

**Characterization of *Aquifex aeolicus* F₁F₀
ATP synthase and its heterologous
production in *Escherichia coli***

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften

vorgelegt beim Fachbereich 14 Biochemie, Chemie und Pharmazie der Johann
Wolfgang Goethe Universitaet in Frankfurt am Main, Deutschland

Von Chunli ZHANG
aus Qingdao, China

Frankfurt am Main (2013)
(D30)

vom Fachbereich Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe
Universitaet als Dissertation angenommen.

Dekan: Prof. Dr. Thomas Prisner

1. Gutachter: Prof. Dr. Bernd Ludwig
2. Gutachter: Prof. Dr. Hartmut Michel

Datum der Disputation:

Diese Doktorarbeit wurde vom 17. März 2009 bis zum 24. Juni 2013 unter Leitung von Prof. Dr. Hartmut Michel in der Abteilung für Molekulare Membranbiologie am Max-Planck-Institut für Biophysik in Frankfurt am Main durchgeführt.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe und keine weiteren Hilfsmittel und Quellen als die hier aufgeführten verwendet habe.

Chunli Zhang

Frankfurt am Main

To my beloved daughters:

Wendi and Yadi

Publications

C. Zhang, M. Marcia, J. D. Langer, G. Peng and H. Michel (2013). "Role of the N terminal signal peptide in the membrane insertion of *Aquifex aeolicus* F₁F₀ ATP synthase c-subunit." FEBS J **280**(14): 3425-3435.

C. Zhang, M. Allegretti, J. Vonck, J. D. Langer, M. Marcia, G. Peng and H. Michel "Production of fully assembled and active *Aquifex aeolicus* F₁F₀ ATP synthase in *Escherichia coli*". Biochim Biophys Acta **1840**(1): 34-40.

Table of content

Table of content	i
List of Figures	v
List of Tables	vii
Zusammenfassung	ix
Detailed English Summary	xv
Abstract	xxi
Abbreviations	xxiii
1. Introduction	1
1.1. Electron transport and oxidative phosphorylation.....	2
1.2. F₁F₀ ATP synthase (Respiratory complex V)	3
1.2.1. Classification, nomenclature and architecture	4
1.2.2. Subunit composition and sequence conservation	6
1.2.3. The soluble F ₁ subcomplex	7
1.2.3.1. Function of F ₁ subcomplex.....	7
1.2.3.2. Structure of F ₁ subcomplex.....	10
1.2.4. The membrane-embedded F ₀ subcomplex.....	12
1.2.4.1. Function of F ₀ subcomplex	12
1.2.4.2. Structure of F ₀ subcomplex.....	13
1.2.5. ATP synthase biogenesis and assembly	18
1.3. <i>Aquifex aeolicus</i> ATP synthase	20
1.4. Open questions	21
1.5. Aim of this work.....	22
2. Results.....	23
2.1. Bioinformatic and functional characterization of AAF₁F₀.....	23
2.1.1. Bioinformatic characterization	23
2.1.1.1. Sequence conservation	23
2.1.1.2. Subunit a	24
2.1.1.3. Subunits b ₁ and b ₂	26
2.1.1.4. Subunit c	27
2.1.2. Functional characterization of native AAF ₁ F ₀	27
2.1.2.1. Purification of AAF ₁ F ₀	27
2.1.2.2. AAF ₁ F ₀ is H ⁺ -dependent, not Na ⁺ -dependent.....	28
2.1.2.3. The ATP hydrolysis activity of AAF ₁ F ₀ is comparable with the activity of other respiratory enzymes from <i>A. aeolicus</i>	30
2.2. Generation of a heterologous system to produce <i>A. aeolicus</i> ATP synthase	30
2.2.1. Gene composition.....	30
2.2.2. Heterologous expression strategy.....	31
2.2.3. Single-gene and dual-gene expression	32
2.2.3.1. Strategy.....	33
2.2.3.2. Cloning	34
2.2.3.3. Subcomplex b ₁ b ₂ : an <i>A. aeolicus</i> native operon with overlapping genes can be recognized by <i>E. coli</i>	34
2.2.3.4. Membrane subunits a and c cannot be expressed individually in <i>E. coli</i>	37
2.2.4. Expression of subcomplexes F ₁ -αβγ and F ₁ -αβγϵ from artificial operons	38
2.2.4.1. Subcomplex F ₁ -αβγ	38

2.2.4.2.	Subcomplex F ₁ -αβγε.....	39
2.2.5.	Production of the entire EAF ₁ F _O complex.....	42
2.2.5.1.	Construction of the artificial operon: DNA amplification, manipulation and subcloning.....	42
2.2.5.2.	The entire ATP synthase complex: the fully assembled His ₆ -EAF ₁ F _O purified from the membranes of <i>E. coli</i>	45
2.2.6.	EAF ₁ F _O catalyzes ATP hydrolysis at the same rate of AAF ₁ F _O	50
2.2.7.	The structure of EAF ₁ F _O is identical to that of AAF ₁ F _O	51
2.3.	Use of EAF₁F_O to study unique properties of <i>A. aeolicus</i> ATP synthase.....	51
2.3.1.	Characterization of the N-terminus of <i>A. aeolicus</i> ATP synthase subunit c.....	51
2.3.2.	Subunits c can be subdivided into four classes based on their sequences.....	52
2.3.3.	The N-terminal region of the subunits c is not conserved.....	53
2.3.4.	The mature form of the native subunit c of <i>A. aeolicus</i> is 81 amino acids long.....	55
2.3.5.	The N-terminal signal peptide of <i>A. aeolicus</i> subunit c is recognized by <i>E. coli</i> and is crucial for membrane insertion of subunit c.....	56
3.	Discussion.....	59
3.1.	Producing proteins from thermophilic organisms in mesophilic hosts.....	59
3.2.	Strategy for the heterologous production of <i>A. aeolicus</i> ATP synthase in <i>E. coli</i>.....	59
3.3.	Properties of the artificial <i>atp</i> operon.....	60
3.3.1.	Composition of the <i>atp</i> operon genes.....	60
3.3.2.	Regulation of expression: translation initiation regions (TIR).....	61
3.3.2.1.	Intergenic regions.....	63
3.3.2.2.	Ribosome binding sites (RBS).....	63
3.3.2.3.	Start and stop codons.....	64
3.3.3.	Gene overlaps.....	64
3.3.4.	Codon usage optimization.....	65
3.3.5.	Position of purification tag.....	66
3.4.	The structure and function of heterologously expressed <i>A. aeolicus</i> F₁F_O ATP synthase.....	66
3.5.	Novel properties of <i>A. aeolicus</i> ATP synthase discovered using EAF₁F_O.....	67
3.5.1.	A new phylogenetic classification of F ₁ F _O ATP synthase based on the N-terminal sequence of their subunit c.....	69
3.5.2.	A revised assembly pathway of F ₁ F _O ATP synthase.....	73
3.6.	Conclusions and future perspectives.....	74
4.	Material and Methods.....	77
4.1.	Material.....	77
4.1.1.	Chemicals.....	77
4.1.2.	Organisms.....	77
4.1.2.1.	<i>A. aeolicus</i>	77
4.1.2.2.	<i>E. coli</i> strains.....	77
4.1.3.	Plasmids.....	78
4.1.4.	Bacterial media and solutions.....	80
4.1.5.	Enzymes, proteins, markers and kits.....	82
4.1.6.	Antibodies.....	84
4.1.7.	Chromatographic columns and matrices.....	84
4.1.8.	Databases, servers and software.....	85
4.2.	Methods.....	86

4.2.1.	Bioinformatics.....	86
4.2.2.	Molecular Biology.....	87
4.2.2.1.	Isolation of genomic DNA from <i>A. aeolicus</i>	87
4.2.2.2.	Isolation of plasmid DNA.....	87
4.2.2.3.	DNA amplification.....	87
4.2.2.4.	Agarose Gel Electrophoresis of DNA.....	89
4.2.2.5.	Quantification of nucleic acids.....	90
4.2.2.6.	Digestion of DNA with restriction endonucleases.....	90
4.2.2.7.	Ligation.....	91
4.2.2.8.	Preparation of chemically competent cells.....	92
4.2.2.9.	Transformation of competent cells.....	93
4.2.2.10.	Screening of positive transformants.....	93
4.2.2.11.	DNA sequencing.....	94
4.2.2.12.	Site-directed mutagenesis.....	94
4.2.2.13.	Storage of <i>E. coli</i> strains.....	95
4.2.3.	Protein expression and isolation.....	95
4.2.3.1.	Rapid expression screening.....	95
4.2.3.2.	Preparative protein purification.....	96
4.2.3.3.	Determination of protein cellular localization.....	96
4.2.3.4.	Membrane preparation.....	97
4.2.3.4.1.	Small-scale membrane preparation.....	97
4.2.3.4.2.	Large-scale membrane preparation.....	97
4.2.3.5.	Detergent screening for solubilization of membrane proteins.....	98
4.2.3.6.	Protein overproduction and purification.....	98
4.2.3.6.1.	Subcomplex b_1b_2	98
4.2.3.6.2.	Subcomplexes $F_1\text{-}\alpha\beta\gamma$ and $F_1\text{-}\alpha\beta\gamma\epsilon$	99
4.2.3.6.3.	The whole EAF_1F_0 complex.....	99
4.2.3.7.	Protein purification.....	100
4.2.3.7.1.	Affinity purification.....	100
4.2.3.7.2.	Size-exclusion chromatography (SEC).....	100
4.2.3.7.3.	Ion exchange chromatography.....	101
4.2.4.	Protein characterization by biochemical methods.....	101
4.2.4.1.	Determination of protein concentration.....	101
4.2.4.2.	Denaturing gel electrophoresis (SDS-PAGE).....	102
4.2.4.3.	Non denaturing gel electrophoresis (Native PAGE).....	103
4.2.4.4.	Electro-elution.....	104
4.2.4.5.	Two-dimensional electrophoresis (2-D Native/SDS-PAGE).....	104
4.2.4.6.	Gel staining.....	105
4.2.4.7.	Antibody production.....	106
4.2.4.8.	Western blot analysis.....	107
4.2.4.9.	Dot blot analysis.....	108
4.2.4.10.	Colony blot analysis.....	108
4.2.5.	Protein identification by mass spectrometry.....	109
4.2.5.1.	Preparation of subunit c monomers from native <i>A. aeolicus</i> F_1F_0 ATP synthase.....	109
4.2.5.2.	DCCD labeling assay.....	110
4.2.5.3.	Peptide mass fingerprinting (PMF).....	110
4.2.5.4.	MALDI-TOF-MS measurements.....	111
4.2.6.	Enzymatic activity assays.....	111
4.2.6.1.	ATP hydrolysis activity assay by phosphate determination.....	111
4.2.6.2.	In-gel ATP hydrolysis assays.....	112
4.2.6.3.	Preparation of inverted membrane vesicles.....	112

Table of content

4.2.7. Single particle electron microscopy.....	112
4.2.8. Protein crystallization.....	113
References.....	115
Appendix.....	xxv
A.1. Bioinformatics.....	xxv
A.2. Codon usage.....	xxix
A.3. Plasmid maps and DNA sequences.....	xxxiv
A.4. Antibodies generation.....	liv
Acknowledgements.....	lxi
Curriculum Vitae.....	lxiii

List of Figures

Figure 1.1.	Adenosine triphosphate (ATP).....	1
Figure 1.2.	Schematic overview of the respiratory chain complexes.....	3
Figure 1.3.	Architecture and subunit composition of bacterial F ₁ F ₀ ATP synthase	5
Figure 1.4.	The Boyer binding-change mechanism.....	7
Figure 1.5.	Rotation of subunit γ relative to the hexamer $\alpha_3\beta_3$	8
Figure 1.6.	Crystal structure of the F ₁ subcomplex $\alpha_3\beta_3\gamma$ from bovine mitochondrial ATP synthase	11
Figure 1.7.	The two conformational states of subunit ϵ	12
Figure 1.8.	Proposed half-channel model for proton translocation in <i>E. coli</i>	13
Figure 1.9.	Proposed structure of <i>E. coli</i> subunit b	14
Figure 1.10.	The interface between rotor and stator in F ₀ -ATP synthase.....	15
Figure 1.11.	c-ring structures solved by X-ray crystallography.....	16
Figure 1.12.	Structural details of the ion binding sites of Na ⁺ - and H ⁺ -dependent ATP synthases	17
Figure 1.13.	Assembly pathway for F ₁ F ₀ ATP synthase	19
Figure 1.14.	Organization of the F ₁ F ₀ ATP synthase genes in <i>A. aeolicus</i>	20
Figure 2.1.	Sequence alignment of subunit a of F ₁ F ₀ ATP synthases.....	25
Figure 2.2.	Sequence alignment of subunit a of V-type ATP synthases	25
Figure 2.3.	Sequence alignment of subunit b of F ₁ F ₀ ATP synthases	26
Figure 2.4.	Characterization of AAF ₁ F ₀	27
Figure 2.5.	Blue-native PAGE gel of AAF ₁ F ₀ in the presence of different detergents	28
Figure 2.6.	Sequence alignment of subunit c of F ₁ F ₀ ATP synthases.....	29
Figure 2.7.	DCCD-labeling of the subunit c of AAF ₁ F ₀	29
Figure 2.8.	The ATP hydrolysis activity of AAF ₁ F ₀ at different temperatures.....	30
Figure 2.9.	The EAF ₁ F ₀ expression strategy.....	32
Figure 2.10.	Single-gene and dual-gene expression vectors	33
Figure 2.11.	The scheme of <i>A. aeolicus</i> native operon <i>atpF1F2</i> encoding subunits b ₁ and b ₂	34
Figure 2.12.	Expression tests on subcomplex b ₁ b ₂	35
Figure 2.13.	Subunits b ₁ / b ₂ are expressed in <i>E. coli</i> membranes	36
Figure 2.14.	Subunits b ₁ and b ₂ form a subcomplex	37
Figure 2.15.	Subunits b ₁ and b ₂ were identified by PMF.	37
Figure 2.16.	Subcomplex F ₁ - $\alpha\beta\gamma$	38
Figure 2.17.	Identification of subunits α , β and γ by PMF.....	39
Figure 2.18.	Subcomplex F ₁ - $\alpha\beta\gamma\epsilon$	40
Figure 2.19.	Identification of subunits α , β , γ and ϵ by PMF	41
Figure 2.20.	SEC calibration curves.....	41
Figure 2.21.	Scheme of the initial steps of construction of the artificial <i>atp</i> operon	43
Figure 2.22.	Scheme for the construction of intermediate expression vectors pCL11, pCL12, and pCL02 and of the expression vector pCL21 containing the whole <i>atp</i> operon.....	44
Figure 2.23.	The artificial <i>atp</i> operon for heterologous product of EAF ₁ F ₀	45

List of Figures

Figure 2.24.	Expression tests for EAF ₁ F ₀	46
Figure 2.25.	The fully assembled EAF ₁ F ₀ purified from membranes of <i>E. coli</i>	47
Figure 2.26.	Silver-stained SDS-PAGE of F ₁ -c _n subcomplex electro-eluted from BN-PAGE.....	47
Figure 2.27.	Mass-spectrum of subunits a and c.....	48
Figure 2.28.	Identification of all subunits of EAF ₁ F ₀ by PMF.....	49
Figure 2.29.	Functional characterization of EAF ₁ F ₀ by <i>in-gel</i> ATP-hydrolysis activity at 80°C.....	50
Figure 2.30.	Structural characterization of EAF ₁ F ₀ by single-particle EM.....	51
Figure 2.31.	Subunits c cluster in four groups	54
Figure 2.32.	Topology prediction of the N-terminal region of group 2 subunits c.....	54
Figure 2.33.	The subunit c of <i>A. aeolicus</i> F ₁ F ₀ ATP synthase possesses an N-terminal signal peptide as indicated by mass spectrometry	55
Figure 2.34.	Production of the F ₁ subcomplex in <i>E. coli</i>	57
Figure 2.35.	Identification of subunit c monomers in the membranes of <i>E. coli</i>	58
Figure 3.1.	Effect of codon usage on heterologous expression of <i>atp</i> genes in <i>E. coli</i>	65
Figure 3.2.	Silver-stained SDS-PAGE of AAF ₁ F ₀ with (+) or without (-) heat treatment.....	68
Figure 3.3.	Multiple-sequence alignments of subunit c.....	70
Figure 3.4.	Phylogenetic clusters of subunit c	71
Figure 3.5.	Membrane insertion pathways proposed for the 4 different groups of subunit c.....	72
Figure 3.6.	Subunit b ₂ was identified by PMF in the subunit c-deletion mutant pCL21-MEN.....	73
Figure 3.7.	Proposed mechanism for the assembly of subcomplex F ₁ on subunit b in the <i>E. coli</i> membranes.	74
Figure A.1.	Sequence alignment of subunit α of F ₁ F ₀ ATP synthases and V ₁ V ₀ ATPases.....	xxvii
Figure A.2.	Sequence alignment of subunit β of F ₁ F ₀ ATP synthases and V ₁ V ₀ ATPases.....	xxviii
Figure A.3.	Codon usage difference between <i>A. aeolicus</i> and <i>E. coli</i> analyzed by Graphical Codon Usage Analyser (GCAU).....	xxix
Figure A.4.	Codon usage in subunit c.....	xxx
Figure A.5.	Codon usage in subunit β	xxxi
Figure A.6.	Plasmid map of vector pCL21	xxxiv
Figure A.7.	Plasmid map of vector pCL01	xxxiv
Figure A.8.	Plasmid map of vector pCL02.....	xxxv
Figure A.9.	Plasmid map of vector pCL11	xxxv
Figure A.10.	Plasmid map of vector pCL12.....	xxxvi
Figure A.11.	Plasmid map of empty vector pTrc99A.....	xxxvi
Figure A.12.	Antigen profile of subunit α	liv
Figure A.13.	Antigen profile of subunit γ.....	lv
Figure A.14.	Antigen profile of subunit δ.....	lvi
Figure A.15.	Antigen profile of subunit ε.....	lvii
Figure A.16	Antigen profile of subunit c.....	lviii

List of Tables

Table 2.1:	Statistics describing the sequence alignment of selected ATP synthase subunits	23
Table 2.2:	Properties of genes and corresponding subunits of the <i>A. aeolicus</i> F ₁ F ₀ ATP synthase	31
Table 2.3:	Properties of the four groups of subunits c	52
Table 3.2:	Length of TIRs selected for the <i>A. aeolicus atp</i> genes.....	62
Table 3.2:	Sequences of TIRs selected for <i>A. aeolicus atp</i> operon.....	63
Table 3.3:	Start codon and stop codon of the <i>A. aeolicus atp</i> genes.....	64
Table 4.1:	List of <i>E. coli</i> strains.....	77
Table 4.2:	List of empty vectors.....	78
Table 4.3:	List of vectors generated for the single or dual gene expression of subunits a, b ₁ b ₂ and c	79
Table 4.4:	List of vectors generated for the expression of the artificial <i>atp</i> operon.....	80
Table 4.5:	List of bacterial media.....	80
Table 4.6:	List of antibiotics.....	82
Table 4.7:	List of enzymes, proteins, markers and kits.....	83
Table 4.8:	List of antibodies.....	84
Table 4.9:	List of chromatographic columns and matrices.....	84
Table 4.10:	List of database and servers.....	85
Table 4.11:	List of software.....	86
Table 4.12:	List of primers used in this work.....	88
Table 4.13:	Typical PCR reaction mixture.....	89
Table 4.14:	Typical PCR cycling program.....	89
Table 4.15:	List of buffers for agarose gel electrophoresis.....	90
Table 4.16:	Typical double digestion reaction.....	91
Table 4.17:	Typical ligation reaction.....	92
Table 4.18:	List of buffers used to prepare chemically competent cells.....	93
Table 4.19:	List of primers used for sequencing the artificial <i>atp</i> operon.....	94
Table 4.20:	List of detergents used for solubilization screens.....	98
Table 4.21:	List of buffers and chemicals used for casting the 13.2% tricine gels.....	102
Table 4.22:	List of solutions used for SDS-PAGE electrophoresis.....	103
Table 4.23:	List of buffers used for BN-PAGE.....	104
Table 4.24:	List of buffers used for 2D Native/SDS-PAGE.....	105
Table 4.25:	List of buffers used for Coomassie blue staining.....	105
Table 4.26:	List of solutions used for silver staining.....	106
Table 4.27:	List of solutions used for Western blot analysis.....	108
Table 4.28:	List of solutions used for colony-blot analysis.....	109
Table 4.29:	LeBel-reagent solutions.....	112
Table A.1:	Amino acid compositions of the <i>A. aeolicus</i> F ₁ F ₀ ATP synthase deduced from the respective <i>atp</i> genes.....	xxv
Table A.2:	List of the c-subunits used for multiple-sequence alignments.....	xxvi

List of tables

Table A.3: 70 day rabbit immunization protocol against synthesized peptide (from Thermo Fisher)..... lix

Zusammenfassung

Die F_1F_0 ATP-Synthase katalysiert die Synthese von ATP aus ADP und anorganischem Phosphat. Die hierfür benötigte Energie wird durch einen über die Zellmembran bzw. Innere Mitochondrienmembran bestehenden elektrochemischen Ionengradienten geliefert. Die F_1F_0 ATP-Synthase ist sowohl in Bakterien, als auch in Mitochondrien und Chloroplasten zu finden und dabei hoch konserviert.

Das Holoenzym besteht aus zwei größeren Subkomplexen, dem hydrophilen F_1 - und dem hydrophoben F_0 -Komplex. Der F_1 -Subkomplex besteht aus den Untereinheiten α , β , γ , δ und ϵ in der Zusammensetzung 3:3:1:1:1. Der membrangebundene F_0 -Komplex besteht aus den Untereinheiten a, b und c, in den Stöchiometrien 1:2: (8-15). Die Untereinheiten a und c sind für die Ionentranslokation zuständig. Die Untereinheiten γ und ϵ verbinden den F_1 -Subkomplex mit dem c-Ring des F_0 -Subkomplexes.

Neben der Einteilung in den hydrophoben und hydrophilen Teil des Enzyms kann die F_1F_0 ATP-Synthase auch in einen Stator (a, b, δ , α , β) und in einen Rotor (γ , ϵ , c) gegliedert werden. Bisher konnten atomare Strukturen nur von Subkomplexen oder einzelnen Untereinheiten bestimmt werden, wie zum Beispiel dem bovinen F_1 -Subkomplex oder den c-Ringen aus *Ilyobacter tartaricus*, *Bacillus pseudofirmus* und *Arthrospira platensis*. Strukturen für den bovinen Stator-Subkomplex konnten ebenfalls bestimmt werden.

Allerdings ist noch keine Struktur für das Holoenzym oder den membrangebundenen F_0 -Subkomplex bekannt. Die Struktur des Holoenzym in atomarer Auflösung könnte detaillierte Einblicke in den Ionen-Transportmechanismus geben, der bis heute nicht komplett geklärt ist.

Bisher konnte in unserem Labor gezeigt werden, dass die F_1F_0 ATP-Synthase aus *Aquifex aeolicus* aufgrund ihrer thermophilen Herkunft ein hoch stabiles Enzym darstellt, das als Holoenzym in seiner aktiven Form aufgereinigt werden konnte. Zusätzlich konnten neue strukturelle Daten für die F_1F_0 ATP-Synthase gewonnen werden, wie etwa eine Deformation des zentralen Stators (γ und ϵ -Untereinheiten) oder ein möglicher heterodimerer peripherer Stator im Vergleich zum Rinderenzym. Daher stellt die F_1F_0 ATP-Synthase aus *A. aeolicus* ein interessantes Ziel für weitere strukturelle und funktionelle Studien dar.

Es wurden vier Ziele für diese Doktorarbeit formuliert, basierend auf früheren Studien: (i) Ergänzung der bisherigen Charakterisierung der nativen F_1F_0 ATP-Synthase aus *A. aeolicus* (AAF $_1F_0$) durch bioinformatische, biochemische und funktionelle Studien, (ii) Etablierung eines

heterologen Expressionssystem für AAF₁F₀ in *E. coli*, (iii) Charakterisierung der so exprimierten ATP-Synthase (EAF₁F₀), (iv) Untersuchung von Eigenschaften der AAF₁F₀, wie etwa die Rolle des N-Terminus der c-Untereinheit, die nur mit Hilfe eines heterologen Expressionssystems durchgeführt werden können.

1) Charakterisierung der native *A. aeolicus* F₁F₀ ATP-Synthase (AAF₁F₀)

Durch den Einsatz bioinformatischer Methoden, wie dem *Multiple-Sequence Alignment*, der membranständigen Untereinheiten der F₁F₀ ATP-Synthase konnte gezeigt werden, dass: (i) die a-Untereinheit statt sechs nur fünf Membran durchspannende Helices hat; (ii) sowohl die b₁ als auch die b₂-Untereinheit (und nicht nur die b₁-Untereinheit) eine membraninsertierte N-terminale Helix besitzt. Allerdings besitzt nur die b₂-Untereinheit eine putative Signalsequenz vor der N-terminalen Helix, die im Laufe des Maturationsprozesses entfernt werden könnte; (iii) die c-Untereinheit eine bezüglich Hydrophobizität und Länge veränderte N-terminale Region besitzt. Dies kann Konsequenzen für die Membraninsertion und Assemblierung der ATP-Synthase haben.

Neben bioinformatischen Analysen wurde in dieser Arbeit ebenfalls eine Aufreinigung und biochemische Charakterisierung der AAF₁F₀ durchgeführt. Hierfür wurde das frühere Reinigungsprotokoll mit dem Ziel optimiert, die Stabilität des Holoenzym zu verbessern bzw. das Auseinanderbrechen in kleinere Subkomplexe zu verhindern. Die hauptsächlichlicher Optimierung des Protokolls bestand in der Verwendung des neuen Detergenzes *trans*-4-(*trans*-4'-propylcyclohexyl)cyclohexyl- α -D-maltosid (α -PCC). Es zeigte sich, dass dieses Detergenz die Stabilität des Holoenzym deutlich verbesserte im Vergleich zu den im Vorfeld verwendeten Detergenzien n-Dodecyl- β -D-maltosid (DDM) und n-Decyl- β -D-maltosid (DM).

Nachdem dieses Protokoll optimiert wurde, folgten funktionelle Studien, um experimentell AAF₁F₀ als Protonen-abhängige (und nicht Natrium) ATP-Synthase zu charakterisieren, da bioinformatische Studien bereits zeigten, dass das benötigte Natrium-Bindemotif in der c-Untereinheit fehlt. MALDI-TOF massenspektrometrische Messungen ergaben, dass die c-Untereinheit der AAF₁F₀ nicht durch Natrium Ionen vor der Bindung des kovalent bindenden "active site" Liganden *N,N'*-dicyclohexyl-carbodiimid (DCCD) geschützt werden kann. Dieser Befund änderte sich auch unter nicht-physiologischen Natrium-Konzentrationen (150 mM) nicht, was typisch für Protonen-abhängige ATP-Synthasen ist.

Zusätzliche enzymatische Studien zeigten, dass die ATP Hydrolyse der AAF₁F₀ bei Temperaturen niedriger als 60°C vernachlässigbar gering ist. SDS-PAGE Analysen offenbarten, dass die γ -Untereinheit nur bei hohen Temperaturen von dem $\alpha\beta$ -Hexagon dissoziiert werden kann.

Einzelpartikel-Elektronenmikroskopische Untersuchungen zeigten eine leicht gebogene γ -Untereinheit des Enzyms aus *A. aeolicus* im Vergleich zu der γ -Untereinheit der meisten anderen Organismen. Diese strukturelle Besonderheit ist bisher nur von der ATP-Synthase aus *Caldkcalibacillus thermarum* TA2.A1 bekannt, in der sie genutzt wird, um durch die Ausbildung von Salzbrücken die Rotation der γ -Untereinheit bei Raumtemperatur zu unterbinden. Bei Temperaturen über 60°C ist diese Interaktion nicht mehr stark genug, um die Rotation der γ -Untereinheit zu verhindern.

Die Reinheit der AAF₁F₀ erlaubte ebenfalls die Produktion von polyklonalen Antikörpern gegen die Untereinheiten α , β , γ , ϵ und c. Diese wurden zu einem späteren Zeitpunkt genutzt, um das heterologe Expressionssystem zu etablieren.

2) Herstellung eines artifiziellen Operons für die heterologe Expression der ATP-Synthase aus *A. aeolicus*

Um funktionelle Eigenschaften von AAF₁F₀ gründlicher untersuchen zu können, wurde ein heterologes Expressionssystem in *E. coli* etabliert. Die Konstruktion eines Vektorsystems erwies sich als Herausforderung, da: (i) AAF₁F₀ ein heteromultimeres Enzym mit einem Molekulargewicht von mehr als 500 kDa und einer komplexen Stöchiometrie verschiedener Untereinheiten ist; (ii) die *atp*-Gene in *A. aeolicus* nicht in einem Operon organisiert, sondern über vier verschiedene Genloci verteilt sind. Die neun *atp*-Gene sind auf sechs verschiedene DNS-Fragmente verteilt, wobei einige Gene in ihrer Sequenz überlappen. *E. coli* wurde trotz der hyperthermophilen Herkunft der AAF₁F₀ als Wirt für die heterologe Expression ausgewählt, da es ein sehr gut untersuchter Modellorganismus darstellt, der schon für die Expression verschiedener ATP-Synthasen erfolgreich eingesetzt wurde.

Die gewählte Strategie umfasste folgende Schritte: (i) die Expression von Genen für einzelne Untereinheiten (α , c, γ und ϵ) sowie für spezifische Kombinationen von Genen (b_1 - b_2 , a-c, a- b_1 - b_2 , γ - ϵ); (ii) Klonierung der *atp*-Gene in verschiedene kleinere Operons und Expression in verschiedenen Vektoren. Untersucht wurde dabei auch der Einfluss verschiedener nativer Codons und die Fähigkeit von *E. coli*, überlappender Gene zu erkennen und ein funktionelles Holoenzym zu produzieren. Co-Transformation von *E. coli* mit dem Vektor pRARE (codierend für seltene t-RNAs) stellte sich als extrem wichtig heraus, um die Unterschiede in der Codon-Nutzung zwischen *E. coli* und *A. aeolicus* zu überwinden; (iii) Expression zweier kleinerer Subkomplexe (F₁- $\alpha\beta\gamma$ und F₁- $\alpha\beta\gamma\epsilon$) jeweils mit einem N-terminalen His₆-tag an der β -Untereinheit zum Zweck der Detektion und Aufreinigung. Ein dritter Subkomplex bestehend aus F₀-ac b_1b_2d wurde ebenfalls erstellt; und (iv) Klonierung aller nötigen Gene in einen einzelnen Expressionsvektor, der für die komplette

AAF₁F₀ kodierte. Die Gene der ATP-Synthase lagen in derselben Reihenfolge wie im *E. coli* Genom vor (abgesehen von *atpI*, welches in *A. aeolicus* fehlt). Die korrekte Stöchiometrie der einzelnen Untereinheiten ist kritisch für dieses Enzym und wird durch translationale Initiationsregionen (TIR) reguliert. Für die hier beschriebene Expression wurden die nativen TIR von *A. aeolicus* verwendet, inklusive 30 Basenpaare vor den Start-Codons der sechs Gene *atpB*, *atpE*, *atpF1*, *atpA*, *atpG* und *atpC*. Weiterhin wurden die Bereiche zwischen den Genen (um Restriktionsschnittstellen und Tags einzufügen) nur minimal verändert und die originale interzistronische Entfernungen zwischen benachbarten Genen beibehalten.

3) Charakterisierung der EAF₁F₀ und der beschriebenen Subkomplexe

Die während der Etablierung des heterologen Expressionssystems in *E. coli* gewonnenen Untereinheiten und Subkomplexe wurden ebenfalls biochemisch und funktionell charakterisiert, was zu folgenden Ergebnissen führte!

Die erfolgreiche Expression des b₁b₂-Subkomplexes ergab, dass: (i) *E. coli* das native Operon sowie die überlappenden Gene aus *A. aeolicus* erkennen kann, (ii) dass die Untereinheiten b₁ und b₂ in *E. coli* Membranen einen Komplex formen können, (iii) sowie dass dieser Komplex aufgereinigt werden kann und in Detergenz solubilisierter Form über längere Zeit stabil bleibt.

Die gefundene Komplexbildung zwischen den Untereinheiten b₁ und b₂ in *E. coli*-Membranen, unterstützt die Beobachtung, dass sie zu einem heterodimeren, peripheren Stator assoziieren können. Dies konnte bisher nur für photosynthetisch aktive Bakterien gezeigt werden. Im Gegensatz dazu konnten die Untereinheiten a und c nicht exprimiert werden. Dies wurde bereits für andere ATP-Synthasen berichtet.

Die erfolgreiche Expression und Aufreinigung der Subkomplexe F₁-αβγ und F₁-αβγε sowie deren biochemische Charakterisierung zeigte, dass diese Komplexe in katalytisch aktiver Form von einem artifiziellen Operon in *E. coli* exprimiert und aufgereinigt werden können. Alle Untereinheiten konnten durch Peptide Mass Fingerprinting (PMF) in Kombination mit Electro Spray Ionisation Mass Spectrometry (ESI-MS) nachgewiesen werden. Die Subkomplexe zeigten eine ATP-Hydrolyse Aktivität von $1,35 \pm 0,14$ U/mg und $1,73 \pm 0,11$ U/mg.

Rekombinante ATP-Synthase (EAF₁F₀) konnte in einer funktionell aktiven und vollständig assemblierten Form erhalten und aufgereinigt werden. Die Reinigung aus *E. coli* Membranen erfolgte mit einer Affinitätschromatographie und Gelfiltration. Alle Untereinheiten der EAF₁F₀ konnten mit Hilfe native Gelelektrophorese, Western-Blot in Kombination mit polyklonalen Antikörpern gegen die Untereinheiten α, β, γ, ε, δ und c, sowie massenspektrometrisch

nachgewiesen werden. Interessanterweise assemblierten die Untereinheiten a und c der *A. aeolicus* ATP-Synthase erfolgreich zu einem Holoenzym trotz der unterschiedlichen Lipidzusammensetzung in *E. coli*. Weiterhin zeigten "in-gel" Aktivitätsbestimmungen als auch Phosphatbestimmungen eine ATP-Hydrolyse Aktivität von $12,77 \pm 3,96$ U/mg für EAF_1F_O . Dies entspricht derselben Größenordnung der AAF_1F_O ($29,65 \pm 3,66$ U/mg). Die gemessenen Werte beschreiben eine spezifische ATP-Synthase Aktivität, da sie durch Azid (einem gängigen ATP-Synthase Inhibitor) um den Faktor 100 ($0,17 \pm 1,85$ U/mg) verringert werden konnten.

Elektronenmikroskopische Einzelmolekül aufnahmen zeigten, dass die Struktur des gereinigten Enzymkomplexes derjenigen eines voll assemblierten Holoenzym entspricht und identisch mit der Struktur der AAF_1F_O ist. Das nach Elektroelution aus „Blue Native“ Gelen entnommene EAF_1F_O offenbarte in elektronenmikroskopischen Bildern die charakteristische Pilzstruktur. EAF_1F_O ist etwa 200 Å lang und weist zwei definierte Subkomplexe auf. Der erste Subkomplex zeigt einen Durchmesser von 100 Å und kann sehr wahrscheinlich dem F_1 -Komplex zugeordnet werden. Der zweite Subkomplex ist etwa 100 Å breit, liegt parallel zur vermuteten Membranebene, ist 45 Å hoch und kann sehr wahrscheinlich dem F_O -Komplex zugeordnet werden.

4) Charakterisierung des N-terminalen Abschnitts der c-Untereinheit und wahrscheinlicher Assemblierungsmechanismus der AAF_1F_O

Die Etablierung eines heterologen Expressionssystems erlaubt nun die Manipulation der Gene, die für die ATP-Synthase kodieren und damit die bereits bioinformatisch gewonnenen Daten experimentell zu verifizieren. Multiple Sequence Alignments haben gezeigt, dass die c-Untereinheiten der ATP-Synthase in vier verschiedene phylogenetische Gruppen eingeteilt werden können. Die c-Untereinheiten der ATP synthase aus *A. aeolicus* und anderer extremophiler Organismen kann in die Gruppe 2 eingeordnet werden. Einzigartig für diese Gruppe ist das N-terminale Ende der c-Untereinheit. Diese charakteristische Signalsequenz interagiert mit dem "Signal Recognition Particle" (SRP), wird im Laufe des Maturationsprozesses der c-Untereinheit abgeschnitten und ist essentiell für die Membraninsertion. Dies konnte experimentell belegt werden. Deletionen oder Mutationen dieser Signalsequenz resultierten in einem Abbruch der Expression der EAF_1F_O . Die c-Untereinheiten der Gruppen 1, 3 und 4 werden dagegen ohne Mitwirkung des SRP in die Membran inseriert. Demnach unterscheidet sich der Assemblierungsmechanismus der c-Untereinheit der Gruppe 2 von dem mesophiler Prokaryoten. In mesophilen Prokaryoten (z.B. *E. coli*) assemblieren die Untereinheiten c und b, bevor der F_1 -Subkomplex an den restlichen Komplex angelagert wird. In dieser Arbeit konnte nun gezeigt werden, dass in *A. aeolicus* der F_1 -Subkomplex nur mit Hilfe der b_1 und b_2 Untereinheiten an die membranständigen Untereinheiten assoziiert wird. Die c-Untereinheiten, die in *A. aeolicus* mit

Hilfe eines für ATP-Synthasen einzigartigen SRP-abhängigem Mechanismus inseriert werden, werden dafür nicht benötigt.

Im Zuge dieser Arbeit wurde eine in *E. coli* exprimierte, vollständig assemblierte und voll funktionelle ATP-Synthase aus *A. aeolicus* aufgereinigt. Durch das etablierte heterologe Expressionssystem gelang es unter anderem den Insertionsmechanismus der c-Untereinheit der ATP-Synthase aus *A. aeolicus* aufzuklären. Weiterhin kann dieses heterologe Expressionssystem für einfache Mutagenese-, sowie „Cross-Linking“-Experimente genutzt werden. Hierdurch könnten Eigenschaften der ATP-Synthase charakterisiert werden, die bis dato noch unbekannt sind. Weiterhin bietet das hier präsentierte heterologe Expressionssystem eine Plattform, großer, mehrere Untereinheiten umfassender Enzymkomplexe mit komplizierter Stöchiometrie zu exprimieren, wie etwa Atmungskettenkomplexe, Transporter, oder andere makromolekulare Maschinen, die gerade im Fokus der aktuellen Forschung stehen.

Detailed English Summary

F_1F_0 ATP synthases catalyze the synthesis of ATP from ADP and inorganic phosphate driven by ion motive forces across the membrane. F_1F_0 ATP synthases are present in bacteria, mitochondria and chloroplasts and they have been remarkably conserved throughout evolution. The overall enzyme is composed of two distinct subcomplexes. The soluble F_1 subcomplex is formed by the subunits α , β , γ , δ , and ϵ in a 3:3:1:1:1 stoichiometry. It catalyzes ATP synthesis or hydrolysis via a so called binding change mechanism. The hydrophobic membrane-inserted F_0 subcomplex is formed by subunits a, b, and c in a 1:2:(8-15) stoichiometry and is the center of ion translocation. Subunits γ and ϵ form a central stalk and subunits b and δ form a peripheral stalk thus providing a connection between the F_1 and the F_0 subcomplexes, which prevents uncoupling in the reciprocal rotation of F_1 and F_0 subunits. A number of ATP synthases have been characterized to date. High resolution structures are available for different parts of the ATP synthase complex, i.e. the bovine F_1 subcomplex, c-rings from *Ilyobacter tartaricus*, *Bacillus pseudofirmus* and *Arthrospira platensis* and subcomplexes of F_1 and c-rings from yeast and bovine mitochondria. Structures of the peripheral stalk of bovine F_1F_0 ATP synthase were also determined. However, to date, no high resolution structure is available for the entire F_1F_0 ATP synthase complex. Furthermore, the mechanism of ion translocation remains unknown due to a lack of high resolution models for the membrane embedded subcomplex F_0 .

Previous work from our lab reported that F_1F_0 ATP synthase extracted and purified from the native cells of the bacterium *Aquifex aeolicus* is highly stable, due to the hyperthermophilic nature of *A. aeolicus*. The enzyme can be purified in an active, fully assembled form and presents unique structural features such as a structurally bent central stalk and a putatively heterodimeric peripheral stalk. In contrast to other ATP synthases, the peripheral stalk of *A. aeolicus* ATP synthase is more rigid and remains intact during purification. Taken together, such features make *A. aeolicus* ATP synthase a promising candidate for studies aiming to determine the structure of the intact enzyme, which would provide an interesting model system for structural and functional studies on ATP synthases.

Based on these premises, this doctoral work had four main objectives: 1) to extend the previous characterization of native *A. aeolicus* F_1F_0 ATP synthase (hereafter named AAF $_1F_0$) by bioinformatic, biochemical and functional studies, 2) to create a heterologous expression system for producing this enzyme in *E. coli*, 3) to characterize the heterologously produced ATP synthase (hereafter named EAF $_1F_0$) and 4) to study properties of *A. aeolicus* F_1F_0 ATP synthase that could only be addressed using a heterologous expression system, i.e. the role of the N-terminus of subunit c.

1) Characterization of the native *A. aeolicus* F₁F₀ ATP synthase (AAF₁F₀).

Bioinformatic studies based on multiple-sequence alignments on the membrane subunits of *A. aeolicus* F₁F₀ ATP synthase revealed that these subunits possess different properties than previously proposed using topology predictions. Specifically, the alignments suggested that (i) subunit a possesses five (and not six) transmembrane helices, (ii) both subunit b₁ and b₂ possess a membrane-inserted N-terminal helix, while subunit b₂ possesses a putative signal peptide preceding this N-terminal transmembrane helix which may be thus cleaved off in the mature form of subunit b₂, and (iii) subunit c possesses an N-terminal region different in length and hydrophobicity from that of subunits c of other ATP synthases, with possible consequences on its membrane insertion and assembly mechanism.

In addition to studying the properties of its subunits bioinformatically, in this work AAF₁F₀ was also purified and characterized biochemically. Most importantly, the previously established purification protocol was optimized to enhance the stability of the complex and prevent its disassembly into smaller subcomplexes. The major improvement involved the identification of *trans*-4-(*trans*-4'-propylcyclohexyl)cyclohexyl- α -D-maltoside (α -PCC) as a new detergent that can stabilize the entire F₁F₀ complex much better than the previously used maltoside detergents n-dodecyl- β -D-maltoside (DDM) and n-decyl- β -D-maltoside (DM).

After optimizing the purification protocol for AAF₁F₀, functional studies were performed to confirm experimentally that AAF₁F₀ is a proton-dependent (and not a sodium ion-dependent) ATP synthase, which was expected since the Na⁺ binding site signature is not present in the sequence of *A. aeolicus* ATP synthase subunit c. Specifically, MALDI-TOF mass spectrometry (MS) was used to show that AAF₁F₀ subunit c is not protected by sodium and reacts with the covalent active-site ligand *N,N'*-dicyclohexyl-carbodiimide (DCCD) in the presence of over-physiological sodium concentrations (150 mM), a typical feature of proton-dependent ATP synthases.

Further enzymatic studies revealed that the ATP hydrolysis activity of AAF₁F₀ is negligible at temperatures below 60°C. While this observation is recurrent in enzymes from hyperthermophilic organisms, the enzymatic data, taken together with other biochemical and structural results, may be indicative of a specific structural-functional property of AAF₁F₀. In particular, SDS-PAGE analysis revealed that subunit γ requires prolonged heat treatment to detach from subunits α / β , and single-particle electron microscopy (EM) revealed that subunit γ is bent, as mentioned above. In ATP synthase from *Caldalkalibacillus thermarum* TA2.A1, such features were correlated to a functionally important conformational switch of subunit γ , which may therefore also happen in *A. aeolicus* ATP synthase. Specifically, subunit γ may adopt an inactive, bent conformation forming

tight salt-bridges to α/β at room-to-low temperatures, and may then switch to an active, more extended conformation above 60 °C.

Finally, work on AAF₁F_O also led to the generation of polyclonal antibodies against subunits α , β , γ , ϵ and c , which were greatly useful in the design and characterization of the heterologous expression system described below.

2) Cloning of an artificial operon for heterologous expression of *A. aeolicus* ATP synthase.

To enable a more manageable investigation on all unique properties characterizing the *A. aeolicus* ATP synthase, an expression system was developed in this work to produce the enzyme heterologously. The design of a heterologous expression vector for the *A. aeolicus* ATP synthase is very challenging because (i) *A. aeolicus* F₁F_O ATP synthase is a large heteromultimeric enzyme of more than 500 kDa in size, with a complex and uncharacterized subunit stoichiometry and (ii) in *A. aeolicus*, the nine *atp* genes are not clustered in one operon, but they are distributed over four different genomic loci. In total, there are six different DNA fragments harboring the genes for the nine subunits of *A. aeolicus* F₁F_O ATP synthase and some genes overlap. Despite the hyperthermophilic nature of *A. aeolicus*, the mesophilic host *E. coli* was chosen for the expression study because it is a cheap, well studied host that had already been successfully used for the production of many other ATP synthases.

The strategy used to produce the *A. aeolicus* ATP synthase in *E. coli* consisted of the following steps. In the first step, single-gene expression was attempted for individual subunits (a , c , γ and ϵ) and dual-gene expression was attempted for specific combinations of different subunits (b_1 - b_2 , a - c , a - b_1 - b_2 , γ - ϵ). At this step, the *atp* genes were cloned into the open reading frame (ORF) of several expression vectors. Expression tests served as checkpoints for assessing the ability of *E. coli* to recognize native operons, native codons, and overlapping genes and to produce functionally active intermediate complexes of the enzyme. Co-transforming the expression vectors with the commercial vector pRARE – which encodes rare tRNAs – revealed to be very beneficial in overcoming codon usage biases between *A. aeolicus* and *E. coli*. In the second step, expression was attempted for subcomplexes of ATP synthase including subunits located in different loci of *A. aeolicus* genome. With this strategy two subcomplexes were produced: F₁- $\alpha\beta\gamma$ and F₁- $\alpha\beta\gamma\epsilon$, both modified with an N-terminal His₆-tag on subunit β for detection and purification purposes. A third vector including all F_O subunits and subunit δ (F_O- $acbb_2\delta$) was also created. This second step guided the appropriate choice of intergenic regions and of the purification tag for artificial operons. Finally, in the third step, the genes for F₁- $\alpha\beta\gamma\epsilon$ and for F_O- $acb_1b_2\delta$ were combined into a single expression vector, which thus contained all nine genes encoding the entire *A. aeolicus* F₁F_O ATP

synthase. The artificial operon was designed to harbor the nine *atp* genes in the order *atpBEF1F2HAGDC*, the same order of the *atp* genes in the native operon from *E. coli*, except for gene *atpI* that is not present in the *A. aeolicus* genome. The correct stoichiometry of all subunits is regulated by accurate selection of the translation initiation regions (TIR). The native TIRs from *A. aeolicus* were used, including 30 base pairs upstream of the start codon of the six genes *atpB*, *atpE*, *atpF1*, *atpA*, *atpG* and *atpC* and introducing only minimal modifications to the native sequences to insert restriction sites and purification tags, but preserving the original intergenic distance between neighboring genes.

3) Characterization of EAF₁F₀ and its subcomplexes.

At each step in the creation of the heterologous expression system, the subunits or subcomplexes that could be successfully expressed were characterized at a biochemical, functional and/or structural level, and such characterization led to the following significant results.

First, the successful production and purification of subcomplex b₁b₂ by dual-gene expression showed that (i) *E. coli* can recognize *A. aeolicus* native operons and overlapping genes, (ii) subunits b₁ and b₂ form a complex in the *E. coli* membranes and (iii) the complex can be purified to homogeneity and is stable in detergent over time. The observation that subunits b₁ and b₂ can associate to form a complex *in vitro* corroborates the previous hypothesis that these two subunits form a heterodimeric peripheral stalk in the native *A. aeolicus* ATP synthase, which is a unique case among ATP synthases of non-photosynthetic organisms. By contrast, membrane subunits a and c cannot be produced in *E. coli* in isolation, as already reported for other ATP synthases.

Second, the successful production and purification of subcomplex F₁-αβγ and F₁-αβγε indicated that functionally active forms of *A. aeolicus* ATP synthase can be obtained from artificial operons in *E. coli*. All subunits were identified by peptide mass fingerprinting (PMF) followed by ESI-MS and the subcomplexes showed rates of ATP hydrolysis of 1.35 ± 0.14 U/mg and 1.73 ± 0.11 U/mg, respectively.

Finally and most importantly, also the production and purification of the entire ATP synthase complex (EAF₁F₀) was accomplished successfully. EAF₁F₀ was purified from the membranes of *E. coli* by affinity and size-exclusion chromatography. All subunits in the pure ATP synthase were identified by native gel electrophoresis, by Western blot analysis using the polyclonal antibodies specifically generated against subunits α, β, γ, ε, δ and c, and by mass spectrometry. Remarkably, also the membrane subunits a and c were identified, revealing that in the context of the whole enzyme they can be correctly incorporated into ATP synthase, despite the differences in the lipid composition between *A. aeolicus* and *E. coli* membranes. Furthermore, in-gel activity assays and

phosphate determination assays showed that the ATP hydrolysis activity of EAF₁F_O is 12.77 ± 3.96 U/mg, a value of the same order of magnitude (43%) as for AAF₁F_O (29.65 ± 3.66 U/mg) and comparable to other respiratory complexes of *A. aeolicus* (i.e. respiratory complex I, sulfide:quinone oxidoreductase) and to other ATP synthases. Such enzymatic activity was reduced approximately 100 fold (0.17 ± 1.85 U/mg residual activity) by 0.02% (w/v) sodium azide, a common inhibitor of bacterial ATP synthases.

Finally, single-particle electron microscopy (EM) showed that EAF₁F_O is fully assembled and possesses an identical structural organization as AAF₁F_O. After electro-elution from BN-PAGE gels, characteristic “mushroom” shaped-particles of EAF₁F_O could be observed. EAF₁F_O is ~ 200 Å long and has two distinct parts. One part possesses a globular shape with a diameter of 100 Å and likely corresponds to the F₁ subcomplex. The other part is approximately 100 Å wide parallel to the putative membrane plane and ~ 45 Å high and likely corresponds to the F_O subcomplex. Importantly, both the central and peripheral stalks are clearly visible in the EM images.

4) Characterization of the N-terminal segment of subunit c and hypotheses on the assembly mechanism of *A. aeolicus* ATP synthase.

The availability of a heterologous system to produce *A. aeolicus* ATP synthase allows for previously impossible manipulations of this enzyme at the genetic level. Therefore, this system was used to investigate the unique properties which had been identified bioinformatically for the N-terminal region of subunit c (see above). The multiple-sequence alignment revealed the presence of four phylogenetic groups of subunit c (groups 1 to 4). The subunit c from *A. aeolicus* F₁F_O ATP synthase clusters together with subunit c of other early diverging and extremophilic organisms into what is reported here as the group 2. As a unique case for bacterial F₁F_O ATP synthases, the N-terminal segment of group 2 members possesses features typical of signal peptides that interact with signal recognition particle (SRP). In this work, we proved experimentally that the N-terminus of *A. aeolicus* subunit c is indeed a signal peptide that is cleaved off in the mature form of subunit c. By designing mutations in our EAF₁F_O construct, we proved that such a signal peptide is obligatorily required for membrane insertion, because deleting it or replacing it with the N-terminal segment of subunit c from other groups completely abolishes expression in our EAF₁F_O construct. Therefore, we conclude that group 2 subunits c likely follow a SRP-dependent membrane insertion pathway different from that of other subunits c, which are instead known not to require SRP. As a consequence of these considerations and based on the results from our heterologous expression tests, we propose that group 2 ATP synthases may have evolved to follow a unique assembly mechanism different from that of other mesophilic prokaryotic homologues. In mesophilic organisms (i.e. *E. coli*) it is known that subunits b and c preassemble before recruiting the F₁

subcomplex to the membranes. Instead, in our experiments we noted that the F_1 subcomplex is recruited to the membranes by subcomplex b_1b_2 in the absence of subunit c . Therefore, it is possible that in *A. aeolicus* subunit c is incorporated into ATP synthase at a later stage of its assembly process, and this may be potentially correlated to the fact that its membrane insertion follows a unique SRP-dependent pathway.

In conclusion, the successful production of the fully assembled and active F_1F_0 ATP synthase from *A. aeolicus* in *E. coli* provides a novel genetic system to study *A. aeolicus* F_1F_0 ATP synthase. While this system has already enabled the investigation of the subunit c membrane insertion mechanism, many more experiments are now feasible. Genetic manipulation allows for relatively straightforward mutagenesis experiments and cross-linking experiments, with direct application for novel functional and structural studies to address the properties of ATP synthase that are still poorly characterized. At the same time, the heterologous expression system described in this work also constitutes a solid reference for designing strategies aimed at producing other large multi-subunit complexes with complicated stoichiometry, i.e. other respiratory complexes, the nuclear pore complex, transporter systems and many other macromolecular machines that are currently very active research targets.

Abstract

This work presents a biochemical, functional and structural characterization of *Aquifex aeolicus* F₁F₀ ATP synthase obtained using both a native form (AAF₁F₀) and a heterologous form (EAF₁F₀) of this enzyme.

F₁F₀ ATP synthases catalyze the synthesis of ATP from ADP and inorganic phosphate driven by ion motive forces across the membrane and therefore play a key cellular function. Because of their central role in supporting life, F₁F₀ ATP synthases are ubiquitous and have been remarkably conserved throughout evolution. For their biological importance, F₁F₀ ATP synthases have been extensively studied for many decades and many of them were characterized from both a functional and a structural standpoint. However, important properties of ATP synthases – specifically properties pertaining to their membrane embedded subunits – have yet to be determined and no structures are available to date for the intact enzyme complex. Therefore, F₁F₀ ATP synthases are still a major focus of research worldwide. Our research group had previously reported an initial characterization of AAF₁F₀ and had indicated that this enzyme presents unique features, i.e. a bent central stalk and a putatively heterodimeric peripheral stalk. Based on such a characterization, this enzyme revealed promising for structural and functional studies on ATP synthases and became the focus of this doctoral thesis. Two different lines of research were followed in this work.

First, the characterization of AAF₁F₀ was extended by bioinformatic, biochemical and enzymatic analyses. The work on AAF₁F₀ led to the identification of a new detergent that maintains a higher homogeneity and integrity of the complex, namely the detergent *trans*-4-(*trans*-4'-propylcyclohexyl)cyclohexyl- α -D-maltoside (α -PCC). The characterization of AAF₁F₀ in this new detergent showed that AAF₁F₀ is a proton-dependent, not a sodium ion-dependent ATP synthase and that its ATP hydrolysis mechanism needs to be triggered and activated by high temperatures, possibly inducing a conformational switch in subunit γ . Moreover, this approach suggested that AAF₁F₀ may present unusual features in its membrane subunits, i.e. short N-terminal segments in subunits a and c with implications for the membrane insertion mechanism of these subunits.

Investigating on these unique features of *A. aeolicus* F₁F₀ ATP synthase could not be done using *A. aeolicus* cells, because these require a harsh and dangerous environment for growth and they are inaccessible to genetic manipulations. Therefore, a second approach was pursued, in which an expression system was created to produce the enzyme in the heterologous host *E. coli*. This second approach was experimentally challenging, because *A. aeolicus* F₁F₀ ATP synthase is a 500-kDa multimeric membrane enzyme with a complicated and still not entirely determined stoichiometry and because its encoding genes are scattered throughout *A. aeolicus* genome, rather than being

organized in one single operon. However, an artificial operon suitable for expression was created in this work and led to the successful production of an active and fully assembled form of *Aquifex aeolicus* F₁F₀ ATP synthase. Such artificial operon was created using a stepwise approach, in which we expressed and studied first individual subunits, then subcomplexes, and finally the entire F₁F₀ ATP synthase complex. We confirmed experimentally that subunits b₁ and b₂ form a heterodimeric subcomplex in the *E. coli* membranes, which is a unique case among ATP synthases of non-photosynthetic organisms. Moreover, we determined that the b₁b₂ subcomplex is sufficient to recruit the soluble F₁ subcomplex to the membranes, without requiring the presence of the other membrane subunits a and c. The latter subunits can be produced in our expression system only when the whole ATP synthase is expressed, but not in isolation nor in the context of smaller F₀ subcomplexes. These observations led us to propose a novel mechanism for the assembly of ATP synthases, in which first the F₁ subcomplex attaches to the membrane via subunit b₁b₂, and then c-ring and subunits a assemble to complete the F₀ subcomplex. Furthermore, we could purify the heterologous ATP synthase (EAF₁F₀) to homogeneity by chromatography and electro-elution. Enzymatic assays showed that the purified form of EAF₁F₀ is as active as AAF₁F₀. Peptide mass fingerprinting showed that EAF₁F₀ is composed of the same subunits as AAF₁F₀ and all soluble and membrane subunits could be identified. Finally, single-particle electron microscopy analysis revealed that the structure of EAF₁F₀ is identical to that of AAF₁F₀. Therefore, the EAF₁F₀ expression system serves as a reliable platform for investigating on properties of AAF₁F₀.

Specifically, in this work, EAF₁F₀ was used to study the membrane insertion mechanism of rotary subunit c. Subunits c possess different lengths and levels of hydrophobicity across species and by analyzing their N-terminal variability, four phylogenetic groups of subunits c were distinguished (groups 1 to 4). As a member of group 2, the subunit c from *A. aeolicus* F₁F₀ ATP synthase is characterized by an N-terminal segment that functions as a signal peptide with SRP recognition features, a unique case for bacterial F₁F₀ ATP synthases. By accurately designing mutants of EAF₁F₀, we determined that such a signal peptide is strictly necessary for membrane insertion of subunit c and we concluded that *A. aeolicus* subunit c inserts into *E. coli* membranes using a different pathway than *E. coli* subunit c. Such a property may be common to other ATP synthases from extremophilic organisms, which all cluster in the same phylogenetic group.

In conclusion, the successful production of the fully assembled and active F₁F₀ ATP synthase from *A. aeolicus* in *E. coli* reported in this work provides a novel genetic system to study *A. aeolicus* F₁F₀ ATP synthase. To a broader extent, it will also serve in the future as a solid reference for designing strategies aimed at producing large multi-subunit complexes with complicated stoichiometry.

Abbreviations

The abbreviations used in this work are listed in the following table.

Table of abbreviations

1. Symbols for measures and units

Å	angstrom
°C	degrees celsius
Da	dalton
h	hour
L	liter
M	molar
min	minute
Pa	pascal
ppm	parts per million
rpm	rotations per minute
U	(enzymatic) unit
V (as unit)	volt
V (as measure)	volume
v/v	volume/volume
w/v	weight/volume
w/w	weight/weight

2. Biomolecules and non-conventional chemicals

AAF ₁ F ₀	<i>A. aeolicus</i> ATP synthase isolated from native cells of <i>A. aeolicus</i>
ACN	acetonitrile
AHC	ammonium hydrogen carbonate
BCA	bicinchonin acid
BisTris	1,3-bis(tris(hydroxymethyl)methylamino)propane
ddH ₂ O	bidistilled water (Millipore)
DDM (also LM)	n-dodecyl-β-D-maltoside
DM	n-decyl-β-D-maltoside
EDTA	ethylenediaminetetracetic acid
EAF ₁ F ₀	<i>A. aeolicus</i> ATP synthase heterologously produced in <i>E. coli</i>
α-PCC	<i>trans</i> -4-(<i>trans</i> -4'-propylcyclohexyl)cyclohexyl-α-D-maltoside
IPTG	Isopropyl-β-thiogalactopyranoside
βME	2-mercapto-ethanol
MES	2-(N-morpholino)-ethanesulfonic acid
Ni-NTA	Ni-Nitrilotriacetic acid
SDS	sodium dodecyl sulfate
TEMED	N,N,N',N'-tetramethylethylenediamine

Abbreviations

Tris	Tris-hydroxymethyl-aminomethane
------	---------------------------------

3. Techniques and instrumentation

BN	blue-native
CN	clear-native
EM	electron microscopy
ESI	electrospray ionisation
IMAC	immobilized-metal affinity chromatography
LC	liquid chromatography
MALDI	matrix-assisted laser desorption ionisation
MS	mass spectroscopy
PAGE	polyacrylamide gel electrophoresis
PMF	peptide mass fingerprint
SEC	size-exclusion chromatography
SMART	simple modular architecture research tool
TOF	time of flight

4. Databases and software

BLAST	Basic Local Alignment Search Algorithm
FASTA	FAST-All
NCBI	National Centre for Biotechnology Information
PDB	Protein Data Bank
T-COFFEE	Tree-based Consistency Objective Function For AlignmEnt Evaluation
TMHMM	Transmembrane Hidden Markov Model

5. General abbreviations

2-D	bi-dimensional
3-D	tri-dimensional
CMC	critical micelle concentration
conc	concentration
e.g.	exempli gratia (lat., engl.: for example)
<i>et al.</i>	et alii (lat., engl.: and others)
eq.	equation
i.e.	in exemplum (lat., engl.: for example)
M	protein molecular weight marker
MW	molecular weight
pI	isoelectric point
UV/Vis or Vis/UV	ultraviolet/visible

1. Introduction

Cells require energy to drive metabolic processes and maintain their homeostasis. The primary energy source for life on Earth is sunlight. Sunlight is converted into chemical energy through photosynthesis, which produces organic molecules. Besides photosynthetic organisms, chemolithoautotrophic organisms are also able to fix carbon and synthesize organic molecules, but they use chemical energy from inorganic compounds instead of sunlight. Organic molecules are then assembled (in anabolic processes) and disassembled (in catabolic processes) in a variety of enzymatic reactions, some of which are exergonic (energy-yielding) and some endergonic (energy-requiring). For efficient energy transfer from exergonic to endergonic processes, the cells store chemical energy in highly energized chemical bonds of intermediary molecules. The most widely used intermediary molecule for cellular energy transfer is adenosine triphosphate (ATP) (e.g. Voet and Voet (2004)).

Adenosine triphosphate (ATP) was discovered in 1929 (Lohmann, 1929) and was proposed to be the main energy transfer molecule in the cell in 1941 (Lipmann, 1941). The ATP molecule is composed of three distinct moieties, an adenine ring, a ribose and a triphosphate group. The triphosphate group is bound in position 5' of the ribose, and the three phosphates that compose it are termed as α , β , and γ . The α phosphate binds the ribose directly forming an ester bond, while the β phosphate binds to the α phosphate, and the γ phosphate binds to the β phosphate through two consecutive phosphoanhydride bonds (Figure 1.1). These phosphoanhydride bonds are highly energized. They are moderately stable in the absence of enzymes but they can be hydrolyzed rapidly in the presence of enzymes (Westheimer, 1987). Their hydrolysis is exergonic. It releases ~ 30 kJ/mol per bond, and it can therefore be used to drive endergonic metabolic processes. It was calculated that in a resting human, around 40 kg of ATP are produced and consumed per day (Capaldi and Aggeler, 2002). Due to this high turnover of ATP and the impermeability of cell membranes for ATP, each cell has to produce its own ATP. Two main

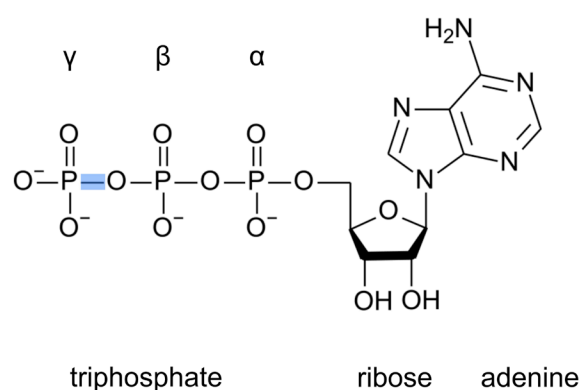


Figure 1.1. Adenosine triphosphate (ATP). ATP composes of adenine, ribose and triphosphate. The first phosphate (α) forms a 5'-ester bond with ribose. The second (β) and third (γ) phosphates are connected through highly energized phosphoanhydride bonds. The hydrolysis of the γ phosphoanhydride bond (shaded blue) releases ~ 30 kJ/mol.

biochemical pathways provide energy for cellular synthesis of ATP, substrate-level phosphorylation, and oxidative or photophosphorylation (e.g. Voet and Voet (2004)). Substrate-level phosphorylation consists of the transfer of highly energetic phosphoryl groups from intermediate molecules formed during catabolic reactions to adenosine diphosphate (ADP). Examples of substrate-level phosphorylation are anaerobic fermentation processes such as glycolysis, which typically generates 2 ATP molecules per molecule of glucose that is converted to pyruvate (e.g. Voet and Voet (2004)). Instead, in oxidative and photophosphorylation processes, chemical or light energy is converted into an electrochemical gradient across a biological membrane generating an ion (H^+ or Na^+) motive force that is then used by the enzyme ATP synthase to generate ATP. Oxidative phosphorylation is the most efficient pathway for producing cellular ATP, because it yields 30 - 36 ATP molecules per molecule of glucose that is converted to carbon dioxide and water (e.g. Voet and Voet (2004)).

1.1. Electron transport and oxidative phosphorylation

Although in net terms oxidative phosphorylation is a combustion reaction, cells have evolved to minimize heat dissipation and maximize conservation of chemical energy by separating the process into a series of separate steps. Throughout these steps, the electrons are gradually transferred from the initial donor (reduced organic or inorganic compounds) to intermediate acceptors (NADH, $FADH_2$, FeS clusters, quinones, heme groups, and metal ions), and subsequently to the final acceptor (molecular oxygen, O_2) by means of sequential oxidoreductive reactions. Depending on the nature of the initial electron donor, many different enzymes may take part in the electron-transport process. In all cases, a central role is played by a series of membrane-inserted enzymes that form the so-called respiratory chain complexes (respiratory complexes I-V). In 1961, Peter Mitchell first proposed that these respiratory complexes are responsible for generating the proton motive force (pmf) required by ATP synthase to synthesize ATP, a theory known as the chemiosmotic theory. Mitchell's chemiosmotic theory explains the mechanism of energy transduction and energy coupling between electron transport and ATP synthesis (Mitchell, 1961). According to Mitchell's theory, the pmf generated by the respiratory complexes consists of a proton (ion) concentration difference (ΔpH) and an electric potential difference ($\Delta\psi$) across biological membranes, and is expressed by Eq. (1.1):

$$\text{pmf} = - \Delta\mu_{H^+} / F = \Delta\psi - (2.3 RT \Delta pH) / F \quad \text{Eq.(1.1)}$$

where F is the Faraday constant ($96\,485\text{ C mol}^{-1}$); R is the molar gas constant ($8.314\text{ J mol}^{-1}\text{ K}^{-1}$), T is the temperature in Kelvin, and $\Delta\psi$ is expressed in volts. The value $\Delta\mu_{H^+}$ indicates how much energy is required (or released, depending on the direction of the transmembrane proton flow) to

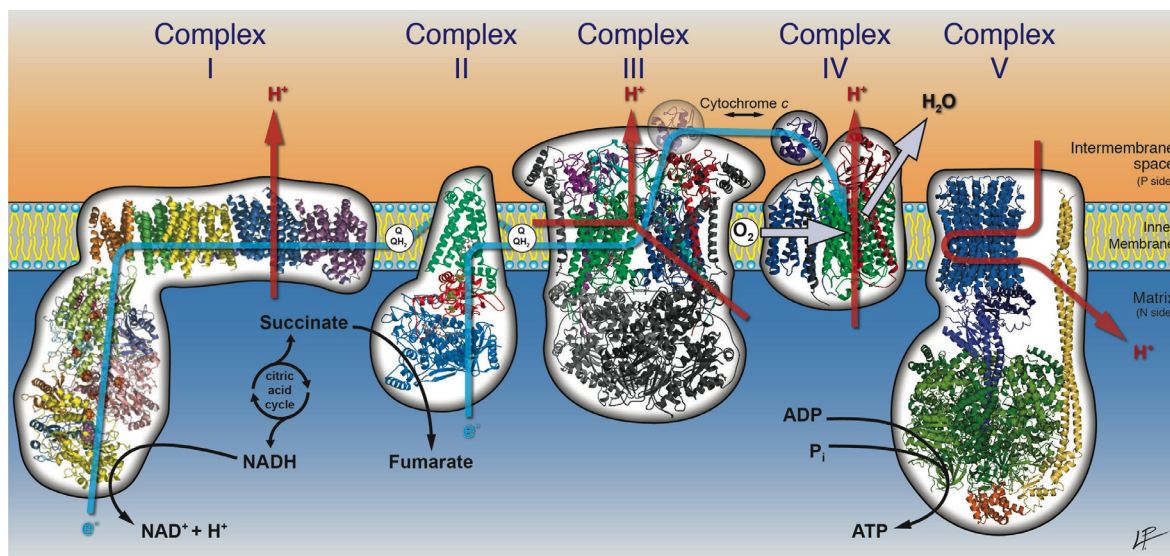


Figure 1.2. Schematic overview of the respiratory chain complexes. The figure represents 3-D structures determined for enzymes of the mitochondrial respiratory chain. Complex I (NADH:quinone oxidoreductase) oxidizes NADH, produced in the citric acid cycle, to NAD^+ and reduces quinone (Q) to quinol (QH_2). The oxidoreduction reaction, which occurs in the soluble part of the enzyme (facing the mitochondrial matrix), is mechanically coupled to the transfer of 4 protons (H^+) from the matrix to the intermembrane space mediated by the membrane-inserted subunits. Complex II (succinate:quinone oxidoreductase) oxidizes another product of the citric acid cycle, succinate, to fumarate, thereby also reducing Q to QH_2 . This reaction is not known to be coupled to proton transfer. Complex III (quinol:cytochrome *c* oxidoreductase) transfer electrons from QH_2 to cytochrome *c* thereby pumping 2 H^+ into the intermembrane space. Complex IV (cytochrome *c* oxidase) transfers electrons from cytochrome *c* to the final electron acceptor, molecular oxygen (O_2), to produce water. Also this last oxidoreduction reaction is associated to the transfer of 4 H^+ across the mitochondrial membrane. Finally, the proton gradient generated by complexes I-IV is used by complex V (ATP synthase) to produce ATP from ADP and inorganic phosphate (P_i). All structures are drawn in cartoon representations assigning different colors to individual subunits. Blue arrows depict the electron pathway. Red arrows depict the transport of protons. This figure was kindly provided by Paolo Lastrico (MPI of Biophysics, Frankfurt, Germany).

transport 1 mol of protons across the membrane. As an example, the mitochondrial respiratory chain complexes are depicted in Figure 1.2.

1.2. F_1F_0 ATP synthase (Respiratory complex V)

Most of the ATP produced in the cell is synthesized by the membrane-inserted heteromultimeric enzyme F_1F_0 ATP synthase, also known as rotary ATP synthase. This protein is an extremely high turnover enzyme that has to be constitutively active, because ATP has to be produced and supplied continuously to the cell. Therefore, the amount of ATP synthesized daily in a living organism by F_1F_0 ATP synthase can be very high. For instance, a resting human being uses his/her own body weight of ATP every day (Capaldi and Aggeler, 2002). As a consequence, ATP synthase is a crucial enzyme for the cell and not surprisingly malfunctions of this enzyme result in severe mitochondrial diseases that are often lethal or manifesting in children very shortly after birth (Houstek *et al.*, 2006).

ATP synthase was first discovered in bacterial crude extracts in 1956 (Brodie and Gray, 1956). Since then, genes encoding ATP synthase have been identified in all sequenced genomes, and it is now known that rotary ATP synthases are present in all three domains of life (Kibak *et al.*, 1992).

1.2.1. Classification, nomenclature and architecture

F₁F₀ ATP synthase is one of three types of membrane-inserted ATPases, according to a classification based on function and taxonomic origin (Cross and Muller, 2004). F₁F₀ ATP synthases, also known as F-type ATP synthases, couple ion-translocation and ATP synthesis in bacteria, mitochondria and chloroplasts (Senior, 1988; Boyer, 1997). The letter "F" was historically chosen to distinguish this type of ATP synthases because these enzymes had been identified as "phosphorylation factors". More specifically, the term F₀ was introduced after the discovery that such phosphorylation factor is sensitive to oligomycin in mitochondria. Other membrane ATPase types are V-type ATPases, present in eukaryotic vacuoles and responsible for pumping protons across the membrane of intracellular compartments driven by ATP hydrolysis (Nelson and Taiz, 1989) and A-type ATP synthases, also called prokaryotic V-ATP synthases, which couple ion-translocation and ATP synthesis in archaea and some bacteria (Yokoyama *et al.*, 2003) and whose function is similar to that of F-type ATP synthases.

F₁F₀ ATP synthase is a sophisticated molecular motor and its three-dimensional structure has been selected throughout evolution to support its conformational dynamism in an efficient manner. As a large mushroom-shaped asymmetric protein complex, ATP synthase is composed of two distinct subcomplexes, F₁ and F₀, that function as opposing motors (Boyer, 1997; Stock *et al.*, 2000; Senior, 2007; von Ballmoos *et al.*, 2008; Junge *et al.*, 2009; von Ballmoos *et al.*, 2009). Under low ionic strength condition, the two subcomplexes F₁ and F₀ can be dissociated into two fully functional entities (Boyer, 1997). Moreover, depending on the physiological demand for ATP in the cell, ATP synthase can reverse its direction of operation (from ATP synthesis to ATP hydrolysis). The two subcomplexes, F₁ and F₀, are connected by a central rotating stalk and by a peripheral stator stalk that anchors the F₁ part to the membrane preventing its rotation. Therefore, F₁F₀ ATP synthase can be mechanically divided into two distinct parts, a rotor and a stator, as it was defined by crosslinking experiments (Tsunoda *et al.*, 2001). The stalks are important not only from a structural point of view, because they hold the F₁ and F₀ subcomplexes together, but also from a functional point of view, because they guarantee that the two subcomplexes can reciprocally exchange energy. In its uncoupled state, the interaction between F₁ and F₀ is disrupted and the energy transduction is lost (Capaldi and Aggeler, 2002).

Each of the two subcomplexes F_1 and F_0 is by itself composed of multiple individual subunits. In bacteria the soluble F_1 part possesses the universal subunits composition $\alpha_3\beta_3\gamma\delta\epsilon$. Whereas, the membrane-embedded F_0 part is formed by subunits ab_2c_{8-15} (the number of c subunits varies from 8-15 in different organisms) (Pogoryelov *et al.*, 2012). Therefore, in summary, bacterial F_1F_0

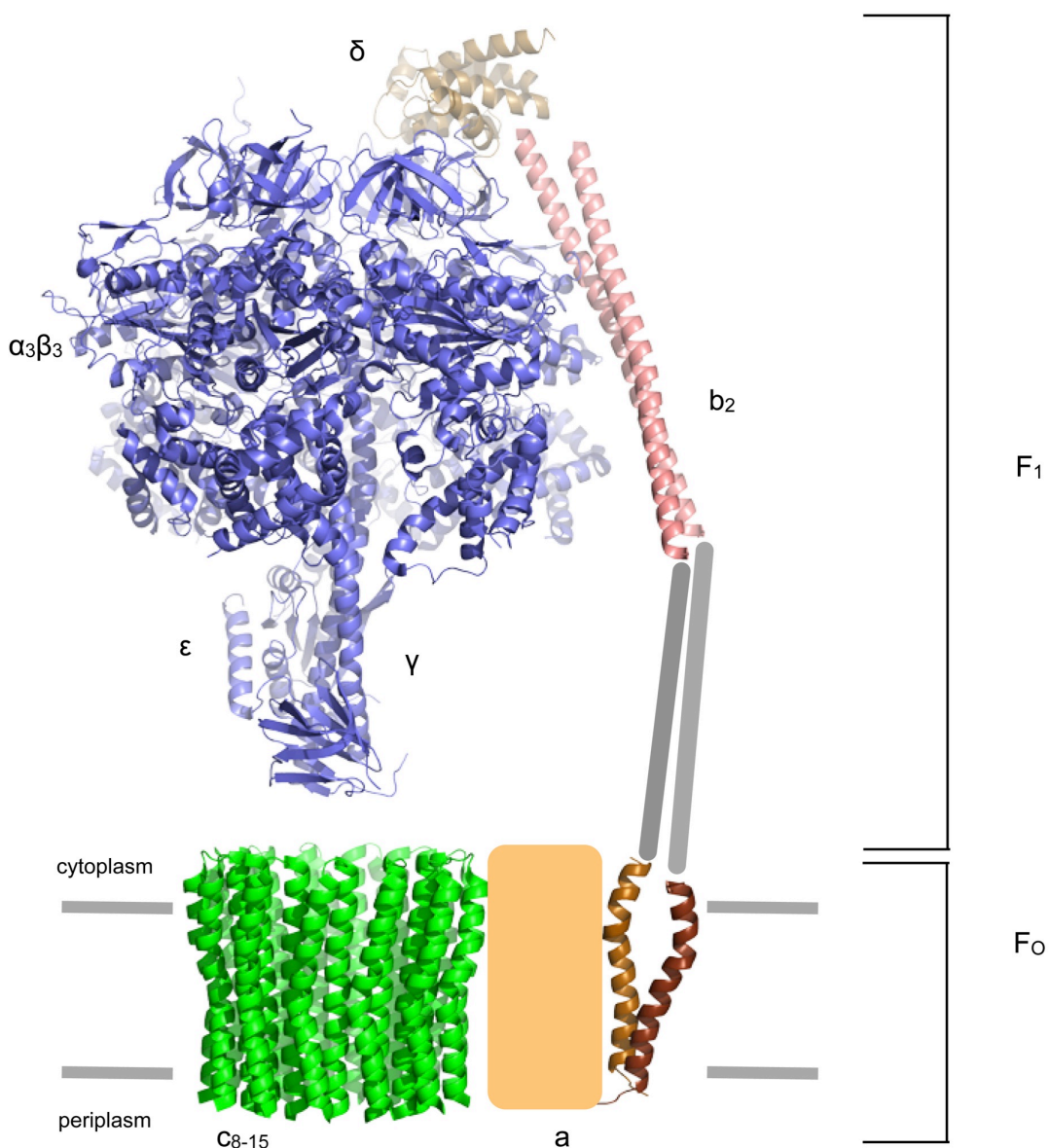


Figure 1.3. Architecture and subunit composition of bacterial F_1F_0 ATP synthase. The figure represents a composite ATP synthase model, manually drawn using the following structures: c -ring of *Ilyobacter tartaricus* ATP synthase (PDB id. 1CYE, green), F_1 subcomplex of *Escherichia coli* ATP synthase (PDB id. 1JNV, purple), subunit δ of *E. coli* ATP synthase (NMR structure, PDB id. 2A7U, light yellow), the dimerization domain of subunit b (PDB id. 1L2P, pink), and the membrane-inserted part of subunit b of *E. coli* ATP synthase (NMR structure, PDB id. 1B9U, brown). No high-resolution structural data is available for subunit a (thus represented as an orange rectangle) or for the hinge region of subunit b (represented as grey rods).

synthase is a multi-subunit membrane protein complex with a molecular mass of > 500 kDa, with stoichiometry $\alpha_3\beta_3\gamma\epsilon\delta a b_2 c_{8-15}$. In photosynthetic organisms (chloroplasts and cyanobacteria), the ATP synthase composition is the same as in the bacterial ATP synthase except for the presence of two, instead of one, isoforms of subunit b. In contrast, mitochondrial ATP synthase is much more complex and contains some additional subunits: F_1 comprises of α , β , γ , ϵ , δ and F_0 contains subunits c, a, b, d, F_6 , OSCP, and the accessory subunits e, f, g and A6L (Jonckheere *et al.*, 2012). Mitochondrial subunit δ is homologous to the bacterial subunit ϵ , whereas the homologue of the bacterial subunit δ in mitochondria is the oligomycinsensitivity conferring protein (OSCP) (Walker and Dickson, 2006). The mitochondrial subunit ϵ and the accessory subunits e, f, g, and A6L have no counterpart in bacteria. Finally, subunit 9 in mitochondrial F_1F_0 ATP synthase (Yan *et al.*, 1994) corresponds to the bacterial subunit c, which is also called proteolipid since it can be extracted from membranes with organic solvents (Folch and Lees, 1951).

This work describes a bacterial F_1F_0 ATP synthase, therefore, the following description of more detailed properties of ATP synthase will focus specifically on the bacterial enzyme, unless otherwise stated (Figure 1.3).

1.2.2. Subunit composition and sequence conservation

In general, ATP synthase genes are well conserved throughout evolution (Boyer, 1997). However, specific considerations can be made to describe the specific level of sequence conservation of the individual subunits.

F_1 subunits α and β share significant homology among themselves (α with α , β with β in different organisms) and between each other (α with β in the same organism). For instance, subunits α and β in bovine are about 20 % identical when comparing their amino acid sequences (Walker *et al.*, 1982). The subunits β from different organisms show exceptionally high sequence homology (i.e. 70 % identity between bovine and *E. coli* subunits β (Runswick and Walker, 1983)). This is likely explained by their determinant role as catalytic subunits (*vide infra*). In contrast, the other F_1 subunits γ , δ and ϵ , which are part of the central and peripheral stalk of the enzyme, show more variation in sequence and size (Boyer, 1997).

In addition, the F_0 subunits show different level of sequence conservation. For instance, in subunit c the C-terminal region plays a central functional role in ion translocation, and is thus highly conserved, while the N-terminal region is not (see Results). The poor level of sequence conservation in the N-terminus of subunit c results in great variation in the length of the sequence of this subunit, ranging from a minimum of 66 amino acids in *Streptococcus pneumoniae* (Tettelin

et al., 2001) to a maximum of 1021 amino acids in *Methanopyrus kandleri* (Lolkema and Boekema, 2003). Subunit α also shows more variation in sequence and size (i.e. 23% overall identity between bovine and *E. coli* subunit α). Last but not least, the stalk subunit of F_0 , subunit b , like those stalk subunits of subcomplex F_1 , is poorly conserved in evolution (i.e. 6% identity between bovine and *E. coli* subunit b (Walker *et al.*, 1987)).

1.2.3. The soluble F_1 subcomplex

1.2.3.1. Function of F_1 subcomplex

The F_1 subcomplex is an ATP-driven rotary motor in which subunit γ rotates against the $\alpha_3\beta_3$ -hexamer, which constitutes the catalytic core. Subunits α and β harbor six nucleotide-binding sites. The three catalytic sites reside in subunits β at the three α - β interfaces, whereas the other three nucleotide binding sites are non-catalytic and located in subunits α . It was determined that the minimal functional unit of the F_1 subcomplex is the $\alpha_1\beta_1$ subcomplex (Boyer, 1997).

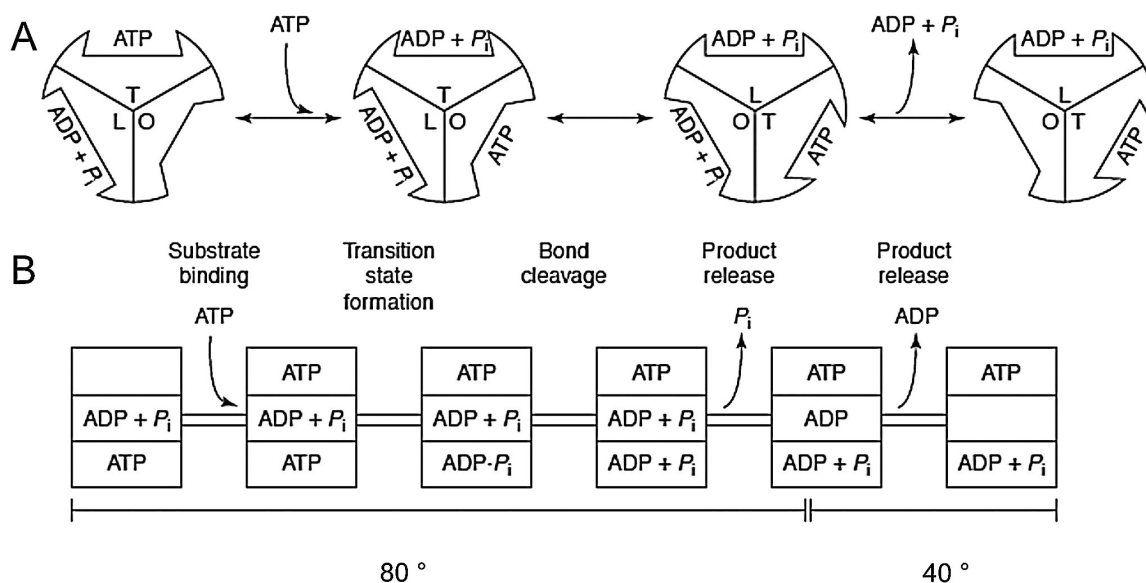


Figure 1.4. The Boyer binding-change mechanism. The figure depicts the mechanism of ATP hydrolysis. (A) Each catalytic site in the three β subunits cycles through three states, tight (T, ATP-bound), loose (L, ADP + P_i -bound) and open (O, empty), which possess different affinity to nucleotides. ATP binds to the O state to convert it into a T state. After hydrolysis, the T state is converted into the L state, from which the products can be released to recover the O state. The concerted switching of states in each of the sites in the subunit β is driven by a 120° rotation of subunit γ . (B) Substeps in the hydrolysis of one ATP molecule. ATP binding in the empty site (top rectangle) leads to the formation of the transition state in the ATP-bound site (bottom rectangle), where hydrolysis occurs. Hydrolysis is followed by the release of P_i first and then of ADP from the ADP+ P_i -bound site (middle rectangle), which becomes empty. Each step is accompanied by two substeps of rotation of subunit γ , first by 80° ('ATP-waiting dwell') and then by 40° ('catalytic dwell') (Shimabukuro *et al.*, 2003). The figure is adapted from (Capaldi and Aggeler, 2002).

1. Introduction

The mechanism of ATP synthesis is cooperative and explained by the binding-change mechanism proposed by Paul Boyer in 1993 (Boyer, 1993). The Boyer binding-change mechanism postulated that ATP synthesis is coupled with the alternation of three conformational states (open, loose and tight) of catalytic subunit β during the reaction cycle. The conformational changes in subunits β are accompanied by the rotation of subunit γ , which is driven by ion-translocation in the F_0 part. The key feature of this hypothesis is that at any given time of the reaction cycle, the three different catalytic sites on subunits β are each in a different conformation and possess each different nucleotide affinity. The open conformation has the lowest affinity for ADP and P_i and is the state that precedes binding of the reaction substrates. The loose conformation has a higher affinity for ADP than for ATP. Finally the tight conformation has the highest affinity for ATP. Each site alternates between the three states as the reaction proceeds. As a consequence of this mechanism, the actual energy-requiring steps for ATP synthesis are the binding of substrates to the open conformation and the release of product from the tight conformation, and not the chemical reaction of ATP synthesis itself (Boyer, 1997). Moreover, this mechanistic hypothesis also implies that the ATP synthase is regulated by a bi-site activation mechanism, meaning that ATP release from a given subunit β is only possible when ADP and P_i are bound to one of the other subunits (catalytic cooperativity of F_1). The Boyer binding-change mechanism is depicted in Figure 1.4.

The Boyer binding-change mechanism hypothesis is strongly supported by the available 3-D structures of F_1 subcomplexes (*vide infra*) and by various biochemical and spectroscopic

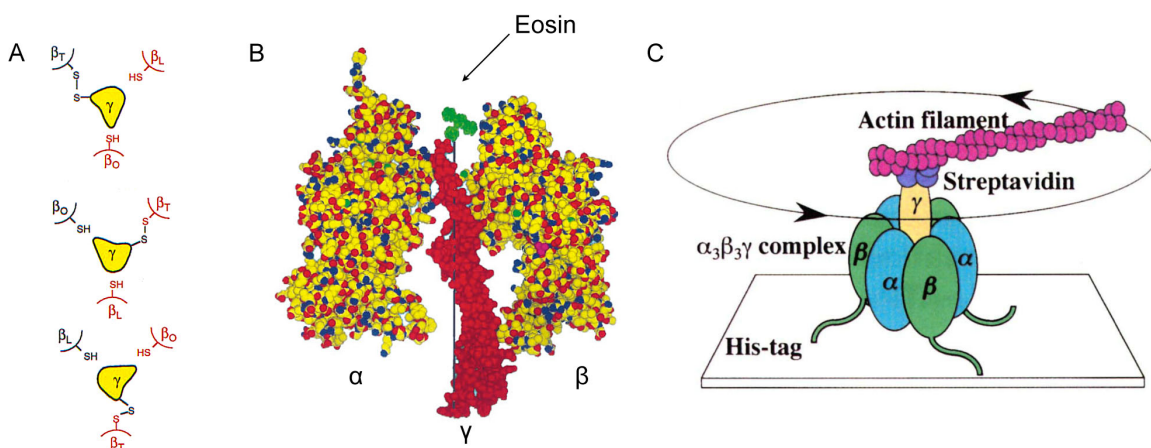


Figure 1.5. Rotation of subunit γ relative to the hexamer $\alpha_3\beta_3$. Rotation of subunit γ was proven for ATP hydrolysis with the following techniques: (A) Crosslinking: schematic diagram showing the cleavage and reformation of disulfide bridges engineered between subunits γ and β during ATP hydrolysis; (B) Polarized absorption recovery after photobleaching: a probe (eosin) is attached to subunit γ and its absorption is recorded keeping the hexamer $\alpha_3\beta_3$ immobilized on an anion-exchange resin; (C) Single-molecule rotation: the rotation of a fluorescent actin filament attached to subunit γ bound to His-tagged hexamer $\alpha_3\beta_3$ immobilized on a resin is directly videographed by fluorescence microscopy. The figure is adapted from Junge *et al.* (1997).

experiments, such as: 1) cross-linking of subunit γ to the C-terminal domain of subunit β which blocks activity of *E. coli* F_1 -ATPase (Duncan *et al.*, 1995), 2) application of polarized absorption recovery after photobleaching to immobilized and γ -fluorescently labeled F_1 -ATPase of chloroplasts (Sabbert *et al.*, 1997), and 3) attachment of a fluorescently labeled actin filament to subunit γ and subsequent observation of the rotation by video fluorescence microscopy of single F_1 molecule (Noji *et al.*, 1997) (Figure 1.5). The latter time-resolved single-molecule rotation experiment also enabled the direct observation of the rotation of subunit γ relative to the hexamer $\alpha_3\beta_3$ (Yasuda *et al.*, 2001). Specifically, this experiment proved that ATP hydrolysis drives the stepped rotation of subunit γ with a period of 120° under saturating concentration of Mg-ATP. In contrast, at substrate-limiting concentrations, two substeps of the rotation mechanism can be distinguished: first, the so-called ‘ATP-waiting dwell’ is accompanied by a rotation of 80° , and second, the ‘catalytic dwell’ is accompanied by a rotation of 40° (Junge *et al.*, 2009).

While subunits α , β and γ possess direct and well-characterized functional roles, F_1 -subunits ϵ and δ are also important. In all ATP synthases, subunit ϵ binds to subunit γ by its N-terminal domain to form the central stalk and is in direct connection to the c-ring of F_O . The role of subunit ϵ is yet to be established precisely, but it is believed that this subunit is involved in the regulation of enzyme activity (Feniouk and Yoshida, 2008), in particular as an inhibitor of ATP hydrolysis (Kato *et al.*, 1997; Kato-Yamada *et al.*, 1999; Keis *et al.*, 2006). In bacteria and in chloroplasts, but not in mitochondria, the C-terminal domain of subunit ϵ functions as a mobile regulatory element that can change conformation from an up-state to a down-state (Tsunoda *et al.*, 2001). When subunit ϵ is in the up-state promoted by a higher $\Delta\mu_{H^+}$ or by ADP binding, the enzyme can catalyze ATP synthesis (Tsunoda *et al.*, 2001). High concentrations of ATP induce the transition to the down-state, in which the enzyme reverses its mode of operation and catalyzes ATP hydrolysis (Suzuki *et al.*, 2003). The presence of these two conformations was also demonstrated by cross-linking experiments for subunit ϵ of *E. coli* (Tsunoda *et al.*, 2001) and of *Bacillus* PS3 (Suzuki *et al.*, 2003). The inhibitory effect of subunit ϵ is caused by the direct electrostatic interaction between the positive charges in the C-terminal region of subunit ϵ and the negative charges in the conserved DELSEED motif of subunits β (Hara *et al.*, 2001). Some bacterial subunits ϵ bind ATP and may act as a built-in cellular sensor of ATP concentration (Kato-Yamada and Yoshida, 2003).

Finally, subunit δ together with the F_O -subunit b dimer form the peripheral stalk of bacterial ATP synthase, connecting F_1 -subunits $\alpha_3\beta_3$ with F_O -subunit a. Besides this structural role, subunit δ has no other known function.

1.2.3.2. Structure of F₁ subcomplex

The first high-resolution 3-D structure of the F₁ subcomplex ($\alpha_3\beta_3\gamma$) was solved at 2.8 Å resolution in 1994 for the ATP synthase from bovine heart mitochondria (Abrahams *et al.*, 1994). Since then, more crystal structures of the F₁ subcomplex were determined in different conformations and with different inhibitors, including aurovertin (van Raaij *et al.*, 1996), efrapeptin (Abrahams *et al.*, 1996), 4-chloro-7-nitrobenzofurazan (NBD) (Orriss *et al.*, 1998), azide (Bowler *et al.*, 2006), N, N'-dicyclohexylcarbodiimide (DCCD) (Gibbons *et al.*, 2000), AlF₃ (Braig *et al.*, 2000), and AlF₄⁻ (Menz *et al.*, 2001). Structures of the F₁ subcomplex from other organisms were also determined, including the native F₁ structure from rat (Bianchet *et al.*, 1998), native (Groth and Pohl, 2001) and tentoxin-inhibited F₁ structures (Groth, 2002) from spinach chloroplast, and the F₁ structure from yeast (Kabaleeswaran *et al.*, 2006; Kabaleeswaran *et al.*, 2009; Dautant *et al.*, 2010). Moreover, structures of F₁ subcomplexes in association with c-rings were solved for yeast (Stock *et al.*, 1999; Dautant *et al.*, 2010) and bovine (Watt *et al.*, 2010) ATP synthases, and the structure of the F₁ subcomplex in association with the peripheral stalk was determined for bovine ATP synthase (Dickson *et al.*, 2006; Rees *et al.*, 2009). Finally, several structures of the bacterial F₁ subcomplex were also determined, including those of *E. coli* (Cingolani and Duncan 2011), *Bacillus* PS3 (Shirakihara *et al.*, 1997), and *Bacillus sp.* TA2. A1 ATP synthases (Stocker *et al.*, 2007). Among all the F₁ structures, the highest resolution obtained was for the F₁ structure from bovine heart mitochondria at 1.9 Å resolution (Bowler *et al.*, 2007).

All these structures show that the overall dimension of the F₁ subcomplex is ~125 Å in height and ~115 Å in diameter. Subunits α and β are arranged as a trimer of heterodimers ($\alpha_3\beta_3$) and are each composed of three domains, an N-terminal β -barrel, a central α -helical- β -sheet domain, and a C-terminal α -helical domain. Furthermore, the structures reveal the architecture of the central nucleotide-binding site domain. The latter is composed of a nine-stranded β -sheet with nine associated α -helices, and is characterized by common nucleotide-binding folds, the Walker A motif (GxxxxGKT/S) and the Walker B motif (R/KxxxGxxxL/VhhhhD) (Walker *et al.*, 1982). The structures of the F₁ subcomplex show the asymmetric features characteristic of the $\alpha_3\beta_3$ hexamers and predicted by the Boyer binding-change mechanism (Figure 1.6). Subunit β is typically trapped in the tight conformation binding to AMP-PNP (a non-hydrolyzable ATP analog) (β_{TP}), in the loose conformation binding ADP (β_{DP}), and in the open, empty conformation (β_E) (Abrahams *et al.*, 1994). Subunit γ extends through the $\alpha_3\beta_3$ hexamers forming two long coiled-coil α -helices and making only limited contacts with subunits α and β . The crystal structures reveal the asymmetry in the contacts between subunit γ and each of the catalytic subunits β (Abrahams *et al.*, 1994).

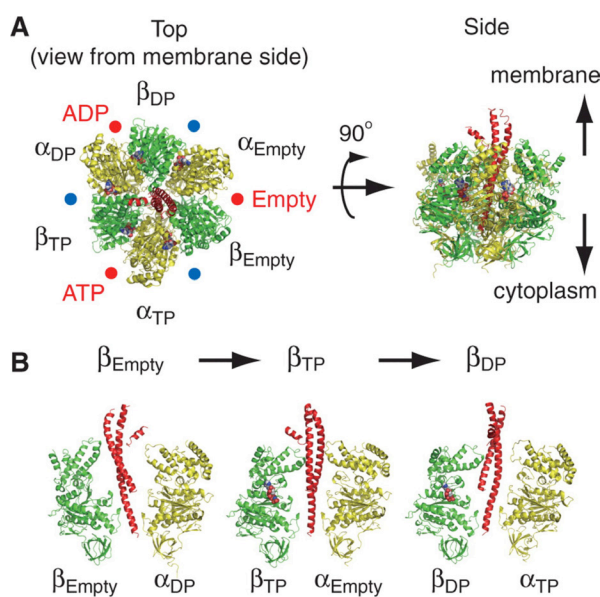


Figure 1.6. Crystal structure of the F_1 subcomplex $\alpha_3\beta_3\gamma$ from bovine mitochondrial ATP synthase (PDB id. 1BMF). Subunits α , β and γ are shown in yellow, green and red, respectively. (A) Three pairs of $\alpha\beta$ heterodimers are arranged around subunit γ . Three catalytic sites (red circles) are in subunit β , at the subunits α/β interface. Each site is occupied by AMP-PNP, ADP, or is empty (β_{TP} , β_{DP} and β_{Empty} , respectively). Each subunit α forming a catalytic site with one subunit β is designated as α_{TP} , α_{DP} and α_{Empty} , respectively. The other three non-catalytic sites in subunit α bind AMP-PNP (blue circles). The protruding part of subunit γ is directed towards the membrane side. (B) Conformational states of subunit β and of the catalytic α/β interface. β_{Empty} is in an open conformation in which the α -helical C-terminal domain is rotated upwards to open the cleft of the nucleotide-binding pocket. Both β_{TP} and β_{DP} have a closed conformation binding the nucleotide. Subunit γ makes limited contacts with α and β , sufficient to generate the torque that causes the conformational changes in subunit β . The figure is adapted from (Okuno *et al.*, 2011).

Subunit ϵ binds to the protruding part of subunit γ (Gibbons *et al.*, 2000). The structure of subunit ϵ is composed of two domains: an N-terminal 10-stranded β -sandwich, and a C-terminal two α -helical hairpin (Wilkens *et al.*, 1995; Uhlin *et al.*, 1997). Two conformations of subunit ϵ in the up and down states, were revealed by comparing the X-ray crystallographic structures, the bovine DCCD-inhibited F_1 -ATPase at 2.4 Å resolution and the *E. coli* $\gamma\epsilon$ -complex at 2.1 Å resolution (Gibbons *et al.*, 2000; Rodgers and Wilce, 2000), and are relevant to understand its regulatory role (Tsunoda *et al.*, 2001). A more recent crystal structure of the *E. coli* F_1 complex revealed the structural basis for the inhibitory role of subunit ϵ to an even greater level of detail, showing that the C-terminal domain of subunit ϵ can adopt a highly extended conformation inserting deeply into the central cavity of the enzyme. In this conformation, subunit ϵ engages interactions with both the rotor and stator subunits preventing the functional rotation (Cingolani and Duncan, 2011). Finally, the structure of ATP-bound subunit ϵ from *Bacillus* PS3 at 1.9 Å resolution provides evidence for the hypothesis that the C-terminal helices of subunit ϵ can undergo an arm-like conformational change from the up-state to the down-state in response to changes in ATP concentration (Yagi *et al.*, 2007) (Figure 1.7).

Finally, subunit δ is located at the top of the $\alpha_3\beta_3$ hexamer, but so far the structure of bacterial subunit δ has not yet been solved by crystallographic methods, either in isolation or in complex with other subunits. It is known that subunit δ forms contacts with both subunit α and β , as confirmed by the EM reconstructions of *E. coli* F_1F_0 ATP synthase decorated with a monoclonal

antibody against subunit δ (Wilkins *et al.*, 2000), by NMR (Wilkins *et al.*, 2005) and by cross-linking data (Lill *et al.*, 1996; Ogilvie *et al.*, 1997; Rodgers and Capaldi, 1998). Cross-linking data also suggested that subunits δ and α interact via their N-terminal region (Lill *et al.*, 1996; Ogilvie *et al.*, 1997; Rodgers and Capaldi, 1998) and that subunits δ and b interact via their C-terminal region (McLachlin *et al.*, 1998). Instead, there is structural information for the mitochondrial homologue of subunit δ , namely subunit OSCP. The latter, which is predicted to be very similar to the bacterial subunit δ (Walker and Dickson, 2006), possesses an N-terminal domain consisting of a six- α -helix

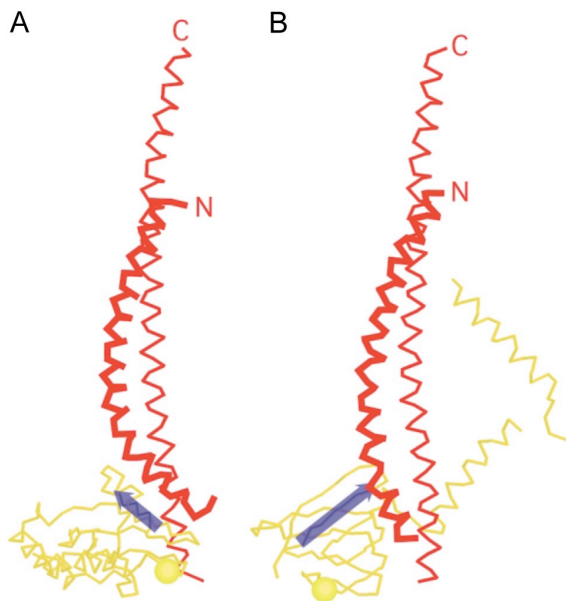


Figure 1.7. The two conformational states of subunit ϵ . Subunits γ and ϵ from structures 1E79 (A) (Gibbons *et al.*, 2000) and 1FS0 (B) (Rodgers and Wilce, 2000) are shown in red and yellow, respectively. Subunit ϵ is composed of an N-terminal β -sandwich and a C-terminal two α -helical hairpin. The latter can change from a down-state conformation (A) to an up-state conformation (B) to inhibit ATP hydrolysis (blue arrow). The figure is adapted from Capaldi and Aggeler (2002).

bundle, as revealed by NMR spectroscopy (Wilkins *et al.*, 1997; Wilkins *et al.*, 1997; Carbajo *et al.*, 2005). Recently, the structure of the complex between bovine F_1 -ATPase and a stator subcomplex has been determined at a resolution of 3.2 Å (Rees *et al.*, 2009). The structure reveals that the N-terminal domain of OSCP links the stator with F_1 -ATPase via α -helical interactions with the N-terminal region of subunit α_E . Its C-terminal domain makes extensive helix-helix interactions with the C-terminal α -helix of subunit b from residues 190–207. The linker region between the two domains of OSCP also appears to be flexible, enabling the stator to adjust its shape as the F_1 domain changes conformation during the catalytic cycle. A similar association may also occur between the subunit δ and the stator in the bacterial enzyme.

1.2.4. The membrane-embedded F_0 subcomplex

1.2.4.1. Function of F_0 subcomplex

The ATP synthase F_0 -subcomplex translocates ions (H^+ or Na^+) across the membrane using an electrochemical gradient to generate a rotary torque. Ion translocation takes place at the interface of subunits a and c. Based on mutagenesis studies, highly conserved residues Asp or Glu in subunit c (i.e. Asp61 in *E. coli*) and Arg in subunit a (i.e. Arg210 in *E. coli*) are directly involved in ion

translocation (Vik and Antonio, 1994; Cain, 2000).

The most widely accepted model for ion translocation is the so-called half-channel model, first proposed by Wolfgang Junge (Junge *et al.*, 1997). According to this model, subunit a interacts with two subunits c at every rotation step and this interaction opens two access channels for ions, one reaching from one side of the membrane to the middle, the other from the middle to the opposite side of the membrane. When an ion enters the former half channel, it binds to the carboxyl group of the functional Asp/Glu residue in subunit c. Neutralization of this negatively charged residue allows subunit c to rotate away from subunit a. Upon such rotation, the neighboring subunit c approaches subunit a and releases its bound ion through the second half channel. After ion release, the functional Asp/Glu residue returns to a negatively charged state and forms an ionic contact with Arg in subunit a to start a new translocation cycle. The direction of rotary motion is determined by the ion concentration gradient established by the respiratory complexes between the two sides of the membrane (see above and Figure 1.8).

1.2.4.2. Structure of F_0 subcomplex

A high-resolution structure of the complete F_0 complex is not yet available, which prevents a clear understanding of the ion-translocation mechanism. However, the general architecture of the F_0 subcomplex is known from low resolution electron microscopy and AFM experiments, and from biochemical data. The simplest bacterial F_0 motor is a membrane-inserted protein complex consisting of an oligomeric c-ring adjacent to subunit a and to the homodimer of subunit b.

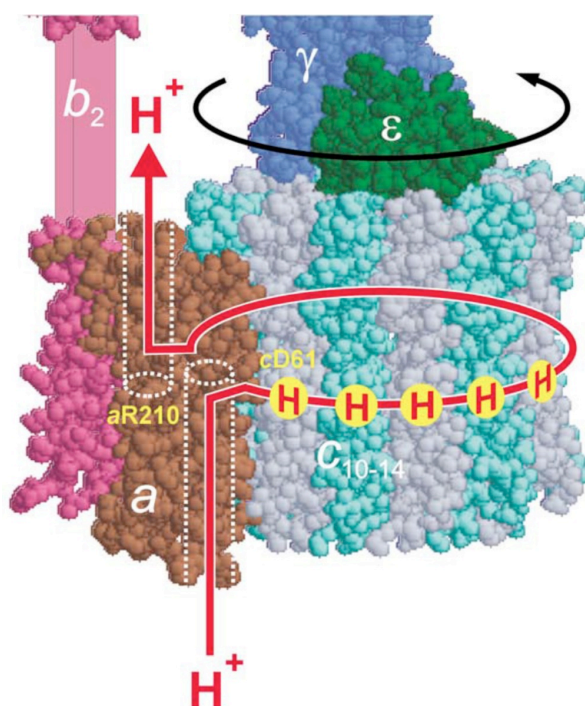


Figure 1.8. Proposed half-channel model for proton translocation in *E. coli*. Highly conserved residues Asp61 (cD61) of subunit c (grey and blue for the neighboring c monomers) and Arg210 (aR210) of subunit a (brown) are located in the middle of the membrane bilayer at the interface between subunits a/c, where they interact enabling the proton translocation. Putative access channels for proton translocation facing the periplasm and with the cytoplasm are shown as white dotted lines (the proton moves from cytoplasm to periplasm when ATP is hydrolyzed and vice versa when ATP is synthesized). During each cycle, the electronic c-ring rotor turns counterclockwise when ATP is hydrolyzed and clockwise when ATP is synthesized. This rotation is coupled to the rotation of the central stalk subunits γ and ϵ (black arrow) that generates torque to power the chemically active F_1 motor by elastic mechanical-power transmission (not shown). The figure is adapted from Weber and Senior (2003).

1. Introduction

The subunit b homodimer together with subunit δ of the F_1 subcomplex forms the so-called peripheral stalk. The interaction between subunits b and δ provides an important connection between the F_0 and the F_1 subcomplexes, which is then further reinforced by the interactions between F_1 subunits γ and ϵ with F_0 subunits c (Walker *et al.*, 1982; Walker and Dickson, 2006). Subunits b are typically between 130 and 180 amino acid residues in length. They are characterized by four regions (Dunn *et al.*, 2000): an N-terminal hydrophobic membrane-spanning region (residues 1-24), a tether region (residues 25-52), a dimerization region (residues 53-122) and a C-terminal binding region. The latter three regions are predominantly hydrophilic, which protrude from the membrane in a highly charged α -helical structure (Walker *et al.*, 1982), and interact with the F_1 part via subunits δ and α (Dunn *et al.*, 2000; Walker and Dickson, 2006). The structure of the N-terminal membrane-spanning region (residues 1-34) of subunit b from *E. coli* was solved by NMR in an organic solvent. The NMR structure showed that residues 4 – 22 form an α -helix, and probably represents the membrane anchor of this subunit (Dmitriev *et al.*, 1999). Cross-linking data showed that the two transmembrane helices in the b_2 homodimer are in close proximity (Dmitriev *et al.*, 1999) and interact with subunits a and c (Fillingame *et al.*, 2000). The structure of residues

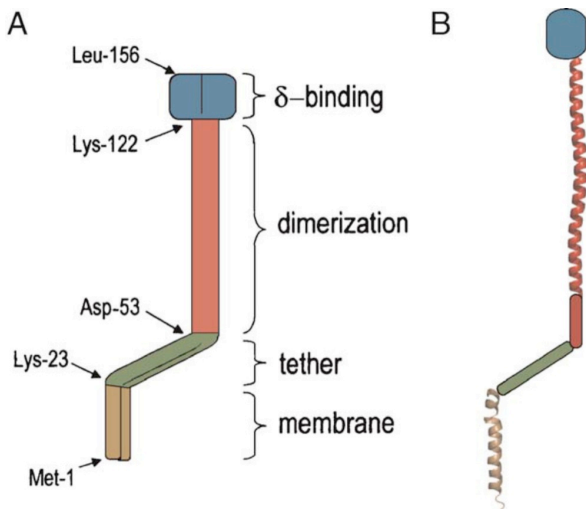


Figure 1.9. Proposed structure of *E. coli* subunit b. (A) Subunit b is composed of four regions: an N-terminal membrane-spanning region, a tether region, a dimerization domain and a C-terminal δ -binding region. Residues at the junctions between the different regions are labeled. (B) Composite structure model of a monomer of subunit b. The model was drawn from the NMR structure of the membrane domain (PDB id. 1B9U), the X-ray crystallographic structure of the dimerization domain (PDB id. 1L2P) and manually drawn rods that represent the tether and δ -binding regions. The figure is adapted from Walker and Dickson (2006).

62-122, which contains the so-called “dimerization domain”, has been solved by X-ray crystallography at 1.5 Å resolution (Del Rizzo *et al.*, 2002). Combined with small-angle X-ray scattering data, the structure indicates that this region forms a right-handed coiled-coil. Finally, the subunit δ -binding region is reported to have a more globular structure (Dunn *et al.*, 2000) (Figure 1.9). Detailed structural information about the soluble part of the peripheral stalk has recently become available for bovine (Dickson *et al.*, 2006; Rees *et al.*, 2009) and bacterial V-type ATP synthase (Lee *et al.*, 2010). Both structures confirmed that the stalk possesses a predominantly helical structure, even though the peripheral stalk of bacterial ATP synthase adopts a straighter conformation than that of its bovine counterpart (Del Rizzo *et al.*, 2002; Del

Rizzo *et al.*, 2006).

The highly hydrophobic subunit a is the largest polypeptide of the F_O complex. It has been studied most extensively in *E. coli*, but unfortunately this subunit is extremely difficult to overproduce *in vivo* (von Meyenburg *et al.*, 1985; Eya *et al.*, 1989), and highly unstable in isolation (Akiyama *et al.*, 1996), and thus it has not been possible to crystallize it. Therefore, besides information on the functional role of the conserved Arg residue described above, little information about subunit a is available. Surface labeling studies using cysteine mutants led to the proposal that *E. coli* subunit a, which is 271 amino acids long, consists of five transmembrane helices (aTMH1-5), with the N-terminal loop facing the periplasm and a C-terminal loop facing the cytoplasm (Vik *et al.*, 2000; Zhang and Vik, 2003). Cross-linking studies have shown that residues 74 and 91 of the first cytoplasmic loop are in close proximity to subunit b (Long *et al.*, 2002). Moreover, cross-linking data also revealed that aTMH2, 3, 4 and 5 form a four-helix bundle and that key functional residues are located in aTMH4 (Ser206, Arg210, and Asn214) near the periplasmic end and close to subunit c (Schwem and Fillingame, 2006). On the basis of an extensive cysteine scanning analysis and accessibility studies with cysteine-reactive chemicals (*N*-ethylmaleimide, Ag⁺), two distinct aqueous access pathways were mapped to aTMH2-5 (Angevine *et al.*, 2007; Steed and Fillingame, 2009). Other experiments indicated that residues Glu219 and His245 are in close proximity to each other and possibly near the periplasmic entrance site for the substrate ion (Cain and Simoni, 1986; Cain and Simoni, 1988; Lightowlers *et al.*, 1988; Hartzog and Cain, 1994). At present, the highest resolution structural information pertaining to subunit a was obtained by cryo-EM single particle analysis on *Thermus thermophilus* V-type ATPase, which was visualized at 9.7 Å resolution (Lau and Rubinstein, 2010; Lau and Rubinstein, 2012). This results showed that the contact interface of subunit a with subunit c is rather minimal, limited to a narrow region in the middle of the membrane, and possibly mediated by lipid molecules (Lau and Rubinstein, 2010; Lau and Rubinstein, 2012) (Figure 1.10).

Similar to subunit a, subunit c is also highly hydrophobic. The structure of subunit c monomers was first investigated by biochemical methods (Fillingame *et al.*, 1990) and NMR spectroscopy (PDB id.:

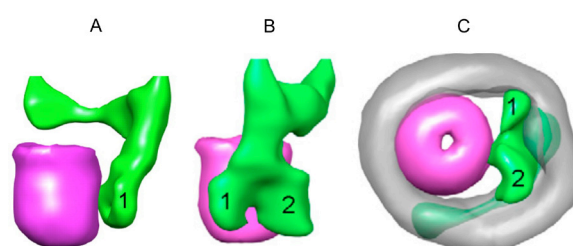


Figure 1.10. The interface between rotor and stator in F_O-ATP synthase. The picture shows three different views of V₀/A₀ subcomplex of *Thermus thermophilus* V-type ATP synthase obtained by EM (Lau and Rubinstein, 2010). The rotor ring, subunit L (homologous to subunit a in F₁F₀ ATP synthase) and the surrounding detergent micelle are shown in magenta, green and grey, respectively. (A) and (B) Views along the membrane plane. (C) View from the cytoplasmic side of the membrane. The figure is reproduced from Lau and Rubinstein (2010).

1A91) (Girvin *et al.*, 1998), which determined that generally subunit c is composed of a hairpin of two transmembrane helices connected by a short highly conserved cytoplasmic loop (RQPE) responsible for interacting with F₁-subunits γ and ϵ (e.g. Pogoryelov *et al.* (2008)). Both the N- and C-terminal ends of subunit c are in the periplasm. In the context of the assembled F₁F₀ ATP synthase, subunit c monomers associate to form oligomeric rings (Meier 2003). Crystal structures available for different rings of subunit c show that these share a high similarity, possessing a cylindrical shape formed by an inner ring of N-terminal helices concentric to an outer ring of C-terminal helices, and by a central pore, filled with phospholipids to prevent ion leakage through the membrane (Meier *et al.*, 2001). Some c subunits form hourglass-shaped cylinders because their transmembrane helices are bent (i.e. in yeast, *Ilyobacter tartaricus*, *S. platensis* (Stock *et al.*, 1999; Meier *et al.*, 2005; Pogoryelov *et al.*, 2005)). Whereas, others form tulip-beer glass-shaped rings as their helices are straighter (i.e. in *Bacillus pseudofirmus* OF4 (Preiss *et al.*, 2010)).

Interestingly, c-rings possess constant stoichiometry in each given species, but different stoichiometries in different organisms. For example, cross-linking data together with atomic force

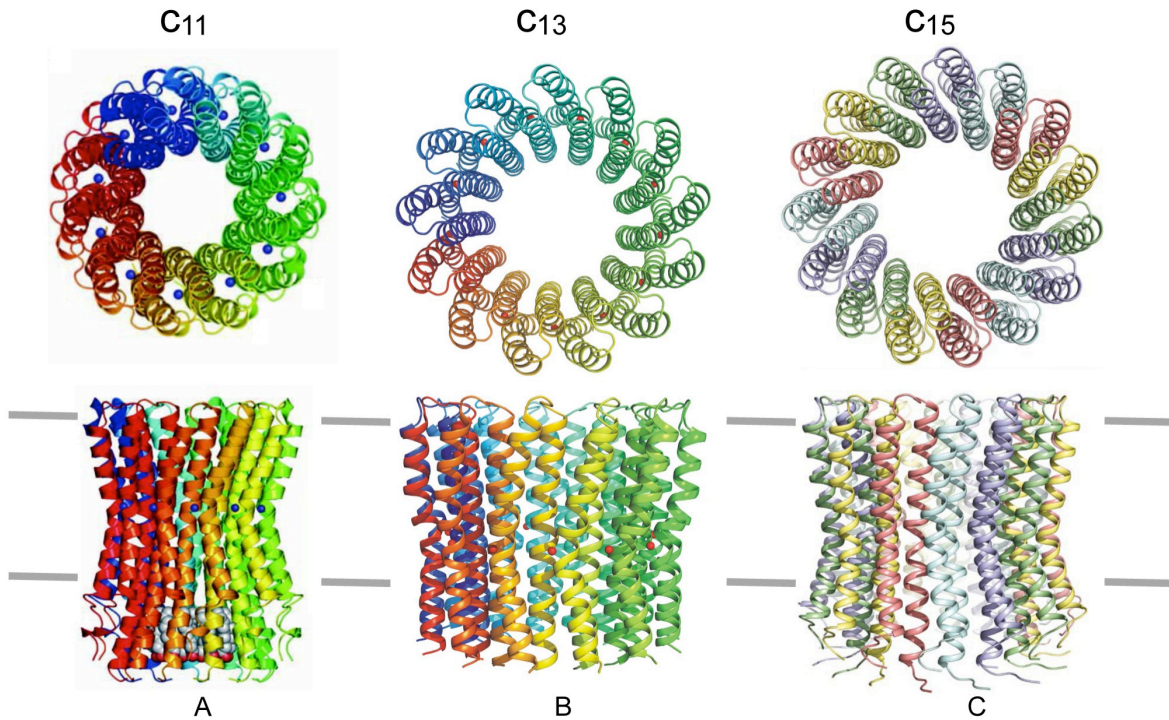


Figure 1.11. C-ring structures determined by X-ray crystallography. (A) *I. tartaricus* c₁₁ ring (left, PDB id. 1YCE, (Meier *et al.*, 2005)), (B) *B. pseudofirmus* OF4 c₁₃ ring (middle, PDB id. 2X2V, (Preiss *et al.*, 2010)) and (C) *S. platensis* c₁₅ ring (right, PDB id. 2WIE, (Pogoryelov *et al.*, 2009)). The figure represents the view of the c-rings from the cytoplasm (top) and parallel to the membrane plane (bottom).

microscopy (AFM) (Singh *et al.*, 1996; Takeyasu *et al.*, 1996) and electron microscopic reconstructions (Birkenhager *et al.*, 1995) showed that the c-ring of *E. coli* F₁F₀ ATP synthase is composed of 10 subunits, similarly to that of yeast ATP synthase (PDB id.: 2XOK and 2WPD Stock *et al.* 1999; Dautant *et al.* 2010). In contrast, the c-ring of the bovine ATP synthase is composed of 8 subunits (PDB id.: 2XND (Stock *et al.*, 1999; Dautant *et al.*, 2010; Watt *et al.*, 2010)), the prokaryotic F-type sodium-pumping ATP synthase from *I. tartaricus* by 11 (PDB id: 1YCE, (Meier *et al.*, 2005)), the proton-pumping ATP synthase of *B. pseudofirmus* OF4 by 13 (PDB id: 2X2V, (Preiss *et al.*, 2010)), and that of the ATP synthase of the cyanobacterium *Arthrospira platensis* by 15 (PDB id: 2WIE, (Pogoryelov *et al.*, 2005)). The stoichiometry of the c-rings is dictated by specific conserved sequence features (i.e. a glycine-motif GxGxGxG), which determines the ion-to-ATP ratio of ATP synthase. This ratio was evolutionarily selected for based on the bioenergetic requirements of each individual species, i.e. the ion motif force generated across the membrane (Pogoryelov *et al.*, 2012). For example, in fungi, eubacteria and plants chloroplasts, c-ring sizes of 10-15 subunits imply that these enzymes need 3.3-5 ions to make one molecule of ATP, while the c₈-ring stoichiometry of bovine ATP synthase implies higher efficiency and a lower bioenergetic cost of 2.7 protons per ATP molecule.

The conserved Asp/Glu residues involved in ion translocations are localized at the middle of the membrane plane on the outer C-terminal helix. In most c-rings structures, the ion binding-site is either in a H⁺ or Na⁺ locked conformation (Meier *et al.*, 2005; Murata *et al.*, 2005; Pogoryelov *et al.*, 2009; Preiss *et al.*, 2010) or in an open conformation bound to the inhibitor N,N'-dicyclohexylcarbodiimide (DCCD) (Pogoryelov *et al.*, 2010; Mizutani *et al.*, 2011). Only one apo-

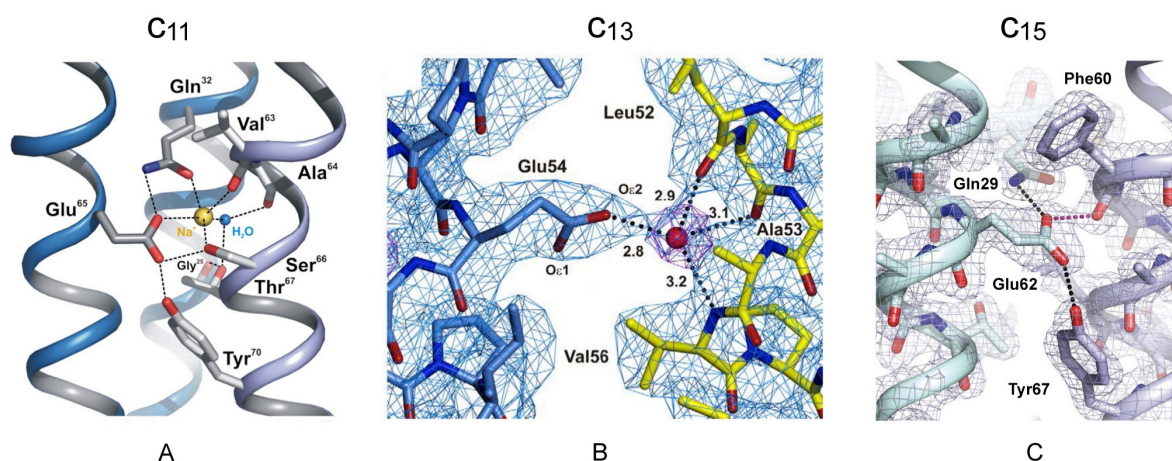


Figure 1.12. Structural details of the ion binding sites of Na⁺- and H⁺-dependent ATP synthases. (A) The Na⁺ binding site in *I. tartaricus* c₁₁ ring (PDB id. 1YCE, (Meier *et al.*, 2005)). (B) The H⁺ binding site in *B. pseudofirmus* OF4 c₁₃ ring (PDB id. 2X2V, (Preiss *et al.*, 2010)). (C) The H⁺ binding site in the *S. platensis* c₁₅ ring (PDB id. 2WIE, (Pogoryelov *et al.*, 2009)).

structure is in an open conformation (Symersky *et al.*, 2012). Besides the conserved Asp/Glu residue, other residues are also involved in ion coordination but these vary from species to species. For example, in the c-15 ring from *S. platensis*, the proton is coordinated by Gln29, Tyr67, and Phe60, while in the c-13 ring from *B. pseudofirmus* OF4, a water molecule coordinated by Glu54 from one helix and Leu52, Ala53 and Val56 from the adjacent helix coordinates the proton. The crystal structure of *I. tartaricus* c-11 ring shows that the binding signature for Na⁺ coordination involves four amino acids (Gln32, Glu65, Ser66, and Val63) and one structural water coordinated by Thr67 (Meier *et al.*, 2009). Thr67 is the only residue involved in binding that distinguishes Na⁺ ATP synthases from H⁺ ATP synthase (Meier *et al.*, 2009). An almost identical Na⁺ coordination was found in the crystal structure of the K-ring of the V-type ATPase from *Enterococcus hirae* (Murata *et al.*, 2005) (Figure 1.11 and 1.12).

1.2.5. ATP synthase biogenesis and assembly

Biogenesis of the F₁F₀ ATP synthase has been studied in *E. coli* and yeast, and a similar coordinated stepwise assembly mechanism was proposed. Assembly starts with the formation of subcomplex F₁, and it is then followed by a sequential assembly of the membrane F₀ subcomplex, and finally, by the coupling of the subcomplexes F₀ to F₁ (Price and Driessen, 2010). Interestingly, the F₁ and F₀ sectors are not yet fully assembled before coupling occurs (Steffens *et al.*, 1984) (Figure 1.13).

In the assembly process, proteins of the Oxa family, namely YidC in *E. coli*, its homologues SpoIIIJ and YqjG in *Bacillus subtilis* (Saller *et al.*, 2009), its homologue Alb4 in chloroplasts of *Arabidopsis thaliana* (Benz *et al.*, 2009) and its homologue Oxa1 in mitochondria (Jia *et al.*, 2007), facilitate the insertion of the membrane subunits into the membrane and act as chaperones to support the correct formation of the ATP synthase complex.

A particularly critical step in the membrane biogenesis of ATP synthase is the assembly of the c-rings, a step that may even determine the total amount of active ATP synthase in mammals (Houstek *et al.*, 2006). Besides YidC, protein AtpI may be involved in the assembly of c-rings (Suzuki *et al.*, 2007; Brandt *et al.*, 2013). In bacterial genomes, the *atpI* gene is the first gene in the *atp* operon (Walker *et al.*, 1984; Santana *et al.*, 1994; Yokoyama *et al.*, 2000; Meier *et al.*, 2003; Keis *et al.*, 2004; Brandt *et al.*, 2013), but the protein AtpI is not present in the assembled holoenzyme. The AtpI's chaperone role in the assembly of the c-rings was proposed based on work on a hybrid F₁F₀ complex (F₁ from *Bacillus* PS3 and F₀ from *Propionigenium modestum*) expressed in *E. coli* (Suzuki *et al.*, 2007), and on work on *Bacillus* PS3 and *Acetobacterium woodii* ATP synthases (Suzuki *et al.*, 2007; Ozaki *et al.*, 2008; Brandt *et al.*, 2013).

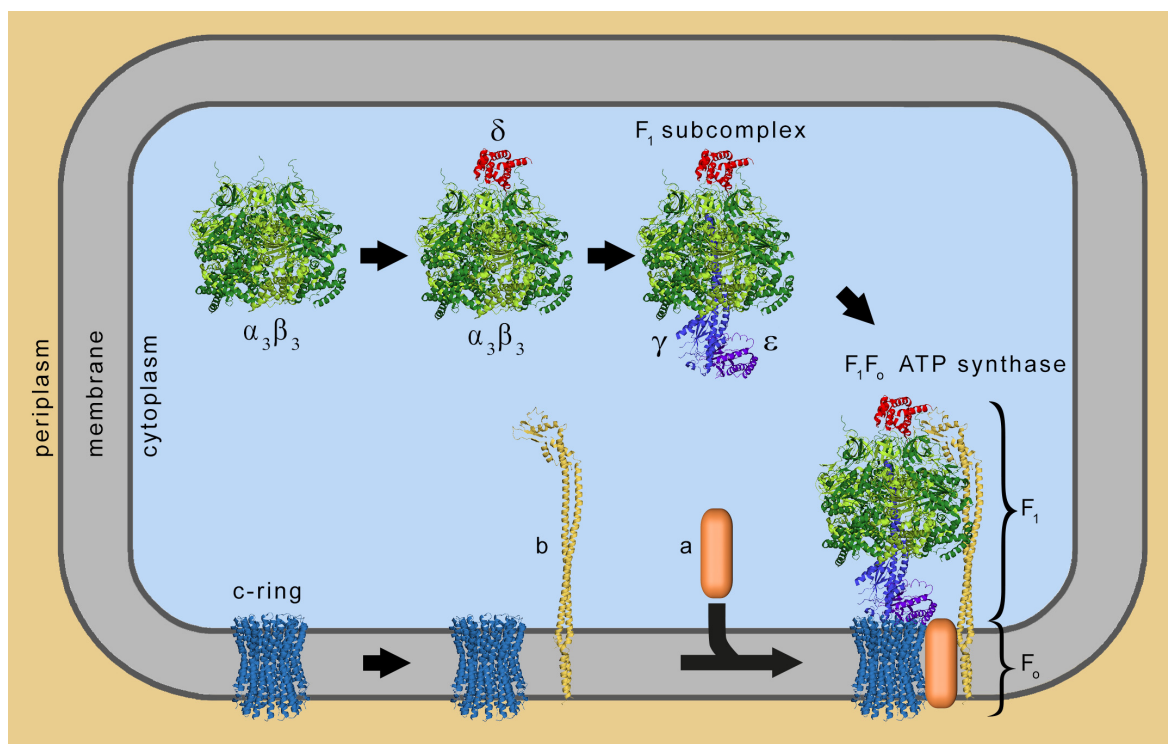


Figure 1.13. Assembly pathway for the F₁F₀ ATP synthase. The assembly of F₁ occurs in the cytoplasm where the subunits α , β , γ , δ , and ϵ form a globular structure (top). The assembly of F₀ occurs in a second assembly pathway (down), in which subunit c forms an oligomeric c-ring in the membrane onto which subunit b assembles. In the final step, subunit a is added to the complex $\alpha_3\beta_3\gamma\epsilon cb_2$. The structures used to depict the figure are F₁- $\alpha_3\beta_3\gamma\epsilon$ subcomplex from bovine (PDB id. 1E79), subunit δ from *E. coli* (PDB id. 2A7U, N-terminal segment only), the peripheral stalk (representing subunit b) from *T. thermophilus* (PDB id. 3V6I) and c-ring from *I. tartaricus* (PDB id. 1CYE). The figure was adapted from (Dalbey *et al.*, 2011) by Paolo Lastrico (MPI of Biophysics, Frankfurt, Germany).

Also subunits a and b require additional factors besides YidC for their membrane insertion and assembly. In *E. coli*, it was proposed that the Sec translocon and the signal recognition particle (SRP) are involved (Yi *et al.*, 2004). The membrane insertion of subunit a requires that both subunits b and c are present in the *E. coli* membrane (Hermolin and Fillingame, 1995). It was shown that subunit a is actually not essential for the correct assembly of stable F₁F₀ complexes in *Bacillus* PS3 and *E. coli* (Vik and Simoni, 1987; Krebstakies *et al.*, 2005), and it can be added after purification to restore the activity of subunit-a-deprived ATP synthase complexes (Ono *et al.*, 2004). Therefore, it is plausible that this protein is added at the last step of ATP synthase biogenesis (Price and Driessen, 2010).

Finally, protein factors may also be required for the assembly of components of the mitochondrial F₁ subcomplex. For instance, in yeast, chaperones Atp12 and Atp11p were identified to bind subunits α and β (Wang and Ackerman, 2000; Wang *et al.*, 2000).

1.3. *Aquifex aeolicus* ATP synthase

Aquifex aeolicus VF5 is a hyperthermophilic bacterium, which grows optimally at 85 – 95 °C. It is one of the earliest diverging bacteria based on phylogenetic analysis of the 16S ribosomal RNA sequence. The cells of this bacteria are rod-shaped, with a length of 2 – 6 μm and a diameter of around 0.5 μm (Huber and Eder, 2006).

A. aeolicus is an obligate chemolithoautotroph, and it uses inorganic carbon as a source for both biosynthesis and inorganic chemical energy (Deckert *et al.*, 1998). It is an obligate aerobic bacterium, but it can grow at very low concentrations of oxygen (as low as 7.5 ppm). *A. aeolicus* is cultured at 85 °C in liquid medium SME under an H₂/CO₂/O₂ (79.5:19.5:1.0), making its cells very expensive and non-manipulatable (Huber and Eder, 2006). *A. aeolicus* preferentially produces cellular energy through the so-called “Knallgas” reaction, where molecular hydrogen is oxidized to produce H₂O using molecular oxygen as the electron acceptor (Deckert *et al.*, 1998). Finally, *A. aeolicus* fixes CO₂ through the reductive tricarboxylic acid cycle (Deckert *et al.*, 1998).

The complete genome of *A. aeolicus* was sequenced in 1998 (Deckert *et al.*, 1998) and it is rather small (1,551,335 base pairs), corresponding to only one-third of the *E. coli* genome. Two pronounced features characterize *A. aeolicus* genome. First, the genome is densely packed: most

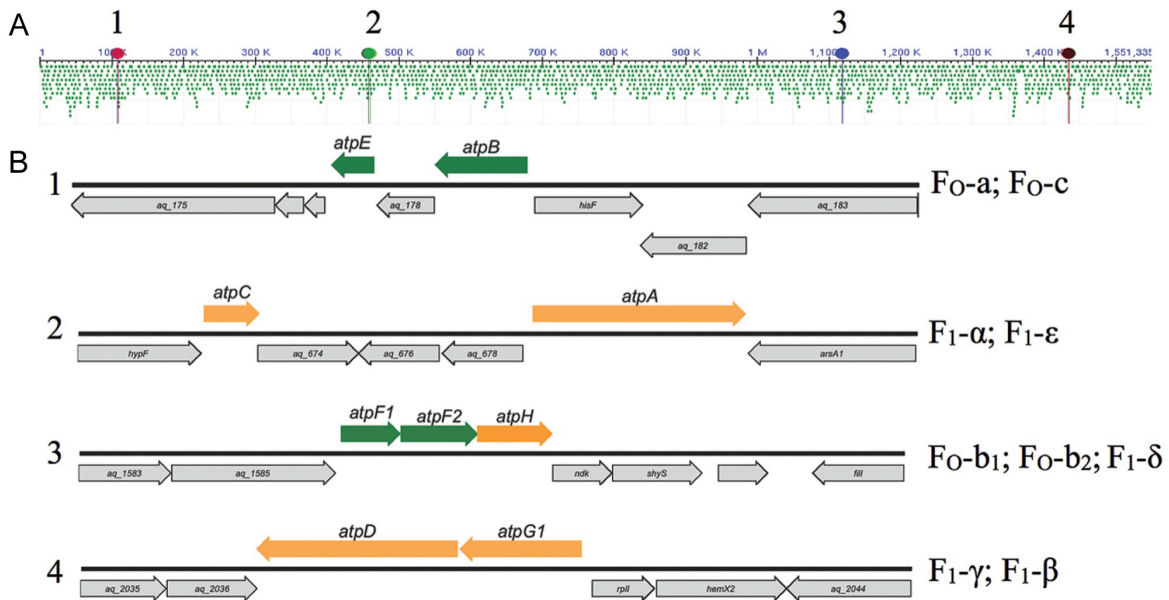


Figure 1.14. Organization of the F₁F₀ ATP synthase genes in *A. aeolicus*. (A) View of the whole genome of *A. aeolicus*. The nine genes encoding the ATP synthase subunits are split into four loci in the genome. (B) Enlarged view of the four loci. The nine *atp* genes are separated into six DNA fragments located in both the plus and minus DNA strands. Only *atpF1F2H* and *atpDG1* form operons. Genes encoding F₁ subunits are shown in orange, those encoding F₀ subunits are shown in green.

genes are apparently expressed in polycistronic operons and many convergently transcribed genes overlap slightly. Second, many genes that are functionally grouped within operons in other organisms are dispersed throughout the *A. aeolicus* genome. This is the case even for genes encoding subunits of the same enzyme. Specifically, this is true for the nine genes encoding subunits of ATP synthase (Deckert *et al.*, 1998; Peng *et al.*, 2006), which are distributed in four different genomic loci. At locus 1, the gene *atpB* encoding subunit a and the gene *atpE* encoding subunit c are separated by *aq_178* encoding a universal stress protein. At locus 2, the gene *atpC* encoding subunit ϵ and the gene *atpA* encoding subunit α are separated by three genes encoding three hypothetical proteins, AQ_674, AQ_676 and AQ_678. At locus 3, a native operon includes the genes *atpF1*, *atpF2* and *atpH* encoding the peripheral stalk subunits b_1 , b_2 and δ , respectively. These three genes are overlapping. Finally, at locus 4, another operon includes the genes *atpD* and *atpG* encoding subunits γ and β , respectively (Figure 1.14).

A. aeolicus F₁F₀ ATP synthase has been studied previously after isolation from the native membranes of the organism (Peng *et al.*, 2006). Strikingly, that work determined that *A. aeolicus* ATP synthase differs from the F₁F₀ ATP synthases of other non-photosynthetic organisms, because it possesses a hetero-, and not homo- dimeric peripheral stalk, composed of subunits b_1 and b_2 (Peng *et al.*, 2006). Furthermore, that work provided a first low resolution visualization of the enzyme by single particle electron microscopy raising the question whether its two stalks may be tilted and/or kinked (Peng *et al.*, 2006). Finally, another striking property of *A. aeolicus* ATP synthase is that its subunit c possesses a short, unusually positively charged N-terminal region (see Results). Despite this previous study, many properties of *A. aeolicus* ATP synthase still need to be determined. For instance, the stoichiometry of its c-ring is not yet characterized, its enzymatic properties were not studied, and its high resolution 3-D structure has not been obtained yet.

1.4. Open questions

Despite the wealth of functional and structural information on ATP synthase, several questions are still open for fully understanding this enzyme at a molecular level (Houstek *et al.*, 2006; Junge *et al.*, 2009). From a functional perspective, full kinetic and mechanistic characterization of the different ATP synthase reactions is needed. Particularly, it is important to integrate and make uniform all data obtained so far using enzymes from different organisms. Possibly, new theoretical approaches will need to be developed to extend the description of ATP synthase activity from the nanosecond time scale to the millisecond time scale, thereby explaining the activity of different types of ATP synthases in the context of cellular physiology. From a structural perspective, the mechanism of how ATP synthase assembles from its individual subunits needs to be understood better. Furthermore, a high-resolution structure of the complete F₀ subcomplex is required to

visualize the ion-translocation pathway. Finally and most importantly, a high-resolution structure of the complete F_1F_0 ATP synthase is needed.

1.5. Aim of this work

In order to contribute to a more complete understanding of the structure and function of ATP synthase, and to complement the characterization of native *A. aeolicus* ATP synthase (hereafter AAF₁F₀) described above (Peng *et al.*, 2006), in this work we set out to produce *A. aeolicus* F₁F₀ ATP synthase in the heterologous host *E. coli* (hereafter EAF₁F₀).

This work specifically aimed to: i) clone individual subunits of *A. aeolicus* ATP synthase in *E. coli*; ii) produce, purify and functionally characterize those subunits in isolation or in the context of small soluble and membrane-inserted subcomplexes; iii) design and clone an artificial operon encoding all subunits of *A. aeolicus* ATP synthase; iv) express such operon in *E. coli*; v) purify the fully assembled EAF₁F₀ from the membranes of *E. coli*; and vi) undertake structural and enzymatic studies to assess its overall morphology and its functionality.

Heterologous production of such a large, multi-subunit, membrane-inserted complex as *A. aeolicus* ATP synthase is a challenging task. However, producing *A. aeolicus* ATP synthase in a heterologous genetic system would allow genetic and functional studies to characterize the unusual properties of its subunits, i.e. the heterodimerization of subunits b₁ and b₂, the unusual N-terminal segment of subunit c, or the putatively tilted structure of its stalks, and would open up the possibility to undertake experiments to support the ongoing structural investigation on the native enzyme. Moreover, to a broader extent, it will also serve in the future as a solid reference for designing strategies aimed to co-express large multi-subunit complexes with complicated stoichiometry and whose genes that are not grouped in an operon.

2. Results

To follow up on a previous report about *Aquifex aeolicus* ATP synthase extracted and purified from native cells (AAF₁F₀, (Peng *et al.*, 2006)), in this work we have set out to characterize this important enzyme further. We have pursued three different lines of research. First, we have extended the previous characterization of AAF₁F₀, determining that it is a proton-dependent ATP synthase, studying its enzymatic activity and assessing its stability in different detergents (see chapter 2.1). Second, we have created an expression system to produce the enzyme in the heterologous host *Escherichia coli*, obtaining a purified form of *A. aeolicus* ATP synthase (EAF₁F₀) identical in structure and enzymatic properties to AAF₁F₀ (see chapter 2.2). Finally, we have used our heterologous system to determine important and unique properties of *A. aeolicus* ATP synthase, i.e. that its subunit c possesses a previously uncharacterized N-terminal signal peptide (see chapter 2.3).

2.1. Bioinformatic and functional characterization of AAF₁F₀

2.1.1. Bioinformatic characterization

2.1.1.1. Sequence conservation

Multiple sequence alignments were performed using PSI-BLAST on sets of different homologues of subunits α , β , a, b and c, using sequences obtained from the SwissProt sequence database. A large variation was detected in the lengths of subunits a, b and c, compared to the length of the functional subunits α and β (Table 2.1). Accordingly, when the sequences used in the alignment were clustered using CD-HIT (Fu *et al.*, 2012), the numbers of clusters identified differed dramatically. We identified 114 clusters for subunit b, 55 clusters for subunit a, and 48 clusters for subunit c, but only 3 clusters for subunits α and 1 cluster for subunit β , respectively. This result indicates that during evolution the functionally important subunits α and β were highly conserved whereas the membrane subunits a and c were less conserved, and b₁-b₂ poorly conserved.

Table 2.1: Statistics describing the sequence alignment of selected ATP synthase subunits

Subunit	Number of seq.	Length	Clusters	Clusters/Number of seq.
α	251	498 - 541	3	1.2%
β	251	457 - 574	1	0.4%
a	271	207 - 905	55	20.3%
c	218	66 - 196	48	22.0%
b	276	142 - 443	114	41.3%

2.1.1.2. Subunit a

A. aeolicus F₁F₀ ATP synthase subunit a is encoded by the *atpB* gene with a predicted length of 216 amino acids (M_r = 24.01 kDa). No signal peptide is predicted for subunit a by SignalP. Topology prediction by TMHMM suggests the presence of six transmembrane helices in subunit a. These predictions are contradicted the results of a multiple-sequence alignment performed in this work, which suggests that *A. aeolicus* subunit a possesses 5 transmembrane helices (aTMH1-5, see Figure 2.1). The alignment results showed that aTMH2-5 are conserved: a functionally important arginine (Arg148) is located in aTMH4, together with other conserved residues (Leu149, Asn 152, Ala155 and Gly156). The highly conserved residues Gln195, Leu202 and Tyr 206 are located in aTMH5, and Asn74 and Pro80 are located in aTMH2. The alignment results also suggested that the N-terminal segment of subunit a is variable: (i) there are differences in the length of the N-termini of subunit a from V-type ATP synthases and F-type F₁F₀ ATP synthases: the V-type subunits a typically possess a 50-kDa N-terminal extension, while F-type subunits a are generally shorter (Figure 2.2); (ii) there are also differences in the length and topology of the N-termini of subunits a within F-type ATP synthases: in subunit a from *A. aeolicus* ATP synthase the N-terminus consists of a short N-terminal peptide followed by a transmembrane helix (aTMH1), while in the subunit a from *E. coli* it consists of a long polar peptide of ~ 40 amino acids. Moreover, an additional transmembrane helix (helix-0) is present in the subunit a from *I. tartaricus*, which is therefore composed of 6 transmembrane helices in total (Hakulinen *et al.*, 2012) (see Figure 2.1). Furthermore, the subunit a of nuclear-encoded eukaryotic ATP synthases (i.e. yeast) possesses an N-terminal mitochondrial targeting sequence while the mitochondrial-encoded subunit a of eukaryotic ATP synthases (i.e. that from *Bos taurus*) possesses a short N-terminal peptide (see Figure 2.1). (iii) Finally, the localization of the N-terminal segment is probably also different in ATP synthases from different organisms. For instance, the N-terminal peptide of subunit a from *I. tartaricus* was suggested to be localized in the cytoplasm (Hakulinen *et al.*, 2012), while that of *E. coli* to be located in the periplasm (Vik *et al.*, 2000). A topology model of subunit a from *A. aeolicus* based on the prediction by the software SOSUI suggests that its N-terminal peptide is located in the periplasm.

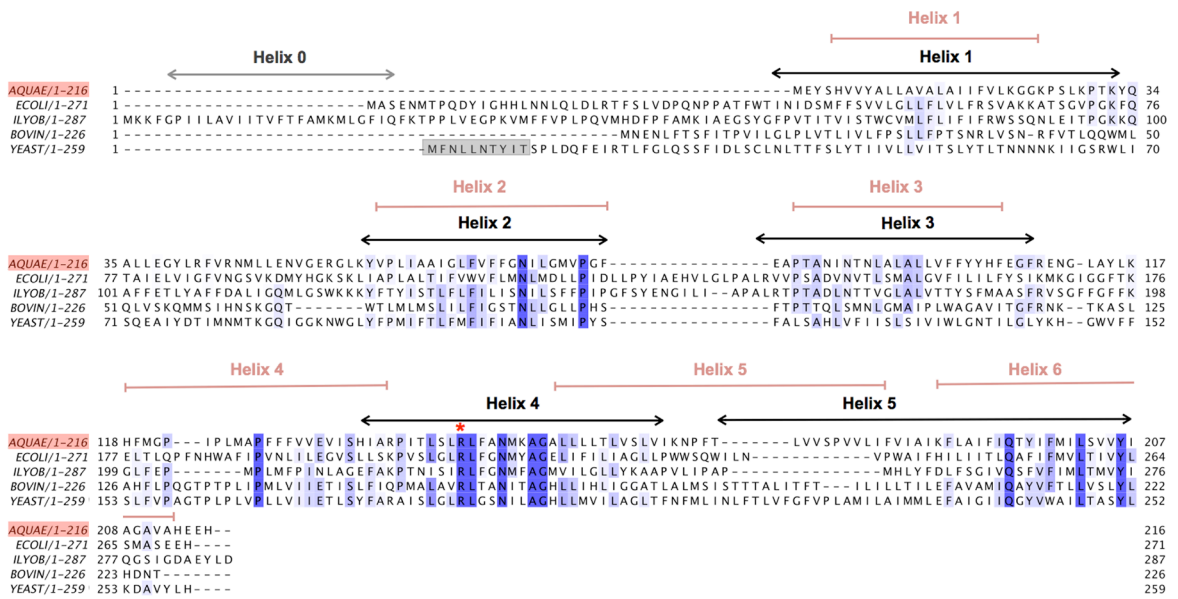


Figure 2.1. Sequence alignment of subunit a of F₁F₀ ATP synthases. Conserved residues are highlighted in blue. The functionally conserved arginine (Arg148 from *A. aeolicus* (AQUAE) subunit a) is marked with a red star. The 5 helices of *E. coli* (ECOLI) subunit a (Helix 1 – 5) are indicated by black arrows (Vik *et al.*, 2000). The six helices of subunit a (Helix 1 - 6) from *A. aeolicus* predicted by TMHMM are indicated by light-red lines. The additional helix 0 of the *I. tartaricus* (ILYOB) subunit a is indicated by a gray arrow (Hakulinen *et al.*, 2012). The mitochondrial targeting peptide of subunit a of yeast ATP synthase is highlighted in gray.

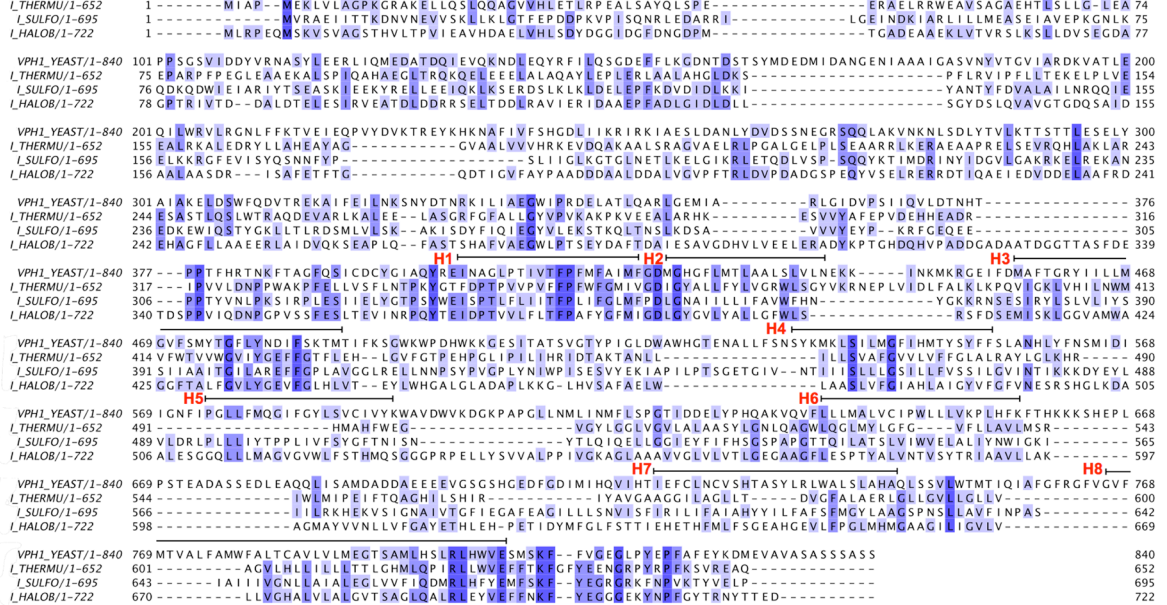


Figure 2.2. Sequence alignment of subunit a of V-type ATP synthases. The C-terminus of subunit a from V-type ATP synthases possesses an 8-helix (H1-H8) bundle (each helix is indicated by black lines). Whereas, the N-terminus is 50 kDa long and is soluble. The sequences from *T. thermophilus* (I_THERMU), *S. acidocaldarius* (I_SULFO), *H. salinarum* (I_HALOB) and yeast (VP11_YEAST) are compared.

2.1.1.3. Subunits b_1 and b_2

A. aeolicus F_1F_0 ATP synthase subunits b_1 and b_2 are encoded by the genes *atpF1* and *atpF2*. They are 144 and 185 amino acids long, with a calculated molecular weight of 16.71 and 21.23 kDa, respectively. Topology prediction by TMHMM suggests the presence of one transmembrane helix in subunit b_1 and of no transmembrane helices in subunit b_2 . These predictions are in contrast to the results of a multiple-sequence alignment performed in this work, which suggests that in *A. aeolicus* both subunits b_1 and b_2 possess all 4 characteristic regions described for the subunit b from *E. coli*: a membrane-anchoring domain, a tether domain, a dimerization domain and an F_1 -binding domain. In other words, the alignment suggests that both subunits b_1 and b_2 possess one putative transmembrane helix. However, considering that the N-terminal 17 amino acids of subunit b_2 are predicted to form a signal peptide by the software SignalP, it is possible that such region is not present in the mature form of the ATP synthase (see Figure 2.3). These observations are very similar to those reported for the ATP synthase of the photosynthetic bacterium *Rhodospirillum rubrum*, for which an N-terminal leader sequence (7 amino acid long) was found to be cleaved off from the mature subunit b' (Falk and Walker, 1988). Besides the presence of this putative N-terminal signal peptide in subunit b_2 , the alignment results also suggest that the N-terminal segment of subunit b is as variable as that of other membrane subunits (i.e. subunits a and c, see chapter

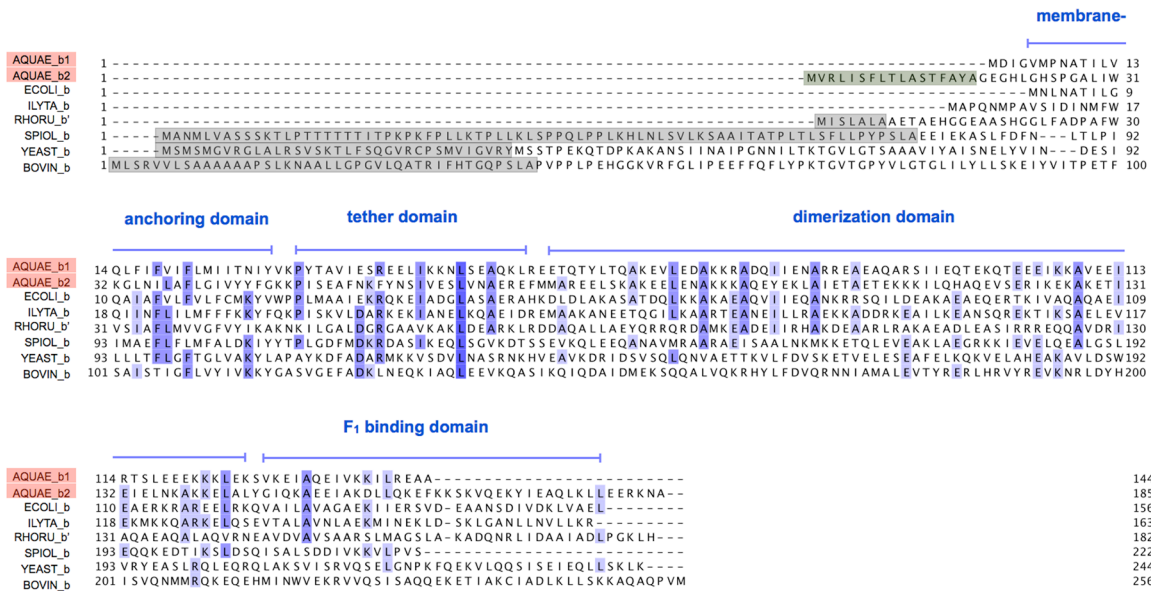


Figure 2.3. Sequence alignment of subunit b of F_1F_0 ATP synthases. Conserved residues are highlighted in blue. The characteristic four regions of subunit b described for *E. coli* are indicated by blue lines. They are: a membrane-anchoring domain, a tether domain, a dimerization domain and an F_1 -binding domain. *A. aeolicus* subunit b_1 (AQUAE_b1) and b_2 (AQUAE_b2) are highlighted in red. The predicted signal peptide of subunit b_2 is highlighted in green. The mitochondrial targeting peptide for subunits b from yeast (YEAST_b) (Velours *et al.*, 1987) and bovine (BOVIN_b) (Walker *et al.*, 1987) mitochondria, the chloroplast targeting peptide for subunit b from spinach chloroplasts (SPIOL_b) (Herrmann *et al.*, 1993), and the cleave-off signal peptide for the subunit b from *R. rubrum* (RHORU_b) (Falk and Walker, 1988) are highlighted in gray.

2.1.1.2 and 2.3.3, respectively). For example a targeting peptide is required for nuclear-encoded mitochondrial (i.e. bovine and yeast) and chloroplast (i.e. spinach and wheat) subunits b to direct the cellular localization of the protein.

2.1.1.4. Subunit c

A more detailed bioinformatics analysis of subunit c is reported in chapters 2.3.2 and 2.3.3.

2.1.2. Functional characterization of native AAF₁F₀

2.1.2.1. Purification of AAF₁F₀

A. aeolicus cells were obtained from Archaeenzentrum (Regensburg, Germany) and treated as described (Peng *et al.*, 2006; Marcia, 2010). AAF₁F₀ was purified as previously reported (Peng *et al.*, 2006). Briefly, *A. aeolicus* membranes were solubilized in the detergent n-dodecyl- β -D-maltoside (DDM) and fractionated on a MonoQ column. Fractions containing AAF₁F₀ were identified by PMF-MS. From these fractions, AAF₁F₀ was further purified by a polishing SEC step in the detergent DDM. Finally AAF₁F₀ was aliquoted in buffers containing Tris-HCl pH 7.4, 150 mM NaCl, 10 mM MgCl₂, 0.05% NaN₃ and one of the following three detergents: 0.05% DDM, 0.2% n-decyl- β -D-maltoside (DM) and 0.2% DM + 0.05% *trans*-4-(*trans*-4'-propylcyclohexyl)cyclohexyl- α -D-maltoside (α -PCC), respectively. Instead, other detergents (i.e. glucoside detergents, such as OG) are not suitable (data not shown). The sample was either used immediately or stored at 4 °C. The quality of the preparation was comparable to that of our previous report (Peng *et al.*, 2006), as judged by size exclusion chromatography (SEC) (Figure 2.4 A), SDS-PAGE (Figure 2.4 B), and electron microscopy (EM) (see Figure 2.4 C). BN-PAGE and SEC suggested that a subcomplex formed by the hetero-dimer $\alpha_1\beta_1$ tends to dissociate from

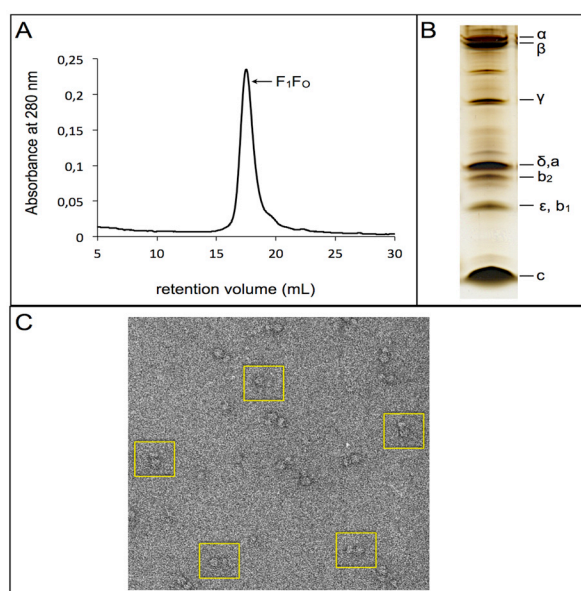


Figure 2.4. Characterization of AAF₁F₀. (A) The SEC profile of AAF₁F₀ in the detergent DM + α -PCC. (B) Silver-stained SDS-PAGE of AAF₁F₀. The subunits were characterized by PMF-MS. The c-ring dissociated into monomeric subunit c during SDS-PAGE. (C) Negative staining EM of AAF₁F₀. Five typical single molecules of AAF₁F₀ are indicated in yellow boxes.

2. Results

the entire complex when AAF₁F₀ is stored in 0.05% DDM and 0.2% DM. In the buffer containing 0.2% DM, the subcomplex containing the peripheral stalk b₁b₂δ and subunit a also tends to dissociate from the entire complex. The stability of the complex is greatly enhanced by adding 0.05% α-PCC. Therefore, 0.05% α-PCC was used in our preparations (Figure 2.5).

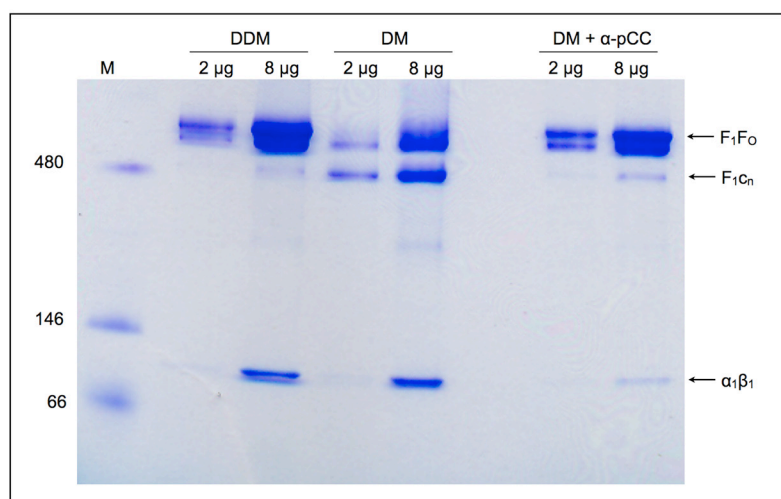


Figure 2.5. Blue-native PAGE of AAF₁F₀ in the presence of different detergents. Subcomplexes α₁β₁ and F₁C_n, and the entire F₁F₀ complex were identified by PMF-MS and their position is indicated by black arrows on the right side of the gel. The gel shows that the stability of AA F₁F₀ in α-PCC is greatly enhanced. A native protein size marker (M) was used as a reference (the molecular weight of the standard proteins is indicated in kDa on the left).

2.1.2.2. AAF₁F₀ is H⁺-dependent, not Na⁺-dependent

We experimentally determined that AAF₁F₀ is a proton and not a sodium ion dependent ATP synthase, as expected from its sequence conservation pattern. To date, Na⁺-translocating ATP synthases rotor rings were found from *I. tartaricus*, *P. modestum*, *A. woodii* and *Enterococcus hirae*. The Na⁺-binding site was identified between each pair of adjacent subunit c (subunit K for *E. hirae*). The Na⁺ binding signature comprises a five-group coordination shell, including four residues (Gln32, Glu65, Ser66 and Val63, numbering is from the subunit c of *I. tartaricus* ATP synthase) and one structural H₂O coordinated by Thr67. Thr67 is the only residue that distinguishes Na⁺-dependent ATP synthases from H⁺-dependent ATP synthases (Meier *et al.*, 2009). In subunit c from *A. aeolicus*, Gln32 is replaced by Met50, Val63 is replaced by Phe81, Ser66 is replaced by Thr84 and most importantly Thr67 is replaced by Ile85. Therefore, the analysis of the amino acid sequence suggests that AAF₁F₀ is a H⁺-dependent ATP synthase (Figure 2.6). We confirmed this hypothesis experimentally by MALDI-TOF MS. We determined that AAF₁F₀ subunit c is modified by *N,N'*-dicyclohexyl-carbodiimide (DCCD). DCCD is a common covalent inhibitor of H⁺-dependent F₁F₀ ATP synthases but not of Na⁺-dependent ATP synthases (Pisa *et al.*, 2007),

because Na^+ and DCCD compete for a common binding site, namely the active-site carboxylate of subunit c (Kluge and Dimroth, 1993). In our study we subjected AAF_1F_0 to DCCD treatment (see 4.2.5.2) in the presence of up to 150 mM NaCl, we extracted subunit c by common procedures (see 4.2.5.1) and we analyzed the sample by MALDI-TOF MS. The MS analysis indicated that within 15 min ~ 30% subunit c was modified by DCCD and that the modification reaction reached a plateau after 30 min when ~ 80% subunit c was modified (Figure 2.7 A and B). This behavior is typical for proton-dependent ATP synthases (Kluge and Dimroth, 1993).

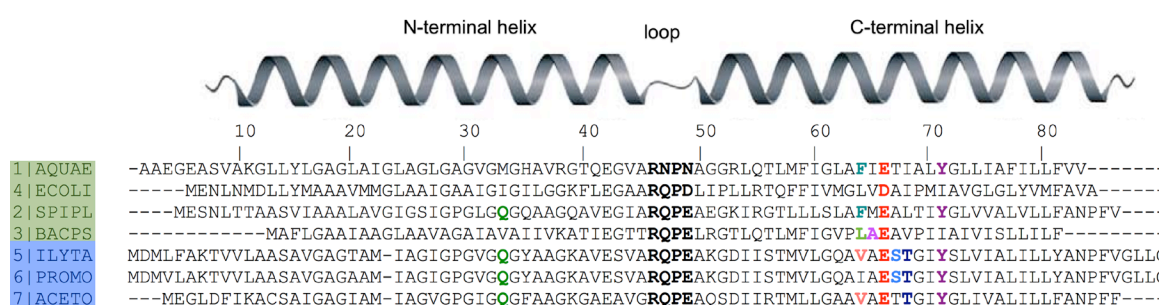


Figure 2.6. Sequence alignment of the subunit c of F_1F_0 ATP synthase. H^+ - and Na^+ -dependent ATP synthases are highlighted in green and blue, respectively. The conserved residues involved in H^+ or Na^+ binding are colored in bold. The cytoplasmic loop, connecting the N-terminal and C-terminal helices is shown in bold. The secondary structure is indicated by the schematics on top. The sequences from *A. aeolicus* (1), *E. coli* (4), *S. platensis* (2), *B. pseudofirmus* OF4 (3), *I. tartaricus* (5), *P. modestum* (6) and *A. woodii* (7) were used. The residue numbering corresponds to the *A. aeolicus* mature sequence (see Chapter 2.3.4).

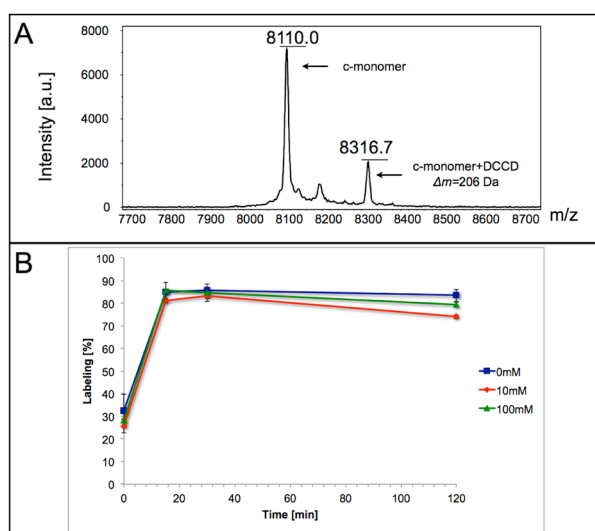


Figure 2.7. DCCD-labeling of the subunit c of AAF_1F_0 . (A) MALDI-TOF MS spectra of subunit c labeled with DCCD. Two peaks centered at 8110.0 and 8316.7 Da correspond to unlabeled and DCCD-modified c monomers. Spectra recorded after 10 min incubation of AAF_1F_0 with 100 μM DCCD, at pH 7.4, at room temperature. (B) Kinetics of DCCD labeling of the subunit c at different Na^+ concentration. Purified AAF_1F_0 was incubated with 100 μM DCCD at room temperature, pH 7.4 at different NaCl concentrations of 0 mM (blue line), 10 mM (red line) and 100 mM (green line). DCCD labeling of the subunit c is independent of Na^+ concentration. Each sample was measured in triplicates and data represents the mean of all measurements for each sample.

2.1.2.3. The ATP hydrolysis activity of AAF₁F₀ is comparable with the activity of other respiratory enzymes from *A. aeolicus*

The enzymatic ATP hydrolysis activity of purified AAF₁F₀ was determined via a phosphate determination assay (see Chapter 4.2.6.1). We determined that the ATP hydrolysis activity of

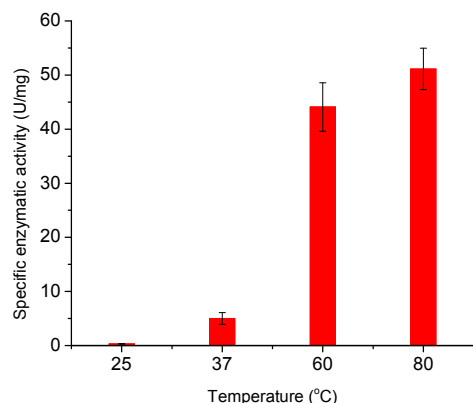


Figure 2.8. The ATP hydrolysis activity of AAF₁F₀ at different temperatures. AAF₁F₀ shows the highest ATP hydrolysis activity at 80 °C. Error bars indicate the standard deviation calculated from 3 independent measurements.

AAF₁F₀ is essentially undetectable at room temperatures, increases substantially at temperature above 60 °C and reaches its maximum at 80 °C (Figure 2.8). At 80 °C the ATP hydrolysis activity of AAF₁F₀ is 51.1 ± 3.8 U/mg. The pattern of temperature sensitivity and the values of specific enzymatic activity of the enzymes are comparable with the same parameters determined for other respiratory complexes of *A. aeolicus* (Peng *et al.*, 2003; Marcia *et al.*, 2010).

2.2. Generation of a heterologous system to produce *A. aeolicus* ATP synthase

2.2.1. Gene composition

A. aeolicus VF5 complete genome sequence (NC_000918) suggests that this organism possesses nine *atp* genes dispersed over six DNA fragments scattered throughout the genome (see Table 2.2 and Figure 1.14 for details).

Table 2.2: Properties of genes and corresponding subunits of the *A. aeolicus* F₁F₀ ATP synthase

Subunit	Gene	Gene length (bp)	Gene locus	MW (kDa)	Length (aa)	Stoichiometry
F ₀ -a	<i>atpB</i>	651	aq_179	24.01	216	1
F ₀ -b ₁	<i>atpF1</i>	435	aq_1586	16.71	144	1
F ₀ -b ₂	<i>atpF2</i>	558	aq_1587	21.23	185	1
F ₀ -c	<i>atpE</i>	303	aq_177	10.17	100	?
F ₁ -α	<i>atpA</i>	1512	aq_679	55.57	503	3
F ₁ -β	<i>atpD</i>	1437	aq_2038	53.32	478	3
F ₁ -γ	<i>atpG</i>	876	aq_2041	33.55	291	1
F ₁ -δ	<i>atpH</i>	546	aq_1588	20.73	181	1
F ₁ -ε	<i>atpC</i>	399	aq_673	14.81	132	1

2.2.2. Heterologous expression strategy for producing EAF₁F₀

The strategy used to produce *A. aeolicus* ATP synthase in the heterologous host *E. coli* consisted of the following steps (Figure 2.9). In the first step, single-gene expression was attempted for individual subunits (a, c, γ and ε) and dual-gene expression was attempted for specific combinations of different subunits (b₁-b₂, a-c, a-b₁-b₂, γ-ε). At this step, the *atp* genes were cloned into the open reading frame (ORF) of several different expression vectors (Surade *et al.*, 2006). Expression tests served as checkpoints for assessing the ability of different host strains to use native operons, native codons, and gene overlaps and to produce functionally active intermediate complexes of the enzyme. In the second step, expression was attempted for subcomplexes of the ATP synthase including subunits located in different loci of *A. aeolicus* genome. The first subcomplex created with this strategy coded for subunits F₁-αβγ (vector pCL11), in which an N-terminal His₆-tag was fused to subunit β for detection and purification purposes. The artificial operon encoding F₁-αβγ was then enlarged including subunit ε (F₁-αβγε, vector pCL12). In parallel, a third vector was created including all subunits composing the F₀ subcomplex and subunit δ (F₀-acb₁b₂δ, vector pCL02). This second step served to assess the importance of native intergenic regions, which were also cloned sequentially between the corresponding genes. In the final step, the genes for F₁-αβγε and for F₀-acb₁b₂δ were combined into a single expression vector (pCL21), which thus contained all nine genes encoding the entire *A. aeolicus* F₁F₀ ATP synthase.

2. Results

In the following sections we report the results of relevant individual subunits (see chapter 2.2.3), subcomplexes (see chapter 2.2.4) and of the complete ATP synthase (see chapter 2.2.5) produced according to the strategy described.

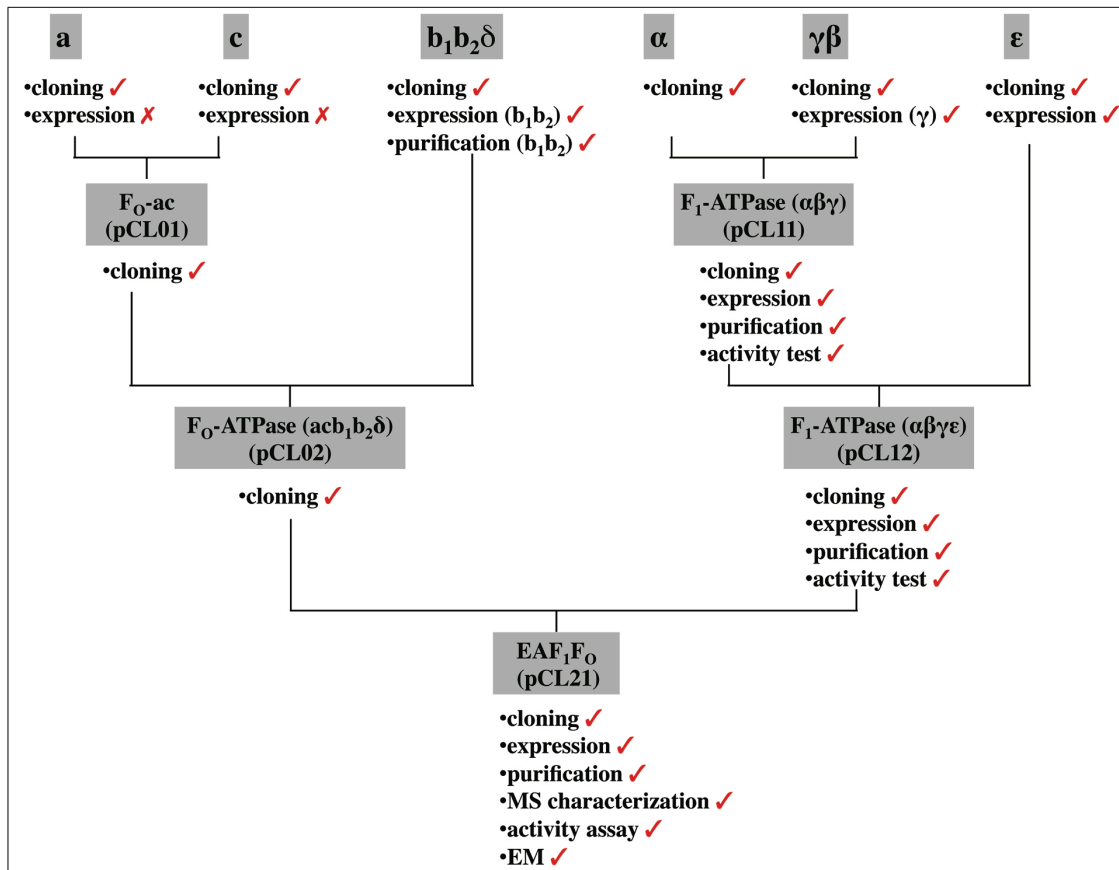


Figure 2.9. The EAF₁F_o expression strategy. Expression tests on the individual subunits (top line, successful steps highlighted by red ticks, unsuccessful steps by red crosses) showed that *E. coli* can correctly process native *A. aeolicus* operons, including those characterized by gene overlaps (i.e. b₁b₂), while other subunits (i.e. a and c) cannot be expressed individually. Furthermore, combinations of individual subunits in artificial operons (middle lines) led to the successful heterologous production of active and assembled subcomplexes of *A. aeolicus* ATP synthase. Finally, the operons encoding such subcomplexes were combined into a single artificial operon to produce the entire *A. aeolicus* ATP synthase in *E. coli* (EAF₁F_o, bottom line). This complex was produced and purified in an active form and consisted of all individual subunits correctly assembled into the characteristic ATP synthase structure as observed by single-particle EM.

2.2.3. Single-gene and dual-gene expression

Single-gene expression was attempted for subunits a, c, γ and ε and dual-gene expression was attempted for the following combinations of subunits: b₁-b₂, a-c, a-b₁-b₂ and γ-ε.

2.2.3.1. Strategy

Three different vector systems have been used for single-gene and dual-gene expression (Figure 2.10).

- i) The in-house modified pTTQ18 / pBAD / pQE expression vectors. These vectors contain only one multiple cloning sites (MCS), which is preceded by a ribosome binding site (RBS) and by the respective promoters, namely, a moderately strong hybrid *trp-lac* (*tac*) promoter for pTTQ18 vectors (Stark, 1987), a strong T5 promoter for pQE vector (Qiagen) and an *araBAD* promoter for pBAD vectors (Invitrogen). The vectors encode two sets of tags. The first set consists of a C-terminal fusion His₁₀ tag (A2 version). The second set consists of an N-terminal fusion His₁₀ tag and a C-terminal fusion strepII tag (C3 version). The additional set of vectors consist of an N-terminal MBP tag and C-terminal fusion His₁₀ tag, which was modified from A2-version vectors (MBP-fusion). These vectors were used to attempt to heterologously produce single subunits a and c, and the subcomplex consisting of subunits b₁ and b₂.
- ii) The modified bicistronic pBAD-CM1 expression vector. This vector contains two MCS, two optimal *E. coli* RBS preceding each MCS and one tightly regulated *araBAD* promoter preceding the first MCS. The vector additionally encodes a fusion His₆- and a StrepII-tags at the C-terminal of each target protein, respectively. The vector has been used to attempt to heterologously produce subcomplexes consisting of subunits a/c and subunits a/b₁/b₂.
- iii) The pETDuet-1 expression vector (Novagen). pETDuet-1 contains two MCS, each of which is preceded by a T7 promoter / *lac* operator and a RBS. It encodes a His₆ tag fused at the N-terminus of the first target protein and an S-tag fused at the C-terminus of the second target

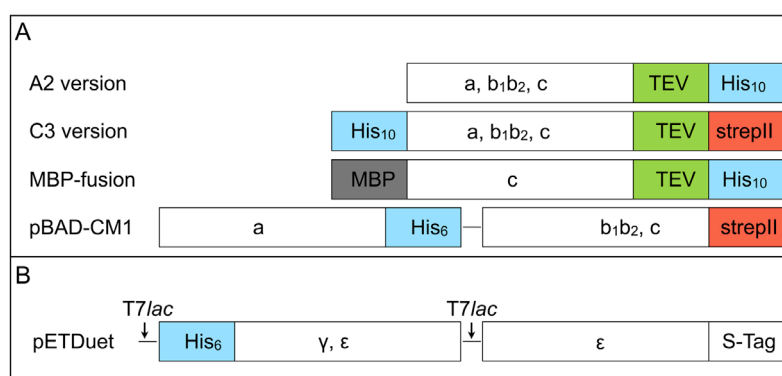


Figure 2.10. Single-gene and dual-gene expression vectors. Tags fused to the N- or C- termini of target proteins are shown in light blue (His₁₀ or His₆-tag), red (StrepII-tag), grey (MBP-tag) and white (S-tag), respectively. TEV protease recognition site is shown in green. (A) Based on the in-house modified pTTQ18 / pBAD / pQE expression vectors, expression vectors were constructed for subunits a, b₁b₂ and c. (B) Based on pETDuet-1, expression vectors were constructed for subunits γ and ε. T7lac indicates T7 promoter / *lac* operator.

protein. This vector has been used to produce the central stalk subunits γ and ϵ individually or in association.

2.2.3.2. Cloning

All primer sequences used for constructing the cloning and expression vectors for single subunits and dual subunits in *E. coli* are included in Table 4.12, while the individual empty vectors and the corresponding resulting vector are listed in Table 4.2 and 4.3 (see also Figure 2.10).

The genes *atpF1F2*, *atpB* and *atpE* were amplified from *A. aeolicus* genome using primers bb2-F/R, a-F/R and c-F/R, respectively. The genes were inserted into the ORFs of vectors pTTQ18-A2/C3, pQE-A2/C3, pBAD-A2/C3 at the BamHI/EcoRI site (Surade *et al.*, 2006). *atpE* was also inserted into the BamHI/EcoRI site of pET-G modified from pET-26 (+) (Novagen). In addition, *atpB* was also cut from the corresponding pBAD-A2 vector by restriction enzymes NcoI and EcoRI and inserted into the NcoI/EcoRI site of pBAD-CM1, resulting in the bicistronic vector a-pBCM. The gene *atpF1F2* amplified by the primer pairs bb2-pBCM-F/R were inserted into XbaI/SalI site of a-pBCM, resulting in vector a-bb2-pBCM for co-expressing subunit a, b_1 and b_2 . The same procedure was applied to the gene *atpE* amplified by the primer pairs c-pBCM-F/R to generate vector a-c-pBCM for co-expression of subunits a and c. Finally, in vector c-pQE-A-MBP, *atpE* was fused with N-terminal MBP derived from vector pMAL (NEB).

2.2.3.3. Subcomplex b_1b_2 : an *A. aeolicus* native operon with overlapping genes can be recognized by *E. coli*

In *A. aeolicus* genome, the genes *atpF1* (subunit b_1), *atpF2* (subunit b_2) and *atpH* (subunit δ) form an operon that is characterized by the presence of overlaps. In particular, *atpF1* and *atpF2* overlap by 1 bp and *atpF2* and *atpH* overlap by 8 bp (Figure 2.11).

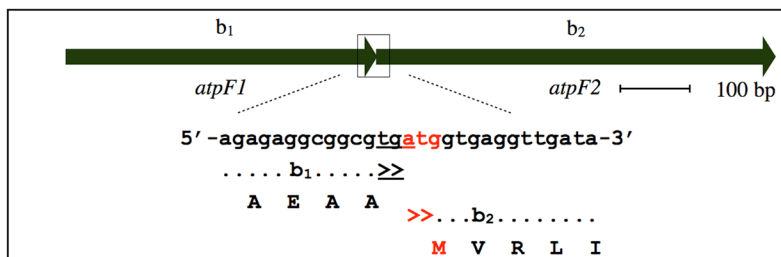


Figure 2.11. The scheme of *A. aeolicus* native operon *atpF1F2* encoding subunits b_1 and b_2 . Two genes overlap by 1 bp (underlined red adenosine).

By producing subcomplex b_1b_2 , we attempted to test (i) whether the native operon *atpF1F2* can be translated into subunits b_1 and b_2 by the heterologous host *E. coli*, and (ii) whether subunits b_1 and b_2 form a complex in *E. coli*.

Three expression vectors (pTTQ18, pQE and pBAD) with different promoters and two different sets of tags were chosen for producing b_1b_2 (see 2.2.3.1). The His₁₀-tagged subunit b_2 was detected by dot blot analysis and Western blot analysis using the monoclonal α -poly-histidine-alkaline phosphatase conjugated antibody. The heterologous production of subunit b_2 was detected from all vectors containing a C-terminal His₁₀-tag fused to subunit b_2 , but not from the vectors containing an N-terminal His₁₀-tag fused to subunit b_1 . Different promoters were associated with different production levels of the protein (Figure 2.12).

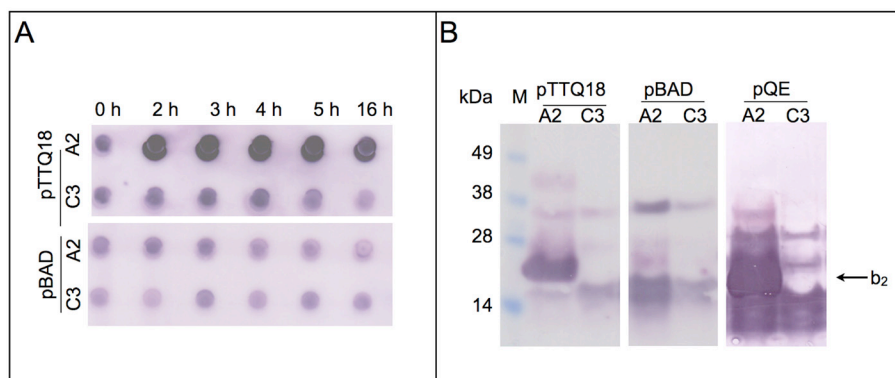


Figure 2.12. Expression tests on subcomplex b_1b_2 . His₁₀-tagged subunit b_2 was detected by dot blot analysis (A) and Western blot analysis (B) using anti-polyHistidine antibody.

The production level was further optimized by screening (i) combinations of different host strains (BL21, C43 and NM554 for pTTQ18 and pQE vectors, TOP10 for pBAD vector), (ii) different media (2× YT, TB, M9 and auto-induction TB (Novagen)), (iii) at various temperatures (18°C, 25°C, 30°C and 37°C), (iv) different IPTG concentrations (0.2 mM – 1 mM), and (v) different induction time (1 h – 16 h). His₁₀-tagged b_2 were detected by Western blot analysis in all these conditions with only slight variations in production levels. However, medium and induction time influenced significantly the amount of host cell mass that could be recovered. For example for bb_2 -pTTQ18-A expressed in *E. coli* C43 (DE3), the wet cell pellet increased from 0.83 g per liter (LB culture at 37°C induced by 1 mM IPTG for 2 h) to 7.7 g per liter (2× YT culture at 30°C induced by 0.02 mM IPTG for 16 h). Therefore, the highest yield was achieved using vector pTTQ18-A2 (Surade *et al.*, 2006) and using *E. coli* strain C43 (DE3) grown in 2× YT medium and induced by 0.2 mM IPTG at 30°C for 16 h.

2. Results

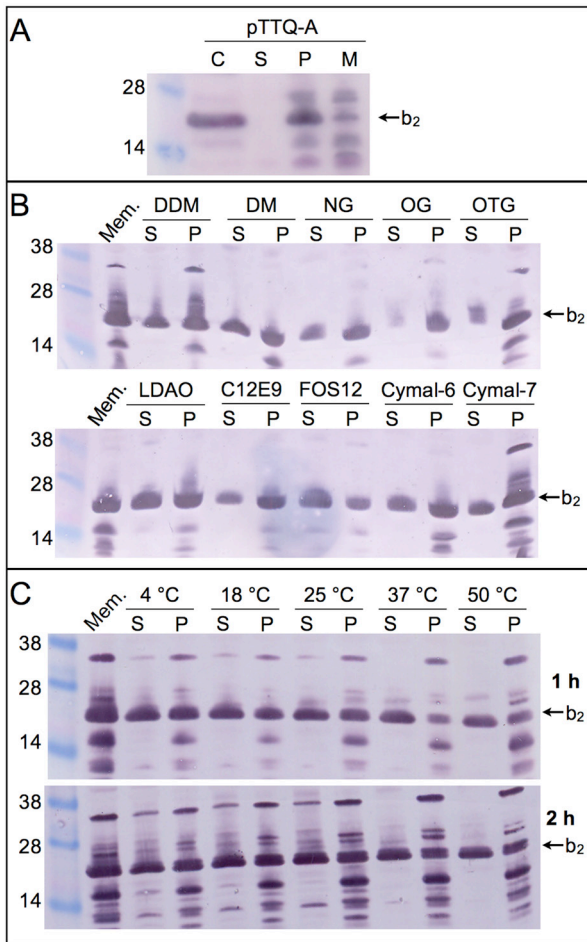


Figure 2.13. Subunit b_2 is expressed in *E. coli* membranes. After membrane preparation, subunit b_2 was detected in the membrane fraction (M) and in inclusion bodies (P), but not in the soluble fraction (S). The whole-cell lysate (C) served as positive control (Panel A). The protein b_2 was solubilized from membrane by various detergents (Panel B). The optimal yield of b_2 was obtained using DM at 50 °C for 1 h (Panel C). For panel (B) and (C), S indicates the solubilized fraction and P indicates the insolubilized fraction.

Subunits b_1 / b_2 could be identified in *E. coli* membranes by MS and Western blot and could be solubilized with various detergents, i.e. DDM, DM, NG, OG, NTG, LDAO, FOS-12, C12E9, Cymal-6 and Cymal-7. Two parameters of solubilization were further screened: temperature (4 °C, 18 °C, 25 °C, 37 °C and 50 °C) and length (1 h and 2 h). The highest yield was obtained using DM at 50 °C for 1 h. These conditions were thus used for further characterization (Figure 2.13).

Finally, b_1b_2 could be purified to homogeneity as a subcomplex by IMAC and size-exclusion chromatography and the subcomplex was stable at 4°C for at least 5 days (Figure 2.14 A and B). Two dimensional blue native / SDS polyacrylamide gel electrophoresis (2D-BN/SDS-PAGE) further confirmed that subunits b_1 / b_2 form a stable complex (Figure 2.14 C). The subunits b_1 and b_2 were identified by mass spectrometry and peptide mass fingerprinting (Figure 2.15). The successful expression of subunits b_1 and b_2 as a complex showed that *E. coli* can use a native operon of *A. aeolicus* with gene

overlaps and that *A. aeolicus* ATP synthase subunits can associate into complexes in the *E. coli* membranes.

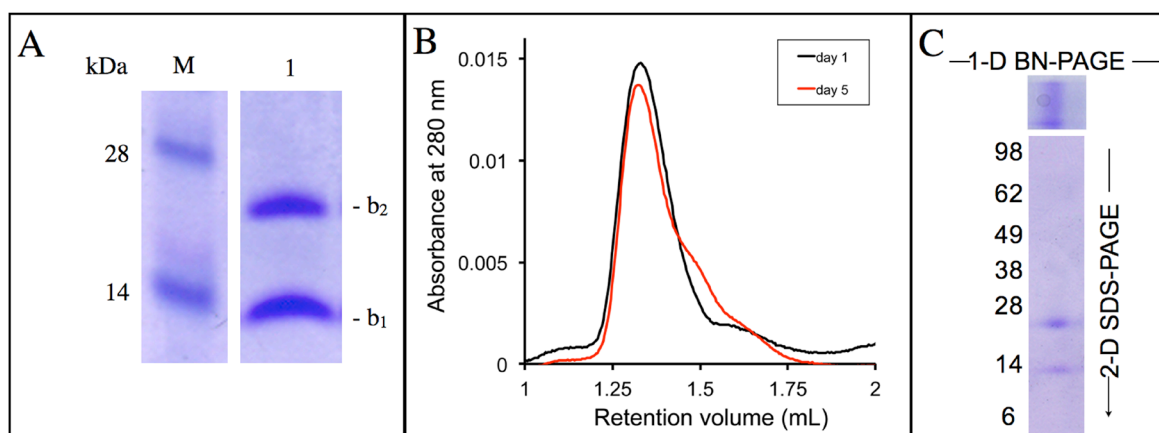


Figure 2.14. Subunits b₁ and b₂ form a subcomplex. Subunits b₁ and b₂ were purified from *E. coli* membranes and co-eluted as a single peak from a size-exclusion chromatographic column run in the presence of 0.25% (w/v) DM. Panel A shows a representative SDS-PAGE gel (M is a protein marker, the MW of the marker proteins is shown in kDa on the left of the gel). Panel B shows a representative SEC profile. The b₁b₂ subcomplex was also stable in detergent for at least 5 days at 4°C (panel B, red line). Panel C shows a representative 2D BN/SDS PAGE gel.

Subunit b ₁				
1	MDIGVMPNAT	ILVQLFIFVI	FLMIITNIYV	KPYTAVIESR EELIKKNLSE
51	AQKLREETQT	YLTQAKEVLE	DAKKRADQII	ENARREAEAQ ARSIIEQTEK
101	QTEEEIKKAV	EEIRTSLEEE	KKKLEKSVKE	IAQEIVKKIL REAA
Subunit b ₂				
1	MVRLISFLTL	ASTFAYAGEG	HLGHSPGALI	WKGLNILAFL GIVYYFGKKP
51	ISEAFNKFYN	SIVESLVNAE	REFMMAREEL	SKAKEELENA KKKAQYEKLL
101	AIETAETEEK	KILQHAQEV	ERIKEKAKET	IEIELNKAKK ELALYGIQKA
151	EEIAKDLLQK	EFKSKVQEK	YIEAQLKLE	ERKNA

Figure 2.15. Subunits b₁ and b₂ were identified by PMF. Residues in red were identified by PMF followed by ESI-MS.

2.2.3.4. Membrane subunits a and c cannot be expressed individually in *E. coli*.

Unlike subunits b₁b₂, membrane subunits a and c could not be obtained in isolation in *E. coli* using the same expression and optimization strategy applied to *atpF1F2* as described above. More sophisticated strategies were also attempted on these subunits. For instance, *atpE* (subunit c) was cloned into vectors containing either stronger promoters (i.e. T7), or encoding an N-terminal maltose-binding protein (MBP). Furthermore, the bicistronic vector pBAD-CM1 (Surade, 2007) was used to produce subunit a together with subunit c, or subunit a together with subunits b₁b₂ in *E. coli*. In all resulting vectors, a His₁₀-tag or a strepII-tag was fused to subunits a and c (see 2.2.3.1 and Figure 2.10). However, none of the above strategies allowed expression of subunits a and c, confirming previous observations about the difficulty of producing the membrane components of

ATP synthases in isolation (von Meyenburg *et al.*, 1985; Arechaga *et al.*, 2003). These results made it necessary to obtain an artificial operon to express the membrane subunits in the context of the whole ATP synthase complex.

2.2.4. Expression of subcomplexes $F_1\text{-}\alpha\beta\gamma$ and $F_1\text{-}\alpha\beta\gamma\epsilon$ from artificial operons

To test the ability of *E. coli* to express ATP synthase complexes from artificial operons, we cloned the His₆-tagged subcomplexes $F_1\text{-}\alpha\beta\gamma$ and $F_1\text{-}\alpha\beta\gamma\epsilon$. We chose these subcomplexes because (i) they are soluble and not membrane-inserted, (ii) we had raised polyclonal antibodies against their subunits, and (iii) we could perform enzymatic activity assays to test their functionality.

2.2.4.1. Subcomplex $F_1\text{-}\alpha\beta\gamma$

Fusing *atpA* upstream of *atpDG*, an *A. aeolicus* native operon, we formed the first artificial operon *atpAGD* to express subcomplex $F_1\text{-}\alpha\beta\gamma$. The expression of $F_1\text{-}\alpha\beta\gamma$ (pCL11 vector) was tested in different host strains with and without the pRARE vector at different temperatures and at different

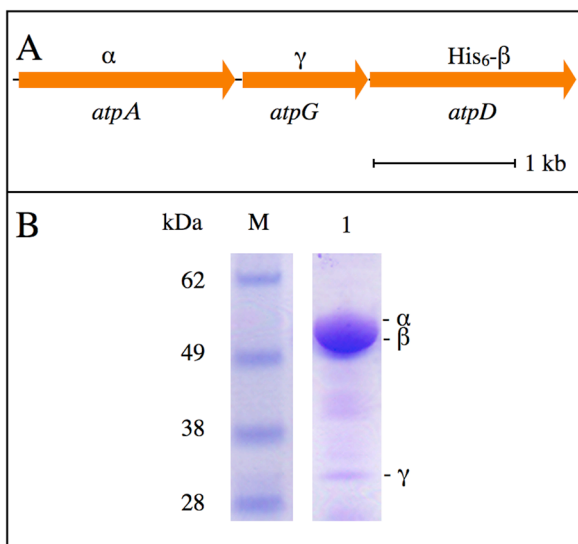


Figure 2.16. Subcomplex $F_1\text{-}\alpha\beta\gamma$. The production of a complex of the subunits α , β and γ was achieved with an artificial operon obtained by fusing gene *atpA* upstream of the native operon *atpGD* and by including the respective native intergenic regions (A). The three subunits formed a complex in *E. coli*, which could be purified by affinity chromatography (corresponding SDS-PAGE gel shown in panel B).

time points after induction with 1 mM IPTG (see 4.2.3.6.2). Under most conditions in which the pRARE vector was used, Western blot analysis indicated the presence of subunit β in whole cell lysates and the maximal levels were obtained in strain C43 (DE3) at 37°C, induced with 1 mM IPTG for 6 h. All three subunits were isolated from the cytoplasmic fraction of *E. coli* cells and purified in a single IMAC step. The subunits were identified by *in gel* trypsin digestion followed by ESI-MS. The three subunits co-eluted in the same chromatographic fractions indicating that they form a complex in *E. coli* (Figure 2.16 and Figure 2.17). Furthermore, the complex also showed activity, with a rate of ATP hydrolysis of 1.35 ± 0.14 U/mg.

Subunit α					
1	MATLTYEEAL	EILRQQIKDF	EPEAKMEEVG	VVYYVGDGVA	RAYGLENVMA
51	MEIVEFQGGQ	QGIAPNLEED	NVGIIILGSE	TGIEEGHIVK	RTGRILDAPV
101	GEGLVGRVID	PLGNPLDGKG	PIQFEYRSPV	EKIAPGVVKR	KPVHEPLQTG
151	IKAIDAMIPI	GRGQRELIIG	DRATGKTTVA	IDTILAQKNS	DVYCIYVAVG
201	QKRAAIARLI	ELLEREGAME	YTTVVVASAS	DPASLQYLAP	FVGCTIGEYF
251	RDNGKHALII	YDDLKHAEEA	YRQLSLLMR	PPGREAYPGD	VFYLHSRLLLE
301	RAAKLNDDL	AGSLTALPII	ETKAGDVAAY	IPTNVISITD	GQIYLEADLF
351	NKGIRPAINV	GLSVSRVGG	AQIKAMKQVA	GTLRLELAQF	RELEAFVQFA
401	SELDKATQQQ	INRGLRLVEL	LKQEPYNPIP	VEKQIVLIYA	GTHGYLDDIP
451	VESVRKFEKE	LYAYLDNERP	DILKEISEKK	KLDEELEKKI	KEALDAFKQK
501	FVP				
Subunit β					
1	MAEVIKGVV	QVIGPVVDVE	FEGVKELPKI	KDGLKTIRRA	IDDRGNWFEE
51	VLMEVAQHI	GEHRVRAIAM	GPTDGLVRGQ	EVEYLGGP	IPVKGKVLGR
101	IFNVAGQPID	EQQPVEAKEY	WPMFRNPPEL	VEQSTKVEIL	ETGIKVIDLL
151	QPIIKGGKVG	LFGGAGVGKT	VLMQELIHNI	ARFHEGYSV	VGVGERTREG
201	NDLWLEMKES	GVLPTYVMVY	GQMNPPGVR	FRVAHTGLTM	AEYFRDVEGQ
251	DVLIFIDNIF	RFVQAGAIEVS	TLLGRLPSAV	GYQPTLNTDV	GEVQERITST
301	KKSITAIQA	VYVPADDITD	PAPWSIFAHL	DATTVLTRL	AELGIYPAD
351	PLESTSKYLA	PEYVGEHHE	VAMEVKRILQ	RYKELQEI	ILGMEELSDE
401	DKAIVNRARR	IQKFLSQPFH	VAEQFTGMPG	KYVKLEDTIR	SFKEVLTGKY
451	DHLPENAFYM	VGTIEDVIEK	AKQMGAKV		
Subunit γ					
1	MAKLSPRDIK	RKIQGIKNTK	RITNAMKVVS	AAKLRKAQEL	VYASRPYSEK
51	LYELVGHAAA	HVDTEDNPLF	DVREERNVDV	ILVTADRGLA	GAFNSNVIRT
101	AENLIREKEE	KGVKVSLILV	GRKGFQYFTK	RGYNVIKGYD	EVERKTVNFN
151	VAKEVAEIVK	ERFLNGETDR	VYLINNEVMT	RASYKPQVRV	FLPFEAQEKE
201	VEELGTYEFE	VSEEEFFDYI	VNLYLNYQVY	RAMVESNAAE	HFARMIAMDN
251	ATKNAEDLIR	QWTLVFNKAR	QEAITTELID	ITNAVEALKA	Q

Figure 2.17. Identification of subunits α , β and γ by PMF. Residues in red were identified by ESI-PMF.

2.2.4.2. Subcomplex $F_1\text{-}\alpha\beta\gamma\epsilon$

The pCL11 vector was extended by inserting the *atpC* gene (subunit ϵ) downstream of gene *atpD* (subunit β), resulting in vector pCL12. The four subunits α , β , γ and ϵ encoded by pCL12 were successfully expressed and detected by ESI-MS and by Western blot analysis against the corresponding antibodies. The four subunits co-eluted from chromatographic columns indicating that they formed a complex. The $F_1\text{-}\alpha\beta\gamma\epsilon$ complex showed an ATP hydrolysis activity of 1.73 ± 0.11 U/mg. BN-PAGE and MS analysis showed the presence of lower order subcomplexes ($F_1\text{-}\alpha_2\beta_2$ and $F_1\text{-}\alpha_1\beta_1$) in the preparation, besides the fully assembled F_1 -subcomplex (Figure 2.18 and Figure 2.19). Finally, the apparent molecular weight of the $F_1\text{-}\alpha_3\beta_3\gamma\epsilon$ complex was determined to be 375 kDa by the SEC calibration curve, in good agreement with the expected MW of 378 kDa calculated from the sequence of the proteins (Figure 2.20).

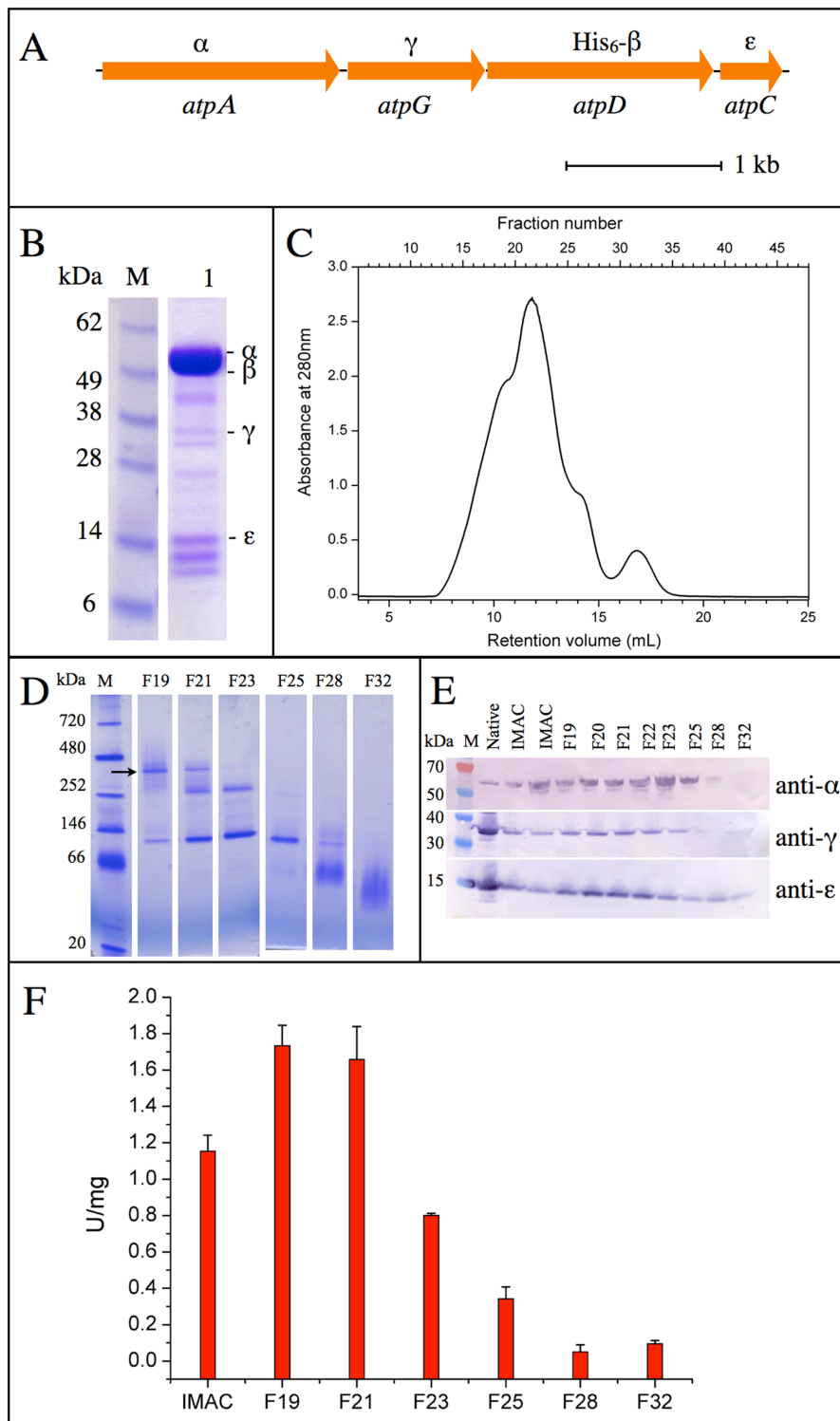


Figure 2.18. Subcomplex F₁-αβγϵ. The co-expression of the genes of subunits α, β, γ and ε was achieved by constructing expression vector pCL12. pCL12 harbors the artificial operon *atpAGDC* (panel A) formed by the DNAs of the corresponding *atp* genes and by their respective intergenic regions. All four genes were expressed. The encoded proteins were purified as a complex and identified by *in-gel* trypsin digestion followed by MS (representative SDS-PAGE gel in panel B, size-exclusion chromatographic profile in panel C, BN-PAGE of chromatographic fractions F19-F32 in panel D with an arrow indicating the fully assembled F₁-α₃β₃γϵ complex, Western blot analysis in panel E). The complex showed ATP hydrolysis activity (panel F).

Subunit α					
1	MATLTYYEAL	EILRQQIKDF	EPEAKMEEVG	VVYYVGDGVA	RAYGLENVMA
51	MEIVEFQGGQ	QGIAFNLEED	NVGIIILGSE	TGIEEGHIVK	RTGRILDAPV
101	GEGLVGRVID	PLGNPLDGKG	PIQFEYRSPV	EKIAPGVVKR	KPVHEPLQGTG
151	IKAIDAMIPI	GRGQRELIIG	DRATGKTTVA	IDTILAQKNS	DVYCIYVAVG
201	QKRAAIARLI	ELLEREGAME	YTTVVVASAS	DPASLQYLAP	FVGCTIGEYF
251	RDNGKHALII	YDDLKSHAEA	YRQLSLLMR	PPGREAYPGD	VFYLHSRLLLE
301	RAAKLNDDL	AGSLTALPII	ETKAGDVAAY	IPTNVISITD	GQIYLEADLF
351	NKGIRPAINV	GLSVSRVGG	AQIKAMKQVA	GTLRLELAQF	RELEAFVQFA
401	SELDKATQQQ	INRGLRLVEL	LKQEPYNPIP	VEKQIVLIYA	GTHGYLDDIP
451	VESVRKFEKE	LYAYLDNERP	DILKEISEKK	KLDEELEKKI	KEALDAFKQK
501	FVP				
Subunit β					
1	MAEVIKGVV	QVIGPVVDVE	FEGVKELPKI	KDGLKTIRRA	IDDRGNWFEE
51	VLFMEVAQHI	GEHRVRAIAM	GPTDGLVRGQ	EVEYLG GPIK	IPVGKEVLGR
101	IFNVAGQPID	EQGPVEAKEY	WPMFRNPPEL	VEQSTKVEIL	ETGIKVIDLL
151	QPIIKGGKVG	LFGGAGVGKT	VLMQELIHNI	ARFHEGYSVV	VGVGERTREG
201	NDLWLEMKES	GVLPTYVMVY	GQMNPPGVR	FRVAHTGLTM	AEYFRDVEGQ
251	DVLIFIDNIF	RFVQAGA EVS	TLLGR LPSAV	GYQPTLNTDV	GEVQERITST
301	KKGSITAIQA	VYVPADDITD	PAPWSIFAHL	DATTVLTRRL	AELGIYP AID
351	PLESTSKYLA	PEYVGEEHYE	VAMEVKRILQ	RYKELQEIIA	ILGMEELSDE
401	DKAIVNRARR	IQKFLSQPFH	VAEQFTGMPG	KYVKLEDTIR	SFKEVLTGKY
451	DHLPENAFYM	VGTIEDVIEK	AKQMGAKV		
Subunit γ					
1	MAKLSPRDIK	RKIQGIKNTK	RITNAMKVVS	AAKLRKAQEL	VYASRPYSEK
51	LYELVGH LAA	HVDTEDNPLF	DVREERNVDV	ILVTADRGLA	GAFN SNVIRT
101	AENLIREKEE	KGVKVSLILV	GRKGFQYFTK	RGYNVIKGYD	EVERKTVNFN
151	VAKEVAEIVK	ERFLNGETDR	VYLINNEMVT	RASYKPQVRV	FLPFEAQEKE
201	VEELGTYEFE	VSEEEFFDYI	VNLYLNYQVY	RAMVESNAAE	HFARMIAMDN
251	ATKNAEDLIR	QWTLVFNKAR	QEAITTELID	ITNAVEALKA	Q
Subunit ϵ					
1	MIQVEIVSPQ	GMVYSGEVES	VNVPTVEGEV	GILENHMYLM	TLLKPGLVYF
51	NGDDKNGIAV	TYGVL DVTPQ	KVLILAE EAY	EVGKLP PASK	LKEEFEEAVK
101	KMATAQT MEE	LKEWEKEAEK	ARTLLELVEK	YR	

Figure 2.19. Identification of subunits α , β , γ and ϵ by PMF. Residues in red were identified by ESI-PMF.

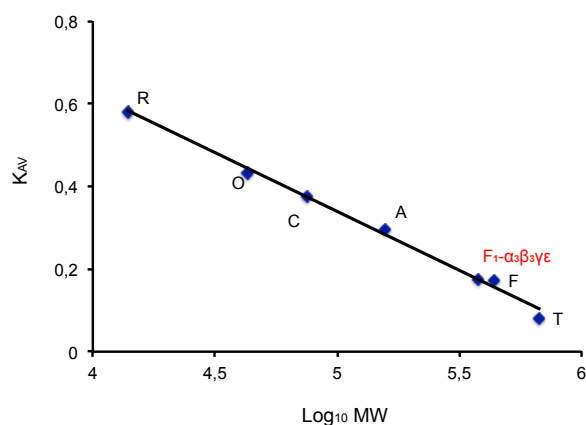


Figure 2.20. SEC calibration curves. The gel filtration column Supdex200 was calibrated in buffer D (20 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 150 mM NaCl, 2.5% (v/v) glycerol). The standards used are ribonuclease A (R, MW=13.7 kDa), ovalbumin (O, 43.0 kDa), conalbumin (C, 75.0 kDa), aldolase (A, 158 kDa), ferritin (F, 440 kDa) and thyroglobulin (T, 669 kDa). The x-axis reports the logarithm of the protein's MW (log₁₀ MW). The y-axis reports a function (K_{AV}) of the retention volume (V). $K_{AV} = (V - V_0) / (V_{tot} - V_0)$, V_{tot} corresponding to the total volume and V₀ to the void volume of the column. F_{1-α₃β₃γ_ε} possesses an approximate MW of 375 kDa in very good agreement with the calculated MW of 378 kDa.

2.2.5. Expression of the entire EAF₁F_O complex

2.2.5.1. Construction of the artificial operon: DNA amplification, manipulation and subcloning

Scheme of the successive steps of DNA amplification and manipulation used in this work for the construction of the artificial operon for production of EAF₁F_O is shown in Figure 2.21. These steps are summarized below.

DNA fragment F1(+) was amplified using primers P1/P2 and inserted into cloning vector pJET1.2 (Clontech), which resulted in vector pJET01(+) containing genes *atpB*, *aq_178* and *atpE* encoding subunit a, universal stress protein AQ_178 and subunit c. To remove *aq_178*, vector pJET01(-) containing only *atpB* (subunit a) and *atpE* (subunit c) was also cloned by overlap extension PCR. For this purpose, gene *atpB* and gene *atpE* were first amplified from pJET01(+) using primers P1/P17 and P18/P2, respectively (P17 and P18 are reverse complements, and are composed of 17 bp from 3' end of *atpB* and 17 bp from 5' end of *atpE*, see Table 4.12). Then the two amplified DNA fragments were used as templates to form DNA fragment F1, which contains only *atpB* and *atpE*. F1 was inserted into cloning vector pJET1.2 resulting in pJET01(-) (see Figure 2.21)

DNA fragments 2, 3, 4 and 5 (F2, F3, F4 and F5, see Figure 2.21) were amplified using primer pairs P3/P4, P5/P6, P7/P8 and P9/10, respectively. All fragments were then inserted into cloning vector pJET1.2. The resulting vectors were named pJET02 (containing *atpF1*, *atpF2* and *atpH* encoding subunit b₁, b₂ and δ), pJET3 (containing *atpA* encoding subunit α), pJET4 (containing *atpG* and *atpD* encoding subunit γ and β), pJET5 (containing *atpC* encoding subunit ε).

pJET4 was successively modified to insert a His₆-tag at the N-terminus of subunit β by site-directed mutagenesis using primers P19/P20, resulting in vector pJET4 (+His₆) (see Figure 2.21).

To combine the genes of the different DNA fragments, series of subcloning steps were carried out, which generated the following intermediate vectors (Figure 2.22).

pCL11 (*atpAGD*), used for the production of subcomplex F₁-αβγ, was constructed by three-way ligation. First, DNA fragment F3 (*atpA*, subunit α) was prepared by BamHI and XbaI double digestion from pJET3. Second, DNA fragment F4-His₆ (*atpGD*, subunits γ and His₆-β) was prepared by XbaI and Sall double digestion from pJET4 (+His₆). Finally, F3 and F4 were ligated and inserted into the BamHI/Sall sites of vector pTrc99A.

pCL12 (*atpAGDC*), used for producing the subcomplex F_1 - $\alpha\beta\gamma\epsilon$, was obtained by inserting DNA fragment F5 (*atpC*, subunit ϵ), into the *SalI*/*PstI* site of vector pCL11. F5 was prepared by *SalI* and *PstI* double digestion from pJET5.

pCL02 (*atpBEF1F2H*), encoding membrane F_0 subunits a, c, b_1 , and b_2 and for soluble F_1 subunit δ , was also constructed by three-way ligation. First, DNA fragment F1 (*atpBE*, subunit a and c) was prepared by *KpnI* and *BglIII* double digestion from pJET01(-). Second, DNA fragment F2 (*atpF1F2H*, subunit b_1 , b_2 and δ) was prepared by *BglIII* and *BamHI* digestion from pJET02. Finally, F1 and F2 were ligated and inserted into the *KpnI*/*BamHI* restriction sites of pTrc99A.

Finally, vector pCL21 (*atpBEF1F2HAGD*) was used to obtain the entire *A. aeolicus* F_1F_0 complex. pCL21 was derived by cutting the DNA fragment encoding F_0 - $\alpha b_1 b_2 \delta$ from pCL02 and inserting it into the *KpnI*/*BamHI* site of pCL12.

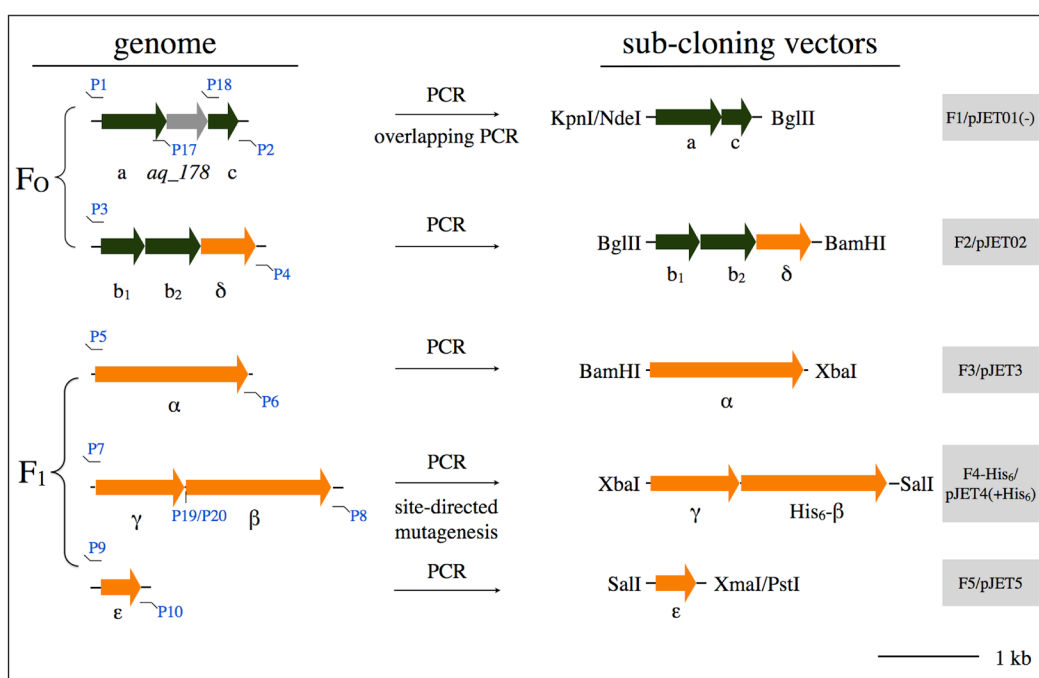


Figure 2.21. Scheme of the initial steps of construction of the artificial *atp* operon. These steps include the amplification of five DNA fragments (F1-F5) from the *A. aeolicus* genome (left), and the introduction of unique restriction sites (indicated by the corresponding restriction endonucleases in the middle) for ligation and sub-cloning into cloning vector pJET1.2 for DNA propagation. The vectors resulting from this process are shown on the right in gray boxes.

2. Results

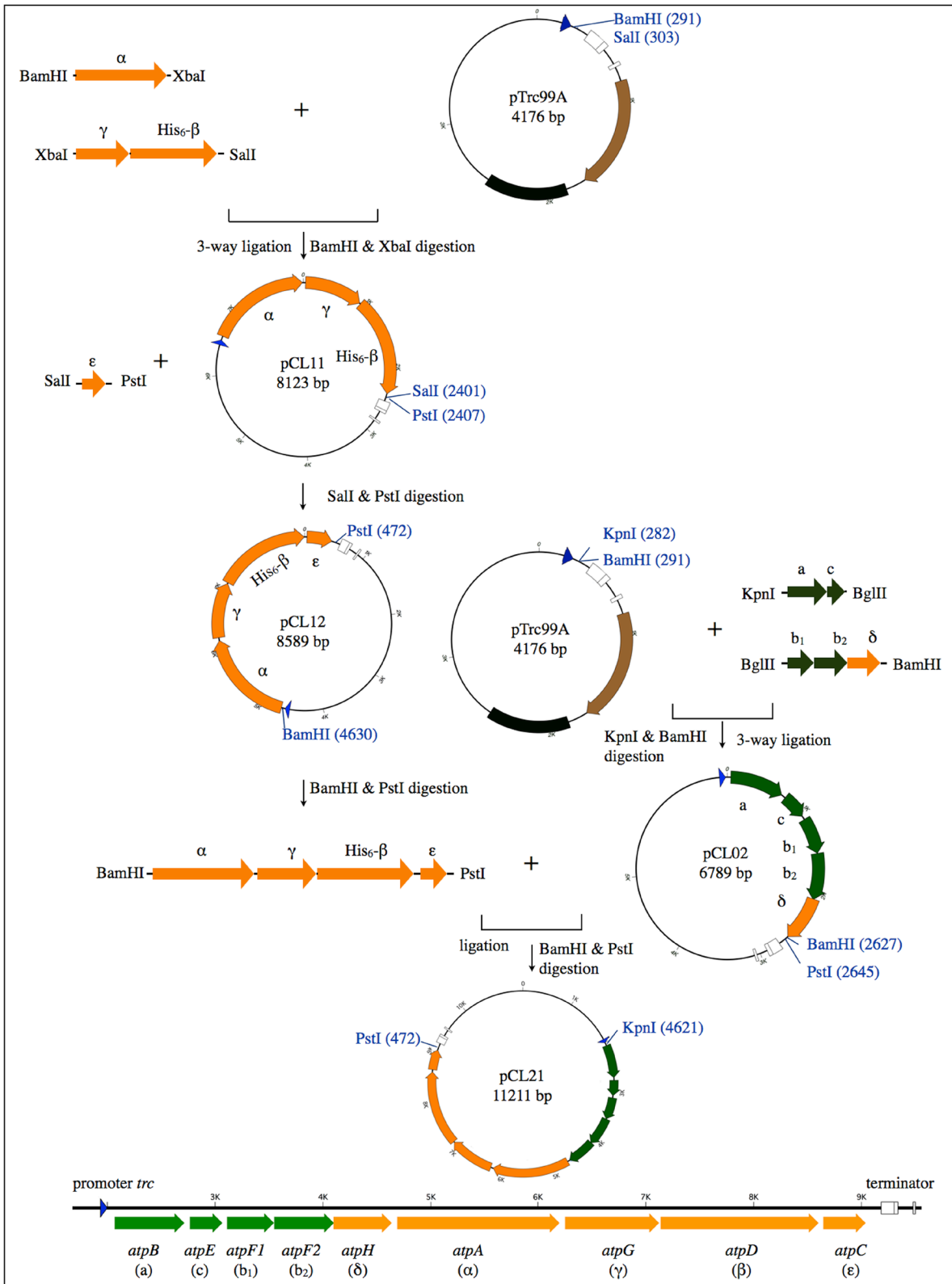


Figure 2.22. Scheme for the construction of intermediate expression vectors pCL11, pCL12, and pCL02 and of the expression vector pCL21 containing the whole *atp* operon.

2.2.5.2. The entire ATP synthase complex: the fully assembled His₆-EAF₁F₀ purified from the membranes of *E. coli*

Vector pCL21 harbors the nine *atp* genes in the order *atpBEF1F2HAGDC*, the same order of the *atp* genes in the native operon from *E. coli*, and resulting in an artificial operon of 7068 bp. The scaffold of pCL21 corresponds to vector pTrc99A, which had already been used for production of ATP synthases in the *E. coli* strain DK8 (Δatp) (McMillan *et al.*, 2007; Hakulinen *et al.*, 2012; Brandt *et al.*, 2013) (Figure 2.23). It contains an inducible *trc* promoter (74 bp) located 60 bp upstream of the start codon of gene *atpB*. The His₆-tag was fused to the N-terminus of subunit β .

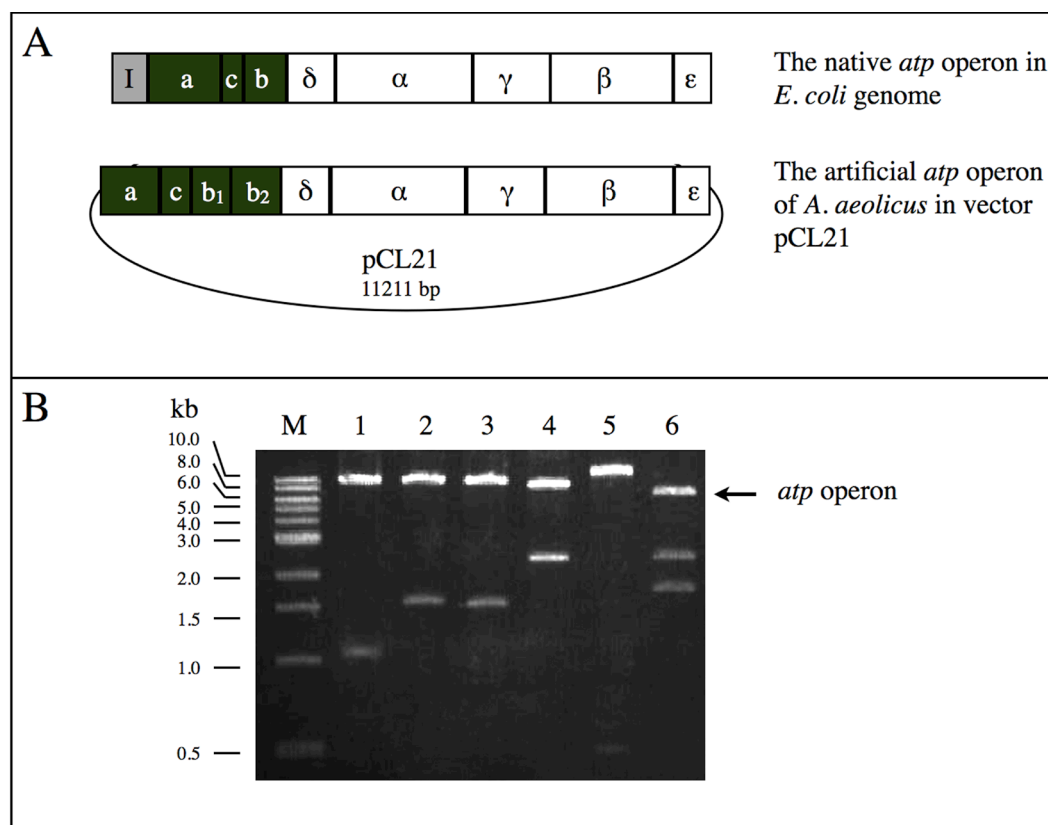


Figure 2.23. The artificial *atp* operon for heterologous production of EAF₁F₀. (A) Comparison of the organization of the native *atp* operon in *E. coli* (Walker *et al.*, 1984) and of the artificial *atp* operon designed for the *A. aeolicus* ATP synthase (vector pCL21). (B) 1% (w/v) agarose gel representing the DNA fragments composing the artificial operon obtained from pCL21 after digestion with different restriction endonucleases. M is a 1 kb DNA ladder, 1-6 indicate the DNA fragments obtained by digesting pCL21 with KpnI/BglIII (*atpBE*, 1051 bp, lane 1), BglIII/BamHI (*atpF1F2H*, 1576 bp, lane 2), BamHI/XbaI (*atpA*, 1558 bp, lane 3), XbaI/SalI (*atpGD*, 2401 bp, lane 4), SalI/PstI (*atpC*, 476 bp, lane 5), and NdeI/SmaI (entire *atp* operon, 7051 bp, indicated by an arrow, lane 6), respectively.

2. Results

Expression of EAF_1F_O from pCL21 was tested at various temperatures and time points after induction with 1 mM IPTG, and Western blot analysis revealed the highest levels for His₆-tagged subunit β in *E. coli* membranes when the cells were grown at 37°C for 6 h after induction (Figure 2.24). The identification of subunit β in *E. coli* membranes provided a first hint that the membrane components of the ATP synthase were also expressed and interacted to form a complex. We therefore attempted to purify the ATP synthase complex from *E. coli* membranes to assess whether it was composed of all subunits and was functional.

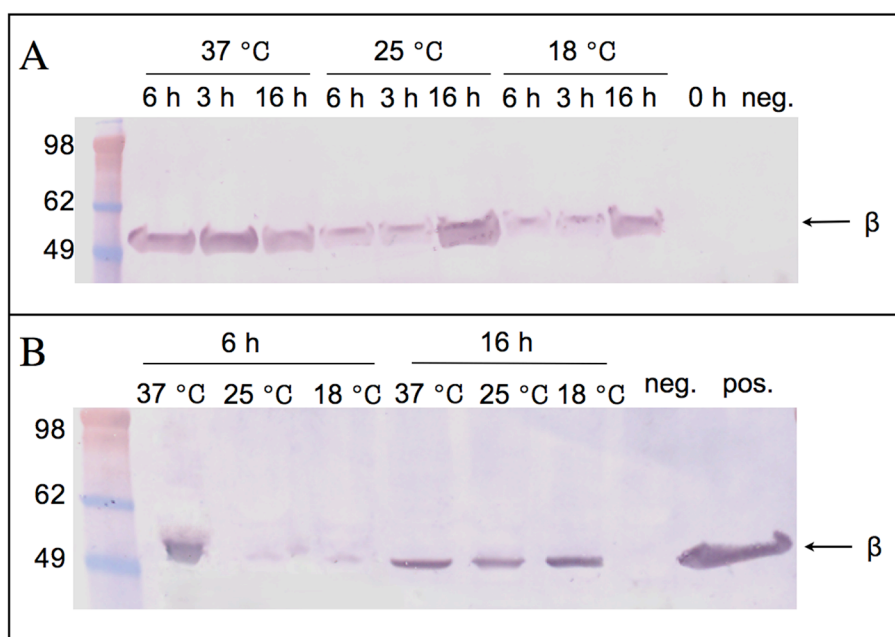


Figure 2.24. Western blot analysis to detect expression of EAF_1F_O . His₆-tagged subunit β was detected in the *E. coli* DK8 with pCL21 by Western blot analysis using anti-polyHistidine antibody in whole cell lysates (A) and membranes (B).

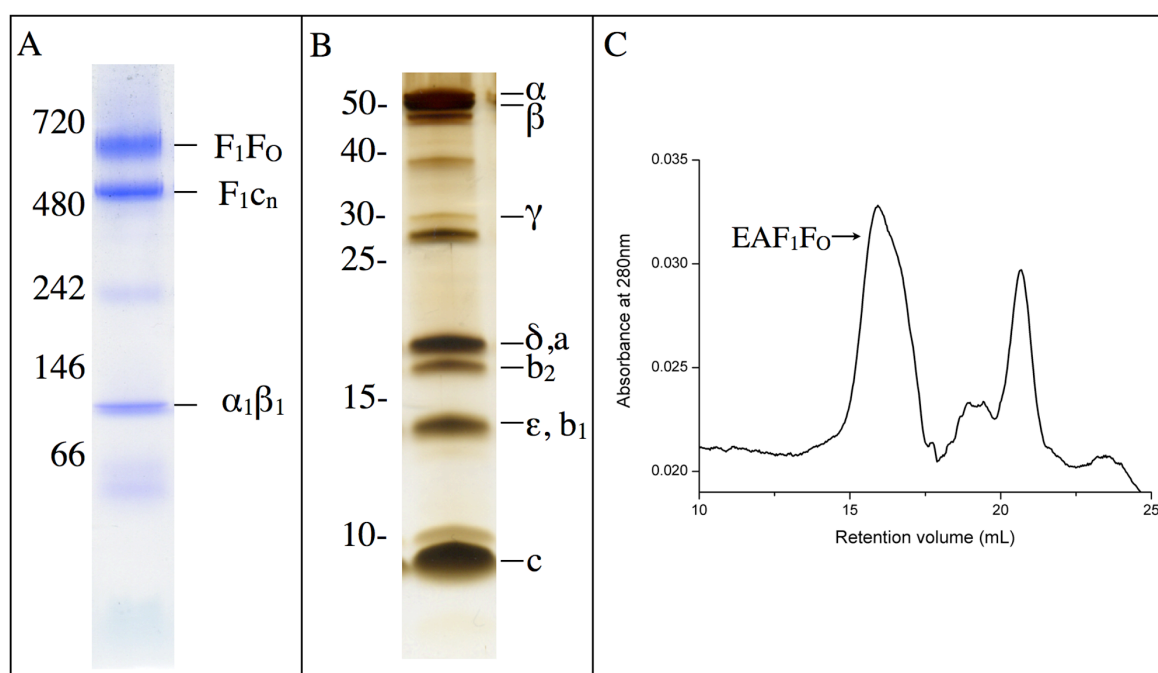


Figure 2.25. The fully assembled EAF_1F_0 purified from membranes of *E. coli*. The figure shows the characterization of EAF_1F_0 by BN-PAGE (A), silver-stained SDS-PAGE (B) and size exclusion chromatography (C).

The optimized purification protocol consisted of the following steps. *E. coli* membranes were solubilized by 3% (w/v) DDM and 3% (w/v) DM and subjected to heat treatment, which is a common procedure used to separate thermophilic proteins from the proteins of *E. coli* (e.g. (Imamura *et al.*, 2006)). After heat treatment, the soluble components of the sample were fractionated by Ni-NTA affinity chromatography followed by SEC in the presence of a detergent mixture composed of 0.05% (w/v) DDM + 0.05% (w/v) α -PCC. All subunits co-eluted in the same fractions of the IMAC column, indicating that they form a membrane complex, which requires detergent for stability and assembly. The complex (EAF_1F_0) was observed as a single electrophoretic band on BN-PAGE gels on which it shows the same electrophoretic mobility as AAF_1F_0 (Figure 2.25), although in both preparation subcomplexes corresponding to F_1c_n (Figure 2.26) and $\alpha_1\beta_1$

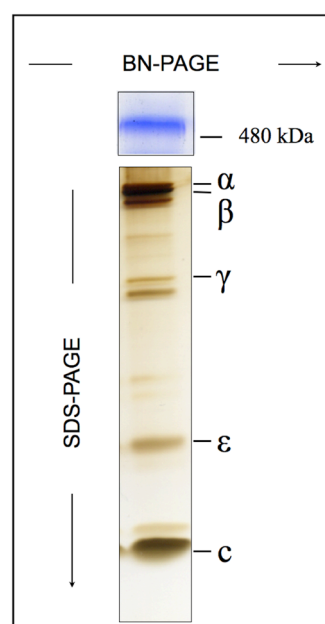


Figure 2.26. Silver-stained SDS-PAGE gel of F_1c_n subcomplex electro-eluted from BN-PAGE gel.

2. Results

were also identified. In the complex all subunits of EAF_1F_0 were detected by ESI-PMF, including membrane subunits a and c (Figure 2.27 and Figure 2.28), which could not be obtained in isolation (see chapter 2.2.3.4).

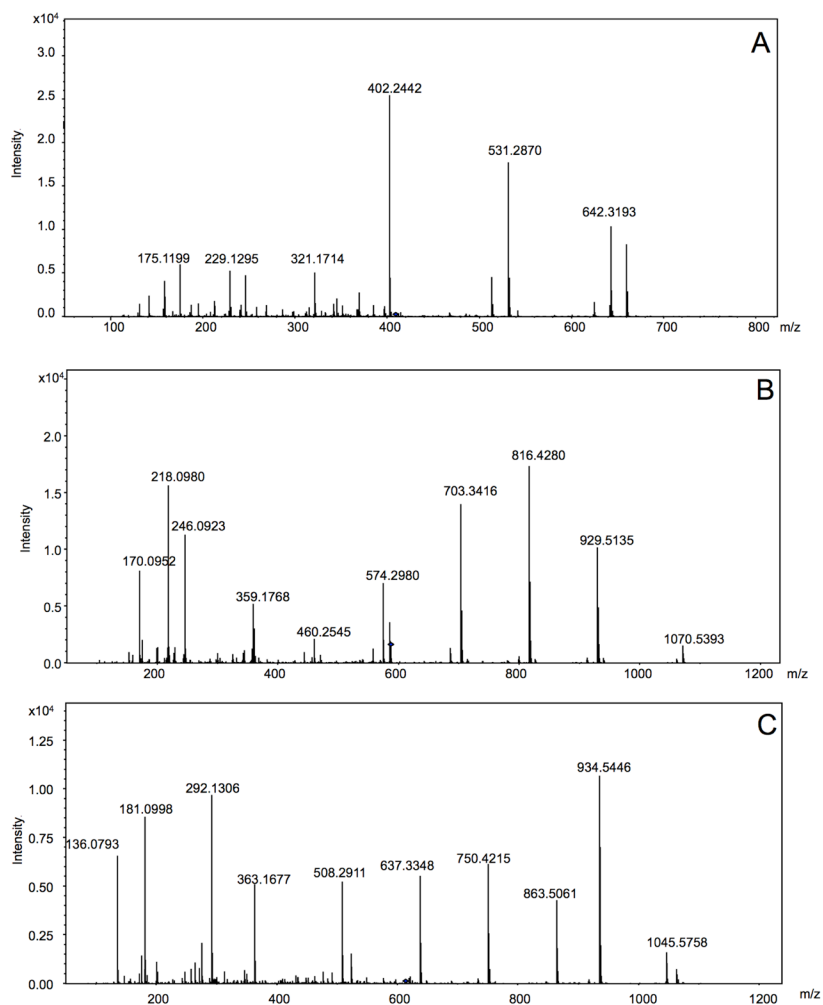


Figure 2.27. Mass-spectra of subunits a and c. Subunits a and c were digested with trypsin in the SDS-PAGE gel. MS/MS spectra of $m/z = 409.2$ (ion score 42), 587.8 (ion score 69) and 613.3 (ion score 72), identified as peptides GTQEGVR of subunit c (A), NMLLENVGER of subunit a (B), and YQALLEGYLR of subunit a (C), respectively.

Subunit a					
1	MEYSHVVYAL	LAVALAIIFV	LKGGKPSLKP	TKYQALLEGY	LRFVRNMLLE
51	NVGERGLKYV	PLIAAIGLFV	FFGNILGMVP	GFEAPTANIN	TNLALALLVF
101	FYYHFEGFRE	NGLAYLKHFM	GPIPLMAPFF	FVVEVISHIA	RPITLSLRLF
151	ANMKAGALLL	LTLVSLVIKN	PFTLVVSPVV	LIFVIAIKFL	AIFIQTYIFM
201	ILSVYIAGA	VAHEEH			
Subunit b₁					
1	MDIGVMPNAT	ILVQLFIFVI	FLMIITNIYV	KPYTAVIESR	EELIKKNLSE
51	AQKLREETQT	YLTQAKEVLE	DAKKRADQII	ENARREAEAQ	ARSIIEQTEK
101	QTEEEIKKAV	EIIRTSLEEE	KKKLEKSVKE	IAQEIVKKIL	REAA
Subunit b₂					
1	MVRLISFRTL	ASTFAYAGEG	HLGHSPGALI	WKGLNILAFL	GIVVYFGKKP
51	ISEAFNKFYN	SIVESLVNAE	REFMMAREEL	SKAKEELENA	KKKAQYEYKL
101	AIETAETEKK	KILQHAQEVS	ERIKEKAKET	IEIELNKAKK	ELALYGIQKA
151	EETAKDLLQK	EFKSKVQEK	YIEAQLKLE	ERKNA	
Subunit c					
1	MMKRLMAILT	AIMPAIAMAA	EGEASVAKGL	LYLGAGLAIG	LAGLGAGVGM
51	GHAVRGTQEG	VARNPNAGGR	LQTLMFIGLA	FIETIALYGL	LIAFILLFVV
Subunit α					
1	MATLTYYEAL	EILRQQIKDF	EPEAKMEEVG	VVYYVGDGVA	RAYGLENVMA
51	MEIVEFQGGQ	QGIAFNLEED	NVGIIILGSE	TGIEEGHIVK	RTGRILDAPV
101	GEGLVGRVID	PLGNPLDGKG	PIQFEYRSPV	EKIAPGVVVKR	KPVHEPLQTG
151	IKAIDAMIPI	GRGQRELIIG	DRATGKTTVA	IDTILAQKNS	DVYCIYVAVG
201	QKRAAIARLI	ELLEREGAME	YTTVVVASAS	DPASLQYLAP	FVGCTIGEYF
251	RDNGKHALII	YDDLKSHAEA	YRQLSLLMR	PPGREAYPGD	VFYLHSRLLLE
301	RAAKLNDDLG	AGSLTALPII	ETKAGDVAAY	IPTNVISITD	GQIYLEADLF
351	NKGIRPAINV	GLSVSRVGG	AQIKAMQVA	GTLRLELAQF	RELEAFVQFA
401	SELDKATQQQ	INRGLRLVEL	LKQEPYNPIP	VEKQIVLIYA	GTHGYLDDIP
451	VESVRKFEKE	LYAYLDNERP	DILKEISEKK	KLDEELEKKI	KEALDAFKQK
501	FVP				
Subunit β					
1	MAEVIKGVV	QVIGPVVDVE	FEGVKELPKI	KDGLKTIRRA	IDDRGNWFEE
51	VLMEVAQHI	GEHRVRAIAM	GPTDGLVRGQ	EVEYLGGPVK	IPVGKEVLGR
101	IFNVAGQPID	EQQPVEAKEY	WPMFRNPPEL	VEQSTKVEIL	ETGIKVIDLL
151	QPIIKGGKVG	LFGGAGVGKT	VLMQELIHNI	ARFHEGYSVV	VGVGERTREG
201	NDLWLEMKES	GVLPTYVMVY	GQMNPPGVR	FRVAHTGLTM	AEYFRDVEGQ
251	DVLIFIDNIF	RFVQAGAEVS	TLLGRLPSAV	GYQPTLNTDV	GEVQERITST
301	KKGSITAIQA	VYVPADDITD	PAPWSIFAHL	DATTVLTRRL	AELGIYPAD
351	PLESTSKYLA	PEYVGEEHYE	VAMEVKRILQ	RYKELQEIIA	ILGMEELSDE
401	DKAIVNRARR	IQKFLSQPFH	VAEQFTGMPG	KYVKLEDTIR	SFKEVLTGKY
451	DHLPENAFYM	VGTIEDVIEK	AKQMGAKV		
Subunit γ					
1	MAKLSPRDIK	RKIQQIKNTK	RITNAMKVVS	AAKLRKAQEL	VYASRPYSEK
51	LYELVGHAAA	HVDTEDNPLF	DVREERNVDV	ILVTADRGLA	GAFNSNVIRT
101	AENLIREKEE	KGVKVSLILV	GRKGFQYFTK	RGYNVIKGYD	EVFRKTVNFN
151	VAKEVAEIVK	ERFLNGETDR	VYLINNEVMT	RASYKPQVRV	FLPFEAQEK
251	ATKNAEDLIR	QWTLVFNKAR	QEAITTELID	ITNAVEALKA	Q
Subunit δ					
1	MLKRKELARK	AVRLIVKKVP	KEKESILKVD	EFLGTLSTAY	RKDKLLRNFF
51	LSPQIDRNAK	VKALESIAKK	YDVPKEVLEV	LEYLIDINAM	ALIPKIKRLY
101	ELELEKLMGM	LKGELELAKK	PSKKLLEKIT	KTINDILNRQ	IEIEVKEDPS
151	LIGGFVFKTQ	AFVLDTSVKT	QLEKLARVGG	V	
Subunit ε					
1	MIQVEIVSPQ	GMVYSGEVES	VNVPTVEGEV	GILENHMYLM	TLLKPGLVYF
51	NGDDKNGIAV	TYGVLDVTPQ	KVLILAEAY	EVGKLPPASK	LKEEFEEAVK
101	KMATAQTMEE	LKEWEKEAEK	ARTLLELVEK	YR	

Figure 2.28. Identification of all subunits of EAF₁F₀ by PMF. Residues in red were identified by ESI-PMF.

2.2.6. EAF₁F₀ catalyzes ATP hydrolysis at the same rate of AAF₁F₀

A series of activity assays were performed to assess the enzymatic activity of the isolated EAF₁F₀ complex.

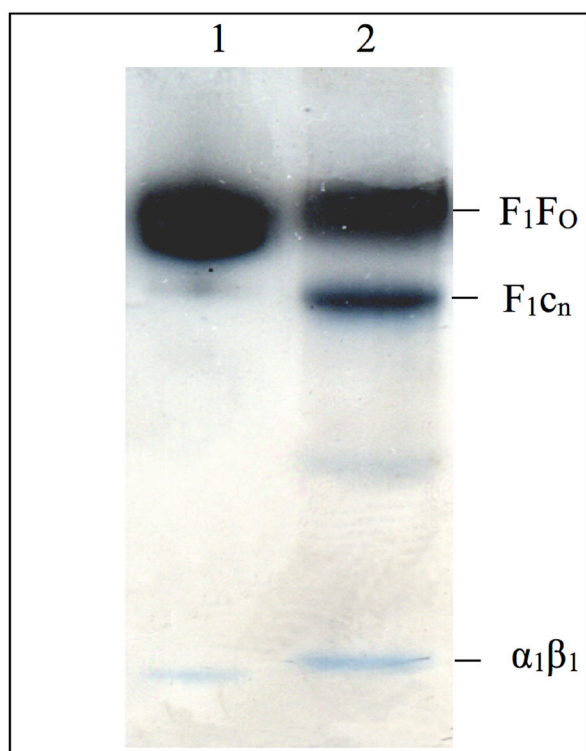


Figure 2.29. Functional characterization of EAF₁F₀ by *in-gel* ATP-hydrolysis activity at 80°C. With the production of phosphate, brownish lead sulfide precipitates form in the presence of Na₂S. Like AAF₁F₀ (lane 1), the fully assembled EAF₁F₀ (lane 2) also shows ATP hydrolysis activity. Some activity is also detectable for the subcomplex F₁c_n.

First, an *in gel* activity assay at 80°C was used to directly visualize ATP hydrolysis. With the production of phosphate, lead phosphate precipitation that turned brownish in the presence of Na₂S was directly observed in BN-PAGE gel bands of both fully assembled AAF₁F₀ and EAF₁F₀ (Figure 2.29). Additionally, the subcomplex F₁c_n purified from both native *A. aeolicus* and from *E. coli* also showed ATP hydrolysis activity, while the F₁-subcomplex $\alpha_1\beta_1$ was inactive. This observation suggests that the minimum ATP hydrolytic unit in *A. aeolicus* ATP synthase is not the $\alpha_1\beta_1$ heterodimer as in the thermophilic bacterium PS3 (Harada *et al.*, 1991) but probably a subcomplex including other subunits, as for *E. coli* ATP synthase (Futai *et al.*, 1988).

Second, phosphate determination assays were used to determine the ATP hydrolysis activity of AAF₁F₀ and EAF₁F₀. Also this assay was performed at 80°C, which was determined to be the ideal working temperature for a number of other respiratory membrane complexes of *A. aeolicus* (Peng *et al.*, 2003; Marcia *et al.*, 2010). The results showed that ATP hydrolysis activity of EAF₁F₀ was 12.77 ± 3.96 U/mg, a value of the same order of magnitude (43%) as that of native AAF₁F₀ (29.65 ± 3.66 U/mg, see chapter 2.1.2.4). Furthermore, the activity was reduced approximately 100 fold (0.17 ± 1.85 U/mg residual activity) by 0.02% (w/v) sodium azide, a common inhibitor of bacterial ATP synthases. These results showed that EAF₁F₀ and AAF₁F₀ possess a comparable activity at 80 °C.

2.2.7. The structure of EAF_1F_O is identical to that of AAF_1F_O

To obtain direct structural evidence that EAF_1F_O was fully assembled, we used single-particle electron microscopy. After electro-elution from the BN-PAGE, characteristic “mushroom” shaped-particles of EAF_1F_O could be observed by single particle analysis after negative staining. EAF_1F_O is ~ 220 Å long and has two distinct parts. One part possesses a globular shape with a diameter of 110 Å and probably corresponds to the F_1 subcomplex. The other part is approximately 100 Å wide parallel to the putative membrane plane and ~ 60 Å high perpendicular to the putative membrane plane and probably corresponds to the F_O subcomplex. Importantly, both the central and peripheral stalks are clearly visible in the EM images. Therefore, we conclude that EAF_1F_O is fully assembled and shows a similar structural organization as AAF_1F_O (Peng *et al.*, 2006) (Figure 2.30).

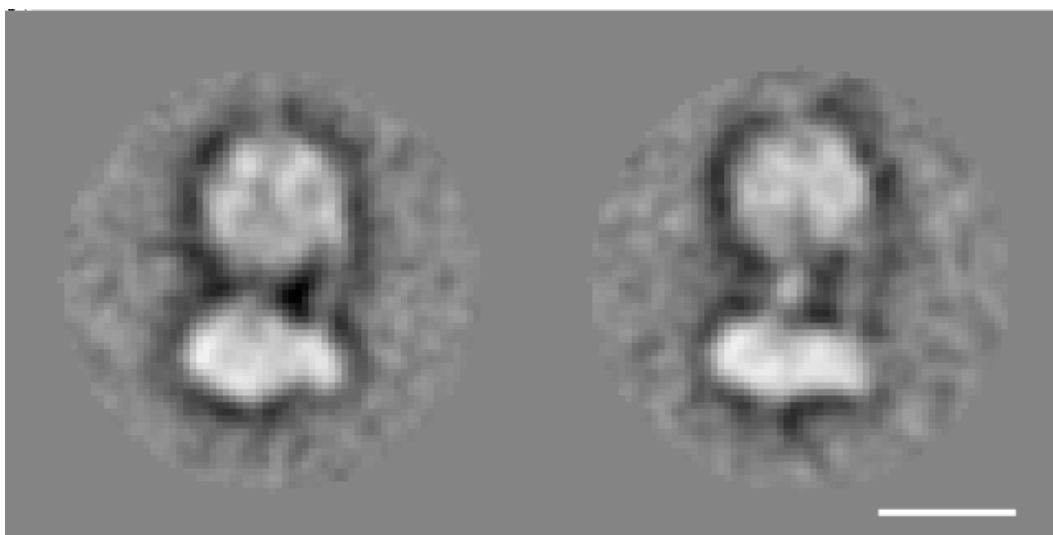


Figure 2.30. Structural characterization of EAF_1F_O by single-particle EM. Two class averages obtained by single particle analysis after negative staining show that EAF_1F_O is fully assembled. The peripheral stalk and the central stalk of the complex are clearly visible. EAF_1F_O shows an identical structure to the native AAF_1F_O (Peng *et al.*, 2006). The scale bar is 10 nm.

2.3. Use of EAF_1F_O to study unique properties of *A. aeolicus* ATP synthase

2.3.1. Characterization of the N-terminus of *A. aeolicus* ATP synthase subunit c

Overall, rotary ATPases (including V-type ATPase and F-type ATP synthases) are evolutionarily well-conserved, but the N-terminal segments of their rotary subunits (subunit c) possess different

lengths and levels of hydrophobicity across species. By analyzing the N-terminal variability, we distinguished four phylogenetic groups of subunits c (groups 1 to 4), with *A. aeolicus* ATP synthase belonging to group 2, which also contains poorly characterized homologues from other predominantly thermophilic organisms. In order to characterize the subunit c from *Aquifex aeolicus* F₁F₀ ATP synthase, we made use of the heterologous expression system described above and we demonstrate that its N-terminal segment forms a signal peptide with SRP-recognition features, and is obligatorily required for membrane insertion.

2.3.2. Subunits c can be subdivided into four classes based on their sequences

During sequence alignments using PSI-BLAST and 218 subunit c sequences from different organisms, we detected a large variation in the lengths of these subunits. Based on the length distribution identified with PSI-BLAST and on available 3-D structures, we propose a 4-group classification scheme for subunit c (Table 2.3).

Group 1 subunits c are typically 66-92 amino acids long. They possess the most common topology among subunits c, with a short N-terminal tail preceding two transmembrane helices, connected by a cytoplasmic loop. This group includes the yeast mitochondrial encoded c subunit, prokaryotic subunits c from proton-pumping ATPases of *E. coli*, *B. pseudofirmus* OF4, and *A. platensis*, and the subunit c from sodium ion-pumping ATPase of *I. tartaricus*.

Group 2 subunits c are typically 96-118 amino acids long. This group comprises sequences of prokaryotic subunits c. The topology of these subunits c is unclear. Their length is too short to form four transmembrane helices (as in certain V-type ATPases, see below) but longer than required for the formation of two transmembrane helices. However, the multiple amino acid sequence

Table 2.3: Properties of the four groups of subunits c

Group	Length (AAs)	Average (AAs)	Size (kDa)	Characteristics	Type
1	66-92	81	~8	two helices arranged in one hairpin	typical F-type ATPase
2	96-118	104	~8	putative cleaved-off N-terminal signal peptide and mature chain of two helices	some prokaryotic F-type and V-type ATPases
3	131-147	140	~8	cleaved-off mitochondrial targeting peptides and mature chain of two helices	mitochondrial F-type ATPase precursors
4	154-196	163	~16	four helices arranged in two hairpins, as results of gene duplication and gene fusion events	typical V-type ATPase

alignment and phylogenetic tree analysis show that this group may possess an N-terminal signal peptide, which has been shown to be removed from some mature V-type subunits c (Denda *et al.*, 1989; Ihara *et al.*, 1997; Yokoyama *et al.*, 2000). Without counting the signal peptide, the length of the group 2 subunits c is similar to that of the subunits belonging to group 1, typically 66-92 amino acids. Therefore, it is likely that after removal of the N-terminal peptides, the mature subunits c possess two transmembrane helices.

Group 3 subunits c are 131-147 amino acids long. This group comprises the precursors of the mitochondrial nuclear-encoded subunits c that possess an N-terminal mitochondrial targeting peptide. Their mature chain possesses two transmembrane helices and its topology resembles that of group 1 subunits c (see above). Most likely, as a nuclear gene product, a longer precursor is synthesized on free ribosomes and imported into the mitochondrion in a post-translational process that involves removal of the targeting peptide from the N-terminal region of the precursor (Schatz and Butow, 1983; Hay *et al.*, 1984). In *Neurospora crassa*, the targeting peptide is 66 amino acids long (Viebrock *et al.*, 1982). The mitochondrial targeting peptides of three mammalian isoforms of the nuclear-encoded subunit c are 61, 68 and 67 amino acids long, respectively, and they are expressed in a tissue-specific manner, whereas the mature subunits c are identical (Gay and Walker, 1985; Yan *et al.*, 1994; Vives-Bauza *et al.*, 2010).

Group 4 subunits c are 154-196 amino acids long. This group includes large V-ATPase subunits c found in eukaryotic vacuoles and some bacteria. They are characterized by possessing two hairpins adding up to four transmembrane helices per monomer, as a result of gene duplication and gene fusion events (Mandel *et al.*, 1988; Hirata *et al.*, 1997).

Finally, in few cases even larger subunits c can be identified, i.e. the *Methanopyrus kandleri* A-type subunit c is 1021 amino acids long and possesses 13 helical hairpins (Lolkema and Boekema, 2003). The discussion of these subunits is beyond the scope of our work.

2.3.3. The N-terminal region of the subunits c is not conserved

After reducing redundancies of the sequences, a set of 53 sequences was selected for multiple-sequence alignment. The alignment shows that the C-terminal two-helix-motif of the ATPase subunits c – including the two functional helices and the functionally important Glu/Asp residues – are highly conserved across all 4 groups described above. In contrast there is no sequence similarity at the N-terminus of subunits c (Figure 2.31 and see Figure 3.3). The subunits c of the different groups can also be clustered in different branches of a phylogenetic tree (see Figure 3.4). Group 1 possesses a short N-terminal tail and group 4 possesses a second copy of the two-helical region as a result of gene duplication and fusion events. In group 3 the N-terminal segment corresponds to the mitochondrial targeting peptide. Whereas, Group 2 is exceptional because it possesses a longer N-terminal segment, as described above. Secondary structure prediction servers,

Therefore, on the basis of this bioinformatics analysis, we propose that group 2 subunits c possess a putative N-terminal signal peptide that is cleaved off from the mature subunits c after translation. Following this consideration, two questions arise. First, what is the role of the N-terminal segment of group 2 subunits c *in vivo*? Second, are there signal peptides also in prokaryotic F-type subunits c, as identified in V-type subunits c? To answer these questions, we characterized the F-type subunit c of the hyperthermophilic bacterium *A. aeolicus*.

2.3.4. The mature form of the native subunit c of *A. aeolicus* is 81 amino acids long

A. aeolicus F₁F₀ ATP synthase subunit c is encoded by the *atpE* gene with a predicted length of 100 amino acids ($M_r = 10.2$ kDa). Therefore it belongs to group 2 as described above (typical length between 96-118 amino acids). Based on our multiple-sequence alignment, it possesses an N-terminal peptide of 29 amino acids preceding the conserved 2-helix C-terminal motif. If such an N-terminal peptide was a signal peptide and was cleaved off post-translationally, as expected from our classification, the mature subunit c would be composed of 81 amino acids and possess a $M_r = 8.1$ kDa.

We determined the true size of the mature subunit c of *A. aeolicus* using the isolated form of native *A. aeolicus* F₁F₀ ATP synthase (Peng *et al.*, 2006). After chloroform/methanol extraction (Cattell *et al.*, 1971; von Ballmoos *et al.*, 2002), the subunit c monomers were subjected to full-length MALDI-TOF mass spectrometric analysis. The results show that the molecular weight of subunit c is 8.1 kDa (Figure. 2.33), which is in very good agreement with our expectations, that the N-terminal 19 amino acids are not present in the mature subunits c and that mature subunit c of *A. aeolicus* possesses 2 transmembrane helices.

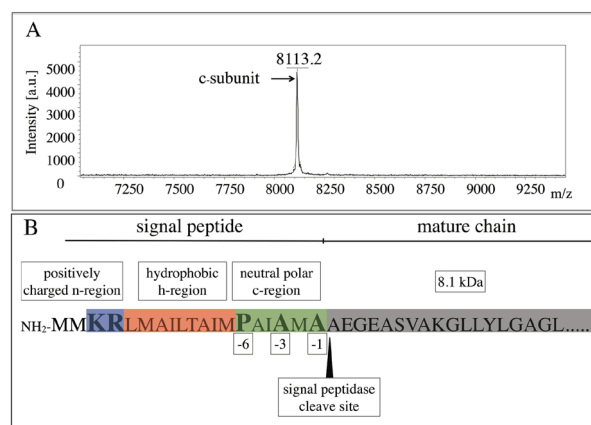


Figure 2.33. The subunit c of *A. aeolicus* F₁F₀ ATP synthase possesses an N-terminal signal peptide as indicated by mass spectrometry. (A) Full-length MALDI-TOF MS analysis of the native subunit c extracted from the membranes of *A. aeolicus*. The arrow indicates the peak at $m/z = 8.1$ kDa, corresponding to the size of the mature c-subunit (without the N-terminal 19 amino acids, which form a signal peptide). (B) The motifs characterizing the signal peptide of subunit c from *A. aeolicus* F₁F₀ ATP synthase. The signal peptide of the subunit c of *A. aeolicus* F₁F₀ ATP synthase is a bacterial signal peptide possessing typical SRP recognition features. These include positively charged n-region (in blue), hydrophobic h-region (in red) and neutral polar c-region (in green). Alanine (at -3 and -1 position) and proline (at -6 position), which are involved in signal peptide cleavage are marked in bold. N-terminal positively charged residues Arg and Lys, which interact with negatively charged phospholipid groups on the membrane during SRP- or Sec-mediated insertion, are also marked in bold.

2.3.5. The N-terminal signal peptide of *A. aeolicus* subunit c is recognized by *E. coli* and is crucial for membrane insertion of subunit c

Considering that homologous recombination is not feasible in *A. aeolicus* given the extreme environmental conditions that this organism requires to grow (Deckert *et al.*, 1998), we investigated the biological role of the N-terminal signal peptide of the *A. aeolicus* subunit c using an *E. coli*-based heterologous system. In particular, we used *E. coli* DK8 (*Δunc*) cells and the pCL21 vector into which we inserted an artificial operon encoding the whole 9-subunit *A. aeolicus* ATP synthase. We designed three different constructs. The first one contains the native subunit c encoding gene *atpE* including its N-terminal signal peptide (total of 100 amino acids). The second and third ones are two different signal-peptide-deletion variants. One, named here pCL21-ΔSP, does not code for the signal peptide of 18 amino acids after the start codon. Its total length is 82 amino acids, corresponding to a calculated molecular size of 8112.62/8243.1 Da (-/+ methionine M1). This construct should successfully be expressed in *E. coli* if the role of the N-terminal signal peptide was not biologically relevant. The other construct, named here pCL21-MEN, was designed by replacing the DNA encoding the N-terminal peptide of the *A. aeolicus* subunit c (a group 2 member) by that of the N-terminal segment of the *E. coli* subunit c (a group 1 member). The former protein possesses 29 amino acids and the latter 7 amino acids preceding the 2-helix conserved C-terminal motif, respectively. Therefore, the first 7 amino acids of the *E. coli* subunit c (MENLNMD) were designed to replace the native N-terminal 29 amino acids of *A. aeolicus*. The total length of the resulting subunit c encoded by pCL21-MEN is 78 amino acids, corresponding to a calculated molecular size of 7929.42/8060.62 Da (-/+ methionine M1). This third construct should be expressed in *E. coli* following the insertion pathway typical of the *E. coli* subunit c if the N-terminus of the latter were also important for its membrane insertion (see chapter 3.5.1).

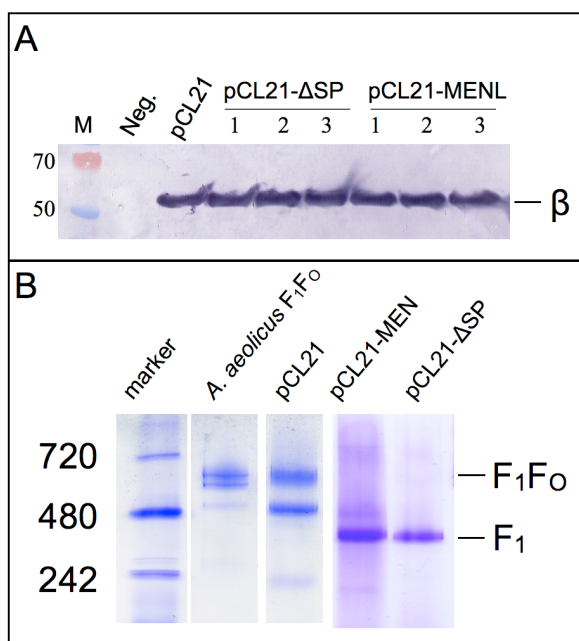


Figure 2.34. Production of the F₁ subcomplex in *E. coli*. (A) Identification of subunit β in the whole cell lysates of *E. coli* by western blot analysis using polyclonal polyHistidine antibodies. From left to right: Marker, Negative control, pCL21, pCL21-ΔSP (colonies 1 - 3), pCL21-MEN (colonies 1 - 3). (B) The assembled F₁F₀ and F₁ complexes purified from membranes of *E. coli*. The figure shows the characterization of F₁F₀ and F₁ complexes by BN-PAGE. From left to right: Marker, the purified F₁F₀ ATP synthase from *A. aeolicus*, pCL21, pCL21-MEN and pCL21-ΔSP. Arrows indicate the expected position of F₁F₀ and F₁ complexes. The results show that only the F₁ complex from the constructs pCL21-MEN and pCL21-ΔSP was successfully produced, whereas the entire EAF₁F₀ from the pCL21 construct was successfully produced.

All three constructs were transformed into *E. coli* and resulted in successful production of the F₁ part of the ATP synthase complex (Figure 2.34). However, when we analyzed the F₀ membrane-embedded part, we found strong differences in the production level of the three constructs. In particular, SDS-PAGE followed by Western blot analysis using a polyclonal antibody against the cytoplasmic loop of the subunit c revealed a clear signal for the subunit c in the pCL21 sample, but not in pCL21-ΔSP or pCL21-MEN (Figure. 2.35A). This result was confirmed by chloroform/methanol extraction followed by MALDI-TOF MS (Figure. 2.35B). In this experiment, the pCL21 sample shows a clear peak at an *m/z* value of 8112.7 Da corresponding to the size of the mature subunit c (8.1 kDa), while we could not detect any signal for the pCL21-ΔSP and pCL21-MEN constructs. Interestingly, the MS result also shows that the N-terminal 19 amino acids of subunit c were recognized and removed in *E. coli*, as a signal peptide. Furthermore, it suggests that the N-terminal signal peptide is crucial for the insertion of the *A. aeolicus* subunit c into the *E. coli* membranes and consequently that the *A. aeolicus* subunit c follows a membrane insertion pathway existing in *E. coli*, which is different from the one used by the native *E. coli* subunit c.

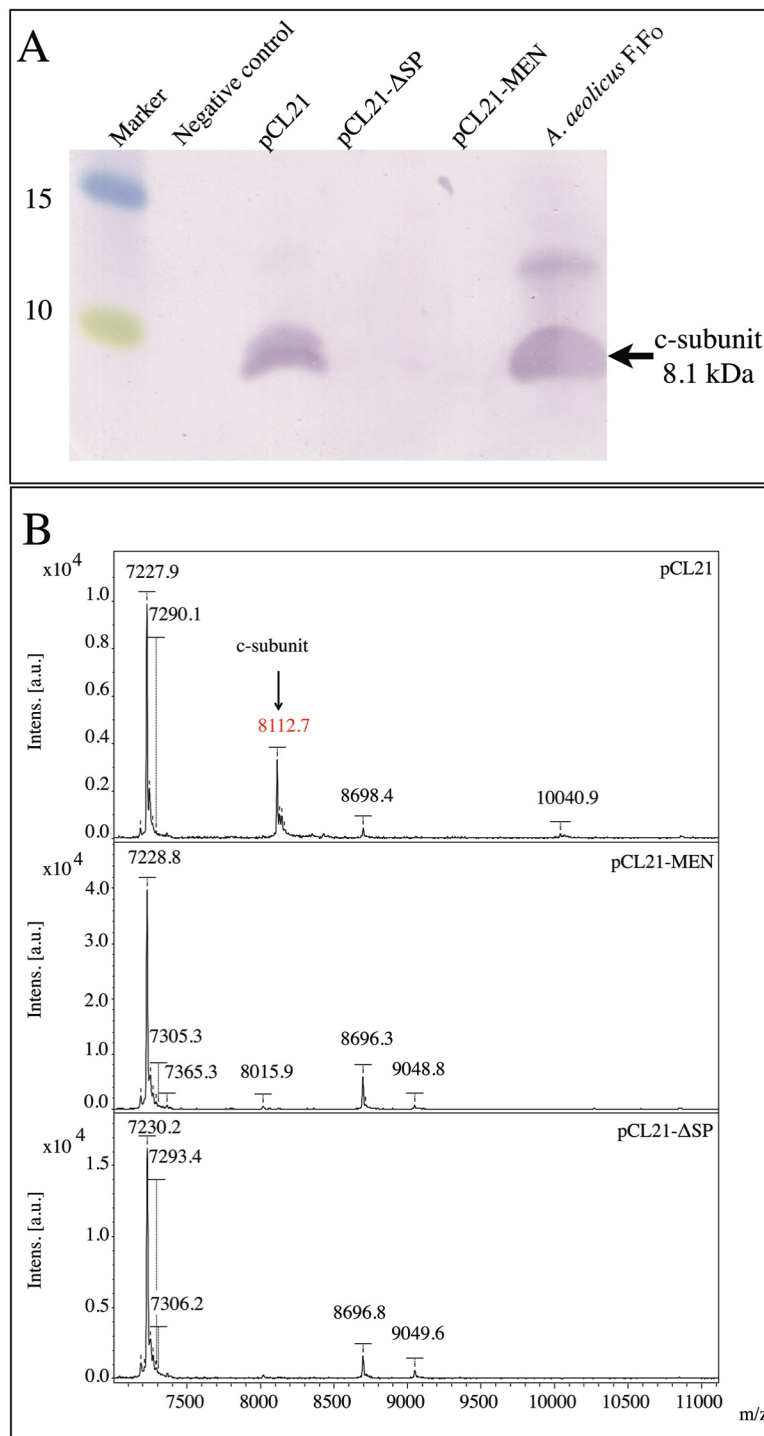


Figure 2.35. Identification of subunit c monomers in the membranes of *E. coli*. (A) Western blot analysis using polyclonal antibodies against the cytoplasmic loop of the subunit c from the *A. aeolicus* ATP synthase. Arrows indicate the expected position of the subunit c. From left to right: Marker, Negative control, pCL21, pCL21-ΔSP, pCL21-MEN, the purified F₁F₀ ATP synthase from *A. aeolicus*. (B) MALDI-TOF MS analysis of full-length subunit c from *A. aeolicus* ATP synthase, heterologously produced in *E. coli*. From top to bottom: pCL21, without SP: 8112.62 Da, with SP: 10041.16/10172.35 Da (-/+ methionine), pCL21-MEN, 7929.42/8060.62 Da (-/+ methionine), pCL21-ΔSP, 8112.62/8243.1 Da (-/+ methionine). The results show that only the subunit c from the pCL21 construct was successfully expressed, and that it was expressed in the mature form, lacking the signal peptide.

3. Discussion

3.1. Producing proteins from thermophilic organisms in mesophilic hosts

Hyperthermophiles, i.e. *A. aeolicus*, live at temperatures above 80 °C (Horikoshi, 1998; Huber and Eder, 2006). At such high temperatures, proteins and nucleic acids normally denature and the membrane fluidity increases to lethal levels. The solubility of gasses is also altered affecting O₂ or CO₂ dependent metabolic pathways. Therefore, hyperthermophilic organisms have developed a number of different strategies to cope with the high temperature stress (Rothschild and Mancinelli, 2001; Jaenicke and Sterner, 2006; Gerday and Glansdorff, 2007). (i) Their membrane composition is different from that of mesophilic organisms to maintain optimal fluidity and reduced permeability (Jaenicke and Sterner, 2006; Gerday and Glansdorff, 2007). (ii) Their DNA stability is enhanced by elevated salt concentrations, polyamines, cationic proteins and supercoiling (Horikoshi, 1998). (iii) Their proteins are stabilized by increasing ion-pair content and compactness (Rothschild and Mancinelli, 2001; Jaenicke and Sterner, 2006; Gerday and Glansdorff, 2007).

Some of these strategies may interfere with the production of proteins from thermophilic organisms in mesophilic hosts, i.e. the different lipid composition of the membranes. However, many reports demonstrate the successful production of proteins from thermophilic organisms in mesophilic hosts (Klinger *et al.*, 2003; Kohlstadt *et al.*, 2008; Obuchi *et al.*, 2009; Stewart *et al.*, 2012), including membrane proteins (Schutz *et al.*, 2003; Rollauer *et al.*, 2012) and large multimeric complexes (Matsui and Yoshida, 1995; McMillan *et al.*, 2007). In this work, we succeeded for the first time to produce a large multimeric membrane protein complex from a hyperthermophilic organism in a mesophilic host using an artificial operon.

3.2. The strategy for the heterologous production of A. aeolicus ATP synthase in E. coli

To obtain the nine-subunit *A. aeolicus* ATP synthase complex in *E. coli*, two different strategies could theoretically be used, namely (i) production of the individual subunits followed by *in vitro* reconstitution (Futai *et al.*, 1988; Imamura *et al.*, 2006), or (ii) co-production of all the subunits in the same host. However, in practice, the first strategy is not applicable for the whole ATP synthase complex, because certain subunits (i.e. a) cannot be obtained well or at all in isolation [see Results and (von Meyenburg *et al.*, 1985; Arechaga *et al.*, 2003)]. Therefore, the second strategy was followed.

Little is known about transcription, translation and assembly of the *A. aeolicus* F₁F₀ ATP synthase in native *A. aeolicus* cells, but these processes must be rather complex in respect to other organisms because the *atp* genes are dispersed over four loci throughout the *A. aeolicus* genome (see chapter 1.3).

The *E. coli* F₁F₀ ATP synthase is better-studied and extensive characterization has been carried out to understand its expression and assembly (Futai *et al.*, 1988). The *atp* operon of *E. coli* consists of nine genes ordered *atpIBEFHAGDC* (Kanazawa and Futai, 1982; Walker *et al.*, 1984). It has a single promoter which initiates transcription 73 bp upstream of the start codon of gene I (*atpI*) (Kanazawa *et al.*, 1981; von Meyenburg *et al.*, 1982; Jones *et al.*, 1983; Porter *et al.*, 1983; Nielsen *et al.*, 1984). The *atp* operon is transcribed to produce a single, large polycistronic mRNA containing all nine cistrons, which are then translated at different rates to match the F₁F₀ ATP synthase stoichiometry (McCarthy *et al.*, 1985; McCarthy, 1988).

Furthermore, recombinant *E. coli* F₁F₀ ATP synthase was overproduced homologously in the *E. coli* DK8 (Δatp) strain (Klionsky *et al.*, 1984) with a recombinant plasmid containing the entire *atp* operon (Noji *et al.*, 1999; Ishmukhametov *et al.*, 2005). The *E. coli* strain DK8 (Δatp) was also used as a host strain for producing heterologously recombinant F₁- $\alpha_3\beta_3\gamma$ of *Bacillus* PS3 (Matsui and Yoshida, 1995), the entire F₁F₀ complexes of *I. tartaricus* (Hakulinen *et al.*, 2012), *C. thermarum* TA2.A1 (McMillan *et al.*, 2007) and *A. woodii* (Brandt *et al.*, 2013).

Because *E. coli* is well studied, and because it was used successfully for so many other ATP synthases, we chose it as a host organism also for our study. Given that *A. aeolicus* does not possess one *atp* operon, we preferred to design an artificial operon as similar as possible to the *E. coli* native one, rather than co-expressing the genes from several expression vectors, each possessing compatible origins of replication and independent antibiotic selection for maintenance (Tolia and Joshua-Tor, 2006). A number of limitations had to be taken into account for obtaining successful production and correct assembly of the whole complex, namely (i) composition of *atp* genes (see chapter 3.3.1), (ii) choice of intergenic regions and regulatory elements (see chapter 3.3.2), (iii) handling of gene overlaps (see chapter 3.3.3), and (iv) optimization of codon usage (see chapter 3.3.4). These limitations and the corresponding strategies used to overcome them will be discussed hereafter.

3.3. Properties of the artificial *atp* operon

3.3.1. Composition of the *atp* operon genes

When the artificial operon strategy was chosen, constructing an operon implied considering the following points related to the composition of the *atp* genes in *A. aeolicus*. First, genes for different isoforms of various subunits are present in the *A. aeolicus* genome (Deckert *et al.*, 1998), but only

some of them are to form the mature ATP synthase in native cells (Peng *et al.*, 2006; Guiral *et al.*, 2009). Therefore, such isoforms were not used in our work, i.e. gene *atpG2* encoding an isoform of subunit γ . In contrast, subunits present in different isoforms were all included in our artificial operon. This is specifically the case for subunits b_1 and b_2 that are encoded by a native operon together with *atpH* and form a hetero- and not a homo-dimeric peripheral stalk (Peng *et al.*, 2006; Guiral *et al.*, 2009). Moreover, *A. aeolicus* does not possess a homologue of the *atpI* gene, which is found in many other ATP synthase operons, including all those used for expression to date (Walker *et al.*, 1984; Santana *et al.*, 1994; Yokoyama *et al.*, 2000; Meier *et al.*, 2003; Keis *et al.*, 2004; Brandt *et al.*, 2013). Although not present in the assembled holoenzyme, the product of the gene *atpI* is involved in c-ring assembly (Suzuki *et al.*, 2007; Ozaki *et al.*, 2008; Brandt *et al.*, 2013) and it was proposed to be species-specific, since *atpI* from *E. coli* could not substitute *atpI* from *A. woodii* (Brandt *et al.*, 2013). However, compared to the other *atp* genes, sequence conservation of the *atpI* gene is low. For example for *C. pasteurianum* and *E. coli*, the sequence identity of the product of gene *atpI* is 19%, while for all other subunits it is 22-72% (Das and Ljungdahl, 2003). It is possible that *atpI* is so divergent in *A. aeolicus* that it was not identified to date, or that it is not conserved across species because it is not essential for all organisms (Das and Ljungdahl, 2003; Liu *et al.*, 2013), and thus not present at all in *A. aeolicus*. We therefore attempted to produce *A. aeolicus* ATP synthase without the *atpI* gene in our artificial operon. Our successful expression results suggest that either *atpI* is not present in *A. aeolicus*, or it is not as species-specific as the one from *A. woodii* (Brandt *et al.*, 2013), so that the correct assembly of our EAF₁F_O could be facilitated by *E. coli atpI* gene, which is present in the DK8 strain that we used.

3.3.2. Regulation of expression: translation initiation regions (TIR)

ATP synthase is a heteromultimeric complex in which the correct reciprocal stoichiometry of all subunits needs to be carefully regulated. It is known that differential expression of the *E. coli atp* genes is controlled at two levels, post-transcriptionally via different types of translation initiation regions (TIR), and at the mRNA level due to the different mRNA stabilities (Schramm *et al.*, 1996). The post-transcriptional level of control via TIR is the predominant one and the one for which we had better control in the design of the artificial operon. The TIR of a gene includes the translation start codon, the Shine-Dalgarno (SD) region and the N-terminal region of the structural gene (McCarthy, 1988). For our *A. aeolicus atp* artificial operon, we designed TIRs on the basis of the following considerations.

First, given that the N-terminal regions of each gene are part of TIRs, we preferred to use complete TIRs from the *A. aeolicus* genes, instead of using mixed TIRs composed of segments from the host organism *E. coli* and segments contributed by *A. aeolicus* genes. This strategy simplified also the handling of gene overlaps (*vide infra*).

Second, we delimited *A. aeolicus* TIRs to include 30 bp upstream of the start codon of six genes, *atpB*, *atpE*, *atpF1*, *atpA*, *atpG* and *atpC* (see Table 3.1 and Table 3.2). For *atpF2* and *atpH*, we did not modify their TIRs because these genes are encoded in a native operon together with *atpF1*. Finally, for *atpD* (subunit β), encoded in a native operon with *atpG*, 27 bp were introduced after its start codon to form an His₆ purification tag. We chose to include 30 bp upstream of each gene because it had been previously observed that certain *atp* genes possess important translational regulators in this region, i.e. a translational enhancer of *atpE* (subunit c) in *E. coli* (McCarthy *et al.*, 1985). In addition, these 30-bp regions also include all other regulatory elements, such as the Shine-Dalgarno (SD) sequences and the start codons (see Table 3.2 and Table 3.3).

Third, we limited as much as possible all modifications of TIRs affecting the intercistronic distance between neighboring genes, because it is known that coupled translation of neighbouring genes also regulates translation initiation in *E. coli* (Gerstel and McCarthy, 1989). For instance, in the *E. coli atp* operon, three genes are more tightly coupled (*atpFH* and *atpHA*) because they are separated by short intercistronic sequences (Hellmuth *et al.*, 1991). Increasing the distance between them decreases the degree of coupling, even if the original *atpA* TIR structure is maintained (Gerstel and McCarthy, 1989). Therefore, the intercistronic distance of genes belonging to operons in the *A. aeolicus* genome were not modified in our artificial *atp* operon. However, restriction sites had to be introduced for the other genes, *atpB*, *atpF1*, *atpA*, *atpG* and *atpC*, for cloning purposes. Our results show that the introduction of such restriction sites did not preclude the expression of EAF₁F₀.

Detailed considerations on the choice of different regulating elements are reported separately in the following paragraphs, which discuss features of the intergenic regions (see chapter 3.2.2.1), the Shine-Dalgarno (SD) regions (see chapter 3.2.2.2), and the translation start codons (see chapter 3.2.2.3).

Table 3.2: Length of TIRs selected for the *A. aeolicus atp* genes

Gene	Length (bp)	Upstream (bp)	Downstream (bp)	5'restriction site	3'restriction site
<i>atpB</i>	651	-30	+20	KpnI&NdeI	-
<i>atpE</i>	303	-30	+6	-	BglII
<i>atpF1/F2/H</i>	435/558/546	-30	+10	BglII	BamHI
<i>atpA</i>	1512	-30	+10	BamHI	XbaI
<i>atpG/D</i>	876/1437	-30	+6	XbaI	SaII
<i>atpC</i>	399	-30	+31	SaII	SmaI&PstI

Table 3.2: Sequences of TIRs selected for *A. aeolicus* *atp* operon

Subunit	DNA sequence	Stoichiometry
a	TCTGAGCCAATTGCAAAAAGAGGTAAGGGAAATGGAGTACTCGCACGTAGT	1
c	CTTATAGTTAAATAAGCTTTAAGGAGGTAGGTGATGAAGAGGTTAATGGC	?
b ₁	ATTGCTATAATTGTTTAGCGGAGGAGAAGAATGGACATAGGAGTAATGCC	1
b ₂	TGTAAAGAAAATTTTGAGAGAGGGCGGCGTGTGGTGAGGTTGATAAGTTT	1
δ	CTCAGTTAAAGCTCCTGGAGGAGAGGAAGAATGCTTAAAGAGGAAAGAACT	1
α	AAACCTTTAAAGAAGGTTAGGAGGTAGAGTATGGCTACACTGACTTATGA	3
γ	TTAGACATTAGTTTATAATAAGTAGCGTTATGGCGAAACTTCTCCAG	1
β	CTCTTAAAGCACAATAAAGGAGGTTTATAGATGGCGGAAGTGATTAAGGG	3
ε	TTGGACTTCTCTGGTATAATTTAGGGATTATGATACAGGTTGAAATAGT	1

N.B.: start codons are in red, stop codons are underlined and the putative RBSs are highlighted by gray boxes.

3.3.2.1. Intergenic regions

To construct the artificial *atp* operon for EAF₁F_O we included intergenic regions taken from *A. aeolicus* genome. We selected these regions manually (see Table 3.1 and Table 3.2), following a similar pattern as in the *E. coli* native *atp* operon (expression vector pKH7 (Noji *et al.*, 1999)). In the latter operon, the TIR of the first gene, *atpB*, is located +84 bp downstream of the promoter and the intergenic region between each pair of neighbouring genes are: *atpB-atpE* (46 bp), *atpE-atpF1* (58 bp), *atpF2-atpH* (14bp), *atpH-atpA* (12bp), *atpA-atpG* (50 bp), *atpG-atpD* (26bp), and *atpD-atpC* (12bp). +6 bp downstream of the stop codon of *atpC* there is a transcription terminator (46 bp). Correspondingly, in our artificial *atp* operon, the translation initiation site of *atpB* is located +60 bp downstream of the *trc* promoter and the intergenic regions between pairs of neighbour genes are *atpB-atpE* (50 bp), *atpE-atpF1* (42 bp), *atpF1-atpF2* (overlapping 1 bp), *atpF2-atpH* (overlapping 8 bp), *atpH-atpA* (46 bp), *atpA-atpG* (46 bp), *atpG-atpD* (13 bp), and *atpD-atpC* (42 bp). +94 bp downstream of the stop codon of *atpC*, there are three other terminators, *rrnB*, *rrnB-T1* and *rrnB-T2* belonging to vector pTrc99A.

3.3.2.2. Ribosome binding sites (RBS)

SD sequences are important elements in TIRs. Therefore, the ribosome binding sites (RBS) and the spacing regions between them and the start codons of the respective genes were designed considering their high level of conservation for all subunits (except for genes *atpG* and *atpC*, see Table 3.2). They are: AGAGG for *atpB* (5 bp from -13 bp to -9 bp), AAGGAGG for *atpE* (7 bp from -10 bp to -4 bp), GGAGG for *atpF1* (5 bp from -11 bp to -7 bp), AGAGG for *atpF2* (5 bp from -12 bp to -8 bp), two constitutive RBS for *atpH* (GGAGG, 5 bp from -14 bp to -10 bp and AGAGG, 5 bp from -9 bp to -5 bp), AGGAGG for *atpA* (6 bp from -12 bp to -7 bp), AGGAGG for *atpD* (6 bp from -13 bp to -8 bp). The rare putative RBS of *atpG* and *atpC* are AAG (3 bp from -10 bp to -8 bp), and AGGG (4 bp from -7 bp to -4 bp), respectively.

3.3.2.3. Start and stop codons

ATG is the most frequently used start codon. In *A. aeolicus*, eight of nine genes begin with the typical ATG start codon. Only *atpE*, represents an exception, beginning with GTG (see Table 3.3). However, *atp* genes in *E. coli*, *B. megaterium*, *B. subtilis*, *I. tartaricus* also use GTG as a start codon (Walker *et al.*, 1984; Brusilow *et al.*, 1989; Santana *et al.*, 1994; Meier *et al.*, 2003). Therefore, the native start codon of *atpE* was not changed in our artificial operon. No rare start codon TTG was observed in *A. aeolicus atp* genes. TTG was instead proposed to play a role in regulating the respective stoichiometry of subunits c and a in *C. thermarum* TA2.A1 (Keis *et al.*, 2004), it is a shared feature of the *atpB* genes from *B. subtilis* (Brusilow *et al.*, 1989), *B. megaterium* (Santana *et al.*, 1994) and *C. thermarum* TA2.A1 (Keis *et al.*, 2004), and it is also present in subunit b of *I. tartaricus* ATP synthase (Meier *et al.*, 2003) and subunits α and β in *C. thermoaceticum* TA2.A1 (Das and Ljungdahl, 1997).

In *A. aeolicus*, all genes possess TAA as a stop codon, except for genes encoding subunits a and b₁, which use TGA as a stop codon. All native stop codons were maintained in our artificial operon. Table 3.3 reports a summary of the start and stop codons used for the artificial operon.

3.3.3. Gene overlaps

Another feature that distinguishes *A. aeolicus* ATP synthase is the presence of overlaps in genes *atpF1*, *atpF2* and *atpH*. In *E. coli*, there are no overlaps in genes of *atp* operon. However, overlaps are found in other operons, i.e. in *trp* (Platt and Yanofsky, 1975; Oppenheim and Yanofsky, 1980), *his* (Barnes and Tuley, 1983), *gal* (McKenney *et al.*, 1981), *frd* (Cole *et al.*, 1982), and *tox* (Yamamoto *et al.*, 1982). Therefore, *E. coli* must be able to use gene overlaps. Furthermore, other *atp* operons with gene overlaps had already been successfully produced in *E. coli*, i.e. *C. thermarum* TA2 (McMillan *et al.*, 2007), suggesting that foreign *atp* genes with overlaps can be used by *E. coli*. Our results confirm these observations, showing that gene overlaps between *atpF1* and *atpF2*, *atpF2* and *atpH* can be recognized by *E. coli* to coordinate gene expression of subunit b₁, b₂ and δ .

Table 3.3: Start codon and stop codon of the *A. aeolicus atp* genes

Gene	Start codon	Stop codon
<i>atpB</i>	ATG	TGA
<i>atpE</i>	GTG	TAA
<i>atpF₁</i>	ATG	TGA
<i>atpF₂</i>	ATG	TAA
<i>atpH</i>	ATG	TAA
<i>atpA</i>	ATG	TAA
<i>atpG</i>	ATG	TAA
<i>atpD</i>	ATG	TAA
<i>atpC</i>	ATG	TAA

3.3.4. Codon usage optimization

The mean difference of codon usage between *E. coli* and *A. aeolicus* is 30.91%. To obtain the best expression in *E. coli*, codon usage bias has been taken into account.

The codon usage bias can be overcome by two common approaches. First, synthetic genes optimized for host codon usage may be employed. Subunit c of spinach chloroplast ATP synthase was successfully expressed and purified in *E. coli* using a synthetic gene with codon optimization (Lawrence *et al.*, 2011). However, in general, this approach is more expensive especially for operons of large size. Second, the copy number of the tRNAs specific for rare codons in the host may be increased, i.e. co-expressing such tRNAs with the target gene using the commercial pRARE vector (Novagen).

Therefore, for expressing the 7 kb long artificial *atp* operon of EAF₁F₀, we selected this second approach. Notably, the replication origin pBR322 of vector pTrc99A, used to create vectors pCL11 and pCL21 is compatible with the p15A origin of the pRARE vector. The benefit of using pRARE was clearly demonstrated by the increase in expression levels of the F₁- $\alpha_3\beta_3\gamma$ subcomplex in *E. coli* C43(DE3) cell and of subunit c in *E. coli* DK8 cells (Figure 3.1).

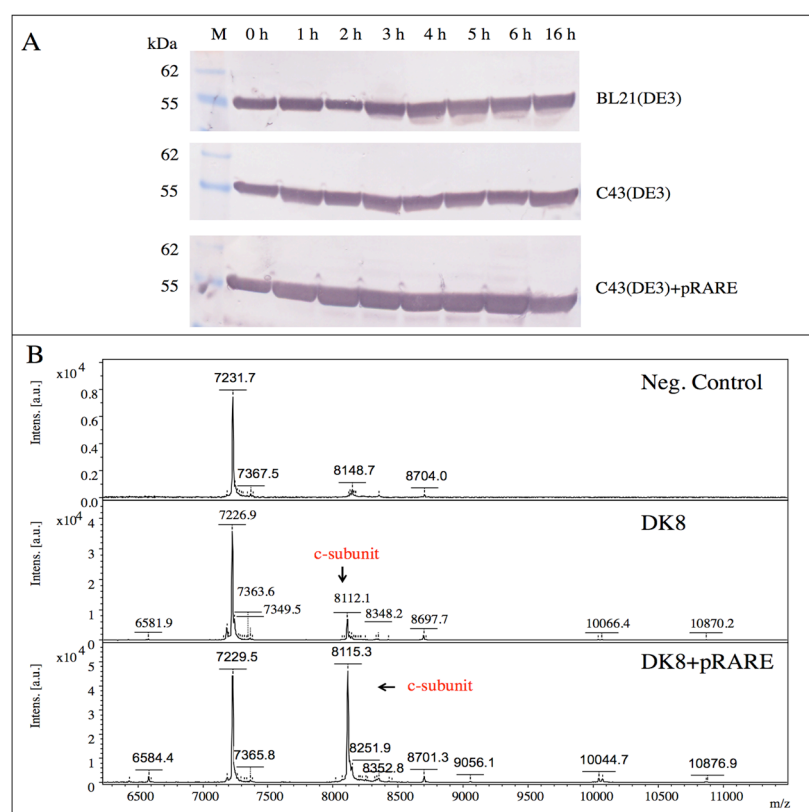


Figure 3.1. Effect of codon usage on heterologous expression of *atp* genes in *E. coli* shown by Western blot analysis and mass spectrometry. Co-expression of the pRare vector with the target genes can increase the levels of subunit β , as detected by Western blot analysis against poly-histidine antibody (A), and of subunit c, as detected by MALDI-TOF after chloroform-methanol extraction.

3.3.5. Position of purification tag

The design of intermediate constructs pCL11 and pCL12 ($F_1\text{-}\alpha\beta\gamma$ and $F_1\text{-}\alpha\beta\gamma\epsilon$) allowed us to define a favourable position for inserting a purification and detection tag, which was also used successively for the characterization of the full ATP synthase complex. We decided to fuse a His₆-tag to the N-terminus of subunit β on the basis of two observations. First, this peptide is exposed to solvent in all known 3-D structures of the F_1 complex (Abrahams *et al.*, 1994), so that a tag in that position should not interfere with structural assembly or with substrate binding. A tag in this position had also been successfully used for purification and single molecule rotation experiments on other ATP synthases (Noji *et al.*, 1997; Noji *et al.*, 1999; Ishmukhametov *et al.*, 2005; Matthies *et al.*, 2011). Second the *atpD* gene encoding subunit β is located at the 3' end of artificial operon *atpAGD*, so that detection of subunit β using Western blots provides an easy estimation of whether the genes of all other subunits are expressed. The same holds for the expression of the entire F_1F_0 ATP synthase complex from the artificial operon, in which gene *atpD* is also located at the 3' end.

3.4. The structure and function of heterologously produced *A. aeolicus* F_1F_0 ATP synthase

Using the expression strategy described above (see chapters 2.2.2 and 3.2), we were able to produce *A. aeolicus* ATP synthase heterologously in *E. coli* with a yield of approximately 0.06 – 0.15 mg of pure EAF₁F₀ per liter of cell culture (or 0.03 mg EAF₁F₀ per gram of *E. coli* wet cell pellet). This yield is comparable to that obtained for F_1F_0 ATP synthase of *I. tartaricus* heterologously produced in *E. coli* (Hakulinen, 2012), and slightly lower than that obtained for homologously recombinant V₁V₀ ATP synthase from *T. thermophilus* (0.15 mg V₁V₀ ATP synthase per gram of wet cell pellet) (Yokoyama *et al.*, 2003). While expression is still low for crystallographic studies, it did allow us to undertake studies and to analyse the structural and functional properties of *A. aeolicus* ATP synthase.

On the one hand, our enzymatic studies proved that the ATP hydrolysis activity of the pure EAF₁F₀ is comparable to that of the pure AAF₁F₀. The temperature dependence of EAF₁F₀ follows a similar trend as that of AAF₁F₀ and of other respiratory membrane complexes of *A. aeolicus* (Peng *et al.*, 2003; Marcia *et al.*, 2010) increasing substantially at temperatures above 60 °C and reaching its maximum at 80 °C. This result indicates that the F_1F_0 ATP synthase, produced from mesophilic organism *E. coli*, could cope with the high temperature stress. The protein complex might be stabilized by different strategies: (i) the presence of additional inter-subunit interactions (Russell *et al.*, 1997); (ii) the secondary structural elements stabilized by shorter surface loop connections, by optimized electrostatic and hydrophobic interactions, by disulfide bridges, by a pronounced hydrophobicity in the protein core and by a higher helical propensity of residues in α -helices

(Rothschild and Mancinelli, 2001; Jaenicke and Sterner, 2006; Gerday and Glansdorff, 2007). Moreover, in general, the rate of ATP hydrolysis of EAF₁F₀ and AAF₁F₀ is in the same range as that of the F₁F₀ ATP synthase from *E. coli* (Ishmukhametov *et al.*, 2005). EAF₁F₀ is actually slightly less active than AAF₁F₀ (43 %), but such a trend is not unusual when comparing the activity of heterologously produced *vs* native enzymes. For instance, the F₁F₀ ATP synthase of *I. tartaricus* is ~3 fold less active when produced heterologously in *E. coli* (Hakulinen, 2012) in respect to the same enzyme isolated from native cells (Neumann *et al.*, 1998). Therefore, our expression system opens the way to previously impossible functional studies on *A. aeolicus* ATP synthase, i.e. by site-directed mutagenesis, cell viability studies, *in vivo* complementation experiments and *in vitro* enzymatic assays.

On the other hand, our single-particle EM reconstruction shows that EAF₁F₀ has an identical structure to AAF₁F₀. AAF₁F₀ is 230 Å long, its globular F₁ subcomplex possesses a diameter of 110 Å, and the c-ring in the F₀ subcomplex is ~ 90 Å wide and ~ 70 Å high (Peng *et al.*, 2006). These dimensions match well with those measured for EAF₁F₀ (see Figure 2.30). More importantly, the AAF₁F₀ single-particle reconstruction shows much stronger density for its peripheral stalk subunits b₁b₂ (Peng *et al.*, 2006) than F₁F₀ ATP synthases from other bacteria, e.g. *E. coli* (Bottcher *et al.*, 2000) and *C. thermarum* strain TA2. A1 (Matthies *et al.*, 2011). In this work, we observe the same strong signal for subunits b₁b₂ in EAF₁F₀, suggesting that in respect to other ATP synthases the peripheral stalk is more rigid and it remains intact during purification in *A. aeolicus* ATP synthase, independent if the enzyme is produced from native source or heterologously.

3.5. Novel properties of A. aeolicus ATP synthase discovered using EAF₁F₀

The heterologous expression system implemented with this work enabled us to define new properties of ATP synthase, which would not have been possible to discover without the heterologous expression of EAF₁F₀. Such properties specifically pertain to subunits b₁b₂, γ and c, which present unique features in *A. aeolicus* ATP synthase.

Subunits b₁ and b₂ had previously been co-identified in the pure AAF₁F₀ (Peng *et al.*, 2006), but their mode of association in the enzyme complex was unclear. Our heterologous expression study now clarifies that subunits b₁ and b₂ can associate to form a complex *in vitro* and thus corroborates the previous hypothesis that *A. aeolicus* ATP synthase possesses a heterodimeric peripheral stalk and is therefore unique among ATP synthases of non-photosynthetic organisms (Peng *et al.*, 2006). Moreover, single-particle EM shows that EAF₁F₀ possesses a bent central stalk, comprised of subunits ε and γ, as it had already been observed also in AAF₁F₀ (Peng *et al.*, 2006). The bent

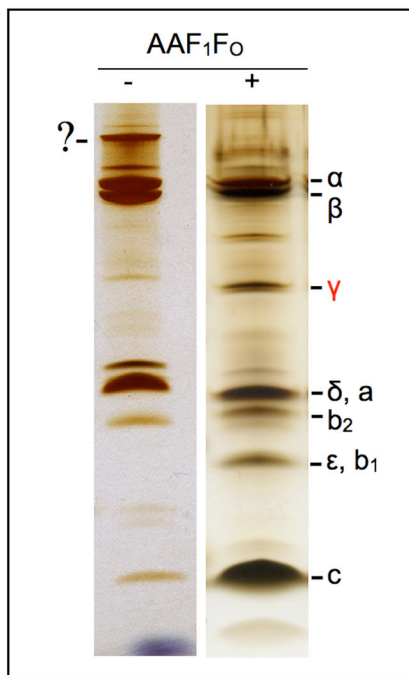


Figure 3.2. Silver-stained SDS-PAGE of AAF₁F_O with (+) or without (-) heat treatment. The heat treatment is required for clear resolution of subunit γ band (red label) by SDS-PAGE. The arrow corresponds to an uncharacterized protein band with molecular weight higher than that of subunits α and β , and present only in the unheated sample. Possibly, such band corresponds to an SDS resistant assembly of ATP synthase subunits (including subunit γ), which fully dissociate only upon heat treatment.

conformation of the stalk, especially the bent subunit γ may be indicative of a specific functional property of *A. aeolicus* ATP synthase, because in other organisms, i.e. the thermoalkaliphic bacterium *C. thermarum* TA2.A1, such a feature was associated to a blockage in ATP hydrolysis activity (Stocker *et al.*, 2007). It was observed that in the bent conformation subunit γ is stabilized by two salt bridges with subunit β_{Empty} (with no bound-nucleotide) (γ R10 and β D372, γ R20 and β D375), and thus prevents rotation during ATP hydrolysis (Stocker *et al.*, 2007). Interestingly, ATP hydrolysis activity of *A. aeolicus* F₁F_O ATP synthase is only observed at temperatures above 60 °C, while purification and EM were performed at low-to-room temperature. Additionally the subcomplexes of *A. aeolicus* ATP synthase are SDS-resistant and can only be fully dissociated by heat treatment. Specifically,

such heat treatment is necessary for visualizing subunit γ (Figure 3.2). Therefore, it may be that we have trapped ATP synthase in a blocked state with our purification and that higher temperatures are necessary to trigger the mobilization of subunit γ from a tight association with subunits α/β and to switch the ATP synthase from a resting state (typical of low-to-room temperatures) to the active state (formed at higher temperatures).

Additionally, such heat treatment is also necessary for visualizing subunit ϵ in SDS-PAGE. Interestingly, also subunit ϵ was described as an inhibitor of ATP hydrolysis activity in many bacteria i.e. *E. coli*, *Bacillus* PS3 and *C. thermarum* TA2.A1 (Tsunoda *et al.*, 2001; Suzuki *et al.*, 2003; Keis *et al.*, 2006).

Finally, we discovered that subunit c unexpectedly contains an N-terminal signal peptide, which is strictly necessary for its correct membrane insertion. Since *A. aeolicus* subunit c could not be obtained in isolation (see chapter 2.2.3.4), the study of its N-terminal properties would not have been possible without the heterologous expression of EAF₁F_O. Considering that such a study on subunit c gave us broader phylogenetic insights into subunits c and led us to propose a new

assembly pathway for the whole ATP synthase complex, a more detailed discussion of these topics is reported separately in the following paragraphs.

3.5.1. A new phylogenetic classification of F₁F₀ ATP synthase based on the N-terminal sequence of their subunit c

Two major determinants influence the membrane insertion mechanism of membrane proteins. The first determinant consists of the interaction with the translocation machinery. To this respect, it is known that most membrane proteins are targeted to the Sec translocase (SecYEG) either by the signal recognition particle (SRP) and its receptor FstY (Ulbrandt *et al.*, 1997; Herskovits *et al.*, 2000; de Gier and Luirink, 2001) or by the translocase subunit SecB (Valent *et al.*, 1998; Koch *et al.*, 1999). Other membrane proteins follow a Sec-independent mechanism, mediated by YidC (Samuelson *et al.*, 2000; van Bloois *et al.*, 2004; van der Laan *et al.*, 2004; Yi *et al.*, 2004). The second determinant consists of the distribution of the charges along the protein, because this distribution determines the topology of the protein, i.e. in accordance to the positive-inside rule (von Heijne and Gavel, 1988). To date, the membrane insertion pathway used by the ATP synthase subunits c was characterized in *E. coli* and yeast mitochondria. In *E. coli*, the subunit c does not possess a signal peptide and is a substrate of the YidC pathway (van der Laan *et al.*, 2004). In mitochondria, the subunit c is inserted into the membranes by Oxa 1, a YidC-homolog (Jia *et al.*, 2007). Furthermore, it is known that at least two of three positively charged residues (K34, R41, R50) in the cytoplasmic loop of *E. coli* subunits c are required for YidC-mediated membrane targeting (Kol *et al.*, 2008), although they do not need to occupy conserved positions in the loop region (Kol *et al.*, 2008). However, it is not known whether these residues are the only requirement for regulating membrane insertion and topology of subunits c, or whether other regions of the protein are also important (Dalbey *et al.*, 1995). For instance, the role of the N-terminal region of subunits c is unclear, despite this region being known to control the topology of proteins with two transmembrane segments (Nilsson and von Heijne, 1990; Gafvelin *et al.*, 1997), like most subunits c, in particular via charged amino acids.

By performing an alignment of a large set of subunits c from different organisms, we noticed that their C-terminal two-helix-motif is well conserved, and so are the three positively charged residues involved in YidC recognition (see Figure 3.3). However, we noticed that subunits c are characterized by a remarkable heterogeneity at their N-terminus (see Figure 3.3). Based on this heterogeneity we have defined 4 groups of subunits c (see Figure 3.4). While the N-terminus of group 4 has relatively complex features, being composed of two hydrophobic N-terminal helices derived from gene duplication of the C-terminal two-helix-motif, the N-terminus of group 1 and 3 and that of group 2 subunits c are simpler and possess opposite characteristics.

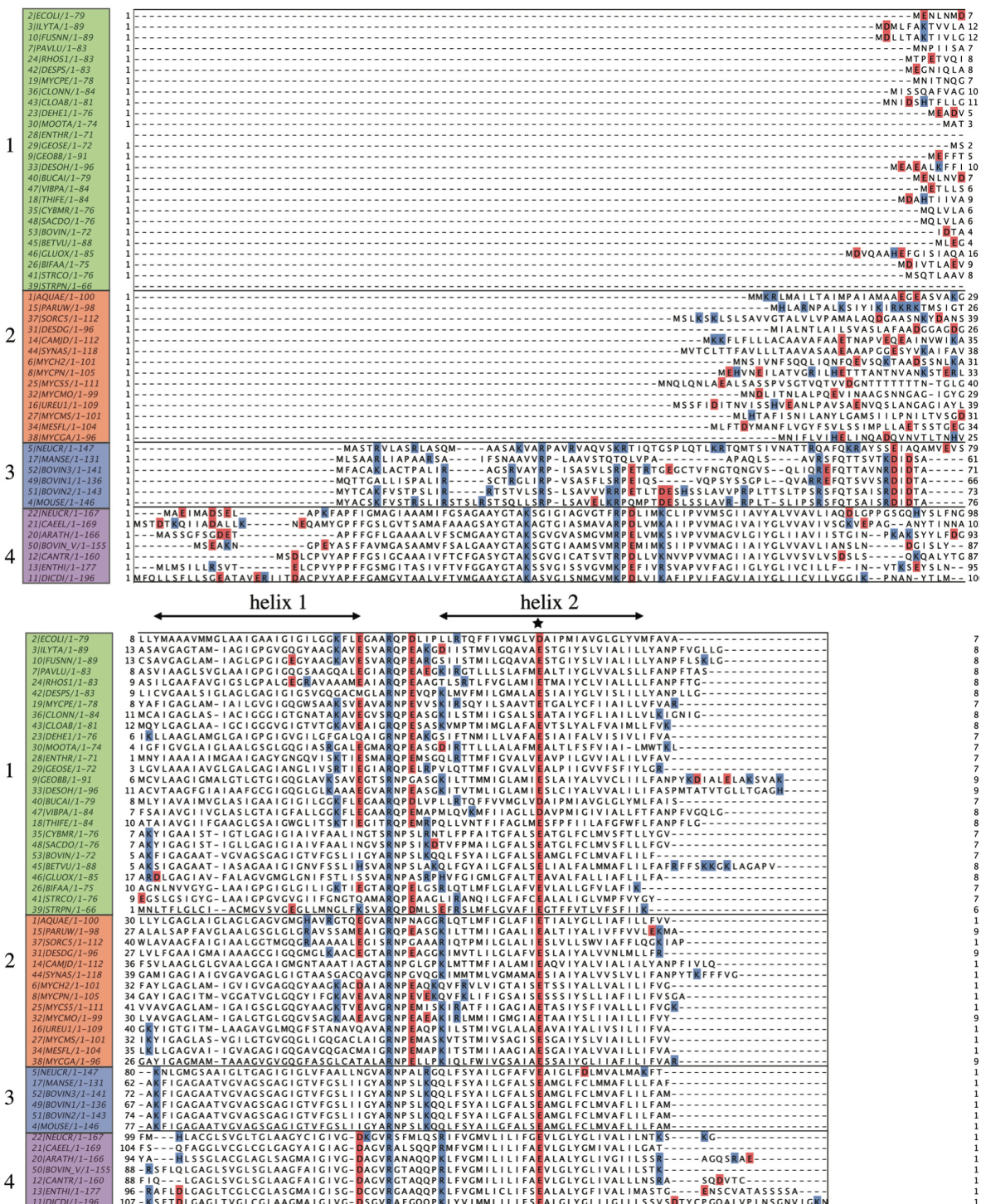


Figure 3.3. Multiple-sequence alignments of the subunits c. 4 groups of subunits c can be classified according to the variability of their N-terminal regions, while their C-terminal regions are highly conserved and possess 2 functional transmembrane helices (black arrows) connected by a positively charged cytoplasmic loop. The star marks the functional Asp/Glu residue in helix-2. Group 1 subunits c are highlighted in green, group 2 in red, group 3 in blue, and group 4 in purple. Positively charged residues (Arg, His and Lys) are shown in blue and negatively charged residues (Glu and Asp) are shown in red, respectively. A detailed list of the subunits c used is reported in Appendix Table A1.

Members of group 1 and of group 3 (in their mature form, after removal of the mitochondrial targeting sequence specific to this group) possess short, predominantly negatively charged N-terminal tails. In contrast, members of group 2 are unique in that they possess a long, and – in most cases – positively charged N-terminal tail (see Figure 3.3).

To investigate the biological role of the N-terminus of subunits c we have (i) studied the long, positively charged N-terminus of a group 2 target (*A. aeolicus*) both in native cells as well as in a heterologous expression system (*E. coli*), and (ii) exchanged the long, positive N-terminus of the same group 2 target by a group 1 short, negatively charged N-terminus. We found that the N-terminus of group 2 subunits c is necessary for their correct membrane insertion both in homologous as well as in recombinant expression systems. Additionally, we found that such an N-terminus acts as a signal peptide not only in V-type subunits c (Denda *et al.*, 1989; Ihara *et al.*, 1997; Yokoyama *et al.*, 2000), but also in F-type subunits c, which was not known to date. Finally, we found that the long N-terminus of the group 2 subunits c cannot be replaced by the short N-terminus of group 1 subunits c, because such a replacement completely abolishes membrane insertion. Therefore, we suggest that the biological role of the N-terminus of group 1 subunits c is different from that of the N-terminus of group 2 subunits c. In particular, based on our results, we propose the following role for the N-terminus of each group (see Figure 3.5).

In group 1, the N-terminus does not play any role in membrane insertion. As proposed earlier, the prokaryotic subunits c of this group insert into the membrane through the YidC pathway, which is independent of the N-terminal peptide of membrane proteins, and they do not require the help of the SRP. Similarly, group 1 yeast subunits c – which are encoded in the mitochondrion – are inserted into the membrane by Oxa1, and do not require SRP either, which correlates well with the fact that there are no SRP homologs, nor homologs of its receptor or of any other Sec-components in yeast mitochondria (Glick and Von Heijne, 1996). Possibly, the same considerations might be also valid for group 4, however, the complication of the N-terminus of these subunits and the lack of experimental evidence prevents formulating clear hypotheses.

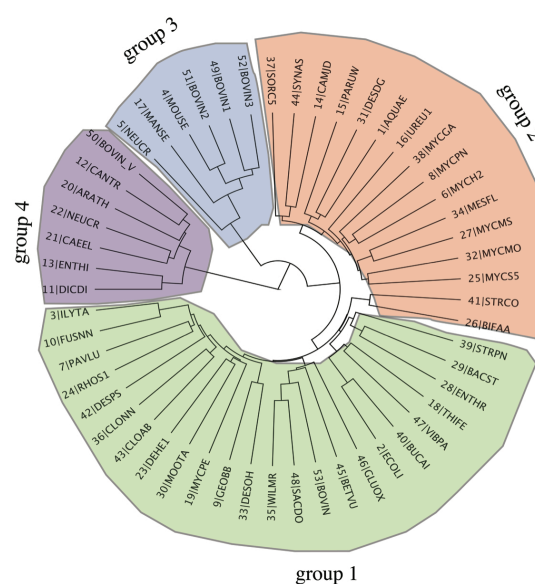


Figure 3.4. Phylogenetic clusters of subunit c. The groups of subunits c identified based on the N-terminal variability cluster in different branches of the phylogenetic tree. Group 1 are highlighted in green, group 2 in red, group 3 in blue, and group 4 in purple. The tree was derived from the multiple-sequence alignment reported in Figure 2.35 with the software Geneious.

In group 3, which includes nuclear-encoded subunit c precursors, the N-terminus functions as a targeting sequence to guide the subunits c to their target organelle, the mitochondrion. Mitochondrial targeting occurs in two independent steps. First, protein precursors are imported into the mitochondrion by the Tim17-23 machinery (Bauer *et al.*, 2000). Successively, their N-terminal targeting peptides are proteolytically removed by the mitochondrial processing peptidase (MPP) and the mature protein is inserted into the inner mitochondrial membrane by the YidC homolog Oxa 1 (Stuart and Neupert, 1996;

Stuart, 2002). The latter step is coupled to the translocation of the mature, short N-terminal tail across the membrane, which occurs in a similar manner as in *E. coli* (Rojo *et al.*, 1995). Therefore, also for group 3 subunits c the mature N-terminal tail does not play a direct role in membrane insertion.

Finally and uniquely among subunits c, the N-terminus of group 2 subunits c does play a role in membrane insertion, acting as an essential N-terminal signal peptide. In our target organism, *A. aeolicus*, this signal peptide possesses the following characteristics typical of SRP-recognition and SRP-mediated membrane insertion (von Heijne, 1985). First, it possesses a positively charged n-region, a hydrophobic central h-region and a neutral, polar c-region. The n-region includes two positively charged amino acids, lysine K3 and arginine R4 (see Figure 2.34B). Second, it follows the “-3, -1 rule”, possessing an alanine residue at both -3 and -1 positions relative to the signal peptidase cleavage site, which is important for recognition and processing by signal peptidase I (Perlman and Halvorson, 1983; von Heijne, 1983; von Heijne, 1986). Third, a proline occupies position -6 relative to the cleavage site. This proline also facilitates the formation of the cleavage site, by breaking the α -helical structure of the peptide inducing the formation of a β -turn (Barkocygallagher *et al.*, 1994). Deleting this signal peptide or substituting it with the N-terminus of subunits c from other groups completely abolishes membrane insertion.

We thus suggest that group 2 subunits c follow a membrane insertion pathway different from the other subunits c and possibly consisting of two steps. First, group 2 subunits c are likely to be recognized by SRP. Successively, they may be inserted into the membranes by either one of three pathways: 1) by YidC only. In this case, YidC recognition would be mediated by the three

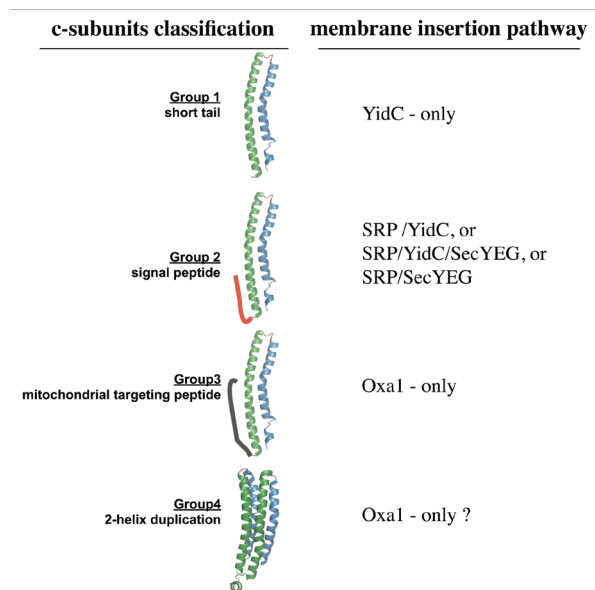


Figure 3.5. Membrane insertion pathways proposed for the 4 different groups of subunit c.

positively charged cytoplasmic loop residues, and membrane insertion would follow a mixed SRP/YidC pathway, whose existence was suggested earlier for an artificial integral membrane protein (IMP) construct (Froderberg *et al.*, 2003); 2) by YidC and SecYEG. In this case, membrane insertion would probably follow a mixed SPR/Sec/YidC pathway that was interestingly also proposed for the ATP synthase subunits a and b (Yi *et al.*, 2004); 3) by SecYEG only. In this case, SecYEG recognition would likely be mediated by a classical SRP/SecYEG interaction.

3.5.2. A revised assembly pathway of F₁F_O ATP synthase

The considerations discussed above about the N-terminal features of subunit c and the results of our expression trials with the two subunit-c deletion mutants have implications on the biogenesis and assembly of the F₁F_O ATP synthase complex. Such processes have been studied in *E. coli* and yeast and showed a similar stepwise and coordinated assembly mechanism (see chapter 1.2.5). Assembly starts with the formation of the subcomplex F₁, and it is then followed by a sequential assembly of the membrane F_O subcomplex, and finally, by the coupling of the subcomplexes F_O to F₁ (see Figure 1.13 and (Price and Driessen, 2010; Dalbey *et al.*, 2011)). However, in our expression system, we noticed a different behaviour for the *A. aeolicus* ATP synthase.

In our studies on the two subunit-c-deletion mutants, the assembled F₁ complex localized at the *E. coli* membranes despite the fact that these mutants failed to express subunit c (see chapter 2.3.5 and Figure 2.35). Subunit a was not detected in these mutants either, which is expected given that the membrane insertion of subunit a requires the presence of both subunits b and c (Hermolin and Fillingame, 1995) and that uncomplexed subunit a is readily degraded by FtsH (Akiyama *et al.*, 1996). Instead, subunits b₁ and b₂ were correctly expressed (Figure 3.6). Therefore, our results suggest that the intact F_O subcomplex is not required for the heterologously produced F₁ subcomplex to associate with *E. coli* membranes. Rather, it seems that the presence of subunits b₁ and b₂ (see Figure 3.6) is sufficient to mediate the recruitment of F₁ to the membranes. Such recruitment could be mediated by the hydrophilic part of subunit b, which in *E. coli* was indeed suggested to be essential for the association of subcomplex F₁ with F_O (Price and Driessen, 2010). Therefore, our results suggest that the assembly of subcomplex F₁ to subunits of F_O might happen earlier than previously proposed (Price and Driessen, 2010; Dalbey *et al.*, 2011) and may follow the mechanism proposed in Figure 3.7.

MVRLISFLTL	ASTFAYAGEG	HLGHSPGALI
WKGLNILAFL	GIVYYFGKKP	ISEAFNKFYN
SIVESLVNAE	REFMMAREEL	SKAKEELENA
KKKAQEYEKL	AIETAETEK	KILQHAQEV
ERIKEKAKET	IEIELNKAKK	ELALYGIQKA
EETAKDLLQK	EFKSKVQEK	YIEAQLKLE
ERKNA		

Figure 3.6. Subunit b₂ was identified by PMF in the subunit c-deletion mutant pCL21-MEN. Residues in red were identified by PMF followed by ESI-MS.

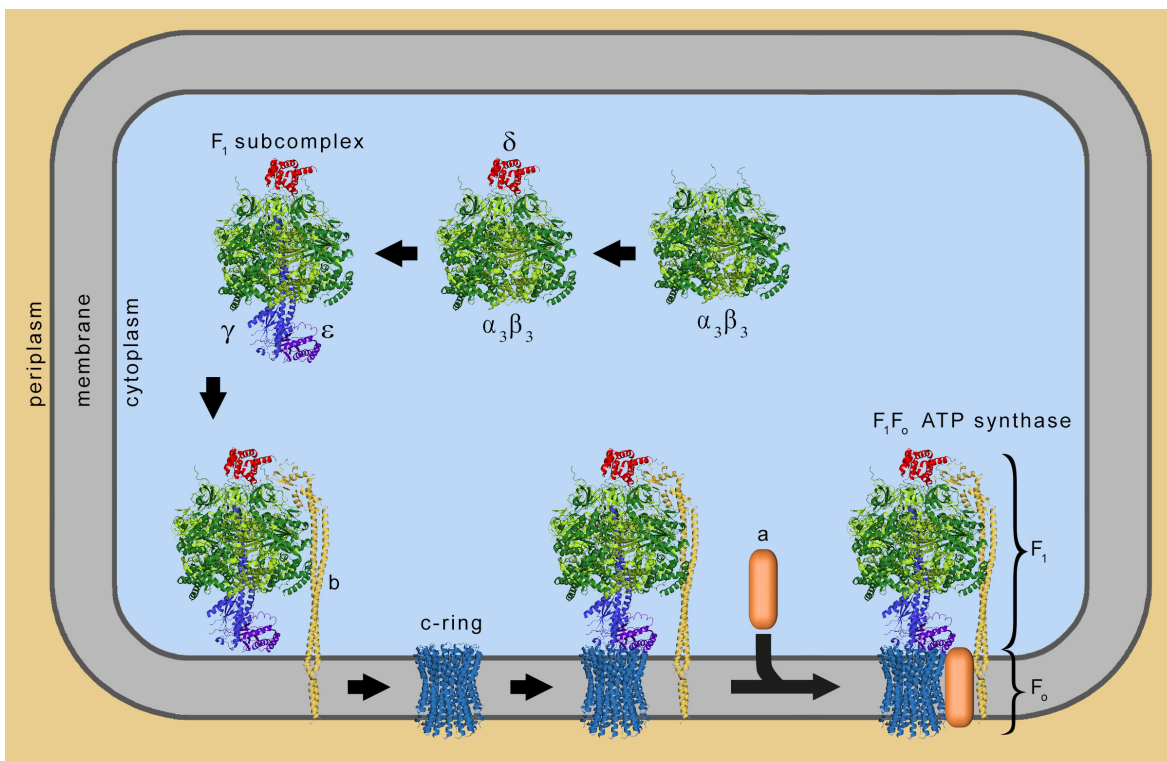


Figure 3.7. Proposed mechanism for the assembly of subcomplex F_1 to subunit b in the *E. coli* membranes. The assembly of F_1 occurs in the cytoplasm where the subunits α , β , γ , δ , and ϵ form a globular structure (top). The assembly of subunit b into *E. coli* membrane is c-ring independent (down). Assembly of the subcomplex F_1 to the *E. coli* via subunit b does not require the presence of the c-ring. Without the c-ring, subunit a cannot assemble into the subcomplex. The figures is modified from Figure 1.13 to reflect the different assembly pathway proposed for *A. aeolicus* ATP synthase. The figure was generated by Paolo Lastrico (MPI of Biophysics, Frankfurt, Germany).

3.6. Conclusions and future perspectives

In conclusion, this work has achieved (i) a deeper characterization of native *A. aeolicus* F_1F_0 ATP synthase by bioinformatic, biochemical and enzymatic analyses, (ii) the creation of a heterologous expression system to produce the enzyme in an active and fully assembled state in the heterologous host *E. coli* using an artificial operon, (iii) the discovery of new features of ATP synthase using such heterologous expression system, and (iv) the development of new hypotheses to describe the different membrane insertion and assembly mechanisms of ATP synthase.

Functional and enzymatic characterization demonstrated that *A. aeolicus* ATP synthase is H^+ -dependent, and not Na^+ -dependent. Its ATP hydrolysis mechanism needs to be triggered and activated by high temperatures, possibly inducing a conformational switch in subunit γ . Finally, unusual features were identified for membrane subunits a , b and c with important implications for the membrane insertion and assembly mechanism of ATP synthase.

Most importantly, the attempts to construct an heterologous expression system to produce the entire *A. aeolicus* F_1F_0 ATP synthase in the mesophilic host *E. coli* were successful. The enzymatic

assays and single-particle electron microscopy showed that the heterologously-produced ATP synthase is fully active and possesses the same structure as the one extracted from native *A. aeolicus* cells. Therefore, the project opens up new and exciting directions for future research on *A. aeolicus* ATP synthase. Such system has already enabled us to propose a new phylogenetic insight into subunit c and certainly many more experiments are now feasible. Genetic manipulation allows now for relatively straightforward functional studies, both *in vivo* and *in vitro*. For instance, by cross-linking and/or site-directed mutagenesis, one could study the binding mode of subunits b₁ and b₂ in the unusual heterodimeric peripheral stalk, the interface of subunit γ and subunits α/β in the F₁ subcomplex, or the interface of subunit a and c, which regulate ion translocation in the F_o subcomplex and which is one of the least understood regions of ATP synthases. Moreover, as we demonstrated for subunit c, also the role of the putative signal peptide of subunit b₂ should be investigated, i.e. by deletion and replacements of such a peptide. In parallel, more experimental evidence should be gathered to support our current hypothesis that the assembly pathway of *A. aeolicus* F₁F_o ATP synthase may be unique and to determine whether the phylogenetic group 2 ATP synthases identified in this work possess other unique functional properties, besides the signal peptide in subunit c. Additionally, site-directed mutagenesis, crosslinking, tags and/or specific antibodies should be used to study the interaction between the peripheral stalk and subunit δ , which was previously suggested to present unusual features (Peng *et al.*, 2006). Furthermore, many of the subcomplexes produced in this work are ready to be subjected to crystallization trials. For instance, the structure of subcomplex b₁b₂ would be very valuable because it would represent the first structure of a heterodimeric peripheral stalk of bacterial F₁F_o ATP synthase and it could thus be compared to the available structure of the peripheral stalk of bovine F₁F_o ATP synthase (Dickson *et al.*, 2006; Rees *et al.*, 2009) and V₁V_o ATP synthase from *T. thermophilus* (Stewart *et al.*, 2012). Additionally, such a structure could potentially also provide interesting insights into the mode of interaction of b₁b₂ with other elements of ATP synthase, i.e. subunit δ in the F₁ subcomplex or subunit c in the F_o subcomplex. Besides subcomplex b₁b₂, the subcomplex F₁ is ready for structural investigation. Obtaining its structure would be very valuable to understand the mechanism of temperature activation of *A. aeolicus* F₁F_o ATP synthase and specifically the relevance of the bent conformation of subunit γ . For all structural studies, the heterologous expression system that was developed in this work would be very helpful not only for producing the targets, but also for introducing artificial amino acid modifications, i.e. mutations at surface residues, which are often required for improving the formation of crystal contacts and thus the resolution of diffraction as well as selenomethionine expression for facilitating the determination of the crystallographic phases (i.e. by multiple-wavelength anomalous dispersion, MAD).

Finally, to a broader extent, the heterologous expression system described in this work will also serve in the future as a solid reference for designing strategies aimed at producing large multi-subunit complexes with complicated stoichiometry, i.e. other respiratory complexes, the nuclear pore complex, transporter systems and many other macromolecular machines that are all very active research targets nowadays.

4. Material and Methods

4.1. Material

4.1.1. Chemicals

All chemicals and enzymes were obtained in highest purity from Carl Roth GmbH (Karlsruhe Germany), Sigma-Aldrich (Taufkirchen Germany), New England Biolabs (Ipswich, USA), Thermo Scientific (Bonn, Germany), and Invitrogen (Carlsbad, USA) unless stated otherwise. Primers were obtained from Eurofins MWG operon (Ebersberg, Germany) and Sigma-Aldrich (Taufkirchen, Germany). DNA Sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany) and SeqLab (Göttingen, Germany).

4.1.2. Organisms

4.1.2.1. *Aquifex aeolicus*

A. aeolicus VF5 cells were obtained from the Archaeenzentrum (Regensburg, Germany) and stored at -20 °C before use.

4.1.2.2. *Escherichia coli* strains

All *E. coli* strains used in this work are listed in Table 4.1.

Table 4.1: List of *E. coli* strains.

Strains	Genotype	Reference/company
DK8	<i>bglR thi rel1 HfrPO1 1100A(uncB-uncC)ilv::Tn10(Tet^R)</i>	(Klionsky <i>et al.</i> , 1984)
DH5 α	<i>F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ-</i>	(Hanahan, 1985)
BL21 (DE3)	<i>F- ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])</i>	(Studier and Moffatt, 1986)
C43 (DE3)	<i>F- ompT gal dcm hsdSB(rB- mB-)(DE3)</i>	(Miroux and Walker, 1996)
NM554	<i>recA13 araD139 Δ(ara-leu)7696 Δ(lac)l7A galU galK hsdR rpsL (Strr) mcrA mcrB</i>	(Ludtke <i>et al.</i> , 1999)
TOP10	<i>F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-</i>	Invitrogen
GM2163	<i>ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10)TetS endA1 rpsL136 dam13::Tn9 xylA-5 mtl-1 thi-1 mcrB1 hsdR2</i>	(Woodcock <i>et al.</i> , 1989)

4.1.3. Plasmids

All empty vectors used in this work are listed in Table 4.2.

Table 4.2: List of empty vectors.

Vectors	Features	Purpose	Reference
pTTQ18A2 ^a	Tac promoter ^c , pMB1, Amp ^R , C-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pTTQ18C3 ^a	Tac promoter ^c , pMB1, Amp ^R , N-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pQEA2 ^b	T5lac2 ^f promoter, pColE1, Amp ^R , C-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pQEC3 ^b	T5lac2 ^f promoter, pColE1, Amp ^R , N-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pETA2 ^c	T7lac ^g promoter, pBR322, Kan ^R , C-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pETC3 ^c	T7lac ^g promoter, pBR322, Kan ^R , N-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pBADA2 ^d	<i>araBAD</i> promoter, pBR322, Amp ^R , C-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pBAD3 ^d	<i>araBAD</i> promoter, pBR322, Amp ^R , N-terminal His ₁₀ -tag	Single gene expression	(Surade <i>et al.</i> , 2006)
pBAD-CM1 ^d	<i>araBAD</i> promoter, pBR322, Amp ^R , C-terminal His ₆ -tag and C-terminal strepII-tag	Dual gene expression	(Surade <i>et al.</i> , 2006)
pET22b(+)	T7lac ^g promoter, pBR322, Amp ^R , C-terminal His ₆ -tag	Single gene expression	Novagen
pETDeut	T7lac ^g promoter, pBR322, Amp ^R , N-terminal His ₆ -tag or C-terminal S-tag	Dual gene expression	Novagen
pTrc99A	pTrc ^e promoter, pBR322, Amp ^R	Operon expression	Amashem
pJET1.2	T7 promoter, pMB1, Amp ^R	Cloning	Clontech
pIVEX2.3d	T7 promoter, Amp ^R	<i>In vitro</i> expression	Roche
pRARE	Cm ^R , p15A	Codon usage optimization	Novagen

- Derived from pTTQ18, regulated and repressed by the lac repressor protein (lacI^q), induced by IPTG. Any *E. coli* host strain can be used because lacI^q allele is included in the plasmid. Basal expression.
- Derived from pQE80L (Qiagen). Regulated and repressed by the lac repressor protein, induced by IPTG. Any *E. coli* host strain can be used because lacI^q is included in the plasmid.
- Derived from pET26(+) (Novagen). Host for expression need to contain DE3 chromosomal copy of the gene for T7 RNA polymerase lysogens of bacteriophage (LacUV5 promoter).
- Derived from pBAD/His (invitrogen).
- Hybrid trp-lac promoter (*tac*) including tryptophan promoter and lac operator sequences.
- Hybrid promoter including Phage T5 promoter (recognized by *E. coli* polymerase) and two lac operator sequences.
- Hybrid T7-lac promoter including phage T7 promoter and lac operator sequences.

All vectors used in this work for single and dual gene expression are listed in Table 4.3.

Table 4.3: List of vectors generated for the single or dual gene expression of subunits a, b₁b₂ and c.

Name	Parental vectors	Cloned genes	Subunits
a-pTTQ-A	pTTQ18A2	<i>atpB</i>	a
a-pTTQ-C	pTTQ18C3	<i>atpB</i>	a
a-pQE-A	pQEA2	<i>atpB</i>	a
a-pQE-C	pQEC3	<i>atpB</i>	a
a-pBAD-A	pBADA2	<i>atpB</i>	a
a-pBAD-C	pBADC3	<i>atpB</i>	a
bb2-pTTQ-A	pTTQ18A2	<i>atpF1F2</i>	b ₁ b ₂
bb2-pTTQ-C	pTTQ18C3	<i>atpF1F2</i>	b ₁ b ₂
bb2-pQE-A	pQEA2	<i>atpF1F2</i>	b ₁ b ₂
bb2-pQE-C	pQEC3	<i>atpF1F2</i>	b ₁ b ₂
bb2-pBAD-A	pBADA2	<i>atpF1F2</i>	b ₁ b ₂
bb2-pBAD-C	pBADC3	<i>atpF1F2</i>	b ₁ b ₂
c-pTTQ-A	pTTQ18A2	<i>atpE</i>	c
c-pTTQ-C	pTTQ18C3	<i>atpE</i>	c
c-pQE-A	pQEA2	<i>atpE</i>	c
c-pQE-C	pQEC3	<i>atpE</i>	c
c-pBAD-A	pBADA2	<i>atpE</i>	c
c-pBAD-C	pBADC3	<i>atpE</i>	c
c-pET-G	pET-G	<i>atpE</i>	c
c-pQE-A-MBP	pQEA2	<i>atpE</i>	c+MBP
a-bb2-pBCM	pBAD-CM1	<i>atpB-atpE</i>	a, b ₁ b ₂
a-c-pBCM	pBAD-CM1	<i>atpB-atpF1F2</i>	a, c

All plasmids used for single artificial *atp* operon expression are listed in Table 4.4.

Table 4.4: List of vectors generated for the expression of the artificial *atp* operon.

Name	Parent vector	Cloned genes	Resistance ^a , origin
pJET01(+)	pJET1.2	<i>atpB</i> , <i>aq_178</i> , <i>atpE</i>	Ap ^R , pMB1
pJET01(-)	pJET1.2	<i>atpBE</i>	Ap ^R , pMB1
pJET02	pJET1.2	<i>atpF1F2H</i>	Ap ^R , pMB1
pJET3	pJET1.2	<i>atpA</i>	Ap ^R , pMB1
pJET4	pJET1.2	<i>atpGD</i>	Ap ^R , pMB1
pJET4(+His ₆)	pJET1.2	<i>atpGD</i>	Ap ^R , pMB1
pJET5	pJET1.2	<i>atpC</i>	Ap ^R , pMB1
pCL01	pTrc99A	<i>atpBE</i>	Ap ^R , pBR322
pCL02	pTrc99A	<i>atpBEF1F2H</i>	Ap ^R , pBR322
pCL11	pTrc99A	<i>atpAGD</i>	Ap ^R , pBR322
pCL12	pTrc99A	<i>atpAGDC</i>	Ap ^R , pBR322
pCL21	pTrc99A	<i>atpBEF1F2HAGDC</i>	Ap ^R , pBR322
pCL21(+strepII)	pTrc99A	<i>atpBEF1F2HAGDC</i>	Ap ^R , pBR322
pCL21-ΔSP	pTrc99A	<i>atpBEF1F2HAGDC</i> (Δc2-19)	Ap ^R , pBR322
pCL21-MEN	pTrc99A	<i>atpBEF1F2HAGDC</i> (Δc1-29: MENLNMD)	Ap ^R , pBR322

Ap^R, resistance to ampicillin; Cm^R, resistance to chloramphenicol.

4.1.4. Bacterial media and solutions

Bacterial media were autoclaved at 121°C for 20 min. All bacterial media used in this work are listed in Table 4.5.

Table 4.5: List of bacterial media.

Medium	Composition	Preparation
Luria-Bertani (LB) broth medium	1% (w/v) tryptone 0.5% (w/v) yeast extract 1% (w/v) NaCl H ₂ O Adjust pH to 7.0 with NaOH	10 g 5 g 10 g Add to 1 L
Terrific Broth (TB) medium	1.2% (w/v) tryptone 2.4% (w/v) yeast extract 0.4% (v/v) glycerol 0.17 M KH ₂ PO ₄ 0.72 M K ₂ HPO ₄ H ₂ O	12 g 24 g 4 mL 23.1 g 125.4 g Add to 1 L
2× YT medium	1.6% (w/v) tryptone 1% (w/v) yeast extract	16 g 10 g

	0.5% (w/v) NaCl. H ₂ O Adjust pH to 7.0 with NaOH	5 g Add to 1 L
TB (Auto- induction)	Overnight Express™ Instant TB Medium (Novagen) 1% glycerol H ₂ O	60 g 10 mL Add to 1 L
M9	42.26 mM Na ₂ HPO ₄ 22.05 mM KH ₂ PO ₄ 8.55 mM NaCl 18.7 mM NH ₄ Cl 100 mM CaCl ₂ 2 mM MgSO ₄ 22.2 mM glucose Centrum Vitamin Mix (Whitehall- Much GmbH) H ₂ O	6 g 3 g 0.5 g 1 g 1 M CaCl ₂ / 0.1 mL 1 M MgSO ₄ / 2 mL 40% glucose/ 10 mL 7.5% (w/v) Centrum Vitamin Mix stock/ 2 mL Add to 1 L
SOC medium	2% (w/v) tryptone 0.5% (w/v) yeast extract 10 mM NaCl 2.5 mM KCl 10 mM MgCl ₂ 10 mM MgSO ₄ 20 mM glucose H ₂ O Adjust pH to 7.0 with NaOH	5 g 1.25 g 0.146 g 0.047 g 1 M MgCl ₂ /2.5 mL 1 M MgSO ₄ /2.5 mL 1 M glucose/5 mL Add to 250 mL
Succinate	34 mM KH ₂ PO ₄ 64 mM K ₂ HPO ₄ 20 mM (NH ₄) ₂ SO ₄ 0.4% (w/v) sodium succinate·6H ₂ O 1 μM ZnCl ₂ 10 μM CaCl ₂ 0.3 mM MgSO ₄ ·7H ₂ O 1 μM FeSO ₄ ·7H ₂ O 2 μg/mL thiamine hydrochloride 50 μg/mL isoleucine 50 μg/mL valine 50 μg/mL thymine 50 μg/mL asparagine H ₂ O	4.6 g 14.6 g 2.6 g 4 g 10 mM ZnCl ₂ /100 μL 100 mM CaCl ₂ /100 μL 1M MgSO ₄ ·7H ₂ O/300 μL 10 mM FeSO ₄ ·7H ₂ O/100 μL 10 mg/mL thiamine hydrochloride/200 μL 20 mg/mL isoleucine/2.5 mL 20 mg/mL valine/2.5 mL 20 mg/mL thymine/2.5 mL 20 mg/mL asparagine/2.5 mL Add to 1 L

All antibiotics used in this work are listed in Table 4.6. Antibiotics were sterilized by filtering through 0.2 µm membranes (Sarstedt Aktiengesellschaft & Co.).

Table 4.6: List of antibiotics.

Antibiotic	Stock solution	Working concentration (dilution)
Ampicillin	100 mg/mL in H ₂ O ^a	100 µg/mL (1:1000)
Carbenicillin	100 mg/mL in H ₂ O ^a	100 µg/mL (1:1000)
Chloramphenicol	34 mg/mL in ethanol	170 µg/mL (1:200)
Kanamycin	50 mg/mL in H ₂ O ^a	50 µg/mL (1:1000)
Streptomycin	10 mg/mL in H ₂ O ^a	50 µg/mL (1:200)
Tetracycline	10 mg/mL in ethanol	30 µg/mL (1:300)

- a. These stock solutions were prepared in H₂O/glycerol (1:1) to avoid repeated freezing/thawing cycles when stored at -20 °C.

4.1.5. Enzymes, proteins, markers and kits

All the restriction enzymes and DNA modifying enzymes were purchased from Fermentas and New England Biolabs (NEB). Other enzymes, proteins, markers and kits used in this work are listed in Table 4.7.

Table 4.7: List of enzymes, proteins, markers and kits.

Name	Supplier	Application
Albumin Standard (bovine serum)	Pierce	Protein concentration assay
Albumin fraction V (bovine serum, biotin-free)	Carl Roth	Western blot
Avidin (egg white)	Gerbu	Avidin-biotin interaction
CloneJET™ PCR Cloning Kit	Fermentas	Molecular cloning
DNA Clean & Concentrator™-25 Kit	Zymo Research	DNA purification
Complete protease inhibitor cocktail tablets	Roche	Protease inhibitors
Quick Ligation™ Kit	NEB	Molecular cloning
GeneRuler™ 1 kb Plus DNA Ladder	Fermentas	DNA standard
GeneRuler™ 100 bp Plus DNA Ladder	Fermentas	DNA standard
G-spin™ Genomic DNA Extraction Kit (for bacterial)	iNtRON	Genomic DNA extraction
Lambda DNA-Mono Cut Mix	NEB	DNA standard
Lysozyme (egg white)	Fluka	Protein purification
NativeMark™ Protein Standard	Invitrogen	Protein standard
PageRuler™ Prestained Protein Ladder	Fermentas	Protein standard
PageRuler™ Prestained Protein Ladder Plus	Fermentas	Protein standard
Phusion® High-Fidelity DNA Polymerase	Finnzymes	PCR
QIAGEN® Plasmid <i>Plus</i> Midi Kit	Qiagen	DNA extraction
QIAprep® Spin Miniprep Kit	Qiagen	DNA extraction
QIAquick® Gel Extraction Kit	Qiagen	DNA purification
QIAquick® PCR Purification Kit	Qiagen	DNA purification
QuikChange® Lightning Site-Directed Mutagenesis Kit	Stratagene	Mutagenesis
Ribonuclease A	Roche	DNA extraction
SeeBlue® Plus2 Prestained Standard	Invitrogen	Protein standard
Supercoiled DNA Ladder	NEB	DNA standard
SilverQuest™ Staining Kit	Invitrogen	Protein staining
Zymoclean™ Gel DNA Recovery Kit	Zymo Research	DNA purification

4.1.6. Antibodies

All the antibodies and conjugated proteins used for Western blot analysis in this work are listed in Table 4.8.

Table 4.8: List of antibodies.

Type	Target protein for detections	Antibodies / Conjugated proteins	Company
1 st antibodies	Poly-His-tagged proteins	A monoclonal α -poly-histidine-alkaline phosphatase conjugated antibody	Sigma-Aldrich
	StrepII-tagged proteins	Streptavidin coupled to alkaline phosphatase	Sigma-Aldrich
	Subunit F ₁ - β	KHL-conjugated synthetic peptide polyclonal antibody	Purchased from Agrisera
	Subunit F ₁ - α	<i>AQUEA_AtpA</i>	The custom peptide polyclonal antibodies generated by Thermo Fisher Scientific for this work
	Subunit F ₁ - γ	<i>AQUEA_AtpG</i>	
	Subunit F ₁ - ϵ	<i>AQUEA_AtpC</i>	
	Subunit F ₁ - δ	<i>AQUEA_AtpH</i>	
Subunit F _O -c	<i>AQUEA_AtpL</i>		
2 nd antibody	IgG from rabbit	Monoclonal mouse anti-rabbit IgG conjugated with alkaline phosphatase	Sigma-Aldrich

4.1.7. Chromatographic columns and matrices

All chromatographic columns and matrices used in this work are listed in Table 4.9.

Table 4.9: List of chromatographic columns and matrices.

Material	Separation type	Supplier
Mini Q 4.6/50 PE	Ion exchange	GE Healthcare
Mono Q 5/50 GL	Ion exchange	GE Healthcare
Superdex 200 3.2/30 GL	Size-exclusion	GE Healthcare
Superdex 200 5/150 GL	Size-exclusion	GE Healthcare
Superdex 200 10/300 GL	Size-exclusion	GE Healthcare
Sucrose 6 5/150 GL	Size-exclusion	GE Healthcare
TSK-GEL G4000SW	Size-exclusion	TOSOH Bioscience
HisTrap HP 1 ml	IMAC affinity chromatography	GE Healthcare
Ni-NTA agarose	IMAC affinity chromatography	Qiagen
Strep-Tactin Superflow [®] high capacity column	IMAC affinity chromatography	IBA
Disposable PD-10 desalting column	Size-exclusion	GE Health
C18	Reverse phase	NanoSeparation

4.1.8. Database, servers and software

The databases and servers used in this work are listed in Table 4.10.

Table 4.10: List of database and servers.

Database and server	URL
BLAST	http://blast.ncbi.nlm.nih.gov/Blast.cgi
Compute pI/Mw	http://web.expasy.org/compute_pi/
EMBL-EBI	http://www.ebi.ac.uk/services
ExPASy: SIB Bioinformatics Resource Portal	http://www.expasy.org/
Mascot	http://www.matrixscience.com/
Membrane Protein Data Bank MPDB	http://www.mpdb.tcd.ie/
MEMSAT-SVM	http://bioinf.cs.ucl.ac.uk/psipred/
NCBI: National Center for Biotechnology Information	http://www.ncbi.nlm.nih.gov/
PDB: Protein Data Bank	http://www.rcsb.org/
PSIPRED v3.0	http://bioinf.cs.ucl.ac.uk/psipred/
PubMed	http://www.ncbi.nlm.nih.gov/pubmed
SignalP Server	http://www.cbs.dtu.dk/services/SignalP/
Signal Peptide Database	http://www.signalpeptide.de/
SOSUI WWW Server	http://bp.nuap.nagoya-u.ac.jp/sosui/
TMHMM Server	http://www.cbs.dtu.dk/services/TMHMM/
UniProt: Universal Protein Resource	http://www.uniprot.org/
Web of Knowledge	http://www.webofknowledge.com/

The software used in this work is listed in Table 4.11.

Table 4.11: List of software.

Software	Version	Application	Company or reference
BioTools	3.1 (build 2.22)	MS data analysis	Bruker Daltonics
Chromas Lite	2.01	Chromatogram editor	Technelysium
Clone Manager	Professional 9	Cloning simulation	Sci-Ed Software
CID-HIT	4.6	Sequence clustering	(Fu <i>et al.</i> , 2012)
ClustalX	2.1	Multiple-sequence alignment	(Thompson <i>et al.</i> , 1997)
Compass/Hystar	3.2	MS data analysis	Bruker Daltonics
Endnote	X5	Reference management	Thomson Reuters
Geneious	Basic 5.6.5	Phylogenic tree viewing	Biomatters Ltd
iWork	09 version 4.1	Apple office software	Apple
JalView	2.8	Sequence alignment viewing and editor	(Waterhouse <i>et al.</i> , 2009)
MacVector	11.0.4	Cloning simulation	MacVector
Microsoft Office	2010	Microsoft office software	Microsoft
Origin	8.6.0	Data processing and analysis	OriginLab
Photoshop	CS5 extended 12.0.4	Image editor	Adobe
PyMOL	1.3	3D molecular visualizer	Schrödinger
T-Coffee	9.03	Multiple-sequence alignment	(Keller <i>et al.</i> , 2011)
Unicorn	5.11	Äkta control system	GE Healthcare

4.2. Methods

4.2.1. Bioinformatics

The position of transmembrane helices in the analyzed sequences was predicted using the TMHMM server v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) (Moller *et al.*, 2001). The secondary structure segments were predicted by PSIPRED v3.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>) (Jones, 1999). Signal peptides were predicted by SignalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP/>) (Petersen *et al.*, 2011) and MEMSAT-SVM (<http://bioinf.cs.ucl.ac.uk/psipred/>) (Nugent and Jones, 2009).

A set of 218 homologous sequences of the *A. aeolicus* subunit c were obtained after 3 iterations of a PSI-BLAST search (Position-Specific Iterated Basic Local Alignment Search Tool) (Schaffer *et al.*, 2001) with the SwissProt sequence database (Bairoch and Apweiler, 2000) using an E-value of 0.001. To reduce the redundancy, 48 clusters were identified using CD-HIT (the Cluster Database at High Identity with Tolerance tool) (Fu *et al.*, 2012) with a percentage of identity of 60%. An initial set of 48 sequences representative of the 48 clusters was obtained. 5 additional sequences were successively integrated into the initial set, which correspond to subunit c homologues of known 3-D structures, possessing previously characterized signal peptides. In total, the full set of

53 sequences was selected for multiple-sequence alignment. Multiple-sequence alignment was performed with the program T-Coffee version 9.03 (Keller *et al.*, 2011) and was manually adjusted. The final score of the alignment is 77, confidently above 40, the cut-off suggested by the developers of T-Coffee to be the minimum for a reliable result. The alignment was visualized and edited in Jalview v11.0 (Waterhouse *et al.*, 2009). The phylogenetic tree was visualized in Geneious basic 5.6.5 (<http://www.geneious.com>).

4.2.2. Molecular Biology

All molecular biology procedures followed the standard protocols recommended by the manufacturers, unless stated otherwise. In particular, detailed protocols are listed hereafter.

4.2.2.1. Isolation of genomic DNA from *A. aeolicus*

A. aeolicus genomic DNA was extracted using the G-spinTM Genomic DNA Extraction Kit for Bacteria (iNtRON Biotechnology), according to the manufacturer's protocol for Gram-negative bacteria. The typical yield of genomic DNA was 2-5 µg per 10 - 20 mg of wet *A. aeolicus* cell pellets. The genomic DNA was stored at 4 °C.

4.2.2.2. Isolation of plasmid DNA

Isolation of plasmid DNA from *E. coli* cells was carried out using suitable kits (Qiagen). Briefly, the Qiagen MiniPrep kit was used for purification of high-copy and low-copy plasmid DNA from 5 mL and 10 mL overnight cultures of *E. coli* grown in LB medium at 37 °C with vigorously shaking at 220 rpm. For large-scale preparations, the Qiagen MidiPrep kit was used for purification of plasmid DNA from 50 mL overnight cultures. The plasmid DNA was stored in 10 mM Tris-HCl pH 8.5 at -20 °C.

4.2.2.3. DNA amplification

Coding sequences for each subunit of AAF₁F_O were obtained from NCBI (NC_000918) and used for primer design. Primers were custom synthesized by Sigma-Aldrich and are listed in Table 4.12. The stock and working concentrations of primers solutions was 100 µM and 10 µM, respectively. DNA fragments were amplified by polymerase chain reaction (PCR) using the Phusion DNA polymerase (Finnzymes) from *A. aeolicus* genomic DNA and using a T-Gradient thermocycler (Biometra). Gradient PCR was used to determine the optimal annealing temperature between 50 °C and 70 °C. Analytical and preparative PCR cycles were performed in a volume of 10 µL and 50 µL, respectively. All reactions were set up on ice and a typical reaction included the components listed

in Table 4.13 and used the program listed in Table 4.14. Phusion DNA polymerase was added last into the reaction to prevent primer degradation caused by 3'->5' exonuclease activity. As a negative control, all components except the DNA template were used.

Table 4.12: List of primers used in this work.

	Primers	Sequence (5'-3')	Restriction sites
single or dual gene	a-F	CGGGATCCTGGAGTACTCGCACGTAGTTTACG	BamHI
	a-R	CGGAATTCGTGTGCTCCTCGTGTGCTACAGC	EcoRI
	bb2-F	CGGGATCCTGATGGACATAGGAGTAATGCCTAATG	BamHI
	bb2-R	CGGGTCTCGAATTCGTAGCATTCTTCTCTCTCCAG	BsaI&EcoRI
	c-F	CGGGATCCTGGTGTGATGAAGAGGTTAATGGCTATCTT	BamHI
	c-R	CGGAATTCGTAACCACGAAGAGCAGTATGAAAG	EcoRI
	bb2_pBCM-F	CGTCTAGAATGGACATAGGAGTAATGCCTAATG	XbaI
	bb2_pBCM-R	CGCGTCGACAGCATTCTTCTCTCTCCAG	SaII
	c_pBCM-F	CGTCTAGAGTGTGATGAAGAGGTTAATGGCTATCTT	XbaI
	c_pBCM-R	CGCGTCGACAACCACGAAGAGCAGTATGAAAG	SaII
artificial operon	P1	GGGGAGGTACCCATATGTAATCTGAGCCAATTGCAAAGAG	KpnI&NdeI
	P2	CGCGCGAGATCTAAGGGCTTAAACCACGAAGAG	BglII
	P3	CGCGCGAGATCTATTGCTATAATTGTTTAGCGGAGG	BglII
	P4	GGGAAAGGATCCCTACGGCCATTAAACACCTC	BamHI
	P5	GGGAAAGGATCCAAACCTTAAAGAAGGTTAGGAGGTAG	BamHI
	P6	GGGAAATCTAGAGGAGGGAGAGTTAGGGAACG	XbaI
	P7	CGCGCGTCTAGATTTAGACATTAGTTTATAATAAGTAGCGTTATG	XbaI
	P8	GAAAAAGTTCGACAGGGGCTTAAACTTAGCCC	SaII
	P9	CGCGCGGTCGACTTGGACTTCTCTGGTATAATTTAGGG	SaII
	P10	GGGAAACTGCAGCCCGGGTTTCCCCGAAAGAAGGG	SamI&PstI
	P17	GCTTATTTAACTATAAGAACTAGCAGTACCTTC ¹	-
	P18	GAAGGTA CTGCTAGTTTCTTATAGTTAAATAAGC ¹	-
	P19	GGAGGTTTATAGATGAGAGGATCGCATCATCATCATCATGGTATGGCGG AAGTGATTAAGGG	-
P20	CCCTTAATCACTTCCGCCATACCATGATGATGATGATGATGATGCGATCCTCTCAT CTATAAACCTCC	-	

¹ 17 bp from 3' end of the gene *atpB* is in red, 17 bp from 5' end of the gene *atpE* is in black.

Table 4.13: Typical PCR reaction mixture.

Component	Volume/50 μ L	Volume/10 μ L	Final concentration
H ₂ O	add to 50 μ L	add to 10 μ L	
5 \times Phusion® HF Buffer	10 μ L	4 μ L	1 \times
2.5 mM dNTPs	4 μ L	1 μ L	200 μ M
primer A (10 μ M)	2.5 μ L	0.5 μ L	0.5 μ M
primer B (10 μ M)	2.5 μ L	0.5 μ L	0.5 μ M
template DNA	\times μ L (250 ng)	\times μ L (50 ng)	250 ng
Phusion® Hot Start DNA Polymerase (2 U/ μ L)	0.5 μ L	0.2 μ L	0.02 U/ μ L

Table 4.14: Typical PCR cycling program.

Cycle step	Temperature	Time	Cycles
Initial denaturation	98°C	30 s	1
Denaturation	98°C	10 s	25-30
Annealing	58-65°C	30 s	
Extension	72°C	15-30 s/1 kb	
Final extension	72°C	10 min	1
	4°C	hold	

The resulting PCR products were analyzed by agarose gel electrophoresis. For downstream applications (i.e subcloning and restriction digestion), the DNA fragments of interest were either purified directly using the QIAquick PCR purification kit (Qiagen) or extracted from the gel using the QIAquick gel extraction (Qiagen) or Zymoclean gel DNA recovery (Zymo Research) kits following to the manufacturer's guidelines.

4.2.2.4. Agarose Gel Electrophoresis of DNA

Agarose gel electrophoresis was typically performed in 1% agarose (NEEO ultra quality, Carl Roth) gels in 0.5 \times TBE to separate DNA fragments of 0.5 kb – 10 kb or in 1 \times TAE to separate supercoiled DNA. DNA samples were prepared by addition of 10 \times loading buffer before the electrophoretic run. The run was performed using PowerPacTM Basic power suppliers (Bio-Rad), a self-made horizontal gel chamber, and 0.5 \times TBE or 1 \times TAE as running buffers, at room temperature. Typically the DNA gel was run at 120 V (6V/cm) for 60-90 min and the progress of the separation was monitored using colored dyes in the loading buffer. 100 bp or 1 kb DNA ladders (NEB) were used to identify the size of the DNA fragments. After electrophoresis, the gel was

stained in 0.5 µg/mL ethidium bromide (Roth) solution for 10-30 min at room temperature and destained in water for 5 - 10 min. The gels were subsequently visualized and photographed using a Bio-Rad gel documentation system with a 302 nm UV transilluminator or visualized under UV light (312 nm, Biometra TI1 transilluminator) to excise the DNA bands of interest. Buffers used for electrophoresis of DNA are listed in Table 4.15.

Table 4.15: List of buffers for agarose gel electrophoresis.

Buffer	Components	Preparations
TAE (50×)	2 M Tris 1 M glacial acetic acid 50 mM EDTA H ₂ O	242 g 57.1 mL 100 mL 0.5 M EDTA (pH8.0) Add to 1 L
TBE (5×)	445 mM Tris 445 mM boric acid 10 mM EDTA H ₂ O	54 g 27.5 g 20 mL 0.5 M EDTA (pH8.0) Add to 1 L
DNA loading buffer (10×)	50 % glycerol 10 mM EDTA 0.25% (w/v) Bromophenol Blue 0.25% (w/v) xylene cyanolFE H ₂ O	2.5 mL glycerol 100 µL 0.5 M EDTA 1.25 mg (<i>dsDNA</i> of 300 bp) 1.25 mg (<i>dsDNA</i> of 4 kb) Add to 5 mL

4.2.2.5. Quantification of nucleic acids

DNA concentration was determined using a NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies), according to the manufacturer's guidelines. Typically, 1 - 2 µL of DNA sample was used for measurement and constant 50 was chosen for calculating *dsDNA* concentration. The purity of DNA samples was estimated by the ratio of absorbance at 260 nm vs 280 nm and 230 nm. Pure DNA generally has a 260/280 absorption ratio of approximately 1.8, and a 260/230 absorption ratio of approximately 2.0-2.2.

4.2.2.6. Digestion of DNA with restriction endonucleases

A typical DNA double digestion was prepared as reported in Table 4.16 using 1.5 mL tubes (Eppendorf) for conventional restriction enzymes and 0.2 mL PCR thin-walled tubes (Thermo Scientific) for the FastDigest restriction enzymes (Thermo Scientific), as described in manufacturer's guidelines. The preparative scale protocol was used if the DNA had to be used in downstream applications (i.e. ligation), while the analytical scale was used for screening purposes.

Table 4.16: Typical double digestion reaction.

	Preparative scale		Analytical scale
	PCR product	Plasmid DNA	Plasmid DNA
Buffer (10×)	1×	1×	1×
DNA	1 µg	1-5 µg	50-200 ng
Restriction enzyme 1	2.5 µL	2.5 µL	1 µL
Restriction enzyme 2	2.5 µL	2.5 µL	1 µL
H ₂ O	Add to 50 µL	Add to 50 µL	Add to 20 µL

The restriction digestion reaction was incubated at 37°C for 1-2 hours (conventional restriction enzymes in Thermomixer[®] compact, Eppendorf) or for 15 min (FastDigest restriction enzymes in TRIO-Thermoblock, Biometra).

For vector preparation, dephosphorylation of the DNA ends was performed using calf intestine (CIAP) or shrimp alkaline phosphatase (SAP) to reduce the background of non-recombinant species due to self-ligation of the vector. The results of restriction digestion were analyzed by gel electrophoresis of DNA. For preparative purposes, the DNA of interest was purified using the DNA gel or QIAquick PCR purification kits and the MinElute reaction cleanup kit (Qiagen). Duplex DNA isolated from bacteriophage lambda (cI857*ind* 1 *Sam* 7) (λ DNA) was used as a control of DNA digestion.

4.2.2.7. Ligation

A typical DNA ligation reaction was prepared as reported in Table 4.17 using 1.5 mL tubes (Eppendorf). The reaction mixes were incubated at 16°C overnight or at 25°C for 1-2 h when conventional T4 DNA ligase (Epicentre) was used, and at 25 °C for 5 – 15 min according to the manufacture's guidelines when the Quick Ligation kit (NEB) was used. A molar ratio of 1:3 (vector:insert) was used in typical DNA ligation reactions and a ratio of 1:1:1 (vector: insert-1: insert-2) was used in 3-way ligation reactions. Heat inactivation was optionally performed when the conventional T4 DNA ligase was used but was avoided when the Quick ligation protocol was used because it dramatically reduced transformation efficiency.

Table 4.17: Typical ligation reaction.

	Conventional T4 ligase	Quick ligation Kit
T4 DNA ligase Buffer (10×) / Quick ligation Buffer (2×)	2 µL	10 µL
Vector DNA	50 - 200 ng	50 - 200 ng
Insert DNA	50 - 200 ng	50 - 200 ng
T4 DNA ligase / Quick T4 DNA ligase	1 µL	1 µL
H ₂ O	Add to 20 µL	Add to 20 µL

4.2.2.8. Preparation of chemically competent cells

Chemically competent cells were prepared according to one of the following two methods:

- a) **The calcium chloride method.** *E. coli* DH5 competent cells were prepared as described (Inoue *et al.*, 1990) but with minor modifications. A 1-2 mL of LB medium was inoculated with a single colony and cells were grown at 37°C overnight as a pre-culture. 1 mL of pre-culture was added into 100 mL of LB medium (supplemented with 10 mM MgCl₂) and shaken (100 rpm) at 18 °C to an OD₆₀₀ of 0.1 (approximately 1.5 h – 3 h). The culture was then chilled on ice for 10 min, cells were pelleted at 4,400 × *g* for 10 min at 4 °C (Sigma 4K15 centrifuge). Cell pellets were gently resuspended in 33 mL of cold TB buffer (1/3 of culture volume) for washing. Afterwards, the cells were re-pelleted as described above (4 °C, 4,400 × *g* for 10 min) and gently resuspended in 8 mL of cold TB buffer. After adding 600 µL DMSO as a cryo-protecting agent, 250 µL aliquots of the resulting competent cells were transferred into sterile and pre-chilled microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80°C. This method was the more efficient in preparing highly competent cells.
- b) **The rubidium chloride method.** 5 mL of LB medium was inoculated with a single colony and cells were grown at 37°C overnight as a pre-culture. 2.5 mL of pre-culture was added into 250 mL of SOC medium in a 1 L baffled flask and shaken (180 rpm) at 37°C to an OD₆₀₀ of 0.4 – 0.6. The culture was then chilled on ice, cells were pelleted at 4,400 × *g* for 15 min at 4 °C (Sigma 4K15 centrifuge, 6 × 40 mL culture in 50 mL tubes). Cell pellets were gently resuspended in 96 mL of cold RF1 buffer and incubated on ice for 20 min. The cells were then re-pelleted at 4 °C, 1,600 × *g* for 5 min, gently resuspended in 9.6 mL of cold RF2 buffer and incubated on ice for 5 min. 250 µL aliquots of the resulting competent cells were transferred into sterile and pre-chilled microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80°C.

All buffers and solutions used for competent cell preparation are listed in Table 4.18.

Table 4.18: List of buffers used to prepare chemically competent cells.

Transformation Buffer	Working solution	Stock solution	Preparation
TB buffer	15 mM CaCl ₂ 55 mM MnCl ₂ 250 mM KCl 10 mM PIPES-KOH pH 6.7 H ₂ O	1 M CaCl ₂ 1 M MnCl ₂ 2 M KCl 200 mM PIPES	1.5 mL 5.5 mL 12.5 mL 5 mL Add to 100 mL
RF1	30 mM KOAc, pH 5.8 100 mM RbCl 50 mM MnCl ₂ 10 mM CaCl ₂ 15 % glycerol (v/v) H ₂ O	1 M KOAc 1 M MnCl ₂ 1 M CaCl ₂	15 mL 6.046 g 25 mL 5 mL 75 mL Add to 500 mL
RF2	100 mM MOPS/KOH, pH 6.8 10 mM RbCl 75 mM CaCl ₂ 15 % glycerol (v/v) H ₂ O	1 M MOPS, pH6.8 1 M RbCl 1 M CaCl ₂	0.5 mL 0.5 mL 3.75 mL 7.5 mL Add to 50 mL

4.2.2.9. Transformation of competent cells

Transformation of chemically competent cells was performed by heat shock. An aliquot of frozen competent cells were thawed on ice. 1 - 10 ng plasmid DNA or an aliquot of the ligation mixture (10 µL or less) was mixed with 100 µL competent cells in a cold sterile 1.5 mL microcentrifuge tube and incubated for 5 – 10 min on ice. The cells were then subjected to a heat shock at 42°C for 90 sec without shaking. Afterwards, the cells were transferred and incubated for 2 – 10 min on ice before adding 900 µL SOC medium. The cells were outgrown at 37°C for 1 h. An aliquot (10 – 50 µL for plasmid DNA and 50 µL – 1000 µL for ligation mix) was plated on LB agar plates containing the appropriate antibiotics. The plates were incubated at 37°C overnight.

4.2.2.10. Screening of positive transformants

For subcloning, typically 2 – 5 single colonies were screened for positive transformants by performing Minipreps (see 4.2.2.2) and restriction digestion (see 4.2.2.6). As an alternative, the Rusconi prep was also used for identifying for positive clones (especially in the case of 3- way ligation reactions). The Rusconi prep is a rapid method for screening positive transformants based on the size of resulting plasmids. Cells from a single colony grown on solid agar or in liquid cultures were transferred into a microcentrifuge tube and mixed vigorously by vortexing with 20 µl

of Rusconi mix (25 mM Tris/HCl, pH 7.5, 25 mM EDTA, 0.5 mg/mL lysozyme, 0.1 mg/ mL RNase, 10% (v/v) glycerol, and 0.025% (w/v) bromophenol Blue). The mixture was then incubated for 30 min at room temperature. Total DNA was extracted by adding 10 μ l of phenol/chloroform (1:1) mixture and mixing vigorously by vortexing. The cell debris was removed by centrifugation at $16,000 \times g$, at room temperature for 5 min (Centrifuge 5415D, Eppendorf). The supernatant containing genomic DNA and plasmid DNA was analyzed directly on a DNA agarose gel. A supercoiled DNA ladder (NEB) was used to estimate the size of the supercoiled plasmid DNA.

The vector sequence of all positive transformants was further confirmed by DNA sequencing (see 4.2.2.11). For the identification of clones expressing the target protein, the colony blot procedure was used (see 4.2.4.10).

4.2.2.11. DNA sequencing

DNA sequencing was performed by SeqLab and Eurofin MWG Operon. Typically 600 – 1000 ng plasmid DNAs are required for one reaction. Sequencing primers used for the whole artificial *atp* operon are listed in Table 4.19.

Table 4.19: List of primers used for sequencing the artificial *atp* operon

Primers	Sequence (5' – 3')
SP1_4414	GGCTGTGCAGGTCGTAATC
SP2_5096	CCCTCTCCCTCAGGTTATTC
SP3_5696	TTAGCGGAGGAGAAGAATGG
SP4_6424	GGCACAGGAATACGAGAAAC
SP5_7133	GGAGGACCCTTCCCTTATAG
SP6_7814	AGACCACTGTTGCGATAGAC
SP7_8556	AGAACCCTACAACCCGATAC
SP8_9337	AAACGGAGAAACCGATAGGG
SP9_10078	TTGCGGGACAACCCATAGAC
SP10_10738	GGTCCATATTCGCGCACCTC
SP11_112	TTGAAGGAGAGGTGGGAATC

4.2.2.12. Site-directed mutagenesis

Site-directed mutagenesis was performed by the following two methods:

- a) **Quikchange methods.** The QuikChange®lightning Site-Directed Mutagenesis kit or The QuikChange®II Site-Directed Mutagenesis kit (Agilent Technologies) were used for site-

directed mutagenesis i.e inserting and/or deleting multiple amino acids, following the manufacturer's specifications. Site-directed mutagenesis was used for generating pCL21-MEN using pCL21 as the template and primers 5'-GTAAATAAGCTTTAAGGAGGTAGATGGAAACTGAATATGGATCTTCTGTACCTTGAGCAGGACTT-3' and its corresponding reverse complement as the mutagenic primers.

- b) **PCR site-directed mutagenesis.** PCR site-directed mutagenesis was used for the pCL21-ΔSP construct, using pCL21 as the template with 5'-phosphorylated primers 5'-GCGGAAGGAGAGGCTTCC-3' (forward) and 5'-CACCTACCTCCTTAAAGC-3' (reverse). PCR products were purified using the QIAquick PCR purification kit (Qiagen), and linearized DNA was self-ligated using Quick ligation (NEB), and transformed into host cells using standard transformation protocol (see 4.2.2.9).

After site-directed mutagenesis, all constructs were confirmed by DNA sequencing (see 4.2.2.11).

4.2.2.13. Storage of *E. coli* strains

E. coli strains were stored at -80°C in 15% glycerol (v/v). Typically, 150 μL of pre-sterilized glycerol (100%) were mixed vigorously by vortexing with 850 μL of a logarithmic-phase *E. coli* culture suspension in LB medium. The glycerol stocks were then flash frozen in liquid nitrogen and stored at -80°C.

4.2.3. Protein expression and isolation

4.2.3.1. Rapid expression screening

Culture medium was supplemented with the appropriate antibiotics in Erlenmeyer flasks, (at a culture/flask volume ratio of 1/5) and pre-warmed. The pre-warmed medium was then inoculated with an aliquot of overnight pre-culture (at a pre-culture /culture volume ratio of 1/20), and incubated at 37°C with vigorous shaking at 180 rpm until an OD₆₀₀ of 0.6 was reached. 1 mL of cell suspension was harvested before induction for use as a control and was pelleted by centrifugation at 16,000 × g, 4°C for 5 min (Centrifuge 5415D, Eppendorf) and stored at -20°C. The cells were then induced with 1 mM IPTG (for vectors pTTQ18, pQE, pET, pCL) or 0.2% (w/v) arabinose (for vector pBAD) and incubated for different times (1, 2, 3, 4, 5, 6 and 16 h) at various temperatures (18, 25, 30 and 37°C), in different media (see 4.1.4). 1 mL of cells were harvested by centrifugation at each time point as described above and stored at -20°C. Protein expression was first analyzed in whole cell lysate under denaturing condition. To prepare the whole

cell lysates, the corresponding cell pellets were resuspended in 100 μ L of 20% (w/v) SDS in water for 15 min at room temperature. The suspension was centrifuged at $16,000 \times g$, at room temperature for 20 min (Centrifuge 5415D, Eppendorf) to remove cell debris and genomic DNA. The supernatant was analyzed by SDS-PAGE, Western blot and dot blot analysis (see 4.2.2.8 and 4.2.2.9).

The rapid expression screening procedure was also used for isolating membrane fraction in a small scale (see 4.2.3.4.1) to determine the solubility and localization of target proteins expressed in *E. coli* (see 4.2.3.3).

4.2.3.2. Preparative protein purification

1 – 2 L of appropriate culture medium was supplemented with the appropriate antibiotics in 5 L baffled flasks. The medium was then inoculated with 50 mL overnight pre-culture (at a pre-culture/culture volume ratio of 1/40), and incubated at 37°C with vigorous shaking at 150 rpm until an appropriate OD₆₀₀. The cells were then induced by adding IPTG depending on the protein complexes produced and incubating for a specific time and at specific temperatures (see 4.2.3.6 for details of producing subcomplexes b₁b₂, F₁- $\alpha\beta\gamma$, F₁- $\alpha\beta\gamma\epsilon$ and the whole EAF₁F_O in *E. coli*). After induction, the cells were harvested by centrifugation at 4 °C, at $10,540 \times g$ for 30 min (Centrifuge Avanti J-26XP, Rotor JLA-8.1000, Beckman Coulter). Finally, the cell pellets were flash frozen in liquid nitrogen and stored at -80°C.

4.2.3.3. Determination of protein cellular localization

To determine the protein cellular localization (membrane-inserted, cytoplasm, or inclusion bodies) cells were disrupted by mechanical methods using either tiny glass beads, 200 μ g/mL lysozyme, a French Press or a microfluidizer depending on the preparation scale (see 4.2.3.4). The insoluble cytoplasmic fraction (cell debris and inclusion bodies) was collected by centrifugation at 4°C, at $23,000 \times g$ for 30 min (Rotor GSA/SS34, Sorvall RC5B superspeed centrifuge). The membrane fraction was collected by centrifugation at 4 °C at $200,000 \times g$ for 60 min (Rotor 70 Ti) or at $150,000 \times g$ for 90 min (Rotor 45 Ti) (Ultracentrifuge Optima L-90K, Beckman Coulter) (see 4.2.2.4). The soluble cytoplasmic fraction was collected as the supernatant of the latter centrifugation step. The protein concentration of the three fractions was determined by BCA methods (see 4.2.4.1) and the fractions were analyzed by SDS-PAGE using a protein concentration of 10 μ g protein per well.

4.2.3.4. Membrane preparation

4.2.3.4.1. Small-scale membrane preparation

For small-scale membrane preparation, cell pellets obtained from 50 mL of culture were resuspended in 1 mL of cell lysis buffer (20 mM HEPES-NaOH, pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM PMDF, 1 U/mL Benzonase, 3 mM MgCl₂, 200 µg/mL lysozyme). 1 mL of glass beads (0.1 - 0.2 mm diameter) was added and mixed by vortexing for 30 min at 4°C. The glass beads were removed using filtration and cell lysates were collected in a 13 mL tube by centrifugation at 800 rpm (Sigma 4K15 centrifuge). The lysate was centrifuged at 4°C, at 16,000 × g for 20 min (Centrifuge 5415D, Eppendorf) to remove the cell debris. The membranes were harvested by centrifugation at 4°C, 100,000 × g for 1 h (Rotor TLA55, Optima™Max Ultracentrifuge, Beckman Coulter) and resuspended in membrane resuspension buffer (20 mM HEPES-NaOH, pH 8, 100 mM NaCl, 1 mM EDTA) to a final protein concentration of 10 mg/mL.

4.2.3.4.2. Large-scale membrane preparation

For large-scale membrane preparation, cell pellets from 1 L – 12 L of culture were resuspended in cell lysis buffer in a ratio of 1 g cells to 6 mL lysis buffer (20 mM HEPES-NaOH, pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM PMDF, DNAase grade II, 3 mM MgCl₂). The cell suspension was homogenized using a disperser tool (Ultra-Turrax® T25 basic, IKA) and filtered through a porous membrane to remove cell clumps. The cell suspension was then passed through a French Press at a pressure of 19,000 psi (40K cell, Thermo Fisher Scientific) for 3 cycles. Alternatively, the cell suspension (200 mL or more) was lysed in a microfluidizer at a pressure of 12,000 psi (Microfluidics Corp) for 3 times. After cell disruption, the cell lysate was centrifuged at 4°C at 13,000 × g and 23,000 × g for 30 min, respectively, to remove the cell debris (Rotor GSA/SS34, Sorvall RC5B superspeed centrifuge). The supernatant containing the membranes was then ultracentrifuged at 4°C at 200,000 × g for 60 min (Rotor 70 Ti) or at 150,000 × g for 90 min (Rotor 45 Ti) (Ultracentrifuge Optima L-90K, Beckman Coulter). The membrane fraction was pelleted and washed with the membrane resuspension buffer one time. The membrane suspension was ultracentrifuged again as described above. The membrane pellet was resuspended in the membrane resuspension buffer (20 mM HEPES-NaOH, pH 8, 100 mM NaCl, 1 mM EDTA, or in a buffer specified) to a typical protein concentration of 10 mg/mL and flash frozen in 10 mL aliquots (unless otherwise stated) using liquid nitrogen before being stored at -80°C.

4.2.3.5. Detergent screening for solubilization of membrane proteins

For membrane protein solubilization trials, 70 μ L of the membrane suspension buffer supplemented with 4% of the appropriate detergents were added to 70 μ L of membrane suspension containing with 1 mM PMSF. The solubilization was carried out at 4°C for 1 h. Insolubilized material was then separated from the solubilized membrane proteins by ultracentrifugation at 4°C for 1 h, at 100,000 $\times g$ (Rotor TLA100, OptimaTMMax Ultracentrifuge, Beckman Coulter). The protein concentration of the total membrane suspension, of the solubilized supernatant and of the insolubilized pellets was determined by the BCA method and the three fractions were then analyzed by SDS-PAGE and western blot after loading 10 μ g protein per well. The detergents used in this screening procedure are listed in Table 4.20.

Table 4.20: List of detergents used for solubilization screens.

Detergent	Abbreviation	CMC% (w/v)	Concentration for solubilization % (w/v)	Concentration for purification % (w/v)
n-dodecyl- β -D-maltoside	DDM	0.0087	1-2%	0.02-0.05
n-decyl - β -D-maltoside	DM	0.087	1-2%	0.1
n-octyl- β -D-glucopyranoside	OG	0.53	2-6%	0.9
n-nonyl- β -D-glucopyranoside	NG	0.2	1-2%	0.2
n-octyl- β -D-thioglucopyranoside	OTG	0.28	2-3%	0.3
n-dodecyl-N,N-dimethylamine-N-oxide	LDAO	0.026	1-2%	0.1
nonaethylene glycol monododecyl ether	C12E9	0.006	1-2%	0.02
dodecyl-phosphorylcholine	FOS-12	0.053	1-2%	0.08
cyclohexyl-hexyl- β -D-maltoside	Cymal-6	0.028	1-2%	0.05
Cyclohexyl-heptyl- β -D-maltoside	Cymal-7	0.01	1-2%	0.02

4.2.3.6. Protein overproduction and purification

4.2.3.6.1. Subcomplex b_1b_2

The sub-complex b_1b_2 was obtained using vector bb2-pTTQ-A in the host strain *E. coli* C43(DE3) cultured in 2 \times YT medium (supplemented with 0.5% (w/v) glucose, 50 μ g/mL carbenicillin). Cells were induced by 0.2 mM IPTG to an OD₆₀₀ of 0.7-0.8 and further grown at 30°C for 16 h. The cell

pellet was resuspended in buffer A supplemented with 1 mM PMSF and DNAaseI. Cells were lysed using the French Press method at 10,000 p.s.i. (for 3 cycles). Cell membranes were prepared as described above (see 4.2.3.4.2). Membranes were successively resuspended in buffer A (20 mM HEPES, pH 8.0, 3 mM MgCl₂, 1 mM EDTA, 100 mM NaCl) to a protein concentration of 10 mg/mL and divided into 10 mL aliquots. Membranes were solubilized by adding buffer B (20 mM HEPES, pH 7.5, 500 mM NaCl) supplemented with 20 mM imidazole, 1 mM PMSF, and 2% (w/v) DM at a ratio of 20 mL buffer per 100 mg total membrane proteins at 50°C for 1 h, and then ultracentrifuged at 200,000 × g for 1 h to remove the unsolubilized materials. The supernatant was incubated with Ni-NTA resin (Qiagen) pre-equilibrated with buffer B supplemented by 0.25% (w/v) DM at 4°C for 2 h (2 mL resin per 100 mg total membrane proteins). The resin was washed with buffer B supplemented by 0.25% (w/v) DM and imidazole concentrations increasing from 20 mM to 80 mM and then the sub-complex b₁b₂ was eluted by 300 mM imidazole. A final size-exclusion polishing chromatographic run was performed using a Superdex 200 PC3.2/30 column (GE healthcare) and the SMART system (Pharmacia) in buffer B with 0.25% (w/v) DM. Proteins were concentrated by Amicon Ultra devices with a molecular weight cut-off of 10 kDa (Millipore).

4.2.3.6.2. *Subcomplexes F₁-αβγ and F₁-αβγϵ*

The subcomplexes F₁-αβγ and F₁-αβγϵ were expressed from pCL11 and pCL12 vectors, respectively in the host strain *E. coli* C43 (DE3) already containing the pRARE vector for codon usage optimization. Cells were cultured in TB medium (supplemented with 50 µg/mL carbenicillin and 34 µg/mL chloramphenicol) to a cell density of 0.3-0.4 (OD₆₀₀), and then induced with 1 mM IPTG and cultured for further 6 h. Cells were harvested and proteins were extracted using a similar procedure as for subcomplex b₁b₂ but in buffer C (20 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 0.2 mM EDTA, 100 mM NaCl, 10% (v/v) glycerol, 1 mM PMSF, DNaseI). Proteins from the cytosolic fraction were then immediately separated on a Ni-NTA resin following a similar procedure as for subcomplex b₁b₂ but using buffer D (20 mM Tris-HCl, pH 7.5, 5 mM MgCl₂, 300 mM NaCl, 2.5% (v/v) glycerol) with imidazole concentrations of 30 mM (washing buffer) to 150 mM (elution buffer). Imidazole was then immediately removed from the eluate using a PD-10 column (GE healthcare) in buffer D with 150 mM NaCl only. Finally, the proteins were further separated on a Superdex 200 10/300 GL size-exclusion column (GE healthcare) on Äkta purifier systems (GE Healthcare) using the same buffer. Proteins were concentrated by Amicon Ultra devices with a molecular weight cut-off of 50 kDa (Millipore).

4.2.3.6.3. *The whole EAF₁F₀ complex*

The expression vector pCL21 was co-transformed in *E. coli* DK8 with the pRARE plasmid (Novagen). Transformants were selected on LB agar plates containing 50 µg/mL carbenicillin, 34 µg/mL chloramphenicol and 30 µg/mL tetracycline. For protein production, *E. coli* cells were grown in 2× YT medium (supplemented with 50 µg/mL carbenicillin and 34 µg/mL chloramphenicol) to a cell density of 0.3-0.4 (OD₆₀₀), and then induced with 1 mM IPTG and incubated for 6 h. Membranes were prepared following the same procedures used as subcomplex b₁b₂, but using buffer E (20 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 0.2 mM EDTA, 100 mM NaCl, 10% (v/v) glycerol) and resuspended to a protein concentration of 30 mg/mL in 8 mL aliquots. Membrane solubilization was carried out at 4 °C for 2 h by mixing 100 mg total membrane proteins with 15 mL buffer F (50 mM Tris-HCl, pH 7.5, 2.3 mM MgCl₂, 300 mM NaCl, 46 mM imidazole, pH 7.5, 3% (w/v) DDM, 0.05% (w/v) α-PCC) supplemented with 1× complete protease inhibitor cocktail tablets, EDTA-free (Roche). After heat treatment at 50 °C for 30 min (Imamura *et al.*, 2006), the extract was centrifuged at 200,000 × *g* for 1 h. The supernatant fraction was incubated at 4 °C for 2 h with Ni-NTA resin equilibrated with buffer G (20 mM Tris-HCl, pH 7.4, 5 mM MgCl₂, 230 mM NaCl, 2.5% (v/v) glycerol, 0.05% (w/v) DDM, 0.05% (w/v) α-PCC) and supplemented with 30 mM imidazole (ratio of 4 mL resin per 100 mg total membrane proteins). The resin was washed with increasing imidazole concentrations up to 50 mM and EAF₁F₀ was then eluted in buffer G supplemented with 150 mM imidazole. Imidazole was then removed using a PD-10 column and the proteins were maintained in buffer G using only 150 mM NaCl and further purified on a TSK-GEL G4000SW column (TOSOH Bioscience) using Äkta purifier systems (GE Healthcare). Proteins were concentrated by Amicon Ultra devices with a molecular weight cut-off of 100 kDa (Millipore).

4.2.3.7. Protein purification

4.2.3.7.1. Affinity purification

For His-tagged protein purification, Ni-NTA agarose (Qiagen) was used in a batch procedure. The proteins were eluted in the relevant buffers containing an appropriated imidazole concentration (See details in 4.2.3.6). Alternatively HisTrap HP 1 ml (GE healthcare) was used with an Äkta purifier systems (GE Healthcare), following the manufacturer guidelines.

4.2.3.7.2. Size-exclusion chromatography (SEC)

For preparative purposes, the Superdex 200 10/300 GL column (GE healthcare) or the TSK-GEL G4000SW (TOSOH Bioscience) were used with an Äkta purifier systems (GE Healthcare). For analytical purposes, either the Superdex 200 PC 3.2/30 was used on the SMART system

(Pharmacia) or the Superdex 200 5/150 GL (GE healthcare) or the Superose 6 5/150GL (GE healthcare) were used on the Äkta purifier systems (GE Healthcare). Columns were equilibrated with 3 column volumes of running buffer, prior to each run that was performed according to the manufacturer guidelines. Size-exclusion column calibration was performed using the Gel Filtration Calibration Kit (GE Healthcare) with the appropriate running buffer.

4.2.3.7.3. Ion exchange chromatography

For preparative protein purification, the Mono Q 5/50 GL column (GE healthcare) was used on Äkta purifier systems (GE Healthcare). For analytical purposes, the Mini Q 4.6/50 PE column was used on the SMART system (Pharmacia). The column was pre-equilibrated with the relevant buffer but without NaCl, and the proteins were eluted with a segmented gradient to a final NaCl concentration of 1 M. The runs were performed according to the manufacturer guidelines.

4.2.4. Protein characterization by biochemical methods

4.2.4.1. Determination of protein concentration

Protein concentration was determined according to the following methods:

- a) BCA protein assay.** In the bicinchoninic acid (BCA) protein assay the reagent was prepared as a 1:50 (v/v) mixture of BCA assay reagents A and B (Pierce) respectively. BSA samples at concentration ranges from 0.01 to 0.1 mg/mL or from 0.1 to 1 mg/mL were used as standards. 10 μ L of water, of BSA standard or of protein sample were mixed with 200 μ L reagent in a 96- well reaction plate (Nunc) and incubated for 30 min at 37 or 60°C, according to standard procedures. The absorbance at 562 nm was measured with a microtiter plate reader (Trista LB 941 Multimode, Berthold). The concentration of protein was calculated as an average of duplicates. This assay is compatible with relatively high (1%) concentrations of both ionic and non-ionic detergents, but not with imidazole. Therefore, this assay was used to determine the total protein concentration in the membrane and the concentration of the purified proteins (in buffer without imidazole).
- b) Bradford protein assay.** In the Bradford protein assay, the dye reagent (Bio-Rad) was prepared with five-fold dilution in water. BSA samples at concentration ranges from 0.05 to 0.5 mg/mL were used as standards. As for the BCA protein assay, 10 μ L of each standard and sample solution were mixed with 200 μ L reagent, incubated at room temperature for 5 min and the absorbance was measured at 595 nm. The concentration of

protein was calculated as an average of duplicates. This assay was used to determine the purified protein in a buffer containing imidazole.

Finally, when the amount of the purified protein was very limited, the concentration was approximately estimated by measuring the absorbance of the protein at 280 nm in a NanoDrop ND-1000 spectrophotometer (one A_{280} unit = 1 mg/mL).

4.2.4.2. Denaturing gel electrophoresis (SDS-PAGE)

Denaturing gel electrophoresis was performed using the anionic detergent sodium dodecylsulfate (SDS). The protein samples mixed with sample buffers were typically incubated for 5 min at 90 °C. The samples were run in a commercial gel chamber (XCell SureLock Mini-Cell, Invitrogen) on NuPAGE® 4-16% Bis-tris gel (Invitrogen), at constant 200 V, for 45 min, using MES/SDS running buffer. Alternatively for proteins in the molecular weight range of 2 – 20 kDa, the samples were run on Novex® 10% tricine gels (Invitrogen), at constant 125 V, 90 min, using Tricine/SDS running buffer. Additionally, for optimal resolution of EAF₁F₀ on SDS-PAGE, 13.2 % tricine gels were self-cast (Schagger and von Jagow, 1987) using 1.0 mm thick gel cassettes (Invitrogen) as described in Table 4.21. The electrophoresis was performed using 1× anode and 1× cathode running buffer at constant 40 mA/ gel at 4°C for 2.5 h (Power HC power supply, Bio-Rad).

Table 4.21: List of buffers and chemicals used for casting the 13.2% tricine gels.

Chemicals	5% stacking gel (1 gel)	13.2% separating gel (1 gel)
30% (w/v) acrylamide/bis solution	533 µL	4.467 mL
3 M Tris/HCl pH 8.45	1.033 mL	3.333 mL
60 % glycerol	-	2.2 mL
10 % (w/v) SDS	31 µL	100 µL
TEMED	6 µL	5 µL
10% (w/v) APS	50 µL	50 µL
ddH ₂ O	2.533 mL	- mL
Total volume	4.186 mL	10.155 mL

Table 4.22: List of solutions used for SDS-PAGE electrophoresis.

Gel type	Solutions	Components
13.2% self- cast tricine gel	1× Sample buffer	50 mM Tris/HCl, pH 6.8 2% (w/v) SDS 12% (v/v) glycerol 0.01% (w/v) bromophenol blue 0.01% servablue G
	1× Anode running buffer	0.2 M Tris/HCl, pH 8.9
	1× Cathode running buffer	0.1 M Tris 0.1 M Tricine 0.1% (w/v) SDS pH 8.25
4-16% Bis-tris (Invitrogen)	5× Sample buffer	0.25 mM Tris/HCl, pH 8.0 25% (v/v) glycerol 12.5% (v/v) β-ME 7.5% (w/v) SDS 0.05% (w/v) bromophenol blue
	1× MES/SDS running buffer	50 mM MES 50 mM Tris 0.1% (w/v) SDS 1 mM EDTA pH 7.3
10% Tricine (Invitrogen)	4× Sample buffer	900 mM Tris HCl 24% (v/v) glycerol 8% (w/v) SDS 0.005% (w/v) Coomassie Blue G 0.005% (w/v) Phenol Red pH 8.45
	1× Tricine/SDS running buffer	100 mM Tris 100 mM Tricine 0.1% (w/v) SDS pH 8.3

4.2.4.3. Non denaturing gel electrophoresis (Native PAGE)

Native PAGE was performed with precast NativePAGE™ Novex® 4-16% Bis-Tris Mini in an XCell SureLock™ Mini-Cell chamber (Invitrogen) according to the manufacturer's specifications. The runs were all performed at 4°C, at 150 V for the first 60 min and then at 250 V for 30 min (Pharmacia LKB ECPS 3000/ 150).

Blue Native PAGE (BN-PAGE) was performed using 50 mM BisTris and 50 mM Tricine as the anode buffer and the same buffer supplemented with 0.02% (w/v) or 0.002% Coomassie brilliant blue G-250 as the dark or light blue cathode buffer, respectively. For BN-PAGE, the samples with

detergents were supplied with 5% (w/v) Coomassie brilliant blue G-250 as an additive prior to loading. For the Clear Native PAGE (CN-PAGE), the cathode buffer was supplied with 0.05% sodium deoxycholate (DOC).

Table 4.23: List of buffers used for BN-PAGE.

Buffers and solutions	Compositions
20× Cathode buffer additive	0.4% (w/v) Coomassie G-250 in H ₂ O
1× Running buffer	50 mM BisTris 50 mM Tricine
Light blue cathode buffer	0.02% Coomassie G-250 in 1× running buffer
Dark blue cathode buffer	0.002% Coomassie G-250 in 1× running buffer
Clear PAGE cathode buffer	0.05% DOC in 1× running buffer
1× sample buffer	50 mM BisTris 6N HCl 50 mM NaCl 3.17 % (v/v) glycerol 0.001% (w/v) Ponceau S pH 7.2

4.2.4.4. Electro-elution

After the electrophoretic run, stained protein bands containing EAF₁F₀ were excised from dark BN-PAGE gels. Proteins were eluted into a membrane cap with a molecular weight cut off (MWCO) of 12 – 15 kDa (Bio-Rad), using the Electro-Eluter Model 422 (Bio-Rad). The electro-elution was performed at 4°C overnight, in a buffer of 50 mM BisTris and 50 mM Tricine, at constant 8 – 10 mA/glass tube (Pharmacia LKB ECPS 3000/ 150). Typically 400 – 600 µL of eluted proteins were collected and then concentrated by Amicon Ultra devices with a MWCO of 100 kDa (Millipore) and rebuffered in 20 mM Tris-HCl, pH 7.4, 5 mM MgCl₂, 150 mM NaCl, 2.5% (v/v) glycerol, 0.05% (w/v) DDM, 0.05% (w/v) α -PCC.

4.2.4.5. Two-dimensional electrophoresis (2-D Native/SDS-PAGE)

2-D Native/SDS-PAGE was performed to investigate the subunit composition of the purified subcomplexes and whole complex. In the first dimension the gel was run as a light BN-PAGE (see 4.2.4.4). After electrophoresis, the gel strip was cut from the light BN-PAGE and transferred into a sterile 15 mL conical tube. The gel was successively incubated at room temperature in 5 mL

Reducing solution for 15 – 30 min, in 5 mL Alkylating solution for 15 – 30 min, and in 5 mL Quenching solution for 15 min. After decanting the Quenching solution, the gel strip was loaded on Novex®4-16% Bis-tris gels with 2D-well (Invitrogen). The gel strip was overlaid with 60 µL of 1 × NuPAGE® LDS sample buffer (Invitrogen), and the run was performed in 1× MES/SDS running buffer at 200 V at 4°C.

Alternatively, the gel lane was incubated in 1× cathode buffer (0.1 M Tris, 0.1 M Tricine, 0.1% (w/v) SDS, pH 8.25) and loaded on 13.2% tricine self-cast gel (in place of the stacking portion, see 4.2.4.5), preequilibrated by electrophoresis at 40 mV/gel at 4°C.

Table 4.24: List of buffers used for 2D Native/SDS-PAGE.

Buffers and solutions	Components
Reducing solution	1× NuPAGE® LDS sample buffer with 50 mM DTT
Alkylating solution	1× NuPAGE® LDS sample buffer with 50 mM N,N-Dimethylacrylamide (DMA)
Quenching solution	1× NuPAGE® LDS sample buffer with 50 mM DTT and 20% (v/v) ethanol

4.2.4.6. Gel staining

Gels were stained by either one of the following two methods:

- a) **Coomassie blue staining.** After electrophoresis, the gels were directly stained with Coomassie blue staining solution (detection range 5 – 10 µg protein). After 30 – 60 min at room temperature, the gels were destained using destaining solution.
- b) **Silver staining.** For higher detection sensitivity (0.05 – 0.1 µg protein), the gels were

Table 4.25: List of buffers used for Coomassie blue staining.

	Components	Preparation
Coomassie blue staining solution	0.04% (w/v) Coomassie brilliant blue R-250 40% (v/v) ethanol 10% (v/v) acetic acid H ₂ O	1 g 150 mL 50 mL Add to 500 mL
Coomassie blue destaining solution	30% (v/v) ethanol 10% (v/v) acetic acid H ₂ O	150 mL 50 mL Add to 500 mL

stained with silver nitrate (Nesterenko *et al.*, 1994). After electrophoresis, the gel was incubated in fixing solution for 5 min, then rinsed 3 times with water for 5 s and then washed in water for 5 min. The gel was then sequentially treated with 50% acetone for 5 min and 10% Na₂S₂O₃·5H₂O for 1 min, before being rinsed 3 times in water, stained with silver nitrate for 8 min, rinsed 2 more times with water, and finally developed. When the appropriate contrast was achieved, the developing solution was decanted, the gel was washed in acetic acid for 30 sec, and finally rinsed with water.

Alternatively silver staining was performed with the SilverQuest™ silver staining Kit (Invitrogen) following the manufacturer's specifications.

Table 4.26: List of solutions used for silver staining.

Step	Components	Volume
Fixation	50% acetone	60 mL
	50% TCA	1.5 mL
	37% HCHO	25 µL
Pretreatment	50% acetone	60 mL
Pretreatment	10% Na ₂ S ₂ O ₃ ·5H ₂ O	100 µL
	H ₂ O	Add to 60 mL
Staining	20% AgNO ₃	0.8 mL
	37% HCHO	0.6 mL
	H ₂ O	Add to 60 mL
Development	Na ₂ CO ₃	2 g
	37% HCHO	25 µL
	10% Na ₂ S ₂ O ₃ ·5H ₂ O	25 µL
	H ₂ O	Add to 60 mL
Stop	1% acetic acid	6 mL
	H ₂ O	Add to 60 mL

After Coomassie blue or silver staining, the gels were dried under vacuum at 80°C for 90 min (Membrane vacuum pump MP40 & Mididry D62, Biometra) to obtain xerogels on chromatography paper (0.34 mm 3MM Chr, Whatman).

4.2.4.7. Antibody production

In this work, five polyclonal antibodies (*AQUEA_AtpA*, *AtpC*, *AtpG*, *AtpH*, *AtpL*) were generated against *A. aeolicus* subunit F₁-α, F₁-ε, F₁-γ, F₁-δ and F₀-c. These antibodies were raised in rabbits against chemically synthesized peptides coupled to keyhole limpet hemocyanin (KLH), followed by the 70 days immunization protocol. The peptides used were KEALDAFKQKFVP (*AtpA*),

EWEKEAEKARTLLELVEKYR (*AtpC*), KLSPRDIKRKIQGIKNTKR (*AtpG*), KTINDILNRQIEIEVKEDP (*AtpH*), and RGTQEGVARNPNAGGRLQ (*AtpL*), respectively (Thermo Fisher Scientific). The polyclonal antibodies were obtained from the crude sera by affinity purification (Thermo Fisher Scientific). See Appendix for the antigen profile and immunization protocol.

4.2.4.8. Western blot analysis

The proteins separated on SDS-PAGE gels were transferred to the PVDF membrane using the iBlot® 7-minute Blotting System (Invitrogen) and iBlot® Transfer Stack, PVDF Regular or Mini (Invitrogen). After the transfer was completed, the membrane was incubated in blocking buffer for 1 h at room temperature or overnight at 4 °C. After blocking, the membrane was subjected to immuno-detection using conventional methods. Briefly, the membrane was washed in 1× TBST buffer for 5 min with gentle agitation for 3 times, incubated with the primary antibodies for 1 - 2 h, then washed 3 times with 1× TBST buffer, incubated with the second antibodies conjugated to alkaline phosphatase (AP) for 1 h, washed for 5 min (3 times) with 1× TBST buffer, and washed with AP buffer 3 times for 5 min. Finally, the signals were detected using the colorimetric BCIP-NBT detection system (Thermo Fisher Scientific). Alternatively, for faster analysis, the membrane was subjected to immune-detection by the SNAP i.d.® system (Millipore) according to the manufacturer's instructions.

The poly-His-tagged proteins were detected using a monoclonal α -poly-histidine-alkaline phosphatase conjugated antibody (Sigma-Aldrich) (1:2,000 dilution). StrepII-tagged proteins were detected by streptavidin coupled to alkaline phosphatase (Sigma-Aldrich). For detection of subunits F_1 - α , F_1 - γ , F_1 - ϵ , F_1 - δ and F_0 -c, the custom peptide polyclonal antibodies *AQUEA_AtpA*, *AQUEA_AtpG*, *AQUEA_AtpC*, *AQUEA_AtpH*, *AQUEA_AtpL* (1:2,000 dilution) were generated and used as the primary antibodies (Thermo Fisher Scientific, see 4.2.4.7) and monoclonal mouse anti-rabbit IgG conjugated with alkaline phosphatase (Sigma-Aldrich) as the secondary antibody. Finally, for detection of subunit F_1 - β , KHL-conjugated synthetic peptide polyclonal antibody (Agrisera) was used with a 1:2,000 dilution as the primary antibody.

Table 4.27: List of solutions used for Western blot analysis.

Solutions	Components
Transfer buffer (1×)	25 mM Tris/HCl, pH 8.3 150 mM glycine 10% (v/v) methanol
TBST buffer (10×)	100 mM Tris/HCl, pH 8.0, 1.5 M NaCl 0.5% (v/v) Tween-20
Blocking buffer	1% (w/v) BSA in 1× TBST buffer
Blocking buffer (strep II)	2 µg avidin in 1× TBST buffer
AP buffer	100 mM Tris/HCl, pH 9.5, 100 mM NaCl, 5 mM MgCl ₂
BCIP-NBT detection buffer	250 µg/mL BCIP (solubilized in DMF) 500 µg/mL NBT (solubilized in 70 % DMF) in 1× AP buffer
Ponceau S solution	0.1% (w/v) Ponceau S 5% (v/v) acetic acid

4.2.4.9. Dot blot analysis

Dot blotting was used for preliminary screening of protein expression in crude lysates. For dot blot analysis, PVDF membrane (Immobilon-P, Millipore) strips of 8 × 11 cm in size were soaked in methanol for 30 sec, washed in water for 5 min, washed with 150 µL of 1× transfer buffer (25 mM Tris/HCl, pH 8.3, 150 mM glycine, 10% (v/v) methanol) with the use of 96-well dot blot apparatus (Bio-Rad) and decanted by applying vacuum. 30 µL of samples (the whole cell lysates prepared in 20% SDS) were then applied on the PVDF membrane and incubated at room temperature for 30 min and decanted by applying vacuum. The PVDF membrane was washed 3 times with 150 µL of TBST buffer and the buffer decanted by applying a vacuum. Finally, the membrane was developed as described for standard Western blot analyses (see 4.2.4.8).

4.2.4.10. Colony blot analysis

The colony-blot procedure was used to select the clones characterized by the highest protein overexpression phenotypes. Freshly transformed cells were plated on LB-agar plates (master plates) containing the appropriate antibiotics and grown overnight. Successively, the plates were equilibrated at room temperature for 15-30 min, and a sterile nitrocellulose transfer membrane (HATF 0.45 µm, 82 mm in diameter, Millipore) was placed on the top of the colonies. The nitrocellulose membrane was then transferred to a fresh LB-agar plate (containing 250 mM IPTG and relevant antibiotics) with the colonies side up and expression was carried out at 37 °C for 4 h.

In parallel, the master plates were incubated at 30°C for 4 h to allow the colonies to regrow. The nitrocellulose membrane was then placed in a Petri dish on top of filter paper soaked with the solutions listed in Table 4.28 and incubated at room temperature as follows: 1) 10% SDS solution: 10 min; 2) Denaturing solution: 5 min; 3) Neutralization solution: 5 min; 4) Neutralization solution: 5 min; 5) 2 × SSC: 15 min. Finally, the membrane was developed as described for standard Western blot analyses (see 4.2.4.8).

Table 4.28: List of solution used for colony-blot analysis.

Solution for colony-blot analysis	Compositions
SDS solution	10 % (w/v) SDS
Denaturing solution	0.5 M NaOH 1.5 M NaCl
Neutralization solution	0.5 M Tris/HCl, pH 7.4 1.5 M NaCl
20 × SSC	87.65 g NaCl 50.25 g trisodium citrate·2H ₂ O H ₂ O up to 500 mL

4.2.5. Protein identification by mass spectrometry

4.2.5.1. Preparation of subunit c monomers from native *A. aeolicus*

F₁F₀ ATP synthase

A. aeolicus F₁F₀ ATP synthase was purified as described (Peng *et al.*, 2006). Its subunit c was extracted with organic solvents as described previously (Cattell *et al.*, 1971; von Ballmoos *et al.*, 2002), but with the following modifications. Typically, the sample (250 µg of protein or membrane) was mixed with a 10-fold excess of chloroform/methanol (1:1, v:v) to precipitate insoluble proteins. After centrifugation at room temperature for 5 min at 14,000 × g (Centrifuge 5415D, Eppendorf), 2.5 volumes of 10 mM Tris/HCl pH 8.0 were added to one volume of supernatant to achieve phase separation. The organic phase containing the hydrophobic subunits c was collected. The sample was then evaporated to dryness in a speed vacuum concentrator and stored at -20°C before analysis.

4.2.5.2. DCCD labeling assay

50 μ L of the native AAF₁F₀ at a concentration of 10 mg/mL was dialyzed with Slide-A-lyzer Mini Dialysis Units (MWCO 10 kDa, Pierce) against 1 L of buffer (20 mM Tris/HCl, pH 7.4, 20 mM MgCl₂, 0.05% (w/v) DDM) at 4 °C for 2 - 3 days. After dialysis, the sample was diluted to 2 mg/mL in the same buffer. For screening different NaCl concentrations, AAF₁F₀ (2 mg/mL) was diluted to 1 mg/mL with the same buffer supplemented with appropriate NaCl concentrations (final NaCl concentrations were 0 mM, 10 mM and 100 mM). Then 0.5 μ L of DCCD (100 μ M) was added to 50 μ L of AAF₁F₀ (1 mg/mL) to a final concentration of 1 μ M, and incubated at room temperature. 10 μ L aliquots were taken at different time points (0 min (as the unlabeled control), 15 min, 30 min and 120 min, respectively), transferred into new microcentrifuge tubes and mixed with 10-fold excess of chloroform/methanol (1:1, v:v) to prepare the subunit c monomers for MALDI-MS as described above (see 4.2.5.1). Each measurement was performed in triplicate.

4.2.5.3. Peptide mass fingerprinting (PMF)

Identification of SDS-PAGE-separated proteins was performed on reduced and alkylated samples digested with trypsin and chymotrypsin. Instead, proteins to be identified in solution were treated with up to 5 M urea prior to reduction, alkylation and digestion. Using a nano-HPLC (Proxeon easy-nLC), the proteolytic peptides were loaded on reverse phase columns (trapping column: particle size 3 μ m, C18, L=20mm; analytical column: particle size 3 μ m, C18, L=15cm; NanoSeparations, Nieuwkoop, The Netherlands), and eluted in gradients of water (0.1% (v/v) formic acid, buffer MS(A)) and acetonitrile (0.1% (v/v) formic acid, buffer MS(B)) with a ramp of 5% to 65% MS(B) in up to 120 minutes at flow rates of 300 nl/min. Peptides eluting from the column were ionized online using a chip-based nano-electrospray source (Advion Triversa NanoMate, 2.5 μ m nozzles, "G"-chips, Advion Biosciences, UK) and analysed in a quadrupole time-of-flight mass spectrometer (Bruker maXis). Mass spectra were acquired over the mass range 50-2200 m/z, and sequence information was acquired by computer-controlled, data-dependent automated switching to MS/MS using collision energies based on mass and charge state of the candidate ions.

The data sets were processed using a standard proteomics script with the software Bruker DataAnalysis 4.0 Service Pack 5 Build 283 and exported as Mascot generic files. Spectra were internally recalibrated using autoproteolytic trypsin fragments when applicable.

Proteins were identified by matching the derived mass lists against the NCBI nr database (downloaded from <http://www.ncbi.nlm.nih.gov/>) on a local Mascot server (Version 2.3.02, Matrix

Sciences, UK). In general, a mass tolerance ± 0.02 Da for parent ion and fragment spectra, two missed cleavages, oxidation of Met and fixed modification of carbamidomethyl cysteine were selected as matching parameters in the search program. PMF was performed by Dr. Julian D Langer and Imke Wuellenweber, at Department of Molecular Membrane Biology, the Max Planck Institute of Biophysics, Frankfurt.

4.2.5.4. MALDI-TOF-MS measurements

Chloroform/methanol extracts were mixed in a 1:1 (v/v) ratio with matrix 2,5-dihydroxyacetophenone (15 mg/mL 2,5-dihydroxyacetophenone in 75% ethanol in 20 mM sodium citrate; Bruker Daltonics) or 2,5-dihydroxybenzoic acid (30 mg of 2,5-dihydroxybenzoic acid/100 μ l of TA solution (0.1% trifluoroacetic acid/acetonitrile, 1:2 (v/v); Bruker Daltonics) and spotted on ground steel target plates (Bruker Daltonics). MALDI mass spectra were recorded in a mass range of 5–20 kDa using a Bruker Autoflex III Smartbeam mass spectrometer. Detection was optimized for m/z values between 5 and 20 kDa and calibrated using calibration standards (protein molecular weight calibration standard 1; Bruker Daltonics). MALDI-TOF-MS was performed by Dr. Julian D Langer, Imke Wuellenweber and myself, in Department of Molecular Membrane Biology, at the Max Planck Institute of Biophysics, Frankfurt.

4.2.6. Enzymatic activity assays

4.2.6.1. ATP hydrolysis activity assay by phosphate determination

ATP hydrolysis activity was measured monitoring phosphate production using the LeBel method (LeBel *et al.*, 1978) with minor modifications. The standard curve was prepared using 1 mg/mL K_2HPO_4 , corresponding to 0, 0.022, 0.044, 0.088, 0.131 and 0.175 μ mol phosphate (Pi). The end volume of all samples was 400 μ L. After incubation of samples in 100 μ L of reaction buffer (50 mM Tris/HCl, pH 8.0, 3 mM $MgCl_2$, 10 mM ATP, pH 7.0) for 5 min at 80°C, the reaction was stopped by adding 500 μ L 0.5 M trichloroacetic acid (TCA) and placing on ice. 500 μ L freshly prepared LeBel-reagent (3.6 M acetic acid, 0.66 M sodium acetate, 20 mM copper sulfate, mixing LeBel-reagent A:B:C with a ratio of 6:1:1 in Table 4.29) was added to the reaction mixture and further incubated for 5 min at room temperature. The absorption of the copper-reduced phosphomolybdate complex was then measured in polystyrene cuvettes (10 \times 4 \times 45 mm, optical pathway 10 mm, Sarstedt) at room temperature at 745 nm using the Agilent 8453 UV-Vis spectroscopy.

Table 4.29: LeBel-reagent solutions.

Solutions	Components
LeBel-reagent A	0.25% (w/v) CuSO ₄ ·5H ₂ O 4.6% (w/v) NaAc·3H ₂ O, pH 4.0
LeBel-reagent B	5%(w/v) Ammonium molybdate tetrahydrate
LeBel-reagent C	2% (w/v) 4-methyl-aminophenolhemisulfate 5% (w/v) Na ₂ SO ₃

4.2.6.2. *In-gel* ATP hydrolysis assays

In-gel ATP hydrolysis assays were performed according to Peters (Peters *et al.*, 1992) with the following modifications. After electrophoresis, the BN-PAGE gel were incubated for 3 h at 80 °C in buffer I (50 mM Tris-HCl, pH 8.0, 3 mM MgCl₂, 10 mM ATP, 0.05% (w/v) DDM, 2 mM Pb (NO₃)₂). After brief washing in water, the gels were stained with 1% (w/v) fresh sodium sulfide. The formation of brown lead sulfide precipitates was observed visually.

4.2.6.3. Preparation of inverted membrane vesicles

The cell suspension from 1 L culture was passed through French Press cell at a pressure of 16,000 psi (40K cell, Thermo Fisher Scientific) for 3 cycles. After cell disruption, the cell debris was removed by two-step centrifugation at 4°C for 30 min at 13,000 × *g* and 23,000 × *g*, respectively. The supernatant containing the membrane vesicles was then ultracentrifuged at 4°C for 60 min at 200,000 × *g* (Rotor 70 Ti) or for 90 min at 150,000 × *g* (Rotor 45 Ti) (Ultracentrifuge Optima L-90K, Beckman Coulter). The membrane vesicles were resuspended in 1 mL of buffer E (20 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 0.2 mM EDTA, 100 mM NaCl, 10% (v/v) glycerol), passed through a 10 mL Sephadex G-50 column (Sigma-Aldrich) prepared by soaking in water overnight, and equilibrated with the same buffer. The inverted membrane vesicles were flash-frozen as drops in liquid nitrogen and stored in liquid nitrogen until use. The inverted membrane vesicles were used for ATP synthesis measurement.

4.2.7. Single particle electron microscopy

EAF₁F₀ eluted from Native PAGE gels was negatively stained with 1% (w/v) uranyl acetate. Electron micrographs were collected using a Philips CM120 (FEI, Eindhoven) at an accelerating voltage of 120 kV under low dose conditions. Images were taken at a magnification of 44,000× on Kodak SO-163 electron image film. The negatives were developed in full-strength D-19 developer for 12 min. Negatives were digitized on a PhotoScan scanner (Z/I Imaging, Aalen, Germany) at a

pixel size of 7 μm . Adjacent pixels were averaged to yield a pixel size on the specimen of 4.77 \AA . Approximately 2000 particle images were selected using the boxer program from EMAN (Ludtke *et al.*, 1999) and aligned, classified and averaged using Imagic V (van Heel *et al.*, 1996). Single particle electron microscopy was performed by Dr. Janet Vonck and Matteo Allegretti, in Department of Structural Biology, at the Max Planck Institute of Biophysics, Frankfurt.

4.2.8. Protein crystallization

Crystallization was attempted for subcomplex b_1b_2 . Briefly, the homogeneous and purified protein solutions were concentrated to a final protein concentration of 6 – 8 mg/mL. Crystallization was then performed following the sitting-drop vapor diffusion method at 18°C. For random screening, 100 nL + 100 nL or 200 nL + 200 nL protein / reservoir drops were equilibrated against 100 μL reservoir solution using nanoliter dispensing robots Honeybee 963 (Cartesian Technologies) and Mosquito (TTP Labtech) in CrystalQuickTM 96-well plates (Greiner Bio-one). All buffers were prepared with ddH₂O (18 Ω) and filtered (0.2 μm membrane cut-off) to ensure the highest homogeneity of the particles in solution. The crystallization plates were incubated at 18°C in the robotic incubator Crystal Farm (Bruker) and monitored by automatic imaging and the software crystal farm navigator. No successful crystallization hit was identified to date.

References:

1. Abrahams, J. P., A. G. Leslie, R. Lutter and J. E. Walker (1994). "Structure at 2.8 Å resolution of F₁-ATPase from bovine heart mitochondria." *Nature* **370**(6491): 621-628.
2. Abrahams, J. P., S. K. Buchanan, M. J. Van Raaij, I. M. Fearnley, A. G. Leslie and J. E. Walker (1996). "The structure of bovine F₁-ATPase complexed with the peptide antibiotic efrapeptin." *Proc Natl Acad Sci U S A* **93**(18): 9420-9424.
3. Akiyama, Y., A. Kihara and K. Ito (1996). "Subunit a of proton ATPase F₀ sector is a substrate of the FtsH protease in *Escherichia coli*." *FEBS Lett* **399**(1-2): 26-28.
4. Angevine, C. M., K. A. Herold, O. D. Vincent and R. H. Fillingame (2007). "Aqueous access pathways in ATP synthase subunit a. Reactivity of cysteine substituted into transmembrane helices 1, 3, and 5." *J Biol Chem* **282**(12): 9001-9007.
5. Arechaga, I., B. Miroux, M. J. Runswick and J. E. Walker (2003). "Over-expression of *Escherichia coli* F₁F₀-ATPase subunit a is inhibited by instability of the *uncB* gene transcript." *FEBS Letters* **547**(1-3): 97-100.
6. Bairoch, A. and R. Apweiler (2000). "The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000." *Nucleic Acids Res* **28**(1): 45-48.
7. Barkocygallagher, G. A., J. G. Cannon and P. J. Bassford (1994). "Beta-turn formation in the processing region is important for efficient maturation of *Escherichia Coli* maltose-binding protein by signal peptidase-I *in-vivo*." *Journal of Biological Chemistry* **269**(18): 13609-13613.
8. Barnes, W. M. and E. Tuley (1983). "DNA sequence changes of mutations in the histidine operon control region that decrease attenuation." *J Mol Biol* **165**(3): 443-459.
9. Bauer, M. F., S. Hofmann, W. Neupert and M. Brunner (2000). "Protein translocation into mitochondria: the role of TIM complexes." *Trends Cell Biol* **10**(1): 25-31.
10. Benz, M., T. Bals, I. L. Gugel, M. Piotrowski, A. Kuhn, D. Schunemann, J. Soll and E. Ankele (2009). "Alb4 of Arabidopsis promotes assembly and stabilization of a non chlorophyll-binding photosynthetic complex, the CF₁CF₀-ATP synthase." *Mol Plant* **2**(6): 1410-1424.
11. Bianchet, M. A., J. Hullihen, P. L. Pedersen and L. M. Amzel (1998). "The 2.8-Å structure of rat liver F₁-ATPase: configuration of a critical intermediate in ATP synthesis/hydrolysis." *Proc Natl Acad Sci U S A* **95**(19): 11065-11070.
12. Birkenhager, R., M. Hoppert, G. Deckers-Hebestreit, F. Mayer and K. Altendorf (1995). "The F₀ complex of the *Escherichia coli* ATP synthase. Investigation by electron spectroscopic imaging and immunoelectron microscopy." *Eur J Biochem* **230**(1): 58-67.
13. Bottcher, B., I. Bertsche, R. Reuter and P. Graber (2000). "Direct visualisation of conformational changes in EF₀F₁ by electron microscopy." *J Mol Biol* **296**(2): 449-457.
14. Bowler, M. W., M. G. Montgomery, A. G. Leslie and J. E. Walker (2006). "How azide inhibits ATP hydrolysis by the F-ATPases." *Proc Natl Acad Sci U S A* **103**(23): 8646-8649.
15. Bowler, M. W., M. G. Montgomery, A. G. Leslie and J. E. Walker (2007). "Ground state structure of F₁-ATPase from bovine heart mitochondria at 1.9 Å resolution." *J Biol Chem* **282**(19): 14238-14242.
16. Boyer, P. D. (1993). "The binding change mechanism for ATP synthase--some probabilities and possibilities." *Biochim Biophys Acta* **1140**(3): 215-250.

References

17. Boyer, P. D. (1997). "The ATP synthase--a splendid molecular machine." *Annu Rev Biochem* **66**: 717-749.
18. Braig, K., R. I. Menz, M. G. Montgomery, A. G. Leslie and J. E. Walker (2000). "Structure of bovine mitochondrial F₁-ATPase inhibited by Mg²⁺ ADP and aluminium fluoride." *Structure* **8**(6): 567-573.
19. Brandt, K., D. B. Muller, J. Hoffmann, C. Hubert, B. Brutschy, G. Deckers-Hebestreit and V. Muller (2013). "Functional production of the Na⁺ F₁F_o ATP synthase from *Acetobacterium woodii* in *Escherichia coli* requires the native *AtpI*." *J Bioenerg Biomembr* **45**(1-2): 15-23.
20. Brodie, A. F. and C. T. Gray (1956). "Activation of coupled oxidative phosphorylation in bacterial particulates by a soluble factor (s)." *Biochim Biophys Acta* **19**(2): 384-386.
21. Brusilow, W. S., M. A. Scarpetta, C. A. Hawthorne and W. P. Clark (1989). "Organization and sequence of the genes coding for the proton-translocating ATPase of *Bacillus megaterium*." *J Biol Chem* **264**(3): 1528-1533.
22. Cain, B. D. and R. D. Simoni (1986). "Impaired proton conductivity resulting from mutations in the a subunit of F₁F_o ATPase in *Escherichia coli*." *J Biol Chem* **261**(22): 10043-10050.
23. Cain, B. D. and R. D. Simoni (1988). "Interaction between Glu-219 and His-245 within the a subunit of F₁F_o-ATPase in *Escherichia coli*." *J Biol Chem* **263**(14): 6606-6612.
24. Cain, B. D. (2000). "Mutagenic analysis of the F_o stator subunits." *J Bioenerg Biomembr* **32**(4): 365-371.
25. Capaldi, R. A. and R. Aggeler (2002). "Mechanism of the F₁F_o-type ATP synthase, a biological rotary motor." *Trends Biochem Sci* **27**(3): 154-160.
26. Carbajo, R. J., F. A. Kellas, M. J. Runswick, M. G. Montgomery, J. E. Walker and D. Neuhaus (2005). "Structure of the F₁-binding domain of the stator of bovine F₁F_o-ATPase and how it binds an alpha-subunit." *J Mol Biol* **351**(4): 824-838.
27. Cattell, K. J., C. R. Lindop, I. G. Knight and R. B. Beechey (1971). "The identification of the site of action of NN'-dicyclohexylcarbodi-imide as a proteolipid in mitochondrial membranes." *Biochem J* **125**(1): 169-177.
28. Cingolani, G. and T. M. Duncan (2011). "Structure of the ATP synthase catalytic complex F₁ from *Escherichia coli* in an autoinhibited conformation." *Nat Struct Mol Biol* **18**(6): 701-707.
29. Cole, S. T., T. Grundstrom, B. Jaurin, J. J. Robinson and J. H. Weiner (1982). "Location and nucleotide sequence of *frdB*, the gene coding for the iron-sulphur protein subunit of the fumarate reductase of *Escherichia coli*." *Eur J Biochem* **126**(1): 211-216.
30. Cross, R. L. and V. Muller (2004). "The evolution of A-, F-, and V-type ATP synthases and ATPases: reversals in function and changes in the H⁺/ATP coupling ratio." *FEBS Lett* **576**(1-2): 1-4.
31. Dalbey, R. E., A. Kuhn and G. von Heijne (1995). "Directionality in protein translocation across membranes: the N-tail phenomenon." *Trends Cell Biol* **5**(10): 380-383.
32. Dalbey, R. E., P. Wang and A. Kuhn (2011). "Assembly of bacterial inner membrane proteins." *Annu Rev Biochem* **80**: 161-187.
33. Das, A. and L. G. Ljungdahl (1997). "Composition and primary structure of the F₁F_o ATP synthase from the obligately anaerobic bacterium *Clostridium thermoaceticum*." *J Bacteriol* **179**(11): 3746-3755.
34. Das, A. and L. G. Ljungdahl (2003). "*Clostridium pasteurianum* F₁F_o ATP synthase: operon, composition, and some properties." *J Bacteriol* **185**(18): 5527-5535.

35. Dautant, A., J. Velours and M. F. Giraud (2010). "Crystal structure of the Mg-ADP-inhibited state of the yeast F₁c₁₀-ATP synthase." *J Biol Chem* **285**(38): 29502-29510.
36. de Gier, J. W. and J. Luirink (2001). "Biogenesis of inner membrane proteins in *Escherichia coli*." *Mol Microbiol* **40**(2): 314-322.
37. Deckert, G., P. V. Warren, T. Gaasterland, W. G. Young, A. L. Lenox, D. E. Graham, R. Overbeek, M. A. Snead, M. Keller, M. Aujay, R. Huber, R. A. Feldman, J. M. Short, G. J. Olsen and R. V. Swanson (1998). "The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*." *Nature* **392**(6674): 353-358.
38. Del Rizzo, P. A., Y. Bi, S. D. Dunn and B. H. Shilton (2002). "The "second stalk" of *Escherichia coli* ATP synthase: structure of the isolated dimerization domain." *Biochemistry* **41**(21): 6875-6884.
39. Del Rizzo, P. A., Y. Bi and S. D. Dunn (2006). "ATP synthase b subunit dimerization domain: a right-handed coiled coil with offset helices." *J Mol Biol* **364**(4): 735-746.
40. Denda, K., J. Konishi, T. Oshima, T. Date and M. Yoshida (1989). "A gene encoding the proteolipid subunit of *sulfolobus acidocaldarius* ATPase complex." *Journal of Biological Chemistry* **264**(13): 7119-7121.
41. Dickson, V. K., J. A. Silvester, I. M. Fearnley, A. G. Leslie and J. E. Walker (2006). "On the structure of the stator of the mitochondrial ATP synthase." *EMBO J* **25**(12): 2911-2918.
42. Dmitriev, O., P. C. Jones, W. Jiang and R. H. Fillingame (1999). "Structure of the membrane domain of subunit b of the *Escherichia coli* F₀F₁ ATP synthase." *J Biol Chem* **274**(22): 15598-15604.
43. Duncan, T. M., V. V. Bulygin, Y. Zhou, M. L. Hutcheon and R. L. Cross (1995). "Rotation of subunits during catalysis by *Escherichia coli* F₁-ATPase." *Proc Natl Acad Sci U S A* **92**(24): 10964-10968.
44. Dunn, S. D., D. T. McLachlin and M. Revington (2000). "The second stalk of *Escherichia coli* ATP synthase." *Biochim Biophys Acta* **1458**(2-3): 356-363.
45. Eya, S., M. Maeda, K. Tomochika, Y. Kanemasa and M. Futai (1989). "Overproduction of truncated subunit a of H⁺-ATPase causes growth inhibition of *Escherichia coli*." *J Bacteriol* **171**(12): 6853-6858.
46. Falk, G. and J. E. Walker (1988). "DNA sequence of a gene cluster coding for subunits of the F₀ membrane sector of ATP synthase in *Rhodospirillum rubrum*. Support for modular evolution of the F₁ and F₀ sectors." *Biochem J* **254**(1): 109-122.
47. Feniouk, B. A. and M. Yoshida (2008). "Regulatory mechanisms of proton-translocating F₀F₁-ATP synthase." *Results Probl Cell Differ* **45**: 279-308.
48. Fillingame, R. H., M. J. Miller, D. Fraga and M. E. Girvin (1990). "Helix-Helix Interaction in the Transmembrane F₀ Sector of *Escherichia Coli* ATP Synthase." *Biophys J* **57**(2): A201-A201.
49. Fillingame, R. H., W. Jiang and O. Y. Dmitriev (2000). "Coupling H⁺ transport to rotary catalysis in F-type ATP synthases: structure and organization of the transmembrane rotary motor." *J Exp Biol* **203**(Pt 1): 9-17.
50. Folch, J. and M. Lees (1951). "Proteolipides, a new type of tissue lipoproteins; their isolation from brain." *J Biol Chem* **191**(2): 807-817.
51. Froderberg, L., E. Houben, J. C. Samuelson, M. Chen, S. K. Park, G. J. Phillips, R. Dalbey, J. Luirink and J. W. De Gier (2003). "Versatility of inner membrane protein biogenesis in *Escherichia coli*." *Mol Microbiol* **47**(4): 1015-1027.
52. Fu, L., B. Niu, Z. Zhu, S. Wu and W. Li (2012). "CD-HIT: accelerated for clustering the next-generation sequencing data." *Bioinformatics* **28**(23): 3150-3152.

References

53. Futai, M., T. Noumi and M. Maeda (1988). "Molecular genetics of F₁-ATPase from *Escherichia coli*." J Bioenerg Biomembr **20**(1): 41-58.
54. Gafvelin, G., M. Sakaguchi, H. Andersson and G. von Heijne (1997). "Topological rules for membrane protein assembly in eukaryotic cells." J Biol Chem **272**(10): 6119-6127.
55. Gay, N. J. and J. E. Walker (1985). "Two genes encoding the bovine mitochondrial ATP synthase proteolipid specify precursors with different import sequences and are expressed in a tissue-specific manner." EMBO J **4**(13A): 3519-3524.
56. Gerday, C. and N. Glansdorff (2007). Physiology and biochemistry of Extremophiles. Washington D.C., USA, ASM Press.
57. Gerstel, B. and J. E. McCarthy (1989). "Independent and coupled translational initiation of *atp* genes in *Escherichia coli*: experiments using chromosomal and plasmid-borne *lacZ* fusions." Mol Microbiol **3**(7): 851-859.
58. Gibbons, C., M. G. Montgomery, A. G. Leslie and J. E. Walker (2000). "The structure of the central stalk in bovine F₁-ATPase at 2.4 Å resolution." Nat Struct Biol **7**(11): 1055-1061.
59. Girvin, M. E., V. K. Rastogi, F. Abildgaard, J. L. Markley and R. H. Fillingame (1998). "Solution structure of the transmembrane H⁺-transporting subunit c of the F₁F₀ ATP synthase." Biochemistry **37**(25): 8817-8824.
60. Glick, B. S. and G. Von Heijne (1996). "*Saccharomyces cerevisiae* mitochondria lack a bacterial-type sec machinery." Protein Sci **5**(12): 2651-2652.
61. Groth, G. and E. Pohl (2001). "The structure of the chloroplast F₁-ATPase at 3.2 Å resolution." J Biol Chem **276**(2): 1345-1352.
62. Groth, G. (2002). "Structure of spinach chloroplast F₁-ATPase complexed with the phytopathogenic inhibitor tentoxin." Proc Natl Acad Sci U S A **99**(6): 3464-3468.
63. Guiral, M., L. Prunetti, S. Lignon, R. Lebrun, D. Moinier and M. T. Giudici-Orticoni (2009). "New insights into the respiratory chains of the chemolithoautotrophic and hyperthermophilic bacterium *Aquifex aeolicus*." J Proteome Res **8**(4): 1717-1730.
64. Hakulinen, J. (2012). "Biochemical and structural investigations on architecture of the F₀ complex from *Ilyobacter tartaricus* ATP synthase." Ph.D. Theses. Max-Planck-Institute of biophysics, Frankfurt, Germany.
65. Hakulinen, J. K., A. L. Klyszejko, J. Hoffmann, L. Eckhardt-Strelau, B. Brutschy, J. Vonck and T. Meier (2012). "Structural study on the architecture of the bacterial ATP synthase F₀ motor." Proc Natl Acad Sci U S A **109**(30): E2050-2056.
66. Hanahan, D. (1985). DNA Cloning: a practical approach. McLean, Virginia, USA, IRL Press Ltd.
67. Hara, K. Y., Y. Kato-Yamada, Y. Kikuchi, T. Hisabori and M. Yoshida (2001). "The role of the betaDELSEED motif of F₁-ATPase: propagation of the inhibitory effect of the epsilon subunit." J Biol Chem **276**(26): 23969-23973.
68. Harada, M., S. Ohta, M. Sato, Y. Ito, Y. Kobayashi, N. Sone, T. Ohta and Y. Kagawa (1991). "The Alpha-1-Beta-1 Heterodimer, the unit of ATP Synthase." Biochim Biophys Acta **1056**(3): 279-284.
69. Hartzog, P. E. and B. D. Cain (1994). "Second-site suppressor mutations at glycine 218 and histidine 245 in the alpha subunit of F₁F₀ ATP synthase in *Escherichia coli*." J Biol Chem **269**(51): 32313-32317.

70. Hay, R., P. Bohni and S. Gasser (1984). "How Mitochondria Import proteins." Biochim Biophys Acta **779**(1): 65-87.
71. Hellmuth, K., G. Rex, B. Surin, R. Zinck and J. E. McCarthy (1991). "Translational coupling varying in efficiency between different pairs of genes in the central region of the *atp* operon of *Escherichia coli*." Mol Microbiol **5**(4): 813-824.
72. Hermolin, J. and R. H. Fillingame (1995). "Assembly of F₀ sector of *Escherichia coli* H⁺ ATP synthase. Interdependence of subunit insertion into the membrane." J Biol Chem **270**(6): 2815-2817.
73. Herrmann, R. G., J. Steppuhn, G. S. Herrmann and N. Nelson (1993). "The nuclear-encoded polypeptide Cfo-II from spinach is a real, ninth subunit of chloroplast ATP synthase." FEBS Lett **326**(1-3): 192-198.
74. Herskovits, A. A., E. S. Bochkareva and E. Bibi (2000). "New prospects in studying the bacterial signal recognition particle pathway." Mol Microbiol **38**(5): 927-939.
75. Hirata, R., L. A. Graham, A. Takatsuki, T. H. Stevens and Y. Anraku (1997). "VMA11 and VMA16 encode second and third proteolipid subunits of the *Saccharomyces cerevisiae* vacuolar membrane H⁺-ATPase." J Biol Chem **272**(8): 4795-4803.
76. Horikoshi, K. (1998). Extremophiles. Microbial life in extreme environments. California, USA, Wiley-Liss.
77. Houstek, J., A. Pickova, A. Vojtiskova, T. Mracek, P. Pecina and P. Jesina (2006). "Mitochondrial diseases and genetic defects of ATP synthase." Biochim Biophys Acta **1757**(9-10): 1400-1405.
78. Huber, R. and W. Eder (2006). The Prokaryotes: ecophysiology and biochemistry. Aquificales. Dworkin M. and F. S. Heidelberg, Germany, Springer Verlag. **7**: 925-938.
79. Ihara, K., S. Watanabe, K. Sugimura and Y. Mukohata (1997). "Identification of proteolipid from an extremely halophilic archaeon *Halobacterium salinarum* as an N,N'-dicyclohexyl-carbodiimide binding subunit of ATP synthase." Arch Biochem Biophys **341**(2): 267-272.
80. Imamura, H., S. Funamoto, M. Yoshida and K. Yokoyama (2006). "Reconstitution in vitro of V₁ complex of *Thermus thermophilus* V-ATPase revealed that ATP binding to the A subunit is crucial for V₁ formation." J Biol Chem **281**(50): 38582-38591.
81. Inoue, H., H. Nojima and H. Okayama (1990). "High efficiency transformation of *Escherichia coli* with plasmids." Gene **96**(1): 23-28.
82. Ishmukhametov, R. R., M. A. Galkin and S. B. Vik (2005). "Ultrafast purification and reconstitution of His-tagged cysteine-less *Escherichia coli* F₁F_s ATP synthase." Biochim Biophys Acta **1706**(1-2): 110-116.
83. Jaenicke, R. and R. Sterner (2006). Life at high temperatures. The prokaryotes: ecophysiology and biochemistry. Heidelberg, Germany, Springer Verlag. **2**: 925-938.
84. Jia, L., M. K. Dienhart and R. A. Stuart (2007). "Oxa1 directly interacts with Atp9 and mediates its assembly into the mitochondrial F₁F_o-ATP synthase complex." Mol Biol Cell **18**(5): 1897-1908.
85. Jonckheere, A. I., J. A. Smeitink and R. J. Rodenburg (2012). "Mitochondrial ATP synthase: architecture, function and pathology." J Inherit Metab Dis **35**(2): 211-225.
86. Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol **292**(2): 195-202.
87. Jones, H. M., C. M. Brajkovich and R. P. Gunsalus (1983). "In vivo 5' terminus and length of the mRNA for the proton-translocating ATPase (*unc*) operon of *Escherichia coli*." J Bacteriol **155**(3): 1279-1287.

References

88. Junge, W., H. Lill and S. Engelbrecht (1997). "ATP synthase: an electrochemical transducer with rotatory mechanics." *Trends Biochem Sci* **22**(11): 420-423.
89. Junge, W., H. Sialaff and S. Engelbrecht (2009). "Torque generation and elastic power transmission in the rotary F₀F₁-ATPase." *Nature* **459**(7245): 364-370.
90. Kabaleeswaran, V., N. Puri, J. E. Walker, A. G. Leslie and D. M. Mueller (2006). "Novel features of the rotary catalytic mechanism revealed in the structure of yeast F₁ ATPase." *EMBO J* **25**(22): 5433-5442.
91. Kabaleeswaran, V., H. Shen, J. Symersky, J. E. Walker, A. G. Leslie and D. M. Mueller (2009). "Asymmetric structure of the yeast F₁ ATPase in the absence of bound nucleotides." *J Biol Chem* **284**(16): 10546-10551.
92. Kanazawa, H., K. Mabuchi, T. Kayano, T. Noumi, T. Sekiya and M. Futai (1981). "Nucleotide sequence of the genes for F₀ components of the proton-translocating ATPase from *Escherichia coli*: prediction of the primary structure of F₀ subunits." *Biochem Biophys Res Commun* **103**(2): 613-620.
93. Kanazawa, H. and M. Futai (1982). "Structure and function of H⁺-ATPase: what we have learned from *Escherichia coli* H⁺-ATPase." *Ann N Y Acad Sci* **402**: 45-64.
94. Kato, Y., T. Matsui, N. Tanaka, E. Muneyuki, T. Hisabori and M. Yoshida (1997). "Thermophilic F₁-ATPase is activated without dissociation of an endogenous inhibitor, epsilon subunit." *J Biol Chem* **272**(40): 24906-24912.
95. Kato-Yamada, Y., D. Bald, M. Koike, K. Motohashi, T. Hisabori and M. Yoshida (1999). "Epsilon subunit, an endogenous inhibitor of bacterial F₁-ATPase, also inhibits F₀F₁-ATPase." *J Biol Chem* **274**(48): 33991-33994.
96. Kato-Yamada, Y. and M. Yoshida (2003). "Isolated epsilon subunit of thermophilic F₁-ATPase binds ATP." *J Biol Chem* **278**(38): 36013-36016.
97. Keis, S., G. Kaim, P. Dimroth and G. M. Cook (2004). "Cloning and molecular characterization of the *atp* operon encoding for the F₁F₀-ATP synthase from a thermoalkaliphilic *Bacillus* sp. strain TA2.A1." *Biochim Biophys Acta* **1676**(1): 112-117.
98. Keis, S., A. Stocker, P. Dimroth and G. M. Cook (2006). "Inhibition of ATP hydrolysis by thermoalkaliphilic F₁F₀-ATP synthase is controlled by the C terminus of the epsilon subunit." *J Bacteriol* **188**(11): 3796-3804.
99. Keller, O., M. Kollmar, M. Stanke and S. Waack (2011). "A novel hybrid gene prediction method employing protein multiple sequence alignments." *Bioinformatics* **27**(6): 757-763.
100. Kibak, H., L. Taiz, T. Starke, P. Bernasconi and J. P. Gogarten (1992). "Evolution of structure and function of V-ATPases." *J Bioenerg Biomembr* **24**(4): 415-424.
101. Klinger, C., M. Rossbach, R. Howe and M. Kaufmann (2003). "Thermophile-specific proteins: the gene product of *aq_1292* from *Aquifex aeolicus* is an NTPase." *BMC Biochem* **4**: 12.
102. Klionsky, D. J., W. S. Brusilow and R. D. Simoni (1984). "In vivo evidence for the role of the epsilon subunit as an inhibitor of the proton-translocating ATPase of *Escherichia coli*." *J Bacteriol* **160**(3): 1055-1060.
103. Kluge, C. and P. Dimroth (1993). "Specific protection by Na⁺ or Li⁺ of the F₁F₀-ATPase of *Propionigenium modestum* from the reaction with dicyclohexylcarbodiimide." *J Biol Chem* **268**(20): 14557-14560.
104. Kluge, C. and P. Dimroth (1993). "Kinetics of inactivation of the F₁F₀ ATPase of *Propionigenium modestum* by dicyclohexylcarbodiimide in relationship to H⁺ and Na⁺ concentration: probing the binding site for the coupling ions." *Biochemistry* **32**(39): 10378-10386.

105. Koch, H. G., T. Hengelage, C. Neumann-Haefelin, J. MacFarlane, H. K. Hoffschulte, K. L. Schimz, B. Mechler and M. Muller (1999). "In vitro studies with purified components reveal signal recognition particle (SRP) and SecA/SecB as constituents of two independent protein-targeting pathways of *Escherichia coli*." Mol Biol Cell **10**(7): 2163-2173.
106. Kohlstadt, M., K. Dorner, R. Labatzke, C. Koc, R. Hielscher, E. Schiltz, O. Einsle, P. Hellwig and T. Friedrich (2008). "Heterologous production, isolation, characterization and crystallization of a soluble fragment of the NADH:ubiquinone oxidoreductase (complex I) from *Aquifex aeolicus*." Biochemistry.
107. Kol, S., N. Nouwen and A. J. Driessen (2008). "The charge distribution in the cytoplasmic loop of subunit C of the F₁F₀ ATPase is a determinant for YidC targeting." J Biol Chem **283**(15): 9871-9877.
108. Krebstakies, T., B. Zimmermann, P. Graber, K. Altendorf, M. Borsch and J. C. Greie (2005). "Both rotor and stator subunits are necessary for efficient binding of F₁ to F₀ in functionally assembled *Escherichia coli* ATP synthase." J Biol Chem **280**(39): 33338-33345.
109. Lau, W. C. and J. L. Rubinstein (2010). "Structure of intact *Thermus thermophilus* V-ATPase by cryo-EM reveals organization of the membrane-bound V_O motor." Proc Natl Acad Sci U S A **107**(4): 1367-1372.
110. Lau, W. C. and J. L. Rubinstein (2012). "Subnanometre-resolution structure of the intact *Thermus thermophilus* H⁺-driven ATP synthase." Nature **481**(7380): 214-218.
111. Lawrence, R. M., B. Varco-Merth, C. J. Bley, J. J. Chen and P. Fromme (2011). "Recombinant production and purification of the subunit c of chloroplast ATP synthase." Protein Expr Purif **76**(1): 15-24.
112. LeBel, D., G. G. Poirier and A. R. Beaudoin (1978). "A convenient method for the ATPase assay." Anal Biochem **85**(1): 86-89.
113. Lee, L. K., A. G. Stewart, M. Donohoe, R. A. Bernal and D. Stock (2010). "The structure of the peripheral stalk of *Thermus thermophilus* H⁺-ATPase/synthase." Nat Struct Mol Biol **17**(3): 373-378.
114. Lightowers, R. N., S. M. Howitt, L. Hatch, F. Gibson and G. Cox (1988). "The proton pore in the *Escherichia coli* F₀F₁-ATPase: substitution of glutamate by glutamine at position 219 of the alpha-subunit prevents F₀-mediated proton permeability." Biochim Biophys Acta **933**(2): 241-248.
115. Lill, H., F. Hensel, W. Junge and S. Engelbrecht (1996). "Cross-linking of engineered subunit delta to (alpha/beta)₃ in chloroplast F-ATPase." J Biol Chem **271**(51): 32737-32742.
116. Lipmann, F. (1941). "Metabolic generation and utilization of phosphate bond energy." Advances in Enzymology and Related Subjects of Biochemistry **1**: 99-162.
117. Liu, J., D. B. Hicks and T. A. Krulwich (2013). "Roles of atpi and two yidc-type proteins from alkaliphilic *Bacillus pseudofirmus* OF4 in ATP synthase assembly and nonfermentative growth." J Bacteriol **195**(2): 220-230.
118. Lohmann, K. (1929). "On the pyrophosphate fraction in muscle." Naturwissenschaften **17**: 624-625.
119. Lolkema, J. S. and E. J. Boekema (2003). "The A-type ATP synthase subunit K of *Methanopyrus kandleri* is deduced from its sequence to form a monomeric rotor comprising 13 hairpin domains." FEBS Lett **543**(1-3): 47-50.
120. Long, J. C., J. DeLeon-Rangel and S. B. Vik (2002). "Characterization of the first cytoplasmic loop of subunit a of the *Escherichia coli* ATP synthase by surface labeling, cross-linking, and mutagenesis." J Biol Chem **277**(30): 27288-27293.
121. Ludtke, S. J., P. R. Baldwin and W. Chiu (1999). "EMAN: semiautomated software for high-resolution single-particle reconstructions." J Struct Biol **128**(1): 82-97.

References

122. Mandel, M., Y. Moriyama, J. D. Hulmes, Y. C. Pan, H. Nelson and N. Nelson (1988). "cDNA sequence encoding the 16-kDa proteolipid of chromaffin granules implies gene duplication in the evolution of H⁺-ATPases." Proc Natl Acad Sci U S A **85**(15): 5521-5524.
123. Marcia, M. (2010). "Functional and structural characterization of *Aquifex aeolicus* sulfide:quinone oxidoreductase." Ph.D. Theses. Max-Planck-Institute of biophysics, Frankfurt, Germany.
124. Marcia, M., J. D. Langer, D. Parcej, V. Vogel, G. Peng and H. Michel (2010). "Characterizing a monotopic membrane enzyme. Biochemical, enzymatic and crystallization studies on *Aquifex aeolicus* sulfide:quinone oxidoreductase." Biochim Biophys Acta **1798**(11): 2114-2123.
125. Matsui, T. and M. Yoshida (1995). "Expression of the wild-type and the Cys-/Trp-less alpha 3 beta 3 gamma complex of thermophilic F₁-ATPase in *Escherichia coli*." Biochim Biophys Acta **1231**(2): 139-146.
126. Matthies, D., S. Habersack, F. Joos, V. Dotsch, J. Vonck, F. Bernhard and T. Meier (2011). "Cell-free expression and assembly of ATP synthase." J Mol Biol **413**(3): 593-603.
127. McCarthy, J. E., H. U. Schairer and W. Sebald (1985). "Translational initiation frequency of *atp* genes from *Escherichia coli*: identification of an intercistronic sequence that enhances translation." EMBO J **4**(2): 519-526.
128. McCarthy, J. E. (1988). "Expression of the *unc* genes in *Escherichia coli*." J Bioenerg Biomembr **20**(1): 19-39.
129. McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady and M. Rosenberg (1981). "A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase." Gene Amplif Anal **2**: 383-415.
130. McLachlin, D. T., J. A. Bestard and S. D. Dunn (1998). "The b and delta subunits of the *Escherichia coli* ATP synthase interact via residues in their C-terminal regions." J Biol Chem **273**(24): 15162-15168.
131. McMillan, D. G., S. Keis, P. Dimroth and G. M. Cook (2007). "A specific adaptation in the a subunit of thermoalkaliphilic F₁F₀-ATP synthase enables ATP synthesis at high pH but not at neutral pH values." J Biol Chem **282**(24): 17395-17404.
132. Meier, T., U. Matthey, F. Henzen, P. Dimroth and D. J. Muller (2001). "The central plug in the reconstituted undecameric c cylinder of a bacterial ATP synthase consists of phospholipids." FEBS Lett **505**(3): 353-356.
133. Meier, T., C. von Ballmoos, S. Neumann and G. Kaim (2003). "Complete DNA sequence of the *atp* operon of the sodium-dependent F₁F₀ ATP synthase from *Ilyobacter tartaricus* and identification of the encoded subunits." Biochim Biophys Acta **1625**(2): 221-226.
134. Meier, T., P. Polzer, K. Diederichs, W. Welte and P. Dimroth (2005). "Structure of the rotor ring of F-Type Na⁺-ATPase from *Ilyobacter tartaricus*." Science **308**(5722): 659-662.
135. Meier, T., J. Yu, T. Raschle, F. Henzen, P. Dimroth and D. J. Muller (2005). "Structural evidence for a constant c11 ring stoichiometry in the sodium F-ATP synthase." FEBS J **272**(21): 5474-5483.
136. Meier, T., A. Krah, P. J. Bond, D. Pogoryelov, K. Diederichs and J. D. Faraldo-Gomez (2009). "Complete ion-coordination structure in the rotor ring of Na⁺-dependent F-ATP synthases." J Mol Biol **391**(2): 498-507.
137. Menz, R. I., J. E. Walker and A. G. Leslie (2001). "Structure of bovine mitochondrial F₁-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis." Cell **106**(3): 331-341.

138. Miroux, B. and J. E. Walker (1996). "Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." J Mol Biol **260**(3): 289-298.
139. Mitchell, P. (1961). "Chemiosmotic coupling in oxidative and photosynthetic phosphorylation." Biochemical Journal **79**(3): P23-&.
140. Mizutani, K., M. Yamamoto, K. Suzuki, I. Yamato, Y. Kakinuma, M. Shirouzu, J. E. Walker, S. Yokoyama, S. Iwata and T. Murata (2011). "Structure of the rotor ring modified with N,N'-dicyclohexylcarbodiimide of the Na⁺-transporting vacuolar ATPase." Proc Natl Acad Sci U S A **108**(33): 13474-13479.
141. Moller, S., M. D. Croning and R. Apweiler (2001). "Evaluation of methods for the prediction of membrane spanning regions." Bioinformatics **17**(7): 646-653.
142. Murata, T., I. Yamato, Y. Kakinuma, A. G. Leslie and J. E. Walker (2005). "Structure of the rotor of the V-Type Na⁺-ATPase from *Enterococcus hirae*." Science **308**(5722): 654-659.
143. Nelson, N. and L. Taiz (1989). "The evolution of H⁺-ATPases." Trends Biochem Sci **14**(3): 113-116.
144. Nesterenko, M. V., M. Tilley and S. J. Upton (1994). "A simple modification of blums silver stain method allows for 30 minute detection of proteins in polyacrylamide gels." J Biochem Biophys Methods **28**(3): 239-242.
145. Neumann, S., U. Matthey, G. Kaim and P. Dimroth (1998). "Purification and properties of the F₁F_o ATPase of *Ilyobacter tartaricus*, a sodium ion pump." J Bacteriol **180**(13): 3312-3316.
146. Nielsen, J., B. B. Jorgensen, K. V. van Meyenburg and F. G. Hansen (1984). "The promoters of the *atp* operon of *Escherichia coli* K12." Mol Gen Genet **193**(1): 64-71.
147. Nilsson, I. and G. von Heijne (1990). "Fine-tuning the topology of a polytopic membrane protein: role of positively and negatively charged amino acids." Cell **62**(6): 1135-1141.
148. Noji, H., R. Yasuda, M. Yoshida and K. Kinosita (1997). "Direct observation of the rotation of F₁-ATPase." Nature **386**(6622): 299-302.
149. Noji, H., K. Hasler, W. Junge, K. Kinosita, Jr., M. Yoshida and S. Engelbrecht (1999). "Rotation of *Escherichia coli* F₁-ATPase." Biochem Biophys Res Commun **260**(3): 597-599.
150. Nugent, T. and D. T. Jones (2009). "Transmembrane protein topology prediction using support vector machines." BMC Bioinformatics **10**: 159.
151. Obuchi, M., K. Kawahara, D. Motooka, S. Nakamura, M. Yamanaka, T. Takeda, S. Uchiyama, Y. Kobayashi, T. Ohkubo and Y. Sambongi (2009). "Hyperstability and crystal structure of cytochrome c₅₅₅ from hyperthermophilic *Aquifex aeolicus*." Acta Crystallogr D Biol Crystallogr **65**(Pt 8): 804-813.
152. Ogilvie, I., R. Aggeler and R. A. Capaldi (1997). "Cross-linking of the delta subunit to one of the three alpha subunits has no effect on functioning, as expected if delta is a part of the stator that links the F₁ and F₀ parts of the *Escherichia coli* ATP synthase." J Biol Chem **272**(26): 16652-16656.
153. Okuno, D., R. Iino and H. Noji (2011). "Rotation and structure of F₀F₁-ATP synthase." J Biochem **149**(6): 655-664.
154. Ono, S., N. Sone, M. Yoshida and T. Suzuki (2004). "ATP synthase that lacks F₀a-subunit: isolation, properties, and indication of F₀b₂-subunits as an anchor rail of a rotating c-ring." J Biol Chem **279**(32): 33409-33412.

References

155. Oppenheim, D. S. and C. Yanofsky (1980). "Translational coupling during expression of the tryptophan operon of *Escherichia coli*." *Genetics* **95**(4): 785-795.
156. Orriss, G. L., A. G. Leslie, K. Braig and J. E. Walker (1998). "Bovine F₁-ATPase covalently inhibited with 4-chloro-7-nitrobenzofurazan: the structure provides further support for a rotary catalytic mechanism." *Structure* **6**(7): 831-837.
157. Ozaki, Y., T. Suzuki, Y. Kuruma, T. Ueda and M. Yoshida (2008). "UncI protein can mediate ring-assembly of c-subunits of F₀F₁-ATP synthase *in vitro*." *Biochem Biophys Res Commun* **367**(3): 663-666.
158. Peng, G., G. Fritzsche, V. Zickermann, H. Schagger, R. Mentele, F. Lottspeich, M. Bostina, M. Radermacher, R. Huber, K. O. Stetter and H. Michel (2003). "Isolation, characterization and electron microscopic single particle analysis of the NADH:ubiquinone oxidoreductase (complex I) from the hyperthermophilic eubacterium *Aquifex aeolicus*." *Biochemistry* **42**(10): 3032-3039.
159. Peng, G., M. Bostina, M. Radermacher, I. Rais, M. Karas and H. Michel (2006). "Biochemical and electron microscopic characterization of the F₁F₀ ATP synthase from the hyperthermophilic eubacterium *Aquifex aeolicus*." *FEBS Lett* **580**(25): 5934-5940.
160. Perlman, D. and H. O. Halvorson (1983). "A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides." *J Mol Biol* **167**(2): 391-409.
161. Peters, J. M., J. R. Harris, A. Lustig, S. Muller, A. Engel, S. Volker and W. W. Franke (1992). "Ubiquitous soluble Mg²⁺-ATPase complex. A structural study." *J Mol Biol* **223**(2): 557-571.
162. Petersen, T. N., S. Brunak, G. von Heijne and H. Nielsen (2011). "SignalP 4.0: discriminating signal peptides from transmembrane regions." *Nat Methods* **8**(10): 785-786.
163. Pisa, K. Y., H. Huber, M. Thomm and V. Muller (2007). "A sodium ion-dependent A₁A₀ ATP synthase from the hyperthermophilic archaeon *Pyrococcus furiosus*." *FEBS J* **274**(15): 3928-3938.
164. Platt, T. and C. Yanofsky (1975). "An intercistronic region and ribosome-binding site in bacterial messenger RNA." *Proc Natl Acad Sci U S A* **72**(6): 2399-2403.
165. Pogoryelov, D., J. Yu, T. Meier, J. Vonck, P. Dimroth and D. J. Muller (2005). "The c15 ring of the *Spirulina platensis* F-ATP synthase: F₁/F₀ symmetry mismatch is not obligatory." *EMBO Rep* **6**(11): 1040-1044.
166. Pogoryelov, D., Y. Nikolaev, U. Schlattner, K. Pervushin, P. Dimroth and T. Meier (2008). "Probing the rotor subunit interface of the ATP synthase from *Ilyobacter tartaricus*." *FEBS J* **275**(19): 4850-4862.
167. Pogoryelov, D., O. Yildiz, J. D. Faraldo-Gomez and T. Meier (2009). "High-resolution structure of the rotor ring of a proton-dependent ATP synthase." *Nat Struct Mol Biol* **16**(10): 1068-1073.
168. Pogoryelov, D., A. Krah, J. D. Langer, O. Yildiz, J. D. Faraldo-Gomez and T. Meier (2010). "Microscopic rotary mechanism of ion translocation in the F₀ complex of ATP synthases." *Nat Chem Biol* **6**(12): 891-899.
169. Pogoryelov, D., A. L. Klyszejko, G. O. Krasnoselska, E. M. Heller, V. Leone, J. D. Langer, J. Vonck, D. J. Muller, J. D. Faraldo-Gomez and T. Meier (2012). "Engineering rotor ring stoichiometries in the ATP synthase." *Proc Natl Acad Sci U S A* **109**(25): E1599-E1608.
170. Porter, A. C., W. S. Brusilow and R. D. Simoni (1983). "Promoter for the *unc* operon of *Escherichia coli*." *J Bacteriol* **155**(3): 1271-1278.
171. Preiss, L., O. Yildiz, D. B. Hicks, T. A. Krulwich and T. Meier (2010). "A new type of proton coordination in an F₁F₀-ATP synthase rotor ring." *PLoS Biol* **8**(8): e1000443.

172. Price, C. E. and A. J. Driessen (2010). "Biogenesis of membrane bound respiratory complexes in *Escherichia coli*." Biochim Biophys Acta **1803**(6): 748-766.
173. Rees, D. M., A. G. Leslie and J. E. Walker (2009). "The structure of the membrane extrinsic region of bovine ATP synthase." Proc Natl Acad Sci U S A **106**(51): 21597-21601.
174. Rodgers, A. J. and R. A. Capaldi (1998). "The second stalk composed of the b- and delta-subunits connects F₀ to F₁ via an alpha-subunit in the *Escherichia coli* ATP synthase." J Biol Chem **273**(45): 29406-29410.
175. Rodgers, A. J. and M. C. Wilce (2000). "Structure of the gamma-epsilon complex of ATP synthase." Nat Struct Biol **7**(11): 1051-1054.
176. Rojo, E. E., R. A. Stuart and W. Neupert (1995). "Conservative sorting of F₀-ATPase subunit 9: export from matrix requires delta pH across inner membrane and matrix ATP." EMBO J **14**(14): 3445-3451.
177. Rollauer, S. E., M. J. Tarry, J. E. Graham, M. Jaaskelainen, F. Jager, S. Johnson, M. Krehenbrink, S. M. Liu, M. J. Lukey, J. Marcoux, M. A. McDowell, F. Rodriguez, P. Roversi, P. J. Stansfeld, C. V. Robinson, M. S. Sansom, T. Palmer, M. Hogbom, B. C. Berks and S. M. Lea (2012). "Structure of the TatC core of the twin-arginine protein transport system." Nature **492**(7428): 210-214.
178. Rothschild, L. J. and R. L. Mancinelli (2001). "Life in extreme environments." Nature **409**(6823): 1092-1101.
179. Runswick, M. J. and J. E. Walker (1983). "The amino acid sequence of the beta-subunit of ATP synthase from bovine heart mitochondria." J Biol Chem **258**(5): 3081-3089.
180. Russell, R. J., J. M. Ferguson, D. W. Hough, M. J. Danson and G. L. Taylor (1997). "The crystal structure of citrate synthase from the hyperthermophilic archaeon *pyrococcus furiosus* at 1.9 Å resolution." Biochemistry **36**(33): 9983-9994.
181. Sabbert, D., S. Engelbrecht and W. Junge (1997). "Functional and idling rotatory motion within F₁-ATPase." Proc Natl Acad Sci U S A **94**(9): 4401-4405.
182. Saller, M. J., F. Fusetti and A. J. Driessen (2009). "*Bacillus subtilis* SpoIIJ and YqjG function in membrane protein biogenesis." J Bacteriol **191**(21): 6749-6757.
183. Samuelson, J. C., M. Chen, F. Jiang, I. Moller, M. Wiedmann, A. Kuhn, G. J. Phillips and R. E. Dalbey (2000). "YidC mediates membrane protein insertion in bacteria." Nature **406**(6796): 637-641.
184. Santana, M., M. S. Ionescu, A. Vertes, R. Longin, F. Kunst, A. Danchin and P. Glaser (1994). "*Bacillus subtilis* F₀F₁ ATPase: DNA sequence of the atp operon and characterization of *atp* mutants." J Bacteriol **176**(22): 6802-6811.
185. Schaffer, A. A., L. Aravind, T. L. Madden, S. Shavirin, J. L. Spouge, Y. I. Wolf, E. V. Koonin and S. F. Altschul (2001). "Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements." Nucleic Acids Res **29**(14): 2994-3005.
186. Schagger, H. and G. von Jagow (1987). "Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa." Anal Biochem **166**(2): 368-379.
187. Schatz, G. and R. A. Butow (1983). "How are proteins imported into mitochondria." Cell **32**(2): 316-318.
188. Schramm, H. C., B. Schneppe, R. Birkenhager and J. E. McCarthy (1996). "The promoter-proximal, unstable IB region of the atp mRNA of *Escherichia coli*: an independently degraded region that can act as a destabilizing element." Biochim Biophys Acta **1307**(2): 162-170.

References

189. Schutz, M., B. Schoepp-Cothenet, E. Lojou, M. Woodstra, D. Lexa, P. Tron, A. Dolla, M. C. Durand, K. O. Stetter and F. Baymann (2003). "The naphthoquinol oxidizing cytochrome bc₁ complex of the hyperthermophilic knallgasbacterium *Aquifex aeolicus*: properties and phylogenetic relationships." Biochemistry **42**(36): 10800-10808.
190. Schwem, B. E. and R. H. Fillingame (2006). "Cross-linking between helices within subunit a of *Escherichia coli* ATP synthase defines the transmembrane packing of a four-helix bundle." J Biol Chem **281**(49): 37861-37867.
191. Senior, A. E. (1988). "ATP synthesis by oxidative phosphorylation." Physiol Rev **68**(1): 177-231.
192. Senior, A. E. (2007). "ATP synthase: motoring to the finish line." Cell **130**(2): 220-221.
193. Shimabukuro, K., R. Yasuda, E. Muneyuki, K. Y. Hara, K. Kinoshita, Jr. and M. Yoshida (2003). "Catalysis and rotation of F₁ motor: cleavage of ATP at the catalytic site occurs in 1 ms before 40 degree substep rotation." Proc Natl Acad Sci U S A **100**(25): 14731-14736.
194. Shirakihara, Y., A. G. Leslie, J. P. Abrahams, J. E. Walker, T. Ueda, Y. Sekimoto, M. Kambara, K. Saika, Y. Kagawa and M. Yoshida (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 **5**(6): 825-836.
195. Singh, S., P. Turina, C. J. Bustamante, D. J. Keller and R. Capaldi (1996). "Topographical structure of membrane-bound *Escherichia coli* F₁F₀ ATP synthase in aqueous buffer." FEBS Lett **397**(1): 30-34.
196. Stark, M. J. (1987). "Multicopy expression vectors carrying the lac repressor gene for regulated high-level expression of genes in *Escherichia coli*." Gene **51**(2-3): 255-267.
197. Steed, P. R. and R. H. Fillingame (2009). "Aqueous accessibility to the transmembrane regions of subunit c of the *Escherichia coli* F₁F₀ ATP synthase." J Biol Chem **284**(35): 23243-23250.
198. Steffens, K., E. Schneider, B. Herkenhoff, R. Schmid and K. Altendorf (1984). "Chemical modification of the F₀ part of the ATP synthase F₁F₀ from *Escherichia coli*. Effects on proton conduction and F₁ binding." Eur J Biochem **138**(3): 617-622.
199. Stewart, A. G., L. K. Lee, M. Donohoe, J. J. Chaston and D. Stock (2012). "The dynamic stator stalk of rotary ATPases." Nat Commun **3**: 687.
200. Stock, D., A. G. Leslie and J. E. Walker (1999). "Molecular architecture of the rotary motor in ATP synthase." Science **286**(5445): 1700-1705.
201. Stock, D., C. Gibbons, I. Arechaga, A. G. Leslie and J. E. Walker (2000). "The rotary mechanism of ATP synthase." Curr Opin Struct Biol **10**(6): 672-679.
202. Stocker, A., S. Keis, J. Vonck, G. M. Cook and P. Dimroth (2007). "The structural basis for unidirectional rotation of thermoalkaliphilic F₁-ATPase." Structure **15**(8): 904-914.
203. Stuart, R. (2002). "Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex." Biochim Biophys Acta **1592**(1): 79-87.
204. Stuart, R. A. and W. Neupert (1996). "Topogenesis of inner membrane proteins of mitochondria." Trends Biochem Sci **21**(7): 261-267.
205. Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol **189**(1): 113-130.
206. Surade, S., M. Klein, P. C. Stolt-Bergner, C. Muenke, A. Roy and H. Michel (2006). "Comparative analysis and "expression space" coverage of the production of prokaryotic membrane proteins for structural genomics." Protein Sci **15**(9): 2178-2189.

207. Surade, S. (2007). "Structural genomics on prokaryotic membrane proteins." Ph.D. Theses. Max-Planck-Institute of biophysics, Frankfurt, Germany.
208. Suzuki, T., T. Murakami, R. Iino, J. Suzuki, S. Ono, Y. Shirakihara and M. Yoshida (2003). "F₀F₁-ATPase/synthase is geared to the synthesis mode by conformational rearrangement of epsilon subunit in response to proton motive force and ADP/ATP balance." J Biol Chem **278**(47): 46840-46846.
209. Suzuki, T., Y. Ozaki, N. Sone, B. A. Feniouk and M. Yoshida (2007). "The product of uncl gene in F₁F₀-ATP synthase operon plays a chaperone-like role to assist c-ring assembly." Proc Natl Acad Sci U S A **104**(52): 20776-20781.
210. Symersky, J., V. Pagadala, D. Osowski, A. Krah, T. Meier, J. D. Faraldo-Gomez and D. M. Mueller (2012). "Structure of the c₁₀ ring of the yeast mitochondrial ATP synthase in the open conformation." Nat Struct Mol Biol.
211. Takeyasu, K., H. Omote, S. Nettikadan, F. Tokumasu, A. Iwamoto-Kihara and M. Futai (1996). "Molecular imaging of *Escherichia coli* F₀F₁-ATPase in reconstituted membranes using atomic force microscopy." FEBS Lett **392**(2): 110-113.
212. Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead and C. M. Fraser (2001). "Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*." Science **293**(5529): 498-506.
213. Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." Nucleic Acids Res **25**(24): 4876-4882.
214. Tolia, N. H. and L. Joshua-Tor (2006). "Strategies for protein coexpression in *Escherichia coli*." Nat Methods **3**(1): 55-64.
215. Tsunoda, S. P., R. Aggeler, M. Yoshida and R. A. Capaldi (2001). "Rotation of the c subunit oligomer in fully functional F₁F₀ ATP synthase." Proc Natl Acad Sci U S A **98**(3): 898-902.
216. Tsunoda, S. P., A. J. Rodgers, R. Aggeler, M. C. Wilce, M. Yoshida and R. A. Capaldi (2001). "Large conformational changes of the epsilon subunit in the bacterial F₁F₀ ATP synthase provide a ratchet action to regulate this rotary motor enzyme." Proc Natl Acad Sci U S A **98**(12): 6560-6564.
217. Uhlin, U., G. B. Cox and J. M. Guss (1997). "Crystal structure of the epsilon subunit of the proton-translocating ATP synthase from *Escherichia coli*." Structure **5**(9): 1219-1230.
218. Ulbrandt, N. D., J. A. Newitt and H. D. Bernstein (1997). "The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins." Cell **88**(2): 187-196.
219. Valent, Q. A., P. A. Scotti, S. High, J. W. de Gier, G. von Heijne, G. Lentzen, W. Wintermeyer, B. Oudega and J. Luirink (1998). "The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon." EMBO J **17**(9): 2504-2512.
220. van Bloois, E., G. Jan Haan, J. W. de Gier, B. Oudega and J. Luirink (2004). "F₁F₀ ATP synthase subunit c is targeted by the SRP to YidC in the *E. coli* inner membrane." FEBS Lett **576**(1-2): 97-100.
221. van der Laan, M., P. Bechtluft, S. Kol, N. Nouwen and A. J. Driessen (2004). "F₁F₀ ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis." J Cell Biol **165**(2): 213-222.

References

222. van Heel, M., G. Harauz, E. V. Orlova, R. Schmidt and M. Schatz (1996). "A new generation of the IMAGIC image processing system." J. Struct. Biol. **116**: 17-24.
223. van Raaij, M. J., J. P. Abrahams, A. G. Leslie and J. E. Walker (1996). "The structure of bovine F₁-ATPase complexed with the antibiotic inhibitor aurovertin B." Proc Natl Acad Sci U S A **93**(14): 6913-6917.
224. Velours, J., G. Arselin de Chateaubodeau, M. Galante and B. Guerin (1987). "Subunit 4 of ATP synthase F₀F₁ from yeast mitochondria. Purification, amino-acid composition and partial N-terminal sequence." Eur J Biochem **164**(3): 579-584.
225. Viebrock, A., A. Perz and W. Sebald (1982). "The imported pre-protein of the proteolipid subunit of the mitochondrial ATP synthase from *Neurospora crassa* - molecular-cloning and sequencing of the messenger-RNA." Embo Journal **1**(5): 565-571.
226. Vik, S. B. and R. D. Simoni (1987). "F₁F₀-ATPase from *Escherichia coli* with mutant F₀ subunits. Partial purification and immunoprecipitation of F₁F₀ complexes." J Biol Chem **262**(17): 8340-8346.
227. Vik, S. B. and B. J. Antonio (1994). "A mechanism of proton translocation by F₁F₀ ATP synthases suggested by double mutants of the a subunit." J Biol Chem **269**(48): 30364-30369.
228. Vik, S. B., J. C. Long, T. Wada and D. Zhang (2000). "A model for the structure of subunit a of the *Escherichia coli* ATP synthase and its role in proton translocation." Biochim Biophys Acta **1458**(2-3): 457-466.
229. Vives-Bauza, C., J. Magrane, A. L. Andreu and G. Manfredi (2010). "Novel role of ATPase subunit c targeting peptides beyond mitochondrial protein import." Mol Biol Cell **21**(1): 131-139.
230. Voet, D. and J. G. Voet (2004). Electron transport and oxidative phosphorylation. Biochemistry. P. F. David Harris, John Wiley & Sons: 797-842.
231. von Ballmoos, C., Y. Appoldt, J. Brunner, T. Granier, A. Vasella and P. Dimroth (2002). "Membrane topography of the coupling ion binding site in Na⁺-translocating F₁F₀ ATP synthase." J Biol Chem **277**(5): 3504-3510.
232. von Ballmoos, C., G. M. Cook and P. Dimroth (2008). "Unique rotary ATP synthase and its biological diversity." Annu Rev Biophys **37**: 43-64.
233. von Ballmoos, C., A. Wiedenmann and P. Dimroth (2009). "Essentials for ATP synthesis by F₁F₀ ATP synthases." Annu Rev Biochem **78**: 649-672.
234. von Heijne, G. (1983). "Patterns of amino acids near signal-sequence cleavage sites." Eur J Biochem **133**(1): 17-21.
235. von Heijne, G. (1985). "Signal sequences - the limits of variation." J Mol Biol **184**(1): 99-105.
236. von Heijne, G. (1986). "A new method for predicting signal sequence cleavage sites." Nucleic Acids Res **14**(11): 4683-4690.
237. von Heijne, G. and Y. Gavel (1988). "Topogenic signals in integral membrane proteins." Eur J Biochem **174**(4): 671-678.
238. von Meyenburg, K., B. B. Jorgensen, J. Nielsen and F. G. Hansen (1982). "Promoters of the *atp* operon coding for the membrane-bound ATP synthase of *Escherichia coli* mapped by Tn10 insertion mutations." Mol Gen Genet **188**(2): 240-248.
239. von Meyenburg, K., B. B. Jorgensen, O. Michelsen, L. Sorensen and J. E. McCarthy (1985). "Proton conduction by subunit a of the membrane-bound ATP synthase of *Escherichia coli* revealed after induced overproduction." EMBO J **4**(9): 2357-2363.

240. Walker, J. E., M. Saraste and N. J. Gay (1982). "*E. coli* F₁-ATPase interacts with a membrane protein component of a proton channel." Nature **298**(5877): 867-869.
241. Walker, J. E., M. Saraste, M. J. Runswick and N. J. Gay (1982). "Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." EMBO J **1**(8): 945-951.
242. Walker, J. E., M. Saraste and N. J. Gay (1984). "The *unc* operon. Nucleotide sequence, regulation and structure of ATP-synthase." Biochim Biophys Acta **768**(2): 164-200.
243. Walker, J. E., M. J. Runswick and L. Poulter (1987). "ATP synthase from bovine mitochondria. The characterization and sequence analysis of two membrane-associated sub-units and of the corresponding cDNAs." J Mol Biol **197**(1): 89-100.
244. Walker, J. E. and V. K. Dickson (2006). "The peripheral stalk of the mitochondrial ATP synthase." Biochim Biophys Acta **1757**(5-6): 286-296.
245. Wang, Z. G. and S. H. Ackerman (2000). "The assembly factor Atp11p binds to the beta-subunit of the mitochondrial F₁-ATPase." J Biol Chem **275**(8): 5767-5772.
246. Wang, Z. G., D. Sheluho, D. L. Gatti and S. H. Ackerman (2000). "The alpha-subunit of the mitochondrial F₁ ATPase interacts directly with the assembly factor Atp12p." EMBO J **19**(7): 1486-1493.
247. Waterhouse, A. M., J. B. Procter, D. M. Martin, M. Clamp and G. J. Barton (2009). "Jalview Version 2--a multiple sequence alignment editor and analysis workbench." Bioinformatics **25**(9): 1189-1191.
248. Watt, I. N., M. G. Montgomery, M. J. Runswick, A. G. W. Leslie and J. E. Walker (2010). "Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria." Proc Natl Acad Sci U S A **107**(39): 16823-16827.
249. Weber, J. and A. E. Senior (2003). "ATP synthesis driven by proton transport in F₁F₀-ATP synthase." FEBS Lett **545**(1): 61-70.
250. Westheimer, F. H. (1987). "Why nature chose phosphates." Science **235**(4793): 1173-1178.
251. Wilkens, S., F. W. Dahlquist, L. P. McIntosh, L. W. Donaldson and R. A. Capaldi (1995). "Structural features of the epsilon subunit of the *Escherichia coli* ATP synthase determined by NMR spectroscopy." Nat Struct Biol **2**(11): 961-967.
252. Wilkens, S., S. D. Dunn, J. Chandler, F. W. Dahlquist and R. A. Capaldi (1997). "Solution structure of the N-terminal domain of the delta subunit of the *E. coli* ATP synthase." Nat Struct Biol **4**(3): 198-201.
253. Wilkens, S., A. Rodgers, I. Ogilvie and R. A. Capaldi (1997). "Structure and arrangement of the delta subunit in the *E. coli* ATP synthase (ECF₁F₀)." Biophys Chem **68**(1-3): 95-102.
254. Wilkens, S., J. Zhou, R. Nakayama, S. D. Dunn and R. A. Capaldi (2000). "Localization of the delta subunit in the *Escherichia coli* F₁F₀-ATP synthase by immuno electron microscopy: the delta subunit binds on top of the F₁." J Mol Biol **295**(3): 387-391.
255. Wilkens, S., D. Borchardt, J. Weber and A. E. Senior (2005). "Structural characterization of the interaction of the delta and alpha subunits of the *Escherichia coli* F₁F₀-ATP synthase by NMR spectroscopy." Biochemistry **44**(35): 11786-11794.
256. Woodcock, D. M., P. J. Crowther, J. Doherty, S. Jefferson, E. DeCruz, M. Noyer-Weidner, S. S. Smith, M. Z. Michael and M. W. Graham (1989). "Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants." Nucleic Acids Res **17**(9): 3469-3478.

References

257. Yagi, H., N. Kajiwara, H. Tanaka, T. Tsukihara, Y. Kato-Yamada, M. Yoshida and H. Akutsu (2007). "Structures of the thermophilic F₁-ATPase epsilon subunit suggesting ATP-regulated arm motion of its C-terminal domain in F₁." Proc Natl Acad Sci U S A **104**(27): 11233-11238.
258. Yamamoto, T., T. Tamura, T. Yokota and T. Takano (1982). "Overlapping genes in the heat-labile enterotoxin operon originating from *Escherichia coli* human strain." Mol Gen Genet **188**(2): 356-359.
259. Yan, W. L., T. J. Lerner, J. L. Haines and J. F. Gusella (1994). "Sequence analysis and mapping of a novel human mitochondrial ATP synthase subunit 9 cDNA (ATP5G3)." Genomics **24**(2): 375-377.
260. Yasuda, R., H. Noji, M. Yoshida, K. Kinoshita, Jr. and H. Itoh (2001). "Resolution of distinct rotational substeps by submillisecond kinetic analysis of F₁-ATPase." Nature **410**(6831): 898-904.
261. Yi, L., N. Celebi, M. Y. Chen and R. E. Dalbey (2004). "Sec/SRP requirements and energetics of membrane insertion of subunits a, b, and c of the *Escherichia coli* F₁F₀ ATP synthase." Journal of Biological Chemistry **279**(38): 39260-39267.
262. Yokoyama, K., S. Ohkuma, H. Taguchi, T. Yasunaga, T. Wakabayashi and M. Yoshida (2000). "V-Type H⁺-ATPase/synthase from a thermophilic eubacterium, *Thermus thermophilus*. Subunit structure and operon." J Biol Chem **275**(18): 13955-13961.
263. Yokoyama, K., K. Nagata, H. Imamura, S. Ohkuma, M. Yoshida and M. Tamakoshi (2003). "Subunit arrangement in V-ATPase from *Thermus thermophilus*." J Biol Chem **278**(43): 42686-42691.
264. Zhang, D. and S. B. Vik (2003). "Helix packing in subunit a of the *Escherichia coli* ATP synthase as determined by chemical labeling and proteolysis of the cysteine-substituted protein." Biochemistry **42**(2): 331-337.

Appendix

A.1. Bioinformatics

Table A.1: Amino acid compositions of the *A. aeolicus* F₁F₀ ATP synthase deduced from the respective *atp* genes

Amino acid	a	c	c ^a	b ₁	b ₂	δ	α	γ	β	ε
Alanine	22	18	13	14	21	11	48	28	31	9
Arginine	6	4	3	8	5	9	28	21	25	2
Asparagine	9	2	2	4	7	5	13	20	10	4
Aspartic Acid	-	-	-	3	1	8	27	12	22	3
Cysteine	-	-	-	-	-	-	2	-	-	-
Glutamic Acid	10	4	4	25	29	18	47	32	48	21
Glutamine	2	2	2	9	7	4	23	9	20	4
Glycine	14	16	16	1	8	7	43	11	45	9
Histidine	6	1	1	-	3	-	6	3	9	1
Isoleucine	19	9	6	18	14	14	44	17	37	5
Leucine	34	16	14	11	21	29	50	22	36	14
Lysine	9	2	1	17	28	28	32	23	29	12
Methionine	7	7	2	3	3	4	8	6	14	6
Phenylalanine	20	4	4	3	8	6	14	14	18	2
Proline	11	2	1	2	2	6	24	5	25	6
Serine	7	1	1	5	9	7	16	9	13	4
Threonine	7	4	3	9	5	7	19	15	24	7
Tryptophan	-	-	-	-	1	-	-	1	4	1
Tyrosine	10	2	2	3	7	4	21	13	16	6
Valine	23	6	6	9	6	14	38	30	52	16
Total	216	100	81	144	185	181	558	399	503	132

^a: mature subunit c without the signal peptide

Table A.2: List of the c-subunits used for multiple-sequence alignments

Group	Label	Organism	Accession number
1	>2 ECOLI	<i>Escherichia coli</i> K-12	gi 56404993
	>3 ILYTA	<i>Ilyobacter tartaricus</i>	gi 75526948
	>7 PAVLU	<i>chloroplast Pavlova lutheri</i>	gi 114650
	>9 GEOBB	<i>Geobacter bemijsiensis</i> Bem	gi 224487646
	>10 FUSNN	<i>Fusobacterium nucleatum</i> subsp. <i>nucleatum</i> ATCC 25586	gi 81763577
	>18 THIFE	<i>Acidithiobacillus ferrooxidans</i>	gi 728934
	>19 MYCPE	<i>Mycoplasma penetrans</i> HF-2	gi 81748107
	>23 DEHE1	<i>Dehalococcoides ethenogenes</i> 195	gi 123732469
	>24 RHOS1	<i>Rhodobacter sphaeroides</i> ATCC 17029	gi 224487618
	>26 BIFAA	<i>Bifidobacterium adolescentis</i> ATCC 15703	gi 224487626
	>28 ENTHR	<i>Enterococcus hirae</i> ATCC 9790	gi 114669
	>29 GEOSE	<i>Geobacillus stearothermophilus</i>	gi 1168601
	>30 MOOTA	<i>Moorella thermoacetica</i> ATCC 39073	gi 123739208
	>33 DESOH	<i>Desulfococcus oleovorans</i> Hxd3	gi 224487660
	>35 CYBMR	mitochondrion <i>Cyberlindnera mrakii</i>	gi 3121815
	>36 CLONN	<i>Clostridium novyi</i> NT	gi 224487643
	>39 STRPN	<i>Streptococcus pneumoniae</i> TIGR4	gi 61219626
	>40 BUCAI	<i>Buchnera aphidicola</i> str. <i>APS (Acyrtosiphon pisum)</i>	gi 11131203
	>41 STRCO	<i>Streptomyces coelicolor</i> A3(2)	gi 61219624
	>42 DESPS	<i>Desulfotalea psychrophila</i> LSv54	gi 81692932
	>43 CLOAB	<i>Clostridium acetobutylicum</i> ATCC 824	gi 5915735
	>45 BETVU	mitochondrion <i>Beta vulgaris</i>	gi 114496
	>46 GLUOX	<i>Gluconobacter oxydans</i> 621H	gi 81352056
>47 VIBPA	<i>Vibrio parahaemolyticus</i> RIMD 2210633	gi 60391832	
>48 SACDO	mitochondrion <i>Saccharomyces douglasii</i>	gi 48428794	
>53 BOVIN	<i>Bos taurus</i> (cattle)	PDB:2XND	
2	>1 AQUAE	<i>Aquifex aeolicus</i> VF5	gi 3913149
	>6 MYCH2	<i>Mycoplasma hyopneumoniae</i> 232	gi 81378799
	>8 MYCPN	<i>Mycoplasma pneumoniae</i> M129	gi 2493074
	>14 CAMJD	<i>Campylobacter jejuni</i> subsp. <i>doylei</i> 269.97	gi 224487634
	>15 PARUW	<i>Candidatus Protochlamydia amoebophila</i> UWE25	gi 81697604
	>16 UREU1	<i>Ureaplasma urealyticum</i> serovar 10 str. ATCC 33699	gi 224487686
	>25 MYCS5	<i>Mycoplasma synoviae</i> 53	gi 224487717
	>27 MYCMS	<i>Mycoplasma mycoides</i> subsp. <i>mycoides</i> SC str. PG1	gi 81697959
	>31 DESDG	<i>Desulfovibrio alaskensis</i> G20	gi 224487719
	>32 MYCMO	<i>Mycoplasma mobile</i> 163K	gi 81614341
	>34 MESFL	<i>Mesoplasma florum</i> L1	gi 81695704
	>37 SORC5	<i>Sorangium cellulosum</i> So ce56	gi 224487664
	>38 MYCGA	<i>Mycoplasma gallisepticum</i> str. R(low)	gi 33860136
	>44 SYNAS	<i>Syntrophus aciditrophicus</i> SB	gi 224487712
3	>4 MOUSE	<i>Mus musculus</i> (house mouse)	gi 51338784
	>5 NEUCR	<i>Neurospora crassa</i> OR74A	gi 114671
	>17 MANSE	<i>Manduca sexta</i> (tobacco hornworm)	gi 12585194
	>49 BOVIN1	<i>Bos taurus</i> (cattle)	gi 416684
	>51 BOVIN2	<i>Bos taurus</i> (cattle)	gi 114680
	>52 BOVIN3	<i>Bos taurus</i> (cattle)	gi 109940311
4	>11 DICDI	<i>Dictyostelium discoideum</i>	gi 1718094
	>12 CANTR	<i>Candida tropicalis</i>	gi 2493143
	>13 ENTHI	<i>Entamoeba histolytica</i>	gi 3915253
	>20 ARATH	<i>Arabidopsis thaliana</i> (thale cress)	gi 27923954
	>21 CAEEL	<i>Caenorhabditis elegans</i>	gi 3334407
	>22 NEUCR	<i>Neurospora crassa</i> OR74A	gi 74626388
	>50 BOVIN_V	<i>Bos taurus</i> (cattle)	gi 137477

1 gii13913127 ATPA_AQUAE/1-503	1-----MATLTYE---EALILRQIKDFEP EAKMEEVGVVYVGDVA40
gii1175143 ATPA_ECOLI/1-513	1-----MQLNST---EISELIKQIAQFNVSSEAHNCTIVSVSDVI139
27 gii22266797 ATPA_ILYOB/1-469	1-----MKIRPE---EISSILKTEIENYKSLDIKTSQSVVEGCDIA39
5 gii114402 ATPA_BOVIN/1-553	1-----EISSILKTEIENYKSLDIKTSQSVVEGCDIA39
4 gii134190649 ATPA_YEAST/1-545	1-----MLAR---TAAIRLSRSLINSTKAARPAALASTRRLASTKAQPT---EVSLLERIKGVSEANLNLEGRVLAJGDIA76
20 gii32172455 VATA_THET8/1-578	1-----MIQVGIQKLAGPAVIAKGMGLARMYDICKVGEGLVGEIIRLDGD---TAFVQVYEDTSLGKLVGPEVVSGLPLAVLELPG78
22 gii74053565 VATA_SULAC/1-592	1-----MAGEGRVVRVNGPILVADGMRNAQMFVEVVEVGLRVLVEITRIEED---RAYIQVYEAATDGIKPGKEAYRTGSLLSLVELEPG79
23 gii12229704 VATA_HALSA/1-585	1-----MSQAELITDTEIETSVSGPVTATG-LDAQMNDVYVVGDEGLMGEVIEIEED---VTTIQVYEEETSGIPGQPVNTGEPITLDLPG84
25 gii124053334 VATA_BOVIN/1-617	1-----MDFSS---KLPKIRDFE-KESTFGYVHGSGPVVTAIDMAGAAMYELVRVGHSELVGEIIRLEGD---MATTIQVYEEETSGVSGDPLRIGKPLVEPG93
gii1390405181 VATA_YEAST/1-617	1MAGAIENARKIEKRISLEDHAESEVGIATVSVSGPVVTAIDMAGAAMYELVRVGHSELVGEIIRLEGD---XATIQVYEEETSGVSGDPLRIGKPLVEPG100
1 gii13913127 ATPA_AQUAE/1-503	41RAYGLENVMAMEIYEFGGQQ---GIAFLNLEDNVGIIILGSETGIEEGHIVKRTGRIIDAPVGEGLVGRVIDPLGNPLDGGKPIQFEY---126
gii1175143 ATPA_ECOLI/1-513	40RIHGLADCMQGMISLPGNRY-----GIALNLERDSVGVAVMGPYADLAEMKVKCTGRILEVPPVGRGLGRVNTLGPALDGGKPLDHDG---125
27 gii22266797 ATPA_ILYOB/1-469	80RIYGLSNAKAGEIYELFPNCIK-----AMLNLENNVGAVLLGDP TCVKEGDEVVRATGEVAAPAGCGLGRVNLGEPIDGGKDAKIEK---125
5 gii114402 ATPA_BOVIN/1-553	83RVHGLRNVQAEEEMVEFSSGLGLE---VDRKTPTE---GMSLNLLEPNMGVYVFGDKLIEKEDIVIKRTGAIYDVPVGEELGRVDALGNALDGGKPIEISKA---168
4 gii134190649 ATPA_YEAST/1-545	77RVFGLNLIQAEELVEFSSGVK---GMLNLEPGQGVIVLFGSDRLVKEGELVKRTGNIYDVPVGPGLGRVDALGNALDGGKPIEISKA---162
20 gii32172455 VATA_THET8/1-578	79MLNGLYDGIQRP LERIREKTG-IYITRGVVHVALDREKKWATPMV--KPGDEVRGGMVLGTVPEFG-FT-HKILVPPDVRGRVKEVKAAGEYVEEPPVVV176
22 gii74053565 VATA_SULAC/1-592	80LMCGIFDGLQRP LDRIAESVKSPPFVTRGVKVPALERNKWHYIPVA--KKGDKVSPDGIKAVNETDLIE-HRIIVPPNVHGTLEKESPDYEDVEDVIARVD179
23 gii12229704 VATA_HALSA/1-585	85MLDSIYDGVQRP LQVLEDEMG-AFLDRGVADPGLDIDTWFEPFV--EAGDEVAAAGDVGTVDETVSIE-HKVLPVPRSDGGEVVAESGTFVVDVDTLE183
25 gii124053334 VATA_BOVIN/1-617	94IMGAIFDGIQRP LQVLSQTSQSYVIRPGVNVSA LSRDVKWDFPCKNLRVGSIIIQDVIYVNEENSLIK-HKIMPEPRNKGTVTYEAPPNAYDTSVVLLE195
gii1390405181 VATA_YEAST/1-617	101LMEITDGIQRP LKAIKEESQSYVIRPGIDTALDRTIKWQFPFGK-FQVGDHISGQDIYGSVFENSLISHKILPPRSRGTITWIIAPAGEYTLDEKILEVE202
1 gii13913127 ATPA_AQUAE/1-503	127-----RSPVEKIPGVVKKRPVHEFLQTCIKAIIDAMIIGRQRELLIGDRATCKTVAIDTILAQK---NSB---VYCTIVYAVGQKRAA IARLI210
gii1175143 ATPA_ECOLI/1-513	126-----FSAVEAIAPGIEROSDDPVTGYAVDSMIIGRQRELLIGDRATCKTVAIDTILNQR---DSC---KCIYVAIQKASTISNW209
27 gii22266797 ATPA_ILYOB/1-469	169-----RRRVGLKAPCIPRI SVREPQGTIKAVDSLVP IGRQRELLIGDRATCKTVAIDTILNQRKFNDCDEKKLICYVAIQKRSVTAQVL260
5 gii114402 ATPA_BOVIN/1-553	163-----RRAQVKAPCIPRRSVHEPQGTGLKAVDVLV IGRQRELLIGDRATCKTVAIDTILNQRKFNDCDEKKLICYVAIQKRSVTAQVL254
4 gii134190649 ATPA_YEAST/1-545	177DG---TELKMYHTWVRRARVQRKLDNPTFLTGMRILDLVLFVAMGTAAIPGPGFGSKTQVQSLAKWSN---AD---VVVYVCCGERNEMTDV1266
20 gii32172455 VATA_THET8/1-578	180MGDQVLEKLYQRWVRIIPRFKLEKLEPTELLCTRVDVTFIIRAKGTAAPGPGFGSKTQVQSLAKWSN---AK---VVVYVCCGERNEMTDV1272
22 gii74053565 VATA_SULAC/1-592	184TG---EIQMHQWVQRQRPVTDKQTPTEFLVSGQRILDLGFLIRAKGTAAPGPGFGSKTQVQSLAKFAD---AB---IVVYVCCGERNEMTDV1273
23 gii12229704 VATA_HALSA/1-585	196FGKIEKFSMVQVWVPRVQRVPTKLANPHLLTQQRVLDALFCVQGTCTTAPAFGCKTVISQSLSKYSN---SD---VITYVCCGERNEMTDV1288
25 gii124053334 VATA_BOVIN/1-617	205FDGKSDFTLYHTWVPRVPRVTEKLSADYLLTQQRVLDALFCVQGTCTTAPAFGCKTVISQSLSKYSN---SD---VITYVCCGERNEMTDV1295
1 gii13913127 ATPA_AQUAE/1-503	211ELLERE-----GAMEYTVVVASASDASLQYLAPFVCTIIEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK304
gii1175143 ATPA_ECOLI/1-513	210RKLLEH-----GALANIIVVVAESAASLQYLAPFVCTIIEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK303
27 gii22266797 ATPA_ILYOB/1-469	210KLENA-----GAMEYTVVVAATASAPALQYMAPYCVSMECEYEMDKHEVLLIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK303
5 gii114402 ATPA_BOVIN/1-553	261KR LTA-----DAMKYSIIVVVAATASDAAPALQYMAPYCVSMECEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK354
4 gii134190649 ATPA_YEAST/1-545	252QTELEH-----DAMKYSIIVVVAATASDAAPALQYMAPYCVSMECEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK348
20 gii32172455 VATA_THET8/1-578	267VEFPLETPDKTGG---GPLMHRVILIANISNMPVAAREASIVYDVTIAEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK367
22 gii74053565 VATA_SULAC/1-592	273RQPLKLDPWTE---KPLQRIILWANTSNMPVAAREASIVYDVTIAEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK373
23 gii12229704 VATA_HALSA/1-585	274EDFPELPDQGT---NPLMARITLIANTSNMPVAAREASIVYDVTIAEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK374
25 gii124053334 VATA_BOVIN/1-617	289RDFPELTMEVDGKVESIMKRILIANISNMPVAAREASIVYDVTIAEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK391
gii1390405181 VATA_YEAST/1-617	296MEFPELYTMSGTEPIKMKRILIANISNMPVAAREASIVYDVTIAEYFRDNGKHALIIVDDLSKHAEVYRQLSLLMRRPGR EAYFGDVFYLSRLLERAAK398
1 gii13913127 ATPA_AQUAE/1-503	305LNDDL-----GASLTALPIIETKAGDAAPVAPVPTNVSITDQGIFFLETLNAGIRPVAVNPAGISVSRVGE-----GAAQIKAMKQVAGT LRL E386
gii1175143 ATPA_ECOLI/1-513	304VNAEYVEAFTKGVKKTGSLTALPIIETQAGDVSFAVPTNVSITDQGIFFLETLNAGIRPVAVNPAGISVSRVGE-----GAAQIKAMKQVAGT LRL E396
27 gii22266797 ATPA_ILYOB/1-469	304ISDEL-----GGOSLTALPIIETQAGDVSFAVPTNVSITDQGIFFLETLNAGIRPVAVNPAGISVSRVGE-----GSAQIKAMKQVAAKVKRL E385
5 gii114402 ATPA_BOVIN/1-553	355MNDAF-----GGOSLTALPIIETQAGDVSFAVPTNVSITDQGIFFLETLNAGIRPVAVNPAGISVSRVGE-----SAAQTRAMKQVAGT LRL E436
4 gii134190649 ATPA_YEAST/1-545	349LSEKE-----GGOSLTALPIIETQAGDVSFAVPTNVSITDQGIFFLETLNAGIRPVAVNPAGISVSRVGE-----SAAQKALQKQVAGS LK L E430
20 gii32172455 VATA_THET8/1-578	368VITLCGE-----AVTIVGAVSPPGDMSEPVQTSITLRFVGFAPRWDASLSAFRRHPFAIHWNCSLFTSALDPWYENVAEDYFELDAISELQR460
22 gii74053565 VATA_SULAC/1-592	374VIALGKPER-----FOSVIVGAVSPPGDMSEPVQTSITLRFVGFAPRWDASLSAFRRHPFAIHWNCSLFTSALDPWYENVAEDYFELDAISELQR460
23 gii12229704 VATA_HALSA/1-585	375FENFNGTE-----CSISVIGAVSPPGDMSEPVQTSITLRFVGFAPRWDASLSAFRRHPFAIHWNCSLFTSALDPWYENVAEDYFELDAISELQR468
25 gii124053334 VATA_BOVIN/1-617	392VKLCNPER-----EOSVIVGAVSPPGDMSEPVQTSITLRFVGFAPRWDASLSAFRRHPFAIHWNCSLFTSALDPWYENVAEDYFELDAISELQR485
gii1390405181 VATA_YEAST/1-617	399AVALGSPDR-----TOSVIVGAVSPPGDMSEPVQTSITLRFVGFAPRWDASLSAFRRHPFAIHWNCSLFTSALDPWYENVAEDYFELDAISELQR492
1 gii13913127 ATPA_AQUAE/1-503	387LAQFRELEAFVQFASLEDKATQOQINRGLRVLLEKQEPYNPVPEKQIVLVIYAGTHGYLDDIPVESVRKFEKEL-----YAYLDNERDPLKEISEKKKL D483
gii1175143 ATPA_ECOLI/1-513	397LAQYRELAASFQFASLDDATRQLDHDQKVTLEKQKQYAPMSVAQSLVLFPAERGVYLDLVEKISIGSFEAAL-----LAYVRDHPALMQEINQTTGGV N493
27 gii22266797 ATPA_ILYOB/1-469	386LAQYSELLTFAQFCSLDLKDAMQALERGRHIMEIKDQGYPKYKVEEQVVSFMVINGYLDLDAIGDVRKFEFEEI-----KDLRQRHN-----469
5 gii114402 ATPA_BOVIN/1-553	437LAQYREVAAFQAFCSLDDAATQQLSLRGRVRLLEKQKQYSPMAIEEQVAVIYAGVRGLDKLEP SKITRFEFNAF-----LSYLSQHOALLSKIRTDGK I E533
4 gii134190649 ATPA_YEAST/1-545	431LAQYREVAAFQAFCSLDDASIKQTLVREGERLLEKQKQYSPMIAIEEQVPLIYAGVNGHLDGCELSRIGCEFSF-----LSYLSKHNEMLLEIRKGL E527
20 gii32172455 VATA_THET8/1-578	461EACLOEIVQLVGPDAQDAERLVEIENGRIREDPQONAHVEVYQSMKKAQIKMKILALPYKEEAIAIKRQVS-----IDELQLP---VLERIGRARRVVS555
22 gii74053565 VATA_SULAC/1-592	469EDELQIVRLVGPESLSDKDLLEASKLRDAFKQNAFDDBAFSSPQKQAKIMRLIYDFYTNASQLLDKGLT-----LKKILEKQVSEFPD I VRVYKVT565
23 gii12229704 VATA_HALSA/1-585	468EELLEIVQLVCKDAPLEQDQLTLEWARYREAWLQONAHVDRYCYPEKTYALLSGKILTHEESFEALDAGVP-----VEEITSID---AAPRLNR LGT7562
25 gii124053334 VATA_BOVIN/1-617	486EEDLAEIVQLVCKDASLAETDKITLLEWAKLKDQIQNQCYPYDRFCFPYKTYVGLMSNMIAFYDMARRAETQAQSDITWSIIRHEHMDILYKLSMKFKD588
gii1390405181 VATA_YEAST/1-617	493AELEEQVQLVCKSALSDDKITLDWATLLEKEDIQNQCYPYDRFCFPYKTYVGLMSNMIAFYDMARRAETQAQSDITWSIIRHEHMDILYKLSMKFKD589
1 gii13913127 ATPA_AQUAE/1-503	484EEL-----EKKIKKELDAFKQKIVP-----503
gii1175143 ATPA_ECOLI/1-513	494DEI-----EGKLGKILDSFKATQSW-----513
27 gii22266797 ATPA_ILYOB/1-469	534EES-----DAKLEKIVTNFLACFEA-----533
5 gii114402 ATPA_BOVIN/1-553	528KEL-----LALSLSATSESVFAT-----545
4 gii134190649 ATPA_YEAST/1-545	556EEE-----FPAYFEAMKEIQGAFKALA-----578
20 gii32172455 VATA_THET8/1-578	566KND-----ELNKIDELDNLKEAFDLSLLKEVA-----592
22 gii74053565 VATA_SULAC/1-592	563DDE-----HEAEVAEKIQKITEQLRLEL-----585
23 gii12229704 VATA_HALSA/1-585	589PVKDGAEKIKADYALKLEDMQNAFRLSLED-----617
25 gii124053334 VATA_BOVIN/1-617	590PSR---GKEVHGFYFKLLSTMQERFAESTD---617

Figure A1. Sequence alignment of subunit α of F₁F₀ ATP synthases and V₁V₀ ATPases. The sequences of F₁F₀ ATP synthases from *A. aeolicus*, *E. coli*, *I. tartaricus*, bovine and yeast and those of V₁V₀ ATPase from *T. thermophilus*, *H. salinarum*, *S. acidocaldarius*, yeast and bovine are compared. Subunits α are conserved and the conserved residues are highlighted in blue. Nuclear-encoded subunits α possess a mitochondrial targeting sequence. Such N-terminal amino acid sequence is removed from mature subunits α as mitochondrial targeting sequence. It corresponds to the first 43 amino acids of bovine ATP synthase.

Appendix

2 gij 3913 128 ATPB_AQUAE/1-478	1	-----MAEVIKGVVQVIGPVVDVEFEGVKELPKIKDGLKTIIRRAIDDRGNWFEEVLVME55	
gij 81175147 ATPB_ECOLI/1-460	1	-----MATGKIIVQVIGAVVDVEFPQD-APVPRVYDALEVQNGN-----ERLVLE42	
28 gij 22266799 ATPB_ILYOB/1-467	1	-----MENKGI LQTIIGPVVDVSDFN--ELPKIYNAIKIDRGNQ-----EILVAE43	
15 gij 114543 ATPB_BOVIN/1-528	1	MLCLVGRVVAASASGALRGLSPAPLPQAQLLLRAAPAAALQPARDYAAQASPSPKAGATTGRIVAVIGAVVDVQFDE--GLPPI LNALEV88	
18 gij 84028178 ATPB_YEAST/1-441	1	-----EIKTPQ-----GKLVLE12	
12 gij 32172456 ATPB_THET8/1-478	1	-----MDLLKKEYTGIITYISGPIILFVFNKADLAVGALVVDIKDGTGRV--GGQ46	
17 gij 73920440 ATPB_SULAC/1-466	1	-----MSTLMNIREYNSISMIKGPLMAIEGVTDAAYNEIIEVEMPDGSKR--RG148	
13 gij 12229705 VATB_HALSA/1-471	1	-----MKVEYTIIEVSGPLVYVETDEPIGVDEIVIEITPNDGDK-----RGG42	
11 gij 401318 VATB1_BOVIN/1-513	1	MAAEVDSRPRCLPGGGASLGAAREHVQAVTRNYITHPRIYTRVTCVSNGLVVDQVKFAQYAEIVNFTLPGNGTQR-----SGQ79	
9 gij 586211 VATB_YEAST/1-517	1	MVLSDKELFAIN-----KKAVEQGFNVKPLNNTVSGVNGPLVILEKVKFRYNEIVNLTLPDGTVR-----QQG66	
2 gij 3913 128 ATPB_AQUAE/1-478	56	VAQHI GEHRVRAIAMGPTDGLVRGQVEYELGPKIPIVGVKEVLRGIFNVACQPIIDEGQPVEAKYEWPMFRNPPPELVEQSTKVIEILETGIK145	
gij 81175147 ATPB_ECOLI/1-460	43	VQQLGGGIVRTIAMGSSDGLRRLDVKLEHPIEVVPYKATLGRIMNVLCPEVDMKGEIIEEERWAHRAAPSYEELNSQELLETGKIK132	
28 gij 22266799 ATPB_ILYOB/1-467	44	VQQLGNVNVRAICMDSGELGRQMEIITDGPITVVPYKAVLGRILNVLCEAIDQKEELNAEEFAPHRAEPAFEDQGTVDIEIFETGKIK133	
15 gij 114543 ATPB_BOVIN/1-528	89	QGR-----TRLVLEVAQHLCGSESTVRIAMDGTEGLVGRQKVLDSGAPIRIPVPGTELRGIMNVIGEPIDERGPIKTKGFAAIIHAE169	
18 gij 84028178 ATPB_YEAST/1-441	13	VQHLGENTVRTIAMDGTEGLVGRQKVLDSGAPISVPGVRETLRGIIINVI GEPIDERGPIKSKLRKPIHADPPSFAEQSTSAEILETGKIK102	
12 gij 32172456 ATPB_THET8/1-478	47	VI EVSEYAVIQVFEETGLDLATLTVSLSVEDVARLGMSEMLGRGRFNIGIKPIDGLPPIIPEKRLPIITGLPLNPAVARRKPEEQFIQTGVS136	
17 gij 73920440 VATB_SULAC/1-466	49	VDSQSGVAIQVFEETGVSPDTQSKVRFLGRCLVEKISEMLGRGFIFTPLECEPLDNCQVLSGCKRDNGLSNPSPVEPVEEPIQTGVS138	
13 gij 12229705 VATB_HALSA/1-471	43	VLES DGFVAIQVFEETGEGV--KDAKRWFLFGETLKMPTIEDLGRVLDGSGNPIIDGCPDIIVPDRVDVGEAINPHAREYEPVEEPIQTGVS131	
11 gij 401318 VATB1_BOVIN/1-513	80	VLESGTKAIQVFEETSGTDAQSTCEFTDGLRTPVSEMLGRVFNCSGKPKIDKCPVMAEDFDLNGCPIINPHDIRIPEEMIETGIVS169	
9 gij 586211 VATB_YEAST/1-517	67	VLEIRGDRAIQVFEETSGIDVKKTTVEFTGSELRIPVSEMDLGRIFDGSGRPIDNCPKVFAYEDYLDNGSPINPYARIPEEMIETGIVS156	
2 gij 3913 128 ATPB_AQUAE/1-478	146	VLDLQPIIKGKGVLFEGAGVGTKVMQELIHNIARFHEG-----YSVYVGVGERTRREGNDLWLEMKESGVLP-----YVM218	
gij 81175147 ATPB_ECOLI/1-460	133	VIDLMPFAKGGKGVLFEGAGVGTKVMMLERINIAIEHSG-----YSVYVGVGERTRREGNDFYHEMTDSNVI-----DKTSL205	
28 gij 22266799 ATPB_ILYOB/1-467	134	VLDLAPVYKGGKIGLFGAGVGTKVMIMELINIAIKGHGG-----LSVYVGVGERTRREGNDLWLEMKESGVLP-----DKTSL206	
15 gij 114543 ATPB_BOVIN/1-528	170	APFEVEMSVDEQELIIVTGIKVVLDLAPYAKGGKIGLFGAGVGTKVMIMELINNVAKAHGG-----YSVYVGVGERTRREGNDLWLEMKESGVLP-----DKTSL206	
18 gij 84028178 ATPB_YEAST/1-441	103	VVDLAPYVYKGGKIGLFGAGVGTKVMVITQELINIAIKAHGG-----FSVYVGVGERTRREGNDLWLEMKETGVINLEGE-SKVAL180	
12 gij 32172456 VATB_THET8/1-478	137	VDMNTLVRQKLPISGSGLPANEAADIAQRAQATVPRLDSEGEKE-EPFAVYFAAMGIRQBELSYFIOEFERTGALS-----RSVL219	
17 gij 73920440 VATB_SULAC/1-466	139	VGLDLSLRQKLPISGSGLPANEAADIAQRAQATVPRLDSEGEKE-EPFAVYFAAMGIRQBELSYFIOEFERTGALS-----RVAM214	
13 gij 12229705 VATB_HALSA/1-471	132	GDGMNTLVRQKLPISGSGLPANEAADIAQRAQATVPRLDSEGEKE-EPFAVYFAAMGIRQBELSYFIOEFERTGALS-----RSV215	
11 gij 401318 VATB1_BOVIN/1-513	170	PIDVIMNIEIARQKIPISASGLPHNEIAAQICRQAGLVKRPKFDVHGHHEENFISVFAAMVNNLETARFFKQDFEENGSL-----RTSL240	
9 gij 586211 VATB_YEAST/1-517	157	ADTMSIARQKIPISASGLPHNEIAAQICRQAGLVKRPKFDVHGHHEENFISVFAAMVNNLETARFFKQDFEENGSL-----RTSL240	
2 gij 3913 128 ATPB_AQUAE/1-478	219	VYQGMNEPVGVRFRVAHTGLTMAEYFRDVEQDVLIFIDNIFRFVQAGAEVSTLLGRLP SAVGYQPTLAEEDVGEVQERITSTK--KGSIT306	
gij 81175147 ATPB_ECOLI/1-460	206	VYQGMNEPVGNNLRVALTGLTMAEYFRDVEQDVLIFIDNIFRFVQAGAEVSTLLGRLP SAVGYQPTLAEEDVGEVQERITSTK--KGSIT292	
28 gij 22266799 ATPB_ILYOB/1-467	207	VYQGMNEPVGARLRVALTGLTMAEYFRDVEQDVLIFVDNIFRFVQAGAEVSTLLGRLP SAVGYQPTLATDMGALQERITSTK--KGSIT294	
15 gij 114543 ATPB_BOVIN/1-528	249	EMIESCVINLKDATSKVALVYQGMNEPVGARLRVALTGLTVAEYFRDVEQDVLIFIDNIFRFVQAGAEVSTLLGRIP SAVGYQPTLATD338	
18 gij 84028178 ATPB_YEAST/1-441	181	VYQGMNEPVGARLRVALTGLTVAEYFRDVEQDVLIFIDNIFRFVQAGAEVSTLLGRIP SAVGYQPTLATDMGALQERITSTK--KGSIT268	
12 gij 32172456 VATB_THET8/1-478	220	FLNKADDTIERILTPRMALVAEYLAFFHNDYHMLVILTDMTNCALREICARIEICRRRGYGYMYTDLATYERAGVVEGKGSVIT309	
17 gij 73920440 VATB_SULAC/1-466	215	FVTLANDPPLSKILTPKALTALREYLAEKDMMHIALILDMTNCALRELSASKSEVPGRGYGYMYTDLATYERAGVVEGKGSVIT304	
13 gij 12229705 VATB_HALSA/1-471	216	FMLNADDPAVERTVPRMALTTAEYLAEKDHYVILTDMTNCALRLQICARIEEVEPGRRGYGYMYTDLATYERAGVVEGKGSVIT305	
11 gij 401318 VATB1_BOVIN/1-513	254	FLNLANDTIERIITPRMLTTAEYLAQCEHVVLITDMSSYADALREVSAAREVEPGRRGYGYMYTDLATYERAGVVEGKGSVIT343	
9 gij 586211 VATB_YEAST/1-517	241	FLNLANDTIERIITPRMLTTAEYLAQCEHVVLITDMSSYADALREVSAAREVEPGRRGYGYMYTDLATYERAGVVEGKGSVIT330	
2 gij 3913 128 ATPB_AQUAE/1-478	307	AIQAVYVYVADDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---EEHYEVAMEVKSILQRYKELQEI IAI L392	
gij 81175147 ATPB_ECOLI/1-460	293	SVQAVYVYVADDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---EEHYDTARGVKSILQRYKELQEI IAI L378	
28 gij 22266799 ATPB_ILYOB/1-467	295	SVQAVYVYVADDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---EEHYTTAREVQVSLQRYKELQEI IAI L380	
15 gij 114543 ATPB_BOVIN/1-528	339	MGTMQERITTTK--KGSITSVQALYVPADDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---SEHYDVAR422	
18 gij 84028178 ATPB_YEAST/1-441	269	SVQAVYVYVADDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---QEVYDVAASKVQETLQTYKSLQDI IAI L354	
12 gij 32172456 VATB_THET8/1-478	310	QIPILTMPDDDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---QEVYDVAASKVQETLQTYKSLQDI IAI L399	
17 gij 73920440 VATB_SULAC/1-466	305	QMPILTMPDDDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---QEVYDVAASKVQETLQTYKSLQDI IAI L394	
13 gij 12229705 VATB_HALSA/1-471	306	QIPILTMPDDDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---QEVYDVAASKVQETLQTYKSLQDI IAI L395	
11 gij 401318 VATB1_BOVIN/1-513	344	QIPILTMPDDDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---QEVYDVAASKVQETLQTYKSLQDI IAI L433	
9 gij 586211 VATB_YEAST/1-517	331	QIPILTMPDDDDITDPAWPSIFAHLDAITVLTIRRLAELGIYVADPDEESTSKYLAPYVVG---QEVYDVAASKVQETLQTYKSLQDI IAI L420	
2 gij 3913 128 ATPB_AQUAE/1-478	393	GMEELSDDEKAI VNRARRIQK-FLSOPFVAAEVFTGMPGKYVVKEDTIRSFKEVLTGKYDHLPEAFYMGV-----TIED466	
gij 81175147 ATPB_ECOLI/1-460	379	GMDLSEDEKLVVARARRIQK-FLSOPFVAAEVFTGSPGKYVSKDTRIRGFKGIMEGEYDHLPEAFYMGV-----SIEE452	
28 gij 22266799 ATPB_ILYOB/1-467	381	GMDLSEDEKTVSRRARRIQK-FLSOPFVAAEVFTGMGKYVVKETITGFKMEIEGKHDALEQAFYMGV-----TIDE454	
15 gij 114543 ATPB_BOVIN/1-528	423	GVKILQDYKSLQDI IAI LGMDELSEDEKLTVSRARRIQK-FLSOPFVAAEVFTGHLGKLVLPKETEIKGFOQLAGEYDHLPEAFYMGV-----GIED428	
18 gij 84028178 ATPB_YEAST/1-441	355	GMDLSEDEKLTVERARRIQK-FLSOPFVAAEVFTGIPKGLVLRKDTVASFAVEGKYDNIPEAFYMGV-----GIED428	
12 gij 32172456 VATB_THET8/1-478	400	GEDALTENDRRYLQFAEFERFINOQQQ-NRSEESLIQAWALMSLMP--QGEIKRI SKDHIKGYVYQK-----LEE469	
17 gij 73920440 VATB_SULAC/1-466	395	GEDLSDVDRRYLQFAEFERFINOQINENRDIATLFGHWEVLSMLP--ESELSLIRTEYIKKYHPNYR-----VKK466	
13 gij 12229705 VATB_HALSA/1-471	396	GEALSERDNRYLQFAEFERFIDGCFKTNRDIETEDLQWELMSLMP--KTELNRVDEDLIEDHY-----VED463	
11 gij 401318 VATB1_BOVIN/1-513	434	GEALTSDELLEYLQKFEKFFINOGPYEKR SVFESLDLQWKLRTFP--KEMLRIPQNIIDEFFSRE-----GAPQ505	
9 gij 586211 VATB_YEAST/1-517	421	GEALSIDKLSLELQKFEKFFITGAYEDRTVFEISLDQAWSLLRIYR--KEMLRISP KILDEFYDRARDDAEDEEDPDRS5GKKK508	
2 gij 3913 128 ATPB_AQUAE/1-478	467	VIEKAKQMGAKV-----	478
gij 81175147 ATPB_ECOLI/1-460	453	AVEKAKKLE-----	460
28 gij 22266799 ATPB_ILYOB/1-467	455	AIAKARELMKGD-----PIIEAVAKADKLAEEHS	467
15 gij 114543 ATPB_BOVIN/1-528	512	-----PIIEAVAKADKLAEEHS	528
18 gij 84028178 ATPB_YEAST/1-441	429	VVAKAEKLAEEAN-----	441
12 gij 32172456 VATB_THET8/1-478	470	IWGAPQALD-----	478
17 gij 73920440 VATB_SULAC/1-466	464	VADAEATAD-----	471
13 gij 12229705 VATB_HALSA/1-471	506	DEADTAL-----	513
11 gij 401318 VATB1_BOVIN/1-513	509	DAQSEESLI-----	517

Figure A2. Sequence alignment of subunit β of F_1F_0 ATP synthases and V_1V_0 ATPases. The sequences of F_1F_0 ATP synthases from *A. aeolicus*, *E. coli*, *I. tartaricus*, bovine and yeast and those of V_1V_0 ATPase from *T. thermophilus*, *H. salinarum*, *S. acidocaldarius*, yeast and bovine are compared. Subunits β are highly conserved and the conserved residues are highlighted in blue. Nuclear-encoded subunits β possess a mitochondrial targeting sequence, which corresponds to the first 48 amino acids of bovine heart ATP synthase..

A.2. Codon usage

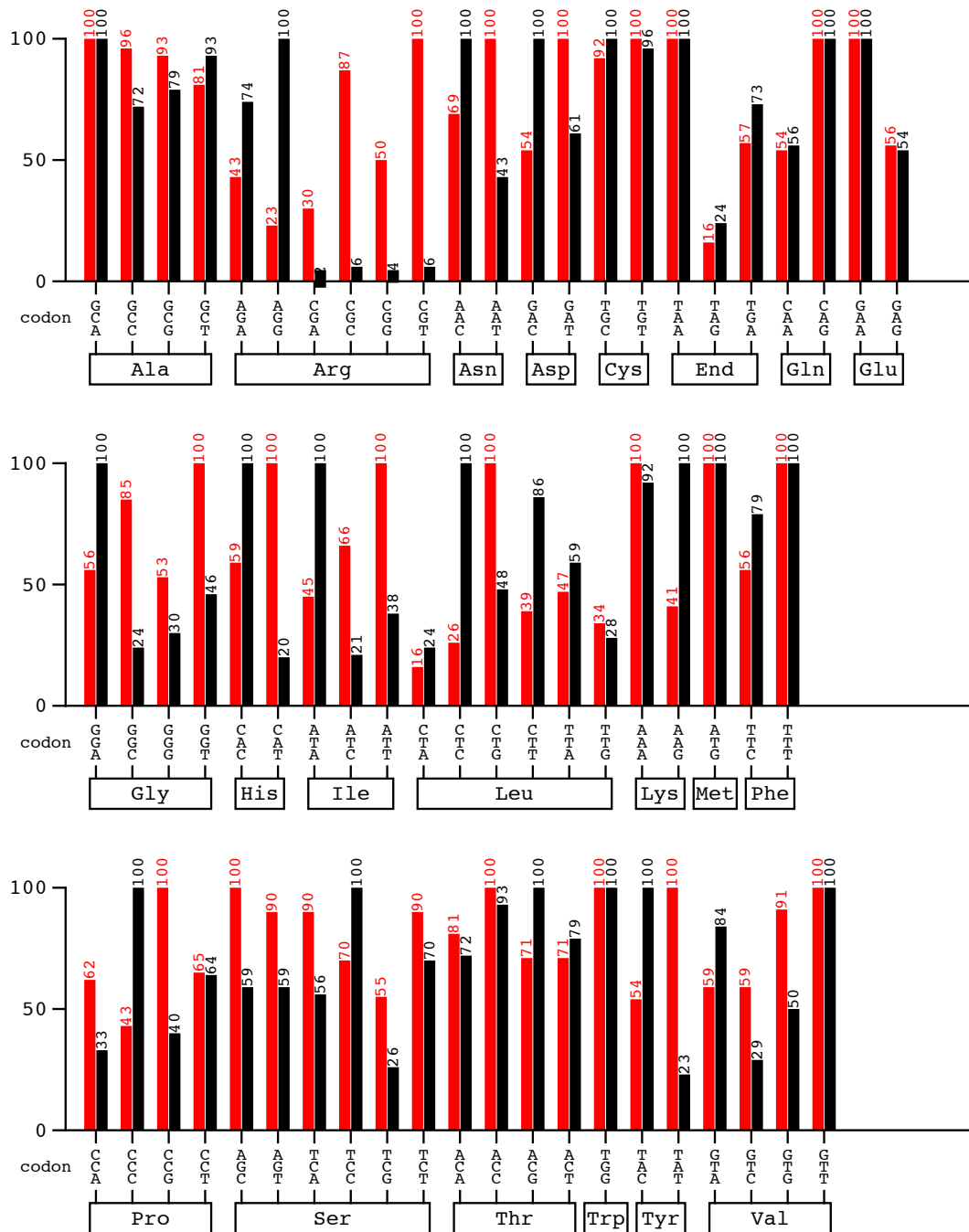


Figure A3. Codon usage difference between *A. aeolicus* and *E. coli* analyzed by Graphical Codon Usage Analyser (GCAU). Codon tables of *E. coli* and *A. aeolicus* are shown in red and black, respectively.

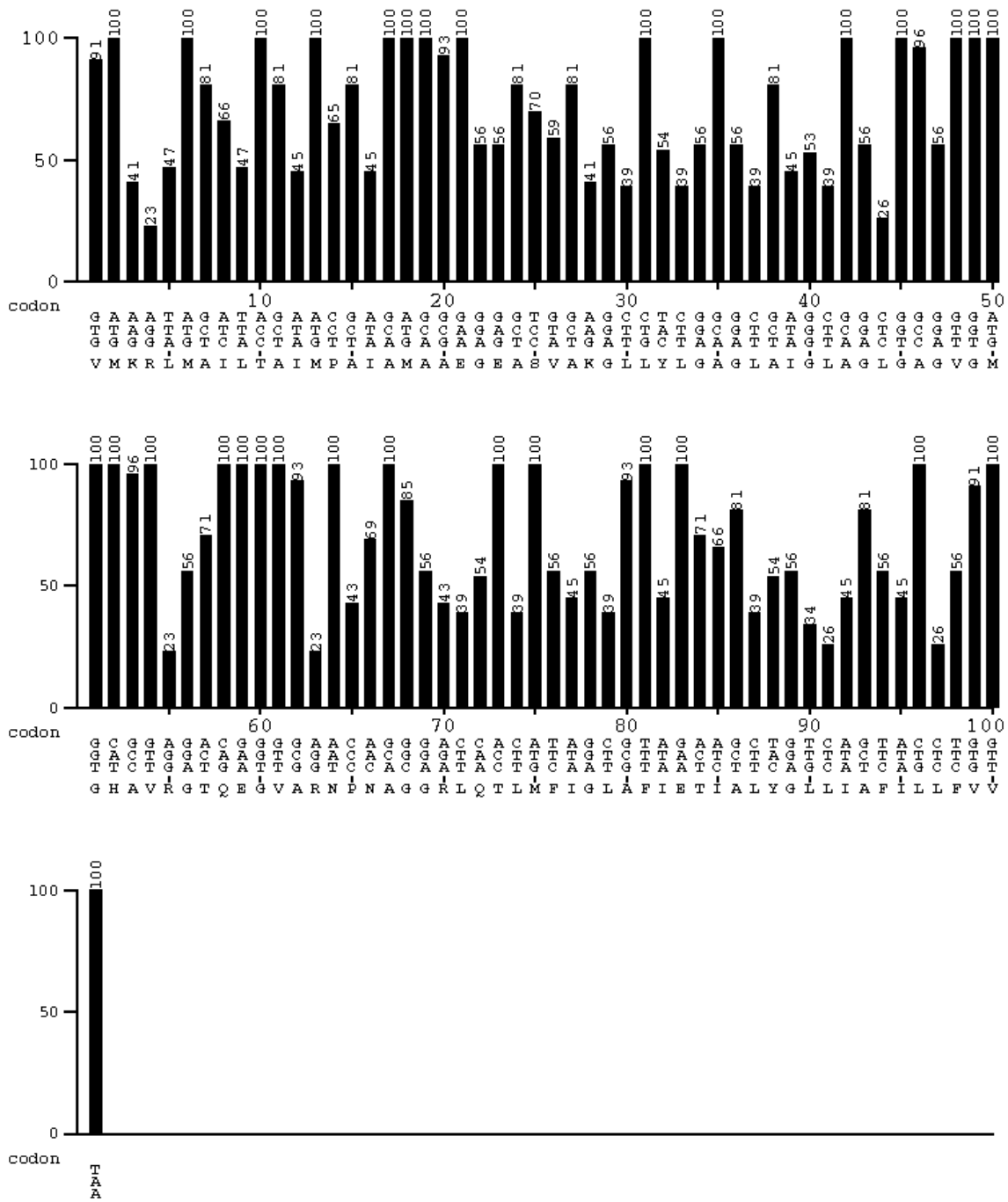


Figure A4. Codon usage in subunit c. The figure reports the codons of subunit c from *A. aeolicus* ATP synthase compared to the codon usage table of *E. coli* analyzed by Graphical Codon Usage Analyser (GCAU).

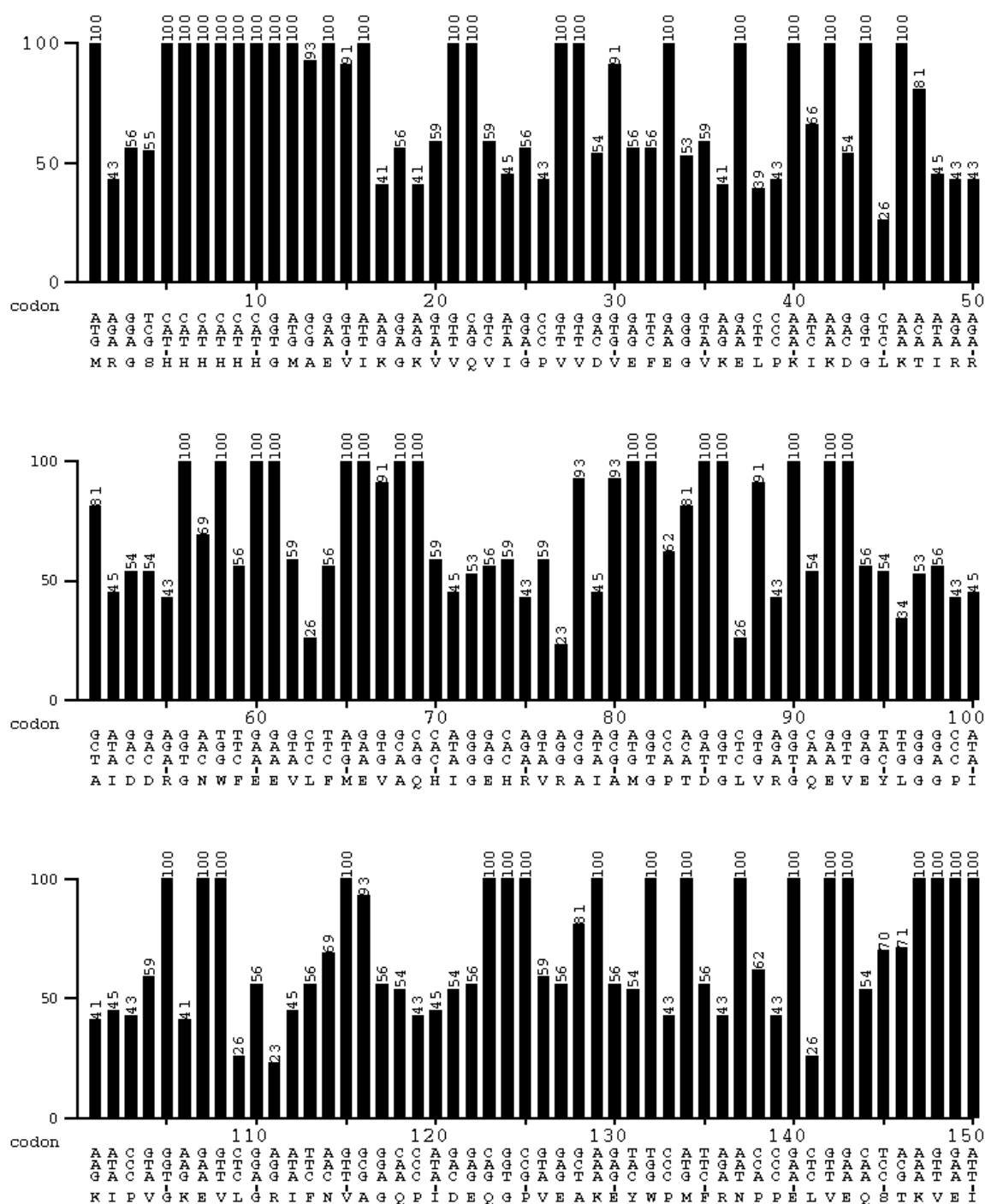


Figure A5. Codon usage in subunit β . The figure reports the codons of subunit β from *A. aeolicus* ATP synthase compared to the codon usage table of *E. coli* analyzed by Graphical Codon Usage Analyser (GCAU) (continues).

A.3. Plasmid maps and DNA sequences

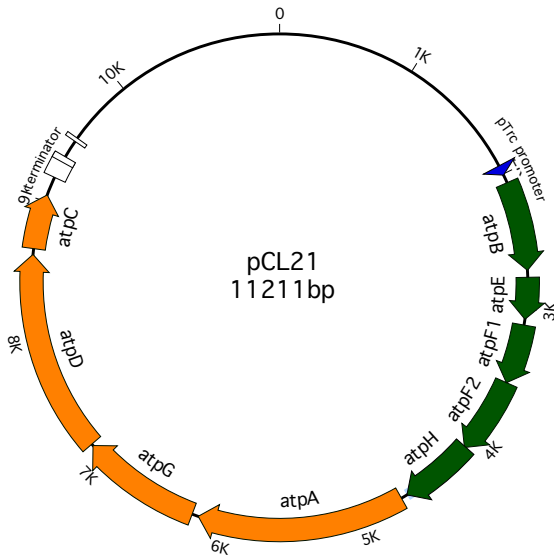


Figure A6. Plasmid map of vector pCL21

Features	Location
pTrc promoter	1935 - 2008
<i>atpB</i> / subunit a (<i>aq_179</i>)	2069 - 2719
<i>atpE</i> / subunit c (<i>aq_177</i>)	2770 - 3072
<i>atpF1</i> / Subunit b ₁ (<i>aq_1586</i>)	3115 - 3549
<i>atpF2</i> / Subunit b ₂ (<i>aq_1587</i>)	3549 - 4106
<i>atpH</i> / Subunit δ (<i>aq_1588</i>)	4099 - 4644
<i>atpA</i> / Subunit α (<i>aq_679</i>)	4691 - 6202
<i>atpG</i> / Subunit γ (<i>aq_2041</i>)	6249 - 7124
<i>atpD</i> / Subunit β (<i>aq_2038</i>)	7138 - 8607
<i>atpC</i> / Subunit ε (<i>aq_673</i>)	8650 - 9048
rrnB-terminator	9186 - 9343
rrnB_T1_terminator	9309 - 9352
rrnB_T2_terminator	9484 - 9511

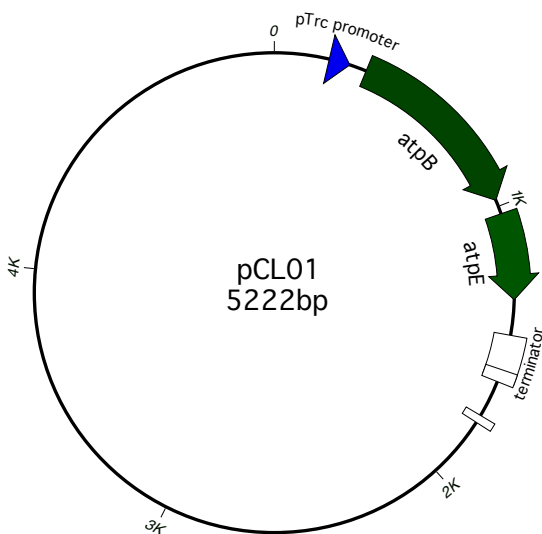


Figure A7. Plasmid map of vector pCL01

Features	Location
pTrc promoter	193 - 266
<i>atpB</i> / subunit a (<i>aq_179</i>)	327 - 977
<i>atpE</i> / subunit c (<i>aq_177</i>)	1022 - 1330
rrnB-terminator	1455 - 1612
rrnB_T1_terminator	1578 - 1621
rrnB_T2_terminator	1753 - 1780
rrnB_T2_terminator	1753 - 1780

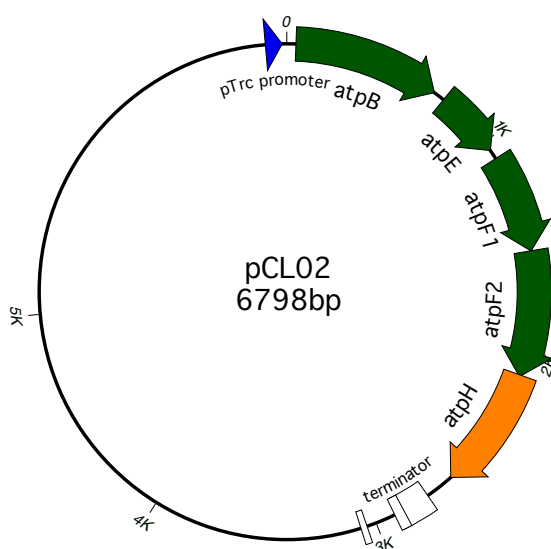


Figure A8. Plasmid map of vector pCL02

Features	Location
pTrc promoter	6705 - 6778
<i>atpB</i> / subunit a (<i>aq_179</i>)	41 - 691
<i>atpE</i> / subunit c (<i>aq_177</i>)	742 - 1044
<i>atpF1</i> / Subunit b ₁ (<i>aq_1586</i>)	1087 - 1521
<i>atpF2</i> / Subunit b ₂ (<i>aq_1587</i>)	1521 - 2078
<i>atpH</i> / Subunit δ (<i>aq_1588</i>)	2071 - 2616
rrnB-terminator	2745 - 2902
rrnB_T1_terminator	2868 - 2911
rrnB_T2_terminator	3043 - 3070

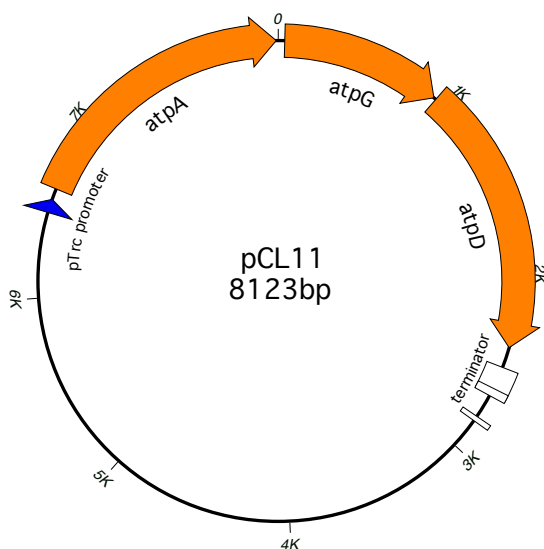
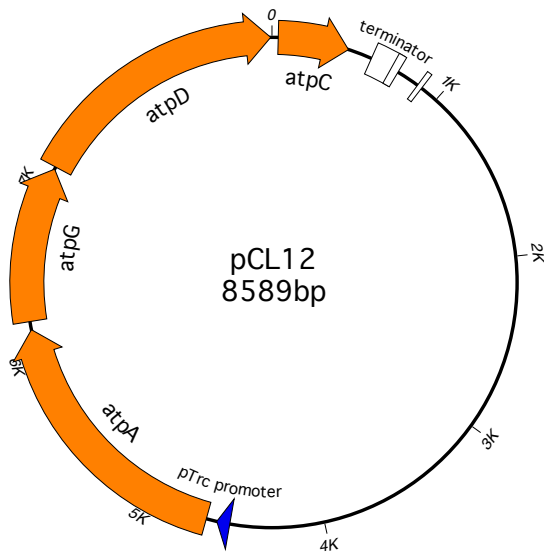


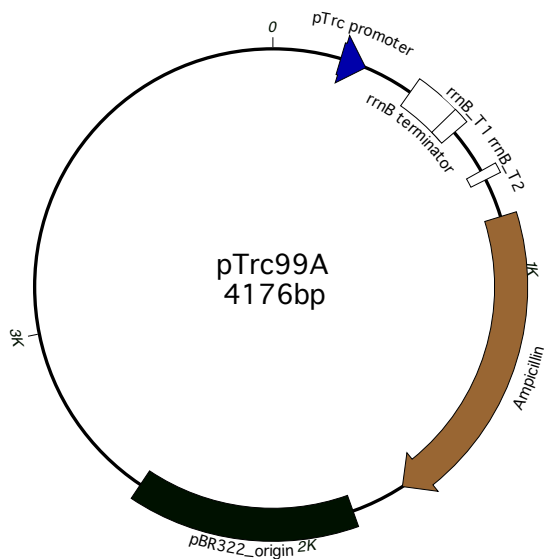
Figure A9. Plasmid map of vector pCL11

Features	Location
pTrc promoter	6467 - 6540
<i>atpA</i> / Subunit α (<i>aq_679</i>)	6601 - 8112
<i>atpG</i> / Subunit γ (<i>aq_2041</i>)	36 - 911
<i>atpD</i> / Subunit β (<i>aq_2038</i>)	925 - 2394
<i>atpC</i> / Subunit ε (<i>aq_673</i>)	8650 - 9048
rrnB-terminator	2507 - 2664
rrnB_T1_terminator	2630 - 2673
rrnB_T2_terminator	2805 - 2832



Features	Location
pTrc promoter	4532 - 4605
<i>atpA</i> / Subunit α (<i>aq_679</i>)	4666 - 6177
<i>atpG</i> / Subunit γ (<i>aq_2041</i>)	6224 - 7099
<i>atpD</i> / Subunit β (<i>aq_2038</i>)	7113 - 8582
<i>atpC</i> / Subunit ϵ (<i>aq_673</i>)	36 - 434
rrnB-terminator	572 - 729
rrnB_T1_terminator	695 - 738
rrnB_T2_terminator	870 - 897

Figure A10. Plasmid map of vector pCL12



Features	Location
pTrc promoter	193 - 266
Ampicillin	846 - 1706
pBR322_origin	1861 - 2480
rrnB-terminator	409 - 566
rrnB_T1_terminator	532 - 575
rrnB_T2_terminator	707 - 734

Figure A11. Plasmid map of empty vector pTrc99A

DNA Sequence of pCL21

```

1  gogcaacgca  attaatgtga  gttagcgcgga  attgatctgg  tttgacagct  tatcatcgac
   cgcggttgctg  taattacact  caatcgcgct  taactagacc  aaactgtcga  atagtagctg

                                                                SPI
                                                                GGCTGTG
61  tgcacgggtgc  accaatgctt  ctggcgctcag  gcagccatcg  gaagctgtgg  tatggctgtg
   acgtgccacg  tggttacgaa  gaccgcagtc  cgtcggtagc  cttcgacacc  ataccgacac

                                                                SPI ▶
                                                                CAGGTCGTAA ATC
121  caggtcgtaa  atcaactgcat  aattcgtgtc  gctcaaggcg  cactcccgtt  ctggataatg
   gtccagcatt  tagtgacgta  ttaagcacag  cgagttccgc  gtgagggcaa  gacctattac

181  ttttttgcg  cgacatcata  acggttctgg  caaatattct  gaaatgagct  gttgacaatt
   aaaaaacg  gctgtagtat  tgccaagacc  gtttataaga  ctttactcga  caactgttaa

241  aatcatccgg  ctcgataaat  gtgtggaatt  gtgagcggat  aacaatttca  cacaggaaac
   ttagtaggcc  gagcatatta  cacaccttaa  cactcgccta  ttgttaaagt  gtgtcctttg
                                                                pTrc promoter

                                                                KpnI
                                                                NdeI
301  agaccatgga  attcgagctc  ggtaccata  tgtaatctga  gccaatgca  aaagagtaa
   tctggtacct  taagctcgag  ccatgggtat  acattagact  cggttaacgt  tttctccatt

361  gggaaatgga  gtactcgcac  gtagtttacg  cactcttagc  ggtggcactt  gcgataat
   ccctttacct  catgagcgtg  catcaaagtc  gtgagaatcg  ccaccgtgaa  cgctattaa
   >>.....subunit a (aq 179).....>
   m e y s h v v y a l l a v a l a i i

421  tcgtcctcaa  aggaggaaaa  ccttccttta  aacctacaaa  gtatcaagct  ctactggaag
   agcaggagtt  tcctcctttt  ggaagggaa  ttggatgttt  catagttcga  gatgacctc
   >.....subunit a (aq 179).....>
   f v l k g g k p s l k p t k y q a l l e

481  gatacctgag  attcgtcagg  aatatgctcc  tggaaaacgt  aggcgaaaga  ggacttaaat
   ctatggactc  taagcagtc  ttatacgagg  accttttga  tccgctttct  cctgaattta
   >.....subunit a (aq 179).....>
   g y l r f v r n m l l e n v g e r g l k

541  acgtaccctt  gatagcagcc  ataggactct  tcgttttctt  cgggaacata  cttagaatgg
   tgcattggaa  ctatcgtcgg  tatcctgaga  agcaaaagaa  gcccttgtat  gatccttacc
   >.....subunit a (aq 179).....>
   y v p l i a a i g l f v f f g n i l g m

601  ttcccggtt  tgaagctcct  accgcaaaca  taaacaccaa  cctagcttta  gcacttttag
   aagggccgaa  acttcgagga  tggcgtttgt  atttggtgtt  ggatcgaaat  cgtgaaatc
   >.....subunit a (aq 179).....>
   v p g f e a p t a n i n t n l a l a l l

661  tgttctttta  ttaccacttt  gaaggcttca  gggaaaacg  tcttgcttat  ttgaaacact
   acaagaaaat  aatggtgaaa  cttccgaagt  cccttttgc  agaacgaata  aactttgtga
   >.....subunit a (aq 179).....>
   v f f y y h f e g f r e n g l a y l k h

721  ttatgggacc  catacctcta  atggctcctt  ttttcttctg  ggttgaagtt  atatctcaca
   aataccctgg  gtatggagat  taccgaggg  aaagaagca  ccaacttcaa  tatagagtg
   >.....subunit a (aq 179).....>
   f m g p i p l m a p f f f v v e v i s h

```

SP2 ▶

CCCTC TCCCTCAGGT TATTC

```

781 tagcaagacc aatcaccctc tccctcaggt tattcgcaa catgaaagcg ggagcactcc
    atcgttctgg ttagtgggag agggagtcca ataagcgtt gtactttcgc cctcgtgagg
    >.....subunit a (aq 179).....>
    i a r p i t l s l r l f a n m k a g a l

841 tacttcttac tttagtaagc ctggttatca agaatccatt cacgctggta gtatcaccgg
    atgaagaatg aatcattcgc gaccaatagt tcttaggtaa gtgcgacat catagtggcc
    >.....subunit a (aq 179).....>
    l l l t l v s l v i k n p f t l v v s p

901 ttgtgcttat attcgttata gctataaagt tcctcgccat attcatacag acttacatct
    aacacgaata taagcaatat cgatatttca aggagcggta taagtatgtc tgaatgtaga
    >.....subunit a (aq 179).....>
    v v l i f v i a i k f l a i f i q t y i

961 ttatgatact ctcggtgggt tacatagccg gagctgtagc acacgaggag cactgatatg
    aatactatga gagccacca atgtatcggc ctgcacatcg tgtgctcctc gtgactatac
    >.....subunit a (aq 179).....>>
    f m i l s v v y i a g a v a h e e h -

1021 aaggtactgc tagtttctta tagttaaata agctttaagg aggtagggtga tgaagaggtt
    ttccatgacg atcaaagaat atcaatttat tcgaaattcc tccatccact acttctccaa
    subunit c (aq_177) >>.....>
    v m k r

1081 aatggctatc ttaaccgcta taatgcctgc tatagcaatg gcagcgggaag gagaggcttc
    ttaccgatag aattggcgat attacggacg atatcgttac cgtcgccttc ctctccgaag
    >.....subunit c (aq 177).....>
    l m a i l t a i m p a i a m a a e g e a

1141 cgtagctaag ggacttctgt accttgagc aggacttgc atagggcttg caggactcgg
    gcatcgattc cctgaagaca tggaaacctc tcctgaacga tatccogaac gtctcgagcc
    >.....subunit c (aq 177).....>
    s v a k g l l y l g a g l a i g l a g l

1201 tgccggagtt ggtatgggtc atgccgtag gggaactcag gaaggtgttg cgaggaatcc
    acggcctcaa ccatacccag tacggcaatc cccttgagtc cttccacaac gctccttagg
    >.....subunit c (aq_177).....>
    g a g v g m g h a v r g t q e g v a r n

1261 caacgcaggc ggaagacttc aaacccttat gttcatagga cttgcgttta tagaaactat
    gttgcgtccg ccttctgaag tttgggaata caagtatcct gaacgcaaat atctttgata
    >.....subunit c (aq 177).....>
    p n a g g r l q t l m f i g l a f i e t

```

BglII

```

1321 cgctctttac ggattgctca tagctttcat actgctcttc gtggtttaag cccttagatc
    gcgagaaatg cctaacgagt atcgaaagta tgacgagaag caccaaattc gggaaatctag
    >.....subunit c (aq 177).....>>
    i a l y g l l i a f i l l f v v -

```

SP3 ▶

TTAGC GGAGGAGAAG AATGG

```

1381 tattgctata attgtttagc ggaggagaag aatggacata ggagtaatgc ctaatgcaac
    ataacgatat taacaaatcg cctcctcttc ttacctgtat cctcattacg gattacgttg
    >>...Subunit b1 (aq 1586).....>
    m d i g v m p n a

1441 aatcctcgtt caattgttca tcttcgtaat attcctaagc ataataccta acatctacgt
    ttaggagcaa gttaacaagt agaagcatta taaggattac tattagtgat tgtagatgca
    >.....Subunit b1 (aq 1586).....>
    t i l v q l f i f v i f l m i i t n i y

```



```

1501 aaagccctac accgcggtga tagaatccag agaagaactc attaagaaga acctctctga
tttcgggatg tggcgccact atcttaggtc tcttcttgag taattcttct tggagagact
>.....Subunit b1 (aq 1586).....>
v k p y t a v i e s r e e l i k k n l s

1561 agcacaaaag ttaagggaag aaactcaaac ctacctcact caggccaaag aagttctoga
tcgtgttttc aattcccttc tttgagtttg gatggagtga gtcgggtttc ttcaagagct
>.....Subunit b1 (aq 1586).....>
e a q k l r e e t q t y l t q a k e v l

1621 agatgcgaag aagagggcgg atcaataaat tgaaaacgca ggaagggagg cggaagctca
tctacgcttc ttctcccgcc tagtttatta acttttgctt ccttccctcc gccttcgagt
>.....Subunit b1 (aq 1586).....>
e d a k k r a d q i i e n a g r e a e a

1681 ggcgagaagc ataatcgagc agacggaaaa acaaaccgaa gaagagatta agaaagcagt
ccgctcttcg tatttagctcg tctgcctttt tgtttggctt cttctctaata tctttogtca
>.....Subunit b1 (aq_1586).....>
q a r s i i e q t e k q t e e e i k k a

1741 ggaggaaatc agaacctcct tagaagaaga gaagaagaag ctgaaaagt ccgtaaagga
cctcctttag tcttggagga atcttcttct cttcttcttc gagcttttca ggcatttctc
>.....Subunit b1 (aq 1586).....>
v e e i r t s l e e e k k k l e k s v k

1801 aatagctcag gaaattgtaa agaaaatttt gagagaggcg gcgtgatggt gaggttgata
ttatcgagtc ctttaacatt tcttttaaaa ctctctccgc cgcactacca ctccaactat
>.....Subunit b1 (aq 1586).....>>
e i a q e i v k k i l r e a a -
Subunit b2 (aq 1587) >>.....>
m v r l i

1861 agtttcttaa ctctggcttc tacttttgct tacgcgggtg aaggacattt gggacactcc
tcaaagaatt gagaccgaag atgaaaacga atgcgcccac ttctgtaaa cctgtgagg
>.....Subunit b2 (aq 1587).....>
s f l t l a s t f a y a g e g h l g h s

1921 cccggagcgc tgatctggaa agggctcaac atactcgcgt tcctcggaat agtttactac
ggcctcgcg actagacctt tcccgagttg tatgagcgca aggagcctta tcaaatgatg
>.....Subunit b2 (aq_1587).....>
p g a l i w k g l n i l a f l g i v y y

1981 tttgaaaaaa aaccataag cgaagccttt aacaagttct acaactcaat agtggagagc
aaaccttttt ttgggtattc gtttcggaaa ttgttcaaga tgttgagtta tcacctctcg
>.....Subunit b2 (aq 1587).....>
f g k k p i s e a f n k f y n s i v e s

2041 ctcgtaaacg cagaaagaga gttcatgatg gcaagggagg aactttcaaa agctaagag
gagcatttgc gtctttctct caagtactac cgttccctcc ttgaaagttt togatttctc
>.....Subunit b2 (aq 1587).....>
l v n a e r e f m m a r e e l s k a k e

SP4_new ►
GGCACAG GAATACGAGA AAC
2101 gaactcgaag atgcgaagaa aaaggcacag gaatacgaga aactcgcaat agaaaccgcg
cttgagcttt tacgcttctt tttccgtgtc cttatgctct ttgagcgita tctttggcgc
>.....Subunit b2 (aq 1587).....>
e l e n a k k k a q e y e k l a i e t a

2161 gaaacggaaa agaaaaagat actccagcac gcccgaggag tttccgaaag gataaaggaa
ctttgccttt tctttttcta tgaggtcgtg cgggtccttc aaaggctttc ctatttctc
>.....Subunit b2 (aq_1587).....>
e t e k k k i l q h a q e v s e r i k e

```

SP4 ▶

AGGCTAAGG AGACGATAGA G

2221 aaggctaagg agacgataga gattgaactg aataaagcta agaaagaact cgccctttac
 ttccgattcc tctgctatct ctaacttgac ttatttcgat tctttcttga gcgggaaatg
 >.....Subunit b2 (aq 1587).....>
 k a k e t i e i e l n k a k k e l a l y

2281 ggaatacaga aggctgaaga aatagcaaag gatcttctcc aaaaagaatt caagaagtcc
 ccttatgtct tccgacttct ttatcgtttc ctagaagagg tttttcttaa gttcttcagg
 >.....Subunit b2 (aq 1587).....>
 g i q k a e e i a k d l l q k e f k k s

2341 aaagttcagg aaaagtacat agaggctcag ttaaagctcc tggaggagag gaagaatgct
 tttcaagtcc ttttcatgta tctccgagtc aatttcgagg acctcctctc cttcttacga
 >.....Subunit b2 (aq 1587).....>
 k v q e k y i e a q l k l l e e r k n a
 Subunit delta (aq 1588) >>...>
 m

2401 taagaggaaa gaactcgcaa ggaaggctgt aaggctcata gtaaagaagg ttccaaagga
 attctccttt cttgagcgtt ccttccgaca ttccgagtat catttcttcc aaggtttcct
 >>> Subunit b2 (aq 1587)
 -
 >.....Subunit delta (aq 1588).....>
 l k r k e l a r k a v r l i v k k v p k

2461 aaaggaaagc atcttaaagg ttgacgagtt cttaggaacc ctttccacag cttacaggaa
 tttcctttcg tagaatttcc aactgctcaa gaatccttgg gaaaggtgtc gaatgtcctt
 >.....Subunit delta (aq 1588).....>
 e k e s i l k v d e f l g t l s t a y r

2521 ggacaaactc ctgagaaact tcttctgtgc gccccaata gacagaaacg caaaggtaaa
 cctgttttgag gactctttga agaaggacag cggggtttat ctgtctttgc gtttccattt
 >.....Subunit delta (aq 1588).....>
 k d k l l r n f f l s p q i d r n a k v

2581 agccctcgag tcaacttgoga agaagtacga cgttccgaag gaagttctcg aagttctcga
 tcgggagctc agtgaacgct tcttcatgct gcaaggcttc cttcaagagc ttcaagagct
 >.....Subunit delta (aq_1588).....>
 k a l e s l a k k y d v p k e v l e v l

2641 gtacctcata gatataaacg ccatggctct tattccggag ataaagagac tatacgaatt
 catggagtat ctatatattgc ggtaccgaga ataaggcttc tatttctctg atatgcttaa
 >.....Subunit delta (aq 1588).....>
 e y l i d i n a m a l i p e i k r l y e

2701 agaactcgaa aagctcatgg gaatgcttaa aggggaactc atacttgcaa agaaaccag
 tcttgagctt ttcgagtagc cttacgaatt tccccttgag tatgaacgct tctttgggtc
 >.....Subunit delta (aq 1588).....>
 l e l e k l m g m l k g e l i l a k k p

2761 taaaaaactc ctagaaaaga ttacaaagac cataaacgat atcctaaaca gacagataga
 attttttgag gatcttttct aatgtttctg gtatttgcta taggatttgt ctgtctatct
 >.....Subunit delta (aq 1588).....>
 s k k l l e k i t k t i n d i l n r q i

SP5_new ▶

GGAGGACC CTTCCCTTAT AG

2821 aattgaagta aaggaggacc cttcccttat aggtggtttt gtcttcaaga cgcaggcttt
 ttaacttcat ttctcctgga gaagggaata tccaccaaaa cagaagttct gcgtccgaaa
 >.....Subunit delta (aq 1588).....>
 e i e v k e d p s l i g g f v f k t q a

```

                                                                    [S5] Primer
                                                                    GTGTTTA
2881  cgttctggac acttctgtta aaaccagct tgaaaaactc gcaagagttg gaggtgttta
      gcaagacctg tgaagacaat tttgggtcga actttttgag cgttctcaac ctccacaaat
      >.....Subunit delta (aq 1588).....>
      f v l d t s v k t q l e k l a r v g g v

      BamHI
      [S5] Primer ▶
      AATGGCCGTA GGG
2941  aatggccgta gggatccaaa cctttaaaga aggttaggag gtagagtatg gctacactga
      ttaccggcat ccctaggttt ggaaatttct tccaatctc catctcatic cgatgtgact
      > Subunit delta (aq 1588)
      -
      Subunit alpha (aq 679) >>.....>
      m a t l

3001  cttatgagga agcccttgag atactaagac aacagataaa ggatttcgaa cctgaagcca
      gaatactcct tcgggaactc tatgattctg ttgtctatct cctaaagctt ggacttcggt
      >.....Subunit alpha (aq 679).....>
      t y e e a l e i l r q q i k d f e p e a

3061  aatggaaga agtaggtgta gtttactacg tcgggtgatgg tgtagcaagg gottacggtc
      ttaccttctt tcatccacat caaatgatgc agccactacc acatcgtcc cgaatgccag
      >.....Subunit alpha (aq_679).....>
      k m e e v g v v y y v g d g v a r a y g

3121  ttgaaaacgt aatggcgatg gaaatagtag agtttcaggg agggcaacag ggaatagcct
      aacttttgca ttaccgctac ctttatcatic tcaaagtccc tcccgttgtc cttatcgga
      >.....Subunit alpha (aq 679).....>
      l e n v m a m e i v e f q g g q q g i a

3181  tcaacctcga agaggacaac gttggtatca taatcctcgg ttctgaaacg ggaatagaag
      agttggagct tctcctgttg caaccatagt attaggagcc aagactttgc cttatcttc
      >.....Subunit alpha (aq 679).....>
      f n l e e d n v g i i i l g s e t g i e

3241  aagggcacat agtaaagaga acgggcagga tattggacgc tcccgttga gaaggactcg
      ttcccgtgta tcatttctct tgcccgtcct ataacctgag agggcaacct ctctctgagc
      >.....Subunit alpha (aq_679).....>
      e g h i v k r t g r i l d a p v g e g l

3301  ttggaagggt tatcgacccc ctcgaaacc cctcgtatgg taaaggacc attcagtttg
      aaccttccca atagctgggg gagcctttgg gggagctacc atttctctgg taagtcaaac
      >.....Subunit alpha (aq 679).....>
      v g r v i d p l g n p l d g k g p i q f

3361  aataccgttc cccagttgaa aagatcgcac ccggtgttgt aaagagaaa cccgttcacg
      ttatggcaag ggtcaactt ttctagcgtg ggccacaaca tttctctttt gggcaagtgc
      >.....Subunit alpha (aq 679).....>
      e y r s p v e k i a p g v v k r k p v h

3421  aacccttca aacaggtatt aaagctatag acgctatgat tccaatagga aggggacaga
      ttggggaagt ttgtccataa tttogatatc tgcgatacta agttatcct tcccctgtct
      >.....Subunit alpha (aq_679).....>
      e p l q t g i k a i d a m i p i g r g q

      SP6_new ▶
      AGACCAC TGTTCGATA GAC
3481  gagagcttat catcggtgac agggctacgg gtaagaccac tgttgcgata gacaccatac
      ctctcgaata gtagccactg tccgatgcc cattctgggtg acaacgctat ctgtgggatg
      >.....Subunit alpha (aq_679).....>
      r e l i i g d r a t g k t t v a i d t i

```

SP6
GG

3541 tcgctcaaaa gaacagtgat gtttactgta ttacgtagc cgtaggacag aaaagagcgg
 agcgagtttt cttgtcacta caaatgacat aaatgcatcg gcatcctgtc ttttctcgcc
 >.....Subunit alpha (aq 679).....>
 l a q k n s d v y c i y v a v g q k r a

SP6 ▶

CGATAGCGAG ACTCATTG

3601 cgatagcgag actcattgag ctccttgaaa gagaaggagc tatggaatac accacagttg
 gctatcgctc tgagtaactc gaggaacttt ctcttctcgc atacccttatg tgggtgcaac
 >.....Subunit alpha (aq 679).....>
 a i a r l i e l l e r e g a m e y t t v

3661 ttgtagcttc agcatcagat cccgcatcac tccagtacct cgcacccttt gttggatgta
 aacatcgaag tcgtagtcta gggcgtagtg aggtcatgga gcgtgggaaa caacctacat
 >.....Subunit alpha (aq 679).....>
 v v a s a s d p a s l q y l a p f v g c

3721 cgatagggga gtacttcaga gacaacggaa agcacgcact catcatatac gacgacctgt
 gctatccctt catgaagtct ctgttgccctt tcgtgcgtga gtagtatatg ctgctggaca
 >.....Subunit alpha (aq 679).....>
 t i g e y f r d n g k h a l i i y d d l

3781 ccaagcacgc ggaagcttac agacagctct cactcctcat gagaagacct ccggtagag
 ggttcgtgcg ccttcgaatg tctgtcgaga gtgaggagta ctcttctgga gggccatctc
 >.....Subunit alpha (aq 679).....>
 s k h a e a y r q l s l l m r r p p g r

3841 aagcttacc cggtgacgtg ttctacctcc actcaagact ccttgaaaga gctgcaaaac
 ttcgaatggg gccactgcac aagatggagg tgagttctga ggaactttct cgacgttttg
 >.....Subunit alpha (aq_679).....>
 e a y p g d v f y l h s r l l e r a a k

3901 ttaacgacga cctcggggca ggttctctca cggcattgcc cataattgaa acgaaagcgg
 aattgctgct ggagccccgt ccaagagagt gccgtaacgg gtattaactt tgctttcgcc
 >.....Subunit alpha (aq 679).....>
 l n d d l g a g s l t a l p i i e t k a

3961 gtgacgtcgc ggcttacatt cccacgaacg ttatctccat tacagacgga cagatatacc
 cactgcagcg ccgaatgtaa ggggtgcttc aatagagga atgtctgcct gtctatatgg
 >.....Subunit alpha (aq 679).....>
 g d v a a y i p t n v i s i t d g q i y

4021 tcgaagcggga cctcttcaac aaaggtataa gacctgctat taacgtaggt ctttcggttt
 agcttcgcct ggagaagttg tttccatatt ctggacgata attgcatcca gaaagccaaa
 >.....Subunit alpha (aq_679).....>
 l e a d l f n k g i r p a i n v g l s v

4081 ccagagtcgg tgggtcggca cagataaagg ctatgaaaca ggttgcgggga accctcagac
 ggtctcagcc accacgcgct gtctatttcc gatactttgt ccaacgcctt tgggagtcgt
 >.....Subunit alpha (aq 679).....>
 s r v g g a a q i k a m k q v a g t l r

4141 tcgaacttgc tcagttcaga gaacttgaag ctttcgttca gttcgttcg gaacttgata
 agcttgaacg agtcaagtct cttgaacttc gaaagcaagt caagcgaagc cttgaactat
 >.....Subunit alpha (aq 679).....>
 l e l a q f r e l e a f v q f a s e l d

SP7
AGAAC

4201 aggcaaccca gcaacaaatc aacagaggtc tgagactcgt agaactcctg aagcaagaac
 tccgttgggt cgttgtttag ttgtctccag actctgagca tcttgaggac ttctgtcttg
 >.....Subunit alpha (aq 679).....>
 k a t q q q i n r g l r l v e l l k q e

SP7 ▶

CCTACAACCC GATAC

```

4261 cctacaaccc gatacccgtt gaaaaacaaa tcgttctcat atacgccgga acgcacggat
ggatgttggg ctatgggcaa ctttttgttt agcaagagta tatgcggcct tgcgtgccta
>.....Subunit alpha (aq 679).....>
  p y n p i p v e k q i v l i y a g t h g

4321 acctcgacga cattcccgta gagtctgtaa gaaagtttga aaaggaactc tacgcttacc
tgagctgct gtaagggcat ctacagacatt ctttcaaact tttccttgag atgcgaatgg
>.....Subunit alpha (aq 679).....>
  y l d d i p v e s v r k f e k e l y a y

4381 tagacaacga aagaccggac atactcaagg agataagtga aaagaagaaa ctcgacgaag
atctgttgc tttctggcctg tatgagttcc tctattcact tttctcttt gagctgcttc
>.....Subunit alpha (aq 679).....>
  l d n e r p d i l k e i s e k k k l d e

4441 aactagagaa gaagataaaa gaggggctcg acgccttcaa gcaaaagttc gttccctaac
ttgatctctt cttctatctt ctccgcgagc tgcggaagtt cgttttcaag caagggattg
>.....Subunit alpha (aq 679).....>>
  e l e k k i k e a l d a f k q k f v p -

4501 tctccctcct ctagatttag acattagttt ataataagta gcgttatggc gaaactttct
agagggagga gatctaaatc tgtaatcaaa tattattcat cgcaataccg ctttgaaaga
      Subunit gamma (aq 2041) >>.....>
                                     m a k l s

4561 cccagggaca taaagagaaa gatacagggg ataaagaaca cgaagagaat aacgaacgcg
gggtccctgt atttctcttt ctatgtccct tatttcttgt gcttctctta ttgcttgccg
>.....Subunit gamma (aq 2041).....>
  p r d i k r k i q g i k n t k r i t n a

4621 atgaaagtcg tttccgccc aaaactcagg aaagctcagg aactcgitta cgcttcccgt
tactttcagc aaaggcggcg ttttgagtcc tttcgagtcc ttgagcaaat gcgaagggca
>.....Subunit gamma (aq 2041).....>
  m k v v s a a k l r k a q e l v y a s r

4681 ccctactcgg agaaactcta cgaactcgtg ggacatctcg ctgccatgt ggacacggaa
gggatgagcc tctttgagat gcttgagcat cctgtagagc gacgggtaca cctgtgcctt
>.....Subunit gamma (aq 2041).....>
  p y s e k l y e l v g h l a a h v d t e

4741 gataatcccc tctttgacgt gaggggaagaa agaaacgttg acgttatcct cgttaccgca
ctattagggg agaaactgca ctcccttctt tctttgcaac tgcaatagga gcaatggcgt
>.....Subunit gamma (aq_2041).....>
  d n p l f d v r e e r n v d v i l v t a

4801 gacaggggtc tcgcccggagc tttcaattca aacgtaatca gaacagcgga aaatttaata
ctgtcccag agcgcctcog aaagttaagt ttgcattagt cttgtcgcct tttaaattat
>.....Subunit gamma (aq 2041).....>
  d r g l a g a f n s n v i r t a e n l i

4861 agggagaagg aagaaaaggg tgtaaggtt agccttatac ttgtggggag aaagggcttt
tcctcttcc ttcttttccc acaattccaa tcggaatag aacaccctc tttcccga
>.....Subunit gamma (aq 2041).....>
  r e k e e k g v k v s l i l v g r k g f

4921 cagtacttta cgaagagggg atacaacgta ataaaggat acgatgaagt ttttagaaag
gtcatgaaat gcttctccc tatgttgc ttttcccta tgctacttca aaaatcttc
>.....Subunit gamma (aq 2041).....>
  q y f t k r g y n v i k g y d e v f r k

```

XbaI

```

SP8
AAAC
4981 accgtaaact tcaatgtggc taaagaggtg gcagaaatag taaaggagag gttcttaaac
    tggcatttga agttacaccg atttctccac cgtctttatc atttcctctc caagaatttg
>.....Subunit gamma (aq 2041).....>
    t v n f n v a k e v a e i v k e r f l n

SP8 ▶
GGAGAAACCG ATAGGG
5041 ggagaaaccg ataggggttta cttgataaac aacgagatgg tcacgagggc gagctacaaa
    cctctttggc tatcccaaat gaactatttg ttgctctacc agtgctcccg ctcgatggtt
>.....Subunit gamma (aq 2041).....>
    g e t d r v y l i n n e m v t r a s y k

5101 cctcaggtaa gggcttctct gccttttgaa gcccaagaaa aagaagtgga agagcttggga
    ggagtccatt cccagaagga cggaaaactt cgggttcttt ttcttcacct tctcgaacct
>.....Subunit gamma (aq 2041).....>
    p q v r v f l p f e a q e k e v e e l g

5161 acttacgagt ttgaagtctc agaagaggag ttctttgact acatagtata cctgtacctt
    tgaatgctca aacttcagag tcttctcctc aagaaactga tgtatcattt ggacatggaa
>.....Subunit gamma (aq 2041).....>
    t y e f e v s e e e f f d y i v n l y l

5221 aactaccagg tatacagggc tatgggttgag tccaacgcgg cggagcactt cgogaggatg
    ttgatggtcc atatgtcccg ataccaactc aggttgccgc gcctcgtgaa gcgctcctac
>.....Subunit gamma (aq 2041).....>
    n y q v y r a m v e s n a a e h f a r m

5281 atagcgatgg acaacgcaac caagaacgca gaggaccta taaggcagtg gaccctcgtg
    tatcgctacc tgttgcggtg gttcttgctc ctcttgatt attcgcacac ctgggagcac
>.....Subunit gamma (aq 2041).....>
    i a m d n a t k n a e d l i r q w t l v

5341 ttcaacaagg caaggcagga agctattaca accgaactta tagatataac caacgctggt
    aagttgttcc gttccgtcct tcgataatgt tggcttgaat atctatattg gttgogacaa
>.....Subunit gamma (aq 2041).....>
    f n k a r q e a i t t e l i d i t n a v

5401 gaagctctta aagcacaata aaggaggttt atagatgaga ggatcgcac atcatcatca
    cttcgagaat ttctgtgtat ttcttccaaa tatctactct cctagcgtag tagtagtagt
>.....>> Subunit gamma (aq 2041)
    e a l k a q -
>>.Subunit beta (aq 2038)..>
m r g s h h h h

5461 tcatggtatg gcggaagtga ttaagggaaa ggtagttcag gtcataaggac ccggttgtga
    agtaccatac cgccttcaat aattcccttt ccatcaagtc cagtatcctg ggcaacaact
>.....Subunit beta (aq 2038).....>
    h h g m a e v i k g k v v q v i g p v v

5521 cgtggagttc gaaggggtaa aggaacttcc caaatcaaa gacggtctca aaacaataag
    gcacctcaag cttccccatt tccttgaagg gttttagttt ctgccagagt tttgttattc
>.....Subunit beta (aq 2038).....>
    d v e f e g v k e l p k i k d g l k t i

5581 aagagctata gacgacagag gtaactgggt cgaagaagta ctcttcatgg aagtggcaca
    ttctcgatat ctgctgtctc cattgaccaa gcttcttcat gagaagtacc ttcaccgtgt
>.....Subunit beta (aq 2038).....>
    r r a i d d r g n w f e e v l f m e v a

5641 gcacataggg gagcacagag taaggcgat agcgatgggt ccaacagatg gtctcgtgag
    cgtgtatccc ctctgtctc attcccgcta tcgctaccca gttgtctac cagagcactc
>.....Subunit beta (aq 2038).....>
    q h i g e h r v r a i a m g p t d g l v

```

5701 aggtcaagaa gttgagtact tggggggacc cataaagata cccgtaggta aggaagttct
tccagttctt caactcoatga acccccctgg gtatttctat gggcatccat tocttcaaga
>.....Subunit beta (aq 2038).....>
r g q e v e y l g g p i k i p v g k e v

SP9 ▶

TTG CGGGACAACC CATAGAC

5761 cggaaggata ttcaacgttg cgggacaacc catagacgag caggggccgg tagaggctaa
gccttcctat aagttgcaac gccctgttgg gtatctgctc gtcccaggcc atctccgatt
>.....Subunit beta (aq 2038).....>
l g r i f n v a g q p i d e q g p v e a

5821 agagtactgg cccatgttca gaaatccacc cgaactcgtt gaacaatcca cgaaagttga
tctcatgacc gggtagaagt ctttaggtgg gcttgagcaa cttgttaggt gctttcaact
>.....Subunit beta (aq 2038).....>
k e y w p m f r n p p e l v e q s t k v

5881 aattcttgaa acgggtatta aagttataga cctcctccag cccatcataa agggtggtaa
ttaagaactt tgcccataat ttcaatatct ggaggaggtc ggtagtatt tcccaccatt
>.....Subunit beta (aq 2038).....>
e i l e t g i k v i d l l q p i i k g g

5941 ggttgactc ttggcggtg cgggagttgg aaagaccggt ctcatgcagg agctcatcca
ccaacctgag aagccgccac gccctcaacc tttctggcaa gactacgtcc togagtaggt
>.....Subunit beta (aq_2038).....>
k v g l f g g a g v g k t v l m q e l i

6001 caacatagcg cgtttccaacg aagggtatc cgttgcggtt ggagtggcg agagaacaag
gttgatcgc gcaaagggtc ttcccataag gcaacagcaa cctcaccgc tctcttgttc
>.....Subunit beta (aq 2038).....>
h n i a r f h e g y s v v v g v g e r t

6061 agaaggaac gacctctggc tggaaatgaa ggagtcgga gttctccctt acacggttat
tcttctttg ctggagaccg agctttactt cctcaggcct caagaggaa tgtgccaata
>.....Subunit beta (aq 2038).....>
r e g n d l w l e m k e s g v l p y t v

6121 ggtttacgga cagatgaacg agcctccggg agttaggttc agggtgccac acaccggact
ccaatgcct gtctacttgc toggaggccc tcaatccaag tcccaccgtg tgtggcctga
>.....Subunit beta (aq_2038).....>
m v y g q m n e p p g v r f r v a h t g

6181 tacaatggct gagtacttca gagacgtgga aggtcaggac gttctcatat tcatagacia
atgttaccga ctcatgaagt ctctgcacct tccagtctg caagagtata agtatctgtt
>.....Subunit beta (aq 2038).....>
l t m a e y f r d v e g q d v l i f i d

6241 catattcagg ttctgtcagg caggtgcgga agtttcaacg cttcttgaa gactaccctc
gtataagtcc aagcaagtcc gtccacgcct tcaaagttgc gaagaacctt ctgatgggag
>.....Subunit beta (aq 2038).....>
n i f r f v q a g a e v s t l l g r l p

6301 cgcggttgggt taccagcca cctcaatac tgacgtcggg gaagttcagg aaagaattac
gcgccaacca atggtcgggt gggagttatg actgcagcca cttcaagtcc tttcttaatg
>.....Subunit beta (aq_2038).....>
s a v g y q p t l n t d v g e v q e r i

6361 ttcaaccaag aaagttteta ttacogcaat tcaggcgggt tacgttcccg cagaogacat
aagttggttc tttccaagat aatggcggtta agtccgcaa atgcaagggc gtctgctgta
>.....Subunit beta (aq 2038).....>
t s t k k g s i t a i q a v y v p a d d

SP10 ▶

GGT CCATATTCGC GCACCTC

6421 aactgacccc gcaccttggg ccatattcgc gcacctcgac gctacgaccg ttctcacaag

```

6481 aagactcgct gagctcggaa tatatcccg c aatagatccc ctcgaatcta catctaagta
ttctgagcga ctcgagcctt atatagggcg ttatctaggg gagcttagat gtagattcat
>.....Subunit beta (aq 2038).....>
r r l a e l g i y p a i d p l e s t s k

6541 cctcgctccc gagtatgtcg gagaagagca ctacgaagtt gcaatggaag taaagaggat
ggagcgaggg ctcatacagc ctcttctcgt gatgcttcaa cgttaccttc atttctccta
>.....Subunit beta (aq 2038).....>
y l a p e y v g e e h y e v a m e v k r

6601 tctccaaagg tacaagaac ttcaagaat catcgcaatc ctcggtatgg aagaactctc
agaggttcc atgtttcttg aagttcttta gtagcgttag gagccatacc ttcttgagag
>.....Subunit beta (aq 2038).....>
i l q r y k e l q e i i a i l g m e e l

6661 tgacgaggac aaggctatcg ttaacagggc gagaagaata cagaagttcc tctctcaacc
actgctcctg ttccgatagc aattgtcccg ctcttcttat gtcttcaagg agagagttgg
>.....Subunit beta (aq_2038).....>
s d e d k a i v n r a r r i q k f l s q

6721 cttccacggt gcgagcaat ttacaggtat gcccgtaag tacgtaaaac togaggatac
gaaggtgcaa cgctcgtta aatgtccata cgggccattc atgcattttg agctcctatg
>.....Subunit beta (aq 2038).....>
p f h v a e q f t g m p g k y v k l e d

6781 cataaggtcc ttcaaggaag ttctcacagg aaagtacgac caccttcccg aaaacgcctt
gtattccagg aagttccttc aagagtgtcc ttctatgctg gtggaagggc ttttgcgaa
>.....Subunit beta (aq 2038).....>
t i r s f k e v l t g k y d h l p e n a

6841 ctacatggtg ggaacaatag aggacgttat agaaaaggca aaacaaatgg gggctaaagt
gatgtaccac ccttgttatc tcttgcaata tcttttccgt tttgtttacc cccgatttca
>.....Subunit beta (aq_2038).....>
f y m v g t i e d v i e k a k q m g a k

6901 SalI
ttaaagccct gtcgacttgg acttttctctg gtataattta gggattatga tacaggttga
aattcgggga cagctgaacc tgaagagac catattaaat ccctaatact atgtccaact
>.>> Subunit beta (aq_2038)
v -
Subunit epsilon (aq 673) >>.....>
m i q v

6961 aatagtttct ccgcagggaa tggtttactc gggagaagta gagagcgtaa acgtccccac
ttatcaaaga ggcgtccctt accaaatgag cctcttcat ctctcgcatt tgcaggggtg
>.....Subunit epsilon (aq_673).....>
e i v s p q g m v y s g e v e s v n v p

SP11 ►
TTGAAGGA GAGGTGGGAA TC

7021 ggttgaagga gaggtgggaa tccttgaaaa ccacatgtac ctgatgacct tcttgaaacc
ccaacttctt ctccaccctt aggaactttt ggtgtacatg gactactggg agaactttgg
>.....Subunit epsilon (aq 673).....>
t v e g e v g i l e n h m y l m t l l k

7081 tggacttgtt tacttcaacg gtgacgacaa aaacggaata gctgtaacct acggcgttct
acctgaacaa atgaagttgc cactgctgtt tttgccttat cgacattgga tgccgcaaga
>.....Subunit epsilon (aq 673).....>
p g l v y f n g d d k n g i a v t y g v

7141 ggacgtcacc ccccaaaagg ttctcattct tgcggaagaa gcttacgaag tcggaaaact
cctgcagtgg ggggttttcc aagagtaaga acgccttctt cgaatgcttc agccttttga
>.....Subunit epsilon (aq 673).....>
l d v t p q k v l i l a e e a y e v g k

```



```

7201  tcctccagca agcaagctaa aagaagagtt tgaagaagcc gtgaagaaaa tggcaaccgc
aggaggtcgt tcgttcgatt ttcttctcaa acttcttcgg cacttctttt accgttggcg
>.....Subunit epsilon (aq 673).....>
l p p a s k l k e e f e e a v k k m a t

7261  ccaaaactatg gaagagttaa aagagtggga gaaggaagca gaaaaggcaa gaactctctt
ggtttgatac cttctcaatt ttctcaccct cttccttcgt cttttccggt cttgagagaa
>.....Subunit epsilon (aq 673).....>
a q t m e e l k e w e k e a e k a r t l

7321  agaactcgtt gaaaagtaca gataactcaa aaccgcgtccc cttctttcgg ggaaaaccog
tcttgagcaa cttttcatgt ctattgagtt ttgggcaggg gaagaaagcc ctttttgggc
>.....>> Subunit epsilon (aq 673)
l e l v e k y r -

7381  ggctgcaggc atgcaagctt ggctgttttg gcggatgaga gaagattttc agcctgatac
cgcaggtccg tacgttcgaa ccgacaaaac cgcctactct cttctaaaag tcggactatg

7441  agattaaatc agaacgcaga agcgggtctga taaaacagaa ttgacctggc ggcagtagcg
tctaatttag tcttgctctc tcgccagact attttgtctt aaacggaccg cgcctatcgc

7501  cgggtgtccc acctgacccc atgccgaact cagaagtga aacccgtagc gccgatggta
gccaccaggg tggactgggg tacggcttga gtcttcactt tgcggcatcg oggctaccat

7561  gtgtggggtc tccccatgcg agagtaggga actgccaggc atcaaataaa acgaaaggct
cacaccccag aggggtacgc tctcatcctc tgacgggtccg tagtttattt tgctttccga
rrnB-terminator

7621  cagtcgaaag actgggcctt tcgttttatc tgtttgtttg cggtgaacgc tctcctgagt
gtcagctttc tgacccggaa agcaaaatag acaacaaaca gccacttgcg agaggactca
rrnB_T1_terminator

7681  aggacaaatc gcgccgggagc ggatttgaac gttgcgaagc aacggcccgg aggggtggcgg
tctgttttag gcggccctcg octaaacttg caacgcttcg ttgccgggccc tcccaccgcc

7741  gcaggacgcc gccataaac tgccaggcat caaattaagc agaaggccat cctgacggat
cgtcctgcgg gcggtatttg acggtcgta gtttaattcg tcttccggta ggactgccta
rrnB_T2_terminator

7801  ggcttttttg cgtttctaca aactcttttt gtttattttt ctaaatacat tcaaataatgt
ccggaaaaac gcaaagatgt ttgagaaaaa caaataaaaa gatttatgta agtttataca

7861  atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaa aggaagagta
taggcgagta ctctgttatt gggactatct acgaagttat tataactttt tccttctcat
Ampicillin promoter

7921  tgagtattca acatttcogt gtgcgcctta ttcccttttt tgccgcattt tgcttctctg
actcataagt tgtaaaggca cagcgggaat aagggaaaaa acgccgtaaa acggaaggac
>.....Ampicillin.....>
m s i q h f r v a l i p f f a a f c l p

7981  tttttgctca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac
aaaaacgagt ggttctttgc gaccactttc attttctacg acttctagtc aacccacgtg
>.....Ampicillin.....>
v f a h p e t l v k v k d a e d q l g a

8041  gagtgggta catcgaactg gatctcaaca gcggtaagat ccttgagagt tttcgccccg

```

SmaI

XmaI

PstI

Ampicillin >

Appendix

```
8101 aagaacgttt tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc
ttcttgcaaa aggttactac tcgtgaaaat ttcaagacga tacaccgcgc cataataggg
>.....Ampicillin.....>
e e r f p m m s t f k v l l c g a v l s

8161 gtgttgacgc cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg
cacaactgcg gcccgttctc gttgagccag cggcgtatgt gataagagtc ttaactgaacc
>.....Ampicillin.....>
r v d a g q e q l g r r i h y s q n d l

8221 ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat
aactcatgag tggtcagtgt cttttcgtag aatgcctacc gtactgtcat tctcttaata
>.....Ampicillin.....>
v e y s p v t e k h l t d g m t v r e l

8281 gcagtgtctc cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg
cgtcacgacg gtattggtac tcaactattgt gacgcgggtt gaatgaagac tgttgctagc
>.....Ampicillin.....>
c s a a i t m s d n t a a n l l l t t i

8341 gaggaccgaa ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg
ctcctggctt cctcgattgg cgaaaaaacg tgttgtaacc cctagtacat tgagcggaac
>.....Ampicillin.....>
g g p k e l t a f l h n m g d h v t r l

8401 atcgttgggg accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc
tagcaaccct tggcctcgac ttacttcggt atggtttgc tctcgcactg tgggtgctagc
>.....Ampicillin.....>
d r w e p e l n e a i p n d e r d t t m

8461 ctacagcaat ggcaacaacg ttgocgcaac tattaactgg cgaactactt actctagctt
gatgtcgta cggttgttgc aacgcgttg ataattgacc gcttgatgaa tgagatcgaa
>.....Ampicillin.....>
p t a m a t t l r k l l t g e l l t l a

8521 cccggcaaca attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgt
gggccgttgt taattatctg acctacctcc gcctatttca acgtcctggt gaagacgcga
>.....Ampicillin.....>
s r q q l i d w m e a d k v a g p l l r

8581 cggcccttcc ggctggctgg tttattgctg ataatctgg agccggtgag cgtgggtctc
gccgggaagg ccgaccgacc aaataacgac tatttagacc tcggccactc gcaccagag
>.....Ampicillin.....>
s a l p a g w f i a d k s g a g e r g s

8641 gcggtatcat tgcagcactg gggccagatg gtaagccctc cgtatcgta gttatctaca
cgccatagta acgtcgtgac cccggtctac cattcgggag ggcatagcat caatagatgt
>.....Ampicillin.....>
r g i i a a l g p d g k p s r i v v i y

8701 cgacggggag tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgctt
gctgccctc agtcogttga tacctacttg ctttatctgt ctacgcactc tatccacgga
>.....Ampicillin.....>
t t g s q a t m d e r n r q i a e i g a

8761 cactgattaa gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt
gtgactaatt cgtaaccatt gacagtctgg ttcaaatgag tataatatgaa atctaactaa
>....Ampicillin....>>
s l i k h w -

8821 taaaacttca tttttaattt aaaaggatct aggtgaagat cttttttgat aatctcatga
attttgaagt aaaaattaaa ttttctaga tccacttcta ggaaaaacta ttagagtact

8881 ccaaaatccc ttaacgtgag ttttcgctcc actgagcgtc agaccccgta gaaaagatca
ggttttaggg aattgcactc aaaagcaagg tgactcgcag tctggggcat cttttctagt
```

```

9001  caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg
      gtggcgatgg tcgccaccaa acaaacggcc tagttctcga tgggtgagaa aaaggcttcc

9061  taactggcct cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag
      attgaccgaa gtcgtctcgc gtctatggtt tatgacagga agatcacatc ggcatcaatc

9121  gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac
      cggtggtgaa gttcttgaga catcgtggcg gatgtatgga gcgagacgat taggacaatg

9181  cagtggctgc tgccagtggc gataagtctg gtcttaccgg gttggactca agacgatagt
      gtcaccgacg acggtcaccc ctattcagca cagaatggcc caacctgagt tctgctatca

9241  taccggataa ggcgcagcgg tggggtgaa cgggggggtc gtgcacacag cccagcttgg
      atggcctatt ccgcgtcgcg agcccgactt gcccccaag cacgtgtgtc gggtcgaacc

9301  agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcggccacgc
      tcgcttgctg gatgtggctt gactctatgg atgtcgcact cgatactctt tcgcggtgcg

9361  ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcggg acaggagagc
      aagggttcc ctctttccgc ctgtccatag gccattcgcc gtcccagcct tgtcctctcg

9421  gcacgagggg gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc
      cgtgctccct cgaaggctcc cctttgcgga ccatagaaat atcaggacag cccaaagcgg

9481  acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa
      tggagactga actcgcagct aaaaacacta cgagcagtcc cccgcctcg gatacccttt

9541  acgccagcaa cgcggccttt ttaocggtcc tggccttttg ctggcctttt gctcacatgt
      tgcggtcgtt gcgccgaaaa aatgcccaagg accggaaaaac gaccggaaaa cgagtgtaca

9601  tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg
      agaaaggacg caatagggga ctaagacacc tattggcata atggcggaaa ctcaactcgac

9661  ataccgctcg ccgcagccga acgaccgagc gcagcagatc agtgagcgag gaagcgggaag
      tatggcgagc ggcgtcggct tgctggctcg cgtcgctcag tcaactcgctc cttcgccttc

9721  agcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac cgcatatggg
      tcgcggaacta cgccataaaa gaggaatgcg tagacacgcc ataaagtgtg gcgatatacca
      ↓
      NdeI

9781  gcaactctcag tacaatctgc tctgatgccg catagttaag ccagatataca ctccgctatc
      cgtgagagtc atgttagacg agactacggc gtatcaattc ggtcatatgt gaggcगतग

9841  gctacgtgac tgggtcatgg ctgogccccg acaccgcga acaccgcgtg acgcgcctg
      cgatgcactg acccagtacc gacgcggggc tgtgggcggt tgtgggcgac tgcgcgggac

9901  acgggcttgt ctgctcccgg catocgotta cagacaagct gtgaccgtct ccgggagctg
      tgcccgaaca gacgagggcc gtaggcgaat gtctgttcga cactggcaga ggccctcgac

9961  catgtgtcag aggtttttcac cgtcatcacc gaaacgcgcg aggcagcaga tcaattcgcg
      gtacacagtc tccaaaagtg gcagtagtgg ctttgcgcgc tccgtcgtct agttaagcgc

10021 cgcgaaggcg aagcggcatg catttacggt gacaccatcg aatggtgcaa aacctttcgc
      gcgcttccgc ttcgccgtac gtaaagtcaa ctgtggtagc ttaccacggt ttggaaagcg

10081 ggtatggcat gatagcgcc ggaagagagt caattcaggg tggatgaatgt gaaaccagta
      ccataccgta ctatcgcggg ccttctctca gttaagtccc accacttaca ctttggteat
      >>.....lacI.....>
      v v n v k p v

10141 acgttatacg atgtcgcaga gtatgccggt gtctcttata agaccgtttc ccgcgtgggtg
      tgcaatatgc tacagcgtct catacggcca cagagaatag tctggcaaaag ggcgcaccac
      >.....lacI.....>
      t l y d v a e y a g v s y q t v s r v v

```

Appendix

```
10201 aaccaggcca gccacgtttc tgcgaaaacg cgggaaaaag tggaaagcggc gatggcggag
ttggtccggt cgggtgcaaag acgcttttgc gccctttttc accttcgccg ctaccgcctc
>.....lacI.....>
n q a s h v s a k t r e k v e a a m a e

10261 ctgaattaca ttccaaccg cgtggcacia caactggcgg gcaaacagtc gttgctgatt
gacttaaatgt aagggttggc gcaccgtgtt gttgaccgcc cgtttgtcag caacgactaa
>.....lacI.....>
l n y i p n r v a q q l a g k q s l l i

10321 ggcgttgcca cctccagtct ggccctgcac ggcctgcgc aaattgtcgc ggcgattaaa
ccgcaacggt ggaggtcaga ccgggacgtg cgcggcagcg ttaacagcg ccgctaattt
>.....lacI.....>
g v a t s s l a l h a p s q i v a a i k

10381 tctcgcgccg atcaactggg tgccagcgtg gtgggtgcga tggtagaacg aagcggcgtc
agagcgcggc tagttgacct acggtcgcac caccacagct accatcttgc ttgcgcccag
>.....lacI.....>
s r a d q l g a s v v v s m v e r s g v

10441 gaagcctgta aagcggcggg gcacaatctt ctgcgcgaac gcgtcagtgg gctgatcatt
cttcggacat ttgcgccca cgtgttagaa gagcgcgttg cgcagtcacc cgactagtaa
>.....lacI.....>
e a c k a a v h n l l a q r v s g l i i

10501 aactatccgc tggatgacca ggatgccatt gctgtggaag ctgcctgcac taatgttccg
ttgataggcg acctactggt cctacggtaa cgacacctc gacggacgtg attacaaggc
>.....lacI.....>
n y p l d d q d a i a v e a a c t n v p

10561 gcgttatttc ttgatgtctc tgaccagaca cccatcaaca gtattatttt ctcccatgaa
cgcaataaag aactacagag actgggtctgt gggtagtgtg cataataaaa gagggtactt
>.....lacI.....>
a l f l d v s d q t p i n s i i f s h e

10621 gacggtacgc gactgggctg ggagcatctg gtcgcattgg gtcaccagca aatcgcgctg
ctgccatgcg ctgaccgca cctcgtagac cagcgtaac cagtggctgt ttagcgcgac
>.....lacI.....>
d g t r l g v e h l v a l g h q q i a l

10681 ttagcgggcc cattaagtct tgtctcggcg cgtctgcgtc tggctggctg gcataaatat
aatcgcggcg gtaattcaag acagagccgc gcagacgag accgaccgac cgtatttata
>.....lacI.....>
l a g p l s s v s a r l r l a g w h k y

10741 ctcaactcga atcaaatca gccgatagcg gaacgggaag gcgactggag tgccatgtcc
gagtgagcgt tagtttaagt cggctatcgc cttgcccttc cgctgacctc acggtacagg
>.....lacI.....>
l t r n q i q p i a e r e g d w s a m s

10801 ggttttcaac aaaccatgca aatgctgaat gagggcatcg ttccaactgc gatgctgggt
ccaaaagttg tttggtagct ttacgactta ctcccgtagc aagggtgacg ctacgaccaa
>.....lacI.....>
g f q q t m q m l n e g i v p t a m l v

10861 gccaacgatc agatggcgtt gggcgcaatg cgcgccatta ccgagtcggg gctgcgcgtt
cgttgctag tctaccgca cccgcgttac ggcggtaat ggctcaggcc cgacgcgcaa
>.....lacI.....>
a n d q m a l g a m r a i t e s g l r v

10921 ggtgcggata tctcggtagt gggatacgc gataccgaag acagctcatg ttatatcccg
ccacgcctat agagccatca ccctatgctg ctatggcttc tctcagtagc aatatagggc
>.....lacI.....>
g a d i s v v g y d d t e d s s c y i p
```

```
10981 ccgtaacca ccatcaaaca ggattttcgc ctgctggggc aaaccagcgt ggaccgcttg
      ggcagttggt ggtagtttgt cctaaaagcg gacgaccccg tttggtcgca cctggcgaac
      >.....lacI.....>
      p s t t i k q d f r l l g q t s v d r l

11041 ctgcaactct ctcagggcca ggcgggtgaag ggcaatcagc tgttgcccgt ctcactggtg
      gacgttgaga gagtcccggg cggccacttc ccgtagtgcg acaacgggca gagtgaccac
      >.....lacI.....>
      l q l s q g q a v k g n q l l p v s l v

11101 aaaagaaaa ccaccctggc gcccaatagc caaacgcct ctccccgcgc gttggccgat
      tttcttttt ggtgggaccg cgggttatgc gtttggcgga gaggggcgcg caaccggcta
      >.....lacI.....>
      k r k t t l a p n t q t a s p r a l a d

11161 tcattaatgc agctggcacg acaggtttcc cgactggaaa gcgggcagtg a
      agtaattacg tcgaccgtgc tgtccaaagg gctgaccttt cgcccgtcac t
      >.....lacI.....>>
      s l m q l a r q v s r l e s g q -
```

DNA sequence of subunit c in the vector pCL21-ΔSP

```

1  ttgtgcttat attcgttata gctataaagt tcctcgccat attcatacag acttacatct
   aacacgaata taagcaatat cgatatttca aggagcggta taagtatgtc tgaatgtaga
   >>.....'aq 179.....>
      v l i f v i a i k f l a i f i q t y i

61  ttatgatact ctcggtggtt tacatagccg gagctgtagc acacgaggag cactgatatg
   aatactatga gagccaccaa atgtatcggc ctgcacatcg tgtgctcctc gtgactatac
   >.....'aq 179.....>>
      f m i l s v v y i a g a v a h e e h -

                                           Del2-19-Fw
                                           G CGGAAGGAGA

                               ◀ Del2-19-Rev
                               CGAAATTCC TCCATCCAC
                               Del_A2-19_Quickchange
                               GTTAAATA AGCTTTAAGG AGGTAGGTGG CGGAAGGAGA

121  aaggtactgc tagtttotta tagttaaata agctttaagg aggtaggtgg cggaaggaga
   ttccatgacg atcaaagaat atcaatttat tcgaaattcc tccatccacc gccttctctc
   Subunit c ('aq 177') >>>
                                           v
   Subunit c ('aq_177') >>.....>
                                           a e g

Del2-19-Fw ▶
GGCTTCC
Del_A2-19_Quickchange ▶
GGCTTCCGTA GCTAAGGG

181  ggcttcgcta gctaaggac ttctgtacct tggagcagga cttgctatag ggcttcgagg
   ccgaaggcat cgattcctcg aagacatgga acctcgtcct gaacgatatc ccgaacgtcc
   >.....Subunit c ('aq 177').....>
   e a s v a k g l l y l g a g l a i g l a

241  actcgggtgcc ggagttgta tgggtcatgc cgttagggga actcaggaag gtgttgcgag
   tgagccacg cctcaacat acccagtaag gcaatcccct tgagtccttc cacaaacgctc
   >.....Subunit c ('aq_177').....>
   g l g a g v g m g h a v r g t q e g v a

301  gaatcccaac gcaggcggaa gacttcaaac ccttatgttc ataggacttg cgtttataga
   cttagggttg cgtccgcctt ctgaagtttg ggaatacaag taccctgaac gcaaatatct
   >.....Subunit c ('aq 177').....>
   r n p n a g g r l q t l m f i g l a f i

361  aactatcgct ctttacggat tgctcatagc tttcatactg ctcttcgtgg tttaaaccc
   ttgatagcga gaaatgccta acgagtatcg aaagtatgac gagaagcacc aaattcggga
   >.....Subunit c ('aq 177').....>>
   e t i a l y g l l i a f i l l f v v -

BglII
421  tagatctatt gctataattg tttagcggag gagaagaatg gacataggag taatgcctaa
   atctagataa cgatattaac aaatgcctc ctcttcttac ctgtatcctc attacggatt
   >>.....aq 1586.....>
   m d i g v m p

481  tgcaacaatc ctcgttcaat tgttcatctt cgtaatatc ctaatgataa tcaactaacat
   acgttggttag gagcaagta acaagtagaa gcattataag gattactatt agtgattgta
   >.....aq_1586.....>
   n a t i l v q l f i f v i f l m i i t n

541  ctacgtaaag ccctacaccg cggatgataa atccagagaa gaactcatta agaagaacct
   gatgcatttc gggatgtggc gccactatct taggtctctt cttgagtaat tcttcttgga
   >.....aq 1586.....>
   i y v k p y t a v i e s r e e l i k k n

```

DNA sequence of subunit c of the vector pCL-MEN

721 tatcaagaat ccattcacgc tggtagtacc accggttggtg cttatattcg ttatagctat
 atagttctta ggtaagtgcg accatcatag tggccaacac gaatataagc aatatcgata
 >.....aq 179.....>
 v i k n p f t l v v s p v v l i f v i a

781 aaagttcctc gccatattca tacagactta catctttatg atactctcgg tggtttacat
 tttcaaggag cggataaagt atgtctgaat gtagaaatac tatgagagcc accaaatgta
 >.....aq 179.....>
 i k f l a i f i q t y i f m i l s v v y

subunit c_MEN_Quickchange

841 agccggagct gtagcacacg aggagcactg atatgaaggc actgctagtt tcttatagtt
 tcggcctcga catcgtgtgc tctcgtgac tatacttcca tgacgatcaa agaatatcaa
 >.....aq 179.....>>
 i a g a v a h e e h -

subunit c_MEN_Quickchange

AAATAAGCTT TAAGGAGGTA GATGGAAAAC CTGAATATGG ATCTTCTGTA CCTTGGAGCA

901 aaataagcct taaggagcta gatggaaaac ctgaatatgg atcttctgta ccttggagca
 ttattogcaa attcctccat ctaccttttg gacttatacc tagaagacat ggaacctcgt
 >>.....subunit c (aq 177) MENL.....>
 m e n l n m d l l y l g a

subunit c_MEN_Quickchange ►

GGACTT

961 ggacttgcta tagggcttgc aggactcggc gccggagttg gtatgggtca tgccgttagg
 cctgaacgat atcccgaacg tctgagcca cggcctcaac catacccagt acggcaatcc
 >.....subunit c (aq 177) MENL.....>
 g l a i g l a g l g a g v g m g h a v r

1021 ggaactcagg aagggtgttc gaggaatccc aacgcaggcg gaagacttca aacccttatg
 ccttgagtcc ttccacaacg ctcttagggg ttgcgtccgc cttctgaagt ttgggaatac
 >.....subunit c (aq 177) MENL.....>
 g t q e g v a r n p n a g g r l q t l m

1081 ttcataggac ttgcgtttat agaaactatc gctctttacg gattgctcat agctttcata
 aagtatctcg aacgcaaata tctttgatag cgagaaatgc ctaacgagta tcgaaagtat
 >.....subunit c (aq 177) MENL.....>
 f i g l a f i e t i a l y g l l i a f i

BglII

1141 ctgctcttcg tggtttaagc ccttagatct attgctataa ttgtttagcg gaggagaaga
 gacgagaagc accaaattcg ggaatctaga taacgatatt aacaaatcgc ctctcttct
 >.....>> subunit c (aq 177) MENL
 l l f v v -

1201 atggacatag gagtaatgcc taatgcaaca atcctcgttc aattgttcat cttcgtaata
 tacctgtatc ctcttaccg attacgttgt taggagcaag ttaacaagta gaagcattat
 >>.....aq 1586.....>
 m d i g v m p n a t i l v q l f i f v i

1261 ttctaataga taatcactaa catctacgta aagccctaca ccgcggtgat agaatccaga
 aaggattact attagtgatt gtagatgcat ttcgggatgt ggcgccacta tcttaggtct
 >.....aq 1586.....>
 f l m i i t n i y v k p y t a v i e s r

1321 gaagaactca ttaagaagaa cctctctgaa gcacaaaagt taagggaaga aactcaaacc
 cttcttgagt aattcttctt ggagagactt cgtgttttca attcccttct ttgagtttgg
 >.....aq 1586.....>
 e e l i k k n l s e a q k l r e e t q t

A.4. Antibodies generation

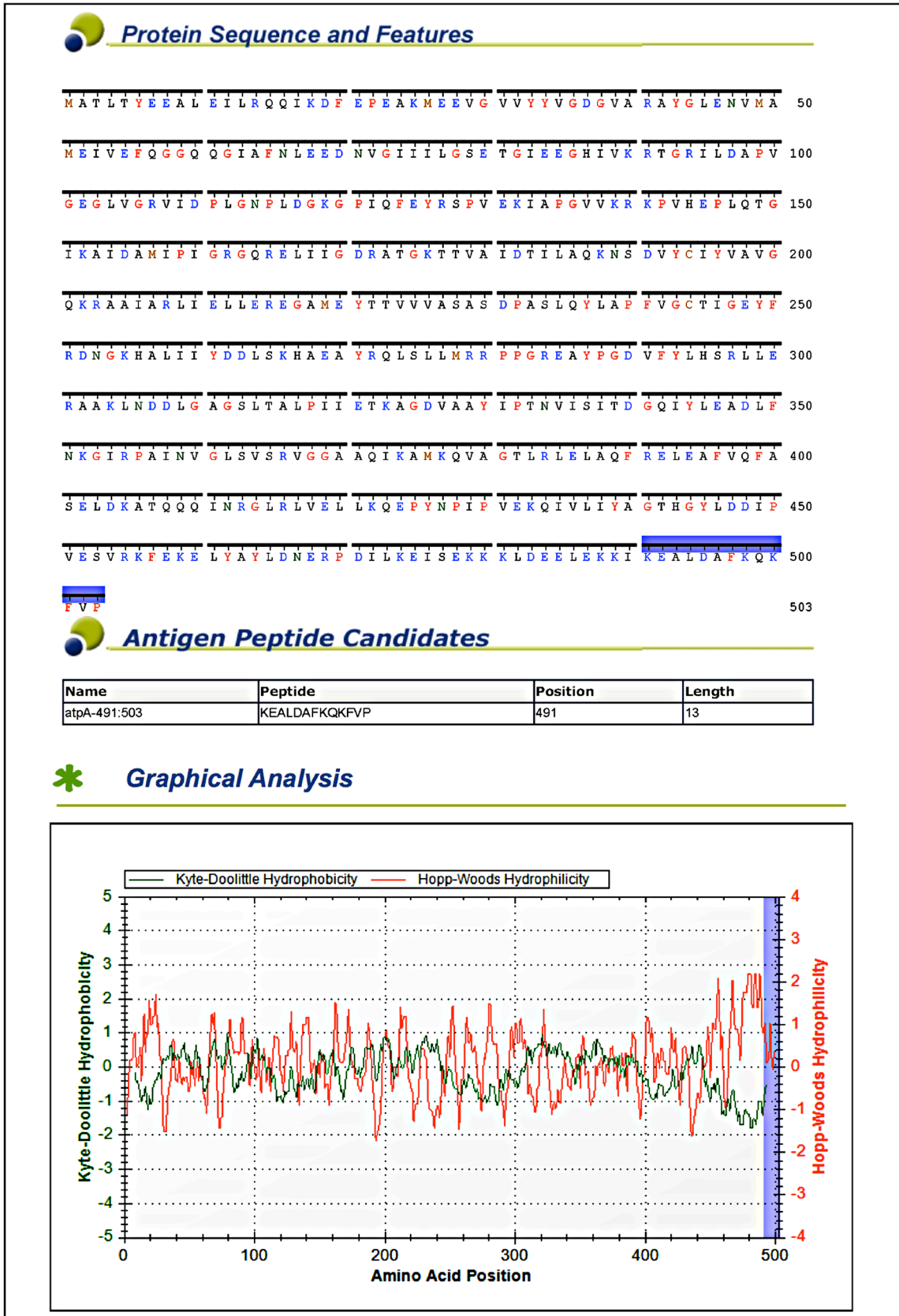
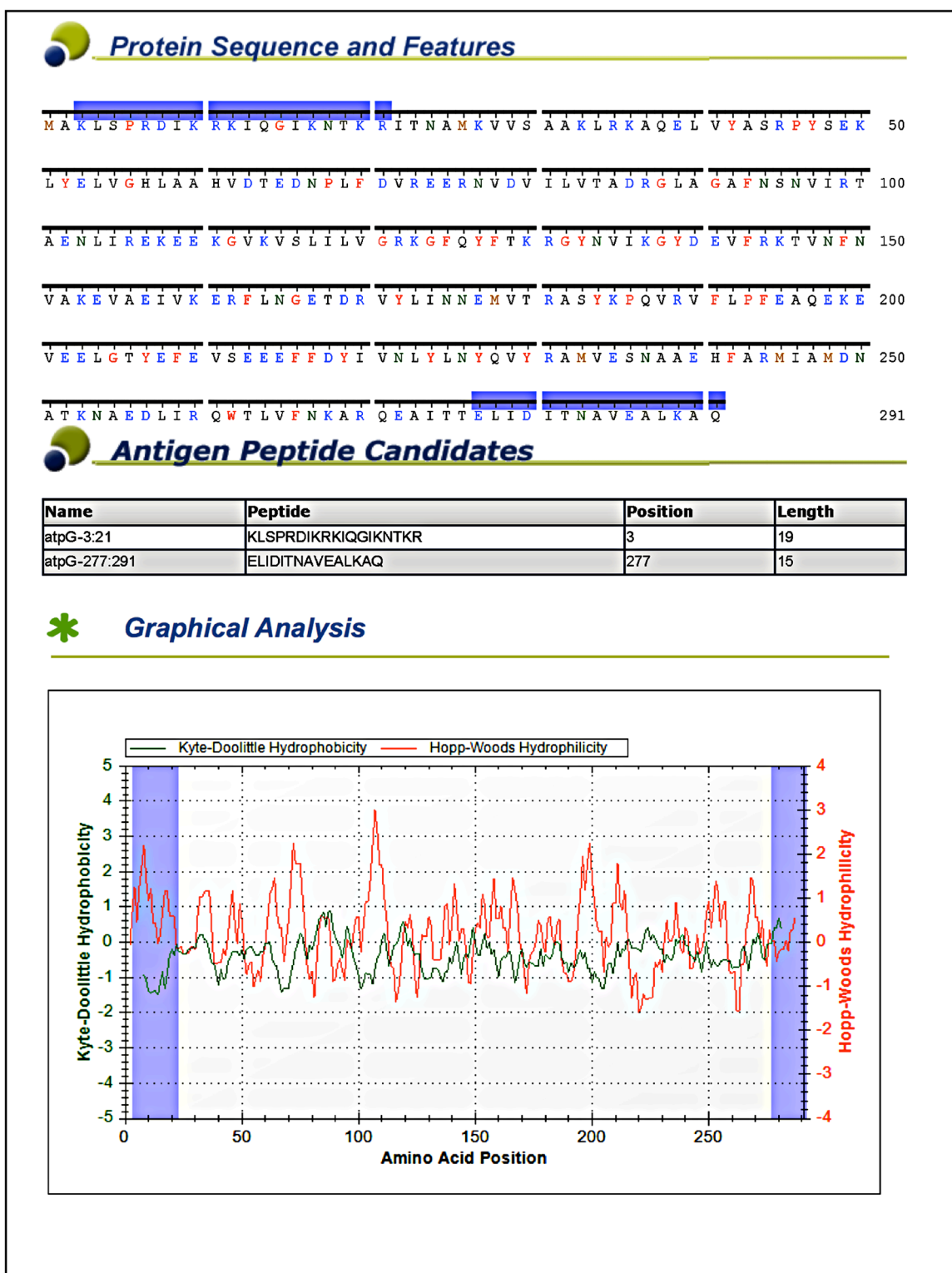
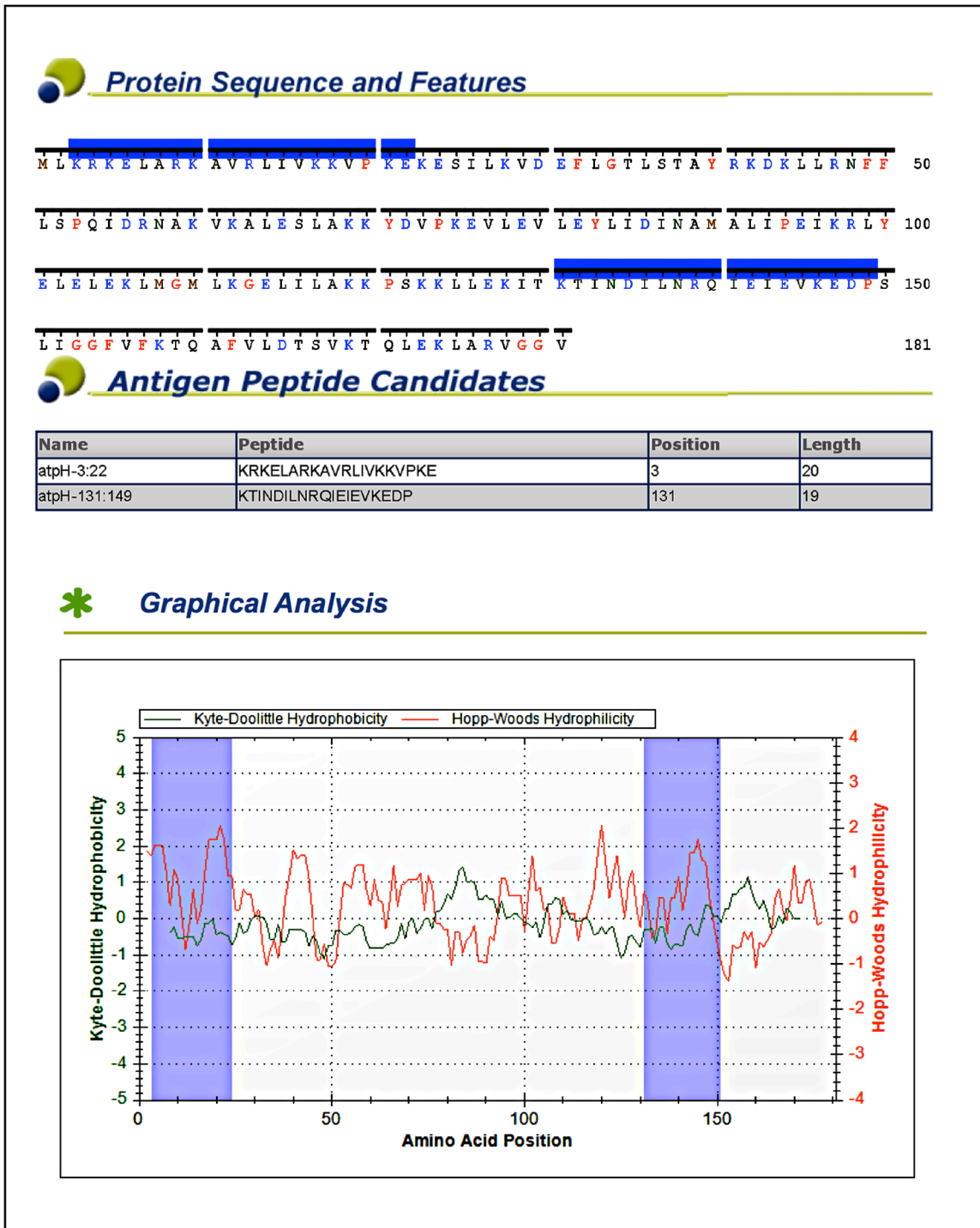
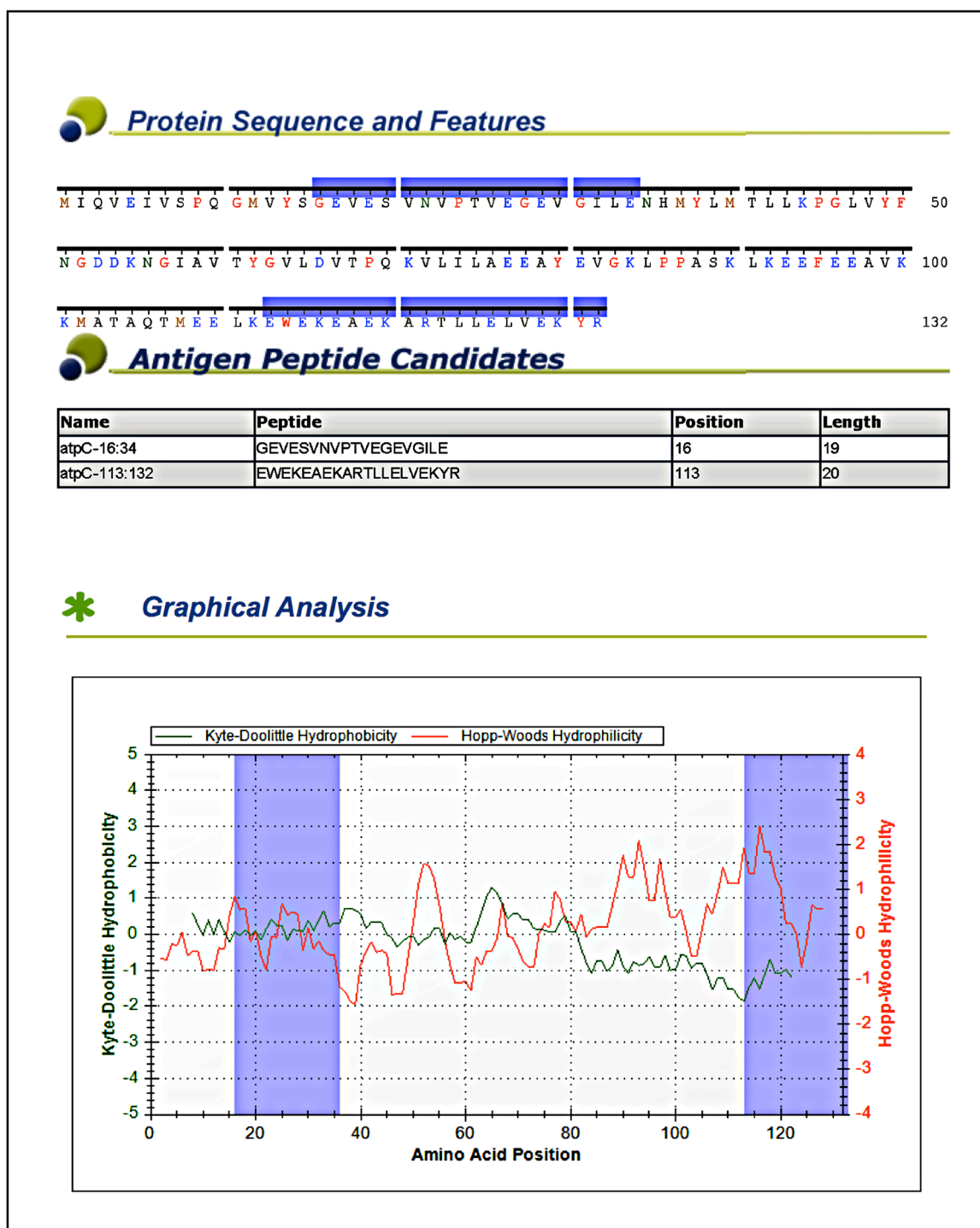


Figure A.12. Antigen profile of subunit α

Figure A.13. Antigen profile of subunit γ



Figure A.15. Antigen profile of subunit ϵ

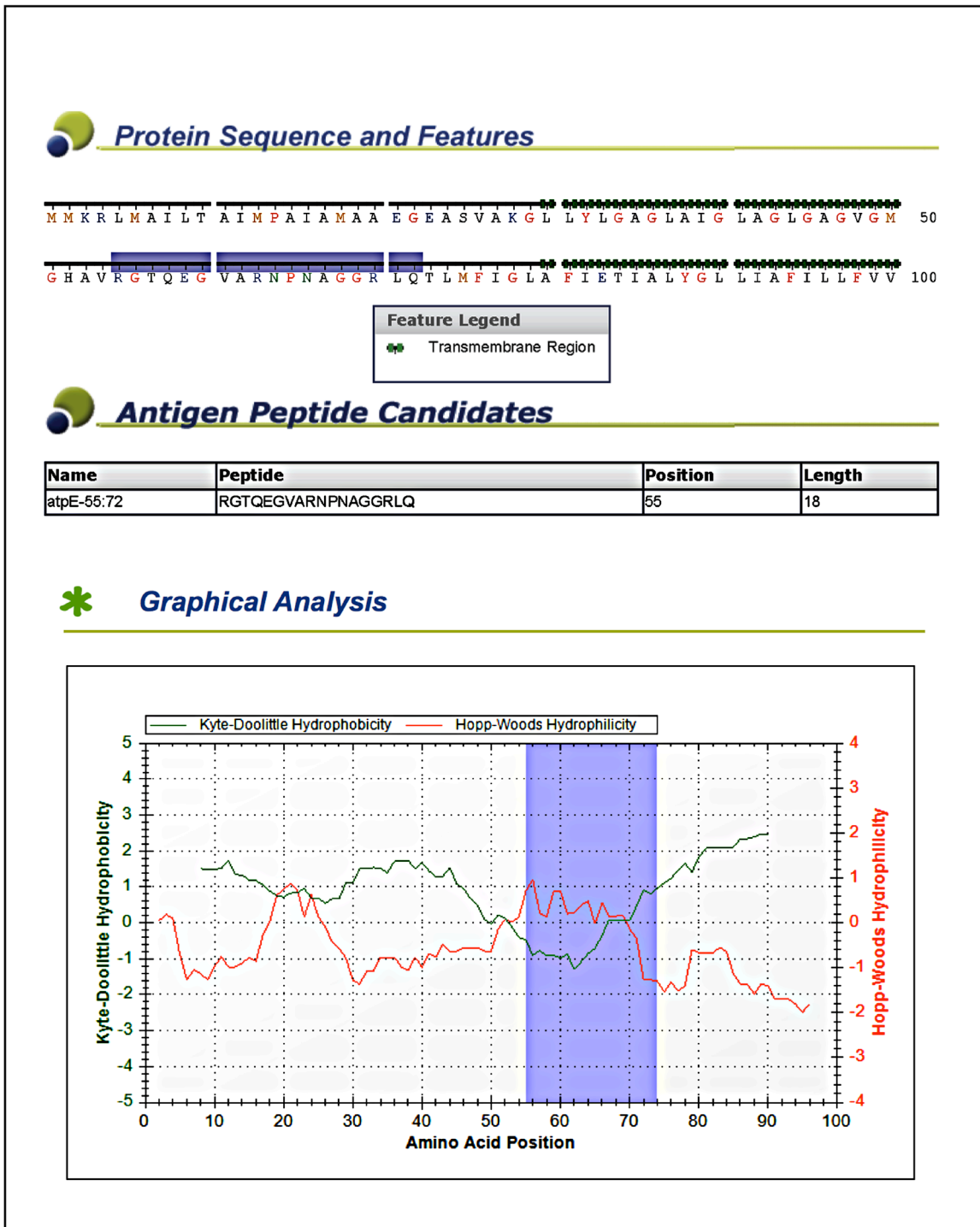


Figure A. 16. Antigen profile of subunit c

Table A3: 70 days rabbit immunization protocol against synthesized peptide (from Thermo Fisher)

Procedure	Protocol Day	Description
Control serum Collection	Day 0	Bleed 5mL per rabbit
Primary injection	Day 1	Primary Immunization with 500ug of antigen in 10 sites, SQ
1 st Booster	Day 14	Boost with 250ug antigen in 4 SQ sites
2 nd Booster	Day 28	Boost with 250ug antigen in 4 SQ sites
Serum Collection	Day 35	~25mL per rabbit
3 rd Booster	Day 42	Boost with 250ug antigen in 4 SQ sites
Serum Collection	Day 56, 58	Two production bleeds (~50mL total per rabbit)
ELISA and Shipping	Day 60	Verify disposition of rabbits, continue or terminate
Instruction Due Date	Day 72	Maintenance charges begin to accrue if instructions not received within time allowed

Acknowledgements

I would like to thank Prof. Hartmut Michel for accepting and supervising me as a graduate student in his laboratories and for providing excellent working conditions, constructive suggestions, critical decisions, and economical as well as scientific support in the project.

I am grateful to Prof. Bernd Ludwig at the Johann Wolfgang Goethe University in Frankfurt for taking the responsibility of internal supervision and for his availability for discussion.

I thank Dr. Guohong Peng for her suggestion to apply for the PhD position in the International Max Planck Research School (IMPreS, Frankfurt) and initial supervision and support in the project.

Many fruitful collaborations were established during this work and led to some of the results presented here.

In the Department of Prof. Hartmut Michel at the Max Planck Institute of Biophysics, Frankfurt, Dr. Julian Langer, Imke Wuellenweber and Falco Peter performed full-length MALDI-TOF and ESI-TOF mass spectrometric experiments. Cornelia Muenke prepared all the modified in-house vectors and established protocols for heterologously producing the membrane proteins.

In the Department of Prof. Werner Kühlbrandt at the Max Planck Institute of Biophysics, Frankfurt, Dr. Janet Vonck and Matteo Allegretti performed single-particle electron microscopy experiments. Dr. Thomas Meier provided me with the *E. coli* DK8 strain and pTrc99A vectors and shared me protocols from his own lab.

Furthermore, I would like to acknowledge Dr. Cristina Fenollar-Ferrer and Dr. Lucy Forrest at Max Planck Institute of Biophysics, Frankfurt for help with the bioinformatic analysis; Paolo Lastrico, Max Planck Institute of Biophysics, Frankfurt for help with the graphics; Dr. Ulrich Ermler, Max Planck Institute of Biophysics, Frankfurt for revising the German summaries of this thesis, as well as discussion and encouragement; Dr. Karen Davies, Max Planck Institute of Biophysics, Frankfurt for revising this thesis.

Sincere thanks to all current and previous members of the Molecular Membrane Biology Department, Max Planck Institute of Biophysics, Frankfurt for the nice working atmosphere, the scientific discussion and the kind technical assistance.

My work has been supported by the Deutsche Forschungsgemeinschaft (SFB 472), the Max Planck Gesellschaft, the Cluster of Excellence “Macromolecular Complexes” (Frankfurt am Main, Germany), and the International Max Planck Research School (IMPreS) on Structure and Function of Biological Membranes (Frankfurt am Main, Germany).

In conclusion, I would like to dedicate some words to people that have been particularly special to me, helped me and supported me for these years in Frankfurt.

Marco Marcia, a half-my-close friend and half-my-personal supervisor. From beginning to end of my PhD, even after he moved to Yale University as a postdoc, he is always being there to support

Acknowledgements

me for both the project and the life in Frankfurt. As a very good friend, he becomes to my daughters' "uncle Marco". Meanwhile, it is from whom that I learned to how to improve my slides and later my writing my papers and thesis. I felt so grateful that I get to know him, get along with him, work with him and especially write together with him.

Tanja Hedderich, another very close irreplaceable friend who is being there always for me. Throughout these years, she has offered me excellent and competent assistance in the lab. More importantly, without her support, some experiments cannot be possible to go as well as it should have. Later even though she was at home with her new born-son Timmy, she still is so supportive for me.

Hao Xie, an important friend and a member of my "personal discussion committee". I thank him for sharing his protocols and experiences on ATP synthase, for his lecture notes on the topics I did not learned before, for many helps with the experiments and criticizing the preparations of Figures. Florian Hilbers, another member of "my discussion committee". Besides all the discussions on the project and troubleshooting for the experiments, I also would like to thank him for listening to my "complains", writing Germany Summary for the thesis, correction on the manuscript.

Laura Preiss, another very important member of "my personal discussion committee". I thank her for all the helps in the experiments, detailed discussions on the project, and comments on my manuscripts.

Sabine Buschmann, my special tea partner. I really enjoy the teatime with her and it becomes part of my life, more recently. I am very impressed by the project and life experiences she shared with me.

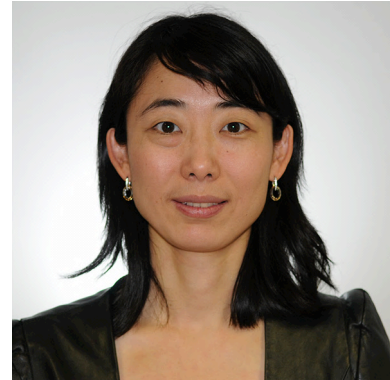
I thank Hanne Mueller for her care about me especially when I was sick and for many gifts she put in front of my doors; Jennifer Witt for her excellent lab support and nice chatting time in the lab; my friends, Yeanxin Sin, Yi-ching Hsueh and Hsiu-chen Lin, for accompanying and spending so many nice weekends and lunch together.

Last but most importantly, I would like to thank my family, my parents, my husband Qi and my daughters Wendi and Yadi. However, my acknowledgement to them is beyond my words. I only can say that I will devote the rest of my life to them, my family.

Curriculum Vitae

Personal information

Name	Chunli ZHANG
Date of birth	14.02.1978
Place of birth	Qingdao, P.R.CHINA



Education

April 2008	PhD position at Molecular Membrane Biology, Max Planck Institute of Biophysics, Frankfurt, Germany Internal supervisor: Prof. Dr. Bernd Ludwig External supervisor: Prof. Dr. Hartmut Michel
July 2002	Master Degree , Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China
July 1999	University degree , China Ocean University, Qingdao, China

Working Experiences

Oct. 2003 – Jun. 2005	Research assistant at Department of Physiology, Faculty of Medicine, National University of Singapore
-----------------------	--

Invited Lectures

May 2009 Joint Workshop on Current Advances of Membrane Protein structure Determinations, Jinan and Qingdao, China

Conferences and courses

Apr. 2013 Conference on Bioenergetics in mitochondria, bacteria and chloroplasts, Marburg

Sep. 2012 EBEC 2012, Freiburg

Sep. 2011 Symposium on Molecular life Science 2011, Frankfurt

Mar. 2011 4th Max Planck Symposium, Life under extreme conditions, Berlin

Jul. 2010 EBEC 2010, Warsaw

Aug. 2008 Bruker Mass Spectrometry training, Bremen
