Photosynthesis Research **33**: 137–146, 1992.

Minireview

Evolution of proton pumping ATPases: Rooting the tree of life

Johann Peter Gogarten[1] & Lincoln Taiz[2]
[1]*Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA*;
[2]*Biology Department, University of California, Santa Cruz, CA 95064, USA*

Received 1 September 1991; accepted in revised form 12 March 1992

Key words: ATPase, progenote, origin of life, archaebacteria, membrane transport

Abstract

Proton pumping ATPases are found in all groups of present day organisms. The F-ATPases of eubacteria, mitochondria and chloroplasts also function as ATP synthases, i.e., they catalyze the final step that transforms the energy available from reduction/oxidation reactions (e.g., in photosynthesis) into ATP, the usual energy currency of modern cells. The primary structure of these ATPases/ATP synthases was found to be much more conserved between different groups of bacteria than other parts of the photosynthetic machinery, e.g., reaction center proteins and redox carrier complexes.

These F-ATPases and the vacuolar type ATPase, which is found on many of the endomembranes of eukaryotic cells, were shown to be homologous to each other; i.e., these two groups of ATPases evolved from the same enzyme present in the common ancestor. (The term eubacteria is used here to denote the phylogenetic group containing all bacteria except the archaebacteria.) Sequences obtained for the plasmamembrane ATPase of various archaebacteria revealed that this ATPase is much more similar to the eukaryotic than to the eubacterial counterpart. The eukaryotic cell of higher organisms evolved from a symbiosis between eubacteria (that evolved into mitochondria and chloroplasts) and a host organism. Using the vacuolar type ATPase as a molecular marker for the cytoplasmic component of the eukaryotic cell reveals that this host organism was a close relative of the archaebacteria.

A unique feature of the evolution of the ATPases is the presence of a non-catalytic subunit that is paralogous to the catalytic subunit, i.e., the two types of subunits evolved from a common ancestral gene. Since the gene duplication that gave rise to these two types of subunits had already occurred in the last common ancestor of all living organisms, this non-catalytic subunit can be used to root the tree of life by means of an outgroup; that is, the location of the last common ancestor of the major domains of living organisms (archaebacteria, eubacteria and eukaryotes) can be located in the tree of life without assuming constant or equal rates of change in the different branches.

A correlation between structure and function of ATPases has been established for present day organisms. Implications resulting from this correlation for biochemical pathways, especially photosynthesis, that were operative in the last common ancestor and preceding life forms are discussed.

The concept of chemi-osmotic coupling

Photosynthetic energy transduction requires the formation of an energy-rich compound that can be utilized by other biochemical reactions. Peter Mitchell (1966) and many others since then have shown that ion gradients across biological membranes can act as energy-rich metabolic intermediate. Most importantly, the electrochemical (i.e., electrical plus concentration) gradient for

protons (H^+) links ATP synthesis to the electron transport chains operating in respiration and photosynthesis. In these electron transport chains the downhill (with regard to the free energy) flow of electrons is linked to the uphill transport of protons across the membrane. The electrochemical gradient for protons thus generated is then utilized by a multi-subunit enzyme called ATP synthase to form ATP from ADP and phosphate, while allowing the downhill flow of protons. The free energy released during the passage of the electrons through the electron transport chain is stored in the *trans*membrane gradient for protons; in turn, the free energy released during the back flow of protons from one site of the membrane to the other is stored in the ATP molecule. Many cellular processes requiring energization use this ATP; only a few, like the movement of the bacterial flagellum, utilize the proton gradient directly. The primary photosynthetic reactions in halobacterial photosynthesis are very different from the oxygenic and non-oxygenic photosynthesis found in eubacteria; however, the final steps in both are the same, i.e., formation of a *trans*membrane gradient for protons and utilization of this gradient by an ATP synthase.

Endosymbiosis and the origin of higher eukaryotes

This mode of ATP synthesis utilizing transmembrane proton gradients is found in all major bacterial groups and in eukaryotes. In the latter, it occurs exclusively in mitochondria and chloroplasts. These organelles were shown to have developed from bacterial endosymbionts that were taken up into the cytoplasm of the eukaryotic ancestor. The eubacterial origin of these cell organelles has been verified by many molecular and physiological characters that are shared between these organelles and 'free' living eubacteria, but that are not shared by the remainder of the eukaryotic cytoplasm and the nucleus (e.g., membrane lipids, ribosome type, etc.).

Present day eukaryotes depend on chloroplasts and/or mitochondria to synthesize ATP from ion gradients; outside these organelles, ATP is synthesized only via substrate level phosphorylation (e.g., pyruvate kinase). Thus, at first sight, ATP synthases do not appear to be a promising marker molecule to trace the origin of the nuclear/cytoplasmic component of the eukaryotes.

ATP synthases–ATPases

However, many of the membranes inside and surrounding the eukaryotic cell are energized by ion gradients; for example: H^+ gradients in lysosomes, vacuoles, storage vesicles, and the plasmamembrane surrounding the cytoplasm of plants and fungi, Na^+ gradients in the case of the plasmamembrane of animal cells; Ca^{2+} in the case of the sarcoplasmic reticulum. These ion gradients are primarily formed by ion transport ATPases. These enzymes use the energy stored in the ATP molecule to transport ions against their electrochemical gradient across the membrane. In many bacteria, under conditions which do not allow energization via electron transport (usually because of the absence of an electron acceptor such as O_2), the ATP synthases can also function as an ATP-driven proton pump; i.e., they catalyze the reverse reaction and provide the energization of the plasmamembrane which is needed for movement and transport of other substances. The ATP synthases in mitochondria and chloroplast can also function as H^+ pump in vitro; although, in situ this is prevented by regulatory mechanisms. As ATP synthases also hydrolyze ATP under some conditions, the ATP synthases found in mitochondria, chloroplasts and eubacteria are also termed coupling factor ATPases, cF_0cF_1, or F-ATPases.

Considering the reversibility of the ATP synthase reaction, it was only a minor surprise, when it was shown that one group of proton pumping ATPases that occurs in eukaryotes was homologous to the F-ATPases (Zimniak et al. 1988). This type of ATPase occurs on many of the endomembranes of the eukaryotic cell. Since it was first characterized on vacuolar membranes, it is called vacuolar type ATPase or V-ATPase. However, this ATPase type is also found in the membranes surrounding lysosomes, chromaffin granules, other storage vesicles, protein sorting organelles known under a variety of names

(endosomes, compartment of uncoupling of receptor and ligand or CURL, *trans* Golgi network or TGN, and the *trans* Golgi cisternae itself), and in clathrin coated vesicles. In the latter only the presence of V-ATPase subunits, but not their function, has been demonstrated (Forgac 1989). Categorizing the V-ATPases as the proton pump of the endomembrane system is appealing; however, there is a growing list of cases where V-type ATPases were found as functioning enzymes also on the plasmamembrane in specialized cells of vertebrates (e.g., osteoclasts, specialized cells in kidney and bladder epithelia, Blair et al. 1989, Gluck and Cadwell 1987).

V- and F-type ATPases

Both V- and F-type ATPases are multi subunit enzymes that in vitro can be dissociated into two parts: a water soluble F_1 or V_1 portion that contains the ATP binding sites, and the F_0 or V_0 portion that is embedded in the membrane. The V_1 and F_1 portions are peripheral protein complexes comprised of three copies each of a catalytic and a regulatory subunit. In addition, single copies of three minor subunits form the connection between the catalytic complex and the membrane. In negatively stained preparations of F- and V-ATPases the catalytic complex resembles a ball attached to the membrane by a narrow stalk (Bowman et al. 1989, Taiz and Taiz 1991). During the catalytic cycle the ATP molecule binds to the so-called catalytic subunit (part of V_1 or F_1), the bound ATP is hydrolyzed directly by a water molecule, the products (ADP and phosphate) are released into the medium and the catalytic subunit is ready to bind the next ATP molecule. At substoichiometric levels of substrate, catalysis can occur at a slow rate at a single catalytic site (uni-site catalysis). At higher substrate concentrations the rate is much higher because all three catalytic sites are acting cooperatively (multisite catalysis). This cycle of ATP hydrolysis is linked to conformational changes that store the free energy of ATP hydrolysis; these conformational changes are propagated through the enzyme to the F_0/V_0 portions of the enzymes, where they are utilized to transport a proton from one site of the membrane to the other. The proton that is to be transported binds to a subunit which has been termed proteolipid, because it is a protein that has chemical properties similar to a lipid. The crucial H^+-binding site is located in the middle of a membrane spanning alpha helix. A proton coming from one site of the membrane has access through a proton channel to this site and after binding and a small relative movement of two membrane spanning helixes, this site can release the proton to the other site of the membrane. At least in the case of F-ATPases there are no steps involved that lead to a large dissipation of free energy. If a strong proton gradient is provided by other means, the whole process can run in reverse and synthesize ATP from ADP and phosphate.

In Fig. 1 the proteolipids are depicted as dark membrane imbedded subunits. In F-ATPases there are 9–12 proteolipids present per functional enzyme, in V-ATPases there are only 6 proteolipids/V-ATPase (Forgac 1989); however, in the latter the proteolipid has twice the size as in the F-ATPases. Although the proteolipids of V-ATPases have 4 membrane spanning helixes and can be thought of as a head to tail fusion of the two smaller F-ATPase type proteolipids, they have only one binding site for the proton that is to be transported.

Orthologous vs. paralogous subunits

The water soluble parts of V- and F-ATPases contain three catalytic subunits each and three non-catalytic subunits (cf. Fig. 1). In addition, both ATPase types have several other subunits that are not recognizably conserved between the F- and V-ATPases. The non-catalytic subunits are very similar to the catalytic subunits; they also contain an ATP binding site, but the ATP that binds to the non-catalytic subunit is not hydrolyzed. The sequence similarity between catalytic and non-catalytic subunits is so high that it can be only explained by a common evolutionary origin of these subunits; i.e., in the past there has been an ATPase which had only one type of subunit. After a gene duplication had occurred, the two genes underwent separate changes and diverged from each other. One of

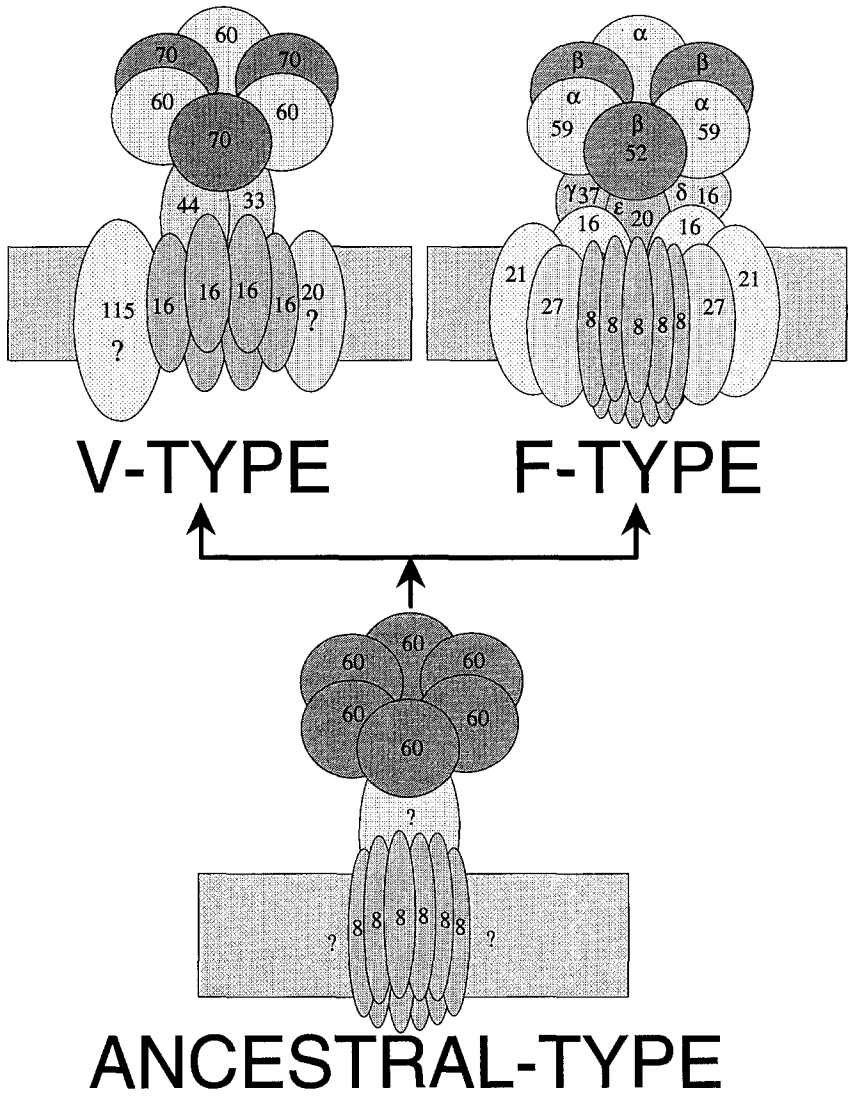

Fig. 1. Comparison of the subunit composition of V- and F-type ATPases. Numbers denote the approximate molecular weight in kDalton. The major subunits of the head groups are symbolized as spheres. The catalytic subunits are depicted as dark colored (70 and 52 kDa, respectively), the non-catalytic subunit are depicted with a lighter coloring (60 and 59 kDa, respectively). The catalytic (dark) and the non-catalytic subunits (light) are paralogous, i.e., they evolved from the same ancestral subunit. The proteolipids are depicted as dark ellipsoids that span the membrane plane. The depicted ancestral type corresponds to the ATPase before the evolution of the non-catalytic subunit (points 1 and 2 in Fig. 4). The depicted structure follows from the assumption of an overall similar structure and is the one most parsimonious with regard to gene duplication events (cf. Fig. 4). Only subunits that have been homologized between V- and F-ATPases have been depicted for the ancestral type. For further explanation and discussion see text.

the encoded subunits lost its catalytic activity and became the non-catalytic subunit. In evolutionary terms, catalytic subunits from different organisms are orthologous to each other; if in the last common ancestor there was only one catalytic subunit encoding gene, then the evolution of this gene reflects the evolutionary history of the organisms. In contrast, the catalytic and the non-catalytic subunits, although they have a common ancestral origin, are not orthologous but paralogous. Comparison between paralogous subunits picked from different species does not

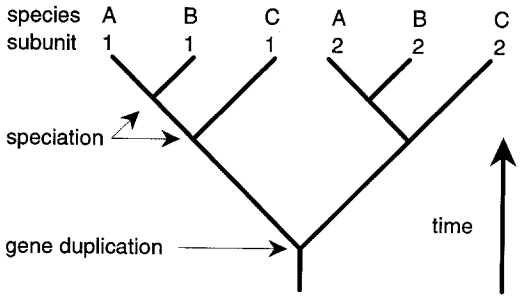

Fig. 2. Schematic diagram depicting the evolution of orthologous (subunits 1 vs. 1 or 2 vs. 2) and paralogous subunits (subunits 1 vs. 2). The evolution of orthologous subunits traces the evolution of the species (at least over the time scales considered in this paper); the evolution of paralogous subunits reflects the event of the gene duplication, which in the depicted case, as well as in case of the paralogous ATPase subunits precedes the speciation. For further discussion see text.

reflect the evolution of the species, but points to the much earlier event of the gene duplication that gave rise to the two subunit types (see Fig. 2). Usually paralogies complicate evolutionary studies. If they go unnoticed, i.e., they are taken as orthologies, they can lead to drastically wrong conclusions. Considering Fig. 2 as an example and analyzing only subunit 1 from species A and C and subunit 2 from species B, if one mistakes the paralogy for orthology, one would conclude that A and C are more closely related among each other than to species B. However, realizing the paralogy, one can use the paralogous subunit to pinpoint the gene duplication event on the phylogenetic tree.

The archaebacterial H^+-ATPase/ATP synthase

The presence of an F-ATPase homolog in the cytoplasm of eukaryotes, the presence of a paralogous subunit present in the same enzyme, and the fact that the gene duplication giving rise to the paralogous subunits apparently had already occurred in the enzyme ancestral to both V- and F-ATPases, point to the usefulness of the ATPase subunits as universal molecular marker. However, the full potential of ATPases as molecular marker molecules became apparent to us only after sequences encoding archaebacterial ATPase subunits became known. Archaebacteria were defined as an independent Urkingdom mainly due to the work of Carl Woese and collaborators on the 16S ribosomal RNA (Woese and Fox 1977a). This second prokaryotic domain was supposed to be equally 'distant' to eubacteria and eukaryotes. The grouping of the archaebacteria into a separate domain is corroborated by many other characters such as sensitivity to antibiotics, occurrence of certain membrane lipids and unique enzymatic cofactors. However, a recent re-analyses of the data by Jim Lake (1988) opened up a different interpretation of the similarity among the archaebacteria: members of this group of bacteria are all slowly evolving organisms; therefore, they remained more similar to the ancestral prokaryote. However, according to Lake's analysis, one group of archaebacteria, the eocytes, has a more recent ancestor with the eukaryotes, whereas, the rest of the archaebacteria are more closely related to the other bacteria. The catalytic and non-catalytic subunits of an ATPase from the archaebacterium *Sulfolobus acidocaldarius* were sequenced (Denda et al. 1988a,b) at the same time the first V-ATPase sequences were obtained (Zimniak et al. 1988, Bowman et al. 1989, Manolson et al. 1989). To the mutual surprise of the researchers involved, these sequences of the *Sulfolobus* ATPase catalytic and non-catalytic subunits turned out to be nearly identical (50–70% identity) to the eukaryotic V-ATPase, and about as distantly related (25%) to the eubacterial F-ATPase as the eukaryotic enzyme (Gogarten et al. 1989).

The proteolipid from the *Sulfolobus* ATPase, which is the only other subunit that has been shown to be homologous between the F and V-ATPases (Nelson and Nelson 1989), is intermediate between the eukaryotic and the eubacterial enzymes. The molecular weight of the encoded proteolipid is 12 kDa. According to hydrophibicity it consists of three membrane spanning regions (Denda et al. 1989). The two helices at the carboxyterminal end of the *Sulfolobus* proteolipid show slight similarities to the two helices of the proteolipid from eubacteria and about the same level of similarities to the front and the back half of the proteolipid of the eukaryotic V-ATPase. According to results obtained by Lübben and Schäfer (pers. communi-

cation by Matthias Lübben) the encoded amino terminal membrane-spanning helix of the proteolipid is not part of the membrane-imbedded proteolipid, but likely to constitute a signal peptide directing the membrane anchoring portion to the target membrane during or after synthesis.

Evolution of the catalytic subunit

Given the nucleotide sequences that encode the catalytic and non-catalytic subunits, one can use the sequence information to reconstruct the evolution of the subunits. An example is depicted in Fig. 3. Many different algorithms can be used for this phylogenetic analyses; all of these algorithms are based on different assumptions, and sometimes they give different results. However, in case of the ATPase subunits, all the algorithms tested so far (including Lake's rate invariant evolutionary parsimony algorithm) result in gene phylogenies with the following features (Gogarten et al. 1989a,b, Iwabe et al. 1989): (1) If analyzed separately the trees for the catalytic and the non-catalytic subunits have the same topology. (2) All F-ATPase catalytic subunits form one monophyletic group (i.e., this grouping has an ancestor that is only ancestor to this group). (3) All archaebacterial catalytic subunits are located on a branch that leads from the juncture to the non-catalytic subunit to the eukaryotes, no archaebacterial subunits are found on the branch leading to the eubacterial subunits (cf. Fig. 3).

Rooting the tree of life

There are different ways to root a phylogenetic tree; however, without assuming a molecular clock ticking at the same rate in the different branches, the algorithms available to analyze molecular data only calculate unrooted trees. Clearly the assumption of constant and equal rates of change over the huge time span covered in Fig. 3 is not warranted. A rate-independent way of rooting that is often employed in phylogenetic analyses is the use of an outgroup, i.e., one includes a species in the analysis that is only distantly related to the organisms under consid-

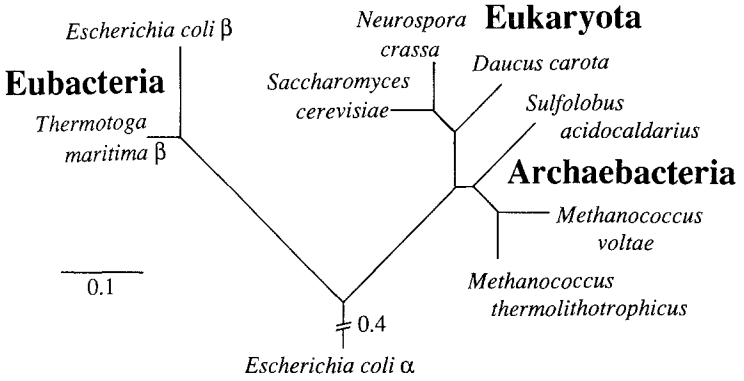

Fig. 3. Phylogenetic tree based on the catalytic subunits of V- and F-ATPases. This tree was calculated using a global alignment of the sufficiently conserved regions from DNA sequences encoding the catalytic subunits from the indicated organisms. In order to locate the root of the tree, the non-catalytic subunit of the *E. coli* F-ATPase is used as an outgroup. The topology and the branch lengths were calculated using Felsenstein's maximum likelihood method (1981). Parameters for the algorithm, sequences and their alignment were as described in Gogarten et al. (1989b); branches are scaled in terms of the probability for change of the first base of the codon. All branches were positive at the 1% significance level. In 100% of bootstrapped samples analyzed with parsimony the *M. thermolithotrophicus* and the *Sulfolobus acidocaldarius* sequences grouped together with the eukaryotic sequences. In 96% of the bootstrapped samples *Methanococcus thermolithotrophicus* and the *Sulfolobus acidocaldarius* formed a monophyletic group. When the sequence encoding the *Methanosarcina barkeri* plasmalemma ATPase (Inatomi et al. 1989) is included in the alignment, in 98 out of 100 bootstrapped samples the *Methanosarcina* sequence constitutes a separate branch that separates from the branch leading to the eukaryotes after the bifurcation leading to the thermophilic archaebacteria (*Sulfolobus acidocaldarius* and *Methanococcus thermolithotrophicus*). Thus, not all archaebacteria but only the mesophilic appear to be the sister group of the eukaryotes.

eration; for example, when studying mammals one could use a bird or crocodile as an outgroup.

If one considers the tree of life containing all groups of living species, obviously there is no species available that can be used as an outgroup. However, in case of duplicated genes that underwent the duplication before the last cominon ancestor split off into different lines, one can use the duplicated gene as an outgroup that from the time of the duplication underwent a separate development. For example in Fig. 2 one can use subunit 1 as a marker for the species and use subunit 2 as an outgroup. In Fig. 3 the non-catalytic subunit of the *E. coli* F-ATPase is used as an outgroup to root the tree calculated for the catalytic subunits. Clearly, the archaebacteria form the sister group to the eukaryotes, whereas the eubacteria constitute a second independent lineage coming from the root.

Are the archaebacteria monophyletic?

All archaebacterial subunits known so far are located on the branch that leads from the root to the eukaryotes (see Fig. 3); however, not in all cases do the archaebacteria form a monophyletic group, i.e., they do not branch off from the main branch in a single node. In the case of the *Methanosarcina barkeri* catalytic subunit (sequence from Inatomi et al. 1989) we found that with some alignments and algorithms, this species branched off later, i.e., closer to the eukaryotes, than *Sulfolobus* and the *Methanococci*. A similar result was obtained for the halobacterial catalytic subunit (Ihara and Mukohata 1991). Thus, contrary to the eocyte tree proposal (Lake 1988), the ATPase data suggest that the mesophilic (moderate temperature growth optimum) archaebacteria are more closely related to the eukaryotes than are the thermophilic archaebacteria. However, as the obtained significance levels for this topology are low, at the present time the question as to whether the archaebacteria are monophyletic or not has to remain open. In contrast, the finding that all the archaebacteria diverged from the branch that leads from the root to the eukaryotes has been obtained reproducibly and with high significance levels (Gogarten et al. 1989a,b).

Supporting evidence for the proposed phylogeny

The location of the root between eubacteria on one side and the archaebacteria and eukaryotes on the other is further supported by the insertion of a so-called non-homologous region (Zimniak et al. 1988) into the catalytic subunit of the archaebacterial and eukaryotic V-ATPase (Gogarten et al. 1989a,b). This non-homologous region is absent in the eubacterial catalytic subunit and in the outgroup (i.e., all non-catalytic subunits, cf. Figs. 3 and 4). Thus, the non-homologous region can be used as a character defining the group comprising archaebacteria and eukaryotes as a monophyletic grouping in the rooted tree of life.

To base a phylogenetic proposal on a single enzyme (although on two separate genes) opens up the possibility of lateral gene transfer as a possible source of error. It was therefore of great value when Iwabe et al. (1989) analyzing other genes that supposedly had undergone an ancient gene duplication (elongation factors, tRNAs and dehydrogenases) deduced the same location of the root based on these different genes. The phylogenies based on 16S rRNA have topologies that are compatible with the one depicted in Figs. 3 and 4, only the branch lengths are different. However, as the rRNA genes did not undergo an ancient gene duplication that lead to two separately evolving genes, there is no outgroup available, and assuming equal rates of evolution is obviously impossible in case of the rRNA genes. (In an unrooted tree based on 16S rRNA, the archaebacteria are located on much shorter branches in the center of the tree. In this case there exists no location of the root that would result in approximately equal branch lengths from this universal root to the different present day species.)

H^+/ATP and proteolipid/catalytic subunit ratios

A theory relating H^+/ATP ratios to gene duplications was put forward by Cross and Taiz (1990). Based on sequence data and comparison to the better studied F-ATPases, it was suggested that the number of protons pumped per ATP is equal to the number of proteolipids divided by

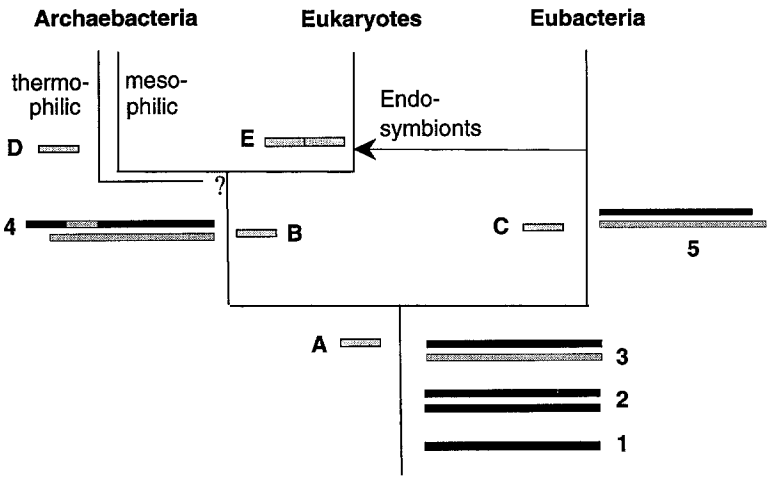

Fig. 4. Phylogenetic relations among the three kingdoms as derived from the analysis of ATPase subunits. The evolution of the genes encoding the proteolipid, catalytic and noncatalytic subunits is depicted schematically at the side of the tree. (1) The catalytic complex consists of a single subunit type encoded by a single gene. The dark color indicates that this subunit is catalytic. (2) A gene duplication occurs, leading to two catalytic subunits. (3) One subunit loses its catalytic activity and becomes a regulatory or non-catalytic subunit. The evolution from state 1 to 3 already had occurred in the last common ancestor. (4) In the line leading to archaebacteria and eukaryotes, the catalytic subunit gene gains a large region near the amino terminal end; this region has no homology to any of the other subunits. (5) In the line leading to the eubacteria and the endosymbionts, the non-catalytic subunit increases in size relative to the catalytic subunit. The proteolipid of the ATPase in the last common ancestor (A) like the one of eubacteria (B) and archaebacteria (D) has a molecular weight of 8 kDa and contains two membrane spanning alpha helices. This type is the one most likely also to be present in the last common ancestor of eukaryotes and archaebacteria (C). During the evolution of the eukaryotes a duplication of the gene encoding the proteolipid and fusion of the resulting genes occurred (E).

the number of catalytic subunits. In accordance with this the H^+/ATP ratio in F-ATPases is 3–4, whereas the ratio in V-ATPases is $2H^+$/ATP (Bennett and Spanswick 1984). In order to function as an effective proton pump, a smaller H^+/ATP ratio is desirable, whereas a larger ratio would favor ATP synthesis. The authors conclude that a proteolipid/catalytic subunit ratio of 2 indicates the exclusive functioning as a proton pump, whereas a proteolipid/catalytic subunit ratio of 3–4 indicates a predominant functioning as an ATP synthase. Present day ATPases fit this pattern exactly, if one assumes an overall similar structure for the ancestral enzymes the following two conclusions can be drawn from this theorem:

1. The last common ancestor already had efficient membrane linked ATP synthesis. The ATPase present in the last common ancestor at the time of the split into the two lineages leading to eubacteria on the one hand and archaebacteria and eukaryotes on the other (point 3 in Fig. 4) already had catalytic and non-catalytic subunits and the proteolipid was of the smaller type. From this, a proteolipid/catalytic subunit ratio and a H^+/ATP ratio of 3–4 and a predominant functioning as ATP synthase can be deduced. A correlary to this conclusion is that the last common ancestor had to have other means to generate the proton gradient in the first place, i.e., some electron transport chains were likely to be present that generated the proton gradient that then could be used for ATP synthesis.
2. H^+–ATPases first evolved as proton pumps, not as ATP synthases. If one extrapolates the subunit structure even further into the past to the time when the non-catalytic subunit had not yet evolved (points 1 and 2 in Fig. 4), a subunit composition as depicted in Fig. 1 for the ancestral ATPase results. The proteolipid/catalytic subunit ratio of 2 could be interpreted to indicate that the ATPases evolved first as proton pumping ATPases that used metabolic energy to maintain a more alkaline intracellular pH, and that the use as ATP synthase evolved only later when other means

to generate a transmembrane proton gradient were available. The ATP synthases that couple modern photosynthesis (including archaebacteria (halobacteria) and oxygenic and non-oxygenic photosynthesis in eubacteria) evolved first as proton pumping ATPases. If photosynthetic processes were functioning at these early stages of evolution (points 1 and 2 in Fig. 4), they did not involve the generation of transmembrane ionic gradients whose energy was conserved in the synthesis of ATP.

The last common ancestor

Very often the term progenote is used to denote the last common ancestor of all currently existing life forms. However, as originally defined (Woese and Fox 1977b) the progenote characterizes a hypothetical primitive stage of development that existed before a defined relation between genotype and phenotype of an organism had been established. Woese and Fox (1977b) had this primitive entity suggested to be the last common ancestor: 'It is at this progenotic state, not the procaryote stage, that the line of descent leading to the eucaryotic cytoplasm diverged from the bacterial lines of descent.'

The picture of the common ancestor that emerges from the study of the molecular evolution of conserved genes is very different from this primitive progenote. The last common ancestor was a highly developed organism that used DNA for the storage of genetic information; it had a sophisticated protein synthesis machinery not much different from today's organisms that used ribosomes, tRNAs, a variety of elongation factors etc.; in addition, it had membranes that separated this organism from the environment, and these membranes were already energized by proton gradients. If we follow the argument sketched above relating subunit composition to function, we can conclude that this last common ancestor already used its ATPases to synthesize ATP and that other means (e.g., photo- or chemosynthetic electron transport chains or other light driven transport reactions) were available to generate the proton gradients necessary to drive ATP synthesis.

The last common ancestor thus appears to have been a sophisticated organism already well advanced in the evolution of life. At first sight this is bad news for attempts to trace the origins and early evolution of life by means of phylogenetic analyses. As there is only one well advanced survivor of the earlier stages, we cannot use the comparison between organisms, and the determination of primitive and derived characters to elucidate the earlier stages of life and the forces that directed their evolution. However, the one surviving organism that subsequently gave rise to all the other life forms contained already several hundreds to thousands of different genes. As exemplified by the evolution of ATPase subunits, the study of the evolution of the genes can reach back further into the past than the comparison between organisms. The only surviving targets for phylogenetic comparison of the early evolution are the genes; their evolution is at least to some degree conserved in the molecular record. Although this record is not perfect but distorted by various selective pressures and scrambled by multiple hits, advances in the study of molecular evolution are likely to further contribute also to the evolution that led from the origin of life to the last common ancestor, especially if the secondary, tertiary and quaternary structure of the gene products, in addition to their linear sequences, are taken into account.

References

Bennett AB and Spanswick RM (1984) H^+-ATPase activity from storage tissue of *Beta vulgaris*; II. H^+/ATP stoichiometry of an anion-sensitive H^+-ATPase. Plant Physiol 74: 545–548

Blair HC, Teichelbaum SL, Ghiselli R and Gluck S (1989) Osteoclastic bone resorption by a polarized vacuolar proton pump. Science 245: 855–857

Bowman BJ, Allen R, Wechser MA and Bowman EJ (1988) Isolation of genes encoding the Neurospora vacuolar ATPase: Analysis of vma-2 encoding the 57 kDa polypeptide and comparison to vma-1. J Biol Chem 263: 14002–14007

Bowman BJ, Dschida WJ, Harris T and Bowman EJ (1989) The vacuolar ATPase of *Neurospora crassa* contains an F_1-like structure. J Biol Chem 264: 15606–15612

Cross RL and Taiz L (1990) Gene duplication as a means for altering H^+/ATP ratios during the evolution. FEBS Lett 259: 227–229

Denda K, Konishi J, Oshima T, Date T and Yoshida M (1988a) Molecular cloning of the b subunit of a possible non-F_0F_1 type ATP synthase from the acidothermophilic archaebacterium *Sulfolobus acidocaldarius*. J Biol Chem 263: 17251–17254

Denda K, Konishi J, Oshima T, Date T and Yoshida M (1988b) The membran associated ATPase from *Sulfolobus acidocatdarius* is distantly related to F_1-ATPase as addressed from the primary structure of its A-subunit. J Biol Chem 263: 6012–6015

Denda K, Konishi J, Oshima T, Date T and Yoshida M (1989) A gene encoding the proteolipid subunit of *Sulfolobus acidocaldarius* ATPase complex. J Biol Chem 264: 7119–7121

Felsenstein J (1981) Evolutionary trees from DNA sequences: A maximum likelihood approch. J Mol Evol 17: 368–376

Forgac M (1989) Structure and function of vacuolar class of ATP driven proton pumps. Physiological Rev 69: 765–796

Gluck S and Cadwell J (1987) Immunoaffinity purification of vacuolar H^+ ATPase. J Biol Chem 262: 15780–15787

Gogarten JP, Kibak H, Taiz L, Bowman EJ, Bowman BJ, Manolson MF, Poole RJ, Date T, Oshima T, Konishi J, Denda K and Yoshida M (1989a) The evolution of the vacuolar H^+-ATPase: Implications for the origin of eukaryotes. Proc Natl Acad Sci USA 86: 6661–6665

Gogarten JP, Rausch T, Bernasconi P, Kibak H and Taiz L (1989b) Molecular evolution of H^+-ATPases. I. *Methanococcus* and *Sulfolobus* are monophyletic with respect to eukaryotes and eubacteria. Z Naturforsch 44c: 641–650

Ihara K and Mukohata Y (1991) The ATP Synthase of *Halobacterium salinarium* (*halobium*) is an archaebacterial type as revealed from the amino acid sequences of its two major subunits. Arch Biochem Biophys 286: 111–116

Inatomi K-I, Eya S, Maeda M and Futai M (1989) Amino acid sequence of the alpha and beta subunits of *Methanosarcina barkeri* ATPase deduced from cloned genes. Similarity to subunits of eukaryotic vacuolar and F_0F_1-ATPases. J Biol Chem 264: 10954–10959

Iwabe N, Kuma K-I, Hasegawa M, Osawa S and Miyata T (1989) Evolutionary relationships of archaebacteria, eubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci USA 86: 9355–9359

Lake JA (1988) Origin of the eukaryotic nucleus determined by rate invariant analysis of rRNA sequences. Nature 331: 184–186

Manolson MF, Ouellette BF, Filion M and Poole RJ (1988) cDNA sequence and homologies of the '57-kDa' nucleotide-binding subunit of the vacuolar ATPase from *Arabidopsis*. J Biol Chem 263: 17987–17994

Mitchell P (1966) Chemiosmotic Coupling in Oxidative and Photosynthetic Coupling. Glynn Research Ltd, Bodmin, UK

Nelson H and Nelson N (1989) The progenitor of ATP synthases was closely related to the current vacuolar H^+-ATPase. FEBS Lett 247: 147–153

Pedersen PL and Carafioli E (1987) Ion motive ATPases. Trends Biochem Sci 12: 146–150 and 186–189

Taiz SL and Taiz L (1991) Ultrastructural comparison of the vacuolar and mitochondrial H^+-ATPases of *Daucus carota*. Bot Acta 104: 117–121

Woese CR (1987) Bacterial evolution. Microbiol Rev 51: 221–271

Woese CR and Fox GE (1977a) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc Natl Acad Sci USA 74: 5088–5090

Woese CR and Fox GE (1977b) The concept of cellular evolution. J Mol Evol 10: 1–6

Zimniak L, Dittrich P, Gogarten JP, Kibak H and Taiz L (1988) The cDNA sequence of the 69 kDa subunit of the carrot vacuolar H^+-ATPase: Homology to the beta-chain of F_0F_1-ATPases. J Biol Chem 263: 9102–9112