THE AMINO ACID SEQUENCE OF CYTOCHROME *c* FROM *SOLANUM TUBEROSUM* (POTATO)

Gérard MARTINEZ and Hervé ROCHAT
Laboratoire de Biochimie, Faculté de Médecine, Secteur Nord, Marseille, France

and

Gaston DUCET
Laboratoire de Physiologie Cellulaire, Centre Universitaire de Luminy, Marseille, France

Received 1 August 1974

1. Introduction

The primary structure of potato (*Solanum tuberosum*) cytochrome *c* has been determined using 1.5 μmole of protein. When compared to the sequence of tomato (*Lycopersicum esculentum*) cytochrome *c* which had been determined by R. Scogin et al. [1], only five differences are observed in the positions 4, 29, 58, 66 and 98. These two proteins belong to two genus of the same family (*Solanaceae*). The homology is strong with other higher plants cytochromes *c* yet described.

2. Materials and methods

The cytochrome *c* was isolated from potato tubercles. As its concentration was found to be low (0.2 mg/kg), the classical method of extraction using seeds, which was described by M. Richardson et al. [2] could not be used. The method for isolating cytochrome *c* from purified mitochondria has been described elsewhere [3]. Its intensive use has given a preparation showing two slight contaminants when submitted to polyacrylamide gel electrophoresis. This preparation was applied for ion-exchange chromatography on a CM_{52} cellulose (Whatman – W. & R. Balston Ltd. England) column, equilibrated with a 15 mM sodium phosphate buffer, pH 7.20, and the proteins were eluted with a linear gradient of molarity up to 200 mM.

The micromethods used for sequence determination have been described previously by E. W. Thompson et al. [4].

After chromatography, the pure protein thus obtained, was denaturated by heating at 100°C for 15 min. TPCK–Trypsin (2 × crist.), α-chymotrypsin, carboxypeptidases A and B (DFP treated), were obtained from Worthington Biochem. Corp., Freehold, N.J., USA. The denaturated cytochrome *c* (0.5 μmole) was submitted to tryptic digestion (5% w/w enzyme/substrate) using a Radiometer pH-Stat (pH 8.0, at 37°C, for 3 hr, under nitrogen). Another 0.5 μmole of cytochrome *c* was digested by α-chymotrypsin, in the same experimental conditions, but during 4 hr. Hydrolyses by carboxypeptidases A or B were performed using a 0.2 M $(NH_4)HCO_3$ buffer, pH 8.6, at 37°C. Enzymatic digests were centrifuged (6000 *g*; 1 hr) and the precipitates discarded. The peptides were purified using horizontal high voltage electrophoresis (42 V/cm; 3 hr), on Whatman 3 MM paper, in a pyridine–acetic acid–water (25 v/1v/225 v) buffer, pH 6.5. This technique allows also the determination of electrophoretic mobilities [5]. When further purification was necessary, a second electrophoresis was performed using the same conditions but at pH 1.9 (acetic acid–formic acid–water 4v/1v/45v). Protein and peptides were hydrolyzed in 6 N HCl for 20 hr in evacuated tubes. Amino acids were analyzed using either a Technicon Amino Acid Analyzer (for ε,*N*-trimethyl-lysine estimation according to R. J. Delange [6]) or a Spinco Automatic Amino Acid Analyzer, model 120 C (for

North-Holland Publishing Company – Amsterdam

routine peptides investigation). Sequences of isolated peptides were established using the so-called 'Dansyl–Edman' method of Gray and Hartley [7]. Dns–amino acids were identified by chromatography on polyamide sheets [8] obtained from Cheng Chin Trading Co. Ltd., Taipei, Taiwan.

3. Results

The chromatography of the preparation of CM cellulose gave two fractions 1 and 2 (fig. 1). These two fractions, when submitted to polyacrylamide disc electrophoresis, showed only one single band with both the same migration. Their purity ratio (estimated as E_{410}/E_{280}) were for both 4.05, whereas it was only 3.20 for the preparation before chromatography. Moreover the amino acid compositions of the two fractions were identical. Consequently they were bulked and used for sequence determination. The discrimination between these two peaks may be explained by a possible deamidation of the native protein as suggested by Flatmark [9]. The cytochrome–amino acid composition (table 1) strongly suggests a high homology in comparison with the tomato cytochrome *c*. The tryptic and chymotryptic hydrolyses gave 14 and 16 peptides respectively (table 2 and 3). Generally, we obtained the same

Table 1
Amino acid composition (residues/molecule)

	Potato cytochrome *c*		Tomato cytochrome *c*	
	Hydrolysis	Sequence	Hydrolysis	Sequence
Try	*	1	*	1
Tml**	1.86	2	1.29	2
Lys	11.00	11	10.35	11
His	2.00	2	1.68	2
Arg	2.20	2	2.22	2
Asp	11.78	11	11.84	11
Thr	7.60	8	8.02	7
Ser	4.15	4	5.01	4
Glu	8.93	9	10.13	10
Pro	8.85	8	7.84	8
Gly	12.99	13	12.02	12
Ala	12.20	12	13.31	13
Cys	0.64	2	2.50	2
Val	3.41	3	4.02	3
Met	1.95	2	2.18	2
Ile	3.09	3	3.68	3
Leu	8.04	8	8.00	8
Tyr	5.72	6	5.00	6
Phe	4.07	4	3.85	4
		111		111

* Not determined
** Tml = ϵ,N–trimethyl-lysine

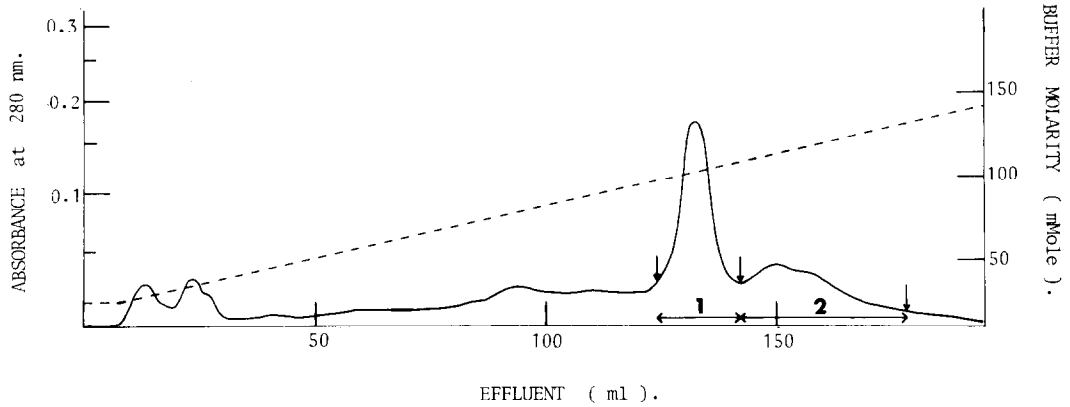

Fig. 1. Pattern of purification of oxidized cytochrome *c* (obtained from mitochondrial preparations [3]) on a CM_{52} cellulose column. Sample amount = 57 O.D. (at 410 nM). Column: 13 cm × 1.4 cm. Buffer: 15 mM NaH_2PO_4–NaOH; pH 7.20. Linear gradient: 15 mM → 200 mM. Flow rate: 11.3 ml/hr.

Table 2
Amino acid composition of tryptic peptides

Amino acid	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9††	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}
Lys* or X	1.0 (1)	0.9 (1)	1.0 (1)	1.0 (1)	0.9 (1)	0.9 (1)		1.0 (1)	1.7 (2)	1.0 (1)	1.7 (2)		1.0 (1)	
His					0.9 (1)	0.9 (1)								
Arg							0.7 (1)				0.9 (1)			
Asp	1.1 (1)				1.0 (1)		2.0 (2)	2.0 (2)	3.2 (4)		0.4 (0–1)	1.1 (1)		
Thr				1.1 (1)	0.8 (1)			1.7 (2)	1.5 (2)	0.9 (1)				1.0 (1)
Ser	0.9 (1)							2.6 (3)						
Glu	1.1 (1)	1.0 (1)			1.1 (1)		1.0 (1)	1.2 (1)	1.0 (1)		1.5 (1–2)			0.9 (1)
Pro	2.8 (3)						1.0 (1)		0.9 (1)	0.9 (1)	1.8 (2)			
Gly	2.1 (2)	1.1 (1)				1.9 (2)	3.0 (3)	2.2 (2)	1.0 (1)	1.2 (1)	1.2 (1)			
Ala	2.1 (2)	0.9 (1)			1.0 (1)	1.2 (1)		1.8 (2)	1.0 (1)			1.8 (2)		2.1 (2)
Cys					1.2 (2)									
Val					0.8 (1)				0.9 (1)		0.8 (1)			
Met									0.7 (1)		0.6 (1)			
Ile			1.0 (1)							0.9 (1)		1.0 (1)		
Leu							1.7 (2)		2.6 (3)		0.9 (1)	1.0 (1)	1.0 (1)	
Tyr								1.6 (2)	1.7 (2)	0.8 (1)		1.0 (1)		
Phe	1.1 (1)		1.0 (1)				0.8 (1)				0.9 (1)			
e.m.**	−0.70	0.00	1.53	1.93	−0.55	2.57	0.73	0.62	0.00	1.09	1.26	−1.00	1.93	−1.58
Final yield†	47%	14%	32%	5%	9%	17%	8%	23%	20%	40%	34%	15%	5%	30%

* Lys or X (ε, *N*–trimethyl-lysine) were not separated using Spinco Automatic Amino Acid Analyzer model 120 C.
** e.m. = Electrophoretic mobility at pH 6.5.
† Final yield corresponds to fully purified peptide.
†† Ehrlich positive.

Table 3
Amino acid composition of chymotryptic peptides

Amino acid	C_1	C_2	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{14}	C_{16}
Lys* or X		1.8 (2)	0.9 (1)			1.1 (1)				2.0 (2)	1.0 (1)	1.9 (2)	1.0 (1)
His													
Arg				0.8 (1)								1.0 (1)	
Asp		1.0 (1)	1.9 (2)			2.6 (3)		1.0 (1)	0.8 (1)	1.2 (1)		1.4 (1–2)	
Thr				1.7 (2)			0.8 (1)	0.9 (1)			1.0 (1)		1.0 (1)
Ser	0.9 (1)			1.0 (1)	1.0 (1)	0.5 (1)							
Glu		2.0 (2)	1.0 (1)	1.0 (1)				1.1 (1)				1.5 (1–2)	1.0 (1)
Pro		2.5 (3)	0.9 (1)							1.2 (1)	0.9 (1)	1.0 (1)	
Gly		2.7 (3)	2.0 (2)	2.7 (3)				0.9 (1)			1.2 (1)		
Ala	1.1 (1)	2.0 (2)		1.1 (1)		1.0 (1)	1.1 (1)					0.9 (1)	2.1 (2)
Cys													
Val							0.5 (1)						
Met						0.5 (1)					0.9 (1)		
Ile		0.9 (1)									0.9 (1)		
Leu			1.9 (2)					0.9 (1)	1.0 (1)	0.9 (1)		0.9 (1)	0.9 (1)
Tyr				0.9 (1)	0.9 (1)			1.0 (1)	0.9 (1)	0.7 (1)			
Phe	0.9 (1)	1.0 (1)	0.9 (1)										
e.m.**	−1.86	0.00	0.00	0.88	0.00	1.04	0.00	−0.97	1.38	2.00	1.23	0.88	0.00
Final yield †	28%	4%	15%	15%	5%	3%	1%	35%	3%	17%	46%	2%	12%

* Lys or X (ϵ, N–trimethyl-lysine) were not separated using Spinco Automatic Amino Acid Analyzer model 120 C.
** e.m. = Electrophoretic mobility at pH 6.5.
† Final yield corresponds to fully purified peptide.

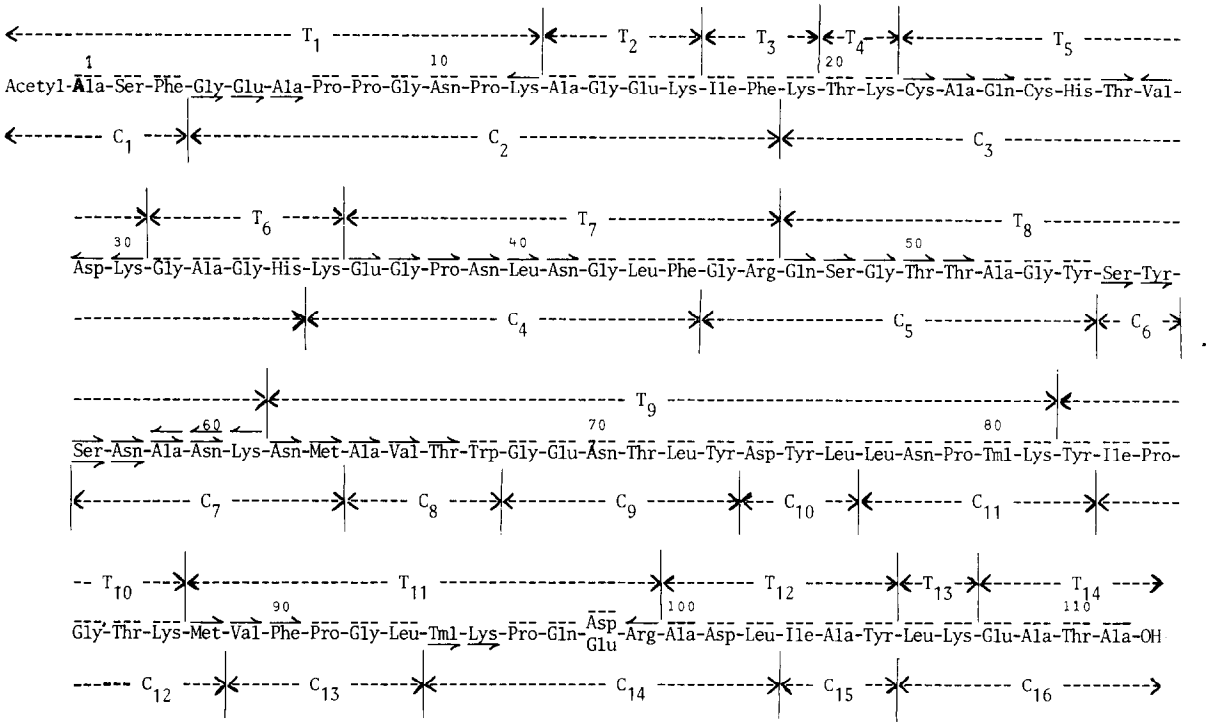

Fig. 2. Amino acid sequence of potato cytochrome *c*. Most of the whole sequence is given by homology with tomato cytochrome *c* as discussed in the text. (– – –) Residues determined by amino acid compositions of purified tryptic or chymotryptic peptides; (↓↑) major enzymic cleavages (→) and (↠) sequence determined by the Dansyl–Edman method on tryptic or chymotryptic peptides; (←) sequence determined by using carboxypeptidases A and B. Abbreviations: Tml = ε, *N*–trimethyl-lysine; T = tryptic; C = chymotryptic.

equivalent peptides as those liberated in tomato cytochrome *c* because of the identical enzymatic splits in the two molecules. However, chymotryptic peptides C_3, C_{13} and C_{15} were not recovered probably due to their low yield. Taking into account their very high homology, it was assumed that peptides belonging to the two cytochromes *c* (from tomato and potato) showing the same electrophoretical mobility and the same amino acid composition should be considered as identical. This assumption was strongly supported by overlappings between tryptic and chymotryptic peptides. The peptides that behaved differently were subjected to structural trials (amino acid sequence determinations, carboxypeptidases digestions, enzymatic hydrolyses) in order to establish the location of the mutations (fig. 2). These were found in positions 4, 29, 58, 66 and 98. The N-terminal residues of peptides C_1, T_1 and native protein gave no Edman reaction and subsequently were considered to be blocked, probably by *N*-acetylation, as are all cytochromes *c* from higher plants yet investigated.

4. Discussion

The tryptic and chymotryptic cleavages were found to be identical to those observed in tomato cytochrome *c* including the unexpected T_{12} release. The high ratio of proline in the whole molecule probably inhibited some splits (for example Asn in position 10).

One can recognize in the amino acid sequence several main caracteristics of higher plants cytochrome *c*: a single chain consisting of 111 residues, N-terminal end blocked (probably by acetylation), Ala–Ser–Phe as N-terminal sequence, Thr in positions 50 and 51, two ε, *N*-trimethyl-lysine residues in the whole amino

acid composition (placed by homology at positions 80 and 94). When looking to all higher plants cytochrome *c* yet described [10] position 97 is always Gln. Taking into account electrophoretic mobilities of peptides T_{11} and C_{14}, one of the two residues 97 or 98 should be acidic and the other neutral (amide). The action of both CpB and CpA on peptide T_{11}, which resulted in the liberation of the only Arg, supports the presence of a acidic residue in position 98. Moreover according to the amino acid compositions of peptides T_{11} and C_{14}, two iso-cytochromes *c* might exist: one with an Asp, the second with a Glu in position 98. Although (due to its low yield) it was not possible to determine the entire sequence of peptide C_{14}, a difference (amide → acidic form) must exist in position 98 between the tomato and potato cytochromes *c*. Residue number 4 is Gly and not Asn or Gln as it was found very often in cytochrome *c* of other higher plants.

Astonishly also we discovered (without ambiguity) Asn instead of Ala in position 58 as it was encountered in quite all other higher plants cytochromes *c* described. The position 66 gave Thr instead of Asn found in tomato cytochrome *c* but the fact seems less interesting as this position corresponds to a 'variation' possible amino acid when looking to all other higher plants cytochromes *c* discovered. By homology considerations with mammalian cytochromes *c*, the residue 24 was considered systematically as Gln in all higher plants. The use of CpA + CpB on peptide T_5 permitted to demonstrate without any ambiguity that residue 29 was Asp (and not Asn). So according to electrophoretic mobilities of T_5, residue 24 must be Gln which is in agreement with the assumptions of E. W. Thompson et al. [11].

Five mutations exist between tomato and potato cytochromes *c* belonging to the same family (*Solanaceae*). This seems rather surprising when taking into account the very low variation ability of cytochromes *c*. However the same phenomenon was elsewhere described [11], showing three differences between two genus (*Abutilon theophrasti* and *Gossypium barbadense*) in the same family (*Malvaceae*). Furthermore, for two species belonging to the same genus, the variation level seems almost naught [12]. All these facts support the hypothesis of Fitch and Margoliash [13] and allow the construction, by E. W. Thompson et al., of a cytochrome *c* phylogenetic tree in higher plants.

Acknowledgements

We wish to thank Miss Monique Diano (Laboratoire de Physiologie Cellulaire, Luminy, Marseille) for the first steps of the purification of cytochrome *c* involving the pure mitochondria preparations and Mr Louis Delbecchi (Laboratoire de Virologie, Faculté de Médecine, Secteur Nord, Marseille) for all polyacrylamide gel electrophoresis controls.

References

[1] Scogin, R., Richardson, M. and Boulter, D. (1972) Archives of Biochemistry and Biophysics 150, 489–492.
[2] Richardson, M., Laycock, M. V., Ramshaw, J. A. M., Thompson, E. W. and Boulter, D. (1970) Phytochemistry 9, 2271.
[3] C.R. Acad. Sc. Paris, t. 272, p 2692–2694 (24 mai 1971) (Note de Mlle. M. Diano et Mr. G. Martinez présentée par Mr Buvat).
[4] Thompson, E. W., Laycock, M. V., Ramshaw, J. A. M. and Boulter, D. (1970) Biochem. J. 117, 183–192.
[5] Offord, R. E. (1966) Nature 211, 591.
[6] Delange, R. J. (1969) J. Biol. Chem. 244, 5, 1385.
[7] Hartley, B. S. (1963) Biochem. J. 89, 379.
[8] Woods, K. R. and Wang, K. T. (1967) Biochim. Biophys. Acta 133, 369.
[9] Flatmark, T. (1966) Acta Chem. Scand. 20, 1487.
[10] Dayhoff, M. O. (1972) Atlas of Protein Sequence and Structure, Vol. 5, Silver Spring, Md., US. National Biomedical Research Foundation.
[11] Thompson, E. W., Notton, B. A., Richardson, M. and Boulter, D. (1971) Biochem. J. 124, 787–791.
[12] Richardson, M. Ramshaw, J. A. M. and Boulter, D. (1971) Biochem. Biophys. Acta 251, 331–333.
[13] Fitch, W. M. and Margoliash, E. (1967) Science 155, 279.