The complete inventory of the yeast *Saccharomyces cerevisiae* P-type transport ATPases

Patrice Catty, Alban de Kerchove d'Exaerde, André Goffeau *

Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix du Sud 2-20, B-1348 Louvain-La-Neuve, Belgium

Received 2 April 1997

Abstract A total of sixteen open reading frames encoding for P-type ATPases have been identified in the complete genome sequence of *Saccharomyces cerevisiae*. Phylogenetic analysis distinguishes 6 distinct families. Topology predictions, identification of aminoacid sequence motifs and phenotype analysis of the available mutants suggest that these families correspond to ATPases transporting either H^+ (2 members), Ca^{2+} (2 members), Na^+ (3 members), heavy metals (2 members), possibly aminophospholipids (5 members including 4 new ones) or unknown substrates (2 new members). It is proposed that the latter family which has homologs in *Tetrahymena thermophila*, *Plasmodium falciparum* and *Caenorhabditis elegans* constitutes a new group called P4-ATPases with characteristic topology and aminoacid signatures.

1. Introduction

In all living cells, a variety of P-type ATPases transport cations against their electrochemical gradient at the energetic expense of ATP hydrolysis [1].

All the P-type ATPases form an aspartyl-phosphate as a catalytic intermediate during ATP hydrolysis. The analysis of their primary sequence has led to the identification of 5 highly conserved sequences located in the 2 major hydrophilic segments of these proteins [2–4]. These signatures comprise the LTGES motif located in the smaller hydrophilic segment of unidentified function and 4 other motifs located in the major hydrophilic ATP-binding segment: the phosphorylation site DKTGTLT, the KGA and MLTGD motifs directly involved in ATP binding and the GDGxND sequence in the hinge which connects the ATP-binding domain to a transmembrane segment involved in ion translocation. Recently 3 groups of P-type ATPases have been distinguished [5,6]. These groups differ by their predicted topology. They bear distinct sequence signatures and transport different ions. The P1-ATPases [5] (or CPx-ATPases [6]) are involved in the transport of heavy metals such as cadmium, copper and mercury. The P2-ATPases involved in the transport of other cations such as proton, calcium, sodium and potassium are ubiquitous whereas *Kdp*, the unique (so far) P3-ATPase, is part of a complex structure involved in potassium uptake by *Escherichia coli*.

The P-type ATPase signatures have enabled us to index the 16 P-type ATPase open reading frames (ORFs) (listed in Table 1) encoded by the now completely decrypted *Saccharomyces cerevisiae* genome [7]. Among them, 8 ORFs correspond to new P-type ATPases which were unravelled by systematic genome sequencing and had not been previously characterized by specific mutant phenotypes.

In this paper, these 16 yeast P-type ATPases have been clustered in 6 different families, using the phylogenetic relations determined by the PHYLIP program (Fig. 1). The prediction of the transmembrane topology of these proteins given in Fig. 2 supports the grouping into the 6 families identified by the phylogenetic approach. On the basis of predicted topology features and of characteristic sequencing motifs, we propose that one of these families represents a new group, the P4-ATPases.

2. The P1-ATPases

The 2 *S. cerevisiae* ORFs Ydr270wp and Ybr295wp have the structural features which allow to classify them unambiguously in the P1-heavy-metal transporting ATPases [5,6]. They have 4 predicted transmembrane spans on the N-terminal side (Fig. 2) of the (I/M)TGES motif. As in all P1-ATPases a proline residue located in the 4th transmembrane span about 43 residues upstream the phosphorylation site is flanked on both sides by cysteine residues (Table 2). The SEHPL motif with invariant H and P located about 40 residues downstream the phosphorylation site was reported to be also characteristic of P1-ATPases [5,6]. It is represented as SDHPV and IKHPV in the yeast Ydr270wp and Ybr295wp respectively. Both yeast proteins display another essential characteristic of heavy metals pumps which is the aminoterminal GMTCxxC sequence (2 copies in Ccc2p, 1 in Pca1p, see Table 3), proposed to be involved in metal binding (review in [6]).

The 2 *S. cerevisiae* P1-ATPases are encoded on one hand by the *CCC2* gene, found as suppressor of the calcium sensitivity of *csg1* mutants [8], and on the other hand by the *PCA1* gene, discovered during the systematic genome sequencing [9]. Significantly related to each other (BLAST P-value of 1.9×10^{-83}), the 2 yeast ORFs, Ydr270wp and Ybr295wp are similar to the human copper ATPases, ATP7Ap and ATP7Bp, involved in the Menkes and the Wilson diseases respectively (BLAST P-values lower than 10^{-155} and 10^{-86} for Ccc2p and Pca1p respectively).

CCC2 and *PCA1* are not essential genes and their disruptions lead to growth defects attributed to an impaired respiration [9,10]. In addition, the *ccc2* null mutant exhibits an increased sensitivity to copper depletion [8], while the *pca1* null mutant shows only a slightly decreased cell density at the stationary phase in 2 mM $CuSO_4$ [9]. Although the respiratory defect of the *ccc2* null mutant is suppressed by high concentrations of iron or copper, the mutant does not exhibit a decreased cytosolic concentration of copper. It has been suggested that Ccc2p mediates copper import into an intracellular compartment containing the copper-dependent ferro-

*Corresponding author. Fax: (32) (10) 473872.
E-mail: goffeau@fysa.ucl.ac.be

PII S0014-5793(97)00446-8

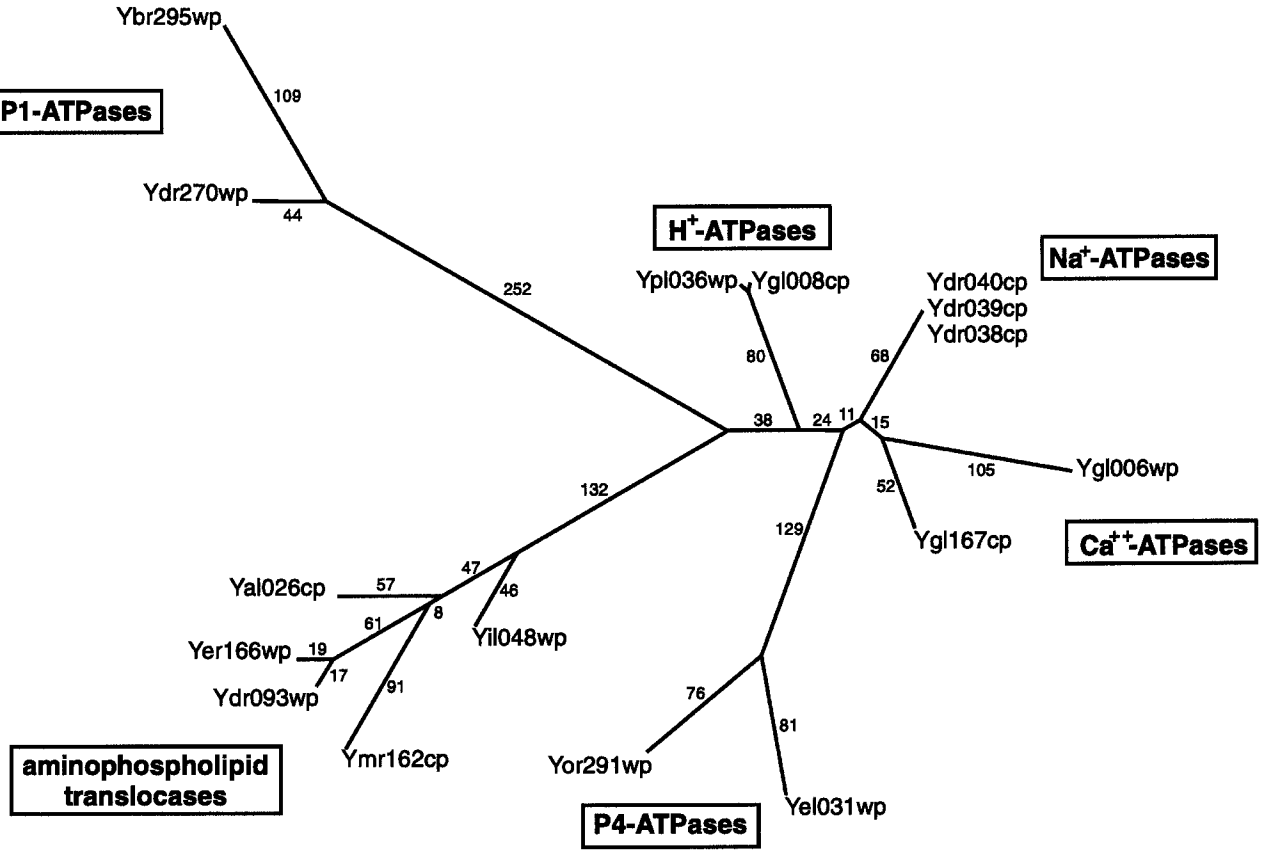

Fig. 1. Phylogenetic tree of the 16 *S. cereviseae* P-type ATPases (strain 288C). Building of the phylogenetic tree was made using the PILEUP algorithm (GCG package, version 8.0) and the PHYLIP program (Felsenstein: E-mail: joe@genetics.washington.edu). Relative evolution distances are indicated to the branches.

oxidase Fet3p, known to be required for iron uptake [10]. The human ATPase ATP7Bp and the yeast ATPase Ccc2p probably ensure comparable physiological functions [11]. Indeed, in the Wilson disease, a loss of ceruloplasmin ferro-oxidase activity linked to copper metabolism disorders, is associated with iron mobilization deficiency.

3. The P2-ATPases

3.1. The H^+-ATPases

The plasma membrane *PMA1* H^+-ATPase gene (*YGL008C*) was the first sequenced yeast P-type ATPase encoding gene [12]. The well-characterized gene product (review

Table 1
The P-type ATPases of *Saccharomyces cerevisiae*

Presumed function	ORF name	Protein name	Length	CAI	TMS	P-type ATPase consensus motifs				
						1	2	3	4	5
Cu^{++}-ATPase	*YDR270w*	Ccc2p	1004	0.135	8	MTGES	FDKTGTLT	KGIVS	MITGD	GDGIND
	YBR295w	Pca1p	1216	0.146	8	ITGES	FDKTGTLT	KGVSA	ILSGD	GDGTND
H^+-ATPase	*YGL008c*	Pma1p	918	0.734	10	ITGES	SDKTGTLT	KGAPL	MLTGD	GDGVND
	YPL036w	Pma2p	947	0.300	10	ITGES	SDKTGTLT	KGAPL	MLTGD	GDGVND
Ca^{++}-ATPase	*YGL006w*	Pmc1p	1173	0.159	10	ITGES	SDKTGTLT	KGAAE	MVTGD	GDGTND
	YGL167c	Pmr1p	950	0.196	10	LTGEN	SDKTGTLT	KGAFE	MITGD	GDGVND
APL-Translocase	*YAL026C*	Drs2P	1355	0.199	10	LDGET	SDKTGTLT	KGADT	VLTGD	ASGAND
	YER166w		1571	0.163	10	LDGET	SDKTGTLT	KGADS	VLTGD	GDGSND
	YIL048w		1151	0.164	10	LDGET	SDKTGTLT	KGADT	MLTGD	GDGGND
	YMR162c		1656	0.140	10	LDGET	SDKTGTLT	KGADN	MLTGD	GDGAND
	YDR093w		1612	0.160	10	LDGET	SDKTGTLT	KGADS	VLTGD	GDGSND
Na^+-ATPases	*YDR040c*	Pmr2ap	1091	0.197	8	LDGES	SDKTGTLT	KGAFE	MLTGD	GDGVND
	YDR039c	Pmr2bp	1091	0.190	8	LDGES	SDKTGTLT	KGAFE	MLTGD	GDGVND
	YDR038c	Ena5p	1091	0.190	8	LDGES	SDKTGTLT	KGAFE	MLTGD	GDGVND
Unknown	*YOR291w*		1472	0.135	10	LTGGS	FDKTGTLT	KGAPE	MCTGD	GDGAND
	YEL031w		1215	0.257	8	LTGES	FDKTGTLT	KGAPE	MITGD	GDGTND

Length is given in number of aminoacids. CAI: codon adaptation index. TMS: number of putative transmembrane segments (from PredictProtein server at EMBL). APL: aminophospholipid. All open reading frames (ORFs) have been identified from yeast genome database of MIPS (Martinreid Institute for Protein Sequences http://www.mip.biochem.mpg.de/yeast) as containing 5 highly conserved P-type motifs.

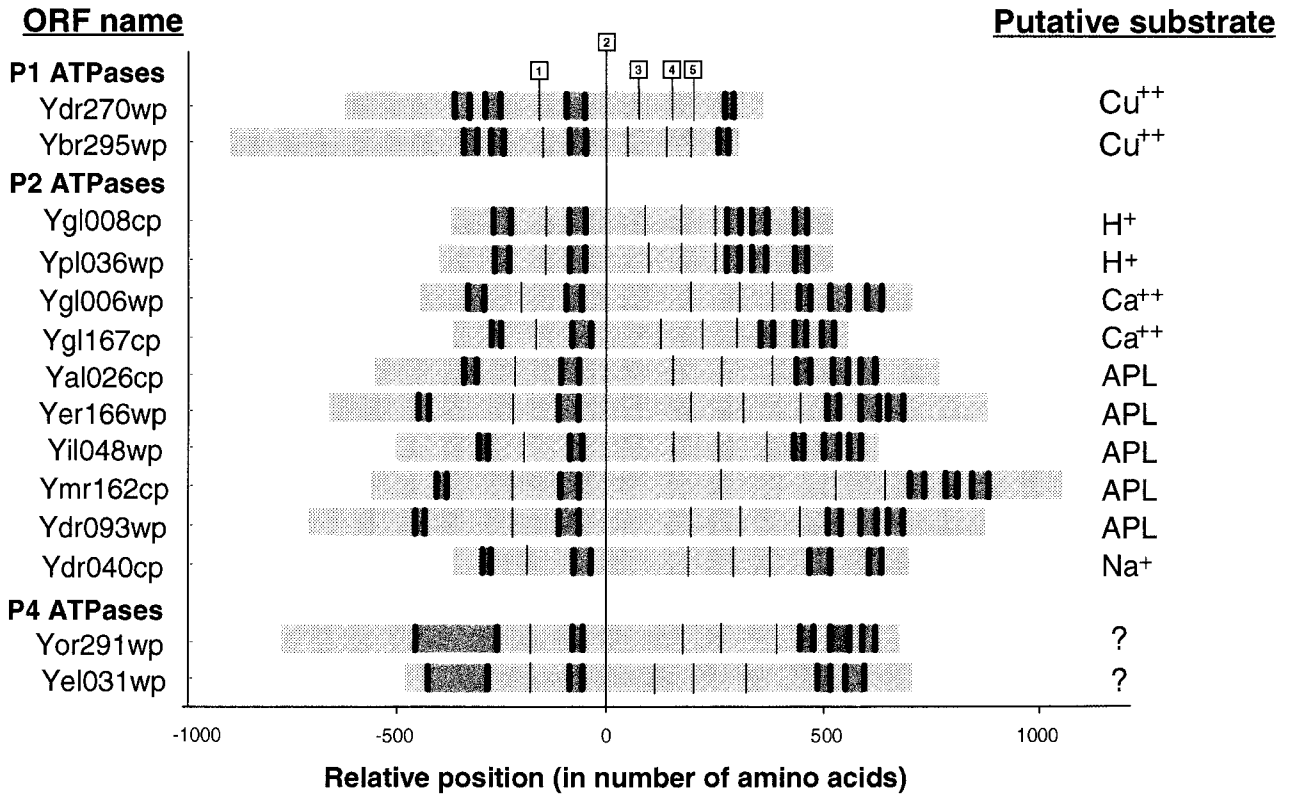

Fig. 2. Predicted topology of the 14 different yeast P-type ATPases. The PredictProtein program (refined prediction of transmembrane helices and topology, PHDtopology, [56,57] was used). Transmembrane spans are indicated by thick black lines. The P-type consensus motifs shown in Table 1 are numbered 1–5 and indicated by thin black lines. Cytoplasmic segments are in light grey. Extracytoplasmic segments are in deep grey. The abscissa gives the aminoacid position relative to the phosphorylation motif (motif 2).

in [13]) is essential for cell growth because it generates the proton motive force driving the transport of metabolites across the plasma membrane.

A second H^+-ATPase gene, *PMA2*, (*YPL036W*) was found to code for a very lowly expressed protein, with minor role in cell growth [14]. While Pma2p can replace Pma1p when expressed from the strong *PMA1* promoter [15], its kinetic properties slightly differ from those of Pma1p [16].

Biochemical analysis of the native or of the mutated Pma1p (review in [17]) reveals global mechanistic similarities between the yeast plasma membrane proton pump and the mammalian sarcoplasmic reticulum Ca^{2+}-ATPase [18]. However, precise definition of the functional domains of the H^+-ATPase involved in ATP hydrolysis and in H^+ transport as well as the identification of the aminoacids that determine the ionic specificity of the pump are still missing. Full understanding of the molecular chemiosmotic mechanisms of Pma1p will require the elucidation of its three-dimensional crystalline structure, a challenge that its abundance and ease of purification allow to envisage now that two-dimensional crystals have been reported for the *Neurospora crassa* H^+-ATPase [19].

Fig. 2 shows that the predicted topology of the 2 H^+-ATPases is canonical for the P2-ATPases; short N and C termini, 1 pair of transmembrane spans before the ITGES motif, an-

Table 2
Specific proline-containing motifs

ORF name	Motif upstream the phosphorylation site	Motif downstream the phosphorylation site	Putative substrate
Ydr270wp	C**P**C (43)	SD**HP**V (42)	Cu^{2+}
Ybr295wp	C**P**C (43)	IK**HP**V (36)	Cu^{2+}
Yg1008cp	V**P**V (43)	–	H^+
Ypl036wp	V**P**V (43)	–	H^+
Yg1006wp	V**P**E (43)	–	Ca^{2+}
Yg1167cp	I**P**E (43)	–	Ca^{2+}
Ya026cp	V**P**I (53)	–	APL
Yer166wp	V**P**I (53)	–	APL
Yil048wp	I**P**V (47)	–	APL
Ymr162cp	I**P**L (53)	–	APL
Ydr093wp	V**P**I (53)	–	APL
Ydr040cp	I**P**S (43)	–	Na^+
Yor291wp	V**P**P (43)	–	unknown
Yel031wp	V**P**P (43)	–	unknown

The number in parentheses corresponds to the distance in aminoacids between the aspartic residue of the DKTGT sequence and the considered proline.

Table 3
Putative copper-binding domain

Consensus	•-GM•C--C---•--------$^{G}_{C}$•-----$^{S}_{N}$L
Pca1p (*S. cer.*)	VSGMSCTGCESKLKKSFGALKCVHGLKTSL
Ccc2p^{A} (*S. cer.*)	VHGMTCSACTNTINTQLRALKGVTKCDISL
Ccc2p^{B} (*S. cer.*)	VQGMTCGSCVSTVTKQVEGIEGVESVVVSL
WD^{A} (Human)	ILGMTCQSCVKSIEDRISNLKGIISMKVSL
MNK^{A} (Human)	VEGMTCNSCVWTIEQQIGKVNGVHHIKVSL
CopA (*E. hirae*)	ITGMTCANCSARIEKELNEQPGVMSATVNL

Bold: conserved aminoacids. Underlined: conservative replacement indicated by • in the consensus sequence. Letters in indice and exposant indicate 2 non-equivalent aminoacids at the same position. All sequences shown are located in the N-terminus of the following P-type ATPases. Pca1p is a yeast ATPase.Ccc2p^{A} and Ccc2p^{B} are the 1st and 2nd putative metal binding domains of the yeast Ccc2p ATPase. WD^{A} en MNK^{A} are the first of the 6 putative metal binding domains of the Wilson and Menkes ATPases respectively [46,47]. CopA, is a copper-ATPase from *Enterococcus hirae* [48].

other pair upstream the phosphorylation site and 3 pairs of transmembrane spans downstream the GDGxND motif. The transmembrane proline residue typically located 43 residues upstream the phosphorylation site of P-type ATPases [6] is flanked by valine residues on both sides in both members of this H^+-ATPase family (Table 2).

3.2. The Ca^{2+}-ATPases

This family contains 2 members Ygl006wp (Pmr1p) and Ygl167cp (Pmc1p) which present structural features similar to the mammalian sarcoplasmic reticulum calcium-ATPase such as 10 predicted transmembrane spans, and characteristic sequence similarities. These sequence similarities are most pronounced in the transmembrane spans 4 and 6 which contain a total of 3 residues (E, N, D) believed to be specifically involved in calcium binding (Table 4). The E residue located in transmembrane 4 makes with its flanking P, a typical PE doublet signature present in all known Ca^{2+}-ATpases (Tables 2 and 4).

The phenotypic properties of the yeast deletion mutants confirm the prediction that both yeast Pmr1p and Pmc1p are involved in calcium transport.

The *PMR1* gene [20] is identical to the *SSC1* gene, a mutation of which is associated to a supersecreting phenotype [21]. The *pmr1* null mutant shows multiple secretory defects: supersecretion of some heterologously produced proteins, partial outer chain glycosylation and impaired proteolytic processing [20,22,23]. In agreement with these phenotypes, Pmr1p was immunolocalized in Golgi, the major processing compartment along the secretory pathway [22,24]. Although ensuring calcium tolerance of the *vcx1 cnb1 pmc1* triple mutant (*VCX1*, *CNB1* and *PMC1* genes encode for Ca^{2+}/H^+ exchanger, calcineurin B subunit and vacuolar Ca^{2+}-ATPase respectively), Pmr1p probably contributes moderately to calcium homeostasis in normal growth conditions [25].

The *pmr1* null mutant exhibits altered growth sensitivities not only to calcium but also to manganese [26,27], 2 ions that induce the expression of PMR1 through a calmodulin/calcineurin-dependent process. Other data suggest that Pmr1p could play a role in tolerance to manganese demonstrated as a calcium substitute for several essential physiological functions in yeast [28]. Indeed, certain *pmr1* mutants are rescued through a manganese-dependent mechanism [27] and growth inhibition of *pmr1* null mutant in low calcium is suppressed by micromolar concentrations of manganese (S. Loukin, personal communication).

The *PMC1* gene, expression of which is regulated by calcium through a calmodulin/calcineurin-dependent mechanism [25], encodes for a vacuolar membrane protein that exhibits 40% homologies with the mammalian plasma membrane Ca^{2+}-ATPases (PMCA). However, Pmc1p does not contain the carboxyterminal autoinhibitory domain found in the PMCA pumps (review in [29]). The *pmc1* null mutant displays decreased calcium tolerance, associated to decreased sequestration of calcium [26,30]. Analysis of the phenotypes associated to various combinations of mutations in *PMC1*, *PMR1*, *VCX1* or *CNB1* shows that Pmc1p is the major determinant of calcium homeostasis in yeast cell [25,26,30].

3.3. The Drs2p ATPases

The *S. cerevisiae* genome contains 4 genes, *YMR162C*, *YDR093W*, *YER166W* and *YIL048W*, encoding for P-type ATPases related to the *YAL026C* gene encoding Drs2p. The Drs2p protein, shown to be required for proper ribosomes assembly, had been tentatively classified in the group of calcium translocating ATPases [31,32]. This possibility is not supported by sequence analysis of transmembrane spans 4 and 6 which do not contain the calcium-binding residues E, N or D and show limited sequence similarity to the same regions of the mammalian calcium ATPases from sarcoplasmic reticulum or plasma membranes (Table 4). Recently, it has been reported that Drs2p, as well as the bovine ATPase II, the ATPase 2 from *Plasmodium falciparum*, and the *Caenorhabditis elegans* CET24H7.5 protein, might be members of a new family of P-type pumps, involved in aminophospholipid translocation [33].

We have searched for aminoacid motifs specific for this group of ATPases, to which also belong 2 newly discovered *Schizosaccharomyces pombe* proteins. The 10 putative aminophospholipid translocases of this P-type ATPase family known today, exhibit several minor variations in consensual P-type motifs (Table 5). For example, the LDGET variant of the highly conserved LxGEx sequence is found only in the group of ATPases related to Drs2p. Also the (Y/H) and (F/L) residues upstream the SDKTGTLT phosphorylation site and the D residue of the KGAD nucleotide-binding sequence are particularly well conserved in the Drs2p family. The isoleucine in the conserved IGDGxND sequence, as well as the flanking MIQ triplet seem also specific to this class of ATPases. In addition, Drs2p and its homologs contain specifically conserved motifs such as the EGLRTL and the VxxCRxxPxQK sequences found in the PMCA and the Na^+/K^+ ATPases respectively. More strikingly, the large cytoplasmic segment of all the Drs2p-like ATPases contains two highly specific sequences: GxT(A/G)(I/V)ED(K/R)LQxxV and (A/S)xSPDExA(L/I)(V/I).

All the members of this family have 10 predicted transmembrane spans. They all contain the (V/I)P(IVL) motif in transmembrane span 4 (Table 2) where the invariant proline residue is located further upstream from the phosphorylated

Table 4
Comparison of the 4th and 6th transmembrane spans of 2 mammalian Ca^{++}-ATPases and 9 yeast P-type ATPases

	TM4	% similarity with SERCA1a	TM6	% similarity with SERCA1a
SERCA1a	FKIAVALAVAAIPEGLPAVIT	100	ALIPVQLLWVNLVTDGLPATA	100
PMCA1a	FIIGVTVLVVAVPEGLPLAVT	71.4	PLKAVQMLWVNLIMDTLASLA	66.7
Pmr1p	FQISVSLAVAAIPEGLPIIVT	90.5	NPLNAMILWINILMDGPPAQS	61.9
Pmc1p	FITSITVIVVAVPEGLPLAVT	71.4	VLTAVQLLWINLIMDTLAALA	66.7
Yal026cp	FLTFWILFSNLVPISLFVTVE	28.6	TMSFYNLFFTVWPPFVIGVFD	28.6
Yer166wp	FWVAVILYQSLVPISLYISVE	47.6	YMMFYNLAFTSLPVIFLGDIL	33.3
Yil048wp	ILRYLILFSTIIPVSLRVNLD	33.3	LMVGYATCYTMAPVFSLTLDH	19.1
Ymr162cp	IMSFIIMYNTVIPLSLYVTME	33.3	SLSMFNTLFTSLPVLCIGMFE	33.3
Ydr093wp	FWVAVILYQSLVPISLYISVE	47.6	YLTFYNLAFTSVPVILLAVLD	28.6
Yor291wp	ILRALDIITIVVPPALPATLT	47.6	FLYIDLLLIVPIAICMSWSKS	33.3
Yel031wp	ILDCILIITSVVPPELPMELT	42.9	ATVSGLLLSVCFLSISRGKPL	33.3

Shaded boxes contain aminoacids of SERCA1a and PMCA1a proposed to be directly involved in calcium binding [18,49].Bold: conserved aminoacids.Underlined: conservative replacement.TM: transmembrane domain.% similarity is the sum of % identity and % of conservative replacement.SERCA1a: fast twitch skeletal muscle sarcoplasmic reticulum Ca^{2+}-ATPase [50].PMCA1a: rat plasma membrane Ca^{2+}-ATPase [51].

aspartate (46 or 52 residues) than in all other P-type ATPases (43 residues). Future biochemical studies are needed to determine whether or not the substrates of these yeast transport ATPases are aminophospholipids.

3.4. The Na^+-ATPases

In *S. cerevisiae*, the sodium pumps are encoded by several identical genes (*YDR040C*, *YDR039C*, *YDR038C*) organized in tandem repeats. Their number varies from 1 to 5, depending upon the strain and the yeast specie [34]. Clustered on the right arm of chromosome IV, these genes are not essential for normal growth but their disruption increases growth sensitivity to sodium and lithium [34–37]. One gene predominantly accounts for cell adaptation to high sodium concentration or high media pH, the others being constitutively weakly expressed [38]. The major sodium pump, called *Pmr2ap* [20,34], or *Ena1p* [37], is located in the plasma membrane and contains a carboxyterminal domain probably involved in a positive calcium/calmodulin mediated posttranscriptional regulation [37].

Pmr2ap/Ena1p activity is also positively modulated at the transcriptional level in response to high salt exposure, through a calmodulin calcineurin-dependent process [25,38–40]. Two mechanisms have been proposed to repress the *PMR2/ENA1* gene expression [41,42], one of them involves the cAMP-dependent protein kinase, the other involves the PPZ serine/threonine phosphatases. Recently, Pmr2p/Ena1p was shown to confer sodium tolerance to a *S. pombe* mutant deficient in the putative Na^+/H^+ transporter Sod2p [43].

The Ydr040ep predicted topology is characterized by the possibility to comprise only 4 transmembrane spans in its C-terminal (Fig. 2). Since this gene has no homolog so far, it is not possible either to confirm this observation or to identify specific sequence signatures for this family.

4. The P4-ATPases

The two genes *YEL031W* and *YOR291W* encode for unusual ATPases, unrelated to the other families of P-type pumps on the basis of their primary sequences identity. Yel031wp exhibits high homology to the ATPase I (X65738 in Table 6) from *P. falciparum* (BLAST P-value of

Table 5
Drs2p family signatures

Variations in the consensual P-type sequences

Specific consensus	T--*D**T	$^{Y}_{H}$•$^{F}_{L}$********	***$^{D}_{G}$	I***_**$^{V}_{R}$-MI•
Yal026cp (*S. cer.*)	TANLDGET	YIFSDKTGTLT	KGAD	IASGANDVSMIQ
Ydr093wp (*S. cer.*)	TKNLDGET	YIFSDKTGTLT	KGAD	IGDGSNDVAMIQ
Yer166wp (*S. cer.*)	TKNLDGET	YIFSDKTGTLT	KGAD	IGDGSNDVAMIQ
Yil048wp (*S. cer.*)	TDQLDGET	YLLSDKTGTLT	KGAD	IGDGGNDVSMIQ
Ymr162cp (*S. cer.*)	TMALDGET	YIFSDKTGTLT	KGAD	IGDGANDVSMIQ
ATCX (*S. pombe*)	TKNLDGET	YIFSDKTGTLT	KGAD	IGDGANDVAMIQ
SPAC6C3 (*S. pombe*)	TDQLDGET	YVLTDKTGTLT	KGAD	IGDGGNDVGMIQ
ATPase II (Bovine)	TSNLDGET	YIFSDKTGTLT	KGAD	IGDGANDVSMIQ
ATPase 2 (*P. falc.*)	TSSLDGET	YIFSDKTGTLT	KGAG	IGDGANDRNMIN
CET24H7.5 (*C. eleg.*)	TCNLDGET	HVLSDKTGTLT	KGAD	IGDGANDVPMIQ

Drs2p family specific motifs

Specific consensus	$^{E}_{K}$GLRTL	G-T$^{A}_{G}$•E*•LQ--V	•-SPDE-A••	V•--R-•P-$^{Q}_{E}$*
Yal026cp (*S. cer.*)	EGLRTL	GATAIEDKLQDGV	AASPDEGALV	VICCRVSPLQK
Ydr093wp (*S. cer.*)	EGLRTL	GGTAIEDRLQDGV	AQSPDESALV	VLCCRVSPAQK
Yer166wp (*S. cer.*)	EGLRTL	GGTAIEDRLQDGV	AQSPDEAALV	VLCCRVSPSQK
Yil048wp (*S. cer.*)	EGLRTL	GLTGVEDKLQKDV	AASPDEIAIV	VIACRCTPQQK
Ymr162cp (*S. cer.*)	EGLRTL	GVTAIEDKLQDGV	SSSPDELALV	VICCRASPSQK
ATCX (*S. pombe*)	EGLRTL	GGTAIEDRLQEGV	AQSPDEAALV	VLCCRVSPAQK
SPAC6C3 (*S. pombe*)	EGLRTL	GLTGVEDKLQKDV	AASPDEVAIV	VVICRCTPTQK
ATPase II (Bovine)	EGLRTL	GATAIEDKLQDQV	AASPDEGALV	VICCRVSPLQK
ATPase 2 (*P. falc.*)	EGLRTL	GITGIEDKLQEGV	SSSPDEEALV	VICGRVSPYQK
CET24H7.5 (*C. eleg.*)	KGLRTL	GVTGIEDRLQDGV	AESPDELALI	VLCYRMTPSEK

The different motifs were obtained using Block Maker Server [52]. Highly conserved aminoacids in P-type ATPases sequences are indicated by * in the consensus. Underlined: conservative replacement, indicated by • in the consensus. Letters in indices and exposant indicate two non-equivalent aminoacids at the same position. Bold: conserved aminoacids in the Drs2p family specific motifs. ATCX and SPAC6C3: sequences obtained from the systematic sequencing of the *Schizosaccharomyces pombe* genome. Bovine ATPases II [33]. ATPase 2 (or Isotype=P) from *Plasmodium falciparum* [53,54]. CET24H7.5 from *Caenorhabditis elegans* [55].

2.4×10^{-77}), while Yor291wp that displays a large aminoterminal domain is closely related to the CEWO8D2 (Z83217) ATPase from *C. elegans* (BLAST P-value of 1.2 10^{-193}).

A total of 7 P-type ATPases from *S. cerevisiae*, *P. falciparum*, *T. thermophila* and *C. elegans* can be clustered in this new ubiquitous family. All those which are complete share an unusual structural feature which had not been noticed yet among the P-type ATPases: that is a long extra membranous segment of 100–200 residues located between transmembrane spans 1 and 2 extruding on the opposite side of the other hydrophilic segments (Fig. 2).

Specific aminoacid sequence motifs have been identified in

Table 6
P4-ATPases signatures

Specific consensus	$^{D}_{L}$I•T--•PP$^{A}_{E}$LP	$^{M}_{F}$C***•**$^{C}_{V}$	••S-$^{A}_{C}$•$^{P}_{S}$FTS$^{K}_{N}$---•--•--••-$^{E}_{Q}$GR$^{C}_{A}$•LV$^{T}_{N}$•
U41552 (*C. eleg.*)	DIITIVVPPALP	MCGDGANDC	EASIAAPFTSNVPDIRCVPTVIKEGRCALVTS
Z70271 (*C. eleg.*)	DIITITVPPALP	MCGDGANDC	EASIAAPFTSKVPDIRCVPTVISEGRAALVTS
Z83217 (*C. eleg.*)	LILTSVIPPELP	MCGDGTNDV	DASIAAPFTSKYTSIASICHVIKQGRCTLVTT
Yor291wp (*S. cer.*)	DIITIVVPPALP	FCGDGANDC	EASVAAPFTSKIFNISCVLDVIREGRAALVTS
Yel031wp (*S. cer.*)	LIITSVVPPELP	MCGDGTNDV	DASCAAPFTSKLANVSAVTNIIRQGRCALVNT
A44396 (*P. falc.*)	DIITDAIPPALP	MCGDGANDC	ESSICSSFTSNKLCLHSIVHILIEGRASLVNS
U41063 (*T. therm.*)	DILIYSAPPGMP	MVGDGANDC	DGQFSSSYVSLSTSLSCVKRVLLEGRVNLSNS

The different motifs were obtained using Block Maker Server [52]. Highly conserved aminoacids in P-type ATPase sequences are indicated by * in the consensus. Underlined: conservative replacement, indicated by • in the consensus. Letters in indices and exposant indicate two non-equivalent aminoacids at the same position. Bold: conserved aminoacids in the P4-ATPase specific motifs.The A44396 accession number is from PIR, Z70271 and Z83217 from EMBL, U41552 and U41063 from GenBank.

all 7 members. A doublet of proline residues is present in predicted transmembrane span 4 (Table 2 and the first motif in Table 6). Given the helix breaking properties of proline, this putative transmembrane span must have an unusual structure. Other sequence motifs reported in Table 6 are specific and not detected in any other family of P-type ATPase: the two cysteine residues flanking the consensual hinge motif GDGxND and the xxS4xFTSx14GRxxLVxx sequence located just upstream the fifth transmembrane span.

The transported substrate(s) of these P-type ATPases are unknown; calcium is unlikely to be involved since transmembrane spans 4 and 6 show limited similarity to those of mammalian Ca^{2+}-ATPase (Table 4) and since calcium-specific binding residues in transmembranes span 4 and 6 are not detected with the possible exception of a nearby E residue in transmembrane span 4 of Yel031wp (*S. cerevisiae*) and CEW08D2 (*C. elegans*). Because of their unique properties, we propose to group these ORFs in a new group: the P4-type ATPases.

5. Conclusion

P-type ATPases are ubiquitous: one putative Ca^{2+}-ATPase have been identified in the 0.6 Mb genome of *Mycoplasma genitalium* considered as a minimal form of life [44] and one putative H^+-ATPase in the 1.7 Mb genome of the archebacteria *Methanococcus Jannaschii* [45].

Phylogenetic approach, based on aminoacid sequence similarity, allows to cluster the 16 different yeast ORFs encoding P-type ATPases in 6 families. Yeast contain 2 members of the group of P1-ATPases transporting heavy metals.

Ten other P-type ATPases seem to belong to the group of P2-ATPases transporting either H^+, Na^+ or Ca^{2+}. Among them, the 2 proton pumps, Pma1p and Pma2p, are the only yeast P-type ATPases which have been characterized biochemically.

Two novel features are revealed by the analysis of the complete yeast genome sequence. One of them is the existence of a 5-member family of the P2-ATPase group including Drs2p. This family has been recently proposed to translocate aminophospholipids [33]. The other new observation is the existence of 2 yeast ORFs of unknown transported substrate. Because of their unusual topology and characteristics motifs, we propose to classify them in a new group which today contains so far 8 members: the P4-ATPases.

Eventhough, the genetic tools provided by the yeast often allow a rapid understanding of some of the physiological functions of new genes ATPases, the limiting step for the biochemical study of the P-type ATPases reported here is likely to remain the development of appropriated overproduction and purification systems of these membrane proteins.

Acknowledgements: This work was supported by the Fonds National de la Recherche Scientifique and the Pôles d'Attraction Inter-Universitaire.

References

[1] Pedersen, P.L., Carafoli, E., Trends Biochem. Sci. 12 (1987) 146–150.
[2] Fagan, M.J., Saier Jr., M.H., J. Mol. Evol. 38 (1994) 57–99.
[3] Serrano, R., Biochim. Biophys. Acta 947 (1988) 1–28.
[4] Taylor, W.R., Green, N.M., Eur. J. Biochem. 179 (1989) 241–248.
[5] Lutsenko, S., Kaplan, J.H., Biochemistry 34 (1995) 15607–15613.
[6] Solioz, M., Vulpe, C., TIBS 21 (1996) 237–241.
[7] Goffeau, A., et al. Science 274 (1996) 546–567.
[8] Fu, D., Beeler, T., Dunn, T., Yeast 11 (1995) 283–292.
[9] Rad, R.M., Kirchrath, L., Hollenberg, C.P., Yeast 10 (1994) 1217–1225.

[10] Yuan, D.S., Stearman, R., Dancis, A., Dunn, T., Beeler, T., Klausner, R.D., Proc. Nat. Acad. Sci. USA 92 (1995) 2632–2636.
[11] Bassett Jr., D.E., Boguski, M.S., Hieter, P., Nature 379 (1996) 589–590.
[12] Serrano, R., Kielland-Brandt, M.C., Fink, G.R., Nature 319 (1986) 689–693.
[13] Goffeau, A., Slayman, C.W., Biochim. Biophys. Acta 639 (1981) 197–223.
[14] Schlesser, A., Ulaszewski, S., Ghislain, M., Goffeau, A., J. Biol. Chem. 263 (1988) 19480–19487.
[15] Supply, P., Wach, A., Thinès-Sempoux, D., Goffeau, A., J. Biol. Chem. 268 (1993) 19744–19752.
[16] Supply, P., Wach, A., Goffeau, A., J. Biol. Chem. 268 (1993) 19753–19759.
[17] Gaber, R., Int. Rev. Cytol. 137 (1992) 299–353.
[18] Andersen, J.P., Biosci. Reports 15 (1995) 243–261.
[19] Cyrklaff, M., Auer, M., Kühlbrandt, W., Scarborough, G.E., EMBO J. 14 (1995) 1854–1857.
[20] Rudolph, H.K., Antebi, A., Fink, G.R., Buckley, C.M., Dorman, T.E., LeVitre, J., Davidow, L.S., Mao, J.I., Moir, D.T., Cell 58 (1989) 133–145.
[21] Smith, R.A., Duncan, M.J., Moir, D.T., Science 229 (1985) 1219–1224.
[22] Antebi, A., Fink, G.R., Mol. Biol. Cell 3 (1992) 633–654.
[23] Harmsen, M.M., Langedijk, A.C., Van Tuinen, E., Geerse, R.H., Raue, H.A., Maat, J., Gene 125 (1993) 115–123.
[24] Schröder, S., Schimmöller, F., Singer-Krüger, B., Riezman, H., J. Cell Biol. 131 (1995) 895–912.
[25] Cunningham, K.W., Fink, G.R., Mol. Cell Biol. 16 (1996) 2226–2237.
[26] Cunningham, K.W., Fink, G.R., J. Cell Biol. 124 (1994) 351–365.
[27] Lapinskas, P.J., Cunningham, K.W., Liu, X.F., Fink, G.R., Culotta, V.C., Mol. Cell Biol. 15 (1995) 1382–1388.
[28] Loukin, S., Kung, C., J. Cell Biol. 131 (1995) 1025–1037.
[29] Carafoli, E., FASEB J. 8 (1994) 993–1002.
[30] Pozos, T.M., Sekler, I., Cyert, M.S., Mol. Cell Biol. 16 (1996) 3730–3741.
[31] Ripmaster, T.L., Vaughn, G.P., Woolford Jr., J.L., Mol. Cell Biol. 13 (1993) 7901–7912.
[32] Catty, P., Goffeau, A., Biosci. Reports 16 (1996) 75–85.
[33] Tang, X., Halleck, M.S., Schlegel, R.A., Williamson, P., Science 272 (1996) 1495–1497.
[34] Wieland, J., Nitsche, A.M., Strayle, J., Steiner, H., Rudolph, H.K., EMBO J. 14 (1995) 3870–3882.
[35] Rudolph, H.K., Fink, G.R., Yeast 6 (1990) S561.
[36] Martinez, R., Latreille, M.-T., Mirande, M., Mol. Gen. Genet. 227 (1991) 149–154.
[37] Haro, R., Garciadeblas, B., Rodriguez-Navarro, A., FEBS Lett. 291 (1991) 189–191.
[38] Garciadeblas, B., Rubio, F., Quintero, F.J., Banuelos, M.A., Haro, R., Rodriguez-Navarro, A., Mol. Gen. Genet. 236 (1993) 363–368.
[39] Mendoza, I., Rubio, F., Rodriguez-Navarro, A., Pardo, J.M., J. Biol. Chem. 269 (1994) 8792–8796.
[40] Nakamura, T., Liu, Y., Hirata, D., Namba, H., Harada, S., Hirokawa, T., Miyakawa, T., EMBO J. 12 (1993) 4063–4071.
[41] Hirata, D., Harada, S.-I., Namba, H., Miyakawa, T., Mol. Gen. Genet. 249 (1995) 257–264.
[42] Posas, F., Camps, M., Arino, J., J. Biol. Chem. 270 (1995) 13036–13041.
[43] Banuelos, M.A., Quintero, F.J., Rodriguez-Navarro, A., Biochim. Biophys. Acta 1229 (1995) 233–238.
[44] Frazer, C.M., et al. Science 270 (1995) 397–403.
[45] Bult, C.J., et al. Science 273 (1996) 1058–1073.
[46] Bull, P.C., Thomas, G.R., Rommens, J.M., Forbes, J.R., Cox, D.W., Nature Genet. 5 (1993) 327–336.
[47] Chelly, J., Tumer, Z., Tonnesen, T., Petterson, A., Ishikawa-Brush, Y., Tommerup, N., Horn, N., Monaco, A.P., Nature Genet. 3 (1993) 14–19.
[48] Odermatt, A., Suter, H., Krapf, R., Solioz, M., J. Biol. Chem. 268 (1993) 12775–12779.
[49] Guerini, D., Foletti, D., Vellani, F., Carafoli, E., Biochemistry 35 (1996) 3290–3296.
[50] Brandl, C.J., Green, N.M., Korczak, B., MacLennan, D.H., Cell 44 (1986) 597–607.
[51] Shull, G.E., Greeb, J., J. Biol. Chem. 263 (1988) 8646–8657.
[52] Henikoff, S., Henikoff, J.G., Alford, W.J., Pietrokovski, S., Gene 163 (1995) 17–26.
[53] Krishna, S., Cowman, G., Maede, J.C., Wells, R.A., Stringer, J.R., Robson, K.J., J. Cell Biol. 120 (1993) 385–398.
[54] Trottein, F., Cowman, A.F., Eur. J. Biochem. 227 (1995) 214–225.
[55] Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., et al. Nature 368 (1994) 32–38.
[56] Rost, B., Casadio, R., Fariselli, P., Sander, C., Prot. Sci. 4 (1995) 521–533.
[57] Rost, B., Fariselli, P., Casadio, R., Prot. Sci. 7 (1996) 1704–1718.