The Amino Acid Sequence of Cytochrome *c* from *Cucurbita maxima* L. (Pumpkin)

By E. W. THOMPSON, M. RICHARDSON AND D. BOULTER
Department of Botany, University of Durham, Durham City, U.K.

(Received 17 March 1971)

The amino acid sequence of pumpkin cytochrome *c* was determined on 2 μ mol of protein. Some evidence was found for the occurrence of two forms of cytochrome *c*, whose sequences differed in three positions. Pumpkin cytochrome *c* consists of 111 residues and is homologous with mitochondrial cytochromes *c* from other plants. Experimental details are given in a supplementary paper that has been deposited as Supplementary Publication SUP 50005 at the National Lending Library for Science and Technology, Boston Spa, Yorks. LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in *Biochem. J.* (1971), 121, 7.

The determination of the amino acid sequence of pumpkin cytochrome *c* is a continuation of the investigation of the sequences of a range of plant cytochromes *c* (Boulter, Thompson, Ramshaw & Richardson, 1970).

MATERIALS

Pumpkin (*Cucurbita maxima* L. var. Flat White Boer) seeds were obtained from Gunson Seeds South Africa (Pty) Ltd., Johannesburg, South Africa. Other materials were as previously described (Thompson, Richardson & Boulter, 1971a).

METHODS

The isolation and purification of cytochrome *c* from pumpkin was described by Richardson, Laycock, Ramshaw, Thompson & Boulter (1970) and Richardson, Richardson, Ramshaw, Thompson & Boulter (1971). The E_{550} (reduced)/ E_{280} (oxidized) ratio was 1.10.

The methods used for the determination of the amino acid sequence were as previously described (Thompson *et al.* 1971a). Haem groups were removed from haem peptides as described by Ramshaw, Thompson & Boulter (1970).

RESULTS

Table 1 gives the amino acid composition of pumpkin cytochrome *c*. Some evidence was found from sequential analysis which suggests that there are two forms of cytochrome *c* in pumpkin. The sequence of the major form is shown in Fig. 1; the minor form had a lysine instead of a glutamine residue in position 47, alanine for proline in position 52, and serine for alanine in position 109. Among the peptides from the original chymotryptic digest, which were separated by electrophoresis at pH 6.5

and pH 1.9, were two peptides whose sequences

were $\begin{smallmatrix} 45 & 54 \\ \text{Gly-Arg-Gln-Ser-Gly-Thr-Thr-Pro-Gly-Tyr} \end{smallmatrix}$

and $\begin{smallmatrix} 45 & 54 \\ \text{Gly-Arg-Lys-Ser-Gly-Thr-Thr-Ala-Gly-Tyr} \end{smallmatrix}$ respectively. Both peptides could only be fitted into the complete cytochrome *c* sequence at positions 45-54. Particularly significant is the fact that they both contain the distinctive plant residues $\begin{smallmatrix} 50 & 51 \end{smallmatrix}$

Thr-Thr. Evidence from the tryptic digestion also supported the occurrence of both peptide sequences. A second example of peptide polymorphism was also found, but in this case the evidence rested on equal yields of two *N*-terminal amino acids at one position during dansyl-Edman analysis. The

peptide $\begin{smallmatrix} 106 & 111 \\ \text{Leu-Lys-Gln-Ala-Thr-Ala} \end{smallmatrix}$ after three Edman steps (Leu-Lys-Glu) was found to have alanine and serine in equal yield, followed by threonine and alanine residues, indicating the

presence of an additional peptide $\begin{smallmatrix} 106 \\ \text{Leu-Lys-Gln-Ser-} \end{smallmatrix}$

$\begin{smallmatrix} 111 \\ \text{Thr-Ala} \end{smallmatrix}$. The two peptides could not be separated, but this result is to be expected, since there would be no difference of charge between them. A similar polymorphism was found in the equivalent tryptic

peptides $\begin{smallmatrix} 108 & 111 \\ \text{Glu-Ala-Thr-Ala} \end{smallmatrix}$, which had equal amounts of alanine and serine at position 109.

If two forms of cytochrome do occur, then it should be possible to separate them by gradient elution on CM-cellulose, since they would differ by a unit charge, owing to the change from glutamine to lysine. On rechromatography of a sample from the remainder of the preparation used above, cytochrome *c* was eluted as a single peak from the CM-cellulose column; it was collected in two

fractions, which were desalted and digested separately with chymotrypsin. The peptide 'maps' obtained for the two fractions after electrophoresis

Table 1. *Amino acid composition of pumpkin cytochrome c*

The values for the amino acid analysis are the averages for one 24 h and two 72 h hydrolysates. Abbreviations: n.d., not determined; Me₃Lys, ϵ -trimethyl-lysine.

Amino acid composition (residues/molecule)		
	Amino acid analysis	Calc. from sequence
Asp	11.44	11
Thr	7.37*	7
Ser	4.92*	5
Glu	8.98	9
Pro	7.71	8
Gly	11.12	11
Ala	11.32	12
Val	3.18†	3
Met	1.27	1
Cys	n.d.	2
Ile	4.02†	4
Leu	7.59	8
Tyr	6.06	6
Phe	4.38	4
Me ₃ Lys	1.97	2
Lys	10.88	12
His	1.90	2
Arg	2.32	3
Trp	n.d.	1

* Calculated, assuming first-order kinetics for destruction.

† Maximal yield.

at pH 6.5 were identical, as were the two sets of peptides identified by their *N*-termini and amino acid compositions. These peptides were identical with those from which the sequence given in Fig. 1 was obtained, and they provided no evidence for the alternative form. (The experimental evidence for the two sequences is given in the supplementary paper SUP 50005.)

Consideration of the mobilities of the peptides indicated that there were ten amide groups, seven of which were placed by the mobilities of the peptides in which they occurred; the remaining three were placed by homology. Comparison with all other known cytochrome *c* sequences indicates that glutamine occurs in position 24 placing aspartic acid in position 29; homology considerations indicate that residue 98 is acidic, and therefore 97 must be glutamine. Reference to other plant cytochrome *c* sequences places glutamic acid in position 5, and therefore either position 4 or 10 is an amide.

DISCUSSION

Some evidence for the occurrence of two forms of cytochrome *c* have been presented by the demonstration of polymorphic forms of two peptides. Attempts to separate two proteins by gradient elution on CM-cellulose have not proved successful, even though this should be theoretically possible, in view of the unit charge difference due to a change from glutamine to lysine; until two proteins can be separated, the possible occurrence of two forms of cytochrome *c* must be considered tentative.

In view of the peptide polymorphism demonstrated, it is noteworthy that a lysine residue is

Fig. 1. Amino acid sequence of pumpkin cytochrome *c*. ---, Composition determined qualitatively; —, composition determined by quantitative amino acid analysis; —, sequence determined by the dansyl-Edman method; —, sequence determined by using carboxypeptidase; ↓, major enzymic cleavages; ↓, partial enzymic cleavages. Abbreviations: Me₃Lys, ϵ -trimethyl-lysine; T, tryptic; C, chymotryptic; for full details of individual peptides see the supplementary paper SUP 50005.

normally found in position 47 in most mammalian cytochromes *c*, whereas in plant cytochromes *c* this residue is glutamine, and also that plant cytochromes *c* have either a serine or an alanine residue at position 109, which in pumpkin is polymorphic for these two amino acids.

The sequence given in Fig. 1 shows that pumpkin cytochrome *c* is closely related to other plant cytochromes *c* of the mitochondrial type (Boulter *et al.* 1970). Its close relationship to that of cauliflower should be noted (see Thompson, Richardson & Boulter, 1971*b*), but the full phylogenetic implications of these data must await the determination of sequences of further plant cytochromes *c*.

We thank Dr M. E. Boulter and Mrs L. Mayo for invaluable assistance with the dansyl-Edman analysis,

and the Nuffield Foundation and the Science Research Council for financial support.

REFERENCES

- Boulter, D., Thompson, E. W., Ramshaw, J. A. M. & Richardson, M. (1970). *Nature, Lond.*, **228**, 552.
Ramshaw, J. A. M., Thompson, E. W. & Boulter, D. (1970). *Biochem. J.* **119**, 535.
Richardson, M., Laycock, M. V., Ramshaw, J. A. M., Thompson, E. W. & Boulter, D. (1970). *Phytochemistry*, **9**, 2271.
Richardson, M., Richardson, D. L., Ramshaw, J. A. M., Thompson, E. W. & Boulter, D. (1971). *J. Biochem., Tokyo*, **69**, 811.
Thompson, E. W., Richardson, M. & Boulter, D. (1971*a*). *Biochem. J.* **121**, 439.
Thompson, E. W., Richardson, M. & Boulter, D. (1971*b*). *Biochem. J.* **124**, 783.