NEUROSPORA CRASSA AND *HUMICOLA LANUGINOSA* CYTOCHROMES C : MORE HOMOLOGY IN THE HEME REGION

Florence Lederer and Anne Marie Simon

Centre de Génétique Moléculaire du C.N.R.S.
91 - Gif-sur-Yvette, France

Received October 23,1973

Summary : *Neurospora crassa* and *Humicola lanuginosa* cytochromes c were submitted to an automatic Edman degradation. It was found that residue 16 is a glutamine, as we had predicted (1) and not a glutamic acid, as published for both proteins (4,7). Moreover, residues 19 to 26 were found to have been placed in a wrong order in both cases. The corrected order shows more homology with other cytochromes c in this area.

Introduction

We recently reported an error in the published sequence of iso-1 cytochrome c from *Saccharomyces cerevisiae* (1) : residues 15 and 16 were found to be -Leu-Gln- and not -Glu-Leu- as previously determined (2,3). After this correction, it appeared that out of about fourty sequences published to that date for naturally occurring eukaryotic cytochromes c, only two had a glutamic acid at position 16, and not a glutamine, as all others ; they were those of *Candida krusei* and *Neurospora crassa* (4,5). Considering the well known lability of glutaminyl residues, we suggested that those two proteins might actually also have a glutamine at position 16.

Subsequent experimentation showed the correctness of our prediction concerning *Candida krusei* cytochrome c (6). We report here the results of a reinvestigation of the N-terminal sequence in *Neurospora crassa* and *Humicola lanuginosa* cytochromes c. The sequence of the latter was published after our first report with a glutamic acid at position 16 (7). In both cases, we were led to a more important correction than expected.

Materials and Methods

The N-terminal sequences were determined by the automatic degradation method (8) with a Socosi sequenator PS 100 (94-Saint Maur, France). Solvents and reagents were purified according to Edman and Begg (8) except for quadrol, which came from Pierce. The phenylhydantoins were identified by thin layer chromatography and occasionally by amino acid analysis after rege-

neration, as previously reported (1,6). Protein hydrolyses and amino acid analyses were carried out as described elsewhere (6).

Humicola lanuginosa cytochrome c was a gift from Dr. W.T. Morgan.

Neurospora crassa cytochrome c was purified in our laboratory from a partly purified lyophilized extract kindly prepared for us in Dr. J.R.S. Fincham's laboratory. The extract had been obtained as follows. Freeze-dried mycelium (grown for 24 h from a heavy inoculum in 1/2 l lots of modified Volgel's medium in shake flasks) was powdered and extracted with 50 mM sodium phosphate buffer, 1 mM EDTA pH 7.4 (about 4 l per kg damp weight of mycelium). The insoluble cell wall material was spun down and successive precipitates with ammonium sulfate were removed by centrifuging until 55 % saturation. The 55 % supernatant was almost saturated with ammonium sulfate, the precipitate spun down, redissolved in 0.1M ammonium bicarbonate and dialysed against the same buffer. On one or two occasions this solution was further purified by a heat step (about 5 to 10 mn at 50°) which enabled more non cytochrome protein to be spun out. The cytochrome was then adsorbed on a column of CM-cellulose equilibrated with 0.1 M ammonium bicarbonate, eluted with 1 M ammonium bicarbonate and freeze-dried.

We received three such extracts containing altogether about 5 µmoles of cytochrome c in about 1 g total protein. Part of the contaminating proteins was removed by precipitation with ammonium sulfate (65 % of saturation in about 20 ml of 0.1 M sodium phosphate buffer, pH 7). The supernatant was diluted 5 fold and adsorbed on about 10 g (wet weight) of Amberlite CG 50 by dialysing against 3 times 6 l of 5 mM sodium phosphate buffer, 0.1 mM ferricyanide, pH 7. The resin with adsorbed cytochrome c was then directly layered on top of an Amberlite CG 50 column (see Results for further details).

Results

1 - Purification of *Neurospora crassa* cytochrome c

Figure 1 shows the elution profile of the protein from an Amberlite column. The first shoulder was reduced cytochrome, which migrated ahead of the main oxidized peak, as always. The amount of cytochrome c in pooled fractions 60 to 90 was 4 µmoles ; the absorbance ratio of reduced band to oxidized form at 280 nm, which was 1.36, indicated a high degree of purity (4). Amino acid analyses, reproduced in Table 1, showed good agreement with published values (4,10). Alongside is shown the composition of tube 96. It indicates that the shoulder observed after the main peak represents an unmethylated cytochrome c, the existence of which was first described by Scott and Mitchell (11).

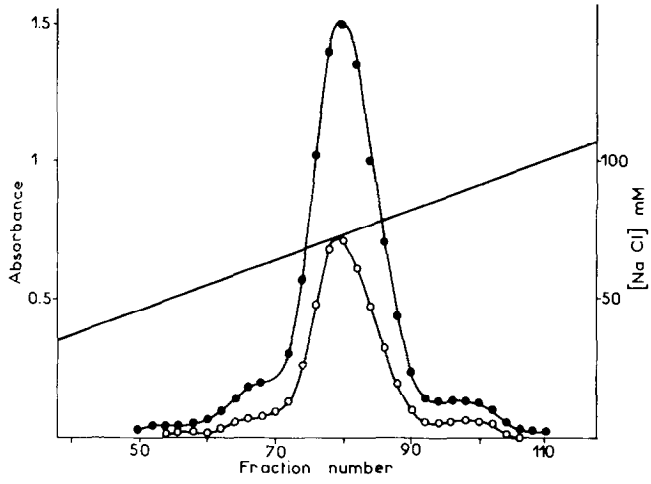

Figure 1. Purification of *Neurospora* cytochrome c.
The column (2.5 x 30 cm) was equilibrated with 0.1 M sodium phosphate buffer pH 7 at room temperature. The elution was carried out at a flow rate of 21 ml/h with a linear gradient between 0 and 0.25 M NaCl in 0.1 M phosphate buffer pH 7 (total volume of the gradient 1200 ml). Fractions of 4.2 ml were collected. (●—● —●) Absorbance at 280 nm ; (○—○—) absorbance at 530 nm.

Table 1. Composition of the purified cytochrome from *Neurospora*.

The results are expressed as number of residues per 7 leucines. The values are the average of 2 analyses of a 24 h hydrolysate, except the figure for εN-Me3-lysine (a single analysis).

	Theory (4)	Experimental	
		Major peak	Tube 96
Lysine	13	13.3	14.5
Histidine	2	1.7	2.0
Arginine	3	2.9	3.5
εN-Me3-lysine	1	0.8	0
Aspartic acid	13	12.2	12.5
Threonine	9	8.4	8.3
Serine	3	3.7	3.9
Glutamic acid	8	8.0	8.5
Proline	3	3.3	3.6
Glycine	15	14.1	13.8
Alanine	9	9.0	9.1
Valine	1	1.0	1.3
Methionine	2	1.4	1.5
Isoleucine	5	3.5	3.3
Leucine	7	7	7
Tyrosine	4	3.8	3.4
Phenylalanine	6	5.9	5.5

2 - N-terminal sequence of *Neurospora crassa* cytochrome c

Edman degradation of 0.32 µmole of cytochrome in the automatic sequenator yielded the N-terminal sequence up to the 43rd residue. The results are given in Table 2. The residues in italics in the top line are those

Table 2. N-terminal sequence of *Neurospora crassa* cytochrome c.

H_2N-Gly-Phe-Ser-Ala-Gly-Asp-Ser-Lys-Lys-Gly-Ala-Asn-Leu-Phe-Lys-Thr-Arg-
-4 -1 1 5 10

Cys *Glu* *Cys* *Gly* *Glu* *Gly* *Gly* *Asn* *Leu* *Thr* *Gln* *Ile*
-Ala-Gln- -His-Thr-Leu-Glu-Glu-Gly-Gly-Gly-Asn-Lys-Ile/Leu-Gly-Pro-
15 20 25 30

Leu *Leu* *Arg*
Ala-Ile/Leu-His-Gly-Ile/Leu-Phe-Gly- -Lys-
39

found by Heller and Smith (4) when they differ from our results or when we have a hole. For residues 28, 32 and 35 no distinction could be made between leucine and isoleucine as not enough material was left for amino acid analyses. As the heme had not been removed for the experiment, the half-cystines 14 and 17 could not have been identified. It is not surprising that residue 38 failed to give a positive reaction with the phenanthrene quinone test : PTH-arginine and -histidine are always difficult to detect at the end of a sequence because of low yields in the conversion of thiazolinones.

3 - N-terminal sequence of *Humicola lanuginosa* cytochrome c

Table 3 shows the order of the first 37 amino acids of *Humicola* cytochrome c obtained in an automatic degradation experiment, carried out with the 4 mg sample kindly sent to us by Dr. W.T. Morgan. The lop line indicates in italics the residues of the published sequence at the corresponding positions (7). In this experiment, no attempt was made to identify the expected arginine and histidine residues at positions 13 and 18 respectively. Half-cystines 14 and 17 were not identified for the reason stated in paragraph 2. Our sequence has holes at positions 19 and 26.

Table 3. N-terminal sequence of *Humicola lanuginosa* cytochrome c.

Ser Ala

H_2N-Ala-Lys-Gly-Gly-Ser-Phe-Glu-Pro-Gly-Asp-Ala-Ser-Lys-Gly-Ala-Asn-Leu-Phe-Lys-

-8 -1 1 5 10

Arg Cys *Glu Cys His Gly Glu Gly Ala Asn Val Ser Gln*

Thr- - -Ala-Gln- - - -Val-Glu-Gln-Gly-Gly-Ala- -Lys-Ile-Gly-

15 20 25

Discussion

It can be seen that, besides the changes we expected to see at position 16, the partial sequences we determined differ from the published ones in the region 19-26. For *Neurospora crassa*, the amino acids correspond to those found in the same area by Heller and Smith (4) but they are in a different order. For *Humicola lanuginosa*, the same holds true, except that a serine and an asparagine residue remain to be placed ; in addition, we found an inverted order at positions 3-4.

The work of Heller and Smith (4) rested on the sequence determination of the peptide obtained by a tryptic split of the chymotryptic hemopeptide. The authors deduced its C-terminal sequence from a carboxypeptidase digestion. Actually, when plotted on a graph, the results of the digestion appear totally uninterpretable. Moreover, assuming some contamination and a wrong interpretation of an ambiguous stoichiometry, the rest of the results published by the authors is perfectly compatible with our sequence, except for the N-terminal of a papain peptide which we assume to have been Thr-Leu-Glu (the authors reported PTH-Leu as N-terminal). On the other hand, inspection of the data or Morgan et al. (7) does not reveal obvious weak points in their determination. However, although we had just enough material for one sequenator run in the case of *Humicola* cytochrome c, we believe that generally speaking automatic sequences are less liable to error than manual ones. Our confidence in our results in strengthened by homology considerations which are exposed in the following paragraph.

Sequences in the heme region published to this date for fungal cytochromes c are listed in Table 4. One can see a general pattern to which our corrected sequences conform remarkably well. This pattern is even apparent when one looks at over fifty sequences of eukaryotic cytochromes c published

Table 4. Sequences of fungal cytochromes c in the heme region.

Candida (5,6)	-Cys-	Ala-	Gln-	Cys-	His-	Thr-	Ile-	Glu-	Ala-	Gly-	Gly-	Pro-	His-
Debaryomyces (17)		Leu					Val	Glu					
Saccharomyces iso-1 (1,3)								Lys					
Saccharomyces iso-2 (1,18)		Gln					Ile[b]	Glu				Asn	
Ustilago (19)		Ala					Leu	Gly	Ala		Glu		
Neurospora (4,a)								Glu	Glu		Gly	Gly	
Humicola (7,a)						Ser?Val			Gln			Ala	Asn?

a : this work
b : in reference 1, a valine is printed at this position. This comes from a heretofore unnoticed typing error. We actually found an isoleucine at this position, in agreement (as stated in (1)) with work of Stewart et al. (18).

to date (33 given in (12) and the rest published since then). For example, one notes only two exceptions to the fact that residues 19 and 23 are generally threonine and glycine respectively : serine was found instead of threonine in *Euglena gracilis* cytochrome c 558 (13,14) and alanine instead of glycine in the Pacific lamprey (15). Furthermore, with few exceptions, residue 20 is generally hydrophobic, residue 21 an acidic one ; position 24 is alanine in plants and glycine in all other sequences. The only residue which seems to be highly variable as to the nature of the side chain is found at position 22.

Thus, the proposed order for *Neurospora crassa* and *Humicola lanuginosa* is much more satisfying. In the case of the latter, on the basis of homology, if one assumes that the total composition of the hemopeptide, as published by Morgan et al. (7), is correct, one can predict that the residues we could not identify are serine and asparagine at positions 19 and 26 respectively. It is reasonable that we could not identify precisely those residues because, in our hands at least, PTH-serine and -asparagine tend to give weaker spots than most other residues.

Neurospora crassa cytochrome c has been included in philogenetic trees built by various authors in terms of evolutionary distance between cytochromes c (12, 16). It is clear that the position of the branching point which leads to *Neurospora crassa* will have to be altered since the number of differences with bakers'yeast and human cytochromes c, for example, is now reduced to 33 and 39 respectively, instead of 38 and 44 (12). Finally, glutamine 16 has now to be put on the list of invariant residues. As the

number of known sequences increases, it may not remain there, especially since the work of Sherman et al. (20,21) on yeast mutants showed that the position can be occupied by several other amino acids without detriment to the function.

Acknowledgements : Our deepest thanks are due to Dr. W.T. Morgan, and to Dr. J.R.S. Fincham and his collaborators, whose cooperation provided us with the indispensable starting materials. We also want to thank Dr. R.P. Ambler for the interest taken in this work.
This investigation was supported by grant D.G.R.S.T. n° 72 7 0500.

References

1. Lederer, F., Simon, A.M. and Verdière, J. (1972). Biochem. Biophys. Res. Commun., 47, 55-58.
2. Narita, K., Titani, K., Yaoi, Y. and Murakami, H. (1963). Biochim. Biophys. Acta, 77, 688-690.
3. Narita, K. and Titani, K. (1969). J. Biochem. (Tokyo), 65, 259-267.
4. Heller, J. and Smith, E.L. (1966). J. Biol. Chem., 241, 3165-3180.
5. Narita, K. and Titani, K. (1965). Proc. Jap. Acad., 41, 831-836.
6. Lederer, F. (1972). Eur. J. Biochem., 31, 144-147.
7. Morgan, W.T., Hensley, C.P. Jr., and Riehm, J.P. (1972). J. Biol.Chem., 247, 6555-6565.
8. Edman, P. and Begg, G. (1967). Eur. J. Biochem., 1, 80-91.
9. Heller, J. and Smith, E.L. (1966). J. Biol. Chem., 241, 3158-3164.
10. DeLange, R.J., Glazer, A.N. and Smith, E.L. (1969). J. Biol. Chem., 244, 1385-1388.
11. Scott, W.A. and Mitchell, K.H. (1969). Biochemistry, 8, 4282-4289.
12. Atlas of Protein Sequence and Structure, Dayhoff, M.O., Ed. (1972). 5.
13. Pettigrew, G.W. (1973). Nature, 241, 531-533.
14. Lin, D.K., Niece, R.L., Fitch, W.M. (1973). Nature, 241, 533-535.
15. Uzzell, T., Nolan, C., Fitch, W.M. and Margoliash, E., quoted in 12.
16. Fitch, W.M. and Margoliash, E. (1967). Science, 155, 279-284.
17. Sugeno, K., Narita, K. and Titani, K.(1971). J. Biochem. (Tokyo),70, 659-682.
18. Stewart, J.W., Margoliash, E. and Sherman, F. (1966). Fed. Proc., 25,2587.
19. Bitar, K.G. and Vinogradov, S.N., Nolan, C., Weiss, L.J. and Margoliash,E. (1972). Biochem. J., 129, 561-569.
20. Sherman, F. and Stewart, J.W. (1971). Annu. Rev. Genet. 5, 257-296.
21. Gilmore, R.A., Stewart, J.W. and Sherman, F. (1971). J. Mol. Biol. 61, 157-173.