

Determination of the 3D structure of the protein PepT1

I. Introduction

The protein Solute carrier Family 15 member 1 (PepT1, HpepT1), is a human intestinal dipeptide transporter (UNIPROT accessibily number : P46059). This protein is a symporter, that means proton-coupled intake of oligopeptides of 2 to 4 amino acids but PepT1 have a preference for dipetides.

PepT1 is encode by the gene SLC15A1 in the human genome, more specifically in the 13 chromosome. It is a protein composed by 708 amino acid and known to have 12 transmembrane segment [11].

This protein is a part of the Major Facilitator Superfamily (MFS). MFS is one of the biggest group of secondary active transporters conserved in deffrents species (bacteria, plant, human) [9]. MFS is a superfamily of membrane transport proteins that permit the movement of a large spectrum of small solutes (like glucose) to cross the membrane (to the cytoplasme or to the outside) according to a chemiostic gradients. MFS play also a crucial role in multiple physical process because the movement of solutes in or out the cell is needed to the survival of all cells[9].

Some MFS active transports are called symporters. PepT1 is a symporter this mean that this protein allows the crossing of two or more substates in the same directions.

The Solute carrier family 15 member 1 is also a part of the subfamily POT (Proton-dependant oligopeptide transporter), a proton will be involve every time a solute need to cross the membrane.

The solute carrier family 15 member 1 is located un the bross membrane of the intertinal epithelium and have a crucial role in the absorption et the digestion of food protein. This protein facilitate the absorption od a lot of drugs.

II. Materiels and Methods:

A. Determination of the Template

The first step is to determinate the sequence template usable. For that a BLASTp is executed directly from the uniprot page of P46059. A blastp is a blast of protein sequence. Blast is a basic local alignment search tool, and permit to find sequences having areas of similarity with the input sequence. This will give an alignment between the sequence PepT1 and other sequences which have a 3D strutures associated (pdb file). For that the parameters « with 3D structure » in target database is required to aligne our sequence with only sequence having a pub file associated, and a blossum 62 (BLOcks SUbsitution Matrix) is use as matrix.

B. Secondary Structure

In the second step, we search the secondary structure of the PepT1. For that PsiPred is used. PsiPred give secondary structure of the query sequence, here the fasta sequence of PepT1.

We realize a PsiPred and also a MEMSAT3 & MEMSAT-SVM.

MEMSAT3 & MEMSAT-SVM [10] make prediction of the helix in the membrane and MEMSAT-SVM improved transmembrane protein topology prediction using SVMs. This method is also able to determinate the pore lining, the loop in the cytoplasm compartment and in the extracellular compartment.

A other method to determinate the secondary structure is HCA (Hydrophobic Cluster Analysis). This program take only in input the fasta sequence of PepT1 and gives information of the secondary structure by creating cluster.

An other tool to detect conserved areas is Tcoffee (Multiple Sequence Alignment Server). More specifically PSI/TM-Coffee is use because is specify for transmembrane protein. This program permit to find conserved areas between sequences. As input a file containing alignment between the query sequence (PepT1), and 5 homologues sequences with a identity pourcentage about 40%.

And we finish this part with OCTOPUS. OCTOPUS will take in input de sequence fasta of PepT1 and will generate a topology prediction.

C. Models

The next step is to generate models. For doing that, two methods exist, comparative homology and threading. Comparative homology can be used when the template have a poucentage identity superior to 30%. With the sequence previously choose as template both methods can be used.

a) Comparative Modelling

First MEDELLER. Medeller is a Homology-Based Coordinate Generation for Membrane Proteins, it is specialized for membrane protein [7].It is taking as input a fasta file containing the query sequence and the sequence of the template.

And other comparative modeling program can be used, MEMOIR [6]. Is an homology modeling algorithme designed for membrane protein. It integrate different software (4) like MEDELLER. In input MEMOIR take the query fasta sequence and the pub file of the templete choose previously.

Than Swiss model can be use. Swiss Model is a fully automated protein structure homology -modalling serveur. It can be use for comparative modeling. In our case Swiss Model is use with only one input sequence the query fasta sequence of PepT1.

Phyre2 is a protein/homology/Analogy Recognition Engine, that take in input se sequence fasta od pepT1, as parameter intensive ls choose, even if is take more time this allows to have more accurate result.

Next, HHpred. It is a homology model using HMM/Modeller. It is taking in input the fast sequence of pepT1.

b) Threading

A program to do threading it is Itasser (Iterative Threading ASSEmbly Refinement). With only in input the fasta sequence of PepT1, iTasser identifies structural templates which have pdb by using LOMETS, a multiple threading.

Furthermore, RaptorX Structure Prediction is used. This created template based on tertiary structure prediction. As input it takes de fast sequence of pept1.

Muster is used, is a Multi-Source ThreaDER program. Muster take in input the fasta sequence of PepT1 and it will identify the differents sequences which can be template from the PDB library.

D. Classification ans Refinement

To find which is the best model we need to order the difference models created previously. For that OREMPRO is used because is adapted to the transmembrane protein. Indeed OREMPRO take in input the pdb of the model and return a MAIDEN Score and the position of the model in a membrane. More the score is low, more the model is accurate.

Once the best model is selected it is needed to affine it. For that Galaxy refine is use [8]. This method can improve the structure quality. This step will be follow by a other step of classification (with OREMPRO) because Galaxy Refine Web will create several models.

E. Normal Mode

For the determination of the normal mode of the best model, NOMAD-ref will be use. As input only the pdb of the model is needed, the parameters used are the default ones except the number of mode choose (25 instead 16).

III. Results**A. Determination of the Template**

- BLASTp : 2 sequences are interesting, Q63424 and Q8EKT7. The first one (Q63424) have a high coverage with the sequence input and a identity percentage of 48%. The second one (Q8EKT7) have a coverage shorter than the previous and a identity percentage of 29,9%. But the first sequence template Q63424 have a pdb with only the position 410 to 601 so it can not be used like template. Instead Q8EKT7 have a pqr with the position 4-516. So Q8EKT7 is chosen to be the sequence template more specifically the pub 4uvn is chosen. So comparative modeling and threading can be tried.

B. Secondary Structure

- PsiPred : PsiPred detect 16 helix (figure1 in pink) and 30 sheet (figure1 in yellow). This is correlated to the distribution of the extracellular region, transmembrane helix and cytoplasmic region shown in the figure2. In the figure2 two major loops can be seen, one on the cytoplasmic region (221-277) and one on the extracellular region (332-587). These two loops are also found with MEMSAT-SVM.

- MEMSAT-SVM : This method determinate 12 helix and the Nter and Cter in the cytoplasmic side of the membrane (figure3). The two big loop find previously are also present. The loop compose by almost 200 amino acid is find between the segment 9 and 10 [5]

So MEMSAT-SVM and PsiPred do not find the same number of helix, but PepT1 is known to have 12 transmembrane segments. Furthermore the Nter and Cter [9] are in the same compartment and a big loop is expected between the helix9 and helix10. So the prediction of MEMSAT-SVM seems to be more accurate than the one does by PsiPred.

Helix (position) figure1	MEMSAT-SVM Transmembrane Helix Prediction Figure2
Ile19-Phe28	16-38
Ile38-Phe41	
Phe59-Ser76	56-74
Leu87-Ile93	83-102
Val120-Gly135	124-142
Gly154-Ile179	164-186
Phe202-Leu213	198-219
Lys235-Arg249	277-297
Glu271-Asp298	
Arg303-Thr310	326-348
Tyr345-Ala439	
Leu358-Asp383	359-380
	587-603
Lys615-Val626	619-637
Gln645-Ala667	648-668
Asp685-Lys688	
Table1 : Abstract of the position of trasmembrane helix found by PsiPred and MEMSAT_SVM	

•

Secondary Structure Map

Feature predictions are colour coded onto the sequence according to the sequence feature key shown below.

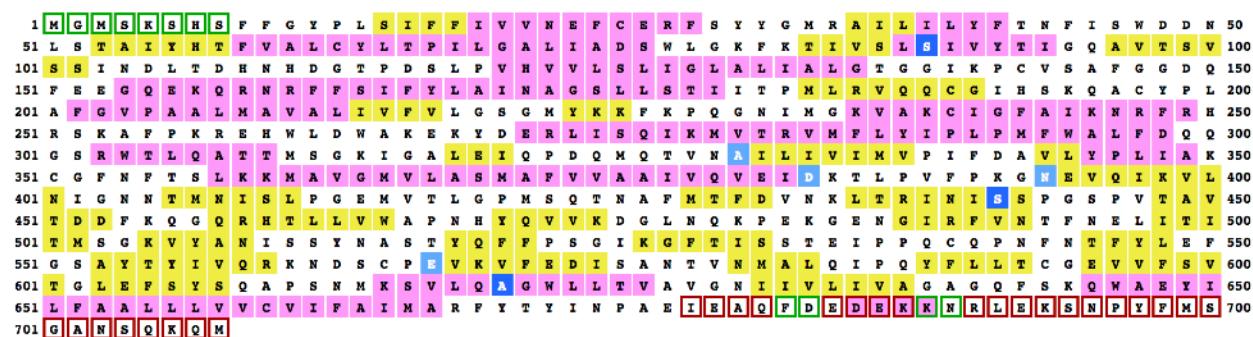


Figure1 : Secondary Structure Map (PsiPred Map) of the PepT1 protein

TM Helix Map

Feature predictions are colour coded onto the sequence according to the sequence feature key shown below.

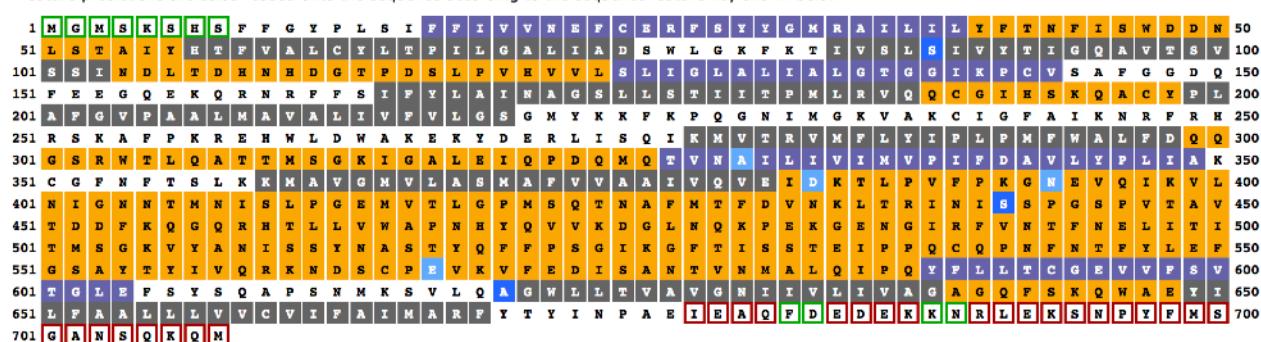


Figure2 : Transmembrane Helix Map (MEMSAT Map) of the PepT1 protein

7df1dbb0-6d46-4b85-8569-5d41e3bf68be.seq.job

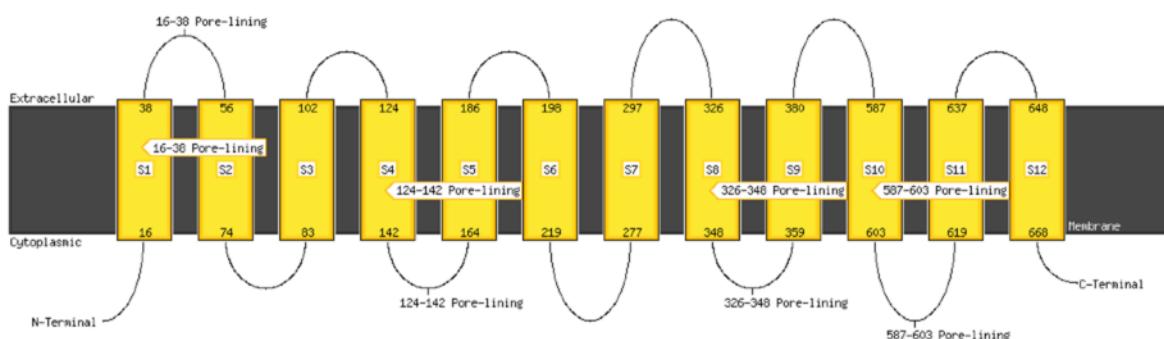


Figure3 : MEMSAT-SVM Cartoon of the PepT1 protein

- OCTOPUS give a topology prediction (figure4). This prediction show 12 TM-helix, the Nter and Cter on the same side which mean on the inside on the cell. This figure 4 show also the 2 big loop previously enunciated, on the inside et one very big loop on this outside and between the TM-helix 9 and 10.

This result can be correlated to the MEMSAT-SVM Cartoon of the PepT1 protein.

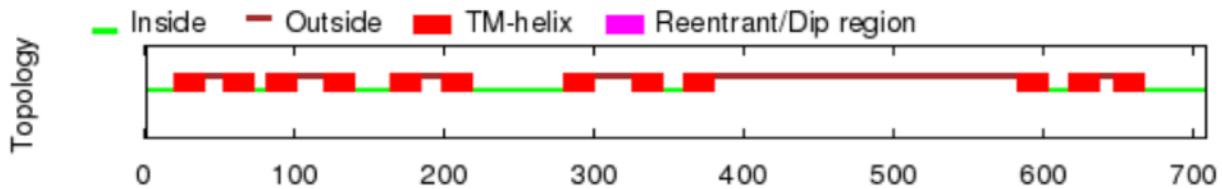


Figure4 : Topologie Prediction of PepT1 (OCTOPUS)

- HCA (figure5) : The form of the amino acid heap is indictable of the nature of the secondary structure which is associated. A green vertical heap is associated to a sheet and a chorizontal heap (yellow and red) is associated to a helix. the sequence that separates heaps correspond to loop region, connecting secondary structure.

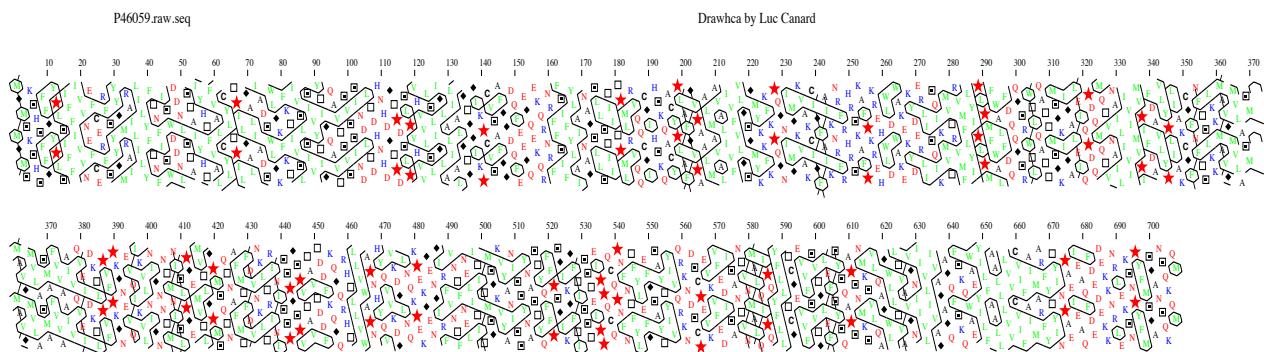


Figure5 : HCA of the PepT1 protein

- Tcoffee : The alignment between PepT1 et 5 other homologue sequences (**A0A151I3X9** : Solute carrier family 15 member 1 of *Atta colombica*; **A0A087UA70** : Solute carrier family 15 member 2 of *Stegodyphus mimosarum*; **A0A0P5Y532** : Solute carrier family 15 member of *Daphnia magna*; **W8C2V3** : Peptide transporter family 1 of *Ceratitis capitata*; **A0A1D2N6G6** : Peptide transporter family 1 of *Orchesella cincta*) in the figure 5 permit to show the conserved areas. Some areas are known to be conserved in different species like the transmembrane segment 5 [2], which in regards of the result of MEMSAT-SVM shown previously should be around the position 164 and 186 and we can see 14 amino acid conserved (* means conserved). There is also a conserved motif, 22ExxERFxYY 30, shown in the article [9] which is found in the figure6 (in the red box).The figure6 show that a lot of conserved amino acid are located inside

T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)
Cedric Notredame

IN	HEL	OUT	
P46059	1 MGM-		
A0A151I3X9	1 MSDNLHLVLRTKR-	RDOCPSYHRGNKLN	GAESLKMST
A0A087UA70	1 MOKLYVCW - KFN-	ENCCR - M	
A0A0P5Y532	1 MSSNN-		
W8C2V3	1 MTEENEHK - QNNGDVQTTDPEIAANEKSSANDVQESRV	VEPEYGLEKAIVPALGLH	5
A0A1d2N6G6	1 MVK-		
	1 *		

P46059	395	V01KVLNIGNNTMNSLPGEMT	---	LGPMSOTNAFMFTDVNLKLTR	--INISPPGSPVT	A	-	VT	45
A0A151I3X9	434	A0IRIFNTMNCVTSVKLDONVLDRNNIISPLMSSE	---	VLEVKVNESIKISYTAFDKDCYDNGYITLLKKNE	-----	-----	-----	-----	50
A0A087UA70	415	TEMIIINSPCSIKINKVGNLTV	-DTTEIAFGNKI	VGNINTRNLN-----DWEFTVNCTAL-----	-----	-----	-----	-----	47
A0A0P5Y532	384	TOLNFINTLPCPNLISYINNNQKWIIEINATSYF	-----	ERNLNNLSIDTIAOLISSPACOOMNFT	-IPTWT	-----	-----	-----	45
W8C2V3	454	AOLRVYNGEPCDYAITTNLTG	-L	GFDVATLDMFV	-----	NRSISIDYTNKLVYOFESOKSGVNCYEFPKOT	-----	-----	51
A0A1D2N6G6	381	AHVNFINGLPCPVKMELEGNNST	-Q	-IEAFGQHQI	-----	LKQIEAGKTYATKFTPAGEGES	--YK	--PFD	44
	484 * :	:	:	55

P46059	452	DDF--KOGORHTLLVWAPNHY--QVV--KDGLNOKPEKGENGIRFVNFT-NELI--TITMSG--	50
A0A1511X39	502	GTMVEEASATSWITPLGLE--YY--YKDTVKNPKEGKPMVRSVIY-ES-ESE--NATLKLKYKG	56
A0A087UA70	477	KTFN-STTDIESMITHDAGKLEVNL-TNDTKIAKSKGKPRVIRRFFNT-EYDFSKOPNSMMVK--KE	53
A0A0P5Y532	451	GT1KGASTKAFAVIVTGHNSSLRMRMNEEPLDKSSTGOPRVGFVF0--DODVPNQNQGNITLKGAG	51
W8C2V3	520	FVF--NISTSHTHLYFMMSNKNTAOPLVW--AEDEISKPSRGYPLARTLANV-OSSR--TIEWRNK	57
A0A01D2N6G6	441	YQLNARNEKEYRSYLIMNSEGTIASIKEATSSDKLKGEGASKIKFVLFGDS1RSKESIQFKHG	50
	553	:	62

P46059	505	-KVK-		-ANISSY-	-NASTY	YOFFPSGIKGF	-T-	-ISSTEIPPOC	535
A0A151I3X9	561	NTKTVS	IPV	SFNFSESKLKEIEDPSEYDVYLD-					E-59
A0A087UA70	540	DKRY	LIETENLEPTLGLTSYSEF	-APGTYDHTPL			-N-	ATS-HN	589
A0A0P5Y532	515	ERIVSFTR	KLIAS	SSLFIOTDVREI	DAGTYDVFM		LRNGDV-	VM	557
W8C2V3	576	KGSL	EH	TEPANRNRVTEL	KSGYYDILV-			-D	603
A0A1D2N6G6	506	DSFMKVWQEPEVEDDGHGKKA		AVSVGDGPVSEF	DSGNYKTPHEQLVLVMTPAIESKDIS-KK				568
	622			.	*				690

P46059	540	OPNFNTFYLEFGSAYTYIVOR-KN--DS--	-CPEVKVFDISANTVNMLAOIPOYFLTCGEVVF	59
A0A151I3X9	594	LLDKNPVFPKSGGVTYVGSRVMYNNKP	-KIVGKTVTVPPTPPNSLHMAWMLPOVIITMAEVMS	65
A0A087UA70	581	EMPGVNRTRVRSGGAYIVSIYONSP-SDG	-NFSVGTTIMIEEINSIHFFIOPIPOVYIMTAGEIMFS	64
A0A0P5Y532	558	ENLIGNITVVOGGSYTIVAO0SDF	-NAQLMNLNNNLLPLIEVTPPSSVHMLWLVP0YFVMTVAEVMF	62
W8C2V3	604	NVVANAEYLKVGGGIYTIIINE-ET-AG	-SYKSXIVTVDPMNSMSIFWLWIPOYFVMTLGEVMF	66
A0A1D2N6G6	569	YVTIGNIKPVQGGNYYYIKKESG-AD	-KFDIKEYVITQPNSIHLMLWLVP0YVIIIGEVMFS	62
	691	.	.*.*	.*.*:.*.*:***:***

Figure6 : T-Coffee realised on the alignment between the sequence of PepT1 and 5 homologues sequences (40% identity with PepT1 from different species)

the membrane so inside the helix. This is coherent by the fact that helix have a crucial role to the transport of dipeptide.

C. Models

- MEDELLER : None pbb were generated, this is maybe due to the fact that the template only covered the first 66% of the query sequence. So utilization of other method is needed.
- MEMOIR : As output 4 pdb are created but when they are visualized in pymol the loop are absent. So there are not good models.
- SWISS MODE : As output 5 models are created. The fist 3 have the higher coverage. The fist model is created by using as template 4uvm, which is the same template choose with the BLASTp. The second and third model have an alignment shorter comparated to the model1. But the model1 have a Qmean very low (table2), this can be explain by the fact that PepT1 is known to have a big loop (190 amino acids) between helix 9 and 10 [1] [4].

Pdb name		Pourcentage identity	Sequence alignement	Qmean
Model1	4uvm	34.46%	10-672	-7.23
Model2	4ikv	35.85%	7-373	-5.39
Model3	4w6v	30.0%	6-374	-5.92

Table2: Result SwissModel

- PHYRE2: One model is created by using 5 templates, selected by phrey2 (table3). We can see that the templates used in model 2 and 3 created by Swiss model, respectively 4ikv and 4w6v, are used by Pyre2.

Pbd name	Confidance (in %)	Identity pourcentage
4iky_A	100	28%
5cl5_B	100	30%
2xut_C	100	38%
4w6v_A	100	28%
4aps_B	100	26%

Table3 : Template use by PHYRE2

- raptorX : Give one model, but it used several templates to created it (table4). In fact it divided the query sequence in 3 domains and find template to each domain. This allows to have a model more accurate and more precise to define the big loop of 190 amino acide present between the positions 380 and 580 . We can see that some template are the same previously seen.

	Positions	Templates
Domain 1	386-583	5a9d
Domain 2	1-385	2xut_A
Domain 3	584-708	6ei3_A 4oh3_A 4ikx_A 2xut_A

Table4 : Templates associated to the different domains find by raptorX

- hhpred : First this sort a lot of sequence matching to the fasta given in input (PepT1 sequence). I choose 4 models that had about 25/30% identity with pepT1 (4uvn, 4oh3, 4ikv, 4ikw) and a pdb associated containing almost all the position for the all sequence. HHpred find 12 helix in the query sequence. And this selection, MODELLER is launch which the query sequence and the sequences selected. This give a model.
- MUSTER : 10 models were generated but 3 sequences have been characterize « bad » and a few have a pourcentage identify superior to 20% and a coverage superior to 60%. Only the first 4 (figure7) will be verify because they have the best coverage (>60%) and the best identity sequence (>23%)

Rank	Template	Align_length	Coverage	Zscore	Seq_id	Type
1	4cl4A	449	0.634	14.185	0.263	Good
2	4ikvA	452	0.638	13.093	0.246	Good
3	2xutA	419	0.591	12.804	0.337	Good
4	4w6vA	449	0.634	12.336	0.238	Good

Figure7 : The first four models generated by MUSTER

D. Classification ans Refinement

OREMPRO : Two other algorithm had to be used but there are not as accurate than OREMPRO with the MAIDEN score.

Methods	Pdb / template use	MAIDEN score	OREMPRO	
			Transmenbranaire	Segment
Modeller	None	None	None	
MEMOIR	4 pdb but loop absent	None	None	
Itasser	Model1	2.0073		
	Model2	Do not work		
	Model3	Do not work		
	Model4	Do not work		
	Model5	Do not work		
Swiss model	Model1: 4uvm	-113.7263		12
	Model2 : 4ikv	-4.6409		7
	Model3: 4w6v	-11.2206		8
Phyre2	1 pub & 5 templates used : 4ikv 5cl5 2xut_C 4w6v_A 5aps_B	-73.7177		11
HHpred	1 pdb & 4 templates 4uvm 4oh3 4ikw 4ikv	-11.4485		4
RaptorX	1 pdb (division in domains)	-12.317		4
Muster	Model1 template 4cl4_A	Do not work		
	Model2 template 4ikv_A	Do not work		
	Model3 template 2xut_A	Do not work		
	Model4 template 4w6v_A	Do not work		

Table5 : Classification of the models created by the different programs used previously.

Some of the models generated can not be classified due to the fact there are problems in the sequence (like an absent of loop).

The lowest MAIDEN score is -113,7263. Or more the score given by MAIDEN is low more the model is accurate. And it is the only model where 12 transmembrane segments were found with OREMPRO (figure8). Or the protein PepT1 is known to have 12 transmembrane segments. And the Nter and Cter are in the same compartment, the

cytoplasmatic which is also expected. So the more accurate is the model 1 generated by SwissModel (figure9) by using as template 4uvm.

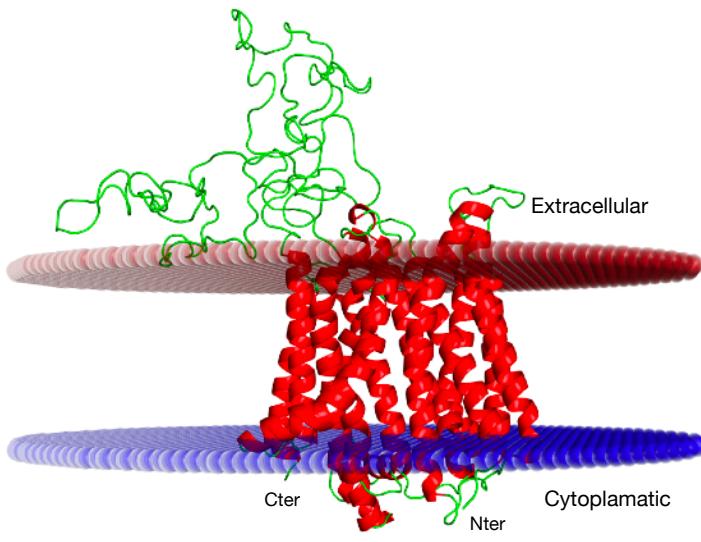


Figure8 : Representation of the model 1 generated by Swiss Model with the template 4uvm after utilisation of OREMPRED



Figure9 : Representation of the model 1 generated by Swiss Model with the template 4uvm

Galaxy Refine Web:

Once the best model from all the models generated is selected a step of refinement needs to be done. For that Galaxy Refine Web is used. It gives in output 5 models refine.

First each model newly created is align to the model 1 of Swiss Model before the refinement, to see where the refinement was done (figures 10 to 14). The Root Mean Square Deviation between the model before refinement et the models after refinement are of the order of 0.3. And for all the 5 models the refinement show the apparition of 1 or 2 (depends the model) helix in the big loop extracellular. This can be explain by the fact that this portion are highly mobile and can at some point adopt the helix conformation as a transition conformation.

Then this 5 models are range in function on the MAIDEN score (table6).

Models	MAIDEN Score	TM
Model1	-120.2927	12
Model2	-119.7790	12
Model3	-104.566	12
Model4	-109.3487	12
Model5	-116.7741	12

Table6: Comparaison MAIDEN score (OREMPRO) of the 5 models generated by Galaxy Refine Web.

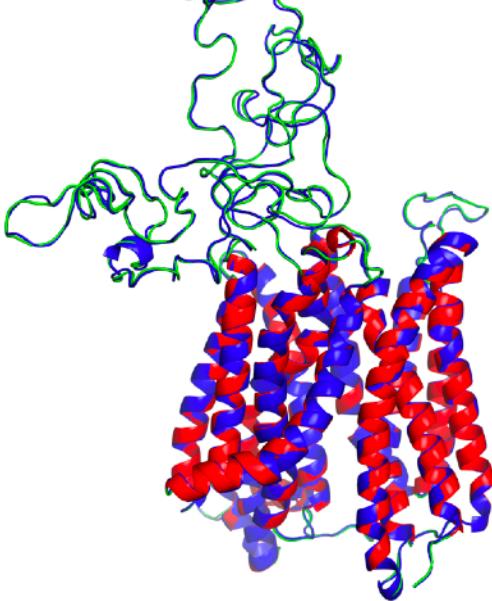


Figure 10 : Superposition between the model 1 generated by Swiss Model (red and green) and the same model but after refinement, model1 of the 5 (blue), RMSD of 0.323 find with pymol, 0,412 with Galaxy Refine Web

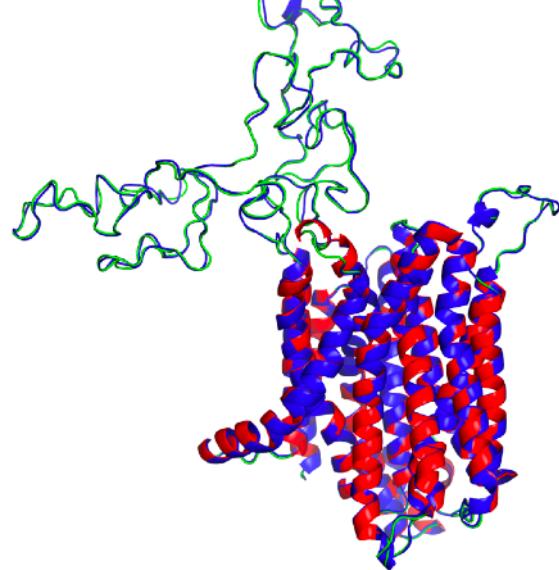


Figure 11 : Superposition between the model 1 generated by Swiss Model (red and green) and the same model but after refinement, model2 of the 5 (blue), RMSD of 0.330 find with pymol, 0,421 with Galaxy Refine Web

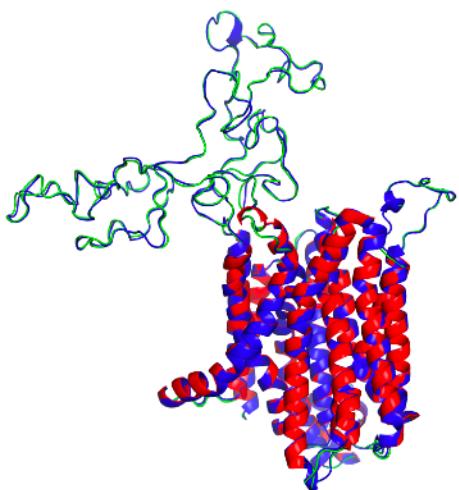


Figure12: Superposition between the model 1 generated by Swiss Model (red and green) and the same model but after refinement, model3 of the 5 (blue), RMSD of 0.328 find with pymol, 0,415 with Galaxy Refine Web

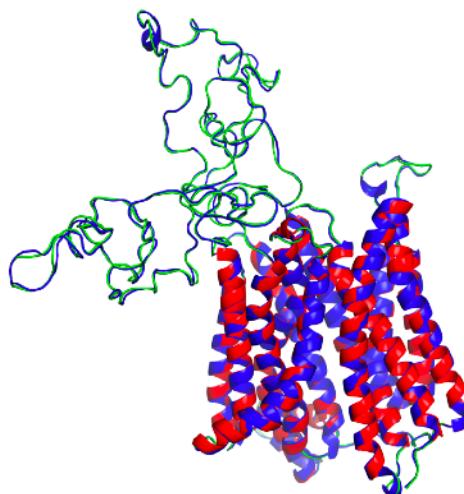


Figure 13: Superposition between the model 1 generated by Swiss Model (red and green) and the same model but after refinement, model4 of the 5 (blue), RMSD of 0.322 find with pymol, 0,414 with Galaxy Refine Web

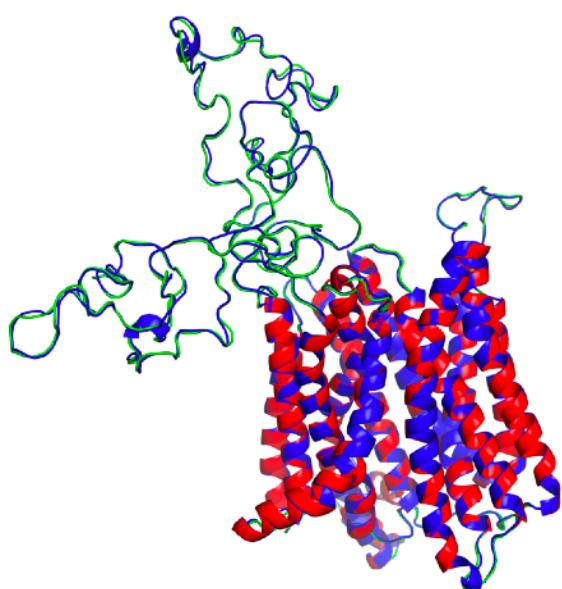


Figure14: Superposition between the model 1 generated by Swiss Model (red and green) and the same model but after refinement, model5 of the 5 (blue), RMSD of 0.349 find with pymol, 0,430 with Galaxy Refine Web

.ter 2 BI

Before the refinement, the MAIDEN score of the model choose (model1 SwissModel) was -113,7263. After the refinement 2 of the 5 models presents a MAIDEN score even higher but the rest of them have a MAIDEN score improved by the refinement. So the model 1 present the lower MAIDEN score (table6) so it is the best model to use for the rest of the analyse (figure15).

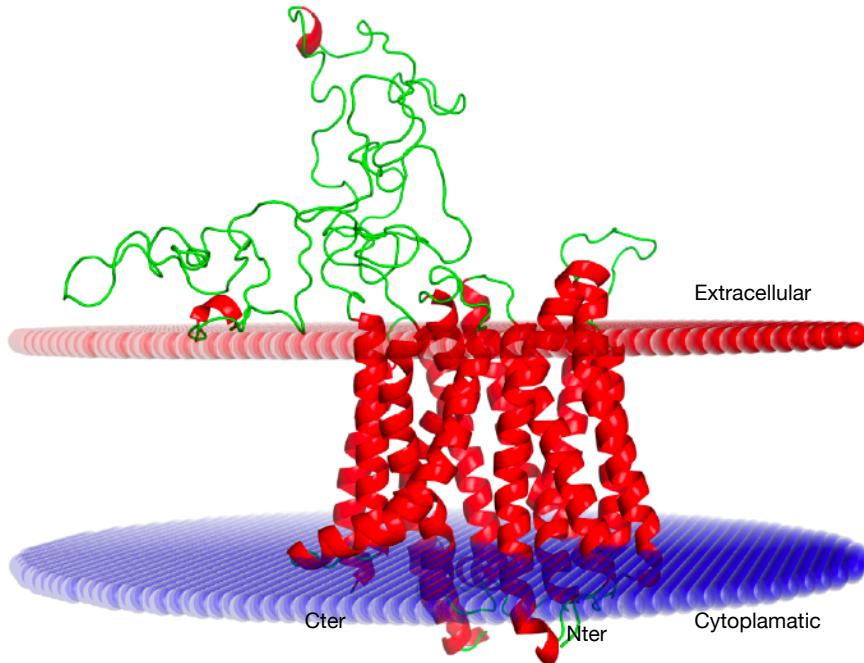


Figure15 : Representation of the model 1 generated by Galaxy Web Serveur (from the model 1 of Swiss Model) after utilisation of OREMPRED

->Comparison with Structure 2D predicted at the beginning and data in the literature
ORMPRO give information on the segment position.

The table7 show that the positions of the 12 transmembranes helix are almost the exact same that the one predicted by MEMSAT-SVM or even OCTOPUS.

Furthermore, some amino acid are known to have a specific place in the secondary structure. Indeed the His57 was found able to bind protons and more specifically His57 seems to be the principal proton-binding site in PepT1[12].
The figure16 show where is located the His57. His57 is, in the model final, in one helix and facing the port. This result is in agreement with the litterature.

There are some amino acid which are critical to the activity of PepT1 find in different helix [12](table8)

Number TM segments	Positions in the final model	MEMSAT-SVM Transmembrane Helix Prediction Figure2
1	16-38	16-38
2	50-74	56-74
3	80-101	83-102
4	123-144	124-142
5	162-184	164-186
6	199-220	198-219
7	283-305	277-297
8	323-351	326-348
9	357-378	359-380
10	581-603	587-603
11	618-640	619-637
12	645-667	648-668

Table7 : Comparaison between the transmembrane helix observed in the final model and predicted by MEMSAT-SVM

Amino Acid	Localisation in the literature	Localisation in the model final
Tyr56	TM2	TM2
Tyr64	TM2	TM2
Tyr167	TM5	TM5
Trp294	TM7	TM7
Glu595	TM10	TM10

Table8 : Verification of the location of critical amino acid for the activity of PepT1

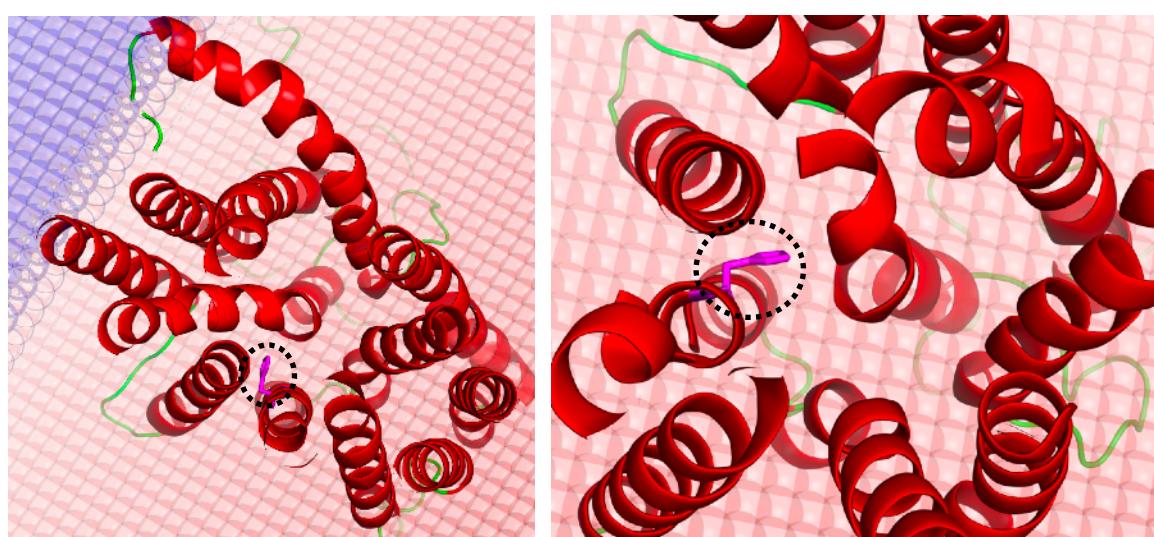


Figure16 : Localisation of His57 (pink) in the 3D representation of the final model (black cercle)

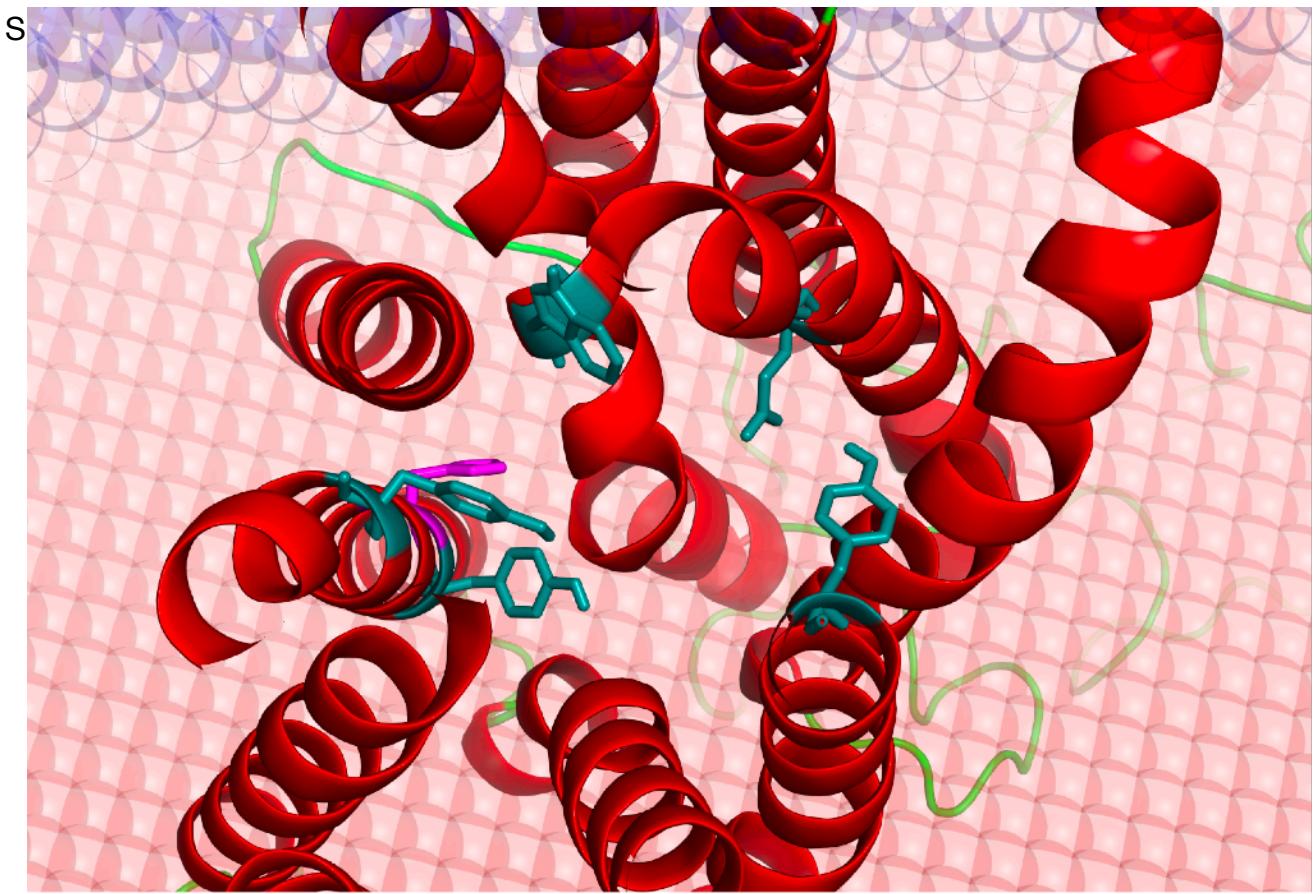


Figure17 : Localisation of His57 (pink) and the amino acid see in table X in the 3D representation of the final model

This amino acid, which are critical to the activity of PepT1 are find at the same position and in the same helix at expected (table8). This residue are all facing the port of the symporter (figure17), so maybe they can release or block the proton and solute. That is in accord which the fact that there are important to the transport activity of PepT1.

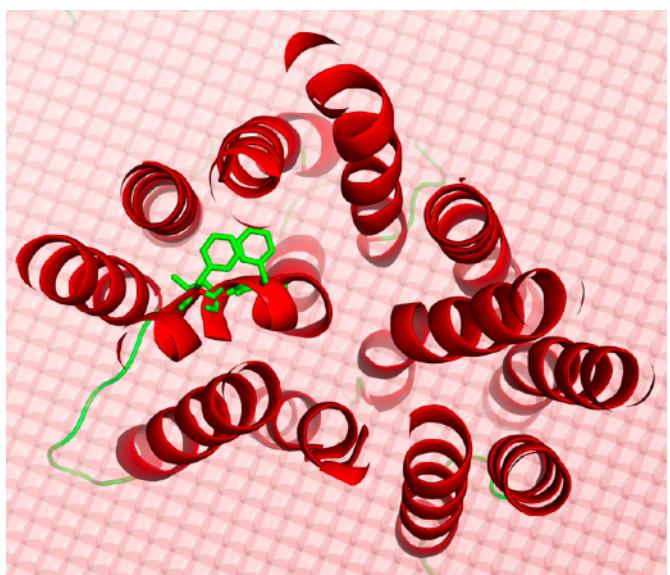


Figure18 : Localisation of the 3 amino acid hydrophobe

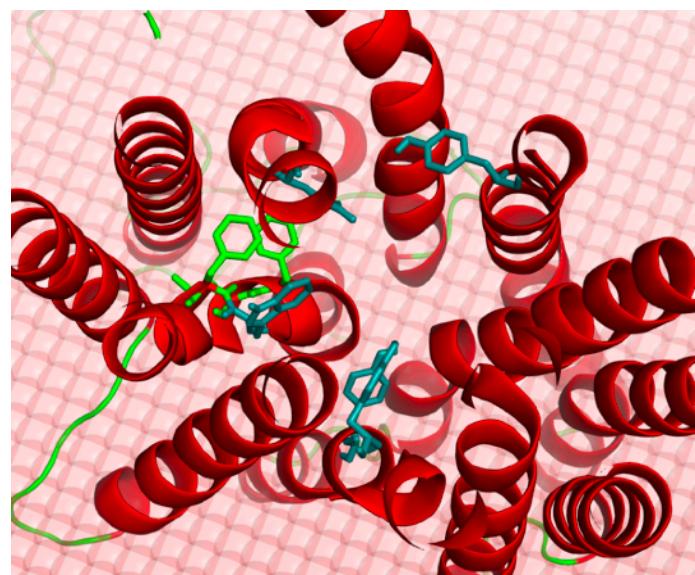


Figure19 : Localisation of the 3 amino acid hydrophobe and the amino acid show previously to have a critical role in the activity of PepT1

Other amino acid are known to be hydrophobic like Phe293, Leu296, Phe297 in the TM7 [3]. This amino acid are located not facing the port put facing other helix(figure18).

The comparision of location between this 3 amino acid and the amino acid show previously (table8) is observed in the figure19. This figure show well the location of each type of amino acid, the amino acid hydrophobic inside a helix groupe and the critical amino acid inside the port allowing the proton and the solute to cross.

E. The normal mode.

Utilisation of NOMAD-ref on the final model was done. AS output 16 normal mode are calculated, but the 6 first mode are rigid body motion. 25 modes were needed to see some movement of the transmembrane helix because the big extracellular loop shown to be very mobile with can explain the formation after the refinement of helix alpha, this helix can be only created during the movement of the big loop. For the transmembrane helix 2 helix are very mobiles the Helix8 and the helix5. The helix 5 is more mobile than the 8. This concord with the importance of the helix5 and in the fact that is the helix the most conserved between species.

IV. Conclusion :

The final model choose is the one with the MAIDEN score the lower. It is SwissModel which generated it with the template 4uvm. This model concord which some fact find in the literature like the 12 helix, the Cter end Nter at the same compartment (inside the cell), the position of some amino acid, and its position in the membrane is like expected.

V. Discussion:

Finding the 3D structure of protein more particularly the protein PepT1 is a challenge because there are a lot of programs which can be use but each programm give some models. So it is important to classify the difference models obtain with the best tool.

The molecular dynamic is not done yet, this can maybe details the difference movements of the protein, and maybe concord some find for the transmembrane helix with the novel mode. And the study of this movement is needed to try to understand how this protein allows peptide and proton to cross.

Bibliographie :

- [1]Anderle P., Carsten Uhd Nielsen, Julia Pinsonneault, Pernille Lindskov Krog, Birger Brodin, and Wolfgang Sade'e
Genetic Variants of the Human Dipeptide Transporter PEPT1
- [2]Ashutosh A. Kulkarni, Ian S. Haworth, and Vincent H.L. Lee
Transmembrane segment 5 of the dipeptide transporter hPepT1 forms a part of the substrate translocation pathway
- [3]Ashutosh A. Kulkarni, Ian S. Haworth, Tomomi Uchiyama, and Vincent H. L. Lee
Analysis of Transmembrane Segment 7 of the Dipeptide Transporter hPepT1 by Cysteine-scanning Mutagenesis*
- [4]Beale John H., Joanne L. Parker, Firdaus Samsudin, ..., David Meredith, Philip W. Fowler, Simon Newstead
Crystal Structures of the Extracellular Domain from PepT1 and PepT2 Provide Novel Insights into Mammalian Peptide Transport
- [5]Doering F., C. Martini, J. Walter, H. Daniel
Importance of a Small N-Terminal Region in Mammalian Peptide Transporters for Substrate Activity and Function
- [6]Ebejer JP., Jamie R. Hill, Sebastian Kelm, Jiye Shi and Charlotte M. Deane
Memoir: template-based structure prediction for membrane proteins
- [7]Kelm S., Jiye Shi and Charlotte M. Deane
MEDELLER: homology-based coordinate generation for membrane proteins
- [8]Lim Heo, Hahnbeom Park and Chaok Seok
GalaxyRefine: protein structure refinement driven by side-chain repacking
- [9]Nieng Yan
Structural advances for the major facilitator superfamily(MFS) transporters
- [10]Nugent T., Sean Ward and David T. Jones
The MEMPACK alpha-helical transmembrane protein structure prediction server
- [11]Rong Liang and all.
Human intestinal H⁺/Peptide Cotransporter
- [12]Uchiyama T., Ashutosh A. Kulkarni, Daryl L. Davies and Vincent H. L. Lee
Biophysical Evidence for His⁵⁷ as a Proton-Binding Site in the Mammalian Intestinal Transporter hPepT1