Sequence Alignments of the H^+-Dependent Oligopeptide Transporter Family PTR: Inferences on Structure and Function of the Intestinal PET1 Transporter

Richard C. Graul[1] and Wolfgang Sadée[1,2]

Received December 24, 1996; accepted February 4, 1997

Purpose. To study the structure and function of the intestinal H^+/peptide transporter PET1, we compared its amino acid sequence with those of related transporters belonging to the oligopeptide transporter family PTR, and with more distant transporter families.

Methods. We have developed a new approach to the sequence analysis of proteins with multiple transmembrane domains (TMDs) which takes into account the repeated TMD-loop topology. In addition to conventional analyses of the entire sequence, each TMD and its adjacent loop residues (=TMD segments) were analyzed separately as independent structural units. In combination with hydropathy analysis, this approach reveals any changes in the order of the TMD segments in the primary structure and permits TMD alignments among divergent structures even if rearrangements of the order of TMD segments have occurred in the course of evolution.

Results. Alignments of TMD segments indicate that the TMD order in PTR transporters may have changed in the process of evolution. Consideration of such changes permits the alignment of homologous TMD segments from PTR transporters belonging to distant akaryotic and eukaryotic phyla. Multiple alignments of TMDs reveal several highly conserved regions that may play a role in transporter function. In comparing the PTR transporters with other transporter gene families, alignment scores using the entire primary structure are too low to support a finding of probable homology. However, statistically significant alignments were observed among individual TMD segments if one disregards the order in which they occur in the primary structure.

Conclusions. Our results support the hypothesis that the PTR transporters may have evolved by rearrangement, duplication, or insertions and deletions of TMD segments as independent modules. This modular structure suggests new alignment strategies for determining functional domains and testing relationships among distant transporter families.

KEY WORDS: PET1; oligopeptide transporter; sequence analysis.

INTRODUCTION

Delivering essential nutrients to cells, oligopeptide transporters are expressed in many organisms of diverse phyla. Of particular pharmaceutical interest is the role of the di-/tripeptide transporters in the bioavailability of peptoid drugs, including β-lactam and cephalosporin antibiotics, renin inhibitors, and ACE inhibitors (1–4). Recently, two closely related, mammalian H^+-coupled dipeptide transporters have been cloned, the intestinal PET1 and the renal PET2 transporter (5–8). The PET1 gene product can account for part of the transporter-dependent oral bioavailability of peptoid drugs, displaying the expected substrate selectivities (9,10). Sequence analysis indicated that these transporters are homologs of a family of peptide transporters distributed over distantly related phyla, including prokaryotes. The proposed PTR family (11) consists of H^+-dependent peptide transporters and a H^+-dependent chlorate/nitrate transporter in *Arabidopsis thaliana*, while this same family termed POT (proton-dependent oligopeptide transporters) in a separate report (12) was shown to contain an additional putative transporter of *E. coli*. We will refer to this group of transporters as the PTR family, including additional homologs identified in the SwissProt database in the present study, shown in Table 1A. No clear relationships of the PTR family have been established to other transporter families, examples of which are shown in Table 1B (13,14), because of substantial dissimilarities in their primary structures (11).

Hydropathy analysis suggested that PET1 and PET2 have 12 membrane spanning α-helical domains (TMDs) (5–8), whereas in the SwissProt database, 13 TMDs are proposed for PET2. Thus, they share a structure common to a large number of other mammalian transporters (13), even though they differ in their primary sequences. However, physical evidence for the number and orientation of the TMDs is lacking. Moreover, because there are no further mammalian homologs presently known with which PET1 and PET2 can be aligned, their structure and function cannot be deduced from previous analyses of 12-TMD transporters. Thus, inferences from sequence analysis on structure and function are limited to alignments with homologs in much more ancient phyla, such as the DTPT_LACLA peptide transporter in *Lactococcus* for which some structural information is available (15). Yet, because of the large evolutionary distances, sequence alignments are successful with only a portion of the primary sequence, which could lead to ambiguities in the alignments of homologous TMDs and their relative position within the primary structure.

This study addresses the relationships among the primary structures of the PTR transporter family (Table 1A). Recognizing that these proteins consist functionally of multiple repeat elements, i.e., the TMDs and their adjacent loops or C- and N-terminal tails (termed TMD segments), we have developed a new approach to sequence analysis which focuses on each of the TMD segments individually. This approach accounts for possible insertions or deletions, exon shuffling, cross-overs, intragenic repeats, and other processes that could have altered the order of the TMD segments in the primary structure during the course of evolution. If used in conjunction with common alignment algorithms involving the entire linear sequence of the protein, we can determine whether the order of TMD segments has changed during evolution and identify conserved domains with possible relevance to PTR transporter function.

Relying on individual analysis of TMD segments in addition to alignments of the entire primary structure, we also address the question whether relationships can be detected between the PTR transporters and other transporters with a proposed structure comprised of 9–15 TMDs (Table 1B).

[1] Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143-0446.

[2] To whom all correspondence should be addressed. (e-mail.sadee@cgl.ucsf.edu)

0724-8741/97/0400-0388$12.50/0

The results suggest that the PTR transporters may represent modular proteins in which the order of TMD segments could have changed by several possible evolutionary mechanisms. Alignments of corresponding TMDs in members of the PTR family from distant phyla reveals the location of highly conserved regions that may play a role in transport function. Lastly, alignments in several TMD segments suggests possible relationships between PTR transporters with other transporter families. These findings define a new approach to determining the functional domains within the peptide transporter proteins and suggest pathways in the evolution of the superfamily of transporter proteins that need to be explored on the basis of their proposed modular structure.

METHODS

Alignments of PET1_HUMAN with Related Transporters

Using the PET1_HUMAN protein sequence, a BLAST-based search of protein databases was performed to identify and align related sequences (16). The nomenclature is that used

Table 1. List of Transporters with 9–15 TMDs Analyzed in this Study

A. Proton-dependent and putative transporters of the PTR gene family

CHL1_ARATH, nitrate/chlorate transporter, *Arabidopsis thaliana* (plants)* (31)
DTPT_LACLA, di-tripeptide transporter, *Lactococcus lactis* (bacteria)*,** (15)
PET1_HUMAN, di-tripeptide transporter, human, intestinal isoform* (6)
PET1_RABIT, di-tripeptide transporter, rabbit, intestinal isoform*,** (7)
PET2_RABIT, di-tripeptide transporter, rabbit, kidney isoform (8)
PT2A_ARATH, peptide/histidine transporter, *Arabidopsis thaliana* (plants)* (32)
PT2B_ARATH, peptide/histidine transporter, *Arabidopsis thaliana* (plants)* (33)
PTR2_CANAL, small peptide transporter, *Candida albicans* (fungi)*,** (34)
PTR2_YEAST, small peptide transporter, *Saccharomyces cerevisiae* (yeast)* (35)
YHIP_ECOLI, putative transporter, *E.coli* (bacteria)** (36)
YJDL_ECOLI, putative transporter, *E.coli* (bacteria) (37)

B. Other Transporter Groups

group	members	starter sequence
acrd	4	ACRD_ECOLI, acriflavin resistance protein.
alp1	9	ALP1_YEAST, basic amino-acid permease.
arae	2	ARAE_ECOLI, arabinose-proton symport (arabinose transporter).
arop	2	AROP_ECOLI, aromatic amino acid transport protein arop.
b3at	4	B3AT_HUMAN, band 3 anion transporter.
bmrp	2	BMRP_CANAL, benomyl/methotrexate resistance protein.
cftr	4	CFTR_BOVIN, cystic fibrosis transmembrane conductance regulator (cftr).
cita	2	CITA_SALTY, citrate-proton symporter.
dal4	2	DAL4_YEAST, allantoin permease (allantoin transport protein).
dcta	3	DCTA_ECOLI, c4-dicarboxylate transport protein.
gltp	2	GLTP_BACSU, proton glutamate symport protein.
gtr1	7	GTR1_BOVIN, glucose transporter type 1, erythrocyte/brain.
hxt1	5	GAL2_YEAST, galactose transporter (galactose permease).
itr1	2	ITR1_YEAST, myo-inositol transporter.
lacy	3	LACY_ECOLI, lactose permease (lactose-proton symport).
mdr1	3	MDR1_CRIGR, multidrug resistance protein 1 (p-glycoprotein 1).
melb	2	MELB_ECOLI, melibiose carrier protein.
nah1	1	NAH1_HUMAN, Na^+/H^+ exchanger 1(amiloride-sensitive).
nanu	1	NANU_RABIT, sodium/nucleoside cotransporter.
ntg1	11	NTBE_CANFA, sodium- and chloride-dependent betaine transporter.
nupg	1	NUPG_ECOLI, nucleoside permease nupg.
putp	1	PUTP_ECOLI, sodium/proline symporter (proline permease).
tcr1	5	TCR1_ECOLI, tetracycline resistance protein (transposon tn10).
tcr2	3	TCR2_BACSU, tetracycline resistance protein.
th11	2	TH11_TRYBB, glucose transporter.
trk1	2	TRK1_SACUV, potassium transport protein, high-affinity.
vmt1	2	VMT1_RAT, chromaffin granule amine transporter (vesicular amine transporter 1).
ycs7	2	YCS7_YEAST, hypothetical 102.5 kd membrane protein in rbk1 3'region.
yim0	2	YIM0_YEAST, hypothetical 61.8 kd protein in kgd1-rpi1 intergenic region.

Note: Individual sequences are denoted in capital letters. Members of the PTR family of proton coupled receptors defined in (11) are labeled with *, and members of the POT family defined in (12) are labeled with **. These 29 groups were established on the basis of similarity among only the TMDs. The group name appears in small letters, and the starter sequence which served to search for additional members to be included with a group is given in capital letters. The number of members in the group is also provided. Homologs with sequence identities $\geq$70% were excluded. A full list of sequences is available upon request from the authors.

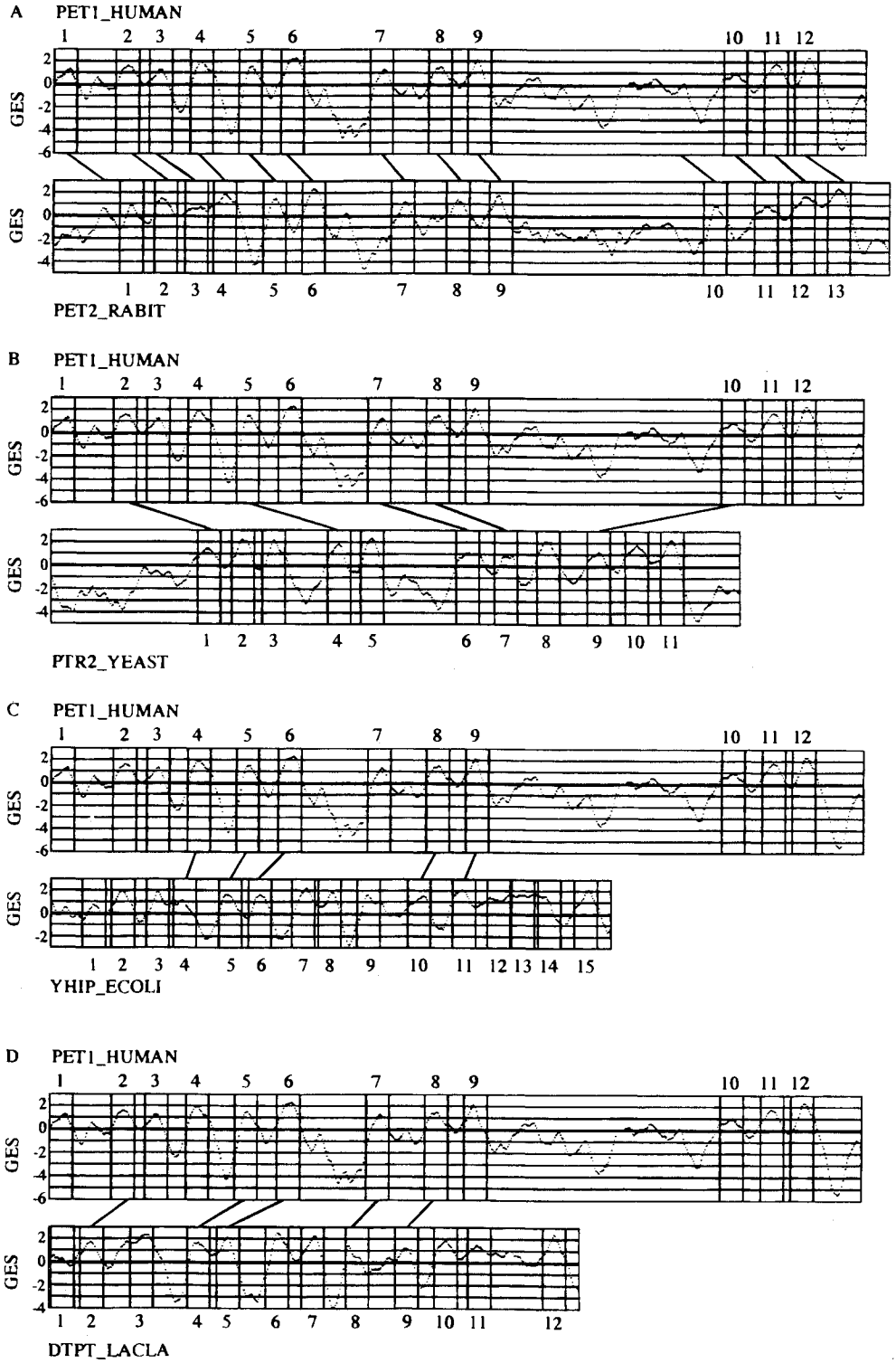


Fig. 2. Hydropathy profile of PET1 compared against profiles from selected other transporters of the PTR family. Hydropathy was calculated using a 7-residue window and the GES scale (20). Profiles of four representative members of the PTR family are compared to that of PET1_HUMAN. The lines connecting TMDs between two profiles indicate where a pairwise alignment of the two TMD segments scored with PAM120 > 20. These lines were allowed to connect any TMD to only one other TMD, and they were not allowed to cross. As a consequence, alignments between one TMD segment and another segment in an entirely different location (e.g., TMD 1 with TMD 12), or any possible internal repeats, are not shown even if they reach PAM120 > 20. The lines therefore connect the most likely homologous TMDs in corresponding locations of the primary strucutre, as they occur in sequential order. In some cases, a preassigned TMD aligned with a region in another sequence that was not considered a TMD in the SwissProt database (see panel A). Such a region was similarly analyzed by pairwise alignment, and the connecting line again indicates a PAM120 score above 20 (example TMD1 of PET1 and YMD10 of PET2 in Panel A).

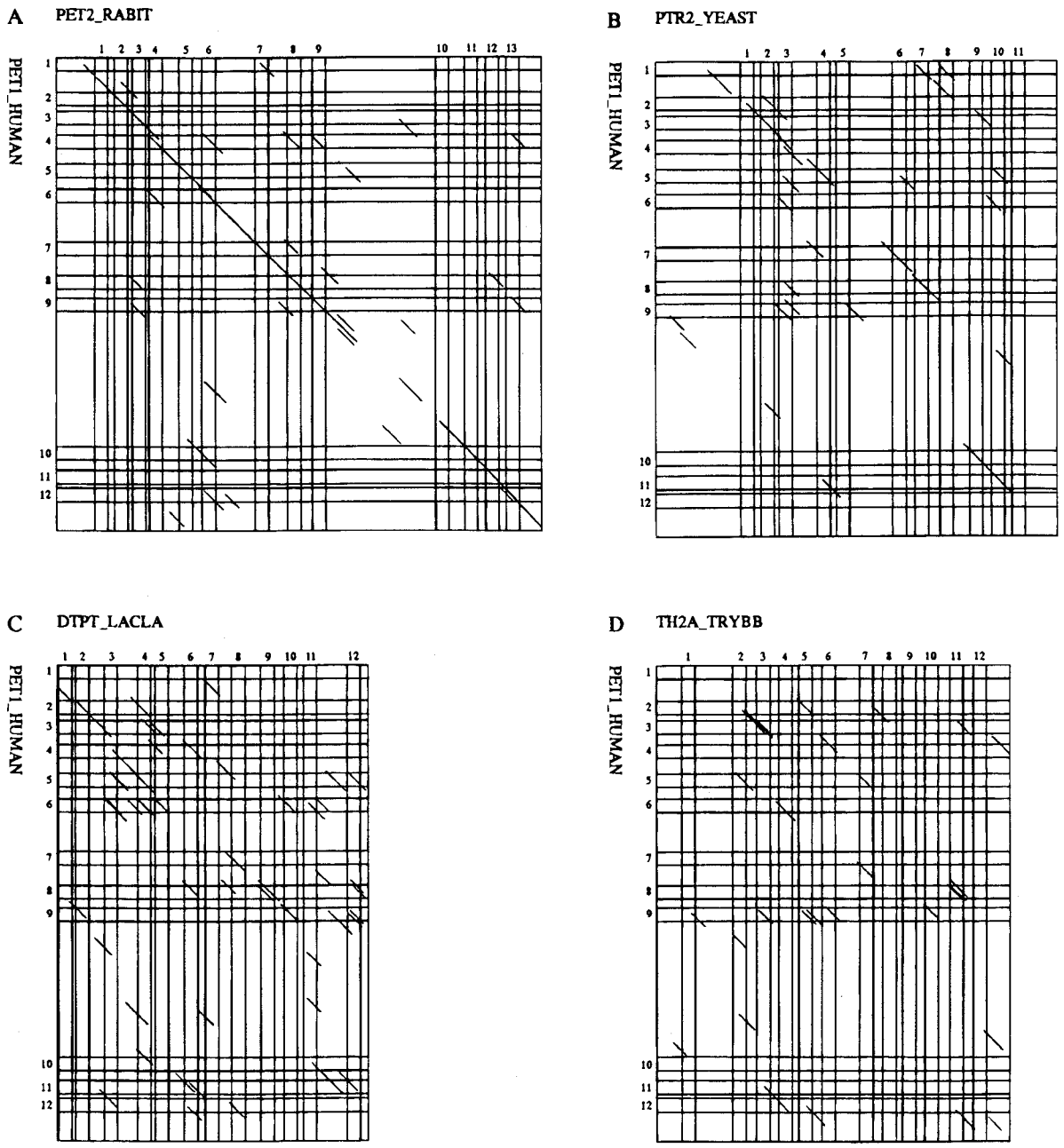


Fig. 1. 2D-Matrix sequence comparisons of PET1 with selected members of the PTR transporter family, and with the trypanosome glucose transporter TH11_TRYBB. Any alignments of sequence fragments (21–41 residues) scoring with PAM120 > 20 are represented by diagonal lines. Parallel lines horizontally or vertically indicate the presence of possible repeat structures.

nSD_{MC} (number of standard deviations separating the PAM120 score of the aligned two sequences from the mean of the PAM120 scores obtained from all randomized sequence comparisons). This provided an independent method for assessing the significance of a match where $nSD_{MC} > 3$ indicates possible homology and $nSD_{MC} > 6$ indicates probable homology (18,25,26).

ALIGNMENTS OF LINKED TMD SEGMENTS

Linking aligned sequence domains in the order they occur in the primary structure was previously used to test for homology between distantly related transporter proteins with a 12-TMD structure, where an nSD_{MC} value of 9 or more for an alignment consisting of several sequence fragments was considered a strong indicator of homology (26). We required a minimum length of 21 residues per fragment and PAM120 of 0 as the minimum score for inclusion in the linked structure. However, it was uncertain as to which TMD segment in one sequence should be aligned with which TMD segment in the second sequence because of possible rearrangements of the order of TMD segments during evolution. In order to automate the task of finding the optimal linkage sequence of the TMD segments, we designed a grid of TMDs in a 2D matrix in which we searched

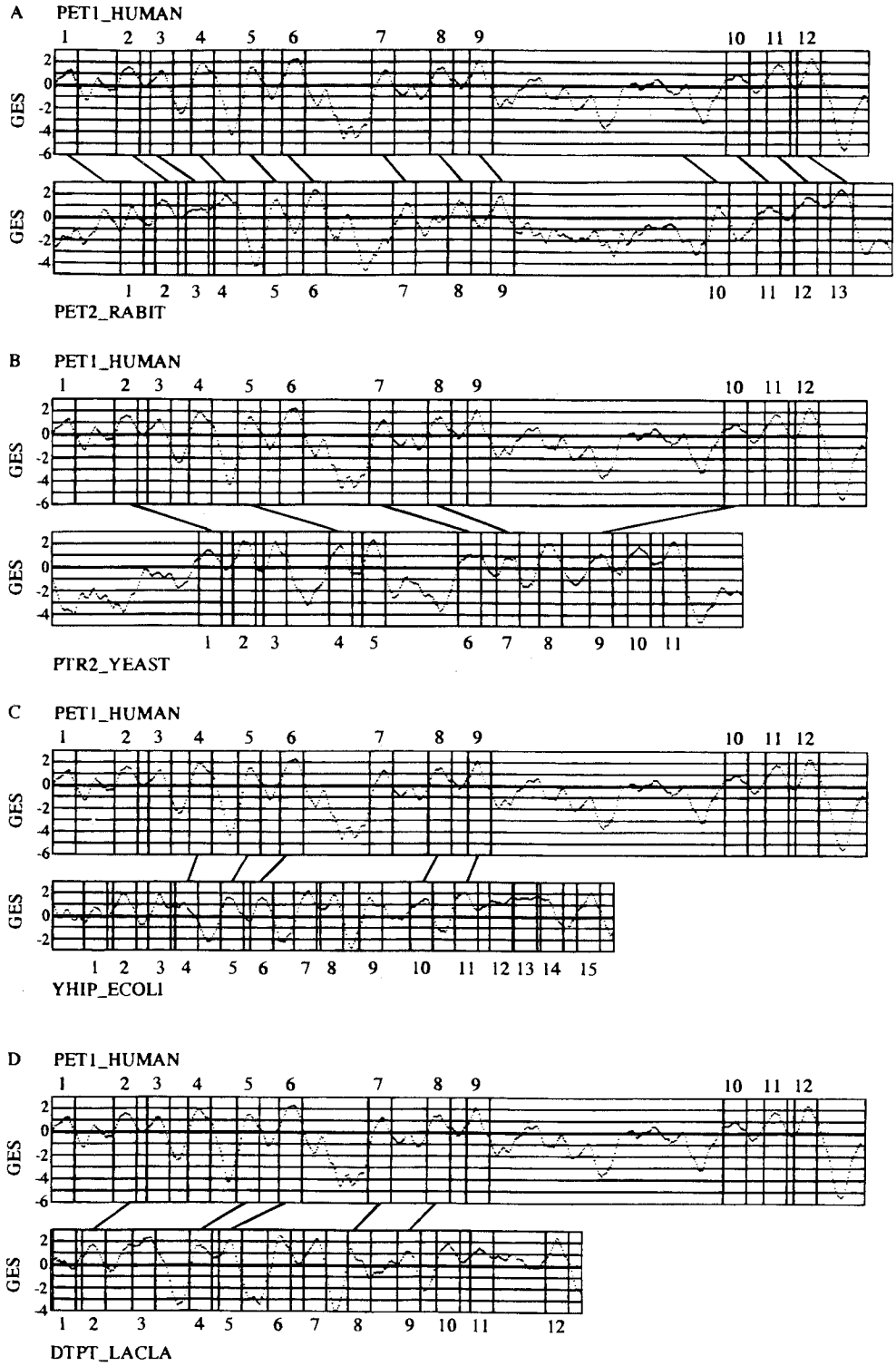


Fig. 2. Hydropathy profile of PET1 compared against profiles from selected other transporters of the PTR family. Hydropathy was calculated using a 7-residue window and the GES scale (20). Profiles of four representative members of the PTR family are compared to that of PET1_HUMAN. The lines connecting TMDs between two profiles indicate where a pairwise alignment of the two TMD segments scored with PAM120 > 20. These lines were allowed to connect any TMD to only one other TMD, and they were not allowed to cross. As a consequence, alignments between one TMD segment and another segment in an entirely different location (e.g., TMD 1 with TMD 12), or any possible internal repeats, are not shown even if they reach PAM120 > 20. The lines therefore connect the most likely homologous TMDs in corresponding locations of the primary strucutre, as they occur in sequential order. In some cases, a preassigned TMD aligned with a region in another sequence that was not considered a TMD in the SwissProt database (see panel A). Such a region was similarly analyzed by pairwise alignment, and the connecting line again indicates a PAM120 score above 20 (example TMD1 of PET1 and YMD10 of PET2 in Panel A).

for the best combination of linked TMDs which would give the highest combined PAM120 value. The only requirement for the search algorithm was that the search had to begin at TMD1 and proceed to the next TMD segments in the primary order, but could not go back towards the beginning of the sequence. The next higher TMDs can be skipped in either sequence, thereby, allowing for gaps and shifts in the order of the aligned TMD segments. For the resultant linked sequences, we calculated the number of standard deviations separating the true alignment from 1,000 serial randomizations as described above for single TMD segments (nSD_{MC}).

Analysis of Aligned Sequences

The SeqVu program (Garvan Institute of Medical Research, Sydney, Australia) was used to analyze aligned individual sequences (27). Amino acid identities are boxed, and homologies (>85%, Kyte Doolittle scale) are shaded. Alternatively, the degree of hydrophobicity, calculated using the GES scale (22) and 7-residue window, is indicated by a gray-scale where indicated.

RESULTS

Multiple Alignment and Hydrophobicity Profiles of Members of the PTR Family

A search of protein data banks confirmed sequence similarities of the PET1 and PET2 with several other transporters of the recently described PTR or POT family (11,12). Using ENTREZ, 51 nearest neighbor sequences can be identified for PET1; these include closely related homologs from various species, and unidentified open reading frames. The eleven structures analyzed in this report include H^+-dependent oligopeptide transporters of mammals, plants, fungi, and bacteria, a nitrate/chlorate transporter in *Arabidopsis*, and two putative transporters in *E. coli.* (Table 1A). These structures were multiply aligned using PIMA (17) (results not shown, for a similar alignment see ref. 11). According to assignments in SwissProt, the PTR structures contain 1–15 TMDs each. Insertion of substantial gaps required for the multiple alignment of all sequences in the PTR family raises the question whether the number and relative order of aligned TMDs varies among these transporters.

Locations of the TMDs can been deduced from hydropathy profiles, such as those shown in Fig. 2. Peaks above an arbitrarily determined threshold value (for example >1 on the GES scale used in our analysis) suggest the presence of a hydrophobic TMD. The vertical lines sectioning the profiles shown in Fig. 2 indicate the locations of the preassigned TMD boundaries identified in the SwissProt database, except for PTR2_YEAST, for which some of the TMDs appear to have been improperly assigned and were corrected for the present study. In some instances the preassigned TMDs barely reach the threshold of a GES hydropathy value of 1 (Fig. 2). This result demonstrates ambiguities in assigning the location and number of TMDs.

The hydropathy profile of PET1 (Fig. 2A) contains a large hydrophilic loop between putative TMDs 9 and 10, thought to be extracellular (5–7). Following TMD1 is a hydrophobic peak in the profile of PET1 which however does not reach the threshold for assigning a separate TMD. To identify the TMDs that correspond to each other in the various PTR transporters, we divided each sequence into its predicted TMD segments and determined which TMD in the PET1 structure aligns with which TMD in the other PTR members. In Fig. 2, we show representative examples of these profile comparisons. Any lines connecting hydropathy peaks between two profiles (Fig. 2) indicate that these hydrophobic domains (TMDs) are similar, and the respective TMD segments can be aligned, scoring with PAM120 > 20.

The closest homologue of PET1 is the mammalian dipeptide transporter, PET2_RABIT, which has been assigned 13 TMDs in SwissProt. Figure 2 shows that the additional putative TMD in PET2 is inserted into the large extracellular loop located between TMD9 and 10 of PET1, so that TMDs 10–12 of PET1 are now aligned with TMDs 11–13 of PET2 (Fig. 2A). This additional putative TMD10 of PET2 would impose the opposite orientation onto the three C-terminal TMDs for PET2 relative to PET1. Interestingly, TMD10 of PET2 aligns with a relatively hydrophobic domain in the large extracellular loop of PET1 (PAM120 23.5, nSD_{MC} 4.1), suggesting that mutational drift could have resulted in a higher hydrophobicity in this region of PET2 relative to PET1. Since reversal of the orientation of the three C-terminal TMDs in PET1 and PET2 would be expected to disrupt the structure and function of the transporter, it is possible that the assignment of TMD10 in PET2 is in error.

Comparison of the profiles of PET1 and PET2 reveals another ambiguity in the assignment of TMDs in SwissProt: the N-terminal portions of both PET1 and PET2 contain two hydrophobic regions that vary somewhat in the GES scores between PET1 and PET2. Only one of these has been assigned TMD1 in either case, but these assignments differ for PET1 and PET2 (Fig. 2A). Therefore, the location and number of these TMDs need to be verified experimentally.

When comparing more distant structures to that of PET1, fewer TMD segments can be aligned with PAM120 scores > 20 (Fig. 1B–D). Moreover, the assigned TMDs are no longer in the same nominal location of the primary structure, e.g., TMD8 of PET1 aligns with TMD7 of PTR2_YEAST (Fig. 2B). In the case of YHIP_ECOLI, TMDs 12–15 were inferred from the hydropathy profile, but none of these show significant similarity to C-terminal PET1 domains. The large insertion between TMDs 9 and 10 observed in PET1 and PET2 is absent in these structures of more ancient phyla which therefore tend to be considerably shorter. On the other hand, YHIP_ECOLI and DTPT_LACLA, contain two additional putative TMDs in the center of their primary structures (TMDs 7 and 8, and TMDs 6 and 7, respectively; Fig. 2C, D) which are absent in other PTR structures (except YJDL_ECOLI). The hydropathy scores are high (GES scale ≥ 2), strongly supporting the assignment of both TMDs. These additional two TMDs appear to represent inserts in these sequences of more ancient phyla that cannot be readily aligned with each other, nor with any other TMD segment in the PTR family. Further, a BLAST search of protein databases (16) identified only a few possibly related sequences. Among these, one alignment stands out, i.e., TMD6 of YJDL_ECOLI with an 18-residue sequence belonging to a hypothetical amino acid permease, C11D3.08C, of *S.Pombe* (identity 50%, score 59, similar to the PAM120 scale used here). This result suggests possible relationships of the two inserted TMDs in ancient PTR structures with TMD segments of other transporters.

Because of the insertion of TMDs, C-terminal TMDs of transporters containing these inserts are shifted by two TMDs relative to those of PET1 (Fig. 2C, D). This shift of apparently homologous TMDs by the insertion of two TMDs allows for the same orientation of the conserved TMDs in crossing the membrane. This result suggests that the function of the aligned TMDs has been conserved even though the order of TMDs has changed. Further, it raises the question as to the function of these two inserted TMDs in some of the PTR structures. Because of the structural data on trans-membrane topology available for DTPT_LACLA (15), we will subsequently emphasize alignments between PET1 or PET2 and DTPT for each type of analysis performed in this study.

Sequence Analysis with 2D-Matrix Plots

Alignments between TMDs suggested in Fig. 2 do not account for the possibility that TMD segments might be related to each other without any regard to the order in which they appear in the primary structure. We therefore used a 2D-matrix approach which permits the analysis of relationships among all TMD segments, independent of their order (Fig. 1). In the comparison of PET1 with itself (not shown), a contiguous diagonal line was observed as expected. However, a number of parallel lines suggested additional relationships among non-corresponding TMDs (e.g., between TMD4 and TMD6, PAM120 score 29.4, nSD_{MC} 3.7). In comparing PET1_HUMAN with PET2_RABIT, again a strong diagonal line of identity is observed, however, with several gaps and shifts indicating the presence of variable domains, insertions and deletions (Fig. 1A). Both PET1 and PET2 contain a large loop between TMD9 and TMD10, but these loops are relatively dissimilar, except for several domains at the junctions and in the middle of the loop. At the C-terminal junction, the line of identity includes the alignment of the putative TMD10 in PET2 with a loop domain of PET1 (discussed above). The function of these conserved loop regions should be probed with the use of mutagenesis studies. Further, multiple parallel diagonal lines indicate similarities between non-corresponding TMDs. As observed with PET1 compared against itself, TMD4 of PET1 again aligns with TMD6 of PET2 (PAM120 score 28.4, nSD_{MC} 3.8), and vice versa. The scores approximate or exceed the threshold values indicative of possible homology in our database. The finding that this TMD4/TMD6 alignment occurred in both PET1 and PET2 suggests the possibility that these two TMDs might represent intragenic repeats.

Comparing PET1 with its more distantly related transporter homologs, PTR2_YEAST and DTPT_LACLA (Fig. 1B, C), reveals an increasingly scattered pattern of diagonal lines, representing similarities among TMD segments that are frequently interrupted close to their boundaries. Whereas in the 2D-plot of PET1 versus PTR2_YEAST, the main diagonal line can still be followed with relative ease, this is no longer the case when aligning PET1 with DTPT_LACLA. Moreover, multiple parallel lines again suggest the possible presence of repeat structures. These findings suggest a modular organization of the primary structures, where each module consists of a single or a few TMD segments that can be rearranged or duplicated during evolution.

Alignments of Individual TMD Segments in the PTR Family

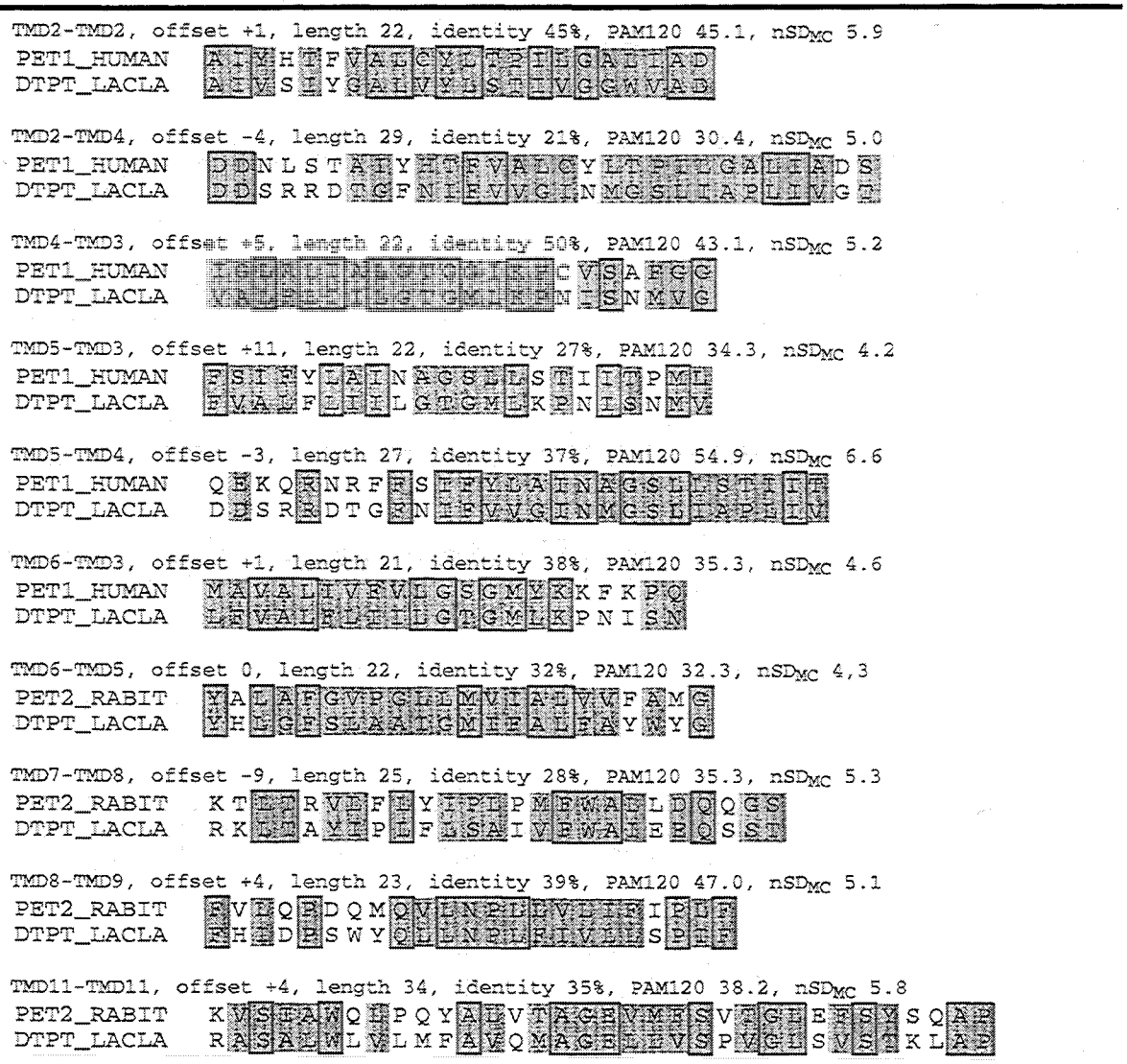
The 2D-matrix plots shown in Fig. 1 indicate that relatively high scoring alignments can be found among TMD segments regardless of their order in the primary sequence. To assess these alignments quantitatively, we determined the best alignment windows for each of the TMD segments of PET1_HUMAN or PET2_RABIT (ignoring the putative TMD10 of PET2) versus each TMD segment of all other members of the PTR family and calculated the corresponding PAM120 and nSD_{MC} values. Even though PET1 and PET2 are quite similar to each other, we included both with the comparison between these mammalian transporters and those of more ancient phyla. The use of more than one sequence of closely related homologs enhances the chance of finding statistically significant alignments in any given TMD segment. As an example of the results obtained by this type of analysis, alignments of TMD segments of PET1 or PET2 with TMD segments of DTPT_LACLA are shown in Table 2. Since our negative control, i.e., comparisons of transporter TMDs with those of G protein coupled receptors, suggested PAM120 threshold values of ~30 and ~54 for indicating possible and probable homology, respectively, only TMD alignments scoring above 30 are shown in Table 2. Most of the TMD alignments in Table 2 are consistent with the alignments identified in the hydropathy profiles in Fig. 1; however, in several instances, the same TMD in PET1 or PET2 is aligned with two different TMDs in DTPT (TMD2/TMD2 and TMD2/TMD4, TMD4/TMD3 and TMD5/TMD3, TMD6/TMD3 and TMD6/TMD5). These alignments would suggest the possible presence of internal repeats. Previous studies indicated that TMDs 1–6 and TMDs 7–12 appear to represent intragenic repeat in some transporters, such as those of the mdr1 group (28). However, alignments of individual TMD segments yielded no compelling evidence for intragenic duplications of the first and second portion of any of the PTR transporters, even though our method readily detected the mdr1 intragenic duplication by matching TMD1 with TMD7, TMD2 with TMD8, etc. (data not shown). Rather, it appears possible that the primary structure of the PTR transporters may be modular, consisting of TMD segments as independent units that can be rearranged, inserted, deleted, or duplicated.

Multiple Alignments of Corresponding TMDs Among Eight PTR Structures

Successful alignments of corresponding TMD segments within the PTR family permits one to determine the most highly conserved domains. Multiple alignments of the 12 TMDs of PET1 with the corresponding TMDs in seven other transporters of the PTR group are shown in Fig. 3. It should be noted that the aligned TMDs are not always from TMDs in exactly the same *nominal* location of the primary structure, but rather represent those that yield the best alignments (as determined in Fig. 1). Multiple alignment of TMD1 yielded the lowest PAM120 scores, indicating that this putative TMD is rather poorly conserved. Highly conserved residues are boxed in Fig. 3. Moreover, to illustrate the hydropathy distribution throughout these TMDs, the alignments are shaded by a gray-scale indicating the degree of hydrophobicity over a 7-residue window.

All TMD segments contain highly conserved residues or regions, interspersed with less well conserved regions. Several residue positions are conserved in most structures, particularly in two highly conserved domains at the end of TMD2 (and extending into the loop, not shown) and the first portion of TMD5 of PET1, representing the two consensus motifs described earlier for the smaller POT family (12). However,

Table 2. Alignments Between TMD Segments of PET1_HUMAN or PET2_RABIT And the Lactococcus Peptide Transporter DTPT_LACLA



Note: The highest scoring alignments for each comparison of TMD segments in any location of the primary structures were determined for either PET1 or PET2 versus DTPT, and those scoring with PAM120 > 30 are shown. Note the alignments of TMDs 2, 3, 4, 5, and 6 each with two different TMDs, suggesting the possible presence of repeat structures. The offset gives the shift of the optimal alignment (in number of residues) relative to an alignment that coincides with the predefined TMD boundaries (having an offset of 0). Minimum length of the alignment window is 21 residues. The nSD_{MC} value provides the number of standard deviations separating the alignment shown from the mean score of scrambled sequences (1000 serial randomizations of the second sequence).

several additional domains are also highly conserved, e.g., in TMD 4, 8, 10, and 11. The Gly-rich domain in TMD4, GLALIALGTGGI in PET1, is most highly conserved, suggesting that it represents a region of functional significance. Further, the multiple alignment reveals several conserved proline residues, important to TMD structure because of their ability to induce bends in the α-helix; highly conserved proline residues are found in TMDs 1, 2, 7, and 8.

Since all structures in Fig. 3 represent H^+-dependent transporters, we searched for conserved charged residues that could play a role in proton translocation. Only one charged residue is fully conserved, namely Glu in TMD10. This Glu residue is located towards the middle of the TMD, suggesting a possible role in charge translocation. Other highly conserved positions include Gly residues in several TMDs, capable of introducing bends into the α-helical TMD structure. These conserved positions suggest a common structure for portions of these transporters proteins.

Recently, introduction of point mutations into PET1 has provided the first evidence of residues involved in peptide transporter function. Thus, replacement of His-57 in TMD2 and His-121 in TMD4 of rat PET1 with glutamine abolished peptide transport in either case (29). Both His residues, however, have an abundance of only 38% in the consensus sequences for TMD2 and TMD4 (Fig. 3). This result could suggest that these His residues are involved in substrate recognition, rather

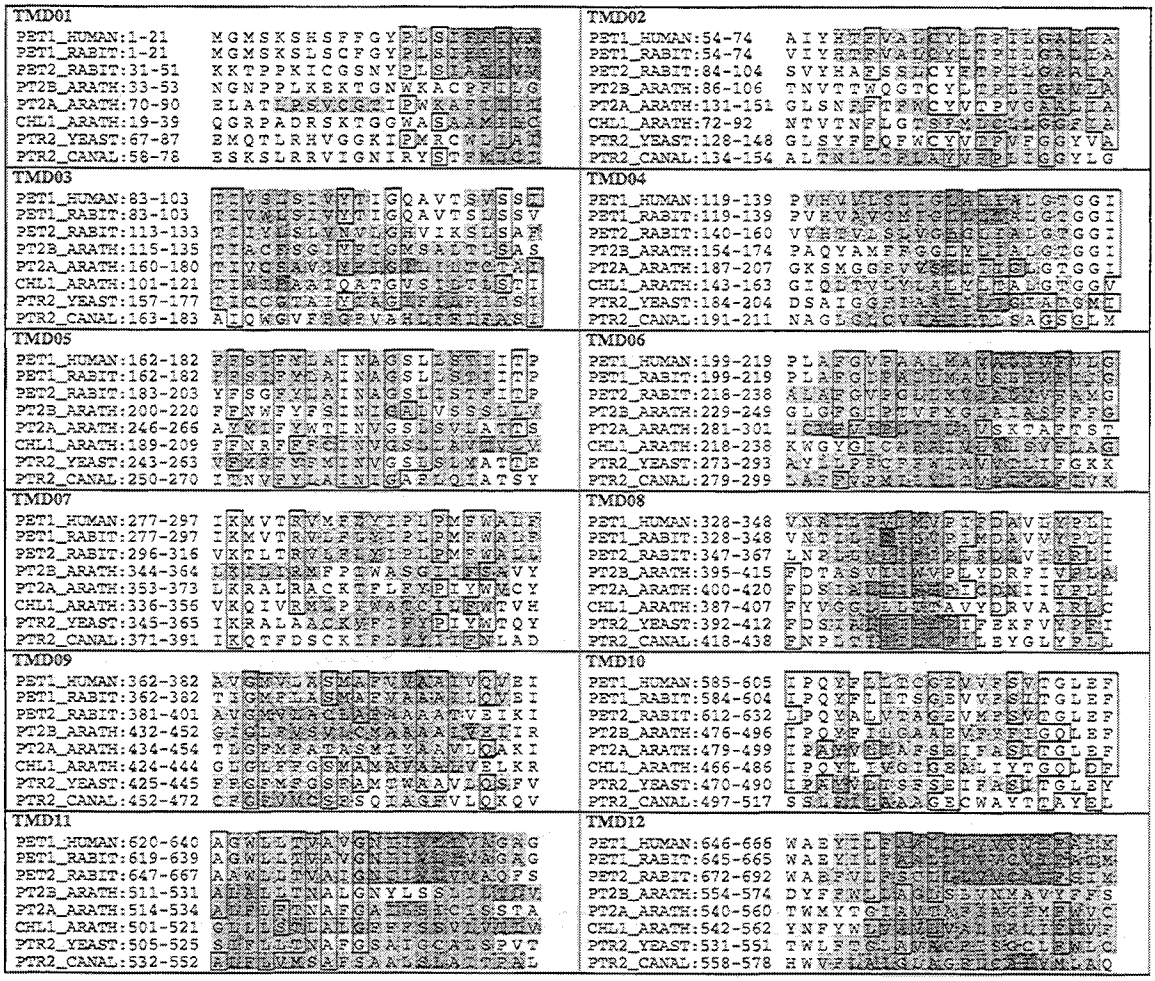


Fig. 3. Multiple alignments of 12 TMDs derived from eight eukaryotic PTRs. The TMDs (21-residues) were aligned using PIMA (17). The TMDs are ordered according to the 12 TMDs assigned to PET1_HUMAN. To account for changes in the order of TMDs during evolution, alignments between PET1 and other PTRs did not have to occur in the same nominal position in their respective structures. Rather alignments of corresponding TMDs were guided by the analysis illustrated in Figs. 1 and 2. For example, in the alignment between PET1 and PET2, we ignored the additional TMD10 of PET2, as assigned in the SwissProt database. Identical residues present at the same alignment position in 5 or more sequences are boxed. To illustrate the distribution of hydrophobic residues in these putative TMDs, the hydrophobicity average was calculated at each position for a 7-residue window, as described in Fig. 1. The hydrophobicity (GES) (20) is represented by an arbitrary gray-scale, where the darker gray shades indicate higher hydrophobicity.

than H^+ translocation which might be expected to be more highly conserved among all PTR transporters. In another study, replacement of a Tyr in TMD5 (88% conserved, Fig. 3) also suppressed transport activity of PET1, whereas Ala replacement of the conserved Arg in TMD7 (75% conserved, Fig. 3) failed to affect transport (30). The alignments presented in Fig. 3 suggest further residues that should be evaluated by site-directed mutagenesis for a role in PET1 function.

Gray-shading in Fig. 3 illustrates the distribution of hydrophobicity in each of the putative TMDs, for all structures analyzed. As seen in Fig. 2, several of the assigned TMDs are less hydrophobic, in particular TMD1. As discussed above, it is doubtful that this region indeed represents a separate TMD. Further, the shaded areas reveal an uneven distribution of the most hydrophobic domains within the TMDs. Taken together with the hydrophobic moment common to amphipathic TMDs, the distribution of hydrophobicity can serve to suggest location and relative orientation of TMDs in the tertiary structure of multiple TMD membrane proteins. The reliability of such structure predictions may be enhanced by considering the hydrophobic profiles as a function of evolution among related structures, as shown in Fig. 3.

Linking TMD Segments for Evaluation of Similarity Among PTR Members

Whereas alignments of individual TMD segments between members of the PTR family suggest possible homology, assessment of homology on the basis of rather short sequences may be problematic. For a more definitive assessment of homology, we linked alignments between TMD segments of two sequences, allowing for the introduction of gaps between the linked TMDs, and determined the nSD_{MC} values of the linked sequences. For this type of analysis, $nSD_{MC} > 9$ has been

proposed as a threshold value for suggesting probable homology (26); however, the obtained linked nSD_{MC} values depend on the conditions of the alignment algorithms, such as minimum length and PAM score of fragments included with the linked sequence. Therefore, we determined the nSD_{MC} values for linked alignments among each of the PTR proteins to establish benchmark values for established homologs in our analysis. These benchmark values subsequently served in evaluating relationships to more distant transporter groups.

For the sequences contained in the PTR family, our results suggest a modular structure in which the order of TMD segments may have changed in the course of evolution. As a result, it is not always clear which TMD segments, and in what order, should be linked to assess homology. This problem increases with increasing evolutionary distance, and therefore, its resolution is critical for defining distant relationships with the use of alignments relying on individual TMD segments. We have implemented a novel procedure by which two structures with multiple TMDs are compared using a 2D-grid similar to those shown in Fig. 1. Starting with TMD1, a search algorithm will determine which TMD segments are to be linked so that the highest possible PAM120 score for the linked sequence pair is attained. The search progresses always to the next TMD with a higher number, so that any repeats are excluded from this analysis. This automated approach improved the nSD_{MC} score for comparing PET1_HUMAN with DTPT_LACLA from below 9, obtained by aligning only TMDs *nominally* in the same location, to an nSD_{MC} value of 16.1. This latter score indicates that homology between these two structures is highly probable, consistent with previous reports (11,12). The high score obtained with the search procedure for linking TMD segments in an optimal order is consistent with the highest scoring alignments between PET1 or PET2 and DTPT_LACLA shown in Table 2, and the lines connecting individual TMDs between PET1 and DTPT_LACLA, shown in Fig. 2D. Other possible alignments of the same TMD (Table 2) are ignored; yet, these alternative alignments of TMD segments in different locations suggest possible rearrangements or intragenic duplications that may have occurred during evolution of these transporters.

Using linked TMD sequences obtained by this procedure, we calculated the overall similarity of PET1 to all other members of the PTR family (Table 3A). Our search algorithm linked between 9–12 TMD segments for calculating the overall nSD_{MC} value. In each case, the nSD_{MC} value exceeded 15; therefore, this analysis of linked TMD segments confirms the high similarity and probable homology among these structures, by taking into account changes in the order of TMD segments in the primary structure. To determine the range of nSD_{MC} values among all members of the PTR family, we aligned each PTR structure with each other PTR structure in a similar fashion (data not shown). A number of linked alignments scored with nSD_{MC} values of 11–12, and the lowest score was observed for PTR2_CANAL versus YJDL_ECOLI (nSD_{MC} 10.9). Since all members of the PTR family can be assumed to be homologous, a threshold value of $nSD_{MC} > 11$ can be considered as an indicator of possible homology in this analysis.

Alignments of Individual TMD Segments of PTR Sequences with Those of Other Transporter Groups

We have established a database of 29 additional groups containing 91 selected transporters with 9–15 proposed TMDs (Table 1B). Using the entire primary structure, the PTR members cannot be readily aligned with other transporter families, because their sequences are dissimilar (11). To test any relationships of the PTR family to any of the other groups in individual TMD segments, we aligned each TMD segment of the PTR members with each TMD segment of all other transporters and calculated the PAM120 and nSD_{MC} values. On the basis of the proposed modular organization of these proteins described above, we searched for alignments between individual TMD segments regardless of the order as they appear in the primary structure.

Table 4A summarizes the highest scoring alignments of an individual TMD segment derived from any member of the PTR family with that from other transporter groups. A number of alignments yielded scores between PTR-TMD segments and those of the other groups that approach our threshold values suggestive of probable homology (PAM120 54 and nSD_{MC} 6). Some of these alignments are in the same nominal TMD location (e.g., TMD8 versus TMD8), whereas others are out of order (TMD8 versus TMD12) (Table 4A). Further, relative alignment positions among the TMD segments deviated in most cases from the preassigned TMD boundaries, indicated by the offsets different from 0. Therefore, in comparing these very distantly, if at all, related TMD segments, one must ignore the preassigned TMD boundaries to find the optimal alignment positions. This result further suggests that for these alignments, the biased amino acid composition of the TMDs may not be a major factor because any shift in the alignment relative to preassigned TMD boundaries juxtaposes TMD and loop residues.

Table 4B shows the actual TMD alignments (with PAM120 > 40) between PET1 or PET2 only and transporters in the other groups. The highest scoring alignments approached nSD_{MC} values of 6. These results demonstrate similarities among TMD segments of the PTR family and the other groups suggestive of possible homology in at least a portion of the primary structure.

Linking TMD Segments for Comparing PTR Members to Other Transporters

To evaluate the relationships of PTR sequences to other, rather distant transporter families, we used the linked sequence alignments, based on our search algorithm to establish the optimal order of linking TMD segments, and determined the linked sequences scoring with the highest nSD_{MC}. For this approach, we have established a benchmark nSD_{MC} value of 11 for suggesting possible homology in our analysis. The linked sequence results obtained by aligning PET1 with all other transporters are provided in Table 3B. Several alignments of linked sequences between PET1 with sequences from the other groups attained nSD_{MC} scores > 11, the threshold value suggestive of possible homology. When testing each member of the PTR family against the other groups, three alignments were found with linked nSD_{MC} scores > 12 (12.1: PT2A_ARATH versus ACRB_ECOLI; 12.2: PET2_RABIT versus NANU_RABIT; 12.3: PET2_RABIT versus TRK1_SACUV). These results further suggest relationships between the PTR family and other transporter groups. On the other hand, aligning YHIP_ECOLI with HXT1_YEAST yielded a nSD_{MC} value of only 6.1, the lowest value attained in this type of analysis within our database. This result indicates that the procedure involving linked TMD segments can discriminate between structures of relatively low

Table 3. A. Alignment of Linked TMD Segments of PET1_HUMAN with Other Members of the PTR Group

	No. of segments	length	nSD_{MC}
PET1_HUMAN - CHL1_ARATH	12	305	18.1
PET1_HUMAN - DTPT_LACLA	9	208	16.1
PET1_HUMAN - PET1_HUMAN	12	532	58.6
PET1_HUMAN - PET2_RABIT	12	443	42.1
PET1_HUMAN - PT2A_ARATH	12	299	19.9
PET1_HUMAN - PT2B_ARATH	11	312	22.2
PET1_HUMAN - PTR2_CANAL	9	232	15.1
PET1_HUMAN - PTR2_YEAST	11	290	18.6
PET1_HUMAN - YHIP_ECOLI	11	255	15.5
PET1_HUMAN - YJDL_ECOLI	11	259	16.2

B. Alignments of Linked TMD Segments of PET1_HUMAN with All Other Groups

	No. of segments	length	nSD_{MC}
PET1_HUMAN - acrd, ACRB_ECOLI	11	249	10.7
PET1_HUMAN - alp1, VAL1_YEAST	7	155	9.8
PET1_HUMAN - arae, ARAE_ECOLI	9	196	9.9
PET1_HUMAN - arop, PHEP_ECOLI	10	221	11.0
PET1_HUMAN - b3at, B3A2_HUMAN	8	171	10.5
PET1_HUMAN - bmrp, BMRP_CANAL	8	175	9.8
PET1_HUMAN - cftr, CFTR_BOVIN	8	177	9.8
PET1_HUMAN - cita, CIT_KLEPN	10	223	10.8
PET1_HUMAN - dal4, DAL4_YEAST	9	194	9.8
PET1_HUMAN - dcta, DCTA_RHIME	9	203	10.2
PET1_HUMAN - gltp, GLTP_BACSU	7	161	9.3
PET1_HUMAN - gtr1, GTR1_BOVIN	8	187	10.8
PET1_HUMAN - hxt1, HXT1_YEAST	5	111	8.5
PET1_HUMAN - itr1, GTR1_LEIDO	8	173	9.7
PET1_HUMAN - lacy, LACY_KLEPN	9	190	9.2
PET1_HUMAN - mdr1, MDR1_MOUSE	9	201	10.0
PET1_HUMAN - melb, MELB_KLEPN	9	201	10.8
PET1_HUMAN - nah1, NAH1_HUMAN	10	229	11.3
PET1_HUMAN - nanu, NANU_RABIT	8	172	10.8
PET1_HUMAN - ntg1, NTGL_MOUSE	5	109	6.9
PET1_HUMAN - nupg, NUPG_ECOLI	9	200	10.6
PET1_HUMAN - putp, PUTP_ECOLI	8	175	10.5
PET1_HUMAN - tcr1, TCR4_ECOLI	8	177	9.3
PET1_HUMAN - tcr2, TCRB_BACSU	9	204	10.5
PET1_HUMAN - th11, TH2A_TRYBB	10	219	11.6
PET1_HUMAN - trk1, TRK1_SACUV	10	224	11.6
PET1_HUMAN - vmt1, VMT2_HUMAN	7	163	10.3
PET1_HUMAN - ycs7, YCS7_YEAST	11	249	11.0
PET1_HUMAN - yim0, YIM0_YEAST	7	152	9.0

Note: The Table provides the number of linked TMD segments, the number of residues in the linked sequences, and the nSD_{MC} value for the linked sequence (1000 serial randomizations).

and high similarity even if they share the same multiple-TMD topology.

By averaging the linked nSD_{MC} scores for the best alignments of each of the PTR members with the other groups, we can determine a rank order of similarity. Highest similarity of PTR was observed with (in decreasing order) the nucleoside transporter group nanu (*mean* nSD_{MC} 10.8), the group of acriflavin resistance proteins, acrd (10.7), the glucose transporter group th11 (10.7), the proline transporter group putp (10.6), the multiple drug resistance group, mdr1 (10.6), and the cystic fibrosis transmembrane regulator group, cftr (10.5). Least similar was the glutamate transporter group, gltp (*mean* nSD_{MC} 8.4). Alignments among individual TMD segments had already identified similarities between PTR transporters and the same transporter groups identified by the linked sequence analysis (compare Tables 3 and 4). Therefore, results obtained from alignments of both individual TMD segments and linked sequences support possible relationships among these transporter groups.

In comparing the members of the PTR family with each other using 2D-matrix plots (Fig. 1A–C), we have seen that the expected diagonal line of identity becomes more fragmented with increasing evolutionary distance. We also applied this 2D-

Table 4. Alignments Between Individual TMD Segments of the PTR Family with Those of All Other Transporter Groups

A. Highest scoring TMD segments with with PAM120 ≥ 46 or nSD_{MC} ≥ 5.5)

PTR family	other groups	TMD/TMD	offset	PAM120	nSD_{MC}
YHIP_ECOLI	mdr1, MDR1_CRIGR	TMD8/TMD8	-5	53.9	5.6
YHIP_ECOLI	tcr2, TCR2_BACSU	TMD8/TMD12	±0	50.0	5.5
CHL1_ARATH	acrd, ACRD_ECOLI	TMD9/TMD11	+4	49.0	5.9
PT2A_ARATH	hxt1, HXT2_YEAST	TMD4/TMD4	-11	47.0	5.8
PT2A_ARATH	acrd, ACRD_ECOLI	TMD9/TMD11	+1	47.0	5.4
PTR2_YEAST	bmrp, CYHR_CANMA	TMD2/TMD3	-4	47.0	5.3
DTPT_LACLA	itr1, ITR1_YEAST	TMD4/TMD4	+1	47.0	4.7
PTR2_CANAL	trk1, TRK2_YEAST	TMD4/TMD4	-1	46.1	6.0
PET2_RABIT	dal4, FUR4_YEAST	TMD2/TMD12	+10	46.1	5.3
PET2_RABIT	itr1, GTR1_LEIDO	TMD6/TMD9	+1	46.1	5.1
YJDL_ECOLI	gtr1, GTR7_RAT	TMD2/TMD2	-1	44.1	5.9
CHL1_ARATH	arop, PHEP_ECOLI	TMD8/TMD7	-6	44.1	5.8
PTR2_YEAST	acrd, YHIV_ECOLI	TMD4/TMD5	+11	42.1	5.6
DTPT_LACLA	acrd, YHIV_ECOLI	TMD11/TMD10	-13	42.1	5.5
PT2A_ARATH	hxt1, HXT2_YEAST	TMD3/TMD3	+1	42.1	5.5
PT2A_ARATH	nah1, NAH1_HUMAN	TMD10/TMD2	±0	41.2	6.1
PTR2_YEAST	ntg1, NTS1_RAT	TMD6/TMD4	+8	41.2	5.9
PET2_RABIT	mdr1, MDR1_MOUSE	TMD5/TMD2	-5	41.2	5.8

B. Examples of alignments between TMD segments of PET1_HUMAN or PET2_RABIT with those of all other transporter groups (nSD_{MC} ≥ 5)

acrd
TMD12-TMD7, offset -6, length 28, identity 21%, PAM120 44.1, nSD_{MC} 5.7
PET2_RABIT
VAL1_YEAST

bmrp
TMD6-TMD2, offset -9, length 21, identity 33%, PAM120 44.1, nSD_{MC} 5.1
PET1_HUMAN
CYHR_CANMA

dal4
TMD2-TMD12, offset +6, length 21, identity 29%, PAM120 46.1, nSD_{MC} 5.4
PET2_RABIT Y H A F S S L G Y F T P T L G A A T A D S
FUR4_YEAST F S S Y T A L G Y F F P V P G G E V N N I

itr1
TMD6-TMD9, offset +1, length 21, identity 43%, PAM120 46.1, nSD_{MC} 5.0
PET2_RABIT
GTR1_LEIDO

mdr1
TMD5-TMD2, offset -6, length 29, identity 34%, PAM120 41.2, nSD_{MC} 5.8
PET2_RABIT
MDR1_MOUSE

Note: The offset indicates any shifts of the optimal alignment with respect to the preassigned TMD boundaries.

matrix analysis to comparisons of PTR sequences with those of the other groups. In the linked sequence analysis using PET1 compared against all other transporter families, the trypanosome glucose transporter TH2A_TRYBB scored best with an nSD_{MC} value of 11.6 (Table 4B). Therefore, we analyzed an alignment of PET1_HUMAN with TH2A_TRYBB using a 2D-matrix plot, to illustrate the resulting plots comparing very distant structures (Fig. 1D). Numerous alignments represented by short diagonal lines in the 2D-matrix plot shown in Fig. 1D indicate regions of sequence similarity scoring with PAM120 > 20, but no clear diagonal line of identity is detectable for the entire sequence. Sequential regions of relative similarity can be observed in TMDs 5, 6, and 7 of TH2A_TRYBB (x-axis), which might have contributed to the rather high nSD_{MC} score of the linked structures. However, the fragmented nature of the alignment regions shown in Fig. 1D highlights the difficulty in establishing homology among these structures.

DISCUSSION

In this study we have analyzed protein sequences of the PTR family to assess their TMD topology, conserved regions of possible functional relevance, and relationships to other transporter families. For this purpose, we have developed new

approaches to sequence analysis designed to account for the recurrent loop-TMD structure, by focusing on pairwise alignments of individual TMD segments, in addition to aligning the entire primary structure. The results suggest that these proteins are modular, with each module consisting of one or more TMD segments that can be rearranged by insertions or deletion, possible duplications, or other evolutionary mechanisms that can affect the order of TMD segments in the primary structure.

Comparing hydropathy profiles along with alignments of individual TMD segments enabled us to align the corresponding TMDs for the members of the PTR family. This approach confirmed the presence of changes in the order of TMDs between PET1 and transporters of more ancient phyla. Moreover, our analysis raised questions about the TMD assignments in the SwissProt database, determined on the basis of hydropathy alone. Multiple alignments of corresponding TMD segments demonstrate variable degrees of conservation for different TMDs and revealed the presence of conserved regions. The most highly conserved residues are primary targets for site-directed mutagenesis experiments designed to probe the functional domains of the transporters.

Alignments of TMD segments regardless of their order in the primary structure revealed the possible presence of intragenic repeats of individual TMD segments. However, the PAM120 and nSD_{MC} scores of such relatively short sequences, while indicating a surprising degree of similarity, are insufficient to provide strong support for the presence of intragenic repeats. Alternatively, the observed similarity could have arisen from convergent evolution by mutational drift, reflecting the functional constraints imposed by the amphipathic α-helical TMDs. A larger database and extensive statistical analysis will be required to resolve this question.

We also attempted to test relationships among the PTR family with other transporters consisting of a multiple-TMD topology. By aligning individual TMD segments, and using a novel search algorithm to link several aligned TMD segments, we have identified similarities that suggest possible relationships among the PTR family and other transporter families. However, because of the fragmented nature of these alignments, the obtained scores are insufficient to support firmly a finding of homology, and a much larger database may be required. Further, our hypothesis that the transporter sequences consist of gene modules that can rearrange during evolution must be resolved before the question of homology between these transporters can be answered. This study provides possible strategies that may be generally applicable to the sequence analysis of membrane proteins consisting of multiple TMD segments.

ACKNOWLEDGMENTS

This work was supported by Public Health Services research grants GM37188 and GM43102 from the National Institute of Health, General Medical Sciences.

REFERENCES

1. G. L. Amidon and H. J. Lee. *Annu. Rev. Pharmacol. Toxicol.* **34**:321–341 (1994).
2. W. Kramer, F. Girbig, U. Gutjahr, H.-W. Kleemann, I. Leipe, H. Urbach, and A. Wagner. *Biochim. et Biophys. Acta* **1027**:25–30 (1990).
3. A. H. Dantzig, and L. Bergin. *Biochim. et Biophys. Acta* **1027**:211–217 (1990).
4. S.-I. Matsumoto, H. Saito, and K.-I. Inui. *J. Pharmacol. Exp. Ther.* **270**:498–504 (1994).
5. Y. J. Fei, Y. Kanal, S. Nussberger, V. Ganapathy, F. H. Leibach, M. F. Romero, S. K. Singh, W. F. Boron, and M. A. Hediger. *Nature* **368**:563–566 (1994).
6. R. Liang, Y.-J. Fei, P. D. Prasad, S. Ramamoorthy, H. Han, T. L. Yang-Feng, M. A. Hediger, V. Ganapathy, and F. H. Leibach. *J. Biol. Chem.* **240**:6456–6463 (1995).
7. M. Boll, D. Markovich, W.-M. Weber, H. Korte, H. Daniel, and H. Murer. *Eur. J. Physiol.* **429**:146–149 (1994).
8. M. Boll, M. Herget, W. M. Weber, D. Markovich, J. Biber, W. Clauss, H. Murer, and H. Daniel. *Proc. Natl. Acad. Sci. USA* **93**:294–289 (1996).
9. U. Wenzel, I. Gevert, H. Weintraut, W.-M. Weber, W. Claub, and H. Daniel. *J. Pharmacol. Exp. Ther.* **277**:831–839 (1996).
10. K.-M. Yu-Covitz, G. L. Amidon, and W. Sadée. *Pharm. Res.* **13**:1631–1634 (1996).
11. H.-Y. Steiner, F. Naider, and J. M. Becker. *Mol. Microbiol.* **16**:825–834 (1995).
12. I. T. Paulsen, and R. A. Skurray. *Trends Biochem. Sci.* **19**:404 (1994).
13. W. Sadée, V. Drübbisch, and G. L. Amidon. *Pharm. Res.* **12**:1823–1837 (1995).
14. P. J. F. Henderson. *Curr. Opin. Cell Biol* **5**:708–721 (1993).
15. A. Hagting, E. R. S. Kunji, K. J. Leenhouts, B. Poolman, and W. N. Konings. *J. Biol. Chem.* **269**:11391–11399 (1994).
16. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. *J. Mol. Biol.* **21**:403–410 (1990).
17. R. F. Smith, and T. F. Smith. *Prot. Engin.* **5**:35–41 (1992).
18. S. F. Altschul. *J. Molec. Biol.* **219**:555–565 (1991).
19. M. Gribskov, and J. Devereux. *Sequence Analysis Primer*, W. H. Freeman & Co., N. Y., 1992.
20. M. G. Claros, and G. von Heijne. *CABIOS* **10**:685–686 (1994).
21. G. von Heijne. *J. Mol. Biol.* **225**:487–494 (1992).
22. D. M. Engleman, T. A. Steitz, and A. Goldman. *Ann. Rev. Biophys. Chem.* **15**:321–353 (1986).
23. B. Persson, and P. Argos. *J. Mol. Biol.* **237**:182–192 (1994).
24. D. C. Rees, L. DeAntonio, and D. Eisenberg. *Science* **245**:510–513 (1992).
25. S. Karlin, and V. Brendel. *Science* **257**:39–49 (1992).
26. M. D. Marger, M. H. Jr. Saier. *TIBS* **18**:13–20 (1993).
27. R. P. Riek, M. D. Handschumacher, S. S. Sung, M. Tan, M. J. Glynias, M. D. Schluchter, J. Novotny, and R. M. Graham. *J. Theor. Biol.* **172**:245–258 (1995).
28. J. K. Griffith, M. E. Baker, D. A. Rouch, M. G. P. Page, R. A. Skurray, I. T. Paulsen, K. F. Chater, S. A. Baldwin, and P. J. F. Henderson. *Curr. Opin. Cell Biol* **4**:684–695 (1992).
29. T. Terada, H. Saito, M. Mukai, and K.-I. Inui. *FEBS Lett.* **394**:196–200 (1996).
30. A. K. Yeung, D. Ann, M. B. Bolger, H. von Grafenstein, S. Hamm-Alvarez, W. Shen, C. T. Okamoto, K. J. Kim, S. K. Basu, I. S. Haworth, and V. H. L. Lee. *Pharm.Res.* **13**:S-243 (1996).
31. Y. F. Tsay, J. I. Schroeder, K. A. Feldmann, and N. M. Crawford. *Cell* **72**:705–713 (1993).
32. H. Y. Steiner, W. Song, L. Zhang, F. Naider, J. M. Becker, and G. Stacey. *Plant Cell* **6**:1289–1299 (1994).
33. W. B. Frommer, S. Hummel, and D. Rentsch. *Febs Let.* **347**:185–189 (1994).
34. M. A. Basrai, M. A. Lubkowitz, J. R. Perry, D. Miller, E. Krainer, F. Naider, and J. M. Becker. *Microbiol.* **141**:1147–1156 (1995).
35. J. R. Perry, M. A. Basrai, H.-Y. Steiner, F. Naider, and J. M. Becker. *Mol. Cell. Biol.* **14**:104–115 (1994).
36. H. J. Sofia, V. Burland, D. L. Daniels, G. Plunkett III, and F. R. Blattner. *Nucleic Acids Res.* **22**:2576–2586 (1994).
37. V. Burland, G. Plunkett III, H. J. Sofia, D. L. Daniels, and F. R. Blattner. *Nucl. Acids Res.* **23**:2105–2119 (1995).