

兔出血症病毒 (RHDV) VP60蛋白 B细胞抗原 表位的预测

Prediction of the B cell epitope
for RHDV VP60 antigen

报告人：刘永相

组 员：胡 亮 秦立得 任建乐

2013.06.21

主要内容

一.研究背景

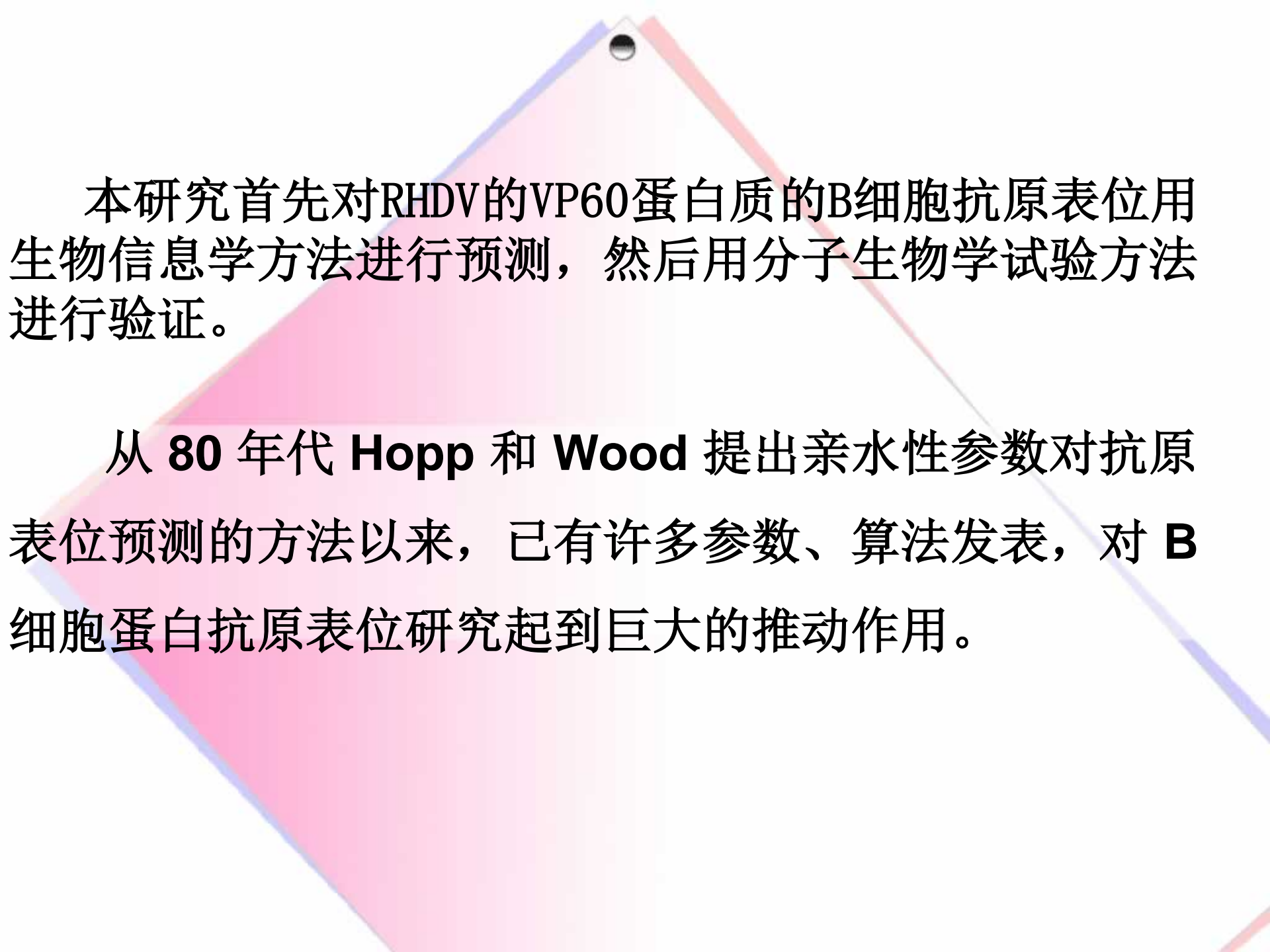
二.研究流程

三.研究结论与讨论

一 研究背景

兔出血症是由兔出血症病毒(Rabbit hemorrhagic disease virus, RHDV)引起的危害3月龄以上青年兔和成年兔的一种急性、烈性、高度接触性传染病。

VP60为RHDV的主要结构蛋白，是组成RHDV衣壳的最基本单位。大量研究结果表明VP60蛋白与病毒的致病性和免疫原性有着密切关系，是病毒的免疫保护性抗原，在诱导抗病毒感染免疫反应中起重要作用。抗原表位（**epitope**）是蛋白质抗原性的基础，是抗原分子中诱生特异性免疫应答的基本结构和功能单位。VP60结构蛋白上的抗原位点或表位的结构及定位，对于阐明RHDV的结构、功能及RHDV基因工程疫苗的研制、病原的检测至关重要。



本研究首先对RHDV的VP60蛋白质的B细胞抗原表位用生物信息学方法进行预测，然后用分子生物学试验方法进行验证。

从 **80 年代 Hopp 和 Wood** 提出亲水性参数对抗原表位预测的方法以来，已有许多参数、算法发表，对 **B 细胞蛋白抗原表位** 研究起到巨大的推动作用。

现已被大众认可并具有较好预测效果的方法主要有：

(1) 亲水性方案 (Hydrophilicity)：在机体内，疏水性残基一般埋在蛋白内部，而亲水性残基位于表面，因此蛋白的亲水部位与蛋白抗原表位有密切的联系。

(2) 可及性方案 (Accessibility)：指蛋白质抗原中氨基酸残基被溶剂分子接触的可能性。它反映了蛋白质抗原内、外各层残基的分布情况。

(3) 抗原性方案 (Antigenicity)：每个氨基酸用出现在抗原区的频率描述，此频率除以各氨基酸在所有蛋白质中的频率就可推出抗原性刻度值。

(4)可塑性方案(Flexibility): 认为蛋白抗原的多肽骨架有一定程度的活动性，这种局部的活动性与其部位的抗原性密切相关，活动性强即柔韧性大的氨基酸残基容易形成抗原表位。

(5)电荷分布方案(Charge distribution): 认为对碱性抗原特异的抗体多趋于酸性，对酸性抗原特异的抗体多趋于碱性。

(6)二级结构预测方案(Secondary structure): 认为 β 转角结构为凸出结构，多出现在蛋白质抗原表面，利于与抗体嵌合，较可能成为抗原表位。

二 研究流程

Protein

[Save search](#) [Limits](#) [Advanced](#)

[Help](#)

[Display Settings:](#) Summary, 20 per page, Sorted by Default order

[Send to:](#) **Filter your results:**

Results: 1 to 20 of 430

<< First < Prev Page 1 of 22 Next > Last >>

All (430)

Bacteria (0)

[Related Structures \(192\)](#)

[RefSeq \(2\)](#)

[Manage Filters](#)

[VP60 \[Rabbit hemorrhagic disease virus\]](#)

1. 32 aa protein

Accession: CAA51524.1 GI: 312507

[GenPept](#) [FASTA](#) [Graphics](#) [Related Sequences](#)

[VP60, partial \[Rabbit hemorrhagic disease virus\]](#)

2. 579 aa protein

Accession: AFN69440.1 GI: 395628882

[GenPept](#) [FASTA](#) [Graphics](#) [Related Sequences](#) [Identical Proteins](#)

[VP60 \[recombinant virus 6918VP60-T2\]](#)

▼ Top Organisms [\[Tree\]](#)

- Rabbit hemorrhagic disease virus (229)
- recombinant virus 6918VP60-T2 (172)

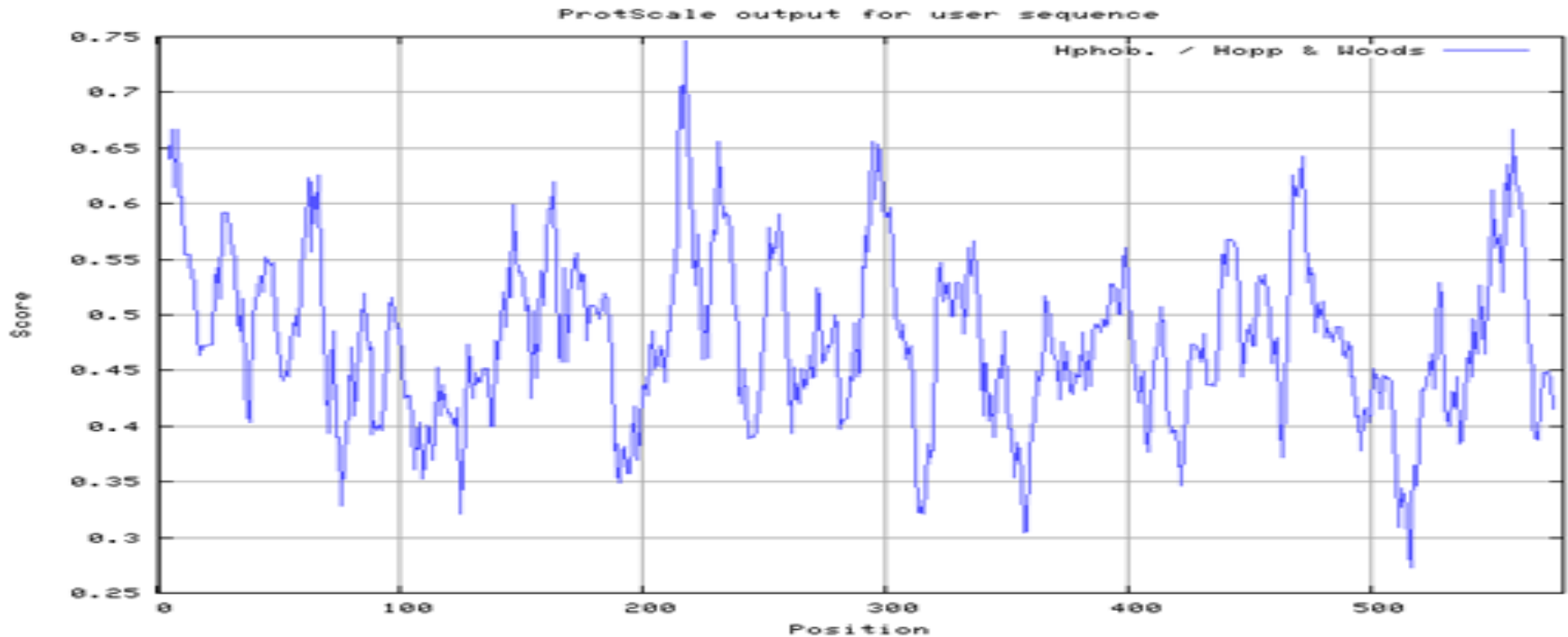
Vp60蛋白质氨基酸序列:

MEGKARTAPQGEAAGTATTASVPGTTTDGMDPGVVATTSSVVTAE NSSASI
ATAGIGGPPQQVDQQETWRTNFYYNDVFTWSVADAPGSILYTVQHSPQNN
PFTAVLSQMYAGWAGGMQFRFIVAGIGVFGGRLVAAVIPP GIEIGPGLEV
RQFP HVVIDARSLEPVTITMPDLRPNMYHPTGDPGLVPTLVLSVYNNLIN
PFGGSTSAIQVTVETRPS EDFEFVMIRAPSSKTVDSISPAGLLTTPVLTG
VGNDNRWNGQIVGLQPVPGGFSTCNRHWNLNGSTYGWSSPRFGDIGHRRG
SASYPGNNATNVLQFWYANAGSAIDNPISQVAPDGFDPMSFVPFNGPGIP
AAGWVGFGAIWNSNSGAPNVTTVQAYELGFATGAPGNLQPTTNTSGSQTV
AKSIYAVVTGTAQNPAGL FVMASGVISTPSANAITYTPQPDRIVTTPGTP
AAPVGKNTPIMFASVRRRTGDVNATAGSANGTQYGTGSQPLPVTIGLSL
NNYSSALMPGQFFVWQLTFASGFMEIGLSVDGYFYAGTGASTTLIDLTEL
IDVRPVGPRPSKSTLVFN LGGTANGFSYV

(1)亲水性方案(Hydrophilicity)

Expasy预测: [ProtScale](#) • protein profile computation and representation • [Hphob.](#) / [Hopp & Woods](#)

<http://web.expasy.org/protscale/>



结果见图,VP60 亲水性区域如图, 预测亲水性较高的为1-15, 24-30, 45-46, 48, 56-77, 94-102, 147-153, 171-185, 197, 199-200, 203-204, 213-221, 229-236, 252-260, 269, 272-311, 318, 322-323, 333-334, 336-337, 361-368, 372-373, 385-399, 411-414, 432-446, 454-458, 468-475, 478-491, 501-507, 555-564

(2) 可及性方案(Accessibility)

NetSurfP - Protein Surface Accessibility and Secondary Structure Predictions

<http://genome.cbs.dtu.dk/services/NetSurfP/>

可能暴露漏在表面的区域有:

1-7, 10-19, 21, 23, 24, 26, 28, 30, 31, 34, 37, 39-45, 49, 59-60, 63, 65, 66, 77, 79, 81, 83-88, 94, 98, 108, 118, 140-141, 143-146, 148, 149, 161-165, 167, 175, 177, 180, 182-186, 201, 203-204, 206, 219-220, 230-254, 256, 258, 260, 268-269, 277, 281, 283, 289, 291, 293-294, 301, 303, 305-310, 316, 319, 321, 325-326, 330, 340-341, 345, 347, 351-352, 358-359, 361, 364-366, 368, 372, 385-386, 389, 391, 393-396, 398-400, 402, 410, 412-413, 415, 423-425, 428, 430, 439-442, 445, 447, 455, 457, 466, 468-472, 474, 479, 481-483, 496, 500-502, 504-506, 509-510, 518-523, 540, 546, 548-550, 552, 554, 562, 564, 566, 568-569, 571-572, 574, 576-577, 579

(3) 抗原性方案 (Antigenicity)

1. Wblab预测: [antigenic\(v6.0.1\) - Finds antigenic sites in proteins](http://weblab.cbi.pku.edu.cn/program.inputForm.do?program=antigenic(v6.0.1))

[http://weblab.cbi.pku.edu.cn/program.inputForm.do?program=antigenic\(v6.0.1\)](http://weblab.cbi.pku.edu.cn/program.inputForm.do?program=antigenic(v6.0.1))

	Score	length	at residues	Sequence	Max_score_pos
1	1.213	14	184->197	PGLVPTLVLSVYNN	189
2	1.194	24	120->143	RFIVAGIGVFGGRLVAAVIPPGE	136
3	1.170	15	147->161	GLEVRQFPHWVIDAR	155
4	1.160	11	261->271	IVGLQPVPGGF	264
5	1.155	14	397->410	SQTVAKSIYAVVTG	405
6	1.149	13	32->44	PGWATTSWVTAE	37
7	1.145	20	231->250	SKTVDSISPAGLLTTPVLTG	245
8	1.139	18	79->96	TWSVADAPGSILYTVQHS	92
9	1.139	18	542->559	TTLIDLTELIDVRPVGPR	553

2.ABCpred Prediction Server (http://www.imtech.res.in/raghava/abcpred/ABC_submission.html)

Rank	Sequence	Start position	Score
1	PGSILYTVQHSPQNNP	86	0.96
1	ANAITYTPQPDRIVTT	431	0.96
1	LRPNMYHPTGDPGLVP	173	0.96
2	TGSQPLPVTIGLSLNN	487	0.94
3	SGVISTPSANAITYTP	423	0.93
4	EAAGTATTASVPGTTT	12	0.92
5	YELGFATGAPGNLQPT	376	0.91
5	GPGIPAAGWVGFGAIW	346	0.91
5	PVTITMPDLRPNMYHP	165	0.91
6	YSSALMPGQFFVWQLT	503	0.90
6	FGAIWNSNSGAPNVTT	357	0.90
6	GHRRGSASYPGNNATN	296	0.90
6	AGLLTTPVLTGVGNDN	240	0.90

3.BepiPred 1.0b Server - prediction results Technical University of Denmark

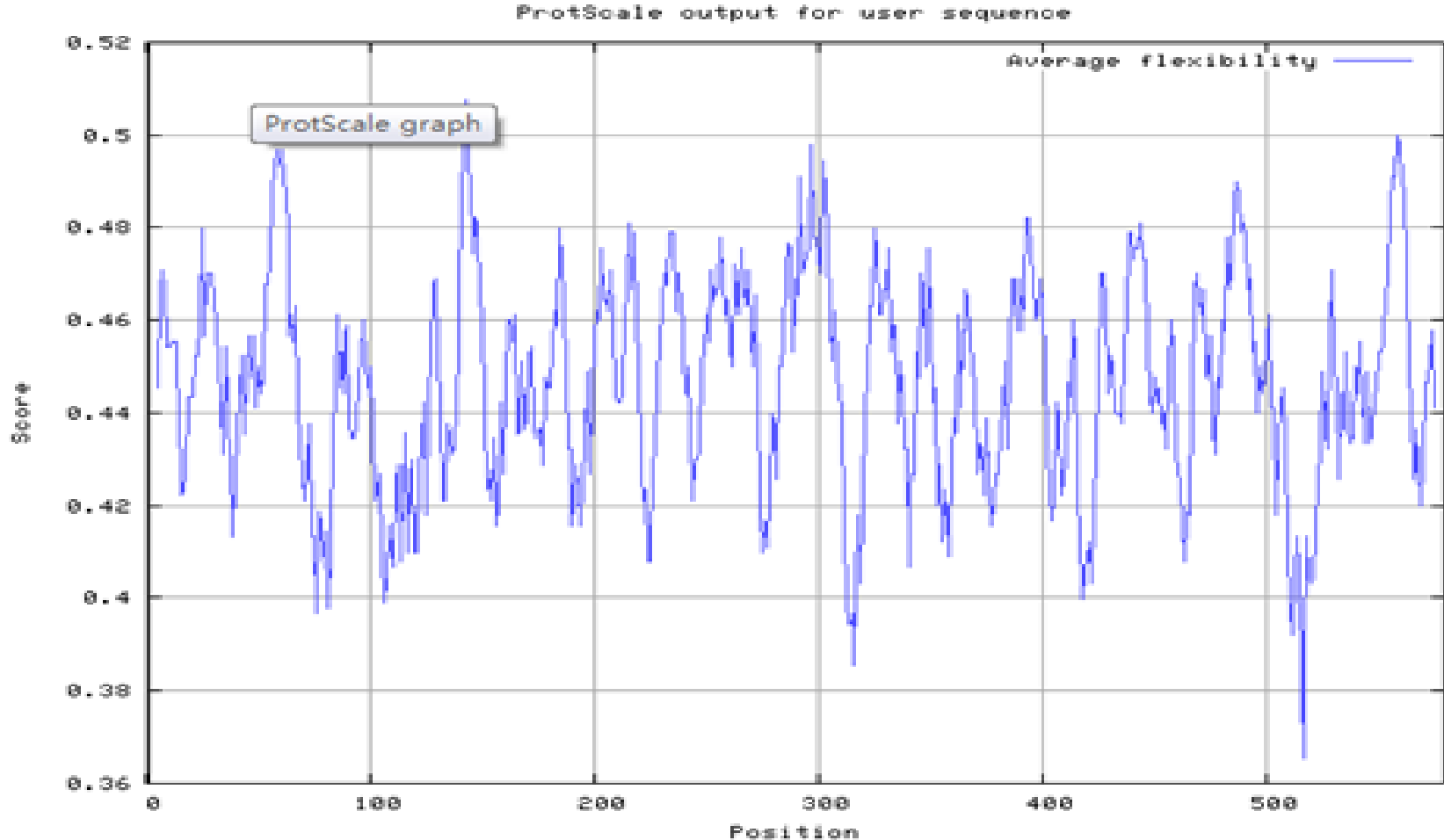


<http://www.cbs.dtu.dk/services/BepiPred/>

#	sequence	source	feature	start	end	score	N/A	?
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	1	1	-1.140	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	2	2	-0.960	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	3	3	-1.252	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	4	4	-1.055	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	5	5	-1.095	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	6	6	-2.226	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	7	7	-2.750	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	8	8	-2.961	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	9	9	-3.149	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	10	10	-3.149	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	11	11	-3.262	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	12	12	-2.765	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	13	13	-2.904	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	14	14	-1.482	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	15	15	-1.086	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	16	16	-0.518	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	17	17	0.094	.	.
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	18	18	0.780	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	19	19	1.013	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	20	20	1.221	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	21	21	1.006	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	22	22	1.047	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	23	23	0.817	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	24	24	1.004	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	25	25	1.078	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	26	26	0.928	.	弱
ep_001233	CGHS_HUMAN	BepiPred-1.0b	epitope	27	27	0.211	.	弱

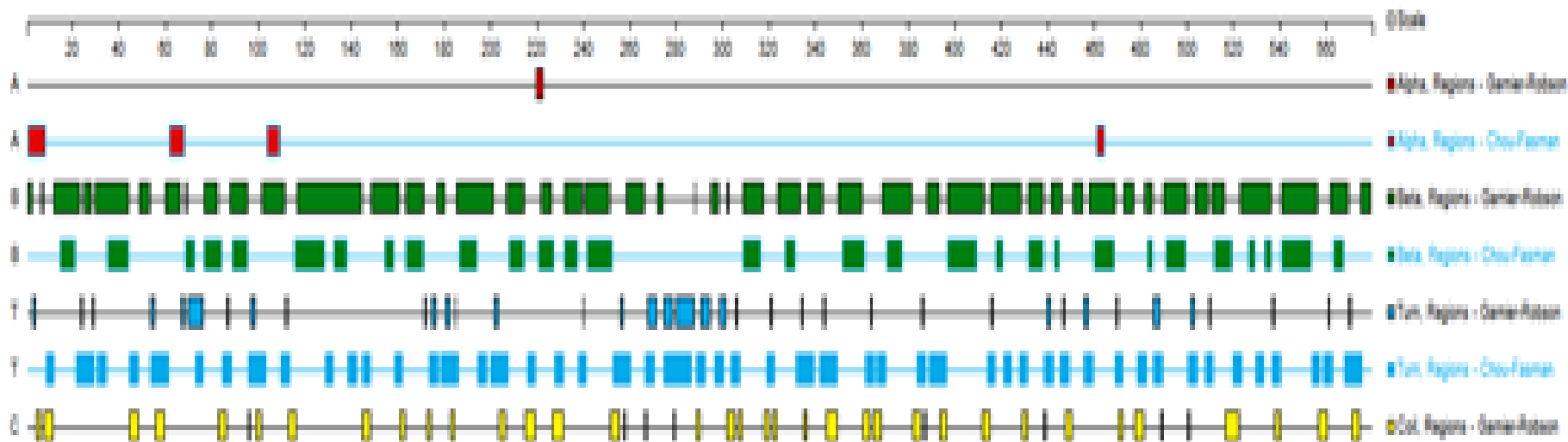
预测的抗原表位: 1-36, 40-68, 82-88, 95-102, 139-149, 172-186, 201-208, 213-220, 229-238, 248-258, 265-273, 282-310, 321-338, 344-352, 362-373, 381-400, 410-415, 426-458, 470-493, 535-542, 555-562, 571-575.

(4)可塑性方案(Flexibility)



可塑性较高的区域: 6-9, 24-25, 27-31, 55-63, 127-128, 140-150, 182, 184-187, 205-208, 213-219, 229-239, 255-259, 261-269, 273, 284-287, 289-296, 307, 323-332, 346, 362, 366-368, 386-387, 389-390, 392-397, 426-428, 437-446, 467-473, 482-494, 528-529, 553-563

(5) 二级结构预测方案 (Secondary structure)



β 转角结构

45-48, 56-72, 83-90, 96-100, 115, 125-130, 139-155, 161-165, 171-178, 182-187, 196-206, 219-221, 228-238, 242-245, 250-252, 256-259, 268-294, 298-300, 306-312, 321-337, 341-352, 362-373, 382-396, 405-416, 422-424-, 429-441, 445-458, 467, 471-485, 489, 501-509, 521, 530-532, 537-544, 548, 562-565,

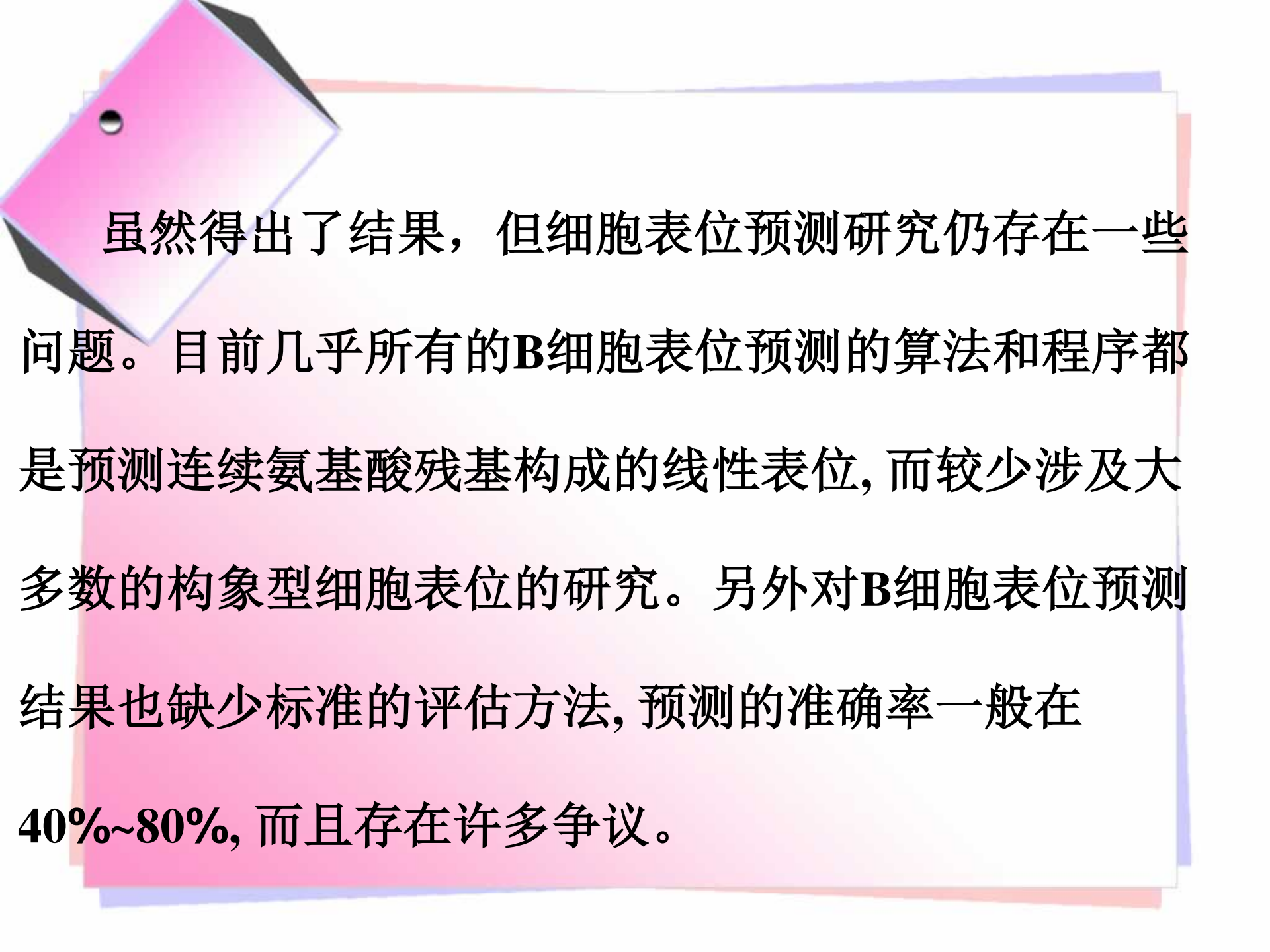
三 研究结论与讨论

通过以上的方案综合预测的潜在的抗原表位区域为：

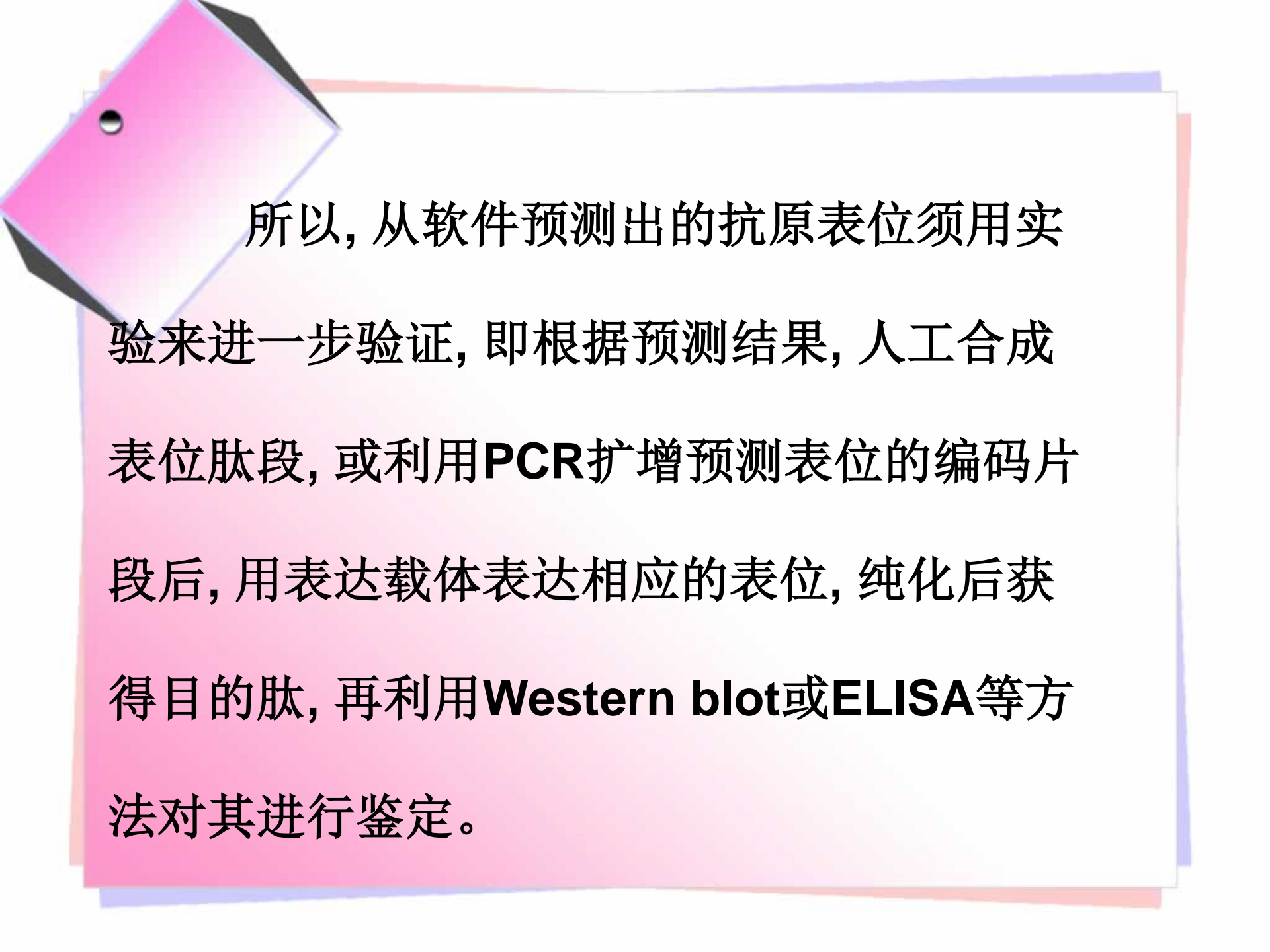
184-197,120-143, 261-271,397-410,32-44,231-

250,79-96,542-559,207-215,462-470,325-334,502-


520562-569,450-458,58-64



虽然得出了结果，但细胞表位预测研究仍存在一些
问题。目前几乎所有的B细胞表位预测的算法和程序都
是预测连续氨基酸残基构成的线性表位，而较少涉及大
多数的构象型细胞表位的研究。另外对B细胞表位预测
结果也缺少标准的评估方法，预测的准确率一般在
40%~80%，而且存在许多争议。



所以,从软件预测出的抗原表位须用实验来进一步验证,即根据预测结果,人工合成表位肽段,或利用**PCR**扩增预测表位的编码片段后,用表达载体表达相应的表位,纯化后获得目的肽,再利用**Western blot**或**ELISA**等方法对其进行鉴定。



由于时间仓促以及本团队能力有限，
以上报告中难免有不周之处，欢迎大家批
评指正，同时由于有大量的数据需要处理，
抗原表位的位点预测不够准确精确，我们
会进一步努力，争取获得更加准确的预测
位点。

致谢

感谢罗老师的谆谆教导，言传身教。

感谢罗老师的ABC网站给大家提供的便利，
希望N年后依然看到ABC。



谢 谢