(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116829166 A
(43)申请公布日 2023.09.29

(21)申请号 202180068658.4
(22)申请日 2021.10.29
(30)优先权数据
63/108,163 2020.10.30 US
(85)PCT国际申请进入国家阶段日
2023.04.06
(86)PCT国际申请的申请数据
PCT/US2021/057197 2021.10.29
(87)PCT国际申请的公布数据
WO2022/094177 EN 2022.05.05
(71)申请人 华盛顿大学
地址 美国华盛顿
申请人 加利福尼亚大学董事会
(72)发明人 I·S·普利茨 C·沃夫 J·B·西格尔 C·E·廷伯格 L·斯特沃特 D·贝克尔
(74)专利代理机构 北京市中伦律师事务所 11410
专利代理师 杨黎峰 刘烽
(51)Int.Cl.
A61K 38/00(2006.01)

权利要求书2页 说明书62页
序列表19页

(54)发明名称

用于治疗乳糜泻的组合物和方法

(57)摘要

本公开涉及能够裂解例如麦胶蛋白的麸质蛋白的多肽、编码所述多肽的核酸分子、包含它们的药物组合物以及使用它们治疗乳糜泻和/或非乳糜泻型麸质敏感(NCGS)的方法。

CN 116829166 A

1.一种多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列,其中所述多肽的N末端处的第一氨基酸为Ser(S)。

2.一种多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列,其中在所述多肽的N末端处,所述多肽不包含Met(M)。

3.一种多肽,所述多肽包含与SEQ ID NO:23中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列,其中SEQ ID NO:23中的Xaa不为Met(M)。

4.一种多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列,其中所述多肽的N末端处的第一氨基酸为Ser(S);其中所述多肽包含SEQ ID NO:8中所列的氨基酸序列。

5.如权利要求1-4中任一项所述的多肽,其中所述多肽从N末端至C末端的前两个N末端氨基酸为Ser-Asp(SD)。

6.如权利要求1-5中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少85%序列同一性的氨基酸序列。

7.如权利要求1-6中任一项所述的多肽,所述多肽包含与SEQ IDNO:1中所列的氨基酸序列具有至少90%序列同一性的氨基酸序列。

8.如权利要求1-7中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少95%序列同一性的氨基酸序列。

9.如权利要求1-8中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少99%序列同一性的氨基酸序列。

10.如权利要求1-97中任一项所述的多肽,所述多肽包含SEQ ID NO:1中所列的氨基酸序列。

11.如权利要求1-10中任一项所述的多肽,其中对应于SEQ ID NO:1的氨基酸467的氨基酸残基为Ser。

12.如权利要求1-11中任一项所述的多肽,其中对应于SEQ ID NO:1的氨基酸267的氨基酸残基为Glu。

13.如权利要求1-10中任一项所述的多肽,其中对应于SEQ ID NO:1的氨基酸271的氨基酸残基为Asp。

14.如权利要求1-13中任一项所述的多肽,其中所述多肽能够裂解麦胶蛋白。

15.如权利要求1-14中任一项所述的多肽,所述多肽还包含组氨酸标签,其中所述组氨酸标签在所述多肽的C末端处融合。

16.如权利要求15所述的多肽,其中所述组氨酸标签包含SEQ ID NO:17中所列的氨基酸序列(GSTENLYFQSGALEHHHHHH)。

17.如权利要求15或16所述的多肽,其中所述组氨酸标签包含可裂解组氨酸标签,包括但不限于包含SEQ ID NO:15中所列的氨基酸序列(X_NPQ(L/Q)PX_NHHHHHH)的可裂解组氨酸标

签,其中X_N为介于1-25个之间氨基酸残基的接头。

18.如权利要求17所述的多肽,其中所述可裂解组氨酸标签包含SEQ ID NO:16中所列的氨基酸序列(GSSGSSGSQPQLPYGSSGSSGSHHHHHH)。

19.一种核酸分子,所述核酸分子编码如权利要求1-18中任一项所述的多肽。

20.一种核酸表达载体,所述核酸表达载体包含如权利要求19所述的核酸分子。

21.一种重组宿主细胞,所述重组宿主细胞包含如权利要求19所述的核酸分子或如权利要求20所述的核酸表达载体。

22.如权利要求21所述的重组宿主细胞,其中所述宿主细胞为原核的。

23.如权利要求21所述的重组宿主细胞,其中所述宿主细胞为真核的。

24.一种药物组合物,所述药物组合物包含如权利要求1至18中任一项所述的多肽、如权利要求19所述的核酸分子、如权利要求20所述的核酸表达载体、如权利要求21-23中任一项所述的重组宿主细胞,或其任何组合以及药学上可接受的载体。

25.一种用于治疗受试者的乳糜泻或非乳糜泻型麸质敏感(NCGS)的方法,所述方法包括向患有乳糜泻或NCGS的所述受试者施用有效治疗所述乳糜泻或NCGS的量的如权利要求1至18中任一项所述的多肽、如权利要求19所述的核酸分子、如权利要求20所述的核酸表达载体、如权利要求21-23中任一项所述的重组宿主细胞,或如权利要求24所述的药物组合物,从而治疗所述乳糜泻或NCGS。

26.一种用于减轻受试者的乳糜泻或非乳糜泻型麸质敏感(NCGS)相关炎症的方法,所述方法包括向患有乳糜泻或NCGS的所述受试者施用有效减轻所述乳糜泻或NCGS相关炎症的量的如权利要求1至18中任一项所述的多肽、如权利要求19所述的核酸分子、如权利要求20所述的核酸表达载体、如权利要求21-23中任一项所述的重组宿主细胞或如权利要求24所述的药物组合物,从而减轻所述炎症。

27.如权利要求26所述的方法,其中所述多肽、所述核酸分子、所述核酸表达载体、所述重组宿主细胞或所述药物组合物经口施用。

28.一种用于降解食品中的麸质的方法,所述方法包括使所述食品与有效降解所述麸质的量的如权利要求1至18中任一项所述的多肽或如权利要求24所述的药物组合物接触,从而降解所述食品中的所述麸质。

29.一种用于降解食品中的麦胶蛋白的方法,所述方法包括使所述食品与有效降解所述麦胶蛋白的量的如权利要求1至18中任一项所述的多肽或如权利要求24所述的药物组合物接触,从而降解所述食品中的所述麦胶蛋白。

30.如权利要求28或29所述的方法,其中所述方法降解所述食品中的至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约98%、至少约99%或约100%的所述麸质或麦胶蛋白。

用于治疗乳糜泻的组合物和方法

[0001] 相关申请

[0002] 本申请要求2020年10月30日提交的美国临时申请第63/108,163号的优先权权益，其以引用的方式整体并入本文。

[0003] 对序列表的引用

[0004] 本申请含有已以ASCII格式以电子方式提交且在此以引用的方式整体并入的序列表。所述ASCII拷贝创建于2021年10月15日，名为7281_50WO2_Seqlisting_ST25.txt并且大小为31,599字节。

技术领域

[0005] 本公开涉及能够裂解例如麦胶蛋白的麸质肽的组合物，以及其用于治疗包括乳糜泻在内的麸质敏感的用途。

背景技术

[0006] 乳糜泻是一种非常流行的疾病，其中在小麦、大麦和黑麦产品中发现的被称为“麸质”的膳食蛋白质在遗传易感个体的小肠中引起免疫反应。由此产生的炎症可导致小肠绒毛退化，阻碍营养物质的吸收。症状可能出现在早期儿童时代或晚年，并且其严重程度范围很广，从腹泻、疲劳和体重减轻到腹胀、贫血和神经系统症状。除了从饮食中完全消除麸质外，目前没有针对这种终生疾病的有效疗法。尽管乳糜泻在很大程度上无法加以全面诊断，但其在美国和欧洲的患病率估计为人口的0.5%-1.0%。除了乳糜泻外，很大一部分人群被认为患有非乳糜泻型麸质敏感(NCGS)疾患，其由摄入麸质引起，尽管症状通常与乳糜泻的症状无法区分，但在机制上与乳糜泻不同。由于在人类消化道的严酷和高酸性环境中特异性地且有效地降解麸质衍生肽的物理和化学要求严格，因此难以鉴定出合适的天然存在的酶作为乳糜泻和NCGS的口服治疗剂。由于麸质肽在进入肠道后立即起始免疫反应，因此任何治疗乳糜泻的口服酶治疗剂均必须使这些免疫原性麸质区域在胃隔室中分解，从而防止这些麸质肽引起因炎症所致的肠道损伤。

发明内容

[0007] 本公开的某些方面涉及一种多肽，所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少85%序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少90%序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少95%序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少99%序列同一性的氨基酸序列。在一些方面中，多肽包含SEQ ID NO:1中所列的氨基酸序列。

[0008] 在一些方面中，对应于SEQ ID NO:6的氨基酸467的氨基酸残基为Ser。在一些方面中，对应于SEQ ID NO:6的氨基酸267的氨基酸残基为Glu。在一些方面中，对应于SEQ ID NO:6的氨基酸271的氨基酸残基为Asp。

[0009] 在一些方面中，多肽能够裂解麦胶蛋白。

[0010] 本公开的某些方面涉及一种多肽，所述多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少85％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少90％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少95％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少99％序列同一性的氨基酸序列。在一些方面中，多肽包含SEQ ID NO:8中所列的氨基酸序列。

[0011] 在一些方面中，对应于SEQ ID NO:3的氨基酸278的氨基酸残基为Ser。在一些方面中，对应于SEQ ID NO:3的氨基酸78的氨基酸残基为Glu。在一些方面中，对应于SEQ ID NO:3的氨基酸82的氨基酸残基为Asp。

[0012] 在一些方面中，多肽能够裂解麦胶蛋白。

[0013] 本公开的某些方面涉及一种多肽，所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列；其中多肽包含SEQ ID NO:8中所列的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少85％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少90％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少95％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少99％序列同一性的氨基酸序列。在一些方面中，多肽包含SEQ ID NO:1中所列的氨基酸序列。

[0014] 在一些方面中，对应于SEQ ID NO:6的氨基酸467的氨基酸残基为Ser。在一些方面中，对应于SEQ ID NO:6的氨基酸267的氨基酸残基为Glu。在一些方面中，对应于SEQ ID NO:6的氨基酸271的氨基酸残基为Asp。

[0015] 在一些方面中，多肽能够裂解麦胶蛋白。

[0016] 在一些方面中，多肽包含组氨酸标签，其中组氨酸标签在多肽的C末端融合。在一些方面中，组氨酸标签包含SEQ ID NO:17中所列的氨基酸序列(GSTENLYFQSGALEHHHHHH)。在一些方面中，组氨酸标签包含可裂解组氨酸标签，包括但不限于包含SEQ ID NO:15中所列的氨基酸序列的可裂解组氨酸标签(X_NPQ(L/Q)PX_NHHHHHH)，其中X_N为介于1-25个之间氨基酸残基的接头。在一些方面中，可裂解组氨酸标签包含SEQ ID NO:16中所列的氨基酸序列(GSSGSSGSQPQLPYGSSGSSGSHHHHHH)。

[0017] 本公开的某些方面涉及一种编码本文公开的多肽的核酸分子。

[0018] 本公开的某些方面涉及一种包含本文公开的核酸分子的核酸表达载体。

[0019] 本公开的某些方面涉及一种包含本文公开的核酸分子或核酸表达载体的重组宿

主细胞。

[0020] 本公开的某些方面涉及一种药物组合物，所述药物组合物包含本文公开的多肽、本文公开的核酸分子、本文公开的核酸表达载体、本文公开的重组宿主细胞或其任何组合以及药学上可接受的载体。

[0021] 本公开的某些方面涉及一种用于治疗乳糜泻或非乳糜泻型麸质敏感(NCGS)的方法，所述方法包括向患有乳糜泻或NCGS的个体施用有效治疗乳糜泻或NCGS的量的本文公开的多肽、本文公开的核酸分子、本文公开的核酸表达载体、本文公开的重组宿主细胞或本文公开的药物组合物。在一些方面中，多肽、核酸分子、核酸表达载体、重组宿主细胞或药物组合物经口施用。

[0022] 在一些方面中，本公开涉及一种多肽，所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列，其中多肽的N末端处的第一氨基酸为Ser(S)。在一些方面中，多肽具有麦胶蛋白酶活性。

[0023] 在一些方面中，本公开涉及一种多肽，所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列，其中多肽在多肽的N末端处不包含Met(M)。

[0024] 在一些方面中，本公开涉及一种多肽，所述多肽包含与SEQ ID NO:23中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列，其中SEQ ID NO:23中的Xaa不为Met(M)。

[0025] 在一些方面中，本公开涉及一种多肽，所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列，其中多肽的N末端处的第一氨基酸为Ser(S)；其中多肽包含SEQ ID NO:8中所列的氨基酸序列。

[0026] 在一些方面中，多肽从N末端至C末端的前两个N末端氨基酸为Ser-Asp(SD)。在一些方面中，多肽从N末端至C末端的前三个N末端氨基酸为Ser-Asp-Met(SDM)。在一些方面中，多肽从N末端至C末端的前四个N末端氨基酸为Ser-Asp-Met-Glu(SDME)。

[0027] 在一些方面中，本文公开的多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少85％序列同一性的氨基酸序列。在一些方面中，本文公开的多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少90％序列同一性的氨基酸序列。在一些方面中，本文公开的多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少95％序列同一性的氨基酸序列。在一些方面中，本文公开的多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少99％序列同一性的氨基酸序列。在一些方面中，本文公开的多肽包含SEQ ID NO:1中所列的氨基酸序列。

[0028] 在本文公开的多肽的一些方面中，对应于SEQ ID NO:1的氨基酸467的氨基酸残基为Ser。在本文公开的多肽的一些方面中，对应于SEQ ID NO:1的氨基酸267的氨基酸残基为Glu。在本文公开的多肽的一些方面中，对应于SEQ ID NO:1的氨基酸271的氨基酸残基为Asp。

[0029] 在本公开的一些方面中，多肽能够裂解麦胶蛋白。在一些方面中，与Kuma011相比，

多肽具有改进的酶活性。

[0030] 在一些方面中，本文公开的多肽还包含组氨酸标签，其中组氨酸标签在多肽的C末端处融合。在一些方面中，组氨酸标签包含SEQ ID NO:17中所列的氨基酸序列(GSTENLYFQSGALEHHHHHH)。在一些方面中，组氨酸标签包含可裂解组氨酸标签，包括但不限于包含SEQ ID NO:15中所列的氨基酸序列的可裂解组氨酸标签(XNPQ(L/Q)PXNHHHHHH)，其中XN为介于1-25个之间氨基酸残基的接头。在一些方面中，可裂解组氨酸标签包含SEQ ID NO:16中所列的氨基酸序列(GSSGSSGSQPQLPYGSSGSSGSHHHHHH)。

[0031] 在一些方面中，本公开涉及一种编码本文所述的多肽的核酸分子。在一些方面中，本公开涉及一种包含本文所述的核酸分子的核酸表达载体。

[0032] 在一些方面中，本公开涉及一种包含本文所述的核酸分子或核酸表达载体的重组宿主细胞。在一些方面中，宿主细胞为原核的。在一些方面中，宿主细胞为真核的。

[0033] 在一些方面中，本公开涉及一种药物组合物，所述药物组合物包含本文所述的多肽、核酸分子、核酸表达载体或重组宿主细胞或其任何组合以及药学上可接受的载体。

[0034] 在一些方面中，本公开涉及一种用于治疗受试者的乳糜泻或非乳糜泻型麸质敏感(NCGS)的方法，所述方法包括向患有乳糜泻或NCGS的受试者施用有效治疗乳糜泻或NCGS的量的本文所述的多肽、核酸分子、核酸表达载体、重组宿主细胞或药物组合物，从而治疗乳糜泻或NCGS。

[0035] 在一些方面中，本公开涉及一种用于减轻受试者的乳糜泻或非乳糜泻型麸质敏感(NCGS)相关炎症的方法，所述方法包括向患有乳糜泻或NCGS的受试者施用有效减轻乳糜泻或NCGS相关炎症的本文所述的多肽、核酸分子、核酸表达载体、重组宿主细胞或药物组合物，从而减轻炎症。在一些方面中，多肽、核酸分子、核酸表达载体、重组宿主细胞或药物组合物经口施用。

[0036] 在一些方面中，本公开涉及一种用于降解食品中的麸质的方法，所述方法包括使食品与有效降解麸质的量的本文所述的多肽或药物组合物接触，从而降解食品中的麸质。

[0037] 在一些方面中，本公开涉及一种用于降解食品中的麦胶蛋白的方法，所述方法包括使食品与有效降解麦胶蛋白的量的本文所述的多肽或药物组合物接触，从而降解食品中的麦胶蛋白。

[0038] 在一些方面中，所述方法降解食品中的至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约98％、至少约99％或约100％的麸质或麦胶蛋白。在一些方面中，所述方法在不到约1.5小时、不到约1小时、不到约45分钟、不到约40分钟、不到约30分钟、不到约25分钟、不到约20分钟、不到约15分钟、不到约10分钟或不到约5分钟降解食品中的麸质或麦胶蛋白。在一些方面中，所述方法在小于约6.5、小于约6.0、小于约5.5、小于约5.0、小于约4.5、小于约4.0、小于约3.5、小于约3.0、小于约2.5、小于约2.0或小于约1.5的pH值下降解食品中的麸质或麦胶蛋白。

具体实施方式

[0039] 本公开提供能够降解麦胶蛋白肽的麦胶蛋白酶。本公开的一些方面涉及包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨

基酸序列的多肽，其中多肽的N末端处的第一氨基酸为Ser(S)。在一些方面中，多肽在多肽的N末端处不包含Met(M)。在一些方面中，多肽包含SEQ ID NO:8中所列的氨基酸序列。

[0040] 1.定义

[0041] 为了可以更容易地理解本公开，首先定义某些术语。除非本文另有定义，否则与本公开相关的科学和技术术语应具有本领域普通技术人员通常理解的含义。术语的含义和范围应该是明确的，然而，在任何潜在歧义的情况下，本文提供的定义优先于任何字典或外部定义。

[0042] 此外，应当注意，每当引用参数的值或值的范围时，旨在引用值的中间值和范围也为本公开的一部分。

[0043] 如本文所用，单数形式“一个/种(a/an)”和“the(所述)”包括复数指示物，除非上下文另有明确规定。除非另有明确规定，否则本文所用的“和”与“或”可互换使用。除非另有说明，否则术语“包含”、“具有”、“包括”和“含有”应被解释为开放式术语(即，意为“包括但不限于”)。除非在本文中另有说明，否则本文对数值范围的引用仅旨在作为引用的或落入所述范围内的各单独值的简写方法，并且每个单独的值均并入说明书中，如同其被单独引用一样。

[0044] 术语“约”或“大约”通常意指在给定值或范围的10％以内、5％以内或更优选1％以内。

[0045] 术语“氨基酸”是指二十种常见的天然存在的氨基酸。天然存在的氨基酸包括：丙氨酸(Ala;A)、天冬酰胺(Asn;N)、天冬氨酸(Asp;D)、精氨酸(Arg;R)、半胱氨酸(Cys;C)、谷氨酸(Glu;E)、谷氨酰胺(Gln;Q)、甘氨酸(Gly;G)、组氨酸(His;H)、异亮氨酸(Ile;I)、亮氨酸(Leu;L)、赖氨酸(Lys;K)、甲硫氨酸(Met;M)、苯丙氨酸(Phe;F)、脯氨酸(Pro;P)、丝氨酸(Ser;S)、苏氨酸(Thr;T)、色氨酸(Trp;W)、酪氨酸(Tyr;Y)以及缬氨酸(Val;V)。

[0046] 术语“乳糜泻(Celiac disease和celiac sprue disease)”可互换使用，并且是指以对麸质(小麦粉中的主要蛋白质)中的免疫原性肽和相关蛋白质发生炎症反应为特征的疾患。摄入后，α-麦胶蛋白被胃和肠蛋白酶部分降解为寡肽，本文中称为“麦胶蛋白”。由于麦胶蛋白异常高的脯氨酸和谷氨酰胺含量，其对在胃条件下进一步的蛋白水解具有抗性。

[0047] 如本文所用，“保守氨基酸取代”是其中一个氨基酸残基被另一个具有相似化学特性(例如电荷或疏水性)的侧链(R基团)的氨基酸残基取代的氨基酸残基。一般而言，保守氨基酸取代基本上不会改变蛋白质的功能特性。在两个或更多个氨基酸序列因保守取代而彼此不同的情况下，可以上调序列同一性百分比或相似度以校正取代的保守性质。进行这种调整的手段是本领域技术人员熟知的。参见例如，Pearson(1994)Methods Mol.Biol.24:307-331，其以引用的方式并入本文中。具有类似化学特性的侧链的氨基酸组的实例包括(1)脂肪族侧链：甘氨酸、丙氨酸、缬氨酸、亮氨酸和异亮氨酸；(2)脂肪族羟基侧链：丝氨酸和苏氨酸；(3)含酰胺的侧链：天冬酰胺和谷氨酰胺；(4)芳香族侧链：苯丙氨酸、酪氨酸和色氨酸；(5)碱性侧链：赖氨酸、精氨酸和组氨酸；(6)酸性侧链：天冬氨酸和谷氨酸，和(7)含硫侧链为半胱氨酸和甲硫氨酸。优选的保守氨基酸取代基为：缬氨酸-亮氨酸-异亮氨酸、苯丙氨酸-酪氨酸、赖氨酸-精氨酸、丙氨酸-缬氨酸、谷氨酸-天冬氨酸和天冬酰胺-谷氨酰胺。或者，保守置换是在以引用的方式并入本文的Gonnet等人.(1992)Science 256:1443-1445中公开的PAM250对数似然矩阵中具有正值的任何变化。“中等保守”置换是在PAM250对数似然

矩阵中具有非负值的任何变化。

[0048] 如本文所用,术语“降解(degrade和degradation)”是指将靶标,例如多肽,例如麸质、麦胶蛋白和相关蛋白质破坏或分解成较小的寡肽。在某些实施方案中,麦胶蛋白的降解导致与乳糜泻相关的免疫原性肽的减少和/或移除。

[0049] 如本文所用,术语“麦胶蛋白酶”是指可有效降解一种或多种麦胶蛋白的多肽(酶)。如本文所用,术语“麦胶蛋白”是指麸质的富含脯氨酸(P)和谷氨酰胺(Q)的肽组分。示例性麦胶蛋白包含PQLP(SEQ ID NO:9)或PQQP(SEQ ID NO:10)基序(诸如PFPQPQLPY(SEQ ID NO:11)和/或PFPQPQQPF(SEQ ID NO:12))。在某些方面中,麦胶蛋白酶在酸性条件下,例如在pH 4或更低下降解一种或多种麦胶蛋白。

[0050] 如本文所用,术语“突变”是指多肽中一个或多个氨基酸或多核苷酸中一个或多个核苷酸的插入、缺失或取代。

[0051] 如本文所用,术语“变体”是指相对于参考多肽或多核苷酸包含一个或多个氨基酸或核苷酸插入、取代或缺失的多肽或多核苷酸。在某些方面中,变体多肽或多核苷酸与参考多肽或多核苷酸序列具有至少约75%氨基酸或核苷酸序列同一性,例如至少约80%、至少约85%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性。在一些方面中,参考多肽或多核苷酸的变体保持参考多肽或多核苷酸的一种或多种功能、活性和/或结构。例如,本文公开的麦胶蛋白酶的变体保持有效降解麸质和/或麦胶蛋白的功能。在另一个实例中,编码麦胶蛋白酶的多核苷酸的变体编码功能性麦胶蛋白酶。

[0052] 通常使用序列分析软件测量序列同一性。蛋白质分析软件使用分配给各种取代、缺失和其它修饰(包括保守氨基酸取代)的相似性度量来匹配相似序列。例如,GCG软件含有诸如Gap和Bestfit的程序,其可使用默认参数来确定密切相关多肽,诸如来自不同生物体物种的同源多肽之间或野生型蛋白质与其突变蛋白之间的序列同源性或序列同一性。参见例如GCG版本6.1。还可使用FASTA使用默认或推荐参数(GCG版本6.1中的程序)比较多肽序列。FASTA(例如FASTA2和FASTA3)提供查询序列与搜索序列之间的最佳重叠区域的比对和序列同一性百分比(Pearson(2000)同上)。可用于将本公开的序列与含有来自不同生物体的大量序列的数据库进行比较的算法的另一个非限制性实例系使用默认参数的计算机程序BLAST,例如BLASTP或TBLASTN。参见例如Altschul等人.(1990)J.Mol.Biol.215:403-410和Altschul等人.(1997)Nucleic Acids Res.25:3389-402,所述文献各自以引用的方式整体并入本文。

[0053] 如本文所用,“治疗(treatment或treating)”是指产生有益效果的作用,例如改善疾病或病症的至少一种症状。有益效果可采取与基线相比改进的形式,即与根据所述方法在开始治疗之前进行的测量或观察相比的改进。有益效果还可采取阻止、减缓、推迟或稳定损伤(例如炎症)的形式,其可导致小肠绒毛降解(包括增生和绒毛萎缩),其特征为乳糜泻或非乳糜泻型麸质敏感(NCGS)。有效治疗可指减轻或预防乳糜泻或NCGS的至少一种症状。这种有效治疗可减轻乳糜泻或NCGS的肠内和/或肠外临床表现,诸如腹泻、腹痛、营养不良、贫血、骨质疏松症或任何已知症状,抑制症状恶化;限制或预防先前患有所述病症的患者的乳糜泻复发;限制或预防先前有乳糜泻或NCGS症状的患者的症状复发;和/或限制有发展乳糜泻或NCGS风险或尚未显示出乳糜泻或NCGS的临床作用的受试者的乳糜泻或NCGS的发展。

[0054] 在一些方面中,治疗减轻小肠中的炎症。与治疗之前的炎症相比,炎症的有效减轻可包括炎症减轻至少约1%、至少约5%、至少约10%、至少约20%、至少约30%、至少约40%、至少约50%、至少约60%、至少约70%、至少约80%、至少约90%、至少约95%、至少约99%或约100%。可通过任何方式测量炎症的减轻。

[0055] 任何对麸质敏感的受试者均可根据本公开的方法进行治疗。在某些方面中,受试者正经历乳糜泻。在某些方面中,受试者患有NCGS。在某些方面中,受试者为人类受试者。在某些方面中,受试者正经历一种或多种与麸质敏感相关的症状。在某些方面中,受试者是无症状的。

[0056] 如本文所用,"有效量"是指足以引起一种或多种麸质敏感症状(例如乳糜泻或NCGS)的严重程度或频率降低的多肽的量。

[0057] 本文公开的多肽可配制成药物组合物,诸如上文公开的那些,并且可经由任何合适的途径施用,包括经口、肠胃外、吸入喷雾或以含有常规药学上可接受的载体、佐剂和媒介物的单位剂量制剂局部施用。

[0058] 本公开的所有方面均可组合使用,除非上下文另有明确规定。引用的所有参考文献以引用的方式整体并入本文。在本申请中,除非另有说明,否则所使用的技术可在几个熟知的参考文献中找到,诸如:Molecular Cloning:A Laboratory Manual(Sambrook等人,1989,Cold Spring Harbor Laboratory Press),Gene Expression Technology(Methods in Enzymology,第185卷,D.Goeddel编,1991.Academic Press,San Diego,CA),"Guide to Protein Purification"in Methods in Enzymology(M.P.Deutshcer编,(1990)Academic Press,Inc.);PCR Protocols:AGuide to Methods and Applications(Innis等人,1990.Academic Press,San Diego,CA),Culture of Animal Cells:A Manual of Basic Technique,第2版.(R.I.Freshney.1987.Liss,Inc.New York,NY),Gene Transfer and Expression Protocols,第109-128页,E.J.Murray编,The Humana Press Inc.,Clifton,N.J.),and the Ambion 1998Catalog(Ambion,Austin,TX)。

[0059] 2.本公开的组合物

[0060] 本公开提供有效降解麦胶蛋白的麦胶蛋白酶。本公开至少部分基于以下发现:如本文所述,含有一个或多个相对于Kuma011的突变的各种多肽具有相对于Kuma011和其它已知的麦胶蛋白酶,诸如SC-PEP(鞘氨醇单胞菌荚膜肽酶)和内切蛋白酶EPB2改进的活性,包括增加的麦胶蛋白降解活性。在某些实施方案中,本文描述的各种多肽在酸性条件下具有优于Kuma011和其它已知麦胶蛋白酶的改进的麦胶蛋白酶活性。

[0061] 在一些方面中,本公开提供包含与SEQ ID NO:6中所列的氨基酸序列至少75%同一的氨基酸序列的多肽,其中(a)残基467为Ser,残基267为Glu,并且残基271为Asp;并且(b)多肽在一个或多个选自由以下组成的组的残基处包含相对于SEQ ID NO:6的氨基酸取代:221、262E、268、269、270、319A、320、354E/Q/R/Y、358S/Q/T、368F/Q、399、402、406、424、449、461、463、105、171、172、173、174和456。在一些方面中,多肽在一个或多个选自由以下组成的组的残基处包含相对于SEQ ID NO:6的氨基酸取代:221、262E、268、269、270、319A、320、354E/Q/R/Y、358S/Q/T、368F/Q、399、402、406、424、449、461和463。

[0062] 表1:Kuma序列

[0063]

Kuma011 (全长) SEQ ID NO: 6 (粗体 = 前蛋白质结构域)	**MSDMEKPWKEGEEARAVLQGHARAQAPQAVDK GPVAGDERMAVTVVLRRQRAGELAAHVERQAAI APHAREHLKREAFAASHGASLDDFAELRRFADA HGLALDRANVAAGTAVLSGPDDAINRAFGVELR HFDHPDGSYRSYLGEVTVPASIAPMIEAVLGLDT RPVARPHFRMQRRAEGGFEARSQA**AAPTAYTPLD VAQAYQFPEGLDGQGQCIAIIELGGGYDEASLAQYF ASLGVPAPQVVSVSVDGASNQPTGDPKGPDGEVEL DIEVAGALAPGAKFAVYFAPDTTAGFLDAITTAIHD PTLKPSVVSISWSGPEDSWTSAAIAAMNRAFLDAAA LGVTVLAAAGDSGSTGGEQDGLYHVHFPAASPYVL ACGGTRLVASGGRIAQETVWNDGPDGGATGGGVS RIFPLPAWQEHANVPPSANPGASSGRGVPDLAGNA DPATGYEVVIDGEATVIGGTSAVAPLFAALVARINQ KLGKAVGYLNPTLYQLPADVFHDITEGNNDIANRA QIYQAGPGWDPCTGLGSPIGVRLLQALLPSASQPQP
Kuma011 前蛋白质结构域 SEQ ID NO: 2	MSDMEKPWKEGEEARAVLQGHARAQAPQAVDKG PVAGDERMAVTVVLRRQRAGELAAHVERQAAIAP HAREHLKREAFAASHGASLDDFAELRRFADAHGLA LDRANVAAGTAVLSGPDDAINRAFGVELRHFDHPD GSYRSYLGEVTVPASIAPMIEAVLGLDTRPVARPHF RMQRRAEGGFEARSQA
Kuma011 成	AAPTAYTPLDVAQAYQFPEGLDGQGQCIAIIELGGG YDEASLAQYFASLGVPAPQVVSVSVDGASNQPTGD

[0064]

熟肽 SEQ ID NO: 3	PKGPDGEVELDIEVAGALAPGAKFAVYFAPDTTAGF LDAITTAIHDPTLKPSVVSISWSGPEDSWTSAAIAAM NRAFLDAAALGVTVLAAAGDSGSTGGEQDGLYHV HFPAASPYVLACGGTRLVASGGRIAQETVWNDGPD GGATGGGVSRIFPLPAWQEHANVPPSANPGASSGR GVPDLAGNADPATGYEVVIDGEATVIGGTSAVAPLF AALVARINQKLGKAVGYLNPTLYQLPADVFHDITE GNNDIANRAQIYQAGPGWDPCTGLGSPIGVRLLQAL LPSASQPQP
Kuma010 (全长) SEQ ID NO: 4 (粗体 = 前蛋白质结构域)	**MSDMEKPWKEGEEARAVLQGHARAQAPQAVDK GPVAGDERMAVTVVLRRQRAGELAAHVERQAAI APHAREHLKREAFAASHGASLDDFAELRRFADA HGLALDRANVAAGTAVLSGPDDAINRAFGVELR HFDHPDGSYRSYLGEVTVPASIAPMIEAVLGLDT RPVARPHFRMQRRAEGGFEARSQA**AAPTAYTPLD VAQAYQFPEGLDGQGQCIAIIELGGGYDEASLAQYF ASLGVPAPQVVSVSVDGASNQPTGDPKGPDGEVEL DIEVAGALAPGAKFAVYFAPDTTAGFLDAITTAIHD PTLKPSVVSISWSGPEDSWTSAAIAAMNRAFLDAAA LGVTVLAAAGDSGSTGGEQDGLYHVHFPAASPYVL ACGGTRLVASGGRIAQETVWNDGPDGGATGGGVS RIFPLPAWQEHANVPPSANPGASSGRGVPDLAGNA DPATGYEVVIDGEATVIGGTSAVAPLFAALVARINQ KLGKAVGYLNPTLYQLPADVFHDITEGNNDIANRA QIYQAGPGWDPCTGLGSPIGVRLLQALLPSASQPQP GSTENLYFQSGALEHHHHHH
Kuma010 成熟肽 SEQ ID NO: 5	AAPTAYTPLDVAQAYQFPEGLDGQGQCIAIIELGGG YDEASLAQYFASLGVPAPQVVSVSVDGASNQPTGD PKGPDGEVELDIEVAGALAPGAKFAVYFAPDTTAGF LDAITTAIHDPTLKPSVVSISWSGPEDSWTSAAIAAM NRAFLDAAALGVTVLAAAGDSGSTGGEQDGLYHV HFPAASPYVLACGGTRLVASGGRIAQETVWNDGPD

[0065]

	GGATGGGVSRIFPLPAWQEHANVPPSANPGASSGR GVPDLAGNADPATGYEVVIDGEATVIGGTSAVAPLF AALVARINQKLGKAVGYLNPTLYQLPADVFHDITE GNNDIANRAQIYQAGPGWDPCTGLGSPIGVRLLQAL LPSASQPQPGSTENLYFQSGALEHHHHHH
Kuma062-M (全长) SEQ ID NO: 1 (粗体 = 前蛋白质结构域)	**SDMEKPWKEGEEARAVLQGHARAQAPQAVDKG PVAGDERMAVTVVLRRQRAGELAAHVERQAAIA PHAREHLKREAFAASHGASLDDFAELRRFADAH GLALDRANVAAGTAVLSGPDDAINRAFGVELRHF DHPDGSYRSYLGEVTVPASIAPMIEAVLGLDTRP VARRRFRMQRRAEGGFEARSQA**AAPTAYTPLDV AQAYQFPEGLDGQGQCIAIIELGGGYDEASLAQYFA SLGVPAPQVVSVSVDGASNQPTGDPEGPDGEVTLDI EVAGALAPGAKFAVYFAPDTTAGFLDAITTAIHDPT LKPSVVSISWSMPEDSWTSAAIAAMNRAFLDAAAL GVTVLAAAGDQGSTSGEQDGLYHVHFPAASPYVLA CGGTRLVASGGRIAQETVWNQGPDGGATGGGVSRI FPLPAWQEHANVPPSANPGASSGRGVPDLAGNADP QTGYEVVIDGEATVTGGTSAVAPLFAALVARINQKL GKAVGYLNPTLYQLPADVFHDITEGNNDIANRAQIY QAGPGWDPCTGLGSPIGVRLLQALLPSASQPQP
Kuma062-M 前蛋白质结构域 SEQ ID NO: 7	SDMEKPWKEGEEARAVLQGHARAQAPQAVDKGPV AGDERMAVTVVLRRQRAGELAAHVERQAAIAPHA REHLKREAFAASHGASLDDFAELRRFADAHGLALD RANVAAGTAVLSGPDDAINRAFGVELRHFDHPDGS YRSYLGEVTVPASIAPMIEAVLGLDTRPVARRRFRM QRRAEGGFEARSQA
Kuma062-M 成熟肽 SEQ ID NO: 8	AAPTAYTPLDVAQAYQFPEGLDGQGQCIAIIELGGG YDEASLAQYFASLGVPAPQVVSVSVDGASNQPTGD PEGPDGEVTLDIEVAGALAPGAKFAVYFAPDTTAGF LDAITTAIHDPTLKPSVVSISWSMPEDSWTSAAIAAM NRAFLDAAALGVTVLAAAGDQGSTSGEQDGLYHV

[0066]

	HFPAASPYVLACGGTRLVASGGRIAQETVWNQGPD GGATGGGVSRIFPLPAWQEHANVPPSANPGASSGR GVPDLAGNADPQTGYEVVIDGEATVTGGTSAVAPL FAALVARINQKLGKAVGYLNPTLYQLPADVFHDITE GNNDIANRAQIYQAGPGWDPCTGLGSPIGVRLLQAL LPSASQPQ

[0067] 如本文所述,Kuma010包含通过氨基键连接至Kuma010序列的C末端处的组氨酸标签序列GSTENLYFQSGALEHHHHHH(SEQ ID NO:17)的Kuma011。

[0068] 表1中提供的序列中的粗体残基代表未加工多肽中存在的N末端部分(即在加工期间被裂解除去);并且非粗体字体代表多肽的经加工型式(即成熟肽序列)中存在的残基。括号内的数字表示残基数;并且其中有两个以“/”隔开的数字,左边的数字为未加工型式的残基数,右边的数字为加工型式的残基数。SEQ ID NO:6为Kuma011的未加工型式;SEQ ID NO:3为Kuma011的加工型式。因此,包含SEQ ID NO:6中所列的氨基酸序列的多肽(全长Kuma011多肽)也必然包含SEQ ID NO:3中所列的氨基酸序列(成熟Kuma011多肽)。SEQ ID NO:1为Kuma062-M的未加工型式;SEQ ID NO:8为Kuma062-M的加工型式。因此,包含SEQ ID NO:1中所列的氨基酸序列的多肽(全长Kuma062-M多肽)也必然包含SEQ ID NO:8中所列的氨基酸序列(成熟Kuma062-M多肽)。

[0069] 在一些方面中,本公开的麦胶蛋白酶在其N末端处具有丝氨酸(Ser或S)。在一些方面中,本公开的麦胶蛋白酶在其N末端处具有SD基序。在一些方面中,本公开的麦胶蛋白酶在其N末端处具有SDM基序。在一些方面中,本公开的麦胶蛋白酶在其N末端处具有SDME(SEQ ID NO:21)。在此类方面中,第一氨基酸(多肽从其N末端开始的位置1为S;第二氨基酸(多肽从其N末端开始的位置2为D;第三氨基酸(多肽从其N末端开始的位置3为M);并且第四氨基酸(多肽从其N末端开始的位置4为E。在一些方面中,寡肽在其N末端连接至N末端S,其中在其N末端与S相邻的氨基酸不为甲硫氨酸(M)。

[0070] 在一些方面中,多肽(例如麦胶蛋白酶)包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列。在一些方面中,多肽包含SEQ ID NO:1中所列的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约80%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约85%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约90%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约95%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约96%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约97%序列同一性的氨基酸序列。在一些方面中,多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约98%序列同一性的氨基酸序列。在一些方面中,多肽

包含与SEQ ID NO:1中所列的氨基酸序列具有至少约99％序列同一性的氨基酸序列。在一些方面中，多肽在对应于SEQ ID NO:1中的氨基酸467的氨基酸残基处包含Ser。在一些方面中，多肽在对应于SEQ ID NO:1中的氨基酸267的氨基酸残基处包含Glu。在一些方面中，多肽在对应于SEQ ID NO:1中的氨基酸271的氨基酸残基处包含Asp。

[0071] 在一些方面中，多肽(例如麦胶蛋白酶)包含与SEQ ID NO:8中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列。在一些方面中，多肽包含SEQ ID NO:8中所列的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约75％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约80％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约85％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约90％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约95％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约96％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约97％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约98％序列同一性的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约99％序列同一性的氨基酸序列。在一些方面中，多肽在对应于SEQ ID NO:3中的氨基酸278的氨基酸残基处包含Ser。在一些方面中，多肽在对应于SEQ ID NO:3中的氨基酸78的氨基酸残基处包含Glu。在一些方面中，多肽在对应于SEQ ID NO:3中的氨基酸82的氨基酸残基处包含Asp。

[0072] 在一些方面中，多肽(例如麦胶蛋白酶)包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列，其中多肽包含SEQ ID NO:8中所列的氨基酸序列。在一些方面中，多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列；其中多肽包含SEQ ID NO:8中所列的氨基酸序列；并且其中多肽在对应于SEQ ID NO:3中的氨基酸278的氨基酸残基处包含Ser、在对应于SEQ ID NO:3中的氨基酸78的氨基酸残基处包含Glu、以及在对应于SEQ ID NO:3中的氨基酸82的氨基酸残基处包含Asp。

[0073] 在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自N末端或C末端的一个或多个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自N末端的至少一个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自N末端的至少两个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自N末端的至少三个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自N末端的至少四个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自N末端的至少五个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自C末端的至少一个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自C

末端的至少两个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自C末端的至少三个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自C末端的至少四个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:1或6中所列的氨基酸序列自C末端的至少五个氨基酸缺失。

[0074] 在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自N末端或C末端的一个或多个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自N末端的至少一个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自N末端的至少两个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自N末端的至少三个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自N末端的至少四个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自N末端的至少五个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自C末端的至少一个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自C末端的至少两个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自C末端的至少三个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自C末端的至少四个氨基酸缺失。在一些方面中，多肽包含相对于SEQ ID NO:3或8中所列的氨基酸序列自C末端的至少五个氨基酸缺失。

[0075] 如以下实施例中所公开，根据本公开的一些方面的多肽是用于例如治疗乳糜泻的改进的多肽。多肽是对应于SEQ ID NO:4(KUMAMAX™，下文称为Kuma010；参见WO2013/023151，其以引用的方式整体并入本文中)的经加工(即成熟)多肽或预加工(即全长)多肽的变体。用于治疗乳糜泻的多肽能够降解富含脯氨酸(P)和谷氨酰胺(Q)的麸质组分，称为“麦胶蛋白”，据信这些组分是大多数乳糜泻患者免疫反应的主要来源。与公开为可用于治疗乳糜泻(参见例如WO2015/023728和WO2016/200880，其各自以引用的方式整体并入本文)和/或显示改进多肽产生的Kuma011和其它多肽相比，本公开的多肽在pH 4下在降解具有PQLP(SEQ ID NO:9)或QQP(SEQ ID NO:10)基序的肽(诸如PFPQPQLPY(SEQ ID NO:11)和/或PFPQPQQPF(SEQ ID NO:12))，其为代表麦胶蛋白的底物)时显示出优异的活性。因此，本公开的多肽构成了显著改进的用于治疗乳糜泻的治疗剂。

[0076] 在一些方面中，本文公开的多肽能够在pH 4下降解包含选自PFPQPQLPY(SEQ ID NO:11)、PFPQPQQPF(SEQ ID NO:12)、LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF(SEQ ID NO:13)和/或FLQPQQPFPQQPQQPYPQQPQQPFPQ(SEQ ID NO:14)的氨基酸序列的多肽。

[0077] 本公开的第一方面的多肽包含本公开的多肽酶的预加工型式。

[0078] 本公开的第一方面的多肽包含本公开的多肽酶的加工型式，并且也在pH 4下降解PFPQPQLPY(SEQ ID NO:11)肽和/或PFPQPQQPF(SEQ ID NO:12)肽，以及LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF(SEQ ID NO:13)和/或FLQPQQPFPQQPQQPYPQQPQQPFPQ(SEQ ID NO:14)。

[0079] 如本文所用，“至少75％同一”或“具有至少75％序列同一性”意指多肽相对于参考序列，例如相对于选自SEQ ID NO:1-8的氨基酸序列在其全长氨基酸序列中相差25％或更少(包括任何氨基酸取代、缺失、添加或插入)。在一些方面中，多肽包含与根据SEQ ID NO:1(预加工)或SEQ ID NO:8(加工)的氨基酸序列具有至少76％、77％、78％、79％、80％、81％、82％、83％、84％、85％、86％、87％、88％、89％、90％、91％、92％、92％、93％、94％、95％、

96％、97％、98％或99％同一性的氨基酸序列或由其组成。

[0080] 本公开的多肽的任何方面的多肽可包含来自SEQ ID NO:1或SEQ ID NO:8在2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23或全部24个(取决于方面)所述残基处的氨基酸取代。

[0081] 在本公开的第一方面的多肽的一个方面中，多肽在一个或多个选自由221D/N/Q/H、262E、268S/T/A、269L/T、270A/T/V、319A、354E/Q/R/Y、358S/Q/T、368F/Q、399Q、402S/Q、406S、424K、449E/N/Q、461R和463A/L/M/Q/R/T/V组成的组的残基处包含来自SEQ ID NO:6的一个或多个氨基酸取代。通篇使用的数字表示SEQ ID NO:6或SEQ ID NO:3多肽序列中的残基编号，并且数字右侧的单字母氨基酸缩写表示与存在于SEQ ID NO:6或3中所述位置的氨基酸残基相比可能的氨基酸取代。

[0082] 在本公开的第一方面的多肽的另一方面中，多肽在残基399和449处包含来自SEQ ID NO:6的氨基酸取代。在一个方面中，多肽包含氨基酸取代399Q和449Q。在一些方面中，基于SEQ ID NO:6的编号，多肽在位置399处包含Q且在位置449处包含Q。

[0083] 在本公开的第一方面的多肽的另一方面中，多肽包含358S和463T。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置358处的S，和(ii)位置463处的T，或(i)-(ii)的任何组合。

[0084] 在本公开的第一方面的多肽的一个方面中，多肽包含262E、269T、354Q、358S、399Q、449Q和463T。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置262处的E，(ii)位置269处的T，(iii)位置354处的Q，(iv)位置358处的S，(v)位置399处的Q，(vi)位置449处的Q，以及(vii)位置463处的T，或(i)-(vii)的任何组合。这些多肽在WO2016/200880中公开的实施例中得到广泛表征，如命名为Kuma030的多肽及其变体所例示。在本公开的第一方面的多肽的另一方面中，多肽包含319A、368F、399Q、449Q和I463T。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置319处的A，(ii)位置368处的F，(iii)位置399处的Q，(iv)位置449处的Q，以及(v)位置463处的T，或(i)-(v)的任何组合。这些多肽在WO2016/200880中公开的实施例中得到广泛表征，如命名为Kuma040的多肽及其变体所例示。在本公开的第一方面的多肽的另一方面中，多肽包含262E、269T、270V、354Q、358S、399Q和A449Q。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置262处的E，(ii)位置269处的T，(iii)位置270处的V，(vi)位置354处的Q，(v)位置358处的S、(vi)位置399处的Q，以及(vii)位置449处的Q，或(i)-(vii)的任何组合。这些多肽在WO2016/200880中公开的实施例中得到广泛表征，如命名为Kuma050的多肽及其变体所例示。在本公开的第一方面的多肽的一个方面中，多肽包含262E、269T、320M、354Q、358S、399Q、449Q和463T。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置262处的E，(ii)位置269处的T，(iii)位置320处的M，(vi)位置354处的Q，(v)位置358处的S、(vi)位置399处的Q、以及(vii)位置449处的Q，或(i)-(vii)的任何组合。这些多肽在WO2016/200880中公开的实施例中得到广泛表征，如命名为Kuma060的多肽及其变体所例示。在本公开的第一方面的多肽的又一方面中，多肽包含319A、320M、368F、399Q、449Q和463T。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置319处的A，(ii)位置320处的M，(iii)位置368处的F，(v)位置399处的Q，以及(v)位置449处的Q，或(i)-(v)的任何组合。这些多肽在WO2016/200880中公开的实施例中得到广泛表征，如命名为Kuma070的多肽及其变体所例示。如本文所用，术语“Kuma020”、“Kuma030”、

“Kuma040”、“Kuma050”和“Kuma070”是指具有与在WO2016/200880中公开的相同名称的相同多肽。

[0085] 在本公开的第一方面的多肽的另一方面中，多肽在一个或多个选自由105、171、172、173、174和456组成的组的氨基酸位置处包含来自SEQ ID NO:6的氨基酸取代。在一个方面中，氨基酸取代为105H;171R A或S;172R、A或S;173R或S、174S和/或456V。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置105处的H;(ii)位置171处的R、A或S;(iii)位置172处的R、A或S;(iv)和位置173处的R或S;(v)位置174处的S;(vi)位置456处的V;或(vii)(i)-(vi)的任何组合。在另一方面中，氨基酸取代为171R、172R和/或456V。在一些方面中，基于SEQ ID NO:6的编号，多肽包含(i)位置171处的R，(ii)位置172处的R，(iii)位置456处的V，或(iv)(i)-(iii)的任何组合。

[0086] 在本公开的第二方面的多肽的一个方面中，多肽在一个或多个选自由32D/N/Q/H、73E、79S/T/A、80L/T、81A/T/V、130A、165E/Q/R/Y、169S/Q/T、179F/Q、210Q、213S/Q、217S、235K、260E/N/Q、272R和274A/L/M/Q/R/T/V组成的组的残基处包含一个或多个来自SEQ ID NO:3的氨基酸取代。在本公开的第二方面的多肽的另一方面中，多肽在残基210和260处包含来自SEQ ID NO:3的氨基酸取代。在本公开的第二方面的多肽的另一方面中，多肽包含氨基酸取代210Q和260Q。在一些方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置210处的Q，(ii)位置260处的Q，或(i)-(ii)的任何组合。在本公开的第二方面的多肽的一个方面中，多肽包含169S和274T。(Kuma020属)。在此类方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置169处的S，(ii)位置274处的T，或(iv)(i)-(ii)的任何组合。在本公开的第二方面的多肽的另一方面中，多肽包含73E、80T、165Q、169S、210Q、260Q和274T。(Kuma030属)。在此类方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置73处的E，(ii)位置80处的T，(iii)位置165处的Q，(iv)位置169处的S，(v)位置210处的Q，(vi)位置260处的Q和(vii)位置274处的T或(i)-(vii)的任何组合。在本公开的第二方面的多肽的另一方面中，多肽包含130A、179F、210Q、260Q和274T。(Kuma040属)。在此类方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置130处的A，(ii)位置179处的F，(iii)位置210处的Q，(iv)位置260处的Q，(v)位置274处的T，或(i)-(v)的任何组合。在本公开的第二方面的多肽的另一方面中，多肽包含73E、80T、81V、165Q、169S、210Q和260Q。(Kuma050属)。在此类方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置73处的E，(ii)位置80处的T，(iii)位置81处的V，(iv)位置165处的Q，(v)位置169处的S，(vi)位置210处的Q，(vii)位置260处的Q或(i)-(vii)的任何组合。在本公开的第二方面的多肽的一个方面中，多肽包含73E、80T、320M、165Q、169S、210Q、260Q和274T。(Kuma060属)。在此类方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置73处的E，(ii)位置80处的T，(iii)位置320处的M，(iv)位置165处的Q，(v)位置169处的S，(vi)位置210处的Q，(vii)位置260处的Q，(viii)位置274处的T，或(i)-(vii)的任何组合。在本公开的第二方面的多肽的另一方面中，多肽包含130A、131M、179F、210Q、260Q和274T。(Kuma070属)。在此类方面中，基于SEQ ID NO:3的编号，多肽包含(i)位置130处的A，(ii)位置131处的M，(iii)位置179处的F，(iv)位置210处的Q，(v)位置260处的Q，(vi)位置274处的T或(i)-(vi)的任何组合。在本公开的第二方面的多肽的另一方面中，多肽在一个或多个选自由267组成的组的氨基酸位置处包含来自SEQ ID NO:3的氨基酸取代。在一个方面中，氨基酸取代为267V。在此类方面中，基于SEQ ID NO:3的编号，多肽在位置267处包含V。

[0087] 在本公开的任何方面的多肽的另一方面中，多肽在多肽的C末端还包含组氨酸标签，以促进多肽的分离。可使用任何合适的组氨酸标签；在一个方面中，标签与TEV蛋白酶切割位点(ENLYFQS)(SEQ ID NO:18)连接以允许在纯化后用TEV蛋白酶有效移除，例如标签可包含氨基酸序列GSTENLYFQSGALEHHHHHH(SEQ ID NO:17)或由其组成。在另一方面中，组氨酸标签为a.可裂解组氨酸标签，其允许更容易地移除His标签。在一个方面中，可裂解的组氨酸标签包含氨基酸序列X_NPQ(L/Q)PX_NHHHHHH(SEQ ID NO:15)，其中X_N为介于1-25个之间氨基酸残基的接头。在一个非限制性实例中，可裂解组氨酸标签包含氨基酸序列GSSGSSGSQPQLPYGSSGSSGSHHHHHH(SEQ ID NO:16)。

[0088] 在本公开的多肽的任何方面的一个方面中，与SEQ ID NO:6或SEQ ID NO:3相比的氨基酸取代可包含表2或3中所指出的一个或多个取代。发现这些位置处的取代通常良好耐受(即通常对活性产生轻微影响或无影响)，并且在一些情况下使本公开的多肽的活性增加不超过20％。

[0089] 表2.相对于Kuma010位置处的可能氨基酸取代。

[0090]

残基数(预加工/加工)	残基	残基数(预加工/加工)	残基
221/32	D、N、Q、H	358/169	A、S、N、Q、T
261/72	A、R、N、D、C、Q、E、G、H、I、L、K、M、S、T、W、Y、V	368/179	A、R、N、C、Q、E、G、K、M、F、S、T、W、Y
262/73	A、R、N、D、C、	397/208	A、C、F、Y

[0091]

	Q、E、G、H、I、L、M、F、T、W、Y、V		
264/75	A、N、D、C、Q、E、G、S、T、Y	399/210	Q、N
266/77	A、C、S	402/213	Q、N、S
268/79	S、T	406/217	S
269/80	L、T	424/235	K
270/81	A、R、N、D、C、Q、E、G、I、K、S、T、V	446/257	G、S
317/128	A、N、C、G、T、V	448/259	A、R、N、D、C、Q、E、G、H、I、L、K、M、F、S、T、W、Y、V
318/129	A、R、N、D、C、Q、E、G、H、L、K、M、F、S、T、Y、V	449/260	Q、E、G、N
319/130	A、N、D、C、Q、H、M、T	456/267	A、N、D、C、Q、E、G、H、L、S、T、V
320/131	A、R、N、D、C、Q、K、M、S	461/272	R
350/161	N、D、C、G、S、	463/274	A、R、N、D、

	T		C、Q、E、G、H、L、K、M、F、S、T、W、Y、V
351/162	G、S	464/277	A、N、D、C、S，
353/164	A、R、N、C、Q、E、G、I、K、M、S、T、V	466/279	D、C、G、S
354/165	A、R、N、D、C、Q、E、G、H、L、K、M、F、T、W、Y		

[0092]

[0093] 在本公开的多肽的任何方面的另一实施方案中，与SEQ ID NO:6或SEQ ID NO:3相比的氨基酸取代可包含表3中指出的一个或多个取代。

[0094] 表3

残基数(预加工/加工)	残基	残基数(预加工/加工)	残基
221/32	D、N、Q、H	358/169	A、S、N、Q、T
261/72	S	368/179	A、N、D、Q、E、S、T
262/73	A、R、N、D、Q、E、G、L、M、T	402/213	Q、S
264/75	A	406/217	S

[0095]

[0096]

268/79	S、T	424/235	K
269/80	L、T	446/257	S
270/81	A、T、V	449/260	Q、N、A
317/128	A、T	456/267	V
319/130	A	461/272	R
354/165	A、R、N、D、Q、E、K、T、Y	463/274	A、R、Q、L、M、T、V

[0097] 在本公开的多肽的任何方面的另一实施方案中，本公开的多肽的每个残基处的氨基酸可如表4中所指出，所述表列出了通过计算诱变分析预测的在多肽酶的每个位置处的所有可能突变。如在WO2016/200880中公开的实施例中所述，使用简并引物在活性位点中存在的每个位置(残基261-264、266-267、270、317-320、350-354、368、397、403-404、446、448、456和463-468)处测试突变，以测试各种氨基酸取代对活性的影响；不干扰活性的那些可并入本公开的多肽中，如表4所示。

[0098] 表4：与Kuma 010相关的残基处的可能氨基酸

[0099]

全长	成熟	氨基酸可能性
190	1	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
191	2	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PHE、PRO、SER、THR、TRP、VAL
192	3	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、PRO、SER、TRP、

		TYR
193	4	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
194	5	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
195	6	ALA、ASN、CYS、GLN、HIS、LEU、MET、PHE、THR、TYR
196	7	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、THR、TRP、TYR
197	8	ALA、GLY、PRO、SER
198	9	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
199	10	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
200	11	ALA、ASN、ASP、CYS、GLY、ILE、SER、THR、VAL
201	12	ALA、CYS、GLY、SER
202	13	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
203	14	ALA、GLY、SER
204	15	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、

[0100]

		ILE、LEU、LYS、MET、PHE、SER、THR、TYR
205	16	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
206	17	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TYR、VAL
207	18	ALA、CYS、GLN、GLU、GLY、LYS、PRO、SER、THR、TRP
208	19	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
209	20	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
210	21	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、LEU、MET、SER、THR、VAL
211	22	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、THR、TYR
212	23	GLY
213	24	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
214	25	GLY
215	26	ALA、ASN、ASP、CYS、GLN、GLU、GLY、SER、THR

[0101]

[0102]

216	27	ALA、ASN、ASP、CYS、GLN、GLY、SER、THR、VAL
217	28	ALA、CYS、ILE、LEU、SER、THR、VAL
218	29	ALA、GLY、SER
219	30	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、MET、SER、THR、VAL
220	31	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、SER、THR、VAL
221	32	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、SER、THR、VAL
222	33	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、SER、THR、VAL
223	34	ALA、ARG、ASN、ASP、CYS、GLU、GLY、LYS、MET、SER
224	35	GLY
225	36	GLY
226	37	ALA、ARG、ASN、ASP、CYS、GLU、GLY、HIS、LEU、PHE、SER、THR、TRP、TYR
227	38	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、LYS、MET、SER
228	39	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
229	40	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、

[0103]

230	41	ALA、GLY、SER
231	42	ALA、ASN、ASP、CYS、GLN、GLU、GLY、LEU、SER、THR
232	43	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
233	44	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
234	45	ALA、ASN、CYS、GLY、HIS、PHE、SER、TYR
235	46	ALA、ASN、ASP、CYS、HIS、MET、PHE、SER、THR、TRP、TYR
236	47	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
237	48	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
238	49	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、MET、SER、THR、VAL
239	50	GLY
240	51	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、SER、THR、TYR、VAL
241	52	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL

[0104]

242	53	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PRO、SER、THR、VAL
243	54	ALA、GLY、PRO、SER
244	55	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
245	56	ALA、ASN、CYS、GLY、SER、THR、VAL
246	57	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
247	58	ALA、ARG、ASP、CYS、GLY、ILE、LYS、MET、PRO、SER
248	59	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
249	60	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、PRO、SER、THR
250	61	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、SER、THR、VAL
251	62	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TYR、VAL
252	63	ASN、ASP、GLY、SER
253	64	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、THR、TRP
254	65	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、

		HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
255	66	ALA、ARG、ASN、ASP、CYS、MET、SER、THR
256	67	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
257	68	ALA、ARG、ASN、CYS、GLN、GLU、GLY、ILE、LYS、MET、PRO、SER、THR、VAL
258	69	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
259	70	GLY
260	71	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
261	72	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PRO、SER、THR、TRP、TYR、VAL
262	73	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
263	74	GLY
264	75	ALA、ASN、ASP、CYS、GLN、GLU、GLY、PRO、SER、THR、TRP
265	76	ALA、ASN、ASP、CYS、GLN、GLU、GLY、SER、THR、VAL

[0105]

[0106]

266	77	ALA、CYS、GLY、SER
267	78	GLU
268	79	ALA、ASN、ASP、CYS、GLY、SER、THR、VAL
269	80	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
270	81	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、SER、THR、VAL
271	82	ASP
272	83	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、MET、SER、THR、VAL
273	84	ALA、ASN、ASP、CYS、GLN、GLU、GLY、SER、THR
274	85	ALA、ASN、ASP、CYS、GLY、ILE、SER、THR、VAL
275	86	ALA、CYS、GLY、SER
276	87	GLY
277	88	ALA、GLY、SER
278	89	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、MET、SER、THR、VAL
279	90	ALA、GLY、SER
280	91	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、MET、PHE、PRO、SER、TRP、TYR
281	92	GLY
282	93	ALA、GLY、SER
283	94	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、

[0107]

		HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
284	95	CYS、HIS、ILE、LEU、MET、PHE、THR、TYR、VAL
285	96	ALA、GLY、SER
286	97	ALA、ASN、ASP、CYS、GLY、SER、THR、VAL
287	98	ALA、ASN、ASP、CYS、GLN、HIS、LEU、PHE、SER、TYR
288	99	HIS、PHE
289	100	ALA、GLY、SER
290	101	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
291	102	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL、
292	103	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
293	104	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PRO、SER、THR、VAL
294	105	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
295	106	GLY
296	107	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、

[0108]

		LEU、LYS、MET、PHE、SER、THR、VAL
297	108	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
298	109	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
299	110	ALA、GLY、SER
300	111	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
301	112	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、PHE、SER、THR、VAL
302	113	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、SER、THR、TRP、VAL
303	114	ALA、GLY、SER
304	115	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、SER、THR、VAL
305	116	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
306	117	ALA、ASN、ASP、SER
307	118	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
308	119	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、

[0109]

		TYR、VAL、
309	120	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
310	121	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
311	122	ALA、CYS、GLY、PRO、SER
312	123	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、THR、TRP、TYR
313	124	ALA、CYS、GLY、ILE、SER、THR、VAL
314	125	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、SER、THR、VAL
315	126	ALA、CYS、GLY、SER、THR
316	127	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、MET、SER、THR、VAL
317	128	ALA、ASN、CYS、GLY、SER、THR、VAL
318	129	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
319	130	ALA、ASN、ASP、CYS、GLN、GLY、HIS、MET、SER、THR
320	131	ALA、ARG、ASN、ASP、CYS、GLN、GLY、LYS、MET、SER
321	132	ALA、CYS、GLY、PRO、SER

[0110]

322	133	ALA、ASP、CYS、GLN、GLU、GLY、LEU、SER
323	134	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、VAL、
324	135	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
325	136	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、TRP、TYR
326	137	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
327	138	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
328	139	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
329	140	ALA、ASP、CYS、GLY、SER
330	141	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
331	142	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
332	143	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、

[0111]

		TYR、VAL、
333	144	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、SER、THR、VAL
334	145	ALA、ARG、ASN、ASP、CYS、GLU、GLY、MET、SER、THR、VAL
335	146	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
336	147	ALA、ARG、CYS、GLN、GLU、GLY、MET、SER
337	148	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
338	149	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
339	150	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
340	151	ALA、ASN、ASP、GLY、SER
341	152	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、SER、THR、VAL
342	153	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
343	154	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、

		TYR、VAL、
344	155	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
345	156	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TYR、VAL
346	157	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、THR
347	158	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LYS、MET、PRO、SER、THR、VAL
348	159	ALA、ASN、ASP、CYS、GLN、GLU、GLY、LEU、SER、THR、VAL
349	160	ALA、CYS、GLY、SER、THR
350	161	ALA、ASN、ASP、CYS、GLY、SER、THR
351	162	ALA、GLY、SER
352	163	GLY
353	164	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LYS、MET、SER、THR、VAL
354	165	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
355	166	GLY
356	167	ALA、GLY、SER
357	168	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、MET、SER、THR、VAL

[0112]

[0113]

358	169	ALA、GLY、SER
359	170	ASN、GLY
360	171	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
361	172	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
362	173	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
363	174	ASN、ASP、GLY、SER
364	175	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
365	176	ALA、ARG、ASN、ASP、CYS、GLY、HIS、MET、PHE、SER、THR、TRP、TYR
366	177	ALA、ASN、ASP、CYS、HIS、LYS、SER
367	178	ALA、ASP、CYS、GLY、SER、THR、VAL
368	179	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、THR、TRP、TYR
369	180	ALA、CYS、HIS、PHE、SER、TYR
370	181	ALA、ASP、CYS、GLY、PRO、SER
371	182	ALA、GLY、SER
372	183	ALA、CYS、GLY、SER
373	184	ALA、GLY、SER

[0114]

374	185	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PRO、SER、THR、TRP、VAL
375	186	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
376	187	ALA、ASN、ASP、CYS、GLY、HIS、ILE、LEU、SER、THR、VAL
377	188	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、SER、THR、VAL
378	189	ALA、GLY、SER
379	190	ALA、ASP、CYS、GLY、SER、THR
380	191	GLY
381	192	GLY
382	193	ALA、CYS、GLY、SER、THR
383	194	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
384	195	ALA、ASN、ASP、CYS、GLN、GLU、GLY、LEU、SER、THR
385	196	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、SER、THR、TRP、VAL
386	197	ALA、CYS、GLY、MET、SER、THR
387	198	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、

[0115]

388	199	ASN、ASP、GLY、LYS、SER
389	200	GLY
390	201	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
391	202	ALA、ASN、ASP、CYS、GLN、GLY、ILE、MET、PRO、SER、THR、VAL
392	203	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
393	204	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
394	205	ALA、CYS、GLN、GLU、GLY、SER、THR
395	206	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LYS、MET、SER、THR、VAL
396	207	ALA、CYS、GLY、SER、THR、VAL
397	208	ALA、CYS、PHE、TRP、TYR
398	209	ARG、ASN、ASP、CYS、GLN、MET、SER
399	210	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、LEU、LYS、MET、SER
400	211	GLY
401	212	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL

[0116]

402	213	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
403	214	GLY
404	215	GLY
405	216	ALA、GLY、SER
406	217	ALA、CYS、GLY、SER、THR
407	218	GLY
408	219	GLY
409	220	GLY
410	221	ALA、ASN、CYS、GLY、ILE、SER、THR、VAL
411	222	ALA、GLY、SER
412	223	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、SER、THR、VAL
413	224	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LYS、MET、PHE、SER、THR、TYR、VAL
414	225	ALA、ASN、CYS、GLN、GLU、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TYR、VAL
415	226	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
416	227	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、PRO、SER、THR、VAL
417	228	ALA、CYS、GLN、GLU、GLY、MET、PRO、SER、THR

[0117]

418	229	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
419	230	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、PHE、SER、TRP、TYR
420	231	GLN、GLU
421	232	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
422	233	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
423	234	ALA、GLY、SER
424	235	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
425	236	ALA、CYS、GLY、PRO、SER、THR、VAL
426	237	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
427	238	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
428	239	ALA、ASN、ASP、CYS、GLN、GLU、GLY、SER、THR、VAL
429	240	ALA、ASN、ASP、CYS、GLY、SER

[0118]

430	241	ALA、ASN、ASP、CYS、GLY、SER、THR
431	242	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
432	243	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
433	244	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
434	245	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
435	246	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
436	247	GLY
437	248	ALA、ARG、ASN、CYS、GLN、SER、THR
438	249	GLY
439	250	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、MET、SER、THR、VAL
440	251	ALA、GLY、PRO、SER
441	252	ASP
442	253	ALA、ASN、ASP、CYS、GLN、GLU、GLY、LEU、MET、SER、THR
443	254	ALA、GLY、SER

444	255	ALA、GLY
445	256	ALA、ASN、ASP、CYS、GLY、SER
446	257	ALA、GLY、SER
447	258	ALA、ASN、ASP、CYS、GLY、SER、THR
448	259	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
449	260	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
450	261	ALA、ASN、ASP、CYS、GLY、HIS、SER、THR
451	262	GLY
452	263	ALA、ASN、CYS、GLN、HIS、ILE、LEU、PHE、SER、THR、TYR、VAL
453	264	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
454	265	ALA、ASN、ASP、CYS、GLY、SER、THR、VAL
455	266	ALA、ARG、ASN、ASP、CYS、GLU、GLY、HIS、ILE、MET、PHE、SER、THR、TRP、TYR、VAL
456	267	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、SER、THR、VAL
457	268	ALA、ASN、ASP、CYS、GLY、ILE、MET、SER、THR、TRP、VAL
458	269	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、

[0119]

[0120]

		LYS、MET、SER
459	270	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
460	271	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TYR、VAL、
461	272	ALA、ASN、ASP、CYS、GLN、GLY、HIS、LYS、MET、SER、THR
462	273	ALA、ARG、ASN、ASP、CYS、GLN、GLY、HIS、ILE、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
463	274	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL、
464	275	GLY
465	276	GLY
466	277	ALA、ASN、ASP、CYS、GLY、SER、THR
467	278	**SER**
468	279	ALA、ASP、CYS、GLY、SER
469	280	ALA、ASN、ASP、CYS、GLY、SER、THR、VAL
470	281	ALA、GLY、SER
471	282	ALA、CYS、GLY、PRO、SER
472	283	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、MET、SER、THR、VAL

[0121]

473	284	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
474	285	ALA、GLY、SER
475	286	ALA、GLY、SER
476	287	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、SER、THR、VAL
477	288	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、SER、THR、VAL
478	289	ALA、GLY、SER
479	290	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、SER、THR、TRP、TYR
480	291	ALA、ARG、ASN、ASP、CYS、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
481	292	ALA、ASN、ASP、CYS、GLN、GLU、GLY、MET、SER
482	293	ALA、GLN、GLU、HIS、LYS、THR
483	294	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、PHE、SER、TRP、TYR
484	295	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
485	296	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、TRP、TYR、VAL
486	297	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、

[0122]

		TYR、VAL
487	298	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PRO、SER、THR、TRP、VAL
488	299	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、SER、THR、VAL
489	300	GLY
490	301	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
491	302	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、VAL
492	303	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PRO、SER、THR、VAL
493	304	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
494	305	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
495	306	ALA、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、MET、SER、THR
496	307	ALA、HIS、PHE、SER、THR、TYR
497	308	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、

[0123]

		TYR
498	309	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、LEU、MET、SER、THR
499	310	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR
500	311	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
501	312	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
502	313	ALA、ASN、ASP、CYS、GLY、ILE、MET、SER、THR、VAL
503	314	ALA、ASN、ASP、CYS、HIS、LEU、MET、PHE、SER、THR、TYR、VAL
504	315	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
505	316	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PRO、SER、THR、TRP、VAL
506	317	ALA、ASN、ASP、CYS、GLN、GLY、ILE、SER、THR、VAL
507	318	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PHE、SER、THR、TRP、TYR、VAL

[0124]

508	319	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
509	320	GLY
510	321	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、SER、THR、TRP、TYR
511	322	ALA、ASN、ASP、CYS、GLY、SER
512	323	ALA、ASN、ASP、CYS
513	324	ALA、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LYS、MET、SER、THR、VAL
514	325	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
515	326	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LYS、MET、SER
516	327	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
517	328	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
518	329	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
519	330	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、

[0125]

		TYR、VAL
520	331	HIS、PHE、THR、TRP、TYR
521	332	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TYR、VAL
522	333	ALA、GLY、SER
523	334	CYS、GLY、HIS、LYS、MET、PHE、SER、TYR
524	335	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、PRO、SER、THR、TRP、TYR、VAL
525	336	GLY
526	337	HIS、PHE、TRP
527	338	ALA、ASN、ASP、CYS、SER
528	339	ALA、GLY、PRO、SER
529	340	ALA、ASP、CYS、GLY、SER、THR
530	341	ALA、ASN、CYS、GLY、SER、THR、VAL
531	342	GLY
532	343	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、LEU、LYS、MET、SER
533	344	GLY
534	345	ALA、CYS、GLY、SER、THR
535	346	ALA、CYS、GLY、PRO、SER、THR
536	347	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、ILE、LEU、LYS、MET、PHE、SER、THR、TYR、VAL

[0126]

537	348	GLY
538	349	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
539	350	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、LEU、LYS、MET、PHE、SER、THR、TRP、TYR
540	351	ALA、ASN、ASP、CYS、GLN、GLU、GLY、LEU、LYS、SER、THR、VAL
541	352	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、LEU、LYS、MET、SER、THR
542	353	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP、TYR、VAL
543	354	ALA、ARG、CYS、GLN、GLU、GLY、MET、SER、THR
544	355	ALA、ASN、ASP、CYS、GLN、GLU、GLY、LEU、MET、SER、THR
545	356	ALA、ARG、ASN、ASP、CYS、GLN、GLU、GLY、HIS、ILE、LEU、LYS、MET、PHE、SER、THR、TRP
546	357	任何残基
547	358	任何残基
548	359	任何残基
549	360	任何残基
550	361	任何残基

[0127]

551	362	任何残基
552	363	任何残基
553	364	任何残基

[0128] 在一些方面中，本文公开的多肽序列还包含组氨酸标签。在一些方面中，组氨酸标签在多肽的C末端与多肽融合。可以使用任何合适的组氨酸标签。在一些方面中，组氨酸标签与TEV蛋白酶切割位点(ENLYFQS)(SEQ ID NO:18)连接以允许在纯化后用TEV蛋白酶有效移除，例如，标签可包含氨基酸序列GSTENLYFQSGALEHHHHHH(SEQ ID NO:17)或由其组成。在另一方面中，在多肽序列的C末端并入可裂解组氨酸标签，其包含氨基酸序列X_NPQ(L/Q)PX_NHHHHHH(SEQ ID NO:15)，其中X_N是介于1-25个之间氨基酸残基的接头。在一个非限制性实例中，可裂解组氨酸标签包含氨基酸序列GSSGSSGSQPQLPYGSSGSSGSHHHHHH(SEQ ID NO:16)。

[0129] 如表5所示，相对于Kuma010/011氨基酸序列的点取代可影响催化活性。表5列出了单个突变在催化各种麦胶蛋白肽序列降解方面的有效性。WO2016/200880中公开的实施例提供了关于特定个体和组合突变体的进一步数据。

[0130] 表5

[0131]

位置(全长)	位置(截短)	Kuma010 A.A.	A.A.相对于Kuma010/011	PFPQPQLPY(SEQ ID NO:11)的改进百分比	PFPQPQQPF(SEQ ID NO:12)的改进百分比
221	32	E	D、N、Q、H	105%	ND
262	73	K	E	109%	110%
268	79	V	A	107%	89%
268	79	V	S	104%	83%

[0132]

268	79	V	T	127%	105%
269	80	E	L	113%	84%
269	80	E	T	263%	191%
270	81	L	A	203%	92%
270	81	L	T	307%	29%
270	81	L	V	474%	61%
319	130	S	A	154%	184%
354	165	S	A	152%	140%
354	165	S	E	124%	120%
354	165	S	Q	145%	141%
354	165	S	R	109%	82%
354	165	S	Y	46%	105%
358	169	G	N	120%	99%
358	169	G	S	331%	224%
358	169	G	Q	147%	149%
358	169	G	T	283%	128%
368	179	H	F	334%	104%
368	179	H	Q	199%	195%
399	210	D	Q	149%	208%
402	213	D	S	94%	108%
402	213	D	Q	164%	111%
406	217	T	S	84%	101%
424	235	N	K	285%	ND
449	260	A	E	149%	208%
449	260	A	N	119%	118%

[0133]

461	272	T	R	120%	86%
463	274	I	A	51%	234%
463	274	I	L	124%	22%
463	274	I	M	123%	53%
463	274	I	Q	129%	69%
463	274	I	R	29%	110%
463	274	I	T	130%	239%
463	274	I	V	256%	141%

[0134] 在某些方面中，本公开提供了包含至少一个提高多肽产量的突变的多肽。在某些方面中，提高产量的突变提供以下三类之一的改进：1.改变纯化方法；2.产量增加；和3.降低纯化期间发生酶促自加工的可能性，从而简化分析。可通过本文公开的多肽的蛋白水解活性移除的His标签的添加属于第1类；R105H突变体似乎将产量提高了约2倍，将此突变归入第2类；并且位置171-174处的突变将这些突变体归入第3类。

[0135] 在本申请全文中使用的术语“多肽”以其最广泛的意义使用，是指亚基氨基酸序列，无论是天然存在的还是合成来源的。本公开的多肽可包含L-氨基酸、D-氨基酸(其在体内对L-氨基酸特异性蛋白酶具有抗性)或D-和L-氨基酸的组合。本文所述的多肽可为化学合成的或重组表达的。多肽可与其它化合物连接以促进体内半衰期的延长，诸如通过PEG化、HES化、PAS化或糖基化。如本领域技术人员所理解，此类键联可为共价的或非共价的。在一些方面中，多肽与任何其它合适的接头连接，包括但不限于可用于纯化或检测的任何接头(诸如FLAG或His标签)。

[0136] A.核酸

[0137] 在另一方面中，本公开提供编码本公开的任何方面的多肽的分离的核酸。下文显示编码Kuma062-M的示例性核酸。

[0138]

AGTGATATGGAAAAACCGTGGAAAGAAGGTGAAGAAGCCCGCGCAGTGCTGCAAGGT
CATGCTCGTGCGCAGGCACCGCAAGCAGTCGATAAAGGCCCGGTGGCAGGTGACGAA
CGCATGGCTGTTACCGTGGTTCTGCGTCGCCAGCGTGCAGGTGAACTGGCGGCCCAC
GTGGAACGTCAAGCAGCTATTGCTCCGCATGCGCGCGAACACCTGAAACGTGAAGCG
TTTGCGGCCAGTCATGGTGCGTCCCTGGATGACTTTGCCGAACTGCGTCGCTTCGCA
GATGCTCACGGCCTGGCGCTGGACCGTGCAAACGTTGCAGCTGGCACCGCCGTTCTG
TCTGGTCCGGACGATGCAATCAATCGCGCTTTTGGTGTGGAACTGCGTCATTTCGAT
CACCCGGACGGCTCATATCGTTCGTACCTGGGTGAAGTCACCGTGCCGGCCAGTATT
GCACCGATGATCGAAGCGGTTCTGGGCCTGGATACGCGTCCGGTCGCCCGCCGTCGT
TTTCGTATGCAGCGTCGCGCAGAAGGCGGTTTCGAAGCTCGTTCCCAAGCGGCGGCA
CCGACCGCATATACGCCGCTGGATGTTGCGCAGGCCTACCAATTTCCGGAAGGTCTG
GACGGCCAGGGTCAATGCATTGCCATTATCGAACTGGGCGGTGGCTATGATGAAGCT
TCACTGGCGCAGTACTTCGCGTCGCTGGGCGTGCCGGCACCGCAAGTCGTGAGTGTT
TCCGTCGATGGTGCGAGCAACCAGCCGACCGGTGATCCGGAAGGTCCGGACGGTGAA
GTGACCCTGGATATCGAAGTTGCAGGCGCTCTGGCGCCGGGTGCCAAATTTGCAGTG
TATTTCGCGCCGGATACCACTGCCGGTTTTCTGGACGCGATTACCACGGCCATCCAC
GATCCGACGCTGAAACCGAGCGTTGTCTCAATTTCGTGGAGCATGCCGGAAGACAGC
TGGACCTCTGCTGCGATCGCCGCAATGAACCGTGCGTTTCTGGATGCTGCGGCCCTG
GGTGTGACCGTTCTGGCAGCTGCGGGCGACCAGGGTTCTACGAGCGGCGAACAGGAC
GGTCTGTATCATGTGCATTTCCCGGCCGCATCACCGTACGTTCTGGCGTGCGGTGGC
ACGCGCCTGGTCGCATCGGGTGGCCGTATTGCGCAGGAAACCGTCTGGAACCAGGGT
CCGGACGGTGGTGCAACGGGTGGCGGTGTGAGCCGCATCTTCCCGCTGCCGGCATGG
CAGGAACACGCTAACGTTCCGCCGTCTGCAAATCCGGGCGCGAGCAGCGGCCGTGGT
GTCCCGGATCTGGCTGGTAATGCGGACCCGCAGACCGGTTATGAAGTGGTTATTGAT

GGCGAAGCAACCGTCACCGGCGGTACGAGCGCCGTGGCACCGCTGTTTGCTGCGCTG
GTTGCGCGTATTAACCAGAAACTGGGCAAAGCAGTTGGTTATCTGAATCCGACCCTG
TACCAACTGCCGGCAGATGTTTTCCATGACATCACGGAGGGTAACAATGATATTGCA
AACCGTGCGCAGATTTATCAAGCAGGTCCGGGCTGGGACCCGTGTACCGGTCTGGGT
TCACCGATTGGTGTGCGTCTGCTGCAAGCACTGTTGCCGAGTGCCTCCCAGCCGCAA
CCGTGA

SEQ ID NO: 22

[0139] 分离的核酸序列可包含RNA或DNA。如本文所用,“分离的核酸”为已从基因体或cDNA序列中其正常周围核酸序列中移除的那些。此类分离的核酸序列可包含可用于促进所编码蛋白的表达和/或纯化的额外序列,包括但不限于polyA序列、经修饰的Kozak序列和编码表位标签、输出信号和分泌信号、核定位信号和质膜定位信号的序列。基于本文的教示,

本领域技术人员将显而易知什么核酸序列将编码本公开的多肽。

[0140] 在另一方面中，本公开提供包含与合适的控制序列可操作地连接的本公开的任何方面的分离的核酸的核酸表达载体。“重组表达载体”包括将核酸编码区或基因可操作地连接至能够影响基因产物表达的任何控制序列的载体。与本公开的核酸序列可操作地连接的“控制序列”是能够影响核酸分子表达的核酸序列。控制序列不需要与核酸序列相邻，只要它们起到指导其表达的作用。因此，例如，插入的未翻译但转录的序列可存在于启动子序列与核酸序列之间，并且启动子序列仍然可以认为与编码序列“可操作地连接”。其它此类控制序列包括但不限于聚腺苷酸化信号、终止信号和核糖体结合位点。此类表达载体可为本领域中已知的任何类型，包括但不限于基于质粒和病毒的表达载体。用于驱动哺乳动物系统中公开的核酸序列表达的控制序列可为组成型的(由多种启动子中的任一种驱动，包括但不限于CMV、SV40、RSV、肌动蛋白、EF)或诱导型的(由多种诱导型启动子中的任一种驱动，包括但不限于四环素、蜕皮激素、类固醇反应性启动子)。用于转染原核细胞的表达载体的构建也为本领域中熟知的，因此可通过标准技术完成。(参见例如Sambrook,Fritsch和Maniatis,in:Molecular Cloning,A Laboratory Manual,Cold Spring Harbor Laboratory Press,1989;Gene Transfer and Expression Protocols,第109页-第128页,E.J.Murray编,The Humana Press Inc.,Clifton,N.J.),和the Ambion 1998Catalog(Ambion,Austin,TX)。表达载体必须可在宿主生物体中作为附加体或通过整合至宿主染色体DNA中进行复制。在一个优选方面中，表达载体包含质粒。然而，本公开旨在包括具有等效功能的其它表达载体，诸如病毒载体。

[0141] B.宿主细胞

[0142] 在另一方面中，本公开提供包含本公开的核酸表达载体的重组宿主细胞。任何能够产生重组蛋白的宿主细胞均可用于本文公开的方法中。宿主细胞可为原核的或真核的。在一些方面中，宿主细胞为原核细胞。合适的原核宿主细胞的非限制性实例包括大肠杆菌(Escherichia coli)、枯草杆菌(Bacillus subtilis)、新月柄杆菌(Caulobacter crescentus)、类球红细菌(Rodhobacter sphaeroides)、假交替单胞菌(Pseudoalteromonas haloplanktis)、希瓦氏菌属Ac10菌株(Shewanellasp.strain Ac10)、荧光假单胞菌恶臭假单孢菌(Pseudomonas fluorescensi Pseudomonas putida)、铜绿假单胞菌(Pseudomonas aeruginosa)、高嗜盐菌(Halomonas elongata)、需盐色盐杆菌(Chromohalobacter salexigens)、变铅青链霉菌(Streptomyces lividans)、灰色链霉菌(Streptomyces griseus)、耐内酰胺诺卡氏菌(Nocardia lactamdurans)、耻垢分枝杆菌(Mycobacterium smegmatis)、谷氨酸棒状杆菌(Corynebacterium glutamicum)、产氨棒杆菌(Corynebacterium ammoniagenes)、乳酸发酵短杆菌(Brevibacterium lactofermentum)、枯草芽孢杆菌(Bacillus subtilis)、短芽孢杆菌(Bacillus brevis)、巨大芽孢杆菌(Bacillus megaterium)、地衣芽孢杆菌(Bacillus licheniformis)、解淀粉芽孢杆菌(Bacillus amyloliquefaciens)、乳酸乳球菌(Lactococcus lactis)、植物乳杆菌(Lactobacillus plantarum)、干酪乳杆菌(Lactobacillus casei)、罗伊氏乳杆菌(Lactobacillus reuteri)和格氏乳杆菌(Lactobacillus gasseri)。在一些方面中，宿主细胞为真核细胞。合适的真核宿主细胞的非限制性实例包括酿酒酵母(Saccharomyces cerevisiae)和构巢曲霉(Aspergillus nidulans)。细胞可瞬时或稳定转染或转导。可经由

本领域中已知的任何技术，包括但不限于标准细菌转化、磷酸钙共沉淀、电穿孔或脂质体介导的、DEAE葡聚糖介导的、聚阳离子介导的或病毒介导的转染。(参见例如Molecular Cloning:ALaboratory Manual(Sambrook等人,1989,Cold Spring Harbor Laboratory Press;Culture of Animal Cells:A Manual of Basic Technique,第2版.(R.I.Freshney.1987.Liss,Inc.New York,NY)。产生根据本公开的多肽的方法是本公开的附加部分。所述方法包括以下步骤：(a)在有助于表达多肽的条件下培养根据本公开的此方面的宿主，和(b)任选地，回收所表达的多肽。所表达的多肽可从无细胞提取物、细胞团块中回收，或从培养基中回收。纯化重组表达的多肽的方法为本领域技术人员熟知的。

[0143] C.药物组合物

[0144] 在另一方面中，本公开提供药物组合物，所述药物组合物包含本公开的任何方面或方面的多肽、核酸、核酸表达载体和/或重组宿主细胞，以及药学上可接受的载体。本公开的药物组合物可用于例如下文所述的本公开的方法中。除了本公开的多肽、核酸等之外，药物组合物还可包含(a)冻干保护剂；(b)表面活性剂；(c)增积剂；(d)张力调节剂；(e)稳定剂；(f)防腐剂和/或(g)缓冲剂。

[0145] 在一些方面中，药物组合物中的缓冲剂为Tris缓冲剂、组氨酸缓冲剂、磷酸盐缓冲剂、柠檬酸盐缓冲剂或乙酸盐缓冲剂。药物组合物也可包含冻干保护剂，例如蔗糖、山梨糖醇或海藻糖。在某些方面中，药物组合物包含防腐剂，例如苯扎氯铵、苄索氯铵、氯己定、苯酚、间甲酚、苯甲醇、对羟基苯甲酸甲酯、对羟基苯甲酸丙酯、氯丁醇、邻甲酚、对甲酚、氯甲酚、硝酸苯汞、硫柳汞、苯甲酸以及其各种混合物。在其它方面中，药物组合物包含增积剂，如甘氨酸。在其它方面中，药物组合物包含表面活性剂，例如聚山梨醇酯-20、聚山梨醇酯-40、聚山梨醇酯-60、聚山梨醇酯-65、聚山梨醇酯-80、聚山梨醇酯-85、泊洛沙姆-188、脱水山梨糖醇单月桂酸酯、脱水山梨糖醇单棕榈酸酯、脱水山梨糖醇单硬脂酸酯、脱水山梨糖醇单油酸酯、脱水山梨糖醇三月桂酸酯、脱水山梨糖醇三硬脂酸酯、脱水山梨糖醇三油酸酯或其组合。药物组合物还可包含张力调节剂，例如使制剂与人血液基本上等张或等渗的化合物。示例性张力调节剂包括蔗糖、山梨糖醇、甘氨酸、甲硫氨酸、甘露糖醇、右旋糖、肌醇、氯化钠、精氨酸和精氨酸盐酸盐。在其它方面中，药物组合物另外包含稳定剂，例如当与目标蛋白质组合时基本上防止或降低冻干形式或液体形式的目标蛋白质的化学和/或物理不稳定性的分子。示例性稳定剂包括蔗糖、山梨糖醇、甘氨酸、肌醇、氯化钠、甲硫氨酸、精氨酸和精氨酸盐酸盐。

[0146] 本公开的多肽、核酸等可为药物组合物中唯一的活性剂，或组合物还可包含一种或多种适用于预期用途的其它活性剂。

[0147] 本文所述的药物组合物通常包含本文所述的化合物与药学上可接受的载体、稀释剂或赋形剂的组合。此类组合物基本上不含非药学上可接受的组分，即含有低于提交本申请时美国监管要求所允许的量的非药学上可接受的组分。在此方面的一些方面中，如果化合物溶解或悬浮于水中，则组合物进一步任选地包含额外的药学上可接受的载体、稀释剂或赋形剂。在其它方面中，本文所述的药物组合物为固体药物组合物(例如片剂、胶囊等)。

[0148] 如美国监管机构所描述，本文所述的组合物还可作为膳食补充剂提供。

[0149] 这些组合物可以药物领域熟知的方式制备，并且可通过任何合适的途径施用。在一个优选方面中，药物组合物和制剂被设计用于经口施用。常规药物载体、水溶液、粉末或

油性基剂、增稠剂和其类似物可能是必要的或期望的。

[0150] 药物组合物可呈任何合适的形式，包括但不限于片剂、丸剂、粉末、锭剂、药囊、扁囊剂、酏剂、混悬剂、乳剂、溶液、糖浆、气雾剂(作为固体或在液体介质中)、含有例如至多10重量％活性化合物的软膏、软质和硬质明胶胶囊、无菌注射液和无菌封装粉末。

[0151] 3.本公开的方法

[0152] 在另一方面中，本公开提供用于治疗乳糜泻或非乳糜泻型麸质敏感(NCGS)的方法，所述方法包括向患有乳糜泻或NCGS的个体施用有效治疗乳糜泻或NCGS的量的一种或多种选自由本公开的多肽组成的组的多肽，或使用这些多肽中的一种或多种加工食物以供患有乳糜泻或NCGS的个体食用。

[0153] 在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约90％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约95％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约96％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约97％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约98％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约99％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含SEQ ID NO:1中所列的氨基酸序列的多肽。

[0154] 在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约90％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约95％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约96％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约97％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约98％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:8中所列的氨基酸序列具有至少约99％序列同一性的氨基酸序列的多肽。在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含SEQ ID NO:8中所列的氨基酸序列的多肽。

[0155] 在某些方面中，所述方法包括向患有乳糜泻或NCGS的受试者施用包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％、至少约99％或约100％序列同一性的氨基酸序列的多肽；其中多肽包含SEQ ID NO:8中所列的氨基酸序列；并且其中多肽在对应于SEQ ID NO:3中的氨基酸278的氨基酸残基处包含Ser、在对应于SEQ ID NO:3中的氨基酸78的氨基酸残基处包含Glu、以及在对应于SEQ ID NO:3中的氨基酸82的氨基酸残基处包含Asp。

[0156] 在某些方面中，本公开提供一种用于降解食品中的麸质的方法，所述方法包括使食品与有效降解麸质的量的上文所述的多肽接触，从而降解食品中的麸质。

[0157] 在某些方面中，本公开提供一种用于降解食品中的麸质的方法，所述方法包括使食品与有效降解麸质的量的上文所述的药物组合物接触，从而降解食品中的麸质。

[0158] 在某些方面中，本公开提供一种用于降解食品中的麦胶蛋白的方法，所述方法包括使食品与有效降解麦胶蛋白的量的上文所述的多肽或药物组合物接触，从而降解食品中的麸质。在一些方面中，所述方法降解食品中的至少约75％、至少约80％、至少约85％、至少约90％、至少约95％、至少约96％、至少约98％、至少约99％或约100％的麸质或麦胶蛋白。在一些方面中，本文公开的方法可在不到约1.5小时、不到约1小时、不到约45分钟、不到约40分钟、不到约30分钟、不到约25分钟、不到约20分钟、不到约15分钟、不到约10分钟或不到约5分钟降解食品中的麸质或麦胶蛋白。在一些方面中，本文公开的方法可在小于约6.5、小于约6.0、小于约5.5、小于约5.0、小于约4.5、小于约4.0、小于约3.5或小于约3.0的pH值下降解食品中的麸质或麦胶蛋白。

[0159] 本公开的发明者已发现，本公开的多肽能够降解麸质中称为‘麦胶蛋白’的富含脯氨酸(P)和谷氨酰胺(Q)的组分，据信这些组分负责大多数乳糜泻患者的大部分免疫反应。与Kuma010/011和公开为可用于治疗乳糜泻的其它多肽(WO2015/023728)相比，本公开的多肽在pH4下降解具有PQLP(SEQ ID NO:9)或PQQP(SEQ ID NO:10)基序(诸如PFPQPQLPY(SEQ ID NO:11)和/或PFPQPQQPF(SEQ ID NO:12))的肽时具有优异活性，其为代表麦胶蛋白的底物)。因此，本公开的多肽构成用于治疗乳糜泻和NCGS的显著改进的治疗剂。

[0160] 在某个方面中，本文公开的多肽的药物组合物和/或制剂经口施用。经口施用途径的非限制性实例包括使用片剂、丸剂、锭剂、酏剂、混悬剂、乳剂、溶液、糖浆或其任何组合。在某些方面中，包含本文公开的多肽的药物组合物在受试者摄取包含一种或多种麸质蛋白的物质，例如食物之前施用于受试者。在一些方面中，包含本文公开的多肽的药物组合物在受试者摄取包含一种或多种麸质蛋白的物质，例如食物的同时施用于受试者。在一些方面中，包含本文公开的多肽的药物组合物在受试者摄取包含一种或多种麸质蛋白的物质，例如食物之后施用于受试者。

[0161] 可调整剂量方案以提供最佳的所需反应(例如治疗或预防反应)。合适的剂量范围可为例如0.1ug/kg-100mg/kg体重；或者，其可为0.5ug/kg至50mg/kg；1ug/kg至25mg/kg，或5ug/kg至10mg/kg体重。多肽可以单次推注递送，或可如主治医师所确定施用超过一次(例如，2、3、4、5次或更多次)。

[0162] 本公开通过以下实施例进一步说明，所述实施例不应被解释为限制性的。所有引用的来源，例如在本文引用的参考文献、出版物、数据库、数据库条目和技术，均以引用的方式并入本申请，即使在引文中没有明确说明。如果引用的来源和本申请的陈述发生冲突，则

以本申请中的陈述为准。

[0163] 章节和表格标题并非旨在为限制性的。

[0164] 实施例

[0165] 实施例1:通过Kuma062-M降解全面包中的麸质

[0166] 本研究旨在证明Kuma062-M可有效降解麸质。

[0167] 胃消化的实验室模拟被设计成代表人类的胃消化。面包样品首先在人工唾液中捣碎以模拟咀嚼,随后通过添加盐酸酸化。除非另有说明,否则胃消化的pH值为3.6-4.5。混合样品以确保能够通过狭窄的移液器吸头抽出适当的材料代表(因为ELISA方法必须使用极小的体积);然而,在指明的情况下,样品仅被捣碎。膳食样品的最终总体积为400-800mL,随后将膳食的等分试样分配到个别管中以开始消化过程。通过添加胃蛋白酶和/或麦胶蛋白酶Kuma062-M开始消化。随后将样品在体温(37℃)下孵育指定的时间点。在大多数全麦面包/膳食消化实验中,使样品消化30分钟,因为食物在开始通过幽门瓣进入十二指肠之前在胃中搅动的平均滞留时间为30-60分钟。在消化期结束时,通过加热至使所有存在的酶不可逆地失活的温度来停止酶活性。

[0168] 按照制造商提供的说明书,通过R5 Ridascreen™ ELISA试剂盒(R-Biopharm)或G12 Gluentox®ELISA试剂盒(Biomedal)对消化样品中的麸质进行定量。这些试剂盒是基于单克隆抗体,R5(识别QQPFP)或G12(识别QPQLPY)(分别为SEQ ID NO:19和SEQ ID NO:20)。这些表位存在于麸质的大多数免疫原性片段中,包括所有免疫显性片段。G12抗体检测α-麦胶蛋白的免疫原性区域,而R5抗体检测ω-麦胶蛋白和γ-麦胶蛋白的免疫原性区域。虽然R5 ELISA方法已显示有效估计未加工食物的麸质浓度,但我们发现在麸质与胃蛋白酶一起孵育后,R5抗体识别的麸质部分会部分减少。胃蛋白酶已显示对G12抗体识别的部分,即33聚体片段LQLQPFPQPQLPYPQPQLPYPQPQLPYPQPQPF8(SEQ ID NO:13)的效果较差。与R5抗体不同,G12抗体对麸质表位的检测通常被观察到通过胃蛋白酶消化不受影响或甚至略有增加,表明用胃蛋白酶处理可使麸质的含有QPQLPY(SEQ ID NO:20)表位的区域更适用于G12抗体。在此实施例中,两种基于ELISA的方法均用于评估麦胶蛋白酶降低所有三个免疫原性麦胶蛋白家族:α-、ω-和γ-麦胶蛋白的量的能力。在下文详述的实施例之一中,使用基于G12的内部ELISA方法。这种内部开发的方法虽然比市售试剂盒便宜,但在定量低浓度麸质方面的可靠性较低。因此,这种方法仅用于评估样品之间的相对差异。

[0169] 表6显示Kuma062-M在模拟胃消化中可有效降解麸质。胃蛋白酶可在模拟胃消化中以低水平降解麸质。

[0170] 表6:Kuma062-M在刺激胃消化过程中对麸质的降解*

[0171]

酶	时间点	麸质 ppm 剩余	St Dev	降解%	St Dev%
胃蛋白酶	30	17920	640	4.55	3.41
Kuma062-M	5	200	14	98.93	0.07
Kuma062-M	30	48	4	99.75	0.02

[0172] 酶浓度:100μg/ml;面包混合物:16mg/ml;St Dev:标准偏差实施例2:Kuma062-M在不同pH值下降解全面包中的麸质本研究旨在评估Kuma062-M在不同pH值下降解麸质的能力。

[0173] 模拟胃消化的方案与实施例1中的方案基本上类似。产生具有以下pH水平的面包浆:3.9、4.5、5.0、5.5和5.9。pH 5.9是在用人工唾液捣碎后仅向面包浆中添加水而非HCl时面包浆的pH值。

[0174] 表7显示Kuma062-M在各种pH值下均可有效降解麸质。

[0175] 表7:Kuma062-M在不同pH值下对麸质的降解*

[0176]

		RS ELISA				G12 ELISA			
酶浓度	pH	平均ppm	标准Dev	麸质降解%	标准Dev	平均ppm	标准Dev	麸质降解的$	标准Dev
1000 ug/mL	3.9	5.9	0.8	99.93%	0.01%	13.8	1.2	99.84%	0.01%
	4.5	11.2	4.7	99.88%	0.05%	19.2	6.5	99.80%	0.08%
	5.0	16.2	2.6	99.85%	0.02%	26.2	7.5	99.75%	0.09%
	5.5	15.5	3.3	99.86%	0.03%	30.5	8.2	99.72%	0.10%
	5.9	24.1	6.2	99.80%	0.05%	56.0	4.1	99.54%	0.03%
400 ug/mL	3.9	11.7	3.8	99.86%	0.04%	20.3	4.3	99.76%	0.05%
	4.5	11.9	0.3	99.88%	0.00%	19.4	3.1	99.80%	0.04%
	5.0	13.3	3.9	99.88%	0.04%	19.6	2.3	99.82%	0.03%
	5.5	9.9	1.4	99.91%	0.01%	22.4	2.4	99.80%	0.03%
	5.9	15.9	1.2	99.87%	0.01%	29.5	4.1	99.75%	0.03%
200 ug/mL	3.9	9.6	1.8	99.89%	0.02%	19.0	2.0	99.78%	0.02%
	4.5	12.4	4.5	99.87%	0.05%	22.8	3.2	99.76%	0.04%
	5.0	5.5	0.8	99.95%	0.01%	25.0	6.7	99.77%	0.08%
	5.5	11.7	1.9	99.89%	0.02%	24.9	2.6	99.77%	0.03%
	5.9	15.3	1.5	99.87%	0.01%	37.5	5.0	99.69%	0.04%

[0177] 麸质浓度:10mg/ml

[0178] 实施例3:通过Kuma062-M降解快餐中的麸质

[0179] 本研究旨在评估Kuma062M是否能够在存在其它膳食蛋白质的情况下保持显著的抗麸质活性。

[0180] 模拟胃消化的方案与实施例1中的方案基本上类似。估计香草奶昔(粗略地,通过与McDonalds®的类似大小的奶昔进行比较)含有10千克蛋白质,而汉堡肉饼估计含有7

千克蛋白质。胃消化中膳食的pH值为4.0-4.5。将对照餐中的汉堡包的量调整为与汉堡的小圆面包和奶昔餐中的量相同。汉堡包和奶昔的胃消化体积为500mL;对照餐也调整至500mL。将捣碎和混合后的膳食浆液的等分试样分装至较小的试管中,并且将麸质酶和胃蛋白酶添加至这些等分试样中。Kuma062-M的酶浓度为700μg/mL或70μg/mL。膳食消化30分钟或5分钟。本研究还包括黑曲霉来源的脯胺酰内切蛋白酶(AN-PEP)和EPB2/SCPEP。

[0181] 表8和9表明,Kuma062-M可在其它膳食蛋白质存在的情况下有效降解麸质。表8显示使用G12 ELISA测定的结果。表9显示使用R5 ELISA测定的结果。

[0182] 表8:在快餐G12 ELISA测定中通过Kuma062-M降解麸质

[0183]

酶	μg/ml	时间点	膳食	剩余麸质(ppm)	St Dev	降解%	St Dev%	剩余当量mg	mg St Dev
胃蛋白酶	700	30	仅小圆面包	13380	1004	8.03	6.41	6690	502
胃蛋白酶	700	30	汉堡包	8056	464	55.61	2.54	4028	232
ANPEP	700	30	仅小圆面包	434	23	97.06	0.15	217	12
ANPEP	700	30	汉堡包	4261	263	77.23	1.44	2131	132

[0184]

EP/SC	700	30	仅小圆面包	2394	97	85.08	0.62	1197	48
EP/SC	700	30	汉堡包	9401	940	54.47	5.15	4701	470
Kuma062	700	5	仅小圆面包	69	4	99.53	0.03	35	2
Kuma062	700	30	仅小圆面包	30	2	99.82	0.02	15	1
Kuma062	700	5	汉堡包	83	7	99.59	0.04	42	4
Kuma062	700	30	汉堡包	54	4	99.68	0.02	27	2
Kuma062	70	30	仅小圆面包	56	3	99.62	0.02	28	2
Kuma062	70	30	汉堡包	151	6	99.14	0.03	75	3

[0185] 表9:在快餐R5 ELISA测定中通过Kuma062-M消化麸质

[0186]

酶	μg/ml	时间点	膳食	剩余麸质(ppm)	St Dev	降解%	St Dev%	剩余当量mg	mg St Dev
胃蛋白酶	700	30	仅小圆面包	9680	554	38.27	3.53	4840	277
胃蛋白酶	700	30	汉堡包	9493	1398	48.03	7.65	4747	699
EP/SC	700	30	仅小圆面包	747	92	95.24	0.59	373	46
EP/SC	700	30	汉堡包	10400	604	43.07	3.31	5200	302
Kuma062	700	5	仅小圆	23	3	99.86	0.02	11	2

[0187]

			面包						
Kuma062	700	30	仅小圆面包	9	2	99.94	0.01	4	1
Kuma062	700	5	汉堡包	113	17	99.38	0.09	56	9
Kuma062	700	30	汉堡包	35	13	99.81	0.07	18	6
Kuma062	70	30	仅小圆面包	23	1	99.86	0.01	11	1
Kuma062	70	30	汉堡包	147	6	99.20	0.03	73	3

[0188] 等效物

[0189] 本领域技术人员将认识到或能够仅使用不超过一种常规实验来确定本公开的特定方面的许多等效物。此类等效物旨在由以下权利要求书涵盖。

[0190] 方面

[0191] E1.一种多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列。

[0192] E2.如E1所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少85%序列同一性的氨基酸序列。

[0193] E3.如E1或E2所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少90%序列同一性的氨基酸序列。

[0194] E4.如E1至E3中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少95%序列同一性的氨基酸序列。

[0195] E5.如E1至E4中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少99%序列同一性的氨基酸序列。

[0196] E6.如E1至5中任一项所述的多肽,所述多肽包含SEQ ID NO:1中所列的氨基酸序列。

[0197] E7.如E1至E6中任一项所述的多肽,其中对应于SEQ ID NO:6的氨基酸467的氨基酸残基为Ser。

[0198] E8.如E1至E7中任一项所述的多肽,其中对应于SEQ ID NO:6的氨基酸267的氨基酸残基为Glu。

[0199] E9.如E1至E8中任一项所述的多肽,其中对应于SEQ ID NO:6的氨基酸271的氨基酸残基为Asp。

[0200] E10.如E1至E9中任一项所述的多肽,所述多肽能够裂解麦胶蛋白。

[0201] E11.一种多肽,所述多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约

98%、至少约99%或约100%序列同一性的氨基酸序列。

[0202] E12.如E11所述的多肽,所述多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少85%序列同一性的氨基酸序列。

[0203] E13.如E11或E12所述的多肽,所述多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少90%序列同一性的氨基酸序列。

[0204] E14.如E11至E13中任一项所述的多肽,所述多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少95%序列同一性的氨基酸序列。

[0205] E15.如E11至14中任一项所述的多肽,所述多肽包含与SEQ ID NO:8中所列的氨基酸序列具有至少99%序列同一性的氨基酸序列。

[0206] E16.如E11至E15中任一项所述的多肽,所述多肽包含SEQ ID NO:8中所列的氨基酸序列。

[0207] E17.如E11至E16中任一项所述的多肽,其中对应于SEQ ID NO:3的氨基酸278的氨基酸残基为Ser。

[0208] E18.如E11至E17中任一项所述的多肽,其中对应于SEQ ID NO:3的氨基酸78的氨基酸残基为Glu。

[0209] E19.如E11至E18中任一项所述的多肽,其中对应于SEQ ID NO:3的氨基酸82的氨基酸残基为Asp。

[0210] E20.如E11至E19中任一项所述的多肽,所述多肽能够裂解麦胶蛋白。

[0211] E21.一种多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%序列同一性的氨基酸序列;其中多肽包含SEQ ID NO:8中所列的氨基酸序列。

[0212] E22.如E21所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少85%序列同一性的氨基酸序列。

[0213] E23.如E21或E22所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少90%序列同一性的氨基酸序列。

[0214] E24.如E21至E23中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少95%序列同一性的氨基酸序列。

[0215] E25.如E 21至24中任一项所述的多肽,所述多肽包含与SEQ ID NO:1中所列的氨基酸序列具有至少99%序列同一性的氨基酸序列。

[0216] E26.如E21至E25中任一项所述的多肽,所述多肽包含SEQ ID NO:1中所列的氨基酸序列。

[0217] 27.如E21至E26中任一项所述的多肽,其中对应于SEQ ID NO:6的氨基酸467的氨基酸残基为Ser。

[0218] E28.如E21至E27中任一项所述的多肽,其中对应于SEQ ID NO:6的氨基酸267的氨基酸残基为Glu。

[0219] E29.如E21至E28中任一项所述的多肽,其中对应于SEQ ID NO:6的氨基酸271的氨基酸残基为Asp。

[0220] E30.如E21至E29中任一项所述的多肽,所述多肽能够裂解麦胶蛋白。

[0221] E31.如E1至E30中任一项所述的多肽，所述多肽还包含组氨酸标签，其中组氨酸标签在多肽的C末端处融合。

[0222] E32.如E所述的多肽，其中组氨酸标签包含SEQ ID NO:17中所列的氨基酸序列(GSTENLYFQSGALEHHHHHH)。

[0223] E33.如E32或E33所述的多肽，其中组氨酸标签包含可裂解组氨酸标签，包括但不限于包含SEQ ID NO:15中所列的氨基酸序列的可裂解组氨酸标签(X_NPQ(L/Q)PX_NHHHHHH)，其中X_N为介于1-25个之间氨基酸残基的接头。

[0224] E34.如E31至E33中任一项所述的多肽，其中可裂解组氨酸标签包含SEQ ID NO:16中所列的氨基酸序列(GSSGSSGSQPQLPYGSSGSSGSHHHHHH)。

[0225] E35.一种核酸分子，所述核酸分子编码如E1至E34中任一项所述的多肽。

[0226] E36.一种核酸表达载体，所述核酸表达载体包含如E35所述的核酸分子。

[0227] E37.一种重组宿主细胞，所述重组宿主细胞包含如E35所述的核酸分子或如E36所述的核酸表达载体。

[0228] E38.一种药物组合物，所述药物组合物包含如E1至E34中任一项所述的多肽、如E35所述的核酸分子、如E36所述的核酸表达载体、如E37所述的重组宿主细胞或其任何组合以及药学上可接受的载体。

[0229] E39.一种治疗乳糜泻或非乳糜泻型麸质敏感(NCGS)的方法，所述方法包括向患有乳糜泻或NCGS的个体施用有效治疗乳糜泻或NCGS的量的如E1至E34中任一项所述的多肽、如权利要求35所述的核酸分子、如权利要求36所述的核酸表达载体、如权利要求37所述的重组宿主细胞或如权利要求38所述的药物组合物。

[0230] E40.如E39所述的方法，其中多肽、核酸分子、核酸表达载体、重组宿主细胞或药物组合物经口施用。

序列表

<110> 华盛顿大学和加州大学董事会

<120> 用于治疗乳糜泻的组合物和方法

<130> 20-1718-WO

<150> US 63/108,163

<151> 2020-10-30

<160> 23

<170> PatentIn version 3.5

<210> 1

<211> 552

<212> PRT

<213> 人工序列

<220>

<223> Kuma062-M

<400> 1

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

180 185 190

Ala Tyr Thr Pro Leu Asp Val Ala Gln Ala Tyr Gln Phe Pro Glu Gly

195 200 205

Leu Asp Gly Gln Gly Gln Cys Ile Ala Ile Ile Glu Leu Gly Gly Gly

210 215 220

Tyr Asp Glu Ala Ser Leu Ala Gln Tyr Phe Ala Ser Leu Gly Val Pro

225 230 235 240

Ala Pro Gln Val Val Ser Val Ser Val Asp Gly Ala Ser Asn Gln Pro

245 250 255

Thr Gly Asp Pro Glu Gly Pro Asp Gly Glu Val Thr Leu Asp Ile Glu

260 265 270

Val Ala Gly Ala Leu Ala Pro Gly Ala Lys Phe Ala Val Tyr Phe Ala

275 280 285

Pro Asp Thr Thr Ala Gly Phe Leu Asp Ala Ile Thr Thr Ala Ile His

290 295 300

Asp Pro Thr Leu Lys Pro Ser Val Val Ser Ile Ser Trp Ser Met Pro

305 310 315 320

Glu Asp Ser Trp Thr Ser Ala Ala Ile Ala Ala Met Asn Arg Ala Phe

325 330 335

Leu Asp Ala Ala Ala Leu Gly Val Thr Val Leu Ala Ala Ala Gly Asp

340 345 350

Gln Gly Ser Thr Ser Gly Glu Gln Asp Gly Leu Tyr His Val His Phe

355 360 365

Pro Ala Ala Ser Pro Tyr Val Leu Ala Cys Gly Gly Thr Arg Leu Val

370 375 380

Ala Ser Gly Gly Arg Ile Ala Gln Glu Thr Val Trp Asn Gln Gly Pro

385 390 395 400

Asp Gly Gly Ala Thr Gly Gly Gly Val Ser Arg Ile Phe Pro Leu Pro

405 410 415

Ala Trp Gln Glu His Ala Asn Val Pro Pro Ser Ala Asn Pro Gly Ala

420 425 430

Ser Ser Gly Arg Gly Val Pro Asp Leu Ala Gly Asn Ala Asp Pro Gln

435 440 445

Thr Gly Tyr Glu Val Val Ile Asp Gly Glu Ala Thr Val Thr Gly Gly

450 455 460

Thr Ser Ala Val Ala Pro Leu Phe Ala Ala Leu Val Ala Arg Ile Asn

465 470 475 480

Gln Lys Leu Gly Lys Ala Val Gly Tyr Leu Asn Pro Thr Leu Tyr Gln

485 490 495

Leu Pro Ala Asp Val Phe His Asp Ile Thr Glu Gly Asn Asn Asp Ile
500 505 510
Ala Asn Arg Ala Gln Ile Tyr Gln Ala Gly Pro Gly Trp Asp Pro Cys
515 520 525
Thr Gly Leu Gly Ser Pro Ile Gly Val Arg Leu Leu Gln Ala Leu Leu
530 535 540
Pro Ser Ala Ser Gln Pro Gln Pro
545 550

<210> 2
<211> 189
<212> PRT
<213> 人工序列
<220>
<223> Kuma011前蛋白质结构域
<400> 2

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

<210> 3
<211> 364
<212> PRT
<213> 人工序列
<220>
<223> Kuma011成熟肽
<400> 3

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

Ala Asp Pro Ala Thr Gly Tyr Glu Val Val Ile Asp Gly Glu Ala Thr
260 265 270
Val Ile Gly Gly Thr Ser Ala Val Ala Pro Leu Phe Ala Ala Leu Val
275 280 285
Ala Arg Ile Asn Gln Lys Leu Gly Lys Ala Val Gly Tyr Leu Asn Pro
290 295 300
Thr Leu Tyr Gln Leu Pro Ala Asp Val Phe His Asp Ile Thr Glu Gly
305 310 315 320
Asn Asn Asp Ile Ala Asn Arg Ala Gln Ile Tyr Gln Ala Gly Pro Gly
325 330 335
Trp Asp Pro Cys Thr Gly Leu Gly Ser Pro Ile Gly Val Arg Leu Leu
340 345 350
Gln Ala Leu Leu Pro Ser Ala Ser Gln Pro Gln Pro
355 360

<210> 4
<211> 573
<212> PRT
<213> 人工序列
<220>
<223> Kuma010 (全长)
<400> 4

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

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

450 455 460

Gly Thr Ser Ala Val Ala Pro Leu Phe Ala Ala Leu Val Ala Arg Ile
465 470 475 480

Asn Gln Lys Leu Gly Lys Ala Val Gly Tyr Leu Asn Pro Thr Leu Tyr
485 490 495

Gln Leu Pro Ala Asp Val Phe His Asp Ile Thr Glu Gly Asn Asn Asp
500 505 510

Ile Ala Asn Arg Ala Gln Ile Tyr Gln Ala Gly Pro Gly Trp Asp Pro
515 520 525

Cys Thr Gly Leu Gly Ser Pro Ile Gly Val Arg Leu Leu Gln Ala Leu
530 535 540

Leu Pro Ser Ala Ser Gln Pro Gln Pro Gly Ser Thr Glu Asn Leu Tyr
545 550 555 560

Phe Gln Ser Gly Ala Leu Glu His His His His His His
565 570

<210> 5

<211> 384

<212> PRT

<213> 人工序列

<220>

<223> Kuma010成熟肽

<400> 5

Ala Ala Pro Thr Ala Tyr Thr Pro Leu Asp Val Ala Gln Ala Tyr Gln
1 5 10 15

Phe Pro Glu Gly Leu Asp Gly Gln Gly Gln Cys Ile Ala Ile Ile Glu
20 25 30

Leu Gly Gly Gly Tyr Asp Glu Ala Ser Leu Ala Gln Tyr Phe Ala Ser
35 40 45

Leu Gly Val Pro Ala Pro Gln Val Val Ser Val Ser Val Asp Gly Ala
50 55 60

Ser Asn Gln Pro Thr Gly Asp Pro Lys Gly Pro Asp Gly Glu Val Glu
65 70 75 80

Leu Asp Ile Glu Val Ala Gly Ala Leu Ala Pro Gly Ala Lys Phe Ala
85 90 95

Val Tyr Phe Ala Pro Asp Thr Thr Ala Gly Phe Leu Asp Ala Ile Thr
100 105 110

Thr Ala Ile His Asp Pro Thr Leu Lys Pro Ser Val Val Ser Ile Ser
115 120 125

Trp Ser Gly Pro Glu Asp Ser Trp Thr Ser Ala Ala Ile Ala Ala Met

130 135 140

Asn Arg Ala Phe Leu Asp Ala Ala Ala Leu Gly Val Thr Val Leu Ala
145 150 155 160
Ala Ala Gly Asp Ser Gly Ser Thr Gly Gly Glu Gln Asp Gly Leu Tyr
165 170 175
His Val His Phe Pro Ala Ala Ser Pro Tyr Val Leu Ala Cys Gly Gly
180 185 190
Thr Arg Leu Val Ala Ser Gly Gly Arg Ile Ala Gln Glu Thr Val Trp
195 200 205
Asn Asp Gly Pro Asp Gly Gly Ala Thr Gly Gly Gly Val Ser Arg Ile
210 215 220
Phe Pro Leu Pro Ala Trp Gln Glu His Ala Asn Val Pro Pro Ser Ala
225 230 235 240
Asn Pro Gly Ala Ser Ser Gly Arg Gly Val Pro Asp Leu Ala Gly Asn
245 250 255
Ala Asp Pro Ala Thr Gly Tyr Glu Val Val Ile Asp Gly Glu Ala Thr
260 265 270
Val Ile Gly Gly Thr Ser Ala Val Ala Pro Leu Phe Ala Ala Leu Val
275 280 285
Ala Arg Ile Asn Gln Lys Leu Gly Lys Ala Val Gly Tyr Leu Asn Pro
290 295 300
Thr Leu Tyr Gln Leu Pro Ala Asp Val Phe His Asp Ile Thr Glu Gly
305 310 315 320
Asn Asn Asp Ile Ala Asn Arg Ala Gln Ile Tyr Gln Ala Gly Pro Gly
325 330 335
Trp Asp Pro Cys Thr Gly Leu Gly Ser Pro Ile Gly Val Arg Leu Leu
340 345 350
Gln Ala Leu Leu Pro Ser Ala Ser Gln Pro Gln Pro Gly Ser Thr Glu
355 360 365
Asn Leu Tyr Phe Gln Ser Gly Ala Leu Glu His His His His His His
370 375 380

<210> 6
<211> 553
<212> PRT
<213> 人工序列
<220>
<223> Kuma011
(全长)
<400> 6

Met Ser Asp Met Glu Lys Pro Trp Lys Glu Gly Glu Glu Ala Arg Ala
1 5 10 15
Val Leu Gln Gly His Ala Arg Ala Gln Ala Pro Gln Ala Val Asp Lys
20 25 30
Gly Pro Val Ala Gly Asp Glu Arg Met Ala Val Thr Val Val Leu Arg
35 40 45
Arg Gln Arg Ala Gly Glu Leu Ala Ala His Val Glu Arg Gln Ala Ala
50 55 60
Ile Ala Pro His Ala Arg Glu His Leu Lys Arg Glu Ala Phe Ala Ala
65 70 75 80
Ser His Gly Ala Ser Leu Asp Asp Phe Ala Glu Leu Arg Arg Phe Ala
85 90 95
Asp Ala His Gly Leu Ala Leu Asp Arg Ala Asn Val Ala Ala Gly Thr
100 105 110
Ala Val Leu Ser Gly Pro Asp Asp Ala Ile Asn Arg Ala Phe Gly Val
115 120 125
Glu Leu Arg His Phe Asp His Pro Asp Gly Ser Tyr Arg Ser Tyr Leu
130 135 140
Gly Glu Val Thr Val Pro Ala Ser Ile Ala Pro Met Ile Glu Ala Val
145 150 155 160
Leu Gly Leu Asp Thr Arg Pro Val Ala Arg Pro His Phe Arg Met Gln
165 170 175
Arg Arg Ala Glu Gly Gly Phe Glu Ala Arg Ser Gln Ala Ala Ala Pro
180 185 190
Thr Ala Tyr Thr Pro Leu Asp Val Ala Gln Ala Tyr Gln Phe Pro Glu
195 200 205
Gly Leu Asp Gly Gln Gly Gln Cys Ile Ala Ile Ile Glu Leu Gly Gly
210 215 220
Gly Tyr Asp Glu Ala Ser Leu Ala Gln Tyr Phe Ala Ser Leu Gly Val
225 230 235 240
Pro Ala Pro Gln Val Val Ser Val Ser Val Asp Gly Ala Ser Asn Gln
245 250 255
Pro Thr Gly Asp Pro Lys Gly Pro Asp Gly Glu Val Glu Leu Asp Ile
260 265 270
Glu Val Ala Gly Ala Leu Ala Pro Gly Ala Lys Phe Ala Val Tyr Phe
275 280 285
Ala Pro Asp Thr Thr Ala Gly Phe Leu Asp Ala Ile Thr Thr Ala Ile
290 295 300
His Asp Pro Thr Leu Lys Pro Ser Val Val Ser Ile Ser Trp Ser Gly

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

<210> 7
<211> 188
<212> PRT
<213> 人工序列
<220>
<223> Kuma062-M前蛋白质结构域
<400> 7
Ser Asp Met Glu Lys Pro Trp Lys Glu Gly Glu Glu Ala Arg Ala Val

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

<210> 8
<211> 363
<212> PRT
<213> 人工序列
<220>
<223> Kuma062-M成熟肽
<400> 8

Ala Ala Pro Thr Ala Tyr Thr Pro Leu Asp Val Ala Gln Ala Tyr Gln
1 5 10 15
Phe Pro Glu Gly Leu Asp Gly Gln Gly Gln Cys Ile Ala Ile Ile Glu
20 25 30
Leu Gly Gly Gly Tyr Asp Glu Ala Ser Leu Ala Gln Tyr Phe Ala Ser
35 40 45
Leu Gly Val Pro Ala Pro Gln Val Val Ser Val Ser Val Asp Gly Ala
50 55 60
Ser Asn Gln Pro Thr Gly Asp Pro Glu Gly Pro Asp Gly Glu Val Thr

65 70 75 80
Leu Asp Ile Glu Val Ala Gly Ala Leu Ala Pro Gly Ala Lys Phe Ala
85 90 95
Val Tyr Phe Ala Pro Asp Thr Thr Ala Gly Phe Leu Asp Ala Ile Thr
100 105 110
Thr Ala Ile His Asp Pro Thr Leu Lys Pro Ser Val Val Ser Ile Ser
115 120 125
Trp Ser Met Pro Glu Asp Ser Trp Thr Ser Ala Ala Ile Ala Ala Met
130 135 140
Asn Arg Ala Phe Leu Asp Ala Ala Ala Leu Gly Val Thr Val Leu Ala
145 150 155 160
Ala Ala Gly Asp Gln Gly Ser Thr Ser Gly Glu Gln Asp Gly Leu Tyr
165 170 175
His Val His Phe Pro Ala Ala Ser Pro Tyr Val Leu Ala Cys Gly Gly
180 185 190
Thr Arg Leu Val Ala Ser Gly Gly Arg Ile Ala Gln Glu Thr Val Trp
195 200 205
Asn Gln Gly Pro Asp Gly Gly Ala Thr Gly Gly Gly Val Ser Arg Ile
210 215 220
Phe Pro Leu Pro Ala Trp Gln Glu His Ala Asn Val Pro Pro Ser Ala
225 230 235 240
Asn Pro Gly Ala Ser Ser Gly Arg Gly Val Pro Asp Leu Ala Gly Asn
245 250 255
Ala Asp Pro Gln Thr Gly Tyr Glu Val Val Ile Asp Gly Glu Ala Thr
260 265 270
Val Thr Gly Gly Thr Ser Ala Val Ala Pro Leu Phe Ala Ala Leu Val
275 280 285
Ala Arg Ile Asn Gln Lys Leu Gly Lys Ala Val Gly Tyr Leu Asn Pro
290 295 300
Thr Leu Tyr Gln Leu Pro Ala Asp Val Phe His Asp Ile Thr Glu Gly
305 310 315 320
Asn Asn Asp Ile Ala Asn Arg Ala Gln Ile Tyr Gln Ala Gly Pro Gly
325 330 335
Trp Asp Pro Cys Thr Gly Leu Gly Ser Pro Ile Gly Val Arg Leu Leu
340 345 350
Gln Ala Leu Leu Pro Ser Ala Ser Gln Pro Gln
355 360

<210> 9
<211> 4

<212> PRT
<213> 人工序列
<220>
<223> 麦胶蛋白
<400> 9
Pro Gln Leu Pro
1
<210> 10
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 麦胶蛋白
<400> 10
Pro Gln Gln Pro
1
<210> 11
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 麦胶蛋白
<400> 11
Pro Phe Pro Gln Pro Gln Leu Pro Tyr
1 5
<210> 12
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> 麦胶蛋白
<400> 12
Pro Phe Pro Gln Pro Gln Gln Pro Phe
1 5
<210> 13
<211> 33
<212> PRT
<213> 人工序列
<220>

<223> 多肽
<400> 13
Leu Gln Leu Gln Pro Phe Pro Gln Pro Gln Leu Pro Tyr Pro Gln Pro
1 5 10 15
Gln Leu Pro Tyr Pro Gln Pro Gln Leu Pro Tyr Pro Gln Pro Gln Pro
20 25 30
Phe
<210> 14
<211> 26
<212> PRT
<213> 人工序列
<220>
<223> 多肽
<400> 14
Phe Leu Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Tyr
1 5 10 15
Pro Gln Gln Pro Gln Gln Pro Phe Pro Gln
20 25
<210> 15
<211> 60
<212> PRT
<213> 人工序列
<220>
<223> 可裂解组氨酸标签
<220>
<221> misc_feature
<222> (1)..(25)
<223> Xaa为任何氨基酸
<220>
<221> misc_feature
<222> (28)..(28)
<223> Xaa为Lys或Gln
<220>
<221> misc_feature
<222> (30)..(55)
<223> Xaa为任何氨基酸
<400> 15
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Gln Xaa Pro Xaa Xaa Xaa
20 25 30
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45
Xaa Xaa Xaa Xaa Xaa Xaa His His His His His His
50 55 60

<210> 16
<211> 28
<212> PRT
<213> 人工序列
<220>
<223> 可裂解组氨酸标签
<400> 16

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

<210> 17
<211> 20
<212> PRT
<213> 人工序列
<220>
<223> 组氨酸标签
<400> 17

Gly Ser Thr Glu Asn Leu Tyr Phe Gln Ser Gly Ala Leu Glu His His
1 5 10 15
His His His His
20

<210> 18
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> TEV蛋白酶切割位点
<400> 18

Glu Asn Leu Tyr Phe Gln Ser
1 5

<210> 19
<211> 5

<212> PRT
<213> 人工序列
<220>
<223> R5
<400> 19
Gln Gln Pro Phe Pro
1 5
<210> 20
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> G12
<400> 20
Gln Pro Gln Leu Pro Tyr
1 5
<210> 21
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 麦胶酶
<400> 21
Ser Asp Met Glu
1
<210> 22
<211> 1659
<212> DNA
<213> 人工序列
<220>
<223> 编码Kuma062-M蛋白的核酸
<400> 22
agtgatatgg aaaaaccgtg gaaagaaggt gaagaagccc gcgcagtgct gcaaggtcat 60
gctcgtgcgc aggcaccgca agcagtcgat aaaggcccgg tggcaggtga cgaacgcatg 120
gctgttaccg tggttctgcg tcgccagcgt gcaggtgaac tggcggccca cgtggaacgt 180
caagcagcta ttgctccgca tgcgcgcgaa cacctgaaac gtgaagcgtt tgcggccagt 240
catggtgcgt ccctggatga ctttgccgaa ctgcgtcgct tcgcagatgc tcacggcctg 300
gcgctggacc gtgcaaacgt tgcagctggc accgccgttc tgtctggtcc ggacgatgca 360
atcaatcgcg cttttggtgt ggaactgcgt catttcgatc acccggacgg ctcatatcgt 420

tcgtacctgg gtgaagtcac cgtgccggcc agtattgcac cgatgatcga agcggttctg 480
ggcctggata cgcgtccggt cgcccgccgt cgttttcgta tgcagcgtcg cgcagaaggc 540
ggtttcgaag ctcgttccca agcggcggca ccgaccgcat atacgccgct ggatgttgcg 600
caggcctacc aatttccgga aggtctggac ggccagggtc aatgcattgc cattatcgaa 660
ctgggcggtg gctatgatga agcttcactg gcgcagtact tcgcgtcgct gggcgtgccg 720
gcaccgcaag tcgtgagtgt ttccgtcgat ggtgcgagca accagccgac cggtgatccg 780
gaaggtccgg acggtgaagt gaccctggat atcgaagttg caggcgctct ggcgccgggt 840
gccaaatttg cagtgtattt cgcgccggat accactgccg gttttctgga cgcgattacc 900
acggccatcc acgatccgac gctgaaaccg agcgttgtct caatttcgtg gagcatgccg 960
gaagacagct ggacctctgc tgcgatcgcc gcaatgaacc gtgcgtttct ggatgctgcg 1020
gccctgggtg tgaccgttct ggcagctgcg ggcgaccagg gttctacgag cggcgaacag 1080
gacggtctgt atcatgtgca tttcccggcc gcatcaccgt acgttctggc gtgcggtggc 1140
acgcgcctgg tcgcatcggg tggccgtatt gcgcaggaaa ccgtctggaa ccagggtccg 1200
gacggtggtg caacgggtgg cggtgtgagc cgcatcttcc cgctgccggc atggcaggaa 1260
cacgctaacg ttccgccgtc tgcaaatccg ggcgcgagca gcggccgtgg tgtcccggat 1320
ctggctggta atgcggaccc gcagaccggt tatgaagtgg ttattgatgg cgaagcaacc 1380
gtcaccggcg gtacgagcgc cgtggcaccg ctgtttgctg cgctggttgc gcgtattaac 1440
cagaaactgg gcaaagcagt tggttatctg aatccgaccc tgtaccaact gccggcagat 1500
gttttccatg acatcacgga gggtaacaat gatattgcaa accgtgcgca gatttatcaa 1560
gcaggtccgg gctgggaccc gtgtaccggt ctgggttcac cgattggtgt gcgtctgctg 1620
caagcactgt tgccgagtgc ctcccagccg caaccgtga 1659

<210> 23
<211> 553
<212> PRT
<213> 人工序列
<220>
<223> 多肽
<220>
<221> misc_feature
<222> (1)..(1)
<223> 其中Xaa不为Met。
<400> 23

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

50 55 60

Ile Ala Pro His Ala Arg Glu His Leu Lys Arg Glu Ala Phe Ala Ala
65 70 75 80

Ser His Gly Ala Ser Leu Asp Asp Phe Ala Glu Leu Arg Arg Phe Ala
85 90 95

Asp Ala His Gly Leu Ala Leu Asp Arg Ala Asn Val Ala Ala Gly Thr
100 105 110

Ala Val Leu Ser Gly Pro Asp Asp Ala Ile Asn Arg Ala Phe Gly Val
115 120 125

Glu Leu Arg His Phe Asp His Pro Asp Gly Ser Tyr Arg Ser Tyr Leu
130 135 140

Gly Glu Val Thr Val Pro Ala Ser Ile Ala Pro Met Ile Glu Ala Val
145 150 155 160

Leu Gly Leu Asp Thr Arg Pro Val Ala Arg Arg Arg Phe Arg Met Gln
165 170 175

Arg Arg Ala Glu Gly Gly Phe Glu Ala Arg Ser Gln Ala Ala Ala Pro
180 185 190

Thr Ala Tyr Thr Pro Leu Asp Val Ala Gln Ala Tyr Gln Phe Pro Glu
195 200 205

Gly Leu Asp Gly Gln Gly Gln Cys Ile Ala Ile Ile Glu Leu Gly Gly
210 215 220

Gly Tyr Asp Glu Ala Ser Leu Ala Gln Tyr Phe Ala Ser Leu Gly Val
225 230 235 240

Pro Ala Pro Gln Val Val Ser Val Ser Val Asp Gly Ala Ser Asn Gln
245 250 255

Pro Thr Gly Asp Pro Glu Gly Pro Asp Gly Glu Val Thr Leu Asp Ile
260 265 270

Glu Val Ala Gly Ala Leu Ala Pro Gly Ala Lys Phe Ala Val Tyr Phe
275 280 285

Ala Pro Asp Thr Thr Ala Gly Phe Leu Asp Ala Ile Thr Thr Ala Ile
290 295 300

His Asp Pro Thr Leu Lys Pro Ser Val Val Ser Ile Ser Trp Ser Met
305 310 315 320

Pro Glu Asp Ser Trp Thr Ser Ala Ala Ile Ala Ala Met Asn Arg Ala
325 330 335

Phe Leu Asp Ala Ala Ala Leu Gly Val Thr Val Leu Ala Ala Ala Gly
340 345 350

Asp Gln Gly Ser Thr Ser Gly Glu Gln Asp Gly Leu Tyr His Val His
355 360 365

Phe Pro Ala Ala Ser Pro Tyr Val Leu Ala Cys Gly Gly Thr Arg Leu
370 375 380
Val Ala Ser Gly Gly Arg Ile Ala Gln Glu Thr Val Trp Asn Gln Gly
385 390 395 400
Pro Asp Gly Gly Ala Thr Gly Gly Gly Val Ser Arg Ile Phe Pro Leu
405 410 415
Pro Ala Trp Gln Glu His Ala Asn Val Pro Pro Ser Ala Asn Pro Gly
420 425 430
Ala Ser Ser Gly Arg Gly Val Pro Asp Leu Ala Gly Asn Ala Asp Pro
435 440 445
Gln Thr Gly Tyr Glu Val Val Ile Asp Gly Glu Ala Thr Val Thr Gly
450 455 460
Gly Thr Ser Ala Val Ala Pro Leu Phe Ala Ala Leu Val Ala Arg Ile
465 470 475 480
Asn Gln Lys Leu Gly Lys Ala Val Gly Tyr Leu Asn Pro Thr Leu Tyr
485 490 495
Gln Leu Pro Ala Asp Val Phe His Asp Ile Thr Glu Gly Asn Asn Asp
500 505 510
Ile Ala Asn Arg Ala Gln Ile Tyr Gln Ala Gly Pro Gly Trp Asp Pro
515 520 525
Cys Thr Gly Leu Gly Ser Pro Ile Gly Val Arg Leu Leu Gln Ala Leu
530 535 540
Leu Pro Ser Ala Ser Gln Pro Gln Pro
545 550