β -Subunit of ATP-synthase: a Useful Marker for Studying the Phylogenetic Relationship of Eubacteria

By RUDOLF AMANN, WOLFGANG LUDWIG AND
KARL HEINZ SCHLEIFER*

*Lehrstuhl für Mikrobiologie, Technische Universität München, Arcisstrasse 21,
D-8000 München 2, FRG*

(Received 16 May 1988; revised 4 July 1988)

The genes encoding the β -subunits of ATP-synthases (ATPases) from *Bacteroides fragilis* DSM 2151, *Cytophaga lytica* DSM 2039 and '*Taxebacter ocellatus*' were cloned. The nucleotide sequences were determined completely for the genes of the first two organisms and to a major part for that of '*T. ocellatus*'. The predicted amino acid sequences were compared with previously published amino acid sequences of β -subunits. Two characteristic insertions were found in genes from organisms belonging to the so-called bacteroides–cytophaga–flavobacterium group. The remaining structure shows a high degree of sequence similarity within this group. These data support the conclusions drawn from comparative 16S rRNA sequence analyses that organisms in this phenotypically heterogeneous group are phylogenetically related. A phylogenetic tree was constructed based on a distance matrix of optimally aligned amino acid sequences of β -subunits of ATPases of various eubacteria, chloroplasts and mitochondria. It is in good agreement with a tree derived from 16S rRNA sequence analyses.

INTRODUCTION

Animals and plants are rich in complex morphological detail and fossil records which can serve as a basis for phylogenetic classification. Prokaryotes, on the other hand, have a rather simple morphology and lack in general fossil records. The phylogenetic relationships of prokaryotes have to be deduced from sequence comparisons of macromolecules that are widely distributed, show a high degree of functional constancy and are sufficiently conserved to span the full evolutionary spectrum. The present view of the phylogenetic relationship of prokaryotes is almost exclusively based on data from one class of macromolecules, the ribosomal ribonucleic acids (Woese, 1987). However, in order to manifest and complete the current picture it is necessary to give it a broader base by incorporating data from other conserved macromolecules.

Eubacteria, mitochondria and chloroplasts contain a proton-translocating ATP-synthase (ATPase) complex (Futai & Kanazawa, 1983). The complex is composed of two portions, designated F_1 and F_0 . F_0 is intrinsic to the membrane and forms a proton channel. F_1 is an extrinsic membrane protein complex composed of five subunits of which one, the β -subunit, contains the catalytic site of the enzyme. The primary structure of the β -subunit seems to be highly conserved throughout evolution (Walker *et al.*, 1984; Tybulewicz *et al.*, 1984; Falk *et al.*, 1985; Zilberstein *et al.*, 1986; Kagawa *et al.*, 1986; Curtis, 1987; Cozens & Walker, 1987; Amann *et al.*, 1988). Since it is also widely distributed among eubacteria and shows functional constancy, it should be an ideal macromolecule for deducing the phylogenetic relationship of bacteria.

Many microbial taxonomists are not convinced that phenotypically dissimilar groups of bacteria, such as the anaerobic, generally pleomorphic, non-motile bacteroides and the aerobic, rod-shaped flavobacteria and gliding cytophaga should form a phylogenetically coherent unit (Paster *et al.*, 1985; Weisburg *et al.*, 1985). These organisms represent, according to the 16S

rRNA data, an ancient divergence in the eubacterial line of descent and one of the 11 eubacterial phyla (Woese, 1987). In the present study, we cloned, sequenced and compared genes encoding the β -subunit of ATPase from organisms belonging to three different genera of the bacteroides-cytophaga-flavobacterium group.

METHODS

Organisms. *Bacteroides fragilis* DSM 2151 was obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, FRG. Cells of *Cytophaga lytica* DSM 2039 and '*Taxebacter ocellatus*' were gifts from Dr Reichenbach, Gesellschaft für Biologische Forschung, Braunschweig, FRG. *Bacteroides fragilis* was grown anaerobically in a medium containing (g l⁻¹): peptone from meat (5); yeast extract (5); meat extract (5); glucose (5); Tween 80 (8); cysteine (0.5); Na₂HPO₄ (2) adjusted to pH 7.0.

Preparation of DNA and cloning of ATPase β -subunit genes. DNA was prepared using the method of Marmur (1961) modified by Meyer & Schleifer (1978). DNAs were digested with restriction endonuclease *Pst*I, separated by agar gel electrophoresis and probed by Southern (1975) hybridization. A cloned fragment of the β -subunit gene of *Escherichia coli* ATPase coding for amino acids 16–340 (Saraste *et al.*, 1981; kindly provided by J. E. Walker, Cambridge, UK) was separated from vector DNA (M13 mp7), nick-translated with [α -³²P]dATP (New England Nuclear) and used for probing. Prehybridization and hybridization were done as described by Amann *et al.* (1988). DNA fractions which hybridized to the *E. coli* β -subunit probe were cloned using the vector pBR322 (Bolivar *et al.*, 1977) and *E. coli* RR28 (Hennecke *et al.*, 1982). Transformants were screened by Southern hybridization of prepared plasmid DNA to the *E. coli* probe. pAC30 contains a 5 kb DNA fragment from *Cytophaga lytica*, pAB16 contains a 4.1 kb DNA fragment from *Bacteroides fragilis*.

DNA sequencing. Following restriction analysis of the clones, subclones were constructed for DNA sequencing using pUC vectors (Yanisch-Perron *et al.*, 1985). Subclones used for sequencing are listed in Table 1. Nucleotide sequences were determined by the chain termination method (Sanger *et al.*, 1977; Chen & Seeburg, 1985) using [α -³²P]dATP (New England Nuclear) and M13 universal sequencing primers (BRL).

Construction of a distance-matrix tree. The amino acid sequences were derived from the nucleotide sequences and were aligned for optimal similarity. Weighing only positions present in all sequences, a distance matrix containing the numbers of amino acid replacements was established. From these data the tree was constructed according to Fitch & Margoliash (1967) using the Phylip computer program from J. Felsenstein (University of Washington, Seattle, USA, personal communication). The mean percent standard deviation for this was 2.591. The lengths of the branches in the trees reflect the phylogenetic distances. The branching topography has been proven by using the protein parsimony (Protpar) program of J. Felsenstein.

RESULTS AND DISCUSSION

We recently demonstrated that the gene encoding the β -subunit of the ATPase from *Flavobacterium ferrugineum* exhibits an additional 99 bases as insertions at two different locations hitherto not found in corresponding genes from other eubacteria (Amann *et al.*, 1988). In the present study, we cloned and sequenced the genes encoding the β -subunit of ATPase from *Bacteroides fragilis* DSM 2151, *Cytophaga lytica* DSM 2039 and '*Taxebacter ocellatus*' (Reichenbach, Braunschweig, FRG; data not shown). According to 16S rRNA cataloging all these organisms belong to the so-called bacteroides-cytophaga-flavobacterium phylum (Woese, 1987).

Fig. 1 shows an alignment of the nucleotide sequences of the β -subunit genes of *Bacteroides fragilis* and *Cytophaga lytica* and of their predicted amino sequences in comparison to that of *Flavobacterium ferrugineum* and *E. coli*. The nucleotide sequences of the β -subunits of the two latter organisms have been published previously (Amann *et al.*, 1988; Saraste *et al.*, 1981). The sequences of the organisms belonging to the bacteroides-cytophaga-flavobacterium group are longer than their *E. coli* counterpart. In particular, they contain a large insertion (position 631–712, Fig. 1), and a small insertion (position 808–822) in the same regions when compared to those of *E. coli* and other eubacteria. Hydrophilicity profiles (data not shown) are very similar, indicating the insertions represent rather hydrophilic regions. The insertions cluster within or between the postulated nucleotide binding regions without changing their conserved sequences (Kagawa *et al.*, 1986; Curtis, 1987). The hydrophilic nature of the insertions is an indication that these regions may occur on the surface of the protein. The lengths of the two insertions (27 and 5 amino acids, respectively) are identical within the β -subunit genes of *Bacteroides*, *Cytophaga* and

Table 1. Subclones constructed from *pAB16* and *pAC30*

Organism	Subclone	Position of the β -subunit gene fragment*†	Cloned fragment	
			Restriction sites	
			5'	3'
<i>Bacteroides fragilis</i>	pAB16 β 1	(-100)-1242	<i>EcoRV</i>	<i>EcoRV</i>
	pAB16 β 2	(-100)-515	<i>EcoRV</i>	<i>SacI</i>
	pAB16 β 3	(-100)-308	<i>EcoRV</i>	<i>ClaI</i>
	pAB16 β 4	309-848	<i>ClaI</i>	<i>ClaI</i>
	pAB16 β 5	849-(1724)	<i>ClaI</i>	<i>HindIII</i>
	pAB16 β 6	1243-(1724)	<i>EcoRV</i>	<i>HindIII</i>
	pAB16 β 7	1128-(1724)	<i>BamHI</i>	<i>HindIII</i>
<i>Cytophaga lytica</i>	pAC30 β 1	(-370)-1147	<i>XbaI</i>	<i>XbaI</i>
	pAC30 β 2	1148-(3518)	<i>XbaI</i>	<i>XbaI</i>
	pAC30 β 3	(-110)-1147	‡	<i>XbaI</i>
	pAC30 β 4	(-370)-379	<i>XbaI</i>	‡
	pAC30 β 5	(-370)-659	<i>XbaI</i>	‡
	pAC30 β 6	(-370)-1009	<i>XbaI</i>	‡
	pAC30 β 7	1383-(3518)	<i>EcoRV</i>	<i>XbaI</i>
	pAC30 β 8	(-370)-1038	<i>XbaI</i>	<i>BamHI</i>
	pAC30 β 9	(-370)-858	<i>XbaI</i>	‡

* Nucleotide numbering according to Fig. 1. Positions of nucleotides upstream of the gene are indicated by a minus.

† Positions of nucleotides upstream or downstream of the β -subunit gene were not exactly determined and therefore put in parentheses.

‡ Subclone pAC30 β 1 shortened by treatment with exonuclease III.

Flavobacterium (Fig. 1). In the case of the large insertion, remarkable sequence similarity can be seen (44.4–70.4%). The same characteristic insertions were also found in the β -subunit gene of '*Taxebacter ocellatus*' (data not shown). The sequence idiosyncrasies in the β -subunit genes together with the high degree of sequence similarity of the remaining structure support the previous proposal that the aerobic *Flavobacterium*, *Cytophaga* and '*Taxebacter*' form a phylogenetically coherent unit with the anaerobic *Bacteroides*.

Previously published amino acid sequences of β -subunits of ATPase from various eubacteria, mitochondria and chloroplasts (Amann *et al.*, 1988; Runswick & Walker, 1983; Shinozaki *et al.*, 1983) were aligned with our data and a phylogenetic (distance matrix) tree was constructed (Fig. 2). This tree resembles those derived from 16S rRNA sequence (Woese, 1987) or oligonucleotide analyses. A clear separation of three eubacterial phyla (purple bacteria, cyanobacteria, and bacteroides plus relatives) can be seen. The mitochondrial sequence clusters with the purple bacteria. The latter exhibit a deep branching between representatives of the alpha- (*Rhodospirillum rubrum*) and the gamma-subgroups (*E. coli*, *Enterobacter aerogenes*). The chloroplast gene clusters with the cyanobacteria. The Gram-positive spore-forming thermophilic strain PS3 shows a deep branching within the cyanobacterial cluster. *Flavobacterium ferrugineum* and *Cytophaga lytica* group together somewhat distant from *Bacteroides fragilis*, as is also seen from rRNA data (Woese, 1987).

Phylogenetic trees were also derived from analysis of the nucleotide sequences. They were very similar to that derived from the amino acid sequences. Slight differences in the tree topology were mainly caused by differences in the G + C content of the various genes.

Our study shows that comparative sequence analysis of genes of β -subunits of ATPases provides valuable information on the phylogenetic relationship of bacteria. The β -subunit genes of ATPase are ubiquitous among eubacteria, the gene product is functionally constant and the sequence is highly conserved. Moreover, it can have sequence idiosyncrasies that are only found

Fig. 1. Alignment of four ATPase β -subunit sequences. Line 1, nucleotide numbering; lines 2 and 4, nucleotide sequences; lines 3 and 5–7, amino acid sequences. Gaps are indicated by dashes, stop codons by asterisks. Bfr, *Bacteroides fragilis*; Cly, *Cytophaga lytica*; Eco, *Escherichia coli*; Ffe, *Flavobacterium ferrugineum*.

Fig. 2. Phylogenetic tree derived from comparative analysis of the β -subunit of ATPases. The lengths of the internodes correspond to the average number of amino acid replacements. The bar at the bottom represents 10 amino acid replacements (10 AR). Ana, *Anabaena* sp. strain PCC 7120; Bfr, *Bacteroides fragilis*; bMit, bovine heart mitochondria; Cly, *Cytophaga lytica*; Eae, *Enterobacter aerogenes*; Eco, *Escherichia coli*; Ffe, *Flavobacterium ferrugineum*; PS3, Gram-positive spore-forming thermophilic strain PS3; Rbl, *Rhodopseudomonas blastica*; Rru, *Rhodospirillum rubrum*; Syn, *Synechococcus* sp. strain 6301; tCh, tobacco chloroplasts.

in a distinct line of descent. The β -subunit of ATPase also has the advantage that it is functionally unrelated to rRNA. If analyses of such unrelated macromolecules yield similar phylogenetic trees, it is a strong indication that the trees reflect the evolution not only of certain macromolecules but of the organisms themselves.

This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

REFERENCES

- AMANN, R., SOSTAK, P., LUDWIG, W. & SCHLEIFER, K. H. (1988). Cloning and sequencing of genes encoding the beta subunits of the ATP-synthases of *Enterobacter aerogenes* and *Flavobacterium ferrugineum*. *FEMS Microbiology Letters* **50**, 101–106.
- BOLIVAR, F., RODRIGUEZ, R., GREENE, P. J., BETLACH, M., HEYNECKER, H. L., BAYER, H. W., CROSA, J. & FALKOW, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2**, 95–113.
- CHEN, E. Y. & SEEBURG, P. H. (1985). Supercoiled sequencing: a fast and simple method for sequencing plasmid DNA. *DNA* **4**, 165–170.
- COZENS, A. L. & WALKER, J. E. (1987). The organization and sequence of the genes for ATP-synthase subunits in the cyanobacterium *Synechococcus* 6301. *Journal of Molecular Biology* **194**, 359–383.
- CURTIS, S. E. (1987). Genes encoding Beta and Epsilon subunits of the proton-translocating ATPase from *Anabaena* sp. strain PCC 7120. *Journal of Bacteriology* **169**, 80–86.
- FALK, G., HAMPE, A. & WALKER, J. E. (1985). Nucleotide sequence of the *Rhodospirillum rubrum* *atp* operon. *Biochemical Journal* **228**, 391–407.
- FITCH, W. M. & MARGOLIAS, E. (1967). Construction of phylogenetic trees. *Science* **155**, 279–284.

- FUTAI, M. & KANAZAWA, H. (1983). Structure and function of proton-translocating adenosine triphosphatase (F_0F_1): biochemical and molecular biological approach. *Microbiological Reviews* **47**, 285–312.
- HENNECKE, H., GÜNTHER, I. & BINDER, F. (1982). A novel cloning vector for the direct selection of recombinant DNA in *Escherichia coli*. *Gene* **19**, 231–234.
- KAWAGA, Y., ISHIZUKA, M., SAISHU, T. & NAKAO, S. (1986). Stable structure of thermophilic proton ATPase beta-subunit. *Journal of Biochemistry* **100**, 923–934.
- MARMUR, J. (1961). A procedure for the isolation of DNA from microorganisms. *Journal of Molecular Biology* **3**, 208–218.
- MEYER, S. A. & SCHLEIFER, K. H. (1978). Deoxyribonucleic acid reassociation in the classification of coagulase-positive staphylococci. *Archives of Microbiology* **117**, 183–188.
- PASTER, B. J., LUDWIG, W., WEISBURG, W. G., STACKEBRANDT, E., HESPELL, R. B., HAHN, C. M., REICHENBACH, H., STETTER, K. O. & WOESE, C. R. (1985). A phylogenetic grouping of the bacteroides, cytophagas and certain flavobacteria. *Systematic and Applied Microbiology* **6**, 34–42.
- RUNSWICK, M. J. & WALKER, J. E. (1983). The amino acid sequence of the beta subunit of ATP-synthase from bovine heart mitochondria. *Journal of Biological Chemistry* **258**, 9307–9313.
- SANGER, F., NICKLER, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SARASTE, M., GAY, N. J., EBERLE, A., RUNSWICK, M. J. & WALKER, J. E. (1981). The *atp* operon: nucleotide sequence of the genes for the γ , β and ϵ subunits of *Escherichia coli* ATP-synthase. *Nucleic Acids Research* **9**, 5287–5296.
- SHINOZAKI, K., YANADA, C., TAKAHATA, N. & SUGIURA, M. (1983). Overlap and cotranscription of the genes for the beta and epsilon subunits of tobacco chloroplast ATPase. *Gene* **24**, 147–155.
- SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.
- TYBULEWICZ, V. L. J., FALK, G. & WALKER, J. E. (1984). *Rhodopseudomonas blastica atp* operon: nucleotide sequence and transcription. *Journal of Molecular Biology* **178**, 185–214.
- WALKER, J. E., SARASTE, M. & GAY, N. J. (1984). The *unc* operon: nucleotide sequence, regulation and structure of ATP-synthase. *Biochimica et biophysica acta* **768**, 164–200.
- WEISBURG, W. G., OYAIZU, Y., OYAIZU, H. & WOESE, C. R. (1985). Natural relationship between *Bacteroides* and *Flavobacteria*. *Journal of Bacteriology* **164**, 230–236.
- WOESE, C. R. (1987). Bacterial evolution. *Microbiological Reviews* **51**, 221–271.
- YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.
- ZILBERSTEIN, D., SHIRVAN, M. H., BARILE, M. F. & ROTTEM, S. (1986). The β -subunit of the F_1F_0 -ATPase is conserved in mycoplasmas. *Journal of Biochemistry* **261**, 7109–7111.