

# Mechanistic Studies on Thiaminase I

## OVEREXPRESSION AND IDENTIFICATION OF THE ACTIVE SITE NUCLEOPHILE\*

(Received for publication, September 12, 1995, and in revised form, November 8, 1995)

Colleen A. Costello‡, Neil L. Kelleher‡, Mitsuko Abe§, Fred W. McLafferty‡,  
and Tadhg P. Begley‡¶

From the ‡Department of Chemistry, Cornell University, Ithaca, New York 14853 and the §Department of Microbiology,  
Yamaguchi University School of Medicine, Ube 755, Yamaguchi-ken, Japan

**Thiaminase I (EC 2.5.1.2) catalyzes the replacement of the thiazole moiety of thiamin with a wide variety of nucleophiles. Here we report the sequencing of a thiaminase I clone from *Bacillus thiaminolyticus*, the overexpression of the cloned gene in *Escherichia coli*, and the purification and characterization of the enzyme. Recombinant thiaminase I functions as a monomer with a  $K_m$  for thiamin of  $3.7 \pm 0.6 \mu\text{M}$  and a  $k_{\text{cat}}$  of  $34 \text{ s}^{-1}$ . Electrospray ionization Fourier-transform mass spectrometry identified a single sequencing error and demonstrated heterogeneity, finding molecular weights of 42,127, 42,198, and 42,255 due to added Ala and Gly-Ala at the amino terminus. Similar analysis of the 4-amino-2-methyl-6-chloropyrimidine (8) inactivated enzyme indicated that the active site nucleophile involved in catalysis of the substitution reaction is located between Pro<sup>79</sup> and Thr<sup>177</sup>. Subsequent cysteine-specific labeling and site-directed mutagenesis identified Cys<sup>113</sup> as the active site nucleophile.**

During the course of studies to determine the thiamin (1) content of various foods in Japan during World War II, it was observed that some foods contained a thiamin-degrading enzymatic activity. This activity is fairly widespread and has been detected in several bacteria, marine organisms, and plants (Fujita, 1954, 1972; Murata, 1982). Animals and humans who ingest food containing thiaminase may develop symptoms of thiamin deficiency (Hayashi, 1957; Duffy *et al.*, 1981; Earl *et al.*, 1994; Evans, 1975). Two thiamin-degrading enzymes have been isolated. Thiaminase I (EC 2.5.1.2) catalyzes the replacement of the thiazole moiety of thiamin with a wide range of nucleophiles; thiaminase II is specific for the use of water as the nucleophile (Reaction 1).

Thiaminase I is the better characterized system (Abe *et al.*, 1987; Lienhard, 1970; Pearson and Lipman, 1988). It is an extracellular, 42-kDa enzyme which has been purified from *Bacillus thiaminolyticus* (Douthit and Airth, 1966), and its gene has also been cloned (Abe *et al.*, 1987). Thiaminase I has

been reported to have a wide tolerance to the nature of the nucleophile (Lienhard, 1970).

A reliable source of large quantities of this enzyme was essential for a comprehensive series of mechanistic studies. Towards this end, we have sequenced the thiaminase I gene from *B. thiaminolyticus* and overexpressed it in *Escherichia coli*. This paper describes the sequencing and overexpression strategy, as well as the purification, characterization, and identification of the active site nucleophile for the recombinant enzyme.

### EXPERIMENTAL PROCEDURES

**General Methods**—Ampicillin (Amp), isopropyl-1-thio- $\beta$ -D-galactopyranoside, Sigma Dalton VII SDS low molecular weight standards, and Cibacron Blue dye affinity resin were obtained from Sigma. Thiamin hydrochloride, aniline, and veratrylamine were obtained from Aldrich Chemical Co. 2-Nitro-5-thiosulfobenzoate (NTSB)<sup>1</sup> was a gift from H. Scheraga (Thannhauser *et al.* 1984). The Wizard<sup>TM</sup> PCR purification kit, and the Wizard<sup>TM</sup> MiniPrep kit were obtained from Promega. Restriction enzymes and T4 DNA ligase were from New England Biolabs. Protein concentrations were determined with the PIERCE<sup>TM</sup> Coomassie Plus Protein Assay Reagent (Bradford, 1976) using bovine serum albumin as a standard. High performance liquid chromatography was carried out on a Waters<sup>TM</sup> 650E Protein Purification System with a Waters<sup>TM</sup> 490E variable wavelength detector. Ultrafiltration was done using an Amicon concentrator, Model 8050 (Beverly, MA). UV spectra were recorded on a Hewlett-Packard 8451A diode array spectrophotometer. Freshly distilled aniline (b.p. 183–184 °C) and veratrylamine were used for thiaminase I activity assays. Oligonucleotide primers were synthesized at the Cornell University Oligonucleotide Synthesis Facility.

**Bacterial Strains and Plasmids**—Plasmids pET17b and pET22b(+) and *E. coli* strain BL21(DE3) were obtained from Novagen, Inc. (Madison, WI). *E. coli* strain SURE was obtained from Stratagene. Plasmid pGEM3z<sup>(+)</sup> was obtained from Promega. Plasmids pAN13 and H104 were gifts from Dr. Nishimune (Higashinasi, Osaka, Japan). For propagation of the recombinant plasmids, *E. coli* K-12 strains RK4353 and VJS533 were used and were a gift from Dr. V. Stewart (Dept. of Microbiology, Cornell University).

**Assay for Thiaminase I Activity**—Thiaminase I activity was determined, at 25 °C, using two different secondary nucleophiles. In both cases, the absorbance change at 252 nm, resulting from the reaction of thiamin with the secondary nucleophile, was monitored (Lienhard, 1970). The assay mixture contained 0.93 mM nucleophile, 0.093 mM thiamin, 100 mM sodium phosphate buffer (pH 6.5), and 2 mM DTT in a 3-ml volume at 25 °C. The assay was initiated by addition of enzyme, and the absorbance was measured at 252 nm every 30 s over a 6-min period. In each case, the recording of absorbance against time was linear over the assay time period. In assays done to track activity during purification, aniline was used as the nucleophile. Assays done to determine kinetic parameters used veratrylamine as the nucleophile. In determining the kinetic parameters for thiamin, the veratrylamine concentration was maintained at 0.93 mM, and the thiamin concentra-

\* This work was supported by National Institutes of Health Grants DK44083 (to T. B.) and GM16609 (to F. M.), National Institutes of Health Training Grants T32 GM08384 (to C. C.) and T2 GM07273 (to N. K.), and by a Department of Education Fellowship (to C. C.). This is Paper II in the series, "Mechanistic Studies on Thiaminase I." Nicewonger *et al.* (1995) is Paper I in the series. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank<sup>TM</sup>/EMBL Data Bank with accession number(s) U17168.

¶ To whom correspondence should be addressed: Dept. of Chemistry, 120 Baker Laboratory, Cornell University, Ithaca, NY 14853. Tel.: 607-255-7133; Fax: 607-255-4137.

<sup>1</sup> The abbreviations used are: NTSB, 2-nitro-5-thiosulfobenzoate; PCR, polymerase chain reaction; DTT, dithiothreitol; bp, base pair(s); ESI, electrospray ionization; FT-MS, Fourier transform-mass spectrometry.

dures. The membrane was washed extensively with water (10 × 5 ml aliquots), stained with 0.06% Coomassie R-250 in 50% MeOH for 5 min, and destained with 30% aqueous MeOH, 10% acetic acid for 30 min. The portion of the membrane containing the major protein band (42 kDa) was cut out, and the amino-terminal sequence was determined at the Sequencing and Analytical Facility, Cornell University, using an ABI

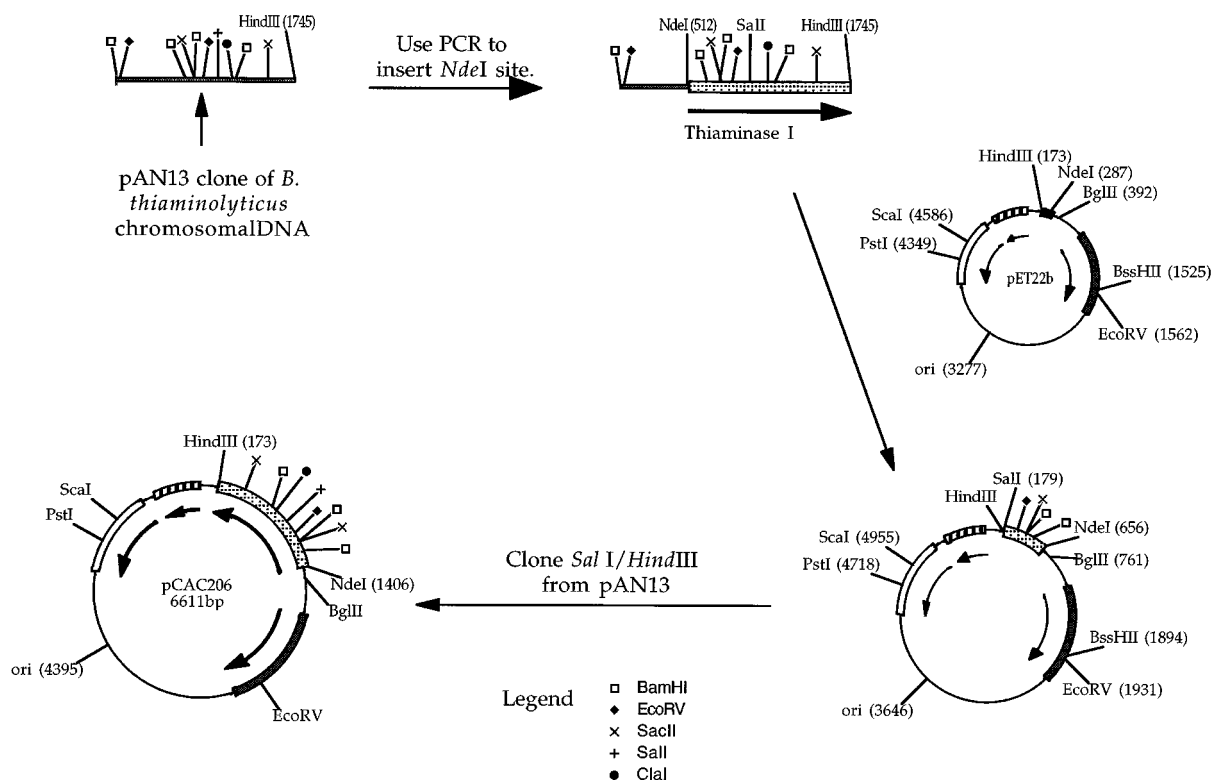


FIG. 4. Strategy for cloning the thiaminase I gene into the overexpression plasmid, pET22b(+). This strategy minimized the quantity of PCR-derived DNA in the overexpression construct.

Model 470A gas phase protein Sequencer with a Model 120A on-line PTH-AA analyzer.

**Construction of the Thiaminase I Expression Vector**—Using pAN13 linear double-stranded DNA as a template and 5'-AGC TTA TCA TCG ATA AGC TT-3' and 5'-CAT TTT AAT AGA AGC GGG GGA CAA GCA TAT GTC AAA GGT AAA AGC CTT CAT T-3' as 3' and 5' primers, respectively. 30 PCR cycles were performed using cycle conditions of 94 °C (denaturing), 40 °C (annealing), and 72 °C (elongation) in standard buffer (10 mM KCl, 20 mM Tris-Cl (pH 8.8 at 25 °C), 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1% Triton X-100, and 3 mM MgSO<sub>4</sub>). The resulting 1.2-kilobase PCR product was isolated using the Wizard<sup>TM</sup> PCR Purification Kit (Promega) and further purified by electrophoresis on a 0.7% low melting agarose gel before it was digested with *NdeI* and *SalI* and cloned into the *NdeI*-*SalI* sites of the pET22b(+) expression vector (Novagen, Inc.). The remainder of the thiaminase I gene was then inserted into the expression vector from pAN13. The PCR-derived DNA was sequenced and shown to contain a one base mutation in codon 6 (GGC → GCC). A representative plasmid, designated pCAC206, was used to transform *E. coli* BL21(DE3) to give the overexpression strain (Fig. 4).

**Purification of Thiaminase I**—The *E. coli* thiaminase I overexpression strain was grown at 37 °C in Lenox broth media (Life Technologies, Inc.) containing 200 µg/ml ampicillin until an  $A_{600\text{ nm}}$  of 0.45 was reached. The temperature of the culture was reduced to 25 °C, and the cells were grown for an additional 30–60 min ( $A_{600\text{ nm}} = 0.6$ ), induced by making the culture 1.0 mM in isopropyl-1-thio- $\beta$ -D-galactopyranoside, and grown for an additional 3.5 h at 25 °C. The cells were harvested by centrifugation (3000  $\times g$ , 10 min) and stored at -80 °C.

The enzyme isolated from this expression system is stable at room temperature, maintaining more than 85% of its original activity over a 12-h period at 23 °C. Therefore, the enzyme was isolated at room temperature unless otherwise stated using extensive modifications of the procedure of Wittliff and Airth (1968). Loss of activity early in the purification scheme was seen when 2 mM DTT was omitted from the buffers. The cells were suspended in 50 mM Tris-HCl (pH 7.5), 2 mM DTT, and 2 mM EDTA (12 ml/g cell paste), ruptured in a French press at 12,000 p.s.i., and the cell debris removed by centrifugation for 20 min at 25,000  $\times g$ . The resulting clarified cell extract was adjusted to 50% saturation by the slow addition of solid ammonium sulfate (291 g/liter) at 0 °C, stirred for 1 h at 0 °C and centrifuged at 15,000  $\times g$  for 20 min. The supernatant was then brought to 70% saturation by the addition of solid ammonium sulfate (125 g/liter) at 0 °C and stirred for 1 h before

centrifugation at 15,000  $\times g$  for 20 min at 0 °C. The resulting 50–70% ammonium sulfate pellet was dissolved in 10–30 ml of 50 mM Tris-HCl (pH 7.5), 2 mM DTT, 2 mM EDTA, and the residual ammonium sulfate was removed by ultrafiltration (Amicon, YM-10, 250-fold dilution in the same buffer) before loading the sample directly onto a Waters<sup>TM</sup> AP-1, 8-µm HR DEAE-ion exchange column (25  $\times$  100 mm) equilibrated in the same buffer. The protein was eluted from the column at a flow rate of 1.5 ml/min using the sodium chloride gradient shown in Fig. 5. This consisted of 8 min at 0 M NaCl, 12.25 min from 0–150 mM NaCl, 20 min at 150 mM NaCl, and 20 min from 150 mM–1 M NaCl. Thiaminase I was eluted at 150 mM NaCl. The fractions containing thiaminase I activity were pooled, and the buffer was changed to 50 mM phosphate buffer (pH 7.0), 2 mM DTT, 2 mM EDTA by ultrafiltration (Amicon, YM-10). The sample was loaded onto a Cibacron Blue dye affinity column (2.5  $\times$  11 cm) which had been equilibrated in the same buffer. Thiaminase I eluted from the column in the same buffer at a flow rate of 0.8 ml/min (Fig. 6). The fractions containing activity were pooled, concentrated (5 mg/ml), and stored in 100-µl aliquots as 10% glycerol stocks at -80 °C.

**Electrophoresis and Gel Filtration of Thiaminase I**—The subunit molecular mass of the recombinant thiaminase I was determined by its mobility on a 12% SDS-polyacrylamide gel relative to Sigma Dalton VII SDS-PAGE molecular mass standards (bovine albumin, 66 kDa; egg albumin, 45 kDa; glyceraldehyde-3-phosphate dehydrogenase, 36 kDa; carbonic anhydrase, 29 kDa; trypsinogen, 24 kDa; trypsin inhibitor, 20.1 kDa;  $\alpha$ -lactalbumin, 14.2 kDa). To determine the native molecular mass, thiaminase I was also loaded onto a Sepharose CL-6B gel filtration column (1.5  $\times$  91.5 cm) in 50 mM phosphate buffer (pH 7.0), 2 mM DTT, 2 mM EDTA, 150 mM NaCl. The elution of thiaminase I relative to blue dextran suggests that the enzyme is monomeric. The column was calibrated using Sigma GF-200 molecular mass markers ( $\alpha$ -amylase, 200 kDa; alcohol dehydrogenase, 150 kDa; bovine serum albumin, 66 kDa; carbonic anhydrase, 29 kDa; cytochrome *c*, 12.4 kDa).

**pI Determination of Thiaminase I**—Isoelectric focusing of purified thiaminase I was done using a Pharmacia Biotech broad range isoelectric focusing gel (pH 3–9). The isoelectric focusing gel was run on a Pharmacia PhastSystem using the manufacturer's procedures.

**Determination of the Extinction Coefficient**—The enzyme was desalted on a Sephadex G-50 gel filtration column (1.5 cm  $\times$  98.4 cm) equilibrated in water. The protein was eluted from the column in water and, prior to analysis, was concentrated by ultrafiltration (Amicon, YM-10). The  $A_{280\text{ nm}}$  was determined from the UV/Vis spectrum. The protein concentration was determined by quantitative amino acid anal-

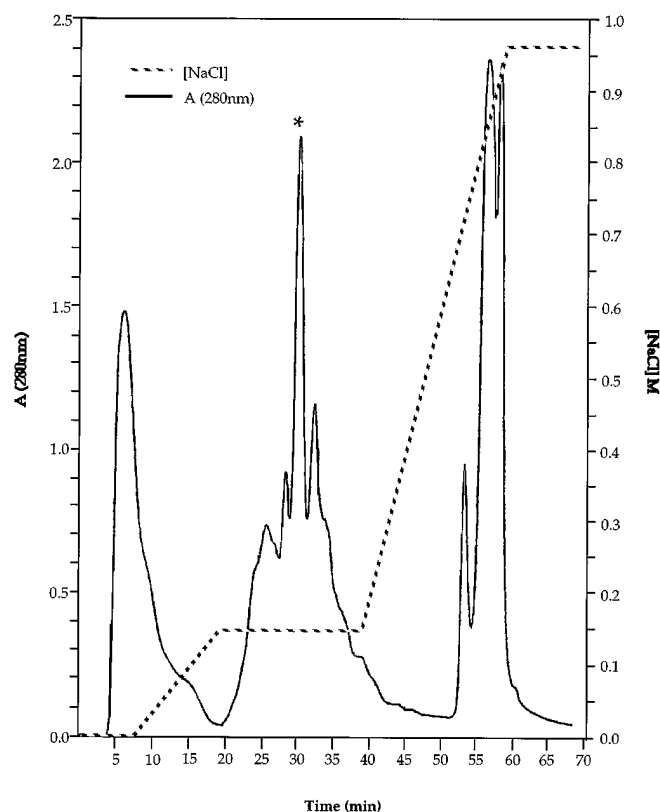


FIG. 5. Purification profile for thiaminase I using Waters™ 8- $\mu$ m HR DEAE-ion exchange chromatography. The thiaminase I-containing peak is marked with an asterisk and was eluted from the column with a NaCl gradient.

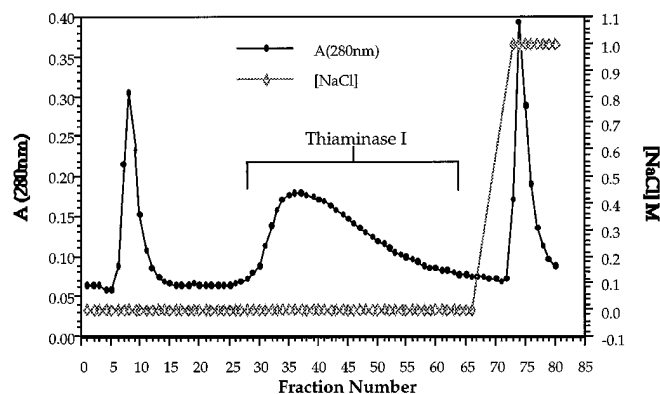


FIG. 6. Purification profile for thiaminase I using Cibacron Blue dye affinity chromatography. Thiaminase I eluted from the column in loading buffer (50 mM phosphate, pH 7.0, 2 mM DTT, AND 2 mM EDTA) as the second of two protein peaks.

ysis, performed in triplicate at the Sequencing and Analytical Facility, Cornell University, using the Waters™ Pico-Tag chemistry.

**NTSB Modification of Thiaminase I and Inactivated Thiaminase I** (Thannhauser *et al.*, 1987)—The standard protocol for denaturation and labeling was used with no modifications. Thiaminase I and fully inactivated thiaminase I were desalted using a Sephadex G-50 column (1.5  $\times$  98.4 cm) equilibrated in water followed by ultrafiltration (Amicon, YM-10, 500-fold in water), respectively. The NTSB assay solution (3.0 ml) was added to a solution of thiaminase I (200  $\mu$ l). In conjunction with each assay, a negative control was run containing water (200  $\mu$ l). The reaction mixture was incubated in the dark for 25 min. The absorbance at 412 nm was recorded against the negative control.

**Inactivation of Thiaminase I with 8**—A purified sample of thiaminase I (1.0 mg) was inactivated to 25% of its original activity by incubation in the presence of **8** (5.5 mM) at 25 °C for 24 h in 50 mM phosphate buffer, 2 mM EDTA, 2 mM DTT (pH 7.0). Excess inhibitor was removed

by ultrafiltration (YM-10, 500-fold dilution) in the same buffer.

**Electrospray Ionization Fourier Transform-Mass Spectrometry (ESI/FT-MS) of Recombinant, Inhibited, and Mutated Thiaminase I**—Thoroughly desalted thiaminase I at  $\sim 10$   $\mu$ M in 76:20:4 MeOH/H<sub>2</sub>O/AcOH was infused at 1  $\mu$ l/min into a modified 6T FT-MS (Beu *et al.*, 1993) through an electrospray needle held at 3kV. The resulting ions were electrostatically guided through a heated metal capillary (110 °C), a skimmer, and three quadrupole ion guides (five stages of differential pumping) into a trapped ion cell at  $10^{-9}$  Torr. Trapped ions were coherently excited into high cyclotron orbits about the magnet field axis for dipolar detection of image current, digitization, and Fourier transformation, yielding a spectrum of all frequencies (*i.e.*  $m/z$  values) simultaneously.

**Site-directed Mutagenesis on Thiaminase I**—Thiaminase I mutants were generated according to the method of Vandeyar *et al.* (1988) using the T7-Gen *In Vitro* Mutagenesis System (U. S. Biochemical Corp.) with the exception that SURE cells were transformed with mutagenesis reaction mixtures. Single-stranded DNA was prepared from *E. coli* JM101 cells (Stratagene) transformed by a M13mp18 clone containing the 410-bp *Bam*HI fragment from pCAC306 which served as a template for mutagenesis. For the C113S mutation, the following oligonucleotide was used: 5'-CG GTA AAA AAG CAG GTT CGT GGA TAA AAT TTG CGG CAG ACC-3'. Mutant clones were confirmed by DNA sequencing. The *Bam*HI fragment containing the mutation was then moved to pGEM3z<sup>(+)</sup> to produce a representative plasmid, pCAC7-1. The *Eco*RV-*Cla*I fragment from pCAC7-1, containing the mutagenized site was used to replace the corresponding fragment in an overexpression plasmid, pCAC406 (a pET17b ( $\Delta$ BstXI) derivative cloned in a similar fashion as shown in Fig. 4), to produce a representative plasmid, pCAC501. The *Nde*I-*Hind*III fragment of pCAC501, containing the entire C113S-thiaminase I gene, was cloned into pET22b(+) to produce the C113S mutant overexpression plasmid, pCAC511 (Fig. 7).

## RESULTS

**Sequencing and Overexpression of Thiaminase I in *E. coli***—The sequencing strategy for the 1,750-bp region of pAN13, which contains the thiaminase I gene, is summarized in Fig. 2. Two open reading frames were identified in the resulting sequence. The second of these was assigned to the thiaminase I gene because it contained the amino-terminal sequence of the purified enzyme from *B. subtilis* 168 (H104) (AHSDASXX-ITLKVAIYPYVP) and coded for a protein of the predicted molecular mass (45 kDa). The leader sequence (codons 1–30), expected for an extracellular protein, had the usual organization: consisting of a positively charged amino-terminal domain, an extended hydrophobic core, and a hydrophilic carboxyl-terminal domain. In addition, alanine was found at the –1 and –3 positions directly preceding what was the apparent signal peptide processing site (Chen and Nagarajan, 1994). The results of a FASTA search using the thiaminase I sequence gave no significant similarity with other known proteins.

The overexpression strategy involved introducing an *Nde*I site at the start codon of the thiaminase I gene using PCR mutagenesis. This facilitated the insertion of the gene into the pET22b(+) overexpression vector. The PCR-derived DNA was sequenced and was shown to contain a 1-base mutation in codon 6 (GGC  $\rightarrow$  GCC). This mutation changed Gly<sup>6</sup> to Ala<sup>6</sup> in the leader peptide. Since this mutation had no apparent effect on signal peptide processing as shown by ESI/FT-MS, no additional efforts were made to correct it.

**Purification and Characterization of Recombinant Thiaminase I**—The purification strategy is summarized in Table I. SDS-polyacrylamide gel electrophoresis analysis of each stage of the purification showed the progressive enrichment of a 42-kDa protein (Fig. 8). The purified enzyme appears to be monomeric by gel filtration analysis. It has a pI of 4.6 and an extinction coefficient at 280 nm of  $6.2 \pm 0.3 \times 10^4$  M<sup>-1</sup> cm<sup>-1</sup>. Amino-terminal sequencing of the recombinant protein indicated that the putative leader sequence had been removed. However, ESI/FT mass spectra of recombinant thiaminase I revealed three species of molecular mass = 42,127, 42,198,

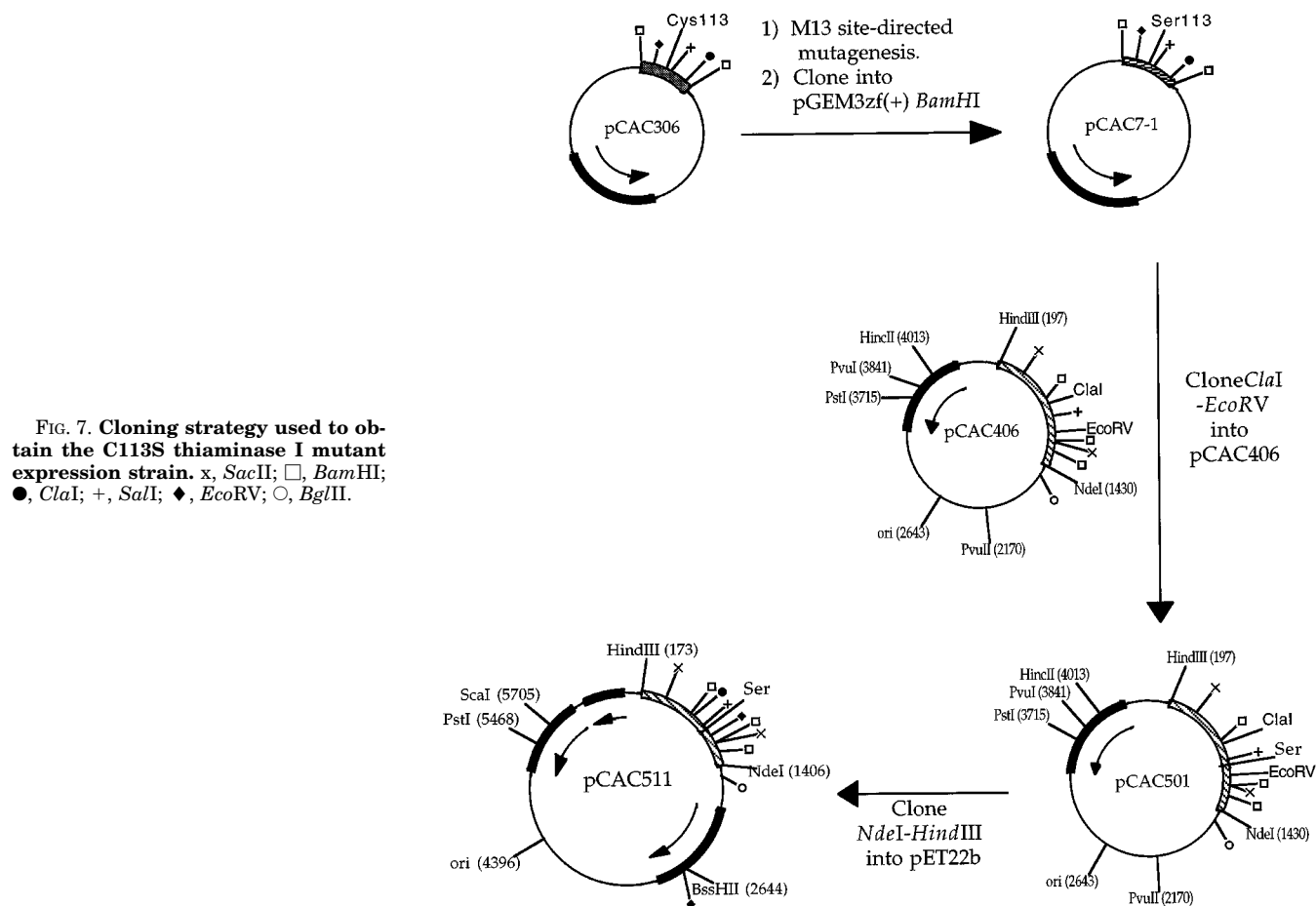


TABLE I  
Purification of recombinant thiaminase I from *E. coli*

See text for details.

Step	Total protein	Total activity	Specific activity	% Total	Purification
	mg	units	units/mg protein		-fold
Cell lysate	1344	1035	0.8	100	1
50–70% (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	300	990	3.3	96	4
DEAE-ion exchange	35	180	5.1	17	7
Cibacron Blue dye affinity chromatography	8.4	69.6	8.3	7	11

42,255 ± 1 Da in a 1:2:1 ratio (Fig. 9a). This heterogeneity, which requires resolving power greater than 600 for its detection, could not have been detected using ESI/quadrupole MS or matrix-assisted laser desorption MS in this mass range. Nozzle-skimmer dissociation (*i.e.* collisions with small neutrals in the ESI source) of the molecular ions produced ~165 fragment ions, some exhibiting the same heterogeneity as the intact protein; parts per million mass accuracy MS spectra of these demonstrated that the mass differences between species were actually 71.02 and 57.02 Da, consistent with an extra Ala (71.04) and Gly (57.02) on the amino terminus (Kelleher *et al.*, 1995).

The  $K_m$  for thiamin was determined to be  $3.6 \pm 0.7 \mu\text{M}$  and the  $k_{\text{cat}}$   $34 \text{ s}^{-1}$  (pH 6.5, 25 °C) using veratrylamine (4) (935  $\mu\text{M}$ ) as a co-substrate. The corresponding kinetic parameters for the enzyme isolated from *B. subtilis* 168 (H104) were  $6.6 \pm 1.2 \mu\text{M}$  and  $30 \text{ s}^{-1}$ . Although aniline had been used previously as the co-substrate (Wittliff and Airth, 1968), we find it to be an unsatisfactory substrate for kinetic studies due to its high extinction coefficient and its facile oxidation to highly chro-

mophoric products. We have found that the use of veratrylamine as the co-substrate gives high quality kinetic data. Although highly sensitive to thiamin concentration, the rate of the enzymatic reaction is independent of veratrylamine concentration.

In characterizing the enzymatic reaction, we have demonstrated that the enzyme reacts with thiamin to liberate 4-methyl-5-(2-hydroxyethyl)thiazole (3) in the absence of a secondary nucleophile. Also, in the presence of limiting amounts of veratrylamine (4) as the co-substrate, a disubstituted veratrylamine product, bis-(4-amino-2-methyl)pyrimidinyl-3,4-dimethoxybenzylamine (6), is formed preferentially to the hydrolyzed product, 4-amino-2-methyl-5-hydroxymethylpyrimidine (7).

**Active Site Analysis of Thiaminase I**—Thiaminase I was inactivated ( $K_I = 7.7 \text{ mM}$ ,  $k_{\text{inact}} = 0.7 \text{ h}^{-1}$ ) by treatment with 2-methyl-4-amino-6-chloropyrimidine (8, Hutter and Slama, 1987). ESI/FT-MS analysis of the inactivated enzyme indicated that all three species were labeled with the pyrimidine moiety and that chloride was lost during the inactivation ( $\Delta$ molecular

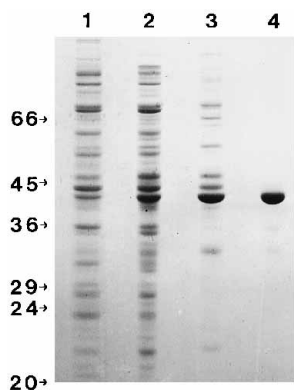


FIG. 8. SDS-polyacrylamide gel electrophoresis analysis of thiaminase I during the purification (12 %) gel. Lane 1, cell lysate; lane 2, 50–70% ammonium sulfate precipitate; lane 3, pooled fractions from the DEAE-ion exchange chromatography; lane 4, pooled fractions from the Cibacron Blue dye affinity chromatography.

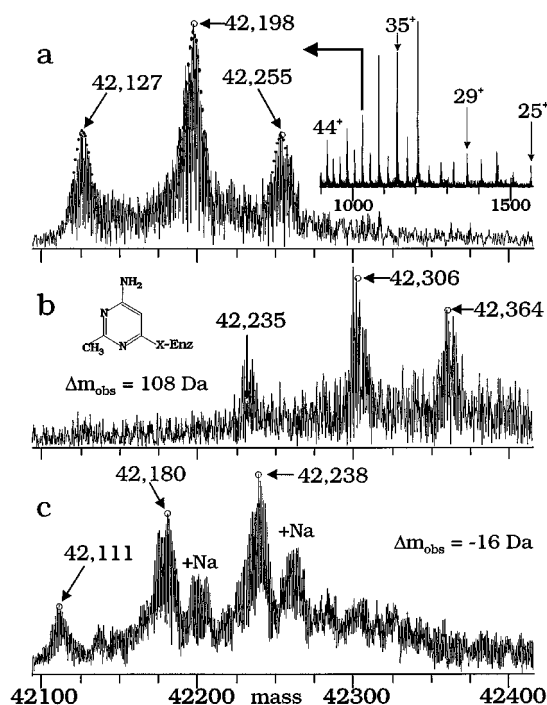
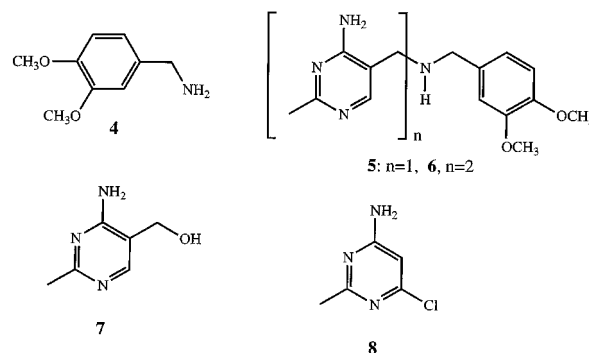


FIG. 9. Electrospray FT-MS of thiaminase I. a, the 39<sup>+</sup> molecular ion region of a broad band ESI/FT mass spectrum (inset) of purified recombinant thiaminase I (20 scans). b, the 38<sup>+</sup> region after incubation of thiaminase I with 4-amino-6-chloro-2-methylpyrimidine (10 scans). c, ESI/FT mass of the thiaminase I C113S mutant (20 scans).

mass = +108, Fig. 9b). Comparison of fragment ion masses from labeled and unlabeled enzyme showed a large carboxyl-terminal fragment and the 99-residue fragment Pro<sup>79</sup>-Thr<sup>177</sup> shifted +108 Da, thus localizing the active site nucleophile to this region. Cys<sup>113</sup> is the most nucleophilic residue contained within this region and was a likely candidate for the active site nucleophile because thiaminase I is inhibited by organomercurials and iodoacetate (Wittliff and Airth, 1970). This was further supported by treating the enzyme with NTSB, a cysteine-specific reagent, which produces the intensely yellow 3-carboxy-4-nitrothiophenylate chromophore ( $\lambda_{\max}$  = 412 nm) after reaction with a thiol. The native enzyme reacted with NTSB with the release of 2 eq of 3-carboxy-4-nitrothiophenylate. Since both the ESI/FT-MS and the DNA sequence data demonstrate that thiaminase I contains only one cysteine residue, we assume that the high thiol titration is due to incomplete removal of tightly bound DTT. The denatured, pyrimi-



dine-labeled, enzyme did not react with this reagent.

Site-directed mutagenesis to make a C113S thiaminase mutant was done according to the procedure of Vandeyar *et al.* (1988) using the T7 *In Vitro* Mutagenesis Kit (U. S. Biochemical Corp.). Due to the cloning sites present in pCAC206, the C113S mutant-containing DNA fragment could not be cloned back directly into pCAC206. The C113S-containing DNA was first cloned into pCAC406, a thiaminase-pET17b derivative, and then the entire gene cloned back into pET22b(+) which yielded pCAC510, the C113S mutant overexpression plasmid (Fig. 7). The C113S mutant was purified in a way similar to the native recombinant enzyme. The mutation was confirmed by electrospray FT-MS which showed the expected 16-Da decrease in molecular mass at isotopic resolution (Fig. 9c) for all three molecular species. Activity assays run on the C113S mutant were done at enzyme concentrations 20-fold in excess of that required to see turnover with native recombinant enzyme (71 nM  $\rightarrow$  356 nM). At this large excess of enzyme, no time-dependent change in absorbance was detected at 252 nm in the presence of thiamin (93  $\mu$ M) and veratrylamine (935  $\mu$ M).

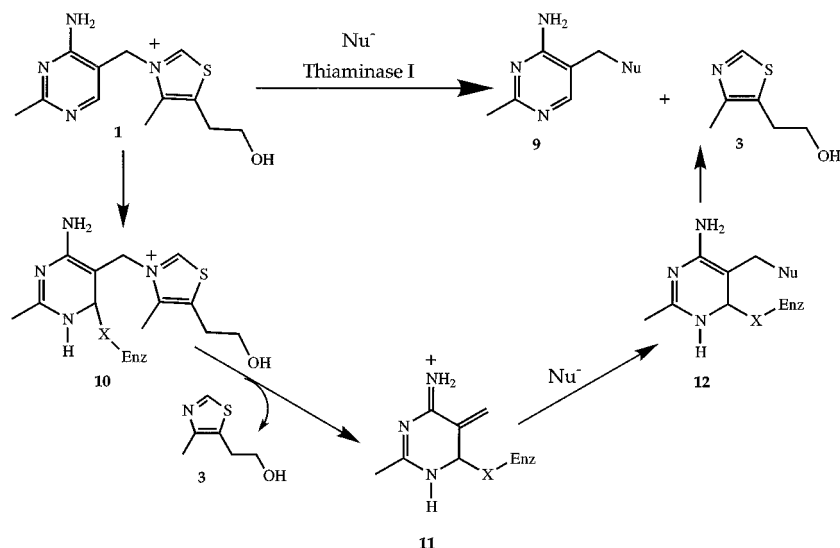
#### DISCUSSION

While thiaminase I from *B. thiaminolyticus* has been previously cloned and overexpressed in *B. subtilis* 168 (Abe *et al.*, 1987), we found the yield of pure enzyme to be low and to vary substantially between preparations. The *E. coli* thiaminase I overexpression strain described here provides a reliable source of large quantities of enzyme for mechanistic studies. Although this enzyme catalyzes the degradation of thiamin, toxicity in *E. coli* is reduced by its secretion into the periplasmic space of the cell. Comparison of the kinetic parameters for the purified recombinant enzyme and thiaminase I purified from *B. subtilis* 168 (H104) shows that the two enzymes are very similar.

ESI/FT-MS on the recombinant enzyme demonstrated that the enzyme was heterogeneous at the amino terminus and localized a single sequencing error to a 150-base pair region (Kelleher *et al.* 1995). The refined sequence was consistent with the molecular mass values, and, moreover, 61 fragment ions ranging from 0.5 to 36 kDa gave extensive sequence verification and could localize any covalent modifications to a small sequence window. The heterogeneity is presumably caused by a signal peptidase that cleaves between residues -3/-2, -2/-1, and -1/+1 during post-translational processing. This has been observed in other systems using Edman sequencing (Hitzeman *et al.*, 1983; Lingappa *et al.*, 1977). Each cleavage site conforms to the "(-3, -1) rule" that defines preferred residues in the -3 (G, A, V, L, I, S, T, M) and -1 (G, A, S, C, T) positions (von Heijne *et al.*, 1984; Perlman and Halvorson, 1983).

Previous studies demonstrated that thiaminase I follows a ping pong mechanism, that catalysis requires the presence of both substrates (Lienhard, 1970; Puzach *et al.*, 1984), and that the substitution reaction occurs with overall retention of configuration at the methylene position (Nicewonger *et al.*, 1995).

FIG. 10. **Mechanistic proposal for thiaminase I.** Addition of an active site nucleophile across the C<sub>6</sub>-N<sub>1</sub> bond would give **10**. Loss of thiazole (**3**) would give **11**. Nucleophilic addition to **11** followed by extrusion of the active site nucleophile completes the reaction. The studies described here demonstrate that cysteine 113 is the active site nucleophile.



From these studies, and the inactivation of the enzyme with 2-methyl-4-amino-6-chloropyrimidine, a mechanism involving the addition of an active site nucleophile to C<sub>6</sub> of the pyrimidine has been proposed (Fig. 10) (Hutter and Slama, 1987; Lienhard, 1970). This mechanistic proposal is also supported by extensive studies on the mechanism of thiamin cleavage with bisulfite (Uray *et al.*, 1993). The independence of the rate on the concentration of veratrylamine suggests that the rate of formation of the enzyme pyrimidine adduct is much slower than the rate of addition of the amine and prevented us from determining the kinetic mechanism of the recombinant enzyme using this assay.

Thiaminase I shows remarkable tolerance to the structure of the nucleophile (Lienhard, 1970). When the reaction is carried out under limiting concentrations of veratrylamine, we have found that the reaction product **5** can function as the nucleophile. We have also found that when thiaminase is treated with thiamin in the absence of any secondary nucleophile, the reaction proceeds for several turnovers with the release of the thiazole (**3**) moiety. We observed no formation of 4-amino-2-methyl-5-hydroxymethylpyrimidine (**7**) suggesting that water is either excluded from the active site or is not nucleophilic enough to react with the proposed intermediate (**11**, Fig. 10). We do not observe any pyrimidine products that migrate from the origin of the TLC plate. One possibility is that the N<sub>1</sub> nitrogen of thiamin is functioning as the nucleophile for this reaction.

The conventional approach to the identification of key active site residues, using radiolabeled suicide substrates, requires kinetic characterization of the inactivation, demonstration of stoichiometric labeling, proteolysis of the labeled enzyme, isolation of the labeled peptide, and sequencing. ESI/FT-MS offers a very attractive alternative strategy. This high resolution technique gives parts per million (ppm) mass accuracy for the intact labeled enzyme which can provide significant insight into the active site chemistry involved in the inactivation event. The stoichiometry of labeling is also evident from the mass shift. Fragmentation of the molecule is carried out in the gas phase by collisional activation (Senko *et al.*, 1994) or irradiation with an IR laser (Little *et al.*, 1994), thus removing the need to radiolabel the substrate, to find suitable proteases, and to purify the labeled fragment. Using this approach, we localized the pyrimidine-labeled residue between Pro<sup>79</sup> and Thr<sup>177</sup>. This peptide contained the enzyme's only cysteine. The observation that the denatured inactivated enzyme did not react with NTSB, while the native enzyme did react strongly sug-

gested that this cysteine was the active site nucleophile. This was confirmed by preparing the C113S mutant which was inactive.

The biological function of thiaminase is unknown (Kimura and Iwashima, 1987). We have considered several possibilities. The proposal that thiaminase serves a defense role by destroying the thiamin in the medium or by catalyzing the biosynthesis of a thiamin antimetabolite fails at least in the case of *E. coli* where the enzyme can be overexpressed with minimal toxic effects on the host. Another possibility is that thiaminase may catalyze the coupling of the pyrimidine and the thiazole moieties to form thiamin. We have demonstrated that thiaminase I was able to overcome a mutation in the coupling enzyme in *E. coli* VJS1391 (*thiE*<sup>-</sup> strain) when grown on minimal media supplemented with ampicillin (200 µg/ml) (Backstrom *et al.*, 1995; Vander Horn *et al.*, 1993). This coupling activity is very low because we have not been able to detect any thiamin phosphate synthase activity with purified thiaminase using the thiochrome assay (Leader, 1970) which is less sensitive than the bioassay. A final possibility is that thiaminase I is involved in the salvage of the thiazole moiety from the medium. This hypothesis remains to be tested.

**Acknowledgment**—We thank Prof. Nishimune for giving us his thiaminase I clone, H.-J. Chiu for running the thiochrome assays on thiaminase I. We thank A. Backstrom, E. DiBella, T. W. Thannhauser, D. P. Little, and R. Nicewonger for their helpful discussions.

#### REFERENCES

- Abe, M., Ito, S., Kimoto, M., Hayashi, R. & Nishimune, T. (1987) *Biochim. Biophys. Acta* **909**, 213–221
- Backstrom, A., McMordie, R. A. & Begley, T. P. (1995) *J. Am. Chem. Soc.* **117**, 2351–2352
- Beu, S. C., Senko, M. W., Quinn, J. P., Wampler, F. M. & McLafferty, F. W. (1993) *J. Am. Soc. Mass Spectrom.* **4**, 557–565
- Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
- Chen, M. & Nagarajan, V. (1994) *J. Bacteriol.* **176**, 5796–5801
- Douthit, H. & Airth, R. (1966) *Arch. Biochem. Biophys.* **113**, 331–337
- Duffy, P., Morris, H. & Neilson, G. (1981) *Am. J. Clin. Nutr.* **34**, 1584–1592
- Earl, J. W. & McCleary, B. V. (1994) *Nature* **368**, 683–684
- Evans, W. C. (1975) *Vitam. Horm.* **33**, 467–504
- Fujita, A. (1954) *Adv. Enzymol.* **15**, 389–421
- Fujita, A. (1972) *J. Vitaminol. (Kyoto)* **18**, 67–72
- Hayashi, R. (1957) *Nutrition Rev.* **15**, 65–67
- Hitzeman, R. A., Leung, R. A., Perry, L. J., Kohr, W. J., Levine, H. L. & Goeddel, D. V. (1983) *Science* **219**, 620–625
- Hutter, J. A. & Slama, J. T. (1987) *Biochemistry* **26**, 1969–1973
- Kelleher, N. L., Costello, C. A., Begley, T. P. & McLafferty, F. W. (1995) *J. Am. Soc. Mass Spectrom.* **6**, 981–984
- Kimura, Y. & Iwashima, A. (1987) *Experientia* **43**, 888–890
- Laemmli, U. K. (1970) *Nature* **227**, 680–685
- Leader, I. G. (1970) *Methods Enzymol.* **18A**, 207–212
- Lienhard, G. (1970) *Biochemistry* **9**, 3011–3020
- Lingappa, V. R., Devillers-Thiery, A. & Blobel, G. (1977) *Proc. Natl. Acad. Sci.*

- U. S. A. **74**, 2432–2436
- Little, D. P., Speir, J. P., Senko, M. W., O'Connor, P. B. & McLafferty, F. W. (1994) *Anal. Chem.* **66**, 2809–2815
- Murata, K. (1982) *Ann. N. Y. Acad. Sci.* **378**, 146–156
- Nicewonger, R., Rammelsberg, A., Costello C. A. & Begley, T. P. (1995) *Bioorg. Chem.*, in press
- Perlman, D. & Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391–409
- Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2444–2448
- Puzach, S., Gorbach, Z. & Ostrovskii, Y. (1984) *Biochemistry (Engl. Transl. Biokhimiya)* **49**, 1010–1016
- Senko, M. W., Speir, J. P. & McLafferty, F. W. (1994) *Anal. Chem.* **66**, 2801–2809
- Thannhauser, T. W., Konishi, Y. & Scheraga, H. A. (1984) *Anal. Biochem.* **138**, 181–188
- Thannhauser, T. W., Konishi, Y. & Scheraga, H. A. (1987) *Methods Enzymol.* **143**, 115–119
- Uray, G., Kriessmann, I. & Zoltewicz, J. A. (1993) *Bioorg. Chem.* **21**, 294–308
- Vander Horn, P. B., Backstrom, A. D., Stewart, V. & Begley, T. P. (1993) *J. Bacteriol.* **175**, 982–992
- Vandeyar, M., Weiner, M., Hutton, C. & Batt, C. (1988) *Gene (Amst.)* **65**, 129–133
- von Heijne, G. (1984) *J. Mol. Biol.* **173**, 243–251
- Wittliff, J. & Airth, R. (1968) *Biochemistry* **7**, 736–744
- Wittliff, J. & Airth, R. (1970) *Methods Enzymol.* **8**, 229–234