DOI:10.16016/j.1000-5404.2004.12.013
第26卷第12期 第三军医大学学报 Vol. 26, No. 12
2004年6月 ACTA ACADEMIAE MEDICINAE MILITARIS TERTIAE Jun. 2004

文章编号:1000-5404(2004)12-1065-03

论著

SARS冠状病毒E蛋白HLA-A*0201限制性CTL表位的预测

王 晴,吴玉章 (第三军医大学基础医学部全军免疫学研究所,重庆 400038)

提 要:目的 预测SARS冠状病毒E蛋白的HLA-A*0201限制性细胞毒性T淋巴细胞(CTL)表位。方法 联合应用超基序法和量化矩阵法。结果 预测出13个九肽CTL表位。结论 通过对SARS冠状病毒E蛋白抗原CTL表位进行预测,从而为SARS冠状病毒E蛋白CTL表位的实验探测和鉴定提供了线索,为认识CTL介导的细胞免疫机制以及基于CTL表位的疫苗研制提供了基础资料。

关键词:E蛋白;HLA;CTL表位;超基序;量化矩阵
中图法分类号:R373.1;R392.13 文献标识码:A

Prediction of HLA-A*0201-restricted CTL epitopes in SARS-CoV E protein

WANG Qing, WU Yu-zhang (Institute of Immunology, College of Medicine, Third Military Medical University, Chongqing 400038, China)

Abstract: Objective To predict the HLA-A*0201-restricted CTL epitopes (nonamers) in SARS coronavirus (SARS-CoV) E protein. Methods The CTL epitopes of E protein were predicted by using supermotif method combined with quantitative matrix method. Results Thirteen HLA-A*0201-restricted CTL epitope candidates were predicted in SARS-CoV E protein. Conclusion The prediction of the CTL epitopes in SARS-CoV E protein will benefit the identification of CTL epitopes by experiment. Those results are of importance for immune recognition and vaccine design for SARS-CoV.

Key words: E protein; HLA; CTL epitope; supermotif; quantitative matrix

严重急性呼吸综合征(severe acute respiratory syndrome, SARS)是一种突发的严重传染性疾病,短期内在多个国家均有感染及死亡报道,严重威胁人们的生命健康。其病原体为SARS冠状病毒(SARS coronavirus, SARS-CoV)。目前尚无疫苗应用。研制基于细胞毒性T淋巴细胞(cytotoxic T lymphocyte, CTL)表位的SARS冠状病毒疫苗,通过应用疫苗使机体产生CTL介导的细胞免疫反应,可实现对该病预防和治疗的目的。对于一个已知序列的蛋白质抗原,其CTL表位的确定通常采用"反向免疫学(reverse immunology)"方法,先通过CTL表位的预测使待选多肽的范围缩小,再用相关实验来鉴定。E蛋白(envelope protein, E protein)是SARS冠状病毒最小的结构蛋白,其散在分布于病毒包膜上,在冠状病毒的自我组装过程中起至关重要的作用。本研究拟对其CTL表位进行预测。由于HLA-A*0201阳性人群的比例很高[1],故HLA-A*0201限制性CTL表位在疫苗设计中更具应用价值。已知MHC-Ⅰ类分子所结合的抗原肽一般由8~12个氨基酸组成,大多数为九肽[2]。故本研究联合应用超基序(supermotif)法和量化矩阵(quantitative matrix)法,预测由9个氨基酸残基组成的HLA-A*0201限制性CTL表位。

1 材料和方法

1.1 目的抗原

SARS冠状病毒E蛋白为由76个氨基酸组成的蛋白质,其氨基酸顺序见图1:

```
          10        20        30        40        50
MYSFVSEETG TLIVNSVLIF LAFVVFLLVT LAILTALRLC AYCCNIVNVS
          60        70
LVKPTVYVYS RVKNLNSSEG VPDLLV
```

图1 SARS冠状病毒E蛋白的氨基酸顺序

Fig 1 Amino acid sequence of SARS-CoV E protein

1.2 CTL表位预测

1.2.1 超基序法 许多不同等位基因甚至不同血清型的HLA-Ⅰ类分子所结合的抗原肽具有相同或相似的锚着残基(anchor residue),这些HLA分子总称为HLA超型(HLA super-

基金项目:国家重点基础研究发展规划资助项目(2003CB514108)
Supported by the National Key Basic Research and Development Plan of China (2003CB514108)
作者简介:王 晴(1970-),女,重庆市人,硕士,讲师,主要从事免疫学基础方面的研究。电话:(023)68752238-801,E-mail:immuno10@mail.tmmu.com.cn
通信作者:吴玉章,电话:(023)68752680;E-mail:wuyuzhang@yahoo.com
收稿日期:2004-02-24;修回日期:2004-05-15

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

type)；而相应的与HLA超型结合的抗原肽所共同具有的氨基酸序列特征称为超基序。目前实验证实或经理论推测的HLA超型有A2、A3、B7、B44、A1、A24、B27、B58、B62、DR等。HLA-A*0201属于HLA-A2超型。HLA-A2超型至少由血清型为A2(等位基因为A*0201、A*0202、A*0203、A*0204、A*0205、A*0206、A*0207)与A28(A*6802、A*6901)的两种HLA抗原共9种HLA等位基因分子组成。HLA-A2超型所结合抗原肽的超基序为第2位与羧基末端(第9位)氨基酸残基均为疏水的脂肪族或小氨基酸：亮氨酸(Leucine, L)、异亮氨酸(Isoleucine, I)、缬氨酸(Valine, V)、蛋氨酸(Methionine, M)、丙氨酸(Alanine, A)或苏氨酸(Threonine, T)[3,4]。

1.2.2 量化矩阵法 在抗原肽与MHC-Ⅰ类分子的结合中，抗原肽的侧链效应(锚定、抑制或中性)主要依赖于相应氨基酸残基在肽中所处的位置，而受邻近氨基酸残基的影响较小。据此人们提出了“非依赖性侧链结合(independent binding of side chains, IBS)”假设，即在一个抗原肽序列中，每个氨基酸残基相互独立地以一定的结合能影响该肽与MHC-Ⅰ类分子的亲和力，而抗原肽与MHC-Ⅰ类分子的结合是该肽中每个位置上的氨基酸残基结合能的综合[5~9]。量化矩阵法的原理则是基于该假设。在此假设前提下，Parker等[8]用多肽结合实验分析了154个九肽与HLA-A*0201分子的亲和力，并得出一个包含180个系数的矩阵(20 amino acids× 9 positions)(表1)，其中每个系数反映了相应氨基酸在该位置时与HLA-A*0201分子的结合能。对于某待定九肽，各位置氨基酸相应系数的乘积再乘以标准系数0.151，即为该肽与HLA-A*0201分子的“理论”结合稳定性。若该值大于或等于5，则认为该肽与HLA-A*0201分子结合稳定；反之则认为结合不稳定。本研究对超基序法预测出的CTL表位进一步采用量化矩阵法进行筛选，以提高预测的准确率。

表1 HLA-A*0201分子限制性九肽结合系数

Tab 1 Nonapeptides binding coefficients restricted by HLA-A*0201

Amino acid	Residue position within the peptide								
	1	2	3	4	5	6	7	8	9
A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
C	0.597	0.500	1.000	1.000	1.000	1.000	1.000	1.000	0.010
D	0.041	0.500	0.726	1.000	1.000	1.000	1.000	1.000	0.010
E	0.578	2.840	0.054	1.000	0.756	3.246	0.820	1.000	0.015
F	1.000	0.500	11.383	0.027	6.004	4.369	6.383	1.000	0.009
G	0.578	0.500	0.088	1.000	0.804	1.060	0.105	1.000	0.009
H	0.044	0.500	1.000	1.000	1.000	1.000	1.000	1.000	0.010
I	1.000	10.151	1.849	0.078	1.000	1.000	1.000	1.000	0.534
K	3.465	20.524	0.024	1.000	2.085	0.239	0.603	1.000	0.015
L	1.000	103.183	3.685	0.646	1.000	1.000	1.000	1.000	2.357
M	1.000	57.920	1.000	1.000	1.000	1.000	1.000	1.000	1.501
N	1.000	0.542	1.000	1.000	1.000	1.000	1.000	1.000	0.009
P	1.000	0.500	1.261	1.000	1.000	1.000	1.000	1.000	0.010
Q	1.000	10.006	1.000	1.000	1.000	1.000	1.000	1.000	0.010
R	1.000	0.500	0.033	1.000	1.000	1.000	0.277	1.000	0.010
S	1.000	0.500	1.000	1.000	1.000	1.000	1.000	1.000	0.009
T	1.000	6.080	1.000	1.000	1.000	1.000	1.000	1.000	0.009
V	1.000	5.919	2.173	1.000	1.000	2.588	1.120	0.293	4.884
W	1.000	0.500	12.978	1.000	2.680	0.251	5.951	1.000	0.010
Y	1.000	0.500	7.613	1.000	8.002	3.470	1.000	1.000	0.009

2 结果

2.1 超基序法预测的结果

采用超基序法对SARS冠状病毒E蛋白抗原进行CTL表位预测，预测出21个候选表位，分别是：FVSEETGTL(4~12)、GTLIVNSVL(10~18)、TLIVNSVLL(11~19)、IVNSVLLFL(13~21)、SVLLFLAFV(16~24)、VLLFLAFVV(17~25)、FLAFVVFLL(20~28)、LAFVVFLLV(21~29)、FVVFLLVTL(23~31)、VVFLLVTLA(24~32)、FLLVTLAIL(26~34)、LLVTLAILT(27~35)、LVTLAILTA(28~36)、VTLAILTAL(29~37)、LAILTALRL(31~39)、ILTALRLCA(33~41)、RLCAYCCNI(38~46)、NVSLVKPTV(48~56)、SLVKPTVYV(50~58)、PTVYVYSRV(54~62)、YVYSRVKNL(57~65)。

2.2 量化矩阵法预测的结果

以上述超基序法预测出的21个候选表位为目标，用量化矩阵法从中筛选出13个候选九肽，分别是：SLVKPTVYV(50~58)(185.20)、FLLVTLAIL(26~34)(135.33)、TLIVNSVLL(11~19)(76.05)、RLCAYCCNI(38~46)(66.58)、SVLLFLAFV(16~24)(62.81)、LAFVVFLLV(21~29)(36.68)、PTVYVYSRV(54~62)(33.81)、YVYSRVKNL(57~65)(25.03)、VVFLLVTLA(24~32)(17.01)、FLAFVVFLL(20~28)(16.38)、ILTALRLCA(33~41)(15.58)、VLLFLAFVV(17~25)(14.15)、VTLAILTAL(7.97)。

3 讨论

人们已认识到蛋白质抗原并非通过其完整分子发挥功能，而是通过其表位体现其特异性。另一方面，由于许多天然抗原引发的免疫反应不能满足预防感染或预防发病的需要，为了提高蛋白质抗原的保护性，就必须在表位水平上做出选择、改造以得到更理想的疫苗分子。而这一切都必须建立在表位筛选、鉴定的基础上。CTL表位的快速筛选和鉴定是研制基于CTL表位疫苗的关键。对于一个已知序列的蛋白质抗原，其CTL表位的确定通常先通过表位预测方法使待选多肽的范围缩小，再用相关实验来鉴定。目前，关于CTL表位的预测方法有很多，如简单基序法、延展基序法、超基序法、量化矩阵法、人工神经网络算法以及基于结构的一些方法等。本研究采用其中的两种方法：超基序法和量化矩阵法。

早期人们认识到各种不同类型MHC-Ⅰ类分子结合抗原肽的锚着残基不同。对于某型别的MHC-Ⅰ类分子，它所结合抗原肽的锚着残基常为某一种或与之接近的几种氨基酸残基。Falk及Ruppert等[10,11]采用肽池测序法和氨基酸置换实验，发现大部分与HLA-A*0201分子有较高亲和力的九肽有两个主要锚着残基(primary anchor residue)：第2位和第9位氨基酸残

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

基, 第 2 位氨基酸残基为 L 或 M, 第 9 位为 L、V 或 I。现在认为, 上述基序只是 MHC-Ⅰ类分子与抗原肽复杂的相互作用中的一种简单规律, 故称为简单基序(simple motif)。基于此建立的表位预测法准确率为 30%左右, 其漏检率及假阳性率均较高。以后人们发现许多不同等位基因甚至不同血清型的 HLA-Ⅰ类分子所结合的抗原肽具有相同或相似的锚着残基, 这种共同具有的氨基酸序列特征称为超基序。HLA-A *0201 属于 HLA-A2 超型。HLA-A2 超型所结合抗原肽的超基序为第 2 位与第 9 位氨基酸残基均为 L、I、V、M、A 或 T[3,4]。基于此建立的表位预测法大大降低了简单基序法的漏检率, 但由于其仅仅考虑了第 2 位、第 9 位氨基酸残基的影响, 而忽略了其他位点的残基, 因此其保留了简单基序法假阳性率较高的缺点。实验资料表明大多数已知的 CTL 表位以高亲和力与 MHC-Ⅰ类分子结合, 因此可通过亲和力来预测 CTL 表位[5,8,9,11,12]。而且最好不仅要考虑到第 2 位、第 9 位主要锚着残基与 MHC-Ⅰ类分子的结合情况, 还要考虑到其他次级锚着残基(secondary anchor residue)与 MHC-Ⅰ类分子的结合情况。而最简单的方法就是假定每个氨基酸残基相互独立地以一定的结合能影响该肽与 MHC-Ⅰ类分子的亲和力(即 IBS 假设)。在此假设前提下, Parker 等[8]通过多肽结合实验得到了一个包含 180 个系数的 HLA-A *0201 限制性矩阵, 其中每个系数反映了相应氨基酸在该位置时与 HLA-A *0201 分子的结合能。某待定九肽各位置氨基酸相应系数相乘, 即得到一个所映两者亲和力的数值, 借此便可判断该多肽是否可能为 CTL 表位。基于此建立的预测方法即为量化矩阵法。用该方法进行预测, 已知的天然提呈的 CTL 表位通常都在预测值最高的 2%预测九肽之内, 从而证实该方法是可行的、有效的。该量化矩阵法能对每一个位点氨基酸残基的结合力进行量化分析, 从而较之超基序法大大提高了预测的准确率。故本研究采用超基序法和量化矩阵法相结合进行预测。

本研究首先采用超基序法对所有可能为 HLA-A *0201限制性的 CTL 表位进行初步预测, 发现符合的有 21 个。然后再采用量化矩阵法进一步筛选, 最后预测出 13 个九肽表位。因为用该量化矩阵法进行预测, 已知的天然提呈的 CTL 表位通常都在预测值最高的 2%预测九肽之内, 而本研究中被预测的目的抗原 E 蛋白由 76 个氨基酸组成(其所有可能的九肽则为 68 个), 所以预测值最高的前 2 个九肽(SLVKPTVYV、FLLVTLAIL)最有可能为真正的 CTL 表位。通过对 SARS 冠状病毒 E 蛋白抗原 CTL 表位进行预测, 从而为 SARS 冠状病毒 E 蛋白 CTL 表位的实验探测和鉴定提供了线索, 为认识 CTL 介导的细胞免疫机制以及基于 CTL 表位的疫苗研制提供了基础资料。

参考文献:

[1] Gulukota K, Delisi C. HLA allele selection for designing peptide vaccines [J]. Genet Anal, 1996, 13(3): 81－86.

[2] Schirle M, Weinschenk T, Stevanovic S. Combining computer algorithms with experimental approaches permits the rapid and accurate identification of T cell epitopes from defined antigens[J]. J Immunol Methods, 2001, 257(1-2): 1－16.

[3] Sette A, Sidney J. HLA supertypes and supermotifs: a functional perspective on HLA polymorphism[J]. Curr Opin Immunol, 1998, 10(4): 478－482.

[4] Sette A, Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism[J]. Immunogenetics, 1999, 50(3-4): 201－212.

[5] Hammer J, Belunis C, Bolin D, *et al*. High- affinity binding of short peptides to major histocompatibility complex class Ⅱ molecules by anchor combinations[J]. Proc Natl Acad Sci USA, 1994, 91(10): 4456－4460.

[6] Hammer J, Bono E, Gallazzi F, *et al*. Precise prediction of major histocompatibility complex class Ⅱ-peptide interaction based on peptide side chain scanning[J]. J Exp Med, 1994, 180(6): 2353－2358.

[7] Hammer J, Gallazzi F, Bono E, *et al*. Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association[J]. J Exp Med, 1995, 181(5): 1847－1855.

[8] Parker K C, Bednarek M A, Coligan J E. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains[J]. J Immunol, 1994, 152(1): 163－175.

[9] Reay P A, Kantor R M, Davis M M. Use of global amino acid replacements to define the requirements for MHC binding and T cell recognition of moth cytochrome c(93-103)[J]. J Immunol, 1994, 152(8): 3946－3957.

[10] Falk K, Rotzschke O, Stevanovic S, *et al*. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules[J]. Nature, 1991, 351(6324): 290－296.

[11] Ruppert J, Sidney J, Celis E, *et al*. Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules[J]. Cell, 1993, 74(5): 929－937.

[12] Chen W, Khilko S, Fecondo J, *et al*. Determinant selection of major histocompatibility complex class I-restricted antigenic peptides is explained by class I-peptide affinity and is strongly influenced by nondominant anchor residues[J]. J Exp Med, 1994, 180(4): 1471－1483.

(编辑　薛国文)

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net