The Multiple Evolutionary Histories of Dioxygen Reductases: Implications for the Origin and Evolution of Aerobic Respiration

Celine Brochier-Armanet,*† Emmanuel Talla,‡‡ and Simonetta Gribaldo§

*Université de Provence, Aix-Marseille I, Marseille, France; †Laboratoire de Chimie Bactérienne, Institut de Microbiologie de la Méditerranée (IF88), Marseille, France; ‡Université de la Méditerranée, Aix-Marseille II, Marseille, France; and §Unité de Biologie Moléculaire Chez les Extremophiles, Département de Microbiologie, Institut Pasteur, Paris, France

Understanding the origin and evolution of cellular processes is fundamental to understand how biological activity has shaped the history of our planet. Among these, aerobic respiration is probably one of the most debated. We have applied a phylogenomics approach to investigate the origin and evolution of dioxygen reductases (O$_2$Red), the key enzymes of aerobic respiratory chains. The distribution and phylogenetic analysis of the four types of O$_2$Red (Cyt-bd and the A, B, and C families of heme–copper O$_2$Red) from 673 complete bacterial and archaeal genomes show that these enzymes have very different evolutionary histories: Cyt-bd are of bacterial origin and were transferred to a few archaea; C-O$_2$Red are of proteobacterial origin and were transferred to a few other bacteria; B-O$_2$Red are of archael origin and were transferred to a few bacteria; and A-O$_2$Red are the most ancient O$_2$Red and were already present prior to the divergence of major present-day bacterial and archael phyla, thus before the emergence of Cyanobacteria and oxygenic photosynthesis. Implications for the origin and the evolution of aerobic respiration are discussed.

Introduction

Respiration is a fundamental cellular process that consists in the transfer of electrons to a final acceptor through membrane-associated redox cofactors and protein complexes. The free energy released during the electron transfer process is coupled to active proton translocation across the membrane, generating an electrochemical gradient that can be used for adenosine triphosphate synthesis. Today, many bacteria and archaea are capable of reducing dioxygen (O$_2$), whereas eukaryotes reduce O$_2$ by means of their mitochondria, which are of proteobacterial origin. The terminal electron acceptors of aerobic respiratory chains are dioxygen reductases (hereafter referred to as O$_2$Red). These membrane-bound enzymes catalyze the reduction of O$_2$ to water by using electrons provided by either a quinol derivate or a cytochrome c (Wikstrom 1977). This reaction is generally coupled to active proton translocation across the membrane (Branden et al. 2006). O$_2$Red belong to two unrelated protein superfamilies: the cytochrome bd (Cyt-bd) (Watanabe et al. 1979) and the heme–copper superfamily.

The heme–copper superfamily contains three families of O$_2$Red named A, B, and C according to a recent classification based on sequence similarity and phylogenetic analysis proposed by Pereira et al. (2001). These O$_2$Red are generally able to pump protons through the membrane (Pitcher and Watmough 2004; Branden et al. 2006). They are also called cytochrome c oxidases or quinol oxidases according to their electron donor: cytochrome bc$_1$ (complex III) via cytochrome c or a quinol derivate, respectively. The number of subunits composing heme–copper O$_2$Red is variable, but all contain at least a catalytic subunit. The catalytic subunits of A-, B-, and C-O$_2$Red are homologous and are characterized by a binuclear center including a Cu$_{B}$ (Kappler et al. 2005) that is electronically coupled to a high-spin heme and a low-spin heme (Pereira et al. 2001). The low-spin heme is the immediate electron donor for the binuclear center where the catalytic reaction takes place (Pereira et al. 2001). The catalytic subunits of heme–copper O$_2$Red can contain various types of hemes (A, B, O, and derivates of A and O hemes). However, it has been shown that it is quite inappropriate to classify O$_2$Red based on the type of heme carried by the catalytic subunit, nor on the type of electron donor (i.e., cytochrome c or quinol derivate) (Pereira et al. 2001).

The A-O$_2$Red is the largest and best-studied subfamily. It was first identified in bacteria and includes mitochondrial O$_2$Red. It includes most of the so-called Cox-type enzymes, such as, for example, Escherichia coli b$_{5}$O$_2$Red, Bacillus subtilis caa$_{3}$ and ada$_{5}$O$_2$Red, and Sulfolobus acidocaldarius SoxM (Pereira et al. 2001). A-O$_2$Red form a complex called complex IV that is composed of a variable number of subunits, 3 or 4 in bacteria (Ostermeier et al. 1996) and up to more than 10 in mammalian mitochondria (Tsukihara et al. 1996). However, all harbor at least a catalytic and a noncatalytic subunit (sometimes called Cox1 and Cox2, respectively, although other names can be used in different organisms, see Castresana et al. [1994] for examples of alternative names). Most aerobic bacteria possess either one or several A-O$_2$Red. For example, E. coli harbors only one quinol oxidase (Chepuri et al. 1990), whereas B. subtilis harbors both a cytochrome c and a quinol oxidase (Winstedt and von Wachenfeldt 2000). A-O$_2$Red have also been described in some archaea such as S. acidocaldarius (Lubben et al. 1994, see also Hemp and Gennis 2008).

B-O$_2$Red were first identified in the archaeon S. acidocaldarius (Lubben et al. 1992) and have been subsequently described in other archaea and in a few bacteria. Similarly to A-O$_2$Red, B-O$_2$Red also include Cox-type enzymes but form a distinct family based on phylogenetic analysis including, for example, S. acidocaldarius SoxB, Thermus thermophilus and Natronomonas pharaonis b$_{5}$ O$_2$Red, Acidianus ambivalens aa$_{2}$O$_2$Red, and Bacillus stearothermophilus ba(o) O$_2$Red (Pereira et al. 2001). Interestingly, some bacteria that possess a B-O$_2$Red have also an A-O$_2$Red. For example, T. thermophilus contains a B-O$_2$Red, which is expressed in microaerobic conditions, in addition to an A-O$_2$Red that is expressed at high oxygen levels (Mather et al. 1993; Keightley et al. 1995).
C-O2Red include cbb3 enzymes (Pereira et al. 2001). The first representative of this family was characterized in the N2-fixing proteobacterium Bradyrhizobium japonicum, where it is expressed under microaerophilic conditions and has a higher affinity for O2 compared with the other O2Red present in this organism (Preisig et al. 1993; Poole and Cook 2000, see also Pitcher et al. 2002; Pitcher and Watmough 2004 and references therein). C-O2Red are present in some Proteobacteria, including a number of human pathogens (e.g., Campylobacter jejuni, Helicobacter pylori, Pseudomonas aeruginosa, and Vibrio cholerae) where they may be required for the successful colonization of anoxic tissues and for pathogenicity (see Pitcher et al. 2002 and references therein). Homologues have also been found in a few nonproteobacterial genomes such as Gemmatia obscuriglobus (Planctomycetales) and Cytophaga hutchinsonii (Bacteroidetes/Chlorobi) (Pitcher et al. 2002; Pitcher and Watmough 2004).

Five new O2Red families (D, E, F, G, and H) have recently been proposed (Hemp and Gennis 2008). However, because most of these families contain only one or very few sequences that are very likely related to the B-O2Red family (Pereira et al. 2008) and have not been biochemically fully characterized, we reckon that this classification is premature, and we will not adopt it in this work. Cyt-bd are quinol oxidases composed of two subunits (CydA and CydB) that contain three types of heme redox centers: heme b-558 in CydA and heme b-595 and heme d in CydB (Junemann 1997). In contrast to heme–copper O2Red, Cyt-bd are not capable of pumping protons but generate a membrane potential because they reduce O2 to water by using protons from the cytoplasm (Hart et al. 2005). Cyt-bd are mainly found in bacteria, and their affinity for O2 depends on the considered organisms (Junemann 1997; Pils and Schmetterer 2001). Although homologues of Cyt-bd have been detected in a few archaea, little or no experimental data on their function are currently available.

Understanding the origin of O2Red can provide precious hints for the emergence of aerobic respiration. However, this is still debated. Following the discovery and characterization of O2Red in some archaea in the 1990s (Lubben et al. 1992, 1994), phylogenetic analyses suggested that both Cyt-bd and heme–copper O2Red are very ancient and were already present in the ancestor of bacteria and archaea (Castresana et al. 1994, 1995; Castresana 2001). Conversely, subsequent analyses proposed that Cyt-bd, A-O2Red, and B-O2Red originated in bacteria and were later acquired by archaea through horizontal gene transfer (HGT) (Hao and Golding 2006; Pereira et al. 2001; Hemp and Gennis 2008, respectively).

Here we have investigated the distribution and phylogeny of the catalytic subunits of all four O2Red types (Cyt-bd, heme–copper A-, B-, and C-O2Red) in 673 complete archaeal and bacterial genomes. Our results show that O2Red are not all ancient nor all bacterial in origin but that each has a singular evolutionary history: Cyt-bd are of bacterial origin and were extensively transferred within bacteria and to a few archaea; C-O2Red are even more recent because they likely originated in Proteobacteria and were subsequently transferred to a few other bacteria. Conversely, B-O2Red are of likely archaeal origin and were transferred to a few bacteria. Finally, A-O2Red are the most ancient of the four O2Red types because they likely originated prior to the divergence of archaea and bacteria. Importantly, clearly identifiable HGT events do not confound the subsequent evolutionary history of A-O2Red, which indicates that these enzymes were largely maintained throughout bacterial and archael diversification, suggesting a selection pressure for the conservation of the native copy during evolution.

Material and Methods

For preliminary analyses, the phylogenies of the four O2Red types were constructed using all homologous sequences retrieved from the nr database at the National Center for Biotechnology Information (NCBI). In a second time, we restricted the phylogenetic analyses to homologues from the 673 completely sequenced archaeal and bacterial genomes available at the NCBI in April 2008 (ftp://ftp.ncbi.nih.gov/). The hmmer package (Durbin et al. 1998) was used to identify functional domains in each O2Red. The Pfam hmm profile PF00115.10 (COX1, cytochrome c, and quinol oxidase polypeptide I) is present in the catalytic subunits of A-, B-, and C-O2Red, whereas the Pfam hmm profile PF01654.8 (bacterial cytochrome ubiquinol oxidase) is found in the catalytic subunit of Cyt-bd. These profiles were used to search for protein sequences from complete genomes containing the corresponding domains (Bateman et al. 2004; Finn et al. 2006). Alignments with the Pfam hmm profile domains having an E value less than 0.1 were considered significant. Additional BlastP searches using various seeds were performed to complete each data set. A total of 623, 718, 64, and 216 homologues of Cyt-bd, A-, B-, and C-O2Red catalytic subunits were retrieved, respectively.

For phylogenetic analyses, sequences were aligned by ClustalW and MUSCLE (Thompson et al. 1994; Edgar 2004) and manually refined using ed from the MUST package (Philippe 1993). Regions where homology was doubtful were removed from further analysis. For phylogenetic analyses of Cyt-bd, A-, B-, and C-O2Red catalytic subunits, respectively, 328, 401, 283, and 400 position were kept. Maximum likelihood (ML) phylogenetic trees were inferred using phylm (Guindon and Gascuel 2003), with a Whelan and Goldman (WAG) evolutionary model, an estimated proportion of invariant sites and a gamma correction to take into account the heterogeneity of the evolutionary rates between sites (four categories). Bootstrap analyses were performed using phylm on 1,000 replicates of each original data set (Guindon and Gascuel 2003).

Results

In April 2008, 673 complete genomes were available from representatives of 13 major bacterial phyla (Proteobacteria [with the five main divisions α-Proteobacteria, β-Proteobacteria, δ-Proteobacteria, ε-Proteobacteria, and γ-Proteobacteria], Firmicutes, Actinobacteria, Bacteroidetes/Chlorobi, PVC [Planctomycetales, Verrucomicrobia, Chlamydiales], Cyanobacteria, Chloroflexi, Spirochaetes, Thermotogae, Deinococcus/Thermus, and the still poorly
represented Acidobacteria [two genomes], Aquificae [one genome], and Fusobacteria [one genome]) and the 3 archaeal phyla (Crenarchaeota including the recently sequenced korarchaeon [Elkins et al. 2008], Euryarchaeota including Nanoarchaeota [Brochier et al. 2005], and Thaumarchaeota [Brochier-Armanet et al. 2008]) (Supplementary Material 1, Supplementary Material online). The growing large number of complete genome sequences allows applying accurate phylogenomics approaches to clarify the evolutionary history of prokaryotic cellular processes/structures (Raymond et al. 2004; Bapteste et al. 2005; Desmond et al. 2007). The strategy is to analyze the distribution and phylogeny of genes coding for the components of the studied system in complete genomes in order to detect their timing and lineage of origin as well as their subsequent evolution (e.g., vertical inheritance, gene duplication, gene loss, and HGT) (Eisen and Fraser 2003).

Although the order of emergence of the main prokaryotic phyla is not fully resolved, it is possible to apply phylogenomics criteria to infer the presence or the absence of each O2Red type in the ancestors of each of the 16 main bacterial and archaeal phyla. We inferred the presence of an O2Red type in the most recent ancestor of a bacterial or archaeal phylum when 1) homologues of its catalytic subunit are widely distributed among the members of the phylum and 2) their sequences form a well-supported monophyletic group in the corresponding phylogenetic gene tree. Conversely, we inferred the absence of a given O2Red type in the most recent ancestor of a bacterial or archaeal phylum when 1) no homologues of its catalytic subunit are present in any of its members, 2) homologues are restricted only to a particular subgroup, or 3) their phylogenetic analysis indicates acquisition by HGT after the diversification of the phylum (e.g., the presence of strongly supported monophyletic clusters mixing up representatives from different phyla).

We retrieved a total of 1,640 O2Red homologues, which highlights the great abundance and redundancy of O2Red in prokaryotic genomes. In fact, only 122 genomes of the 673 analyzed do not contain any homologues of the catalytic subunit of known O2Red types (26 of 54 archaeal and 96 of 619 bacterial genomes, Supplementary Material 1, Supplementary Material online). The identification of a gene coding for the catalytic subunit of a particular O2Red type was nearly always accompanied by that of a gene coding for the corresponding noncatalytic subunit, generally lying close by in the genome (data not shown), indicating the presence of a functional enzyme. The use of complete genomes is essential to phylogenomic approaches because it allows determining the exact distribution of homologues. However, for each O2Red type, we compared the taxonomic distribution and phylogenies with those built on sequences from the nr database (data not shown). No important differences were evident between the two analyses, meaning that the use of complete genomes only does not bias the interpretation of results.

Origin and Evolution of Cyt-bd

We detected 639 homologues of the Cyt-bd catalytic subunit (CydA) (Supplementary Material 1, Supplementary Material online). These are present in 418 of 619 bacterial genomes (~68%) and in 15 of 54 archaeal genomes (~28%). Homologues of CydA are found in members of all bacterial phyla, except Thermotogales, Spirochaetes, and Fusobacteria (Supplementary Material 1, Supplementary Material online). In archaea, CydA homologues are present in two euryarchaeal orders, Halobacteriales (but only two Halobacterium sp. among the six complete genomes available) and all available Thermoplasmatales (Supplementary Material 1, Supplementary Material online). Interestingly, Halobacterium sp. NRC-1 has four copies of the CydA-coding gene, which are located on two different plasmids (pNRC100 and pRNC200). Thermoplasmatales have two copies each of CydA. Interestingly, the two Thermoplasma appear to use only Cyt-bd for aerobic respiration because they lack homologues of all other O2Red (Supplementary Material 1, Supplementary Material online). Intriguingly, although methanogens are described as strict anaerobes, Methanosarcina Barkeri and Methanosarcina acetivorans (but not their close relatives Methanosarcina mazei and Methanococcoides burtonii) harbor homologues of CydA. Because they also harbor a CydB homologue, this suggests the presence of a functional enzyme that would be very interesting to study. Similarly, the genomes of some other archaea described as anaerobes contain highly divergent homologues of CydA (Caldovirga maquilensis, Hyperthermus butylicus, Pyrobaculum arsenaticum, Pyrobaculum islandicum, Thermofilum pendens, Thermoproteus neutrophilus, Candidatus Korarchaeum cryptofilum, and Archaeoglobus fulgidus, indicated with “d” in Supplementary Material 1, Supplementary Material online). However, no CydB homologues are detectable in these archaea, and these organisms may not harbor a functional Cyt-bd.

According to the large number of sequences and the limited number of unambiguously aligned positions available for phylogenetic analysis (328 amino acids), the ML unrooted tree of CydA is only partially resolved, especially at the most basal nodes (fig. 1A and Supplementary Material 2, Supplementary Material online). However, the presence of well-supported clusters composed of CydA homologues from different bacterial phyla (gray/white-striped triangles, fig. 1A) suggests that recent HGT has played a major role in the evolution of Cyt-bd in bacteria. Nevertheless, three monophyletic clusters corresponding to three major bacterial phyla are apparent: the PVC group (bootstrap value [BV] = 569), including a well-supported monophyletic group of Chlamydia (~870); Actinobacteria (BV = 849); and Cyanobacteria (BV = 1,000) (fig. 1A and Supplementary Material 2, Supplementary Material online). In the case of PVC and Actinobacteria, the species included in these groups are widely distributed within the corresponding phyla (i.e., 14 of 15 genomes for PVC and 39 of 50 genomes for Actinobacteria [Supplementary Material 2, Supplementary Material online]) and are not restricted to a particular subgroup (Supplementary Material 1, Supplementary Material online). According to phylogenomics criteria, this suggests that a Cyt-bd was present at least in the ancestor of each of these two groups (fig. 1B). Unfortunately, the poor resolution of the tree leaves the situation less clear for the remaining bacterial phyla, such as Firmicutes or Proteobacteria, because their sequences do not form strongly supported monophyletic
clusters (fig. 1A and B and Supplementary Material 2, Supplementary Material online).

Interestingly, archaeal sequences do not group together (fig. 1A). Moreover, CydA homologues from Thermoplasmatales and Methanosarcinaceae emerge from within two different clusters of Firmicutes sequences, suggesting that these archaea independently acquired their CydA from Firmicutes (fig. 1A and Supplementary Material 2, Supplementary Material online). The multiple copies of CydA from Halobacterium sp. cluster together (BV = 1,0000, fig. 1A), indicating that they arose from species-specific duplications. The fact that they lie on plasmids and that no other member of Haloarchaea has a CydA homologue with respect to the number of genomes harboring at least one CydA homologue with respect to the number of genomes available. Asterisks indicate the presence in the corresponding groups of highly divergent homologues of CydA that were not included in the phylogenetic analysis. Filled triangles indicate phyla whose ancestor harbored a Cyt-bd according to phylogenomics criteria.

We detected 216 homologues of the C-O2Red catalytic subunit but exclusively in bacteria. Moreover, these enzyme was already present in their ancestor (Castresana 2004). However, our phylogenomic analysis of CydA highlights a different picture. In fact, the abundance of CydA homologues in all major bacterial phyla but not in archaea, combined with the phylogenomic inference of a Cyt-bd in the ancestors of at least two major bacterial phyla (i.e., Actinobacteria and PVC), indicates that Cyt-bd originated likely early in bacteria and were subsequently largely transferred among bacteria and to a few archaea, in agreement with a recent analysis (Hao and Golding 2006).

Origin and Evolution of C-O2Red

We detected 216 homologues of the C-O2Red catalytic subunit but exclusively in bacteria. Moreover, these
are mainly from Proteobacteria. Indeed, 48% of available proteobacterial genomes (18 of 19 e-Proteobacteria, 4 of 19 δ-Proteobacteria, 45 of 83 α-Proteobacteria, 34 of 53 β-Proteobacteria, and 58 of 157 γ-Proteobacteria) contain at least one homologue of the C-O 2Red catalytic subunit (Supplementary Material 1, Supplementary Material online). In contrast, only 17 homologues are present in 15 of the 288 (~5%) other bacterial genomes (1 of 2 Acidobacteria, 7 of 17 Bacteroidetes/Chlorobi, 2 of 15 PVC, 2 of 32 Cyanobacteria, 1 of 144 Firmicutes, and 2 of 9 Spirochaetes) (Supplementary Material 1, Supplementary Material online). In the corresponding unrooted ML phylogeny (fig. 2A), all homologues from e-Proteobacteria form a monophyletic group (BV = 1,000 %), as well as those from all α-, γ-, and most β-Proteobacteria (BV = 959 %), within which all α-proteobacterial sequences cluster together (BV = 960 %, Supplementary Material 3, Supplementary Material online). Because the relationships among these proteobacterial lineages are in agreement with the species phylogeny, phylogenomics criteria strongly suggest that the most recent ancestor of Proteobacteria had a C-O 2Red (fig. 2B). This implies that its absence in some proteobacterial lineages (such as Enterobacteriales or Rickettsiales, Supplementary Material 1, Supplementary Material online) likely results from secondary losses. On the contrary, the 17 nonproteobacterial sequences form heterogenous groups that are not consistent with species phylogeny (fig. 2A and B and Supplementary Material 3, Supplementary Material online). For example, the nine sequences from Bacteroidetes/Chlorobi do not group together, suggesting that they have different evolutionary origins, and the same is observed for the PVC sequences (fig. 1B and Supplementary Material 3, Supplementary Material online). This strongly suggests that these C-O 2Red sequences were acquired by HGT. Moreover, most of them are highly divergent (fig. 1B and Supplementary Material 3, Supplementary Material online), consistently with a shift in evolutionary rate following their acquisition by HGT.

Ducluzeau et al. (2008) recently published an analysis focusing on C-O 2Red. However, although they obtained a tree very similar to ours, these authors interpreted it as favoring an early—but undetermined—origin in bacteria.

**Fig. 2.**—(A) Schematic representation of the unrooted ML tree of the catalytic subunit of C-O2Red (216 sequences and 400 amino acid positions). The complete tree is provided as Supplementary Material 3 (Supplementary Material online). Numbers at nodes indicate bootstrap values for 1,000 replicates of the original data set. For additional details, see the legend to figure 1. (B) Schematic representation of the number and taxonomic distribution of complete prokaryotic genomes. For additional details, see the legend to figure 1.
(Ducluzeau et al. 2008). An even more ancient origin of C-O2Red, before the divergence of archaea and bacteria, was put forward (Castresana et al. 1995; Castresana 2004). However, the very limited taxonomic distribution of C-O2Red—virtually restricted to Proteobacteria—and its phylogenetic analysis strongly suggest that these enzymes emerged in Proteobacteria and were then transferred by HGT to a few other bacteria. In fact, if a C-O2Red was ancestrally present in these few nonproteobacterial lineages or even in the ancestor of archaea and bacteria as previously suggested (Castresana and Saraste 1995), this would imply the presence of multiple paralogues in these organisms and massive, independent, and recent losses in the majority of archaeal and bacterial phyla. In contrast, the incongruence of the C-O2Red tree with the reference species phylogeny is much more easily explained with a few recent HGTs.

Origin and Evolution of B-O2Red

We identified a relatively large number of B-O2Red catalytic subunits in archaea (25 homologues in 8 of the 15 available crenarchaeal genomes and in 5 of the 35 available euryarchaeal genomes), whereas these are absent in the two Thaumarchaeota (Supplementary Material 1, Supplementary Material online). More precisely, at least one B-O2Red is present in members of three of four crenarchaeal orders (with multiples copies in Sulfolobales) but only in one euryarchaeal order (Halobacteria) (Supplementary Material 1, Supplementary Material online). In contrast, bacterial B-O2Red are scarce (only 39 homologues) and belong to very distantly related genomes (1 in Aquifex aeolicus, 1 of 50 Actinobacteria, 1 of 17 Bacteroidetes/Chlorobi, 3 of 7 Chloroflexi, 3 of 4 Deinococci/Thermus, 5 of 136 Firmicutes, 6 of 83 z-Proteobacteria, 13 of 53 β-Proteobacteria, 2 of 19 e-Proteobacteria, and 2 of 157 γ-Proteobacteria, Supplementary Material 1, Supplementary Material online), whereas they are completely absent in the seven remaining bacterial phyla (Supplementary Material 1, Supplementary Material online).

The ML unrooted tree of the B-O2Red catalytic subunit indicates the presence of two copies in the ancestor of Sulfolobales, resulting from an ancient gene duplication event (indicated by a black diamond symbol, fig. 3A). This duplication event was followed by multiple species-specific duplications leading to the large number of paralogues in present-day Sulfolobales (Supplementary Materials 1 and 4, Supplementary Material online). Interestingly, these paralogues represent the recently proposed D- to H-O2Red units (with multiples copies in Sulfolobales) but only in one euryarchaeal order (Halobacteria) (fig. 3A). Thus, similarly to C-O2Red, phylogenomic of B-O2Red does not indicate -as previously suggested- the presence of this enzyme in the ancestor of bacteria(ref), hence, even less in the ancestor of archaea and bacteria (Castresana and Saraste 1995; Castresana 2001, 2004). Indeed, our data suggest that B-O2Red probably originated in Crenarchaeota and subsequently went through interdomain HGT (i.e., from crenarchaeota to bacteria and then possibly from bacteria to euryarchaeota [Halobacteria]).

Origin and Evolution of A-O2Red

Contrarily to B- and C-O2Red, we identified homologues of the A-O2Red catalytic subunit, sometimes in several copies, in members of nearly all bacterial phyla (except in Thermotogales, e-Proteobacteria, and the poorly represented Fusobacteria) and in all archaeal phyla for which complete genome sequences are available (Supplementary Material 1, Supplementary Material online). In fact, these are present in 432 of the 673 genomes analyzed (~64%), which confirms that members of the A family are the most widespread O2Red.

In contrast to the partially resolved phylogenies of the catalytic subunits of the other O2Red, the ML tree of the A-O2Red catalytic subunit shows a number of robustly supported monophyletic groups that correspond to major bacterial phyla (fig. 4A). These are Actinobacteria (BV = 1,000%, Cyanobacteria (BV = 908%), Deinococcus/Thermus (BV = 1,000%), z-Proteobacteria (BV = 1,000%), and β + γ-Proteobacteria (BV = 996%), the last two phyla being sister groups (BV = 991%). Because these
groups include sequences from species that are widely distributed within the corresponding phyla, phylogenomics criteria strongly suggest that the ancestor of each of these phyla had an A-O2Red (fig. 4B). The presence of an A-O2Red can also be inferred in Firmicutes and δ-Proteobacteria. In fact, Firmicutes form two sister groups (Firmicutes I and II, BV = 916%o, fig. 4A). The Firmicutes I group (Firmicutes I, BV = 779%o) include the quinol oxidase of B. subtilis (Supplementary Material 5, Supplementary Material online), whereas the Firmicutes II group (Firmicutes II, BV = 955%o) is composed of additional A-O2Red catalytic subunit sequences from Bacillales and includes the cytochrome c oxidase from B. subtilis (Supplementary Material 5, Supplementary Material online). This suggests a single duplication event in Bacillales followed by a functional shift from a cytochrome c oxidase to a quinol oxidase activity, confirming previous hypotheses (see Castresana et al. 1994 for a detailed discussion). Firmicutes I are sister to a well-supported monophyletic group of α-β-γ-Proteobacteria (including the quinol oxidase of E. coli) plus a sequence from the PVC Proteochlamydia amoebophila (α-β-γ-Proteobacteria I, BV = 999%o, fig. 4A and Supplementary Material 5, Supplementary Material online). This can be explained by invoking a single HGT of an A-O2Red from Firmicutes to a proteobacterium, followed by HGT both among Proteobacteria and to Proteochlamydia. Interestingly, this HGT was likely followed by the loss of the native copy in Enterobacteria.

δ-Proteobacterial homologues also form a well-supported group (BV = 1,000%o, fig. 4) that also includes a few representatives of other bacterial lineages (i.e., eight α-β-γ-Proteobacteria, two Acidobacteria, four Spirochaetes, one representative of Chloroflexi, and one representative of Bacteroidetes/Chlorobi, Supplementary Material 5, Supplementary Material online). Here, a similar interpretation as for Firmicutes (i.e., presence of an A-O2Red in the last ancestor of δ-Proteobacteria) can be put forward but involving more than one initial HGT to other bacteria. Interestingly, this is in agreement with the recent proposition that the ancestor of δ-Proteobacteria was an aerobic organism (Thomas et al. 2008). The presence of an A-O2Red in the ancestors of Acidobacteria, Chloroflexi, Bacteroidetes/Chlorobi, and Aquificales seems possible but needs to be confirmed when the complete genomes of additional members of these groups will be available. In contrast, the absence of an A-O2Red in all representatives of ε-Proteobacteria, Thermotogales, and Fusobacteria suggests that the ancestors of these lineages did not possess this enzyme (fig. 4B). The same conclusion
can be reached for Spirochaetes and the PVC group because homologues are only restricted to a few representatives of these phyla and were probably acquired by independent HGT from different proteobacterial groups (fig. 4A and Supplementary Material 5, Supplementary Material online).

However, the presence of an A-O2Red in the ancestors of some of these phyla cannot be definitively discarded because of the limited number of representatives that are currently available. Finally, two well-supported additional groups of α-β-γ-Proteobacteria are observed (fig. 4A). The origin of the first additional group (α-β-γ-Proteobacteria II, BV = 995%) is probably linked to an HGT from Bacteroidetes/Chlorobi, whereas the origin of the second additional group (α-β-γ-Proteobacteria III, BV = 969%) is less clear. To sum up, the phylogenomic analysis of A-O2Red indicates that this enzyme was present at least in the most recent ancestors of each major bacterial phylum (i.e., Deinococci/Thermus, Cyanobacteria, Actinobacteria, Firmicutes, δ-Proteobacteria, α-β-γ-Proteobacteria, and possibly in the ancestors of Aquificales, Bacteroidetes/Chlorobi, Chloroflexi, and Acidobacteria, fig. 4B). Accordingly, its absence in some members of these phyla reflects secondary gene losses. This is compatible with the presence of the enzyme in the most recent bacterial ancestor and that its absence in ε-Proteobacteria, Thermotogales, Fusobacteria, Spirochaetes, and the PVC group may be the result of ancient losses.

In archaea, homologues of A-O2Red catalytic subunits are present in the genomes from representatives of three of four crenarchaeal orders (Sulfolobales, Desulfurococcales, and Thermoproteales), in the two genomes from Thaumarchaeota, and in two euryarchaeal orders (some Thermoplasmatales and all Halobacteriales) (Supplementary Material 1, Supplementary Material online). However, in contrast to Cyt-bd and B-O2Red, archaeal homologues do not appear to have been recently acquired from bacteria as it was previously suggested (Pereira et al. 2001; Hemp and Gennis 2008). Moreover, the presence of an A-O2Red in Thaumarchaeota and Crenarchaeota suggests that this enzyme was present in the common ancestor to these two archaeal phyla, which may correspond to the ancestors of all archaea (Brochier-Armanet et al. 2008). The scarcity of available genomes from aerobic representatives of Eurarchaeota (restricted to Thermoplasmatales and Halobacteriales) prevents us to settle definitively if an A-O2Red was present in the ancestor of Eurarchaeota and was independently lost in anaerobic lineages (such as Thermococcales).
and methanogens) or if this enzyme was acquired by Thermoplasmatales and Halobacteriales through HGT from crenarchaeota.

Finally, the presence of an A-O₂Red in the ancestor of bacteria and the possible presence of an A-O₂Red in the ancestor of archaea is compatible with the proposition that an A-O₂Red was already present in the ancestor of bacteria and archaea (Castresana and Saraste 1995; Castresana 2004). Importantly, phylogenetic analysis indicates that, despite a few HGTs, A-O₂Red were mainly vertically inherited from the ancestors of most bacterial phyla through their subsequent diversification up to today. This suggests a strong selective pressure to maintain this enzyme during evolution.

Discussion

Aerobic respiration is a fundamental and widespread biological process. Indeed, phylogenomic analysis indicates the possible ancestral absence of any O₂Red type in only three bacterial phyla (i.e., Thermotogales, Fusobacteria, and Spirochaetes), albeit these are still poorly represented among complete genome sequences. Moreover, our phylogenomic analysis highlights the high level of redundancy of O₂Red in prokaryotes. For example, ~20% and ~58% of the 673 genomes analyzed contain at least three or two of the four types of O₂Red, respectively. Interestingly, some organisms exhibit a high level of redundancy for O₂Red. For example, 96 genomes harbor from 5 to 10 O₂Red-coding genes (Supplementary Material 1, Supplementary Material online). Surprisingly, O₂Red homologues are present in organisms described as strict anaerobes (e.g., methanogens, some δ-Proteobacteria), suggesting that these may be capable of using oxygen in particular conditions, as it was in some cases experimentally proven (Lemos et al. 2001; Baughn and Malamy 2004; Das et al. 2005). Present-day atmospheric oxygen originates mainly from water photolysis through the oxygenic photosynthetic activity of Cyanobacteria and photosynthetic eukaryotes (with their chloroplasts of cyanobacterial origin), whereas the abiotic production of O₂ from water photolysis by UV light is considered negligible. Following this logic, it is generally assumed that aerobic respiration arose only after the appearance of oxygenic photosynthesis, in response to the accumulation of biologically produced O₂ (Battistuzzi et al. 2004). This is at odds with the hypothesis of the presence of all four O₂Red types in the ancestors of bacteria and archaea (Castresana 2004). Consequently, the hypothesis that O₂Red originated in bacteria and were later transferred to archaea may appear more reasonable. However, our results indicate that the four O₂Red types are not all ancestral to the divergence between bacteria and archaea, nor do they all have a bacterial origin but have, instead, very different evolutionary histories: Cyt-bd are of bacterial origin, possibly ancient, and were transferred many times among bacteria and to archaea (fig. 5A). Interestingly, and contrary to what is largely assumed (Castresana et al. 1994; Pereira et al. 2001; Castresana 2004), B-O₂Red are likely of crenarchaeal origin and were transferred to a few bacteria (fig. 5B). These interdomain HGT events show that, in contrast to what was previously proposed (Castresana 2001), difference in lipid membrane compositions between bacteria and archaea is not a barrier to HGT of membrane-bound proteins. This has indeed been shown, for example, with bacteriorhodopsins (Frigaard et al. 2006). C-O₂Red have a recent origin because they likely emerged in the proteobacterial lineage and were later transferred to a few bacteria (fig. 5C). Finally, A-O₂Red are the most ancient O₂Red because they likely originated prior to the divergence between archaea and bacteria, thus before the emergence of oxygenic photosynthesis in the cyanoarchaeal lineage and were largely maintained through the subsequent diversification of bacterial and archaeal phyla (fig. 5D).

To sum up, the origin of A-O₂Red preceded that of B-O₂Red and C-O₂Red, as well as of Cyt-bd (fig. 6). If the ancestral function of A-O₂Red was O₂ reduction (fig. 6A), this would indicate that the divergence and subsequent diversification of present-day prokaryotic phyla occurred in a fairly regular—albeit low—exposure to O₂. Nevertheless, we wish to stress that the presence of a single O₂Red does not imply an obligate aerobic lifestyle but is compatible with a microaerophilic, facultative aerobic or even anaerobic but aerotolerant lifestyle. For example, the first O₂Red might have been used to scavenge O₂. The antiquity of A-O₂Red with respect to B- and C-types might appear incongruent with their respective affinity for O₂. In fact, it might be expected that O₂Red with higher affinity for O₂ would have coped better with the initial low O₂ levels in the early Earth atmosphere and oceans. This is in fact a usual argument for inferring the antiquity of B- and C-types with respect to A-O₂Red (Castresana and Saraste 1995). Nevertheless, the antiquity of A-O₂Red does not exclude that they might have had initially higher O₂ affinity. The possibility of a divergence of early microbiota in the presence of low quantities of O₂ should be taken into account in theoretical models of early archaean atmosphere and life evolution. Moreover, this implies that the ability for O₂ reduction did not emerge in response to the growing availability of O₂ produced by Cyanobacteria but rather that it was an important key adaptation that would have allowed the very emergence of oxygenic photosynthesis. Finally, the ancestral presence of an O₂Red type indicates that present-day anaerobes for which complete genomes are available are only secondarily so.

However, it cannot be excluded that A-O₂Red originally reduced other compounds (Pereira et al. 2001) (fig. 6B). Indeed, some O₂Red can use NO instead of O₂ under particular laboratory conditions (Giuffre et al. 1999; Hendriks et al. 2000; Forte et al. 2001). An additional link between O₂Red and NO reduction is suggested by the fact that NO reductases (NORs) are distant homologues of O₂Red (Castresana and Saraste 1995; Zumft 2005). Accordingly, O₂Red have been proposed to derive from NORs (Castresana and Saraste 1995; de Vries and Schroder 2002). If A-O₂Red originally reduced other compounds, a functional shift to O₂ reduction would have had to occur many times independently in most archaean and bacterial phyla after the emergence of oxygenic photosynthesis in Cyanobacteria and the rise of O₂ (fig. 6B). However, this would not explain the loss of A-O₂Red in ancestrally anaerobic phyla such as Thermotogales or Spirochaetes instead of being kept with their original function. Moreover, the
phylogenetic analysis of NORs strongly suggests an evolutionary history similar to that of C-O2Red, that is, it originated within Proteobacteria and spread by HGT among Proteobacteria, to a few representatives of nonproteobacterial phyla and to a few archaea (Supplementary Material 6, Supplementary Material online). NORs are thus very likely more recent than A-O2Red. Finally, it may be argued that ancestral A-O2Red possessed a broad substrate range (perhaps including both NO and O2). In this case, no functional shifts would have occurred but rather multiple specializations to reduce O2. Nevertheless, an ancestral broad specificity also implies that an amount of O2 would have been available during early microbial evolution. In order to get more insights on this issue, it will be essential to explore further the functional diversity and function of O2Red from a broad sampling of bacteria and archaea.

Conclusions

Elucidating the evolutionary history of major biological processes is a major challenge to understand the evolution of...
life in the distant past of our planet. However, although phylogenomics approaches are potentially very useful, their utilization to study the origin and evolution of cellular processes—including metabolic pathways—remains currently limited. One of the principal reasons is linked to the a priori idea that understanding the origin and evolution of “operational” genes (i.e., coding for proteins not involved in replication, transcription, and translation) is bound to fail due to the supposed high rate of HGT, poor sequence conservation, and gene losses affecting these genes. However, some recent studies suggest that this may indeed not be the case and that the study of the evolution of operational genes is a feasible task (Pereto et al. 2004; Raymond et al. 2004; Bapteste et al. 2005; Desmond et al. 2007). Moreover, the growing availability of genomic data opens up a new era where accurate phylogenomics analyses become possible. Our study of O2Red is a good example and allowed bringing an updated insight on a long-standing debate around the origin and evolution of enzymes involved in a fundamental energy conversion process such as aerobic respiration.

Supplementary Material

Supplementary Materials 1–6 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We wish to thank Joyce Hebbo, Sylvain Buffet, and Marielle Bedotto for contributing to preliminary analyses.

This work was supported by an Action Thématique et Incitative sur Programme of the French Centre National de la Recherche Scientifique, section “Environnement et Développement Durable.”

Literature Cited


2CP-C suggests an aerobic common ancestor to the delta-proteobacteria. PLoS ONE. 3:e2103.


Andrew Roger, Associate Editor

Accepted October 22, 2008