Biochem. J. (1976) **159**, 193–199
Printed in Great Britain

Amino Acid Sequence of Cytochrome *c* from *Tetrahymena pyriformis* Phenoset A

By GEORGE E. TARR and WALTER M. FITCH
Department of Physiological Chemistry, University of Wisconsin Medical School, Madison, WI 53706, *U.S.A.*

(*Received* 23 *February* 1976)

The cytochrome *c* of *Tetrahymena pyriformis* GL (Phenoset A) had an isoelectric point of 6.5 and by sequence the following composition: $Asp_{(7)}$ $Asn_{(2)}$ $Thr_{(4)}$ $Ser_{(8)}$ $Glu_{(6)}$ $Gln_{(2)}$ $Pro_{(7)}$ $Gly_{(13)}$ $Ala_{(13)}$ $Val_{(7)}$ $Met_{(2)}$ $Ile_{(5)}$ $Leu_{(6)}$ $Tyr_{(2)}$ $Phe_{(5)}$ $Lys_{(11)}$ $His_{(3)}$ $Trp_{(1)}$ $Arg_{(3)}$ $Cys_{(2)}$ (total 109 residues). The peptides derived from the protein afforded complete overlap, so a complete sequence could be determined without reference to homologous proteins. Alignment with other mitochondrial cytochromes *c* required two internal deletions totalling three residues and an *N*-terminal region two residues longer than, and a *C*-terminal region one residue shorter than, the previously known limits. The sequence was the most divergent of the known mitochondrial cytochromes *c*, suggesting a distant relationship of ciliates to other eukaryotes. Details of the sequence data have been deposited as Supplementary Publication no. SUP 50068 (37 pages) at The British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms given in *Biochem. J.* (1976) **153**, 5.

The comparative study of protein sequences, particularly of cytochrome *c*, has yielded dendrograms taken to be approximate representations of the phylogenies of the organisms sampled (Fitch & Margoliash, 1967; Brown *et al.*, 1972; McLaughlin & Dayhoff, 1973). Application of this technique to the Protista would seem to be especially appropriate, as classical approaches such as comparative morphology have not provided any widely accepted phylogenetic scheme. Unfortunately, although the primary structures have been determined for about 50 cytochromes *c* from plants, animals and fungi, sequences for mitochondrial cytochromes are available only for three protists, namely the euglenid *Euglena gracilis* (Pettigrew, 1973; Lin *et al.*, 1973), the trypanosome *Crithidia oncopelti* (Pettigrew, 1972), and the myxomycete *Physarum polycephalum* (Lin, 1971). It was therefore decided to sequence comparable proteins from other protists that could be obtained in sufficient quantities. *Tetrahymena pyriformis* was chosen because it is easy to culture and because of the important and disputed relationship between ciliates and the Metazoa (Dougherty, 1963). Physical and biochemical characteristics of the cytochromes *b*-560 and *c*-553 of this organism have been determined by Yamanaka *et al.* (1968).

Materials and Methods

T. pyriformis used in the present study was cloned from a strain obtained from Dr. John Moner (Department of Zoology, University of Massachusetts at Amherst), who has identified it as strain GL-7 of Phenoset A (Borden *et al.*, 1973). The organism was cultured at 28°C in 6-litre Erlenmeyer flasks (1.5–2 litres of medium) with shaking, 5-gallon carboys (16 litres) with sparging, or a 50-gallon stainless-steel fermenter (160 litres) with sparging and stirring. Cultures were harvested by centrifugation after reaching the stationary phase as the cytochrome *c* content was greatest at this time (Shrago *et al.*, 1971). The culture medium initially used was 2% (w/v) proteose peptone plus liver extract and salts [as described by Kidder & Dewey (1951); phosphates omitted], but this was gradually changed to the less expensive 0.5% peptone, 0.2% yeast extract, 2% (w/v) glucose plus salts, also described by Kidder & Dewey (1951); all components were autoclaved together. The high-glucose medium gave higher yields (up to about 1.5×10^6 cells/ml) without apparent change in the concentration of cytochrome *c*.

Harvested cells were frozen, and stored at −20°C. In a typical extraction of cytochrome *c*, which was similar to that previously used for *Euglena* (Lin *et al.*, 1973), partly thawed lumps of cells were homogenized in a Waring Blendor for 2min with 1.5vol. of water. The pH was adjusted to 4.0 with acetic acid. The homogenate was stirred for 1h at 22°C, the pH was brought back to 6.5 with conc. NH_3 (sp.gr. 0.90) and solid $(NH_4)_2SO_4$ was added to produce a 50% saturated solution which was left to stand at 0–5°C for several hours. The supernatant fluid obtained by centrifugation at 10000*g* for 30min was adjusted to 100% $(NH_4)_2SO_4$ saturation to precipitate the cytochrome. Low-molecular-weight solutes were removed by concentrating to 20–30ml in an Amicon Diaflo apparatus equipped with a UM-10 filter, then washing

continuously with 20–30vol. of 0.005M-Tris/HCl buffer, pH8.5, at 6–7kg/cm pressure. The oxidized cytochrome was adsorbed to a column of DEAE-Sephadex A-25 equilibrated with 0.01M-Tris/HCl, pH8.4. Cytochrome B-560 was only slightly retarded, whereas cytochrome *c* required about 0.03M-NaCl for elution. Final purification was achieved by filtration through Bio Gel P-60 (100–200mesh) in 0.1M-ammonium bicarbonate. About 15mg of pure cytochrome *c* was obtained per kg of packed cells (see the Results section).

If the homogenate was more dilute, or if the $(NH_4)_2SO_4$ saturation was first brought to 70–75% saturation to remove the large amount of glycogen (see Levy, 1973) and other material, the cytochrome *c* did not usually precipitate at 100% saturation. In this case a prolonged concentration in the Diaflo apparatus (several weeks), followed by washing with dilute buffer, was required before proceeding to the ion-exchange column.

Finally, it was found that *T. pyriformis* cytochrome *c* could be collected from a dilute homogenate on a short, wide column of DEAE-cellulose (coarse grade), equilibrated with 0.01M-Tris/HCl, pH8.4, which was then washed with dilute buffer and extracted with 0.2M-NaCl.

The isoelectric point was determined with an LKB isoelectric-focusing unit over a pH range of 5–8. Visible and u.v. spectra were taken with a Shimadzu MPS-50L spectrophotometer.

Peptides were produced from 0.2–1.0μmol samples of cytochrome *c* by enzymic digestion with trypsin (L-1-chloro-4-phenyl-3-tosylamidobutan-2-one-treated, from Worthington Biochemical Corp., Freehold, NJ, U.S.A.) or chymotrypsin in 0.2M-*N*-methylmorpholine adjusted to pH7.6 with acetic acid. To block tryptic cleavage at lysine residues in one digest, the cytochrome was first acetylated by adding 1vol. of *N*-methylmorpholine to 1vol. of protein solution, cooling to 0°C, then adding 1/3vol. (a 200-fold excess over reactive groups) of acetic anhydride with mixing. After 30min at 22°C, the acetylated protein was precipitated with ethyl acetate/acetone (2:1, v/v) and washed with acetone. Cleavage at methionine residues was effected with a 100-fold excess of CNBr in trifluoroacetic acid for 40h at 22°C in the dark; excess of reagent was removed by freeze-drying. The single tryptophan residue was attacked in other samples of the protein for 48h with 10 equivalents of 2-(2-nitrophenylsulphenyl)-3-methyl-3-bromoindolenine in 67% acetic acid (Fontana *et al.*, 1973*b*). Haem was removed as required with *o*-nitrobenzenesulphenyl chloride (Fontana *et al.*, 1973*a*).

Peptides were generally purified on cellulose thin-layer plates (Brinkman) by electrophoresis (t.l.e.) in 10% pyridine adjusted to pH6.5 with acetic acid and/or chromatography (t.l.c.) in butan-1-ol/pyridine/water/acetic acid (6:4:4:1, by vol.). Alternatives occasionally used were column chromatography on DEAE-Sephadex A-25 in 5% (v/v) pyridine with a linear acetic acid gradient from 0.05 to 1M and gel filtration through Bio Gel P-10 or P-30 in 0.1M-ammonium bicarbonate.

Peptides were detected on thin-layer plates by the phenanthrenequinone spray for arginine, followed by acidified cadmium–ninhydrin (Yamada & Itano, 1966), and sometimes also by the Ehrlich reagent for tryptophan. Column effluent was monitored at E_{280}.

Peptides were hydrolysed with 6M-HCl vapours in an evacuated desiccator for 18h at 110°C. Amino acids were determined with a Technicon automatic analyser by the procedure of Niece (1975).

The procedure of Edman & Begg (1967) was used initially for the sequence determination of peptides, but this was gradually modified to improve sensitivity and versatility. The final 'fully improved' procedure was substantially different from the initial procedure and has been described elsewhere (Tarr, 1975), as has the special apparatus which was used in all but the initial stage. Intermediate stages of improvement were used for most of the degradations described in the Results section. Although these changes were gradual and not always progressive, it will be convenient to recognize three stages in addition to the initial and final conditions. First-stage improvements consisted of raising the reaction and cleavage temperatures to 50°C, decreasing coupling time to 90min and cleavage time (with trifluoroacetic acid) to 5min, and extracting thiazolinones from a roughly neutral medium (20–40% pyridine) instead of a strongly acidic medium. In stage 2, coupling time was further decreased to 30min, ethyl acetate or dioxan was added to the benzene wash solvent to improve the extraction of unreacted phenylisothiocyanate, cleavage was usually effected with trifluoroacetic acid, but sometimes instead with 90% (v/v) formic acid (50°C for 5min) as this preserved threonine and serine derivatives much better than trifluoroacetic acid and extraction of derivatives was accomplished with benzene/dioxan (1:1, v/v) to avoid the excessive volatility of diethyl ether. Stage-3 improvements substituted conc. HCl for previous cleavage agents (5min at 22°C) and benzene/ethyl acetate (1:2, v/v) for the peroxide-prone dioxan mixture. The fully improved procedure compounded the wash solvents with heptane instead of benzene, giving more-rapid phase separation and thus better recoveries of emulsion-forming peptides; also, a more extractive wash step was added for large and/or polar peptides, which produced cleaner analyses and somewhat more extended degradations. The modification described in Tarr (1975) involving acetone washes of dried material would have aided the degradation of long

peptides, but was not introduced in time for these studies.

Residues were identified directly by t.l.c. of phenylthiohydantoins, the procedure for which was modified to a minor extent during the course of these experiments (see Tarr, 1975), and indirectly by hydrolysis of these derivatives with HI or NaOH (Smithies *et al.*, 1971), followed by amino acid analysis. The intact protein and one peptide were partially sequenced with an Illitron automatic sequenator (Smithies *et al.*, 1971), after which hydrolysis was the only useful method of identification.

Results

Tetrahymena cytochrome *c* was eluted from the gel-filtration column in the same volume as horse heart cytochrome *c* and was pure by the following criteria. It was electrophoretically homogeneous, exhibited only one sequence on Edman degradation, and had an absorbance ratio $\alpha_{red.}$ (552.5)/$280_{ox.}$ of 1.26, consistent with there being one tryptophan residue (Margoliash & Walasek, 1967). Spectral characteristics were the same as previously reported (Yamanaka *et al.*, 1968).

Although oxidized cytochrome was loaded for isoelectric focusing, much reduction occurred during the run, causing extensive streaking toward lower pH values. The principal peak of material, containing mainly the oxidized form, was centred at pH 6.5, and there was a very minor peak of material at pH 5.3. As judged by electrophoresis on cellulose acetate strips at pH 8, no basic cytochrome *c* could be extracted from cultures at any stage of growth.

During the effort to sequence this protein, a large number of peptides were produced, purified to various degrees, and structurally investigated with variable efficiency. For the sake of clarity and brevity a minimum set of peptides necessary to demonstrate the protein sequence was selected for presentation here. Moreover, this set is presented neither in chronological order of determination nor grouped according to type of digest, but rather in a logical progression from *N*-terminus to *C*-terminus. The other peptides, which confirm most of this sequence and contain no unaccountable residues, are described in the Supplementary Publication along with a much more complicated version of Fig. 1 and compositional information on the total protein and parts thereof. Although the mass of information is variable in quality, and those parts least convincing are indicated in the Supplementary Publication, all parts of the sequence are considered reliable.

Several extended partial sequences from the *N*-terminus were established with the automatic sequenator, yielding, except for the absence of information differentiating between acids and amides, the first 24 residues plus a few probable assignments up to residue 33 (Fig. 1). Sequence similarity to other known eukaryotic cytochromes *c* and also cytochrome c_2 of *Rhodospirillum rubrum* was clear, permitting alignment of the *N*-terminal regions of these proteins (Fig. 2). Manual degradation of the protein through seven cycles confirmed the sequenator results. The first two attempts (using the initial and stage-1 sequencing procedures) indicated that residue 6 was at least in part a glutamine, but all of many subsequent analyses of this position have shown only glutamic acid. The tryptic peptide covering residues 4–18 was purified by t.l.c. and t.l.e. in several separate experiments and finally degraded in its entirety by using fully improved Edman procedures; no amides were detected, and the serine at position 15, indicated by low recovery of alanine from an alkaline hydrolysis of the sequenator sample, was confirmed.

The next large sequence block was deciphered in later experiments from the large tryptic haem peptide (residues 19–47). After removal of the haem from the peptide degradation with fully improved procedures was carried all the way to the *C*-terminus; the first 24 residues (19–42) and residue 44 were identified unambiguously (except for the two cysteine residues) by both t.l.c. and hydrolysis with HI, residues 43 and 45 were clear by hydrolysis but obscure by t.l.c. and the identity of residue 46 was probable. The *C*-terminal section of this peptide was established from the products of partial acid hydrolysis (Schultz, 1967). The first attempt to cleave at the three aspartic acid residues (about 400 nmol, 6 h at 110°C in 1 ml of 0.03 M-HCl) produced a large number of products, which were partially purified by t.l.c. Among the relatively non-polar peptides was the major arginine-positive peptide; this was pure, and comprised residues 43–47 (see Fig. 1) thus providing extensive overlap with the results of the extended degradation. All of the other non-polar peptides identified through fully improved sequencing were also from the *C*-terminal half of the haem peptide and could be accounted for by complete cleavage at each side of Asp-34 and between Gly-42 and Gly-43 and partial cleavage of the remaining bonds adjacent to the glycine residues and Ala-38.

The second fragmentation of the haem peptide was performed on a much smaller amount under milder conditions (2 h at 100°C in 0.03 M-HCl) and yielded only a few peptides detectable on t.l.c. The least polar began, as expected, after Asp-34 and was sequenced through residue 46. Because of limited quantity (about 4 nmol of peptide), only t.l.c. identification was attempted, but this was reasonably clear for all except Ile-45, which was obscured by a constant by-product of the degradation. A small amount of arginine was identified by amino acid analysis of the remaining material. Sequence analysis of the second least-polar peptide which began at residue 30 gave the surprising result that

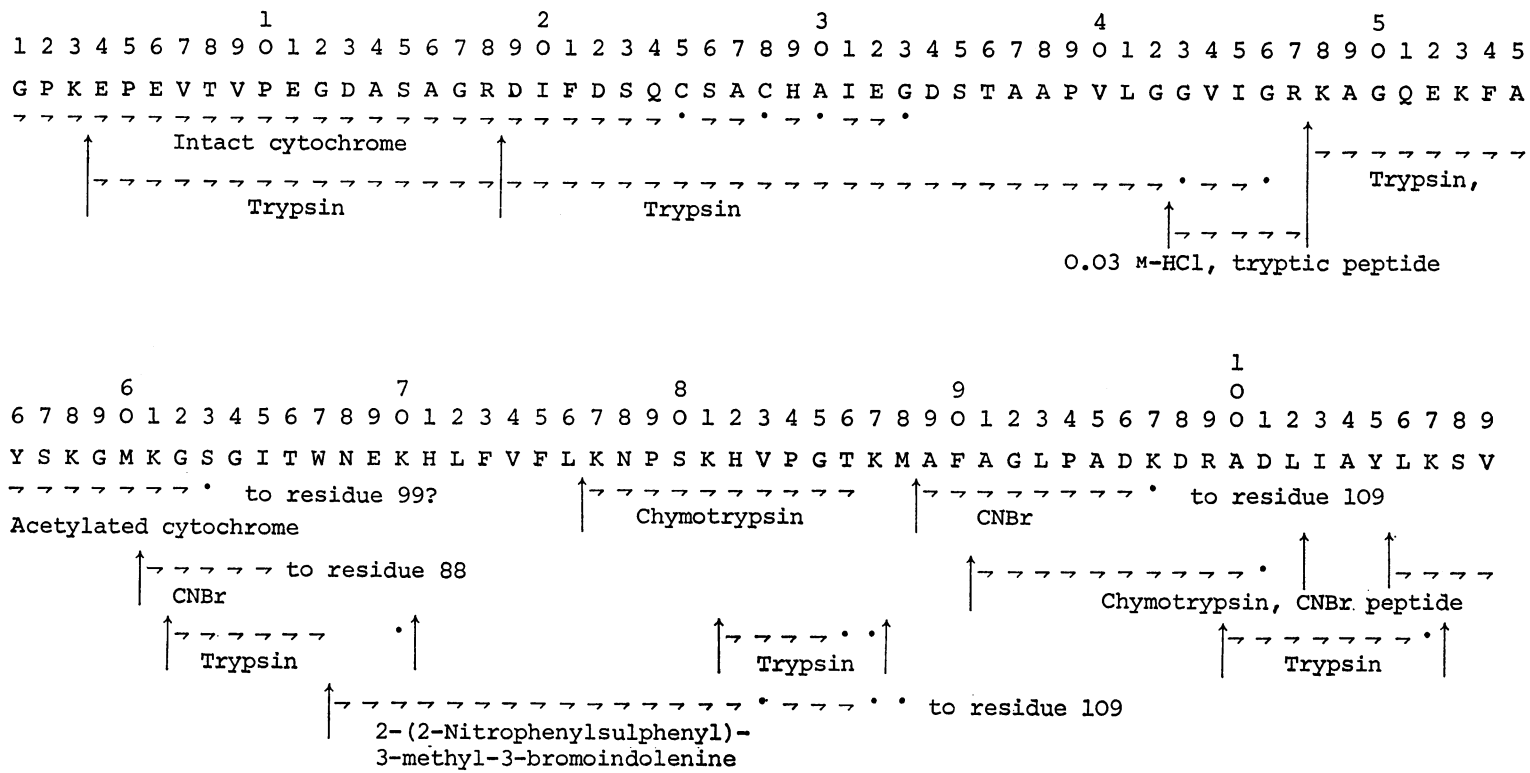

Fig. 1. *Amino acid sequence of cytochrome c from Tetrahymena pyriformis*

Indicated between the limits of the peptide are the cleaving agent and substrate, if other than intact cytochrome. —, Edman degradation with positive identification of residue; ·, sequencing with probable identification; ↑ site of cleavage. Overlap with R 47 was achieved with the major haem-containing chymotryptic peptide from whole cytochrome; this peptide gave a composition consistent with residues 22 through 54; sequencing stopped as usual at the site of haem attachment (residue 25), but after digestion of this degraded material with trypsin, the sequencing resumed at residue 48, although only residues 49–52 could be positively identified. A peptide that is not part of this Figure, but which shows an unusual arrangement, is discussed in the text. Key to the one-letter symbols for amino acids: A, alanine; R, arginine; N, asparagine; D, aspartic acid; C, cysteine; Q, glutamine; E, glutamic acid; G, glycine; H, histidine; I, isoleucine; L, leucine; K, lysine; M, methionine; F, phenylalanine; P, proline; S, serine; T, threonine; W, tryptophan; Y, tyrosine; V, valine; (B, aspartic acid or asparagine; Z, glutamic acid or glutamine; X, unknown).

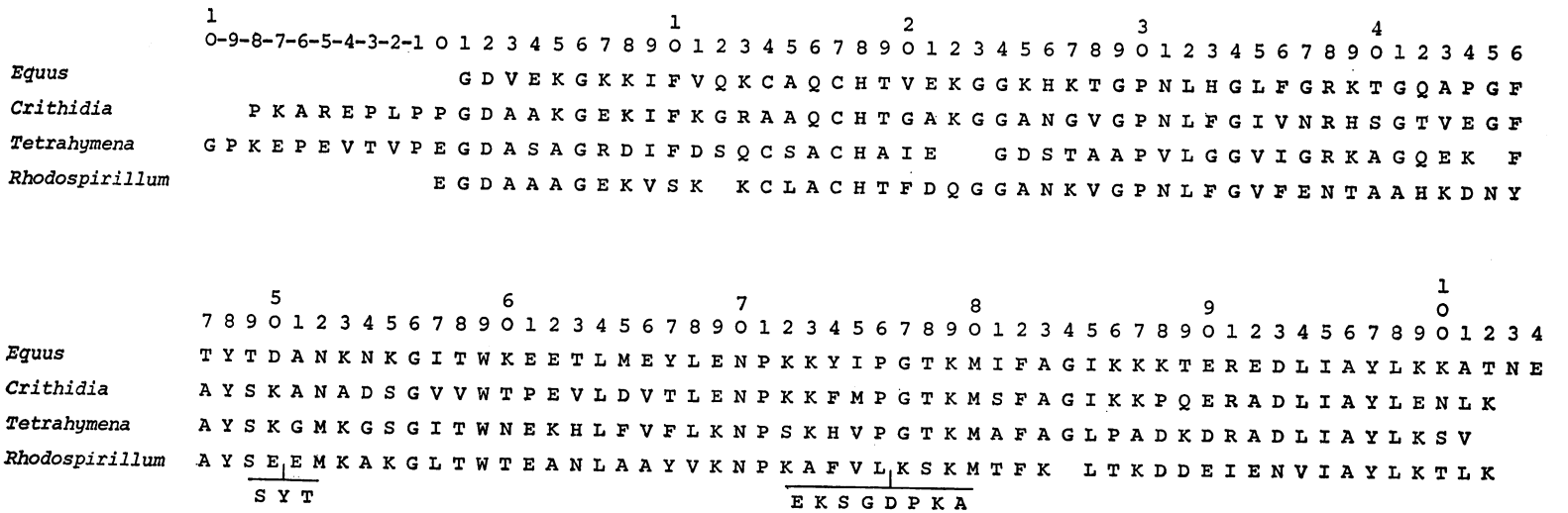

Fig. 2. *Sequence of cytochrome c of T. pyriformis is compared with those for Equus caballus (a mammal), Crithidia oncopelti (a trypanosome) and cytochrome c_2 of Rhodospirillum* rubrum *(a prokaryote)*

The 'universal' numbering system takes the first residue of vertebrate cytochromes *c* as 1, as is customary, but does not omit zero when going from the positive to the negative numbers, which is not customary. Gaps in the *Tetrahymena* and *Rhodospirillum* sequences, and insertions in the latter, are indicated. An additional gap, between residues E-4 and P-3 of *Crithidia*, moving the five *N*-terminal residues to the left, would improve the fit with *Tetrahymena* by five nucleotide replacements but the variability of this region and the time separation involved leave us uncertain whether insertion of the additional gap is appropriate. One-letter symbols for amino acids are defined in the legend to Fig. 1.

Glu-32 was followed directly by Asp-34. As there is no doubt about the true sequence Gly-33 must have been split out by the acid treatment through an unknown type of rearrangement. The composition of the haem peptide was in excellent agreement with the above results. Alignment of this section of the cytochrome with other cytochromes *c* requires a gap of two residues, which was located as shown in Fig. 2 by minimizing mutation distance to other sequences.

Overlap with the next section of the protein was achieved by using the major chymotryptic haem peptide isolated by ion-exchange chromatography (residues 22–54) and sequencing it, with stage-2 improvements (see the Materials and Methods section), up to the haem, beyond which nothing could be obtained. The remaining material was treated with trypsin and the entire product re-sequenced. As cleavage was restricted to the single arginine residue, the observed sequence must begin at residue 48; however, lysine residues were not recovered, so only cycles 2–5 could be identified.

Residues 48–62 were unambiguously established by sequencing with fully improved procedures the insoluble material from the tryptic digest of completely acetylated cytochrome. Minor peptides corresponding to residues 19–47 and 100–109 were also observed. Residues were identified by both t.l.c. and HI hydrolysis. A gap of one residue between 53 and 54 is required to align this portion of the cytochrome with other cytochromes *c*.

The CNBr-cleaved fragment corresponding to residues 61–88 was lost during final purification, but the complete mixture of CNBr-cleaved products and the partially purified mixture of peptide 61–88 with the *N*-terminal peptide were both sequenced several cycles with first-stage improvements. By subtracting the sequence of the known peptides (see evidence elsewhere), the first five residues of this peptide could be determined. Overlap and sequence up to the next large block of residues was provided by the single tryptophan-positive peptide isolated from an unrestricted tryptic digest of the cytochrome. The automatic sequenator was used here, with the first cycle using 4-sulphophenylisothiocyanate to make a hydrophilic residue out of the *C*-terminal lysine. Quantities fell rapidly, however, and identifications could be made after alkaline hydrolysis only through Trp-67.

2-(2-Nitrophenylsulphenyl)-3-methyl-3-bromoindolenine cleaved at Trp-67 with about 40% efficiency and removed the haem (Fontana *et al.*, 1973*b*). As expected, the first two of three fractions from gel filtration through Bio-Gel P-30 possessed *N*-terminal sequences corresponding to the *N*-terminus of the protein, so the third fraction contained residues 68 to the *C*-terminus. This last peptide was degraded 22 cycles with fully improved procedures, except that cleavage time was not long enough to keep the degradation in phase at glycine and proline residues. The first 15 cycles were unambiguous by t.l.c., as were cycles 17–19 (residues 84–86 of the cytochrome). The sequence up through Met-88 was clarified with the chymotryptic peptide comprising residues 77–88, which was isolated by ion-exchange chromatography and t.l.c., and the tryptic peptide containing residues 82–87, purified by t.l.e. and t.l.c. The first ten residues of the former peptide and the first four of the latter, obtained through stage-3 and stage-1 improved Edman procedures respectively, were identified by t.l.c. and HI hydrolysis respectively. Residues 87 and 88 could then be established from the compositions of the two peptides.

The *C*-terminal peptide from CNBr cleavage of the cytochrome was purified by gel filtration through Bio-Gel P-10 and a sample of the peptide sequenced 12 cycles with stage-2 improvements. Identification was clear only through Asp-96. The remaining sample was digested with chymotrypsin, and the products were purified by t.l.c. The sequence of the nearly neutral arginine-positive fraction was unambiguous through ten cycles (residues 91–100), whereas the positive peptide comprising the last four residues of the protein was completely sequenced. The tryptic peptides covering residues 100–107 were also sequenced completely by using stage-3 procedures. The composition of the *C*-terminal CNBr-cleaved peptide was exactly accounted for by the sequence thus established, which therefore requires Val-109 to be the *C*-terminal residue.

Discussion

A comparison of the composition of this cytochrome sequence was made with that determined from a 20h/110°C/HCl hydrolysis. Small corrections (large for Tyr) for recoveries were made using a hydrolysate of horse cytochrome *c* as a standard. Rounding off to the nearest integer left six amino acids differing by one from the sequence value. Similar degrees of ambiguity applied to other large peptides and no significance beyond analytical error and minor contamination is attached to these discrepancies. Therefore except in two or three cases where composition provided useful confirmation, the evidence for the sequence resides entirely in the Edman results themselves. However, the validity of the sequence is thought not to be compromised: every section of the protein has been sequenced at least twice, and two different methods of identifying residues have been used with identical assignments. The reader may not share our confidence in direct t.l.c. identification because of its 'subjective' elements, namely experience required for proper evaluation, results not intrinsically quantitative, and absence of direct documentation. Analysis of

HI hydrolyses had different properties, not all advantageous: poor discrimination of acids/amides and of serine/alanine, loss of methionine, interference by extracted peptide, and one-order-of-magnitude lower sensitivity than t.l.c. Our confidence in the correct determination of this protein sequence is given with full awareness of the sometimes extensive errors made by others and an overriding concern with presenting our phylogenetic studies with absolutely correct information.

Cytochrome *c* of *T. pyriformis* differs considerably from that of other known eukaryotic mitochondrial cytochrome *c*. Specific items include the acidic isoelectric point [which Yamanaka *et al.* (1968) interpret as an affinity with bacterial cytochromes], the two gaps required for alignment with other cytochromes *c* (only the *Euglena* protein possesses a similar internal deletion), the extra long *N*-terminal portion and short *C*-terminal portion, the absence of a basic residue just before the haem attachment (residue 23), the absence of glutamine within the haem attachment (residue 26), the absence of lysines at residues 72 and 86, and the absence of an aromatic residue at 74 (residue numbering as in Fig. 2). Also, this *Tetrahymena* cytochrome is more similar to cytochrome c_2 of *Rhodospirillum* than are other eukaryotic cytochromes *c*. A more precise and detailed comparison will be published separately, but from those items alone it would seem likely that all other eukaryotes whose cytochromes *c* have been sequenced are more closely related to each other than to *Tetrahymena*.

There is, of course, the alternative possibility that the *Tetrahymena* cytochrome *c* is not orthologously related to the other eukaryotic cytochromes *c*, that is, not descended from a single gene in the primaeval eukaryote. That it is, instead, paralogously related (i.e., descended from separate genes produced by gene duplication within the preceding prokaryotes) would seem unlikely, since its counterpart in other eukaryotic organisms has not, apparently, been observed despite the isolation and sequencing of more than 60 other cytochromes *c* and the isolation and functional testing of many others. Neither has a diligent search for a more basic cytochrome *c* in *Tetrahymena* proved fruitful. There are, of course, two known instances of paralogous cytochromes *c* in *Neurospora* (iso-1 and iso-2) and mouse (testes) (Stewart *et al.*, 1966; Borden & Margoliash, 1976; Hennig, 1975),

No sequence polymorphism was observed in these experiments, but this may be due only to the care taken to ensure genetic homogeneity. Comparative electrophoresis of many enzymes from different strains of amicronucleate *T. pyriformis* (Borden *et al.*, 1973) strongly suggests that other strains will be found to possess cytochromes *c* with different sequences. This probability plus ease of its culture and its high cytochrome *c* concentration make this organism an attractive subject for future sequence study.

We thank Dr. Ronald L. Niece for his invaluable assistance, and acknowledge Public Health Service Fellowship no. 5 FO2 GM51532 for financial support to G. E. T., and NSF grant GB32274X2 and NIH grant AM15282 to W. M. F.

References

Borden, D. & Margoliash, E. (1976) in *Handbook of Biochemistry: Proteins* (Fasman, G. D., ed.), vol. 3, pp. 268–268, Chemical Rubber Co., Cleveland, OH

Borden, D., Whitt, G. S., & Nanney, D. L. (1973) *J. Protozool.* **20**, 693–700

Brown, R. H., Richardson, M., Boulter, D., Ramshaw, J. A. M. & Jeffries, R. P. S. (1972) *Biochem. J.* **128**, 971–974

Dougherty, E. C. (ed.) (1963) *The Lower Metazoa*, University of California Press, Berkeley

Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91

Fitch, W. M. & Margoliash, E. (1967) *Science* **155**, 279–284

Fontana, A., Veronese, F. M. & Boccu, E. (1973*a*) *FEBS Lett.* **32**, 135–138

Fontana, A., Vita, C. & Toniolo, C. (1973*b*) *FEBS Lett.* **32**, 139–142

Hennig, B. (1975) *Eur. J. Biochem.* **55**, 167–183

Kidder, G. W. & Dewey, V. C. (1951) in *Biochemistry and Physiology of Protozoa* (Lwoff, A., ed.), pp. 324–400, Academic Press, New York

Levy, M. R. (1973) in *Biology of Tetrahymena* (Elliott, A. M., ed.), pp. 227–757, Dowden, Hutchinson and Ross, Stroudsburg, PA

Lin, D. K. (1971) Ph.D. Thesis, University of Wisconsin

Lin, D. K., Niece, R. L. & Fitch, W. M. (1973) *Nature (London)* **241**, 533–535

Margoliash, E. & Walasek, O. F. (1967) *Methods Enzymol.* **10**, 339–348

McLaughlin, P. J. & Dayhoff, M. O. (1973) *J. Mol. Evol.* **2**, 99–116

Niece, R. L. (1975) *J. Chromatogr.* **103**, 25–32

Pettigrew, G. W. (1972) *FEBS Lett.* **22**, 64–66

Pettigrew, G. W. (1973) *Nature (London)* **241**, 531–533

Schultz, J. (1967) *Methods Enzymol.* **11**, 255–263

Shrago, E., Shug, A. L. & Ferguson, S. M. (1971) *Int. J. Biochem.* **2**, 312–318

Smithies, O., Gibson, D., Fanning, E. M., Goodfliesch, R. M., Gilman, J. G. & Ballantyne, D. L. (1971) *Biochemistry* **10**, 4912–4921

Stewart, J. W., Margoliash, E. & Sherman, F. (1966) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **25**, 647

Tarr, G. E. (1975) *Anal. Biochem.* **63**, 361–370

Yamada, S. & Itano, H. A. (1966) *Biochim. Biophys. Acta* **130**, 538–540

Yamanaka, T., Nagata, Y. & Okunuki, K. (1968) *J. Biochem. (Tokyo)* **63**, 753–760