Phytochemistry, 1972, Vol. 11, pp. 553 to 561. Pergamon Press. Printed in England.

STRUCTURE-FUNCTION RELATIONSHIPS IN PLANT CYTOCHROME *c*

D. BOULTER and J. A. M. RAMSHAW

Department of Botany, University of Durham, Durham

(*Received* 17 *June* 1971)

Abstract—Structure-function relationships in plant cytochrome *c* are discussed in the light of Dickerson's horse heart and bonito ferricytochromes *c* model, in conjunction with the complete amino acid sequences of 14 plant cytochromes *c*.

INTRODUCTION

IN A RECENT paper, Dickerson *et al.*[1] reported on the structure of crystalline horse heart and bonito ferricytochromes *c*, determined by X-ray methods to a resolution of 2·8 Å. The structures of both cytochromes were found to be identical apart from minor changes associated with side-chain substitutions, showing that the essential structure of the protein has been conserved for at least four hundred million years.

Cytochrome *c* is a good example of the 'oil-drop' model of a protein, with buried hydrophobic side-chains and polar charged side-chains on the surface. Four surface features are probably important for the function of the molecule in the electron transport chain: the partially exposed heme, which sits in an hydrophobic crevice, two channels which run from the interior to the surface to the left and the right of the heme, and a negative surface patch. The channels are hydrophobic, and positively charged lysine side-chains are associated with each at the surface. It is suggested that these surface features are involved in interactions with cytochrome oxidase, cytochrome reductase and the mitochondrial membrane.

From stereoscopic α carbon diagrams,[1] we have constructed an approximate 3-dimensional α carbon model of horse heart ferricytochrome *c*, which we have used in conjunction with the complete amino acid sequences of 14 plant cytochromes *c* (see Table 1[2–11]) to make reasonable predictions about some of the structure-function relationships of the plant cytochrome *c* molecule. Whilst these predictions do not have the authority of those based upon precise X-ray diffraction patterns, they would appear reasonable in the light of the X-ray data obtained by Dickerson's group.[1]

[1] R. E. DICKERSON, T. TAKANO, D. EISENBERG, O. B. KALLAI and L. SAMSON *J. Biol Chem.* **246**, 1511 (1971).
[2] J. A. M. RAMSHAW, M. RICHARDSON and D. BOULTER, *Europ. J. Biochem.* (in press).
[3] F. C. STEVENS, A. N. GLAZER and E. L. SMITH, *J. Biol. Chem.* **242**, 2764 (1967).
[4] D. RICHARDSON and D. BOULTER, Unpublished.
[5] E. W. THOMPSON, M. RICHARDSON and D. BOULTER, *Biochem. J.* **124**, 783 (1971).
[6] E. W. THOMPSON, M. RICHARDSON and D. BOULTER, *Biochem. J.* **124**, 779 (1971).
[7] E. W. THOMPSON, B. A. NOTTON, M. RICHARDSON and D. BOULTER, *Biochem. J.* **124**, 787 (1971).
[8] E. W. THOMPSON, M. V. LAYCOCK, J. A. M. RAMSHAW and D. BOULTER, *Biochem. J.* **117**, 183 (1970).
[9] J. A. M. RAMSHAW, E. W. THOMPSON and D. BOULTER, *Biochem. J.* **119**, 535 (1970).
[10] J. A. M. RAMSHAW and D. BOULTER, Unpublished.
[11] E. W. THOMPSON, M. RICHARDSON and D. BOULTER, *Biochem. J.* **121**, 439 (1970).

A COMPARISON OF PLANT AND HORSE HEART CYTOCHROME *c*

Overall Description of Plant Cytochrome c*

Dickerson's description of the horse heart model considers the molecule as consisting of two halves: residues 1–47 to the right of the heme crevice, and residues 48–91 to the left, with residues 92–104 coming back across the top to the right-hand side of the molecule (see Fig. 1).

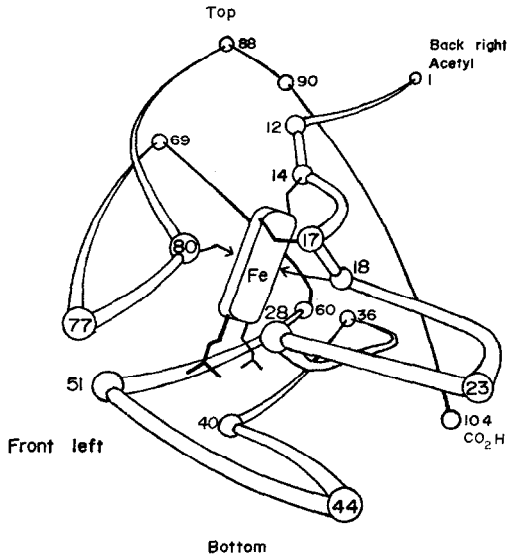

FIG. 1. DIAGRAMMATIC REPRESENTATION OF THE CYTOCHROME *c* MOLECULE TO ILLUSTRATE THE DESCRIPTIVE TERMS, LEFT, RIGHT, etc. USED IN THE TEXT.

The plant sequences (see Fig. 2) when compared with horse heart, all have an additional eight residues at the *N*-terminal end with an acetylated *N*-terminal residue. Identical stretches of sequence are found associated with the heme attachment sites. Residues 8–47 are invariant in the plant sequences, apart from positions 21, 22, 33 and 44; position 21, which in animals is always glu, is also acidic in plants but here can be either asp or glu. Lys occurs in position 22 in all plants except wheat, and this residue is also variable in animals; in horse heart cytochrome *c*, both these residues have side-chains to the outside of the molecule. The other two variable residues, 33 and 44, have two alternatives at each position in plants and are therefore, less variable than these positions in animals.

The second half of the molecule consists of residues 48–91. In plants, positions 48, 49, 52–54 are invariant, position 50 is ala or thr, a position also variable in animals, and position 51 is invariant ala in all cytochromes examined to date, except in *Ginkgo*, where this residue is replaced by gly.[2] Then comes a more variable region of the plant sequences starting with residue 55 and finishing with residue 66, in which six variable positions are found, although only conservative changes occur in positions 63 and 66. Thus, residue 63 is thr in all plant sequences except niger,[10] where it is replaced by ser, and 66 is always an acidic residue either asp or glu. The region, residues 67 to 91 inclusive, is invariant in plants except that position 90 can be either of the acidic residues asp or glu; this region contains the eleven residue sequence 70–80, which has been found to be invariant in all cytochromes *c*.

* Residues throughout are numbered to complement with the horse heart sequence.

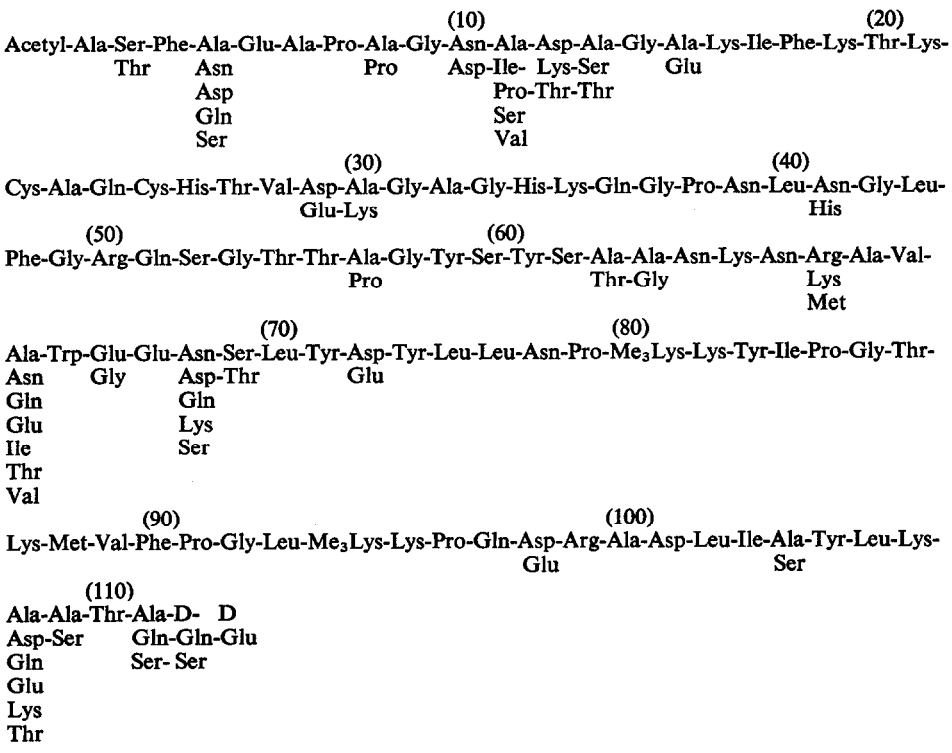

FIG. 2. SUMMARY OF THE VARIABLE AND INVARIANT POSITIONS OF PLANT CYTOCHROME *c*. Data from the cytochrome *c* sequences of the fourteen species listed in Table 1. D = deletion, Me_3Lys = ϵ-*N*-trimethyllysine.

The final region in the plant sequences is also invariant up to position 99, except for a substitution of ser for ala in *Ginkgo* in position 96;[2] the *C*-terminal residues 100–104 are highly variable, apart from thr in position 102.

Attachment of Heme Group

Plants like animals have cys 14 and cys 17 covalently linked to the heme via the vinyl side-chains of its α-carbon atoms. Thus, acid/acetone treatment is not sufficient for removal of the heme, and performic oxidation is necessary.[8] The heme chymotryptic and tryptic peptides always have the sequences: NH_2-lys thr lys cys ala gln cys his thr val $^{asp\ ala}_{glu\ lys}$ gly ala gly his CO_2H and NH_2-cys ala gln cys his thr val $^{asp\ ala\ gly\ ala\ gly\ his\ lys\ CO_2H}_{glu\ lys\ CO_2H}$, respectively. Evidence that his 18 and met 80 are also liganded to the heme as in horse heart cytochrome *c*, follows from the invariance of these positions in all plant cytochromes *c*, and from the absorption spectrum of the plant molecules.

Hydrophobic Residues

Of the 16 hydrophobic residues which point inward to the heme and pack tightly round it in horse heart, identical residues occur in these positions in plants, except that phe (46) is replaced by tyr, ile (81) by val, and ile (85) by leu. These replacements are hydrophobic, and chemically and sterically very similar; similar replacements have also occurred in these positions in some animals.

In horse heart cytochrome *c*, a right hydrophobic channel is bordered by residues 6–20 and the α-helix 92–101. In plants this hydrophobic channel probably exists also, involving the residues, phe 10; ile 9 and 95; leu 94 and 98; and tyr 97. In addition, in plants leu 32 and 35, val 20, have hydrophobic side-chains pointing into the cavity; val 11, whose side-chain points outwards in horse heart is replaced by lys in the plant sequences.

Residues 52–74 enclose the hydrophobic left channel in the horse heart model, which is again more or less filled with hydrophobic side-chains, particularly prominent being the invariant trp 59, tyr 67 and tyr 74, and these residues also occur in these positions in all plant cytochromes.

Aromatic Residues

In horse heart there are nine aromatic residues, of which only one, phe 36, is ever replaced in vertebrates by a non-aromatic group. In plants there are eleven aromatic residues and these are always invariant; the additional ones to those in horse heart cytochrome *c* are phe at minus 6 in the plant *N*-terminal 'tail', and tyr (65) associated with the hydrophobic left channel; the parallel orientation of aromatic rings is the postulated mechanism for electron transport through overlapping π electron orbitals.[1]

Basic Residues

In horse heart the basic residues are not randomly distributed, and their side-chains all point outwards.[1] On the left side of the molecule, residues 86, 87 and 88, 72, 73 and 79, 53, 39, 55, form three basic straps, and these basic areas also occur in plants, although positions 88 and 39 are not basic in plants and position 55 is not always so. On the right side of the animal molecule, residues 5, 7, 8, 13, 22, 25, 27, 99 and 100 are basic. In plants, the corresponding basic region consists of residues 4, 7, 11, 13, 22, 27, 99 and 100.

TABLE 1. PLANT CYTOCHROMES *c*

Species	Common name	Family	Residues	Molecular weight of cytochrome *c*
Ginkgo biloba L.[2]	Ginkgo	Ginkgoaceae	113	13107
Triticum spp.[3]	Wheat	Poaceae	112	12792
Zea mays L.[4]	Maize	Poaceae	111	12793
Arum maculatum L.[4]	Arum	Araceae	111	12783
Fagopyrum esculentum Moench.[5]	Buckwheat	Polygonaceae	111	12843
Cucurbita maxima Duchesne[6]	Pumpkin	Cucurbitaceae	111	12866
Abutilon theophrasti Medic.[7]	Abutilon	Malvaceae	111	12782
Gossypium barbadense L.[7]	Cotton	Malvaceae	111	12756
Phaseolus aureus L.[8]	Mung bean	Fabaceae	111	12877
Helianthus annuus L.[9]	Sunflower	Asteraceae	111	12702
Guizotia abyssinica Cass.[10]	Niger	Asteraceae	111	12684
Brassica oleracea L.[5]	Cauliflower	Brassicaceae	111	12827
Sesamum indicum L.[11]	Sesame	Pedaliaceae	111	12796
Ricinus communis L.[11]	Castor	Euphorbiaceae	111	12725

The plant molecule is not as basic as the animal; thus, the isoelectric point of mung bean cytochrome *c* has been determined by iso-electric focusing to be pH 9·5,[12] compared with

[12] M. V. LAYCOCK, Unpublished.

near pH 10 for animal cytochrome *c*.[13] Thus, horse heart has 19 lysine, 2 arginine and 12 acidic residues, compared with plant cytochromes which on average have 12 lysine, 2 arginine and 9 acidic residues.

Acidic Residues

Acidic residues are also not distributed randomly. The acidic patch residues 2–4 in horse heart[1] is not evident in plants, except possibly in wheat. Residue 21 however, is acidic in both plants and animals. Acidic patches occur in the horse heart sequence from residues 60–69 and 89–93;[1] similar acidic patches occur here in plants and acidic residues are not found elsewhere.

Hydrogen Bonding

The plant molecules have a larger number of residues capable of forming hydrogen bonds than do those of animals. In horse heart the δ nitrogen atom of his 18 is hydrogen-bonded to the carbonyl oxygen of pro 30,[1] and plants also have his and pro in these positions. Also, in horse heart,[1] the hydroxyl of tyr 67 is hydrogen-bonded to thr 78, and hydrogen bonding also occurs between tyr 48, trp 59 and the main chain carbonyl of residue 40, with the buried propionic group of the heme. Since these residues also occur in these positions in plants, it is probable that hydrogen-bonding occurs there also. In horse heart[1] the other heme propionic group is probably hydrogen-bonded to thr 49; in plants this residue is ser. Thr 63, which bonds to the carbonyl of residue 60 in animals, also occurs in all plants except niger, where it is replaced by ser.[10] Thr 78, which is an invariant residue, is bonded to invariant tyr 67 in animals and presumably also in plants. Of the four threonyl residues, 28, 47, 58 and 89, which are on the surface and not apparently hydrogen-bonded in horse heart,[1] residues 28 and 89 are gln in plants, a residue known to form a similar hydrogen bond to thr; residue 47 is ser, and position 58 is not occupied by a hydrogen-bonding residue. Thr 40 in horse heart connects two polypeptide chains by a hydrogen bond with the main chain carbonyl of residue 55;[1] it is replaced in plants by ser 40. Invariant thr 102, which also occurs in this position in plants, has no obvious hydrogen-bonded partner in animals. In horse heart only one asn residue, asn 52, is not on the surface. This residue is invariant and probably involved in hydrogen bonding with thr 49 (ser 49 in plants) and the outer heme propionic group. Gln 16 in horse heart also occurs in plants; in animals it points out and is apparently unbonded. Gln 12 and gln 42, which in plants are thr residues, interact with arg 91 and arg 38 in horse cytochrome, the former bond closing the top of the heme crevice and latter the bottom of the molecule. Although potential hydrogen-bonding residues may occur in plants at positions minus 5, 2, 3, 5, 33, 50, 62, 100, 101, 103 and 104, they do not invariably do so; positions minus 7, 39 and 43 however, contain invariant hydrogen-bonding residues.

Glycine Residues

In animals there are twelve invariant glycine residues, nine of which occur in the first half of the molecule, and in most instances, the steric reasons for gly in these positions is clear.[1] In plants there are eleven invariant glycine residues and nine occur in the first half of the molecule; gly (56) in animals is replaced in all plants by ala. There are three type II 3_{10} bonds in the horse cytochrome *c*,[1] and the sequence constraints required for this type of bond also occur in the plant molecules.

[13] G. H. BARLOW and E. MARGOLIASH, *J. Biol. Chem.* **241**, 1473 (1966).

Invariants

If one considers all cytochrome *c* sequences except those of higher plants, there are 34 residues which have been observed to be invariant. In addition, residue 72 is also invariantly lys as synthesized, although this residue is subsequently tri-methylated in fungi and higher plants. The only change in the situation when one considers the higher plant cytochromes, is that *Ginkgo* has gly[2] instead of the hitherto invariant ala at position 51, reducing the invariant residues to an overall total of 33.

α-Helical Regions

Present methods for predicting an α-helix do not give good results for the cytochrome *c* molecule.[1] In horse heart only residues 91–101 exist in α-helical conformation;[1] the other parts predicted as α-helical are distorted by the dominance of the heme prosthetic group.

Considering some of the predictive methods, the regions which they forecast as helical in plants, are essentially the same as those predicted in horse cytochrome. The results in Table 2 give the predicted regions using the Kotelchuck and Scheraga[14] H_4C_2 model. Residue 90, glu or asp, is critical in determining whether the prediction is positive or negative for the *C*-terminal region. When glu occurs here, the method gives a helical prediction and when asp, non-helical as in the case of wheat. Thus, the statement[1] that the prediction breaks down in this region for plants, is not generally true and refers only to wheat.

TABLE 2. KOTELCHUCK AND SCHERAGA'S H_4C_2 AND PROTHERO'S METHODS FOR PREDICTION OF HELICAL REGIONS APPLIED TO PLANT CYTOCHROME *c*

Species	Predicted helical regions				
	Kotelchuck and Scheraga		Prothero		
	1	2	A	B	C
Ginkgo		(89–95)	(15–24)	(56–70)	(88–98)
Wheat	(56–61)		(18–24)	(54–69)	(90–101)
Maize	(56–61)	(89–103)	(15–24)	(54–69)	(88–100)
Arum	(55–61)		(15–24)	(54–69)	(88–100)
Buckwheat		(89–98)	(15–24)	(55–70)	(88–100)
Pumpkin	(56–61)			(54–69)	(90–100)
Abutilon	None		(15–24)	(63–69)	(90–100)
Cotton	(55–61)			(55–62/63–69)	(90–100)
Mungbean	(55–61)			(54–69)	(90–100)
Sunflower		(89–98)	(15–24)	(54–69)	(88–100)
Niger		(89–100)	(15–24)	(56–64)	(88–100)
Cauliflower	(56–61)			(54–69)	(90–100)
Sesame	(55–61)	(89–103)		(55–62/63–69)	(88–100)
Castor	(55–61)		(15–24)	(55–62/63–69)	(90–100)

In the *C*-terminal region, relaxation to the H_3C_2 rule gives the region from residue 94→ as helical in all species, but does not extend the prediction any further towards *N*-terminus. No region is consistently predicted as helical in all species by this method.

[14] D. KOTELCHUCK and H. A. SCHERAGA, *Proc. Nat. Acad. Sci. U.S.* **61**, 1163 (1968); D. KOTELCHUCK and H. A. SCHERAGA, *Proc. Nat. Acad. Sci. U.S.* **62**, 14 (1969).

Prothero's method[15] (Table 2) predicts region 15–24 as helical in most species, in addition to the two regions generally predicted by Kotelchuck and Scheraga's method. A prediction in region A (see Table 2) depends on the nature of residue 21 (asp or glu), and one in region B on residue 60 (gly/glu).

The method of Schiffer and Edmundson[16] using helical wheels predicts as α-helix, regions 31–38 and 62–70 as for horse cytochrome. However, in the other regions which this method predicts for horse cytochrome, an α-helix is not generally predicted in plants because of the substitution of lys (11) in region 9–19, of pro (83) in region 80–87 and ser (101) in region 95–101.

Overall, the three methods give a similar result, but they are not sufficiently discriminatory to make accurate predictions. It is unlikely that the additional residues of the *N*-terminal plant-tail have an α-helical conformation, since this is not predicted by any method.

DISCUSSION

Examination of the sequences of fourteen higher plant cytochromes *c* (see Fig. 2), suggests that all the main features described by Dickerson *et al.*[1] for the animal molecules, also occur in higher plant cytochrome *c*. These include, (1) the binding of the heme and its interactions with the polypeptide chain; (2) the hydrophobic environment of the heme; (3) the two hydrophobic channels and the role of aromatic side-chains; (4) the importance of the basic surface patches; (5) the involvement of several residues in hydrogen-bonding, and the importance of the eleven invariant glycine residues.

There are however, differences between the construction of the animal and plant molecules. Thus, (1) in vertebrates, residues 44, 60, 89 and 92 are highly variable, whereas in plants these residues are either invariant (89, 92) or with only two alterations (44, 60), (2) variable regions of the sequence occur in different parts of the vertebrate and plant molecules,[17] although in both, the variable positions are exposed (Fig. 1). In plant, the variable *N*-terminal region occurs at the top back right-hand side of the molecule with the variable *C*-terminal region below it. Residues 21 and 22 are variable and are to the front right-hand side of the molecule level with the heme. The region 55–63 is variable as in animals it is level with the heme at the back bottom left-hand side of the molecule. The lateral segments and the solid centre of the molecule, where the polypeptide chain bends back repeatedly to allow hydrophobic side-chains to pack against the heme, are invariant, apart from slight variation in one or two residues. In contrast, the variable sections in the vertebrate sequences occur at the top right-hand lateral side of the molecule, at the front right-hand side level with the heme, and the top back left-hand side, (3) there are twelve positions, invariant in higher plant cytochromes *c*, the amino acid residues of which, do not occur in these positions in any non-higher plant sequences, (4) plant cytochromes *c*, as compared to those of vertebrates, are readily digested by trypsin and chymotrypsin. In this respect they resemble the non-vertebrate animals and yeast,[18] although the reasons for the differences are not known. However, the ratio 416 (red.)/550 (red.) is consistently greater in plant cytochromes than in mammals, reflecting the more open structure of the former.

In mammals,[19] residues 58 and 60 are two of the immunological sites of the molecule, and

[15] J. W. Prothero, *Biophys. J.* **6**, 367 (1966).
[16] M. Schiffer and A. B. Edmundson, *Biophys. J.* **7**, 121 (1967).
[17] E. Margoliash, G. H. Barlow and V. Byers, *Nature, Lond.* **228**, 723 (1970).
[18] E. Margoliash and A. Schejter, *Adv. Prot. Chem.* **21**, 113 (1966).
[19] E. Margoliash, A. Nisonoff and M. Reichlin, *J. Biol. Chem.* **245**, 931 (1970); A. Nisonoff, M. Reichlin, and E. Margoliash, *J. Biol. Chem.* **245**, 940 (1970); M. Reichlin, A. Nisonoff and E. Margoliash, *J. Biol. Chem.* **245**, 947 (1970).

Dickerson[20] has pointed out the exceptional nature of the region 56–60 containing as it does, hydrophobic residues in a hydrophilic area. In plants, residue 58 is also highly variable, and the region 56–60 contain hydrophobic and hydrophilic residues.

Two trimethyllysine residues (Me_3lys) are found in all higher plant sequences in positions 72 and 86, whereas in fungi only lys in position 72 is trimethylated, and no animal sequences contain this amino acid derivative.[21] From a consideration of the horse heart cytochrome *c* structure,[1] it would appear that in the plant molecule all the lysine residues are exposed and therefore, direct availability of different lysine residues to a methylating enzyme does not explain why only residues 72 and 86 are substituted. It is possible that proline in position 83 of higher plants is important in this connection, but speculation in the absence of accurate 3-dimensional data is probably unprofitable, although the involvement of lys 13 in cytochrome oxidase activity[22] clearly indicates the undesirability of methylation of this residue.

At first sight it might appear that there has been a progressive shortening of the molecule from the *N*-terminal end during the evolution of animals and fungi.[23] However, there is now strong evidence that in mammals and fungi, cytochrome *c* is synthesized on 80s cytoplasmic ribosomes.[24] Also, it has been clearly demonstrated that the mechanism for the initiation of protein synthesis on 80*s* ribosomes involves $tRNA_{F*}^{Met}$,[25] implying that when originally synthesized the *N*-terminus of the cytochrome *c* molecule was methionine and therefore, that subsequently, a hydrolytic enzyme, although not yet isolated, must have cleaved off some of the *N*-terminal residues. The fact that the cytochrome *c* molecule of plants, insects, fungi and animals, differs in its overall length, might therefore reflect the specificity of this postulated enzyme, rather than the length of the gene responsible for coding for the molecule. Invariance of gly 1 in animals may be related to the specificity of the enzyme, as it is not clear otherwise why this residue is invariant.[1] The variable length at the *C*-terminal region in plants, which in *Ginkgo* cytochrome *c* has two[2] and in wheat one[3] residue more than the other plant sequences, probably reflects the possibility of mutation to a chain terminating triplet.[23]

The limited variation found in the plant sequences considering that members of both monocotyledons, dicotyledons and a gymnosperm, are included in the investigation, fully substantiates the view that cytochrome *c* is a conservative molecule in evolution.[26] This is understandable since it is a relatively small molecule as far as proteins go, and has at least three interacting functions which involve surface features. In addition, a relatively large prosthetic group which has remained constant during the evolution of the cytochrome *c* molecule, dictates the manner of folding of the molecule and imposes severe structural constraints. It is of interest to note that using the terminology proposed by Zuckerkandl and Pauling,[27] cytochrome *c* is a mixture of a semantide molecule (the protein) and an episemantide molecule (prosthetic group). The substitution in *Ginkgo* cytochrome *c* of gly for ala at an hitherto invariant position, i.e. 51, has led us to look at the invariant position in the light

[20] R. E. DICKERSON, Discussion to E. MARGOLIASH, M. REICHLIN and A. NISONOFF, in *Structure and Function of Cytochromes* edited by (K. OKUNUKI, D. K. MARTIN and I. SEKUZU), p. 281, University of Tokyo Press, Tokyo (1968).
[21] R. J. DELANGE, A. N. GLAZER and E. L. SMITH, *J. Biol. Chem.* **245**, 3325 (1970).
[22] K. WADA and K. OKUNUKI, *J. Biochem. Tokyo* **66**, 249 (1969).
[23] E. L. SMITH, in *Harvey Lectures*, p. 231, series 62, Academic Press, New York (1968).
[24] D. BOULTER, *Ann. Rev. Pl. Physiol.* **21**, 91 (1970).
[25] A. E. SMITH and K. MARCKER, *Nature, Lond.* **226**, 607 (1970).
[26] M. O. DAYHOFF, *Atlas of Protein Sequence and Structure*, Vol. 4. Silver Spring, Md.: National Biomedical Research Foundation (1969).
[27] E. ZUCKERKANDL and L. PAULING, *J. Theoret. Biol.* **8**, 357 (1965).

of the Dickerson model.[1] Thus, the side-chain residue 51 is not totally internal. The latest prediction for the total number of invariant residues irrespective of the number of sequences examined is 32.[28] At present 34 residues are totally invariant in all cytochromes and of these only gly 45 and 84 and ile 75 have not been specifically correlated with a structural aspect of the horse molecule, indicating that so far as the mitochondrial eukaryotic cytochrome *c* molecule is concerned, the statistical estimate of invariants is the same as that from structural considerations.

Consideration of the plant and animal sequences indicates that essentially the 3-dimensional structure of cytochrome *c* has been conserved at least since the divergence of the plant and animal lines from their common ancestor.[23] Using the method described by Nolan and Margoliash,[29] this time is calculated to be in excess of one thousand million years.

Acknowledgements—We thank Dr. R. E. Dickerson for allowing us to see his X-ray data prior to publication, and Drs. M. Richardson and E. W. Thompson and Mr. D. Richardson for allowing us to use their data, some of which is unpublished. We also thank the Nuffield Foundation and the S.R.C. for financial support.

[28] W. M. FITCH and E. MARKOWITZ, *Biochem. Genetics* **4**, 559 (1970).
[29] C. NOLAN and E. MARGOLIASH, *Ann. Rev. Biochem.* **37**, 727 (1968).

Key Word Index—Cytochromes *c*; amino acid sequences; structure-function relationships; protein evolution.