*Acta Cryst.* (1995). D**51**, 731–739

# Expression, Characterization and Crystallographic Analysis of Telluromethionyl Dihydrofolate Reductase*

BY JEFFREY O. BOLES

*Department of Chemistry, Tennessee Technological University, Cookeville, TN* 38505, *USA*

KRZYSZTOF LEWINSKI

*Jagiellonian University, Cracow, Poland*

MARCI G. KUNCKLE

*Department of Chemistry and Biochemistry, The University of South Carolina, Columbia, SC* 29208, *USA*

MARCOS HATADA

*Ariad Pharmaceuticals, Cambridge, MA* 02139-4234, *USA*

AND LUKASZ LEBIODA,† R. BRUCE DUNLAP† AND JEROME D. ODOM†

*Department of Chemistry and Biochemistry, The University of South Carolina, SC* 29208, *USA*

(*Received* 11 *October* 1994; *accepted* 24 *January* 1995)

## Abstract

Selenomethionine-containing proteins analyzed by multi-wavelength anomalous diffraction provide a facile means of addressing the phase problem, whose solution is necessary to determine protein structures by X-ray crystallography [Hendrickson (1991). *Science*, **254**, 51–58]. Since this method requires synchrotron radiation, we sought to incorporate a true heavy atom into protein, allowing the solution of the phase problem by more traditional methods of data collection. Media containing TeMet alone or TeMet with low levels of Met failed to sustain growth of a methione auxotroph of *Escherichia coli* carrying the dihydrofolate reductase expression vector. Growth of the organism to stationary phase and incorporation of TeMet was observed when the culture was initiated in media containing minimal Met levels and TeMet was added after induction with isopropyl-1-thio-β-D-galactopyranoside. The purified enzyme exhibited properties similar to those of the native enzyme. Atomic absorption spectroscopy and amino-acid analysis indicated that 40% of the methionines were replaced with TeMet. Sequence analysis did not indicate significant levels of replacement in the first three sites (1, 16 and 20), suggesting that TeMet was present only in the last two sites (42 and 92). Crystals of this enzyme were grown in the presence of methotrexate and were isomorphous with crystals of wild-type dihydrofolate reductase. Difference Fourier maps and restrained least-squares refinement showed no substitution at the first three methionines, while incorporation was seen at positions 42 and 92.

## Abbreviations

DHFR, dihydrofolate reductase; TS, thymidylate synthase; SeMet, selenomethione; TeMet, telluromethionine, $H_2$folate, 7,8-dihydrofolate; $H_4$folate, 5,6,7,8-tetrahydrofolate; wt, wild type; IPTG, isopropyl-1-thio-β-D-galactopyranoside; NADPH, β-nicotinamide adenine dinucleotide phosphate; BSA, bovine serum albumin; PEG, polyethylene glycol; Tris, tris(hydroxymethyl)aminomethane; Hepes, *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; MTX, methotrexate; CFE, cell free extract; NMR, nuclear magnetic resonance; AAS, atomic absorption spectroscopy; MIR, multiple isomorphous replacement; MAD, multiwavelength anomalous dispersion.

## Introduction

Dihydrofolate reductase (DHFR, tetrahydrofolate: $NADP^+$ oxidoreductase, E.C. 1.5.1.3) catalyzes the NADPH-dependent reduction of $H_2$folate to $H_4$folate. This enzyme is a key target for chemotherapeutic agents such as trimethoprim, pyrimethamine and methotrexate (Blakley, 1969; Hughes, 1984; Montgomery & Piper, 1984) because it is a key enzyme in the thymidylate biosynthesis pathway. This pathway serves as the sole *de novo* source of thymidine for a cell; thus, if this cycle is

* This research was supported by NIH Grant GM 42907 and NSF MCB 901814.

† To whom correspondence should be addressed.

interrupted, DNA synthesis is halted and the cell perishes. The *E. coli* enzyme ($M_r$ = 18 kDa) is a monomer of 159 residues, five of which are Met (Baccanari, Stone & Kuyper, 1989). The crystal structure of the *E. coli* enzyme (Bolin, Filman, Matthews, Hamlin & Kraut, 1982) complexed with MTX and NADPH indicates no essential role in binding or catalysis for any of the Met residues in the enzyme.

The crystal structure of DHFR was solved by classical MIR methods (Matthews *et al.*, 1977). Hendrickson and coworkers (Hendrickson, Smith & Sheriff, 1985; Hendrickson *et al.*, 1989; Hendrickson, Horton & LeMaster, 1990; Yang, Hendrickson, Crouch & Satow, 1990) have pioneered innovative approaches for incorporation of heavy atoms into protein in bacterial systems. SeMet is known to produce strong anomalous scattering at the Se absorption edge, thus enabling the solution of crystal structures *via* MAD techniques (Hendrickson *et al.*, 1985, 1989). The potential of this approach has been demonstrated by solving the structures of several SeMet-containing proteins (Hendrickson *et al.*, 1990; Graves *et al.*, 1990; Yang *et al.*, 1990; Hädener *et al.*, 1993). These proteins can also be studied by NMR spectroscopy since the $^{77}$Se isotope exhibits properties that make it an attractive NMR probe (Dunlap & Odom, 1988).

We extended this technology to *E. coli* thymidylate synthase by comparing wt- and SeMet-TS and found there were no perturbations to binding or catalysis as a result of the replacement of 14 Met by SeMet (Boles *et al.*, 1991). In addition, a Fourier difference map of wt- and SeMet-TS revealed no significant positive density other than at the Se positions, providing strong evidence that this replacement does not perturb the protein (Boles *et al.*, 1991). Similarly, no perturbations to structural or binding features were found due to the complete replacement of the five Met residues in *E. coli* DHFR with SeMet (Boles, Tolleson, Schmidt, Dunlap & Odom, 1992). $^{77}$Se NMR analysis of the SeMet DHFR revealed a single resonance for each of the SeMet residues (Boles *et al.*, 1992).

While MAD methodology represents a novel approach to the solution of the phase problem, the requirement for a synchrotron source is a serious drawback. We sought a solution to this problem in incorporating TeMet into proteins, as it should permit data collection without a synchrotron radiation source. $^{125}$Te NMR spectroscopy is also possible for such proteins. However, the use of TeMet for incorporation studies presents a new set of obstacles. First, tellurium is not a trace bioelement and, as a result, few if any tellurium-containing biomacromolecules have been isolated (Liangyau, Kangming, Cangmin & Zheng, 1993). Also, the relative size of the Te atom (1.37 Å) as compared to the Se atom (1.17 Å) in SeMet or the S atom (1.04 Å) in Met could cause structural alterations in a protein. A preliminary communication of these results has been published (Boles, Tolleson, Schmidt, Dunlap & Odom, 1994). This paper describes the incorporation of TeMet into *E. coli* DHFR. We have shown that TeMet partially replaces Met in the enzyme and that such incorporation yields homogeneous protein whose specific activity is 93% that of the wt enzyme.

## Experimental procedures

### *Materials*

Seleno-D,L-methionine, the 20 amino acids commonly found in protein, ampicillin, β-nicotinamide adenine dinuclotide phosphate reduced form, 2-mercaptoethanol, uridine, cytidine, guanosine, adenosine and polyethylene glycol were purchased from Sigma. Tris, potassium chloride and isopropyl 1-thio-β-D-galactopyranoside were purchased from Research Organics. PhastGel gradient 8–25%, PhastGel Blue R, PhastGel native buffer strips, phenyl sepharose CL-4B and Q-sepharose Fast Flow were products of Pharmacia LKB Biotechnology Inc. Methotrexate was purchased from Aldrich. Folic acid was purchased from Calbiochem. $H_2$folate was prepared *via* the reduction of folic acid by sodium hydrosulfite (Futterman, 1957).

### *Telluromethionine synthesis*

L-Telluromethionine was prepared by reaction of optically active α-amino-γ-butyrolactone with lithium methyl tellurolate as described previously (Silks, Boles, Modi, Dunlap & Odom, 1990).

### *Cell cultivation*

*E. coli* DL41 (ΔMet), an *E. coli* K-12 derivative with a lesion in the *met A* gene, was developed by Dr David LeMaster (Hendrickson *et al.*, 1990). The DHFR expression vector pWT8 with an IPTG inducible promotor (amp$^r$) was transfected into DL41 using standard protocols to form *E. coli* DL41 (pWT8) (Maniatis, Fritsch & Sambrook, 1982). This system overproduced DHFR at levels approaching 15% total soluble protein. Media components and inoculation procedures are described elsewhere (Boles *et al.*, 1991, 1992; Hendrickson *et al.*, 1990). Culture *A* was a wt control grown only with L-Met (25 μg ml$^{-1}$). Culture *B* was a control grown only on SeMet (Boles *et al.*, 1992). Conditions for growths *C–F* were modified in the following manner. Each of five flasks containing 1 l growth medium was made 15 μg ml$^{-1}$ in L-Met following innoculation with a 10 ml starter culture of *E. coli* DL41 (pWT8). All cultures were grown at 310 K in these flasks in an environmental shaker. Because of the sensitivity of TeMet to oxygen, shaking was terminated 30 min prior to the stationary phase. After 30 min, IPTG was added to each of the cultures to a final concentration of 1 m*M*. Approximately 10 min later, culture *C* was made at 50

and 5 μg ml$^{-1}$ in TeMet and L-Met, respectively, representing a 10:1 ratio of TeMet to L-Met. Similarly, 20:1 (50 and 2.5 μg ml$^{-1}$) and 50:1 (50 and 1 μg ml$^{-1}$) ratios of TeMet to L-Met were provided to cultures *D* and *E*, respectively, while only TeMet (50 μg ml$^{-1}$) was added to culture *F*. Cells were harvested by centrifugation after 4 additional hours.

*Enzyme assays*

DHFR catalytic activity was monitored at 297 K for 30 s by the decrease in absorbance at 340 nm due to the reduction of $H_2$folate to $H_4$folate after the addition of the enzyme to the substrate premix (Stone & Morrison, 1982). Data were collected with a Shimadzu 2101-PC UV/VIS spectrophotometer and processed with an on-line computer.

*Protein assays*

Protein concentration was determined by the Bradford method (Bradford, 1976) using BSA as the standard. Concentrations of homogeneous DHFR were measured using an extinction coefficient of 27 000 $M^{-1}$ cm$^{-1}$ at 278 nm for the *E. coli* enzyme (Boles *et al.*, 1992).

*Enzyme purification*

All forms of DHFR were purified by chromatography on Q-sepharose followed by phenyl sepharose (Boles *et al.*, 1992). DHFR-containing fractions were determined by spectrophotometric assay. Purity was assessed by native polyacrylamide gel electrophoresis. Purified DHFR was activated by dialysis by reducing the concentration of salts, and thus, restoring full activity.

*Gel electrophoretic procedures*

Both the purity of DHFR and its ability to form a complex with NADPH and methotrexate were assessed by native polyacrylamide gel electrophoresis on PhastGel 8–25% gels using a Pharmacia PhastSystem. For native lanes, 1 μl of 1 mg ml$^{-1}$ DHFR was mixed with 2 μl of 0.1% bromophenol blue. Inhibitory ternary complexes were prepared by incubating 1 ul of 1 mg ml$^{-1}$ DHFR with 1 μl of 1 m*M* NADPH and 1 μl of 2.5 m*M* MTX for 5 min.

*Amino-acid analysis*

DHFR samples were hydrolyzed under nitrogen with 6 *M* HCl for 24 and 48 h at 383 K. The hydrolysis was carried out in the presence of either 3% phenol to protect against Met oxidation, or 1% dodecanethiol to protect tryptophan from degradation. Hydrolysates were dried under vacuum at 298 K, redissolved in sodium citrate buffer, pH 2.20, and analyzed on a Beckman System 6300 analyzer using ninhydrin as reagent. A Beckman System Gold integrator was used for quantitation and data analysis.

*Atomic absorption analysis*

The incorporation of TeMet was determined by AAS using a Perkin-Elmer 4000 flame atomic absorption spectrometer equipped with a Te hollow cathode lamp. Standard curves for Te analysis were developed using solutions of telluric acid [$Te(OH)_6$] or TeMet in water. Standards between 0 and 20 p.p.m. Te were used to generate a linear range for the technique. Matrix effects due to enzyme were tested by creating standard curves in the presence and absence of either 1 or 5 mg ml$^{-1}$ wt DHFR.

*Crystallization and intensity data collection*

Crystals of TeMet DHFR complexed with MTX and $CaCl_2$ were obtained by vapor diffusion in a hanging-drop setup using 10% 2-propanol and 20% PEG 4000 in 0.1 *M* Na–Hepes, pH 7.5 (Crystal Screen No. 41 from Hampton Research, Jancarik & Kim, 1991). Two data sets were collected on crystals from the same crystallization batch. The first data set was collected four months and the second 14 months after protein expression. The first set of diffraction data was collected at the National Synchrotron Light Source, beamline X8C in the Brookhaven National Laboratory, using a 0.98 Å wavelength and a 1024 × 1024 pixel CCD sensor mounted on a Huber four-circle goniostat. The measurements were processed using *MADNES* (Messerschmidt & Pflugrath, 1987) and *PROCOR* (Kabsch, 1988). The data set to 2.5 Å resolution, limited by the aperture of the detector, was obtained using two crystals. The data contained 48 880 observations which, in the averaging procedure, yielded 11 234 unique reflections (88% complete). The $R_{merge}$ values were 0.065 for the full set and 0.058 and 0.043 for each crystal.

Ten months after the initial data set was measured, another set of diffraction data was collected on a crystal from the same crystallization batch. Intensities were measured on a R-AXIS II system with rotating-anode generator as the X-ray source. The averaging procedure yielded $R_{merge}$ = 0.054 and 27 496 independent reflections, of which 11 737, to 2.5 Å resolution, were used in the refinement.

*Refinement*

The restrained least-squares refinement (Hendrickson & Konnert, 1980) was carried out using the *PROLSQ* program. The structure of wt-DHFR was determined initially by Matthews *et al.* (1977). The crystal coordinates of wt-DHFR (Bolin *et al.*, 1982), obtained as entry 4DFR from the Protein Data Bank (Bernstein *et al.*, 1977; Abola, Bernstein, Bryant, Koetzle & Weng,

1987) were the starting model. During the refinement all water molecules with temperature factors exceeding $B = 70\,Å^2$ or with refined occupancy below 0.3 were removed, and water molecules found in difference Fourier maps were added. No Te atoms were in the initial model. The typical values of restraints on the isotropic thermal parameters, 1.0 and $1.5\,Å^2$ for the first and second neighbors, respectively, were used throughout the refinement. The next set of refinements was carried out with Met42 and Met92 fully replaced by TeMet. To estimate the substitution, another series of refinements with fixed values for temperature factors and free occupancies for Te atoms was performed. The initial estimate of Te content in positions 42 and 92 was further verified by refinement of a model that contained Te, O and S in the $\delta$ position, with occupancies equalling 1.0. We were convinced that the most realistic model was 0.1 occupancy of S, which reflects the initial content of Met, and complementary occupancies of Te from TeMet and O from hydrolysis of TeMet. Since the O-atom scattering power is about half of that of the S atom, $O_\delta$ scattering was approximated by using $(S_{\delta\,\text{occupancy}}) = 0.1 + 0.5(O_{\delta\,\text{occupancy}})$. Occupancies and temperature factors were refined in alternate cycles and dumping factors 0.2–0.3 were applied to the shifts. The refinement was stable and the parameters of the four independent Te atoms converged to very similar values as shown by the standard deviations. The distances from Te and S to $C_\gamma$ and $C_\varepsilon$ were restrained while van der Waals interactions between S and Te in the same residue were disabled.

## Results

### *Initial incorporation studies of TeMet into DHFR*

In our previous studies, in which DL41(pWT8) was grown on media containing only SeMet, the growth curve featured a 12 h lag phase (Fig. 1, curve B), which was not found when L-Met-containing media was used (Fig. 1, curve A). Induction with IPTG was most efficient when added to an L-Met-containing culture medium which had reached an optical density of 0.7, followed by 4 h additional growth. However, when SeMet-containing media was employed, overexpression of DHFR was only observed when addition of IPTG was delayed until the onset of stationary phase (2.5–3.0 OD) followed by 4 h of incubation. In our first attempt to produce TeMet-containing protein, we used a medium containing TeMet as the sole methionine source, as we had done with SeMet. However, little or no growth of the bacterium occurred. As an alternative, a culture grown on an L-Met-containing medium was used to inoculate several samples of chemically defined culture media containing increasing levels of TeMet and decreasing levels of L-Met. Unfortunately, these attempts to adapt the bacterium to the presence of TeMet also failed, since no growth was observed. These results indicated that the cells had difficulty tolerating TeMet; thus, we concluded that TeMet, or a metabolite thereof, was somehow toxic to the cell, and that another protocol would have to be developed.

Table 1. *Properties of culture and cell-free extracts of E. coli DL41(pWT8) grown on various amounts of* L-*Met and TeMet*

| | *A** | *B** | *C*† | *D* | *E* | *F* |
|---|---|---|---|---|---|---|
| $OD_{600}$ at Stationary phase | 3.0 | 3.0 | 2.2 | 1.3 | 1.2 | 1.1 |
| Protein‡ (mg) | 540 | 428 | 277 | 250 | 240 | 220 |
| DHFR activity‡ (units) | 3456 | 1071 | 1639 | 500 | <10 | <10 |
| Specific activity (units $mg^{-1}$) | 6.4 | 2.5 | 5.9 | 2.0 | 0.04 | 0.05 |

* Boles *et al.* (1992).

† The values for this culture represent the average of three independent trials.

‡ 2 g *E. Coli* DL41 (pWT8) cells were sonicated and centrifuged to form the CFE in each culture.

### *Cell cultures of E. coli DL41 (pWT8)*

As an alternative, we sought to use our IPTG inducible system to biosynthetically incorporate TeMet at the time when DHFR was overexpressed, thus directing its incorporation into DHFR. A 10 ml inoculum was grown adding only L-Met to the medium. The 1 l culture was then given enough L-Met to permit the bacteria to grow only to mid log phase, at which time IPTG was added. We reasoned that adding TeMet at this point would be most effective, since the bacteria were growing rapidly,

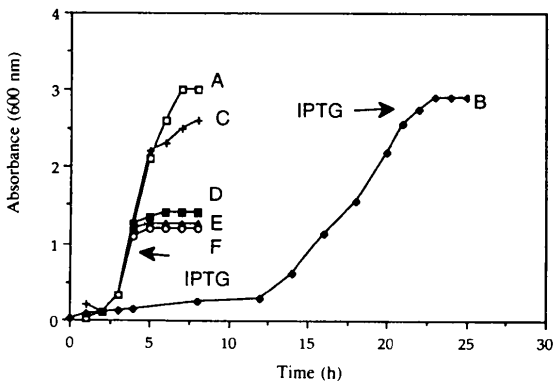


Fig. 1 Growth curves of *E. coli* DL41 (pWT8) on chemically defined media. Culture *A* (□) is a wt control grown only on L-Met containing media. Culture *B* (◆) is grown on defined media containing D,L-SeMet. Cultures *C–F* were grown to mid-log phase on limiting levels of Met, and, following indution with IPTG, were provided various ratios of Met to TeMet (see *Methods* for details). Induction of each culture with IPTG (1 m*M*) is indicated in the figure. Cells were harvested 4 h after induction.

Table 2. *E. coli dihydrofolate reductase purification*

| | Purification step* | Units (μmol $min^{-1}$) | Volume (ml) | Specific activity (units $mg^{-1}$) | Total protein (mg) | Recovery (%) | Purification (fold) |
|---|---|---|---|---|---|---|---|
| wt DHFR | Cell-free extract | 3456 | 7.5 | 6.4 | 540 | 100 | 1 |
| | Q-sepharose | 3110 | 50 | 39 | 80 | 92 | 6.0 |
| | Phenyl sepharose | 3041 | 80 | 45 | 68 | 89 | 7.0 |
| SeMet DHFR | Cell-free extract | 1071 | 7.5 | 2.5 | 428 | 100 | 1 |
| | Q-sepharose | 980 | 40 | 38 | 26 | 90 | 15.2 |
| | Phenyl sepharose | 950 | 80 | 44 | 23 | 88 | 18.0 |
| TeMet DHFR† | Cell-free extract | 1639 | 20 | 5.9 | 277 | 100 | 1 |
| | Q-sepharose | 1466 | 96 | 16.7 | 87.6 | 87 | 2.8 |
| | Phenyl sepharose | 871 | 35 | 40 | 21.8 | 53 | 6.8 |

* 2 g *E. coli* DL41(pWT8) cells were sonicated and centrifuged to form the CFE in each purification.
† These results represent the average of three independent trials.

were primed to overexpress DHFR, and were in need of Met.

The growth curves for cultures *C–F* (Fig. 1) are representative of the extent of cell proliferation under each condition studied and can readily be compared to those for Met (curve A) and SeMet (curve B). It is clear from the low $OD_{600}$ values that little or no further growth occurred in cultures *D*, *E* and *F* after induction and addition of TeMet–L-Met mixtures (Fig. 1 and Table 1). However, in culture *C*, where IPTG induction was followed by the addition of a 10:1 ratio of TeMet to L-Met, the cell density increased to within 90% of that of the wt control. These data alone suggest that the cells in culture *C* were successfully undergoing growth, cell division, and the process of protein biosynthesis of TeMet.

Approximately 1 h after addition of the TeMet–L-Met mixture, cultures *C–F* turned from a faint yellow color to one of greyish black. Following centrifugation of these cultures the supernatant exhibited a faint yellow color typically obtained with Met- or SeMet-containing cultures, while the cell pellets were greyish black. By comparison, cells harvested from media containing either L-Met or SeMet gave light tan cell pellets. This suggested that cultures *C–F* had assimilated at least some of the TeMet and that some of the TeMet inside the bacterial cells had decomposed or was metabolized to yield elemental Te, which has a greyish black color. When this protocol was repeated but with the cessation of shaking near stationary phase, the cell pellets were light tan.

As a more direct measure of the extent of protein biosynthesis in cultures *C–F*, we employed measurements of protein levels and the specific activity of DHFR in the cell-free extracts of these cultures. 2 g samples of *E. coli* DL41(pWT8) were sonicated and centrifuged to produce the CFE for each culture. As seen in Table 1, the specific activity of the CFE from culture *C* was 6 units $mg^{-1}$. This was much higher than the specific activity of DHFR in cultures *D*, *E* and *F*, indicating a lower level of overexpression of DHFR in these cultures. This illustrates that higher ratios of TeMet to Met interfere with protein synthesis and cell growth. The specific activity of DHFR in the CFE from culture *C* was comparable with that from culture *A*, but was 2.4 times the value for culture *B*. Also, the diminished levels of total protein in the CFE of cultures *C–F* indicate that the TeMet has some adverse effect on the ability of the cell to not only overexpress DHFR, but to synthesize other proteins as well. There is a 50% decrease in total protein and approximately twofold less DHFR units in the CFE of culture *C* compared to that for the wt culture *A*. TeMet also seems to make the cells express variable amounts of protein in cultures grown with a 10:1 ratio of TeMet to L-Met. In particular, overall protein levels in *C*-type cultures ranged from 46 to 56% of total protein observed in culture *A* while DHFR units ranged from 45 to 51% compared to the CFE for culture *A*. These results further confirm the deleterious effect that TeMet has on protein biosynthesis in cells.

### *Purification of dihydrofolate reductase*

For this report we chose to purify and perform characterization studies on the DHFR produced from culture *C* since levels of DHFR in the CFE of this culture were of comparable specific activity to enzyme obtained from media enriched with only L-Met. Cell pellets from culture *C* were taken up in buffer, sonicated, and centrifuged to yield the CFE, which was then subjected to successive chromatography on Q-sepharose and phenyl sepharose resins (Boles *et al.*, 1992). Just as SeMet DHFR eluted under the same conditions as wt DHFR from both columns, DHFR isolated from culture *C* did also (Table 2). Enzyme purified from culture *C*

exhibited a specific activity of 42 units $mg^{-1}$ while pure SeMet- and wt-DHFR yielded specific activities of 44 and 45 units $mg^{-1}$, respectively (Table 2) (Boles *et al.*, 1992). The resulting *E. coli* TeMet DHFR was shown to be pure by native polyacrylamide gel electrophoresis and was found to form the usual ternary complex with NADPH and MTX, just as the Met- and SeMet-containing enzymes do (data not shown).

*Atomic absorption analysis of tellurium*

To test for the presence of Te, a standard curve from 0 to 20 p.p.m. Te was developed using TeMet and $Te(OH)_6$ as standards. Both standards gave linear results, indicating that the chemical status of Te in the sample did not affect the readings obtained from this technique. After addition of either 1 or 5 mg $ml^{-1}$ wt-DHFR, no adverse matrix effects were observed, as the AAS readings remained linear (data not shown). Also, samples of wt-DHFR were tested alone and these readings were negligible, indicating the absence of Te in the wt enzyme. The fractions from the phenyl sepharose column containing DHFR activity were subjected to AAS. Molar ratios of Te:DHFR between 1.6 and 2.2 were observed, suggesting the association of Te with the protein. Additionally, the profiles of DHFR (both activity and protein) and Te in the fractions were symmetrical with one another (Fig. 2), further supporting the notion that incorporation of Te had occurred. Following a 24 h dialysis of the enzyme against 3 l of 50 m*M* Tris buffer, we determined that the molar ratio Te:protein had not changed, indicating that the Te was at least tightly associated with the protein and not in equilibrium with free Te in solution. The enzyme was then subjected to denaturation by guanidine–HCl, followed by AAS analysis. The results showed that the molar ratio of Te:protein remained in the range initially determined, thus strongly supporting the covalent association of Te with the enzyme. Furthermore, AAS measurements of the used dialysis buffers from the previous experiments showed the absence of tellurium, indicating that the association between Te and the protein was both strong and stable (covalent), whether the protein was in its native state or unfolded.

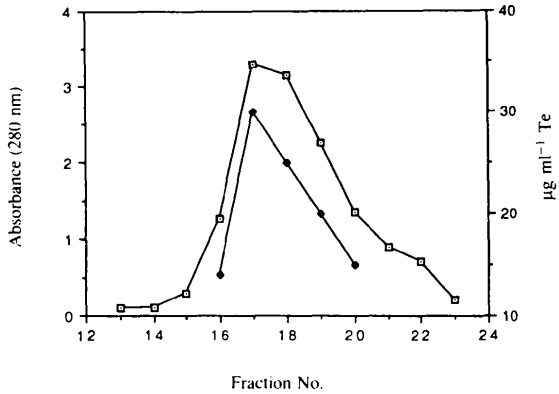


Fig. 2. Molar ratios of Te: DHFR between 1.6 and 2.2 were observed in the DHFR fractions eluting from the phenyl sepharose column. Absorbance (280 nm) (□), mol Te/mol (◆).

Table 3. *Methionine analysis of wt- and TeMet-DHFR (moles of Met per mole of pure DHFR)*

| | 3% Phenol 24 h | 3% Phenol 48 h | 1% DT 24 h | 1% DT 48 h |
|---|---|---|---|---|
| wt-DHFR | 4.90 | 5.00 | 4.10 | 3.30 |
| Te-DHFR | 2.95 | 3.06 | 2.83 | 2.80 |

*Amino-acid and N-terminal sequence analysis of DHFR*

Amino-acid analysis of DHFR purified from culture *C* revealed the presence of 3.2±0.2 mol L-Met per mol DHFR (TeMet) as compared to 4.9 mol Met per mol DHFR (L-Met). These data (Table 3) suggest the presence of 1.8 to 2.0 mol TeMet, which is coincident with the amount determined by AAS analysis. Both SeMet and TeMet are degraded during amino-acid analysis; thus comparison of the values obtained for a SeMet- or TeMet-containing sample with a wt sample indicates how much of the heavy-atom derivative is present by the difference in the two values. This analysis is a good indirect method for TeMet detection in that the absence of L-Met suggests the incorporation of TeMet into these positions. The results of initial N-terminal sequence analysis studies reveal that the first 22 residues of the TeMet-containing enzyme are the same as those of the wt enzyme (Smith & Calvo, 1980) and that similar yields of Met phenylthiohydantoins were obtained for both enzyme samples (data not shown), thus suggesting that it is only the last two Met sites (42 and 92) which contain TeMet.

*Crystallization and crystal data*

TeMet-DHFR crystals grew as hexagonal bipyramids of size up to 0.4 mm. Despite different crystallization conditions than those used for wt-DHFR (Bolin *et al.*, 1982), the crystals obtained were isomorphous with those of wt-DHFR. The unit-cell parameters, $a = b = 93.0$, $c = 74.4$ Å, compare well to those reported for wt-DHFR, $a = b = 93.27$, $c = 73.56$ Å. The crystals belong to space group $P6_1$ and the asymmetric part of the unit cell contains two molecules.

*Refinement*

In the refinement of the wt-DHFR model against the TeMet-DHFR data, the number of water molecules decreased from 428 to 343 used in the final refinement.

The refinement converged at $R = 0.132$ and the refined structure showed no significant conformational differences when compared to the wt-DHFR model. However, the S and C atoms in Met42 and Met92 showed much lower temperature factors than the average values obtained for these and adjacent residues (Fig. 3). In contrast, there were no anomalies in the temperature factors of S atoms in Met1, Met16 or Met20. The difference Fourier maps, $F_o - F_c$, showed pronounced electron densities around the S-atom positions of Met42 and Met92 in both independent molecules, and no density higher than the $2\sigma$ level at the other Met sites. These results indicated at least partial substitution of Met by TeMet in residues 42 and 92 and little or no substitution for Met1, 16 and 20. The results were consistent for both independent molecules.

The next series of refinements was carried out with a model in which Met42 and Met92 were fully replaced by TeMet. This refinement resulted in an increase of the temperature factors for Te atoms to 50–55 Å$^2$, also indicating partial occupancy. Refinements of the model with variable occupancies of Te atoms converged at the average occupancy for the four Te atoms, 0.61 (3) for the first data set and 0.33 (3) for the second data set. A difference Fourier, $F_o$(data set I) $- F_o$(data set II), showed positive peaks at the $5\sigma$ and $3\sigma$ levels at the sulfur positions of Met42 and Met92, respectively. This indicated slow degradation, probably hydrolysis, of TeMet in the crystals. Thus, Te atoms, O atoms from the hydrolysis product, and probably S atoms from normal Met occupy the same positions. On the basis of the AAS measurements, amino-acid composition, and sequence analysis of the protein, the content of Met at positions 42 and 92 must be $\sim$ 10% if it reflects the ratio of Met to TeMet in the medium. The lowest estimate of Te content for the first data set was 0.38 (3), obtained assuming complementary occupancies of Te and S. The most realistic model, 0.1 occupancy of S and complementary occupancies of O and Te, converged at 0.48 (3) Te for the first data set. The $R$ value was 0.124 and deviations of bond lengths and bond angles from the ideal values were 0.015 Å and 2°, respectively.

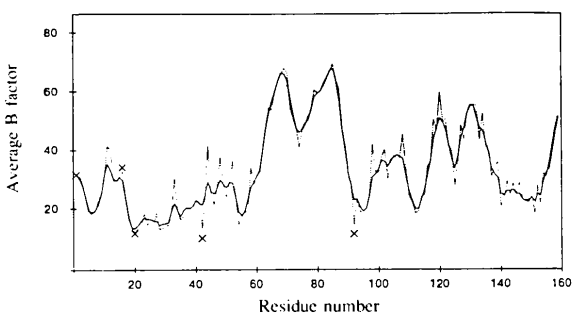


Fig. 3. Average temperature factors as a function of residue number after refinement without Te. The solid lines is for main-chain atoms and the broken line is for side-chain atoms. Crosses represent temperature factors for the S atoms of Met residues. It is apparent that the S atoms in Met 42 and 92 have abnormally low temperature factors indicating additional scattering power.

Table 4. *Summary of parameters and results of restrained refinement*

| | $\sigma$ | $\Delta$ |
|---|---|---|
| Distances (Å) | | |
| Bond lengths (1–2) neighbors) | 0.020 | 0.018 |
| Bond angles (1–3 neighbors) | 0.040 | 0.059 |
| Planes (1–4 neighbors) | 0.050 | 0.062 |
| Planar groups (Å) | 0.020 | 0.013 |
| Chiral volume (Å$^2$) | 0.150 | 0.178 |
| Non-bonded contacts (Å) | | |
| Single-torsion contacts | 0.300 | 0.220 |
| Multiple-torsions | 0.300 | 0.226 |
| Possible hydrogen contacts | 0.300 | 0.207 |
| No. of water molecules | 343 | |
| Final $R$ factor | 0.124 | |
| Torsion angles (°) | | |
| Peptide plane | 3.0 | 2.2 |
| Staggered | 15.0 | 21.5 |
| Orthonormal | 20.0 | 23.5 |

The temperature factors for Met42 and Met92 side chains are close to the average values in their environment. The final temperature factors (not shown) were very close to those presented, except for positions 42 and 92. The final difference Fourier map did not have any significant features at either positive or negative $2\sigma$ contouring levels at Met1, Met16 and Met20. The final values of the refinement parameters are given in Table 4.

## Discussion

The most fascinating aspects of TeMet bioincorporation into proteins are that it actually occurs and the serendipitous finding that the isolated protein has an uneven distribution of TeMet in the peptide chain. Why is TeMet detected in positions 42 and 92 while it is not at positions 1, 16 or 20 in purified active DHFR? We must assume that TeMet incorporation occurs randomly at positions normally occupied by Met with the possible exception of the N-terminal Met. The issue of whether or not TeMet can be utilized to initiate protein synthesis is unresolved, since there is no direct evidence that TeMet can be successfully *N*-formylated and that *N*-formylated TeMet can be loaded on the specific tRNA employed to begin translation. Obviously following induction with IPTG, the overexpression of the DHFR gene and the consequent synthesis of DHFR occur throughout the course of the remaining culture period, presumably in a nonsynchronized fashion. There is no evidence to support the notion that the *E. coli* exhausted the supply of authentic Met in the middle of translation for each DHFR molecule, leaving only TeMet to occupy the remaining Met positions. If this were occurring, one

would expect to find DHFR with TeMet in the first three positions, especially when a tenfold excess of TeMet was used. Initial studies with methionyl tRNA synthetase suggest that Met, SeMet and TeMet are activated and loaded with similar efficiencies onto the tRNA (C. W. Carter Jr, personal communication). Thus, once formed, Met tRNA and TeMet tRNA should be employed indiscriminately by the translation system, yielding a mixture of DHFR's featuring random substitution of Met residues by TeMet, perhaps with the exception of Met1, which uses a different tRNA.

It could be speculated that DHFR molecules with TeMet in positions 1 and/or 16 and/or 20 were produced, but not isolated by our purification process. This would be possible if TeMet at these positions led to improperly folded molecules. Such molecules might be marked for degradation, possess no enzymatic activity, or not purify with active DHFR. It has been proposed that the protein folding process is more sensitive to subtle alterations occurring near the N-terminus. On the other hand, the survey of DHFR sequences (Blakley, 1984) indicates that none of the Met residues is conserved, even if only bacterial proteins are analyzed. Thus, it appears that the folding is not very sensitive to sequence alteration in these positions. Even more drastic changes in this region of the molecule were introduced in the DL1 mutant of the *E. coli* DHFR where the loop Met16–Ala19 was replaced with a glycine residue. Yet the enzyme retained some activity and, as indicated by $^1$H NMR spectoscopy, was properly folded (Li, Falzone, Wright & Benkovic, 1992). Only very small conformational changes around buried TeMet are observed and one would expect that it would be even easier to accommodate TeMet in surface residues. In our opinion, there is no obvious hypothesis explaining the observation that the presence of accessible TeMet, or perhaps a product of its hydrolysis, marks DHFR molecules for fast degradation. Studies of other proteins with incorporated TeMet should provide more data and enable us to solve this fascinating problem.

The slow degradation of incorporated TeMet with time is observed. Clearly, for future structure determinations it is essential to proceed from protein expression to crystallographic data collection as fast as possible. It would perhaps be inappropriate to extrapolate back to the contents of TeMet in fresh crystals on the basis of two measurements established by crystallography. However, there is no TeMet in the first three Met residues and there is full occupancy of L-Met. Taking the contents of TeMet measured by AAS in fresh protein (2 mol Te per mol DHFR), it is likely that the TeMet occupancy in fresh DHFR at Met42 and Met92 is close to one. This would significantly increase chances of obtaining useful heavy-atom derivatives for the MIR method.

These exciting new results indicate that TeMet can be successfully incorporated into DHFR with little or no alteration of its structure or function. We think that this scheme for TeMet incorporation could be applied to many systems and could generate materials that will greatly facilitate the analysis of protein structure and function by X-ray crystallography and NMR spectroscopy. Indeed, we think our procedure for incorporation of TeMet following IPTG induction could be extended to the incorporation of other amino acids bearing reporter groups, including stable isotopic species or radiolabels. Obviously, the extent of incorporation may be enhanced under different growth conditions in various medium enrichments. We are currently optimizing TeMet incorporation into selected proteins and determining the feasibility of observing $^{125}$Te-NMR signals from TeMet-containing proteins.

## References


ABOLA, E. E., BERNSTEIN, F. C., BRYANT, S. H., KOETZEL, T. F. & WENG, J. (1987). *Crystallographic Databases – Information Content, Software Systems, Scientific Applications*, edited by F. H. ALLEN, G. BERGERHOFF & R. SIEVERS, pp. 107–132. Bonn/Cambridge/Chester: IUCr.

BACCANARI, D. P., STONE, D. & KUYPER, L. (1981). *J. Biol. Chem.* **256**, 1738–1747.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BLAKLEY, R. L. (1969). *The Biochemistry of Folic Acid and Related Pteridines*, edited by A. NEUBERGER & E. L. TATUM, pp. 464–508. New York: John Wiley.

BLAKLEY, R. L. (1984). *Folates and Pterins*, Vol. 1, *Chemistry and Biochemistry of Folates*, edited by R. L. BLAKLEY & S. J. BENKOVIC, pp. 191–253. New York: John Wiley.

BOLES, J. O., CISNEROS, R. J., WEIR, M. S., ODOM, J. D., VILLAFRANCA, J. E. & DUNLAP, R. B. (1991). *Biochemistry*, **30**, 11073–11080.

BOLES, J. O., LEWINSKI, K., KUNKLE, M., HATADA, M., ODOM, J. D., DUNLAP, R. B. & LEBIODA, L. (1994). *Nature Struct. Biol.* **1**, 283–284.

BOLES, J. O., TOLLESON, W. H., SCHMIDT, J. C., DUNLAP, R. B. & ODOM, J. D. (1992). *J. Biol. Chem.* **267**, 22217–22223.

BOLIN, J. T., FILMAN, D. J. MATTHEWS, D. A., HAMLIN, R. C. & KRAUT, J. (1982). *J. Biol. Chem.* **257**, 13650–13662.

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

DUNLAP, R. B. & ODOM, J. D. (1988). *Phosphorus Sulfur*, **38**, 217–229.

FUTTERMAN, S. (1957). *J. Biol. Chem.* **228**, 1031–1038.

GRAVES, B. J., HATADA, M. H., HENDRICKSON, W. A., MILLER, J. K., MADISON, V. S. & SATOW, Y. (1990). *Biochemistry*, **29**, 2679–2684.

HÄDENER, A., MATZINGER, P. K., MALESHKEVICH, V. N., LOUIE, G. V., WOOD, S. P., OLIVER, P., ALEOUNDER, P. R., PITT, A. R., ABELL, C. & BATTERSBY, A. R. (1993). *Eur. J. Biochem.* **211**, 615–624.

HENDRICKSON, W. A. (1991). *Science*, **254**, 51–58.

HENDRICKSON, W. A., HORTON, J. R. & LEMASTER, D. M. (1990). *EMBO J.* **9**, 1665–1672.

HENDRICKSON, W. A., PAHLER, A., SMITH J. L., SATOW, Y., MERRIT, E. A. & PHIZACKERLY, R. P. (1989). *Proc. Natl Acad Sci. USA*, **86**, 2190–2194.

HENDRICKSON, W. A., SMITH, J. L. & SHERIFF, S. (1985). *Methods Enzymol.* **115**, 41–55.

HUGHES, W. T. (1984). *Folate Antagonists as Therapeutic Agents*, Vol. 2, edited by F. M. SIROTNAK, J. J. BURCHALL, W. B. ENSMINGER & J. A. MONTGOMERY, pp. 253–265. New York: Academic Press.

JANCARIK, J. & KIM, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

KABSCH, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

LI, L., FALZONE, C. J., WRIGHT, P. E. & BENKOVIC, S. J. (1992). *Biochemistry*, **31**, 7826–7833.

LIANGYAU, Y., KANGMING, H., CANGMIN, Y. & ZHENG, O. (1993). *Anal. Biochem.* **209**, 318–322.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning*, edited by T. MANIATIS, pp. 75–86. New York: Cold Spring Harbor Laboratory Press.

MATTHEWS, D. A., ALDEN, R. A., BOLIN, J. T., FREER, S. T., HAMLIN, R., XUONG, N., KRAUT, J., POE, M., WILLIAMS, M. & HOOGSTEEN, K. (1977). *Science*, **197**, 452–455.

MESSERSCHMIDT, A. & PFLUGRATH, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.

MONTGOMERY, J. A. & PIPER, J. R. (1984). *Folate Antagonists as Therapeutic Agents*, Vol. 1, edited by F. M. SIROTNAK, J. J. BURCHALL, W. B. ENSMINGER & J. A. MONTGOMERY, pp. 109–116. New York: Academic Press.

SILKS, L. A., BOLES, J. O., MODI, B. P., DUNLAP, R. B. & ODOM, J. D. (1990). *Syn. Comm.* **20**(10), 1555–1562.

SMITH, D. R. & CALVO, J. M. (1980). *Nucl. Acids Res.* **8**, 2255–2274.

STONE, R. S. & MORRISON, J. F. (1982). *Biochemistry*, **21**, 3757–3765.

YANG, W., HENDRICKSON, W. A., CROUCH, R. J. & SATOW, Y. (1990). *Science*, **249**, 1398–1405.