Choline monooxygenase, an unusual iron-sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants: Prosthetic group characterization and cDNA cloning

(metallobiochemistry/osmoprotectant biosynthesis/Rieske center/salinity/*Spinacia oleracea* L.)

Bala Rathinasabapathi[*], Michael Burnet[*†], Brenda L. Russell[*], Douglas A. Gage[‡], Pao-Chi Liao[‡], Gordon J. Nye[§], Paul Scott[¶‖], John H. Golbeck[¶**], and Andrew D. Hanson[*††]

[*]Horticultural Sciences Department, University of Florida, Gainesville, FL 32611; [‡]Biochemistry Department, Michigan State University, East Lansing, MI 48824; [§]Agricultural Biotechnology, Ciba–Geigy Corporation, Research Triangle Park, NC 27709; and [¶]Department of Biochemistry, University of Nebraska, Lincoln, NE 68583

Communicated by Eric E. Conn, University of California, Davis, CA, January 28, 1997 (received for review December 5, 1996)

ABSTRACT Plants synthesize the osmoprotectant glycine betaine via the route choline → betaine aldehyde → glycine betaine. In spinach, the first step is catalyzed by choline monooxygenase (CMO), a ferredoxin-dependent stromal enzyme that has been hypothesized to be an oligomer of identical subunits and to be an Fe-S protein. Analysis by HPLC and matrix-assisted laser desorption ionization MS confirmed that native CMO contains only one type of subunit (M_r 42,864). Determination of acid-labile sulfur and nonheme iron demonstrated that there is one [2Fe-2S] cluster per subunit, and EPR spectral data indicated that this cluster is of the Rieske type—i.e., coordinated by two Cys and two His ligands. A full-length CMO cDNA (1,622 bp) was cloned from spinach using a probe generated by PCR amplification for which the primers were based on internal peptide sequences. The ORF encoded a 440-amino acid polypeptide that included a 60-residue transit peptide. The deduced amino acid sequence included two Cys-His pairs spaced 16 residues apart, a motif characteristic of Rieske-type Fe-S proteins. Larger regions that included this motif also showed some sequence similarity (≈40%) to Rieske-type proteins, particularly bacterial oxygenases. Otherwise there was very little similarity between CMO and proteins from plants or other organisms. RNA and immunoblot analyses showed that the expression of CMO in leaves increased several-fold during salinization. We conclude that CMO is a stress-inducible representative of a new class of plant oxygenases.

Organic solutes such as betaines, polyols, and proline are accumulated by many bacteria, animals and plants as an adaptation to salt or water stress (1, 2). These compounds, termed compatible osmolytes or osmoprotectants, are compatible with metabolism and can reduce ion toxicity (1). Their biosynthetic pathways were accordingly identified more than a decade ago as targets for metabolic engineering to improve stress tolerance. The potential of this approach has now been demonstrated by engineering the accumulation of glycine betaine in microorganisms (3, 4) and mannitol or proline in tobacco (5, 6).

Glycine betaine is a particularly widespread and effective osmoprotectant (1, 2, 7). In plants, it is synthesized by a two-step oxidation of choline: choline → betaine aldehyde → glycine betaine. In spinach, the first step is catalyzed by choline monooxygenase (CMO), a ferredoxin-dependent enzyme that

produces the hydrate form of betaine aldehyde (8). The second step is mediated by betaine aldehyde dehydrogenase (BADH) (8). Both CMO and BADH are located in the chloroplast stroma and increase in activity in response to salt stress (8). It proved to be straightforward to purify BADH and to use protein sequence data to isolate BADH cDNAs from various plants (see refs. 9 and 10). Some of these have now been expressed in transgenic tobacco (10, 11). In contrast, cDNA cloning of CMO has been refractory because the enzyme itself is most unusual.

CMO is in fact unique to plants. In other organisms, flavoprotein dehydrogenases or oxidases catalyze the choline oxidation reaction (3, 12–15). CMO may also be unique among plant oxygenases. Unlike many of these, it appears not to be a P450 enzyme: it is not membrane-bound, it is not sensitive to CO, and its optical spectrum is not that of a heme protein (16, 17). The features of this spectrum suggest rather that CMO could be an Fe-S protein (17). The quaternary structure of CMO is not clear. Gel filtration and SDS/PAGE analyses suggested that CMO might be a homodimer of subunit $M_r \approx$ 45,000, but did not rule out additional subunits (16, 17). These considerations led us to study the subunit composition and prosthetic group of CMO and to isolate and characterize CMO cDNAs.

MATERIALS AND METHODS

Protein Isolation and Analyses. Enzymatically active CMO was purified as described from leaves of spinach (*Spinacia oleracea* L. cv. Savoy Hybrid 612) plants that had been grown and salinized with 200 mM NaCl (17). CMO was assayed radiometrically (17). Nonheme iron (18) and acid-labile sulfur (19) were determined colorimetrically. Protein was assayed by the bicinchoninic acid method (20). The native M_r of CMO was estimated by the PAGE method of Hedrick and Smith (21) and by gel filtration chromatography as described (17), except that a Superose 12 column (Pharmacia) was placed in tandem ahead of the Superdex 200 column (Pharmacia) to increase resolution. The tandem column was calibrated using (M_r

standards (Sigma MW-GF-200) with the addition of transferrin (M_r 80,000).

HPLC and Mass Spectrometry. CMO isolated as above was fractionated using a microbore HPLC system equipped with a Reliasil C_{18} column (Michrom BioResources, Auburn, CA). Elution (50 μl·min^{-1}) was with a 5–95% gradient of acetonitrile in water, with 0.1% trifluoroacetic acid present throughout; the gradient was completed in 20 min. The eluate was monitored by absorbance at 214 nm. M_r values of peaks were determined by matrix-assisted laser desorption ionization MS (MALDI-MS) (22, 23). Spectra were acquired with a Voyager Elite time-of-flight mass spectrometer (PerSeptive Biosystems, Framingham, MA) operated in linear mode. The matrix was sinapinic acid. CMO mass determinations were internally calibrated using trypsinogen (M_r 23,979) and bovine serum albumin (M_r 66,431).

EPR Spectroscopy. Samples of purified CMO ($\approx$250 μg in 250 μl) were adjusted to pH 10 with 50 μl of 1 M glycine-NaOH and reduced by adding 1 mg Na dithionite. They were analyzed using a Bruker ECS 106 EPR x-band spectrometer with ER 4116 DM resonator and an Oxford liquid helium cryostat. Temperature was controlled by an Oxford Intelligent controller and monitored with a thermocouple 3 mm beneath the sample tube with liquid nitrogen as the reference. The microwave frequency was sampled by a Hewlett–Packard model 5340A frequency counter. Data manipulations were carried out using the program IGORPRO 2.04 (Wavemetrics, Lake Oswego, OR).

Peptide Microsequencing. Purified CMO was subjected to SDS/PAGE and the $M_r \approx 45,000$ band was blotted to polyvinylidene difluoride membrane. Tryptic peptides were generated by the procedure of Fernandez *et al.* (24) and purified by reverse-phase HPLC using an Aquapore RP-300 (C_8, 2.1 × 220 mm) column developed with a trifluoroacetic acid–acetonitrile gradient. Isolated peptides were subjected to sequence analysis on an Applied Biosystems model 476A protein/peptide sequencer (Perkin–Elmer). The N-terminal sequence of the intact protein was determined on a sample further purified by reverse-phase HPLC.

cDNA Cloning. Total RNA from salinized spinach leaves was extracted as described (25), except that a step to precipitate carbohydrates with 75 mM $BaCl_2$ was added prior to LiCl precipitation. Poly(A)$^+$ RNA was isolated using poly(U) Sephadex (26) and used to construct a cDNA library (9 × 10^6 plaque-forming units) in λUniZap XR (Stratagene). A 532-bp DNA fragment was generated by reverse transcription–PCR with primers based on CMO peptides; the (+) and (−) primers were, respectively, 5'-CCIGARCARAAYYTNGAYC-CIAARG-3' and 5'-CCATCATRTTYTCYTCDATRTART-ARTC-3'. The PCR (100 μl) contained 3 ng first-strand cDNA, 40 pmol of each primer, 200 μM each of all four dNTPs, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% (wt/vol) gelatin, and 2.5 units of AmpliTaq DNA polymerase (Perkin–Elmer) in 10 mM Tris·HCl (pH 8.3). Forty cycles of 0.5 min each at 94°C and 30°C and 1 min at 50°C were carried out. The 532-bp fragment was isolated from low-melting agarose (27), and labeled with [α-^{32}P]dCTP (>3 × 10^9 cpm/μg) by the random primer method. Library screening and *in vivo* excision were according to the supplier's protocols. Screening the amplified library ($\approx$2 × 10^5 plaques) with this probe yielded 18 clones of which the longest was 1,189 bp. The unamplified library ($\approx$2 × 10^5 plaques) was then screened with a 223-bp *Eco*RI fragment from the 5' region of this clone. Of eight positive clones, two (pRS3 and pRS5) were sequenced in both strands using the fluorescent chain terminating dideoxynucleotides method (28). They were identical except that pRS5 lacked 150 bp at the 5' end and had one base change in the 3' noncoding region. Sequences were analyzed with the GCG sequence analysis package (29).

DNA and RNA Blot Analyses. Genomic DNA was prepared from leaves as described (27). Total RNA was isolated from control and salinized leaves (30), denatured and subjected to electrophoresis in formaldehyde/1.2% agarose gels (27). RNA was quantified by the orcinol method (31). Blotting and hybridizations were performed using standard protocols (27). Molecular size markers were an RNA ladder (0.24–9.5 kb; GIBCO/BRL) for RNA blots, and *Hin*dIII-digested λ DNA
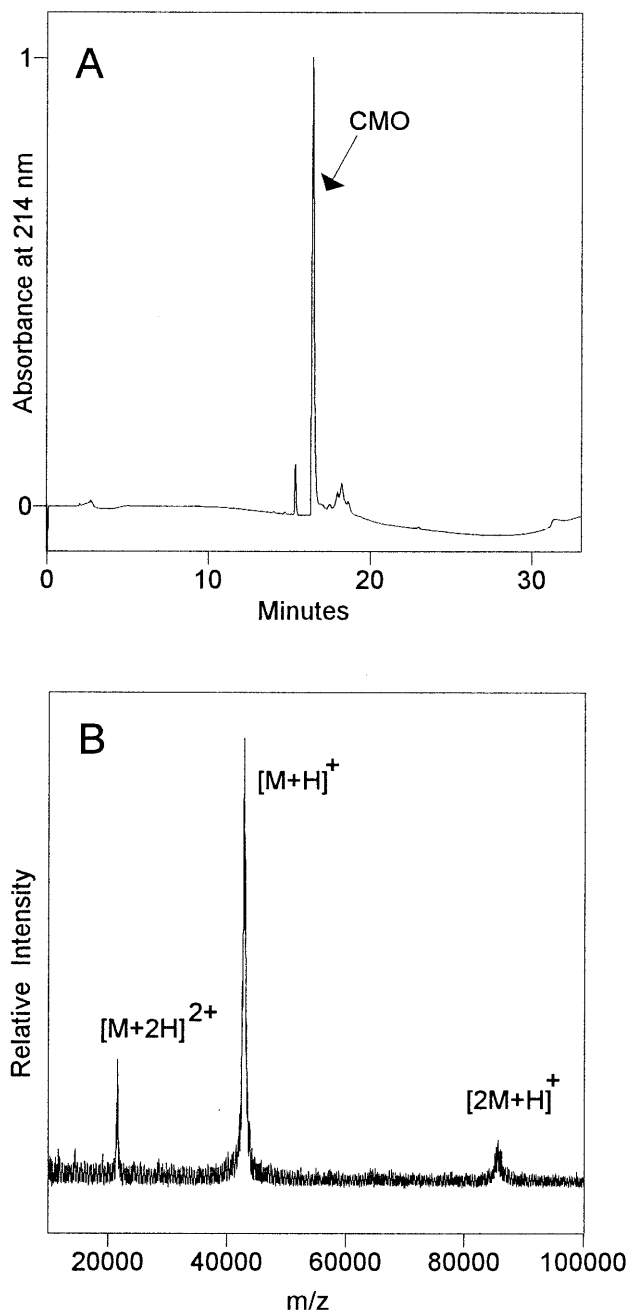

FIG. 1. Subunit composition of CMO and M_r of the monomer. (*A*) HPLC elution profile of a native CMO preparation (5.4 μg of protein). The cluster of small peaks eluting at $\approx$17–19 min contained no polypeptides detectable by MALDI-MS. Between analyses of the same preparation, the contaminant peak eluting at $\approx$15 min varied in size from <1% to 6% (shown here) of the CMO peak. (*B*) MALDI mass spectrum of the CMO peak from the HPLC separation above. The major signal corresponds to the singly protonated CMO monomer [M+H]$^+$; other peaks are the singly charged CMO dimer [2M+H]$^+$ and the doubly charged monomer [M+2H]$^{2+}$. The spectrum is unsmoothed to maximize the detection of any heterogeneity.

fragments for DNA blots. DNA blots were probed with an 1151-bp *Acc*I–*Eco*RV fragment of clone pRS3.

Antibody Production and Immunoblot Analysis. Rabbit antibodies were raised against CMO purified by SDS/PAGE. Mouse antibodies were raised against CMO purified by reverse-phase HPLC on a Delta-Pak C_{18} column (3.9×150 mm; Waters) using acetonitrile/water as the mobile phase. To determine the effect of salinization, spinach leaf proteins were precipitated with PEG 8000 (17), separated by SDS/PAGE, and transferred to nitrocellulose (32). Prestained M_r markers (Bio-Rad) were run simultaneously. Blots were probed with a 1:500 dilution of rabbit serum (32). CMO was detected after native PAGE using a 1:500 dilution of mouse ascites fluid.

RESULTS

Subunit Composition of CMO. In previous work, the purest CMO preparations that still had enzyme activity contained one major polypeptide ($M_r \approx 45,000$ by SDS/PAGE) and a few minor ones (17). To clarify whether the latter were contaminants or additional subunits, a highly purified, enzymatically active CMO preparation was fractionated by HPLC (Fig. 1*A*) and the CMO peak (identified by its immunoreactivity) and minor peaks were analyzed by MALDI-MS. The mass spectrum of the CMO peak showed a strong signal at m/z 42,864 $\pm$ 22 SD ($n = 13$ measurements); the symmetry of this signal suggested that just one polypeptide was present (Fig. 1*B*). The small peaks eluting at $\approx$17–19 min gave no MALDI-MS signals, suggesting that they were small-molecule contaminants. The peak eluting at $\approx$15 min gave a signal at m/z 11,698. However, the size of this peak relative to the CMO peak varied greatly between analyses of the same preparation (Fig. 1), making it unlikely to be a component of the CMO holoenzyme. Taken together, these data indicate that CMO has only one type of subunit. From this it follows that one gene is sufficient to encode an active CMO enzyme.

M_r of Native CMO. The M_r estimated from native PAGE for the CMO band was 135,000 (Fig. 2). A very similar value (M_r 133,000) was obtained from gel filtration on a tandem Superose 12–Superdex 200 column (not shown). These values are higher than the earlier estimate of M_r 98,000 from gel filtration on Ultrogel AcA 34 (16) and indicate that CMO may exist as a homotrimer rather than a homodimer. However, molecular
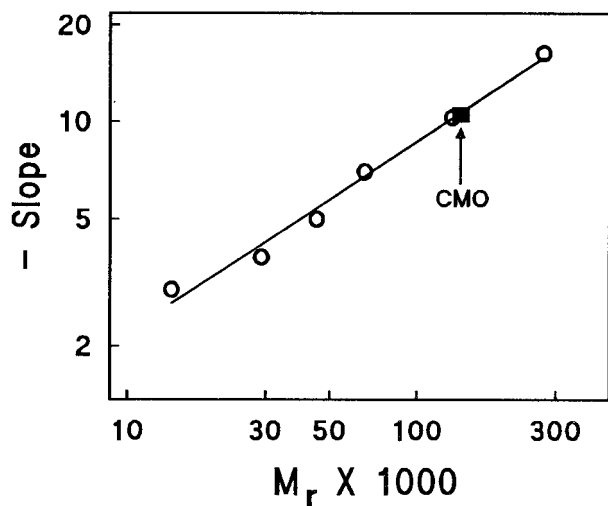
Fig. 2. Estimation of the M_r of native CMO by nondenaturing PAGE. The standard proteins were α-lactalbumin (M_r 14,200), carbonic anhydrase (M_r 29,000), chicken egg albumin (M_r 45,000), bovine serum albumin monomer (M_r 66,000), dimer (M_r 132,000), and jack bean urease trimer (M_r 272,000). The coefficient of linear correlation for the regression line shown was 0.994. The interpolated M_r of CMO was 135,000.
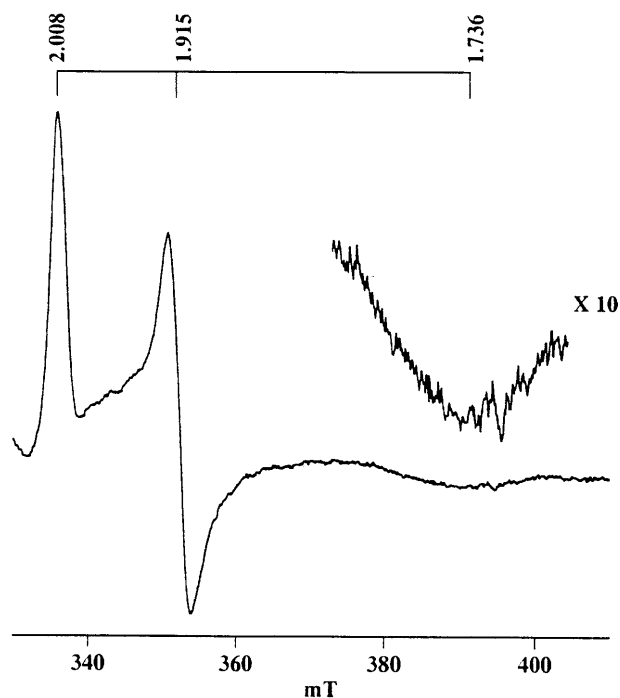
Fig. 3. EPR spectrum of CMO. Spectra of CMO reduced by sodium dithionite were acquired at 15 K with a microwave power of 20 mW and a modulation amplitude of 10 G. Apparent g values are indicated. The spectrum shown is the average of 16 measurements.

shape and interactions with the media affect the mobility of proteins in gel electrophoresis and gel filtration chromatography so that none of these native M_r values is precise enough to distinguish with certainty between a trimer and a dimer.

Evidence for a Rieske-Type [2Fe-2S] Center. Definitive evidence for an Fe-S center was sought by EPR spectroscopy and by chemical assays of acid-labile sulfide and nonheme iron. In EPR studies, upon reduction by dithionite, a rhombic spectrum with apparent g values of 2.008, 1.915, and 1.736 was observed (Fig. 3). The narrow low field peak and broad high field trough resemble those reported for certain types of [2Fe-2S] cluster (33). The g_{avg} (sum of g values/3) of this spectrum was 1.89, similar to that of many 2 His-, 2 Cys-liganded Rieske-type [2Fe-2S] clusters (33, 34). By contrast, [2Fe-2S] clusters liganded by 4 Cys residues typically have g_{avg} = 1.94 (33, 34). The spectrum for CMO reached maximum intensity at 15 K, somewhat lower than is typical of Rieske-type clusters, but still within the expected range. Consistent with these data, CMO was found to contain approximately 2Fe and 2S per subunit (Table 1).

cDNA Cloning. Amino acid sequences were obtained for the N terminus of CMO and for 12 tryptic peptides. Two internal sequences were used to design primers for reverse transcription–PCR, which gave a 532-bp DNA. Screening a library with this fragment yielded several truncated CMO cDNAs; the 5′ region from the longest of these was then used as a probe to

Table 1. Nonheme iron and acid-labile sulfur content of CMO

Analyte	nmol·ml^{-1}
CMO subunit	57 (0.3)
FE	105 (5.0)
S	99 (6.0)

Enzymatically active CMO was prepared as described (17). The reverse-phase HPLC elution profile (absorbance at 280 nm) of the preparation indicated that CMO was 51% of the total protein. This value, together with a M_r of 42,864 was used to calculate the molar concentration of CMO subunit. Values are means and SE (in parentheses) for three or four determinations.

```
TAGTAAGGGTGTAGCTAATTAGCAAAATAAACAAAAAGGAAGTGTTGAGTTGGTTAATGA    60
                                               M  M                2
TGGCAGCAAGCGCAAGCGCAACCACAATGTTGCTAAAATACCCAACTACAGTTTGTGGTA   120
 A  A  S  A  S  A  T  T  M  L  L  K  Y  P  T  T  V  C  G  I     22
TTCCAAATCCTTCATCAAACAATAATAATGATCCTTCAAACAATATAGTTTCTATTCCAC   180
 P  N  P  S  S  N  N  N  N  D  P  S  N  N  I  V  S  I  P  Q     42
AAAATACTACTAATCCAACACTTAAGTCCCGTACACCTAATAAAATCACCACCAACGCCG   240
 N  T  T  N  P  T  L  K  S  R  T  P  N  K  I  T  T  N  A* V     62
TCGCGGCACCGTCCTTCCTTCTTAACCACCACTACACCGTCGTCCATCCAATCACTTG    300
 A  A  P  S  F  P  S  L  T  T  T  T  P  S  S  I  Q  S  L  V     82
TCCACGAATTCGACCCTCAAATTCCCCCTGAAGACGCTCATCACCTCCTAGCTCTTGGT   360
 H  E  F  D  P  Q  I  P  P  E  D  A  H  T  P  P  S  S  W  V    102
ATACCGAACCTGCCTTCTATTCCCATGAACTTGAGCGTATCTTTTATAAAGGATGGCAAG   420
 T  E  P  A  F  Y  S  H  E  L  E  R  I  F  Y  K  G  W  Q  V    122
TTGCAGGGATCAGCGATCAAATAAAAGAGCCTAACCAATATTTCACTGGCAGCTTAGGAA   480
 A  G  I  S  D  Q  I  K  E  P  N  Q  Y  F  T  G  S  L  G  N    142
ATGTTGAATATTTGGTGTCTCGAGATGGTGAAGGGAAAGTTCATGCATTTCACAATGTTT   540
 V  E  Y  L  V  S  R  D  G  E  G  K  V  H  A  F  H  N  V  C†   162
GCACCCATCGTCGTCATCTATTCTTGCTTGCGGTAGTGGCAAAAAGTCGTGTTTCGTGTGCC   600
 T  H† R  A  S  I  L  A  C  G  S  G  K  K  S  C  F  V  C† P    182
CTTACCATGGATGGGTATATGGCATGGACGGATCACTTGCGAAAGCCTCCAAAGCAAAAC   660
 Y  H† G  W  V  Y  G  M  D  G  S  L  A  K  A  S  K  A  K  P    202
CTGAACAAAACTTGGATCCTAAAGAACTTGGGCTTGTACCCCTAAAAGTTGCAGTAGGG    720
 E  Q  N  L  D  P  K  E  L  G  L  V  P  L  K  V  A  V  W  G    222
GGCGGTTCTTATCAGCTTGGACAGATCACTTGAAGAAGGTGGTGATGTTGGAACTG    780
 P  F  V  L  I  S  L  D  R  S  L  E  E  G  G  D  V  G  T  E    242
AGTGGCTTGGTACTTCTGCTGAAGATGTTAAGGCCCATGCTTTTGATCCTTCACTTCAAT   840
 W  L  G  T  S  A  E  D  V  K  A  H  A  F  D  P  S  L  Q  F    262
TCATTCACGAAGTGAATTCCCAATGGAATCTAATTGGAAGATTTTCAGTGACAACTACT   900
 I  H  S  E  F  P  M  E  S  N  W  K  I  F  S  D  N  Y  L    282
TGGATAGCTCATATCATGTTCCTTATGCACACAAATACTATGCAACTGAACTCAACTTTG   960
 D  S  S  Y  H  V  P  Y  A  H  K  Y  Y  A  T  E  L  N  F  D   302
ACACTTACGATACCCAAATGATCGAAAACGTTACAATTCAAAGAGTGGAAGGAAGTTCAA  1020
 T  Y  D  T  Q  M  I  E  N  V  T  I  Q  R  V  E  G  S  S  Q   322
ACAAGCCTGATGGTTTTGATAGAGTTGGAATTCAAGCATTCTATGCTTTCGCGTATCCAA  1080
 K  P  D  G  F  D  R  V  G  I  Q  A  F  Y  A  F  A  Y  P  N   342
ATTTCGCTGTGGAAAGGTATGGCCCTTGGATGACTACAATGCATATTCACCCATTAGGAC  1140
 F  A  V  E  R  Y  G  P  W  M  T  T  M  H  I  H  P  L  G  P   362
CAAGGAAATGCAAACTAGTGGTGGACTATTATATTGAAAATTCTATGTTGGATGACAAGG  1200
 R  K  C  K  L  V  V  D  Y  Y  I  E  N  S  M  L  D  D  K  D   382
ATTACATCGAGAAGGCATAGCAATCAATGATAACGTACAGAGGGAAGATGTGGTGCTGT  1260
 Y  I  E  K  G  I  A  I  N  D  N  V  Q  R  E  D  V  V  L  C   402
GTGAAAGTGTACAAAGAGGTTTGGAGACACCAGCATATCGTAGTGGGAGATATGTGATGC  1320
 E  S  V  Q  R  G  L  E  T  P  A  Y  R  S  G  R  Y  V  M  P   422
CTATTGAGAAAGGAATCCACCATTTCCACTGCTGGTTGCAACAAACTTTGAAGTAATGT  1380
 I  E  K  G  I  H  H  F  H  C  W  L  O  O  T  L  K          439
TGACTTGACTTCACCGGACGTTTGCGTGCTCCTCTTGGCTCTTGCAGTTAGTTATGTGTG  1440
TATGCTCATGGCCTAAACTTATAACATTTTATACTCAATTTATAATAAACACCATAGTAA  1500
GTACCTCAGATATCCCGTGCATTTCATTTTCAGGGGAAATAATGACATTGTCTATCAAT  1560
ATCAATATCAATATCAATATCAATATTACCAGTAATTTAAACAAAAAAAAAAAAAAAAAA  1620
AA                                                            1622
```

FIG. 4. Nucleotide and deduced amino acid sequences of CMO cDNA clone pRS3. The amino acid sequences determined for tryptic peptides are underlined; overlaps between peptide sequences are underlined twice. The N terminus of the processed polypeptide is indicated with an asterisk. The Cys-His pairs conserved in Rieske-type iron-sulfur proteins are marked with daggers. The stop codon and putative polyadenylylation signal are boxed.

isolate a full-length cDNA (Fig. 4). This cDNA (1622 bp) had 5′ and 3′ noncoding regions of 56 and 246 bp, respectively. A putative polyadenylylation signal (AAATAAT) preceded the poly(A) sequence by 58 bp. The ORF (1320 bp) encoded 440 amino acids that included a 60-residue transit peptide. As the ORF begins with two adjacent ATG codons, either could be the translational start. However, the sequences flanking the second ATG match the consensus translational initiation motif in plants (35). The coding region included all the amino acid sequences determined for purified CMO. The size and composition of the deduced transit peptide were typical for a chloroplast stromal targeting peptide (36), consistent with the stromal location of CMO (16). The predicted M_r of the processed polypeptide was 42,884. As this value differs from that obtained by MALDI-MS (M_r 42,864) by less than the experimental error of the method ($\approx 0.1\%$) the CMO polypeptide can have very few if any posttranslational modifications.

Primary Structure Comparisons. No sequence in the data bases had significant homology with the entire CMO sequence so that no oxygenase of this kind is known to date. Rieske-type iron-sulfur proteins share a consensus sequence Cys-Xaa-His (15–17 amino acids) Cys-Xaa-Xaa-His, where Xaa = any amino acid. This motif, which is considered to be involved in binding the [2Fe-2S] cluster (34), was conserved in CMO (Fig. 4). This finding strongly supports the chemical and EPR data indicating that CMO has a [2Fe-2S] center. Between the Cys-His pairs, the CMO sequence had more residues in common with bacterial aromatic-ring dioxygenases than with other Rieske-type iron-sulfur proteins (37). Consistent with this, homologies in the region around the consensus sequence were somewhat stronger with these dioxygenases than with bacterial alkyl-group hydroxylases or Rieske iron-sulfur proteins of mitochondria and chloroplasts. Representative data for each of these families (38) are shown in Table 2.

CMO Induction by Salinity. In unsalinized plants, CMO mRNA levels were low in both expanding and expanded leaves. Salinization increased these levels, especially in expanded leaves (Fig. 5A). CMO protein levels (Fig. 5B) and enzyme activity (Fig. 5C) broadly paralleled these changes. Most of the rise in CMO activity in salinized plants can therefore be ascribed to increased gene expression rather than to enzyme activation. The magnitude of the salt induction of CMO mRNA is comparable to that reported for BADH (8).

Analysis of Genomic DNA. Following digestion of spinach genomic DNA with *Hin*dIII, *Eco*RV, or *Eco*RI, blot analysis revealed single bands of about 18, 9, and 3.7 kb, respectively (data not shown). This is consistent with there being one CMO gene containing a large intron(s). Reconstruction experiments also suggested a single copy of CMO per haploid genome (not shown).

DISCUSSION

Our results indicate that CMO is an oligomer of identical subunits, each of which has a Rieske-type [2Fe-2S] cluster— i.e., one that is liganded by two Cys and two His residues. The presence of this cluster, and the amino acid sequence of the protein, place CMO in a new class of plant oxygenases. Although without known relatives among plant oxygenases, CMO appears to be distantly related in structure to various bacterial oxygenases that, like CMO, require ferredoxin (or a ferredoxin-like domain in a larger protein) as electron donor (34). It is therefore reasonable to speculate that CMO is, in common with other nuclear-encoded chloroplastic enzymes (39), of prokaryotic origin.

The cDNA cloning of CMO and the finding that salt stress enhances CMO expression have implications for agriculture. As stated in the Introduction, it has been thought for some time that engineering the synthesis and accumulation of glycine betaine into crops naturally lacking it could improve their stress tolerance (40, 41). Although bacterial choline oxidases (3) or dehydrogenases (42) are being explored for this purpose, use of CMO (in conjunction with BADH) may be preferable

Table 2. Amino acid sequence homology between CMO and other proteins with Rieske-type iron-sulfur centers

Sequence	Species	Size of region compared residues	Identity, %	Similarity, %
Naphthalene dioxygenase	*Pseudomonas putida*	195	29.2	43.0
Vanillate demethylase (hydroxylase component)	*Pseudomonas* sp.	58	31.0	43.0
Mitochondrial Rieske Fe-S protein	*Zea mays*	72	29.2	36.1

The data shown are for representative members of the bacterial aromatic-ring dioxygenase and alkyl-group hydroxylase families (8) and of the Rieske protein family. The regions of homology that were compared included the conserved [2Fe-2S] cluster binding motif. Local homologies were first identified using BLASTP with nonredundant sequences in the National Center for Biotechnology Information (NCBI) database and individual entries were compared by FASTA.
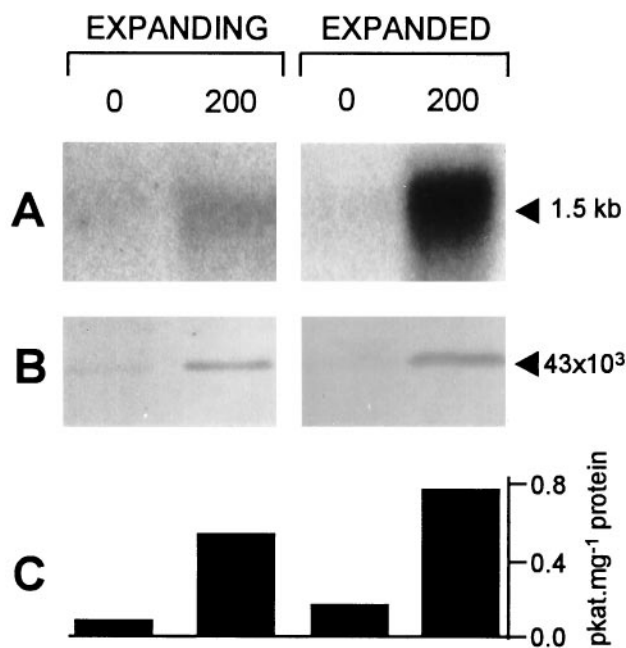
FIG. 5. CMO expression in expanding and expanded spinach leaves. Plants had been irrigated with nutrient solution (lane 0) or for 10 days before the experiment with nutrient solution containing 200 mM NaCl (lane 200). (*A*) RNA blot analysis. Lanes contained 5 μg of total RNA. The probe was a 532-bp DNA fragment [positions 660–1,191 of pRS3 (Fig. 4)]. Ethidium bromide staining demonstrated that all lanes contained equivalent amounts of RNA. Densitometry of autoradiographs indicated that salinization increased CMO mRNA levels 2-fold in expanding leaves and 7-fold in expanded leaves. (*B*) Immunoblot analysis. Lanes contained 40 μg of total leaf protein. Rabbit antibodies against SDS-denatured CMO were used for immunodetection. (*C*) Enzyme activity. CMO was assayed following precipitation of leaf proteins with PEG 8000 (17). Bars are means of three determinations; SE values were $\leq$16% of the means.

for two reasons. First, the reduced ferredoxin requirement of CMO links glycine betaine synthesis with the light reactions of photosynthesis. This could help to match the supply of glycine betaine with the demand for osmotic adjustment and osmoprotection, which climbs rapidly after sunrise as the water potential and water content of salt- or drought-stressed leaves start falling (43). Second, the spinach CMO gene, like the BADH gene (8), presumably has stress-responsive cis regulatory elements. These may be essential to reproduce the natural pattern of stress-induced glycine betaine accumulation in engineered crops. It is now possible to isolate CMO genes and to begin identifying these elements.

1. Yancey, P. H. (1994) in *Cellular and Molecular Physiology of Cell Volume Regulation,* ed. Strange, K. (CRC, Boca Raton, FL), pp. 81–109.
2. Gorham, J. (1995) in *Amino acids and Their Derivatives in Higher Plants,* ed. Wallsgrove, R. M. (Cambridge Univ. Press, Cambridge, U.K.), pp. 173–203.
3. Rozwadowski, K. L., Khachatourians, G. G. & Selvaraj, G. (1991) *J. Bacteriol.* **173,** 472–478.
4. Nomura, M., Ishitani, M., Takabe, T., Rai, A. K. & Takabe, T. (1995) *Plant Physiol.* **107,** 703–708.
5. Tarczynski, M. C., Jensen, R. G. & Bohnert, H. J. (1993) *Science* **259,** 508–510.
6. Kishor, P. B. K., Hong, Z., Miao, G., Hu, C. & Verma, D. P. S. (1995) *Plant Physiol.* **108,** 1387–1394.
7. Papageorgiou, G. C. & Murata, N. (1995) *Photosynth. Res.* **44,** 243–252.
8. Rhodes, D. & Hanson, A. D. (1993) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **44,** 357–384.
9. McCue, K. F. & Hanson, A. D. (1992) *Plant Mol. Biol.* **18,** 1–11.
10. Ishitani, M., Nakamura, T., Han, S. Y. & Takabe, T. (1995) *Plant Mol. Biol.* **27,** 307–315.
11. Rathinasabapathi, B., McCue, K. F., Gage, D. A. & Hanson, A. D. (1994) *Planta* **193,** 155–162.
12. Landfald, B. & Strom, A. R. (1986) *J. Bacteriol.* **165,** 849–855.
13. Nagasawa, T., Mori, N., Tani, Y. & Ogata, K. (1976) *Agric. Biol. Chem.* **40,** 2077–2084.
14. Tsuge, H., Nakano, Y., Onishi, H., Futamura, Y. & Ohashi, K. (1980) *Biochim. Biophys. Acta* **614,** 274–284.
15. Yamada, H., Mori, N. & Tani, Y. (1979) *Agric. Biol. Chem.* **43,** 2173–2177.
16. Brouquisse, R., Weigel, P., Rhodes, D., Yocum, C. F. & Hanson, A. D. (1989) *Plant Physiol.* **90,** 322–329.
17. Burnet, M., Lafontaine, P. J. & Hanson, A. D. (1995) *Plant Physiol.* **108,** 581–588.
18. Atkin, C. L., Thelander, L. & Reichard, P. (1973) *J. Biol. Chem.* **248,** 7464–7472.
19. Beinert, H. (1983) *Anal. Biochem.* **131,** 373–378.
20. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985) *Anal. Biochem.* **150,** 76–85.
21. Hedrick, J. L. & Smith, A. J. (1968) *Arch. Biochem. Biophys.* **126,** 155–164.
22. Suizdak, G. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11290–11297.
23. Geisow, N. J. (1992) *Trends Biotechnol.* **10,** 432–441.
24. Fernandez, J., Andrews, L. & Mische, S. M. (1994) in *Techniques in Protein Chemistry,* ed. Crabb, J. W. (Academic, New York), 5th Ed., pp. 215–222.
25. Hall, T. C., Ma, Y., Buchbinder, B. U., Pyne, J. W., Sun, S. M. & Bliss, F. A. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 3196–3200.
26. Hondred, D., Wadle, D-W., Titus, D. E. & Becker, W. M. (1987) *Plant Mol. Biol.* **9,** 259–275.
27. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1995) *Current Protocols in Molecular Biology* (Wiley, New York), Vols. 1–3.
28. Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., Cocuzza, A. J., Jensen, M. A. & Baumeister, K. (1987) *Science* **238,** 336–341.
29. Devereux, J. R. (1994) *GCG Sequence Analysis Software Package* (Genetics Computer Group, Madison, WI), Version 8.0.
30. Puissant, C. & Houdebine, L.-M. (1990) *BioTechniques* **8,** 148–149.
31. Dawson, R. M. C., Elliott, D. C., Elliott, W. H. & Jones, K. M. (1986) *Data for Biochemical Research* (Clarendon, Oxford), 3rd Ed., pp. 543–544.
32. Tokuhisa, J. G., Daniels, S. M. & Quail, P. H. (1985) *Planta* **164,** 321–332.
33. Johnson, M. K. (1994) in *Encyclopedia of Inorganic Chemistry*, ed. King, R. B. (Wiley, New York), Vol. 4, pp. 1896–1915.
34. Mason, J. R. & Cammack, R. (1992) *Annu. Rev. Microbiol.* **46,** 277–305.
35. Joshi, C. P. (1987) *Nucleic Acids Res.* **15,** 6643–6653.
36. Cline, K. & Henry, R. (1996) *Annu. Rev. Cell. Dev. Biol.* **12,** 1–26.
37. Neidle, E. L., Hartnett, C., Ornston, L. N., Bairoch, A., Rekik, M. & Harayama, S. (1991) *J. Bacteriol.* **173,** 5385–5395.
38. Harayama, S., Kok, M. & Neidle, E. L. (1992) *Annu. Rev. Microbiol.* **46,** 565–601.
39. Brinkmann, H. & Martin, W. (1995) *Plant Mol. Biol.* **30,** 65–75.
40. LeRudulier, D., Strom, A. R., Dandekar, A. M., Smith, L. T. & Valentine, R. C. (1984) *Science* **224,** 1064–1068.
41. McCue, K. F. & Hanson, A. D. (1990) *Trends Biotechnol.* **8,** 358–362.
42. Lilius, G., Holmberg, N. & Bulow, L. (1996) *Bio/Technology* **14,** 177–180.
43. Hanson, A. D. & Hitz, W. D. (1982) *Annu. Rev. Plant Physiol.* **33,** 163–203.