The EMBO Journal vol.13 no.11 pp.2516–2525, 1994

Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen

Jose Castresana[1], Mathias Lübben[2], Matti Saraste and Desmond G.Higgins

European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

[2]Present address: Ruhr-Universität Bochum, Institut für Biophysik, Universitätstrasse 150, D-44780 Bochum, Germany

[1]Corresponding author

Communicated by W.Kühlbrandt

Cytochrome oxidase is a key enzyme in aerobic metabolism. All the recorded eubacterial (domain *Bacteria*) and archaebacterial (*Archaea*) sequences of subunits 1 and 2 of this protein complex have been used for a comprehensive evolutionary analysis. The phylogenetic trees reveal several processes of gene duplication. Some of these are ancient, having occurred in the common ancestor of *Bacteria* and *Archaea*, whereas others have occurred in specific lines of *Bacteria*. We show that eubacterial quinol oxidase was derived from cytochrome *c* oxidase in Gram-positive bacteria and that archaebacterial quinol oxidase has an independent origin. A considerable amount of evidence suggests that Proteobacteria (Purple bacteria) acquired quinol oxidase through a lateral gene transfer from Gram-positive bacteria. The prevalent hypothesis that aerobic metabolism arose several times in evolution after oxygenic photosynthesis, is not sustained by two aspects of the molecular data. First, cytochrome oxidase was present in the common ancestor of *Archaea* and *Bacteria* whereas oxygenic photosynthesis appeared in *Bacteria*. Second, an extant cytochrome oxidase in nitrogen-fixing bacteria shows that aerobic metabolism is possible in an environment with a very low level of oxygen, such as the root nodules of leguminous plants. Therefore, we propose that aerobic metabolism in organisms with cytochrome oxidase has a monophyletic and ancient origin, prior to the appearance of eubacterial oxygenic photosynthetic organisms.

Key words: aerobic metabolism/common ancestor/cytochrome oxidase/evolution/quinol oxidase

Introduction

Aerobic organisms generate energy by means of a respiratory chain that uses oxygen as the terminal acceptor of electrons. Cytochrome oxidase is the enzyme that catalyzes the reduction of oxygen to water and therefore it is an essential enzyme for aerobic metabolism. It creates a proton gradient as an intermediate step in the conversion of redox energy to ATP (Capaldi, 1990; Chan and Li, 1990; Saraste, 1990; Babcock and Wikström, 1992)

Organisms with aerobic metabolism are distributed among the three major domains in which living systems can be classified: *Bacteria*, *Archaea* and *Eucaria* (eubacteria, archaebacteria and eukaryotes) (Woese and Fox, 1977; Woese *et al.*, 1990). Many eubacterial species are facultative or obligate aerobes, although most of them are anaerobic. The majority of archaebacterial species live in extreme environments which are often highly anaerobic. There are, however, several cases of aerobic archaebacteria such as *Halobacterium*, *Sulfolobus* and *Thermoplasma* (Woese, 1987). Most of the known eukaryotes possess aerobic metabolism that is performed within mitochondria. Only a few eukaryotic species lack this organelle and are strictly anaerobic (Müller, 1988).

Several kinds of electron transfer chains are present in aerobic bacteria (Anraku and Gennis, 1987; Anraku, 1988). They differ in the number and type of electron carriers and terminal oxidases. Some terminal oxidases receive electrons from cytochrome *c* (cytochrome *c* oxidases) whereas others take them from quinols (quinol oxidases) (Figure 1A). Amino acid sequences of cytochrome *c* oxidases and a subclass of quinol oxidases are homologous (Chepuri *et al.*, 1990). Therefore, they are grouped together in the same superfamily of 'cytochrome oxidases' (Saraste, 1990; Saraste *et al.*, 1991a). In eubacteria there is another subclass of quinol oxidases that have no homology with this family, the cytochrome *bd* complex, found for example in *Escherichia*

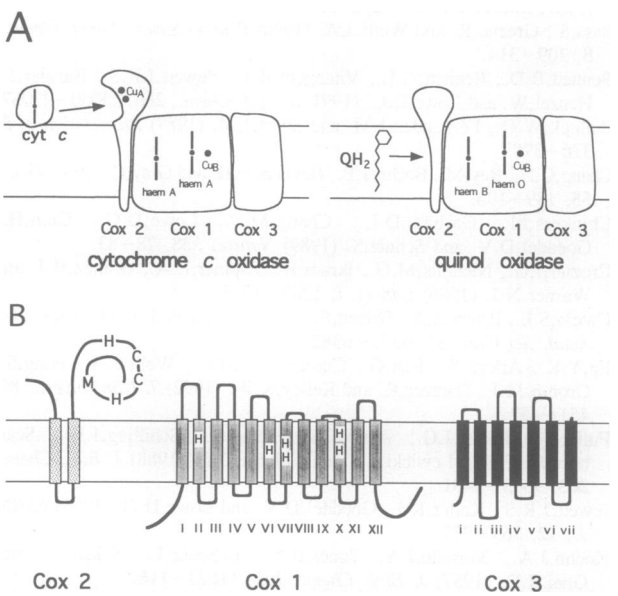

Fig. 1. Structural model of cytochrome oxidase. (**A**) Subunit composition and redox centres of a cytochrome *c* oxidase (*P.denitrificans*) and a quinol oxidase (*E.coli*). The electron donors, cytochrome *c* (cyt *c*) and ubiquinol (QH_2), respectively, are also represented. (**B**) Putative transmembrane helices of the subunits of the cytochrome *c* oxidase of *P.denitrificans*. The amino acids involved in binding redox centres are also shown (one letter code).

coli (Green *et al.*, 1988) and *Azotobacter vinelandii* (Moshiri *et al.*, 1991).

Members of the cytochrome oxidase family are multi-subunit, membrane-bound complexes. The eubacterial enzymes are the best known (Figure 1). They consist of three main common subunits, of which subunit 1 (Cox 1) is the largest. In a typical cytochrome *c* oxidase, such as the enzyme of *Paracoccus denitrificans*, Cox 1 consists of 12 putative transmembrane helices (Figure 1B). The reduction of oxygen takes place in the bimetallic cytochrome a_3-Cu_B active site bound to Cox 1 through four histidines. Cox 1 has another haem A centre, cytochrome *a*, involved in intramolecular electron transfer, which is co-ordinated to two histidines. In the quinol oxidase of *E.coli*, the bimetallic centre contains a haem O and the other redox centre is a haem B. The kind of haems in Cox 1 may differ in other species, even between enzymes of very similar sequences. The six histidines that bind the redox centres (Figure 1B) are completely conserved in all the cytochrome oxidases sequenced so far and are highly diagnostic of the family. Subunit 2 (Cox 2) is anchored to the membrane by two putative transmembrane helices. It has an aqueous exposed region which, in the case of cytochrome *c* oxidases, is supposed to be the site of interaction with cytochrome *c*. A redox centre, Cu_A, is placed in a cupredoxin domain (Adman, 1991; van der Oost *et al.*, 1992) and involved in the transfer of electrons to the bimetallic active site of Cox 1. Two cysteines, two histidines and one methionine are the ligands of this copper centre (Kelly *et al.*, 1993) (Figure 1B). Four of these amino acids are absent from quinol oxidases correlating to the lack of Cu_A centre. Subunit 3 (Cox 3) has seven transmembrane helices in the cytochrome *c* oxidase of *P.denitrificans* (Figure 1B) but this number can be lower in other oxidases. Its function remains enigmatic but a role in the assembly of the complex has been proposed (Haltia *et al.*, 1991).

A notable eubacterial oxidase complex has recently been described in *Bradyrhizobium japonicum* (Preisig *et al.*, 1993). When this bacterium is living as a nitrogen-fixing endosymbiont, it uses an alternative oxidase complex with a high affinity for oxygen. It consists of three membrane-anchored subunits, Cox 1 and two cytochromes *c*. Cox 2 and Cox 3 are not present in this complex. Cox 1 is very divergent from the corresponding subunits of other oxidases, but all the important metal-binding residues are conserved and situated in the same positions of the transmembrane helices.

Archaebacterial cytochrome oxidases have a different subunit composition. They all have Cox 1 but some of them lack Cox 2 or Cox 3 and have instead other types of subunits (Lübben *et al.*, 1992; M.Lübben, S.Arnaud, J.Castresana, A.Warne, S.P.J.Albracht and M.Saraste, in preparation).

The eukaryotic (mitochondrial) enzyme has a more complex subunit composition. For example, cytochrome *c* oxidase in mammalian mitochondria contains 13 subunits. Three of them are homologous to the eubacterial counterparts described above and constitute the catalytic core; they are coded by the mitochondrial DNA. The rest of the subunits are nuclear encoded and supposed to have regulatory or structural functions (Kadenbach *et al.*, 1991).

Aerobic respiratory chains are more complex in archaebacteria and eubacteria than in mitochondria. In particular, the electron transfer chain is often branched to different terminal oxidases. Depending on the growth conditions, one of them may predominate. As an example, in *Bacillus subtilis* four terminal oxidases have been detected spectroscopically: a cytochrome *c* oxidase, a quinol oxidase and two more oxidases whose function and amino acid sequence are still unknown (von Wachenfeldt and Hederstedt, 1992). Similar situations are found in other bacteria (Anraku and Gennis, 1987; Anraku, 1988).

In order to trace the evolutionary history of the cytochrome oxidase family, we have carried out a phylogenetic analysis of the sequences available in the databases. The corresponding trees exhibit several interesting and unexpected features. The existing eubacterial and archaebacterial sequences are still insufficient to unravel completely the evolution of this enzyme, but it is already possible to give a clear answer to some evolutionary events and to advance some speculations that can be tested when more sequences become available.

The origin of aerobic metabolism had very important consequences for the evolution of life. Organisms which could adapt to use oxygen as the oxidant in their respiratory chains had a big metabolic advantage as much more energy could be released from this reaction in comparison with the use of less effective oxidants such as nitrogen oxides (Harold, 1986). The prevalent view is that aerobic metabolism arose several times independently during evolution because aerobic organisms do not form a monophyletic clade (Fox *et al.*, 1980; Woese, 1987; Buse and Steffens, 1991). In contrast, our sequence analysis of cytochrome oxidase, the key enzyme in aerobic metabolism, gives support to a monophyletic and ancient origin of respiration.

Results

All the eubacterial and archaebacterial sequences in the EMBL database and three representative sequences from the major eukaryotic kingdoms have been used in this study (see Table I). It is not necessary to consider more eukaryotic sequences as they all are very similar and group closely together in the trees. Three separate multiple alignments and corresponding phylogenetic trees were produced for the Cox 1, Cox 2 and Cox 3 proteins.

The multiple alignment of the Cox 1 sequences contains a conserved core of ~400 residues which can be aligned across all 24 proteins. Several sequences have N- and C-terminal extensions, which reflects the different number of transmembrane helices in these subunits. We have used hydrophobicity plots to infer the number of putative transmembrane helices in the Cox 1 subunits. The schematic alignment of the homologous transmembrane helices (Figure 2) shows that the common core of Cox 1 is composed of 12 helices. Some Cox 1 have just these 12 helices, for example, cytochrome *c* oxidases of Proteobacteria (Purple bacteria), Cyanobacteria and eukaryotes. Other proteins like cytochrome *c* oxidases of *Bacillus* have a C-terminal extension of two more helices. The Cox 3 proteins which form complexes with these extended Cox 1 subunits have lost two N-terminal helices. The quinol oxidases of *Bacillus* and Proteobacteria have these two helices which normally belong to Cox 3 and another helix at the N-terminus. *B.japonicum* and *Rhizobium meliloti* FixN proteins have two extra helices at the N-terminus, neither of them showing homology to the N-terminal extra helices of quinol oxidases

Table I. Relation of the Cox 1 and Cox 2 subunits of the cytochrome oxidases used in this study

Species	Cox 1	Cox 2	Activity	Domain	Classification
B.subtilis	CtaD[a]	CtaC[a]	C	Bacteria	D. Gram-positive bacteria
B.subtilis	QoxB[b]	QoxA[b]	Q	Bacteria	D. Gram-positive bacteria
B.firmus	CtaD[c]	CtaC[c]	C	Bacteria	D. Gram-positive bacteria
Bacillus PS3	CaaB[d]	CaaA[d]	C	Bacteria	D. Gram-positive bacteria
B.japonicum	CoxA[e]		C	Bacteria	D. Proteobacteria (Sc. Alpha)
B.japonicum	CoxN[f]	CoxM[f]	C	Bacteria	D. Proteobacteria (Sc. Alpha)
B.japonicum	FixN[g]		C	Bacteria	D. Proteobacteria (Sc. Alpha)
R.leguminosarum	CoxA		C	Bacteria	D. Proteobacteria (Sc. Alpha)
R.meliloti	FixN		C	Bacteria	D. Proteobacteria (Sc. Alpha)
P.denitrificans	CtaDI[h]	CtaC[i]	C	Bacteria	D. Proteobacteria (Sc. Alpha)
P.denitrificans	CtaDII[j]		C	Bacteria	D. Proteobacteria (Sc. Alpha)
R.sphaeroides	CtaD[k]	Cox2[l]	C	Bacteria	D. Proteobacteria (Sc. Alpha)
A.aceti	CyaA[m]	CyaB	Q	Bacteria	D. Proteobacteria (Sc. Alpha)
E.coli	CyoB[n]	CyoA[n]	Q	Bacteria	D. Proteobacteria (Sc. Gamma)
S.vulcanus	CtaD[o]	CtaC[o]	C	Bacteria	D. Cyanobacteria
Synechocystis PCC 6803	CtaD[p]	CtaC[p]	C	Bacteria	D. Cyanobacteria
T.thermophilus	CaaB[q]	CaaA[r]	C	Bacteria	D. Deinococcaceae and Thermus
T.thermophilus	CbaA	CbaB	C?	Bacteria	D. Deinococcaceae and Thermus
H.halobium	Cox1[s]		C	Archaea	F. Halobacteriaceae
S.acidocaldarius	SoxM[t]		Q	Archaea	O. Sulfolobales
S.acidocaldarius	SoxB[u]	SoxA[u]	Q	Archaea	O. Sulfolobales
Z.mays	Cox1	Cox2	C	Eucaria	K. Plantae
A.nidulans	Cox1	Cox2	C	Eucaria	K. Fungi
H.sapiens	Cox1	Cox2	C	Eucaria	K. Animalia

The activity, when known, is cytochrome *c* oxidase (C) or quinol oxidase (Q). The *T.thermophilus* CbaAB is most probably a cytochrome *c* oxidase as it has a Cu_A centre and oxidizes cytochrome *c* (Zimmermann *et al.*, 1988). The classification given for *Bacteria* and *Archaea* is based on Woese (1987, 1992). We give the level of classification which is most useful for the discussion in the text. The abbreviations used in the classification stand for Division (D.), Subclass (Sc.), Family (F.), Order (O.) and Kingdom (K.).
References: [a]Saraste *et al.* (1991b); [b]Santana *et al.* (1992); [c]Quirk *et al.* (1993); [d]Ishizuka *et al.* (1990); [e]Bott *et al.* (1990); [f]Bott *et al.* (1992); [g]Preisig *et al.* (1993); [h]Raitio *et al.* (1987); [i]Steinrücke *et al.* (1987); [j]Raitio *et al.* (1990); [k]Shapleigh *et al.* (1992); [l]Cao *et al.* (1991); [m]Fukaya *et al.* (1993); [n]Chepuri *et al.* (1990); [o]Sone *et al.* (1993); [p]Alge and Peschek (1993); [q]Mather *et al.* (1993); [r]Mather *et al.* (1991); [s]Denda *et al.* (1991); [t]M.Lübben *et al.*, in preparation; [u]Lübben *et al.* (1992). Accession numbers for the DNA sequences are as follows: *B.subtilis* CtaD and CtaC, X54140; *B.subtilis* QoxB and QoxA, M86548; *B.firmus*, M94110; *Bacillus* PS3, D00728; *B.japonicum*, CoxA, X54800; *B.japonicum* CoxN and CoxM, X68547; *B.japonicum* FixN, L07487; *R.leguminosarum* CoxA, X74341; *R.meliloti* FixN, Z21854; *P.denitrificans* CtaDI, X05829; *P.denitrificans* CtaC, X05828; *P.denitrificans* CtaDII, Y07533; *R.sphaeroides* CtaD, X62645; *R.sphaeroides* Cox2, M57680; *E.coli*, J05492; *A.aceti*, D13185; *Synechococcus vulcanus*, S67470; *Synechocystis*, X53746; *T.thermophilus* CaaB, M84341; *T.thermophilus* CaaA, M59180; *T.thermophilus* CbaA and CbaB, L09121; *H.halobium*, D10611; *S.acidocaldarius* SoxM, X73567; *S.acidocaldarius* SoxB, X62643; *Zea mays*, X02660; *Aspergillus nidulans*, X00790; *Homo sapiens*, V00662.

(in Figure 2 they are aligned only for simplicity). In *Thermus thermophilus* CaaB and *Sulfolobus acidocaldarius* SoxM, Cox 1 and Cox 3 are merged.

Helices II–XI are the most conserved and therefore they have been used for the subsequent phylogenetic analysis (alignment A in Figure 2). A small number of insertions and deletions in some of the helices are required for optimal alignment. These are not normally expected in transmembrane helices. In most cases, in the present alignments, they appear to be real as they are flanked by conserved residues. When the FixN proteins, which are the most divergent, are included, only helices VI–XI can be easily aligned and used for the analysis (alignment B in Figure 2). The six histidines involved in binding the redox centres (Figure 1B) can be unambiguously aligned in all sequences, including those of FixN.

Figure 3 shows the tree obtained with the neighbour-joining algorithm and the Kimura's distance correction from the alignment A (Figure 2) of all Cox 1 except the most divergent ones, *B.japonicum* and *R.meliloti* FixN. It was rooted in the middle of the tree assuming that the rate of evolution in all lineages is similar. Two groups (both including eubacterial and archaebacterial sequences) diverge from this root. Bootstrap confidence levels (expressed as percentages) are given for all groupings in the tree. High values (e.g. 95% or more) indicate that the particular grouping is well supported by the data. The bootstrap values are high for many of the groupings in the tree, indicating that the overall branching order is stable. Some parts of the tree cannot be resolved clearly (indicated by low bootstrap support); for example, the branching order of the eubacterial divisions (Cyanobacteria, Deinococcaceae and Thermus, Gram-positive bacteria and Proteobacteria) is ambiguous, but this does not affect our further conclusions.

When the sequences of the two divergent FixN oxidases are included, the alignment becomes less certain and the bootstrap values are lower. Therefore, an independent tree (Figure 4) was constructed with the FixN oxidases and two representatives of each main group of the first tree. Only helices VI–XI, which are the best conserved, were used for this alignment (alignment B in Figure 2). In the new tree the root is placed in the middle of the tree (branch leading to the FixN proteins), assuming, again, a constant rate of evolution of bacterial genomes. This further justifies the position of the root in the previous tree. It should be noted that the exact position of the root does not affect the main conclusions of this paper.

There are fewer bacterial sequences of Cox 2 than those

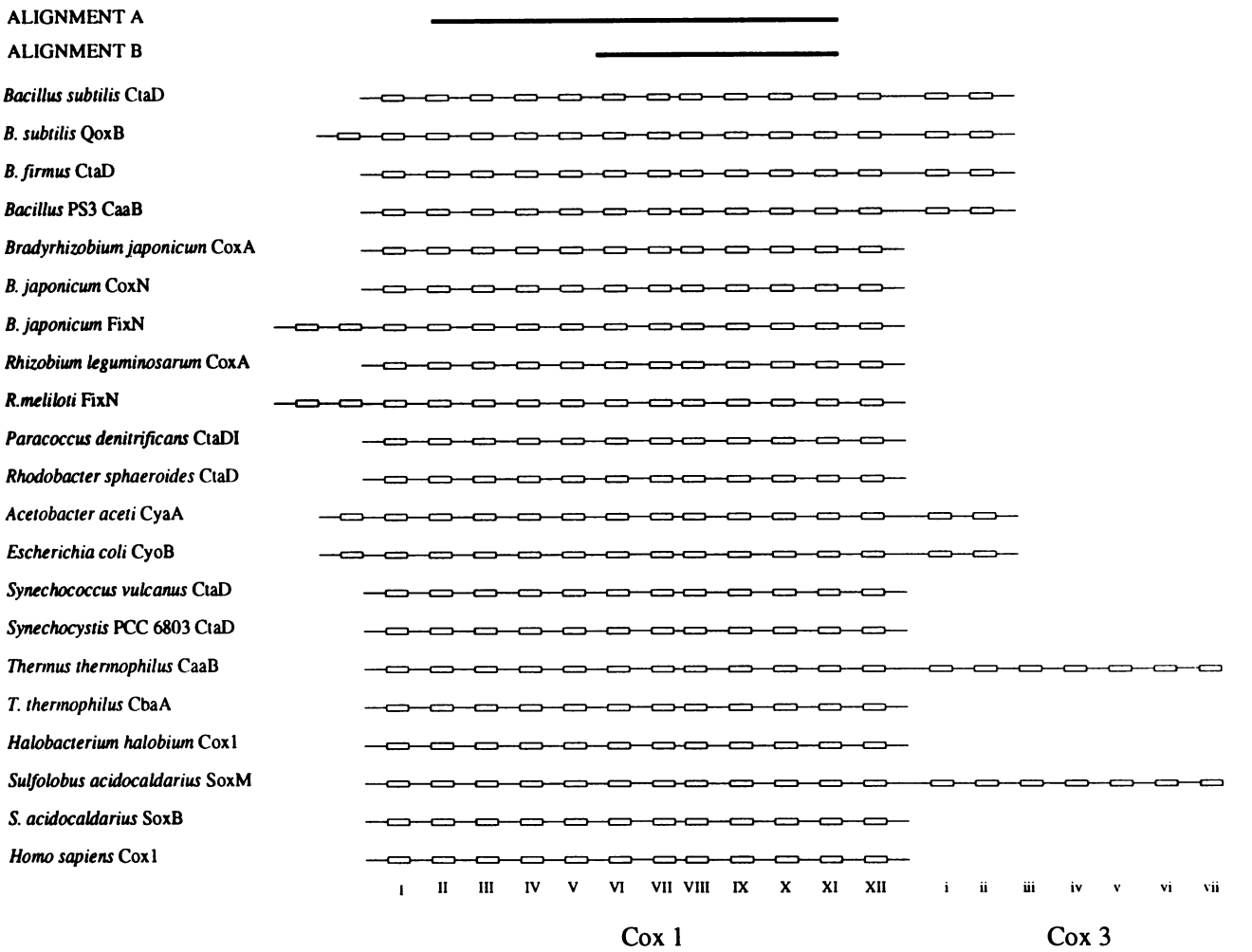

Fig. 2. Schematic alignment of the putative transmembrane helices of the most representative Cox 1 subunits used in this work. The two heavy lines in the upper part span the sections of alignment used for the phylogenetic trees in Figures 3 (alignment A) and 4 (alignment B). Roman numbering in the lower part refers to the helices for canonical Cox 1 (capitals) and Cox 3 (lower case) subunits.

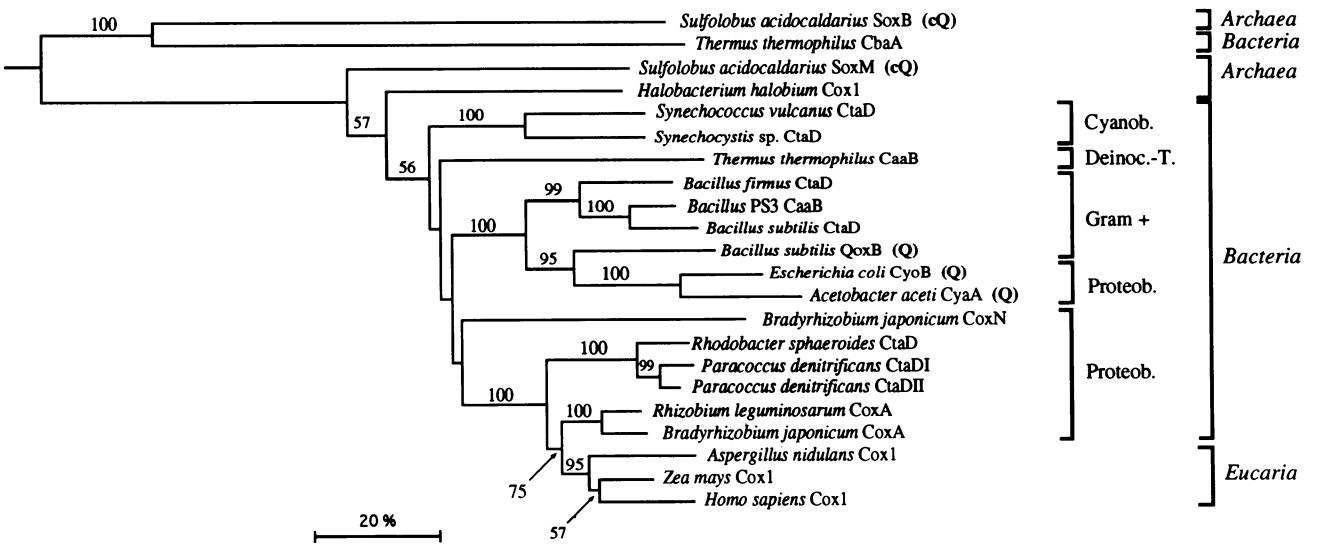

Fig. 3. Phylogenetic tree of the Cox 1 subunit of cytochrome oxidase inferred by the neighbour-joining method with Kimura's distance correction. The distantly related oxidases *B.japonicum* and *R.meliloti* FixN are not included. Eubacterial quinol oxidases are marked with (Q). Archaebacterial quinol oxidases are marked with (cQ). The abbreviations for the eubacterial divisions are: Cyanob., Cyanobacteria; Deinoc.-T., Deinococcaceae and Thermus; Gram +, Gram-positive bacteria; Proteob., Proteobacteria. Figures represent bootstrap confidence levels from 2000 bootstrap samples for the groupings; only values >50% support are shown. The scale bar represents a distance of 20%.

of Cox 1 in the databases. For some oxidase complexes only Cox 1 has been sequenced and for others Cox 2 may not be a component of the complex (Table I). Furthermore, Cox 2 sequences are shorter and less conserved. Only the second transmembrane helix (the first one is lacking in *S.acidocaldarius* SoxA and *T.thermophilus* CbaB), the cupredoxin-like domain of the membrane-exposed region (van der Oost *et al.*, 1992) and a short connection between both could be consistently aligned across all sequences. Consequently, with the low number of positions in the multiple alignment, the Cox 2 tree (Figure 5) has very low bootstrap values. The topology is, however, identical to the Cox 1 tree in all important respects. The differences correspond to groupings with poor bootstrap support in the Cox 1 tree, e.g. the positions of *T.thermophilus* CaaA and *B.japonicum* CoxM.

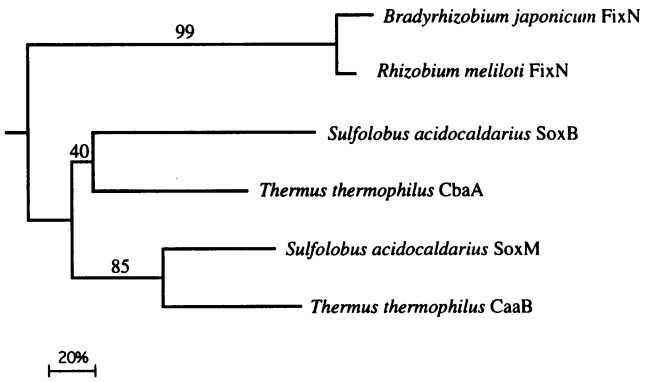

Fig. 4. Phylogenetic tree of Cox 1 subunits that includes the two FixN oxidases and four selected oxidases from the tree of Figure 3. The tree is based on the shorter sequence alignment (Figure 2). The neighbour-joining method with Kimura's distance correction was used. Values represent percentage bootstrap support from 2000 bootstrap samples. The scale bar represents a distance of 20%.

There are only 11 eubacterial and one archaebacterial sequences available for Cox 3. The phylogenetic tree of these sequences (data not shown) agrees with those of the Cox 1 and Cox 2 trees and does not add any new information.

The same tree topology was obtained in all cases when the Dayhoff model for multiple substitutions correction, together with the neighbour-joining algorithm, was applied to the above data sets. Only some of the groupings with lowest bootstrap values (e.g. the branching order of eubacterial phyla) were slightly altered. The maximum-parsimony analysis gave the same basic tree topology for every data set except for the Cox 2 tree, in which eubacterial quinol oxidases are not grouped with *Bacillus* oxidases. However, this can be explained because the number of positions used for this tree is very low, and the rate of evolution in different branches is variable (being especially fast amongst the quinol oxidases), which greatly affects the maximum-parsimony method (Saitou and Imanishi, 1989).

Discussion

The cytochrome oxidase trees do not follow the phylogeny of the organisms which is based on comparison of ribosomal RNA (Woese, 1987; Cedergren *et al.*, 1988). This implies that several gene duplication or lateral transfer events, or both, have happened during the evolution of the protein. In the following, we will try to give explanations of some features seen in the oxidase trees. Finally, we will show the implications that the analysis of the cytochrome oxidase trees have for the understanding of the evolution of respiration.

Two early gene duplication processes

The first gene duplication, as deduced from Figure 4, led to separate evolution of the FixN oxidases and the rest of the enzymes. The FixN-containing complex is an oxidase that works in the microaerobic environment of the symbiotic

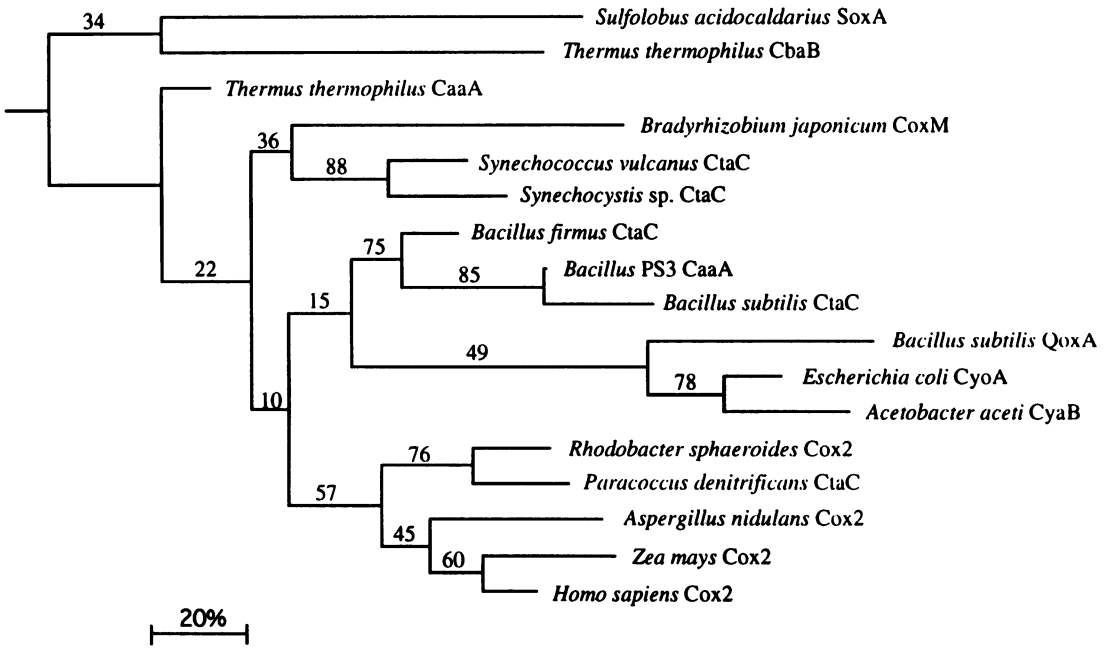

Fig. 5. Phylogenetic tree of the Cox 2 subunit of cytochrome oxidase inferred by the neighbour-joining method with Kimura's distance correction. Numbers represent percentage bootstrap support values for the groupings. The scale bar represents a distance of 20%.

bacteroids, i.e. in the root nodules (O'Brian and Maier, 1989; Bergersen and Turner, 1990; Preisig *et al.*, 1993). Biochemical analysis has recently revealed the presence of a FixN type oxidase in a non-endosymbiotic Proteobacteria (*Rhodobacter sphaeroides*) when grown under microaerobic conditions (Garcia-Horsman *et al.*, 1994). The ancestral oxidase presumably also functioned in a biosphere with a very low pressure of oxygen, which gives additional support to the position of the root of the Cox 1 tree along the FixN branch. The first gene duplication pre-dates the split between *Archaea* and *Bacteria*, and thus it can be predicted that a similar gene (i.e., more similar to FixN than to any other Cox 1 sequence) could be found in *Archaea*.

The second early gene duplication, also pre-dating the *Archaea*–*Bacteria* split, resulted in separate evolution of the two main kinds of oxidases that can be distinguished in the Cox 1 tree (Figure 3). It separated the group containing the archaebacterial *S.acidocaldarius* SoxB and the eubacterial *T.thermophilus* CbaA from the rest of archaebacterial and eubacterial oxidases. The same grouping is corroborated by the Cox 2 tree (Figure 5).

An important biochemical differentiation could have followed the early gene duplications, for example vectorial proton pumping activity may not be a property of all branches of oxidases (Lübben *et al.*, 1994).

Evolution of quinol oxidase

The presence of quinol and cytochrome *c* oxidases in eubacteria has led to many studies on the differences and similarities in the catalytic mechanisms that both enzymes use for oxygen reduction and proton pumping (Babcock and Wikström, 1992; Hosler *et al.*, 1993; Musser *et al.*, 1993a). An important question underlying these discussions is the evolutionary origin of both enzymes (Musser *et al.*, 1993a, b; Haltia, 1993). Was the first eubacterial oxidase a quinol oxidase or a cytochrome *c* oxidase?

Sequences of eubacterial quinol oxidases are known from *B.subtilis, Acetobacter aceti* and *E.coli*. Cytochrome *c* oxidases of *Bacillus* (Gram-positive bacteria) are more closely related to these quinol oxidases than to cytochrome *c* oxidases of other eubacteria (see trees of Figure 3 for Cox 1 and Figure 5 for Cox 2). It means that a gene duplication event happened during the evolution of Gram-positive bacteria and gave rise to quinol oxidases (Figure 6). The original, ancestral gene that was duplicated is predicted to be a cytochrome *c* oxidase, not a quinol oxidase. This is because all the other eubacterial oxidases, branching before the quinol–cytochrome *c* oxidase duplication, are cytochrome *c* oxidases. Therefore, the quinol oxidase activity was derived from a more primitive cytochrome *c* oxidase activity. The other possibility, that a quinol oxidase gave rise to a cytochrome *c* oxidase in the *Bacillus* line would imply that the same switch happened independently in the other eubacterial lines (Cyanobacteria and Proteobacteria), which is highly improbable.

The hypothesis that a quinol oxidase evolved from a cytochrome *c* oxidase in eubacterial evolution is congruent with the difference in redox centres of both enzymes. A copper atom centre is bound to an extramembrane domain in Cox 2 of cytochrome *c* oxidase by five residues, two cysteines, two histidines and one methionine (van der Oost *et al.*, 1992; Kelly *et al.*, 1993), leading to the centre called Cu_A (Figure 1), which might be directly involved in

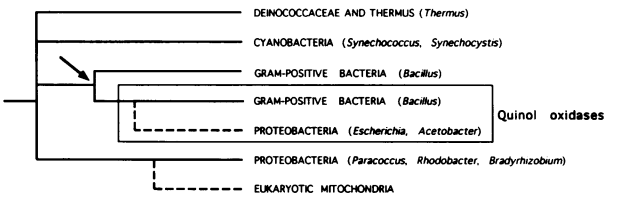

Fig. 6. Scheme of the evolution of eubacterial cytochrome *c* oxidase and quinol oxidase. The branching order of the eubacterial divisions is represented as a star-like phylogeny. The arrow points to the gene duplication in Gram-positive bacteria that gave rise to quinol oxidase. The proposed lateral gene transfer to Proteobacteria is represented with a dashed line. Similarly, a dashed line represents the endosymbiotic process that gave rise to eukaryotic mitochondria. Branch lengths are not drawn to scale.

accepting electrons from cytochrome *c* (see Lappalainen *et al.*, 1993). Four of these residues are absent in quinol oxidases, resulting in the lack of this metal centre. However, van der Oost *et al.* (1992) were able to restore the 'lost' metal-binding site of the quinol oxidase by simultaneously changing six amino acids, including four potential Cu_A ligands. Therefore, evolution of quinol oxidase required the loss of the Cu_A centre of the cytochrome *c* oxidase.

SoxB and SoxM in *S.acidocaldarius* belong to the group of quinol oxidase complexes, probably using the endogenous sulfur containing caldariellaquinol as electron donor (Anemüller and Shäfer, 1989; Lübben *et al.*, 1992). Their position in the tree indicates that the evolutionary pathway followed by these quinol oxidases is different from the one followed by the *B.subtilis, A.aceti* and *E.coli* quinol oxidases. Therefore, quinol oxidases in *Bacteria* and *Archaea* have different and independent origins. Substantial differences between the quinol oxidases of *Archaea* and *Bacteria*, at least in the quinol binding site (i.e. amino acids participating in quinol binding or even the subunit location), can be predicted.

Lateral gene transfer of quinol oxidase from Gram-positive bacteria to Proteobacteria

Ribosomal RNA phylogenetic analysis of eubacteria (Woese, 1987) puts both *E.coli* and *A.aceti* in the division Proteobacteria, whereas *Bacillus* is placed in the different division Gram-positive bacteria (Table I). However, in the cytochrome oxidase trees (Figures 3 and 5) the quinol oxidases of *E.coli* and *A.aceti* are more closely related to the quinol and cytochrome *c* oxidases of *Bacillus* than to the cytochrome *c* oxidases of the other Proteobacteria, such as *Paracoccus* and *Rhodobacter*. The most plausible explanation for this is the occurrence of lateral gene transfer (see Smith *et al.*, 1992 for other cases) of the quinol oxidase gene from Gram-positive bacteria to Proteobacteria (Figure 6). The close similarity of the three quinol oxidases and the cytochrome *c* oxidases of *Bacillus* exclude the possibility that the quinol oxidase genes were present before the split of Gram-positive bacteria and Proteobacteria (note the long internal branch leading to the quinol oxidases and cytochrome c oxidases of *Bacillus* in Figure 3, together with maximum bootstrap support for this grouping). Two additional facts support lateral gene transfer and eliminate any possibility of convergent evolution.

First, the number of putative transmembrane helices is the same in the quinol oxidases of Gram-positive bacteria

(*Bacillus*) and Proteobacteria. One extra helix is present in the N-terminus and two in the C-terminus, and these are conserved in the three quinol oxidase sequences (Figure 2). The last two helices may derive from the Cox 3 subunit through a fusion of the genes coding for Cox 1 and Cox 3, a further split between helices ii and iii of Cox 3 and creation of a translational initiation site for the new Cox 3 gene. This succession of events is unlikely to have happened more than once in evolution, supporting the notion of lateral transfer. (Note, however, that the complete fusion of Cox 1 and Cox 3, which is a simpler process, also happened independently in *S.acidocaldarius* SoxM and *T.thermophilus* CaaB).

The second argument supporting the gene transfer refers to the arrangement of the cytochrome oxidase genes in operons in Gram-positive bacteria and Proteobacteria. Figure 7 shows selected operons in both divisions as well as in a cyanobacterium for comparison. The main characteristic of the genes encoding cytochrome *c* oxidase in Proteobacteria is that the gene coding for Cox 1 is in a separate operon from those coding for the other subunits. In Gram-positive bacteria a single operon codes for cytochrome oxidase and contains, in this order, the genes coding for Cox 2, Cox 1, Cox 3 and a fourth subunit, with very small intergenic sequences separating them. Exactly the same arrangement of genes is observed in the quinol oxidases of Proteobacteria (*E.coli* and *A.aceti*) favouring the hypothesis that the entire operon for quinol oxidase was transferred from Gram-positive bacteria to Proteobacteria.

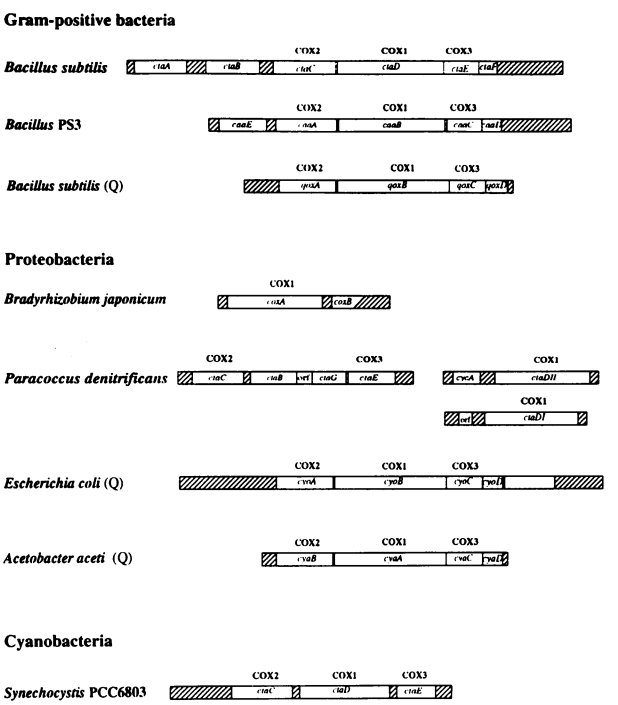

Fig. 7. Operon structure of genes coding for cytochrome oxidase subunits in Gram-positive bacteria, Proteobacteria and Cyanobacteria. Dashed squares represent non-coding regions. Adapted and extended from Saraste *et al.* (1991a). The three quinol oxidases are indicated by (Q).

Endosymbiosis leading to mitochondria

The branching of the eukaryotic cytochrome oxidases from the Alpha subclass of Proteobacteria (Figure 3) is another example of non-correspondence of the protein tree with the true phylogeny of the organisms. This is, however, a well documented situation as there are overwhelming data indicating the endosymbiotic origin of mitochondria from an ancestor of the Alpha subclass of Proteobacteria (Margulis, 1981; Yang *et al.*, 1985; Cedergren *et al.*, 1988). In our trees, members of the family Rhizobiaceae are the closest relatives of mitochondria.

Implications for the evolution of respiration

Many authors have proposed that aerobic metabolism was possible only after the oxygen released by cyanobacterial photosynthesis reached a level that made it useful as a final electron acceptor in the electron transfer chains (Dickerson *et al.*, 1976; Broda and Peschek, 1979; Moore and Pettigrew, 1990; Alge and Peschek, 1993). Another generally accepted hypothesis is the polyphyletic origin of aerobic metabolism, i.e. it arose independently in several evolutionary lines (Broda and Peschek, 1979; Fox *et al.*, 1980; Margulis, 1981; Woese, 1987; Buse and Steffens, 1991). This is due to the fact that extant aerobic organisms do not form a monophyletic clade. As suggested by Harold (1986), the sequences of cytochrome oxidases allow us to check these hypotheses.

It is generally believed that the first increase in the atmospheric level of oxygen happened ~2400–2800 million years ago (Knoll, 1992). Evidence for this date is found in the sedimentary record, in which a higher proportion of oxidized iron (red beds) started to appear during this period. It is also accepted that this increase in atmospheric oxygen occurred as a result of oxygenic photosynthesis, and many authors (for example, Margulis *et al.*, 1976) propose that the first oxygenic photosynthetic organisms were ancestors of Cyanobacteria. What is important for the following arguments is that oxygenic photosynthesis is only present in eubacteria (Cyanobacteria and products of their endosymbiosis, plastids), but not in archaebacteria. For this reason, it is generally accepted that oxygenic photosynthesis, i.e. photosystem II, has developed in the eubacterial line, *after* the *Archaea–Bacteria* split (see Figure 8).

In contrast to the unique affiliation of oxygenic photosynthesis to *Bacteria*, cytochrome oxidase sequences are found in *Archaea* and *Bacteria*. It means that an ancestral oxidase (uroxidase) evolved before the *Archaea–Bacteria* split, i.e. it was present in the last common ancestor of both domains of life (more accurately, at least three different types of oxidases existed in the common ancestor due to the proposed early duplication processes; see above). Therefore, the uroxidase was present before the appearance of the first photosynthetic eubacteria and thus before the increase of oxygen in the atmosphere due to oxygenic photosynthesis (Figure 8).

It might be hard to accept that this uroxidase had the same function that it has today, i.e. the reduction of oxygen to produce energy. However, small quantities of oxygen could have been produced as a result of water photolysis, making possible the existence of oxygen oases in localized regions of the ocean surface (Kasting, 1993). The organisms having the uroxidase could have lived in these restricted environments and could have started to develop a rudimentary aerobic metabolism. The *B.japonicum* and *R.meliloti* FixN-containing complexes, which work in the extremely low oxygen concentration environment of root

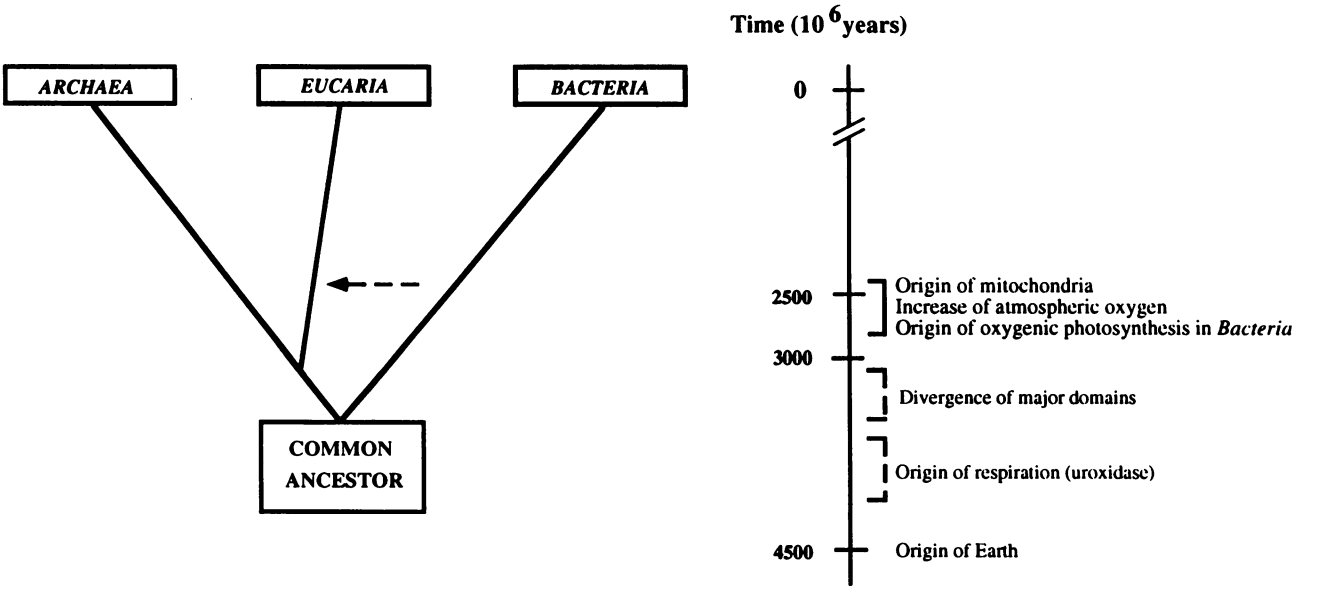

Fig. 8. Scenario of the evolution of aerobic metabolism. Dates for the origin of oxygenic photosynthesis, atmospheric oxygen and mitochondria ($2400-2800\times10^6$ years) are taken from Knoll (1992). The dashed arrow represents the endosymbiotic process that gave rise to eukaryotic mitochondria. The phylogeny of *Archaea*, *Eucaria* and *Bacteria* is based on Woese *et al.* (1990) and references therein. Only the split between *Archaea* and *Bacteria* is important for this discussion; the position of *Eucaria* in the tree does not affect it because its cytochrome oxidase was acquired at a later stage. The dates for the divergence of the major domains of life and the origin of respiration are not known; only the order of these events can be predicted from this work.

nodules (Preisig *et al.*, 1993) and are situated directly off the root of the cytochrome oxidase phylogenetic tree, support the possibility of the existence of an uroxidase working in a primitive biosphere poor in oxygen.

The idea that respiration could have pre-dated oxygenic photosynthesis has been pointed out by some authors based on other molecular data (Schwartz and Dayhoff, 1978; 1979; but see also Demoulin, 1979; Buse *et al.*, 1989) and the biogeochemical record (Towe, 1978). The present molecular analysis on cytochrome oxidase gives more direct support for this hypothesis.

There are still other possible functions for the product of the uroxidase gene in the common ancestor. For example, it could have been part of an electron transfer chain that, instead of using oxygen as the oxidant, could have used another molecule with a similar electronic configuration (for example, nitric or nitrous oxide). Then, when the atmosphere turned oxidizing as a result of oxygenic photosynthesis, the uroxidase acquired the capacity to use oxygen as the oxidant, independently in *Archaea* and *Bacteria*. Another possibility is that the primitive function of this enzyme was just to get rid of the minor amounts of oxygen that would be poisonous for cells but not to produce energy. Later on, when the oxygen level increased, selective pressure made this enzyme evolve into a component of the respiratory chain, accepting electrons of other components to reduce oxygen and create a proton gradient in the membrane. Both hypotheses have the serious inconvenience of involving multiple independent appearances of a proton pumping oxidase in *Archaea* and in *Bacteria*, giving rise to a convergence in function. The two gene duplications that appear to have occurred in the common ancestor (see above) make even more unlikely the possibility of convergent evolution from an oxidase which is not involved in aerobic metabolism to a respiratory cytochrome oxidase. This should have happened independently in at least five different oxidases, three in eubacteria and two in archaebacteria (FixN has not been detected in the latter).

The knowledge of the catalytic mechanism of oxygen reduction in both the eubacterial and archaebacterial enzymes would be of utmost importance to finally assert this hypothesis. If the mechanisms are equivalent, as the high sequence similarity of the active sites predicts, the existence of an aerobic respiratory chain in the times of the common ancestor would be strongly supported.

The lack of a close phylogenetic relationship between extant aerobic organisms, which has hidden the notion of monophyly of aerobic metabolism, remains to be explained. The same lack of phylogenetic relation occurs among photosynthetic bacteria (Woese *et al.*, 1985). However, the homology of the integrating parts of photosystem I is well recognized, and it is presupposed that such a complex mechanism of photon harvesting and energy production could have evolved only once, i.e. it was present in the common ancestor of all eubacteria. Only some of the eubacterial lines maintained the photosynthetic apparatus whereas others lost it and adapted to other methods of producing energy (Woese *et al.*, 1985; Stouthamer, 1992). We believe that a similar succession of events happened during the evolution of aerobic metabolism. An aerobic electron transfer chain was present in the common ancestor of *Archaea* and *Bacteria*. This metabolic capacity was subsequently lost in many evolutionary lines of *Archaea* and *Bacteria* so that it presently only exists in some of them. Earlier eukaryotes did not have aerobic metabolism (Müller, 1988), either because they lost it in a primitive stage or because they derived from an anaerobic bacterium. When the atmospheric levels of oxygen increased, some eukaryotes acquired aerobic metabolism by means of a eubacterial endosymbiont which gave rise to mitochondria. We conclude that the origin of aerobic metabolism in organisms with cytochrome oxidase is monophyletic.

Materials and methods

Sequences

Sequences were extracted from the EMBL nucleotide sequence database, release 37 (Rice *et al.*, 1993), and translated to proteins. Sequences were identified by keyword search and sequence similarity search (Gibson *et al.*, 1993). In all, 21 eubacterial and archaebacterial complete amino acid sequences for Cox 1, 14 for Cox 2 and 12 for Cox 3, together with three representative eukaryotic sequences for each subunit were analysed (Table I).

Sequence alignments

Sequence alignments were carried out with the PILEUP program of the GCG package (Devereux *et al.*, 1984) for the most conserved sequences and the profile alignment option of the CLUSTALV program (Higgins *et al.*, 1992) to introduce the divergent sequences, followed by manual adjustment where necessary. For the subsequent phylogenetic analysis only the most conserved sections, where the alignment is unambiguous, were used in each case. In the case of the most conserved Cox 1 subunits (see Results), the region encompassing the putative transmembrane helices II–XI was used (alignment A in Figure 2). A total of 322 sites from the alignment were analysed, after removal of all positions with a gap in any sequence. When the two divergent FixN sequences were introduced, only the region of helices VI–XI was used (alignment B). This left 193 positions in the alignment for phylogenetic analysis. For Cox 2, the second putative transmembrane helix, the extramembraneous domain and a short region connecting them were used for the alignment, giving a total of 89 sites for analysis. For Cox3, 144 sites from an automatic alignment were analysed. All the alignments are available from J.Castresana upon request (please send an electronic mail message to Castresana@EMBL-Heidelberg.DE).

Data analysis

Hydrophobicity profiles were calculated with the PEPPLOT program of the GCG package (Devereux *et al.*, 1984) that uses simultaneously the methods of Engelman *et al.* (1986) and Kyte and Doolittle (1982). These profiles were used together with the sequence alignments to produce Figure 2.

Phylogenetic trees were calculated using the neighbour-joining method of Saitou and Nei (1987). Distance matrices were calculated from the three multiple alignments by calculating percentage amino acid difference values between all pairs of sequences. As mentioned, all alignment positions that contained a gap in any sequence were removed from the analysis. Distances were corrected for multiple substitutions using equation 4.8 from Kimura (1983). Confidence intervals were calculated using a bootstrap procedure (Felsenstein, 1985) with 2000 replications. All of the above calculations were carried out using the CLUSTALV program. As some of the sequences are very divergent, the Dayhoff model of evolutionary change (Dayhoff *et al.*, 1978) was also applied for distance correction. For that purpose, the PROTDIST and NEIGHBOR programs of the PHYLIP package (Felsenstein, 1989) were used. In order to compare results, maximum-parsimony analysis using the PROTPARS program of the PHYLIP package was also carried out with the same data sets. The trees were drawn using the DRAWTREE, RETREE and DRAWGRAM programs of the PHYLIP package.

Acknowledgements

We are grateful to Toby Gibson and Julie Thompson for their help with the sequence similarity search and alignments and their useful suggestions and to Alfonso Valencia for reading the manuscript. J.Castresana thanks the Basque Government for a postdoctoral fellowship. M.Lübben has received support from an EMBO long term fellowship and the CEE grant SC1-CT91-0698 (CAESAR project).

References

Adman,E.T. (1991) *Curr. Opin. Struct. Biol.*, **1**, 895–904.
Alge,D. and Peschek,G.A. (1993) *Biochem. Biophys. Res. Commun.*, **191**, 9–17.
Anemüller,S. and Shäfer,G. (1989) *FEBS Lett.*, **244**, 451–455.
Anraku,Y. (1988) *Annu. Rev. Biochem.*, **57**, 101–132.
Anraku,Y. and Gennis,R.B. (1987) *Trends Biochem. Sci.*, **12**, 262–266.
Babcock,G.T. and Wikström,M. (1992) *Nature*, **356**, 301–309.
Bergersen,F.J. and Turner,G.L. (1990) *Proc. R. Soc. London B*, **238**, 295–320.
Bott,M., Bolliger,M. and Hennecke,H. (1990) *Mol. Microbiol.*, **4**, 2147–2157.
Bott,M., Preisig,O. and Hennecke,H. (1992) *Arch. Microbiol.*, **158**, 335–343.
Broda,E. and Peschek,G.A. (1979) *J. Theor. Biol.*, **81**, 201–212.
Buse,G. and Steffens,G.C.M. (1991) *J. Bioenerg. Biomembr.*, **23**, 269–289.
Buse,G., Hensel,S. and Fee,J.A. (1989) *Eur. J. Biochem.*, **181**, 261–268.
Cao,J., Shapleigh,J., Gennis,R.B., Revzin,A. and Ferguson-Miller,S. (1991) *Gene*, **101**, 133–137.
Capaldi,R.A. (1990) *Annu. Rev. Biochem.*, **59**, 569–596.
Cedergren,R., Gray,M.W., Abel,Y. and Sankoff,D. (1988) *J. Mol. Evol.*, **28**, 98–112.
Chan,S.I. and Li,P.M. (1990) *Biochemistry*, **29**, 1–12.
Chepuri,V., Lemieux,L., Au,D.C.T. and Gennis,R.B. (1990) *J. Biol. Chem.*, **265**, 11185–11192.
Dayhoff,M.O., Schwartz,R.M. and Orcutt,B.C. (1978) In Dayhoff,M.O. (ed.), *Atlas of Protein Sequence and Structure*. National Biomedical Research Foundation, Silver Spring, Maryland, Vol 5, Supplement 3, pp. 345–352.
Demoulin,V. (1979) *Science*, **205**, 1036–1038.
Denda,K., Fujiwara,T., Seki,M., Yoshida,M., Fukumori,Y. and Yamanaka,T. (1991) *Biochem. Biophys. Res. Commun.*, **181**, 316–322.
Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
Dickerson,R.E., Timkovich,R. and Almassy,R.J. (1976) *J. Mol. Biol.*, **100**, 473–491.
Engelman,D.M., Steiz,T.A. and Goldman,A. (1986) *Annu. Rev. Biophys. Biophys. Chem.*, **15**, 321–353.
Felsenstein,J. (1985) *Evolution*, **39**, 783–791.
Felsenstein,J. (1989) *Cladistics*, **5**, 164–166.
Fox,G.E. *et al. (1980)* Science, **209**, 457–463.
Fukaya,M., Tayama,K., Tamaki,T., Ebisuya,H., Okumura,H., Kawamura,Y., Horinouchi,S. and Beppu,T. (1993) *J. Bacteriol.*, **175**, 4307–4314.
Garcia-Horsman,J.A., Berry,E., Shapleigh,J.P., Alben,J.O. and Gennis,R.B. (1994) *Biochemistry*, **33**, 3113–3119.
Gibson,T.J., Rice,P.M., Thompson,J.D. and Heringa,J. (1993) *Trends Biochem. Sci.*, **18**, 331–333.
Green,G.N., Fang,H., Lin,R.J., Newton,G., Mather,M., Georgiou,C.D. and Gennis,R.B. (1988) *J. Biol. Chem.*, **263**, 13138–13143.
Haltia,T. (1993) *FEBS Lett.*, **335**, 294–295.
Haltia,H., Saraste,M. and Wikström,M. (1991) *EMBO J.*, **10**, 2015–2021.
Harold,F.M. (1986) *The Vital Force: a Study of Bioenergetics*. W.H.Freeman and Company, New York.
Higgins,D.G., Bleasby,A.J. and Fuchs,R. (1992) *Comput. Applic. Biosci.*, **8**, 189–191.
Hosler,J.P. *et al.* (1993) *J. Bioenerg. Biomembr.*, **25**, 121–136.
Ishizuka,M., Machida,K., Shimada,S., Mogi,A., Tsuchiya,T., Ohmori,T., Souma,Y., Gonda,M. and Sone,N. (1990) *J. Biochem.*, **108**, 866–873.
Kadenbach,B., Stroh,A., Hüther,F.-J., Reimann,A. and Steverding,D. (1991) *J. Bioenerg. Biomembr.*, **23**, 321–334.
Kasting,J.F. (1993) *Science*, **259**, 920–926.
Kelly,M., Lappalainen,P., Talbo,G., Haltia,T., van der Oost,J. and Saraste,M. (1993) *J. Biol. Chem.*, **22**, 16781–16787.
Kimura,M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.
Knoll,A.H. (1992) *Science*, **256**, 622–627.
Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.
Lappalainen,P., Aasa,R., Malmström,B.G. and Saraste,M. (1993) *J. Biol. Chem.*, **35**, 26416–26421.
Lübben,M., Kolmerer,B. and Saraste,M. (1992) *EMBO J.*, **11**, 805–812.
Lübben,M., Castresana,J. and Warne,A. (1994) *Syst. Appl. Microbiol.*, **16**, 556–559.
Margulis,L. (1981) *Symbiosis in Cell Evolution*. W.H.Freeman and Company, San Francisco.
Margulis,L., Walker,J.C.G. and Rambler,M. (1976) *Nature*, **264**, 620–624.
Mather,M.W., Springer,P. and Fee,J.A. (1991) *J. Biol. Chem.*, **266**, 5025–5035.
Mather,M.W., Springer,P., Hensel,S., Buse,G. and Fee,J.A. (1993) *J. Biol. Chem.*, **268**, 5395–5408.
Moore,G.R. and Pettigrew,G.W. (1990) *Cytochromes c. Evolutionary, Structural and Physicochemical Aspects*. Springer-Verlag, Berlin.
Moshiri,F., Chawla,A. and Maier,R.J. (1991) *J. Bacteriol.*, **173**, 6230–6241.
Müller,M. (1988) *Annu. Rev. Microbiol.*, **42**, 465–488.
Musser,S.M., Stowell,M.H. and Chan,S.I. (1993a) *FEBS Lett.*, **327**, 131–136.
Musser,S.M., Stowell,M.H. and Chan,S.I. (1993b) *FEBS Lett.*, **335**,

296–298.
O'Brian,M.R. and Maier,R.J. (1989) *Biochim. Biophys. Acta*, **974**, 229–246.
Preisig,O., Anthamatten,D. and Hennecke,H. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 3309–3313.
Quirk,P.G., Hicks,D.B. and Krulwich,T.A. (1993) *J. Biol. Chem.*, **268**, 678–685.
Raitio,M., Jalli,T. and Saraste,M. (1987) *EMBO J.*, **6**, 2825–2833.
Raitio,M., Pispa,J.M., Metso,T. and Saraste,M. (1990) *FEBS Lett.*, **261**, 431–435.
Rice,C.M., Fuchs,R., Higgins,D.G., Stoehr,P.J. and Cameron,G.N. (1993) *Nucleic Acids Res.*, **13**, 2967–2971.
Saitou,N. and Imanishi,T. (1989) *Mol. Biol. Evol.*, **6**, 514–525.
Saitou,N. and Nei,M. (1987) *Mol. Biol. Evol.*, **4**, 406–425.
Santana,M., Kunst,F., Hullo,M.F., Rapoport,G., Danchin,A. and Glaser,P. (1992) *J. Biol. Chem.*, **267**, 10225–10231.
Saraste,M. (1990) *Q. Rev. Biophys.*, **23**, 331–366.
Saraste,M., Holm,L., Lemieux,L., Lübben,M. and van der Oost,J. (1991a) *Biochem. Soc. Trans.*, **19**, 608–612.
Saraste,M., Metso,T., Nakari,T., Jalli,T., Laureaus,M. and van der Oost,J. (1991b) *Eur. J. Biochem.*, **195**, 517–525.
Schwartz,R.M. and Dayhoff,M.O. (1978) *Science*, **199**, 395–403.
Schwartz,R.M. and Dayhoff,M.O. (1979) *Science*, **205**, 1038–1039.
Shapleigh,J.P. and Gennis,R.B. (1992) *Mol. Microbiol.*, **6**, 635–642.
Smith,M.W., Feng,D.-F. and Doolittle,R.F. (1992) *Trends Biochem. Sci.*, **17**, 489–493.
Sone,N., Tano,H. and Ishizuka,M. (1993) *Biochim. Biophys. Acta*, **1183**, 130–138.
Steinrücke,P., Steffens,G.C., Panskus,G., Buse,G. and Ludwig,B. (1987) *Eur. J. Biochem.*, **167**, 431–439.
Stouthamer,A.H. (1992) *Antonie van Leeuwenhoek*, **61**, 1–33.
Towe,K.M. (1978) *Nature*, **274**, 657–661.
van der Oost,J. *et al. (1992) EMBO J.*, **11**, 3209–3217.
von Wachenfeldt,C. and Hederstedt,L. (1992) *FEMS Microbiol. Lett.*, **100**, 91–100.
Woese,C.R. (1987) *Microbiol. Rev.*, **51**, 221–271.
Woese,C.R. (1992) In Balows,A., Trüper,H.G., Dworkin,M., Harder,W. and Schleifer,K.H. (eds), *The Prokaryotes*. 2nd edition, vol. 1. Springer Verlag, New York.
Woese,C.R. and Fox,G.E. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5088–5090.
Woese,C.R., Debrunner-Vossbrinck,B.A., Oyaizu,H., Stackebrandt,E. and Ludwig,W. (1985) *Science*, **229**, 762–765.
Woese,C.R., Kandler,O. and Wheelis,M.L. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 4576–4579.
Yang,D., Oyaizu,H., Olsen,G.J. and Woese,C.R. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 4443–4447.
Zimmermann,B.H., Nitsche,C.I., Fee,J.A., Rusnak,F. and Münck,E. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 5779–5783.

Received on January 27, 1994; revised on March 4, 1994