

Molecular cloning of the Na,K-ATPase α -subunit in developing brine shrimp and sequence comparison with higher organisms

Lee Ann Baxter-Lowe*, Jian Zhong Guo^o, Ellen E. Bergstrom and Lowell E. Hokin

Department of Pharmacology, University of Wisconsin Medical School, 1300 University Avenue, Madison, WI 53706, USA

Received 3 August 1989; revised version received 9 September 1989

We report here the molecular cloning, nucleotide sequence, and predicted amino acid sequence of an α -subunit of the developmentally useful model, *Artemia*. The amino acid sequence shows divergence from that of mammals, birds, *Torpedo*, and *Drosophila*. However, regions in the putative ATP binding and transmembrane domains show absolute or high levels of conservation. Major differences occur in the amino-terminal domain and several other hypervariable regions. These differences are consistent with the suggestion that the brine shrimp is a 'fast clock' organism which diverged from the precursors of vertebrates 0.5–1 billion years ago.

Na,K-ATPase; Molecular cloning; DNA, complementary; Amino acid sequence; (*Artemia*)

1. INTRODUCTION

The sodium- and potassium-activated adenosine triphosphatase (Na,K-ATPase) is a plasma membrane protein which plays an essential role in the maintenance of intracellular concentrations of Na⁺ and K⁺ in all animal cells (see reviews [1–3]). These ion gradients subserve important physiological functions such as nerve excitation and conduction, muscle excitation, and epithelial ion transport. Na,K-ATPase consists of two subunits, designated α and β [1–3]. The catalytic functions of the enzyme have been attributed to the α -subunit, while the function of the β -subunit remains unclear. Species- or isoform-specific properties have been correlated with sequence differences in order to identify conserved or variable functional domains.

Comparisons of α -subunits from several species have shown high homology between the α -subunits from human [4], pig [5], sheep [6], *Torpedo* [7], rat [8], chicken [9], and *Drosophila* [10]. The α -subunit from *Drosophila* is the most divergent, exhibiting about 80% identity to mammalian forms [10].

Our laboratory has been studying the structure and

developmental regulation of the Na,K-ATPase in the brine shrimp, *Artemia* [11–13]. In addition to being a useful developmental model, the brine shrimp is of interest from an evolutionary standpoint. It is an ancient organism, apparently diverging from the precursors of the vertebrates 0.5–1 billion years ago. At the same time, the brine shrimp demonstrates a high rate of mutation, resulting in its classification as a 'fast clock' organism [14]. Thus, comparison of the amino acid sequences of the Na,K-ATPase of brine shrimp and higher organisms has potential for differentiating functionally restricted (conserved) domains and non-essential (divergent) domains not already revealed by sequence comparison of the highly homologous vertebrate enzymes.

We report here the molecular cloning, the nucleotide sequence, and the predicted amino acid sequence of an α -subunit of the brine shrimp as well as a comparison of Na,K-ATPase sequences from all species so far cloned.

2. MATERIALS AND METHODS

2.1. Construction of libraries

Brine shrimp cysts (San Francisco Bay Brand, Lot 1521) were hydrated and grown in half-strength sea water, as described previously [15]. After 18 h, the brine shrimp were sacrificed, total RNA was isolated [16], and polyadenylated RNA was enriched by chromatography through oligo(dT) cellulose (Type 7, Pharmacia). Complementary DNA (cDNA) was synthesized according to the method of Gubler and Hoffman [17]. *Eco*RI sites were methylated [18], and *Eco*RI linkers (New England BioLabs) were ligated to the ends of the cDNA [18]. Separation of excess linkers and size fractionation of the cDNA were accomplished by electrophoresis on agarose gels and electroelution (IBI electroeluter) of cDNA from 1–2-mm thick slices of the gel. Several size-fractionated cDNA libraries were constructed by ligation of the electroeluted cDNA fractions into λ gt10 or λ gt11 arms

Correspondence address: L.E. Hokin, Department of Pharmacology, University of Wisconsin Medical School, 1300 University Avenue, Madison, WI 53706, USA

*Present address: The Blood Center, 1701 W. Wisconsin Avenue, Milwaukee, WI 53233, USA

^oPresent address: Department of Physiology, University of Manitoba, 770 Bannatyne Avenue, Winnipeg, Manitoba R3E 0W3, Canada

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession no. Y07513

(Stratagene). Packaging and transfection were performed using a packaging kit according to the methods recommended by the manufacturer (Stratagene).

2.2. Screening of libraries

The λ gt11 expression library was screened using polyclonal antisera raised against highly purified brine shrimp α -subunits [18,19]. The λ gt10 library was screened using standard DNA hybridization methods [20] with cDNA probes that were labelled with ^{32}P by nick-translation [20] or a random primer method (BMB).

2.3. cDNA analysis and sequencing

Plaque purification, restriction mapping, and subcloning were performed by standard methods [20]. DNA sequencing was accomplished by a dideoxy termination method using a sequencing kit (Sequenase, USB). The templates for the sequencing reactions were derived from subclones in M13mp18 or M13mp19 vectors (single-stranded) or double-stranded DNA fragments isolated from pBR322 subclones. Primers for initiating the extension reactions within the vectors (M13 and pBR322) were purchased from Pharmacia, and cDNA-specific oligonucleotide primers (5'-dGCTGCTGTATTGCT and 5'-dGAAGCACAGAATGCA) were purchased from the University of Wisconsin Biotechnology Center. The entire sequence was determined for both strands of DNA, and the data were analyzed using the UWGCG sequence analysis programs [21].

3. RESULTS AND DISCUSSION

3.1. Nucleotide and predicted amino acid sequence

The composite nucleotide sequence determined from α 2850 and 273 bases at the 5'-end of α 1290 (fig.1a) is shown in fig.1b. The predicted amino acid sequence is shown below the nucleotide sequence. This sequence contains a 1008 amino acid open reading frame, beginning at the 5'-end of the sequence. Several observations, apart from maintaining an open reading frame, are consistent with the assignment of the initiation methionine at ATG, 38-40. The amino terminal sequence of our brine shrimp strain aligns with that of a small N-terminal end of an α -subunit determined by amino acid sequencing [22]. The coding sequence that precedes this methionine shows no homology to the 5'-end of α -subunits from other species and also lacks the characteristic high content of basic residues in the amino termini of α -subunits. The most convincing evidence for our assignment of this methionine as the start methionine is that the N-terminal tripeptide, MGK, is the same in our strain of brine shrimp as that of *Torpedo*, human, rat I (the predominant isoform), sheep, and chicken. Only *Drosophila*, which has an ex-

tended N-terminal end and, surprisingly, the brine shrimp strain used by Morohashi and Kawamura [22] (fig.2) break this rule. If this methionine is assumed to be the amino terminus, the brine shrimp strain α -subunit contains 996 amino acids with a molecular mass of 111021 Da. This sequence contains five potential glycosylation sites. The polyadenylation signal, AAUAAA [23], is found 18 residues upstream of the poly(A) tail, within the expected range of 11-30 residues [24].

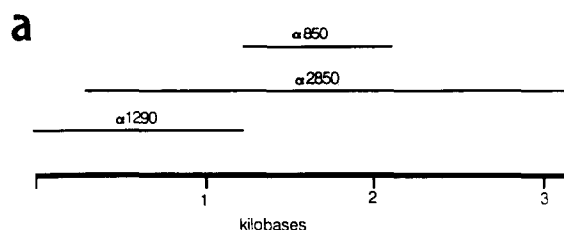
3.2. Homology between other Na,K-ATPase α -subunits

The amino acid sequence shows only 62% identity with a sequence which was previously reported by Morohashi and Kawamura for 26 residues at the amino-terminal end of the brine shrimp α -subunit (fig.2) [22]. We believe the most likely explanation for the low homology is that the sequences were derived from different strains of brine shrimp. The protein sequence determined by the Japanese workers [22] was derived from a 'Tetra Brand' strain of *Artemia* and to our knowledge has not been typed. Our brine shrimp was obtained from San Francisco Bay Brand Co., and a DNA sample of our strain was shown by restriction map analysis to be derived from *Artemia sanfranciscana* by J.C. Bagshaw (personal communication). Brine shrimp exhibit a high rate of mutation [14] which could result in substantial differences between the sequences derived from different strains. Restriction fragment length polymorphism in several genes has been observed in different strains of brine shrimp throughout the world (Bagshaw, personal communication).

Isoforms of the brine shrimp α -subunit have been identified [12,13,15,19,22] and derivation of sequences from different isoforms could conceivably explain the lack of identity between the sequence reported here and that reported by Morohashi and Kawamura [22], but it is unlikely that substantial amounts of isoforms other than α_1 and α_2 would have escaped detection in both studies.

If the ATG at positions 38-40 is selected as the initiator methionine, the brine shrimp α -subunit lacks 23 N-terminal residues present in mammalian and *Torpedo* α -subunits and 41 present in *Drosophila*

Fig.1. Brine shrimp Na,K-ATPase cDNA. A brine shrimp cDNA library was constructed in the expression vector, λ gt11. Thirty plaques gave positive signals in an initial screening with polyclonal α -subunit antibody. The nucleotide sequence of one cDNA insert (α 850) was determined, and comparison with previously reported sequences from other species confirmed that this cDNA encoded an α -subunit of the Na,K-ATPase. This cDNA was purified from the vector, labelled, and used as a hybridization probe for screening a λ gt10 library containing cDNA inserts larger than 1500 bp. An initial screening identified 42 plaques with positive hybridization to the cDNA probe; 17 were plaque purified, and the nucleotide sequence of the longest cDNA (α 2850) was determined. Since this cDNA was not full-length, additional screening was undertaken to locate the missing portion of the coding sequence. Hybridization probes derived from the 5'-end of α 2850 were used to select another clone (α 1290) which overlapped α 2850 bp and contained an additional 273 bp of sequence. (a) Map of the three clones used for sequence analysis. The entire lengths of α 850 and α 2850 were sequenced on both strands. The first 5' non-overlapping portion derived from α 1290 was also sequenced on both strands. (b) Nucleotide and predicted amino acid sequences of the composite sequence from the three clones. Numbering pertains to the nucleotide sequence, beginning with the first nucleotide in the cDNA clone. The selection of the translated region is described in the text. Potential glycosylation sites are underlined.



1 GAATTCCTGTAACAGTGGTTCACATCTTCACGGATGGGGAAAAAACAAGGAACGCTGTAGTGATCTGAAGAAGAGTGTGACATCTGCATACGACATAAAATCCACTTGAGGAATATGCACAGACCTTGGTACAAATACGGAAC 150
M G K K Q G K Q L S D L K K E L E L D Q R K I P L E E L C R R L G T N T E T

151 TGGCTTAACATCGTCCGACGAAAGCTCTCAATTTAGAAAAATATGGTGTCCAAATGACCTTACTTCCCAAGAACTACTCAGCAATGATCAAAATTTTGAAGAGCTTTTTCAGAGATTTCAGATGCTTCTATGGATGGATCAATTTTGTG 300
G L T S S Q A K S H L E K Y G P N A L T P P R T T P E W I K F C K Q L F G G F Q M L L W I G S I L C

301 CTTCATTCATATCAATGAAAAATATAAAATCCAGATCTTCTAGGTGATAATCTTTACCTTGGTCTGCTCTCTTTTTCGTGATATGACAGGATGTTTGCATCTATCAAGATCACAAGCGCTCAAAAATATGGACCTCTTT 450
F I A Y T M E K Y K N P D V L G D N L Y L G L A L L F V V I M T G C F A Y Y Q D H N A S K I M D S F

451 CAAGAACCCTATGCCACAGTTTGTCTTTGTATCCGAGATGGAAAAAATTAACCTGAAGCCGAGAGCTAACAGTAGTGATCTAGTGTGAAGTAAATTTGGTGATAGAAATCCAGCTGATATCAGAAATCTCTTCTGCCAAGATAT 600
K N L M P Q F A F V I R D G K K I Q L K A E E V T V G D L V E V K F G D R I P A D I R I T S C Q S M

601 GAAGGTGTGAATCTCTTCTCAACAGGAGATCCGAACTCTCAATCTAGAGACACAGAAATGACATTAATGACAAATCTTTTGAAGAACAAAGATTTAGGTTTTCCTTCCACAACTTTTGAAGGACCTGGAAGAGGTATGTATCAATGT 750
K V D N S S L T G E S E P Q S R S T E C T N D N P L E T K N L A F F F T N T L E G T G R C I V I N V

751 CGGAGACGACTCTGAATGGCTAGAAATTCATGTTTAGCTTCTAGTTTGAATCTGTGTAAGTCCAAATGCCACAGAAATGAACATTTTATCCATATATTATCTGCTATGGCTGTATCCCTAGCTGCTGATTTTCTGCTAAATTCAT 900
G D D S V M G R I A C L A S S L D S G K T P I A R E I E H F I R I I T A M A V S L A A V F A V I S F

901 TTTATATGGATACACTTGGCTGAAGCTGCTATTTTCATGTTGGTATTTATCGTCCGAAAGTACCAGAGGCTTACTAGCTACTGTAAACAGCTGTTTAACTATGACAGCTAAGCGTATGCCAAGAAAACTGTTTGGTCTGAACT 1050
L Y G Y T W L E A A I F M I G I I V A K V P E G L L A T V T V C L T L T A K R M A K K N C L V R N L

1051 TGAAGCTGTGGAACCTTAGGTTCTACATCAACCTTTTGTCTAGATAAGACTGCGACATTAACCAAAACCGAATGACAGCTGTCTCATATGTGGTTCGACACAGAAATGTCAACGACACACCTGAAACCACTGGTGAACCACT 1200
E A V E T L G S T S T I C S D K T G T L T Q N R M T V A H M W F D Q K I V T A D T T E N Q S G N Q L

1201 ATATCGAGTAGTAAGAGATTTCTGAGCTTATCCGAGTTGCATCACTTTGACATCGGGCTGAATTCAGAGCTGAACATGCTCATTTGCCGTGTGTAAGCGGTGATGTTAAGCGGATGGCTAGAGAGCTCTATTCTGAATTTGCAGA 1350
Y R G S K G F P E L I R V A S L C S R A E F K T E H A L P V L K R D V N G D A S E A A I L K F A E

1351 GATGTCACCTGGGTCTGAATGAACATAAGAGCAGCAGAAAGGTTTCTGAATTCCTTTCAATTCGTGCCAATAAATCAAGTTTCTGTCACGAAGACAGACAAATCTGGTTATTTCTGTTATGAAGGTGCTCCGGAACG 1500
M S T G S V M N I R S K Q K K V S E I P F N S A N K Y Q V S V H E R E D K S G Y F L V M K G A P E R

1501 AATTCAGAACGGTGTTCACCATCTTAATAGAGCCACAGAAATACCACTTGACACACATGAAGGAGCTTTTCAACAATGCCATATGAGCTAGGAGGACCTGGGTGAAGAGCTTTTAGGTTTCTGTGACTTCGAAATACCACTGA 1650
I L E R C S T I L I D G T E I P L D N H M K E C F N N A Y M E L G G M G E R V L G F C D F E L P S D

1651 TCAATATCCAGAGGCTATGATTTTATGATCTGATGAACCTAATTTCCCAATAAGCGGTCTAGTTTGTGTGCTCATGTGCTATGATGATGCTCCCAAGAGCTGCCGTCTCTGATGCGGTTTGGAAATAGATCTGCTGGTATTAAGT 1800
Q Y P R G Y V F D A D E P N F P I S G L R P V G L M S M I D P P R A A V P D A V S K C R S A G I K V

1801 AATTATGGTAAGTGTATCAACCAATTAATGCAAAAGCTATAGCCAGACAGTGTGAATCATATCTGAAGGTCAACCAAGCTGTATGATATGCTGTGCTTTAAACATTCCTGTATCGAAGTTAACCACAGATCTGCGCAGCAGC 1950
I M V T G D H P I T A K A I A R Q V G I I S E G H E T V D D I A A R L N I P V S E V N P R S A Q A A

1951 CGTCATCATGGAATGATCTTAAGATATGAATAGCGACCAATAGATGACATCTTAGACATATACAGAGAAATGTGTGTTGCCAGAACTTCTCCGACGACAACTTATCATGTTGAAGGTGTTTCAAGACAGAGGAAATCTGTTCG 2100
V I H G N D L K D M N S D Q L D D I L R H Y R E I V F A R T S P Q Q K L I I V E G V Q R Q G E F V A

2101 TGTACTGTGATGGCGTAAGTACTACACAGCTCTAAGAAGCTGATATGGTGTGTCTATGGGTATGCTGGCTCAGATGTGCTAAACAGCTGCCGATATGATCTTCTTGTATGATATTTTGCATCTATGTAACGTGTGTGA 2250
V T G D G V N D S P A L K K A D I G V A M G I A G S D V S K Q A A D M I L L D D N F A S I V T G V E

2251 AGAAGGGCGTCTCATTTTGCAGCAATATAGAAATCTATGGCTACACTTACTCTTAAATTCCTGAATCTTCCCTTCTGATGATATCTGTTTGAATTTACCTCTTGTATTTGGTACTGTGATCTATTTTGCATTTGATCTTGG 2400
E G R L I F D N I K K S I A Y T L T S K I P E L S P F L M Y I L F D L P L A I G T V T I L C I D L G

2401 AACCAGTGTGTACTGCATTTCTATGGCTTATGAAGCTCCAGAGCAGATCAAGAAAGCCACAGACCGGCTTAAGAAAAAATCTGTCATGAAGCTTGATTTCAATGGCTTACGACAAATTCGGCTTATGCAGCTTTTGGAGG 2550
T D V V P A I S M A Y E G P E A D P R K P R D P V K E K L V N E R L I S M A Y G Q I G V M Q A F G G

2551 ATTCTTCACATACTCTGTATCTGGCGAGTGGGAGTTTTTACCAATGACATTTTGGACCTAAGAAAAATGGTGGGAGCTTAAGGCTATATAGCTTAAACAGATCTTATGACAAAGATGGATCTGGACGCAAGCAAGACAGCTCGA 2700
F F T Y F V I N G E C G F L P N R L F G L R K M W E S K A Y N D L T D S Y G Q E W T W D A R K Q L E

2701 ATATACCTGCACACTGCCTTTTCATATCTATTGTAACTGTCCAACTGACAGATCTAATTTATTTAGACGCGCTGTTTATCACTCTTCCAGCAAGGAATGAAGTGAACCTTTAACTTGGCTCTGTTTTCGACAGCTGTGTGTG 2850
Y T C H T A F F I S I V I V Q M T D L I I C K T R R L S L F Q Q G M K N G T L N F A L V F E T C V A

2851 AGCTTCTCTAGCTACACCCCTGGAATGATAGGCTCTCAGGATGTATCCATTAAGATATGGTGGTGTCTCCACAGATGCCCTTCTCACTTCTATTTCTGCTATGACAGAGTGGCGTAAATTCCTCATGGGAGAAATCCCGGTGG 3000
A F L S Y T P G M D K G L R M Y P L K I W W F P P M P F S L L I L V Y D E C R K F L M R R N P G G

3001 TTTCTCTGAAGCTGAACTATTATTATCACTTAAATGTAATGAGAAATTTTGGATTTTGTCTGTGTTGTAATAAAAAAATGCACTTTTAAAAAAA 3106
F L E R E T Y Y

ARTEMIA NSANKYQVSVHEREDKSG..YFLVMKGAPERILERCSTILIDGTEIPLDNHMKECFNWAYMELGGMGERVLFCDPSPDQYPRGYVFDADPNFPISG
 DROSOPHILA NSTNKYQVSIHETEDTNDPRYLLVMKGAPERILERCSTIFINGKEKVLDEEMKEAFNNAYMELGGGLGERVLGFCDPMLPSDKYPNGFKFNTDDINFPIDN
 TORPEDO NSTNKYQVSIHE..NDKADSRYLVMKGAPERILDRCSITLLNGEDKPLNEEMKEAFQWAYLELGGGLGERVLGFCHLKLSTSKGFPDFDVEEPNFTD
 HUMAN NSTNKYQVSIHKNPNTSEPHLLVMKGAPERILDRCSITLLHGKEQPLDEELKDAFQWAYLELGGGLGERVLGFCHLFLPDEQFPEGFQFDDVNFPIDN
 RAT I NSTNKYQVSIHKNPNASEPKHLLVMKGAPERILDRCSITLLHGKEQPLDEELKDAFQWAYLELGGGLGERVLGFCHLFLPDEQFPEGFQFDDVNFPIDN
 RAT II NSTNKYQVSIHERED..SPQSHVLMKGAPERILDRCSITLLVQGEIPLDKEMQDAFQWAYMELGGGLGERVLGFCDPMLPSGKFPGRGKFDDELNFPTEK
 RAT III NSTNKYQVSIHETEDPNDNRLLVMKGAPERILDRCSITLLQGEQPLDEEMKEAFQWAYLELGGGLGERVLGFCHLFLPDEQFPEGFQFDDVNFPIDN
 SHEEP NSTNKYQVSIHKNANAGEPRHLLVMKGAPERILDRCSITLLHGKEQPLDEELKDAFQWAYLELGGGLGERVLGFCHLFLPDEQFPEGFQFDDVNFPIDN
 CHICKEN NSTNKYQVSIHKNANAGESRHLLVMKGAPERILDRCSITLLHGKVQPLDEETKDAFQWAYLELGGGLGERVLGFCHLFLPDDQFPEGFQFDDVNFPIDN
 0* *0*

ATP

ARTEMIA LRFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIARQVGI I SEGHETVDDIAARLNIPVSEVNPRSAQAQAVIHGNDLKDMSDQLDDIL
 DROSOPHILA LRFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKSVGI I SEGHETVEDIAARLNIPVSEVNPREAKAAVHGAELRDVSSDQLDEIL
 TORPEDO LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SEGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDDIL
 HUMAN LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SEGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDDIL
 RAT I LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SEGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDDIL
 RAT II LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SEGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDEIL
 RAT III LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SEGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDEIL
 SHEEP LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SEGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDDIL
 CHICKEN LCFVGLMSMIDPPRAAVPDVAGKCRSAGIKVIMVTGDHPITAKAIKGVGI I SDGHETVEDIAARLNIPVQVNPRAKACVHGSGLKDMTSEQLDDIL
 * **** *****0 **** *0***0**** ***** **** * * * * * * *

ARTEMIA RHYREIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 DROSOPHILA RYHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 TORPEDO HYHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 HUMAN KYHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 RAT I RYHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 RAT II RDHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 RAT III QNHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 SHEEP KYHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 CHICKEN LHTEIVFARTSPQKKLI IVEGQORQGAIVAVTGDGVNDSPALKKADIGVAMGIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAITLTS
 00*****0** *00*****0**** *****0*****

ARTEMIA KIPELSPFLMYILFDLPLAIGTVTILCIDLGTDMVPAISLAYEGEAD..PRKPRDPVKEKLVNERLISMAYGQIGVMAQAGGFFTYFVIMGECGFLPNR
 DROSOPHILA NIPETSPFLASILCDIPLPLGTVTILCIDLGTDMVPAISLAYDHAEADIMKRPDRDPFNDKLVNSRLISMAYGQIGMIAQAGGFFTYFVIMGECGFLPKK
 TORPEDO NIPETIPFLVFI IANVPLPLGTVTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPID
 HUMAN NIPETIPFLIFI IANIPPLGTVTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPIH
 RAT I NIPETIPFLIFI IANIPPLGTVTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPFH
 RAT II NIPETIPFLIFI IANIPPLGTVTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPSR
 RAT III NIPETIPFLIFI IANIPPLGTVTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPGN
 SHEEP NIPETIPFLIFI IANIPPLGTVTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPNH
 CHICKEN NIPETIPFLIFI IANIPPLGTCTILCIDLGTDMVPAISLAYEAEADIMKRPDRPNPKTKLVNERLISMAYGQIGMIAQAGGFFTYFVILAEENGFLPSG
 0**0 *** * *00** *****0****0** 0* *000* ** 0**** *****00** **** * * 0*0****

H5

H6

ARTEMIA LFGRLKWKESKAYNDLDSYGQEWTDARKQLEYTCHTAFFSIVVQWADLI ICKTRRLSLFQQGMKNGTLNFAVFETCVAAFLSYTPGMGDKGLRMYP
 DROSOPHILA LFGIRKMDKSAVNDLDSYGQEWTDARKQLEYTCHTAFFSIVVQWADLI ICKTRRNSIFQQGMNRNALNFAVFETVLAFLSYCPGMGKGLRMYP
 TORPEDO LIGIREKWDDELWTDLEDYSGQWTEYQKIVEFTCHTAFFSIVVQWADLI ICKTRRNSIFQQGMKNKILIFGLFEETALAFLSYTPGTDIALRMYP
 HUMAN LLGLRVDWDRWINDVEDSYGQWTEYQKIVEFTCHTAFFSIVVQWADLVICKTRRNSVFQQGMKNKILIFGLFEETALAFLSYCPGMGVALRMYP
 RAT I LLGIRETWDRWINDVEDSYGQWTEYQKIVEFTCHTAFFSIVVQWADLVICKTRRNSVFQQGMKNKILIFGLFEETALAFLSYCPGMGAALRMYP
 RAT II LLGIRLDWDRRTNDLEDYSGQWTEYQKIVEFTCHTAFFSIVVQWADLI ICKTRRNSVFQQGMKNKILIFGLFEETALAFLSYCPGMGVALRMYP
 RAT III LVGIRLNDWDRRTNDLEDYSGQWTEYQKIVEFTCHTAFFSIVVQWADLI ICKTRRNSVFQQGMKNKILIFGLFEETALAFLSYCPGMGVALRMYP
 SHEEP LLGIRLNDWDRWINDVEDSYGQWTEYQKIVEFTCHTAFFSIVVQWADLVICKTRRNSVFQQGMKNKILIFGLFEETALAFLSYCPGMGVALRMYP
 CHICKEN LVGIRLNDWDRWINDVEDSYGQWTEYQKIVEFTCHTAFFSIVVQWADLI ICKTRRNSVFQQGMKNKILIFGLFEETALAFLSYCPGMGVALRMYP
 * * * 0 * * * * * * * * * * * 0 * * * * * * * * * * * 0 * * * * * * * * * * *

H7

H8

ARTEMIA LKIWWFPPMPFSLILVYDECRKFLMRRNPGGFLERETYY
 DROSOPHILA LKLVWFPAIPFALAI FVYDETRRFYLRNPGGWLEQETYY
 TORPEDO LKPSWFCAPYSLLIFVYDEARRFIRNPGGWVEQETYY
 HUMAN LKPTWFCAPYSLLIFVYDEVKLIIRRRPGGWVEKETYY
 RAT I LKPTWFCAPYSLLIFVYDEVKLIIRRRPGGWVEKETYY
 RAT II LKVTWFCAPYSLLIFVYDEVKLIIRRRPGGWVEKETYY
 RAT III LKPSWFCAPYSLLIFVYDEIRKLIIRRRPGGWVEKETYY
 SHEEP LKPTWFCAPYSLLIFVYDEVKLIIRRRPGGWVEKETYY
 CHICKEN LKPTWFCAPYSLLIFVYDEIRKLIIRRRPGGWVERETYY
 ** *** 0 * * * 0 *** * * * * 0 * * * *

the larger α -subunits are not essential for enzymatic activity. This concept is supported by the high degree of variability in this region. However, the N-terminal tripeptide appears to be important since it is conserved from brine shrimp to man with the exception of *Drosophila* and the Tetra Brand of *Artemia*. Also, the functional importance of at least part of the 30 residues from the amino terminus of the mammalian α -subunit is suggested by proteolytic cleavage studies that showed that cleavage at position 30 altered binding of ions and conformational changes that occur during the transport cycle [26].

Since the brine shrimp α -subunit has about 69–72% identity with previously reported sequences from other species (the largest degree of divergence identified to date) identification of conserved regions in the brine shrimp may be useful for identification of essential structural components of the enzyme.

Models of the structure of the α -subunit suggest four transmembrane domains at the amino terminus, a large cytoplasmic loop which has been implicated in the hydrolysis of ATP, and 3–4 transmembrane segments at the carboxy terminus. Transmembrane domains have been predicted based upon hydrophobicity plots [7] (fig.3), and their high homology is consistent with their functional importance in selective ion transport. A striking exception to this is H3, in which the brine shrimp sequence shows considerable difference from that of other species. These changes, however, maintain the general hydrophobic nature of the region, as shown by hydrophobicity plots of the α -subunit of the brine shrimp (data not shown). This observation suggests that this domain may be important for maintaining the general structure of the α -subunit but may not play an important role in the transport of ions.

Several segments of the large cytoplasmic loop have been implicated in the phosphorylation of the α -subunit and the hydrolysis of ATP. These show absolute conservation in all species, despite the large divergence between the sequences of the brine shrimp and other species studied. The 75 residues surrounding the phosphorylation site (fig.3) show complete identity between *Torpedo*, sheep, rat, and human, with a single R→S change in *Drosophila* [11] (fig.3). Although the brine shrimp sequence shows a relatively high level of conservation in this region, there are some substitutions (V→A, L→M, N→K, K→R, N→Q, Q→K) (fig.3), which conserve charge. Twenty-nine residues surrounding the phosphorylated aspartic acid residue show absolute conservation in all available sequences (fig.3).

Another region that has been implicated in ATP binding [27,28] is absolutely conserved in all α -subunits (LVMKGAPERIL) (fig.3), but the surrounding regions show low homology. This peptide contains a fluorescein isothiocyanate-reaction region which is also conserved in other ion-transporting ATPases [29,30]. There is also an amazing conservation of sequence in

the 200 residues that precede H5 (fig.3), consistent with the concept that this region may be involved in hydrolysis of ATP.

The molecular cloning of the α -subunit of the brine shrimp Na,K-ATPase is not only important for its unique contribution to the compendium of amino acid sequences, but also for its usefulness in studying developmental regulation of the Na,K-ATPase and evolutionary aspects. The brine shrimp was selected as a model system for studying developmental regulation of the Na,K-ATPase because its enzymatic activity rapidly increases following hydration of dormant cysts. Since the Na,K-ATPase is strictly an animal enzyme, this model system is one of the few such systems where development and synthesis of a vital enzyme can be turned on by manipulation of the environment. The increase in enzymatic activity during development is associated with increased levels of enzyme protein [13,14] and mRNA [13,15]. Molecular cloning, structure analysis, and utilization of hybridization probes provide the basis for further characterization of this developmental regulation [13].

Acknowledgements: The authors wish to thank Kim Huston and Anne Weege-Kern for their technical assistance and Karen Wipperfurth for preparation of the manuscript. This work was supported by Grant GM33850 from the National Institutes of Health. J.Z.G. was supported in part by a Clinical Nutrition Fellowship, University of Wisconsin.

REFERENCES

- [1] Skou, J.C. (1988) *Methods Enzymol.* 157, 1–25.
- [2] Baxter-Lowe, L.A. and Hokin, L.E. (1989) in: *The Red Cell Membrane: A Model for Solute Transport* (Raess, B.U. and Tunnick, G. eds) pp. 185–280, Humana Press, New Jersey.
- [3] Sweadner, K.J. (1989) *Biochim. Biophys. Acta* 988, 185–220.
- [4] Kawakami, K., Ohta, T., Nojima, H. and Nagano, K. (1986) *J. Biochem.* 100, 389–397.
- [5] Ovchinnikov, Y.A., Modyanov, N.N., Broude, N.E., Petrukhin, K.E., Grishin, A.V., Arzamazova, N.M., Aldanova, N.A., Monastyrskaya, G.S. and Sverdlov, E.D. (1986) *FEBS Lett.* 201, 237–245.
- [6] Shull, G.E., Schwartz, A. and Lingrel, J.B. (1985) *Nature* 316, 691–695.
- [7] 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.
- [8] Shull, G.E. and Greer, J. (1988) *J. Biol. Chem.* 263, 8646–8657.
- [9] Takeyasu, K., Tamkun, M.M., Renaud, K.J. and Fambrough, D.M. (1988) *J. Biol. Chem.* 263, 4347–4354.
- [10] Lebovitz, R.M., Takeyasu, K. and Fambrough, D.M. (1989) *EMBO J.* 8, 193–202.
- [11] Fisher, J.A., Baxter-Lowe, L.A. and Hokin, L.E. (1986) *J. Biol. Chem.* 261, 515–519.
- [12] Peterson, G.L., Churchill, L., Fisher, J.A. and Hokin, L.E. (1982) *J. Exp. Zool.* 221, 295–308.
- [13] Guo, J.Z. and Hokin, L.E. (1989) in: *Cell and Molecular Biology of Artemia Development* (Warner, A.H., MacRae, T.H. and Bagshaw, J.C. eds) Plenum, New York, in press.

- [14] Field, K.G., Olsen, G.J., Lane, D.J., Giovannoni, S.J., Ghiselin, M.T., Raff, E.C., Pace, N.R. and Raff, R.A. (1988) *Science* 239, 748-753.
- [15] Peterson, G.L. and Hokin, L.E. (1980) *Biochem. J.* 192, 107-118.
- [16] Baxter-Lowe, L.A., Yohanan, J.M. and Hokin, L.E. (1988) *Biochim. Biophys. Acta* 943, 343-348.
- [17] Gubler, U. and Hoffman, B.J. (1983) *Gene* 25, 263-269.
- [18] Huynh, T.V., Young, R.A. and Davis, R.W. (1985) in: *DNA Cloning: A Practical Approach*, vol. 2 (Glover, D.M. ed.) pp. 49-78, IRL Press, Oxford.
- [19] Fisher, J.A., Baxter-Lowe, L.A. and Hokin, L.E. (1984) *J. Biol. Chem.* 259, 14217-14221.
- [20] Maniatis, T., Fritsch, F.F. and Sambrook, J.O. (1982) *Molecular Cloning*, Cold Spring Harbor, NY.
- [21] Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* 12, 387-395.
- [22] Morohashi, M. and Kawamura, M. (1984) *J. Biol. Chem.* 259, 14928-14934.
- [23] Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature* 263, 211-214.
- [24] Fitzgerald, M. and Schenk, T. (1981) *Cell* 24, 251-260.
- [25] Shull, G.E., Greeb, J. and Lingrel, J.B. (1986) *Biochemistry* 25, 8125-8132.
- [26] Jorgensen, P.L. and Collins, J. (1986) *Biochim. Biophys. Acta* 860, 570-576.
- [27] Farley, R.A., Tran, C.M., Carilli, C.T., Hawke, D. and Shively, J.E. (1984) *J. Biol. Chem.* 259, 9532-9535.
- [28] Kirley, T.L., Wallick, E.T. and Lane, L.K. (1984) *Biochem. Biophys. Res. Commun.* 125, 767-773.
- [29] Filoteo, A.G., Gorski, J.P. and Penniston, J.T. (1987) *J. Biol. Chem.* 262, 6526-6530.
- [30] Shull, G.E. and Greeb, J. (1988) *J. Biol. Chem.* 263, 8646-8657.