ELSEVIER

Biochimica et Biophysica Acta 1467 (2000) 207–218

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

www.elsevier.com/locate/bba

The yeast mitochondrial transport proteins: new sequences and consensus residues, lack of direct relation between consensus residues and transmembrane helices, expression patterns of the transport protein genes, and protein–protein interactions with other proteins

Roman Belenkiy [a], Amanda Haefele [a], Michael B. Eisen [b], Hartmut Wohlrab [a,*]

[a] *Boston Biomedical Research Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 64 Grove Street, Watertown, MA 02472, USA*

[b] *Lawrence Berkeley National Laboratories and Department of Molecular and Cellular Biology, University of California at Berkeley, Berkeley, CA, USA*

Received 2 September 1999; received in revised form 7 March 2000; accepted 13 April 2000

Abstract

Mitochondrial transport proteins (MTP) typically are homodimeric with a 30-kDa subunit with six transmembrane helices. The subunit possesses a sequence motif highly similar to *Pro X Asp/Glu X X Lys/Arg X Arg* within each of its three similar 10-kDa segments. Four (YNL083W, YFR045W, YPR021C, YDR470C) of the 35 yeast (S. cerevisiae) MTP genes were resequenced since the masses of their proteins deviate significantly from the typical 30 kDa. We now find these four proteins to have 545, 285, 902, and 502 residues, respectively. Together with only four other MTPs, the sequences of YPR021C and YDR470C show substitutions of some of the five residues that are absolutely conserved among the 12 MTPs with identified transport function and 17 other MTPs. We do now find these five consensus residues also in the new sequences of YNL083W and YFR045W. Additional analyses of the 35 yeast MTPs show that the location of transmembrane helix sequences do not correlate with the general consensus residues of the MTP family; protein segments connecting the six transmembrane helices and facing the intermembrane space are not uniformly short (about 20 residues) or long (about 40 residues) when facing the matrix; most MTPs have at least one transmembrane helix for which the sum of the negative hydropathy values of all residues yields a very small negative value, suggesting a membrane location bordering polar faces of other transmembrane helices or a non-transmembrane location. The extra residues of the three large MTPs are hydrophilic and at the N-terminal. The 200-residue N-terminal segment of YNL083W has four putative Ca^{2+}-binding sites. The 500-residue N-terminal segment of YPR021C shows sequence similarity to enzymes of nucleic acid metabolism. cDNA microarray data show that YNL083W is expressed solely during sporulation, while the expressions of YFR045W, YPR021C, and YDR470C are induced by various stress situations. These results also show that the 35 MTP genes are expressed under a rather diverse set of metabolic conditions that may help identify the function of the proteins. Interestingly, yeast two-hybrid screens, that will also be useful in identifying the function of MTPs, indicate that MIR1, AAC3, YOR100C, and YPR011C do interact with non-MTPs.

Keywords: Mitochondrion; Transport; Protein; Carrier; Expression; Gene

Abbreviations: PTP (or MIR1), phosphate transport protein; AAC, ADP/ATP translocase; MTP, mitochondrial transport protein or mitochondrial solute carrier protein

* Corresponding author. Fax: +1-617-972-1753; E-mail: wohlrab@bbri.org

0005-2736/00/$ – see front matter
PII: S0005-2736(00)00222-4

1. Introduction

Mitochondria have to be able to transport various ions (organic and inorganic) across the inner membrane to sustain metabolic processes and to permit oxidative phosphorylation with its requirement for a proton gradient across the membrane to proceed. Membrane proteins that catalyze some of these transport functions with a high degree of substrate specificity have been identified. Their sequences and some of their substrate specificities have been determined. Those that have been identified have sequence motifs similar to *Pro X Glu/Asp X X Lys X Arg* [1] within each of three highly similar sequence segments (about 100 residues each) within a protein of about 30 kDa [2,3], and have 12 transmembrane helices in their typical homodimeric structure [3]. Searching the entire yeast (*Saccharomyces cerevisiae*) genome for genes of proteins with these primary characteristics, several investigators have identified a group of such open reading frames (ORF) [4–7]. Among these, the largest group has 35 such ORFs [6].

We have now taken these 35 ORFs, identified consensus residues within the sequences of their proteins and determined how their transmembrane helix sequences relate to these consensus residues. We used published yeast DNA microarray results [8–11] to identify expression patterns of these 35 genes and to identify other mitochondria-related yeast genes with similar expression patterns. Of these 35 proteins, 12 have been shown to possess transport function. These 12 possess five absolutely conserved consensus residues. Seventeen other MTPs also have these five consensus residues. However six MTPs have substitutions in a subset of three of the five absolutely conserved residues. Our analysis of the results of two recent large-scale yeast two-hybrid screens [12] that will be helpful in identifying the function of MTPs and in characterizing their topology show that four MTPs interact with non-MTP proteins.

2. Materials and methods

2.1. Yeast genomic DNA preparation

Yeast (CG379) was grown in YPD medium and its DNA liberated by vortexing with glass beads. The DNA was separated/precipitated with phenol/chloroform/isoamyl alcohol and 100% ethanol [13]. The DNAse-free RNAse treatment step was dropped for higher fidelity of subsequent PCR reactions. The final DNA preparation was quantitated spectrophotometrically.

2.2. PCR, cloning, and sequencing

PCR amplification was performed with yeast GENEPAIRS primers (Research Genetics). Subcloning of PCR products was carried out with the TOPO TA cloning kit (Invitrogen) for sequencing. Plasmids were purified with the Plasmid Midi Kit (Qiagen). The genes were sequenced with an ABI373A automated DNA sequencer equipped with Stretch upgrade (Tufts DNA Sequencing Facility).

2.3. Computer programs

Fig. 1 was prepared with Microsoft Powerpoint.

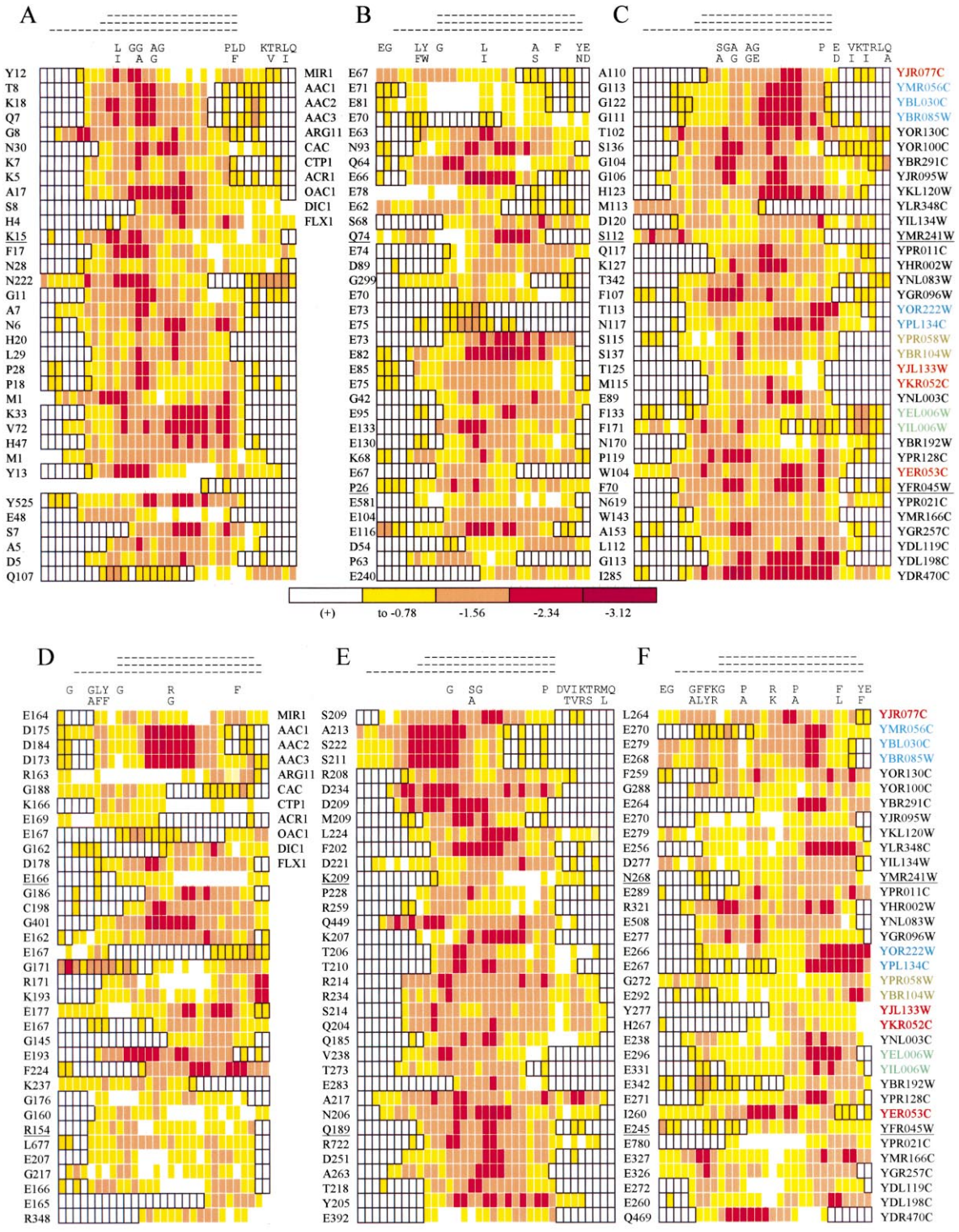
→

Fig. 1. Negative hydropathy values at residue positions within six (A–F) partial sequences of the 35 yeast MTPs. Each partial sequence is indicated to harbor one transmembrane helix. Color codes refer to negative hydropathy [14] (nine residue window) value ranges, e.g. white is a positive (+) hydropathy value, dark red indicates a negative hydropathy value from −2.35 to −3.12. Consensus residues (single letter code) (Fig. 2) are indicated at the top of the figure. The three dashed lines (top of figure) identify transmembrane helix residue segments suggested by reviewers (top [6], middle [7], bottom [32]) to apply uniformly to all MTPs and thus to correlate directly with the MTP consensus residues. Colored/white rectangles without black frame indicate likely residues of transmembrane helices as identified with hydropathy plots of each individual MTP. Names of the ORFs and names of some transport proteins are indicated. Sequences with names of the same color are most highly similar [6]. Top 12 sequences in the figure indicate proteins with identified transport function. The top 29 sequences possess five absolutely conserved residues (Fig. 2). The last six sequences have one or two residue substitutions at three of these five conserved residues (Table 2). The N-terminal residue (single letter code) of each partial sequence is shown with the number indicating its position with respect to the N-terminal residue of the intact protein translated from each ORF.

Hydropathy values were obtained with DNAStar (version 3.02) and hydrophilicity plots [14] with a nine-residue window.

2.4. MTP gene expression

Expression data was obtained from yeast cell cycle



experiments [8] (genome-www.stanford.edu/cellcycle/), sporulation experiments [9] (cmgm.stanford.edu/pbrown/sporulation/), diauxic shift experiments [10] (cmgm.stanford.edu/pbrown/explore/), and cluster analysis results [11] (rana.stanford.edu/clustering/).

3. Results

3.1. Sequences of YNL083W, YFR045W, YPR021C, YDR470C

The typical sequence of an MTP subunit consists of about 300 residues (30 kDa). The sequences of four (YNL083W, YFR045W, YPR021C, YDR470C) yeast genes suggest proteins of size sig-

Table 1
Expression of MTP genes[a]

Gene	Chromosome	Transport[b]	Residues	ORF	Cell cycle	Diauxic shift	Sporulation
MIR1	X	phosphate [15,16]	311	YJR077C	–	U(7)[c]	D(9,11.5)
AAC1	XIII	ADP/ATP [17]	309	YMR056C	–	U(6,7)	U(2,5,7,9,11.5)
AAC2	II	ADP/ATP [18]	318	YBL030C	Y (M)[d]	U(7)	–
AAC3	II	ADP/ATP [19]	307	YBR085W	–	–	–
ARG11	XV	ornithine [20,21]	292	YOR130C	Y (none)	–	–
CAC	XV	carnitine [22]	327	YOR100C	–	U(6,7)	U(0.5,2,5,7,11.5)
CTP1	II	tricarboxylate [23]	299	YBR291C	–	D(7)	–
ACR1	X	dicarboxylate [24]	322	YJR095W	–	U(3,6,7)	U(0.5) D(5,7,9,11.5)
OAC1	XI	oxaloacetate [4,25]	324	YKL120W	Y (none)	–	U(0.5,2,5,7,9,11.5)
DIC1	XII	dicarboxylate [26–28]	298	YLR348C	–	–	U(0.5) D(9,11.5)
FLX1	IX	flavin [29]	311	YIL134W	–	–	D(0.5,5,7)
nn	XIII	tricarboxylate [4]	314	YMR241W	–	D(6,7)	D(11.5)
nn	XVI		326	YPR011C	–	–	D(0.5)
nn	VIII		357	YHR002W	–	–	D(7,9,11.5)
nn	XIV		545	YNL083W	–	–	U(7,9,11.5)
nn	VII		314	YGR096W	–	–	–
nn	XV		307	YOR222W	–	–	–
nn	XVI		310	YPL134C	–	U(7)	D(7,9,11.5)
YMC1	XVI		307	YPR058W	–	D(6)	U(0.5,2)
YMC2	II		329	YBR104W	–	–	–
MRS3	X		314	YJL133W	–	–	–
MRS4	XI		304	YKR052C	–	–	D(5,7,9,11.5)
PET8	XIV		284	YNL003C	Y (G2)	–	–
nn	V		335	YEL006W	–	–	U(11.5)
nn	IX		373	YIL006W	–	–	D(5,7,9,11.5)
RIM2	II		377	YBR192W	Y (none)	–	–
nn	XVI		328	YPR128C	Y (M)	–	–
nn	V		300	YER053C	–	U(3,4,5,6,7)	D(11.5)
nn	VI		285	YFR045W	–	–	–
nn	XVI		902	YPR021C	–	–	–
nn	XIII		368	YMR166C	–	?	?
nn	VII		366	YGR257C	–	–	–
nn	IV		307	YDL119C	–	–	–
YHM1	IV		300	YDL198C	Y (none)	–	–
nn	IV		502	YDR470C	–	–	–

[a]See Section 2 for source of data.
[b]Dominant substrate transported by MTP.
[c]Increased expression (U) or decreased expression (D) at indicated sampling time points.
[d]Y means this gene shows increased expression in cell cycle phase M, G2, or phase cannot clearly be identified (none).

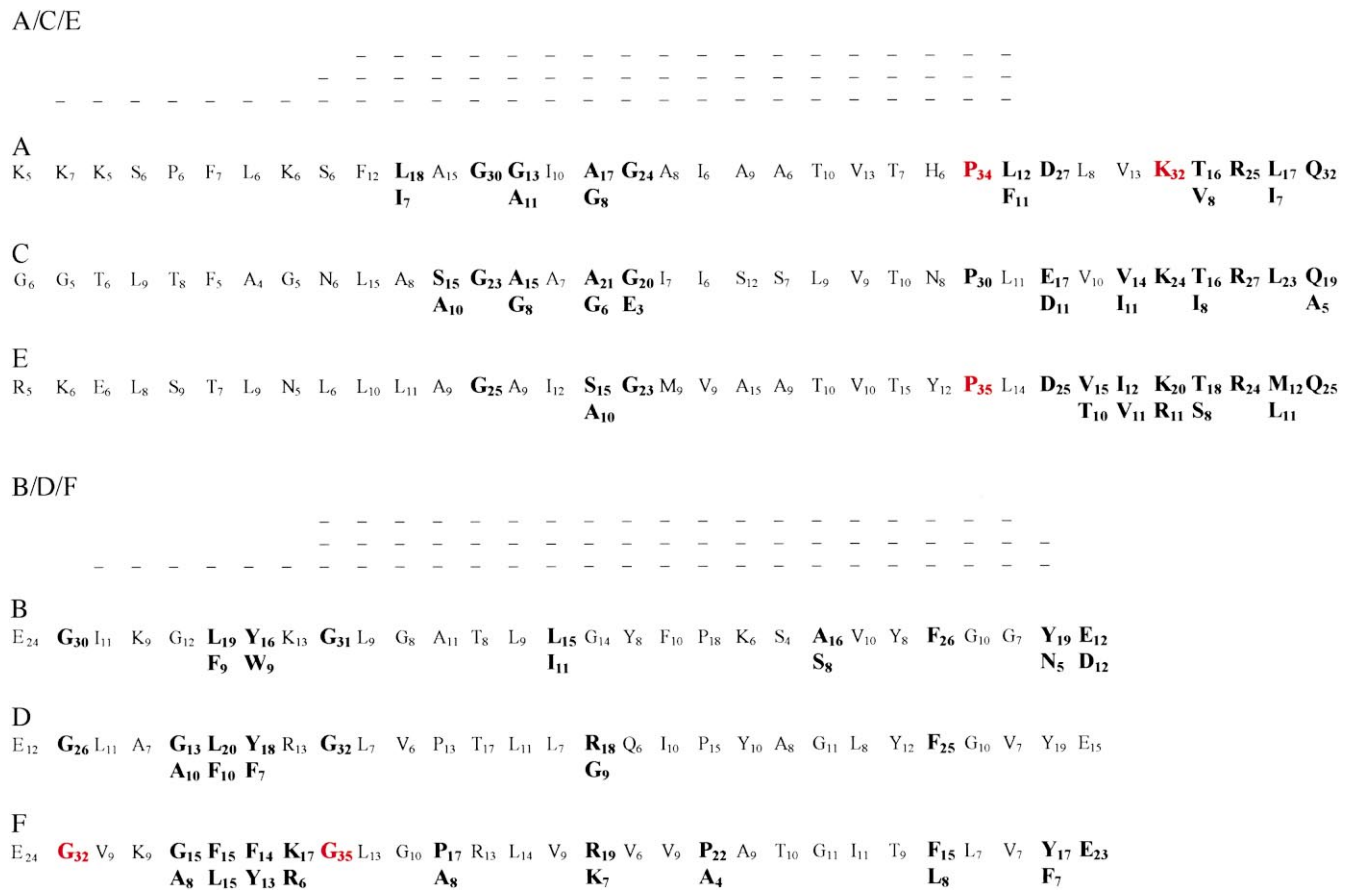


Fig. 2. Consensus sequences of the six (A–F) partial sequences (Fig. 1) of the 35 MTPs. Red residues are the only residues identical at identical consensus locations among the 12 MTPs with established transport function (top 12 sequences in Fig. 1). Bold black residues are the same for at least 23 MTPs; bold black double (two vertically aligned residues) are those where at least 23 MTPs have one of the two residues. The number associated with the residue indicates the number of sequences of the largest group with the same residue (residue shown) at this location. Dashed lines indicate generalized transmembrane helix sequences as suggested in several reviews [6,7,32] (see Fig. 1).

nificantly different from 300 residues, i.e. 178 (YFR045W) to 902 (YPR021C) residues. We therefore resequenced these four genes. Our sequence of YFR045W (chromosome VI) shows a G insertion between 242 181 (G) and 242 182 (T) (aggGtac) and a silent mutation of C to G (242 830) in Thr codons ACC to ACG (tacGtta). Both of these changes are not present in the GenBank sequence. Four independent sequencing runs yielded the same sequence. The resulting frameshift yields a 285-residue protein, clearly in line with the typical MTP (Table 1). Our YNL083W sequence yields a deletion of the G at position 472 581 (chromosome XIV) of the GenBank sequence (gggTgggC-tca, see also below for substitutions). Eight independent sequencing runs yielded this result. The result is a frameshift, which adds 153 nucleotides or 51 residues to the protein sequence of 494 to make it 545 residues. In addition, we find that a T replaces G at 472 576, C replaces the G at 472 580, and A replaces G at 471 764 (aagAtga).

Fig. 1 shows that YNL083W and YFR045W possess hydropathy profiles not too different from most other MTP sequences except that partial sequence A of YFR045W lacks all of the potential transmembrane helix residues, but does possess the consensus residues starting with Pro. Deleting the stop codon (TAA) located 29 codons upstream from the N-terminal Met, yields a 332-residue protein with a hydrophobic residue sequence of only 11 residues, too short for a transmembrane helix A. The hydrophobic

```
                x  y  z-y-x  -z
D 51EAIKMLFTAMDVNKDSVVDLSDFKKYASNA   (I)
E 82SQIWNGFQRIDLDHDGKIGINEINRYLSDL   (II)
L151RGQASHKKNTDNDRSKKTTDSDLYVTYDQW   (III)
T198AYSYFYLFNEDVDLSSEGDVTLINGFIRGF   (IV)
A   DDVKKAFAIIDQDKSGFIEEDELKLFLNQF   (Cparv)
I   AEFKEAFSLFDKDGDGTITTKELGTVMRSL   (TCalm)
E   EELAECFRIFDRNADGYIDAEELAEIFRAS   (RSTnC)
```

Fig. 3. Ca^{2+}-binding sites within the N-terminal of YNL083W. Roman numerals indicate the four sequences of the four sites. The residue number within the YNL083W sequence is indicated. *x,y,z,-x,-y,-z* are the calcium coordinating positions of the calcium binding loop [33]. L at *-z* in IV is the only calcium coordinating ligand that does not seem to be similar to the motifs of other such calcium binding sites. Cparv is carp parvalbumin [34], TCalm is *Tetrahymena* calmodulin [35], RSTnC is rabbit skeletal troponin C [36].

sequence segment of partial sequence F of the protein (YFR045W) is shifted towards the C-terminal end. Neither of the original GenBank sequences of these two genes possess the five absolutely conserved residues of Fig. 2. Both new sequences, however, do possess these five residues.

The YNL083W protein has an N-terminal extension of about 200 residues as indicated in Fig. 1, i.e. the N-terminal residue of partial sequence A is Asn222. This N-terminal extension is hydrophilic and harbors four putative Ca^{2+} binding sites (Fig. 3). Ca^{2+} binding and tissue expression has been shown for similar proteins from rabbit small intestine [30] and human heart [31].

Our resequencing of YPR021C (chromosome XVI) yielded a sequence identical to that in GenBank. Our YDR470C (chromosome IV) sequence had only one substitution at 1 400 863: GenBank's T is replaced with a C (tgtCcaa). This results in a replacement of a Glu (GAA) with a Gly (GGA) on the complementary strand where the gene is located.

3.2. Consensus residues and transmembrane helices

Thirty-five ORFs have been identified in the entire yeast genome [6]. Since it has not been shown whether transmembrane helix sequences, established by hydropathy plots, do indeed correlate with consensus residues, we aligned the 35 proteins with each other using the DNAStar program and optimized the alignments manually. Fig. 1A–F shows the alignments of the six partial sequences that have been suggested to harbor sequences for transmembrane helices. At the top of the figures we show the consensus residues that are presented in more detail in Fig. 2. Dramatic differences in hydrophobicity and thus location of hypothetical transmembrane helices within different MTPs are quite apparent (Fig. 1A–F). We have placed the 12 transport-identified proteins at the top, followed by the 17 proteins that also possess the five consensus residues as indicated in Table 2 and in Fig. 2 (bold red). The six MTPs at the bottom have substitutions at one or more of three of the five absolute consensus residue positions (see Table 2).

We have quantitated the hydrophobic value of the potential transmembrane helices by summing the negative hydropathy values of the positions. These sums are shown in Table 3. It is clear that while most sums are similar, exceptions are clearly present. Most MTPs have at least one transmembrane helix sequence with a very small sum (bold numbers in Table 3). Among the MTPs, these small sums are not associated with only one particular region of the protein. None of the MTPs (Table 3) shows such a small sum for transmembrane helix A.

The consensus sequence of Fig. 2 is presented in such a way that the bold black residues represent at least 23 of the 35 MTPs. The subscript indicates the number of MTPs with that residue at that location.

Table 2
The five absolutely conserved residues within the 12 functionally established MTPs (top 12 of Fig. 1)

Name	A[a]	A	E	F	F
Consensus[b]	P26	K31	P26	G2	G9
YPR021C	P26	K31	P26	**S**	G9
YMR166C	**S**	K31	P26	G2	G9
YGR257C	P26	**R**	P26	G2	G9
YDL119C	P26	K31	P26	**N**	G9
YDL198C	P26	**S**	P26	G2	G9
YDR470C	P26	**R**	P26	**L**	G9

The MTPs (bottom six in Fig. 1) shown are the only ones with one or more of three of these five residues replaced (replacements are shown bold in single letter code).

[a]Letters refer to partial sequences (Fig. 2).

[b]Consensus refers to the 12 MTPs with established transport function. Only A, E, and F possess residues at locations absolutely conserved among these 12 MTPs. Residue numbers refer to residue position with respect to the N-terminal residue in the sequences of Fig. 2.

Table 3
Sum of only negative hydropathy values [14] (nine-residue window) of residue positions of transmembrane helices as indicated from hydropathy plots and Table 4

Name	A[a]	B	C	D	E	F
MIR1	16.8	20.4	20.4	13.5	26.7	33.2
AAC1	17.5	**1.6**[b]	21.2	23.6	34.6	20.7
AAC2	20.3	**2.6**	21.9	21.5	29.4	26.5
AAC3	19.5	**2.9**	23.4	21.8	28.5	24.9
ARG11	23.2	19.3	35.4	14.8	18.0	26.7
CAC	21.5	24.3	20.1	12.3	32.9	17.8
CTP1	17.1	22.8	28.4	**6.1**	26.4	29.0
ACR1	17.2	26.4	25.6	**8.0**	14.9	**7.1**
OAC1	31.4	**9.5**	34.2	11.0	24.8	25.8
DIC1	14.8	20.7	23.8	12.2	22.5	18.9
FLX1	21.9	18.6	24.0	20.0	18.8	35.2
YMR241W	21.4	19.9	22.9	**4.8**	8.4	16.2
YPR011C	23.5	13.5	20.0	17.4	13.4	21.6
YHR002W	17.1	14.3	24.7	19.7	10.3	27.8
YNL083W	24.4	10.0	15.7	26.8	30.8	21.2
YGR096W	20.6	**6.9**	24.8	20.3	19.2	17.9
YOR222W	24.0	**6.8**	24.3	**2.5**	14.0	25.4
YPL134C	25.5	12.2	24.9	13.3	18.5	22.8
YPR058W	17.1	21.0	14.9	12.9	31.7	23.3
YBR104W	18.4	28.1	16.6	18.9	26.4	15.7
YJL133W	19.3	16.1	**8.6**	19.3	23.4	**8.4**
YKR052C	18.7	15.4	14.9	14.3	26.4	**9.3**
YNL003C	25.6	21.4	19.2	**7.0**	18.6	15.7
YEL006W	28.1	17.3	16.9	28.2	15.7	22.9
YIL006W	30.2	21.2	16.5	27.8	14.5	15.1
YBR192W	25.9	12.5	15.4	10.4	11.6	19.3
YPR128C	26.0	19.9	27.5	**8.4**	34.8	18.3
YER053C	18.1	**7.9**	29.8	**9.4**	31.8	23.3
YFR045W		22.1	28.3	**7.7**	20.2	21.2
YPR021C	25.2	**9.5**	23.4	15.6	21.2	**9.3**
YMR166C	11.1	18.9	18.3	**9.3**	21.6	24.2
YGR257C	21.4	22.0	26.0	**8.1**	24.0	21.5
YDL119C	19.3	14.5	20.5	12.6	16.4	13.1
YDL198C	20.3	**9.0**	28.5	**9.1**	26.7	20.1
YDR470C	12.2	21.2	42.7	16.6	8.8	17.2

[a]Letters refer to the six transmembrane helices associated in most cases with the partial sequences of Fig. 1.
[b]Bold numbers refer to transmembrane helices with an exceptionally small sum of negative hydropathy values.

Only those residues with the largest number of MTPs are shown. Thus a V_4 (21st position in partial sequence B) indicates that at that location only four MTPs have a valine and that there is no group with more than four proteins and the same residue at that location. Such a location is least restrictive for type of residue requirement. On the other hand, P_{34} (20th position in partial sequence A) indicates that 34 of the 35 proteins as aligned in Fig. 1 have a proline at this location.

3.3. Intertransmembrane helix segments

Much controversy exists as to the role of the protein segments that connect the transmembrane helices and whether they face the extramembrane space or are embedded in the membrane. The original MTP model [3] that proposed six transmembrane helices for the MTP subunit, based on the ATP/ADP translocase sequence [2], suggested long segments connecting the matrix ends of transmembrane helices A and B, C and D, E and F [3]. Those segments connecting the intermembrane space ends of transmembrane helices B and C, D and E were suggested to be short [3]. We have determined (Table 4) the number of residues connecting the hypothetical transmembrane helices as suggested by hydropathy plots and mostly shown in Fig. 1. Table 4 also shows the number of residues N-terminal to transmembrane helix A (Nt) and C-terminal to transmembrane helix F (Ct). Among the 12 proteins that have been shown to catalyze transport, only the CTP1, the YMR241W and the ADP/ATP translocases (AAC1, AAC2, AAC3) have intertransmembrane helix segment length distributions similar to those of the original model [3]. Beyond this, several intertransmembrane helix segments are exceptionally short (AAC1,2,3) or exceptionally long (AAC3, YNL083W, YBR192W, YGR257C, YDR470C).

3.4. Expression patterns of MTP genes

The new technology of DNA microarrays provides a unique opportunity to group genes according to their expression patterns. Genes with similar expression patterns may be functionally related. In this manner, a new gene with unknown gene product function may be identified by correlating its expression pattern with genes of a known metabolic pathway.

Table 1 shows expression patterns of the 35 ORFs. The results from the cell cycle experiments [8] identify AAC2 with increased expression in the M-phase. This is also true for YPR128C, which, however, shows no change in expression level during the dia-

Table 4
Number of residues within segments connecting transmembrane helices, derived from hydropathy plot, of the MTPs

Name	Nt[a]	A–B[b]	B–C	C–D	D–E	E–F	Ct	Walker profile[c]
MIR1	17	(T43) 17	(Q86) 35	(A138) 32	(A194) 19	(D236) 17	21	−
AAC1	14	(A31) 45	(N94) 26	(L139) 38	(Y198) 8	(S233) 51	4	+
AAC2	23	(A41) 42	(N104) 27	(L148) 41	(Y207) 8	(S242) 31	14	+
AAC3	12	(A30) 58	(G105) 15	(D138) 41	(F200) 7	(S231) 30	12	+
ARG11	10	(F34) 36	(N92) 15	(L135) 20	(Y190) 26	(V237) 31	0	−
CAC	37	(V61) 38	(Y120) 23	(T162) 18	(R203) 36	(T262) 31	9	−
CTP1	12	(F33) 35	(K95) 15	(L137) 35	(K197) 18	(L235) 42	0	+
ACR1	10	(L31) 40	(I88) 23	(L139) 23	(A183) 36	(N232) 44	26	−
OAC1	22	(M50) 21	(L97) 32	(T154) 37	(Y212) 20	(I257) 12	18	−
DIC1	20	(A44) 11	(C79) 27	(N130) 48	(D201) 12	(D229) 35	14	−
FLX1	14	(T42) 36	(E100) 27	(I153) 32	(D205) 19	(L249) 18	8	−
YMR241W	19	(M48) 25	(E97) 12	(V138) 37	(E197) 24	(V240) 41	13	+
YPR011C	23	(S54) 31	(A103) 25	(L150) 49	(Y213) 24	(D255) 45	6	+
YHR002W	33	(F61) 37	(Y116) 23	(L160) 48	(H229) 40	(E286) 40	10	−
YNL083W	219	(T245) 58	(S320) 29	(T370) 37	(A430) 23	(R479) 35	11	+
YGR096W	16	(T39) 43	(Y97) 17	(D134) 31	(Y189) 27	(E234) 49	10	+
YOR222W	8	(V35) 12	(K68) 53	(I144) 13	(K174) 43	(I239) 34	10	−
YPL134C	10	(V34) 16	(K70) 55	(L150) 35	(K209) 12	(K240) 43	9	−
YPR058W	25	(F46) 37	(N100) 22	(H143) 36	(A203) 18	(M247) 14	0	−
YBR104W	35	(F55) 31	(N109) 35	(Q165) 34	(G229) 12	(I267) 34	4	−
YJL133W	33	(A56) 35	(Y112) 23	(F151) 27	(Y204) 17	(Q248) 45	0	+
YKR052C	24	(A46) 35	(Y102) 24	(F141) 40	(S198) 13	(Q238) 42	6	+
YNL003C	0	(T27) 24	(Y71) 29	(E116) 45	(Y172) 23	(L218) 37	9	−
YEL006W	38	(V61) 44	(D123) 19	(V161) 38	(P217) 28	(H264) 38	13	+
YIL006W	74	(V100) 42	(Y160) 13	(T192) 45	(Y257) 23	(T296) 45	15	+
YBR192W	52	(L75) 64	(Y157) 21	(R202) 34	(L255) 42	(H309) 51	0	−
YPR128C	0	(T29) 44	(T96) 29	(A149) 33	(F202) 18	(Q251) 27	30	+
YER053C	19	(L46) 30	(Q86) 19	(I133) 33	(S196) 14	(I239) 22	17	+
YFR045W		(M56) 30	(C57) 20	(N98) 61	(Y181) 17	(D216) 42	4	−
YPR021C	529	(F553) 34	(N608) 17	(Q653) 29	(Y704) 28	(L755) 35	89	+
YMR166C	52	(L74) 40	(Y132) 20	(V171) 41	(Y234) 24	(D278) 49	11	+
YGR257C	18	(I38) 84	(F140) 21	(K183) 37	(Y244) 24	(F289) 38	14	+
YDL119C	13	(K35) 32	(S86) 36	(V143) 28	(Y193) 35	(D245) 32	8	−
YDL198C	11	(D32) 36	(K82) 42	(L142) 44	(K203) 12	(T236) 29	10	−
YDR470C	127	(T150) 105	(P279) 14	(F323) 37	(Q388) 14	(E419) 56	11	+

Value in parentheses is the first residue (single letter amino acid code) of that segment with its position counted from the N-terminal of the protein.

[a]Nt is the number of residues between N-terminal of protein and N-terminal of transmembrane helix A. Ct is number of residues between C-terminal of protein and C-terminal of transmembrane helix F.

[b]Indicates protein segment connecting transmembrane helices A and B.

[c]+ indicates that this MTP profile corresponds to the classical model with long A-B, C-D, and E-F segments and short B-C and D-E segments [3].

uxic shift experiment during which AAC2 does show a change.

3.5. Cluster analysis of MTP gene expression patterns

Eisen and coworkers [11] have developed an algorithm for identifying genes with similar expression patterns in several types of expression experiments. In Table 5, we have identified the MTP genes that have been cluster-analyzed [11] and we cite for each one of these genes other yeast mitochondria-related genes with similar expression patterns. We note here only that from among the three AAC genes, AAC2 appears to be expressed similar to more mitochon-

Table 5
Examples of mitochondria-related yeast genes with expression patterns (cluster analysis) similar to those of a select number of MTP genes[a]

Name	Cluster sector[c]	Mitochondria-related yeast genes with similar expression patterns
AAC1	5	ribonuclease P protein component (mitochondrial) (YML091C); ATP10 protein (essential for assembly of F_1–F_0 complex) (YLR393W)
YHM1/ SHM1	8	asparaginyl-tRNA synthetase (mitochondrial) (YCR024C)
AAC3 (expression subset of YHM1/ SHM1)[b]	8	dihydrolipoamide dehydrogenase (mitochondrial) (YFL018C); pyruvate dehydrogenase E1 component β-subunit (mitochondrial) (YBR221C); pyruvate dehydrogenase E1 component α-subunit (mitochondrial) (YER178W)
ACR1	10	cytochrome *c* oxidase polypeptide VA (YNL052W)
DIC1	10	YME1 protein (causes increased escape of DNA from mitochondria) (YPR024W)
FLX1 (expression subset of DIC1)	10	ABC1 protein (essential for electron transfer in bc_1 complex) (mitochondrial) (YGL119W)
MRS4	13	
MIR1 (expression subset of MRS4)	13	CDP-diacylglycerol-serine *O*-phosphatidyltransferase (microsomal/ outer mitochondrial membrane) (YER026C); cytochrome *c* oxidase biogenesis protein OXA1 (YER154W); D-lactate dehydrogenase (mitochondrial) (YDL174C)
PET8 (expression subset of ARG11)	16	glutamyl-tRNA synthetase (mitochondrial) (YOL033W)
ARG11	16	
CTP1	17	
YMC1 (expression subset of CTP1)	18	
MRS3	20	mitochondrial import inner membrane translocase subunit TIM23 (YNR017W)
YMC2	28	mitochondrial 40S ribosomal protein MRP4 (YHL004W)
RIM2	31	mitochondrial precursor proteins import receptor (TOM70) (YNL121C); mitochondrial peptide chain release factor 1 (YGL143C); CBP4 protein (assembly/stability of ubiquinol–cytochrome *c* reductase) (YGR174C); ATP11 protein (assembly of mitochondrial F_1–F_0 complex) (YNL315C); cytochrome *c* oxidase assembly protein COX14 (YML129C)
AAC2	35	outer mitochondrial membrane protein porin 1 (VDAC1) (YNL055C); succinate dehydrogenase (ubiquinone) flavoprotein subunit (YKL148C); rotenone-insensitive NADH-ubiquinone oxidoreductase (YML120C); ubiquinol-cytochrome *c* reductase complex 14-kDa protein (YDR529C); succinate dehydrogenase (ubiquinone) iron–sulfur protein (YLL041C); cytochrome c_1 (YOR065W); citrate synthase (mitochondrial) (YNR001C)

[a]MTP genes are arranged sequentially as they appear in Fig. 2 (http://rana.stanford.edu/clustering/) of the cluster analysis [11] to facilitate comparisons of expression patterns. Detailed information on the expressed genes can be found at http://www.expasy.ch/cgi-bin/sprot-search-de. Cluster analysis [11] is based on diauxic shift [10], sporulation [9], and cell cycle [8] experiments. Analysis is available for only 16 of the 35 MTP genes.
[b]Expression pattern of AAC3 is a subset of expression pattern associated with YHM1/SHM1 [11].
[c]Cluster analysis (Fig. 2) (http://rana.stanford.edu/clustering/) was divided into 35 sequential sectors to yield additional relational expressions of the genes.

drial genes than the other MTPs. It should be noted that AAC2 is in fact the major AAC in aerobically grown yeast [37]. Dramatic differences are apparent among the three ADP/ATP translocase genes since they are located in different sectors of the cluster analysis figure [11].

Table 6
Non-MTP interacting with MTP[a]

MTP	Interacting protein						
	Name	ORF	Function	Gene expression[b]			
				Cell cycle	Diauxic shift	Sporulation	Clustering sector[c]
MIR1	UBC6 (physical)	YER100W	ubiquitin conjugating enzyme, ATP required, selective degradation of misfolded membrane proteins	–	–	U(7,9)	3
AAC3	SAC6 (genetic)	YDR129C	fimbrin homolog, development and maintenance of cell polarity	–	U(6)	D(2,5,7,9,11.5)	12
AAC3	CCT4 (genetic)	YDL143W	cytoplasmic chaperonin subunit, ATP required, mitotic spindle formation, folding actin, tubulin	–	D(7)	–	27
CAC	RSC1 (physical)	YGR056W	member of remodel structure of chromatin complex, DNA-dependent ATPase activity	–	D(7)	–	11
YPR011C	RAD51 (physical)	YER095W	recombination and repair of DNA damage caused by X-rays, RecA homolog	Y (G1)	–	U(0.5,2,5,7,9,11.5)	5

[a]Data based on results obtained from yeast two-hybrid analyses [12].
[b]Refers to data as cited in Table 1.
[c]Refers to sectors as defined in Table 5.

3.6. Protein–protein interactions

It is generally assumed that the MTP proteins possess no protein segments that extend significantly out of the membrane region. These assumptions are based on hydrodynamic studies [38,39] and the behavior of these proteins in column chromatography [40]. Exceptions of course are YNL083W, YPR021C, and YDR470C with their large hydrophilic N-terminal and/or C-terminal domains. Thus it is not expected that the MTPs interact with non-MTPs. Yeast two-hybrid analyses suggest that four MTPs may interact with non-MTPs [12]. Table 6 lists the four MTPs and the proteins that have been suggested to interact with them. We have added the expression patterns of the genes of these new proteins to permit a comparison with those of the MTPs (Tables 1 and 5) to which they are suggested to bind.

4. Discussion

Fig. 1 summarizes the first detailed analysis of the individual MTPs with respect to their potential transmembrane helices. Among the 12 functionally identified transport proteins, the Pro of the consensus sequence of the partial sequences A, C, and E is at the C-terminal end of the hydrophobic transmembrane helix region. DIC1 has this hydrophobic region in C shifted dramatically towards the N-terminal. YMR241W has very little hydrophobicity in E compared to the other MTPs. ACR1 has very much hydrophobicity in B and so do CAC and YMR241W. The hydrophobic sequence associated with B of YOR222W and YPL134W have been shifted completely out of the partial sequence B region (see Table 3). However, as indicated in Table 3, the hydrophobicity of these shifted sequences is still small. Results of this analysis are differently presented in Table 3. All of the 35 proteins have a normal A hydrophobic region.

Among absolute consensus residues, K31 (partial sequence A) (Table 2) is critical for the phosphate transport protein (PTP or MIR1) function [41]. The two Pro and two Gly residues have not yet been replaced and thus the critical nature of these locations has not yet been established. The locations of G2 and G9 in partial sequence F (Table 2) are such that they are on the same side of a transmembrane helix cylinder and similar to the collagen triple helix

may be at locations where interactions with other transmembrane helices (rather than the membrane lipid phase) may sterically prevent the presence of a larger residue side group.

Table 3 shows that most MTPs have one transmembrane helix region with a very small hydrophobic value. It is likely that these transmembrane helices are located next to other transmembrane helices and provide ligands for the transmembrane path of the transported polar molecules. MTPs like PTP (MIR1) do not show such a small-valued transmembrane helix sequence. The summed values of Table 3, to be sure, are crude evaluations of transmembrane helices. Thus partial sequence A of MIR1 has a hydrophilic region (His-32) between two hydrophobic regions. The transmembrane helix sequence in A may provide a ligands for the cotransported proton [42].

Fig. 1 presents data that can be used to determine that the partial sequences C and D of OAC1 are connected by a short protein segment in the intact protein compared to the very long segment of AAC at the same location. However, this difference disappears when this segment refers to the segment that connects the transmembrane helices as determined from the hydropathy plot (Table 4). This conclusion can also be drawn from the segment connecting partial sequences D and E compared to segment length connecting the hypothetical transmembrane helices (Table 4). These results emphasize clearly that it is not appropriate to assign transmembrane helices based on the consensus residues (dashed lines in Fig. 1) as has been done by the authors of the three reviews [6,7,32]. It should also be noted that the ratios of the lengths of protein segments connecting transmembrane helices differ for many MTPs from those for AAC [3] and cannot be generalized (Table 4).

A first conclusion from the expression data of Table 1 is that the MTP genes do not behave alike. In fact they are spread out over the whole spectrum of yeast gene expression patterns (Table 5) [11]. A close look at the MTP sequences shows that several are very similar (see colored names in Fig. 1). Comparative sequence analyses were carried out by Nelson and coworkers [6]. Similar sequences might imply isozymes and thus reflect different expression patterns. This holds true for the three AACs (Table 1). Yet the MIR1 and YER053C with similar sequences do have a similar expression pattern. YER053C is upregulated earlier than MIR1 in the diauxic shift experiment and MIR1 is earlier down regulated in the sporulation experiment. MRS3 and MRS4 with similar hydrophobic sequence profiles (see Fig. 1) and sequences [6] have different expression patterns (Table 1). This holds also (Table 1) for sequence pairs of YMC1/YMC2, YOR222W/YPL134C, and YEL006W/YIL006W.

A much more thorough analysis of expression patterns, taking advantage of results as shown in Table 5, will be carried out in the future. We expect that similar expression patterns of genes within different metabolic pathways will help identify the functions of all MTPs.

Since evidence exists that all MTPs purified from mitochondria behave as if none of their protein domains face to a significant extent out of the mitochondrial membrane, it was indeed surprising that the yeast two-hybrid analyses [12] suggest that four MTPs interact with non-MTPs. This is shown in Table 6. The authors do acknowledge that some reproducible two-hybrid signals are unlikely to reflect true interactions. The suggested interactions are provocative and should be investigated. One should also remember, however, that the MTPs are located in the inner membrane with obvious access to the cytosol limited by the outer membrane. The gene expression patterns of the non-MTPs (Table 6) do not correlate with those of the four MTPs (Tables 1 and 5). This of course is not an absolute requirement for interacting proteins.

Acknowledgements

This research was supported by NIH Grant GM 57563.

References

[1] H.V. Kolbe, H. Wohlrab, J. Biol. Chem. 260 (1985) 15899–15906.
[2] H. Aquila, D. Misra, M. Eulitz, M. Klingenberg, Hoppe Seylers Z. Physiol. Chem. 363 (1982) 345–349.
[3] M. Saraste, J.E. Walker, FEBS Lett. 144 (1982) 250–254.
[4] J.A. Mayor, D. Kakhniashvili, D.A. Gremse, C. Campbell,

R. Kramer, A. Schroers, R.S. Kaplan, J. Bioenerg. Biomembr. 29 (1997) 541–547.
[5] B. el Moualij, C. Duyckaerts, J. Lamotte-Brasseur, F.E. Sluse, Yeast 13 (1997) 573–581.
[6] D.R. Nelson, C.M. Felix, J.M. Swanson, J. Mol. Biol. 277 (1998) 285–308.
[7] F. Palmieri, B. van Ommen, in: S. Papa, F. Guerrieri, J.M. Tager (Eds.), Frontiers in Cellular Bioenergetics, Kluwer Academic/Plenum Publisher, New York, 1999, pp. 489–519.
[8] P.T. Spellman, G. Sherlock, M.Q. Zhang, V.R. Iyer, K. Anders, M.B. Eisen, P.O. Brown, D. Botstein, B. Futcher, Mol. Biol. Cell 9 (1998) 3273–3297.
[9] S. Chu, J. DeRisi, M. Eisen, J. Mulholland, D. Botstein, P.O. Brown, I. Herskowitz, Science 282 (1998) 699–705.
[10] J.L. DeRisi, V.R. Iyer, P.O. Brown, Science 278 (1997) 680–686.
[11] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Proc. Natl. Acad. Sci. USA 95 (1998) 14863–14868.
[12] P. Uetz, L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J.M. Rothberg, Nature 403 (2000) 623–627.
[13] C.S. Hoffman, in: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds.), Current Protocols in Molecular Biology, John Wiley and Sons, Boston, 1994, pp. 13.11.2–13.11.3.
[14] J. Kyte, R.F. Doolittle, J. Mol. Biol. 157 (1982) 105–132.
[15] A. Phelps, H. Wohlrab, J. Biol. Chem. 266 (1991) 19882–19885.
[16] A. Phelps, C.T. Schobert, H. Wohlrab, Biochemistry 30 (1991) 248–252.
[17] G.S. Adrian, M.T. McCammon, D.L. Montgomery, M.G. Douglas, Mol. Cell. Biol. 6 (1986) 626–634.
[18] J.E. Lawson, M.G. Douglas, J. Biol. Chem. 263 (1988) 14812–14818.
[19] J. Kolarov, N. Kolarova, N. Nelson, J. Biol. Chem. 265 (1990) 12711–12716.
[20] L. Palmieri, V. De Marco, V. Iacobazzi, F. Palmieri, M.J. Runswick, J.E. Walker, FEBS Lett. 410 (1997) 447–451.
[21] M. Crabeel, O. Soetens, M. De Rijcke, R. Pratiwi, R. Pankiewicz, J. Biol. Chem. 271 (1996) 25011–25018.
[22] L. Palmieri, F.M. Lasorsa, A. Vozza, M.J. Runswick, F. Palmieri, J.E. Walker, in: International Symposium on Molecular Basis of Biomembrane Transport, Bari, Italy, 1999, p. 22.
[23] R.S. Kaplan, J.A. Mayor, D.A. Gremse, D.O. Wood, J. Biol. Chem. 270 (1995) 4108–4114.
[24] L. Palmieri, F.M. Lasorsa, A. De Palma, F. Palmieri, M.J. Runswick, J.E. Walker, FEBS Lett. 417 (1997) 114–118.
[25] L. Palmieri, A. Vozza, G. Agrimi, V. De Marco, M.J. Runswick, F. Palmieri, J.E. Walker, J. Biol. Chem. 274 (1999) 22184–22190.
[26] D. Kakhniashvili, J.A. Mayor, D.A. Gremse, Y. Xu, R.S. Kaplan, J. Biol. Chem. 272 (1997) 4516–4521.
[27] L. Palmieri, F. Palmieri, M.J. Runswick, J.E. Walker, FEBS Lett. 399 (1996) 299–302.
[28] J. Lancar-Benba, B. Foucher, M. Saint-Macary, Biochimie 78 (1996) 195–200.
[29] A. Tzagoloff, J. Jang, D.M. Glerum, M. Wu, J. Biol. Chem. 271 (1996) 7392–7397.
[30] F.E. Webert, G. Minestrini, J.H. Dyer, M. Werder, D. Boffelli, S. Compassi, E. Wehrli, R.M. Thomas, G. Schultheiss, H. Hauser, Proc. Natl. Acad. Sci. USA 94 (1997) 8509–8514.
[31] A.D. Arco, J. Satrustegui, J. Biol. Chem. 273 (1998) 23327–23334.
[32] J.E. Walker, Curr. Opin. Struct. Biol. 2 (1992) 519–526.
[33] J. Gariepy, R.S. Hodges, FEBS Lett. 160 (1983) 1–6.
[34] C.J. Coffee, R.A. Bradshaw, J. Biol. Chem. 248 (1973) 3305–3312.
[35] D.R. Marshak, D.M. Watterson, L.J. Van Eldik, Proc. Natl. Acad. Sci. USA 78 (1981) 6793–6797.
[36] J.P. Van Eerd, K. Takahashi, Biochem. Biophys. Res. Commun. 64 (1975) 122–127.
[37] M. Knirsch, M.P. Gawaz, M. Klingenberg, FEBS Lett. 244 (1989) 427–432.
[38] M. Klingenberg, C.S. Lin, Methods Enzymol. 126 (1986) 490–498.
[39] H. Hackenberg, M. Klingenberg, Biochemistry 19 (1980) 548–555.
[40] H. Wohlrab, A. Collins, D. Costello, Biochemistry 23 (1984) 1057–1064.
[41] C. Briggs, L. Mincone, H. Wohlrab, Biochemistry 38 (1999) 5096–5102.
[42] A. Phelps, C. Briggs, L. Mincone, H. Wohlrab, Biochemistry 35 (1996) 10757–10762.