



(12)发明专利申请

(10)申请公布号 CN 108290935 A

(43)申请公布日 2018.07.17

(21)申请号 201680058626.5

(74)专利代理机构 北京纪凯知识产权代理有限公司 11245

(22)申请日 2016.08.10

代理人 王永伟 赵蓉民

(30)优先权数据

62/203,102 2015.08.10 US

62/293,880 2016.02.11 US

62/317,572 2016.04.03 US

(51)Int.Cl.

G07K 14/435(2006.01)

C12N 15/866(2006.01)

(85)PCT国际申请进入国家阶段日

2018.04.08

(86)PCT国际申请的申请数据

PCT/IL2016/050874 2016.08.10

(87)PCT国际申请的公布数据

W02017/025964 EN 2017.02.16

(71)申请人 赛威克斯材料科学公司

地址 以色列,耶路撒冷

权利要求书2页 说明书27页

(72)发明人 S·艾特塔禾 M·史梅 U·盖特

序列表49页 附图9页

(54)发明名称

用于制造合成的拖丝蜘蛛丝的组合物和方法

(57)摘要

提供了包含源自Ma SP的蛋白质的混合物的组合物、编码其的核酸和制备合成拖丝蜘蛛丝的方法。本发明的组合物包含不同分子量的蛋白质的混合物,其中所述混合物的每种蛋白质独立地包含多个重复的Ma SP(大壶状腺丝蛋白)蛋白的重复区或其功能同系物、变体、衍生物或片段。

1. 组合物,其包含蛋白质混合物,所述蛋白质混合物包含m种类型的不同分子量的蛋白质,其中所述混合物中的每种蛋白质独立地包含n个重复的大壶状腺丝蛋白(MaSP)的重复区、或其功能同系物、变体、衍生物或片段,其中m和n独立地为2-70之间的整数。

2. 权利要求1所述的组合物,其中所述n对于所述混合物中的每种类型的蛋白质是相同的。

3. 权利要求1所述的组合物,其中每个重复具有2kDa至3.5kDa范围内的分子量。

4. 权利要求3所述的组合物,其中每个重复具有2.6kDa至3kDa范围内的分子量。

5. 权利要求1所述的组合物,其包含分子量增量为2kDa至3.5kDa的所述混合物的两种或更多种蛋白质。

6. 权利要求1所述的组合物,其包含分子量增量为2.6kDa至3kDa的所述混合物的两种或更多种蛋白质。

7. 权利要求1所述的组合物,其中n是等于4和32或在4和32之间的整数。

8. 权利要求1所述的组合物,其中m是等于4和32或在4和32之间的整数。

9. 权利要求1所述的组合物,其中'n'与'm'的比在2:1-1:2的范围内。

10. 权利要求1所述的组合物,其中'n'和'm'相等。

11. 权利要求1所述的组合物,其中所述重复具有ADF-4或MaSP1蛋白或其片段的重复区的同系物、变体、衍生物。

12. 权利要求1所述的组合物,其中所述重复区具有第一部分和与其邻接的第二部分,所述第一部分是包含至少50%丙氨酸残基的5-30个氨基酸的氨基酸序列,所述第二部分是包含至少80%选自甘氨酸、丝氨酸、脯氨酸和酪氨酸的残基的20-60个氨基酸的氨基酸序列。

13. 权利要求12所述的组合物,其中所述重复区的所述第二部分包含至多两个谷氨酰胺残基。

14. 权利要求1所述的组合物,其中所述重复区具有如SEQ ID NO:1所示的氨基酸序列:

(X₁) zX₂GPGGYGPX₃X₄X₅GPX₆GX₇GGX₈GPGGPGX₉X₁₀

其中X₁独立地在每种情况下为A或G,其中至少50%的(X₁) z为A,Z为5至30之间的整数;X₂为S或G;X₃为G或E;X₄为G、S或N;X₅为Q或Y;X₆为G或S;X₇为P或R;X₈为Y或Q;X₉为G或S;并且X₁₀为S或G。

15. 权利要求1所述的组合物,其中所述重复区具有如SEQ ID NO:33所示的氨基酸序列:

(X₁) zX₂SGPX₃GGYGX₄PX₅QGPX₆GGYGP

其中X₁独立地在每种情况下为A或G,其中至少50%的(X₁) z为A,Z为5至30之间的整数;X₂为S-G或不存在;X₃为G-Q或不存在;X₄为G或不存在;X₅为S或G;X₆为S-P、G-R或不存在。

16. 权利要求1所述的组合物,其中所述重复区具有如SEQ ID NO:2-4和35-44中任一个所示的氨基酸序列。

17. 权利要求1所述的组合物,其中所述重复区包含如SEQ ID NO:2 (SGPGGYGPGSQGPSGPGGYGPGPGSS)所示的氨基酸序列。

18. 权利要求1所述的组合物,其中所述重复区包含如SEQ ID NO:3 (AAAAAASGPGGYGPGSQGPSGPGGYGPGPGSS)所示的氨基酸序列。

19. 权利要求1所述的组合物,其中所述同系物与SEQ ID NO:1-3、33、35-44中的任一个共享至少70%同源性。

20. 权利要求1所述的组合物,其中所述混合物的每种蛋白还包含选自SEQ ID NO:5 (MSYYHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLV); SEQ ID NO:6 (MSYYHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLVRPLSNLDNA); SEQ ID NO:7 (MSYYHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLVDPGCRNSARAGSS)的单个N-端区域、或其任何功能同系物、变体、衍生物或片段。

21. 权利要求20所述的组合物,其中所述N-端区域的所述同系物与SEQ ID NO:5-7中的任一个共享至少70%同源性。

22. 权利要求1所述的组合物,其中所述混合物的每种蛋白质还包含如SEQ ID NO:9 (GPSGPGAYGPSASASVAASRLSSPAASSRVSSAVSSLVSSGPTNGAAVSGALNSLVSQISASNPGLSGCDALVQALLELVSALVAILSSASIGQVNVSSVSQSTQMISQALS)所示的单个C-端区域、或其任何功能同系物、变体、衍生物、片段或突变体。

23. 权利要求22所述的组合物,其中所述C-端区域的所述同系物与SEQ ID NO:9共享至少70%同源性。

24. 权利要求1所述的组合物,其中所述混合物的一种或多种蛋白质还包含至少一个标签序列。

25. 权利要求1所述的组合物,其中所述蛋白质混合物还包含ADF-3或MASP-2蛋白质。

26. 权利要求1所述的组合物,其特征在于显示至少在250℃至330℃范围内的吸热峰的DSC图谱。

27. 权利要求26所述的组合物,其中所述DSC图谱还包含在220℃至250℃范围内的吸热峰。

28. 权利要求1-27中任一项所述的组合物,其还包含载体、稀释剂或赋形剂。

29. 物品,其包含权利要求1-28中任一项所述的组合物。

30. 纤维,其包含权利要求1-28中任一项所述的组合物。

31. 分离的核酸序列,其编码权利要求1所述的蛋白质混合物中的两种或多种蛋白质。

32. 表达载体,其包含权利要求31所述的核酸序列,其中所述核酸序列处于可操作地连接的启动子和任选地调节序列的表达控制下。

33. 宿主细胞,其用根据权利要求32所述的表达载体转化。

用于制造合成的拖丝蜘蛛丝的组合物和方法

[0001] 本申请要求于2015年8月10日提交的美国临时专利申请号62/203,102,于2016年2月11日提交的美国临时专利申请号62/293,880以及于2016年4月3日提交的美国临时专利申请号62/317,572的优先权。上述文件的内容通过引用以其全部并入,如同在此完全阐述一样。

技术领域

[0002] 本发明涉及用于制备合成的拖丝蜘蛛丝(dragline spider silk)的包含源自MaSP(大壶状腺丝蛋白(major ampullate spidroin))蛋白的蛋白质混合物的组合物。

[0003] 发明背景

[0004] 拖丝蜘蛛丝在本领域中被称为由织网蜘蛛(orb-web weaving spider)在它们跌落或逃离危险时构建它们的网的框架和幅射线以及生命线所使用的丝。为了能够完成这些任务,由于高弹性和强度的组合,拖丝纤维(dragline fiber)显示出非常高的韧性,这使其成为最坚韧的纤维,无论是天然的还是人造的。例如,拖丝的直径比高强度钢强6倍,比凯夫拉(Kevlar)——其是曾经制备的最强的合成纤维之一——坚韧3倍。

[0005] 拖丝由两种主要多肽组成,通常被称为大壶状腺丝蛋白(MaSp) 1和2,在十字园蛛(*Araneus diadematus*)中也被称为ADF-3和ADF-4。这些蛋白质的表观分子量在200-720kDa范围内,这取决于样品的年龄和分析条件。已知的拖丝蜘蛛丝蛋白由交替的富含丙氨酸的片段——在纤维中形成结晶的 β -片层——和富含甘氨酸的片段——其更具柔性并且主要缺乏有序结构——的高度重复的区段组成。C-端区域不重复,在物种之间高度保守,并且采用 α -螺旋构象。拖丝蛋白的N-端区域也被发现在不同的蛛丝蛋白之间以及不同的蜘蛛物种之间高度保守。

[0006] 已经进行了许多尝试来合成地产生蜘蛛丝,如通过在组织培养中使用细菌、酵母、植物和哺乳动物细胞以及甚至转基因山羊进行基因工程。

[0007] 美国专利号8,461,301尤其涉及包含半合成蜘蛛丝蛋白结构域或其任何功能同系物、变体、衍生物、片段或突变体的多个重复的分离的氨基酸序列。该出版物通过引用以其全部并入本文。

[0008] 有关拖丝蜘蛛丝的其他出版物包括但不限于Ittah,S.,et al.*Biopolymers*,93(5),458-468,2010;Ittah,S.,et al.*Biomacromolecules*,8(9),2768-2773,2007;Ittah,S.,et al.,*Biomacromolecules*,7(6),1790-1795,2006;和Huemmerich,D.,Ittah,S.,et al.,*Current Biology*,14,2070-2074,2004。这些出版物通过引用以其全部并入本文。

[0009] 对用于生产具有类似于天然蜘蛛丝机械性能的纤维的改进组合物和方法存在未满足的需求。

发明内容

[0010] 本发明涉及包含具有不同分子量并源自MaSP蛋白的蛋白质混合物的组合物,诸如用于制造合成的拖丝蜘蛛丝。

[0011] 根据一些方面,提供了一种组合物,其包含含有m种类型的不同分子量的蛋白质的蛋白质混合物,其中所述混合物中的每种蛋白质独立地包含n个重复的源自MaSP蛋白的重复区、或其功能同系物、变体、衍生物或片段,其中m和n独立地为2-70之间的整数。

[0012] 在一些实施方式中,所述MaSP蛋白是选自MaSP-1、MaSP-2、ADF-4和ADF-3的蛋白质。在一些实施方式中,所述MaSP蛋白是选自MaSP-1和ADF-4的蛋白质。

[0013] 在一些实施方式中,所述n对于所述混合物中的每种类型的蛋白质是相同的。在另一个实施方式中,n是等于4和32或在4和32之间的整数。在另一个实施方式中,m是等于4和32或在4和32之间的整数。在另一个实施方式中,'n'与'm'的比在2:1至1:2的范围内。在另一个实施方式中,'n'和'm'是相等的。

[0014] 在另一个实施方式中,每个重复具有2kDa至3.5kDa范围内的分子量。在另一个实施方式中,每个重复序列具有2.6kDa至3kDa范围内的分子量。

[0015] 在另一个实施方式中,组合物包含分子量增量为2kDa至3.5kDa的所述混合物的两种或更多种蛋白质。在另一个实施方式中,组合物包含分子量增量为2.6kDa至3kDa的所述混合物的两种或更多种蛋白质。

[0016] 在一些实施方式中,所述重复具有MaSP蛋白或其片段的重复区的同系物、变体、衍生物。在一些实施方式中,所述重复具有MaSP1蛋白或其片段的重复区的同系物、变体、衍生物。在一些实施方式中,所述重复具有ADF-4蛋白或其片段的重复区的同系物、变体、衍生物。

[0017] 在一些实施方式中,所述重复区具有第一部分和与其邻接的第二部分,第一部分是包含至少50%丙氨酸残基的5-30个氨基酸的氨基酸序列,第二部分是包含至少80%选自甘氨酸、丝氨酸、脯氨酸和酪氨酸的残基的20-60个氨基酸的氨基酸序列。

[0018] 在一些实施方式中,所述重复区的第二部分包含至多两个谷氨酰胺残基。

[0019] 在一些实施方式中,所述重复区具有如SEQ ID NO:1所示的氨基酸序列:

[0020] $(X_1)_Z X_2 G P G G Y G P X_3 X_4 X_5 G P X_6 G X_7 G G X_8 G P G G P G X_9 X_{10}$

[0021] 其中 X_1 独立地在每种情况下为A或G,其中至少50%的 $(X_1)_Z$ 为A,Z为5至30之间的整数; X_2 为S或G; X_3 为G或E; X_4 为G、S或N; X_5 为Q或Y; X_6 为G或S; X_7 为P或R; X_8 为Y或Q; X_9 为G或S;并且 X_{10} 为S或G。

[0022] 在一些实施方式中,所述重复区包含如SEQ ID NO:33所示的氨基酸序列:

[0023] $(X_1)_Z X_2 S G P X_3 G G Y G X_4 P X_5 Q G P X_6 G G Y G P$

[0024] 其中 X_1 独立地在每种情况下为A或G,其中至少50%的 $(X_1)_Z$ 为A,Z为5至30之间的整数; X_2 为S-G或不存在; X_3 为G-Q或不存在; X_4 为G或不存在; X_5 为S或G; X_6 为S-P、G-R或不存在。

[0025] 在另一个实施方式中,所述重复区包含如SEQ ID NO:34 (PGGYGP)所示的氨基酸序列。

[0026] 在另一个实施方式中,所述重复区具有如SEQ ID NO:2-4中任一个所示的氨基酸序列。在另一个实施方式中,所述同系物与SEQ ID NO:1-3中的任一个共享至少70%的同源性。

[0027] 在另一个实施方式中,所述重复区具有如SEQ ID NO:2-4和35-44中任一个所示的氨基酸序列。在另一个实施方式中,所述同系物与SEQ ID NO:1-3、33和35-44中的任一个共享至少70%的同源性。

[0028] 在另一个实施方式中,所述重复区包含如SEQ ID NO:2 (SGPGGYGPSQGPSGPGGYPGGPGSS)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:3 (AAAAAAAAASGPGGYGPSQGPSGPGGYPGGPGSS)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:35 (SGPGGYGPSQGPSGPGGYPGGPGSSAGAGAGAAA)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:36 (SGPGGYGPSQGPSGPGGYPGGPGSSAGAGAGAGA)所示的氨基酸序列。在另一个实施方式中,所述重复区包含SEQ ID NO:35 (SGPGGYGPSQGPSGPGGYPGGPGSSAGAGAGAAA)和SEQ ID NO:36 (SGPGGYGPSQGPSGPGGYPGGPGSSAGAGAGAGA)之间1:2-1:16、1:2-1:8或1:4的比。

[0029] 在另一个实施方式中,所述重复区包含如SEQ ID NO:37 (SGSGPGGYPGGQPGGYPGGQGPYGP)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:38 (SGSGPGGYPGGQPGGYPGGQGPYGPAAAAAAAA)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:39 (GGSGPGGYPGGQPGGYPGGQGPYGP)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:40 (GGSGPGGYPGGQPGGYPGGQGPYGPAAAAAAAA)所示的氨基酸序列。在另一个实施方式中,所述重复区包含SEQ ID NO:38 (SGSGPGGYPGGQPGGYPGGQGPYGPAAAAAAAA)和SEQ ID NO:40 (GGSGPGGYPGGQPGGYPGGQGPYGPAAAAAAAA)之间1:2-1:16、1:2-1:8或1:4的比。

[0030] 在另一个实施方式中,所述重复区包含如SEQ ID NO:41 (SGPGQGGYGGPGGQGPGRGGYGPAGS)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:42 (SGPGQGGYGGPGGQGPGRGGYGPAGSAAAAAAAA)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:43 (GGPGQGGYGGPGGQGPGRGGYGPAGS)所示的氨基酸序列。在另一个实施方式中,所述重复区包含如SEQ ID NO:44 (GGPGQGGYGGPGGQGPGRGGYGPAGSAAAAAAAA)所示的氨基酸序列。在另一个实施方式中,所述重复区包含SEQ ID NO:42 (SGPGQGGYGGPGGQGPGRGGYGPAGSAAAAAAAA)和SEQ ID NO:44 (GGPGQGGYGGPGGQGPGRGGYGPAGSAAAAAAAA)之间1:2-1:16、1:2-1:8或1:4的比。

[0031] 在另一个实施方式中,所述重复区包含如SEQ ID NO:4 (AAAAAAAAASGPGGYGPENQGPSGPGGYPGGPGSS)所示的氨基酸序列。

[0032] 在另一个实施方式中,所述混合物的每种蛋白质还包含选自SEQ ID NO:5 (MSYYHHHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLV);SEQ ID NO:6 (MSYYHHHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLVRPLSNLDNA);SEQ ID NO:7 (MSYYHHHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLVDPGCRNSARAGSS)的单个N-端区域、或其任何功能同系物、变体、衍生物或片段。在另一个实施方式中,所述N-端区域的所述同系物与SEQ ID NO:5-7中的任一个共享至少70%的同源性。

[0033] 在另一个实施方式中,所述混合物的每种蛋白质还包含具有如SEQ ID NO:9 (GPSGPGAYGPSASASVAASRLSSPAASSRVSSAVSSLVSSGPTNGAAVSGALNSLVQSISASNPGLSGCDALVQALLELVSALVAILSSASIGQVNVSSVSQSTQMISQALS)所示的氨基酸序列的单个C-端区域、或其任何功能同系物、变体、衍生物、片段或突变体。在另一个实施方式中,所述C-端区域的所述同系物与SEQ ID NO:9共享至少70%的同源性。

[0034] 在一些实施方式中,所述混合物的一种或多种蛋白质还包含至少一个标签序列。

[0035] 在一些实施方式中,所述蛋白质混合物还包含ADF-3或MASP-2蛋白或其功能同系

物、变体、衍生物或片段。在一些实施方式中,所述ADF-3或MASP-2蛋白构成所述蛋白质混合物的分子量的约1-50%,或其间的任何整数。在一个实施方式中,所述ADF-3蛋白具有GenBank登录号AAC47010.1。

[0036] 在一些实施方式中,组合物还包含载体、稀释剂或赋形剂。

[0037] 根据一些方面,本发明提供了编码本发明所述蛋白质混合物的两种或更多种蛋白质的分离的核酸序列。根据一些方面,本发明提供了包含本发明的核酸序列的表达载体,其中所述核酸序列处于可操作地连接的启动子和任选地调节序列的表达控制下。根据一些方面,本发明提供了用本发明的表达载体转化的宿主细胞。

[0038] 根据一些方面,本发明提供了包含本发明的组合物的纤维。根据一些方面,本发明提供了包含本发明的组合物和/或纤维的物品。

[0039] 根据下文给出的详细描述,本发明的进一步实施方式和全部适用范围将变得明显。然而,应该理解的是,尽管指出了本发明的优选实施方式,但是详细描述和具体实施例仅仅是以示例的方式给出的,因为根据该详细描述,在本发明的精神和范围内的各种变化和修改对于本领域技术人员将变得明显。

附图说明

[0040] 图1A-I使用IMMUNO-TEM(透射电子显微镜)表征(图1A-C),采用三种不同的构建体:C1、C2和C3(分别是图1D-F)的组装纤维的光学显微镜检查和共聚焦显微镜检查(图1G-I)描绘了本发明的纳米纤维。

[0041] 图2A-C示出了使用考马斯蓝染色和蛋白质印迹分析的蛋白质重复单元的梯状图案(laddering pattern)。

[0042] 图3A-B是电纺纤维(electrospun fiber)的高分辨率扫描电子显微镜检查(HR-SEM)(3A)和差示扫描量热法(DSC)(3B)。

[0043] 图4A-B显示了嵌入透明基质(4A)中的冻干纤维和相应的DSC曲线(4B)。

[0044] 图5A-C显示了本发明纤维上的多层HEK293细胞生长。

[0045] 图6显示了描绘本发明的纤维指纹的DSC曲线。

具体实施方式

[0046] 在一些实施方式中,本发明提供包含具有不同分子量的蛋白质的混合物的组合物、编码其的核酸序列、表达载体和细胞,其可用于制备合成的拖丝蜘蛛丝。本发明还提供了包含所述组合物的物品和纤维。

[0047] 本发明部分基于意想不到的发现,即使用不同分子量并源自MaSP蛋白的蛋白质混合物合成的人造拖丝蜘蛛丝具有异常的机械性能——其类似于并且在一些性能上优于天然拖丝蜘蛛丝。

[0048] 如下文所显示的,本发明的人造拖丝蜘蛛丝显示出意想不到的热性质(参见图6)。具体而言,本发明的纤维在约265°C至320±5°C显示DSC峰,表明相对于天然拖丝蜘蛛丝有利的热性质。

[0049] 根据一些方面,提供了包含含有m种类型的不同分子量蛋白质的蛋白质混合物的组合物,其中所述混合物中的每种蛋白质独立地包含n个重复的MaSP蛋白的重复区或其功

能同系物、变体、衍生物或片段,其中m和n独立地为2至70之间的整数。

[0050] 如本文所用,术语“蛋白质的混合物”或“蛋白质混合物”是指多种蛋白质,如至少2、至少3、至少4、至少5、至少6、至少7、至少8、至少9或至少10种类型的蛋白质,其中每种类型的蛋白质具有相对独特的和均匀的分子量。如本文所用,术语“独特的”是指所述蛋白质混合物中的每种类型的蛋白质的分子量与所述混合物中的其他类型的蛋白质不同。如本文所用,术语“均匀的”是指所述蛋白质混合物中每种类型的蛋白质的分子量与所述混合物中相同类型的蛋白质至少95%相同。如本文所用,术语“相对”是指每种类型的蛋白质中不超过一个氨基酸残基的变异。

[0051] 术语“大壶状腺丝蛋白”和“蛛丝蛋白”在整个说明书中可互换使用,并且包括所有已知的大壶状腺丝蛋白,在十字园蛛的情况下通常缩写为“MaSP”或“ADF”。这些大壶状腺丝蛋白通常有两种类型,1和2。这些术语还包括如本文所公开的非天然蛋白质,其与至少已知的大壶状腺丝蛋白的重复区具有高度同一性和/或相似性。其他合适的蜘蛛丝蛋白包括MaSP2、MiSP、MiSp2、AcSP、FLYS、FLAS、和梨形(piriform)。

[0052] 如本文所用,术语“重复区”、“重复序列”或“重复”是指源自在蜘蛛丝氨基酸序列中(例如,在MaSP-1肽中)自然地发生多次的重复单元的重组蛋白质序列。本领域技术人员将会理解,蜘蛛丝蛋白的一级结构被认为主要由一系列单元重复的小的变异组成。在天然存在的蛋白质中的单元重复通常彼此不同。也就是说,沿着蛋白质长度很少或没有单元重复的确切重复。

[0053] 在一些实施方式中,制备本发明的合成蜘蛛丝,其中蛋白质的一级结构包含许多单一单元重复或直接重复的确切重复。本文使用的术语“直接重复”是具有类似重复的串联重复(头对尾排列)。在另一个实施方式中,用于形成本发明的合成蜘蛛丝的所述重复是直接重复。在一些实施方式中,所述重复在自然界中未发现(即,不是天然存在的氨基酸序列)。在一些实施方式中,所述单一单元重复(或可选地,直接重复)在其氨基酸序列内包含不多于一个变异。

[0054] 在另外的实施方式中,本发明的合成蜘蛛丝包含多个一个单元重复的重复以及多个第二个单元重复的重复。这样的结构将类似于典型的嵌段共聚物。几个不同序列的单元重复也可以组合以提供具有适用于特定应用的性质的合成蜘蛛丝蛋白。

[0055] 包含重复序列的示例性序列是ADF-4:
AAAAAASGSGGYGPENQGPSGPVAYGPGGPVSSAAAAAASGPGGYGPENQGPSGPGGYGPGGSGSSAAAAAASGPGGYGPGSQGPSGPGGSGGYGPGSQGPSGPGASSAAAAAASGPGGYGPGSQGPSGPGAYGPGGPGSSAAASGPGGYGPGSQGPSGPGGSGGYGPGSQGPSGPGGPGASAAAAAASGPGGYGPGSQGPSGPGAYGPGGPGSSAAASGPGGYGPGSQGPSGPGAYGPGGPGSSAAAAAASGPGGYGPGNQGSPGPGGYGPGGPGSSAAAAAASGPGGYGPGSQGPSGPGVYGPGGPGSSAAAAAASGPGGYGPGNQGSPGPGGYGPGGSGSSAAAAAASGPGGYGPGSQGPSGPGGSGGYGPGSQGPSGPGASSAAAAAASGPGGYGPGSQGPSGPGAYGPGGPGSSAAASGPGGYGPGSQGPSGPGAYGPGGPGSSAAAAAASGPGGYGPGSQGPSGPGGSRGYGPGSQGPGGPGASAAAAAASGPGGYGPGSQGPSGPGYQGSPGPGAYGPSASAS (SEQ ID NO:10)。

[0056] 在一些实施方式中,本发明的合成重复序列基于(例如,如下文所定义具有高百分比同一性)源自ADF-4 (SEQ ID NO:10)的一个或多个重复序列。在一些实施方式中,本发明的合成重复序列基于(例如,具有高百分比同一性)源自SEQ ID NO:18-32中任一个的一个

或多个重复序列。如本文所用,术语“基于”是指与重复序列具有高百分比的同源性的序列。如本文所用,高百分比的同源性包括以下任何一种:在ADF-4序列的特定区域上至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或99%的同一性。

[0057] 在一些实施方式中,每个重复序列包含多达60个氨基酸、多达55个氨基酸、多达50个氨基酸、多达49个氨基酸、多达48个氨基酸、多达47个氨基酸、多达46个氨基酸、多达45个氨基酸、多达44个氨基酸、多达43个氨基酸、多达42个氨基酸、多达41个氨基酸、多达40个氨基酸、多达39个氨基酸、多达38个氨基酸、多达37个氨基酸、多达36个氨基酸或多达35个氨基酸,其中可能性表示本发明的单独实施方式。在一些实施方式中,每个重复序列包含5至60个氨基酸、10至55个氨基酸、15至50个氨基酸、20至45个氨基酸、25至40个氨基酸、25至39个氨基酸或28至36个氨基酸,其中可能性表示本发明的单独实施方式。在一些实施方式中,每个重复序列包含30至40个氨基酸、31至39个氨基酸、32至38个氨基酸、33至37个氨基酸、34至36个氨基酸,其中可能性表示本发明的单独实施方式。在另外的实施方式中,每个重复序列包含35个氨基酸。

[0058] 在一些实施方式中,n是等于4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69和70中任一个的整数。

[0059] 在一些实施方式中,m是等于4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69和70中任一个的整数。

[0060] 在另一个实施方式中,'n'与'm'的比在2:1至1:2的范围内。在另一个实施例中,'n'和'm'是相等的。

[0061] 在一些实施方式中,所述'n'对于所述混合物中的每种类型的蛋白质是相同的。如本文所用,术语“n对于所述混合物中的每种类型的蛋白质是相同的”涉及对于每种类型的蛋白质即对于具有相同分子量的一种或多种蛋白质的重复序列的数目。作为非限制性实例,对于具有不同分子量的16种类型蛋白质的蛋白质混合物,每组蛋白质具有不同数量的重复序列。

[0062] 在一些实施方式中,所述混合物的各种蛋白质组在每种类型的蛋白质的重复序列数与所述组的摩尔比之间具有反比。在一些实施方式中,对于每组蛋白质(例如,具有相同数量的重复),所述蛋白质的分子量越低,所述组的摩尔比越高。

[0063] 在另一个实施方式中,每个重复具有2kDa至3.5kDa范围内、2.1kDa至3.4kDa范围内、2.2kDa至3.3kDa范围内、2.4kDa至3.2kDa范围内、2.5kDa至3.1kDa范围内、2.6kDa至3kDa范围内或2.7kDa至2.9kDa范围内的分子量,其中每种可能性表示本发明的单独实施方式。在另一个实施方式中,每个重复具有约2.8kDa范围内的分子量。

[0064] 在另一个实施方式中,组合物包含所述混合物的两种或更多种蛋白质,其分子量增量为2kDa至3.5kDa、2.1kDa至3.4kDa、2.2kDa至3.3kDa、2.4kDa至3.2kDa、2.5kDa至3.1kDa、2.6kDa至3kDa、或2.7kDa至2.9kDa,其中每种可能性代表本发明的单独实施方式。

在另一个实施方式中,组合物包含分子量增量为约2.8kDa的所述混合物的两种或更多种蛋白质。

[0065] 在一些实施方式中,所述重复区具有第一部分和第二部分,其中第一部分和第二部分是邻接的(即彼此直接相邻)。通常,第一部分和第二部分通过肽键连接。

[0066] 在一些实施方式中,所述重复区的第一部分是5-30个氨基酸的氨基酸序列,其包含至少95%、至少90%、至少85%、至少80%、至少75%、至少60%、至少55%或至少50%的丙氨酸残基。

[0067] 在一些实施方式中,第一部分可以包含一个或多个甘氨酸残基。在一些实施方式中,第一部分包含多达5%、多达10%、多达15%、多达20%、多达25%、多达30%、多达45%或多达50%的甘氨酸残基。

[0068] 在一些实施方式中,第一部分包含丙氨酸-甘氨酸二肽的1至15个重复,如在式(AG)₁₋₁₅中。

[0069] 在一些实施方式中,第一部分包含甘氨酸-丙氨酸二肽的1至15个重复,如在式(GA)₁₋₁₅中。

[0070] 在一些实施方式中,所述重复区的第二部分是20-60个氨基酸的氨基酸序列,其包含至少80%的选自甘氨酸、丝氨酸、脯氨酸和酪氨酸的残基。

[0071] 在一些实施方式中,所述重复区的第二部分是20-60个氨基酸的氨基酸序列,其包含至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%的选自甘氨酸、丝氨酸、脯氨酸和酪氨酸的残基。在一些实施方式中,所述重复区的第二部分包含不超过一个或两个谷氨酰胺残基。本领域技术人员将会理解,只要序列形成自组装纤维,重复区中的甘氨酸、丝氨酸、脯氨酸和酪氨酸残基的确切数量和顺序可以不同。

[0072] 在一些实施例中,所述重复区包含:

[0073] (i) 10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%或50%的丙氨酸残基或其间的任何范围;

[0074] (ii) 20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%或50%、51%、52%、53%、54%、55%、56%、57%、58%、59%或60%的甘氨酸残基或其间的任何范围;

[0075] (iii) 10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%或30%的丝氨酸残基或其间的任何范围;

[0076] (iv) 10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%或30%的脯氨酸残基或其间的任何范围;

[0077] (v) 1%、2%、3%、4%、5%、6%、7%、8%、9%或10%的酪氨酸残基或其间的任何范围;

[0078] (vi) 1%、2%、3%、4%、5%、6%、7%、8%、9%或10%的谷氨酰胺残基或其间的任何范围;和

[0079] (vii) 0%、1%、2%、3%、4%、5%的精氨酸残基或其间的任何范围。

[0080] 在一些实施方式中,所述重复区包含13-42%的丙氨酸残基、25-55%的甘氨酸残基、10-18%的丝氨酸残基、12-21%的脯氨酸残基、4-7%的酪氨酸残基、4-7%的谷氨酰胺残基、和0-3%的精氨酸残基。

[0081] 在一些实施方式中, MaSP1蛋白的所述重复区具有如SEQ ID NO:1所示的氨基酸序列

[0082] $(X_1)_z X_2 G P G G Y G P X_3 X_4 X_5 G P X_6 G X_7 G G X_8 G P G G P G X_9 X_{10}$

[0083] 其中 X_1 独立地在每种情况下为A或G,其中至少50%的 $(X_1)_z$ 为A, Z为5至30之间的整数; X_2 为S或G; X_3 为G或E; X_4 为G、S或N; X_5 为Q或Y; X_6 为G或S; X_7 为P或R; X_8 为Y或Q; X_9 为G或S; 并且 X_{10} 为S或G。

[0084] 在一些实施方式中,所述重复区具有如SEQ ID NO:8所示的氨基酸序列

[0085] $(X_1)_z X_2 G P G G Y G P G X_3 X_4 G P X_5 G X_6 G G X_7 G P G G P G X_8 X_9$

[0086] 其中 X_1 独立地在每种情况下为A或G,其中至少50%的 $(X_1)_z$ 为A, Z为5至30之间的整数; X_2 为S或G; X_3 为G或S; X_4 为Q或Y; X_5 为G或S; X_6 为P或R; X_7 为Y或Q; X_8 为G或S; 并且 X_9 为S或G。

[0087] 在一些实施方式中, (SEQ ID NO:1或8的) Z是6至11之间的整数、6至10之间的整数或7至9之间的整数。在一个实施方式中, Z是选自5、6、7、8、9、10、11和12的整数。在另一个实施方式中, Z是8。

[0088] 在另一个实施方式中, MaSP1蛋白的所述重复区包含如SEQ ID NO:2 (SGPGGYGPGSQGPSGPGGYGPGGPGSS) 所示的氨基酸序列。在另一个实施方式中, MaSP1蛋白的所述重复区包含如SEQ ID NO:3 (AAAAAAAASGPGGYGPGSQGPSGPGGYGPGGPGSS) 所示的氨基酸序列。

[0089] 在另一个实施方式中,所述同系物与SEQ ID NO:1共享至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的同源性。

[0090] 在另一个实施方式中,所述同系物与SEQ ID NO:2共享至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的同源性。

[0091] 在另一个实施方式中,所述同系物与SEQ ID NO:3共享至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%的同源性。

[0092] 在另一个实施方式中, MaSP1蛋白的所述重复区包含如SEQ ID NO:4所示的氨基酸序列。在另一个实施方式中, MaSP1蛋白的所述重复区具有如SEQ ID NO:10所示的氨基酸序列。

[0093] 在另一个实施方式中,所述混合物的每种蛋白质还包含选自SEQ ID NO:5 (MSYYHHHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLV); SEQ ID NO:6 (MSYYHHHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLVRPLSNLDNA); SEQ ID NO:7 (MSYYHHHHHHHDYDIPTTENLYFQGAMDPEFKGLRRRAQLVDPGCRNSARAGSS) 的单个N-端区域、或其任何功能同系物、变体、衍生物或片段。在另一个实施方式中,所述C-端区域的所述同系物与SEQ ID NO:5-7中的任一个共享至少70%的同源性。

[0094] 在另一个实施方式中,所述混合物的每种蛋白质还包含单个C-端区域SEQ ID NO: 9

[0095]

(GPSGPGAYGPSPSASASVAASRLSSPAASSRVSSAVSSLVSSGPTNGAAVSGALNSLVSQISASNPGLSGCDALVQ ALLELV SALVAILSSASIGQVNVSSVSQSTQMISQALS)、或其任何功能同系物、变体、衍生物、片段或突变体。在另一个实施方式中,所述N-端区域的所述同系物与SEQ ID NO:9共享至少70%的同源性。

[0096] 在一些实施方式中,所述混合物的一种或多种蛋白质还包含至少一个标签序列。可用于本发明的标签的非限制性例子包括His标签、HA标签、T7标签等。示例性的His标签包含如SEQ ID NO:11 (HHHHHH) 所示六个His残基或由如SEQ ID NO:11 (HHHHHH) 所示六个His残基组成。在另一个实施方式中,标签是包含如SEQ ID NO:12 (YPYDVPDYA) 所示氨基酸序列或由其组成的HA标签。在另一个实施方式中,标签是包含如SEQ ID NO:13 (MASMTGGQQMG) 所示氨基酸序列或由其组成的T7标签。技术人员非常清楚可选的合适标签或其他融合伴侣。

[0097] 如本文所用,“氨基酸”是指天然存在的和合成的氨基酸,以及以与天然存在的氨基酸类似的方式起作用的氨基酸类似物和氨基酸模拟物。天然存在的氨基酸是那些由遗传密码编码的氨基酸,以及随后修饰的那些氨基酸,例如羟脯氨酸、 γ -羧基谷氨酸和O-磷酸丝氨酸。“氨基酸类似物”是指具有与天然存在的氨基酸相同的基本化学结构(即与氢、羧基、氨基和R基团结合的 α 碳)的化合物,例如高丝氨酸、正亮氨酸、甲硫氨酸亚砷、甲硫氨酸甲基砷。此类类似物具有修饰的R基团或修饰的肽骨架,但保留与天然存在的氨基酸相同的基本化学结构。“氨基酸模拟物”是指具有与氨基酸的一般化学结构不同的结构,但以与天然存在的氨基酸类似的方式起作用的化学化合物。氨基酸在本文中可以通过其公知的三字母符号或由IUPAC-IUB生物化学命名委员会推荐的单字母符号来表示。

[0098] “氨基酸序列”或“肽序列”是通过肽键连接的氨基酸残基位于肽和蛋白质链中的顺序。该序列通常从含有游离氨基的N-端到含有游离羧基的C-端报道。如果氨基酸序列代表蛋白质的一级结构,则该序列通常称为肽、蛋白质序列,然而必须辨别术语“氨基酸序列”或“肽序列”和“蛋白质”,因为蛋白质被定义为折叠成特定三维构型并且通常经历了翻译后修饰,如磷酸化、乙酰化、糖基化、巯基键形成、裂解等的氨基酸序列。

[0099] 如本文所用,在本发明所例示的合成蜘蛛丝氨基酸序列或编码其的核酸分子的上下文中,“分离的”或“基本上纯化的”是指氨基酸序列或多核苷酸已从它们的天然环境移除或已经从它们的自然状态改变。这样的“分离的”不一定反映氨基酸序列或核酸分子已被纯化的程度。然而,应该理解的是,已经被纯化到一定程度的这种分子是“分离的”。如果所述分子不存在于天然环境中,即它不存在于自然界中,则无论分子存在于何处,该分子都是“分离的”。举例来说,甚至当氨基酸序列或多核苷酸存在于人类中时,没有天然存在于人类中的氨基酸序列或多核苷酸也是“分离的”。

[0100] 当应用于氨基酸序列或核酸时,术语“分离的”或“基本上纯化的”表示氨基酸序列或核酸基本上不含天然状态中它们结合的其他细胞组分。它可以处于均一状态,或者可选地在干燥或水溶液中。通常使用分析化学技术如聚丙烯酰胺凝胶电泳或高效液相色谱来确定纯度和均一性。存在于制剂中的主要种类的氨基酸序列或核酸是基本纯化的。

[0101] 在一些实施方式中,所述重复具有MaSP1蛋白或其片段的重复区的同系物、变体、

衍生物。在一些实施方式中,所述重复具有ADF-4蛋白或其片段的重复区的同系物、变体、衍生物。在一个示例性实施方式中,所述重复包含SEQ ID NO:45-47中任一个所示的多核苷酸序列或由其组成。

[0102] 如本文所用,如在“功能同系物、变体、衍生物或片段”中的术语“功能”是指具有通过限定的功能测定而鉴定的生物学功能或活性的氨基酸序列。更具体地,限定的功能测定是在表达所述功能同系物、变体、衍生物或片段的细胞中形成自组装纤维。

[0103] 当同源性被确定为至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少92%、至少94%、至少96%、至少98%或至少99%时,氨基酸序列或核酸序列被认为是相应氨基酸序列或核酸的同系物。

[0104] 如本文所用,可以基于两个氨基酸(肽)或DNA序列之间的百分比同一性来确定同源性。一般来说,将要比较的两个序列进行比对以给出序列之间的最大相关性。检查两个序列的比对,确定两个序列之间给出确切氨基酸(或核苷酸)对应的位置数量,除以比对的总长度,乘以100,得到百分比同一性数字。该百分比同一性数字可以在待比较的序列的全长上确定——这特别适合于长度相同或非常相似并且高度同源的序列,或者在更短的限定长度上确定——这更适合于长度不等或具有较低同源性水平的序列。用于比较两个或更多个序列的同一性的方法在本领域中是公知的。因此,例如,威斯康星序列分析包版本9.1中可用的程序,例如程序GAP和BESTFIT,可用于确定两个氨基酸序列之间的百分比同一性和两个多核苷酸序列之间的百分比同一性。BESTFIT使用Smith和Waterman的“局部同源性”算法,找到两个序列之间最好的单一相似区域。BESTFIT更适合于比较长度不同的两个多肽或两个多核苷酸序列,程序假定较短的序列代表较长的一部分。相比之下,根据Needleman和Wunsch的算法,GAP比对找到“最大相似性”的两个序列。GAP更适合于比较大体相同长度的序列,并预期整个长度上的比对。优选地,在每个程序中使用的参数“空位权重”和“长度权重”分别对于多核苷酸序列是50和3,对于多肽序列是12和4。优选地,当被比较的两个序列被最佳比对时确定百分比同一性和相似性。

[0105] 在两个或更多个氨基酸或核酸序列的上下文中,术语“同一”、“基本同一性”、“基本同源性”或百分比“同一性”是指相同或具有特定百分比的相同氨基酸残基或核苷酸的两个或更多个序列或子序列(即,当在比较窗口或指定区域上针对最大对应而进行比较和比对时,在指定区域(例如氨基酸序列SEQ ID NO:2或3)上约60%同一性,或至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或99%同一性),如使用具有下述默认参数的BLAST或BLAST 2.0序列比较算法,或通过手动比对和视觉检查所测量的。然后这样的序列被认为是“基本同一的”。该定义还涉及或可以应用于测试序列的互补序列(compliment)。该定义还包括具有缺失和/或添加的序列,以及具有取代的序列。优选的算法可以考虑空位等。

[0106] 对于序列比较,通常一个序列作为参考序列,测试序列与其进行比较。当使用序列比较算法时,将测试序列和参考序列输入计算机,必要时指定子序列坐标,并指定序列算法程序参数。优选地,可以使用默认程序参数,或者可以指定可选参数。然后序列比较算法基于程序参数计算测试序列相对于参考序列的序列同一性百分比。

[0107] 应该理解的是,本发明进一步包括包含n个重复的SEQ ID NO:1、2或3任一个变体

的氨基酸序列。如本文所用,术语“变体”或“基本上相似的”包含不同于具体鉴定的序列的氨基酸或核苷酸序列,其中一个或多个(例如1、2、3、4、5、6、7、8、9、10、20或25个)氨基酸残基或核苷酸缺失、取代或添加。变体可以是天然存在的等位基因变体或非天然来源的变体。变体或基本上相似的序列是指氨基酸序列或核酸的片段,其可以通过其氨基酸或核苷酸序列与本文所述的氨基酸或核苷酸序列的同一性百分比来表征,如通过在现有技术中使用的常用算法所确定的。氨基酸或核酸的优选片段是当与参考序列比较时具有至少约40或45%的序列同一性,优选约50%或55%的序列同一性,更优选约60%或65%的序列同一性,更优选约70%或75%的序列同一性,更优选约80%或85%的序列同一性,更优选约90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性的氨基酸或核苷酸序列的那些片段。

[0108] 本文使用的术语衍生物和功能衍生物是指具有任何插入、缺失、取代和修饰的本发明的氨基酸序列。

[0109] 应该理解的是,术语“插入”如本文所用,意指在本发明的序列中添加1至50个氨基酸残基,具体地20至1个氨基酸残基,更具体地,1至10个氨基酸残基的氨基酸残基。最具体地,1、2、3、4、5、6、7、8、9和10个氨基酸残基。此外,本发明的氨基酸序列可以在其N-端和/或C-端用各种相同或不同的氨基酸残基延伸。

[0110] 氨基酸“取代”是用具有相似结构和/或化学性质的另一个氨基酸取代一个氨基酸的结果,即保守氨基酸置换。氨基酸取代可以基于所涉及残基的极性、电荷、溶解度、疏水性、亲水性和/或两亲性的相似性来进行。例如,非极性(疏水)氨基酸包括丙氨酸、亮氨酸、异亮氨酸、缬氨酸、脯氨酸、苯丙氨酸、色氨酸和甲硫氨酸;极性中性氨基酸包括甘氨酸、丝氨酸、苏氨酸、半胱氨酸、酪氨酸、天冬酰胺和谷氨酰胺;带正电的(碱性)氨基酸包括精氨酸、赖氨酸和组氨酸;和带负电的(酸性)氨基酸包括天冬氨酸和谷氨酸。

[0111] 在另一个实施方式中,本发明的重复序列具有17个或更少、16个或更少、15个或更少、14个或更少、13个或更少、12个或更少、11个或更少、10个或更少、9个或更少、8个或更少、或者7个或更少的对SEQ ID NO:2或3中任一个序列的氨基酸取代。在一个实施方式中,本发明的重复序列具有至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个或至少13个对SEQ ID NO:2或3中任一个序列的氨基酸取代。

[0112] 就氨基酸序列而言,技术人员将认识到对氨基酸、核酸、肽、多肽或蛋白质序列的个别取代、缺失或添加——其改变、添加或缺失编码序列中的单个氨基酸或少部分氨基酸——是“保守性修饰的变体”,其中所述改变导致用化学上相似的氨基酸取代氨基酸。提供功能上相似的氨基酸的保守取代表是本领域众所周知的。这些保守修饰的变体除了并且不排除本发明的多态变体、种间同系物和等位基因。

[0113] 例如,可以进行取代,其中脂肪族氨基酸(G、A、I、L或V)用该组的另一个成员取代,或取代如一个极性残基取代另一个极性残基如精氨酸取代赖氨酸,谷氨酸取代天冬氨酸,或谷氨酰胺取代天冬酰胺。以下八组中的每组含有其他示例性氨基酸,这些氨基酸是彼此的保守取代:1) 丙氨酸(A)、甘氨酸(G);2) 天冬氨酸(D)、谷氨酸(E);3) 天冬酰胺(N)、谷氨酰胺(Q);4) 精氨酸(R)、赖氨酸(K);5) 异亮氨酸(I)、亮氨酸(L)、甲硫氨酸(M)、缬氨酸(V);6) 苯丙氨酸(F)、酪氨酸(Y)、色氨酸(W);7) 丝氨酸(S)、苏氨酸(T);和8) 半胱氨酸(C)、甲硫氨

酸(M)。

[0114] 保守核酸取代是导致如上所定义的保守氨基酸取代的核酸取代。

[0115] 本发明的氨基酸序列的变体可具有与由SEQ ID NO:2或3中的任一个表示的重复单元在氨基酸水平上至少80%序列相似性、至少85%序列相似性、90%序列相似性或至少95%、96%、97%、98%或99%序列相似性。

[0116] 本发明的氨基酸序列可以包含n个重复的SEQ ID NO.1或SEQ ID NO.3或其任何片段。“片段”构成特定区域的氨基酸或DNA序列的一部分。肽序列的片段比特定区域短至少一个氨基酸,并且DNA序列的片段比特定区域短至少一个碱基对。该片段可以在C-端或N-端或两端截短。氨基酸片段可以包含SEQ ID NO.1或3的至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12、至少13、至少14、至少15、至少16、至少17、至少18、至少19、至少20、至少21、至少22、至少23、至少24、至少24、至少26、至少27、至少28、至少29、至少30、至少31、至少32、至少33或至少34个氨基酸。

[0117] 本发明氨基酸序列的突变体的特征在于其一个(点突变体)或更多,约多达10个氨基酸与一个或多个另一个氨基酸的交换。它们是DNA水平上相应突变——导致不同密码子——的结果。

[0118] 另外,本发明涉及本发明氨基酸序列的衍生物。本发明的氨基酸序列的衍生物是,例如,其中官能团如氨基、羟基、巯基或羧基被衍生化,例如分别被糖基化、酰化、酰胺化或酯化。在糖基化衍生物中,寡糖通常与天冬酰胺、丝氨酸、苏氨酸和/或赖氨酸连接。酰化衍生物特别是被天然存在的有机或无机酸如乙酸、磷酸或硫酸酰化,其通常分别发生在N-端氨基处或羟基处,特别是酪氨酸或丝氨酸的N-端氨基处或羟基处。酯是天然存在的醇例如甲醇或乙醇的酯。其他衍生物是盐,尤其是药学上可接受的盐,例如金属盐,如碱金属和碱土金属盐,例如钠盐、钾盐、镁盐、钙盐或锌盐,或与氨或合适的有机胺,如低级烷基胺例如三乙胺、羟基-低级烷基胺例如2-羟乙胺等形成的铵盐。

[0119] 根据一些方面,本发明提供了编码本发明所述蛋白质混合物的两种或更多种蛋白质的分离的核酸序列。根据一些实施方式,本发明提供了编码本发明蛋白质混合物的分离的核酸序列。

[0120] “核酸”是指可以是单链或双链的,由含有糖、磷酸和嘌呤或嘧啶的单体(核苷酸)组成的分子。在细菌、低等真核生物以及高等动物和植物中,“脱氧核糖核酸”(DNA)是指遗传物质,而“核糖核酸”(RNA)参与将信息从DNA翻译为蛋白质。

[0121] 由于遗传密码的简并性质,清楚的是可以使用多种不同的核酸序列来编码本发明的氨基酸序列。应该理解的是,包含在本发明的核酸序列中的密码子可以针对在Sf9宿主细胞中的表达进行优化。

[0122] 当指用于转化各种宿主的核酸分子的基因或编码区时,术语“密码子优化”指核酸分子的基因或编码区中的密码子的改变,以反映宿主生物的典型密码子使用而不改变DNA编码的多肽。在本发明的上下文内,基因和DNA编码区经密码子优化以在宿主细胞中以及在特定实例中,在Sf9草地贪夜蛾(*Spodoptera frugiperda*)昆虫细胞中最佳表达。

[0123] 本文使用的术语“表达”意指从编码基因产物序列的基因转录和翻译成基因产物。在该表达中,首先将编码基因产物序列的DNA链转录为通常为信使RNA的互补RNA,然后,如果基因产物是蛋白质,则将如此转录的信使RNA翻译成上述基因产物。

[0124] 在一些实施方式中,本发明涉及一种或多种包含编码本发明的蛋白质混合物的核酸序列的表达载体。在一些实施方式中,本发明涉及一种或多种表达载体,其包含编码本发明的蛋白质混合物的至少一部分(例如,两组或更多组具有不同分子量的蛋白质)的核酸序列。由包含在本发明的表达载体内的核酸序列编码的氨基酸序列可以任选地进一步包含C-端区域(例如,表示为SEQ ID NO:9);和N-端区域(例如,选自SEQ ID NO:5-7)中的至少一个。应该注意的是,核酸序列处于可操作连接的启动子和任选地调节序列的表达控制之下。

[0125] 如本文所用,本文所指的“载体”、“表达载体”或“质粒”是一种染色体外元件,通常携带不是细胞中心代谢的部分的基因,并且通常是环状双-链DNA分子的形式。它可以是许多核酸中的任何核酸,通过限制和连接可将期望的序列插入其中,以便在不同遗传环境之间转运或在宿主细胞中表达。载体通常由DNA组成,尽管RNA载体也是可用的。载体包括但不限于质粒和噬菌粒。克隆载体是能够在宿主细胞中复制的载体,并且其特征还在于一个或多个内切核酸酶限制位点,在该限制位点可以以可测定的方式切割载体,并且可以将期望的DNA序列连接到其中,使得新的重组载体保留其在宿主细胞中复制的能力。在质粒的情况下,当质粒在宿主细菌内的拷贝数增加时期望的序列的复制可以发生多次或者在宿主通过有丝分裂繁殖之前每个宿主期望的序列的复制可以仅发生一次。在噬菌体的情况下,复制可以在裂解阶段期间主动地发生或在溶源阶段期间被动地发生。表达载体是可通过限制和连接将期望的DNA序列插入其中的表达载体,使得其可与调节序列可操作地连接并可表达为RNA转录物。载体可以进一步含有一种或多种适用于鉴定和选择用载体已转化或转染的细胞的标记序列。如本文所用,“转化”或“转染”是通过掺入核酸获得细胞中的新基因。标记包括例如编码增加或减少对抗生素或其他化合物的抗性或敏感性的蛋白质的基因,编码其活性可通过本领域已知的标准测定可检测的酶(例如 β -半乳糖苷酶或碱性磷酸酶)的基因,和明显影响转化或转染的细胞、宿主、集落或噬斑的表型的基因。优选的载体是能够自主复制和表达存在于它们可操作地连接的DNA区段中的结构基因产物,即表达合成蜘蛛丝蛋白的载体。

[0126] 在具体的实施方式中,载体是病毒载体,最具体地是杆状病毒载体系统或痘苗病毒载体系统。这种可商购的杆状病毒系统的例子Baculo-Gold®、Flash-Bac®和bac to bac系统。另外的病毒载体系统也可以用于本发明。在个案的情况下,可能需要对载体进行修饰。另外的病毒载体的例子是腺病毒和所有负链RNA病毒,例如,狂犬病、麻疹、RSV等。

[0127] 在一个实施方式中,杆状病毒系统用于表达本发明的合成丝蛋白。杆状病毒是可分为两个属:核多角体病毒和颗粒病毒的大型杆状病毒家族。它们具有有限的它们可以感染的宿主范围,通常限于有限数量的密切相关的昆虫种类。因为杆状病毒对人体无害,所以它们是用于研究和商业或工业应用的安全选择。昆虫细胞中的杆状病毒表达代表生产重组糖蛋白的稳健方法,相对于原核表达的重要优势,其缺乏糖基化,因此蛋白质折叠恰当。

[0128] 如上所述,本发明的表达载体可操作地连接至启动子。术语“启动子”和“启动子区”是指通常在结构基因的蛋白质编码序列的上游(5')的DNA序列,其通过提供对RNA聚合酶和/或转录在正确位置开始所需的其他因子的识别来控制编码区的表达。启动子序列是必要的,但并不总是足以驱动基因的表达。术语“合适的启动子”将指能够驱动合成蜘蛛丝变异基因表达的任何真核或原核启动子。

[0129] 用于驱动Sf9中异源DNA片段表达的启动子是众多的且本领域技术人员熟悉的。几

乎任何能够驱动编码丝变异蛋白的基因的启动子都适用于本发明。例如,多角体蛋白、碱性蛋白质、p10、OpIE2和gp4启动子可以是用于所述表达的合适启动子。

[0130] 当编码序列和调节序列共价连接时,编码序列和调节序列被称为“可操作地连接”或“可操作地结合”,从而将编码序列的表达或转录置于调节序列的影响或控制下。如果调节序列相对于基因定位使得调节序列能够对所产生的基因产物的量施加可测量的影响,那么调节序列与基因可操作地连接。如果期望将编码序列翻译成功能蛋白质,则两个DNA序列被称为可操作地结合,如果5'调节序列中的启动子的诱导导致编码序列的转录并且如果两个DNA序列之间连接的性质不(1)导致引入移码突变,(2)干扰启动子区指导编码序列转录的能力,或(3)干扰相应的RNA转录物被翻译成蛋白质。因此,如果启动子区能够影响该DNA序列的转录,则启动子区将可操作地结合至编码序列,使得所得转录物可被翻译成期望的蛋白质或多肽。

[0131] 基因表达所需的调节序列的确切性质可以在物种或细胞类型之间变化,但是按照需要,通常应该包括分别涉及转录和翻译起始的5'非转录和5'非翻译序列,如TATA盒、加帽序列、CAAT序列。特别地,这种5'非转录调节序列将包括启动子区,其包括用于可操作地结合的基因的转录控制的启动子序列。根据需要,调节序列还可以包括增强子序列或上游激活子序列。

[0132] “调节(Regulation)”和“调节(regulate)”是指由主要但不唯一位于基因转录起点上游(5')的DNA序列元件控制的基因表达的调节。调节可能导致对刺激的全或无响应,或者可能导致基因表达水平的变化。

[0133] 另一方面,本发明提供了用根据本发明的表达载体转化的宿主细胞。

[0134] “细胞”、“宿主细胞”或“重组的宿主细胞”是本文可互换使用的术语。应该理解的是,这样的术语不仅指特定的受试细胞,而且指这样的细胞的后代或潜在后代。因为某些修饰可能由于突变或环境影响而在后代中发生,所以这样的后代事实上可能不与亲本细胞相同,但仍包括在本文所用的术语的范围内。

[0135] 如本文所用,“宿主细胞”是指可以用裸DNA或使用重组DNA技术构建的表达载体重组转化的细胞。抗药性或其他选择性标记部分地旨在便于选择转化体。此外,选择性标记如抗药性标记的存在可用于防止污染性微生物在培养基中繁殖。转化的宿主细胞的这种纯培养物将通过在需要诱导的存活表型的条件下培养细胞来获得。

[0136] 本发明的宿主细胞用本文描述的表达载体转化或转染以表达本发明的合成蜘蛛丝蛋白。如本文所用,“转化”是指细胞的基因型由于细胞摄入外源DNA或RNA而改变的过程,并且例如转化的细胞表达重组形式的期望的合成蜘蛛丝蛋白。术语“转染”意指通过核酸介导的基因转移将核酸,例如裸DNA或表达载体引入受体细胞。

[0137] 在一个具体的实施方式中,用根据本发明的表达载体转化的宿主细胞是昆虫细胞。作为昆虫细胞,可以使用鳞翅目昆虫细胞,更具体地,可以使用来自草地贪夜蛾和来自粉纹夜蛾(*Trichoplusia ni*)的细胞。最具体地,昆虫细胞是Sf9、Sf21或high 5细胞。

[0138] 在一些实施方式中,本发明的丝蛋白缺乏翻译后修饰。

[0139] 在一些实施方式中,本发明的丝蛋白是可生物降解的。例如,在医学领域中,每当丝蛋白打算用于体内使用——其中期望生物降解——时,该特征可能具有重要性。该特征可能特别适用于缝合材料和伤口闭合和覆盖系统。

[0140] 根据一些方面,本发明提供了包含本发明的核酸序列的表达载体,其中所述核酸序列处于可操作地连接的启动子和任选地调节序列的表达控制下。

[0141] 纤维

[0142] 根据一些方面,本发明提供了包含本发明的组合物的纤维。

[0143] 如本文所用,“纤维”是指由两根或更多根捻合在一起的细丝组成的纤维材料的细线。“细丝”是指细长的、伸长的、线状的物体或长度不定的结构,范围从微观长度到英里或更大的长度。具体地,合成蜘蛛丝细丝是微观的,并且是蛋白质的。“生物丝(biofilament)”是指由蛋白质产生的细丝,包括重组产生的蜘蛛丝蛋白。术语“纤维”不包含非结构化聚集体或沉淀物。

[0144] 在一些实施方式中,纤维具有至少50nm的厚度直径。在一些实施方式中,纤维具有至多350nm的厚度直径。在一些实施方式中,纤维具有至少50-350nm或其间的任何数值的厚度直径。如本文所证明的(参见图1),最终的纤维由直径为5-10nm的纳米纤维构成。

[0145] 在一些实施方式中,与其厚度相比,纤维具有相当大的延伸长度,优选20 μ m以上。

[0146] “微纤维”是指细度小于1旦尼尔的细丝(旦尼尔被定义为每9000米的质量克数)。

[0147] 在一些实施方式中,蛋白质的纤维的特征在于其至少一个维度(例如直径,长度)的大小。

[0148] 例如但没有限制,纤维的直径在10nm-1 μ m、20-100nm或10-50nm之间。

[0149] 在一些实施方式中,纤维由纳米原纤维构成。在一些实施方式中,纳米原纤维的直径为例如1nm、约2nm、约3nm、约4nm、约5nm、约6nm、约7nm、约8nm、约9nm、约10nm、约11nm、约12nm、约13nm、约14nm、约15nm、约16nm、约17nm、约18nm、约19nm、约20nm、约21nm、约22nm、约23nm、约24nm、约25nm、约26nm、约27nm、约28nm、约29nm、约30nm、约31nm、约32nm、约33nm、约34nm、约35nm、约36nm、约37nm、约38nm、约40nm、约42nm、约44nm、约46nm、约48nm或50nm,包括其间的任何值或范围。在一个实施方式中,纳米原纤维具有3-7nm的直径。在一个实施方式中,纳米原纤维具有4-6nm的直径。

[0150] 在一些实施方式中,所公开的纤维的长度在1-200 μ m、10-100 μ m、100-500 μ m或200-500 μ m之间。

[0151] 通过一个实施方式,本发明的纤维通过自组装而组装。“自组装”是指所述纤维的单体即本发明的合成蜘蛛丝蛋白在正常生理条件下或在室温下以能量有利的方式自发地彼此结合以产生具有本文所述性质的大分子纤维结构。此外,本发明的纤维非常有弹性,并且一经组装就可承受极端的化学攻击,如在10%SDS中溶解并煮沸至少1小时。

[0152] “韧度(tenacity)”或“拉伸强度”是指细丝在断裂之前可承受的重量。所产生的最大比应力通常是在细丝、纱线或织物中通过拉伸试验来断裂材料。根据具体实施方式,本发明的纤维具有约100-3000MPa(MPa=N/mm²)、约300-3000MPa、约500-2700MPa、约700-2500MPa、约900-2300MPa、约1100-2000MPa、约1200-1800MPa、约1300-1700MPa或约1400-1600MPa的拉伸强度。更具体地,约1500MPa。

[0153] “韧性(toughness)”是指断裂纤维所需的能量。这是应力应变曲线下的面积,有时称为“断裂能量”或破裂功。根据具体的实施方式,本发明的纤维的韧性为约20-1000MJ/m³、约50-950MJ/m³、约100-900MJ/m³、约120-850MJ/m³、约150-800MJ/m³、约180-700MJ/m³、约180-750MJ/m³、约250-700MJ/m³、约280-600MJ/m³、约300-580MJ/m³、约310-560MJ/m³、约

320–540MJ/m³或约350–520MJ/m³,最具体地约350–520MJ/m³。

[0154] “弹性”是指在变形后物体倾向于恢复其原始尺寸和形状的性质。塑性,变形而不恢复,与弹性相反。在纤维的分子构型上,通过原子间和分子间结构键的拉伸(再定向),可恢复或弹性变形是可能的。相反,将分子间键断裂并重新形成新的稳定位置导致不可恢复或塑性变形。

[0155] “延伸”是指以初始长度的百分比或分数表示的长度增加。

[0156] “细度”是指纤维或细丝(例如生物丝)的平均直径,其通常以微米表示。

[0157] 在一些实施方式中,公开的组合物通过限定的差示扫描量热法(DSC)图谱来表征。在一些实施方式中,“DSC图谱(DSC pattern)”意指峰的位置。在一些实施方式中,“峰”意指放热峰。遍及本文,“峰的位置”或“峰位置”是指温谱图图谱(thermogram pattern)中沿温度轴的峰,并且在一些实施方式中,可以指任何峰强度处的峰位置。本领域技术人员将会理解,在DSC测量中获得的数据部分取决于所使用的仪器和进行测量时的环境条件(例如湿度)。

[0158] 在一些实施方式中,公开的组合物通过显示至少在250℃至330℃范围内的吸热峰的DSC图谱表征。在一些实施方式中,公开的组合物通过显示至少在290℃至310℃范围内的吸热峰的DSC图谱表征。在一些实施方式中,公开的组合物通过显示至少在295℃至305℃范围内的吸热峰的DSC图谱表征。

[0159] 在一些实施方式中,公开的组合物通过显示至少在260℃至320℃和220℃至250℃范围内的吸热峰的DSC图谱表征。在一些实施方式中,公开的组合物通过在120℃至160℃范围内的另外的DSC放热峰来进一步表征。

[0160] 在一些实施方式中,公开的组合物没有在约–100℃至约220℃范围内的DSC峰。在一些实施方式中,公开的化合物没有在约–100℃至约25℃范围内的DSC峰。在一些实施方式中,公开的组合物通过至少显示在40℃至70℃范围内没有放热峰的DSC图谱表征。

[0161] 在一些实施方式中,公开的化合物没有在约–100℃至约–50℃范围内的DSC峰。在一些实施方式中,公开的化合物没有在约–50℃至约0℃范围内的DSC峰。在一些实施方式中,公开的化合物没有在约–0℃至约–25℃范围内的DSC峰。

[0162] 组合物

[0163] 根据具体的实施方式,组合物可以以凝胶、泡沫或用于涂覆支架和植入物的涂层的形式或以用于组织工程目的的形式提供。在其他实施方式中,本发明的组合物是药物组合物。

[0164] 应该注意,本发明的药物组合物可以包含本发明的氨基酸序列、重组蛋白和纤维中的至少一种,并且可以直接给予待治疗的受试者。制剂通常包含至少一种如上定义的活性成分以及一种或多种其可接受的载体。

[0165] 制剂特别适用于局部给药,或用作侵入性医疗设备的涂层或用作组织工程的支架,然而皮下、皮内、肌肉、腹膜内、静脉内和甚至口服、直肠、鼻或肠胃外给药途径不被忽视。

[0166] 用于局部给药的药物组合物和制剂可以包括透皮贴剂、软膏、洗剂、霜剂、凝胶剂、滴剂、栓剂、喷雾剂、液体和粉剂。常规的药物载体、水性、粉末或油性基质、增稠剂等可能是必要的或需要的。

[0167] 本发明的药物组合物通常包含缓冲剂、调节其渗透性的剂和任选地本领域已知的一种或多种药学上可接受的载体、赋形剂和/或添加剂。补充活性成分也可以掺入组合物中。载体可以是含有例如水、乙醇、多元醇(例如甘油、丙二醇和液体聚乙二醇等)、其合适的混合物、和植物油的溶剂或分散介质。可以例如,通过使用诸如卵磷脂的涂层,通过在分散的情况下维持所需的粒度和通过使用表面活性剂来维持适当的流动性。

[0168] 如本文所用,“药学上可接受的载体”包括任何和所有溶剂、分散介质、涂层、抗菌剂和抗真菌剂等。此类介质和剂用于药物活性物质的用途在本领域中是公知的。除了任何常规介质或剂与活性成分不相容之外,还考虑其在治疗组合物中的用途。

[0169] 包含本发明的氨基酸序列、重组蛋白和纤维中的至少一种的本发明的药物组合物可以特别适合于生产透皮贴剂即透皮递送系统、弹性伤口敷料、缝合线、涂层或医用织物——其需要弹力、弹性、韧度且无毒。

[0170] 在又一些实施方式中,本发明的组合物可以是化妆品组合物。术语“化妆品组合物”涉及具有有益的皮肤或其他表面组织美学特性如改善或增强皮肤色调和颜色、头发颜色和光泽、隐藏表面组织缺陷如瑕疵和疤痕、或防止未来或累积的损伤如阳光损伤和皮肤老化的组合物。

[0171] 根据本发明的用于治疗皮肤病学或化妆品组合物作为软膏发膏剂、洗剂、霜剂和凝胶剂局部施加于表皮上以及作为水乳剂如乳膏、洗剂或凝胶剂局部施加于粘膜上。化妆品可以使用这样的组合物来生产,包括产品如剃须膏、护手霜、香波、肥皂、调理剂(conditioner)、身体乳霜、皮肤防晒(sun skin-protection)、面霜或身体乳液。根据本发明的化妆品组合物中组分的比可根据化妆品组合物的预期应用进行调节。

[0172] 组织支架(Tissue scaffolding)

[0173] 另一方面,本发明提供组合物,所述组合物包含(i)细胞;和(ii)包含至少一种本发明的纤维的细胞支架材料。

[0174] 另一方面,本发明提供培养细胞的方法,该方法包括:

[0175] 提供待培养的细胞样品;

[0176] 将所述样品施加到包含至少一种本发明的纤维的细胞支架材料;和

[0177] 在适于细胞培养的条件下保持所述细胞支架材料——具有施加于其上的细胞。

[0178] 在本发明的上下文中,术语细胞的“培养”、“细胞培养”应被广泛地解释,使得它们包括例如细胞分裂和/或增殖的情况、细胞维持在分化状态——保留在其天然环境中存在时由细胞类型表现的至少一种功能特征——的情况、以及干细胞维持在未分化状态的情况。

[0179] 在一些实施方式中,培养方法或细胞组合物可以在包含含有特定生长因子或细胞外基质(ECM)组分的细胞培养和/或培养基的条件下进行。在其他实施方式中,培养或制备方法可以在包含将施加有细胞的细胞支架材料保持在无血清培养基中的条件下进行。在无血清培养基中培养细胞的可能性提供了对含血清培养基和/或含有特定生长因子或细胞外基质(ECM)组分的培养基的使用的成本有效且可控的替代方案。

[0180] 在本文所述的方法和细胞组合物的一些实施方式中,细胞是真核细胞。在本文所述的方法和细胞组合物的一些实施方式中,真核细胞是哺乳动物细胞,例如人类细胞。在其他实施方式中,真核细胞是非哺乳动物细胞,如昆虫或酵母细胞。

[0181] 可以通过所述方法培养或制备或包含在根据本发明的细胞组合物中的哺乳动物细胞的非限制性例子。在一些实施方式中,所述细胞是肝细胞、成纤维细胞、角质化细胞和内皮细胞。在一些实施方式中,所述细胞是干细胞和来自包括 β 细胞的胰岛的细胞。在一些实施方式中,所述细胞是选自神经祖细胞、间充质祖细胞和造血祖细胞的祖细胞。在一些实施方式中,所述细胞是选自造血干细胞、神经干细胞、间充质干细胞、乳腺干细胞、内皮干细胞、上皮干细胞和嗅干细胞,特别选自造血干细胞、神经干细胞和间充质干细胞的成体干细胞。

[0182] 在一些实施方式中,本发明的组合物用作组织粘合剂。如本文所用,“组织粘合剂(也称为组织密封剂或组织胶)”允许彼此连接,特别是重新连接组织层,例如至少两个组织层。特别地,组织粘合剂可以在组织层之间提供紧密的,特别是形式适合(form-fit)的连接,或者在组织层彼此远离的情况下,组织粘合剂可以填充组织层之间的间隙,替换缺失的组织层和/或桥接缺失的组织层。

[0183] 可以通过本所述方法或使用根据本发明的细胞组合物培养或制备的组织的非限制性例子包括结缔组织、肌肉组织、神经组织、上皮组织、及其组合,例如多个(不同的)组织或任何器官,例如胃、小肠、大肠、肠、直肠、食道、肺、脾、脑、心、肾、肝、皮肤、腺如淋巴和甲状腺、眼或胰。

[0184] 在一些实施方式中,组合物还包含细胞结合基序。关于在某些条件下培养某些细胞,细胞结合基序的存在可以改善或维持细胞活力。在一些实施方式中,细胞结合基序是通过至少一个肽键与本发明的纤维偶联的寡肽。例如,它可以偶联到本发明纤维内的蛋白质的N-端或C-端,或者在本文所述的蛋白质混合物的其余部分的氨基酸序列内的任何位置。关于寡肽细胞结合基序的选择,技术人员知道几种替代方案。所述寡肽可以例如包含选自RGD、RGE、IKVAV、YIGSR、EPDIM和NKDIL的氨基酸序列。RGD、IKVAV和YIGSR是普通的细胞结合基序,而EPDIM和NKDIL被称为角质化细胞特异性基序,这在角质化细胞培养的情况下特别有用。使用标准基因工程或化学偶联技术,本领域技术人员容易地实现寡肽细胞结合基序与纤维内蛋白质的偶联。因此,在一些实施方式中,细胞结合基序通过基因工程引入,即形成编码“野生型”蛋白质的核酸与细胞结合基序之间的遗传融合的部分。

[0185] 在一些实施方式中,与本发明的纤维接触的细胞呈多层形式。多层细胞培养或3D细胞培养包括至少2层细胞,如一层中至少10%的细胞与另一层中至少10%的细胞接触。在一些实施方式中,多层细胞培养或3D细胞培养包括至少3层细胞。

[0186] 在一些实施方式中,多层细胞培养或3D细胞培养中一层中至少10%的细胞与同一多层细胞培养或3D细胞培养中另一层中至少10%的细胞接触细胞。在一些实施方式中,多层细胞培养或3D细胞培养中一层中至少20%的细胞与同一多层细胞培养或3D细胞培养中另一层中至少20%的细胞接触。在一些实施方式中,多层细胞培养或3D细胞培养中一层中至少30%的细胞与同一多层细胞培养或3D细胞培养中另一层中至少30%的细胞接触。在一些实施方式中,多层细胞培养或3D细胞培养中一层中至少40%的细胞与同一多层细胞培养或3D细胞培养中另一层中至少40%的细胞接触。在一些实施方式中,多层细胞培养或3D细胞培养中一层中至少50%的细胞与同一多层细胞培养或3D细胞培养中另一层中至少50%的细胞接触。在一些实施方式中,多层细胞培养或3D细胞培养中一层中至少60%的细胞与同一多层细胞培养或3D细胞培养中另一层中至少60%的细胞接触。在另一个实施方式中,

短语“接触”是物理接触。在另一个实施方式中,短语“接触”是细胞与细胞的相互作用。

[0187] 在另一个实施方式中,短语“3D培养(三维培养)”是指其中细胞被置于与细胞生长相容的条件,同时允许细胞在多于一个层中生长的培养。

[0188] 物品

[0189] 在另外的方面,本发明提供包含至少一种由根据本发明的重组蛋白构成的纤维的物品。

[0190] 术语“物品”或“制造品”包括为有形、可移动和独立物体的制造物品。更具体地,本文术语“物品”是指这样的制造物品,其包含或掺入本发明的氨基酸序列、重组蛋白质和纤维中的至少一种。这些物品的非限制性实例包括:合成蜘蛛丝涂布的支架和缝合线、皮肤贴片、组织支架材料、织物、背心、防弹背心、绳索、线、化妆品等。

[0191] 这种物品的实例是用于手术缝合的线、或用于编织服装的线、或物品可以是用于各种组织工程方面的支架。

[0192] 根据本发明的物品的其它实例包括医疗装置,如医用粘合带、皮肤移植物、替换韧带和外科网;以及各种工业和商业产品,如服装面料、防弹背心衬里、容器织物、包或钱包带、电缆、绳索、钓丝、粘合性粘合材料、非粘合性粘合材料、捆扎材料、汽车覆盖件和部件、飞机建筑材料、耐候性材料、柔性隔断材料、运动设备;事实上,高拉伸强度和弹性是期望特性的纤维或织物的几乎任何用途。本发明还考虑其他形式的稳定纤维产品如干喷涂层、珠状颗粒或与其他组合物混合使用的适应性和用途。

[0193] 本发明的重组蜘蛛丝蛋白可以添加到纤维素和角蛋白和胶原产品中,因此,本发明还涉及包含纤维素和/或角蛋白和/或胶原和本发明的蜘蛛丝蛋白的纸或皮肤护理和头发护理产品。其中掺入了本发明的蛋白质的纸和皮肤护理和头发护理产品显示出改善的特征,特别是改善的拉伸强度或撕裂强度。

[0194] 复合材料

[0195] 在一些实施方式中,本发明提供了复合材料,其包含:(a)具有不同分子量的用于制备合成拖丝蜘蛛丝的蛋白质混合物;和(b)聚合物。

[0196] 在一些实施方式中,术语“复合材料”是指由两种或更多种具有不同特征的物质构成并且其中每种物质保持其本性同时为整体提供所需特性的材料。

[0197] 在一些实施方式中,术语“材料”是指固体材料。在一些实施方式中,术语“材料”是指半固体材料(例如,凝胶)。

[0198] 在一些实施方式中,公开的复合材料表现出优异的机械性能

[0199] 在一些实施方式中,提供了包含蛋白质混合物的纤维。

[0200] 在一些实施方式中,多个纤维通过接头彼此附接。

[0201] 在一些实施方式中,如本文通篇所用的术语“聚合物”描述了由多个重复结构单元(可互换地称为主链单元或单体单元)共价地相互连接并形成聚合物的聚合物主链构成的物质,例如有机物,但可选地无机物。如本文所用,术语“聚合物”包括有机和无机聚合物,并且还包括均聚物、共聚物或其混合物(例如掺合物)中的一种或多种。如本文所用,术语“均聚物”描述了由一种类型的单体单元制成并因此由同种主链单元构成的聚合物。如本文所用,术语“共聚物”描述了由多于一种类型的单体单元制成并因此由异种主链单元构成的聚合物。异种主链单元可以通过其侧基而彼此不同。

[0202] 为了简单起见,如本文通篇可互换使用的术语“聚合物”和“聚合物主链”涉及均聚物、共聚物和它们的混合物。

[0203] 在一些实施方式中,聚合物是疏水的。在一些实施方式中,聚合物是UV固化的。

[0204] 在一些实施方式中,公开的复合材料是生物稳定的。在一些实施方式中,公开的复合材料是生物可裂解的。

[0205] 在一些实施方式中,术语“生物稳定的”描述了在生理条件下保持完整(例如,在体内不降解,因此不可生物降解或不可生物裂解)的化合物或聚合物。

[0206] 在一些实施方式中,术语“生物可降解的”描述了可在生理和/或环境条件下分解成分解产物的物质。这样的生理和/或环境条件包括例如水解(通过水解切割而分解)、酶促催化(酶促降解)和机械相互作用。该术语通常是指在这些条件下分解的物质,使得50重量百分比的物质在短于一年的时间内分解。

[0207] 在一些实施方式中,如在本发明的实施方式的上下文中使用的术语“生物可降解的”还包括术语“生物可吸收的”,其描述了在生理条件下分解以将经历生物吸收的产物分解为宿主生物体即变成宿主生物体的生物化学系统的代谢物的物质。

[0208] 在一些实施方式中,聚合物是或包含合成聚合物。在一些实施方式中,聚合物是或包含天然聚合物。天然聚合物可以指由但不限于天然来源如植物、动物和矿物来源制成的聚合物,或者可以由天然纤维如棉花、亚麻、黄麻、苧麻、剑麻和大麻、毛发、和羊毛织成。

[0209] 进一步的示例性天然聚合物包括聚丙交酯、胶原角蛋白、纤维素、肌动蛋白(actine)、肌球蛋白(myosine)、壳多糖、家蚕丝。

[0210] 在一些实施方式中,聚合物是热塑性聚合物。在一些实施方式中,聚合物是热固性塑料。在一些实施方式中,聚合物是环氧树脂。在一些实施方式中,聚合物是聚酯(例如脂肪族聚酯)。在一些实施方式中,聚合物选自聚酰胺、聚氨酯和尼龙。在一些实施方式中,聚合物是交联聚合物。在一些实施方式中,聚合物是共聚物。在一些实施方式中,聚合物呈水凝胶形式。

[0211] 在一些实施方式中,聚合物材料是两种或更多种组分材料(例如共聚物)。如以下实施例部分所示,组分(也称为“部分”)A可以是主要(基础)聚合物,并且部分B可以是例如硬化剂或催化剂。

[0212] 硬化剂化学族随着聚合物基础而变化,但包括胺、异氰酸酯、过氧化物和其他少数几种。

[0213] 共聚物可以通过选自自由基聚合过程(例如使用偶氮二异丁腈(缩写为AIBN)),逐步生长聚合和链增长聚合的机制来生产。

[0214] 如本文所用,术语“环氧”是指具有一个氧和两个亚甲基的三元杂环分子的反应基团,其分子式为-C₂H₃O。

[0215] 生产方法

[0216] 在一些实施方式中,提供了用于生产本发明的蛋白质混合物的方法。在具体实施方式中,本发明的方法包括以下步骤:

[0217] a. 提供包含编码所述氨基酸序列的核酸序列的表达载体,其中所述核酸在可操作地连接的启动子和任选地调节序列的表达控制下;

[0218] b.用(a)的表达载体转化宿主细胞;

[0219] c.提供(b)的宿主细胞表达异种蛋白的条件;和

[0220] d.分离表达的蛋白质,从而获得本发明的合成氨基酸序列。

[0221] 可以理解的是,尽管本发明一般涉及源自十字园蛛拖丝的合成蜘蛛丝蛋白或其任何片段或部分,但是许多其他蜘蛛物种也可用于以类似方式获得合成蜘蛛丝。更优选地,拖丝蛋白源自以下蜘蛛中的一种或多种:希氏尾园蛛(*Arachnura higginsi*)、*Araneus circuli*sparsus、十字园蛛、*Argiope picta*、条带园蛛(Banded Garden Spider) (三带蜘蛛(*Argiope trifasciata*))、Batik Golden Web Spider(*Nephila antipodiana*)、Beccari's Tent Spider(*Cyrtophora beccarii*)、鸟粪蛛(*Celaenia excavata*)、黑白棘蛛(Black-and-White Spiny Spider) (库氏棘腹蛛(*Gasteracantha kuhlii*))、黑黄园蛛(Black-and-yellow Garden Spider) (*Argiope aurantia*)、流星锤蛛(Bolas Spider) (*Ordgarius furcatus*)、流星锤蛛巨蜘蛛(Bolas Spiders Magnificent Spider) (*Ordgarius magnificus*)、棕色水手蛛(Brown Sailor Spider) (嗜水新园蛛(*Neoscona nautica*))、棕腿蛛(Brown-Legged Spider) (*Neoscona rufofemorata*)、带帽黑头蜘蛛(Capped Black-Headed Spider) (帆楚蛛(*Zygiella calyptrata*))、普通园蛛(Common Garden Spider) (*Parawixia dehaani*)、Common Orb Weaver(*Neoscona oxancensis*)、蟹样棘园蛛(Crab-like Spiny Orb Weaver) (*Gasteracantha cancriformis*(*elipsoides*))、Curved Spiny Spider(*Gasteracantha arcuata*)、皿云斑蛛(*Cyrtophora moluccensis*)、*Cyrtophora parnasia*、*Dolophones conifera*、*Dolophones turrigera*、Doria's Spiny Spider(*Gasteracantha doriae*)、双点棘蛛(Double-Spotted Spiny Spider) (*Gasteracantha mammosa*)、双尾帐篷蜘蛛(Double-Tailed Tent Spider) (方格云斑蛛(*Cyrtophora exanthematica*))、塞若尖腹蛛(*Aculeperia ceropegia*)、*Eriophora pustulosa*、Flat Anepsion(*Anepsion depressum*)、四棘宝石蜘蛛(Four-spined Jewel Spider) (*Gasteracantha quadrispinosa*)、园网蛛(Garden Orb Web Spider) (*Eriophora transmarina*)、Giant Lichen Orbweaver(*Araneus bicentenarius*)、金色网蛛(Golden Web Spider) (*Nephila maculata*)、Hasselt's Spiny Spider(*Gasteracantha hasseltii*)、*Tegenaria atrica*、*Heurodes turrita*、Island Cyclosa Spider(岛艾蛛(*Cyclosa insulana*))、珠宝或棘蜘蛛(Jewel or Spiny Spider) (*Astracantha minax*)、肾形园蛛(Kidney Garden Spider) (丽园蛛(*Araneus mitificus*))、Laglaise's Garden Spider(*Eriovixia laglaisei*)、Long-Bellied Cyclosa Spider(*Cyclosa bifida*)、Malabar Spider(*Nephilengys malabarensis*)、多色圣安德鲁十字蜘蛛(Multi-Coloured St Andrew's Cross Spider) (多色金蛛(*Argiope versicolor*))、观赏性树干蛛(Ornamental Tree-Trunk Spider) (裂腹蛛(*Herennia ornatissima*))、椭圆形圣安德鲁十字蜘蛛(Oval St.Andrew's Cross Spider) (好胜金蛛(*Argiope aemula*))、红色帐篷蜘蛛(Red Tent Spider) (单色云斑蛛(*Cyrtophora unicolor*))、俄罗斯帐篷蜘蛛(Russian Tent Spider) (*Cyrtophora hirta*)、圣安德鲁十字蜘蛛(凯氏金蛛(*Argiope keyserlingi*))、猩红阿秋蛛(Scarlet Acusilas) (*Acusilas coccineus*)、银色金蛛(Silver Argiope) (*Argiope argentata*)、Spinybacked Orbweaver(*Gasteracantha cancriformis*)、斑点园蛛(Spotted Orbweaver) (*Neoscona domiciliorum*)、圣安德鲁十字

(Argiope aetheria)、圣安德鲁十字蜘蛛(Argiope Keyserlingi)、树干蜘蛛(Tree-Stump Spider)(无鳞波蛛(Poltys illepidus))、三角蜘蛛(Triangular Spider)(Arkys clavatus)、三角蜘蛛(Arkys lancearius)、双棘蜘蛛(Two-spined Spider)(Poecilopachys australasia)、络新妇蛛属(Nephila)物种,例如Nephila clavipes、Nephila senegalensis、Nephila madagascariensis和更多。

[0222] 此外,合成蜘蛛丝不仅可以通过选择源自的不同的蜘蛛物种来增强,而且可以通过使用除蛋白质以外的各种化合物来增强。吡咯烷具有吸湿性,有助于保持线湿润。它以特别高的浓度出现在胶线中。磷酸氢钾在水溶液中释放质子,导致pH值约为4,使丝呈酸性,从而保护它免受真菌和细菌的伤害,否则这些细菌会消化蛋白质。据信硝酸钾防止蛋白质在酸性环境中变性。

[0223] 如本文所用,术语“约”是指 $\pm 10\%$ 。

[0224] 本文使用词语“示例性”来表示“用作例子、实例或说明”。被描述为“示例性”的任何实施方式不一定被解释为优选或优于其他实施方式和/或不包括从其他实施方式引入特征。本文使用术语“任选地”表示“在一些实施方式中提供而在其他实施方式中不提供”。本发明的任何特定实施方式可以包括多个“任选的”特征,除非这些特征相冲突。

[0225] 如本文和所附权利要求中所使用的,除非上下文另外明确指出,否则单数形式“一”、“和”以及“该”包括复数指示物。因此,例如,提及“MaSp”包括多个这样的基因和变体,并且提及“该肽”包括本领域技术人员已知的一种或多种肽,等等。

[0226] 另外,除非另有说明,否则“或”的使用意味着“和/或”。类似地,“包含”、“包括”是可互换的,并且不意欲是限制性的。还应理解,在各种实施方式的描述使用术语“包含”的情况下,本领域技术人员将理解,在一些特定情况下,可以使用语言“基本上由...组成”或“由...组成”来可选地描述实施方式。

[0227] 贯穿本申请,本发明的各种实施方式可以以范围形式呈现。应该理解,范围形式的描述仅仅是为了方便和简洁,而不应该被解释为对本发明范围的不灵活的限制。因此,范围的描述应被认为已具体公开了所有可能的子范围以及该范围内的单个数值。例如,1到6的范围的描述应该被认为已具体公开了子范围,如1到3、1到4、1到5、2到4、2到6、3到6等,以及该范围内的单个数字,例如1、2、3、4、5和6。无论范围的宽度如何,这都适用。

[0228] 在使用类似于“A、B和C等中的至少一个”的惯例的那些情况下,一般而言,在本领域技术人员将理解该惯例的意义上(例如,“具有A、B和C中的至少一个的系统”将包括但不限于具有以下系统:单独A、单独B、单独C、A和B一起、A和C一起、B和C一起、和/或A、B和C一起等),这样的构造是预期的。在使用类似于“A、B或C等中的至少一个”的惯例的那些情况下,一般而言,在本领域技术人员将理解该惯例的意义上(例如,“具有A、B或C中的至少一个的系统”将包括但不限于具有以下系统:单独A、单独B、单独C、A和B一起、A和C一起、B和C一起、和/或A、B和C一起等),这样的构造是预期的。本领域技术人员将会进一步理解,无论在说明书、权利要求书还是附图中,实际上任何呈现两个或更多个可选术语的分离性词语和/或短语应当被理解为考虑包括该术语中的一个、该术语中的任一个、或两个术语的可能性。例如,短语“A或B”将被理解为包括“A”或“B”或“A和B”的可能性。

[0229] 除非另外定义,否则本文使用的所有技术和科学术语具有与本公开所属领域的普通技术人员通常理解的相同的含义。尽管与本文所述相似或等同的方法和材料可以用于所

公开的方法和组合物的实践中,但是本文描述了示例性方法、装置和材料。

[0230] 实施例

[0231] 材料和方法

[0232] 质粒:从Genearth (Regensburg, Germany) 获得的PCR-ScriptAmpSK (+) 质粒中的DNA序列。从Invitrogen获得的pFastBacHTa。

[0233] 限制酶:从(New England Biolabs, MA, USA) 获得的PstI、HindIII、NsiI。

[0234] 转染和转化:含有杆粒和辅助质粒的感受态大肠杆菌DH10BAC细胞来自Invitrogen。ESCORT转染试剂从Sigma-Aldrich获得。

[0235] 培养基:ESF 921昆虫细胞培养基,无蛋白质,从Expression Systems获得。

[0236] 细胞:在悬浮液中生长的SF9-草地贪夜蛾昆虫细胞(ATCC#:CRL-1711)。

[0237] 抗体:从Roche获得的小鼠抗-His6单克隆抗体。从Jackson实验室获得的德克萨斯红缀合的抗小鼠二级IgG。

[0238] 染料:NanoVan (Nanoprobes, NY, USA)。

[0239] 成像:Olympus BX51荧光显微镜。MagnaFire SP照相机来自Optonics。

[0240] 实验程序

[0241] 编码拖丝蜘蛛丝蛋白的单个重复单元的序列的合成:设计代表构成ADF-4 (Genbank登录U47856) 的重复区的15个重复的平均共有序列的35个氨基酸长的序列。平均共有序列肽序列是:SGPGGYGPGSQGPSGPGGYGPGGPGSSAAAAAAAA (SEQ ID NO 14), 其由105个DNA碱基对序列:5'-TCTGGTCCTGGAGGTTATGGCCCAGGAAGCCAAGGACCATCTGGTCCAGGAGGATATGGTCCAGGCGGACCTGGCTCTAGTGCAGCAGCTGCCGAGCAGCTGCA-3' (SEQ ID NO:15) 编码。上述合成的DNA在PCR-ScriptAmpSK (+) 质粒中获得。根据草地贪夜蛾的密码子使用, 序列被优化用于表达, 其细胞用于合成蜘蛛丝蛋白和纤维。

[0242] 在某些实验中,进行了具有不同氨基酸序列的三种构建体,每种具有独特的氨基酸序列:

[0243] C1 (SEQ ID NO:35和36):SGPGGYGPGSQGPSGPGGYGPGGPGSSAGAGAGAXaaA (Xaa描绘Ala或Gly)

[0244] C2 (SEQ ID NO:38和40):XaaGSGPGGYGPGGQGPGGYGPGGQGPYGPAAAAAAAA (Xaa描绘Ser或Gly)

[0245] C3 (SEQ ID NO:42和44):XaaGPGQGGYGGPGGQGPGRGGYGPAGSAAAAAAAA (Xaa描绘Ser或Gly)。

[0246] 编码三个(C1-C3) 构建体的单个重复的多核苷酸序列分别以SEQ ID NO:45-47提供。

[0247] 供体质粒构建:用Xba I和Xho I切除ScriptAmpSK (+) 质粒,将含有Nsi I和Pst I限制位点位于两侧的基本重复序列的136-bp序列分离和克隆到杆状病毒供体质粒pFastBacHTa的多克隆位点(MCS)中。因此,产生了编码人工49个氨基酸N-端结构域和35个氨基酸核心结构域的基本供体质粒。

[0248] 单个重复的多聚化:限制酶位点Nsi I和Pst I位于编码蜘蛛丝蛋白的一个重复(单体)的基本模块两侧,它们是相容的。在第一步中,单体通过双重限制来释放并符合读框地(in frame)插入用Pst I切割的相同供体质粒中。只有如果插入以正确的有义方向连接时,

双切割才能释放二聚体(两个重复之间的限制位点在连接时被消除)。在第二步中释放二聚体,然后以相同的方式重新插入以获得具有四个重复的载体。在接下来的步骤中,重复该程序以获得含有多个合成重复的供体质粒。由使用的分子生物学工具产生的约束和序列的重复性质限制了相同重复的最大可实现数目。

[0249] 将天然C-端结构域连接至合成重复的下游:使用具有以下引物的PCR进行ADF4 1 1 4 个氨基酸的C-端结构域的插入:有义引物,具有序列5'-ATATGCTGCAGGCCCTAGTGGTCCTGGA-3'(SEQ ID NO:16),含有PstI限制位点(加下划线的),和反义引物,具有序列5'-TCGACAAGCTTGGTACCGCA-3'(SEQ ID NO:17),编码3'HindIII限制位点(加下划线的)。将具有不同重复数目的供体质粒载体和PCR产物用PstI和HindIII切除,纯化并连接,产生编码His6标签的pFastBacHTa供体质粒,所述His6标签是人工N端结构域的一部分,随后是不同的数目相同重复(本发明人获得了含有核酸序列的1、2、4、8、12、16、20、24、32个重复的构建体)和天然C末端结构域。

[0250] 细胞培养:将Sf9细胞在27℃下在ESF 921无血清昆虫细胞培养基中繁殖。Sf9细胞或者作为单层在6孔板上的盖玻片上或者在130rpm搅拌的摇瓶中生长。

[0251] 重组杆状病毒的产生:使用含有杆粒(杆状病毒穿梭载体质粒)和辅助质粒的感受态大肠杆菌DH10BAC细胞根据制造商的方案(Invitrogen)产生重组杆粒。通过PCR验证基因插入杆粒中。使用ESCORT转染试剂在6孔板中用重组杆粒DNA转染Sf9细胞。将细胞在27℃温育5小时,冲洗并再温育72小时。收获培养基,离心,含有病毒的上清液用于连续2-3次感染,导致病毒体滴度的放大。

[0252] 基于合成的ADF-4的蛋白质的表达:用重组病毒以0.1至10的各种MOI(感染复数)范围感染Sf9细胞(3×10^6 个细胞/ml)。感染四天后,通过以16000g离心10min收获细胞。

[0253] 合成纤维的纯化:感染后4天收获感染的细胞,并以16000g离心10min。将细胞沉淀重悬于0.25%SDS溶液中,在室温下温育30分钟,如上沉淀蛋白质组装体。纯化纤维的典型产率约为150mg/LSf9昆虫细胞培养物。将纯化的纤维以所需溶液和体积再悬浮。

[0254] 差示扫描量热计(DSC)测量:使用Mettler DSC 822e热分析仪和铝样品盘在惰性氮气氛围下对3-6mg丝心蛋白样品进行热分析。在5℃/min的加热速率下,温谱图(thermograms)的范围为从10℃到400℃。

[0255] 免疫细胞化学:以MOI=10用重组病毒感染在50%汇合的盖玻片上生长的细胞。感染后三天,细胞在-20℃用甲醇固定。将盖玻片与1:300稀释度的小鼠抗-His6单克隆抗体温育,然后与1:500稀释度的德克萨斯红缀合的抗小鼠二级IgG温育。用Olympus BX51荧光显微镜观察细胞,并用MagnaFire SP照相机拍摄图像或通过共聚焦显微镜分析。

[0256] 透射电子显微镜检查(TEM):为了超结构分析,将纯化的细丝直接吸附到300目铜多孔碳栅上,或用钒(NanoVan1, Nanoprobes)负染,用在120千伏下运行的Tecnai T12显微镜观察和拍照。

[0257] 实施例1

[0258] 表征组成微纤维的纳米纤维

[0259] 将直径为1.8纳米的金颗粒(Ni-NTA-纳米探针)结合到本发明每种单体的N-端His6标签上。免疫TEM观察揭示了两步装配过程,其中单体首先以头尾方式相互作用,从而产生直径约为5纳米的纳米原纤维作为第一阶段(图1A)。因此,纳米原纤维以非定向的方式

相互作用(图1B),从而产生直径约为150纳米的纤维(图1C)。

[0260] 此外,使用光学显微镜检查基于三种独立构建体(C1、C2和C3)的各种组装纤维的表型和使用共聚焦显微镜检查荧光标记的纤维。如图1D-I所示,构建体C1的单体自组装以在表达细胞质溶胶内产生卷曲纤维。该纤维的直径范围在100至200nm之间,长度范围在10至150 μ m之间。构建体C2的单体自组装以产生直径范围在100至200nm之间而它们的长度范围在10至50 μ m之间的纤维。构建体C3的单体自组装成各种表型:(i)平均直径为1 μ m的球状聚集体;(ii)覆盖有平均直径为500nm的聚集体的纤维;(iii)具有光滑表面的纤维,类似于构建体C1和C2。

[0261] 实施例2

[0262] 独特的梯状现象表征了本发明的纤维

[0263] 不重复的N-端和C-端结构域位于两侧的含有不同数目的重复(0、1、2、3、4、8、12、16、20、24、32)的单体倾向于自组装并变成不溶。因此,为了确定单体的分子量(MW),将纤维纯化并用6M胍SCN拆解。顺序地,将胍溶液针对8M尿素透析(使用MW截止值为10KDa的穿孔透析盒),加入样品缓冲液,并在变性10%丙烯酰胺凝胶中分析样品。

[0264] 在来自三种不同来源的拆解纤维上进行考马斯蓝染色:

[0265] (1)来自自由N-端结构域、24个相同重复(2.7KDa)、C-端结构域组成的基因的带梯;

[0266] (2)来自自由N-端结构域、16个相同重复(2.7KDa)、C-端结构域组成的基因的带梯;
和

[0267] (3)来自自由N-端结构域、15个非相同重复(取自十字园蛛的天然序列)、C-端结构域组成的基因的带梯。

[0268] 如可以观察到的(图2A),虽然在人工构建体情况(1、2)中,存在与有利于较轻版本的带的不同带的摩尔比相关的恒定梯度,但是在天然序列的梯处不是这种情况(3)。例如,在源于天然序列的梯处与38KDa相关的带比周围带(上方和下方)密集得多。

[0269] 这一观察结果强有力地表明了不同重复区来源的重要性。虽然在人工序列情况下重复是相同的,但是在天然序列情况下,重复相差至少它们的整体大小和聚丙烯酰胺拉伸长度。此外,所述差异与源自天然重复序列的纤维与源自本文公开的人工重复序列的纤维之间存在的不同组装性质、不同热稳定性和不同机械性质相关。

[0270] SDS-PAGE和蛋白质印迹分析:

[0271] 如上所述,Sf9细胞用杆粒DNA构建体转染并在转染后72小时收获。

[0272] 细胞沉淀通过使用10%凝胶的SDS-PAGE分离并通过半干印迹转移至硝化纤维素。用奶粉在PBSx1、0.5%Tween20中封闭1小时。将膜与小鼠抗-His6单克隆抗体(1:2000)和作为二级抗体的HRP缀合的山羊抗小鼠(1:5000)温育。ECL被用来直接检测带HIS标签的构建体。

[0273] 使用针对N-端His6标签的一级抗体的蛋白质印迹验证了在每个观察到的带中存在N-端结构域(图2B)。所有三种蛋白质(描述为C1、C2和C3)引起梯状现象。所有三种蛋白质的最强带(由箭头表示)对应于源自构建体1、2和3的蛋白质的计算的M.W(分别为61.41、62.84和62.16KDa)

[0274] 通过二硫键特异性结合半胱氨酸(只有一个半胱氨酸存在于我们的蛋白质中,位于C-端结构域)的荧光团用于验证在所有观察到的梯带处C-端结构域的存在(图2C)。

[0275] 从在多角体蛋白启动子下编码全长蛋白质的杆状病毒感染的Sf9细胞提取的mRNA的RNA印迹揭示了与全长蛋白质大小相关的大小的mRNA。接下来将提取的mRNA用作RT-PCR的模板,产生cDNA,其随后用作PCR的模板,产生DNA带梯。使用具有对完全蛋白质完全序列的第一个和最后18个碱基的互补序列(complement sequence)的引物对来自上述梯的分离的DNA带进行测序。测序显示存在不同数量的重复,完全的N和C-端结构域序列位于其两侧。

[0276] 上述发现与RNA二级结构一起导致令人惊讶地发现一种新的蛋白质合成控制,其中mRNA的重复区采用发夹结构,而核糖体在该发夹上滑动。由于这个发夹的交替大小,一系列蛋白质由一个mRNA合成。这些蛋白质彼此不同之处仅在于它们所含的重复数目,从而使雌性蜘蛛能够基于一个基因表达一系列丝心蛋白单体。

[0277] 实施例3

[0278] 溶解的纯化纤维可以被电纺成连续纤维

[0279] 将从SF9感染的细胞中分离出的合成纤维脱水(在55摄氏度(°C)的条件下过夜)。向干纤维中加入六氟-2-丙醇(HFIP)至最终浓度为纺丝原液的23%重量/体积。电纺方案如下进行:注射速率:0.5ml/小时,电压:18千伏,喷嘴末端与收集器之间的距离:16厘米,湿度:40%,温度:28°C。

[0280] 上述方案产生了平均直径为100纳米的连续纤维网(图3A)。该纤维的DSC检查显示在289°C的吸热峰,表明在合成纤维中存在基于聚丙烯酸拉伸的纳米晶体(图3B)。

[0281] 实施例4

[0282] 纯化纤维的冻干和线性表征

[0283] 将纯化纤维以0.5%重量/体积悬浮在双蒸水(DDW)中,用液氮冷冻(温度-200°C),然后冻干约24小时。上述过程产生了含有干纤维的白色粉末。对这些纤维的深入检查显示,它们被线性化,保持其直径,并且容易分散在各种基质中,同时采用线性结构(图4A)以及保持其机械性质。

[0284] 值得注意的是,与天然拖丝熔点(230°C)相比,测量本文公开的合成纤维的熔点是高的,在约235°C(图4B)。

[0285] 实施例5

[0286] 本发明纤维上的多层细胞生长

[0287] 将本发明的纤维(PBS中1:25)加入无菌96孔组织培养板的孔中。在温育(在4°C)和洗涤步骤后,将HEK293细胞铺板到每个纤维涂布的孔中,随后进行温育步骤(在37°C,5%CO₂,72小时)。

[0288] 如图5A-C所见,发现HEK293细胞优选粘附于本发明的纤维而非组织培养板。此外,本发明的纤维改变了从层生长到3D生长的细胞偏好。

[0289] 实施例6

[0290] 本发明纤维的热指纹

[0291] 进行若干测试以表征本发明的纤维的热指纹。

[0292] 每个测试时纤维称重为5-10mg。每个测试在配备有液氮冷却罐的Mettler-Toledo DSC 2系统上在40μl体积的穿孔铝盘上进行。在加热至100°C并保持该温度5min的脱水过程后,以5°C/min的加热速率从25-350°C运行测试。

[0293] 结果显示,出乎意料的是,热指纹显示峰开始于265±5°C,直到320±5°C,在~300

℃具有峰值。

[0294] 额外的测试显示具有3个峰的热指纹：

[0295] 1. 在-120-160℃的放热峰；

[0296] 2. 在230-260℃的吸热峰；

[0297] 3. 在 $265 \pm 5^\circ\text{C}$ 直至 $320 \pm 5^\circ\text{C}$ 的分解峰。

[0298] 虽然已经结合其具体实施例描述了本发明，但是明显的是，对于本领域技术人员而言，许多替代、修改和变化将是显而易见的。因此，旨在包括落入所附权利要求的精神和广泛范围内的所有这些替代、修改和变化。

[0299] 本说明书中提及的所有出版物、专利和专利申请在此通过引用以其全部并入本说明书，其程度如同每个单独的出版物、专利或专利申请被具体和单独地指出通过引用并入本文。另外，本申请中任何参考文献的引用或识别不应被解释为承认这种参考文献可作为本发明的现有技术。就使用章节标题而言，不应将其解释为必需的限制。

序列表

<110> 赛威克斯材料科学公司
<120> 用于制造合成的拖丝蜘蛛丝的组合物和方法
<130> SVX-P-01-PCT
<150> 62/203,102
<151> 2015-08-10
<150> 62/293,880
<151> 2016-02-11
<150> 62/317,572
<151> 2016-04-03
<160> 47
<170> PatentIn version 3.5
<210> 1
<211> 28
<212> PRT
<213> 人工的
<220>
<223> 合成的
<220>
<221> MISC_FEATURE
<222> (1) .. (1)
<223> X是A或G
<220>
<221> MISC_FEATURE
<222> (2) .. (2)
<223> X是S或G
<220>
<221> MISC_FEATURE
<222> (10) .. (10)
<223> X是G或E
<220>
<221> MISC_FEATURE
<222> (11) .. (11)
<223> X是G, S或N
<220>
<221> MISC_FEATURE
<222> (12) .. (12)
<223> X是Q或Y

<220>

<221> MISC_FEATURE

<222> (15) .. (15)

<223> X是G或S

<220>

<221> MISC_FEATURE

<222> (17) .. (17)

<223> X是P或R

<220>

<221> MISC_FEATURE

<222> (20) .. (20)

<223> X是Y或Q

<220>

<221> MISC_FEATURE

<222> (27) .. (27)

<223> X是G或S

<220>

<221> MISC_FEATURE

<222> (28) .. (28)

<223> X是S或G

<400> 1

Xaa Xaa Gly Pro Gly Gly Tyr Gly Pro Xaa Xaa Xaa Gly Pro Xaa Gly

1 5 10 15

Xaa Gly Gly Xaa Gly Pro Gly Gly Pro Gly Xaa Xaa

20 25

<210> 2

<211> 27

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 2

Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro

1 5 10 15

Gly Gly Tyr Gly Pro Gly Gly Pro Gly Ser Ser

20 25

<210> 3

<211> 35

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 3

Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro
1				5					10					15	
Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	Gly	Pro
			20					25					30		
Gly	Ser	Ser													
			35												

<210> 4

<211> 35

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 4

Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro
1				5					10					15	
Glu	Asn	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	Gly	Pro
			20					25					30		
Gly	Ser	Ser													
			35												

<210> 5

<211> 40

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 5

Met	Ser	Tyr	Tyr	His	His	His	His	His	His	Asp	Tyr	Asp	Ile	Pro	Thr
1				5					10					15	
Thr	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Ala	Met	Asp	Pro	Glu	Phe	Lys	Gly
			20					25					30		
Leu	Arg	Arg	Arg	Ala	Gln	Leu	Val								
			35				40								

<210> 6

<211> 49

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 6

Met	Ser	Tyr	Tyr	His	His	His	His	His	His	Asp	Tyr	Asp	Ile	Pro	Thr
1				5					10					15	
Thr	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Ala	Met	Asp	Pro	Glu	Phe	Lys	Gly
				20				25					30		
Leu	Arg	Arg	Arg	Ala	Gln	Leu	Val	Arg	Pro	Leu	Ser	Asn	Leu	Asp	Asn
				35				40					45		

Ala

<210> 7

<211> 54

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 7

Met	Ser	Tyr	Tyr	His	His	His	His	His	His	Asp	Tyr	Asp	Ile	Pro	Thr
1				5					10					15	
Thr	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Ala	Met	Asp	Pro	Glu	Phe	Lys	Gly
				20				25					30		
Leu	Arg	Arg	Arg	Ala	Gln	Leu	Val	Asp	Pro	Pro	Gly	Cys	Arg	Asn	Ser
				35				40					45		

Ala Arg Ala Gly Ser Ser
50

<210> 8

<211> 28

<212> PRT

<213> 人工的

<220>

<223> 合成的

<220>

<221> MISC_FEATURE

<222> (1) .. (1)

<223> X是A或G

<220>

<221> MISC_FEATURE

<222> (2) .. (2)

<223> X是S或G
<220>
<221> misc_feature
<222> (10) .. (10)
<223> Xaa可以是任何天然存在的氨基酸
<220>
<221> MISC_FEATURE
<222> (11) .. (11)
<223> X是S或G
<220>
<221> MISC_FEATURE
<222> (12) .. (12)
<223> X是Q或Y
<220>
<221> MISC_FEATURE
<222> (15) .. (15)
<223> X是S或G
<220>
<221> MISC_FEATURE
<222> (17) .. (17)
<223> X是P或R
<220>
<221> MISC_FEATURE
<222> (20) .. (20)
<223> X是Y或Q
<220>
<221> MISC_FEATURE
<222> (27) .. (27)
<223> X是G或S
<220>
<221> MISC_FEATURE
<222> (28) .. (28)
<223> X是G或S
<400> 8
Xaa Xaa Gly Pro Gly Gly Tyr Gly Pro Xaa Xaa Xaa Gly Pro Xaa Gly
1 5 10 15
Xaa Gly Gly Xaa Gly Pro Gly Gly Pro Gly Xaa Xaa
20 25
<210> 9

<211> 114

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 9

Gly	Pro	Ser	Gly	Pro	Gly	Ala	Tyr	Gly	Pro	Ser	Pro	Ser	Ala	Ser	Ala
1				5					10					15	
Ser	Val	Ala	Ala	Ser	Arg	Leu	Ser	Ser	Pro	Ala	Ala	Ser	Ser	Arg	Val
			20					25					30		
Ser	Ser	Ala	Val	Ser	Ser	Leu	Val	Ser	Ser	Gly	Pro	Thr	Asn	Gly	Ala
		35					40					45			
Ala	Val	Ser	Gly	Ala	Leu	Asn	Ser	Leu	Val	Ser	Gln	Ile	Ser	Ala	Ser
		50				55					60				
Asn	Pro	Gly	Leu	Ser	Gly	Cys	Asp	Ala	Leu	Val	Gln	Ala	Leu	Leu	Glu
65					70					75					80
Leu	Val	Ser	Ala	Leu	Val	Ala	Ile	Leu	Ser	Ser	Ala	Ser	Ile	Gly	Gln
				85					90					95	
Val	Asn	Val	Ser	Ser	Val	Ser	Gln	Ser	Thr	Gln	Met	Ile	Ser	Gln	Ala
			100					105						110	

Leu Ser

<210> 10

<211> 555

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 10

Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Ser	Gly	Gly	Tyr	Gly	Pro	Glu
1				5					10					15	
Asn	Gln	Gly	Pro	Ser	Gly	Pro	Val	Ala	Tyr	Gly	Pro	Gly	Gly	Pro	Val
			20						25					30	
Ser	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Ser	Gly	Pro	Gly	Gly	Tyr
		35					40					45			
Gly	Pro	Glu	Asn	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly
		50				55					60				
Gly	Ser	Gly	Ser	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Pro
65					70					75					80
Gly	Gly	Tyr	Gly	Pro	Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Ser

85							90						95			
Gly	Gly	Tyr	Gly	Pro	Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Ala	Ser	
100							105						110			
Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	
115							120						125			
Pro	Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Ala	Tyr	Gly	Pro	Gly	Gly	
130							135						140			
Pro	Gly	Ser	Ser	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	
145							150						155			
Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Ser	Gly	Gly	Tyr	Gly	Pro	Gly	
165							170						175			
Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Pro	Gly	Ala	Ser	Ala	Ala	Ala	
180							185						190			
Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	Ser	
195							200						205			
Gln	Gly	Pro	Ser	Gly	Pro	Gly	Ala	Tyr	Gly	Pro	Gly	Gly	Pro	Gly	Ser	
210							215						220			
Ser	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	Ser	Gln	Gly	
225							230						235			
Pro	Ser	Gly	Pro	Gly	Ala	Tyr	Gly	Pro	Gly	Gly	Pro	Gly	Ser	Ser	Ala	
245							250						255			
Ala	Ala	Ala	Ala	Ala	Ala	Gly	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	
260							265						270			
Asn	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	Gly	Pro	Gly	
275							280						285			
Ser	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	
290							295						300			
Pro	Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Val	Tyr	Gly	Pro	Gly	Gly	
305							310						315			
Pro	Gly	Ser	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Ser	Gly	Pro	Gly	
325							330						335			
Gly	Tyr	Gly	Pro	Gly	Asn	Gln	Gly	Pro	Ser	Gly	Pro	Gly	Gly	Tyr	Gly	
340							345						350			
Pro	Gly	Gly	Ser	Gly	Ser	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ser	
355							360						365			
Gly	Pro	Gly	Gly	Tyr	Gly	Pro	Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	
370							375						380			
Gly	Ser	Gly	Gly	Tyr	Gly	Pro	Gly	Ser	Gln	Gly	Pro	Ser	Gly	Pro	Gly	
385							390						395			
													400			

Ala Ser Ser Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly
 405 410 415
 Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro Gly Ala Tyr Gly Pro
 420 425 430
 Gly Gly Pro Gly Ser Ser Ala Ala Ala Ser Gly Pro Gly Gly Tyr Gly
 435 440 445
 Pro Gly Ser Gln Gly Pro Ser Gly Pro Gly Ala Tyr Gly Pro Gly Gly
 450 455 460
 Pro Gly Ser Ser Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly
 465 470 475 480
 Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro Gly Gly Ser Arg Gly
 485 490 495
 Tyr Gly Pro Gly Ser Gln Gly Pro Gly Gly Pro Gly Ala Ser Ala Ala
 500 505 510
 Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gly Tyr Gly Pro Gly
 515 520 525
 Ser Gln Gly Pro Ser Gly Pro Gly Tyr Gln Gly Pro Ser Gly Pro Gly
 530 535 540
 Ala Tyr Gly Pro Ser Pro Ser Ala Ser Ala Ser
 545 550 555
 <210> 11
 <211> 6
 <212> PRT
 <213> 人工的
 <220>
 <223> 合成的
 <400> 11
 His His His His His His
 1 5
 <210> 12
 <211> 9
 <212> PRT
 <213> 人工的
 <220>
 <223> 合成的
 <400> 12
 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
 1 5
 <210> 13

<211> 11

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 13

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly

1

5

10

<210> 14

<211> 35

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 14

Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro

1

5

10

15

Gly Gly Tyr Gly Pro Gly Gly Pro Gly Ser Ser Ala Ala Ala Ala Ala

20

25

30

Ala Ala Ala

35

<210> 15

<211> 105

<212> DNA

<213> 人工的

<220>

<223> 合成的

<400> 15

tctggtcctg gaggttatgg cccaggaagc caaggacat ctggtccagg aggatatggt 60

ccaggcggac ctggctctag tgcagcagct gccgcagcag ctgca 105

<210> 16

<211> 28

<212> DNA

<213> 人工的

<220>

<223> 合成的

<400> 16

atatgctgca ggccctagtg gtcctgga 28

<210> 17

<211> 20

<212> DNA

<213> 人工的

<220>

<223> 合成的

<400> 17

tcgacaagct tggtagcgca 20

<210> 18

<211> 266

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 18

Gly	Ser	Ala	Met	Gly	Tyr	Leu	Trp	Ile	Gln	Gly	Gln	Gly	Gly	Tyr	Gly
1				5					10					15	
Gly	Leu	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Ala	Gly	Ser	Ser	Ala	Ala
			20					25					30		
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Gln	Gly	Gly	Gln
			35					40					45		
Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Ser	Gly	Gly	Ser	Ala	Ala	Ala	Ala
			50					55				60			
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Arg	Gly	Gln	Gly
65						70					75				80
Gly	Tyr	Gly	Gln	Gly	Ser	Gly	Gly	Asn	Ala	Ala	Ala	Ala	Ala	Ala	Ala
						85					90				95
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gln	Gly	Gly	Gln	Gly	Gly	Tyr
						100					105				110
Gly	Arg	Gln	Ser	Gln	Gly	Ala	Gly	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Ala
						115					120				125
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Ser	Gly	Gln	Gly	Gly	Tyr	Gly
						130					135				140
Gln	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Ser	Ser	Ala	Ser	Ala	Ser	Ala	Ala
145						150					155				160
Ala	Ser	Ala	Ala	Ser	Thr	Val	Ala	Asn	Ser	Val	Ser	Arg	Leu	Ser	Ser
						165					170				175
Pro	Ser	Ala	Val	Ser	Arg	Val	Ser	Ser	Ala	Val	Ser	Ser	Leu	Val	Ser
						180					185				190
Asn	Gly	Gln	Val	Asn	Met	Ala	Ala	Leu	Pro	Asn	Ile	Ile	Ser	Asn	Ile

195	200	205
Ser Ser Ser Val Ser Ala Ser Ala Pro Gly Cys Glu Val Ile Val Gln		
210	215	220
Ala Leu Leu Glu Val Ile Thr Ala Leu Val Gln Ile Val Ser Ser Ser		
225	230	235
Ser Val Gly Tyr Ile Asn Pro Ser Ala Val Asn Gln Ile Thr Asn Val		
245	250	255
Val Ala Asn Ala Met Ala Gln Val Met Gly		
260	265	
<210> 19		
<211> 257		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的		
<400> 19		
Gly Ser Ala Met Ala Ser Gly Gln Gly Gly Gln Gly Gly Gln Gly Gln		
1	5	10
Gly Gly Tyr Gly Gln Gly Ala Gly Ile Ser Ala Ala Ala Ala Ala Ala		
20	25	30
Ala Ala Ala Ala Ala Gly Ala Ala Ala Gly Arg Gly Gln Gly Gly Tyr		
35	40	45
Gly Gln Gly Ala Gly Gly Asn Ala Ala Ala Ala Ala Ala Ala Ala Ala		
50	55	60
Ala Ala Ala Ala Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gly Gln		
65	70	75
Gly Leu Gly Gly Tyr Gly Gln Gly Ala Gly Ser Ser Ala Ala Ala Ala		
85	90	95
Ala Ala Ala Ala Gly Arg Gly Gln Ser Val Tyr Ala Ser Gly Gly Ala		
100	105	110
Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala Gly Arg Gly		
115	120	125
Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly Ser Gly Gln		
130	135	140
Gly Gly Tyr Gly Gly Val Gly Ser Gly Ala Ser Ala Ala Ser Ala Ala		
145	150	155
Ala Ser Arg Leu Ser Ser Pro Glu Ala Ser Ser Arg Val Ser Ser Ala		
165	170	175
Val Ser Asn Leu Val Ser Ser Gly Pro Thr Asn Ser Ala Ala Leu Ser		

180	185	190
Ser Thr Ile Ser Asn Val Val Ser Gln Ile Gly Ala Ser Asn Pro Gly		
195	200	205
Leu Ser Gly Cys Asp Val Leu Val Gln Ala Leu Leu Glu Val Val Ser		
210	215	220
Ala Leu Ile His Ile Leu Gly Ser Ser Ser Ile Gly Gln Val Asn Tyr		
225	230	235
Gly Ser Ala Gly Gln Ala Thr Gln Leu Val Gly Gln Ser Val Tyr Gln		
245	250	255
Ala		
<210> 20		
<211> 276		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的		
<400> 20		
Gly Ser Ala Met Gly Gln Gly Ser Gly Gly Asn Ala Ala Ala Ala Ala		
1	5	10
Ala Ala Ala Ala Ala Ala Ala Ala Ser Gly Gln Gly Gly Gln Gly Gly		
20	25	30
Gln Gly Gln Gly Gly Tyr Gly Gln Gly Ala Gly Ile Ser Ala Ala Ala		
35	40	45
Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala Ala Ala Gly Arg Gly Gln		
50	55	60
Gly Gly Tyr Gly Gln Gly Ala Gly Gly Asn Ala Ala Ala Ala Ala Ala		
65	70	75
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gln Gly Gly Gln Gly Gly Tyr		
85	90	95
Gly Gly Gln Gly Leu Gly Gly Tyr Gly Gln Gly Ala Gly Ser Ser Ala		
100	105	110
Ala Ala Ala Ala Ala Ala Ala Gly Arg Gly Gln Ser Val Tyr Ala Ser		
115	120	125
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Gly Gln Gly Ala		
130	135	140
Gly Arg Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Gly Gly		
145	150	155
Ser Gly Gln Gly Gly Tyr Gly Gly Val Gly Ser Gly Ala Ser Ala Ala		
165	170	175

Ser Ala Ala Ala Ser Arg Leu Ser Ser Pro Glu Ala Ser Ser Arg Val
 180 185 190
 Ser Ser Ala Val Ser Asn Leu Val Ser Ser Gly Pro Thr Asn Ser Ala
 195 200 205
 Ala Leu Ser Ser Thr Ile Ser Asn Val Val Ser Gln Ile Gly Ala Ser
 210 215 220
 Asn Pro Gly Leu Ser Gly Cys Asp Val Leu Val Gln Ala Leu Leu Glu
 225 230 235 240
 Val Val Ser Ala Leu Ile His Ile Leu Gly Ser Ser Ser Ile Gly Gln
 245 250 255
 Val Asn Tyr Gly Ser Ala Gly Gln Ala Thr Gln Leu Val Gly Gln Ser
 260 265 270
 Val Tyr Gln Ala
 275

<210> 21

<211> 269

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 21

Gly Ser Ala Met Gly Tyr Leu Trp Ile Gln Gly Gln Gly Gly Tyr Gly
 1 5 10 15
 Gly Leu Gly Gln Gly Gly Tyr Gly Gln Gly Ala Gly Ser Ser Ala Ala
 20 25 30
 Ala Ala Ala Cys Cys Ala Ala Ala Ala Gly Gly Gln Gly Gly Gln
 35 40 45
 Gly Gln Gly Gly Tyr Gly Gln Gly Ser Gly Gly Ser Ala Ala Ala Ala
 50 55 60
 Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Arg Gly Gln Gly
 65 70 75 80
 Gly Tyr Gly Gln Gly Ser Gly Gly Asn Ala Ala Ala Ala Ala Ala Ala
 85 90 95
 Ala Ala Ala Ala Ala Ala Ala Ala Gly Gln Gly Gly Gln Gly Gly Tyr
 100 105 110
 Gly Arg Gln Ser Gln Gly Ala Gly Ser Ala Ala Ala Ala Ala Ala Ala
 115 120 125
 Ala Ala Ala Ala Ala Ala Ala Gly Ser Gly Gln Gly Gly Tyr Gly Gly
 130 135 140

Gln	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Ser	Ser	Ala	Ser	Ala	Ser	Ala	Ala	
145					150					155					160	
Ala	Ser	Ala	Ala	Ser	Thr	Val	Ala	Asn	Ser	Val	Ser	Arg	Leu	Ser	Ser	
				165				170						175		
Pro	Ser	Ala	Val	Ser	Arg	Val	Ser	Ser	Ala	Val	Ser	Ser	Leu	Val	Ser	
			180					185					190			
Asn	Gly	Gln	Val	Asn	Met	Ala	Ala	Leu	Pro	Asn	Ile	Ile	Ser	Asn	Ile	
		195					200					205				
Ser	Ser	Ser	Val	Ser	Ala	Ser	Ala	Pro	Gly	Ala	Ser	Gly	Cys	Glu	Val	
	210					215					220					
Ile	Val	Gln	Ala	Leu	Leu	Glu	Val	Ile	Thr	Ala	Leu	Val	Gln	Ile	Val	
225					230					235					240	
Ser	Ser	Ser	Ser	Val	Gly	Tyr	Ile	Asn	Pro	Ser	Ala	Val	Asn	Gln	Ile	
				245				250						255		
Thr	Asn	Val	Val	Ala	Asn	Ala	Met	Ala	Gln	Val	Met	Gly				
		260						265								

<400> 22

Gly	Ser	Ala	Met	Gly	Tyr	Leu	Trp	Ile	Gln	Gly	Gln	Gly	Gly	Tyr	Gly	
				5									15			
Gly	Leu	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Ala	Gly	Ser	Ser	Ala	Ala	
						20								30		
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Gln	Gly	Gly	Gln	
						35						45				
Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Ser	Gly	Gly	Ser	Ala	Ala	Ala	Ala	
				50							60					
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Arg	Gly	Gln	Gly	
						70								80		
Gly	Tyr	Gly	Gln	Gly	Ser	Gly	Gly	Asn	Ala	Ala	Ala	Ala	Ala	Ala	Ala	
				85										95		
Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gln	Gly	Gly	Gln	Gly	Gly	Tyr	
						100							110			
Gly	Arg	Gln	Ser	Gln	Gly	Ala	Gly	Ser	Ala	Ala	Ala	Ala	Ala	Ala	Cys	
						115							125			

Cys Ala Ala Ala Ala Ala Gly Ser Gly Gln Gly Gly Tyr Gly Gly
 130 135 140
 Gln Gly Gln Gly Gly Tyr Gly Gln Ser Ser Ala Ser Ala Ser Ala Ala
 145 150 155 160
 Ala Ser Ala Ala Ser Thr Val Ala Asn Ser Val Ser Arg Leu Ser Ser
 165 170 175
 Pro Ser Ala Val Ser Arg Val Ser Ser Ala Val Ser Ser Leu Val Ser
 180 185 190
 Asn Gly Gln Val Asn Met Ala Ala Leu Pro Asn Ile Ile Ser Asn Ile
 195 200 205
 Ser Ser Ser Val Ser Ala Ser Ala Pro Gly Ala Ser Gly Cys Glu Val
 210 215 220
 Ile Val Gln Ala Leu Leu Glu Val Ile Thr Ala Leu Val Gln Ile Val
 225 230 235 240
 Ser Ser Ser Ser Val Gly Tyr Ile Asn Pro Ser Ala Val Asn Gln Ile
 245 250 255
 Thr Asn Val Val Ala Asn Ala Met Ala Gln Val Met Gly
 260 265

<210> 23

<211> 217

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 23

Gly Ser Ala Met Gly Tyr Leu Trp Ile Gln Gly Gln Gly Gly Tyr Gly
 1 5 10 15
 Gly Leu Gly Gln Gly Gly Tyr Gly Gln Gly Ala Gly Ser Ser Ala Ala
 20 25 30
 Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Gln Gly Gly Gln
 35 40 45
 Gly Gln Gly Gly Tyr Gly Gln Gly Ser Gly Ala Ala Ala Ala Ala Ala
 50 55 60
 Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Arg Gly Gln Gly Gly Tyr
 65 70 75 80
 Gly Gln Gly Ser Gly Gly Asn Ala Ala Ala Ala Ala Ala Ala Ala
 85 90 95
 Ala Ala Ala Ala Ala Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg
 100 105 110

Gln Ser Gln Gly Ala Gly Ser Ala Ala Ala Ala Ala Ala Ala Ala
 115 120 125
 Ala Ala Ala Ala Ala Gly Ser Gly Gln Gly Gly Tyr Gly Gly Gln Gly
 130 135 140
 Gln Gly Gly Tyr Gly Gln Ser Ser Ala Ser Ala Ser Ala Ala Ala Ser
 145 150 155 160
 Ala Ala Ser Thr Val Ala Asn Ser Val Ser Arg Leu Ser Ser Pro Ser
 165 170 175
 Ala Val Ser Arg Val Ser Ser Ala Val Ser Ser Leu Val Ser Asn Gly
 180 185 190
 Gln Val Asn Met Ala Ala Leu Pro Asn Ile Ile Ser Val Asn Val Val
 195 200 205
 Ala Asn Ala Met Ala Gln Val Met Gly
 210 215
 <210> 24
 <211> 3129
 <212> PRT
 <213> 人工的
 <220>
 <223> 合成的
 <400> 24
 Met Thr Trp Ser Thr Arg Leu Ala Leu Ser Phe Leu Phe Val Leu Cys
 1 5 10 15
 Thr Gln Ser Leu Tyr Ala Leu Ala Gln Ala Asn Thr Pro Trp Ser Ser
 20 25 30
 Lys Ala Asn Ala Asp Ala Phe Ile Asn Ser Phe Ile Ser Ala Ala Ser
 35 40 45
 Asn Thr Gly Ser Phe Ser Gln Asp Gln Met Glu Asp Met Ser Leu Ile
 50 55 60
 Gly Asn Thr Leu Met Ala Ala Met Asp Asn Met Gly Gly Arg Ile Thr
 65 70 75 80
 Pro Ser Lys Leu Gln Ala Leu Asp Met Ala Phe Ala Ser Ser Val Ala
 85 90 95
 Glu Ile Ala Ala Ser Glu Gly Gly Asp Leu Gly Val Thr Thr Asn Ala
 100 105 110
 Ile Ala Asp Ala Leu Thr Ser Ala Phe Tyr Gln Thr Thr Gly Val Val
 115 120 125
 Asn Ser Arg Phe Ile Ser Glu Ile Arg Ser Leu Ile Gly Met Phe Ala
 130 135 140

Gln	Ala	Ser	Ala	Asn	Asp	Val	Tyr	Ala	Ser	Ala	Gly	Ser	Ser	Gly	Gly
145						150				155					160
Gly	Gly	Tyr	Gly	Ala	Ser	Ser	Ala	Ser	Ala	Ala	Ser	Ala	Ser	Ala	Ala
				165					170					175	
Ala	Pro	Ser	Gly	Val	Ala	Tyr	Gln	Ala	Pro	Ala	Gln	Ala	Gln	Ile	Ser
			180					185					190		
Phe	Thr	Leu	Arg	Gly	Gln	Gln	Pro	Val	Ser	Tyr	Gly	Gln	Gly	Gly	Ala
		195					200					205			
Gly	Pro	Gly	Gly	Ala	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly
	210					215					220				
Gly	Ala	Gly	Gln	Gly	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Tyr	Gly
225				230						235					240
Gln	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Ser	Gly	Ala	Ala	Ala	Ala	Ala	Ala
				245					250					255	
Ala	Ala	Ala	Gly	Gly	Thr	Gly	Gln	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Ala
			260					265					270		
Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Ala	Gly	Gln	Gly
	275						280					285			
Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Thr	Gly
	290					295					300				
Gln	Gly	Gly	Ala	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Ala
305				310						315					320
Gly	Gln	Gly	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly
				325					330					335	
Gly	Tyr	Gly	Gln	Gly	Gly	Ser	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala
		340						345						350	
Ala	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly
	355					360					365				
Tyr	Gly	Gln	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Ala	Gly	Ala	Ala	Ala	Ala
	370					375				380					
Ala	Ala	Ala	Ala	Ala	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Tyr	Gly	Arg	Gly
385				390						395					400
Gly	Ala	Gly	Gln	Gly	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly
				405					410					415	
Ala	Gly	Gln	Gly	Gly	Tyr	Gly	Gly	Gln	Gly	Ala	Gly	Gln	Gly	Gly	Ser
		420						425					430		
Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Ala	Gly	Gln	Gly	Gly
	435						440					445			
Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Ser	Gly	Ala

450	455	460
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly		
465	470	475
Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly		480
	485	490
Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly		495
	500	505
Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly		510
	515	520
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		525
	530	535
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly		540
545	550	555
Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala		560
	565	570
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr		575
	580	585
Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala		590
	595	600
Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly		605
	610	615
Gln Gly Gly Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly		620
625	630	635
Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala		640
	645	650
Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly		655
	660	665
Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala		670
	675	680
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr		685
	690	695
Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala		700
705	710	715
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg		720
	725	730
Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala		735
	740	745
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly		750
	755	760
		765

Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly			
770	775	780	
Gly Gln Gly Asp Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala Gly			
785	790	795	800
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala			
	805	810	815
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly			
	820	825	830
Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ser Ala Ala Ala Ala			
	835	840	845
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly			
850	855	860	
Gly Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln			
865	870	875	880
Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala			
	885	890	895
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr			
	900	905	910
Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala			
	915	920	925
Thr Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln			
930	935	940	
Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala			
945	950	955	960
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly			
	965	970	975
Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly			
	980	985	990
Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala			
	995	1000	1005
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln			
1010	1015	1020	
Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln			
1025	1030	1035	
Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala			
1040	1045	1050	
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln			
1055	1060	1065	
Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala			

1070	1075	1080
Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly		
1085	1090	1095
Gln Gly Gly Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr		
1100	1105	1110
Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala		
1115	1120	1125
Ala Ser Arg Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg		
1130	1135	1140
Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala		
1145	1150	1155
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln		
1160	1165	1170
Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln		
1175	1180	1185
Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln		
1190	1195	1200
Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala		
1205	1210	1215
Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly		
1220	1225	1230
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly		
1235	1240	1245
Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly		
1250	1255	1260
Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala		
1265	1270	1275
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly		
1280	1285	1290
Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala		
1295	1300	1305
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr		
1310	1315	1320
Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala		
1325	1330	1335
Ala Ala Ala Gly Ser Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala		
1340	1345	1350
Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly		
1355	1360	1365

Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr		
1370	1375	1380
Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala		
1385	1390	1395
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr		
1400	1405	1410
Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala		
1415	1420	1425
Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly		
1430	1435	1440
Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala		
1445	1450	1455
Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln		
1460	1465	1470
Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		
1475	1480	1485
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln		
1490	1495	1500
Gly Gly Ala Gly Gln Gly Gly Ala Gly Thr Ala Ala Ala Ala Ala		
1505	1510	1515
Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln		
1520	1525	1530
Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala		
1535	1540	1545
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly		
1550	1555	1560
Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala		
1565	1570	1575
Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly		
1580	1585	1590
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly		
1595	1600	1605
Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly		
1610	1615	1620
Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala		
1625	1630	1635
Ala Ala Ala Gly Gly Ala Ser Gln Gly Gly Gln Gly Gly Tyr Gly		
1640	1645	1650
Gln Gly Asp Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala		

1655	1660	1665
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly		
1670	1675	1680
Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Gly Ala		
1685	1690	1695
Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala		
1700	1705	1710
Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Arg Gly		
1715	1720	1725
Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala		
1730	1735	1740
Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly		
1745	1750	1755
Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr		
1760	1765	1770
Gly Gln Gly Gly Thr Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala		
1775	1780	1785
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly		
1790	1795	1800
Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly		
1805	1810	1815
Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly		
1820	1825	1830
Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln		
1835	1840	1845
Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly		
1850	1855	1860
Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala		
1865	1870	1875
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly		
1880	1885	1890
Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly		
1895	1900	1905
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly		
1910	1915	1920
Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala		
1925	1930	1935
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly		
1940	1945	1950

Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly		
1955	1960	1965
Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly		
1970	1975	1980
Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala		
1985	1990	1995
Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly		
2000	2005	2010
Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala		
2015	2020	2025
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg		
2030	2035	2040
Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala		
2045	2050	2055
Ala Gly Ser Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln		
2060	2065	2070
Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		
2075	2080	2085
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln		
2090	2095	2100
Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala		
2105	2110	2115
Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln		
2120	2125	2130
Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln		
2135	2140	2145
Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		
2150	2155	2160
Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala		
2165	2170	2175
Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly		
2180	2185	2190
Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala		
2195	2200	2205
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly		
2210	2215	2220
Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala		
2225	2230	2235
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly		

2240	2245	2250
Tyr Gly Gln Gly Gly Asn Gly Gln Gly Gly Ala Gly Gln Gly Gly		
2255	2260	2265
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly		
2270	2275	2280
Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala		
2285	2290	2295
Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly		
2300	2305	2310
Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala		
2315	2320	2325
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly		
2330	2335	2340
Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala		
2345	2350	2355
Ala Ala Ala Ala Ala Gly Gly Ala Ser Gln Gly Gly Gln Gly Gly		
2360	2365	2370
Tyr Gly Gln Gly Asp Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly		
2375	2380	2385
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly		
2390	2395	2400
Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala		
2405	2410	2415
Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly		
2420	2425	2430
Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly		
2435	2440	2445
Arg Gly Gly Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly		
2450	2455	2460
Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala		
2465	2470	2475
Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly		
2480	2485	2490
Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala		
2495	2500	2505
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg		
2510	2515	2520
Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Gly Ala Gly		
2525	2530	2535

Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly		
2540	2545	2550
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Arg Gly Gly		
2555	2560	2565
Gln Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala Gly		
2570	2575	2580
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly		
2585	2590	2595
Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly		
2600	2605	2610
Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala		
2615	2620	2625
Ala Val Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala		
2630	2635	2640
Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly		
2645	2650	2655
Ser Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly		
2660	2665	2670
Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln		
2675	2680	2685
Gly Gly Gln Gly Gly Tyr Gly Gly Gly Gly Tyr Gly Gln Gly Gly		
2690	2695	2700
Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala		
2705	2710	2715
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly		
2720	2725	2730
Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala		
2735	2740	2745
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly		
2750	2755	2760
Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Thr Gly Ala Gly Gln		
2765	2770	2775
Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly Ala		
2780	2785	2790
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln		
2795	2800	2805
Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln		
2810	2815	2820
Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		

2825	2830	2835
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln		
2840	2845	2850
Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala		
2855	2860	2865
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln		
2870	2875	2880
Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln		
2885	2890	2895
Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly Ala		
2900	2905	2910
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln		
2915	2920	2925
Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln		
2930	2935	2940
Gly Gly Ala Gly Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly		
2945	2950	2955
Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly		
2960	2965	2970
Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly		
2975	2980	2985
Ala Gly Gln Gly Gly Tyr Gly Gly Tyr Gly Gln Gln Gly Gly Ala		
2990	2995	3000
Gly Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly Gln Ile Tyr Tyr		
3005	3010	3015
Gly Pro Gln Ser Val Ala Ala Pro Ala Ala Ala Ala Ala Ser Ala		
3020	3025	3030
Leu Ala Ala Pro Ala Thr Ser Ala Arg Ile Ser Ser His Ala Ser		
3035	3040	3045
Ala Leu Leu Ser Asn Gly Pro Thr Asn Pro Ala Ser Ile Ser Asn		
3050	3055	3060
Val Ile Ser Asn Ala Val Ser Gln Ile Ser Ser Ser Asn Pro Gly		
3065	3070	3075
Ala Ser Ala Cys Asp Val Leu Val Gln Ala Leu Leu Glu Leu Val		
3080	3085	3090
Thr Ala Leu Leu Thr Ile Ile Gly Ser Ser Asn Ile Gly Ser Val		
3095	3100	3105
Asn Tyr Asp Ser Ser Gly Gln Tyr Ala Gln Val Val Thr Gln Ser		
3110	3115	3120

Val Gln Asn Ala Phe Ala

3125

<210> 25

<211> 751

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 25

Met Arg Arg Met Tyr Ser Leu Ser Ile Gln Ser Asp Phe Pro Thr Thr

1 5 10 15

Thr Met Thr Trp Ser Thr Arg Leu Ala Leu Ser Phe Phe Ala Val Ile

20 25 30

Cys Thr Gln Ser Ile Tyr Ala Leu Gly Gln Gly Asn Thr Pro Trp Ser

35 40 45

Thr Lys Ala Asn Ala Asp Asn Phe Met Asn Gly Phe Leu Ser Ala Cys

50 55 60

Ala Gln Ser Gly Val Phe Ser Ala Asp Gln Val Asp Asp Met Thr Thr

65 70 75 80

Ile Gly Lys Thr Leu Met Ile Ala Met Asp Lys Met Gly Gly Lys Ile

85 90 95

Ser Ser Ser Lys Leu Gln Ala Leu Asp Met Ala Phe Ala Ser Ser Val

100 105 110

Ala Glu Ile Ala Thr Ala Glu Gly Gly Ala Asn Ile Asn Asp Ile Thr

115 120 125

Asp Ala Ile Arg Tyr Ala Leu Gln Asn Ala Phe Tyr Gln Thr Thr Gly

130 135 140

Ala Val Asn Ser Lys Phe Ile Asn Glu Ile Ser Asn Leu Ile Tyr Met

145 150 155 160

Phe Ala Gln Thr Asn Ile Asn Asp Val Asn Gly Gly Gly Gly Tyr Gly

165 170 175

Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Gly Gly

180 185 190

Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Arg Gln Gly Gly Ala

195 200 205

Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Asp Gln Gly Ala

210 215 220

Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Thr Ala Ser Gly

225 230 235 240

Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly		
245	250	255
Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr		
260	265	270
Gly Gly Gln Gly Ala Ala Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala		
275	280	285
Ala Ala Ala Gly Gly Ala Gly Leu Gly Gly Leu Gly Gly Tyr Gly Gln		
290	295	300
Gly Gly Ser Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly		
305	310	315
Glu Gly Gly Val Gly Gln Gly Gly Tyr Gly Gln Arg Gly Ala Gly Gln		
325	330	335
Gly Gly Val Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		
340	345	350
Gly Gln Gly Glu Tyr Gly Arg Gly Gly Ala Gly Lys Gly Gly Ala Ala		
355	360	365
Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly		
370	375	380
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		
385	390	395
Gly Leu Gly Gly Lys Gly Gly Tyr Gly Gln Gly Gly Ala Gly Ala Ala		
405	410	415
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly		
420	425	430
Tyr Gly Arg Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala		
435	440	445
Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Val Gly		
450	455	460
Gln Ser Gly Tyr Gly Gln Glu Gly Tyr Gly Gln Gly Gly Ala Gly Gln		
465	470	475
Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala		
485	490	495
Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Glu Ala Ala		
500	505	510
Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly		
515	520	525
Gln Gly Ala Gly Gln Gly Arg Ala Gly Ala Ala Ala Ala Ala Ala Ala		
530	535	540
Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly		

545	550	555	560
Tyr Gly Gln Gly Gly Ser Gly Ala Ala Ala Ala Ala Ala Ala Ala			
	565	570	575
Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr			
	580	585	590
Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Val Ala Ala Ala			
	595	600	605
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Pro Gln Gly Gly			
	610	615	620
Ala Gly Ala Thr Ala Ala Ser Ala Ser Gly Pro Val Gln Ile Tyr Tyr			
625	630	635	640
Gly Pro Gln Ser Val Val Ala Pro Ala Ala Ala Ala Ala Ser Ala Leu			
	645	650	655
Ala Ala Pro Ala Thr Ser Ala Arg Ile Ser Ser His Ala Ser Ala Leu			
	660	665	670
Leu Ser Asn Gly Pro Thr Asn Pro Ala Ser Ile Ser Asn Val Ile Ser			
	675	680	685
Asn Ala Val Ser Gln Ile Ser Ser Ser Asn Pro Gly Ala Ser Ala Cys			
	690	695	700
Asp Val Leu Val Gln Ala Leu Leu Glu Leu Val Thr Ala Leu Leu Thr			
705	710	715	720
Ile Ile Gly Ser Ser Asn Ile Gly Ser Val Asn Tyr Asp Ser Ser Gly			
	725	730	735
Gln Tyr Ala Gln Val Val Thr Gln Ser Val Gln Asn Ala Phe Ala			
	740	745	750
<210> 26			
<211> 748			
<212> PRT			
<213> 人工的			
<220>			
<223> 合成的			
<400> 26			
Met Arg Arg Met Tyr Ser Leu Ser Ile Gln Ser Asp Phe Pro Thr Thr			
1	5	10	15
Thr Met Thr Trp Ser Thr Arg Leu Ala Leu Ser Phe Phe Ala Val Ile			
	20	25	30
Cys Thr Gln Ser Ile Tyr Ala Leu Gly Gln Gly Asn Thr Pro Trp Ser			
	35	40	45
Thr Lys Ala Asn Ala Asp Asn Phe Met Asn Gly Phe Leu Ser Ala Cys			

50	55	60
Ala Gln Ser Gly Val Phe	Ser Ala Asp Gln Val Asp	Asp Met Thr Thr
65	70	75
Ile Gly Lys Thr Leu Met	Ile Ala Met Asp Lys Met	Gly Gly Lys Ile
85	90	95
Ser Ser Ser Lys Leu Gln	Ala Leu Asp Met Ala Phe	Ala Ser Ser Val
100	105	110
Ala Glu Ile Ala Thr Ala	Glu Gly Gly Ala Asn Ile	Asn Asp Ile Thr
115	120	125
Asp Ala Ile Arg Tyr Ala	Leu Gln Asn Ala Phe Tyr	Gln Thr Thr Gly
130	135	140
Ala Val Asn Ser Lys Phe	Ile Asn Glu Ile Ser Asn	Leu Ile Tyr Phe
145	150	155
Ala Gln Thr Asn Ile Asn	Asp Val Asn Gly Gly Gly	Gly Tyr Gly Gln
165	170	175
Gly Gly Ala Gly Gln Gly	Gly Ala Ala Ala Ala Ala	Gly Gly Ala
180	185	190
Gly Gln Gly Gly Tyr Gly	Arg Gly Gly Ala Arg Gln	Gly Gly Ala Ala
195	200	205
Ala Ala Ala Gly Ala Gly	Gln Gly Gly Tyr Gln Gly	Ala Gly Gln Gly
210	215	220
Gly Ala Gly Ala Ala Ala	Ala Ala Ala Thr Ala Ser	Gly Gly Ala Gly
225	230	235
Gln Gly Gly Tyr Gly Arg	Gly Gly Ala Gly Gln Gly	Gly Ala Ala Ala
245	250	255
Ala Ala Ala Ala Ala Ala	Gly Ala Gly Gln Gly Gly	Tyr Gly Gly Gln
260	265	270
Gly Ala Ala Gln Gly Gly	Ala Gly Ala Ala Ala Ala	Ala Ala Ala Ala
275	280	285
Gly Gly Ala Gly Leu Gly	Gly Leu Gly Gly Tyr Gly	Gln Gly Gly Ser
290	295	300
Gly Ala Ala Ala Ala Ala	Gly Gly Ala Gly Gln Gly	Gly Glu Gly Gly
305	310	315
Val Gly Gln Gly Gly Tyr	Gly Gln Arg Gly Ala Gly	Gln Gly Gly Val
325	330	335
Gly Ala Ala Ala Ala Ala	Ala Ala Ala Gly Gly Ala	Gly Gln Gly
340	345	350
Glu Tyr Gly Arg Gly Gly	Ala Gly Lys Gly Gly Ala	Ala Ala Ala Ala
355	360	365

Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Gly			
370	375	380	
Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Leu Gly			
385	390	395	400
Gly Lys Gly Gly Tyr Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala			
405	410	415	
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Arg			
420	425	430	
Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala			
435	440	445	
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Val Gly Gln Ser Gly			
450	455	460	
Tyr Gly Gln Glu Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala			
465	470	475	480
Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly			
485	490	495	
Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Glu Ala Ala Ala Ala Ala			
500	505	510	
Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala			
515	520	525	
Gly Gln Gly Arg Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly			
530	535	540	
Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln			
545	550	555	560
Gly Gly Ser Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala			
565	570	575	
Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly			
580	585	590	
Gly Ala Gly Gln Gly Gly Ala Ala Ala Val Ala Ala Ala Ala Ala Ala			
595	600	605	
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Pro Gln Gly Gly Ala Gly Ala			
610	615	620	
Thr Ala Ala Ser Ala Ser Gly Pro Val Gln Ile Tyr Tyr Gly Pro Gln			
625	630	635	640
Ser Val Val Ala Pro Ala Ala Ala Ala Ala Ser Ala Leu Ala Ala Pro			
645	650	655	
Ala Thr Ser Ala Arg Ile Ser Ser His Ala Ser Ala Leu Leu Ser Asn			
660	665	670	
Gly Pro Thr Asn Pro Ala Ser Ile Ser Asn Val Ile Ser Asn Ala Val			

675	680	685
Ser Gln Ile Ser Ser Ser Asn Pro Gly Ala Ser Ala Cys Asp Val Leu		
690	695	700
Val Gln Ala Leu Leu Glu Leu Val Thr Ala Leu Leu Thr Ile Ile Gly		
705	710	715
Ser Ser Asn Ile Gly Ser Val Asn Tyr Asp Ser Ser Gly Gln Tyr Ala		
725	730	735
Gln Val Val Thr Gln Ser Val Gln Asn Ala Phe Ala		
740	745	
<210> 27		
<211> 631		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的		
<400> 27		
Ser Arg Asp Phe Pro Thr Thr Thr Met Thr Trp Ser Thr Arg Leu Ala		
1	5	10
Leu Ser Phe Leu Phe Val Leu Cys Thr Gln Ser Leu Tyr Ala Leu Ala		
20	25	30
Gln Ala Asn Thr Pro Trp Ser Ser Lys Ala Asn Ala Asp Ala Phe Ile		
35	40	45
Asn Ser Phe Ile Ser Ala Ala Ser Asn Thr Gly Ser Phe Ser Gln Asp		
50	55	60
Gln Met Glu Asp Met Ser Leu Ile Gly Asn Thr Leu Met Ala Ala Met		
65	70	75
Asp Asn Met Gly Gly Arg Ile Thr Pro Ser Lys Leu Gln Ala Leu Asp		
85	90	95
Met Ala Phe Ala Ser Ser Val Ala Glu Ile Ala Ala Ser Glu Gly Gly		
100	105	110
Asp Leu Gly Val Thr Thr Asn Ala Ile Ala Asp Ala Leu Thr Ser Ala		
115	120	125
Phe Tyr Gln Thr Thr Gly Val Val Asn Ser Arg Phe Ile Ser Glu Ile		
130	135	140
Arg Ser Leu Ile Gly Met Phe Ala Gln Ala Ser Ala Asn Asp Val Tyr		
145	150	155
Ala Ser Ala Gly Ser Ser Gly Gly Gly Gly Tyr Gly Ala Ser Ser Ala		
165	170	175
Ser Ala Ala Ser Ala Ser Ala Ala Ala Pro Ser Gly Val Ala Tyr Gln		

180	185	190
Ala Pro Ala Gln Ala Gln Ile Ser Phe Thr Leu Arg Gly Gln Gln Pro		
195	200	205
Val Ser Tyr Gly Gln Gly Gly Ala Gly Pro Gly Gly Ala Gly Ala Ala		
210	215	220
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly		
225	230	235
Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly		
245	250	255
Ser Gly Ala Ala Ala Ala Ala Gly Gly Thr Gly Gln Gly Gly Ala Gly		
260	265	270
Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly		
275	280	285
Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln		
290	295	300
Gly Gly Thr Gly Gln Gly Gly Ala Gly Ala Ala Ala Arg Gly Gly Tyr		
305	310	315
Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala		
325	330	335
Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln		
340	345	350
Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala		
355	360	365
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly		
370	375	380
Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala		
385	390	395
Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly		
405	410	415
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln		
420	425	430
Gly Gly Tyr Gly Arg Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly		
435	440	445
Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly		
450	455	460
Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly		
465	470	475
Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly		
485	490	495

Gly Ala Gly Gln Gly Gly Tyr Gly Gly Tyr Gly Gln Gln Gly Gly Ala
 500 505 510
 Gly Ala Ala Ala Ala Ala Ser Gly Pro Gly Gln Ile Tyr Tyr Gly
 515 520 525
 Pro Gln Ser Val Ala Ala Pro Ala Ala Ala Ala Ala Ser Ala Leu Ala
 530 535 540
 Ala Pro Ala Thr Ser Ala Arg Ile Ser Ser His Ala Ser Ala Leu Leu
 545 550 555 560
 Ser Asn Gly Pro Thr Asn Pro Ala Ser Ile Ser Asn Val Ile Ser Asn
 565 570 575
 Ala Val Ser Gln Ile Ser Ser Ser Asn Pro Gly Ala Ser Ala Cys Asp
 580 585 590
 Val Leu Val Gln Ala Leu Leu Glu Leu Val Thr Ala Leu Leu Thr Ile
 595 600 605
 Ile Gly Ser Ser Asn Ile Gly Ser Val Asn Tyr Asp Ser Ser Gly Gln
 610 615 620
 Tyr Ala Gln Val Val Thr Gln
 625 630

<210> 28

<211> 697

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 28

Met Tyr Ser Leu Ser Ile Gln Ser Asp Phe Pro Thr Thr Thr Met Thr
 1 5 10 15
 Trp Ser Thr Arg Leu Ala Leu Ser Phe Phe Ala Val Ile Cys Thr Gln
 20 25 30
 Ser Ile Tyr Ala Leu Gly Gln Gly Asn Thr Pro Trp Ser Thr Lys Ala
 35 40 45
 Asn Ala Asp Asn Phe Met Asn Gly Phe Leu Ser Ala Cys Ala Gln Ser
 50 55 60
 Gly Val Phe Ser Ala Asp Gln Val Asp Asp Met Thr Thr Ile Gly Lys
 65 70 75 80
 Thr Leu Met Ile Ala Met Asp Lys Met Gly Gly Lys Ile Ser Ser Ser
 85 90 95
 Lys Leu Gln Ala Leu Asp Met Ala Phe Ala Ser Ser Val Ala Glu Ile
 100 105 110

Ala Thr Ala Glu Gly Gly Ala Asn Ile Asn Asp Ile Thr Asp Ala Ile		
115	120	125
Arg Tyr Ala Leu Gln Asn Ala Phe Tyr Gln Thr Thr Gly Ala Val Asn		
130	135	140
Ser Lys Phe Ile Asn Glu Ile Ser Asn Leu Ile Tyr Met Phe Ala Gln		
145	150	155
Thr Asn Ile Asn Asp Val Asn Gly Gly Gly Gly Tyr Gly Gln Gly Gly		
165	170	175
Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln		
180	185	190
Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala		
195	200	205
Ala Gly Ala Gly Gln Gly Gly Tyr Gly Asp Gln Gly Ala Gly Gln Gly		
210	215	220
Gly Ala Gly Ala Ala Ala Ala Ala Ala Thr Ala Ser Gly Gly Ala Gly		
225	230	235
Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Glu Ala Ala		
245	250	255
Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln		
260	265	270
Glu Ala Ala Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala		
275	280	285
Gly Gly Ala Gly Leu Gly Gly Leu Gly Gly Tyr Gly Gln Gly Gly Ser		
290	295	300
Gly Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Glu Gly Gly		
305	310	315
Val Gly Gln Gly Gly Tyr Gly Gln Arg Gly Ala Gly Gln Gly Gly Ala		
325	330	335
Gly Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Glu Tyr Gly		
340	345	350
Arg Gly Gly Ala Gly Gln Gly Arg Ala Ala Ala Ala Ala Ala Ala Ala		
355	360	365
Ala Gly Ala Gly Gln Gly Ser Tyr Gly Gly Gln Gly Ala Gly Gly Tyr		
370	375	380
Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala		
385	390	395
Ala Ala Ala Gly Gly Ala Gly Gln Val Gly Gln Ser Gly Tyr Gly Gln		
405	410	415
Glu Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala		

420	425	430
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly		
435	440	445
Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala Ala		
450	455	460
Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln Gly Ala Gly Gln Gly		
465	470	475
Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Trp Ala Gly Gln		
485	490	495
Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ser		
500	505	510
Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly		
515	520	525
Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Asp Gln Gly Gly Ala Gly		
530	535	540
Gln Gly Gly Ala Ala Ala Val Ala Ala Ala Ala Ala Gly Gly Ala		
545	550	555
Gly Gln Gly Gly Tyr Gly Pro Gln Gly Gly Ala Gly Ala Ala Ala Ala		
565	570	575
Ser Ala Ser Gly Pro Val Gln Ile Tyr Tyr Gly Pro Gln Ser Val Ala		
580	585	590
Ala Pro Ala Ala Ala Ala Ala Ser Ala Leu Ala Ala Pro Ala Thr Ser		
595	600	605
Ala Arg Ile Ser Ser His Ala Ser Ala Leu Leu Ser Asn Gly Pro Thr		
610	615	620
Asn Pro Ala Ser Ile Ser Asn Val Ile Ser Asn Ala Val Ser Gln Ile		
625	630	635
Ser Ser Ser Asn Pro Gly Ala Ser Ala Cys Asp Val Leu Val Gln Ala		
645	650	655
Leu Leu Glu Leu Val Thr Ala Leu Leu Thr Ile Ile Gly Ser Ser Asn		
660	665	670
Ile Gly Ser Val Asn Tyr Asp Ser Ser Gly Gln Tyr Ala Gln Val Val		
675	680	685
Thr Gln Ser Val Gln Asn Ala Phe Ala		
690	695	

<210> 29

<211> 675

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 29

Met	Thr	Trp	Ser	Thr	Arg	Leu	Ala	Leu	Ser	Phe	Leu	Leu	Val	Leu	Cys
1				5					10					15	
Thr	Gln	Ser	Ile	Tyr	Ala	Leu	Ala	Gln	Ala	Asn	Thr	Pro	Trp	Ser	Ser
			20					25					30		
Lys	Ala	Asn	Ala	Asp	Ala	Phe	Ile	Asn	Ser	Phe	Ile	Ser	Ala	Ala	Ser
		35				40						45			
Asn	Thr	Gly	Ser	Phe	Ser	Gln	Asp	Gln	Met	Glu	Asp	Met	Ser	Leu	Ile
	50					55					60				
Gly	Asn	Thr	Leu	Met	Ala	Ala	Met	Asp	Asn	Met	Gly	Gly	Arg	Ile	Thr
65					70					75					80
Pro	Ser	Lys	Leu	Gln	Ala	Leu	Asp	Met	Ala	Phe	Ala	Ser	Ser	Val	Ala
				85					90						95
Glu	Ile	Ala	Ala	Ser	Glu	Gly	Gly	Asp	Leu	Gly	Val	Thr	Thr	Asn	Ala
			100					105						110	
Ile	Ala	Asp	Ala	Leu	Thr	Ser	Ala	Phe	Tyr	Gln	Thr	Thr	Gly	Val	Val
		115					120						125		
Asn	Ser	Arg	Phe	Ile	Ser	Glu	Ile	Arg	Ser	Leu	Ile	Gly	Met	Phe	Ala
		130					135					140			
Gln	Ala	Ser	Ala	Asn	Asp	Val	Tyr	Ala	Ser	Ala	Gly	Ser	Ser	Gly	Gly
145					150						155				160
Gly	Gly	Tyr	Gly	Ala	Ser	Ser	Ala	Ser	Ala	Ala	Ser	Ala	Ser	Ala	Ala
				165						170					175
Ala	Pro	Ser	Gly	Val	Ala	Tyr	Gln	Ala	Pro	Ala	Gln	Ala	Gln	Ile	Ser
			180					185						190	
Phe	Ser	Leu	Arg	Gly	Gln	Gln	Pro	Val	Gly	Tyr	Gly	Gln	Gly	Gly	Ala
		195					200						205		
Ser	Ala	Ala	Ser	Gly	Ala	Ala	Ala	Gly	Gln	Gly	Gly	Ala	Gly	Pro	Gly
		210					215					220			
Gly	Ala	Gly	Ala	Ala	Ala	Ala	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Gln	Gly
225					230					235					240
Gly	Tyr	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Ala	Gly	Gln	Gly	Gly
				245						250					255
Ser	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Thr	Gly	Gln
			260						265					270	
Gly	Gly	Ala	Gly	Gln	Gly	Gly	Ala	Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala
		275						280							285

Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly			
290	295	300	
Gly Tyr Gly Gln Gly Gly Ser Gly Ala Ala Ala Ala Ala Ala Ala			
305	310	315	320
Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Asp Tyr Gly Gln Gly Gly			
	325	330	335
Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala			
	340	345	350
Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly			
	355	360	365
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln			
370	375	380	
Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ser Gly Ala Ala			
385	390	395	400
Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly			
	405	410	415
Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ser Gly Ala Ala Ala			
	420	425	430
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly			
	435	440	445
Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala			
450	455	460	
Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly			
465	470	475	480
Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala			
	485	490	495
Ala Ala Ala Ala Gly Ala Gly Gln Arg Gly Tyr Gly Gly Gln Gly Ala			
	500	505	510
Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly			
	515	520	525
Ala Gly Gln Gly Gly Gln Arg Gly Tyr Gly Gln Gly Gly Tyr Gly Gln			
530	535	540	
Gly Gly Ser Gly Ala Ala Ala Ala Ala Ala Ala Ser Gly Pro Gly			
545	550	555	560
Gln Val Tyr Tyr Gly Pro Gln Ser Phe Ala Ala Pro Ala Ala Ala Ala			
	565	570	575
Ala Ser Ala Leu Ser Ala Pro Ala Thr Ser Ala Arg Ile Ser Ser His			
	580	585	590
Ala Ser Ala Leu Leu Ser Ser Gly Pro Thr Asn Pro Ala Ser Ile Ser			

595	600	605
Asn Val Ile Ser Asn Ala Val Ser Gln Ile Ser Ser Ser Asn Pro Gly		
610	615	620
Ala Ser Ala Cys Asp Val Leu Val Gln Ala Leu Leu Glu Leu Val Thr		
625	630	635
Ala Leu Leu Thr Ile Ile Gly Ser Ser Asn Ile Gly Ser Val Asn Tyr		
645	650	655
Asp Ser Ser Gly Gln Tyr Ala Gln Val Val Thr Gln Ser Val Gln Asn		
660	665	670
Val Phe Gly		
675		
<210> 30		
<211> 605		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的		
<400> 30		
Met Thr Trp Ser Thr Arg Leu Ala Leu Ser Phe Leu Phe Val Leu Cys		
1	5	10
Thr Gln Ser Leu Tyr Ala Leu Ala Gln Ala Asn Thr Pro Trp Ser Ser		
20	25	30
Lys Ala Asn Ala Asp Ala Phe Ile Asn Ser Phe Ile Ser Ala Ala Ser		
35	40	45
Asn Thr Gly Ser Phe Ser Gln Asp Gln Met Glu Asp Met Ser Leu Ile		
50	55	60
Gly Asn Thr Leu Met Ala Ala Met Asp Asn Met Gly Gly Arg Ile Thr		
65	70	75
Pro Ser Lys Leu Gln Ala Leu Asp Met Ala Phe Ala Ser Ser Val Ala		
85	90	95
Glu Ile Ala Ala Ser Glu Gly Gly Asp Leu Gly Val Thr Thr Asn Ala		
100	105	110
Ile Ala Asp Ala Leu Thr Ser Ala Phe Tyr Gln Thr Thr Gly Val Val		
115	120	125
Asn Ser Arg Phe Ile Ser Glu Ile Arg Ser Leu Ile Gly Met Phe Ala		
130	135	140
Gln Ala Ser Ala Asn Asp Val Tyr Ala Ser Ala Gly Ser Ser Gly Gly		
145	150	155
Gly Gly Tyr Gly Ala Ser Ser Ala Ser Ala Ala Ser Ala Ser Ala Ala		

	165		170		175
Ala Pro Ser Gly Val Ala Tyr Gln	Ala Pro Ala Gln Ala Gln Ile Ser				
180	185	190			
Phe Thr Leu Arg Gly Gln Gln Pro Val Ser Tyr Gly Gln Gly Gly Ala					
195	200	205			
Gly Pro Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly					
210	215	220			
Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly					
225	230	235	240		
Gln Gly Gly Ala Gly Gln Gly Gly Ser Gly Ala Ala Ala Ala Ala Gly					
245	250	255			
Gly Thr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala					
260	265	270			
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly					
275	280	285			
Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Thr Gly Gln Gly Gly Ala					
290	295	300			
Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly					
305	310	315	320		
Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ser Gly Ala					
325	330	335			
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln					
340	345	350			
Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly					
355	360	365			
Gly Ala Gly Ala Ala Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly					
370	375	380			
Gln Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala Ala Ala Ala Ala					
385	390	395	400		
Gly Gly Ala Gly Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly					
405	410	415			
Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly					
420	425	430			
Tyr Gly Gly Gln Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala Ala					
435	440	445			
Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly					
450	455	460			
Arg Gly Gly Tyr Asp Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala					
465	470	475	480		

Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln		
485	490	495
Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly		
500	505	510
Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln		
515	520	525
Gly Gly Tyr Gly Gly Tyr Gly Gln Gln Gly Gly Ala Gly Ala Ala Ala		
530	535	540
Ala Ala Ala Ser Gly Pro Gly Gln Ile Tyr Tyr Gly Pro Gln Ser Val		
545	550	555
Ala Ala Pro Ala Ala Ala Ala Ala Ser Ala Leu Ala Ala Pro Ala Thr		
565	570	575
Ser Ala Arg Ile Ser Ser His Ala Ser Ala Leu Leu Ser Asn Gly Pro		
580	585	590
Thr Asn Pro Ala Ser Ile Ser Asn Val Ile Ser Asn Ala		
595	600	605

<210> 31

<211> 729

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 31

Met Thr Trp Ser Thr Arg Leu Ala Leu Ser Phe Leu Phe Val Leu Cys		
1	5	10
Thr Gln Ser Leu Tyr Ala Leu Ala Gln Ala Asn Thr Pro Trp Ser Ser		
20	25	30
Lys Ala Asn Ala Asp Ala Phe Ile Asn Ser Phe Ile Ser Ala Ala Ser		
35	40	45
Asn Thr Gly Ser Phe Ser Gln Asp Gln Met Glu Asp Met Ser Leu Ile		
50	55	60
Gly Asn Thr Leu Met Ala Ala Met Asp Asn Met Gly Gly Arg Ile Thr		
65	70	75
Pro Ser Lys Leu Gln Ala Leu Asp Met Ala Phe Ala Ser Ser Val Ala		
85	90	95
Glu Ile Ala Ala Ser Glu Gly Gly Asp Leu Gly Val Thr Thr Asn Ala		
100	105	110
Ile Ala Asp Ala Leu Thr Ser Ala Phe Tyr Gln Thr Thr Gly Val Val		
115	120	125

Asn Ser Arg Phe Ile Ser Glu Ile Arg Ser Leu Ile Gly Met Phe Ala		
130	135	140
Gln Ala Ser Ala Asn Asp Val Tyr Ala Ser Ala Gly Ser Ser Gly Gly		
145	150	155
Gly Gly Tyr Gly Ala Ser Ser Ala Ser Ala Ala Ser Ala Ser Ala Ala		
	165	170
Ala Pro Ser Gly Val Ala Tyr Gln Ala Pro Ala Gln Ala Gln Ile Ser		
	180	185
Phe Thr Leu Arg Gly Gln Gln Pro Val Ser Tyr Gly Gln Gly Gly Ala		
	195	200
Gly Pro Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly		
	210	215
Gly Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly		
225	230	235
Gln Gly Gly Ala Gly Gln Gly Gly Ser Gly Ala Ala Ala Ala Ala Gly		
	245	250
Gly Thr Gly Gln Gly Gly Ala Gly Gln Gly Gly Ala Gly Ala Ala Ala		
	260	265
Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln Gly Gly		
	275	280
Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Thr Gly Gln Gly Gly Ala		
	290	295
Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly		
305	310	315
Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ser Gly Ala		
	325	330
Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln		
	340	345
Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln Gly Gly Ala Gly Gln Gly		
	355	360
Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly		
	370	375
Gln Gly Gly Tyr Gly Arg Gly Gly Ala Gly Gln Gly Gly Ala Ala Ala		
385	390	395
Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly Gly Tyr Gly Gly Gln		
	405	410
Gly Ala Gly Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly		
	420	425
Ala Gly Gln Gly Gly Gln Gly Gly Tyr Gly Gln Gly Gly Tyr Gly Gln		
		430

435	440	445
Gly Gly Ala Gly Gln Gly Gly	Ala Ala Ala Ala Ala Ala	Ala Ala Ala Gly
450	455	460
Gly Ala Gly Gln Gly Gly Tyr	Gly Arg Gly Gly Ala Gly	Gln Gly Gly
465	470	475
Ala Ala Ala Ala Ala Ala Ala	Ala Ala Gly Ala Gly Gln	Gly Gly Tyr
485	490	495
Gly Gly Gln Gly Ala Gly Gln	Gly Gly Ala Gly Ala Ala	Ala Ala Ala
500	505	510
Ala Ala Ala Gly Gly Ala Gly	Gln Gly Gly Gln Gly Gly	Tyr Gly Arg
515	520	525
Gly Gly Tyr Asp Gln Gly Gly	Ala Gly Gln Gly Gly Ala	Gly Ala Ala
530	535	540
Ala Ala Ala Ala Ala Ala Ala	Gly Gly Ala Gly Gln Gly	Gly Gln Gly
545	550	555
Gly Tyr Gly Gln Gly Gly Tyr	Gly Gln Gly Gly Ala Gly	Gln Gly Gly
565	570	575
Ala Ala Ala Ala Ala Ala Ala	Ala Ala Gly Gly Ala Gly	Gln Gly
580	585	590
Gly Tyr Gly Gly Tyr Gly Gln	Gln Gly Gly Ala Gly Ala	Ala Ala Ala
595	600	605
Ala Ala Ser Gly Pro Gly Gln	Ile Tyr Tyr Gly Pro Gln	Ser Val Ala
610	615	620
Ala Pro Ala Ala Ala Ala Ala	Ser Ala Leu Ala Ala Pro	Ala Thr Ser
625	630	635
Ala Arg Ile Ser Ser His Ala	Ser Ala Leu Leu Ser Asn	Gly Pro Thr
645	650	655
Asn Pro Ala Ser Ile Ser Asn	Val Ile Ser Asn Ala Val	Ser Gln Ile
660	665	670
Ser Ser Ser Asn Pro Gly Ala	Ser Ala Cys Asp Val Leu	Val Gln Ala
675	680	685
Leu Leu Glu Leu Val Thr Ala	Leu Leu Thr Ile Ile Gly	Ser Ser Asn
690	695	700
Ile Gly Ser Val Asn Tyr Asp	Ser Ser Gly Gln Tyr Ala	Gln Val Val
705	710	715
Thr Gln Ser Val Gln Asn Ala	Phe Ala	
725		

<210> 32

<211> 285

<212> PRT

<213> 人工的

<220>

<223> 合成的

<400> 32

Met	Thr	Trp	Ser	Thr	Arg	Leu	Ala	Leu	Ser	Phe	Leu	Leu	Val	Leu	Cys
1				5					10					15	
Thr	Gln	Ser	Ile	Tyr	Ala	Leu	Ala	Gln	Ala	Asn	Thr	Pro	Trp	Ser	Ser
			20					25					30		
Lys	Ala	Asn	Ala	Asp	Ala	Phe	Ile	Gly	Ser	Phe	Ile	Ser	Ser	Ala	Gln
		35					40					45			
Asn	Thr	Gly	Ala	Phe	Ser	Thr	Asp	Gln	Met	Asp	Asp	Met	Ser	Leu	Ile
	50					55				60					
Gly	Asn	Thr	Leu	Met	Ala	Ala	Met	Asp	Asn	Met	Gly	Gly	Arg	Ile	Thr
65				70				75						80	
Pro	Ser	Lys	Leu	Gln	Ala	Leu	Asp	Met	Ala	Phe	Ala	Ser	Ser	Val	Ala
			85					90						95	
Glu	Ile	Ala	Ala	Ala	Glu	Gly	Gly	Asp	Leu	Gly	Val	Thr	Thr	Asn	Ala
		100						105						110	
Ile	Ala	Asp	Ala	Leu	Thr	Ser	Ala	Phe	Tyr	Gln	Thr	Thr	Gly	Val	Val
		115					120						125		
Asn	Ser	Arg	Phe	Ile	Ser	Glu	Ile	Arg	Ser	Leu	Ile	Asn	Met	Phe	Ala
		130					135					140			
Gln	Ala	Ser	Ala	Asn	Asp	Val	Tyr	Ala	Ser	Ala	Gly	Ser	Ser	Gly	Gly
145				150						155				160	
Gly	Gly	Tyr	Gly	Ala	Ala	Ser	Ser	Ser	Ala	Ser	Ala	Ala	Ala	Ser	Ser
			165					170						175	
Ser	Ala	Ser	Ala	Ala	Ala	Pro	Ser	Gly	Val	Ser	Tyr	Gln	Ala	Pro	Ala
		180						185					190		
Gln	Ala	Gln	Ile	Ser	Phe	Ser	Leu	Thr	Arg	Gln	Gln	Gln	Pro	Val	Asn
		195					200					205			
Tyr	Gly	Gln	Ser	Gly	Ala	Ser	Ala	Ala	Ser	Ala	Ala	Ala	Ala	Gly	Gly
	210					215						220			
Ala	Gly	Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Ala
225				230						235				240	
Gly	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Gly	Ala	Gly	Gln	Gly	Gly
			245						250					255	
Gln	Gly	Gly	Tyr	Gly	Gln	Gly	Gly	Ala	Gly	Gln	Gly	Gly	Ala	Ala	Ala
		260						265						270	

Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Gln
 275 280 285

<210> 33

<211> 22

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<220>

<221> MISC_FEATURE

<222> (1) .. (1)

<223> X是A或G

<220>

<221> MISC_FEATURE

<222> (2) .. (2)

<223> X是SG或GG

<220>

<221> MISC_FEATURE

<222> (6) .. (6)

<223> X是GQ或不存在

<220>

<221> MISC_FEATURE

<222> (11) .. (11)

<223> X是G或不存在

<220>

<221> MISC_FEATURE

<222> (13) .. (13)

<223> X是G或S

<220>

<221> MISC_FEATURE

<222> (17) .. (17)

<223> X是SP或GR或不存在

<400> 33

Xaa Xaa Ser Gly Pro Xaa Gly Gly Tyr Gly Xaa Pro Xaa Gln Gly Pro
 1 5 10 15
 Xaa Gly Gly Tyr Gly Pro
 20

<210> 34

<211> 6

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 34

Pro Gly Gly Tyr Gly Pro

1 5

<210> 35

<211> 36

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 35

Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro

1 5 10 15

Gly Gly Tyr Gly Pro Gly Gly Pro Gly Ser Ser Ala Gly Ala Gly Ala

20 25 30

Gly Ala Ala Ala

35

<210> 36

<211> 36

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 36

Ser Gly Pro Gly Gly Tyr Gly Pro Gly Ser Gln Gly Pro Ser Gly Pro

1 5 10 15

Gly Gly Tyr Gly Pro Gly Gly Pro Gly Ser Ser Ala Gly Ala Gly Ala

20 25 30

Gly Ala Gly Ala

35

<210> 37

<211> 29

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 37

Ser Gly Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gly Gln Gly Pro Gly
 1 5 10 15
 Gly Tyr Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly
 20 25

<210> 38

<211> 36

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 38

Ser Gly Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gly Gln Gly Pro Gly
 1 5 10 15
 Gly Tyr Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ala Ala
 20 25 30
 Ala Ala Ala Ala
 35

<210> 39

<211> 29

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 39

Gly Gly Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gly Gln Gly Pro Gly
 1 5 10 15
 Gly Tyr Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly
 20 25

<210> 40

<211> 36

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 40

Gly Gly Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gly Gln Gly Pro Gly
 1 5 10 15
 Gly Tyr Gly Pro Gly Gly Gln Gly Pro Tyr Gly Pro Gly Ala Ala Ala

20	25	30
Ala Ala Ala Ala		
35		
<210> 41		
<211> 27		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的肽		
<400> 41		
Ser Gly Pro Gly Gln Gly Gly Tyr Gly Gly Pro Gly Gly Gln Gly Pro		
1 5 10 15		
Gly Arg Gly Gly Tyr Gly Pro Gly Ala Gly Ser		
20 25		
<210> 42		
<211> 36		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的肽		
<400> 42		
Ser Gly Pro Gly Gln Gly Gly Tyr Gly Gly Pro Gly Gly Gln Gly Pro		
1 5 10 15		
Gly Arg Gly Gly Tyr Gly Pro Gly Ala Gly Ser Ala Ala Ala Ala Ala		
20 25 30		
Ala Ala Ala Ala		
35		
<210> 43		
<211> 37		
<212> PRT		
<213> 人工的		
<220>		
<223> 合成的肽		
<400> 43		
Ala Arg Thr Ile Phe Ile Cys Ile Ala Leu Gly Gly Pro Gly Gln Gly		
1 5 10 15		
Gly Tyr Gly Gly Pro Gly Gly Gln Gly Pro Gly Arg Gly Gly Tyr Gly		
20 25 30		
Pro Gly Ala Gly Ser		

35

<210> 44

<211> 27

<212> PRT

<213> 人工的

<220>

<223> 合成的肽

<400> 44

Gly Gly Pro Gly Gln Gly Gly Tyr Gly Gly Pro Gly Gly Gln Gly Pro

1

5

10

15

Gly Arg Gly Gly Tyr Gly Pro Gly Ala Gly Ser

20

25

<210> 45

<211> 108

<212> DNA

<213> 人工的

<220>

<223> 合成的

<400> 45

tccggacccg gtggctacgg acctggttct cagggcccca gcggaccagg ttgctacgga 60

cctggtggcc caggatcctc agctggtgcc ggcgctggtg ccgctgca 108

<210> 46

<211> 108

<212> DNA

<213> 人工的

<220>

<223> 合成的

<400> 46

tccggttctg gccccggagg ttacggcct ggtggccagg gaccaggtgg ctacggaccg 60

ggtggccaag gaccctacgg tcctggcgct gccgctgccg cggtgca 108

<210> 47

<211> 108

<212> DNA

<213> 人工的

<220>

<223> 合成的

<400> 47

tccggtccag gccaaggagg ttacggcgga cctggtggcc agggacccgg tcgtggcgga 60

tacggtccgg gcgctggatc tgctgccgcg gcagctgccg cggtgca 108

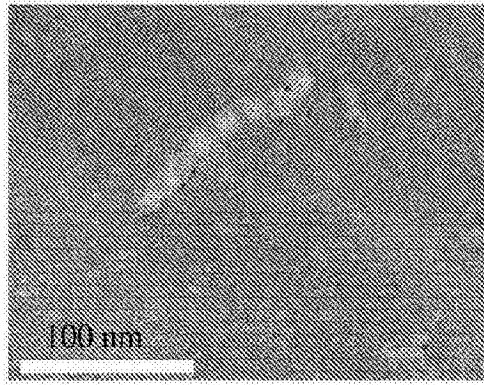


图1A

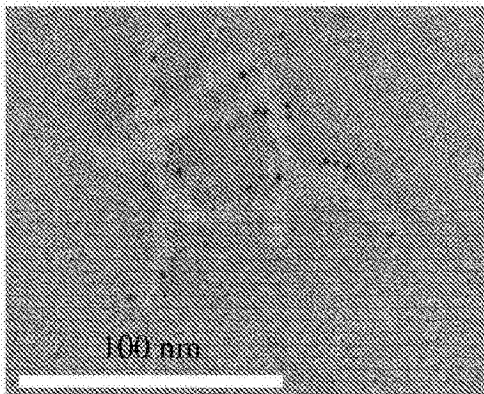


图1B

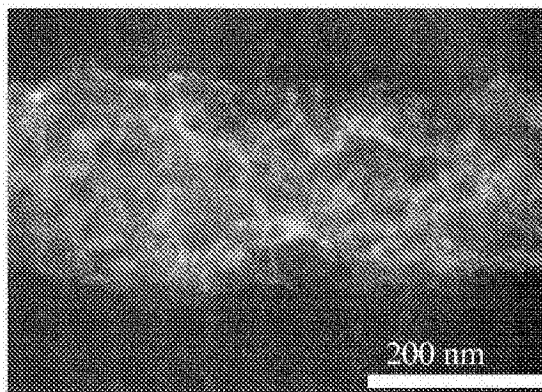


图1C

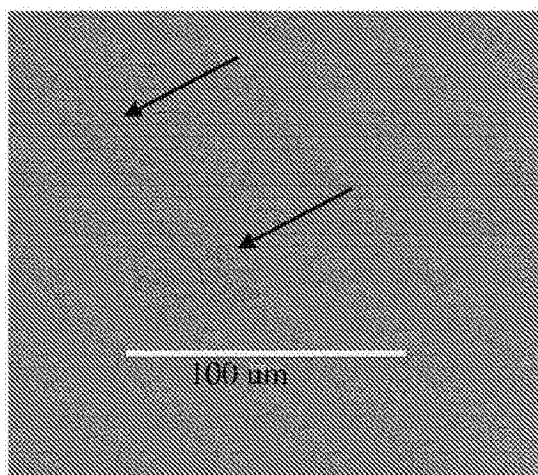


图1D

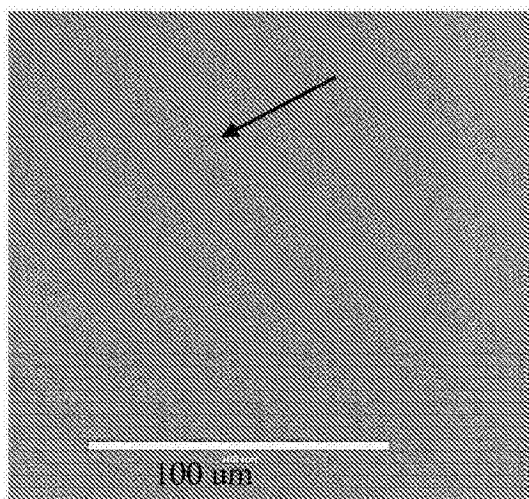


图1E

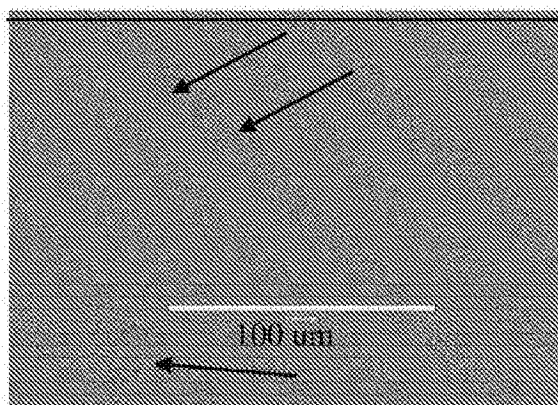


图1F

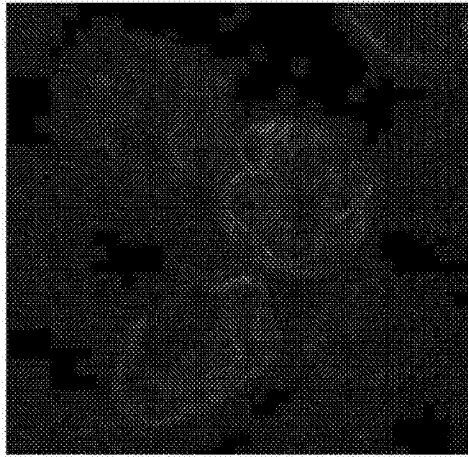


图1G

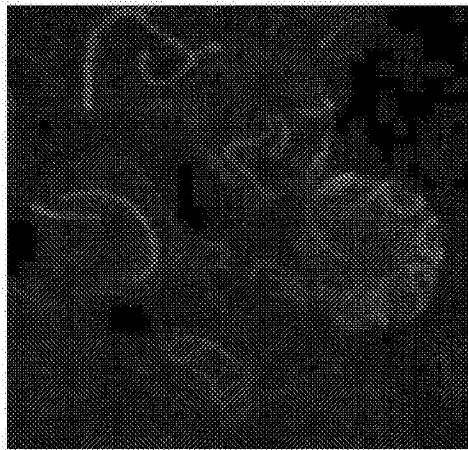


图1H

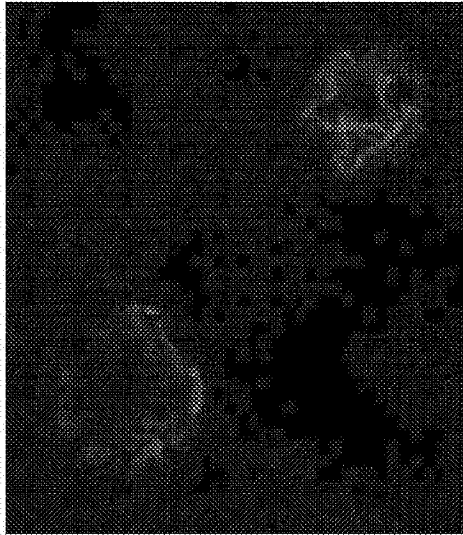


图1I

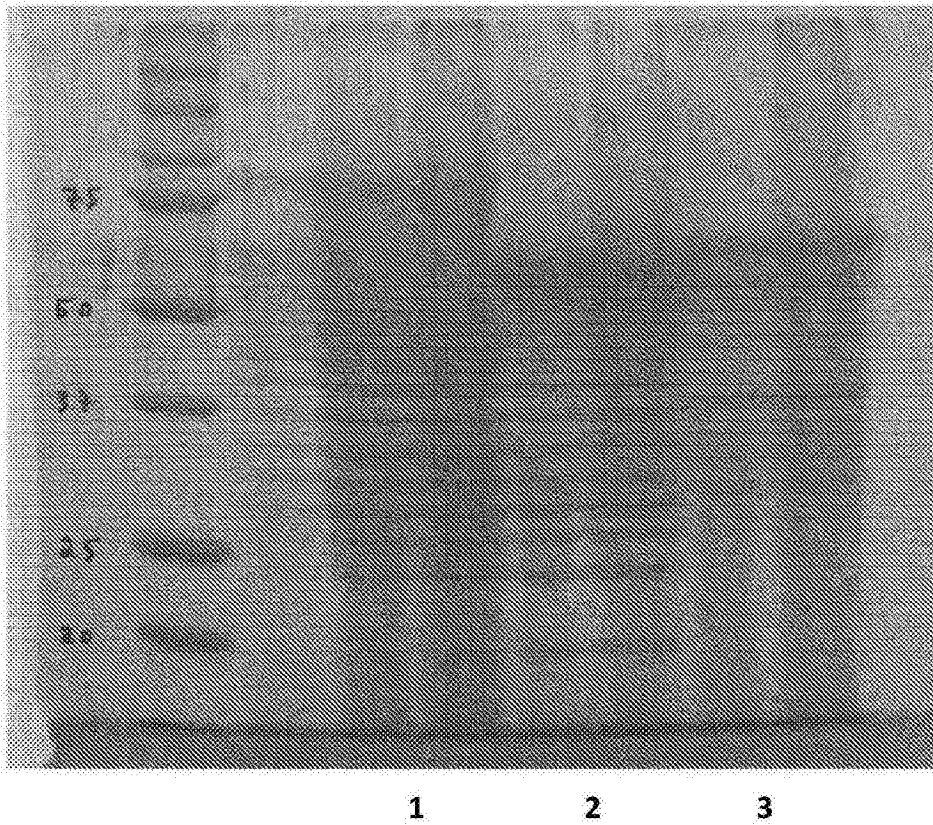


图2A

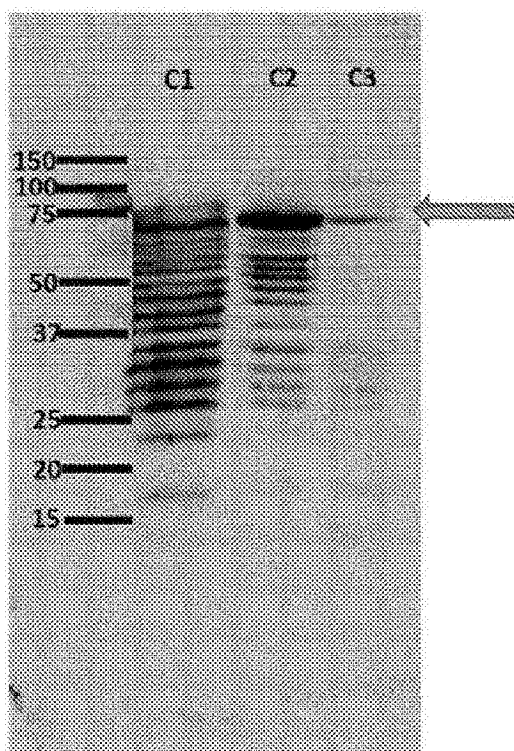


图2B

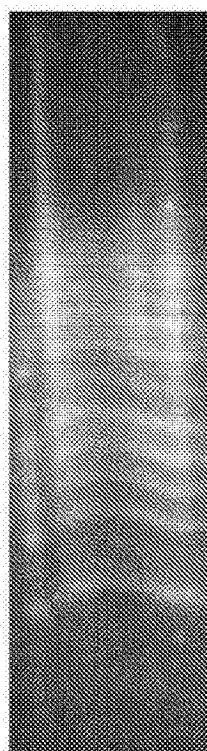


图2C

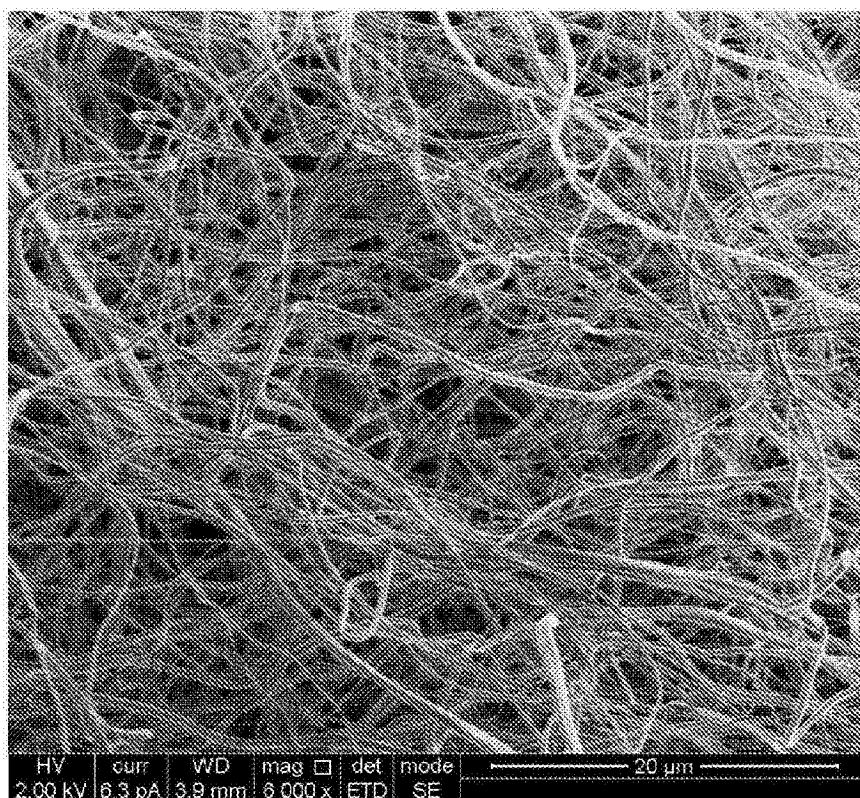


图3A

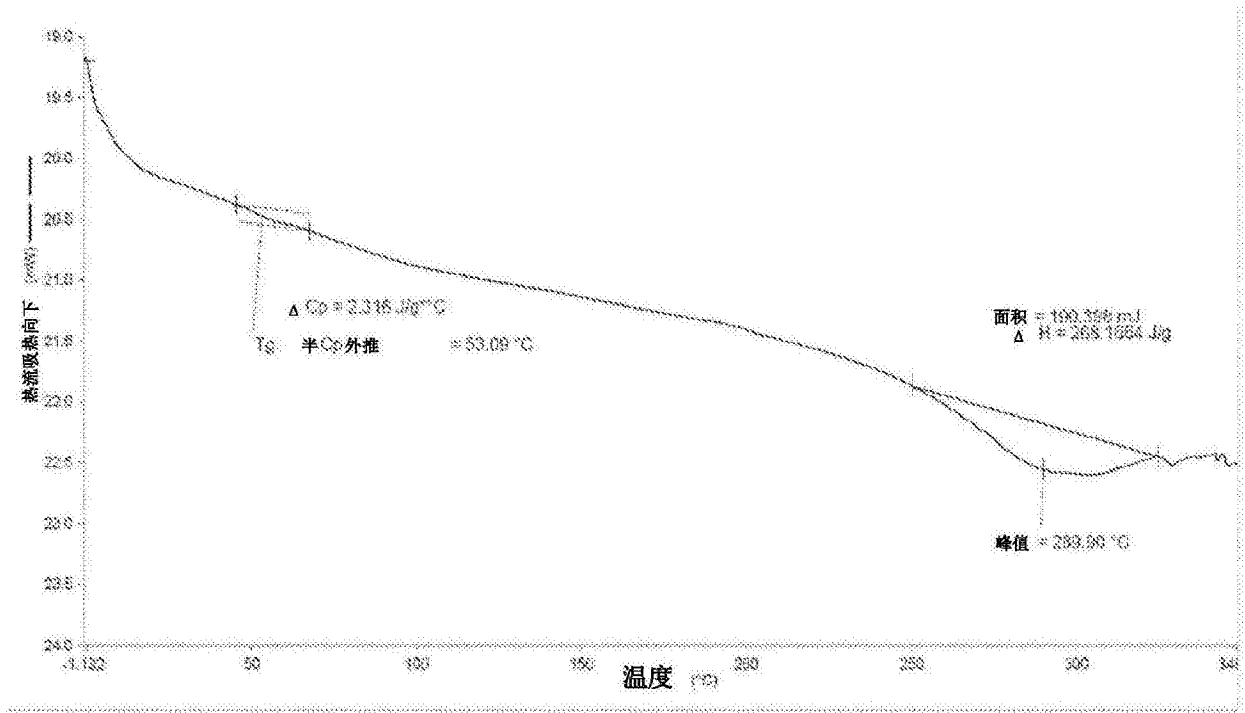


图3B

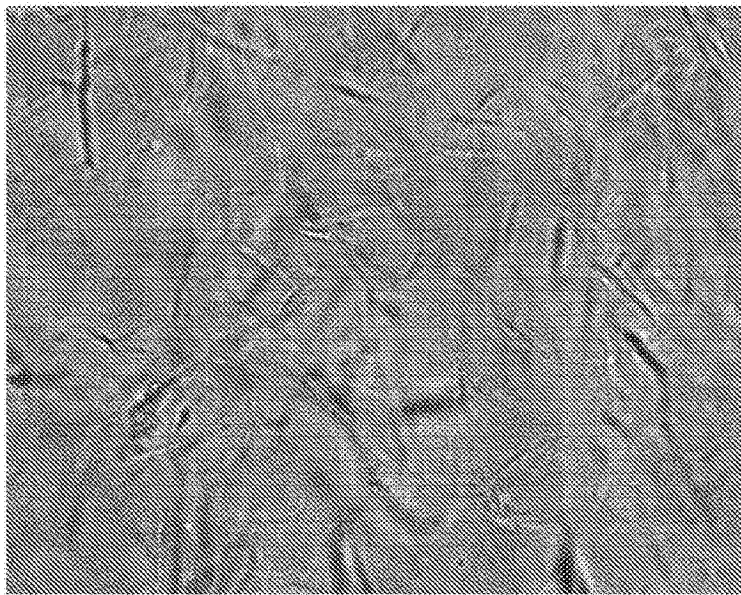


图4A

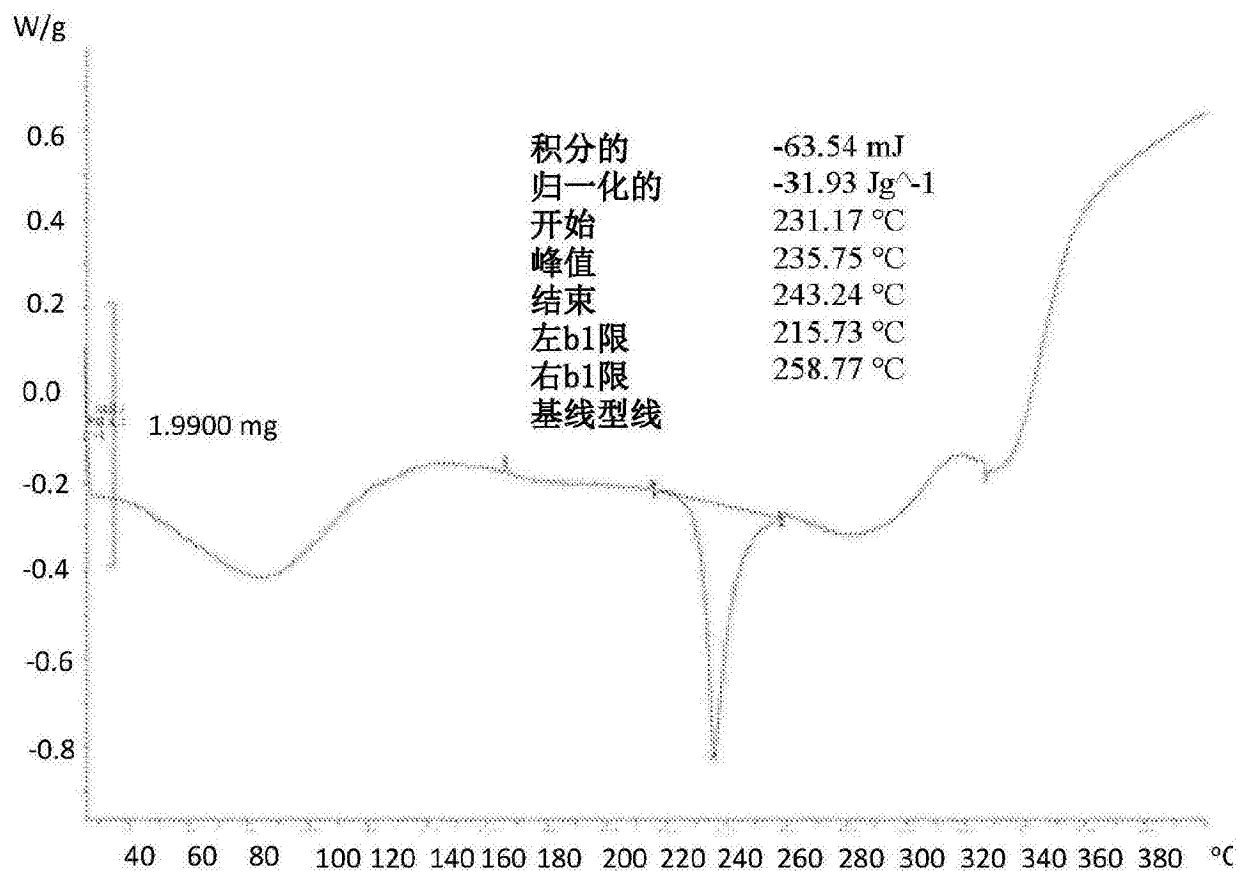


图4B

低聚焦视野

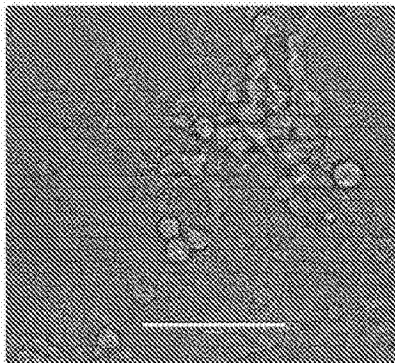


图5A

中聚焦视野

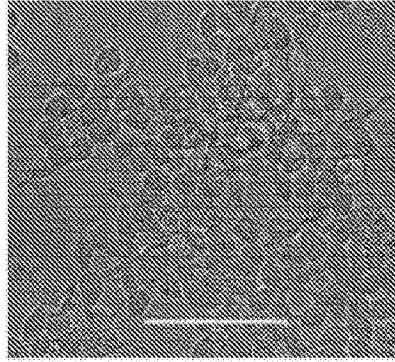


图5B

高聚焦视野

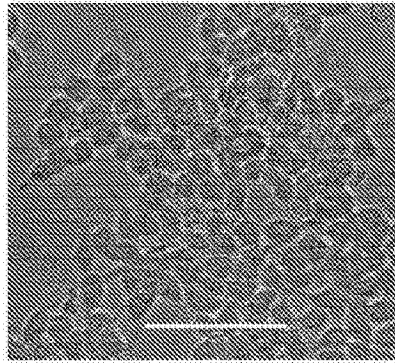


图5C

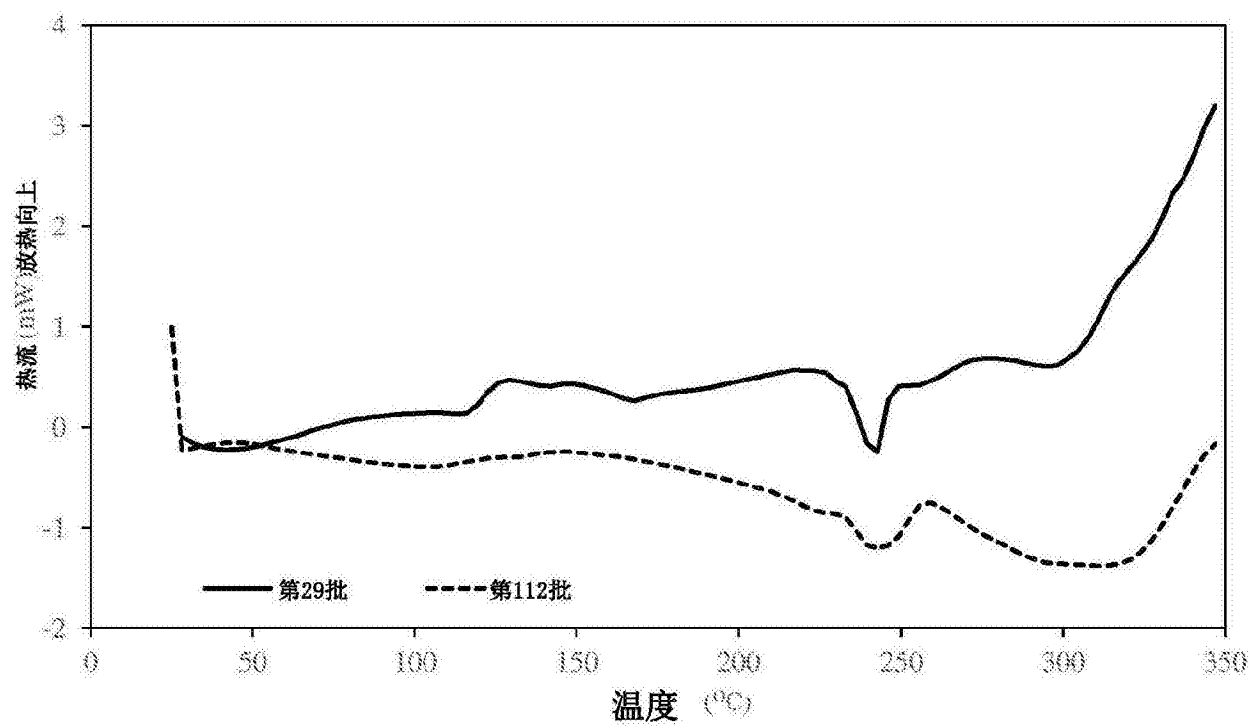


图6