

烟草K⁺通道蛋白NK14 的生物信息学分析

报告人：向小华

小组：G07

成员：向小华 陈俊宇 刘健伟 王若焱

选题意义

研究背景

进化分析

蛋白结构预测

同源建模

致谢



选题意义

1. 钾是植物生长发育所必需的矿质营养元素之一,是作产量与品质形成的限制因素。植物中钾的吸收与转运依赖膜蛋白,主要包括钾离子通道和高亲和性的钾转运体。
2. 烟草是一种耗钾量很大的作物,烟叶含钾量又是衡量烟叶品质的重要指标之一。 K^+ 通道是允许 K^+ 特异通透质膜的离子通道,是烟草 K^+ 吸收的重要途径之一。
3. 为获得烟草 K^+ 通道基因,为通过基因工程手段利用 K^+ 通道基因,进而提高烟叶品质、增强烟草抗逆性及开展烟草钾高效育种提供理论与技术支撑。



研究背景

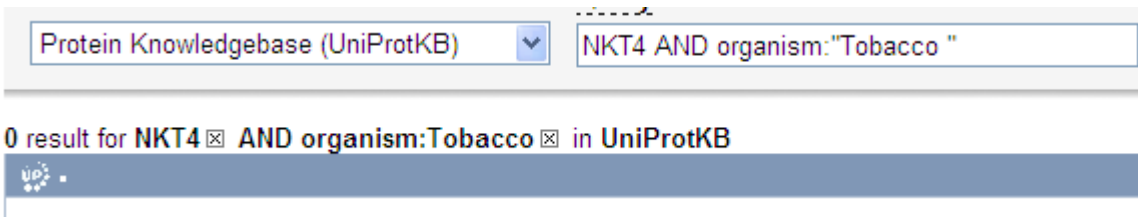
1.2007年Sano等利用同源克隆法首次克隆出烟草K⁺通道基因*NKT1*、*NTORK1*、*NKT2*和*NtKC1*，其中*NKT1*的主要功能是吸收K⁺。近几年，随着分子生物学技术的日益成熟和K⁺通道研究的不断深入，K⁺通道基因的研究以及烟草K⁺吸收机制的研究成为热点。但是对于烟草K⁺通道蛋白的分析研究较少。

2.本文拟通过相关文献的搜索对烟草K⁺通道基因*NKT4*的比对，从而进行烟草K⁺通道蛋白*NKT4*预测分析。

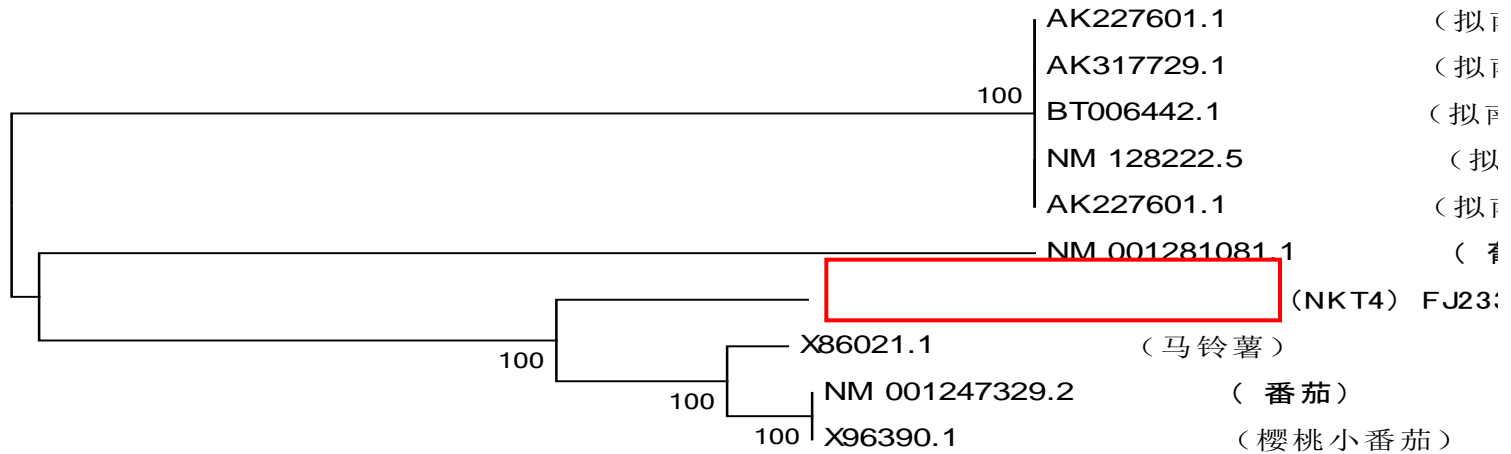


NKT4的获得与基本信息

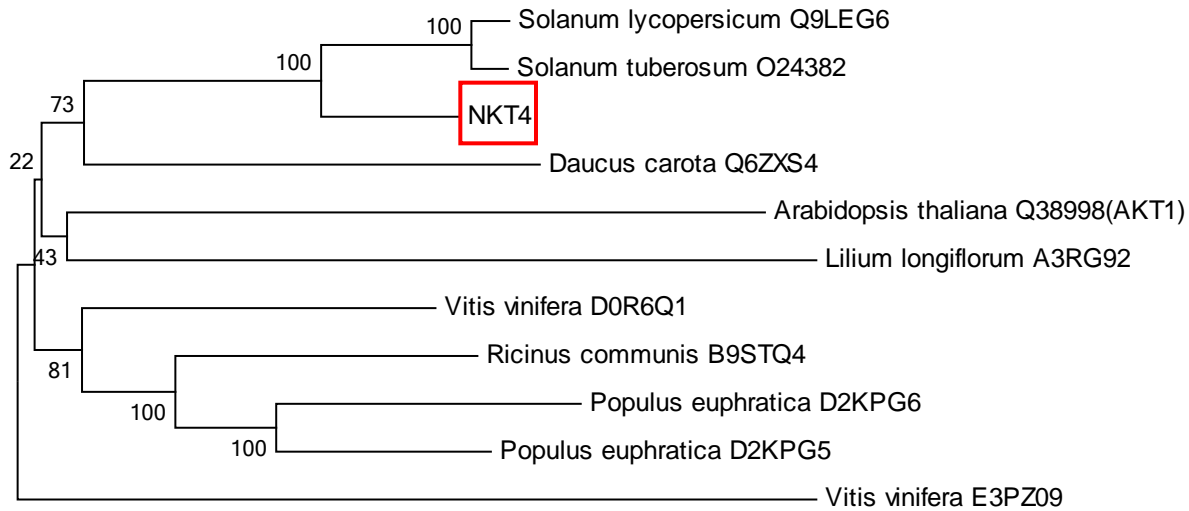
- 通过比对拟南芥、胡萝卜、番茄和马铃薯的K⁺ 通道氨基酸序列得到了保守序列, 设计简并引物, 利用RT-PCR获得大小为490bp的普通烟草K⁺ 通道基因中间片段。设计特异性引物, 应用RACE方法得到5'末端和3'末端cDNA序列。通过拼接并结合全长克隆及测序验证, 获得一个未报道的普通烟草K⁺ 通道基因, 并将其命名为NKT4 (GenBank登录号为FJ233071)。
- 在Uniport 中无相关信息。



系统进化分析-MEGA5.0



0.05



0.05



K⁺通道蛋白的特征序列

NKT4	236	IGASMG-DFLHQSLWIRYITSIIYWSIT	TLTTVGYGD	LHPENTREMIFDIFYMLFNLGLTA
Solanum lycopersicum Q9LEG6	233	IGVAMD-DFLNQSLWIRYVTSIIYWSIT	TLTTVGYGD	LHPENTREMIFDIFYMLFNLGLTA
Solanum tuberosum O24382	233	IGASMD-DFLNQSLWIRYITSIIYWSIT	TLTTVGYGD	LHPENTREMIFDIFYMLFNLGLTA
Vitis vinifera D0R6Q1	232	IGASM-NNFLEQSLWIRYVTAIIYWSIT	TLTTVGYGD	LHPENTREMIFDIFYMVFNGLTA
Daucus carota Q6ZXS4	227	IGASIT-DFKNQSLWIRYVTSIIYWSIT	TLTTVGYGD	LHAQNTGEMIVDIFYMLFNLGLTA
Ricinus communis B9STQ4	232	IGQSLGDNFLEQSLWIRYVTSIIYWSIT	TLTTVGYGD	LHPVNTREMIFDIFYMLFNLGLTA
Arabidopsis thaliana Q38998(AKT1)	224	IGANVA-NFLEESLWMRYVTSIIYWSIT	TLTTVGYGD	LHPVNTKEMIFDIFYMLFNLGLTA
Populus euphratica D2KPG6	233	IGAALGDNFLEQGIWIRYVTSIIYWSIT	TLTTVGYGD	LHPVNTSEMVF DIFYMLFNLGLTA
Populus euphratica D2KPG5	232	IGASLGDNFLEQSLWRRYVTSIIYWSIT	TLTTVGYGD	LHPVNTREMLFDIFYMLFNLGLTA
Lilium longiflorum A3RG92	220	IGASMP-DFHEQSLWVRYVTSIIYWSIT	TLTTVGYGD	LHAQNTGEMIFDIAYMLFDLGLTA
Vitis vinifera E3PZ09	238	LGLVMG-DLNTHSLWQRYVTSIIYWSIT	TLTTTGYGD	LHAVNTREMVF DIFYMVFNGLTSS

TxxTxGYGD



蛋白质理化性质分析(PROTPARAM)

亲疏水性分析(PROTSYSCALE)

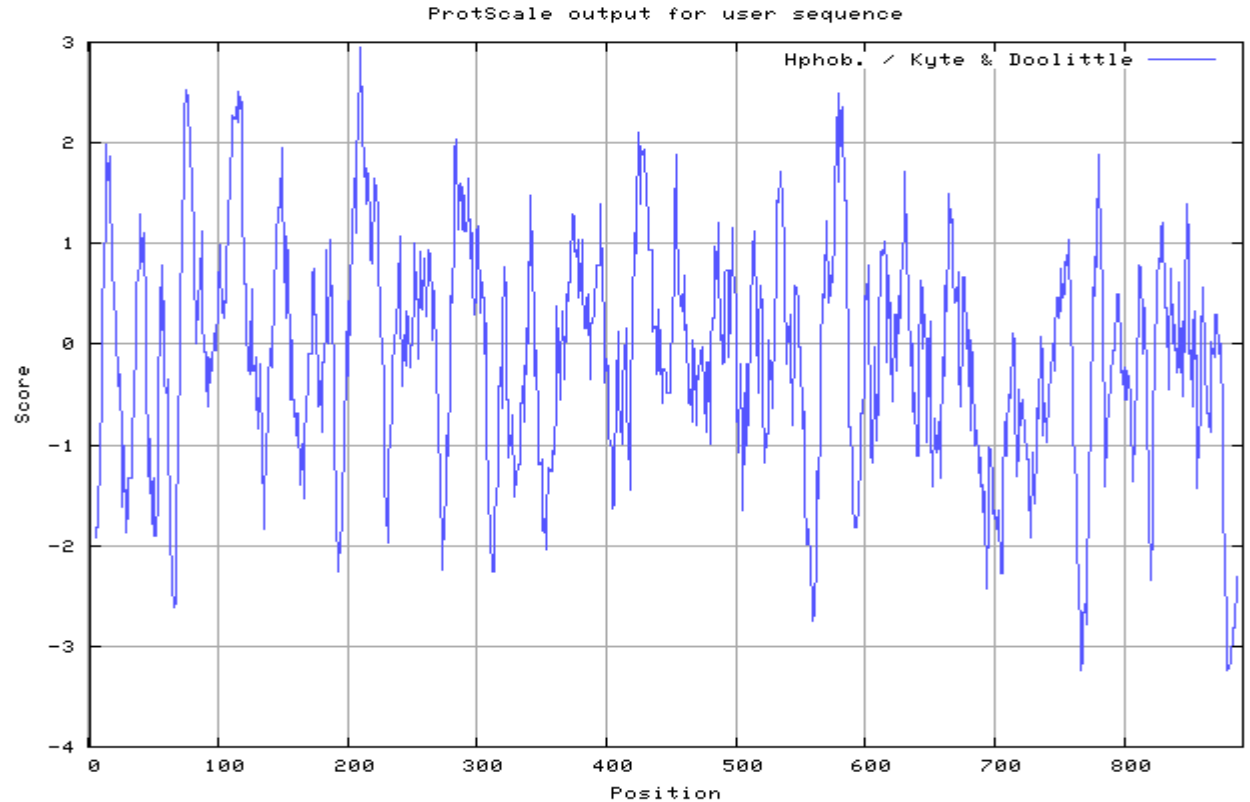
Number of amino acids: 892

Molecular weight: 100340.9

Theoretical pI: 6.64

Amino acid composition:

Ala (A)	56	6.3%
Arg (R)	60	6.7%
Asn (N)	49	5.5%
Asp (D)	52	5.8%
Cys (C)	15	1.7%
Gln (Q)	30	3.4%
Glu (E)	48	5.4%
Gly (G)	63	7.1%
His (H)	22	2.5%
Ile (I)	46	5.2%
Leu (L)	107	12.0%
Lys (K)	36	4.0%
Met (M)	17	1.9%
Phe (F)	46	5.2%
Pro (P)	33	3.7%
Ser (S)	56	6.3%
Thr (T)	50	5.6%
Trp (W)	11	1.2%
Tyr (Y)	26	2.9%
Val (V)	69	7.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%



MIN: -3.233, MAX: 2.956

Scale: Hydropathicity.

Window size: 9

Scale not normalized.

Relative weight for window edges: 100 %

Weight variation model: linear

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%



蛋白质二级结构预测(PSIPRED)

Secondary Structure Map

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

```

1 M G D V R R R N H F G V L G V S M C G A A Q E I E Q L S R D S S H Y S L S T G I L P S L G A R S N R 50
51 R V K L Q R F I I S P Y D R H Y R S W E T F L V A L V V Y T A W V S P F E F G F L E K P T G P L A V 100
101 T D H V V H G F F A I D I V L T F F V A Y L D R T T Y L L V D N H K K I A W K Y A S T W F L F D V I 150
151 S T I P S E L A R K I S P K P L R Q Y G L F H M L R L W R L R R V S A L F A R L E K D R N F H Y F W 200
201 V R C A K L V C V T L F A V H C A G C F Y Y L I A A N Y P N P T K T W I G A S M G D F L H Q S L W I 250
251 R Y I T S I Y W S I T T L T I V G Y G D L H P E N T R E M I F D I F Y M L F N L G L T A Y L I G N M 300
301 T H L V V H G T S R T R K F R D T I Q A A S S F A Q R N Q L P A R L Q D Q M L A H L C L K F R T D S 350
351 E G L Q Q Q E T L E S L P K A I R S S I S H F L F Y S L V D K V Y L F R G V S N D L L F Q L V S E M 400
401 K A E Y F P P K E D V I L Q N E A P T D F Y I L V T G A V D L L V L K N G V E Q V V G E A K T G D L 450
451 C G E I G V L C Y R P Q L F T V R T K R L C Q L L R M N R T T F L N I V Q A N V G D G T I I M N N L 500
501 L Q H L K D I K D P I M E G V L L E T E R M L A R G R M D L P L T L C F A T L R G D D L L L H Q L L 550
551 K R G L D P N E S D H H G R S A L H V A A S T G I E S C V V L L L D F G A D V N S R D S E G N V P L 600
601 W E A I S G K H E P V I K L L V D N G A K L S A G D V G H F A C V A A E Q N N L N L L K D I V R Y G 650
651 G D V T R P K V N G S S A L H V A V C E G N M E I V K Y L L D R G A N V D Q V D E H G W T P R D L A 700
701 E Q Q G H E D I K E L F E S G E V M R T R S V D P I P E E R H G V R F L G R F K S E P T I F P A S H 750
751 G A S F L A S G G S L G R S R P R R R T N H F H N S L F G I M S A V Q T N E H D V V L S A N E A N V 800
801 S A T T T K T Y A P R V T V C C P E K G D N G G K L V L L P P S F Q E L L Q I G F N R Y I G I L Q V K 850
851 V I S K D G A E I D D I E L I R D G D R L I F V S D K E N N E T D N H Q N G D E L R
    
```

Key
Annotations

Helix
M

Sheet
L

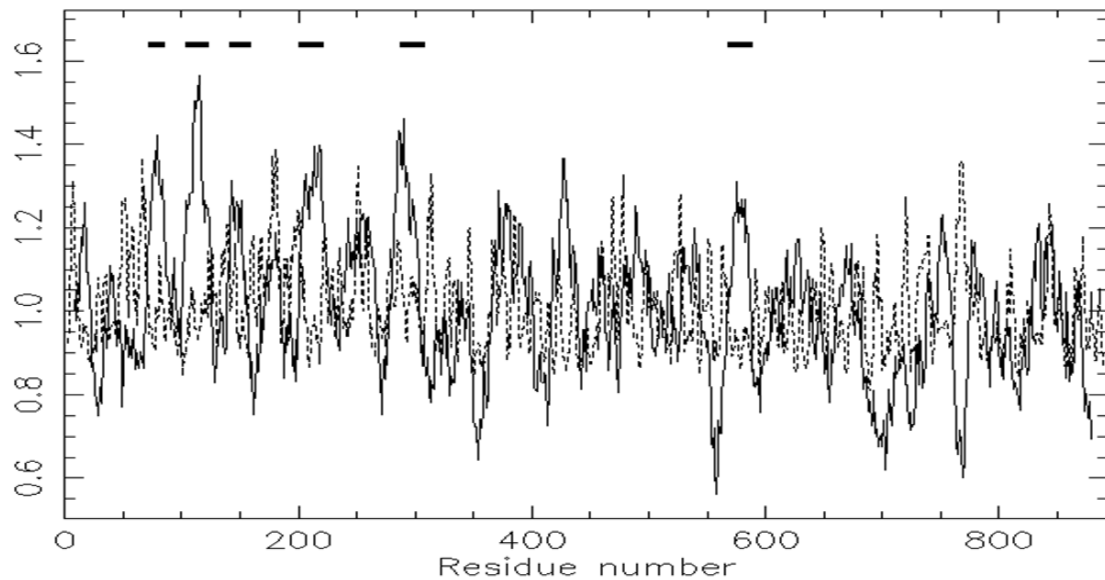
Disordered
E

Dompred Boundary
A

DomSSEA Boundary
D

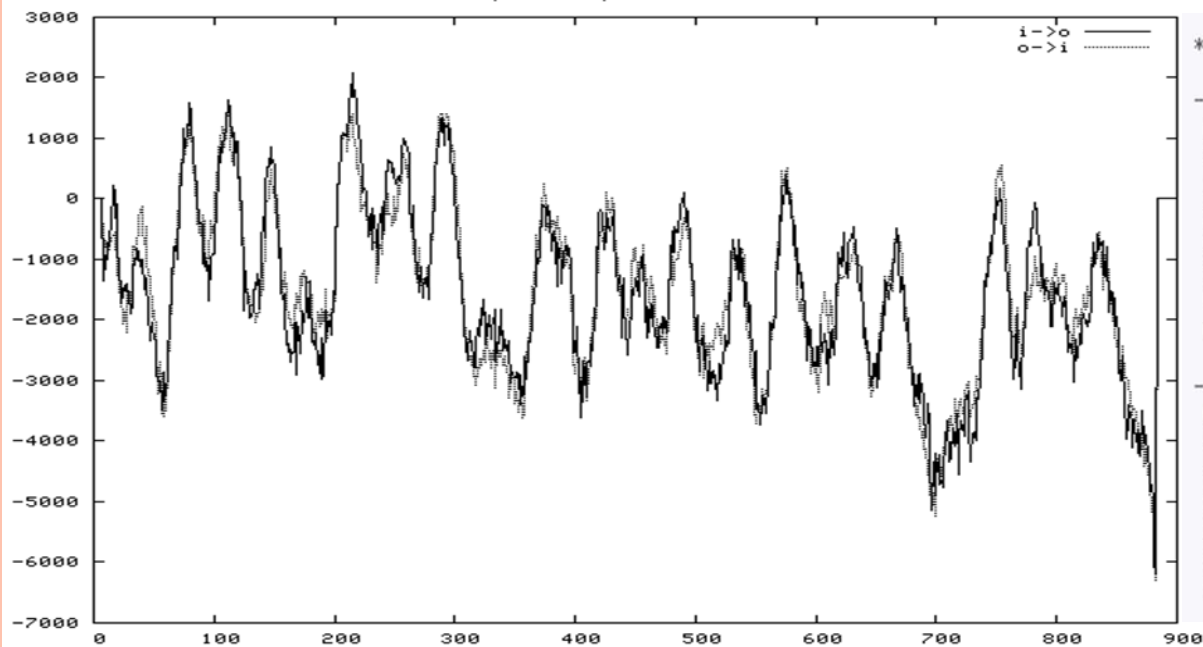


Tmap



蛋白质跨膜区分析 (TMAP&TMPRED)

TMpred output for unknown



*** the models differ in the number of TM-helices ! ***

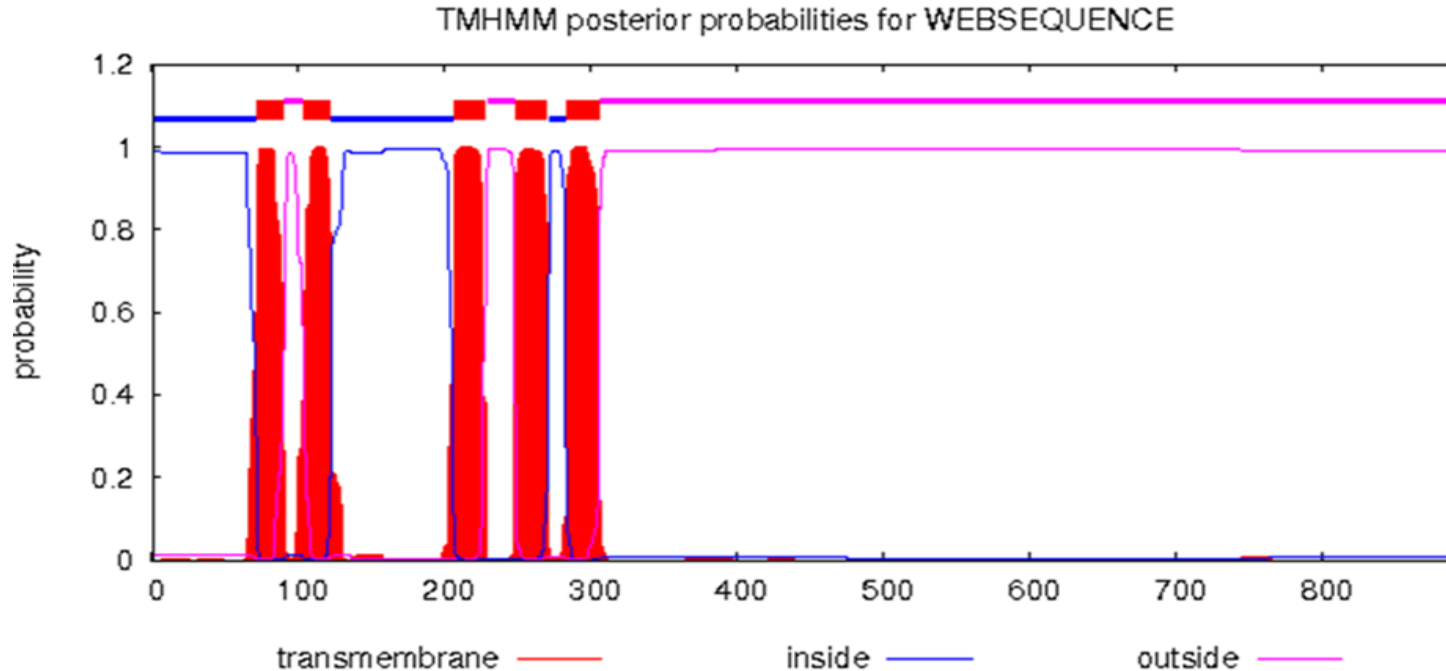
----> STRONGLY preferred model: N-terminus inside
6 strong transmembrane helices, total score : 7862

#	from	to	length	score	orientation
1	71	91	(21)	1596	i-o
2	104	122	(19)	1539	o-i
3	140	158	(19)	862	i-o
4	206	225	(20)	1428	o-i
5	249	269	(21)	1003	i-o
6	283	305	(23)	1434	o-i

----> alternative model
5 strong transmembrane helices, total score : 6645

#	from	to	length	score	orientation
1	71	91	(21)	1142	o-i
2	104	122	(19)	1638	i-o
3	206	225	(20)	1428	o-i
4	249	269	(21)	1003	i-o
5	283	305	(23)	1434	o-i

蛋白质跨膜区分析(TMHMM&SOSUI)



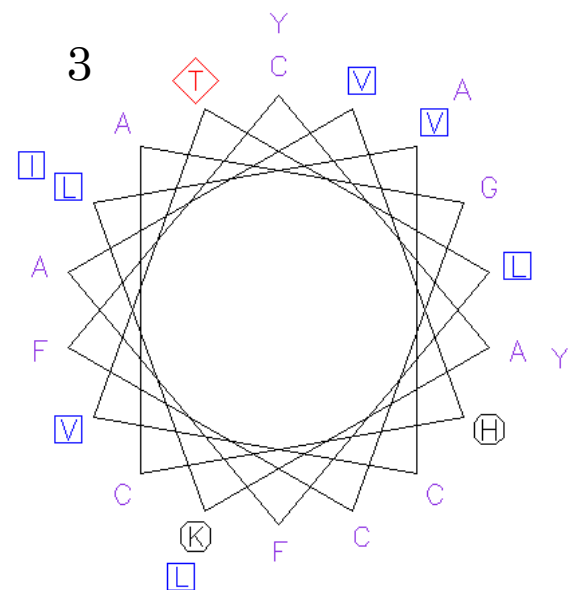
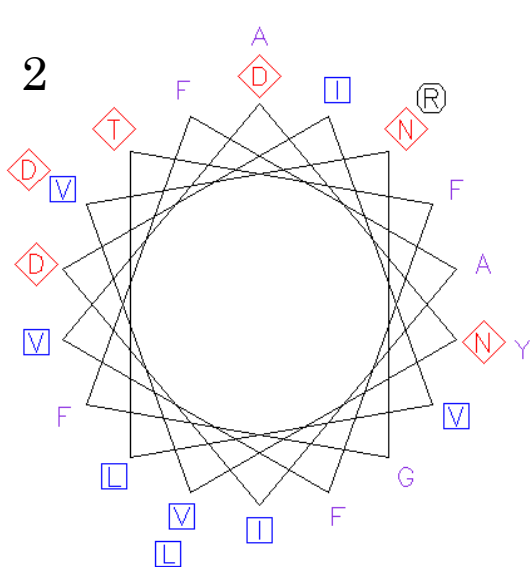
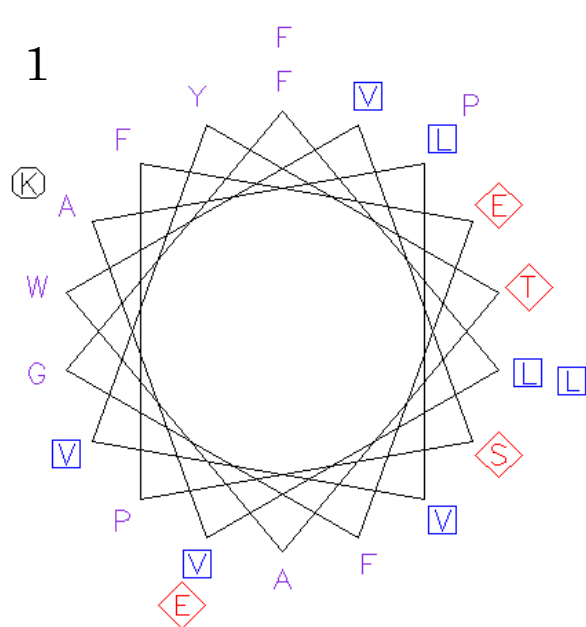
This amino acid sequence is of a MEMBRANE PROTEIN which have 3 transmembrane helices.

No.	N terminal	transmembrane region	C terminal	type	length
1	72	FLVALVVYTAWVSPFEGFLEKP	94	SECONDARY	23
2	102	DNVVNGFFAIDIVLTFFVAYLDR	124	PRIMARY	23
3	203	CAKLVCVTLFAVHCAGCFYYLIA	225	SECONDARY	23



蛋白质跨膜区分析并绘制螺旋图(PEPWHEEL)

程序	预测的螺旋						
	1	2	3	4	5	6	7
Tmap	68-89	99-127	137-163	197-225		283-311	564-592
Tmpred	71-91	104-122	140-158	206-225	249-269	283-305	
TMHMM	72-89	104-122		206-228	248-270	283-305	
SOSUI	72-94	102-124		203-225			

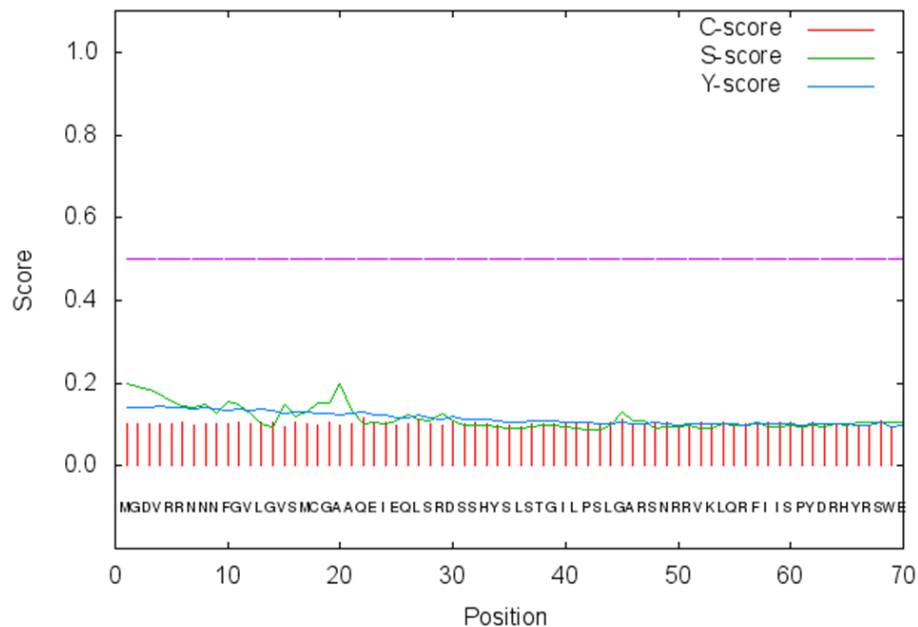


亚细胞定位分析 (TARGETP 1.1SERVER) 信号肽预测 (SIGNALP 4.1SERVER)

```
### targetp v1.1 prediction results #####
Number of query sequences: 1
Cleavage site predictions included.
Using PLANT networks.
```

Name	Len	cTP	mTP	SP	other	Loc	RC	IPlen
Sequence	892	0.072	0.213	0.016	0.570	_	4	-
cutoff		0.000	0.000	0.000	0.000			

SignalP-4.1 prediction (euk networks): Sequence

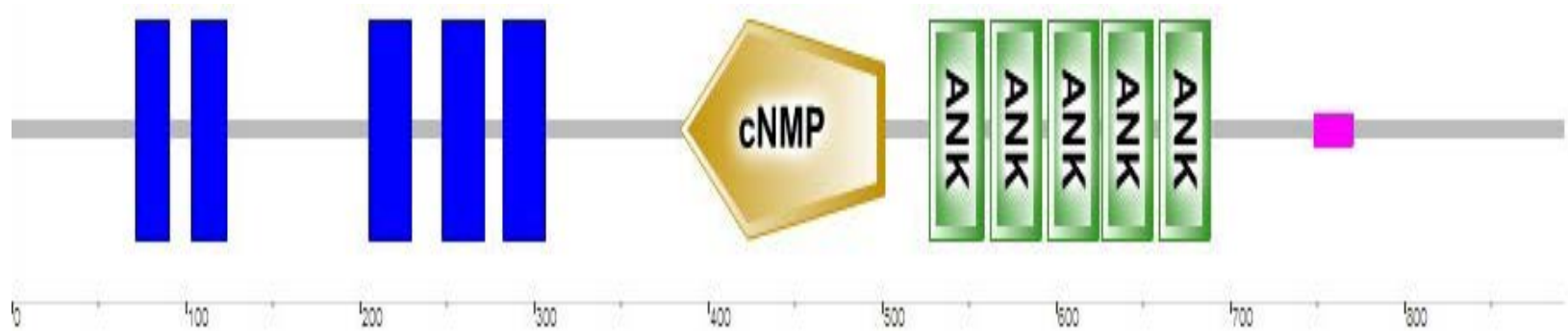


# Measure	Position	Value	Cutoff	signal peptide?
max. C	22	0.116		
max. Y	11	0.138		
max. S	20	0.200		
mean S	1-10	0.162		
D	1-10	0.151	0.450	NO

Name=Sequence SP='NO' D=0.151 D-cutoff=0.450 Networks=SignalP-noTM



结果域预测 (SMART&NCBI)



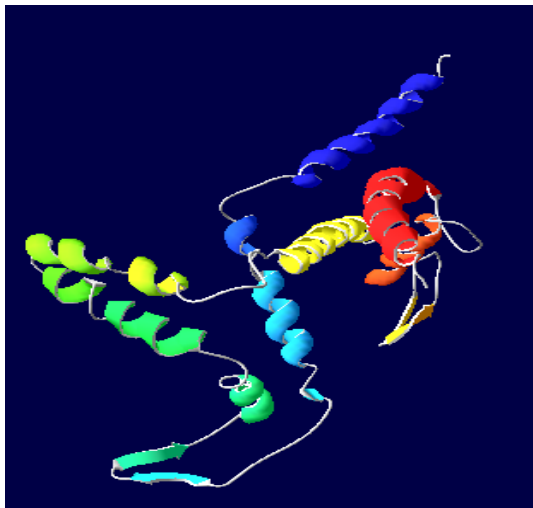
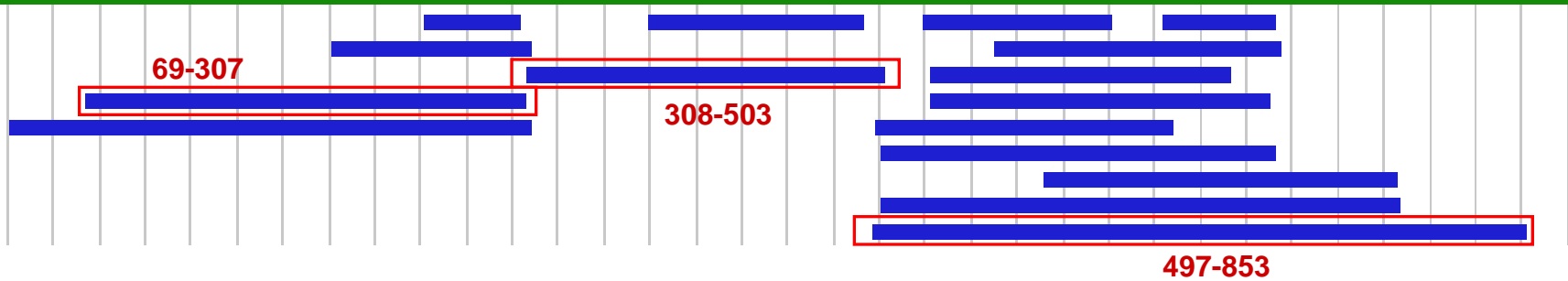
Putative conserved domains have been detected, click on the image below for detailed results.



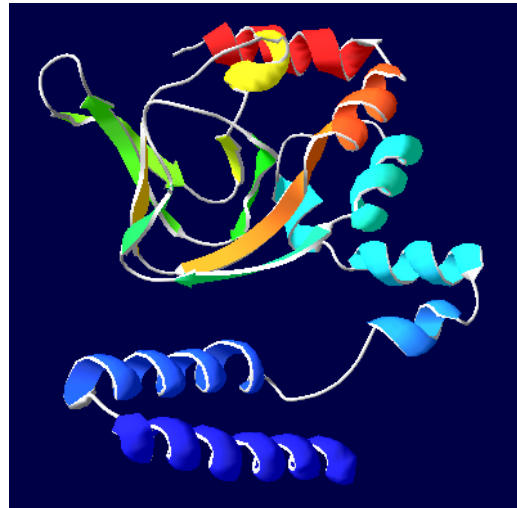
同源建模(SWISS-MODEL)

892

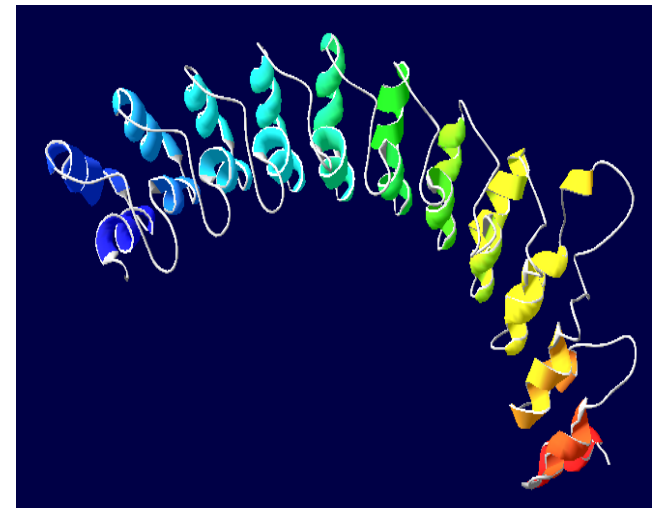
1



69-307

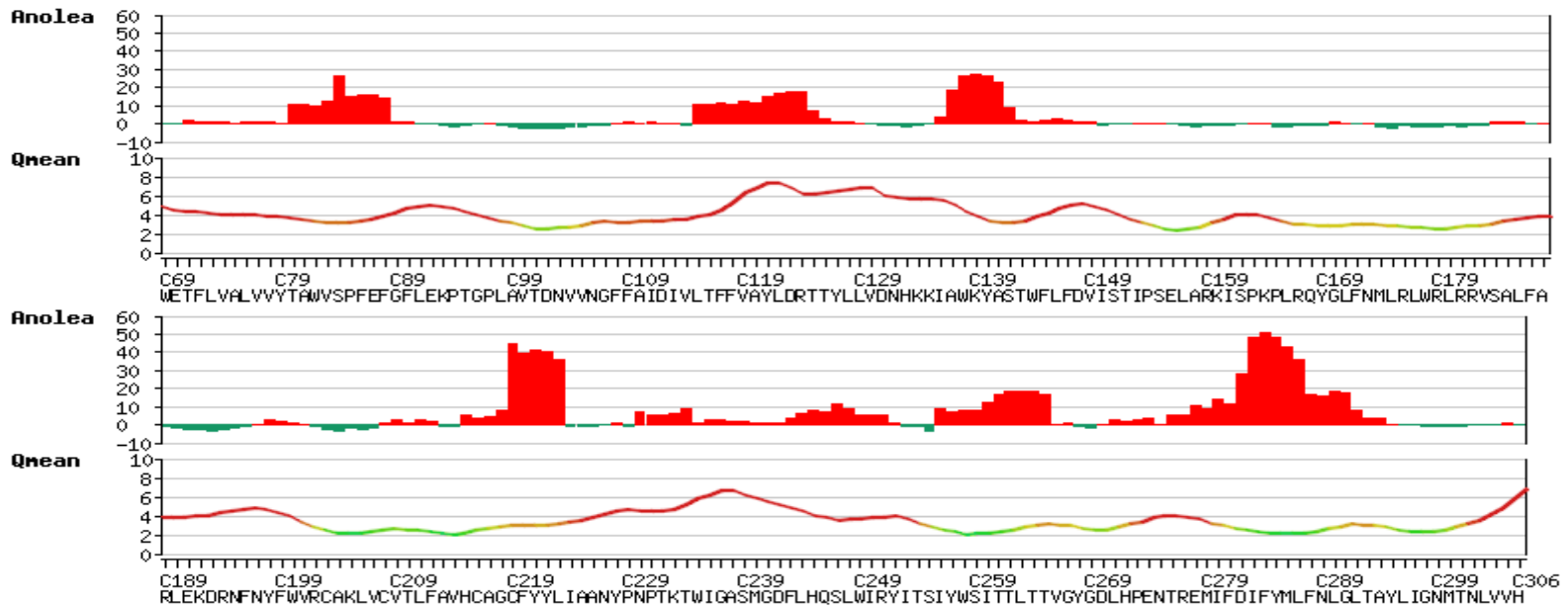


308-503

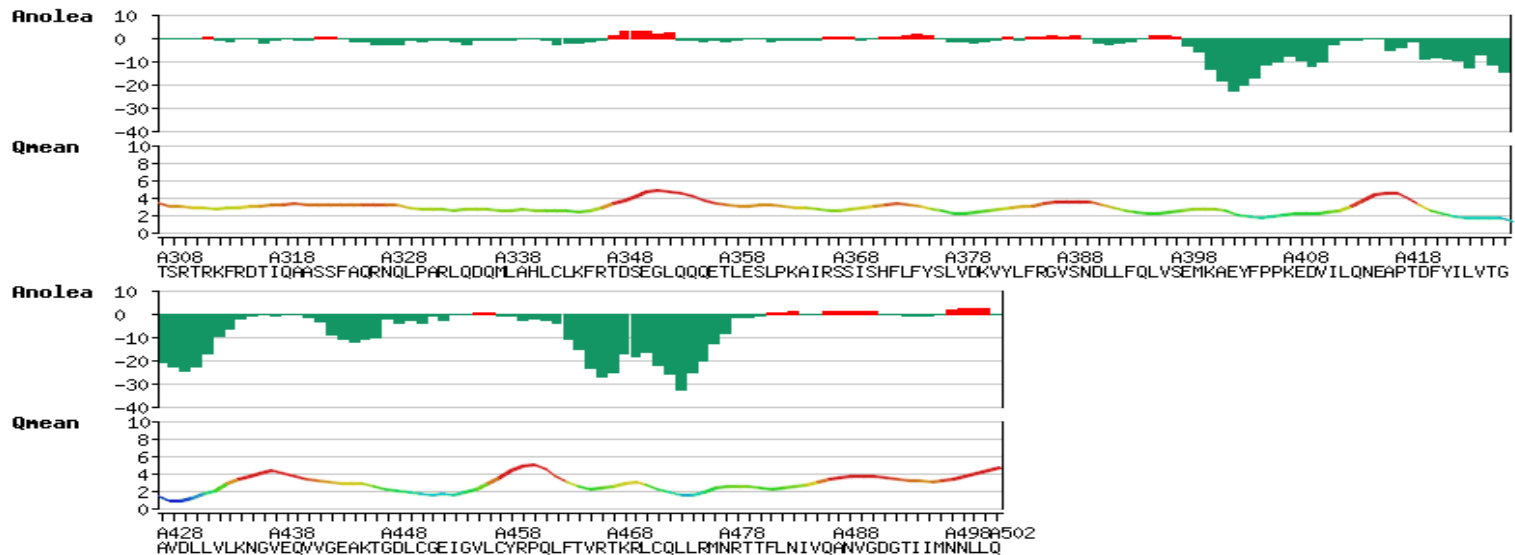


497-853



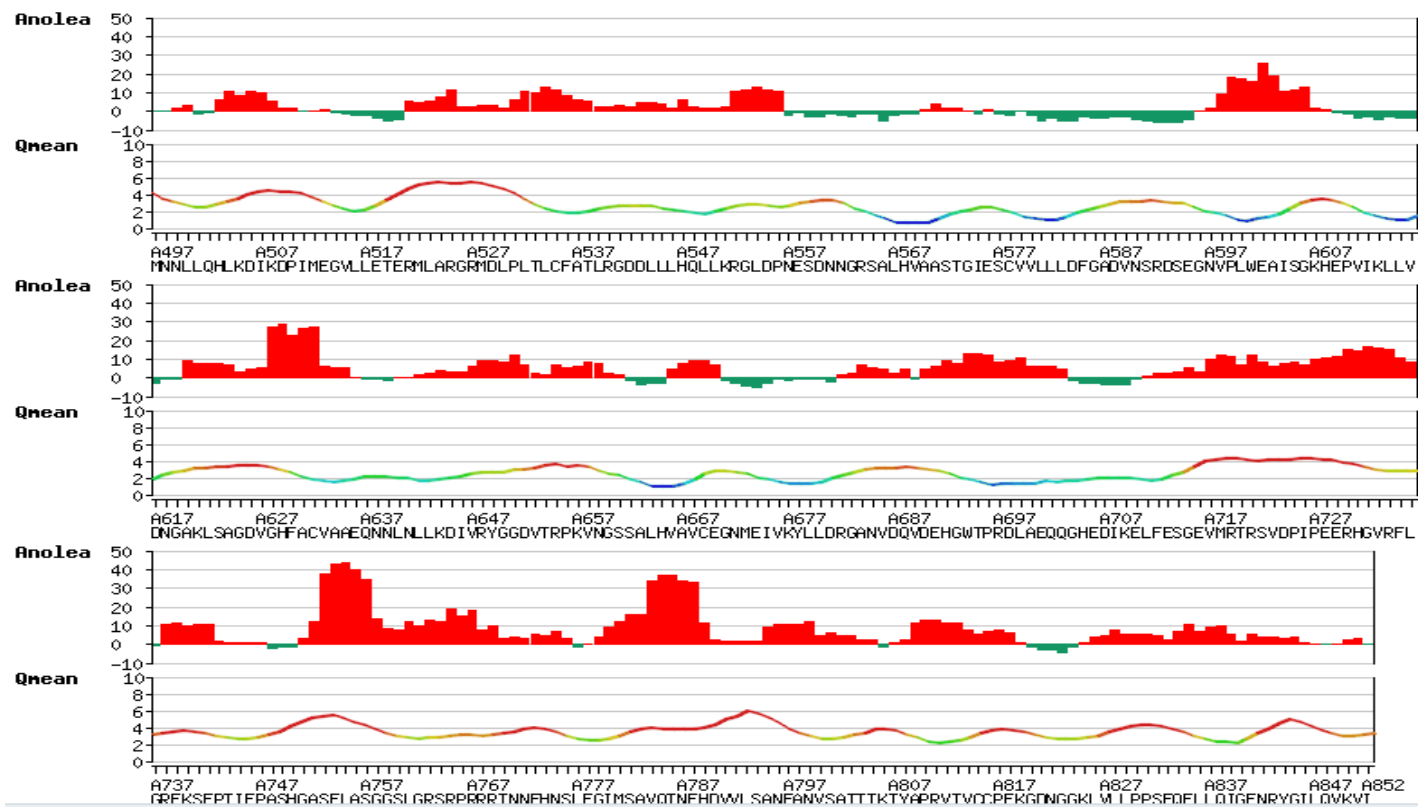


69-307



308-503

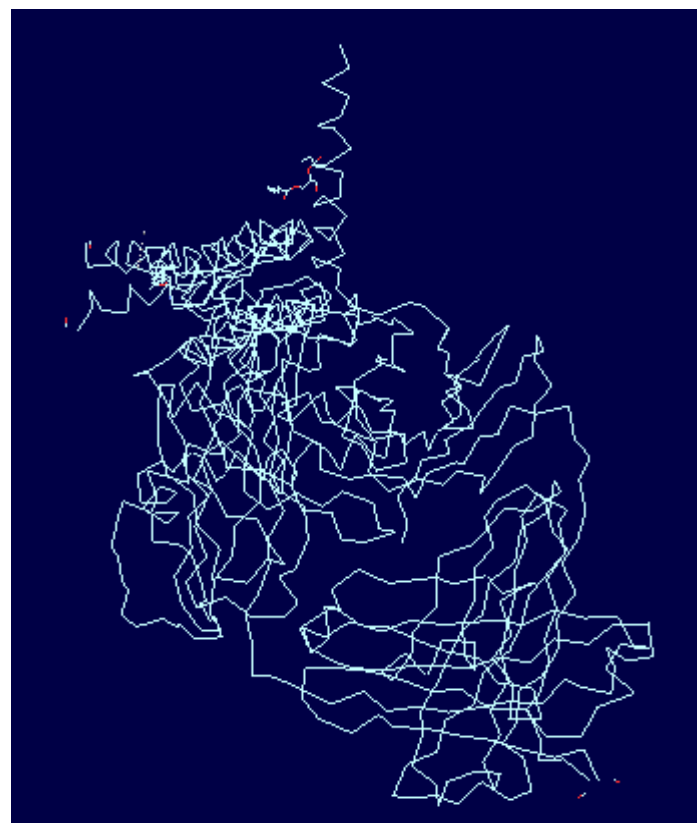
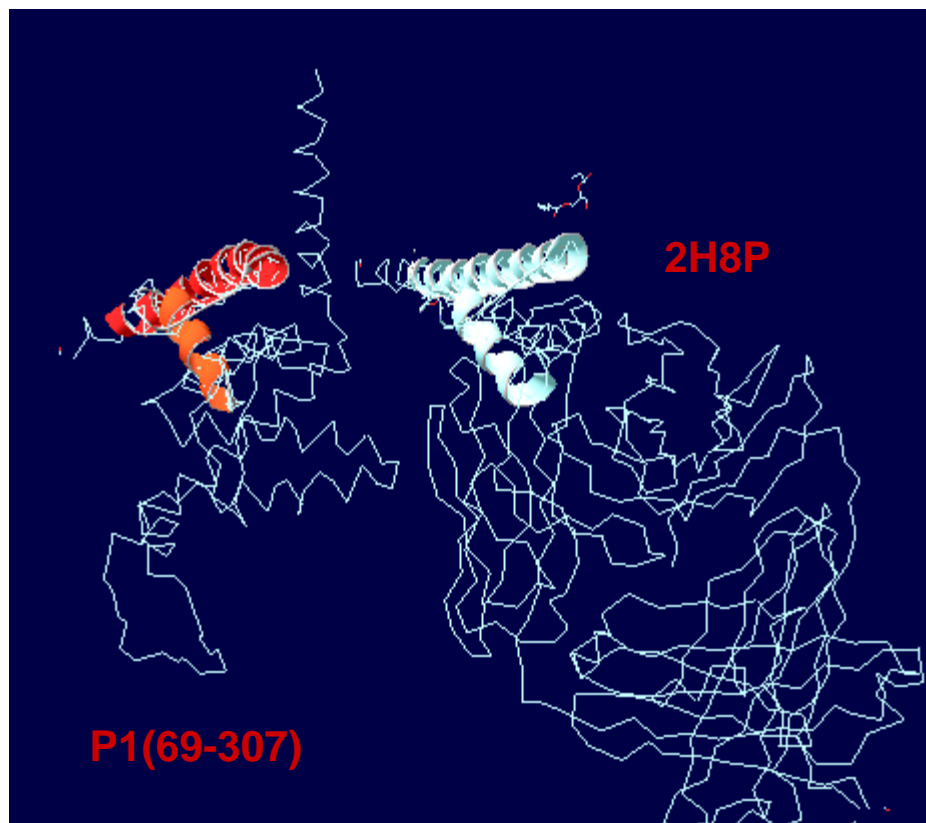




497-853

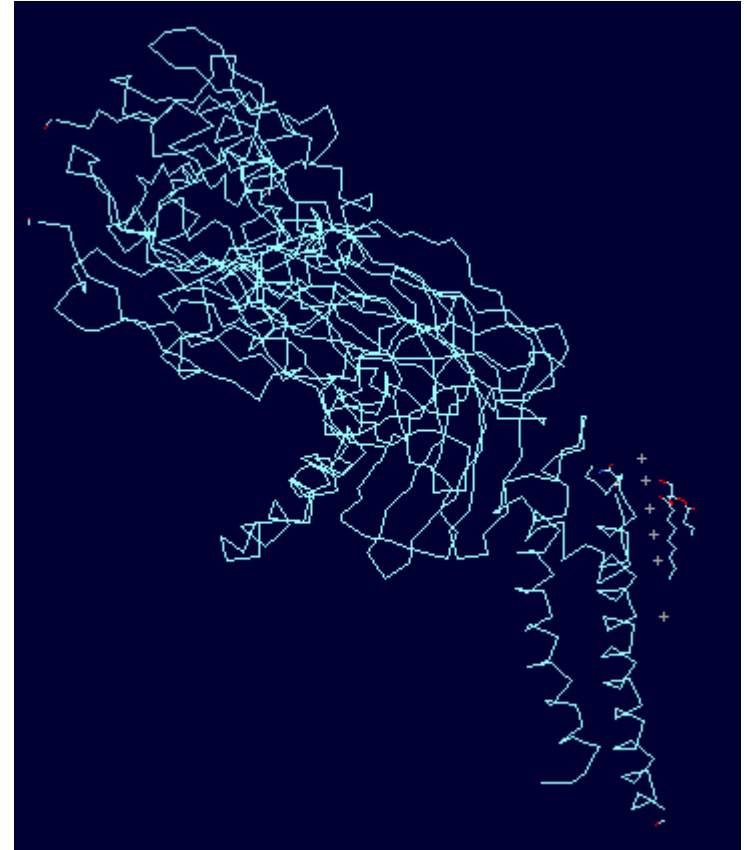
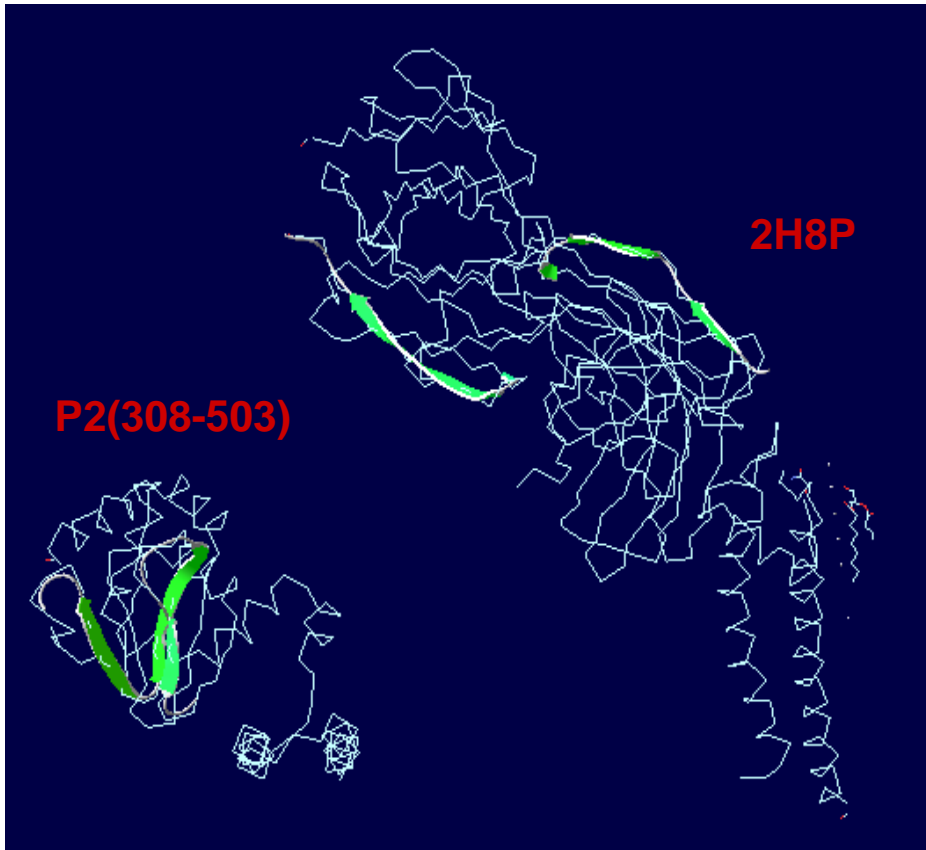


预测模体与同源模体 (2H8P) 的FIT比较



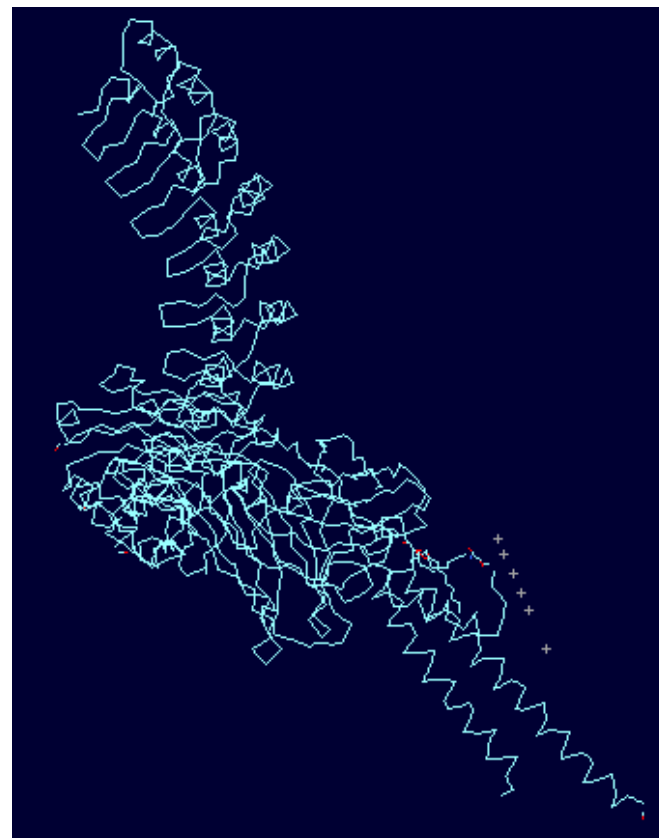
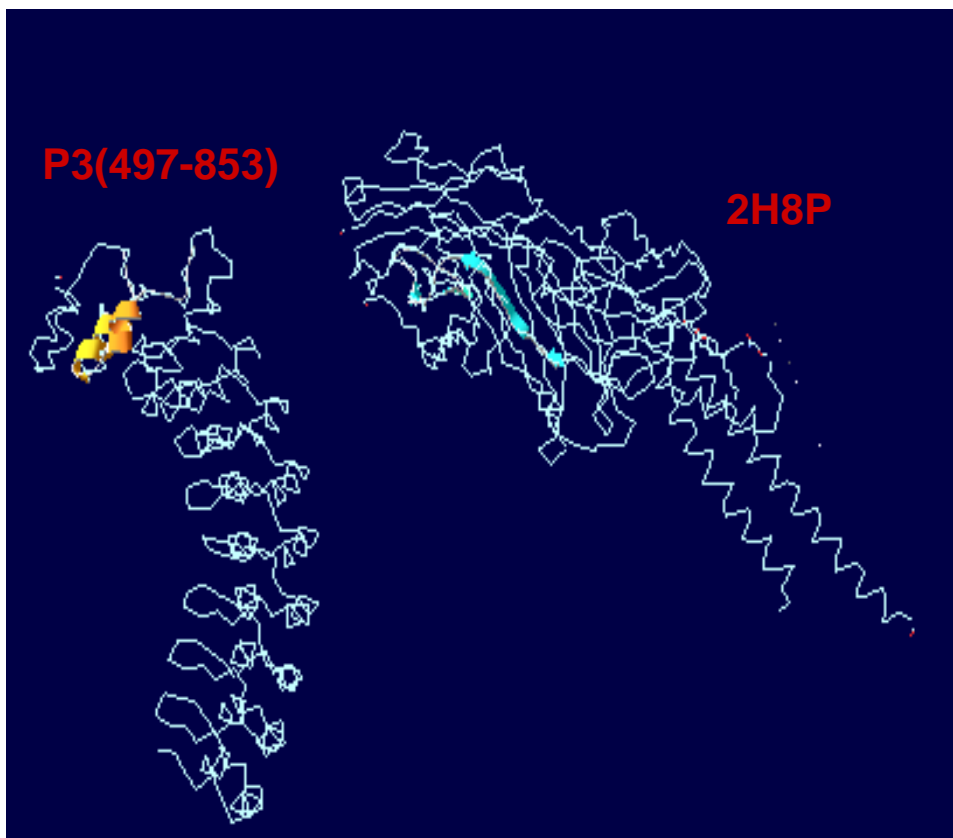
同源模体2H8P中的C链60-78位及D链的79-111位共52个氨基酸与P1(69-307)预测模体中250-301为氨基酸匹配。





同源模体A链205-219及B链中1-16为共31个氨基酸氨基酸与P2(308-503)预测模体中的A链406-437位匹配。





同源模体A链109-143共35个氨基酸与P3(497-853)预测模体中的A链795-829位匹配。



主要参考文献

[1] Karena L Waller, Characterization of two putative potassium channels in Plasmodium Falciparum, Malaria Journal 2008, 7 :19.

[2] Asher *et al* . World Journal of Surgical Oncology 2010, 8:113.

[3] Ashley M K, Grant M, Grabov A ,Plant responses to potassium deficiencies: a role for potassium transport proteins[J] J Exp Bot, 2006, 57: 425-436.

[4] KalimS, LuthraYP, Gandhi, SK. Roleof zincandmanganesein, resistanceof cowpearoot rot [J]. Journal of Plant Diseases and Protection (Germany), 2003, 110(3): 235-243.

[5] 刘晓燕, 何萍等, 钾在植物抗病性中的作用及机理的研究进展, 植物营养与肥料学报2006,12(3) :445- 450.

[6] 严蔚东,王校常等, 利用外源钾通道基因改良水稻钾素营养, 中国水稻科学(ChineseJ RiceSci), 2002, 16(1) :77~79.



致谢

- 由衷的感谢静初先生对我们课程的悉心指导和谆谆教诲，让我们学习到了更加全面的生物信息学知识！从中也体会了更多的乐趣。
- 非常感谢组员及同学们的支持和帮助，使我们的学习更轻松！
- 特别感谢班长任磊同学及杨春红同学的大力帮助。

