

# 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

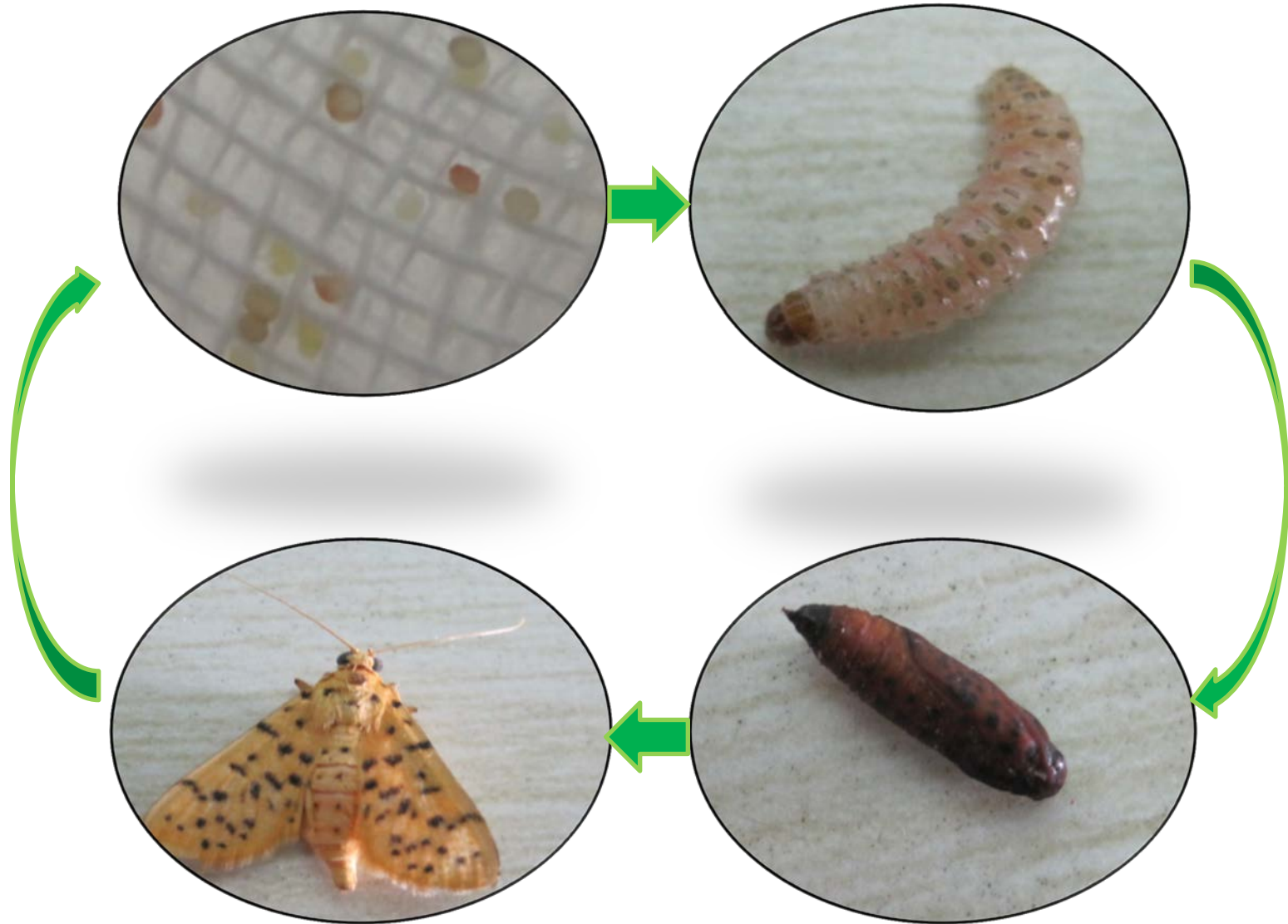
序列的获得

序列的生物信息分析

后续试验

# 研究背景-桃蛀螟的生活史

---



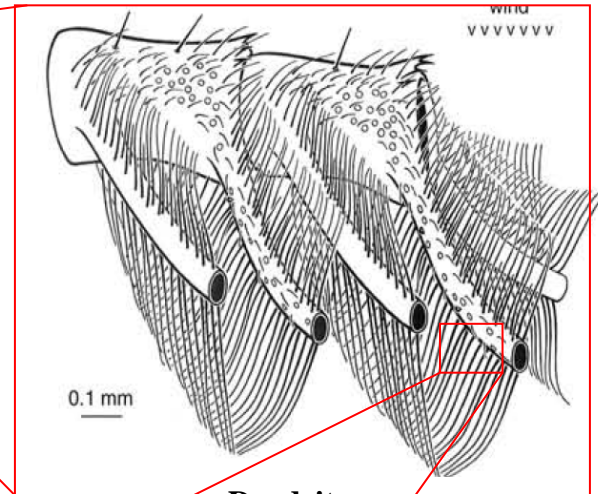
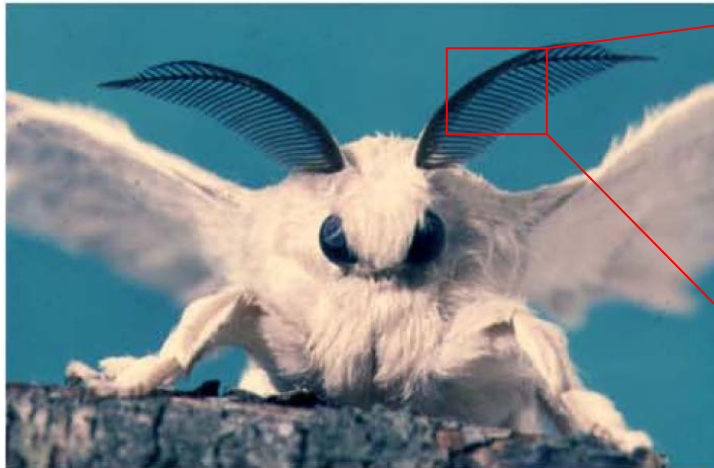
# 研究背景-桃蛀螟的危害

---

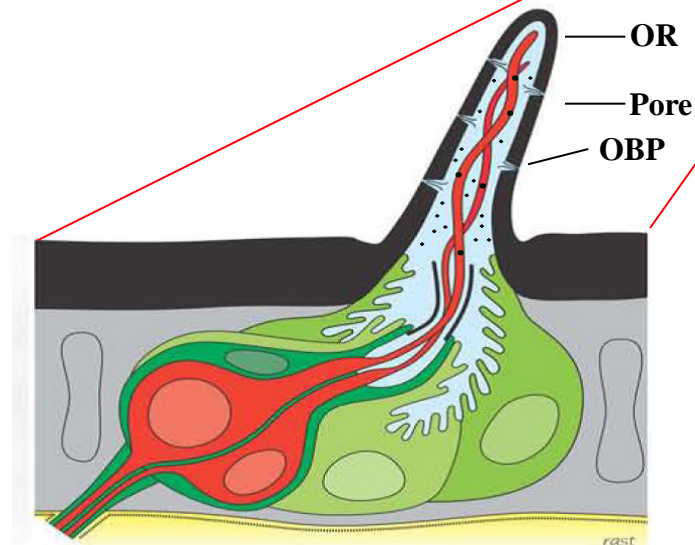




# 研究背景-嗅觉感器



Dendrite



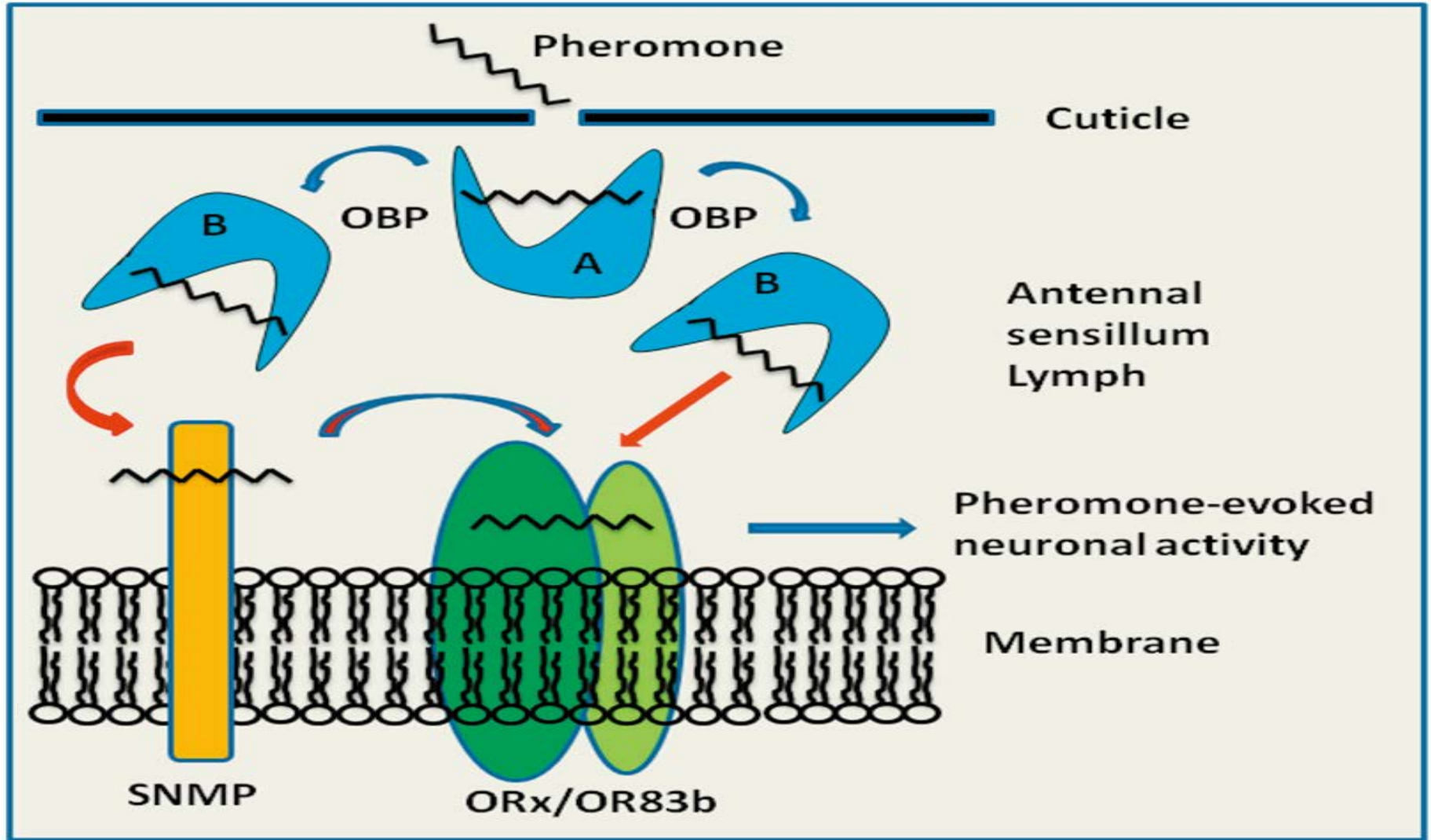
# 研究背景-气味结合蛋白（OBP）

---

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。



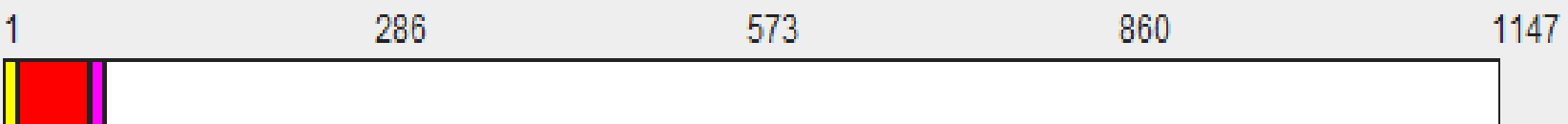
# 序列的获得

---

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DNAs<sub>tar</sub>软件查找ORF

---

ATGGCGGCGATATTCAAATGGCGGGTTG  
GTAGCCATCTTGGTTTTGGGTTTAGCTG  
TCAACGTACGAGTGAAAGCCTCTCAGG  
AGGTGATGAAGAAAA...GCCTAG

# 利用DNASTAR软件将CDs区翻译成氨基酸序列

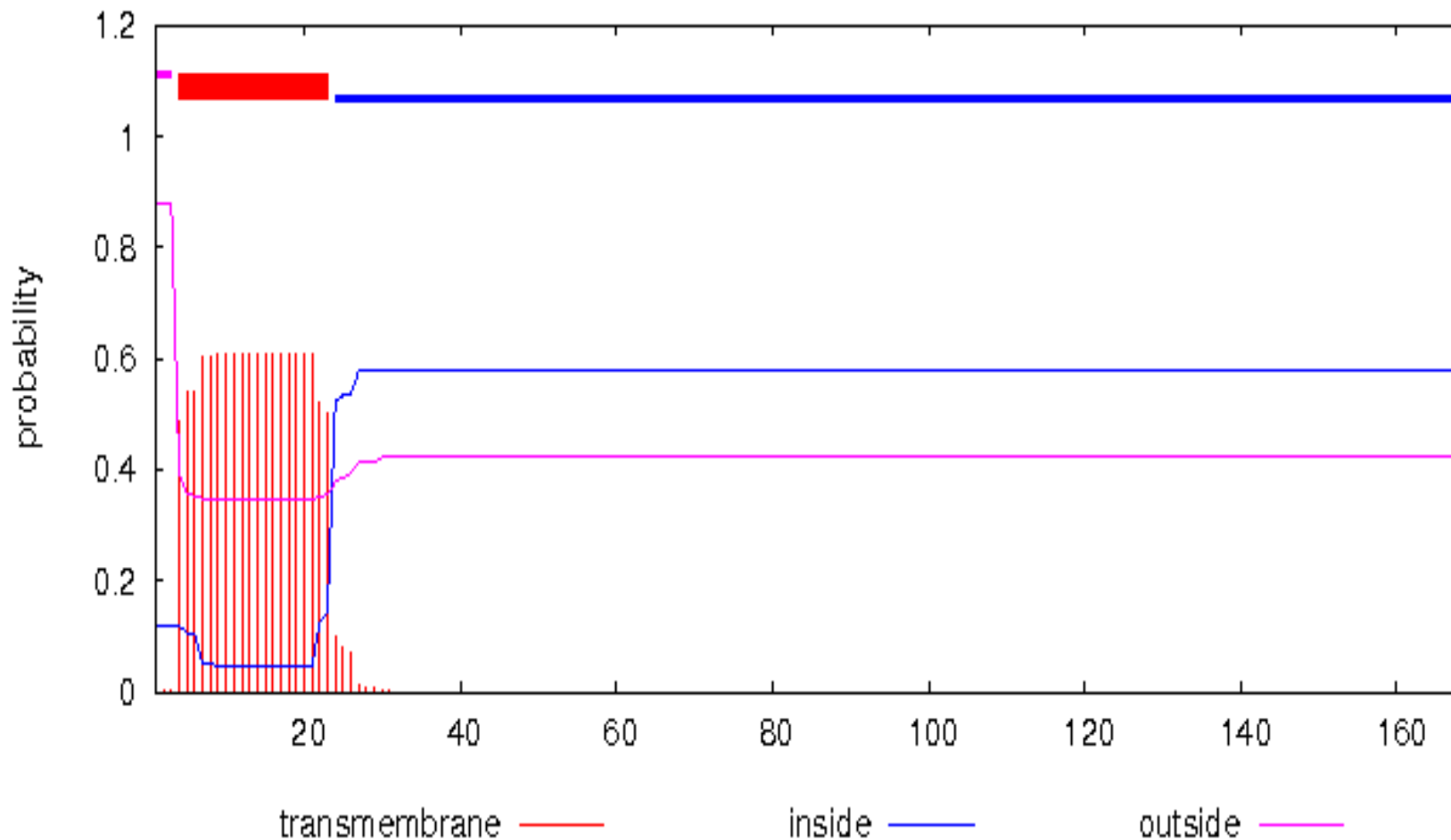
---

MAAIFKWRLVAILVLGLAVNVRVKASQEVMKK

MSATFFKL...EA.

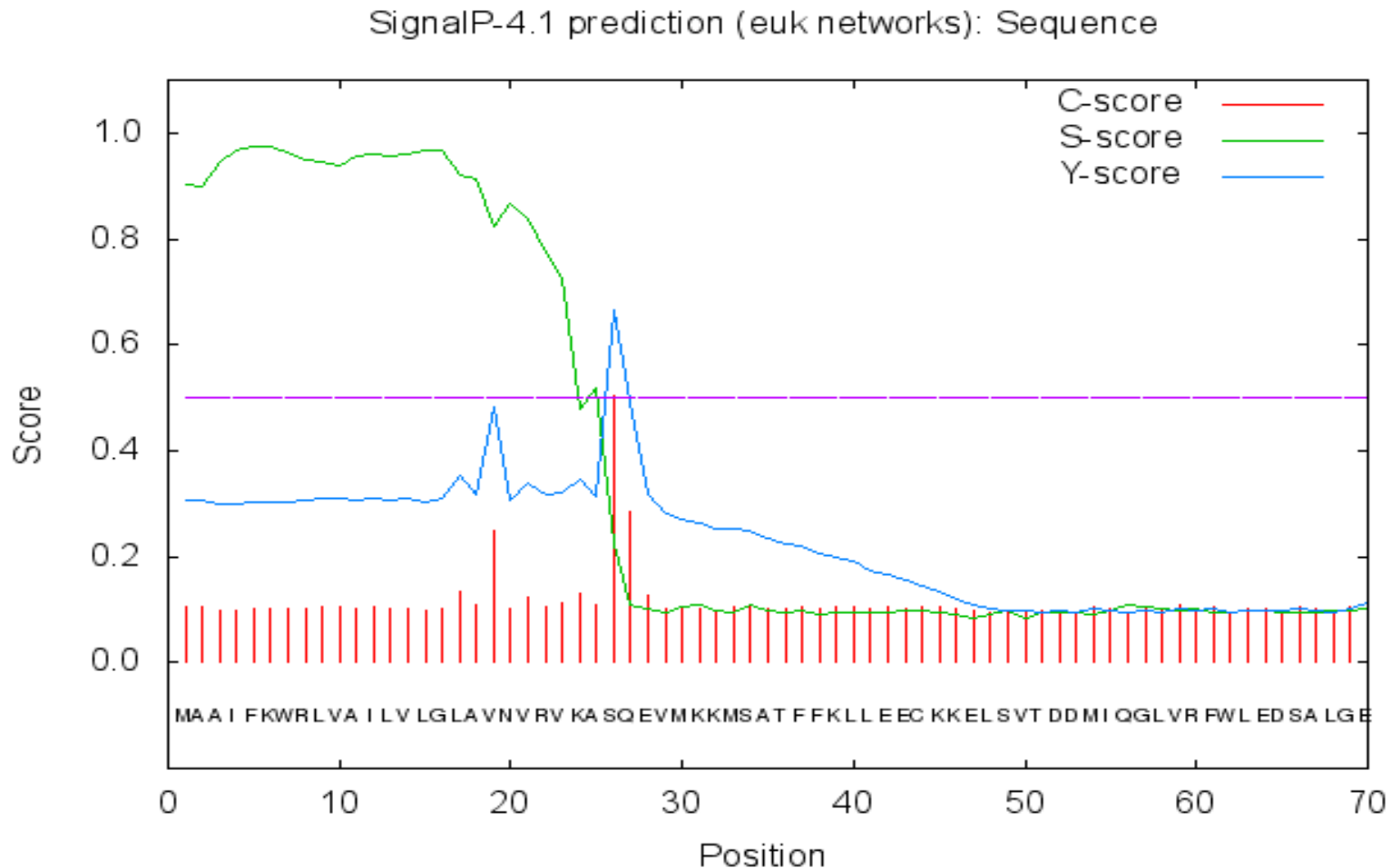
# 利用TMHMM程序对序列跨膜结构进行预测

TMHMM posterior probabilities for WEBSEQUENCE

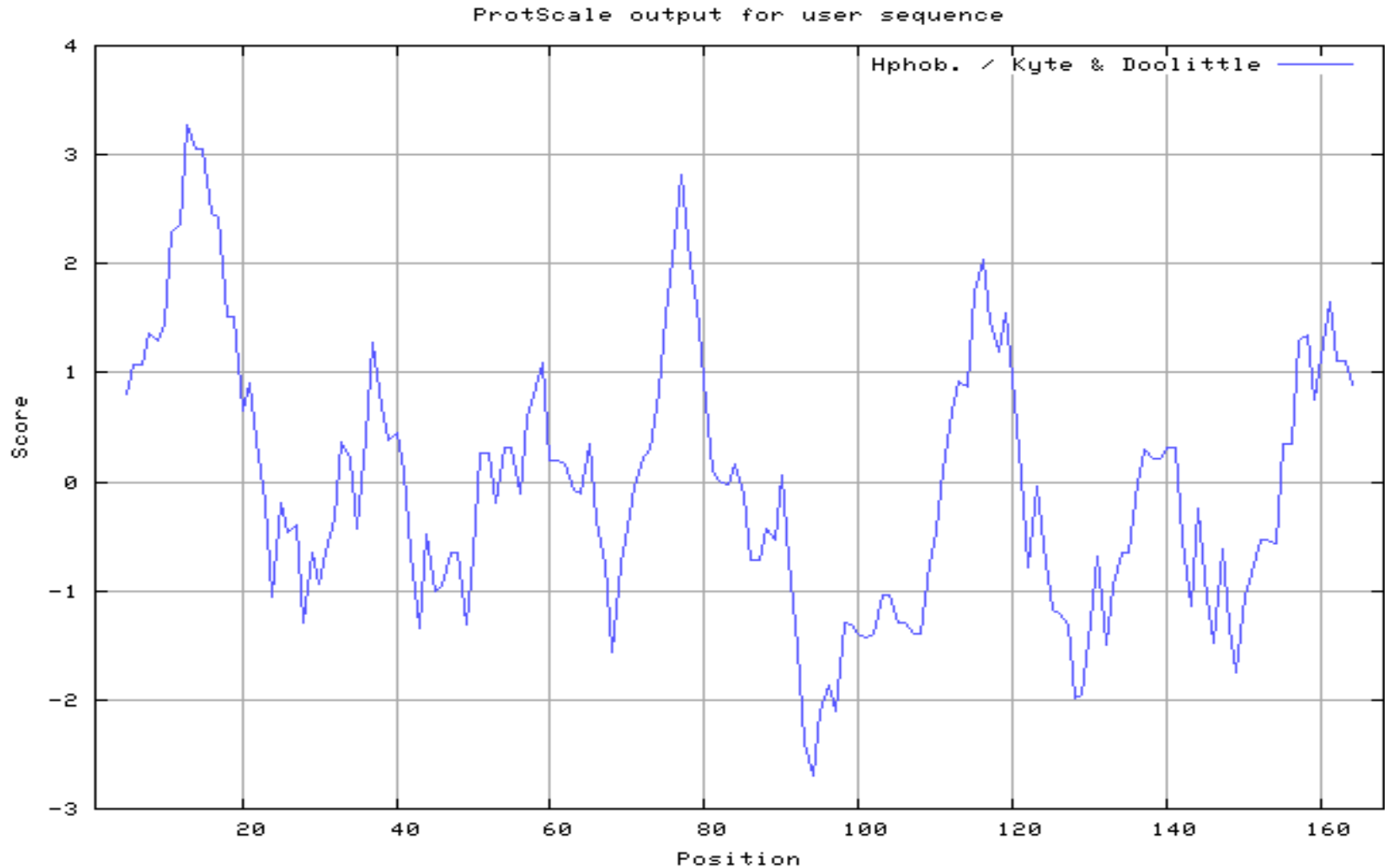




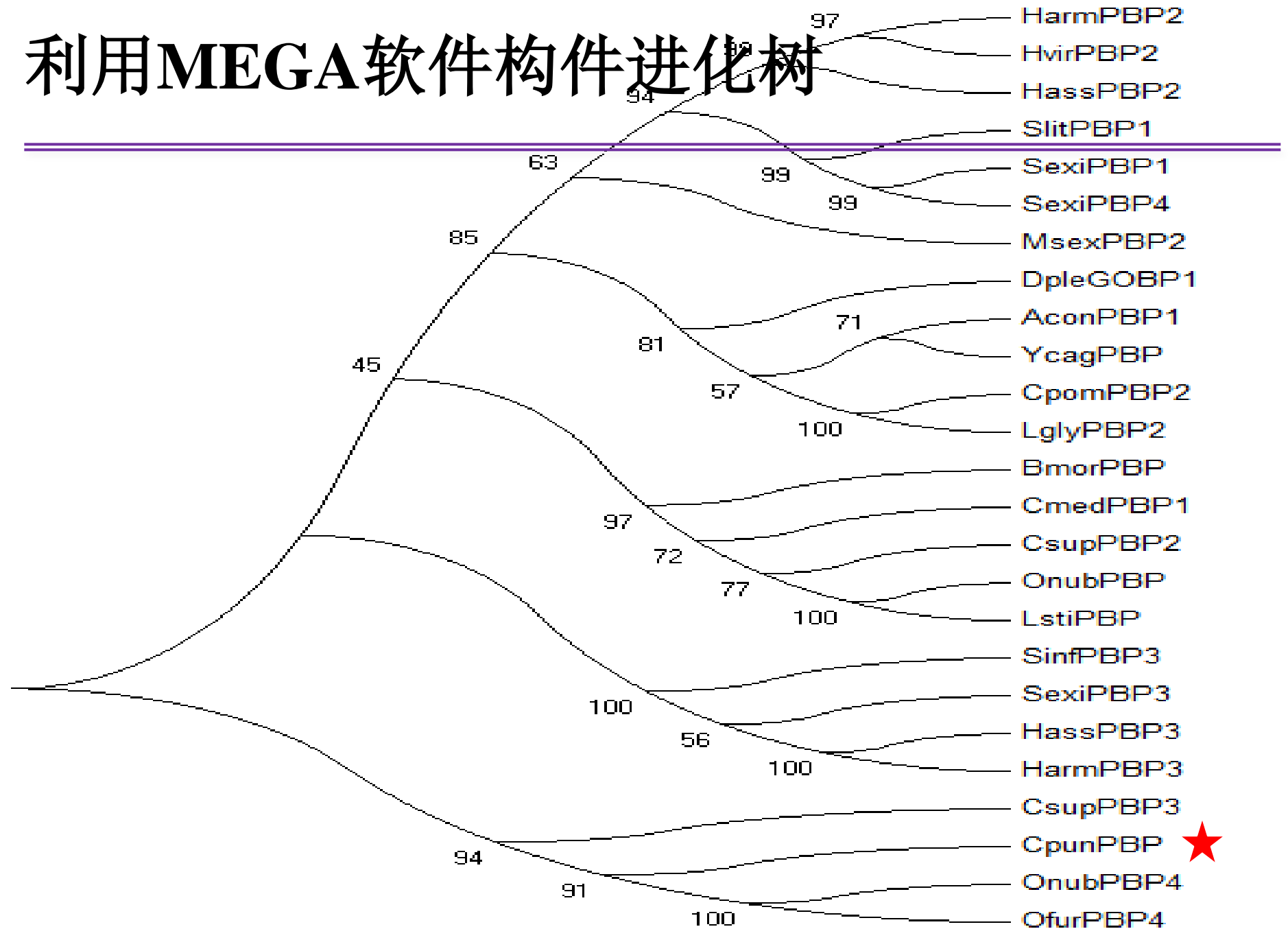
# 利用SignalP-4.1 Server程序预测序列信号肽



# 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	QFITITD	DCSR	ALEVAR	CFE	QAHMQRI	QWAE	SMEV	MVB	EBILA																																														
Ofury	VTTFNVLE	CKKELKVT	TNINEGLVRF	WSQGAAP	ERE	IGCVF	IQMAHKKD	ILEDQK	RLHHENAH	CFAKGHG	ADDKATE	IVSLLR	ICEQ	QFITITD	DCSR	ALEVAR	CFE	QAHMQRI	QWAE	SMEV	MVB	EBILA																																														
DpleG0BP1	TCFVVKAME	CKKELNLGDH	IQDFMN	WRREEY	ELLNR	TGCAI	IQMA	SKHDI	ITEDM	KIHHENAH	EFAKSHG	ADDLAKQ	LVQMIH	ICEK	QFTDIT	DCSK	LEISK	CFE	RTKI	HED	KWAE	SMET	ILBBLMT																																													
AconPBP4	AGASTALD	CKKELNIQDHI	IQDFYN	WRREDY	ALVNR	MGCVI	IQMAVQ	LITEDI	KMHHGKA	EEFAKTHG	ADDLAKQ	LVSMIH	ICEK	LHSGOCT	DECAT	TLEIA	KCFE	ASKI	HED	KWAE	SMEV	WLBVMT																																														
CpomPBP2	TCFATALD	CKKELNVQEN	VMDQFYN	WRREDY	ELLNR	TGCVI	IQMA	LKFDI	IDEDA	KLHKKNA	EFAKTHG	ADDLAKQ	LVGMIH	ICEK	ATTEA	DCI	RTLDIA	KCFE	RTKI	HED	KWAE	SMET	VLBVMT																																													
HarmPBP2	SGFTKVV	CKKELSVGDH	IQDMYN	WRREEY	QIVNR	TGCVI	IQMA	AKLDI	IGDDQ	RMHHGKA	EEFAKSHG	ADDALAKQ	LVGLIH	CE	TQHQ	AE	DCSR	ALEIA	KCFE	RTKI	HED	KWAE	SMEV	IMBIMT																																												
MsexPBP2	VGF	SKVLQ	CKKELSVGDH	IQDFYN	WRREDY	LLNR	TGCMV	IQMA	VKHD	INDQL	TMHHGNA	HAFVKT	HGADD	TQQLV	TILR	CEAK	HQS	VE	DCN	RALEMA	KCFE	RTKI	HED	KWAE	EAMEV	VLBIM																																										
HaasPBP3	SGFVKVLE	CKKELNLNEQ	ILADLFH	WKLEYS	ILGR	TGCAI	IQMS	KKLD	ILDANG	RMHHGNAA	EFAKKHG	ADVEVASK	IVTIIH	ICEK	KHEQDGD	ECL	RVLEVA	KCFE	RTGI	HED	DWQ	EKVEV	IVSBVLT																																													
HaasPBP2	EGFTKVV	CKKELNVGDH	IQDMYN	WRREEY	QIVNR	TGCVI	IQMA	VAKLDI	IGDDQ	KMHHGKA	EEFAKSHG	ADDVLAKQ	LVSLIH	CE	TQHQ	AE	DCSR	VLEIA	KCFE	RTKI	HED	KWAE	SMEV	MVBIMT																																												
SexiPBP3	SGFVKVLE	CKKELNMNDHI	IADLFH	WKLEYA	LLSR	TGCVI	IQMS	KKLD	ILDANG	RMHHGNAA	EFAKRHG	ADDVASK	IVQIIH	ICEK	KHER	DDD	ECL	RVLEVA	KCFE	RTGI	HED	DWQ	EKVEV	IVSBVLT																																												
YcagPBP	VGF	SKALD	CKKELAIQEN	VLDQFYN	WRREDY	TLVNR	MGCVL	IQMA	SKFDI	ITEDM	KVHHKNA	EFAKTHG	ADDEMA	KQVSMIH	ICEK	THEG	VV	DCGR	VLEMA	KCFE	RTKI	HED	KWAE	SMEV	VLBVMT																																											
LglyPBP2	TCFAMALD	CKKELIGIQD	NIQDFFN	WRREDY	ELLNR	TGCVI	IQMA	LKFDI	IDEDS	KLDHKNA	EFAKTHG	ADDLAKQ	LVMTIH	ICEK	QNT	EA	DCI	RTINIA	KCFE	RTKI	HED	KWAE	SMET	VLBVMT																																												
SinfPBP3	SGFVKVLE	CKKELNMNDHI	IADLFH	WKLDY	TLNR	TGCAI	IQMS	KKLD	ILDVNG	RMHHGNAA	EFAKKHG	ADVEVAK	IVDIIH	ICEK	KFER	DDD	ECL	RVLEVA	KCFE	RTGI	HED	DWQ	EKVEV	IVSBVFT																																												
HarmPBP3	SGFVKVLE	CKKELDLNEQ	ILADLFH	WKLEYS	ILGR	TGCAI	IQMS	KKLD	ILDANG	RMHHGNAA	EFAKKHG	ADVEVASK	IVTIIH	ICEK	KHEQDGD	ECL	RVLEVA	KCFE	RTGI	HED	DWQ	EKVEV	IVSBVLT																																													
SlitPBP1	KCFTRVVD	CKKELNVGDH	IQDMYN	WRREDY	QILNR	MGCML	IQMA	KKLD	IM	DDQ	TMHHGK	TEDFAKSHG	ADDVAK	KLVS	VIH	ICEQ	QHTG	IA	DCM	RVLEVA	KCFE	RTKI	HED	KWAE	SMEV	IMBVM																																										
SexiPBP4	KCFTKVV	CKKELNAVEH	IQDMYN	WRCEDH	QILNR	TGCVI	IQMA	KNLD	IM	EDQ	KMHHGK	TTEFAKSHG	ADDEVA	KKLV	SIH	ICEQ	QHAG	IA	DCM	RVLEIS	KCFE	RTKI	HED	KWAE	NMEV	IMBVM																																										
CsupPBP3	ISFLKVLQ	CKKELSVPEE	VLSLMT	FWNQD	TLSHR	TGCVI	IQV	SKLD	ILE	ETYK	LHPD	NAEYV	KKHG	ADDE	TASQ	IMN	ILR	CEI	KNEA	ITD	HCD	RVREIA	KCFE	HGHM	HED	KWAE	NMEV	IINLVA																																								
OnubPBP	INFGKALD	CKKELDL	PDSINAD	FYNWR	KEGYEL	SNR	TGCAI	IQLS	SKLD	IVDPEG	KLHHGN	THEFAKKHG	ADDSMA	KQV	ELIH	CEG	SVADDPD	ACM	KVINIA	KCFE	KA	EIHK	INWAE	SMDL	IVAB	VLA																																										
SexiPBP1	KCFTKVV	CKKELNAGEH	IQDMYN	WRREDY	QILNR	TGCVI	IQMA	KKLD	IM	EDQ	KMHHGK	TTEFAKSHG	ADDEVA	KKLV	SIH	ICEQ	QHAG	IA	DCM	RVLEIS	KCFE	RTKI	HED	KWAE	NMEV	IMBVM																																										
CsupPBP2	LNFGKAYD	CKKELLELP	NEVD	TDFN	WKDDY	AVTNR	TGCAI	IQMS	KNLD	ILD	PDG	KMHHGNAA	EFAKKHG	ADDSMA	QQL	IDILH	CEK	GAS	PGD	GLA	CV	VLEIS	KCFE	KVEI	HK	INWAE	SMDL	IMAB	VLA																																							
HvirPBP2	SGFTKVV	CKKELNVGDH	IQDMYN	WRREEY	QIVNR	TGCVI	IQMA	AKLDI	VGDDQ	KMHHGKA	EEFAKSHG	ADDALAKQ	LVGLIH	CE	TQHQ	AE	DCSR	RTLEVA	KCFE	RTKI	HED	KWAE	SMEV	IMBIMT																																												
LstiPBP	INFGKALD	CKKELDL	PDSINAD	FYNWR	KEGYEL	SNR	TGCAI	IQLS	SKLD	IVDPEG	KLHHGN	THEFAKKHG	ADDA	MAKQ	VDLIH	CE	S	V	P	D	P	D	E	C	L	K	V	I	N	A	E	S	M	D	L	I	M	A	B	V	L	A																										
BmorPBP	IQFAKPLE	CKKEMGLT	TETVLK	DYFN	WIEDY	EFTDR	TGCAI	IQMS	KKLD	IM	DGDY	NLHHGKA	EEFAKKHG	ADDE	MAKQ	VDLIH	CE	S	V	A	T	M	P	D	E	C	E	R	L	K	V	A	K	C	F	E	A	E	I	H	K	I	N	W	A	E	D	V	E	L	L	M	A	B	V	L	N											
CpunPBP	ATDFKLE	CKKELSV	TDDI	QGLV	RWLED	SALGER	TGCVI	IQMA	EKQD	IVV	TEDY	RMH	ENAY	NEFAK	NHG	ADDA	MA	T	A	I	V	K	V	I	H	CE	Q	F	T	S	N	P	D	E	C	A	R	V	M	E	V	S	K	C	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
      TDDMIQGLVRFWLEDSALGERELGCVIICMAEKQDLVVTEDYRMHHENAY
helix  HHHH      HHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
sheet          EEEEE
turns
coil          CC


      .  110   .  120   .  130   .  140   .  150
      NFAKNHGADDAMATAIVKVIHTCEEQFTSNPDHCCARVMEVSKCFRDEIHR
helix  HHHH      HHHHHHHHHHHHHH      HH      HHHHHHHHHHHHHHHHHHH
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.com


Modelling Tools Repository Documentation

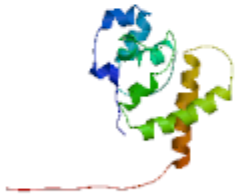
[ myWorkspace ] [ Settings ] [ logout ]

Workunit: P000002 - Overview



Print/Save this page as 

Model Summary 



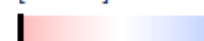
#### Model information:

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.

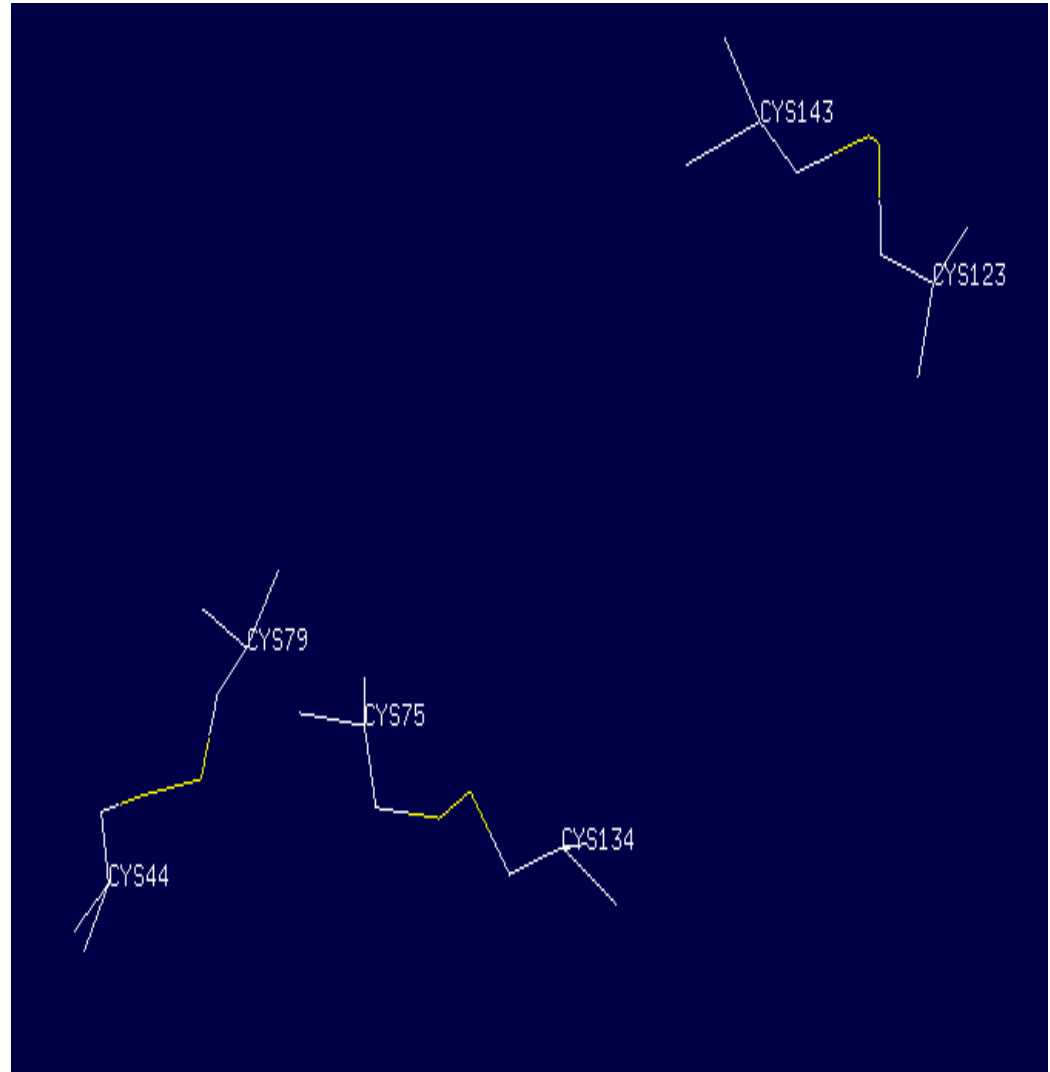
**Warning: Low QMEAN Z-scores!** Only membrane proteins or models of poor quality are expected to reach such low scores.

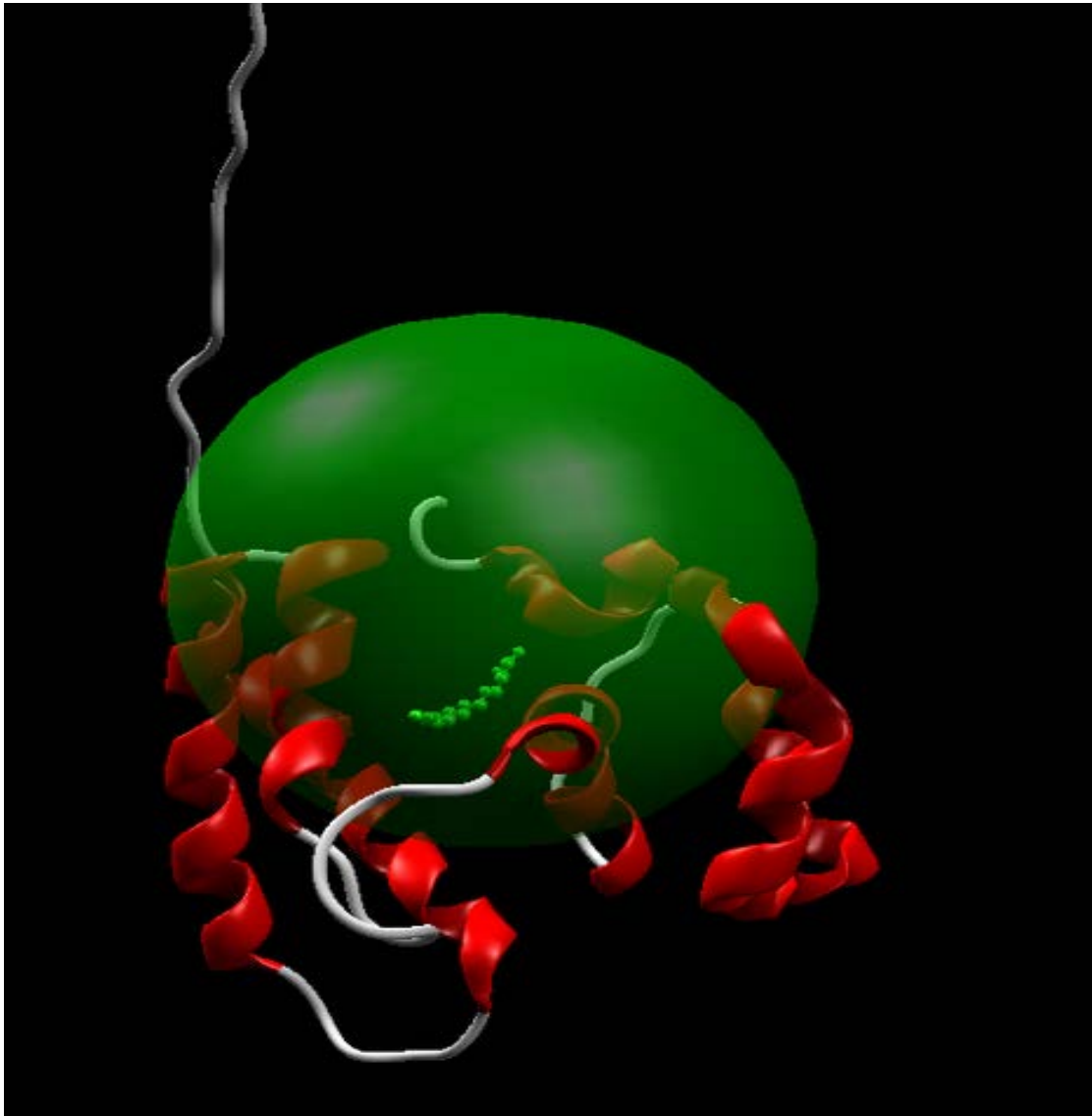
logs: [Templates] [Alignment] [Modeling]

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!**

