in a volume of 1 mL containing 50 mM Tris, pH 7.5 or 8.5 (the rate being maximal at pH 8.5), 0.5 mM EDTA, 45 μM N^{10}-formyl-DDF, and 35 μM β-GAR (α,β mixture) at 26 °C. The assay can be initiated with cofactor, GAR, or enzyme. Initiation with enzyme, however, is followed by a slight lag phase. The course of the reaction was monitored at 295 nm, $\Delta\epsilon = 18.9$ mM^{-1} cm^{-1} (Smith et al., 1981b).

Strains and Media. Strain TX635 [F′ lacZ^+ cI857 (Mieschendahl & Muller-Hill, 1985)] contains an episome-borne temperature-sensitive λ repressor and was used as a host for the λ pL plasmids. Strains were made competent and transformed by the procedure of Dagert and Ehrlich (1979). The minimal medium of Neidhardt et al. (1974) and the rich media described by Miller (1972) were used for the growth of the *E. coli* K12 strains. The recombinant DNA techniques employed have previously been described (Tiedeman et al., 1985).

Construction of AIR Synthase and GAR TFase Expression Vector: pJS119. Plasmid pJS119 (Figure 1A) was constructed by two successive subclonings of restriction fragments that covered the nucleotide sequence 732–2746 [the numbering scheme refers to the published sequence (Smith & Daum, 1986 and 1987)] which removes the *pur*MN promoter and *pur*R binding site. The first restriction fragment subcloned was a 186-bp *Hin*P1 fragment (nucleotide 732–919) that had been treated with T4 DNA polymerase to create blunt ends and cloned into the *Sma*I site of M13mp18 (Yanisch-Perron et al., 1985). After DNA sequencing to verify fragment identity and determine the orientation, the restriction fragment (in the correct orientation to maintain *pur*MN expression from the lac promoter) was transferred to plasmid pUC18 (Yanisch-Perron et al., 1985) by restriction digest and ligation to form plasmid pJS117. The remainder of the *pur*MN operon was added as a *Ppu*MI–*Xho*II restriction fragment into the *Ppu*MI–*Bam*HI sites of plasmid pJS117 to form plasmid pJS118. An *Eco*RI–*Sal*I restriction digest was then used to transfer the promoterless *pur*MN operon into the λ pL expression vector, plasmid pJS88 (A. A. Tiedeman and J. M. Smith, unpublished data) to form plasmid pJS119, followed by transformation of strain TX635.

Construction of a GAR TFase Expression Vector: pJS167. This plasmid was created by a series of manipulations equivalent to the deletion of the *pur*M coding region for AIR

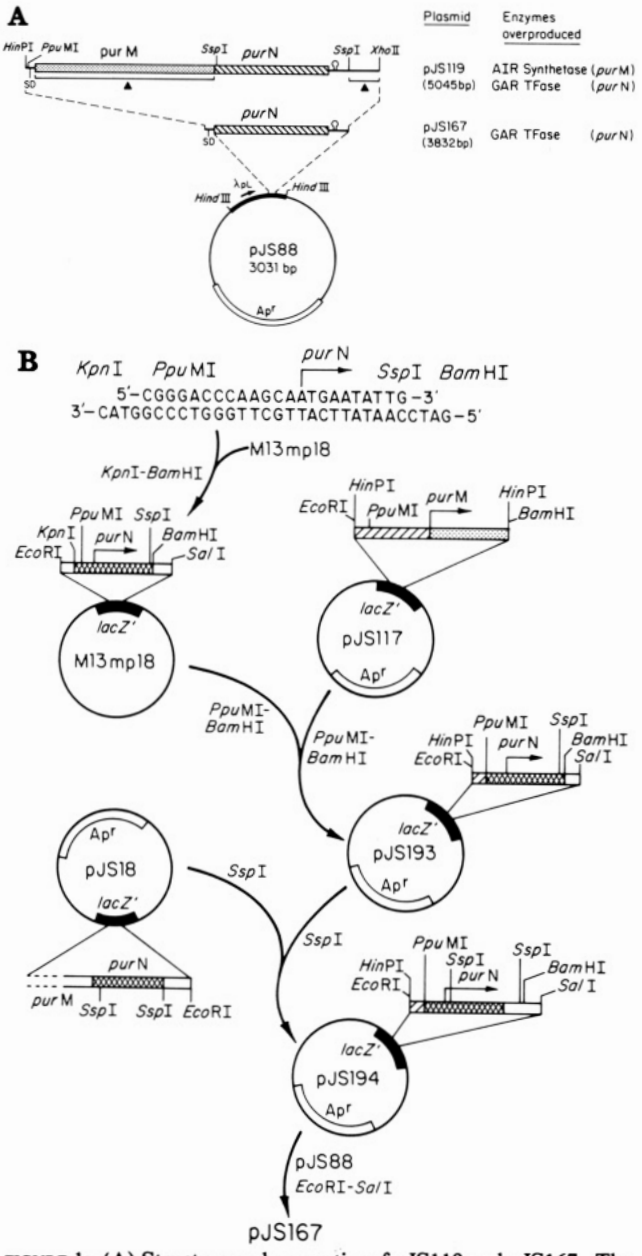


FIGURE 1: (A) Structure and properties of pJS119 and pJS167. The relationship of the two plasmids and the enzymes overproduced are indicated. The indicated modifications (▲, deletions) to plasmid pJS119 were made in vitro to create plasmid pJS167. Flanking *Eco*RI and *Sal*I restriction sites in the intermediate construction plasmids were used to move the indicated restriction fragments into the λ pL expression vector pJS88. (B) Construction of the GAR TFase expression vector. The expression vector pJS167 was constructed as indicated. The source of the *pur*N structural gene was plasmid pJS18 (Smith & Daum, 1986). During the intermediate steps in the construction of the expression vector, clonings were carried out in such a way as to maintain expression via the lac promoter on the host plasmids. All of the clonings utilized purified DNA fragments. Plasmid pJS194 was isolated by its ability to complement a strain lacking GAR TFase activity, strain TX680 (Smith & Daum, 1987).

synthetase activity (Figure 1A). Two complementary (23- and 31-mer) oligonucleotides were synthesized (Figure 1B) with the 23-mer oligonucleotide containing the altered *pur*N initiation sequence and the complementary 31-mer additionally incorporating *Kpn*I and *Bam*HI linkers for cloning. The length of the linkers was adjusted to include translational stop codons that ensured disruption of the lacZ reading frame when the fragment was cloned into the M13mp18 cloning region. The oligonucleotides were synthesized on a Milligen/Biosearch Cyclone DNA synthesizer using a 0.2-μmol scale and purified by the OPC cartridge method of Applied Biosystems. The two oligonucleotides were annealed and ligated into the *Kpn*I/*Bam*HI sites of M13mp18. After transformation of strain JM101, colorless plaques were picked at random and sequenced to verify the structure of the insert.

The synthesized DNA fragment containing the *pur*N translational initiation site was then transferred to the plasmid pJS117 via the *Ppu*MI and *Bam*HI sites to form plasmid pJS193. The former contains a 187-bp sequence (nucleotides 732–919) (Smith & Daum, 1986, 1987) that is part of the *pur*M/*pur*N operon. The remainder of the *pur*N structural gene was added as a 747-bp *Ssp*I DNA fragment derived from pJS18. Because of the presence of another *Ssp*I site elsewhere in plasmid pJS193, DNA from a partial *Ssp*I digest of pJS193 was ligated with the 747-bp *Ssp*I DNA fragment containing the *pur*N structural gene. The ligation mixture was transformed into strain TX680 with selection for purine prototrophy. Following characterization, a representative plasmid was designated pJS194. The resulting *pur*N structural gene was transferred to the expression plasmid pJS88 via the *Eco*RI and *Sal*I restriction sites to create pJS167.

E. coli GAR TFase Purification. *E. coli* strain TX393 containing the high copy number plasmid pJS18 with a DNA insert encoding GAR TFase (Smith & Daum, 1987) was grown in M9CA media (Maniatis et al., 1982) supplemented with 30 mg/L ampicillin. Growth was initiated with a 1% culture inoculum and maintained at 37 °C. The cells were harvested in the late log phase by centrifugation to yield typically 2.5 g/L.

E. coli strain TX635 containing either plasmid pJS119 or pJS167 was grown in the rich media described (vide supra) containing 30 mg/L ampicillin. The cells were grown at 30 °C to maximum density and protein overexpression induced by raising the temperature to 42 °C for up to 9 h to obtain maximum protein production. Cells were harvested by centrifugation with yields of 5–6 g/L.

Unless otherwise designated all purification steps were performed at 4 °C in a buffer (buffer A) which contained 50 mM Tris, pH 7.5, and 1 mM EDTA in addition to other specified components. The cells (22.7 g) were resuspended in 125 mL of buffer A that contained 10 mg of PMSF (dissolved in 50 μL of DMF). The cells were disrupted by adding 72 mg of egg white lysozyme in 2 mL of buffer and 5.3 mL of Triton X-100/glycerol (25 mL of glycerol/1 mL of 10% Triton X-100). The suspension was mixed and allowed to stand at 4 °C for 40 min. The lysed cells were passed through a syringe (17-gauge needle) five times to shear DNA. The lysate was cleared by centrifugation at (16000 rpm) 30900*g* for 20 min and the supernatant recovered.

To this supernatant (110 mL) was added 600 mg of streptomycin sulfate in 5 mL of buffer A via a syringe driver over 10 h with gentle stirring. The resulting milky white suspension was centrifuged at (16000 rpm) 30900*g* for 20 min. The supernatant was dialyzed ($1\text{-}^1/_8$ in., 12000 M_r cutoff tubing) against 2 × 2.5 L of buffer A with a final recovered volume of 130 mL.

This protein solution was diluted to 300 mL with buffer A and applied (2–3 mL/min) to a column of QAE-A25 Sephadex (5 × 25 cm) previously equilibrated with buffer A. The column was washed with buffer A until the absorbance (280 nm) at the column outlet was less than 0.1 (~1 L) and developed with a 3-L linear gradient of KCl (0.05–0.5 M KCl in buffer A). Fractions (14 mL) from the QAE-Sephadex column that contained GAR TFase activity, which eluted at approximately 250 mM KCl, were pooled (600 mL) and

Table I: Purification of *E. coli* GAR TFase

step	volume (mL)	total protein (mg)	sp act.[a] (units/mg)[b]	total act. (units)	purifn (x-fold)	purifn yield (%)
cell-free extract	110	2169	0.16	347	1	100
streptomycin sulfate	130	2145	0.175	375	1.1	108
QAE-A25	24	441	0.73	322	4.6	93
G-100	25	43	6.92	297	43	86
FPLC (Mono Q)	12	22.8	11.0	251	69	72
G-100	5.5	20.5	11.6	238	72.5	69

[a] Assayed as described under Methods by utilizing 20 μM N^{10}-CHO-DDF and 10 μM β-GAR (α,β mixture). [b] Unit: μmol min^{-1} mL^{-1}.

concentrated to 24 mL by using an Amicon ultrafiltration apparatus with a YM10 membrane.

Half of this concentrated protein solution was applied to a column (2.5 × 55 cm) of Sephadex G-100 equilibrated with buffer A at a flow rate of 1 mL/min. Fractions (5 mL) were collected, and those containing GAR TFase activity were pooled (60 mL) and concentrated (12 mL). This step was repeated for the remaining 12 mL of protein concentrate.

The resulting >90% pure protein obtained was further purified at room temperature (25 °C) on a Mono Q HR5/5 column (Pharmacia, FPLC) using a linear gradient of KCl (0–500 mM in buffer A) at a rate of 1% 1 M KCl/min, loading about 5 mg of total protein per injection. The elution was monitored at 280 nm by using an in-line detector (Pharmacia). Active fractions from the center of the major peak were pooled, concentrated to 12 mL, and passed over the G-100 column described above. Active fractions were pooled, on the basis of SDS–PAGE assessed purity, and concentrated. Aliquots were frozen in liquid nitrogen and stored at −80 °C as 2–4 mg/mL solutions. Table I summarizes the purification for strain TX635 containing plasmid pJS167.

Steady-State Kinetics. Initial velocity studies were carried out by using a modification of the procedure of Young et al. (1984). Data were analyzed by using the computer programs of Cleland (1979). Studies were performed in 50 mM Hepes and 0.5 mM EDTA at 26 °C. Values were obtained at pH 7.5 and 8.5. The concentration of enzyme used per assay was between 1 and 5 nM (stock solution was 1 μM enzyme in 50 mM Hepes, 20% glycerol, and 1 mM EDTA, pH 7.5).

Reactions utilizing N^{10}-formyl-DDF as substrate were assayed in 5–200 μM β-GAR (stock solutions were 2–5 mM β-GAR (α,β mixture) in 25 mM K_2HPO_4, pH 6.8) and 1–100 μM N^{10}-formyl-DDF (stock solutions were 2–4 mM N^{10}-formyl-DDF in 25 mM Hepes, pH 7.5). The concentration of N^{10}-formyl-DDF was determined from the absorbance at 310 nm (ϵ = 3.8 mM^{-1} cm^{-1}) and its conversion to DDF in the presence of enzyme ($\Delta\epsilon$ = 18.9 mM^{-1} cm^{-1}) (Smith et al., 1981b). The concentration of β-GAR was determined enzymatically by measuring the disappearance of N^{10}-formyl-DDF with limiting GAR.

Reactions utilizing (6*R*)-N^{10}-formyl-H_4folate as substrate were assayed by using 5–100 μM β-GAR (α,β mixture) and 10–100 μM (6*R*)-N^{10}-formyl-H_4folate. The substrate for this study was prepared from (6*R*)-5,10-methenyl-H_4folate as described by Smith et al. (1981a). The spectrophotometric assay was the same as above except the reaction was monitored at 298 nm in argon-saturated buffers containing 10 mM β-mercaptoethanol. The concentration of (6*R*)-N^{10}-formyl-H_4folate was determined at pH 7.5 from the extinction coefficient ϵ(298 nm) = 9.54 mM^{-1} cm^{-1} and its conversion to (6*S*)-H_4folate in the presence of enzyme using $\Delta\epsilon$(298) = 19.7 mM^{-1} cm^{-1} (Black et al., 1978).

Inactivation Kinetics. Inactivation kinetics with N^{10}-(bromoacetyl)-DDF were carried out as described by Daubner et al. (1986) with some modifications. Inactivations were performed in 1.0 mL of 50 mM Hepes, pH 7.5, at 26 °C at an enzyme concentration of 500 nM. A zero time enzyme activity was determined prior to addition of inactivator. After addition of N^{10}-(bromoacetyl)-DDF, 10-μL aliquots of the incubation mixture were withdrawn at specified times and assayed for enzymatic activity.

Protection experiments were carried out as above except the incubation mixture contained the appropriate additional ligands. Enzyme–ligand mixtures were incubated for 5 min before addition of inactivator.

The stoichiometry of inactivation was performed by adding subequivalent aliquots of N^{10}-(bromoacetyl)-DDF (microliter volumes of a 3.11 mM solution of the inactivator in 20 mM K_2HPO_4, pH 7.5) to 1 mL of GAR TFase (0.067 μmol). An equal volume of 20 mM K_2HPO_4, pH 7.5, was added to a control sample. Both protein samples were incubated at 37 °C for 3 h. A 1-μL aliquot, one from each sample, was diluted 100-fold into 50 mM Hepes and 20% glycerol, and 10-μL aliquots of this dilution were assayed for activity. An additional subequivalent of inactivator and phosphate buffer were added to reaction and control, respectively, and the above procedure was repeated until no activity remained.

UV Difference Spectroscopy. The enzyme (10 μM) and the ligand (20 μM) were contained within separate compartments of a quartz split cuvette with a total path length of 1.0 cm (0.5 cm per compartment). The cuvette containing the unmixed components was used to blank the instrument, and following complete mixing of the contents of the two compartments by inversion, the difference spectrum was obtained after 30 s. Experiments with enzyme–ligand mixtures were performed in MTEN buffer containing 0.1 M NaCl similar to the method of Stone and Morrison (1983). The pH-dependent UV difference spectra were obtained as described above at a constant ionic strength of 0.15 in the ligand-containing (20 μM ligand) compartment with 1 M NaOH in the second compartment. The spectra were recorded 30 s after mixing.

Results and Discussion

Subcloning and Purification. GAR TFase from *E. coli* has been purified to homogeneity for the first time. The enzyme was first overproduced by using the high copy plasmid pJS18 (Smith & Daum, 1987) that contained the genes for both AIR synthetase and GAR TFase. This clone produces high levels of AIR synthetase (Schrimsher et al., 1986) but significantly lower levels of GAR TFase. In this paper we describe the construction of a second expression vector containing only the gene for GAR TFase.

The initial expression vector constructed consisted of a promoterless *pur*MN operon cloned into the λ pL expression vector, pJS88 (A. A. Tiedeman and J. M. Smith, unpublished data) to create plasmid pJS119 (Figure 1A). Plasmid pJS88 is a λ pL expression vector similar to the ones described by Remault et al. (1981). Upon characterization, plasmid pJS119 was found to overproduce both AIR synthetase (*pur*M) and GAR TFase (*pur*N), but the overexpression of GAR TFase

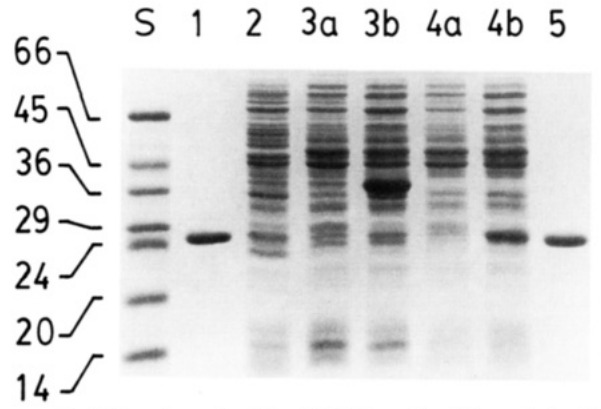
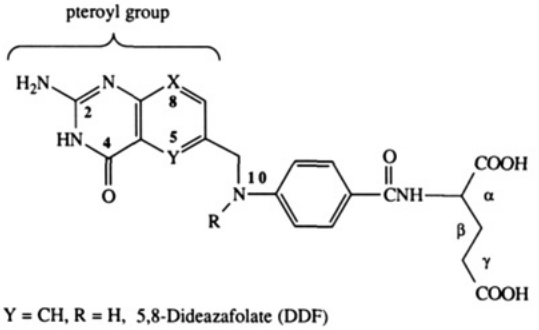


FIGURE 2: SDS–polyacrylamide gel (12% acrylamide) analysis of *E. coli* GAR TFase and clone crude lysates. (Lane S) Molecular weight markers; (lanes 1 and 5) 3 μg of purified *E. coli* GAR TFase; (lane 2) crude lysate from pJS18 grown at 37 °C for 10 h; (lane 3) crude lysate from pJS119 (a) grown at 30 °C for 10 h and (b) after induction at 42 °C for 8 h (the dark band of molecular weight 38 500 is AIR synthetase); (lane 4) crude lysate from pJS167 (a) grown at 30 °C for 10 h and (b) after induction at 42 °C for 8 h. Lanes 2–4 were loaded with protein from equal quantities of cells on the basis of OD 660 nm.



X = Y = CH, R = H, 5,8-Dideazafolate (DDF)

R = CHO, N-10-formyl-DDF
R = $COCH_2Br$, N-10-(bromoacetyl)-DDF
R = CHO, 4 oxo —4 amino, 2,4-Diamino-N-10-formyl-DDF
R = CHO, pteroyl group —aryl group, [p-(Benzylformylamino)benzoyl]-L-glutamate

X = CH, Y = N, R = H, 8-Deazafolate (8-DF); R = CHO, N-10-formyl-8-DF

X = N, Y = CH, R = H, 5-Deazafolate (5-DF); R = CHO, N-10-formyl-5-DF

X = Y = N, R = H, Folate

R = CHO, N-10-formyl-folate
R = CH_3, 4 oxo —4 amino, Methotrexate (MTX)

FIGURE 3: Chemical structures of cofactor analogues.

was not coordinate with AIR synthetase (Figure 2, lanes 3a and 3b). This may be due to the translational coupling between the *pur*M and *pur*N genes (Smith & Daum, 1987), reflecting a lower efficiency of ribosomal reinitiation for the translation of GAR TFase.

Because of the noncoordinate expression of AIR synthetase and GAR TFase in plasmid pJS119, an expression vector designed to maximize the overproduction of GAR TFase was constructed. In a series of manipulations functionally equivalent to the deletion of the AIR synthetase (*pur*M) coding region, plasmid pJS167 was created (Figure 1A). This expression system produces approximately 27-fold the amount of GAR TFase per cell as pJS18 on the basis of measured activity (Figure 2, lane 4b vs lane 2).

This investigation is the first account of studies directed toward this enzyme species since the demonstration by Dev and Harvey (1978) that *E. coli* GAR TFase uses N^{10}-formyl-H_4folate as the formyl donor source. As with GAR TFase from avian (Smith et al., 1981b), HeLa O, murine leukemia L1210 (Daubner & Benkovic, 1985; Daubner et al., 1986), murine lymphoma L5178Y (Caperelli, 1985), Chinese hamster ovary, and human (Hards et al., 1986) sources, the *E. coli* enzyme can also utilize the DDF analogue of N^{10}-formyl-H_4folate as a cofactor (Figure 3). Our work has, for the most part, exploited the air stability of the quinazoline analogue and the relative ease of its synthetic manipulation for the construction of the affinity label N^{10}-(bromoacetyl)-DDF (Daubner et al., 1986).

The enzyme was initially purified from the pJS18 clone, the original *pur*MN clone previously described (Smith & Daum, 1987), by utilizing a combination of conventional liquid chromatography techniques that allowed the purification of milligram quantities of enzyme. The same purification was used to obtain GAR TFase from cells transformed with pJS167. The final purification scheme is outlined for pJS167 in Table I. The Sepharose–N^{10}-(formyl)-DDF affinity column employed by Young et al. (1984) for the purification of chicken liver GAR TFase was not used due to the large quantity of recoverable enzyme and the weaker binding of the bacterial enzyme to the affinity resin (the K_m of N^{10}-formyl-DDF is approximately 10-fold greater for *E. coli* than for chicken GAR TFase at pH 7.5).

Physical Properties. Unlike the avian trifunctional GAR TFase, the *E. coli* species has a sole function. The genetic evidence of Smith and Daum (1987) from *E. coli*, as well as the work of Schrimsher et al. (1986), who purified and characterized *E. coli* AIR synthetase, suggests that the relevant activities, in *E. coli*, are found as monofunctional enzymes, whereas in chicken they belong to a trifunctional protein containing the activities GAR synthetase, AIR synthetase, and GAR TFase. Figure 4 illustrates these points for several species and depicts the monofunctional nature of the bacterial

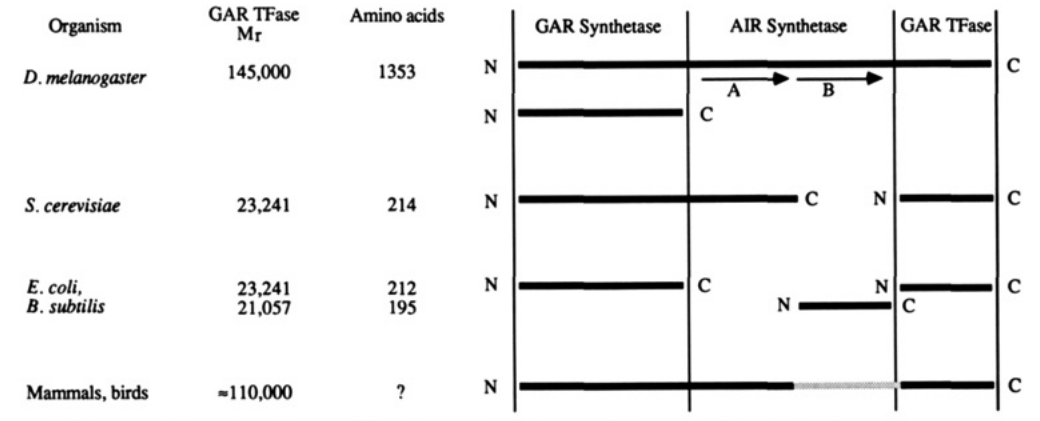


FIGURE 4: Structural characteristics of species-specific GAR TFase's [adapted from Henikoff (1987)].

Table II: Kinetic Constants of *E. coli* GAR TFase Substrates and Inhibitors

substrate or inhibitor	pH	k_{cat} (s^{-1})	K_m (μM)	K_i (μM)	k_{cat}/K_m
N^{10}-formyl-DDF	7.5	16.1 ± 1.3	16.7 ± 3		0.96
	8.5	40 ± 8	36.5 ± 6.4		1.1
(6*R*)-N^{10}-formyl-H_4folate	7.5	13.5 ± 0.8	84.8 ± 5.9		0.16
	8.5	20.7 ± 0.5	77.5 ± 3.9		0.27
β-GAR[a]	7.5	-	19.2 ± 3.7		-
	8.5	-	23.5 ± 4.6		-
β-GAR[b]	7.5	-	12.2 ± 1.2		-
	8.5	-	8.1 ● 0.8		-
carbocyclic β-GAR[c]	8.5	33 ± 2	46.5 ± 4.5		-
N^{10}-formylfolate[d]	8.5	0.016 ± 0.012	600 ± 400		0.00002
N^{10}-formyl-5-DF[d]	8.5	1.52 ± 0.11	190 ± 20		0.008
N^{10}-formyl-8-DF[d]	8.5	0.069 ± 0.013	150 ± 40		0.0005
(6*S*)-N^{10}-formyl-H_4folate[e]	8.5			12 ± 1	
DDF	8.5			28 ± 3	
	7.5			18 ± 2	
N^{10}-formyl-2,4-diamino-DDF[d,f]	8.5			96 ± 10	
MTX[d]	8.5			>100	
[*p*-(benzylformylamino)benzoyl]-L-glutamate[d]	8.5			>200	

[a] The constants for β-GAR were determined by using the DDF cofactor or [b] H_4folate. [c] These are apparent values obtained with the DDF cofactor at a concentration of 100 μM. [d] The constants obtained for these compounds are apparent values where [β-GAR] = 200 μM. A mixture of α,β-GAR was used; however, all experiments were conducted by using the concentration determined for the β anomer. [e] Determined by using a mixture of (6*RS*)-formyl-H_4folate; inhibition constant determined from the equation $K_i = [I]_t/(1 - v_i/v_0)/(1 + [S]/K_m)v_0/v_i$ (Segel, 1975). [f] This compound is a noncompetitive inhibitor [determined by using the method as described by Segel (1975)].

and yeast GAR TFase. In several higher organisms GAR TFase constitutes the C-terminal domain of a trifunctional polypeptide (Henikoff, 1987).

Evidence suggests that *E. coli* GAR TFase is a small, single-subunit protein with a molecular weight of 23 241. This weight was calculated from the peptide sequence deduced from the cDNA (Smith & Daum, 1987). The molecular weight of the purified protein obtained from SDS–PAGE is approximately 25 000 (Figure 2). Ultracentrifugation sedimentation velocity experiments performed on the protein under reducing and nonreducing conditions gave an average molecular weight of 24 000, indicating that under the concentrations and conditions studied *E. coli* GAR TFase behaves as a monomer in solution.

The N-terminal sequence of GAR TFase has been determined by automated Edman degradation and agrees with the sequence predicted from the gene (Smith & Daum, 1987):

M N I V V L ... (gene)
M N I V V L ... (protein)

The N terminus begins with methionine that has probably undergone posttranslational processing (fMet → Met) because the yield on the first cycle of the sequencer was very high (i.e., a protein beginning with fMet would have sequenced poorly or not at all).

Kinetic Studies. Initial velocity studies preformed with N^{10}-formyl-DDF gave a series of intersecting lines when the data were analyzed by using Cleland's programs (Cleland, 1979), suggesting a sequential mechanism. This study revealed the kinetic parameters listed in Table II. The pH–rate profile has a bell shape ($pK_1 = 7.6$, $pK_2 = 9.4$) with an optimum rate at pH 8.5, a full pH unit above that reported for the chicken enzyme (Young, 1985).

A complete initial velocity study was also performed with the natural cofactor (6*R*)-N^{10}-formyl-H_4folate, indicating a sequential-type mechanism. The kinetic constants are reported in Table II. Again with the natural cofactor the pH optimum appears to be 8.5. This is consistent with results reported by Dev and Harvey (1978) for their crude *E. coli* GAR TFase preparation when the extract was assayed with the natural cofactor. In comparison to the avian enzyme, both the natural and analogue cofactors bind about 10-fold less tightly to the bacterial enzyme as indicated by their higher K_m values. A noteworthy finding is the fact that a racemic mixture of (6*R,S*)-N^{10}-formyl-H_4folate can be used to assay the bacterial enzyme with only the 6*R* isomer being turned over. The $K_{i,app}$ of the 6*S* isomer is approximately 6-fold lower than the K_m of the 6*R* isomer. Smith et al. (1981b) observed a similar result for chicken liver GAR TFase with the K_i of the 6*S* isomer 8-fold smaller than the K_m of the cofactor.

A recent study of the steady-state kinetics of a murine lymphoma GAR TFase has demonstrated an ordered sequential mechanism, with N^{10}-formyl-DDF binding first (Caperelli, 1989). Current experimental efforts are directed toward the definition of the binding mechanism for *E. coli* GAR TFase in this laboratory and should reveal if this type of kinetic scheme is also employed by the bacterial species.

Inactivation Kinetics. Consistent with the chicken and mammalian GAR TFase (Daubner et al., 1986), the *E. coli* enzyme is inactivated by the cofactor analogue affinity label, N^{10}-(bromoacetyl)-DDF, although binding (K_i) is weaker and the inactivation rate (k_{inact}) slower than with the other species of GAR TFase. The enzyme was inactivated in a time- and concentration-dependent fashion following the mechanism

$$E + I \overset{K_i}{\rightleftharpoons} E{\cdot}I \xrightarrow{k_{inact}} E{\sim}I$$

where E∼I is the inactivated enzyme. The process is described by

$$v = \frac{k_{inact}[I]_0}{K_I + [I]_0}([E]_0 - [E{\sim}I]) \qquad (2)$$

The values for K_i and k_{inact} (±10%) were 23.4 μM and 0.027 min^{-1}, respectively. DDF (K_i = 18 μM) was shown to protect the enzyme from inactivation and suggests an active-site-directed inactivation mechanism. Because the enzyme is not inactivated by iodoacetamide at millimolar concentrations, even though the reactive portion of the affinity label is an α-haloacetamide, implies that the DDF structure is crucial for binding and positioning of the α-haloacetamide group near an active-site nucleophile.

Inactivation Stoichiometry. The stoichiometry of inactivation was determined by an inactivation titration of affinity label vs enzyme activity as shown in Figure 5. The stoichiometry from the titration indicates approximately one inactivator binds per enzyme monomer. This situation contrasts with that of the avian enzyme, where up to three labels per

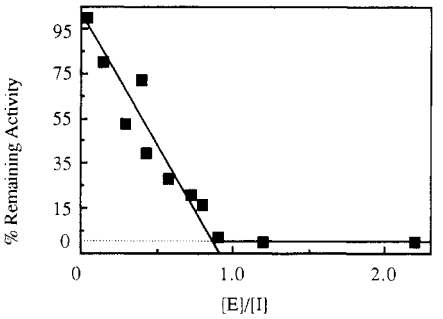
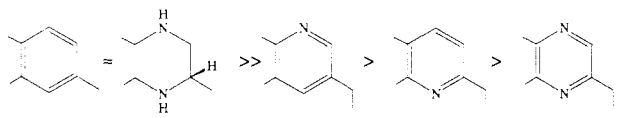


FIGURE 5: Inactivation titration of *E. coli* GAR TFase with N^{10}-(bromoacetyl)-DDF. (*X* axis) percent remaining GAR TFase activity; (*Y* axis) equivalents of inactivator. Experiment as described under Methods.

Scheme II: Cofactor Specificity



$\frac{k_{cat}}{K_m}$ 1.1 0.27 8×10^{-3} 5×10^{-4} 2×10^{-5}

enzyme monomer were observed (Daubner et al., 1986; Young, 1985) and should allow the determination of the active-site region of *E. coli* GAR TFase by peptide mapping techniques.

Cofactor and GAR Analogues. (*a*) *Pyrazine Ring Oxidation States and Deaza Analogues.* Table II compiles the results of studies with folate analogues whose structures are given in Figure 3. The effects of oxidation state and alterations in the oxidized pyrazine ring system of folate have been examined previously by Benkovic and co-workers (Smith et al., 1981b) on the transformylase enzymes of chicken and parallel the results reported here for *E. coli* GAR TFase. The order of decreasing efficiency (k_{cat}/K_m) of the pyrazine ring analogues of N^{10}-formylfolate as cofactors for the *E. coli* enzyme is as shown in Scheme II. As with chicken, the *E. coli* enzyme utilizes the dideaza cofactor approximately 3-fold more efficiently [based on (k_{cat}/K_m)] than (6*R*)-N^{10}-formyl-H_4folate. The two deazafolates, 5-DF and 8-DF, are utilized 170- and 3000-fold less efficiently than DDF with N^{10}-formylfolate 70000 times less. A noteworthy related observation was made by Brixner et al. (1987) in regard to the inhibitory properties of various aza derivatives of N^{10}-propargyl-DDF (PDDF), a potent inhibitor of thymidylate synthase (TS). PDDF was observed to bind 2 orders of magnitude more tightly than the pyridopyrimidine (8-DF) analogue, which was a slightly better inhibitor (2,3-fold) than N^{10}-propargylfolate. This similar trend in the binding of the pyrazine deaza derivatives of folate in GAR TFase and TS may suggest potential similarities in the folate binding sites of the two enzymes.

(*b*) *Pyrimidine Ring Specificity.* As with avian GAR TFase the DDF analogue lacking the pyrimidine moiety [[*p*-(benzylformylamino)benzoyl]-L-glutamate; Smith et al., 1981b] had no inhibitory effect on *E. coli* GAR TFase up to 200 μM. Further, the 4-amino congener of N^{10}-formyl-DDF, 2,4-diamino-N^{10}-formyl-DDF, is not a cofactor for the transformylation reaction but binds with a K_i of 96 μM, acting as a noncompetitive inhibitor. Methotrexate (MTX), a potent 2,4-diaminofolate inhibitor of DHFR, has no measurable effect on the activity of GAR TFase up to 100 μM. MTX contains a fully oxidized pyrazine ring (shown to have low affinity for the tetrahydrofolate binding site) and lacks a 4-oxo group on the pyrimidine moiety (apparently important in binding and/or catalysis). These results indicate certain functional requirements of the pyrimidine group for the effective binding of the cofactor to GAR TFase.

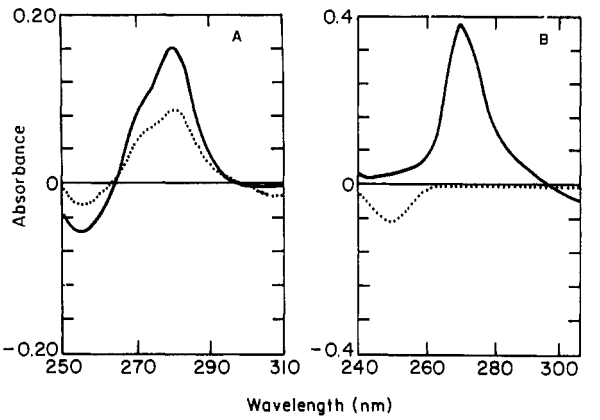


FIGURE 6: UV difference spectra of E–cofactor complexes and pH-dependent cofactor UV difference spectra. (A) (···) E-N^{10}-formyl-DDF difference spectrum; (—) pH-dependent UV difference spectrum of N^{10}-formyl-DDF. (B) (···) E-(6*R*)-N^{10}-formyl-H_4folate difference spectrum; (—) pH-dependent UV difference spectrum of (6*R*)-N^{10}-formyl-H_4folate. Both E–ligand complexes (···) contained 10 μM E and 20 μM ligand, and the difference spectra of ligand alone (—) contain 20 μM ligand.

(*c*) *GAR Analogues.* In accordance with previous observations (Caperelli & Price, 1988) a β-GAR analogue in which the C1′-C4′ bridging oxygen was replaced with a methylene group was found to be a substrate for the *E. coli* enzyme. The K_m of this carbocyclic GAR analogue was approximately 2-fold higher than the K_m of β-GAR, with the k_{cat} being roughly the same.

(*d*) *UV Difference Study.* As a further analysis of why N^{10}-formyl-DDF may efficiently serve as a cofactor for GAR TFase, UV difference spectra were obtained for enzyme–cofactor complexes. Analysis of these spectra (Figure 6A) indicated the enzyme-dependent formation of an absorbance peak at ca. 280 nm when N^{10}-formyl-DDF is the ligand. This spectrum is strikingly similar to the pH-dependent deprotonation UV difference spectra of N^{10}-formyl-DDF (the proton undergoing ionization in this case is the N3 lactam proton). A similar absorbance at ca. 280 nm that can be observed in the pH-dependent deprotonation spectrum of (6*R*)-N^{10}-formyl-H_4folate, however, is not observed in enzyme–ligand complexes (Figure 6B). These experiments corroborate initial observations indicating the potential ionization of the DDF cofactor when bound covalently to GAR TFase (Inglese, 1989) and may be important in understanding the mechanism by which GAR TFase can utilize the dideaza cofactor as a substrate. Additional information must be obtained, however, before we can fully rationalize these results in terms of a mechanism.

Recently we have obtained several different types of crystals of *E. coli* GAR TFase including heavy atom derivatives suitable for X-ray diffraction analysis (Stura et al., 1989). Through the use of cocrystals with the potent specific multisubstrate inhibitor of GAR TFase, β-TGDDF (Inglese et al., 1989), essential binding site residues and geometries should be revealed, whereas catalytic active site residues are currently being determined by using the N^{10}-(bromoacetyl)-DDF affinity label and peptide mapping techniques.

Although there are some differences between the bacterial and avian enzymes, for example, molecular size, pH optimum, and reduced binding affinity for the natural cofactor or analogues, the similarities are more important. These include a sequential kinetic mechanism, identical selectivity toward utilization of natural and cofactor analogues, sensitivity to

N^{10}-(bromoacetyl)-DDF, and potent inhibition by β-TGDDF. Such common features allow the *E. coli* enzyme to be used as a model for the more complex and pharmacologically important mammalian transformylase in terms of mechanism, structure, and drug design.

ACKNOWLEDGMENTS

We appreciate the collaboration of L. Bloom during the development of the enzyme purification and the guidance of P. Jennings during the analytical ultracentrifugation experiments.

Registry No. GAR TFase, 9032-02-4; β-GAR, 10074-18-7; MTX, 59-05-2; DDF, 5854-11-5; (6*R*)-*N*-10-formyl-H_4folate, 74644-66-9; carbocyclic β-GAR, 116925-84-9; *N*-10-formylfolate, 134-05-4; *N*-10-formyl-5-DF, 76622-75-8; *N*-10-formyl-8-DF, 76622-76-9; (6*S*)-*N*-10-formyl-H_4folate, 73650-39-2; *N*-10-formyl-DDF, 61038-31-1; *N*-10-formyl-2,4-diamino-DDF, 124512-43-2; [*p*-(benzylformylamino)benzoyl]-L-glutamate, 76622-77-0; *N*-10-(bromoacetyl)-5,8-DDF, 101375-70-6.

REFERENCES

Achanya, S. P., & Hynes, J. B. (1975) *J. Heterocycl. Chem. 12*, 1283–1286.

Black, S. L., Black, M. J., & Mangum, J. H. (1978) *Anal. Biochem. 90*, 397–401.

Brixner, D. I., Ueda, T., Cheng, Y.-C., Hynes, J. B., & Broom, A. D. (1987) *J. Med. Chem. 30*, 675–678.

Buchanan, J. M. (1982) in *Amino Acids, Fermentations, and Nucleic Acids: A Symposium* (Snell, E. E., Ed.) pp 57–92, Annual Reviews, Palo Alto, CA.

Buchanan, J. M., & Hartman, S. C. (1959) *Adv. Enzymol. Relat. Areas Mol. Biol. 21*, 199–261.

Calvert, A. H., Jones, T. R., Jackman, A. L., Brown, S. J., & Harrap, K. R. (1980) in *Advances in Tumor Prevention, Detection and Characterization* (Davis, W., Harrap, K. R., & Stathopoulos, G., Eds.) pp 272–283, Excerpta Medica, Amsterdam.

Caperelli, C. A. (1985) *Biochemistry 24*, 1316–1320.

Caperelli, C. A. (1989) *J. Biol. Chem. 264*, 5053–5057.

Caperelli, C. A., & Price, M. A. (1988) *Arch. Biochem. Biophys. 264*, 340–342.

Chervenka, C. H. (1969) in *A Manual of Methods for the Analytical Ultracentrifuge*, pp 42–55, Spinco Division of Beckman Instruments.

Chettur, C. A., & Benkovic, S. J. (1977) *Carbohydr. Res. 56*, 75–86.

Cleland, W. W. (1979) *Methods Enzymol. 63*, 103–138.

Dagert, M., & Ehrlich, S. D. (1979) *Gene 6*, 23–28.

Daubner, S. C., & Benkovic, S. J. (1985) *Cancer Res. 45*, 4990–4997.

Daubner, S. C., Schrimsher, J. L., Schendel, F. J., Young, M., Henikoff, S., Patterson, D., Stubbe, J., & Benkovic, S. J. (1985) *Biochemistry 24*, 7059–7062.

Daubner, S. C., Young, M., Sammons, R. D., Courtney, L. F., & Benkovic, S. J. (1986) *Biochemistry 25*, 2951–2957.

Dev, I. K., & Harvey, R. J. (1978) *J. Biol. Chem. 256*, 4242–4244.

Hards, R. G., Benkovic, S. J., Van Keuren, M. L., Graw, S. L., Drabkin, H. A., & Patterson, D. (1986) *Am. J. Hum. Genet. 39*, 179–185.

Henikoff, S. (1986) *J. Mol. Biol. 190*, 519–528.

Henikoff, S. (1987) *BioEssays 6*, 8–13.

Inglese, J. (1989) Ph.D. Dissertation, The Pennsylvania State University.

Inglese, J., Blatchley, R. A., & Benkovic, S. J. (1989) *J. Med. Chem. 32*, 937–940.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, pp 440–441, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

McMeekin, T. L., & Marshall, L. (1952) *Science 116*, 142.

Mieschendahl, M., & Muller-Hill, B. (1985) *J. Bacteriol. 164*, 1366–1369.

Miller, J. H. (1972) in *Experiments in Molecular Genetics*, pp 1–466, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Neidhardt, F. C., Bloch, P. L., & Smith, D. F. (1974) *J. Bacteriol. 119*, 736–747.

Perrin, D. D., Armarego, W. L. F., & Perrin, D. R. (1980) in *Purification of Laboratory Chemicals*, Pergamon Press, New York.

Remault, E., Stanssens, P., & Fiers, W. (1981) *Gene 15*, 81–93.

Rowe, P. B. (1984) in *Folates and Pterins* (Blakley, R. L., & Benkovic, S. J., Eds.) Vol. 1, pp 329–344, Wiley, New York.

Schachman, H. K. (1957) *Methods Enzymol. 4*, 32–103.

Schrimsher, J. L., Schendel, F. J., Stubbe, J., & Smith, J. M. (1986) *Biochemistry 25*, 4366–4371.

Segel, I. H. (1975) *Enzyme Kinetics*, Wiley, New York.

Smith, G. K., Benkovic, P. A., & Benkovic, S. J. (1981a) *Biochemistry 20*, 4034–4036.

Smith, G. K., Mueller, W. T., Benkovic, P. A., & Benkovic, S. J. (1981b) *Biochemistry 20*, 1241–1245.

Smith, J. M., & Daum III, H. A. (1986) *J. Biol. Chem. 261*, 10632–10636.

Smith, J. M., & Daum III, H. A. (1987) *J. Biol. Chem. 262*, 10565–10569.

Srinivasan, A., & Broom, A. D. (1981) *J. Org. Chem. 46*, 1777–1781.

Stone, S. R., & Morrison, J. F. (1983) *Biochim. Biophys. Acta 745*, 247–258.

Stura, E., Johnson, D. L., Inglese, J., Smith, J. M., Benkovic, S. J., & Wilson, I. A. (1989) *J. Biol. Chem. 264*, 9703–9706.

Tiedeman, A. A., Smith, J. M., & Zalkin, H. (1985) *J. Biol. Chem. 260*, 8676–8679.

Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103–119.

Young, M. (1985) Ph.D. Dissertation, The Pennsylvania State University.

Young, M., Sammons, R. D., Mueller, W. T., & Benkovic, S. J. (1984) *Biochemistry 23*, 3979–3986.