Journal of Bioenergetics and Biomembranes, Vol. 20, No. 2, 1988

MINI-REVIEW

Amino Acid Identities in the Three Redox Center-Carrying Polypeptides of Cytochrome bc_1/b_6f Complexes

G. Hauska,[1] W. Nitschke,[1] and R. G. Herrmann[2]

Received August 18, 1987

Abstract

The comparison of primary structures is extended to 22 cytochromes *b* or b_6, 12 cytochromes c_1 or *f*, and 8 Rieske FeS proteins. Conclusions are drawn as to their phylogenetic relationship as well as on conserved, functionally important amino acids and secondary structures. The results are in favor of two independent quinone binding sites at opposite surfaces of the membrane, topping one of the two hemes of cytochrome *b* each.

Key Words: Cytochrome bc_1 complex; cytochrome b_6f complex; Rieske FeS protein; amino acid sequences; quinone-binding peptides; membrane proteins; organelle gene evolution.

Introduction

Cyt bc_1/b_6f complexes function as quinol-cyt *c* or plastocyanin oxidoreductases in many respiratory and all photosynthetic electron transport chains so far studied (Hauska *et al.*, 1983; Hauska, 1985, 1986a). Active b_6f complexes from chloroplasts and cyanobacteria, and bc_1 complexes from photosynthetic bacteria, can be isolated by a general procedure (Hauska, 1986b), which has

[1]Institut für Botanik, Universität Regensburg, Universitätsstrasse 31, D-8400 Regensburg, FRG.
[2]Botanisches Institut der Ludwig-Maximilians-Universität, Menzingerstrasse 67, D-800 München 19, FRG.

0145-479X/88/0400-0211$06.00/0 © 1988 Plenum Publishing Corporation

been improved recently (Ljungdahl *et al.*, 1986). These complexes are remarkably simpler than their mitochondrial counterparts (Schägger *et al.*, 1986). Essentially only three polypeptides are found, which represent cyt *f* or c_1, cyt *b*, and the Rieske FeS protein (Hauska *et al.*, 1983; Hauska, 1985, 1986a). In the b_6f complexes the cyt *b* component (ca. 42 kD) is split into cyt b_6 (23 kD) and subunit IV (15.2 kD) (Heinemeyer *et al.*, 1984). In addition to electron transport from quinol to cyt *c* or plastocyanin, these complexes translocate protons charging the membrane. The mechanism of this reaction is complex, and different concepts—Q cycle, *b* cycle, semiquinone cycle—are discussed (Hauska *et al.*, 1983; Crofts, 1985). Recently, a whole issue of *J. Bioenerg. Biomembr.* (No. 3, 1986) has been devoted to this matter.

Earlier studies on the accessibility for antibodies, and analyses of the primary structures, led to a universal topographical model for the bc_1/b_6f complexes (Hauska *et al.*, 1985; Hauska, 1985, 1986a; see Gabellini and Link *et al.*, in Colson, 1987). It is especially detailed for the cyt *b* component, because several sequences were compared, and because hydrophobic spans suggest that it folds nine times through the membrane (Saraste, 1984; Widger *et al.*, 1984; see Crofts in Colson, 1987). This model is in striking agreement with current functional concepts.

Recently this study has been extended from six (Saraste, 1984; Widger *et al.*, 1984) to eleven different cyt b/b_6 sequences, with the aim of predicting functional centers, especially quinone binding sites, and to find indications for the different inhibitor sensitivities of bc_1 and b_6f complexes (Nitschke and Hauska, 1987a; Lang *et al.*, 1985; Schuster and Brennicke, 1985). The sequence of cyt *b* from *Trypanosoma* (Benne *et al.*, 1983) was of particular importance in pinpointing essential residues. It shows only about 20% identity to any of the other sequences, but the functional characteristics, like antimycin sensitivity (Markos *et al.*, 1986), seem to be retained.

Here we extend the sequence comparison to 22 cyt *b* sequences. In addition we compare 12 cytochromes c_1 or *f*, and 8 Rieske FeS proteins. The sequences are aligned for maximal identity, without penalties for deletions. This might somewhat distort evolutionary distances. We hope that besides its contribution to our understanding of the function and evolution of the bc_1/b_6f complexes, such a study will be of value for directed mutagenesis.

Cytochrome $b = b_6$ + subunit IV

Cytochrome *b* leads membrane proteins with respect to the number of elucidated primary structures (Fig. 1). By now these are known for 14 mitochondria—man = H (Anderson *et al.*, 1981a), beef = B (Anderson

et al., 1981b), mouse = M (Bibb *et al.*, 1981), rat = R (Koike *et al.*, 1982), toad = T (Roe *et al.*, 1985), the protozoa *Trypanosoma brucei* = Tr (Benne *et al.*, 1983) and *Leishmania tarantolae* (de la Cruz *et al.*, 1984; not shown), the yeasts *Saccharomyces cerevisiae* = Y (Nobrega and Tzagoloff, 1980) and *Schizosaccharomyces pombe* (Lang *et al.*, 1985; not shown), the fungi *Neurospora crassa* = Nc (Fig. 1 shows the partial sequence of Burke *et al.*, 1984, because of higher identity, completed by the data from Helmer-Citterich *et al.*, 1983) and *Aspergillus nidulans* = A (Waring *et al.*, 1981), and the plants maize = Mz (Dawson *et al.*, 1984), wheat = Wh (Boer *et al.*, 1985), and Oenothera = Oe (Schuster and Brennicke, 1985). Furthermore, Fig. 1 shows the sequences of cytochromes *b* for three bacteria—*Paracoccus denitrificans* = Pc (Kurowski and Ludwig, 1987) and two strains of *Rhodobacter capsulatus* = Rc1, Rc2 (Gabellini and Sebald, 1986; Davidson and Daldal, 1987a); originally Rc1 was considered to be *Rhodopseudomonas sphaeroides* GA (Gabellini and Sebald, 1986), but later was identified as *R. capsulatus* (Davidson and Daldal, 1987b; see also Gabellini, 1988)—and for three chloroplasts (b_6 + subunit IV)—spinach = Sp (Heinemeyer *et al.*, 1984), tobacco = Tb (Shinozaki *et al.*, 1986), and liverwort = Lw (Ohyama *et al.*, 1986). In addition, a partial sequence for b_6 + subunit IV of the cyanobacterium *Nostoc* (Kallas *et al.*, 1986; R. Malkin, personal communication; not shown) and the sequence of subunit IV from pea (Phillips and Gray, 1984) are known.

In Fig. 1, 19 of these sequences are aligned in blocks of 75 amino acids in single-letter code. Mistakes found in earlier alignments (eight in Saraste, 1984 and two in Widger *et al.*, 1984) have been corrected for the original publications. The N-terminal sequences of spinach b_6 and subunit IV have been changed from previous presentations (Heinemeyer *et al.*, 1984; Widger *et al.*, 1984; Nitschke and Hauska, 1987a) after the discovery of introns (Shinozaki *et al.*, 1986; Ohyama *et al.*, 1986; Westhoff and Herrmann, unpublished). This correction has been confirmed for spinach by amino acid sequencing (Pfefferkorn and Meyer, 1987), the initial M being cleaved off in b_6 as well as in subunit IV. Pair identities and nonconservation exchanges are indicated by vertical and horizontal lines, respectively, leaving conservative exchanges within the groups QNED, RKH, VLIMF, GAP, FYW, ST, and C (Dayhoff, 1978) unmarked. Complete identity (large dots) and complete conservation of related residues (small dots), including *Trypanosoma*, is shown at the bottom, and that excluding *Trypanosoma* is shown between Rc_2 and Lw. For the bc_1 complexes only, the same is shown at the top and between Y and Nc. Identity of the b_6/subunit IV sequence is about 90% (Table I) and useless to mark. The nine hydrophobic spans, probably representing transmembrane helices (Saraste, 1984; Widger *et al.*, 1984; Nitschke and Hauska, 1987a), are lined on top and bottom, indicated by

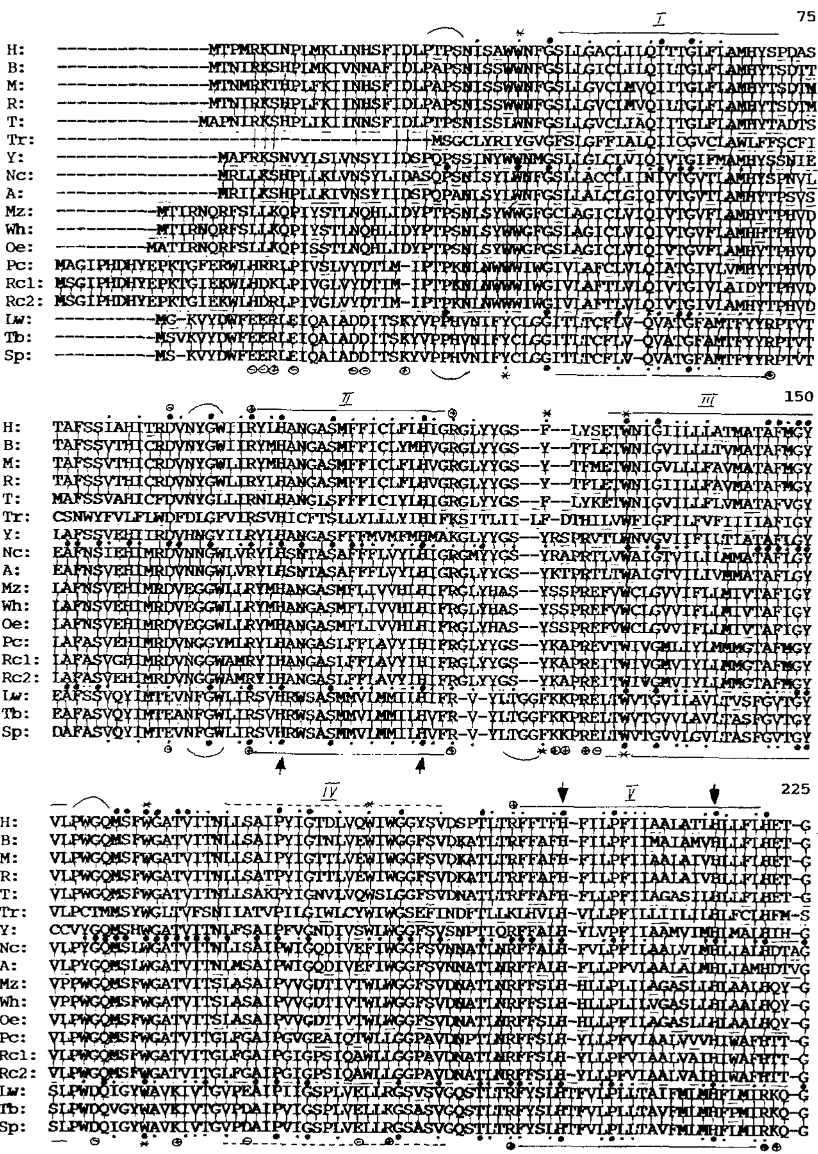

Fig. 1. Alignment of cyt b/b_6 + subunit IV (see text for description).

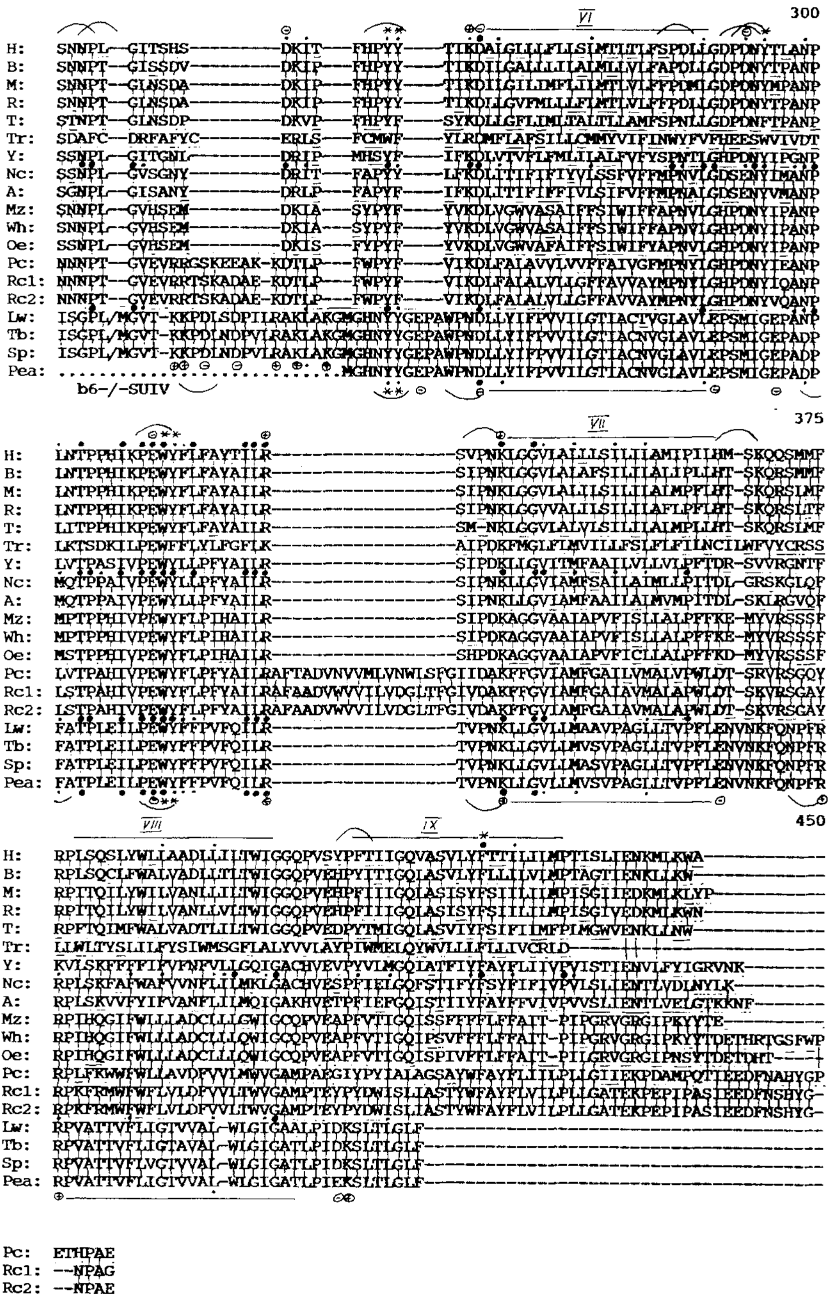

Fig. 1. Continued.

roman numbers. The four conserved histidines probably binding the two hemes between helices II and V (Saraste, 1984; Widger *et al.*, 1984; Nitschke and Hauska, 1987a) are marked by arrows. The rather polar span IV, possibly an amphipathic helix, is dashed and might actually not span the membrane. This is favorable in view of inhibitor resistance sites on the cyt *b* gene in yeast mutants (Colson, 1987, and below). If this is true, the heme binding helices II and V would run in parallel and not antiparallel (see Link *et al.* in Colson, 1987) through the membrane. In addition, conserved charges are marked in Fig. 1, and conserved aromatic residues and β-turn tetrapeptides of high probability (Chou and Fasman, 1979) are indicated by asterisks and arcs, respectively—on the top for the bc_1, on the bottom for the b_6f complexes.

Besides the hydrophobic spans and the four conserved histidines in helix II and V, which are stabilized by positive charges at the membrane surfaces, possibly interacting with the carboxyethyl groups of the heme rings, the other essentials of the earlier comparisons (Saraste, 1984; Widger *et al.*, 1984; Nitschke and Hauska, 1987a) are also retained in Fig. 1. There are eight conserved glycines six of them within hydrophobic spans. These are the positions 59 and 83 in I, 91 in the β-turn before II, 135 and 149 in III, 176 and 185 in IV, and 347 in VII. Three conserved prolines are found: one in span IV at 173, one in the heme binding helix V at 206, and one in a β-turn at 309. The one in helix V is of particular interest because it has been suggested that *cis/trans*-isomerization of prolines might play a role in the vectorial function of membrane proteins (Brandle and Deber, 1986). Hydroxylated amino acids are conserved at position 103 in the heme binding helix II, and at positions 193, and 303. Conserved aromatic residues are at positions 56, 124, 132, and 150, both in III, and 160, 258, 259, 311, and 312. Besides the helix stabilizing positive charges 95 and 115 at the boundaries of II, 196 at V and 344 at VII, and the negative charge 267 at VII, one additional positive charge at 321 and two negative charges at 87 and 310 are found. A striking region of identity includes the peptides PEWY312 and ILR321, between helix VI and VII. We suggest that this might form part of the quinol oxidizing site Qo (Nitschke and Hauska, 1987a), with E310 stabilizing the positive charge at R321 which might bind the semiquinone anion (Rich, 1984). In addition, we consider YY260 to be involved in the quinone reduction site Qi (Nitschke and Hauska, 1987a), which is sensitive to antimycin in bc_1 complexes. The insensitivity of the b_6f complexes to this inhibitor could be related to the split of cytochrome *b* into b_6 and subunit IV in this region. Further support comes from the mapping of diuron resistance at Y259 in yeast (Colson, 1987). A second diuron resistance site is at the aromatic residue 56, and unless long distance interactions are invoked, positions 56 and YY260 should be close to the Qi-binding pocket. Conse-

quently, as mentioned above, previous folding models of cyt *b* need revision, span IV possibly not traversing the membrane (Saraste, 1984; Widger *et al.*, 1984; Nitschke and Hauska, 1987a; see Crofts in Colson, 1987). Still, our observation that all conserved aromats are found close to the two membrane surfaces (Nitschke and Hauska, 1987a) holds, and is in favor of two independent sites Qo and Qi, as pictured in the Q cycle (Hauska *et al.*, 1983; Crofts, 1985). Support for this notion comes also from a mutant of *Rhodobacter capsulatus* with an inactive Qo-site, but an unchanged Qi-site (Robertson *et al.*, 1986). It might be important for the Qi-site, as well as for the Qo-site, that YY260 and PEWY312 both lie in β-turns. Other conserved β-turns are at P40, at G91 just before the heme carrying span II, at K344 before span VII, and after span VII.

Cyt *b* from *Trypanosoma* is very different, showing not more than 25% identity with any of the other sequences (Table I). It is very rich in cysteins (Benne *et al.*, 1983). Special residues that are different from conserved positions in all the others, including the b_6f complexes, are: S40, F229, T300, S304, and I362, all instead of prolines; instead of aromats there are T118, C154, and LH198 instead of FF in helix V; a conserved tetrapeptide AFXS80 is SNWY in *Trypanosoma*. An important finding is that F105 in all the other bc_1 complexes is L in *Trypanosoma*. This F was thought to facilitate electron transfer between the two heme groups (Widger *et al.*, 1984), and was considered to explain the differences in the two types of complexes since it is replaced by M in cyt b_6.

The value of comparing more and more sequences for exclusion of putatively important conservations is shown also by other examples: Plant mitochondria (Mz, Wh, and Oe) have a histidine at position 317 where all the others have an aromat. Another H is found at position 110 in the heme binding helix II, where the rest of the bc_1 complexes have an aromat, and the b_6f complexes have an I. Even more surprising is H at position 70 in wheat (Wh), where all the other sequences have an aromat. Could this be a mistake in sequencing? The same question arises for L92 in the sequence from toad (T).

Cyt b_6 + subunit IV is shorter at the C-terminus than cyt *b*. The hydrophobic span IX is missing. The bc_1 complexes show a conserved F427 in this span, which also might be involved in the difference at the Qi site between the two types of complexes. If span IV indeed is outside the membrane (see above), the cyt *b*-complement of the b_6f complex again might have the seven-helix motif generally postulated for proteins with transmembrane function (Dixon *et al.*, 1986). The N-terminus of subunit IV is highly charged, like the N-termini of cyt *b*. It contributes to a comparatively extended region between span V and VI in cyt b_6 + subunit IV which might additionally affect the Qi site. Interestingly, this extension is partially shared by bacterial cyt *b*,

Table I. Percent Identity in Pairs of Cyt b/b_6 + IV (see Text for Description)

		bc_1														b_6f			
		H	B	M	R	T	Tr	Y	Nc	A	Mz	Wh	Oe	Pc	Rc	No	Lw	Tb	Sp
bc_1	H	100																	
	B	81	100																
	M	79	88	100															
	R	81	88	94	100														
	T	74	78	78	78	100													
	Tr	23	23	22	23	25	100												
	Y	51	52	53	53	50	21	100											
	Nc	55	56	57	57	53	22	61	100										
	A	54	56	58	57	55	24	61	86	100									
	Mz	53	56	55	55	54	24	54	55	56	100								
	Wh	54	56	55	55	54	24	54	54	55	99	100							
	Oe	53	56	55	55	54	25	55	56	56	97	96	100						
	Pc	46	48	47	47	47	20	46	49	49	53	52	53	100					
	Rc	46	48	48	48	48	21	44	50	50	52	51	53	84	100				
b_6f	No	28	29	29	28	28	16	23	29	29	29	28	30	28	33	100			
	Lw	28	28	28	28	27	17	26	31	31	31	31	32	28	28	86	100		
	Tb	28	28	29	28	27	18	25	31	31	31	30	32	28	28	86	95	100	
	Sp	28	28	28	27	27	17	25	31	31	31	30	32	28	28	86	95	98	100

which additionally has an insertion after ILR321, close to the suggested Qo site between VI and VII. As pointed out before (Gabellini and Sebald, 1986), these insertions might reflect the lack of further subunits in plastidal and bacterial complexes.

Further, possibly important, differences between cyt b and b_6 + IV are: b_6 has a positive charge, R99, next to the heme binding H98 in span II; there is IG instead of MS158 in III, K instead of T163, L182 instead of an aromat in IV, which interestingly also is L in *Schizosaccharomyces pombe* (Lang *et al.*, 1985; not shown), and R instead of H221. A difference emphasized before (Widger *et al.*, 1984; Nitschke and Hauska, 1987a) is the insertion of T202 after H201 in helix V. This additional residue between the two hemes must exert an additional helical torsion and might explain why cyt b_6 has a lower redox potential and is less stable than cyt b (Nitschke and Hauska, 1987a, 1987b).

Table I gives the percentage identities for the cyt b sequences counted from position 38 to 411 in Fig. 1, deletions included. The percentages are

similar to the ones presented before for a smaller number of sequences (Lang *et al.*, 1985). The partial sequence (about 80%) for cyt b_6 + subunit IV from the cyanobacterium *Nostoc* (Malkin, personal communication; not shown in Fig. 1) has also been evaluated (No in Table I). The abbreviations are as for Fig. 1. The main observations are: (1) Cyt b_6 + subunit IV, including the cyanobacterial sequence, show about 30% identical residues to any other cyt *b*. (2) The protozoon *Trypanosoma* shows even less identity, which points to a very high mutation rate of mitochondrial genes in this parasitic organism with its special organization of the mitochondrial genome (Benne *et al.*, 1983). Interestingly, cyt *b* from the related organism *Leishmania tarantolae* (de la Cruz *et al.*, 1984) is 84% identical to the one from *Trypanosoma*. (3) Percentages of identity are significantly lower in vertebrate mitochondria than in plant chloroplasts, although species radiation occurred later, especially if one considers liverwort (Lw). (4) Even more pronounced is the higher conservation between cyanobacteria and plastids (86%) compared to bacteria and mitochondria (around 50%). Is the evolutionary clock of plastidal genes that come from cyanobacteria slower than the one of mitochondrial genes originating from bacteria? This is not the case because the sequences of mitochondrial cyt *b* of higher plants are as conserved as plastidal cyt b_6 (Table I). Probably plant organellar DNA has a comparatively slow mutation rate, possibly being better protected from mutagens like UV light. (5) Yeast is rather distant from filamentous fungi (Nc and A), as noted before (Lang *et al.*, 1985).

How has cyt *b* evolved? Is there any indication of domain shuffling (Gilbert *et al.*, 1986)? The splitting into cyt b_6 and subunit IV might indicate the evolutionary fusion of a heme-binding domain with a quinone-reaction domain. Indeed, the cyt *b* gene of *Schizosaccharomyces pombe* shows an intron in this region (Lang *et al.*, 1985), after K266. The gene in *Saccharomyces cerevisiae* is split by five introns, however. These are after Q161, G166, G186, G290, and, just before the conserved tetrapeptide, PEWY321 (Colson, 1987). The *Neurospora* gene is split twice, after G149 and E181, and the one from *Aspergillus* has an intron after G186, at the same site as the third intron of yeast (Burke *et al.*, 1984; Waring *et al.*, 1981). These introns separate the two heme binding helices II and V, reminiscent of the two peptides forming plastidal cyt *b*559 (Brown *et al.*, 1982). So the origin from ancestral domains might be more complicated than indicated by cyt b_6 and subunit IV. Interestingly the N-terminal introns found in the plastidal genes of cyt b_6 and subunit IV (Shinozaki *et al.*, 1986; Ohyama *et al.*, 1986) are absent in the cyanobacterium (Kallas *et al.*, 1986). There are no particularly conserved regions in cyt *b*, like in the Rieske FeS protein, although less identity is observed at the N- and C-terminus, as in other proteins.

The Rieske FeS Protein

The primary structure of the Rieske FeS protein from the b_6f complex of spinach chloroplasts has been obtained by sequencing genomic DNA (Steppuhn *et al.*, 1987). It confirms the partial sequence obtained directly from the protein (Pfefferkorn and Meyer, 1986). Figure 2 shows its alignment with three Rieske FeS proteins from mitochondria—beef = B (Shägger *et al.*, 1987), *Neurospora crassa* = Nc (Harnisch *et al.*, 1985), and yeast = Y (Beckmann *et al.*, 1987), and four from bacteria—*Paracoccus denitrificans* (Waring *et al.*, 1981), two strains of *Rhodobacter capsulatus* = Rc1 and Rc2 (Gabellini and Sebald, 1986; Davidson and Daldal, 1987a), and a partial sequence of *Rhodobacter sphaeroides* (Davidson and Daldal, 1987b). The arrangement and the indications of special residues and structures are as

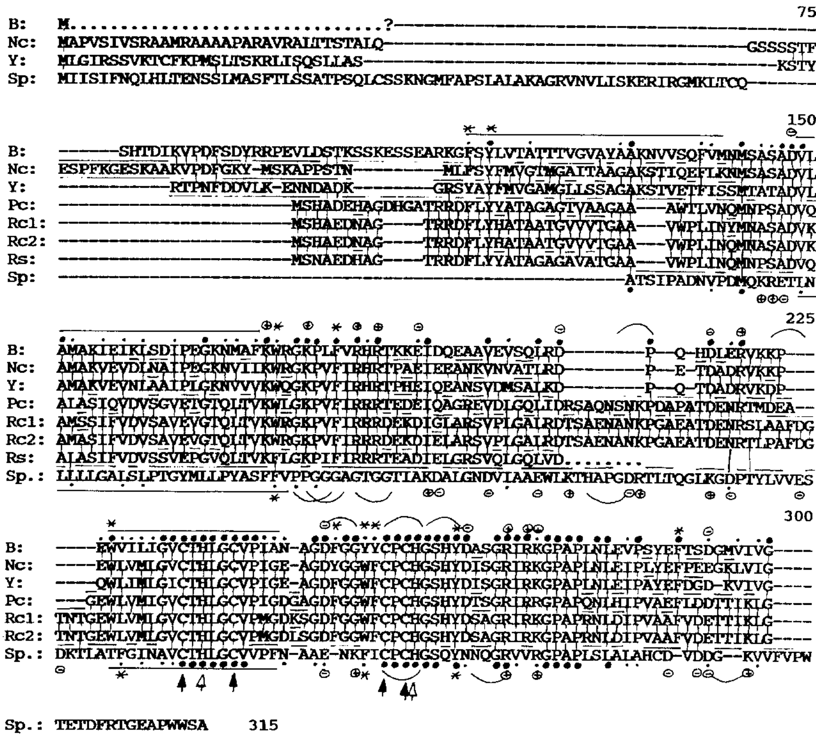

Fig. 2. Alignment of Rieske FeS protein (see text for description).

for Fig. 1. Identical residues (large dots) and conservative exchanges (small dots) are shown at the bottom for all eight sequences, on top for the seven bc_1 types.

The chloroplast protein (Sp) is considerably shorter at the N-terminus and longer at the C-terminus, but also the bacterial proteins are shorter at the N-terminus than the mitochondrial ones, which are rather variable in this region. In addition they have deletions after D200, in the middle part. The eukaryotic proteins are synthesized on cytoplasmic ribosomes, and leader peptides are shown in the first block of Fig. 2. The bovine leader peptide is not known because the sequence has been obtained directly from the mature protein (Schägger *et al.*, 1987). The double processing of the protein from *Neurospora* (Nc) on its path to its mitochondrial subcompartment (Hartl *et al.*, 1986), and of the chloroplast protein to the intrathylakoid space (Steppuhn *et al.*, 1987), will not be discussed here. Noteworthy is the fact that the leader peptide for the plastidal target is more than double the length of the mitochondrial counterparts. Perhaps this reflects the sorting of related protein precursors for different organelles in the plant cell. The bacterial Rieske FeS proteins probably lack leader peptides.

There are two hydrophobic regions, close to the N-terminus of the bc_1 types, each of them long enough to span the membrane. These might form a transmembrane hairpin (Schägger *et al.*, 1987; see also Link *et al.* in Colson, 1987). The hydrophobicity is not very high, though, in line with the relatively easy dissociation of the Rieske FeS proteins from the complexes. Thus these proteins might actually not be transmembranous (see Gabellini and Neupert in Colson, 1987). Because of N-terminal shortening, the chloroplast protein shares only the second hydrophobic span. This is followed by a P and G-rich span instead of a highly positively charged stretch. Another hydrophobic region, found in all the cases, includes the conserved heptapeptide CTHLGCV249, which binds the Fe2S2 cluster from one side. The other side is bound to the conserved hexapeptide CPCHGS268 downstream, which probably forms a β-turn in a rather hydrophilic region, containing at least one additional β-turn. This highly conserved part has been pointed out before (Gabellini and Sebald, 1986) and additionally includes two aromats at positions 261 and 270, and one negative plus two positive charges at positions 257, 275, and 278. Further conserved features of possible significance are M143, the aromatic residue 172, after the N-terminal hydrophobic region, and a β-turn after D200, in the region of the mitochondrial deletions. The bc_1 types have additional aromats at positions 116, 118, 180, 258, 262, and 287. These might contribute to differences at the quinol oxidizing site Qo, resulting in different inhibitor sensitivities of the bc_1 and b_6f complexes (Hauska *et al.*, 1983).

Table II. Percent Identity in Pairs of Rieske FES Proteins (see Text for Description)

	bc_1					b_6f
	B	Nc	Y	Pc	Rc	Sp
B	100^{100}_{100}					
Nc	64^{57}_{75}	100^{100}_{100}				
Y	63^{53}_{78}	74^{68}_{83}	100^{100}_{100}			
Pc	47^{34}_{68}	49^{36}_{70}	51^{38}_{72}	100^{100}_{100}		
Rc	44^{33}_{62}	48^{37}_{68}	48^{37}_{67}	69^{62}_{82}	100^{100}_{100}	
Sp	20^{6}_{46}	19^{6}_{42}	18^{6}_{41}	19^{5}_{45}	18^{4}_{42}	100^{100}_{100}

The four cysteines holding the Fe2S2 center are shown by full arrows in Fig. 2. ENDOR measurements indicate that nitrogen is also a strong ligand to Fe in the bc_1 complex (Telser *et al.*, 1987). This could not be confirmed for the b_6f complex so far (Nitschke and Hauska, 1987a), but the two conserved histidines, one in each of the cluster binding peptides (open arrows), are very intriguing. These are absent in the Fe-binding peptides of ferredoxins which also otherwise are very different (Dayhoff, 1978; Smeekens *et al.*, 1985).

There is a clearly different degree of conservation in the N- and C-terminal part of the Rieske FeS proteins. This is documented in Table II, which gives the overall percent amino acid identities in large numbers, the one of the N-terminal part, from position 131 to 236, in superscripts, and the C-terminal one, from 237 to 296, in subscripts. Most striking is the difference when comparing the b_6f with the bc_1 types, but also, within the bc_1 types, the C-terminal part is less conserved. Possibly the Rieske FeS protein has different ancestral domains (Gilbert *et al.*, 1986), although no introns have been found so far.

Cytochrome c_1 or f

Figure 3 shows an attempt to align six, cyt c_1-sequences, three mitochondrial—beef = B (Wakabayashi *et al.*, 1980), yeast = Y (Sadler *et al.*, 1984),

and *Neurospora crassa* = Nc (Römisch *et al.*, 1987)—and three bacterial ones—*Paracoccus* = Pc (Kurowski and Ludwig, 1987) and two strains of *Rhodobacter capsulatus* = Rc_1, Rc_2 (Gabellini and Sebald, 1986; Davidson and Daldal, 1987a), with six cyt *f*-sequences—the partial sequence until 391 from the *cyanobacterium Spirulina* maxima = Sm (Alam and Krogmann, 1984), and five sequences from chloroplasts—liverwort = Lw (Ohyama *et al.*, 1986), spinach = Sp (Alt and Herrmann, 1984), tobacco = Tb (Shinozaki *et al.*, 1986), pea (Willey *et al.*, 1984a), and wheat = Wh (Willey *et al.*, 1984b). The sequences of beef and *Spirulina* have been obtained directly from the mature proteins, the rest from DNA. Organization of the figure and symbols are as for Fig. 1. The indication of complete identity (large dots), and of conservative exchanges (small dots), is shown on top for the bc_1 types, on the bottom for the b_6f types, and for all the sequences between Rc2 and Sm. Aromatic residues (asterisks) and charges are indicated on top for c_1 and on the bottom for *f*. Cyt c_1 from four species (Römisch *et al.*, 1987), and of beef with cyt *f* (Davis, 1987), have been aligned before in a different way.

The first block shows the leader peptides. In *Paracoccus* this is followed by an enigmatic sequence of about 180 amino acids, highly acidic and rich in A, P, and T (Kurowski and Ludwig, 1987). The conserved heme-binding hexapeptide VCXXCH, which forms a β-turn close to the C-terminus, is indicated by arrows. Another feature found in all cases is the hydrophobic span at the C-terminus, probably forming a transmembrane helix stabilized by positive charges on either surface (Alt and Herrmann, 1984; Willey *et al.*, 1984a). In addition, several β-turns occur in the same regions in cyt c_1 and *f*. Besides the heme-binding site only the region before the transmembrane helix shows some overall conservation, including the aromat F506. Further conservations comprise several aromats and charged amino acids, but overall identity is very low between cyt c_1 and *f*. Only 13 conservations and 21 conservative exchanges out of 308 residues are seen. As found also for the comparison of cyt *b* with cyt b_6 (see above), the plastidal protein Cyt *f* is much more conserved than cyt c_1, although relatively less than cyt b_6. This is compiled in Table III, which gives the percent pair identities from the N- to the C-terminus of Cyt *f* (position 248 to 556).

Cyt *f* shows several insertions, which in part are also found for bacterial cyt c_1. In particular, the region around position 430, rich in β-turns, is partially identical between bacteria and chloroplasts.

Especially interesting is the lack of a conserved methionine in cyt *f* as a sixth ligand to the heme-Fe. This M is at position 477 for cyt c_1 (arrow). The only conserved M in cyt *f* is at the C-terminus, spaced from the heme by the membrane. An absorption of oxidized cyt *f* around 690 nm, characteristic for

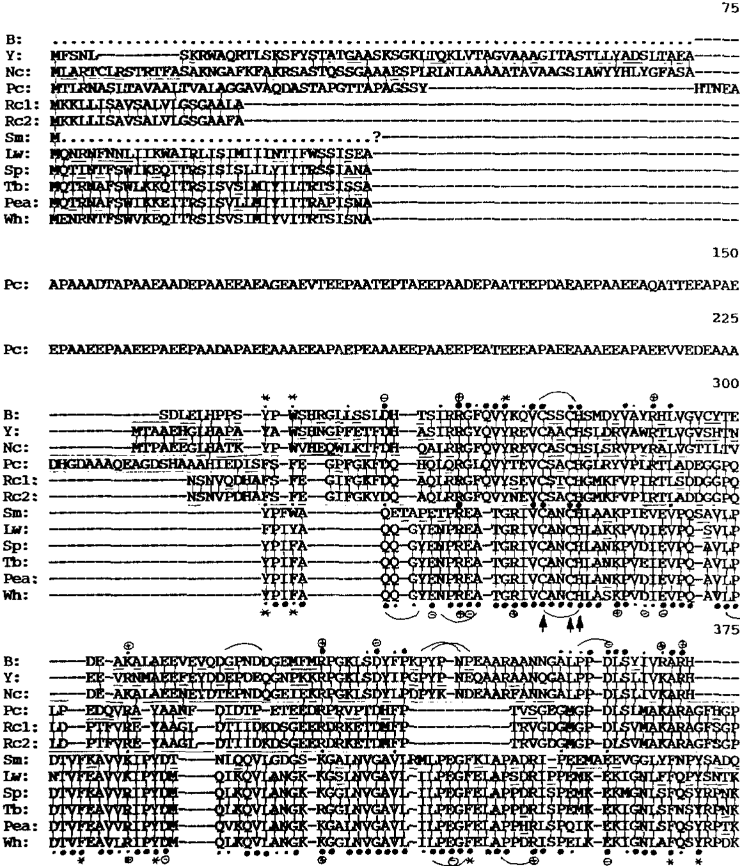

Fig. 3. Alignment of cyt c_1/f (see text for description).

the sulfur ligation of heme *c* (Yu *et al.*, 1986), is seen, although with unusual shape (W. Nitschke, unpublished). The sixth heme ligand of cyt *f* thus is unclear.

Conclusions

1. Cyt *b* is the most highly conserved of the three redox center-carrying subunits of cyt bc_1/b_6f complexes. It probably is of monophyletic origin. Cyt b_6 + subunit IV of cyanobacteria and of chloroplasts resulted from gene splitting, or formed cyt *b* by gene fusion. They are very distant to any other cyt *b*, with less than 30% identity. This evolutionary distance is also

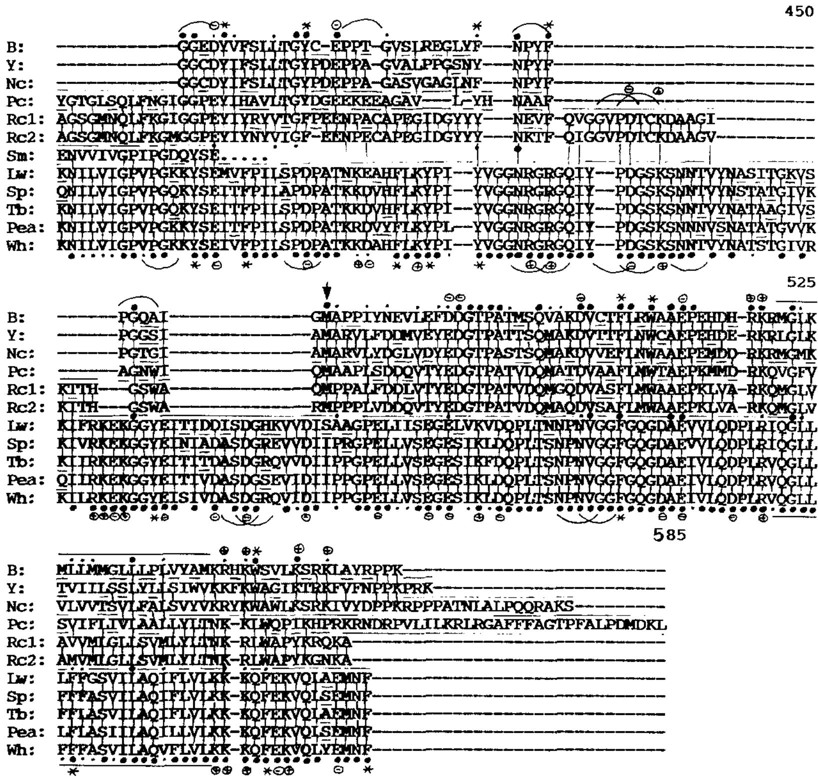

Fig. 3. Continued.

documented by cyanobacterial cyt c_6 and mitochondrial cyt c (Dayhoff, 1978). Noteworthy is the fact that the β-subunit of the coupling factor complex shows over 80% identity between chloroplasts, mitochondria, and bacteria (Walker *et al.*, 1985). This cannot be merely explained by higher functional stringency, because even then the evolutionary distance should be seen, but is not.

2. Cyt b of plant mitochondria, like cyt b_6 + subunit IV of chloroplasts, is significantly more conserved than cyt b of animal mitochondria, although cyt c, which is encoded in the nucleus, is similarly conserved within plants and animals (Dayhoff, 1976). Thus the organellar genomes of plants change more slowly than the nucleus, which is the opposite from animals (Brown *et al.*, 1982).

3. Cyt b_6 + subunit IV, as well as cyt f, are more closely related in cyanobacteria and chloroplasts than cyt b and cyt c_1 from bacteria and mitochondria.

4. Cyt c_1 and Cyt f arose by convergence of different ancestral proteins.

Table III. Percent Identity in Pairs of Cyt c_1/f

		bc_1					b_6f					
		B	Y	Nc	Pc	Rc	Sm	Lw	Sp	Tb	Pea	Wh
bc_1	B	100										
	Y	59	100									
	Nc	61	66	100								
	Pc	26	35	32	100							
	Rc	26	30	28	53	100						
b_6f	Sm	9	8	8	8	7	100					
	Lw	8	10	9	11	13	63	100				
	Sp	8	11	9	10	12	64	84	100			
	Tb	9	12	9	11	13	65	83	94	100		
	Pea	8	11	9	11	13	62	81	91	91	100	
	Wh	8	11	9	10	12	64	85	95	93	90	100

5. The Rieske FeS proteins have a common origin of the C-terminal part, only.

6. All three proteins carry insertions in chloroplasts and bacteria that might compensate for the lack of the additional sununits present in mitochondria.

7. The peptides carrying the redox centers can be defined for all three proteins. In addition, conserved aromatic residues indicate the sites for electron transfer, in particular with quinone/quinol. These putative reaction sites in cyt *b*, as well as the redox centers of cyt c_1 or *f*, and of the Rieske FeS protein, are formed by β-turns of the peptide backbone.

8. Differences in the conserved structures of bc_1 and b_6f-polypeptides are observed and might explain the different inhibitor sensitivity.

9. We believe that this comparative approach could finally lead to a realistic, detailed model of the bc_1/b_6f complexes, when extended by computer-aided secondary structure predictions and combined with experimental test.

Acknowledgments

We thank F. Daldal (Cold Spring Harbor), B. Ludwig (Luebeck), G. von Jagow (Munich), B. L. Trumpower (Hanover), R. Malkin (Berkeley),

and H. Weiss (Dusseldorf) for providing unpublished material. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 43/C2).

References

Alam, J., and Krogmann, D. W. (1984). In *Advances in Photosynthesis Research*, Vol. 1 (Sybesma, C., ed.), Martinus Nijhoff/Dr. W. Junk Publishers, Dordecht, Netherlands, pp. 1.4.521–524.

Alt, J., and Hermann, R. G. (1984). *Curr. Genet.* **8**, 551–557.

Anderson, S., Bankeir, A. T., Barrell, B. C., de Bruijn, M. H. L., Coulson, A. R., Dronin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schrier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981a). *Nature* (*London*) **290**, 456–465.

Anderson, S., de Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F., and Young, I. G. (1981b). *J. Mol. Biol.* **156**, 683–717.

Beckmann, J. D., Ljungdahl, P. O., Lopez, J. L., and Trumpower, B. L. (1987). *J. Biol. Chem.* **262**, 129–000.

Benne, R., De Vries, B. F., Van den Burg, J., and Klaver, B. (1983). *Nucleic Acids Res.* **11**, 6925–6941.

Bibb, M. J., Van Elten, R. A., Wright, C. T., Walberg, M. W., and Clayton, D. A. (1981). *Cell* **26**, 167–180.

Boer, P. H., McIntosh, J. E., Gray, M. W., and Bonen, L. (1985). *Nucleic Acids Res.* **13**, 2281–2292.

Brandle, C. J., and Deber, C. M. (1986). *Proc. Natl. Acad. Sci. U.S.A.* **83**, 917–921.

Brown, W. M., Prager, E. M., Wang, A., and Wilson, A. C. (1982). *J. Mol. Evol.* **18**, 225–239.

Burke, J. M., Breitenberger, C., Heckman, J. E., Dujon, B., and Rajbhandary, U. L. (1984). *J. Biol. Chem.* **259**, 504–511.

Chou, P. Y., and Fasman, G. D. (1979). *Biophys. J.* **26**, 367–383.

Colson, S. (1987). In *Cytochrome Systems: Molecular Biology and Bioenergetics* (Papa, S., ed.), Proc. Int. Symp. 1987, Fasano, Plenum Press, New York, in press.

Crofts, A. R. (1985). In *The Enzymes of Biological Membranes* (Martonosi, A. N., ed.), Plenum Press, New York, pp. 347–382.

Davidson, E., and Daldal, F. (1987a). *J. Mol. Biol.* **195**, 13–24.

Davidson, E., and Daldal, F. (1987b). *J. Mol. Biol.* **195**, 23–30.

Davis, D. J. (1987). *Progress in Photosynthesis Research* (Biggins, J., ed.), Martinus Nijhoff Publishers, Dordrecht, Netherlands, pp. II.7.473–476.

Dawson, A. J., Jones, V. P., and Leaver, C. J. (1984). *EMBO J.* **3**, 2107–2113.

Dayhoff, M. O. (1976). *Atlas of Protein Sequence and Structure*, Suppl. 2, Natl. Biomed. Res. Found., Washington, D.C., p. 33.

Dayhoff, M. O. (1978). *Atlas of Proetin Sequence and Structure*, Suppl. 3, Natl. Biomed. Res. Found., Washington, D.C.

de la Cruz, V. F., Neckelmann, N., and Simpson, L. (1984). *J. Biol. Chem.* **259**, 15136–15147.

Dixon, *et al.* (1986). *Nature* (*London*) **321**, 75–79.

Gabellini, N., and Sebald, W. (1986). *Eur. J. Biochem.* **154**, 569–579.

Gilbert, W., Marchionni, M., and McKnight, G. (1986). *Cell* **46**, 151–154.

Harnisch, U., Weiss, H., and Sebald, W. (1985). *Eur. J. Biochem.* **149**, 95–99.

Hartl, F. U., Schmidt, B., Wachter, E., Weiss, H., and Neupert, W. (1986). *Cell* **47**, 939–951.

Hauska, G. (1985). In *Molecular Biology of the Photosynthectic Apparatus* (Steinback, K. E., Bonitz, S., Arntzen, C. J., and Bogorad, L., eds.), Cold Spring Harbor Laboratory, pp. 79–87.

Hauska, G. (1986a). *Encycl. Plant Physiol., New Ser.* **19**, 496–507.

Hauska, G. (1986b). *Methods Enzymol.* **126**, 271–285.

Hauska, G., Hurt, E., Gabellini, N., and Lockau, W. (1983). *BIochim. Biophys. Acta* **726**, 97–133.

Heinemeyer, W., Alt, J., and Herrmann, R. G. (1984). *Curr. Genet.* **8**, 543–606.

Helmer-Citterich, M., Morelli, G., and Macino, G. (1983). *EMBO J.* **2**, 1235–1242.
Herrmann, R. G., Alt, J., Schiller, B., Widger, W. R., and Cramer, W. A. (1984). *FEBS Lett.* **176**, 239–244.
Kallas, T., Spiller, S., and Malkin, R. (1986). *Progress in Photosynthesis Research* (Biggins, J., ed.), Martinus Nijhoff Publishers, Dordrecht, Netherlands, pp. IV.12.801–803.
Koike, K., Kobayashi, M., Yaginuma, K., Taira, M., Yoshida, E., and Imai, M. (1982). *Gene* **20**, 177–185.
Kurowski, B. and Ludwig, B. (1987). *J. Biol. Chem.* **262**, 13805–13811.
Lang, B. F., Ahne, F., and Bonen, L. (1985). *J. Mol. Biol.* **184**, 353–366.
Ljungdahl, P. O., Pennoyer, J. D., and Trumpower, B. L. (1986). *Methods Enzymol.* **126**, 181–190.
Markos, A., Blahuskova, A., and Kalous, M. (1986). 4th Eur. Bioenerg. Conf. (EBEC), Prague, 1986, reports, Vol. 4, 35.
Nitschke, W., and Hauska, G. (1987a). *Progress in Photosynthesis Research* (Biggins, J., ed.), Martinus Nijhoff Publishers, Dordrecht, Netherlands, Vol. II, pp. 165–172.
Nitschke, W., and Hauska, G. (1987b). *Biochim. Biophys. Acta,* **892**, 314–319.
Nobrega, F. G., and Tzagoloff, A. (1980). *J. Biol. Chem.* **255**, 9828–9837.
Ohyama, *et al.* (1986). *Nature (London)* **322**, 572–575.
Pfefferkorn, B., and Meyer, H. E. (1986). *FEBS Lett.* **206**, 233–237.
Pfefferkorn, B., and Meyer, H. E. (1987). *Hoppe-Seylers Biochem.* **368**, 561.
Phillips, A. C., and Gray, J. C. (1984). *Mol. Gen. Genet.* **194**, 477–484.
Rich, P. (1984). *Biochim. Biophys. Acta* **768**, 53–79.
Robertson, D. E., Davidson, E., Prince, R. C., van den Berg, W. H., Marrs, B. L., and Dutton, D. L. (1986). *J. Biol. Chem.* **261**, 584–591.
Roe, B. A., Ma, D. P., Wilson, R. K., and Wong, J. F. H. (1985). *J. Biol. Chem.* **260**, 9759–9774.
Römisch, J., Tropschung, M., Sebald, W., and Weiss, H. (1987). *Eur. J. Biochem.*, in press.
Sadler, I., Suda, K., Schatz, G., Kaudewitz, F., and Haid, A. (1984). *EMBO J.* **3**, 2137–2143.
Saraste, M. (1984). *FEBS Lett.* **166**, 367–372.
Schägger, H., Link, T. A., Engel, W. D., and von Jagow, G. (1986). *Methods Enzymol.* **126**, 224–237.
Schägger, H., Borchart, U., Machleidt, W., Link, T. A., and von Jagow, G. (1987). *FEBS Lett.*, in press.
Schuster, W., and Brennicke, A. (1985). *Curr. Genet.* **9**, 157–163.
Shinozaki, *et al.* (1986). *EMBO J.* **5**, 2043–2049.
Smeekens, S., van Binsbergen, J., and Weisbeek, P. (1985). *Nucleic Acids Res.* **13**, 3179–3194.
Steppuhn, J., Hermans, J., Janson, T., Vater, J., Hauska, G., and Herrmann, R. G. (1987). *Mol. Gen. Genet.*, in press.
Telser, J., Hoffman, B. M., LoBrutto, R., Ohnishi, T., Tsai, A. L., Simpkin, D., and Palmer, G. (1987). *FEBS Lett.* **214**, 117–121.
Wakabayashi, S., Matsubara, H., Kim, C. H., Kawai, K., and King, T. E. (1980). *Biochem. Biophys. Res. Commun.* **97**, 1548–1554.
Walker, J. E., Fearnley, I. M., Gray, N. J., Gibson, B. W., Northrop, F. D., Powell, S. J., Runswick, M. J., Saraste, M., and Tybulowicz, V. L. Y. (1985). *J. Mol. Biol.* **184**, 677–701.
Waring, R. B., Davis, R. W., Lee, S., Grisi, E., McPhail-Berks, M., and Scazzocchio, C. (1981). *Cell* **27**, 4–11.
Widger, W. R., Cramer, W. A., Herrmann, R. G., and Trebst, A. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 674–678.
Willey, D. L., Auffret, A. D., and Gray, J. C. (1984a). *Cell* **36**, 555–562.
Willey, D. L., Howe, C. J., Auffret, A. D., Bowman, C. M., Dyer, T. A., and Gray, J. C. (1984b). *Mol. Gen. Genet.* **194**, 416–422.
Yu, L., Dong, J. H., and Yu, C. A. (1986). *Biochim. Biophys. Acta* **852**, 203.