

[19]中华人民共和国国家知识产权局

[51]Int. Cl<sup>6</sup>

# [12] 发明专利申请公开说明书

A61K 38/03  
A61K 39/002 A61K 39/02  
A61K 39/12

[21] 申请号 97194430.X

[43]公开日 1999年8月11日

[11]公开号 CN 1225590A

[22]申请日 97.3.10 [21]申请号 97194430.X

[30]优先权

[32]96.3.11 [33]US[31]60/013,113

[86]国际申请 PCT/US97/03778 97.3.10

[87]国际公布 WO97/33602 英 97.9.18

[85]进入国家阶段日期 98.11.6

[71]申请人 埃皮曼尼公司

地址 美国加利福尼亚州

[72]发明人 A·瑟特 R·W·彻斯特鲁特

J·悉尼

[74]专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 齐曾度

权利要求书 5 页 说明书 48 页 附图页数 10 页

[54]发明名称 对 HLA 分子结合亲和力提高的肽

[57]摘要

本发明提供包含 HLA - A3 超基元的免疫原性肽。  
此肽可用于治疗、诊断或检测许多病理状态。

I S S N 1 0 0 8 - 4 2 7 4

## 权 利 要 求 书

1. 一种诱导病人对预选的抗原发生细胞毒 T 细胞反应的方法，此方法包括将细胞毒 T 细胞与约 9 至 15 个残基的免疫原性肽接触，这些肽以小于约 500nM 的解离常数结合至少两种 HLA-A3 样分子，并诱导细胞毒 T 细胞反应，免疫原性肽具有含有一个结合基元的 9 残基序列，从 N 末端至 C 末端如下：

位于第 2 位的第 1 个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第 2 个初级锚定残基，选自 R 和 K；和

一个或多个次级锚定残基，选自第 3 位的 Y、F 或 W，第 6 位的 Y、F 或 W，第 7 位 Y、F 或 W，第 8 位的 P，和任何这些的组合。

2. 权利要求 1 的方法，其中肽由 9 至 10 个残基组成。

3. 权利要求 1 的方法，其中肽来自病毒抗原、肿瘤相关抗原、寄生虫抗原，或真菌抗原。

4. 权利要求 1 的方法，其中接触步骤在体外进行。

5. 权利要求 1 的方法，其中接触步骤通过向病人给予包含编码此免疫原性肽的序列的核酸分子而进行。

6. 一种诱导病人对预选的抗原发生细胞毒 T 细胞反应的方法，此方法包括将细胞毒 T 细胞与约 9 至 15 个残基的免疫原性肽接触，这些肽以小于约 500nM 的解离常数结合 HLA-A\*0301 基因产物，并诱导细胞毒 T 细胞反应，免疫原性肽具有含有一个结合基元的 9 残基序列，从 N 末端至 C 末端如下：

位于第 2 位的第 1 个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第 2 个初级锚定残基，选自 R 和 K；和

一个或多个次级锚定残基，选自第 1 位的 R、H 或 K，第 3 位的 Y、F 或 W，第 4 位的 P、R、H、K、Y、F 或 W，第 5 位的 A，第 6 位的 Y、F 或 W，第 8 位的 P，和任何这些的组合。

7. 权利要求 6 的方法，其中肽由 9 至 10 个残基组成。

8. 权利要求 6 的方法，其中肽来自病毒抗原、肿瘤相关抗原、寄生虫抗原，或真菌抗原。

9. 权利要求 6 的方法，其中接触步骤在体外进行。

10. 权利要求 6 的方法，其中接触步骤通过向病人给予包含编码此免疫原性肽的序列的核酸分子而进行。



11. 一种诱导病人对预选的抗原发生细胞毒 T 细胞反应的方法，此方法包括将细胞毒 T 细胞与约 9 至 15 个残基的免疫原性肽接触，这些肽以小于约 500nM 的解离常数结合 HLA - A\*1101 基因产物，并诱导细胞毒 T 细胞反应，免疫原性肽具有含有一个结合基元的 9 残基序列，从 N 末端至 C 末端如下：

位于第 2 位的第 1 个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第 2 个初级锚定残基，选自 R 和 K；

一个次级锚定残基，选自第 1 位的 A，第 3 位的 Y、F 或 W，第 4 位的 Y、F 或 W，第 5 位的 A，第 6 位的 Y、F 或 W，第 7 位的 Y、F 或 W，第 8 位的 P，和任何这些的组合。

12. 权利要求 11 的方法，其中肽由 9 至 10 个残基组成。

13. 权利要求 11 的方法，其中肽来自病毒抗原、肿瘤相关抗原、寄生虫抗原，或真菌抗原。

14. 权利要求 11 的方法，其中接触步骤在体外进行。

15. 权利要求 11 的方法，其中接触步骤通过向病人给予包含编码此免疫原性肽的序列的核酸分子而进行。

16. 一种诱导病人对预选的抗原发生细胞毒 T 细胞反应的方法，此方法包括将细胞毒 T 细胞与约 9 至 15 个残基的免疫原性肽接触，这些肽以小于约 500nM 的解离常数结合 HLA - A\*3101 基因产物，并诱导细胞毒 T 细胞反应，免疫原性肽含有 9 到约 15 个残基，具有含有一个结合基元的 9 残基序列，从 N 末端至 C 末端如下：

位于第 2 位的第 1 个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第 2 个初级锚定残基，选自 R 和 K；

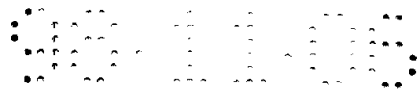
一个次级锚定残基，选自第 1 位的 R、H 或 K，第 3 位的 Y、F 或 W，第 4 位的 P，第 6 位的 Y、F 或 W，第 7 位的 Y、F 或 W，或第 8 位的 A 或 P，和任何这些的组合。

17. 权利要求 16 的方法，其中肽由 9 至 10 个残基组成。

18. 权利要求 16 的方法，其中肽来自病毒抗原、肿瘤相关抗原、寄生虫抗原，或真菌抗原。

19. 权利要求 16 的方法，其中接触步骤在体外进行。

20. 权利要求 16 的方法，其中接触步骤通过向病人给予包含编码此免疫原性肽的序列的核酸分子而进行。



21. 一种诱导病人对预选的抗原发生细胞毒 T 细胞反应的方法，此方法包括将细胞毒 T 细胞与约 9 至 15 个残基的免疫原性肽接触，这些肽以小于约 500nM 的解离常数结合 HLA - A\*3301 基因产物，并诱导细胞毒 T 细胞反应，免疫原性肽包括含有一个结合基元的 9 残基序列，  
5 从 N 末端至 C 末端如下：  
    位于第 2 位的第 1 个初级锚定残基，选自 A、L、I、V、M、S 和 T，  
    位于第 9 位的第 2 个初级锚定残基，选自 R 和 K；  
    一个次级锚定残基，选自第 3 位的 Y、F 或 W，第 7 位的 A、Y、F 或 W，和任何这些的组合。
22. 权利要求 21 的方法，其中肽由 9 至 10 个残基组成。
23. 权利要求 21 的方法，其中肽来自病毒抗原、肿瘤相关抗原、  
10 寄生虫抗原，或真菌抗原。
24. 权利要求 21 的方法，其中接触步骤在体外进行。
25. 权利要求 21 的方法，其中接触步骤通过向病人给予包含编码  
15 此免疫原性肽的序列的核酸分子而进行。
26. 一种诱导病人对预选的抗原发生细胞毒 T 细胞反应的方法，此方法包括将细胞毒 T 细胞与约 9 至 15 个残基的免疫原性肽接触，这些肽以小于约 500nM 的解离常数结合 HLA - A\*6801 基因产物，并诱导细胞毒 T 细胞反应，免疫原性肽包括含有一个结合基元的 9 残基序列，  
20 从 N 末端至 C 末端如下：  
    位于第 2 位的第 1 个初级锚定残基，选自 A、L、I、V、M、S 和 T，  
    位于第 9 位的第 2 个初级锚定残基，选自 R 和 K；  
    一个次级锚定残基，选自第 1 位的 Y、F、W、S、T 或 C，第 5 位的 Y、F、W、L、I、V 或 M，第 7 位的 Y、F 或 W，第 8 位的 P，和任何  
25 这些的组合。
27. 权利要求 26 的方法，其中肽由 9 至 10 个残基组成。
28. 权利要求 26 的方法，其中肽来自病毒抗原、肿瘤相关抗原、  
寄生虫抗原，或真菌抗原。
29. 权利要求 26 的方法，其中接触步骤在体外进行。
30. 权利要求 26 的方法，其中接触步骤通过向病人给予包含编码  
30 此免疫原性肽的序列的核酸分子而进行。
31. 一种包含约 9 至 15 个残基的免疫原性肽的组合物，这些肽以



小于约 500nM 的解离常数结合至少两个 HLA - A3 样分子, 并诱导细胞毒 T 细胞反应, 免疫原性肽具有包含一个结合基元的 9 残基序列, 从 N 末端到 C 末端如下:

位于第 2 位的第 1 个初级锚定残基, 选自 A、L、I、V、M、S 和 T,  
5 位于第 9 位的第 2 个初级锚定残基, 选自 R 和 K;

一个或多个次级锚定残基, 选自第 3 位的 Y、F 或 W, 第 6 位的 Y、F 或 W, 第 7 位的 Y、F 或 W, 第 8 位的 P, 和任何这些的组合。

32. 一种包含约 9 至 15 个残基的免疫原性肽的组合物, 这些肽以  
10 小于约 500nM 的解离常数结合 HLA - A\*0301 基因产物, 并诱导细胞毒 T 细胞反应, 免疫原性肽具有包含一个结合基元的 9 残基序列, 从 N 末端到 C 末端如下:

位于第 2 位的第 1 个初级锚定残基, 选自 A、L、I、V、M、S 和 T,  
位于第 9 位的第 2 个初级锚定残基, 选自 R 和 K;

15 一个或多个次级锚定残基, 选自第 1 位的 R、H 或 K, 第 3 位的 Y、F 或 W, 第 4 位的 P、R、H、K、Y、F 或 W, 第 5 位的 A, 第 6 位的 Y、F 或 W, 第 8 位的 P, 和任何这些的组合。

33. 一种包含约 9 至 15 个残基的免疫原性肽的组合物, 这些肽以  
20 小于约 500nM 的解离常数结合 HLA - A\*1101 基因产物, 并诱导细胞毒 T 细胞反应, 免疫原性肽具有包含一个结合基元的 9 残基序列, 从 N 末端到 C 末端如下:

位于第 2 位的第 1 个初级锚定残基, 选自 A、L、I、V、M、S 和 T,  
位于第 9 位的第 2 个初级锚定残基, 选自 R 和 K;

25 一个次级锚定残基, 选自第 1 位的 A, 第 3 位的 Y、F 或 W, 第 4 位的 Y、F 或 W, 第 5 位的 A, 第 6 位的 Y、F 或 W, 第 7 位的 Y、F 或 W, 第 8 位的 P, 和任何这些的组合。

34. 一种包含约 9 至 15 个残基的免疫原性肽的组合物, 这些肽以  
小于约 500nM 的解离常数结合 HLA - A\*3101 基因产物, 并诱导细胞毒 T 细胞反应, 免疫原性肽含有 9 到约 15 个残基, 具有包含一个结合基元的 9 残基序列, 从 N 末端到 C 末端如下:

30 位于第 2 位的第 1 个初级锚定残基, 选自 A、L、I、V、M、S 和 T,  
位于第 9 位的第 2 个初级锚定残基, 选自 R 和 K;

一个次级锚定残基, 选自第 1 位的 R、H 或 K, 第 3 位的 Y、F 或 W,



第 4 位的 P, 第 6 位的 Y、F 或 W, 第 7 位的 Y、F 或 W, 第 8 位的 A 或 P, 和任何这些的组合。

35. 一种包含约 9 至 15 个残基的免疫原性肽的组合物, 这些肽以小于约 500nM 的解离常数结合 HLA-A\*3301 基因产物, 并诱导细胞毒 T 细胞反应, 免疫原性肽具有包含一个结合基元的 9 残基序列, 从 N 末端到 C 末端如下:

位于第 2 位的第 1 个初级锚定残基, 选自 A、L、I、V、M、S 和 T, 位于第 9 位的第 2 个初级锚定残基, 选自 R 和 K;

一个次级锚定残基, 选自第 3 位的 Y、F 或 W, 第 7 位的 A、Y、F 或 W 和任何这些的组合。

36. 一种包含约 9 至 15 个残基的免疫原性肽的组合物, 这些肽以小于约 500nM 的解离常数结合 HLA-A\*6801 基因产物, 并诱导细胞毒 T 细胞反应, 免疫原性肽包括含有一个结合基元的 9 残基序列, 从 N 末端到 C 末端如下:

位于第 2 位的第 1 个初级锚定残基, 选自 A、L、I、V、M、S 和 T, 位于第 9 位的第 2 个初级锚定残基, 选自 R 和 K;

一个次级锚定残基, 选自第 1 位的 Y、F、W、S、T 或 C, 第 5 位的 Y、F、W、L、I、V 或 M, 第 7 位的 Y、F 或 W, 第 8 位的 P, 和任何这些的组合。

37. 一种鉴定以小于约 500nM 的解离常数结合 HLA-A3 样分子的免疫原性肽的方法, 此方法包括以下步骤:

筛选抗原蛋白中存在权利要求 31、32、33、34、45 和 36 的结合基元的氨基酸序列;

在抗原蛋白中选择有结合基元的一个或多个亚序列;

制备包含选定亚序列的约 8 个和约 11 个残基的试验肽;

确定试验肽结合基因产物的能力;

鉴定解离常数小于 500nM 的肽。



# 说明书

## 对 HLA 分子结合亲和力提高的肽

### 发明背景

5 本发明涉及预防、治疗或诊断许多病理状态的组合物和方法。特别提供能够结合选择的主要组织相容复合物 (MHC) 分子并诱导免疫反应的新肽。

MHC 分子可分成 I 型或 II 型。II 型 MHC 分子基本上表达于参与激发和维持免疫反应的细胞, 如 T 淋巴细胞、B 淋巴细胞、巨噬细胞等。

10 II 型 MHC 分子被辅助 T 淋巴细胞识别, 并诱导辅助 T 淋巴细胞的增殖, 和对呈现的特定免疫原性肽的免疫反应的放大。I 型 MHC 分子表达于几乎所有的有核细胞上, 并被细胞毒性 T 淋巴细胞 (CTLs) 识别, 这种 T 淋巴细胞然后破坏带有抗原的细胞。CTLs 在肿瘤排斥和抵抗病毒、真菌和寄生虫的感染中是十分重要的。

15 用两种不同的试验方法已经分析了 I 型 MHC 分子的结合亲和力与分散的肽表位的免疫原性的关系 (Sette 等, 免疫学杂志., 153: 5586-5592(1994))。在第一种方法中, MHC 分子结合亲和力在 10000 倍范围内变化的潜在表位的免疫原性在 HLA-A\*0201 转基因小鼠中进行了分析。在第二种方法中, 用急性乙型肝炎病人的 PBL 评价了大约 100  
20 种乙型肝炎病毒来源的不同的潜在表位的抗原性, 这些表位都带有 A\*0201 结合基元。在这两种方法中, 发现结合阈值约为 500nM (优选 500nM 或更低) 决定肽表位引起 CTL 反应的能力。这些结果与天然加工的肽或前面所述的 T 细胞表位的 I 型结合亲和力的测定值十分相符。这些结果说明, 在 T 细胞反应形成中决定基选择具有重要作用。

25 通常, CTL 反应并不直接针对所有可能的表位。而是限于少数几个免疫优势决定基 (Zinkernagel 等, 免疫学进展 27, 51-59(1979); Bennink 等, 实验医学杂志 168, 1935-1939(1988); Rawle 等, 免疫学杂志 146, 3977-3984(1991))。早就认识到免疫优势 (Benacerraf 等, 科学 175, 273-279(1972)) 可以解释为一个给定的表位选择性地  
30 结合一个特定的 MHC 分子的能力 (决定基选择理论) (Vitiello 等, 免疫学杂志 131: 1635(1983); Rosenthal 等, 自然 267, 156-158 (1977)), 或被存在的 TCR 特异性选择性地识别 (全能理论) (Klein,



J., Immunology, the Science of Self-Nonself Discrimination, John Wiley & Sons, New York, pp270-310(1982)。已经证实,另外的因素,大多数与处理过程相关,也在诱发严格的免疫原性以外的反应中起关键作用,这些因素中的多数潜在的決定基将以免疫优势的形式存在 (Sercarz 等, 免疫学年评 11, 729-766(1993))。

调节一个特定的免疫原性肽与一个或多个 HLA 分子结合亲和力的能力,以及由此而调节由此肽引发的免疫反应的能力,将大大提高肽基疫苗和治疗制剂的用途。本发明提供这些和其它的优势。

### 发明概述

10 本发明提供用于疫苗和治疗学的肽和编码这些肽的核酸。本发明提供在病人身上诱导对预选的抗原产生细胞毒 T 细胞反应的方法。此方法包括将细胞毒 T 细胞与本发明的免疫原性肽接触。本发明的肽可能来源于许多抗原,包括病毒抗原、肿瘤相关抗原、寄生虫抗原、真菌抗原等。本发明的方法可以在体外和体内进行。在优选实施方案中,向病人给予含有编码免疫原性肽序列的核苷酸分子,使肽与细胞毒 T 细胞接触。

20 在一个实施方案中,肽有约 9-15 个残基,以小于约 500nM 的解离常数结合至少 2 个 HLA-A3 样分子,并诱导细胞毒 T 细胞反应。此免疫原性肽具有包含一个结合基元的 9 残基序列,从 N 末端到 C 末端如下:

位于第 2 位的第一个初级锚定残基,选自 A、L、I、V、M、S 和 T,位于第 9 位的第二个初级锚定残基,选自 R 和 K; 和

25 一个或多个次级锚定残基,选自位于第 3 位的 Y、F 或 W,位于第 6 位的 Y、F 或 W,位于第 7 位的 Y、F 或 W,位于第 8 位的 P,以及这些的任意组合。

本发明进一步提供结合 HLA-A\*0301 基因产物的免疫原性肽。这些肽含有一个 9 残基的结合基元,从 N 末端到 C 末端如下:

位于第 2 位的第一个初级锚定残基,选自 A、L、I、V、M、S 和 T,位于第 9 位的第二个初级锚定残基,选自 R 和 K; 和

30 一个或多个次级锚定残基,选自位于第 1 位的 R、H 或 K,位于第 3 位的 Y、F 或 W,位于第 4 位的 P、R、H、K、Y、F 或 W,位于第 5 位的 A,位于第 6 位的 Y、F 或 W,位于第 8 位的 P,以及这些的任意组合。



本发明也提供结合 HLA-A\*1101 基因产物的免疫原性肽。这些肽含有一个 9 残基的结合基元，从 N 末端到 C 末端如下：

位于第 2 位的第一个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第二个初级锚定残基，选自 R 和 K；和

5 一个次级锚定残基，选自位于第 1 位的 A，位于第 3 位的 Y、F 或 W，位于第 4 位的 Y、F 或 W，位于第 5 位的 A，位于第 6 位的 Y、F 或 W，位于第 7 位的 Y、F 或 W，位于第 8 位的 P，以及这些的任意组合。

本发明也提供结合 HLA-A\*3101 基因产物的免疫原性肽。这些肽含有一个 9 残基的基元，从 N 末端到 C 末端如下：

10 位于第 2 位的第一个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第二个初级锚定残基，选自 R 和 K；和

一个次级锚定残基，选自位于第 1 位的 R、H 或 K，位于第 3 位的 Y、F 或 W，位于第 4 位的 P，位于第 6 位的 Y、F 或 W，位于第 7 位的 Y、F 或 W，位于第 8 位的 A 或 P，以及这些的任意组合。

15 本发明也提供结合 HLA-A\*3301 基因产物的免疫原性肽。这些肽含有一个 9 残基的基元，从 N 末端到 C 末端如下：

位于第 2 位的第一个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第二个初级锚定残基，选自 R 和 K；和

20 一个次级锚定残基，选自位于第 3 位的 Y、F 或 W，位于第 7 位的 A、Y、F 或 W，以及这些的任意组合。

本发明也提供结合 HLA-A\*6801 基因产物的免疫原性肽。这些肽含有一个 9 残基的基元，从 N 末端到 C 末端如下：

位于第 2 位的第一个初级锚定残基，选自 A、L、I、V、M、S 和 T，位于第 9 位的第二个初级锚定残基，选自 R 和 K；和

25 一个次级锚定残基，选自位于第 1 位的 Y、F、W、S、T 或 C，位于第 5 位的 Y、F、W、L、I、V 或 M，位于第 7 位的 Y、F 或 W，位于第 8 位的 P，以及这些的任意组合。

30 本发明也提供鉴定以小于约 500nM 的解离常数结合 HLA-A3 样分子的免疫原性肽的方法。此方法包括筛选抗原蛋白中存在本发明结合基元的氨基酸序列，在抗原蛋白中选择有结合基元的一个或多个亚序列。制备和检测包含选定亚序列的约 8 个和 11 个残基的试验肽。然后鉴定解离常数小于 500nM 的肽。



## 定义

“肽”这个词与“寡肽”在本说明书中交换使用，以指示一系列残基，典型的是 L 氨基酸，一般是相邻氨基酸的 $\alpha$ 氨基与羧基之间通过肽键相互连接。本发明的寡肽长度小于 15 个残基，通常包含 8 个至 11 个残基，优选 9 或 10 个残基。

一种“免疫原性肽”是一种含有等位基因特异的基元或超基元的肽，这样此肽能结合一个 MHC 分子，并诱导 CTL 反应。本发明的免疫原性肽能结合到合适的 HLA-A3 样分子上，并诱导对产生此免疫原性肽的抗原的细胞毒 T 细胞反应。

“初级锚定残基”是在肽序列上一个特定位置的氨基酸，它可以在免疫原性肽和 MHC 分子之间提供一个接触点。一个限定长度的肽中 1~3 个，通常是 2 个初级锚定残基决定一个免疫原性肽的基元。这些残基通常与肽结合槽紧密接触，它们的侧链埋在槽本身的特定口袋中。典型的初级锚定残基位于 9 聚体肽的第 2 位和第 9 位。

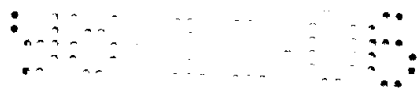
“次级锚定残基”是指一个氨基酸，它在肽中除初级锚定位置外的一个特定位置上，出现频率显著高于随机出现频率。另外次级锚定残基可以鉴定为高亲和力结合肽中出现频率较高的氨基酸或与高亲和力结合相关的氨基酸。在这些位置上特定残基的存在与否，可以用来精细地调节含有特定基元的肽的结合亲和力。

这里使用的“负结合残基”是指如果在一定的位置（通常不是初级锚定位置）上存在，将会引起对靶 HLA 分子结合亲和力降低的氨基酸。

“基元”是指限定长度的肽，通常是 8 到 15 个氨基酸，中残基的模式，它被特定的 MHC 分子识别。对于每种人的 MHC 等位基因，肽基元通常是不同的，并且在初级和次级锚定残基的模式上，肽基元也是不同的。

“超基元”是指存在于免疫原性肽上时，使肽结合一个以上的 HLA 抗原的基元。超基元优选被至少一个 HLA 等位基因以高或中等亲和力（见下面的限定）识别，优选被至少 2 个等位基因识别，更优选被至少 3 个等位基因识别，最优选被 3 个以上等位基因识别。

共有某些相似的肽结合基元的 HLA I 型分子归类为 HLA 超型。这里所用的“HLA 超型或家族”表示基于共同的肽结合特异性而归类的分



子组，而不是基于共同的抗原决定基的血清超型。

这里所用的“HLA-A3样”HLA分子（也称为一个等位基因）是指由HLA-A等位基因编码的一组HLA分子，它们与这里所公开的HLA-A3超基元共有交叉的肽结合基元。这些等位基因共有的9个残基的超基元包括以下初级锚定残基：在第2位的A、L、I、V、M、S或T和第9位（9聚体的C末端）的正电荷残基，如R和K。用血清学或DNA分型鉴定的此家族的典型成员包括：A3（A\*0301）、A11（A\*1101）、A31（A\*3101）、A\*3301和A\*6801。此家族的其它成员包括A34、A66和A\*7401。正如下面的详细解释，通过在次级锚定位置的置换，可以精细的调节与每一种等位基因的结合。

“HLA-A2样”超型的特征是在肽配体的第2位和C末端偏爱小的或脂肪族氨基酸（L、I、V、M、A和T）。此家族包含至少8种HLA-A等位基因（A\*0201、A\*0202、A\*0203、A\*0204、A\*0205、A\*0206、A\*6802和A\*6901）。

“HLA-B7样”超型包括至少一打HLA-B等位基因编码的产物（B7、B\*3501、B51、B\*5301、B\*5401、B\*5501、B\*5502、B\*5601、BB\*6701和B\*7801）（Sidney等，免疫学杂志154:247(1995)；Barber等，Curr Biol. 5:179(1995)；Hill等，自然360:434(1992)；Rammensee等，免疫遗传学41:178(1995)），其特征是识别在第2位带脯氨酸并在C末端带疏水性或脂肪族氨基酸（L、I、V、M、A、F、W和Y）的肽。

“分离的”和“生物纯的”这个词是指基本上没有在天然状态经常与之相伴的组分的物质。因此，本发明的肽不含有正常与原位环境相关的物质，如抗原呈递细胞上的MHC I分子。一种蛋白即使分离到均一或优势带，仍然有天然蛋白的5-10%的痕量污染物与目的蛋白一起纯化。本发明的分离肽不含有这种共纯化的内源性蛋白。

“残基”是指通过酰胺键或酰胺键模拟物整合到寡肽中的氨基酸或氨基酸模拟物。

#### 附图简述

图1显示5个A3样等位基因的精确等位基因特异的基元：A\*0301、A\*1101、A\*3101、A\*3301和A\*6801。显示了与每一个非锚定位置相关的，以良好的或弱的结合能力结合到每个单独的等位基因上的单个



残基，或化学性质相似的残基族。

图 2 显示 A3 样超基元。圆括号内的数字指示优选的或有害的残基或残基族的分子数。

图 3 概括了影响肽对 a)B\*0702、b)B\*3501、c)B51、d)B\*5301 和 e)B\*5401 结合能力的次级作用。这些图随后用来确定 B7 样超基元 (f)。圆括号内的数值指示优选的或有害的残基或残基族的频率。

图 4 显示由于在每一个非锚定位置上不同残基的作用，A\*0101 亚基元 9 聚体肽的相对平均结合能力。2-9 残基 (a) 3-9 残基 (b) 亚基元肽的数据组，象在材料与方法中介绍的那样分析和列表显示。2-9 和 3-9 组分别含有 101 和 85 个不同的肽。也显示了影响带 2-9 (c) 3-9 (d) 残基的 A\*0101 亚基元的 9 聚体肽的结合能力的次级作用图。

图 5 显示由于在每一个非锚定位置上不同残基的作用，A\*0101 亚基元 10 聚体肽的相对平均结合能力。分析和列表显示了 2-10 残基 (a) 或 3-10 残基 (b) 亚基元肽的数据组。2-10 和 3-10 组分别含有 91 和 89 个不同的肽。也显示了影响带 2-10 (c) 残基和 (1) 和或 3-10 残基 (d) A1 亚基元的 10 聚体肽的结合能力的次级作用图。

#### 优选实施方案的描述

本发明涉及对人 I 型 MHC (称为 HLA) 等位基因，特别是 HLA-A3 样等位基因的等位基因特异的肽基元和超基元的确定。然后这些基元用来制备和修饰任何目的抗原的 T 细胞表位，这些抗原特别包括与人病毒病、寄生虫病、真菌病或癌症相关的抗原。

正如上面所提到的，高 HLA 结合亲和力与更高的免疫原性相关。更高的免疫原性可以用数种不同的方法加以证明。例如，更高结合的肽会更经常地产生免疫。接近 90% 的高结合肽是免疫原性的，相比之下，以中等亲和力结合的肽只有 50% 是免疫原性的。更高结合的肽也导致更强烈的反应。结果，就需要更少的肽产生相似的生物学作用。因此在本发明的某些实施方案中，特别希望高结合的表位。

在本发明的某些实施方案中，需要鉴定与优势表位相对的亚优势表位。这里采用的术语中 (见 Sercarz 等, (1993), 同上)，一个“优势表位”诱导用完全天然抗原免疫产生的反应。这样一种反应在体外与肽表位交叉。一个“隐藏表位”激发由肽免疫引起的反应，但在体外不与完整蛋白用作抗原时交叉。最后，一个“亚优势表位”是这样



一种表位，即用完整的抗原免疫几乎不引起任何免疫反应，但用肽免疫可以获得一种反应，而且这种反应（不象隐藏表位那样）在用完整蛋白在体外回忆这种反应时被检测。

5 优势和亚优势的概念是与病毒病和癌症的治疗有关的。在慢性病毒病过程中，亚优势表位的参与对感染的成功清除是十分重要的，特别是在由于功能性耐受、抑制、病毒突变和其它机制引起优势 CTL 特异性失活时 (Franco 等, *Current Opinion in Immunology*, 7: 524-531, (1995))。而且，在癌症和肿瘤抗原的情况下，识别至少某些最高结合的肽的 CTL，似乎由于耐受和抑制而功能性失活，而较低结合亲和力的肽就被优先识别。

10 特别是已经提到的，大量的来自已知的非病毒肿瘤相关抗原 (TAA) 的肽，以中等亲和力 (IC50% 在 50 - 500nM 之间) 结合 HLA I 型。已经发现 15 个已知的 TAA 肽中的 8 个，被肿瘤浸润淋巴细胞 (TIL) 或 CTL 识别，以 50 - 500nM 结合。这些结果与估计的组成对照，即 90% 的已知病毒抗原被识别为以 50nM 的 IC50% 或更小结合 HLA 的肽，而大约只有 10% 以 50 - 500nM 之间结合 (Sette 等, *J. Immunol.*, 153: 5586-5592 (1994))。这种现象可能是由于在癌症状态下，因 T 细胞的耐受，而消除或功能性抑制识别数种最高结合的肽的 CTL 所致。

20 本发明提供通过初级和次级锚定位置上所需残基的选择，而调节免疫原性肽的结合亲和力的方法。正如以下详细解释的那样，这里提供了增强结合到 A3 样等位基因的超基元。依靠对结合亲和力的理想作用，置换合适肽上的锚定残基。用本发明可以获得的修饰的例子包括：增加特定等位基因的亲和力（如，通过对等位基因特异的次级锚定残基的置换）、提高不同等位基因之间的交叉反应性（如，通过一个以上的等位基因共有的次级锚定残基的置换）、和产生亚优势表位（如，通过能增加亲和力，但在免疫优势表位不存在的残基的置换）。

30 许多潜在的靶蛋白上的表位可以用于本发明。合适抗原的例子包括：前列腺特异抗原 (PSA)、乙型肝炎核心和表面抗原 (HBVc、HBVs)、丙型肝炎抗原、EB 病毒抗原、黑色素瘤抗原（如 MAGE-1）、人免疫缺陷病毒 (HIV) 抗原和人乳头瘤病毒 (HPV) 抗原。真菌抗原的例子包括来自白色念珠菌、新型隐球菌、球孢菌属的种类、组织胞浆菌属的种类和烟曲霉的抗原。寄生虫抗原包括来自疟虫属的种类、锥虫属的



种类、血吸虫属的种类、利氏曼虫等的抗原。

本发明肽的制备和评价在 PCT 出版物 W094/20127 和 W094/03205 中有描述。简言之就是：合成含有来自特定抗原的表位的肽，并检测其结合到合适 MHC 分子上的能力，使用的检测方法包括：用纯化的 I 型分子和放射碘标记的肽和/或表达空载 I 型分子的细胞，进行免疫荧光染色和流动显微荧光测定法；肽依赖的 I 型装配测定法；和肽竞争性 CTL 识别抑制。进一步评价了那些结合到 I 型分子上的肽作为 CTLs 靶子的能力，这些 CTLs 来自感染的或免疫过的个体；也评价它们诱导初级体外或体内 CTL 反应的能力，这些 CTL 反应可以产生能与疾病相关的选择性靶细胞反应的 CTL 群体。

在整个这篇公开的内容中，用 IC<sub>50</sub>'s 这个词来表示结果。设定测定法进行的条件（即限制 MHC 和标记的肽的浓度），这些值大约是 K<sub>D</sub> 值。应该指出的是，如果测定的条件改变了，IC<sub>50</sub> 的值可以改变，而且常常依据所用的特定的试剂而变化很大（如 MHC 制剂）。例如 MHC 过浓会增加给定配体的表现的 IC<sub>50</sub> 检测值。正如这里所用的，“高亲和力”定义为以小于 50nM 的 IC<sub>50</sub>（或 K<sub>D</sub>）结合。“中等亲和力”是以 50 - 500nM 之间的 IC<sub>50</sub>（或 K<sub>D</sub>）结合。测定结合的检测法在 PCT 出版物 W094/20127 和 W094/03205 中有详细描述。

另一个表示结合数据的方法，是以与一个参照肽的相对值来表示。这个参照肽在每一次检测中都用。随着特定检测法的敏感度的升降，被试肽的 IC<sub>50</sub>'s 也有一些变化。然而，与参照肽的相对的结合力是不会变的。例如在一定条件下进行检测，参照肽的 IC<sub>50</sub> 增加 10 倍，所有的 IC<sub>50</sub> 值也将升高约 10 倍。因此，为了避免意思不清楚，评价一个肽是一个好的、中等的、弱的或负的结合物，应该根据它与标准肽的 IC<sub>50</sub> 相对的 IC<sub>50</sub>。

用来描绘肽化合物的名称遵循常规的方法，即每个氨基酸残基的氨基在左边（N 末端），羧基在右边（C 末端）。在代表本发明选择的特殊实施方案的表达式中，尽管没有特别指明，但氨基和羧基基团是处于生理 pH 值条件下的状态，除非有特别的说明。在氨基酸结构式中，每种残基通常以标准的 3 字母或单字母表示法表示。氨基酸残基的 L 型以一个大写的单字母符号或大写 3 字母符号中的第一个字母表示，而有 D 型的氨基酸残基的 D 型以小写的单字母符号或小写的 3 字母符



号表示。甘氨酸没有不对称的碳原子，就简单地以“Gly”或G表示。

5 免疫原性肽可以通过合成、DNA重组技术或从诸如全病毒或肿瘤等天然来源而制备。虽然这种肽优选基本上没有其它天然存在的宿主细胞蛋白及其碎片，但在某些实施方案中，此肽能够被合成连接到天然的片段或颗粒上。

多肽或肽可以有不同的长度，不管是在中性（无电荷）状态，还是其盐，也不管是在没有被糖基化、侧链氧化或磷酸化的修饰状态，还是在有这些修饰状态，只要在这些修饰状态下，多肽的生物活性不被破坏。

10 理想的条件是，肽尽可能小，而仍然保持大肽的全部生物活性。如果可能，将本发明的肽优化到8至20个氨基酸残基的长度比较理想，典型的是9到15个，优选9到10个，与结合到细胞表面MHC I型分子上的、内源性处理的病毒肽或肿瘤细胞肽的大小相当。这里公开的优选的超基元适合于长度约9个残基的肽。其它长度的肽（如8、10  
15 和11个残基）的超基元、初级锚定和次级锚定的鉴定，可以用这里介绍的技术进行。

有所需活性的肽可以按需要被修饰成含有某些所需的特征，如：改进的药理学特性，而增加或至少基本上保留未修饰肽的全部生物活性，以结合理想的MHC分子并活化合适的T细胞。例如肽可以被不同  
20 地改变，如保守的或非保守的置换，这种改变为其使用提供了一定的优势，如改进的MHC结合。保守置换是指一个氨基酸残基被另一个生物学和/或化学性质相似的残基所置换，如一个疏水残基被另一个疏水残基置换，或一个极性残基被另一个极性残基置换。这些置换包括如下的组合：Gly, Ala; Val, Ile, Leu., Met; Asp, Glu; Asn, Gln;  
25 Ser, Thr; Lys, Arg; 和 Phe, Tyr。单个氨基酸置换的效果也可以用D-氨基酸检测。这种修饰可以用熟知的肽合成方法完成，正如在Merrifield, 科学 232: 341-347(1986), 和 Stewart and Young, Solid Phase Peptide Synthesis, (Rockford, Ill., Pierce), 2d ED. (1984) 介绍的那样。

30 肽也可以通过延伸或缩短此化合物的氨基酸序列而被修饰，如氨基酸的增加或删除。本发明的肽或类似物也可以通过改变某些残基的次序或组成而被修饰，但首先应该清楚某些氨基酸残基对生物活性是



必需的，如那些在关键结合位点的或保守的残基，这些残基通常不被改变，从而对生物活性没有不利的作用。非关键的氨基酸不必限定于在蛋白中天然存在的那些，如 L- $\alpha$ 氨基酸或其 D 异构体，也可以包括非天然的氨基酸，如  $\beta$ - $\gamma$ - $\delta$ 氨基酸，和许多 L- $\alpha$ 氨基酸的衍生物。

5 通常，单个氨基酸置换的一系列的肽被用来测定静电荷、疏水性对结合的影响。例如，沿着肽的长度制备一系列的正电荷（如 Lys 或 Arg）或负电荷（如 Glu）氨基酸置换物，揭示对各种 MHC 分子和 T 细胞受体敏感性的不同模式。另外，可以用小的、相对中性的残基如 Ala、Gly、Pro、或类似的残基进行多个置换。置换物可以是同源寡聚体或  
10 异源寡聚体。被置换或添加的残基的数量和类型，取决于基本结合点之间的必需间距与所寻求的特定功能特征（如疏水性对亲水性）。与母本肽的亲合力相比，这种置换也可以获得对 MHC 分子或 T 细胞受体增加的结合亲和力。在任何情况下，这种置换所选用的氨基酸残基或其它分子片段，应该避免破坏结合的空间和电荷的干扰。

15 典型的氨基酸置换是单个残基置换。置换、删除、插入或这些的任何组合可以合并以获得最终的肽。置换的变体是指那些在肽上至少一个残基被删除而在其位置上插入另一个不同的残基。这些置换通常按照以下的表 1 进行，此时希望精确调节肽的特征。



表 1

原始残基	示范置换
Ala	Ser
Arg	Lys, His
Asn	Gln
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Lys; Arg
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; His
Met	Leu; Ile
Phe	Tyr; Trp
Ser	Thr
Thr	Ser
Trp	Tyr; Phe
Tyr	Trp; Phe
Val	Ile; Leu



通过选择比表 1 中的置换更不保守的置换得到功能上（如对 MHC 分子或 T 细胞受体的亲和力）的明显改变，即选择对维持以下特征有显著不同作用的残基，（a）在置换区域内肽主链的结构，如片层或螺旋构象，（b）在靶位点分子的电荷或疏水性，或（c）侧链的大小。通常预计对肽的性质产生最大改变的置换是：（a）亲水残基，如丝氨酰基，置换（或被）疏水残基（置换），如亮氨酰基、异亮氨酰基、苯丙氨酰基、缬氨酰基或丙氨酰基；（b）带正电荷侧链的残基，如赖氨酰基、精氨酰基、或组氨酰基，置换（或被）负电荷残基（置换），如谷氨酰基或天冬氨酰基；或（c）有庞大侧链的残基，如苯丙氨酸，置换（或被）没有侧链的残基（置换），如甘氨酸。

此肽也可以包含免疫原性肽中两个或更多个残基的等配物。这里所说的等配物是指由于第一段序列的空间构象符合第二段序列特异的结合位点，而能够置换第二段序列的两个或更多残基的一段序列。这个词特别包括本领域技术人员熟知的肽主链修饰。这些修饰包括酰胺氮的修饰、 $\alpha$ 碳的修饰、酰胺羧基的修饰、酰胺键的完全取代、延伸、删除或主链交联。一般参见 Spatola, *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins*, Vol. VII (Weinstein ed., 1983)。

用各种不同的氨基酸的模拟物或非天然的氨基酸进行肽修饰在提高肽在体内的稳定性是特别有用的。稳定性可以用许多方法检测。例如，肽酶和各种生物介质，如人血浆和血清，已被用于检测稳定性。见如 Verhoef 等, *Eur. J. Drug Metab. Pharmacokin* 11: 291-302 (1986)。本发明的肽的半衰期用 25% 人血清 (V/V) 检测法方便地确定。方法大致如下：收集的血清 (AB 型, 非热灭活) 在使用前离心去除脂肪。然后用 RPMI 组织培养基稀释到 25%，用于检测肽的稳定性。在预定时间间隔，取出小量的反应液，加入 6% 的三氯乙酸或乙醇水溶液。絮状反应样品冷却 (4°C) 15 分钟，然后旋转使沉淀的血清蛋白形成小球。然后用稳定特异的层析条件，通过反相 HPLC 确定肽的存在。

本发明的肽或有 CTL 刺激活性的类似物，可以被修饰而提供所需的特性，而不是提高血清半衰期。例如肽诱导 CTL 活性的能力可以通过连接一段序列而加强，此序列含有至少一个能诱导 T 辅助细胞反应的表位。特别优选的免疫原性肽/T 辅助肽结合物通过一个间隔分子连



接。典型的间隔分子包括相对小的中性分子，如氨基酸或氨基酸模拟物，它们在生理条件下基本不带电。间隔物一般选自 Ala、Gly，或其它非极性氨基酸或中性极性氨基酸的中性间隔物。应当知道，任选的现有的间隔物不必包含同一个残基，因此可以是异源或同源寡聚物。

5 现有的间隔物通常是至少一个或两个残基，更常见的是 3-6 个残基。另外，CTL 肽可以不用间隔物连接到 T 辅助肽上。

免疫原性肽可以直接地，或借助一个间隔物在 CTL 肽的氨基末端或羧基末端，连接到 T 辅助肽上。免疫原性肽或 T 辅助肽的氨基末端可以被酰化。本发明所用的 T 辅助肽可以以 CTL 肽同样的方式修饰。

10 例如，它们可以被修饰成含有 D 氨基酸或连接到诸如脂、蛋白质、糖等其它分子上。典型的 T 辅助肽包括破伤风类毒素 830-843、流感 307-319、疟疾环孢子 382-398 和 378-389。

另外，T 辅助肽是一种被大多数 T 辅助细胞识别的肽。这可以通过选择结合许多、大多数或全部的 II 型 MHC 分子的氨基酸序列而实现。

15 这些是被认为“MHC 不严格限定的”或“杂乱的”T 辅助序列。氨基酸序列杂乱的例子，包括来自抗原的序列，如破伤风毒素 830-843 位 (QYIKANSKFIGITE)，Plasmodium falciparum CS 蛋白 378-398 位 (DIEKKIAKMEKASSVFNVNS) 和链球菌 18KD 蛋白 1-16 位 (GAVDSILGGVATYGAA)。

20 另外，用自然界未发现的氨基酸序列，以 MHC 不严格限定的方式，可以制备能刺激 T 辅助淋巴细胞的合成肽(见 PCT 出版物 W095/07707)。这些称为 Pan-DR 结合表位 (PADRE) 的合成化合物是根据其对大多数 HLA-DR (人 II 型 MHC) 分子的结合活性而设计的。例如，发现含有通式：aKXVWANTLKAAa 的肽，这里 X = 环己基丙氨酸、苯丙氨酸、或酪氨酸，a = D-丙氨酸或 L 丙氨酸，结合大多数 HLA-DR 等位基因，并刺激大多数个体的 T 辅助淋巴细胞的反应，不管个体的 HLA 类型。T 辅助表位也可以被修饰而提高其生物学作用。例如，存在于 T 辅助表位的肽可以含有 D-氨基酸而提高其对蛋白酶的抵抗力，因此延长其血清半衰期。T 辅助表位也可以结合到其它分子上，如脂质、蛋白质或糖，或任何其它的合成化合物，而提高其生物活性。尤其是 T 辅助肽可以  
30 以氨基或羧基末端结合到一个或多个棕榈酸链上。

在某些实施方案中，可能需要在本发明的药用组合物中包含至少



一种引发 CTL 的组分。脂质已被鉴定为能引发体外抗病毒抗原 CTL 的试剂。例如，棕榈酸残基能结合到 Lys 残基的 $\alpha$ 和 $\epsilon$ 氨基基团上，然后借助一个或多个连接残基，如 Gly、Gly-Gly-、Ser、Ser-Ser 等，连接到免疫原性肽上。然后这种脂化的肽能以分子团的形式、整合到脂质体中或与不完全弗氏佐剂等佐剂乳化直接注射。在优选实施方案中，特别有效的免疫原包括结合到 Lys 残基的 $\alpha$ 和 $\epsilon$ 氨基基团上的棕榈酸，这是借助 Ser-Ser 的连接结合到免疫原性肽的氨基端上。

作为 CTL 反应的脂质引发的另一个例子，大肠杆菌脂蛋白，如三棕榈酰-S-甘油酰半胱氨酰丝氨酰-丝氨酸 ( $P_3$ CSS) I，当共价结合到一个适合的肽上，可以用来引发病毒特异的 CTL。见 Deres 等，自然 342: 561-564 (1989)，在此引用作为参考。本发明的肽可以偶联到  $P_3$ CSS 上，例如给一个个体给予脂肽以特异地引发对靶抗原的 CTL 反应。而且，由于中和抗体的诱导也可以用结合到呈现合适表位的肽上的  $P_3$ CSS 引发，所以这两种组合物可以合并使用，以更加有效地激发对抗感染的体液和细胞介导的应答。

另外，额外的氨基酸可以加到肽的末端，而使一个肽容易连接到另一个肽上、使肽偶联到载体或大肽上、修饰肽或寡肽的理化性质等。氨基酸如酪氨酸、半胱氨酸、赖氨酸、谷氨酸或天冬氨酸等可以被引入到肽或寡肽的 C 或 N 末端。在某些例子中，C 末端的修饰会改变肽的结合特性。而且肽和寡肽的序列由于被末端-NH<sub>2</sub> 酰化，如链烷基 (alkaboyl) ( $C_1 - C_{20}$ )、或巯基乙酰化，末端羧基酰胺化，如氨、甲胺等修饰，而与天然序列不同。在某些例子中，这些修饰能提供结合到一个支持物或其它分子的位点。

本发明的肽能够以许多途径制备。由于较短，此肽可以按常规技术在溶液中或在固体支持物上合成。有许多商品化的自动合成仪，可以按已知的程序使用。例如见：Stewart and Young, Solid Phase Peptide Synthesis, 2d. ed., Pierce Chemical Co. (1984)，同上。

另外，可以用重组 DNA 技术，将编码一个感兴趣的免疫原性肽的核苷酸序列插入到一个表达载体中，转化或转染到一个合适的宿主细胞中，在适合表达的条件下培养。这些步骤在本领域是众所周知的，正如在 Sambrook 著的 Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, New York (1982)



所一般介绍的，在此引用作为参考。因此，包含一个或多个本发明的肽序列的融合蛋白可以用来呈递合适的 T 细胞表位。

5 由于这里期望长度的肽的编码序列可以通过化学技术合成，例如 Matteucci 等，J. Am. Chem. Soc. 103: 3185(1981)介绍的磷酸三酯法，可以通过简单地将编码天然肽序列的合适碱基置换而进行修饰。然后编码的序列可以提供合适的连接物，并连接到本领域常用的表达载体中，此载体用来转化合适的宿主产生所需要的融合蛋白。现在已有许多这样的载体和合适的宿主系统。对于融合蛋白的表达，编码序列带有可操作连接的起始密码子和终止密码子、启动子和终止子区以及通常的复制系统，以提供在所需的细胞系统中表达的表达载体。例如，与细菌宿主相容的启动子序列装在含有合适的限制位点以插入所需编码序列的质粒中。得到的表达载体转化到合适的细菌宿主中。当然，合适的载体和对照序列也可以使用酵母或哺乳动物细胞宿主。

15 本发明的肽和本发明的药用和疫苗组合物，对给哺乳动物给药，特别是人给药，以治疗和/或预防病毒感染和癌症是有用的。用本发明的免疫原性肽可以治疗的疾病的例子包括前列腺癌、乙型肝炎、丙型肝炎、AIDS、肾癌、颈癌、淋巴瘤、CMT 和尖锐湿疣。另外，此肽可用于治疗任何感染性疾病，如病毒、真菌和寄生虫感染。例如在 W094/20127 和 W094/03205 中公开了合适的抗原。

20 正如上面所提到的。本发明的肽与对肽上表位特异的 CTL 接触时，诱导 CTL 免疫反应。然而，肽与 CTL 接触的方式对本发明并不关键。例如肽可以在体内或体外与 CTL 接触。如果接触发生在体内，肽本身可直接对病人给药，或用以下介绍的其它的载体（如编码一种或多种肽的 DNA 载体、编码此肽的病毒载体、脂质体等）给药。

25 对于药用组合物，免疫原性肽或编码它们的 DNA 给药给已经患有癌症或感染有目的病原体的病人。肽或编码它们的 DNA 可以单独给药，或以这里公开的一个或多个肽序列融合给药。处于潜伏期或急性感染期的病人可以用免疫原性肽单独治疗，或其它合适的治疗方法联合。在治疗应用中，向病人给予足够量的组合物，以引发对病毒或癌症抗原有效的 CTL 反应，来治疗或至少部分控制症状和/或并发症。达到此目的足够的量称为“治疗有效量”。对这种用途的有效量取决于诸如所给予的特定的组合物、给药的方式、疾病治疗时的阶段和严重程度、

30



病人的体重和大概健康状态和处方医生的判断。但初次免疫（治疗或预防给药）的剂量范围大概是：70kg 的病人  $1.0\mu\text{g} \sim 5000\mu\text{g}$  的肽。加强剂量为  $1.0\mu\text{g} \sim 1000\mu\text{g}$  的肽，根据检测病人血液中特异的 CTL 的活性而确定的病人的反应和条件，决定加强期为数周或数月。必须牢记本发明的肽和组合物通常是用于严重的疾病状态，即威胁生命或潜在威胁生命的状态。在这种情况下，鉴于此肽的外源物质最少和相对无毒性，治疗医生可能和希望给予超量的肽组合物。

对于治疗应用，应该在病毒感染的第一症状出现时，或肿瘤检出或手术切除时，或急性感染诊断后不久，开始给药。然后加强剂量直到至少症状基本消退并维持一段时间。在慢性感染中需要在加强剂量后补充剂量。

用本发明的组合物治疗感染的病人可以促进急性感染病人感染的消除。对于那些易于（或倾向于）发展成慢性感染的病人，这些组合物对于预防从急性转变成慢性是特别有用的。例如一旦在感染前或在感染过程中，确定了易于转变的病人，正如这里介绍的，就向他们给予组合物，尽量减少给更大人群给药的必要。

此肽和其它组合物也可以用来治疗慢性感染，和刺激免疫系统来清除携带者中病原体感染的细胞。重要的是，以一种制剂和方式提供的免疫增强肽的量，要足以有效地刺激 T 细胞反应。因此，为了治疗慢性感染，典型的剂量是，70kg 病人每次约  $1.0\mu\text{g} \sim 5000\mu\text{g}$  的范围，优选约  $5\mu\text{g} \sim 1000\mu\text{g}$ 。可能需要在确定的间隔期，如 1 到 4 周，给予免疫剂量继之以加强剂量，持续一个延长期至有效地免疫病人。在慢性感染的情况下，应该连续给药，直到临床症状或实验室检测显示病毒感染被清除或基本消退，并维持一段时间。

用于治疗的药用组合物打算通过非肠道、表面、口服或局部给药。优选地，此药用组合物通过非肠道给药，如静脉、皮下、真皮内或肌肉内。因此本发明提供用于非肠道给药的组合物，此组合物包括免疫原性肽溶于或悬浮于一个可接受的载体，优选水性载体的溶液。可以使用各种水性载体，如水、缓冲水、0.8% 的盐水、0.3% 的甘氨酸、透明质酸等。这些组合物可以通过常规的、熟知的灭菌技术除菌，也可以通过过滤除菌。得到的液体溶液可以分装备用，或冻干。冻干的制剂在给药前与无菌溶液联合。这些组合物可以含有药用可接受的、



因需要接近生理条件而加的辅助物质，如 pH 调节和缓冲剂、张力调节剂、润湿剂等，例如乙酸钠、乳酸钠、氯化钠、氯化钾、氯化钙、单月桂酸山梨聚糖、三乙醇胺油酸酯等。

5 本发明的 CTL 刺激肽在药用制剂中的浓度变化很大，即从重量小于 0.1%，通常在或小于 2%，到高至 20% 至 50% 或更高。浓度主要通过液体体积、粘性等选择，与所选择的特定的给药方式一致。

10 本发明的肽也可以借助脂质体给药，脂质体把肽送到特定的组织，如淋巴组织、或选择性地送到感染的细胞，还可以延长肽组合物的半衰期。脂质体包括乳剂、发泡剂、微胶粒、不溶性单分子层、液晶、磷脂分散体、片层等。在这些制剂中，释放的肽单独地、或与一个分子联合整合到脂质体中，这个分子结合到淋巴细胞中普遍存在的受体，如结合 CD45 抗原的单克隆抗体上，或与其它治疗的或免疫原性的组合物联合整合到脂质体中。因此填充有或包裹有本发明肽的脂质体，可以被直接送到淋巴细胞的部位，在这里脂质体然后释放选择的治疗性/免疫原性肽组合物。本发明所用的脂质体是由标准的发泡脂制备的，通常包含中性的和带负电荷的磷脂和甾醇，如胆甾醇。脂质的选择通常以以下的考虑为指导，如脂质体的大小、耐酸性和脂质体在血流中的稳定性。已有许多方法制备脂质体，如 Szoka 等，生物物理与生物工程年评 9: 467(1980)，U.S. Patent Nos. 4235871, 4501728, 15 4837028, 和 5019369 中介绍的。在此引入作为参考。

20 为了运送到免疫细胞上，整合到脂质体中的配体可以包括，诸如对所需的免疫系统细胞的表面决定基特异的抗体或其片段。含有肽的脂质体悬液可以静脉给药，表面给药或局部给药，给药的剂量具体根据给药的方式、所给的肽的种类和所治疗疾病的阶段而变化很大。

25 对于固体组合物可以使用常规的无毒的固相载体，例如包括药用级的甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、滑石、纤维素、葡萄糖、蔗糖、碳酸镁等。对于口服给药，药用可接受的无毒组合物是通过混合任何正常使用的赋形剂，如前面列出的那些载体，和一般 10 - 95% 的活性成分，即一种或多种本发明的肽，更优选的浓度为 25 - 75% 而形成。

30 对于气溶胶给药，免疫原性肽优选地与表面活性剂和推进剂一起以精细分散的形式提供。肽的典型的百分比是重量的 0.01% - 20%，



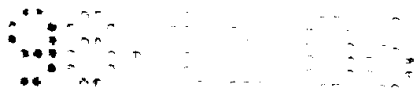
5 优选 1% - 10%。表面活性剂必须是无毒的，而且最好在推进剂中是可溶的。这些试剂的代表是含 6 至 22 个碳原子的脂肪酸的酯或不完全酯，如己酸、辛酸、月桂酸、棕榈酸、硬脂酸、亚油酸、亚麻酸、油硬脂酸 (olestERIC) 和油酸与脂肪族多元醇或其环状酐形成的酯。可以使用混合酯如混合或天然的甘油酯。表面活性剂可以占组合物重量的 0.1% - 20%，优选 0.25-5%。组合物的平衡是用通常的推进剂。如果需要，也可包括载体，如鼻内给药的卵磷脂。

10 本发明的另一方面涉及疫苗，疫苗含有一种活性组分，即免疫有效量的这里介绍的一种或多种免疫原性肽。这种肽可以通过连接到自身的载体或以活性肽单元的同源多聚物或异源多聚物的形式引入到包括人的宿主体内。这种多聚物具有提高免疫反应的优势，而且如果用不同的肽制备多聚物，就有另外的能力来诱导与病毒或肿瘤细胞不同的抗原决定基反应的抗体和/或 CTLs 的产生。有用的载体在本领域是熟知的，包括诸如甲状腺球蛋白、白蛋白如人血清白蛋白、破伤风类毒素、多聚氨基酸如多聚(赖氨酸: 谷氨酸)、流感病毒、乙型肝炎核  
15 心蛋白、乙型肝炎重组疫苗等。疫苗也可以含有生理耐受的(可接受的)稀释剂，如水、磷酸缓冲盐水、或盐水，而且进一步一般包括一种佐剂。佐剂，如不完全弗氏佐剂、磷酸铝、氢氧化铝、或明矾是本领域熟知的物质。而且，如上所述，CTL 反应可以通过将本发明的肽结  
20 合到脂质，如 P<sub>3</sub>CSS 上而引发。用这里介绍的肽组合物，通过注射、气溶胶、口服、经皮或其它途径免疫，宿主的免疫系统通过产生大量的对目的抗原特异的 CTLs 而对疫苗反应，而且宿主对以后的感染至少部分免疫，或对发展成慢性感染有抵抗力。

25 含有本发明肽的疫苗组合物给予易于患、或正处于病毒感染或癌症危险的病人，以激发对抗原的免疫反应，因此提高病人自身的免疫反应能力。这样一种量称为“免疫有效量”。在这种使用中，精确的剂量再次取决于病人的健康状况和体重、给药的方式和制剂的性质等。但通常的范围是 70kg 的病人约 1.0μg 至大约 5000μg，更常用的是每 70kg 体重约 10μg 至大约 500μg。

30 在某些例子中，需要将本发明肽疫苗与能诱导对目的病毒，特别是对病毒被膜抗原，产生中和抗体反应的疫苗联合。

为了治疗或免疫的目的，本发明的肽也可以用减毒的病毒宿主，



如牛痘或禽痘病毒表达。这种方法涉及将牛痘病毒用作载体来表达编码本发明肽的核苷酸序列。如果将其引入急性或慢性感染的宿主或引入未感染的宿主，重组的牛痘病毒会表达免疫原性肽，因而激发宿主的 CTL 反应。牛痘载体和免疫程序有用的方法如 U.S Patent No. 4722848 的描述。在此引入作为参考。另一个载体是 BCG (Bacille Calmette Guerin)。BCG 载体在 Stover 等, (Nature 351: 456-460 (1991)) 中有描述, 在此引入作为参考。对本发明肽的治疗给药或免疫有用的大量其它的载体, 如伤寒沙门氏菌载体等, 从这里的描述, 本领域的技术人员是明了的。

抗原肽也可用来引发离体 CTL。产生的 CTL 能用来治疗对其它常规的治疗方法没有反应, 或对肽疫苗的治疗方法没有反应的病人的慢性感染 (病毒性或细菌性) 或癌症。通过将病人的 CTL 前体细胞 (CTLp) 与一种来源的抗原呈递细胞 (APC) 以及合适的免疫原性肽在组织培养中共孵, 可以诱发对特定病原体 (感染因子或肿瘤抗原) 的离体 CTL 反应。经过合适的孵育时间 (通常 1-4 周), CTLp 被激活并成熟和发展成效应 CTL。这些细胞重新输回给病人, 这样它们将破坏特异的靶细胞 (感染的细胞或肿瘤细胞)。

编码一个或多个本发明肽的 DNA 也可以给病人给药。这种方法, 例如在 Wolff 等, 科学 247: 1465-1468 (1990) 和 U.S. Patent Nos. 5580859 和 5589466 中有描述。

编码本发明肽的核苷酸给药的一个优选方法是使用编码多个本发明表位的小基因构建体。为了构建在人细胞中表达的编码选择的 CTL 表位的 DNA 序列 (小基因), 反翻译此表位的氨基酸序列。使用人密码子使用表指导每一个氨基酸的密码的选择。这些编码表位的 DNA 序列直接毗连, 构建一个连续的多肽序列。为了优化表达和/或免疫原性, 将另外的元件整合到小基因的设计上。能够被反翻译并包括在小基因序列中的氨基酸序列的例子包括: 辅助 T 淋巴细胞表位、引导 (信号) 肽和内质网滞留信号。另外, CTL 表位的 MHC 呈递, 能够通过包含与 CTL 表位毗连的合成的 (如多聚丙氨酸) 或天然存在的旁侧序列而加强。

小基因序列通过装配寡核苷酸而转到 DNA 中, 这些寡核苷酸编码小基因的正链和负链。用熟知的技术, 在合适的条件下合成、磷酸化、纯化和退火重叠的寡核苷酸 (30-100 个碱基长)。用 T4 DNA 连接酶



连接寡核苷酸的末端。这种合成的小基因，编码 CTL 表位多肽，能够克隆到所需的表达载体中。

5 本技术领域熟知的标准调节序列包含在载体中，以保证在靶细胞中表达。还需要几种载体元件：一个带有下游克隆位点以便小基因的插入的启动子；一个为了有效转录终止的聚腺苷化信号；一个大肠杆菌复制起始区；和一个大肠杆菌的选择标记（如氨苄青霉素或卡那霉素抗性）。许多启动子可以用于此目的，如人巨细胞病毒（hCMV）启动子。其它合适的启动子序列，见 U. S. Patent Nos. 5580859 和 5589466。

10 可能需要另外的载体修饰来优化小基因表达和免疫原性。在某些情况下，有效的基因表达需要内含子，一个或多个合成的、或天然存在的内含子可以整合到小基因的转录区。也可以考虑包含 mRNA 稳定序列以提高小基因的表达。近来已经提出免疫刺激序列（ISSs 或 CpGs）对 DNA 疫苗的免疫原性起作用。如果发现可以提高免疫原性，这些序列可以包含在载体中，在小基因编码序列之外。

15 在某些实施方案中，可以使用双顺反子，使产生小基因编码的表位和第二蛋白，后者用来加强或减弱免疫原性。蛋白或多肽如果协同表达，能够有助于加强免疫反应的例子包括细胞因子（如 IL2、IL12、GM-CSF）、细胞因子诱导分子（如 LeIF）或协同刺激分子。辅助（HTL）表位可以结合到细胞内靶信号上，并与 CTL 表位分开表达。这就使 HTL 表位针对与 CTL 表位不同的细胞区。如果需要，还可以使 HTL 表位更有效地进入到 II 型 MHC 途径，因此加强 CTL 诱导。与 CTL 诱导相比，通过免疫抑制分子（如 TGF- $\beta$ ）的协同表达特异性地减弱免疫反应在某些疾病是有益的。

25 一旦选择了一个表达载体，小基因就克隆到启动子下游的多克隆位点中。这个质粒转化到合适的大肠杆菌菌株中，用标准技术制备 DNA。小基因的方向和 DNA 序列，以及包含在载体中的其它元件，用限制性酶切图谱和 DNA 序列分析来证实。装有正确质粒的细菌细胞可以以主细胞库和工作细胞库保存。

30 治疗量的质粒 DNA 通过发酵大肠杆菌生产，接着纯化。用工作细胞库中的小量菌种，接种到发酵培养基中（如 Terrific 肉汤），按熟知的技术在摇瓶或生物反应器中生长至饱和。用标准的生物分离技术，如 Quiagen 公司提供的固相阴离子交换树脂，纯化质粒 DNA。如果需



要，可以用凝胶电泳或其它方法从开环和线性 DNA 中分离超螺旋 DNA。

纯化的质粒 DNA 可以用许多制剂注射。最简单的是冻干的 DNA 在灭菌的磷酸缓冲盐水 (PBS) 中重配。这种方法称为“裸 DNA”，近来正用于临床实验的肌肉内 (IM) 注射。为了使小基因 DNA 疫苗的免疫治疗效果达到最大，可能需要另一种方法配制纯化的 DNA。已经介绍了许多方法，并且有新方法。阳离子脂质也可以用于这种制剂 (如见 Debs and Zhu(1993) WO 93/24640; Mannino and Gould-Fogerite(1988) 生物技术 6(7): 682-691; Rose U. S. Pat No. 5279833; Brigham(1991) WO 91/06309; 和 Felgner 等, (1987) 美国国家科学院院刊 84: 7413-7414 中介绍的)。另外，糖脂、融合脂质体、肽和统称为保护性、相互作用、非浓缩 (PINC) 的化合物也可以与纯化的质粒 DNA 混合，以影响变量，如稳定性、肌肉内分散、或结合到特异的器官或细胞类型上。

核酸也可以用弹道发送的方式给药，例如，象 U. S. Patent No. 5204253 介绍的那样。可以给予只包含 DNA 的颗粒。另外，DNA 可以粘附到颗粒，如金粒上。

靶细胞的致敏可用作一种功能性检测法，检测编码 CTL 表位的小基因的表达和 I 型 MHC 的呈递。质粒 DNA 引入到哺乳动物细胞系中，此细胞系合适于作为标准 CTL 铬释放检测法的靶子。使用的转染法取决于最终的配方。对于“裸”DNA 可以用电穿孔法，而阳离子脂质就可以在体外直接转染。表达绿荧光蛋白 (GFP) 的质粒可以共转染，而通过荧光活化细胞分类法 (FACS) 富集转染的细胞。然后这些细胞用铬-51 标记，用作表位特异的 CTL 系的靶细胞。细胞裂解，通过 <sup>51</sup>Cr 的释放检测，指示编码 CTL 表位的小基因的 MHC 呈递的产生。

体内免疫原性是第二种功能性检测小基因 DNA 制剂的方法。表达合适人 MHC 分子的转基因小鼠用 DNA 产物免疫。给药剂量和途径取决于制剂 (如，IM 适合溶于 PBS 的 DNA，IP 适合与脂质混合的 DNA)。免疫后 21 天，收集脾细胞，在有编码每种被试表位的肽存在下再刺激 1 周。用标准的技术检测这些效应细胞 (CTLs) 对带有肽、铬-51 标记的靶细胞的细胞裂解作用。由于 MHC 载有相应于小基因编码的表位的肽而致敏的靶细胞裂解，就证实了 DNA 疫苗体内诱导 CTLs 的功能。

本发明的肽也发现用作诊断试剂。例如本发明的肽可以用来确定



特定的个体对使用肽或相关肽的治疗方案的敏感性，因此对改进现有的治疗方案、或确定感染病人的预后是有帮助的。另外，此肽也可以用来预测哪些病人会有发展成慢性感染的危险。

下面提供的例子是为了说明，并非为了限制。

5

### 实例 1

#### A3 样超型结合

此实例提供每一种 A3 样等位基因 A3、A11、A\*3101、A\*3301 和 A\*6801 的精细的基元，概要的次级锚定结合特异性。基元是通过计算此肽序列上每一个非锚定位置上肽的平均相对结合能力而得到的，这些肽带有 20 个常见氨基酸的每一个。按照每一个的化学相似性分组。

#### 材料与方法

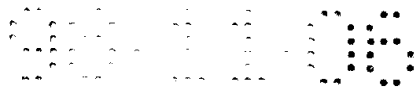
I 型纯化。以下的 Epstein-Barr 病毒 (EBV) 转化的纯合细胞系用作 I 型分子的来源: GM3107 (A3、B7; 人遗传突变资源库); BVR (A11、B35.3、Cw4; 人遗传突变资源库); SPACH (A31、B62、Cw1/3; ASHI 资源库总汇); 和 LWAGS (A\*3301、B14、Cw8; ASHI 资源库总汇) (Bodmer 等, 人类免疫学 43: 149 (1995))。由 Dr. Walter Storkus (匹兹堡大学) 鉴定的 C1R 转染子用来分离 A\*6801。细胞系按以前的介绍维持 (Sidney 等, 免疫学杂志 154: 247 (1995); Sette 等, 分子免疫学 31: 813 (1994))。

20

按以前的介绍 (Sidney 等, 免疫学杂志 154: 247 (1995); Sette 等, 分子免疫学 31: 813 (1994)) 制备细胞裂解物并纯化 I 型分子。

简言之, 细胞在含 1% NP-40 (Fluka Biochemika, Buchs, Switzerland)、150mM NaCl、5mM EDTA 和 2mM PMSF 的 50mM Tris-HCl, pH8.5 中以  $10^8$  细胞/ml 的浓度裂解。裂解物过 0.45 $\mu$ m 的滤膜, 10000g 离心 20 分钟除去细胞核和碎片。然后用亲和层析纯化主要组织相容复合物 (MHC) 分子。非活化的 Sepharose CL4B 和蛋白 A Sepharose 柱用作预柱。细胞裂解物反复通过连接有抗 HLA (B、C) 抗体 B1.23.2 (Rebai 等, 组织抗原 22: 107 (1983)) 的蛋白 A Sepharose 珠, 而清除 HLA-B 和 HLA-C 分子。为了有效地清除, 一般需要 2 至 4 次通过。接着, 用抗 HLA (A、B、C) 抗体 W6/32 (Barnstable 等, 细胞 14: 9 (1978)) 来结合 HLA-A 分子。蛋白的纯度、浓度、和清除步骤的效果由十二烷基硫酸钠-聚丙烯酰胺凝胶电泳 (SDS-PAGE) 检测。

30



结合检测和超基元确定。基于放射标记的标准探针肽结合到溶于去污剂的 MHC 分子的抑制, 肽与可溶 I 型分子结合的定量检测按以前介绍的进行 (Kubo 等, 免疫学杂志 152: 3913(1994); Kast 等, 免疫学杂志 152: 3904(1994); Sidney 等, 免疫学杂志 154: 247(1995); Sette 等, 分子免疫学 31: 813(1994); Ruppert 等, 细胞 74: 929(1993))。简言之, 1-10nM 放射标记的探针肽, 氯胺 T 法标记碘 (Greenwood 等, 生物化学杂志 89: 114(1963)), 在有 1 $\mu$ M 人 $\beta_2$ 微球蛋白 (Scripps Laboratories, San Diego, CA, USA) 和混合的蛋白酶抑制剂存在下, 与不同量的 MHC 在室温下共孵。共孵 2 天后, MHC 结合的放射活性的百分率用 TSK2000 柱上通过大小排除凝胶过滤层析确定。

A3CON1 肽 (序列 KVFPYALINK) (Kubo 等, 免疫学杂志 152: 3913(1994)) 用作检测 A3、A11、A31 和 A\*6801 的放射标记探针。HBVc141-151 (序列 STLPETYVRR) 的 T<sub>7</sub>→Y 类似物 (Missale 等, 实验医学杂志 177: 751(1993)) 用作检测 A\*3301 的放射标记探针。在竞争检测法的情况中, 计算产生放射标记探针肽结合抑制 50% 的肽浓度 (IC<sub>50</sub>)。常常在一或两个高剂量检测肽, 在随后的实验中确定产生阳性抑制的肽的 IC<sub>50</sub>, 在此实验中如果需要, 进一步检测 2-6 个稀释度。使放射标记探针肽大约 15% 结合的 MHC 浓度用于所有的竞争抑制检测。因为在这些条件下 (标记的) < (MHC) 而且 IC<sub>50</sub> > (MHC), 检测的 IC<sub>50</sub> 是真实 K<sub>d</sub> 值的合理的近似值。每一种竞争肽在 2-4 个完全独立的实验中检测。作为一个阳性对照, 在每一次实验中都要检测与放射标记探针有关的未标记的肽, 并测定其 IC<sub>50</sub>。A3CON1 检测 A3、A11、A31 和 A\*6801 的平均 IC<sub>50</sub> 分别是 11、6、18 和 8nM。eh HBVc 141-151 肽在 A\*3301 检测法中的平均 IC<sub>50</sub> 是 29nM。

HLA-A 特异的次级锚定基元的确定。使用 Ruppert 等 (Ruppert 等, 细胞 74: 929(1993)) 介绍的方法的改良法来确定 A\*0201 的次级锚定基元。简言之, 等位基因特异的次级锚定基元, 通过评价对 20 个常见氨基酸在 9 聚体序列的每一个非锚定位置上的 HLA 结合的作用而确定。通过计算每一位置氨基酸组合 (如 1 位丙氨酸; 2 位丙氨酸等) 的平均相对结合值而进行评估。为克服某些氨基酸的出现频率低的问题, 如以前介绍的 (Ruppert 等, 细胞 74: 929(1993)) 那样, 将一些



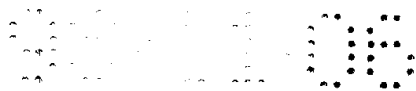
氨基酸与其它性质类似的归成一组。在特定位置上以平均结合能力 4 倍高于（或低于）全部 200 肽组合的总的平均结合能力结合的残基类型，就被认为以好（或弱）结合能力结合。

5 肽合成。按以前介绍的方法（Ruppert 等，细胞 74: 929(1993)）合成肽，或从 Chiron Mimotopes(Chiron Corp., Australia)以原料购买。合成的肽通过反相高压液相层析（HPLC）纯化到 >95% 的均一性。这些合成肽的纯度在分析反相柱上检测，其组成用氨基酸分析、测序和/或质谱分析确定。

在不同的人种背景和设计的人群覆盖度中 HLA 超型的表型频率的计算。用双向分布公式  $gf=1-(\text{SQRT}(1-af))$  (Tiwari 等, The HLA Complex, 见 HLA and Disease Associates, NY, Springer-Verlag (1985)) 从抗原或等位基因的频率计算每一种 HLA 等位基因的基因频率 (Imanishi 等, 第 11 届国际组织相容性专题讨论会论文汇编, Vol, 1, Tokyo, Oxford University Press (1992); Fernandez-Vina 等, 15 人类免疫学 33: 163(1992))。为获得总的表型频率, 计算累积基因频率, 用逆公式 ( $af=1-(1-Cgf)^2$ ) 推算累积抗原频率。如下所讨论的, 如果在 DNA 分型水平没有频率数据, 就假定对应的血清学鉴定的抗原频率。为了获得总的潜在的人群覆盖度, 假定没有连锁不平衡并且只包括经鉴定属于每一种超型的等位基因 (最少估计值)。由基因座相互 20 组合取得的总的潜在覆盖度的估计值, 是通过把非 A 覆盖人群的比例加到 A 的覆盖度, 非 A 覆盖人群的比例可以认为是 B 等位基因所覆盖的部分 (如, 总数 =  $A + B^*(1-A)$ )。A3 系超型的确定成员是 A3、A11、A31、A\*3301 和 A\*6801。虽然 A3 样超型可能潜在地包括 A32、A66 和 A\*7401, 但这些等位基因并不包括在总的频率计算值中。同样地, A2 样超型家族的确定成员包括 A\*0201、A\*0202、A\*0203、A\*0204、A\* 25 0205、A\*0206、A\*6802 和 A\*6901 (可能也包括 A\*3001)。最后, B7 样超型的确定等位基因是 B7、B\*3501-03、B51、B\*5301、B\*5401、B\*5501-2、B\*5601、B\*6701 和 B\*7801 (可能也包括 B\*1401、B\*3504 - 06、B\*4201 和 B\*5602)。

30 结果

不同的 HLA 等位基因的肽结合口袋的结构分析。如上所述, 以前的研究显示, HLA 分子 A3、A11 和 A\*6801 与第二位带有小的或疏水



的残基、并在 C 末端带正电荷的配体的特异性有关 (Kubo 等, 免疫学杂志 152: 3913(1994); Guo 等, 自然 360: 364(1994); Falk 等, 免疫遗传学 40: 238(1994); Dibrino 等 免疫学杂志 151: 5930(1993); Dibrino 等, 美国国家科学院院刊 90: 1508(1993); Zhang 等, 美国国家科学院院刊 90: 2217(1993); Sette 等, 分子免疫学 31: 813(1994)). 我们决定更详细地评价在配体特异性上这些明显的相似性的结构基础。为这个目的, 由于已知抗原肽在第二位和 C 末端的残基侧链与形成 HLA I 型分子的 B 和 F 口袋的残基接触 (Madden 等, 细胞 75: 693(1993); Saper 等, J Mol Biol 219: 277(1991)), 因此列表显示对于各种 HLA I 型分子形成这些多形袋口的残基。发现已知识别第二位上小的或疏水的残基的 HLA 型 (如, A\*0101、A\*0201、A\*0301、A\*1101、A\*6801、和 A\*6802), 和识别配基肽的 C 末端带正电荷残基的 HLA 型 (如, A\*0301、A\*1101、A\*6801、和 A\*2705) 共有某些关键的结构特征。尤其是, 结合第二位小的和疏水残基的 HLA 分子分别在第 45 位和 67 位带有脂肪族残基 (M 或 V), 在 66 位和 70 位分别带有潜在的氢键形成残基如 N 和 K, 或 H 和 Q。所有的分子在 99 位都带有 Y 残基。相比之下, 表现出不同结合特异性的 I 型分子在一个或多个位置不同。同样, 只有优选的 C 末端带正电荷的 I 型分子在 77、80、81 和 116 位分别带有 D、T、L 和 D。总而言之, 这种分析显示一类 HLA I 型分子 (A3、A11 和 A\*6801) 在其 B 和 F 口袋共有某些关键的结构特征, 以及一个常见的肽配体基元, 其特征在于第二位的小的或疏水的残基和 C 末端带正电荷的残基。在这一点上, 我们暂时把这些 HLA I 型分子称为 A3 样超型的一部分, 相应的基元称为 A3 样超基元。

对其它基元不清楚的 I 型分子的分析显示, A\*3101、A\*3301、A\*3401、A\*6601 和 A\*7401 在 B 和 F 口袋也共有这些同样的共有序列。因此预测这些分子是 A3 样超型的一部分。最近的结果 (Falk 等, 免疫遗传学 40: 238(1994)) 独立地证实 A31 和 A33 的确有 A3 样肽基元的特征。

表现重叠初级锚定特异性和肽结合全部特性的 A3 样分子。为了比较被某些最常见的 A3 样分子 (A3、A11、A31、A\*3301 和 A\*6801) 识别的基元的范围, 对这些分子的主要锚定 (第二位和 C 末端) 结合

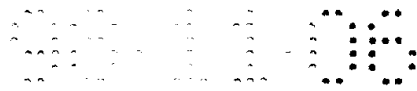


特异性进行了更详细的分子分析。以前已经介绍了测定非标记合成肽对放射标记探针肽结合到亲和纯化的 HLA I 型分子的抑制能力的 A3 和 A11 特异的肽结合检测法 (Kubo 等, 免疫学杂志 152: 3913(1994); Kast 等, 免疫学杂志 152: 3904(1994); Sette 等, 分子免疫学 31: 813(1994) )。用类似的方法建立了对 A31、A\*3301 和 A\*6801 特异的结合检测法 (Kubo 等, 免疫学杂志 152: 3913 (1994); Kast 等, 免疫学杂志 152: 3904(1994); del Guercio 等, 免疫学杂志 154: 685(1995); Sidney 等, 免疫学杂志 154: 247(1995); Sette 等, 分子免疫学 31: 813(1994); Ruppert 等, 细胞 74: 929(1993) )。

然后, 通过检测一个原型多聚 A 肽 AXAAAAAX 的第 2 位或第 9 位被置换的一组肽对 A3、A11、A31、A\*3301 和 A\*6801 的抑制能力, 考察 A3 样分子的初级锚定特异性。每一个分子显示偏爱各自特定的残基, 但在大多数所考虑的例子中, 第 2 位被 A、I、L、M、S、T 或 V 占据和 C 末端被 R 或 K 占据时, 获得显著的结合。这些结果与由我们 (Kubo 等, 免疫学杂志 152: 3913 (1994) ) 和其它人 (Falk 等, 免疫遗传学 40: 238(1994); Dibrino 等 免疫学杂志 151: 5930(1993); Dibrino 等, 美国国家科学院院刊 90: 1508(1993) ) 得出的集中测序结果完全相符, 并且在 A31、A\*3301 和 A\*6801 的初级锚定基元的确定中也相符。总而言之, 这些结果显示, A3 样初级锚定基元可定义为第 2 位的 A、I、L、M、S、T 或 V 和 C 末端的 R 或 K。

A3 样分子共有重叠的肽结合全部特征。 下一步检查 A3 样超型初级锚定基元使 A3 样超型分子之间的结合减弱的程度。收集了一组 200 个, 第 2 位带有 A、I、L、M、S、T 或 V 和 C 末端带有 R 或 K 的, 天然存在的 9 聚体肽序列。与限制每一种可能的锚定组合以单个氨基酸的天然频率的比例出现不同, 含有此组合的肽是从病毒和肿瘤抗原序列中随机选择。当检测每种肽结合纯化的 A3、A11、A31、A\*3301 和 A\*6801 分子的能力时, 显然, 独特的结合模式与每一种等位基因型相关 (结果未列出)。例如, 有些肽更选择性地只结合一种 I 型分子, 而另一些肽则更广泛的交叉结合 4 种或 5 种被试分子。

我们发现, 一般大约 10% (5% - 16%) 的肽 - HLA 组合与良好结合 ( $IC_{50} \leq 50nM$ ) 到任何给定的等位基因相关, 大约 17% (11% - 24%) 与中等结合 ( $IC_{50} 50-500nM$ ) 到任何给定的等位基因相关。高结



合和中等结合的频率与以前提到的含 A\*0201 集中测序基元的肽的频率相当 (Ruppert 等, 细胞 74: 929(1993))。然而, 最值得注意的是观察到相对高的交叉反应性。在能够结合最少一种 A3 样分子的 127 种肽中, 有 43 种 (34%) 结合 3 种或更多个 A3 样超型分子。4 种肽结合所有 5 个被测试的 A3 样分子。比较起来, 在被测试的结合 5 个无关的 I 型分子 (A\*0101、A3、A24 和 B7) 的 39 种肽中, 只有 3 种 (8%) 结合两个分子, 没有一种结合 3 个或更多的分子。5 种测试的 A3 样分子中的至少 4 种被鉴定为高或中等结合物的肽, 见表 2。在此表中, 好或中等结合能力是指  $IC_{50} \leq 500nM$ , 并用阴影加以突出。综合在一起, 这些结果证实, 在 A3 样超型分子的结合全部特征上有显著的重叠, 而且证实了 A3 样超型的初级锚定基元。

A3 样超型确定对潜在的肽配体提供简并性的次级锚定残基。如上所述, 虽然 A3 样超型分子的结合全部特征上的重叠是显著的, 但也远不是完全的; 每一种 A3 样分子保留相当程度的特异性。为了了解所观察到的交叉反应性的基础, 我们试图限定一个 A3 样超基元, 此基元能描绘与 A3 样超型简并性有关的肽配体的特征, 并在预测 A3 样简并的结合物时是有用的。

首先, 这里分析的概括了次级锚定结合特异性的每一种 A3 样等位基因的精确的基元 (A3、A11、A31、A\*3301 和 A\*6801), 象在材料与方法中介绍的那样得到。方法与以前用于确定一个精确的 A\*0201 基元的方法 (Ruppert 等, 细胞 74: 929(1993)) 类似。通过计算带有 20 种常见氨基酸的每一种的肽序列上每一个非锚定位置的平均相对结合能力而得到这些基元, 按照单个的化学相似性分组。用这种方法产生的结果的代表、A3 的计算值见表 3A。接着, 正如一个例子, 检测在其序列的第 3 位带有一个芳香族残基 (F、W、Y) 的 21 种不同的肽。这些肽对 A3 的平均相对结合能力, 是 200 个肽的总的平均值的 31.7 倍。与以前在 A\*0201 介绍的相类似, 优选的和有害的残基是指平均结合能力比总的平均值分别高 4 倍或低 4 倍的残基。因此, 在第 3 位的芳香族残基被认为是对 A3 结合“优选的”残基。

表 2. A3样超基元交叉反应性肽

肽	序列						结合能力 (IC50 nM) 结合的等位基因				
	123456	7	8	9	A3	A11		A31	A*33	A*68	
HIV pol 1227	YLAWVP	A	H	K	14	196	327	33	01	01	5
MAGE-2 299	ISYPPEL	H	E	R	32	214	23	375	81		5
HIV pol 1075	IVTWCK	T	P	K	12	16	18	97	242		5
HIV pol 1434	AVFIHN	F	K	R	67	3.3	3.2	109	42		5
HBV syw 713	IMPARE	Y	P	K	11	41	34	59	4103		4
HBV adr pol 601	RLADEG	L	N	R	10	245	144	197	3137		4
HBV ade "X" 1548	KVFLVG	G	C	R	26	73	30	608	2712		4
P. falciparum TRAP 611	LACAGL	A	Y	K	46	207	124	92	71		4
HBV adr pol 621	NVSIPW	T	H	K	15	79	211	260	92		4
MAGE-1 96	SIFRAV	I	T	K	2	2.2	2.8	1250	43		4
P. falciparum LSAI 94	QTNFKS	L	L	R	34	138	220	1657	63		4
HBV pol 398	GSTHVS	W	P	K	49	3.1	225	6744	77		4
HIV env 370	RAKANN	T	L	K	42	188	25	--	500		4
PSA 299	YTKVWH	Y	R	K	18	104	15	10	6.3		4
MAGE-1 66	TTINFT	R	Q	R	33	3	36	38	8.9		4

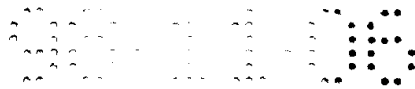
A.在初始分析中鉴定的A3样超基元交叉反应性肽

HBV X 69	CALPFT S A R	30	316	41	40	17	4
		14					
HBV adr pol 711	AVHHYF K T R	15	61	10	50	70	4
		57					
HIV pol 859	MTKILE P F R	--	378	82	70	25	4
<b>B. 另外的A3样超基元交叉反应性肽</b>							
c-ERB2 669	VWFGIL I K R	96	8	12	73	4	5
FLU A NP 166	LMQGST L P R	35	58	67	43	421	5
		5					
MAGE 1 66	TTINFT R Q R	25	16	21	47	5	5
		9					
MAGE 2 66	TTINYT L W R	12	5	71	245	6	5
		4					
HCV ENV1290	QLFTFS P R R	15	182	480	4496	3	4
MAGE3 226	SVLEVF E G R	--	44	129	50	69	4
HCV CORE 51	KTSEFS Q P R	69	94	81	1871	225	4
HBV adr POL 1257	HLYPVA R Q R	20	3077	92	330	410	4
c-ERB2 754	VIRENT S P K	28	462	180	363	--	4
C-ERB2 852	LVKSPN H V K	23	86	300	659	90	4



表 3 由相对结合值确定的等位基因特异性次级锚定优先位点

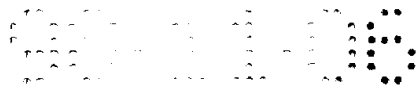
组别	位置								
	1	2	3	4	5	6	7	8	9
<b>A. A3</b>									
A	2.8		1.8	0.73	4.5	1.2	0.54	0.75	
G	0.52	A	0.53	2.4	0.59	1.2	1.4	0.64	A
DE	0.22	N	0.16	0.36	0.29	0.28	0.35	0.51	N
RHK	7.8	C	0.45	6.2	0.95	0.42	0.59	2.3	C
LIVM	1.1	H	1.0	0.43	1.4	1.7	1.6	0.41	H
YFW	0.45	O	32	4.2	3.3	29	2.2	0.90	O
QN	0.57	R	0.64	0.36	0.66	0.90	0.67	0.96	R
STC	1.3		0.82	0.50	1.1	2.0	1.1	3.2	
P	0.21		0.55	14	3.6	0.49	0.91	6.6	
<b>B. A11</b>									
A	7.9		0.86	0.85	5.0	2.0	0.18	0.30	
G	1.5	A	0.44	0.68	0.37	1.5	1.4	0.22	A
DE	0.12	N	0.30	0.37	0.48	0.30	0.46	1.4	N
RHK	1.3	C	0.35	1.9	0.47	0.37	0.37	2.1	C
LIVM	1.3	H	1.8	0.52	3.1	1.5	2.9	0.38	H
YFW	0.31	O	6.2	5.1	3.6	13	14	3.9	O
QN	0.65	R	0.92	0.73	0.48	0.88	0.59	0.98	R
STC	3.8		1.3	1.	1.4	1.2	0.38	1.8	
P	0.17		0.45	3.2	1.6	1.5	0.77	18	
<b>C. A31</b>									
A	3.9		1.3	0.28	0.11	0.38	1.1	7.1	
G	0.34	A	0.34	0.99	0.85	1.2	0.44	1.9	A
DE	0.06	N	0.10	0.50	0.24	0.24	0.16	0.22	N
	0								
RHK	11	H	3.3	1.5	2.0	2.0	1.7	0.79	H
LIVM	1.4	O	1.2	0.71	3.3	1.1	0.45	0.60	O
YFW	0.59	R	7.2	1.5	2.4	6.6	7.5	0.70	R
QN	0.57		0.49	1.7	0.44	1.4	0.82	1.6	
STC	2.1		1.1	0.67	1.5	1.1	2.7	2.2	
P	0.07		0.40	21	0.55	0.72	0.80	5.8	
	3								
<b>D. A*3301</b>									
A	0.48		0.62	1.4	0.74	0.65	4.9	1.2	
G	0.22	A	0.39	1.0	1.7	0.59	0.47	1.3	A
DE	1.5	N	0.24	1.1	0.36	0.43	0.54	0.41	N
RHK	0.64	O	2.3	1.0	0.88	1.4	1.1	0.76	O
LIVM	1.3	H	1.5	0.90	2.8	1.4	0.45	0.87	H
YFW	2.2	O	4.0	1.1	2.5	3.6	4.0	1.6	O
QN	1.3	R	1.7	0.77	0.44	0.72	1.2	1.4	R
STC	1.9		0.71	0.82	1.0	0.93	1.8	1.2	
P	0.21		0.41	2.7	1.4	1.2	0.85	3.6	



对每一种等位基因 (A3、A11、A31、A\*3301) 进行类似的分析，并用来得出每一位置上等位基因特异的次级锚定所需图 (表 3B-E)。检测的 A3 样超型的每种等位基因所获得的修饰的基元总结于图 1。每种分子表现出各自独特的次级锚定需要。例如，在第 4 位带正电荷残基 (R、H、K) 的被 A3 优选，而不被任何其它 A3 样分子优选。同样，在第 8 位，甘氨酸 (G) 与 A11 的低结合能力有关，而负电荷残基 (D、E) 只对 A31 有害。除了这些类型的独特的等位基因特异的特性外，某些残基与大多数的 A3 超型分子的低结合力或高结合力有关。例如，第 1 位的脯氨酸 (P) 对所有被检测的 5 种 A3 样分子都是有害的。第 7 位的芳香族残基 (F、W、Y) 和第 8 位的脯氨酸被 5 种被检测分子中的 4 种优选 (图 1)。

基于这些不同的单个精确的基元，构建了 A3 样超基元。对所考虑的 5 种等位基因的至少 3 种是有害的残基，确定为在超基元上有害的残基。相反，被所考虑的 5 种等位基因的至少 3 种优选，但对任何等位基因无害的残基，确定为优选残基。按这种方法得到的 A3 样超基元见图 2。

预测高交叉反应性肽的 A3 超基元功效的实验。为了检测以上确定的 A3 样超基元的正确性，检测了以前未用于分析超基元的另外一组 108 个肽对 A3、A11、A31、A\*3301 和 A\*6801 的结合。这组肽包括至少有一个优选的超基元残基而没有有害的残基 (超基元阳性) 的 30 个肽，至少有一个有害残基的 43 个肽 (超基元阴性)，和既没有优选的也没有有害的残基的 35 个肽 (超基元中性)。在 30 个超基元阳性的肽中，27 个 (90%) 结合到 2 个或更多的 A3 样分子上，16 个 (53%) 结合到 3 个或更多的分子上。相比之下，35 个超基元中性的肽中的 18 个 (51%) 结合 2 个或更多 A3 型，8 个 (23%) 结合 3 个或更多的分子。最后，超基元阴性的肽结合多个等位基因的能力低很多，6 个肽 (14%) 结合 2 个 A3 样分子，没有肽结合 3 个或更多的分子。这些结果与原来那组用来确定超基元的肽用同一类型的分析得到的结果性质相似，而与以前提到的对照肽结合无关的 HLA 等位基因所观察到的交叉反应性水平形成鲜明的对照。在以前的实验中，只有少数几个肽 (8%) 结合到一个等位基因上，而不是原始的限制元件上。从用于证实 A3 样超基元的那组肽中，证实了 10 个另外的肽以高或中等亲和力结合 5 个被试



A3 样分子中的至少 4 个 (表 2B)。

在所有主要人种群中保守的 HLA 超型的高表型频率。为了评价 HLA 超型大体上的潜在关联,特别是 A3 样超型的潜在关联,检测了各种 HLA I 型等位基因或抗原的发生率。到目前为止,现有的大多数 HLA-A 和 B 的数据是根据血清学分型得到的。这些数据在 DNA 序列确定的等位基因水平没有分辨率,因此不能区分亚型。然而,通过肽结合的研究 (del Guercio 等, 免疫学杂志 154: 685(1995); Tanigaki 等, 人类免疫学 39: 155(1994)), 或集中测序分析 (Fleischer 等, 组织抗原 44: 311(1994); Rotzschke 等, Eur 免疫学杂志 22: 2453(1992)), 或基于一级序列的口袋结构的分析, 比较亚型的肽结合特异性显示, 大多数例子的亚型有非常相似的, 即使不完全一致的, 肽的主要锚定特异性。因此在以下的分析中, 如果没有 DNA 亚型水平的群体数据, 但结合的数据、发表的基元或序列分析认为亚型有重叠的肽结合特异性, 就可以假定在亚型等位基因和血清学确定的抗原之间有一对一的相关。

检测不同的人种背景中各种 A3 样等位基因的发生率时, 很明显每一种单个的等位基因或抗原的频率在人种群之间变化很大 (Imanishi 等, 第 11 届国际组织相容性专题讨论会论文汇编, Vol. 1, Tokyo, Oxford University Press(1992)), 5 种 A3 样等位基因的累积频率是非常恒定的 (在 37% 至 53% 之间)。例如, A3 在高加索人、北美黑人和西班牙人中常见, 但在日本人中几乎没有。相反, A31 在日本人中常见而在高加索人和北美黑人中很少见。相比之下, 在 5 种被检测的人群中的每一种中, A3 样超型在至少 37%、高的达 53% 的个体中存在。

在本文中也有兴趣指出 A3 样超型的存在不是一个孤立的事情, 因为最近我们已经报道了 A2 样 (del Guercio 等, 免疫学杂志 154: 685(1995)) 和 B7 样 (Sidney 等, 免疫学杂志 154: 247(1995)) 超型的存在。每种这些另外的超型也是非常突出的, 在不同人种背景中以十分恒定的累积频率 (在 40% 至 60% 之间) 存在。事实上, 能够容易地在基因水平计算出, 现存的 HLA-A 或 B 基因总拷贝的至少一半, 似乎属于这 3 种 HLA 超型中的一种或另一种。

#### 讨论

这里发表的结果证实来自至少 5 种不同的 HLA 等位基因 (A3、A11、

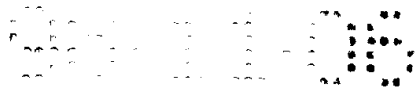


A31、A\*3301 和 A\*6801), 和依据口袋结果分析预示的可能的至少 3 种其它的等位基因 (A\*3401、A\*6601 和 A\*7401) 产物, 可以归类为一个功能性 A3 样超型。这个确定是基于许多观察作出的。作为一个组, 这些分子 (a) 在它们的肽结合区内共有一定的关键结构特征, (b) 它们偏爱类似的初级锚定, (c) 共有大部分重叠的结合性能。通过检查 5 这些等位基因分子优选的、大量的带锚定残基的肽的结合活性, 也确定了 A3 样超基元。这种超基元, 是基于每一种 A3 样超型分子的次级锚定需要的详细图得到的, 可以有效地预测 A3 样简并的结合肽。最后, A3 样超型, 和通用的超型, 在所有主要人种群中以非常高的表型频率 10 出现。照此, 基于肽结合特异性的 HLA I 型超型, 代表了解 HLA I 型分子之间相互关系的血清学和系谱分类的一种功能性的替代。

除了用于 A3 样超基元的产生外, 本研究公开的单个次级锚定图本身对了解肽结合到 I 型分子上有显著作用。由于这些图是用大小均一的肽得到的, 因此在每个次级位置上的优选决定基可能比天然加工肽的集中测序得到的更加准确。这里确定的基元也可以确定对肽结合起 15 有害作用的残基。

Barber 和他的同事 (Barber 等, 当代生物学 5: 179(1995)) 已经证实, 肽可以被我们归于 HLA-B7 样超型中的两个分子识别, 两个其它的肽被一个以上的 A3 样等位基因识别 (Missale 等, 实验医学杂志 177: 751(1993); Koenig 等, 免疫学杂志 145: 127(1990); Culmann 等, 免疫学杂志 146: 1560(1991)) (见表 4)。使用一种体内诱导初级 CTLs 的方法 (Wentworth 等, 分子免疫学 32: 603(1995)), 我们观察到能够被 A3 和 A11 识别的肽的几个例子。我们检测了 A3 样超型限制性表位对 A3 样超型分子的结合能力, 并注意到相对高水平的简并性。 20 在表 4 所列的 7 个表位中, 只有一个是九碱基顺序, 能够分析在图 2A 中提出的超基元 (将来的研究目标是把超基元扩展到长于 9 个残基的肽)。这个肽是超基元阳性的, 并结合 5 个 A3 样分子中的 3 个。然而, 每种表位与 A3 样超型的初级锚定特异性相符是重要的。

比较我们基于肽结合而提出的超型的分类与基于 DNA 序列 (和血清学反应性) 的 HLA-A 等位基因的分类的关系 (Ishikawa 等, 人类免疫学 39: 220(1994); Firgaira. 等, 免疫遗传学 40: 445(1994); Karo 等, 免疫学杂志 143: 3371(1989)) 揭示了相似和差异。例如, 30



HLA - A3 和 A11 似乎关系密切, 而且来自一个共同的祖先基因 (48 - 50)。然而, A31 和 A33 来自含有 A2/A10/A19 组的古老谱系, 这个谱系与 A3 和 A11 的谱系不同。最后, HLA - A\*6901 属于 A28 HLA 进化组 (Fernandez-Vina 等, 人类免疫学 33: 163(1992); Ishikawa 等, 人类免疫学 39: 220(1994); Lawlor 等, 免疫学年评 8: 23(1190)), 也含有 HLA - A\*6802 和 - A\*6901 等位基因。然而, 基于它们的肽结合特异性, HLA A\*6801 是 A3 样超型的一员, 而 A\*6802 和 A\*6901 已证实属于 A2 样超型 (del Guercio 等, 免疫学杂志 154: 685(1995))。因此, 基于现有的 HLA 等位基因的表型树 (Ishikawa 等, 人类免疫学 39: 220(1994); Firgaira. 等, 免疫遗传学 40: 445(1994); Karo 等, 免疫学杂志 143: 3371(1989)), 在两种古代 HLA 谱系中发现 A3 样等位基因: 含有 A3 和 A11 的 A1/A9, 和含有 A31、A33 和 A\*6801 的 A2/A10/A19。如果 HLA - A3 样超型的存在是共同祖先的反映, 那么 A3 样基元事实上代表原始人 HLA I 型肽结合特异性, 其它的特异性代表对变化的致病环境的适应。

表 4 A3样限制肽的结合

肽	序列	结合能力 (IC50 nM)	
		A All A31 A*3301	A*6801
	1 2 3 4 5 6 7 8 9 11 0 1	A All A31 A*3301 3	A*6801
HBVc 141-151'	STLPETTIVRR A31 Aw68	3 4 181 181	26
HIV nef 73-82 <sup>b</sup>	QVPLRPMTYK A3, A11	1 10 183 2164 8 7	133
MAGE 1 96-104	SLFRAVITK A3, A11	3 2 248 10943 3	113
HIV env 43-52	TVYYGVPPVK A3, A11	1 4 163 10357 1 6	15
HCV lorf 1858-1867	GVAGALVAFK A3, A11	2 4 330 27188 8 3	119
HIV env 42-52	VTVYYGVPPVK A3, A11	8 11 461 38667 5 5	172
HBV pol 152-161	TLWKAGILYK A3, A11	2 17 292 30526 7	533

<sup>b</sup> 良好结合能力是指<50nM, 以阴影突出。



然而, 这些观察也激起引起兴趣的可能性, 即可能在群体水平, 存在一个维持各种 HLA I 型超型等位基因的高频率的选择优势, 这种选择优势由共同的祖先或趋同进化产生。这种优势可能与群体水平有效肽结合特性的产生相关。这种观察明显来自这个事实, 即来自同一个系统发育家族的不同分子可能属于不同的结合超型家族。也有可能这些现象部分地与人转运相关抗原加工分子 (TAP) 的肽特异性的最佳利用有关 (Androlewicz 等, 美国国家科学院院刊 90: 9130(1993); Androlewicz 等, 免疫 1: 7(1994); van Endert 等, 免疫 1: 491(1994); Heemels 等, 免疫 1: 775(1994); Momburg 等, 当代免疫学视点 6: 32(1994); Neefjes 等, 科学 261: 769 (1993) ), 这些 TAP 分子优先转运带有一定序列特征的肽, 如疏水的、芳香族的、或 C 端带正电荷的。最近由 van Endert 和他的同事与我们合作进行的研究, 评价了大量肽的 TAP 相对亲和力, 并且介绍了扩展的 TAP 结合基元。令人惊奇的是, 这个基元含有许多与 A3 样超基元相关的结构特征, 如在 9 聚体肽的第 3 位和第 7 位优选芳香族残基, 而且在第 1 位和第 3 位没有负电荷残基, 在第 1 位没有 P。

基于这里提供的实验证据, 似乎除了基于血清学或系统发育关系的分类外, HLA I 型等位基因也可以根据它们配体的特异性归类到超型。最近已鉴定了 3 个超型: A2 样、A3 样和 B7 样, 其它的, 如 B44 样超型已在文献中提出 (Fleischer 等, 组织抗原 44: 311(1994); Harris 等, 免疫学杂志 151: 5966(1993); Thorpe 等, 免疫遗传学 40: 303 (1994); Falk 等, 免疫遗传学 41: 162(1994); Falk 等, 免疫遗传学 41: 165 (1994) )。虽然尚不知道还有多少超型将被鉴定, 也不知道它们包括哪些种类, 但现有的结果证实肽结合特异性简并的现象, 以前认为是限制于 II 型 HLA (Panina-Bordignon 等, Eur 免疫学杂志 19: 2237(1989); O'Sullivan 等, 免疫学杂志 145: 1799(1990); Busch 等, 国际免疫学 2: 443(1990) ), 也是肽结合到 I 型 HLA 的特征。不管 I 型超型出现的原因是什么, 都应该强调它们潜在的实际的关联。定量结合检测法的可行性和详细的超基元可以鉴定高交叉反应的肽。接着这个可以用几种混合的 CTL 表位广泛地覆盖, 也可能对潜在地使用以肽为基础的方法发展疫苗有重大的意义 (Vitiello 等, J Clin Invest 95: 341(1995) )。



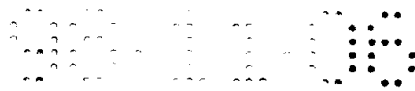
## 实例 2

### B7 样超型结合

以前的工作 (Sidney 等, 免疫学杂志., 154, 247-159(1995); Hill 等, 自然 360: 434-439(1992); Falk 等, 免疫遗传学 38: 161-162(1993b); Barber 等, 当代生物学 5: 179(1995); Schonbach 等, 免疫学杂志., 154: 5951-5958(1995)) 显示, 统称为 B7 样超型的 HLA B 特异性的较大家族, 以常见的肽结合基元为特征 (第 2 位的 P、C 末端的疏水残基或芳香族残基)。在这个例子中, 用以上介绍的分子结合检测法详细检测 5 种最常见的 B7 样等位基因 (B\*0702、B\*3501、B51、B\*5301 和 B\*5401) 初级锚定 (第 2 位和 C 末端) 的特异性。

为完成这个, 我们合成并检测了一系列 FHV nef 84-92 肽 (序列 FPVRPQVPL) 的单个置换类似物的结合。HIV nef 84-92 以高 (IC<sub>50</sub>≤50nM) 或中等 (IC<sub>50</sub> 50-500nM) 亲和力结合 B\*0702、B\*3501、B51、B\*5301 和 B\*5401。发现 B7 样超型分子有共同的而更严格的第 2 位特异性。对于所有的 5 个等位基因, 脯氨酸都是优选的残基。除了一个例外 (在 B\*3501 中的 A), 所有被检测的第 2 位置换物比在 C 末端带脯氨酸锚定的亲本肽的结合亲和力降低 10 倍以上。相比之下, 当分析结合特异性时, 每一种 HLA-B 型在 C 末端表现出完全不同的特异性模式。例如, B\*0702 优选 M、F 和 L, 而 B\*5101 优选 L、I 和 V。除了这些差异以外, 所有的 C 末端特异性模式呈现很大程度的重叠。脂肪族残基 I 和 V 至少被 5 种分子中的 4 种优选, A、L、M、F 和 W 在大多数例子中优选或耐受。其它残基, 如 Y 或 T, 只被个别的例子耐受, 有些 (如 K 或 D) 根本不被耐受。

这些关于初级锚定特异性的结果与 Sidney 等, 免疫学杂志, 154, 247-259(1995) 公开的结果完全相符。能够简并 B7 样超型结合的肽, 在第 2 位带有脯氨酸, 在 C 末端带有疏水或芳香族残基 (V、I、L、M、F、W、A)。在正式确定 B7 样超型初级锚定基元时, 我们也保守地包括 Y, 尽管它相对缺乏简并性, 因为 Y 在集中测序分析时构成 B\*3501 的优势信号 (Hill 等, 自然 360, 434-439(1992); Falk 等, 免疫遗传学 38.161-162(1993b); Schonbach 等, 免疫学杂志 154: 5951-5958(1995))。总而言之, B7 样超型的初级锚定基元确定为第 2 位的 P, 和 C 末端的 A、I、L、M、V、F、W 和 Y。



### B7 样超型配体的优选大小

I 型分子通常优选长度为 8 至 10 个残基的肽 (Falk 等, 自然 351, 290-296 (1991)), 虽然更长的肽也表现结合 (Massale 等, 实验医学杂志 177: 751(1993); Chen 等, 免疫学杂志. 152: 2874(1994); Collins 等, 自然 371-626(1994)). 为了确定对 B7 样超型最佳的肽长度, 合成了几组 8-、9-、10- 和 11- 残基的肽, 它们代表了天然存在的病毒、肿瘤或细菌序列, 并且每种都带有上述的 B7 样超型初级锚定特异性, 对它们的 B7 样每个类型 6 种结合能力进行了分析。结论是 9 残基代表所有被检测的 B7 样分子的最佳肽长度。这个评估在任何分子结合每种大小的肽的百分率方面, 和在观察到的交叉反应程度方面都是真实的 (结果未列出)。

### B7 样等位基因的次级锚定基元和 B7 样超基元

其它的残基可以作为次级锚定, 因此对肽提供补充的结合能量 (Ruppert 等, 细胞 74: 929-937(1993); Madden 等, 细胞 75, 693-708(1993); Saito 等, 生物与化学杂志 268, 21309 (1993); Sidney 等, 人类免疫学 45, 79-93(1996); Kondo 等, 免疫学杂志 155: 4307-4312 (1995); Parker 等, 免疫学杂志 152, 163-175(1994)). 也已经证实, 一些残基可能对肽结合到 I 型分子有负作用 (Ruppert 等, 细胞 74: 929-937(1993); Sidney 等, 人类免疫学 45, 79-93(1996); Kondo 等, 免疫学杂志 155: 4307-4312 (1995); Boehncke 等, 免疫学杂志 150, 331-341(1993)).

为了建立能有效地选择有简并 B7 样结合能力的肽的 B7 样超基元, 我们用这里介绍的方法, 试图确定参与肽结合到 B7 样分子的次级锚定和次级作用。确定了 5 种最常见的 B7 样分子, B\*0702、B\*3501、B51、B\*5301 和 B\*5401 的结合能力, 这 5 种分子包含 199 个 9 聚体肽, 代表含有 B7 样超型初级锚定 (第 2 位的脯氨酸、C 末端的 A、V、I、L、M、F 和 W) 的天然存在的病毒序列; 并分析了结果。计算了带有 20 种氨基酸中每一种的肽的每一位置的平均相对结合能力 (ARBC), 并与完整的肽的 ARBC 进行比较。因为一些氨基酸很少出现, 因此, 按照以前介绍的 (Ruppert 等, 细胞 74: 929-937(1993)) 按单个化学相似性对残基分组。这个分析对 B\*0702、B\*3501、B51、B\*5301 和 B\*5401 分别进行。



发现优选的和排斥的模式，对于次级锚定而言，每个分子表现的相当不一致。例如，在实验组中，18个肽在第1位有正电荷残基（R、H或K）。这些肽作为一组，是很好的 B\*0702 结合物，ARBC 为 21。然而，对于 B51，同样的肽是相对弱的结合物，ARBC 为 0.25。然而也可以注意到优选的很大的相似性。例如，在第1位带有芳香族残基（F、W和Y）的肽作为一组，是 B7 样超型分子组的良好结合物，对 B\*0702、B\*3501、B51、B\*5301 和 B\*5401，ARBC 分别为 4.2、17、16、20 和 70。

以上讨论的数值，接着用来获得每个位置等位基因特异的次级锚定所需条件的图谱。为完成这个，优选的和有害的残基定义为 ARBCs 比总的平均值分别高 3 倍或低 3 倍的残基。这些优选的和有害的作用总结于图 3。次级锚定图更清楚的揭示，尽管每个分子表现自己独特的次级锚定需要，但一些特征在 B7 样分子之间高度保守。例如，如上所示，在第1位的芳香族残基（F、W和Y）被所有 5 种 B7 样分子优选。相反，第8位的酸性残基（D、E）与所有 5 种分子的弱结合能力相关。

被所考虑的 5 种等位基因中 3 种或更多优选，但对任何分子无害的次级效应，或另一方面，对 5 种分子中的 3 种或更多种有害的次级效应称为共有的效应。然后这些共有的特征，整合到扩展的 B7 样超基元限定的，与大多数 B7 样超型分子弱或好结合相关的残基中。

按照这个逻辑，预计带有优选的次级残基超基元的肽，将比不带优选残基或带有害残基的肽表现出更大程度的 B7 样超型简并性。通过确定一组独立的带有 B7 样初级锚定特异性的肽的结合交叉反应性而检测这种假设。正如预计的，超基元阳性的肽（即，至少一个超基元带优选的次级残基，并没有有害残基的肽），比超基元阴性的肽（一个或更多的超基元有有害的残基）在 B7 样超型内表现出明显更大程度的交叉反应性（结果未列出）。

#### 提高结合能力的超基元的完成

正如这里证实的，HLA 超基元在预测高度简并的肽可能是有价值的。然而，最有趣的是，HLA 超基元的确定，应该也能够通过鉴定肽序列内哪些残基可能被归类或“固定”，而设计高度交叉反应的表位，并在一个超型内使肽有更大的简并性。

为评估这种可能性，选择了在 B7 样超型内显示高度简并性的 6 个肽（表 5）。每一种肽以高（ $IC_{50} < 50 \text{ nM}$ ）或中等（ $IC_{50} 50-500 \text{ nM}$ ）



亲和力结合 5 种最常见的 B7 样分子中的至少 3 种。这些肽在以上介绍的 B7 样超基元和等位基因特异的次级锚定基元的范围内进行了分析，以确定在其序列内的特定的残基是否会被“固定”，而进一步加强它们结合到 B7 样超型分子上。这个评估发现所考虑的特定的肽中，没有一个含有超基元阴性残基。3 种肽(HCV 核心 168、MAGE2 170 和 MAGE3 196) 5 每种带有一个对单个 B7 样分子有害的残基(表 5)。

下一步，合成一组单个置换的类似物。某些含有次级锚定置换的类似物根据这里公开的数值选择，这些置换为超基元阳性，或在特定的等位基因是阳性，对任何其它置换无害。因为 C 末端初级锚定的优选对每种等位是独特的，因此也考虑在这个位置的置换。因此，例如 10 为了检测是否能够增加简并性，用超基元阳性 F 置换第 1 位的天然残基制备许多类似物。其它的置换，如 HBV pol 541 C 末端的 L 置换 Y，用以说明亲本肽对 B\*0702 和 B\*5401 的弱结合。

当检测这组肽与 B7 样超型分子的结合能力时，产生表 5 所示的结果。在每种情况下，第 1 位被 F 置换时，表现出比亲本序列增强的结合和/或简并性。例如，MAGE 2 170 以高的亲和力结合到 B\*0702 上， 15 以中等亲和力结合到 B\*3501、B51 和 B\*5301 上，但以弱亲和力结合到 B\*5401 上。此肽的 F1 类似物以高亲和力结合所有 5 种分子。

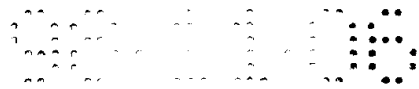
针对特异分子的置换的成功很难加以概括。例如，HBV pol 541 C 20 末端 L 置换天然的 Y 在结合到 B\*0702 上是成功的，同时增加结合到其它分子上的亲和力(在 B51 和 B\*5401 最显著)。在另一些例子中，观察的效果不象预期的那样，正如在 HBV env 313 证实的那样。此肽以高亲和力结合到 B\*0702、B\*3501、B51 和 B\*5301 上，但只很弱地结合到 B\*5401 上。基于这个观察，即第 5 位的脂肪族残基(L、I、V 和 25 M)对 B\*5401 是阳性的，而对其它分子是相对中性的，因此制备第 5 位是 M 的类似物以提高 B\*5401 的结合。然而，正如表 5 所示，用 M5 类似物获得的显著提高的 B\*5401 的结合能力是以降低对 B\*0702、B51 和 B\*5301 的结合为代价的。

虽然单个类似物的成功是可变的，但值得注意的是，对于每一种 30 情况，至少一种类似物能够提高结合亲和力，或扩展亲本肽的简并性。因此，已经简并的肽能够被独立地“固定”以提高其结合能力并扩展其简并性。

表 5

独立的置换提高肽配体的B7样超型结合能力和简并性

来源	结合							x-rxn							
	1	2	3	4	5	6	7		8	9	B*0701	B*3501	B*5101	B*5301	B*5401
HIV nef 84	F	P	V	R	P	Q	V	P	L	16	43	12	481	71	5
	F	P	V	R	P	Q	F	P	L	9.5	23	207	319	241	5
	F	P	V	R	P	Q	V	P	I	22	120.0	15	31	1.6	5
HBV ENV 313	I	P	I	P	S	S	W	A	F	42	2.6	2.3	12	2970	4
	F	P	I	P	S	S	W	A	F	24	1.2	305	1.7	105	5
	I	P	I	P	M	S	W	A	F	229	2.2	580	930	5.9	3
	I	P	F	P	S	S	W	A	F	186	1.7	16	3.8	826	4
	I	P	I	P	S	S	W	A	I	31	54	15	24	7.7	5
HBV POL 541	F	P	H	C	L	A	F	S	Y	-	14	83	17	503	3
	F	P	H	C	L	A	F	A	L	25	2.7	28	5.0	24	5
	F	P	H	C	L	A	F	S	L	74	2.4	4.5	15	7.7	5
	F	P	F	C	L	A	F	S	Y	-	6.5	27	4.8	5.1	4
	F	P	H	C	L	A	F	S	I	675	29	6.3	3.8	1.0	4
	F	P	H	C	L	A	F	S	A	3667	6.5	250	137	0.6	4
HCV Core 168	L	P	G	C	S	F	S	I	F	28	90	100	114	6897	4
	F	P	G	C	S	F	S	I	F	19	1.6	132	3.2	67	5
	L	P	V	C	S	F	S	I	F	8.7	2.5	13	2.7	5128	4
	L	P	G	C	M	F	S	I	F	24	28	53	39	2778	4
MAGE2 170	V	P	I	S	H	L	Y	I	L	22	171	96	238	3175	4
	F	P	I	S	H	L	Y	I	L	16	7.3	6.4	7.0	28	5
	V	P	I	S	M	L	Y	I	L	164	273	73	1075	1493	3
	V	P	I	S	H	L	Y	I	I	108	5333	8.3	326	91	4
MAGE3 196	M	P	K	A	G	L	L	I	I	940	5039	393	90	248	3
	F	P	K	A	G	L	L	I	I	162	1303	5.8	60	150	4
	M	P	V	A	G	L	L	I	I	86	66	1.0	2.3	112	5
	M	P	K	A	M	L	L	I	I	1528	-	5.8	186	24	3
	M	P	F	A	G	L	L	I	I	229	1.0	0.9	2.3	0.27	5



总之，根据这里显示的结果，很明显更高亲和力的结合肽可以通过用中性或优选的残基置换序列中的“阴性”或中性的残基而产生。这些肽也可能具有独特的免疫学特性，即尽管与野生型序列仍然有交叉反应，但不会发生耐受、删除或抑制机制，不会灭活对野生型序列有反应的 CTL，并由于癌症的慢性感染而呈递。同样的技能可以用来产生有更高程度交叉反应性，因此更有利群体覆盖的肽。

### 实例 3

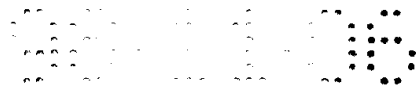
#### A2 样超型结合

我们也进一步获得了关于 A2.1 结合的结构要求的信息。为完成这个，我们试图确定第 2 位和第 9 位锚定的可行程度。为了这个目的，合成了一组在第 2 位或第 9 位带单个置换的多聚 (A) 9 聚体模型肽的类似物，此模型肽含有以前报道的 (Ruppert 等，细胞 74: 929-937(1993)) 在第 2 位是 L、第 9 位是 V 的 A2.1 基元，并测定了结合能力。合成了第 2 位和第 9 位锚定的一共 13 个不同的类似物。

与以前报道的 A2.1 基元十分相符，在第 2 位带有 L 或甲硫氨酸 (M) 的肽是最好的结合物。即使用相对保守的置换，如异亮氨酸 (I)、V、丙氨酸 (A) 和苏氨酸 (T) 也引起结合能力的显著降低 (10 至 100 倍)。更加根本的改变 (即 D、K、F、C、P、G、N 和 S 残基) 会完全消除结合能力。在第 9 位得到相似的结果，在此位置上只有保守的置换，如 L 和 I，在 10 倍以内结合未置换的模型多聚 A A2.1 肽结合物。带有 A 或 M 置换的类似物也结合，但强度较低 (降低 10- 至 100- 倍)。最后，所有其它被测置换 (T、C、N、F、S、G、P 和 R) 与 A2.1 结合能力的完全丧失相关。因此，基于这些结果并与以前的研究 (19-20) 十分相符，“典型的” A2.1 基元可以被鉴定为在第 2 位是 L 或 M，第 9 位是 L、V 或 I。

从这些结果我们推测，任何在第 2 位或第 9 位 (或第 10 位) 带有“非典型的” (但仍能接受的) 残基的肽 (例如第 2 位是 I、V、A 或 T，第 9 或 10 位是 A、M 或 T) 的 A2 结合能力，可以通过用更“典型的”锚定取代而提高。下面举出一些例子说明实际上这是怎样完成的。

例如，序列为 LWVTVYYGV 的 FHV Env 2181 肽以 12500nM 的 IC50 % 结合 A2.1，而第 2 位的锚定被置换的类似物 LMVTVYYGV 以 3.3nM 的 IC50 % 结合。天然存在序列 FLPSDFFPSI 的 HBVc 以 22nM 的 IC50 % 结



合 A2.1, 但其 C 末端锚定被置换的  $V_{10}$  变型以 2.5nM 的  $K_d$  结合 A2.1。最后, HBV pol 538 肽 (YMDDVVLGA) 以 200nM 的 IC50% 结合 A2.1, 而  $V_9$  变型以 5.1 nM 的 IC50% 结合。其它固定的锚定肽的例子见表 6。检测了某些固定肽诱导 CTL 反应的能力。例如, 用初级 CTL 检测法 (21) 检测了 HIV Env 2181 肽和 HBV pol 721 肽, 发现是阳性。阳性 CTL 识别结果也见于 HBVc18-27 和 HBV pol 538 肽。

进一步的实验揭示了非锚定残基在确定结合能力方面的突出作用。这些分析的结果也在 Ruppert 等 (同上) 中有介绍。按照我们的分析, 一个给定的氨基酸组别在 A2.1 结合物上的频率被非结合物的频率相除得到频率比率。这个比率说明一个给定的残基组别是否优先存在于结合物 (比率 > 1) 或非结合物 (比率 < 1) 的给定的位置上。为了便于分析, 设定了比率的临界水平, 使在结合物上的频率比非结合物上高 4 倍以上的残基视为是有利的或优选的残基, 而在结合物的频率比非结合物上低 4 倍以上的残基视为是不利的或有害的残基。按照这个方法, 鉴定了与 A2.1 结合能力或结合缺失有显著关系的残基组别。通常, 最不利的作用见于带电氨基酸。在第 1 位, P 和酸性残基 (E 和 D) 在 A2.1 结合肽上不常见。在第 6 位非结合肽是碱性 (R、H 或 K) 残基, 而酸性和碱性残基在好的结合物的第 3 位和第 7 位不常见。相反, 第 1、3 和 5 位的芳香族残基与高亲和力结合相关。而且带有侧链的 OH- 或 SH- 残基, 如 S、T 或 C 偏爱第 4 位, 而 A 偏爱第 7 位, P 偏爱第 8 位。总之, 不同氨基酸组别的频率分析使我们能够更准确地确定 A2.1 基元, 这个基元考虑了第 2 和第 9 锚定位置之外的其它位置对 A2.1 结合亲和力的影响。

#### 10 聚体 A2.1 配体的分析

以上介绍的用于 9 聚体肽的同样的方法也用于分析由一组 10 聚体肽获得的结果。在 10 聚体肽的 N- 和 C- 末端, 观察到的模式与 9 聚体肽观察到的十分相似。例如, 象在 9 聚体肽中的一样, 在 10 聚体肽组中, 第 1 位的特征是, 在结合物组中芳香族残基的频率升高, 而负电荷残基和 P 同样与弱结合相关。在第 3 位, 也是负电荷与弱结合相关。有趣的是在这个位置上, 脂肪族 (不是芳香族) 残基与高亲和力结合相关。在此肽的 C 末端也观察到一定的相似性。在 10 聚体肽中, 倒数第二位即第 9 位 (与 9 聚体肽的第 8 位对应) 是相当随意的, 只



有碱性残基在非结合物上出现频率较高。与 9 聚体肽的第 7 位的情况相似，10 聚体的倒数第 3 位，即第 8 位既不容忍正电荷残基也不容忍负电荷残基。同样，在 10 聚体的第 7 位不喜欢正电荷残基，这与以前观察的 9 聚体的第 6 位一样。与第 3 位观察到的相比，与高结合相关的残基是不同的。芳香族和疏水性残基在高亲和力结合物的第 8 位常见（这与 9 聚体上第 7 位只有 A 常见不同）。

最后，在肽的中间观察到一个明显有特色的模式。在第 4 位，G 被高结合物偏爱，而 A 和正电荷残基非常常见于非结合物。在 A2.1 结合物的第 5 位完全没有 P。值得注意的是在 10 聚体肽的第 4 和 5 位观察到的趋势，没有一个与 9 聚体肽的第 4 或 5 位相对应。

概括地说，按照以上在 9 聚体肽中介绍的相似的策略，可以产生 A2.1 10 聚体肽的详细的基元。比较 9 聚体肽和 10 聚体肽的非锚定位置，可以发现重要的差别和显著的相似性。

总之，根据这里显示的结果，很明显更高亲和力的结合肽可以通过用中性或优选的残基置换序列中的“阴性”或中性的残基而产生。这些肽也可能具有独特的免疫学特性，即尽管与野生型序列仍然有交叉反应，但不会发生耐受、删除或抑制机制，不会灭活对野生型序列有反应的 CTL，并且由于癌症的慢性感染而呈递。

肽抗原结合到多个 HLA 等位基因上以确定 A2 样超型。

已经建立了用放射标记的肽和表达 HLA I 型的哺乳动物细胞，如 EBV 转化的 B 细胞系和 PHA 活化的母细胞的直接 MHC 结合检测法。如果靶细胞在有  $\beta 2$  微球蛋白存在时  $26^{\circ}\text{C}$  预孵过夜，可以获得放射标记探针的显著结合。在这些条件下，每个细胞类型表达的 HLA 分子的小部分都可以被标记的肽结合。用这些检测法可以检测  $A^{*}0201$  限制的乙型肝炎病毒核心 18-27 肽与其它 A2 亚型的交叉反应程度。已经确定此肽表位也结合  $A^{*}0202$ 、 $A^{*}0205$  和  $A^{*}0206$  但不结合  $A^{*}0207$  亚型。用数组合成的肽类似物进行的抑制实验强调了 HLA -  $A^{*}0201$ 、 $A^{*}0202$  和  $A^{*}0205$  等位基因的类似配体特异性。有助于形成不同 HLA 等位基因的 B 和 F 口袋的多形残基的分析可以预测乙型肝炎病毒核心 18-27 表位与 2 个其它的 HLA 等位基因的结合（HLA -  $A^{*}6802$  和  $A^{*}6901$ ）。因此，统称为 A2 超型的至少 6 个不同的 HLA - A 分子组成的一个家族，似乎共有重叠的配体特异性（在第 2 位和 C 末端的脂肪族残基）。这些结果显



示广泛交叉反应的肽表位可以被鉴定，并可以极大地提高肽基接种方法的预期可行性。

而且，这些结果显示置换的类似物的另外的用途。特别是同时提高 A2 超型的数个成员的结合亲和力。

5 用 HPV16EF. 86 - 93 肽、TLGIVCPI 及其类似物 TLGIVXPI (这里 X 代表  $\alpha$  氨基丁酸) 为例，说明了这是怎样完成的。然后检测了这两种肽对 A2 样等位基因的结合模式。显然 X 置换极大地提高了对所有 A2 样等位基因的结合亲和力，而得到一个更有用的肽。其特征是比其原来的亲本序列有更高的结合能力和更广泛的交叉反应性。用 A2/Kb 转  
10 基因小鼠随后进行的实验证实，X 置换肽诱导的 CTL 与野生型序列完全交叉反应，而且 X 肽，由于它的更高的结合亲和力，是一个更有效的免疫原。

#### 实例 4

#### A24 和 A1 结合

#### 15 次级作用和 A24 结合

我们使用一个多聚 (A) 9 聚体模型肽，此肽含有在第 2 位是 Y、第 9 位是 F 的 A24 特异的基元。发现第 2 位不仅接受 Y 也接受 F、M，可能还接受 W。在 9 或 10 残基配体的 C 末端，F 和 W 是最优选的，但 L 和 I 也能接受。预计 M 在这个位置也能接受 (Kubo 等，免疫学杂志 152: 3913 (1994))。从这些结果，我们得出结论，在第 2 位或第 9 位 (或第  
20 10 位) 带有一个非典型的，但仍能被接受的残基 (例如在第 2 位是 F、M、W 或第 9 (或 10) 位是 L、I、M) 的任何肽的 A24 结合力，可以通过用一个更典型的锚定取代这些残基而提高。

Kondo 等，免疫学杂志 155: 4307-4312 (1995) 介绍了非锚定残基  
25 在决定 A24 结合能力的突出作用的进一步实验的结果。收集了所有 A24 结合的数据库，并且计算了带有特定残基的肽的每个位置的相对平均结合亲和力。

在 9 聚体肽的情况下发现，例如在第 1 位带 G 或负电荷残基 (D、E) 的肽倾向弱结合，平均亲和力比所分析的总共 141 个 9 聚体肽的平均亲和力低 10 倍。相比之下，在第 1 位带芳香族残基 (F、Y、W) 的  
30 肽结合良好，平均亲和力比总的平均亲和力高 11.8 倍。在第 1 位带正电荷残基 (R、K、M) 的肽也倾向良好结合，平均亲和力比总的平均亲



和力高 4.6 倍。

也检测了数个其它位置对 A24 结合能力的消极的次级作用：第 3 和 6 位的 D 或 E，第 4 和 7 位的 G，第 6 位的正电荷残基 (K、R、H)，第 8 位的 A，第 5 位的 P，和第 5 和 8 位的酰胺 (Q 和 N)。相反，发现第 7 或 8 位的芳香族残基 (Y、F、W) 对 A24 结合有利，第 4 位的小的氢键残基如 (S、T、C) 有积极的作用。

显然，9 聚体序列上的每个单个的位置都可能影响 A24 的结合。还有趣的是疏水残基 (F、W、Y、L、I、V 和 M) 从不与弱结合相关。

用 10 聚体肽获得的结果也进行了类似的分析。与上一节的结果相似，在 10 聚体肽中也发现数个次级作用。

象 9 聚体肽的情况那样，第 3 和第 6 位的负电荷残基 (D、E) 与弱结合相关。然而，一般来说，10 聚体的次级作用图与 9 聚体的完全不同。例如，在 9 聚体肽中，P 在任何位置都与显著提高结合不相关，但在第 5 位时，与降低结合相关。对于 10 聚体，P 在 10 聚体肽配体第 4、5 或 7 位与提高结合能力相关。

在 10 聚体中，第 5 位似乎对次级作用最重要，因为 Y、F 和 W (除了已经提到的 P) 与好的 A24 结合相关，而 R、H 和 K 与弱的结合能力相关。第 7 和 9 位存在 A 和第 4 和 8 位存在酰胺 (Q、N) 也与弱结合能力相关。

#### 20 次级作用和 A1 结合

与以上介绍的对 A24 类似的分析也应用于 HLA-A\*0101 分子。简言之，以上的研究已经确定了两个对 HLA-A\*0101 特异的不同的肽结合基元。获得了对 A\*0101 亚基元的 9 聚体和 10 聚体次级相互作用图。为得到这个次级相互作用图，收集了与这两个基元的每个相对应的肽组有关的 A\*0101 结合数据。对于每个位置，计算了带每个特定残基的肽的相对平均结合亲和力。为了补偿某些特定残基的低出现率，并获得更有意义的抽样，按 Ruppert 等 (同上) 的介绍，组合带有类似化学侧链的氨基酸。对 9 聚体肽进行这种分析获得的，2-9 和 3-9 亚基元的结果分别见图 4a 和 4b。由这种分析检测的次级作用的说明图见图 4c 和 4d (分别代表 2-9 和 3-9 亚基元)。正如上面所介绍的，平均亲和力提高或降低 4 倍以上被断定为显著的。

一般说来，特定残基类型的存在会消极地或积极地影响大多数位



置的结合能力。例如，在 2-9 亚基元的情况下，发现在第 1 位带有 D 或 E 的肽很弱地结合到 A\*0101 分子上，平均相对结合能力 (ARBC) 为 0.20。相反，在同一位置 (第 1 位) 带有芳香族残基 (Y、F 或 W) 的肽，平均亲和力比整个肽组的总平均结合能力高 4 倍 (ARBC 4-0)。

5 检查这个图发现肽结合 A\*0101 的肽的一些有趣的特征。首先，如上面所提到的，在第 2 和 3 位上的两个锚定相互之间起协同作用。在第 2 位带有 M、S 或 T 锚定的肽的亲和力，被第 3 位存在的 D 或 E 显著提高 (第 3 位是 A 时提高程度弱一些)。反之，在第 3 位带有 D 或 E 锚定的肽的亲和力，被 S、T 和 M 的存在显著提高 (但也可以是其它的疏水或短链分子如 L、V、I、C 和 A)。

通过检测其它位置揭示了带两个亚基元的肽与同 A\*0101 分子相互作用的差异程度。比较图 4a 和 4b 的数值，清楚地显示有许多例子在一个基元上起中性作用的残基，在另一个基元上起积极的或消极的作用。例如，在 2-9 基元的第 1 位，G 和芳香族残基 (Y、F 和 W) 是优选的 (ARBC > 4.0)，A 和带正电荷的残基 (R、H 和 K) 是相对中性的 (ARBCs 在 4.0 和 0.25 之间)，而带负电荷的残基 (D 和 E) 是有害的 (ARBC < 0.25)。在带 3-9 亚基元的肽中注意到有一个不同的模式，除了 G 仍然是优选的外，其它的优选被打乱了。在第 1 位带正电荷的残基对肽结合有显著的正影响 (ARBC 为 8.3)，带负电荷的残基和芳香族残基是中性的 (ARBCs 分别为 1.3 和 0.61)，A 是有害的 (ARBC 为 0.15)。在基元内的每个位置都观察到类似的修饰类型。总的说来，次级锚定优选残基在亚基元之间不同，在图 3 的总结图中可以清楚地看到。在这一节里可以看到，除了第 1 位的 G 是共有的优选残基这个唯一的例外，和第 2 和 3 位的协同锚定外，两个 A\*0101 9 聚体亚基元的次级基元事实上是完全不同的。因此，定量而言，两个 9 聚体基元的 27 个次级作用中只有一个是共有 (3.7%)。与见于 A24、A\*0201 扩展的基元和 A3 样分子 (20) 之间的多种相似性相比，这些 A\*0101 基元差异的程度是显著的，在那些基元之间，任何两个扩展的基元之间可见共有 3 到 5 (13-26%) 个次级作用。

30 10 聚体配体次级锚定残基的确定。

与上一节关于 9 聚体配体的介绍相类似，也确定了 2-10 和 3-10 亚基元的次级锚定残基和次级作用。这些分析的结果见图 5a-d。第 2 和



3 位的锚定再一次表现出相互协同作用。在 2-10 亚基元中，第 3 位的 D、E (A、Q 和 N 的程度低很多)，在 3-10 亚基元中，第 2 位的疏水 (L、I、V、M) 或短链 (S、T、C) 残基与显著提高平均结合能力相关。

5 在第 2、3 位和 C 末端以外的位置比较两个 10 聚体基元显示，与 9 聚体中的情况一样，次级锚定特异性的修饰取决于主要锚定残基。例如，在第 7 位，A 和 S，T 和 C 在 2-10 基元上优选，而在 3-10 基元上是中性的。相反，G 在 3-10 基元上优选，而在 2-10 基元上是中性的。但与 9 聚体亚基元对比，在 10 聚体中观察到的这些差异明显没有那么显著。事实上，两个 10 聚体亚基元共有许多优选残基。第 1 和  
10 5 位的 Y、F 和 W，第 4 位的 A，第 7 位的 P，和第 8 位的 G 对两种亚基元都有积极的作用。同样，第 8 位的 R、H 和 K 对两种 10 聚体亚基元都是有害的 (图 5c 和 5d)。总的说来，两种 10 聚体基元共有 25 个次级作用中的 6 个 (24%)。

15 提供以上的例子用来说明本发明，而非限制其范围。本发明的其它改变对本领域的技术人员是显而易见的，并包括在所附权利要求中。本文引用的所有出版物、专利和专利申请均引作参考文献。



说明书附图

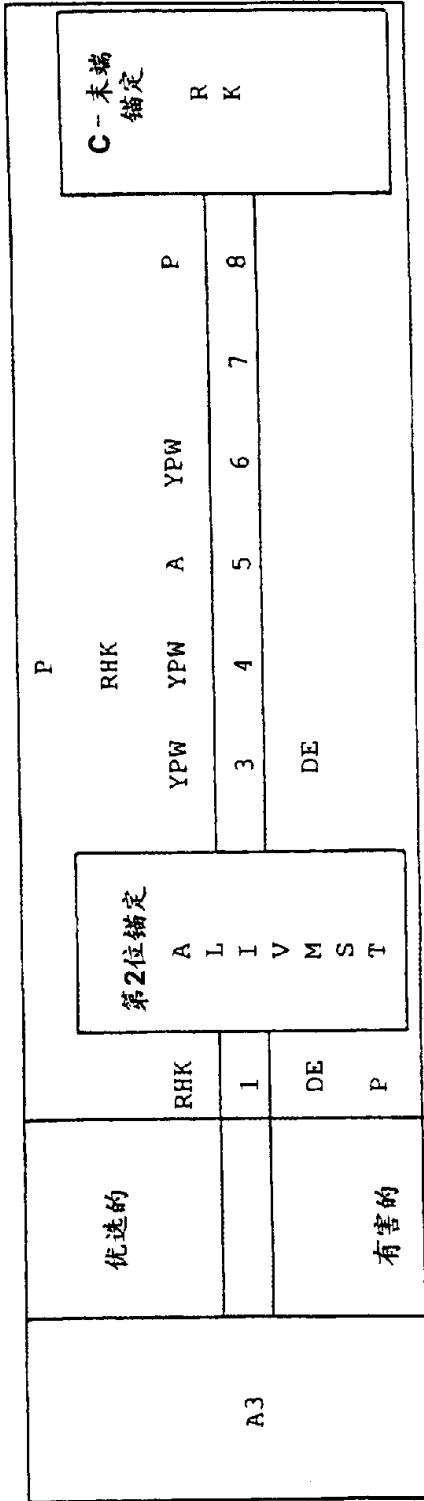


图1A

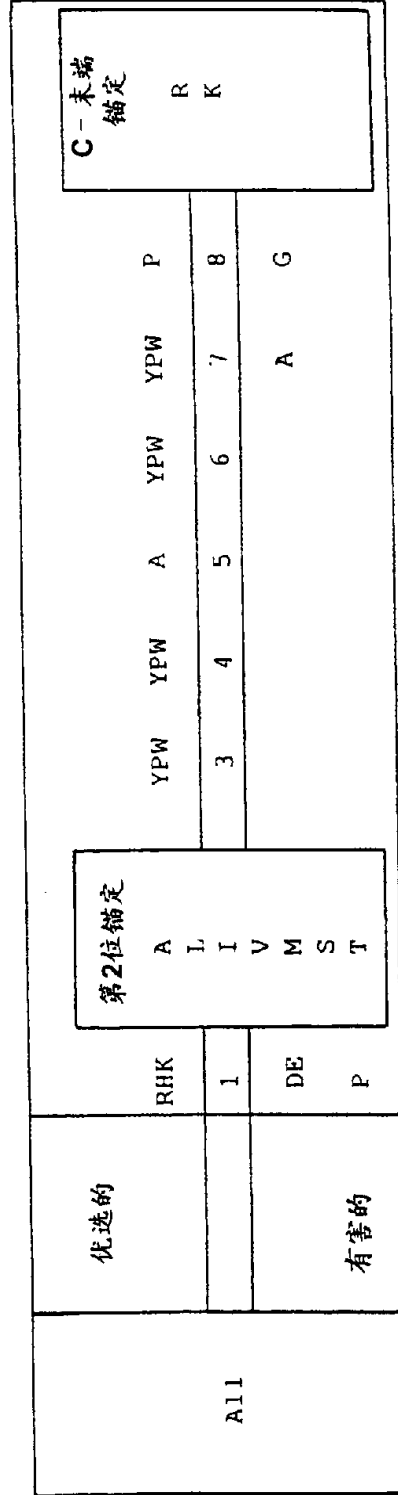


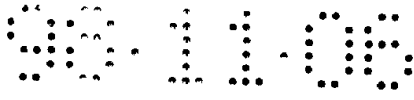
图1B

A*3101	优选的	RHK	第2位锚定								C-末端锚定				
		1	A	L	I	V	M	S	T	Y	P	W	A	R	K
		DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE
有害的	P														

图 1C

A*3301	优选的	RHK	第2位锚定								C-末端锚定				
		1	A	L	I	V	M	S	T	Y	P	W	A	R	K
		G	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE	DE
有害的	P														

图 1D



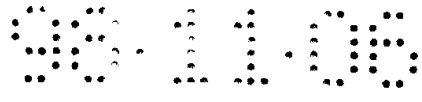
A*6801	优选的	YPW	第2位锚定								C-末端锚定				
		STC	A	L	I	V	M	S	T	YPW	YPW	P	R	K	
	有害的	1								3	4	5	6	7	8
		G								DE	RHK				A
		P							G						

图 1E

A3-样超单元	优选的	YPW	第2位锚定								C-末端锚定				
		STC	A	L	I	V	M	S	T	YPW(4)	YPW(3)	YPW(4)	P(4)	R	K
	有害的	1								3	4	5	6	7	8
		DE(3)								DE(4)					
		P(5)													

图 2





B51	优选的	LIVM	第2位锚定					C-末端 锚定 A L I V M F W Y				
		FWY		FWY	STC	FWY	G	FWY				
		1	P	3	4	5	6	7	8			
	有害的	AGPDE		DE	G	DE	G					
		RHK				QN	DE					
		STC										

图 3C

B*5301	优选的	LIVM	第2位锚定					C-末端 锚定 A L I V M F W Y				
		FWY		FWY	STC	FWY	I.IVM	FWY				
		1	P	3	4	5	6	7	8			
	有害的	A				G	RHK	DE				
		GP					QN					
		QN										

图 3D



**HLA A1平均相对结合能力: 9聚体肽**

组别	位置								
	1	2	3	4	5	6	7	8	9
A	1.1	4.5	0.09	0.64	0.19	0.30	0.27	<b>C-末端 锚定</b>  Y	
G	8.4	0.69	0.34	0.10	1.4	0.31	0.55		
DE	0.20	33	0.36	0.35	0.37	12	0.97		
RHK	0.78	0.13	1.1	1.5	0.33	0.32	0.83		
LIVM	0.50	0.17	0.97	2.2	1.7	1.2	1.7		
YFW	4.0	1.8	4.5	2.6	3.4	2.4	6.2		
QN	1.3	1.8	0.53	1.9	1.0	7.7	0.79		
STC	1.4	2.4	2.6	1.0	2.0	0.74	0.97		
P	0.60	0.17	1.2	0.57	22	0.27	0.87		

图 4A

**HLA A1平均相对结合能力: 9聚体肽**

组别	位置								
	1	2	3	4	5	6	7	8	9
A	0.15	34	0.58	0.58	1.5	421	2.3	0.34	<b>C-末端 锚定</b>  Y
G	27	0.63	63	0.35	3.2	0.49	0.16	0.052	
DE	1.3	0.033	0.035	0.38	0.37	2.5	4.0	0.95	
RHK	8.3	0.029	0.58	0.48	0.20	0.58	2.3	0.95	
LIVM	0.72	16	2.1	1.7	0.92	5.4	2.3	0.41	
YFW	0.61	0.035	0.75	1.1	1.8	1.3	0.41	2.4	
QN	0.27	0.43	0.47	0.16	1.4	1.4	2.4	1.9	
STC	0.71	30	4.7	3.8	11	0.80	1.9	0.073	
P	3.1	0.047	0.91	0.032	0.76	0.052	0.073	0.073	

图 4B



**HLA A1 平均相对结合能力: 10 聚体肽**

组别	位置									
	1	2	3	4	5	6	7	8	9	10
锚定:	2.3	8.5	15	2.3	0.23	0.78	0.27	0.058	C-末端 锚定 Y	
第2位	0.26	0.16	0.31	0.67	1.8	0.65	14	0.87		
和C末端	0.42	56	0.22	0.28	2.1	0.46	6.2	0.36		
	0.61	0.080	0.49	0.19	1.8	0.13	0.077	0.99		
	2.1	0.15	0.89	0.75	1.1	0.66	2.3	1.5		
	10	0.94	0.26	5.5	1.4	0.13	2.5	2.0		
	0.94	4.4	3.3	4.5	0.13	2.2	0.70	1.2		
	1.2	0.71	2.8	0.83	1.0	4.5	1.1	0.71		
	0.21	2.2	0.77	0.72	0.77	18	1.1	4.3		

图 5A

**HLA A1 平均相对结合能力: 10 聚体肽**

组别	位置									
	1	2	3	4	5	6	7	8	9	10
锚定:	0.67	0.80	11	0.67	1.5	1.7	0.67	0.26	C-末端 锚定 Y	
第3位	0.27	0.33	2.7	3.2	0.093	7.1	6.5	1.0		
和C末端	1.4	0.068	0.40	1.4	0.50	0.34	1.0	1.0		
	0.16	0.066	0.52	0.63	0.98	0.40	0.20	0.66		
	2.2	48	1.0	1.7	1.2	1.0	0.76	1.0		
	13	0.066	0.59	7.5	3.6	3.4	2.8	19		
	0.66	0.76	1.6	0.48	0.57	1.9	1.4	0.25		
	0.52	90	0.34	0.44	2.1	0.34	0.78	0.72		
	0.53	0.076	1.3	0.098	3.1	15	0.15	1.9		

图 5B

**HLA A1 次级锚定基元: 10聚体肽**

锚定: 第2位 和C末端	优选的	DE										C-末端 锚定
		P										
有害的	YFW	A										Y
		G										
第2位 锚定	1	QN										A
		STC										
第3位 锚定	2	3										A
		4										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	3	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	4	5										A
		6										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	5	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	6	7										A
		8										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	7	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	8	9										A
		10										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	9	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	10	11										A
		12										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	11	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	12	13										A
		14										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	13	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	14	15										A
		16										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	15	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	16	17										A
		18										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	17	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	18	19										A
		20										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	19	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	20	21										A
		22										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	21	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	22	23										A
		24										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	23	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	24	25										A
		26										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	25	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	26	27										A
		28										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	27	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	28	29										A
		30										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	29	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	30	31										A
		32										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	31	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	32	33										A
		34										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	33	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	34	35										A
		36										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	35	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	36	37										A
		38										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	37	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	38	39										A
		40										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	39	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	40	41										A
		42										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	41	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	42	43										A
		44										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	43	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	44	45										A
		46										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	45	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	46	47										A
		48										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	47	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	48	49										A
		50										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	49	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	50	51										A
		52										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	51	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	52	53										A
		54										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	53	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	54	55										A
		56										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	55	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	56	57										A
		58										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	57	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	58	59										A
		60										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	59	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	60	61										A
		62										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	61	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	62	63										A
		64										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	63	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	64	65										A
		66										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	65	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	66	67										A
		68										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	67	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	68	69										A
		70										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	69	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	70	71										A
		72										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	71	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	72	73										A
		74										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	73	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	74	75										A
		76										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	75	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	76	77										A
		78										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	77	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	78	79										A
		80										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	79	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	80	81										A
		82										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	81	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	82	83										A
		84										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	83	RHK										RHK
		QN										
有害的	YFW	G										A
		YFW										
第3位 锚定	84	85										A
		86										
有害的	YFW	RHK										RHK
		DE										
第3位 锚定	85	RHK										RHK
		QN										