

P-type ATPases in *Caenorhabditis* and *Drosophila*: Implications for Evolution of the P-type ATPase Subunit Families with Special Reference to the Na,K-ATPase and H,K-ATPase Subgroup

H. Okamura¹, J.C. Yasuhara¹, D.M. Fambrough², K. Takeyasu¹

¹Graduate School of Biostudies, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan

²Department of Biology, The Johns Hopkins University, Baltimore, MD 21218, USA

Received: 1 April 2002/Revised: 14 August 2002

Abstract. Here we show a complete list of the P-type ATPase genes in *Caenorhabditis elegans* and *Drosophila melanogaster*. A detailed comparison of the deduced amino-acid sequences in combination with phylogenetic and chromosomal analyses has revealed the following: (1) The diversity of this gene family has been achieved by two major evolutionary steps; the establishment of the major P-type ATPase subgroups with distinct substrate (ion) specificities in a common ancestor of vertebrate and invertebrate, followed by the evolution of multiple isoforms occurring independently in vertebrate and invertebrate phyla. (2) Pairs of genes that have intimate phylogenetic relationship are frequently found in proximity on the same chromosome. (3) Some of the Na,K- and H,K-ATPase isoforms in *D. melanogaster* and *C. elegans* lack motifs shown to be important for α/β -subunit assembly, suggesting that such α - and β -subunits might exist by themselves (lonely subunits). The mutation rates for these subunits are much faster than those for the subunits with recognizable assembly domains. (4) The lonely α -subunits also lack the major site for ouabain binding that apparently arose before the separation of vertebrates and invertebrates and thus well before the separation of vertebrate Na,K-ATPases and H,K-ATPases. These findings support the idea that a relaxation of functional constraints would increase the rate of evolution and provide clues for identifying the origins of inhibitor sensitivity, subunit assembly, and separation of Na,K- and H,K-ATPases.

Key words: P-type ATPase superfamily — Sodium pump — Isoform — Molecular evolution — *Drosophila* — *Caenorhabditis*

Introduction

P-type ATPases constitute a large family of membrane proteins (enzymes) that actively transport charged substrates (e.g., cations and phospholipids) across biological membranes. The transport process carried out by the P-type ATPase family is distinguished from that by the other ATP-dependent ion-pumps in that it involves formation of a phosphointermediate of the enzyme itself [25]. Each member of the P-type ATPase family has as part of its structure a catalytic (or α) subunit that contains the conserved amino-acid motifs for phosphorylation of an aspartyl carboxyl group and for ATP binding. The Na,K-ATPase (sodium pump) and the H,K-ATPase (proton pump), in particular, have an additional, glycosylated β -subunit.

Axelsen and Palmgren [3, 41] have classified P-type ATPases into ten categories based on their sequence similarities and substrate specificities (*see* the legend for Table 1). Phylogenetic analyses of the large collection of deduced amino-acid sequences for P-type ATPase catalytic subunits have shown that this categorization can be applied universally to the P-type ATPases of vertebrate and invertebrate animals, plants, fungi, lower eukaryotes, and even prokaryotes [3, 20, 56; for alignments and references to the deduced amino-acid sequences, *see* 13, 36]. Phylogenetic analysis led to the conclusion that the divergence of substrate-specificity (ion selectivity) occurred early in the evolution of P-type ATPases and has been conserved ever since [24].

Table 1. List of *C. elegans* and *D. melanogaster* P-type ATPase genes identified in the present and previous studies.

Species	Type	Isoform	Accession number	Gene name	Ref.
<i>C. elegans</i>	IB	Ce IB	AL032665 Z93382 ^a D83665	Y76A2A F45G2 ^a :cua-1	46
	IIA	Ce IIA-1	Z92807 AJ012296	K11D9.2:mca-4	
	IIA	Ce IIA-2	Z82088	ZK256.1	
	IIB	Ce IIB-1	Z68221 AJ223616	W09C2.3: mca-1	29
	IIB	Ce IIB-2	AF125446 AJ010708	R05C11.3: mca-2	29
	IIB	Ce IIB-3	AC025724 AJ010646	Y67D8C.b: mca-3	29
	IIC	Ce IIC-1	Z81028 U18546	B0365.3: eat-6	11
	IIC	Ce IIC-2	CAB07586	C01G12.8	
	IIC	Ce IIC-3	AF016433	C09H5.2	
	IIC	Ce IIC-4	AF016446	C02E7.1	
	IIC	Ce IIC-5	AL132876	Y105E8A12	
	IV	Ce IV-1	U80025	F02C9.3	
	IV	Ce IV-2	Z81078	F36H2.1	
	IV	Ce IV-3	U28940	T24H7.5	
	IV	Ce IV-4	Z98866	Y49E10.11	
	IV	Ce IV-5	AF099920	H06H21.10	
	IV	Ce IV-6	Z93785	W09D10.2	
	V	Ce V-1	Z70271	W08D2.5	
	V	Ce V-2	U41552	K07E3.7	
	V	Ce V-3	Z83217	C10C6.6	
V	Ce V-4	AC024842	Y59H11AR.2		
	β -subunit	Ce β -1	Z81037	C17E4.9	
	β -subunit	Ce β -2	Z82262	C43F9.6	
	β -subunit	Ce β -3	Z81550	F53F3.3	
<i>D. melanogaster</i>	IB	Dm IB	AE003487	CG1886	
	IIA	Dm IIA-1	AE003462 M62892	Ca-P60A	32
	IIA	Dm IIA-2	AE003598	CG7651	
	IIB	Dm IIB	AE003844	CG2165	
	IIC	Dm IIC-1	AE003732 AF044974	Atpalpha	51
	IIC	Dm IIC-2	AE002610 AF202632	CG 17923	
	IIC	Dm IIC-3	AE003457	CG3701	
	IV	Dm IV-1	AE003818	CG17034	
	IV	Dm IV-2	AE003639	CG6263	
	IV	Dm IV-3	AE003625	CG18419	
	IV	Dm IV-4	AE003694	CG14741	
	IV	Dm IV-5	AE003502	CG9981	
	IV	Dm IV-6	AE003502	CG4301	
	V	Dm V-1	AE003845	_____b	
	V	Dm V-2	AE003631	CG6230	
	β -subunit	Dm β -1	AE003615 U22438	CG9258: Nervana 1	50
	β -subunit	Dm β -2	AE003615 U22439	CG9261: Nervana 2	50
	β -subunit	Dm β -3	AE003670	CG8663	
β -subunit	Dm β -4	AE003725	CG11703		
β -subunit	Dm β -5	AE003725	CG5250		
β -subunit	Dm β -6	AE003646	CG15274		

Categorization of the catalytic subunit genes is after Axelsen and Palmgren [3]; i.e., type-IA (represented by bacterial K-ATPase), type-IB (Cu^{2+} - or Cd^{2+} -ATPase), type-IIA (organelle membrane Ca^{2+} -ATPase (SERCA)), type-IIB (plasma membrane Ca^{2+} -ATPase (PMCA)), type-IIC (Na,K-ATPase or H,K-ATPase), type-IV (phospholipid translocase), and type-V (ATPase with no assigned specificity). The accession numbers show the genome project data in the NCBI database. The numbers and names in bold show the data given in the references before the genome project.

^aThese isoforms step across two cosmid clones and are not shown as CDS in database.

^bThis isoform is not shown as CDS, but detected by BLAST search.

Within the vertebrates, a single type of ion pump often occurs as multiple isoforms, in some cases as the result of alternative splicing and in other cases as products of divergent genes that arose by gene duplications. For the Na,K- and the H,K-ATPase catalytic subunit, several isoforms with ~90% amino acid sequence identity are found [56], and a given isoform from one species has an or-

thologue in other species. For example, human Na,K-ATPase isoforms α 1, α 2, and α 3 genetically correspond to chicken Na,K-ATPase isoforms α 1, α 2, and α 3, respectively [52; for the vertebrate β -subunit isoforms, see 57]. Isoforms also occur in invertebrates, but these isoforms never correspond one-to-one with the vertebrate isoforms, suggesting that the divergence of the P-type ATPase isoforms

occurred after vertebrate and invertebrate lineages separated [56].

Completion of genome sequencing projects for the nematode, *Caenorhabditis elegans* [4], and the fruit fly, *Drosophila melanogaster* [12], has made possible a detailed exploration of phylogenetic relationships among various isoform genes. In the present study, we analyzed a complete list of P-type ATPase sequences extracted from these two model organisms and found novel isoforms that retain what appears to be an ancestral characteristic of the Na,K- and H,K-ATPases.

Materials and Methods

PHYLOGENETIC ANALYSIS

Dm2C2 full-length cDNA clones were isolated from *Drosophila* cDNA libraries by hybridization screening with the Dm2C1 cDNA as probe. Dm2C3, Dm β -5, and Dm β -6 were isolated by reverse transcription PCR of tissues. Searching the genome and EST databases released by the Sanger Centre and by the Berkeley *Drosophila* Genome Project identified the other novel P-type ATPase genes. The previously reported vertebrate and invertebrate P-type ATPase sequences were used as probes to carry out BLAST searches [1]. In the BLAST search at the National Center for Biotechnology Information web site, many P-type ATPase amino-acid sequences of mammalian and invertebrate have been used as query [3]. These query sequences include all of human, *Drosophila*, *C. elegans* and *Tetrahymena* enzymes such as SERCA, PMCA, NaK, HK and β -subunit, Menkes disease-associated copper transporter, ATPase II, *Drosophila* NaK1, *C. elegans* NaK, and *Tetrahymena* TPA. These sequences were also used in this paper, and many distinct subject (answer) sequences were identified. From the sequences identified by BLAST search, the ones that have the unique conserved motifs of P-type ATPase were defined as the putative P-type ATPase isoforms. The conserved motifs are as follows: X T/D GE T/S (unknown), DKTGTLT (phosphorylation site), KGAXX (FITC binding in NaK), M/V XTGD (unknown), GDGXND (FSBA binding in NaK).

The nucleotide sequences of the identified EST clones were determined, open reading frames were identified, and putative coding sequences were converted to amino-acid sequences. Where full-length cDNA sequences were not available, the putative coding sequences of most isoforms were extracted from the CDS sequences in the database directly. For some isoforms (e.g., Ce IB or DmV-1, see Table 1 and legend), the putative sequences were obtained by the FGENES program in the Computational Genomics Group web page (in the Sanger Centre [49]) and by eye. The alignments of the amino-acid sequences were achieved with the CLUSTAL W program [55] in Vector NTI suite6 software. Neighbor-joining phylogenetic trees [45] were constructed based on the conserved regions of the alignment, using the SEQBOOT, PROTDIST, NEIGHBOR, and CONSENSE of the PHYLIP package version 3.5 [15] and visualized using the TREEVIEW program [40]. The number of the amino-acid residues used for the trees are 664(IIA), 976(IIB), 879(IIC α) and 171(IIC β). Bootstrap analyses for 1000 replications were performed to provide confidence estimates for tree topologies.

REVERSE TRANSCRIPTION PCR (RT-PCR)

RT-PCR was employed to verify transcription of putative genes encoding Na,K- and H,K-ATPase α - and β -subunits. Wild-type

adult *D. melanogaster* and *C. elegans* were homogenized and processed for total RNA isolation with RNeasy kits (QIAGEN). The RNA samples were treated with RNase-free DNaseI (New England BioLabs) to remove any residual genomic DNA. Two μ g RNA from each sample were used for poly-T primed reverse transcription. Each PCR reaction was performed with isoform-specific primers and one-twentieth volume of the synthesized cDNA as a template. Primers used for RT-PCR were as follows:

Dm α 2-5'[AGGTATTGGTGTCTTTTCGCCA],
 Dm α 2-3'[GCGTCTAATTCTTCAACTTTCA],
 Dm α 3-5'[ACGGATCAGCTCCACTGGAT],
 Dm α 3-3'[GTGGCCAGGCAATAGGCCAG],
 Dm β 3-5'[TAGTCGACTAGCCGATAAAAAAATTGGT],
 Dm β 3-3'[TACTGCAGCATTGGCTAGTTTCACT],
 Dm β 4-5'[ATGTGCGACTACCGGAGGACGTAATGGTG],
 Dm β 4-3'[TACTGCAGCAATTTGTGCTCGTCTAT],
 Dm β 5-5'[ATAAGCTTTCCATGCCTAACGACGCGATA],
 Dm β 5-3'[ATAAGCTTCTGATCTACTCGTACCCTC],
 Dm β 6-5'[ATAAGCTTAAGATGCCAACCAAGAAGAT],
 Dm β 6-3'[ATAAGCTTTTATAATGAGTAGTTCGGTTCCG],
 Ce α 2-5'[AGGAAAACGTGCAGGTCACA],
 Ce α 2-3'[ACTTCATTCACAAACGGAGTGT],
 Ce α 3-5'[GACCAGAACTTCTGGATTAAGTG],
 Ce α 3-3'[GCTTGAGCAGCCATGAAGAC],
 Ce α 4-5'[GGTCAGAACTCCAGCATTGACCC],
 Ce α 4-3'[GCCTGATTGGCCATATACAT],
 Ce α 5-5'[ATAGGACAGAGTGTCCAGGCC],
 Ce α 5-3'[TTCGAGATCTTGGCGGAGTG],
 Ce β 1-5'[ATGGAGAAAAAGGTGCGACG],
 Ce β 1-3'[CTACAACCTTTTCTTCTTCT],
 Ce β 2-5'[ATGGTTGGGACCAATTCCT],
 Ce β 2-3'[CTAACCAACAGTTTCCACTG],
 Ce β 3-5'[ATGACAAACGGCGGTAGAAATG],
 Ce β 3-3'[TTAAGCGGCTGGTGCAGC].

Equal volumes of the PCR products were subjected to agarose-gel electrophoresis and visualized with ethidium bromide dye. No amplification was observed when reverse transcription was omitted.

Results

P-TYPE ATPASES IN *C. ELEGANS* AND *D. MELANOGASTER*

There are twenty-one P-type ATPase catalytic (α -) subunit genes and three β -subunit genes in the *C. elegans* genome and fifteen catalytic subunit genes and six β -subunit genes in the *D. melanogaster* genome (Table 1). Newly evolved isoforms sometimes lose their ability to be expressed, remaining in the genome as nonfunctional “fossils”, or pseudogenes, that may evolve rapidly [27]. However, it is unlikely that any of the type-IIc ATPase genes identified in this study are pseudogenes, because all of them are transcribed, as confirmed by RT-PCR (*data not shown*).

THE CATALYTIC SUBUNITS

All of the *C. elegans* and *D. melanogaster* P-type ATPase catalytic subunit genes can be classified into six subgroups (Table 1) by multiple alignment of the deduced amino-acid sequences with the corresponding vertebrate sequences (*not shown*). These six sub-

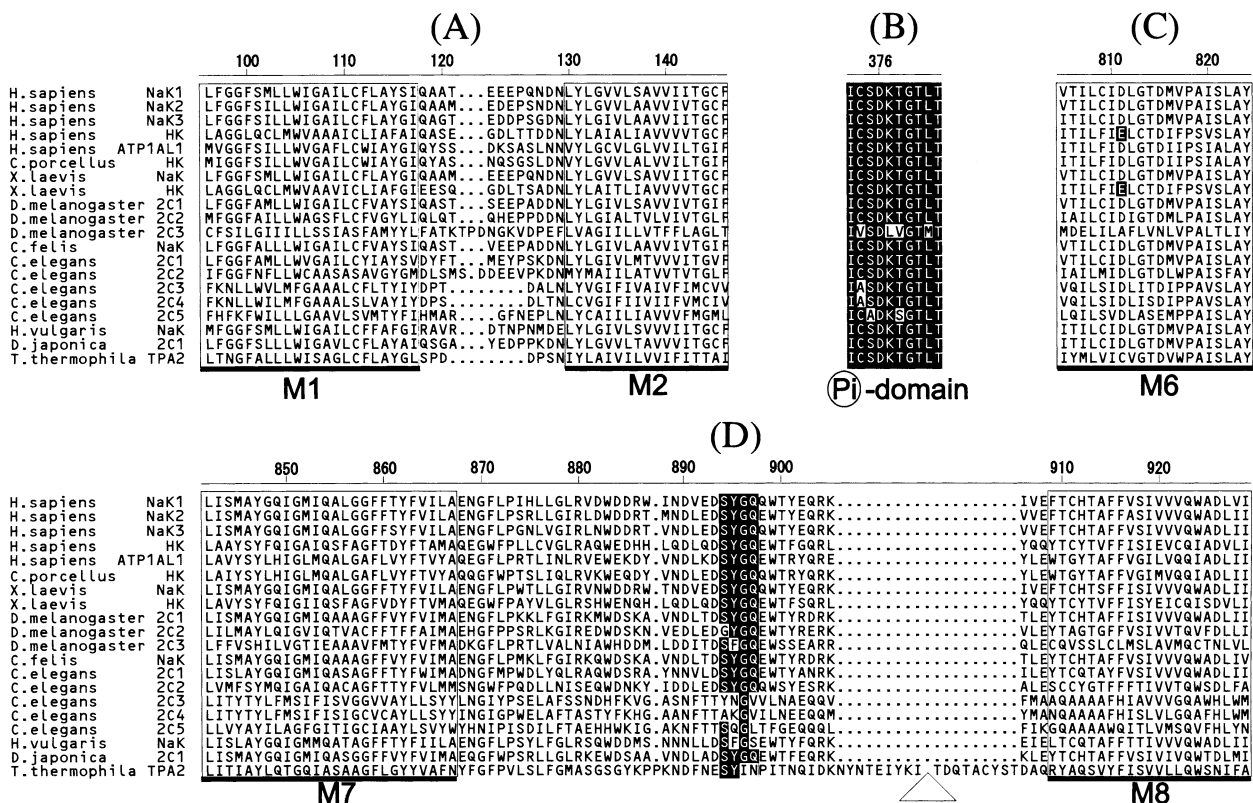


Fig. 1. Amino-acid sequence comparison among the Type-IIC P-type ATPase α -subunits. The amino-acid numbers and the numbers of transmembrane regions are according to the human Na,K-ATPase α -1 subunit sequence. (A) Ouabain-binding extracellular region between the transmembrane M1 and M2. (B) Phosphorylation region in the large cytoplasmic domain. (C) The specific residue (Glu) that binds SCH28080 in the gastric H,K-ATPase is in the

black boxes. (D) Subunit-assembly M7-M8 region. One set of critical amino acids for α - β subunit assembly is shown in black. The white triangle shows the omitted 70 amino-acid residues of TPA2. The GenBank accession numbers not included in Fig. 3 are as follows: D21854 for *C. porcellus* H,K-ATPase, P35317 for *H. vulgaris* Na,K-ATPase, S66043 for *C. felis* Na,K-ATPase, AB012391 for *D. japonica*.

groups (in the nomenclature of Axelsen and Palmgren [3]) are the heavy-metal ATPases (type IB), the SERCA ATPases (type IIA), the PM Ca^{2+} -ATPases (type IIB), the Na,K- and H,K-ATPases (type IIC), the phospholipid translocases (type IV), and ATPases with no assigned specificity (type V). Sequence similarity is rather low between the ATPases from different subgroups within a single invertebrate species, whereas sequence similarity between invertebrate and vertebrate ATPases from a single subgroup is quite high. For this reason, assignment of invertebrate ATPases to particular subgroups, based upon homology with the corresponding vertebrate ATPases, was easily achieved. For example, the amino-acid sequence identity between *Drosophila* Na,K-ATPase (α -subunit) and *Drosophila* Ca^{2+} -ATPase (SERCA) is merely $\sim 22\%$, while sequence identities between homologous *Drosophila* and human proteins are up to 77% for the Na,K-ATPase α -subunits and 72% for the Ca^{2+} -ATPases. The fact that the P-type ATPase catalytic subunits of invertebrate organisms are readily classified into the same subgroups that are

found in vertebrate organisms supports the conclusion [3, 20, 56] that the divergence of the P-type ATPase subgroups occurred early in evolution, preceding the split of invertebrates and vertebrates. As expected, type-IA, type-IID, and type-III ATPases, which are thought to be specific to prokaryotes, fungi or plants, are not found in these invertebrate organisms.

The existence of multiple genes for a single class of P-type ATPase in both *C. elegans* and *D. melanogaster* (Table 1) indicates that a multiplicity of P-type ATPase isoforms is a general feature rather than one specific to vertebrates, where isoform diversity has been well documented. Among the subgroups, the diversity of the type-IIC ATPases is of particular interest, because this group is known to function only when assembled with a β -subunit, and different combinations of the α - and β -subunit isoforms could generate a further variety of enzymatic function.

Amino-acid sequence comparison (Fig. 1) and the phylogenetic relationships (Fig. 2A) among the invertebrate and vertebrate type-IIC ATPases reveal

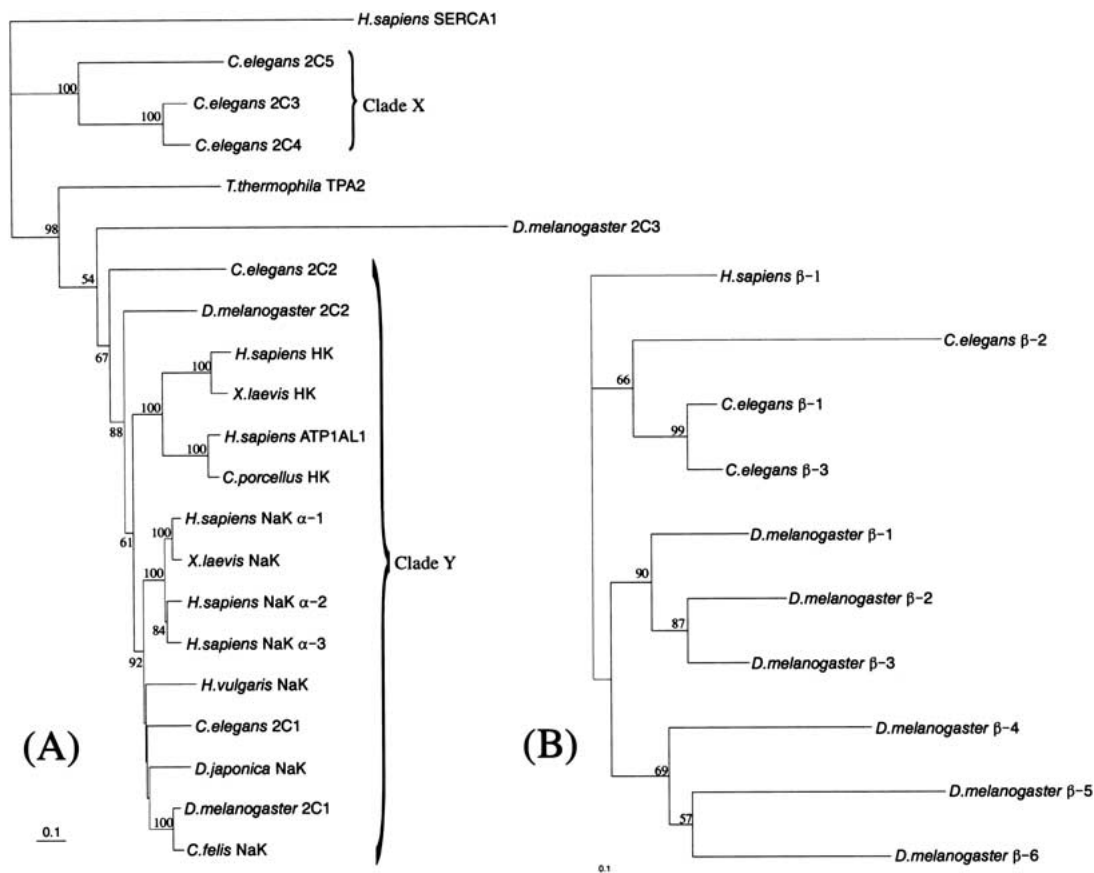


Fig. 2. Rooted phylogenetic trees of the Type IIC P-type ATPase subunits. (A) Alpha-subunit tree constructed by using the same sequence data as in Fig. 1. The root was determined by taking the human SERCA1 sequence as an out-group. Number of used residues is 879. (B) Beta subunit tree constructed by using the sequence data for invertebrate β -subunits. The root was determined by taking the

human β 1-subunit sequence as an out-group. Number of used residues is 171. Both trees are inferred by the neighbor-joining method as described in Materials and Methods, based on the alignment of conserved regions. Numbers at the nodes represent the bootstrap value in 1000 replications. Values less than 50 are not shown. The scale bar represents 0.1 substitutions per amino-acid position.

that three isoforms of *C. elegans* (Ce2C3, 2C4, and 2C5) retain characteristics that may reflect an ancestral form of the Na,K/H,K-ATPase that has not yet acquired the binding sites for ouabain or SCH28080 (see Discussion). These isoforms also appear to lack motifs correlated with α/β -subunit assembly, and we speculate that they may exist without β -subunits (see Discussion). Moreover, the rate of evolution for these subunits, based upon analysis of rooted molecular phylogenetic trees, is much higher than those for the ordinary α -subunits that contain the consensus motif for assembly with β -subunits (Fig. 2A). On the other hand, one isoform of *D. melanogaster* (Dm2C3), which has both the M1-M2 extracellular loop and the α/β subunit assembly motifs (see Discussion), has the highest evolutionary rate (Fig. 2A). The amino-acid sequence for the phosphorylation site (DKTGTLLT) in the catalytic subunits is almost 100% conserved throughout evolution. However, Dm2C3, which is expressed as detected by RT-PCR, has a mutation (Lys³⁷⁰ in the sheep Na,K-ATPase α -subunit to Leu)

in this region. Since no mutagenesis study has been reported on this substitution, the functionality of Dm2C3 remains unknown. This residue has been shown to be critical for Asp to be phosphorylated [34]. Therefore, the highest evolutionary rate of Dm2C3 might come from a relaxation of functional constraint.

THE β -SUBUNITS

The molecular evolution of the vertebrate β -subunit isoforms of the type IIC ATPases has been analyzed by Yu et al. [57], who found that the three β -subunit isoforms for the Na,K-ATPase and one unique β -subunit for the H,K-ATPase appear to be equidistant from a common ancestral sequence. Pestova et al. [42] have recently identified an additional β -subunit in a human genomic DNA library, resulting in a total of five distinct β -subunit isoforms. This is evident in the unrooted tree presented in Fig. 3. In contrast, the invertebrate β -subunit isoforms are simply different

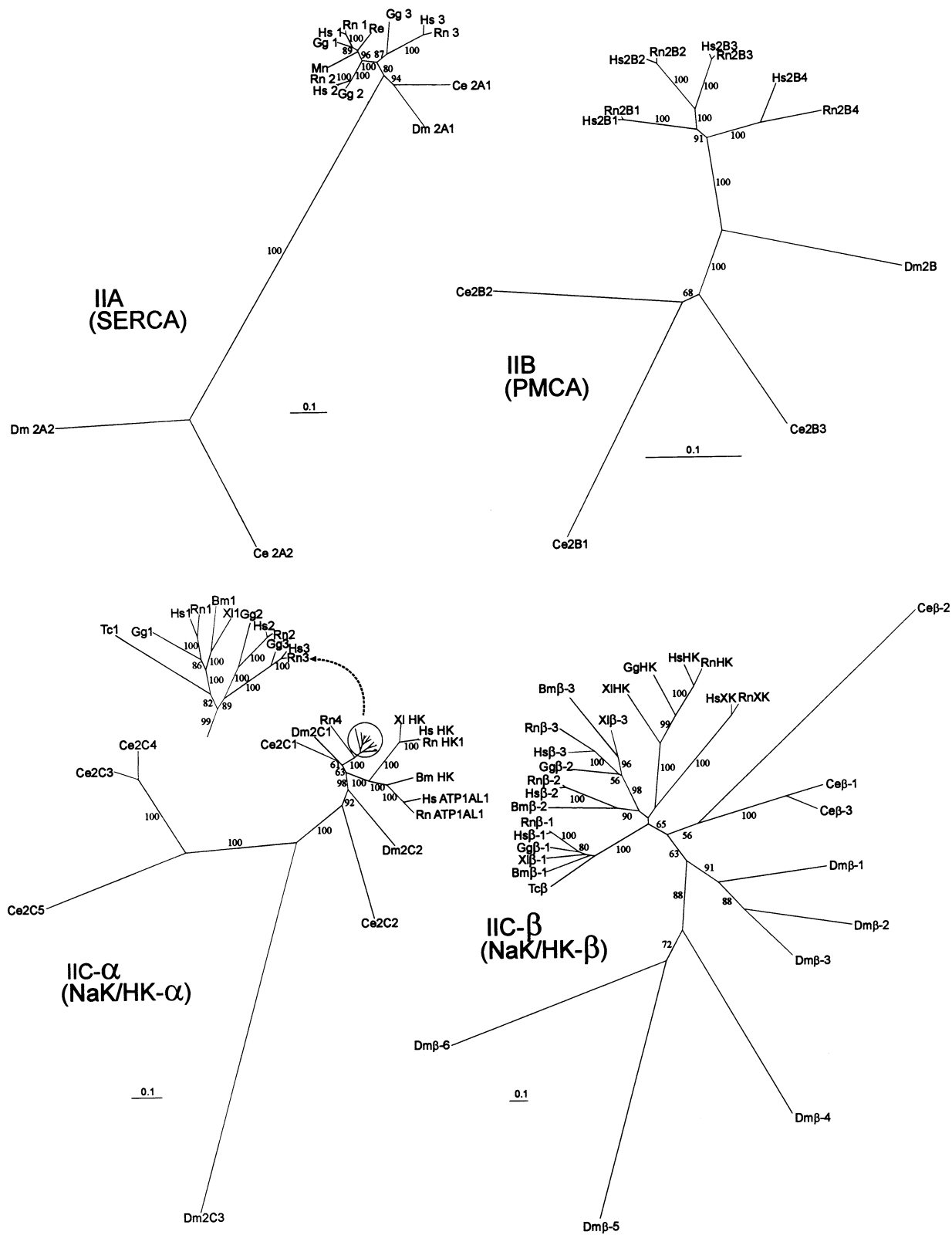


Fig. 3.

from the vertebrate isoforms, as supported by the bootstrap value of nearly 70 (Fig. 3). The *C. elegans* and *D. melanogaster* β -subunits respectively form a unique single cluster. From our analysis of the branching points and the inferred evolutionary rates of the invertebrate β -subunits, it appears that *D. melanogaster* possesses two distinct types of β -subunits, one type (β -1, β -2, and β -3) closer to the vertebrate β -subunits and the other type (β -4, β -5, and β -6) sharing more homology with the *C. elegans* β -subunits (Fig. 2B).

The β -subunits contain several Asn-linked oligosaccharides in their extracellular domain [17, 19, 53, 54], but Asn-linked glycosylation is not required for biosynthesis, assembly or intracellular transport of the β -subunit of the Na,K-ATPase [54]. In the case of the gastric H/K-ATPase, this issue has been addressed by introducing point mutations [2]. Removal of several glycosylation sites resulted in retention of α/β -subunit complexes in the ER, while removal of all seven sites resulted in no assembly, possibly due to misfolding and rapid degradation. Given these divergent results and the high degree of variability both in number and position of glycosylation sites in β -subunits, we did not address this issue in the present analysis. On the other hand, the formation of three S-S bonds is required for proper expression and function of the ATPase on cell surface [26]. All known vertebrate β -subunits contain six conserved cysteines [28, 35] that form three disulfide bonds (Cys¹²⁶-Cys¹⁴⁹, Cys¹⁵⁹-Cys¹⁷⁵, and Cys²¹⁴-Cys²⁷⁷ in the rat β -1 subunit). The assembly of α - and β -subunits in the ER [54] does not require all three S-S bonds, because truncated β -subunits that have only a single disulfide bridge (Cys¹²⁶-Cys¹⁴⁹) assemble with α -subunits. However, these complexes are not able to get to the plasma membrane [21]. In this sense, three of the *Drosophila* β -subunits (Dm β -4, Dm β -5, and Dm β -6) are particularly interesting because these isoforms lack the disulfide bridge (Cys¹²⁶-Cys¹⁴⁹) and contain only one or both S-S bonds further toward the carboxy terminal end (Fig. 4). Therefore, they can be expected not to assemble with the α -subunit.

CHROMOSOMAL DISTRIBUTION OF P-TYPE ATPase GENES

Figure 5 shows the chromosomal distribution of the P-type ATPase genes of *C. elegans* and *D. melanogaster*. In *Drosophila*, two of the type-IV ATPase genes cluster at position 14B-C on the X chromosome, and two each of the four β -subunit genes cluster at positions 27B and 91F on the second and third chromosome, respectively. The β -subunit genes both at 27B and 91F are oriented in a head-to-tail manner, while the α -subunit genes at 14B-C are oriented head-to-head. Interestingly, as seen in the *Tetrahymena* P-type ATPase genes [56], the *Drosophila* genes that are close together on the chromosomes are also positioned close to each other in the phylogenetic trees (Fig. 3). In *C. elegans*, clustering of the phylogenetically related genes is less evident. However, multiple genes that belong to the same category tend to be contained in the same chromosome (e.g., type IIB and type V on chromosome IV, type IIC on chromosome V). It is likely that in these cases the isoforms arose through a local rather than a trans-chromosomal duplication event.

Sun and Salvaterra [50] reported that Nervana (*Drosophila* Na,K-ATPase β -subunit; Dm β -1 and -2 in this study) genes mapped to position 92C-D by in situ hybridization. However, we found all the genomic sequences that contain Nervana genes at position 27B. At position 91F, the stop codon of the first gene (Dm β -5) and the initial codon of the second gene (Dm β -4) are separated by only 184 base pairs. The length of the spacer sequence between the two genes is 4224 bp for the cluster at 27B (between the stop codon of Dm β -2 gene and the initial codon of Dm β -1 gene). Between the nucleotides encoding the phosphorylation sites of Dm IV-5 and IV-6 genes at 14B-C, there are about 10,000 bp.

One interesting finding is that the *Drosophila* β -6 gene has no introns and is localized in the ninth intron of the GABAb receptor gene. In the genome database, Dm β -6 is recognized as a part of one CDS (CG15274), but not as a complete CDS, by BLAST search. There is a potential TATA box 252 bp up-

Fig. 3. Unrooted neighbor-joining trees of type II P-type ATPase α -subunits and β -subunits. Numbers of residues used are 664 (IIA), 976 (IIB), 879 (IIC α -subunits) and 171 (IIC β -subunits). Numbers at the branches represent the bootstrap value in 1000 replications. Values less than 50 are not shown. The scale bar represents 0.1 substitutions per amino acid position. Abbreviation for the vertebrate species are; Bm: *Bufo marinus*, Gg: *Gallus galus*, Hs: *Homo sapiens*, Mn: *Makaira nigricans*, Re: *Rana esculenta*, Rn: *Rattus norvegicus*, Tc: *Torpedo californica*, Xl: *Xenopus laevis*. Genbank accession numbers are as follows: Type-IIA: M26064 (Gg1), M66385 (Gg2), Y18063 (Gg3), U96781 (Hs1), M23114 (Hs2), Z69881 (Hs3), U65229 (Mn), X63009 (Re), M99223 (Rn1), X15635 (Rn2), M30581 (Rn3). Type-IIB: J04027

(Hs2B1), X63575 (Hs2B2), U60414 (Hs2B3), M25874 (Hs2B4), J03753 (Rn2B1), J03754 (Rn2B2), J05087 (Rn2B3), U15408 (Rn2B4). Type-IIC α -subunits: Z11798 (Bm1), Z25809 (Bm HK), J03230 (Gg1), M59959 (Gg2), P24798 (Gg3), X04297 (Hs1), J05096 (Hs2), P13637 (Hs3), M63962 (HsHK), U02076 (Hs ATP1A1), X05882 (Rn1), M14512 (Rn2), M14513 (Rn3), J02649 (Rn HK1), X02810 (Tc), U49238 (Xl1), U17249 (Xl HK). Type-IIC β -subunits: P30715 (Bm1), P43002 (Bm2), P30716 (Bm3), P08251 (Gg1), P33879 (Gg3), I50160 (Gg HK), P05026 (Hs1), P14415 (Hs2), P54709 (Hs3), P51164 (Hs HK), AAD49692 (Hs XK), P07340 (Rn1), P13638 (Rn2), Q63377 (Rn3), P18598 (Rn HK), AAD49694 (Rn XK), P05029 (Tc), B60444 (Xl1), P21188 (Xl3), AAB97472 (Xl HK).

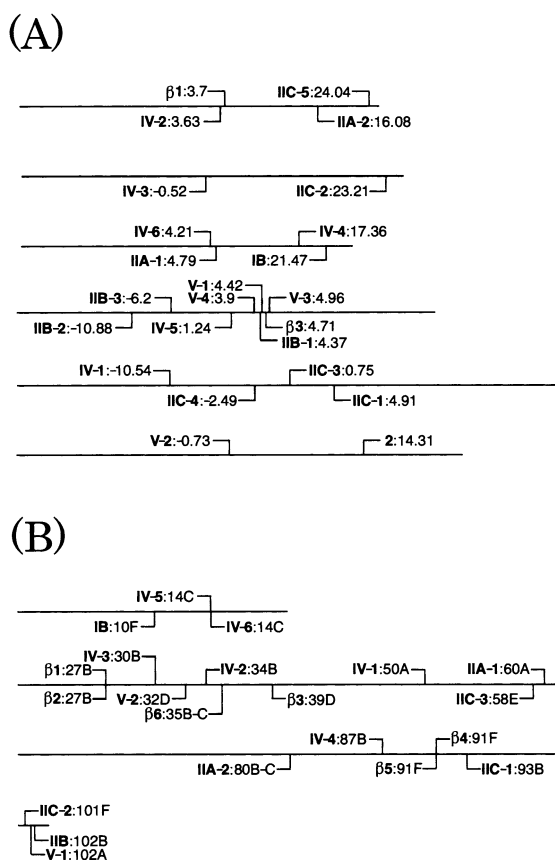


Fig. 5. Distribution of P-type ATPase genes on the *C. elegans* (A) and *Drosophila* (B) chromosomes. Note that the numbers represent centi-Morgan (for *C. elegans*) or polytene chromosome band-number (for *Drosophila*). Relative lengths of the chromosomes shown here reflect the actual sizes of the chromosomes among each organism but not between *C. elegans* and *Drosophila*.

tained this original nature. A similar characteristic is also seen in the *T. thermophila* TPA2.

ORIGIN OF H,K-ATPASES

The rooted tree in Fig. 2A indicates that the H,K-ATPases, which are seen in mammalian genomes, form a single cluster supported by high bootstrap value in Clade Y, and that they do not relate to any invertebrate isoforms. In addition, in the H,K-ATPase cluster, gastric and non-gastric H,K-ATPases separate from each other. These data indicate that the H,K-ATPase evolved only in the vertebrate lineage independently of invertebrates, and that the first evolutionary event was the separation of the ouabain-insensitive type (gastric H,K) from the ouabain-sensitive type (non-gastric H,K). Subsequently, the ouabain-insensitive type acquired a sensitivity to SCH28080 (which correlated with the presence of Glu⁸⁸¹ in the rabbit gastric H,K-ATPase α -subunit [44] (Fig. 1)). In fact, ouabain-sensitive H,K-ATPases are not inhibited by SCH28080.

ORIGIN OF α/β -SUBUNIT ASSEMBLY

At least under conditions of high expression or expression in heterologous species, any β -subunit isoforms can assemble with any α -subunit isoforms [22, 31, 38, 47, 53]. Therefore, common mechanisms for P-type ATPase subunit assembly must exist. A direct interaction essential for α/β -subunit assembly has been found to occur between extracellular elements of the α - and β -subunits [14]. The critical domain of the Na,K-ATPase α -subunit required for such interaction resides in a short amino-acid stretch localized in the extracellular loop between M7 and M8 of the α -subunit [Fig. 1; see also 30, 31]. The corresponding region of the H,K-ATPase α -subunit has likewise been shown to participate in subunit assembly [48]. Yeast two-hybrid experiments have further defined the crucial region of the α -subunit for the assembly, particularly the residues Ser-Tyr-Gly-Gln/Glu within the M7-M8 loop [9]. These four amino acids can be found in all Na,K- and H,K-ATPase α -subunits that have been well characterized in terms of the association with a β -subunit and in terms of their ion selectivities.

On the other hand, the Ser-Tyr-Gly-Gln/Glu motif is lacking in the potential Na,K- (H,K-) ATPases of *C. elegans* (Ce2C3, 2C4, and 2C5) and *Tetrahymena* (TPA2) (Fig. 1). These isoforms probably do not assemble with β -subunits. On the other hand, *Drosophila* Dm2C2 and Dm2C3 have one substitution in this assembly motif (Fig. 1), and might assemble with β -subunits. The *Hydra* Na/K-ATPase has two substitutions in this domain and whether or not this pump assembles with the β -subunit is not known. However, the *Hydra* Na/K-ATPase is known to bind ouabain [5].

Several lines of evidence indicate that the proper functional expression of vertebrate Na,K-ATPase requires the assembly of the α - and β -subunits [8, 14, 16, 33]. The assembly is necessary for (i) stabilization of the enzyme in the ER membrane [6, 17], (ii) acquisition of inhibitor- (ouabain-) binding ability [17], (iii) transportation of the enzyme to the plasma membrane [7, 18, 37, 39], and (iv) ion-pumping activity [58]. These functions are possibly associated with the specific conformation that is achieved upon α/β -subunit interaction. In this regard, the newly identified invertebrate subunit isoforms (*C. elegans* α -subunit isoforms Ce2C3, 2C4, and 2C5, and *D. melanogaster* β -subunit isoforms Dm β -4-6) are particularly interesting, because they lack features that are thought to be critical for subunit assembly. Their evolutionary rates are much faster than those of the isoforms that have the features associated with subunit assembly (Fig. 2A and B).

We propose that the *C. elegans* α -subunit isoforms Ce2C3, 2C4, and 2C5 never acquired the assembly domain in their evolutionary lineage and are

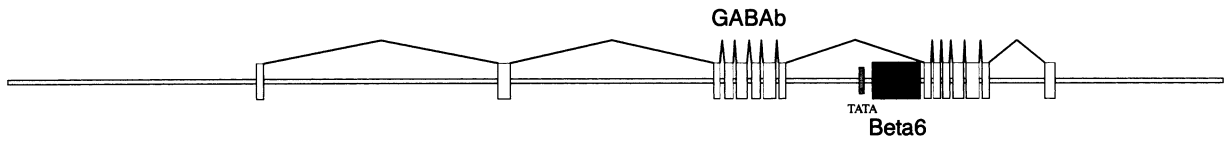


Fig. 6. The schematic diagram of the *Drosophila* β -6 and GABAAb receptor gene structures. The GABAAb receptor gene consists of 15 exons (white boxes), and Dm β -6 has only one exon (large black box) located between the 8th and 9th exons of the GABAAb receptor gene. The small black box indicates the putative TATA box for Dm β -6.

evolving without the evolutionary constraints associated with an obligatory protein-protein interaction. On the other hand, the three *D. melanogaster* β -isoforms (Dm β -4–6) appear to be derived from assembly-competent β -subunit isoforms (Dm β -1–3) (Fig. 2B) and appear to have lost their ability to assemble with the α -subunit during their evolution. Therefore, it is concluded that the loss of assembly competence in these β -isoforms resulted in a relaxation of functional constraints that was favorable for a higher rate of evolution.

Although the mutation rate of one of the *C. elegans* β -isoforms (Ce β -2) is also higher (Fig. 2B), this protein has three disulfide bonds. In the amino-acid sequence comparison between the Ce β -2 and the other invertebrate β -isoforms, the amino-terminal region including the transmembrane segment shows more sequence divergence than the extracellular region; the amino acid-sequence similarities of the amino-terminal and of the extracellular region are 10.2 vs 20.4 and 18.9 vs 32.0, respectively. Amino-acid residues critical for participation of the Ce β -2 subunit in Na,K-ATPase function might have been substituted, the relaxation of functional constraints resulting in a higher rate of evolution for the Ce β -2 lineage.

Conclusion

Figure 7 shows plausible pathways for the evolution of the Na,K/H,K-ATPase α -subunits, including the points of appearance or disappearance of sites for ouabain binding and assembly with β -subunits. If the ancestral form already possessed the ouabain-binding and subunit-assembly sites (Model 1), then during subsequent evolution, the sites in the ancestor were lost either in a single step or in multiple steps. If we choose Model 1, our analysis (Fig. 2) supports a multi-step loss of these sites. On the other hand, if the sites arose during evolution from the ancestral form, as in Model 2, the scenario we favor, the results in Fig. 2 are consistent with a rapid, virtually simultaneous acquisition of both ouabain sensitivity and subunit assembly. We favor Model 2 because primitive organisms (e.g., *Tetrahymena*) appear to have Na,K-ATPases that lack both a high-affinity ouabain-binding site and a consensus site for α/β -subunit assembly.

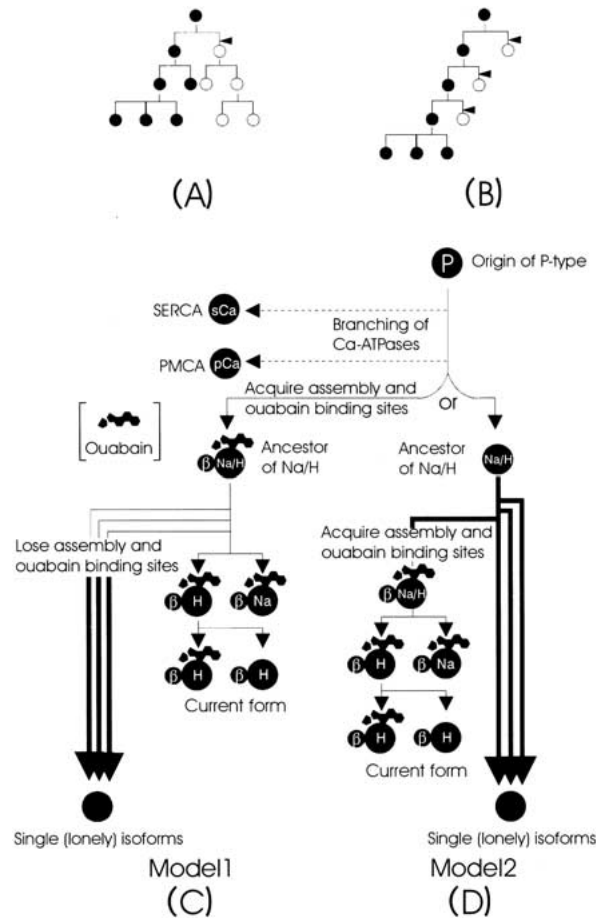


Fig. 7. Models of the Na,K-(H,K)-ATPase α -subunit evolution in terms of the sites for ouabain-binding and subunit assembly. The models deal with the possibilities of either acquiring or losing these sites. With the assumption that the ancestral form already possessed the sites, our analysis (Fig. 2) supports the idea that these sites might have been lost in multiple steps (B), but not in a single step (A). In this case, some of the isoforms in *Tetrahymena* and nematoda must have lost these sites step by step in the evolution of their ancestral forms (C; Model 1). The loss of assembly sites would result in a relaxation of functional constraints and result in an increase in the evolutionary rate (*thick arrows*). On the other hand, with the assumption that the sites were acquired early in evolution, the results in Fig. 2 suggest a virtually simultaneous acquisition (D; Model 2). The resultant subunit assembly posed functional constraints on the pump molecule and led to a decrease in the evolutionary rate (*thin arrows*). Model 2 is more plausible than Model 1 because of its simplicity.

In summary, we propose the following evolutionary sequence. First, an ancestral member of the Na,K and H,K-ATPase family lacked the inhibitor-binding and the subunit-assembly domains, and it existed as a single-subunit P-type ATPase in Protists and in lower invertebrate organisms. Second, in its evolution, this ancestral form acquired the abilities to bind ouabain and to assemble with a β -subunit, becoming an immediate ancestor of the currently well-known Na,K- and H,K-ATPase family (Clade Y in Fig. 2A). Third, from this family, the Na,K- and the H,K-ATPases diverged and became established in vertebrate organisms. Fourth, after the establishment of the H,K-ATPase subfamily, some of the members lost their ouabain-binding ability. In addition to these evolutionary steps, gene duplication of both α - and β -subunit genes within various animal phyla has independently led to the occurrence of α - and β -subunit isoform diversity in these phyla. Also, during the evolution of the β -subunit, at least in Arthropods, some of the isoforms have lost disulfide bonds thought to be necessary for proper assembly with α -subunits. Such forms may play some function other than in ATP-driven ion transport.

We thank Dr. Jonathan Pevsner (Kennedy Krieger Institute, The Johns Hopkins University) for extensive comments on the manuscript. We also thank Drs. K. Ozaki and A.K. Satoh for providing the cDNA library, and Dr. Futai for *C. elegans* stock. This work was supported by Grant-in-Aid for International Scientific Joint Research and Grant-in-Aid for Scientific Research (B) from Mombusho (Japanese Ministry of Education, Science, Sports and Culture) and grant NS-23241 from the N.I.H.

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410
- Asano, S., Kawada, K., Kimura, T., Grishin, A.V., Caplan, M.J., Takeguchi, N. 2000. The roles of carbohydrate chains of the β -subunit on the functional expression of Gastric H^+ , K^+ -ATPase. *J. Biol. Chem.* **275**:8324–8330
- Axelsen, K.B., Palmgren, M.G. 1998. Evolution of substrate specificities in the P-type superfamily. *J. Mol. Evol.* **46**:84–101
- The *C. elegans* Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**:2012–2018
- Canfield, V.A., Xu, K.-Y., D'Aquila, T., Shyjan, A.W., Levenson, R. 1992. Molecular cloning and characterization of Na,K-ATPase from *Hydra vulgaris*: Implications for enzyme evolution and ouabain sensitivity. *The New Biologist* **4**:339–348
- Caplan, M.J., Forbush III, B., Palade, G.E., Jamieson, J.D. 1990. Biosynthesis of the Na,K-ATPase in Madin-darby canine kidney cells. *J. Biol. Chem.* **265**:3528–3534
- Caplan, M.J., Palade, G.E., Jamieson, J.D. 1986. Newly synthesized Na, K-ATPase α subunit has no cytosolic intermediate in MDCK cells. *J. Biol. Chem.* **261**:2860–2865
- Cho, D.C., Forte, J.G. 1995. Functional significance of the β -subunit for heterodimeric P-type ATPases. *J. Exp. Biol.* **198**:1–17
- Colonna, T.E., Huynh, L., Fambrough, D.M. 1997. Subunit interaction in the Na,K-ATPase explored with the yeast two-hybrid system. *J. Biol. Chem.* **272**:12366–12372
- Croyle, M.L., Woo, A.L., Lingrel, J.B. 1997. Extensive random mutagenesis analysis of the Na^+/K^+ -ATPase α subunit identifies known and previously unidentified amino acid residues that alter ouabain sensitivity: Implications for ouabain binding. *Eur. J. Biochem.* **248**:488–495
- Davis, M.W., Somerville, D., Lee, R.Y., Lockery, S., Avery, L., Fambrough, D.M. 1995. Mutations in the *Caenorhabditis elegans* Na,K-ATPase alpha-subunit gene, eat-6, disrupt excitable cell function. *J. Neurosci.* **15**:8408–8418
- The *D. melanogaster* Sequencing Consortium 2000. Genome sequence of the insect *D. melanogaster*: A platform for investigating biology. *Science* **282**:2012–2018
- Fagan, M.J., Saier, M.H. 1994. P-type ATPases of eukaryotes and bacteria: Sequence analyses and construction of phylogenetic trees. *J. Mol. Evol.* **38**:57–99
- Fambrough, D.M., Lemas, M.V., Hamrick, M., Emerick, M., Renaud, K.J., Inman, E.M., Hwang, B., Takeyasu, K. 1994. Analysis of subunit assembly of the Na,K-ATPase. *Am. J. Physiol.* **266**:C579–C589
- Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package), version 3.57c. Distributed by the author: <http://evolution.genetics.washington.edu/phylip.html>. Department of Genetics, University of Washington, Seattle
- Geering, K. 1990. Subunit assembly and functional maturation of Na,K-ATPase. *J. Membrane Biol.* **115**:109–121
- Geering, K., Kraehenbuhl, J.P., Rossier, B.C. 1987. Maturation of the catalytic alpha-subunit of Na,K-ATPase during intracellular transport. *J. Cell Biol.* **105**:2613–2619
- Geering, K., Meyer, D.I., Paccolat, M.P., Kraehenbuhl, J.P., Rossier, B.C. 1985. Membrane insertion of α - and β -subunits of Na^+ , K^+ -ATPase. *J. Biol. Chem.* **260**:5154–5160
- Gloor, S., Antonicek, H., Sweadner, K.J., Pagliusi, S., Frank, R., Moose, M., Schachner, M. 1990. The adhesion molecule on glia (AMOG) is a homologue of the beta subunit of the Na,K-ATPase. *J. Cell Biol.* **110**:165–174
- Green, N.M. 1992. Evolutionary relationships within the family of P-type cation pumps. *Ann. N.Y. Acad. Sci.* **671**:104–112
- Hamrick, M., Renaud, K.J., Fambrough, D.M. 1993. Assembly of the extracellular domain of the Na,K-ATPase beta subunit with the alpha subunit: Analysis of beta subunit chimeras and carboxyl-terminal deletions. *J. Biol. Chem.* **268**:24367–24373
- Horisberger, J.-D., Jaunin, P., Reuben, M.A., Lasater, L.S., Chow, D.C., Forte, J.G., Sachs, G., Rossier, B.C., Geering, K. 1991. The H,K-ATPase β -subunit can act as a surrogate for the β -subunit of Na,K-pumps. *J. Biol. Chem.* **266**:19131–19134
- Ishii, T., Takeyasu, K. 1993. The amino-terminal 200 amino acids of the plasma membrane Na,K-ATPase α -subunit confer ouabain-sensitivity on the sarcoplasmic reticulum Ca-ATPase. *Proc. Natl. Acad. Sci. USA* **90**:8881–8885
- Iwabe, N., Kuma, K.I., Miyata, T. 1996. Evolution of gene families and relationship with organismal evolution: Rapid divergence of tissue-specific genes in the early evolution of chordates. *J. Mol. Evol.* **13**:483–493
- Jørgensen, P.L., Anderson, J.P. 1988. Structural basis for E1-E2 conformational transitions in $Na^+ + K^+$ -pump and Ca^{2+} -pump proteins. *J. Membrane Biol.* **103**:95–120
- Kawamura, M., Noguchi, S., Ueno, S., Kusaba, M., Takeda, K. 1994. Functional analysis of the disulfide bonds of Na^+/K^+ -ATPase β -subunit by site-directed mutagenesis. In: *The Sodium Pump*. E. Bamberg and W. Schoner, ed. pp. 287–296. Darmstadt, Germany
- Kimura, M. 1983. *The neutral theory of molecular evolution*. Cambridge University Press, Cambridge

- Kirley, T.L. 1989. Determination of three disulfide bonds and one free sulfhydryl in the β subunit of (Na,K)-ATPase. *J. Biol. Chem.* **264**:7185–7192
- Kraev, A., Kraev, N., Carafoli, E. 1999. Identification and functional expression of the plasma membrane calcium ATPase gene family from *Caenorhabditis elegans*. *J. Biol. Chem.* **274**:4254–4258
- Lemas, M.V., Hamrick, M., Takeyasu, K., Fambrough, D.M. 1994a. 26 amino acids of an extracellular domain of the Na,K-ATPase α -subunit are sufficient for assembly with the Na,K-ATPase β -subunit. *J. Biol. Chem.* **269**:8255–8259
- Lemas, M.V., Yu, H.-Y., Takeyasu, K., Kone, B., Fambrough, D.M. 1994b. Assembly of the avian Na,K-ATPase α -isoforms with the avian Na,K-ATPase β -isoforms and the avian H,K-ATPase β -subunit. *J. Biol. Chem.* **269**:18651–18655
- MacLennan, D.H., Clarke, D.M., Loo, T.W., Skerjanc, I.S. 1992. Site-directed mutagenesis of the Ca^{2+} ATPase of sarcoplasmic reticulum. *Acta Physiol. Scand. Suppl.* **607**:141–150
- Magyar, A., Varadi, A. 1990. Molecular cloning and chromosomal localization of a sarco/endoplasmic reticulum-type Ca^{2+} -ATPase of *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun.* **173**:872–877
- McDonough, A.A., Geering, K., Farley, R.A. 1990. The sodium pump needs its β subunit. *FASEB J.* **4**:1598–1605
- Miller, R.P., Farley, R.A. 1990. β subunit of ($\text{Na}^{+} + \text{K}^{+}$)-ATPase contains three disulfide bonds. *Biochemistry* **29**:1524–1532
- Møller, J.V., Juul, B., le Maire M. 1996. Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim. Biophys. Acta.* **1286**:1–51
- Noguchi, S., Higashi, K., Kawamura, M. 1990. A possible role of the β subunit of (Na,K)-ATPase in facilitating correct assembly of the α subunit into the membrane. *J. Biol. Chem.* **265**:15991–15995
- Noguchi, S., Maeda, M., Futai, M., Kawamura, M. 1992. Assembly of a hybrid from the α subunit of Na,K-ATPase and the β subunit of H,K-ATPase. *Biochem. Biophys. Res. Commun.* **182**:659–666
- Noguchi, S., Mishina, M., Kawamura, M., Numa, S. 1987. Expression of functional ($\text{Na}^{+} + \text{K}^{+}$)-ATPase from cloned cDNAs. *FEBS Lett.* **225**:27–32
- Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**:357–358
- Palmgren, M.G., Axelsen, K.B. 1998. Evolution of P-type ATPases. *Biochim. Biophys. Acta* **1365**:37–45
- Pestova, N.B., Adamsa, G., Shakhparonov, M.I., Modyanov, N.N. 1999. Identification of a novel gene of the X,K-ATPase β -subunit family that is predominantly expressed in skeletal and heart muscles. *FEBS Lett.* **456**:243–248
- Price, E.M., Lingrel, J.B. 1988. Structure-function relationships in the Na,K-ATPase α subunit: site-directed mutagenesis of glutamine-111 to arginine and asparagine-122 to asparatic acid generates a ouabain-resistant enzyme. *Biochemistry* **27**:8400–8408
- Sachs, G., Shin, J.M., Briving, C., Wallmark, B., Hersey, S. 1995. The pharmacology of the gastric acid pump: The H^{+} , K^{+} ATPase. *Annu. Rev. Pharmacol. Toxicol.* **35**:277–305
- Saitou, N., Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425
- Sambongi, Y., Wakabayashi, T., Yoshimizu, T., Omote, H., Oka, T., Futai, M. 1997. *Caenorhabditis elegans* cDNA for a Menkes/Wilson disease gene homologue and its function in a yeast CCC2 gene deletion mutant. *J. Biochem.* **121**:1169–1175
- Schmalzing, G., Gloor, S., Omay, H., Kroner, S., Appelhans, H., Schwarz, W. 1991. Up-regulation of sodium pump activity in *Xenopus laevis* oocytes by expression of heterologous β 1 subunits of the sodium pump. *Biochem. J.* **279**:329–336
- Shin, J.M., Sachs, G. 1994. Identification of a region of the H,K-ATPase alpha subunit associated with the beta subunit. *J. Biol. Chem.* **269**:8642–8646
- Solovyev, V.V. 1997. FGENES program. The URL of the Web page: <http://genomic.sanger.ac.uk/gf/gf.shtml>. Computational Genomics Group, The Sanger Centre, Cambridge
- Sun, B., Salvaterra, P.M. 1995. Two *Drosophila* nervous system antigens, Nervana 1 and 2, are homologous to the β subunit of Na^{+} , K^{+} -ATPase. *Proc. Natl. Acad. Sci. USA* **92**:5396–5400
- Sun, B., Wang, W., Salvaterra, P.M. 1998. Functional analysis and tissue-specific expression of *Drosophila* Na^{+} , K^{+} -ATPase subunits. *J. Neurochem.* **71**:142–151
- Takeyasu, K., Lemas, V., Fambrough, D.M. 1990. Stability of Na^{+} , K^{+} -ATPase α -subunit isoforms in evolution. *Am. J. Physiol.* **259**:C619–C630
- Takeyasu, K., Tamkun, M.M., Siegel, N.R., Fambrough, D.M. 1987. Expression of hybrid ($\text{Na}^{+} + \text{K}^{+}$)-ATPase molecules after transfection of mouse Ltk⁻ cells with DNA encoding the β -subunit of an avian brain sodium pump. *J. Biol. Chem.* **262**:10733–10740
- Tamkun, M.M., Fambrough, D.M. 1986. The ($\text{Na}^{+} + \text{K}^{+}$)-ATPase of chick sensory neurons: Studies on biosynthesis and intracellular transport. *J. Biol. Chem.* **261**:1009–1019
- Thompson, J.D., Higgins, D.G., Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680
- Wang, S., Takeyasu, K. 1997. Primary structure and evolution of the ATP-binding domains of the P-type ATPases in *Tetrahymena thermophila*. *Am. J. Physiol.* **272**:C715–C728
- Yu, H., Ishii, T., Pearson, W.R., Takeyasu, K. 1994. Primary structure of H^{+} / K^{+} -ATPase β -subunit. *Biochim. Biophys. Acta* **1190**:189–192
- Zhao, J., Vasilets, L.A., Yoshimura, S.H., Gu, Q., Ishii, T., Takeyasu, K., Schwarz, W. 1997. The Ca^{2+} /calmodulin binding domain of the Ca^{2+} -ATPase linked to the Na^{+} , K^{+} -ATPase alters transport stoichiometry. *FEBS Lett.* **408**:271–275