CEA及其实例（CEAM5）分析

王群 孟艳艳 王圣惠 李世贵 刘少卿

2008年1月13日

报告内容

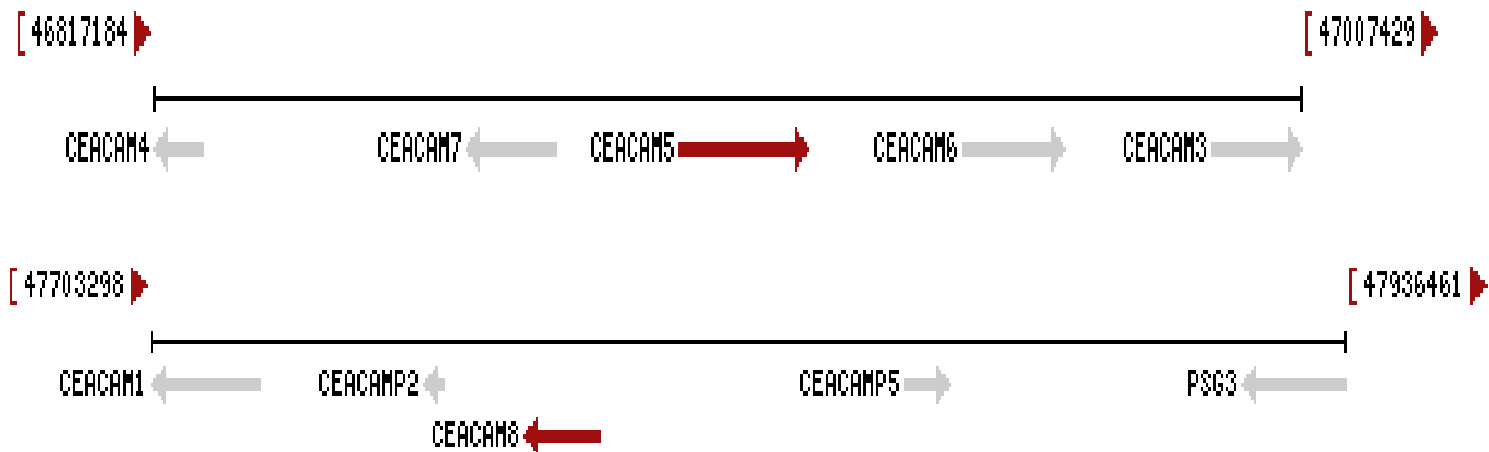- CEA简介

- CEACAM基本序列分析

- 实例（CEAM5）分析

CEA简介

- Carcinoembryonic antigen,CEA癌胚抗原

- 位于第19号染色体的长臂端

- 29个基因串联

- 两个亚家族

两个亚家族

- CEA-related cell adhesion molecule, CEACAM
  癌胚抗原相关细胞黏附分子

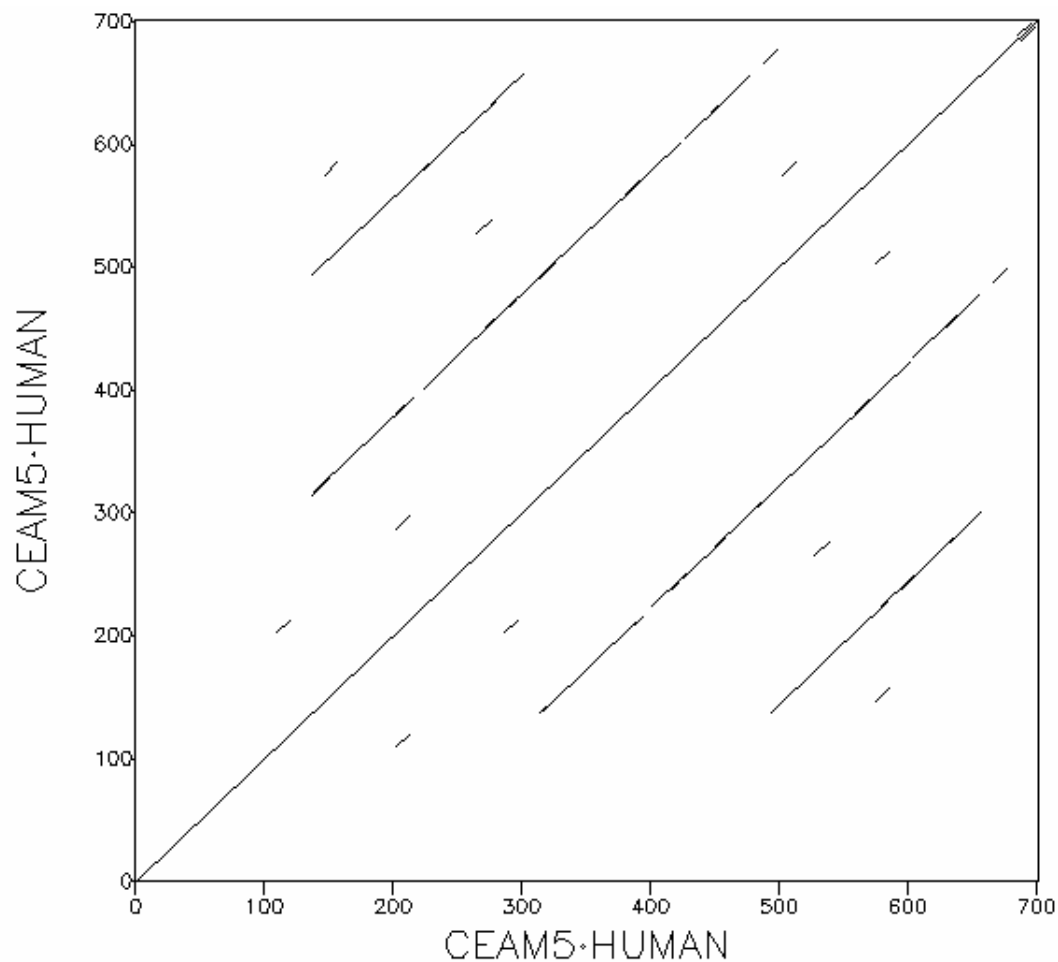- pregnancy specific-glycoprotein, PSG
  妊娠特异性糖蛋白
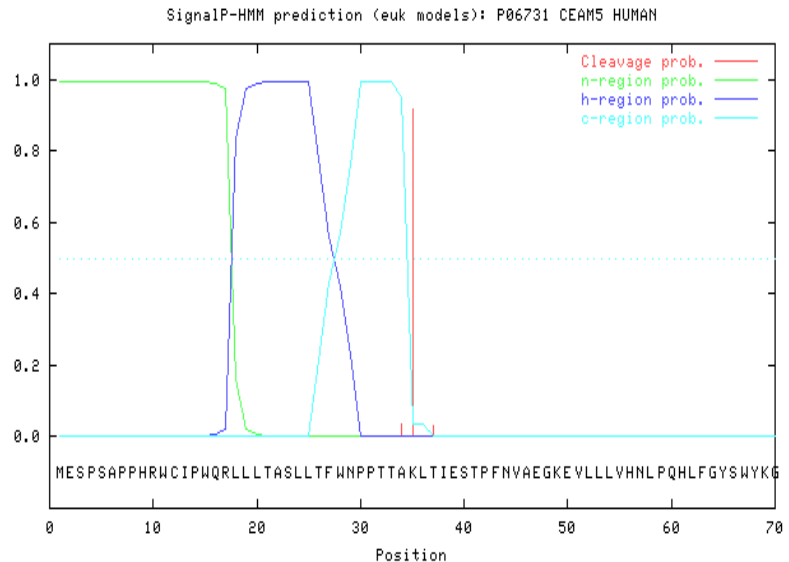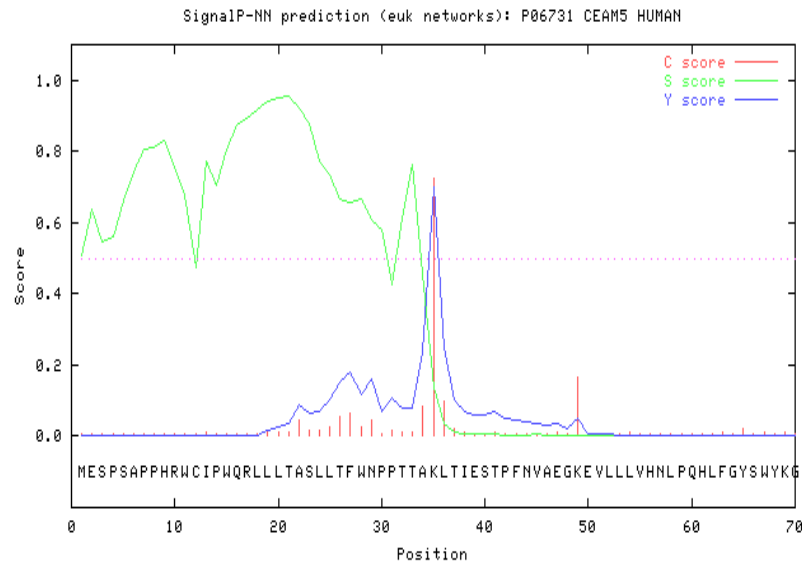
CEACAM基本序列分析

- 属于免疫球蛋白超家族

- 具有1-7个免疫球蛋白样结构域

实例（CEAM5）分析

- 通常所说的**CEA** 主要指**CEAM5**
- 肿瘤细胞抗原
- **9**个外显子
- 成熟体有**668**个氨基酸

CEAM5分析

- **1、Dotmatcher**序列比对
- **2、**信号肽预测
- **3、**跨膜分析
- **4、**糖基化位点预测
- **5、**保守序列分析
- **6、**结构功能域预测
- **7、3D**结构预测

1、Dotmatcher序列比对

2、信号肽(SignalP)分析

SignalP-NN prediction (euk networks): P06731 CEAM5 HUMAN

SignalP-HMM prediction (euk models): P06731 CEAM5 HUMAN

```
>P06731_CEAM5_HUMAN     length = 70
# Measure   Position   Value   Cutoff   signal peptide?
  max. C      35        0.723   0.32      YES
  max. Y      35        0.699   0.33      YES
  max. S      21        0.957   0.87      YES
  mean S     1-34       0.723   0.48      YES
       D     1-34       0.711   0.43      YES
# Most likely cleavage site between pos. 34 and 35: TTA-KL
```
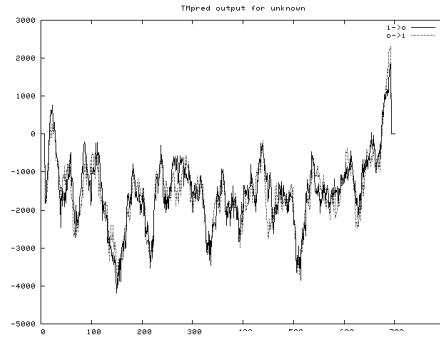
3、跨膜分析

HMMTOP预测结果

TMpred output的预测结果

TMHMM预测结果
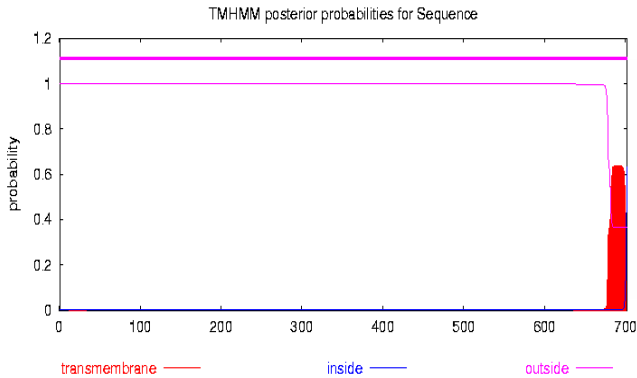
DAS预测结果

4、糖基化位点预测

```
Name:   P06731_CEAM5_HUMAN        Length:   702
MESPSAPPHRWCIPWQRLLLTASLLTFWNPPTTAKLTIESTPFNVAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQII          80
GYVIGTQQATPGPAYSGREIIYPNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYPELPKPSISSNNSKPVEDK         160
DAVAFTCEPETQDATYLWWVNNQSLPVSPRLQLSNGNRTLTLFNVTRNDTASYKCETQNPVSARRSDSVILNVLYGPDAP         240
TISPLNTSYRSGENLNLSCHAASNPPAQYSWFVNGTFQQSTQELFIPNITVNNSGSYTCQAHNSDTGLNRTTVTTITVYA         320
EPPKPFITSNNSNPVEDEDAVALTCEPEIQNTTYLWWVNNQSLPVSPRLQLSNDNRTLTLLSVTRNDVGPYECGIQNELS         400
VDHSDPVILNVLYGPDDPTISPSYTYYRPGVNLSLSCHAASNPPAQYSWLIDGNIQQHTQELFISNITEKNSGLYTCQAN         480
NSASGHSRTTVKTITVSAELPKPSISSNNSKPVEDKDAVAFTCEPEAQNTTYLWWVNGQSLPVSPRLQLSNGNRTLTLFN         560
VTRNDARAYVCGIQNSVSANRSDPVTLDVLYGPDTPIISPPDSSYLSGANLNLSCHSASNPSPQYSWRINGIPQQHTQVL         640
FIAKITPNNNGTYACFVSNLATGRNNSIVKSITVSASGTSPGLSAGATVGIMIGVLVGVALI

................................................................................          80
.........................N...........N.................................N........         160
...................N...........N......N...N........         240
.....N........N........N........N...N............N.........         320
.......N...........N...........N.........         400
......................N......................N......         480
...................N......................N......N         560
..........................N......         640
........N......         720
```
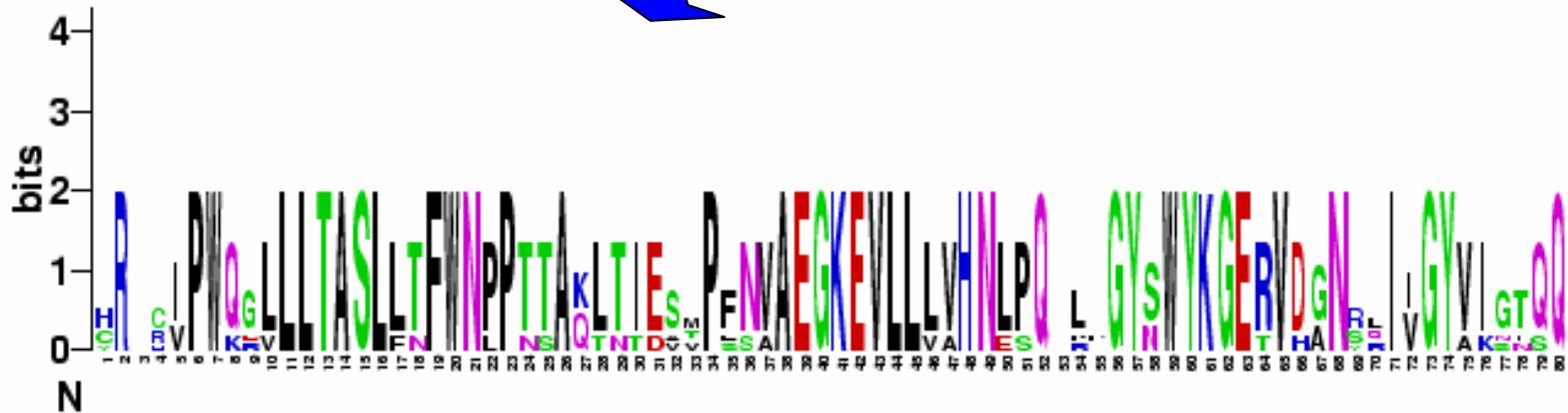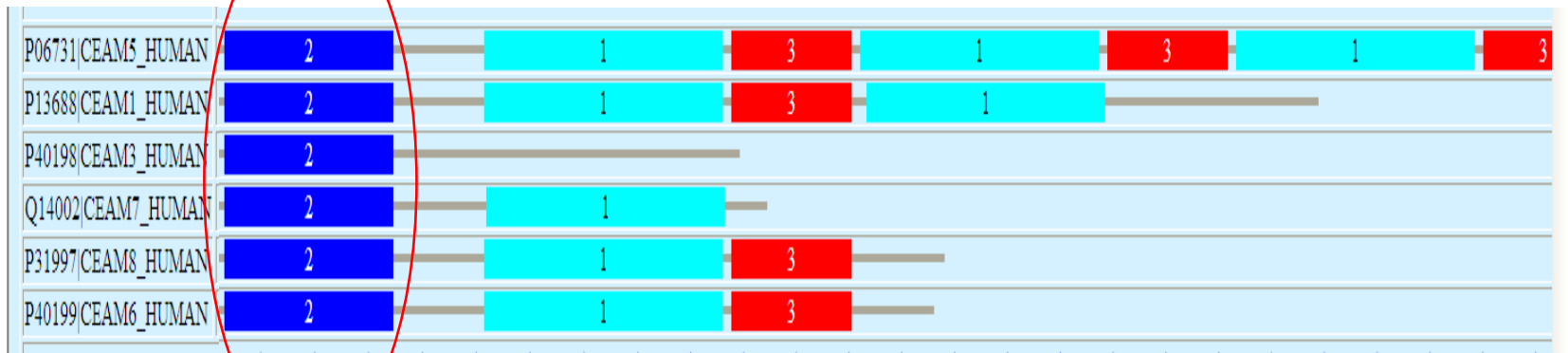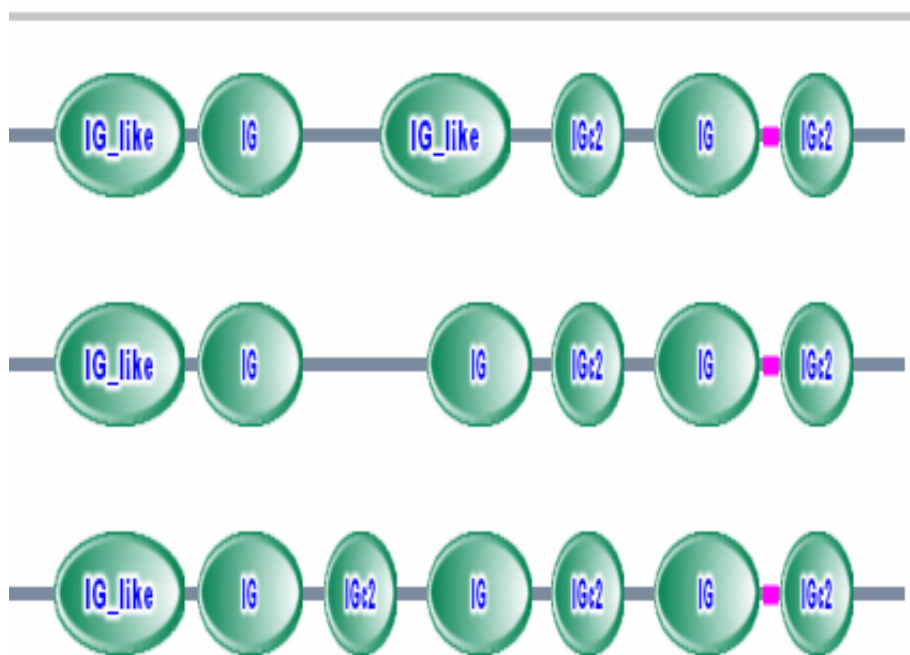
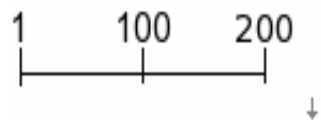**Asn-Xaa-Ser/Thr**序列（蓝色） 糖基化的**Asn**位点（红色）

5、保守结构域分析

6、SMART结构功能域预测

Confidently predicted domains, repeats, motifs and features:

Name	Begin	End	E-value
IG_like	40	141	6.35e+00
IG	152	234	2.37e-05
IGc2	250	306	6.65e-09
IG_like	294	395	6.15e-01
IG	330	412	2.64e-03
IGc2	428	484	3.56e-09
IG	508	590	4.36e-04
low complexity	593	604	-
IGc2	606	662	9.49e-05

7、Swiss Model 3D结构预测
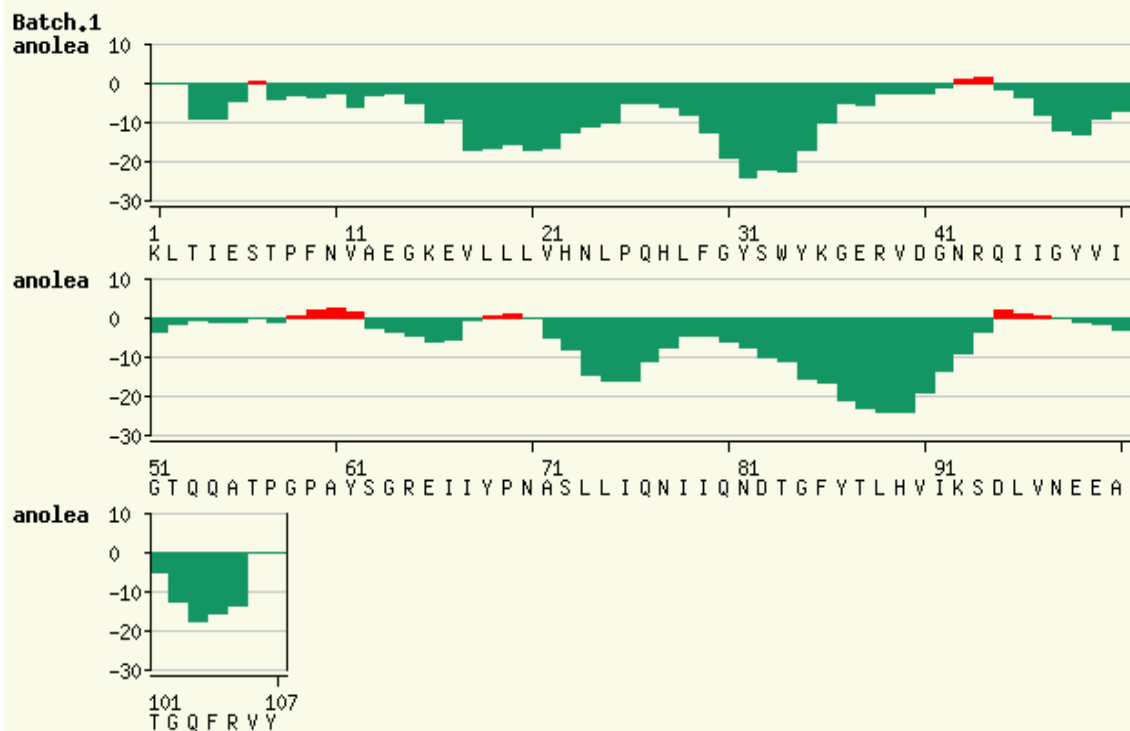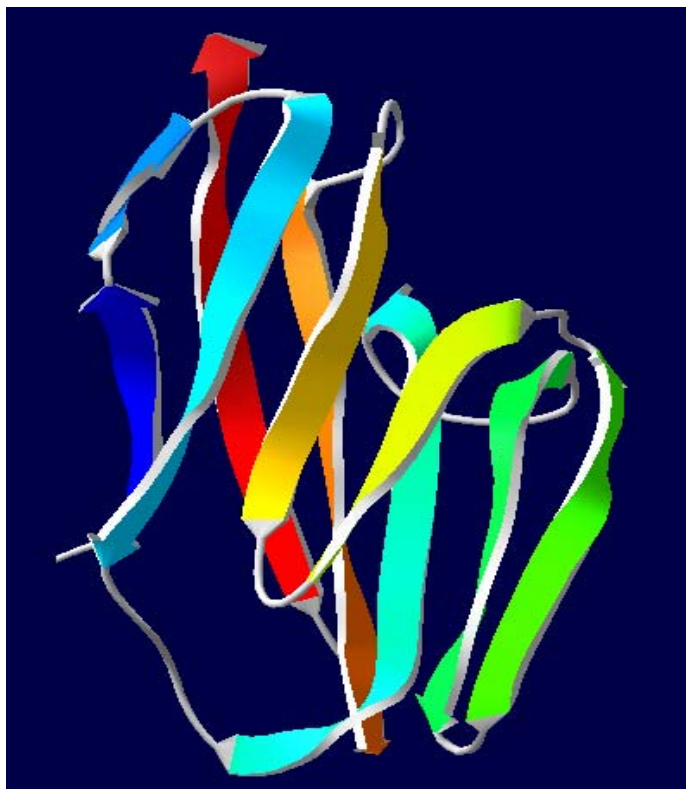
谢 谢 大 家

请 多 指 教

祝大家春节愉快！

王群 孟艳艳 王圣惠 李世贵 刘少卿