

## EMBOSS 命令实例 - 蛋白质序列分析

**pepstats**

```
pepstats PPF1_PEA.FASTA
```

功能: 统计蛋白质序列“PPF1\_PEA.FASTA”中各种氨基酸残基的含量, 输出结果如下:

```
PEPSTATS of PPF1_PEA from 1 to 442
```

Molecular weight = 48238.30                      Residues = 442  
Average Residue Weight = 109.136      Charge = 12.5  
Isoelectric Point = 9.9504  
Improbability of expression in inclusion bodies = 0.910

Residue	Number	Mole%	DayhoffStat
A = Ala	39	8.824	1.026
B = Asx	0	0.000	0.000
C = Cys	1	0.226	0.078
D = Asp	11	2.489	0.452
E = Glu	24	5.430	0.905
F = Phe	19	4.299	1.194
G = Gly	31	7.014	0.835
H = His	5	1.131	0.566
I = Ile	23	5.204	1.156
K = Lys	28	6.335	0.960
L = Leu	50	11.312	1.529
M = Met	6	1.357	0.799
N = Asn	18	4.072	0.947
P = Pro	27	6.109	1.175
Q = Gln	25	5.656	1.450
R = Arg	17	3.846	0.785
S = Ser	42	9.502	1.357
T = Thr	27	6.109	1.001
V = Val	31	7.014	1.063
W = Trp	7	1.584	1.218
X = Xaa	0	0.000	0.000
Y = Tyr	11	2.489	0.732
Z = Glx	0	0.000	0.000

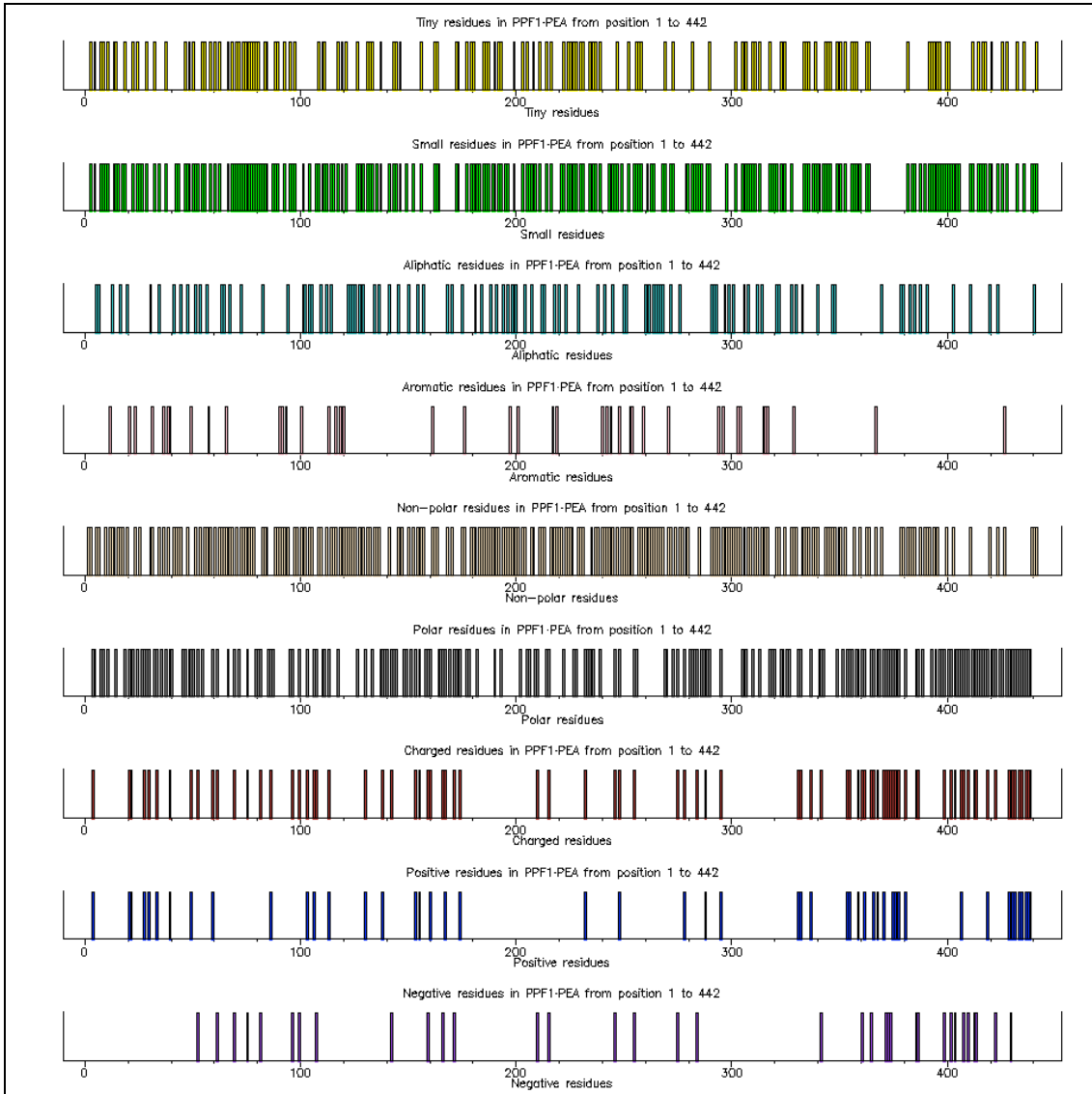
  

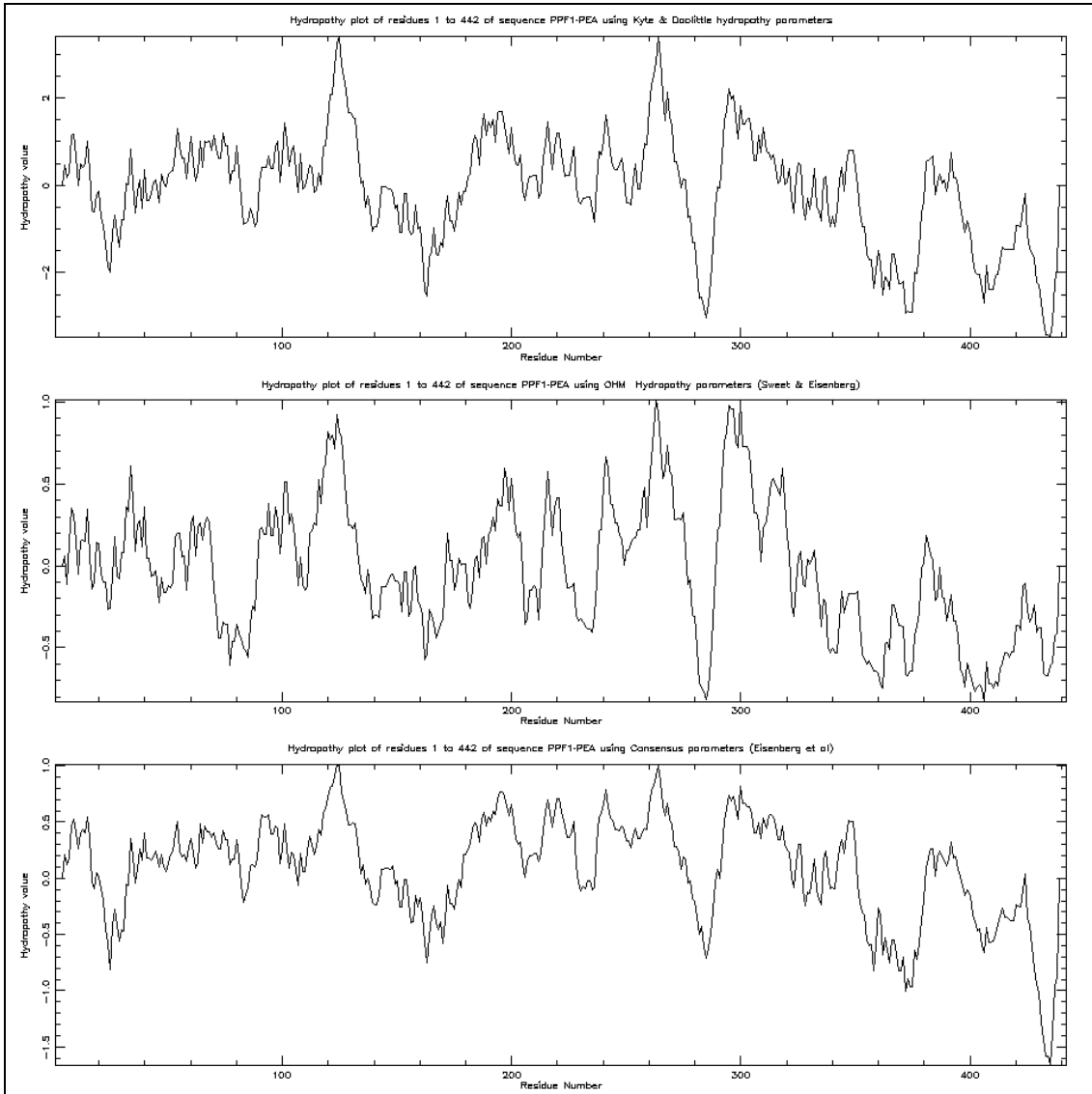
Property	Residues	Number	Mole%
Tiny	(A+C+G+S+T)	140	31.674
Small	(A+B+C+D+G+N+P+S+T+V)	227	51.357
Aliphatic	(I+L+V)	104	23.529
Aromatic	(F+H+W+Y)	42	9.502
Non-polar	(A+C+F+G+I+L+M+P+V+W+Y)	245	55.430
Polar	(D+E+H+K+N+Q+R+S+T+Z)	197	44.570
Charged	(B+D+E+H+K+R+Z)	85	19.231
Basic	(H+K+R)	50	11.312
Acidic	(B+D+E+Z)	35	7.919

**pepinfo**

```
pepinfo PPF1_PEA.FASTA
```

功能: 以图形方式显示蛋白质序列“PPF1\_PEA.FASTA”中各种不同性质的氨基酸残基的含量, 输出结果如下:





**garnier**

garnier PPF1\_PEA.FASTA

功能: 用 GOR 方法预测蛋白质序列 “PPF1\_PEA.FASTA” 中二级结构, 结果如下:

```
#####
# Program: garnier
# Rundate: Mon Apr 05 15:32:11 2004
# Report_format: tagseq
# Report_file: ppf1_pea.garnier
#####

#=====
#
# Sequence: PPF1_PEA      from: 1   to: 442
# HitCount: 116
#
```

```

# DCH = 0, DCS = 0
#
# Please cite:
# Garnier, Osguthorpe and Robson (1978) J. Mol. Biol. 120:97-120
#
#=====
      . 10   . 20   . 30   . 40   . 50
      MAKTLISSPSFLGTPLPSLHRTFSPNRTLFTKVQFSFHQLPPIQSVSHS
helix HHH
sheet  EEEE          EE          EEEE E          EEEEE
turns          TTT   T  TT   TTTT   TT TTT   TT
coil      CC  CCCCC   CCC          CCC   C
      . 60   . 70   . 80   . 90   . 100
      VDLSGIFARAEGLLYTLADATVAADAAASTDVAAQKNGWFGFISDGMEF
helix  HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
sheet  E  EE          E          EEE
turns  T          TTTT
coil   CC          C  CCCC
      . 110  . 120  . 130  . 140  . 150
      VLKVLKDGLSSVHPYSYGFALLLTIVIVKAATLPLTKQVESTLAMQNL
helix  HHHHHH          H  HHHH  HHHHHH  HHHHHHH
sheet  EEEEEE  EEEEEE  EE          EEE
turns  TT          TTTT
coil   C          C
      . 160  . 170  . 180  . 190  . 200
      QPKIKAIQERYAGNQERIQLETSRLYTQAGVNPLAGCLPTLATIPVWIGL
helix  HH  H
sheet  EEEEE          EEEE          EE  EE  EEEEEEEEEEEEEEE
turns  T          TTT          TTTTT  TT  TTT
coil   CCCCCC          C  CCCC          C
      . 210  . 220  . 230  . 240  . 250
      YQALSNVANEGLLTEGFLWIPSLGGPTSIAARQSGSGISWLFPFVDGHPL
helix  HHHHHH
sheet  EE          EE  EE          EEEEE          EEEEE
turns          T  TT          TTTT          TTT
coil   CCCCCC          C  CCCC          CC  CCCC
      . 260  . 270  . 280  . 290  . 300
      LGWYDTAAAYLVLPVLLIVSQYVSMEIMKPPQTNDPNQKNTLLIFKFLPLM

helix  H
sheet  E EEEEEEEEEEEEEEE EEEEEEE          EEEEEEEEEEE E
turns  TT T          TTT  TTT          T
coil  C          CC  CCC
      . 310  . 320  . 330  . 340  . 350
      IGYFSLSPVPSGLTIYWFTNNVLSTAQQVWLRKLGAKPAVNENAGGIITA
helix  HHH          HH          H
sheet  EEEEEE          EEEEE          EEEEE          EEE          EEEE
turns          TT          TTT          TTT
coil   CCC          CCCCCC          CCCC          C
      . 360  . 370  . 380  . 390  . 400
      GQAKRSASKPEKGGERFRLKEEEKKKLIKALPVEEVQPLASASASNDG
helix  HH          HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
sheet
turns  T          TTT          T
coil   CC CCCCC  C          C          CCCC C
      . 410  . 420  . 430  . 440
      SDVENNKEQEVTEESNTSKVSQEVQSFSTRERRSKRSKRKPVA
helix  HHHHHHHHHH          HHHH
sheet          EEEE          EE
turns          T          TTTTTTT
coil  CCCCC          CCCC  CCC          CC
#-----

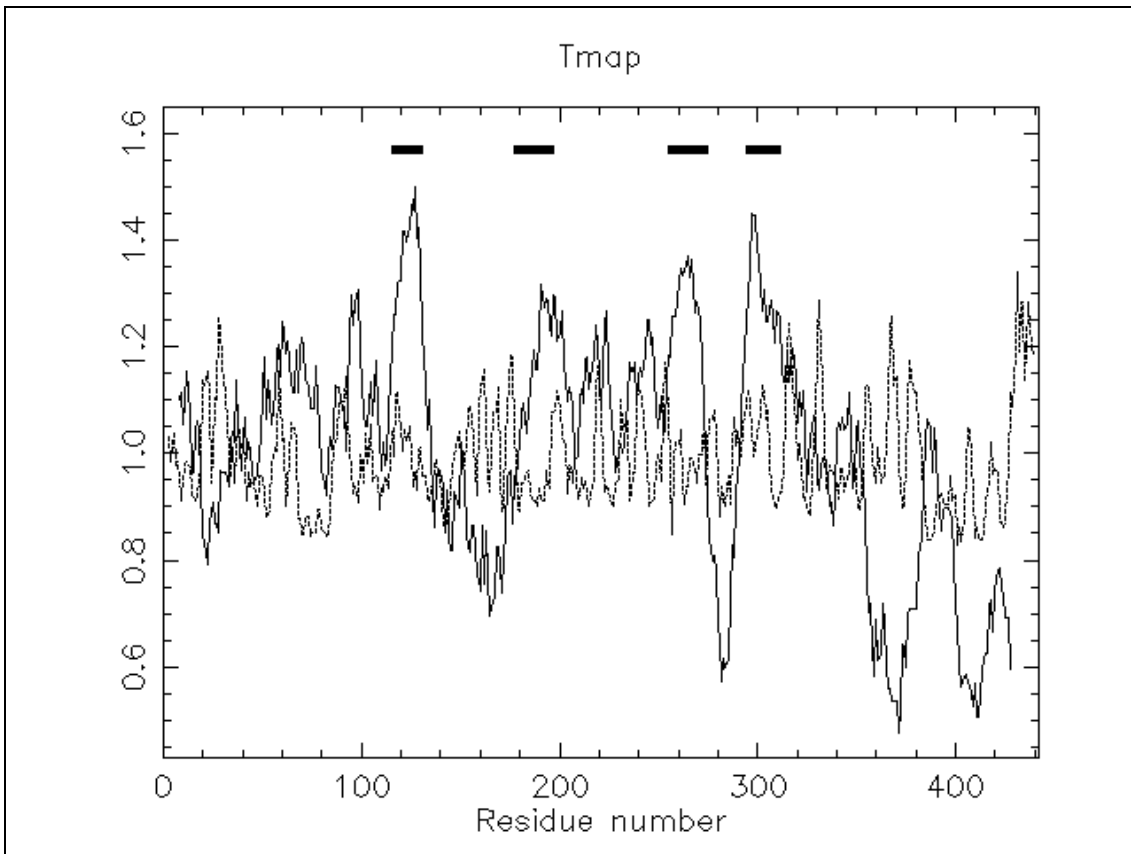
```

```
#  
# Residue totals: H:115 E:147 T: 83 C: 97  
# percent: H: 27.0 E: 34.5 T: 19.5 C: 22.8  
#  
#-----
```

### tmap

tmap PPF1\_PEA.FASTA graph x11

功能: 预测蛋白质序列“PPF1\_PEA.FASTA”中跨膜区, 结果如下:



输出文件 tmap.res

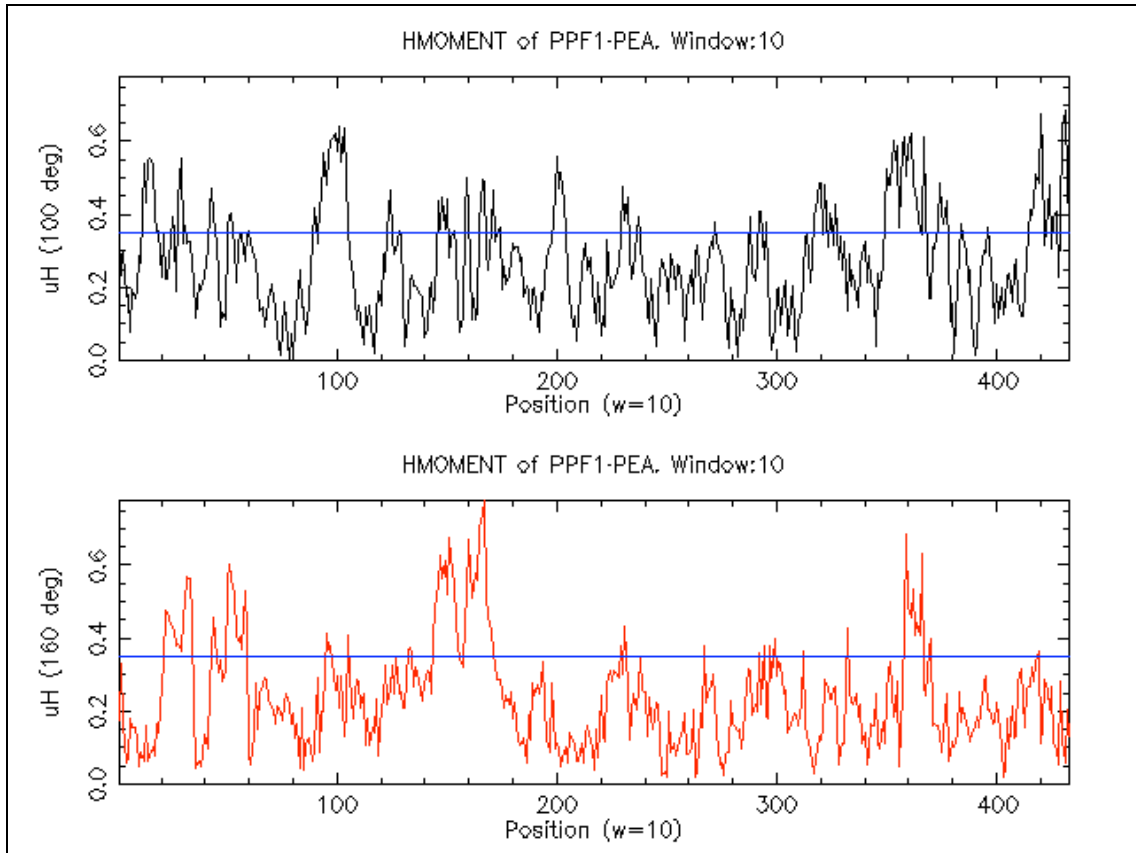
```
PREDICTED TRANSMEMBRANE SEGMENTS FOR PROTEIN PPF1_PEA_1  
  
TM 1: 111 - 135 (25)  
TM 2: 173 - 201 (29)  
TM 3: 251 - 279 (29)  
TM 4: 290 - 315 (26)
```

### hmoment

hmoment PPF1\_PEA.FASTA -plot -double -graph png

功能: 以图形方式显示蛋白质序列“PPF1\_PEA.FASTA”可能形成 alpha 螺旋和 beta

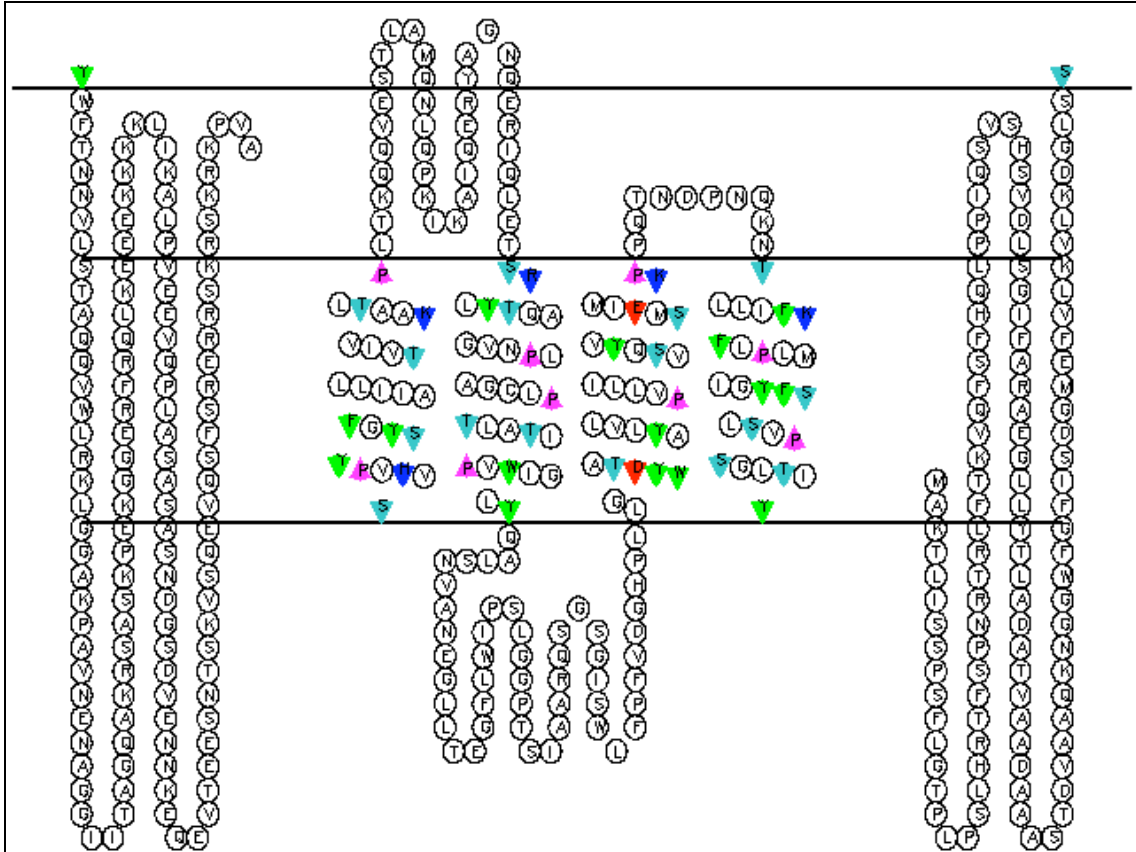
折叠的片段，输出结果如下：



**topo**

```
topo PPF1_PEA.FASTA -prol=um -cys=cg -gly=sr -acid=dr -basic=db -ohaa=da  
-araa=dg -section "111-135 173-201 251-279 290-315" -nomem -graph png
```

功能：以图形方式显示蛋白质序列“PPF1\_PEA.FASTA”跨膜拓扑结构，蓝色表示碱性残基，红色表示酸性残基，绿色表示芳香族残基，粉色表示脯氨酸，浅蓝色表示丝氨酸和苏氨酸，输出结果如下：



### pepwheel

```
pepwheel PPF1_PEA.FASTA -sbegin 111 -send 135 -graph x11
```

功能：以图形方式显示蛋白质序列“PPF1\_PEA.FASTA”中 111-135 残基的螺旋轮，输出结果如下：

