THE WIDE EXTENT OF INTER-STRAIN DIVERSITY IN VIRULENT AND VACCINE STRAINS OF PRV

猪伪狂犬病毒（PRV）疫苗株和野毒株间出现大量变异的生物信息学分析

Reporter：叶超

Members：张家林 白洁 齐宾宾

SYLLABUS

- Ⅰ. Background
- Ⅱ. Main relative virulent gene analysis
- -TK gene （白洁）
- -gH gene （齐宾宾）
- Ⅲ. Main antigen epitope gene analysis
- -gB gene （叶超）
- Ⅳ. gB amino acid sequence BLAST(张家林)

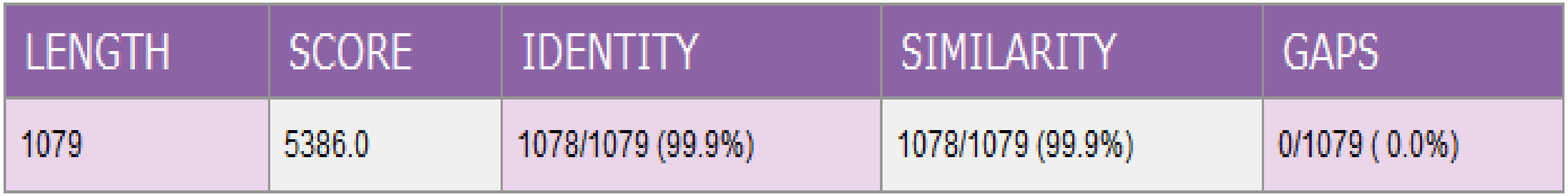
BACKGROUND

- Before 2011, Bartha (PRV vaccine strain) could provide very high protective efficacy .
- But nowadays, Bartha doesn’t work well.

- Bartha (PRV vaccine strain)
- Becker (PRV wild strain)
- we analyse the differences between vaccine and virulent strains of PRV.

TK GENE

- -Needle
- TK gene complete genome

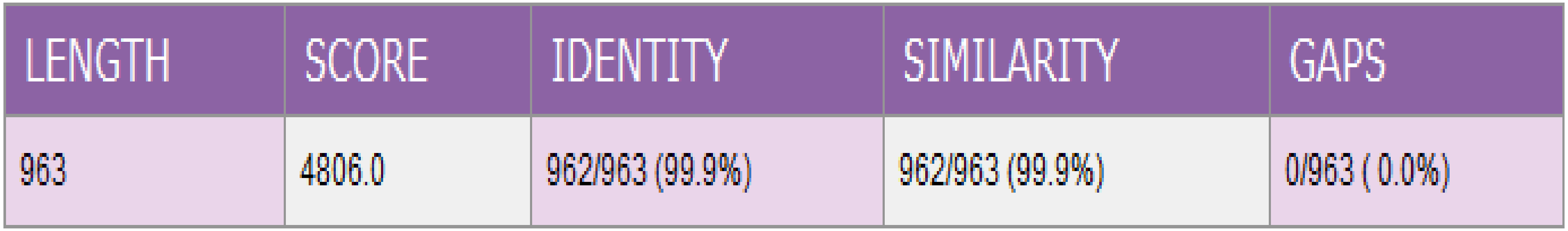


LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
1079	5386.0	1078/1079 (99.9%)	1078/1079 (99.9%)	0/1079 (0.0%)

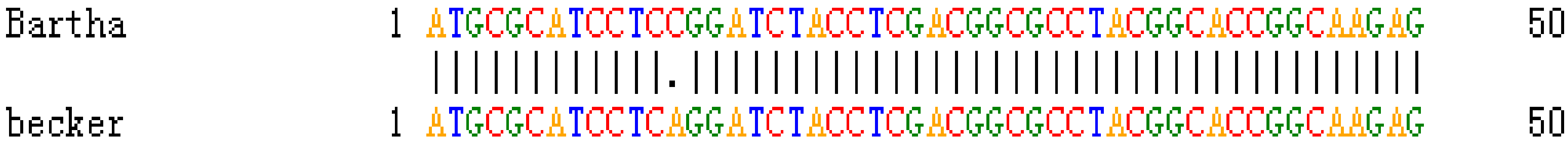
```
Bartha      1 ATGCGCATCCTCCGGATCTACCTCGACGGCGCCTACGGCACCGGCAAGAG     50
              ||||||||||||.|||||||||||||||||||||||||||||||||||||
Becker      1 ATGCGCATCCTCAGGATCTACCTCGACGGCGCCTACGGCACCGGCAAGAG     50
```

TK GENE

- Needle
- CDS



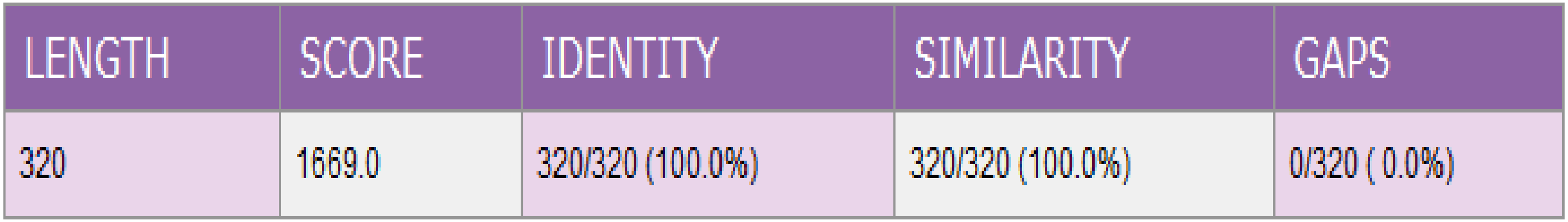
LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
963	4806.0	962/963 (99.9%)	962/963 (99.9%)	0/963 (0.0%)



```
Bartha               1 ATGCGCATCCTCCGGATCTACCTCGACGGCGCCTACGGCACCGGCAAGAG     50
                       ||||||||||||.|||||||||||||||||||||||||||||||||||||
becker               1 ATGCGCATCCTCAGGATCTACCTCGACGGCGCCTACGGCACCGGCAAGAG     50
```

TK GENE

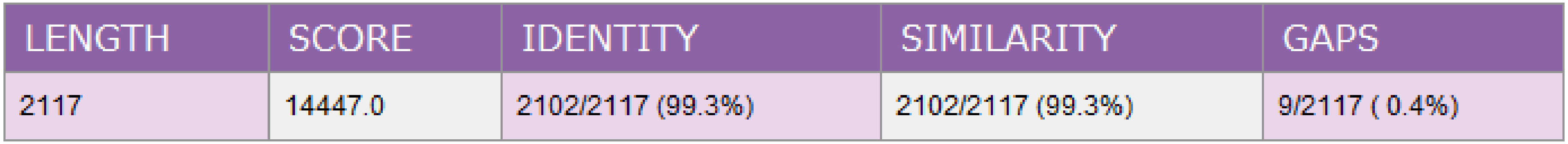
- Needle
- Amino Acid



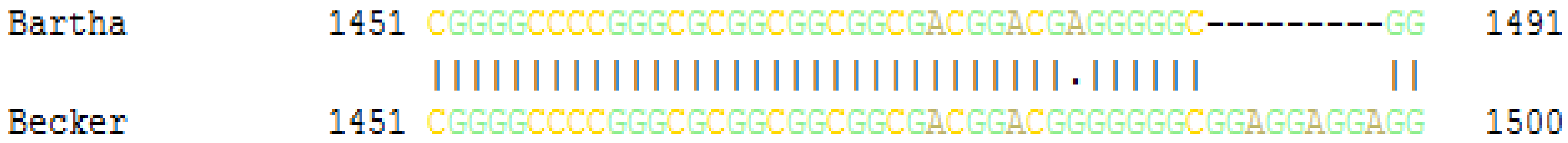
LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
320	1669.0	320/320 (100.0%)	320/320 (100.0%)	0/320 (0.0%)

GH GENE

- -Needle
- gH gene complete genome



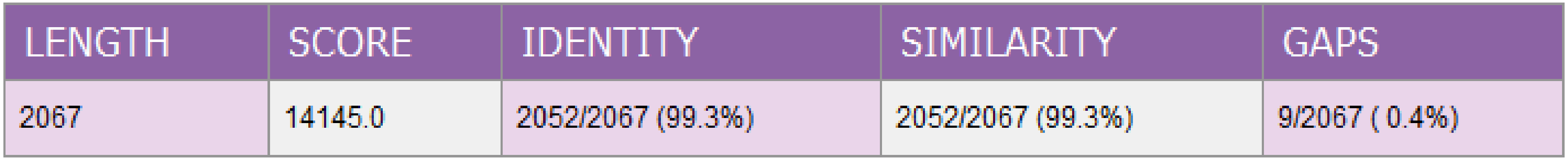
LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
2117	14447.0	2102/2117 (99.3%)	2102/2117 (99.3%)	9/2117 (0.4%)



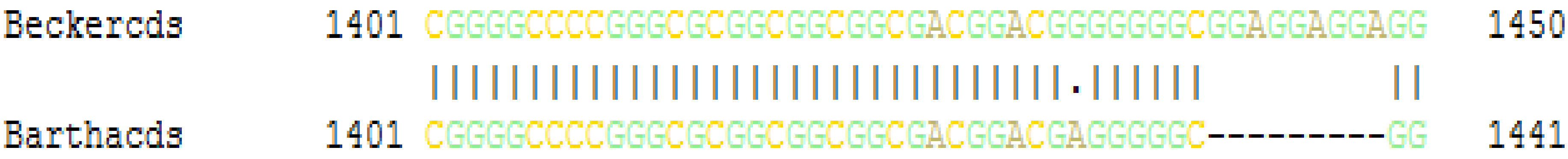
```
Bartha          1451 CGGGGCCCCGGGCGCGGCGGCGGCGACGGACGAGGGGGC---------GG   1491
                     ||||||||||||||||||||||||||||||||.||||||         ||
Becker          1451 CGGGGCCCCGGGCGCGGCGGCGGCGACGGACGGGGGGGCGGAGGAGGAGG   1500
```

gH GENE

- -Needle
- gH CDS



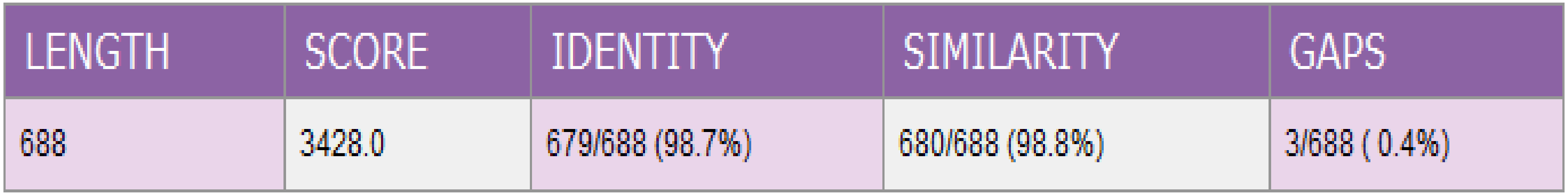
LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
2067	14145.0	2052/2067 (99.3%)	2052/2067 (99.3%)	9/2067 (0.4%)



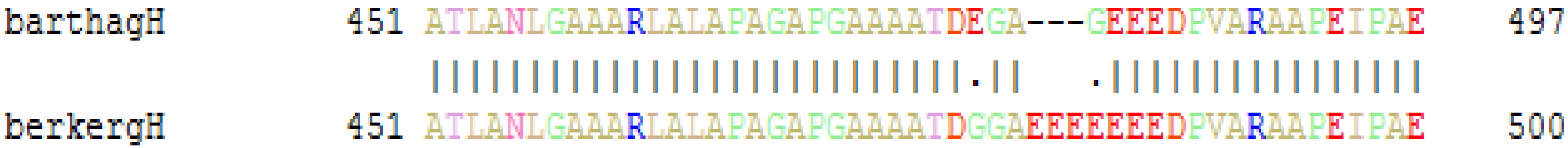
```
Beckercds        1401 CGGGGCCCCGGGCGCGGCGGCGGCGACGGACGGGGGGGCGGAGGAGGAGG   1450
                      ||||||||||||||||||||||||||||||||.||||||         ||
Barthacds        1401 CGGGGCCCCGGGCGCGGCGGCGGCGACGGACGAGGGGGC---------GG   1441
```

GH GENE

- -Needle
- gH Amino Acid



LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
688	3428.0	679/688 (98.7%)	680/688 (98.8%)	3/688 (0.4%)

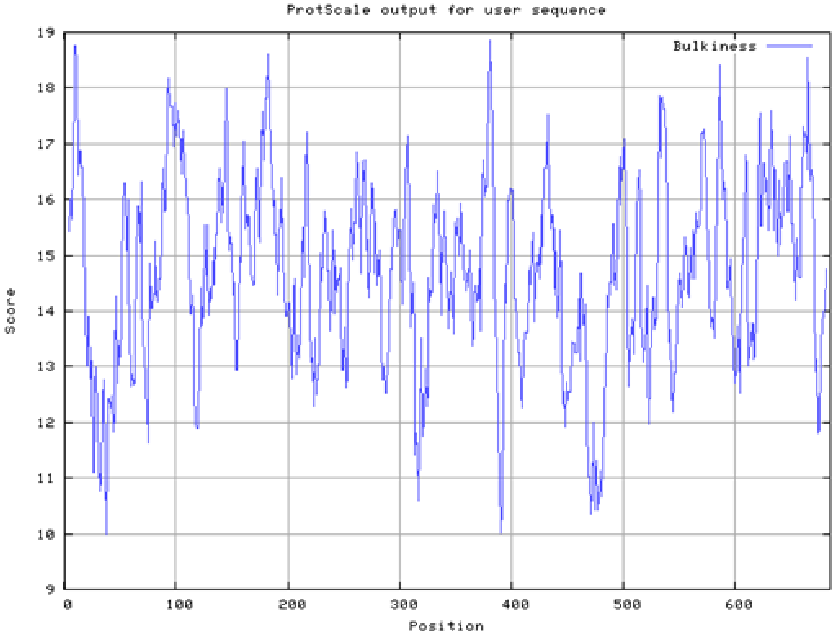


```
barthagH         451 ATLANLGAAARLALAPAGAPGAAAATDEGA---GEEEDPVARAAPEIPAE    497
                     |||||||||||||||||||||||||||.||   .||||||||||||||||
berkergH         451 ATLANLGAAARLALAPAGAPGAAAATDGGAEEEEEEEDPVARAAPEIPAE    500
```

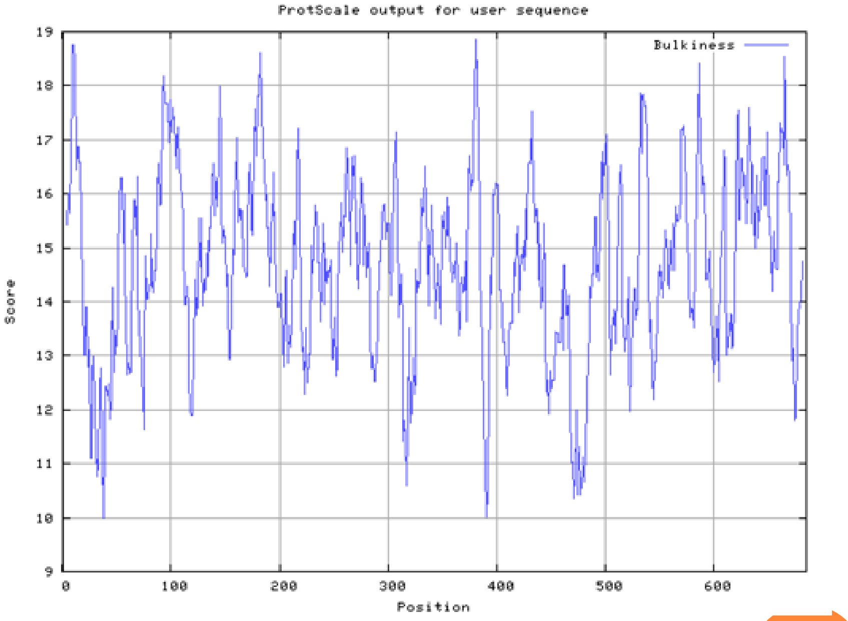
gH AMINO ACID ANALYSIS

Bulkiness

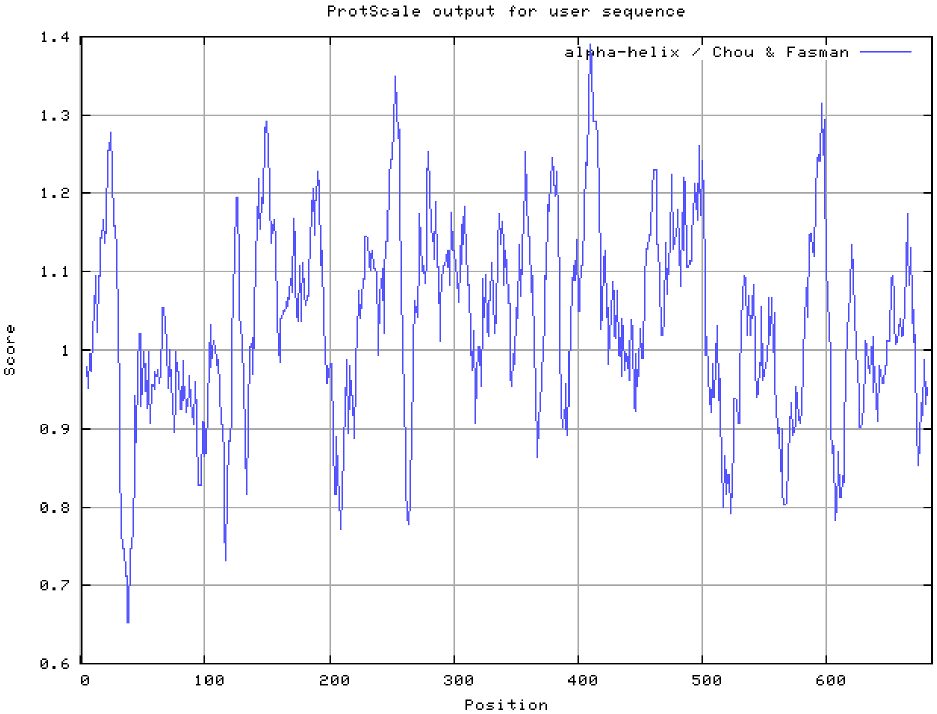
Bartha

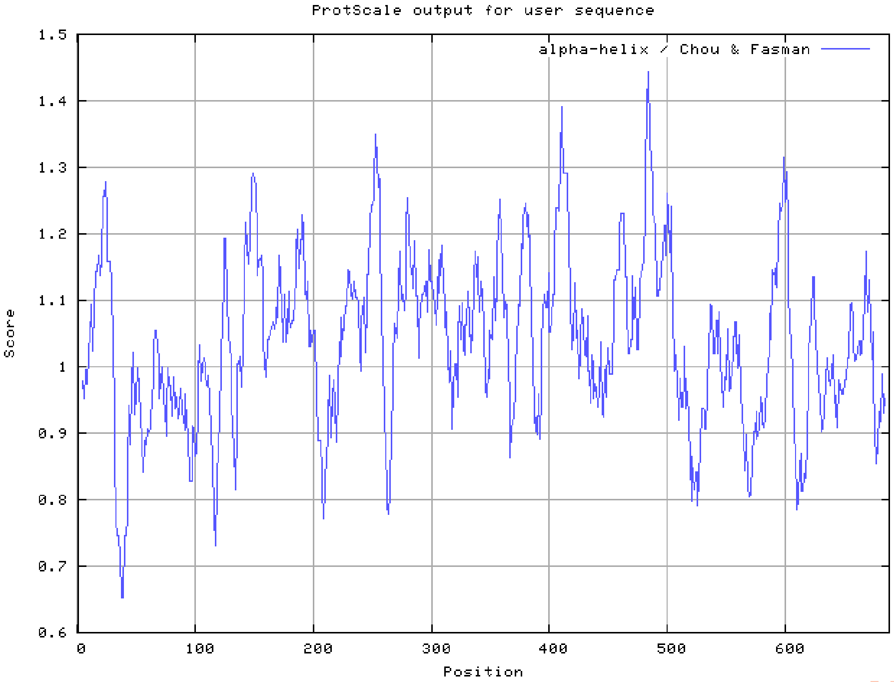


Becker



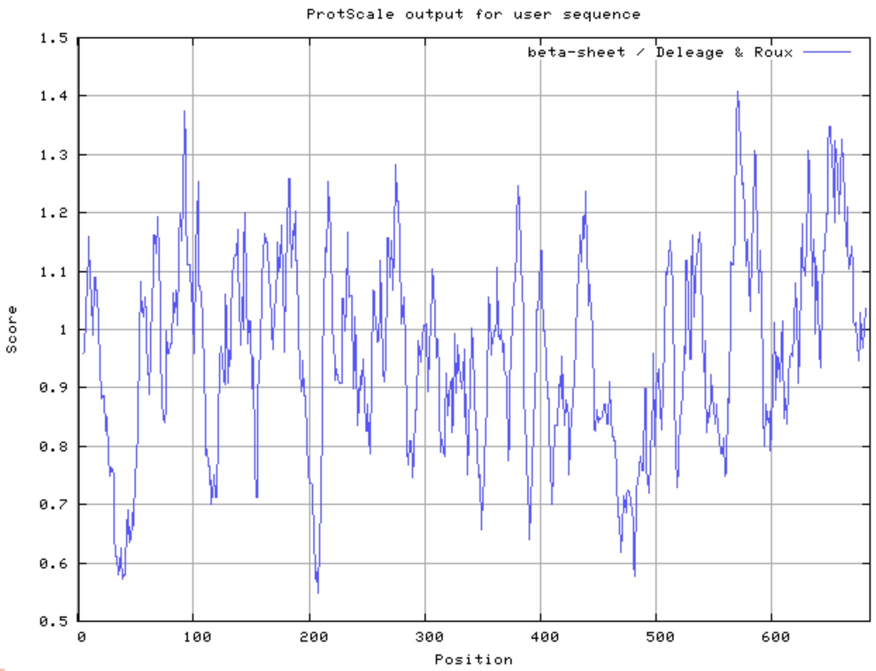
- Alpha helix
- Bartha Becker

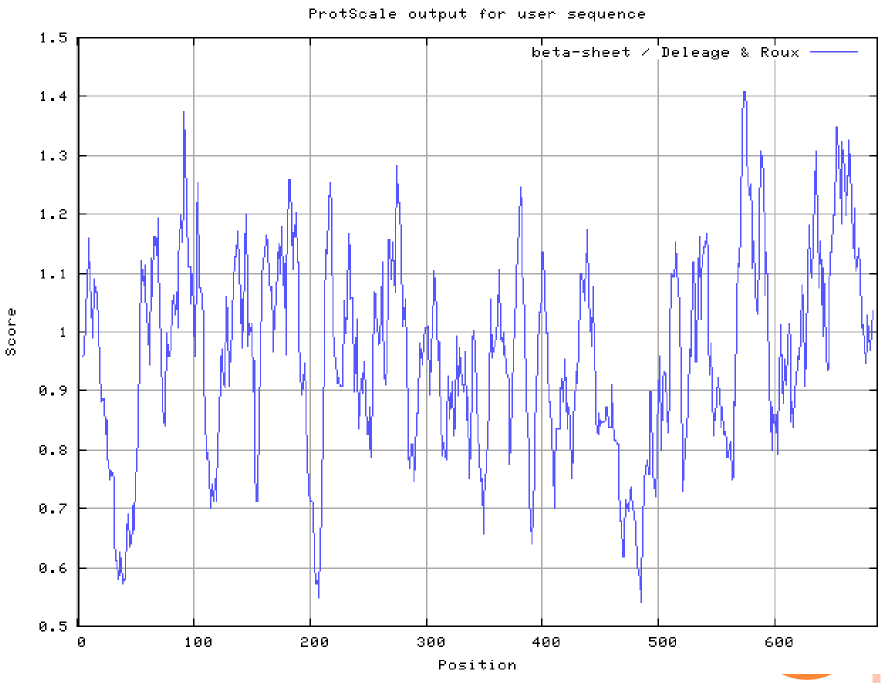




- Beta-sheet

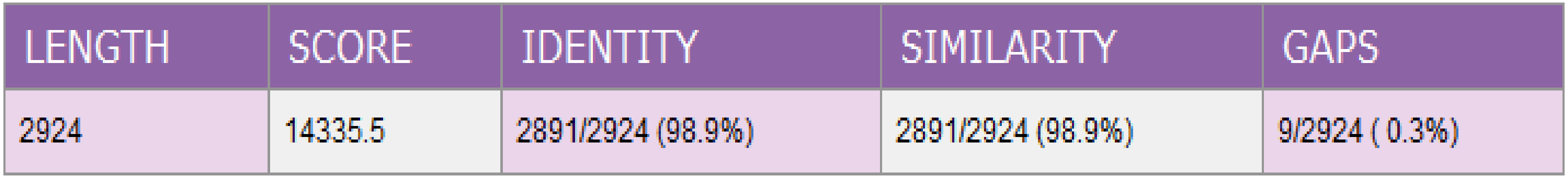
- Bartha Becker



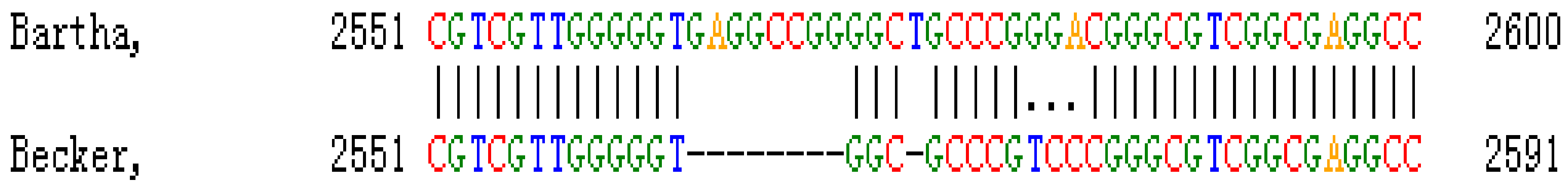


GB GENE

- -Needle
- gB gene complete genome



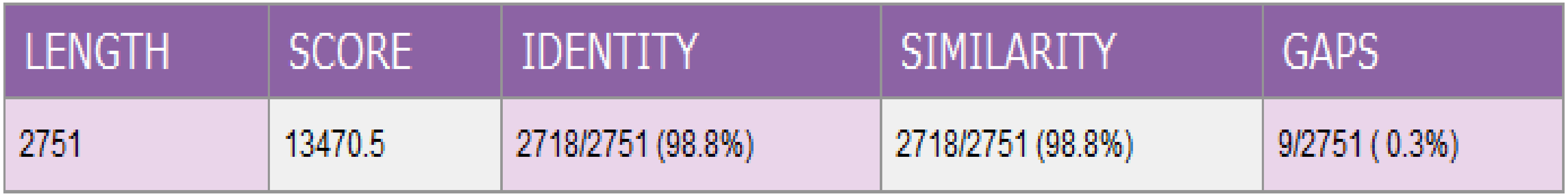
LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
2924	14335.5	2891/2924 (98.9%)	2891/2924 (98.9%)	9/2924 (0.3%)



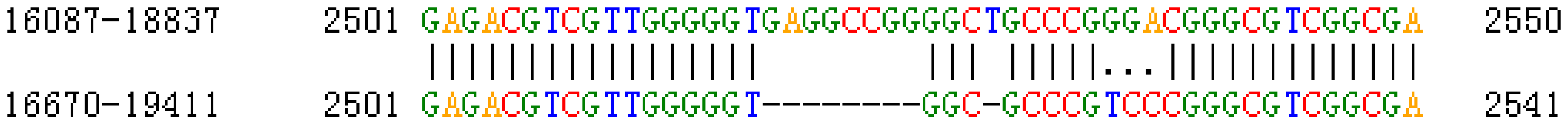
```
Bartha,          2551 CGTCGTTGGGGGTGAGGCCGGGGCTGCCCGGGACGGGCGTCGGCGAGGCC   2600
                      |||||||||||||        ||| |||||...|||||||||||||||||
Becker,          2551 CGTCGTTGGGGGT--------GGC-GCCCGTCCCGGGCGTCGGCGAGGCC   2591
```

GB GENE

- -Needle
- gB CDS



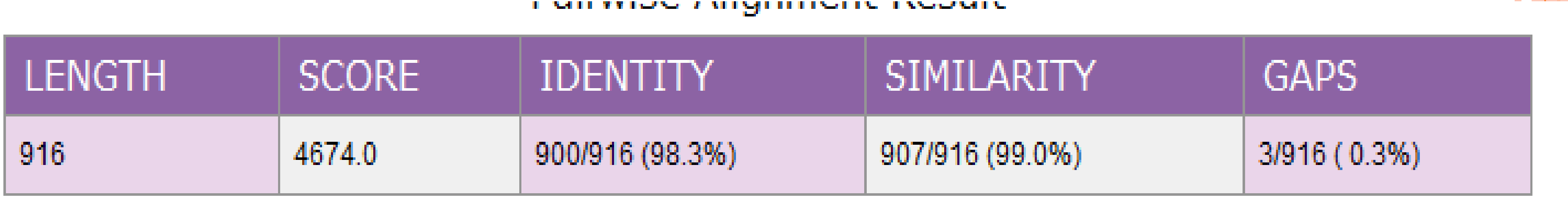
LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
2751	13470.5	2718/2751 (98.8%)	2718/2751 (98.8%)	9/2751 (0.3%)



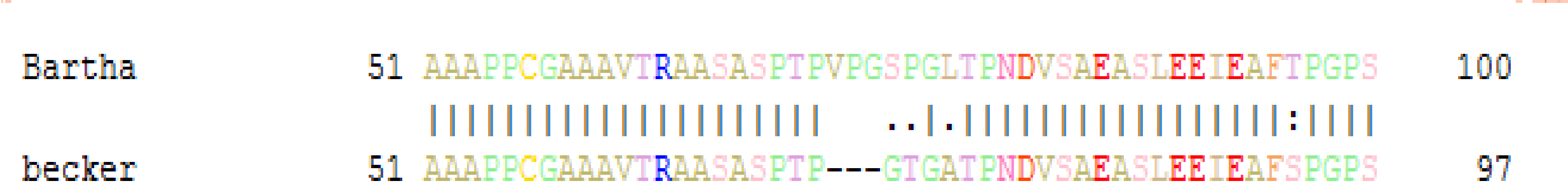
```
16087-18837       2501 GAGACGTCGTTGGGGGTGAGGCCGGGGCTGCCCGGGACGGGCGTCGGCGA    2550
                       |||||||||||||||||        ||| |||||...|||||||||||||
16670-19411       2501 GAGACGTCGTTGGGGGT--------GGC-GCCCGTCCCGGGCGTCGGCGA    2541
```

GB GENE

- -Needle
- gB Amino Acid



LENGTH	SCORE	IDENTITY	SIMILARITY	GAPS
916	4674.0	900/916 (98.3%)	907/916 (99.0%)	3/916 (0.3%)

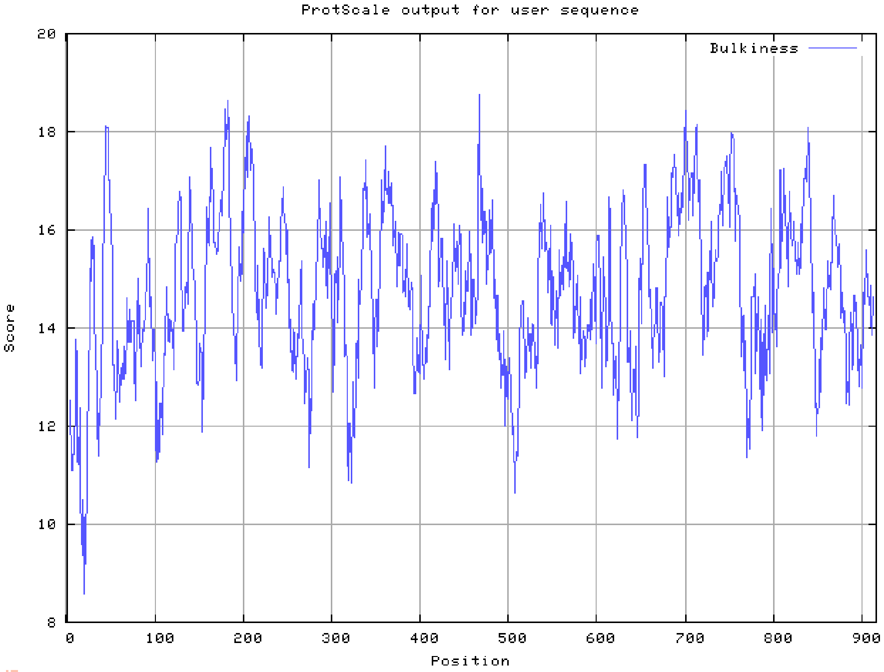


```
Bartha               51 AAAPPCGAAAVTRAASASPTPVPGSPGLTPNDVSAEASLEEIEAFTPGPS    100
                        |||||||||||||||||||||   ..|.|||||||||||||||||:||||
becker               51 AAAPPCGAAAVTRAASASPTP---GTGATPNDVSAEASLEEIEAFSPGPS     97
```

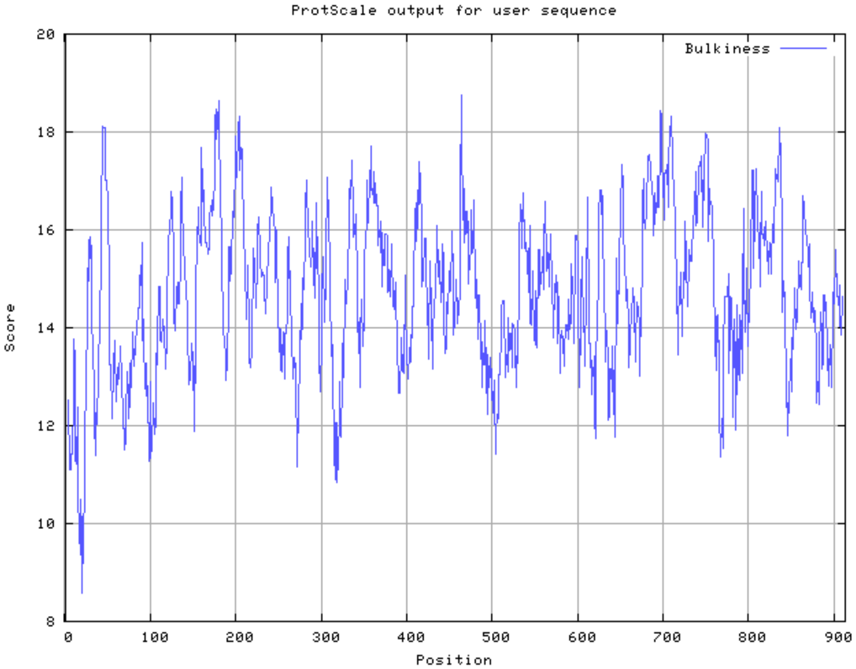
GB AMINO ACID ANALYSIS

Bulkiness

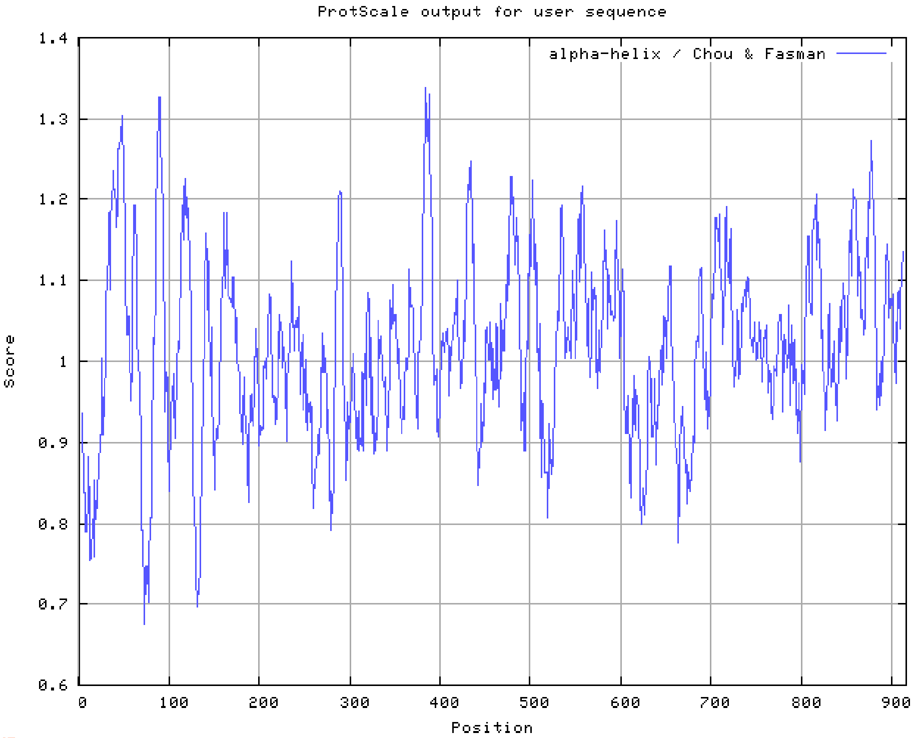
Bartha

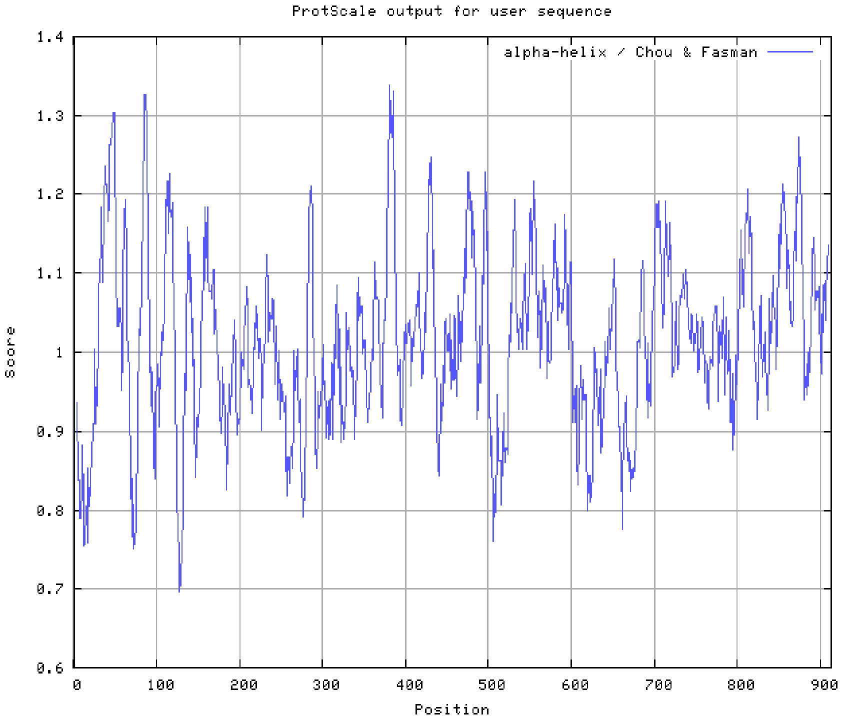


Becker



- Alpha helix
- Bartha Becker

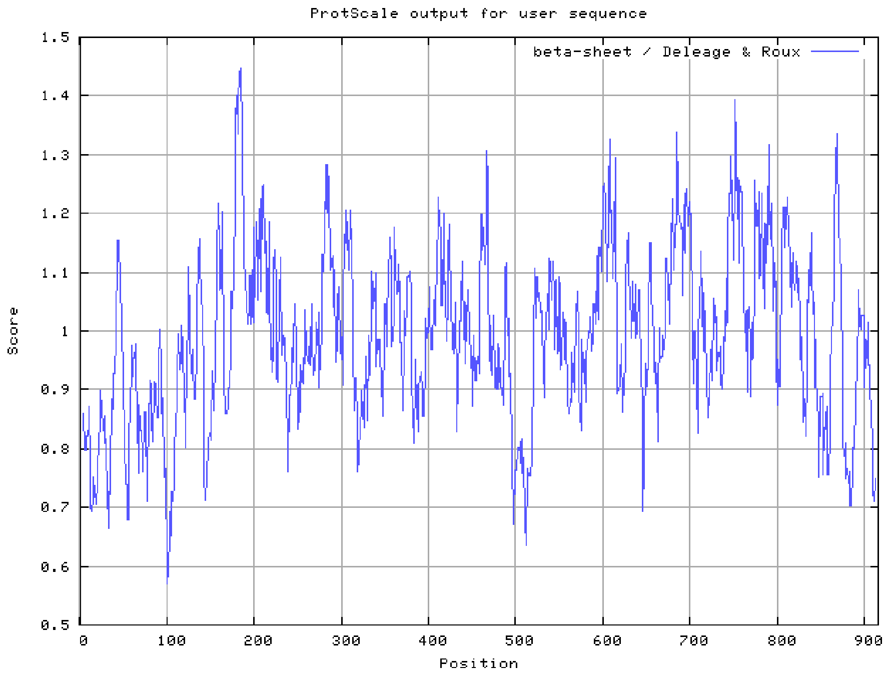


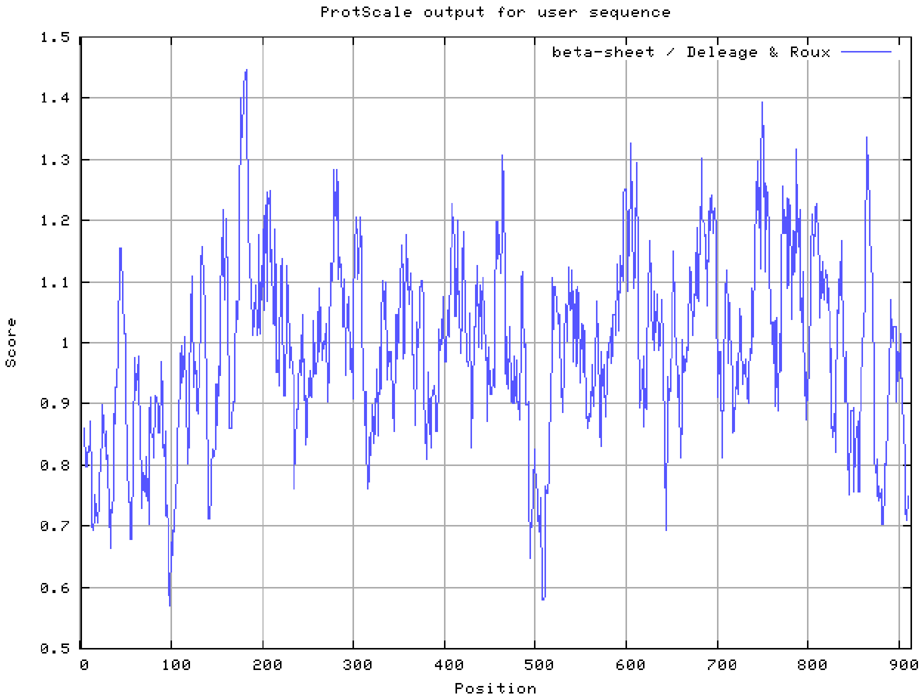


- Beta-sheet

Bartha

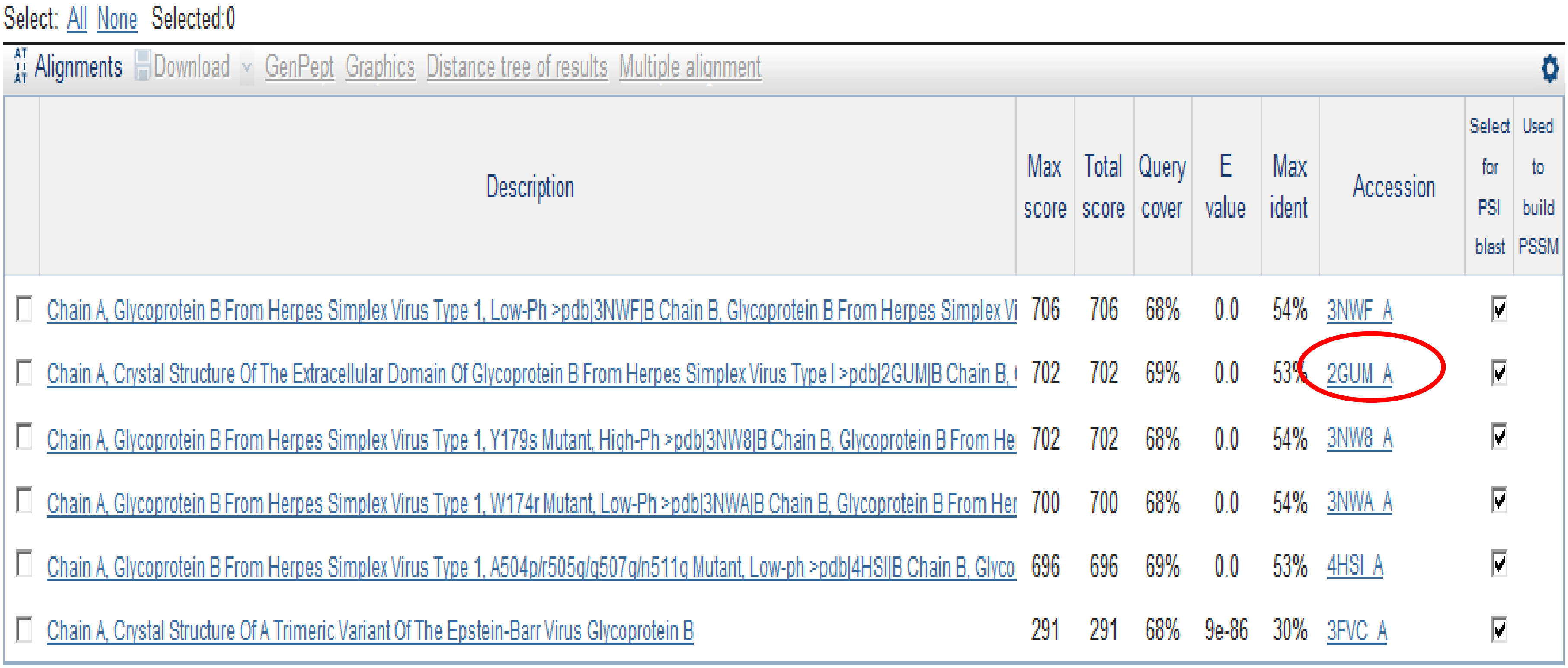
Becker





The result of Becker gB amino acid sequence Blast

- select PDB database



	Description	Max score	Total score	Query cover	E value	Max ident	Accession	Select for PSI blast	Used to build PSSM
☐	Chain A, Glycoprotein B From Herpes Simplex Virus Type 1, Low-Ph >pdb|3NWF|B Chain B, Glycoprotein B From Herpes Simplex Vi	706	706	68%	0.0	54%	3NWF_A	☑	
☐	Chain A, Crystal Structure Of The Extracellular Domain Of Glycoprotein B From Herpes Simplex Virus Type I >pdb|2GUM|B Chain B,	702	702	69%	0.0	53%	2GUM_A	☑	
☐	Chain A, Glycoprotein B From Herpes Simplex Virus Type 1, Y179s Mutant, High-Ph >pdb|3NW8|B Chain B, Glycoprotein B From He	702	702	68%	0.0	54%	3NW8_A	☑	
☐	Chain A, Glycoprotein B From Herpes Simplex Virus Type 1, W174r Mutant, Low-Ph >pdb|3NWA|B Chain B, Glycoprotein B From Her	700	700	68%	0.0	54%	3NWA_A	☑	
☐	Chain A, Glycoprotein B From Herpes Simplex Virus Type 1, A504p/r505g/g507g/n511g Mutant, Low-ph >pdb|4HSI|B Chain B, Glyco	696	696	69%	0.0	53%	4HSI_A	☑	
☐	Chain A, Crystal Structure Of A Trimeric Variant Of The Epstein-Barr Virus Glycoprotein B	291	291	68%	9e-86	30%	3FVC_A	☑	

THE TAKE HOME MESSAGE

The consecutive deletions of some bases in gB and gH gene are apparent,but the relevant analysis show that the variation may not change their corresponding proteins' nature and function.

Still, More genes need to be analysed.

The BLAST results of gB protein show that there are no high identical (>90%) sequence in the PDB database.So,relative research is to be well developed.

Thanks for your attention!