James, P., Inui, M., Tada, M., Chiesi, M., & Carafoli, E. (1989) *Nature 342*, 90–92.
Kataoka, M., Head, J. F., Seaton, B. A., & Engelman, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6944–6948.
Kuchinka, E., & Seelig, J. (1989) *Biochemistry 28*, 4216–4221.
Lafleur, M., Faucon, J.-F., Dufourcq, J., & Pézolet, M. (1989) *Biochim. Biophys. Acta 980*, 85–92.
Laggner, P., & Barratt, M. D. (1975) *Arch. Biochem. Biophys. 170*, 92–101.
Lewis, S. M., & Thomas, D. D. (1986) *Biochemistry 25*, 4615–4621.
Malencik, D. A., & Anderson, S. R. (1985) *Biochem. Biophys. Res. Commun. 130*, 22–29.
Malencik, D. A., & Anderson, S. R. (1988) *Bicohemistry 27*, 1941–1949.
Marsh, D. (1981) in *Membrane Spectroscopy* (Grell, E., Ed.) Vol. 31, pp 51–142, Springer-Verlag, Berlin.
Maurer, T., Lücke, C., & Rüterjans, H. (1991) *Eur. J. Biochem. 196*, 135–141.
O'Brian, C. A., & Ward, N. E. (1989) *Mol. Pharmacol. 36*, 355–359.
Popp, C. A., & Hyde, J. S. (1981) *J. Magn. Reson. 43*, 249–258.
Prendergast, F. G., Lu, J., Wei, G. J., & Bloomfield, V. A. (1982) *Biochemistry 21*, 6963–6970.
Raynor, R. L., Zheng, B., & Kuo, J. F. (1991) *J. Biol. Chem. 266*, 2753–2758.
Saffman, P. J., & Delbrück, M. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3111–3113.
Schulze, J., Mischeck, U., Wigand, S., & Galla, H.-J. (1987) *Biochim. Biophys. Acta 901*, 101–111.
Schwarz, G., & Beschiaschvili, G. (1989) *Biochim. Biophys. Acta 979*, 82–90.
Shinitzky, M. (1984) in *Physiology of Membrane Fluidity* (Shinitzky, M., Ed.) Vol. 1, pp 1–51, CRC Press, Inc., Boca Raton, FL.
Squier, T. C., & Thomas, D. D. (1986a) *Biophys. J. 49*, 921–935.
Squier, T. C., & Thomas, D. D. (1986b) *Biophys. J. 49*, 937–942.
Squier, T. C., & Thomas, D. D. (1988) *J. Biol. Chem. 263*, 9171–9177.
Squier, T. C., & Thomas, D. D. (1989) *Biophys. J. 56*, 735–748.
Squier, T. C., Hughes, S. E., & Thomas, D. D. (1988a) *J. Biol. Chem. 263*, 9162–9170.
Squier, T. C., Bigelow, D. J., & Thomas, D. D. (1988b) *J. Biol. Chem. 263*, 9178–9186.
Squier, T. C., Lakowicz, J. R., Mahaney, J. E., Yin, J.-J., and Lai, C.-S. (1991) *Biophys. J. 59*, 654–669.
Tada, M., & Katz, A. M. (1982) *Annu. Rev. Physiol. 44*, 401–423.
Thomas, D. D. (1986) in *Techniques for the Analysis of Membrane Proteins* (Ragan, C. I., & Cherry, R. J., Eds.) pp 377–431, Chapman and Hall, London.
Voss, J., Birmachu, W., Hussey, D., & Thomas, D. D. (1991) *Biochemistry* (in press).
Wille, B. (1989) *Anal. Biochem. 178*, 118–120.
Yeagle, P. L. (1985) *Biochim. Biophys. Acta 822*, 267–287.

Amino Acid Sequence of Nitrite Reductase: A Copper Protein from *Achromobacter cycloclastes*†

Faith F. Fenderson,‡ Santosh Kumar,§ Elinor T. Adman,*,‡ M.-Y. Liu,∥ W. J. Payne,⊥ and Jean LeGall∥
Departments of Biological Structure and Biochemistry, University of Washington, Seattle, Washington 98195, and Departments of Biochemistry and Microbiology, University of Georgia, Athens, Georgia 30406
Received January 23, 1991; Revised Manuscript Received March 13, 1991

ABSTRACT: The amino acid sequence of the copper-containing nitrite reductase (EC 1.7.99.3) from *Achromobacter cycloclastes* strain IAM 1013 has been determined by using peptides derived from digestion with *Achromobacter* protease I (Lys), *Staphylococcus aureus* V8 protease (Glu), cyanogen bromide, and BNPS-skatole in acetic acid. The subunit contains 340 amino acids. The identity of the first seven amino acids is tentative. The sequence has been instrumental in the X-ray structure determination of this molecule; in conjunction with the X-ray structure, ligands to a type I copper atom and a type II copper atom (one of each per subunit) have been identified. Comparison of the sequence to those of multi-copper oxidases such as ascorbate oxidase, laccase, and ceruloplasmin [Messerschmidt, A., & Huber, R. (1990) *Eur. J. Biochem. 187*, 341–352] reveals that each of two domains seen in the X-ray structure is similar to the oxidases and also to the small blue copper-containing proteins such as plastocyanin. The *combination* of sequence and structural similarity to ascorbate oxidase and sequence similarity to ceruloplasmin leads to a plausible model for the domain structure of ceruloplasmin.

Nitrite reductase (NIR) catalyzes the reduction of nitrite as part of the dissimilatory pathway in which dentrifying bacteria convert nitrate to nitrogen (Payne, 1985):

$$NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$$

The reaction product of NIR in vivo has been shown to be nitric oxide (NO) (Liu et al., 1986; Iwasaki & Matsubara, 1972). The reduction of nitrite to nitric oxide and the reduction of nitric oxide to nitrous oxide are two independent reactions in *Achromobacter cycloclastes* (Shapleigh & Payne,

† This work was supported by NIH Grants GM31770 and GM15731 and by NSF Grant DMB-8718646.
* Author to whom correspondence should be addressed.
‡ Department of Biological Structure, University of Washington.
§ Department of Biochemistry, University of Washington.
∥ Department of Biochemistry, University Georgia.
⊥ Department of Microbiology, University Georgia.

0006-2960/91/0430-7180$02.50/0

1985; Shapleigh et al., 1987).

The nitrite reductases of the denitrifying pathway are represented by two classes depending upon whether a heme group or copper is used in the active site. Copper-containing nitrite reductases have been isolated from at least six organisms in addition to this one. These include *Achromobacter xylosoxidans* (formerly *Alcaligines* sp. NCIB 11015 and *Pseudomonas denitrificans*) (Masuko et al., 1984), *Alcaligenes faecalis* strain S-6 (Kakutani et al., 1981), *Rhodobacter sphaeroides* (Sawada et al., 1978), *Pseudomonas aureofaciens* (Zumft et al., 1987), *Haloferax* (formerly *Halobacterium*) *denitrificans* (Tomlinson & Hochstein, 1988), and most recently *Bacillus halodenitrificans* (Denariaz et al., 1991). The enzyme from *Achromobacter cycloclastes* has been reported to bind either two (Iwasaki & Matsubara, 1972) or three (Liu et al., 1986) copper atoms per dimer of molecular weight 69K. The copper atoms have been characterized as two kinds of type I copper sites [having four protein ligands; a cysteine, two histidines, and a methionine in a distorted tetrahedral array (Dooley et al., 1988)] or a type I and a type II site [tetragonal environment (Iwasaki et al., 1975)].

Nitrite reductase from *Achromobacter cycloclastes* has been crystallized in this laboratory (Turley et al., 1988), and the interpretation of the 2.3-Å MIR electron density map is now complete. Amino acid sequence analysis was undertaken both to facilitate interpretation of the map and to clarify some of the physical properties of the molecule. The amino acid sequence of NIR from *Achromobacter cycloclastes* and the analysis of its relationship to other copper-containing proteins are reported here.

Materials and Methods

Nitrite reductase from *Achromobacter cycloclastes* was prepared as described previously (Liu et al., 1986), then reduced, and either S-carboxymethylated as described by Takio et al. (1983) or S-pyridylethylated (Andrews & Dixon, 1987). Cleavage at lysine residues was accomplished by incubation of nitrite reductase with *Achromobacter* protease I [a gift of Dr. T. Masaki, Ibaraki, Japan (Masaki et al., 1981)] in 50 mM Tris-HCl and 2 M urea, pH 9.0, for 24 h at 37 °C. Cleavage at glutamic acid residues was performed with the *Staphylococcus aureus* V8 protease in 0.1 M NH_4HCO_3 and 2 M urea, pH 8, for 18 h at 37 °C. Cyanogen bromide (2 mg) was added to the protein (25 nmol) in 72% formic acid for 15 h at room temperature in the dark. Asn–Gly bonds were cleaved with 2 M NH_2OH containing 6 M guanidine and 0.1 M K_2CO_3, pH 9, for 5 h at 45 °C. Finally, peptides were generated by cleavage at tryptophan residues with the addition of 1 mg of BNPS-skatole to 2 mg of protein in 0.25 mL of 80% glacial acetic acid. The mixture was stirred at room temperature for 24 h.

Peptide K9 was subdigested at arginine residues with trypsin in 0.1 M NH_4HCO_3 and 2 M urea for 6 h at 37 °C. Endoproteinase Asp-N cleavage of peptide N1 was accomplished in 0.1 M NH_4HCO_3 and 1 M urea for 10 h at 37 °C. Asp bonds in peptides N1 and E3 were selectively cleaved in 2 M formic acid for 4 h at 100 °C. For all enzymatic cleavages, an enzyme to substrate ratio of 1:100 (w/w) was used.

The peptides generated by cleavage at lysyl and glutamyl residues were first fractionated by exclusion chromatography on tandem TSK columns (Figure 1a,b). Pooled fractions were then purified to homogeneity by reversed-phase HPLC. Peptide mixtures generated by cleavage at methionine, tryptophan, and Asn-Gly residues were purified by using only the reversed-phase HPLC columns.

Amino acid compositions were determined from 20-h hydrolysates with a Waters Picotag system. Amino acid sequences were determined with an Applied Biosystems 470A sequencer with a Model 120A on-line HPLC system.

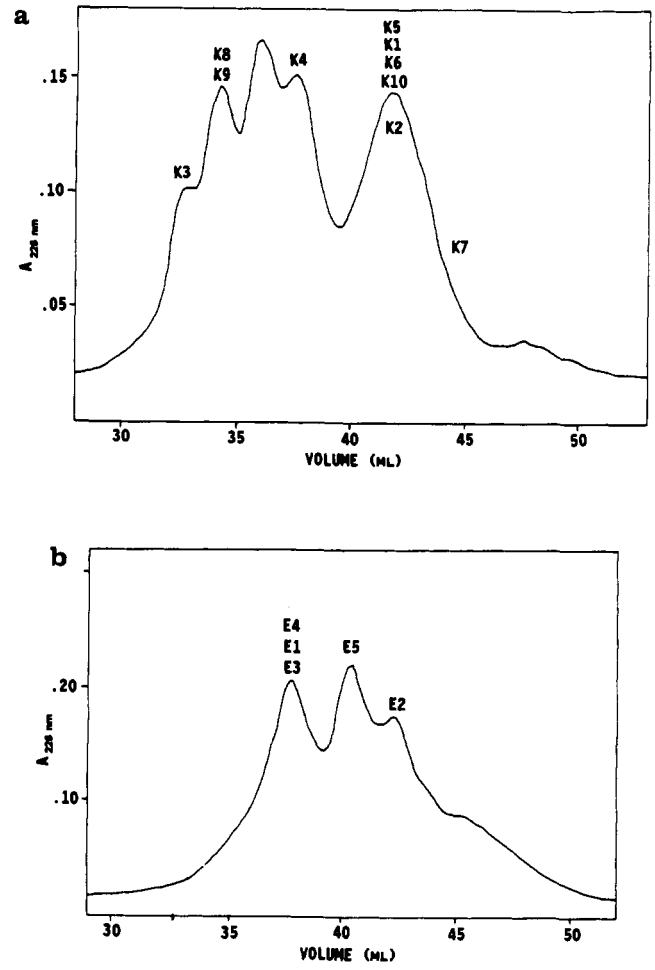

FIGURE 1: Primary separation by size-exclusion chromatography after enzymatic cleavages of nitrite reductase. (a) Cleavage of 150 nM nitrite reductase with *Achromobacter* protease I. (b) Cleavage of ~20 nM nitrite reductase at glutamic acids with *Staphylococcus aureus* V8 protease. The separations were performed in 6 M guanidine hydrochloride/10 mM sodium phosphate, pH 6.0, at a flow rate of 1 mL/min on tandem TSK-2000 SW (7.5 × 600 mm) columns. Fractions were pooled as indicated and further separated by reverse-phase HPLC using acetonitrile gradients and Synchropak RP-P or RP-8 columns. Peptides are listed from the top in increasing elution order on reverse-phase HPLC.

Searches for homologous sequences and alignments of homologous sequences were performed on a VAX computer using the PROFILE (DeHaen et al., 1976) and SEQCOMP (Gribskov et al., 1987) programs.

Results

The sequence data essential for proof of the primary structure of NIR are summarized in Figure 2. A large portion of the sequence (280 of 340 residues) was readily identified after the initial digest of the protein at lysyl residues. Peptides isolated from cleavage at glutamic acid residues were used to align most of the lysyl fragments. Only two additional peptides were required to establish a continuous sequence, one resulting from cleavage at methionine residues and the other from cleavage at aspartic acid residues.

Examination of Figure 2 indicates that only two lysyl fragments of significant length were not identified in the first digest. These fragments represent residues 1–16 and residues 34–48. The absence of the amino-terminal peptide was

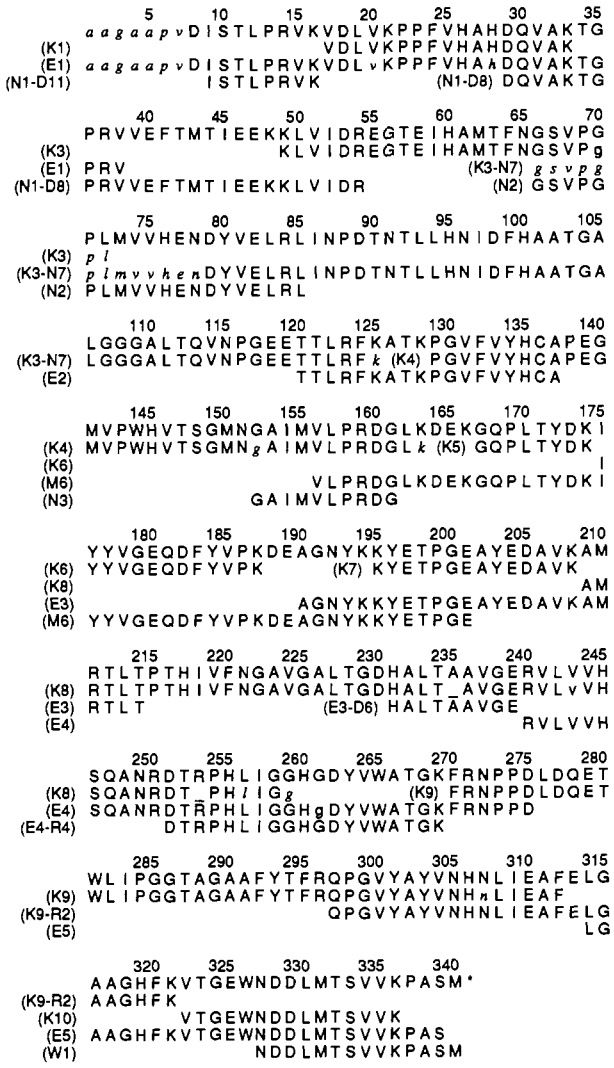

FIGURE 2: Amino acid sequence of nitrite reductase. The upper sequence in each group is the final assignment. Supporting data from Edman degradation of individual peptides are listed in one-letter code. K*n* denotes peptides from cleavage at Lys, E*n*, at Glu, M*n* at Met, N*n* at Asn–Gly, W*n* at tryptophan, and R*n* at arginine. Hyphenated symbol denotes a subdigestion product. Lower case letters indicate tentative identifications.

probably a result of proteolysis, as indicated by multiple sequences during direct sequence analysis of the undigested protein. This problem was significantly diminished by completing subsequent digests with a fresh preparation of protein. The second missing peptide (K2) was observed, but it was isolated with an equimolar amount of a contaminating peptide apparently not of nitrite reductase origin.

Subdigests were performed on two of the larger fragments (K3 and K9) to optimize the information gained from the digest. Edman degradation of the longest peptide (K3) had been prematurely terminated by an Asn-Gly sequence at residues 65–66. To circumvent this problem, K3 was treated with hydroxylamine, and the largest peptide purified was then sequenced. Surprisingly, there was a 3:1 mixture of Asp and Gly residues in the first sequencing turn. This apparently had resulted from a preferential cleavage of the Asn–Asp bond at residues 78–79 over the Asn–Gly bond at residues 65–66. The minor sequence (beginning with a Gly) was easily identified and provided a tentative assignment of the residues between the two cleavage sites, as indicated at the start of peptide K3–N7 (Figure 2). Hydroxylamine cleavage of the entire protein was later performed to confirm the identification of the residues located C-terminal to the Asn–Gly bonds, and peptides N2 and N3 corroborated these initial assignments. Analysis of *major* peptides of hydroxylamine cleavage of the entire protein did not show cleavage at Asn-Asp. [A total of three Asn–Gly bonds was observed in this protein (65–66, 151–152, 221–222).] In a separate subdigest, arginine bonds in peptide K9 were cleaved with trypsin to produce a fragment that was used to extend the sequence of this major lysyl peptide. Thus, the entire sequence of all isolated lysyl peptides was known except for the final four residues of peptide K10.

Overlaps provided by glutamyl and methionyl fragments established a continuous sequence from residue 49 to the carboxy terminus. All that remained was to link peptide E1 to the rest of the molecule. Examination of the residues flanking the region of the missing overlap (residues 39–48) indicated that the sequence could be completed with a peptide produced by cleavage at aspartic acid residues. Endoproteinase Asp-N was added to peptide N1 (residues 1–65) rather than the whole protein to simplify purification of the desired peptide. The sequence of peptide N1–D8 provided the remaining overlap.

Finally, it was necessary to confirm the sequence at both ends of the molecule. Proteolysis at the amino terminus precluded a straightforward interpretation of the sequence in this region. Sequence analysis of peptide E1 (~600 pmol) showed in general three to four residues at each turn of Edman degradation. After four turns, it became evident that the sequence of the second largest peaks paralleled the largest, but displaced by three turns, whereas the third largest was displaced by only one turn. Because of the high alanine content of the initial part of the sequence, it was not until the first unique amino acid came along (Val) that the ratio of peptide was determined to be 8:4:1:trace, which in turn led to the hypothesis that the ragged ends each started with alanines at positions, 1, 2, and 4. Confirmation of the sequence in this region was extended from residues 9 to 16 with peptide N1–D11, isolated after partial acid hydrolysis of N1. Although we are reasonably confident of the remaining sequence, assignment of the first eight residues must remain tentative due to the presence of multiple sequences.

The sequence of the carboxy terminus of the molecule was verified with peptide W1, obtained after treating the protein with BNPS-skatole. Sequence analysis of peptide W1 revealed a previously unidentified methionine as the final residue in the sequence. Mass spectrometric analysis of W1 unambiguously confirmed the indicated sequence for this peptide (data not shown).

Discussion

The sequence of this copper-containing protein is the first of its kind in the denitrification pathway. The fact that it contains type I copper, and all type I copper proteins have some sequence similarity, led us to expect to find related sequences in this protein. A type I copper center characteristically has Cys, His, and Met ligands to the copper, usually in the pattern CxxxHxxxM, preceded by a sequentially remote "upstream" histidine ligand. The finding of a *single* cysteine in the NIR sequence suggests there is only one type I site, since Cys is required for this kind of copper site. Candidates for the other type I ligands in the vicinity of Cys136 are His145, and Met150. His95 or His100 are in reasonable locations for the upstream histidine ligand. In a type II site, the copper is characteristically bound by histidine ligands, as in superoxide dismutase. There are many histidines present in the sequence suitable for ligating the second copper, so that it is not possible to deduce the ligands from the sequence alone.

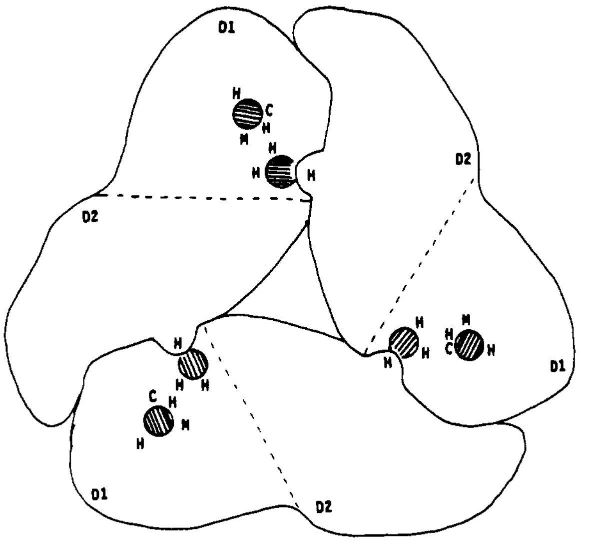

FIGURE 3: Cartoon of the subunit arrangement of NIR showing six copper sites, based on the 2.3-Å MIR X-ray structure (Godden et al., 1991). Domains are denoted D1 and D2; Cu sites are shaded spheres, and Cu ligands are denoted by the one-letter amino acid code.

Previous studies of the molecular weight of this protein have suggested that it is isolated as a dimer. Nevertheless, the X-ray structure of nitrite reductase (Godden et al., 1991) has revealed that the molecule is associated as a trimer, not a dimer. The chain tracing completed in the 2.3-Å NMR electron density map (greatly facilitated by the amino acid sequence) shows that within a monomer of 36 kDa, one domain (residues 8–175) folds into a Greek-key β-barrel containing a type I copper and that the ligands are indeed Cys136, His145, Met150, and His95. The folding of the second domain (residues 175–340) is also a β-barrel similar to the type I copper binding domains. However, the second copper is not a type I copper and is bound *between* molecules in the trimer. Its ligands are His100 and His135 from domain I of one molecule and His306 from domain 2 of the second. A cartoon of the trimeric arrangement is shown in Figure 3.

Initial attempts to align the sequences of the two domains of NIR against the type I copper-containing protein plastocyanin or with each other were unsuccessful using only the sequences and no three-dimensional information other than that the two domains folded similarly. However, domain 1 could be aligned against the blue copper-containing domains of multicopper oxidases [laccase, ascorbate oxidase (AO), and ceruloplasmin]. When the latter were grouped as a family to optimize the alignment, domain 1 *and* domain 2 of NIR could each be aligned with the family and therefore indirectly with each other. The alignment is shown in Figure 4. This observation suggests that both domains are indeed related to a blue copper domain even though domain 2 lacks the usual type I Cu ligands.

Although these alignments are statistically significant (domain 1 is 10 standard deviations above an alignment of a randomly generated sequence of the same amino acid composition, whereas domain 2 is 5 standard deviations above), the relationship is much more apparent in the three-dimensional structure. Superposition of the three-dimensional structures of the two domains indicates that the primary structure alignment in large part parallels the three-dimensional alignment, with most differences lying in loop regions. For example, residues 136–150 of domain 1 contain the CxxxHxxxM ligands to the type I copper center in domain 1. The analogous loop in domain 2, which should contain the trio of ligands CxxxHxxxM (roughly 305–315), not only does not contain any of them but also is shorter by three residues. A second noteworthy difference lies in the vicinity of the

```
                   . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . |
1: NIR      17-67: V D L V K P P F V H A H D Q V A K T G P R V V E F T M T I E E K K L V I D R E G T E I H A M T F N G
2: NIR    190-223: - - - - - - - - - - - - - - - - A G N Y K K Y E T P G E A Y E D A V K A M R T L T P T H I V F N G A
3: AO     368-418: V S L A L P P T P Y L A A M K M R L N T A F N Q N P P P E T F P L N Y D I N N P P P N P E T T T G N
4: LAC    401-450: N I V Q V E G V N Q W K Y W L I E N D P D G A F S L P H P I H L H G H D F L I L G R S P D V T A I -
5: CP     198-247: F V V M F S V V D E N F S W Y L E D N I K T Y C S E P E K V D K D N E D F Q E S N R M Y S V N G Y T
6: CP     559-608: F Y L F P T V F D E N E S L L L E D N I R M F T T A P D Q V D K E D E D F Q E S N K M H S M N G F M
7: CP     897-946: F A L L F L V F D E N E S W Y L D D N I K T Y S D H P E K V N K D D E E F I E S N K M H A I N G R M
                       4                               5               5         4       4 4   4         4

                   . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . |
1: NIR     68-103: S V P G P L - - - M V V H E N D Y V E L R L I N P D T N T L L H N I D F H A A T - - - - - - - - - -
2: NIR    224-268: V G A L T G D H A L T A A V G E R V L V V H S Q A N R D T R P H L I G G H G D Y V W A T G - - - - -
3: AO     419-466: G V Y K F N - - - M G E T V D V I L Q N A N M L N P N M S E I H P W H L H G H D F W V L G Y G E K
4: LAC    451-464: - - - - - - - - - - - - - - - - S Q T R Y V F D P A V D M A - - - - - - - - - - - - - - - - - - - -
5: CP     248-285: F G S L P G - - - L S M C A E D R V K W Y L F G M G N E V D V H A A F F H G Q A L - - - - - - - - -
6: CP     609-646: Y G N Q P G - - - L T M C K G D S V V W Y L F S A G N E A D V H G I Y F S G N T Y - - - - - - - - -
7: CP     947-984: F G N L Q G - - - L T M H V G D E V N W Y L M G M G N E I D L H T V H F H G H S F - - - - - - - - -
                     4       4       4           4   5     4 4         4         6       4 5 5

                   . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . |
1: NIR    104-141: - - - - - - - - - - - - G A L G G G A L T Q V N P G E E T T L R F K A T K P G V F V Y H C A P E G M
2: NIR    269-310: - - - - - - - - K F R N P P D L D Q E T W L I P G G T A G A A F Y T F R Q P G V Y A Y V N H N L I E
3: AO     467-513: F Y A P E D E K K L N L K N P P L R N T V V I F P Y G W T A I R F V A D N P G V W A F H C - - - - H
4: LAC    465-501: - - - - - - - - R L N G N N P T R R D T A M L P A K G W L L I A F R T D N P G S W L M H C - - - - H
5: CP     286-320: - - - - - - - - T N K N - - - Y R I D T I N L F P A T L F D A Y M V A Q N P G E W M L S C - - - - Q
6: CP     647-682: - - - - - - - - L W R G - - - E R R D T A N L F P Q T S L T L H M W P D T E G T F N V E C - - - - L
7: CP    985-1022: - - - - - - - - Q Y K H R G V Y S S D V F D I F P G T Y Q T L E M F P R T P G I W L L H C - - - - H
                                                       4 5       4 5   4                     6 7   4     4 6

                   . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . | . . . . . . . . . |
1: NIR    148-165: V P W H V T S G M N G A I M V L P R D G L K D E - - - - - - - - - - - - - - - - - - - - - - - - - -
2: NIR    311-330: A F E L G A A G H F K V T G E W N D D L - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
3: AO     513-557: I E P H L H M G M G V V F A E G V H M V G M I P P K A L A C G S T A L V K N Y P R L P - - - - - - -
4: LAC    501-545: I A W H V S G G L S N Q F L E R A Q D L R N S I S P A D K K A F N D N C D A W R A Y F P D N A P F P
5: CP     321-350: N L N H L K A G L Q A F F Q V Q E C N K S S S K D N I R G K - - - - - - - - - - - - - - - - - - - -
6: CP     683-710: T T D H Y T G G M K Q K Y T V N Q C R R Q S E D S T F Y L - - - - - - - - - - - - - - - - - - - - -
7: CP   1022-1046: V T D H I H A G M E T T Y T V L Q N E D T K S G - - - - - - - - - - - - - - - - - - - - - - - - - -
                         6       7 4       3   4
```

FIGURE 4: Alignment of sequences of portions of domain 1 and domain 2 of NIR against the blue copper domains of the multi-copper oxidases *Neurospora crassa* laccase, zucchini ascorbate oxidase, and human ceruloplasmin. The symbols below the alignment signify the number of identities at that position. Stretches of sequence known to be part of β-sheet in NIR and in ascorbate oxidase are in enlarged, boldface letters.

```
CP  D1   99 T - - F H S H G I T Y - - -  155 C V T R I Y H S H - I D A P K(4)G L
    D3  456 S - - I E P I G V R F - - -  515 C L L A K Y Y S A - V D P T K(4)G L
    D5  824 S - - I H A H G V Q - - - -  855 C I P W A Y Y S T - V D Q V K(4)G L

CP  D2  275 V H A A F F H G Q A L - - -  312 P G E W M L S C Q N L N H L K A G L
    D4  636 V H G I Y F S G N T Y - - -  673 E G T F N V E C L L T T H Y T G G M
    D6  974 L H T V H F H G H S F - - -  1014 P G I W L L H C H V T D H I H A G M

LAC D1   91 G T S I H W H G M H - - - -  134 Y G T S W Y H S H F S A Q Y G N G I
    D2  265 G L L L G I G Q R Y - - - -  300 C G T S D N K Y P A A I F R (12) G V
    D3  427 P H P I H L H G H D F - - -  493 P G S W L M H C H I A W H V S G G L

                2 3                                 3 3
AO  D1   58 G V V I H W H G I - - - - -  100 A G T Y F Y H G H L G M Q R S A G L
    D2  228 S L N F A I G N H E L - - -  274 P - L E N Y W V S I G V R A R (6) G L
    D3  446 I H P W H L H G H D F - - -  503 P G V W A F H C H I E P H L H M G M
              1   2 3                               3 1 3     1       1

              1       3                             3 1       1       1
NIR D1   94 L H N I D F H A A T - - - -  129 P G V F V Y H C A (5) P W H V T S G M
    D2  251 D T R P H L I G G H - - - -  298 P G V Y A Y V N H N L I (2) F E L G A
                                                    3

NOR         D V S H G F V V V N H G - - -   P G L H W Y Y C S W F C H A L H M E M V
COT         D V I H S F W V P G L A - - -   P G L Y Y G F C A E L C G A S H A R M L
COH         D V L H S W A V P S L G - - -   P G V Y Y G Q C S E I C G A N H S F M P
```

FIGURE 5: Alignment of the copper binding regions of domain 1 and domain 2 of NIR with the structural alignment of Messerschmidt and Huber (1990) of ascorbate oxidase, laccase, and ceruloplasmin. Ligand residues are identified by the number of the type of copper to which they bind, where known. Putative ligands are in boldface. Abbreviations: CP, ceruloplasmin; LAC, laccase; AO, ascorbate oxidase; NIR, nitrite reductase; NOR, N_2O reductase (Viebrock & Zumft, 1988); COT, cytochrome oxidase subunit IIc from *Thermus thermophilus* (Mather et al., 1991); COH, cytochrome oxidase subunit II from human (Anderson et al., 1981).

"upstream" histidine ligand. The alignment of the primary sequence suggests His255 in domain 2 is in a position corresponding to the "upstream" His ligand in domain 1 (His95), but the three-dimensional comparison reveals that the loop preceding His255 is shortened by three residues and that the 3D alignment is actually the following:

```
      95      103
(D1) ...HNIDFHAAT...
(D2) ...TRPHLIGGH...
     252        260
```

The alignment comes back in register further along in a loop subsequent to the β-strand containing these histidines (109–115 vs 280–285). Possibly each half of the monomer did evolve from gene duplication, and these histidines were once equivalent, but in subsequent changes, the folding of the protein forced these residues to lie in different locations with respect to the three-dimensional structure, consistent with different function.

The alignment of the sequences shown in Figure 4 of domains 1 and 2 with the family of blue copper oxidases was not straightforward or unique. Messerschmidt and Huber have published an alignment (MH) of the blue copper oxidases using information from the X-ray structure of zucchini ascorbate oxidase (Messerschmidt & Huber, 1990). As is the case with NIR, the oxidases contain domains which fold in the same way as the type I Cu-containing domain, but lack a type I copper. It is possible to align portions of NIR domains 1 and 2 with the MH alignment, and the portions surrounding the copper ligands are shown in Figure 5. Aside from the use of structure to aid in the alignment, an important difference from the alignment of Figure 4 is the inclusion of the oxidase domains *not* binding type I Cu. An interesting twist is that it is the C-terminal domain which binds the type I Cu in the oxidases, whereas in NIR it is the N-terminal domain.

The X-ray structure of ascorbate oxidase shows that it has a novel trinuclear copper center in which a type III pair of copper atoms and a single type II copper are bound by eight histidines, four from each of two domains within a monomer. It can be seen from Figure 5 that certain key residues in the vicinity of the ascorbate oxidase trinuclear cluster are not His in the NIR sequence, so that it is not likely that NIR could bind a trinuclear cluster, and indeed it does not. Some of the

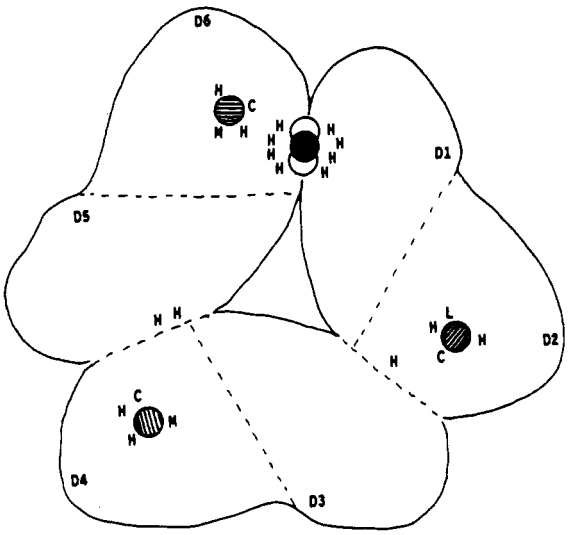

FIGURE 6: Proposed arrangement of six domains of ceruloplasmin based on sequence similarities to ascorbate oxidase and sequence and structural similarities between ascorbate oxidase and nitrite reductase. Domains are designed D1–D6; dotted lines represent domain boundaries; copper sites are shaded spheres, and copper ligands and histidines which appear to be sequentially correlated to histidines in NIR are denoted by the one-letter amino acid code.

histidines in NIR are homologous to some of these eight, so that by analogy with ascorbate oxidase, the second copper of NIR would correspond to 1/2 of the type III pair. Moreover, the locations of the histidines in residues 252–260 appear to be more similar to histidines of laccase domain 1 than ascorbate oxidase domain 1.

Figure 5 also includes sequences of other copper proteins, such as N_2O reductase of *Pseudomonas stutzerei* and subunit II of mammalian cytochrome oxidase, previously suggested to be related to each other (Scott et al., 1989) and to the small blue copper proteins. The recently determined sequence of subunit IIc of the two-subunit cytochrome oxidase of *Thermus thermophilus* (Mather et al., 1991) is also included. Certain portions of the NIR sequence are similar to these sequences as well. Although the alignment is fragmentary, the presence of the CxxxHxxxM ligands in these proteins does suggest that these too will have Greek-key domains. It is extremely interesting that the N_2O reductase, a protein found further along in this pathway, appears to have some similarity.

Messerschmidt and Huber (1990) have suggested from the structurally based alignment of ceruloplasmin to ascorbate oxidase that ceruloplasmin is likely to consist of six domains and that at least two of these domains will bind a type I copper, while a trinuclear copper center will be bound between domains 1 and 6. Since ceruloplasmin domains 2, 4, and 6 are more similar to each other than they are to domains 1, 3, and 5, it is likely that a gene duplication event preceded the gene triplication event, supporting the idea that ceruloplasmin could have evolved from a trimeric arrangement of a two-domain monomer such as NIR. Because the sequence of NIR domains 1 and 2 can be aligned with the ascorbate oxidase and ceruloplasmin sequences, and because there are six domains in the trimer of two-domain NIR molecules, a plausible three-dimensional arrangement of the six domains of ceruloplasmin is as shown in Figure 6. The presence of six different domains in ceruloplasmin, at least three of which may participate in copper binding, has resulted in the possibility of new activities relative to NIR. It has been shown that ceruloplasmin has superoxide dismutase activity, ferroxidase activity, and electron transfer from the two different type I copper centers (Ryden, 1984).

ACKNOWLEDGMENTS

We are grateful for the help of Maria Harrylock with the Edman instrumentation. We thank Dr. Dave Teller for his advice and guidance in carrying out the sequence comparisons and Dr. Ken Walsh for his helpful advice on sequencing and on the manuscript.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C.; Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., & Young, I. G. (1981) *Nature 290*, 457–465.

Andrews & Dixon (1987) *Anal. Biochem. 161*, 5–24.

DeHaen, C., Swanson, E., & Teller, D. C. (1976) *J. Mol. Biol. 106*, 639–661.

Denariaz, G., Payne, W. J., & LeGall, J. (1991) *Biochim. Biophys. Acta 1056*, 255–262.

Dooley, D. M., Moog, R. S., Liu, M.-Y., Payne, W. J., & LeGall, J. (1988) *J. Biol. Chem. 263*, 14625–14628.

Godden, J. W., Turley, S., Teller, D. C., Adman, E. T., Liu, M.-Y., Payne, W. J., & LeGall, J. (1991) *Science*, (in press).

Gribskov, M., McLachlan, A. D., & Eisenberg, D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4355–4358.

Iwasaki, H., & Matsubara, T. (1972) *J. Biochem. 71*, 645–652.

Iwasaki, H., Noji, S., & Shidara, S. (1975) *J. Biochem. 78*, 355–361.

Kashem, M. A., Dunford, H. B., Liu, M.-Y., Payne, W. J., & LeGall, J. (1987) *Biochem. Biophys. Res. Commun. 145*, 563–568.

Liu, M.-Y., Liu, M.-C., Payne, W. J., & LeGall, J. (1986) *J. Bacteriol. 166*, 604–608.

Masaki, T., Tanabe, M., Nakamura, K., & Soejima, M. (1981) *Biochim. Biophys. Acta 660*, 44.

Mather, M. W., Springer, P., & Fee, J. A. (1991) *J. Biol. Chem. 266*, 5025–5035.

Messerschmidt, A., & Huber, R. (1990) *Eur. J. Biochem. 187*, 341–352.

Messerschmidt, A., Rossi, A., Ladenstein, R., Huber, R., Bolognesi, M., Gati, G., Marchesini, A., Petruzzelli, R., & Finazzi-Agro, A. (1989) *J. Mol. Biol. 206*, 513–529.

Payne, W. (1985) in *Denitrification in the Nitrogen Cycle* (Golterman, H. L., Ed.) pp 47–65, Plenum Publishing Corp., New York.

Ryden, L. (1984) in *Copper Proteins and Copper Enzymes* (Lontie, R., Ed.) Vol. III, pp 37–100, CRC Press, Boca Raton, FL.

Sawada, E., Satoh, T., & Kitamura, H. (1978) *Plant Cell Physiol. 19*, 1339–1351.

Scott, R. A., Zumft, W. G., Coyle, C. L., & Dooley, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4082–4086.

Shapleigh, J. P., & Payne, W. J. (1985) *FEMS Microbiol. Lett. 26*, 275–279.

Shapleigh, J. P., Davies, K. J. P., & Payne, W. J. (1987) *Biochim. Biophys. Acta 911*, 334–340.

Takio, K., Smith, S. B., Walsh, K. W., Krebs, E. G., & Titani, K. (1983) *J. Biol. Chem. 258*, 5531–5536.

Tomlinson, G. A., & Hochstein, L. I. (1988) Abstracts of the Annual Meeting of the American Society of Microbiology Miami, FL, p 209.

Turley, S., Adman, E. T., Sieker, L. C., Liu, M.-Y., Payne, W. J., & LeGall, J. (1988) *J. Mol. Biol. 200*, 417–419.

Viebrock, A., & Zumft, W. G. (1988) *J. Bacteriol. 170*, 4658–4668.

Zumft, W. G., Gotzman, D. J., & Kroneck, P. M. H. (1987) *Eur. J. Biochem. 168*, 301–307.