Primary structure of the β -subunit of *Torpedo californica* (Na⁺ + K⁺)-ATPase deduced from the cDNA sequence

Shunsuke Noguchi*†, Masaharu Noda†, Hideo Takahashi†, Kiyoshi Kawakami†, Toshiko Ohta†, Kei Nagano†, Tadaaki Hirose, Seiichi Inayama, Masaru Kawamura*§ and Shosaku Numa†°

[†]Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Kyoto 606, *Department of Biology, Chiba University Faculty of Science, Chiba 260, *Department of Biology, Jichi Medical School, Yakushiji, Tochigi-ken 329-04 and Pharmaceutical Institute, Keio University School of Medicine, Tokyo 160, Japan

Received 9 December 1985

DNA complementary to the *Torpedo californica* electroplax mRNA coding for the β -subunit of (Na⁺ + K⁺)-ATPase has been cloned by screening a cDNA library with an oligodeoxyribonucleotide probe. Nucleotide sequence analysis of the cloned cDNA has revealed that this polypeptide consists of 305 amino acid residues (including the initiating methionine). The transmembrane topology and the potential *N*-glycosylation sites of this polypeptide are discussed.

(Na⁺+K⁺)-ATPase cDNA cloning Transmembrane topology Nucleotide sequence N-Glycosylation site RNA blot hybridization analysis Primary structure

1. INTRODUCTION

Sodium- and potassium-dependent adenosine triphosphatase [(Na⁺ + K⁺)-ATPase] mediates the active transport of Na⁺ and K⁺ across the plasma membrane and is known to consist of 2 kinds of subunits, α and β [1-4]. The primary structure of the larger α -subunit, which plays the principal catalytic role, has recently been elucidated by cloning and sequencing DNAs complementary to the *Torpedo* electroplax and sheep kidney mRNAs coding for the polypeptide [5,6]. The β -subunit is a glycoprotein with an M_r of 40000-60000. In this study, the complete amino acid sequence of the

Abbreviations: (Na⁺ + K⁺)-ATPase, sodium- and potassium-dependent adenosine triphosphatase; PTH-amino acid, phenylthiohydantoin-amino acid

 β -subunit of *Torpedo californica* (Na⁺ + K⁺)-ATPase has been deduced by cloning and sequencing the cDNA.

2. MATERIALS AND METHODS

 $(Na^+ + K^+)$ -ATPase from the electric organ of T. californica was purified as in [5]. The purified enzyme was treated with 1% SDS and fractionated by gel filtration on a Biogel A-1.5 m column $(2.5 \times 100 \text{ cm})$ [7]. The fraction consisting mainly of the β -subunit was further purified by electrophoresis on 0.1% SDS/8% polyacrylamide gels according to [8]; approx. 0.4 mg of the protein was loaded per slab gel $(0.2 \times 14 \times 14 \text{ cm})$. The protein band corresponding to the β -subunit, visualized with 0.25 M KCl and 1 mM dithiothreitol [9], was excised and the gel piece was crushed and subjected to electroelution [10] with an ISCO model 1750 sample concentrator in 40 mM (outer chamber) and 4 mM (inner chamber) Tris-acetate buffer, pH 8.6, containing 0.1% SDS.

[°] To whom correspondence should be addressed

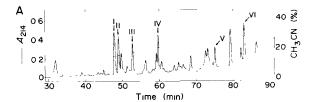
[§] Present address: Department of Biology, University of Occupational and Environmental Health, Kitakyushu 807, Japan

The Okayama-Berg cDNA library [11] screened for $(Na^+ + K^+)$ -ATPase β -subunit cDNA was the same as that used for cloning cDNAs encoding the acetylcholine receptor subunits [12]. The transformation and screening procedures were as in [13]. The oligodeoxyribonucleotide probe (see below), synthesized by the triester method [14], was labelled with 32P at the 5'-end and used for hybridization at 37°C. The 118 bp HaeIII(-35)-Sau96I(84) fragment excised from clone pNKA\$1 was labelled with ³²P at the 5'-end and used for hybridization at 60°C; the restriction sites are identified by numbers indicating the 5'-terminal nucleotide generated by cleavage. DNA sequencing was carried out according to [15]. Reagents were obtained as in [5].

3. RESULTS AND DISCUSSION

Tryptic peptides from the purified β -subunit of T. californica $(Na^+ + K^+)$ -ATPase were isolated by reverse-phase HPLC and analysed for amino acid sequence with a gas-phase sequencer (fig.1). Six peptide sequences were thus determined. The synthetic oligodeoxyribonucleotide 5'-ACATGAT- $T_C^T T C_G^A T A - 3'$, corresponding to the pentapeptide sequence Tyr-Glu-Asn-His-Val contained in fraction IVa (fig.1), was used as hybridization probe to screen a cDNA library derived from poly(A)+ RNA from an electric organ of T. californica. Three hybridization-positive clones were thus isolated from about 1×10^4 transformants. One of them (pNKA\beta1), which apparently carried the largest cDNA insert, was subjected to nucleotide sequence analysis. Further screening of about 8 × 10⁴ transformants from the same cDNA library with a hybridization probe derived from the 5'-terminal region of clone pNKAβ1 yielded 24 positive clones. One of them (pNKA\beta144), which apparently harboured the largest cDNA insert, was analysed for nucleotide sequence.

Fig.2 shows the 2105 nucleotide sequence (excluding the poly(dA) tract) of the cDNA encoding the β -subunit of T. californica (Na⁺ + K⁺)-ATPase, determined with clones pNKA β 1 and pNKA β 144. All the 6 partial amino acid sequences determined by peptide analysis were found to be encoded by the cDNA sequence in the same reading frame (amino acid residues 16–22, 73–82, 87–106, 152–158, 226–245 and 280–290). This



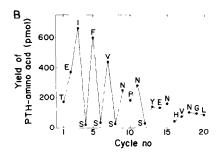


Fig.1. Partial amino acid sequence analysis of the β subunit of T. californica $(Na^+ + K^+)$ -ATPase. (A) Reverse-phase HPLC of tryptic peptides. After dialysis to remove SDS, approx. 1 mg of the purified β -subunit was digested by trypsin and then subjected to reverse-phase HPLC. Six fractions (I-VI) corresponding to absorbance peaks were collected and rechromatographed on the same column to yield fractions Ia-VIa, respectively. (B) Sequence analysis of peptides. The yield of PTH-amino acids at each cycle of Edman degradation is shown for fraction IVa; the oneletter amino acid notation is used. The sequences determined for the other fractions (not shown) were as follows: Ia, VFGDNIAYSEK; IIa, FLLQWLK; IIIa, FLWDSEK: Va. VAPPGLSHSP; VIa, IGTIEYFG-MGGVGGFPLQYY. For experimental details, see [5].

reading frame was used to deduce the primary structure of the $(Na^+ + K^+)$ -ATPase β -subunit (fig.2). The translational initiation site was assigned to the methionine codon composed of nucleotide residues 1-3 because this is the first ATG triplet that appears downstream of a nonsense codon (TAA at positions -51 to -49) found in frame and because the second ATG triplet encodes amino acid residue 58, which is preceded by one of the peptide sequences determined. The 305th codon specifying serine is followed by 2 successive translational termination codons (TGA and TAA). Thus, it is concluded that the β -subunit of T. californica $(Na^+ + K^+)$ -ATPase consists of 305 amino acid residues (including the

5'ACCTTCCTCGAAACTCCTCACACCTGGAGCTTCTACTCAGTCCTCCCTTTCACCCCTAGCCCCCCCC	-241
	-121
ACGCGGGCGAGGGCGCGGCCGATCGGCTCCGCGAGCAAGGCTCGAAGGGGGCACCTCGCCTCCAGCGTAACAGGGGCGCCTTGGCCTCTCCGCCAACTGCCACCGCAAAGTCAGCGAAG	-1
1 10 20 30 Met Ala Arg Glu Lys Ser Thr Asp Asp Gly Gly Gly Trp Lys Lys Phe Leu Trp Asp Ser Glu Lys Lys Gln Val Leu Gly Arg Thr Gly ATG GCA CGA GAA AAG TCC ACG GAT GAT GGT GGA GGA TGG AAG AAA TTC CTC TGG GAT TCC GAG AAG AAA CAG GTT TTG GGC AGG ACC GGC	90
40 50 Thr Ser Trp Phe Lys Ile Phe Val Phe Tyr Leu Ile Phe Tyr Gly Cys Leu Ala Gly Ile Phe Ile Gly Thr Ile Gln Val Met Leu Leu ACG AGC TGG TTC AAG ATA TTT GTA TTC TAC TTG ATC TTT TAC GGT TGC CTC GCT GGA ATC TTC ATC GGT ACC ATC CAA GTA ATG TTA CTG	180
70 80 90 Thr Ile Ser Asp Phe Glu Pro Lys Tyr Gln Asp Arg Val Ala Pro Pro Gly Leu Ser His Ser Pro Tyr Ala Val Lys Thr Glu Ile Ser ACC ATA AGT GAT TTC GAA CCA AAA TAC CAG ACA AGA GTT GCA CCT CCA GGT CTA TCG CAT AGT CCA TAT GCT GTA AAA ACT GAG ATA AGC	270
100 120 Phe Ser Val Ser Asn Pro Asn Ser Tyr Glu Asn His Val Asn Gly Leu Lys Glu Leu Leu Lys Asn Tyr Asn Glu Ser Lys Gln Asp Gly TTT AGT GTT TCA AAC CCA AAC TCC TAT GAA AAT CAT GTG AAT GGT CTG AAG GAA CTG TTG AAG AAT TAC AAT GAA TCA AAG CAA GAT GGC	360
130 140 150 Asn Thr Pro Phe Glu Asp Cys Gly Val Ile Pro Ala Asp Tyr Ile Thr Arg Gly Pro Ile Glu Glu Ser Gln Gly Gln Lys Arg Val Cys AAC ACT CCT ITT GAG GAT TGT GGC GTA ATT CCT GCA GAC TAC ATA ACT AGA GGT CCA ATA GAG GAA TCG CAG GGA CAG AAG AGA GTA TGC	450
160 170 180 Arg Phe Leu Leu Gln Trp Leu Lys Asn Cys Ser Gly Ile Asp Asp Pro Ser Tyr Gly Tyr Ser Glu Gly Lys Pro Cys Ile Ile Ala Lys AGA TTT TTG CTT CAG TGG CTT AAA AAT TGC TCT GGA ATT GAT GAT CCT AGT TAT GGC TAT TCT GAA GGA AAA CCC TGC ATC ATT GCC AAG	540
190 200 210 Leu Asn Arg Ile Leu Gly Phe Tyr Pro Lys Pro Pro Lys Asn Gly Thr Asp Leu Pro Glu Ala Leu Gln Ala Asn Tyr Asn Gln Tyr Val CTC AAC AGG ATC CTT GGT TTC TAT CCT AAG CCT CCG AAA AAT GGC ACT GAT CTT CCT GAA GCA TTG CAA GCA AAT TAT AAT CAG TAT GTC	630
220 230 240 Leu Pro Ile His Cys Gln Ala Lys Lys Glu Glu Asp Lys Val Arg Ile Gly Thr Ile Glu Tyr Phe Gly Met Gly Gly Val Gly Gly Phe CTT CCT ATT CAT TGT CAA GCA AAG AAA GAA GAA GAC AAA GTC AGA ATT GGA ACC ATC GAA TAT TTT GGT ATG GGT GGA GTT GGT GGC TTT	720
250 270 Pro Leu Gln Tyr Tyr Pro Tyr Tyr Gly Lys Arg Leu Gln Lys Asn Tyr Leu Gln Pro Leu Val Gly Ile Gln Phe Thr Asn Leu Thr His CCC TTG CAG TAC TAT CCA TAC TAT GGA AAG CGC CTG CAG AAA AAT TAC CTT CAA CCC TTG GGT GGT ATT CAG TTC ACC AAC CTC ACA CAC	810
280 290 300 Asn Val Glu Leu Arg Val Glu Cys Lys Val Phe Gly Asp Asn Ile Ala Tyr Ser Glu Lys Asp Arg Ser Leu Gly Arg Phe Glu Val Lys AAT GTG GAG CTG CGT GTT GAG TGT AAA GTG TTT GGT GAC AAT ATT GCA TAT AGT GAA AAA GAT CGC TCA CTG GGG CGC TTT GAA GTA AAA	900
Ile Glu Val Lys Ser ATT GAA GTT AAA AGC TGA TAACTGTAATAGAATTTACCCCACATTACAAATTAGTCTTGAACAAACTGTCATACATA	1014
	1134
	1254
	1374
	1494 1614
	1734

Fig. 2. Nucleotide sequence of the cDNA encoding the β -subunit of T. californica (Na⁺ + K⁺)-ATPase. Nucleotide residues are numbered in the 5' to 3' direction, beginning with the first residue of the ATG triplet encoding the initiating methionine, and the nucleotides on the 5'-side of residue 1 are indicated by negative numbers; the number of the nucleotide residue at the right-hand end of each line is given. The deduced amino acid sequence of the (Na⁺ + K⁺)-ATPase β -subunit is shown above the nucleotide sequence, and amino acid residues are numbered beginning with the initiating methionine. The 5'-terminal sequence presented does not extend to the 5'-end of the mRNA. The 3'-terminal sequence of the cDNA insert of clone pNKA β 1 (carrying nucleotides -67 to 1777) as well as clone pNKA β 144 (carrying nucleotides -328 to 1751) is followed by a poly(dA) tract connected with the vector DNA sequence [11]. The 3'-noncoding region contains 2 copies of the polyadenylation signal AATAAA [16] (nucleotides 1035–1040 and 1734–1739). No nucleotide difference was found between the 2 clones.

initiating methionine) and has a calculated M_r of 34671, which agrees with the reported M_r of the protein moiety of this subunit (32000–38000) [4,17–20]. The sequence of amino acid residues 2–10 or 2–35 is homologous with the reported

ATAAAACTTAAGAACCCATTCTTGTGGGACTTCTCTGGAAATT----3'

amino-terminal sequence of the lamb [21] or dog kidney (Na⁺ + K⁺)-ATPase β -subunit [22]; in the latter case, residue 12 or 13 in the *Torpedo* sequence is assumed to be an insertion. Furthermore, the amino-terminal residue has been shown to be

alanine for all the mammalian, avian and fish β subunits hitherto examined [23–25]. These findings suggest that the alanine residue following the
initiating methionine represents the amino terminus of the β -subunit.

The blot hybridization analysis of T. californica electroplax RNA with $(Na^+ + K^+)$ -ATPase β -subunit cDNA as probe exhibited a single hybridizable RNA species with an estimated size of approx. 2600 nucleotides (fig.3). This implies that a 5'-terminal sequence of about 400 nucleotide is missing from the cDNA sequence presented in fig.2.

Fig.4 shows the hydropathy profile [29] and predicted secondary structures [30] of the

Origin—

28 S-

23 S-

18 S-

16 S-

Fig. 3. Autoradiogram of the blot hybridization analysis of RNA from *T. californica* electric organ. Total RNA was extracted as in [26], and was analysed by the procedure in [27], using Biodyne membrane (Pall, East Hills, NY); the amount used was 5 μg. The hybridization probe was the *Ban*I(158)-*Fnu*4HI(821) fragment excised from clone pNKAβ1 and labelled by nick-translation [28] with [α-32P]dCTP. The size markers were bovine and *Escherichia coli* rRNA.

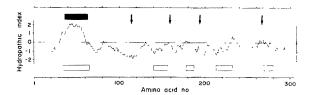


Fig.4. Hydropathy profile and predicted secondary structures of the $(Na^+ + K^+)$ -ATPase β -subunit. The averaged hydropathic index of a nonadecapeptide composed of amino acid residues i-9 to i+9 is plotted against i, where i represents amino acid number; the hydropathy indices of individual amino acids have been taken from [29]. The positions of the predicted structures of α -helix and/or β -sheet [30] that have a length of 10 or more residues are shown by open boxes. The location of the putative transmembrane segment is indicated by a closed box and the potential N-glycosylation sites by arrows.

 $(Na^+ + K^+)$ -ATPase β -subunit. A highly hydrophobic region with predicted secondary structure is observed near the amino terminus. It corresponds to a continuous stretch of 28 uncharged amino acids (residues 36-63) including many nonpolar residues and may represent a transmembrane segment. The $(Na^+ + K^+)$ -ATPase β -subunit does not possess a hydrophobic amino-terminal sequence characteristic of the signal peptide [31]. Furthermore, all 4 potential N-glycosylation sites [32] present in the β -subunit (asparagine residues 114, 159, 194 and 267) are located on the carboxy-terminal side of the putative transmembrane segment. On the basis of these observations, it is assumed that the large portion of the β -subunit molecule following the putative transmembrane segment is located on the extracellular side of the membrane, whereas the small amino-terminal portion resides on the cytoplasmic side. This transmembrane topology of the β -subunit is consistent with previous findings as follows: (i) immunochemical [33] and chemical labelling studies [34] indicate that the β -subunit spans the membrane; (ii) chemical labelling [35,36] and cell-free translation studies [19] indicate that virtually all of the hydrophilic part of the β subunit is located on the extracellular side of the membrane: (iii) studies by pulse-chase labelling and endoglycosidase digestion suggest that the β subunit has 2 or more N-linked oligosaccharide chains [20,37]; (iv) the results of papain fragmentation suggest that the carboxy-terminal fragment of M_r 16000 as well as the preceding fragment of M_r 31000 contains covalently bound carbohydrate, whereas the amino-terminal fragment of M_r 9000 apparently has no carbohydrate [22].

Three additional segments comprising almost exclusively uncharged residues (residues 201-217, 231-249 and 255-272) are present in the carboxyterminal region of the β -subunit. However, they seem not to be sufficiently hydrophobic to form transmembrane segments (fig.4). It is also to be noted that the sequence Gly-X-Gly-X-X-Gly (X can be any amino acid), which is characteristic of the nucleotide binding region of nucleotide binding proteins [38], is found in the β -subunit (residues 233–238). The functional implication of this sequence is not known because no information is available as to the role of the β -subunit in nucleotide binding. In any event, a second transmembrane segment would be required to locate this sequence on the cytoplasmic side of the membrane.

ACKNOWLEDGEMENTS

We thank Dr Paul Berg and Dr Hiroto Okayama for providing their high-efficiency cloning system and Dr Takashi Miyata and Mr Hidenori Hayashida for computer analysis. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

REFERENCES

- [1] Kyte, J. (1981) Nature 292, 201-204.
- [2] Schuurmans-Stekhoven, F. and Bonting, S.L. (1981) Physiol. Rev. 61, 1-76.
- [3] Jørgensen, P.L. (1982) Biochim. Biophys. Acta 694, 27-68.
- [4] Post, R.L. (1983) Curr. Top. Membrane Transp. 19, 53-65.
- [5] Kawakami, K., Noguchi, S., Noda, M., Takahashi, H., Ohta, T., Kawamura, M., Nojima, H., Nagano, K., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. and Numa, S. (1985) Nature 316, 733-736.
- [6] Shull, G.E., Schwartz, A. and Lingrel, J.B. (1985) Nature 316, 691-695.

- [7] Marshall, P.J. and Hokin, L.E. (1979) Biochem. Biophys. Res. Commun. 87, 476-482.
- [8] Laemmli, U.K. (1970) Nature 227, 680-685.
- [9] Hager, D.A. and Burgess, R.R. (1980) Anal. Biochem. 109, 76-86.
- [10] Bhown, A.S. and Bennett, J.C. (1983) Methods Enzymol. 91, 450-455.
- [11] Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161-170.
- [12] Noda, M., Takahashi, H., Tanabe, T., Toyosato, M., Furutani, Y., Hirose, T., Asai, M., Inayama, S., Miyata, T. and Numa, S. (1982) Nature 299, 793-797.
- [13] Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Hirose, T., Inayama, S., Nakanishi, S. and Numa, S. (1982) Nature 295, 202-206.
- [14] Ito, H., Ike, Y., Ikuta, S. and Itakura, K. (1982) Nucleic Acids Res. 10, 1755-1769.
- [15] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
- [16] Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211-214.
- [17] Hastings, D.F. and Reynolds, J.A. (1979) Biochemistry 18, 817-821.
- [18] Sabatini, D., Colman, D., Sabban, E., Sherman, J., Morimoto, T., Kreibich, G. and Adesnik, M. (1981) Cold Spring Harbor Symp. Quant. Biol. 46, 807-818.
- [19] Geering, K., Meyer, D.I., Paccolat, M.-P., Kraehenbühl, J.-P. and Rossier, B.C. (1985) J. Biol. Chem. 260, 5154-5160.
- [20] Sweadner, K.J. and Gilkeson, R.C. (1985) J. Biol. Chem. 260, 9016-9022.
- [21] Collins, J.H. and Zot, A.S. (1983) IRCS Med. Sci. Libr. Compend. 11, 799.
- [22] Chin, G.J. (1985) Biochemistry 24, 5943-5947.
- [23] Perrone, J.R., Hackney, J.F., Dixon, J.F. and Hokin, L.E. (1975) J. Biol. Chem. 250, 4178-4184.
- [24] Hopkins, B.E., Wagner, H. jr and Smith, T.W. (1976) J. Biol. Chem. 251, 4365-4371.
- [25] Dzhandzhugazyan, K.N., Modyanov, N.N. and Ovchinnikov, Yu.A. (1981) Bioorg. Khim. 7, 847-857.
- [26] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.
- [27] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
- [28] Weinstock, R., Sweet, R., Weiss, M., Ceder, H. and Axel, R. (1978) Proc. Natl. Acad. Sci. USA 75, 1299-1303.
- [29] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.
- [30] Chou, P.Y. and Fasman, G.D. (1978) Annu. Rev. Biochem. 47, 251-276.

- [31] Blobel, G. and Dobberstein, B. (1975) J. Cell. Biol. 67, 852-862.
- [32] Marshall, R.D. (1974) Biochem. Soc. Symp. 40, 17-26.
- [33] Girardet, M., Geering, K., Frantes, J.M., Geser, D., Rossier, B.C., Kraehenbuhl, J.-P. and Bron, C. (1981) Biochemistry 20, 6684-6691.
- [34] Jørgensen, P.L. and Brunner, J. (1983) Biochim. Biophys. Acta 735, 291-296.
- [35] Dzhandzhugazyan, K.N. and Jørgensen, P.L. (1985) Biochim. Biophys. Acta 817, 165-173.
- [36] Ovchinnikov, Yu.A., Demin, V.V., Barnakov, A.N., Kuzin, A.P., Lunev, A.V., Modyanov, N.N. and Dzhandzhugazyan, K.N. (1985) FEBS Lett. 190, 73-76.
- [37] Fambrough, D.M. (1983) Cold Spring Harbor Symp. Quant. Biol. 48, 297-304.
- [38] Möller, W. and Amons, R. (1985) FEBS Lett. 186, 1-7.