The A-type ATP synthase subunit K of *Methanopyrus kandleri* is deduced from its sequence to form a monomeric rotor comprising 13 hairpin domains

Juke S. Lolkema[a,*], Egbert J. Boekema[b]

[a] *Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands*

[b] *Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands*

Received 4 March 2003; revised 8 April 2003; accepted 8 April 2003

First published online 28 April 2003

Edited by Stuart Ferguson

Abstract **The *ntpK* gene of the archaeon *Methanopyrus kandleri* encodes the equivalent of the c subunit of ATP synthase. The gene product contains 1021 residues and consists of 13 homologous domains, each one corresponding to a single helical hairpin. The amino acid sequence of the domains is highly conserved, ranging between 50 and 80% sequence identity. Each of the 13 domains contains a conserved Gln and Glu residue in the N- and C-terminal helix, respectively, both of which are believed to be involved in cation binding. The protein is likely to form the monomeric rotor of the ATP synthase that consists of 13 hairpin domains.**

Key words: A-ATP synthase; Rotor; Subunit c; Ring structure; Hyperthermophile; *Methanopyrus kandleri*

1. Introduction

F-type ATP synthases, found in bacteria, mitochondria and thylakoids, and, by analogy, A-type ATPases, found in archaea, and V-type ATPases, found in vacuoles, are membrane-bound molecular motors that interconvert the free energy stored in the ATP molecule and the electrochemical gradient of cations, H^+ or Na^+, across the membrane. The ATP hydrolysis/synthesis and the ion translocation activities reside in different parts of the complex, termed F_1 (V_1, A_1) and F_0 (V_0, A_0), respectively. F_1, the headpiece, is soluble and connected to membrane-bound F_0 by a central stalk and by peripheral stalks, the number of which varies between the different types (e.g. [1–3]). F_1 consists of a hexagonal arrangement of alternating α and β subunits ($\alpha_3\beta_3$; F-type nomenclature). The elongated γ subunit, the main constituent of the central stalk, sticks in the central cavity of the $\alpha_3\beta_3$ complex. ATP synthesis/hydrolysis in F_1 is coupled to physical rotation of the γ subunit. The central stalk is connected to a multimer of subunit c that is part of F_0. The subunit c complex, or the rotor, rotates against the static part of F_0 that is fixed to F_1 via the peripheral stalk(s). Cation translocation takes place at the interface of the rotor and the static part of F_0.

While the structure of F_1 has been resolved to atomic resolution [4], much less is known about the structure of membrane-bound F_0. The rotor part of F_0 is a cylindrically shaped multimeric arrangement of c subunits. Subunit c is a small 8 kDa protein in F-ATPases, which forms a helical hairpin orientated perpendicular to the membrane. The C-terminal transmembrane α-helix contains the cation binding site. The c subunits interact to form a ring-shaped structure in the plane of the membrane. A low resolution crystal structure of the F_1c complex of yeast [5] and atomic force microscopy and cryo-electron microscopy of isolated bacterial [6,7] and chloroplast [8] rotors indicate that the rotor consists of an inner and an outer ring formed by the N-terminal and C-terminal transmembrane α-helices of the c subunits, respectively. This organization accounts for the cation binding site in the C-terminal helix to be in contact with the static part of F_0. The number of helical hairpins that form the rotor determines the ATP/cation stoichiometry during catalysis and was long believed to be a multiple of three, conforming to the rotational symmetry in F_1. Experiments proved otherwise. The rotor part of the F-type enzyme complexes of yeast, the bacteria *Ilyobacter tartaricus* and *Propiogenium modestum*, and spinach chloroplasts contained 10, 11, 11 and 14 c subunits, respectively [5–8]. The stoichiometry of subunit c in V-ATPases and A-ATPases has not been directly determined. Here, we provide strong evidence that the rotor of the A-type complex of the archaeon *Methanopyrus kandleri* consists of 13 helical hairpins.

2. Results and discussion

2.1. *The* ntp *genes on the genome of* M. kandleri

A-type ATPases/synthases consist of nine subunits G, I, K, E, C, F, A, B, D that usually are encoded in a single operon in the indicated order. The genome sequence of the archaeon *M. kandleri* revealed that the genes encoding the subunits of the ATP synthase (the *ntp* genes) were clustered in two separate groups [9]. The genes encoding subunits I, K, E, C, F, and A were clustered together in the indicated order, while the genes encoding subunits B and D were encoded elsewhere on the genome in the order B-D. A homologous gene encoding subunit G could not be detected by a BLAST search [10]. The number of residues of subunits I (656), E (200), C (375), F (113), A (593), and D (233) was very similar as observed for other A-ATPases/synthases. The *ntpB* gene encodes 991 resi-

*Corresponding author. Fax: (31)-50-3632154.
E-mail address: j.s.lolkema@biol.rug.nl (J.S. Lolkema).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00398-3

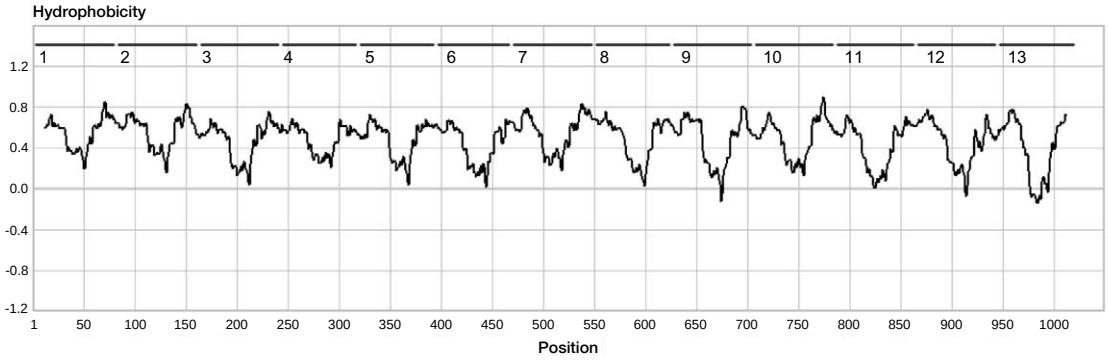

Fig. 1. Hydropathy profile of subunit K of *M. kandleri*. The bars in the upper part indicate the positions of the 13 domains. The window was 20 residues.

dues which is much more than normal, but the gene product is a preprotein that contains an intein [11]. Remarkably, subunit K, corresponding to the F-type c subunit, contains 1021 residues.

2.2. *The K subunit of* M. kandleri

The hydropathy profile of the K subunit of *M. kandleri* revealed a remarkable repetition of hydrophobic regions separated by shorter stretches of hydrophilic residues (Fig. 1). In total, the sequence contained 26 hydrophobic regions long enough to traverse the membrane in α-helical conformation. Analysis of the amino acid sequence revealed that the protein consists of a concatenation of 13 homologous domains, each one representing one helical hairpin. Apparently, the rotor of the ATP synthase of *M. kandleri* consists of a single protein in which 13 helical hairpins form a ring-shaped structure. Multiple sequence alignment [12] of the 13 repeats (Fig. 2) showed that each putative C-terminal helix contains a conserved glutamate residue that is believed to be the cation binding site. The site is characterized by the conserved sequence motif PET. Each putative N-terminal helix contains the conserved motif GQG, in which the glutamine residue is also believed to be part of the cation binding site [13]. The 13 domains are well conserved with pairwise sequence identities between 50 and 80%. The high conservation may reflect strong structural constraints on the domains. Recent studies of the undecamer rotor of *I. tartaricus* and the 14-mer rotor of spinach chloroplasts showed that the curvature of incomplete rings was the same as observed for the complete rotors, strongly suggesting that the diameter of the rotor is an intrinsic property of the structure of the helical hairpin [14]. The curvature of the ring may be very sensitive to mutations in the interface between the domains and, therefore, mutations are unwanted.

2.3. *Comparative analysis of the 13 domains*

The organization of 13 helical hairpins in a single rotor molecule allows for an analysis of the constraints set by the structure and function of the complex on the amino acid sequence of the units. The region around the conserved Gln residue in the N-terminal half of the domains contains an unusually high fraction of amino acid residues with small side chains, i.e. G, A, S. Six glycine residues are conserved in all 13 domains. In contrast, the region around the conserved Glu residue in the C-terminal half of the domains contains mainly residues with bulky, hydrophobic side chains like I, L, V. In a helical wheel presentation the polar Glu is at a face of the helix that otherwise is very hydrophobic. A smaller diameter of the N-terminal helix as compared to the C-terminal helix would be consistent with a compact structure in which the former would form the inner ring and the latter the outer ring of the rotor (see also [15]).

The pairwise sequence identity profile of the alignment clearly shows two well conserved regions centered on the GQG and PET motifs in the N- and C-terminal helices, respectively (Fig. 3). Comparison with the averaged hydropathy profile of the domains shows that the peaks in the identity profiles are shifted towards the center of the domains relative to the peaks in the hydropathy profile. If the latter indicate the position of the transmembrane helices this would indicate not only that the Q and E sites are situated in the cytoplasmic halves of the transmembrane segments, but also that the cytoplasmic halves are particularly well conserved. In such a model, the cytoplasmic half of the hairpin could be involved in the access pathway to the cation binding site from the cytoplasm.

The loop between the two helices is the least conserved region of the helical hairpin (Fig. 3). It contains only one

```
                                ~  *~* *  *  * ~*** ~** *          * ~ ~~ ~~ *** ~***** *~
NTPK1       1 MVSTELTIAAIGAGLAAGVAGVGSGIGQGIAAAAGAGAVAEDEATFGKAIVFSVLPETQAIYGLLTAILIMVGIGLLGAAK-    81
NTPK2      82 AVTVGAALAALGAGLAVGLAGI-SGIGQGIAAASGIGAVLKDEALFGRAIVYAVLPETQAIYGLLVAIIIMVGSGLLGGAGG   162
NTPK3     163 KVSLGAGLAAMGAGLAVGLAGT-SGIGQGIAAASGIHGVLRKEELFGRLIVFSVLPETQAIYGLLTAILIANFVGLLGGPT-   242
NTPK4     243 SVSVGAGLAAMGAGLAVGLAGT-SGIGQGIAAASGIKSLIEEEGVFGRAIVFSVLPETQAIYGLLVAILTLFSLLKP-----   318
NTPK5     319 DLSLAAGLAALGMGLAVGIAGT-SGIGQGIAAASGIAGVLRKEELFGRLIVFSVLPETQAIYGLLTAILAMFFLGAG-----   394
NTPK6     395 KPTLAAGLAAVGAGLAVGFGGT-SGIGQGIAAASGIRAMIERAELFVRGMVLSVLPETRAIYGLLIAILALFMMK-------   468
NTPK7     469 SGSVGAGLALIGAGLAVGLVGV-SGIGQGFTAATGAATLVKNEGFFGRAIIFSVLPETQAIYGLLTAILIMMFAGILGGAGA   549
NTPK8     550 NIGLGAGLAAVGAGLAVGLAGS-SAIGQGIAAAAGVGASAEKEELFGRSVVFSILPETQSIYGLLIGILLAVFAMKA-----   625
NTPK9     626 GSPVGAGLAALGAGLAVGIAGF-SGIGQGIAAAAGIGALKRDPGSFGRSLIFSILPETRSIYGLLVAILVMVGLGLMGGTF-   705
NTPK10    706 SGNEAVGLAALGAGLAIGLAGL-SGVGQGVTAATGISNVVKDPGMFGRSLLFSVFPETQAIYGLLIAILIMMFAGILGGSK-   785
NTPK11    786 SPALGVGLAALGAGIAVGMAGT-SGIGQGISAAAGARATAEDPGNFGRSIVFSILPETQSIYGLLAGILALTPVLTGAGA--   864
NTPK12    865 HLAAAAGLIGIGAGLAVGVAGT-SGIGQGIAAAGGTGALAERTEMFARSLILSILPETRSIYGLLIAILSMSLTGVLGGAG-   944
NTPK13    945 KASLAVGFAAVAAGIAVGFAGL-SGIGQGITAARGSASMVRREQVFGKSLVFSVLPETQAIYGLLTAILIVFAALAAS----- 1021
```

Fig. 2. Multiple sequence alignment of the 13 domains of subunit K of *M. kandleri*. The 13 domains are indicated as NtpK1–13 from the N-terminus to the C-terminus. The sequences were aligned with the ClustalX program using the default settings [12]. The residue numbers of the first and last residue of each domain in the complete protein are indicated on the left and right, respectively. The conserved GQG and PET sequence motifs were printed in bold.

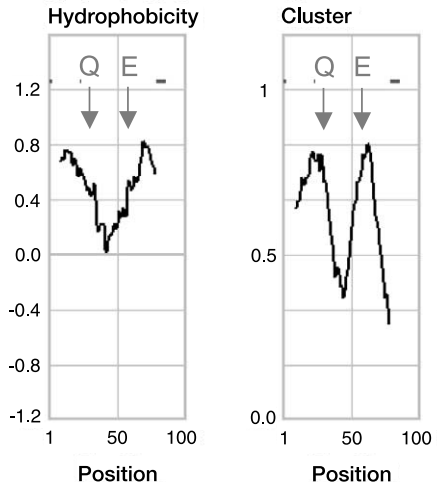

Fig. 3. Averaged hydropathy profile (left) and pairwise sequence identity profile (Cluster; right) of the 13 homologous domains of subunit K of *M. kandleri*. The cluster function measures the fraction of identical pairs at each position averaged over a window of 16 residues. The arrows indicate the positions of the conserved Gln and Glu residues in the first and second transmembrane segments, respectively. Bars in the upper part of the graph indicate positions of gaps in any of the sequences in the alignment.

phenylalanine residue that is conserved in all repeats. This is remarkable since the loop is believed to play an important role in the interaction with the central stalk [15]. Periodicity analysis of the region indicated that even though the sequence is poorly conserved, the loops have a strong and conserved hydrophobic moment when folded in a α-helical structure (Fig. 4). Possibly, the hydrophobic faces of the 13 putative loop helices, containing the conserved phenylalanine residue, form a kind of 'base' by which the rotor part connects to the c subunit that is intermediate between rotor and γ-like subunit in A- and V-type ATPases [16]. The hydrophilic face of the loop helices would contact the water phase.

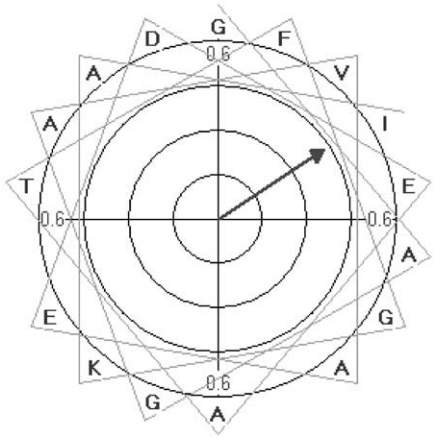

Fig. 4. Helical wheel representation of the loop region between the two transmembrane segments. Printed on the wheel is the sequence of domain NtpK3. The hydrophobic moment indicated by the arrow was calculated using the loop sequences of all 13 domains.

2.4. Comparison to other A- and V-type rotor subunits

Do all the rotors of A-type ATPases/synthases consist of 13 helical hairpins? Highest sequence identity between the domains of the *M. kandleri* subunit K is observed with the corresponding subunits of the related methanogens *Methanothermobacter thermautotrophicus* and *Methanocaldococcus jannaschii*. The c subunits of these organisms do not contain 13 hairpin domains, but two and three, respectively [17]. In fact, subunit K of *M. kandleri* is quite an exception in the archaeal kingdom (see Table 1). It is likely that the number of repeats that make a complete rotor in *M. thermautotrophicus* and *M. jannaschii* consists of a multiple of two and three, respectively, unless an additional gene encoding a single hairpin is encoded on the genome. A precedent for this would be the *atp* operon encoding the F-type ATPase of *Acetobacterium woodii* that contains multiple copies of genes encoding rotor subunits with one and two domains [18]. However, such an additional gene is not present in the operons encoding the

Table 1
Rotor subunits of A-type ATPases/synthases

Archaeon[a]	Gi[b]	Residues	Hairpin domains	Sites/hairpin[c]	Sequence motif
Archaeoglobus fulgidus	11498760	75	1	1/1	PET
Pyrococcus abyssi	14521964	158	2	1/2	PMT/PET
Pyrococcus horikoshii	14591716	162	2	1/2	PMT/PET
Pyrococcus furiosus	18976550	159	2	1/2	PMT/PET
Methanocaldococcus jannaschii	15668394	220	3	2/3	PQT/PET/PET
Methanothermobacter thermautotrophicus	15678977	162	2	1/1	PET/SET
Methanosarcina acetivorans	20092947	82	1	1/1	PET
Methanosarcina mazei	21226886	80	1	1/1	PET
Methanosarcina barkeri	23051714	81	1	1/1	PET
Methanopyrus kandleri	20094449	1021	13	1/1	PET
Halobacterium sp.	15790977	71	1	1/1	PET
Halobacterium salinarum	11362633	89	1	1/1	PET
Thermoplasma volcanium	13540879	75	1	1/1	PET
Thermoplasma acidophilum	16082468	75	1	1/1	PET
Aeropyrum pernix	14601995	131[d]	1	1/1	GEG
Sulfolobus solfataricus	15897488	102[d]	1	1/1	GEG
Sulfolobus tokodaii	15921729	101[d]	1	1/1	GEG
Pyrobaculum aerophilum	18312157	87	1	1/1	AEA

[a]The upper half of the table lists the Euryarchaeota, the lower part the Crenarchaeota.
[b]Gi number of the sequence in the NCBI protein database at http://www.ncbi.nlm.nih.gov/.
[c]A cation binding site is defined as a conserved glutamate residue in the C-terminal helix of the helical hairpin.
[d]Additional putative transmembrane helix at N-terminus.

subunits of the ATPase complex of the two methanogens. Also, BLAST searches [10] did not reveal a second gene encoded elsewhere on the genomes. In conclusion, it is unlikely that the rotor parts of the ATPase complexes of *M. thermautotrophicus* and *M. jannaschii* are built of 13 repeats as well.

Duplication of the c subunit encoding gene is a characteristic of V-type ATPases of eukaryotic origin. The c subunits of V-ATPases have double the mass of those of F-ATPases. Importantly, only the second domain in the V-type c subunits contains the cation binding site, and, consequently, the cation/ATP stoichiometry is half that of an F-type enzyme with the same number of domains in the rotor [19]. The difference provides a mechanistic explanation for the different physiological functions of V- and F-types, ATP hydrolysis and ATP synthesis, respectively [18]. The archaeal complexes are evolutionarily closely related to the V-type ATPases, but their rotor subunits come in a great variety (Table 1). They differ both in the number of hairpin domains per subunit, the number of cation binding sites per hairpin, and, most likely, in the number of hairpins per rotor structure. The c subunit of all of the Crenarchaeota and many of the Euryarchaeota consists of a single hairpin domain. The *Pyrococcus* species and *M. thermautotrophicus* have c subunits with two hairpin domains, but the first hairpin of the former does not contain a cation binding site. The *M. jannaschii* subunit is the only one with three domains, two of which contain a binding site [17]. Finally, the protein of *M. kandleri* contains 13 domains and 13 binding sites. In all cases, the binding site residues come as part of a conserved triad, PET in the Euryarchaeota and GEG in the Crenarchaeota. The neighboring residues are even conserved when the binding site Glu itself is not present. It should be noted that in a helical conformation, the neighboring residues are not likely to be part of the binding pocket. The variability in the archaeal rotors may reflect different physiological functions of the A-ATPases/synthases.

2.5. Conclusions

Analysis of the *ntpK* gene of *M. kandleri* results in two main conclusions. First, the rotor of the ATP synthase consists of a single protein, and second, the ring structure of the rotor is built of 13 homologous domains. *M. kandleri* is a strictly anaerobic methanogen isolated from a 'black smoker' at 2 km below sea level [20]. The hyperthermophile grows optimally at temperatures between 80 and 110°C. These harsh conditions will put high demands on the stability of the cellular proteins which may explain why the rotor consists of a single protein. An additional advantage would be that no regulatory mechanisms during biosynthesis of the complex are necessary to account for a higher stoichiometry of the rotor subunit in the complex. Recent structural data suggested that the stoichiometry of the rotor ring depends on the biological species from which the ATP synthase originates. In the yeast *Saccharomyces cerevisiae* the rotor is a decamer, in the bacteria *I. tartaricus* and *P. modestum* an undecamer, and in chloroplasts a 14-mer. The 13-mer described here in *M. kandleri* appears to further add to the heterogeneity of the rotors. It is to be expected that this stoichiometry is not general in the archaeal kingdom. In general, rotor subunit stoichiometry may be determined conditionally, rather than evolutionally.

References

[1] Wilkens, S. and Capaldi, R.A. (1998) Nature 393, 29.
[2] Boekema, E.J., van Breemen, J.F.L., Brisson, A., Ubbink-Kok, T., Konings, W.N. and Lolkema, J.S. (1999) Nature 401, 37–38.
[3] Domgall, I., Venzke, D., Luttge, U., Ratajczak, R. and Böttcher, B. (2002) J. Biol. Chem. 277, 13115–13121.
[4] Abrahams, J.P., Leslie, A.G., Lutter, R. and Walker, J.E. (1994) Nature 370, 621–628.
[5] Stock, D., Leslie, A.G. and Walker, J.E. (1999) Science 286, 1700–1705.
[6] Stahlberg, H., Müller, D.J., Suda, K., Fotiadis, D., Engel, A., Meier, T., Matthey, U. and Dimroth, P. (2001) EMBO Rep. 2, 229–233.
[7] Meier, T., Matthey, U., von Ballmoos, C., Vonck, J., Krug von Nidda, T., Kühlbrandt, W. and Dimroth, P. (2003) J. Mol. Biol. 325, 389–397.
[8] Seelert, H., Poetsch, A., Dencher, N.A., Engel, A., Stahlberg, H. and Müller, D.L. (2000) Nature 405, 418–419.
[9] Slesarev, A.I., Mezhevaya, K.V., Makarova, K.S., Polushin, N.N., Shcherbinina, O.V., Shakhova, V.V., Belova, G.I., Aravind, L., Natale, D.A., Rogozin, I.B., Tatusov, R.L., Wolf, Y.I., Stetter, K.O., Malykh, A.G., Koonin, E.V. and Kozyavkin, S.A. (2002) Proc. Natl. Acad. Sci. USA 99, 4644–4649.
[10] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389–3402.
[11] Perler, F.B., Davis, E.O., Dean, G.E., Gimble, F.S., Jack, W.E., Neff, N., Noren, C.J., Thorner, J. and Belfort, M. (1994) Nucleic Acids Res. 22, 1125–1127.
[12] Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. (1998) Trends Biochem. Sci. 23, 403–405.
[13] Kaim, G., Wehrle, F., Gerike, U. and Dimroth, P. (1997) Biochemistry 36, 9185–9194.
[14] Müller, D.J., Dencher, N.A., Meier, T., Dimroth, P., Suda, K., Stahlberg, H., Engel, A., Seelert, H. and Matthey, U. (2001) FEBS Lett. 504, 219–222.
[15] Vonck, J., Krug von Nidda, T., Meier, T., Matthey, U., Mills, D.J., Kühlbrandt, W. and Dimroth, P. (2002) J. Mol. Biol. 321, 307–316.
[16] Chaban, Y., Ubbink-Kok, T., Keegstra, W., Lolkema, J.S. and Boekema, E.J. (2003) EMBO Rep. 3, 982–987.
[17] Ruppert, C., Kavermann, H., Wimmers, S, Schmid, R., Kellermann, J., Lottspeich, F., Huber, H., Stetter, K.O. and Muller, V. (1999) J. Biol. Chem. 274, 2581–2584.
[18] Muller, V., Aufurth, S. and Rahlfs, S. (2001) Biochim. Biophys. Acta 1505, 108–120.
[19] Forgac, M. (2000) J. Exp. Biol. 203, 71–80.
[20] Huber, R., Kur, M., Jannasch, H.W. and Stetter, K.O. (1989) Nature 342, 833–836.