

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 02807505.6

[51] Int. Cl.

C07K 16/46 (2006.01)

C07K 19/00 (2006.01)

A61K 39/395 (2006.01)

[45] 授权公告日 2009 年 5 月 20 日

[11] 授权公告号 CN 100488984C

[22] 申请日 2002.3.30 [21] 申请号 02807505.6
[30] 优先权
[32] 2001. 3.30 [33] US [31] 60/280,625
[86] 国际申请 PCT/US2002/009815 2002.3.30
[87] 国际公布 WO2002/079232 英 2002.10.10
[85] 进入国家阶段日期 2003.9.28
[73] 专利权人 默克专利有限公司
地址 德国达姆施塔特
[72] 发明人 S·D·吉利斯
[56] 参考文献
WO9902709 A 1999.1.21
WO9859244 A 1998.12.30
审查员 马 岚

[74] 专利代理机构 北京市中咨律师事务所
代理人 黄革生 林柏楠

权利要求书 1 页 说明书 59 页

[54] 发明名称

降低融合蛋白的免疫原性

[57] 摘要

本发明公开了制备具有降低的免疫原性的融合蛋白的组合物和方法。本发明的融合蛋白包括具有氨基酸改变的连接区，此氨基酸改变降低了连接处表位和 MHC II 类结合的能力，因此降低了它和 T-细胞受体的相互作用。本发明的方法涉及对融合蛋白连接区中一处或多处氨基酸的分析、改变或修饰，目的是鉴定出 T-细胞表位并降低它和 T-细胞受体相互作用的能力。本发明的组合物和方法在治疗中 useful。

1. 具有降低的免疫原性的融合蛋白，其包含 IgG 分子或其片段，该 IgG 分子或其片段在其 C 端通过融合连接处与第二蛋白的 N 端融合；

该融合蛋白的特征在于：通过对包括融合连接处两侧各 1 到 8 个氨基酸的候选 T 细胞表位进行修饰以降低该候选 T 细胞表位与 T 细胞受体相互作用的能力，所述修饰为将 IgG 序列 C 端的候选 T 细胞表位 SEQ ID NO: 5 中的氨基酸序列 LSLS 改变成 ATAT；或者，

该融合蛋白的特征在于：融合连接处包含 N-连接的或 O-连接的糖基化位点。

2. 权利要求 1 的融合蛋白，其中融合连接处包含间隔区或接头。

3. 权利要求 1 或 2-的融合蛋白，其中 IgG 分子片段是 Fc 片段。

降低融合蛋白的免疫原性

相关申请

本申请要求系列号 60/280,625 (2001 年 3 月 30 日提交) 的美国临时专利申请的优先权和利益, 此临时专利申请的所有公开内容都在此通过参考方式并入。

发明领域

本发明一般地涉及用于制造和使用作为治疗剂的具有降低的或没有免疫原性的经修饰融合蛋白的方法和组合物。更具体的, 本发明涉及融合蛋白, 通过鉴定候选 T-细胞表位和修饰氨基酸序列消除此表位, 该融合蛋白具有降低的免疫原性。

发明背景

很多治疗性蛋白是正常的人体蛋白。例如, 白细胞介素-2, 促红细胞生成素, 和生长激素均是人体蛋白, 它们被施用给通常已有内源水平的这些蛋白的人。通常将完全正常的人体蛋白作为治疗剂使用时极少有针对这种蛋白的免疫应答。

最近显示很多有人工活性的融合蛋白可用作治疗性蛋白。例如, Enbrel 是 TNF 受体的细胞外结构域和 IgG1 Fc 区域的融合物。Enbrel 用于治疗类风湿性关节炎, 而且被认为通过滴定 TNF 和阻止 TNF 活性来作用。但是, 已注意到在用 Enbrel 治疗的患者中有显著的抗 Enbrel 抗体发生率。

另一个治疗上有用的融合蛋白类型的例子是免疫细胞因子。这些蛋白包括一个抗体部分和一个细胞因子部分, 可用于使细胞因子靶向疾病细胞, 如癌细胞。

但是，这些融合蛋白中的很多蛋白在治疗上的应用因为它们哺乳动物中，尤其在人类中的免疫原性而受到了限制。

因此，为了在治疗中使用这些蛋白，需要制备具有降低的免疫原性的融合蛋白。

发明概述

本发明的特征在于用于制备治疗中使用的具有降低免疫原性的融合蛋白的方法和组合物。例如，本发明涉及具有降低的免疫原性的免疫细胞因子，免疫融合素（immunofusin），免疫配体，其它抗体和 Fc 融合蛋白，细胞因子-细胞因子融合蛋白，以及清蛋白融合蛋白。

本发明部分涉及融合蛋白含有“异己”序列的认识。例如，即使在两个人体蛋白之间的融合中，围绕融合连接处的区域也含有一段通常不在人体中出现的肽序列。例如，蛋白质药物，如 Enbrel 来源于两种正常的人体蛋白：TNF 受体和 IgG1。但是，在 TNF 受体和 IgG1 之间的连接处是一段通常在人体中不出现的肽序列。

本发明优选的方法涉及通过降低连接处表位（连接肽）和 T-细胞受体相互作用的能力来降低融合蛋白的免疫原性，其中降低连接处表位和 T 细胞受体相互作用的能力是通过降低该表位和 MHC 分子结合的能力（其结合亲和力）实现的。根据本发明，连接处表位或肽优选为“非己”的。通常蛋白质，包括治疗性蛋白，都具有免疫原性，这部分是因为蛋白质通过抗原提呈细胞内吞和蛋白水解后产物肽结合到称为主要组织相容性复合体（MHC）的分子上然后这些分子把肽提呈给 T 细胞。抗原提呈细胞（APC）表面的抗原肽-MHC 复合体刺激 T 细胞增殖，分化并释放细胞因子。同时，B-细胞分化和抗体产生被诱导，这可以进一步通过清除治疗性蛋白而限制治疗性蛋白的效用。这样，来源于治疗性蛋白的抗原肽能够引起一系列不想要的免疫应答。治疗性蛋白的效用由于抗体的中和而受限，而且因为 T-细胞和 B-细胞应答会在患者中引起炎性和变态反应，其诱导经常是有害的。

本发明涉及(1)鉴别具有一个或多个候选 T-细胞表位的免疫球蛋白-靶蛋白连接区域中的新氨基酸序列;(2)对这些氨基酸序列进行修饰以减少或消除来源于连接处序列的具有 T-细胞表位功能的肽。

本发明提供两大类与降低免疫原性有关的组合物和方法。根据本发明的一个实施方案,在跨融合连接处的序列中鉴定潜在的非己 T-细胞表位。例如,可通过计算机方法基于对与 MHC II 类分子结合的肽的模建鉴定出潜在的非己 T-细胞表位。然后进行替代以降低或消除来自连接区的肽与 MHC II 类结合的能力。用于鉴定和修饰与 MHC II 类分子结合的肽的该方法被命名为“去免疫化”而且所获修饰蛋白分子被称作“去免疫化的”。

根据本发明的另一个实施方案,在融合连接处引入一个或多个糖基化位点。虽然可以使用 O-连接的糖基化位点,但优选使用 N-连接的糖基化位点。根据一个优选的实施方案,对围绕融合连接处的连接区域中野生型序列的氨基酸进行突变,使 N 端融合部分的最后一个氨基酸突变成精氨酸,而第二个融合部分的头两个氨基酸突变成甘氨酸及之后的丝氨酸或苏氨酸。

根据本发明, MHC II 类结合的去除了在以下情况下是优选的,即当蛋白质在细菌或者不产生哺乳动物糖基化模式的生物,如酵母或昆虫细胞内产生时。

当蛋白质在哺乳动物细胞系内或者在产生对哺乳动物无害的糖基化模式的细胞系内制备时,可能优选引入糖基化位点。

在一个优选的实施方案中,融合蛋白的一个组分是细胞因子。术语“细胞因子”此处用于描述天然的或重组的蛋白质、其类似物及其片段,它们可以在具有所述细胞因子的受体的细胞中引起特异性反应。优选地,细胞因子是可由细胞产生和分泌的蛋白质。优选地,细胞因子包括白细胞介素如白细胞介素 2 (IL-2)、IL-3、IL-4、IL-5、IL-6、IL-7、IL-10、IL-12、IL-13、IL-14、IL-15、IL-16 和 IL-18,造血细胞因子如粒细胞-巨噬细胞集落刺激因子(GM-CSF)、G-CSF 和促红细胞生成素,肿瘤坏死因子(TNF)如 TNF α 、淋巴因子如淋巴毒素,代谢过程调节物如 leptin 以及干扰素如

干扰素 α 、干扰素 β 和干扰素 γ 和趋化因子。优选地，本发明的抗体-细胞因子融合蛋白显示出细胞因子特异性的生物活性。

在另一个优选的实施方案中，融合蛋白的一个组分是抗肥胖细胞因子。例如，组分为 leptin, CNTF, 或 Acrp30 的一部分。

在一个备选的优选实施方案中，融合蛋白的一个组分是激素。例如，组分可以是胰岛素，生长因子或胰高血糖素样肽（GLP-1）。

在再一个备选的实施方案中，融合蛋白的一个组分是具有生物学活性的配体结合蛋白。在一个优选实施方案中，使用 TNF 受体的细胞外结构域。

根据一系列的实施方案，本发明的融合蛋白包含非抗体部分的 N 端与抗体部分的 C 端的融合。根据另一系列的实施方案，本发明的融合蛋白包含非抗体部分的 C 端与抗体部分的 N 端的融合。本发明中，抗体部分可以是完整的免疫球蛋白或完整免疫球蛋白的一部分。免疫球蛋白的一部分可以包括可变区或恒定区或两者都包括。优选的免疫球蛋白包括 Fc 区或其部分。本发明一个优选的实施方案包括经修饰具有降低的免疫原性和/或更长的血清半衰期的 IgG1 免疫球蛋白同种型，或其一部分。例如，优选在 CH3-细胞因子连接处附近具有氨基酸残基修饰的 IgG1。对特定的应用，优选来自 IgG2 或 IgG4 同种型的抗体部分。

免疫细胞因子仅仅是靶向肿瘤的融合蛋白疗法的一个例子。其他肿瘤-毒性分子也可以通过融合到肿瘤特异性的抗体上而靶向肿瘤。另外，抗体融合蛋白可以攻击其他类型的疾病细胞，如感染病毒的细胞。另一个改造靶向融合蛋白的方法是使用 Fc-X 和 X-Fc 技术，此处的 X 是多肽。这些技术运用了如下知识，即如果目的多肽连接到免疫球蛋白的 Fc 部分上，则靶蛋白的制备和收集将得以改善。对于 Fc-X 融合蛋白，信号肽及后面的免疫球蛋白基因的 Fc 部分是靶蛋白的 N 端融合部分。在一些情况下，以 X-Fc 取向构建融合蛋白是特别有利的。在这样的结构中靶蛋白是 N 端融合部分，Fc 片段位于其后。对一些蛋白，如淋巴细胞表面糖蛋白(LHR)(US 专利 5, 428, 130)和胰高血糖素样肽（GLP-1），这样的方法是有用的。

因此，本发明的方法和组合物提供具有降低的免疫原性的 Fc-X 和

X-Fc 形式融合蛋白。根据本发明，融合蛋白的免疫原性可根据本领域内公知的或此处公开的方法测定。

本发明的方法和组合物还提供了具有降低的免疫原性的清蛋白融合蛋白。人血清清蛋白 (HAS) 因为具有非常长的半衰期、在体内分布广泛以及缺少酶学和免疫学功能而被用作治疗性肽/蛋白的载体 (Yeh 等, PNAS89: 1904-1908, 1992)。生物活性肽与 HAS 的遗传融合对分泌的治疗性 HAS 衍生物的回收是有用的。但是, 根据本发明, 清蛋白融合蛋白如 HAS-CD4 具有一个新的连接处, 其中一般含有一个或多个可被 MHC II 类分子提呈的 T 细胞表位。本发明提供具有降低的免疫原性的清蛋白融合蛋白, 以及用于降低清蛋白融合蛋白的免疫原性的一般方法。本发明中有用的清蛋白包括各种清蛋白, 其等位变体和突变体, 亦包括它们的片段。优选的清蛋白保留野生型清蛋白如 HAS 的结构和功能特性。

另一方面, 本发明提供包含有用突变的去免疫化的抗体融合蛋白, 该蛋白具有正常的、突变的或杂合的同种型。这些突变可能位于连接处附近或位于连接区域之外的位置。

例如, 本发明提供在连接处有修饰的去免疫化的免疫细胞因子, 其在 IgG 与非 IgG 部分之间的连接处有点突变。细胞因子部分包括任何细胞因子, 但是优选 IL-2 或 IL-12。在一个实施方案中, 氨基酸的改变涉及到抗体部分 C 末端赖氨酸变成疏水氨基酸如丙氨酸或亮氨酸。将这种突变和本发明的去免疫化修饰相结合的一个关键优点是这些突变可共同作用以增加血清半衰期及降低免疫原性。此处描述的把融合连接处的去免疫化与改变血清半衰期的突变相结合的方法可用于显著改进这些融合蛋白的临床效用。

另一方面, 本发明提供含有杂合抗体部分以及在融合连接处进行了去免疫化修饰的免疫细胞因子, 其中所述杂合抗体部分包括不同 Ig 同种型来源的结构域, 优选 IgG1 和 IgG2 同种型来源的结构域。例如, 本发明提供去免疫化的免疫细胞因子, 其使用 IgG2 和 IgG2h 杂合体并在连接处具有修饰 (所述 IgG2h 为铰链区修饰成 IgG1 的 IgG2)。在一个

优选的实施方案中，杂合融合蛋白由去免疫化的 IgG 免疫球蛋白部分（由（ $\gamma 1$: CH1-H）（ $\gamma 2$: CH2-CH3）构成）和细胞因子部分组成。

另一方面，本发明提供编码具有降低的免疫原性的融合蛋白或有助于具有降低的免疫原性的融合蛋白表达、制备和分泌的新核酸序列。此核酸序列根据标准的重组 DNA 技术产生。

在一个优选实施方案中，核酸分子编码免疫细胞因子融合蛋白。优选的免疫细胞因子包括细胞因子（如白细胞介素 2）和肿瘤特异性的单克隆抗体（如抗人上皮细胞粘附分子 KSA（EP-CAM）（huKS）的抗体）。

在另一个优选的实施方案中，核酸分子编码各种构型的 Fc 融合蛋白。核酸分子按 5'至 3'方向顺次编码（i）信号序列、免疫球蛋白 Fc 区和靶蛋白序列或（ii）信号序列、靶蛋白和免疫球蛋白 Fc 区，或（iii）信号序列、第一靶蛋白、免疫球蛋白 Fc 区及第二靶蛋白。产生的核酸分子因此编码 Fc-X、X-Fc 或 X-Fc-Y 结构，此处 X 和 Y 是靶蛋白。在一个备择实施方案中，核酸分子编码不带信号序列的 Fc-X、X-Fc 或 X-Fc-Y 蛋白。

在另一个优选实施方案中，本发明的核酸编码具有突变或杂合同种型的 Ig 融合蛋白。具体地，该核酸提供具有杂合同种型或者备择性地具有经修改的铰链区的抗体部分。例如，融合蛋白由在铰链区含有较少二硫键的修饰 IgG2 组成，或者由铰链区来源于其他抗体（优选正常的或突变的 IgG1 铰链区）的 IgG2 CH2 和 CH3 区组成。

本发明的核酸优选以可操作连接方式整合进可复制的表达载体中，然后此载体导入能产生此融合蛋白的哺乳动物宿主细胞中。所得融合蛋白将从哺乳动物宿主细胞中高效地产生并分泌。接着无需裂解哺乳动物宿主细胞可从培养基中收集分泌的融合蛋白。对蛋白产物测定活性和/或按需要用普通试剂进行纯化，和/或与融合部分分离，这些均可使用常规技术实现。

因此，本发明还提供具有降低的免疫原性的融合蛋白的制备方法。

本发明的方法和组合物还可用于提供使用经修饰具有较小免疫原性的融合蛋白的治疗方法。本发明的总目的是提供既有效又廉价的方法以及有更小免疫原性的蛋白。本发明优选的治疗组合物包括治疗有效量的去免疫

化的融合蛋白。优选地，去免疫化的融合蛋白和药物学上可接受的载体一起施用。

从下面的详细说明、附图和权利要求中，本发明上述的和其他的方面、特性和优点将更加清晰。

发明详述

包括抗体在内的所有蛋白在治疗性施用给患者时都有可能在受体宿主内诱导免疫应答。此免疫应答由 T-淋巴细胞 (T-细胞) 介导，然后触发 B-淋巴细胞 (B-细胞) 产生抗体。针对治疗药物的抗体产生是有害的，因为这导致治疗药物更快的清除而且可能导致变态反应。

本发明提供降低融合蛋白免疫原性的方法。根据本发明的一个方法，鉴定融合蛋白中融合连接处的连接区域内的潜在 T-细胞表位。T-细胞表位可以通过多种计算机和非计算机方法进行鉴定，包括基于以结构为基础的计算机模建进行预测或通过合成肽和检验与特定 MHC II 类分子的结合进行鉴定，或通过免疫原性实验进行鉴定。

依据本发明，融合连接处 (junction) 定义为融合蛋白中在第一蛋白或肽的最末 (C 端) 氨基酸与第二蛋白或肽的第一个 (N 端) 氨基酸之间的位置。因此，融合连接处包括位于一个蛋白的最末氨基酸与第二蛋白的第一个氨基酸之间的任何氨基酸。在一个实施方案中，融合连接处包括接头 (linker)。

依据本发明，连接区域 (junction region) 是融合蛋白中围绕和涵盖两个蛋白之间的融合连接处的区域。连接区域优选包括 1 到约 100 个氨基酸，更优选 1 到约 50 个氨基酸，或 1 到约 25 个氨基酸，甚至更优选 1 到约 15 个氨基酸，或 1 到 9 个氨基酸。在一个实施方案中，连接区域包含插在两个蛋白之间的连接点上的间隔肽或接头肽。依据本发明，含有间隔肽或接头肽的连接区域也可被去免疫化以使患者对含有间隔肽或接头肽的融合蛋白的反应最小化。

依据本发明，连接处的 T-细胞表位定义为能够与 MHC II 类分子结合的肽序列，其含有来源于融合蛋白的至少 2 个不同融合部分之每一个的各

至少一个氨基酸。例如，Paul（基础免疫学（Fundamental Immunology），第8章，表8，p.276[2000]第4版）阐述了能够与MHC II类分子结合的10个氨基酸的片段。在连接处的T-细胞表位中，这种10个氨基酸的肽来源于不同的融合部分。依据本发明，跨融合连接处的潜在的或候选的T-细胞表位（连接处的候选T-细胞表位）优选包括连接处两侧每一侧的各1到8个氨基酸，或更优选连接处两侧每一侧的各1到10个或1到11个氨基酸。候选表位优选为9，11或12个氨基酸长。这样，因为本发明的连接处T-细胞表位包括连接处两侧之每一侧的各至少一个氨基酸，故优选的候选T-细胞表位为这样的连接处表位，其含有1-8（或1-10，或11）个源于连接处一侧的氨基酸，而且它还包括互补数目的源于连接处另一侧的氨基酸，结果产生具有9-12个氨基酸，最优选9个氨基酸的表位。

依据本发明，然后突变连接处T-细胞表位内的锚（anchor）残基以防止其与MHC II类分子的结合。通常，应小心不要引入另外的潜在T-细胞表位，而且保留每个融合部分的功能。

依据本发明，野生型序列的融合是指融合连接处N端侧和C端侧的序列直接来源于天然序列的融合。

依据本发明，去免疫化的融合连接处是与野生型序列的连接处相对而言，在一处或多处引入了替换突变的连接处序列。在最优选的实施方案中，融合连接处的去免疫化不涉及接头，如‘非免疫原性’甘氨酸-丝氨酸接头的引入，而且融合部分之间的空间相互关系在去免疫化的融合蛋白中不发生变化。根据本发明，可对连接区域中的一个或多个氨基酸进行替换或改变，这些氨基酸可位于融合连接处的N端、融合连接处的C端或者融合连接处的N端和C端。

依据本发明，潜在T-细胞表位是一种序列，当其作为分离的肽被考虑时，预计可以与MHC II类分子或非人物种内的等价物结合。对潜在T-细胞表位的定义不考虑抗原加工的其他方面，如蛋白被摄入抗原提呈细胞的效率、在完整蛋白的位点上断裂产生可与MHC II类结合的肽的效率等等。因此，在蛋白施用给动物之后在MHC II上实际提呈的T-细胞表位组是潜

在 T-细胞表位的亚组。

依据本发明, T-细胞表位是蛋白上与 MHC II 类分子相互作用的表位。在不希望被理论束缚的情况下, T-细胞表位被理解为蛋白质或融合蛋白中的一段氨基酸序列, 其在 T-细胞发育过程中未能经历 T-细胞负选择过程因此预期将被 MHC II 类分子提呈并被 T-细胞受体识别。在本发明的一个优选实施方案中, 非己 T-细胞表位出现在形成融合蛋白的两个蛋白的融合连接处的连接区域中。

本发明提供用于降低或消除融合蛋白连接处中 T-细胞表位数目的非计算机方法, 其不需要复杂的计算机模拟或蛋白质三维结构。在一个实施方案中, 本发明的方法利用在抗原提呈时 9 氨基酸的核心片段既与 MHC II 类分子又与 T-细胞受体相互作用的事实。最 N 端的氨基酸被称为“锚”位置残基, 其与 MHC II 类分子内的一个深口袋结合。下面氨基酸中的一个通常出现在对与 MHC II 类分子的结合重要的锚位置处: 亮氨酸、缬氨酸、异亮氨酸、蛋氨酸、苯丙氨酸、酪氨酸和色氨酸。依据本发明, 核心 9 氨基酸邻近的另外 2 至 3 个氨基酸也影响与 MHC 分子的相互作用。而且, 在融合蛋白中第一蛋白的最 C 端氨基酸通常被替换。当已知具有活性的 N 端融合部分或第一蛋白与 C 端融合部分或第一蛋白 C 端的第二蛋白融合时, 这特别有用。

本发明的一个一般方法包括对融合蛋白的 N 端融合部分的最 C 端 8 个氨基酸中出现的亮氨酸、缬氨酸、异亮氨酸、蛋氨酸、苯丙氨酸、酪氨酸和色氨酸进行突变。在一个实施方案中, 连接处的候选 T-细胞表位氨基酸中的一个或多个氨基酸优先突变成苏氨酸、丙氨酸或脯氨酸。这保留了被替换氨基酸的一些疏水性质。在本发明的另外实施方案中, 上面提到的一个或多个氨基酸从候选的或潜在的连接处 T-细胞表位中删除, 或被合适的氨基酸类似物替换。依据本发明, 如果删除氨基酸来破坏潜在的 T-细胞表位, 则小心不要产生包括删除位置附近的氨基酸的新 T-细胞表位。

依据本发明, 构建一个通用表达质粒构建中间体常常是有用的, 此中间体包含在最后 8 个氨基酸中一个或多个疏水残基发生突变的 N-端融合部

分的编码序列。通常，此质粒在 N 端融合部分的 C 端编码 DNA 之上或附近有一个或多个方便的限制性酶切位点。

质粒构建中间体的目的是为了构建编码融合蛋白的表达质粒，此融合蛋白的一个或多个 N 端融合部分在 8 个 C 端氨基酸中具有一个或多个亮氨酸、缬氨酸、异亮氨酸、蛋氨酸、苯丙氨酸、酪氨酸或色氨酸向另一氨基酸的替换。此最终表达质粒的构建可用本领域内熟知的各种其他方法进行，如产生 PCR 片段或合成核酸，然后将片段连接入适当的载体或通过熟知的 PCR 技术与其它序列连接。

特别优选的实施方案包括 Fc-X 融合质粒，清蛋白-X 融合质粒，scFv-X 融合质粒以及 Fab-X 融合质粒。对于 Fc(γ)-X，在编码序列中引入突变以引起 IgG1、IgG2、IgG3 或 IgG4 分子的 Fc 区 C 端附近的亮氨酸-丝氨酸-亮氨酸-丝氨酸片段的氨基酸替换是有用的，如此处对于 IgG1 的说明：源于 IgG1，IgG2，IgG3 和 IgG4 的人 Fc 区的氨基酸序列分别在 SEQ ID NO:1,2,3 和 4 中描述。

在一个实例中，KSLSLSPGK (SEQ ID NO:5) 改变成 KSATATPGK (SEQ ID NO:6)。此突变的目的是消除潜在的连接处 T-细胞表位和去掉上游苯丙氨酸或酪氨酸作为位置 1 锚残基的 T-细胞表位。

备择地，有时把去掉候选连接处 T-细胞表位的突变和延长血清半衰期的突变结合起来是有用的。例如，把 KSLSLSPGK (SEQ ID NO:5) 改变成 KSATATPGA (SEQ ID NO:7)。

其他实施方案包括在 LSLS 片段中用其他氨基酸残基如甘氨酸或脯氨酸进行替换。

对于用于制备 IgA 融合蛋白的表达载体，删除一些 C 末端氨基酸使 C 末端附近参与 IgA 寡聚化的半胱氨酸缺失是有用的。例如，可以删除 15 个氨基酸，这样在被融合到第二蛋白上之前 IgA 重链序列末端为脯氨酸-苏氨酸-组氨酸。而且，在 IgA Fc 区域的 CH3 结构域的 C 末端附近引入下面的改变也是有用的：

QKTIDRLAGKPTH(SEQ ID NO:8)改变成 QKTADRTAGKPTH

(SEQ ID NO:9)

IgA-X 融合蛋白中的其他去免疫化序列为:

QKTPTRTAGKPTH (SEQ ID NO: 10)

QKTPTRPAGKPTH (SEQ ID NO: 11)

QKTATRPAGKPTH (SEQ ID NO: 12).

对于清蛋白-X 融合物, 在清蛋白-X 表达质粒中引入下面的改变使得清蛋白的 C-末端被修饰是有用的:

KKLVAASQAALGL(SEQ ID NO:13)改变成 KKLVAASQAATTA(SEQ ID NO:14)

因此, 本发明提供对构建较小免疫原性融合蛋白有用的核酸序列和蛋白质。具体地, 本发明提供在最后 8 个氨基酸中具有任何亮氨酸、缬氨酸、异亮氨酸、蛋氨酸、苯丙氨酸、酪氨酸或色氨酸突变的蛋白。此蛋白优选为其序列通常对应于人体中发现的序列的人蛋白。本发明还提供编码此蛋白的核酸序列。本发明此方面的核酸序列可以为质粒、PCR 产生的片段、或化学合成产生的核酸的形式。

本发明还提供编码融合蛋白的表达质粒, 其中所述融合蛋白的一个或多个 N-端融合部分在 8 个 C 末端氨基酸中有一个或多个亮氨酸、缬氨酸、异亮氨酸、蛋氨酸、苯丙氨酸、酪氨酸或色氨酸突变成其他氨基酸。

例如, 本发明提供编码 Fc 区按如上所述突变的 Fc-IL2 或全抗体-IL2 融合蛋白的质粒。另外, 本发明还提供含有如上所述突变的 Fc 区域的与促红细胞生成素的正常或突变形式(例如 WO01/36489 中描述的促红细胞生成素形式)融合的融合物。

本发明还提供通过在融合连接处附近, 或优选在融合连接处引入 N-连接或 O-连接糖基化位点来降低融合蛋白的免疫原性的方法。例如, 按如下所述引入氨基酸残基天门冬酰胺、丝氨酸或苏氨酸及第三残基。可以考虑如下序列, 其中 X 代表 N 端融合部分的氨基酸而 Z 代表 C-端融合部分的氨基酸。

$$X_1X_2X_3X_4X_5X_6Z_1Z_2Z_3Z_4Z_5Z_6Z_7Z_8Z_9$$

$$X_1X_2X_3X_4X_5NGSZ_3Z_4Z_5Z_6Z_7Z_8Z_9$$

依据本方法,连接肽的结合不一定必须通过引入糖基化位点进行阻断。但是,任何结合在 MHC II 类沟中的肽如果在最 N 端锚残基的 C 端具有糖基化天门冬酰胺,将不具有作为 T-细胞表位的功能。巨大的糖基化部分的出现将在空间上阻碍对 MHC II 类/肽复合物的识别。优选的糖基化位点包括序列 Asn-X-Ser 或 Asn-X-Thr, 此处 X 优选为 Gly, 但也可以为任意氨基酸。

另外,引入甘氨酸和丝氨酸残基的突变不会产生新的 T-细胞表位。甘氨酸和丝氨酸都不能作为锚残基。在抗原加工过程中,原则上融合蛋白在糖基化天门冬酰胺和甘氨酸之间或在甘氨酸和丝氨酸之间断裂。每一种情况下,产物肽都在锚残基 N 端具有突变的甘氨酸和/或丝氨酸,因此突变的甘氨酸和/或丝氨酸不能被 T 细胞受体识别,因为锚残基 N 端的残基处于 TCR 识别区域之外。

本方法的一个变体中,融合连接区域已经含有丝氨酸或苏氨酸,且其前面为氨基酸残基如甘氨酸、丝氨酸、丙氨酸等。当连接区域具柔性且被排除在每个融合部分的疏水核心之外的时候优选使用此第二种方法,以使新的 N-连接糖基化不影响每个融合部分的折叠或功能。

对蛋白质工程领域内的技术人员来说判断何时引入糖基化位点是可行的是件简单的事。例如,各融合部分或融合部分的近同系物的三维结构可能是已知的。通常,蛋白质的 N-末端或 C 末端几个氨基酸在 X 射线结构中不能被解析,或在 NMR 结构中显示出许多可能的构象。当在糖基化位点任何一侧有 3 个或更多个氨基酸处于无序态时,可在某种程度上确信得到的融合蛋白将能正确折叠而且每个部分都有活性。对于确定是否给定的融合蛋白结构具有功能,有些常规的实验是必需的。

在本发明的优选实施方案中,融合蛋白的 N 端和 C 端部分都是人蛋白。在这种融合蛋白中潜在的 T 细胞表位产生于 N 端部分(第一蛋白)的最后

8个氨基酸和C端部分(第二蛋白)的首8个氨基酸的组合。这提供了一系列8个由第一蛋白和第二蛋白产生的杂合9聚体。在第一蛋白最后8个氨基酸中任何脂肪族的或芳香族的残基(亮氨酸、缬氨酸、异亮氨酸、蛋氨酸、苯丙氨酸、酪氨酸或色氨酸)都有产生如下MHC结合肽的高风险性,在此MHC结合肽中此氨基酸位于第一个位置(锚位置)与MHC分子的口袋结合。因此,用非上面提到的任何一个氨基酸,且优选用丙氨酸、脯氨酸或苏氨酸替换上述任一氨基酸,将除去候选T细胞表位。

例如,当Fc融合蛋白含有下面序列时:

HNHYTQKSLSLSPGKGGGGSGGGSGGGGS (SEQ ID NO: 15),

亮氨酸残基产生两个潜在表位。因此,可以通过将L变成A和S变成T使序列去免疫化成为:

HNHYTQKSATATPGKGGGGSGGGSGGGGS (SEQ ID NO: 16),

此改变分别去除了在MHC结合口袋中作为首位氨基酸的亮氨酸和在MHC结合口袋中作为首位氨基酸的酪氨酸。

对于人类而言,这些用于去免疫化的替换对所有Fc融合蛋白,不论有无接头序列都可行,并优选在如下时候实施:1)融合蛋白中的两个蛋白都是人蛋白;2)两个蛋白的天然序列中的MHC结合肽可被忽略;以及3)和这些最初序列相同的9聚体也可被忽略。

本发明的方法一般适用于所有的脊椎动物,优选哺乳动物而且最优选人类。本发明通过下面非限定性的实施例作进一步解释。

实施例

实施例1:对huKS-IL2免疫细胞因子中免疫原反应性表位的推导

huKS-IL2由人源化的V_H和V_L区域结合人H链和L链恒定区组成。

H 链如以前所述在其羧基末端融合人 IL-2 的成熟序列。此 H 链是 $\gamma 1$ 同种型，对 Fc 受体具有高亲和性。此高亲和性使 huKS-IL2 从循环中被迅速清除掉。不被理论所束缚，推测 huKS-IL2 的清除通过肝脏（枯否细胞）和脾脏（抗原提呈细胞）中的 FcR-携带细胞进行。

以前已经知道某些患者对 huKS-IL2 分子的一些部分产生免疫应答，但是，抗体所识别的表位仍未知。为了推导出反应性表位，将患者血清与 huKS-IL2 的相对反应性与其他相关蛋白作比较：

(1) Hu14.18-IL2，一种具有完全不同的人源化 V 区但有完全相同的 C 区和与 IL-2 的融合连接处的分子。

(2) VH1，在 VH 和 VL 区域无 T 细胞表位的 huKS-IL2 的去免疫化形式，其源于表面暴露的小鼠 B 细胞表位被贴面化（veneer）处理为人残基的小鼠 V 区。

(3) VH2，在 CDR3 中剩余一个 T 细胞表位的 huKS-IL2 的去免疫化形式，其源于表面暴露的小鼠 B 细胞表位被贴面化处理为人残基的小鼠 V 区，其中 VH 含有一个 T 细胞表位。

(4) 用 KOL 或 EU C γ 1 区（而不是 KS）构建的 425-IL2（为了比较同种异型反应性）；

(5) huKS-mIL2，具有融合到小鼠 C 区和小鼠 IL2 上的 huKS V 区的分子；

(6) 人 Fc-IL2；

(7) 仅人 Fc；

(8) 仅人 IL2。

免疫球蛋白融合蛋白和片段通过蛋白 A 琼脂糖层析纯化，在重碳酸盐缓冲液中包被 96 孔平板然后用含有 1% BSA 的 1% 山羊血清封闭。对患者血清稀释液进行孵育，然后用 PBS-吐温三次洗涤，去除不结合的材料。根据结合蛋白采用不同的 HRP 偶联抗体检测患者血清中的结合的人抗体。通常，使用山羊抗人 λ 链 HRP 偶联物，因为大多数与平板结合的蛋白由

人 Fc 和人 κ 链构成。

某些患者的血清显示出对 huKS-IL2 清楚的反应性，而这在来自同一患者的注射前血清中没有检测到。使用免疫前抗血清来建立基线非免疫对照。在患者血清中观察到的反应性可归因于（1）抗 IL2 反应性，（2）抗 Fc（同种异型）反应性，（3）对新的连接序列的反应性，或（4）对 KS 独特型的抗独特型反应性，或这些反应性的组合。

没有对重组 IL2 或对 Fc 区（上面的 1 和 2）有显著反应的患者血清。有些患者显示出对 KS V 区的抗独特型反应性。所有的患者血清都显示有对 Fc-IL2 的反应性。四名患者中有三名显示对 Fc-IL2 的反应性。存在针对 Fc-IL2 的反应性而不存在针对 Fc 及 IL2 的反应性的现象提示 Fc 和 IL2 之间的连接处被患者的抗血清识别。

实施例 2：修饰抗体-细胞因子融合蛋白连接处的氨基酸残基以通过消除 MHC II 类结合基序来降低免疫原性

肽穿线分析（peptide threading analysis）法在免疫细胞因子的 Fc 部分和 IL2 部分之间的连接处鉴定出两个交叠的具有强 MHC 结合潜力的肽。肽穿线法和潜在的 T 细胞表位鉴定按 Carr (WO00/34317) 所公开的进行。引入氨基酸改变以消除已存在的潜在的 MHC II 类结合表位，但不引入新的潜在的 MHC II 类表位。

连接序列 LSLSPGK-AP(SEQ ID NO:17)修饰成 ATATPGA-AP(SEQ ID NO:18)（“LSLS 变成 ATAT”），此处的连字符是指免疫细胞因子 huKS-IL2 的连接处，此修饰使连接处来源的肽序列不能够与任何人类 MHC II 类以足够引起免疫原性的高亲和力结合。

实施例 3：在免疫细胞因子融合蛋白的连接处对氨基酸残基进行修饰以降低免疫原性

连接序列 LSLSPGK-AP(SEQ ID NO:17)修饰成 LNLSPGA-AP(SEQ ID NO:19)（“LSLS 变成 LNL”），此处的连字符是指免疫细胞因子

huKS-IL2 的连接处, 经此修饰后连接处来源的肽序列仍能与某些 MHC II 类分子结合。但是, 当 KS-IL2 蛋白在哺乳动物细胞中表达并被分泌时, 蛋白将因为 NXS/T 序列而在连接处附近被 N-糖基化。

由此得到的连接处来源的肽不能有效地作为 T 细胞表位, 因为当连接处来源的肽通过 MHC II 类提呈给 T 细胞时, 巨大的 N-糖基化部分将阻止 T 细胞-受体与 MHC II 类之间的特异性对接 (dock)。

实施例 4: 通过对比以表征抗原提呈细胞对免疫细胞因子 huKS-IL2 和去免疫化的 huKS-IL2 免疫细胞因子的免疫反应性

通过 LSLS 突变成 ATAT 对反应性表位进行修饰, 由此导致的免疫原性的降低按下面所述直接测定。模拟此序列的合成肽改变了经典的抗原提呈细胞如树突细胞 (DC) 的免疫应答。下面的合成肽

KSLSLSPGK-APTS (SEQ ID NO: 20) 和
KSATATPGK-APTS (SEQ ID NO: 21),

被用于刺激 DC-介导的向自体 T 细胞的抗原提呈, 此处连字符是 KS-IL2 的连接处。将这些 T 细胞响应随后此肽抗原攻击而增殖的能力用作此肽的免疫原性的测定结果。

具体地, 外周血单核细胞 (PBMC) 通过标准密度梯度技术从 leukopacks 中分离。单核细胞被重悬在无血清的 Aim V 培养基中, 使之贴壁。37℃, 2h 后去除未贴壁的细胞。贴壁细胞在含有人 GM-CSF(50ng/ml) 和 IL-4 (20 ng/ml) 的培养质中培养 7 天以产生未成熟的树突细胞 (DC)。7 天以后, 收获细胞, 用适当的 FITC 标记的针对 MHC I 类、MHC II 类、CD80 和 CD40 的抗体通过流式细胞计对细胞进行表型分析以确定未成熟的 DC 表型。

用 IL2 和 IL7 培养未贴壁细胞以获得自体效应细胞 (T 细胞) 用于其后的功能研究。为了进行功能研究, 将 T 细胞加入未成熟树突细胞 (10:1 比率) 中并与 huKS、去免疫化的 huKS、13 聚体肽连接处

(KSLSLSPGK-APTS)(SEQ ID NO:20)以及经修饰的去免疫化 13 聚体肽 (KSATATPGK-APTS)(SEQ ID NO:21)共同培养。暴露于各免疫细胞因子或免疫原性肽和经修饰去免疫化的肽之后,通过测定掺入的氘化胸苷比较增殖指数,以证明各分子的免疫原性强度。即,放射活性掺入量的增加大致与各肽与 DC 上的 MHC II 类分子结合的能力以及被提呈给 T 细胞的能力成比例。

实施例 5: 对清蛋白融合蛋白中的免疫原性反应性表位的推导和为了降低免疫原性在融合连接处对氨基酸残基的修饰

人血清清蛋白 (HAS) 因为具有异常长的半衰期、在体内广泛分布。且没有酶学或免疫学的功能,已被用作治疗性肽/蛋白的载体。遗传工程化的 HAS-CD4 杂合体已经显示可阻断人类免疫缺陷病毒进入 CD4 + 细胞,同时展示出类似于可溶性 CD4 的体外抗病毒特性 (Yeh 等, PNSA 89: 1904-1908, 1992)。因此,生物活性肽与 HSA 的遗传融合对于设计和回收分泌的治疗性 HSA 衍生物是有用的。但是,同所有融合蛋白一样, HAS-CD4 具有一个新的连接处,此连接处可能具有免疫原性而且包含能在 MHC II 类分子上被提呈的 T 细胞表位。用实施例 1, 2, 3 和 4 中的方法分析 HAS 和 CD4 之间的连接处,鉴定具有 MHC 结合可能性的肽。修饰这些潜在免疫原性序列来降低或消除潜在的 T 和 B 细胞表位从而降低免疫原性。类似地,可在连接区引入新的糖基化位点以降低免疫原性。

清蛋白序列

CD4 序列

TCFAEEGKKLVAASQAALGL – KKVVLGKKGDTVELTCTAS (SEQ ID NO: 22).

据本发明,预计 HSA-IFN α 融合蛋白连接区域含有 3 个候选 T 细胞表位,

KKLVAASQAALGL (SEQ ID NO: 13);
KLVAASQAALGLC (SEQ ID NO: 23); 和
LGLCDLPQTHSLG (SEQ ID NO: 24).

SEQ ID NO:13 和 23 中所述的 T 细胞表位相互重叠, 并可通过将 LV (粗体) 改变成除 F、I、L、M、V、W 和 Y 以外的任何氨基酸来进行去免疫化。备择地, 可通过将 LG 变成 TT 而显著地降低肽穿线分值。SEQ ID NO: 24 中的 T 细胞表位可通过将第二个 L (粗体) 变成 A 来去免疫化。

而且, 预计对于 HAS-X 融合物, 此处的 X 可以为任意蛋白质, 可通过将氨基酸序列 AALGL(SEQ ID NO:25)改变成 TATTA(SEQ ID NO:26)实现融合连接处的去免疫化。

CFABEGKKLVAAASQTATTA (SEQ ID NO: 27).

实施例 6: X-Fc 融合蛋白以及在融合连接处旨在降低免疫原性的氨基酸残基修饰

有时采用 X-Fc 取向构建融合蛋白具有特别的优越性。在这种结构中, 靶蛋白是 N 端融合蛋白, 其后为 Fc 片段。例如, 胰高血糖素样肽 (GLP-1) 的活性需要游离的 N 末端, 因此 GLP-1-Fc 融合是有用的。

GLP-1-Fc 融合蛋白按照本领域所述的标准技术构建。此融合蛋白具有与 $\gamma 1$ 重链的铰链连接的 GLP-1 的 C 末端。所用的 $\gamma 1$ 铰链序列含有 Cys 到 Ser 的突变 (残基 5), 这消除了与 IgG1 轻链形成二硫键的 Cys 残基 (Lo 等, (1998) 蛋白质工程 11: 495-500)。未突变的 Fc 序列是

EPKSCDKTHTCPPCPAPELLG (SEQ ID NO: 28)

其铰链区域以下划线显示, 其后为 CH2 结构域序列的起始。

GLP-1(7-37)与突变 Fc 之间的融合连接处是:

HAEGTFTSDVSSYLEGQAAKEFIAWLVKGRG -
EPKSDKTHTCPPCPAPELLG (SEQ ID NO: 29).

GLP-1(7-37)与正常 Fc 之间的融合连接处是:

SYLEGQAAKEFI~~AWLVKGRG~~ – EPKSCDKTHTCPPCPAPELLG (SEQ ID NO: 30)

肽穿线分析在 GLP-1-Fc 融合连接处鉴定出 3 个潜在表位。

KEFI~~AWLVKGRG~~E (SEQ ID NO: 31)
 EFI~~AWLVKGRG~~EP (SEQ ID NO: 32)
 A~~WLVKGRG~~EPKSS (SEQ ID NO: 33).

按实施例 1-3 进行 GLP-1 (粗体) 和 Fc (无加粗) 之间的融合连接处的分析, 鉴定出具有 MHC 结合潜力的肽。在通过肽穿线分析进行潜在位点的鉴定之后, 通过氨基酸替换修饰这些潜在的免疫原性序列以降低或消除潜在 T 和 B 细胞结合表位并降低免疫原性。

上面在 SEQ ID NOs:31、32 和 33 中提到的潜在 T 细胞表位可通过单氨基酸替换进行去免疫化。例如, SEQ ID NO:31 所示的肽通过将赖氨酸 (加粗显示) 改变成苏氨酸以及将精氨酸 (加粗显示) 改变成苏氨酸进行去免疫化。SEQ ID NO:32 所示的肽通过将异亮氨酸 (加粗显示) 替换成丙氨酸或脯氨酸进行去免疫化, SEQ ID NO:33 所示的肽通过用丙氨酸或脯氨酸替换亮氨酸进行去免疫化。得到的去免疫化的连接处为:

HAEGTFTSDVSSYLEGQAAKEF~~AAWAVTGTG~~ – EPKSSDKTHTCPPCPAPELLG (SEQ ID NO: 34).

根据用于在融合连接处引入糖基化位点的典型方法引入下面的变化:

SYLEGQAAKEFI~~AWLVKGRN~~ – ~~G~~SKSSDKTHTCPPCPAPELLG (SEQ ID NO: 35).

实施例 7: TNFR-Fc 融合蛋白 Enbrel 的免疫原性反应性表位的推断及在融合连接处旨在降低免疫原性的氨基酸残基修饰

ENBREL 或 etanercept 是 FDA 批准的 X-Fc 融合蛋白, 其是用于治疗

类风湿性关节炎的一种肿瘤坏死因子 (TNF) 抑制剂。ENBREL 是一个二聚体融合蛋白，由 TNF 受体的细胞外配体结合结构域连接人 IgG1 的 Fc 蛋白构成。TNFR-Fc 竞争性抑制 TNF 对其受体的结合使得结合的 TNF 丧失生物学活性，导致炎症活性显著降低。如以上对 GLP-1-Fc 的描述，TNF-Fc 也具有一个含有潜在 T 细胞表位的新连接处。

当 TNF-R 的 C 端部分 (粗体) 直接融合 g1 铰链的 N 端 (带下划线的未加粗字母为铰链区) 时，两者之间的连接处为

STSFLLLPMGSPPAEGSTGD - EPKSCDKTHTCPPCPAPPELLG (SEQ ID NO: 36)

按实施例 1-4 进行 TNF-R 与 Fc 之间的连接处的分析，鉴定出具有 MHC 结合潜力的肽。在通过肽穿线分析进行潜在位点的鉴定之后，通过氨基酸替换修饰潜在的免疫原性序列以降低或消除潜在 T 和 B 细胞结合表位并降低免疫原性。

根据用于在融合连接处引入糖基化位点的典型方法引入下面的变化：

STSFLLLPMGSPPAEGSTGN - GSKSCDKTHTCPPCPAPPELLG (SEQ ID NO: 37).

实施例 8: Fc-X-Y 融合蛋白如 Fc-IL12-IL2 的免疫原性反应性表位的推断及在融合连接处旨在降低免疫原性的氨基酸残基修饰

Fc-X-Y 取向的融合蛋白如 Fc-IL12-IL2 具有多个有潜在免疫原性的融合连接处。例如 Fc-IL12 具有类似其他 Fc-X 融合蛋白或免疫细胞因子 (实施例 1) 的融合连接处，但是因为细胞因子 IL12 的使用，此连接处为新的。因此对融合连接处进行免疫原性结合位点的分析和修饰。其次，由两个不同的细胞因子构成的融合蛋白 X-Y 的融合连接处类似实施例 5 中的描述。对每个融合连接处作肽穿线分析。

对位于 Fc 的 C 末端(加粗)与 IL12p35 的 N 末端 (未加粗) 之间的连接处：

(1) MHEALHNHYTQKSLSLSPGK - RNLPVATPDPGMFPCLLH SQ (SEQ ID NO: 38)

以及位于 IL12p40 的 C 末端(加粗)与 IL2 的 N 末端(未加粗)之间的连接处

(2) RAQDRYYSSSWSEWASVPCS - APTSSSTKKTQLQLEHLLLD (SEQ ID NO: 39)

用肽穿线法分析鉴定出具有 MHC 结合潜力的肽。修饰潜在的免疫原性序列以降低或消除 T 细胞表位。

例如，在上面的序列 (1) 中，进行下面的改变：

MHEALHNHYTQKSATATPGK - RNLPVATPDPGMFPCLLHHSQ (SEQ ID NO: 40).

这些变化降低或消除了位于 Fc 与 IL12 的 p35 亚基的连接处的几个 T 细胞表位的 MHC II 类结合潜力。

在另一个实例中，通过在 IL2 内头两个位置处引入天门冬酰胺和甘氨酸修饰上面的序列 (2)，引入糖基化位点。此策略使用了在成熟 IL2 中第 3 位的天然苏氨酸。而且，重要的是不打断在 p40 部分内的二硫键形成，这样用至少一个或两个氨基酸将 p40 内的半胱氨酸与糖基化位点分隔开是有用的。

RAQDRYYSSSWSEWASVPCS - NGTSSSTKKTQLQLEHLLLD (SEQ ID NO: 41).

对于 IL12p40-IL2 融合，按如上所讨论的引入糖基化位点产生如下潜在的 T 细胞表位：

SEWASVPCSNGTS (SEQ ID NO: 42)

ASVPCSNGTSSST (SEQ ID NO: 43)

但是, T 细胞表位的糖基化防止了 MHC II 类结合, 由此降低了免疫原性。

实施例 9: X-Fc-Y 融合蛋白连接处的免疫原性反应性表位的推断及在融合连接处旨在降低 MHC II 类结合的氨基酸残基修饰

X-Fc-Y 构型的融合蛋白, 如 IL4-Fc-GMCSF 具有多个含有潜在 T 细胞表位的新融合连接处。IL4-Fc 具有与其他 X-Fc 融合蛋白(实施例 6 和 7)类似的连接处, 但因为细胞因子 IL4 的使用其为新的。例如, 采用包含源于人 $\gamma 1$ 的铰链区、CH2 和 CH3 结构域的 Fc 形式。如上所述, pdCs-huFc $\gamma 1$ 中的 $\gamma 1$ 铰链序列可以含有 Cys 向 Ser 的突变(下划线)以消除与 IgG1 轻链形成二硫键的 Cys 残基(Lo 等, (1998) 蛋白质工程 11: 495-500), 由此产生了用于分析的第三潜在免疫原性融合连接处。依据实施例 1-4 的方法分析融合连接处的潜在 T 细胞表位并修饰。

Fc-Y 融合连接处与实施例 1 中对免疫细胞因子 huKS-IL2 的描述相类似, 只是由不同的细胞因子 GMCSF 构成融合蛋白。此融合连接处也按照实施例 1-4 的方法进行潜在 T 细胞表位分析及修饰。

具体地, 对位于 IL4 的 C 末端(加粗)与 Fc 的 N 末端(未加粗)之间的连接处:

(1) ENFLERLKTIMREKYSKCSS – epks~~c~~dktht~~c~~ppcpapellg (SEQ ID NO: 44)

以及位于 Fc 的 C 末端(加粗)与 GMCSF 的 N 末端(未加粗)之间的连接处

(2) MHEALHNHYTQKSLSLSPGK – parspspstqpwehvna~~i~~qe (SEQ ID NO: 45)

用肽穿线法分析鉴定出具有 MHC 结合潜力的肽。修饰潜在的 T 细胞表位以减少或消除潜在 T 细胞表位从而降低免疫原性。

IL4-Fc 融合蛋白的连接处的候选 T 细胞表位是,

EKYSKCSSEPKSC (SEQ ID NO: 46),

此处将 E (粗体) 改变成 T 显著地降低了肽穿线分值或 MHC II 类结合潜力。修饰后 IL4-Fc 融合序列如下:

ENFLERLKTIMREKYSKCSS - tpkscdkthtcppcpapellg (SEQ ID NO: 47).

如下所示, 将序列 LSLS 改变成 ATAT 使 Fc-GMCSF 的融合连接处去免疫化。

MHEALHNHYTQKSATATPGK - parspspstqpwehvnaige (SEQ ID NO: 48).

实施例 10: 用杂合同种型制备的免疫融合素和免疫细胞因子的融合连接处的氨基酸残基修饰以去除 T 细胞表位

构建具杂合同种型的抗体或者基于抗体的融合蛋白常常是有用的, 这样可将不同同种型的有用特性组合到单个分子中。带有杂合同种型的融合蛋白可根据本发明进行修饰以降低免疫原性。

具有下面组分的抗体融合蛋白可用标准重组 DNA 技术进行构建: 轻链和重链及识别肿瘤特异性抗原的 V 区, 其中轻链是典型轻链, 而重链包含源于 IgG2 的 CH1、CH2 和 CH3 结构域和源于 IgG1 的铰链区, 细胞因子通过上述融合连接处融合到重链 C 末端。

此蛋白含有位于 CH1g2 与铰链-g1 以及铰链-g1 与 CH2g2 之间的新连接处。按下面所述对这些连接处进行潜在的 T 细胞表位的鉴定和修饰。对于用 IgG2 或 IgG2h 同种型制备的免疫细胞因子和 Fc-X 融合蛋白, 这些修饰与上面实施例 1, 2, 3 和 8 中给出的一样。对于 X- Fc IgG2h 免疫融合素, 其也具有相同的新连接处, 因为 Fc 的 N 末端位于已经修饰成 IgG1 型

的 IgG2h 蛋白的铰链区内。但是，由 IgG1 铰链插入 IgG2 免疫球蛋白中产生两个新的融合连接处，这两个新的连接处位于 IgG2 CH1 与 IgG1 铰链之间和 IgG1 铰链与 IgG2 CH2 之间。

IgG2 CH1-IgG1 铰链- IgG2 CH2- IgG2 CH3-靶蛋白

因此，对位于 IgG2 CH1 的 C 末端（加粗）与 IgG1 铰链的 N 末端（未加粗）之间的连接处，

qytycnvdhkp^{snt}tkvdktv – epkscdktht^cppcp (SEQ ID NO: 49)

以及位于 IgG1 铰链 C 末端（加粗）与 IgG2 CH2 F 的 N 末端（未加粗）之间的连接处

epkscdktht^cppcp – appvagpsvflfppkpkdtl (SEQ ID NO: 50)

通过肽穿线法进行分析，应能鉴定出有 MHC 结合潜力的肽。对潜在的免疫原性序列进行修饰来降低或消除潜在的 T 和 B 细胞表位，以此降低免疫原性。

位于 IgG2CH1-IgG1 铰链融合连接处的两个潜在的 T 细胞表位为：

TKVDKTVEPKSCD (SEQ ID NO: 51) 和 KTVEPKSCDKTHT (SEQ ID NO: 52).

IgG2CH1-IgG1 铰链融合连接处通过将 V（加粗）改变成 A、T 或 P 而去免疫化。修饰的融合连接处的序列描述在 SEQ ID NO:53 中。

qytycnvdhkp^{snt}tkadk^ta – epkscdktht^cppcp (SEQ ID NO: 53).

如上所述，pdCs-huFc γ 1 中的 γ 1 铰链序列可以含有 Cys 向 Ser 的突变（下划线），此突变消除了与 IgG1 轻链形成二硫键的 Cys 残基（Lo 等,(1998)蛋白质工程 11: 495-500），由此产生 2 个额外的潜在免疫原性融合连接处供分析和修饰：

(3) qytycnvdhkpstktvdktv – epksS₂dkthtcppcp (SEQ ID NO: 54)

(4) epksS₂dkthtcppcp – appvagpsvflfppkpkdtl (SEQ ID NO: 55).

实施例 11: 使用来自 IgG1 和 IgG4 的杂合同种型 Fc 组分制备 Fc-EPO 融合蛋白

为了制备 Fc-促红细胞生成素融合蛋白, 用标准分子生物学技术构建了下列的表达质粒。使用含有 WO01/36489 中公开的人促红细胞生成素编码序列的 XmaI-XhoI DNA 片段, 此编码序列中有导致氨基酸替换 His32Gly、Cys33Pro、Trp88Cys 以及 Pro90Ala 的突变。相应的蛋白质序列如 SEQ ID NO:56 所示。

```
APPRLICDSRVLERYLLEAKEAENITTGCAEGPSLNENITVPDTKVNIFYAWKRMEVGQQAVEV
WQGLALLSEAVLRGQALLVNSSQPCEGLQLHVDKAVSGLRSLTTLRLALGAQKEAISPPDAAS
AAPLRITITADTFRKLFRVYSNFLRGKCLKLYTGEACRTGDR
```

将此 XmaI-XhoI DNA 片段插入编码源于 IgG1 的铰链区以及源于 IgG2 的 CH2 和 CH3 区的质粒载体中, 其中在 CH3 的 C 末端区域中有导致氨基酸替换的两组突变, 因而在 CH3 C 末端与 Epo N 末端的连接处的序列为:

.... TQKSATATPGA-APPRLI....(SEQ ID NO: 57)

第一组突变将 IgG2 CH3 区中的序列 KSLSLSPG (SEQ ID NO:58) 变成 KSATATPG (SEQ ID NO:59), 其在系列号为 60/280, 625 的 U.S. 专利申请中公开。以 Ala-Thr-Ala-Thr(SEQ ID NO:59 中位置 3 至位置 6) 替换 Leu-Ser-Leu-Ser(SEQ ID NO:58 中位置 3 至位置 6) 的效果是除去可能因为人 Fc 与人促红细胞生成素之间的连接处含有非己肽序列而产生的潜在人非己 T 细胞表位。第二组突变由 CH3 区 C 末端氨基酸处的 K 至 A 单

氨基酸替换组成，其在系列号为 09/780, 668 的 U.S.专利申请中公开。

将得到的质粒转染进 NS/0 细胞，对 Fc-Epo 融合蛋白按照本领域公知的方法进行表达和纯化。在基于与蛋白质 A 的结合进行纯化之后，将含有如上所述的 IgG2 CH3 和促红细胞生成素替换的 huFc γ 2h-huEpo 蛋白用大小排阻层析进行表征，发现其在两个独立制备物中由 97% 单体和 90% 单体组成。含有如上所述的 IgG2 CH3 以及促红细胞生成素替换的 huFc γ 2h-huEpo 蛋白被发现在基于细胞的实验中在摩尔基础上具有与人促红细胞生成素大约一样的活性，其中所述基于细胞的实验为测定促红细胞生成素蛋白刺激 TF-1 细胞分裂的能力。实验按照 WO01/36489 所述进行。

另外，还对非突变的人促红细胞生成素与由 IgG1 (铰链-CH2-CH3)、IgG2 (铰链-CH2-CH3) 或 IgG1 (铰链)-IgG2 (CH2-CH3) 组成的 Fc 区的 C 末端的融合物进行了表征。含有非突变的人 Fc 序列和非突变的促红细胞生成素序列的表达质粒类似于上述质粒进行构建。NS/0 细胞用 Fc γ 1-Epo, Fc γ 2-Epo, 以及 Fc γ 2h-Epo 表达质粒进行转染，在对每种质粒进行大致相等克隆数目的筛选之后，分离出稳定的克隆。最佳生产克隆产出 50 μ g/ml 的 Fc γ 1-Epo, 20 μ g/ml 的 Fc γ 2-Epo, 以及 120 μ g/ml 的 Fc γ 2h-Epo。

下面的实施例详细地描述了在本发明公开的融合蛋白的序列中鉴定免疫原性序列区域 (T 细胞表位) 的一种优选方法。但是，应当指出所述的分子可用其他已知的方法获得。

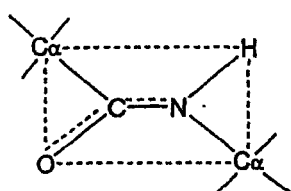
实施例 12 计算机方法对 T 细胞表位的鉴定

依据本发明，融合蛋白连接区域中的表位可通过向蛋白内引入突变以调节它们同免疫系统的相互作用的方法进行修饰。根据本发明，可适用于本发明的本领域内公知的方法包括现有技术中描述的那些 (WO92/10755 和 WO96/40792 (Novo Nordisk), EP 0519 596 (Merck & Co.), EP 0699755 (Centro de Imunologia Molecular), WO 98/52976 和 WO 98/59244 (Biovation Ltd.) 或相关方法。

但是，如果所述表位的鉴定是通过以下详细描述的新方法实施的并应用于本发明融合蛋白的连接区域，则可获得有利的突变蛋白。

有多种因素对决定蛋白或多肽的总体结构起重要作用。首先是肽键，即将氨基酸连接在一起形成链的键，它是一种共价键。这种键是平面结构的，实质上是一种取代的酰胺。“酰胺”指含-CONH-基团的一组有机化合物中的任何一个化合物。

连接相邻氨基酸的 $C\alpha$ 的平面肽键如下所示：



由于 $O=C$ 和 $C-N$ 原子位于一个相对刚性的平面中，所以不会发生沿这些轴的自由旋转。因此，图中虚线所示的平面有时被称作“酰胺”平面或“肽平面”，肽主链中的氧(O)、碳(C)、氮(N)和氢(H)原子位于其中。 $C\alpha$ 原子位于酰胺平面中相对的角上。由于肽或酰胺平面中的 $O=C$ 和 $C-N$ 原子基本上不发生旋转，所以多肽链包含一系列连接 $C\alpha$ 原子的平面肽键。

第二个对决定多肽或蛋白的整体结构或构象起重要作用的因素是每一酰胺平面绕共有 $C\alpha$ 键的转角。此后术语“转角”和“扭转角”是等同的术语。假定 O、C、N 和 H 原子保留在酰胺平面中(这通常是一种正确的假设，尽管在一些构象中这些原子会轻微的偏移平面)，这些转角确定了 N 和 R 多肽主链构象，即相邻残基之间的结构。这两个转角称为 ϕ 和 ψ 。因此，一套 ϕ_i 和 ψ_i 角(其中，脚标 i 代表多肽链中的特定残基)有效地规定了多肽链的二级结构。在文献中定义了用于确定 ϕ 和 ψ 角的惯例，即在给定的多肽中酰胺平面形成 0 度角的参考点，以及哪个角是 ϕ 角，哪个角是 ψ 角的定义。参见 Ramachandran 等, Adv. Prot. Chem. 23: 283-437 (1968), 285-94 页，这些页中的内容在此引入作为参考。

本发明的方法可应用于任何蛋白，并部分基于下述发现，即人 MHC

II 类分子结合沟的主要口袋 1 锚定位点对特定氨基酸侧链具有设计好的特异性。这一口袋的特异性由 MHC II 类分子 β 链第 86 位的氨基酸的身份来确定。这一位点位于口袋 1 的底部并决定可容纳于这一口袋中的氨基酸侧链的大小。Marshall, K. W., J. Immunol., 152: 4946-4956 (1994)。如果这一残基是甘氨酸,则所有的疏水性脂肪族和芳香族氨基酸(疏水性脂肪族氨基酸是:缬氨酸、亮氨酸、异亮氨酸、甲硫氨酸,芳香族氨基酸是:苯丙氨酸、酪氨酸和色氨酸)均可容纳于所述的口袋中,优选芳香族侧链。如果这一口袋残基是缬氨酸,则该氨基酸的侧链伸到口袋中并限制了可容纳的肽侧链的大小,所以只有疏水性脂肪族侧链可容纳进去。因此在氨基酸残基序列中,无论哪里发现了带有疏水性脂肪族或芳香族侧链的氨基酸,即有存在 MHC II 类限制性 T-细胞表位的可能性。但是,如果所述的侧链是疏水性脂肪族侧链,其与 T-细胞表位相关的可能性约是芳香族侧链的两倍(假定 1 型口袋近似平均地分布于全球种群中)。

将本发明具体化的计算机方法描绘出肽区域包含 T-细胞表位的可能性,该方法如下:(1)扫描预定长度肽片段的一级序列,并鉴定存在的所有疏水性脂肪族和芳香族侧链。(2)对疏水性脂肪族侧链赋予比芳香族侧链高的值;优选约两倍于赋予芳香族侧链的值,例如,给疏水性脂肪族侧链赋值为 2,给芳香族侧链赋值为 1。(3)将所述肽中的预定统一长度的每一重叠氨基酸残基片段(窗口)中确定存在的值总和起来,再将某一特定片段(窗口)的总值赋予该片段(窗口)中间位置的某个单个氨基酸残基,优选赋予大约处于抽样片段(窗口)中间点的氨基酸。将这一过程对每一抽样的重叠氨基酸残基片段(窗口)重复进行。因此,所述肽的每一氨基酸残基均被赋予了一个值,该值与 T-细胞表位存在于此特定片段(窗口)中的可能性相关。(4)用按照上述步骤 3 中的描述计算、赋予的值对被评估的整个氨基酸残基序列的氨基酸坐标作图。(5)序列中具有预定值(例如该值为 1)的所有部分均被认为可能包含 T 细胞表位,并且在需要时可进行修饰。在这一方面本发明提供了通用的方法,由此可描述可

能包含 T-细胞表位的肽区域。在这些区域中对所述的肽进行修饰有可能改变 MHC II 类的结合特性。

依照本发明的另一方面，可利用考虑了肽与 MHC II 等位基因模型之间的相互作用的更复杂计算方法来更精确地预测 T-细胞表位。根据这一方面，对肽中存在的 T 细胞表位的计算预测考虑到：基于所有已知 MHC II 类分子的结构构建至少 42 个 MHC II 类等位基因模型；使用这些模型计算鉴定 T 细胞表位的方法；对每一模型构建肽主链文库以允许在相关肽主链 α 碳($C\alpha$)位置具有已知的变异性；在肽和 MHC II 类分子间相互作用关键的位置处，对与每一模型对接 (dock) 的每一主链，相对于 20 种氨基酸选项中每一种构建氨基酸侧链构象文库；以及将这些主链和侧链构象文库与评分函数结合用于选择与特定 MHC II 类分子对接的特定肽的最佳主链和侧链构象并得出该相互作用的结合分值。

MHC II 类分子模型可从 Brookhaven 蛋白数据库("PDB")中的许多类似的结构出发通过同源建模推导得出。它们可通过使用引入了模拟退火算法的半自动同源建模软件(Modeller, Sali A. & Blundell TL., 1993. J. Mol Biol 234 : 779-815)并结合用于能量最小化的 CHARMM 力场(购自 Molecular Simulations Inc., San Diego, Ca.)来制备。也可以应用其他的建模方法。

本发明的方法与下述的其他计算方法有着显著的不同，这些方法在于：利用从实验中得来的关于一小组 MHC II 类分子结合沟中每一位点的每一种氨基酸选项的结合数据文库 (Marshall, K. W.等, Biomed. Pept. Proteins Nucleic Acids, 1 (3): 157-162) (1995)；或利用类似的实验结合数据以定义所述的沟中特定结合口袋类型的结合特性(同样利用相对小的 MHC II 类分子亚组)然后将这一口袋文库中的口袋类型进行‘混合和匹配’以人工构建更“实际的”MHC II 类分子 (Sturniolo T.等, Nat. Biotech, 17 (6): 555-561 (1999)。这两种现有方法的主要缺陷在于实验的复杂性和需要合成大量的肽变体造成仅有少量的 MHC II 类分子可通过实验扫描。因此第一种已知的方法仅能预测少量的 MHC II 类分子。第

二种已知的方法还假设在不同 II 类等位基因的背景下在一个分子中衬有类似氨基酸的口袋将具有相同的结合特性，故其另外的缺陷在于，仅仅可“实际地”地构建出那些包含口袋文库中所包含的口袋的 MHC II 类分子。利用本发明的建模方法可推导出任意数量和类型的 MHC II 类分子的结构，因此可特异性地选择等位基因以代表全球种群的特征。此外，扫描的 MHC II 类分子的数量可通过构建更多的模型而增加而无需通过复杂的实验获得额外的数据。利用主链文库使得被扫描的各种肽在与特定的 MHC II 类分子结合时其 $C\alpha$ 原子位置处可进行变化。这也与上述现有技术中的计算机方法不同，在那些方法中依赖于利用简化的肽主链来扫描结合在特定口袋中的氨基酸。这些简化的主链不可能代表在“真正的”肽中存在的主链构象，从而导致对肽结合的预测不准确。本发明的主链文库是通过叠加蛋白数据库中所有与 MHC II 类分子结合的肽的主链，并考虑到位于结合沟内的 11 个氨基酸的每个氨基酸的 $C\alpha$ 原子之间的均方根(RMS)差而构建的。尽管该文库可来自少量合适的可获得的小鼠和人的结构(当前为 13 种)，但为了允许存在甚至更大变异的可能性，将每一 $C''-\alpha$ 位点的 RMS 数提高 50%。然后确定每一氨基酸的平均 $C\alpha$ 位置，围绕这一点划一个球，其半径等于在该位置的 RMS 差加 50%。该球体代表所有可允许的 $C\alpha$ 位置。

自具有最小 RMS 差的 $C\alpha$ (上述口袋 1 中氨基酸残基的 $C\alpha$ ，等同于结合沟中 11 个残基的位置 2)起运作，将所述的球三维网格化，网格内的每个顶点作为该氨基酸的 $C\alpha$ 的可能位置。将后续的酰胺平面（相应于与后续氨基酸的肽键）移动到这些 $C\alpha$ 的每一个上面，将 ϕ 和 ψ 角以设定的间隔逐步地转动以便于安置后续的 $C\alpha$ 。如果后续的 $C\alpha$ 落入对这一 $C\alpha$ 而言‘可被允许的位置球’中，则此二肽的方向即可被接受，如果其落入所述球之外则所得的二肽不能被接受。对每一后续 $C\alpha$ 位置均重复这一过程，使肽从所述的口袋 1 $C\alpha$ ‘种子’开始生长，直到全部 9 个后续的 $C\alpha$ 的位置均根据之前 $C\alpha$ 的所有可能排列确定下来。然后对口袋 1 前的单个 $C\alpha$ 重复上述步骤 1 次以上以构建定位于结合沟内的主

链 C α 位置文库。

生成的主链数目取决于几种因素：‘可被允许的位置球’的大小；对口袋1位点处‘最初的球’网格化的细度；用于定位后续 C α 的 ϕ 和 ψ 角逐步旋转的细度。利用这一程序可以构建大的主链文库。主链文库越大越可能发现对 MHC II 类分子结合沟内的特定肽而言的最适主链。鉴于和结合结构域的氨基酸可能存在冲突，所以不是所有的主链均适合于与所有 MHC II 类分子模型‘对接’（docking），故对每个等位基因建立亚文库以包含适合被该等位基因容纳的主链。利用所述的主链文库并结合 MHC II 类分子模型可以构建出由与每一容许主链对接的每一 MHC II 类分子结合沟的每一位点中的每一氨基酸的容许侧链构象所组成的详尽数据库。可以利用简单的立体重叠函数构建这一数据组，其中，主链与 MHC II 类分子对接，氨基酸侧链在所需位置被嫁接到主链上。将侧链上可旋转的键以设定的间隔逐步旋转，记录下依赖于该键的原子的最终定位。将所述原子与结合沟侧链原子间的相互作用记录下来，根据下述的标准确定是否接受这些位置：如此定位的所有原子的重叠总量不能超过预定值。因此，构象搜索的严谨度是在键的逐步旋转中所用的间隔及对总重叠的预定限度的函数。如果已知特定的口袋是刚性的，则后一值可较小，但若已知口袋侧链的位置相对灵活则严谨度可放松。这样便可以模拟结合沟口袋内灵活性的变化。针对与每一 MHC II 类分子对接后每一主链的所有位点上的所有氨基酸重复这种构象搜索以建立详尽的侧链构象数据库。

用适当的数学表达式评价 MHC II 类分子模型与肽配体构象的结合能量，所述的肽配体构象需通过扫描上述的主链/侧链构象大数据库根据经验获得。这样，通过对每一长度在 9-20 个氨基酸范围内变化（尽管对于每一次扫描长度是一定的）的可能肽进行下述计算，可以扫描蛋白以搜索潜在的 T-细胞表位：选择 MHC II 类分子及适合于该分子的肽主链，将相应于所需肽序列的侧链移植到其上。对于氨基酸的每一容许构象（由上述数据库获得），收集与主链上特定位点的特定侧链相关的原子

身份和原子间距数据。沿主链对每一侧链重复此过程，利用评分函数推导肽得分。保留该主链的最佳得分，对所选模型的每一容许主链重复该过程。比较所有容许主链的得分，最高的得分被认为是该 MHC II 类模型中所需肽的得分。对每一模型用从扫描的蛋白得到的所有可能肽重复上述过程，列出肽相对于模型的得分。

在本发明中，用于结合亲和力计算的每种配体都是选自上述肽或蛋白的氨基酸片段。因此所述配体为来自已知序列的肽、多肽或蛋白的长度为约 9 到 20 个氨基酸的选定氨基酸链。此后术语“氨基酸”和“残基”视为等同的术语。将移植到选自上述主链文库的主链上的待检测肽中的连续氨基酸形式的配体，通过肽主链上 C^α-α 原子坐标定位到来自 MHC II 类分子模型库的 MHC II 类分子的结合裂缝中，并从容许构象的数据库中选择每一侧链的容许构象。相关的原子身份和原子间距也可以从这一数据库获得并用于计算肽结合分数。将对 MHC II 类结合口袋具有高亲和力的配体作为候选者标记出来用于定点诱变。在标记的配体中（也由此在目的蛋白中）进行氨基酸替代，然后用评分函数重新测定以确定使结合亲和力降低到预定的阈值以下的变化。这些变化即可引入到目的蛋白中以去除 T-细胞表位。

肽配体与 MHC II 类分子结合沟的结合涉及非共价键相互作用，其包括但不限于：氢键、静电相互作用、疏水(亲脂)相互作用和范德华相互作用。它们被包括在下面将详细描述肽评分函数中。应当理解，氢键是非共价键，其可在极性或带电的基团之间形成，由被两个其他原子共享的氢原子构成。氢供体中的氢带正电荷，而氢受体带有部分负电荷。为肽/蛋白相互作用的目的，氢键供体可以是连接氢的氮，或连接在氧或氮上的氢。氢键受体原子可以是没有连接氢的氧、没有连接氢并具有一或两个连接的氮或仅有一个连接的硫。某些原子，如连接了氢的氧或亚胺氮(如 C=NH)，既可以是氢受体也可以是氢供体。氢键的能量在 3-7 Kcal/mol，大大强于范德化键，但弱于共价键。氢键具有高度的方向性，且当供体原子、氢原子和受体原子共线时最强。静电键是在带有相反电

荷的离子对间形成的,根据库仑定律这种相互作用的强度与原子间距离的平方成反比。离子对间的最佳距离是约 $2.8\text{\AA}$ 。在肽/蛋白相互作用中,可在精氨酸、组氨酸或赖氨酸和天冬氨酸或谷氨酸之间形成静电键。该键的强度依赖于电离基团的 pK_a 和介质的介电常数,尽管其与氢键的强度类似。

亲脂相互作用是蛋白和肽配体之间发生的有利疏水-疏水相互作用。这种相互作用通常出现在埋于结合沟口袋中的肽的疏水性氨基酸侧链间,以使它们不暴露在溶剂中。将疏水残基暴露于溶剂中是非常不利的,因为周围的溶剂分子将被迫在彼此间形成氢键而形成笼状结构。所致的熵值降低是非常不利的。亲脂性原子可以是既非极性又不是氢受体的硫和非极性的碳原子。

范德华键是相距 $3\text{-}4\text{\AA}$ 的原子间的非特异性的力。它比氢键和静电键弱、特异性低。原子周围的电荷分布随时间变化,并且在任何瞬间电荷分布均是不对称的。这种瞬间的电荷不对称性诱导临近原子中的类似不对称性。在范德华接触距离中所导致的原子之间的吸引力达到最大,而在约 $1\text{\AA}$ 到 $2\text{\AA}$ 处迅速消失。相反,当原子间隔的距离小于此接触距离时,由于原子外部的电子云重叠使不断增强的斥力成为主导。尽管与静电和氢键相比,此吸引力相对较弱(约 0.6 Kcal/mol),但所述斥力对于决定肽配体是否能与蛋白成功结合可能非常重要。

在一个实施方案中,利用 Böhm 评分函数(SCORE1 方法)评估结合常数(Böhm, H. J., J. Comput Aided Mol. Des., 8 (3): 243-256 (1994), 该文献在此全文引入作为参考)。在另一个实施方案中,用评分函数(SCORE2 方法)评估结合亲和力作为配体含有 T-细胞表位的指示物(Böhm, H. J., J. Comput Aided Mol. Des., 12 (4): 309-323 (1998), 该文献在此全文引入作为参考)。但是上述文献中描述的 Böhm 评分函数是用于评估下述情况中配体对蛋白的结合亲和力的,即已知所述的配体可成功地与所述蛋白结合,且蛋白/配体复合物的结构已解析,这一结构已列于蛋白数据库("PDB")中。因此,利用已知的阳性结合数据对评分函

数作了发展。为了区分阳性和阴性的结合体，需向方程中加入排斥项。此外，可通过以成对的方式计算亲脂相互作用，而非利用上述 Böhmer 函数中基于面积的能量项来进行更理想的结合能量评估。因此，在一个优选实施方案中，用经修饰的 Böhmer 评分函数评估结合能。在经修饰的 Böhmer 评分函数中，在评估蛋白和配体之间的结合能(ΔG_{bind})时考虑了下述参数：由于配体的平移和转动熵的整体损失造成的结合能减低(ΔG_0)；理想氢键的贡献(ΔG_{hb})，其中至少一个配对物是中性的；无扰离子相互作用的贡献(ΔG_{ionic})；亲脂配体原子和亲脂受体原子之间的亲脂相互作用(ΔG_{lipo})；由于配体中内在自由度的冻结，即绕每一 C-C 键的旋转自由度降低而造成的结合能损失(ΔG_{rot})；蛋白和配体之间相互作用的能量(E_{vdw})。考虑到这些项给出等式 1：

$$(\Delta G_{\text{bind}}) = (\Delta G_0) + (\Delta G_{\text{hb}} \times N_{\text{hb}}) + (\Delta G_{\text{ionic}} \times N_{\text{ionic}}) + (\Delta G_{\text{lipo}} \times N_{\text{lipo}}) + (\Delta G_{\text{rot}} + N_{\text{rot}}) + (E_{\text{vdw}})$$

其中 N 是对于特定项符合的相互作用的数目，在一个实施方案中， ΔG_0 、 ΔG_{hb} 、 ΔG_{ionic} 、 ΔG_{lipo} 和 ΔG_{rot} 是常数，其值分别为：5.4、-4.7、-4.7、-0.17 和 1.4。

N_{hb} 项依照等式 2 计算：

$$N_{\text{hb}} = \sum_{\text{h-bond}} f(\Delta R, \Delta \alpha) \times f(N_{\text{neighb}}) \times f_{\text{pcs}}$$

$f(\Delta R, \Delta \alpha)$ 是罚函数，其解决氢键自理想情况的巨大偏离，其依照等式 3 计算：

$$f(\Delta R, \Delta \alpha) = f_1(\Delta R) \times f_2(\Delta \alpha)$$

其中：

如果 $\Delta R \leq \text{TOL}$ 则 $f_1(\Delta R) = 1$ ，或者

如果 $\Delta R \leq 0.4 + \text{TOL}$ 则 $f_1(\Delta R) = 1 - (\Delta R - \text{TOL})/0.4$ ，

或者

如果 $\Delta R > 0.4 + \text{TOL}$ 则 $f_1(\Delta R) = 0$

并且：

如果 $\Delta\alpha < 30^\circ$

则 $f_2(\Delta\alpha) = 1$, 或者

如果 $\Delta\alpha \leq 80^\circ$

则 $f_2(\Delta\alpha) = 1 - (\Delta\alpha - 30)/50$, 或者

如果 $\Delta\alpha > 80^\circ$

则 $f_2(\Delta\alpha) = 0$

TOL 是氢键键长中所能允许的偏差 = $0.25\text{\AA}$

ΔR 是 H-O/N 氢键键长与理想值 = $1.9\text{\AA}$ 的偏差

$\Delta\alpha$ 是氢键键角 $\angle_{N/O-H..O/N}$ 与 180° 理想值的偏差

$f(N_{\text{neighb}})$ 区分蛋白表面的凹凸部分, 并因此赋予口袋中的极性相互作用比蛋白表面的极性相互作用更高的权重。这一函数根据下述等式 4 计算:

$$f(N_{\text{neighb}}) = (N_{\text{neighb}}/N_{\text{neighb},0})^\alpha, \text{ 其中 } \alpha = 0.5$$

N_{neighb} 为蛋白中与任意给定蛋白原子之间的距离小于 $5\text{\AA}$ 的非氢原子的数量。

$N_{\text{neighb},0}$ 是常数 = 25

f_{pcs} 是考虑到每氢键的极性接触表面面积的函数, 由此可以区分强和弱的氢键, 其值由下述的标准确定:

$$\text{当 } A_{\text{polar}}/N_{\text{HB}} < 10\text{\AA}^2 \text{ 时} \quad f_{\text{pcs}} = \beta \text{ 或}$$

$$\text{当 } A_{\text{polar}}/N_{\text{HB}} > 10\text{\AA}^2 \text{ 时} \quad f_{\text{pcs}} = 1$$

A_{polar} 是极性蛋白-配体接触面的大小

N_{HB} 是氢键的数目

β 是常数 = 1.2

由于假定了相同的几何相关性, 在实施经修饰的 Böhm 评分函数时, 离子相互作用的贡献 ΔG_{ionic} 用与上述有关氢键的类似方式计算。

N_{lipo} 项按下述的等式 5 计算:

$$N_{\text{lipo}} = \sum_{\text{IL}} f(r_{\text{IL}})$$

根据下述标准, 对于所有亲脂配体原子 I 和所有亲脂蛋白原子 L, 计算 $f(r_{\text{IL}})$:

$$\text{当 } r_{\text{IL}} \leq R_1 \text{ 时}$$

$$f(r_{\text{IL}}) = 1$$

当 $R2 < r_{IL} > R1$ 时

$$f(r_{IL}) = (r_{IL} - R1)/(R2 - R1)$$

当 $r_{IL} \geq R2$ 时

$$f(r_{IL}) = 0$$

其中: $R1 = r_1^{vdw} + r_L^{vdw} + 0.5$

$$R2 = R1 + 3.0$$

r_1^{vdw} 是原子 I 的范德华半径

r_L^{vdw} 是原子 L 的范德华半径

N_{rot} 项是氨基酸侧链中可旋转的键的数目, 其被视为无环的 sp^3-sp^3 及 sp^3-sp^2 键的数目。末端 $-CH_3$ 或 $-NH_3$ 的旋转未考虑进去。

最终, 项 E_{vdw} 依照如下等式 6 计算:

$$E_{vdw} = \epsilon_1 \epsilon_2 ((r_1^{vdw} + r_2^{vdw})^{12}/r^{12} - (r_1^{vdw} + r_2^{vdw})^6/r^6),$$

其中: ϵ_1 和 ϵ_2 是取决于原子身份的常数

$r_1^{vdw} + r_2^{vdw}$ 是范德华原子半径

r 是原子对间的距离。

关于式 6, 在一个实施方案中, ϵ_1 和 ϵ_2 常数被赋予如下原子值, 分别为: C: 0.245, N: 0.283, O: 0.316, S: 0.316 (即分别对于碳、氮、氧和硫原子)。对于式 5 和 6, 给予范德华半径如下原子值, 分别为 C: 1.85, N: 1.75, O: 1.60, S: 2.00 Å。

应当理解上述等式中所有预定的值和给定的常数都是在现有的对蛋白配体相互作用的理解局限内具体相对于此处所用的计算类型确定的。因此, 随着这种评分函数的进一步精练, 这些值和常数也会因此而改变, 任何能在蛋白和配体结合能的评估方面给出所需结果的适宜数值均可使用, 而且, 其也落入本发明的保护范围内。

如上所述, 所述的评分函数应用于由上述侧链构象、原子身份和原子间距数据库中提取的数据。为本说明书的目的, 该数据库中包含的 MHC II 类分子数是 42 个模型加上 4 个已解析的结构。从上述描述中可清楚地了解到, 本发明的计算构建方法的模块性质意味着, 可简单地添加新的模型, 并利用肽主链文库和侧链构象搜索功能进行扫描以创建其它的可通过上述的肽评分函数处理的数据集。这使得被扫描的 MHC II

类分子库可以很容易地增加，或者如果可以获得相关数据，则可以替换结构和相关数据以创建现有等位基因的更精确的模型。

本发明的预测方法可以相对于包含大量已通过实验确定了其对不同 MHC II 类分子的亲和力的肽的数据集进行校准。将计算值与实验数据相比较，可确定一截断值，已知该值之上所有经实验确定的 T-细胞表位都得以正确的预测。

应当理解，尽管上述评分函数与现有的一些复杂方法相比相对简单，但计算进行得非常迅速。还应当理解的是，其目的并不在于计算出对接到所选择的 MHC II 类蛋白结合沟内的每种肽的真正结合能本身。根本的目的在于获得相对的结合能数据以助于根据所选蛋白的一级结构(即氨基酸序列)预测 T-细胞表位的定位。相对高的结合能或结合能高于选定的阈值意味着在配体中存在 T-细胞表位。然后可以将所述配体进行至少一轮氨基酸替代，并再次计算结合能。由于计算可迅速进行，对肽序列的这些操作可在现有成本划算的计算机硬件上于程序用户界面中互动进行。由此不需要对计算机硬件进行大量投资。

本领域的技术人员应当了解，也可使用其他软件达到相同的目的。特别是可以使用能将配体对接入蛋白结合位点的更复杂的软件，并与能量最小化相结合。对接软件的例子包括：DOCK (Kuntz 等, J. Mol. Biol., 161: 269-288 (1982)), LUDI (Böhm, H. J., J. Comput Aided Mol. Des., 8: 623-632 (1994))和 FLEXX (Rarey M.等, ISMB, 3: 300-308 (1995))。分子建模和操作软件的例子包括：AMBER (Tripos)和 CHARMM (Molecular Simulations Inc.)。使用这些计算方法将严重限制本发明方法的信息吞吐量，这是由于进行必要的计算所需的处理时间导致的。但是可行的方式为，将这些方法作为‘二级筛选’以获得关于通过本发明的方法发现为‘阳性结合体’的肽的更精确的肽结合能计算值。用于复杂的分子机械或分子动力学计算的处理时间的限制性是由进行所述计算的软件设计和目前计算机硬件技术的限制共同确定的。可以预期在将来，随着编写更高效的代码和计算机处理器速度的不断提高，在

更易控制的时间框架内进行上述计算是可行的。有关用于大分子的能量函数的其他信息和有关在折叠蛋白结构内发生的多种相互作用的考虑可参考下述文献:Brooks, B. R., 等., J. Comput. Chem., 4 : 187-217 (1983), 有关蛋白-配体一般相互作用的信息参见: Dauber-Osguthorpe 等, Proteins 4 (1) : 31-47 (1988), 这些文献均全文引入作为参考。其他有用的背景资料也可参见 Fasman, G. D.编, Prediction of Protein Structure and the Principles of Protein Conformation, Plenum Press, New York, ISBN: 0-306 4313-9.

等同方案

本发明可以以没有偏离其精神或基本特征的其他特殊形式体现。因此前面的实施方案无论从哪一方面来看都只是举例说明而非对此处描述的本发明的限定。因此本发明由所附权利要求而非前面的描述指示范围, 而且所有落入权利要求的等同含义和范围内的变化均被包含于其中。

参考文献的并入

上面提到的所有专利、专利申请和科学出版物均以参考文献的形式完整地在此并入本申请。

序列表

<110> 默克专利有限公司
 <120> 降低融合蛋白的免疫原性
 <130> LEX-017PC
 <150> US 60/280,625
 <151> 2001-03-30
 <160> 59
 <170> PatentIn version 3.0
 <210> 1
 <211> 330
 <212> PRT
 <213> 人(Homo sapiens)

<220>
 <221> misc_feature
 <223> 人 IgY 重链 C 区

<400> 1

Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Ser	Ser	Lys	1	5	10	15
Ser	Thr	Ser	Gly	Gly	Thr	Ala	Ala	Leu	Gly	Cys	Leu	Val	Lys	Asp	Tyr	20	25	30	
Phe	Pro	Glu	Pro	Val	Thr	Val	Ser	Trp	Asn	Ser	Gly	Ala	Leu	Thr	Ser	35	40	45	
Gly	Val	His	Thr	Phe	Pro	Ala	Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	50	55	60	
Leu	Ser	Ser	Val	Val	Thr	Val	Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	65	70	75	80
Tyr	Ile	Cys	Asn	Val	Asn	His	Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	85	90	95	
Lys	Val	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	100	105	110	
Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	115	120	125	
Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	130	135	140	
Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	145	150	155	160
Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	165	170	175	

Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu
 180 185 190
 His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
 195 200 205
 Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly
 210 215 220
 Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu
 225 230 235 240
 Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
 245 250 255
 Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn
 260 265 270
 Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
 275 280 285
 Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
 290 295 300
 Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
 305 310 315 320
 Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 325 330

<210> 2

<211> 326

<212> PRT

<213> 人

<220>

<221> misc_feature

<223> 人 Ig γ-2 链 C 区

<400> 2

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
 1 5 10 15

Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
 20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
 35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
 50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Asn Phe Gly Thr Gln Thr
 65 70 75 80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys
 85 90 95

```

Thr Val Glu Arg Lys Cys Cys Val Glu Cys Pro Pro Cys Pro Ala Pro
      100                      105                      110

Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp
      115                      120                      125

Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp
      130                      135                      140

Val Ser His Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp Gly
      145                      150                      155                      160

Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe Asn
      165                      170                      175

Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Val His Gln Asp Trp
      180                      185                      190

Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu Pro
      195                      200                      205

Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln Pro Arg Glu
      210                      215                      220

Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn
      225                      230                      235                      240

Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile
      245                      250                      255

Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr
      260                      265                      270

Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys
      275                      280                      285

Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys
      290                      295                      300

Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu
      305                      310                      315                      320

Ser Leu Ser Pro Gly Lys
      325

```

<210> 3

<211> 362

<212> PRT

<213> 人

<220>

<221> misc_feature

<223> 人 Ig3 恒定区

<400> 3

```

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1                      5                      10                      15

```

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
 20 25 30
 Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
 35 40 45
 Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
 50 55 60
 Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr
 65 70 75 80
 Tyr Thr Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys
 85 90 95
 Arg Val Glu Leu Lys Thr Pro Leu Gly Asp Thr Thr His Thr Cys Pro
 100 105 110
 Arg Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
 115 120 125
 Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg Cys
 130 135 140
 Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
 145 150 155 160
 Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
 165 170 175
 Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Lys Trp
 180 185 190
 Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Trp Glu
 195 200 205
 Glu Gln Tyr Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Leu
 210 215 220
 His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
 225 230 235 240
 Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly
 245 250 255
 Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu
 260 265 270
 Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr
 275 280 285
 Pro Ser Asp Ile Ala Met Glu Trp Glu Ser Ser Gly Gln Pro Glu Asn
 290 295 300
 Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe
 305 310 315 320

Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn
325 330 335

Ile Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr
340 345 350

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
355 360

<210> 4

<211> 327

<212> PRT

<213> 人

<220>

<221> misc_feature

<223> Ig Y-4 链C 区

<400> 4

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
1 5 10 15

Ser Thr Ser Glu Ser Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Lys Thr
65 70 75 80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys
85 90 95

Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro Ala Pro
100 105 110

Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
115 120 125

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val
130 135 140

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp
145 150 155 160

Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe
165 170 175

Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp
180 185 190

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu
195 200 205

```

Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg
 210                      215                      220

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys
225                      230                      235                      240

Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp
                245                      250                      255

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys
                260                      265                      270

Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
                275                      280                      285

Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser
                290                      295                      300

Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser
305                      310                      315                      320

Leu Ser Leu Ser Leu Gly Lys
                325

```

```

<210> 5
<211> 9
<212> PRT
<213> 人工序列

```

```

<220>
<223> 潜在 T 细胞表位

```

```

<400> 5
Lys Ser Leu Ser Leu Ser Pro Gly Lys
1                      5

```

```

<210> 6
<211> 9
<212> PRT
<213> 人工序列

```

```

<220>
<223> 突变的潜在 T 细胞表位

```

```

<400> 6
Lys Ser Ala Thr Ala Thr Pro Gly Lys
1                      5

```

```

<210> 7
<211> 9
<212> PRT
<213> 人工序列

```

```

<220>
<223> 突变的潜在 T 细胞表位

```

```

<400> 7

```

```

Lys Ser Ala Thr Ala Thr Pro Gly Ala
1                5

<210> 8
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> IgA Fc 区 CH3 结构域的 C 末端附近序列

<400> 8
Gln Lys Thr Ile Asp Arg Leu Ala Gly Lys Pro Thr His
1                5                10

<210> 9
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> 突变的 IgA Fc 区 CH3 结构域的 C 末端附近序列

<400> 9
Gln Lys Thr Ala Asp Arg Thr Ala Gly Lys Pro Thr His
1                5                10

<210> 10
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> IgA-X 融合物中的去免疫化序列

<400> 10
Gln Lys Thr Pro Thr Arg Thr Ala Gly Lys Pro Thr His
1                5                10

<210> 11
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> IgA-X 融合物中的去免疫化序列

<400> 11
Gln Lys Thr Pro Thr Arg Pro Ala Gly Lys Pro Thr His
1                5                10

<210> 12
<211> 13
<212> PRT
<213> 人工序列

```

<220>

<223> IgA-X 融合物中的去免疫化序列

<400> 12

Gln Lys Thr Ala Thr Arg Pro Ala Gly Lys Pro Thr His
1 5 10

<210> 13

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> HSA-IFN α 连接处中的潜在 T 细胞表位

<400> 13

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu
1 5 10

<210> 14

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> 修饰的清蛋白 C 末端

<400> 14

Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Thr Thr Ala
1 5 10

<210> 15

<211> 30

<212> PRT

<213> 人工序列

<220>

<223> Fc 融合蛋白中的序列

<400> 15

His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly
1 5 10 15

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25 30

<210> 16

<211> 30

<212> PRT

<213> 人工序列

<220>

<223> 修饰的 Fc 融合蛋白中的序列

<400> 16

His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala Thr Pro Gly Lys Gly

1	5	10	15
---	---	----	----

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 20 25 30

<210> 17
 <211> 9
 <212> PRT
 <213> 人工序列

<220>
 <223> 连接处序列

<400> 17
 Leu Ser Leu Ser Pro Gly Lys Ala Pro
 1 5

<210> 18
 <211> 9
 <212> PRT
 <213> 人工序列

<220>
 <223> 修饰的连接处序列

<400> 18
 Ala Thr Ala Thr Pro Gly Ala Ala Pro
 1 5

<210> 19
 <211> 9
 <212> PRT
 <213> 人工序列

<220>
 <223> 修饰的连接处序列

<400> 19
 Leu Asn Leu Ser Pro Gly Ala Ala Pro
 1 5

<210> 20
 <211> 13
 <212> PRT
 <213> 人工序列

<220>
 <223> 含有反应性表位的合成肽

<400> 20
 Lys Ser Leu Ser Leu Ser Pro Gly Lys Ala Pro Thr Ser
 1 5 10

<210> 21
 <211> 13
 <212> PRT

<213> 人工序列

<220>

<223> 含有反应性表位的修饰合成肽

<400> 21

Lys Ser Ala Thr Ala Thr Pro Gly Lys Ala Pro Thr Ser
1 5 10

<210> 22

<211> 39

<212> PRT

<213> 人工序列

<220>

<223> 清蛋白-CD4 连接处序列

<400> 22

Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala
1 5 10 15

Ala Leu Gly Leu Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val
20 25 30

Glu Leu Thr Cys Thr Ala Ser
35

<210> 23

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> HSA-IFN α 融合物中的潜在 T 细胞表位

<400> 23

Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly Leu Cys
1 5 10

<210> 24

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> HSA-IFN α 融合物中的潜在 T 细胞表位

<400> 24

Leu Gly Leu Cys Asp Leu Pro Gln Thr His Ser Leu Gly
1 5 10

<210> 25

<211> 5

<212> PRT

<213> 人工序列

<220>
<223> 清蛋白 C 末端序列

<400> 25
Ala Ala Leu Gly Leu
1 5

<210> 26
<211> 5
<212> PRT
<213> 人工序列

<220>
<223> 突变的清蛋白 C 末端序列

<400> 26
Thr Ala Thr Thr Ala
1 5

<210> 27
<211> 19
<212> PRT
<213> 人工序列

<220>
<223> 修饰的清蛋白连接区

<400> 27
Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Thr Ala
1 5 10 15
Thr Thr Ala

<210> 28
<211> 21
<212> PRT
<213> 人工序列

<220>
<223> 非突变 Fc 序列

<400> 28
Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1 5 10 15
Pro Glu Leu Leu Gly
20

<210> 29
<211> 52
<212> PRT
<213> 人工序列

<220>
<223> GLP-1-突变 Fc 融合连接处

<400> 29
 His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15
 Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu
 20 25 30
 Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
 35 40 45
 Glu Leu Leu Gly
 50

<210> 30
 <211> 41
 <212> PRT
 <213> 人工序列

<220>
 <223> GLP-1-正常 Fc 融合连接处

<400> 30
 Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val
 1 5 10 15
 Lys Gly Arg Gly Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
 20 25 30
 Pro Cys Pro Ala Pro Glu Leu Leu Gly
 35 40

<210> 31
 <211> 13
 <212> PRT
 <213> 人工序列

<220>
 <223> GLP-1-Fc 融合处的潜在 T 细胞表位

<400> 31
 Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu
 1 5 10

<210> 32
 <211> 13
 <212> PRT
 <213> 人工序列

<220>
 <223> GLP-1-Fc 融合连接处的潜在 T 细胞表位

<400> 32
 Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Glu Pro
 1 5 10

<210> 33
 <211> 13
 <212> PRT
 <213> 人工序列

<220>
 <223> GLP-1-Fc 融合连接处的潜在 T 细胞表位

<400> 33
 Ala Trp Leu Val Lys Gly Arg Gly Glu Pro Lys Ser Ser
 1 5 10

<210> 34
 <211> 52
 <212> PRT
 <213> 人工序列

<220>
 <223> 去免疫化的 GLP-1-Fc 融合连接处

<400> 34
 His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15

Gln Ala Ala Lys Glu Phe Ala Ala Trp Ala Val Thr Gly Thr Gly Glu
 20 25 30

Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro
 35 40 45

Glu Leu Leu Gly
 50

<210> 35
 <211> 41
 <212> PRT
 <213> 人工序列

<220>
 <223> 具有糖基化位点的 GLP-1-Fc 融合连接处

<400> 35
 Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val
 1 5 10 15

Lys Gly Arg Asn Gly Ser Lys Ser Ser Asp Lys Thr His Thr Cys Pro
 20 25 30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
 35 40

<210> 36
 <211> 41
 <212> PRT
 <213> 人工序列

<220>

<223> TNF-R-γ-1 融合连接处

<400> 36

Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly
 1 5 10 15

Ser Thr Gly Asp Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
 20 25 30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
 35 40

<210> 37

<211> 41

<212> PRT

<213> 人工序列

<220>

<223> TNF-R-Fc 融合连接处

<400> 37

Ser Thr Ser Phe Leu Leu Pro Met Gly Pro Ser Pro Pro Ala Glu Gly
 1 5 10 15

Ser Thr Gly Asn Gly Ser Lys Ser Cys Asp Lys Thr His Thr Cys Pro
 20 25 30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
 35 40

<210> 38

<211> 40

<212> PRT

<213> 人工序列

<220>

<223> Fc-IL12p35 融合连接处

<400> 38

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
 1 5 10 15

Ser Pro Gly Lys Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met
 20 25 30

Phe Pro Cys Leu His His Ser Gln
 35 40

<210> 39

<211> 40

<212> PRT

<213> 人工序列

<220>

<223> IL-12p40-IL2 融合连接处

<400> 39

Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser
1 5 10 15

Val Pro Cys Ser Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu
20 25 30

Gln Leu Glu His Leu Leu Leu Asp
35 40

<210> 40

<211> 40

<212> PRT

<213> 人工序列

<220>

<223> 修饰的 Fc-IL12p35 融合连接处

<400> 40

Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala
1 5 10 15

Thr Pro Gly Lys Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met
20 25 30

Phe Pro Cys Leu His His Ser Gln
35 40

<210> 41

<211> 40

<212> PRT

<213> 人工序列

<220>

<223> 修饰的 IL12p40-IL2 融合连接处

<400> 41

Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser
1 5 10 15

Val Pro Cys Ser Asn Gly Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu
20 25 30

Gln Leu Glu His Leu Leu Leu Asp
35 40

<210> 42

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> IL12p40-IL 融合物中的潜在 T 细胞表位

<400> 42

Ser Glu Trp Ala Ser Val Pro Cys Ser Asn Gly Thr Ser
1 5 10

```

<210> 43
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> IL12p40-IL2 融合连接处中的潜在 T 细胞表位

<400> 43
Ala Ser Val Pro Cys Ser Asn Gly Thr Ser Ser Thr
1          5          10

<210> 44
<211> 41
<212> PRT
<213> 人工序列

<220>
<223> IL4-Fc 融合连接处

<400> 44
Glu Asn Phe Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser
1          5          10          15

Lys Cys Ser Ser Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
          20          25          30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
          35          40

<210> 45
<211> 40
<212> PRT
<213> 人工序列

<220>
<223> Fc-GMCSF 融合连接处

<400> 45
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu
1          5          10          15

Ser Pro Gly Lys Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp
          20          25          30

Glu His Val Asn Ala Ile Gln Glu
          35          40

<210> 46
<211> 13
<212> PRT
<213> 人工序列

<220>
<223> IL4-Fc 融合连接处的潜在 T-细胞表位

```

```

<400> 46
Glu Lys Tyr Ser Lys Cys Ser Ser Glu Pro Lys Ser Cys
1          5          10

<210> 47
<211> 41
<212> PRT
<213> 人工序列

<220>
<223> 修饰的 IL4-Fc 融合物

<400> 47
Glu Asn Phe Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser
1          5          10          15

Lys Cys Ser Ser Thr Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
          20          25          30

Pro Cys Pro Ala Pro Glu Leu Leu Gly
          35          40

<210> 48
<211> 40
<212> PRT
<213> 人工序列

<220>
<223> 去免疫化的 Fc-GMCSF 融合连接处

<400> 48
Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Ala Thr Ala
1          5          10          15

Thr Pro Gly Lys Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp
          20          25          30

Glu His Val Asn Ala Ile Gln Glu
          35          40

<210> 49
<211> 35
<212> PRT
<213> 人工序列

<220>
<223> IgG2CH1-IgG1 铰链融合连接处

<400> 49
Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val
1          5          10          15

Asp Lys Thr Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
          20          25          30

Pro Cys Pro
          35

```

<210> 50
 <211> 35
 <212> PRT
 <213> 人工序列

<220>
 <223> IgG1 铰链-IgG2CH2 融合连接处

<400> 50
 Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
 1 5 10 15
 Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
 20 25 30

Asp Thr Leu
 35

<210> 51
 <211> 13
 <212> PRT
 <213> 人工序列

<220>
 <223> IgG2CH1-IgG1 铰链融合连接处中的潜在 T 细胞表位

<400> 51
 Thr Lys Val Asp Lys Thr Val Glu Pro Lys Ser Cys Asp
 1 5 10

<210> 52
 <211> 13
 <212> PRT
 <213> 人工序列

<220>
 <223> IgG2CH1-IgG1 铰链融合连接处中的潜在 T 细胞表位

<400> 52
 Lys Thr Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr
 1 5 10

<210> 53
 <211> 35
 <212> PRT
 <213> 人工序列

<220>
 <223> 修饰的 IgG2CH1-IgG1 铰链融合连接处

<400> 53
 Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Ala
 1 5 10 15

Asp Lys Thr Ala Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro
20 25 30

Pro Cys Pro
35

<210> 54
<211> 35
<212> PRT
<213> 人工序列

<220>
<223> 修饰的 IgG2CH1-IgG1 铰链融合连接处

<400> 54
Gln Thr Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val
1 5 10 15

Asp Lys Thr Val Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro
20 25 30

Pro Cys Pro
35

<210> 55
<211> 35
<212> PRT
<213> 人工序列

<220>
<223> 修饰的 IgG1 铰链-IgG2CH2 融合连接处

<400> 55
Glu Pro Lys Ser Ser Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala
1 5 10 15

Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys
20 25 30

Asp Thr Leu
35

<210> 56
<211> 166
<212> PRT
<213> 人工序列

<220>
<223> 突变 EPO 序列

<400> 56
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu Gly
20 25 30

[illegible]

<210>	57
<211>	17
<212>	PRT
<213>	人工序列

<220>
<223> CH3-EPO 融合连接处

```

<400> 57
Thr Gln Lys Ser Ala Thr Ala Thr Pro Gly Ala Ala Pro Pro Arg Leu
1          5          10          15
Ile

```

<210>	58
<211>	8
<212>	PRT
<213>	人工序列

<220>
<223> IgG2 CH3 序列

<400> 58
Lys Ser Leu Ser Leu Ser Pro Gly
1 5

<210>	59
<211>	8
<212>	PRT

<213> 人工序列

<220>

<223> 修饰的 IgG2CH3 序列

<400> 59

Lys Ser Ala Thr Ala Thr Pro Gly
1 5

专利名称(译)	降低融合蛋白的免疫原性		
公开(公告)号	CN100488984C	公开(公告)日	2009-05-20
申请号	CN02807505.6	申请日	2002-03-30
申请(专利权)人(译)	默克专利有限公司		
当前申请(专利权)人(译)	默克专利有限公司		
[标]发明人	SD吉利斯		
发明人	S·D·吉利斯		
IPC分类号	C07K16/46 C07K19/00 A61K39/395 C12N15/09 A61K38/00 A61K38/22 A61K38/24 A61K38/27 A61K45/00 A61K47/00 C07K1/00 C07K14/00 C07K14/505 C07K14/52 C07K14/54 C07K14/55 C07K14/575 C07K14/705 C07K14/725 C07K14/76 C07K16/00 C07K16/18 C07K17/00 C12N15/62 G01N33/53		
CPC分类号	C07K2319/02 C07K14/505 C07K2319/00 C07K14/55 C07K2319/40 C07K14/5434 C07K14/70578 C07K2319/75 C12N15/62 C07K2319/30 C07K16/30 C07K2319/33 C07K1/1077 Y10S424/809 Y10S530/868		
审查员(译)	马岚		
优先权	60/280625 2001-03-30 US		
其他公开文献	CN1531545A		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了制备具有降低的免疫原性的融合蛋白的组合物和方法。本发明的融合蛋白包括具有氨基酸改变的连接区，此氨基酸改变降低了连接处表位和MHC II类结合的能力，因此降低了它和T-细胞受体的相互作用。本发明的方法涉及对融合蛋白连接区中一处或多处氨基酸的分析、改变或修饰，目的是鉴定出T-细胞表位并降低它和T-细胞受体相互作用的能力。本发明的组合物和方法在治疗中 useful。

