



(12)发明专利申请

(10)申请公布号 CN 109073655 A

(43)申请公布日 2018.12.21

(21)申请号 201780019687.5

(74)专利代理机构 中国专利代理(香港)有限公司 72001

(22)申请日 2017.02.03

代理人 李志强 周齐宏

(30)优先权数据

62/291216 2016.02.04 US

(51)Int.Cl.

G01N 33/68(2006.01)

(85)PCT国际申请进入国家阶段日

2018.09.25

(86)PCT国际申请的申请数据

PCT/US2017/016549 2017.02.03

(87)PCT国际申请的公布数据

W02017/136753 EN 2017.08.10

(71)申请人 安口生物公司

地址 美国新泽西州

(72)发明人 X.肖 L.李

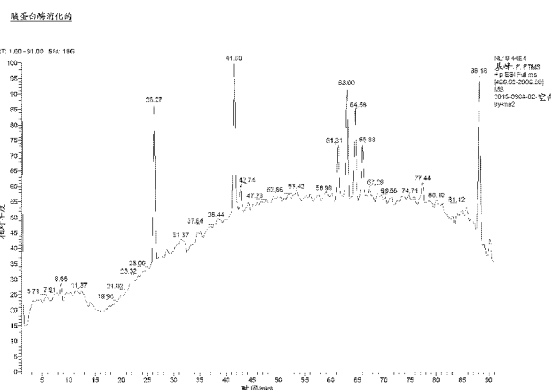
权利要求书1页 说明书13页
序列表68页 附图21页

(54)发明名称

鉴定和分析蛋白的氨基酸序列的方法

(57)摘要

本公开内容提供用于测定受试蛋白相对于靶标生物制品的生物相似性的方法,其中受试蛋白被两种不同的蛋白酶消化,得到的消化肽片段用柱色谱-串联质谱分析,以达到受试蛋白的100%的氨基酸序列覆盖度和100%的氨基酸序列精确性。



1. 一种测定受试蛋白相对于靶标生物制品的生物相似性的方法,所述方法包括以下步骤:

(a) 使用第一蛋白酶消化受试蛋白的第一样本第一孵育时间,和使用第二蛋白酶消化所述受试蛋白的第二样本第二孵育时间,其中所述第一样本和所述第二样本是物理上分开的;

(b) 在足以增强小肽与柱的结合的条件下,将柱色谱和串联质谱应用于所述第一样本,并产生所述第一样本中的所述受试蛋白的序列;

(c) 在足以增强小肽与柱的结合的条件下,将柱色谱和串联质谱应用于所述第二样本,并产生所述第二样本中的所述受试蛋白的所述序列,其中所述第一样本和所述第二样本是物理上分开的;

(d) 当所述受试蛋白包含与所述靶标生物制品的100%序列同一性时,鉴定所述受试蛋白与所述靶标生物制品为生物相似的;和

(e) 当所述受试蛋白不包含与靶标生物制品的100%序列同一性时,鉴定所述受试蛋白与所述靶标生物制品为非生物相似的。

2. 权利要求1的方法,其中所述单克隆抗体包含阿达木单抗。

3. 权利要求1或2的任一项的方法,其中所述第一蛋白酶是胰蛋白酶。

4. 权利要求1-3的任一项的方法,其中所述第二蛋白酶是糜蛋白酶。

5. 权利要求1-4的任一项的方法,其中第一消化时间段是约0.1-约1.0小时。

6. 权利要求5的方法,其中所述第一消化时间段是约0.1-约0.5小时。

7. 权利要求5的方法,其中所述第一消化时间段是约0.6-约1.0小时。

8. 权利要求5的方法,其中所述第一消化时间段是约0.5小时。

9. 权利要求1-8的任一项的方法,其中第二消化时间段是约0.1-约2.0小时。

10. 权利要求9的方法,其中所述第二消化时间段是约0.1-约1.5小时。

11. 权利要求9的方法,其中所述第二消化时间段是约1.5-约2.0小时。

12. 权利要求9的方法,其中所述第二消化时间段是约1.5小时。

鉴定和分析蛋白的氨基酸序列的方法

[0001] 相关申请

本申请要求2016年2月4日提交的美国专利申请号62/291,216的优先权,其内容通过引用以其整体并入本文。

[0002] 序列表的结合

命名为“ONBI-008001WO_SeqList.txt”的文本文件(创建于2017年2月2日且大小为81.7 KB)的内容通过引用以其整体并入本文。

[0003] 公开内容的领域

本公开内容主要涉及:改进的蛋白测序方法,其使用减少的孵育时间用于变性蛋白的蛋白酶消化,并包括在柱色谱和串联质谱(LC-MS/MS)分析期间增加的含水流动相,以提高序列覆盖度和精确性至100%;以及改进的蛋白测序方法,其用于治疗性重组蛋白的开发和经批准的生物制品生产的质量控制分析。

[0004] 背景

包括重组单克隆抗体(mAb)和天然蛋白的重组版本在内的重组蛋白已被用作生物医学研究的试剂,以及用于人的诊断剂和治疗剂。重组蛋白的一个实例包括生物相似分子(也称为“生物制品”)。为了被批准用作人的治疗剂,必须证明生物相似分子具有相同的氨基酸序列,并且在翻译后的修饰方面与亲本创新生物产品是非常密切地相似的,例如,具有“一致性(sameness)”。评估生物相似分子的一致性至关重要,因为重组蛋白本质上是复杂的。重组蛋白是利用遗传修饰的活生物体(例如细菌、酵母、动物或人细胞系)改造的。活生物体产生作为通过复杂机制折叠的长链氨基酸和/或修饰氨基酸的重组蛋白。因此,重组蛋白显示出高分子复杂性且对生产过程中的变化是高度敏感的。

[0005] 重组蛋白的特异性和效应子功能高度依赖于氨基酸序列和是否存在特定的修饰。因此,DNA测序通常用于初步表征生物制品,例如单克隆抗体。然而,重组蛋白(例如单克隆抗体)的蛋白-水平重排例如随后的突变和翻译后修饰(PTM)通过蛋白水平的分析而识别,因为这样的重排只能通过蛋白水平分析来揭示。因此,当抗体的cDNA或原始细胞系不可用时,或当氨基酸序列的表征是必要的,以验证经批准用作治疗剂的重组抗体的相似性,以及用于生产过程的质量控制时,需要单克隆抗体的氨基酸测序。

[0006] 尽管蛋白中氨基酸的序列鉴定的重要性,但是目前还没有开发出提供高水平(100%)的序列精确性和覆盖度的未知蛋白测序方法。尤其测序重组蛋白仍是一个挑战。两种通用方法用于使用质谱对蛋白进行测序。在第一种方法中,完整的蛋白被电离,然后被导入质量分析仪用于质量测量和串联质谱(MS/MS)分析。这种方法被称为“自上而下的”蛋白组学。在第二种方法中,使用蛋白酶例如胰蛋白酶,将蛋白经酶促消化成较小的肽。随后,所述肽被引入质谱仪并通过肽质量指纹图法或串联质谱(MS/MS)鉴定。这后一种方法被称为“自下而上的”蛋白组学,并利用肽水平的鉴定来推断蛋白的存在情况。自下而上的蛋白组学是一种鉴定蛋白和表征其氨基酸序列以及PTM的优选方法。

[0007] 自下而上的蛋白组学的一种熟知方法是Edman降解。在这种方法中,氨基端残基被标记并从肽上切下,而不破坏其它氨基酸残基之间的肽键。因为Edman降解是从蛋白的N-端

开始,所以,如果N-端氨基酸经过了化学修饰或如果它被隐藏在天然蛋白体内,则是不可靠的。为了可辨别的结果,还需要猜测或单独的程序来确定二硫桥的位置,以及需要1皮摩尔或更高的肽浓度。因此,Edman方法不适合对长度超过50个氨基酸的蛋白或具有PTM的蛋白测序。

[0008] 基于质谱的方法通过组装(assembly)从蛋白的多重蛋白水解消化产生的重叠肽的串联质谱(MS/MS)来表征蛋白。每个串联质谱(MS/MS)仅覆盖靶标蛋白的短肽。因此,高覆盖度蛋白测序的关键是找到来自重叠肽的波谱对,以组装串联质谱法(MS/MS)波谱为长的波谱。然而,肽的重叠区域可能太短,以致无法被可信地鉴定。此外,依靠解释单个串联质谱法(MS/MS)波谱的自动化从头测序方法受到限制,因为这些方法通常不能在不会于平均5个氨基酸中错误鉴定1个氨基酸的情况下重建长的(8+氨基酸)序列。从头肽测序的进展提高了测序精确性至超过95%,但是是在有限的序列覆盖度下,例如仅55%的序列覆盖度。当前所有的按波谱从头测序策略都面临着在测序精确性与覆盖度之间的权衡,因为表现出完全的肽片段化的波谱很少覆盖整个靶标蛋白,虽然这些策略是精确地重建全长肽序列所需要的。

[0009] 分别测序单个波谱的替代方法是使用另一种称为鸟枪蛋白测序(Shotgun Protein Sequencing, SPS)的方法同时解释来自重叠肽的多重MS/MS谱。已经发现SPS产生常常覆盖90-95%的靶标蛋白序列的序列,同时在高分辨率的MS/MS波谱上,每20个氨基酸中只有一个被错误鉴定。SPS具有局限性。它产生不单独地覆盖靶标蛋白序列的大区域(更不必说完整蛋白)的片段化序列。SPS序列具有10-15个氨基酸的平均长度,且最长的回收的SPS从头序列长度小于45个氨基酸。

[0010] 为了批准用于人或动物的治疗用途,基于通过临床研究、动物研究和分析研究汇编的数据以及构象状态,生物相似物必须显示为与亲本创新生物产品近乎相同,例如具有“一致性”。自上而下或自下而上的反相色谱方法都不能提供用于测定重组蛋白的生物相似性的可靠且简单的基础(例如100%序列精确性和覆盖度)。

[0011] 因此,目前需要一种测定重组蛋白(例如单克隆抗体)与亲本创新生物产品相比的分析相似性或“一致性”的方法,其中该方法使用显著缩短的用于重组蛋白的蛋白酶消化的时间框(当与广为使用的蛋白酶消化方案相比时),以及增强的肽暴露条件和因此增加的肽对色谱柱的附着,以高置信度精确地分析达100%的氨基酸序列覆盖度。所述方法可用于开发经批准的生物相似物,以及在经批准的生物相似物的生产期间的质量控制分析。

[0012] 概述

本公开内容提供用于评估、选择和/或生产生物制品的方法,生物制品包括例如生物相似物,包括与其相关的可交换组合物(例如药用制剂)。例如,本公开内容提供下述的方法:通过特有的特征(例如氨基酸序列)定义靶标蛋白(例如,在生物制品许可申请(biologics license application, BLA)下批准的亲本创新生物产品),并且这样的特征用于评估、鉴定和/或生产与用于诊断或经批准用作治疗剂的靶标蛋白具有所需“一致性”的生物制品。所公开的方法还用于例如监测产品变化和控制可由于重组技术与活细胞的应用在生产生物制品期间发生的产品偏离(drift)。所述方法包括用于以高可靠性在达生物制品的100%氨基酸序列的覆盖度和精确性下评估受试蛋白与靶标蛋白的相似性的步骤。例如,可以对受试蛋白进行评估,以确定其是否具有与靶标蛋白的预定水平的相似性或“一致性”,所述

靶标蛋白是商业上可获得的和/或经批准用于人或动物的治疗用途。在以下条件的一个或多个或全部存在时,这是特别有益的:(1) 受试蛋白通过与靶标蛋白不同的方法制备,或用来制备靶标蛋白的方法是不为受试蛋白的制备者所知的;(2) 受试蛋白是由与制备靶标蛋白的实体具有不同的上市许可(或者根本无许可)的实体制备的;或(3) 受试蛋白在依赖于或涉及与靶标蛋白的许可相关的临床信息的方法中被批准。

[0013] 本公开内容提供一种测定受试蛋白相对于靶标生物制品的生物相似性的方法,该方法包括以下步骤:(a) 使用第一蛋白酶消化受试蛋白的第一样本第一孵育时间,和使用第二蛋白酶消化受试蛋白的第二样本第二孵育时间,其中第一样本和第二样本是物理上分开的;(b) 在足以增强小肽与柱的结合的条件下,将柱色谱和串联质谱应用于第一样本,并产生第一样本中的受试蛋白的序列;(c) 在足以增强小肽与柱的结合的条件下,将柱色谱和串联质谱应用于第二样本,并产生第二样本中的受试蛋白的序列,其中第一样本和第二样本是物理上分开的;(d) 当受试蛋白包含与靶标生物制品的100%序列同一性时,鉴定受试蛋白与靶标生物制品为生物相似的;和(e) 当受试蛋白不包含与靶标生物制品的100%序列同一性时,鉴定受试蛋白与靶标生物制品为非生物相似的。

[0014] 在本公开内容的测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,单克隆抗体包含阿达木单抗。

[0015] 在本公开内容的测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第一蛋白酶是胰蛋白酶。备选地或另外,在本公开内容的测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第二蛋白酶是糜蛋白酶。

[0016] 在本公开内容的测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第一消化时间段是约0.1-约1.0小时。在某些实施方案中,第一消化时间段是约0.1-约0.5小时。在某些实施方案中,第一消化时间段是约0.6-约1.0小时。在某些实施方案中,第一消化时间段是约0.5小时。

[0017] 在本公开内容的测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第二消化时间段是约0.1-约2.0小时。在某些实施方案中,第二消化时间段是约0.1-约1.5小时。在某些实施方案中,第二消化时间段是约1.5-约2.0小时。在某些实施方案中,第二消化时间段是约1.5小时。本公开内容提供一种测定受试蛋白相对于靶标生物制品的生物相似性的方法,该方法包括以下步骤:使用第一蛋白酶消化受试蛋白的第一样本第一孵育时间,和使用第二蛋白酶消化受试蛋白的第二样本第二孵育时间,其中受试蛋白分别在第一样本和第二样本中被消化成肽序列;和使用柱色谱与第二样本的肽序列分开地分析第一样本的肽序列,以测定受试蛋白的100%的氨基酸序列覆盖度和100%的氨基酸序列精确性,其中柱色谱包括增强小肽与柱的结合的条件。

[0018] 在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,受试蛋白是蛋白、糖蛋白、融合蛋白、生长因子、疫苗、血液因子、溶栓剂、造血蛋白、激素、干扰素、基于白介素的产品、抗体、单特异性(例如单克隆)抗体、聚乙二醇化抗体、抗体药物缀合物、治疗性酶、细胞因子或可溶性受体片段中的一种。

[0019] 在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第一蛋白酶是胰蛋白酶。备选地或另外,在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第二蛋白酶是糜蛋白酶。

[0020] 在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第一蛋白酶是胰蛋白酶。在包括第一蛋白酶是胰蛋白酶的那些实施方案在内的某些实施方案中,第一消化时间段是约0.1-约1.0小时。在包括第一蛋白酶是胰蛋白酶的那些实施方案在内的某些实施方案中,第一消化时间段是约0.1-约0.5小时。在包括第一蛋白酶是胰蛋白酶的那些实施方案在内的某些实施方案中,第一消化时间段是约0.6-约1.0小时。在包括第一蛋白酶是胰蛋白酶的那些实施方案在内的某些实施方案中,第一消化时间段是约0.5小时。

[0021] 在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,第二蛋白酶是糜蛋白酶。在包括第二蛋白酶是糜蛋白酶的那些实施方案在内的某些实施方案中,第二消化时间段是约0.1-约2.0小时。在包括第二蛋白酶是糜蛋白酶的那些实施方案在内的某些实施方案中,第二消化时间段是约0.1-约1.5小时。在包括第二蛋白酶是糜蛋白酶的那些实施方案在内的某些实施方案中,第二消化时间段是约1.5-约2.0小时。在包括第二蛋白酶是糜蛋白酶的那些实施方案在内的某些实施方案中,第二消化时间段是约1.5小时。

[0022] 在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,靶标生物制品是商业可获得的或经批准的用于人或动物的治疗用途的生物制品、对于二级批准程序(secondary approval process)列出的参比药物、蛋白、糖蛋白、融合蛋白、生长因子、疫苗、血液因子、溶栓剂、造血蛋白、激素、干扰素、基于白介素的产品、抗体、单特异性(例如单克隆)抗体、聚乙二醇化抗体、抗体药物缀合物、治疗性酶、细胞因子或可溶性受体片段。在某些实施方案中,靶标生物制品是以下中的一种:阿达木单抗(Humira®)、贝伐单抗(Avastin®)、地舒单抗(Xgeva®)、西妥昔单抗(Erbix®);利妥昔单抗(Rituxan®); Mabthera®; Campath®; Herceptin®; Xolair®; Prolia®; Vectibix®; ReoPro®; Zenapax®; Simulect®; Synagis®; Remicade®; Mylotarg®; Campath®; Raptiva®; Zevalin®; Erbitux®; Tysabri®; Lucentis®; Soliris®; Cimzia®; Ilaris®; Arzerra®; Bexxar®; Simponi®; Actemra®; Benlysta®; Adcetris®; 或 Yervoy®。在测定受试蛋白相对于靶标生物制品的生物相似性的方法的某些实施方案中,靶标生物制品是阿达木单抗(Humira®)。

[0023] 本公开内容提供一种分析重组单克隆抗体相对于阿达木单抗或其生物等效物的生物相似性的方法,该方法包括以下步骤:通过用第一蛋白酶消化重组单克隆抗体的第一样本和分开地用第二蛋白酶消化重组单克隆抗体的第二样本,测定达100%的重组单克隆抗体的氨基酸序列,其中蛋白酶消化步骤包括总共不超过2小时的孵育时间;并比较重组单克隆抗体的氨基酸序列与阿达木单抗或其生物等效物的氨基酸序列以测定一致性。

[0024] 在分析重组单克隆抗体相对于阿达木单抗或其生物等效物的生物相似性的方法的某些实施方案中,一致性包括在重组单克隆抗体的氨基酸序列和阿达木单抗或其生物等效物的氨基酸序列之间的100%相似性。

[0025] 本公开内容提供一种生产包含重组单克隆抗体的药用产品的方法,该方法包括以下步骤:提供重组单克隆抗体,其中重组单克隆抗体在BLA或补充BLA下未被批准;获得重组单克隆抗体的输入值,其中一个或多个输入值是靶标生物制品的氨基酸序列;获得通过比较输入值与靶标生物制品的多个氨基酸序列进行的多个评价,其中靶标生物制品在生物制品许可申请(BLA)或补充BLA下被批准;和如果输入值与靶标生物制品的所述氨基酸序列的靶标值是不能区分的,则将重组单克隆抗体加工成药用产品。

[0026] 在生产包含重组单克隆抗体的药用产品的方法的某些实施方案中,重组单克隆抗体经改造为与以下中的一种是生物相似的:阿达木单抗(Humira®)、贝伐单抗(Avastin®)、地舒单抗(Xgeva®)、西妥昔单抗(Erbitux®);Rituxan®;Mabthera®;Campath®;Herceptin®;Xolair®;Prolia®;Vectibix®;ReoPro®;Zenapax®;Simulect®;Synagis®;Remicade®;Mylotarg®;Campath®;Raptiva®;Zevalin®;Erbitux®;Tysabri®;Lucentis®;Soliris®;Cimzia®;Ilaris®;Arzerra®;Bexxar®;Simponi®;Actemra®;Benlysta®;Adcetris®;或Yervoy®。

[0027] 在生产包含重组单克隆抗体的药用产品的方法的某些实施方案中,重组单克隆抗体经改造为与阿达木单抗(Humira®)是生物相似的。

[0028] 在生产包含重组单克隆抗体的药用产品的方法的某些实施方案中,输入值包含重组单克隆抗体的氨基酸序列的100%覆盖度。

[0029] 本公开内容提供一种分析达100%的重组单克隆抗体的氨基酸序列以测定与药用产品的一致性的方法,该方法包括以下步骤:通过使用第一蛋白酶消化变性重组单克隆抗体的第一样本第一孵育时间,和使用第二蛋白酶消化变性重组单克隆抗体的第二样本第二孵育时间,使变性重组单克隆抗体片段化为离散肽,其中第一孵育时间是约0.1-约1.0小时,此后猝灭第一蛋白酶,且其中第二孵育时间是约1.0-约2.0小时,此后猝灭第二蛋白酶;分析重组单克隆抗体的离散肽以测定形成重组单克隆抗体的氨基酸序列;并比较重组单克隆抗体的氨基酸序列与药用产品的氨基酸序列,其中药用产品在生物制品许可申请(BLA)或补充BLA下被批准。

[0030] 还提供用于产生或评估预定的多个靶标值以产生或评估受试蛋白的特征(例如氨基酸序列)和/或使用或应用这样的信息以获得描述在受试蛋白和靶标蛋白之间的关系(例如结构关系)的一致性/同一性值的方法。在一些实例中,一致性/同一性值可用于评估、鉴定和/或制备(例如生产)受试蛋白。在一些实例中,一致性/同一性值是释放受试蛋白的规格。因此,在此公开可用于评估、鉴定和生产批准的生物制品的方法。

[0031] 所述方法任选地包括在高度纯化制剂(例如受试蛋白制剂)中将受试生物制品制剂与受试生物制品的其它同种型或变体以及来自受试生物制品生产的副产物分开的制备步骤,其中受试生物制品在生物制品许可申请(BLA)、补充BLA或其等同物下未被批准;然后使用靶标生物制品的一个或多个氨基酸序列的输入值处理高度纯化的受试生物制品制剂。

[0032] 在一个实施方案中,受试蛋白被测定为具有与靶标蛋白氨基酸序列相同或几乎相同(例如100%匹配,其中由于翻译错误具有0.5%的序列变化容许量)的氨基酸序列(例如一级氨基酸序列),且靶标蛋白在BLA、补充BLA或其等同物下被批准。

[0033] 在一个实施方案中,所述方法包括以下步骤:(1) 制备富集的受试蛋白制剂,其中受试蛋白可在生物制品许可申请(BLA)、补充BLA或其等同物下批准或可未被批准;和(2) 处理受试蛋白制剂以测定氨基酸序列与靶标蛋白的氨基酸序列是不能区别的,其中受试蛋白具有与靶标蛋白氨基酸序列具有达100%同一性的氨基酸序列(例如一级氨基酸序列),且靶标蛋白在BLA、补充BLA或其等同物下被批准,从而生产包含蛋白例如单克隆抗体(mAb)的药用产品。

[0034] 在一个实施方案中,靶标蛋白是抗体,例如单克隆抗体、人源化抗体或人抗体。在备选的实施方案中,靶标蛋白可以是与聚乙二醇(PEG)聚合物链缀合的抗体,例如聚乙二醇

化抗体。对于聚乙二醇化单克隆抗体,取决于聚乙二醇化程度和聚乙二醇化的质量大小范围,本公开内容的方法可用于在样本准备以进行肽作图(peptide mapping)之前释放PEG的步骤后对单克隆抗体的氨基酸进行测序。在进一步的实施方案中,靶标蛋白可以是抗体-药物缀合物(ADC)复合物分子,其包含经由稳定的具有不稳定键的化学接头与生物活性细胞毒性(抗癌)有效载荷或药物连接的抗体,例如全mAb或抗体片段,例如单链可变片段(scFv)。在带有药物的反应性残基被修饰的ADC复合物分子中,本公开内容的方法可用于通过在序列数据库中包括药物分子量作为修饰以对药物缀合位点进行作图。

[0035] 例如,靶标蛋白可选自作为以下商标上市的产品以及其它生物制品:阿达木单抗(Humira®)、贝伐单抗(Avastin®)、地舒单抗(Xgeva®)、西妥昔单抗(Erbix®); Rituxan®; Mabthera®; Campath®; Herceptin®; Xolair®; Prolia®; Vectibix®; ReoPro®; Zenapax®; Simulect®; Synagis®; Remicade®; Mylotarg®; Campath®; Raptiva®; Zevalin®; Erbix®; Tysabri®; Lucentis®; Soliris®; Cimzia®; Ilaris®; Arzerra®; Bexxar®; Simponi®; Actemra®; Actemra®; Benlysta®; Adcetris®;或Yervoy®。

[0036] 附图简述

当根据以下结合附图进行的示例性实施方案的描述来考虑本公开内容时,本公开内容关于其方法和用途的另外的方面、特征和优点将容易理解并变得更加显而易见,附图中:

图1A-B是一对分别说明针对特异性运行的胰蛋白酶-消化的和糜蛋白酶-消化的色谱基质的色谱概况的图。基质显示在Sequence Discoverer上没有命中任何靶标氨基酸序列。基质色谱图上的小峰显示系统峰和酶峰。

[0037] 图2A-D是分别说明ONS-3010参比标准品(阿达木单抗)的胰蛋白酶-消化的重链(图2A (SEQ ID NO: 1,具有可能的修饰(SEQ ID NO: 2))、糜蛋白酶-消化的重链(图2B (SEQ ID NO: 3,具有可能的修饰(SEQ ID NO: 4))、胰蛋白酶-消化的轻链(图2C (SEQ ID NO: 5,具有可能的修饰(SEQ ID NO: 6)) and 糜蛋白酶-消化的轻链(图2D (SEQ ID NO: 7,具有可能的修饰(SEQ ID NO: 8))的序列覆盖度的一系列比对。这些图证明本公开内容的方法能够实现100%序列覆盖度。

[0038] 图3A-B是一对分别说明针对特异性运行的胰蛋白酶-消化的和糜蛋白酶-消化的ONS-3010参比标准品的色谱概况的图。

[0039] 图4A-D是分别说明阳性对照阿达木单抗(Humira®) (样本试验ID H35)的胰蛋白酶-消化的重链(图4A (SEQ ID NO: 9,具有可能的修饰(SEQ ID NO: 10))、糜蛋白酶-消化的重链(图4B (SEQ ID NO: 11,具有可能的修饰(SEQ ID NO: 12))、胰蛋白酶-消化的轻链(图4C (SEQ ID NO: 13,具有可能的修饰(SEQ ID NO: 14)) and 糜蛋白酶-消化的轻链(图4D (SEQ ID NO: 15,具有可能的修饰(SEQ ID NO: 16))的100%序列覆盖度的一系列比对。这些图证实靶标序列是阿达木单抗(Humira®)的精确氨基酸序列。

[0040] 图5A-B是一对分别说明胰蛋白酶-消化的和糜蛋白酶-消化的阳性对照阿达木单抗(Humira®) (样本试验ID H35)的色谱概况的图。

[0041] 图6A-D是分别说明阴性对照利妥昔单抗(Rituxan®) (样本试验ID M6)的胰蛋白酶-消化的重链(图6A (SEQ ID NO: 17,具有可能的修饰(SEQ ID NO: 18))、糜蛋白酶-消化的重链(图6B (SEQ ID NO: 19,具有可能的修饰(SEQ ID NO: 20))、胰蛋白酶-消化的轻

链(图6C (SEQ ID NO: 21,具有可能的修饰(SEQ ID NO: 22))和糜蛋白酶-消化的轻链(图6D (SEQ ID NO: 23,具有可能的修饰(SEQ ID NO: 24))的序列覆盖度的一系列比对。这证明本公开内容的方法能够精确地鉴定序列。

[0042] 图7A-B是一对分别说明胰蛋白酶-消化的和糜蛋白酶-消化的阴性对照利妥昔单抗(Rituxan®) (样本试验ID M6)的色谱概况的图。

[0043] 图8是说明阿达木单抗(Humira®) ((SEQ ID NO: 26)的轻链((SEQ ID NO: 25)和重链的理论氨基酸序列的示意图。

[0044] 详细描述

如本文所用的,术语“生物制品”(单数或复数)指肽和蛋白产品。例如,生物制品包括在细胞中表达的天然衍生的产品或重组产品,例如蛋白、糖蛋白、融合蛋白、生长因子、疫苗、血液因子、溶栓剂、激素、干扰素、基于白介素的产品、单特异性(例如单克隆)抗体、治疗性酶。生物制品可依据美国公共卫生服务(Public Health Service, PHS)法案351(a)节在生物制品许可申请(BLA)下被批准,而涉及BLA作为参比产品的生物相似的和可互换的生物制品依据PHS法案351(k)节被发许可。美国公共卫生服务(PHS)法案351节被编纂为42 U.S.C. 262。其它生物制品可依据美国联邦食品和化妆品法案(Federal Food and Cosmetic Act)的505(b)(1)节被批准,或依据Hatch Waxman法案(Hatch Waxman Act)的505(b)(2)和505(j)节作为缩略申请被批准,其中505节被编纂为21 U.S.C. 355。

[0045] 如本文所用的,术语“同种型”(单数或复数)指源于单核苷酸多态性、mRNA的差别剪接或翻译后修饰(例如硫酸化、糖基化等)的相同蛋白的几种不同形式的任何一种。

[0046] 如本文所用的,术语“抗体”(单数或复数)在最广义上指任何的IgG、IgM、IgD、IgA和IgE类的单克隆抗体(包括全长单克隆抗体),以及表现所需生物活性的抗体片段。短语“抗体片段”指全长抗体的部分,一般为其抗原结合区或可变区。抗体片段的实例包括:Fab、Fab'、F(ab')₂和Fv片段;双特异性抗体(diabody);线性抗体;单链抗体分子;和从抗体片段形成的多-特异性抗体。

[0047] 如本文所用的,术语“单克隆抗体”(单数或复数)指高特异性的抗体,其针对单个抗原表位。备选地,术语“单克隆抗体”指由单一脾细胞克隆产生的抗体。在一个非限制性实例中,单克隆抗体可以是完全人源化抗体,即其可变区和恒定区均源自人来源。

[0048] 如本文所用的,术语“批准”指监管实体(例如USFDA)批准用于人或动物的治疗用途或诊断用途的候选物的程序。如本文所用的,初级批准程序(primary approval process)是指不涉及先前批准的蛋白的批准程序,例如,它不需要被批准的蛋白与先前批准的蛋白(例如,具有相同一级氨基酸序列或某一级氨基酸序列的先前批准的蛋白)具有结构相似性或功能相似性。在实施方案中,初级批准程序是申请人不依赖来自先前批准的产品数据(例如临床数据)进行批准的过程。示例性初级批准程序包括:在美国,生物制品许可申请(BLA)或补充生物制品许可申请(sBLA)、依据联邦食品和化妆品法案的505(b)(1)节的新药申请(NDA);以及在欧洲,依据欧洲指令(European Directive) 2001/83/EC的8(3)条的规定的批准,或其它国家或司法管辖区的类似程序。

[0049] 如本文所用的,术语“糖蛋白”指包括一个或多个共价连接于其上的寡糖链(例如聚糖)的氨基酸序列。示例性氨基酸序列包括肽、多肽和蛋白。示例性糖蛋白包括糖基化抗体和抗体-样分子(例如Fc融合蛋白)。示例性抗体包括单克隆抗体和/或其片段、多克隆抗

体和/或其片段以及含Fc域的融合蛋白(例如,含有IgG1的Fc区的融合蛋白或其糖基化部分)。糖蛋白制剂是包含至少一种糖蛋白的组合物或混合物。

[0050] 如本文所用的,短语“靶标生物制品”(例如靶标蛋白)指商业可获得的或批准的生物制品,其定义或提供测量或评估受试生物制品的基础。在实施方案中,靶标生物制品是商业可获得的,用于人或动物的治疗用途。在其它实施方案中,靶标生物制品被初级批准程序批准用于人或动物中的用途。在进一步的实施方案中,靶标生物制品是对于二级批准程序列出的参比药物。示例性靶标蛋白是抗体,例如人源化抗体或人抗体。其它靶标蛋白包括糖蛋白、细胞因子、造血蛋白、可溶性受体片段和生长因子。

[0051] 如本文所用的,术语“评估”指审查、考虑、测定、评价、测量和/或检测受试蛋白和/或靶标生物制品中一个或多个参数的存在、不存在、水平和/或比例,以提供与一个或多个参数有关的信息。在一些实例中,评估糖蛋白制剂包括检测受试蛋白和靶标生物制品之间的一个或多个相似点的存在、不存在、水平或比例。

[0052] 如本文所用的,术语“分析”指进行涉及样本或另一种物质(例如起始原料)中的物理变化的过程。示例性变化包括:从两种或更多的起始原料制备物理实体,对物质进行剪切或片段化,分离或纯化物质,将两个或更多个分开的实体混合成混合物,或进行包括破坏或形成共价键或非共价键的化学反应。分析样本可包括:进行包括物质(例如样本、分析物或试剂)的物理变化的分析过程(有时在此称为“物理分析”);进行分析方法,例如包括一个或多个以下步骤的方法:从另一种物质分离或纯化物质,例如分析物或其片段或其它衍生物;混合分析物或其片段或其它衍生物与另一种物质,例如缓冲剂、溶剂或试剂;或改变分析物或其片段或其它衍生物的结构,例如通过断裂或形成在分析物的第一和第二原子之间的共价键或非共价键,或通过改变试剂或其片段或其它衍生物的结构,例如通过断裂或形成在试剂的第一和第二原子之间的共价键或非共价键。

[0053] 如本文所用的,短语“输入值”指与受试生物制品的参数有关的值。值可以是定性的,例如存在、不存在、中间状态,或值可以是定性的,例如,它可以是参数的数值,如单个数字或范围。

[0054] 本公开内容的通用方法

本公开内容的方法可用于在开发和生产生物相似的治疗性分子的整个过程中分析性测定重组蛋白(例如受试蛋白)与亲本创新生物制品产品(例如靶标蛋白)的相似性。

[0055] 该方法的非限制性应用包括在测定重组蛋白与生物制品产品的相似性中的用途,所述生物制品产品包括但不限于:阿达木单抗(Humira®)、贝伐单抗(Avastin®)、地舒单抗(Xgeva®)、西妥昔单抗(Erbitux®);利妥昔单抗(Rituxan®);Mabthera®;Campath®;Herceptin®;Xolair®;Prolia®;Vectibix®;ReoPro®;Zenapax®;Simulect®;Synagis®;Remicade®;Mylotarg®;Campath®;Raptiva®;Zevalin®;Erbitux®;Tysabri®;Lucentis®;Soliris®;Cimzia®;Ilaris®;Arzerra®;Bexxar®;Simponi®;Actemra®;Benlysta®;Adcetris®;和Yervoy®,以及其它生物制品。

[0056] 所述方法提供分析以评估受试蛋白的一级结构(用于分析受试蛋白或者靶标蛋白)和/或分析与靶标蛋白相比的受试蛋白。这种方法提供达100%的氨基酸序列覆盖度和精确性。

在一个实施方案中,可分析受试蛋白,例如可包括变体的重组蛋白。在一个非限制性备

选实施方案中,受试蛋白可任选地通过柱色谱(例如HPLC)初步纯化,以将生物制品与生物制品的碱性和酸性变体和/或生产生物制品的任何其它副产物(例如酶、细胞和细胞碎片等)分开。生物制品、其变体和相关的生产副产物可通过能够高效、高分辨率地分离相近地洗脱的蛋白的色谱系统(例如阳离子交换柱)处理。

[0057] 在一个进一步的非限制性实例中,本公开内容提供用于鉴定和证实受试蛋白 - 一种单克隆抗体(例如ONS-3010,单克隆抗体阿达木单抗(Humira®)的一种生物相似物) - 的一级结构以表征的方法。

[0058] 根据本公开内容的通用方法,将受试蛋白和/或靶标蛋白(例如单克隆抗体(mAb))变性、还原、烷基化并通过10 kDa离心过滤器离心。这包括基于受试蛋白的溶解、受试蛋白的变性和二硫键还原的最佳消化和完全的序列覆盖。通过使用胰蛋白酶(Try)和糜蛋白酶(Chy),选择性地片段化离散肽。将该肽混合物注入到反相超高效液相色谱(RP-UPLC)系统中,以获得独特的概况(肽图)。肽的精确质荷比(m/z)通过高分辨率质谱仪上进行全扫描来测定。然后将肽断裂成离子片段用于MS/MS氨基酸序列分析。可通过使用例如Proteome Discoverer软件针对阿达木单抗氨基酸序列数据库分析MS/MS数据,以鉴定肽序列。报告在参比产品或靶标产品和受试产品之间的氨基酸序列的相似性。

[0059] 关于它鉴定序列的应用,Proteome Discoverer软件从“.raw"文件提取相关的MS/MS波谱,并测定前体电荷状态和片段化波谱的品质。SEQUEST搜索算法通过比较使实验MS/MS波谱与来自蛋白数据库的理论MS/MS波谱相关。Proteome Discoverer使用基于概率的评分系统来评定由SEQUEST算法发现的最佳匹配的相关性。算法颜色编码氨基酸表以显示鉴定的对应肽序列部分。绿色、黄色和粉红色分别表示高、中、低置信度。没有颜色意味着没有命中肽。Protein Results View突出显示了肽MS/MS波谱中匹配预测的片段质量的碎片离子。具体地,在本公开内容的图2A-2D、4A-4D和6A-6D中,高、中、低置信度命中分别以实线(____)、短划线(-----)和点线(……)代替颜色来标示。

[0060] 在一个具体实例中,通过将水转移到1.5 mL聚丙烯离心管中并将300 µg样本加入到管中,制备浓度≥ 3.0 mg/mL的单克隆抗体(靶标蛋白,例如阿达木单抗和/或生物等效物)的两个分开的等分试样。阴性对照(例如,配制缓冲液或HPLC-级水)和阳性对照(例如,参比标准品)也作为参比物制备。用作仪器系统适宜性对照的肽标准混合物是通过将2.5 mL流动相A(见下文)加入到标准混合物(Sigma, Cat # H2016-1VL)的一个小瓶中制备的20 µg/mL HPLC肽标准混合物。

[0061] 通过将500 µL 8N盐酸胍(Fisher, Cat # 24115)、40 µL 2.5 M Tris碱(3.03 g Tris碱,在HPLC水中至最终体积10 mL)和20 µL 1N HCl (Fisher Scientific, Cat # SA48-1或等同物)的混合物加入到各管中,使各等分试样的单克隆抗体变性。

[0062] 在促进变性蛋白的二硫键断裂的条件下,将稳定剂例如二硫苏糖醇(DTT)加入到变性蛋白的各个样本中。然后,通过将20 µL 25 mg/mL二硫苏糖醇(DTT) (Bio-Rad, Cat # 161-0611) (例如,25 mg DTT,在HPLC级水(Fisher Scientific, Cat # W5-4)中至最终体积1.0 mL)加入到各样本中,还原变性单克隆抗体的各样本。将样本在约 $37 \pm 2^\circ\text{C}$ 分开地孵育约0.5小时。

[0063] 在孵育结束时,通过将8 µL 200 mg/mL碘醋酸钠(Sigma, Cat # I2512) (例如,与HPLC水混合至最终体积1 mL的200 mg碘醋酸钠)加入到各样本中,使各样本经历烷基化,

然后将样本在环境温度下在黑暗中孵育约15分钟。

[0064] 在烷基化孵育结束时,对各样本进行脱盐。脱盐过程包括在Millipore Biomax-10 kDa Ultrafree 0.5 Centrifuge过滤器中洗涤各样本。首先通过以10000 rpm离心约300 μ L碳酸氢铵5分钟来湿润过滤器。将各样本转移至预湿过滤器的表面,然后用300 μ L 0.1 M碳酸氢铵洗涤,并以10000 rpm离心10分钟。洗涤步骤可重复达2次或更多次。最终洗涤包括以10000 rpm离心约10-13分钟,使得各样本的最终体积为约100 μ L。

[0065] 然后,通过在包括缩短的消化时间框(例如对于胰蛋白酶达0.5小时和对于糜蛋白酶达1.5小时)的最适孵育条件下,将不同的蛋白酶(例如胰蛋白酶和糜蛋白酶)加入到样本中,酶促消化脱盐的样本。缩短的孵育时间在持续时间上比传统的孵育时间(例如2-4小时和甚至达18小时)短。较短的消化时间段提供了更多情形的氨基酸的特定错误裂解,使得通过消化产生的糖蛋白包括较长的肽,而不是由较长的消化时间产生的较短的肽。消化单独发生于两个脱盐样本中,即,一个样本使用在经过第一孵育时间段例如0.5小时之后猝灭的胰蛋白酶消化,然后第二样本使用在经过第二孵育时间例如1.5小时之后猝灭的糜蛋白酶消化。糜蛋白酶的蛋白酶消化的这种用途支持轻链中小肽EAK和重链中小肽SLR的覆盖度达到100%序列覆盖度。在消化期间,随着酶裂解连接变性蛋白的氨基酸残基的肽键,变性蛋白的多肽链被切割成较短片段。

[0066] 第一样本经历用胰蛋白酶的蛋白水解消化。例如,可将胰蛋白酶(序列级修饰的, Promega, Cat # V5111, 20 μ g)在20 μ L重构缓冲液中重构。将15 μ L重构胰蛋白酶加入到各样本中,以达到比例为约1:20的胰蛋白酶-与-样本比例。将样本与加入的胰蛋白酶于37 $^{\circ}$ C孵育0.5小时,然后将约5 μ L 10%甲酸(v/v)(Thermo, 产品编号28905或等同物,例如,在900 μ L HPLC级水中的100 μ L甲酸)加入到各样本中,以猝灭酶促消化,为第二消化做准备。

[0067] 第二样本经历用糜蛋白酶的蛋白水解消化。例如,可将糜蛋白酶(Promega糜蛋白酶测序级, 25 μ g, Cat # V1062)在20 μ L 1 mM HCl的HPLC-级水溶液(Fisher Scientific, Cat # SA48-1或等同物)的重构缓冲液中重构。将约15 μ L重构糜蛋白酶以约1:16的糜蛋白酶-与-样本比例加入到各样本中。将样本与加入的糜蛋白酶于37 $^{\circ}$ C孵育1.5小时,然后将约5 μ L 10%甲酸(v/v)加入到各样本中以猝灭酶促消化。

[0068] 然后,将消化的单克隆抗体的样本分别通过UPLC-MS/MS运行,以在促进肽吸附至柱的条件下分析,所述肽包括在正常条件下不太可能结合的较小链肽。例如,UPLC柱可以是具有预柱(VanGuard, BEH300 C18, 2.1x5mm, 1.7 μ m, Cat# 186003975)的UPLC柱(Waters BEH300 C18, 2.1x100mm, 1.7 μ m, Cat# 186003555)。于约45 $^{\circ}$ C的温度下,将样本以具有约3 μ g/ μ L蛋白浓度的约10 μ L的体积注入柱中。在样本被加载到柱上后,使由流动相A(0.1% TFA水溶液(*Optima LC/MS*, Fisher Scientific, Cat# LS119))和流动相B(0.085% TFA在95%乙腈(ACN)中的溶液(v/v))(HPLC级, Fisher Scientific, Cat # A-998-4或等同物)组成的梯度以约400 μ L/min的流速通过柱。UPLC系统参数如下:自动采样器设置于约4 $^{\circ}$ C;数据速率20 pts/sec;PMT增益为1;和PDA波长210 nm和280 nm。

[0069] 肽标准混合物的梯度如下运行:

时间 (min)	%流动相A	%流动相B
0	100	0

7	67	33
7.5	59	41
8	30	70
8.5	30	70
9	100	0
10	100	0

样本的梯度如下运行：

时间 (min)	%流动相A	%流动相B
0	100	0
10	100	0
15	98	2
75	67	33
85	60	40
90	30	70
95	30	70
96	100	0
98	100	0

样本批次可如下设置：

描述	方法	注入#
平衡运行	平衡	1
水	肽标准品	2
肽标准品	肽标准品	2
水	肽标准品	2
胰蛋白酶-缓冲液空白	肽98 min运行-MS/MS	1
水	肽标准品	2
胰蛋白酶-ONS3010-参比标准品	肽98 min运行-MS/MS	1
水	肽标准品	2
胰蛋白酶-ONS3010-样本	肽98 min运行-MS/MS	1
水	肽标准品	2
糜蛋白酶-缓冲液空白	肽98 min运行-MS/MS	1
水	肽标准品	2
糜蛋白酶-ONS3010-参比标准品	肽98 min运行-MS/MS	1
水	肽标准品	2
糜蛋白酶-ONS3010-样本	肽98 min运行-MS/MS	1
水	肽标准品	2
肽标准品	肽标准品	1
水	肽标准品	2
水	停止	1

Amino Acid Sequence Coverage Search Proteome Discoverer 1.3用于每次UPLC运

行的数据分析。

[0070] 数据接收、记录和报告包括精确的质量校准,之后是将Xcalibur的肽标准品布局(layout)应用于肽标准品注入。记录肽标准品注入的RT在MS色谱图上的5个峰,单独计算每个峰RT的% RSD。记录选择的肽在m/z 532的精确质量,并计算所有注入的质量精确性。

[0071] 肽标准品布局如下:

$$\text{质量精确性(ppm)} = \frac{|\text{测量质量} - \text{理论质量}|}{\text{理论质量}} \times 10^6$$

实施例

[0072] 实施例1:ONS-3010/ONS-1045

根据公开的方法,通过实施缩短的消化时间(对于胰蛋白酶0.5小时)和通过使用允许肽结合于柱的时间的98分钟方法用质谱测试消化的肽,制备ONS-3010(批号X1302-BDS-0)和ONS-1045(批号1407104501)。使用88分钟方法分析过的ONS-3010和ONS-1045的纯化样本使用98分钟方法分析,并将结果与使用88分钟方法获得的结果比较。改进的序列覆盖度的概述在表1中提供。

[0073] 表1:用0.5小时胰蛋白酶消化和98分钟UPLC方法改进序列覆盖度

ONS-3010	重链	轻链
0.5小时胰蛋白酶消化 + 98分钟instr方法	100.00%	98.60%
1小时胰蛋白酶消化 + 88分钟instr方法	95.79%	97.20%

ONS-1045	重链	轻链
0.5小时胰蛋白酶消化 + 98分钟instr方法	100.00%	98.60%
1小时胰蛋白酶消化 + 88分钟instr方法	96.91%	98.60%

胰蛋白酶消化的ONS-3010重链序列覆盖度从95.79%增加至100%,而胰蛋白酶消化的ONS-1045重链序列覆盖度从96.91%增加至100%。

[0074] 根据本公开内容的方法,使用实施98分钟UPLC方法的质谱,分析ONS-1045样本#1407104501和#1408104502。将样本于-80℃冷冻并再次注入具有相同流动相的相同设备上,但用98分钟UPLC方法,其包括在梯度开始时额外10分钟的0.4 mL/min 100%流动相A。先前的方法实施88分钟运行,在运行开始时没有额外10分钟的流动相A。依据再注入的结果,98分钟UPLC方法增加胰蛋白酶消化的样本的氨基酸序列覆盖度。此外,肽CK被检测为裂解的肽CKVSNK,但在88分钟方法中未得到。

[0075] 为评估该测定的特异性,基质、阳性对照(Humira)和阴性对照(Rituxan)用于特异

性的测定。基质显示没有来自靶标序列的命中。在合并胰蛋白酶和糜蛋白酶消化的样本的分析结果后,阳性对照在重链(HC)和轻链(LC)二者上具有针对阿达木单抗序列的100%序列覆盖度。阴性对照显示在Fab区上没有针对阿达木单抗序列的序列覆盖度。一些肽在阴性对照中检测到,因为不同的IgG1恒定区氨基酸序列是相同的。本文的某些图显示来自Proteome Discover的序列覆盖度结果或总离子色谱图。各样本的总体序列覆盖度在表2中记录。

[0076] 表2:特异性的序列覆盖度概述

	重链			轻链		
	Try	Chy	总计	Try	Chy	总计
基质	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
ONS-3010参比标准品	100.0%	100.0%	100.0%	98.6%	89.3%	100.0%
阳性对照-Humira (H35)	98.9%	84.1%	100.0%	98.6%	89.3%	100.0%
ONS-3010样本	98.9%	87.8%	100.0%	98.6%	89.3%	100.0%
阴性对照-Rituxan (M6)	69.4%	50.8%	72.1%	48.6%	45.8%	50.0%

为评估该测定的重现性,ONS-3010参比标准品在每次运行中测试,并重复3次运行。来自胰蛋白酶消化和糜蛋白酶消化的参比标准品和样本的重链和轻链序列覆盖度在表3中记录。

[0077] 表3:ONS-3010序列方法的重现性

	重链			轻链		
	Try	Chy	总计	Try	Chy	总计
#1参比标准品	100.0%	86.3%	100.0%	98.6%	89.3%	100.0%
#1样本	98.9%	87.8%	100.0%	98.6%	89.3%	100.0%
#2参比标准品	99.3%	81.2%	100.0%	98.6%	89.3%	100.0%
#2样本	99.3%	86.3%	100.0%	98.6%	89.3%	100.0%
#3参比标准品	99.3%	81.2%	100.0%	98.6%	89.3%	100.0%
#3样本	99.3%	81.2%	100.0%	98.6%	89.3%	100.0%

虽然以上已结合具体实施方案描述了本公开内容,但是根据以上描述,替代、修改、变换和变化对本领域技术人员而言将变得显而易见。因此,本发明旨在涵盖落入以下权利要求书范围内的所有这样的替代、修改和变化。

[0078] 所要求保护的是:

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
130 135 140

Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145 150 155 160

[0002]

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
165 170 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His
195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys
210 215 220

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
 260 265 270

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 275 280 285

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

[0003]

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

Pro Gly Lys
 450

<210> 2
 <211> 451
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<220>
 <221> MOD_RES
 <222> (22)..(22)
 [0004] <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (96)..(96)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (148)..(148)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (204)..(204)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (222)..(222)
 <223> Xaa = 赖氨酸损失

- <220>
<221> MOD_RES
<222> (224).. (224)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (230).. (230)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (233).. (233)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (265).. (265)
<223> Xaa = 羧甲基
- [0005] <220>
<221> MOD_RES
<222> (301).. (301)
<223> Xaa = N-连接糖基化
- <220>
<221> MOD_RES
<222> (325).. (325)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (371).. (371)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (429).. (429)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (451).. (451)
<223> Xaa = 赖氨酸损失

<400> 2

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Xaa Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
 65 70 75 80

[0006]

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Xaa
 85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
 100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
 115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
 130 135 140

Ala Leu Gly Xaa Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
 145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala

	165		170		175										
Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val
	180		185		190										
Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Xaa	Asn	Val	Asn	His
	195		200		205										
Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Xaa	Ser	Xaa
	210		215		220										
Asp	Lys	Thr	His	Thr	Xaa	Pro	Pro	Xaa	Pro	Ala	Pro	Glu	Leu	Leu	Gly
225			230		235										240
Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met
	245		250		255										
[0007]															
Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Xaa	Val	Val	Val	Asp	Val	Ser	His
	260		265		270										
Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val
	275		280		285										
His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Xaa	Ser	Thr	Tyr
	290		295		300										
Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly
305			310		315										320
Lys	Glu	Tyr	Lys	Xaa	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile
	325		330		335										
Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val

340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
355 360 365

Leu Thr Xaa Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Xaa Ser Val Met
420 425 430

[0008]

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
435 440 445

Pro Gly Xaa
450

<210> 3

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 3

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

[0009] Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
 100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
 115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
 130 135 140

Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
 145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
 165 170 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
 180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His
 195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys
 210 215 220

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
 225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
 245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
 260 265 270

[0010]

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 275 280 285

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

[0011] Pro Gly Lys
 450

<210> 4

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> MOD_RES

<222> (22)..(22)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (96)..(96)

<223> Xaa = 羧甲基

- <220>
<221> MOD_RES
<222> (148)..(148)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (174)..(174)
<223> Xaa = 赖氨酸损失
- <220>
<221> MOD_RES
<222> (204)..(204)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (224)..(224)
<223> Xaa = 羧甲基
- [0012] <220>
<221> MOD_RES
<222> (230)..(230)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (233)..(233)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (265)..(265)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (429)..(429)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (451)..(451)
<223> Xaa = 赖氨酸损失

<400> 4

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Xaa Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
 65 70 75 80

[0013]

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Xaa
 85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
 100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
 115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
 130 135 140

Ala Leu Gly Xaa Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
 145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Xaa Pro Ala

	165		170		175										
Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val
	180		185		190										
Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Xaa	Asn	Val	Asn	His
	195		200		205										
Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	Ser	Xaa
	210		215		220										
Asp	Lys	Thr	His	Thr	Xaa	Pro	Pro	Xaa	Pro	Ala	Pro	Glu	Leu	Leu	Gly
225			230		235										240
Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met
	245		250		255										

[0014]

Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Xaa	Val	Val	Val	Asp	Val	Ser	His
	260		265		270										
Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val
	275		280		285										
His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr
	290		295		300										
Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly
305			310		315										320
Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile
	325		330		335										
Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val

340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
355 360 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Xaa Ser Val Met
420 425 430

[0015]

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
435 440 445

Pro Gly Xaa
450

<210> 5

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 5

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

[0016] Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys
 210

<210> 6
 <211> 214
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<220>
 <221> MOD_RES
 <222> (23)..(23)
 <223> Xaa = 羧甲基

[0017]

<220>
 <221> MOD_RES
 <222> (88)..(88)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (134)..(134)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (194)..(194)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (214)..(214)
 <223> Xaa = 羧甲基

<400> 6

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Xaa Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

[0018] Glu Asp Val Ala Thr Tyr Tyr Xaa Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Xaa Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Xaa Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Xaa
 210

<210> 7

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 7

[0019]

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr

85

90

95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

[0020]

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys
 210

<210> 8

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<220>
 <221> MOD_RES
 <222> (88)..(88)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (134)..(134)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (194)..(194)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (214)..(214)
 <223> Xaa = 羧甲基

<400> 8

[0021]

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Xaa Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Xaa Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

[0022] Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Xaa Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Xaa
 210

<210> 9

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 9

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
65 70 75 80

[0023] Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
130 135 140

Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
165 170 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
 180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His
 195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys
 210 215 220

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
 225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
 245 250 255

[0024] Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
 260 265 270

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 275 280 285

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

[0025]

Pro Gly Lys
 450

<210> 10

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> MOD_RES

<222> (22) .. (22)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

- <222> (96).. (96)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (148).. (148)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (204).. (204)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (222).. (222)
<223> Xaa = 赖氨酸损失
- <220>
<221> MOD_RES
<222> (224).. (224)
<223> Xaa = 羧甲基
- [0026] <220>
<221> MOD_RES
<222> (230).. (230)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (233).. (233)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (265).. (265)
<223> Xaa = 羧甲基
- <220>
<221> MOD_RES
<222> (301).. (301)
<223> Xaa = N-连接糖基化
- <220>

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Xaa
85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
130 135 140

Ala Leu Gly Xaa Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145 150 155 160

[0028]

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
165 170 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Xaa Asn Val Asn His
195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Xaa Ser Xaa
210 215 220

Asp Lys Thr His Thr Xaa Pro Pro Xaa Pro Ala Pro Glu Leu Leu Gly
225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Xaa Val Val Val Asp Val Ser His
 260 265 270

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 275 280 285

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Xaa Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Xaa Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

[0029] Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Xaa Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Xaa Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

Pro Gly Xaa
 450

<210> 11

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 11

[0030]

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
 100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
 115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
 130 135 140

Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
 145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
 165 170 175

[0031] Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
 180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His
 195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys
 210 215 220

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
 225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
 245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His
 260 265 270

	Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val		
	275	280	285
	His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr		
	290	295	300
	Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly		
	305	310	315 320
	Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile		
	325	330	335
	Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val		
	340	345	350
[0032]	Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser		
	355	360	365
	Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu		
	370	375	380
	Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro		
	385	390	395 400
	Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val		
	405	410	415
	Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met		
	420	425	430
	His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser		
	435	440	445

Pro Gly Lys
450

<210> 12
<211> 451
<212> PRT
<213> 人工序列

<220>
<223> 合成构建体

<220>
<221> MOD_RES
<222> (22)..(22)
<223> Xaa = 羧甲基

<220>
<221> MOD_RES
<222> (96)..(96)
<223> Xaa = 羧甲基

[0033]

<220>
<221> MOD_RES
<222> (148)..(148)
<223> Xaa = 羧甲基

<220>
<221> MOD_RES
<222> (174)..(174)
<223> Xaa = 赖氨酸损失

<220>
<221> MOD_RES
<222> (204)..(204)
<223> Xaa = 羧甲基

<220>
<221> MOD_RES
<222> (224)..(224)
<223> Xaa = 羧甲基

<220>

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Xaa
85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
130 135 140

Ala Leu Gly Xaa Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145 150 155 160

[0035]

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Xaa Pro Ala
165 170 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Xaa Asn Val Asn His
195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Xaa
210 215 220

Asp Lys Thr His Thr Xaa Pro Pro Xaa Pro Ala Pro Glu Leu Leu Gly
225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Xaa Val Val Val Asp Val Ser His
 260 265 270

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 275 280 285

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

[0036]

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Xaa Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

Pro Gly Xaa
 450

<210> 13

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 13

[0037] Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

[0038]

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys
 210

<210> 14

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> MOD RES

<222> (23).. (23)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (88)..(88)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (134)..(134)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (194)..(194)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (214)..(214)

<223> Xaa = 羧甲基

[0039]

<400> 14

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Xaa Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Xaa Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Xaa Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

[0040] Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Xaa Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Xaa
 210

<210> 15

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 15

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

[0041]

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys
 210

<210> 16

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

[0042]

<220>

<221> MOD_RES

<222> (88)..(88)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (134)..(134)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (194)..(194)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (214)..(214)

<223> Xaa = 羧甲基

<400> 16

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

[0043] Glu Asp Val Ala Thr Tyr Tyr Xaa Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Xaa Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Xaa Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Xaa
 210

<210> 17

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 17

[0044]

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys

85

90

95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
 100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr His Gly Pro Ser
 115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
 130 135 140

Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
 145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
 165 170 175

[0045]

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
 180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His
 195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys
 210 215 220

Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly
 225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
 245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val Ser His

	260		265		270
Glu Asp Pro	Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val				
	275		280		285
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr					
	290		295		300
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly					
305		310		315	320
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile					
	325		330		335
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val					
	340		345		350
[0046]					
Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser					
	355		360		365
Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu					
	370		375		380
Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro					
385		390		395	400
Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val					
	405		410		415
Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Pro Ser Cys Ser Val Met					
	420		425		430
His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser					

435

440

445

Pro Gly Lys

450

<210> 18

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> MOD_RES

<222> (148)..(148)

<223> Xaa = 羧甲基

[0047]

<220>

<221> MOD_RES

<222> (204)..(204)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (224)..(224)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (230)..(230)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (233)..(233)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (265)..(265)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (325)..(325)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (371)..(371)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (429)..(429)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (451)..(451)

<223> Xaa = 赖氨酸损失

[0048]

<400> 18

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr His Gly Pro Ser
115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
130 135 140

Ala Leu Gly Xaa Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145 150 155 160

[0049] Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
165 170 175

Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
180 185 190

Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Xaa Asn Val Asn His
195 200 205

Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Xaa
210 215 220

Asp Lys Thr His Thr Xaa Pro Pro Xaa Pro Ala Pro Glu Leu Leu Gly
225 230 235 240

Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met
245 250 255

Ile Ser Arg Thr Pro Glu Val Thr Xaa Val Val Val Asp Val Ser His
 260 265 270

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val
 275 280 285

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Xaa Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

[0050] Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Xaa Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Pro Ser Xaa Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

Pro Gly Xaa
 450

<210> 19

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 19

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
 1 5 10 15

[0051]

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
 20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
 50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
 100 105 110

His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr
 290 295 300

Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly
 305 310 315 320

Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile
 325 330 335

Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 340 345 350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

[0053]

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Pro Ser Cys Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

Pro Gly Lys
 450

<210> 20
 <211> 451
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成构建体

<220>
 <221> MOD_RES
 <222> (22)..(22)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (148)..(148)
 <223> Xaa = 羧甲基

[0054] <220>
 <221> MOD_RES
 <222> (174)..(174)
 <223> Xaa = 赖氨酸损失

<220>
 <221> MOD_RES
 <222> (265)..(265)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (429)..(429)
 <223> Xaa = 羧甲基

<220>
 <221> MOD_RES
 <222> (451)..(451)
 <223> Xaa = 赖氨酸损失

<400> 20

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg

1	5	10	15
Ser Leu Arg Leu Ser Xaa Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr	20	25	30
Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val	35	40	45
Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val	50	55	60
Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr	65	70	75
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys	85	90	95
[0055]			
Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly	100	105	110
Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser	115	120	125
Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala	130	135	140
Ala Leu Gly Xaa Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val	145	150	155
Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Xaa Pro Ala	165	170	175
Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val			

	180		185		190
Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn His					
	195		200		205
Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys Ser Cys					
	210		215		220
Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu Leu Gly					
225		230		235	240
Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met					
	245		250		255
Ile Ser Arg Thr Pro Glu Val Thr Xaa Val Val Val Asp Val Ser His					
	260		265		270

[0056]

Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val Glu Val					
	275		280		285
His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser Thr Tyr					
	290		295		300
Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly					
305		310		315	320
Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile					
	325		330		335
Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val					
	340		345		350
Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser					

355

360

365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Pro Ser Xaa Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

[0057]

Pro Gly Xaa
 450

<210> 21

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 21

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 I 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

[0058]

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys
210

<210> 22

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<220>

<221> MOD_RES

<222> (134)..(134)

<223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (194)..(194)

[0059] <223> Xaa = 羧甲基

<220>

<221> MOD_RES

<222> (214)..(214)

<223> Xaa = 羧甲基

<400> 22

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

Tyr Ala Ala Ser Thr Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly

<211> 214

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 23

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

[0061] Tyr Ala Ala Ser Thr Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
 85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
 100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
 115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
 130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
 145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
 165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Cys
 210

- [0062] <210> 24
 <211> 214
 <212> PRT
 <213> 人工序列
- <220>
 <223> 合成构建体
- <220>
 <221> MOD_RES
 <222> (134)..(134)
 <223> Xaa = 羧甲基
- <220>
 <221> MOD_RES
 <222> (194)..(194)
 <223> Xaa = 羧甲基
- <220>
 <221> MOD_RES
 <222> (214)..(214)
 <223> Xaa = 羧甲基

<400> 24

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65 70 75 80

[0063]

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125

Thr Ala Ser Val Val Xaa Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser

165

170

175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
 180 185 190

Ala Xaa Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
 195 200 205

Phe Asn Arg Gly Glu Xaa
 210

<210> 25

<211> 214

<212> PRT

<213> 人工序列

<220>

[0064] <223> 合成构建体

<400> 25

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Arg Asn Tyr
 20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45

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

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80

Glu Asp Val Ala Thr Tyr Tyr Cys Gln Arg Tyr Asn Arg Ala Pro Tyr
85 90 95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145 150 155 160

[0065] Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195 200 205

Phe Asn Arg Gly Glu Cys
210

<210> 26

<211> 451

<212> PRT

<213> 人工序列

<220>

<223> 合成构建体

<400> 26

Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Arg
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Asp Asp Tyr
20 25 30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Ala Ile Thr Trp Asn Ser Gly His Ile Asp Tyr Ala Asp Ser Val
50 55 60

Glu Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Ser Leu Tyr
65 70 75 80

[0066]

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Lys Val Ser Tyr Leu Ser Thr Ala Ser Ser Leu Asp Tyr Trp Gly
100 105 110

Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
115 120 125

Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala
130 135 140

Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
145 150 155 160

Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala

	165		170		175										
Val	Leu	Gln	Ser	Ser	Gly	Leu	Tyr	Ser	Leu	Ser	Ser	Val	Val	Thr	Val
	180		185		190										
Pro	Ser	Ser	Ser	Leu	Gly	Thr	Gln	Thr	Tyr	Ile	Cys	Asn	Val	Asn	His
	195		200		205										
Lys	Pro	Ser	Asn	Thr	Lys	Val	Asp	Lys	Lys	Val	Glu	Pro	Lys	Ser	Cys
	210		215		220										
Asp	Lys	Thr	His	Thr	Cys	Pro	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly
225			230		235										240
Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met
	245		250		255										
[0067]															
Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	Val	Val	Asp	Val	Ser	His
	260		265		270										
Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val
	275		280		285										
His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr
	290		295		300										
Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly
305			310		315										320
Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile
	325		330		335										
Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val

340

345

350

Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln Val Ser
 355 360 365

Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu
 370 375 380

Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro
 385 390 395 400

[0068]

Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val
 405 410 415

Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met
 420 425 430

His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser
 435 440 445

Pro Gly Lys
 450

胰蛋白酶消化的

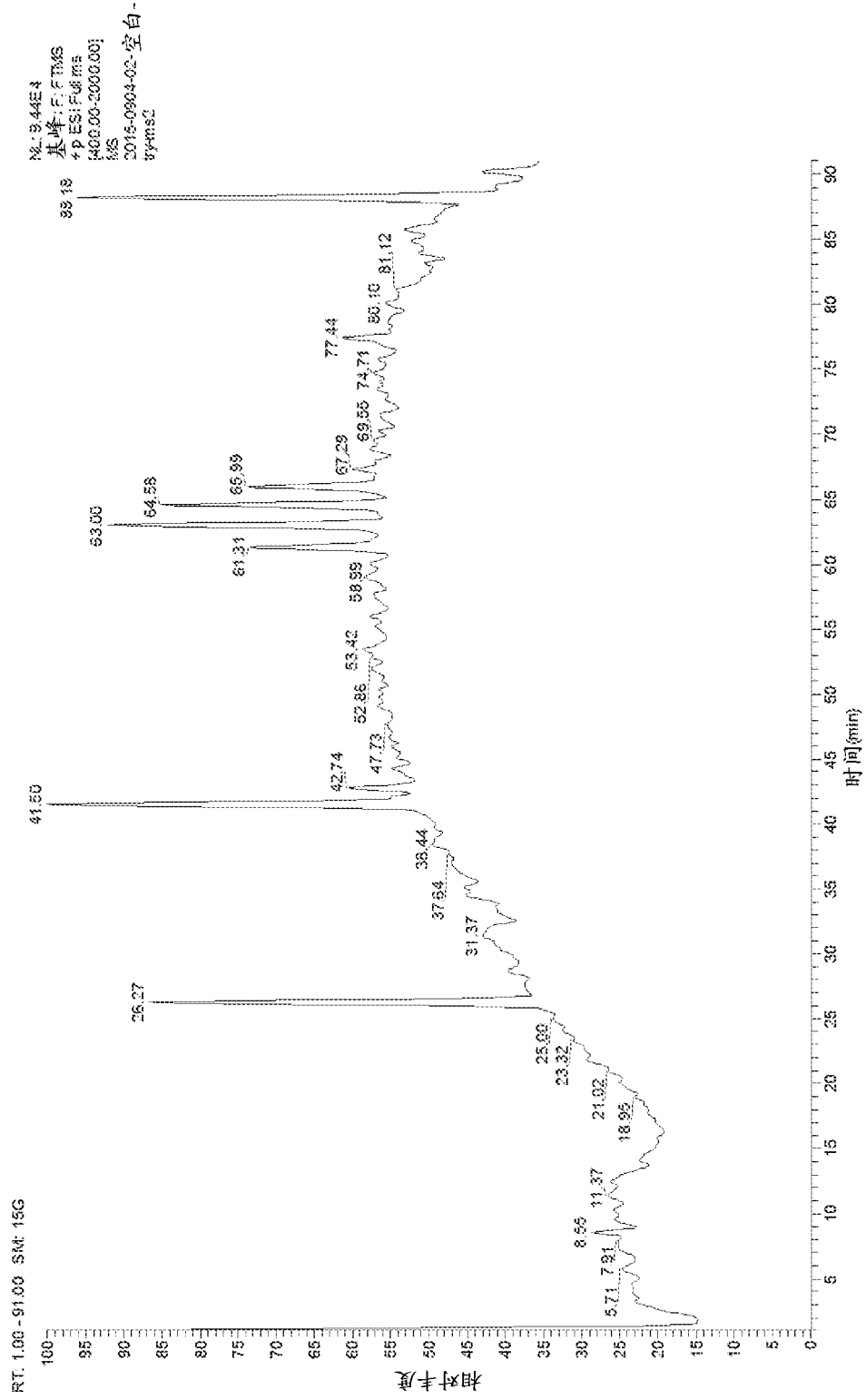


图 1a

糜蛋白酶消化的

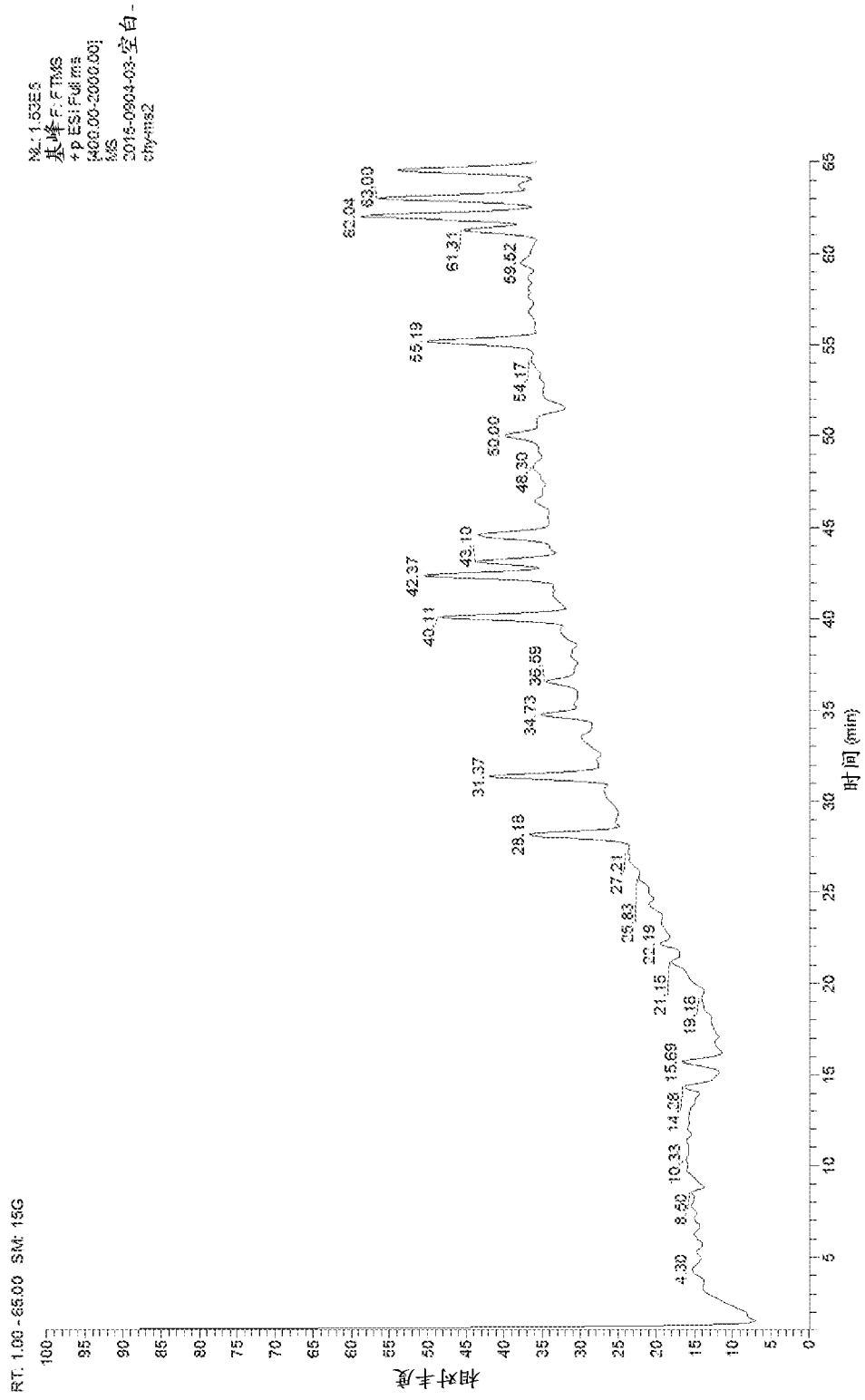


图 1b

覆盖度: 100.00%

发现的修饰:

C 羧甲基(C)
L 赖氨酸损失(C-端)
N N-连接糖基化(N)

序列	修饰列表
1 修饰 Humira	1 11 21 31 41 51 61 71 81 91 C EQVVESEGG LADPERSLRL SCRSSETEB DYAMRWVQA PENGLWNSR ITWNSCHLWY RDSVECHETI SRNANNSLY LQNSLNEB TAVYERKYS C
101 修饰 Humira	101 C YLSTRSLRY WSGCLTPTVS SAKKPEEVE PIRSSSSTS GETRAGSLN KQYPERWY SWSSEALUSG WQEDPAVLQS SGLYSSLSW TPSSSLSTQ C
201 修饰 Humira	201 C LC C C C TYLGNWKEP SNEKIDKWE PESCKTHC PDCAPHLIG EPSVLETPK PDLTALSR DEVICTWVY SHEDPEVEN WYVNGYRNS AKTPEEY C
301N 修饰 Humira	301N C NSTRVWVL TULQWMLG KEVCKYSK ALPAPLETI SKNGQDPEP QUTLQPSD ELTNDQSLT CLVNGTQSD IADWESNQQ PENNKYTP C
401 修饰 Humira	401 C L VLDSESEIL YSELFVWNSK WQNGWVCS VAEERLHNY TQKSLSLSPG K C

----- 高可能性
- - - - - 中可能性
..... 低可能性

图 2a

序列	修饰列表
1	11 21 31 41 51 61 71 81 91
1 修饰 Humira	EVOLVESEGG IYQPERSEL SCHSCEPIED DYMHWYFOR PEKLEWYER ITYNSCHIDY ROSTEGBEFTI SEOMIENSLY LOMGSIERED TRVYCKEYS C
101 修饰 Humira	YLSTRSSLDY WQQTILWVS SAGNEPSTV BLDSESTG CEFRAICELV KYZPEPUDV SONSCHLPG VHTREDAVQS SCLYSLSSTV TPDSSELETO L
201 修饰 Humira	TYICNNEP SNTKTKKVE PASCHEHC PCORPELLG GRSUELEPK FKOALISU PEVTCVWDV SIBBEUKEN WYHGVYEN AKIKPEEQU C C C C
301 修饰 Humira	NSYVVSUML FVIMQWMLG KEVYCVSUK ALPAPIETI SKAKQPEEP CVYVLEPSB ELTENQSLY CLUNGFYSD INAVESNGO PENYKXTPP
401 修饰 Humira	VLDGSGEEL YERLYVYKSR WAGCVFSES VMEHMLNHY NOKSLSLSPG K L

覆盖度: 86.25%

发现的修饰:

- C 羧甲基(C)
- L 赖氨酸损失(C-端)

高可能性

中可能性

低可能性

图 2b

覆盖度: 98.60%

发现的修饰:

C 羧甲基(C)

序列	修饰列表																														
修饰 Humira	<table border="1"> <thead> <tr> <th>1</th> <th>11</th> <th>21</th> <th>31</th> <th>41</th> <th>51</th> <th>61</th> <th>71</th> <th>81</th> <th>91</th> </tr> </thead> <tbody> <tr> <td>1</td> <td></td> <td>C</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td colspan="10">DIOMQSPSS LSSVSRVTF ITCRSQIR NTLNMQDF GRAPKLIYA MTLQSEYD HEGESQTD PHLISSLQ EDVATYYQR YNEPTTQ</td> </tr> </tbody> </table>	1	11	21	31	41	51	61	71	81	91	1		C								DIOMQSPSS LSSVSRVTF ITCRSQIR NTLNMQDF GRAPKLIYA MTLQSEYD HEGESQTD PHLISSLQ EDVATYYQR YNEPTTQ									
1	11	21	31	41	51	61	71	81	91																						
1		C																													
DIOMQSPSS LSSVSRVTF ITCRSQIR NTLNMQDF GRAPKLIYA MTLQSEYD HEGESQTD PHLISSLQ EDVATYYQR YNEPTTQ																															
修饰 Humira	<table border="1"> <thead> <tr> <th>101</th> <th>111</th> <th>121</th> <th>131</th> <th>141</th> <th>151</th> <th>161</th> <th>171</th> <th>181</th> <th>191</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td colspan="10">GKVEIKRV ARSVETTFP SEQLKGTA SPULNMFY FRALPQAV ENLQSEHQ ESTEQSDQ STYLSSTF LSRDYEEK YKCEYTHG</td> </tr> </tbody> </table>	101	111	121	131	141	151	161	171	181	191											GKVEIKRV ARSVETTFP SEQLKGTA SPULNMFY FRALPQAV ENLQSEHQ ESTEQSDQ STYLSSTF LSRDYEEK YKCEYTHG									
101	111	121	131	141	151	161	171	181	191																						
GKVEIKRV ARSVETTFP SEQLKGTA SPULNMFY FRALPQAV ENLQSEHQ ESTEQSDQ STYLSSTF LSRDYEEK YKCEYTHG																															
修饰 Humira	<table border="1"> <thead> <tr> <th>201</th> <th>211</th> <th>221</th> <th>231</th> <th>241</th> <th>251</th> <th>261</th> <th>271</th> <th>281</th> <th>291</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td colspan="10">LSPVTKSN KCEC</td> </tr> </tbody> </table>	201	211	221	231	241	251	261	271	281	291											LSPVTKSN KCEC									
201	211	221	231	241	251	261	271	281	291																						
LSPVTKSN KCEC																															

..... 高可能性

..... 中可能性

..... 低可能性

图 2c

覆盖度: 89.25%

发现的修饰:

C 羧甲基(C)

序列	修饰列表
修饰 Humira	<pre> 1 11 21 31 41 51 61 71 81 91 1 MDTGKSPK LKASVGDVNT ITPAQGQIP NYLAWYQKD GEPKLLIYA KSTLQSGDPS KESGEGSDP EPTLSSQD KQKATYTOR YNRAPYFQK </pre>
修饰 Humira	<pre> 101 GTAVLEKPY NRPVYIEPP SGEQKSGEK SVLLLENTY PRRMTPQKY KATKQSGNQ KSTFQDSEK STYFLSSIT LSKDYTEK KTRKTYTQK </pre>
修饰 Humira	<pre> 201 LSSPVEKSN RQKRC </pre>

高可能性

中可能性

低可能性

图 2d

胰蛋白酶消化的

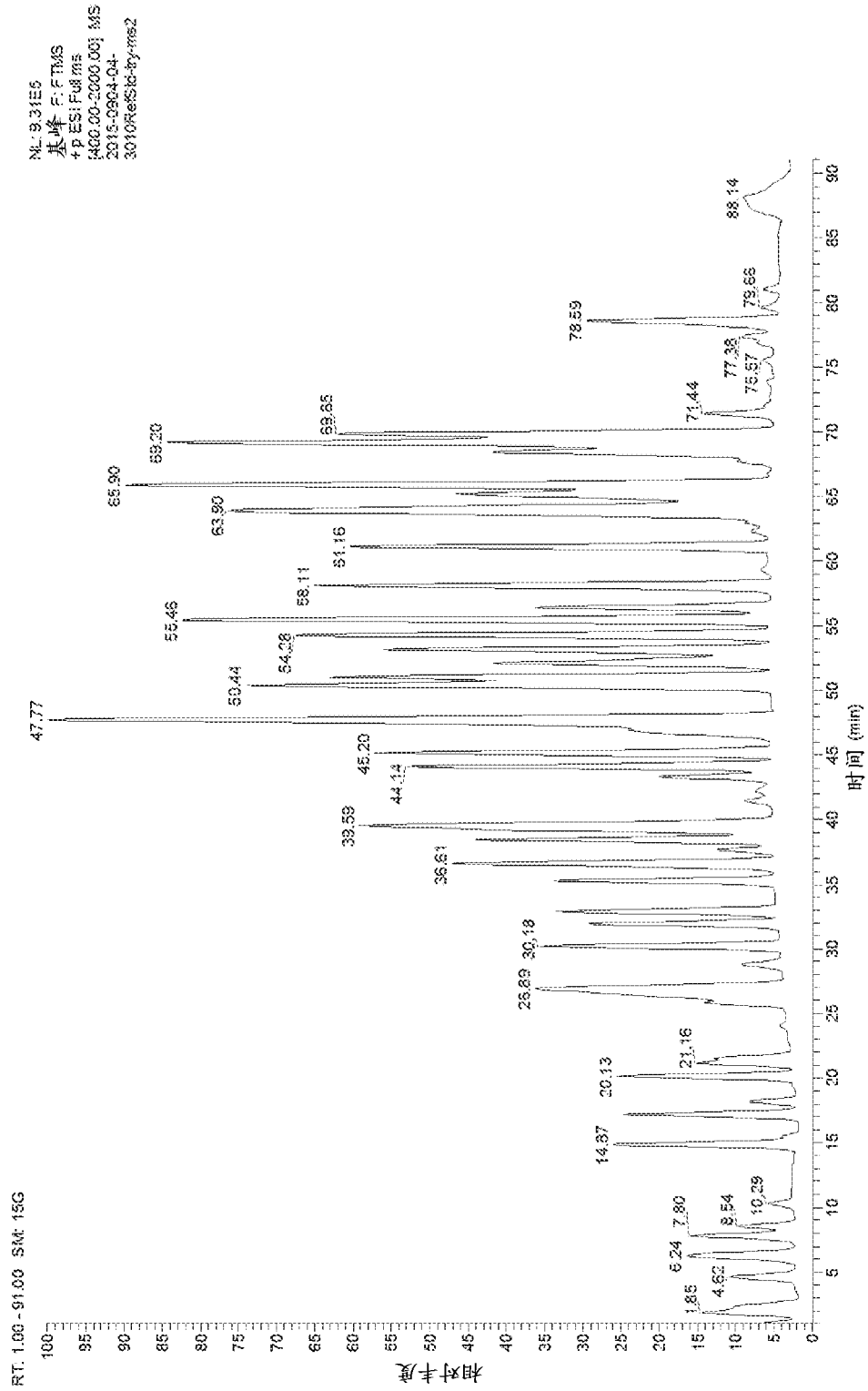


图 3a

糜蛋白酶消化的

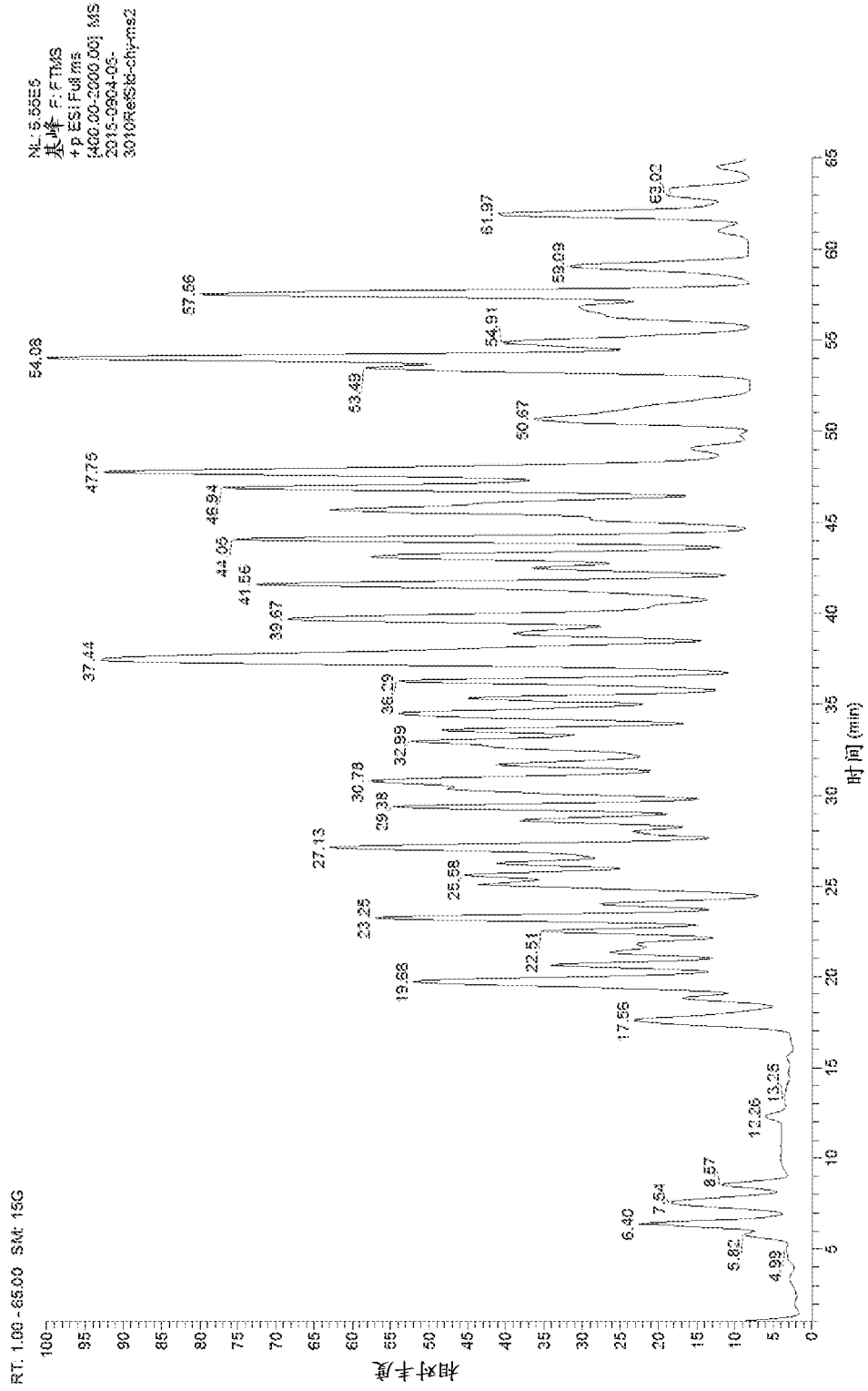


图 3b

覆盖度: 98.69%

发现的修饰:

- C 羧甲基(C)
- L 赖氨酸损失(C-端)
- N N-连接糖基化(N)

序列	修饰列表
1 11 21 31 41 51 61 71 81 91	
1 Humira	EVQLVESGGGLVQPGASLRLSASGQSFDYWGDIYVQA RGSGLIENYSA ITWNSCHIDY RDSYEEKFTI SEDNENGLY LQNSLRAR TRVYTCRYS C
101 Humira	YLSYASSELNY WQGNLTWVS SASYKRPQVE PLAPSSKSY S GEFRALEGLV KNYTPPEPTV SWNSGALISG WITHPKVLQS SGALYSLSWV TWPSSLSGQ C
201 Humira	TYLQWNRKP SNTKPRKVE PASCQKTHC PEPARELLG GDSVELEPPK PAFPLAKSI PFTVYQWIV SHEDPEVEN WYDVEYEVK AKTRPRRQ C
301 Humira	KSTYRFEVI TVLQWHLG KTKLQVSK ALRPIKTI SGRKQKPEP QFTLEPSSR ELIENQSLI CLKGCFYED INVEESNGR RENTFTTTP C
401 Humira	VLSGSETEI YSKLTWDSR WQGGVPSGS WDEALNNY TQKSLSPG K C

----- 高可能性

----- 中可能性

..... 低可能性

图 4a

覆盖度: 85.14%

发现的修饰:

C 羧甲基(C)

L 赖氨酸损失(C-端)

序列	修饰列表
1	11 21 31 41 51 61 71 81 91
修饰	C
Humira	EVLVVEGG LYPGSEEL SCASSETD HYAMVIRQ PGNLEWTA ITWNSCHIF ADVEGHEFI SEWAKESLY LQNSLAKED TRFYCKEYS
修饰	L
Humira	YLTRSSLY MQGTVTVS SMTREPSVF EIRDSSEYS GSTRUGCLV KVEPRPVV SENSCHTSG VRTFAYVDS SGLYSLSSTV FVPSSEIGIQ
修饰	C C C C
Humira	FYIENSHQ SMTATDKEE PSCDENEK PRCYRDLG CSUHELEPK PMLIKISEF PEVICVVIVY SDEEVEN WYDCEYVEN AKTREREQY
修饰	L
Humira	NSYEVVSL FVLHUMLNG KKYKLVSEK ALPAPIETI SEAKQPEEP QVYLLPSSQ ELTKMUSLT CLAKLEYSD IAKNESNQ PENNYKTFP
修饰	C
Humira	VLSGSEEL YSKLTVKSR WQGVTFSES VDEELAHNY FQKSLSPG K

----- 高可能性

----- 中可能性

..... 低可能性

图 4b

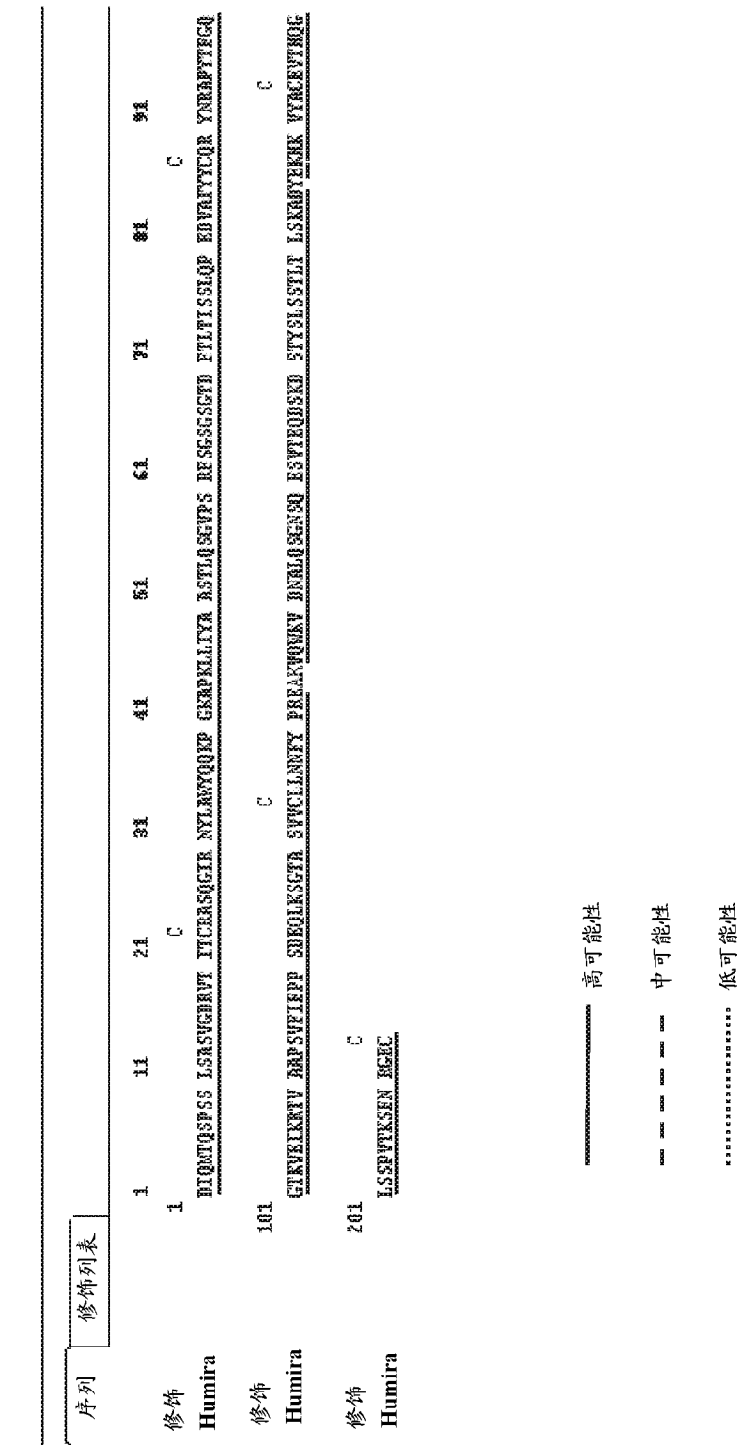


图 4c

覆盖度: 89.25%

发现的修饰:

C 羧甲基(C)

序列	修饰列表
1 Humira 1 NIGHTSPSS LSASVGDVDT IICRASQCIE NYLWVQOE GEPKLLIYA RFLHSGHPS KESGSGSD EHLTSSLOP EDRTFTYCOR YNRKPTTCA	1 11 21 31 41 51 61 71 81 91 C
101 Humira 101 GTFVETKRY RHPVETLPP GDEQLSGFR SVVLLNNFY PKEKTKVNV RHLQSGNQ LSTVQDSEB STYSLSSTLF LSKRHTEEK VTKLFTVHC	101 C
201 Humira 201 LSSFTKSEN REBC	201 C

高可能性

中可能性

低可能性

图 4d

胰蛋白酶消化的

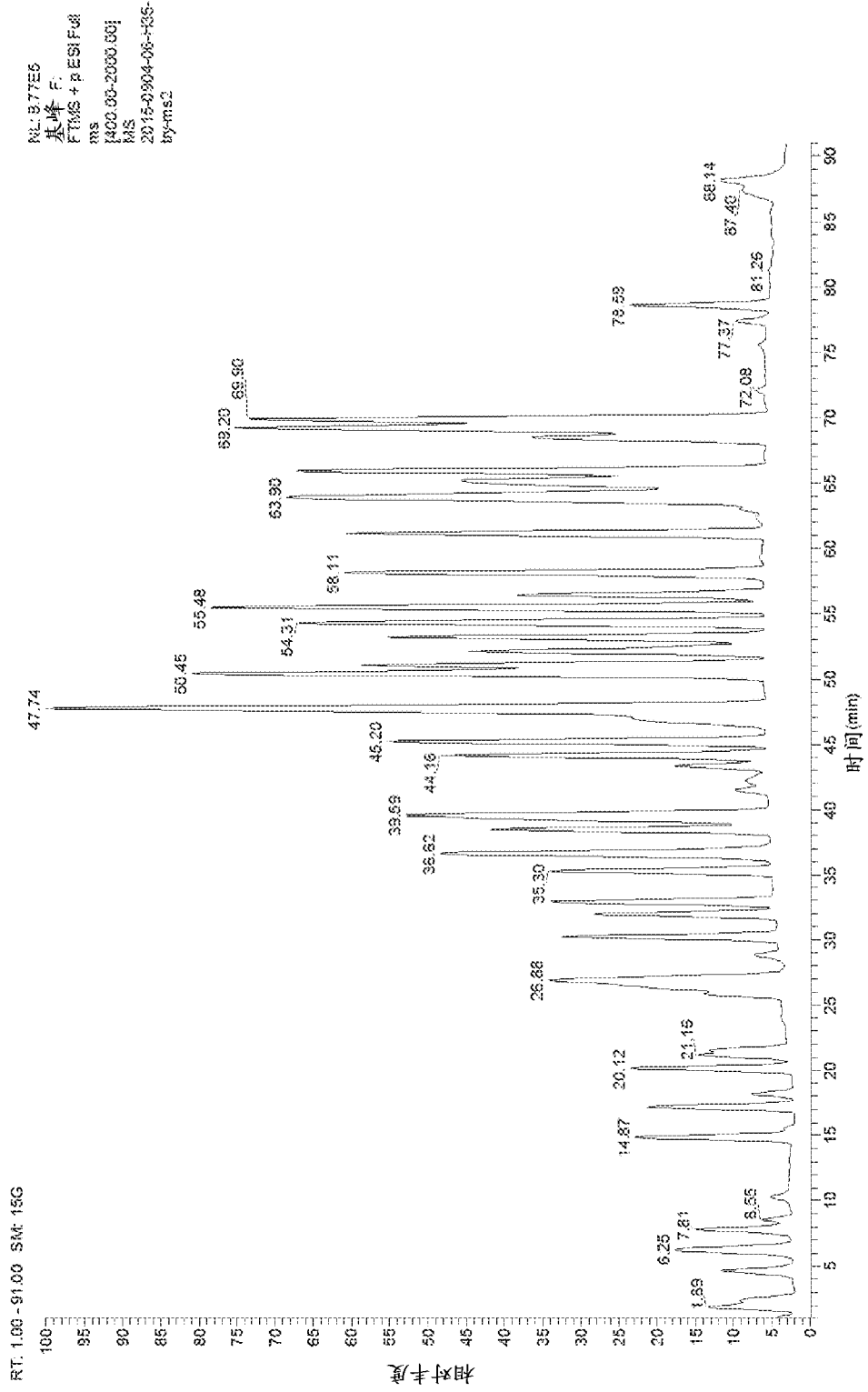


图 5a

糜蛋白酶消化的

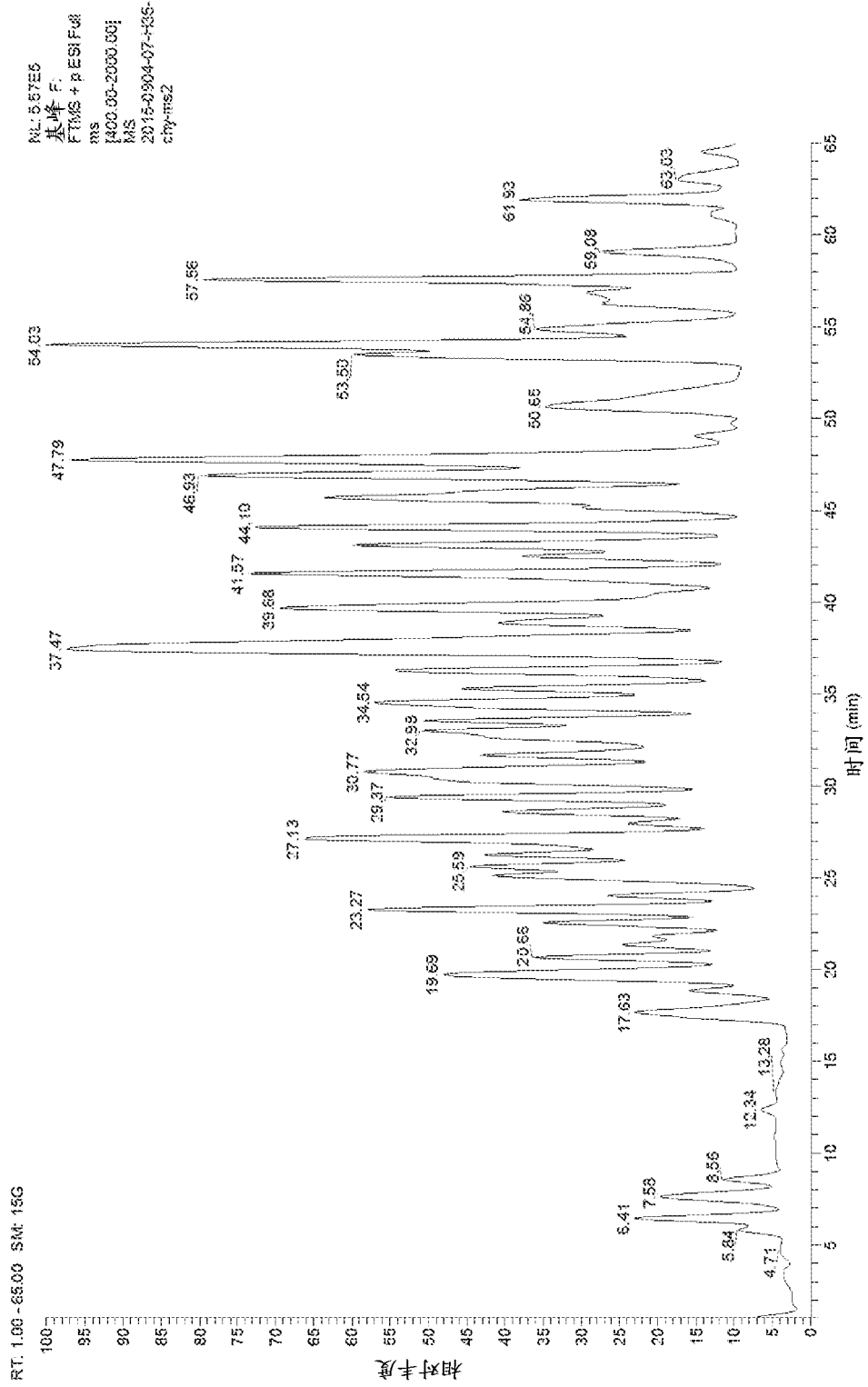


图 5b

覆盖度: 69.40%

发现的修饰:

C 羧甲基(C)
L 赖氨酸损失(C-端)

序列	修饰列表	1	11	21	31	41	51	61	71	81	91
1 Humira	修饰 Humira	1	11	21	31	41	51	61	71	81	91
101 Humira	修饰 Humira	101									
201 Humira	修饰 Humira	201									
301 Humira	修饰 Humira	301									
401 Humira	修饰 Humira	401									

..... 高可能性
 - - - - - 中可能性
 低可能性

图 6a

序列	修饰列表
1	11 21 31 41 51 61 71 81 91
1	C
Humira	KVQLVDSGGG LVQPGRGRLG LSRRAISGTFED DYAEHHUNQA PEGSLELVSA ITNNSCHIDY AHSVNGRFTI ESDMANKSLY LQWNSLRERD TAVTYCAKYS
101	C
Humira	YLSLASELDY WQGLTIVTS SASTREPSWV PLAPSEKYS GSTRNIGELV KNTTEPPTV SANSALNSG VHTFPAVIGS SELYSLSLWV TURSSELSGQ
201	C
Humira	TYICWNEKP ENIKYDKEVE PRCSCNTHC PPGAPAEILG EPSVELEPPK PRTNLEKNI PAVICTYVWV SHEPVEEN WYDQGVVHN AKIKPEEIKY
301	C
Humira	NSITRFVSL YVLRWVWNG KYTKGVENK ALPAPIKTI SSKAQDPEP QVTFEPPD EITKNSVLI CLVMEYTSB IAVVESKGO PERRVYTFP
401	C L
Humira	VLSGSEEL YEKLVVDSR WQGVTFSCS VHEHRLNHY TQKLSLSFG K

覆盖率: 58.78%

发现的修饰:

- C 羧甲基(C)
- L 赖氨酸损失(C-端)

----- 高可能性

----- 中可能性

..... 低可能性

图 6b

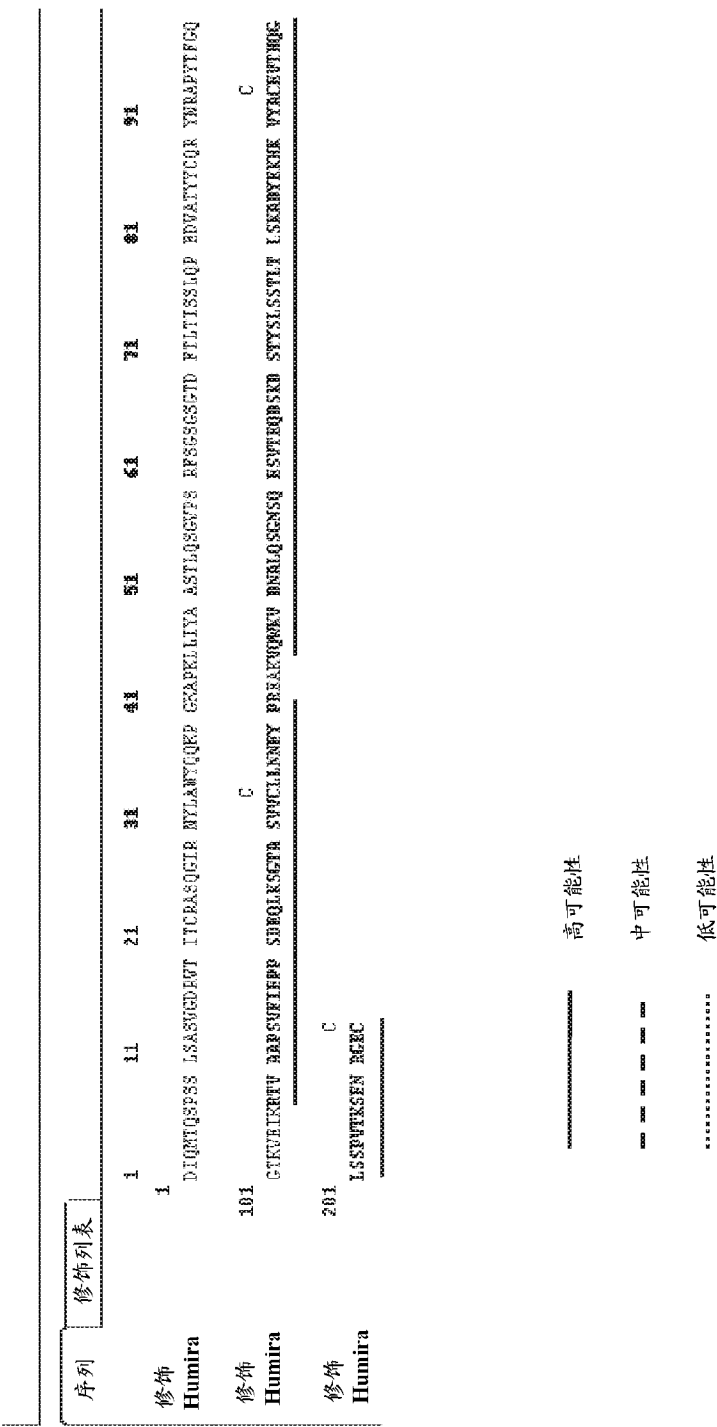


图 6c

覆盖度: 45.74%

发现的修饰:

C 羧甲基(C)

序列	修饰列表
1 11 21 31 41 51 61 71 81 91 1 DIGNTQGESE LGASVDEWVT ITCFASQGLE NYLAWYQOKP GKAPKLLIYA ASTIQSEVES EFGSSGSETD FILLISSIQE EDVATYTCUR YNRKAFYFGQ Humira	
101 C GTEVIEEIV AAPGVFIKPP SDRLKSEFA SVLELNNFY PRERKPKVK ENRIQSENG ESVTEKSKD STYLSSTII LSKRDYEEK YTRKRVING Humira	
201 C LSSPYIKEN RCFEC Humira	

高可能性

中可能性

低可能性

图 6d

胰蛋白酶消化的

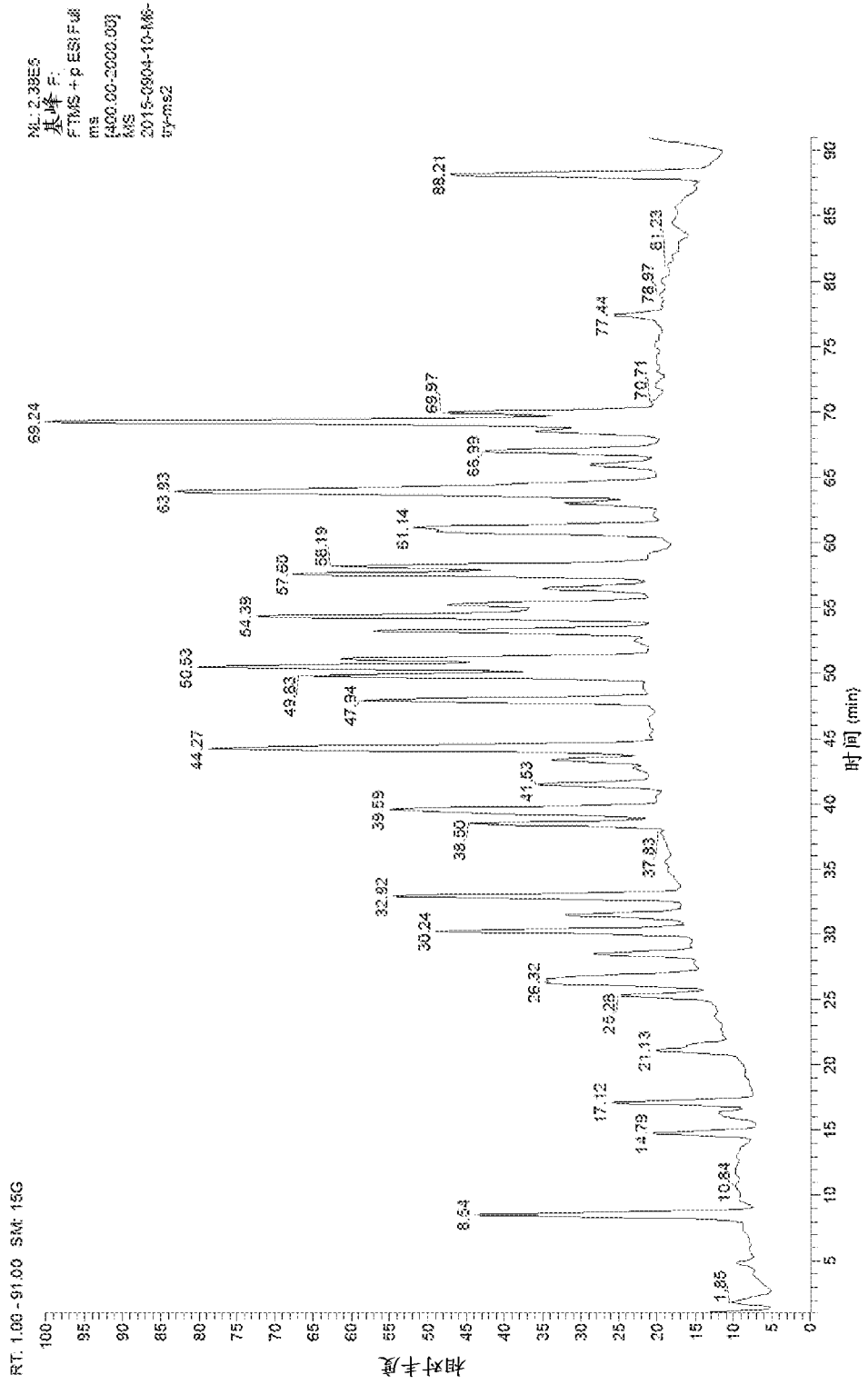


图 7a

糜蛋白酶消化的

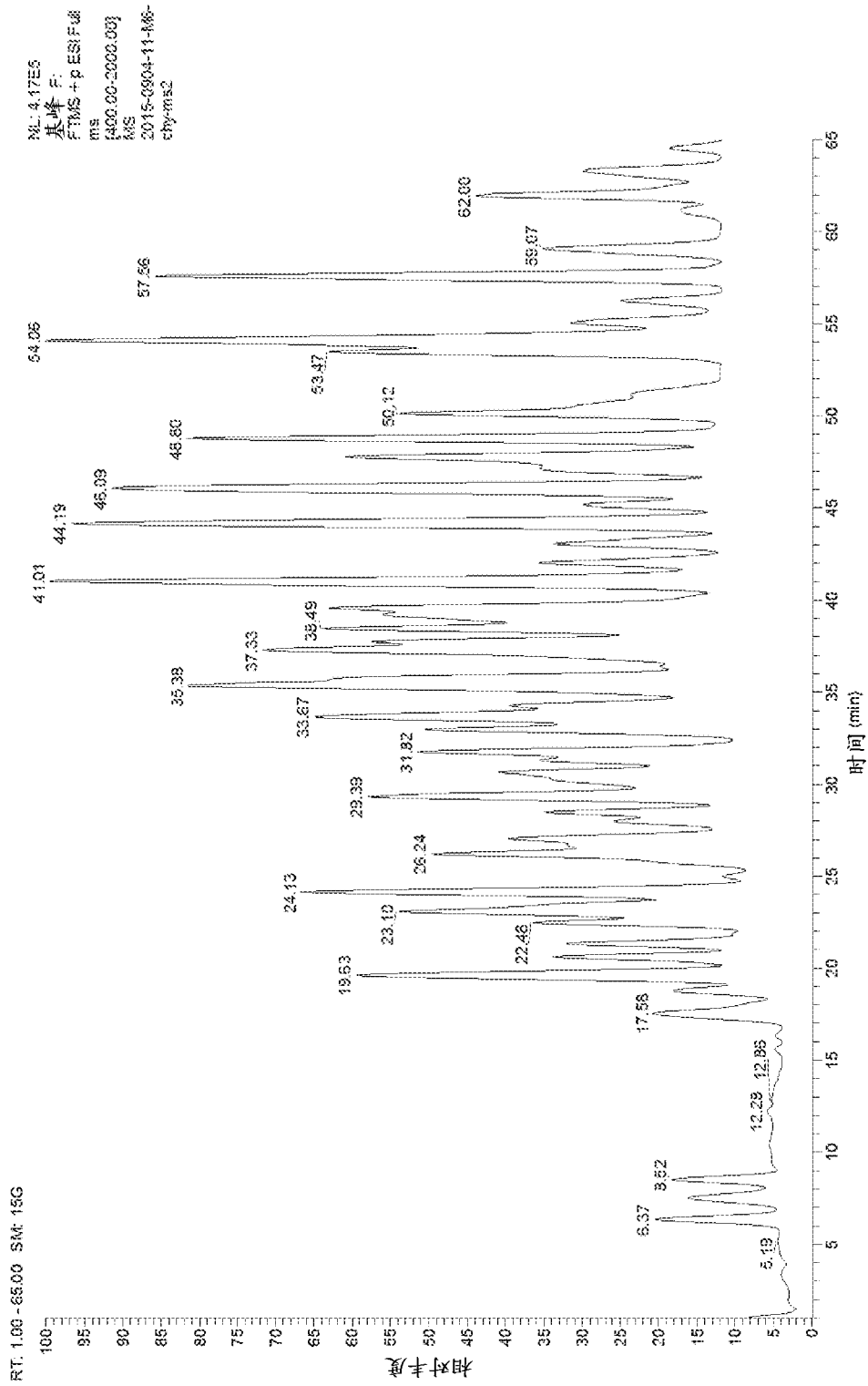


图 7b

说明Humira的轻链和重链的理论氨基酸序列

轻链:

```

10 20 30 40 50 60
DIQMTQSPSS LSASVGDKVT ITCRASQGIT NYLAWYQCKP GKAPKLLIYA ASTLQSGVPS
70 80 90 100 110 120
RFSGCSGTD FTLTISSLPQ EDVAIYCCQR YNRAPYTFGQ GTKVEIKRIV AAPSVFIFPP
130 140 150 160 170 180
SDECIKSGTA SVVCLLNPFY PREAKVQMKV DNALQSCNSQ ESVTFQDSKD STYLSLSSTIE
190 200 210
LSKADYEKHK VYACEVTHQG LSPVTKSFN RCEC

```

重链:

```

10 20 30 40 50 60
EVCLVESGGG LVQPKRSLRL SCAASGFTFD DYAMHWVROA FKGLEWVSA ITWNSGHIDY
70 80 90 100 110 120
ADSVEGRETI SRDNANKSLY LQNSLRAED TAVYICAKVS YLSTASSLDY WGQGTLLVTVS
130 140 150 160 170 180
SASTKGPSVF PLAPSSKSTP GGTAALGCLV KDYFEPVTV SWNSGALTSQ VHIEPAVLQS
190 200 210 220 230 240
SGLYSLSVVV TVPSSSLGTD TYICNVNKKP SNTKVKKVE FKSCDKHTC PCFPAPPELLG
250 260 270 280 290 300
GPSVFLPEK FKDFIMISRT PEVTCVVVDV SHEDPEVKFN WYVDGVEVHN ARTKPREPEQY
310 320 330 340 350 360
NSTYRVSVL TVLHQDWLNG KEYKCKVSNK ALPAPIEKTI SKRAGQPREP QVYTLPPSRD
370 380 390 400 410 420
ELTEKQVSLT CLVKGFTPSD IAVEWESNGQ PENNYKTPPP VLDSGSEFFL YSKLTVDKSR
430 440 450
WQGGNVFCSY VMHEALHMHY TOKSLSLSPCK

```

图 8