

# 牛整合蛋白 $\alpha v \beta 3$ 的生物信息学分析



G09组

组员：蔡虎 宋江伟

张晓 赵帅阳

2012.6.10

- 研究背景
- 牛整联蛋白  $\alpha v \beta 3$  系统分析

# 背景介绍

- 口蹄疫是由口蹄疫病毒感染引起的偶蹄动物共患的急性、热性、接触性传染病。
- 特点：
- 宿主范围广：包括猪牛羊在内所有偶蹄动物都易感；
- 发病急：潜伏期1-14天，快者十几个小时发病；
- 传播方式多样：除接触感染外，还可通过空气源距离传播；
- 感染途径多：依次为呼吸道、消化道、伤口、人工受精、胚胎移植、注射治疗等；
- 病毒变异性强；
- 病毒致病性强：十几个TCID<sub>50</sub>就可以引起发病；
- 病毒免疫原性弱，可以形成持续性感染。

- 牛  $\alpha v \beta 3$  整联蛋白
- 现已确定有两类受体与口蹄疫病毒的感染有关，即硫酸乙酰肝素和整联蛋白。硫酸乙酰肝素作为细胞受体只局限于一些毒株，在介导野毒株进入培养细胞时的作用目前还未得到证实，而整联蛋白作为启动感染的受体已被许多研究所证实。
- 整联蛋白是细胞表面的一种膜蛋白，由  $\alpha$  亚基和  $\beta$  亚基以非共价形式结合而成的异二聚体，其中整联蛋白  $\alpha v \beta 3$  在牛猪等易感动物的上皮细胞，为口蹄疫病毒侵入的主要受体，也是第一个被发现的口蹄疫病毒的整联蛋白受体。目前未有牛整联蛋白三维结构模型建立。

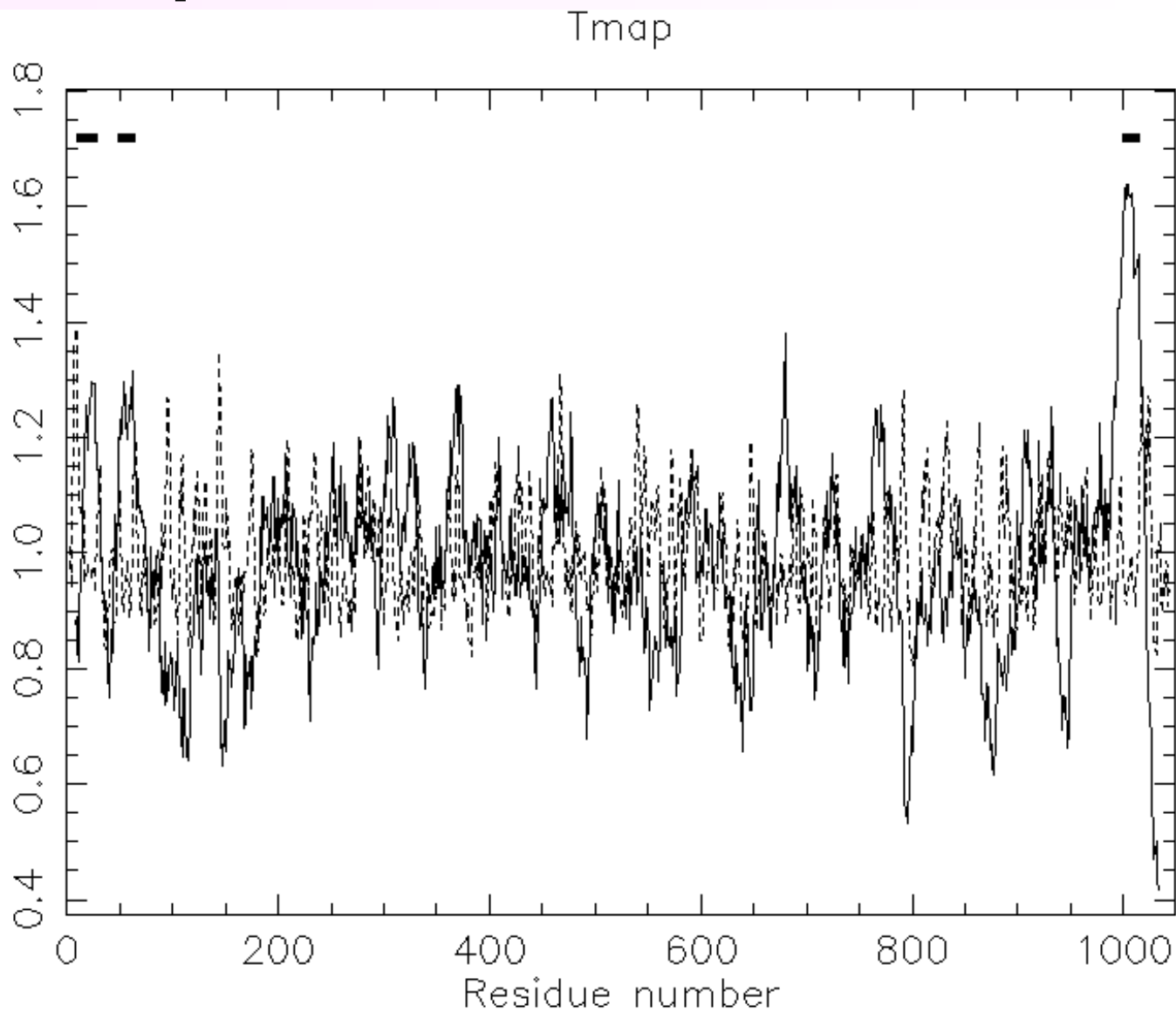
# 牛整联蛋白 $\alpha v$ 序列分析

- 由Uniprot数据库获得整联蛋白  $\alpha v$  序列:

- > BOVIN alpha-V

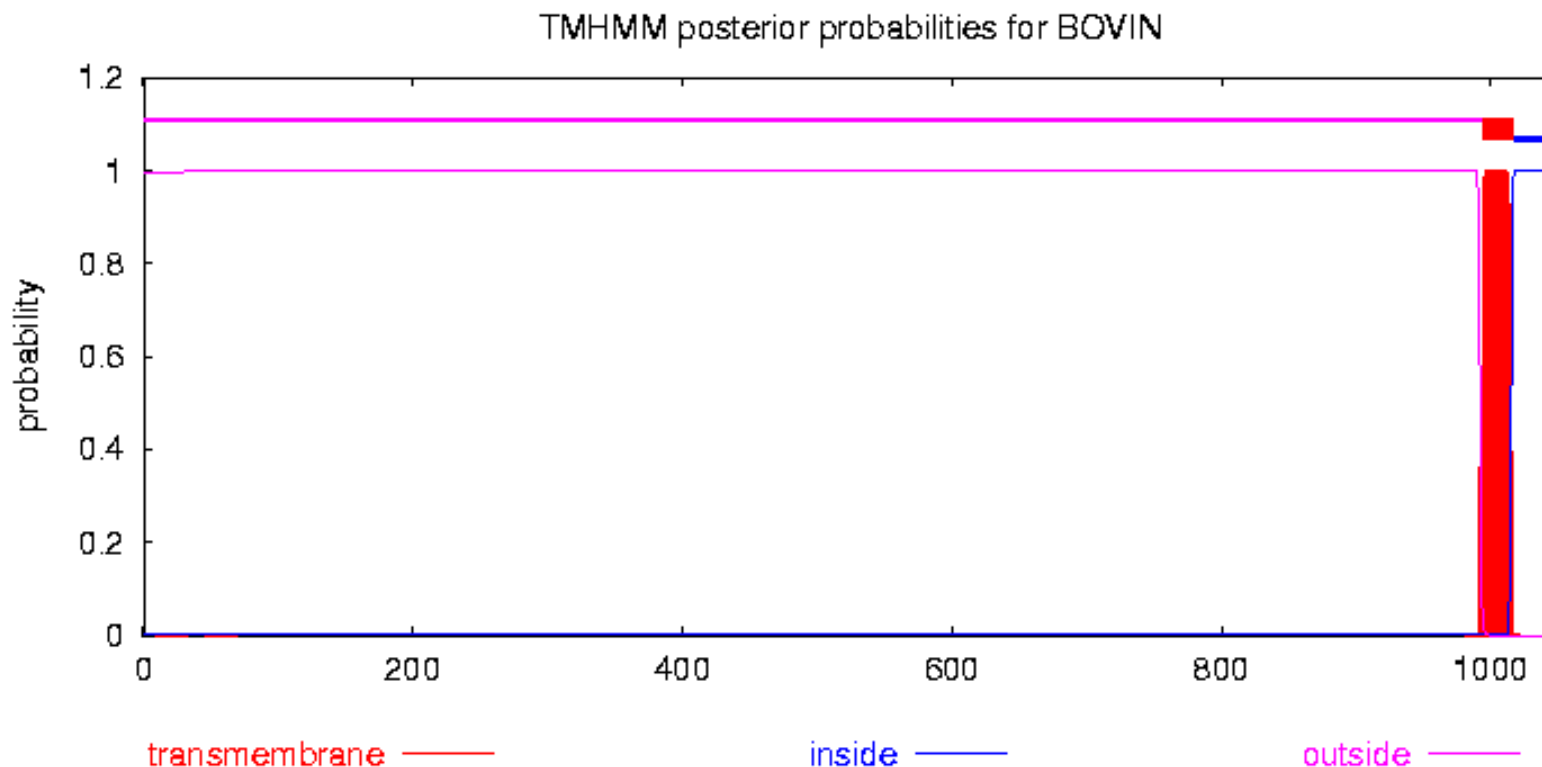
MAFPPRRRLRLGPRGLPLLLSGLLLPLCRAFNLDVESPAEYSGPEGSYFGFAVDFFVPSASSRMF  
LLVGAPKANTTQPGIVEGGQVLKCDWSSHRRRCQP|EFDATGNRDYAKDDPLEFKSHQWFGASVRS  
KQDK|LACAPLYHWRTEMKQEREPVGT CFLQDGTKTVEYAPCRSKN|DADGQGF CQGGFS|DFTK  
ADRVLLGGPGSFYWQQQL|SDQVAE|VSKYDPKVYS|KYNNQLATRQAQ|FDDSYLGYSVAVGD  
FNGDG|DDFVSGVPRAARTLGMVY|YDGKNMSSLHNFTGEQMAAYFGFSVAATD|NGDDYADVFI  
GAPLFMDRGS DGKLQEVGQVSVSLQKASGDFQT|KLNQFEV FARFGSA|APLGDLDQDGFND|A|  
AAPYGGEDKKG|VY|FNQRPTGLNAVPSQ|LEGKWAARSMPPSFGYSMKGATD|DKNGYPDL|VG  
AFGVDRAVLYRARPV|TVNAGLEVYPS|LNQENKTCPLPGTDLKVSCFNVRFC LKADGKGALPTK  
LDFQVELLLDKLKQKGA|RRALFLHNRSPGH SKNMT|SRGGQM QCEEL|AYLRDESEFRDKLTP|  
T|FMEYWLDYRTAADATGLQP|LNQFTPANVSRQAH|LLDCGEDNVCKPKLEVSVDSQKK|Y|G  
DDNPLTL|VKAQNQGE GAYEAE|VS|PLQADF|GVVRNSEALARLSCAFKTENQTRQVVCDLGN  
PMKAGTQLLAGLRFSVHQQSEMDTSVKFDLQ|QSSNLFDKVSPVVS YKVDLAVLAAVE|IRGVSSP  
DH|FLP|PNWKYKENPETEEDVGPVVQH|YELRNNGPSSFSKAMLHLQWPYKYNNNTLLY|LQYD  
IDGPMNCTSDME|NPLR|K|SNSQTSEKNDTVGGQGDRNHL|TKRDLT LNEG DVHTLGCG|AECL  
K|VCQVGRLDRGKSA|LYVRSLLWTETFMNKENQNHSYSLKSSASFNV|IEFPYKNLP|IED|FNST  
LVTTNVTWG|QPAPMPVPVWV|LAVLAGLLLLAVLVFVMYRMGFFKRVRPPQEEQEREQLQPHE  
NGEGNSET

# 利用Tmap分析跨膜区

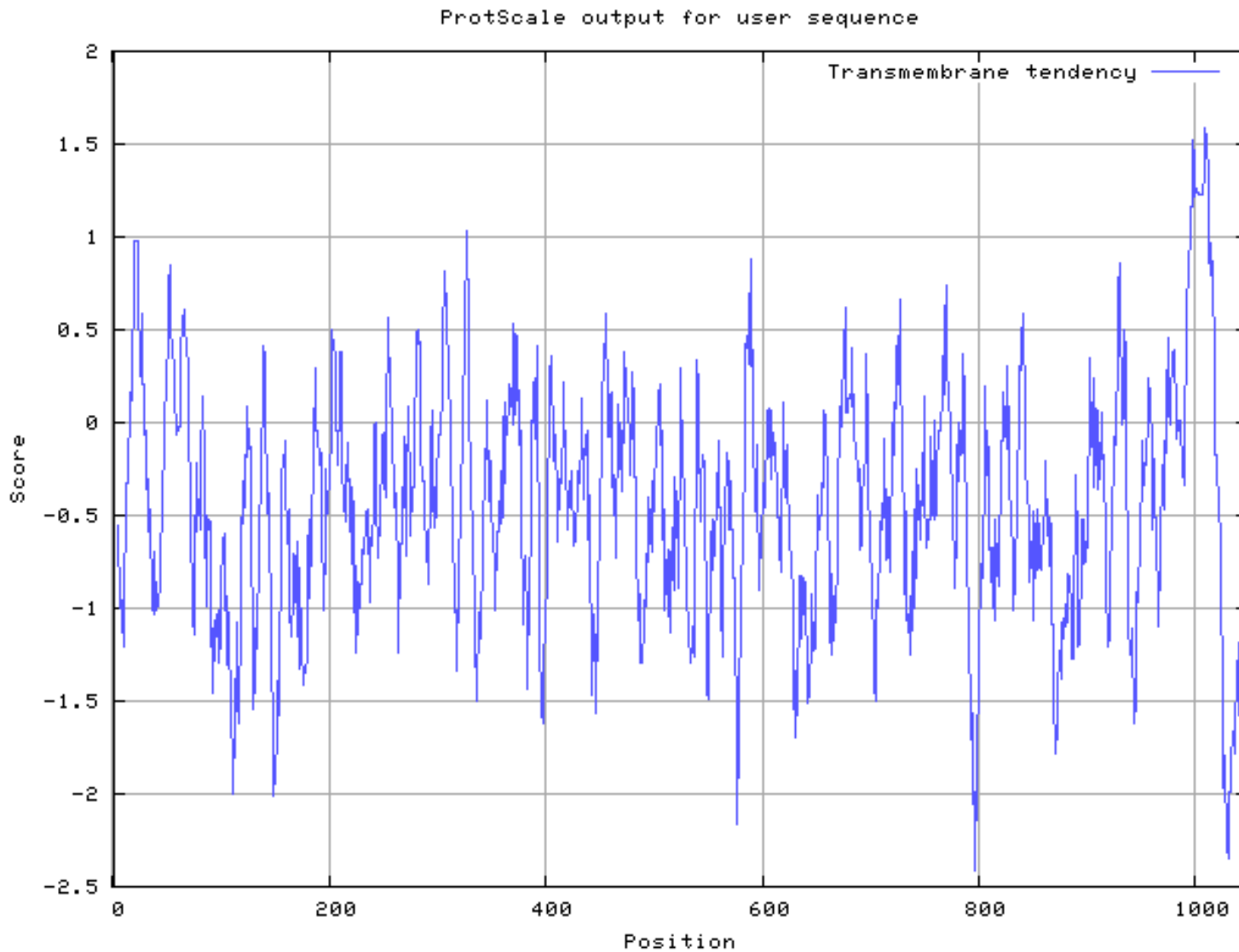


# 利用TMHMM分析跨膜区

```
# BOVIN Length: 1048
# BOVIN Number of predicted TMHs: 1
# BOVIN Exp number of AAs in TMHs: 22.87702
# BOVIN Exp number, first 60 AAs: 0.0524
# BOVIN Total prob of N-in: 0.00255
BOVIN  TMHMM2.0      outside      1   993
BOVIN  TMHMM2.0      TMhelix      994 1016
BOVIN  TMHMM2.0      inside     1017 1048
```



跨膜序列为994-1016为VWVILAVLAGLLLLAVLVFVMY



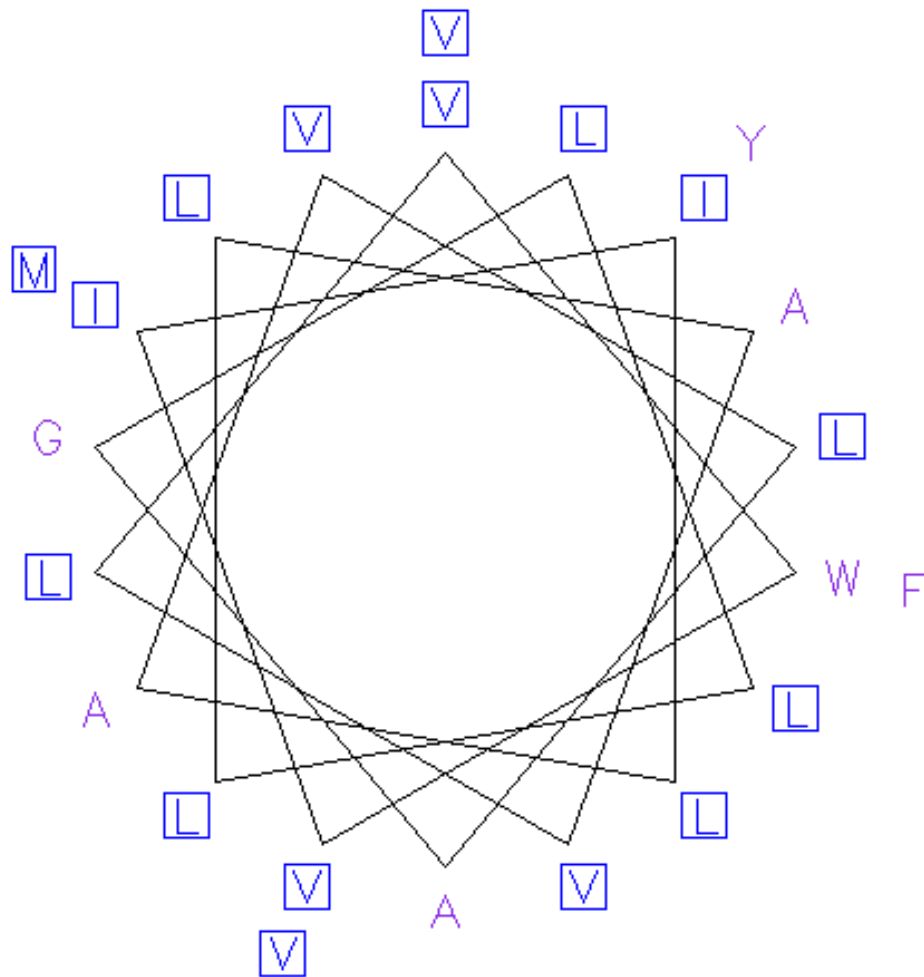
得知有一个跨膜序列在1000左右。跨膜序列VWVILAVLAGLLLLAVLVFVMY



# 用ExPasy分析跨膜区994-1016为 VWVIILAVLAGLLLLAVLVFVMY

Helical wheel of raw::543469

Fri 8 Jun 2012 12:19:32

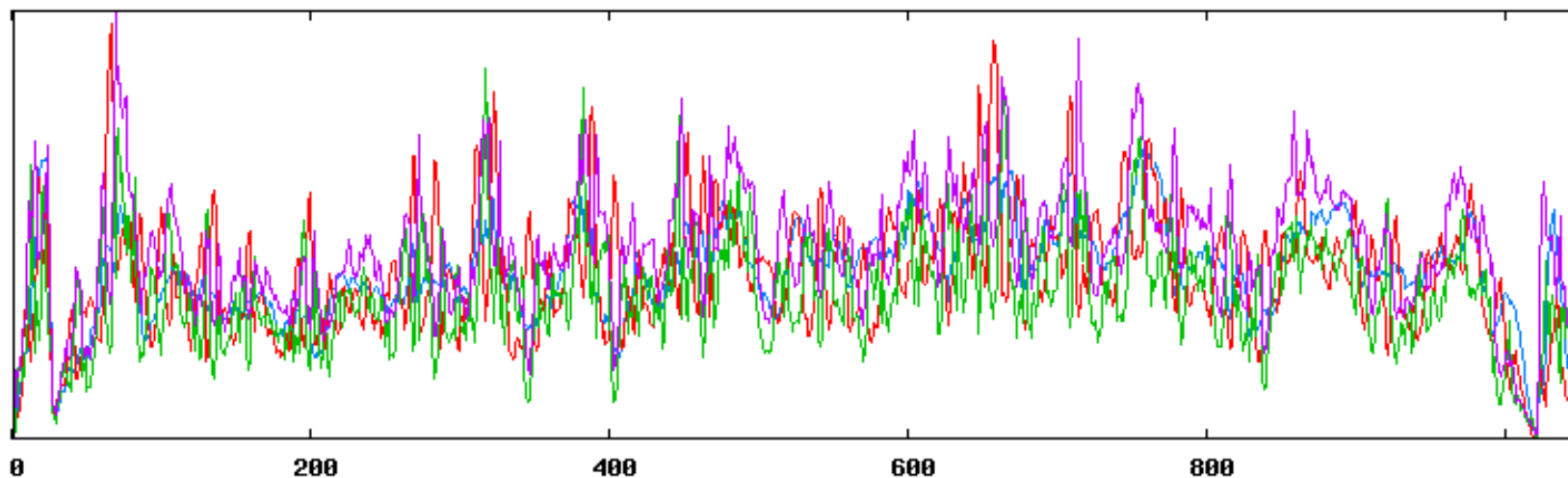
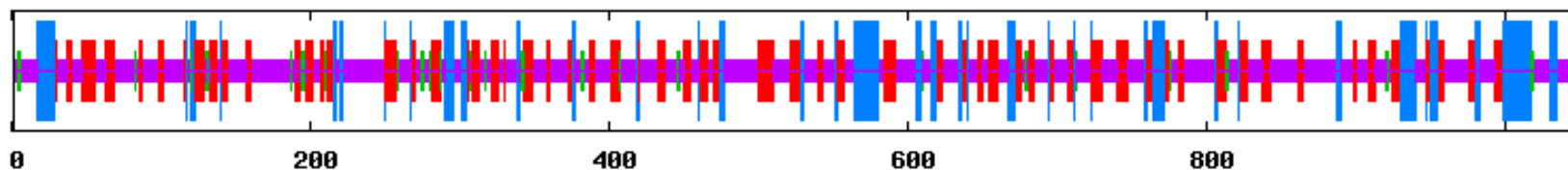


疏水性越强越有可能是跨膜区。氨基酸V、L都是疏水性极强的氨基酸。所以这段为跨膜区。

# SOPMA软件分析二级结构

SOPMA :

Alpha helix	(Hh)	:	150	is	14.31%
$3_{10}$ helix	(Gg)	:	0	is	0.00%
Pi helix	(Ii)	:	0	is	0.00%
Beta bridge	(Bb)	:	0	is	0.00%
Extended strand	(Ee)	:	287	is	27.39%
Beta turn	(Tt)	:	41	is	3.91%
Bend region	(Ss)	:	0	is	0.00%
Random coil	(Cc)	:	570	is	54.39%
Ambiguous states (?)		:	0	is	0.00%
Other states		:	0	is	0.00%



# 使用Blast分析

Entry	Entry name	Status	Protein names	Organism	Length	Identity
P80746	ITAV_BOVIN	★	Integrin alpha-V	Bos taurus (Bovine)	1,048	100.0%
B0FYY5	B0FYY5_SHEEP	★	Integrin subunit alphaV	Ovis aries (Sheep)	1,048	98.0%
D2SYX8	D2SYX8_CAMDR	★	Integrin alpha V	Camelus dromedarius (Dromedary) (Arabian camel)	1,048	96.0%
A4USB0	A4USB0_PIG	★	Integrin alpha V	Sus scrofa (Pig)	1,046	96.0%
P06756	ITAV_HUMAN	★	Integrin alpha-V	Homo sapiens (Human)	1,048	95.0%
B0FYY6	B0FYY6_CAMBA	★	Integrin subunit alphaV	Camelus bactrianus (Bactrian camel)	1,054	95.0%
F6W3W7	F6W3W7_HORSE	★	Uncharacterized protein	Equus caballus (Horse)	1,018	97.0%
G1TE34	G1TE34_RABIT	★	Uncharacterized protein	Oryctolagus cuniculus (Rabbit)	1,047	94.0%
G3SU46	G3SU46_LOXAF	★	Uncharacterized protein	Loxodonta africana (African elephant)	1,045	94.0%
F7IJZ6	F7IJZ6_CALJA	★	Uncharacterized protein	Callithrix jacchus (White-tufted-ear marmoset)	1,047	93.0%
P43406	ITAV_MOUSE	★	Integrin alpha-V	Mus musculus (Mouse)	1,044	92.0%

# 使用needle对牛整联蛋白 $\alpha v$ 序列与人的整联蛋白 $\alpha v$ 序列进行比对

## Pairwise Alignment Result

LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
1048	5295.0	1005/1048 (95.9%)	1031/1048 (98.4%)	0/1048 ( 0.0%)



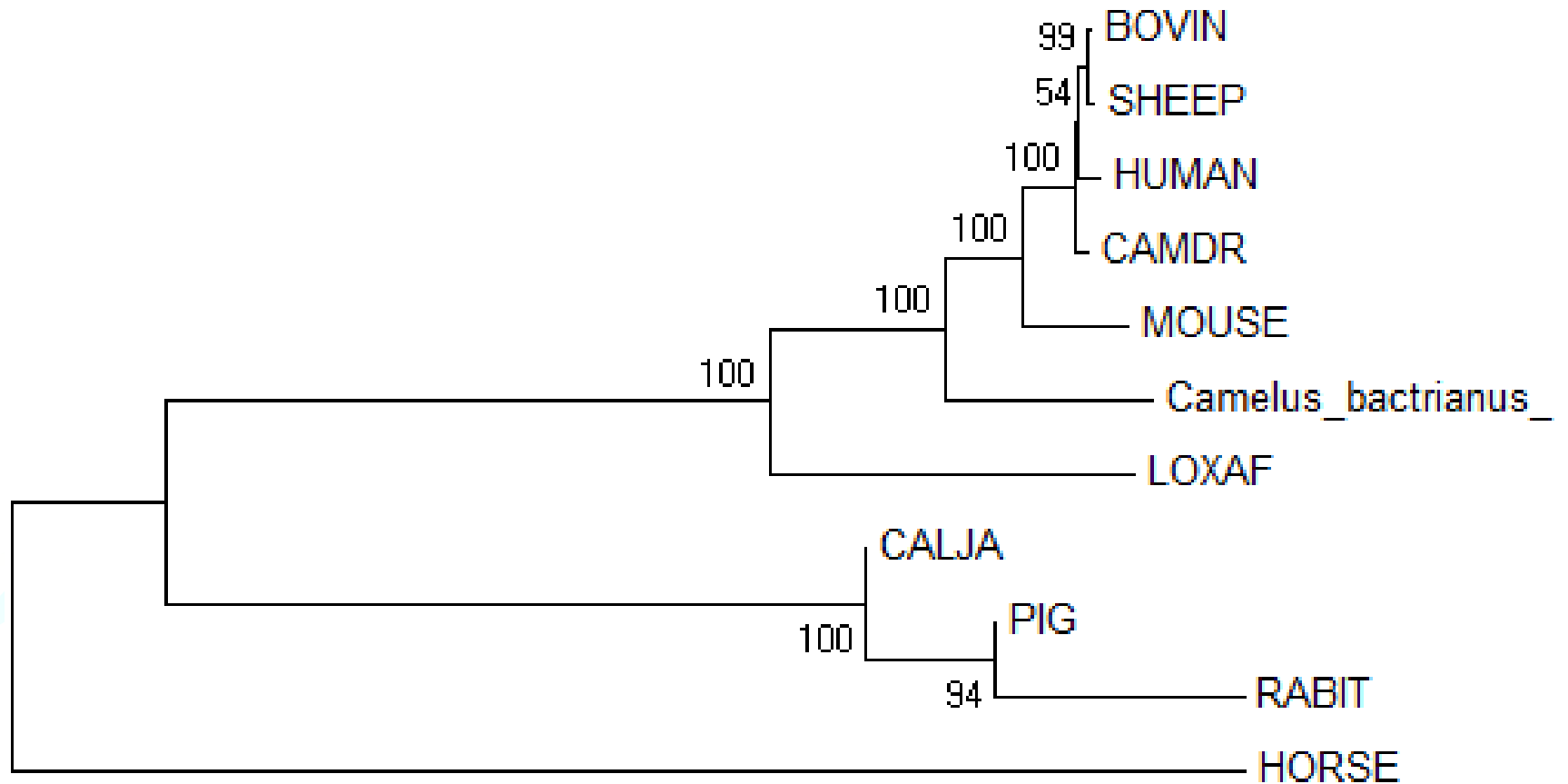






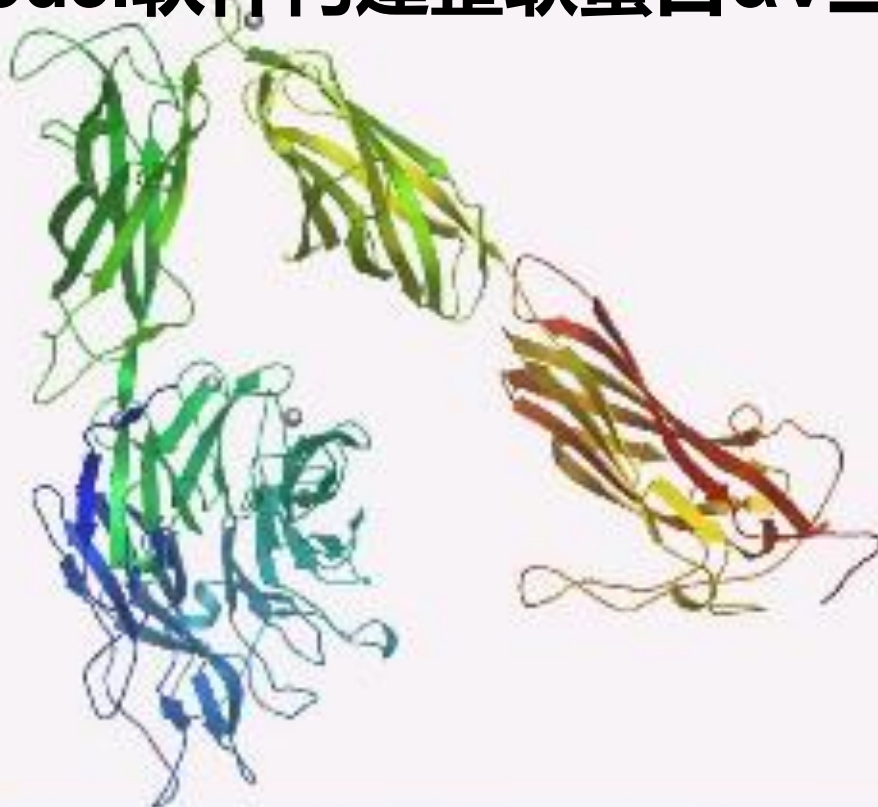


# 用MEGA分析与系统发育树建立



0.2

# 使用Swiss Model软件构建整联蛋白 $\alpha v$ 三维结构



## Model information:

Modelled residue range: 31 to 997  
Based on template: [3ijeA] (2.90 Å)  
Sequence Identity [%]: 93.18  
Evalue: 0.00e-1

## Quality information:

QMEAN Z-Score: -2.49

[details] ▶



## Quaternary structure information: [details] ▶

Template (3ije): HETERO DIMER  
Model built :SINGLE CHAIN

## Ligand information: [details] ▶

Ligands in the template: CA: 5, NAG: 11.  
Ligands in the model: CA: 5

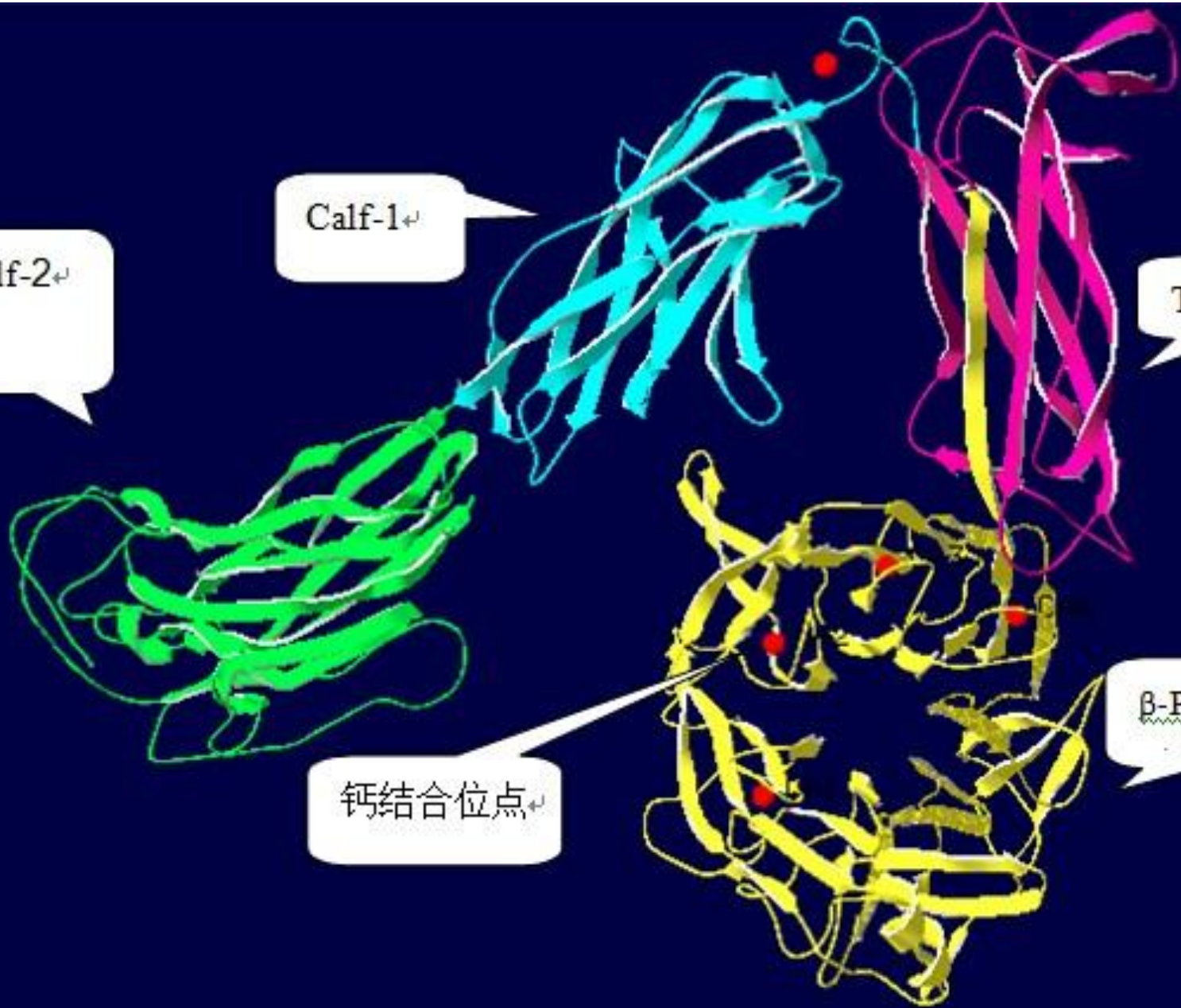
Calf-2<sup>+</sup>

Calf-1<sup>+</sup>

Thigh<sup>+</sup>

β-Propelle

钙结合位点<sup>+</sup>



**Model information:**  
Modelled residue range:  
Based on template:  
Sequence Identity [%]:  
Evalue:

988 to 1034  
[2kncA] (99.9 Å)  
48.94  
3.00e-18

**Quality information:**  
QMEAN Z-Score: -4.09

[details] ▶



**Quaternary structure information:** [details] ▶  
Template (2knc): HETERO DIMER  
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.

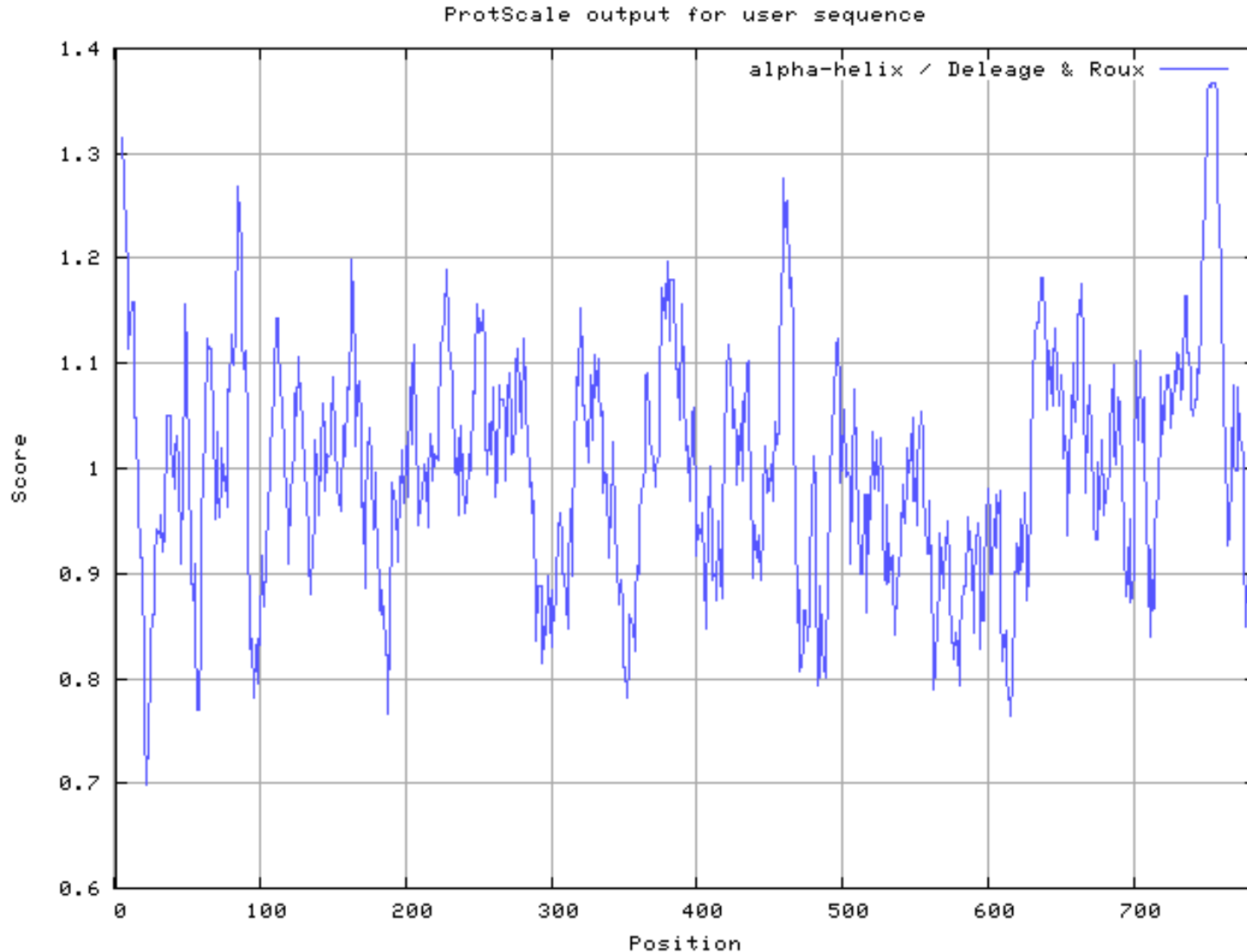
# 牛整联蛋白 $\beta$ 3序列分析

- 从Uniprot获得牛整联蛋白  $\beta$  3序列:

- >BOVIN Integrin beta

MRARRLWAAVLVVGALAGVGVGGPNICATRGVSSCQQCLAVSPTCAWCSDEALPP  
GSPRCNLKENLLKDNCHPESIEFPISEARILEARPLSDKGTGDSSQITQVSPQRI  
ALRLRPDDSKIFSVQVRQVEDYPVDIYYLMDLSYSMKDDLRLNIQNLGTKLASQMR  
KLTSNLRIGFGAFVDKPVSPYMI SPPEAIRNPCYDMKATCLPMFGYKHVLTLD  
QVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRNDASHLLVFTTDAK  
THIALDGRLAGIVQPNDGQCHVGSNDHYSASTTMDYPSLGLMTEKLSQKNINLIF  
AVTESVVNLYQNYSELIPGTTVGVLSTDSSNVLQLIVDAYGKIRSKVELEVRDLP  
EELSLSFNATCLNNEVIPGLKSCVGLKIGDTSVFSIEAKVRGCPQKKEKSFTIKP  
VGFKDSLTVQVTFDCDCACQAEAEPFSHRCNNGNGTFECGVCRCGPGWLGSQCEC  
SEEDYRPSQQDECSREGQPICSORGECLCGQCVCHSSDFGKITGKYCECDDFSC  
VRYKGEMCSGHGQCSCGDCLCSDWTGYCNCCTTRTDTCMSSNGLLCSGRGKCEC  
GSCVCIQPGSYGDTCEKCPTCPDACTFKKECVECKKFNRGALNEENTCNRYCRDE  
IEPVKELKDTGKDAVNCTYKNEDDCVVRFQYYEDSSGKSILYVVEEPECPKGPDI  
LVVLLSVMGAILLIGLATLLIWKLLITIHDRKEFAKFEERARAKWDTANNPLYK  
EATSTFTNITYRGT

# 使用Expasy分析序列的螺旋和折叠



Using the scale **alpha-helix / Deleage & Roux**, the individual values for the 20 amino acids are:

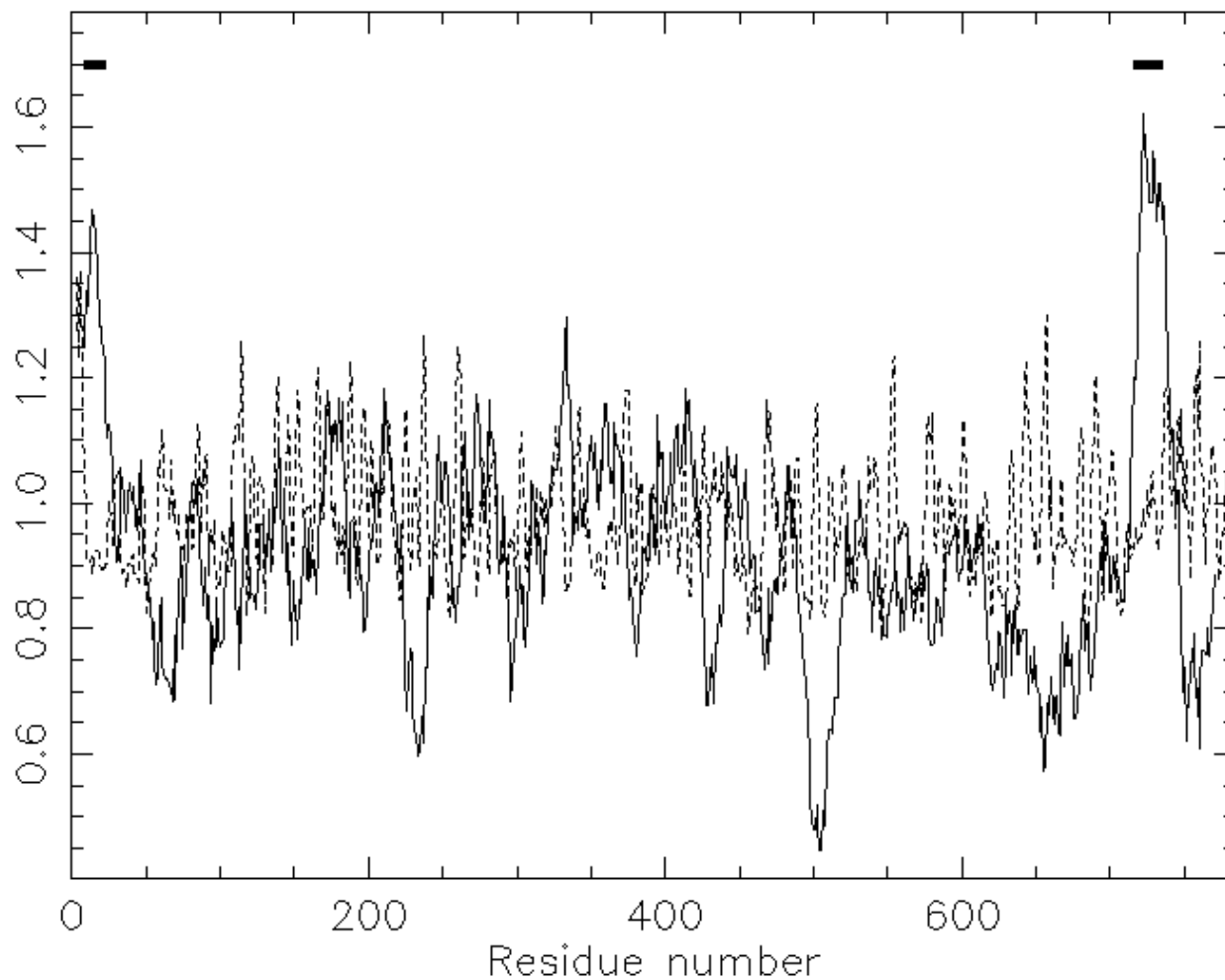
Ala:	1.489	Arg:	1.224	Asn:	0.772	Asp:	0.924	Cys:	0.966	Gln:	1.164
Glu:	1.504	Gly:	0.510	His:	1.003	Ile:	1.003	Leu:	1.236	Lys:	1.172
Met:	1.363	Phe:	1.195	Pro:	0.492	Ser:	0.739	Thr:	0.785	Trp:	1.090
Tyr:	0.787	Val:	0.990	:	0.848	:	1.334	:	1.020		

Weights for window positions 1,...,9, using **linear weight variation model**:

1	2	3	4	5	6	7	8	9
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
edge				center				edge

# 利用Tmap分析跨膜区

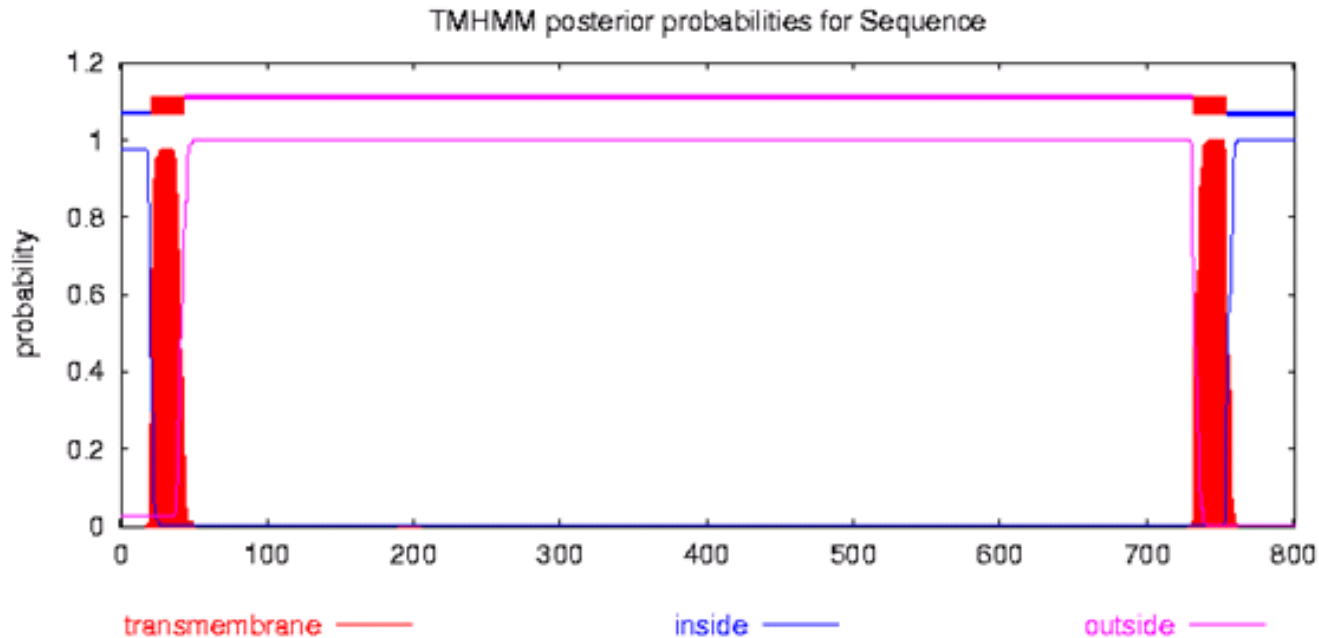
Tmap





# 利用TMHMM分析跨膜区

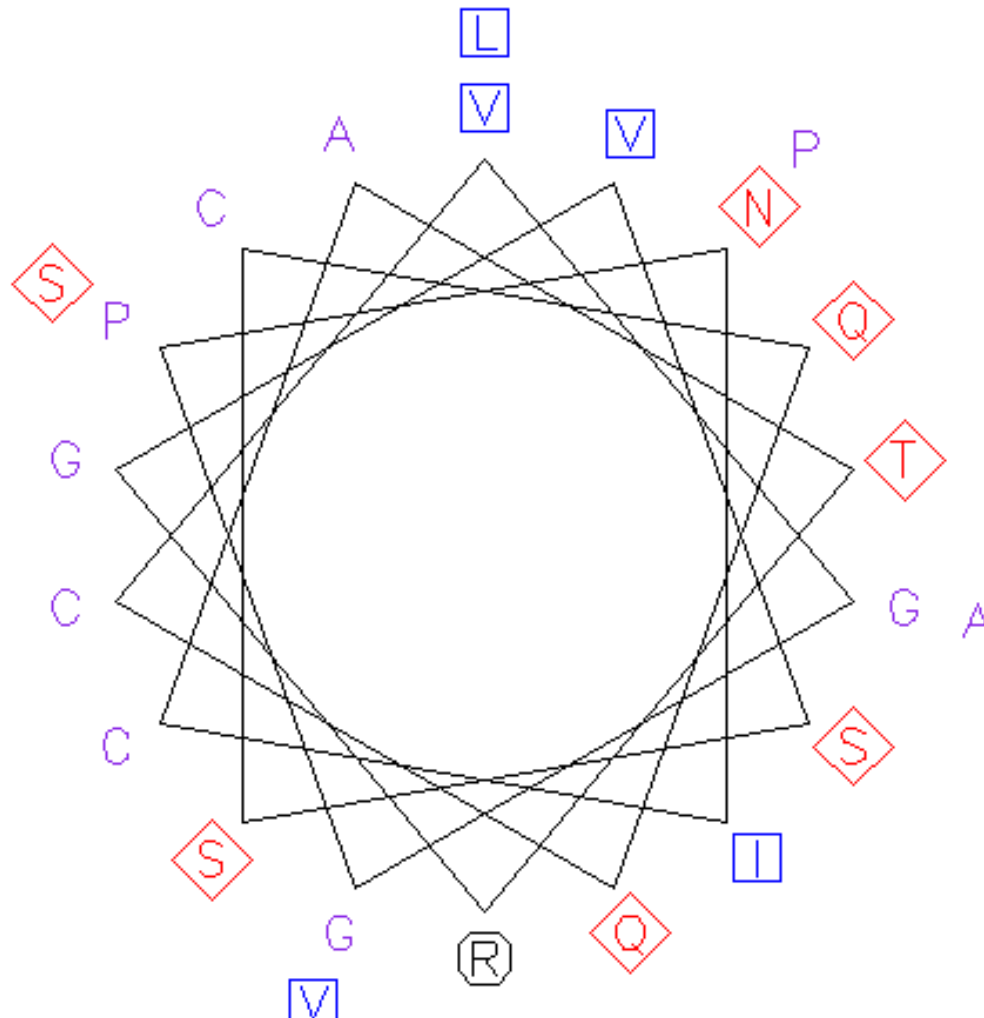
```
# Sequence Length: 801
# Sequence Number of predicted TMHs: 2
# Sequence Exp number of AAs in TMHs: 43.14956
# Sequence Exp number, first 60 AAs: 20.36766
# Sequence Total prob of N-in: 0.97387
# Sequence POSSIBLE N-term signal sequence
Sequence TMHMM2.0 inside 1 20
Sequence TMHMM2.0 TMhelix 21 43
Sequence TMHMM2.0 outside 44 731
Sequence TMHMM2.0 TMhelix 732 754
Sequence TMHMM2.0 inside 755 801
```



# 用ExPasy分析跨膜区第一个21-43位 VGGPNICATRGVSSCQQCLAVSP

Helical wheel of raw::543594

Fri 8 Jun 2012 15:38:04



疏水性不是很强



# 提取牛整联蛋白 $\beta$ 3序列用Blast做同源性比对分析

Alignments	Entry	Entry name	Status	Protein names	Organism	Length	Identity	Score	E-value	Gene names
	<a href="#">F1MTN1</a>	F1MTN1_BOVIN	★	Integrin beta	Bos taurus (Bovine)	784	100.0%	4,257	0.0	INTB3
	<a href="#">D6BRG0</a>	D6BRG0_SHEEP	★	Integrin beta	Ovis aries (Sheep)	762	99.0%	4,116	0.0	
	<a href="#">P05106</a>	ITB3_HUMAN	★	Integrin beta-3	Homo sapiens (Human)	788	95.0%	4,085	0.0	ITGB3 GP3A
	<a href="#">G3QXN4</a>	G3QXN4_GORGO	★	Integrin beta	Gorilla gorilla gorilla (Lowland gorilla)	788	95.0%	4,083	0.0	ENSG00000056345
	<a href="#">G1QXM3</a>	G1QXM3_NOMLE	★	Integrin beta	Nomascus leucogenys (Northern white-cheeked gibbon) (Hylobates leucogenys)	788	95.0%	4,076	0.0	ENSG00000056345
	<a href="#">H2NVP1</a>	H2NVP1_PONAB	★	Integrin beta	Pongo abelii (Sumatran orangutan)	788	95.0%	4,075	0.0	ITGB3
	<a href="#">Q52MQ6</a>	Q52MQ6_HORSE	★	Integrin beta	Equus caballus (Horse)	784	95.0%	4,071	0.0	
	<a href="#">G7PV32</a>	G7PV32_MACFA	★	Integrin beta	Macaca fascicularis (Crab-eating macaque) (Cynomolgus monkey)	788	95.0%	4,069	0.0	EGM_08034
	<a href="#">G7NJ66</a>	G7NJ66_MACMU	★	Integrin beta	Macaca mulatta (Rhesus macaque)	788	95.0%	4,069	0.0	EGK_08849
	<a href="#">G1Q3E4</a>	G1Q3E4_MYOLU	★	Integrin beta	Myotis lucifugus (Little brown bat)	784	94.0%	4,065	0.0	
	<a href="#">O97702</a>	O97702_CANFA	★	Integrin beta	Canis familiaris (Dog) (Canis lupus familiaris)	784	94.0%	4,063	0.0	GPIIIa Cfa.3641

# 使用needle对牛整联蛋白 $\beta$ 3序列与人的整联蛋白 $\beta$ 3序列进行比对

## Pairwise Alignment Result

LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
788	3805.0	696/788 (88.3%)	713/788 (90.5%)	64/788 ( 8.1%)

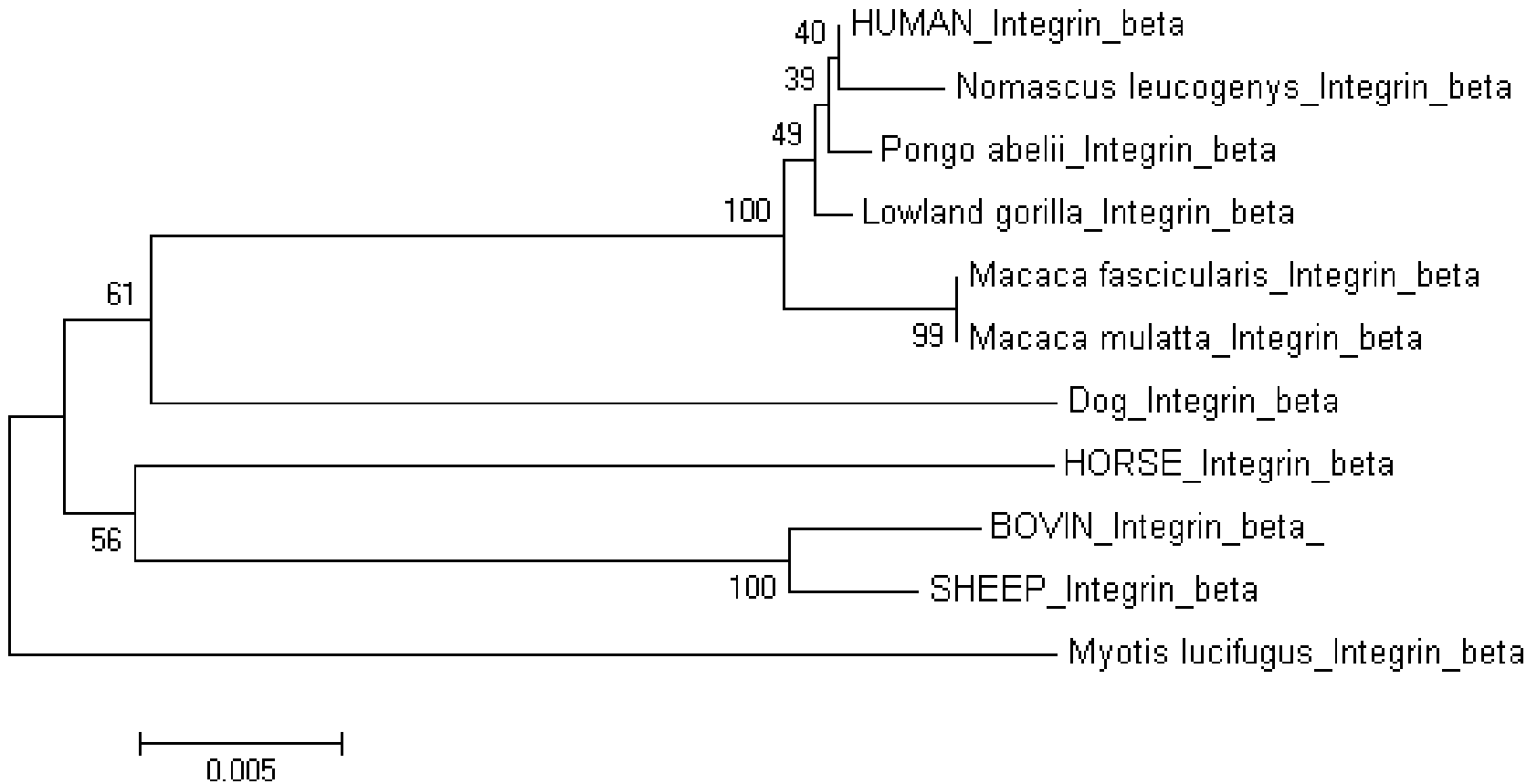




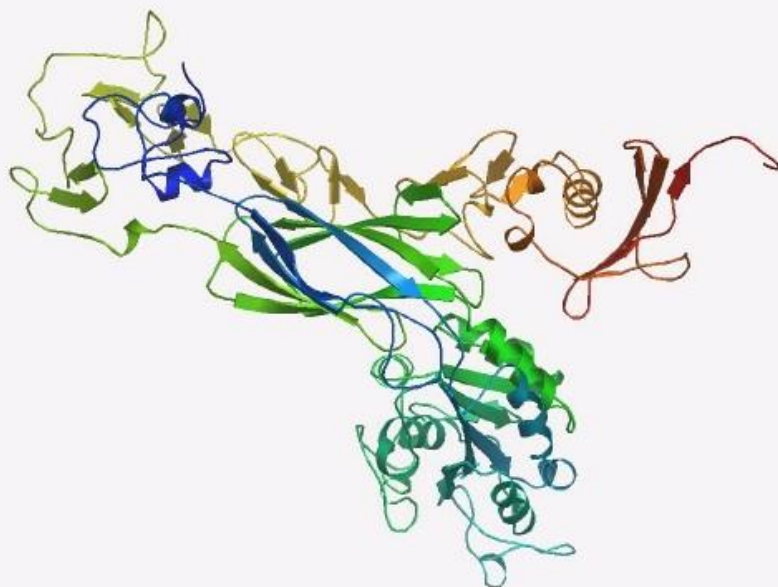




# MEGA分析并建立系统发育树



# 使用Swiss Model软件构建整联蛋白 $\beta 3$ 三维结构



## Model information:

Modelled residue range: 23 to 717  
Based on template: [3ijeB] (2.90 Å)  
Sequence Identity [%]: 95.68  
Evalue: 0.00e-1

## Quality information:

QMEAN Z-Score: -1.83

[details] ▶



## Quaternary structure information:

Template (3ije): HETERO DIMER  
Model built :SINGLE CHAIN

[details] ▶

## Ligand information:

Ligands in the template: CA: 1, NAG: 5.  
Ligands in the model: none.

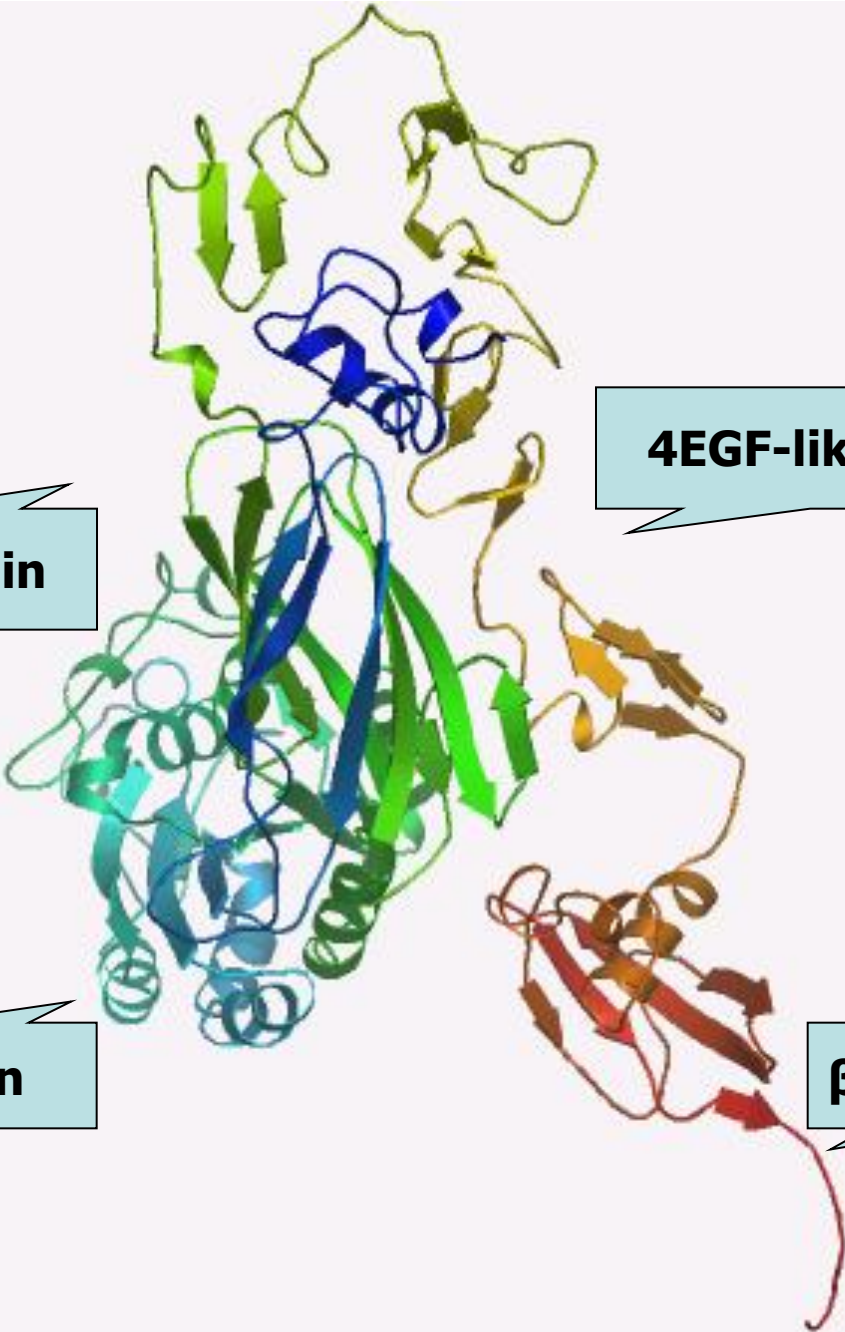
[details] ▶

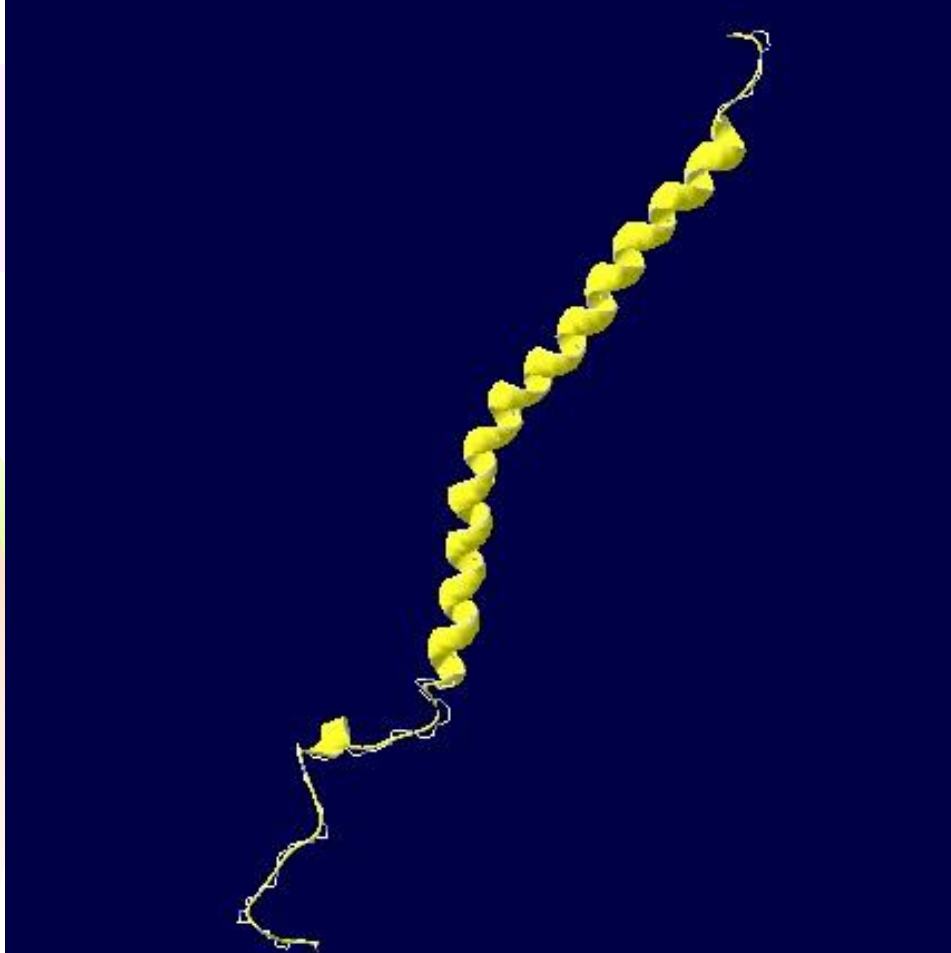
**Hybrid domain**

**4EGF-like domain**

**$\beta$ A-domain**

**$\beta$ -tail domain**





### Model information:

Modelled residue range: 711 to 784  
Based on template: [2kncB] (99.9 Å)  
Sequence Identity [%]: 98.65  
Evalue: 3.39e-29

### Quality information:

QMEAN Z-Score: -3.7

[details] ▶



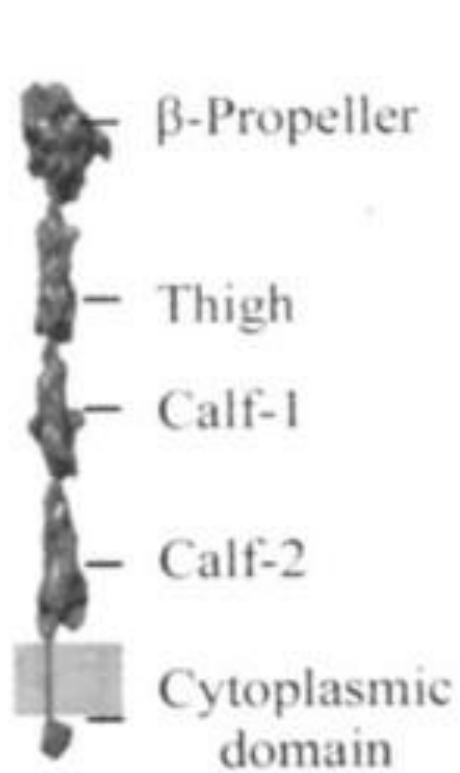
### Quaternary structure information: [details] ▶

Template (2knc): HETERO DIMER  
Model built :SINGLE CHAIN

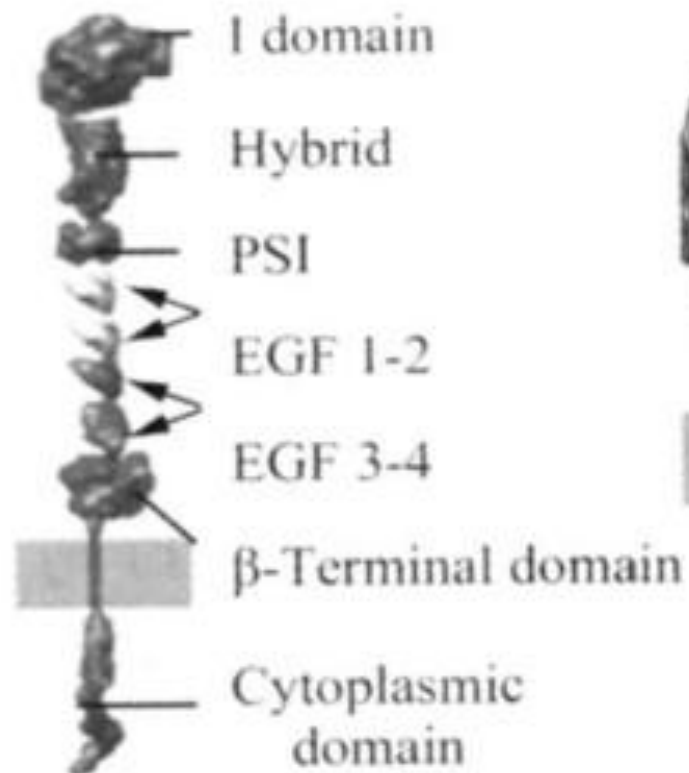
### Ligand information:

[details] ▶

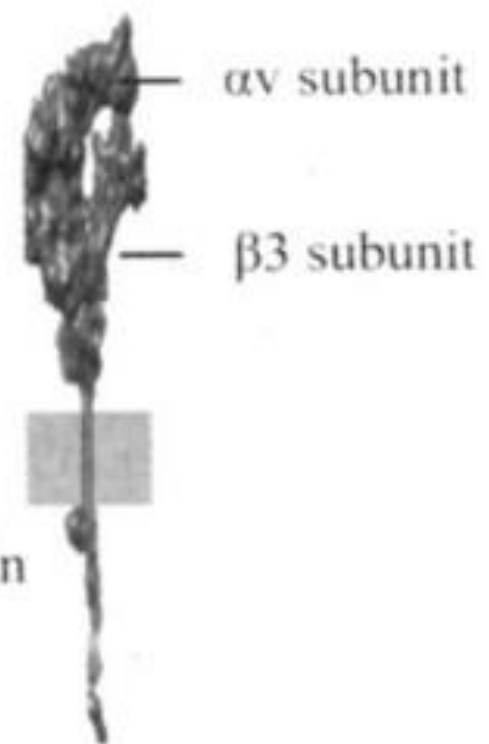
Ligands in the template: none.  
Ligands in the model: none.



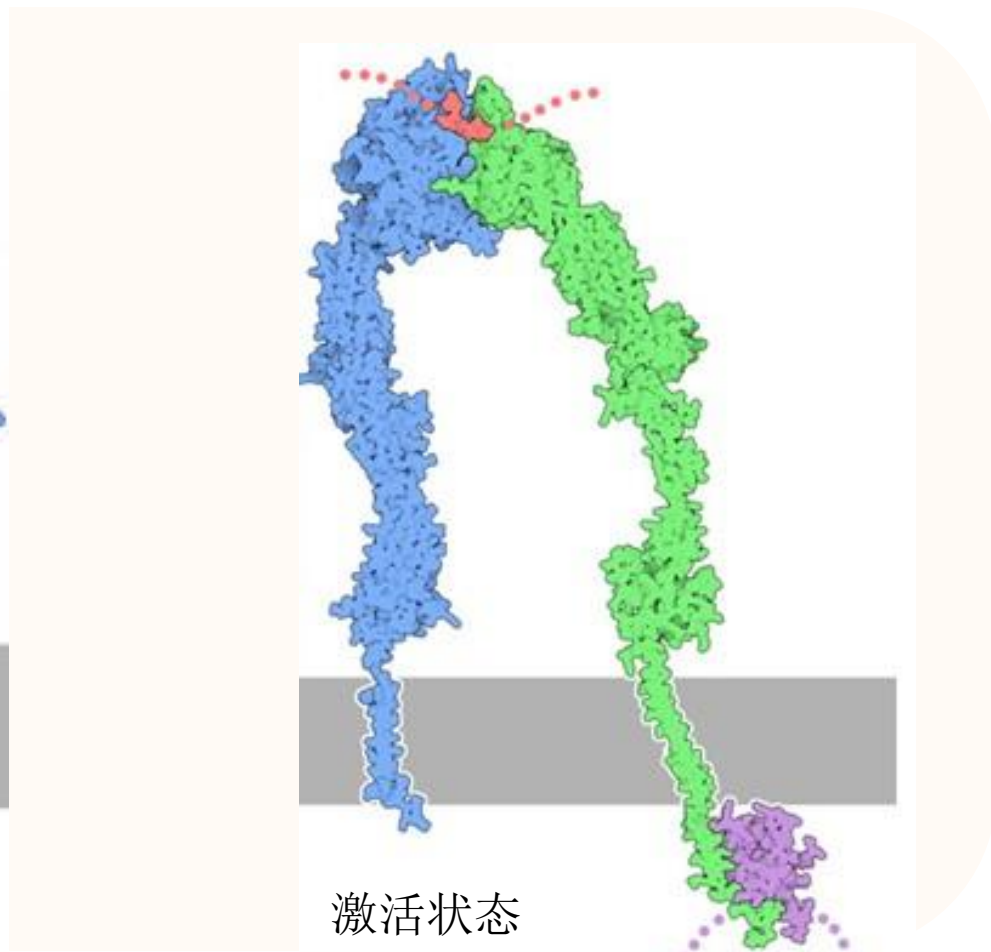
A

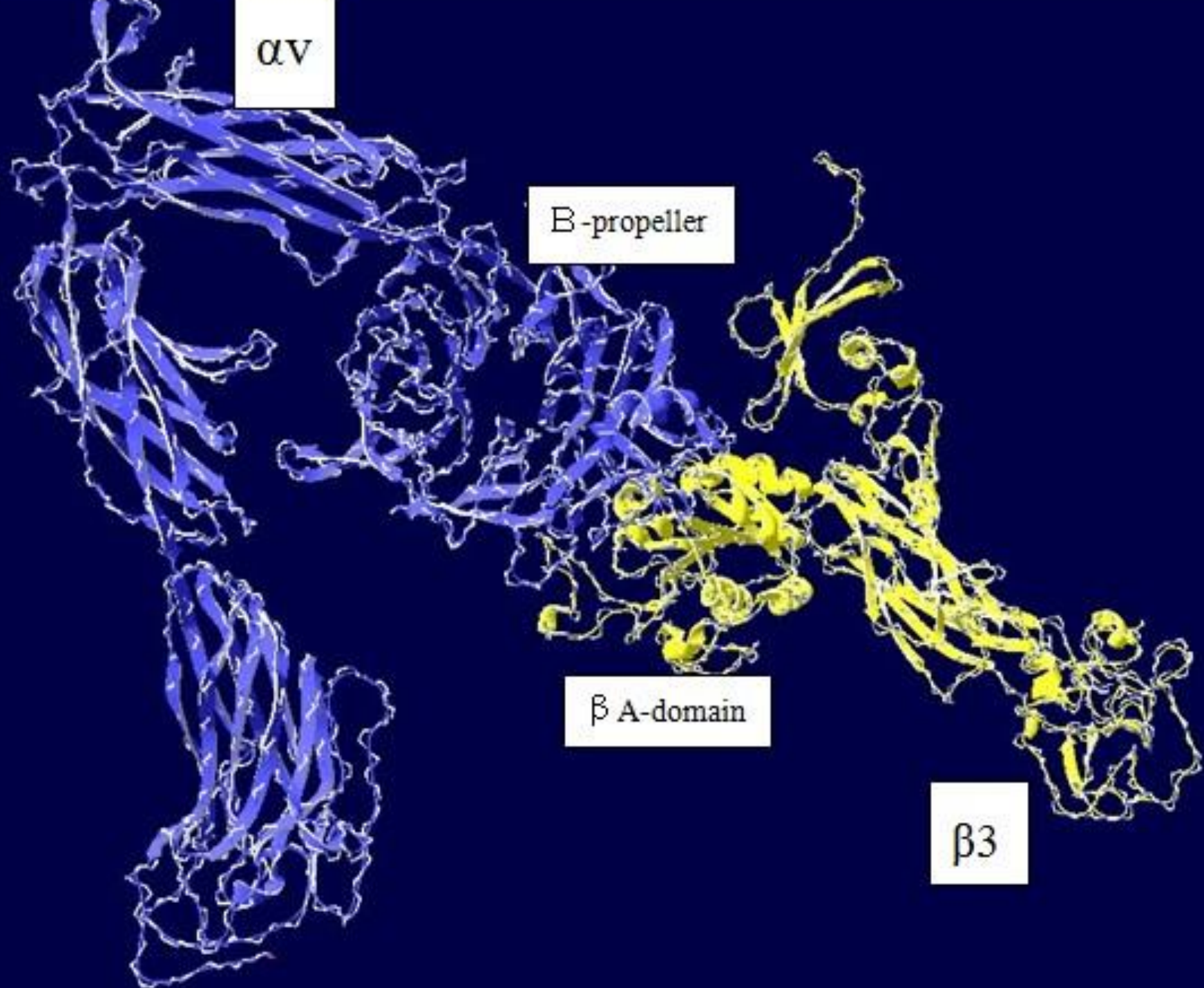


B



C





# 感谢

- 感谢罗老师的辛勤指导！
- 感谢每个组员的积极参与！





Thanks