Comparative structural studies of peptide transporters and their roles in physiological processes.

Shubham Bhojane

A dissertation submitted for the partial fulfilment of
BS-MS dual degree in Science

Indian Institute of Science Education and Research Mohali

April 2019

Certificate of Examination

This to certify that the dissertation titled **"Comparative structural studies of peptide transporters and their roles in physiological processes"** submitted by **Mr. Shubham Bhojane** (Reg. No. MS14020) for the partial fulfilment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Debrina Jana Dr. Sanchita Sengupta Dr. Monika Sharma

(Supervisor)

Dated: April --, 2019

Declaration

The work presented in the dissertation has been carried by me under the guidance of Dr.Monika Sharma at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, or diploma, or a fellowship to any other University or Institute. Whenever contribution of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in bibliography.

Shubham Bhojane

(Candidate)

Dated: April --, 2019

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge

Dr. Monika Sharma

(Supervisor)

Acknowledgement

I would like to take the opportunity to thank to all those who have helped and supported me directly or indirectly throughout my BS-MS time. Without them this thesis would not have been possible.

First and foremost I would like to express my deepest sense of gratitude to Dr. Monika Sharma ma'am for allowing me to complete my final year project under her guidance. It has been such a pleasure to work with her. She always took time for me and gave me invaluable advises.

I also wants to thank my all friends from MS14 batch with particular mention of Ajay, Apoorv, Indrajeet, Neetish, Ravi, Rohit, Suresh, Tinku, Varun, Vishal, Yogendra, for their moral and emotional support during my good and bad time.

And last but not least, I am most thankful to my parents Mayadevi Dnyanobarao bhojane, my sister Dipali & brother Sagar who gave me love, support and strength to get here.

List of figures

List of Tables

Abstract

The POT family of membrane transporter proteins shows remarkable feature of diverse substrate promiscuity. This promiscuity has led to development of peptide based pro-drugs that us PepT1 and PepT2, mammalian homologues, to improve oral drug delivery[1]. Here I tried to classify 46 POT family transporters on the basis of transmembrane α-helical spanners and to establish evolutionary relationship between POT family members and by sequence alignment technique looked into residues conserved in POT family members.

Contents

Chapter 1: Introduction

1.1 Cell membrane its function and composition

Both prokaryotic and eukaryotic cells are enclosed in a semipermeable membrane also known as the cell membrane or the plasma membrane. cell membrane has four important functions function number one is, it creates a protective barrier between the outside and the inside environment of the cell it basically prevents toxins and other pathogenic agents from entering the cell and it prevents molecules from spontaneously exiting the cell in the first place, now function number two is transport, that is it only allows the Selective movement of certain things outside or into that cell so this is a semipermeable membrane and what that means is the cell actually picks and chooses what to allow into the cell and what to not allow into the cell, function number three is signal transduction what that means is certain molecules for instance hormones can actually interact with the outside portion of the cell membrane and that will create many different types of signals and processes inside the cell, function number four energy storage. Therefore cell membrane is an extremely important type of structural component of the cell. Plasma membrane found in prokaryotes and eukaryotes is a phospholipid bilayer means it has two layers of phospholipid. A phospholipid is a molecule that consists of a polar phosphate group attached to a nonpolar fatty acid tail via a group known as the glycerol group. Inside a cell membrane, we also have proteins which give the cell membrane its functionality. They are embedded in the cell membrane. These proteins are of two types' first integral proteins and second peripheral proteins. Integral proteins extend throughout the entire bilayer membrane and their usual function because they extend through the entire membrane is to transport things from the inner portion to the outer portion of our cell and vice versa from the outside to inside of the cell. Peripheral proteins bind via electric forces to either the integral proteins themselves or they bind to phospholipid bilayer. Because these

proteins do not actually span the entire bilayer that means they usually do not act as transport proteins, so they function in cell recognition, cell signalling or cell communication.

In reality, integral proteins embedded in phospholipid bilayer gives the cell membrane its semipermeable character. Which basically means certain materials will pass across the membrane while others will be blocked. Two factors influence the ease with which any external molecule or atom passes across the cell membrane and these factors include the size of a molecule as well as the polarity of that molecule. Size aspect is a bit intuitive so the larger the molecule less likely it will pass through that cell membrane. Since cell membrane is a phospholipid bilayer that means there are two layers of phospholipids and phospholipids themselves are predominantly fatty acids so that means predominant portion of the cell membrane is nonpolar inside, in fact, the entire intermediate space between these heads is basically nonpolar and so that means nonpolar molecules no matter how large will be able to move across and that's exactly why cholesterol which is a large molecule has no problem moving across our cell membrane but even a very small atom that has a full positive or full negative charge such as the sodium atom will not be able to move across the membrane, because it is very polar and contains a full charge.

1.2 Membrane transport

Membrane transport is about how exactly molecules make their way across the cell membrane from one side to the other side of that cell membrane. In doing so molecules use different transport mechanisms or methods and these methods that they use to cross the membrane really depends on the properties of those molecules. To begin let's focus on nonpolar and small molecules. cell membrane consists predominantly of a hydrophobic region so the entire core of the membrane is hydrophobic nonpolar because of the presence of hydrocarbon tails[2], part of the phospholipid molecules now because a non-polar molecule can easily dissolve in a hydrophobic nonpolar solution what that means is if a small non polar molecule wants to make its way across the cell membrane all it has to do is dissolve inside that cell membrane and this process by which a small nonpolar molecule will move down its concentration gradient from a high to low potential through the core of that membrane by dissolving in that cell membrane this is known as simple diffusion. For example, way cell membrane of the cells found inside our lungs facilitates the movement of O_2 and CO_2 molecules easily. Inside the

cells of our lungs we know we have a high concentration of CO_2 inside and a low concentration on the outside conversely we have a low concentration of O_2 on the inside but a high concentration of O_2 on the outside. Both carbon dioxide and oxygen are small nonpolar molecules and therefore these two molecules will have no problem making their way and dissolving into the hydrophobic core of that membrane. These oxygen molecules will naturally move spontaneously from a high potential concentration to a low potential concentration from the outside to the inside. they will pass and dissolve inside and through membrane likewise these carbon dioxide's being nonpolar will also dissolve and move through that membrane via simple diffusion but they will move from the inside to the outside down their concentration gradient.

Now, what about if the molecule is polar? What if it has some type of charge? For instance, let's say we're looking at ions like sodium ions or chloride ions and molecules that don't have a charge yet they're very polar for instance sugar molecules. Sugar molecules are large enough and polar enough to not be able to pass across the membrane via simple diffusion. In this case, if a molecule is polar and large for example sugars and sodium they will not be able to simply dissolve inside the hydrophobic core of cell membrane[3] and in this particular case they have to use another method and what they use is integral proteins that exist inside the membrane of our cells membrane. These are also known as transport membrane proteins because we find these proteins in the plasma membrane and their function is to transport such polar or charged molecules across the two sides of the membrane.

Now there are two types of transport membrane proteins, membrane channels, and membrane pumps. Membrane channels essentially create a passageway in the plasma membrane which basically lack or doesn’t contain the hydrophobic core. In other words, membrane channel removes that hydrophobic core and facilitate the spontaneous diffusion of ions in some cases molecules, down their electrochemical gradient. This type of transport is known as facilitated diffusion. And since membrane channels don't use any energy molecules for instance ATP molecules to carry out the process this type of transport is known as passive transport. Membrane channels and membrane pumps both transport membrane proteins basically move large and polar molecules but the difference is that channels do not use energy

and they always move molecules down their electrochemical gradient from a high to low potential. But membrane pumps actually use energy to move molecules against their electrochemical gradient from a low potential to a high potential. Therefore such type of molecular transport carried out by protein pumps using energy stored in the chemical bonds of ATP molecules is known as active transport.

Chapter 2: Peptide transporter proteins

Peptide transport may be defined as the process by which peptides of two to eight amino acid residues are transported across the cell membrane into a cell. Transport of peptides into cells is a well-documented process which is carried out by specific, energy dependent transporters found in vide variety of organisms[4] as diverse as bacteria and human including fungi, plants and mammals. Peptide transport is an important route for the cellular acquisition of nitrogen and amino acids[5]. Thus obtained nitrogen and amino acids are subsequently used by cell for metabolism and growth. Therefore cellular uptake of peptides is an important physiological process mediated by the PTR family of proton-coupled peptide transporters[6]. PTR family proteins are members of the major facilitator superfamily (MFS) of secondary active transporters[7]. Proteins of proton-dependent oligopeptide transporter (POT) family also known as peptide transport family (PTR) are of about 450-600 amino acyl residues in length with the eukaryotic proteins in general being longer than the prokaryotic bacterial proteins. They exhibit 12 to 14 established or putative transmembrane α-helical spanners. Remarkable characteristic of POT family protein is their diverse substrate promiscuity meaning POT family proteins recognise and transport diverse library of peptides. This promiscuity has led to development of peptide based pro-drug that use PepT1 and PepT2, the mammalian homologues, to improve oral drug delivery[8]. Therefore interest in peptide transporters of mammalian members increased recently after the discovery that both PepT1 and PepT2 are able to recognise and transport a large number of orally administered drug molecules[9], including beta-lactam antibiotics, antiviral and anticancer drugs. Major challenges in drug delivery system are poor intestinal absorption & distribution of drug molecule, missing desired target eventual side effect or breakdown of drug & excretion, poor bioavailability. A successful strategy to tackle this problem has been to develop prodrug molecules, which could target the intestinal peptide transporter, mammalian members of POT family transporters for active transport into the body[10]. Prodrugs are bioreversible derivatives of drug molecules that undergo an enzymatic or chemical transformation in vivo to release the active parent drug[11]. This ability

Of PepT1 and PepT2 to recognise and transport peptide based prodrug molecules resulted into increased attention of mammalian homologues by pharmaceutical industry.

Chapter 3: Methods used

3.1 Protein sequence alignment: aligned protein sequences of 46 peptide transporter family member proteins using Uniprot database.

In bioinformatics sequence alignment is considered as basic fundamental procedure to compare given sequence with other sequence available in biological database either it can be RNA, DNA or Protein. By sequence alignment we can understand about sequence homology whether the sequence is ortholog or paralog? Or whether the given sequence has very good sequence identity with other sequence in database? Or sequence conservation whether the active site residues are conserved? Or motifs are conserved? Such information we get by sequence alignment. Protein sequence alignment gives more information than DNA sequence alignment because protein sequence contain 20 different characters like 20 Amino Acid having each different chemical properties therefore protein sequence alignment is more complicated as compare to DNA sequence alignment[12], where only 4 characters A,T,G,C are present. Which can be easily analysed but information we get by DNA sequence alignment is may not be reliable as compare to that of protein sequence alignment.

Why we need sequence alignment? By sequence alignment we can identify changes or differences at the individual base level or amino acid level between two or more sequences. Sequence alignment helps us to organise and analyse sequence data to check the sequence homology and identity between sequences[13]. Sequence alignment help us to phylogenetic trees from homologues sites. It helps use to estimate and establish evolutionary relationship. And we can identify and highlight the conserved and variable sites and regions. It also used to test and hypotheses about protein fold, motifs, domains and their three dimensional structures.

Uniprot which stands for universal protein resource is an online freely accessible database. It aims to provide various types of information about proteins for example protein sequence including isoforms and variants if present. It also provides detailed information on

protein function, interaction, pathways, involvement in diseases and other such biological areas of interest. It also provides stable identifiers and accessions. Uniprot provides "Align" tool for multiple sequence alignment. End result of sequence alignment depends upon algorithms used for alignment. Uniprot website can be accessed at the URL http://www.uniprot.org/.

3.2 Protein structure prediction:

Used web based tool Phyre2 to predict structure of PepT1 and PepT2 proteins. This server can be accessed at the URL http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index .

Phyre2 or a Protein Homology/Analogy Recognition Engine is an online web service to predict three dimensional structures of proteins[14]. At present we have around fifteen thousand structures of proteins in protein data bank, on the other hand the available information regarding protein sequences in Uniport database is about eighty nine million sequences. If we compare the availability of amino acid sequences and structures available, sequences are about six thousand fold than the availability of the structures. The protein structures is helpful in understanding function, identifying active sites, antigenic sites and different binding sites with other complexes. Therefore it is very important to have the structures of proteins than sequence information. Due to the availability of more number of sequences and less number of structures it is important to predict the structures of proteins from just amino acid sequence. Deciphering the native conformation of protein from amino acid sequence is called a protein folding problem. From the unfolded state, protein finally folds into the stable three dimensional structure. So how to obtain 3D structure from the amino acid sequence. There are various methods to predict the 3D structures just from amino acid sequence. One of the most popular methods is homology modelling. Which is also used by Phyre2 an online 3D structure predicting web service. Homology modelling is based on the principle that if two sequences share high sequence identity or sequence similarity then the assumption is that the structures are also similar.

3.3 Phylogenetic tree: used Uniprot to establish evolutionary relationship between 46 POT family members.

All species are related to one another. Which could be easily observed by looking at the species in itself. There are similarities which are progressing in a hierarchical way from a unicellular organism to multicellular organisms. When we observe, we could see there is a hierarchical progress in several levels which eventually ends up with evolved organisms like humans or other animals. Having explained in a qualitative way it is also possible to explain this relationship between the species in quantitative manner. A tree structure representing this relationship in a quantitative manner is called as a phylogenetic tree. In earlier studies morphological characteristics were recorded from existing or fossilized species which were used to obtain this quantitative relationship and was represented by means of phylogenetic trees. Recently revolution in sequencing technologies and development of efficient bioinformatics algorithms has allowed us to use DNA or protein sequences for phylogenetic analysis. So earlier morphological traits were used for building phylogenetic trees but with the various bioinformatics methods and sequencing methodologies that we have today we can use molecular information and describe similarities and dissimilarities between sequences in terms of phylogeny. With the set of DNA or protein sequences from different species we are able to infer the phylogenetic relationship among the species. This follows an underlying assumption that the sequence considered has evolved from a common ancestral gene such set of sequences which comes out from a common ancestor gene are known as orthologous sequences[15]. It is also possible that two sequences considered are similar because of some gene duplication events and it could be present in the same species such sequences are called as paralog they might even have the same function but they don't have a common ancestor. Therefore phylogenetic informations are usually derived from an ortholog a set of gene.

Chapter 4: Result

1] Classification of POT family proteins based on number of transmembrane α- helical spanners

No of transmembrane α-helical spanners	Protein name	Organism	Gene
8	Uncharacterized protein	Rhizophagus irregularis	GLOINDRAFT_34 3346

Table 1.

Figure 3: structure of U9UKY5_RHIID Uncharacterized protein OS=Rhizophagus irregularis modelled using Phyre2.

No of transmembrane α-helical spanners	Protein name	Organism	Gene
9	Nitrate transporter	Oryza sativa (Rice)	NRT1

Table 2

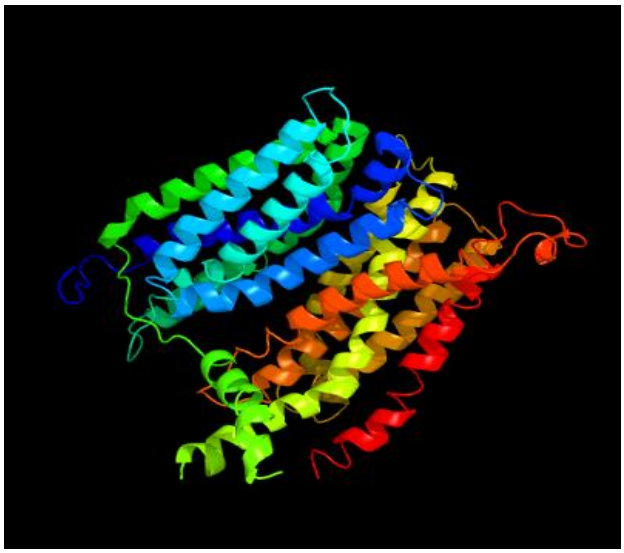


Figure 4: structure of Q9SPU1_ORYSA Nitrate transporter OS= Oryza sativa (Rice) modelled using Phyre2.

No of transmembrane α-helical spanners	Protein name	Organism	Gene
10	Probable peptide transporter ptr2	Schizosaccharomyces pombe (Fungus)	ptr2
	Protein NRT1/ PTR FAMILY 8.1	Arabidopsis thaliana (Mouse-ear cress)	NPF8.1
	Protein NRT1/ PTR FAMILY 8.2	Arabidopsis thaliana (Mouse-ear cress)	NPF8.2
	Protein NRT1/ PTR FAMILY 1.2	Arabidopsis thaliana (Mouse-ear cress)	NPF1.2
	Peptide transporter PTR3-A	Aegilops tauschii (Tausch's goatgrass)	F775_27357
	Peptide transporter family 1	Drosophila melanogaster (Fruit fly)	yin

Table 3

No of transmembrane α-helical spanners	Protein name	Organism	Gene
11	Peptide transporter PTR2	Saccharomyces cerevisiae (Baker's yeast)	PTR2
	Putative peptide transporter	Trichoderma harzianum (Fungus)	ptr2
	Ptr22p	Candida albicans (Yeast)	PTR22
	Protein NRT1/ PTR FAMILY 8.3	Arabidopsis thaliana (Mouse-ear cress)	NPF8.3
	RCH2 protein	Brassica napus (Rape)	N/A
	Protein NRT1/ PTR FAMILY 5.2	Arabidopsis thaliana (Mouse-ear cress)	NPF5.2
	Solute carrier family 15 member 5	Homo sapiens (Human)	SLC15A5
	Protein NRT1/ PTR FAMILY 2.12	Arabidopsis thaliana (Mouse-ear cress)	NPF2.12
	Protein NRT1/ PTR FAMILY 7.3	Arabidopsis thaliana (Mouse-ear cress)	NPF7.3
	Oligopeptide transporter PEPT1	Fundulus heteroclitus macrolepidotus	SLC15A1b
	Peptide transporter family 1	Caenorhabditis elegans (Roundworm)	pept-1

Table 4

No of transmembrane α-helical spanners	Protein name	Organism	Gene
12	Di-/tripeptide transporter	Lactococcus lactis (Bacteria)	dtpT
	Di-tripeptide ABC transporter	Geobacillus kaustophilus	GK2020
	Protein NRT1/ PTR FAMILY 6.3	Arabidopsis thaliana	NPF6.3
	Protein NRT1/ PTR FAMILY 2.7	Arabidopsis thaliana	NPF2.7
	Protein NRT1/ PTR FAMILY 3.1	Arabidopsis thaliana	NPF3.1
	Solute carrier family 15 member 3	Homo sapiens (Human)	SLC15A3
	Solute carrier family 15 member 4	Homo sapiens (Human)	SLC15A4
	Protein NRT1/ PTR FAMILY 4.2	Arabidopsis thaliana	NPF4.2
	Protein NRT1/ PTR FAMILY 2.13	Arabidopsis thaliana	NPF2.13
	Protein NRT1/ PTR FAMILY 4.6	Arabidopsis thaliana	NPF4.6
	Protein NRT1/ PTR FAMILY 2.10	Arabidopsis thaliana	NPF2.10
	Protein NRT1/ PTR FAMILY 6.2	Arabidopsis thaliana	NPF6.2
	Protein NRT1/ PTR FAMILY 2.11	Arabidopsis thaliana	NPF2.11
	Solute carrier family 15 member 1	Rattus norvegicus (Rat)	Slc15a1
	Peptide transporter family 2	Caenorhabditis elegans	pept-2
	Solute carrier family 15 member 2	Mus musculus (Mouse)	Slc15a2
	Solute carrier family 15 member 2	Danio rerio (Zebrafish)	slc15a2
	Solute carrier family 15 member 2	Homo sapiens (Human)	SLC15A2
	Solute carrier family 15 member 1	Homo sapiens (Human)	SLC15A1
	Peptide transporter 3	Caenorhabditis elegans	pept-3

Table 5

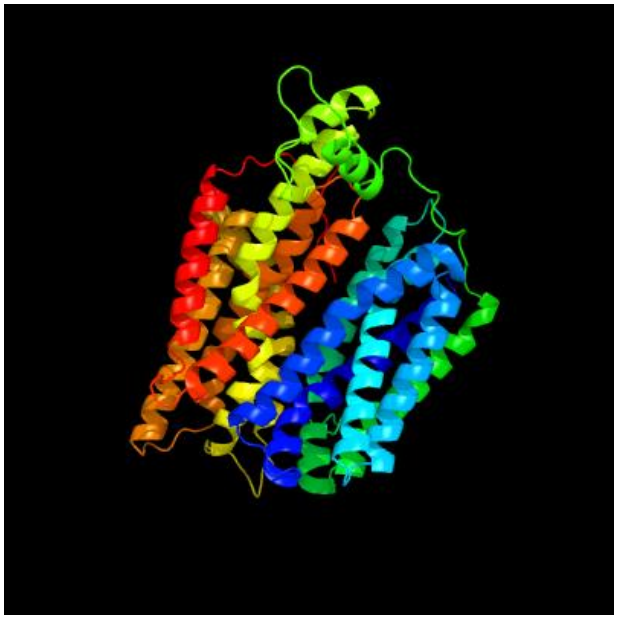


Figure 5: P46059|S15A1_HUMAN Solute carrier family 15 member (PepT1) protein structure modelled using Phyre2.

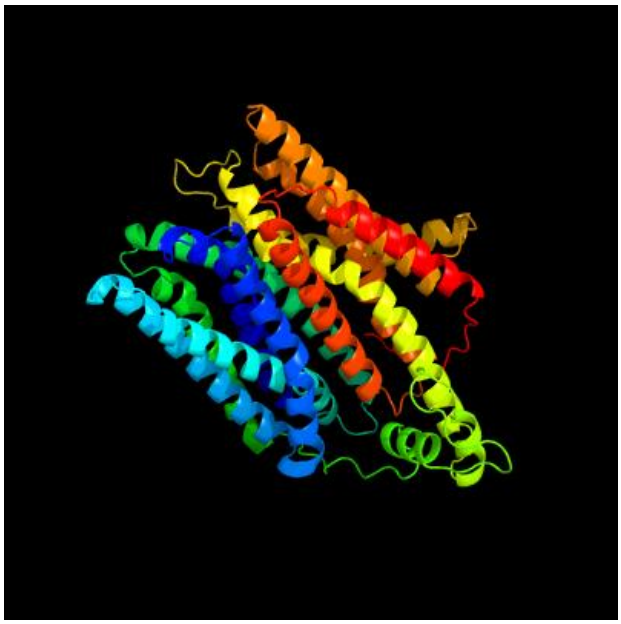


Figure 6: Q16348|S15A2_HUMAN Solute carrier family 15 member 2 (PepT2) protein 3D structure modelled using Phyre2.

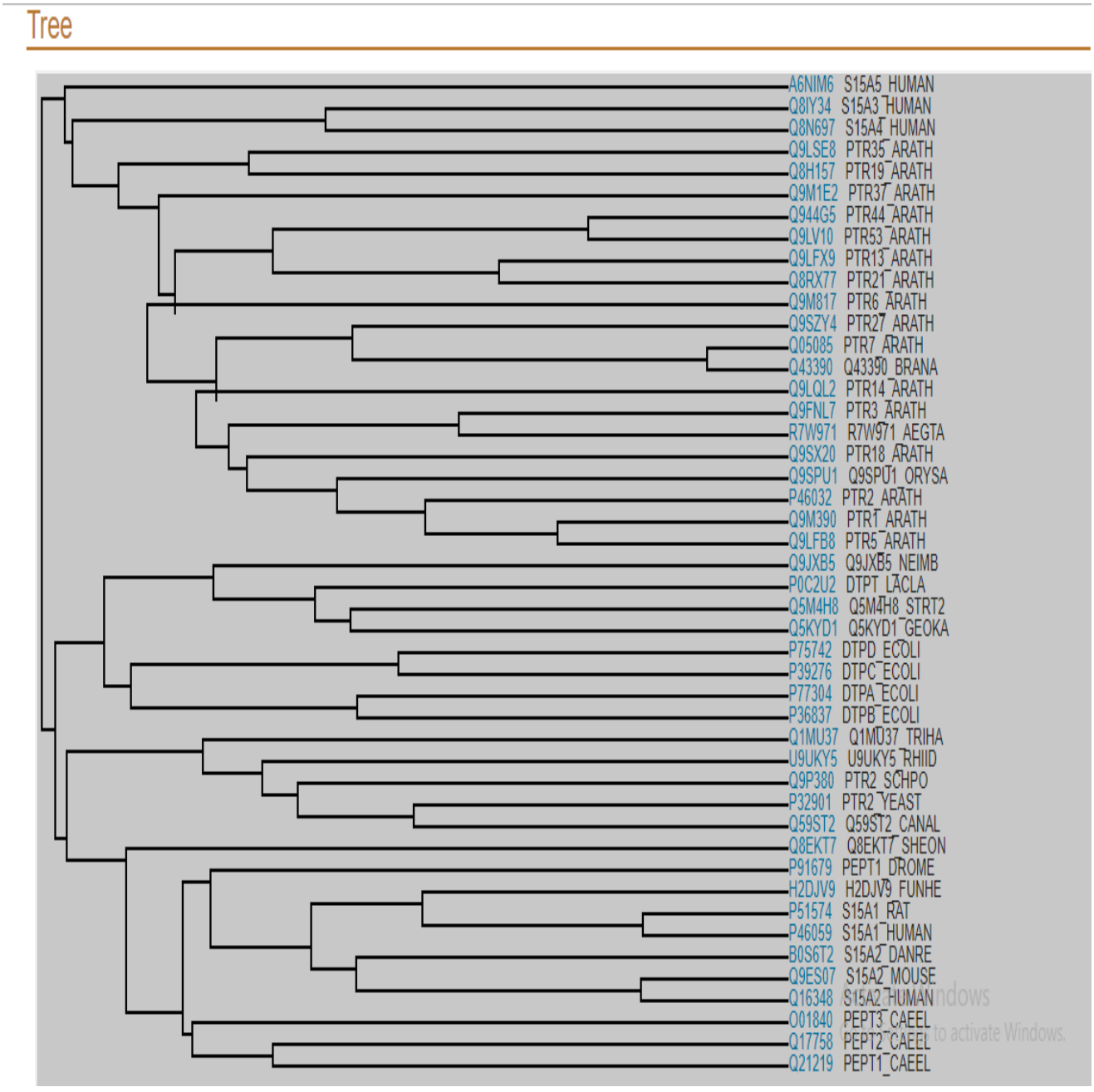
No of transmembrane α-helical spanners	Protein name	Organism	Gene
13	Di-or tripeptide:H+ symporter	Streptococcus thermophiles (Bacteria)	dtpT
	Glutathione uptake transporte	Shewanella oneidensis (Bacteria)	SO_0002

Table 6

No of transmembrane α-helical spanners	Protein name	Organism	Gene
14	Dipeptide and tripeptide permease A	Escherichia coli (Bacteria)	dtpA
	Dipeptide and tripeptide permease B	Escherichia coli (Bacteria)	dtpB
	Dipeptide permease D	Escherichia coli (Bacteria)	dtpD
	Dipeptide and tripeptide permease C	Escherichia coli (Bacteria)	dtpC
	Peptide transporter	Neisseria meningitidis serogroup B (Bacteria)	NMB2136

Table 7

2] Phylogenetic tree establishing evolutionary relationship between POT family members.



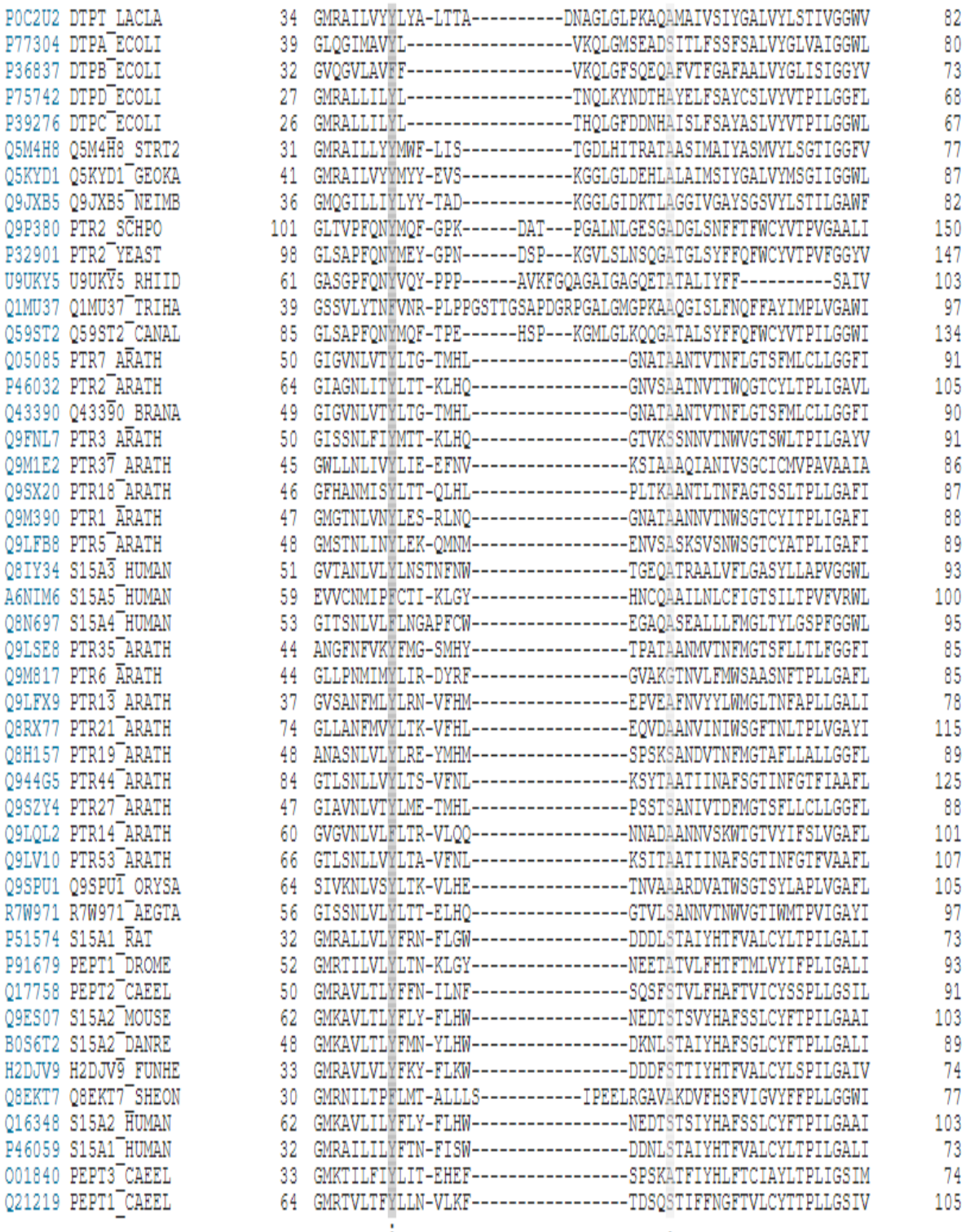
As we can see in above phylogenetic tree, peptide transporter protein which is root protein in this case has evolved mainly into two branches. From tree we can deduce POT members of

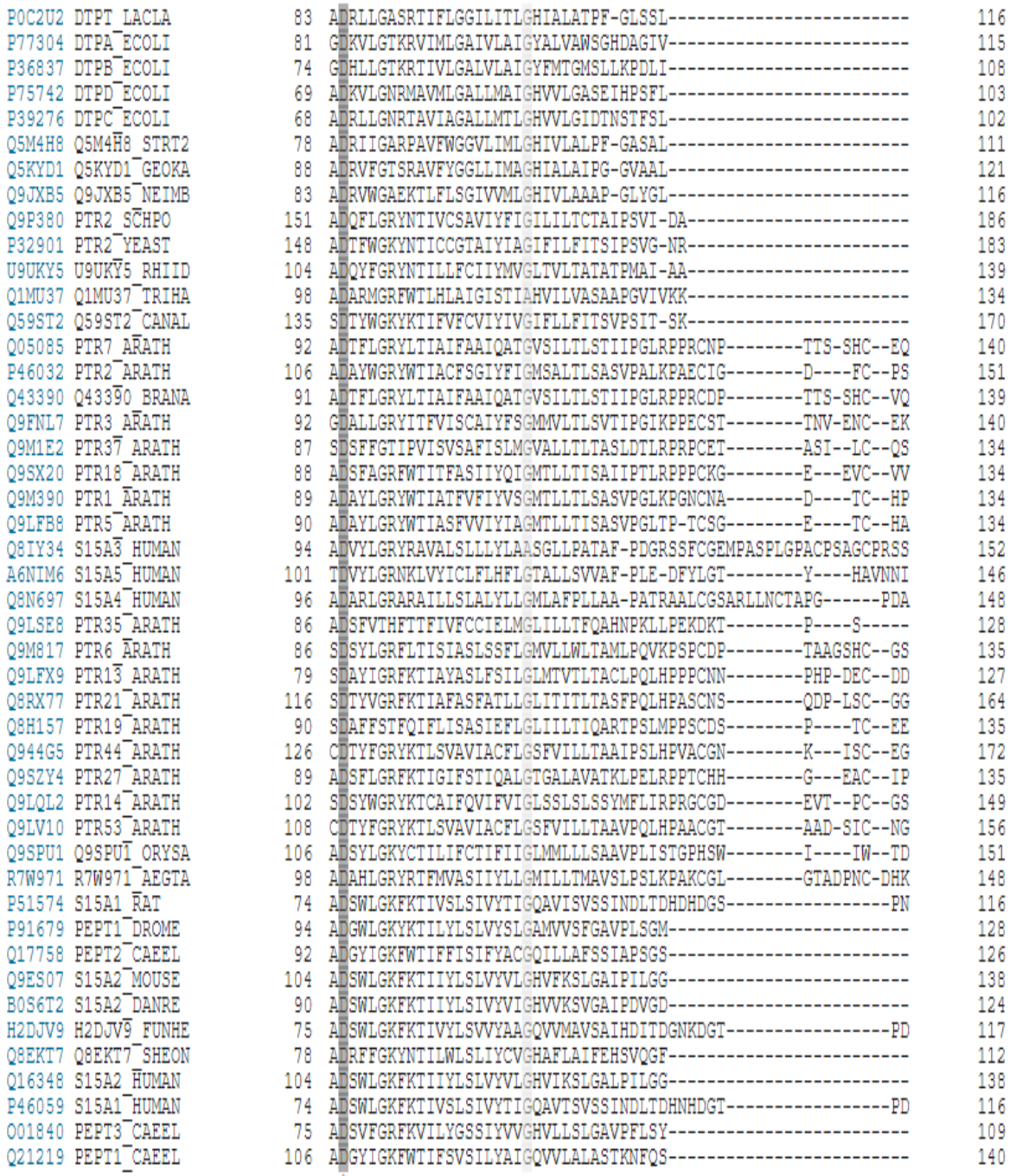
eukaryotic organisms are more closely related to one another as compare to prokaryotic organisms.

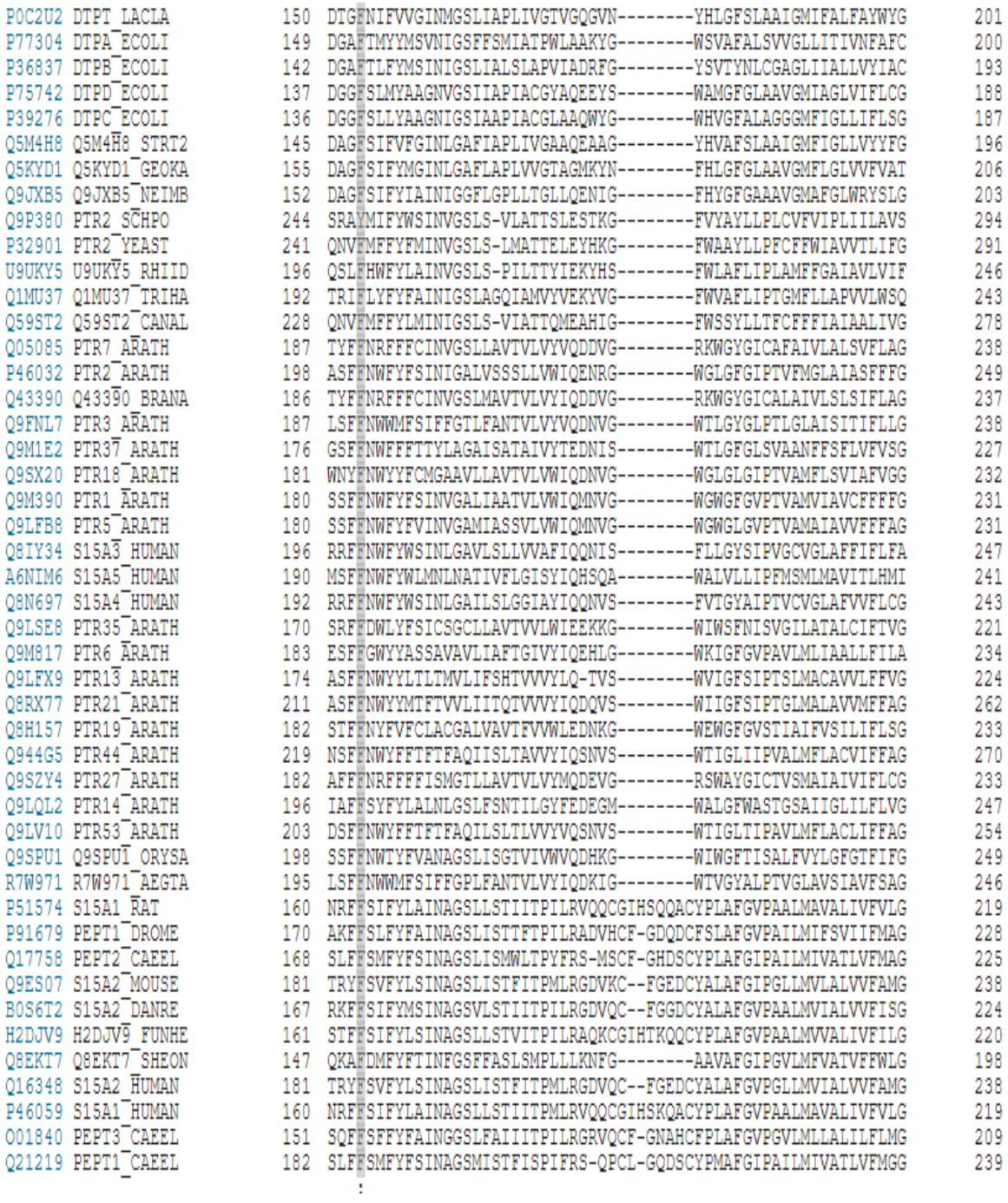
REFERENCE:

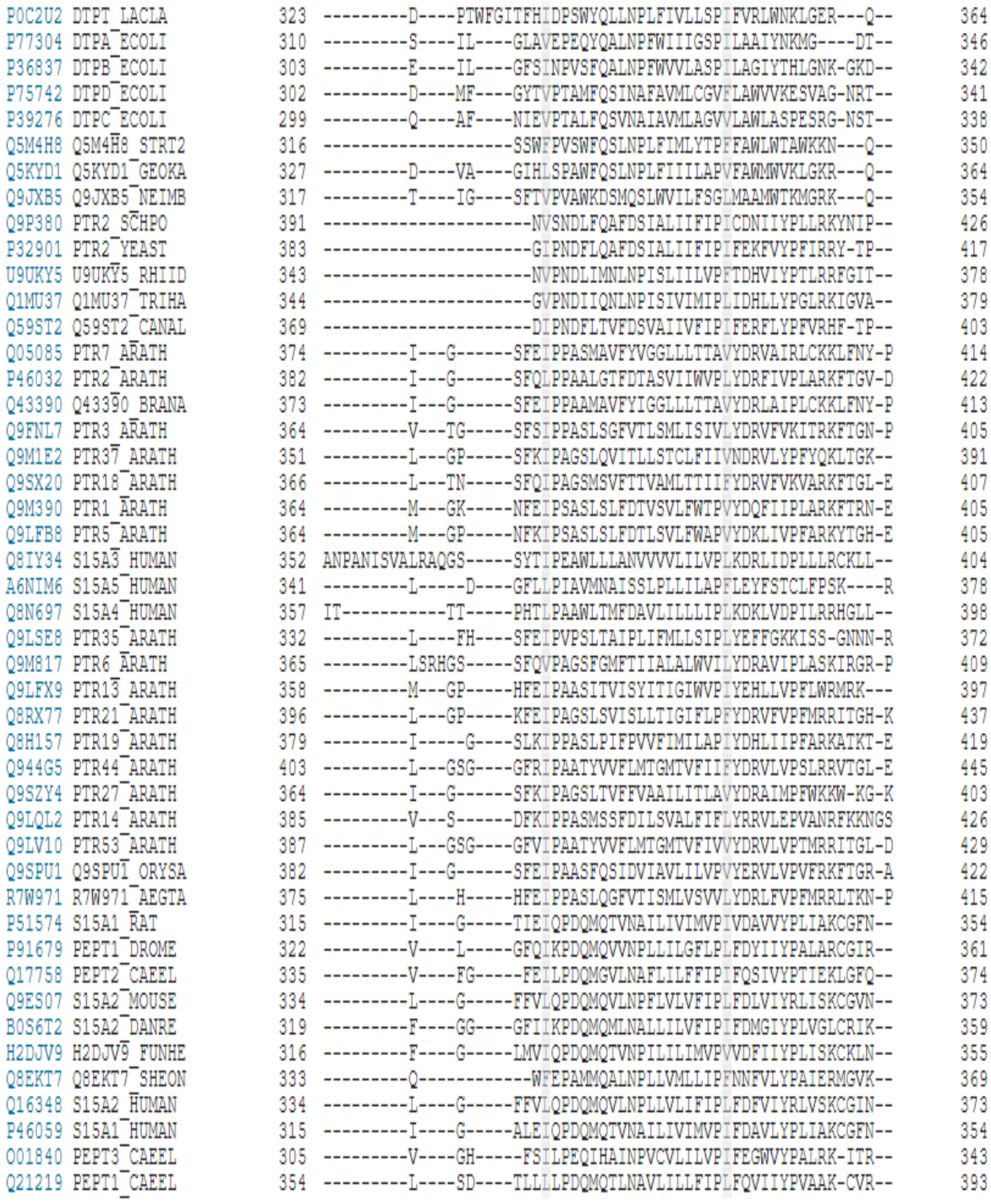
1. Simon newstead *Current Opinion in Structural Biology* **2017**, *45: 17-24*
2. Lodish H, Berk A, Zipursky SL, *et al. New York: W. H. Freeman;* **2000.** Section 3.4Membrane Proteins
3. Lodish H, Berk A, Zipursky SL, *et al. New York: W. H. Freeman;* **2000.** Section 15.1Diffusion of Small Molecules across Phospholipid Bilayers.
4. Steiner HY, Naider F, Becker JM*, molecular microbiology 16(5):825-34,* **July 1995**
5. Simon newstead, *Biochim Biophys Acta.* **2015 Mar***;1850(3):488-99*
6. Simon newstead, *Biochim Biophys Acta* **2015 Mar***; 1850(3): 488–499.*
7. Stephanie S. Pao, Ian T. Paulsen, and Milton H. Saier Jr*. Microbiol Mol Biol Rev.* **1998 Mar***; 62(1): 1–34.*
8. Simon newstead*, Curr Opin Struct Bio***. 2017 Aug***; 45:17-24.*
9. Simon newstead*, Biochim Biophys Acta* **2015 Mar***; 1850(3): 488–499*.
10. Gurdeep S. Minhas and Simon Newstead, *Proc Natl Acad Sci U S A* **2019 Jan 15;** *116(3): 804–809.*
11. Jarkko Rautio, Krista Laine, Mikko Gynther, and Jouko Savolainen *AAPS J***. 2008 Mar**; 10(1): 92–102.
12. Matthias branssch ilka knutter eva bosse-*doenecke journal of pharmacy and pharmacology 60(5):543-85*
13. Simmonds, P. BMC *Res Notes* **(2012)** 5: 50.
14. Vijay Nema, Sudhir Kumar Pal , *Bioinformation 9(15): 796-801* **(2013)**
15. Barry G. Hall *Molecular Biology and Evolution*, Volume 30, Issue 5, **May 2013**, Pages 1229–1235.

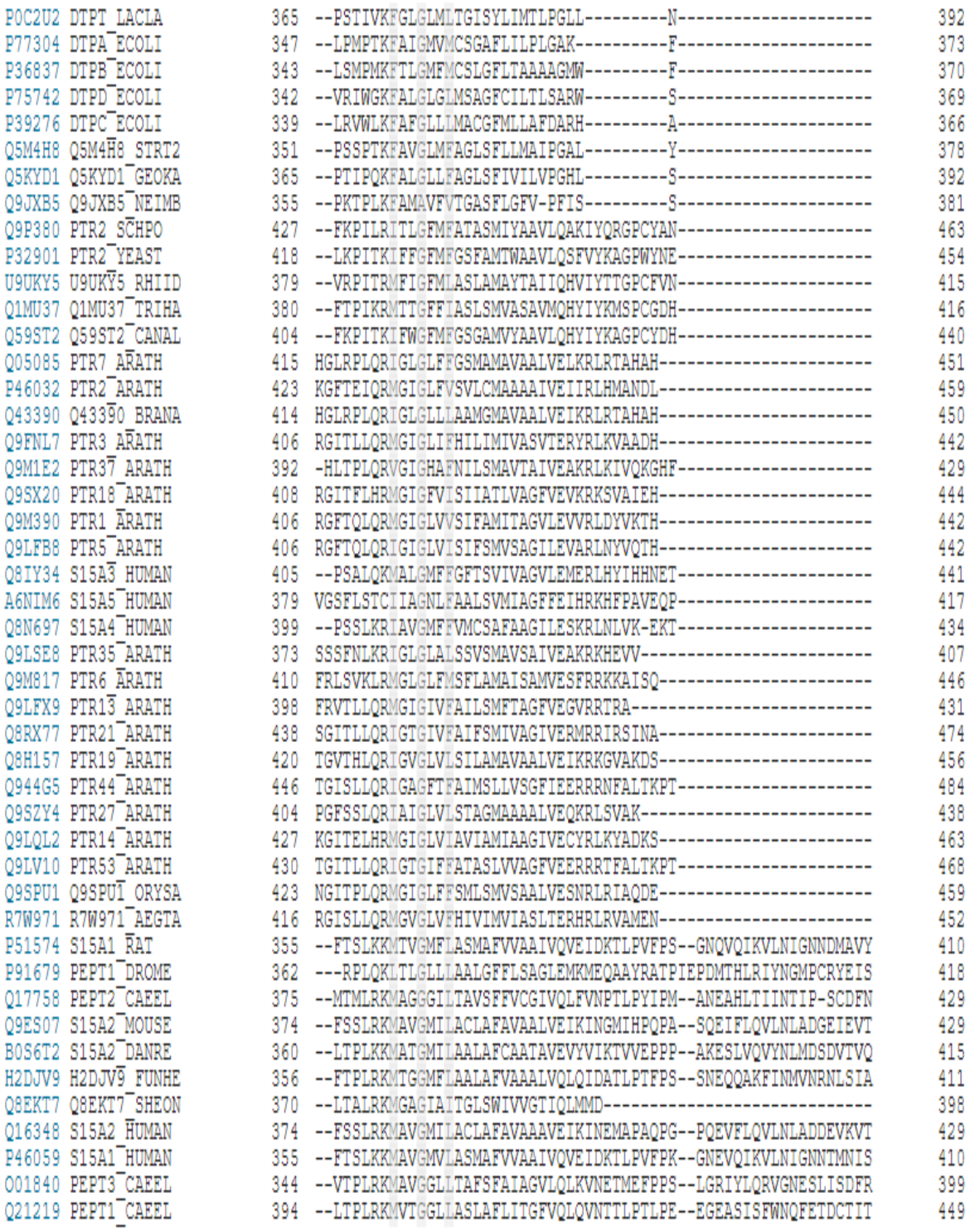
Detailed sequence alignment of 46 POT family transporters highlighting conserved residues.

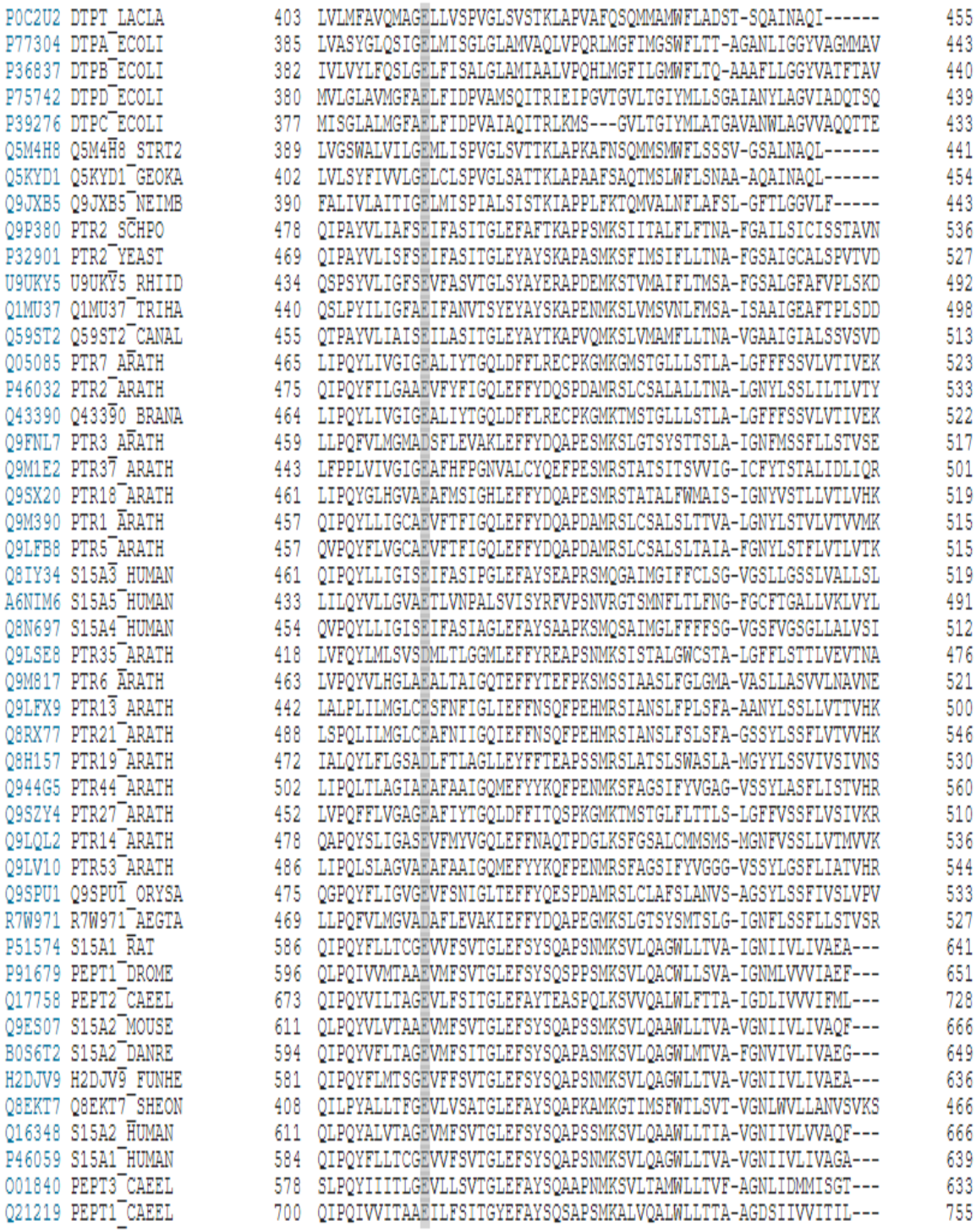



P0C2U2 DTPT_LACLA 83 ADRLLGASRTIFLGGILITLGHIALATPF-GLSSL---------- 116
P77304 DTPA_ECOLI 81 GDKVLGTKRVIMLGAIVLAIGYALVAWSGHDAGIV---------- 115
P36837 DTPB_ECOLI 74 GDHLLGTKRTIVLGALVLAIGYFMTGMSLLKPDLI---------- 108
P75742 DTPD_ECOLI 69 ADKVLGNRMAVMLGALLMAIGHVVLGASEIHPSFL---------- 103
P39276 DTPC_ECOLI 68 ADRLLGNRTAVIAGALLMTLGHVVLGIDTNSTFSL---------- 102
Q5M4H8 Q5M4H8_STRT2 78 ADRIIGARPAVFWGGVLIMLGHIVLALPF-GASAL---------- 111
Q5KYD1 Q5KYD1_GEOKA 88 ADRVFGTSRAVFYGGLLIMAGHIALAIPG-GVAAL---------- 121
Q9JXB5 Q9JXB5_NEIMB 83 ADRVWGAEKTLFLSGIVVMLGHIVLAAAP-GLYGL---------- 116
Q9P380 PTR2_SCHPO 151 ADQFLGRYNTIVCSAVIYFIGILILTCTAIPSVI-DA---------- 186
P32901 PTR2_YEAST 148 ADTFWGKYNTICCGTAIYIAGIFILFITSIPSVG-NR---------- 183
U9UKY5 U9UKY5_RHIID 104 ADQYFGRYNTILLFCIIYMVGLTVLTATATPMAI-AA---------- 139
Q1MU37 Q1MU37_TRIHA 98 ADARMGRFWTLHLAIGISTIAHVILVASAAPGVIVKK---------- 134
Q59ST2 Q59ST2_CANAL 135 SDTYWGKYKTIFVFCVIYIVGIFLLFITSVPSIT-SK---------- 170
Q05085 PTR7_ARATH 92 ADTFLGRYLTIAIFAAIQATGVSILTLSTIIPGLRPPRCNP--------TTS-SHC--EQ 140
P46032 PTR2_ARATH 106 ADAYWGRYWTIACFSGIYFIGMSALTLSASVPALKPAECIG--------D----FC--PS 151
Q43390 Q43390_BRANA 91 ADTFLGRYLTIAIFAAIQATGVSILTLSTIIPGLRPPRCDP--------TTS-SHC--VQ 139
Q9FNL7 PTR3_ARATH 92 GDALLGRYITFVISCAIYFSGMMVLTLSVTIPGIKPPECST--------TNV-ENC--EK 140
Q9M1E2 PTR37_ARATH 87 SDSFFGTIPVISVSAFISLMGVALLTLTASLDTLRPRPCET--------ASI--LC--QS 134
Q9SX20 PTR18_ARATH 88 ADSFAGRFWTITFASIIYQIGMTLLTISAIIPTLRPPPCKG--------E---EVC--VV 134
Q9M390 PTR1_ARATH 89 ADAYLGRYWTIATFVFIYVSGMTLLTLSASVPGLKPGNCNA--------D----TC--HP 134
Q9LFB8 PTR5_ARATH 90 ADAYLGRYWTIASFVVIYIAGMTLLTISASVPGLTP-TCSG--------E----TC--HA 134
Q8IY34 S15A3_HUMAN 94 ADVYLGRYRAVALSLLLYLAASGLLPATAF-PDGRSSFCGEMPASPLGPACPSAGCPRSS 152
A6NIM6 S15A5_HUMAN 101 TDVYLGRNKLVYICLFLHFLGTALLSVVAF-PLE-DFYLGT--------Y----HAVNNI 146
Q8N697 S15A4_HUMAN 96 ADARLGRARAILLSLALYLLGMLAFPLLAA-PATRAALCGSARLLNCTAPG------PDA 148
Q9LSE8 PTR35_ARATH 86 ADSFVTHFTTFIVFCCIELMGLILLTFQAHNPKLLPEKDKT--------P----S----- 128
Q9M817 PTR6_ARATH 86 SDSYLGRFLTISIASLSSFLGMVLLWLTAMLPQVKPSPCDP--------TAAGSHC--GS 135
Q9LFX9 PTR13_ARATH 79 SDAYIGRFKTIAYASLFSILGLMTVTLTACLPQLHPPPCNN--------PHP-DEC--DD 127
Q8RX77 PTR21_ARATH 116 SDTYVGRFKTIAFASFATLLGLITITLTASFPQLHPASCNS--------QDP-LSC--GG 164
Q8H157 PTR19_ARATH 90 SDAFFSTFQIFLISASIEFLGLIILTIQARTPSLMPPSCDS--------P----TC--EE 135
Q944G5 PTR44_ARATH 126 CDTYFGRYKTLSVAVIACFLGSFVILLTAAIPSLHPVACGN--------K---ISC--EG 172
Q9SZY4 PTR27_ARATH 89 ADSFLGRFKTIGIFSTIQALGTGALAVATKLPELRPPTCHH--------G---EAC--IP 135
Q9LQL2 PTR14_ARATH 102 SDSYWGRYKTCAIFQVIFVIGLSSLSLSSYMFLIRPRGCGD--------EVT--PC--GS 149
Q9LV10 PTR53_ARATH 108 CDTYFGRYKTLSVAVIACFLGSFVILLTAAVPQLHPAACGT--------AAD-SIC--NG 156
Q9SPU1 Q9SPU1_ORYSA 106 ADSYLGKYCTILIFCTIFIIGLMMLLLSAAVPLISTGPHSW--------I----IW--TD 151
R7W971 R7W971_AEGTA 98 ADAHLGRYRTFMVASIIYLLGMILLTMAVSLPSLKPAKCGL--------GTADPNC-DHK 148
P51574 S15A1_RAT 74 ADSWLGKFKTIVSLSIVYTIGQAVISVSSINDLTDHDHDGS----------------PN 116
P91679 PEPT1_DROME 94 ADGWLGKYKTILYLSLVYSLGAMVVSFGAVPLSGM---------- 128
Q17758 PEPT2_CAEEL 92 ADGYIGKFWTIFFISIFYACGQILLAFSSIAPSGS---------- 126
Q9ES07 S15A2_MOUSE 104 ADSWLGKFKTIIYLSLVYVLGHVFKSLGAIPILGG---------- 138
B0S6T2 S15A2_DANRE 90 ADSWLGKFKTIIYLSIVYVIGHVVKSVGAIPDVGD---------- 124
H2DJV9 H2DJV9_FUNHE 75 ADSWLGKFKTIVYLSVVYAAGQVVMAVSAIHDITDGNKDGT----------------PD 117
Q8EKT7 Q8EKT7_SHEON 78 ADRFFGKYNTILWLSLIYCVGHAFLAIFEHSVQGF---------- 112
Q16348 S15A2_HUMAN 104 ADSWLGKFKTIIYLSLVYVLGHVIKSLGALPILGG---------- 138
P46059 S15A1_HUMAN 74 ADSWLGKFKTIVSLSIVYTIGQAVTSVSSINDLTDHNHDGT----------------PD 116
O01840 PEPT3_CAEEL 75 ADSVFGRFKVILYGSSIYVVGHVLLSLGAVPFLSY---------- 109
Q21219 PEPT1_CAEEL 106 ADGYIGKFWTIFSVSILYAIGQVVLALASTKNFQS---------- 140


P0C2U2 DTPT_LACLA 150 DTGFNIFVVGINMGSLIAPLIVGTVGQGVN--------YHLGFSLAAIGMIFALFAYWYG 201
P77304 DTPA_ECOLI 149 DGAFTMYYMSVNIGSFFSMIATPWLAAKYG--------WSVAFALSVVGLLITIVNFAFC 200
P36837 DTPB_ECOLI 142 DGAFTLFYMSINIGSLIALSLAPVIADRFG--------YSVTYNLCGAGLIIALLVYIAC 193
P75742 DTPD_ECOLI 137 DGGFSLMYAAGNVGSIIAPIACGYAQEEYS--------WAMGFGLAAVGMIAGLVIFLCG 188
P39276 DTPC_ECOLI 136 DGGFSLLYAAGNIGSIAAPIACGLAAQWYG--------WHVGFALAGGGMFIGLLIFLSG 187
Q5M4H8 Q5M4H8_STRT2 145 DAGFSIFVFGINLGAFIAPLIVGAAQEAAG--------YHVAFSLAAIGMFIGLLVYYFG 196
Q5KYD1 Q5KYD1_GEOKA 155 DAGFSIFYMGINLGAFLAPLVVGTAGMKYN--------FHLGFGLAAVGMFLGLVVFVAT 206
Q9JXB5 Q9JXB5_NEIMB 152 DAGFSIFYIAINIGGFLGPLLTGLLQENIG--------FHYGFGAAAVGMAFGLWRYSLG 203
Q9P380 PTR2_SCHPO 244 SRAYMIFYWSINVGSLS-VLATTSLESTKG--------FVYAYLLPLCVFVIPLIILAVS 294
P32901 PTR2_YEAST 241 QNVFMFFYFMINVGSLS-LMATTELEYHKG--------FWAAYLLPFCFFWIAVVTLIFG 291
U9UKY5 U9UKY5_RHIID 196 QSLFHWFYLAINVGSLS-PILTTYIEKYHS--------FWLAFLIPLAMFFGAIAVLVIF 246
Q1MU37 Q1MU37_TRIHA 192 TRIFLYFYFAINIGSLAGQIAMVYVEKYVG--------FWVAFLIPTGMFLLAPVVLWSQ 243
Q59ST2 Q59ST2_CANAL 228 QNVFMFFYLMINIGSLS-VIATTQMEAHIG--------FWSSYLLTFCFFFIAIAALIVG 278
Q05085 PTR7_ARATH 187 TYFFNRFFFCINVGSLLAVTVLVYVQDDVG--------RKWGYGICAFAIVLALSVFLAG 238
P46032 PTR2_ARATH 198 ASFFNWFYFSINIGALVSSSLLVWIQENRG--------WGLGFGIPTVFMGLAIASFFFG 249
Q43390 Q43390_BRANA 186 TYFFNRFFFCINVGSLMAVTVLVYIQDDVG--------RKWGYGICALAIVLSLSIFLAG 237
Q9FNL7 PTR3_ARATH 187 LSFFNWWMFSIFFGTLFANTVLVYVQDNVG--------WTLGYGLPTLGLAISITIFLLG 238
Q9M1E2 PTR37_ARATH 176 GSFFNWFFFTTYLAGAISATAIVYTEDNIS--------WTLGFGLSVAANFFSFLVFVSG 227
Q9SX20 PTR18_ARATH 181 WNYFNWYYFCMGAAVLLAVTVLVWIQDNVG--------WGLGLGIPTVAMFLSVIAFVGG 232
Q9M390 PTR1_ARATH 180 SSFFNWFYFSINVGALIAATVLVWIQMNVG--------WGWGFGVPTVAMVIAVCFFFFG 231
Q9LFB8 PTR5_ARATH 180 SSFFNWFYFVINVGAMIASSVLVWIQMNVG--------WGWGLGVPTVAMAIAVVFFFAG 231
Q8IY34 S15A3_HUMAN 196 RRFFNWFYWSINLGAVLSLLVVAFIQQNIS--------FLLGYSIPVGCVGLAFFIFLFA 247
A6NIM6 S15A5_HUMAN 190 MSFFNWFYWLMNLNATIVFLGISYIQHSQA--------WALVLLIPFMSMLMAVITLHMI 241
Q8N697 S15A4_HUMAN 192 RRFFNWFYWSINLGAILSLGGIAYIQQNVS--------FVTGYAIPTVCVGLAFVVFLCG 243
Q9LSE8 PTR35_ARATH 170 SRFFDWLYFSICSGCLLAVTVVLWIEEKKG--------WIWSFNISVGILATALCIFTVG 221
Q9M817 PTR6_ARATH 183 ESFFGWYYASSAVAVLIAFTGIVYIQEHLG--------WKIGFGVPAVLMLIAALLFILA 234
Q9LFX9 PTR13_ARATH 174 ASFFNWYYLTLTMVLIFSHTVVVYLQ-TVS--------WVIGFSIPTSLMACAVVLFFVG 224
Q8RX77 PTR21_ARATH 211 ASFFNWYYMTFTVVLIITQTVVVYIQDQVS--------WIIGFSIPTGLMALAVVMFFAG 262
Q8H157 PTR19_ARATH 182 STFFNYFVFCLACGALVAVTFVVWLEDNKG--------WEWGFGVSTIAIFVSILIFLSG 233
Q944G5 PTR44_ARATH 219 NSFFNWYFFTFTFAQIISLTAVVYIQSNVS--------WTIGLIIPVALMFLACVIFFAG 270
Q9SZY4 PTR27_ARATH 182 AFFFNRFFFFISMGTLLAVTVLVYMQDEVG--------RSWAYGICTVSMAIAIVIFLCG 233
Q9LQL2 PTR14_ARATH 196 IAFFSYFYLALNLGSLFSNTILGYFEDEGM--------WALGFWASTGSAIIGLILFLVG 247
Q9LV10 PTR53_ARATH 203 DSFFNWYFFTFTFAQILSLTLVVYVQSNVS--------WTIGLTIPAVLMFLACLIFFAG 254
Q9SPU1 Q9SPU1_ORYSA 198 SSFFNWTYFVANAGSLISGTVIVWVQDHKG--------WIWGFTISALFVYLGFGTFIFG 249
R7W971 R7W971_AEGTA 195 LSFFNWWMFSIFFGPLFANTVLVYIQDKIG--------WTVGYALPTVGLAVSIAVFSAG 246
P51574 S15A1_RAT 160 NRFFSIFYLAINAGSLLSTIITPILRVQQCGIHSQQACYPLAFGVPAALMAVALIVFVLG 219
P91679 PEPT1_DROME 170 AKFFSLFYFAINAGSLISTTFTPILRADVHCF-GDQDCFSLAFGVPAILMIFSVIIFMAG 228
Q17758 PEPT2_CAEEL 168 SLFFSMFYFSINAGSLISMWLTPYFRS-MSCF-GHDSCYPLAFGIPAILMIVATLVFMAG 225
Q9ES07 S15A2_MOUSE 181 TRYFSVFYLSINAGSLISTFITPMLRGDVKC--FGEDCYALAFGIPGLLMVLALVVFAMG 238
B0S6T2 S15A2_DANRE 167 RKFFSIFYMSINAGSVLSTIITPILRGDVQC--FGGDCYALAFGVPAALMVIALVVFISG 224
H2DJV9 H2DJV9_FUNHE 161 STFFSIFYLSINAGSLLSTVITPILRAQKCGIHTKQQCYPLAFGVPAALMVVALIVFILG 220
Q8EKT7 Q8EKT7_SHEON 147 QKAFDMFYFTINFGSFFASLSMPLLLKNFG--------AAVAFGIPGVLMFVATVFFWLG 198
Q16348 S15A2_HUMAN 181 TRYFSVFYLSINAGSLISTFITPMLRGDVQC--FGEDCYALAFGVPGLLMVIALVVFAMG 238
P46059 S15A1_HUMAN 160 NRFFSIFYLAINAGSLLSTIITPMLRVQQCGIHSKQACYPLAFGVPAALMAVALIVFVLG 219
O01840 PEPT3_CAEEL 151 SQFFSFFYFAINGGSLFAIIITPILRGRVQCF-GNAHCFPLAFGVPGVLMLLALILFLMG 209
Q21219 PEPT1_CAEEL 182 SLFFSMFYFSINAGSMISTFISPIFRS-QPCL-GQDSCYPMAFGIPAILMIVATLVFMGG 239


P0C2U2 DTPT_LACLA 323 ---------D----PTWFGITFHIDPSWYQLLNPLFIVLLSPIFVRLWNKLGER---Q-- 364
P77304 DTPA_ECOLI 310 ---------S----IL----GLAVEPEQYQALNPFWIIIGSPILAAIYNKMG----DT-- 346
P36837 DTPB_ECOLI 303 ---------E----IL----GFSINPVSFQALNPFWVVLASPILAGIYTHLGNK-GKD-- 342
P75742 DTPD_ECOLI 302 ---------D----MF----GYTVPTAMFQSINAFAVMLCGVFLAWVVKESVAG-NRT-- 341
P39276 DTPC_ECOLI 299 ---------Q----AF----NIEVPTALFQSVNAIAVMLAGVVLAWLASPESRG-NST-- 338
Q5M4H8 Q5M4H8_STRT2 316 ------------------SSWFPVSWFQSLNPLFIMLYTPFFAWLWTAWKKN---Q-- 350
Q5KYD1 Q5KYD1_GEOKA 327 ---------D----VA----GIHLSPAWFQSLNPLFIIILAPVFAWMWVKLGKR---Q-- 364
Q9JXB5 Q9JXB5_NEIMB 317 ---------T----IG----SFTVPVAWKDSMQSLWVILFSGLMAAMWTKMGRK---Q-- 354
Q9P380 PTR2_SCHPO 391 --------------------NVSNDLFQAFDSIALIIFIPICDNIIYPLLRKYNIP-- 426
P32901 PTR2_YEAST 383 --------------------GIPNDFLQAFDSIALIIFIPIFEKFVYPFIRRY-TP-- 417
U9UKY5 U9UKY5_RHIID 343 --------------------NVPNDLIMNLNPISLIILVPFTDHVIYPTLRRFGIT-- 378
Q1MU37 Q1MU37_TRIHA 344 --------------------GVPNDIIQNLNPISIVIMIPLIDHLLYPGLRKIGVA-- 379
Q59ST2 Q59ST2_CANAL 369 --------------------DIPNDFLTVFDSVAIIVFIPIFERFLYPFVRHF-TP-- 403
Q05085 PTR7_ARATH 374 ---------I---G------SFEIPPASMAVFYVGGLLLTTAVYDRVAIRLCKKLFNY-P 414
P46032 PTR2_ARATH 382 ---------I---G------SFQLPPAALGTFDTASVIIWVPLYDRFIVPLARKFTGV-D 422
Q43390 Q43390_BRANA 373 ---------I---G------SFEIPPAAMAVFYIGGLLLTTAVYDRLAIPLCKKLFNY-P 413
Q9FNL7 PTR3_ARATH 364 ---------V---TG-----SFSIPPASLSGFVTLSMLISIVLYDRVFVKITRKFTGN-P 405
Q9M1E2 PTR37_ARATH 351 ---------L---GP-----SFKIPAGSLQVITLLSTCLFIIVNDRVLYPFYQKLTGK-- 391
Q9SX20 PTR18_ARATH 366 ---------L---TN-----SFQIPAGSMSVFTTVAMLTTIIFYDRVFVKVARKFTGL-E 407
Q9M390 PTR1_ARATH 364 ---------M---GK-----NFEIPSASLSLFDTVSVLFWTPVYDQFIIPLARKFTRN-E 405
Q9LFB8 PTR5_ARATH 364 ---------M---GP-----NFKIPSASLSLFDTLSVLFWAPVYDKLIVPFARKYTGH-E 405
Q8IY34 S15A3_HUMAN 352 ANPANISVALRAQGS-----SYTIPEAWLLLANVVVVLILVPLKDRLIDPLLLRCKLL-- 404
A6NIM6 S15A5_HUMAN 341 ---------L-----D----GFLLPIAVMNAISSLPLLILAPFLEYFSTCLFPSK----R 378
Q8N697 S15A4_HUMAN 357 IT-----------TT-----PHTLPAAWLTMFDAVLILLLIPLKDKLVDPILRRHGLL-- 398
Q9LSE8 PTR35_ARATH 332 ---------L----FH----SFEIPVPSLTAIPLIFMLLSIPLYEFFGKKISS-GNNN-R 372
Q9M817 PTR6_ARATH 365 ---------LSRHGS-----SFQVPAGSFGMFTIIALALWVILYDRAVIPLASKIRGR-P 409
Q9LFX9 PTR13_ARATH 358 ---------M---GP-----HFEIPAASITVISYITIGIWVPIYEHLLVPFLWRMRK--- 397
Q8RX77 PTR21_ARATH 396 ---------L---GP-----KFEIPAGSLSVISLLTIGIFLPFYDRVFVPFMRRITGH-K 437
Q8H157 PTR19_ARATH 379 ---------I-----G----SLKIPPASLPIFPVVFIMILAPIYDHLIIPFARKATKT-E 419
Q944G5 PTR44_ARATH 403 ---------L---GSG----GFRIPAATYVVFLMTGMTVFIIFYDRVLVPSLRRVTGL-E 445
Q9SZY4 PTR27_ARATH 364 ---------I---G------SFKIPAGSLTVFFVAAILITLAVYDRAIMPFWKKW-KG-K 403
Q9LQL2 PTR14_ARATH 385 ---------V---S------DFKIPPASMSSFDILSVALFIFLYRRVLEPVANRFKKNGS 426
Q9LV10 PTR53_ARATH 387 ---------L---GSG----GFVIPAATYVVFLMTGMTVFIVVYDRVLVPTMRRITGL-D 429
Q9SPU1 Q9SPU1_ORYSA 382 ---------I---G------SFEIPAASFQSIDVIAVLILVPVYERVLVPVFRKFTGR-A 422
R7W971 R7W971_AEGTA 375 ---------L----H-----HFEIPPASLQGFVTISMLVSVVLYDRLFVPFMRRLTKN-P 415
P51574 S15A1_RAT 315 ---------I----G-----TIEIQPDQMQTVNAILIVIMVPIVDAVVYPLIAKCGFN-- 354
P91679 PEPT1_DROME 322 ---------V----L-----GFQIKPDQMQVVNPLLILGFLPLFDYIIYPALARCGIR-- 361
Q17758 PEPT2_CAEEL 335 ---------V----FG-----FEILPDQMGVLNAFLILFFIPIFQSIVYPTIEKLGFQ-- 374
Q9ES07 S15A2_MOUSE 334 ---------L----G-----FFVLQPDQMQVLNPFLVLVFIPLFDLVIYRLISKCGVN-- 373
B0S6T2 S15A2_DANRE 319 ---------F----GG----GFIIKPDQMQMLNALLILVFIPIFDMGIYPLVGLCRIK-- 359
H2DJV9 H2DJV9_FUNHE 316 ---------F----G-----LMVIQPDQMQTVNPILILIMVPVVDFIIYPLISKCKLN-- 355
Q8EKT7 Q8EKT7_SHEON 333 ---------Q-----------WFEPAMMQALNPLLVMLLIPFNNFVLYPAIERMGVK-- 369
Q16348 S15A2_HUMAN 334 ---------L----G-----FFVLQPDQMQVLNPLLVLIFIPLFDFVIYRLVSKCGIN-- 373
P46059 S15A1_HUMAN 315 ---------I----G-----ALEIQPDQMQTVNAILIVIMVPIFDAVLYPLIAKCGFN-- 354
O01840 PEPT3_CAEEL 305 ---------V----GH-----FSILPEQIHAINPVCVLILVPIFEGWVYPALRK-ITR-- 343
Q21219 PEPT1_CAEEL 354 ---------L----SD----TLLLPDQMQTLNAVLILLFIPLFQVIIYPVAAK-CVR-- 393


P0C2U2 DTPT_LACLA 365 --PSTIVKFGLGLMLTGISYLIMTLPGLL---------N---------- 392
P77304 DTPA_ECOLI 347 --LPMPTKFAIGMVMCSGAFLILPLGAK---------F---------- 373
P36837 DTPB_ECOLI 343 --LSMPMKFTLGMFMCSLGFLTAAAAGMW---------F---------- 370
P75742 DTPD_ECOLI 342 --VRIWGKFALGLGLMSAGFCILTLSARW---------S---------- 369
P39276 DTPC_ECOLI 339 --LRVWLKFAFGLLLMACGFMLLAFDARH---------A---------- 366
Q5M4H8 Q5M4H8_STRT2 351 --PSSPTKFAVGLMFAGLSFLLMAIPGAL---------Y---------- 378
Q5KYD1 Q5KYD1_GEOKA 365 --PTIPQKFALGLLFAGLSFIVILVPGHL---------S---------- 392
Q9JXB5 Q9JXB5_NEIMB 355 --PKTPLKFAMAVFVTGASFLGFV-PFIS---------S---------- 381
Q9P380 PTR2_SCHPO 427 --FKPILRITLGFMFATASMIYAAVLQAKIYQRGPCYAN---------- 463
P32901 PTR2_YEAST 418 --LKPITKIFFGFMFGSFAMTWAAVLQSFVYKAGPWYNE---------- 454
U9UKY5 U9UKY5_RHIID 379 --VRPITRMFIGFMLASLAMAYTAIIQHVIYTTGPCFVN---------- 415
Q1MU37 Q1MU37_TRIHA 380 --FTPIKRMTTGFFIASLSMVASAVMQHYIYKMSPCGDH---------- 416
Q59ST2 Q59ST2_CANAL 404 --FKPITKIFWGFMFGSGAMVYAAVLQHYIYKAGPCYDH---------- 440
Q05085 PTR7_ARATH 415 HGLRPLQRIGLGLFFGSMAMAVAALVELKRLRTAHAH---------- 451
P46032 PTR2_ARATH 423 KGFTEIQRMGIGLFVSVLCMAAAAIVEIIRLHMANDL---------- 459
Q43390 Q43390_BRANA 414 HGLRPLQRIGLGLLLAAMGMAVAALVEIKRLRTAHAH---------- 450
Q9FNL7 PTR3_ARATH 406 RGITLLQRMGIGLIFHILIMIVASVTERYRLKVAADH---------- 442
Q9M1E2 PTR37_ARATH 392 -HLTPLQRVGIGHAFNILSMAVTAIVEAKRLKIVQKGHF---------- 429
Q9SX20 PTR18_ARATH 408 RGITFLHRMGIGFVISIIATLVAGFVEVKRKSVAIEH---------- 444
Q9M390 PTR1_ARATH 406 RGFTQLQRMGIGLVVSIFAMITAGVLEVVRLDYVKTH---------- 442
Q9LFB8 PTR5_ARATH 406 RGFTQLQRIGIGLVISIFSMVSAGILEVARLNYVQTH---------- 442
Q8IY34 S15A3_HUMAN 405 --PSALQKMALGMFFGFTSVIVAGVLEMERLHYIHHNET---------- 441
A6NIM6 S15A5_HUMAN 379 VGSFLSTCIIAGNLFAALSVMIAGFFEIHRKHFPAVEQP---------- 417
Q8N697 S15A4_HUMAN 399 --PSSLKRIAVGMFFVMCSAFAAGILESKRLNLVK-EKT---------- 434
Q9LSE8 PTR35_ARATH 373 SSSFNLKRIGLGLALSSVSMAVSAIVEAKRKHEVV---------- 407
Q9M817 PTR6_ARATH 410 FRLSVKLRMGLGLFMSFLAMAISAMVESFRRKKAISQ---------- 446
Q9LFX9 PTR13_ARATH 398 FRVTLLQRMGIGIVFAILSMFTAGFVEGVRRTRA---------- 431
Q8RX77 PTR21_ARATH 438 SGITLLQRIGTGIVFAIFSMIVAGIVERMRRIRSINA---------- 474
Q8H157 PTR19_ARATH 420 TGVTHLQRIGVGLVLSILAMAVAALVEIKRKGVAKDS---------- 456
Q944G5 PTR44_ARATH 446 TGISLLQRIGAGFTFAIMSLLVSGFIEERRRNFALTKPT---------- 484
Q9SZY4 PTR27_ARATH 404 PGFSSLQRIAIGLVLSTAGMAAAALVEQKRLSVAK---------- 438
Q9LQL2 PTR14_ARATH 427 KGITELHRMGIGLVIAVIAMIAAGIVECYRLKYADKS---------- 463
Q9LV10 PTR53_ARATH 430 TGITLLQRIGTGIFFATASLVVAGFVEERRRTFALTKPT---------- 468
Q9SPU1 Q9SPU1_ORYSA 423 NGITPLQRMGIGLFFSMLSMVSAALVESNRLRIAQDE---------- 459
R7W971 R7W971_AEGTA 416 RGISLLQRMGVGLVFHIVIMVIASLTERHRLRVAMEN---------- 452
P51574 S15A1_RAT 355 --FTSLKKMTVGMFLASMAFVVAAIVQVEIDKTLPVFPS--GNQVQIKVLNIGNNDMAVY 410
P91679 PEPT1_DROME 362 ---RPLQKLTLGLLLAALGFFLSAGLEMKMEQAAYRATPIEPDMTHLRIYNGMPCRYEIS 418
Q17758 PEPT2_CAEEL 375 --MTMLRKMAGGGILTAVSFFVCGIVQLFVNPTLPYIPM--ANEAHLTIINTIP-SCDFN 429
Q9ES07 S15A2_MOUSE 374 --FSSLRKMAVGMILACLAFAVAALVEIKINGMIHPQPA--SQEIFLQVLNLADGEIEVT 429
B0S6T2 S15A2_DANRE 360 --LTPLKKMATGMILAALAFCAATAVEVYVIKTVVEPPP--AKESLVQVYNLMDSDVTVQ 415
H2DJV9 H2DJV9_FUNHE 356 --FTPLRKMTGGMFLAALAFVAAALVQLQIDATLPTFPS--SNEQQAKFINMVNRNLSIA 411
Q8EKT7 Q8EKT7_SHEON 370 --LTALRKMGAGIAITGLSWIVVGTIQLMMD---------- 398
Q16348 S15A2_HUMAN 374 --FSSLRKMAVGMILACLAFAVAAAVEIKINEMAPAQPG--PQEVFLQVLNLADDEVKVT 429
P46059 S15A1_HUMAN 355 --FTSLKKMAVGMVLASMAFVVAAIVQVEIDKTLPVFPK--GNEVQIKVLNIGNNTMNIS 410
O01840 PEPT3_CAEEL 344 --VTPLRKMAVGGLLTAFSFAIAGVLQLKVNETMEFPPS--LGRIYLQRVGNESLISDFR 399
Q21219 PEPT1_CAEEL 394 --LTPLRKMVTGGLLASLAFLITGFVQLQVNTTLPTLPE--EGEASISFWNQFETDCTIT 449


P0C2U2 DTPT_LACLA 403 LVLMFAVQMAGELLVSPVGLSVSTKLAPVAFQSQMMAMWFLADST-SQAINAQI------ 455
P77304 DTPA_ECOLI 385 LVASYGLQSIGELMISGLGLAMVAQLVPQRLMGFIMGSWFLTT-AGANLIGGYVAGMMAV 443
P36837 DTPB_ECOLI 382 IVLVYLFQSLGELFISALGLAMIAALVPQHLMGFILGMWFLTQ-AAAFLLGGYVATFTAV 440
P75742 DTPD_ECOLI 380 MVLGLAVMGFAELFIDPVAMSQITRIEIPGVTGVLTGIYMLLSGAIANYLAGVIADQTSQ 439
P39276 DTPC_ECOLI 377 MISGLALMGFAELFIDPVAIAQITRLKMS---GVLTGIYMLATGAVANWLAGVVAQQTTE 433
Q5M4H8 Q5M4H8_STRT2 389 LVGSWALVILGEMLISPVGLSVTTKLAPKAFNSQMMSMWFLSSSV-GSALNAQL------ 441
Q5KYD1 Q5KYD1_GEOKA 402 LVLSYFIVVLGELCLSPVGLSATTKLAPAAFSAQTMSLWFLSNAA-AQAINAQL------ 454
Q9JXB5 Q9JXB5_NEIMB 390 FALIVLAITIGELMISPIALSISTKIAPPLFKTQMVALNFLAFSL-GFTLGGVLF----- 443
Q9P380 PTR2_SCHPO 478 QIPAYVLIAFSEIFASITGLEFAFTKAPPSMKSIITALFLFTNA-FGAILSICISSTAVN 536
P32901 PTR2_YEAST 469 QIPAYVLISFSEIFASITGLEYAYSKAPASMKSFIMSIFLLTNA-FGSAIGCALSPVTVD 527
U9UKY5 U9UKY5_RHIID 434 QSPSYVLIGFSEVFASVTGLSYAYERAPDEMKSTVMAIFLTMSA-FGSALGFAFVPLSKD 492
Q1MU37 Q1MU37_TRIHA 440 QSLPYILIGFAEIFANVTSYEYAYSKAPENMKSLVMSVNLFMSA-ISAAIGEAFTPLSDD 498
Q59ST2 Q59ST2_CANAL 455 QTPAYVLIAISEILASITGLEYAYTKAPVQMKSLVMAMFLLTNA-VGAAIGIALSSVSVD 513
Q05085 PTR7_ARATH 465 LIPQYLIVGIGEALIYTGQLDFFLRECPKGMKGMSTGLLLSTLA-LGFFFSSVLVTIVEK 523
P46032 PTR2_ARATH 475 QIPQYFILGAAEVFYFIGQLEFFYDQSPDAMRSLCSALALLTNA-LGNYLSSLILTLVTY 533
Q43390 Q43390_BRANA 464 LIPQYLIVGIGEALIYTGQLDFFLRECPKGMKTMSTGLLLSTLA-LGFFFSSVLVTIVEK 522
Q9FNL7 PTR3_ARATH 459 LLPQFVLMGMADSFLEVAKLEFFYDQAPESMKSLGTSYSTTSLA-IGNFMSSFLLSTVSE 517
Q9M1E2 PTR37_ARATH 443 LFPPLVIVGIGEAFHFPGNVALCYQEFPESMRSTATSITSVVIG-ICFYTSTALIDLIQR 501
Q9SX20 PTR18_ARATH 461 LIPQYGLHGVAEAFMSIGHLEFFYDQAPESMRSTATALFWMAIS-IGNYVSTLLVTLVHK 519
Q9M390 PTR1_ARATH 457 QIPQYLLIGCAEVFTFIGQLEFFYDQAPDAMRSLCSALSLTTVA-LGNYLSTVLVTVVMK 515
Q9LFB8 PTR5_ARATH 457 QVPQYFLVGCAEVFTFIGQLEFFYDQAPDAMRSLCSALSLTAIA-FGNYLSTFLVTLVTK 515
Q8IY34 S15A3_HUMAN 461 QIPQYLLIGISEIFASIPGLEFAYSEAPRSMQGAIMGIFFCLSG-VGSLLGSSLVALLSL 519
A6NIM6 S15A5_HUMAN 433 LILQYVLLGVAETLVNPALSVISYRFVPSNVRGTSMNFLTLFNG-FGCFTGALLVKLVYL 491
Q8N697 S15A4_HUMAN 454 QVPQYLLIGISEIFASIAGLEFAYSAAPKSMQSAIMGLFFFFSG-VGSFVGSGLLALVSI 512
Q9LSE8 PTR35_ARATH 418 LVFQYLMLSVSDMLTLGGMLEFFYREAPSNMKSISTALGWCSTA-LGFFLSTTLVEVTNA 476
Q9M817 PTR6_ARATH 463 LVPQYVLHGLAEALTAIGQTEFFYTEFPKSMSSIAASLFGLGMA-VASLLASVVLNAVNE 521
Q9LFX9 PTR13_ARATH 442 LALPLILMGLCESFNFIGLIEFFNSQFPEHMRSIANSLFPLSFA-AANYLSSLLVTTVHK 500
Q8RX77 PTR21_ARATH 488 LSPQLILMGLCEAFNIIGQIEFFNSQFPEHMRSIANSLFSLSFA-GSSYLSSFLVTVVHK 546
Q8H157 PTR19_ARATH 472 IALQYLFLGSADLFTLAGLLEYFFTEAPSSMRSLATSLSWASLA-MGYYLSSVIVSIVNS 530
Q944G5 PTR44_ARATH 502 LIPQLTLAGIAEAFAAIGQMEFYYKQFPENMKSFAGSIFYVGAG-VSSYLASFLISTVHR 560
Q9SZY4 PTR27_ARATH 452 LVPQFFLVGAGEAFIYTGQLDFFITQSPKGMKTMSTGLFLTTLS-LGFFVSSFLVSIVKR 510
Q9LQL2 PTR14_ARATH 478 QAPQYSLIGASEVFMYVGQLEFFNAQTPDGLKSFGSALCMMSMS-MGNFVSSLLVTMVVK 536
Q9LV10 PTR53_ARATH 486 LIPQLSLAGVAEAFAAIGQMEFYYKQFPENMRSFAGSIFYVGGG-VSSYLGSFLIATVHR 544
Q9SPU1 Q9SPU1_ORYSA 475 QGPQYFLIGVGEVFSNIGLTEFFYQESPDAMRSLCLAFSLANVS-AGSYLSSFIVSLVPV 533
R7W971 R7W971_AEGTA 469 LLPQFVLMGVADAFLEVAKIEFFYDQAPEGMKSLGTSYSMTSLG-IGNFLSSFLLSTVSR 527
P51574 S15A1_RAT 586 QIPQYFLLTCGEVVFSVTGLEFSYSQAPSNMKSVLQAGWLLTVA-IGNIIVLIVAEA--- 641
P91679 PEPT1_DROME 596 QLPQIVVMTAAEVMFSVTGLEFSYSQSPPSMKSVLQACWLLSVA-IGNMLVVVIAEF--- 651
Q17758 PEPT2_CAEEL 673 QIPQYVILTAGEVLFSITGLEFAYTEASPQLKSVVQALWLFTTA-IGDLIVVVIFML--- 728
Q9ES07 S15A2_MOUSE 611 QLPQYVLVTAAEVMFSVTGLEFSYSQAPSSMKSVLQAAWLLTVA-VGNIIVLIVAQF--- 666
B0S6T2 S15A2_DANRE 594 QIPQYVFLTAGEVMFSITGLEFSYSQAPASMKSVLQAGWLMTVA-FGNVIVLIVAEG--- 649
H2DJV9 H2DJV9_FUNHE 581 QIPQYFLMTSGEVFFSVTGLEFSYSQAPSNMKSVLQAGWLLTVA-VGNIIVLIVAEA--- 636
Q8EKT7 Q8EKT7_SHEON 408 QILPYALLTFGEVLVSATGLEFAYSQAPKAMKGTIMSFWTLSVT-VGNLWVLLANVSVKS 466
Q16348 S15A2_HUMAN 611 QLPQYALVTAGEVMFSVTGLEFSYSQAPSSMKSVLQAAWLLTIA-VGNIIVLVVAQF--- 666
P46059 S15A1_HUMAN 584 QIPQYFLLTCGEVVFSVTGLEFSYSQAPSNMKSVLQAGWLLTVA-VGNIIVLIVAGA--- 639
O01840 PEPT3_CAEEL 578 SLPQYIIITLGEVLLSVTGLEFAYSQAAPNMKSVLTAMWLLTVF-AGNLIDMMISGT--- 633
Q21219 PEPT1_CAEEL 700 QIPQIVVITAAEILFSITGYEFAYSQSAPSMKALVQALWLLTTA-AGDSIIVVITIL--- 755