Primary- and secondary-structural analysis of a unique prokaryotic metallothionein from a *Synechococcus* sp. cyanobacterium

Robert W. OLAFSON,* William D. McCUBBIN† and Cyril M. KAY†

*Department of Biochemistry and Microbiology, University of Victoria, Victoria, B.C. V8W 2Y2, Canada, and †Medical Research Council Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

Biochemical and physiological studies of *Synechococcus* cyanobacteria have indicated the presence of a low- M_r heavy-metal-binding protein with marked similarity to eukaryotic metallothioneins (MTs). We report here the characterization of a *Synechococcus* prokaryotic MT isolated by gel-permeation and reverse-phase chromatography. The large number of variants of this molecule found during chromatographic separation could not be attributed to the presence of major isoproteins as assessed by amino acid analysis and amino acid sequencing of isoforms. Two of the latter were shown to have identical primary structures that differed substantially from the well-described eukaryotic MTs. In addition to six long-chain aliphatic residues, two aromatic residues were found adjacent to one another near the centre of the molecule, making this the most hydrophobic MT to be described. Other unusual features included a pair of histidine residues located in repeating Gly-His-Thr-Gly sequences near the C-terminus and a complete lack of association of hydroxylated residues with cysteine residues, as is commonly found in eukaryotes. Similarly, aside from a single lysine residue, no basic amino acid residues were found adjacent to cysteine residues in the sequence. Most importantly, sequence alignment analyses with mammalian, invertebrate and fungal MT sequences showed no statistically significant homology aside from the presence of Cys-Xaa-Cys structures common to all MTs. On the other hand, like other MTs, the prokaryotic molecule appears to be free of α -helical structure but has a considerable amount of β -structure, as predicted by both c.d. measurements and the Chou & Fasman empirical relations. Considered together, these data suggested that some similarity between the metal-thiolate clusters of the prokaryote and eukaryote MTs may exist.

INTRODUCTION

Metallothioneins (MTs) are a group of increasingly well-characterized metalloproteins that complex an array of metal ions in polynuclear thiol cluster arrangements (Otvos *et al.*, 1982). Although the precise physiological role(s) of MTs is still unknown, an abundance of information now exists implicating them in an array of cellular functions served by the above-mentioned metal-complexing capabilities (Olafson, 1980, 1983, 1985; Karin *et al.*, 1980; Leone *et al.*, 1985). The zinc homeostasis role alone is of considerable physiological importance in that the activity of a very large number of metalloproteins depends upon the ready availability of this element. Both DNA polymerases and RNA polymerases, for example, require zinc for activity (Slater *et al.*, 1971; Scrutton *et al.*, 1971). Since these proteins are found in all tissues of mammals (Olafson, 1981; Zelazowski & Piotrowski, 1977), and ubiquitously throughout the eukaryotic phyla (Olafson & Thompson, 1974; Olafson *et al.*, 1979; Kägi & Nordberg, 1979a), it is reasonable to assume that MTs occupy a significant position in cellular metabolism. For this reason, aside from the question of physiological function, research activity involving MTs has focused on structure-function relationships in an attempt to understand the molecular

aspects of the metal-protein interactions. For example, n.m.r. spectroscopy and X-ray-diffraction studies have helped elucidate the chemical nature of the metal-thiolate clusters of several eukaryotic MTs (Otvos & Armitage, 1980; Vasak *et al.*, 1984; Neuhaus *et al.*, 1985; Furey *et al.*, 1986) by taking advantage of a growing number of primary structures (Kägi & Nordberg, 1979b; Lerch *et al.*, 1982).

Until recently, the growing number of eukaryotic MT primary structures reported were found to be highly conserved, such that this was becoming an accepted trait of this group of metalloproteins. Newer findings, however, have suggested that eukaryotic MTs are not a single highly conserved group but exist in several forms (Butt *et al.*, 1984; Nemer *et al.*, 1985). The latter structures offer significant opportunity to improve our understanding of metal-thiolate complexes in proteins and ultimately our appreciation of MT structure-function relationships.

We earlier reported preliminary findings concerning the N-terminal sequence of a cyanobacterial protein (Olafson, 1984) whose physiological characteristics were very similar to those of the eukaryotic MTs (Olafson *et al.*, 1979, 1980). The present paper details the isolation and complete amino acid sequence analysis of a prokaryotic MT and compares this structure with known

Abbreviation used: MT, metallothionein.

* To whom correspondence and reprint requests should be sent.

eukaryotic proteins by sequence alignment analysis and c.d. spectroscopy. These data describe a unique MT with no apparent evolutionary relationship in primary structure to the eukaryotic MTs.

EXPERIMENTAL

Materials

Iodoacetate was obtained from Sigma Chemical Co., and iodo[¹⁴C]acetate was purchased from New England Nuclear. The 75 mm 30 nm-pore-size propyl and 250 mm octyldecylsilane reverse-phase h.p.l.c. columns were purchased from Beckman Instruments, and the 250 mm cyano column was obtained from Hamilton. Carboxypeptidase Y and *Staphylococcus* V8 proteinase were obtained from Boehringer Mannheim.

Isolation of *Synechococcus* MT

Synechococcus strain Tx-20 (Stevens & Myers, 1976) was obtained from Dr. W. F. Doolittle and cultured axenically in BG-11 medium as described previously (Olafson, 1984). French-pressure-cell lysates were made in the presence of 0.5 M-Tris/HCl buffer, pH 8.6, and 40000 g supernatants were chromatographed on a 5 cm × 100 cm Sephadex G-75 column developed with 10 mM-NH₄HCO₃ containing 5 mM-2-mercaptoethanol. Further purification employed a 30 nm-pore-size propylsilane reverse-phase column developed with 20 mM-KH₂PO₄, pH 7.0, with acetonitrile as the organic modifier. Portions (3 mg) of freeze-dried sample were applied to the reverse-phase column, which was developed with a very shallow acetonitrile gradient as indicated in the Results section. Chromatographic elution of MT was monitored spectrophotometrically at 250 nm and polarographically by using the Brdička procedure for thiol groups, as previously described (Olafson, 1981). Reverse-phase fractions were desalted on Sephadex G-25 and freeze-dried. As previously reported, MT purified in this manner was found to be electrophoretically homogeneous on SDS/polyacrylamide-gel electrophoresis (Olafson *et al.*, 1979; Olafson, 1984).

Amino acid analysis

Amino acid analyses were performed on a Beckman analyser with hydrolysed protein depleted of metal ions by brief exposure to 0.1 M-HCl and 6 M-guanidinium chloride followed by Sephadex G-25 chromatography in the presence of 30 % formic acid (Olafson, 1984). Valine, isoleucine and leucine values were determined from 72 h hydrolysates, and threonine and serine values were determined from 18 h hydrolysates. Cysteine was determined independently as the cysteic acid derivative. Since purified protein had no 280 nm absorption, tryptophan was not determined.

Amino acid sequence analysis

Sequence analyses were performed by the use of both standard spinning-cup and gas-phase methodologies on metal-free protein modified by either reduction and carboxymethylation (Crestfield *et al.*, 1963) or reduction and pyridylethylation (Hermodson *et al.*, 1973). Positions of cysteine residues were corroborated by [¹⁴C]carboxymethylation and equivocal residues were verified by determination on both octyldecylsilane and cyano

columns as described elsewhere (Black & Coon, 1982; Hunkapiller & Hood, 1983). All sequence analyses were confirmed in either duplicate or triplicate runs.

Proteinase digestion and peptide isolation

Carboxypeptidase Y digestion was used for determination of the C-terminus. A 90 nmol sample of [¹⁴C]carboxymethylated MT was digested in 0.2 M-pyridine/acetate buffer, pH 5.6 with a 20:1 molar ratio of substrate to enzyme for 3 h at 37 °C. The reaction was halted by the addition of an equal volume of acetic acid, and the sample was applied directly to an amino acid analyser, or after formation of phenylthiohydantoin derivatives to the reverse-phase h.p.l.c. system described above.

Staphylococcus-V8-proteinase digests were conducted on 200 nmol of metal-free pyridylethylated protein dissolved in a 30:1 molar ratio of substrate to enzyme in 50 mM-ammonium acetate buffer, pH 4.0 for 18 h at 25 °C. Peptides were separated by reverse-phase h.p.l.c. with 10 mM-NaClO₄/15 mM-H₃PO₄, pH 2.0, in the aqueous mobile phase. The applied acetonitrile organic modifier gradient was as shown in the Results section.

Spectropolarimetric analysis

C.d. measurements were made on a Jasco 500-C spectropolarimeter interfaced with a DP 500N data processor. The cell was maintained at 25 °C by means of a Lauda K-2/R circulating water bath. Far-u.v. (250–195 nm) scans were performed in a cell with a pathlength of 0.05 cm. The aromatic region (320–250 nm) was monitored by using a 90 µl microcell with a pathlength of 1 cm. Eight scans were normally run. The computer-averaged trace with subtracted solvent baseline was employed in all calculations. The instrument was routinely calibrated with solutions of (+)-camphor-10-sulphonic acid or D(–)-pantoyl-lactone of known concentrations. MT concentrations in solution were established by catalytic wave-differential pulse-polarographic determination of thiol content (Olafson, 1981).

Computer-assisted analyses

In addition to plotting the far-u.v. c.d. data in terms of mean residue ellipticity versus wavelength the secondary structure of MT was determined by using the main-frame-driven FORTRAN program CONTIN, developed by Provencher & Glöckner (1981), which analyses c.d. spectra as a sum of data collected from 16 proteins whose structures are known from X-ray crystallography. The input to the program was the mean residue ellipticities in 1 nm intervals from the minimum value measured to 240 nm. Predictions for the secondary structure of MT were based upon the Dufton & Hider (1977) modification of the Chou & Fasman (1974) method. A program to handle these calculations was written by Mr. D. Ross and was run on an IBM PC XT computer.

A composite surface profile was drawn showing the regions of sequence most likely to lie on the surface of the protein. This computer-generated profile is a composite of three approaches, which include h.p.l.c. hydrophilicity parameters (Parker *et al.*, 1986), accessibility (Janin, 1979) and flexibility (Karplus & Schulez, 1985). The program used to generate these data was obtained from the Alberta Peptide Institute (Department of Biochemistry, University of Alberta).

Fig. 1. Reverse-phase h.p.l.c. chromatogram of protein isolated by Sephadex G-75 chromatography

MT was found to occur in fractions 3–10. The arrow indicates an increase in chart speed. Organic modifier (B) was acetonitrile. Metal concentrations in peak fractions are shown in the inset Table.

RESULTS

MT isolation

Cyanobacterial MT was prepared from cultures exposed to either cadmium or zinc salts to induce synthesis of the protein and improve yields before harvesting (Olafson *et al.*, 1979). Fig. 1 shows the reverse-phase h.p.l.c. profile of a 10000- M_r -MT-rich fraction obtained by Sephadex G-75 chromatography of lysate from a Cd^{2+} -induced culture. The reverse-phase system was capable of resolving no fewer than seven species of MTs (fractions 3–9) at near isocratic conditions without attempting to fractionate an additional group that was eluted at higher organic modifier concentrations (fraction 10). Detection of MT-containing fractions relied on polarographic determination of thiol content along with metal detection, either as the γ -ray-emitting radionuclides by γ -radiation spectrometry or the non-

isotopic element by atomic absorption spectrometry. On this basis fractions 1 and 2 were shown to be contaminants free of high-thiol content metal-binding material. Peaks 3–9 were found to contain MT isoforms only, but peak 10 was an unresolved complex and contained MT and pigmented contaminants.

Metal analysis

In order to investigate the nature of the unexpectedly large number of isoforms, samples were taken across the elution profile in Fig. 1 and assayed for Zn, Cu and Cd by atomic absorption spectrophotometry and for MT by catalytic wave-differential pulse polarography (Olafson, 1981). Even though cultures were only exposed to a single metal on each occasion, other metals were present as nutrients in the growth medium and were therefore found in small amounts in the isolated protein (e.g. Cu and Zn were always present in Cd^{2+} -induced cultures). When metal values were normalized with MT values, Cd was observed to be present in the highest concentration, whereas Cu was present in only two fractions (peak 5 and a minor peak preceding peak 6 in Fig. 1) and Zn was not found in significant amounts in any of the high-thiol-content fractions. Thus the large number of chromatographically separable forms of MT could only partly be attributed to metal speciation and incomplete metal saturation.

Amino acid analysis

Amino acid analyses were performed on each of the high-thiol-content metal-containing peaks eluted from the propylsilane column, and the results of this study are reported in Table 1. These data were calculated from duplicate or triplicate determinations performed on unmodified metal-stripped protein, and only residues that differed from the composition of peak 3 in Table 1 are tabulated. Only single-residue changes were observed, and, since such small differences in composition were close to the error of this technique, it was not possible confidently to attribute the above chromatographic-mobility differences to microheterogeneity in amino acid

Table 1. Amino acid analysis of *Synechococcus* sp. Tx-20 MT isoforms

Amino acid residue	Isoform ...	Composition (mol of residue/mol)						
		3	4	5	6	7	8	9
Asx		5.7						
Thr		4.8						
Ser		3.4						
Glx		2.0						
Pro		1.9						
Gly		7.5						
Ala		4.0						
Cys		10.0						
Val		2.3		0.9				1.3
Met		0						
Ile		1.1						
Leu		2.9						
Tyr		2.3	2.8	1.3		1.1		3.0
Phe		0						
His		2.4						
Lys		2.6						
Arg		1.0						

Table 2. Automated sequence analysis of carboxymethylated *Synechococcus* sp. Tx-20 MT

Dehydrothreonine and dehydroserine are not included in yields for threonine and serine.

Edman cycle	Residue and yield (nmol)		Edman cycle	Residue and yield (nmol)	
	MT-1	MT-2		MT-1	MT-2
1	T 0.59	T 1.2	22	A 0.27	A 3.65
2	S 0.09	S 1.33	23	I 0.22	I 2.15
3	T 0.58	T 1.38	24	D 0.17	D 3.33
4	T 0.49	T 2.06	25	R 0.10	R 0.87
5	L 1.18	L 6.70	26	N 0.13	N 0.52
6	V 1.31	V 4.45	27	G 0.12	G 1.27
7	K 0.75	K 3.57	28	L 0.10	L 1.65
8	C 0.47	C 3.58	29	Y 0.13	Y 2.81
9	A 1.23	A 4.53	30	Y 0.15	Y 3.49
10	C 0.47	C 3.58	31	C 0.06	C 2.66
11	E 0.73	E 4.55	32	X	C 2.78
12	P 0.54	P 3.46	33	E 0.03	E 2.01
13	C 0.36	C 3.55	34	A 0.04	A 2.55
14	L 0.88	L 3.30	35	X	C 2.31
15	C 0.41	C 3.50	36	A 0.06	A 1.82
16	N 0.55	N 1.04	37	X	X
17	V 0.70	V 2.29	38	X	G 0.51
18	D 0.36	D 3.22	39	X	H 0.37
19	P 0.54	P 2.77	40	T 0.04	T 0.83
20	S 0.09	S 0.59	41		G 0.53
21	K 0.09	K 2.23	42		G 0.63

composition. Thus, since evidence for metal speciation and partial metal depletion was found, these data indicate that the variety of forms apparently reflects a complex mixture of metal speciation and disulphide aggregation products, rather than major differences in amino acid composition. Evidence for aggregation was obtained by showing the presence of higher- M_r species on gel-permeation chromatography of h.p.l.c. fractions (results not shown).

The above analyses also showed that, although similar in M_r , the prokaryotic MTs from cyanobacteria had much lower cysteine content in comparison with most eukaryotic proteins (19% compared with 32%) and therefore complex only half as much metal per mol of protein. This, together with the high content of hydrophobic residues relative to eukaryotic MTs, suggested that the structure of the cyanobacterial proteins might differ significantly from that of the eukaryotic proteins.

Amino acid sequence analysis

Since eukaryotic MTs were found to have highly conserved homologous amino acid sequences, it was decided to undertake the primary structural analysis of the *Synechococcus* sp. Tx-20 MT to assess any relatedness that might exist with the former group of metalloproteins. An *N*-terminal sequence analysis was performed on the first two high-thiol-content metal-binding species to be eluted from the propylsilane reverse-phase h.p.l.c. column (peaks 3 and 4, Fig. 1). These MTs were referred to as MT-1 and MT-2. Table 2 shows the automated sequencer results for these analyses together with the yield at each step. Verification of the *N*-terminal amino acid assignments was made on four additional automated sequence analyses.

In order to provide an overlapping *C*-terminal

sequence, *Staphylococcus* V8 proteinase was utilized to cleave at glutamic acid residues. Since only two glutamic acid residues were found in the composition and their positions were determined by the above *N*-terminal sequence analysis, a single overlapping peptide, along with the *N*-terminus, was predicted from the use of this proteinase. This was possible since the first glutamic acid residue at position 11 was adjacent to a proline residue, leaving the remaining glutamic acid residue at position 33 an ideal candidate for generation of a suitable peptide. Digests of the whole metal-stripped, reduced and pyridylethylated MT-1 and MT-2 were fractionated by reverse-phase h.p.l.c. with identical results for both isoforms (Fig. 2). Fraction 2 was subsequently found to have the appropriate amino acid composition for the *C*-terminal peptide after comparison of the *N*-terminal amino acid sequence with the amino acid composition for the intact molecule. Peak 1 contained unresolved peptides, and peaks 3 and 4 were non-peptide contaminants. The small shoulder on this peak fraction 2 was found to be residual 4-vinylpyridine or a related adduct that did not interfere with subsequent work. Peptides from both MT-1 and MT-2 digests were desalted by gel-permeation chromatography and subjected to automated sequence analysis. Table 3 shows the cumulative results of these degradations providing a three-residue overlap with the *C*-terminal segment of the previously sequenced *N*-terminal portion of both molecules. Ambiguities at positions 6 and 15 in the sequence of the *C*-terminal fragment of MT-1 were clearly shown to be histidine residues in MT-2, which also agrees with the previously described amino acid composition for both MT-1 and MT-2 (isoforms 3 and 4, Table 1). Carboxypeptidase Y digests carried out on [^{14}C]-carboxymethylated protein indicated a *C*-terminal

Fig. 2. Reverse-phase h.p.l.c. profile of *Staphylococcus*-V8-proteinase digest of pyridylethylated *Synechococcus* MT

Organic modifier (B) was acetonitrile.

cysteine residue by both amino acid analysis and reverse-phase h.p.l.c., in agreement with the above sequence determinations. Thus it was concluded that the metal-free *Synechococcus* MT-1 and MT-2 proteins had a minimum *M_r* of 5216 and identical primary structures. No evidence was found for the presence of a third tyrosine residue as suggested by earlier amino acid compositions of MT-2.

Amino acid sequence alignment analysis

The cyanobacterial sequence was then compared with a selection of eukaryotic sequences as is shown in Fig. 3.

Fig. 3. Computer alignment of *Synechococcus* MT (top sequence) with selected eukaryotic sequences

Colons indicate identities, stops (periods) indicate conservative replacements.

Sequences were aligned by using the algorithm of Lipman & Pearson (1985). The best alignment score obtainable, allowing for breaks where alignments were improved by 2 or more and allowing for conservative replacements, was 23 % corresponding residues. [Corresponding residues were defined as the sum of identical residues and conservative replacements (Haber & Koshland, 1970) where the following conservative replacements were allowed: A = G, S = T, N = Q, K = R, Y = F, I = L = V, E = D.] The statistical significance of the alignments with eukaryotic molecules was also estimated by using the RDF program of Lipman & Pearson (1985), which compares the query sequence with randomly permuted versions of the potentially related sequence, maintaining length and amino acid composition. Data are compiled after comparing 20 such

Table 3. Automated sequence analysis of pyridylethylated *Synechococcus* sp. Tx-20 MT C-terminus peptide from *Staphylococcus*-V8-proteinase digest

Edman cycle	Residue and yield (nmol)		Edman cycle	Residue and yield (nmol)	
	MT-1	MT-2		MT-1	MT-2
1	A 0.40	A 5.20	11	K 0.01	K 0.24
2	C 0.12	C 5.03	12	G 0.03	G 0.72
3	A 0.23	A 2.00	13	C 0.01	C 0.34
4	D 0.08	D 0.51	14	G 0.03	G 0.59
5	G 0.04	G 1.84	15	X	H 0.06
6	H/T	H 0.20	16	T 0.09	T 0.45
7	T 0.09	T 0.45	17	G 0.02	G 0.38
8	G 0.03	G 0.95	18	C 0.01	C 0.21
9	G 0.04	G 1.00	19	N 0.02	N 0.23
10	S 0.02	S 0.12	20	C 0.01	C 0.26

Table 4. Analysis for statistical significance of eukaryotic sequences aligned with the *Synechococcus* MT sequence

Protein compared	Mean aligned score randomized protein	Aligned score	z value
Sea-urchin MT	54.8 ± 8.0	65	1.27
Yeast MT	42.1 ± 5.4	38	-0.76
Human MT-2	53.2 ± 10.7	51	-0.20
Crab (<i>Scylla</i>) MT-2	50.6 ± 7.2	47	-0.50

randomly generated sequences with results expressed in terms of z values, where:

$$z = \frac{\text{similarity score} - \text{mean of random scores}}{\text{standard deviation of random scores}}$$

This value is then the number of standard deviations above the randomly generated mean score. Experience with such data-base comparisons suggests that z values less than 2 are unlikely to be significant (Lipman & Pearson, 1985). Table 4 summarizes the results obtained in these comparisons. Since all z values were below 2, the statistical significance of any apparent relatedness was very low. In view of the high content of cysteine in both eukaryotic and prokaryotic molecules it was necessary to reduce the significance of these homologies further because of the high probability of chance cysteine alignments. Thus no evidence for significant homology was observed on comparison with any of the eukaryotic sequences.

Other measurable parameters also showed a considerable divergence from the eukaryotic structures. Eukaryotic MTs show a pattern of frequent adjacency of conserved basic residues with cysteine (Lerch *et al.*, 1982). Although the ratio of cysteine to the sum of the lysine and arginine residues appeared similar on comparison with eukaryotic MTs (2.5 in *Synechococcus* and a mean of 2.3 ± 0.3 in the eukaryotes considered), there was only one basic residue found adjacent in sequence to a cysteine residue in the cyanobacterial MT whereas there are seven such adjacent basic residues in both *Scylla* and human MTs. In addition, the usual close association of hydroxylated residues with cysteine residues was not found in the prokaryote. No serine or threonine residues were found adjacent to cysteine residues in the cyanobacterial MT, whereas five are seen in yeast, six in *Scylla* and nine in human MT-2. It is of interest that both the cyanobacterium and yeast MTs have aromatic residues (phenylalanine in yeast and tyrosine in the cyanobacterium) as well as long-chain aliphatic residues. In addition, the prokaryotic MT has a cluster of hydroxylated residues at the *N*-terminus and a repeating Gly-His-Thr-Gly sequence near the *C*-terminus of these molecules. Finally, no evidence was found during investigation of any of the prokaryotic isoforms for an *N*-acetylmethionine residue or otherwise blocked *N*-terminus as is frequently observed in the eukaryotic MTs.

C.d.

The nature of the higher-order structure of the cyanobacterial MT was investigated by using c.d. and compared with predictive algorithms that utilize primary-

Fig. 4. Far-u.v. c.d. spectrum of native *Synechococcus* Zn-thionein in 0.05 M-sodium phosphate buffer, pH 7.0 (spectrum A) and apothionein in 0.05 M-HCl, pH 2.0 (spectrum B)

Protein concentration was 0.184 mg/ml in spectrum A and 0.155 mg/ml in spectrum B, and a 0.0503 cm-pathlength cell was used. Temperature was 25 °C.

structure information. The spectrum of the prokaryotic Zn-thionein in the far u.v. was not unlike that described for human Zn-thionein (Bühler & Kägi, 1979) or rabbit liver Zn-thionein (Pande *et al.*, 1986), although the absolute value of the negative ellipticity at the minimum was only about 30% of that of the mammalian proteins. The c.d. in this region (Fig. 4, spectrum A) showed a single negative maximum near 200 nm, which is observed with random-coil poly(amino acids) and fully denatured proteins (Fasman *et al.*, 1970). Minor positive bands occurred in the region 235–250 nm with a maximum near 238 nm, which compares well with human Zn-thionein reported to have a positive Zn-thionein–apo-thionein difference band at 239 nm (Bühler & Kägi, 1979). By using data from the 190–240 nm region and employing the program CONTIN (Provencher & Glöckner, 1981), estimates of the secondary structure were obtained as shown in Table 5. Metal-depleted cyanobacterial MT was studied at pH 2.0 and the results of this work shown in Fig. 4 (spectrum B). An increase in negative ellipticity of approx. 1000° was noted for the negative trough near 200 nm. In addition, a marked loss was observed in a positive band in the region near 238 nm relative to the native Zn-thionein. The latter finding compared well with results from a similar study of a human MT that showed loss of a 239 nm band after metal depletion. This phenomenon was attributed to disruption of the metal-

Table 5. Secondary-structure predictions from c.d. data by use of the program CONTIN (Provencher & Glöckner, 1981)

	Secondary structure (%)			
	α -Helix	β -Pleat	β -Turn	Remainder
Zn-thionein	0	66	27	6
Thionein (pH 2.0)	0	61	29	4
Zn-thionein + 1.97 mol equiv. of Cd^{2+}	0	67	31	2
Zn-thionein + 1.97 mol equiv. of Cu^{2+}	0	67	29	4

Fig. 5. Near-u.v. c.d. spectra for native Zn-thionein (—) and Zn-thionein titrated with 1.18 mol (-----) and 2.36 mol of (— · —) of Cu^{2+} /mol of protein

The inset in this Figure shows difference c.d. spectra obtained by subtraction of the Zn-thionein spectrum from those of the Cu derivatives after titration with 1.18 mol (—) and 2.36 mol (-----) of Cu^{2+} /mol of protein. Protein concentration was 0.92 mg/ml. The cell pathlength was 1 cm. The temperature was 25 °C.

thiolate clusters by the previous workers (Bühler & Kägi, 1979). The latter region of the spectrum was studied in better detail when the 350–230 nm c.d. region was investigated with use of a longer-pathlength cell. The native Zn-thionein spectrum in Fig. 5 shows the positive transition spoken of above at 239 nm and a negative maximum near 282 nm. Titration experiments with cadmium and copper salts were also conducted on the native Zn-thionein in this region of the u.v., with marked perturbations in the c.d. being observed with copper salts. Fig. 5 inset shows the difference spectrum on addition of 1.18 and 2.36 mol of CuCl_2 /mol of protein, indicating a shift of the 239 nm maximum to higher wavelengths near 248 nm. It may be noted that very similar spectra have been reported by Weser & Rüpp (1979) for native Cu-thioneins from eukaryotes.

Computer-predicted secondary structure

The secondary-structure predictions of the spectroscopic analysis were compared with those obtained by using the Chou & Fasman (1974) empirical analysis with the graphical results depicted in Fig. 6. The composition results are shown in Table 6. By utilizing a composite algorithm for hydrophilicity, accessibility and flexibility

Fig. 6. Graphical display of Chou & Fasman (1974) predicted β -structure (-----) and β -turns (—) as a function of amino acid sequence for cyanobacterial MT

The negligible amount of helix is depicted by the continuous line. The cut-off point for all parameters is 1 on the ordinate axis.

Table 6. Estimation of secondary structure by using the Chou & Fasman (1974) empirical relations

	Secondary structure (%)	
	α -Helix	β -Structure
	0	37.7
	β -Turn	41.5
	Remainder	20.8

(Parker *et al.*, 1986) it would normally be predicted that the β -structure suggested for the regions from residues 13–17 and 28–32 are buried, and that residues 1–10 are partially buried (Fig. 7). Since these regions encompass the hydrophobic sections of the primary structure, save the *N*-terminal region, this would be a reasonable prediction. However, like the better-known eukaryotic MTs, the bacterial MT is probably too small to have significant buried sequence. Therefore all that can reasonably be predicted from this analysis is that extended chains with chiral properties similar to β -structure are associated with the indicated sections of the primary structure. Similarly the β -turns predicted at residues 10–13 and 32–45 are likewise associated with more hydrophobic regions and could be interpreted as

Fig. 7. Composite plot of hydrophilicity, surface accessibility and polypeptide flexibility

The algorithm of Parker *et al.* (1986) was used.

being buried, if not for the relatively small size of the molecule.

DISCUSSION

This study describes the primary structure of a unique prokaryotic MT with superficial features similar to those of eukaryotic MTs. The molecule complexes Cd, Zn and Cu and has a high thiol content (19%) with cysteine-containing sequences similar to those observed in eukaryotic MT structures (Olafson, 1984). Earlier physiological studies with *Synechococcus* strain RRIMP showed that the synthesis of prokaryotic MT was induced in the presence of the former two metals, with induction controlled at the level of transcription, which is also in analogy with the eukaryotic MTs (Olafson *et al.*, 1980). Although the latter similarities initially suggested that the prokaryotic MT would be evolutionarily and structurally related to the MTs from higher organisms, the primary-structural analysis reported here does not support that assumption. Our findings indicate that the bacterial molecule is related only in that there is a high cysteine content with characteristic clusters of Cys-Xaa-Cys, Cys-Xaa-Xaa-Cys and Cys-Cys sequences. Otherwise these data are more generally in support of a convergent evolutionary relationship between molecules with an apparently similar function. This proposal is based on the fact that, although a computer-assisted sequence alignment analysis was performed, no statistically significant homology could be found. In addition to those proteins compared above, two other sequences were also compared in an attempt to find a related sequence. MT from the fungus *Agaricus bisporus* (Munger & Lerch, 1985) and the MT-like E_c protein from wheat germ (Hofmann *et al.*, 1984) were compared with the cyanobacterial molecule by using the FASTP program, but again no significant homologies were found. Thus no sequence homology could be found on comparison with MTs from four different phyla, supporting the above proposal that the cyanobacterial molecule may have a convergent evolutionary relationship with eukaryotic MTs.

A lack of homology within the eukaryotes has also been reported, with the determination of both yeast MT

(Butt *et al.*, 1984) and sea-urchin MT (Nemer *et al.*, 1985) by DNA sequencing. Aside from a few small similarities it is remarkable that both of these sequences show essentially no relationship to previously reported eukaryotic structures or to one another. Moreover, neither showed significant homology with the prokaryotic structure reported here, and no structural evidence could be found on examination of the bacterial protein for a common central segment as proposed for eukaryotic MTs (Nemer *et al.*, 1985). A minor similarity was initially thought to occur in that both the yeast and bacterial sequences were shown to have a relatively high hydrophobicity on comparison with eukaryotic structures. *Saccharomyces* MT has five hydrophobic residues and the *Synechococcus* MT has eight. However, it was noteworthy that all hydrophobic residues in the yeast sequence were concentrated in the *N*-terminus region, whereas in the prokaryotic structure the hydrophobic residues were distributed throughout the first half of the molecule, with a pair of adjacent tyrosine residues situated in the middle of the protein. Moreover, unlike the yeast MT, the first four *N*-terminal residues of the bacterial protein were hydroxylated side chains, conferring an unusually polar nature to this portion of the molecule. Subsequent studies have indicated that the original yeast DNA sequence did not represent the mature protein in that the *N*-terminal eight residues of the yeast molecule are lost during post-translational processing (Hamer, 1986). The difference in distribution of hydrophobic residues is therefore quite significant, leaving the prokaryotic molecule the most hydrophobic MT yet described.

Another difference of significance is the lack of adjacency of hydroxylated and basic residues with cysteine residues in the prokaryotic MT. Although the structural importance of these residues is not yet clear from tertiary-structure data, some evidence exists that ϵ -aminolysine groups hydrogen-bond to the metal-thiolate complexes (Pande *et al.*, 1985). Conservation of these amino acid residues in most eukaryotic sequences also lends strong support to the proposal that they play a fundamental role in the function of eukaryotic MTs (Lerch *et al.*, 1981). The complete lack of serine or threonine residues adjacent to cysteine residues and the presence of only a single adjacent basic residue again suggests that the prokaryotic molecule has arisen from a distinctly different primordial ancestor.

Although the primary-structure evidence is in support of a molecule that is quite dissimilar to the highly conserved eukaryotic MT structures, studies of the secondary structure, on the other hand, indicate that some structural similarities may exist in the metal-thiolate complex region. As with the well-characterized mammalian proteins, c.d. studies of the cyanobacterial molecule gave no evidence for the presence of helical structure but indicated that some 61% of the latter molecule was in β -pleated sheet, with 29% in the form of β -turns. A c.d. study on rat liver MT by Pande *et al.* (1986) has discussed the problems inherent in using the approach of Provencher & Glöckner (1981) to calculate secondary-structural elements in the case of the metal-bound protein. However, for the apo state, where no problems arise from metal-ligand charge-transfer transitions, they have computed some 55% disordered, 21% β -turn and 18% β -sheet with a negligibly small (6%) helical contribution. We suggest that the larger amount

of β -sheet and β -turns present in the structure of the cyanobacterial MT could provide the rationale for the reduced ellipticity near 200 nm compared with the mammalian protein, since pure forms of these structures exhibit strong positive ellipticities in this region (Kawai & Fasman, 1978; Compton & Johnston, 1986). X-ray-diffraction data from rat liver MT crystals have shown that each domain of this eukaryotic structure can be described as an antiparallel β -sandwich, allowing for interactions from cysteine to the metal-thiolate clusters to be made in place of normal main-chain hydrogen-bonding (Furey *et al.*, 1986). Indeed, the general features of both the native Zn-thionein and titrated Cu-thionein prokaryotic MT spectra, in both the near and far u.v., agreed well with chiroptical data from eukaryotic proteins (Bühler & Kägi, 1979; Weser & Rüpp, 1979), suggesting that the prokaryotic protein may have a metal-thiolate cluster structure similar to that found in the latter proteins but in a single domain. On applying the empirically determined values of Chou & Fasman (1974) to the established sequence, helical structure was once again not predicted, but the amount of β -structure was predicted to be considerably less than that determined by using the program CONTIN for c.d. data. The disagreement between the predictions of the two data sets may reflect the contribution of the metal-thiolate cluster that is not accounted for in either method of calculation.

This investigation was supported by a grant from the National Science and Engineering Research Council of Canada, the Medical Research Council of Canada and the Alberta Heritage Foundation for Medical Research. We are grateful to Sandra Keilland for amino acid analysis and automated sequence analysis, to Alison Malacko for excellent technical support in purification and chemical modification and to Kimio Oikawa for his skilful operation of the Jasco spectropolarimeter.

REFERENCES

- Black, S. D. & Coon, M. H. (1982) *Anal. Biochem.* **121**, 281–285
- Bühler, R. H. O. & Kägi, J. H. R. (1979) in *Metallothionein* (Kägi, J. H. R. & Nordberg, M., eds.), pp. 211–220, Birkhäuser, Basel
- Butt, T. R., Sternberg, E. J., Gorman, J. A., Clarke, P., Hamer, D., Rosenberg, M. & Crooke, S. T. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3332–3336
- Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13**, 222–245
- Compton, L. A. & Johnston, W. C., Jr. (1986) *Anal. Biochem.* **155**, 155–167
- Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627
- Dufton, P. Y. & Hider, R. C. (1977) *J. Mol. Biol.* **115**, 177–193
- Fasman, G. D., Hoving, H. & Timasheff, S. N. (1970) *Biochemistry* **9**, 3316–3323
- Furey, W. F., Robbins, A. H., Clancy, L. L., Winge, D. R., Wang, B. C. & Stout, C. D. (1986) *Science* **231**, 704–710
- Haber, J. E. & Koshland, D. E., Jr. (1970) *J. Mol. Biol.* **50**, 17–30
- Hamer, D. A. (1986) *Annu. Rev. Biochem.* **55**, 913–951
- Hermanson, M. A., Ericsson, L. H., Neurath, H. & Walsh, K. A. (1973) *Biochemistry* **12**, 3146–3147
- Hofmann, T., Kells, D. I. C. & Lane, B. G. (1984) *Can. J. Biochem. Cell Biol.* **62**, 908–913
- Hunkapiller, M. W. & Hood, L. E. (1983) *Methods Enzymol.* **91**, 486–495
- Janin, J. (1979) *Nature (London)* **277**, 491–492
- Kägi, J. H. R. & Nordberg, M. (1979a) in *Metallothionein* (Kägi, J. H. R. & Nordberg, M., eds.), pp. 51–55, Birkhäuser, Basel
- Kägi, J. H. R. & Nordberg, M. (1979b) in *Metallothioneins* (Kägi, J. H. R. & Nordberg, M., eds.), pp. 60–64, Birkhäuser, Basel
- Karin, M., Andersen, R. D., Slater, E., Smith, K. & Herschman, H. R. (1980) *Nature (London)* **286**, 295–297
- Karplus, P. A. & Schulez, B. E. (1985) *Naturwissenschaften* **72**, S. 212
- Kawai, K. & Fasman, G. (1978) *J. Am. Chem. Soc.* **100**, 3630–3632
- Leone, A., Pavlakis, G. N. & Hamer, D. H. (1985) *Cell* **40**, 301–309
- Lerch, K., Ammer, D. & Olafson, R. W. (1981) *J. Biol. Chem.* **257**, 2420–2426
- Lipman, D. J. & Pearson, W. R. (1985) *Science* **227**, 1435–1441
- Munger, K. & Lerch, K. (1985) *Biochemistry* **24**, 6751–6756
- Nemer, M., Wilkinson, D. F., Travaglini, E. C., Sternber, E. J. & Butt, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4992–4994
- Neuhaus, D., Wagner, G., Vasak, M., Kägi, J. H. R. & Wuthrich, K. (1985) *Eur. J. Biochem.* **151**, 257–264
- Olafson, R. W. (1980) *J. Biol. Chem.* **256**, 1263–1268
- Olafson, R. W. (1983) *J. Nutr.* **113**, 268–275
- Olafson, R. W. (1984) *Int. J. Pept. Protein Res.* **24**, 303–308
- Olafson, R. W. (1985) *Can. J. Biochem. Cell Biol.* **63**, 91–95
- Olafson, R. W. & Thompson, J. A. J. (1974) *Mar. Biol. (Berlin)* **28**, 83–86
- Olafson, R. W., Abel, K. & Sim, R. G. (1979) *Biochem. Biophys. Res. Commun.* **89**, 36–43
- Olafson, R. W., Loya, S. & Sim, R. G. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1495–1503
- Otvos, J. D. & Armitage, I. M. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7094–7098
- Otvos, J. D., Olafson, R. W. & Armitage, I. M. (1981) *J. Biol. Chem.* **257**, 2427–2431
- Pande, J., Vasak, M. & Kägi, J. H. R. (1985) *Biochemistry* **24**, 6717–6722
- Pande, J., Pande, C., Gilg, D., Vasak, M., Callender, R. & Kägi, J. H. R. (1986) *Biochemistry* **25**, 5526–5532
- Parker, J. M. R., Guo, D. & Hodges, R. S. (1986) *Biochemistry* **25**, 5425–5432
- Provencher, S. & Glöckner, J. (1981) *Biochemistry* **20**, 33–37
- Scrutton, M. C., Wu, C. W. & Goldthwait, D. A. (1971) *Proc. Natl. Acad. Sci. U.S.A.* **68**, 2497–2501
- Slater, J. P., Mildvan, A. S. & Loeb, L. A. (1971) *Biochem. Biophys. Res. Commun.* **44**, 37–43
- Stevens, C. L. R. & Myers, J. (1976) *J. Phys. Chem.* **80**, 350–362
- Vasak, M., Hawkes, G. E., Nicholson, J. K. & Sadler, P. J. (1984) *Biochemistry* **24**, 740–747
- Weser, U. & Rüpp, H. (1979) in *Metallothionein* (Kägi, J. H. R. & Nordberg, M., eds.), pp. 221–230, Birkhäuser, Basel
- Zelazowski, A. J. & Piotrowski, J. K. (1977) *Acta Biochim. Pol.* **24**, 97–103

Received 17 August 1987/21 October 1987; accepted 18 December 1987