Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 6434–6437, November 1982
Biochemistry

Amino acid sequence of a basic blue protein from cucumber seedlings

(copper proteins/azurin/plastocyanin/stellacyanin/sequence homology)

M. MURATA*, G. S. BEGG†, F. LAMBROU†, B. LESLIE†, R. J. SIMPSON†, H. C. FREEMAN*‡, AND F. J. MORGAN†

*Department of Inorganic Chemistry, The University of Sydney, New South Wales 2006, Australia; and †St. Vincent's School of Medical Research, Fitzroy, Victoria 3065, Australia

Communicated by Harry B. Gray, July 12, 1982

ABSTRACT **The amino acid sequence of a type 1 copper protein, the 96-residue basic blue protein from cucumber seedlings, has been determined by Edman degradation of the intact molecule and of fragments produced by cleavage with cyanogen bromide and with trypsin. The cucumber basic blue protein shows a marked sequence homology with stellacyanin, and to a smaller degree with plastocyanin and azurin. The known copper ligands of plastocyanin and azurin (corresponding to histidine-37, cysteine-84, histidine-87, and methionine-92 in plastocyanin) are present in the cucumber basic blue protein. However, the latter also contains a half-cystine residue analogous to the suggested fourth ligand of stellacyanin, where methionine is absent.**

Blue or "type 1" copper proteins are characterized by an intense electronic absorption band near 600 nm, by a small hyperfine splitting constant $A_{\parallel}$ in the EPR spectrum, and in most cases, by exceptionally high reduction potentials. Two of these proteins, azurin and plastocyanin, have relatively well-defined functions in biological electron transfer reactions. The three-dimensional structures of azurin from *Pseudomonas aeruginosa* and plastocyanin from *Populus nigra* var. *italica* have been determined (1, 2), and the amino acid sequences of azurins and plastocyanins from a wide variety of sources are available (see ref. 3).

By contrast, a number of other blue copper proteins have as yet no known biological functions. These proteins are stellacyanin (4–6), umecyanin (7), mavicyanin (8), and blue copper proteins from rice bran (9), mung bean (10), cucumber peelings (11), and cucumber seedlings (12, 13). The amino acid sequence of only one protein in this group, stellacyanin, has been published (14).

In this communication, we report the amino acid sequence of the cucumber basic blue protein (CBP) from cucumber seedlings.§ CBP has previously been characterized by a M_r of 10,100 (gel filtration); by electronic visible absorption bands at 443 nm (ε = 2,030 $M^{-1}cm^{-1}$), 597 nm (ε = 3,400 $M^{-1}cm^{-1}$), and 750 nm (ε = 1,800 $M^{-1}cm^{-1}$); by an EPR spectrum that exhibits a strong rhombic distortion and a small hyperfine splitting constant; and by a reduction potential of E^0 = 317 mV at pH 7 (12, 13). The spectroscopic properties, but not the redox potential, strongly resemble those of stellacyanin (5). Crystals of the protein suitable for x-ray structure analysis have been obtained (13).

MATERIALS AND METHODS

CBP was isolated by using the method described by Vickery (12) with some modifications. The method consists of batchwise extraction of the protein with Bio-Rex 70 carboxylic resin (Bio-Rad) and repeated ion-exchange column chromatography on the same resin and then on CM-Sephadex C-25 (Pharmacia). The buffer used was 0.02 M sodium phosphate at pH 7.6, and elution from ion-exchange columns was achieved by a linear NaCl gradient from 0 to 0.5 M made in the same buffer. The final purification was made on Sephadex G-75 with 0.05 M phosphate buffer at pH 7.6. The protein thus purified was seen as a single band on polyacrylamide gel electrophoresis (7.5%, pH 4.3), and its absorption ratio A_{280}/A_{597} was 5.5.

Table 1. Amino acid composition of CBP

Amino acid	Residues per mole of CBP: From amino acid analysis*	From sequence determination
Asp	9.99	10
Thr	4.76	5
Ser	6.74	6
Glu	5.12	5
Pro	4.54	5
Gly	13.11	13
Ala	6.19	6
Cys†	3.10	3
Val	8.83	9
Met	1.78	2
Ile	3.82	4
Leu	3.87	4
Tyr	3.92	4
Phe	5.83	6
His	1.97	2
Lys	5.93	6
Trp‡	1.61	2
Arg	3.00	3
X§		1
Total		96

* Calculated on the basis of M_r 10,100 (12). Except where noted, values are derived from the average of duplicate 24-, 48-, and 72-hr hydrolyses of *S*-carboxymethyl-CBP in 6 M HCl at 110°C with appropriate corrections for changes associated with prolonged hydrolysis.
† Half-cystine as cysteic acid, after performic acid oxidation.
‡ Hydrolysis in 4 M methanesulfonic acid (19).
§ Unidentified residue at position 37.

Abbreviation: CBP, cucumber basic blue protein.

‡ To whom reprint requests should be addressed.

§ The basic blue protein isolated from cucumber seedlings appears to be the same protein as has been obtained from cucumber peelings (11) and several other plant sources (15, 16). The names "cusacyanin" and "plantacyanin" have been proposed (11). The name plantacyanin is undesirable due to the possibility of confusion with "plastocyanin." To avoid the further trivialization of the nomenclature of blue copper proteins, we propose the use of cucumber basic blue protein until the function or origin of the protein is established.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

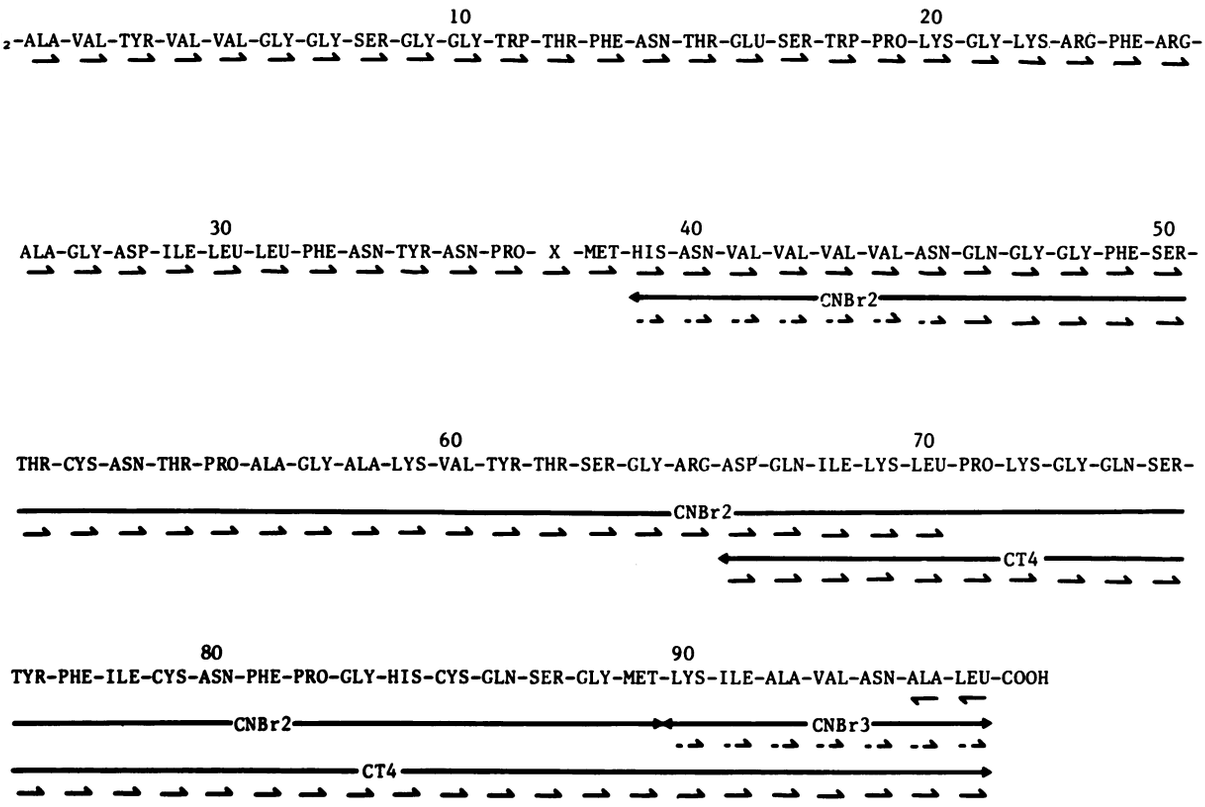

FIG. 1. Amino acid sequence of CBP from seedlings. Peptides used in the structure determination are indicated by double-headed arrows. CNBr-peptides are from CNBr cleavage; CT-peptide is derived from tryptic cleavage of citraconylated, carboxymethyl-CBP; the identifying numbers refer to the theoretical position of the peptide in the final sequence proposal and do not imply isolation of a complete set of peptides from each cleavage. Small arrows underneath intact protein or peptides indicate extent of sequence degradation; →, automated Edman degradation used in sequence proposal; - -→, automated Edman degradation not used in sequence proposal (see text); ←, carboxypeptidase A digestion. X, unidentified residue at position 37.

Performic acid oxidation, S-[^{14}C]carboxymethylation, citraconylation, cyanogen bromide cleavage, and enzymatic digestions with trypsin and carboxypeptidase A were carried out by standard methods as described (17). Acetylation of S-[^{14}C]-carboxymethyl-CBP was performed according to Matsuo and Narita (18). Peptide fragments from enzymatic digests were fractionated by reverse-phase HPLC on Lichrosorb RP-8, 10 μm (25 × 0.46 cm), in a linear water/acetonitrile gradient from 0% to 50% acetonitrile in 0.05% trifluoroacetic acid.

Amino acid analysis was routinely performed on a Dionex D-502 amino acid analyzer after hydrolysis in 6 M HCl at 110°C. For tryptophan determination, hydrolysis in 4 M methanesulfonic acid was employed (19). Half-cystine was determined as cysteic acid after performic acid oxidation of native CBP. Automated amino acid sequence analysis (20) of S-[^{14}C]carboxymethyl-CBP and its peptides was performed on a Beckman 890C sequencer with a modified controller (21) by using essentially Beckman Program no. 121078 with Polybrene as a carrier (22). Phenylthiohydantoin amino acids were identified by HPLC (23). All sequenator runs were performed at least twice. Alignments of protein sequences for purposes of comparison were made by a variant of the method of Needleman and Wunsch (24).

RESULTS AND DISCUSSION

The amino acid composition from analysis of CBP is given in Table 1. Preliminary NH_2-terminal sequence analysis showed that CBP has a single unblocked NH_2-terminal amino acid alanine and that the preparation was homogeneous by this criterion. Carboxypeptidase A digestion of the citraconylated, S-carboxymethyl-CBP gave the COOH-terminal sequence as Ala-Leu, with rapid release of four further residues, whose order could not be reliably determined, as follows: (Asn, Ala, Val, Ile)Ala-Leu-COOH.

The amino acid sequence of CBP is given in Fig. 1 and was derived by using the following strategy. Automated sequence of S-[^{14}C]carboxymethyl-CBP provided a definitive sequence to residue 50 (with the exception of residue 37). The occurrence of three arginine residues in the composition (Table 1) indicated that four peptides were likely to be generated by tryptic cleavage of the citraconylated, carboxymethylated CBP. An arginine-free peptide was isolated by HPLC from the limited tryptic digest and was considered to be the probable COOH-terminal peptide. Its composition was as follows: S-carboxymethyl-$Cys_{1.5}$, $Asp_{3.1}$, $Ser_{2.0}$, $Glu_{3.2}$, $Pro_{2.2}$, $Gly_{3.2}$, $Ala_{2.4}$, $Val_{1.3}$, $Met_{0.9}$, $Ile_{2.6}$, $Leu_{3.0}$, $Tyr_{0.9}$, $Phe_{2.1}$, $His_{1.2}$, $Lys_{3.0}$. Sequence analysis confirmed this composition (Fig. 1) together with its identification as the COOH-terminal peptide. These data provided evidence for the NH_2-terminal (residues 1–50) and COOH-terminal (residues 66–96) regions of CBP, including the positions of the two methionines (residues 38 and 89).

The remaining sequence (residues 51–65) was determined by automated sequence determination of unfractionated cyanogen bromide cleavage products of S-[^{14}C]carboxymethyl-CBP following prior acetylation to block the original NH_2 terminus. A double sequence consistent with residues 39–45 and 90–96 in the final proposal was apparent in the first seven cycles of the degradation; the remaining cycles gave an unambiguous sequence corresponding to residues 46–70 of the final proposal (Fig. 1). With the exception of residue 37, the three degrada-

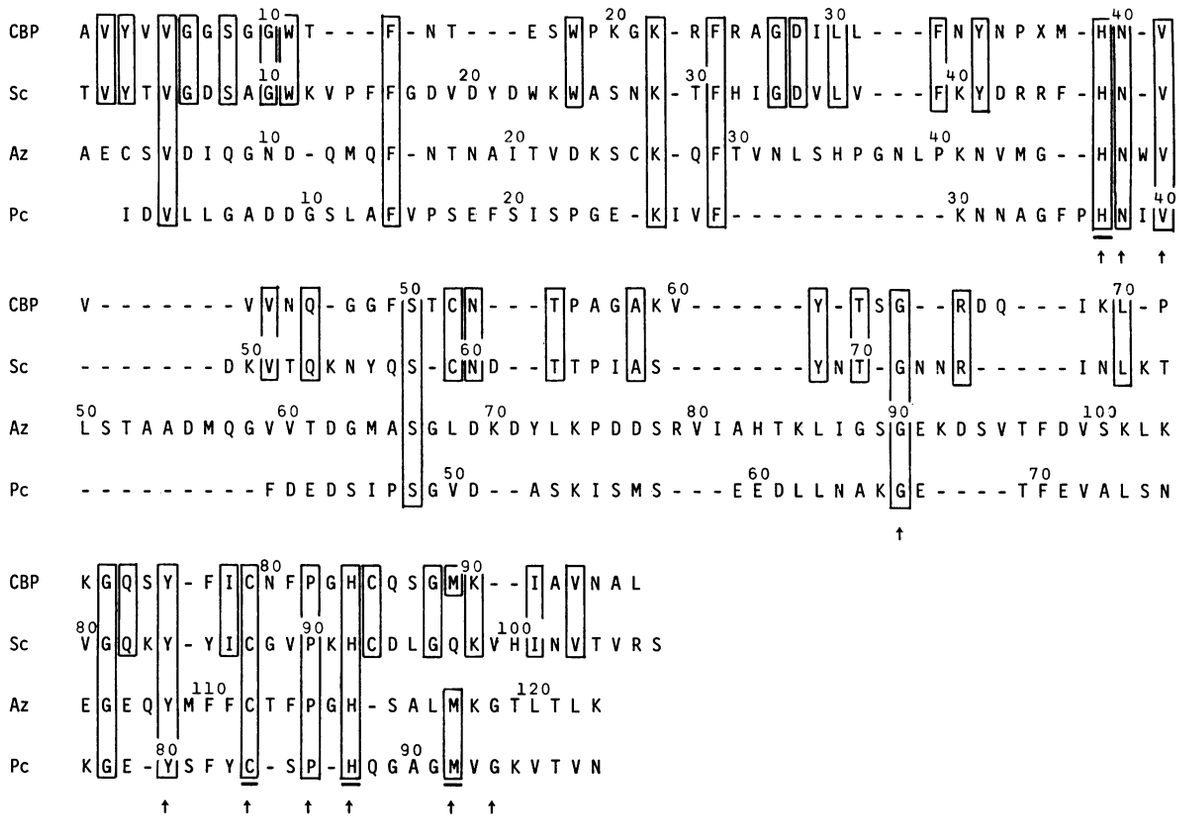

FIG. 2. Amino acid sequence comparison between CBP, stellacyanin (14), *P. aeruginosa* azurin (25), and poplar plastocyanin. Sc, stellacyanin; Az, azurin; and Pc, plastocyanin. Boxes indicate the alignment of residues in CBP with identical residues in Sc or Az and Pc, or both. Arrows signify residues that are conserved in all known complete sequences of Az and Pc (3). The four copper-binding residues in Az and Pc are underlined.

tions described enabled the entire amino acid sequence of CBP to be deduced, the overlaps being extensive enough to ensure that no small peptides had been overlooked. The composition derived from the sequence agrees well with that calculated from amino acid analysis (Table 1). No phenylthiohydantoin amino acid above background was seen at step 37 in the intact CBP on repeated degradations despite clear sequence identification at subsequent residues. This position may be occupied by a modified amino acid whose phenylthiohydantoin is unstable, and the indication of an extra serine residue on amino acid analysis (Table 1) is consistent with this explanation; further detailed structural work in this region of the molecule will be needed to clarify this point.

In Fig. 2, the sequence of CBP is aligned with those of *P. aeruginosa* azurin (25), poplar plastocyanin (R. Ambler, personal communication), and stellacyanin (14). There is remarkable similarity between the sequences of CBP and stellacyanin. In CBP, 43 of the 96 residues can be aligned with identical residues in stellacyanin. However, there are also significant differences. The CBP includes both Met and Glu residues that are absent from stellacyanin (14), whereas the sequence Asn-X-Thr that is the carbohydrate attachment site at three places in stellacyanin is absent from the CBP sequence.

Important sequence homology is evident even when azurin and plastocyanin are included in the comparison. Of 10 residues that are invariant in all known sequences of azurin and plastocyanin (3), 7 can be aligned with identical residues in CBP and stellacyanin. In particular, three of the copper ligands in plastocyanin (and in azurin; in parentheses)—His-37 (46), Cys-84 (112), and His-87 (117)—are found in both CBP and stellacyanin. Thus, it seems likely that three of the copper ligands in all four proteins are identical.

The fourth copper ligand in plastocyanin (azurin; in parentheses) is an invariant Met-92 (121). A corresponding residue, Met-89, occurs in CBP, but as indicated previously, no Met is found in stellacyanin. The nature of the fourth copper ligand in stellacyanin has been the subject of intense speculation and some experimentation (e.g., refs. 26 and 27). Residue Cys-93 of stellacyanin has been implicated, either as a thiolate ligand (26) or as part of a copper-binding cystine disulfide bridge (27). It is somewhat tantalizing that CBP not only has Met-89 at a position corresponding to the copper-binding Met residue in plastocyanin and azurin but also has Cys residues at positions 52 and 85, precisely aligned with the putative copper-binding Cys residue(s) in stellacyanin. Because CBP bears a closer spectroscopic relationship to stellacyanin than to plastocyanin or azurin, it will be interesting to see whether it belongs structurally to the plastocyanin/azurin (fourth ligand = Met) class or to the stellacyanin (fourth ligand ≠ Met) class. The impending structure analysis should resolve this question.

The authors thank Dr. J. M. Guss for helpful discussions and for the computer analyses used in the sequence alignment. This work was supported by grants from the Australian Research Grants Committee (C80/15377) and Esso Australia Ltd. (to H.C.F.) and from the National Health and Medical Research Council of Australia, the Ian Potter Foundation, and Felton Bequest (to F.J.M.).

1. Adman, E. T., Stenkamp, R. E., Sieker, L. C. & Jensen, L. H. (1978) *J. Mol. Biol.* **123**, 35–47.
2. Colman, P. M., Freeman, H. C., Guss, J. M., Murata, M., Norris, V. A., Ramshaw, J. A. M. & Venkatappa, M. P. (1978) *Nature (London)* **272**, 319–324.
3. Schwartz, R. M. & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, pp. 50–51.

4. Omura, T. (1961) *J. Biochem (Tokyo)* **50,** 394–399.
5. Peisach, J., Levine, W. G. & Blumberg, W. E. (1967) *J. Biol. Chem.* **242,** 2847–2858.
6. Malmstrom, B. G., Reinhammar, B. & Vanngard, T. (1970) *Biochim. Biophys. Acta* **205,** 48–57.
7. Stigbrand, T. (1971) *Biochim. Biophys. Acta* **236,** 246–252.
8. Marchesini, A., Minelli, M., Merkle, H. & Kroneck, P. M. H. (1979) *Eur. J. Biochem.* **101,** 77–84.
9. Morita, Y., Wadano, A. & Icla, S. (1971) *Agric. Biol. Chem.* **35,** 255–260.
10. Shichi, H. & Hackett, D. P. (1963) *Arch. Biochem. Biophys.* **100,** 185–191.
11. Markossian, K. A., Aikazyan, V. T. & Nalbandyan, R. M. (1974) *Biochim. Biophys. Acta* **359,** 47–54.
12. Vickery, L. E. (1971) Dissertation (Univ. of California, Santa Barbara).
13. Colman, P. M., Freeman, H. C., Guss, J. M., Murata, M., Norris, V. A., Ramshaw, J. A. M., Venkatappa, M. P. & Vickery, L. E. (1977) *J. Mol. Biol.* **112,** 649–650.
14. Bergman, C., Gandvik, E., Nyman, P. O. & Strid, L. (1977) *Biochem. Biophys. Res. Commun.* **77,** 1052–1059.
15. Aikazyan, V. T. & Nalbandyan, R. M. (1975) *FEBS Lett.* **55,** 272–274.
16. Aikazyan, V. T. & Nalbandyan, R. M. (1981) *Biochim. Biophys. Acta* **667,** 421–432.
17. Begg, G. S., Pepper, D. S., Chesterman, C. N. & Morgan, F. J. (1978) *Biochemistry* **17,** 1739–1744.
18. Matsuo, H. & Narita, K. (1975) in *Protein Sequence Determination,* ed. Needleman, S. B. (Springer, Berlin), pp. 104–113.
19. Simpson, R. J., Neuberger, M. R. & Liu, T.-Y. (1976) *J. Biol. Chem.* **251,** 1936–1940.
20. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1,** 80–91.
21. Begg, G. S. & Morgan, F. J. (1982) *Anal. Biochem.* **119,** 12–16.
22. Klapper, D. G., Wilde, C. E. & Capra, J. D. (1978) *Anal. Biochem.* **85,** 125–131.
23. Zimmerman, C. L., Apella, E. & Pisano, J. J. (1977) *Anal. Biochem.* **77,** 569–573.
24. Needleman, S. B. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453.
25. Ambler, R. P. (1971) in *Recent Developments in the Chemical Studies of Protein Structures,* eds. Previero, A., Pechere, J. F. & Coletti-Previero, M. A. (Inst. Natl. Santé Rech. Med., Paris), pp. 289–305.
26. Hill, H. A. O. & Lee, W. K. (1979) *J. Inorg. Biochem.* **11,** 101–113.
27. McMillin, D. R. & Morris, M. C. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6567–6570.