J Mol Evol (2001) 52:17–28
DOI: 10.1007/s002390010130

JOURNAL OF MOLECULAR EVOLUTION

Detecting Changes in the Functional Constraints of Paralogous Genes

Ignacio Marín, Mario Alí Fares, Fernando González-Candelas, Eladio Barrio, Andrés Moya

Instituto Cavanilles de Biodiversidad y Biología Evolutiva and Departamento de Genética, Universidad de Valencia, Dr. Moliner, 50, Burjassot 46100 (Valencia) Spain

Received: 24 February 2000 / Accepted: 12 August 2000

Abstract. We describe a new procedure to determine whether regional alterations in the evolutionary constraints imposed on paralogous proteins have occurred. We used as models the A and B (alternatively called α and β) subunits of V/F/A-ATPases, originated by a gene duplication more than 3 billion years ago. Changes associated to three major splits (eubacteria versus Archaea-eukaryotes; Archaea versus eukaryotes; and among free-living bacteria and symbiotic mitochondria) were studied. Only in the first case, when we compared eubacterial or mitochondrial F-ATPases versus eukaryotic vacuolar V-ATPases or archaeal A-ATPases, constraint changes were observed. Modifications in the degree of regional constraining were not detected for the other two types of comparisons (V-ATPases versus A-ATPases and within F-ATPases, respectively). When the rates of evolution of the two subunits were compared, it was found that F-ATPases regulatory subunits evolved faster than catalytic subunits, but the opposite was true for A- and V-ATPases. Our results suggest that, even for universal and essential proteins, selective constraints may be occasionally altered. On the other hand, in some cases no changes were detected after periods of more than 2.2 billion years.

Key words: Molecular coevolution — Adaptation — ATPase — Functional constraints — Evolvability

Correspondence to: I. Marín; *email:* ignacio.marin@uv.es

Introduction

For proteins that are physically interacting, we would expect that selective pressure on one of them sometimes influences the evolution of its interacting partners. However, to demonstrate that two proteins coevolved—i.e., that adaptive changes in a protein induced adaptive changes in a second, interacting, protein—is very difficult. Two types of evidence are required. First, it has to be determined that certain interacting proteins have undergone coordinated changes in their sequences. Second, it has to be established that those coordinated changes are indeed due to their mutual interactions and not to correlated responses to external selective processes.

In this work, we present a novel approach to analyzing the first aspect, applied to the particular case of interacting proteins encoded by paralogous genes. We wanted to determine whether homologous regions of two paralogous proteins are more constrained in some organisms than in others. The rationale behind our method is simple: Consider a parental lineage where a gene is duplicated. At some point, such lineage splits into two descendant lineages. Now, we compare the current degree of differentiation of the paralogous genes of two species, one from each of those descendant lineages. If such differentiation is caused by random fixation of mutations since the split occurred, it must be approximately the same in the two species. If, on the contrary, the differentiation of some regions of the duplicated proteins for a species of one lineage is significantly higher than the differentiation in a species of the other lineage, we can conclude that there has been at some point after the separation of both lineages a change in the constraints acting on those genes.

For this type of study, three related data sets are necessary. First, comparative sequence data is needed. Second, a well-established phylogeny is required because it is necessary to know unambiguously how the analyzed sequences are evolutionarily related. Finally, the results obtained by sequence comparison can be interpreted only when precise information on the structure and function of these proteins is available. In addition to these requirements, we are interested in interacting paralogous proteins. In this work, we have analyzed the evolution of two paralogous subunits of the protein complexes that function as A-, V-, or F-ATPases (collectively called V/F/A-ATPases). These two subunits are respectively named A and B or α and β, the nomenclature varying with the type of ATPase. Throughout the text, and to avoid ambiguity, we will name these proteins as "regulatory subunit" (usually called subunit α for F-ATPases and subunit B for V- and A-ATPases) and "catalytic subunit" (corresponding to subunit β for F-ATPases and subunit A for V- or A-ATPases). These names derive from the fact that although both subunits bind ATP, one contains the catalytic center, while the precise function of the nucleotide-binding sites in the regulatory subunits is still poorly understood (reviewed in Stevens and Forgac 1997). The evolutionary history of these proteins has been thoroughly studied. They originated from a gene duplication that occurred before the last universal common ancestor, more than 3 billion years ago (Feng et al. 1997). Thus, they are among the few proteins that provide useful information for the deepest branches of the tree of life, including the splits of eubacteria, Archaea, and eukaryotes (Gogarten et al. 1989, 1992; Gogarten 1994; Hilario and Gogarten 1998). It is also well established that eukaryotic F-ATPases derive from the symbiotic events that gave rise to mitochondria and chloroplasts, whereas it is generally accepted that V/A-ATPases were present in the cell membrane of the progenitor of Archaea and eukaryotes and subsequently internalized in the latter (Gogarten et al. 1992). Accordingly, the subunits of F-ATPases of mitochondria and eubacteria, in particular proteobacteria, are closely related, whereas the subunits of V-ATPases and A-ATPases are quite similar but very different from those in F-ATPases (Gogarten et al. 1989 and many subsequent references).

These two proteins are an excellent model for four reasons. First, being present in all organisms, we can determine the possible effects of profound changes in lifestyles, as well as changes over very long periods of time. Second, the protein complexes that include these subunits are acting either reversibly as ATP synthases/ATPases—as in Archaea (A-ATPases) or in eubacteria and eukaryotic mitochondria (F-ATPases)—or only as ATPases (V-ATPases of eukaryotic vacuolae). Thus, the effect of partial changes in their biological function may also be analyzed. Third, these subunits are intimately interacting in the ATPase complex. There are three catalytic and three regulatory subunits in close contact in the globular, hydrophilic sector of the complex known as F_1, V_1, or A_1 (for F-, V-, or A-ATPases, respectively). Finally, the precise three-dimensional folding of these subunits, and thus their physical interactions, is well understood. It has been determined for bovine and rat mitochondrial F_1-ATPases (Abrahams et al. 1994; Bianchet et al. 1998) and for the bacterial F_1-ATPase of *Bacillus* PS3 (Shirakihara et al. 1997).

Materials and Methods

We have developed a novel method (which we call "constraint analysis") to compare sets of sequences of paralogous genes. It is related to other recently proposed methods for analyzing regional heterogeneity in DNA and protein sequences (e.g., Dorit and Ayala 1995; Goss and Lewontin 1996; Tang and Lewontin 1999 and references therein). Our procedure involves several steps. First, the amino acidic sequences are aligned and conserved regions are chosen. Second, the sequences of the two subunits of each species are compared, one amino acid at a time, and a value of similarity for each residue is assigned. Then, two species are compared, using such similarity values, to determine whether certain regions of the paralogous subunits of one species are more similar than the corresponding regions of the paralogs of the other species.

Obtainment and Validation of Multiple-Sequence Alignments. Sixteen protein sequences were analyzed, corresponding to the catalytic and regulatory subunits of four F-ATPases, two V-ATPases, and two A-ATPases. The sequences were selected to cover an as wide as possible phylogenetic range. They were: (1) F-ATPases from *Drosophila melanogaster* mitochondria (accession numbers Y07894 and Q05825 for α and β subunits, respectively) and (in descending order of relatedness with the mitochondrial endosymbiont, and always citing regulatory subunits first) those of a proteobacterium (*Escherichia coli;* acc. nos. 67814; P00824), a Gram-positive bacterium (*Bacillus subtilis;* P37808; P37809), and a cyanobacterium (*Synechocystis* sp.; P27179; P26527); (2) the V-ATPases of representatives of two of the main eukaryotic groups, the yeast *Schizosaccaromyces pombe* (S25335; P31406) and the fly *D. melanogaster* (P31409; P48602); and (3) A-ATPases of the euryarchaeon *Halobacterium sanitarium* (P25164; P25163) and the crenarchaeon *Sulfolobus acidocaldarius* (P13052; P09639), representing the two main groups of Archaea (Woese 1987).

We used the global, progressive alignment program Clustal X (Thompson et al. 1997) to obtain a multiple-sequence alignment for those 16 sequences. After obtaining that alignment using Clustal X default parameters (pairwise alignment parameters: gap opening penalty = 10.0; gap extension penalty = 0.1; multiple alignment parameters: gap opening penalty = 10.0; gap extension penalty = 0.05), we selected six regions of 25 or more amino acids with a high degree of conservation, which we will call "modules." To define the limits of those six modules, avoiding the inclusion of regions that are too variable, we used a quite conservative method. First, we grouped similar amino acids according to the Blosum62 matrix (Henikoff and Henikoff 1992), as defined by default in the GeneDoc program (Nicholas and Nicholas 1997). The groups are as follows: D–N; E–Q; S–T; K–R; F–Y–W; L–I–V–M; P; A; G; C; H (the last five amino acids were thus considered independently). We then defined as "variable positions" those for which less than 50% of the sequences have amino acids of the same group, and we set a limit for extension of the modules of having at most three variable positions at their ends or six internally. This convention allowed to join two close modules with at most three terminal variable positions each. Gaps of up to two amino acids were allowed within a module, but no gaps were allowed at the ends of the

modules. The alignment obtained and the modules defined are shown in Fig. 1.

We used three different approaches to test the reliability of the defined modules, given the dependence of our subsequent analyses on a proper alignment. First, we varied the parameters of the Clustal X program. Relaxation of the alignment conditions (i.e., pairwise alignment parameters: gap opening penalty = 7.0; gap extension penalty = 0.07; multiple alignment parameters: gap opening penalty = 7.0; gap extension penalty = 0.03) did not affect the modules. As a second independent test, we used the local, segment-based, iterative program DIALIGN 2 (default parameters T = 0; see Morgenstern 1999. The program is available online at the Institute Pasteur Web pages [http://bioweb.pasteur.fr/seqanal/interfaces/dialign2-simple.html]). This program obtains multiple alignments using an algorithm totally different from Clustal X. Third, when discrepancies appeared in the Clustal X and DIALIGN 2 outputs, we used a third program, LALIGN (developed by W. Pearson, based on Huang and Miller 1991, implemented online at the European Molecular Biology Network, Swiss node Web pages [http://www.ch.embnet.org/software/LALIGN_form.html]). This program aligns two sequences using an algorithm unrelated to those implemented by Clustal X or DIALIGN 2.

Constraint Analysis. Once the modules were defined, we eliminated those positions that contained gaps in one or several sequences. Then, we compared one by one the amino acids of the two subunits of each species, assigning a value of conservation for each position according to the Blosum62 matrix. We called Bl_k the Blosum62 value of the pair of amino acids located at position *k*. Notice that, because we were interested in assigning a global level of divergence for the two sequences, we did not consider polarity (i.e., two amino acids receive the same Bl_k value, no matter what subunit contains each amino acid). We then established whether a position was more or less similar *between* species by taking the value for the first species, Bl_{k1}, and subtracting from it the value for the second species, Bl_{k2}, thus obtaining the "constraint value" for position *k* between species 1 and 2, $C(k)_{12}$:

$$C(k)_{12} = Bl_{k1} - Bl_{k2} \quad (1)$$

If the differentiation of the paralogs followed the same dynamics since the two species being compared separated, then $C(k)$ values should be randomly distributed. To determine whether there were nonrandom associations of $C(k)$ values, we used computer simulation. We generated 1000 random sequences by shuffling the $C(k)$ values of a particular module (according to the algorithm described in Weir 1996, p. 386). Then, we calculated the sum of $C(k)$ values for a particular window size (w) starting in amino acid *k*. We called such a sum $S(k, w)$:

$$S(k,w) = C(k) + C(k + 1) + \ldots + C(k + w - 1) \quad (2)$$

A total of $N - w + 1$ different $S(k,w)$ values can be obtained for a sequence of size N and for a window size w. Once all these $S(k,w)$ values were calculated, we determined the extreme values, largest and smallest, of $S(k,w)$ for each of the 1000 randomly shuffled sequences. We determined the largest and smallest values of $S(k,w)$ in each module and compared them with the distribution of extreme values generated in the corresponding simulation. The observed values above the upper 2.5% or below the lower 2.5% of the extreme values in the simulation were considered significant. If multiple nested windows of different sizes occurred that had significant values, the one with the largest (or smallest, if negative) value of $S(k,w)$ was chosen. In this way, we avoided that a few positions with large deviant values induced significant results for large window sizes.

Analyses of Rates of Evolution of Catalytic and Regulatory Subunits. We followed the procedure developed by Takezaki et al. (1995) to compare the rates of evolution of the two subunits. This type of analysis involves two steps: (1) building a tree with a set of sequences, and (2) measuring the distances of the sets of sequences whose rates of evolution want to be compared with respect to an outgroup. Distances obtained using nucleotide-based analyses were large, with some values over 1 (data not shown). Thus, we followed the recommendations by Kumar et al. (1993), performing the analyses with protein sequences, using the six modules defined above. We used the PAML program (Yang 1997) for these analyses but implemented the Blosum62 matrix instead of those available by default in such program. The tree obtained showed the expected results (Gogarten et al. 1989): first a separation of regulatory from catalytic subunits and then, within each type of subunits, a split of the subunits of F-ATPases from those of V- and A-ATPases. We then estimated the average length of the branches for all subunits. Finally, we determined the mean and standard deviation of the mean for the distances among (1) F-ATPases regulatory subunits versus average catalytic subunit; (2) V- and A-ATPases regulatory subunits versus average catalytic subunit; (3) F-ATPases catalytic subunits versus average regulatory subunit; and (4) V- and A-ATPases catalytic subunits versus average regulatory subunit. Comparisons for values (1) and (2) and for values (3) and (4) are presented and discussed below.

Three-Dimensional Representations. Three-dimensional analyses of the location of the significant regions were made using the program RasMol version 2.5 (Sayle and Milner-White 1995) and the coordinates for bovine heart mitochondria F1-ATPase obtained by Abrahams et al. (1994).

Results

Figure 1 shows the Clustal X–based amino acid alignment of the 16 sequences detailed in the previous section. It also shows the six modules defined according to the conventions described above. DIALIGN 2 confirmed a total of 283 out of the 304 positions included in these modules (coincidence: 92.4%). Two of the modules (nos. 2 and 5 in Fig. 1) were aligned identically. Modules 1 and 4 showed only one and three changes, respectively (both gaps in these modules were moved one position to the right and the last two amino acids in module 4 were aligned differently in some sequences). As expected considering these high similarities, we found that the constraint analysis results presented below for modules 1 and 4, based on the Clustal X alignment, are qualitatively identical to those obtained using the DIALIGN 2 alignment (not shown).

Finally, for modules 3 and 6, the four sequences corresponding to regulatory subunits of F-ATPases were aligned differently by DIALIGN 2 and Clustal X. For module 3, this affected the last 10 amino acids of those four sequences that were aligned by DIALIGN 2 outside of the positions that define that module (Fig. 1). For module 6, the differences affected the three most N-terminal amino acids, the four most C-terminal amino acids (these seven amino acids were again aligned by DIALIGN 2 in positions outside of module 6) and the two amino acids situated N-terminally respect to the gap in the center of the module, that were shifted two positions to the right by DIALIGN 2.

LALIGN analyses never confirmed the DIALIGN 2

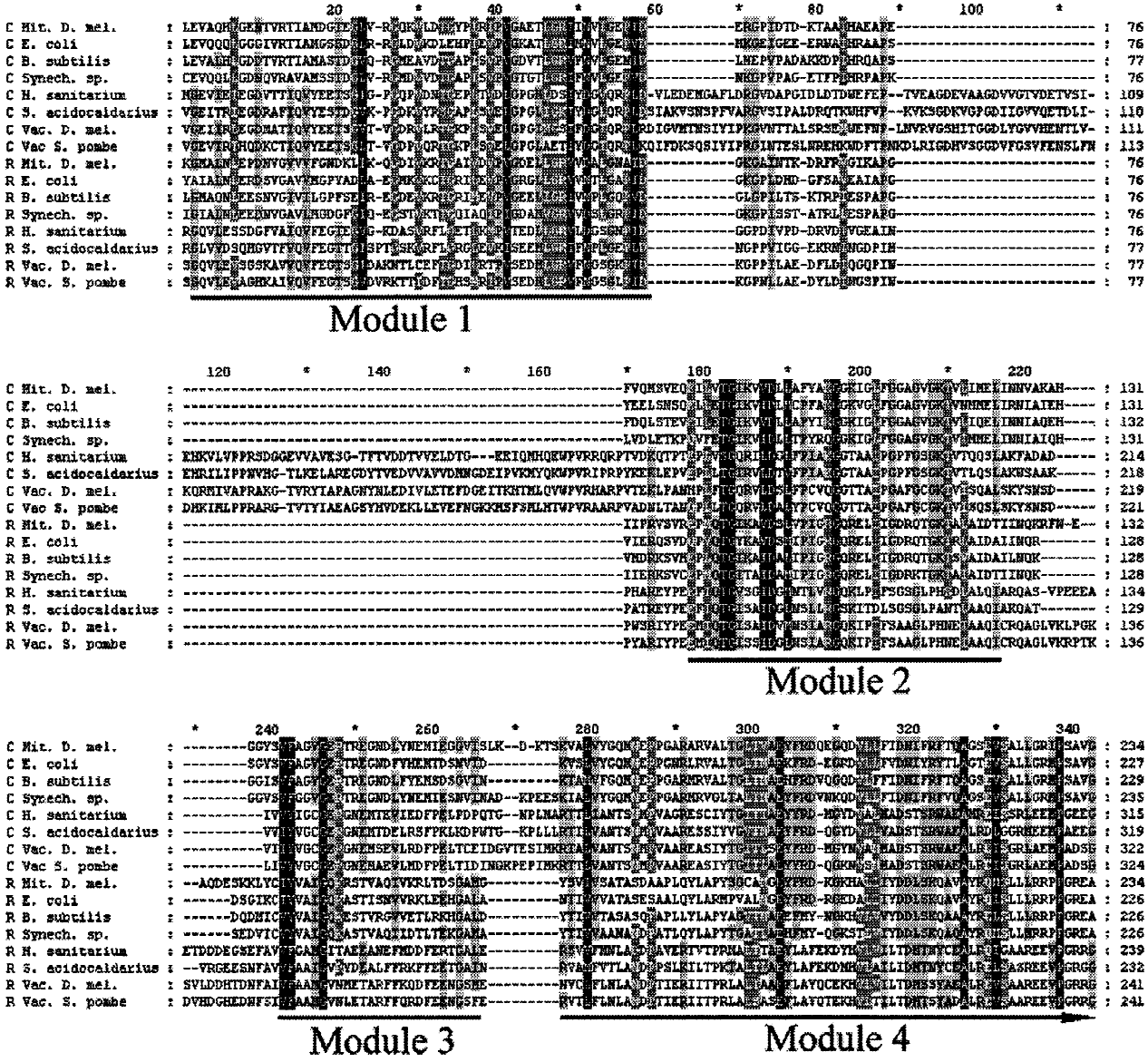

Fig. 1. Alignments of the catalytic and regulatory subunits of selected V/F/A-ATPases. Modules 1–6, used in subsequent analyses, are shown. Similarities have been highlighted using GeneDoc, as detailed in the Materials and Methods section. The highly variable N-terminal region of these proteins has not been included.

alignments for modules 3 and 6. On the other hand, for module 3, all LALIGN alignments comparing regulatory subunits of F-ATPases with regulatory subunits of V-ATPases or with catalytic subunits of F-ATPases confirmed the Clustal X results. However, inconsistent results were obtained for those comparisons between regulatory subunits of F-ATPases and catalytic subunits of A- or V-ATPases. For module 6, several noncongruent results were obtained in LALIGN analyses when F-ATPase subunits were used. Therefore, subsequent analyses for module 3 involving V- or A-ATPases as well as those module 6 results that involve F-ATPases have to be regarded with caution.

Figure 2 summarizes the departures from randomness of $S(k,w)$ values for the 28 possible comparisons among these sequences, for windows with two or more amino acids. Details of the significant regions, including average values of $S(k,w)$ are shown in Table 1. Figure 2 shows that significant results for comparisons involving ATPases of the same class were rare and, in general, limited to windows of small size. In the six comparisons among F-ATPases, only three long windows with significantly deviant results were observed. Only in one case results were consistent, namely, the short significant run in the center of module 4, where both mitochondrial and *E. coli* F-ATPases were more conserved than those of *B. subtilis* or *Synechocystis* sp. The other four significant runs, including the three largest ones, appeared only once. Similarly, for the single comparisons among V-ATPases and A-ATPases, we found in each case only a single, very short run of amino acids with significant values.

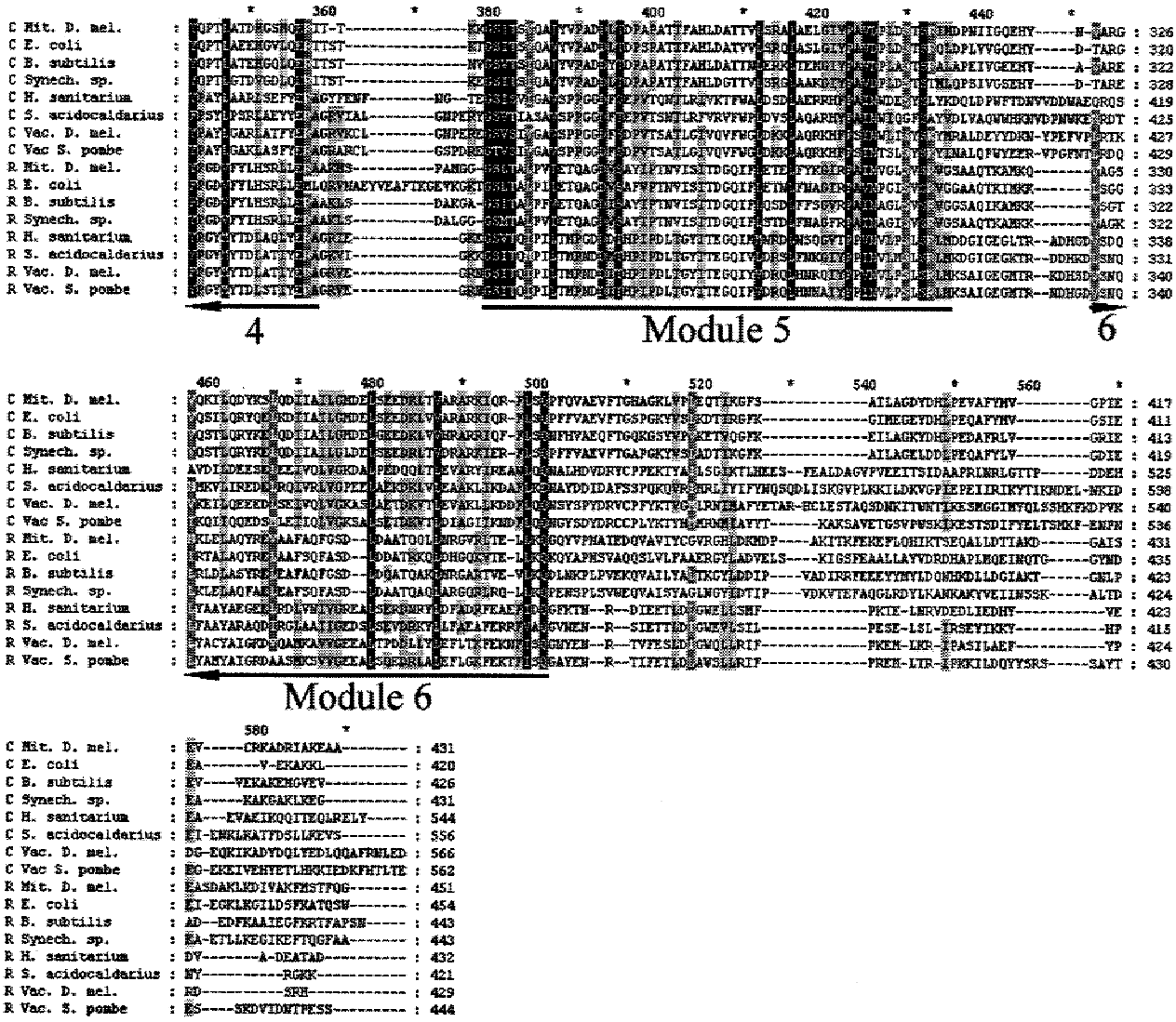

Fig. 1. Continued.

Comparisons of F-ATPases versus A-ATPases or F-ATPases versus V-ATPases gave more varied results. First, the values of modules 2 and 5 were, with a single exception, randomly distributed. However, modules 1, 4, and 6 very often showed significant departures from randomness, and also several significant values were observed in module 3. Results for modules 1 and 4 are especially significant, considering the data supporting the multiple alignment for these modules. Second, the results of all the comparisons were highly congruent, regardless of whether we considered comparisons of F-ATPases versus A-ATPases or of F-ATPases versus V-ATPases. Seven regions (two in module 1, one in module 3, two in module 4 and two in module 6) appeared four or more times (and the nonsignificant cases were often very close to the 2.5% significance level, see Discussion). Six of them were significant in both several V- versus F- and also in A- versus F- comparisons. The exception is the first region of module 6, which gave significant values only in some comparisons between V- and F-ATPases. However, as we explained above, F-ATPases alignments for module 6 are problematic, so this exceptional result may be spurious.

For the four comparisons among A- and V-ATPases, we found a situation that might be considered intermediate between the intragroup results and the results of A- and V-ATPases versus F-ATPases. Up to three significant regions were found. However, only two regions, in modules 3 and 6, appeared more than once. For comparisons among V- and A-ATPases, only module 6 may be unambiguously considered correctly aligned (see Materials and Methods). Thus, the region in module 6 showing significant differences in three out of the four comparisons is the only one clearly differentiated. In conclusion, consistent departures from randomness were also very rare in comparisons of A- versus V-ATPases, as rare as when molecules of the same type (F versus F, A versus A, or V versus V) are compared.

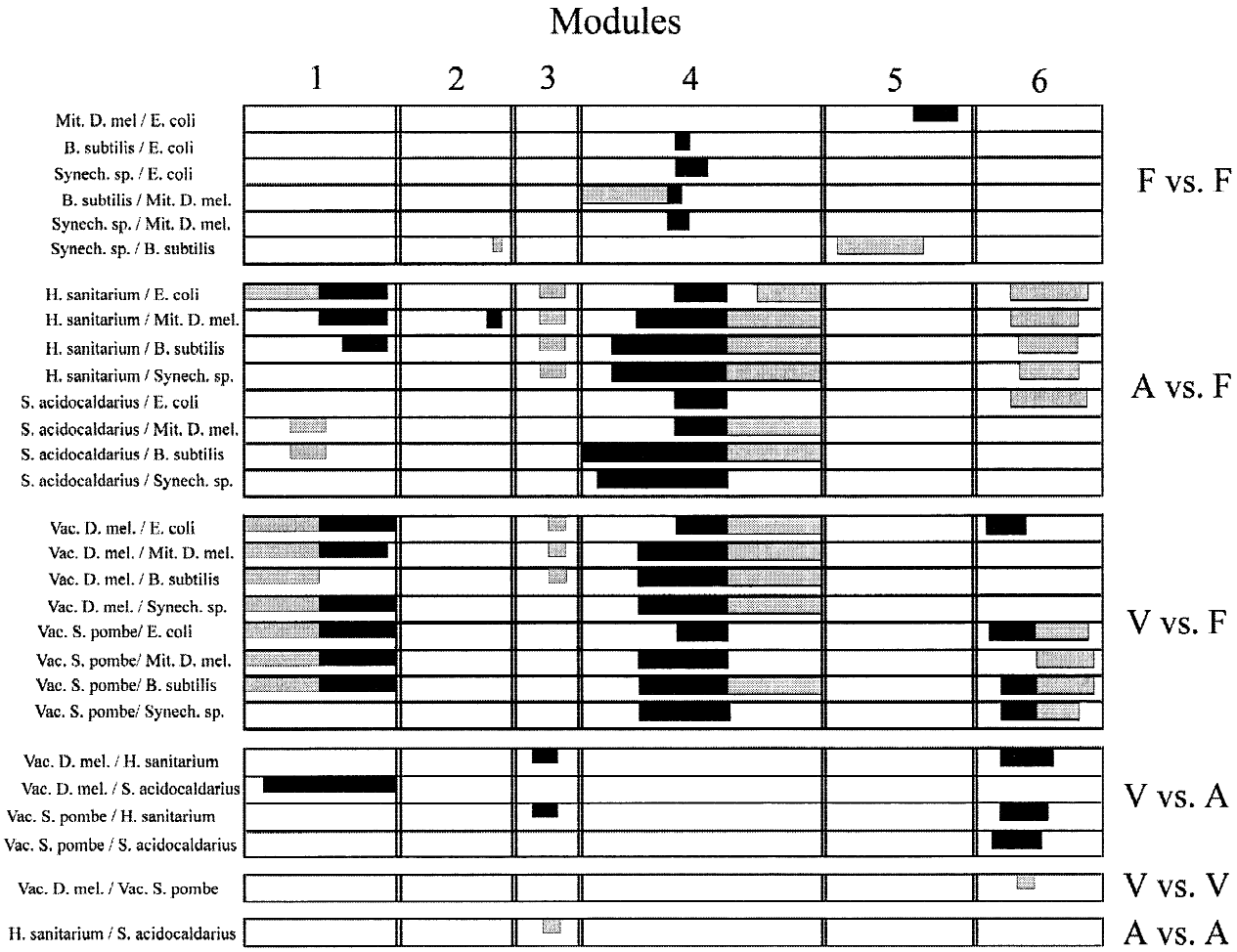

Fig. 2. Graphical summary of the results presented in Table 1. Positive values, that is regions where the two paralogous proteins of the first of the two named species showed a significantly higher degree of conservation that the two paralogous proteins of the second species, are shown here as gray boxes; negative values are depicted as black boxes.

The three-dimensional positions of the seven significant regions found in the comparison of F- versus A- or V-ATPases were analyzed. The spatial distributions of such regions were overimposed on the known structure of the α and β subunits of the F-ATPase of bovine heart mitochondria (Fig. 3; Abrahams et al. 1994). Three main results were obtained. First, the highly variable regions excluded from our modules were mostly in the outside of the barrel-like structure formed by the six subunits, that is, in those regions of the proteins that interact the least. Second, those regions in modules 1 and 4 where A- and V-ATPases showed less conservation than F-ATPases are spatially quite close, surrounding the place where the catalytic subunits of A- and V-ATPases have some 80–90 additional amino acids (see Fig. 1). Finally, the region around the active center was similarly conserved in F-, V-, and A-ATPase paralogs.

To establish whether different rates of evolution in regulatory or catalytic subunits could contribute to explain our results, we followed the procedures described by Takezaki et al. (1995) as detailed in the Materials and Methods section. The estimated shape parameter for the gamma distribution accounting for rate variation among sites was $\alpha = 2.43$. Distances obtained for the comparison F-ATPase regulatory subunits versus average catalytic subunits (1.61 ± 0.04) were significantly ($0.001 < p < 0.01$) larger than those for the comparison V-, A-ATPase regulatory subunits—average catalytic subunits (average distance $= 1.48 \pm 0.04$). Distances obtained for the comparison F-ATPase catalytic subunits versus average regulatory subunits (1.44 ± 0.02) were significantly shorter than those for the comparison V-, A-ATPase catalytic subunits—average regulatory subunits (average distance $= 1.65 \pm 0.04$; $p < 0.001$).

Discussion

Detecting Regional Heterogeneity in Paralogous Proteins

There is substantial interest in developing methods to detect regional heterogeneity in DNA or protein se-

quences (see Dorit and Ayala 1995; McDonald 1996; Goss and Lewontin 1996; Tang and Lewontin 1999). Here, we have developed a new procedure, which we have called constraint analysis, to establish whether runs of amino acids show significant departures from randomness when two pairs of paralogous proteins are compared. This kind of method has two difficulties. First, it is necessary to carefully select the regions included in the analysis. Improperly aligned sequences or sequences that are simply too divergent, thus increasing random noise, may alter the results. In this work, we have followed a series of conservative conventions, selecting regions that, according to the results of two different multiple-alignment algorithms, show a high degree of conservation. Clustal X and DIALIGN 2 were chosen because they use totally different algorithms and, at the same time, they have been evaluated as two of the most accurate programs for obtaining multiple-sequence alignments (Thompson et al. 1999; Morgenstern 1999). In general, the outputs of both programs were highly congruent. However, when differences arose, Clustal X alignments were often confirmed by LALIGN analyses, whereas DIALIGN 2 alignments were never compatible with LALIGN results. This can be explained by DIALIGN 2 occasionally finding spurious local maxima, an effect that results in a liberal addition of gaps to the alignment (discussed by Morgenstern 1999).

Figure 1 clearly shows the qualitative difference among those regions considered in our analysis, in general very similar in the different proteins, and those excluded, where gaps are abundant and containing only a few (if any) conservative positions. Thus, we think that the defined modules (perhaps excepting some regions of modules 3 and 6, for which the alignment results are sometimes ambiguous) correspond to "natural" conserved regions of these proteins. This is also supported by the fact that the excluded, variable regions lay in the outside of the globular structure formed by the six subunits (Fig. 3), in positions potentially less important for the function of these molecules and thus prone to rapid changes.

The second problem of methods devised to detect regional heterogeneity is the evaluation of the statistical significance of the results. When confronted with this problem, we explored two different possibilities. A first option was to compare each $S(k,w)$ value with the whole distribution of $S(k,w)$ values generated in simulations. However, two complications appeared: (1) a correction for the length of the module should be included, because the number of windows per module increases with increasing length, and therefore the number of tests also increases; (2) the different windows are not independent but correlated, and such correlation increases with increasing window size. Any length correction must also consider this fact and, thus, those corrections should be window-size specific. These problems lead us to consider a second possibility, that is, to compare for a particular window size the extreme (maximum and minimum) values of the observed data with the distributions of extreme values obtained in the simulations. This is a similar strategy to that used by Tang and Lewontin (1999) for detecting those regions in a set of sequences that accumulate an unexpected number of changes. The main difference is that they compared the extreme values for their test statistic with values obtained from simulations based on a theoretically expected function. We have compared instead with simulated values obtained from randomly shuffled sequences. This difference is due to the fact that a theoretical expected function for the $S(k,w)$ statistic cannot be simply determined. The use of extreme values overcomes the second problem described above (independence of windows) but not completely the first one, that is, the effect of differences in the number of tests for different module sizes. For two modules of different sizes, more tests for departure of randomness of extreme values are possible for the largest module, simply because there are more window sizes to be tested. However, this statistical problem does not affect the conclusions that follow in the next section, because it is only significant if comparisons among modules of different sizes are attempted. It does not affect comparisons of the same module for different sets of sequences, which are the basis for our discussion.

A final consideration is that the analysis performed in this work is based on primary protein structures. We have used three-dimensional information only to interpret the results obtained from that analysis. However, proximity in primary sequence is not a requirement to use constraint analysis. This method can be also used to analyze regions defined according not to primary but to secondary or tertiary protein structures (e.g., the region to analyze could be defined by characteristics as its closeness to the active center, being part of an exposed surface of the protein, etc.).

Analyzing Long-Term Regional Changes in Constraining in ATPase Subunits

Figure 2 and Table 1 can be quickly summarized. For the modules that we have defined, constraint analysis detected very little heterogeneity within each of the three types of ATPases, as well as in comparisons of A- versus V-ATPases. However, a considerable degree of nonrandomness became evident when A- or V-ATPases were compared to F-ATPases. As these proteins are very ancient, we have been able to test whether their constraints have been significantly altered since three major evolutionary events occurred (differentiation eubacteria versus Archea/eukaryotes; split Archaea versus eukaryotes; and eubacterial diversification, including the establishment of the symbiosis mitochondria/nucleus). Our method detects substantial changes only since the oldest of such

Table 1. Summary of the regions presenting significant deviant values

	No. of amino acids	Position in module	Average value ± SEM of $S(k,w)$ (%)
Module 1			
H. sanitarium/E. coli	23	1–23	1.91 ± 0.75 (2.1)
H. sanitarium/E. coli	29	24–52	−1.83 ± 0.81 (2.0)
H. sanitarium/Mit. *D. mel.*	29	24–52	−1.55 ± 0.79 (1.3)
H. sanitarium/B. subtilis	18	35–52	−2.28 ± 1.04 (2.3)
S. acidocaldarius/Mit. *D. mel.*	16	12–27	3.06 ± 0.92 (0.6)
S. acidocaldarius/B. subtilis	17	12–28	2.65 ± 0.76 (0.7)
Vac. *D. mel./E. coli*	21	1–21	2.48 ± 0.70 (0.3)
Vac. *D. mel./E. coli*	35	22–56	−1.77 ± 0.72 (0.0)
Vac. *D. mel.*/Mit. *D. mel.*	23	1–23	2.35 ± 0.70 (0.6)
Vac. *D. mel.*/Mit. *D. mel.*	29	24–52	−1.62 ± 0.63 (0.5)
Vac. *D. mel./B. subtilis*	21	1–21	2.38 ± 0.69 (2.3)
Vac. *D. mel./Synech.* sp.	23	1–23	2.13 ± 0.62 (1.6)
Vac. *D. mel./Synech.* sp.	33	24–56	−1.28 ± 0.67 (1.1)
Vac. *S. pombe/E. coli*	21	1–21	2.05 ± 0.75 (0.5)
Vac. *S. pombe/E. coli*	35	22–56	−2.23 ± 0.68 (0.3)
Vac. *S. pombe*/Mit. *D. mel.*	21	1–21	2.14 ± 0.78 (0.9)
Vac. *S. pombe*/Mit. *D. mel.*	33	24–56	−1.67 ± 0.70 (1.2)
Vac. *S. pombe/B. subtilis*	21	1–21	2.05 ± 0.78 (1.6)
Vac. *S. pombe/B. subtilis*	35	22–56	−1.74 ± 0.73 (1.3)
Vac. *D. mel./S. acidocaldarius*	45	12–56	−0.75 ± 0.46 (1.9)
Module 2			
Synech. sp./*B. subtilis*	4	34–37	2.25 ± 0.85 (2.3)
H. sanitarium/Mit. *D. mel.*	5	31–35	−5.20 ± 0.86 (1.9)
Module 3			
H. sanitarium/E. coli	10	10–19	2.60 ± 0.70 (0.1)
H. sanitarium/Mit. *D. mel.*	11	9–19	2.55 ± 0.79 (0.1)
H. sanitarium/B. subtilis	10	10–19	3.00 ± 0.94 (0.1)
H. sanitarum/Synech. sp.	10	10–19	2.70 ± 0.92 (0.2)
Vac. *D. mel./E. coli*	5	17–21	2.60 ± 0.93 (1.9)
Vac. *D. mel.*/Mit. *D. mel.*	5	17–21	2.80 ± 1.07 (1.4)
Vac. *D. mel./B. subtilis*	5	17–21	3.20 ± 1.46 (1.8)
Vac. *D. mel./H. sanitarium*	9	9–17	−2.44 ± 0.96 (1.3)
Vac. *S. pombe/H. sanitarium*	9	9–17	−2.78 ± 0.95 (0.4)
H. sanitarium/S. acidocaldarius	6	13–18	3.67 ± 1.43 (1.6)
Module 4			
B. subtilis/E. coli	6	31–36	−3.50 ± 1.77 (1.4)
Synech. sp./*E. coli*	9	31–39	−3.22 ± 1.32 (0.4)
B. subtilis/Mit. *D. mel.*	26	2–27	1.31 ± 0.49 (1.7)
B. subtilis/Mit. *D. mel.*	5	29–33	−4.60 ± 1.75 (0.1)
Synech. sp./Mit. *D. mel.*	6	29–34	−4.83 ± 1.49 (0.1)
H. sanitarium/E. coli	15	30–44	−2.80 ± 1.19 (0.1)
H. sanitarium/E. coli	24	59–82	3.04 ± 0.67 (2.3)
H. sanitarium/Mit. *D. mel.*	26	19–44	−1.50 ± 0.78 (0.2)
H. sanitarium/Mit. *D. mel.*	38	45–82	2.24 ± 0.59 (1.4)
H. sanitarium/B. subtilis	35	10–44	−0.97 ± 0.59 (0.7)
H. sanitarium/B. subtilis	38	45–82	2.08 ± 0.61 (0.3)
H. sanitarium/Synech. sp.	35	10–44	−0.71 ± 0.54 (0.9)
H. sanitarium/Synech. sp.	38	45–82	2.08 ± 0.58 (0.9)
S. acidocaldarius/E. coli	15	30–44	−2.93 ± 1.09 (0.3)
S. acidocaldarius/Mit. *D. mel.*	17	30–46	−2.58 ± 0.92 (0.3)
S. acidocaldarius/Mit. *D. mel.*	36	47–82	2.03 ± 0.64 (2.5)
S. acidocaldarius/B. subtilis	43	2–44	−0.86 ± 0.46 (0.1)
S. acidocaldarius/B. subtilis	17	45–82	−2.58 ± 0.92 (2.0)
S. acidocaldarius/Synech. sp.	40	5–44	−0.60 ± 0.43 (1.2)
Vac. *D. mel./E. coli*	15	30–44	−2.27 ± 0.95 (0.0)
Vac. *D. mel./E. coli*	38	45–82	2.13 ± 0.60 (2.3)
Vac. *D. mel.*/Mit. *D. mel.*	25	20–44	−2.08 ± 0.74 (0.2)
Vac. *D. mel.*/Mit. *D. mel.*	38	45–82	2.21 ± 0.61 (1.4)
Vac. *D. mel./B. subtilis*	25	20–44	−1.72 ± 0.64 (0.0)
Vac. *D. mel./B. subtilis*	38	45–82	2.05 ± 0.59 (0.4)
Vac. *D. mel./Synech.* sp.	25	20–44	−1.36 ± 0.55 (0.4)
Vac. *D. mel./Synech.* sp.	38	45–82	2.05 ± 0.60 (0.7)

Table 1. Continued.

	No. of amino acids	Position in module	Average value ± SEM of $S(k,w)$ (%)
Vac. *S. pombe/E. coli*	15	30–44	−3.06 ± 1.11 (0.4)
Vac. *S. pombe*/Mit. *D. mel.*	25	20–44	−2.12 ± 0.75 (0.1)
Vac. *S. pombe/B. subtilis*	25	20–44	−2.08 ± 0.65 (0.4)
Vac. *S. pombe/B. subtilis*	36	45–82	1.94 ± 0.58 (1.8)
Vac. *S. pombe/Synech.* sp.	25	20–44	−1.40 ± 0.62 (0.3)
Module 5			
Mit. *D. mel./E. coli*	20	32–51	−1.00 ± 0.49 (0.2)
Synech. sp./*B. subtilis*	33	8–40	0.45 ± 0.21 (1.4)
Vac. *D. mel./H. sanitarium*	17	41–57	
Module 6			
H. sanitarium/E. coli.	21	16–36	1.67 ± 0.47 (0.0)
H. sanitarium/Mit. *D. mel.*	21	14–34	2.14 ± 0.55 (0.0)
H. sanitarium/B. subtilis	18	17–34	1.72 ± 0.55 (2.3)
H. sanitarium/Synech. sp.	18	17–34	1.83 ± 0.57 (0.5)
S. acidocaldarius/E. coli	12	16–36	1.62 ± 0.59 (1.7)
Vac. *D. mel./E. coli*	12	4–15	−2.67 ± 0.82 (1.3)
Vac. *S. pombe/E. coli*	12	4–15	−2.83 ± 0.76 (0.3)
Vac. *S. pombe/E. coli*	21	17–37	1.95 ± 0.53 (0.1)
Vac. *S. pombe*/Mit. *D. mel.*	26	17–42	1.77 ± 0.70 (0.2)
Vac. *S. pombe/B. subtilis*	8	9–16	−4.13 ± 0.93 (0.4)
Vac. *S. pombe/B. subtilis*	26	17–42	1.62 ± 0.70 (1.8)
Vac. *S. pombe/Synech.* sp.	8	9–16	−3.25 ± 0.73 (0.6)
Vac. *S. pombe/Synech.* sp.	18	17–34	2.17 ± 0.54 (0.8)
Vac. *D. mel./H. sanitarium*	17	11–27	−1.88 ± 0.66 (0.3)
Vac. *S. pombe/H. sanitarium*	14	8–21	−2.35 ± 0.75 (0.4)
Vac. *S. pombe/S. acidocaldarius*	14	6–19	−1.79 ± 0.71 (2.2)
Vac. *D. mel.*/Vac. *S. pombe*	4	13–16	3.00 ± 1.08 (1.8)

The total number of amino acids in the significant runs and their positions in each module are detailed. The average value and standard error value of $S(k,w)$ for the significant regions are also included, together with the percentage of cases where the extreme cases obtained in the simulations had values higher or lower than the observed ones (in parentheses).

events, the eubacteria versus Archaea/eukaryotes split. We must emphasize now that the variability in the results involving comparisons of F-ATPases versus A- and V-ATPases that can be observed in Fig. 2 is (at least in several cases) more apparent than real. For example, for module 1, most comparisons between V- and F-ATPases show highly significant differences, but not so the comparison *S. pombe* vacuolae versus *Synechocystis* sp. However, the analysis for amino acids 1–21 and 25–56 of module 1 in that comparison shows almost significant values (only 4.4% maximum values of the simulations are higher than the observed ones for amino acids 1–21, only 2.7% are lower than the value observed for amino acids 25–56). The same happens for (1) the second part of module 1 in the comparison *D. melanogaster* vacuolae versus *B. subtilis* (level of significance, 2.6%); (2) the comparisons among A- and F-ATPases and V- and F-ATPases in module 4 (for those five cases in which the second half of the module does not show significantly lower values in F-ATPases, those regions still show deviations, levels of significance ranging from 3.1% to 6.2%); and (3) the three cases where there are no significant differences between A- and F-ATPases in module 6, all them involving *S. acidocaldarius*, actually show levels of significance between 3.7% and 6.9%. Such almost significant deviations are not present in any comparisons within the same group of ATPases.

Philippe and Forterre (1999) have shown that ATPase sequences are mutationally saturated (i.e., multiple substitutions per codon have occurred since the divergence of eubacteria, archaea, and eukaryotes). However, this problem should not affect constraint analysis results. Artifactual positive results due to saturation are highly unlikely, because multiple amino acidic changes in a position may occur without altering the $C(k)$ values. On the other hand, one may envisage artifactual negative results due to extreme divergence of the analyzed proteins (i.e., there have been so many changes that the paralogous sequences have become almost randomized with respect to one other). We think that this type of problem must be generally avoided by the fact that then we would be unable to effectively align the sequences. In any case, spurious negative results caused by saturation should be more frequent the least related the sequences were. In our study, however, we have found that the closest relatives are the ones that more frequently show lack of differentiation. This result strongly argues against a generalized artifact.

Our results have generally interesting implications. First, although it is known that different protein regions

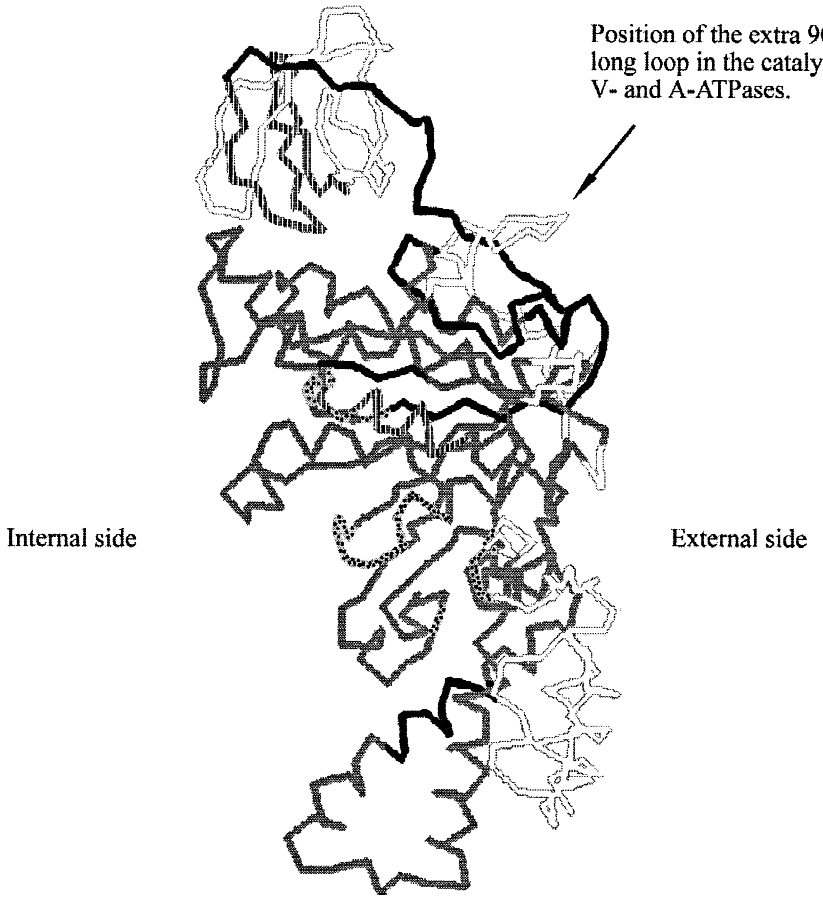

Fig. 3. Three-dimensional structure of F-ATPase subunits, shown in a lateral view. External side, toward the exterior of the ATPase complex, is right. N-terminal end of the molecule is at the top, C-terminus at the bottom. Light gray: nonconserved regions, excluded from the present analyses. Notice that most of these regions are in the external side of the complex. Dotted pattern: amino acids of the active center. Gray: residues where F-ATPases did not show any significant difference in constraining when compared with A- or V-ATPases. Black: regions where F-ATPases were more constrained than A- and V-ATPases. Black and white stripes: regions where F-ATPases were less constrained than A- and V-ATPases. Notice that two large regions in black, corresponding to the regions of modules 1 and 4 where the two subunits of A- and V-ATPases are highly divergent, are very close in space to the position where the catalytic subunits have the extra loop of amino acids (arrow).

have diverse levels of evolutionary constraint, precise answers to what level of constraint and change is associated with major adaptive changes are just emerging. Golding and Dean (1998) reviewed several cases and, among other conclusions, established that major adaptive shifts, including new activities, may be due to a few amino acidic changes. However, they also emphasized that many changes, some far apart from the active center, may be important for adaptation, contributing to the fine-tuning of the activity of the altered proteins. Our results suggest that, apart from the most obvious variable and constant regions that can be detected by simply comparing the sequences of two or several species, long-term evolution implies, at least in some cases, subtle general modifications of the degree of constraining of proteins, in such a way that regions that in some circumstances are conserved may at other times change substantially. Our results show that even universal and essential components of the biochemical machinery of the cells may suffer such dynamics.

Fitch and Markowitz (1970) suggested that amino acidic replacements are tolerated in different regions of the proteins in different lineages (the covarion model). Several empirical results are best explained by the covarion model (Fitch and Ayala 1994; Miyamoto and Fitch 1995; Lockhart et al. 1998; Lopez et al. 1999). Our results are compatible with this view. The local modifications in constraints in some lineages but not in others that we have detected cannot be simply explained by rate variation among positions. Thus, constraint analysis may be used as a method for detection of cases where a covarion model may apply.

Changes in the constraints acting on these proteins are confirmed by the analysis of rates of divergence among regulatory and catalytic subunits. We have observed that, for F-ATPases, regulatory subunits evolve slower than catalytic subunits, and the opposite is true for V- and A-ATPases. Moreover, the differences for both catalytic and regulatory subunits are significant when comparing F- versus V-, A-ATPases. Assuming that catalytic sub-

units should be in general more constrained than regulatory ones, and therefore evolve at a slower rate, the simplest explanation is that some regions of the regulatory subunits are intrinsically more constrained in V- and A-ATPases than in F-ATPases, while some regions of the catalytic subunit in V- and A-ATPases have suffered a drastic increase of their rate of change.

An attractive possibility was the detection of changes in constraining associated with the main functional difference between F- and A-ATPases, which are able to act also as ATP synthases, and V-ATPases, which cannot. Indeed, minor differences in constraining between V- and A-ATPases, especially in module 6 (see Fig. 2 and Results), cannot be ruled out. However, a generalized change when those two types of proteins are compared was not observed. Moreover, although there is a substantial number of significant deviations from randomness in the comparisons among F- and V-ATPases, they correlate well with those found in the comparison of F- and A-ATPases (Figure 2), suggesting that those differences are unrelated with any functional characteristic specific to V-ATPases. In fact, it has been suggested that the ATPase–ATP synthase functional change would be unrelated to the subunits studied here and related to changes in the membrane sector of ATPases (reviewed in Nelson 1995; see also Hilario and Gogarten 1998). Our data are compatible with this hypothesis.

Two models can explain the similarity of the results of the comparisons A- versus F- and V- versus F-ATPases. First, it is possible that the observed changes in constraining were caused by a single, very ancient event that occurred after the eubacteria versus Archaea/eukaryotes split and before the split between these two last groups, leaving a mark in the protein sequences that still can be observed today. A second possibility is that, once a certain unknown cause induced a shift in the requirements of the two subunits, differential constraints have been acting on their sequences for long periods of time and more or less continuously from then on. Because we are only observing the end products of the process, it is impossible to distinguish between these two extreme views. However, we favor the first one. In particular, the heterogeneity in constraining in modules 1 and 4 could be a response to the acquisition, in the catalytic subunits of V- and A-ATPases, of the 80–90-amino-acids-long loop (situated between modules 1 and 2, see Fig. 1). This loop is close in the 3D structure to the regions with significant deviations in modules 1 and 4 (Fig. 3). We think that modifications associated to the appearance of such a loop could also explain the higher rate of change for the catalytic subunits of V- and A-ATPases compared to those of F-ATPases. If this is the case, the time for the observed shifts to occur may have been relatively short. On the other hand, our data suggest that long-term stasis in the constraints acting on these proteins exist, because almost no significant deviations have been observed within groups. For example, considering that the last common ancestor of cyanobacteria, Gram-positive and Gram-negative bacteria may have lived about 2.2 billion years ago (Feng et al. 1997), the apparent absence of constraint changes among F-ATPases of species of these groups is striking. Bearing in mind also the big lifestyle change occurred, it is also very interesting that mitochondrial symbiosis has not apparently had significant effects. Of course, constraint changes in the most variable regions of these proteins, which we have excluded from our analyses, cannot be ruled out.

At the beginning of our work, we pointed out the interest in demonstrating molecular coevolution of interacting proteins. Our data have shown that pairs of interacting proteins may sometimes present changes in their sequences that can be interpreted as shifts in their functional constraints. Also, our data suggest that other similar pairs of proteins do not present alterations, even after more than 2.2 billion years of evolution. The question now is whether any of those changes or stasis may be due to coevolution between the two subunits. We think that the changes observed for the branches that separate F- from A- and V-ATPases may fit a model of independent evolution for the two subunits, with an important acceleration of the rate of change of the catalytic subunits of A-, V-ATPases with respect to those of F-ATPases and the opposite dynamics for the regulatory subunits. On the other hand, the substantial conservation in some regions of the proteins of the same constraints for long periods of time may be an argument for coevolution. The requirement for intimate interactions among these ATPase subunits may help create strict patterns of constraining. We have detected changes in this pattern only once in more than 3 billion years, precisely when F-ATPases and V/A-ATPases differentiated. Comparisons using similar procedures with pairs of paralogous but not interacting proteins may indicate whether such strict conservation of constraining patterns are common only for proteins for which coevolution is possible.

Acknowledgments. We would like to thank Santiago F. Elena for critical reading of a previous version of this manuscript. This work was supported by DGES project PB96-0793-C04-01.

References

Abrahams JP, Leslie AGW, Lutter R, Walker JE (1994) Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria. Nature 370:621–628

Bianchet MA, Hullihen J, Pedersen PL, Amzel LM (1998) The 2.8-Å structure of rat liver F_1-ATPase: configuration of a critical intermediate in ATP synthesis/hydrolysis. Proc Natl Acad Sci USA 95:11065–11070

Dorit RL, Ayala FJ (1995) ADH evolution and the phylogenetic footprint. J Mol Evol 40:658–662

Feng DF, Cho G, Doolittle R (1997) Determining divergence times with a protein clock: update and reevaluation. Proc Natl Acad Sci USA 94:13028–13033

Fitch WM, Ayala FJ (1994) The superoxide dismutase molecular clock revisited. Proc Natl Acad Sci USA 91:6802–6807

Fitch WM, Markowitz E (1970) An improved method for determining codon variability in a gene and its implication to the rate of fixation of mutations in evolution. Biochem Genet 4:579–593

Gogarten JP (1994) Which is the most conserved group of proteins? Homology-orthology, paralogy, xenology, and the fusion of independent lineages. J Mol Evol 39:541–543

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

Gogarten JP, Starke T, Kibak H, Fishmann J, Taiz L (1992) Evolution and isoforms of V-ATPase subunits. J Exp Biol 172:137–147

Golding GB, Dean AM (1998) The structural basis of molecular adaptation. Mol Biol Evol 15:355–369

Goss PJE, Lewontin RC (1996) Detecting heterogeneity of substitution along DNA and protein sequences. Genetics 143:589–602

Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915–10919

Hilario E, Gogarten JP (1998) The prokaryote-to-eukaryote transition reflected in the evolution of the V/F/A-ATPase catalytic and proteolipid subunits. J Mol Evol 46:703–715

Huang X, Miller W (1991) A time-efficient, linear-space local similarity algorithm. Adv Appl Math 12:337–357

Kumar S, Tamura K, Nei M (1993) MEGA: molecular evolutionary genetic analysis, version 1.01. University Park, PA: Pennsylvania State University

Lockhart PJ, Steel MA, Barbrook AC, Huson DH, Charleston MA, Howe CJ (1998) A covariotide model explains apparent phylogenetic structure of oxygenic photosynthetic lineages. Mol Biol Evol 15:1183–1188

Lopez P, Forterre P, Philippe H (1999) The root of the tree of life in the light of the covarion model. J Mol Evol 49:496–508

McDonald JH (1996) Detecting non-neutral heterogeneity across a region of DNA sequence in the ratio of polymorphism to divergence. Mol Biol Evol 13:253–260

Miyamoto MM, Fitch WM (1995) Testing the covarion hypothesis of molecular evolution. Mol Biol Evol 12:503–513

Morgenstern B (1999) DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. Bioinformatics 15:211–218

Nelson N (1995) Molecular and cellular biology of F- and V-ATPases. In: Nelson N (ed) Organellar proton-ATPases. Heidelberg (Germany): Springer-Verlag, pp 1–27

Nicholas KB, Nicholas HB (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors (www.cris.com/~ketchup/genecdoc.shtml)

Philippe H, Forterre P (1999) The rooting of the universal tree of life is not reliable. J Mol Evol 49:509–523

Sayle RA, Milner-White EJ (1995) RASMOL: biomolecular graphics for all. Trends Biochem Sci 20:374

Shirakihara Y, Leslie AG, Abrahams JP, Walker JE, Ueda T, Sekimoto Y, Kambara M, Saika K, Kagawa Y, Yoshida M (1997) The crystal structure of the nucleotide-free alpha 3 beta 3 subcomplex of F1-ATPase from the thermophilic *Bacillus* PS3 is a symmetric trimer. Structure 15:825–836

Stevens TH, Forgac M (1997) Structure, function and regulation of the vacuolar (H^+)-ATPase. Annu Rev Cell Dev Biol 13:779–808

Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized tree. Mol Biol Evol 12:823–833

Tang H, Lewontin RC (1999) Locating regions of differential variability in DNA and protein sequences. Genetics 153:485–495

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X Windows interface: flexible strategies for multiple-sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Thompson JD, Plewniak F, Poch O (1999) A comprehensive comparison of multiple sequence alignment programs. Nucleic Acids Res 277:2682–2690

Weir BS (1996) Genetic data analysis II. Sunderland, MA: Sinauer Associates

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

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. CABIOS 13:555–556