GENOMIC ANALYSIS OF HUMAN AND MOUSE GUANINE-7-METHYLTRANSFERASE WITH ACTIVE SITE CHARACTERIZATION

by

David James Bautz

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

BIOCHEMISTRY

Thomas O. Sitz, Chair

Shirley Luckhart

Glenda Gillaspy

May 10, 2001

Blacksburg, VA

Cap Structure, Methylation, Genomics, NEM

GENOMIC ANALYSIS OF HUMAN AND MOUSE GUANINE-7-METHYLTRANSFERASE WITH ACTIVE SITE CHARACTERIZATION

David James Bautz

(ABSTRACT)

The 5' end of eukaryotic and viral mRNAs contain a "cap" structure with the sequence m^7G(5')pppN(5'). The methylation of the 7-position on the guanine cap is very important to proper mRNA processing and initiation of translation. The enzyme responsible for this methylation, RNA guanine-7-methyltransferase, has been cloned and studied from a number of different species, including human, *X. laevis*, yeast, and *C. elegans*. The sequences for mouse guanine-7-methyltransferase cDNA and protein have been deduced based upon identity of mouse ESTs to the cDNA of the human enzyme. The deduced mouse cDNA encodes an ORF of 465 amino acids and is 76.4% identical to the human enzyme, or 86.5% within the C-terminal domain.

Active site characterization of mouse and human guanine-7-methyltransferase indicates a cysteine residue is important to proper enzyme activity. Enzyme activity was completely eliminated when N-ethylmaleimide (NEM) was added to the assay mixture. When the product of the reaction, S-adenosyl-L-homocysteine (SAH), was added at a concentration of 40μM the mouse enzyme retained 60% activity while enzyme isolated from Human Osteosarcoma (HOS) cells retained 100% of the original activity. SAH demonstrated no protective effects on the cloned human enzyme.

Factors that affect binding of RNA to the active site were also investigated. UV-cross-linking of RNA to the active site of the mouse enzyme was inhibited 35% by NEM. Cap analog, GpppG, at a concentration of 1mM, inhibited cross-linking, but the similar nucleotide GMP, at a concentration of 1mM, did not inhibit cross-linking. These analyses have given a clearer understanding of this very important enzyme.

I. ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Thomas O. Sitz, for giving me the opportunity to work in his laboratory and his advice, encouragement and good sense of humor that helped me immensely in both my undergraduate and graduate education. I would like to thank the other members of my committee, Dr. Shirley Luckhart and Dr. Glenda Gillaspy, for helping me with the genomic analysis and their advice and assistance. I would also like to thank Kim Kimbleton from the Virginia Tech Statistical Counseling Center for her assistance with the statistical analysis. Finally, I would like to thank my mother. Without her support, both financially and emotionally, nothing I have accomplished would have been possible.

II. TABLE OF CONTENTS

III. LIST OF FIGURES

IV. ABBREVIATIONS/SYMBOLS

ATP	Adenosine 5’-triphosphate
bp	Base pairs
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
cDNA	Complimentary deoxyribonucleic acid
CBC	Cap binding complex
CPM	Counts per minute
CPV	Cytoplasmic polyhedrosis virus
CTD	C-terminal domain
CTP	Cytosine 5’-triphosphate
DNA	Deoxyribonucleic acid
DTE	Dithioerythritol
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESTs	Expressed sequence tags
g	Force of gravity
GMP	Guanosine 3’-monophosphate
GMT	Guanine-7-methyltransferase
GpppG	Cap Analog
GST	Glutathione-S-transferase
GTP	Guanosine 5’-triphosphate
HEPES	N-2-hydroxyethylpiperazine-N’-2-ethane sulfonic acid
HOS	Human osteosarcoma
ICR	Institute for cancer research
IPTG	Isopropylthiogalactoside
kDa	Kilodaltons
MEM	Minimum essential media
MOPS	3-[N-morpholine]propanesulfonic acid
mRNA	Messenger ribonucleic acid
MW	Molecular weight
NEM	N-ethylmaleimide
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PMSF	Phenylmethylsulfonyl flouride
RNA	Ribonucleic acid
SAH	S-adenosyl-L-homocysteine

SAM	S-adenosyl-L-methionine
SDS	Sodium dodecyl sulfate
snRNA	Small nuclear ribonucleic acid
Tris	Tris-(hydroxymethyl)aminomethane
UTP	Uradine 5’-triphosphate

V. INTRODUCTION

An organism's genetic data is stored in the DNA. The processes inside a cell are carried out by proteins and to some extent RNA. In order for DNA to direct the production of proteins, mRNA must first be produced. Eukaryotes and prokaryotes differ in how their genetic information is relayed. In prokaryotes, most genes are transcribed in clusters, with the polycistronic mRNAs being translated as they are transcribed (1). In eukaryotes, RNA polymerase II first transcribes monocistronic mRNA as pre-mRNA. This pre-mRNA then goes through a series of post-transcriptional modifications, which includes 5'-terminal capping, cleavage, methylation, 3'-terminal polyadenylation, and splicing (2). These post-transcriptional modifications of mRNA do not occur in prokaryotes, indicating that mRNA synthesis in eukaryotes is more complex than in prokaryotes, which correlates with the increased genetic complexity of eukaryotes as compared to prokaryotes.

5.0 Discovery of the 5'-Terminal Cap Structure

The 5'-cap structure was first discovered in 1974 from studies on snRNAs derived from Novikoff hepatoma ascites cells. The primary sequence of these snRNAs revealed a blocked 5'-terminus with a hypermethylated guanosine residue ($m_3^{2,2,7}G$) linked to a 2'-OH methylated adenosine residue through a 5'-5' pyrophosphate linkage (3). The eventual elucidation of the 5'-terminal methylated cap structure was reported independently in reovirus (4), vaccinia virus (5) and cytoplasmic polyhedrosis virus

(CPV) (6). These reports were followed by a series of reports that demonstrated the existence of the cap structure in eukaryotic mRNAs (7-10).

5.1 The 5'-Terminal Cap

The general structural features of the 5'-cap of eukaryotic mRNAs, which is represented by m^7G(5')ppp(5')XpmYpm, are shown in Figure 1. Most cap structures contain at least the methyl group on the N7 position of the guanosine. This methylation results in the acquisition of a positive charge that partially neutralizes the phosphate groups in m^7GpppGX, which therefore has a net negative charge of –2.5 (11). Under alkaline conditions the imidazole ring is opened at the 8-9 bond, converting m^7G to the ring-opened derivative 2-amino-4-hydroxy-5-(N-methyl)-carboxamide-6-ribosylamino-pyrimidine, which lacks the positive charge.

While capped messages terminate in m^7G, the penultimate residue varies, both in identity and methylation state. Some mRNAs from animal, plant viruses, yeast, and slime mold contain m^7GpppGX, or the "cap 0" that is not methylated further. Other mRNAs have an additional methyl group on the 2'-OH position of the penultimate residue to form m^7GpppGXm, or the "cap 1" structure. In these cap structures the 3'-5' bond between the 2'-O-methylated penultimate nucleotide and the next residue in the mRNA chain is stabilized against hydrolysis by alkali or T2 RNase because formation of the 2'-3' cyclic intermediate required for cleavage is blocked. If the penultimate residue in cap 1 structures is adenosine, base methylation at the 6-position amine usually occurs (12). 2'-O-methylation of the second residue from the 5' end results in an even more extensively modified "cap 2" structure.

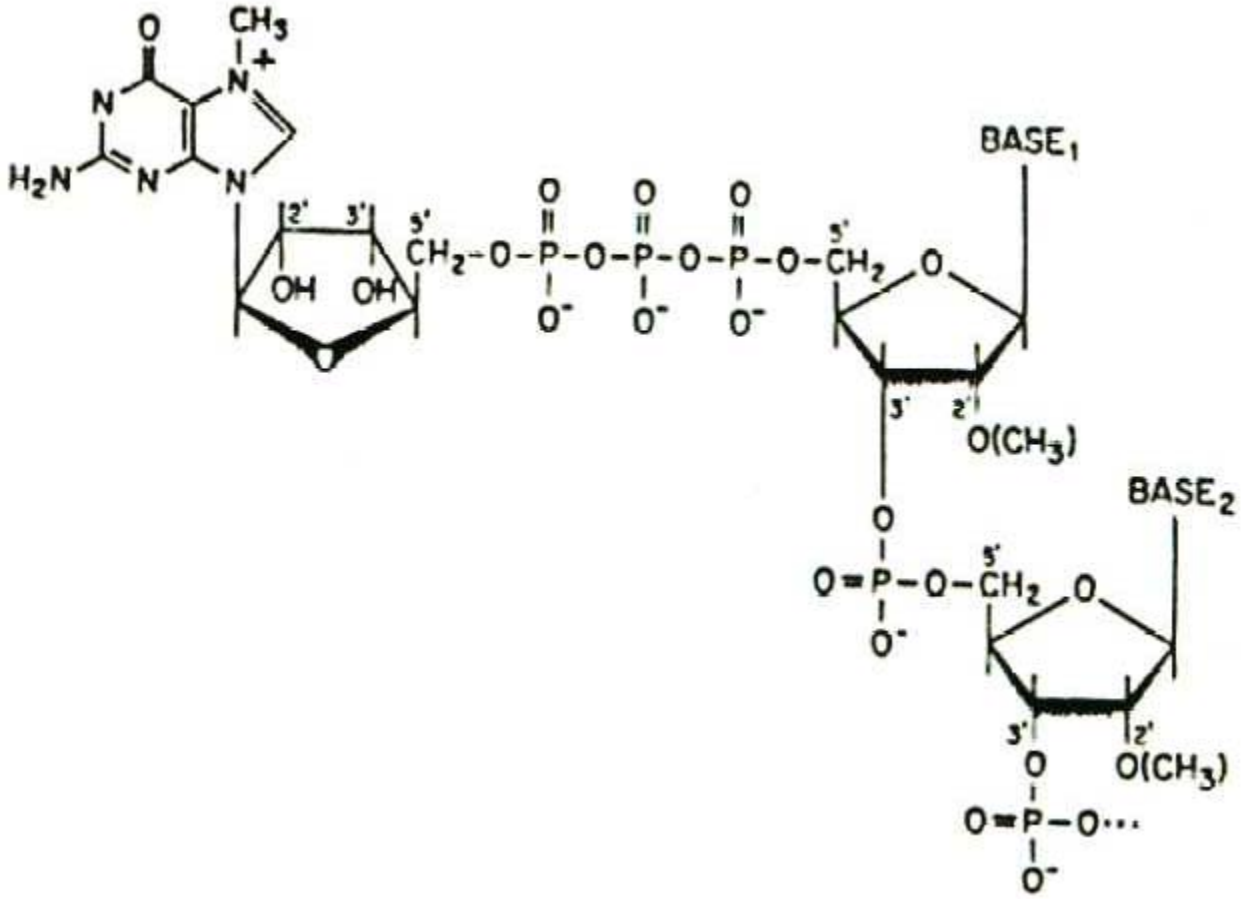


Figure 1: Messenger RNA 5'-cap structure

5.2 Enzymes of Cap Formation

Cap structure formation is catalyzed by a series of enzymes that include RNA triphosphatase, RNA guanylyltransferase, and RNA guanine-7-methyltransferase. A schematic of the enzymes and the reactions they catalyze is presented in Figure 2.

The first step in the mechanism of cap formation is the removal of the γ-phosphate from the triphosphate-terminated mRNAs (27). This is followed by addition of GMP to form the GpppX structure (28). This reaction proceeds through an enzyme-GMP intermediate, whereby GMP is linked to the ε-amino group of a lysine residue of the enzyme through a phosphoamide linkage. The next, and for some final, step of cap formation is the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to the N-7 position of the cap structure, generating the cap 0 structure m^7GpppX (19).

The first set of capping enzymes to be isolated and characterized were from vaccinia virus (13). Vaccinia virus replicates in the cytoplasm of eukaryotic cells and thus its capping apparatus can be isolated separately from the capping apparatus of the cell, which resides in the nucleus. The vaccinia capping enzyme is a heterodimer containing subunit polypeptides of 95 and 33 kDa, encoded by the D1 and D12 genes, respectively (14). The RNA triphosphatase and guanylyltransferase activities have been mapped to the amino-terminal 60 kDa of the D1 subunit while the GMT domain is located in the C-terminal portion of the D12 protein (15). The vaccinia virus is unique in that functional capping and methylation activities are never dissociable from each other, whereas in eukaryotic nuclear extracts the two activities readily separate during chromatography (16).

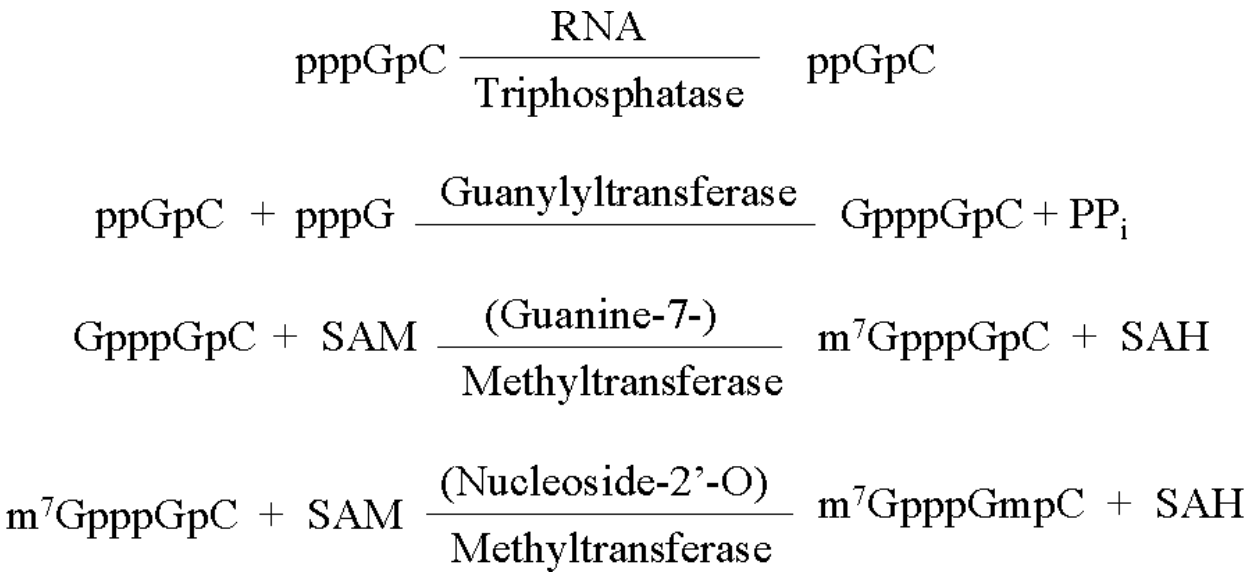


Figure 2: Pathway of 5'-cap Formation

The capping enzymes from *S. cerevisiae* have all been isolated and are the best-characterized (17). The yeast capping enzyme is a bifunctional complex consisting of two polypeptides of 80 and 52 kDa (18). The RNA triphosphatase activity is intrinsic to the 80-kDa subunit, encoded by the Cet1p gene, while the 52-kDa subunit is the guanylyltransferase, which is encoded by the Ceg1p gene. Yeast GMT is a 50-kDa protein encoded by the ABD1 gene (19). Yeast differ from mammals in that separate triphosphatase and guanylyltransferase polypeptides interact to form a heterodimer complex while in mammals the triphosphatase and guanylyltransferase domains are linked within a single polypeptide, which in humans is encoded by the Mce1p gene (20). In humans, the GMT cDNA has been cloned, and is shown to code for a 53-kDa protein that shows identity with other known GMTs (21).

5.3 Mechanism of Cap Formation

In order for mRNAs to be capped, the capping enzymes must be recruited to the pre-mRNA. McCracken *et al.* (22) showed that capping enzymes are bound to the C-terminal domain (CTD) of RNA polymerase II, but only to the hyperphosphorylated form. The CTD of RNA polymerase II is completely conserved from yeast to human (23) and the hyperphosphorylated CTD is associated with the transition of the transcription complex to initiate elongation (24). The cloned human GMT was shown to associate with both the human guanylyltransferase and the CTD of RNA polymerase II (21). This binding of the capping enzymes to RNA polymerase II helps explain why transcripts of RNA polymerase I and III are not capped, while transcripts of RNA polymerase II are capped.

The timing of capping was examined by ribonuclease footprinting of nascent mRNA within ternary complexes of vaccinia RNA polymerase (25). This study revealed an 18-nucleotide long RNA binding domain on RNA Polymerase II and also showed capping of mRNA chains occurred at or before they reached 31 nucleotides in length. This was supported by work on *D. melanogaster* showing that mRNAs coding for heat shock proteins were capped after about 20-30 nucleotides had been transcribed (26).

5.4 Biological Functions of the Cap Structure

The functions of the cap and its importance have been studied extensively. Virus models were tested to examine the essential role of the cap in translation (29-31). Both *et al.* (29) first showed this effect with reovirus mRNA, which required 5'-terminal m^7G in caps for efficient translation in a wheat germ cell-fee, protein synthesis system. Rose (30) observed that *in vivo*, all polysome-associated viral RNA was capped and the small portion of viral mRNA not directing protein formation contained uncapped, 5'-triphosphate ends. The selection by ribosomes of capped over uncapped messenger molecules was demonstrated in reovirus RNA molecules containing a mixture of pppGm and m^7GpppGm ending RNAs (31). They also showed that removing the 5'-terminal m^7G was accompanied by a marked decrease in the ability to direct protein synthesis in wheat germ extract. The cap structure was also shown to facilitate ribosome binding during the initial phase of translation through binding of cap binding proteins (CBPs), most notably eIF-4E (32). The crystal structure of eIF-4E with m^7GDP bound bas been resolved (33). The crystal structure shows the m^7GDP stacked between two tryptophan residues, which were experimentally determined to be necessary for enzyme function.

Adams *et al.* (34) found that 7-methyl GDP, 7-ethyl GDP and 7-benzyl GDP, but not GDP, decreased stable translational initiation complex formation by 80%. This showed that 7-substitution of some form was important. Elimination of the positive charge on the imidazole of the 7-substituted compounds destroyed their inhibitory activity (34). Darzynkiewicz *et al.* (32) did observe that there was a size limit to the N-7 alkyl substituent in order to have the cap structure functional. Any N-7 alkyl or alicyclic substituents larger than ethyl, such as 7-propyl, 7-isopropyl, 7-butyl, 7-isobutyl, 7-cyclopentyl or 7-benzyl, significantly decreased the inhibitory activity of these cap analogs. This was most likely due to a decreased affinity toward the cap binding proteins.

Splicing of pre-mRNAs and transport of them from the nucleus seem to be at least enhanced by the cap structure. Splicing of precursor mRNAs has been shown to only be efficient when they possess a cap structure in a HeLa nuclear extract (35). It was also reported that splicing was highly sensitive to inhibition by cap analogs. However, the cap structure is not absolutely essential to splicing, especially in the splicing of introns other then the cap-proximal intron (36). Hamm and Mattaj (38) reported that snRNAs transcribed by RNA polymerase II were exported into the cytoplasm, while snRNAs transcribed by RNA polymerase III remained in the nucleus. Export of RNA polymerase II transcribed snRNAs was inhibited by cap analog m^7GpppG. It has since been reported that two proteins of 80 and 26 kDa comprise the cap binding complex (CBC) and are responsible for mediating pre-mRNA processing and nuclear export (40).

mRNA stability is also aided by the cap structure (37). Reovirus mRNAs with m^7GpppGm or GpppG at the 5' terminus were shown to be more stable than mRNA containing 5'-ppG ends in *X. laevis* oocytes. In yeast, the capping enzyme is essential to mRNA stability (39). By isolating a conditional mutant of Ceg1p (capping enzyme), it was shown that at the restrictive temperature accumulation of mRNAs was prevented by decreasing their half-lives. In yeast lacking the 5' ribonuclease Xrn1, uncapped poly(A)$^+$ mRNA accumulation was detected. These findings suggest that the 5' cap is responsible for blocking 5' ribonuclease digestion and thus leading to longer mRNA half-lives.

5.5 Guanine-7-methyltransferase

The GMT from mouse Earlich ascites tumor cells has been isolated and characterized (41). The enzyme was purified 270-fold with a 31% yield and was unstable at low protein concentration; storage buffer with 50% glycerol increased stability. From a Superose 6 gel filtration column, an apparent molecular weight of 95 kDa was obtained for the enzyme. By SDS-PAGE, a major protein band at 46 kDa was detected, suggesting that the native protein is a dimer. Yeast GMT is a 50-kDa protein (19) while the cloned human GMT is 53-kDa (21).

The kinetics of the GMT catalyzed reaction were also studied by Bu (41). The enzyme displays a sequential random mechanism. The sequential nature of the mechanism requires that both substrates, SAM and capped unmethylated RNA, be present simultaneously in the enzyme active site. The random mechanism implies that binding of substrate to the enzyme (a) occurs at a different site for each substrate, (b) is

independent of the binding of the other substrate and (c) is freely reversible prior to a reaction involving the triple complex of enzyme, RNA and SAM.

The K_m value for SAM is 0.47μM, which is low compared to other methyltransferases (42). This low K_m value gives the GMT top priority *in vivo* among all the SAM-dependent methyltransferases. The K_m value for a 40-mer RNA substrate was found to be 0.27μM, which is also a low value compared to the K_m values for the RNA substrates for other RNA methyltransferases. The K_i value for S-adenosyl-L-homocysteine (SAH) was 0.37μM while the K_i value for the RNA substrate analog pppG-RNA, a dead-end inhibitor, was 4.1μM.

The chainlength requirement on RNA substrates was also examined (41). It was shown that the cap analogue, GpppG, was a poor substrate for the enzyme, with a V_{max}/K_m of only 0.47ml/min. Adding a phosphate to the 3' end, to generate GpppGp, resulted in a 13-fold increase in binding of RNA to the enzyme. Addition of another nucleoside at the 3' end did not change V_{max}/K_m. However, addition of another phosphate, to now generate GpppGpCp, resulted in a 140-fold increase in binding over just GpppG. This suggests that the phosphate groups are responsible for the proper binding of RNA to the enzyme active site, i.e. that the negative charges on the RNA are interacting with positively charged amino acids in the active site.

While some RNA modification enzymes require magnesium divalent ion for enzyme activity, the GMT was actually inactivated by magnesium at a concentration as low as 5mM (41). Two other divalent ions, calcium and manganese, were also shown to inhibit the enzyme. It is possible that the divalent ions may interact with the

triphosphate group of the RNA cap structure, thus neutralizing the negative charges on the phosphate groups, which are important for RNA binding to the enzyme (41).

Based upon the aforementioned data, a model of the active site has been proposed (41). This model is shown in Figure 3. In it, the SAM and RNA binding domains are separate, however still close enough to one other to allow transfer of the methyl group from SAM to the N-7 position of the capped RNA. There are also numerous positive residues indicated near the cap and RNA binding domains that facilitate RNA binding to the enzyme, as mentioned previously.

5.6 BLAST and Sequence Similarity Searches

In 1990 it was reported that a new approach to rapid sequence comparison had been created, the basic local alignment search tool, or BLAST (43). The basic algorithm used was simple and robust. It could be applied in a variety of contexts including DNA and protein sequence database searches, motif searches, gene identification searches, and in the analysis of multiple regions of similarity in long DNA sequences (43).

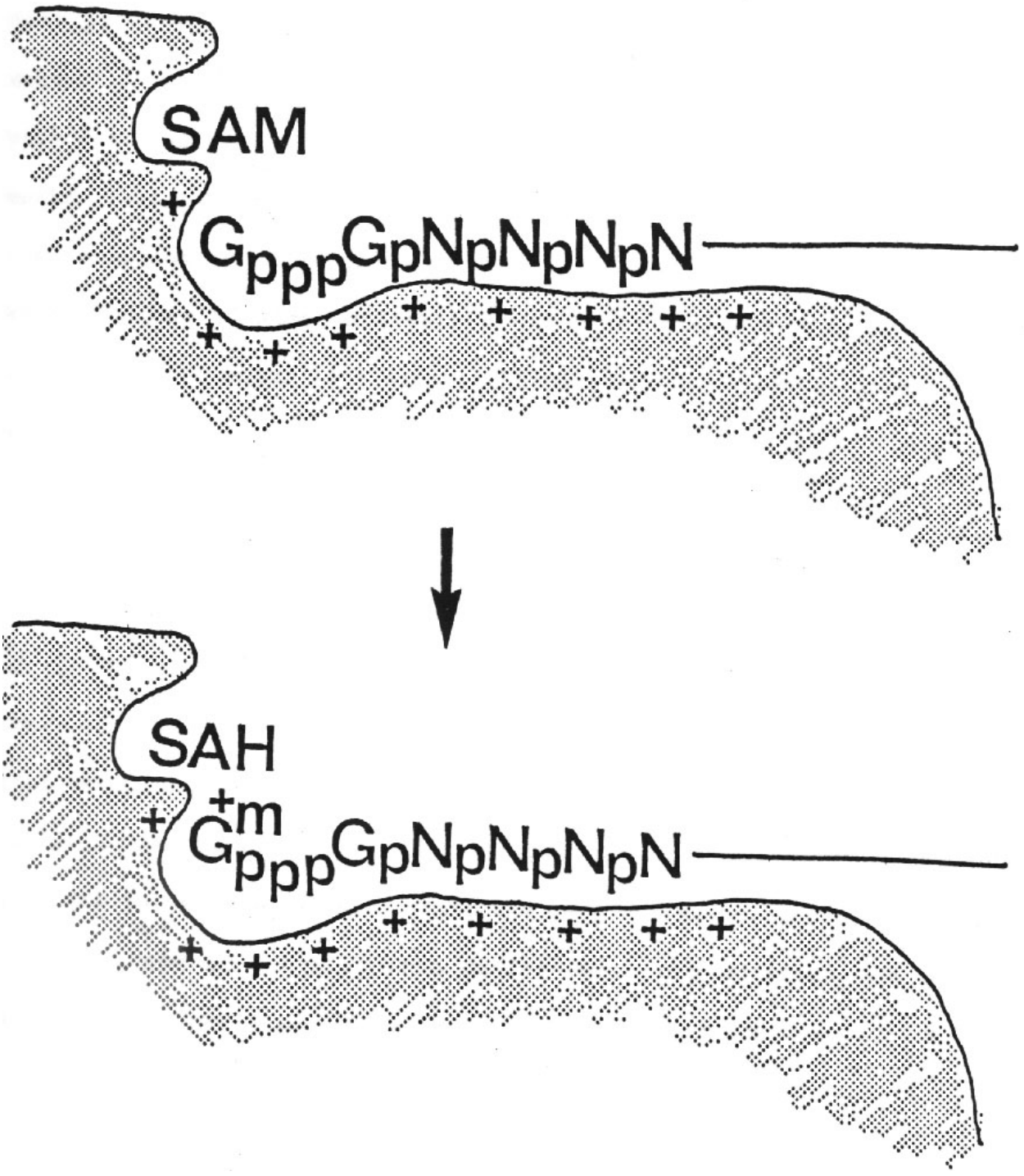


Figure 3: Proposed Model for GMT Active Site

5.7 Sequence Motifs in SAM Dependent Methyltransferases

Kagan and Clarke (45) used multiple sequence alignments to identify three sequence motifs common to a broad group of SAM-dependent methyltransferases. This broad group includes DNA, RNA and tRNA methyltransferases. Motif I, (V/I/L)(L/V)(D/E)(V/I)G(G/C)G(T/P)G, has been identified in GMTs from yeast (46), human (21), *C. elegans* (47), *D. melanogaster* (21), *X. laevis* (48), *C. albicans* (49), and *A. thaliana* (50). Motif II is not observed in RNA methyltransferases and, therefore, is not found in the previously mentioned enzymes. Motif III, LL(R/K)PGG(R/I/L)(L/I)(L/F/I/V)(I/L), is less well conserved, however the central glycines are completely conserved in all the aforementioned enzymes.

Motif I has been shown to be the probable SAM binding domain (53). The crystal structure of the HhaI DNA methyltransferase complexed with SAM has been resolved. It shows SAM binding into a pocket, part of which is made up by Motif I, with the general form for DNA methyltransferases being hh(D/E)hGXGXG, where “h” is a hydrophobic residue (53).

5.8 Essential Amino Acids in GMT

Studies in *S. cerevisiae* have been undertaken to determine essential amino acids in the GMT, especially in and around the conserved motifs (46,51,52). Mao *et al.* (46) identified Gly-174 and Asp-178 as essential residues. Both of these residues lie in Motif I. Asp 178 is completely conserved in cap methyltransferases but not in other methyltransferases (51). The other essential amino acids, as determined by the three mentioned papers, are denoted in Figure 4. Interestingly, in both yeast and humans,

the N-terminal region of ≈ 120 amino acids of the GMT was dispensable; i.e., the enzyme still functioned without it (52). The N-terminal region shows extreme sequence divergence among the GMTs (21, 52).

Previous work in this laboratory (41, 54) has shown that the mouse GMT requires a reducing agent for activity, such as dithiothreitol (DTT) or dithioerythritol (DTE) and the enzyme is susceptible to inhibition by N-ethylmaleimide (NEM). NEM reacts with cysteine residues to form a stable adduct (55, Figure 5). Cysteine residues were shown to be essential to enzyme activity in two bacterial DNA methyltransferases and a CheR methyltransferase (53,56,57). These data led to the idea that the cysteine in Motif I may be essential to enzyme activity. However, Schwer *et al.* (58) have shown that mutation of the cysteine in Motif I to an alanine in yeast GMT did not interfere with enzyme activity.

5.9 Thesis Objective

The sequences for many GMTs have been elucidated (21,46-50). However, the sequence of the mouse GMT has not been determined. One of the objectives of this study was to construct a mouse GMT sequence based upon alignment of mouse expressed sequence tags (ESTs) to the human GMT cDNA.

NEM is known to inactivate the GMT, indicating a cysteine may be either essential for enzyme activity or near a residue that is. It has been suggested that a cysteine may be located in the SAM binding domain of the enzyme, however without an

Figure 4: Essential Amino Acids in Yeast ABD1 Protein

The yeast and human GMT sequences are shown aligned. Arrows above residues indicate those that are required for enzyme activity in the yeast GMT. Dashes denote gaps in the sequences. Motifs I and III are indicated by bold, underlined letters.

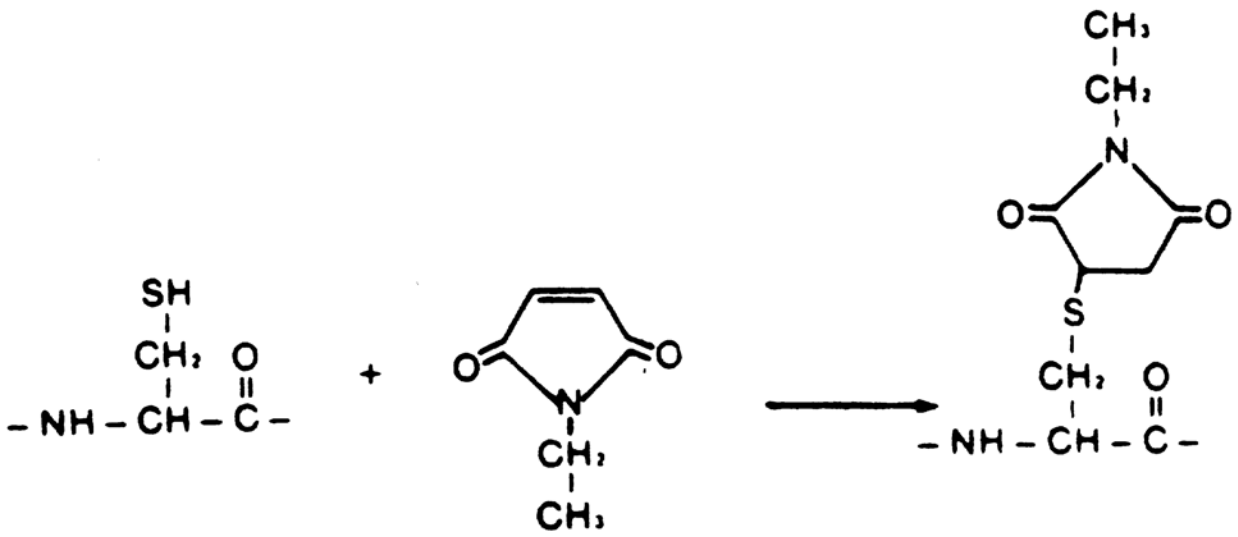


Figure 5: NEM and its Reaction With Cysteine Residues

X-ray structure of a GMT, this is speculative. Activity of the enzyme was evaluated in the presence of NEM with and without SAH to determine if SAH could protect the activity of the enzyme. With these assays, we hoped to show that a cysteine is located in the SAM binding domain.

The factors that affect binding of RNA to the active site of the GMT are not well known. To characterize these factors, radiolabeled RNA was cross-linked to the active site of the enzyme in the presence of NEM, cap analog, SAH, and GMP. With these experiments, we hoped to show how binding of RNA to the active site of GMT can be altered.

VI. EXPERIMENTAL PROCEDURES

6.0 Materials

Tris-(hydroxymethyl)aminomethane (TRIS), 3-[N-morpholino]propanesulfonic acid (MOPS), N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES), ethylenediaminetetraacetic acid, disodium salt (EDTA), ampicillin, chloramphenicol, phenylmethylsulfonyl fluoride (PMSF), Triton-X-100, formamide, DEAE Sephadex, glutathione, GMP, $MgCl_2$, bis-acrylamide, 8-OH-quinoline, S-adenosyl-L-homocysteine (SAH), N-ethylmaleimide (NEM), L-glutamine, Fetal Bovine Serum (FBS), and Trypsin/EDTA were purchased from Sigma. Glycine, glycerin, potassium chloride, sodium chloride, potassium phosphate, glucose, sodium carbonate, LB, agarose, acetic acid, hydrochloric acid, sodium hydroxide, sodium acetate, phenol, potassium acetate, urea, and citric acid were purchased from Fisher. Bovine Serum Albumin (BSA), cap analog (GpppG), T7 RNA Polymerase and buffer, polynucleotide kinase and RNA ligase were purchased from Ambion. Minimum Essential Media (MEM) with Earle's salts and without L-glutamine, and gentamycin were purchased from GibcoBRL. Ammonium persulfate, Quantum Prep Plasmid Miniprep Kit and SDS MW standards were purchased from Bio-Rad. ATP, CTP, UTP, GTP, BamHI, and EcoRI were purchased from Promega. Isopropylthiogalactoside (IPTG), dithioerythritol (DTE), and dithiothreitol (DTT) were purchased from Indofine Chemical Company. Glutathione Sepharose was purchased from Pharmacia Biotech. Sodium dodecyl sulfate (SDS) was purchased from Pierce. Cresol was purchased from Aldrich. S-[^{3}H-methyl]-adenosyl-L-methionine (SAM, specific activity 255μCi/455μl) was purchased from New England Nuclear. ^{32}P-ATP was a gift from Dr. Peter Kennelly and was purchased from New England Nuclear.

95% ethanol was purchased from AAPER Alcohol and Chemical Company. Coomassie Brilliant Blue R-250 was purchased from Kodak. Acrylamide was purchased from Schwartz/Mann Biotech. Ecolume was purchased from ICN Biomedicals.

6.1 Methods

6.1A BLAST sequence homology searches

All of the sequence identity work was performed using BLAST 2.1 at http://www.ncbi.nlm.nih.gov/BLAST. All scores greater than 100 were assumed to be significant while all scores less than 100 were assumed to be insignificant.

6.1B Deduction of mouse GMT cDNA and protein sequence from ESTs

Mouse expressed sequence tags (ESTs) with identity to the human GMT cDNA (Accession XM_012722) were analyzed with the Open Reading Frame (ORF) finder at http://www.ncbi.nlm.nih.gov/gorf/gorf.html. All possible reading frames were manually compared to the human GMT amino acid sequence. ESTs with ORFs homologous to the human protein were compared to one another using the BLAST algorithm at http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html. Based upon EST overlaps, the deduced sequence of the mouse GMT cDNA was constructed. Six ORFs for this deduced mouse sequence were examined. Each amino acid sequence was manually aligned against the human GMT amino acid sequence. The sequence showing the highest degree of identity with the human protein sequence was considered to be the deduced mouse protein sequence.

6.1C Identification of conserved motifs in the mouse protein sequence

The deduced mouse GMT sequence was aligned against known GMT sequences using ClustalW (44) at Biology Workbench 3.2 website, http://workbench.sdsc.edu.

6.1D Growth and maintenance of Ehrlich ascites cells

Ehrlich ascites tumor cells were grown in the peritoneal cavity of ICR mice (12-14 weeks old, approximately 40 grams). The mice were injected with approximately 4 X 10^7 cells (0.1ml packed cell volume) per animal in 1ml GKN buffer (10mM Hepes, pH 7.4, 0.1% glucose, 0.04% KCl, and 0.8% NaCl). After one week, 3-4ml of packed cells per mouse were harvested in ice-cold saline solution (0.9% NaCl) (41).

6.1E Growth and maintenance of human osteocarcoma cells

Human osteosarcoma cells (HOS) were added to 10ml of Minimum Essential Media (MEM; Earle's salts, 25μg/ml gentamycin, 1mM glutamine, and 1% $NaHCO_3$) in a $25cm^2$ plate. After allowing the cells to adhere for 4 hr, the medium was replaced. When the cells reached confluence they were split 1:10 as follows. The medium was poured off and 10ml of phosphate buffer saline (PBS) was added to wash the cells. The cells were then incubated in 5ml of trypsin/EDTA for 5 min at 37°C to release them from the culture plate. 5ml of MEM was added to stop the trypsin reaction.

6.1F Transformation of *E. coli*

An aliquot of competent BL21 (DE3) *E. coli* cells were thawed on ice for transformation. Dr. Aaron Shatkin, from Rutgers University, generously supplied a pGex-4T-3 plasmid

containing the coding sequence for the human GMT. Two μl of plasmid (0.5ng/μl) was added to the cells. After incubating in an ice bath for 30 min, the cells were heat shocked for 2 min at 42°C. Ice-cold LB broth (1ml) was added to the cells and they were incubated at 37°C for 60 min prior to spreading 200μl on an agar plate containing ampicillin (100μg/ml). The plates were placed in 37°C to allow for colony growth.

6.1G Plasmid purification

Plasmid was isolated using Quantum Prep Plasmid Miniprep Kit (#732-6100).

6.1H Sonication

All sonications were performed with a Fisher Sonic Dismembrator (Model 300). The cycle time was five seconds at 60% for a total of 3 min.

6.1I Induction and isolation of cloned human GMT

Transformed *E. coli* were streaked on an agar plate and allowed to grow overnight at 37°C. A single colony was selected and inoculated into 5ml of LB broth with 34μg/ml of chloramphenicol and 100μg/ml of ampicillin and incubated at 37°C overnight. Two ml of this culture was used to inoculate 200ml of LB broth with 34μg/ml of chloramphenicol and 100μg/ml of ampicillin. When the A_{600} was 0.6-0.8, expression of the protein was induced with 1mM isopropylthiogalactoside (IPTG) for six hours at 30°C. Induced cells were harvested and centrifuged at 2,000 x *g* (3,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 10 min. The pellet of cells was washed, resuspended in 10ml of PBS and centrifuged at 2,000 x *g* (3,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 10 min.

The pellet was suspended in 10ml of PBS-2 (PBS with 1mM DTT and Sigma protease inhibitor mix), disrupted by sonication, and centrifuged for 20 min at 20,000 x *g* (15,000 rpm, Sorvall RC-2B with a SS-34 rotor). A glutathione affinity column was prepared by adding 0.7ml of glutathione sepharose 4B to a 0.7 x 5.0 cm column and equilibrated with 10ml of PBS to equilibrate it. The supernatant solution was filtered through a Millex-GV Filter Unit with a 0.22μm pore size and applied to the column. The column was washed with 10ml of PBS-2 to remove any proteins with low affinity for glutathione. The recombinant human methyltransferase (GT-hMT) was eluted with 3ml of 10mM glutathione and stored for further use at -80°C.

6.1J Isolation of nuclei from Ehrlich ascites cells

Ehrlich ascites cells harvested from mice were pelleted at 1,200 x *g* (2,000 rpm, Beckman J-6B) for 3 min and the total weight of the cells was determined. The pelleted cells were suspended in buffer-1 (10mM Tris-HCl, pH 7.6, 10mM NaCl, 5mM $MgCl_2$, and 1mM DTE) to lyse the contaminating red blood cells, and then pelleted at 500 x *g* (1,400 rpm, Beckman J-6B). The tumor cells were pelleted at 1,200 x *g* (2,000 rpm, Beckman J-6B) for 5 min and resuspended in buffer-1 (+0.2% NP-40) at a concentration of 4.5 ml buffer per ml packed cell volume. The cells were homogenized using a motor-driven teflon pestle (1,000 rpm) and a glass Thomas type C homogenizer. The degree of cell breakage was monitored with a phase-contrast microscope. Usually about 60-80 strokes of the pestle were required to break the Ehrlich ascites cells without damaging the nuclei. The homogenate was centrifuged at 1,200 x *g* (2,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 10 min to pellet the nuclei (41).

6.1K Isolation of nuclei from HOS cells

After the HOS cells had reached confluence in a 150 cm^2 dish the cells were harvested using trypsin/EDTA, as described in section 6.1E, to detach the cells. Ten ml of MEM from each dish, with trypsin/EDTA and detached cells, was transferred to a 250ml centrifuge tube and centrifuged for 10 min at 1,200 x *g* (2,000 rpm, Beckman J-6B with a JS-4.2 rotor). The supernate was poured off and the cell pellet was resuspended in 10ml of PBS and transferred to a 15ml tube. The cells were then centrifuged at 1,200 x *g* (2,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 10 min. The mass of the pellet was then determined. Homogenizing buffer (10mM Tris-HCl, pH 7.6, 10mM NaCl, 5mM $MgCl_2$, 1% Triton-X-100, 1mM PMSF, 1mM DTE) was added in a 10:1 w/v ratio. The cells were homogenized with 60-70 strokes; homogenization was considered complete when no intact cells were observed. The homogenate was then centrifuged at 1,200 x *g* (2,000 rpm, Beckman J-6B with a JS-4.2 rotor) to collect the nuclei.

6.1L Purification of GMT from HOS or Ehrich ascites nuclei

The pellet of nuclei from HOS or Ehrlich ascites cells was resuspended in 10 vol of sonication buffer (10mM Tris-HCl, pH 8.0, 10mM KCl, 2mM EDTA, 1mM PMSF, 1mM DTE) and sonicated as described above. The sonicate was centrifuged at 20,000 x *g* (15,000 rpm, Beckman J-21B with an SS-34 rotor) for 15-20 min. The supernate was then centrifuged at 100,000 x *g* (31,000 rpm, in a Beckman L7-55 ultracentrifuge with a 60 Ti rotor) for 2 hr to pellet nuclear membrane debris and adjusted to 0.2M NaCl from a 5M NaCl stock solution. A DEAE-Sephadex column (4 x 0.5cm) was poured and pre-equilibrated with sonication buffer and the supernate was applied to the column. The

material was forced through the column by applying pressure with a syringe and the flow through was collected as a single fraction. This fraction was then dialyzed overnight against CM-buffer (10mM MOPS pH 7.0, 10mM KCl, 2mM EDTA, 1mM DTT) at 4°C. After dialysis, denatured protein was removed by centrifugation at 20,000 x *g* (15,000 rpm, Beckman J-21B with a SS-34 rotor) for 20 min. A CM-Sepharose column (4 x 0.5 cm) was equilibrated with 20ml of CM-buffer. The supernatant material was applied to the CM-Sepharose column and washed with buffer until the A_{280} of the wash returned to the background level. The enzyme was eluted from the column with 0.3M KCl in CM-buffer and 1ml fractions were collected. The absorbance at 280nm was used to determine which fractions had the highest protein concentrations and those fractions were pooled.

Two ml of the HOS enzyme solution were concentrated in a centricon-30 (Micron) at 3,000 x *g* (4,000 rpm, Beckman J-21B with a SS-34 rotor). This process was repeated until approximately 0.9ml of HOS enzyme solution remained.

A Mono-Q anion exchange column (0.5 x 7 cm) with FPLC system at room temperature was used to further purify the Ehrlich ascites enzyme. The column was washed in sequence with 1M NaCl, 1M NaOH and 70% acetic acid to remove residual materials bound to the column. The column was then washed with hypotonic buffer-2 (10mM Tris-HCl, pH 8.0, 10mM KCl, 1mM DTE, 2mM EDTA, 1mM freshly added PMSF, and 10% v/v glycerol) containing 0.3M KCl, followed with hypotonic buffer-2 containing 0.05M KCl. The CM-Sepharose-purified enzyme was diluted with 3 vol of hypotonic buffer-2 applied to the Mono-Q column. The column was washed with the starting buffer at a flow rate of 1ml/min until the A_{280} returned to background level. A 30ml linear

salt gradient (0.05-0.3M KCl) in hypotonic buffer-2 was used to elute the protein in 1ml fractions. The fractions were assayed for GMT activity and two peaks of GMT activity were pooled.

6.1M Activity assay

For activity assays of Ehrlich ascites GMT, the standard assay at 37°C consisted of 25μl of enzyme preparation and 25μl of "assay mix" in a final volume of 50μl. The "assay mix" was 2mM DTT or DTE, 0.32μg/μl BSA, 9ng/μl RNA, 1.8μCi ^{3}H-SAM, and 50mM KCl. Aliquots of 10μl were spotted on 1.5 x 1.5 cm DE-81 filter paper at 0,10, 20, and 30 min after the start of the assay. The filter papers were washed twice in 250ml H_2O for 7 min, and then twice in 250ml of 0.05% K_2HPO_4 for 7 min, blotted on paper and transferred to scintillation vials where 0.5ml NaCl (1M) and 4ml of scintillation fluid (Ecolume) were added. The samples were counted in a Beckman Model LS 5801 for 2 min.

For activity assays of recombinant human GMT the enzyme was first diluted into hypotonic buffer-1 for two separate assays. For low salt assays, 55μl of cloned human enzyme was added to 176μl of hypotonic buffer-1 and 0.64μg/μl of BSA in a final volume of 275μl. For high salt assays, 55μl of cloned human enzyme was added to 176μl of hypotonic buffer-1, 0.64μg/μl of BSA and 127mM KCl in a final volume of 275μl. The enzyme was then assayed as described in the previous paragraph.

6.1N Active site protection experiments

In these experiments the enzyme was pre-treated as follows prior to the assay. For the Ehrlich ascites sample, 100μl of enzyme was adjusted to the appropriate concentration of SAH and incubated for 10 min at 37°C. The solution was then adjusted to 5mM NEM and incubated for 1 minute. Five mM NEM provided an excess to react with both the enzyme and the DTT or DTE present in the enzyme solution. Finally, the solution was adjusted to 9mM DTT to inactivate remaining NEM. The solutions were then dialyzed against 250ml of hypotonic buffer-1 containing 1mM DTT or DTE overnight at 4°C with a buffer exchange every 5 or 6 hr. The dialyzed samples were then assayed as described in section 6.1M.

For the cloned human GMT 50μl of diluted enzyme (as described in section 6.1M) was used and the samples were dialyzed for 2 hr at room temperature, with a buffer exchange every hr. The dialyzed samples were then assayed as described in section 2.1M.

6.1P *In vitro* transcription

A 0.5ml transcription reaction consisted of 40mM Tris-HCl pH 8.0, 1mM spermidine-HCl, 50μg/ml BSA, 6mM $MgCl_2$, 5mM DTT, 0.05% Tween-20, 0.5mM each of ATP, CTP, UTP, and 0.05mM GTP, 0.5mM cap analog (GpppG), 150μl 18/38-mer DNA mixture, and 1000 units of T7 RNA polymerase (41). After incubation for 4 hr at 37°C, 1ml of transcription mixture was mixed with 1ml of SDS buffer (0.14M NaCl, 0.05M NaOAc, 0.3% SDS, pH 5.1). An equivalent amount of phenol-cresol (1625ml phenol, 325ml cresol, and 2g 8-OH quinoline in 2L) was added and the solution was incubated

at 37°C for 15 min with rigorous shaking every 5 min. The solution was then centrifuged at 2,000 x *g* (3,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 15 min. Six vol of 95% ethanol/2% potassium acetate was added to the aqueous layer and stored at -20°C overnight. The RNA was pelleted by centrifugation at 2,000 x *g* (3,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 20 min, dissolved in autoclaved H_2O, and precipitated with 5 vol of 95% ethanol/2% potassium acetate. The procedure was usually repeated an additional time. The final RNA pellet was washed with 95% ethanol and air dried.

The final RNA pellet was then dissolved in 150μl of 7M urea/0.025 M citric acid and incubated for 2 min at 37°C. An aliquot of 20μl of the RNA sample was added to each well of a 20% acrylamide/7M urea gel which had been pre-electrophoresed overnight at 100mV. The gel was stained with methylene blue and de-stained to visualize the RNA bands. The bands of interest were cut out of the gel and homogenized using a motor-driven teflon pestle (1,000 rpm) and a glass Thomas type C homogenizer in 1.5ml of salt buffer-1 (0.15M NaCl, 1mM MOPS, pH 7.2, 1mM EDTA). The homogenized solution was frozen at -80°C for 15 min, thawed at 37°C for 5 min and mixed. This process was repeated twice. After the third thawing the solution was centrifuged at 2,000 x *g* (3,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 15 min. The supernate was poured off and 5 vol of 95% ethanol/2% potassium acetate was added. The solution was placed at -20°C overnight. The RNA was pelleted by centrifugation at 2,000 x *g* (3,000 rpm, Beckman J-6B with a JS-4.2 rotor) for 1 hour. The RNA pellet was dissolved in autoclaved H_2O and precipitated with 5 vol of 95% ethanol/2% potassium acetate. The final RNA pellet was washed with 95% ethanol, air dried, and

dissolved in 300μl of autoclaved H_2O. An absorbance reading at 260nm was used to estimate the amount of RNA present. The RNA solution was then adjusted to 0.1μg/μl.

6.1Q Synthesis of (5'-^{32}P) cytidine bisphosphate (*pCp)

A total of 0.3mCi of (γ-^{32}P) ATP was reacted with 0.4mM Cp, 5mM Tris pH 9.0, 1mM $MgCl_2$, 0.5mM DTT, 2.3μl of H_2O and 2μl of polynucleotide kinase in a final volume of 10μl for one hr at 37°C. Heating the solution at 95°C for 5 min stopped the reaction.

6.1R 3'-end labeling of RNA

Radiolabeling of the 3' end of the RNA was performed by the addition of (5'-^{32}P)pCp to the 3'-end of the RNA chain. One μg of RNA was dried and dissolved in 0.06M HEPES pH 7.5, 0.02M $MgCl_2$, 4mM DTT, 12ng/μl BSA, 31μM ATP, and 10 units of RNA ligase in a final volume of 20μl. The reaction was incubated overnight at 4°C. A total of 40μl of formamide was added to each reaction and 20μl was loaded into each well of a 20% acrylamide(1:19X)/7M urea gel.

6.1S 5'-end labeling of RNA

The 5'-end of the RNA was labeled in an exchange reaction with γ-^{32}P ATP by polynucleotide kinase. One μg of RNA was dried and dissolved in 5mM Tris pH 9.0, 1mM $MgCl_2$, 0.5mM DTT, and 20 units of polynucleotide kinase in a final volume of 10μl. The reaction was incubated for one hr at 37°C and stopped by the addition of 50μl of formamide. A total of 20μl of the reaction mixture was loaded in each well of a 20% acrylamide(1:19X)/7M urea gel.

6.1S Isolation of ^{32}P-labeled RNA

After electrophoresis the bands of interest were visualized by exposing the gel to Kodak X-Omat XAR-5 film. The bands were cut out of the gel and homogenized using a motor-driven teflon pestle (1,000 rpm) and a glass Thomas type C homogenizer as described in section 6.1J.

6.1U UV-cross-linking of RNA to the GMT

The ^{32}P-labeled RNA was cross-linked to the enzyme in the presence of short wavelength UV light (254nm). A reaction mixture of 50μl of Ehrlich ascites GMT, 1μg/μl of BSA, 50mM DTT and approximately 1,000,000 counts of ^{32}P-labeled RNA was dried, dissolved in 5μl of H_2O, and incubated for 10 min at 37°C. To cross-link the RNA to the enzyme, the reaction mixture was exposed on ice to an unfiltered UV lamp (Mineralight Lamp Model UVG-54) for 45 min. Five μl of 2x loading buffer (125mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 10% 2-mercaptoethanol) was added to each reaction mixture and the entire 10μl was added to a 10% acrylamide/2.7% bis-acrylamide resolving gel and 4% acrylamide/2.7% bis-acrylamide stacking gel. The mixture was resolved with 15mA of constant current until the tracking dye was approximately 1cm from the bottom tank buffer. The gels were then stained with Coomassie blue and dried onto 3mm Whatman filter paper. Analysis of NEM inhibition on cross-linking was examined by cutting the bands out of the dried gel. The bands were then counted in a scintillation counter (Beckman Model LS 5801) and compared to one another.

VII. RESULTS

7.0 Sequence Homology Searches

A BLAST search of the mouse EST databank using the human cDNA for the GMT (Accession XM_012722) revealed numerous mouse ESTs with significant identity to the human cDNA. A score was assigned to each EST by the BLAST program based upon the extent of the identity and the length of the sequence. Out of the homologous ESTs, six had scores of 155 (89% identity, length of 134 nucleotides), 174 (91% identity, length of 148 nucleotides), 436 (85% identity, length of 552 nucleotides), 396 (89% identity, length of 341 nucleotides), 182 (88% identity, length of 172 nucleotides) and 69.9 (90% identity, length of 55 nucleotides). The EST with score of 69.9 was selected because it showed evidence of a poly A tail and numerous other ESTs ended in the same sequence. When these six ESTs were aligned against one another overlaps among them were discovered. The overlaps of the six ESTs compared with the human cDNA sequence are shown in Figure 6. The overlaps were used to construct a predicted mouse cDNA sequence, which is shown in Figure 7. The predicted mouse GMT cDNA contained an ATG start codon, a TGA stop codon, and an AATAAA poly A signal.

The predicted cDNA was then put into an ORF finder algorithm, which revealed six different open reading frames (Fig. 8A). The +3 Frame (Fig. 8B) was selected for analysis based upon the fact it was the largest of the six, contained no internal stop codons, and showed the highest level of identity to the human GMT amino acid sequence.

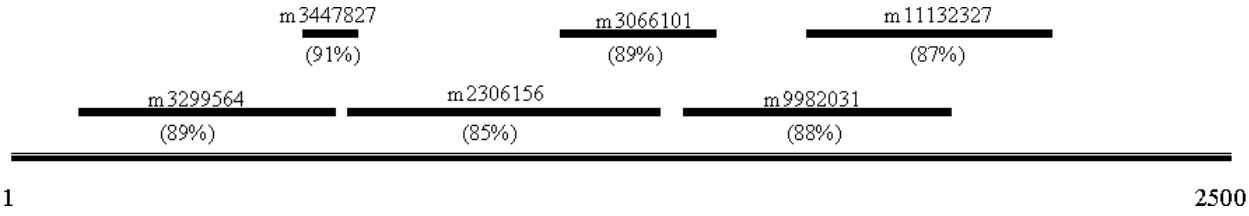


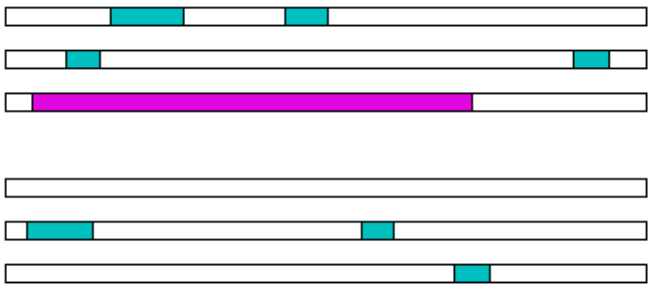
Figure 6: Mouse ESTs Aligned with the Human GMT cDNA
Base pairs 1-2500 of the human cDNA (Accession XM_012722) are shown aligned with the six mouse ESTs. The six mouse ESTs are referenced by their accession number, which is shown above each line. Indicated below each line are the % identities between the mouse ESTs and the human cDNA.

```
   1      TCCGGCCGGCGTTTCCGCAAACAGAATCGTGGATGATTGTGTGGTGTCGATTGATGAGGTTTTGGGGTCA      70
  71      GTTCAAGTAACGAGAAATGGAAGGTTCCGCGAAAGCGTCAGTCGCTTCTGACCCAGAGTCCCCTCCAGGT     140
 141      GGTAATGAGCCCGCTGCTGCCTCTGGGCAGAGGCTTCCAGAAAATACTCCTCCATGCCAGCAAGTAGATC     210
 211      AGCCAAAAATGCAGAAGGAGTTTGGGGAGGATCTTGTAGAGCAAAATTCGAGTTATGTGCAAGATTCTCC     280
 281      ATCCAAGAAAAGGAAACTTGATGTTGAAATCATCTTAGAGGAAAAGCACTCTGAGGATGATGGAGGTTCG     350
 351      GCAAAGAGAAGCAAACTGGAAAGAGGTGATGTGTCGGAAGATGAGCCTTCCCTTGGTAGGCTTAACCAGA     420
 421      CCAAGAGAAAACTGCAGCCTCAGGACGATGAGGTACCGCAGAAGTTGCAGAAGCTGGAGGAAGGACACAG     490
 491      CTCAGCCGTGGCTGCCCACTATAATGAGCTTCAGGAAGTTGGCCTGGCGAAGCGTAGTCAAAGTCGCATT     560
 561      TTTTACCTAAGAAACTTCAATAATTGGATTAAAAGTATTCTGATTGGGGAAATTTTAGAAAAGGTACGAC     630
 631      AAAGAAAAACCCGTGACATCACTGTTTTGGACCTGGGATGTGGAAAAGGTGGAGATTTGCTCAAGTCGAG     700
 701      GAAGGGGAGAATTAGCAGGCTAGTGTGTGCTGATATCGCTGATATTTCTATGAAACAATGTCAACAGCGT     770
 771      TACGAAGACATGAGATGTCGTCGCGATAATGAGCATATTTTTAGTGCAGAATTTATAACTGCTGACTGTT     840
 841      CAAAGGAACTTTTGGTTGAAAAATTCCGTGACCCAGAAATGTACTTTGATGTCTGCAGCTGCCAGTTTGC     910
 911      TTGCCATTACTCCTTTGAGTCCCAGGTGCAGGCGGATACGATGCTTAGAAATGCCTGCGGGAGGCTTAAC     980
 981      CCTGGCGGCTATTTTATTGGCACCACTCCCAATAGCTTTGAACTGATAAGACGCCTTGAAGCTTCAGAAA    1050
1051      CAGAATCATTTGGAAATGAAATATACACTGTGAAATTCCAGAAGAAAGGAAATTATCCTTTATTTGGCTG    1120
1121      CAAATATGACTTTAACTTGGAAGGTGTTGTGGATGTCCCTGAATTCTTGGTCTATTTTCCATTGCTAACT    1190
1191      GAAATGGCAAAGAAGTATAATATGAAACTAATCTACAAAAAAACATTTTTGGAATTCTATGAAGAAAAGA    1260
1261      TTAAAAACAATGAAAACAAAATGCTCTTAAAACGAATGCAGGCCCTGGAGCAATACCCTGCACATGAGAA    1330
1331      CTCTAAACTTGCCTCTGAAAAGGTGGGTGACTACACGCATGCAGCAGAGTACCTGAAGAAGAGCCAAGTG    1400
1401      AGGTTACCTCTGGGAACCTTAAGTAAGTCAGAATGGGAAGCTACAAGTATCTACTTGGTCTTTGCCTTTG    1470
1471      AGAAGCAGCAGTGAGCACGTTGCTTTAGAGCAGCACTCCATCCTCCCAGGTGGAGCAGACCATCTCAGAA    1540
1541      ACATTTGACAGCTGTTTGTTTATTTTAATAGTAAGTTCTCAATGTGTAGGATGCTGCCACAAACTTCAGT    1610
1611      GTATGAATTTGACACTTACTGTCTGTGACAGGTTAGCATAATGTGTGTACATAGGGATGAGTTGTCTTGA    1680
1681      AGATCTATTTTTAAGTACTGTTGTAATTGTTCCCCTCTACTGTCAAAACTCTAGCAAGGCATGTCAGAGC    1750
1751      AGCTGACCTCCCCAGTGCTGTGATGTGTGAGCAGCTGACCTCCCCAGTGCTGGGATGTGTGAATGTGTTT    1820
1821      ACAGAGCTGATTTGACAGTCGCTAGAATTGGCAGAGGAACGTTCACGTTTCCTTACAGTGTTTTTCAAGA    1890
1891      TTTAATGGATTGTGTTCTGTATACTTGAATGTAAGTGTTCAATATATATTGCTAAAGTTATAAGATTAAA    1960
1961      TTTTGATTTTAGATTGTCCTAAGAGCTACCTGGCTAGATTTTGACCGTGTTTCAATAAAATAACTATTTG    2030
2031      CTAAT
```

Figure 7: The Predicted Mouse GMT cDNA Sequence

The predicted mouse cDNA for GMT is shown. Indicated by underline and bold are the ATG start codon, the TGA stop codon and the AATAAA putative poly A tail signal sequence.

A.



Frame	from	to	Length
+3	87	1484	1398
+1	337	570	234
-2	73	282	210
+1	889	1026	138
-3	1428	1544	117
+2	1805	1918	114
+2	194	301	108
-2	1132	1236	105

B.

```
MEGSAKASVASDPESPPGGNEPAAASGQRLPENTPPCQQVDQPKMQKEFGEDLVEQNSSYVQDSPSKKRK
LDVEIILEEKHSEDDGGSAKRSKLERGDVSEDEPSLGRLNQTKRKLQPQDDEVPQKLQKLEEGHSSAVAA
HYNELQEVGLAKRSQSRIFYLRNFNNWIKSILIGEILEKVRQRKTRDITVLDLGCGKGGDLLKSRKGRIS
RLVCADIADISMKQCQQRYEDMRCRRDNEHIFSAEFITADCSKELLVEKFRDPEMYFDVCSCQFACHYSF
ESQVQADTMLRNACGRLNPGGYFIGTTPNSFELIRRLEASETESFGNEIYTVKFQKKGNYPLFGCKYDFN
LEGVVDVPEFLVYFPLLTEMAKKYNMKLIYKKTFLEFYEEKIKNNENKMLLKRMQALEQYPAHENSKLAS
EKVGDYTHAAEYLKKSQVRLPLGTLSKSEWEATSIYLVFAFEKQQ
```

Figure 8: ORF Corresponding to Mouse GMT

ORFs for the deduced mouse GMT cDNA and the predicted protein sequence are shown. (**A**) Six ORFs were generated from the deduced mouse cDNA. Of these, the +3 ORF was selected based upon the fact that it contained no internal stop codons and was the largest. (**B**) The predicted mouse GMT protein sequence corresponding to the selected ORF.

An alignment was then performed using ClustalW between the predicted mouse GMT sequence and the human GMT sequence (Fig. 9). This alignment shows very little identity between the two sequences at the N-terminal region, amino acids 1-120. However, towards the C-terminal region, residues 121-465, there is much higher identity. Overall, the two sequences are 76.4% identical. The identity increases to 86.5% from residues 121-465.

7.1 Identification of conserved motifs

The predicted mouse GMT amino acid sequence was then aligned with other known GMT amino acid sequence (Fig. 10). This alignment reveals evidence of conserved sequences. The previously identified motif I, (V/I/L)(L/V)(D/E)(V/I)G(G/C)G(T/P)G, (43) can be located in all of the sequences. Other highly conserved residues include the two central glycines of motif III, LL(R/K)PGG(R/I/L)(L/I)(L/F/I/V)(I/L).

7.2 Active site protection experiments

NEM completely inactivates GMT. Capped RNA and SAH were used to protect activity in GMT. SAH protected the active site from inactivation by NEM (Fig. 11). This assay was performed with 40μM SAH and 18ng/μl of capped RNA. Samples containing SAH demonstrated some protection from inactivation by 5mM NEM. This established SAH as the compound that would give the highest protection of the active site from inactivation by NEM. SAH/RNA treatment was shown to be significantly greater than SAH treatment alone (Fig. 11).

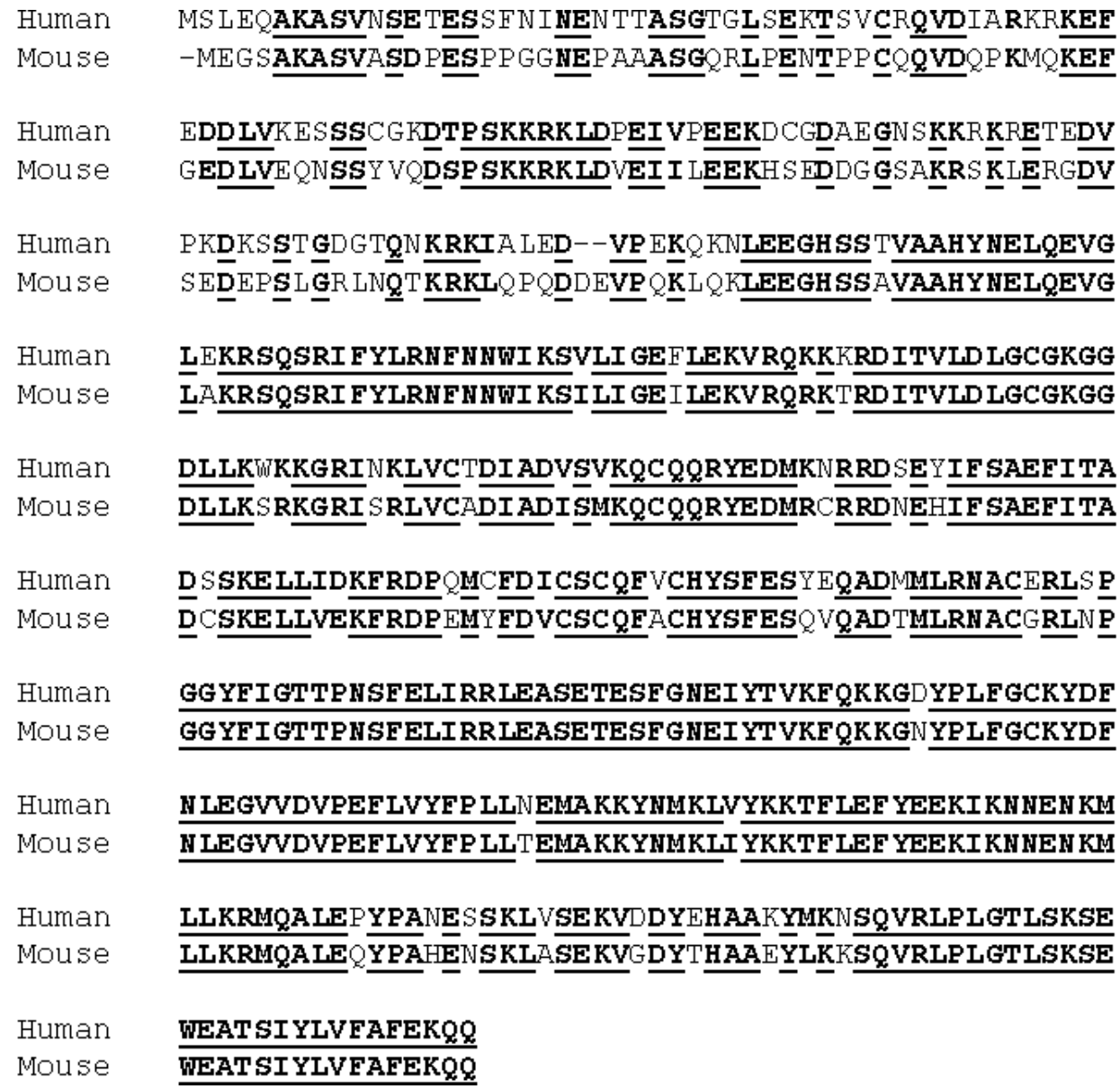


Figure 9: Alignment Between Mouse and Human GMT Sequences
An alignment between the mouse and human GMTs was created using ClustalW (44). The underlined and bolded letters correspond to complete identity while the letters in only bold are conserved mutations.

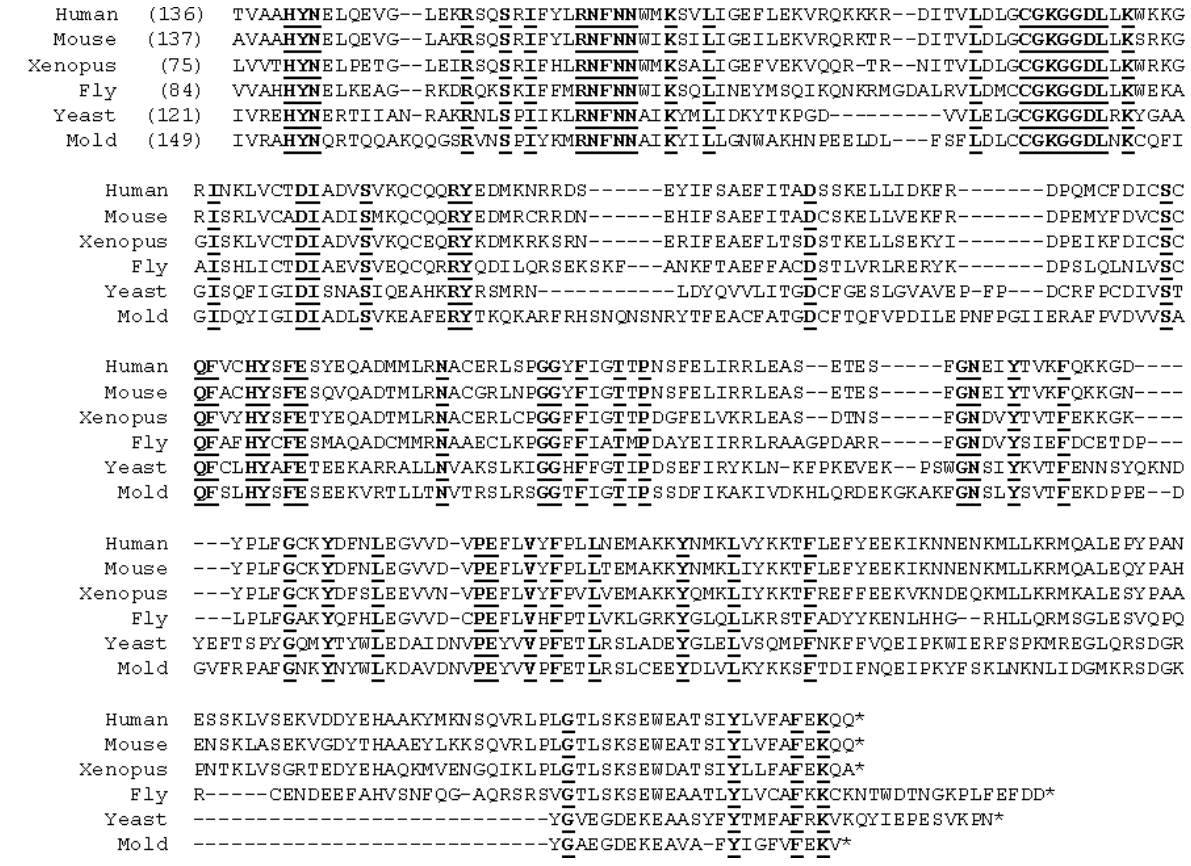


Figure 10: Alignment of Mouse and Other GMT Sequences

An alignment was performed between the predicted mouse GMT protein sequence and the known GMT sequences from human, *X. laevis*, *D. melanogaster* (fly), *S. cerevisiae* (yeast), and *C. albicans* (mold). The conserved sequences are indicated by bold and underline. Dashes represent breaks in the sequence. The * represents the end of the sequence.

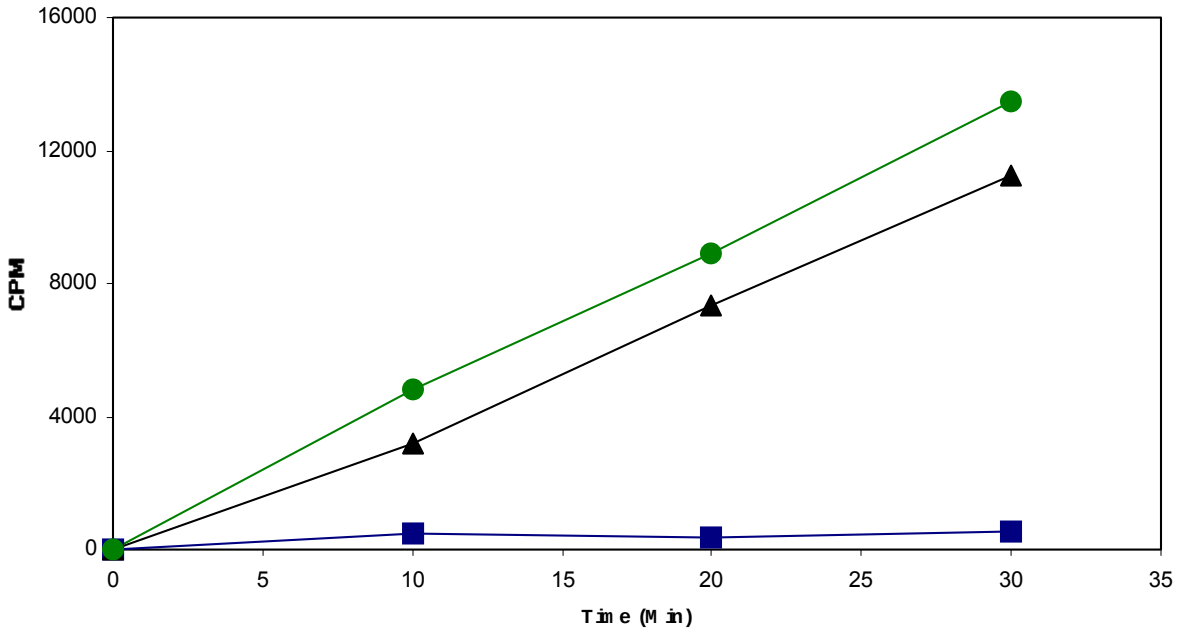


Figure 11: SAH Protects the Active Site of the GMT
A study was performed to determine which compound would protect the enzyme from inactivation by NEM. Standard assay conditions, as described in section 6.1M, were used for this assay. This experiment was performed twice and the best graph is shown. Enzyme treated with 40μM SAH and 5mM NEM (-▲-), enzyme treated with 18ng/μl capped RNA and 5mM NEM (-■-), and enzyme treated with 40μM SAH, 18ng/μl capped RNA and 5mM NEM (-•-) are shown. Activity of the enzyme is represented by radioactivity incorporated at a given time. A repeated measures statistical analysis was performed between the three sets of data at 30 min. SAH treatment (-▲-) compared to RNA treatment (-■-) yields a p-value = <0.0001. SAH treatment coupled with RNA treatment (-•-) yields a p-value = <0.0001 when compared to just RNA treatment (-■-). SAH treatment coupled with RNA treatment (-•-) compared to SAH treatment (-▲-) yields a p-value = 0.0077.

The difficulty with using SAH as a protector of the enzyme active site is that SAH is also a potent inhibitor of GMT (41), with a K_i value of 0.37μM. In order to remove the SAH, so that the enzyme could be assayed, it was necessary to dialyze the enzyme samples. The first sets of experiments to determine the optimal dialysis time for the mouse GMT were done with a two hr dialysis time. However, only 50% of the activity could be recovered in the SAH-treated samples. Multiple treatments were made to the buffer solution in an attempt to increase activity of the SAH-treated samples, including addition of BSA, bubbling Argon gas through the buffer to minimize oxidation, and the addition of DEAE Sephadex. None of these treatments increased the enzyme activity of the SAH-treated GMT. An overnight dialysis was then performed at 4°C, with a buffer exchange every 6-7 hr. This experiment was performed twice with the mouse GMT (Fig. 12). The overnight dialysis resulted in 75% of the enzymatic activity remaining in the SAH-treated sample as compared to the non-treated sample. This decrease was measured to be significant (Fig. 12). This was consistent with the enzymatic experiments conducted with HOS enzyme, whereby HOS enzyme treated with SAH showed only 75% of the enzymatic activity compared to the non-treated sample. However, mouse GMT and cloned human GMT always showed higher activity when only treated with SAH as compared to the nontreated samples in the NEM inactivation experiments. The cause of this phenomenon was not further investigated. The remainder of the figures will indicate GMT treated with SAH alone as the basis for active site protection; untreated GMT will not be shown.

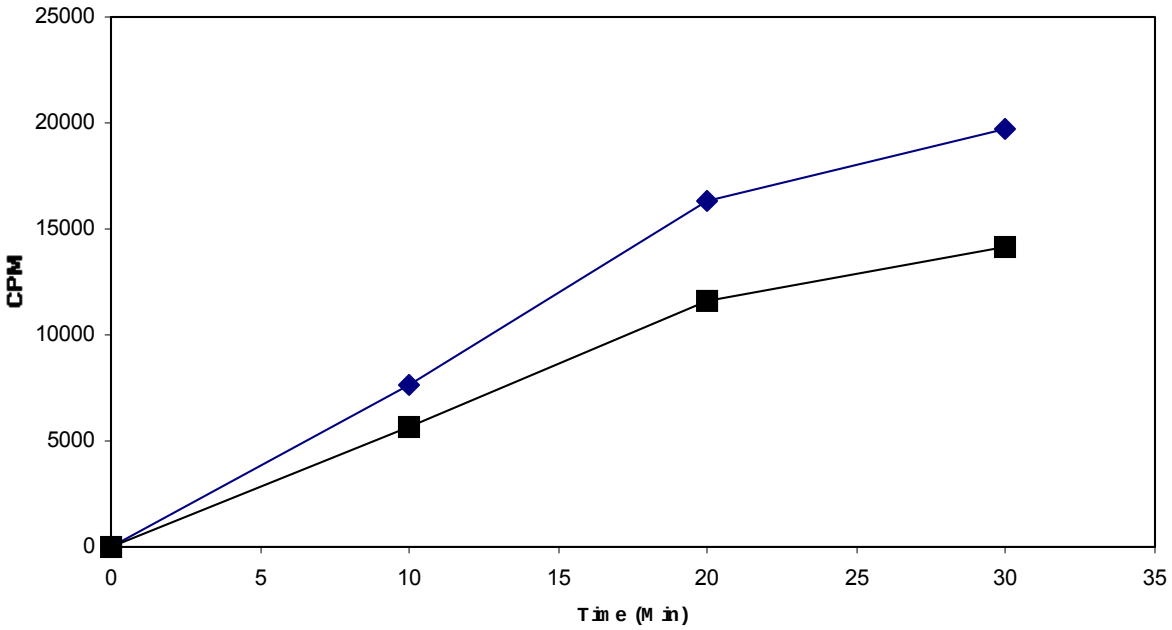


Figure 12: Overnight Dialysis of SAH Treated mouse GMT
A 24-hour dialysis was performed on non-treated (-◆-) and SAH-treated (-■-) enzyme sample at 4°C, changing the hypotonic buffer-1 + 2mM DTT every 6-7 hours. A standard assay was then performed on the samples as described in section 6.1M. Enzyme activity is represented by the amount of radioactivity incorporated for each time point. This experiment was performed twice with the best graph shown. A repeated measures statistical analysis was performed between the two sets of data at 30 min. SAH treatment (-■-) results in a significant decrease in activity of GMT, with a p-value = 0.0361.

NEM completely inactivated GMT and SAH provided protection from this inactivation (Fig. 11). When an experiment was performed with SAH, NEM, and an overnight dialysis, the enzyme sample treated with both SAH and NEM retained 60% of the activity of the non-NEM treated sample (Fig. 13). There is a significant difference between GMT treated with NEM alone and GMT treated with NEM and 40μM SAH. There was no evidence of increased protection of the enzyme at 100μM SAH, suggesting that protection was maximal at 40μM (Fig. 14).

Dialyzing the HOS GMT in the presence of NEM gave somewhat different results from the mouse GMT. Even with an overnight dialysis, only 75% of the enzymatic activity was present in the SAH only treated sample. SAH protected HOS GMT from inactivation by NEM, with nearly 100% of the activity remaining in the SAH- and NEM-treated sample when compared to enzyme treated only with SAH (Fig. 15). NEM treatment of SAH-treated GMT resulted in enzyme activity that was not significantly different from the activity of GMT treated with SAH alone.

A dialysis time of only 2 hours was needed for the cloned human GMT. Recombinant human SAH-treated GMT, like mouse SAH-treated GMT, showed higher activity than untreated enzyme. However, recombinant human GMT, in contrast to both mouse and HOS GMT, was not protected from NEM inactivation by SAH at any concentration up to 500μM (Figs. 16-17).

7.3 Active site labeling

To investigate the binding of RNA to the active site of mouse GMT, radiolabeled RNA was cross-linked to the enzyme by UV light exposure.

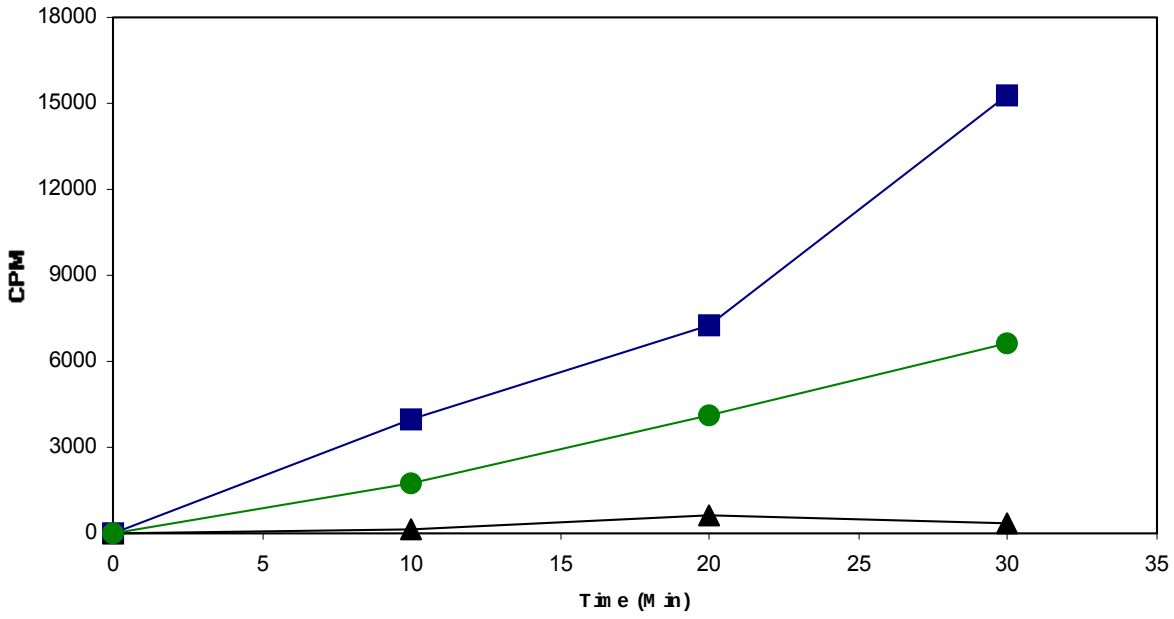


Figure 13: Protection of Mouse GMT by SAH

Mouse GMT was incubated with 40μM SAH (-■-), 40μM SAH and 5mM NEM (-•-) or 5mM NEM alone (-▲-). The samples were dialyzed overnight at 4°C. The hypotonic buffer-1 with 2mM DTT was changed every 6-7 hours and the samples were then assayed as described in section 6.1M. This experiment was performed twice with the best graph shown. A repeated measures statistical analysis was performed between the three sets of data at 30 min. SAH treatment with NEM treatment (-•-) compared to NEM treatment alone (-▲-) yields a p-value = 0.0029. SAH treatment alone (-■-) compared to NEM treatment alone (-▲-) yields a p-value = <0.0001. SAH treatment with NEM treatment (-•-) compared to SAH treatment alone (-■-) yields a p-value = 0.0021.

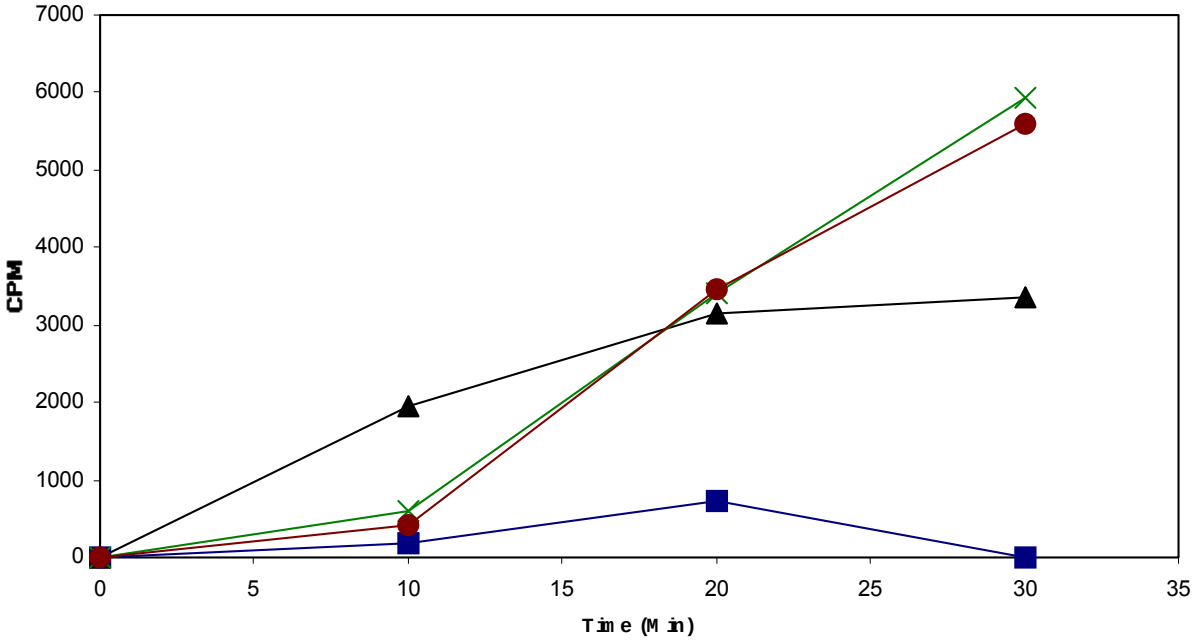


Figure 14: Increasing Concentrations of SAH

SAH was added to mouse GMT samples treated with NEM at concentrations of 0 (-■-), 10 (-▲-), 40(-X-), and 100μM (-•-). The samples were dialyzed overnight at 4°C, changing the hypotonic buffer-1 + 2mM DTT every 6-7 hours, and assayed as described in section 6.1M.

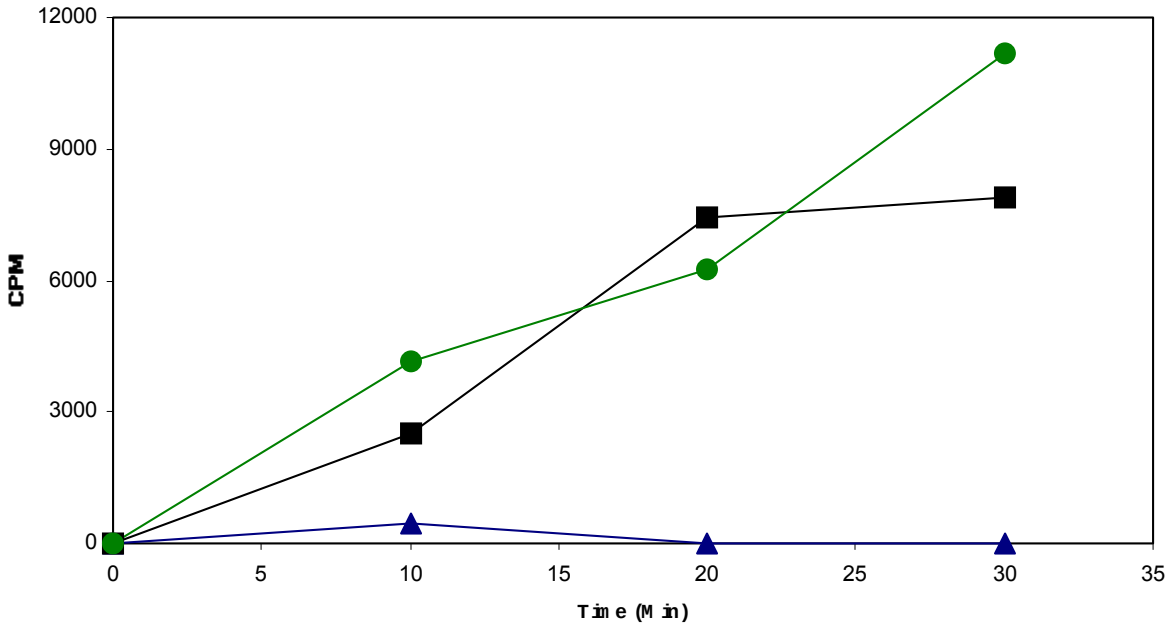


Figure 15: Protection of HOS GMT by SAH

HOS GMT was incubated with 40μM SAH (-■-), 40μM SAH and 5mM NEM (-•-) or 5mM NEM alone (-▲-). The samples were dialyzed overnight in hypotonic buffer-1 + 2mM DTT at 4°C. The buffer was changed every 6-7 hours and the samples were then assayed as described in section 6.1M. This experiment was performed twice with the best graph shown. A repeated measures statistical analysis was performed between the three sets of data at 30 min. SAH treatment with NEM treatment (-•-) compared to NEM treatment alone (-▲-) yields a p-value = 0.0003. SAH treatment alone (-■-) compared to NEM treatment alone (-▲-) yields a p-value = <0.0001. SAH treatment with NEM (-•-) compared to SAH treatment alone (-■-) yields a p-value = 0.4180.

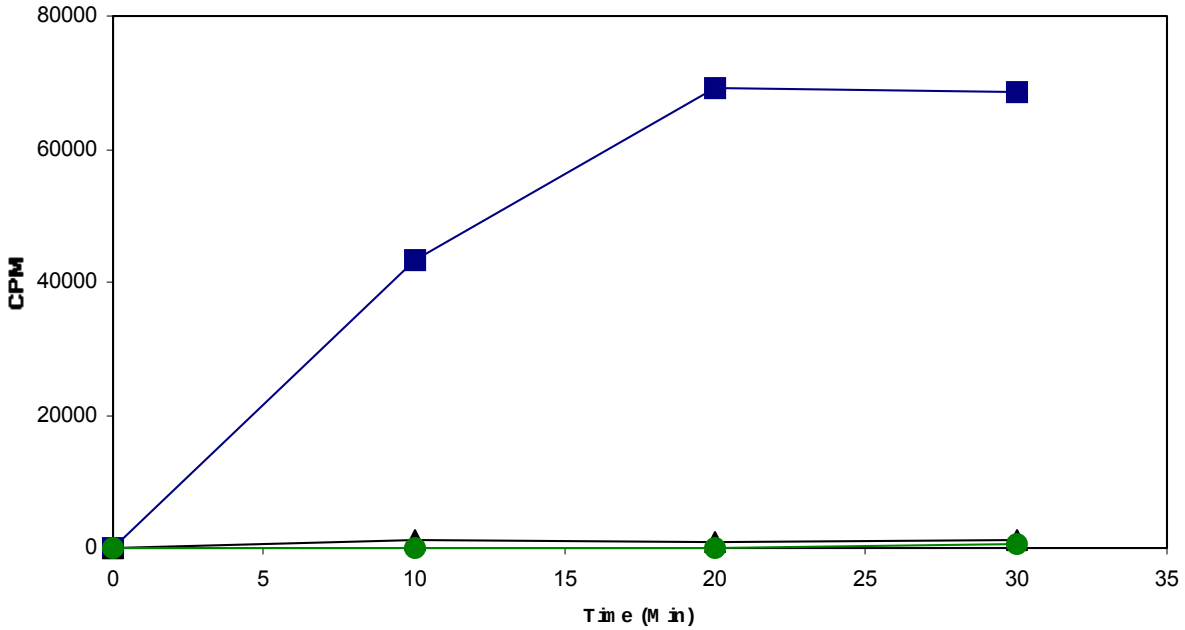


Figure 16: Protection of Cloned Human GMT by SAH
Recombinant human enzyme was incubated with 40μM SAH (-■-), 5mM NEM (-▲-), or with 40μM SAH and then with 5mM NEM (-●-). The samples were dialyzed in hypotonic buffer-1 + 2mM DTT for 2 hours at room temperature, changing the buffer after 1 hour. The samples were then assayed as described in section 6.1M. This experiment was performed twice with the best graph shown. A repeated measures statistical analysis was performed between the three sets of data at 30 min. SAH treatment with NEM treatment (-●-) compared to NEM treatment alone (-▲-) yields a p-value = 0.9569. SAH treatment alone (-■-) compared to NEM treatment alone (-▲-) yields a p-value = <0.0001. SAH treatment with NEM (-●-) compared to SAH treatment alone (-■-) yields a p-value = <0.0001.

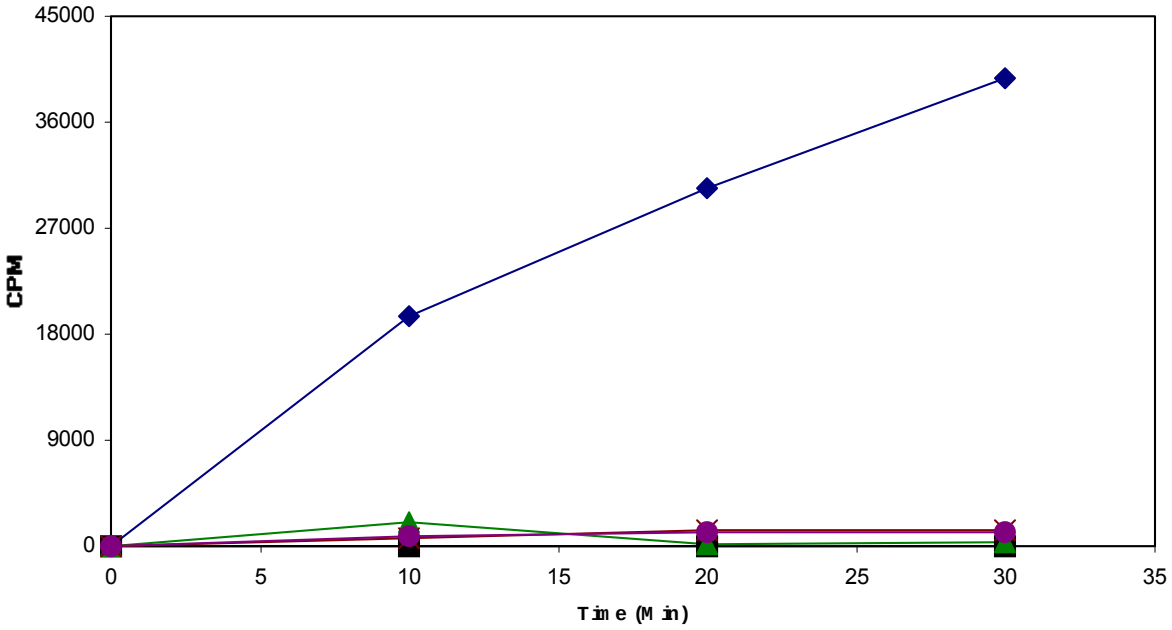


Figure 17: Increasing Concentration of SAH with Cloned Human GMT
Recombinant human enzyme was incubated with SAH at concentrations of 0 (-■-), 40 (-▲-), 100 (-X-), and 500μM (-•-) and then with 5mM NEM. Enzyme not treated with SAH or NEM is denoted by (-◆-). The samples were dialyzed in hypotonic buffer-1 + 2mM DTT for 2 hours at room temperature, changing the buffer after 1 hour, and the samples were then assayed as described in section 6.1M.

NEM inactivation of GMT suggested that NEM could also inhibit cross-linking of RNA to the active site of the enzyme. NEM treatment of the enzyme inhibited cross-linking to either capped or uncapped 11-mer RNA by 35% (Fig. 18). However, NEM did not inhibit the cross-linking of capped 21-mer RNA (Fig. 18). Quantitation of NEM inhibition was performed as described in section 6.1U.

The effect of cap analog on the cross-linking of RNA to the GMT was examined next. Cap analog, GpppG, at a concentration of 1mM maximally inhibits cross-linking (Fig. 19). The cross-linking of RNA to the enzyme decreases proportionally with increasing amounts of cap analog.

To determine if non-specific interactions were responsible for the inhibition of RNA cross-linking at such high concentrations of GpppG, GMP was added to enzyme/radiolabeled RNA samples. GMP is not a substrate for GMT. Even at a concentration of 1mM, GMP was not able to inhibit cross-linking of 5'-labeled RNA (Fig. 20).

SAH protects GMT from inactivation by NEM, however its effects on the cross-linking of RNA were unknown. SAH, at a concentration of 0.5mM, did not have a dramatic effect on cross-linking of RNA to the enzyme (Fig 21). Cap analog, at a concentration of 1mM, inhibited cross-linking of RNA to the enzyme as before (Fig. 21). SAH and cap analog together were only as effective as the cap analog alone (Fig. 21).

The next experiment was performed to examine whether or not the radioactive phosphate at the 5'-end was the reason for cap analogs inhibitory effects. RNA was labeled with ^{32}P at either the 5' or the 3' ends, as described in sections 6.1R and 6.1S. The cross-linking of both the 5' and 3' labeled RNA to the enzyme was examined in the

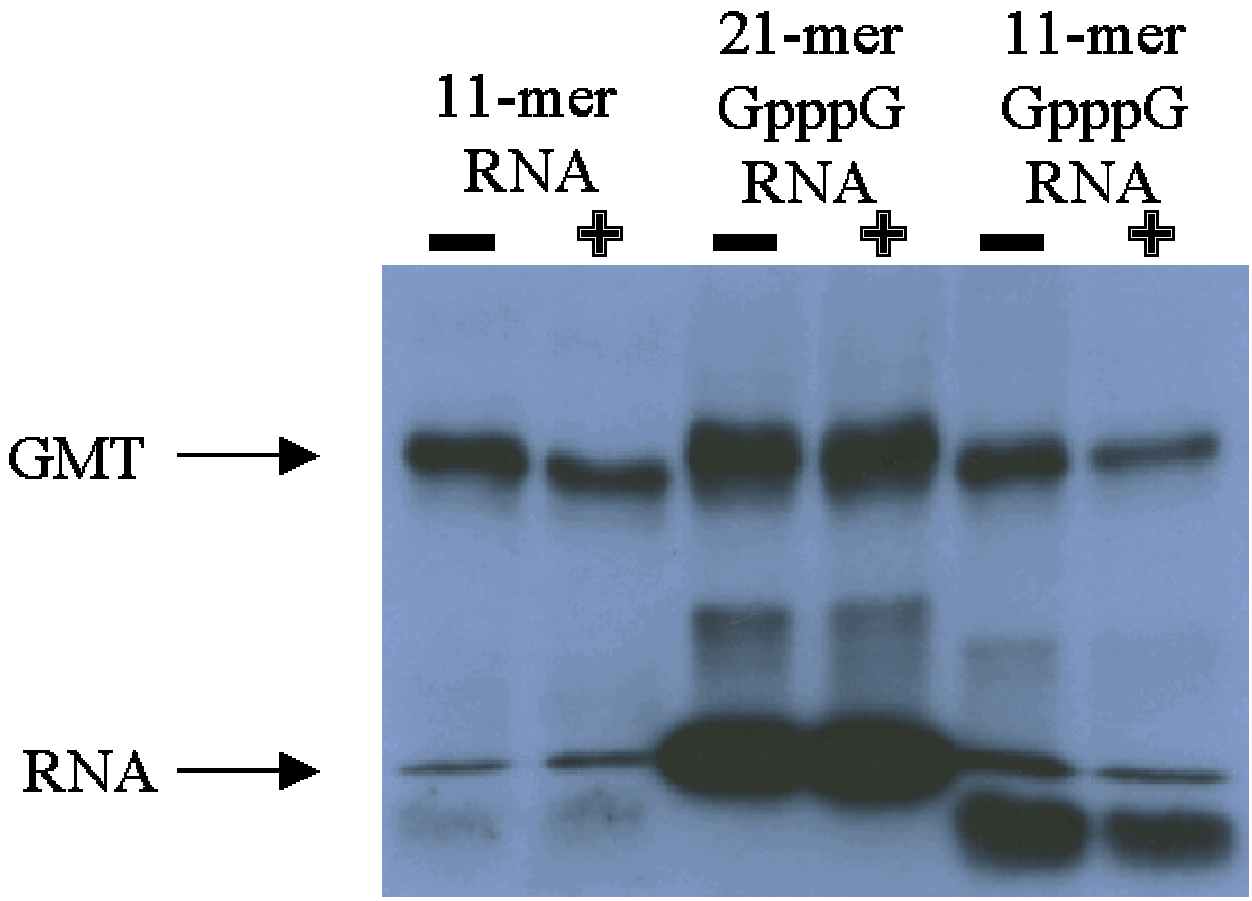


Figure 18: Active Site Labeling with and without NEM
NEM was added to mouse GMT at a concentration of 12mM. 11-mer uncapped, 11-mer capped, and 21-mer capped RNAs were all cross-linked to the enzyme. The (+) and (-) refer to when NEM presence (+) and absence (-). The samples were exposed to UV light for 45 min, separated by SDS-PAGE and visualized by exposure to X-ray film at room temperature for 3 hours.

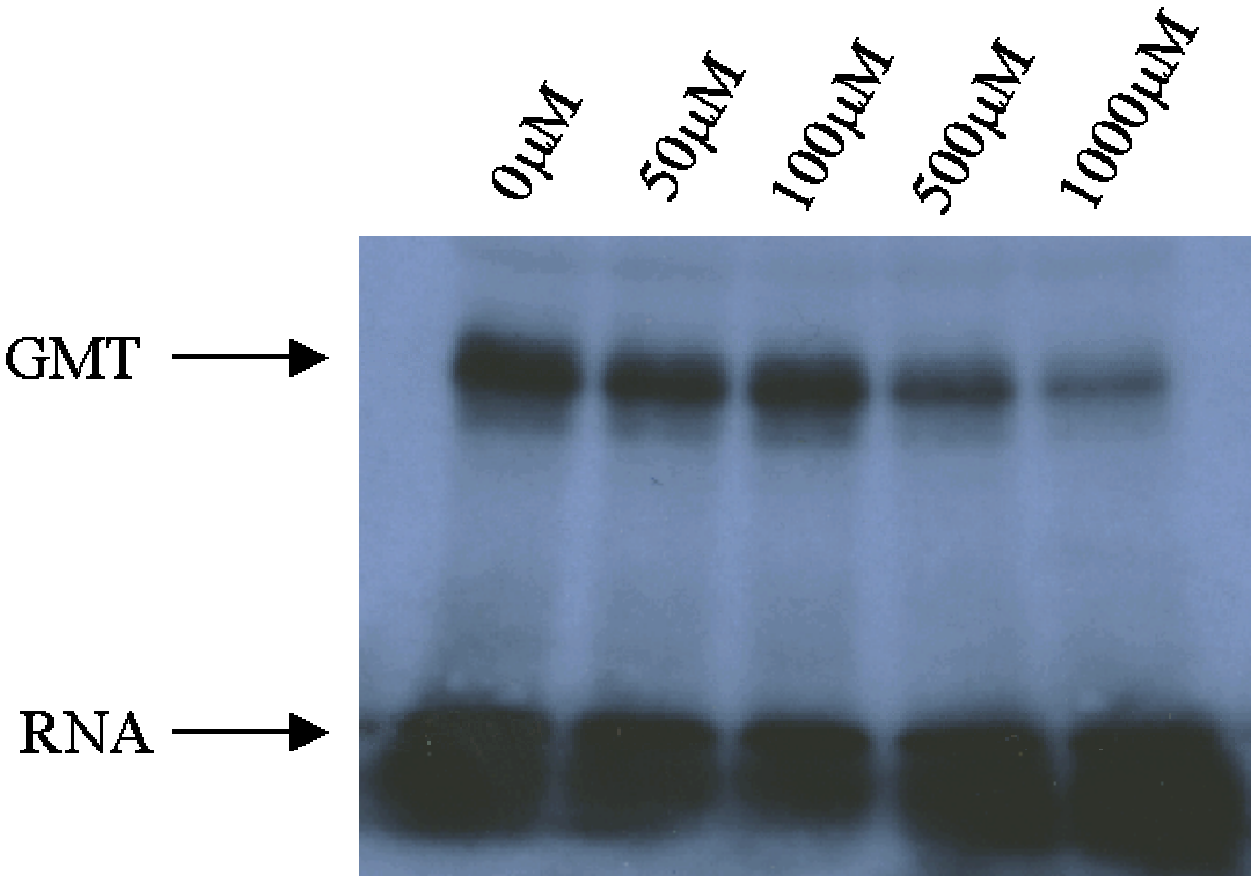


Figure 19: Active Site Labeling with Increasing Concentrations of GpppG
Cap analog, GpppG, was added to mouse enzyme at concentrations of 0, 50, 100, 500, and 1000μM. The samples were exposed to UV light for 45 min, separated by SDS-PAGE and visualized by exposure to X-ray film at room temperature for 3 hours.

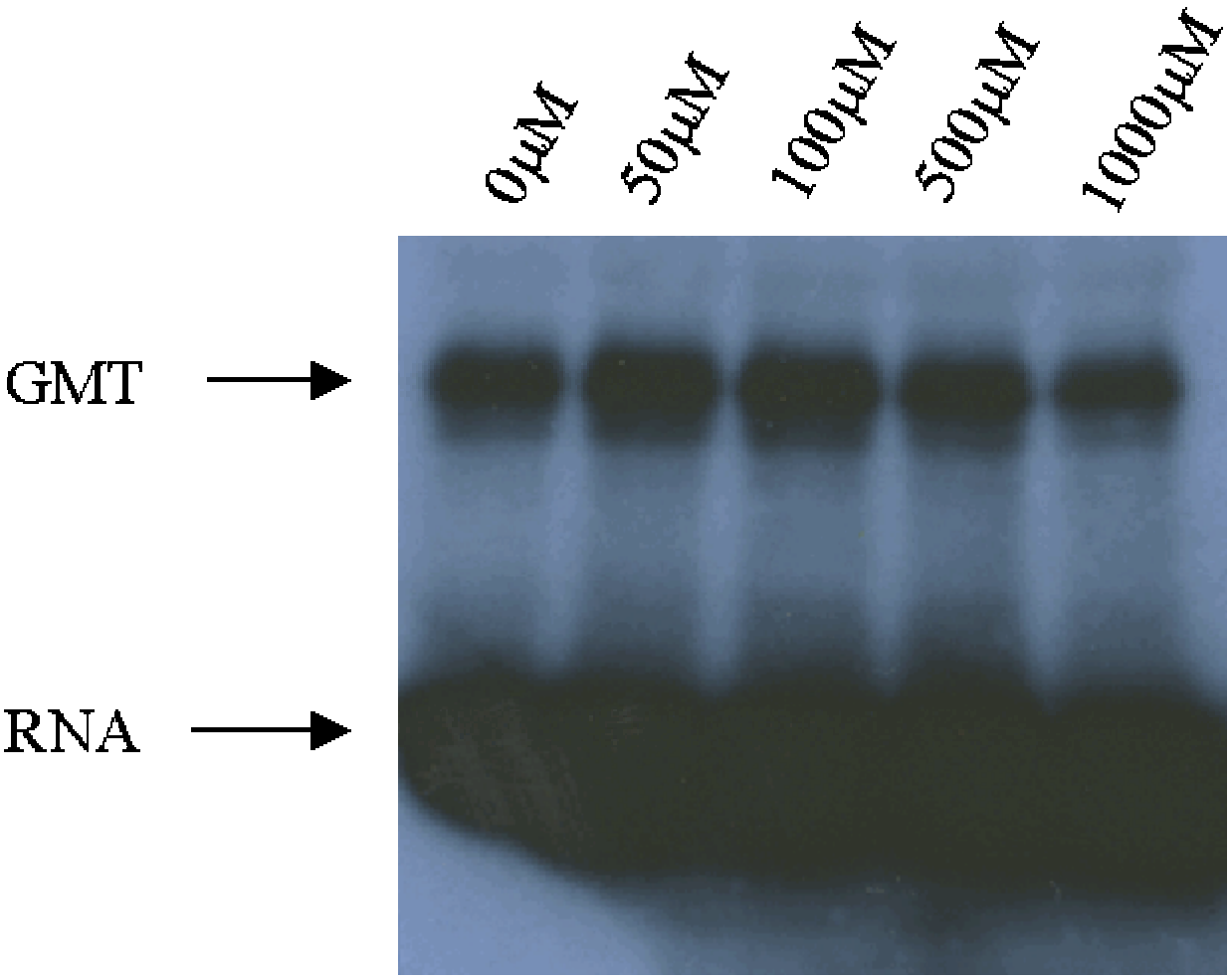


Figure 20: Active Site Labeling with Increasing Concentrations of GMP
GMP was added to mouse GMT at concentrations of 0, 50, 100, 500, and 1000μM. The samples were exposed to UV light for 45 min, separated by SDS-PAGE and visualized by exposure to X-ray film at room temperature for 3 hours.

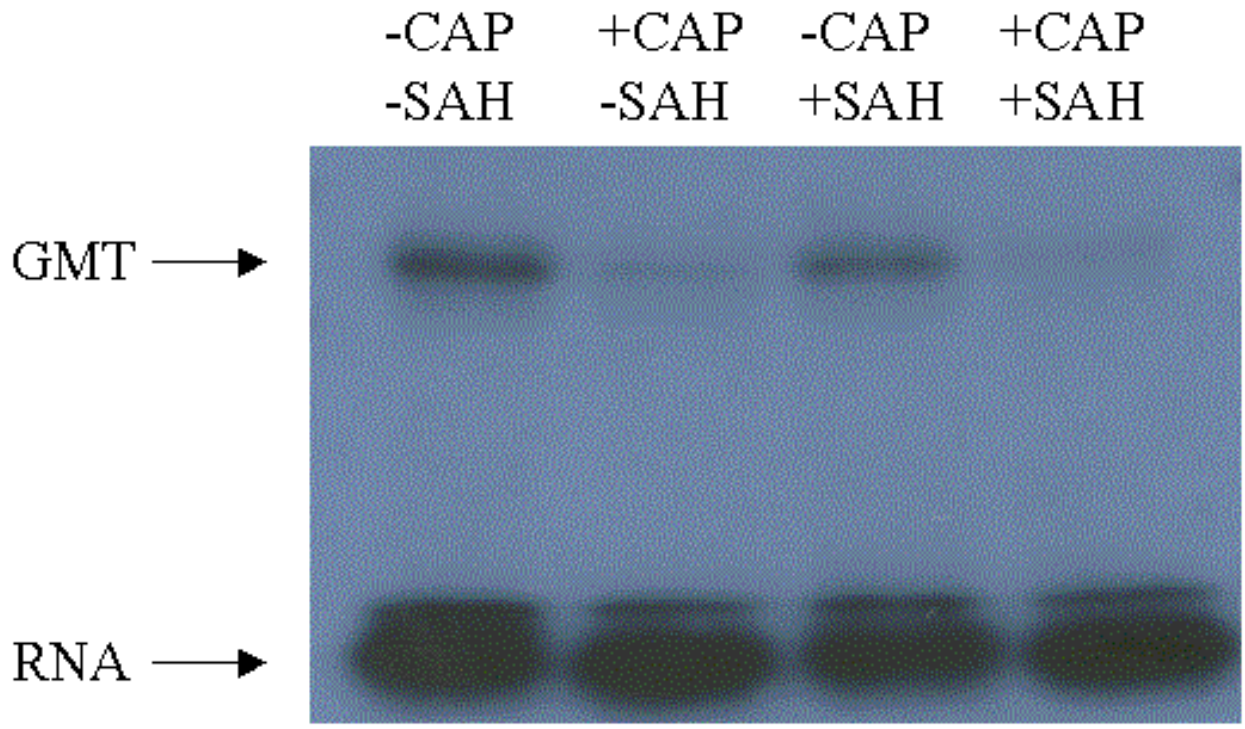


Figure 21: Active Site Labeling with SAH and GpppG
SAH and GpppG were added to mouse enzyme at concentrations of 0.5mM and 1mM, respectively. The samples were exposed to UV light for 45 min, separated by SDS-PAGE and visualized by exposure to X-ray film at room temperature for 3 hours.

presence of 1mM GpppG (Fig. 22). Both the 5' and 3'-labeled RNA were inhibited from cross-linking to the enzyme by cap analog (Fig. 22). Thus, the location of the radioactive phosphate does not impact cap analogs inhibition of RNA binding to GMT.

When GMT is isolated from the Mono-Q column two peaks of activity are recovered. Cross-linking of RNA was performed with both Peak 1 and Peak 2 mouse GMT samples from the Mono-Q column to determine if any differences existed between the two samples. GpppG was added to the samples at a concentration of 1mM. When radiolabeled RNA was cross-linked to the Peak 2 enzyme sample, two very strong bands were apparent, as opposed to one band with the Peak 1 enzyme (Fig. 23). One of the bands appears at the same position on the gel as the band in Peak 1 samples. The other band migrated slower on the gel. Addition of 1mM of the cap analog inhibits cross-linking to the Peak 1 enzyme sample, as shown previously, and also to the Peak 2 sample (Fig. 23). Peak 2 appears to have at the very least another RNA binding protein, which perhaps is an alternate form of GMT.

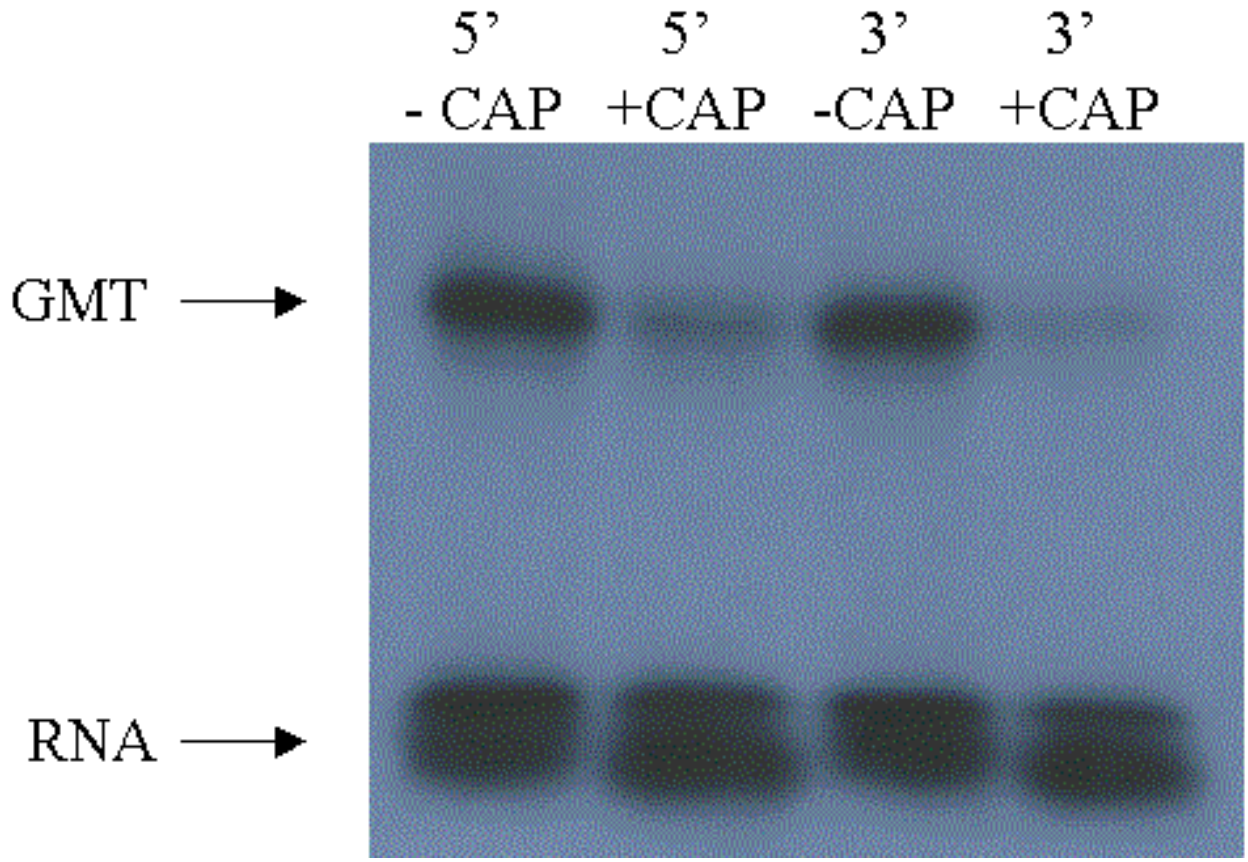


Figure 22: Active Site Labeling with 5' and 3'-Labeled RNA
RNA was labeled with ^{32}P as described in the methods section and incubated with mouse enzyme. Cap analog, GpppG, was added at a concentration of 1mM. The samples were exposed to UV light for 45 min, separated by SDS-PAGE and visualized by exposure to X-ray film at room temperature for 3 hours.

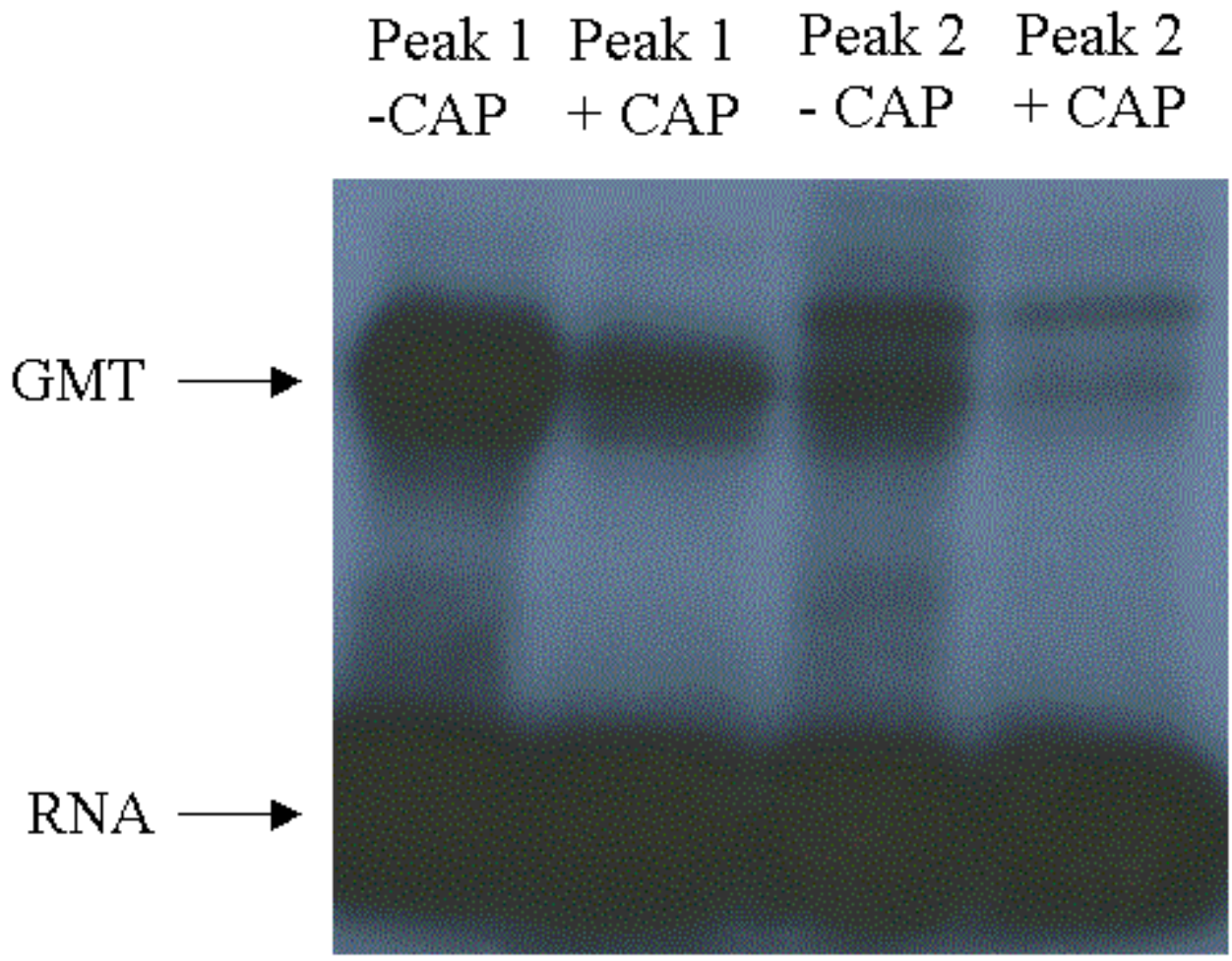


Figure 23: Active Site Labeling with Peak 1 and Peak 2 Mouse GMT
Mouse enzyme samples from Peak 1 and Peak 2 of the Mono-Q column were UV-cross-linked with ^{32}P-5'-labeled RNA. The samples were exposed to UV light for 45 min, separated by SDS-PAGE and visualized by exposure to X-ray film at room temperature for 3 hours.

VIII. DISCUSSION

GMT is known to be essential for cell viability, thus the study of the enzyme is warranted to better understand RNA processing. To date, the sequences of eight other GMTs have been elucidated. Until now, however, the sequence of the mouse GMT had not been determined.

The sequence of the mouse cDNA encoding the GMT was constructed using 6 ESTs that overlapped with one another. This cDNA is 2,035 bp in length and encodes an ATG at 87 base pairs from the first nucleotide and a putative AATAAA poly-A signal sequence 16 nucleotides from end of the cDNA.

When comparing the mouse and the human cDNAs it is interesting to note that the mouse cDNA is 2,035 bp while the human cDNA is 6,189 bp. The 3' untranslated region (UTR) in the mouse cDNA is 551 bp, which is still longer than the usual 200-300 base pairs, but the human GMT 3'-UTR is 4,578 bp, by far the largest 3'-UTR of any of the cDNAs encoding the GMT. The *X. laevis* GMT 3'-UTR is 662 bp, the *C. albicans* GMT 3'-UTR is 528 bp, and the *D. melanogaster* GMT 3'-UTR is 447 bp. The function for the enormous 3'-UTR in the human GMT cDNA is unclear. Tsukamoto *et al.* (59) reported finding 3 different cDNAs encoding for the human GMT, none of which were 6,189 base pairs, which is the length of the human GMT 3'-UTR reported by Pillutla *et al.* (21).

The mouse GMT ORF encodes 465 amino acids. Comparison of the amino acid sequence of the deduced mouse and human GMTs revealed that they are 74.6% identical. Analysis of the deduced mouse and human GMT amino acids 121-465 revealed that they are 86.5% identical in the C-terminal domain. The mouse GMT

cDNA encodes conserved motifs shared among all GMTs. Taken together this evidence indicates that this is most likely the correct full-length sequence of the mouse GMT, deduced solely from available genomic data. Once the mouse genome is sequenced it could confirm these results.

The second objective of this study was to characterize the active site, through both protection of the enzyme from a known inhibitor and blocking of RNA binding to the active site. NEM inactivated the mouse and human GMTs at a concentration of 5mM, while 40μM SAH protected the mouse and HOS enzymes from inactivation. SAH treatment alone, however, resulted in a detectable loss of enzyme activity. It is possible that SAH, a potent inhibitor of GMT, was not fully removed by dialysis or caused a conformational change in the enzyme. The latter of the two possibilities seems more likely because recovery of the recombinant human GMT activity was evident after a 2 hr dialysis. Treatment of the recombinant human GMT with SAH actually made the enzyme sample more active than the untreated sample. Perhaps in this case SAH created a conformational change in the GST-fused enzyme that protected it from inactivation. The fact that the human GMT was fused to glutathione-S-transferase (GST) probably had something to do with the fact it behaved differently from enzyme isolated from either Ehrlich ascites or HOS cells.

SAH protected only the mouse and HOS GMTs, and showed no protective effects on the cloned human GMT. Approximately 60% of the mouse enzyme activity remained following treatment of the sample with 40μM SAH and 5mM NEM, while nearly 100% of the activity of the HOS GMT remained following SAH and NEM

treatment. Cloned human GMT was not protected by SAH, even at a concentration of 500μM.

The fact that only 60% of the mouse enzyme activity was protected by SAH suggests that a cysteine is not essential to the enzymatic activity of the enzyme. Schwer *et al.* (55) mutated the cysteine in Motif I to an alanine and retained activity in the enzyme. However, the glycine next to the cysteine in Motif I is essential for enzyme activity. It is possible that when NEM reacts with the cysteine located next to the essential glycine, the participation of glycine in catalysis is impaired or hindered. NEM may also react with the other GMT cysteines (12 in human GMT, 11 in mouse GMT), altering conformation and enzyme activity. If other cysteines are modified, binding of SAH to the active site would not reverse the effects of these modifications on enzyme activity.

Why the HOS GMT is totally protected from NEM inactivation by SAH, while the mouse enzyme is only approximately 60% protected, is perplexing. It is possible that an exposed cysteine outside of the active site of the mouse enzyme that is not present at the same site in the human enzyme could react with NEM. The fact that the cloned human GMT could not be protected at all by SAH may be due to the fact it is fused to GST.

NEM was able to completely inactivate the enzymatic activity of the GMT but could only inhibit cross-linking of 11-mer capped or uncapped RNA by 35%. NEM did not have any effect on the cross-linking of 21-mer capped RNA. Perhaps a second RNA binding domain is present that is only accessible to longer chain RNAs. SAH did protect the enzymatic activity of the GMT but did not have any significant effect on the

cross-linking of RNA to the active site. This supports the idea that there are distinct binding domains for both SAM and capped RNA in the active site of the enzyme.

UV-cross-linking of RNA was performed with and without cap analog, and with both 5' and 3'-labeled RNA. An increase in the concentration of cap analog, from 0-1mM coincided with a decrease in the cross-linking of RNA to the enzyme. This occurred using both the 5' and the 3'-labeled RNA. This suggests that the active site must be vacant for proper binding of RNA to the active site. To check for any non-specific interactions which may be occurring, the similar nucleotide GMP was incubated with the enzyme at the same concentrations as the cap analog. Even at a concentration of 1mM, GMP had no effect on cross-linking of RNA. A slight enhancement of cross-linking at GMP concentrations of 50μM and 100μM was observed, but this phenomenon was not investigated further.

These studies raised several unanswered questions. One question is related to the proposed molecular weight of the mouse GMT. Bu (41) found the molecular weight of the mouse GMT to be approximately 95,000 on a Superose 6 gel filtration column and approximately 46,000 on an SDS-gel. The calculated molecular weight from the deduced sequence of the mouse GMT is approximately 53,000 molecular weight. This size suggests that an investigation of post-translational processing of GMT may be warranted.

IX. LITERATURE CITED

1. Kozak, M. (1983) Comparison of Initiation of Protein Synthesis in Procaryotes, Eucaryotes, and Organelles. *Microbiological Rev.* **47**, 1-45.

2. Banerjee, A.K. (1980) 5'-Terminal Cap Structure in Eukaryotic Messenger Ribonucleic Acids. *Microbiological Rev.* **44**, 175-205.

3. Reddy, R., Ro-Choi, T.S., Henning, D., and Busch, H. (1974) Primary Sequence of U-1 Nuclear Ribonucleic Acid of Novikoff Hepatoma Ascites Cells. *J. Biol. Chem.* **249**, 6486-6494.

4. Furuichi, Y., Morgan, M., Muthukrishnan, S., and Shatkin, A.J. (1975) Reovirus Messenger RNA Contains a Methylated, Blocked 5'-Terminal Structure, m^7G(5')ppp(5')GpCp. *Proc. Natl. Acad. Sci. USA* **72**, 362-366.

5. Wei, C.M., and Moss, B. (1975) Methylated Nucleotides Block 5'-Terminus of Vaccinia Virus Messenger RNA. *Proc. Natl. Acad. Sci. USA* **72**, 318-322.

6. Furuichi, Y., and Miura, K.I. (1975) A Blocked Structure at the 5'-Terminus of Messenger RNA from Cytoplasmic Polyhedrosis Virus. *Nature (London)* **253**, 374-375.

7. Adams, J.M., and Cory, S. (1975) Modified Nucleosides and Bizarre 5'-Termini in Mouse Myeloma mRNA. *Nature* **255**, 28-33.

8. Perry, R.P., Kelley, D.E., Frederick, K.H., and Rottman, F.M. (1975) Methylated Constituents of Heterogeneous Nuclear RNA: Presence in Blocked 5'-Terminal Structures. *Cell* **4**, 13-19.

9. Furuichi, Y., Morgan, M., Shatkin, A.J., Jelinek W., Salditt-Georgieff, M. and Darnell, J.E. (1975) Methylated, Blocked 5'-Termini in HeLa Cell mRNA. *Proc. Natl. Acad. Sci. USA* **72**, 1904-1908.

10. Muthukrishnan, S., Filipowicz, W., Sierra, J.M., Both, G.W., Shatkin, A.J., and Ochoa, S. (1975) Messenger RNA Methylation and Protein Synthesis in Extracts from Embryos of Brine Shrimp, *Artemia salina*. *J. Biol. Chem.* **250**, 9336-9341.

11. Shatkin, A.J. (1976) Capping of Eukaryotic mRNAs. *Cell* **9**, 645-653.

12. Dubin, D.T., and Taylor, R.H. (1975) The Methylation State of Poly A Containing Messenger RNA from Cultured Hamster Cells. *Nucleic Acids Res.* **2**, 1653-1668.

13. Shuman, S., Surks, M., Furneaux, H., and Hurwitz, J. (1980) Purification and Characterization of a GTP-Pyrophosphatase Exchange Activity from Vaccinia Virus. Association of the GTP-Pyrophosphatase Exchange Activity with Vaccinia mRNA

Guanylyltransferase, RNA (Guanine-7-)Methyltransferase Complex (Capping Enzyme). *J. Biol. Chem.* **255**, 11588-11598.

14. Niles, E.G., Lee-Chen, G., Shuman, S., Moss,B., and Broyles, S. (1989) Vaccinia Virus D12L Encodes the Small Subunit of the Viral mRNA Capping Enzyme. *Virology* **172**, 513-522.

15. Higman, M.A., Bourgeois, N., and Niles, E.G. (1992) The Vaccinia Virus mRNA (Guanine-N7-) Methyltransferase Requires Both Subunits of the mRNA Capping Enzyme for Activity. *J. Biol. Chem.* **267**, 16430-16437.

16. Mizumoto, K., and Lippman, F. (1979) Transmethylation and Transguanylylation in 5'-RNA Capping System Isolated from Rat Liver Nuclei. *Proc. Natl. Acad. Sci. USA* **76**, 4961-4965.

17. Shuman, S. (1995) Capping Enzyme in mRNA Synthesis. *Prog. Nucleic Acids Res.* **50**, 101-129.

18. Itoh, N., Yamada, H., Kaziro, Y., and Mizumoto, K. (1987) Messenger RNA Guanylyltransferase from *Saccharomyces cerevisiae*: Large Scale Purification, Subunit Functions and Subcellular Localization. *J. Biol. Chem.* **262**, 1989-1995.

19. Mao, X., Schwer, B., and Shuman, S. (1995) Yeast mRNA Cap Methyltransferase is a 50-Kilodalton Protein Encoded by an Essential Gene. *Mol. Cell. Biol.* **15**, 4167-4174.

20. Wen, Y., Yue, Z., and Shatkin, A.J. (1998) Mammalian Capping Enzyme Binds RNA and Uses Protein Tyrosine Phosphatase Mechanism. *Proc. Natl. Acad. Sci. USA* **95**, 12226-12231.

21. Pillutla, R.C., Yue, Z., Maldonado, E., and Shatkin, A.J. (1998) Recombinant Human mRNA Cap Methyltransferase Binds Capping Enzyme/RNA Polymerase IIo Complexes. *J. Biol. Chem.* **273**, 21443-21443.

22. McCracken, S., Fang, N., Rosanina, E., Yankulov, K. Brothers, G., and Bentley, D. (1999) 5'-Capping Enzymes are Targeted to Pre-mRNA by Binding to the Phosphorylated Carboxy-terminal Domain of RNA Polymerase II. *Genes Dev.* **11**, 3306-3318.

23. Corden, J.L., Codena, D.L., Ahearn, J.M., and Dahmus, M.E. (1985) A Unique Structure at the Carboxyl Terminus of the Largest Subunit of Eukaryotic RNA Polymerase II. *Proc. Natl. Acad. Sci. USA* **82**, 7934-7938.

24. Dahmus, M.E. (1996) Reversible Phosphorylation of the C-Terminal Domain of RNA Polymerase II. *J. Biol. Chem.* **271**, 19009-19012.

25. Hagler, J., and Shuman, S. (1992) A Freeze Frame View of Eukaryotic Transcription During Elongation and Capping of Nacsent mRNA. *Science* **255**, 983-986.

26. Rasmussen, E., and Lis, J.T. (1993) *In vivo* Transcriptional Pausing and Cap Formation on Three *Drosophila* Heat Shock Genes. *Proc. Natl. Acad. Sci. USA* **90**, 7923-7927.

27. Ho, C.K., Schwer, B., and Shuman, S. (1998) Genetic, Physical, and Functional Interactions Between the Triphosphatase and Guanylyltransferase Components of the Yeast mRNA Capping Apparatus. *Mol. Cell. Biol.* **18**, 5189-5198.

28. Shibagaki, Y., Ituh, N., Yamada, H., Nagata, S., and Mizumoto, K. (1992) mRNA Capping Enzyme: Isolation and Characterization of the Gene Encoding mRNA Guanylyltransferase Subunit from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 9521-9528.

29. Both, G.W., Banerjee, A.K., and Shatkin, A.J. (1975) Methylation Dependent Translation of Viral Messenger RNAs *in vitro*. *Proc. Natl. Acad. Sci. USA* **72**, 1189-1193.

30. Rose, J.K. (1975) Heterogeneous 5'-Terminal Structures Occur on Vesicular Stomatitis Virus mRNAs. *J. Biol. Chem.* **250**, 8089-8104.

31. Muthukrishnan, S., Both, G.W., Furuichi, Y. and Shatkin, A.J. (1976) 5'-Terminal 7-Methylguanosine in Eukaryotic mRNA is Required for Translation. *Nature* **255**, 33-37.

32. Darzynkiewicz, E., Stepinski, J., Ekiel, I., Goyer, C., Sonenberg, N., Temeriusz, A., Jin., Y., Sijuwade, T., Habe, D., and Tahara, S. (1989) Translation by Nucleoside 5'-Monophosphate Analogues of mRNA 5'-Cap: Changes in N7 Substituent Affect Analogue Activity. *Biochemistry* **28**, 4771-4778.

33. Hsu, P.C., Hodel, M.R., Thomas, J.W., Taylor, L.J., Hagedon, C.H., and Hodel, A.E. (2000) Structural Requirements for the Specific Recognition of a m^7G mRNA Cap. *Biochemistry* **39**, 13730-13736.

34. Adams, B.L., Morgan, M., Muthukrishnan, S., Hecht, S.M., and Shatkin, A.J. (1978) The Effect of "Cap" Analogues on Reovirus mRNA Binding to Wheat Germ Ribosomes. *J. Biol. Chem.* **253**, 2589-2595.

35. Edery, I., and Sonenberg, N. (1985) Cap-dependent RNA Splicing in a HeLa Nuclear Extract. *Proc. Natl. Acad. Sci. USA* **82**, 7590-7594.

36. Inoue, K., Ohno, M., Sakamoto, H., and Shimura, Y. (1989) Effect of the Cap Structure on Pre-mRNA Splicing in *Xenopus* Oocyte Nuclei. *Genes Dev.* **3**, 1472-1479.

37. Furuichi, Y., LaFiandra, A., and Shatkin, A.J. (1977) 5'-Terminal Structure and mRNA Stability. *Nature* **266**, 235-239.

38. Hamm, J., and Mattaj, I.W. (1990) Monomethylated Cap Structures Facilitate RNA Export from the Nucleus. *Cell* **63**, 109-118.

39. Schwer, B., Mao, X., and Shuman, S. (1998) Accelerated mRNA Decay in Conditional Mutants of Yeast mRNA Capping Enzyme. *Nucleic Acids Res.* **26**, 2050-2057.

40. Izzauralde, E., Lewis, J., McGuigan, C., Jankowska, M. Darzynkiewicz, E., and Mattaj, I.W. (1994) A Nuclear Cap-Binding Protein Complex Involved in Pre-mRNA Splicing. *Cell* **78**, 657-668.

41. Bu, G. (1990) Purification, Characterization, and Substrate Specificity of a Nuclear mRNA (Guanine-7-) Methyltransferase from Ehrlich Ascites Cells.

42. Shugart, L. (1978) Kinetic Studies of *Escherichia coli* Transfer RNA (Uracil-5-) Methyltransferase. *Biochemistry* **17**, 1068-1072.

43. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**, 403-410.

44. Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994) ClustalW: Improving the Sensitivity of Progressive Multiple Sequence Alignments Through Sequence Weighting, Position-Specific Gap Penalties, and Weight Matrix Choice. *Nucleic Acids Res.* **22**, 4673-4680.

45. Kagan, R.M., and Clarke, S.C. (1994) Widespread Occurrence of Three Sequence Motifs in Diverse S-Adenosylmethionine-dependent Methyltransferases Suggests a Common Structure for these Enzymes. *Arch. Biochem. Biophys.* **310**, 417-427.

46. Mao, X., Schwer, B., and Shuman, S. (1996) Mutational Analyses of the *Saccharomyces cerevisiae* ABD1 Gene: Cap Methyltransferase Activity is Essential for Cell Growth. *Mol. Cell. Biol.* **16**, 475-480.

47. Wang, S.P., Deng, L., Ho, C.K., and Shuman, S. (1997) Phylogeny of mRNA Capping Enzymes. *Proc. Natl. Acad. Sci. USA* **94**, 9573-9578.

48. Yokoska, J., Tsukamoto, T., Miura, K., Shiokawa, K., and Mizumoto, K. (2000) Cloning and Characterization of mRNA Capping Enzyme and mRNA (Guanine-7-) Methyltransferase cDNAs from *Xenopus laevis*. *Bioch. Biophys. Res.* **268**, 617-624.

49. Yamada-Okabe, T., Mio, T., Koshima, Y., Matsui, M., Arisawa, M., and Yamada-Okabe, H. (1999) The *Candida albicans* Gene for mRNA 5'-Cap Methyltransferase: Identification of Additional Residues Essential for Catalysis. *Microbiology* **145** (Pt 11), 3023-3033.

50. Kaneko, T., Katoh, T., Sato, S., Nakamura, A., Asamizu, E., and Tabata, S. (2000) Structural Analysis of *Arabidopsis thaliana* Chromosome 3.11. Sequence Features of the 4,251,695 bp Regions Covered by 90 P1, TAC, and BAC Clones. *DNA Res.* **7**, 217-221.

51. Wang, S.P., Shuman, S. (1997) Structure Function Analysis of the mRNA Cap Methyltransferase of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **272**, 14683-14689.

52. Saha, N., Schwer, B., and Shuman, S. (1999) Characterization of Human, *Schizosaccharomyces pombe*, and *Candida albicans* mRNA Cap Methyltransferase and Complete Replacement of the Yeast Capping Apparatus by Mammalian Enzymes. *J. Biol. Chem.* **274**, 16553-16562.

53. Cheng, X., Kumar, S., Pasfai, J., Pfulgrath, J.W., and Roberts, R. (1993) Crystal Structure of the HhaI DNA Methyltransferase Complexed with S-Adenosyl-L-Methionine. *Cell* **74**, 299-307.

54. Bautz, D.J., Simms, B.L., Huggins, J.D., Raulfs, E.C., and Sitz, T.O. (1999) Inactivation of Guanine-7-Methyltransferase with N-ethylmaleimide (NEM): Essential Sulfhydrals? *Virginia J. Sci.* **50**, 113.

55. Gregory, J.D. (1955) The Stability of N-ethylmaleimide and its Reaction with Sulfhydral Groups. *J. Am. Chem. Soc.* **77**, 3922.

56. Subbaramaiah, K., Charles, H., and Simms, S. (1991) Probing the Role of Cysteine Residues in the CheR Methyltransferase. *J. Biol. Chem.* **266**, 19023-19027.

57. Everett, E., Falick, A.M., Reich, N.O. (1990) Identification of a Critical Cysteine in EcoRI DNA Methyltransferase by Mass Spectrometry. *J. Biol. Chem.* **265**, 17713-17719.

58. Schwer, B., Saha, N., Mao, X., Chen, H.W., Shuman, S. (2000) Structure-Function Analysis of Yeast mRNA Cap Methyltransferase and High Copy Suppression of Conditional Mutants by AdoMet Synthesis and the Ubiquitin Conjugating Enzyme Cdc34p. *Genetics* **155**, 1561-1576.

59. Tsukamoto, T., Shibagaki, Y., Niikura, Y., and Mizumoto, K. (1998) Cloning and Characterization of Three Human cDNAs Encoding mRNA (Guanine-7-) Methyltransferase, an mRNA Cap Methyltransferase. *Biochem. Biophys. Res. Comm.* **251**, 27-34.

X. CURRICULUM VITAE

David J. Bautz

EDUCATION
Master of Science, Biochemistry, May 2001
5-year Bachelor/Master's Degree Program
Virginia Polytechnic Institute and State University, Blacksburg, VA
Overall GPA: 3.6/4.0

Bachelor of Science, Biochemistry, May 2000
Virginia Polytechnic Institute and State University, Blacksburg, VA
Overall GPA: 3.5/4.0

LAB SKILLS

Immunoblotting	Cell Culture
Kinetic Assays	pH Calibration
RNA Isolation	Buffer Preparation
SDS-PAGE	Maintenance of Lab Notebook
Affinity Chromatography	Radioactive Labeling

COMPUTER SKILLS
Windows 95/98 BLAST Sequence Homology Search
Microsoft Word/Excel/Power Point

PRESENTATIONS
D. J. Bautz, K. Newbrough and T. O. Sitz "*Characterization of Human and Mouse Guanine-7-Methyltransferase*" May 2000, 78th Annual Virginia Academy of Science Meeting, Radford, VA

D. J. Bautz, B. L. Simms, J. D. Huggins, E. C. Raulfs and T. O. Sitz "*Inactivation of Guanine-7-Methyltransferase with N-Ethylmaleimide (NEM): Essential Sulfhydrals?*" May 1999, 77th Annual Virginia Academy of Science Meeting, Norfolk, VA

HONORS AND AWARDS
Best Student Paper Award, Chemistry Section, 1999, 2000 VAS Meeting
American Institute of Chemists Foundation 2000 Baccalaureate Student Award in Biochemistry
John Lee Pratt Senior Research Scholarship, 1999-2000
Phi Kappa Phi Honor Society, 1999
Dean's List: Fall '96, Spring '98 – Spring '00