



(12) 发明专利申请

(10) 申请公布号 CN 105247067 A

(43) 申请公布日 2016. 01. 13

(21) 申请号 201480029964. 7

(51) Int. Cl.

(22) 申请日 2014. 05. 23

C12P 21/06(2006. 01)

C12N 15/62(2006. 01)

(30) 优先权数据

13169173. 5 2013. 05. 24 EP

13171191. 3 2013. 06. 10 EP

13194053. 8 2013. 11. 22 EP

61/834100 2013. 06. 12 US

61/909412 2013. 11. 27 US

(85) PCT国际申请进入国家阶段日

2015. 11. 24

(86) PCT国际申请的申请数据

PCT/EP2014/060696 2014. 05. 23

(87) PCT国际申请的公布数据

W02014/187974 EN 2014. 11. 27

(71) 申请人 诺和诺德股份有限公司

地址 丹麦鲍斯韦

(72) 发明人 A. C. 肖

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 李志强 林森

权利要求书1页 说明书31页 附图8页

(54) 发明名称

融合蛋白酶

(57) 摘要

本发明涉及可用于由融合蛋白制造成熟蛋白的新的双功能融合蛋白酶。更具体而言,本发明涉及包含小 RNA 病毒 3C 蛋白酶和 Xaa-Pro- 二肽基氨肽酶的双功能融合蛋白酶。

1. 双功能融合酶,其包含小 RNA 病毒 3C 蛋白酶和 XaaProDAP 的催化结构域。
2. 权利要求 1 的双功能融合蛋白酶,其包含下式的蛋白:
X-Y-Z (I) 或 Z-Y-X (II)
其中
X 是小 RNA 病毒 3C 蛋白酶或其功能变体;
Y 是任选的接头;
Z 是 Xaa-Pro- 二肽基氨肽酶 (XaaProDAP) 或其功能变体;
其中所述融合蛋白酶实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性。
3. 权利要求 1-2 中任一项的双功能融合蛋白酶,其具有式 (I),即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 N- 末端部分。
4. 权利要求 1-3 中任一项的双功能融合蛋白酶,其中 X 是人鼻病毒 3C 蛋白酶或其功能变体。
5. 权利要求 1-4 中任一项的双功能融合蛋白酶,其中 X 包含 SEQ ID NO:2 或其功能变体。
6. 权利要求 1-5 中任一项的双功能融合蛋白酶,其中 Z 是 E. C. 3.4.14.11 酶或其功能变体。
7. 权利要求 6 的双功能融合蛋白酶,其中 Z 是来自乳酸细菌的酶或其功能变体。
8. 权利要求 1-7 中任一项的双功能融合蛋白酶,其中 Z 是 SEQ ID NO:1 或其功能变体。
9. 权利要求 1-6 中任一项的双功能融合蛋白酶,其中 Z 是来自链球菌的酶或其功能变体。
10. 权利要求 10 的双功能融合蛋白酶,其中 Z 是 SEQ ID NO:24 或其功能变体。
11. 权利要求 2-10 中任一项的双功能融合蛋白酶,其中所述功能变体相对于对应的蛋白的天然存在的蛋白或天然存在的子序列包含:1-15 个氨基酸取代、缺失或添加。
12. 权利要求 1-11 中任一项的双功能融合蛋白酶,其包含接头 Y。
13. 权利要求 1-12 中任一项的双功能融合蛋白酶,其包含与 N- 末端连接的标签蛋白。
14. 用于制备权利要求 1-13 中任一项的双功能融合蛋白酶的方法,其包括在宿主细胞中重组表达包含所述双功能融合蛋白酶的蛋白和随后分离所述双功能融合蛋白酶。
15. 权利要求 1-13 中任一项的双功能融合蛋白酶用于由较大的肽或蛋白除去 N- 末端肽或蛋白的用途。

融合蛋白酶

技术领域

[0001] 本发明涉及蛋白表达和蛋白化学的技术领域,其中由融合蛋白释放成熟蛋白。

背景技术

[0002] 重组蛋白技术使得能够生产大量的可用于其生物活性的所需蛋白。这种蛋白通常表达为微生物宿主细胞中的重组融合蛋白。成熟蛋白(目标蛋白)通常与融合伙伴蛋白或较小的氨基酸延伸连接,以便提高表达水平、提高溶解性、促进蛋白折叠或促进纯化和下游加工。

[0003] 对于维持蛋白的完整生物活性以及药物调节目的,由融合蛋白除去融合伙伴蛋白以释放具有天然 N- 和 C- 末端的成熟蛋白可能是关键的。

[0004] 目前,少数可用于由融合蛋白除去融合伙伴蛋白的蛋白酶可用于工业用途的经济上可持续的酶,该蛋白酶在释放的成熟靶蛋白中留下天然 N- 末端。

[0005] 一种这样的酶是肠激酶,然而,其缺少特异性而难以广泛应用。其它的这样的酶有因子 Xa、胰蛋白酶、梭菌蛋白酶、凝血酶、TEV 或鼻病毒 3C 蛋白酶,它们均要么缺少特异性,因为大多数蛋白包含内部二次切割位点,要么在成熟蛋白的 C- 或 N- 末端留下氨基酸延伸。

[0006] Waugh, Protein Expr. Purif. 80:283-293 (2011) 公开了用于除去亲和标签的酶试剂的综述。

[0007] W092/10576 公开了在药物制剂中使用具有 DPP IV 可切割的延伸肽部分的融合蛋白。

[0008] Xin, Protein Expr. Purif. 2002, 24, pp530-538 公开了用于由重组蛋白除去 N- 末端 Pro-Pro 的来自乳酸乳球菌 (*Lactococcus lactis*) 的 X- 脯氨酰基二肽基氨肽酶在大肠杆菌 (*Escherichia coli*) 中的克隆、表达以及应用。

[0009] Bülow, TIBTECH 9:226-231(1991) 公开了一种通过基因融合来制备双功能酶的方法。

[0010] Seo, Appl. Environ. Microbiol. 2000, 66, pp2484-2490 公开了一种海藻糖 -6- 磷酸合酶和海藻糖 -6- 磷酸磷酸酶的双功能融合酶。

[0011] 在制药工业,蛋白药物目前构成竞争市场的相当部分,因此,需要大规模制造这些蛋白药物的高效方法。融合蛋白的工业用途的关键问题仍然是由融合蛋白除去融合蛋白伙伴以释放完整的成熟蛋白。

[0012] 因此,需要一种工业方法,其特异性地除去融合伙伴蛋白,而不成熟蛋白中具有内部切割位点,且不在成熟蛋白上留下任何氨基酸延伸。优选地,该融合伙伴蛋白的除去是应用在工业方法中容易制备的仅单一酶进行。同样需要这样的方法,其可在温和的加工条件下对众多不同的蛋白起到该作用,以防止非计划的成熟蛋白的化学和物理变化。

[0013] 概述

本发明的一个目的在于,提供一种用于由融合蛋白提供成熟蛋白的简单的一步方法。

[0014] 小 RNA 病毒 3C 蛋白酶和 Xaa-Pro- 二肽基氨肽酶 (XaaProDAP) 均是非常特异的酶,

其表现出被发现可用于制造蛋白药物的互补活性。然而,作为蛋白水解酶,它们还引起在融合成为一种双功能融合蛋白酶时自我切割方面的问题。

[0015] 由于两种酶物理上靠近并因此副反应更少,在融合蛋白酶中组合两种酶可具有有利的反应动力学的优点。在融合蛋白酶中组合两种酶还具有的优点是,仅需要提供和使用一种试剂。由于尺寸更大,融合蛋白酶也可容易地通过简单的凝胶过滤方法由成熟蛋白中除去。

[0016] 根据本发明的第一方面,提供一种双功能融合蛋白酶,其包含小 RNA 病毒 3C 蛋白酶和 XaaProDAP 的催化结构域。在一个实施方案中,双功能融合蛋白酶包含小 RNA 病毒 3C 蛋白酶和 XaaProDAP。

[0017] 根据本发明的第二方面,提供一种双功能融合蛋白酶,其包含下式的蛋白:

X-Y-Z (I) 或 Z-Y-X (II)

其中

X 是小 RNA 病毒 3C 蛋白酶或其功能变体;

Y 是任选的接头;

Z 是 Xaa-Pro- 二肽基氨肽酶 (XaaProDAP) 或其功能变体;

其中所述融合蛋白酶实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性。

[0018] 在一个实施方案中,本发明的双功能融合蛋白酶具有式 (I),即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 N- 末端部分。

[0019] 在另一个实施方案中,X 是人鼻病毒 14 型 3C 蛋白酶 (HRV14 3C) 或其功能变体。

[0020] 在另一个实施方案中,Z 是 E. C. 3. 4. 14. 11 酶或其功能变体。

[0021] 根据本发明的第三方面,提供一种制备本发明的双功能融合蛋白酶的方法,其包括在宿主细胞中重组表达包含双功能融合蛋白酶的蛋白和随后分离双功能融合蛋白酶。

[0022] 在一个实施方案中,制备双功能融合蛋白酶的方法包括大肠杆菌作为所述宿主细胞。

[0023] 根据本发明的第四方面,提供本发明的双功能融合蛋白酶用于由较大的肽或蛋白除去 N- 末端肽或蛋白的用途。

[0024] 附图简述

图 1 显示纯化的双功能 HRV14-XaaProDAP 融合蛋白酶 (蛋白酶 20986) 的还原 SDS-PAGE。泳道 1:蛋白标志物。数字表示以 kDa 计的大小。泳道 2:纯化的蛋白酶 20986。

[0025] 图 2 显示在使用 1:20 摩尔的酶与底物比与蛋白酶 20986 于 37°C 孵育 3 小时 (反应 1) 后 RL27_EVLFQGP_PYY (3-36) 的去卷积质谱。X- 轴:以 Da 计的质 / 荷比。Y- 轴:相对强度。

[0026] 图 3 显示在使用 1:40 摩尔的酶与底物比与蛋白酶 20986 于 37°C 孵育 3 小时 (反应 2) 后 RL27_EVLFQGP_PYY (3-36) 的去卷积质谱。X- 轴:以 Da 计的质 / 荷比。Y- 轴:相对强度。

[0027] 图 4 显示在使用 1:20 摩尔的酶与底物比与 RL9-HRV14 3C 蛋白酶于 37°C 孵育 3 小时 (反应 3) 后 RL27_EVLFQGP_PYY (3-36) 的去卷积质谱。X- 轴:以 Da 计的质 / 荷比。Y- 轴:相对强度。

[0028] 图 5 显示在使用 1:40 摩尔的酶与底物比与 RL9-HRV14 3C 蛋白酶于 37℃ 孵育 3 小时（反应 4）后 RL27_EVLFQGP_PYY(3-36) 的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0029] 图 6 显示在使用 1:500 摩尔的酶与底物比与蛋白酶 20986 于 4℃ 孵育过夜（反应 12）后 RL27_EVLFQGP_胰高血糖素的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0030] 图 7 显示在使用 1:100 摩尔的酶与底物比与蛋白酶 28994 于 4℃ 孵育过夜（反应 13）后 RL27_EVLFQGP_胰高血糖素的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0031] 图 8 显示在使用 1:500 摩尔的酶与底物比与蛋白酶 28996 于 4℃ 孵育过夜（反应 16）后 RL27_EVLFQGP_胰高血糖素的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0032] 图 9 显示在使用 1:500 摩尔的酶与底物比与蛋白酶 28997 于 4℃ 孵育过夜（反应 17）后 RL27_EVLFQGP_胰高血糖素的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0033] 图 10 显示在使用 1:20 摩尔的酶与底物比与 RL9-HRV14 3C 蛋白酶于 4℃ 孵育过夜（反应 18, 对照）后 RL27_EVLFQGP_胰高血糖素的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0034] 图 11 显示在使用 1:500 摩尔的酶与底物比与蛋白酶 20986 于 4℃ 孵育过夜（反应 20）后 RL27_EVLFQGP_GLP-1(7-37, K34R) 的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0035] 图 12 显示在使用 1:100 摩尔的酶与底物比与蛋白酶 28994 于 4℃ 孵育过夜（反应 21）后 RL27_EVLFQGP_GLP-1(7-37, K34R) 的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0036] 图 13 显示在使用 1:100 摩尔的酶与底物比与蛋白酶 28996 于 4℃ 孵育过夜（反应 23）后 RL27_EVLFQGP_GLP-1(7-37, K34R) 的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0037] 图 14 显示在使用 1:100 摩尔的酶与底物比与蛋白酶 28997 于 4℃ 孵育过夜（反应 25）后 RL27_EVLFQGP_GLP-1(7-37, K34R) 的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0038] 图 15 显示在使用 1:20 摩尔的酶与底物比与 RL9-HRV14 3C 蛋白酶于 4℃ 孵育过夜（反应 27）后 RL27_EVLFQGP_GLP-1(7-37, K34R) 的去卷积质谱。X-轴：以 Da 计的质 / 荷比。Y-轴：相对强度。

[0039] 描述

根据本发明的第一方面，提供一种双功能融合酶，其包含小 RNA 病毒 3C 蛋白酶和 XaaProDAP 的催化结构域。

[0040] 根据本发明的第二方面，提供一种双功能融合蛋白酶，其包含下式的蛋白：

X-Y-Z (I) 或 Z-Y-X (II)

其中

X 是小 RNA 病毒 3C 蛋白酶或其功能变体；

Y 是任选的接头；

Z 是 Xaa-Pro- 二肽基氨肽酶 (XaaProDAP) 或其功能变体；

其中所述融合蛋白酶实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性。

[0041] 本发明的方法提供优于以前描述的用于由融合蛋白释放成熟蛋白的方法的许多优点。例如,已出乎意料地发现,可获得非常特异性的融合蛋白水解,从而在没有或具有最低水平的相关杂质的情况下以高产率释放具有正确的天然 N- 末端氨基酸的成熟蛋白。存在任何相关杂质(即以具有化学结构的有限差异的方式与成熟蛋白类似的蛋白)显然是不合乎需要的,因为在制造过程中难以将它们除去,因此是昂贵的。另外的实施方案具有下述优点:使得能够在具有低温的反应条件下由融合蛋白释放成熟蛋白。

[0042] 还已出乎意料地发现,本发明的双功能融合蛋白酶可通过在大肠杆菌中重组表达来制备。通常,难以在不引起问题的情况下在大肠杆菌中表达大的蛋白。然而,本发明的双功能融合蛋白酶可通过在大肠杆菌中重组表达来制备,如在公开的本发明实施例中所示。

[0043] 本发明人旨在提供一种融合蛋白酶,其包含功能性 XaaProDAP 和功能性小 RNA 病毒 3C 蛋白酶。这样的双功能融合蛋白酶应当能够在微生物中表达,且应当在表达、纯化以及用于由融合蛋白释放成熟蛋白期间稳定。在制备该双功能融合蛋白酶期间,遇到了许多技术挑战。首先,发现 HRV14 3C 由 HRV14 3C - XaaProDAP 融合蛋白酶切割自身,从而融合蛋白酶不稳定。其次,HRV14 3C 还在来自乳酸乳球菌的 XaaProDAP 中不被识别为典型的 HRV14 3C 切割位点内部切割 HRV14 3C - XaaProDAP 融合蛋白酶。这也使融合蛋白酶不稳定。再次,当 XaaProDAP 位于融合蛋白酶的 C- 末端时,来自乳酸乳球菌的 XaaProDAP 可由 HRV14 3C - XaaProDAP 融合蛋白酶的 N- 末端除去二肽。因此,第一融合蛋白酶表现出在三个不同位点的自我切割,如果双功能融合蛋白酶的表达、纯化、催化功能、稳定性或它们的组合是原因,那么导致没有活性和待解决的挑战性任务。

[0044] 在设计本发明的双功能融合蛋白酶时,可进行下述步骤:

a) 提供不具有蛋白表面上可用的 QG 子序列的 XaaProDAP 或功能变体,

b) 提供小 RNA 病毒 3C 蛋白酶或其功能变体,其若是位于双功能融合蛋白酶的 N- 末端则不在其 N- 末端具有 XaaProDAP 切割位点,且不具有使其能够通过在其 C- 末端的切割而切除自身的切割位点,和

c) 将 XaaProDAP 和小 RNA 病毒 3C 蛋白酶经由任选的氨基酸接头序列连接,从而构成可由单一核酸序列表达的双功能融合蛋白酶。

[0045] 应当理解,术语多肽、肽和蛋白在当前语境中可互换使用。同样,根据 IUPAC 命名法将氨基酸缩写为单字母或三字母名称。

[0046] 本发明的双功能融合蛋白酶优选表现出在低温(例如 2-10°C 或 2-15°C) 的充分的活性,因为从工业制造观点来看这是所需的,例如由于在非无菌工艺条件下控制微生物活性。

[0047] 本文所用的“Xaa-Pro 二肽基氨肽酶”(“XaaProDAP”)旨在意指具有 Xaa-Pro 二肽特异性的二肽酶活性的酶,即易切断的键连接 Xaa-Pro 二肽的 C- 末端与目标肽或蛋白的 N- 末端。根据国际生物化学和分子生物学学会(IUBMB)酶命名法,将 XaaProDAP

分类为来自肽酶家族 S15 的酶 EC 3.4.14.11 和来自肽酶家族 S9B 的酶 EC 3.4.14.5。XaaProDAP 的非限制性实例有来自哺乳动物的二肽基-肽酶 IV (DPP-IV)。XaaProDAP 的其他非限制性实例有来自细胞的 Xaa-脯氨酰基二肽基氨肽酶,例如来自乳酸乳球菌、嗜热链球菌 (*Streptococcus thermophilus*)、德氏乳杆菌 (*Lactobacillus delbrueckii*) 和猪链球菌 (*Streptococcus suis*)。来自乳酸乳球菌乳脂亚种 (*Lactococcus lactis subsp. cremoris*) CNM I-163I 的 Xaa-脯氨酰基二肽基氨肽酶具有下述序列:

MRFNHFSLVDKNFDEQLAELDQLGFRWSVFWDEKKILKDFLIQSPTDMTVLQANTELDVIEFLKSSIELDW
EIFWNITLQLLDFVFNDFEIGKATEFAKKNLNPQRDVEMTTETIISAFYYLLCSRKSGMILVEHWVSEGLPLD
NHYHFFNDKSLATFDSSLLEREVVWVESPDTEQKGNLDLKIQIIRPKSTEKLPVVITASPYHLGINEKANDLALH
EMNVDLEKKDSHKIHVQGLPQKRPSETKELPIVDKAPYRFTHGWYSLNDYFLTRGFASIYVAGVGTGRSNGFQTS
GDYQQIYSMTAVIDWLNTRTRAYTSRKKTHEIKATWANGKVAMTGKSYLGTMAYGAAATGVDGLEVILAEAGISSWY
NYYRENLVRSPPGFGEDLDVLAALTYSRNLDGADYLGKNDYEKRLAEMTTALDRKSGDYNQFWHNRNYLINSQD
VRADVLIHGLQDWNVTPEQAYNFWQALPEGHAKHAFHRGAHIYMNWSQSIDFSETINAYFSAKLLDRDLNLNLPP
VILQENSKEQVWSAVSKFGDDQLKPLGKTAVSFAQFDNHYDDESFKKYSKDFNVFKKDLFENKANEAVIDLLEPS
ELTINGPIELEIRLKLNSKGLLSAQILDFGPKKRLEDKARVKDFKVLDRGRNFMLDDLVELPLVESPYQLVTKGFT
NLQNKDLLTVSDLKADEWFTLKFELQPTIYHLEKADKLRVILYSTDFEHTVRDNRKVTYEIDLSQSKLIPIESVKK
(SEQ ID NO: 1)。

[0048] XaaProDAP 可以是在例如细菌或哺乳动物中天然存在的酶,但是其也可以是这样的酶的功能变体。功能变体的非限制性实例有天然存在的 XaaProDAP 的类似物、延伸或截短形式,该功能变体保留 Xaa-Pro 二肽特异性的二肽酶活性。

[0049] 小 RNA 病毒 3C 蛋白酶 (或蛋白 3C、Picornian 3C 或小 RNA 病毒 3C) 是一组具有丝氨酸蛋白酶样折叠的半胱氨酸蛋白酶,其负责在小 RNA 病毒科病毒中由前体多蛋白产生成熟的病毒蛋白。

[0050] 本文所用的“小 RNA 病毒 3C 蛋白酶”旨在意指来源于小 RNA 病毒科的蛋白酶,包括其功能变体,该蛋白酶切割 P1-P1' Gln-Gly 对之间的肽键,其中易切断的键连接 Gln 与 Gly (其中根据常用记号法, P1 和 P1' 分别表示易切断的键的 N-末端和 C-末端侧的首个氨基酸)。若干小 RNA 病毒 3C 蛋白酶对 P2' 的 Pro 具有另外的优先选择性,其中 P2' 表示易切断的键的 C-末端侧的第二个氨基酸。具有该底物特异性的酶典型地分离自肠道病毒 (enterovirus) 属的病毒,其目前包括柯萨奇病毒 (Coxsackie virus)、埃可病毒 (Echovirus)、肠道病毒、脊髓灰质炎病毒 (Poliovirus) 和鼻病毒。这样的小 RNA 病毒 3C 蛋白酶的非限制性实例有:人鼻病毒 14 型 3C (HRV14 3C) 蛋白酶,其具有序列 GPNTEFALSLLRKNIMTITTSKGEFTGLGIHDRVCVIPHTAQPGDDVLVNGQKIRVKDKYKLVDPENINLELTVLTLDRNEKFRDIRGFISEDLEGVDATLVVHNNFTNTILEVGPVTMAGLINLSSPTNRMIRYDYATKTGQCGVLCATGKIFGIHVGGNGRQGFSAQLKKQYFVEKQ (SEQ ID NO: 2);肠道病毒 71 3C 蛋白酶;柯萨奇病毒 A16 3C 蛋白酶;柯萨奇病毒 B3 3C 蛋白酶;豇豆花叶豇豆花叶病毒型小 RNA 病毒 3C;和人脊髓灰质炎病毒 3C 蛋白酶。这些 3C 蛋白酶能够由大的融合蛋白释放在 N-末端具有 Gly-Pro 的蛋白,且通常可通过具有在其自身的天然 N-末端天然存在的 Gly-Pro 来鉴定。根据本发明,小 RNA 病毒 3C 蛋白酶可以是小 RNA 病毒科中天然存在的酶,但是其也可以是这样的酶的功能变体。功能变体的非限制性实例有天然存在的小 RNA 病毒 3C 蛋白酶的类似物、延伸或截短形式,

该功能变体保留对于 Gln-Gly 对的底物特异性。

[0051] 本文所用的“实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性”旨在意指，在表达条件、纯化条件、贮存条件和切割靶蛋白的前体的制造用途中，双功能融合蛋白酶不切割自身，或者仅以极慢的速率切割自身，而不阻碍其切割靶蛋白的前体的预期用途。

[0052] 在一个实施方案中，通过双功能融合蛋白酶在制造条件下足够稳定用于切割靶蛋白的前体，来确定“实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性”。

[0053] 在另一个实施方案中，确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶，是通过所述双功能融合蛋白酶适合于其预期用途确定。

[0054] 在另一个实施方案中，确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶，是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液，pH 7.4 中于温度 37°C 孵育所述双功能融合蛋白酶 3 小时后至少 50% 的双功能融合蛋白酶是完整的来确定。

[0055] 在另一个实施方案中，确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶，是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液，pH 7.4 中于温度 37°C 孵育所述双功能融合蛋白酶 3 小时后双功能融合蛋白酶的至少 50% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的来确定。

[0056] 在另一个实施方案中，确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶，是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液，pH 7.4 中于温度 37°C 孵育所述双功能融合蛋白酶 3 小时后双功能融合蛋白酶的至少 80% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的来确定。

[0057] 在另一个实施方案中，确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶，是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液，pH 7.4 中于温度 4°C 孵育所述双功能融合蛋白酶 24 小时后双功能融合蛋白酶的至少 50% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的来确定。

[0058] 在另一个实施方案中，确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶，是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液，pH 7.4 中于温度 4°C 孵育所述双功能融合蛋白酶 24 小时后双功能融合蛋白酶的至少 80% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的来确定。本文所用的“成熟蛋白”旨在意指目标蛋白、肽或多肽或其延伸形式，该延伸形式可被 XaaProDAP 在其 N- 末端切割。成熟蛋白在其制造期间通常以融合蛋白存在，例如除了成熟蛋白之外还包含标签序列、任选的接头序列和小 RNA 病毒 3C 蛋白酶位点的蛋白。成熟蛋白的非限制性实例有胰高血糖素、PYY(3-36)、GLP-1(7-37)、Arg34-GLP1(7-37)、Arg34-GLP-1(9-37) 和 Arg34-GLP-1(11-37)。使用常用的氨基酸残基单字母缩写，例如，Arg34-GLP-1(7-37) 是 K34R-GLP-1(7-37) (亦称为 GLP-1(7-37, K34R))。

[0059] 本文所用的“融合蛋白”旨在意指可通过包含编码至少两个不同蛋白的核苷酸序列的核酸分子表达的杂合蛋白。例如，融合蛋白可包含与具有目标药物活性的蛋白融合的

标签蛋白。融合蛋白通常用于提高治疗蛋白的重组表达以及用于提高这样蛋白由细胞培养物的回收和纯化等。融合也可用于将两种不同酶活性组合到一种单一蛋白中。融合蛋白也可包含人工序列,例如接头序列。

[0060] 本文所用的“融合蛋白酶”旨在意指可通过包含编码至少两个均具有蛋白水解活性的不同蛋白的核苷酸序列的核酸分子表达的杂合蛋白。例如,融合蛋白酶可包含两个不同的蛋白酶,例如内肽酶和外切蛋白酶。融合蛋白酶也可包含例如与两个蛋白水解性蛋白融合的标签蛋白。

[0061] 在一个实施方案中,融合蛋白酶所包含的两个不同的蛋白表现出两种不同的蛋白水解活性。在另一个实施方案中,融合蛋白酶所包含的两个不同的蛋白是来源于不同生物体的蛋白酶或其功能变体。

[0062] XaaProDAP 蛋白酶具有包含经由大的蛋白环连接在一起的两个 α 螺旋的蛋白结构。该环暴露于蛋白的表面,因此易于被小 RNA 病毒 3C 蛋白酶切割,尤其是在该小 RNA 病毒 3C 蛋白酶和 XaaProDAP 包含在双功能融合蛋白酶中时。XaaProDAP 的连接两个小的 α -螺旋的环表示 XaaProDAP 蛋白酶中的高度保守区。在 SEQ ID NO: 1 中,环是跨越大约 223 to 270 的残基的子序列。本发明人发现,XaaProDAP 在与 HRV14 3C 融合时是不稳定的,而这是由于 241-242 位的 QG 子序列的 HRV14 3C 切割所引起的。这是非常出乎意料的,因为该环并不包含作为常见的小 RNA 病毒 3C 蛋白酶切割位点的子序列。因此,通过使用将 QG 氨基酸取代成其他氨基酸(例如 ET)的 XaaProDAP 功能变体,解决了该问题。

[0063] 本文所用的“融合伙伴蛋白”或“融合伙伴”旨在意指作为融合蛋白的一部分的蛋白,即融合蛋白所包含的至少两个蛋白中的一个。融合伙伴蛋白的非限制性实例有标签蛋白和增溶结构域,例如 His6- 标签、麦芽糖结合蛋白、硫氧还蛋白等。

[0064] 本文所用的“融合酶”旨在意指包含至少两个均为酶的蛋白的融合蛋白(意思是两个蛋白具有共价连接的骨架序列)。

[0065] 本文所用的“标签蛋白”或“标签”旨在意指蛋白,其与另一个蛋白连接以促进或提高所述另一个蛋白的制造,例如促进或提高所述另一个蛋白的重组表达、回收和/或纯化。标签蛋白的非限制性实例有 His6- 标签、谷胱甘肽 S- 转移酶(GST)、麦芽糖结合蛋白(MBP)、金黄色葡萄球菌蛋白 A、生物素化肽以及 W02006/108826 和 W02008/043847 中所述的来自嗜热细菌的高碱性蛋白。

[0066] 本文所用的“标签序列”旨在意指包含蛋白的序列。标签序列还可任选包含额外的序列,例如接头序列。蛋白标签是遗传接枝到重组蛋白上的肽序列,其可通过化学试剂或通过酶方法(例如蛋白水解)除去。标签与蛋白连接用于多种目的,例如为了促进蛋白的表达或由细胞分泌、为了提高蛋白溶解性或为了助于蛋白的正确折叠。

[0067] 本文所用的“接头”旨在意指通常用于促进融合蛋白的功能、折叠或表达的氨基酸序列。本领域技术人员已知,以融合酶形式存在的两个蛋白可能干扰彼此的酶活性,这是一种一般可通过在两个酶序列之间插入接头来消除或减少的相互作用。

[0068] 本文所用的“类似物”旨在意指通过取代、缺失和/或添加另一个蛋白的一个或多个氨基酸残基而来源于该蛋白的蛋白。GLP-1(7-37)类似物的非限制性实例有残基 34 已取代为精氨酸残基的 K34R-GLP-1(7-37)和残基 34 已取代为精氨酸残基且氨基酸残基 7-8 已缺失的 K34R-GLP-1(9-37)(使用 GLP-1 肽的氨基酸残基的常用编号)。

[0069] 本文所用的“功能变体”旨在意指某蛋白的化学变体,其具有改变的氨基酸序列,但是保留与原始蛋白基本上相同的功能。因此,功能变体通常是蛋白的修饰形式,其中视需要对修饰蛋白引入尽可能少的修饰以获得一些所需的性质,同时保留与原始蛋白基本上相同的功能。功能变体的非限制性实例例如有延伸蛋白、截短蛋白、融合蛋白和类似物。HRV14 3C 功能变体的非限制性实例例如有带 His6 标签的 HRV14 3C、带 GST- 标签的 HRV14 3C 和截短(例如为了不包含 N- 末端 GP 二肽)的 HRV14 3C。GLP-1(7-37) 的非限制性功能变体有 K34R-GLP-1(7-37)。

[0070] 在一个实施方案中,蛋白的功能变体与所述蛋白相比包含 1-2 个氨基酸取代、缺失或添加。在另一个实施方案中,功能变体与所述蛋白相比包含 1-5 个氨基酸取代、缺失或添加。在另一个实施方案中,功能变体相对于对应的蛋白的天然存在的蛋白或天然存在的子序列包含 1-15 个氨基酸取代、缺失或添加。

[0071] 本文所用的“增溶结构域”旨在意指蛋白,其是融合蛋白的一部分,且用于使所述融合蛋白与目标蛋白本身相比在特定条件下溶解性更高。增溶结构域的非限制性实例有 DsbC (硫醇、二硫键互换蛋白)、W02008/043847 中所述的 RL9 (核糖体蛋白 L9)、MPB (麦芽糖结合蛋白)、NusA (转录终止 / 抗终止蛋白) 和 Trx(硫氧还蛋白)。

[0072] 本文所用的“酶处理”旨在意指使底物蛋白与催化至少一个涉及所述底物蛋白的反应的酶接触。一种常见的酶处理是使融合蛋白与具有蛋白水解活性的酶接触,以分离作为融合蛋白的组成的两个蛋白。

[0073] 根据本发明的第四方面,提供本发明的双功能融合蛋白酶用于由较大的肽或蛋白除去 N- 末端肽或蛋白以便获得具有预期的 N- 末端 aa 残基的成熟蛋白。所述较大的肽或蛋白通常是融合蛋白,其包含成熟蛋白和一个或多个用来促进蛋白的重组表达、正确折叠、纯化目的等的标签序列。

[0074] 在一个实施方案中,所述较大的肽或蛋白与所述双功能融合蛋白酶在合适的反应条件下接触足够时间,以释放大数目的所述 N- 末端肽。反应条件例如可包括下述范围的 pH :约 6.0-约 9.0、约 7.0-约 8.5、约 7.5-约 8.5、约 8.0-约 9.0 或约 6.0-约 7.0。反应条件可包括下述范围的温度 :约 0°C -约 50°C、约 30°C -约 37°C、约 0°C -约 15°C、约 0°C -约 10°C、约 2°C -约 10°C、约 5°C -约 15°C、约 0°C -约 5°C 或约 2°C -约 8°C。在另一个实施方案中,反应条件包括范围为约 pH 7.5-约 pH 8.5 的 pH 和范围为约 4°C -约 10°C 的温度。在又一个实施方案中,反应条件包括范围为约 1 分钟 -约 3 小时的反应时间。在又一个实施方案中,反应条件包括范围为约 3 小时 -约 24 小时的反应时间。在又一个实施方案中,反应时间范围为约 3 小时 -约 24 小时、约 3 小时 -约 16 小时、约 6 小时 -约 24 小时、约 10 小时 -约 16 小时。在另一个实施方案中,反应条件包括水性介质,其包含磷酸缓冲盐水,例如 50 mM 磷酸钠+0.9% 氯化钠。磷酸缓冲盐水(缩写为 PBS) 是常用的缓冲溶液,且典型地是水基盐溶液,包含磷酸钠、氯化钠和(在一些溶液中)氯化钾和磷酸钾。用于本发明中酶反应的典型的 1x PBS 缓冲液是(8.05 mM Na₂HPO₄·2H₂O, 1,96 mM KH₂PO₄, 140 mM NaCl, pH 7.4)。

[0075] 其它可用于反应介质的缓冲液可为 TRIS(三(羟基甲基)-氨基甲烷)或 HEPES(4-(2-羟基乙基)-1-哌嗪乙磺酸)缓冲液。

[0076] 在另一个实施方案中,双功能融合蛋白酶与所述较大的肽或蛋白共表达,以在宿

主细胞内表达期间体内释放目标蛋白。在另一个实施方案中,所述较大的肽或蛋白与所述双功能融合蛋白酶在从用于其表达的宿主细胞分离这两种蛋白之后接触。

[0077] 在另一个实施方案中,所述较大的肽或蛋白选自包含肽或蛋白,其包含选自下述的肽:GLP-1(胰高血糖素样肽 1)、胰高血糖素、肽 YY (PYY) 和淀粉样肽 (amylin) 以及它们的功能变体。

[0078] 在又一个实施方案中,所述较大的肽或蛋白的大小少于 200 个氨基酸残基、少于 150 个氨基酸残基、少于 100 个残基或少于 60 个氨基酸残基。

[0079] “施加物 (application)”意指包含融合蛋白的样品,其被加到纯化柱上。

[0080] “流穿液 (flow through)”意指含有宿主细胞蛋白和与纯化柱结合的污染物的施加物部分。

[0081] “主要峰”是指纯化色谱图中的峰,其具有最高的 UV 强度且含有融合蛋白。

[0082] “UV280 强度”是以毫吸光度单位测量的在 280nm 波长处的吸光度,在该波长蛋白会吸收。

[0083] “UV215”是以毫吸光度单位测量的在 215nm 波长处的吸光度,在该波长蛋白会吸收。

[0084] “IPTG”是异丙基 - β -D- 硫代吡喃型半乳糖苷。

[0085] TIC 是总离子计数 (Total Ion Count)。

[0086] HPLC 是高效液相色谱。

[0087] LC-MS 是指液相色谱质谱。

[0088] “% 纯度”定义为特定蛋白的量 / (特定蛋白的量 + 污染物的量) X 100。

[0089] SDS-PAGE 是十二烷基硫酸钠聚丙烯酰胺凝胶电泳。

[0090] 根据本发明的第三方面,提供一种用于制备本发明的双功能融合蛋白酶的方法,其包括在宿主细胞中重组表达包含双功能融合蛋白酶的蛋白和随后分离双功能融合蛋白酶。

[0091] 在一个实施方案中,用于制备双功能融合蛋白酶的方法包括大肠杆菌作为所述宿主细胞。

[0092] 在另一个实施方案中,用于制备双功能融合蛋白酶的方法包括将所述双功能融合蛋白酶分离为可溶性蛋白。

[0093] 在另一个实施方案中,用于制备双功能融合蛋白酶的方法包括在不使用重折叠步骤的情况下将所述双功能融合蛋白酶分离为可溶性蛋白。

[0094] 在另一个实施方案中,用于制备双功能融合蛋白酶的方法包括具有实施方案 2 所示的式 (I) 的双功能融合蛋白酶,即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 N- 末端部分。

[0095] 双功能融合蛋白酶可通过重组蛋白技术产生。通常,对克隆的野生型小 RNA 病毒 3C 蛋白酶和克隆的野生型 XaaProDAP 核酸序列或其功能变体进行修饰,以编码所需的融合蛋白。该修饰包括框内融合编码待表达为融合蛋白的两个以上蛋白的核酸序列。这样的融合蛋白可以是双功能融合蛋白酶,具有或不具有接头肽,以及双功能融合蛋白酶与标签融合,例如 His- 标签或增溶结构域 (例如 DsbC、RL9、MBP、NusA 或 Trx)。然后将该修饰序列插入表达载体,继而转化或转染到表达宿主细胞中。

[0096] 编码双功能融合蛋白酶的核酸构建体可合适地为基因组、cDNA 或合成来源。通过利用公知技术改变遗传密码来完成氨基酸序列改变。

[0097] 通常将编码双功能融合蛋白酶的 DNA 序列插入可为任何载体的重组载体,其可合宜地进行重组 DNA 程序,载体的选择常常取决于其所要引入的宿主细胞。因此,载体可为自主复制载体,即以染色体外实体存在的载体,其复制与染色体复制无关,例如质粒。备选地,载体可为下述载体:在引入宿主细胞中时,其被整合到宿主细胞基因组中,并与整合它的染色体一起复制。

[0098] 载体优选表达载体,其中编码双功能融合蛋白酶的 DNA 序列与 DNA 转录所需的另外的区段有效连接。术语“有效连接”表示区段的排列使得它们对于它们的预期目的协调地起作用,例如转录起始于启动子并行进穿过编码多肽的 DNA 序列,直到其终止于终止子。

[0099] 因此,用于表达双功能融合蛋白酶的表达载体包含能够起始和指导克隆的基因或 cDNA 的转录的启动子。启动子可为任何 DNA 序列,其显示出所选宿主细胞中的转录活性,且可来源于编码对于宿主细胞而言或同源或异源的蛋白的基因。

[0100] 此外,用于双功能融合蛋白酶的表达载体还会包含终止子序列,其是为宿主细胞识别以终止转录的序列。终止子序列与编码多肽的核酸序列的 3' 末端有效连接。任何在所选宿主细胞中有功能的终止子均可用于本发明。

[0101] 双功能融合蛋白酶的表达可瞄准宿主细胞的胞质溶质中的胞内表达,或被导向分泌途径,以胞外表达达到培养基中。

[0102] 胞内表达是默认途径,要求表达载体具有 DNA 序列,其包含启动子,后为编码双功能融合蛋白酶多肽的 DNA 序列,然后是终止子。

[0103] 为了将双功能融合蛋白酶导向宿主细胞的分泌途径,需要分泌信号序列(亦被称为信号肽或前序列)作为双功能融合蛋白酶的 N-末端延伸。将编码信号肽的 DNA 序列在正确的阅读框内连接至编码双功能融合蛋白酶的 DNA 序列的 5' 端。信号肽可以是通常与该蛋白连接的,或可来自编码另一个分泌性蛋白的基因。

[0104] 用于连接分别编码双功能融合蛋白酶、启动子、终止子和 / 或分泌信号序列的 DNA 序列的程序和用于将它们插入含有复制所必需的信息的合适载体中的程序均是本领域技术人员周知的(参见例如 Sambrook 等, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, New York, 1989)。

[0105] 引入编码双功能融合蛋白酶的 DNA 序列的宿主细胞可以是能够胞内或胞外表达双功能融合蛋白酶的任何细胞。如果需要翻译后修饰,那么合适的宿主细胞包括酵母、真菌、昆虫、和高等真核细胞,例如哺乳动物细胞。

[0106] 细菌表达

对于大肠杆菌中的表达,用于指导核酸构建体在细菌宿主细胞中转录的合适启动子实例有可获自以下的启动子:*lac*操纵子、*trp*操纵子及其杂合体 *trc*和 *tac*,均来自大肠杆菌(DeBoer 等, 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25)。用于大肠杆菌的其它甚至更强的启动子是来自 T7 和 T5 噬菌体的噬菌体启动子。T7 启动子需要大肠杆菌宿主中存在 T7 聚合酶(Studier 和 Moffatt, *J. Mol. Biol.* 189, 113, (1986))。所有这些启动子受 IPTG、乳糖或色氨酸诱导的调节,以在细菌生长期的策略性时间点起始转录。大肠杆菌还具有用于连续表达的强启动子,例如 Dalbøge 等, 1987,

Biotechnology 5, 161-164 中用于表达 hGH 的合成启动子。

[0107] 对于杆菌 (*Bacillus*) 中的表达, 合适的实例有来自以下的启动子: 枯草芽孢杆菌 (*Bacillus subtilis*) 果聚糖蔗糖酶基因 (*sacB*)、地衣芽孢杆菌 (*Bacillus licheniformis*) α -淀粉酶基因 (*amyL*)、嗜热脂肪芽孢杆菌 (*Bacillus stearothermophilus*) 产麦芽糖淀粉酶基因 (*amyM*)、解淀粉芽孢杆菌 (*Bacillus amyloliquefaciens*) α -淀粉酶基因 (*amyQ*)、地衣芽孢杆菌青霉素酶基因 (*penP*)、枯草芽孢杆菌 *xyIA* 和 *xyIB* 基因。更多的启动子见述于: "Useful proteins from recombinant bacteria (来自重组细菌的有用蛋白)", *Scientific American*, 1980, 242:74-94; 和 Sambrook 等, 1989, 同上。

[0108] 对于大肠杆菌, 细菌宿主细胞的有效信号肽编码区有获自下述基因的信号肽: *DegP*、*OmpA*、*OmpF*、*OmpT*、*PhoA* 和内毒素 STII, 均来自大肠杆菌。对于杆菌, 信号肽区获自杆菌 NCIB 11837 产麦芽糖淀粉酶、嗜热脂肪芽孢杆菌 α -淀粉酶、地衣芽孢杆菌枯草杆菌蛋白酶、地衣芽孢杆菌 β -内酰胺酶、嗜热脂肪芽孢杆菌中性蛋白酶 (*nprT*、*nprS*、*nprM*) 和枯草芽孢杆菌 *prsA*。更多的信号肽见述于 Simonen 和 Palva, 1993, *Microbiological Reviews* 57:109-137。对于大肠杆菌和杆菌二者, 可根据算法 SignalP (Nielsen 等, 1997, 蛋白 Eng. 10, 1-6., Emanuelsen 等, 2007, *Nature Protocols* 2, 953-971) 中列出的规则从头产生信号肽。使信号序列适于给定环境, 并检查 SignalP 评分。

[0109] 强转录终止子的实例有: Thiofusion 表达系统中的天冬氨酸酶 *aspA*、PET 载体中的 T7 基因 10 终止子 (Studier 等) 和核糖体 RNA 基因 *rrnA*、*rrnD* 的终止子。

[0110] 优选表达宿主的实例有大肠杆菌 K12 W3110、具有 B, MC1061 记号的大肠杆菌 K12、和大肠杆菌 B BL21 DE3, 其通过用噬菌体进行溶源化而含有 T7 聚合酶。当用质粒转化以表达时, 可用抗生素选择这些宿主。对于无抗生素选择, 优选的宿主有例如具有下述缺失的大肠杆菌 B BL21 DE3 3xKO: 缺失 2 D, L-丙氨酸消旋酶基因 $\Delta a1r$ 、 $\Delta dadX$ 和缺失大肠杆菌 B 特有且常常与病原性质相关的 II 组荚膜基因簇 $\Delta (kpsM-kpsF)$ 。缺失 II 组基因簇使大肠杆菌 B BL21 DE3 3xKO 的安全类别与大肠杆菌 K12 相同。选择基于不需要 D-丙氨酸, 这是插入表达质粒中替代 AmpR 基因的 *a1r* 基因所提供的。

[0111] 一旦在宿主生物体中表达了双功能融合蛋白酶, 可通过常规技术将其回收并纯化至所需的纯度。这种常规的回收和纯化技术的非限制性实例有离心、溶解、过滤、沉淀、离子交换色谱、固定化金属亲和色谱 (IMAC)、RP-HPLC、凝胶过滤和冷冻干燥。

[0112] 重组表达和纯化 HRV14 3C 的实例可见于例如: Cordingley 等, *J. Virol.* 1989, 63, pp5037-5045; Birch 等, *Protein Expr Purif.*, 1995, 6, pp609-618; 和 W02008/043847。

[0113] 由乳酸乳球菌进行 XaaProDAP 的微生物表达和纯化的实例可见于例如: Chich 等, *Anal. Biochem.* 1995, 224, pp 245-249; 和 Xin 等, *Protein Expr. Purif.* 2002, 24, pp530-538。

[0114] 本发明进一步通过下述非限制性实施方案说明:

1. 双功能融合酶, 其包含小 RNA 病毒 3C 蛋白酶和 XaaProDAP 的催化结构域。

[0115] 2. 实施方案 1 的双功能融合蛋白酶, 其包含下式的蛋白蛋白:

X-Y-Z (I) 或 Z-Y-X (II)

其中

X 是小 RNA 病毒 3C 蛋白酶或其功能变体；

Y 是任选的接头；

Z 是 Xaa-Pro- 二肽基氨肽酶 (XaaProDAP) 或其功能变体；

其中所述融合蛋白酶实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性。

[0116] 3. 实施方案 1-2 中任一项的双功能融合蛋白酶,其具有式 (I),即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 N- 末端部分。

[0117] 4. 实施方案 1-3 中任一项的双功能融合蛋白酶,其中 X 是鼻病毒蛋白酶或其功能变体。

[0118] 5. 实施方案 1-3 中任一项的双功能融合蛋白酶,其中 X 是小 RNA 病毒蛋白酶或其功能变体。

[0119] 6. 实施方案 1-4 中任一项的双功能融合蛋白酶,其中 X 是 HRV14 3C 或其功能变体。

[0120] 7. 实施方案 1-6 中任一项的双功能融合蛋白酶,其中 X 包含 SEQ ID NO:2 或其功能变体。

[0121] 8. 实施方案 5-6 中任一项的双功能融合蛋白酶,其中 X 是 X₁选自除了 P 之外的遗传编码氨基酸残基的 P2X₁ - SEQ ID NO:2 或 G1P - SEQ ID NO:2 或者其功能变体。

[0122] 9. 实施方案 5-6 中任一项的双功能融合蛋白酶,其中 X 是 CVB3 3C 或其功能变体。

[0123] 10. 实施方案 5 的双功能融合蛋白酶,其中 X 包含 SEQ ID NO:23 或其功能变体。

[0124] 11. 实施方案 1-10 中任一项的双功能融合蛋白酶,其中 X 是 C- 末端截短的功能性小 RNA 病毒 3C 蛋白酶或其功能变体。

[0125] 12. 实施方案 11 的双功能融合蛋白酶,其中所述 C- 末端截短的功能性小 RNA 病毒 3C 蛋白酶截短了不多于 20 个氨基酸残基,例如不多于 10 个氨基酸残基,例如不多于 5 个氨基酸残基,例如不多于 2 个氨基酸残基。

[0126] 13. 实施方案 1-12 中任一项的双功能融合蛋白酶,其中 X 是来自选自以下的病毒的酶:肠道病毒、柯萨奇病毒、埃可病毒、豇豆花叶豇豆花叶病毒、鼻病毒和脊髓灰质炎病毒,或者其功能变体。

[0127] 14. 实施方案 1-13 中任一项的双功能融合蛋白酶,其中 Z 是 E. C. 3.4.14.11 酶或其功能变体。

[0128] 15. 实施方案 14 的双功能融合蛋白酶,其中 Z 是来自乳酸细菌的酶或其功能变体。

[0129] 16. 实施方案 15 的双功能融合蛋白酶,其中 Z 是来自乳球菌 (*Lactococcus spp.*)、链球菌 (*Streptococcus spp.*)、乳杆菌 (*Lactobacillus spp.*)、双歧杆菌 (*Bifidobacterium spp.*) 的酶或其功能变体。

[0130] 17. 实施方案 1-16 中任一项的双功能融合蛋白酶,其中 Z 是 SEQ ID NO:1 或其功能变体。

[0131] 18. 实施方案 1-14 中任一项的双功能融合蛋白酶,其中 Z 是来自 *Bacillus spp.*

的酶或其功能变体。

[0132] 19. 实施方案 1-16 中任一项的双功能融合蛋白酶,其中 Z 是来自 *Streptococcus suis* 的酶或其功能变体。

[0133] 20. 实施方案 17 的双功能融合蛋白酶,其中 Z 是 SEQ ID NO: 24 或其功能变体。

[0134] 21. 实施方案 1-17 中任一项的双功能融合蛋白酶,其中 Z 是来自乳酸乳球菌的酶或其功能变体。

[0135] 22. 实施方案 1-13 中任一项的双功能融合蛋白酶,其中所述 Z 是 E. C. 3. 4. 14. 5 酶或其功能变体。

[0136] 23. 实施方案 1-22 中任一项的双功能融合蛋白酶,其中 Z 是具有暴露的连接两个 α -螺旋的环的蛋白。

[0137] 24. 实施方案 23 的双功能融合蛋白酶,其中所述环不包含任何 QG 子序列。

[0138] 25. 实施方案 23-24 中任一项的双功能融合蛋白酶,其中所述环不包含下述子序列中的任一个:QS、QI、QN、QA 和 QT。

[0139] 26. 实施方案 23-25 中任一项的双功能融合蛋白酶,其中所述环是跨越 SEQ ID NO: 1 中氨基酸残基 223-270 的序列。

[0140] 27. 实施方案 23-26 中任一项的双功能融合蛋白酶,其中所述环是 XaaProDAP 中对应于 SEQ ID NO: 1 中跨越氨基酸残基 223-270 的序列的序列。

[0141] 28. 实施方案 23-27 中任一项的双功能融合蛋白酶,其中所述环是与跨越 SEQ ID NO: 1 中氨基酸残基 223-270 的残基的序列具有至少 70% 氨基酸同一性的序列。

[0142] 29. 实施方案 1-28 中任一项的双功能融合蛋白酶,其中 Z 包含不多于 1 个的 QG 子序列。

[0143] 30. 实施方案 1-29 中任一项的双功能融合蛋白酶,其中 Z 不包含任何 QG 子序列。

[0144] 31. 实施方案 1-17 中任一项的双功能融合蛋白酶,其中 Z 包含 Q241-G242 中氨基酸残基的至少 1 个取代、添加或缺失。

[0145] 32. 实施方案 31 的双功能融合蛋白酶,其中 Z 包含取代 Q241E、G242T。

[0146] 33. 实施方案 1-32 中任一项的双功能融合蛋白酶,其中所述融合蛋白酶中 N-末端起第 2 个氨基酸残基不是 P。

[0147] 34. 实施方案 1-33 中任一项的双功能融合蛋白酶,其中所述融合蛋白酶中 N-末端起第 2 个氨基酸残基不是 G、A 和 T。

[0148] 35. 实施方案 1-33 中任一项的双功能融合蛋白酶,其中所述融合蛋白酶的 N-末端具有氨基酸序列 MX_1P ,其中 X_1 是使 MX_1P 序列为甲硫氨酸氨肽酶的差底物的氨基酸。

[0149] 36. 实施方案 1-34 中任一项的双功能融合蛋白酶,其中所述双功能融合蛋白酶的 N-末端氨基酸残基是 P。

[0150] 37. 实施方案 36 的双功能融合蛋白酶,其中所述融合蛋白酶中 N-末端起第 2 个氨基酸残基不是 P、G、A 或 T。

[0151] 38. 实施方案 1-37 中任一项的双功能融合蛋白酶,其不包含接头 Y。

[0152] 39. 实施方案 1-37 中任一项的双功能融合蛋白酶,其包含接头 Y。

[0153] 40. 实施方案 39 的双功能融合蛋白酶,其中所述接头 Y 的长度为 2-100 个氨基酸残基。

[0154] 41. 实施方案 39-40 中任一项的双功能融合蛋白酶,其中所述接头 Y 的长度为 2-50 个氨基酸残基。

[0155] 42. 实施方案 39-41 中任一项的双功能融合蛋白酶,其中所述接头 Y 的长度为 2-25 个氨基酸残基。

[0156] 43. 实施方案 39-42 中任一项的双功能融合蛋白酶,其中所述接头 Y 的长度为 2-15 个氨基酸残基。

[0157] 44. 实施方案 39-41 中任一项的双功能融合蛋白酶,其中 Y 的长度为约 5- 约 50 个氨基酸残基。

[0158] 45. 实施方案 38-39 中任一项的双功能融合蛋白酶,其中 Y 的长度为约 5- 约 15 个氨基酸残基。

[0159] 46. 实施方案 39-45 中任一项的双功能融合蛋白酶,其中 Y 不包含 Cys 残基。

[0160] 47. 实施方案 39-46 中任一项的双功能融合蛋白酶,其中 Y 不包含 Gln 残基。

[0161] 48. 实施方案 39-47 中任一项的双功能融合蛋白酶,其中 Y 仅包含下述氨基酸残基 :G、S、A、L、P 和 T。

[0162] 49. 实施方案 39-48 中任一项的双功能融合蛋白酶,其中 Y 选自 SEQ ID NO: 3、4 和 12。

[0163] 50. 实施方案 1-49 中任一项的双功能融合蛋白酶,其是式 (I),即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 N- 末端部分。

[0164] 51. 实施方案 50 的双功能融合蛋白酶,其中 X 不具有 C- 末端氨基酸残基 Q。

[0165] 52. 实施方案 1-49 中任一项的双功能融合蛋白酶,其是式 (II),即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 C- 末端部分。

[0166] 53. 实施方案 1-52 中任一项的双功能融合蛋白酶,其包含与 N- 末端连接的标签蛋白。

[0167] 54. 实施方案 53 的双功能融合蛋白酶,其中所述标签蛋白选自 His- 标签、增溶结构域和带 His- 标签的增溶结构域。

[0168] 55. 实施方案 1-54 中任一项的双功能融合蛋白酶,其中所述功能变体相对于对应的天然存在的蛋白或天然存在的子序列包含 :1-2 个氨基酸取代、缺失或添加,或 1-5 个氨基酸取代、缺失或添加,或 1-15 个氨基酸取代、缺失或添加。

[0169] 56. 实施方案 1-55 中任一项的双功能融合蛋白酶,其中确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶,是通过所述双功能融合蛋白酶适合用于其预期用途确定。

[0170] 57. 实施方案 1-55 中任一项的双功能融合蛋白酶,其中确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶,是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液, pH 7.4 中于温度 37°C 孵育所述双功能融合蛋白酶 3 小时后至少 50% 的双功能融合蛋白酶是完整的来确定。

[0171] 58. 实施方案 1-55 中任一项的双功能融合蛋白酶,其中确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶,是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液, pH 7.4 中于温度 37°C 孵育所述双功能融合蛋白酶 3 小时后双功能融合蛋白酶的至少 50% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的来

确定。

[0172] 59. 实施方案 58 的双功能融合蛋白酶,其中在以 0.5 mg/mL 浓度在 1x PBS 缓冲液, pH 7.4 中于温度 37°C 孵育所述双功能融合蛋白酶 3 小时后双功能融合蛋白酶的至少 80% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的。

[0173] 60. 实施方案 1-55 中任一项的双功能融合蛋白酶,其中确定所述实质上不具有能够损害两种蛋白水解活性中的至少一种的自我切割活性的融合蛋白酶,是通过在以 0.5 mg/mL 浓度在 1x PBS 缓冲液, pH 7.4 中于温度 4°C 孵育所述双功能融合蛋白酶 24 小时后双功能融合蛋白酶的至少 50% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的来确定。

[0174] 61. 实施方案 60 的双功能融合蛋白酶,其中在以 0.5 mg/mL 浓度在 1x PBS 缓冲液, pH 7.4 中于温度 4°C 孵育所述双功能融合蛋白酶 24 小时后双功能融合蛋白酶的至少 80% 的小 RNA 病毒 3C 蛋白酶活性和 XaaProDAP 活性是完整的。

[0175] 62. 用于制备实施方案 1-61 中任一项的双功能融合蛋白酶的方法,其包括在宿主细胞中重组表达包含双功能融合蛋白酶的蛋白和随后分离双功能融合蛋白酶。

[0176] 63. 实施方案 62 的方法,其中所述宿主细胞是大肠杆菌。

[0177] 64. 实施方案 62-63 中任一项的方法,其中所述双功能融合蛋白酶分离为可溶性蛋白。

[0178] 65. 实施方案 62-64 中任一项的方法,其中所述双功能融合蛋白酶在使用重折叠步骤的情况下分离为可溶性蛋白。

[0179] 66. 实施方案 62-65 中任一项的方法,其中所述双功能融合蛋白酶具有实施方案 2 所示的式 (I),即所述小 RNA 病毒 3C 蛋白酶或其功能变体位于所述双功能融合蛋白酶的 N-末端部分。

[0180] 67. 实施方案 1-66 中任一项的双功能融合蛋白酶用于由较大的肽或蛋白除去 N-末端肽或蛋白的用途。

[0181] 68. 实施方案 67 的用途,其中所述较大的肽或蛋白与所述双功能融合蛋白酶在合适的反应条件下接触足够时间,以释放大多数的所述 N-末端肽。

[0182] 69. 实施方案 67-68 中任一项的用途,其中双功能融合蛋白酶与所述较大的肽或蛋白共表达,以在宿主细胞内表达期间体内释放目标蛋白。

[0183] 70. 实施方案 67-68 中任一项的用途,其中所述较大的肽或蛋白与所述双功能融合蛋白酶在从用于其表达的宿主细胞分离这两种蛋白之后接触。

[0184] 71. 实施方案 67-70 中任一项的用途,其中所述较大的肽或蛋白选自肽或蛋白,其包含选自以下的肽:GLP-1、胰高血糖素、PYY、淀粉样肽以及它们的功能变体。

[0185] 72. 实施方案 67-71 中任一项的用途,其中所述较大的肽或蛋白的大小少于 200 个氨基酸残基、少于 150 个氨基酸残基、少于 100 个残基或少于 60 个氨基酸残基。

实施例

[0186] 实施例 1

HRV14/XaaProDAP 或 XaaProDAP/HRV14 变体的质粒构建体和表达

pET 系统用于表达酶,因为该系统提供强有力的在大肠杆菌中表达蛋白的方法。在 pET

载体中,在强的噬菌体 T7 转录和释放信号控制下克隆靶基因,并通过在宿主细胞中提供 T7 RNA 聚合酶来源诱导表达。

[0187] 大肠杆菌表达质粒 (pET22b, Novagen) 编码包含 HRV14 3C 和乳酸乳球菌 XaaProDAP 序列的融合物的双功能融合蛋白酶。在一组构建体中,HRV14 3C 部分位于 XaaProDAP 序列的 N-末端,使用间插的接头 GSGSGGS (SEQ ID NO: 3) 分开这两个结构域 (表 1)。

[0188] 表 1. 编码 NH₂-HRV14 3C-XaaProDAP-COOH 融合蛋白酶的 pET22b 质粒构建体

蛋白酶	蛋白名称	融合伙伴	HRV14 3C 结构域 (N-末端)	Gly-Ser 接头	XaaProDAP 酶 (C-末端)	GSS延伸
12756	His- HRV14- XaaProDAP	SEQ ID NO: 5	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 1	GSS
12757	DsbC- HRV14- XaaProDAP	SEQ ID NO: 6	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 1	GSS
12758	RL9- HRV14- XaaProDAP	SEQ ID NO: 7	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 1	GSS
12759	NusA- HRV14- XaaProDAP	SEQ ID NO: 8	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 1	GSS
12760	His-MBP2- HRV14- XaaProDAP	SEQ ID NO: 9	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 1	GSS
12761	His-Trx- HRV14- XaaProDAP	SEQ ID NO: 10	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 1	GSS

在另一组编码融合蛋白酶的质粒中,HRV14 3C 部分置于 XaaProDAP 序列的 C-末端,其
中间插的接头 GSSGSGGS (SEQ ID NO: 4) 分开这两个结构域。

[0189] 表 2. 编码 NH₂-XaaProDAP-HRV14 3C-COOH 融合蛋白酶的 pET22b 质粒构建体

蛋白酶	蛋白名称	融合伙伴	Gly-Ser 接头	XaaProDAP酶 (C-末端)	GS 接头	HRV14 3C结构域 (N-末端)
12768	His- XaaProDAP- HRV14	SEQ ID NO: 5	GS	SEQ ID NO: 1	SEQ ID NO: 4	SEQ ID NO: 2
12769	DsbC-His- XaaProDAP- HRV14 _{3C}	SEQ ID NO: 6	GS	SEQ ID NO: 1	SEQ ID NO: 4	SEQ ID NO: 2
12770	RL9-His- XaaProDAP- HRV14	SEQ ID NO: 7	GS	SEQ ID NO: 1	SEQ ID NO: 4	SEQ ID NO: 2
12771	NusA-His- XaaProDAP- HRV14	SEQ ID NO: 8	GS	SEQ ID NO: 1	SEQ ID NO: 4	SEQ ID NO: 2
12772	MBP2- XaaProDAP- HRV14	SEQ ID NO: 9	GS	SEQ ID NO: 1	SEQ ID NO: 4	SEQ ID NO: 2
12773	His-Trx- XaaProDAP- HRV14	SEQ ID NO: 10	GS	SEQ ID NO: 1	SEQ ID NO: 4	SEQ ID NO: 2

增强表达、纯化或溶解性的融合伙伴置于两种双功能蛋白酶变体的 N- 末端。设计融合伙伴,以包含 His6 标签(位于融合伙伴序列的 N- 或 C- 末端)编码柔性 Gly-Ser- 富的接头的序列,包含在双功能蛋白酶序列的 N- 末端附近引入的甲型肝炎病毒 3C 蛋白酶(HAV)切割位点,其具有序列 GGSSGSGSELRTQS (SEQ ID NO: 22),以便视需要酶分离融合伙伴与蛋白酶部分。

[0190] 对表 1 和 2 中描述的编码融合蛋白酶的基因片段进行密码子优化,以用于在大肠杆菌中表达,并通过基因合成(GenScript)制备。利用本领域技术人员已知的标准克隆技术(获自 GenScript),通过将合成基因片段插入 pET22b 载体来产生表 1 和 2 中详述的质粒构建体。

[0191] *通过小规模表达和纯化来评价融合蛋白酶变体*

将表达质粒转化到大肠杆菌 BL21 (DE3) (Novagen) 中,并小规模表达。

[0192] 根据制造商,利用基于 42°C 的热休克 (Heat Shock) 的程序,用质粒转化大肠杆菌 BL21 (DE3)。将转化的细胞铺板至 LB 琼脂板上,并在 37°C 与 10 mg/L 氨苄青霉素温育过夜。在 30°C 制备含有各转化体的含有 0.5% 葡萄糖和 50 mg/L 羧苄西林的过夜 Terrific broth (TB) 培养物,并使用 Glas-Col 振荡仪 (Glas-Col) 在 700 rpm 振荡。将各转化体的 20 μL 过夜培养物用于接种 96 深孔板 (2 ml) 中的含有 50 mg/L 羧苄西林的 0.95 μL TB 培养基,并使转化体在 700 rpm 下繁殖过夜。在 37°C 温育表达培养物,直到 OD600 达到 1.5。然后将培养物冷却至 20°C,利用 0.3 mM IPTG 进行蛋白诱导过夜。通过在 1800xG 离心,收获含

有表达的蛋白的沉淀物。

[0193] 纯化筛选:利用 IMAC 树脂进行小规模纯化,以评价蛋白酶的组表达和纯化潜力以及完整性。简而言之,将 250 μ L 裂解缓冲液(50 mM NaPO₄, 300 mM NaCl, 10 mM 咪唑, 10mg/ml 溶菌酶, 250U/ μ L Benzoase 和 10% DDM(十二烷基麦芽糖苷))加入各沉淀物,利用冷冻/解冻循环来裂解细胞。离心除去碎片,将上清液过滤(0.45 μ m)并转移至 1.2 μ m 滤板上,其含有加了 Ni²⁺ 的 Sepharose Fast Flow(由在 20% EtOH 中洗涤 30 μ L 50% 浆液来制备)(GE Healthcare)。通过在 400 rpm 与树脂一起振荡,将上清液温育 20 分钟以结合蛋白,通过 100xg 温和离心 1 分钟除去溶质。通过轻轻混合,用 50mM 磷酸钠, 300 mM NaCl, 30 mM 咪唑, pH 7.5 洗涤树脂,并离心干燥。为了洗脱蛋白,将 40 μ L 洗涤缓冲液(50 mM 磷酸钠, 300 mM NaCl, 300 mM 咪唑)加入树脂,在 400 rpm 振荡 10 分钟,并收集含有部分纯化的酶的洗脱物。

[0194] 通过 SDS-PAGE 分析来自融合蛋白酶变体表达的沉淀物的全裂解物。对于表 1 或 2 中描述的所有融合蛋白酶变体,均未能观察到显著量的全长蛋白。但是,对于若干融合蛋白酶变体,观察到明显的大小不同的条带。IMAC 纯化的样品的 SDS-PAGE 分析与这些观察结果一致,因为其也没有显示产生全长蛋白酶,更确切地,显示产生较小大小的条带。观察结果表明,在表达期间和/或捕获到 IMAC 树脂上后,融合蛋白酶被截短或降解。由于对若干融合蛋白酶变体观察到明显不同的条带,且表达水平似乎显著基于凝胶条带强度,所以,更确切地说,不存在全长蛋白是因为融合蛋白酶中特定位置的非所期的水解,导致融合蛋白酶的显著截短。

[0195] 融合蛋白酶的 LC-MS 分析

通过质谱检测能够解释截短形式的融合蛋白酶的被观察到存在的可能的切割位点,使用 Maxis Impact 超高分辨飞行时间(UHR-TOF)质谱仪(Bruker Daltonics),其装备有 Dionex UltiMate3000™液相色谱仪(Dionex),根据制造商的说明书,允许以一般设定在 UV215 nm 进行二极管阵列测量。在 1.7 μ m 孔径的 Waters Aquity BEH300 C4 反相 1.0 X 100 mm 柱上分离酶,利用的柱温为 45°C,流速为 0.2 ml/分钟。所用的溶剂如下:

溶剂 A:0.1% 甲酸 H₂O 溶液

溶剂 B:99.9% MeCN, 0.1% 甲酸(v/v)

利用下列梯度进行液相色谱以分离酶消化物。

[0196] 时间(分钟)	%A	%B
0	90	10
2	90	10
10	10	90
11	10	90
12	90	10
13	90	10
14	50	50

将记录的质谱去卷积,并用 Bruker Compass 数据分析软件版本 4.1(Bruker Daltonics)进行分析,根据制造商说明书,覆盖 10.000 Da-140.000 Da 的质量范围和分辨率(>10.000)。平行评价 UV215 nm 色谱图和总离子计数(TIC)色谱图,以确保所获的 MS 数

据和肽的 UV215 nm 迹线之间一致。示出的实验确定的质量是指平均同位素质量,且质谱数据以好于 200 ppm 的质量精度获得。

[0197] 当分析蛋白酶 12756 时,检测到 22241.54 Da 的质量。该质量对应于 His6 融合伙伴 (SEQ ID NO: 5) 和 HRV14 3C 结构域 (SEQ ID NO:2) 的质量 (计算质量 22242.27 Da)。因此,在连接 HRV14 3C 结构域 (SEQ ID NO:2) 的 C-末端和接头 (SEQ ID NO: 3) 的 N-末端的 Gln/Gly 之间存在切割位点。这表示,3C 蛋白酶能够将其自身从 HRV14 3C 与 XaaProDAP 的 N-末端融合的融合蛋白酶中切除,切除方式类似于已报道的从其天然病毒多蛋白中切除的方式。This was also observed 对于融合蛋白酶 12757、12758、12760 和 12761,也观察到了这种现象,其中大小变化对应于所用的 N-末端融合伙伴的大小的差异。

[0198] 实施例 2

除去 NH₂-HRV14-XaaProDAP-COOH 融合蛋白酶的 HRV14 结构域中的 C-末端 Q182

对于表 1 所示的融合蛋白酶变体,观察到片段的大小常常对应于融合伙伴加 HRV14 3C 结构域 序列。为了除去该切割可能性,设计了新的接头以替换开始的 GS 接头 (SEQ ID NO:3),其位于包含 His6、RL9 或 Trx 融合伙伴的融合蛋白酶中 HRV14 3C 和 XaaProDAP 结构域之间 (表 1, 实施例 1)。连接 HRV14 酶和 SEQ ID NO: 3 起点的 Gln/Gly 切割位点替换为 Ser-Gly。因此,除去了 HRV14 3C 蛋白酶结构域中的最后一个氨基酸 (Gln182),得到具有下述序列的 des182-HRV14 3C :GPNTEFALSLLRKNIMTITTSKGEFTGLGIHDRVCVIPHTAQP GDD VLVNGQKIRVKDKYKLVDPENINLELTVLTLDRNEKFRDIRGFISEDLEGVDATLVVHSNNFTNTILEVGPVTMAGL INLSSTPTNRMIRYDYATKTGQC GGVL CATGKIFGIHVGGNGRQGFSAQLKKQYFVEK (SEQ ID NO: 11),且除去了接头 (SEQ ID NO:3) 最开始的 Gly,因为该位点表示 3C 蛋白酶的切割位点。但是,HRV14 结构域 (SEQ ID NO:11) 和 XaaProDAP 结构域之间的接头替换为 SGSGGSGSGS (SEQ ID NO:12)。新的融合蛋白酶变体示于表 3:

表 3. 编码 des182HRV14 3C-XaaProDAP 融合蛋白酶的 pET22b 质粒构建体

蛋白酶	蛋白名称	融合伙伴	HRV14 3C 结构域 (N-末端)	Gly-Ser 接头	XaaProDAP 酶 (C-末端)	GSS 延伸
20177	His-des182HRV14 XaaProDAP	SEQ ID NO: 5	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 1	GSS
20397	RL9-des182HRV14 XaaProDAP	SEQ ID NO: 7	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 1	GSS
20400	His-Trx-des182HRV14 XaaProDAP	SEQ ID NO: 10	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 1	GSS

按实施例 1 所述,进行这些构建体的小规模表达和纯化。来自 IMAC 纯化的样品的 SDS-PAGE 显示,此时对于这三种融合蛋白酶变体出现了两个清楚可见且主要的大约 50-60 kDa 的条带,表明全长蛋白酶被切成两个片段。按实施例 1 中所述,进行 LC-MS 分析以定位切割位点。蛋白酶 20177 的分析显示,该融合蛋白酶变体被切成两个大的条带,各具有

51091.27 Da 和 59773.49 Da 的质量。这些质量证实,另一个切割位点出现在 XaaProDAP 序列 (SEQ ID NO:1) 的 Gln241 和 Gly242 之间,因为测定的质量与这些片段的计算质量一致,如由蛋白酶 20177 氨基酸序列所推导的 (分别为 51092.15 Da 和 59772.43 Da)。通过分析表 3 中所示所有三个构建体的去卷积谱,清楚地观察到完全相同的切割位点,由此表明,该位点是高度敏感的,而与所用的 N-末端融合伙伴无关。在评价可用的 3D 结构 (Rigolet 等, Structure, 10, pp 1384-1394) 时,可以确定,切割位点存在于乳酸乳球菌 XaaProDAP 的催化结构域中连接两个小的 α -螺旋的非常大的环 (跨越大约氨基酸残基 223 - 270) 的中间。该环高度暴露,并因此对切割敏感, Q/G 序列表明 3C 蛋白酶本身负责该切割。另一个较不主要的非所需的切割位点被观察到位于融合伙伴的 C-末端的 HAV 蛋白酶 (ELRTQ / S) 切割位点中的 Gln/Ser 位置,也可通过分析 IMAC 纯化的样品来检测。

[0199] 实施例 3A

设计全长双功能 NH₂-HRV14-XaaProDAP-COOH 蛋白酶

为了除去实施例 2 中 XaaProDAP 序列的 Gln241 和 Gly242 之间观察到的切割位点,用 Glu241 和 Thr242 取代这两个氨基酸。选择 Glu241-Thr242 取代作为 Gln241-Gly242 的替代,因为其存在作为基于来自乳酸乳球菌的不同分离物的 XaaProDAP 直向同源物的同源性检索的天然氨基酸变化。由于非所需的切割还存在于融合伙伴的 C-末端的 HAV 位点,所以将 HAV 位点替换为小的含 GS 序列。这些融合伙伴具有序列 MHHHHHGGSSGSGSGSGSGS (SEQ ID NO: 13)、MKVILLRDVPGKIGKKGEIKEVSDGYARNYLIPRGFAKEYTEGLERAIKHEKEIEKRKKERERESE KILKELKKRTHVVKV KAGEGKIFGAVTAATVAEEISKTTGLKLDKRWFKLDKPIKELGEYSLEVS LPPGVKDTIKI RVEREEGSGSGHHHHHGGSSGSGSGSGSGS (SEQ ID NO:14) 和 MHHHHHGGSGSGSDKI IHLTDDSFDTDLV LKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGAL SKGQLKEFLDANLAGGSSGSGSGSGSGS (SEQ ID NO: 15)。

[0200] 获得包含 Q241E、G242T 取代且在 HRV14 3C 结构域前方由接头除去了 HAV 位点的质粒构建体 (Genscript)。设计和测试的示于表 4。

[0201] 按实施例 1 中所述,进行新的融合蛋白酶 构建体的小规模表达和 IMAC 纯化。由 SDS-PAGE 分析观察到, Q241E 和 G242T 取代明显地防止了融合蛋白酶被切成两部分。对于所有三个构建体,均观察到非常强的大约 100-120 kDa 的凝胶条带,表明 Q241E、G242T 取代导致产生可溶且完整的全长融合蛋白酶,其包含 HRV14 3C 结构域和 XaaProDAP 结构域二者。在该实验中,除去 ELRTQ 位点 (通过用 GSGSG 取代) 的益处较不显著。

[0202] 按实施例 1 中所述,对表 4 中的融合蛋白酶变体进行 LC-MS,并证实了来自 SDS-PAGE 的观察结果。蛋白酶 20986、20988 和 20990 的测定的质量分别为 110604.97 Da、127867.76 Da 和 122607.21 Da,它们分别与计算质量 110605.18 Da、127867.23 Da 和 122605.91 Da 一致。因此,表 4 中的修饰融合蛋白酶未被显著截短或降解,因为主要检测到的质量对应于全长融合蛋白酶的 计算质量。

[0203] 表 4. 编码 NH₂-des182HRV14 3C-XaaProDAP(Q241E, G242T)-COOH 融合酶的 pET22b 质粒构建体

蛋白酶	蛋白名称	融合伙伴	HRV14 3C结构域 (N-末端)	Gly-Ser接头	XaaProDAP酶 (C-末端)	GSS延伸
20986	His-des182HRV14-XaaProDAP (Q241E,G242T)	SEQ ID NO:13	SEQ ID NO:11	SEQ ID NO:12	SEQ ID NO:1 (Q241E,G242T)	GSS
20988	RL9-des182HRV14-XaaProDAP (Q241E,G242T)	SEQ ID NO:14	SEQ ID NO:11	SEQ ID NO:12	SEQ ID NO:1 (Q241E,G242T)	GSS
20990	His-Trx-des182HRV14-XaaProDAP (Q241E,G242T)	SEQ ID NO:15	SEQ ID NO:11	SEQ ID NO:12	SEQ ID NO:1 (Q241E,G242T)	GSS

实施例 3B

设计全长双功能 NH₂-XaaProDAP-HRV14-COOH 蛋白酶

使用实施例 1 and 3A 中所述的通用设计、克隆和表达程序,我们还评价了是否能够获得在 C-末端包含 HRV14 3C 的功能性可溶融合蛋白酶。评价 3 种融合蛋白质的表达,其包含 C-末端 HRV14 3C 结构域和 N-末端 XaaProDAP (Q241E, G242T),并使用之前说明的 3 个不同的 N-末端标签 (His6,RL9,Trx)。将 HRV14 3C 结构域位于 C-末端的所有 3 个构建体表达为不溶性蛋白,如通过 SDS-PAGE 测定的未诱导的、诱导的、可溶的和不可溶的级分(未示出详细数据)。这证实,在 N-末端包含 HRV14 3C 蛋白酶和在 C-末端包含乳酸乳球菌 XaaProDAP 的融合蛋白酶变体出乎意料地具有更优化的折叠动力学,其产生可溶的稳定融合蛋白酶、更易于生产且不需要任何成本禁止的折叠步骤。总之,蛋白质设计的一些规格使得可以生产包含 HRV14 3C 和 XaaProDAP 蛋白酶的完整融合蛋白酶。

[0204] 实施例 4

放大表达和纯化 NH₂-His-des182HRV14-LLXaaProDAP (Q241E, G242T)-COOH (蛋白 20986)

为了制备更大量的全长融合蛋白酶,放大蛋白酶 20986 用于进一步的活性测试。

[0205] 在含有 50 mg/L 羧苄西林和 0.5% 葡萄糖的 50 ml Terrific Broth 培养基中,通过在 37°C 以 100rpm 振荡 (Multitron Standard shaker, 50mm 振幅, Infors HT),使含有编码蛋白酶 20986 的 pET22b 质粒的 BL21 (DE3) 转化体(来自甘油贮备物)繁殖过夜。次日,使用 7.5 ml 过夜培养物接种 2L 振荡烧瓶中含 50 mg/L 羧苄西林的 750 ml TB 培养基,随后在 37°C 以 100rpm 温育培养物。当 OD₆₀₀ 达到约 1.5 时,将培养物冷却至 20°C 持续 30 分钟,然后加入 0.3 mM IPTG 以诱导蛋白。在 20°C、100 rpm 进行诱导过夜,通过 4000 x g 离心 10 分钟收获细胞。将沉淀细胞冷冻备用。

[0206] 纯化 His-des182HRV14-LLXaaProDAP (Q241E, G242T) (蛋白 20986)

为了获得纯化的双功能融合蛋白酶以进一步分析,进行两个连续的纯化步骤,从而纯化蛋白酶 20986。

[0207] 将 14.7 g 细胞沉淀物重悬于 100 ml 裂解缓冲液,其含有 50mM 磷酸钠 pH 7.5 和

3 μ L benzonase。在细胞匀浆机中于 1.4 kBar 破坏细胞一个循环,将细胞碎片于 18,000g 离心 20 分钟以沉淀。然后对上清液进行无菌过滤 (0.45 微米)。使用 AKTAExpress (GE Healthcare) 进行两个连续纯化步骤,完成蛋白酶 20986 的纯化。在捕获步骤中,在 2X1 ml HisTrap crude 柱 (GE Healthcare) 上纯化来自 100 ml 样品施加物的酶,其中流速为 0.8 ml/分钟,并使用下述缓冲液:

缓冲液 A :50 mM 磷酸钠, 300 mM NaCl, 10 mM 咪唑 pH 7.5

缓冲液 B :50 mM 磷酸钠, 300 mM NaCl, 300 mM 咪唑 pH 7.5

缓冲液 C :50 mM 磷酸钠, 300 mM NaCl, 30 mM 咪唑 pH 7.5

首先,用 10 倍柱体积的缓冲液平衡柱。在加了施加物后,通过使用 7 倍柱体积的缓冲液 C 洗涤来除去未结合的蛋白。使用 5 倍柱体积的 0-100 % 缓冲液 B 分步洗脱,以洗脱蛋白酶 20986,将收集的峰贮于环中并加到 120 ml HiLoad S200 16/600 (GE-Healthcare) 凝胶过滤柱。进行大小分离,其中流速为 1.2 ml/分钟,使用 1X PBS 缓冲液 (磷酸缓冲盐水, pH 7.4,组成 :8.05 mM Na₂HPO₄·2H₂O, 1,96 mM KH₂PO₄, 140 mM NaCl, pH 7.4)。通过 SDS-PAGE 分析收集的主要峰级分,观察到约 100 kDa 预期大小的清晰条带。合并含有最多量的蛋白酶的级分,使用 UV280 measurements (NanoDrop, ThermoScientific) 测量的浓度为 1,6 mg/ml。通过 SDS-PAGE (图 1) 和 HPLC 分析来判断,估计纯度高于 90%。

[0208] 实施例 5

含有碱性标签的模型融合蛋白的质粒构建体和表达

为了测试双功能融合蛋白酶是否能够用于除去 N-末端标签,制备了三种不同的模型融合蛋白作为蛋白底物。使用 W02008/043847 中描述过的包含来自海栖热袍菌 (*T. maritima*) 的核糖体蛋白 L27 的碱性标签作为融合伙伴,其具有下述序列 :MAHKKSGGVAKNGR DSLPKYLGVKVGDGQIVKAGNILVRQRGTRFYPGKNVGMGRDFTLFALKDGRVKFETKNNKKYVSVYEE (SEQ ID NO: 16)。设计融合蛋白,使得 RL27 融合伙伴可通过 HRV14 3C 酶除去,且剩余 GP 序列可通过 XaaProDAP 除去。

[0209] 使用含有 HRV14 切割位点的柔性接头连接碱性标签与模型肽序列,该接头具有序列 SSSGGSEVLFQGP (SEQ ID NO: 17)。所用的模型肽序列为人肽 YY 3-36 (PYY(3-36))、胰高血糖素和胰高血糖素样肽 1 (7-37, K34R) (GLP-1 (7-37, K34R)), 其各自具有下述序列:

PYY(3-36) :IKPEAPGEDASPEELNRYASLRHYLNLVTRQRY (SEQ ID NO: 18)

胰高血糖素 :HSQGTFTSDYSKYLDSSRAQDFVQWLMNT (SEQ ID NO: 19)

GLP-1 (7-37, K34R) :HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG (SEQ ID NO: 20)。

[0210] 制备大肠杆菌表达质粒 (pET22b, Novagen),使它们编码表 5 中详述的三个融合蛋白。

[0211] 表 5. 使用 pET22b 载体的质粒构建体编码的模型融合蛋白

蛋白名称	计算分子量 (无 Met)	RL27	HRV14接头	肽
RL27_EVLFQGP_PYY(3-36)	14354.3 Da	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 18
RL27_EVLFQGP_胰高血糖素	13787.1 Da	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 19
RL27_EVLFQGP_GLP1(7-37,K34R)	13688.1 Da	SEQ ID NO: 16	SEQ ID NO: 17	SEQ ID NO: 20

通过基因合成来产生针对大肠杆菌进行了密码子优化且跨越整个融合蛋白的基因片段,并使用标准克隆技术(获自 GenScript)将其连接到 pET22b 载体的克隆位点中。

[0212] 表达模型融合蛋白

基本上按实施例 4 中对于蛋白酶 20986 所述,进行 RL27_EVLFQGP_PYY(3-36) 的表达。简而言之,如下表达 RL27_EVLFQGP_胰高血糖素和 RL27_EVLFQGP_GLP-1(7-37, K34R):用质料转化大肠杆菌 BL21 (DE3),铺板到含有 100 mg/L 氨苄青霉素的 LB 琼脂板上,将过夜培养物溶解于 10 ml LB 培养物中,并用于接种振荡烧瓶中的含有 50mg/ml 羧苄西林的 750 ml LB。将振荡烧瓶以 100 rpm 于 37°C 温育。当 OD600 达到 0.4 时,通过加入 0.3 mM IPTG 来诱导蛋白表达,并通过在 37°C 温育 3 小时后离心来收获细胞。

[0213] 纯化模型融合蛋白

简而言之,通过阳离子交换色谱从得自细胞破坏的上清液捕获融合蛋白,这基本上按照现有描述 (W02008/043847),使用 AKTA Express 上的 SP FF HiTrap 5 ml (GE Healthcare) 柱,流速为 4 ml/分钟,缓冲液如下:

缓冲液 A :50mM 磷酸钠, pH 7.0

缓冲液 B :50mM 磷酸钠, 1000mM NaCl, pH 7.0。

[0214] 简而言之,在加样和洗涤步骤之后,使用缓冲液 B 从柱上洗脱融合蛋白。为了提高下述捕获的纯度,基本上按实施例 4 中所述通过凝胶过滤来纯化蛋白,不同在于使用 S75 16/600 分离柱 (GE-Healthcare)。纯化的蛋白经 SDS-PAGE 分析评价,正确的完整质量由 LC-MS 核实。UV280 用于测定融合蛋白的浓度。

[0215] 实施例 6

利用蛋白酶 20986 和作为模型蛋白底物的 RL27_EVLFQGP_PYY(3-36) 的酶反应

将浓缩物 RL27-HRV14-PYY(3-36) 调整至 1X PBS, pH 7.4 中的 0.5 mg/ml 浓度。在 22 μ l 反应体积中建立酶反应,使用 PBS, pH 7.4 作为酶反应缓冲液。分别使用 1:20 或 1:40 的酶底物摩尔比,建立蛋白酶 20986 与 RL27-EVLFQGP-PYY(3-36) 底物的温育,在 37°C 进行反应 3 小时(如表 6 所示)。在该实验中还包括 W02008/043847 中描述的纯化 HRV14 3C 蛋白酶变体,其具有 N-末端标签(核糖体 L9,来自海栖热袍菌)。该蛋白酶称为 RL9-HRV14 3C,以与蛋白酶 20986 相同的摩尔比使用,但是仅具有 HRV14 3C 活性。RL9-HRV14 3C 具有下述序列:MKVILLRDVPKIGKKGEIKEVSDGYARNYLIPRGFAKEYTEGLERAIKHEKEIEKRKKERERESEKILK ELKKRTHVVKV KAGEGGKIFGAVTAATVAEEISKTTGLKLDKRWFKLDKPIKELGEYSLEVSLPGGVKDTIKIRVER EESSGSSGSSGSSGNTEFALSLLRKNIMTITTSKGFTGLGIHDRVCVIPHTAQPGDDVLVNGQKIRVKDKYKLV

DPENINLELTVLTLDRNEKFRDIRGFISEDLEGVDATLVVHSNNFTNTILEVGPVTMAGLINLSSTPTNRMIRYDYA
TKTGQCGVLCATGKIFGIHVGGNGRQGFSAQLKKQYFVEKQ (SEQ ID NO: 21)。作为阴性对照,还在
不含蛋白酶的反应缓冲液中温育 RL27-HRV14-PYY(3-36) 底物。通过在 LC-MS 分析之前加
入 >0.5 M AcOH, 终止酶反应。

[0216] 使用 RL27_HRV14_PYY(3-36) 作为融合蛋白模型底物的蛋白酶 20986 结果

基本上按实施例 1 中所述进行酶反应的 LC-MS 分析, 区别仅在于具有 1.7 μm 孔径的 C18
Aquity BEH300 C4 反相 1.0 X 100 mm 柱用于确保充分分离和分辨所评价的较小肽。根据
制造商说明书, 将仪器调整至用于质量范围 (2000-17000 Da) 和分辨率 (>20,000) 的设定。
平行评价 UV215 nm 色谱图和总离子计数 (TIC) 色谱图, 以确保所得 MS 数据和肽的 UV215
nm 迹线之间的一致。下述实施例中示出的实验测定的质量是指最富余的质量, 例如基于所
检测的蛋白的同位素的天然丰度而具有最具代表性的同位素分布的分子质量。下面, 利用
小于 100 ppm 的质量精度获得质谱数据。

[0217] 去卷积质谱的分析显示, RL27_EVLFGQP_PYY(3-36) 融合蛋白 (对照不含酶) 质量
为 14354.17 Da。这与不含起始甲硫氨酸的融合蛋白的计算质量 (14354.5 Da) 一致。

[0218] 不同反应的结果示于表 6。

[0219] 表 6: 酶反应, 使用来自实施例 4 的蛋白酶 20986 和作为底物的 RL27_EVLFGQP_
PYY(3-36), 均在 37°C 下温育 3 小时。示出反应 1-4 的去卷积质谱中检测到的实验确定的主
要峰。

反应编号	酶	摩尔比	检测到的主要峰	测定的分子量 (道尔顿)	计算质量 (道尔顿)	对应于
反应 1	蛋白酶 20986	1:20	峰 #1	4049.98	4050.1	PYY(3-36) (SEQ ID NO:18)
			峰 #2	10168.21	10168.4	RL27 标签
反应 2	蛋白酶 20986	1:40	峰 #1	4050.06	4050.1	PYY(3-36) (SEQ ID NO:18)
			峰 #3	4204	4204.1	GP-PYY(3-36)
			峰 #2	10168.19	10168.4	RL27 标签
反应 3	RL9-HRV14 3C	1:20	峰 #1	4204.05	4204.1	GP-PYY(3-36)
			峰 #2	10168.19	10168.4	RL27 标签
反应 4	RL9-HRV14 3C	1:40	峰 #1	4204.08	4204.1	GP-PYY(3-36)
			峰 #2	10168.23	10168.4	RL27 标签

[0220] 反应 1 显示, 融合蛋白的完全加工在 1:20 的酶底物摩尔比且 37°C 温育小时的酶处
理之后获得 (图 2)。观察到的主要检测质量是 4049.9 Da, 其对应于成熟 PYY(3-36) (峰
#1) 和释放的标签 (峰 #2) 的质量。未观察到剩余的融合蛋白, 但是观察到强度小于峰 #1 的
10% 的峰, 其对应于 GP-PYY(3-36)。反应 2 显示, 1:40 酶底物比导致约一半的 GP-PYY(3-36)
加工成成熟 PYY(3-36) (图 3)。反应 3 (图 4) 和 (图 5) 显示, 反应 1 和 2 中观察到的由
GP-PYY(3-36) 除去 Gly-Pro 是蛋白酶 20986 的 XaaProDAP 部分特异性的, 因为仅含有 HRV14

3C 结构域的 RL9-HRV14 3C 蛋白酶只能释放 GP-PYY(3-36)。

[0221] 实验显示,可通过双功能融合蛋白酶释放完全成熟的 PYY(3-36) 肽(4050 Da),由此实现了本发明的构想。

[0222] 实施例 7

包含来自其它物种的备选 3C 和 XaaProDAP 结构域的全长双功能融合蛋白酶设计

为了证实其它 3C 蛋白酶和 XaaProDAP 酶可融合以获得性质与对蛋白酶 20986 所观察到的相同的融合蛋白酶,使用来自人柯萨奇病毒 B3 的 3C 蛋白酶序列(CVB3 3C)或来自猪链球菌(*Streptococcus suis*)的 XaaProDAP(猪链球菌 XaaProDAP)替换 HRV14 3C 和乳酸乳球菌 XaaProDAP(LLXaaProDAP)序列,并产生新的融合蛋白酶变体。与来自人鼻病毒 14 3C 的 3C 蛋白酶序列一样,人柯萨奇病毒 B3 3C 蛋白酶序列也包含 C-末端 Q,其缺失以获得具有下述序列的 CVB3 3C(des183):

GPAPFEFVAMMKRNSSTVKTEYGEFTMLGIYDRWAVLPRHAKPGPTILMNDQEVGVLDKELVDKDGNTLE
LTLKLNKRNKFRDIRGFLAKEEVEVNEAVLAINTSKFPNMYIPVGVQVTEYGFNLGGTPTKRMLMYNFPTRAGQC
GGVLMSTGKVLGIHVGGNGHQGFSALLKHVFNDE(SEQ ID NO: 23)。

[0223] 在猪链球菌 XaaProDAP 序列的 Q212-G213 位置观察到 QG 位点,其接近对乳酸乳球菌序列所测定的 3C 切割位点(Q241-G242)。引入 Glu212-Thr213 取代以防止任何潜在的 3C 切割,由此得到下述序列:MRFNQFSFIKKETSVYLQELDTLGFQLIPDASSKTNLETFRKCHFLTAN
TDFALSNMIAEWDTDLLTFFQSDRELTDQIFVQVAFQLLGFVPGMDYTDVMDVFVEKSNFPIVYGDIDNLYQLLNTR
TKSGNTLIDQLVSDDLIPEDNHYHFFNGKSMATFSTKNLIREVVYVETPVDTAGTGQTDIVKLSILRPHFDGKIPAV
ITNSPYHETVNDVASKALHKMEGELAEKQVGTIQVKQASITKLDLDQRNLPVSPATEKLGHITSYSLNDYFLARGF
ASLHVSGVGTGSGTYMTSGDYQQVEGYKAVIDWLNDRTKAYTDHTRSLEVKADWANGKVATTGLSYLGTMSNALAT
TGVDGLEVIIEAGISSWYDYRENGLVTSPPGGYDGLDLSLTALYSKSLQAGDFLRNKAAYEKGLAERAALDRT
SGDYNQYWHDRNYLLHADRVKCEVVFTHGSQDWNVQPIHVWNMFHALPSHIKKHLFFHNGAHVYMNWQSIDFRESM
NALLSQKLLGYENNYQLPTVIWQDNSGEQTWTTLDTFGGENETVLPVLTGSGQTVANQYTQEDFERYGKSYSFAHQDL
YAGKANQISIELPVTEGLLLNGQVTLKLRVASSVAKGLLSAQLLDKGNKKRLAPIPAPKARLSLDNGRYHAQENLVE
LPYVEMPQRLVTKGFMNLRNRTDMLTVEEVVPGQWMNLTKLQPTIYQLKKGDVLELILYTTDFECTVRDNSQWQIH
LDLSQSQLILPH(SEQ ID NO: 24)。

[0224] 设计三个新的融合蛋白酶变体,其包含新的 3C 和 XaaProDAP 直向同源物,并使用相同的 His6 融合伙伴(SEQ ID NO:13)和相同的实施例 3 所示的间插接头(SEQ ID NO:12)。蛋白酶 28994 包含实施例 3A 中对蛋白酶 20986 所说明的乳酸乳球菌 XaaProDAP 序列,但是 N-末端 HRV14 3C 结构域替换为来自人柯萨奇病毒 B3 的 3C 结构域(CVB3 3C)。蛋白酶 28996 包含 N-末端的对蛋白酶 20986 所说明的 HRV14 3C 序列和 C-末端的猪链球菌 XaaProDAP 序列。蛋白酶 28997 是全新的融合蛋白酶,其中两个结构域均替换为 3C 和 XaaProDAP 蛋白酶的直向同源物,由此蛋白酶包含 N-末端的 CVB3 3C 序列和 C-末端的猪链球菌 XaaProDAP 序列。

[0225] 由 GenScript 获得质粒构建体,其使用 pET22b 载体骨架且包含新的融合蛋白酶。编码设计的融合蛋白酶变体的序列组合示于表 7。

[0226] 表 7. 编码融合蛋白酶变体的 pET22b 质粒构建体,各变体包含 N-末端 HRV14 3C 或 CVB3 3C 和 C-末端乳酸乳球菌 XaaProDAP(Q241E, G242T) 或猪链球菌

XaaProDAP (Q212E, G213T) 的组合

蛋白酶	蛋白名称	融合伙伴	HRV14 3C 结构域 (N-末端)	Gly-Ser 接头	XaaProDAP 酶 (C-末端)	GSS 延伸
28994	His-CVB3_3C-LLXaaProDAP-(Q241E,G242T)	SEQ ID NO:13	SEQ ID NO:23	SEQ ID NO:12	SEQ ID NO: 1 (Q241E,G242T)	GSS
28996	His-HRV14_3C-SSXaaProDAP-(Q212E,G213T)	SEQ ID NO:13	SEQ ID NO:11	SEQ ID NO:12	SEQ ID NO: 24 (Q212E,G213T)	
28997	His-CVB3_3C-SSXaaProDAP-(Q212E,G213T)	SEQ ID NO:13	SEQ ID NO:23	SEQ ID NO:12	SEQ ID NO: 24 (Q212E,G213T)	

按实施例 1 中所述,进行新的融合蛋白酶构建体的小规模表达和 IMAC 纯化,显示所有三个新的蛋白酶得到可溶的完整融合蛋白酶,其包含新的 3C 和 XaaProDAP 直向同源物序列。

[0227] 按实施例 1 中所述,通过 IMAC 纯化的融合蛋白酶变体的 LC-MS 分析来测定完整质量,结果证实了 SDS-PAGE 的观察结果。蛋白酶 28994、28996 和 28997 的测定质量分别为 107797.8 Da、107687.2 Da 和 107964.2 Da,其分别与计算质量 107798.1 Da、107687.4 Da 和 107964.8 Da 高度一致。因此,如对蛋白酶 20986 所观察到的,新的蛋白酶未显著被截短或降解,因为检测到的主要峰质量对应于全长融合蛋白酶的計算质量。因此,所有的蛋白酶 20986、28994、28996 和 28997 均不具有能够损害两种蛋白水解活性组成中的至少一种的自我切割活性。总之,使用具有明显不同的氨基酸序列的小 RNA 病毒 3C 和 XaaProDAP 酶的其他直向同源物,进一步证实了本发明制备功能性 3C/XaaProDAP 融合蛋白酶的构想。

[0228] 实施例 8

放大表达和纯化包含新的 3C 和 XaaProDAP 结构域的蛋白酶 28994、28996 和 28997

按实施例 4 中所述,使用 BL21 (DE3) 作为表达宿主表达蛋白酶 28994、28996 和 28997。基本上按实施例 4 中所述进行纯化,利用 IMAC 步骤捕获,然后是凝胶过滤步骤。通过实施例 4 中所述的两步方案,成功纯化了所有的蛋白酶 28994、28996 和 28997。按 SDS-PAGE 凝胶检查和 LC-MS 分析期间来自 RP 分离 HPLC 的 UV215 nm 概况评价来判断,估计纯度至少 90%。按实施例 1 中所述进行 MS 分析,显示蛋白酶 28994 的估计质量为 107797.8 Da,与预期质量 (110798.1 Da, 平均同位素质量) 非常一致。蛋白酶 28996 的质量为 107686.9 Da,与预期质量 (107687.4 Da, 平均同位素质量) 非常一致,且蛋白酶 28997 的测定质量为 107964.8 Da,与预期质量 (107964.8, 平均同位素质量) 一致。进行 UV280 吸光度测量,以测定融合蛋白浓度 (NanoDrop)。

[0229] 实施例 9

利用蛋白酶 20986、28994、28996 和 28997 的酶反应

在 30 μ l 反应体积中建立酶反应,使用 1 X PBS, pH 7.4 作为酶反应缓冲液。用于评价切割特异性的模型蛋白底物包含融合蛋白,其被酶正确加工后应当产生人 PYY (3-36) (SEQ ID NO: 18)、野生型胰高血糖素 (SEQ ID NO: 19) 和 GLP-1 (7-37, K34R) (SEQ ID NO: 20)。

将模型蛋白底物的浓度调整为 0.5 mg/ml, 这使用 1xPBS, pH 7.4, 如实施例 6 中所述。在酶底物比以及酶反应持续时间和温度方面评价反应条件的变化。包括对照, 其不包含酶 (1x PBS pH 7.4) 或包含 RL9-HRV14 3C (SEQ ID NO. 21)。在实验结束时通过加入 >0.5 M AcOH 来终止反应。使用实施例 6 中所述的条件和通用设定, 进行酶反应的 LC-MS 分析。

[0230] 作为模型蛋白底物的 RL27_EVLFQGP_PYY(3-36)

分别使用 1:20 或 1:100 的酶底物摩尔比, 建立蛋白酶 28994、28996 和 28997 与 RL27-EVLFQGP-PYY(3-36) 底物的温育, 在 37°C 行反应 3 小时 (如表 8 中所示)。通过 LC-MS 分析完整质量显示, 当使用 1:20 的酶底物摩尔比时, 蛋白酶 28994、28996 和 28997 在 37°C 温育 3 小时后能够将 RL27_EVLFQGP_PYY(3-36) 彻底加工成成熟的 PYY(3-36) (SEQ ID NO: 18) (如实施例 6 中对 20986 所观察到的)。在 1:100 的酶底物比时, 检测到少量的 PYY(3-36) 以及 GP-PYY(3-36) (反应 6、8 和 10), 反应不总是彻底的, 因为检测到了完整融合蛋白。在 1:100 的比时, 蛋白酶 28996 和 28997 提供最有效的切割, 具有最低量的剩余 GP-PYY(3-36), 相对强度分别为成熟 PYY(3-36) 峰强度的约 25% 或约 50%。含有 RL9-HRV14 3C (SEQ ID NO. 21) 的对照仅得到 GP_PYY(3-36) 峰, 显示 XaaProDAP 结构域负责完成反应以产生 PYY(3-36) 的天然 N- 末端, 而不添加酶仅产生未加工的融合蛋白。实验显示, 组合来自人鼻病毒或人柯萨奇病毒的 3C 蛋白酶与来自乳酸乳球菌或猪链球菌的 XaaProDAP 的不同融合蛋白酶变体可成功用于将 RL27_EVLFQGP_PYY(3-36) 加工成成熟 PYY(3-36), 其中 Ile 是正确的 N- 末端氨基酸残基。

[0231] 表 8: 酶反应, 使用来自实施例 8 的蛋白酶 28994、28996 和 28997 和作为底物的 L27_EVLFQGP_PYY(3-36), 均 37°C 下温育 3 小时。示出反应 5-10 的去卷积质谱中检测到的实验测定的主要峰。

反应编号	酶	摩尔比	检测到的主要峰	测定质量 (道尔顿)	计算质量 (道尔顿)	对应于
反应 5	蛋白酶 28994	1:20	峰 #1	4050.09	4050.1	PYY(3-36) (SEQ ID NO: 18)
			峰 #2	10168.47	10168.4	RL27 标签
反应 6	蛋白酶 28994	1:100	峰 #1	4050.07	4050.1	PYY(3-36) (SEQ ID NO: 18)
			峰 #2	10168.42	10168.4	RL27 标签

			峰#3	4204.14	4204.1	GP-PYY(3-36)
			峰#4	14354.54	14354.5	RL27_EVLFGQP_PYY(3-36)
反应7	蛋白酶 28996	1:20	峰 #1	4050.09	4050.1	PYY(3-36) (SEQ ID NO: 18)
			峰 #2	10168.46	10168.4	RL27标签
反应8	蛋白酶 28996	1:100	峰 #1	4050.09	4050.1	PYY(3-36) (SEQ ID NO: 18)
			峰#2	10168.47	10168.4	RL27标签
			峰#3	4204.16	4204.1	GP-PYY(3-36)
反应9	蛋白酶 28997	1:20	峰 #1	4050.10	4050.1	PYY(3-36) (SEQ ID NO: 18)
			峰 #2	10168.49	10168.4	RL27标签
反应10	蛋白酶 28997	1:100	峰 #1	4050.09	4050.1	PYY(3-36) (SEQ ID NO: 18)
			峰#2	10168.47	10168.4	RL27标签
			峰#3	4204.16	4204.1	GP-PYY(3-36)
			峰#4	14354.62	14354.5 Da	RL27_EVLFGQP_PYY(3-36)

[0232] 作为模型蛋白底物的 RL27_EVLFGQP_胰高血糖素

如上所述,建立蛋白酶 20986、28994、28996 和 28997 与 RL27-EVLFGQP-胰高血糖素底物的温育。通过 LC-MS 分析完整质量显示,蛋白酶 20986、28994、28996 和 28997 均能够将 RL27_EVLFGQP_胰高血糖素加工成成熟胰高血糖素,其中在总体效率和特异性方面观察到差异,使用 1:100 或 1:500 酶底物比以及使用 4°C 或 37°C 温育温度(图 6-9)。对于蛋白酶 20986、28996,1:500 的酶底物比和 4°C 的温育温度(表 9, 反应 11 和 16)) 提供最优化的切割条件,其中彻底加工了融合蛋白且没有显著的非特异性切割(图 6 和 8)。释放的胰高血糖素的测定质量与人野生型胰高血糖素的计算质量 3482.8 Da 一致(峰 #1)。蛋白酶 28994 和 28997 的效率较差,在测试的条件下没有彻底加工全部的融合蛋白(对于蛋白酶 28994),低强度的峰(峰 #3 和 #4) 表示非常有限的非特异性切割(表 9, 反应 13(图 7) 14 和 17(图 9))。含有 RL9-HRV14 3C (SEQ ID NO. 21) 的对照仅产生 GP_胰高血糖素(反应 18, 图 10),显示 XaaProDAP 结构域负责完成反应以产生胰高血糖素 (SEQ ID NO: 19) 的天然 N-末端组氨酸。不添加酶仅产生未加工的融合蛋白,测定质量与不含起始甲硫氨酸的 RL27_EVLFGQP_胰高血糖素的计算质量 13787.1 Da 一致。这表明,组合来自人鼻病毒或人柯萨奇病毒的小 RNA 病毒 3C 蛋白酶与来自乳酸乳球菌或猪链球菌的 XaaProDAP 的不同融合蛋白酶变体可被成功优化,以将 RL27_EVLFGQP_胰高血糖素加工成成熟的胰高血糖素,其中 His 为正确的 N-末端氨基酸残基,且不产生或产生极少的融合蛋白相关杂质。

[0233] 表 9. 酶反应,使用蛋白酶 20986、28994、28996 和 28997 以及作为底物的 RL27_EVLFGQP_胰高血糖素,在 4°C 温育过夜。示出反应 11-18 的去卷积质谱中检测到的实验测定的主要峰。

反应编号	酶	摩尔比	主要峰	测定质量 (道尔顿)	计算质量 (道尔顿)	对应于
反应11	20986	1:100	峰#1	3482.61	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10168.37	10168.4	RL27标签
反应12	20986	1:500	峰#1	3481.62	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10168.4	10168.4	RL27标签
反应13	28994	1:100	峰#1	3481.61	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10168.39	10168.4	RL27标签
			峰#3	3257.52	3258.6	胰高血糖素 (3-29)
			峰#4	3072.44	3073.4	胰高血糖素 (5-29)
			峰#5	13787.06	13787.1	RL27_EVLFPQ-胰 高血糖素
反应14	28994	1:500	峰#1	3481.62	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10168.41	10168.4	RL27标签
			峰#3	13787.09	13787.1	RL27_EVLFPQ-胰 高血糖素
反应15	28996	1:100	峰#1	3481.63	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10168.46	10168.4	RL27标签
反应16	28996	1:500	峰#1	3481.65	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10167.48	10168.4	RL27标签
反应17	28997	1:500	峰#1	3481.65	3482.8	胰高血糖素 (SEQ ID NO:19)
			峰#2	10168.48	10168.4	RL27标签
			峰#3	13787.19	13787.1	RL27_EVLFPQ-胰 高血糖素
反应18	RL9- HRV14 3C	1:20	峰#1	3636.72	3636.7	GP-胰高血糖素
			峰#2	10167.44	10168.4	RL27标签

[0234] 作为模型蛋白底物的 RL27_EVLFPQ-GLP-1 (7-37, K34R)

如上所述, 建立蛋白酶 20986、28994、28996 和 28997 与 RL27_EVLFPQ-GLP-1 (7-37, K34R) 底物的温育。通过 LC-MS 分析完整质量显示, 蛋白酶 20986、28994、28996 和 28997 均能够将 RL27_EVLFPQ-GLP-1 完全加工成成熟 GLP-1 (7-37, K34R), 测定分子量对应于计算质量 3382, 7 Da (表 10, 图 11-14)。在总体效率和特异性方面观察到微小的差异, 使用 1:100 或 1:500 的酶底物比以及 4°C 或 37°C 作为温育温度。观察到的非特异性片段主要是 GLP-1 (9-37, K34R) (计算质量 3174, 6 Da), 其中由 GLP-1 序列除去了额外的二肽。在该实验设定中, 最佳切割条件在 4°C 获得, 其中彻底加工融合蛋白且具有非常有限的或不具有非特异性切割。蛋白酶 28994 效率较差 (反应 21 (图 12) 和 22, 表 10), 因为在温育后观察到剩余的融合蛋白。蛋白酶 28996 提供融合蛋白的彻底加工, 并释放成熟 GLP-1 (7-37, K34R), 其中利用 37°C 3h 未观察到非特异性切割 (未示出)。

[0235] 最有效的反应获自蛋白酶 20986, 其优化切割条件使用 1:500 的酶底物比以及 4°C 温育过夜, 没有可检测的来自非特异性或不完整加工的片段分布 (反应 20, 图 11)。蛋白酶 28996 和 28997 获得类似结果 (反应 23 (图 13) & 25 (图 14)), 其几乎排它性地产

生完全加工的成熟 GLP-1(7-37, K34R), 使用 1:100 的酶底物比以及 4°C 温育过夜, 而在以 1:500 的比温育后, 可检测到少但可检测到的量的未加工的 GP-GLP-1(7-37, K34R) (成熟峰强度的约 10%) (反应 24 & 26)。含有 RL9-HRV14 3C (SEQ ID NO:21) 的对照仅产生预期的 GP_GLP-1(7-37, K34R) (反应 27, 图 15), 显示融合蛋白酶的 XaaProDAP 酶结构域负责提供 GLP-1 中的天然 N-末端组氨酸 (7-37, K34R)。不添加酶仅产生未加工的融合蛋白, 测定质量与计算质量 13688.1 Da 一致, 对应于不含起始甲硫氨酸的 RL27_EVLFGQP_GLP-1(7-37, K34R)。因此, 组合来自人鼻病毒或人柯萨奇病毒的小 RNA 病毒 3C 蛋白酶与来自乳酸乳球菌或猪链球菌的 XaaProDAP 的不同融合蛋白酶变体可被优化, 以将 RL27_EVLFGQP_GLP-1(7-37, K34R) 加工成成熟的 GLP-1(7-37, K34R) (SEQ ID NO:20), 其中 His 为正确的 N-末端且不产生或产生非常有限的融合蛋白相关杂质。

[0236] 表 10. 酶反应, 使用蛋白酶 20986、28994、28996 和 28997 以及作为底物的 RL27_EVLFGQP_GLP-1(7-37, K34R), 4°C 温育过夜。示出反应 19-27 的去卷积质谱中检测到的实验测定的主要峰。

反应编号	酶	摩尔比	主要峰	测定质量 (道尔顿)	计算质量 (道尔顿)	对应于
反应 19	20986	1:100	峰#1	3174.6	3174.6	GLP-1(9-37,K34R)
			峰#2	3382.7	3382.7	GLP-1(7-37,K34R) (SEQ ID NO: 20)
			峰#3	10168.46	10168.4	RL27标签
反应 20	20986	1:500	峰#1	3382.71	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:20)
			峰#2	10167.49	10168.4	RL27标签
反应 21	28994	1:100	峰#1	3175.58	3174.6	GLP-1(9-37,K34R)
			峰#2	3382.68	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:20)
			峰#3	10167.39	10168.4	RL27标签
			峰#4	13688.14	13688.1	RL27_EVLFGQP_GLP-1(7-37,K34R)
反应 22	28994	1:500	峰#1	3382.67	3382.7	GLP-1(7-37,K34R)

			峰#2	10168.4	10168.4	(SEQ ID NO: 20) RL27标签
			峰#3	13688.15	13688.1	RL27_EVLFQGP_GLP-1(7-37,K34R)
反应23	28996	1:100	峰#1	3174.6	3174.6	GLP-1(9-37,K34R)
			峰#2	3382.69	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:20)
			峰#3	10167.44	10168.4	RL27标签
反应24	28996	1:500	峰#1	3382.7	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:20)
			峰#2	10168.48	10168.4	RL27标签
			峰#3	3537.77	3537.7	GP-GLP-1(7-37,K34R)
反应25	28997	1:100	峰#1	3382.7	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:20)
			峰#2	10168.47	10168.4	RL27标签
反应26	28997	1:500	峰#1	3382.71	3382.7	GLP-1(7-37,K34R) (SEQ ID NO:20)
			峰#2	10167.49	10168.4	RL27标签
			峰#3	3537.78	3537.7	GP-GLP-1(7-37,K34R)
反应27	RL9 HRV14 3C		峰#1	3537.78	3537.7	GP-GLP-1(7-37,K34R)
			峰#2	10168.48	10168.4	RL27标签

[0237] 虽然在此阐述和说明了本发明的一些特征,但是本领域技术人员现在会想到许多改变、替换、变化和等效物。因此,应当理解所附权利要求旨在覆盖落入本发明的真正精神范围内的所有这种改变和变化。

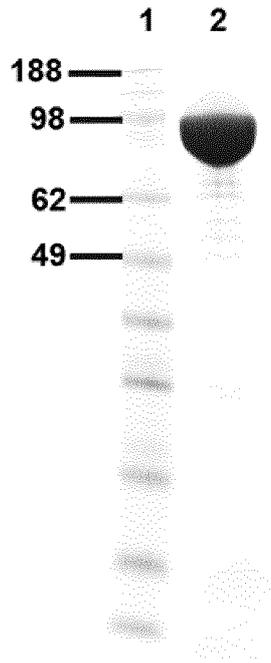


图 1

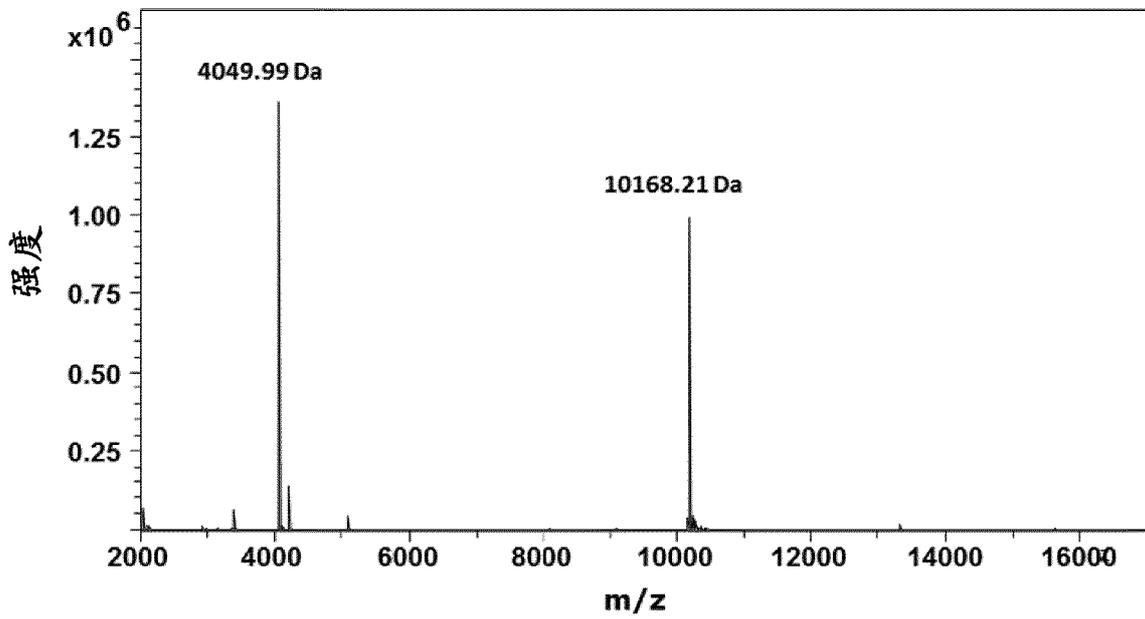


图 2

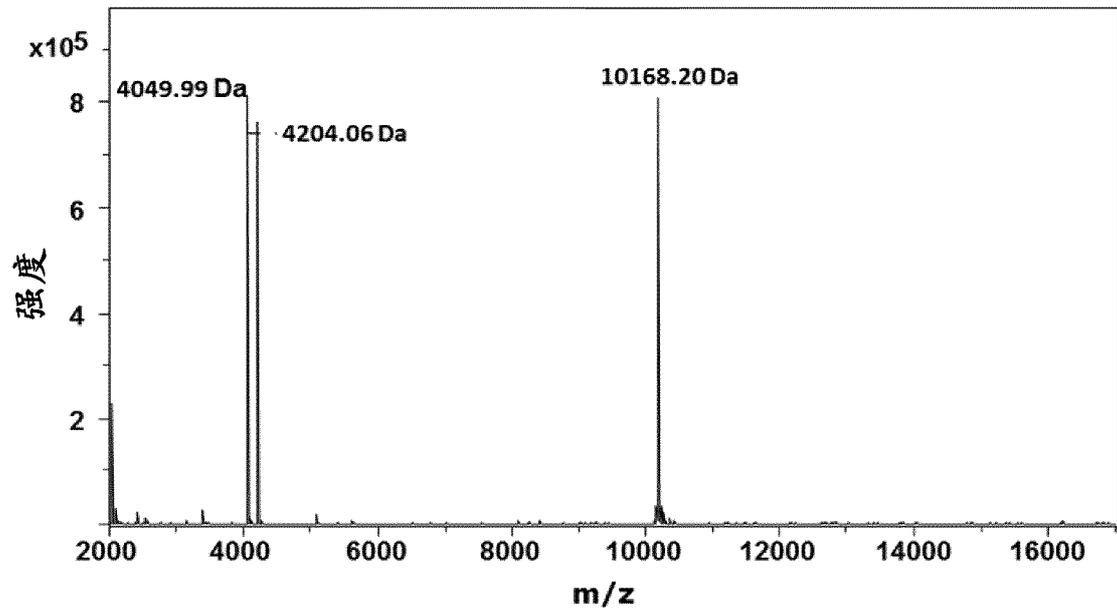


图 3

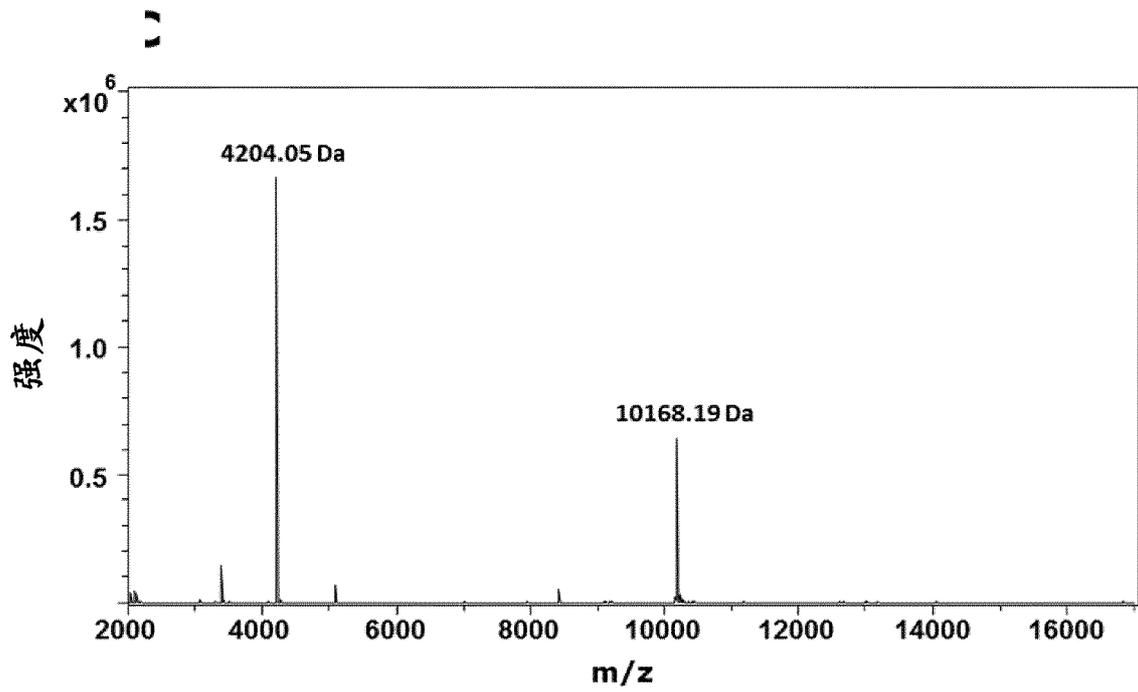


图 4

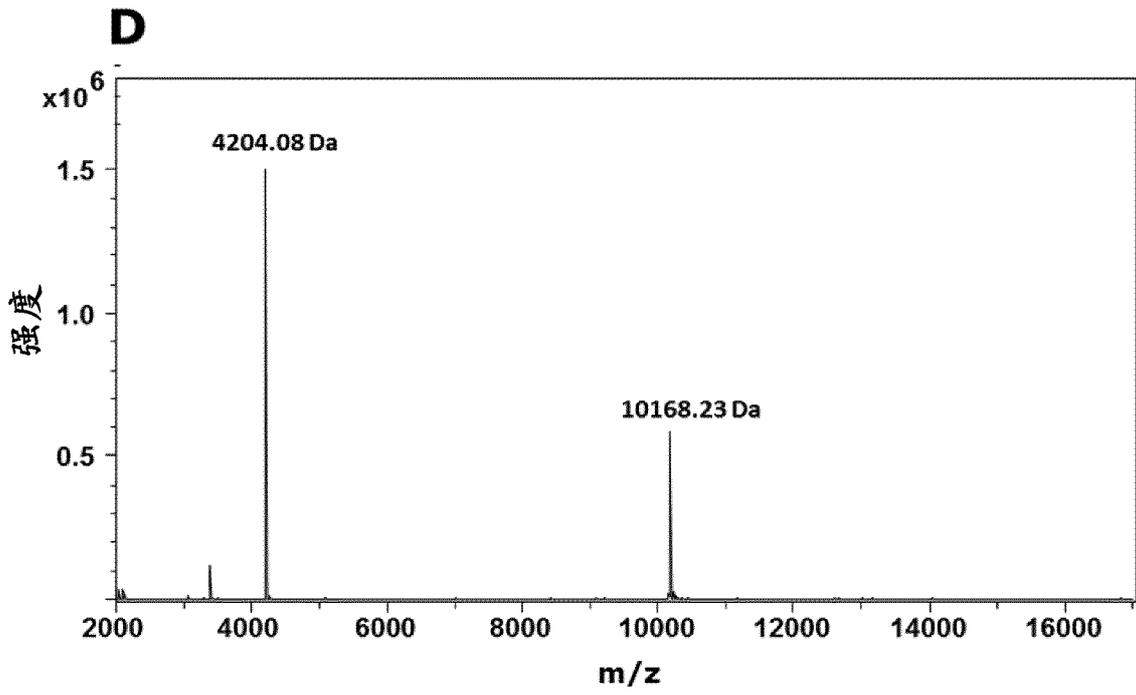


图 5

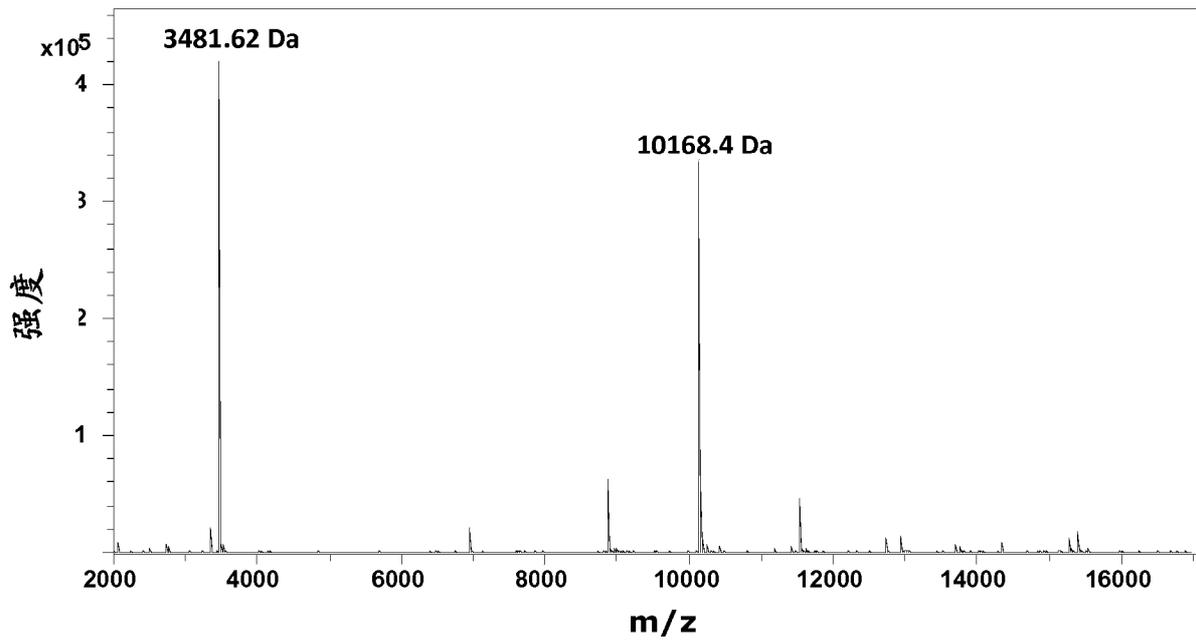


图 6

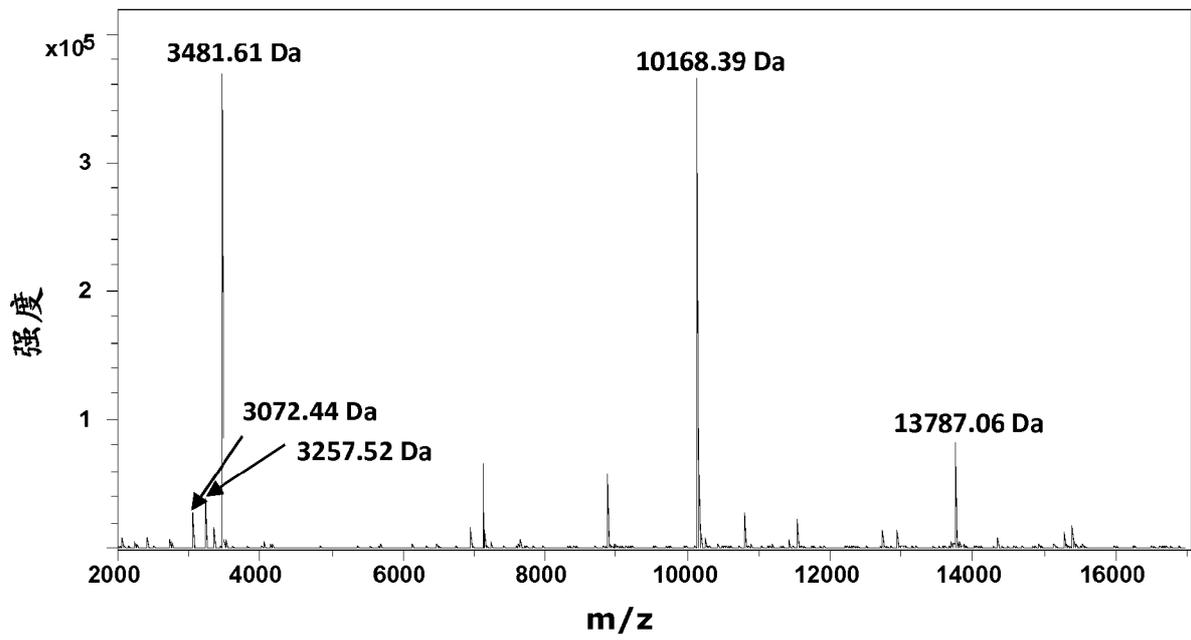


图 7

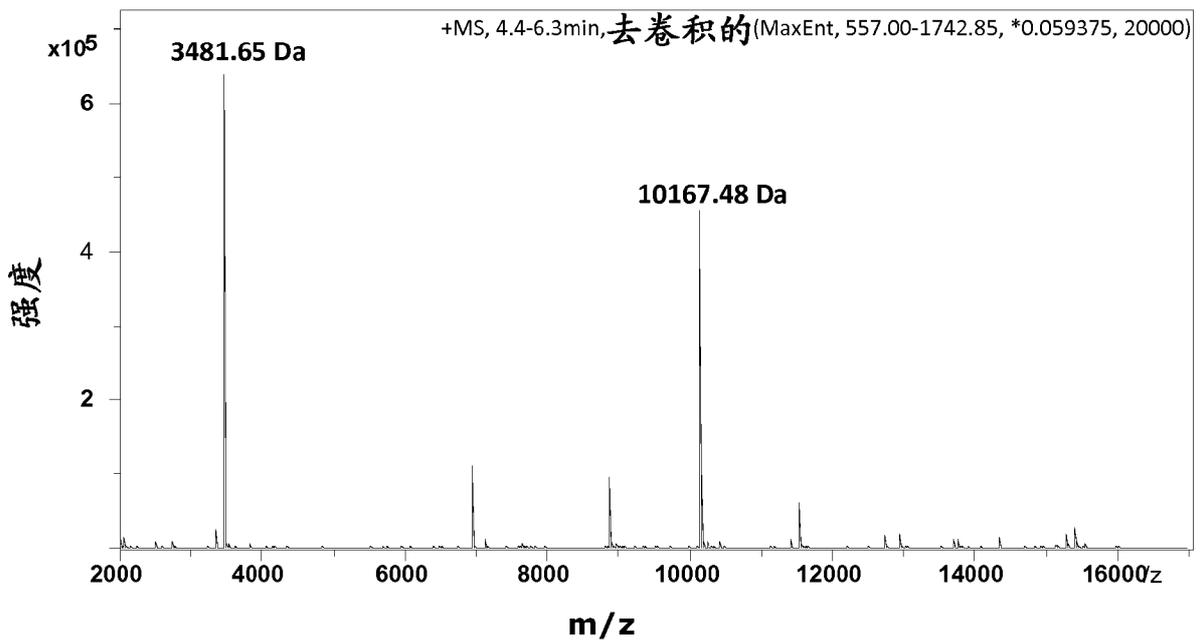


图 8

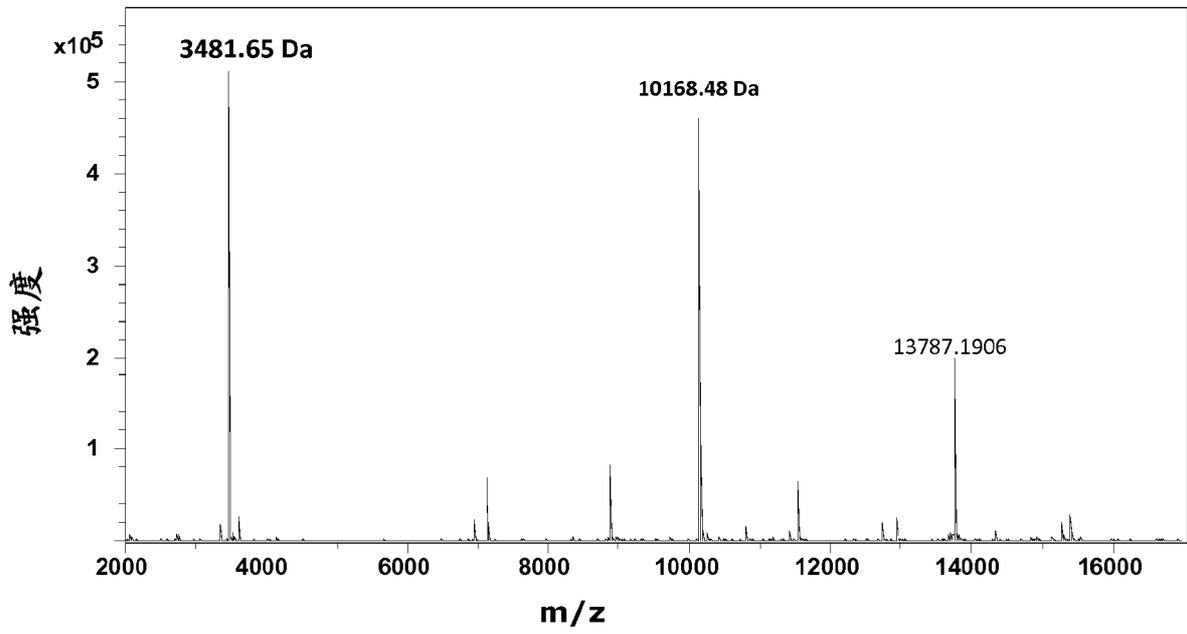


图 9

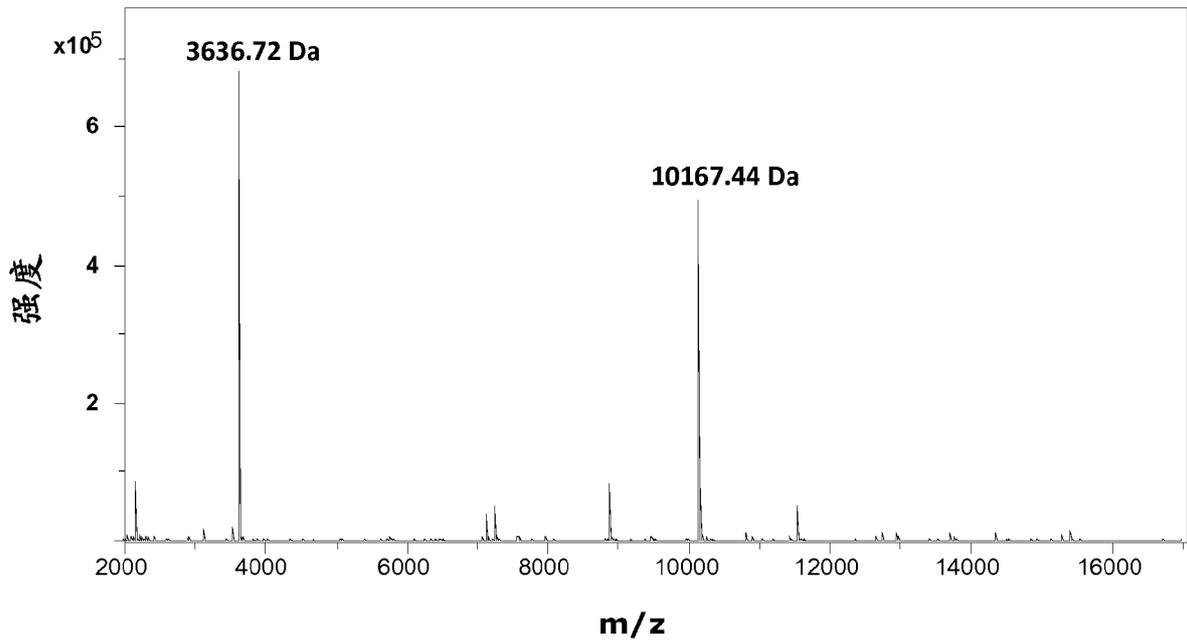


图 10

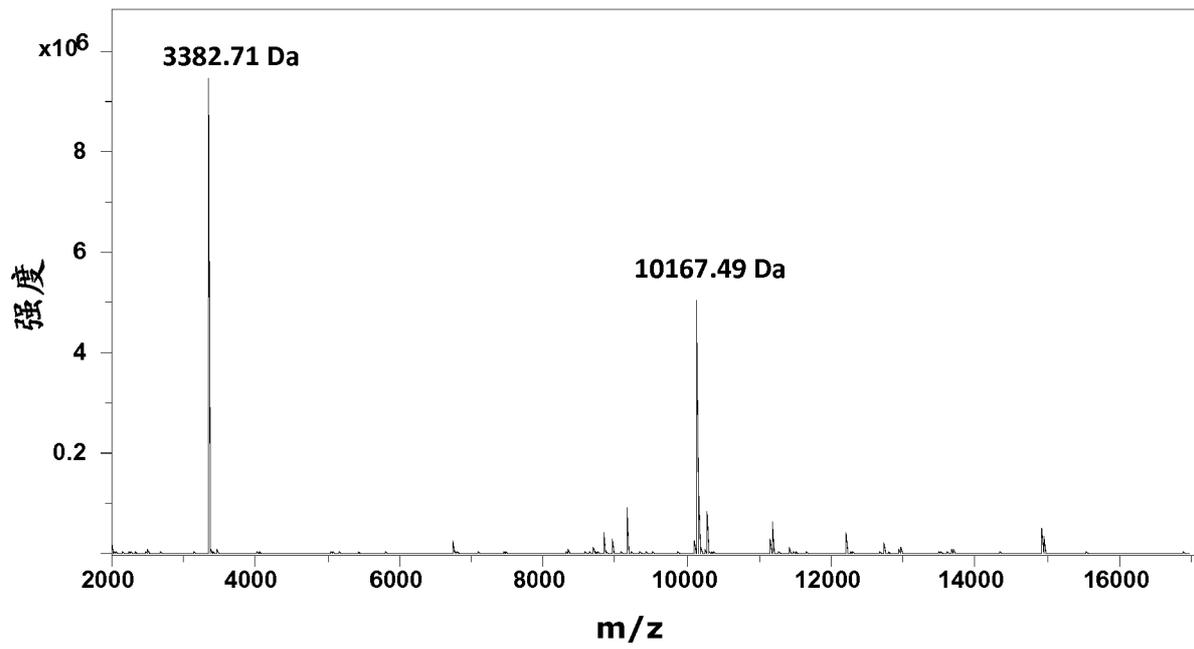


图 11

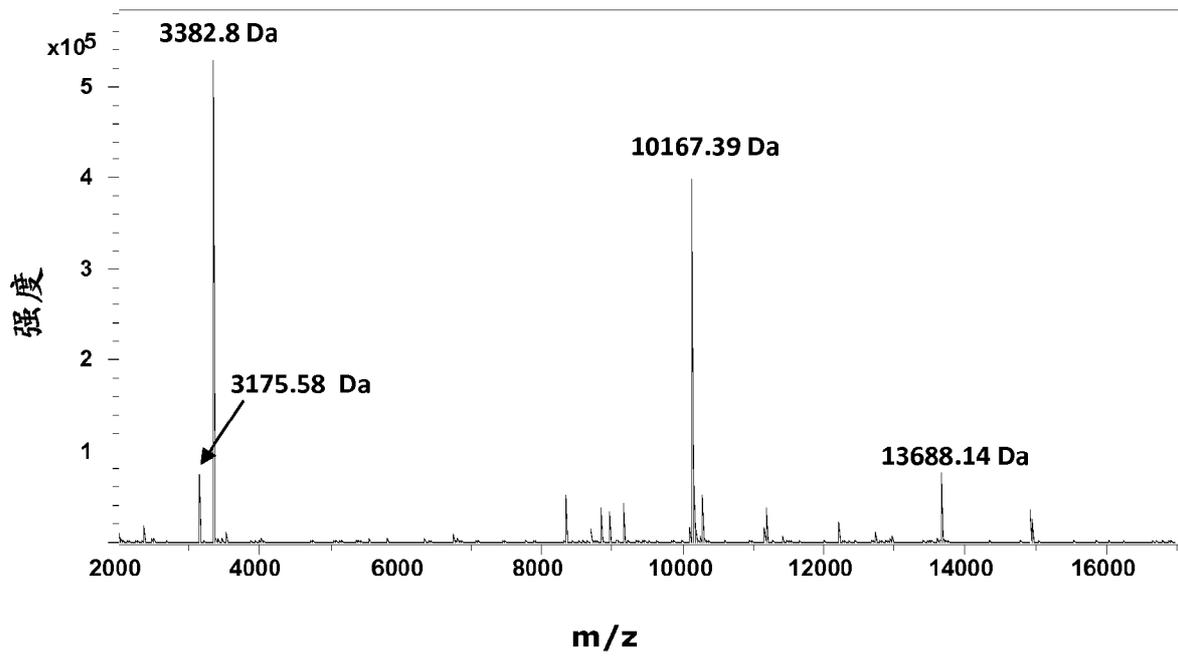


图 12

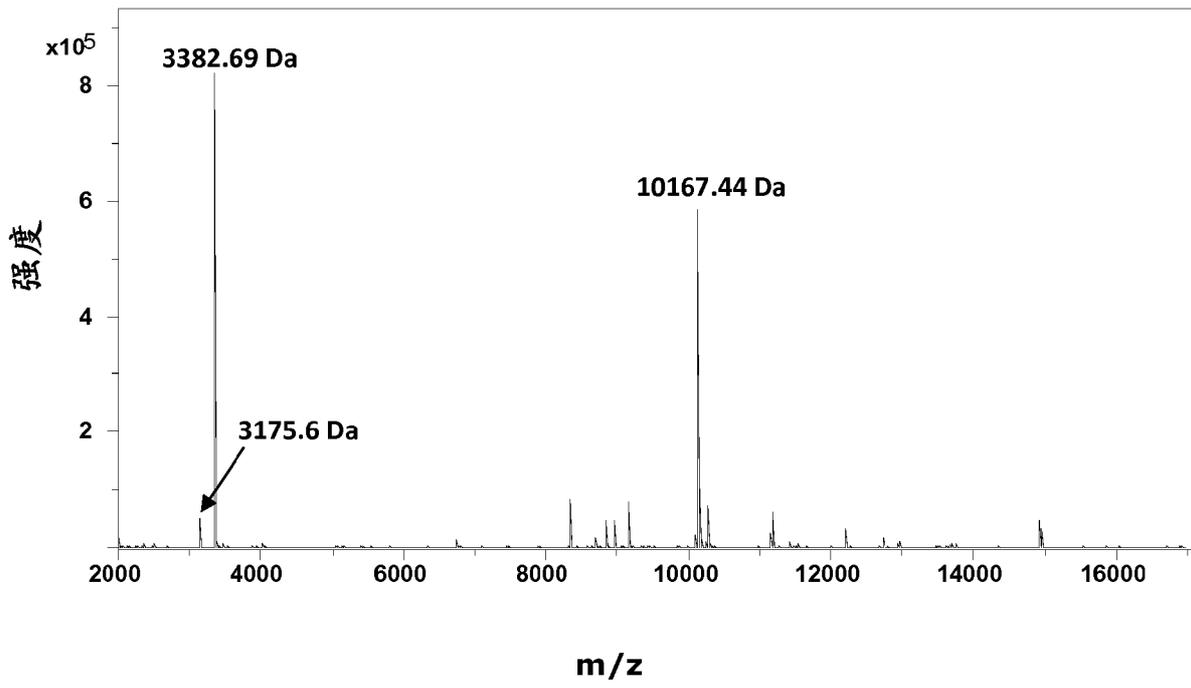


图 13

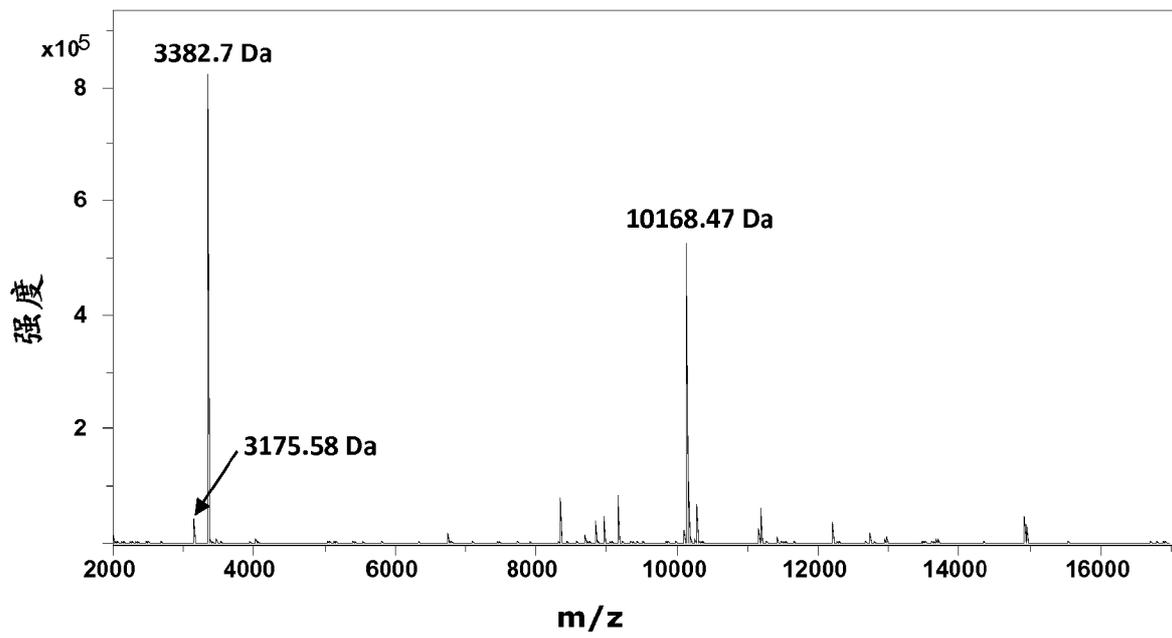


图 14

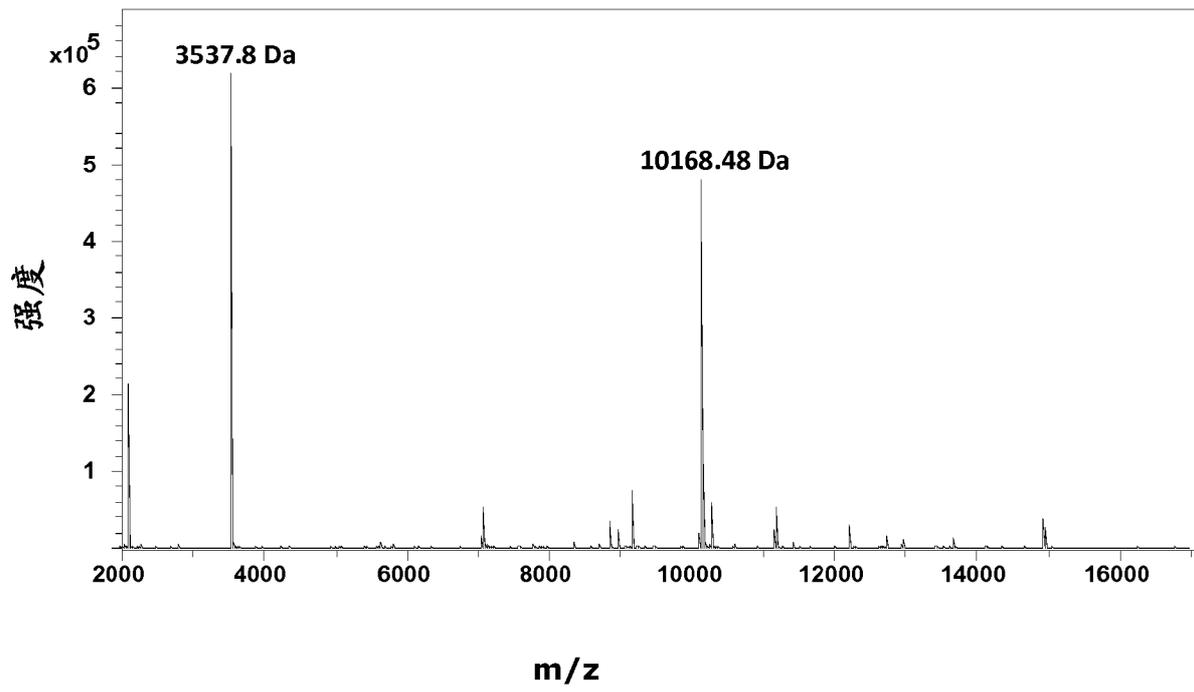


图 15