

桃蛀螟性信息素结合蛋白的 生物信息学分析

第一组汇报人：葛星

组员：张林雅、王娟、葛星、杨春红
植保所

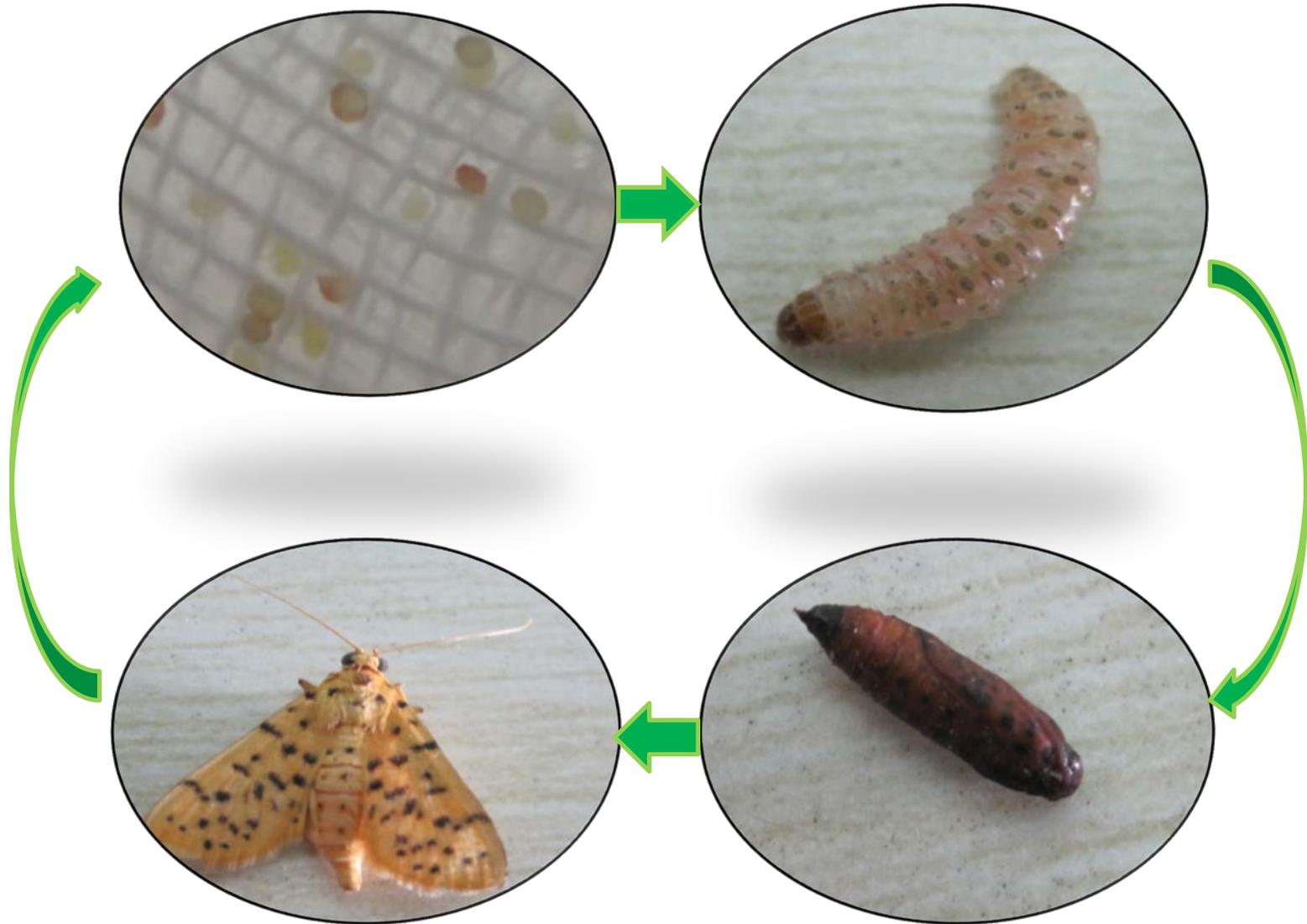
研究背景

序列的获得

序列的生物信息分析

后续试验

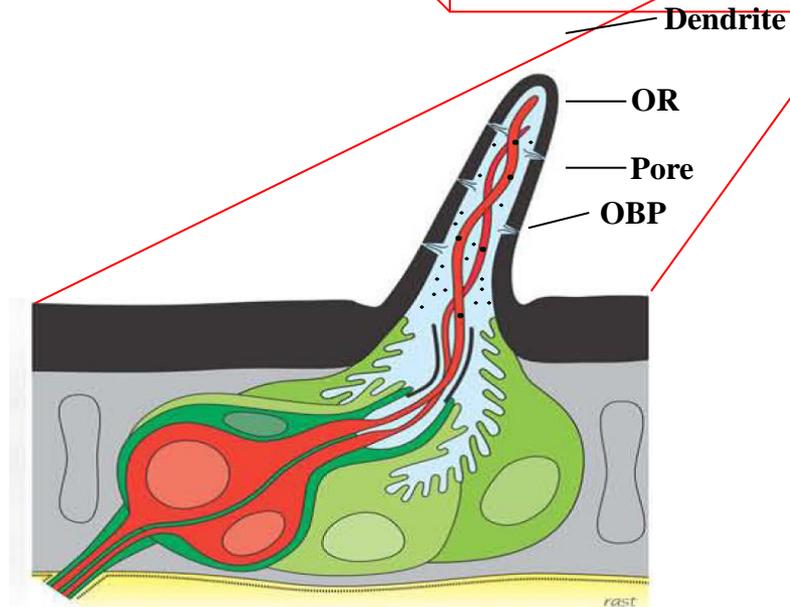
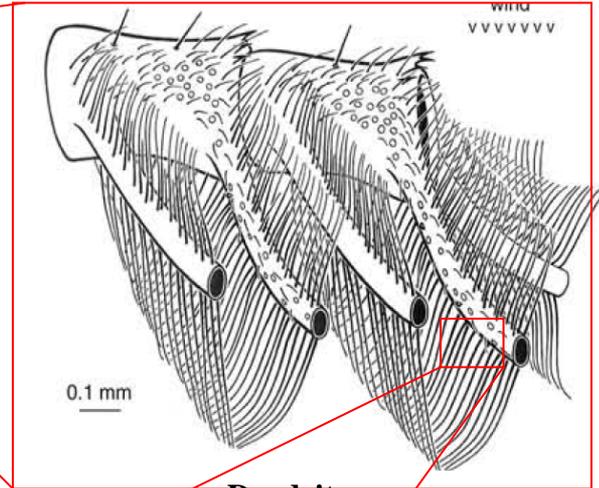
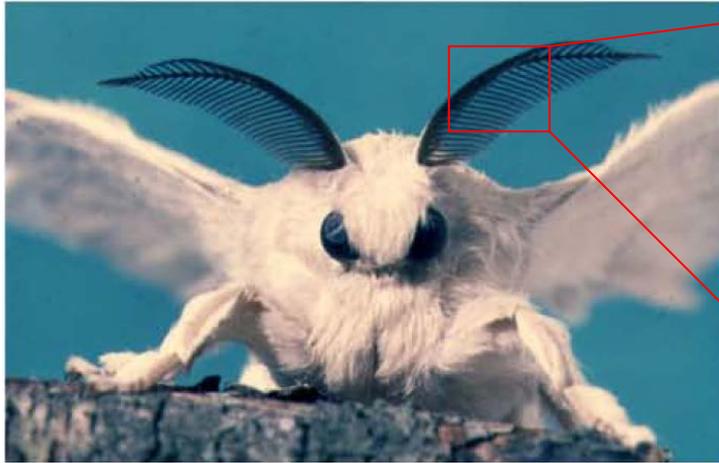
研究背景-桃蛀螟的生活史



研究背景-桃蛀螟的危害



研究背景-嗅觉感器

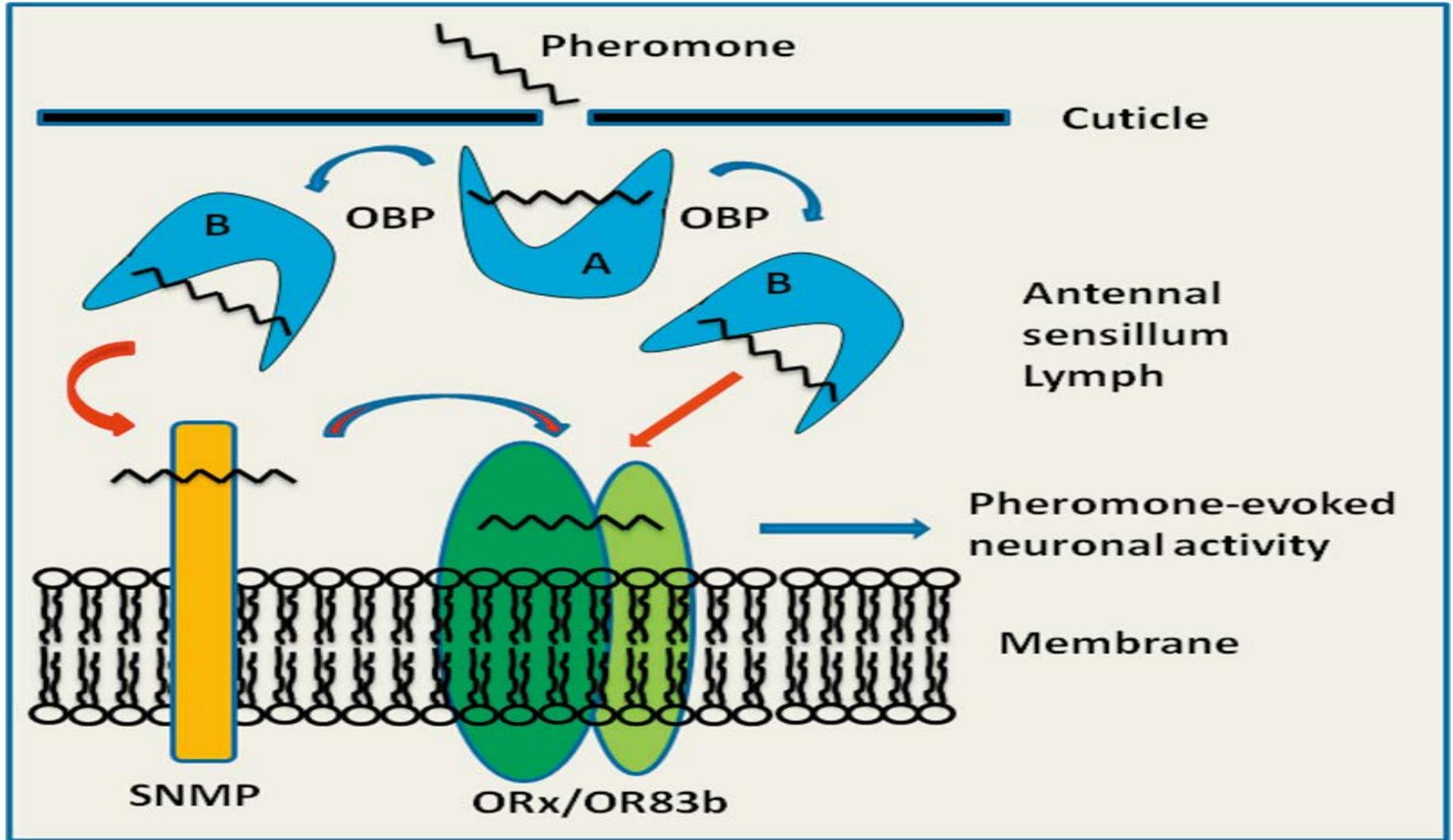


研究背景-气味结合蛋白（OBP）

气味结合蛋白（odorant binding protein）

➤ 信息素结合蛋白（pheromone binding protein）

研究背景-分子嗅觉机制



序列的获得

cDNA文库：以mRNA为模板，经反转录酶催化，在体外反转录成cDNA，与适当的载体连接后，转化到受体菌中，则每个细菌含有一段cDNA，并能繁殖扩增，这样包含着细胞全部mRNA信息的cDNA克隆集合称为该组织细胞的cDNA文库。

序列的获得

CAGGATAATCCAGCTTGCATGCCTGCAGGTCGACTCTAGAGG
ATCCCAAGCAGTGGTATCAACGCAGAGTACAGGGGAGTTTCC
GTTTTCTTTTGAAGAAGACAAGATGGCGGCGATATTCAAATG
GCGGTTGGTAGCCATTTTGGTTTTAGGTTTAGCTGTCAACGTA
CGAGTGGAAGCCTCTCAGGAGGTGATGAAGAAAATGAGCGC
AACGTTTTTT.....

序列分析-利用VecScreen程序查找载体序列

Distribution of Vector Matches on the Query Sequence



Match to Vector: ■ Strong ■ Moderate ■ Weak

Segment of suspect origin: ■

Segments matching vector:

Strong match: 10-62

Moderate match: 63-72

Suspect origin: 1-9

序列分析-利用DNAsstar软件查找ORF

ATGGCGGCGATATTCAAATGGCGGGTTG
GTAGCCATCTTGGTTTTGGGTTTAGCTG
TCAACGTACGAGTGAAAGCCTCTCAGG
AGGTGATGAAGAAA...GCCTAG

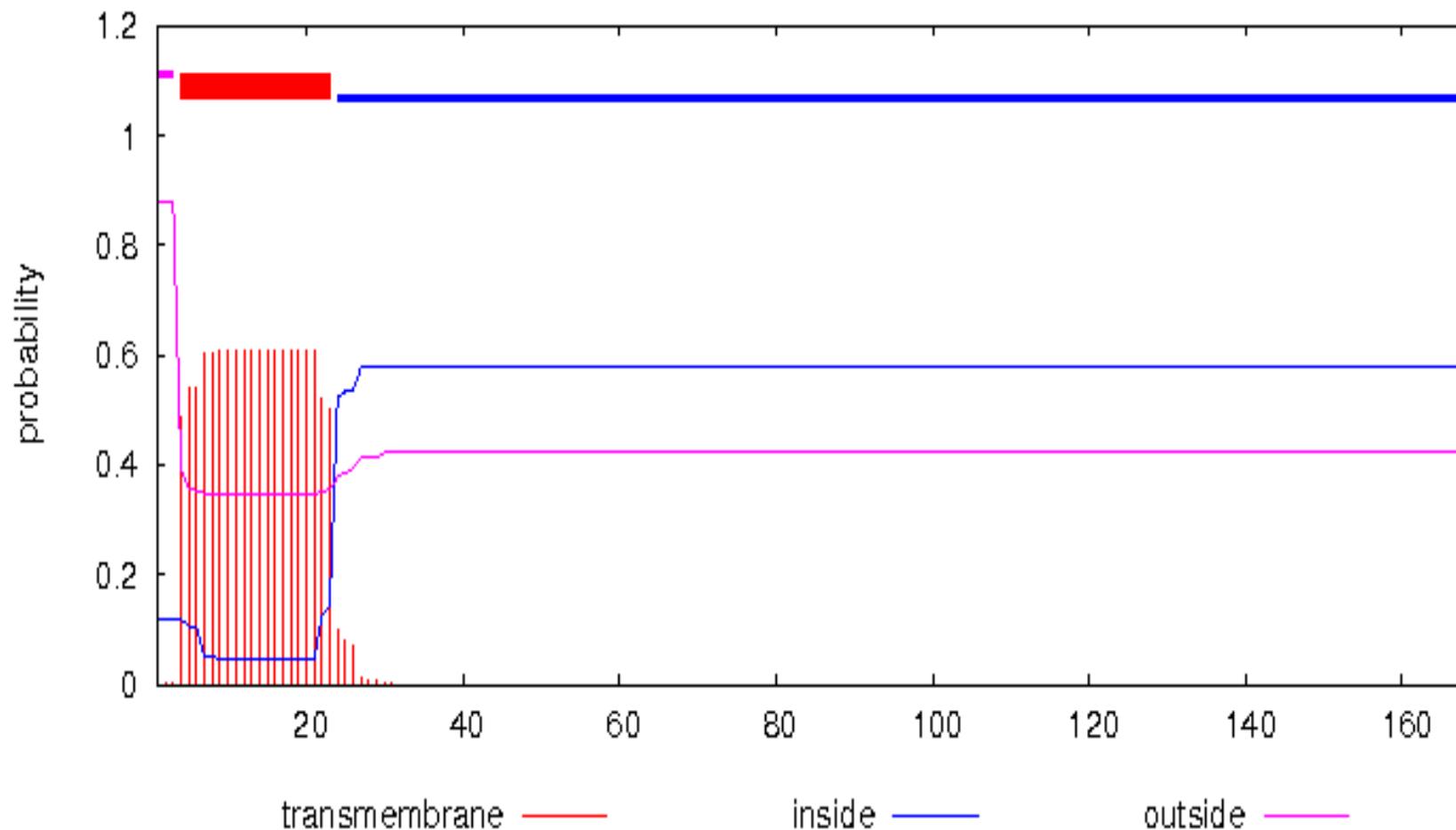
利用DNASTAR软件将CDs区翻译成氨基酸序列

MAAIFKWRLVAILVLGLAVNVRVKASQEVMKK

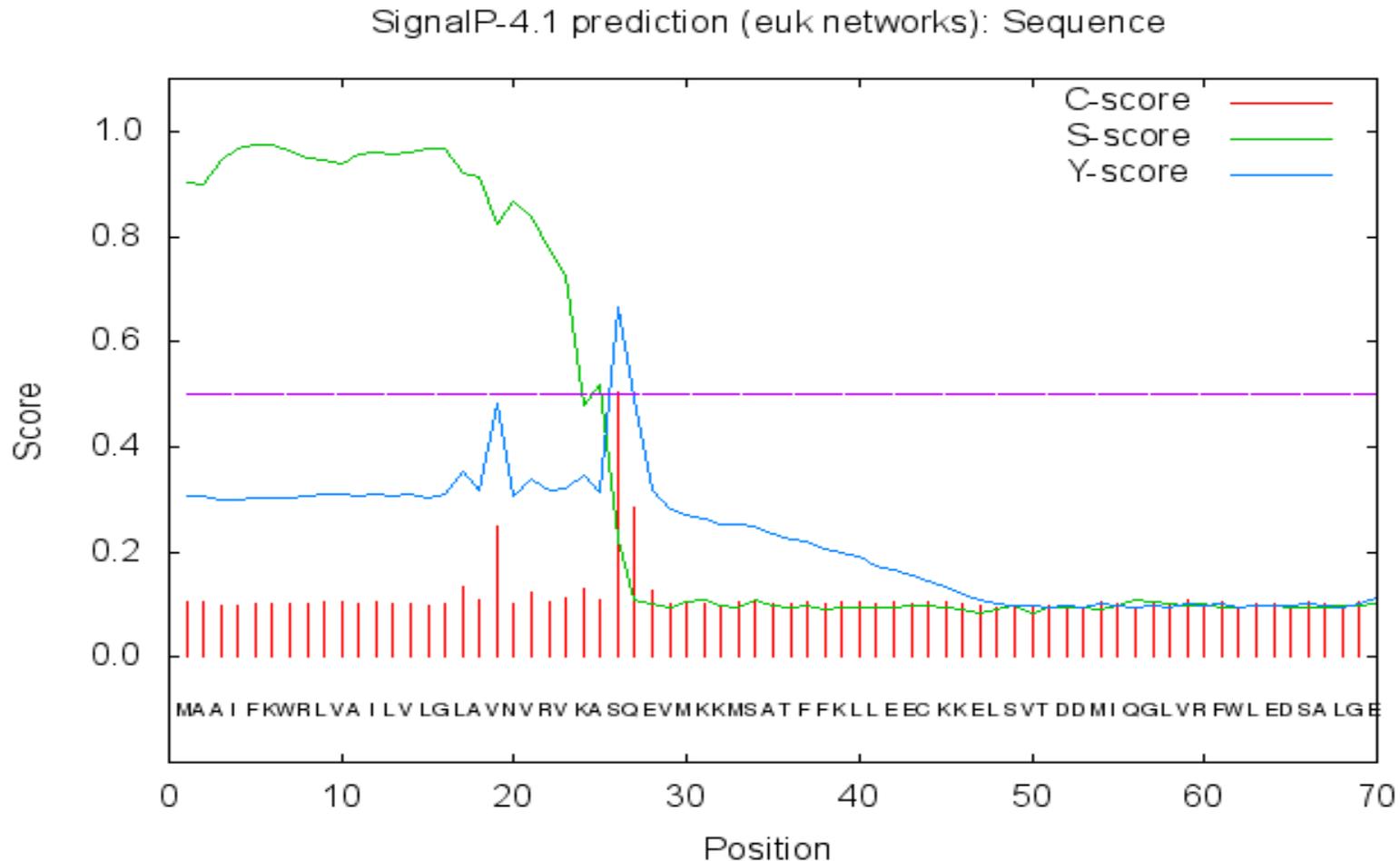
MSATFFKL...EA.

利用TMHMM程序对序列跨膜结构进行预测

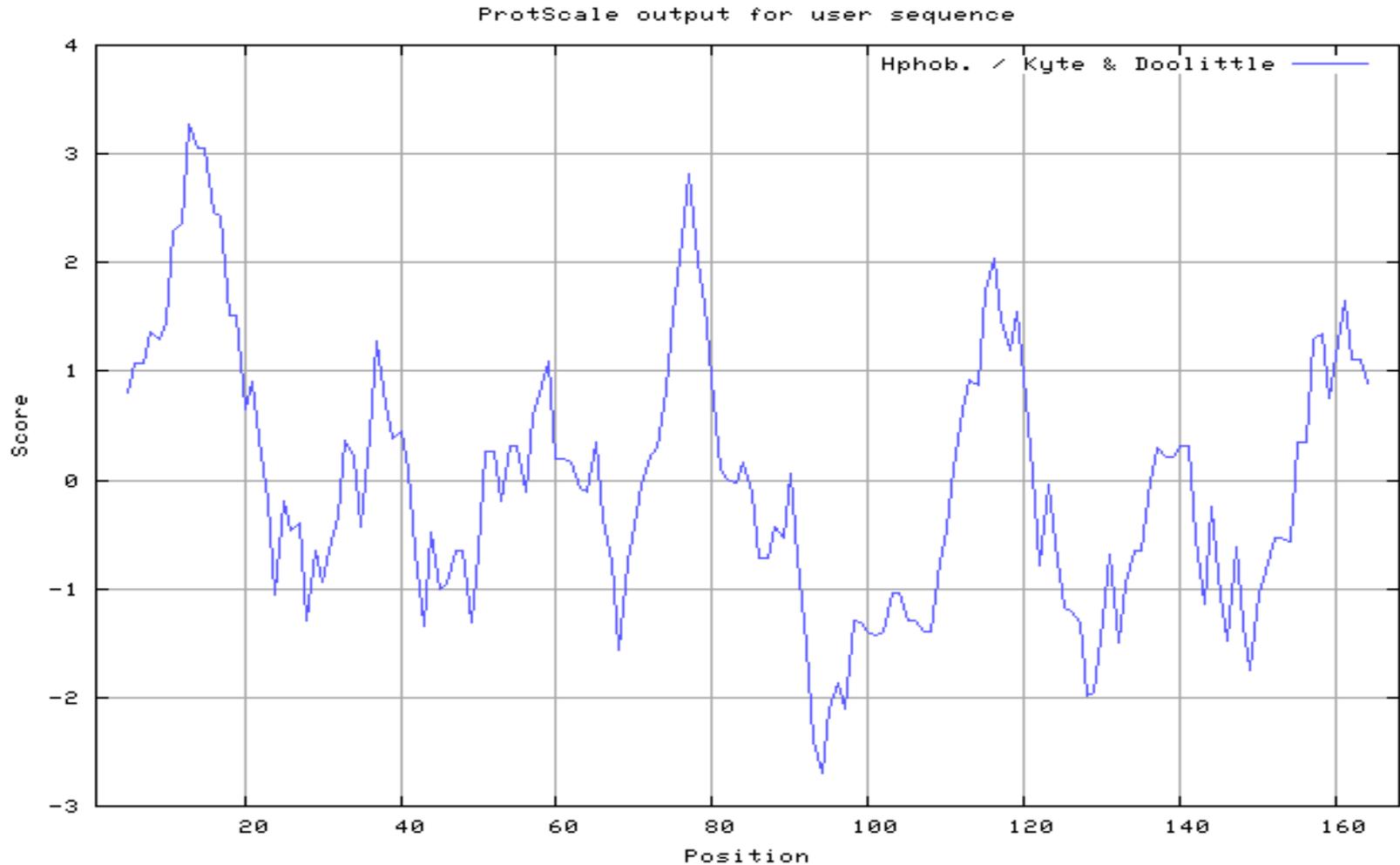
TMHMM posterior probabilities for WEBSEQUENCE



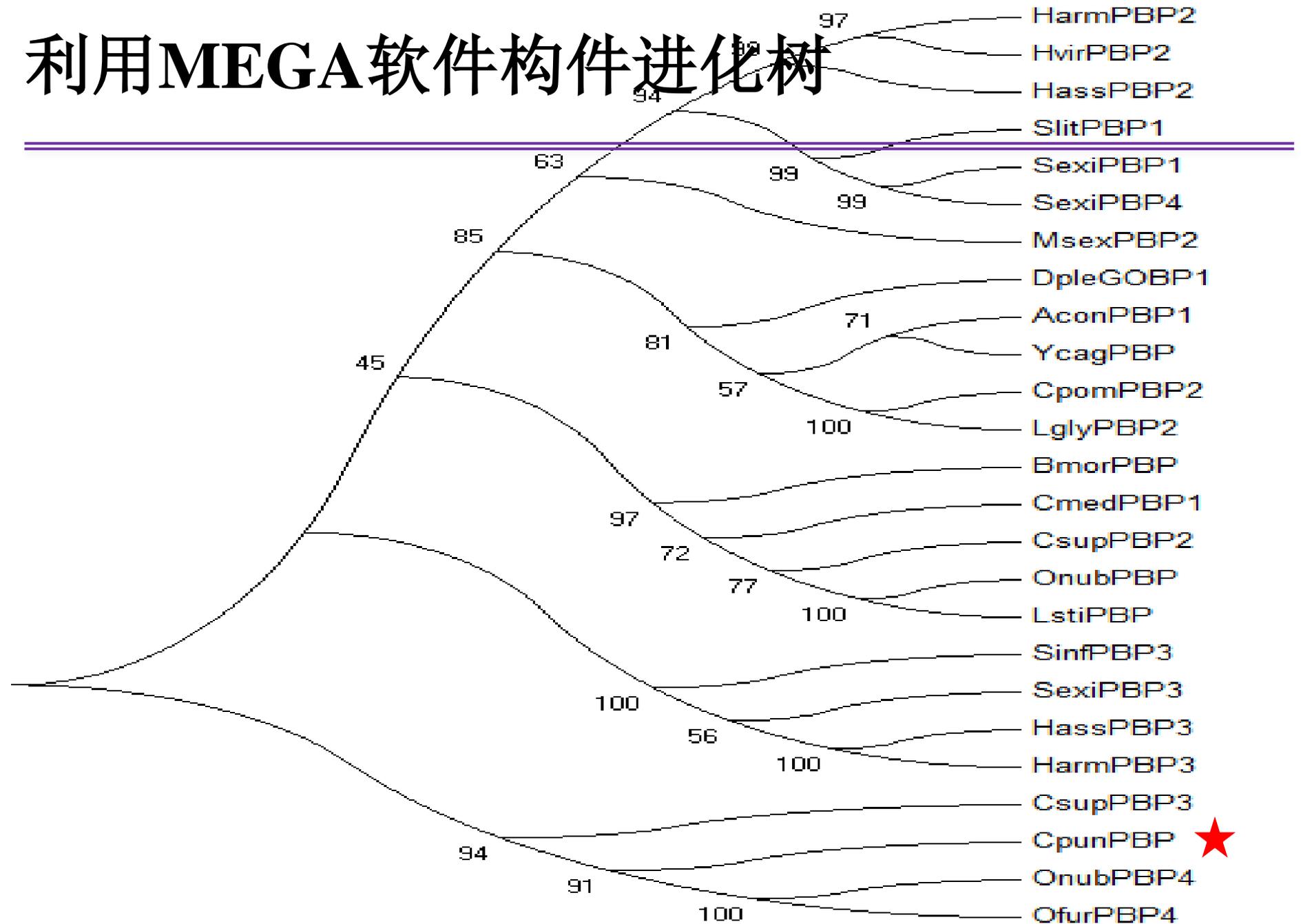
利用SignalP-4.1 Server程序预测序列信号肽



利用ProtScale程序对序列疏水性进行预测



利用MEGA软件构件进化树



利用DNAMAN软件进行多序列比对

OnubPBP4	VTTFNVLE	CKKELKVT	TNINEGLVRF	WSQGAAP	ERE	IGCVF	IQMAHKKD	ILEDQK	RLHHENAH	CFAKGHG	ADDKATE	IVSLLR	ICEQ	QFITT	D	DCSR	ALEVAR	CFE	QAHMQRI	QWAE	SMEV	MVB	EBILA																																																																																			
Ofury	VTTFNVLE	CKKELKVT	TNINEGLVRF	WSQGAAP	ERE	IGCVF	IQMAHKKD	ILEDQK	RLHHENAH	CFAKGHG	ADDKATE	IVSLLR	ICEQ	QFITT	D	DCSR	ALEVAR	CFE	QAHMQRI	QWAE	SMEV	MVB	EBILA																																																																																			
DpleG0BP1	TCFVVKAME	CKKELNLGDHI	IQDFMN	WRREEY	ELLNRI	TGCAI	IQMA	SKHDI	ITEDM	KIHHENAH	EFAKSHG	ADDLAKQ	LVQMIH	ICEK	QFTDI	D	DCSK	LEISK	CFE	RTKI	HDI	KWAE	SMETIL	EBLMT																																																																																		
AconPBP4	AGASTALD	CKKELNIQDHI	IQDFYN	WRREDY	ALLNRI	TGCAI	IQMA	VQDI	ITEDI	KMHGKA	HEFAKTHG	ADDLAKQ	LVSMIH	ICEK	LHSG	CT	IECAT	TL	EIK	CFE	ASKI	HDI	KWAE	SMEV	WLBVMT																																																																																	
CpomPBP2	TCFATALD	CKKELNVQEN	VQDFYN	WRREDY	ELLNRI	TGCVI	IQMA	LKFDI	IDEDA	KLHKN	AHEFAKTHG	ADDLAKQ	LVGMIH	ICEK	ATEA	.	DCI	RTLD	IA	CFE	RTKI	HDI	KWAE	SMETV	LBVMT																																																																																	
HarmPBP2	SGFTKVV	CKKELSVGDH	IQDMYN	WRREEY	QIVNRI	TGCVI	IQMA	AKLDI	IGDDQ	RMHGG	KAEEFAKSHG	ADDALAKQ	LVGLIH	ICEK	QHQ	QA	IE	DHCS	RALE	IA	CFE	RTKI	HDI	KWAE	SMEV	IMB	IMT																																																																															
MsexPBP2	VGF	SKVLQ	CKKELSVGDH	IQDFYN	WRREDY	LLNRI	TGCMV	IQMA	VKHD	I	INDQL	TMHGG	NAHAFV	KTHG	ADDTA	Q	Q	LV	TILR	ICE	A	KHQS	VE	D	V	CN	RA	LE	MA	CFE	RTKI	HDI	KWAE	AMEV	V	LB	BIM																																																																					
HaasPBP3	SGFVKVLE	CKKELN	LNELQ	ILADL	FHWK	LEYS	ILGR	TGCAI	IQMS	KKLD	I	LDANG	RMHGG	NAAE	FAK	KHGA	G	DEV	ASK	IV	T	I	I	H	ICEK	K	HE	Q	D	G	D	E	C	L	R	V	L	E	A	K	F	E	RTGI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																														
HaasPBP2	EGFTKVV	CKKELNVGDH	IQDMYN	WRREEY	QIVNRI	TGCVI	IQMA	VAKLD	IGDDQ	KMHG	KAEEFAKSHG	ADDVLA	KQ	LVSLIH	ICEK	QHQ	QA	IE	DHCS	R	V	L	E	I	A	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																																												
SexiPBP3	SGFVKVLE	CKKELN	MNDHI	IADL	FHWK	LEYS	ALLSRI	TGCVI	IQMS	KKLD	I	LDANG	RMHGG	NAAE	FAK	RHGA	G	DDV	ASK	IV	Q	I	I	H	ICEK	K	H	E	R	D	D	D	E	C	L	R	V	L	E	A	K	F	E	RTGI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																													
YcagPBP	VGF	SKALD	CKKELAI	QENVL	QDFYN	WRREDY	TLVNRI	TGCVL	IQMA	SKFD	I	ITEDM	KVHHK	NAHE	FAKTHG	ADDE	MA	KQ	LV	SMIH	ICEK	T	H	E	G	V	V	D	C	R	V	L	E	M	A	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	L	B	V	M	T																																																			
LglyPBP2	TCFAMALD	CKKELIG	IQDNI	IQDFN	WRREDY	ELLNRI	TGCVI	IQMA	LKFD	I	IDEDS	KLDH	KNAHE	FAKTHG	ADDLAKQ	LV	TMIH	ICEK	Q	N	T	E	A	.	DCI	R	T	I	N	I	A	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	B	V	M	T																																																			
SinfPBP3	SGFVKVLE	CKKELN	MNDHI	IADL	FHWK	LDY	TLNRI	TGCAI	IQMS	KKLD	I	LDVNG	RMHGG	NARE	FALK	HGA	G	DEV	A	K	I	V	D	I	I	H	ICEK	K	F	E	R	D	D	D	E	C	L	R	V	L	E	A	K	F	E	RTGI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																											
HarmPBP3	SGFVKVLE	CKKELDL	NEQ	ILADL	FHWK	LEYS	ILGR	TGCAI	IQMS	KKLD	I	LDANG	RMHGG	NAAE	FAK	KHGA	G	DEV	ASK	IV	T	I	I	H	ICEK	K	HE	Q	D	G	D	E	C	L	R	V	L	E	A	K	F	E	RTGI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																														
SlitPBP1	KCFTRVVD	CKKELNVGDH	IQDMYN	WRREDY	QILNRI	TGCVL	IQMA	AKLD	I	DDQ	TMHGG	K	T	E	D	FAK	SHG	ADDV	A	K	L	V	S	V	I	H	ICEQ	Q	H	T	G	IA	D	C	M	R	V	L	E	A	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																													
SexiPBP4	KCFTRVVD	CKKELNA	VEH	IQDMYN	WR	CE	DH	QILNRI	TGCVI	IQMA	KNLD	I	DDQ	KMHG	K	T	E	E	FAK	SHG	ADDE	V	A	K	L	V	S	I	I	H	ICEQ	Q	H	A	G	IA	D	C	M	R	V	L	E	I	S	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																								
CsupPBP3	ISFLKVLQ	CKKELSV	PEEVL	QSLMT	FWNQ	DTL	SHRI	TGCVI	IQV	SKLD	I	I	E	L	E	T	Y	K	L	H	P	D	N	A	N	E	Y	V	K	K	H	G	A	D	D	E	T	A	S	Q	I	M	N	I	L	R	ICEI	K	N	E	A	I	T	D	H	C	D	R	V	R	E	I	A	K	F	E	H	G	M	H	D	I	K	W	A	E	N	M	E	V	I	I	N	D	L	V	A																			
OnubPBP	INFGKALD	CKKELDL	PDS	INAD	FYN	WR	KEG	YEL	SNRI	TGCAI	IQLS	SKLD	I	V	D	PEG	KLHGG	N	T	H	E	F	A	K	K	H	G	A	D	D	S	M	A	K	Q	L	V	E	L	I	H	ICEG	S	V	A	D	D	P	D	A	C	M	K	V	I	N	I	A	K	F	E	K	A	E	I	H	K	I	N	W	A	E	S	M	D	L	I	V	A	B	V	L	A																							
SexiPBP1	KCFTRVVD	CKKELN	AGEH	IQDMYN	WRREDY	QILNRI	TGCVI	IQMA	AKLD	I	DDQ	KMHG	K	T	E	E	FAK	SHG	ADDE	V	A	K	L	V	S	I	I	H	ICEQ	Q	H	A	G	IA	D	C	M	R	V	L	E	I	S	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																										
CsupPBP2	LNFGKAYD	CKKEL	E	L	P	N	E	V	D	T	D	F	N	W	R	K	D	D	Y	A	V	T	N	R	I	TGCAI	IQMS	N	K	L	D	L	D	P	G	KMHG	N	A	R	E	F	A	K	K	H	G	A	D	D	S	M	A	Q	L	I	D	I	L	H	ICEK	G	A	S	P	G	P	D	G	I	A	C	V	Q	V	L	E	I	S	K	F	E	K	V	E	I	H	K	I	N	W	A	E	S	M	D	L	I	M	A	B	V	L	A			
HvirPBP2	SGFTKVV	CKKELNVGDH	IQDMYN	WRREEY	QIVNRI	TGCVI	IQMA	AKLD	I	V	G	DDQ	KMHG	KAEEFAKSHG	ADDALAKQ	LVGLIH	ICEK	QHQ	QA	IE	DHCS	R	T	L	E	V	A	K	F	E	RTKI	H	D	I	D	W	Q	E	K	V	E	V	I	S	D	V	L	T																																																										
LstiPBP	INFGKALD	CKKELDL	PDS	INAD	FYN	WR	KEG	YEL	SNRI	TGCAI	IQLS	SKLD	I	V	D	PEG	KLHGG	N	T	H	E	F	A	K	K	H	G	A	D	D	A	M	A	K	Q	L	V	D	L	I	H	ICES	D	V	P	D	D	P	D	E	C	L	K	V	I	N	I	A	K	F	E	K	A	E	I	H	K	I	N	W	A	E	S	M	D	L	I	M	A	B	V	L	A																							
BmorPBP	IQFAKPLE	CKKEM	GLT	T	E	T	V	L	K	D	F	Y	N	W	I	E	D	Y	E	F	T	D	R	I	TGCAI	IQMS	KKLD	I	D	M	D	G	D	Y	NL	H	G	K	A	H	E	F	A	R	K	H	G	A	D	E	T	M	A	K	Q	L	V	D	L	I	H	ICES	Q	S	V	A	T	M	P	D	E	C	E	R	T	L	K	V	A	K	F	E	K	A	E	I	H	K	I	N	W	A	E	D	V	E	L	L	M	A	B	V	L	N		
CpunPBP	ATDFKLE	CKKELSV	T	D	D	I	Q	G	L	V	R	W	L	E	D	S	A	L	G	E	R	I	TGCVI	IQMA	E	K	Q	D	I	V	V	T	E	D	Y	R	M	H	E	N	A	Y	N	E	F	A	K	N	H	G	A	D	D	A	M	A	T	A	I	V	K	V	I	H	ICEE	Q	F	T	S	N	P	D	H	C	A	R	V	M	E	V	S	K	F	E	R	D	E	I	H	R	I	K	W	A	E	S	I	E	L	L	I	G	B	M	L	G

利用garnier程序预测氨基酸二级结构

```
      .   10   .   20   .   30   .   40   .   50
      MAAIFKWRLVAILVLGLAVNVRVKASQEVMMKMSATFFKLLLEECKKELSV
helix  HHHHHHHHHH      HHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
sheet          EEEEE   E
turns
coil

      .   60   .   70   .   80   .   90   .  100
      TDDMIQGLVRFWLEDSALGERELGCVIICMAEKQDLVVTEDYRMHHEENAY
helix  HHHHH      HHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
sheet          EEEEEEE
turns
coil          CC

      .  110   .  120   .  130   .  140   .  150
      NFAKNHGADDAMATAIVKVIHTCEEQFTSNPDHCCARVMEVSKCFRDEIHR
helix  HHHH      HHHHHHHHHHHHHH      HH      HHHHHHHHHHHHHHHHHHHHH
sheet          E
turns          TTTTT TTT
coil          CCC      CC      C

      .  160
      LKWAPSIPELLIGEMLGEA
helix          HHHHHHHHHHHHHH
sheet
turns  TT
coil   CCC
```

利用Swiss-model进行蛋白三级结构预测

 **BIOZENTRUM**
Universität Basel
The Center for Molecular Life Sciences

 **SWISS-MODEL Workspace**

welcome ge_xiny_zz@icb.unibas.ch

Modelling Tools Repository Documentation

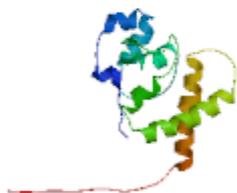
[myWorkspace] [Settings] [logout]

Workunit: P000002 - Overview



Print/Save this page as 

Model Summary 



Model information:

Modelled residue range: 26 to 167
Based on template: [1qwwA] (99.9 Å)
Sequence Identity [%]: 43.66
Evaluate: 0.00e-1

Quality information:

QMEAN Z-Score: -5.77

[details]



Quaternary structure information: [details]

Template (1qww): MONOMER
Model built: SINGLE CHAIN

Ligand information: [details]

Ligands in the template: none.
Ligands in the model: none.

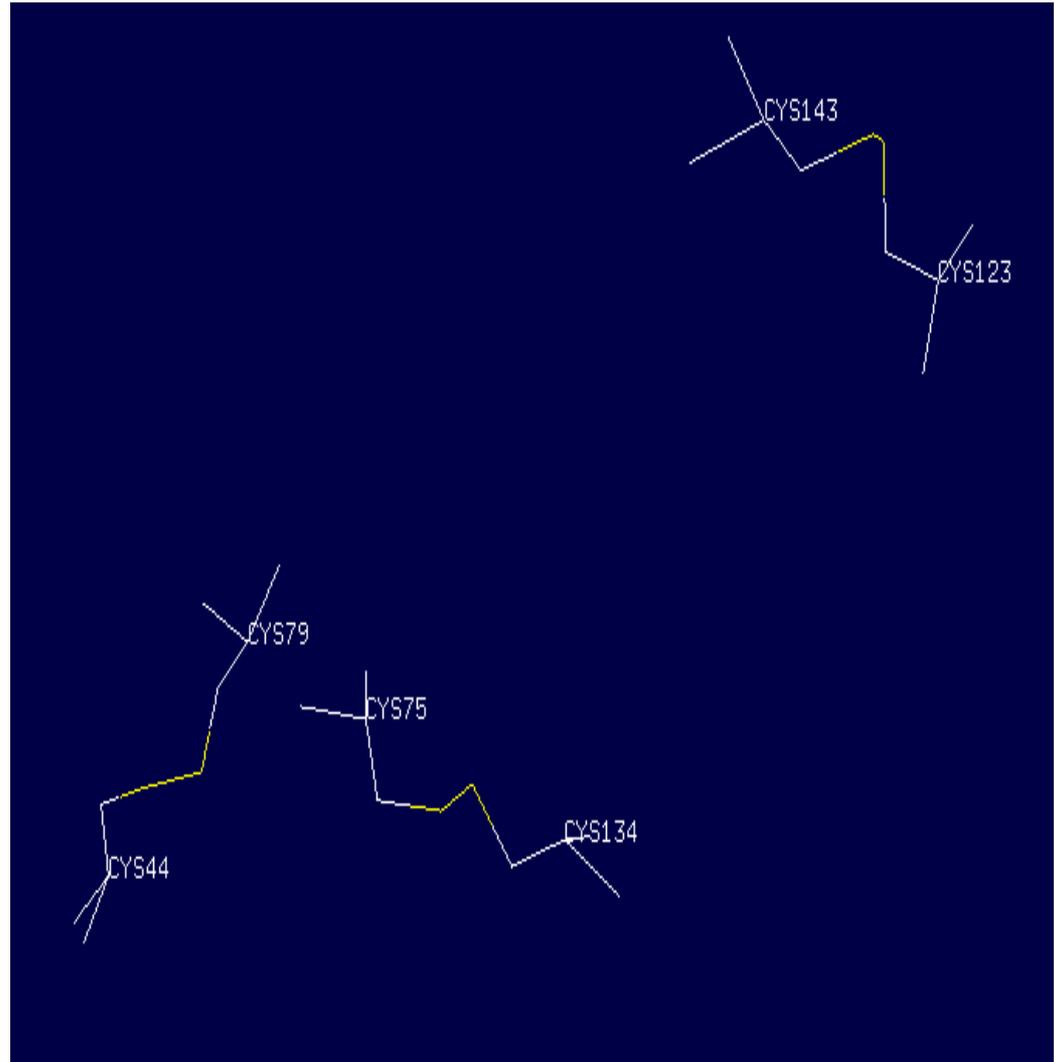
Warning: Low QMEAN Z-scores! Only membrane proteins or models of poor quality are expected to reach such low scores.

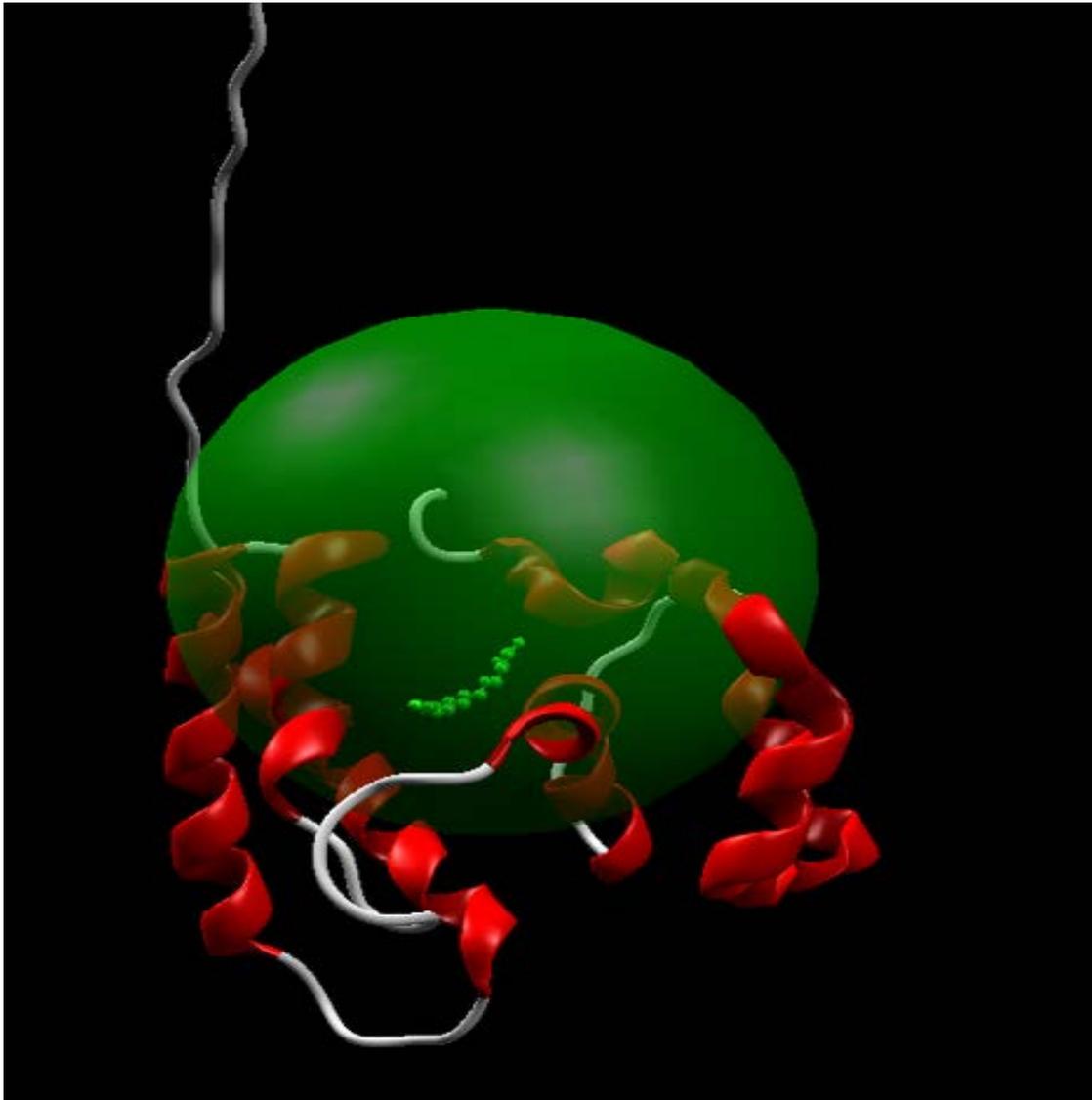
logs: [Templates] [Alignment] [Modeling]

display model: as [pdb] - as [DeepView project] - in [AstexViewer]

download model: as [pdb] - as [Deepview project] - as [text]

利用swiss-prot程序分析蛋白三级结构





本图为molegro
virtual docker软
件分析的结果，
中间绿色的为信
息素，不知道这
个结果对不对

后续试验

- 原位杂交
- 定点突变

Thank you!

