Crystal Structure of Glycine *N*-Methyltransferase from Rat Liver[†,‡]

Zhuji Fu,[§] Yongbo Hu,[§] Kiyoshi Konishi,[||] Yoshimi Takata,[||] Hirofumi Ogawa,[||] Tomoharu Gomi,[||] Motoji Fujioka,*[,||] and Fusao Takusagawa*[,§]

Department of Biochemistry, University of Kansas, Lawrence, Kansas 66045, and Department of Biochemistry, Faculty of Medicine, Toyama Medical and Pharmaceutical University, Sugitani, Toyama 930-01, Japan

Received May 3, 1996; Revised Manuscript Received July 3, 1996[⊗]

ABSTRACT: Glycine *N*-methyltransferase (GNMT) from rat liver is a tetrameric enzyme with 292 amino acid residues in each identical subunit and catalyzes the *S*-adenosylmethionine (AdoMet) dependent methylation of glycine to form sarcosine. The crystal structure of GNMT complexed with AdoMet and acetate, a competitive inhibitor of glycine, has been determined at 2.2 Å resolution. The subunit of GNMT forms a spherical shape with an extended N-terminal region which corks the entrance of active site of the adjacent subunit. The active site is located in the near center of the spherical subunit. As a result, the AdoMet and acetate in the active site are completely surrounded by amino acid residues. Careful examination of the structure reveals several characteristics of GNMT. (1) Although the structure of the AdoMet binding domain of the GNMT is very similar to those of other methyltransferases recently determined by X-ray diffraction method, an additional domain found only in GNMT encloses the active site to form a molecular basket, and consequently the structure of GNMT looks quite different from those of other methyltransferases. (2) This unique molecular structure can explain why GNMT can capture folate and polycyclic aromatic hydrocarbons. (3) The unique N-terminal conformation and the subunit structure can explain why GNMT exhibits positive cooperativity in binding AdoMet. From the structural features of GNMT, we propose that the enzyme might be able to capture yet unidentified molecules in the cytosol and thus participates in various biological processes including detoxification of polycyclic aromatic hydrocarbons. In the active site, acetate binds near the S-CH_3 moiety of AdoMet. Simple modeling indicates that the amino group of the substrate glycine can be placed close to the methyl group of AdoMet within 3.0 Å and form a hydrogen bond with the carboxyl group of Glu^{15} of the adjacent subunit. On the basis of the ternary complex structure, the mechanism of the methyl transfer in GNMT has been proposed.

Biological methylation reactions utilizing *S*-adenosylmethionine (AdoMet) as the methyl donor are widespread in nature and participate in a wide variety of cellular processes through methylation of nucleic acids, proteins, phospholipids, and small molecules. Small molecule methyltransferases may be divided into two categories: (1) the enzymes involved in the biosynthesis and degradation of bioactive amines, *e.g.*, norepinephrine *N*-methyltransferase, acetylserotonin methyltransferase, and catechol *O*-methyltransferase, and (2) the enzymes involved in bulk metabolic transformations such as glycine *N*-methyltransferase and guanidinoacetate methyltransferase.

Glycine *N*-methyltransferase (*S*-adenosyl-L-methionine: glycine methyltransferase, EC 2.1.1.20; GNMT) catalyzes the AdoMet-dependent methylation of glycine to form sarcosine (*N*-methylglycine). GNMT, unlike most AdoMet-dependent methyltransferases, is an oligomeric protein consisting of four identical subunits (Ogawa & Fujioka, 1982a). The primary structure of the enzyme has been deduced from the nucleotide sequence of a cloned cDNA, and each subunit has been shown to contain 292 amino acid residues (Ogawa *et al.*, 1987). Fujioka and his co-workers have characterized chemical and biological properties of the enzyme isolated from rat liver (Ogawa *et al.*, 1987; Konishi & Fujioka, 1987, 1988).

From enzymological and nutritional evidence, GNMT has been implicated as a major enzyme that regulates the AdoMet/AdoHcy ratio in some organisms (Ogawa & Fujioka, 1982a,b; Cook *et al*., 1989; Wagner *et al*., 1985; Balaghi *et al*., 1993). In addition to this role, GNMT has been shown to be a major folate binding protein of rat liver cytosol (Cook & Wagner, 1984) and, more recently, to be a polycyclic aromatic hydrocarbon binding protein (Raha *et al.*, 1994) and a mediator of cytochrome P450 1A1 gene expression (Raha *et al*., 1995).

To date, the crystal structures of only four AdoMet-dependent methyltransferases, catechol *O*-methyltransferase (COMT) (Vidgren *et al*., 1994), *Hha*I DNA methyltransferase (HhaI) (Cheng *et al*., 1993), *Taq*I DNA methyltransferase (TaqI) (Labahn *et al.*, 1994), and *Hae*III DNA methyltransferase (HaeIII) (Reinisch *et al*., 1995), have been reported. These enzymes, despite having low degrees of sequence homology, show strikingly similar structures in the AdoMet binding domains, suggesting that methyltransferases in general have a common structure at the AdoMet binding site. Here we report the crystal structure of GNMT complexed

[†] The work carried out at the University of Kansas has been supported by NIH Grant GM37233 and in part by the Kansas Health Foundation and the Marion Merrell Dow Foundation.

[‡] The atomic coordinates have been deposited with the Brookhaven Protein Data Bank (entry name 1XVA).

* Authors to whom correspondence should be addressed.

[§] University of Kansas.

[||] Toyama Medical and Pharmaceutical University.

[⊗] Abstract published in *Advance ACS Abstracts,* September 1, 1996.

S0006-2960(96)01068-9 CCC: $12.00

Table 1: Experimental Details and Refinement Parameters of Crystal Structure Analyses: Space Group, $P2_12_12$; Cell Dimension (Å), $a = 86.4$, $b = 175.7$, $c = 45.5$; M_r of Subunit, 31 500; Number of Subunits in Unit Cell, 8; Solvent Content, 54%

Experimental Details					
data set	GNMT	GNMT-AdoMet	GNMT-Au	GNMT-Pt	
resolution	20.0−2.2	20.0−3.0	20.0−3.0	20.0−3.0	
no. of crystals	7	1	8	6	
no. of reflections measured	146394	31935	44694	51962	
no. of unique reflections[a]	33778	11657	14203	12850	
% completeness	93.5	80.3	97.8	88.5	
R_{sym}[b]	0.068	0.056	0.074	0.071	
$I/\sigma(I)$[c]	3.4	6.8	4.5	3.8	
no. of heavy atom sites			4	2	
RFD[d]		0.047	0.161	0.142	
R_c[e]			0.383	0.439	
phasing power[f]			1.793	1.640	
FOM of SIR[g]			0.39 (0.80)	0.34 (0.81)	
FOM of MIR[h]			0.54 (0.79)	0.54 (0.79)	
Refinement Parameters					
no. of residues	584	rms deviations		mean B values	
no. of AdoMet molecules	2	bond (Å)	0.014	Cα atoms (Å^3)	19.6
no. of water molecules	94	angle (deg)	3.1	main chain (Å^3)	20.0
R[i]	0.196	torsion angle (deg)	24.9	all atoms (Å^3)	21.6
free R	0.262				

[a] Unique reflections in the range between 10.0 Å and the highest resolution. [b] $R_{sym} = \sum|I - <I>|/\sum I$. [c] $I/\sigma(I)$ in the 2.2−2.4 Å resolution range for GNMT data sets and in the 3.0−3.1 Å resolution range for the other data sets. [d] RFD $= \sum|F_{PH} - F_P|/\sum F_P$. In the phase determination, the GNMT−Au and GNMT−Pt data were cut off at 4.0 Å resolution. [e] $R_c = \sum||F_{PH} \pm F_P| - F_H(\text{calcd})|/\sum|F_{PH} \pm F_P|$. [f] Phasing power $= <F_H>/<E>$: the rms heavy-atom structure factor amplitude divided by the residual lack of closure error. [g] FOM of SIR = figure of merit of single isomorphous replacement. The value after density modifications and phase extension to 3.0 Å resolution is given in parentheses. The phasing statistics listed in the table are for 4.0 Å resolution data. The completeness of the data at the 4.0 Å resolution level is more than 98.7% for all data sets. [h] FOM of MIR = figure of merit of multiple isomorphous replacement. The value after density modifications is given in parentheses. [i] $R = \sum|F_o - F_c|/\sum F_o$.

with AdoMet and acetate, a competitive inhibitor of glycine, determined at 2.2 Å resolution.

EXPERIMENTAL PROCEDURES

Purification and Crystallization Procedures. GNMT used in this study is the recombinant rat enzyme produced in *Escherichia coli* JM109 transformed with pCWGNMT plasmid that contains the coding sequence of rat GNMT cDNA. The enzyme was purified to homogeneity from *E. coli* extracts by gel filtration over Sephacryl S-300 and DEAE-cellulose chromatography. Recombinant GNMT lacks the N-terminal acetyl group but exhibits other structural features similar to those of the liver enzyme (M. Fujioka, unpublished results).

The hanging drop method of vapor diffusion was employed for crystallization of the enzyme. All crystallization experiments were conducted at 4 °C. Crystals were grown in a buffer containing 100 mM sodium citrate buffer (pH = 5.6), 100 mM ammonium acetate, 50 mM NaCl, and 13% (w/v) PEG 4000 with a protein concentration of 10 mg/mL. The thick plate-shaped crystals suitable for X-ray diffraction studies (∼1.0 × 0.5 × 0.2 mm) were grown for 2 weeks. Crystals were also prepared in a similar solution that contained 5 mM AdoMet. The gold ($K[Au(CN)_2]$) and platinum ($K_2[Pt(CN)_4]$) derivative crystals were prepared by the soaking method. For the gold derivative crystals, the crystals prepared in the absence of AdoMet were incubated in an artificial mother liquor containing 100 mM sodium citrate buffer (pH = 5.6), 100 mM ammonium acetate, 50 mM NaCl, 13% (w/v) PEG 4000, and 4 mM $K[Au(CN)_2]$ for 1 week. For the platinum derivative crystals, the crystals were incubated in an artificial mother liquor containing 100 mM sodium citrate buffer (pH = 5.6), 100 mM sodium acetate (pH = 6.5), 50 mM NaCl, 13% (w/v) PEG 4000, and 2 mM $K_2[Pt(CN)_4]$ for 10 h before they were mounted into the capillary.

Data Measurement. The crystals were mounted in the glass capillaries with the mother liquor and placed on a locally modified DIP100S imaging plate X-ray diffractometer (manufactured by MAC Science, Tokyo, Japan) with a rotating anode X-ray generator as X-ray source (Cu Kα radiation operated at 50 kV and 150 mA). The diffraction data of native crystals were measured up to 2.2 Å resolution at room temperature, while those of the heavy-atom derivatives and the crystal grown with AdoMet were measured up to 3.0 Å resolution. The data were processed with the program ELMS (Tanaka *et al.*, 1990). Integrated reflections were scaled and reduced with locally developed programs (Takusagawa, 1992). The data statistics are given in Table 1.

Structure Determination. The gold positions (four sites) and platinum positions (two sites) were located in difference Patterson maps and were refined by the locally developed program HEVY. The positions are listed in Table 2. Since two pairs of the gold atoms and one pair of the platinum atoms were apparently related by a pseudo-2-fold symmetry, the noncrystallographic 2-fold axis was determined by a least-squares method using these pairs of heavy-atom coordinates. The noncrystallographic 2-fold axis makes a 90° angle with the crystallographic 2-fold axis, creating a local pseudo-222 symmetry at (1.0, 0.5, 0.103).

Initial protein phases of 3.0 Å resolution were determined from the multiple isomorphous replacement (MIR) using the gold and platinum derivative data. Although the map calculated with the phases showed the tetrameric feature of the enzyme related by the pseudo-222 symmetry, the map was not interpretable. The phases between 3.0 and 4.0 Å resolution were discarded, and the phases up to 4.0 Å

Table 2: Refined Heavy-Atom Parameters[a]

site	*x*	*y*	*z*	occupancy[b]	*B*[c]	binding site
Au(1a)	0.532	0.424	−0.267	0.164	15.5	Cys^{262}
Au(2a)	0.567	0.338	−0.133	0.223	14.5	Cys^{57}
Au(1b)	0.275	0.296	0.472	0.164	15.5	Cys^{262B}
Au(2b)	0.142	0.239	0.333	0.223	14.5	Cys^{57B}
Pt(1a)	0.793	0.503	−0.008	0.260	16.3	Met^{215}
Pt(1b)	0.158	0.439	0.218	0.260	16.3	Met^{215B}

[a] The 530 and 436 centric reflections in the resolution between 10.0 and 4.0 Å were used to refine the positional, occupancy, and temperature parameters of Au and Pt derivatives, respectively. The function minimized is $\sum w\{|F_{PH}| - k[|F_P| \pm F_H(\text{calcd})]\}^2$, where $w = 1/(\sigma F_{PH}^2 + \sigma F_P^2)$ and k is a scale factor. The *R*-factors as defined, $R = \sum||F_{PH}| - k[|F_P| \pm F_H(\text{calcd})]|/\sum|F_{PH}|$, are 0.107 for the Au derivative and 0.129 for the Pt derivative. [b] Occupancy on arbitrary scale. [c] Isotropic temperature factor (Å^2).

resolution were refined and gradually extended to 3.0 Å resolution by combination of the solvent flattening (Wang, 1985) and the noncrystallographic symmetry averaging methods. The map calculated with the new phases was improved. The entire main chain except for 1−4, 188−192, and 290−293 was traced without ambiguity. The initial model was built on an IRIS workstation using the program TOM/FRODO (Jones, 1985; Cambillau & Horjales, 1987). The model was refined with the positional protocol and then with the simulated annealing procedure of X-PLOR (Brünger, 1993). The model was rebuilt where necessary, and previously undefined residues were built into the electron density map.

During later stages of refinement, difference maps calculated with coefficients of $(F_o - F_c)$ and calculated phases showed a large significant residual electron density peak in the region of the active site (Figure 1). The electron density distribution indicated the presence of what appeared to be an AdoMet or AdoHcy molecule in the active site. To confirm the residual electron density peak in the active site, another difference map was calculated using the data of the crystal prepared in the presence of AdoMet (3.0 Å resolution data). The map clearly showed a residual electron density for the AdoMet molecule (Figure 1). Furthermore, the agreement on F_o^2 between the data of the crystal prepared in presence and absence of AdoMet was very good ($R = 0.047$), indicating that these data are the same within experimental error. As discussed below, since the relative location and orientation of the AdoMet molecule in the maps are quite similar to those found in the COMT structure, the residual electron density was assigned to the AdoMet molecule rather than cluster of solvent molecules. In the active site, the other significant residual electron density peak was assigned to an acetate since the crystals were grown in the solution containing 100 mM ammonium acetate and acetate binds to the glycine site in the E−AdoMet complex (Konishi & Fujioka, 1987). Other well-defined residual electron density peaks in difference maps were assigned to water molecules if peaks were able to bind the protein molecules with hydrogen bonds.

Refinement of isotropic temperature factors for individual atoms was carried out by the individual *B*-factor refinement procedure of X-PLOR using bond and angle restraints. After multiple cycles of model building and refinement, all residues except for a loop were built into the electron density map. The $(2F_o - F_c)$ map showed that the electron density peaks for residues 188−192, which are on the surface of the protein, were disconnected at the 1.25σ contour level. Distribution of the electron density peaks indicated that the loop was 2-fold disordered. Thus, the loop (188−192) was built in two possible routes and assigned a half-occupancy for each route. The noncrystallographic 2-fold symmetry restriction was gradually removed after the crystallographic *R*-factor was below 0.24. The final crystallographic *R*-factor is 0.19 for all data from 10.0 to 2.2 Å resolution. The free *R*, determined from a randomly selected 5% of the data, is 0.26. The coordinates have been deposited with the Brookhaven Protein Data Bank. The molecular diagrams are drawn with ORTEP (Johnson, 1965) and locally developed graphics programs (S. Kamitori and F. Takusagawa, DRAEMO and FPLOT, unpublished work).

RESULTS

The crystallographic refinement parameters (Table 1), final $(2F_o - F_c)$ maps, and conformational analysis by PROCHECK (Laskowski *et al.*, 1993) indicate that the structure of GNMT has been determined successfully. GNMT consists of four subunits related by pseudo-222 symmetry (Figure 2A). A large solvent channel (50 × 15 Å) passes through the center of the tetramer. The four subunits are denoted subunits A, B, C, and D. Subunit A and B (or C and D) are related by a pseudo-2-fold symmetry, and subunit A and C (or B and D) are related by a crystallographic 2-fold symmetry. We use the following convention to distinguish residues in subunits A, B, C, and D. A residue name with no added symbol refers to a residue of subunit A; a residue name with an appended B, C, and D refers to a residue of subunit B, C, and D, respectively. For example, Glu^{15} means Glu^{15} of subunit A, and Glu^{15B} means Glu^{15} of subunit B. The convention also applies to elements of secondary structure (α-helix, H1, $H1^B$; β-strand, B1, $B1^B$, etc.). The rms deviation between the Cα's in two subunit structures related by a pseudo-2-fold symmetry is less than 0.5 Å, indicating that the subunits are structurally very similar (Table 3). For simplicity, the following description mainly refers to subunit A.

The GNMT subunit is approximately spherical in shape except for the N-terminal 20 amino acid residues. Each subunit is composed of three domains, with the peptide chain organized into 9 α-helices and 11 β-strands assigned by the program PROCHECK (Figures 2 and 3). The three domains are denoted the N-domain (residues 1−36), the C-domain (residues 37−175 and 243−292), and the S-domain (residues 176−242). In the N-domain, the N-terminal region (1−20 residues), which is mostly composed of random coil, is far from the body of the subunit and goes into the opposite subunit. The S-domain is composed of an α-helix, followed by a large loop and three antiparallel β-strands. In the C-domain, a typical α/β folding of the polypeptide is observed. A β-sheet composed of seven mixed β-strands is sandwiched by pairs of three α-helices (Figure 3A).

In the tetramer, subunits are mainly connected by hydrophilic interactions. Unlike most oligomeric proteins, hydrophobic interactions make little contribution in GNMT. Except for interactions of the N-terminal regions, subunit A contacts to subunits B and D at two positions. One is between helices H4−H5 and $H4^B$−$H5^B$, where six salt linkages and two hydrogen bonds connect the subunits (Figure 4). The other is between the β-strand B7 and $B7^D$,

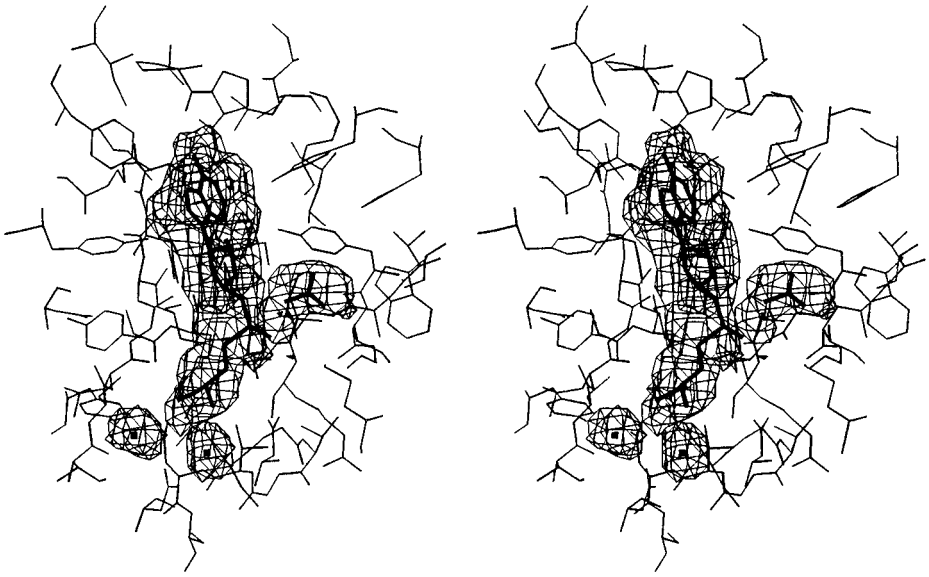

FIGURE 1: Difference electron density map (omit map) calculated after 30 cycles of positional refinement of the protein structure by X-PLOR, showing well-defined residual electron density peaks of the AdoMet, acetate, and two water molecules in the active site. The contours are drawn at the 2.5σ level. The final model is superimposed in the map.

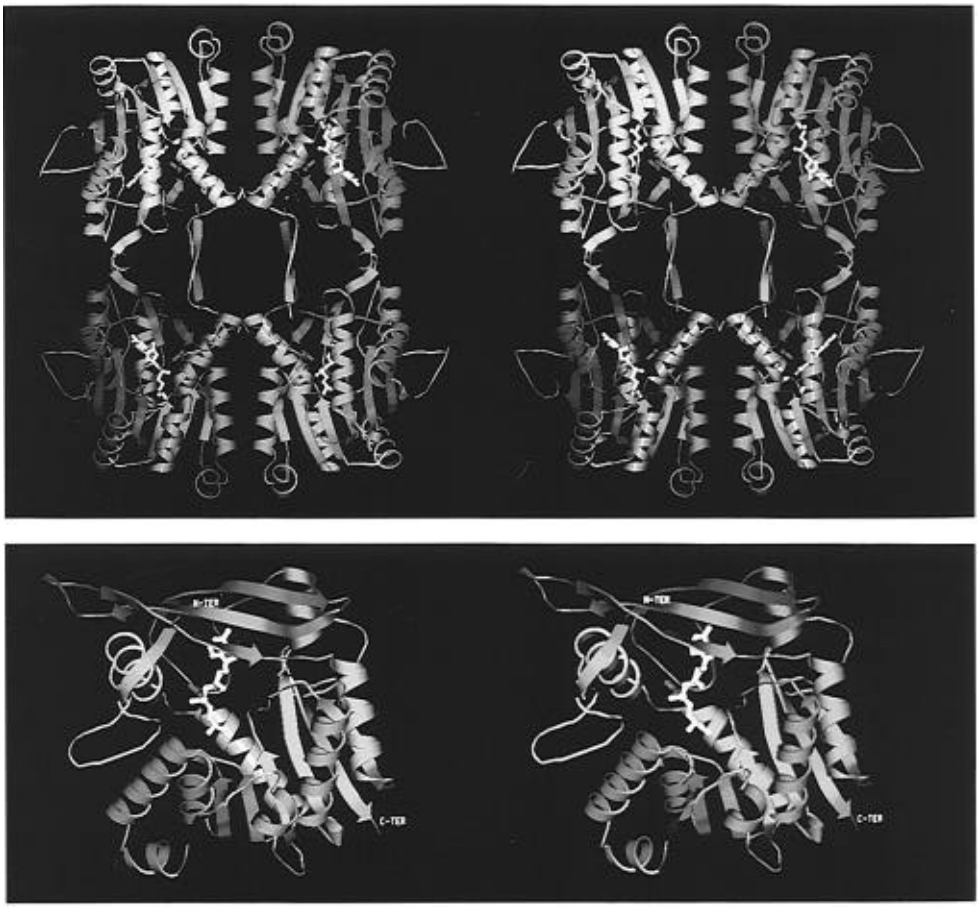

FIGURE 2: (A, top) Tetrameric structure of GNMT. The four subunits A, B, C, and D are colored yellow, blue, magenta, and green, respectively. AdoMet and acetate in the active site are colored with white and red, respectively. It should be noted that the extended N-terminal regions go into adjacent subunits. Residues 10−20 in each subunit form a U-shape loop and hang into the active site of the adjacent subunit. Also, the tail section of residues 2−7 is involved in antiparallel β-sheet hydrogen bonds. (B, bottom) Subunit structure with AdoMet and acetate viewed along the active entrance. The N-, C-, and S-domains are colored yellow, magenta, and green, respectively.

where eight consecutive antiparallel β-sheet hydrogen bonds from residue 204 to 210 connect the subunits. The N-terminal segment (residues 1−20), apart from the main body of the subunit, goes into the entrance of active site of subunit B as if this portion closes the active site of subunit B (Figures 2A and 5B). The tail section of the N-terminal region (residues 2−7) is involved in the antiparallel β-sheet hydrogen bonds with the same section of subunit D.

As shown in Figure 1, the electron density maps clearly show the presence of an AdoMet and an acetate molecule in the near center of each subunit. Since acetate is a competitive inhibitor of glycine, this region must be the

Table 3: Root-Mean-Square (rms) Deviations (Å) between Two Structures

structure 1	structure 2	Cα (292 atoms)	main chain (1168 atoms)	all atoms (4622 atoms)
GNMT-1	GNMT-2	0.46	0.50	1.19
GNMT-1	HhaI[a]	1.40		

[a] The residues in the β-strands in the AdoMet binding domain were compared as described in Figure 8.

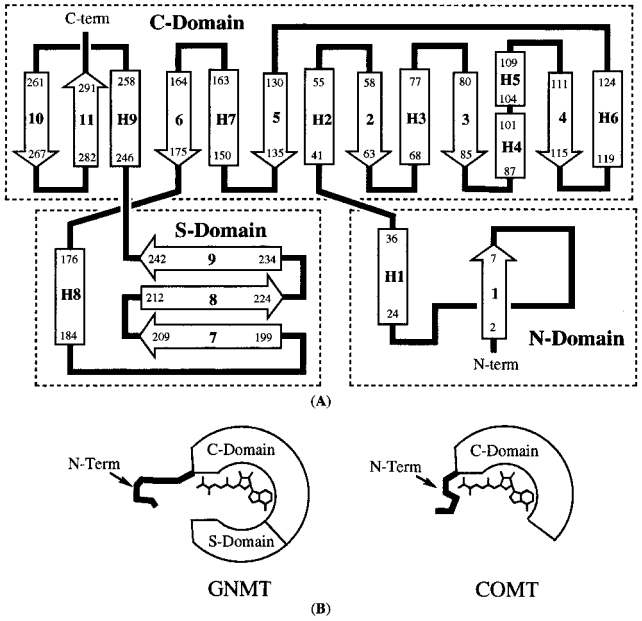

FIGURE 3: (A) Topology diagram showing the N-, C-, and S-domains. The α-helices and β-strands are indicated by rectangles and arrows, respectively. Small numbers in the rectangles and arrows indicate amino acid residue numbers. (B) Schematic diagram showing the relative location of the S-domain in the GNMT and the major structural difference between GNMT and COMT.

active site of the enzyme. The environment around the AdoMet and acetate is shown in Figure 6. Possible interactions between amino acid residues and the AdoMet and acetate are illustrated in Figure 7. The adenine moiety has two hydrogen bonds with GNMT (N6−O[Pro^{187}] and N3−OH[Tyr^{33}]). The hydroxyl group O2′ of ribose is involved in a hydrogen bond with Tyr^{242} (O2′−OH[Tyr^{242}]). The amino and carboxyl groups of the methionine moiety are surrounded by hydrophilic residues within hydrogen bond distance. These are O1−O[Leu^{136}], O1−Oδ1[Asp^{70}], O1−OH_2, O2−Oγ[Ser^{139}], O2−Oϵ1[Glu^{15B}], N−Oδ1[Asp^{70}], and N−OH_2. The carboxyl group of the acetate hydrogen bonds to Trp^{30} and Ala^{13B} (O1−Nϵ1 [Trp^{30}] and O2−O[Ala^{13B}]), and its methyl group is located near the S-CH_3 moiety. Interestingly, the AdoMet and acetate are completely surrounded by amino acid residues, and thus no apparent entrance or exit of the active site is observed (Figure 6B).

DISCUSSION

GNMT, unlike most AdoMet-dependent methyltransferases, is an oligomeric protein consisting of four identical subunits (Ogawa & Fujioka, 1982a). This study reveals that the GNMT tetramer has a flat square shape with a large hole in the middle. Each subunit of GNMT forms approximately spherical shape, and the amino acid residues on the subunit surface are mostly hydrophilic. Four subunits are connected two-dimensionally mainly by hydrogen bonds. The flat and square-shaped tetramer exposes to the solvent not only at its surface but also in the central regions. Thus, although the N-terminal regions (residue 1−20) are apart from the main bodies of subunits and interact with each other in the central region of the tetrameric enzyme in the crystal structure, these regions might be movable in solution. These characteristic features indicate that the tetrameric GNMT has evolved from a monomeric enzyme.

The crystal structures of four methyltransferases with bound AdoMet, catechol *O*-methyltransferase (COMT) (Vidgren *et al.*, 1994), *Hha*I DNA methyltransferase (HhaI) (Cheng *et al.*, 1993), *Taq*I DNA methyltransferase (TaqI) (Labahn *et al.*, 1994), and *Hae*III DNA methyltransferase (HaeIII) (Reinisch *et al.*, 1995), have recently been reported. The AdoMet binding domains of these methyltransferases are strikingly similar to each other, suggesting that many methyltransferases may have a common structure (Schluckebier *et al.*, 1995). Indeed, the polypeptide folding pattern in the C-domain of GNMT is very similar to those of other methyltransferases (Figure 3A). Interestingly, if the S-domain is removed and β-strand B6 is directly connected to the helix H9, then the peptide folding pattern from helix H1 (residue 24) to the C-terminus (residue 292) is exactly the same as the folding of helix α2 (residue 23) through the C-terminus (residue 221) in COMT. Furthermore, as shown in Figure 8, the central regions forming characteristic β-sheet and α-helices can be superimposed on the corresponding regions of DNA methyltransferases, indicating that the tertiary structures of AdoMet binding domains are indeed quite similar throughout AdoMet-dependent methyltransferases. Relative position and orientation of the AdoMet molecule in GNMT are similar to those in COMT. However, it should be noted that the AdoMet binding pocket of the *E. coli* methionine repressor has no resemblance to the AdoMet binding pockets found in the methyltransferases (Rafferty *et al.*, 1989). The structure of AdoMet synthetase determined recently in our laboratory shows no structural similarity either to the methyltransferases or to the repressor (Takusagawa *et al.*, 1996).

Sarcosine is a nonstandard amino acid that is not found in proteins of eukaryotic organisms. Nevertheless, GNMT has been found in the liver and kidney of various mammalian species (Kerr, 1972). In rabbit and rat livers, the relative amount of GNMT is very high in comparison with the other AdoMet-dependent methyltransferases. For example, the concentration of GNMT in rat liver cytosol is 3520 times higher than that of tRNA methyltransferase (Kerr, 1972). Despite its abundance, the role of GNMT is not clear, and the enzyme has been suggested to play a role in modulating the relative level of AdoMet and AdoHcy in the cell (Kerr, 1972; Ogawa & Fujioka, 1982a,b; Cook *et al.*, 1989). The present crystallographic study suggests new roles for GNMT. Whereas the structures of AdoMet binding domains of GNMT and other methyltransferases (COMT, HhaI, TaqI, and HaeIII) are very similar, the overall structure of the AdoMet binding unit of GNMT appears to be unique among methyltransferases. The major difference is that the S-domain found in the GNMT structure is absent in the three methyltransferases mentioned above (Figure 3B). The active sites of the latter methyltransferases are located on the protein surface and thus exposed to the solvent pool as commonly seen in most enzymes. Thus, only limited numbers of specific molecules could bind to the active sites. On the

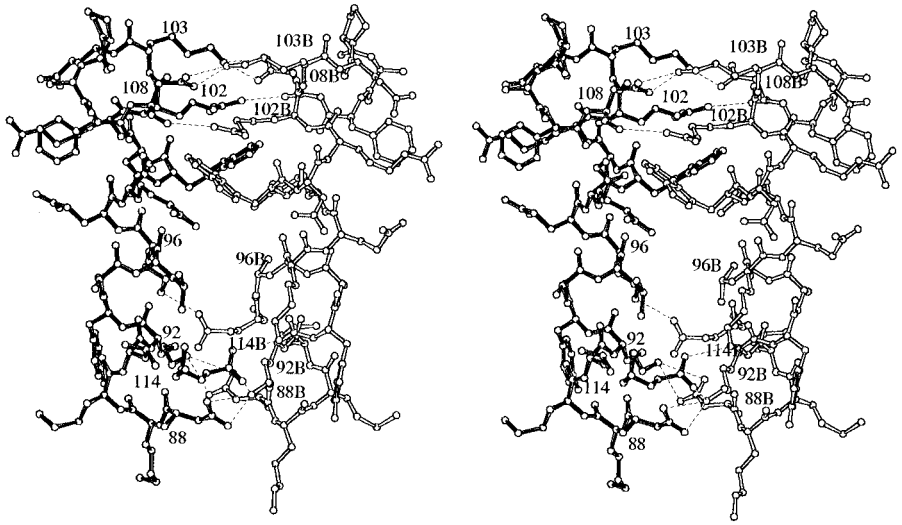

FIGURE 4: Short contact region between subunits A and B. The helices H4−H5 are connected to the H4^{B}−H5^{B} by six salt linkages (Lys^{103}−Asp^{108B}, Lys^{103B}−Asp^{108}, Lys^{92}−Asp^{88B}, Lys^{92B}−Asp^{88}, Lys^{96}−Glu^{114B}, Lys^{96B}−Glu^{114}) and two hydrogen bonds (NH1[Arg^{102}]−O[Asp^{108B}], NH1[Arg^{102B}]−O[Asp^{108}]). In addition to the hydrophilic interaction, Trp^{99} and Trp^{99B} are stacked on Arg^{98B} and Arg^{98}, respectively. The residues of subunit A are shown by solid bonds, and those of subunit B by open bonds.

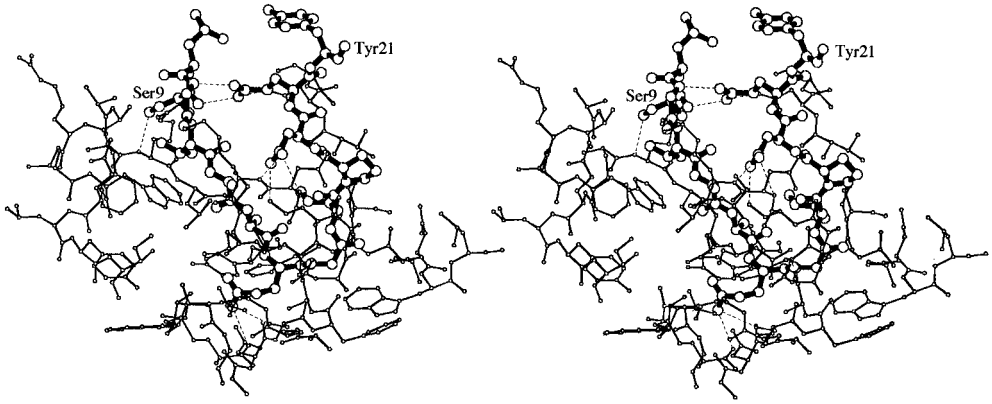

FIGURE 5: Interactions of the U shaped N-terminal segment (^{9}SLGVAAEGIPDQY^{21}) with subunit B, which corks the active site of subunit B. The Oϵ1 and Nϵ2 of Gln^{20} hydrogen-bond to N[Gly^{11}] and O[Ser^{9}], respectively, to form a U-shaped loop. Possible hydrogen bonds between the N-terminal section (residues 9−21) of subunit A and subunit B are as follows: N[Ser^{9}]−O[Arg^{239B}]; O[Gly^{11}]−Nζ [Lys^{89B}]; Oϵ2[Glu^{15}]−Oγ[Ser^{139B}]; O[Glu^{15}]−Oδ2[Asp^{85B}]; Oδ1[Asp^{19}]−Oγ[Ser^{87B}]; Oδ2[Asp^{19}]−N[Lys^{89B}].

other hand, the active site of GNMT is mostly buried in the S-domain and located in the near center of the spherical subunit. Consequently, the GNMT subunit looks like a molecular basket. Thus, various molecules could be held in this molecular basket. Furthermore, as will be discussed below, the molecular basket is corked by the N-terminal regions of the adjacent subunit. This unique structural feature might be consistent with the fact that GNMT can bind not only AdoMet but also tetrahydrofolate (THF) (Cook & Wagner, 1984) and polycyclic aromatic hydrocarbon (PAH) molecules such as benzo[*a*]pyrene (Raha *et al.*, 1994). Although the exact chemical structures of AdoMet and THF are quite different, the overall molecular shapes and charge distributions seem to be quite similar to each other. For example, the adenine, ribose, and methionine moieties of AdoMet roughly correspond to the pteridine, *p*-aminobenzoate, and glutamate moieties of THF, respectively. In GNMT, the adenine moiety of AdoMet lies deep inside the active site, and the methionine moiety is near the entrance. If THF binds in the active site of GNMT in the same manner as AdoMet, even longer tailed THF analogues, such as 5-methyltetrahydropteroylpentaglutamate could be accommodated in the GNMT basket. It is quite reasonable to assume that PAH molecules bind near the adenine pocket of the active site. It should be noted that an unusually large number of tyrosine residues (33, 44, 177, 193, 194, 220, 242, 283) are located at the inner surface of active site. As depicted in Figure 7, Tyr^{33} and Tyr^{242} are the only residues that make direct contact with AdoMet. These and other tyrosine residues might interact with THF and water-soluble PAHs bound at this site. It is also possible that GNMT can capture yet unidentified compounds in the cytosol and thus participates in various biological processes including detoxification of polycyclic aromatic hydrocarbons. In this connection, it should be noted that GNMT has recently been identified to be a mediator of cytochrome P450 1A1 gene expression (Raha *et al.*, 1995). To our knowledge, GNMT is the first example of a molecular basket having a deep active site in the center of a subunit.

Another remarkable feature of the tetrameric GNMT structure is its extended N-terminal regions which go into adjacent subunits. Similar N/C-terminal exchanges between subunits have been observed in several oligomeric protein structures such as interferon-γ (Samudzi *et al.*, 1991) and 6-phosphogluconate dehydrogenase (Adams *et al.*, 1991). In the tetrameric GNMT structure, residues 2−7 of subunit A are tightly connected to residues 7−2 of subunit D, respectively, with six consecutive antiparallel β-sheet hydrogen

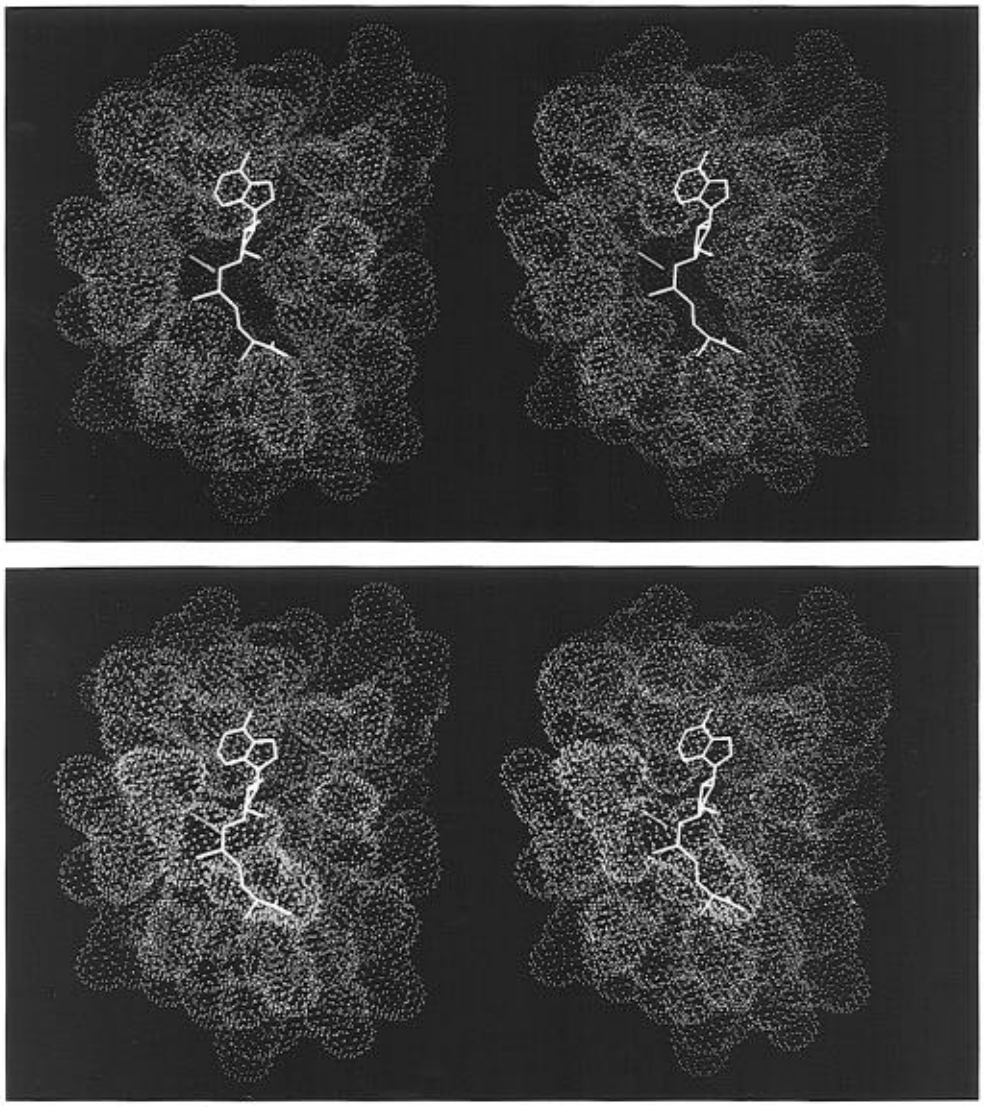

FIGURE 6: Active site with AdoMet (white) and acetate (red) drawn by van der Waals surface, showing the molecular basket structure. (A, top) View excluded the N-terminal regions of the adjacent subunit showing the entrance of the active site. (B, bottom) View included the N-terminal regions of the adjacent subunit showing the completely enclosed AdoMet molecule.

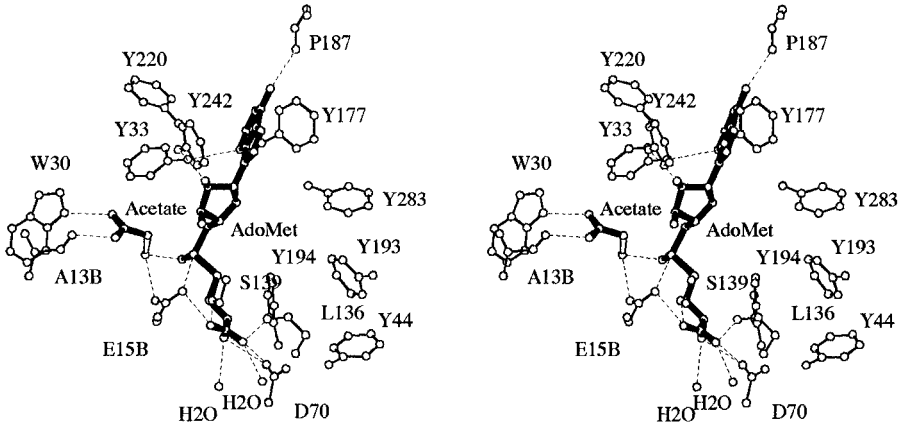

FIGURE 7: Active site geometry with AdoMet (solid bonds) and acetate. The glycine is built by attaching an amino group (open bond) to the acetate (solid bond) found in the active site. Possible polar interactions (O−O and O−N < 3.1 Å and O−S < 4.0 Å) are indicated by dotted lines. Tyrosine residues located at the inner surface of the active site are illustrated.

bonds. Residues 9−20 of subunit A form a "U"-shape loop and hang into the active site of subunit B as if these residues cork the entrance of the active site of subunit B (Figure 5). As a result, the AdoMet molecule when bound in the active site is completely surrounded by amino acid residues, and thus apparent entrance or exit of the active site is not recognizable in each subunit structure (Figure 6). It should be noted that AdoMet was located in the electron density map of the crystals grown in the absence of added AdoMet, indicating that AdoMet molecules were tightly trapped deep

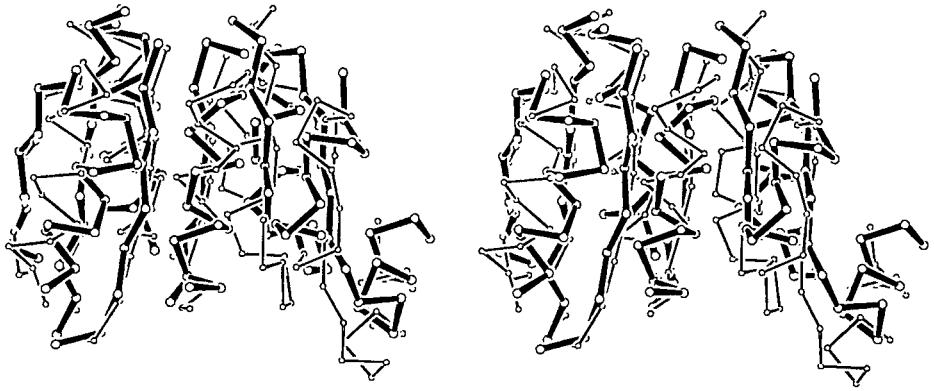

FIGURE 8: Comparison of the central regions of the AdoMet binding domain of GNMT and HhaI (from Brookhaven Protein Data Bank) showing how they are similar to each other. The Cα atoms in the β-strands are fitted by a least-squares method. GNMT and HhaI are drawn with thick and thin lines, respectively.

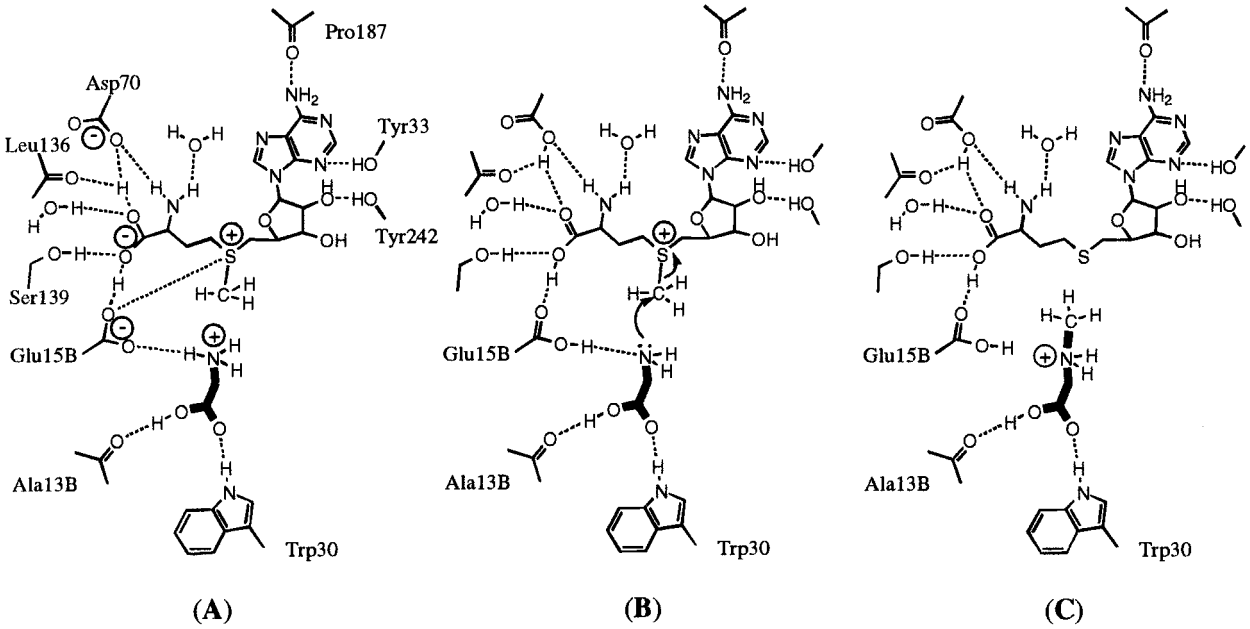

FIGURE 9: Proposed mechanism of the methyl transfer reaction of GNMT: (A) prereaction state, (B) intermediate state, and (C) postreaction state. Possible polar interactions are illustrated by dotted lines. The "+" and "−" in circles indicate a positive and negative charge, respectively. The acetate and sarcosine are drawn by thick bonds.

inside of subunits. Although TaqI does not have a molecular basket structure, the crystallized enzyme contains bound AdoMet (Labahn *et al.*, 1994).

The N-terminal residues 9−20 form a U-shaped loop connected by two hydrogen bonds (Oε1[Gln^{20}]−N[Gly^{11}] and Nε2[Gln^{20}]−O[Ser^{9}]). The carboxylate group of Glu^{15B} is located at the bottom of U (Figure 5), and its Oε1 hydrogen-bonds to O2 of AdoMet (one of the two carboxylates should be protonated to form a hydrogen bond). Interestingly, the Oε1 of Glu^{15B} points to the positively charged S atom of AdoMet at a distance of 4.0 Å, indicating some O^--S^+ interaction. Similar O^--S^+ interactions are seen in the HhaI and TaqI structures although AdoMet binds to the active sites in an opposite head−tail orientation compared with GNMT and COMT. In the active site, an acetate, a competitive inhibitor of glycine, is seen near the S-CH_3 moiety of AdoMet. An amino group was introduced to the methyl carbon of the acetate to mimic the substrate glycine. The C−C bond was rotated until the amino group was able to form a hydrogen bond with Oε2[Glu^{15B}] (N−O = 3.0 Å). In this conformation, the amino group of glycine was placed near the methyl group of AdoMet (N−C = 3.0 Å). On the basis of this structure, the mechanism of the methyl transfer reaction in GNMT may be considered as illustrated in Figure 9. GNMT binds substrates in an obligatory order (Konishi & Fujioka, 1987). AdoMet binds to the free enzyme, and glycine binds near Trp^{30} of the *E*−AdoMet complex. When the U loop of the adjacent subunit corks the active site as the result of substrate binding, the three hydrogen bonds and a weak O^--S^+ interaction could be formed between substrates (AdoMet and Gly) and the enzyme (Oε1[Glu^{15B}]−O2[AdoMet], O[Ala^{13B}]−O2[Gly], Oε2[Glu^{15B}]−N[Gly], and Oε1[Glu^{15B}]−S[AdoMet]). These interactions would force the amino group of glycine to move toward the methyl group of AdoMet (N−C < 3.0 Å), and the following series of events would take place at the active site: The carboxylate of Glu^{15B} neutralizes the positive charge of the glycine amino group to facilitate nucleophilic attack on the methyl carbon of AdoMet. Neutralization of Glu^{15B} should weaken the O^--S^+ interaction, and this would help to enhance the methyl transfer reaction.

GNMT shows sigmoidal rate behavior with respect to AdoMet (Ogawa & Fujioka, 1982a; Ogawa *et al.*, 1993), which is due to cooperative binding of AdoMet to the active

site residing on each subunit (Konishi & Fujioka, 1988). The unusual structural features of the tetrameric GNMT would explain why GNMT exhibits positive cooperativity in binding AdoMet. In the ternary complex with AdoMet and acetate, the N-terminal region of each subunit appears to cork the entrance of AdoMet to the adjoining subunit. In the free enzyme, it is obvious that no cork is present so as to allow binding of AdoMet. By changing a few torsion angles of the main chain bonds, the N-terminal region can be moved drastically. For example, when the ϕ (N−Cα) and ψ (Cα−C) angles of Gly^{24} are changed by 90° and −50°, respectively, the N-terminal region moves in the solvent channel and the entrance of the active site of the adjacent subunit would be completely opened. Whereas the structure of the free enzyme is not known, it would be reasonable to assume that the extended N-terminal regions interact each other to form an intricate structure. Thus, a conformational change brought about by binding of AdoMet to one subunit would be transmitted to the other subunit through the N-terminal network to result in the facilitated binding of AdoMet.

ACKNOWLEDGMENT

The authors express their thanks to Professor R. L. Schowen for valuable comments.

REFERENCES

Adams, M. J., Gover, S., Leaback, R., Phillips, C., & Somers, D. O'N. (1991) *Acta Crystallogr. B47*, 817−820.
Balaghi, M., Horne, D. W., & Wagner, C. (1993) *Biochem. J. 291*, 145−149.
Brünger, A. T. (1993) *X-PLOR 3.1: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
Cambillau, C., & Horjales, E. (1987) *J. Mol. Graphics 5*, 174.
Cheng, X., Kumar, S., Posfai, J., Pflugrath, J. W., & Roberts, R. J. (1993) *Cell 74*, 299−307.
Cook, R. J., & Wagner, C. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3631−3634.
Cook, R. J., Horne, D. W., & Wagner, C. (1989) *J. Nutr. 119*, 612−617.
Johnson, C. K. (1965) *ORTEP: A Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations*, Oak Ridge National Laboratory, Oak Ridge, TN.
Jones, T. A. (1985) *Methods Enzymol. 115*, 157−171.
Kerr, S. J. (1972) *J. Biol. Chem. 247*, 4248−4252.
Konishi, K., & Fujioka, M. (1987) *Biochemistry 26*, 8496−8502.
Konishi, K., & Fujioka, M. (1988) *J. Biol. Chem. 263*, 13381−13385.
Labahn, J., Granzin, J., Schluckebier, G., Robinson, D. P., Jack, W. E., Schildkraut, I., & Saenger, W. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10957−10961.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Monod, J., Changeux, J.-P., & Jacob, F. (1963) *J. Mol. Biol. 6*, 306−329.
Ogawa, H., & Fujioka, M. (1982a) *J. Biol. Chem. 257*, 3447−3452.
Ogawa, H., & Fujioka, M. (1982b) *Biochem. Biophys. Res. Commun. 108*, 227−232.
Ogawa, H., Konishi, K., Takata, Y., Nakashima, H., & Fujioka, M. (1987) *Eur. J. Biochem. 168*, 141−151.
Ogawa, H., Gomi, T., & Fujioka, M. (1993) *Comp. Biochem. Physiol. 106B*, 601−611.
Rafferty, J. M., Somers, W. S., Saint-Girons, I., & Phillips, S. E. V. (1989) *Nature 341*, 705−710.
Raha, A., Wagner, C., MacDonald, R. G., & Bresnick, E. (1994) *J. Biol. Chem. 269*, 5750−5756.
Raha, A., Joyce, T., Gusky, S., & Bresnick, E. (1995) *Arch. Biochem. Biophys. 322*, 395−404.
Reinisch, K. M., Chen, L., Verdine, G. L., & Lipscomb, W. N. (1995) *Cell 82*, 143−153.
Samudzi, C. T., Burton, L. E., & Rubin, J. R. (1991) *J. Biol. Chem. 226*, 791−797.
Schluckebier, G., O'Gara, M., Saenger, W., & Cheng, X. (1995) *J. Mol. Biol. 247*, 16−20.
Takusagawa, F. (1992) *J. Appl. Crystallogr. 25*, 26−30.
Takusagawa, F., Kamitori, S., Misaki, S., & Markham, G. D. (1996) *J. Mol. Biol. 271*, 136−147.
Tanaka, I., Yao, M., Suzuki, M., Hikichi, K., Matsumoto, T., Kozasa, M., & Katayama, C. (1990) *J. Appl. Crystallogr. 23*, 334−339.
Vidgren, J., Svensson, L. A., & Liljas, A. (1994) *Nature 368*, 354−356.
Wagner, C., Briggs, W. T., & Cook, R. J. (1985) *Biochem. Biophys. Res. Commun. 127*, 746−752.
Wang, B.-C. (1985) *Methods Enzymol. 115*, 90−112.

BI961068N