Journal of Bioenergetics and Biomembranes, Vol. 24, No. 4, 1992

Evolution of Structure and Function of V-ATPases

Henrik Kibak,[1] Lincoln Taiz,[1] Thomas Starke,[2] Paul Bernasconi,[1,3] and Johann Peter Gogarten[2,4]

Received February 16, 1992; accepted March 16, 1992

Proton pumping ATPases/ATPsynthases are found in all groups of present-day organisms. The structure of V- and F-type ATPases/ATP synthases is very conserved throughout evolution. Sequence analysis shows that the V- and F-type ATPases evolved from the same enzyme already present in the last common ancestor of all known extant life forms. The catalytic and noncatalytic subunits found in the dissociable head groups of the V/F-type ATPases are paralogous subunits, i.e., these two types of subunits evolved from a common ancestral gene. The gene duplication giving rise to these two genes (i.e., encoding the catalytic and noncatalytic subunits) predates the time of the last common ancestor.

Mapping of gene duplication events that occurred in the evolution of the proteolipid, the noncatalytic and the catalytic subunits, onto the tree of life leads to a prediction for the likely subunit structure of the encoded ATPases. A correlation between structure and function of V/F-ATPases has been established for present-day organisms. Implications resulting from this correlation for the bioenergetics operative in proto-eukaryotes and in the last common ancestor are presented. The similarities of the V/F-ATPase subunits to an ATPase-like protein that was implicated to play a role in flagellar assembly are evaluated.

Different V-ATPase isoforms have been detected in some higher eukaryotes. These data are analyzed with respect to the possible function of the different isoforms (tissue specific, organelle specific) and with respect to the point in their evolution when these gene duplications giving rise to the isoforms had occurred, i.e., how far these isoforms are distributed.

KEY WORDS: ATPases; evolution; eukaryotes; endomembranes; archaebacteria; progenote; bioenergetics; flagella assembly; endosymbiont theory.

INTRODUCTION

Ion gradients play a central role in the bioenergetics of modern cells. Peter Mitchell (1976), and many others since, have shown that ion gradients, in particular the electrochemical gradient for protons, can be used as energy-rich metabolic intermediates. ATP synthases can utilize the proton gradients, for example those generated by the electron transport chains of respiration and photosynthesis, to catalyze the synthesis of ATP from ADP and inorganic phosphate while allowing the "downhill" flow of protons. *In vitro* and in many cases also *in vivo* (e.g., facultative anaerobic bacteria) these ATP synthases can also catalyze the inverse reaction, coupling the uphill flow of protons to the hydrolysis of ATP.

These ATPases, also termed coupling factor ATPases or F-ATPases, were shown to be homologous to the ATPases found on many of the endomembranes

[1]Biology Department, University of California, Santa Cruz, California 95064.
[2]Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269.
[3]Present address: Sandoz Crop Protection, Palo Alto, California 94304.
[4]To whom correspondence should be addressed.

0145-479X/92/0800-0415$06.50/0

of eukaryotic cells (Zimniak *et al.*, 1988). The latter have been termed vacuolar type ATPases or V-ATPases because they were first described as enzymes energizing the vacuolar membrane. This ATPase type is also found in the membranes surrounding lysosomes, chromaffin granules, other storage vesicles, protein sorting organelles, and in clathrin-coated vesicles. In addition, there is a growing list of cases where V-type ATPases were found as functioning enzymes also on the plasma membrane in specialized cells of vertebrates (e.g., osteoclasts, specialized cells in kidney and bladder epithelia; see Blair *et al.*, 1989; Gluck and Cadwell, 1987). In all of the cases studied the vacuolar type ATPase functions exclusively as a proton pumping ATPase, not as an ATP synthase. Its function is to energize the respective membranes. This provides energy for a multitude of secondary active transport systems, and it provides a compartment with an acidic interior. This internal acidity can be regarded as a second messenger that acts on the activity and affinity of enzymes and receptors. Acidic intracellular compartments were found throughout the eukaryotic domain, e.g., acidosomes in *Dictyostelium* (Nolta *et al.*, 1991) and acidic vacuole-like vesicles in *Giardia* (Lindmark, 1988; Feeley and Deyer, 1987); based on its 16S-like rRNA the latter organism is thought to represent one of the earliest branching lines of descent within the eukaryotes (Sogin *et al.*, 1989), whereas the former belongs to a group that has been suggested to be close to the root of the eukaryotic domain based on 5S rRNA analyses (Nanney *et al.*, 1991).

Nucleotide and amino acid sequences of the catalytic and noncatalytic subunits (subunit A and B) of archaebacterial coupling factor ATPases are much more similar to the eukaryotic vacuolar type ATPase (about 50% identical amino acids) than to the homologous F-ATPase subunits (α and β subunits, about 25% identity). These different degrees of similarity reflect the evolutionary relationships of these ATPases (see below); based on this "natural" system (i.e., based on most recent common ancestry), one could regard the archaebacterial ATPases as V-ATPases. However, with respect to function and, as will be discussed below, with respect to quarternary structure, the archaebacterial coupling factor ATPase appears more similar to F-type ATPases. In the following the group comprising the eubacterial *and* the archaebacterial coupling factor ATPases *and* the eukaryotic V-ATPases will be denoted V/F-ATPases, however, the term V-ATPase will be used restrictively to denote the eukaryotic endomembrane ATPase only.

EVOLUTION OF CATALYTIC AND NONCATALYTIC SUBUNITS

The V-, F-, and archaebacterial ATPases are homologs, and the gene duplication which gave rise to the catalytic and noncatalytic subunits had already occurred in the last common ancestor of eubacteria, archaebacteria, and eukaryotes (Gogarten *et al.*, 1989a). Using this ancient gene duplication to place the root in the tree of life (cf. Fig. 1) demonstrated that the archaebacteria (only *Sulfolobus* was utilized in the 1989 study) are more closely related to the eukaryotes than to the eubacteria, i.e., the root was placed between eubacteria on one side and eukaryotes and archaebacteria on the other. Since then, sequences of other archaebacterial and eukaryotic ATPase subunits have been obtained (Gogarten *et al.*, 1989b; Bernasconi *et al.*, 1990; Inatomi *et al.*, 1989; Ihara and Mukohata, 1991). When these sequences had been subjected to phylogenetic analyses, the results confirmed the above conclusion (Gogarten *et al.*, 1989b; Mukohata *et al.*, 1990; Linkkila and Gogarten, 1991). Analysis of other duplicated genes (t-RNA^{met}, dehydrogenases, elongation factors; see Iwabe *et al.*, 1989) agree with the conclusion obtained with the ATPases, namely, all archaebacteria branch off the line leading from the root to the eukaryotes.

The related question "Do all archaebacteria form a monophyletic group or are some archaebacteria more closely related to the eukaryotes than others" has still not been answered to the satisfaction of all workers (Lake, 1991). Most researchers favor a monophyletic grouping (e.g., Iwabe *et al.*, 1991), but the error margins, if analyzed at all, are comparatively large. Furthermore, it is difficult to assess bias introduced by the sequence alignment; most tree-constructing algorithms require a global alignment of the sequences. The tree depicted in Fig. 3 uses only pairwise sequence alignments calculated using the Needleman and Wunsch algorithm (1970); however, dispensing with a global alignment is correlated with a loss in resolution.

THE EARLY EVOLUTION OF THE PROTEOLIPID

The V/F-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

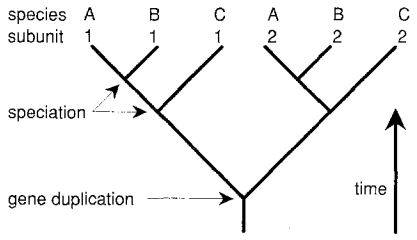

Fig. 1. Schematic diagram depicting the evolution of orthologous (subunits 1 vs. 1 or 2 vs. 2) and paralogous subunits (subunits 1 vs. 2). The study of the evolution of proteins, nucleic acids, and organisms is based on characters (molecular, morphological, or physiological). Two similar characters can be classified using the following three exclusive categories. (A) Analogy—the similarity is due to convergent evolution, meaning that this character has been obtained independently in two separate lines of descent; this can be due to chance or to similar evolutionary pressure caused by comparable function. (B) Orthology—the characters in the present-day organisms are derived from the same character in the species that is the most recent ancestor of both organisms under consideration. (C) Paralogy—this category is similar to the last one in that these characters are derived from an identical ancestral character; however, this ancestral character gave rise to two separate characters both of which were already present in the last ancestor of the two present-day organisms. An example of paralogous proteins is the catalytic (A in V-ATPases, β in F-ATPases) vs. noncatalytic (B and α, respectively) V/F-ATPase subunits. These two subunit types evolved from the same ancestral gene by means of a gene duplication event. Subsequently the two encoded subunits accumulated not only sequence differences but also aquired different functions. Catalytic and noncatalytic subunits are paralogous subunits. However, each of the three major subunits of the V/F ATPases (proteolipid, catalytic, and noncatalytic subunits) are thought to be orthologous, within each group; i.e., the catalytic subunits of ATPases from different organisms are thought to be orthologous to each other. In the case of the eukaryotic cell, the situation is complicated by the uptake of endosymbionts into the cell that evolved into the present-day mitochondria and plastids. These organelles and the genes that they contributed to the symbiosis are clearly eubacterial in character, whereas the cytoplasmic/nuclear component has features that are different from the eubacteria. The vacuolar ATPase is a distinctive marker for the evolution of the endomembrane system of the eukaryotic cell, and as such can contribute to our understanding of its origin and evolution. The evolution of orthologous subunits traces the evolution of the species (at least over the time scales considered); the evolution of paralogous subunits reflects the event of the gene duplication. There are different ways to root phylogenetic trees; 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. 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 organism under consideration; 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 the case of duplicated genes that underwent the duplication before the last common 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. If the subunits of type 1 are considered as markers for the organismal evolution, subunits of type 2 can be used as an outgroup, provided the gene duplication preceded the speciation. The tree gives a schematic representation of the evolution of the catalytic and noncatalytic V/F-ATPase subunits, if one sets A: Archaebacteria, B: Eukaryotes, C: Eubacteria, subunit 1: catalytic subunit, 2; noncatalytic subunit.

membrane. One scenario for the catalytic mechanism is the following: the F/V_1-portion translates a change in chemical energy to a mechanical conformation change; this conformational change is propagated to the V/F_0 portion, where this change is used to move a proton from one side of the membrane to the other. One subunit involved in this proton translocation is the proteolipid. In F-ATPases this proteolipid consists of two membrane-spanning helices; one of these (usually the one closer to the carboxy terminal) contains an amino acid with a carboxylic acid-containing side chain (Asp-61 in *E. coli*). This carboxylic acid group is required for activity, and it is thought to undergo protonation and deprotonation during the catalytic cycle (Fillingame, 1990).

The proteolipid of the eukaryotic V-ATPase consists of four membrane-spanning helices (Mandel *et al.*, 1988; Nelson and Nelson, 1989; Birman *et al.*, 1990; Hanada *et al.*, 1991; Lai *et al.*, 1991). Searching a DNA data bank revealed that the first two and the last two membrane-spanning helices were each similar to the F-ATPase proteolipid (Mandel *et al.*, 1988). Although the degree of similarity is low (many hydrophobic amino acids can substitute for each other without leading to a loss of function), at least some of the pairwise comparisons between the F-proteolipids and the carboxy terminal pair and the amino terminal pair result in significant z values (= distance in terms of standard deviation between the alignment score calculated for the actual sequences and the mean of the alignment scores obtained for randomized sequences). For example, the first two helices from the *Avena* proteolipid (Lai *et al.*, 1991) give a z value of 3.2 when compared to the proteolipid of the spinach chloroplast ATP synthese; the second pair of membrane-spanning helices gives z values of 4 when compared to the yeast mitochondrial ATP synthase proteolipid, and 3.2 when compared to the proteolipid of the ATPase from the thermophilic eubacterium PS3 (using the mutation matrix of Dayhoff *et al.*, 1978,

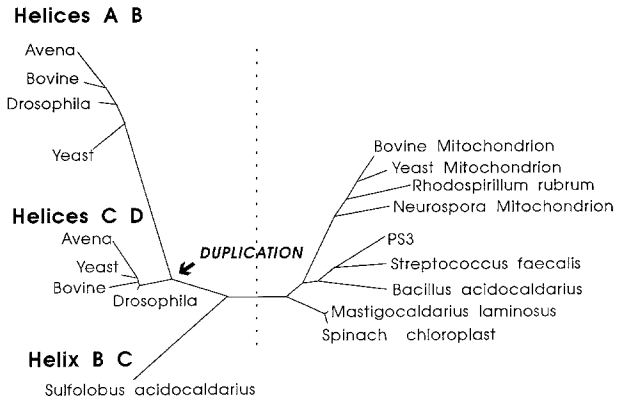

Fig. 2. Evolution of the proteolipids of V/F ATPases depicted as an unrooted tree. The two amino terminal and the two carboxy terminal membrane-spanning helices were treated as separate entities. The duplication event giving rise to these two similar halves is indicated. The tree was calculated using only pairwise comparison between amino acid sequences. Optimum alignment scores were calculated using Needleman and Wunsch's algorithm (1970) and Dayhoff's mutation matrix (Dayhoff *et al.*, 1978; positive scores are also awarded for conservative and often occurring amino acid exchanges). The distances between two sequences were calculated using Feng and Doolittle's (1987) formula; however, the alignment scores for the randomized sequences were actually calculated (mean of more than 10 randomizations each) instead of estimated. Topology and branch lengths of the depicted tree represent the minimum-length tree calculated using the algorithm of Fitch and Margoliash (1967) as implemented by Felsenstein (1988). Trees with the same topology (except for a minor rearrangement within the eubacterial domain) were obtained using the neighbor-joining method as implemented by Felsenstein (1988). The sequences used were retrieved from GenBank or were taken from Hoppe and Sebald (1984). For further discussion, see text.

and 50 randomizations). Only the last membrane-spanning helix of the V-ATPase proteolipid contains a carboxylic acid residue thought to participate in the proton passage through the membrane (Glu 139; Mandel *et al.*, 1988).

For the tree depicting the evolution of the proteolipid subunit (Fig. 2), the first two (amino terminal) and the last two (carboxy terminal) membrane-spanning helices were treated as separate entries. As expected for an evolution that does not involve the repeated lateral transfer of genes, the two halves of the V-ATPase proteolipids are more closely related to each other than to the bacterial type proteolipids. The gene duplication event that gave rise to the two similar portions of the proteolipid is indicated in Fig. 2. In accordance with the finding that the carboxy terminal helices of the V-ATPase proteolipid are more intimately involved with proton translocation (Mandel *et al.*, 1988), the two amino terminal helices have changed much faster since the duplication event occurred. This might be reflective of the changed and probably relaxed selection pressure acting on this portion of the proteolipid. Note that in the depicted tree the topology of the terminal branches is influenced equally by all other entries (minimum overall lengths). The branching orders at the periphery of one part of the tree can be influenced by convergences occurring in a completely different part of the tree. If a better resolution of peripheral branches is desired, it is preferable to consider only a subset of more closely related species.

In contrast to the V/F_1 subunits, the proteolipid of *Sulfolobus* appears to be intermediate between the eukaryotic proteolipid and the eubacterial one. The sequence as published by Denda *et al.* (1989) comprises three membrane-spanning helices; however, the first one appears to constitute a signal sequence that is absent in the assembled ATPase (Schäfer *et al.*, 1990). For the tree depicted in Fig. 2 the membrane-spanning helix corresponding to this putative signal sequence was not included; however, the same location for the *Sulfolobus* proteolipid results, if all three membrane-spanning helices are considered.

No sequence (or group of sequences) that is part of the tree depicted in Fig. 2 can be used as an obvious outgroup, and clearly the assumption of equal rates of change in all branches is not justified. Searches of GenBank (Release 69, using the algorithm of Pearson and Lipman, 1988) suggested some candidates that might have evolved from the same ancestral gene as the ATPase proteolipid following an ancient gene duplication. In particular, similarities to parts of the yeast mitochondrial cytochrome *c* oxidase and to a portion of the FlaA gene of *Campylobacter coli* are intriguing. (58.8% identity in a 109 bp overlap when compared to the spinach chloroplast proteolipid; however, this DNA encodes the carboxy terminal end of the FlaA gene product, including the stop codon.) These putatively paralogous genes might be useful as an outgroup (compare Fig. 1); however, the stretches of similarity are short and the *z* values for pairwise comparison are not significant. Note that the topology of the proteolipid tree (Fig. 2) is compatible with the trees depicted in Figs. 3 and 4. The main difference is that the duplication event in case of the proteolipid is more recent; it occurred in the branch leading to the eukaryotes after the archaebacteria separated. There-

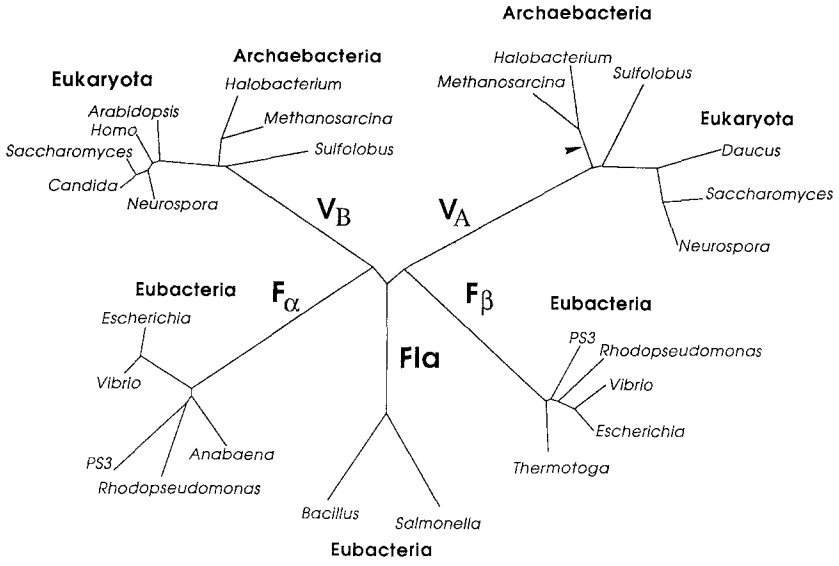

Fig. 3. Evolution of catalytic and noncatalytic V/F-ATPase subunits and the ATPase-like flagellar protein. The depicted "preliminary" tree represents the minimum-length tree calculated from a distance matrix. The tree is preliminary in the sense of Feng and Doolittle (1987) in that it utilizes only pairwise comparisons. It does not require a global alignment of all sequences. This procedure sacrifices accuracy in favor of a bias-free alignment procedure. Distance matrix and tree calculation were performed as described for Fig. 2, except the unitary matrix and 20 randomizations each were used to calculate the distances. The sequence for *Thermotoga maritima* was kindly provided by K. H. Schleifer, Tech. Univ. Munich, Germany, the sequences for *Halobacterium* were taken from Ihara and Mukohata (1991), and all other sequences were retrieved from GenBank.

fore, the size of the archaebacterial proteolipid is identical to the eubacterial proteolipid.

ORIGIN OF THE EUKARYOTIC ENDOMEMBRANE SYSTEM

The fact that the analyses of the ATPase sequences result in trees of the same topology as other markers for the nuclear/cytoplasmic compartments (e.g., Iwabe *et al.*, 1989; Puhler *et al.*, 1989) suggests that the ATPases energizing the eukaryotic endomembranes and the endomembrane system itself evolved within the cell that functioned as the host for the symbioses that evolved into the present-day eukaryotes, i.e., the endomembrane system itself is *not* due to another symbiotic event similar to the ones that gave rise to plastids and mitochondria. The evolution of V-ATPases from the coupling factor ATPase of archaebacterial plasma membrane furthermore suggests that the endomembrane system evolved from invaginations of the plasma membrane of the cell which contributed the bulk of the nuclear genome.

SIMILARITY OF V/F ATPase SUBUNITS TO THE *FLII* GENE PRODUCT

An ATPase-like peptide has been found to be encoded as part of the *fla* operon (ORF 4) of *Bacillus subtilis* (Ying *et al.*, 1991). A thermosensitive mutant of the homologous *fliI* gene in *Salmonella typhimurium* exhibits normal usage of their flagella at the restrictive tive temperature; only the assembly of flagella is interrupted in the mutant. The two gene products from *Bacillus* and *Salmonella* are 47% identical to each other and about 23% identical each to the catalytic and noncatalytic subunits of the V/F-ATPases; that is these ATPase-like proteins have about the same degree of similarity to these subunits as the paralogous V/F-ATPase subunits have among themselves. All pairwise comparisons between these flagellar ATPase-like proteins and the V-ATPase subunits resulted in z values greater than 6 [using the Needleman and Wunsch (1970) comparison with the unitary comparison matrix and 20 randomizations each]. These similarities far exceed the ones observed between other ATPases and ATP handling enzymes (e.g., adenylate cyclase,

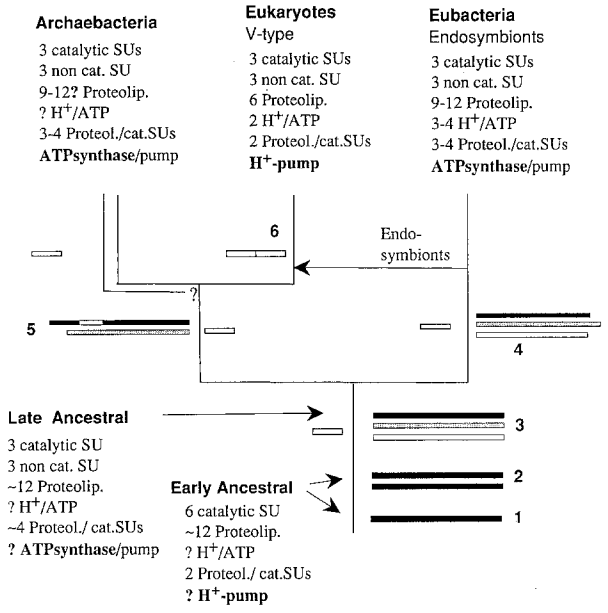

Fig. 4. Phylogenetic relations among the three kingdoms as derived from the analysis of ATPase subunits. The bars beside the tree schematically depict the evolution of the genes encoding the proteolipid (light gray), ATPase-like flagellar protein (white), catalytic (black), and noncatalytic subunits (dark gray). (1) The catalytic complex consists of a single subunit 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 noncatalytic subunit. Another gene duplication gives rise to the ATPase-like flagellar protein. The evolution from state 1 to 3 already had occurred in the last common ancestor. (4) In the line leading to the eubacteria and the endosymbionts, the noncatalytic subunit increases in size relative to the catalytic subunit. (5) 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. The proteolipid of the ATPase in the last common ancestor, like the one of eubacteria and archaebacteria, has a molecular weight of about 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. (6) During the evolution of the eukaryotes a duplication of the gene encoding the proteolipid and fusion of the resulting genes occurred. For discussion of the relation between H^+/ATP and subunit ratios see the text.

P-type ATPases or myosin; cf. Zimniak *et al.*, 1988); note that even these less related proteins are thought to be constructed out of homologous domains (Taylor and Green, 1989; Pederson and Carafoli, 1987).

In accordance with these observed similarities, the flagellar ATPase-like protein joins the tree of the ATPase subunits at the innermost branch (Fig. 3), i.e., the ancestral gene that gave rise to the two paralogous F/V ATPase subunits also gave rise to the flagellar ATPase-like protein. Figure 3 indicates that independent genes for all three proteins (i.e., the catalytic and noncatalytic subunits and the flagellar ATPase-like protein) were already present in the last common ancestor of eubacteria, archaebacteria, and eukaryotes. The evolution of the ATPase-like flagellar protein has been more rapid recently than that of the V/F subunits. While the *Bacillus* and *Salmonella* proteins are only 47% identical, the catalytic F-ATPase subunit from *Thermotoga maritima*, a thermophilic bacterium which represents the deepest known branch in the eubacterial domain (Woese, 1987), is more than 65% identical to the catalytic F-ATPase subunit of the other eubacteria.

Based on the similarity, it is appealing to assume that the ATPase-like flagellar protein is involved in the generation of local proton gradients; however, the phenotype of the thermosensitive mutants (Vogler *et al.*, 1991) is more compatible with the role of an ATPase involved in export of flagella-specific proteins.

ROOTING THE TREE OF LIFE AND THE V-ATPase NONHOMOLOGOUS REGION

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 paralogous genes that are derived from a gene duplication that occurred before the last common ancestor split off into different lines, the paralogous gene can be used as an outgroup, because it underwent a separate development from the time of the duplication (cf. Fig. 1).

In Fig. 3 one can consider the catalytic V/F-ATPase subunits as markers for the organismal evolution; the noncatalytic subunits and the ATPase-like flagellar protein can be employed as an outgroup. The resulting tree depicting the organismal evolution is given schematically in Fig. 4.

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 large, so-called nonhomologous region (Zimniak *et al.*, 1988) into the catalytic subunit of the archaebacterial and eukaryotic V-ATPase (Gogarten *et al.*, 1989a,b). This nonhomologous region is absent in the eubacterial catalytic subunit and in the outgroup (i.e., in all noncatalytic subunits and in the flagellar ATPase-like protein). Thus the nonhomologous region can be used

as a character defining the group composed of archaebacteria and eukaryotes as monophyletic in the rooted tree of life.

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 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).

For a proton-pumping ATPase the change in free energy associated with the hydrolysis of ATP (ΔG_{ATP}) has to be more negative than $n\Delta\mu_H^+$ (n is the number of protons pumped per ATP, μ_H^+ is the transmembrane gradient for protons); for an ATP synthase ΔG_{ATP} has to be less negative than $n\Delta\mu_H^+$. The difference in energy between ΔG_{ATP} and $n\Delta\mu_H$ is dissipated during the reaction. The sign of the difference determines the direction in which the reaction proceeds (proton pumping or ATP synthesis).

Therefore, 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 appear to fit this pattern (cf. Fig. 4). 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 noncatalytic subunits; the proteolipid was of the smaller type. From this, a proteolipid/catalytic subunit ratio and a H^+/ATP ratio of 3–4 corresponding to a predominant functioning as ATP synthase can be deduced. A correlate 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 noncatalytic subunit had not yet evolved (points 1 and 2 in Fig. 4), a V/F_1 portion consisting of six identical catalytic subunits results. The corresponding 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 of ATP synthase evolved only later when other means to generate a transmembrane proton gradient were available. If chemo- or 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 function and distribution of the ATPase-like flagellar protein will be of great interest with regard to this very early evolution. The evolutionary history of other ion-pumping V/F ATPases, including the F-type Na^+-translocating ATPase of the eubacterium *Propionigenium modestum* (Laubinger and Dimroth, 1987), the V-type ATPase of *Thermus thermophilus* (Tsutsumi *et al.*, 1991), and the V-type Na^+-translocating ATPase of *Enterococcus hirae*, also eubacteria (Kakinuma and Igarashi, 1990; Kakinuma *et al.*, 1991), should shed additional light on the early differentiation of these two ATPase types.

LATERAL GENE TRANSFER BETWEEN KINGDOMS

The V-type ATPase of *Thermus thermophilus* is especially interesting since it may be an example of lateral gene transfer from an archaebacterium to a eubacterium. Bacteria of the genus *Thermus*, because of their rRNAs and other biochemical characters (cell wall, lipids, antibiotic resistance), are clearly classified as eubacterial, with closest similarities to the Deinococci (Woese, 1987; Hensel *et al.*, 1986, 1989). Surprisingly, these bacteria contain an ATPase that is very similar to the eukaryotic/archaebacterial V-ATPase (Yokoyama *et al.*, 1990). The recently published sequence of the A subunit of this enzyme (Tsutsumi *et al.*, 1991) has 50–55% identical amino acid residues when aligned to the A subunits of the

```
                              1  2   3
Halobacterium                 GGTAAIPGPFGSGKTVTQQSLAKFA
Methanosarcina                GGTAAIPGPFGSGKTVTQQSLAKWS
Sulfolobus                    GGTAAIPGPFGSGKTVTLQSLAKWS

Archaebact. Consensus         GGTAAIPGPFGSGKTVTqQSLAKws
                              ************************
Thermus                       GGTAAIPGPFGSGKTVTQQSLAKWS
                              *** **** ** ****  *** * *
Eukaryotic Consensus          GGT-aIPGAFGCGKTVISQslSKyS

Saccharomyces                 GGTTCIPGAFGCGKTVISQSLSKYS
Neurospora                    GGTVAIPGAFGCGKTVISQSVSKFS
Daucus                        GGTCAIPGAFGCGKTVISQALSKYS
```

Fig. 5 Alignment of the highly conserved "phosphoanhydride binding site" found in all catalytic subunits of V/F-ATPases as well as many other ATP-binding proteins. In this region the archaebacterial character of the *Thermus* sequence is highlighted by the presence of three signature residues labelled 1, 2, and 3. Residue 1 is a proline residue in all known archaebacteria, and an alanine in eukaryotes. Residue 2 is a serine in all archaebacteria and a cysteine in eukaryotes. Residue 3 is a threonine in all archaebacteria and an isoleucine in eukaryotes.

eukaryotic V-ATPase and 53–60% identical residues when compared to the various archaebacterial subunits. Furthermore, when a phylogeny is constructed using the distance matrix method as described for Fig. 3, the *Thermus* sequence groups with the Archaebacteria, branching from the line leading to the Methanogens and Halophiles (Gogarten *et al.*, unpublished; the location of the *Thermus* sequence is indicated by an arrow in Fig. 3). The unambiguous archaebacterial character of the *Thermus* ATPase can easily be seen in an alignment of the putative phosphoanhydride binding site (Fig. 5). A similar surprising localization for a eubacterial sequence was reported by Iwabe *et al.* (1989) for the malate dehydrogenase from *Thermus flavus*. Also, for this enzyme the sequence determined for the *Thermus* enzyme groups together with its eukaryotic counterparts and not with the malate dehydrogenases from other eubacteria. The chimeric nature of the *Thermus* species seems to be a clear case of gene transfer between an archaebacterial-like organism and a eubacterium.

THE EVOLUTION OF V-ATPases WITHIN THE EUKARYOTIC DOMAIN

The V-ATPases in eukaryotes that have been studied function exclusively as proton pumps. It is not yet known if this is true for all eukaryotes. One could envision an early branching eukaryote, whose V-ATPase (or rather V-ATPase ancestor) is still located in the plasma membrane and can function as both proton pump and ATP synthase. Alternatively, the switch to exclusive proton pump might have occurred already in the archaebacterium like ancestor of the eukaryotic cell. In this case it might be possible to find an archaebacterial group that is a close relative to this eukaryotic ancestor and that already has performed the switch to exlusive ATP-driven proton pump. Likely candidates would be archaebacteria that no longer possess a proton-pumping redox chain in their plasma membrane and that rely exclusively on other means for ATP synthesis.

So far in the case of fungi (*Saccharomyces, Neurospora, Candida*), only one gene per genome has been found for each subunit. Gene disruption experiments that lead to a loss of V-ATPase activity in yeast (Umemoto *et al.*, 1990; Nelson and Nelson, 1990; Foury, 1990) are compatible with only one functional gene present for each subunit.

In the case of humans and in higher plants, different genes encoding the same subunit type have been reported. Südhoff *et al.* (1989) reported a *c*DNA from human kidney encoding the noncatalytic subunit (subunit B). A different, partial *c*DNA encoding the same subunit type has been obtained from hippocampal mRNA (Bernasconi *et al.*, 1990). The part of the encoded proteins for which both sequences have been determined is 91% identical on the amino acid level. The two isoforms are much more similar to each other than either is to the homologous subunit from plants or fungi. 75–80% idential amino acid residues are found for the comparisons between each of the two human isoforms and the sequences from *Arabidopsis, Neurospora, Candida*, and *Saccharomyces*. This suggests that the two isoforms observed in humans evolved during the evolution of the metazoans, but these probably tissue-specific isoforms are not a characteristic of multicellular eukaryotes in general.

Separate genes that encode the same subunit type (i.e., the catalytic subunit A) have been characterized by their sequence. Using primers directed toward the coding and the 3′ noncoding region of the *Daucus* mRNA resulted in more than three different genes being identified and partially sequenced in *Daucus* (Taiz, unpublished). Using two primers directed toward sequences conserved between the V-ATPases resulted in the amplification of two intron-containing DNA fragments from several plant species (Starke *et al.*, 1991). Currently the corresponding sequences from 13 different plants and algal species have been obtained. All of the sequences obtained from higher plants and their likely green algal ancestor contain a

surprisingly conserved intron (Starke *et al.*, 1991; Starke and Gogarten, unpublished). In the case of the angiosperms (with the exception of *Arabidopsis*), two fragments were obtained per species. These fragments could be characterized by their slightly different sized introns. Figure 6 depicts the likely evolutionary history of these fragments as reconstructed from their nucleotide sequences (Starke and Gogarten, unpublished). The time point of the invasion of the gene by this intron, and the duplication event that gave rise to these different genes, is indicated in Fig. 6. It is clear that the duplication had already occurred before the evolution of the angiosperm. Whether both genes were already present before the evolution of the different groups of ferns, or whether separate gene duplications occurred in the different lineages, is not settled. Sequence comparison suggests the latter, but the greater similarity within the different ferns might also be due to gene conversion events acting on the two forms in the different lineages. The sequence comparison as well as the presence of the intron in both genes clearly show that this particular gene duplication is restricted to the plant kingdom; the duplication event as well as the intron invasion occurred after separation from the fungal and metazoan lineages.

Transgenic plant cells (*Daucus carota*) that express antisense mRNA constructs directed against the catalytic subunit of the V_1 portion (subunit A) show a more than 70% reduction of V-ATPase catalytic subunit protein and of V-ATPase activity associated with the membrane of the central vacuole; however, V-ATPase levels and V-ATPase activity remained unchanged or even slightly increased in membrane fractions enriched in Golgi-derived vesicles (Fischmann, Gogarten, and Taiz, submitted). These findings are indicative of organelle-specific isoforms; however, the isoform that is not specific for the membrane of the central vacuole is currently only characterized by its nonreactiveness toward the antisense constructs, one of which was directed against the entire coding region, and by its cross-reactivity to a monospecific polyvalent antiserum raised against a 14 amino acid long peptide located near the amino terminal end of the other isoform, for which the entire sequence is known (Zimniak *et al.*, 1988). Southern blots of *Daucus carota* genomic DNA indicate the presence of several genes (Taiz, unpublished); however, Southern blots of *Arabidopsis* genomic DNA (Ecsedy and Gogarten, unpublished) as well as the above described PCR amplification across intron–exon boundaries gave no

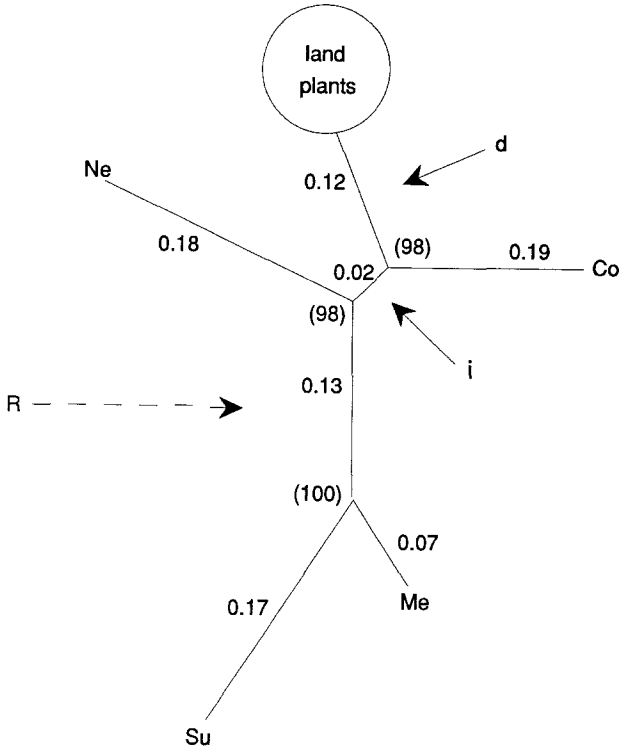

Fig. 6. Phylogenetic tree for the partial sequence of the catalytic subunit of the V-type ATPase. The topology of the tree was calculated using Felsenstein's maximum likelihood method (Felsenstein, 1981) as implemented in PHYLIP (Felsenstein, 1988). Only the coding regions of the sequences were used for the computation. The relative probability for substitution of the first, second, and third codon position (C) was set to 1 : 1 : 4. The transition/transversion ratio (T) was set at 3.3. The same topology was also obtained with a variety of different parameters (e.g., C = 2 : 1 : 10 and T = 2.3). The numbers at the branches give the length (i.e., the expected number of substitutions) as computed by the DNAML (maximum likelihood) program. All branches were determined to be significant at a level of p = 0.01. The numbers in brackets at the nodes give the percentage in which the same group of species was attached to this node during the evolution of 100 bootstrapped samples using parsimony analyses (as implemented in Felsenstein, 1988). The calculated topology was obtained with both fragments, the large and small one. The branch length given for the land plants was calculated for the small fragment of *Chenopodium rubrum*, and the same tree topology was obtained with all sequences (large and small) from land plants as denoted by the circle. From comparison with Fig. 3 the root of this partial tree can be inferred to be on the branch denoted by "R." The likely point of acquisition of the intervening sequence interrupting the coding region is denoted by "i," and the likely location of the duplication is indicated by "d." The abbreviations used for the species are as follows: Su = *Sulfolobus acidocaldarius*, Me = *Methanosarcina barkeri*, Ne = *Neurospora crassa*, Co = *Coleochaete scutata*.

indication for the presence of more than one isoform in *Arabidopsis*. No direct information is currently available as to how widely distributed these organelle-specific isoforms are. A more detailed characterization of the various isoforms will yield insights into developmental and tissue-specific regulation, and into the evolutionary process that gave rise to the highly differentiated endomembrane system of higher eukaryotes.

ACKNOWLEDGMENTS

This research was funded in part by NSF grant BSR-9020868 to JPG.

REFERENCES

Bennett, A. B., and Spanswick, R. M. (1984). *Plant Physiol.* **74**, 545–548.

Bernasconi, P., Rausch, T., Gogarten, J. P., and Taiz, L. (1990). *FEBS Lett.* **259**, 227–229.

Bernasconi, P., Rausch, T., Struve, I., Morgan, L., and Taiz, L. (1990). *J. Biol. Chem.* **265**, 17428–17431.

Birman, S., Meunier, F.-M., Lesbats, B., LeCair, J.-P., Rossier, J., and Israel, M. (1990). *FEBS Lett.* **261**, 303–306.

Blair, H. C., Teichelbaum, S. L., Ghiselli, R., and Gluck, S. (1989). *Science* **245**, 855–857.

Cross, R. L., and Taiz, L. (1990). *FEBS Lett.* **259**, 227–229.

Dayhoff, M. O., Schwartz, R. M., and Orcutt, B. C. (1978). In *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3 (Dayhoff, M. O., ed.), Natl. Biomedical Research Foundation, Washington DC, pp. 345–358.

Denda, K., Konishi, J., Oshima, T., Date, T., and Yoshida, M. (1989). *J. Biol. Chem.* **264**, 7119–7121.

Feeley, D. E., and Deyer, J. K. (1987). *J. Protozool.* **34**, 80–83.

Felsenstein, J. (1981). *J. Mol. Evol.* **17**, 368–376.

Felsenstein, J. (1988). Phylogeny Inference Package, Version 3.2 and 3.3, Department of Genetics SK-50, University of Washington, Seattle, Washington 98195.

Feng, D., and Doolittle, R. F. (1987). *J. Mol. Evol.* **25**, 351.

Fillingame, R. H. (1990). In *The Bacteria: A Treatise on Structure and Function*, Vol. 12 (Krulwich, T. A., ed.), Academic Press, New York, pp. 345–391.

Fitch, W. M., and Margoliash, E. (1967). *Science* **15**, 279–284.

Forgac, M. (1989). *Physiol. Rev.* **69**, 765–796.

Foury, F. (1990). *J. Biol. Chem.* **265**, 18554–18560.

Gluck, S., and Cadwell, J. (1987). *J. Biol. Chem.* **262**, 15780–15787.

Gogarten, J. P., Kibak, H., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, T., Oshima, T., Konishi, J., Denda, K., and Yoshida, M. (1989a). *Proc. Natl. Acad. Sci. USA,* **86**, 6661–6665.

Gogarten, J. P., Rausch, T., Bernasconi, P., Kibak, H., and Taiz, L. (1989b). *Z. Naturforsch.* **44c** 641–650.

Hanada, H., Hasebe, M., Moriyama, Y., Maeda, M., and Futai, M. (1991). *Biochem. Biophys. Res. Commun.* **176**, 1062–1067.

Hensel, R., Demharter, W., Kandler, O., Kroppenstedt, M., and Stackebrandt, E. (1986). *Int. J. Sys. Bacteriol.* **36**, 444–453.

Hensel, R., Zwickl, P., Fabry, S., Lang, G., and Palm, P. (1989). *Can. J. Microbiol.* **35**, 81–85.

Hoppe, J., and Sebald, W. (1984). The proton conducting Fo-part of bacterial ATP synthases, *Biochim. Biophys. Acta* **768**, 1–27.

Ihara, K., and Mukohata, Y. (1991). *Arch. Biochem. Biophys.* **286**, 111–116.

Inatomi, K.-I., Eya, S., Maeda, M., and Futai, M. (1989). *J. Biol. Chem.* **264**, 10954–10959.

Iwabe, N., Kuma, K.-I., Hasegawa, M., Osawa, S., and Miyata, T. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 9355–9359.

Iwabe, N., Kuma, K.-I., Kishino, H., Hasegawa, M., and Miyata, T. (1991). *J. Mol. Evol.* **32**, 70–78.

Kakinuma, Y., and Igarishi, K. (1990). *FEBS Lett.* **271**, 97–101.

Kakinuma, Y., Igarishi, K., Konishi, K., and Yamato, I., (1991). *FEBS Lett.* **292**, 64–68.

Lai, S., Watson, J. C., Hansen, J. N., and Sze, H. (1991). *J. Biol. Chem.* **266**, 16078–16084.

Lake, J. A. (1991). *Trends Biochem. Sci.* **16**, 46–50.

Laubinger, W., Dimroth, P. (1987). *Eur. J. Biochem.* **168**, 475–480.

Lindmark, D. G. (1988). *Exp. Parsitol.* **65**, 141–147.

Linkkila, T. P., and Gogarten, J. P. (1991). *Trends Biochem. Sci.* **16**, 287–288.

Mandel, M., Moriyama, Y., Hulmes, J. D., Pan, Y.-C. E., Nelson, H., and Nelson, N. (1988). *Proc. Natl. Acad. Sci. USA* **85**, 5521–5524.

Meagher, L., McLean, P., and Finbow, M. E. (1990). *Nucleic Acids Res.* **18**, 6712–6712.

Mitchell, P. (1976). *Biochem. Soc. Trans.* **4**, 399–430.

Mukohata, Y., Ihara, K., Kishino, H., Hasegawa, M., Iwabe, N., and Miyata, T. (1990). *Proc. Jpn. Acad.* **66**, 63–67.

Nanney, D. L., Mobley, D. O., Preparata, R. M., Meyer, E. B., and Simon, E. M. (1991). *J. Mol. Evol.* **32**, 316–327.

Needleman, S. B., and Wunsch, C. D. (1970). *J. Mol. Biol.* **48**, 443–453.

Nelson, H., and Nelson, N. (1989). *FEBS Lett.* **247**, 147–153.

Nelson, H., Nelson, N. (1990). *Proc. Natl. Acad. Sci. USA* **87**, 3503–3507.

Nolta, K. V., Padh, H., and Steck, T. L. (1991). *J. Biol. Chem.* **266**, 18318–18323.

Pearson, W. R., and Lipman, D. J. (1988). *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.

Pedersen, P. L., and Carafioli, E. (1987). *Trends Biochem. Sci.* **12**, 146–150, 186–189.

Puhler, G., Leffers, H., Gropp, F., Palm, P., Klenk, H. P., Lottspeich, F., Garrett, R. A., and Zillig, W. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 4569–4573.

Schäfer, G., Anemüller, S., Moll, R., Meyer, W., and Lübben, M. (1990). *FEMS Microbiol. Rev.* **75**, 335–348.

Sogin, M., Gunderson, J. H., Elwood, H. J., Alonso, R. A., and Peatie, D. A. (1989). *Science* **243**, 75–77.

Starke, T., Linkkila, T. P., and Gogarten, J. P. (1991). *Z. Naturforsch.* **46c**, 613–620.

Südhoff, T. C., Fried, V. A., Stone, D. K., Johnston, P. A., and Xie, X.-S. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 6067–6071.

Taylor, W. R., and Green, N. M. (1989). *Eur. J. Biochem.* **179**, 241–248.

Tsutsumi, S., Denda, K., Yokoyama, K., Oshima, T., Datd, T., and Yoshida, M. (1991). *Biochim. Biophys. Acta* **1098**, 13–20.

Umemoto, N., Yoshihisa, T., Hirata, R., and Anraku, Y. (1990). *J. Biol. Chem.* **265**, 18447–18453.

Vogler, P. A., Homma, M., Irikura, V. M., and MacNab, R. M. (1991). *J. Bacteriol* **173**, 3564–3572.

Woese, C. R. (1987). *Microbiol. Rev.* **51**, 221–271.

Ying, C., Scoffone, F., Albertini, A. M., Galizzi, A., and Ordal, G. W. (1991). *J. Bacteriol.* **173**, 3584–3586.

Yokoyama, K., Oshima, T., and Yoshida, M. (1990). *J. Biol. Chem.* **265**, 21946–21950.

Yoshimi, K., and Kazuei, I. (1990). *FEBS Lett.* **271**, 102–105.

Zimniak, L., Dittrich, P., Gogarten, J. P., Kibak, H., and Taiz, L. (1988). *J. Biol. Chem.* **263**, 9102–9112.