Comparison of Cobalamin-Independent and Cobalamin-Dependent Methionine Synthases from *Escherichia coli*: Two Solutions to the Same Chemical Problem[†,‡]

Julio C. González, Ruma V. Banerjee,[§] Sha Huang, James S. Sumner, and Rowena G. Matthews*

Biophysics Research Division and Department of Biological Chemistry, The University of Michigan, Ann Arbor, Michigan 48109

Received February 12, 1992; Revised Manuscript Received March 31, 1992

ABSTRACT: In *Escherichia coli*, two enzymes catalyze the synthesis of methionine from homocysteine using methyltetrahydrofolate as the donor of the required methyl group: cobalamin-dependent and cobalamin-independent methionine synthases. Comparison of the mechanisms of these two enzymes offers the opportunity to examine two different solutions to the same chemical problem. We initiated the research described here to determine whether the two enzymes were evolutionarily related by comparing the deduced amino acid sequences of the two proteins. We have determined the nucleotide sequence for the *metE* gene, encoding the cobalamin-independent methionine synthase. Our results reveal an absence of similarity between the deduced amino acid sequences of the cobalamin-dependent and cobalamin-independent proteins and suggest that the two have arisen by convergent evolution. We have developed a rapid one-step purification of the recombinant cobalamin-independent methionine synthase (MetE) that yields homogeneous protein in high yield for mechanistic and structural studies. In the course of these studies, we identified a highly reactive thiol in MetE that is alkylated by chloromethyl ketones and by iodoacetamide. We demonstrated that alkylation of this residue, shown to be cysteine 726, results in complete loss of activity. While we are unable to deduce the role of cysteine 726 in catalysis at this time, the identification of this reactive residue suggests the possibility that this thiol functions as an intermediate methyl acceptor in catalysis, analogous to the role of cobalamin in the reaction catalyzed by the cobalamin-dependent enzyme.

In *Escherichia coli*, two distinct enzymes catalyze the terminal step in the de novo biosynthesis of methionine (Foster et al., 1961). Cobalamin-dependent methionine synthase (MetH, EC 2.1.1.13) is encoded by the *metH* gene and contains a cobalamin prosthetic group that is required for activity. Cobalamin-independent methionine synthase (MetE, EC 2.1.1.14) is encoded by the *metE* gene and has no known requirement for a vitamin-derived prosthetic group. Both enzymes catalyze essentially the same overall reaction, in which a methyl group from 5-methyltetrahydrofolate (CH_3-$H_4PteGlu_n$)[1] is transferred to the thiolate of homocysteine to generate methionine and tetrahydrofolate ($H_4PteGlu_n$) (eq 1).

$$CH_3\text{-}H_4PteGlu_n + \text{homocysteine (RSH)} \rightarrow H_4PteGlu_n + \text{methionine } (RSCH_3) \quad (1)$$

Comparison of the mechanisms and structures of these two enzymes from the same organism offers the opportunity to examine two different solutions to the same chemical problem. We initiated the research described here to determine whether the two enzymes were evolutionarily related and whether the deduced amino acid sequences of the two enzymes showed motifs that reflected the fact that both enzymes bind methyltetrahydrofolate derivatives and homocysteine. We were also interested in developing a rapid purification of the MetE enzyme that would permit characterization of the structure and catalytic mechanism of this enzyme. The availability of recombinant cobalamin-dependent enzyme (Banerjee et al., 1989) has permitted the isolation of large amounts of this protein for mechanistic and structural studies, and an X-ray crystallographic analysis of the structure of the cobalamin binding domain is now in progress (Luschinsky et al., 1992). In contrast, very little is known about the structure and catalytic mechanism of the cobalamin-independent enzyme.

The catalytic mechanism of MetH has been extensively studied in the laboratories of Weissbach, Taylor, and Huennekens and in our own laboratory [reviewed in Banerjee and Matthews (1990); more recent studies are described in Banerjee et al. (1990a,b)]. These studies have established that the cobalamin-dependent enzyme cycles in catalysis between methylcobalamin and cob(I)alamin. In methylcobalamin, the cobalt in the cobalamin prosthetic group is directly bonded to the methyl group, while in cob(I)alamin, the electrons that formed the bond with the methyl group are left on the cobalamin in an orbital perpendicular to the plane of the corrin ring. Both enzyme-bound methylcobalamin and cob(I)alamin have been shown to be kinetically competent intermediates (Banerjee et al., 1990a). Cob(I)alamin is one of the most potent nucleophiles known, with a reactivity toward methyl iodide that is ~30000-fold greater than that of thiols (Brown, 1982), and this reactivity may be essential for its function as a nucleophile in the displacement of the methyl group from

[†] This research has been supported in part by Research Grant R37 GM24908 from the National Institute of General Medical Sciences, National Institutes of Health. J.C.G. is a trainee in the Medical Scientist Training Program at The University of Michigan funded by Grant GM 07863 and is supported by a MARC predoctoral fellowship from the National Institute of General Medical Sciences (GM14330).

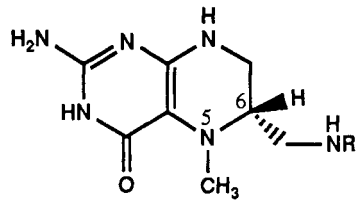
[‡] The nucleotide sequence has been submitted to GenBank under Accession Number M87625.

[§] Present address: Department of Biochemistry, East Campus, The University of Nebraska, Lincoln, NB 68583-0718.

[1] Abbreviations: CH_3-$H_4PteGlu_n$, 5-methyltetrahydropteroylpolyglutamate with *n* glutamyl residues; $H_4PteGlu_n$, tetrahydropteroylpolyglutamate with *n* glutamyl residues; MetE, cobalamin-independent methionine synthase, the product of the *metE* gene (EC 2.1.1.14); MetH, cobalamin-dependent methionine synthase, the product of the *metH* gene (EC 2.1.1.13); $PteGlu_3$, pteroyltriglutamate; TLCK, N^{α}-*p*-tosyl-L-lysine chloromethyl ketone; TPCK, *N*-tosyl-L-phenylalanine chloromethyl ketone; LysC, lysyl endopeptidase or *Achromobacter* protease I.

0006-2960/92/0431-6045$03.00/0

methyltetrahydrofolate (**1**). This reaction requires the removal



CH_3-H_4folate (1)

of a methyl substituent of a tertiary amine. If methyltetrahydrofolate is activated for methyl transfer by protonation at N^5, the incoming nucleophile must distinguish between the hydrogen and alkyl substituents. Cob(I)alamin has chemical properties that make it suitable for such a role, in that it is not only an excellent nucleophile but also an extremely weak base, with a pK_a of 1.0 for the protonation of the d_z^2 orbital that is perpendicular to the plane of the corrin ring (Lexa & Saveant, 1983). The carbon–cobalt bond in the intermediate, methylcobalamin, is weak, with a homolytic bond dissociation energy of ~37 kcal/mol (Martin & Finke, 1990). Although the stability of this bond to homolytic and heterolytic cleavage is probably not identical, the overall weakness of the carbon–cobalt bond may make the nucleophilic attack of homocysteine on methylcobalamin relatively facile.

MetE, unlike MetH, is unable to accept the monoglutamyl derivative of methyltetrahydrofolate (CH_3-H_4PteGlu$_1$) and is unique among folate-dependent enzymes in its absolute requirement for a polyglutamylated folate substrate (Foster et al., 1964). The properties of a homogeneous preparation of MetE were initially described by Whitfield et al. (1970). In addition to requiring a poly(glutamate) folate substrate, MetE is further distinguished by a dependence of activity on phosphate anions and divalent cations such as magnesium or magnanese. In comparison to the cobalamin-dependent enzyme, which has a turnover number of 1500 min^{-1}, the MetE enzyme is catalytically sluggish, with a turnover number of 12.3 min^{-1}. However, very little is known about the catalytic mechanism involved in the methyl-transfer reaction catalyzed by MetE, and in particular, the roles of the phosphate anion and Mg^{2+} or Mn^{2+} have not been elucidated. The mechanism of activation of the methyl substituent of methyltetrahydrofolate in this reaction is also of interest. If the cobalamin-dependent enzyme requires a strong nucleophile that is only very weakly basic to displace the methyl group from protonated methyltetrahydrofolate, how is such a nucleophile generated in the absence of cobalamin? Or is the chemical strategy for activation entirely different?

In this paper, we report the development of a rapid one-step purification of recombinant MetE that provides high yields of homogeneous protein and that will greatly facilitate subsequent mechanistic and structural studies. Nucleotide sequence analysis of the *metE* gene reveals an absence of similarity between the deduced amino acid sequences of the MetE and MetH proteins and suggests that the two proteins have arisen by convergent evolution. In the course of these studies, we discovered an extremely reactive thiol in MetE and demonstrated that alkylation of this residue, shown to be cysteine 726, results in complete loss of activity. While we are unable to deduce the role of cysteine 726 in catalysis at this time, the identification of this reactive residue suggests the possibility that this thiol functions as an intermediate methyl acceptor in catalysis, analogous to the role of cobalamin in the reaction catalyzed by MetH. The thiol group at the active site of papain provides a precedent for the role of an enzyme in stabilizing a highly reactive thiol that is also very weakly basic (Polgar, 1973; Lewis et al., 1976), and our preliminary results suggest that the reactive thiol in MetE may have similar properties.

Materials and Methods

Materials. Lysozyme, ribonuclease, T4 DNA ligase, and restriction enzymes were obtained from Bethesda Research Laboratories. Deoxyadenosine 5′-[α-^{35}S]thiotriphosphate (1000 Ci/mmol), [^{14}C]formaldehyde, and [1-^{14}C]iodoacetamide were obtained from Amersham. PteGlu$_3$ was purchased from Schircks Laboratories. [^{14}C]-CH_3-H_4PteGlu$_3$ was synthesized as described previously (Matthews, 1986). The following items were obtained from Sigma: sodium ampicillin, dithiothreitol, phenylmethanesulfonyl fluoride, N^{α}-*p*-tosyl-L-lysine chloromethyl ketone (TLCK), *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK), DEAE-Sepharose (fast flow), trypsin (treated with TPCK), iodoacetamide, and iodoacetic acid. Lysyl endopeptidase (LysC or *Achromobacter* protease I) was obtained from Wako Laboratories, Richmond, VA.

Bacterial Strains and Plasmids. The *E. coli* strain RR1/pRSE562 was obtained from Herbert Weissbach at the Roche Institute of Molecular Biology. Strain DH5αF′ and plasmid pGEM3B were obtained from Bethesda Research Laboratories and Promega, respectively.

Subcloning and Nucleotide Sequence Analysis of the metE Gene. Plasmid pRSE562 contains genes encoding both cobalamin-independent methionine synthase (MetE) and a regulatory protein required for efficient transcription of the *metE* gene (the *metR* gene product); its construction has been described by Maxon et al. (1989). The *metE* gene was subcloned into the sequencing vector pGEM3B to decrease the amount of flanking DNA prior to sequencing (Figure 1). pRSE562 was digested with *Eco*RI and *Sma*I, and the resulting fragments were ligated to pGEM3B that had been cleaved with the same restriction enzymes. DNA from the ligation mixture was employed to transform *E. coli* strain DH5αF′, and the resulting transformants were screened for resistance to ampicillin. Of the numerous ampicillin-resistant colonies that were identified, 49 were screened by the colony hybridization method with an oligomer that was complementary to the 5′ end of the *metE* gene. The first 65 bases of the MetE coding sequence have been reported by Maxon et al. (1989), and the complementary oligomer GCC TGC GTC GCG AGC TGA was used for colony hybridization. Positive clones could represent either uncut parent plasmid pRSE562 or a subclone carrying a 2.5-kb insert containing the *metE* gene. Ten positive clones were selected by rapid sizing as described in the Erase-A-Base manual (Promega). The population of uncut parent plasmid (9.6 kb) could easily be distinguished from the subclones of interest (5.3 kb). Of the three subclones identified, pME6 was employed for further characterization and sequence analysis.

Plasmid pHS1 contains a *Pst*I–*Bam*HI fragment carrying the *metE* and *metR* genes from pRSE562 inserted into the same sites in the multiple cloning region of pGEM3B. After pGEM 3B was cleaved with *Pst*I and *Bam*HI, the cut plasmid was treated with alkaline phosphatase from calf intestine to prevent religation with the released oligonucleotide from the multiple cloning region. The *Pst*I–*Bam*HI fragment from pRSE562 was then ligated to the treated vector using T4 DNA ligase. The construction was verified by restriction mapping with *Sma*I, *Pst*I, and *Bam*HI.

The strategy employed for nucleotide sequence analysis is shown in Figure 2. Double-stranded nucleotide sequence analysis was accomplished by the dideoxy chain-terminating

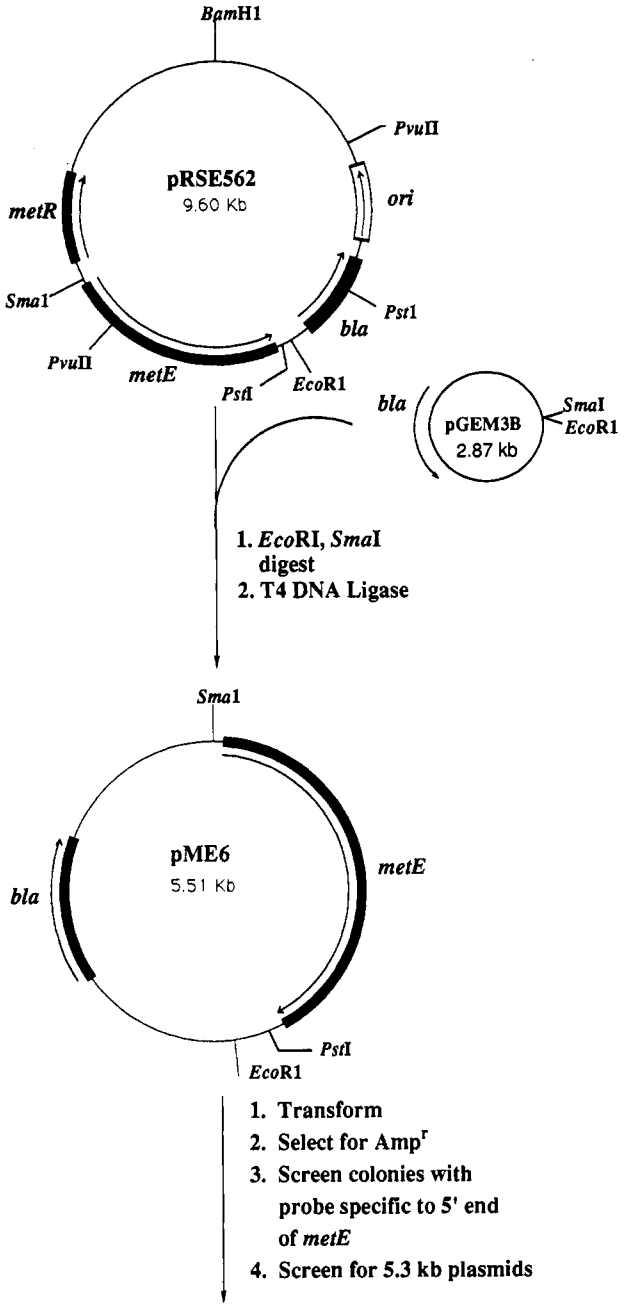


FIGURE 1: Subcloning strategy for sequencing MetE. The restriction map of pRSE562 is based on the work of Maxon et al. (1989). Details of the subcloning procedure are presented in the Results section.

method (Sanger et al., 1977), using deoxyadenosine 5'-[α-^{35}S]thiotriphosphate as recommended by the manufacturer of the Sequenase kit (United States Biochemicals). The sequence of both strands was obtained by primer extension, using synthetic primers (17-mers) that were purchased from Oligos Etc. (Ridgefield, CT). Primer sites were spaced at ~180-bp intervals. The sequences of both strands were completely determined, including 588 bases downstream of the protein coding sequence. The sequence of the region upstream of the *metE* gene was previously determined by Maxon et al. (1989) and was not redetermined in this study. Compilation of sequence overlap was accomplished with the MacVector program. Codon translation, estimate of p*I*, and determination of open reading frames were done using the same program. Codon preference analysis based on the algorithm derived by Gribskov et al. (1984) was plotted with programs in the GCG sequence analysis software package, version 7.0 (Devereaux et al., 1984). The Swiss protein database (release 19.0) was searched with FastA using the GCG package, while the deduced amino acid sequence of MetE was compared with translations of GenBank Release 69.0 using TFastA in the GCG package. Sequence comparisons of MetE and MetH and of the N- and C-terminal halves of MetE were performed in the GCG package using Bestfit, and the statistical significance of the matches was assessed by comparing the quality score for the match with the quality score derived from 10 random permutations of one of the sequences being compared. A *Z* value was then calculated for the comparison using the formula shown in eq 2.

$$Z = [(\text{quality score for best-fit match}) - (\text{mean quality score for 10 randomized sequences})]/(\text{standard deviation of randomized quality scores}) \quad (2)$$

Enzyme Assay and Purification of Recombinant MetE. The enzymatic assay for MetE employs a modification of the protocol described by Whitfield et al. (1970). The reaction mixture contained 66 μM (6*S*)-[*methyl*-5-^{14}C]-CH_3-H_4PteGlu$_3$ (22 000 dpm/nmol), 2 mM L-homocysteine, 10 mM potassium phosphate buffer, pH 7.2, 100 μM $MgSO_4$ or $MnSO_4$, 10 mM dithiothreitol, and enzyme in a total volume of 50 μL. The assay mixture lacking CH_3-H_4PteGlu$_3$ was incubated at 37 °C for 2 min, and the assay was initiated by the addition of radiolabeled CH_3-H_4PteGlu$_3$. After incubation at 37 °C for 5 min, the reaction was terminated by the addition of 0.95 mL of ice-cold water. In our hands, this assay was found to be linear over a 5-min time course. The labeled methionine product was separated from residual labeled CH_3-H_4PteGlu$_3$ by chromatography on a 0.5 × 6 cm column containing Bio-Rad AG1-X8 (chloride form). The eluate was collected in a scintillation counting vial containing 10 mL of Ecolite+ scintillation cocktail (ICN Biochemicals). The assay tube was rinsed with two 1-mL aliquots of distilled water, and

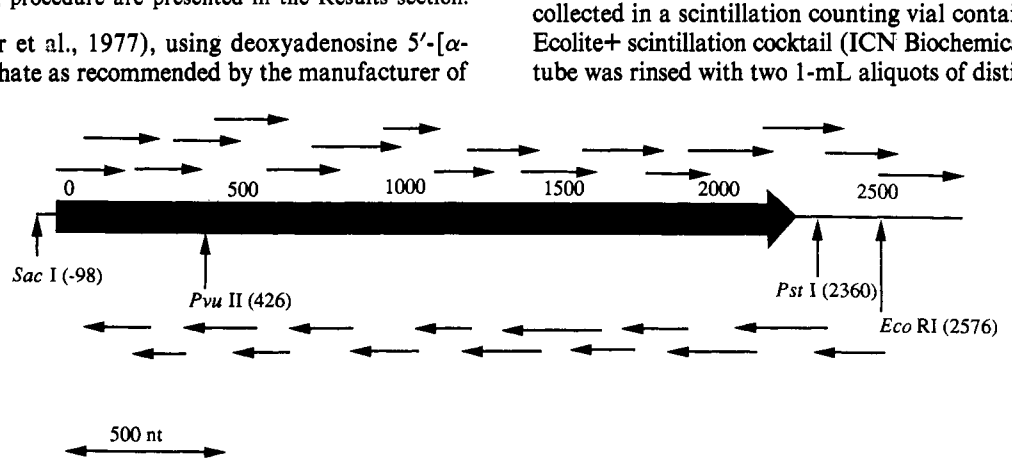


FIGURE 2: Nucleotide sequence analysis. The entire insert in pME6 is shown. From the data in Maxon et al. (1989), it is known that *metR* is divergently transcribed with respect to *metE*, with 238 nucleotides in the intragenic region, and that the *Sma*I site is located 98 base pairs upstream of the origin of translation of *metE*.

the rinses were added to the vial. Samples were counted in a Beckman LS7500 liquid scintillation counter. Values for enzyme activity have been corrected by subtraction of the counts obtained when an identical assay lacking enzyme was analyzed by the same procedure.

MetE was purified from *E. coli* strain DH5αF′/pRSE562. The cells were grown in 6 L of M9 minimal medium (Sambrook et al., 1989) supplemented with 0.4% glucose, 10 μM thiamin, and 100 μg/mL ampicillin. Cells of this strain grow very slowly, with a doubling time of ~3 h at 37 °C, and the slow growth is probably related to gross overproduction of MetE. The cells were harvested at an A_{420} of 4–4.2 by chilling and centrifuging at 12000*g* for 10 min. After being washed with cold 180 mM potassium phosphate buffer, pH 7.2, the cell pellet (~16 g) was suspended in 80 mL of the same buffer. The cells were disrupted with a Branson sonifier, Model 185, using an output setting of 7. To prevent overheating, the beaker containing the cell paste was immersed in an ice water bath and sonicated with four 1-min bursts alternated with 2-min breaks between cycles. The suspension was centrifuged for 1 h at 100000*g* to remove cell debris and unbroken cells. The specific activity of enzyme in the supernatant is ~0.03 μmol min^{-1} mg^{-1}, and densitometric scanning of a gel obtained after electrophoresis of the sonicate supernatant under denaturing conditions indicates that MetE constitutes ~40% of the total soluble protein in the sonicate. Initial studies indicated that the use of the serine protease inhibitors TLCK or TPCK led to loss of enzyme activity, and so it was necessary to devise a rapid purification procedure that would minimize proteolysis during the purification. The supernatant was loaded onto a DEAE-Sepharose column (2.2 × 50 cm) equilibrated with 180 mM potassium phosphate buffer, pH 7.2. The column was washed with 100 mL of the same buffer, supplemented with 500 μM dithiothreitol, and then eluted with two consecutive 400-mL linear gradients from 180 to 340 mM buffer/500 μM dithiothreitol and then from 340 to 500 mM buffer/500 μM dithiothreitol. Fractions were monitored for enzyme activity using the assay described above, and protein contents were determined by the method of Bradford (1976) using Bio-Rad protein assay and bovine serum albumin as the standard protein. This chromatographic step yielded enzyme that was homogeneous as judged by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate and had a specific activity of 0.15 μmol min^{-1} mg^{-1}, quite comparable to the specific activity of 0.17 μmol min^{-1} mg^{-1} reported for the homogeneous enzyme by Whitfield et al. (1970). Enzyme-containing fractions were pooled and concentrated under nitrogen in an Amicon ultrafiltration cell fitted with a PM30 membrane. The concentrated protein solution was desalted by washing twice with 5 volumes of 10 mM potassium phosphate buffer/500 μM dithiothreitol, pH 7.2, in Centricon 30 microconcentrators (Amicon). Purified enzyme was stored at −80 °C.

Quantitation of Protein Bands Seen during Polyacrylamide Gel Electrophoresis. Gels were scanned and digitized using a cold charge-coupled device (CCD) camera (Star 1, Photometrics). The images were converted to optical densities on a Silicon Graphics IRIS workstation. Integrated band intensities were calculated using the EMPRO image processing package. The procedures and equipment are described by Vincenz et al. (1991).

Procedures for Alkylation of Enzyme and Monitoring of Activity Loss during Alkylation. Enzyme (38 nmol) was incubated at 37 °C in 50 μL of 20 mM potassium phosphate buffer, pH 7.2, containing the alkylating agent. The alkylating agent was added to initiate the reaction. A cocktail containing all the ingredients for assay of MetE activity was aliquoted into individual tubes, and at intervals throughout the reaction, 4-μL samples of enzyme were removed from the alkylation reaction and diluted into the assay cocktail to initiate the enzymatic assay. Activities were determined as described above. For labeling, [1-^{14}C]iodoacetamide (100000 dpm/nmol, 200 μM final concentration) was used. For analysis of the incorporation of covalent label, an aliquot of the reaction mixture was quenched by addition of dithiothreitol (10 mM final concentration), and noncovalently bound label was removed by washing in a Centricon 30 microconcentrator prior to counting.

Location of Radiolabeled Peptides in Alkylated MetE. Enzyme (39 nmol), either untreated or pretreated with 200 μM TPCK for 60 min, was alkylated with radiolabeled iodoacetamide as described above. A 2-nmol portion of the radiolabeled enzyme in 6 M urea/200 mM Tris-HCl buffer, pH 8.0, was digested with 1 μg of lysyl endopeptidase (Lys C) for 4 h at 37 °C. Separation of peptides following LysC digestion was accomplished by reversed-phase chromatography on a 25 × 0.21 cm Vydac C18 column, using a gradient of acetonitrile/0.1% trifluoroacetic acid as solvent B and water/0.1% trifluoroacetic acid as solvent A. A series of linear gradients were employed: 0–5% solvent B over 3 min, 5–60% solvent B over 57 min, and finally 60–85% solvent B over 10 min.

Peptide Sequence Analysis. The N-terminal sequences of the enzyme and of peptides generated by lysyl endopeptidase (LysC) digestion were determined in a Model 473 Applied Biosystems liquid-phase sequencer at The University of Michigan Protein Sequencing Facility.

Results and Discussion

Nucleotide Sequence Analysis of the metE Gene. Figure 3 presents the results of nucleotide sequence analysis of the insert in plasmid pME6 containing the *metE* gene. An open reading frame extending for 2259 bases and starting with an ATG initiation codon was found. This reading frame encodes a protein of 753 amino acids with a predicted molecular weight of 84654. The predicted molecular weight is in good agreement with the value of 84000 obtained by ultracentrifugation of the native enzyme (Whitfield et al., 1970) and the value of 88000 obtained by electrophoresis in the presence of sodium dodecyl sulfate under the standard electrophoretic conditions used to assign polypeptides in the *E. coli* database (VanBogelen et al., 1990). This value is, however, slightly lower than the molecular weight of 93000 estimated in our laboratory by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. Codon usage probability is high in only one of the three frames and supports the reading frame assignment deduced by nucleotide sequence determination. It yields a codon preference statistic of 1.02, based on the algorithm derived by Gribskov et al. (1984) and implemented with a program in the GCG sequence analysis software package (Devereaux et al., 1984). This value is only slightly higher than the codon preference statistic of 0.929 for the MetH protein (Banerjee et al., 1989). The relatively low codon preference statistic for MetE is somewhat surprising, because this enzyme is extremely abundant during growth of *E. coli* in the absence of cobalamin and methionine and represents almost 3% of the total soluble protein in the cell (Whitfield et al., 1970). In contrast, MetH represents only ~0.1% of the total soluble protein when cells are grown in glucose minimal medium to which cobalamin has been added (Banerjee et al., 1989). A region of dyad symmetry is present 10

bases downstream from the termination codon and is followed by a stetch of four T's, suggesting ρ-independent termination of transcription (Yager & von Hippel, 1987). This feature has been indicated in Figure 3 by underlining of the associated nucleotide sequence at the 3′ end of the gene. It is estimated that formation of a stem–loop structure in this region would provide ~25.8 kcal of stabilization energy at 25 °C (Tinoco et al., 1973).

The N-terminal sequence of the protein deduced by Edman degradation of the purified recombinant protein yielded a mixture of methionine and threonine in the first cycle, isoleucine in the second cycle, and thereafter the sequence predicted from the deduced amino acid sequence of the protein. Threonine is a low-yielding amino acid, and its presence as a minor peak in the first cycle, and virtual absence in the second, suggests that the majority of the MetE protein has been processed posttranslationally to remove the N-terminal methionine. Threonine is one of the most common N-terminal amino acids in prokaryotic polypeptides (Sarimo & Pine, 1969). The N-terminal sequence establishes the translational start site of the *MetE* gene and confirms the verity of the initiator codon identified by Maxon et al. (1989). Amino acid sequences of other internal peptides obtained from digestion of the native enzyme with lysine endoprotease are also identical to the inferred amino acid sequence shown in Figure 3 and have been indicated by underlining.

Because we were concerned about the discrepancy between the molecular weight of 84 654 for MetE estimated from the deduced amino acid sequence and the molecular weight of ~93 000 estimated by gel electrophoresis, we excised a *Pst*I–*Bam*HI fragment from pRSE562 and ligated it into pGEM3B. As indicated in Figure 2, the *Pst*I site is located at nt 2360, 101 nt downstream of the estimated termination of translation of MetE, and one *Pst*I–*Bam*HI fragment is expected to contain the complete sequences of both the *metE* and *metR* genes. If the *metE* gene were to encode a protein of 93 000 molecular weight, the coding sequence would have to extend more than 200 nt beyond the estimated termination site, and we would expect the *Pst*I–*Bam*HI fragment to express a truncated protein. When cells of strain DH5αF′ were transformed with the pGEM-derived plasmid containing the *Pst*I–*Bam*HI insert, they expressed an active MetE protein that was identical in its mobility on SDS–PAGE to the MetE protein expressed in cells transformed with pRSE562.

Using TFastA, we compared the deduced amino acid sequence of the *metE* gene with translations of the nucleic acid sequences in release 69.0 of GenBank. A highly significant homology (43% identity over a 486 amino acid overlap) was found with the sequence of the P8 promoter from *Saccharomyces cerevisiae* (Ohtake et al., 1988; GenBank Accession Numbers D00365 and X07238). This promoter and its associated coding sequence for the amino terminus of the P8 protein were cloned by screening for their ability to stimulate β-galactosidase expression when cloned upstream of a promoterless *lacZ* gene. The deduced amino acid sequence is of a fusion protein in which amino acid 477 of the P8 gene has been linked with an *Xho*I linker to the glutathione synthetase protein from *E. coli* and does not contain the sequence from the 3′ end of the P8 gene. The extensive homology observed between the P8 coding sequence and the deduced amino acid sequence of the *metE* gene strongly suggests that the P8 gene in *S. cerevisiae* encodes cobalamin-independent methionine synthase in this yeast. Indeed, cobalamin-independent methionine synthase has been purified extensively from *S. cerevisiae* (Burton & Sakami, 1969) and is present in large amounts (~300 mg/lb of commercial bakers' yeast). No other significant homologies were observed in either the GenBank/EMBL or Swiss Protein databases. Attempts to align the deduced amino acid sequences of the *metE* and *metH* genes using the Bestfit program in the GCG sequence analysis software package failed to indicate significant sequence similarities ($Z = 0$).

An interesting property of the *metE* gene emerged when we compared the N- and C-terminal halves of the deduced amino acid sequence. This comparison strongly suggested that the *metE* gene has arisen by internal gene duplication. We used the Align program in MacVector to look for similarities in the two halves of the *metE* gene. The diagonal matrix shown in Figure 4A compares the deduced amino acid sequence of residues 1–369 with that of residues 370–753. Figure 4B shows an alignment of the deduced amino acid sequences, with identical residues shown in bold type. There are 90 identities, or ~24% identity, in the compared sequences. Of the seven cysteinyl residues in the MetE sequence, only one is conserved in the alignment, and that is C726, which aligns opposite C324. In the immediate vicinity of C726, spanning residues 720–728, five amino acids are identical when the N- and C-terminal halves of the protein are aligned. C726 is located in a region that is predicted to be moderately hydrophobic and of low surface probability (data not shown). The significance of the alignment was determined using the Bestfit program of the GCG sequence analysis software package (Devereaux et al., 1984). This program compares the alignment of 10 random amino acid sequences having the same length and composition as residues 370–753 of MetE and calculates a quality score for each alignment that includes both identities and similarities in the compared amino acid sequences. The quality score for alignment of the deduced amino acid sequences was 136.6. Comparisons of the N-terminal half of MetE to 10 randomized permutations of the C-terminal MetE sequence gave an average quality score of 113.5 ± 4.9. The difference between the authentic alignment and alignment to randomized sequences is 4.7 times the standard deviation of the aligned randomized sequences. Thus, significant similarities exist between the N- and C-terminal halves of MetE, and these are especially apparent in the C-terminal two-thirds of each half. Figure 4B also compares the sequence of the N- and C-terminal halves of MetE with the deduced amino acid sequence of the P8 open reading frame. Identities between MetE and P8 have been underlined.

Tryptic Digestion of the Native MetE Enzyme. As described under Experimental Procedures, we have developed a rapid one-step purification of recombinant enzyme that enables us to prepare 75 mg of electrophoretically homogeneous enzyme from 1 L of cultured cells. The specific activity of the purified enzyme, 0.15 μmol min^{-1} mg^{-1}, is quite comparable to the values reported earlier for homogeneous enzyme by Whitfield et al. (1970). To determine whether spatially separate protein domains are present in the native enzyme, the recombinant enzyme was subjected to proteolysis with 0.5 μg of trypsin/mg of enzyme. The results of such an experiment are shown in Figure 5. The intact protein, which on this gel had an apparent molecular mass of ~92 kDa, is initially cleaved into two fragments of nearly equal size, ~48 and ~44 kDa. The 48-kDa band is further degraded until the two fragments migrate at equivalent positions. The ~44-kDa fragments are relatively stable but are further degraded on prolonged digestion. When attempts were made to assay activity as digestion proceeded, using TLCK to quench proteolysis by inactivation of trypsin, we found that undigested

```
        10          20          30          40          50          60          70          80          90
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
ATG ACA ATA TTG AAT CAC ACC CTC GGT TTC CCT CGC GTT GGC CTG CGT CGC GAG CTG AAA AAA GCG CAA GAA AGT TAT TGG GCG GGG AAC
Met Thr Ile Leu Asn His Thr Leu Gly Phe Pro Arg Val Gly Leu Arg Arg Glu Leu Lys Lys Ala Gln Glu Ser Tyr Trp Ala Gly Asn

       100         110         120         130         140         150         160         170         180
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
TCC ACG CGT GAA GAA CTG CTG GCG GTA GGG CGT GAA TTG CGT GCT CGT CAC TGG GAT CAA CAA AAG CAA GCG GGT ATC GAC CTG CTG CCG
Ser Thr Arg Glu Glu Leu Leu Ala Val Gly Arg Glu Leu Arg Ala Arg His Trp Asp Gln Gln Lys Gln Ala Gly Ile Asp Leu Leu Pro

       190         200         210         220         230         240         250         260         270
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GTG GGC GAT TTT GCC TGG TAC GAT CAT GTA CTG ACC ACC AGT CTG CTG CTG GGT AAC GTT CCG GCG CGT CAT CAG AAC AAA GAT GGT TCG
Val Gly Asp Phe Ala Trp Tyr Asp His Val Leu Thr Thr Ser Leu Leu Leu Gly Asn Val Pro Ala Arg His Gln Asn Lys Asp Gly Ser

       280         290         300         310         320         330         340         350         360
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GTA GAT ATC GAC ACC CTG TTC CGT ATT GGT CGT GGA CGT GCG CCG ACT GGC GAA CCT GCG GCG GCA GCG GAA ATG ACC AAA TGG TTT AAC
Val Asp Ile Asp Thr Leu Phe Arg Ile Gly Arg Gly Arg Ala Pro Thr Gly Glu Pro Ala Ala Ala Ala Glu Met Thr Lys Trp Phe Asn

       370         380         390         400         410         420         430         440         450
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
ACC AAC TAT CAC TAC ATG GTG CCG GAG TTC GTT AAA GGC CAA CAG TTC AAA CTG ACC TGG ACG CAG CTG CTG GAC GAA GTG GAC GAG GCG
Thr Asn Tyr His Tyr Met Val Pro Glu Phe Val Lys Gly Gln Gln Phe Lys Leu Thr Trp Thr Gln Leu Leu Asp Glu Val Asp Glu Ala

       460         470         480         490         500         510         520         530         540
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CTG GCG CTG GGC CAC AAG GTG AAA CCT GTG CTG CTG GGG CCG GTT ACC TGG CTG TGG CTG GGG AAA GTG AAA GGT GAA CAA TTT GAC CGC
Leu Ala Leu Gly His Lys Val Lys Pro Val Leu Leu Gly Pro Val Thr Trp Leu Trp Leu Gly Lys Val Lys Gly Glu Gln Phe Asp Arg

       550         560         570         580         590         600         610         620         630
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CTG AGC CTG CTG AAC GAC ATT CTG CCG GTT TAT CAG CAA GTG CTG GCA GAA CTG GCG AAA CGC GGC ATC GAG TGG GTA CAG ATT GAT GAA
Leu Ser Leu Leu Asn Asp Ile Leu Pro Val Tyr Gln Gln Val Leu Ala Glu Leu Ala Lys Arg Gly Ile Glu Trp Val Gln Ile Asp Glu

       640         650         660         670         680         690         700         710         720
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CCC GCG CTG GTA CTG GAA CTA CCA CAG GCG TGG CTG GAC GCA TAC AAA CCC GCT TAC GAC GCG CTC CAG GGA CAG GTG AAA CTG CTG CTG
Pro Ala Leu Val Leu Glu Leu Pro Gln Ala Trp Leu Asp Ala Tyr Lys Pro Ala Tyr Asp Ala Leu Gln Gly Gln Val Lys Leu Leu Leu


       730         740         750         760         770         780         790         800         810
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
ACC ACC TAT TTT GAA GGC GTA ACG CCA AAT CTC GAC ACG ATT ACT GCG CTG CCT GTT CAG GGT CTG CAT GTT GAC CTC GTA CAT GGT AAA
Thr Thr Tyr Phe Glu Gly Val Thr Pro Asn Leu Asp Thr Ile Thr Ala Leu Pro Val Gln Gly Leu His Val Asp Leu Val His Gly Lys

       820         830         840         850         860         870         880         890         900
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GAT GAC GTT GCT GAA CTG CAC AAG CGC CTG CCT TCT GAC TGG TTG CTG TCT GCG GGT CTG ATC AAT GGT CGT AAC GTC TGG CGC GCC GAT
Asp Asp Val Ala Glu Leu His Lys Arg Leu Pro Ser Asp Trp Leu Leu Ser Ala Gly Leu Ile Asn Gly Arg Asn Val Trp Arg Ala Asp

       910         920         930         940         950         960         970         980         990
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CTT ACC GAG AAA TAT GCG CAA ATT AAG GAC ATT GTC GGC AAA CGT GAT TTG TGG GTG GCA TCT TCC TGC TCG TTG CTG CAC AGC CCC ATC
Leu Thr Glu Lys Tyr Ala Gln Ile Lys Asp Ile Val Gly Lys Arg Asp Leu Trp Val Ala Ser Ser Cys Ser Leu Leu His Ser Pro Ile

      1000        1010        1020        1030        1040        1050        1060        1070        1080
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GAC CTG AGC GTG GAA ACG CGT CTT GAT GCA GAA GTG AAA AGC TGG TTT GCC TTC GCC CTA CAA AAA TGC CAT GAA CTG GCA CTG CTG CGC
Asp Leu Ser Val Glu Thr Arg Leu Asp Ala Glu Val Lys Ser Trp Phe Ala Phe Ala Leu Gln Lys Cys His Glu Leu Ala Leu Leu Arg

      1090        1100        1110        1120        1130        1140        1150        1160        1170
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GAT GCG GTG AAC AGT GGT GAC ACG GCA GCT CTG GCA GAG TGG AGC GCC CCG ATT CAG GCA CGT CGT CAC TCT ACC CGC GTA CAT AAT CCG
Asp Ala Val Asn Ser Gly Asp Thr Ala Ala Leu Ala Glu Trp Ser Ala Pro Ile Gln Ala Arg Arg His Ser Thr Arg Val His Asn Pro

      1180        1190        1200        1210        1220        1230        1240        1250        1260
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GCG GTA GAA AAG CGT CTG GCG GCG ATC ACC GCC CAG GAC AGC CAG CGT GCG AAT GTC TAT GAA GTG CGT GCT GAA GCC CAG CGT GCG CGT
Ala Val Glu Lys Arg Leu Ala Ala Ile Thr Ala Gln Asp Ser Gln Arg Ala Asn Val Tyr Glu Val Arg Ala Glu Ala Gln Arg Ala Arg

      1270        1280        1290        1300        1310        1320        1330        1340        1350
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
TTT AAA CTG CCA GCG TGG CCG ACC ACC ACG ATT GGT TCC TTC CCG CAA ACC ACG GAA ATT CGT ACC CTG CGT CTG GAT TTC AAA AAG GGC
Phe Lys Leu Pro Ala Trp Pro Thr Thr Thr Ile Gly Ser Phe Pro Gln Thr Thr Glu Ile Arg Thr Leu Arg Leu Asp Phe Lys Lys Gly

      1360        1370        1380        1390        1400        1410        1420        1430        1440
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AAT CTC GAC GCC AAC AAC TAC CGC ACG GGC ATT GCG GAA CAT ATC AAG CAG GCC ATT GTT GAG CAG GAA CGT TTG GGA CTG GAT GTG CTG
Asn Leu Asp Ala Asn Asn Tyr Arg Thr Gly Ile Ala Glu His Ile Lys Gln Ala Ile Val Glu Gln Glu Arg Leu Gly Leu Asp Val Leu


      1450        1460        1470        1480        1490        1500        1510        1520        1530
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GTA CAT GGC GAG GCC GAG CGT AAT GAC ATG GTG GAA TAC TTT GGC GAG CAC CTC GAC GGA TTT GTC TTT ACG CAA AAC GGT TGG GTA CAG
Val His Gly Glu Ala Glu Arg Asn Asp Met Val Glu Tyr Phe Gly Glu His Leu Asp Gly Phe Val Phe Thr Gln Asn Gly Trp Val Gln

      1540        1550        1560        1570        1580        1590        1600        1610        1620
    *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AGC TAC GGT TCC CGC TGC GTG AAG CCA CCG ATT GTC ATT GGT GAC ATT AGC CGC CCG GCA CCG ATT ACC GTG GAG TGG GCG AAG TAT GCG
Ser Tyr Gly Ser Arg Cys Val Lys Pro Pro Ile Val Ile Gly Asp Ile Ser Arg Pro Ala Pro Ile Thr Val Glu Trp Ala Lys Tyr Ala
```

```
      1630        1640        1650        1660        1670        1680        1690        1700        1710
CAA TCG CTG ACC GAC AAA CCG GTG AAA GGG ATG CTG ACG GGG CCG GTG ACC ATA CTC TGC TGG TCG TTC CCG CGT GAA GAT GTC AGC CGT
Gln Ser Leu Thr Asp Lys Pro Val Lys Gly Met Leu Thr Gly Pro Val Thr Ile Leu Cys Trp Ser Phe Pro Arg Glu Asp Val Ser Arg

      1720        1730        1740        1750        1760        1770        1780        1790        1800
GAA ACC ATC GCC AAA CAG ATT GCG CTG GCG CTG CGT GAT GAA GTG GCC GAT CTG GAA GCC GCT GGA ATT GGC ATC ATC CAG ATT GAC GAA
Glu Thr Ile Ala Lys Gln Ile Ala Leu Ala Leu Arg Asp Glu Val Ala Asp Leu Glu Ala Ala Gly Ile Gly Ile Ile Gln Ile Asp Glu

      1810        1820        1830        1840        1850        1860        1870        1880        1890
CCG GCC CTG CGC GAA GGT TTA CCG CTG CGT CGT AGC GAC TGG GAT GCG TAT CTC CAG TGG GGC GTA GAG GCC TTC CGT ATC AAC GCC GCC
Pro Ala Leu Arg Glu Gly Leu Pro Leu Arg Arg Ser Asp Trp Asp Ala Tyr Leu Gln Trp Gly Val Glu Ala Phe Arg Ile Asn Ala Ala

      1900        1910        1920        1930        1940        1950        1960        1970        1980
GTG GCG AAA GAT GAC ACA CAA ATC CAC ACT CAC ATG TGT TAT TGC GAG TTC AAC GAC ATC ATG GAT TCG ATT GCG GCG CTG GAC GCA GAC
Val Ala Lys Asp Asp Thr Gln Ile His Thr His Met Cys Tyr Cys Glu Phe Asn Asp Ile Met Asp Ser Ile Ala Ala Leu Asp Ala Asp

      1990        2000        2010        2020        2030        2040        2050        2060        2070
GTC ATC ACC ATC GAA ACC TCG CGT TCC GAC ATG GAG TTG CTG GAG TCG TTT GAA GAG TTT GAT TAT CCA AAT GAA ATC GGT CCT GGC GTC
Val Ile Thr Ile Glu Thr Ser Arg Ser Asp Met Glu Leu Leu Glu Ser Phe Glu Glu Phe Asp Tyr Pro Asn Glu Ile Gly Pro Gly Val

      2080        2090        2100        2110        2120        2130        2140        2150        2160
TAT GAC ATT CAC TCG CCA AAC GTA CCG AGC GTG GAA TGG ATT GAA GCC TTG CTG AAG AAA GCG GCA AAA CGC ATT CCG GCA GAG CGC CTG
Tyr Asp Ile His Ser Pro Asn Val Pro Ser Val Glu Trp Ile Glu Ala Leu Leu Lys Lys Ala Ala Lys Arg Ile Pro Ala Glu Arg Leu

      2170        2180        2190        2200        2210        2220        2230        2240        2250
TGG GTC AAC CCG GAC TGT GGC CTG AAA ACG CGC GGC TGG CCA GAA ACC CGC GCG GCA CTG GCG AAC ATG GTG CAG GCG GCG CAG AAC TTG
Trp Val Asn Pro Asp Cys Gly Leu Lys Thr Arg Gly Trp Pro Glu Thr Arg Ala Ala Leu Ala Asn Met Val Gln Ala Ala Gln Asn Leu

      2260        2270        2280        2290        2300        2310        2320        2330        2340
CGT CGG GGG TAA AAT CCA AAC CGG GTG GTA ATA CCA CCC GGT CTT TTC TCA TTA CAG CGA CTT CTT CCC ACC ATA CTG CTT AAA CCA TTC
Arg Arg Gly End

      2350        2360        2370        2380        2390        2400        2410        2420        2430
CAG CAT ACG CTG CCA GCC ATC TTC TGC AGA TGC GGC ATG ATA GCT CGG GCG ATA ATC AGC GTT GAA TGC ATG CCC GGC GTC CGG GTA CAC

      2440        2450        2460        2470        2480        2490        2500        2510        2520
GAT AAT CTC TGC TTT CGC ATT AGC AGC CCG CAG CGC CTG GCG CAT GGT TTC AAC GCT CTC CTG CGG AAT GCT GTT ATC CTG ACC ACC ATA

      2530        2540        2550        2560        2570        2580
TAA GCC GAG AAT CGG CGC GTT AAG ATC GGT TGC GAT ATC AAC AGG TTG TTT CGG TGA ATT C
```

FIGURE 3: Nucleotide sequence of the *metE* gene and the deduced amino acid sequence. The *metE* sequence is numbered so that A in the ATG initiaton codon is in position 1. The sequence upstream of the initiation codon was previously determined and is not shown here (Maxon et al., 1989). The amino acid sequence that has been confirmed by peptide sequencing is underlined. A possible ρ-independent terminator sequence downstream of the stop codon is underlined in the nucleotide sequence.

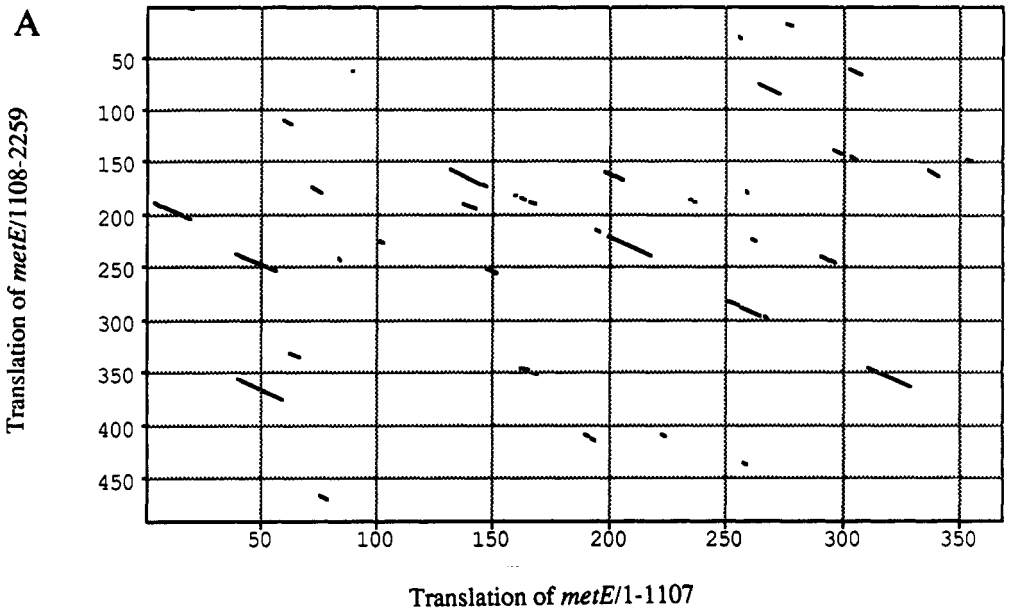
enzyme was inactivated by TLCK. Activity was therefore measured by quenching the aliquots on ice prior to enzymatic assay. The concentrations of the protein fragments as a function of time were determined by gel scanning. Figure 5 shows a comparison of the rate at which activity was lost with the rate of disappearance of the intact polypeptide determined by gel scanning. It can be seen that the cleavage of the intact polypeptide into two fragments is associated with partial, but not complete, loss of activity. The slower phase of activity loss appears to correlate with the loss of the ~44-kDa fragments.

These results corroborate the presence of two distinct structural domains in the MetE protein, as also inferred from analysis of the deduced amino acid sequence. Further experiments will be required to determine whether one or both of these domains show catalytic activity.

Inactivation of MetE by TLCK and Other Alkylating Agents. The observation that TLCK inhibits MetE was of considerable interest. Chloromethyl ketones are generally specific for serine proteases and serine esterases, where they alkylate the histidine in the catalytic triad (Shaw, 1970). Chloromethyl ketones are generally unreactive with thiol residues (Lee & McElroy, 1969) but do react with thiol residues in firefly luciferase (Lee & McElroy, 1969) and papain (Hinkle & Kirsch, 1970). Thus inactivation of MetE by TLCK suggested the presence of an unusually reactive residue in the active site of the enzyme. Accordingly, we surveyed a series of alkylating agents to determine the rate and extent of inactivation of MetE on exposure to these agents. The results of these experiments are summarized in Table I. We found that both TLCK and TPCK, at 200 μM concentrations, caused apparently complete inactivation of the enzyme on sufficiently long exposure but that the rate of inactivation of MetE in the presence of TPCK was markedly greater than the rate observed in the presence of TLCK. Iodoacetamide treatment was also found to result in complete inactivation of MetE, while iodoacetate treatment resulted in rapid partial loss of activity. Propyl iodide led to slow loss of activity, with a half-time of ~110 min. Since TLCK reacts less rapidly than TPCK, and since iodoacetamide treatment of MetE leads to

Table I: Inactivation of MetE by Alkylating Agents

compd	concn (μM)	half-time for inactivation (min)	% residual activity
TLCK	200	8.4	none detectable
TPCK	200	1	none detectable
iodoacetamide	200	30	none detectable
iodoacetate	200	not determined	65% at 60 min
propyl iodide	200	~110	not determined



B

```
                                                                    δP8/1-59
 MVQSAVLGFPRIGPNRELKKATEGYWNGKITVDELFKVGKDLRTQ-NWKLQKEAGVDIIP      δP8/1-59
MTILNHTLGFPRVGLRRELKKAQESYWAGNSTREELLAVGRELRAR-HWDQQKQAGIDLLP     MetE/1-60
 ALAEWSAPIQARRHSTRVHNPAVEKRLAAITAQDSQRANVYEVRAEAQRARFKLPAWPTTT     MetE/370-430
AALEANAKSVESRGKSKFIHDAAVKARVASIDEKMSTRAAPFEQRLPEQQKVENLPLFPTTT     δP8/383-443
                                                                    δP8/60-108
SNDFSFYDQVLDFVFVV--QCHS--------DRYT-KYDLSPIDTL--FAMGRGLQRKATE     δP8/60-108
VGDFAWYDHVLTTSLLL--GNVP--------ARHQNKDGSVDIDTL--FRIGRGRAPT---     MetE/61-106
IGSFPQTTEIRTLRLDFKKGNLDANNYRTGIAEHI-KQAIVEQERLGLDVLVHGEAERNDM     MetE/431-490
IGSFPQTKDIRINRNKFNKGTISAEEYEKFINSEIEK-VIRFQEDLEVIPDVSQALAW>     δP8/444-500
                                                                    δP8/109-163
ETEKAVDVTALEMVK--WFDSNYHY-VRPTESKTTQFKLNGQVPVDFFLEAKELGIH-TRP     δP8/109-163
----GEPAAAAEMTK--WFNTNYHYMV-PEFVKGQQFKLTWTQLLDEVDEALALGHK-VKP     MetE/107-159
VEYFGEHLDGFVFTQNGWVQSYGSRCVKPPIVIGDISRPAPIT-VEWAKYAQSLTDKPVKG     MetE/491-550
                                                                    δP8/164-225
VLLGPVSYLFLGKADKDSLDLEPLSLLEQLLPLYTEILSKLASAGATEVQIDEPVLVLDLPA     δP8/164-225
VLLGPVTWLWLGKVKGE--QFDRLSLLNDILPVYQQVLAELAKRGIEWVQIDEPALVLELPQ     MetE/160-219
MLTGPVTILCWSFPR-E--DVSRETIAKQIALALRDEVADLEAAGIGIIQIDEPALREGLPL     MetE/551-609
                                                                    δP8/226-278
N---AQAAI-KKAYTYFGEQSNL-----PKITLATYFGTVVPNLDAIKGLPVAALHVDFVRA     δP8/226-278
A---WLDAY-KPAYDALQGQV--------KLLLTTYFEGVTPNLDTITALPVQGLHVDLVHG     MetE/220-270
RRSDW-DAYLQWGVEAFRINAAVAKDDTQIHTHMCYCEFN-DIMDSIAALDADVITIETSRS     MetE/610-659
                                                                    δP8/279-334
PEQFDEVVAAIGNKQTLSVGI--VDGRNI----WKNDFKKSSAIVNKAIEKLGADRVVVATS     δP8/279-334
KDDVAELHKRLPSDWLLSAGL--INGRNV----WRADLTEKYAQI-KDI--VGKRDLWVASS     MetE/271-323
DMELLESFEEFDYPNEIGPGVYDIHSPNVPSVEWIEALLKKAA---KRI--PAER-LWVNPD     MetE/660-725
                                                                    δP8/335-382
SSLLHTPVDLNNETKLDAEIKGFFSFATQKLDEVVVITKNVSGQDVA                    δP8/335-382
CSLLHSPIDLSVETRLDAEVKSWFAFALQKCHELALLRDAVNSGDTA                    MetE/323-369
CGL---------KTRGWPETRAALANMVQAAQNLRRG                              MetE/726-753
```

FIGURE 4: (A) Alignment of N- and C-terminal halves of the MetE protein. Alignment was performed in MacVector with Pustell's protein and DNA matrices using a pam250S scoring matrix. The window size was 20, the minimum percent score was set at 20, and the hash value was set at 1. Residues 1–369 are shown aligned to residues 370–753. (B) Alignment of amino acids 1–369 with residues 370–753. After alignment in MacVector, manual gapping was performed to optimize the alignment. The numbers refer to the amino acid residues of the MetE protein. There are 90 identities, or ~24% identity in the compared sequences. Note that C726 and C324 align in this comparison. These are the only cysteinyl residues that are in identical positions in the aligned segments. Also shown in (B) is the alignment of the MetE protein with the deduced amino acid sequence of the open reading frame associated with the P8 promoter of *S. cerevisiae* (Ohtake et al., 1988). Identities between the deduced amino acid sequences of MetE and the P8 open reading frame are underlined.

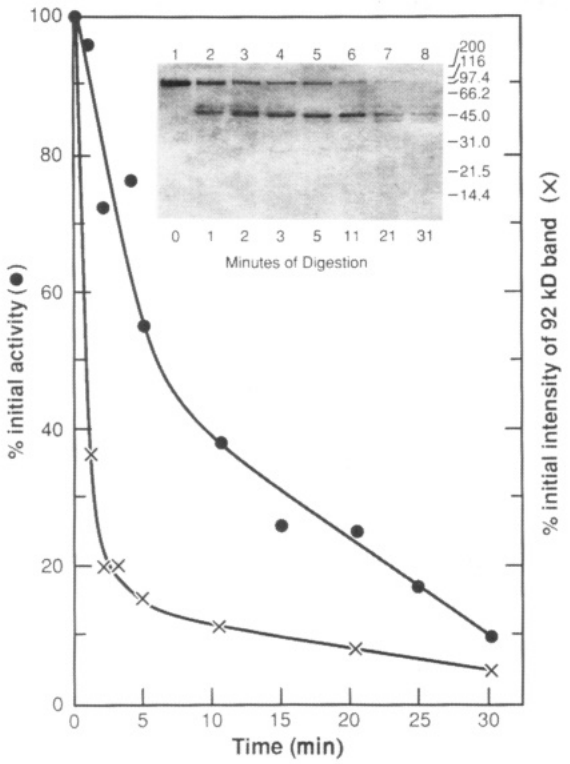


FIGURE 5: Tryptic digestion of the native MetE protein. Aliquots (50 μL) containing 11 nmol of MetE in 20 mM potassium phosphate buffer, pH 7.2, were digested with 0.047% (w/w) trypsin at 37 °C. At indicated time points, portions were removed for activity measurements or quenched by addition of TLCK and used later for polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. The inset shows the progress of the digestion. This figure was produced from a digital representation of the dried gel. The enzymatic activity measured at intervals during the tryptic digestion is plotted as a function of time (●). The intensity of the protein band corresponding to uncleaved MetE (~92 kDa) was determined by gel scanning as described under Materials and Methods and has also been plotted as a function of time (×).

complete inactivation while iodoacetate treatment does not, we tentatively concluded that inactivation is due to the reaction of a residue or residues in a relatively hydrophobic environment. For further characterization of the events associated with MetE inactivation, we chose to study the reaction of MetE with TPCK and iodoacetamide.

Figure 6 shows the time course for inactivation of MetE by 200 μM TPCK. The observed time course for inactivation can be well fitted by a first-order decay curve with a rate constant of 0.58 min^{-1}, as shown by the solid line. The inset to Figure 6 shows a plot of the observed first-order rate constant for inactivation as a function of TPCK concentration. A linear dependence of the rate constant on TPCK concentration is observed between 0 and 400 μM TPCK. Above 400 μM TPCK, the rate of inactivation is too fast to permit accurate determination of the rate constant. The calculated second-order rate constant for TPCK inactivation over this concentration range is 2.0×10^5 M^{-1} s^{-1}. Reaction of MetE with iodoacetamide also showed second-order dependence of the observed first-order rate constant for inactivation on the concentration of iodoacetamide in the range from 0 to 4 mM iodoacetamide (data not shown). The calculated second-order rate constant for reaction of MetE with iodoacetamide is 5.5×10^3 M^{-1} s^{-1}.

Identification of C726 as the Reactive Residue in MetE. Radiolabeled TPCK is not commercially available. Accordingly, we next wished to determine whether TPCK and iodoacetamide modify the same residue(s) on MetE, so that

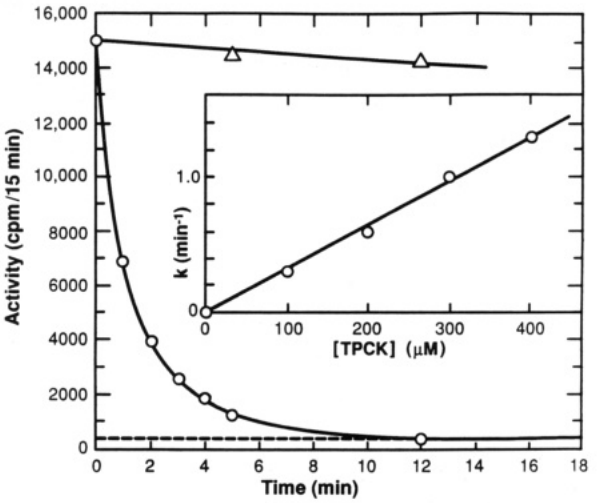


FIGURE 6: Rate of inactivation of MetE by 200 μM TPCK. The enzymatic activity, in arbitrary units of cpm of methionine formed per 15 min, is plotted as a function of the time of incubation of the enzyme with TPCK (○). The solid line represents the theoretical decay of activity associated with a first-order rate constant of 0.58 min^{-1}. Also shown is the activity of a control sample of enzyme incubated in the absence of TPCK (△). The inset shows a plot of the measured first-order rate constant for inactivation as a function of the TPCK concentration in the incubation.

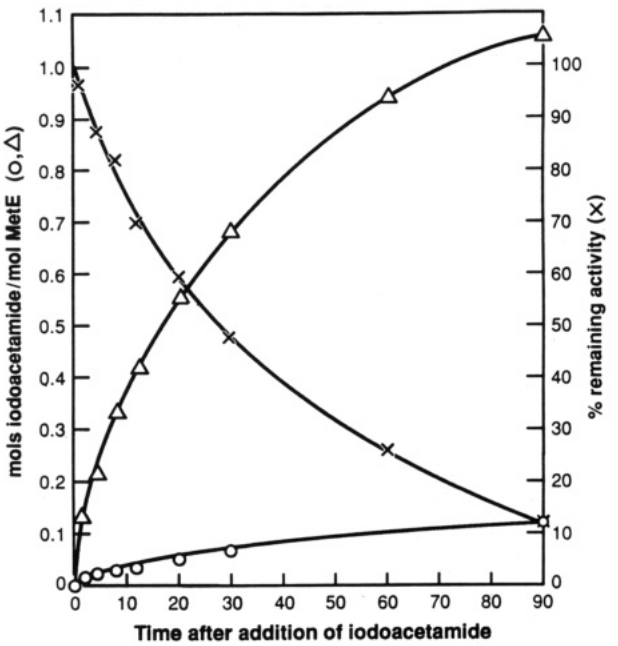


FIGURE 7: Rate of alkylation of MetE by 200 μM iodoacetamide. The covalent incorporation of radiolabeled iodoacetamide into MetE protein is plotted as a function of time for enzyme that was untreated at zero time (△) and for enzyme that had been pretreated for 60 min with 100 μM TPCK (○). Also shown is the rate at which enzymatic activity was lost during the incubation of untreated enzyme with 200 μM iodoacetamide (×). Incubations were performed at 37 °C.

radiolabeled iodoacetamide could be used to identify the reactive residue(s). Our approach was to compare the incorporation of labeled iodoacetamide into untreated enzyme and enzyme that had been pretreated for 60 min with 100 μM TPCK. Figure 7 shows the results of such an experiment. Enzyme that had not been pretreated with TPCK showed a first-order rate of incorporation of label from iodoacetamide into the MetE protein for the first 2 half-lives, and at 90 min, label from 1.06 mol of iodoacetamide/mol of MetE had been incorporated. Enzyme that had been pretreated with TPCK prior to exposure to iodoacetamide contained only 0.125 mol of label/mol of MetE. Also shown in Figure 7 is the rate of

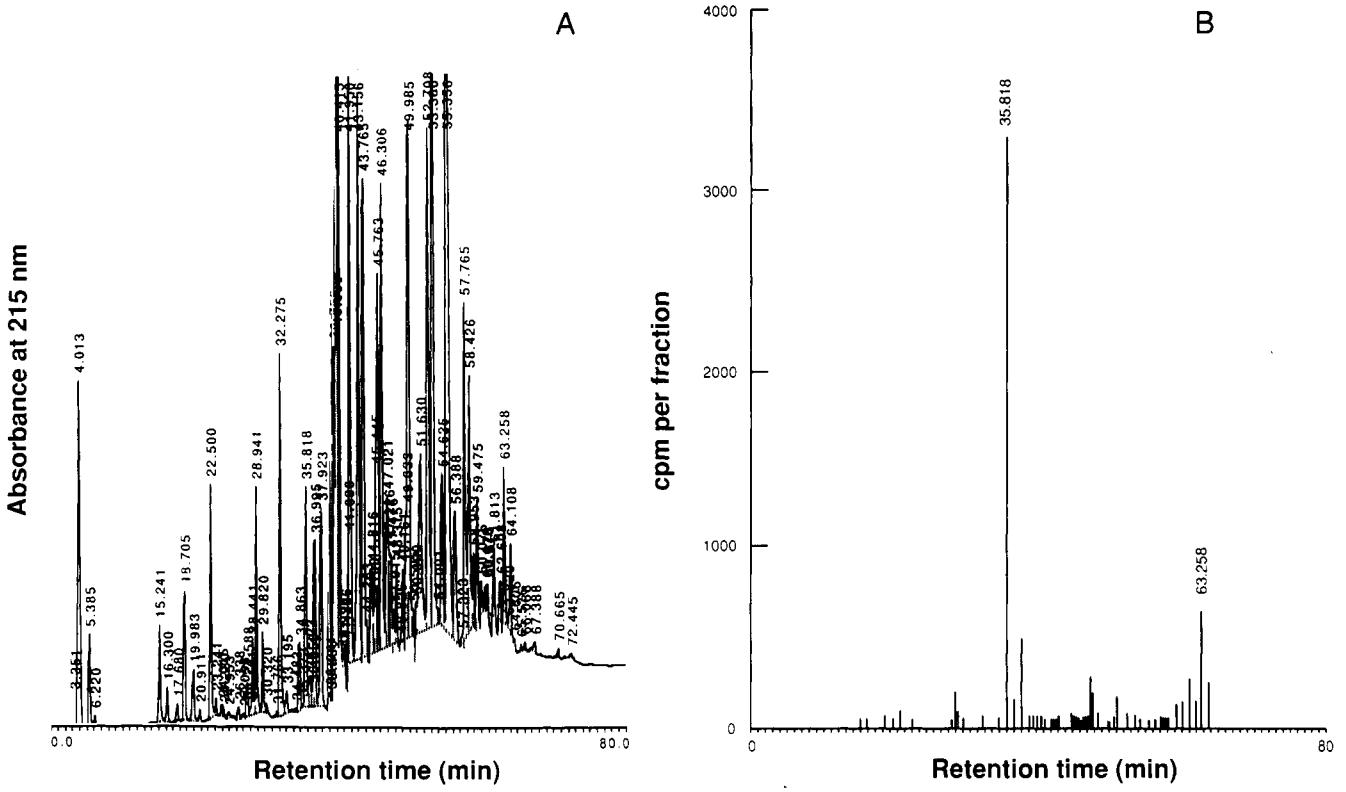


FIGURE 8: Isolation of peptides labeled by [^{14}C]iodoacetamide by digestion with lysyl endopeptidase and HPLC on a Vydac C18 reversed-phase column. In (A), the elution profile of the column, as monitored at 215 nm, is shown. In (B), the radioactivity associated with half of each peak eluting from the column is shown. The exact retention times of the prominent radiolabeled peaks are also indicated above the peaks.

activity loss during iodoacetamide treatment of enzyme that had not been preexposed to TPCK. Loss of activity was also first order, with a half-time of ~30 min, and at 90 min, 12% residual activity was observed. The rate of loss of activity was slightly slower than the rate of incorporation of label; the half-time for label incorporation was ~20 min. We conclude that alkylation of a single residue was responsible for the loss of activity seen on iodoacetamide treatment and that TPCK reacts with the same residue that was alkylated by iodoacetamide. The discrepancy between the observed rates of alkylation and of enzyme inactivation was probably due to some nonspecific alkylation of residues that did not lead to loss of enzyme activity. We estimate that ~18% of such nonspecific alkylation had occurred in 90 min in the enzyme that had not been pretreated with TPCK, based on the comparison of the level of incorporation of label at that time with the residual activity.

Samples of labeled alkylated enzyme prepared by incubation of MetE with 200 μM iodoacetamide, with or without pretreatment with TPCK, were digested with lysyl endopeptidase (LysC), and the peptides were separated by HPLC chromatography according to the protocol described under Experimental Procedures. The peptide separation is shown in Figure 8A. Figure 8B shows a bar graph of the counts in each peak of the chromatogram of the enzyme that was not pretreated with TPCK. The chromatogram of the digest of enzyme that had been pretreated with TPCK is not shown, but no peaks contained over 100 cpm. Thus TPCK appears to prevent labeling of two peptides by iodoacetamide, a major labeled peak eluting at 35.8 min and a minor labeled peak eluting at 63.3 min. The major peak contained 51% of the counts eluting from the column. Protection against labeling of these two peaks by pretreatment with TPCK has been seen in two independent labeling experiments. These two peptides were subjected to sequence analysis by Edman degradation. The peptide eluting at 35.8 min was identified as corresponding to residues 714–729. The amino acid eluted in each cycle of Edman degradation was exactly as predicted by the deduced amino acid sequence except in cycle 13, where a peak coeluting with the phenylhydantoin of glutamate was seen instead of the predicted cysteine derivative. We then measured the radioactivity released at each cycle of the Edman degradation. In the first 12 cycles, background levels of radioactivity were measured (<100 cpm), but 835 cpm were released at cycle 13, and in subsequent cycles the radioactivity levels decayed to background levels. From these observations we conclude that cysteine 726 is the residue that reacts most rapidly with iodoacetamide in MetE and that pretreatment with TPCK blocks alkylation of this residue. Edman degradation of the minor radiolabeled peak indicated the sequence XDTQIHTHMXYXE, identifying this peptide as that corresponding to residues 634–709. This peptide contains two cysteines, at positions 643 and 645. We have not determined the nature or position of the labeled residue(s) in this peptide.

We have shown that C726 is an unusually reactive thiol residue, and even though it has a low predicted surface probability, it reacts more rapidly with iodoacetamide than any other thiol in the protein. We have also shown that TPCK treatment blocks alkylation of C726 by iodoacetamide, although we have not yet established whether TPCK reacts with the same residue. Since alkylation of C726 leads to apparently complete loss of enzyme activity, we tentatively conclude that C726 is located at the active site of MetE. It is possible that C726 is merely located in physical proximity to the active site, such that its alkylation blocks access of the substrates. Alternatively, alkylation might induce a conformational change in the protein that destroys catalytic activity.

Although a thiol, C324, is present at the corresponding position in the N-terminal domain of the protein, this thiol is not alkylated by iodoacetamide to any significant degree.

These results suggest that the two domains, although probably structurally similar, have different functions in catalysis. In support of this conclusion, the residue corresponding to C324 in the P8 protein of *S. cerevisiae* is a serine rather than a cysteine.

Attempt To Label C726 of MetE with Methyl Iodide. We attempted to determine whether methylation of C726 results in the formation of an active or an inactive protein. Enzyme was incubated with 200 μM methyl iodide for 90 min, and activity was assayed at intervals throughout the incubation. Less than 5% loss of activity was observed during the incubation. Since this might indicate that C726 had been methylated, and the methyl group could be removed by homocysteine, we added radiolabeled iodoacetamide after 90 min of incubation with methyl iodide and showed that 1 equiv of iodoacetamide could be incorporated into the protein. These results indicated that C726 had not reacted with methyl iodide. At a much higher concentration of methyl iodide, 10 mM, $\sim$50% loss of enzyme activity was observed during a 90-min incubation, and subsequent reaction with radiolabeled iodoacetamide indicated that C726 had been partially methylated by methyl iodide. However, at this high concentration of methyl iodide, reaction with other residues on the protein is likely to be occurring, and so the experiment cannot be clearly interpreted.

Attempts To Demonstrate Methyl Transfer from Radiolabeled Methyltetrahydrofolate to MetE. We incubated the enzyme with methionine and [*methyl*-5-^{14}C]-CH_3-H_4PteGlu$_3$, precipitated the washed enzyme with trichloroacetic acid, and measured the radiolabel in the precipitate. No significant label was found in the precipitate, indicating that transfer of the methyl group from CH_3-H_4PteGlu$_3$ had not occurred to any significant degree. A negative experiment is not conclusive, because detection of a methylated protein intermediate requires that the equilibrium between methylated tetrahydrofolate and methylated enzyme in a dead-end ternary complex with methionine favor the methylated enzyme species. A definitive determination of the role of C726 in enzyme catalysis will require site-directed mutagenesis of this residue and the determination of the catalytic properties of the mutant enzyme. Stereochemical analysis of the methyl transfer from CH_3-H_4PteGlu$_3$ to homocysteine will also be mechanistically informative, in that a direct attack of homocysteine on the methyl group of CH_3-H_4PteGlu$_3$ should lead to inversion of configuration at the transferred methyl group, while a double-displacement mechanism with a methylated intermediate should result in retention of configuration at the transferred methyl group.

Overview. This report describes initial studies in a long-term program to compare the structures and catalytic mechanisms of two enzymes from *E. coli* that catalyze highly similar reactions. We are interested in examining the flexibility of nature to devise alternate solutions to a difficult chemical problem, namely, the transfer of a methyl group from an unactivated tertiary amine. We would have predicted that a direct methyl transfer from CH_3-H_4PteGlu$_3$ to homocysteine occurs in the reaction catalyzed by MetE. The observation of a highly reactive thiol in MetE, with reactivity toward chloromethyl ketones that is reminiscent of the reactivity of the thiol at the active site of papain (Hinkle & Kirsch, 1973), suggests an alternate possibility. The thiol at the active site of papain not only is highly nucleophilic but also has a pK_a less than 5 (Lewis et al., 1979). This dramatic shift in the pK value for the thiol group associated with free cysteine appears to result from its interaction in a hydrophobic environment with the positively charged imidazolium group of an adjacent histidyl residue (Polgar, 1973; Lewis et al., 1979). While the thiol in papain is activated to facilitate attack on the carbonyl group of the scissile peptide bond of protein substrates, the properties of such an activated thiol resemble those of cob(I)alamin in cobalamin-dependent methionine synthase. That is, both nucleophiles exhibit weak basicity and strong nucleophilicity. Such properties may permit discrimination between hydrogen and alkyl substituents in methyltransfer reactions, as well as facilitating carbonyl addition reactions.

ACKNOWLEDGMENTS

We thank Herbert Weissbach for the generous gift of pRSE562. We are indebted to the personnel in The University of Michigan Protein Sequencing Facility, and especially to Sari Vlahakos, for assistance with the LysC digestion of MetE and peptide separation and for sequence analysis of peptides by Edman degradation. We also thank James Drummond for his thoughtful comments on the manuscript.

REFERENCES

Banerjee, R. V., & Matthews, R. G. (1990) *FASEB J. 4*, 1450–1459.

Banerjee, R. V., Johnston, N. L., Sobeski, J. K., Datta, P., & Matthews, R. G. (1989) *J. Biol. Chem. 264*, 13888–13895.

Banerjee, R. V., Frasca, V., Ballou, D. P., & Matthews, R. G. (1990a) *Biochemistry 29*, 11101–11109.

Banerjee, R. V., Harder, S. R., Ragsdale, S. W., & Matthews, R. G. (1990b) *Biochemistry 29*, 1129–1135.

Bradford, M. (1976) *Anal. Biochem. 72*, 248–254.

Brown, K. L. (1982) in *B_{12}, Volume 1: Chemistry* (Dolphin, D., Ed.) pp 245–294, Wiley, New York.

Burton, E., & Sakami, W. (1969) in *Methods in Enzymology* (Tabor, H., & Tabor, C. W., Eds.) Vol. 17B, pp 388–392, Academic Press, New York.

Devereaux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Foster, M. A., Tejerina, G., & Woods, D. D. (1961) *Biochem. J. 81*, 1P.

Foster, M. A., Tejerina, G., Guest, J. R., & Woods, D. D. (1964) *Biochem. J. 92*, 476–488.

Gribskov, M., Devereux, J., & Burgess, R. R. (1984) *Nucleic Acids Res. 12*, 539–549.

Hinkle, P. M., & Kirsch, J. F. (1970) *Biochemistry 9*, 4633–4643.

Lee, R., & McElroy, W. D. (1969) *Biochemistry 8*, 130–136.

Lewis, S. D., Johnson, F. A., & Shafer, J. A. (1976) *Biochemistry 15*, 5009–5017.

Lexa, D., & Saveant, J.-M. (1983) *Acc. Chem. Res. 16*, 235–243.

Lipman, D. J., Wilbur, W. J., Smith, T. F., & Waterman, M. S. (1984) *Nucleic Acids Res. 12*, 215–226.

Luschinsky, C. L., Drummond, J. T., Matthews, R. G., & Ludwig, M. L. (1992) *J. Mol. Biol.* (in press).

Martin, B. D., & Finke, R. G. (1990) *J. Am. Chem. Soc. 112*, 2419–2420.

Matthews, R. G. (1986) *Methods Enzymol. 122*, 333–339.

Maxon, E. M., Redfield, B., Cai, X.-Y., Shoeman, R., Fujita, K., Fisher, W., Stauffer, G., Weissbach, H., & Brot, N. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 85–89.

Ohtake, Y., Watanabe, K., Tezuka, H., Ogata, T., Yabuuchi, S., Murata, K., & Kimura, A. (1988) *Agric. Biol. Chem. 52*, 2753–2762.

Polgar, L. (1973) *Eur. J. Biochem. 33*, 104–109.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Sarimo, S. S., & Pine, M. J. (1969) *J. Bacteriol. 98*, 368–374.
Shaw, E. (1970) in *The Enzymes* (Boyer, P., Ed.) 3rd ed., Vol. 1, pp 91–146, Academic Press, Orlando, FL.
Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature 246*, 40–41.
VanBogelen, R. A., Hutton, M. E., & Neidhardt, F. C. (1990) *Electrophoresis 11*, 1131–1166.
Vincenz, C., Fronk, J., Tank, G. A., & Langmore, J. P. (1991) *Nucleic Acids Res. 19*, 1325–1336.
Whitfield, C. D., Steers, E. J., Jr., & Weissbach, H. (1970) *J. Biol. Chem. 245*, 390–401.
Yager, T. D., & von Hippel, P. H. (1987) in *Escherichia coli and Salmonella typhimurium* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., & Umbarger, H. E., Eds.) Vol. 2, pp 1241–1275, American Society for Microbiology, Washington, DC.

Basis for the Anomalous Effect of Competitive Inhibitors on the Kinetics of Hydrolysis of Short-Chain Phosphatidylcholines by Phospholipase A_2[†]

Joseph Rogers, Bao-Zhu Yu, and Mahendra Kumar Jain*
Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716
Received January 10, 1992; Revised Manuscript Received April 14, 1992

ABSTRACT: The effect of four specific competitive inhibitors on the kinetics of hydrolysis of short-chain diacyl-*sn*-glycero-3-phosphocholines below their critical micelle concentrations was examined. The kinetics of hydrolysis of short-chain substrates dispersed as solitary monomers were generally consistent with the classical Michaelis–Menten formalism; i.e., hydrolysis began without any latency period, the steady-state rate was observed at higher substrate concentrations, the steady-state initial rate showed a linear dependence on the enzyme concentration, and the hyperbolic dependence of the initial rate on the substrate concentration could be described in terms of K_M and V_{max} parameters. The competitive nature of the inhibitors used in this study has been established by a variety of techniques, and the equilibrium dissociation constants for the inhibitors bound to the enzyme were measured by the protection method [Jain et al. (1991) *Biochemistry 30*, 7306–7317]. The kinetics of hydrolysis in the presence of competitive inhibitors could be described by a single dissociation constant. However, the value of the dissociation constant obtained under the kinetic conditions was comparable to that obtained by the protection method for the inhibitor–enzyme complex bound to a neutral diluent, rather than to the value of the dissociation constant obtained with solitary monomeric inhibitors and the enzyme in the aqueous phase. Spectroscopic methods showed that the effectively lower dissociation constant of an inhibitor bound to PLA2 at the interface is due to the stabilization of the enzyme–inhibitor complex by interaction with other amphiphiles present in the reaction mixture. These results show that the EI complex in the aqueous phase binds other solitary or aggregated amphiphiles to the interfacial recognition region on the enzyme (i-face).

Interfacial catalysis by phospholipase A_2 (PLA2)[1] has been quantitatively described by an adaptation of Michaelis–Menten formalism, where the binding of the enzyme in the aqueous phase to the substrate interface precedes the catalytic turnover by the enzyme in the interface (Berg et al., 1991; Verger & de Haas, 1976). Elsewhere, we have shown that virtually all aspects of the kinetics of action of PLA2 on anionic vesicles are quantitatively described by the catalytic turnover cycle in the interface and the enzyme does not leave the interface between the successive turnover cycles (Jain et al., 1986a; Jain & Berg, 1989). The rate and equilibrium parameters for the catalytic action of PLA2 on DMPM vesicles in this highly processive scooting mode have been determined (Berg et al., 1991; Jain et al., 1991a). The catalysis in the scooting mode has been used to quantitatively characterize the action of PLA2 from different sources (Jain et al., 1991b) as well as the covalently modified enzyme (Ghomashchi et al., 1991), to determine the substrate specificity (Jain & Rogers, 1989; Ghomashchi et al., 1991), and to describe the kinetics in the presence of inhibitors (Jain et al., 1989, 1991a,b) and activators (Jain et al., 1991c). This kinetic model also accounts for interfacial catalysis by PLA2 on micelles and monolayers with the provision that at such interfaces the exchange of the substrate and products must be taken into consideration (Jain & Berg, 1989).

While there appears to be a general consensus about the formation of a specific active-site-directed complex of the enzyme in the interface with the ligands in the interface (Verheij et al., 1981; Jain et al., 1982; Dennis, 1983; Ramirez

[†] This research was supported by a grant from Sterling Inc. and GM29703.
* To whom correspondence should be addressed.

[1] Abbreviations: cmc, critical micelle concentration; deoxy-LPC, 1-hexadecylpropanediol 3-phosphocholine; 2H-GPC, 2-hexadecyl-*sn*-glycero-3-phosphocholine; PC6, PC7, and PC8, diacyl-*sn*-glycero-3-phosphocholines with the indicated acyl chain lengths; MG14, RM2, RM3, MJ33, and MJ72, structures shown in Figure 2; PLA2, phospholipase A_2 from pig pancreas; proPLA, precursor of PLA2.

0006-2960/92/0431-6056$03.00/0