# The amino acid sequence of the blue copper protein of *Alcaligenes faecalis*

Scott Hormel, Elinor Adman+, Kenneth A. Walsh, Teruhiko Beppu* and Koiti Titani

*Departments of Biochemistry and +Biological Structure, University of Washington, Seattle, WA 98195, USA* and **Department of Agricultural Chemistry, University of Tokyo, Tokyo 113, Japan*

Received 10 December 1985

The complete amino acid sequence of a blue copper protein from *Alcaligenes faecalis* S-6 has been determined. This protein is clearly homologous to pseudoazurins in *Achromobacter cycloclastes* and *Pseudomonas* AM1, more distantly related to plant plastocyanins, and markedly different from the azurin of *Pseudomonas aeruginosa*. Yet all of these proteins bind copper, and analogous ligands appear to be involved.

*Amino acid sequence* *X-ray crystallography* *Blue copper protein* *Azurin* *Plastocyanin* *Pseudoazurin*

## 1. INTRODUCTION

Blue copper proteins are low molecular mass (10–15 kDa) soluble electron transfer proteins found in bacteria and plants [1,2]. Numerous spectroscopic investigations [3] and the X-ray crystallographic determination of the structures of plastocyanin and azurin from two sources [4–6] have shown that the copper is bound in each case by one Cys, one Met and two His residues in a distorted tetrahedral array.

A blue protein in *Alcaligenes faecalis* S-6 is required for inactivation (in the presence of $O_2$) of a copper-containing nitrite reductase [7]. It has properties which originally suggested that it might belong to another class of blue proteins including amicyanin [8] and an 'azurin' from *Achromobacter cycloclastes* [9]. We have begun the X-ray diffraction determination of its structure [10] in order to understand how this protein is related to the other blue proteins, and ultimately how it interacts with its redox partner, nitrite reductase, which can also be crystallized [11].

An essential part of this structure determination is knowledge of the amino acid sequence. A preliminary amino-terminal sequenator analysis (by A. Tsugita, communicated by R.P. Ambler) has already helped our initial map interpretation. The following reports the complete sequence determination of this protein. The details of the X-ray investigation will be published subsequently.

## 2. MATERIALS AND METHODS

The entire sequence was determined with less than 5 mg protein crystallized according to Kakutani et al. [7]. The protein was reduced with dithiothreitol, carboxymethylated with iodoacetic acid, and cleaved at methionyl bonds as described in [12]. Lysyl bonds were cleaved with *Achromobacter* protease I [13] (a gift of Dr T. Masaki, Ibaraki, Japan) in 2 M urea at 37°C for 12 h. Peptide M6 was subdigested with *Staphylococcus aureus* V8 protease (Miles) in 0.1 M $NH_4HCO_3$, pH 8, 37°C, for 12 h.

Peptides generated by cleavage at Met, Lys or Glu were purified by reverse-phase HPLC using a SynChropak RP-8 column and an acetonitrile gradient into dilute aqueous trifluoroacetic acid [14]. Peptides K2 and SP1 were purified further on SynChropak RP-P and Cosmosil 5 C18P columns, respectively. Amino acid compositions were determined with a Waters Picotag system [15], sequence analysis with a Beckman 890C sequencer, phen-

*Published by Elsevier Science Publishers B.V. (Biomedical Division)*
00145793/86/$3.50 

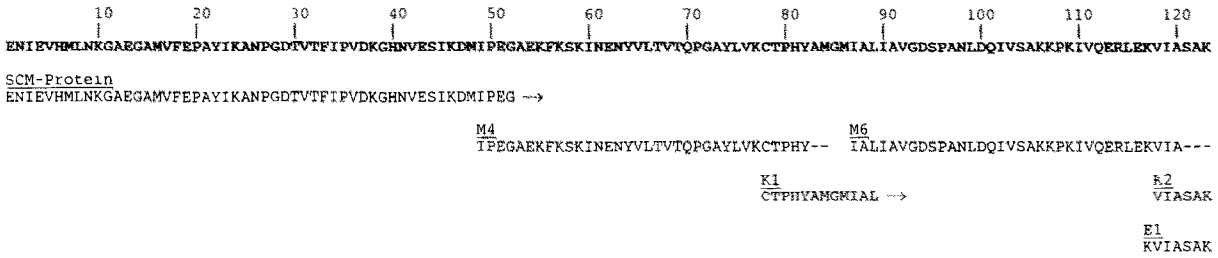


Fig.1. Summary of the proof of sequence. The sequences of specific peptides (names are underlined) are given in one-letter code below the summary sequence (bold type). Prefixes M, K and E denote peptides generated by cleavage at the carboxyl side of Met, Lys and Glu, respectively. A dash indicates an unidentified amino acid; an arrow denotes a longer unidentified sequence.

thiohydantoins by HPLC, and homologies by the SEARCH or ALIGN programs, all previously described [12].

## 3. RESULTS

The results of Edman degradation of the intact *S*-carboxymethylated (S-CM-)protein and of 5 peptides derived therefrom are summarized in fig.1. Together these 6 analyses provide an overlapping set of information that describes the unique sequence of the 123-residue protein.

Fragmentation at methionyl residues with cyanogen bromide provided the two principal fragments M4 and M6, yielding most of the sequence of residues 54–120. The isolation procedure and the amino acid compositions of all of the cyanogen bromide fragments are presented in fig.2 and table 1. A small dipeptide (Gly-Met, residues 85–86) was not recovered in the experiment depicted in fig.2, but was probably located in the break-through fraction. Peptide M4 clearly overlapped the amino-terminal sequence of the protein beginning at Ile 49. Cleavage at lysyl residues provided peptide K1, which overlapped M4 to M6 and included the missing dipeptide. Peptide K2 overlapped the carboxyl-terminus of M6. Since no other peptide in this digest lacked lysine, it appeared that the carboxyl-terminus of the whole protein might be lysine. This was verified by digestion of M6 with *S. aureus* V8 protease at glutamyl residues and analysis of the sequence of peptide E1 (fig.1). The finding of this peptide, with a carboxyl-terminal lysine after digestion at glutamyl residues, indicates that the carboxyl-terminus of the protein is Lys 123.

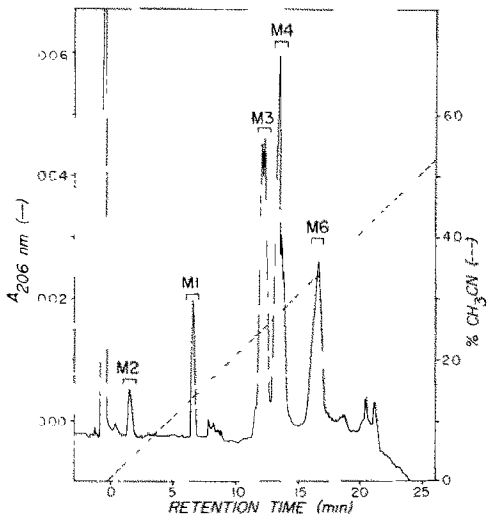


Fig.2. Separation of peptides after digestion of 0.25 mg S-CM-protein with CNBr. The flow rate was 2.0 ml/min through a SynChropak RP-8 column.

| | Alcal. | Pseudoazurin Ps. AM1 | Pseudoazurin Achrom. | Plasto-cyanin Elder | Azurin Ps. Aer. |
|---|---|---|---|---|---|
| Present Work (Alcal.) | -- | | | | |
| Pseudoazurin (Ps. AM1) | 24.3 | | | | |
| Pseudoazurin (Achrom.) | 25.9 | 20.4 | | | |
| Plastocyanin (Elder) | 5.2 | 3.4 | 5.3 | | |
| Azurin (Ps. Aer.) | -0.9 | -0.4 | 0.3 | 7.7 | |
| Amicyanin (Ps. AM1) | 0.7 | 4.2 | 1.8 | 7.4 | 5.1 |

Fig.3. Sequence similarities among selected blue copper proteins, as quantified by the ALIGN program. An alignment score greater than 3 corresponds to a probability of less than 0.0014 that a randomly generated sequence would produce a better score. Alcal, *A. faecalis*; Ps. AM1, *Pseudomonas AM1*; Achrom, *Achrom. cycloclastes*; Ps. Aer, *P. aeruginosa*.

Table 1

Amino acid composition of cyanogen bromide fragments[a]

| | Fragments: | M1 | M2 | M3 | M4 | M5 | M6 | Protein composition |
|---|---|---|---|---|---|---|---|---|
| | Residues: | 1–7 | 8–16 | 17–48 | 49–84 | 85–86 | 87–123 | (from sequence) |
| Asx | (D and N) | 1.0 (1) | 0.7 (1) | 3.0 (5) | 1.4 (2) | | 3.4 (3) | (12) |
| Glx | (E and Q) | 1.9 (2) | 1.0 (1) | 2.0 (2) | 3.2 (4) | | 4.0 (4) | (13) |
| CMC | (C) | | | | 1.0 (1) | | | (1) |
| Ser | (S) | | | 1.1 (1) | 2.5 (1) | | 2.4 (3) | (5) |
| Gly | (G) | | 2.1 (2) | 2.4 (2) | 3.4 (2) | (1) | 1.4 (1) | (8) |
| His | (H) | 0.9 (1) | | 0.8 (1) | 0.8 (1) | | | (3) |
| Arg | (R) | | | | | | 1.0 (1) | (1) |
| Thr | (T) | | | 1.5 (2) | 2.0 (3) | | | (5) |
| Ala | (A) | | 2.0 (2) | 2.0 (2) | 3.1 (3) | | 6.0 (6) | (13) |
| Pro | (P) | | | 3.0 (3) | 2.0 (3) | | 2.0 (2) | (8) |
| Thyr | (Y) | | | 0.8 (1) | 2.2 (3) | | | (4) |
| Val | (V) | 1.0 (1) | | 3.8 (4) | 2.6 (3) | | 3.6 (4) | (12) |
| Met[b] | (M) | 0.3 (1) | 0.3 (1) | 0.3 (1) | 0.3 (1) | (1) | | (5) |
| Ile | (I) | 1.3 (1) | | 3.8 (3) | 3.3 (2) | | 5.6 (5) | (11) |
| Leu | (L) | | 1.0 (1) | | 2.0 (2) | | 3.2 (3) | (6) |
| Phe | (F) | | | 2.0 (2) | 1.0 (1) | | | (3) |
| Lys | (K) | | 1.0 (1) | 2.7 (3) | 3.2 (4) | | 5.2 (5) | (13) |
| Number of residues | | 7 | 9 | 32 | 36 | 2 | 37 | (123) |
| Yield (%) | | 57 | 51 | 57 | 49 | | 52 | |

[a] Residues/molecule by amino acid analysis or (in parentheses) from the sequence. Peptide M5 (Gly-Met, residues 85–86) was not recovered

[b] Recovered as homoserine

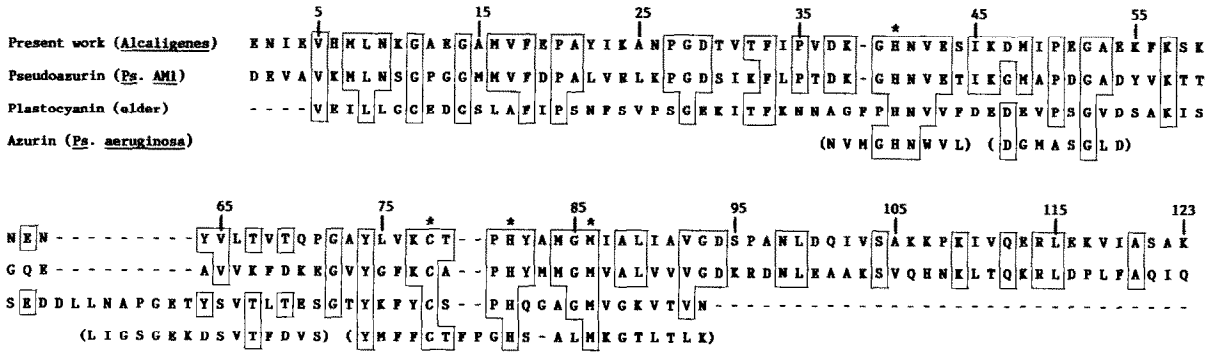


Fig.4. Sequence alignment indicating that the *A. faecalis* blue copper protein is more closely related to a pseudoazurin from *Pseudomonas AM1* than to either a plastocyanin or an azurin (cf. fig.3). Residues identical with the *A. faecalis* sequence are boxed. Asterisks indicate probable copper ligand sites (see text). The 4 segments of azurin demarcated by parentheses depict its residues 42–50, 62–69, 86–100 and 108–128, in that order.

## 4. DISCUSSION

A search of the May 1985 protein databank (National Biomedical Research Foundation) for sequences showing the greatest similarity to the *Alcaligenes* protein yielded all of the known plastocyanins but none of the azurins, indicating that the new microbial protein is more closely related to the plant plastocyanins than to the microbial azurins. Recently, Ambler and Tobari [16] have

described the sequences of two more blue copper proteins from *Pseudomonas AM1,* pseudoazurin (123 residues) and amicyanin (99 residues). Of the various blue copper proteins examined by others, ours clearly bears the greatest resemblance, at the level of sequence, to the pseudoazurin of *Pseudomonas AM1* and to its homolog, the pseudoazurin of *Achrom. cycloclastes* (figs 3 and 4).

Previous sequence comparisons demonstrated the likelihood of analogous copper ligands in plastocyanin and azurin [17], and this has been confirmed in their 3-dimensional structures. The present sequence also contains putative copper ligands at His 40, Cys 78, His 81 and Met 86. It is now clear that while plastocyanin is similar to azurin, and pseudoazurin is similar to plastocyanin, pseudoazurin is much less similar to azurin. The details of the current 3-dimensional structure determination will clarify these relationships.

## ACKNOWLEDGEMENTS

The authors are indebted to Dr H. Watanabe for his help in the isolation of the protein, to Dr Koji Takio for analysis of a fragment on a gas-phase sequencer, to Roger D. Wade for amino acid analysis, and to Santosh Kumar for advice and guidance. The authors also thank Drs A. Tsugita and R.P. Ambler for communicating preliminary sequence data indicating a similarity with the pseudoazurin from *Achrom. cycloclastes.* This work was supported in part by grants from the NIH, GM-15731, GM-31770 and HL-29595.

## REFERENCES

[1] Spiro, T. (1981) Copper Proteins, John Wiley, New York.
[2] Adman, E.T. (1984) in: Topics in Molecular and Structural Biology (Harrison, P. ed.) vol. I, pp. 1–42, Metalloproteins, McMillan, London.
[3] See references summarized in [1,2] and in the series: R. Lontie (1984) Copper Proteins and Copper Enzymes, vols I–III, CRC Press, Boca Raton, FL.
[4] Guss, J.M. and Freeman, H.C. (1983) J. Mol. Biol. 165, 521–563.
[5] Norris, G.E., Anderson, B.F. and Baker, E.N. (1983) J. Mol. Biol. 165,501–521.
[6] Adman, E.T. and Jensen, L.H. (1981) Isr. J. Chem. 21, 8–12.
[7] Kakutani, T., Watanabe, H., Arima, K. and Beppu, T. (1981) J. Biochem. 89, 463–472.
[8] Iwasaki, H. and Matsubara, T. (1973) J. Biochem. 73, 659–661.
[9] Ambler, R.P. (1977) in: The Evolution of Metalloenzymes, Metalloproteins and Related Materials, Proceedings of a Symposium of the Inorganic Biochemistry Discussion Group of the Chemical Society, University of Sussex (Leigh, G.S. ed.) pp. 100–118, Symposium Press, London.
[10] Adman, E.T., Beppu, T. and Watanabe, H. (1984) Acta Crystallogr. A40, Suppl. C-33.
[11] Kakutani, T., Watanabe, H., Arima, K. and Beppu, T. (1981) J. Biochem. 89, 453–461.
[12] Titani, K., Sasagawa, T., Ericsson, L.H., Kumar, S., Smith, S.B., Krebs, E.G. and Walsh, K.A. (1984) Biochemistry 23, 4193–4199.
[13] Masaki, T., Tanabe, M., Nakamura, K. and Soejima, M. (1981) Biochim. Biophys. Acta 660, 44–50.
[14] Mahoney, W.C. and Hermodson, M.A. (1980) J. Biol. Chem. 255, 11199–11203.
[15] Bidlingmeyer, B.A., Cohen, S.A. and Tarvin, T.L. (1984) J. Chromatogr. 336, 93–104.
[16] Ambler, R.P. and Tobari, J. (1985) Biochem. J. in press.
[17] Ryden, L. and Lundgren, J.-O. (1976) Nature 261, 344–346.