A C-TERMINAL HAEM-BOUND CYTOCHROME *c*-556 FROM *AGROBACTERIUM TUMEFACIENS*, STRAIN B_2a

Amino acid composition and N-terminal sequence analysis

J. VAN BEEUMEN, C. VAN DER BRANDEN, P. TEMPST and J. DE LEY
Laboratorium voor Microbiologie en microbiële Genetica, Rijksuniversiteit Gent, K.L. Ledeganckstraat 35, 9000 Gent, Belgium

Received 13 March 1979

1. Introduction

Comparison of amino acid sequences of soluble cytochromes *c* is a successful approach to unravel eukaryotic phylogeny [1]. For the prokaryotes, the same approach has, until now, almost exclusively been applied to the cytochromes of some phototrophic bacteria, of *Paracoccus denitrificans*, of some *Desulfovibrio* and *Pseudomonas* strains and of *Azotobacter vinelandii* [1]. Cytochromes of aerobically grown chemo-organotrophes have not been investigated [1–3]. Because of its peculiar nature, we report here on the N-terminal sequence of cytochrome *c*-556 from *Agrobacterium tumefaciens*, strain B_2a. It is a soluble protein with mol. wt ~12 000. Its purification and properties have been described [4]. In addition to our recent findings on the sequence of the haem-binding region of the protein [5], we conclude that B_2a *c*-556 is a C-terminal haem-bound cytochrome.

2. Materials and methods

The preparation of B_2a *c*-556 was electrophoretically pure and had a purity index of 1.47 [4]. Amino acid analyses were carried out with a Biocal 200 analyser after hydrolysis of the native protein with 6 N HCl at 105°C in sealed, evacuated ampoules. The N-terminal sequence of the protein was determined in a Socosi PS 110 sequencer using the 1 N quadrol programme in [6]. The thiazolinones were converted with 0.2 ml 1 N HCl at 80°C. After 3 extractions with 0.7 ml ethylacetate, the PTH-amino acids in the organic phase were identified on Kieselgel thin-layer plates (Merch 5715, 20 × 20 cm). One-half of the aqueous phase was assayed for PTH-arginine [7], the other half for PTH-histidine [8], both on paper. Three tryptic peptides which cover the N-terminal region of the cytochrome were obtained by digestion of 2 μmol dehaemed protein. Their sequence was determined by manual dansyl-Edman degradation.

3. Results and discussion

The amino acid composition of B_2a *c*-556 is in table 1. Since –Cys–X–Y–Cys–His– is the haem-binding sequence in most cytochromes *c* (except for –Ala–X–Y–Cys–His– in two eukaryotic cytochromes [1]), the occurrence of 2 cysteines and 1 histidine indicates that B_2a *c*-556 is a monohaem cytochrome. The mol. wt 12 540 calculated from the presumed composition is in good agreement with the value obtained by SDS–gel electrophoresis [4].

Figure 1 shows the N-terminal sequence of B_2a *c*-556. The first 16 residues were identified with the native protein in the sequenator. No further residues could be identified, probably because of a nearly complete cyclisation of Gln-13 to pyrrolidone–car-

Abbreviations: PTH, phenyl thiohydantoin amino acids; B_2a *c*-556, cytochrome *c*-556 from *A. tumefaciens*, strain B_2a

Table 1
Amino acid composition of *A. tumefaciens* B_2a *c*-556

Amino acid	Residues/mol protein after hydrolysis at		Nearest integer
	24 h	96 h	
Asp	9.8	8.6	10
Thr[a]	9.8	9.2	10
Ser[a]	2.9	2.9	3
Glu	12.0	12.4	12
Pro	4.5	4.5	5
Gly	13.3	12.9	13
Ala	20.8	20.9	21
Val	4.4	4.5	4
Met	1.8	3.2	3
Ile	4.1	4.1	4
Leu	6.0	5.8	6
Phe	5.0	4.9	5
Lys	10.8	12.0	12
His	0.7	1.1	1
Arg	2.0	2.0	2
Tyr	0.7	1.0	1
Cys[b]	2	2	2
Trp			1[c]

[a] Corrected values for decomposition during acid hydrolysis
[b] Determined as cysteic acid
[c] Estimated from the purity of the native cytochrome

boxylic acid; the repetitive yield for the PTH of Glu-8 and Gly-9 was found to be as high as 95%. Sequence information beyond residue 16 could nevertheless be obtained by analysing those peptides of a tryptic protein digest which fit the N-terminal sequence found with the sequenator. The peptides T3e and T3d3 (fig.1), constitute the region 1 → 12 of the protein. T2b3 is the only other tryptic peptide with N-terminal Glu or Gln (determined as dansyl-Glu). Its subsequent sequence –Ile–Glu–Gly– is identical with the last 3 amino acids determined with the sequenator. T2b3 is thus the neighbouring peptide of T3d3, covering the region 13 → 26.

It is clear from our data that no cysteines or histidine occur within the 26 N-terminal positions of B_2a *c*-556. In most N-terminal haem-bound cytochromes (they are grouped under class I according to the classification in [2]) the first cysteine of the invariant haem-binding sequence –Cys–X–Y–Cys–His– is at or before position 14 [1]. The most remote position in eukaryotic cytochromes of class I is at position 25 (*Tetrahymena pyriformis* [9]), in prokaryotic cytochromes of the same class it is at position 23 (*Prosthecochloris aestuarii* [10,11]). We therefore conclude that B_2a *c*-556 is not an N-terminal haem-bound cytochrome.

Within the range of the presently studied structures, the only other alternative is that B_2a *c*-556 is a member of class II (A classification [2]). This class contains high- (cytochromes *c*′) and low-spin cytochromes with the haem group bound at the C-terminal part of the polypeptide chain. Our conclusion is strengthened by the sequence homology of the haem region of B_2a *c*-556 with the C-terminal region of the following class II cytochromes examined: cytochrome *c*-556 (low-spin) from *Rhodopseudomonas palustris*; the cytochromes *c*′ from *Rps. capsulata, Rps. palustris, Chromatium vinosum, Rhodospirillum rubrum* from one *Alcaligenes* strain, NCIB 11015 (fig.2b). The primary structure of the latter 2 cytochromes *c*′ are the only ones completely elucidated so far. The similarity between the haem regions from the latter 2 proteins and B_2a *c*-556, reported [5], was 35%. As a result of the present work, we also aligned the N-terminal region of the *Agrobacterium* cytochrome with the corresponding region of both cytochromes *c*′ (fig.2a). The similarity is clearly lower than at the C-terminus: it is 27% with *Rhs. rubrum* cytochrome *c*′ and 15% with the *Alcaligenes* protein; it is also 27% between the 2 cytochromes *c*′. Three residues seem to be evolutionary invariant: Val-4, Arg-7 and Lys-26. All these data support the conclusion that

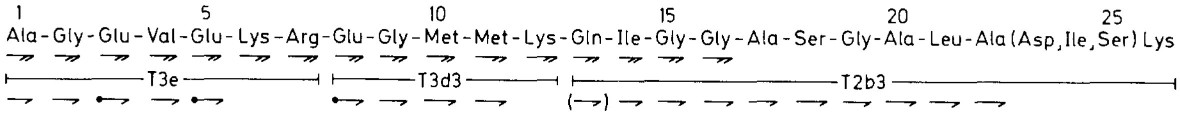

Fig.1. N-terminal amino acid sequence of *A. tumefaciens* B_2a *c*-556. T3e, T3d3 and T2b3: tryptic peptides. (—⇁) Residues identified with the sequenator; (—⇁) residues identified as the dansyl amino acid; (•—⇁) residues identified as the dansyl derivate of the acid but final acid assignment deduced from the mobility of the peptide; ((—⇁)) acid or amide which cannot be deduced from the mobility alone.

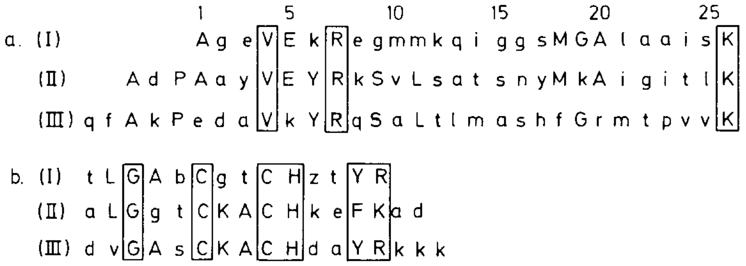

Fig.2. Comparison of the N-terminus (a) and the C-terminal haem region (b) of *A. tumefaciens*, cytochrome *c*-556 (I) and the cytochromes *c*′ from *Rhs. rubrum* (II) and *Alcaligenes* NCIB 11015 (III). The one-letter code follows the recommendations in [14]. Capital letters are used when a residue in any one sequence is identical to a residue in another sequence. Boxed residues, including the substitution Tyr ↔ Phe and Arg ↔ Lys at the C-terminus, are common to the 3 cytochromes. The boxed residues in (b) are also common to the 4 other cytochromes of class II mentioned in the text.

A. tumefaciens B_2a *c*-556 is a class II cytochrome *c*.

As a point of structural and phylogenetic interest, it is remarkable that the first cytochrome *c* studied from an aerobically grown chemo-organotrophe belongs to a cytochrome class of which the members, with exception of the cytochromes *c*′ from *Azotobacter vinelandii* [12], a halotolerant 'micrococcus' [3,13] and one strain of *Alcaligenes*, are found exclusively in photosynthetic bacteria. The complete sequence determination of B_2a *c*-556 is in progress.

Acknowledgements

J.D.L. is indebted to the Fonds voor Kollektief Fundamenteel Onderzoek for research and personnel grants. P.T. is aspirant of the Nationaal Fonds voor Wetenschappelijk Onderzoek.

References

[1] Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure, vol. 5; (1975) suppl. I; (1976) II. Natl. Biomed. Res. Found., Washington, DC.

[2] Ambler, R. P. (1979) in: Electron Transport in Microorganisms (Senez, J. et al., eds) CNRS Int. Symp., Marseille, 1976, in press.

[3] Ambler, R. P. (1978) in: Evolution of Protein Molecules (Matsubara, H. and Yamanaka, T. eds) pp. 311–322, Japan Sci. Soc. Press, Tokyo.

[4] Van den Branden, C., Van Beeumen, J. and De Ley, J. (1975) Hoppe Seyler's Z. Physiol. Chem. 356, 1251–1258.

[5] Tempst, P., François, G. and Van Beeumen, J. (1979) Arch. Int. Physiol. Biochim. 87, in press.

[6] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.

[7] Yamado, S. and Itano, H. A. (1966) Biochim. Biophys. Acta 130, 538–540.

[8] Pauly, H. (1904) Z. Phys. Chemie 42, 508–518.

[9] Tarr, G. E. and Fitch, W. M. (1976) Biochem. J. 159, 193–199.

[10] Van Beeumen, J., Ambler, R. P., Meyer, T. E., Kamen, M. D., Olson, J. M. and Shaw, E. K. (1976) Biochem. J. 159, 757–774.

[11] Olson, J. M. (1978) Int. J. Syst. Bacteriol. 28, 128–129.

[12] Yamanaka, T. and Imai, S. (1972) Biochem. Biophys. Res. Commun. 46, 150–154.

[13] Kono, M. and Taniguchi, S. (1960) Biochim. Biophys. Acta 43, 419–430.

[14] IUPAC–IUB, Tentative Rules (1968) J. Biol. Chem. 243, 3557–3559.