

IgA 蛋白酶截短体、包含 IgA 蛋白酶截短体的融合蛋白及其用途

5 技术领域

本申请涉及生物医药领域，具体地，本申请涉及一种 IgA 蛋白酶截短体、包含 IgA 蛋白酶截短体的融合蛋白、包含所述 IgA 蛋白酶截短体或所述融合蛋白的药物组合物、编码所述 IgA 蛋白酶截短体或所述融合蛋白的核酸、所述 IgA 蛋白酶截短体或所述融合蛋白的制备方法、以及 IgA 蛋白酶截短体或所述融合蛋白在制备用于治疗 IgA 沉积相关疾病的药物中的用途。

背景技术

IgA 肾病是目前世界上最常见的原发性肾小球疾病之一，给患者和社会带来沉重的负担。目前针对 IgA 肾病尚缺乏特异性的治疗。临床上多用 RAS 阻断剂为基础的支持治疗，以减缓肾功能恶化。对于支持治疗无效的患者予以联合激素免疫抑制剂治疗。但激素免疫抑制剂的使用长期疗效不佳且给患者带来严重的副作用。

亟需开发有效并且具有低副作用的治疗药物。

发明简述

在一个方面，本申请提供了一种分离的 IgA 蛋白酶截短体，其包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的非天然截短片段，或者与所述非天然截短片段具有至少 70% 的序列同一性。在某些实施方式中，所述非天然截短片段在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的基础上有氨基酸取代、缺失、插入或修饰，使得所述 IgA 蛋白酶截短体丧失或降低自酶切功能。在某些实施方式中，所述氨基酸取代、缺失、插入或修饰发生在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的天然自酶切位点、所述天然自酶切位点的上游 5 个位点以内和/或下游 5 个位点以内。在某些实施方式中，所述多枝梭菌（*Clostridium ramosum*）为 *Clostridium ramosum* AK183 株。在某些实施方式中，所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的氨基酸序列

如 SEQ ID NO: 1 所示。在某些实施方式中，所述天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 730 位至第 840 位之间（例如，第 792 位至第 797 位之间）。在某些实施方式中，所述天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 790 位、第 791 位、第 792 位、第 793 位、第 794 位、第 795 位、第 796 5 位、第 797 位、第 798 位、第 799 位或第 800 位。

在某些实施方式中，所述非天然截短片段为获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的 N 末端截短片段或 C 末端截短片段。在某些实施方式中，所述 N 末端截短片段包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的 N 末端第 31 位起的至少 760 个连续氨基酸的多肽片段，或者与所述多肽片段具有至少 70% 的序列同一性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的至少 760 个（例如，至少 761 个、至少 762 个、至少 763 个、至少 764 个、至少 765 个、至少 766 个、至少 767 个、至少 768 个、至少 769 个、至少 770 个、至少 771 个、至少 772 个、至少 773 个、至少 774 个、至少 775 个、至少 776 个、至少 777 个、至少 778 个、至少 779 个、至少 780 个、至少 781 个、至少 782 个、至少 783 个、至少 784 个、至少 785 个、至少 786 个、至少 787 个、至少 788 个、至少 789 个、至少 790 个、至少 791 个、至少 792 个、至少 793 个、至少 794 个、至少 795 个、至少 796 个、至少 797 个、至少 798 个、至少 799 个、至少 800 个、至少 801 个、至少 802 个、至少 803 个、至少 804 个、至少 805 个、至少 806 个、至少 807 个、至少 808 个、至少 809 个、至少 810 个、至少 900 个、至少 950 个、至少 1000 个、至少 1100 个、至少 1150 个或者至少 1200 个）连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含选自下组的多肽片段：如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 798 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 807 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 816 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 833 位氨基酸以及与其具有至少 70% 的序列同一性的多肽片段。

在某些实施方式中，所述非天然截短片段包含获得自或衍生自多枝梭菌

(*Clostridium ramosum*) 的野生型 IgA 蛋白酶的 N 末端第 335 位起的至少 456 个连续氨基酸的多肽片段, 或者与所述多肽片段具有至少 90%或至少 95%的序列同一性。在某些实施方式中, 本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位起的至少 456 个 (例如, 至少 457 个、至少 458 个、
5 至少 459 个、至少 460 个、至少 461 个、至少 462 个、至少 463 个、至少 464 个、至少 465 个、至少 466 个、至少 467 个、至少 468 个、至少 469 个、至少 470 个、至少 471 个、至少 472 个、至少 473 个、至少 474 个、至少 475 个、至少 476 个、至少 477 个、至少 478 个、至少 479 个、至少 480 个、至少 481 个、至少 482 个、至少 483 个、至少 484 个、至少 485 个、至少 486 个、至少 487 个、至少 488 个、
10 至少 489 个、至少 490 个、至少 491 个、至少 492 个、至少 493 个、至少 494 个、至少 495 个、至少 496 个、至少 497 个、至少 498 个、至少 499 个、至少 500 个、至少 550 个、至少 600 个、至少 650 个、至少 700 个、至少 750 个、至少 800 个、至少 850 个或者至少 900 个) 连续氨基酸的多肽片段。在某些实施方式中, 本申请提供的 IgA 蛋白酶截短体包含选自下组的多肽片段: 如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 816 位氨基酸以及与其具有至少 90%或至少 95%的序列同一性的多肽片段。
20

在某些实施方式中, 本申请提供的 IgA 蛋白酶截短体在所述多肽片段的氨基酸序列基础上, 在一个或多个位点具有氨基酸的保守替换。在某些实施方式中, 所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置有氨基酸突变。在某些实施方式中, 所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置突变为甘氨酸。
25

在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位有氨基酸突变、第 862 位有氨基酸突变、第 931 位和第 933 位有氨基酸突变、第 978 位有氨基酸突变、或者第 1002 位和第 1004 位有氨基酸突变。在某些实施方式中，所述多肽片段的氨基酸序列如 SEQ ID NO: 53（也称为“PA-GA Mut”）、SEQ ID NO: 54（也
5 称为“PI-GI Mut”）、SEQ ID NO: 55（也称为“PAP-GAG Mut”）、SEQ ID NO: 56（也称为“PAT-GAT Mut”）或者 SEQ ID NO: 57（也称为“PIP-GIG Mut”）所示。

在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA 的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA 重链的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特
10 异性切割人 IgA 重链 CH1 与铰链区交汇处的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA1 的酶活性。

在另一方面，本申请提供了一种融合蛋白，其包含第一多肽和第二多肽，其中所述第一多肽包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的全长、获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶去掉信号肽之后形成的多肽或者如本申请所述的 IgA 蛋白酶截短体，所述第
15 二多肽包含用于延长所述第一多肽在受试者体内半衰期的氨基酸序列。在某些实施方式中，所述第一多肽包含如 SEQ ID NO: 1 或者 SEQ ID NO: 42 所示的序列。在某些实施方式中，所述第二多肽位于所述第一多肽的 N 末端或者 C 末端。

在某些实施方式中，所述第一多肽和所述第二多肽之间通过连接子连接。在某些实施方式中，所述第一多肽和所述第二多肽之间直接连接。在某些实施方式中，所述连接子选自下组：可切割连接子、不可切割连接子、肽连接子、柔性连接子、刚性连接子、螺旋连接子和非螺旋连接子。在某些实施方式中，所述连接子包括肽连接子。在某些实施方式中，所述肽连接子包括含有甘氨酸和丝氨酸的连接子。在某些实施方式中，所述含有甘氨酸和丝氨酸的连接子包括如 SEQ ID NO: 21 (GGGS)、
20 SEQ ID NO: 22 (GGGGS)、SEQ ID NO: 86 (GGGGGS) 或 SEQ ID NO: 87 (GGGGGGGS) 所示的一个、两个、三个、四个或更多个重复。在某些实施方式中，所述连接子包括如 SEQ ID NO: 23 (GGCGGCGGTGGATCC) 所示、如 SEQ ID NO: 58 (EEKKKEKEKEEQEERETK) 所示或者如 SEQ ID NO: 59 (HHHHHHHHHH) 所示的氨基酸序列。

在某些实施方式中，所述第二多肽选自：Fc 结构域和白蛋白。在某些实施方式中，所述 Fc 结构域包含铰链区。在某些实施方式中，所述 Fc 结构域来源于人 IgG Fc 结构域。在某些实施方式中，所述 Fc 结构域来源于人 IgG1 Fc 结构域、人 IgG2 Fc 结构域、人 IgG3 Fc 结构域或人 IgG4 Fc 结构域。在某些实施方式中，所述 Fc 结构域包含与 SEQ ID NO: 24、SEQ ID NO: 25、SEQ ID NO: 32 或 SEQ ID NO: 77 具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%或至少 95%的序列同一性的氨基酸序列。在某些实施方式中，所述 Fc 结构域包含如 SEQ ID NO: 24、SEQ ID NO: 25、SEQ ID NO: 32 或 SEQ ID NO: 77 所示的氨基酸序列。在某些实施方式中，所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位有氨基酸突变。在某些实施方式中，所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位的氨基酸（例如，丙氨酸）突变为缬氨酸、甘氨酸、丝氨酸或者亮氨酸。在某些实施方式中，所述 Fc 结构域包含一个或多个延长所述融合蛋白的半衰期的突变。在某些实施方式中，所述 Fc 结构域与所述第一多肽的 C 末端或者 N 末端连接。在某些实施方式中，所述白蛋白包含人血清白蛋白的一个或多个结构域。在某些实施方式中，所述白蛋白包含人血清白蛋白的 D3 结构域。

在某些实施方式中，本申请所述的融合蛋白进一步包含标签。在某些实施方式中，所述标签选自下组：荧光标签、发光标签、纯化标签和生色标签。在某些实施方式中，所述标签选自下组：c-Myc 标签、HA 标签、VSV-G 标签、FLAG 标签、V5 标签和 HIS 标签。在某些实施方式中，所述标签是包含 6 个、7 个、8 个、9 个或 10 个组氨酸的 HIS 标签。在某些实施方式中，所述第二多肽位于所述第一多肽的 C 末端，所述标签位于所述第二多肽的 C 末端。

在某些实施方式中，本申请所述的融合蛋白在受试者体内的血循环中的半衰期为至少 1 天、至少 2 天、至少 3 天、至少 4 天、至少 5 天、至少 6 天、至少 7 天、至少 8 天、至少 9 天、至少 10 天、至少 11 天、至少 12 天、至少 13 天、至少 14 天。

在另一方面，本申请提供了一种分离的核酸，其包含编码本申请所述的 IgA 蛋白酶截短体的核苷酸序列或包含编码本申请所述的融合蛋白的核苷酸序列。在某些实施方式中，本申请所述的核酸包含选自下组的核苷酸序列：SEQ ID NO: 3、SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ

ID NO: 33、SEQ ID NO: 34、SEQ ID NO: 35、SEQ ID NO: 36、SEQ ID NO: 37、SEQ ID NO: 38 以及与其具有至少 70%的序列同一性的核苷酸序列。

在另一方面，本申请提供了一种载体，其包含本申请所述的核酸。

在另一方面，本申请提供了一种细胞，其包含本申请所述的核酸或如本申请所述的载体。在某些实施方式中，所述细胞是原核生物细胞或真核生物细胞。在某些实施方式中，所述原核生物细胞是大肠杆菌细胞。在某些实施方式中，所述真核生物细胞是哺乳动物细胞。在某些实施方式中，所述哺乳动物细胞是人细胞或中国仓鼠卵巢（CHO）细胞。在某些实施方式中，所述哺乳动物细胞是人胚胎肾细胞 293（HEK293 细胞）。

10 在另一方面，本申请提供了一种药物组合物，其包含本申请所述的 IgA 蛋白酶截短体、包含本申请所述的融合蛋白、包含本申请所述的核酸、包含本申请所述的载体或者包含本申请所述的细胞，以及药学上可接受的载体。

在另一方面，本申请提供了一种产生融合蛋白的方法，其包括培养本申请所述的细胞的步骤。

15 在另一方面，本申请提供了一种治疗或预防 IgA 沉积相关疾病的方法，其包括向需要治疗或预防的受试者施用如本申请所述的 IgA 蛋白酶截短体、如本申请所述的融合蛋白或者如本申请所述的药物组合物。

在另一方面，本申请提供了如本申请所述的 IgA 蛋白酶截短体、如本申请所述的融合蛋白或者如本申请所述的药物组合物在制备用于治疗或预防 IgA 沉积相关疾病的药物中的用途。

20 在另一方面，本申请提供了用于治疗或预防 IgA 沉积相关疾病的如本申请所述的 IgA 蛋白酶截短体、如本申请所述的融合蛋白或者如本申请所述的药物组合物。

在另一方面，本申请提供了一种治疗或预防 IgA 沉积相关疾病的方法，其包括向需要治疗或预防的受试者施用 IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物组合物，其中，所述 IgA 蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、

SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、
SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75、SEQ ID NO: 76
或其组合。

在另一方面，本申请提供了 IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其
5 截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物
组合物在制备用于治疗或预防 IgA 沉积相关疾病的药物中的用途，其中，所述 IgA
蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、
SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、
SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、
10 SEQ ID NO: 74、SEQ ID NO: 75、SEQ ID NO: 76 或其组合。

在另一方面，本申请提供了用于治疗或预防 IgA 沉积相关疾病的 IgA 蛋白酶
或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白
酶或其截短体或者所述融合蛋白的药物组合物，其中，所述 IgA 蛋白酶的氨基酸序
列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、
15 SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、
SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、
SEQ ID NO: 75、SEQ ID NO: 76 或其组合。

在某些实施方式中，所述 IgA 沉积相关疾病包括 IgA 肾病、疱疹样皮炎、亨诺
赫-舍恩莱因紫癜（又称 IgA 血管炎）、川崎病、紫癜性肾炎、IgA 血管炎肾损害、
20 IgA 类风湿因子阳性的类风湿性关节炎、IgA 型抗 GBM 病或 IgA 型 ANCA 相关血
管炎。在某些实施方式中，所述 IgA 沉积相关疾病为 IgA 肾病、IgA 血管炎或川崎
病。

附图描述

图 1 显示了 AK183(31-737)、AK183(31-768)、AK183(31-798)和 AK183(31-833)
25 这四种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 2 显示了 AK183(31-773)、AK183(31-778)、AK183(31-782)、AK183(31-787)
和 AK183(31-792)这五种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 3a-c 显示了 AK183(31-788)、AK183(31-789)、AK183(31-790)和 AK183(31-791)这四种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 4 显示了 PET30a-AK183(31-790)-Fc 质粒的构建流程图。

图 5 显示了 AK183(31-790)-Fc 融合蛋白的表达结果图。

5 图 6a 显示了 AK183(31-792)-Fc 融合蛋白的表达结果图，图 6b 显示了 AK183(31-792)-Fc 融合蛋白对 IgA1 的体外酶切活性实验结果。

图 7 显示了 AK183(31-798)-Fc、AK183(31-807)-Fc、AK183(31-816)-Fc 和 AK183(31-833)-Fc 这四种融合蛋白对 IgA1 的体外酶切活性实验结果。

图 8 显示了 AK183(31-807)-Fc 融合蛋白对 IgA1 的体内酶切活性实验结果。

10 图 9 显示了 AK183(31-792)-Fc 融合蛋白在 HEK293 细胞中的表达结果图。

图 10 显示了 AK183(285-792)、AK183(330-792)、AK183(380-792)、AK183(430-792)、AK183(480-792)、AK183(530-792)和 AK183(580-792)这 7 种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 11 显示了 AK183(335-792)、AK183(340-792)、AK183(345-792)、AK183(350-15 792)、AK183(355-792)、AK183(360-792)、AK183(365-792)、AK183(370-792)和 AK183(375-792)这 9 种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 12 显示了 AK183(336-792)、AK183(337-792)、AK183(338-792)和 AK183(339-792)这四种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 13 显示了再次验证 AK183(285-792)、AK183(330-792)、AK183(335-792)、20 AK183(336-792)、AK183(337-792)、AK183(338-792)、AK183(339-792)、AK183(340-792)、AK183(345-792)和 AK183(350-792)这 10 种 IgA 蛋白酶截短体对 IgA1 的体外酶切活性实验结果。

图 14 显示了 AK183(285-816)-Fc 和 Fc-AK183(285-816)这两种融合蛋白的表达结果图。

25 图 15 显示了 AK183(285-816)-Fc 和 Fc-AK183(285-816)这两种融合蛋白对 IgA1

的体外酶切活性实验结果。

图 16 显示了 Fc-AK183(285-816)融合蛋白、AK183(285-816)-Fc 融合蛋白和 AK183(285-816) IgA 蛋白酶截短体对 IgA1 的酶切活性实验结果。

图 17 显示了 Fc-AK183(285-816)融合蛋白对 IgA1 的体内酶切活性实验结果。

5 图 18 显示了 AK183(285-816)-Fc 融合蛋白和 Fc-AK183(31-1203)融合蛋白对 IgA1 的酶切活性实验结果。

图 19 显示了 AK183(31-816)-IgG1 Fc 融合蛋白、AK183(31-816)-IgG4 Fc 融合蛋白和 AK183(31-816)-白蛋白融合蛋白对 IgA1 的酶切活性实验结果。

10 图 20 显示了具有不同连接子（SEQ ID NO: 59、SEQ ID NO: 58、SEQ ID NO: 22、SEQ ID NO: 78、SEQ ID NO: 79 或 SEQ ID NO: 80）的 AK183(285-816)-Fc 融合蛋白对 IgA1 的酶切活性实验结果。

图 21 显示了如 SEQ ID NO: 53、SEQ ID NO: 54、SEQ ID NO: 55、SEQ ID NO: 56 或 SEQ ID NO: 57 所示的五种 IgA 蛋白酶截短体的突变体对 IgA1 的酶切活性实验结果。

15 图 22 显示了对 AK183(31-816)-Fc 融合蛋白的 Fc 区分别进行 4 种不同突变形成的四种突变体对 IgA1 的酶切活性实验结果。

图 23a 和图 23b 显示了 16 种 AK183 同源酶对 IgA1 的酶切活性实验结果。

发明详述

20 尽管本申请将在以下公开多个方面和实施方式，但是在不违背本申请主题精神和范围的前提下，本领域技术人员显然可以对其进行各种等同改变和修改。本申请公开的多个方面和实施方式仅用于举例说明，其并非旨在限制本申请，本申请的实际保护范围以权利要求为准。除非另外指出，本文中使用的所有技术和科学术语均具有与本申请所属领域中的普通技术人员通常所理解的相同的含义。本申请中引用的所有参考文献、专利、专利申请均通过整体引用并入本文。

25 定义

本申请中使用的术语“多枝梭菌”或“梭状芽孢杆菌”均指的是 *Clostridium ramosum* 菌，也称为 *Ramibacterium ramosum*，它是一种人体肠道共生菌，可以产生 IgA 蛋白酶。

本申请中使用的术语“蛋白酶”指的是具有分解蛋白质和肽的能力的酶。蛋白酶
5 可以通过在形成蛋白质的肽或多肽链中将氨基酸连接在一起的肽键进行水解，从而分解蛋白质。现有技术中已知多种方法测试某种蛋白酶的蛋白水解活性。例如，可以通过分析各种蛋白酶水解合适底物的能力的比较测定来确定蛋白酶的蛋白水解活性。用于蛋白水解活性分析的示例性底物包括，例如，二甲基酪蛋白、牛胶原、牛弹性蛋白等等。使用这些底物的比色测定也是现有技术中已知的（参见，例如
10 WO99/34011 和 US 6,376,450）。

本申请中使用的术语“IgA 蛋白酶”是指能够特异性切割或分解受试者（例如，人）的 IgA 免疫球蛋白分子（例如，IgA1 或 IgA2）的酶。例如，获得自或衍生自多枝梭菌（*Clostridium ramosum*）的 IgA 蛋白酶能够特异性切割 IgA1 和 IgA2 的第 221 位脯氨酸（Pro）和第 222 位缬氨酸（Val）之间的肽键，从而分解 IgA1 和 IgA2。

15 当提及的是多肽或蛋白质时，本申请中使用的术语“野生型”指的是在一个或多个氨基酸位置处不包括人为的取代、插入、缺失或修饰的天然存在的多肽或蛋白质；当提及的是核酸、核苷酸或多核苷酸时，本申请中使用的术语“野生型”指的是在一个或多个核苷酸位置处不包括人为的替换、插入、缺失或修饰的天然存在的核酸、核苷酸或多核苷酸。但是编码野生型多肽的多核苷酸不仅限于天然存在的多核苷酸，
20 也包括编码野生型多肽的任何多核苷酸。

本申请中使用的术语“AK183”是指多枝梭菌（*Clostridium ramosum*）的 AK183 株。*Clostridium ramosum* AK183 株产生的野生型 IgA 蛋白酶的氨基酸序列如 SEQ ID NO: 1 所示（其中，第 1-30 位氨基酸为信号肽）。

	10	20	30	40	50
	MTKKLMTKKI	TAIFLALYMA	ISVLPMTIQA	ASKPDIKVG	YVKMGVYNNA
	60	70	80	90	100
	SILWRCVSID	NNGPLMLADK	IVDTLAYDAK	TNDNSNSKSH	SRSYKRDDYG
5	110	120	130	140	150
	SNYWKDSNMR	SWLNSTAAEG	KVDWLCGNPP	KDGYVSGVGA	YNEKAGFLNA
	160	170	180	190	200
	FSKSEIAAMK	TVTQRLVSH	PEYNGGIVDG	DANSDLLYYT	DISEAVANYD
	210	220	230	240	250
10	SSYFETTTEK	VFLLDVKQAN	AVWKNLKGYY	VAYNNDGMAW	PYWLRTPVTD
	260	270	280	290	300
	CNHDMRYISS	SGQVGRYAPW	YSDLGVRPAF	YLDSEYFVTT	SGSGSQSSPY
	310	320	330	340	350
	IGSAPNKQED	DYTISEPAED	ANPDWNVSTE	QSIQLTLGPW	YSNDGKYSNP
15	360	370	380	390	400
	TIPVYTIQKT	RSDTENMVVV	VCGEYTKSQ	QGKFINDVKR	LWQDAMKYEP
	410	420	430	440	450
	YRSYADRFNV	YALCTASEST	FDNGGSTFFD	VIVDKYNSPV	ISNNLHGSQW
	460	470	480	490	500
20	KNHIFERCIG	PEFIEKIHDA	HIKKKCDPNT	IPSGSEYEPY	YYVHDYIAQF
	510	520	530	540	550
	AMVVNTKSDF	GGAYNNREYG	FHYFISPSDS	YRASKTFAHE	FGHGLLGLGD
	560	570	580	590	600
	EYSNGYLLDD	KELKSLNLSS	VEDPEKIKWR	QLLGFRNTYT	CRNAYGSKML
25	610	620	630	640	650
	VSSYECIMRD	TNYQFCEVCR	LQGFKRMSQL	VKDVDLYVAT	PEVKEYTGAY
	660	670	680	690	700
	SKPSDFTDLE	TSSYYNYTYN	RNDRLLSGNS	KSRFNTNMNG	KKIELRTVIQ
	710	720	730	740	750
30	NISDKNARQL	KFKMWIKHSD	GSVATDSSGN	PLQTVQTFDI	PVWNDKANFW
	760	770	780	790	800
	PLGALDHIKS	DFNSGLKSCS	LIYQIPSDAQ	LKSGDTVAFQ	VLDENGNVLA
	810	820	830	840	850
	DDNTETQRYT	TVSIQYKFED	GSEIPNTAGG	TFTVPYGTKL	DLTPAKTLYD

	860	870	880	890	900
	YEFIKVDGLN	KPIVSDGTVV	TYYYKNKNEE	HTHNLTLVAA	KAATCTTAGN
	910	920	930	940	950
	SAYYTCDGCD	KWFADATGSV	EITDKTSVKI	PAPGHTAGTE	WKSDDTNHWH
5	960	970	980	990	1000
	ECTVAGCGVI	IESTKSAHTA	GEWIVDTPAT	ATTAGTKHKE	CTVCHRVLET
	1010	1020	1030	1040	1050
	QPIPSTGTEL	KIIAGDNQIY	NKASGSDVTI	TCNGDFAKFT	GIKVDGGSVVD
	1060	1070	1080	1090	1100
10	SSNYTAVSGS	TVLTLKASYL	GTLTDGSHTI	TFVYTDGEAN	ANLTVRTAGS
	1110	1120	1130	1140	1150
	GHIHDYGTEW	KSNADNHWHE	CNCGDKKDEA	AHSFKWVVDK	EATATKKGSK
	1160	1170	1180	1190	1200
	HEECKICGYK	RSAVEIPATG	TSTAPTDTTK	PNDTTKPGNT	NGSEKSPQTG
15	1210	1220	1230		
	DNSNIFLWFA	LLFVSAAGVT	GITAYNKKKK	EHAE	
	(SEQ ID NO: 1)				

本申请中使用的术语“信号肽”是指可以参与蛋白质的成熟或前体形式的分泌或定向转运的氨基酸残基序列。信号肽通常位于前体或成熟蛋白序列的 N 末端。信号肽可以是内源的或外源的。成熟蛋白中一般不存在信号肽。通常情况下，在蛋白转运后，信号肽通过信号肽酶从所述蛋白中被切割掉。例如，SEQ ID NO: 1 所示的氨基酸序列去除掉 N 末端的信号肽之后形成的氨基酸序列如 SEQ ID NO: 42 所示。

25 ASKPDIKVGDYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
 NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKDG YV
 SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANS D L L Y Y
 TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
 YWLRTPVTDCNHDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
 30 QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
 TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQDAMKYEPYRS
 YADRFNVYALCTASESTFDNGGSTFFDVIVDKYNPVSNNLHGSQWKNHIFERC

IGPEFIEKIHDAHIKKKCDPNTIPSGSEYEPYVVDYIAQFAMVVNTKSDFGGAY
NNREYGFHYFISPSDSYRASKTFAHEFGHLLGLGDEYSNGYLLDDKELKSLNLS
SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDVLVYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS
5 GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNP
LQTVQTFDIPVWNDKANFWPLGALDHKSDFNSGLKSCSLIYQIPSDAQLKSGDT
VAFQVLDENGVLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTKL
DLTPAKTLYDYEFIKVDGLNKPIVSDGTVVTYYYKNKNEETHNLTLVAAKAAT
CTTAGNSAYYTCDGCDKWFADATGSVEITDKTSVKIPAPGHTAGTEWKSDDTN
10 HWHECTVAGCGVIIESTKSAHTAGEWIVDTPATATTAGTKHKECTVCHRVLETQ
PIPSTGTELKIIAGDNQIYNKASGSDVTITCNGDFAKFTGIKVDGSVVDSSNYTAV
SGSTVLTLKASYLGTLTDGSHTITFVYTDGEANANLTVRTAGSGHIHDYGTWK
SNADNHWHECNCGDKKDEAAHSFKWVVDKEATATKKGSKHEECKICGYKRSA
VEIPATGTSTAPTDTTKPNDTTKPGNTNGSEKSPQTGDNSNIFLWFALLFVSAAG
15 VTGITAYNKKKKEHAE (SEQ ID NO: 42)

本申请中使用的术语“受试者”包括人类和非人类的动物。非人类的动物包括所有的脊椎动物，例如哺乳动物和非哺乳动物。“受试者”也可以是家畜动物，例如牛、猪、羊、家禽和马；或啮齿类动物，例如大鼠、小鼠；或灵长类动物，例如猿（ape）、猴子、黑猩猩（chimpanzee）、大猩猩（gorilla）、猩猩（orangutan）、狒狒（baboon）；
20 或家养动物，例如狗和猫。“受试者”可以是雄性或者雌性，可以是老年、成年、青少年、儿童或者婴儿。人类“受试者”可以是高加索人、非洲人、亚洲人、闪族人，或其他种族或所述种族背景的混合。

本申请中使用的术语“蛋白”、“多肽”以及“肽”可以互换使用，是指氨基酸的聚合物。本申请所述的蛋白、多肽或肽可以含有天然的氨基酸，也可以含有非天然的
25 氨基酸，或氨基酸的类似物、模拟物。本申请所述的蛋白、多肽或肽可以通过本领域公知的任何方法获得，例如但不限于，通过天然分离、重组表达、化学合成等。

本申请所用的术语“氨基酸”是指含有氨基（-NH₂）和羧基（-COOH）官能团以及每个氨基酸特有的侧链的有机化合物。氨基酸名称在本申请中也以标准的单字母或三字母代码表示，总结如下：

名称	三字母代码	单字母代码
丙氨酸	Ala	A
精氨酸	Arg	R
天冬酰胺	Asn	N
天冬氨酸	Asp	D
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
蛋氨酸	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
缬氨酸	Val	V

在本申请中当“保守替换”用于氨基酸序列时，是指将一个氨基酸残基用另一个具有相似理化性质的侧链的氨基酸残基替代。例如，可以在具有疏水侧链的氨基酸残基之间（例如 Met、Ala、Val、Leu 和 Ile）、具有中性亲水侧链的氨基酸残基之间（例如 Cys、Ser、Thr、Asn 和 Gln）、具有酸性侧链的氨基酸残基之间（例如 Asp、Glu）、具有碱性侧链的氨基酸残基之间（例如 His、Lys 和 Arg）或具有芳香侧链的氨基酸残基之间（例如 Trp、Tyr 和 Phe）进行保守替换。本领域已知，保守替换通常不会引起蛋白构象结构的显著变化，因此能够保留蛋白质的生物活性。

本申请所述的术语“同源的”指当最佳比对时核酸序列（或其互补链）或氨基酸序列与另一条序列具有至少 60%（例如，至少 65%、70%、75%、80%、85%、88%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%）的序列同一性。

当“百分比(%)序列同一性”用于氨基酸序列(或核酸序列)时,是指在进行序列比对,并且必要时引入间隔使相同氨基酸(或核酸)数目达到最多后,在候选序列中,与参比序列相同的氨基酸(或核酸)残基占所述候选序列的氨基酸(或核酸)残基的百分比。换言之,可以通过用与其比较的参比序列相同的氨基酸残基(或碱基)数除以候选序列或参比序列(以较短者为准)中的氨基酸残基(或碱基)总数来计算氨基酸序列(或核酸序列)的百分比(%)序列同一性。所述氨基酸残基的保守替换可以认为或不认为是相同残基。可以通过本领域公开的工具,例如 BLASTN、BLASTp(美国国家生物技术信息中心网站(NCBI),也可参见 Altschul S.F. *et al.*, *J. Mol. Biol.*, 215: 403-410 (1990); Stephen F. *et al.*, *Nucleic Acids Res.*, 25: 3389-3402 (1997))、ClustalW2(欧洲生物信息研究所网站,可参见 Higgins D.G. *et al.*, *Methods in Enzymology*, 266: 383-402 (1996); Larkin M.A. *et al.*, *Bioinformatics* (Oxford, England), 23 (21): 2947-8 (2007))和 ALIGN 或 Megalign (DNASTAR) 软件,对序列进行比对以确定氨基酸(或核酸)序列的百分比序列同一性。本领域技术人员可以使用所述工具的默认参数或根据比对的需要适当调整参数,例如通过挑选合适的算法。

“分离的”物质已经经人工由自然状态改变。如果自然界中出现某种“分离的”混合物或物质,那么其已经被改变或脱离其原始状态,或二者均有发生。例如,某一活体动物体内天然存在的多核苷酸或多肽不是“分离的”,但如果这些多核苷酸或多肽与之在天然状态下共存的物质足够分离并以基本上纯的状态存在,则可以认为是“分离的”。“分离的核酸序列”是指分离的核酸分子的序列。在一些实施方式中,“分离的 IgA 蛋白酶截短体”是指纯度为至少 60%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 的 IgA 蛋白酶截短体,其中纯度由电泳方法(例如,SDS-PAGE、等电聚焦、毛细管电泳),或色谱法(例如,离子交换色谱或反相 HPLC)确定。

本申请中的术语“载体”是指可将遗传元件操作性地插入其中并使该遗传元件获得表达的一种运载工具,例如生产由该遗传元件编码的蛋白质、RNA 或 DNA,或者复制所述遗传元件。载体可用于转化、转导或转染宿主细胞,使其携带的遗传元件在宿主细胞内得以表达。举例来说,载体包括:质粒、噬菌粒、柯斯质粒(cosmid)、人工染色体如酵母人工染色体(YAC)、细菌人工染色体(BAC)或 P1 衍生的人

工染色体 (PAC)、噬菌体如 λ 噬菌体或 M13 噬菌体, 以及动物病毒等。载体可含
有多种控制表达的元件, 包括启动子序列、转录起始序列、增强子序列、选择元件
及报告基因。另外, 载体还可含有复制起始位点。载体还可包括协助其进入细胞的
成分, 包括但不限于, 病毒颗粒、脂质体或蛋白外壳。载体可以是表达载体或克隆
5 载体。本申请提供的载体 (例如表达载体) 含有本申请所述的编码 IgA 蛋白酶截短
体或融合蛋白的核酸序列、至少一个可操作地连接至所述核酸序列的启动子 (例如,
SV40、CMV、EF-1 α), 以及至少一个选择标记。

本申请中使用的对某种疾病、病症或状况的“治疗”或“疗法”包括预防或减轻某
种疾病、病症或状况, 降低某种疾病、病症或状况发生或发展的速度, 降低发展出
10 某种疾病、病症或状况的风险, 预防或延迟与某种疾病、病症或状况相关的症状发
展, 减少或终止与某种疾病、病症或状况相关的症状, 产生某种疾病、病症或状况
的完全或部分的逆转, 治愈某种疾病、病症或状况, 或以上的组合。

术语“药学上可接受的”表示指定的载体、媒介、稀释剂、赋形剂和/或盐通常在
化学和/或物理上与组成该制剂的其他成分相容, 并且在生理上与其受体相容。

15 术语“IgA 沉积相关疾病”是指与 IgA 免疫球蛋白以聚集或非聚集形式在受试者
的组织或器官中积累相关的疾病。例如, 包括但不限于, IgA 肾病、疱疹样皮炎、
亨诺赫-舍恩莱因紫癜 (又称 IgA 血管炎)、川崎病、紫癜性肾炎、IgA 血管炎肾损
害、IgA 类风湿因子阳性的类风湿性关节炎、IgA 型抗 GBM 病或 IgA 型 ANCA 相
关血管炎。

20 术语“IgA 肾病”是指肾脏内以 IgA 沉积为特征的肾脏疾病。

IgA 蛋白酶截短体

在一个方面, 本申请提供了一种分离的 IgA 蛋白酶截短体, 其包含获得自或衍
生自多枝梭菌 (*Clostridium ramosum*) 的野生型 IgA 蛋白酶的非天然截短片段, 或
者与所述非天然截短片段具有至少 70% 的序列同一性 (例如, 具有至少 75%、至少
25 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少
87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少
94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性)。在

某些实施方式中，与所述非天然截短片段具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性）的 IgA 蛋白酶截短体仍然保持 IgA 蛋白酶的功能或活性（例如，蛋白水解活性、特异性切割 IgA 的酶活性等）。

本申请中使用的术语“截短体”或“截短片段”是指从野生型多肽的一端或两端去除一个或多个氨基酸之后形成的肽。因此，本申请中的“截短体”或“截短片段”并不包括其对应的野生型多肽的全长，但是可以与野生型多肽的截短形式相比有一个或多个氨基酸取代、缺失、插入或修饰等。例如，“IgA 蛋白酶截短体”或“IgA 蛋白酶截短片段”可以包括从野生型的 IgA 蛋白酶的一端或者两端去除一个或多个氨基酸之后形成的肽，也可以包括与野生型的 IgA 蛋白酶的截短形式相比有一个或多个氨基酸取代、缺失、插入或修饰的肽。

在某些实施方式中，本申请所述的 IgA 蛋白酶截短体与其对应的野生型 IgA 蛋白酶相比有一个或多个氨基酸取代、缺失、插入或修饰。例如，在某些实施方式中，本申请所述的 IgA 蛋白酶截短体包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的非天然截短片段，其中所述非天然截短片段在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的基础上有氨基酸取代、缺失、插入或修饰，使得所述 IgA 蛋白酶截短体丧失或降低自酶切功能。

本申请中使用的术语“获得自”和“衍生自”不仅包括由所提及的生物体生产或可由其生产的蛋白质，而且还包括由从此类生物体分离的 DNA 序列编码并且在含有此类 DNA 序列的宿主生物体中生产的蛋白质，也包括由合成的和/或 cDNA 来源的 DNA 序列编码并且具有所提及的蛋白质的鉴定特征的蛋白质。例如，获得自或衍生自多枝梭菌的野生型 IgA 蛋白酶既包括由多枝梭菌天然生产的 IgA 蛋白酶，也包括通过使用遗传工程技术由用编码 IgA 蛋白酶的核酸转化的其他宿主细胞（例如，大肠杆菌）生产的 IgA 蛋白酶。

本申请中使用的术语“非天然截短片段”是指与多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶在自然环境下自酶切之后形成的截短片段具有不同的氨基酸

序列（例如，不同的氨基酸长度、不同的氨基酸类型等等）的片段。

- 在某些实施方式中，所述氨基酸取代、缺失、插入或修饰发生在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的天然自酶切位点。在某些实施方式中，所述氨基酸取代、缺失、插入或修饰发生在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的天然自酶切位点的上游 5 个位点以内（例如，在所述天然自酶切位点的上游 1 个位点、2 个位点、3 个位点、4 个位点或 5 个位点）。在某些实施方式中，所述氨基酸取代、缺失、插入或修饰发生在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的天然自酶切位点的下游 5 个位点以内（例如，在所述天然自酶切位点的下游 1 个位点、2 个位点、3 个位点、4 个位点或 5 个位点）。
- 在某些实施方式中，所述氨基酸取代、缺失、插入或修饰发生在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的天然自酶切位点的上游 5 个位点（例如，在所述天然自酶切位点的上游 1 个位点、2 个位点、3 个位点、4 个位点或 5 个位点）以内和下游 5 个位点以内（例如，在所述天然自酶切位点的下游 1 个位点、2 个位点、3 个位点、4 个位点或 5 个位点）。
- 在某些实施方式中，所述非天然截短片段为获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的 N 末端截短片段或 C 末端截短片段。

- 本申请中使用的术语“N 末端截短片段”是指包括多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的氨基端的氨基酸序列的截短片段。“氨基端”的起始位置可以是靠近多枝梭菌的野生型 IgA 蛋白酶的氨基酸序列的氨基端的任何位置，例如，可以从氨基端数起的第 1 位，也可以是从氨基端数起的其他位置。再例如，如果野生型 IgA 蛋白酶全长的氨基酸序列由 1000 个氨基酸组成，那么它的 N 末端截短片段的氨基端起始位置可以是它的氨基酸序列从氨基端起第 1 位至第 500 位之间的任何位置。

- 本申请中使用的术语“C 末端截短片段”是指包括多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的羧基端的氨基酸序列的截短片段。“羧基端”的终止位置可以是靠近多枝梭菌的野生型 IgA 蛋白酶的氨基酸序列的羧基端的任何位置，例如，可以从羧基端数起的第 1 位，也可以是从羧基端数起的其他位置。再

例如，如果野生型 IgA 蛋白酶全长的氨基酸序列由 1000 个氨基酸组成，那么它的 C 末端截短片段的羧基端终止位置可以是它的氨基酸序列从氨基端起第 501 位至第 1000 位之间的任何位置。

5 多枝梭菌 (*Clostridium ramosum*) 为 *Clostridium* 属中的众多菌种 (species) 之一，包括多种菌株 (strain)，例如 AK183、VPI-0496A、NCTC 10474 菌株等。在某些实施方式中，所述多枝梭菌 (*Clostridium ramosum*) 为 *Clostridium ramosum* AK183 株。

在某些实施方式中，所述 N 末端截短片段包含获得自或衍生自多枝梭菌 (*Clostridium ramosum*) 的野生型 IgA 蛋白酶的 N 末端第 31 位起的至少 760 个连续氨基酸的多肽片段，或者与所述多肽片段具有至少 70% 的序列同一性 (例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性)。在某些实施方式中，与所述多肽片段具有至少 70% 的序列同一性 (例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性) 的 N 末端截短片段仍然保持 IgA 蛋白酶的功能或活性 (例如，蛋白水解活性、特异性切割 IgA 的酶活性等)。

20 在某些实施方式中，本申请所述的 IgA 蛋白酶的非天然截短片段包含获得自或衍生自多枝梭菌 (*Clostridium ramosum*) 的野生型 IgA 蛋白酶的 N 末端第 335 位起的至少 456 个连续氨基酸的多肽片段，或者与所述多肽片段具有至少 90% 或至少 95% 的序列同一性 (例如，至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性)。在某些实施方式中，与所述多肽片段具有至少 90% 或至少 95% 的序列同一性 (例如，至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性) 的非天然截短片段仍然保持 IgA 蛋白酶的功能或活性 (例如，蛋白水解活性、特异性切割 IgA 的酶活性等)。

在某些实施方式中，所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的氨基酸序列如 SEQ ID NO: 1 所示。

除特殊说明外，本申请所提到的 IgA 蛋白酶的氨基酸位点即为对应于野生型 AK183 IgA 蛋白酶（其氨基酸序列如 SEQ ID NO: 1 所示）的氨基酸位点。例如，
5 本申请所提到的 AK183 IgA 蛋白酶的第 790 位对应于 SEQ ID NO: 1 的第 790 个位点。除特殊说明外，本申请所提到的 AK183 IgA 蛋白酶截短体的命名规则为 AK183(对应于 SEQ ID NO: 1 的起始位点-对应于 SEQ ID NO: 1 的终止位点)。例如，AK183(31-790)指的是 SEQ ID NO: 1 的第 31 位至第 790 位氨基酸形成的 IgA 蛋白酶截短体。

10 在某些实施方式中，申请所述的 IgA 蛋白酶的天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 730 位至第 840 位之间。在某些实施方式中，本申请所述的 IgA 蛋白酶的天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 710 位至 830 位之间、第 720 位至第 820 位之间、第 730 位至 810 位之间、第 740 位至 800 位之间、第 750 位至第 790 位之间、第 791 位至第 780 位之间或者第 792
15 位至第 797 位之间。在某些实施方式中，所述天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 790 位、第 791 位、第 792 位、第 793 位、第 794 位、第 795 位、第 796 位、第 797 位、第 798 位、第 799 位或第 800 位。

在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的至少 760 个连续氨基酸的多肽片段。例如，在某些
20 实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的至少 761 个、至少 762 个、至少 763 个、至少 764 个、至少 765 个、至少 766 个、至少 767 个、至少 768 个、至少 769 个、至少 770 个、至少 771 个、至少 772 个、至少 773 个、至少 774 个、至少 775 个、至少 776 个、至少 777 个、至少 778 个、至少 779 个、至少 780 个、至少 781 个、至少 782 个、至少 783
25 个、至少 784 个、至少 785 个、至少 786 个、至少 787 个、至少 788 个、至少 789 个、至少 790 个、至少 791 个、至少 792 个、至少 793 个、至少 794 个、至少 795 个、至少 796 个、至少 797 个、至少 798 个、至少 799 个、至少 800 个、至少 801 个、至少 802 个、至少 803 个、至少 804 个、至少 805 个、至少 806 个、至少 807 个、至少 808 个、至少 809 个、至少 810 个、至少 850 个、至少 860 个、至少 870

个、至少 880 个、至少 890 个、至少 900 个、至少 910 个、至少 920 个、至少 930 个、至少 940 个、至少 950 个、至少 960 个、至少 970 个、至少 980 个、至少 990 个、至少 1000 个、至少 1050 个、至少 1100 个、至少 1150 个、至少 1200 个连续氨基酸的多肽片段。

- 5 在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 760 个连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 761 个连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 762 个连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 768 个连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 777 个连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 786 个连续氨基酸的多肽片段。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的 803 个连续氨基酸的多肽片段。

- 在某些实施方式中，本申请提供的 IgA 蛋白酶截短体包含选自下组的多肽片段：如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 798 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 807 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 816 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 833 位氨基酸以及与其具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性）的多肽片段。在某些实施方式中，与上述多肽片段具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少

94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性) 的 IgA 蛋白酶截短体仍然保持 IgA 蛋白酶的功能或活性 (例如, 蛋白水解活性、特异性切割 IgA 的酶活性等)。

在某些实施方式中, 本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位起的至少 456 个连续氨基酸的多肽片段。例如, 在某些实施方式中, 本申请提供的 IgA 蛋白酶截短体包含如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位起的至少 457 个、至少 458 个、至少 459 个、至少 460 个、至少 461 个、至少 462 个、至少 463 个、至少 464 个、至少 465 个、至少 466 个、至少 467 个、至少 468 个、至少 469 个、至少 470 个、至少 471 个、至少 472 个、至少 473 个、至少 474 个、至少 475 个、至少 476 个、至少 477 个、至少 478 个、至少 479 个、至少 480 个、至少 481 个、至少 482 个、至少 483 个、至少 484 个、至少 485 个、至少 486 个、至少 487 个、至少 488 个、至少 489 个、至少 490 个、至少 491 个、至少 492 个、至少 493 个、至少 494 个、至少 495 个、至少 496 个、至少 497 个、至少 498 个、至少 499 个、至少 500 个、至少 550 个、至少 600 个、至少 650 个、至少 700 个、至少 750 个、至少 800 个、至少 850 个或者至少 900 个连续氨基酸的多肽片段。

在某些实施方式中, 本申请提供的 IgA 蛋白酶截短体包含选自下组的多肽片段: 如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 816 位氨基酸以及与其具有至少 90% 或至少 95% 的序列同一性 (例如, 至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性) 的多肽片段。在某些实施方式中, 与所述多肽片段具有至少 90% 或至少 95% 的序列同一性 (例如, 至少 91%、至少 92%、至少 93%、

至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性) 的 IgA 蛋白酶截短体仍然保持 IgA 蛋白酶的功能或活性 (例如, 蛋白水解活性、特异性切割 IgA 的酶活性等)。

在某些实施方式中, 本申请提供了 AK183(31-790)截短体, 其氨基酸序列如 SEQ ID NO: 14 所示。

ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWLCGNPPKDG YV
SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANSDLLYY
TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
10 YWLRTPVTDCNHDMRYISSSGQVG RYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS
YADRFNVYALCTASESTFDNGGSTFFDVIVDKYN SPVISNNLHGSQWKNHIFERC
IGPEFIEKIHDAHIKKKCDPNTIPSGSEYEPYYVHDYIAQFAMVVNTKSDFGGAY
15 NNREYGFHYFISPSDSYRASKTFAHEFGHLLGLGDEYSNGYLLDDKELKSLNLS
SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS
GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNP
LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNSGLKSCSLIYQIPSDAQLKSGDT
20 VAFQ (SEQ ID NO: 14)

在某些实施方式中, 本申请提供了 AK183(31-791)截短体, 其氨基酸序列如 SEQ ID NO: 15 所示。

ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWLCGNPPKDG YV
25 SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANSDLLYY
TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
YWLRTPVTDCNHDMRYISSSGQVG RYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS
30 YADRFNVYALCTASESTFDNGGSTFFDVIVDKYN SPVISNNLHGSQWKNHIFERC
IGPEFIEKIHDAHIKKKCDPNTIPSGSEYEPYYVHDYIAQFAMVVNTKSDFGGAY

NNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS
SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS
GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNP
5 LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT
VAFQV (SEQ ID NO: 15)

在某些实施方式中,本申请提供了 AK183(31-792)截短体,其氨基酸序列如 SEQ
ID NO: 16 所示。

ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
10 NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKDG YV
SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANDSLLYY
TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
YWL RTPVTDCNHDMRYISSSGQVG RYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
15 TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQDAMKYEPYRS
YADRFNVYALCTASESTFDNGGSTFFDVIVDKYNPVISNNLHGSQWK NHIFERC
IGPEFIEKIHDAHIKKCDPNTIPSGSEYEPYVYVHDYIAQFAMVVNTKSDFGGAY
NNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS
SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
20 FKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS
GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNP
LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT
VAFQVL (SEQ ID NO: 16)

在某些实施方式中,本申请提供了 AK183(31-798)截短体,其氨基酸序列如 SEQ
25 ID NO: 17 所示。

ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKDG YV
SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANDSLLYY
TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
30 YWL RTPVTDCNHDMRYISSSGQVG RYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP

TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS
YADRFNVYALCTASESTFDNNGGSTFFDVIVDKYNPVISNNLHGSQWKNHIFERC
IGPEFIEKIHDAHIKKKCDPNTIPSGSEYEPYVVDYIAQFAMVVNTKSDFGGAY
NNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS
5 SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDLDLYVATPEVKEYTGAYSKPSDFTDLETSSYNYTYNRNDRLLS
GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNP
LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT
VAFQVLDENGNV (SEQ ID NO: 17)

10 在某些实施方式中,本申请提供了 AK183(31-807)截短体,其氨基酸序列如 SEQ ID NO: 18 所示。

ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWLGCNPPKDG YV
SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANDSLLYY
15 TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
YWLRTPVTD CNHDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS
YADRFNVYALCTASESTFDNNGGSTFFDVIVDKYNPVISNNLHGSQWKNHIFERC
20 IGPEFIEKIHDAHIKKKCDPNTIPSGSEYEPYVVDYIAQFAMVVNTKSDFGGAY
NNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS
SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDLDLYVATPEVKEYTGAYSKPSDFTDLETSSYNYTYNRNDRLLS
GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNP
25 LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT
VAFQVLDENGNVLADDNTETQ (SEQ ID NO: 18)

在某些实施方式中,本申请提供了 AK183(31-816)截短体,其氨基酸序列如 SEQ ID NO: 19 所示。

ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTND
30 NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWLGCNPPKDG YV
SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANDSLLYY

TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
YWL RTPVTDCNHDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS
5 YADRFNVYALCTASESTFDNNGGSTFFDVIVDKYNPVISNNLHGSQWKNHIFERC
IGPEFIEKIHD AHIKKKCDPNTIPSGSEYEPYYYVHDYIAQFAMVVNTKSDFGGAY
NNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS
SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS
10 GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDG SVATDSSGNP
LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT
VAFQVLDENG NVLADDNTETQRYTTVSIQY (SEQ ID NO: 19)

在某些实施方式中,本申请提供了 AK183(31-833)截短体,其氨基酸序列如 SEQ ID NO: 20 所示。

15 ASKPDIKVGDYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTLAYDAKTND
NSNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKDG YV
SGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANS D LLYY
TDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWP
YWL RTPVTDCNHDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGS
20 QSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNP
TIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS
YADRFNVYALCTASESTFDNNGGSTFFDVIVDKYNPVISNNLHGSQWKNHIFERC
IGPEFIEKIHD AHIKKKCDPNTIPSGSEYEPYYYVHDYIAQFAMVVNTKSDFGGAY
NNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS
25 SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQG
FKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS
GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIKHSDG SVATDSSGNP
LQTVQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT
VAFQVLDENG NVLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFT (SEQ ID NO:
30 20)

在某些实施方式中,本申请提供了 AK183(285-790)截短体,其氨基酸序列如

SEQ ID NO: 43 所示。在某些实施方式中，本申请提供了 AK183(285-791)截短体，其氨基酸序列如 SEQ ID NO: 44 所示。在某些实施方式中，本申请提供了 AK183(285-792)截短体，其氨基酸序列如 SEQ ID NO: 45 所示。在某些实施方式中，本申请提供了 AK183(285-816)截短体，其氨基酸序列如 SEQ ID NO: 46 所示。在某些实施方式中，本申请提供了 AK183(330-790)截短体，其氨基酸序列如 SEQ ID NO: 47 所示。在某些实施方式中，本申请提供了 AK183(330-791)截短体，其氨基酸序列如 SEQ ID NO: 48 所示。在某些实施方式中，本申请提供了 AK183(330-792)截短体，其氨基酸序列如 SEQ ID NO: 49 所示。在某些实施方式中，本申请提供了 AK183(335-790)截短体，其氨基酸序列如 SEQ ID NO: 50 所示。在某些实施方式中，本申请提供了 AK183(335-791)截短体，其氨基酸序列如 SEQ ID NO: 51 所示。在某些实施方式中，本申请提供了 AK183(335-792)截短体，其氨基酸序列如 SEQ ID NO: 52 所示。

SEQ ID NO: 43 ~ 52 的序列如下所示。

SEQ ID NO	序列描述	氨基酸序列
43	AK183(285-790)	EYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYEPYYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSVVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNIISKDKNARQLKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDFNSGLKSCSLIYQIPSDAQ

		LKSGDTVAFQ
44	AK183(285-791)	EYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDA NPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTI QKTRSDTENMVVVVCGEYTKSQQGFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGG STFFDVIVDKYNPVISNNLHGSQWKNHIFERCIGPE FIEKIHDAAHIKKKCDPNTIPSGSEYEPYYYVHDYIAQ FAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSY ECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLYV ATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRND RLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQ LKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWN DKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQ LKSGDTVAFQV
45	AK183(285-792)	EYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDA NPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTI QKTRSDTENMVVVVCGEYTKSQQGFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGG STFFDVIVDKYNPVISNNLHGSQWKNHIFERCIGPE FIEKIHDAAHIKKKCDPNTIPSGSEYEPYYYVHDYIAQ FAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS SVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSY ECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLYV ATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRND RLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQ LKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWN DKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQ LKSGDTVAFQVL

46	AK183(285-816)	EYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDA NPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTI QKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGG STFFDVIVDKYNSPVISNNLHGSQWKNHIFERCIGPE FIEKIHDAHIKKKCDPNTIPSGSEYEPYYYVHDYIAQ FAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLS SVEDPEKIKWRQLLGRNTYTCRNAYGSKMLVSSY ECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLYV ATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRND RLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQ LKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWN DKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQ LKSGDTVAFQVLDENG NVLADDNTETQRYTTVSIQ Y
47	AK183(330-790)	EQSIQLTLGPWYSNDGKYSNPTIPVYTIQKTRSDTEN MVVVVCGEGYTKSQQGKFINDVKRLWQDAMKYE PYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDK YNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHI KKKCDPNTIPSGSEYEPYYYVHDYIAQFAMVVNTK SDFGGAYNNREYGFHYFISPSDSYRASKTFAHEFGH GLLGLGDEYSNGYLLDDKELKSLNLSVEDPEKIK WRQLLGRNTYTCRNAYGSKMLVSSYECIMRDTN YQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEY TGAYSKPSDFTDLETSSYYNYTYNRNDRLLSGNSKS RFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPL GALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAF Q
48	AK183(330-791)	EQSIQLTLGPWYSNDGKYSNPTIPVYTIQKTRSDTEN

		MVVVVCGEGYTKSQQGKFINDVKRLWQDAMKYE PYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDK YNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHI KKKCDPNTIPSGSEYEPYYYVHDYIAQFAMVVNTK SDFGGAYNNREYGFHYFISPSDSYRASKTFAHEFGH GLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIK WRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTN YQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEY TGAYSKPSDFTDLETSSYYNYTYNRNDRLLSGNSKS RFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPL GALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAF QV
49	AK183(330-792)	EQSIQLTLGPWYSNDGKYSNPTIPVYTIQKTRSDTEN MVVVVCGEGYTKSQQGKFINDVKRLWQDAMKYE PYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDK YNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHI KKKCDPNTIPSGSEYEPYYYVHDYIAQFAMVVNTK SDFGGAYNNREYGFHYFISPSDSYRASKTFAHEFGH GLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIK WRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTN YQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEY TGAYSKPSDFTDLETSSYYNYTYNRNDRLLSGNSKS RFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPL GALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAF QVL
50	AK183(335-790)	LTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVV VVCGEGYTKSQQGKFINDVKRLWQDAMKYE PYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDK YNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHI KKK

		DPNTIPSGSEYEPYYYYVHDYIAQFAMVVNTKSDFG GAYNNREYGFHYFISPSDSYRASKTFAHEFGHGLLG LGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWRQL LGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCE VCRLQGFKRMSQLVKDVDLYVATPEVKEYTGAYS KPSDFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTN MNGKKIELRTVIQNIISKDNARQLKFKMWIKHSDGS VATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDH IKSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQ
51	AK183(335-791)	LTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVV VVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS YADRFNVYALCTASESTFDNGGSTFFDVIVDKYNSP VISNNLHGSQWKNIHIFERCIGPEFIEKIHDAHIKKKC DPNTIPSGSEYEPYYYYVHDYIAQFAMVVNTKSDFG GAYNNREYGFHYFISPSDSYRASKTFAHEFGHGLLG LGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWRQL LGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCE VCRLQGFKRMSQLVKDVDLYVATPEVKEYTGAYS KPSDFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTN MNGKKIELRTVIQNIISKDNARQLKFKMWIKHSDGS VATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDH IKSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQV
52	AK183(335-792)	LTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVV VVCGEGYTKSQQGFINDVKRLWQDAMKYEPYRS YADRFNVYALCTASESTFDNGGSTFFDVIVDKYNSP VISNNLHGSQWKNIHIFERCIGPEFIEKIHDAHIKKKC DPNTIPSGSEYEPYYYYVHDYIAQFAMVVNTKSDFG GAYNNREYGFHYFISPSDSYRASKTFAHEFGHGLLG LGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWRQL LGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCE VCRLQGFKRMSQLVKDVDLYVATPEVKEYTGAYS

		KPSDFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTN MNGKKIELRTVIQNISDKNARQLKFKMWIKHSDGS VATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDH IKSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQVL
--	--	---

在某些实施方式中，本申请提供的 IgA 蛋白酶截短体在上述多肽片段的氨基酸序列基础上，在一个或多个位点（例如，在 1 个、2 个、3 个、4 个、5 个或更多个位点）具有氨基酸的保守替换。氨基酸残基的保守替换是指性质相似的氨基酸之间的替换，例如极性氨基酸之间的替换（如谷氨酰胺和天冬酰胺之间的替换），疏水性氨基酸之间的替换（如亮氨酸、异亮氨酸、蛋氨酸和缬氨酸之间的替换），以及带相同电荷的氨基酸之间的替换（如精氨酸、赖氨酸和组氨酸之间的替换，或者谷氨酸和天冬氨酸之间的替换）等。在某些实施方式中，本申请所述的 IgA 蛋白酶截短体与 SEQ ID NO: 14、SEQ ID NO: 15、SEQ ID NO: 16、SEQ ID NO: 17、SEQ ID NO: 18、SEQ ID NO: 19、SEQ ID NO: 20、SEQ ID NO: 43、SEQ ID NO: 44、SEQ ID NO: 45、SEQ ID NO: 46、SEQ ID NO: 47、SEQ ID NO: 48、SEQ ID NO: 49、SEQ ID NO: 50、SEQ ID NO: 51 或者 SEQ ID NO: 52 所示的氨基酸序列相比，在 1 个、2 个、3 个、4 个、5 个、6 个、7 个、8 个、9 个、10 个、11 个、15 个、20 个或更多个位点具有氨基酸的保守替换。

在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置有氨基酸突变。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位有氨基酸突变。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 862 位有氨基酸突变。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 931 位和第 933 位有氨基酸突变。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 978 位有氨基酸突变。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 第 1002 位和第 1004 位有氨基酸突变。

在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置突变为甘氨酸。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 844

位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置的脯氨酸（P）突变为甘氨酸（G）。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位的脯氨酸突变为甘氨酸。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 862 位的脯氨酸突变为甘氨酸。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 931 位和第 933 位的脯氨酸突变为甘氨酸。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 的第 978 位的脯氨酸突变为甘氨酸。在某些实施方式中，所述多肽片段在对应于 SEQ ID NO: 1 第 1002 位和第 1004 位的脯氨酸突变为甘氨酸。

在某些实施方式中，所述多肽片段的氨基酸序列如 SEQ ID NO: 53（也称为“PA-GA Mut”）、SEQ ID NO: 54（也称为“PI-GI Mut”）、SEQ ID NO: 55（也称为“PAP-GAG Mut”）、SEQ ID NO: 56（也称为“PAT-GAT Mut”）或者 SEQ ID NO: 57（也称为“PIP-GIG Mut”）所示。

SEQ ID NO: 53 ~ 57 的序列如下所示。

SEQ ID NO	序列描述	氨基酸序列
53	PA-GA Mut	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPL MLADKIVDTLAYDAKTNDNSNSKSHSR SYKRDDY GSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKD GYVSGVGAYNEKAGFLNAFSKSEIAAMKT VQTQ RSL VSHPEYNKGIVDGDANSDLLYYTDISEAV ANYDSS YFETTTEKVFLLDVKQANAVWKNLKGYYVAYNND GMAWPYWL RTPVTDCNHDMRYISSSGQVGRYAPW YSDLGVRPAFYLDSEYFVTTS GSGSQSSPYIGSAPNK QEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSN DGKYSNPTIPVYTIQKTRSDTENMVVVVC GEGYTK SQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYA LCTASESTFDNGGSTFFDVIVDKYNSPVISNNLHGSQ WKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEY EPYYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGF

		HYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYL LDDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTCRN AYGSKMLVSSYECIMRDTNYQFCEVCRLQGFKRMS QLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSS YYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTV IQNISDKNARQLKFKMWIKHSDGSVATDSSGNPLQT VQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSC SLIYQIPSDAQLKSGD TVAFQVLDENGNVLADDNTE TQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTKLD LT<u>G</u>AKTLYDYEFIKVDGLNKPIVSDGTVVTTYYYKN KNEETHNLTLVAAKAATCTTAGNSAYYTCDGCD KWFADATGSVEITDKTSVKIPAPGHTAGTEWKSDD TNHWECTVAGCGVIIESTKSAHTAGEWIVDTPAT ATTAGTKHKECTVCHRVLETQPISTGTTELKIIAGDN QIYNKASGSDVTITCNGDFAKFTGIKVDGSVVDSSN YTAVSGSTVLTLKASYLGTLTDGSHTITFVYTDGEA NANLTVRTAGSGHIHDYGTEWKSADNHWHECNC GDKKDEAAHSFKWVVDKEATATKKGSKHEECKIC GYKRSAVEIPATGTSTAPTDTTKPNDTTKPGNTNGS EKSPQTGDNS
54	PI-GI Mut	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPL MLADKIVDTLAYDAKTNDNSNSKSHSR SYKRDDY GSNYWKDSNMRSWLNSTAAEGKVDWLCGNPPKD GYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSL VSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSS YFETTTEKVFLLDVKQANAVWKNLKGYYVAYNND GMAWPYWL RTPVTDCNHDMRYISSSGQVGRYAPW YSDLGVRPAFYLDSEYFVTTSGSGSQSSPYIGSAPNK QEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSN DGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTK SQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYA LCTASESTFDNGGSTFFDVIVDKYNSPVISNNLHGSQ

		WKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEY EPYVVHDIYAQFAMVVNTKSDFGGAYNNREYGF HYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYL LDDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTCRN AYGSKMLVSSYECIMRDTNYQFCEVCRLQGFKRMS QLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSS YYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTV IQNISDKNARQLKFKMWIKHSDGSVATDSSGNPLQT VQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSC SLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTE TQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTKLD LTPAKTLYDYEFIKVDGLNK<u>G</u>IVSDGTVVTYYYKN KNEETHNLTLVAAKAATCTTAGNSAYYTCDGCD KWFADATGSVEITDKTSVKIPAPGHTAGTEWKSDD TNHWECTVAGCGVIIESTKSAHTAGEWIVDTPAT ATTAGTKHKECTVCHRVLETQPISTGTTELKIIAGDN QIYNKASGSDVTITCNGDFAKFTGIKVDGSVVDSSN YTAVSGSTVLTLKASYLGTLTDGSHTITFVYTDGEA NANLTVRTAGSGHIIHDYGTEWKSADNHWHECNC GDKKDEAAHSFKWVVDKEATATKKGSKHEECKIC GYKRSAVEIPATGTSTAPTDTTKPNDTTKPGNTNGS EKSPQTGDNS
55	PAP-GAG Mut	ASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPL MLADKIVDTLAYDAKTNDNSNSKSHSRSYKRDDY GSNYWKDSNMRSWLNSTAAEGKVDWLCGNPPKD GYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSL VSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSS YFETTTEKVFLLDVKQANAVWKNLKGYYVAYNND GMAWPYWL RTPVTD CNHDMRYISSSGQVGRYAPW YSDLGVRPAFYLDSEYFVTTSGSGSQSSPYIGSAPNK QEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSN DGKYSNPTIPVYTIQKTRSDTENMVVVVCGEYTK

		SQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYA LCTASESTFDNGGSTFFDVIVDKYNPVISNNLHGSQ WKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEY EPYYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGF HYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYL LDDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTCRN AYGSKMLVSSYECIMRDTNYQFCEVCRLQGFKRMS QLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSS YYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTV IQNISDKNARQLKFKMWIKHSDGSVATDSSGNPLQT VQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSC SLIYQIPSDAQLKSGD TVAFQVLDENGNVLADDNTE TQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTKLD LTPAKTLYDYEFIKVDGLNKPIVSDGTVVTTYYYKN KNEETHNLTLVAAKAATCTTAGNSAYYTCDGCD KWFADATGSVEITDKTSVKI<u>CAG</u>GHTAGTEWKSDD TNHWECTVAGCGVIIESTKSAHTAGEWIVDTPAT ATTAGTKHKECTVCHRVLETQPISTGTTELKIIAGDN QIYNKASGSDVTITCNGDFAKFTGIKVDGSVVDSSN YTAVSGSTVLTALKASYLGTLTDGSHTITFVYTDGEA NANLTVRTAGSGHIHDYGTEWKSADNHNWHECNC GDKKDEAAHSFKWVVDKEATATKKGSKHEECKIC GYKRSAVEIPATGTSTAPTDTTKPNDTTKPGNTNGS EKSPQTGDNS
56	PAT-GAT Mut	ASKPDIKVGDYVKMGVYNNASILWRCVSIDNNGPL MLADKIVDTLAYDAKTNDNSNSKSHSRSYKRDDY GSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKD GYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSL VSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSS YFETTTEKVFLLDVKQANAVWKNLKGYYVAYNND GMAWPYWL RTPVTD CNHDMRYISSSGQVGRYAPW YSDLGVRPAFYLDSEYFVTTSGSGSQSSPYIGSAPNK

		QEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSN DGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTK SQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYA LCTASESTFDNGGSTFFDVIVDKYNSPVISNNLHGSQ WKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEY EPYYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGF HYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYL LDDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTCRN AYGSKMLVSSYECIMRDTNYQFCEVCRLQGFKRMS QLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSS YYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTV IQNISDKNARQLKFKMWIKHSDGSVATDSSGNPLQT VQTFDIPVWNDKANFWPLGALDHIKSDFNGLKSC SLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTE TQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTKLD LTPAKTLYDYEFIKVDGLNKPIVSDGTVVTTYYYKN KNEETHNLTLVAAKAATCTTAGNSAYYTCDGCD KWFADATGSVEITDKTSVKIPAPGHTAGTEWKSDD TNHWHECTVAGCGVIIESTKSAHTAGEWIVDT<u>G</u>AT ATTAGTKHKECTVCHRVLETQPIPSTGTTELKIIAGDN QIYNKASGSDVTITCNGDFAKFTGIKVDGSVVDSSN YTAVSGSTVLTLKASYLGTLTDGSHTITFVYTDGEA NANLTVRTAGSGHIIHDYGTEWKSADNHWHECNC GDKKDEAAHSFKWVVDKEATATKKGSKHEECKIC GYKRSAVEIPATGTSTAPTDTTKPNDTTKPGNTNGS EKSPQTGDNS
57	PIP-GIG Mut	ASKPDIKVGDYVKMGVYNNASILWRCVSIDNNGPL MLADKIVDTLAYDAKTNDNSNSKSHSRSYKRDDY GSNYWKDSNMRSWLNSTAAEGKVDWL CGNPPKD GYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQRSL VSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSS YFETTTEKVFLLDVKQANAVWKNLKGYYVAYNND

		GMAWPYWL RTPVTDCNHDMRYISSSGQVGRYAPW YSDLGVRPAFYLDSEYFVTTSGSGSQSSPYIGSAPNK QEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSN DGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTK SQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYA LCTASESTFDNGGSTFFDVIVDKYNPVISNNLHGSQ WKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEY EPYYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGF HYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYL LDDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTCRN AYGSKMLVSSYECIMRDTNYQFCEVCRLQGFKRMS QLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSS YYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTV IQNISDKNARQLKFKMWIKHSDGSVATDSSGNPLQT VQTFDIPVWNDKANFWPLGALDHIKSDFN SGLKSC SLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTE TQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTKLD LTPAKTLYDYEFIKVDGLNKPIVSDGTVVYYYYKN KNEETHNLT LVA AKAATCTTAGNSAYYTCDGCD KWFADATGSVEITDKTSVKIPAPGHTAGTEWKSDD TNHWECTVAGCGVIIESTKSAHTAGEWIVDTPAT ATTAGTKHKECTVCHRVLETQGI ST GTTELKIIAGD NQIYNKASGSDVTITCNGDFAKFTGIKVDGSVVDSS NYTAVSGSTVLT LKASYLGTLTDGSHTITFVYTDGE ANANLTVRTAGSGHIHDYGTEWKSNAADNHWHECN CGDKKDEAAHSFKWVVDKEATATKKGSKHEECKI CGYKRSAVEIPATGTSTAPTDTTKPNDTTKPGNTNG SEKSPQTGDNS
--	--	--

在不影响活性的前提下,本申请提供的 IgA 蛋白酶截短体还可以含有非天然的氨基酸。非天然的氨基酸包括例如,β-氟代丙氨酸、1-甲基组氨酸、γ-亚甲基谷氨酸、α-甲基亮氨酸、4,5-脱氢赖氨酸、羟基脯氨酸、3-氟代苯基丙氨酸、3-氨基酪氨

酸、4-甲基色氨酸等。

本申请提供的 IgA 蛋白酶截短体也可以使用本领域公知的方法进行修饰。例如，但不限于，PEG 化、糖基化、氨基端修饰、脂肪酰化、羧基端修饰、磷酸化、甲基化等。本领域技术人员可以理解，本申请提供的 IgA 蛋白酶截短体使用本领域公知的方法进行修饰之后，仍然保留了与 IgA 蛋白酶或 IgA 蛋白酶截短体基本上相似的功能。

在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA 的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA 重链的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA 重链 CH1 与铰链区交汇处的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体具有特异性切割人 IgA1 的酶活性。

在某些实施方式中，本申请提供的 IgA 蛋白酶截短体在上述多肽片段的氨基酸序列基础上，在一个或多个位点具有氨基酸的保守替换，但仍然具有切割人 IgA（例如，IgA1）的酶活性。在某些实施方式中，本申请提供的 IgA 蛋白酶截短体与上述多肽片段具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性），并且仍然具有切割人 IgA（例如，IgA1）的酶活性。

20 融合蛋白

在另一方面，本申请提供了一种融合蛋白，其包含第一多肽和第二多肽，其中所述第一多肽包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的全长、获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶去掉信号肽之后形成的多肽或者本申请所述的 IgA 蛋白酶截短体，所述第二多肽包含用于延长所述第一多肽在受试者体内半衰期的氨基酸序列。在某些实施方式中，所述第一多肽包含如 SEQ ID NO: 1 或者 SEQ ID NO: 42 所示的序列。在某些实施方式中，所述第二多肽位于所述第一多肽的 N 末端。在某些实施方式中，所述第二多肽位于所述第一多肽的 C 末端。

在某些实施方式中，所述第一多肽和所述第二多肽之间通过连接子连接。在某些实施方式中，所述第一多肽和所述第二多肽之间直接连接（即，不通过连接子连接）。本申请中使用的术语“连接子”或“接头”是指具有 1、2、3、4 或 5 个氨基酸残基，或长度介于 5 和 15、20、30、50 或更多个氨基酸残基之间的人工氨基酸序列，通过肽键连接，并用于连接一个或多个多肽。连接子可能有也可能没有二级结构。连接子序列在本领域是已知的，例如，参见 Holliger *et al.*, *Proc. Natl. Acad. Sci. USA* 90:6444-6448 (1993); Poljak *et al.*, *Structure* 2:1121-1123 (1994)。

在某些实施方式中，所述连接子选自下组：可切割连接子、不可切割连接子、肽连接子、柔性连接子、刚性连接子、螺旋连接子和非螺旋连接子。可以使用本领域已知的任何合适的连接子。在某些实施方式中，所述连接子包含肽连接子。例如，本申请中的有用连接子可能富含甘氨酸和丝氨酸残基。示例包括具有包含苏氨酸/丝氨酸和甘氨酸的单个或重复序列的连接子，例如 GGGS (SEQ ID NO: 21) 或 GGGGS (SEQ ID NO: 22)、GGGGGS (SEQ ID NO: 86) 或 GGGGGGS (SEQ ID NO: 87) 或其串联重复（例如，2、3、4、5、6、7、8、9、10 或更多重复）。在某些实施方式中，本申请中使用的连接子包括 GGCGGCGGTGGATCC (SEQ ID NO: 23)。可选地，所述连接子可能是包含如 GGCGGCGGTGGATCC (SEQ ID NO: 23) 所示的氨基酸序列的一个或多个顺序或串联重复的长肽链。在某些实施方式中，所述连接子包括 SEQ ID NO: 23 的 1、2、3、4、5、6、7、8、9、10 或更多顺序或串联重复。在某些实施方式中，所述连接子包括选自下组的氨基酸序列或由选自下组的氨基酸序列组成：与 SEQ ID No: 21、22、23 中的任何一个具有至少 80%、至少 85%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性的氨基酸序列。

在某些实施方式中，本申请中使用的连接子包括如 SEQ ID NO: 58 (EEKKKEKEKEEQEERETK) 所示的氨基酸序列。可选地，所述连接子可能是包含如 SEQ ID NO: 58 所示的氨基酸序列的一个或多个顺序或串联重复的长肽链。在某些实施方式中，所述连接子包括 SEQ ID NO: 58 的 1、2、3、4、5、6、7、8、9、10 或更多顺序或串联重复。在某些实施方式中，所述连接子包括与 SEQ ID NO: 58 具有至少 80%、至少 85%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性的氨基酸序列。

列。

在某些实施方式中，本申请中使用的连接子包括如 SEQ ID NO: 59 (HHHHHHHHHH)所示的氨基酸序列。在某些实施方式中，所述连接子包括与 SEQ ID NO: 59 具有至少 80%、至少 85%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99%的序列同一性的氨基酸序列。

在某些实施方式中，所述第二多肽选自 Fc 结构域和白蛋白。在某些实施方式中，所述 Fc 结构域包含铰链区。在某些实施方式中，所述 Fc 结构域包含下游铰链区 (lower hinge)。在某些实施方式中，所述 Fc 结构域包含核心铰链区 (core hinge region) 和下游铰链区 (lower hinge)。在某些实施方式中，所述 Fc 结构域包含上游铰链区 (upper hinge region)、核心铰链区 (core hinge region) 和下游铰链区 (lower hinge)。在某些实施方式中，所述 Fc 结构域不包含铰链区。在某些实施方式中，所述 Fc 结构域来源于人 IgG Fc 结构域。在某些实施方式中，所述 Fc 结构域来源于人 IgG1 Fc 结构域、人 IgG2 Fc 结构域、人 IgG3 Fc 结构域或人 IgG4 Fc 结构域。

在某些实施方式中，所述 Fc 结构域包括如 SEQ ID NO: 24 所示的氨基酸序列。在某些实施方式中，所述 Fc 结构域由如 SEQ ID NO: 24 所示的氨基酸序列组成。在某些实施方式中，所述 Fc 结构域的氨基酸序列与如 SEQ ID NO: 24 所示的氨基酸序列具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%或至少 95%的序列同一性。

EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDP EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV SNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV EWESNGQPENNYKTTTPVLDSGDSFFLYSKLTVDKSRWQQGNVFCFSVMHEAL HNHYTQKSLSLSPGK (SEQ ID NO: 24)

在某些实施方式中，所述 Fc 结构域的编码核酸序列包括如 SEQ ID NO: 39 所示的核苷酸序列。在某些实施方式中，所述 Fc 结构域的编码核酸序列由如 SEQ ID NO: 39 所示的核苷酸序列组成。在某些实施方式中，所述 Fc 结构域的编码核酸序列与如 SEQ ID NO: 39 所示的核苷酸序列具有至少 70%、至少 75%、至少 80%、至

少 85%、至少 90%或至少 95%的序列同一性。

GAGCCCAAATCTTGTGACAAACTCACACATGCCCACCGTGCCCAGCACCTG
AACTCCTGGGGGACCGTCAGTCTTCCTCTTCCCCCAAACCCAAGGACACC
CTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCC
5 ACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCA
TAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGT
GGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTAC
AAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCT
CCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATC
10 CCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGC
TTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGA
ACAATAACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTC
TACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCT
CATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCT
15 CTCCTGTCTCCGGGTAAA (SEQ ID NO: 39)

在某些实施方式中,所述 Fc 结构域包括如 SEQ ID NO: 25 所示的氨基酸序列。
在某些实施方式中,所述 Fc 结构域由如 SEQ ID NO: 25 所示的氨基酸序列组成。
在某些实施方式送,所述 Fc 结构域的氨基酸序列与如 SEQ ID NO: 25 所示的氨基
酸序列具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%或至少 95%的
20 序列同一性。

TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV
DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI
EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP
ENNYKTTTPVLDSGDSFFLYSKLTVDKSRWQQGNVFSVSMHEALHNHYTQKS
25 LSLSPGK (SEQ ID NO: 25)

在某些实施方式中,所述 Fc 结构域的编码核酸序列包括如 SEQ ID NO: 40 所
示的核苷酸序列。在某些实施方式中,所述 Fc 结构域的编码核酸序列由如 SEQ ID
NO: 40 所示的核苷酸序列组成。在某些实施方式中,所述 Fc 结构域的编码核酸序
列与如 SEQ ID NO: 40 所示的核苷酸序列具有至少 70%、至少 75%、至少 80%、至
30 少 85%、至少 90%或至少 95%的序列同一性。

ACATGCCCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCT
CTTCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCA
CATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTG
GTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGA
5 GCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCACCGTCCTGCACCAG
GACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCC
CAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACC
ACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTC
AGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGT
10 GGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGC
TGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGC
AGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAA (SEQ ID NO:
40)。

15 在某些实施方式中,所述 Fc 结构域包括如 SEQ ID NO: 32 所示的氨基酸序列。
在某些实施方式中,所述 Fc 结构域由如 SEQ ID NO: 32 所示的氨基酸序列组成。
在某些实施方式送,所述 Fc 结构域的氨基酸序列与如 SEQ ID NO: 32 所示的氨基
酸序列具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%或至少 95%的
序列同一性。

20 ELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN
AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK
GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP
PVLDSGDGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK
(SEQ ID NO: 32)。

25 在某些实施方式中,所述 Fc 结构域包括如 SEQ ID NO: 77 所示的氨基酸序列。
在某些实施方式中,所述 Fc 结构域由如 SEQ ID NO: 77 所示的氨基酸序列组成。
在某些实施方式送,所述 Fc 结构域的氨基酸序列与如 SEQ ID NO: 77 所示的氨基
酸序列具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%或至少 95%的
序列同一性。

30 ESKYGPPCPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDPEVQ

FNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK
GLPSSIEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWE
SNGQPENNYKTTTPVLDSGDGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALHNHY
TQKSLSLSLGK (SEQ ID NO: 77)

- 5 在某些实施方式中，所述 Fc 结构域具有一个或多个氨基酸突变。在某些实施方式中，所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位有氨基酸突变。在某些实施方式中，所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位的氨基酸（例如，丙氨酸）突变为缬氨酸。在某些实施方式中，所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位的氨基酸（例如，丙氨酸）突变为甘氨酸。在某些实施方式中，所述 Fc
10 结构域在对应于 SEQ ID NO: 25 的第 7 位的氨基酸（例如，丙氨酸）突变为丝氨酸。在某些实施方式中，所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位的氨基酸（例如，丙氨酸）突变为亮氨酸。

- 在某些实施方式中，所述 Fc 结构域包含一个或多个延长所述融合蛋白的半衰期的突变。在某些实施方式中，所述 Fc 结构域与所述第一多肽的 C 末端连接。在
15 某些实施方式中，所述 Fc 结构域与所述第一多肽的 N 末端连接。

在某些实施方式中，所述第二多肽为白蛋白。在某些实施方式中，所述白蛋白的氨基酸序列如 SEQ ID NO: 60 所示。在某些实施方式中，所述白蛋白包含人血清白蛋白的一个或多个结构域。在某些实施方式中，所述白蛋白包含人血清白蛋白的 D3 结构域。

- 20 DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFEDHVKLVNEVTEFAKTCV
ADESAENCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKD
DNPNLPRPVRPEVDVMCTAFHDNEETFLKKYLYEIRRHYPYFYAPELLFFAKRY
KAAFTECCQAADKAAACLLPKLDEL RDEGKASSAKQRLKCASLQKFGERAFAKAW
AVARLSQRFPKAEFAEVSKLVTDLT KVHTECCHGDLLECADD RADLAKYICENQ
25 DSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAK
DVFLGMFLYEYARRHPDYSVVLRLAKTYETTLEKCCAAADPHECYAKVFDEF
KPLVEEPQNLIKQNCLEFEQLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK
VGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRP
CFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKATK
30 EQLKAVMDDFAAFVEKCKADDKETCFAEEGKKLVAASQAALGL (SEQ ID NO:

60)

在某些实施方式中，本申请提供的融合蛋白进一步包含标签。在某些实施方式中，所述标签选自下组：荧光标签、发光标签、纯化标签和生色标签。在某些实施方式中，所述标签选自下组：c-Myc 标签、HA 标签、VSV-G 标签、FLAG 标签、V5 标签和 HIS 标签。在某些实施方式中，所述标签为 HIS 标签。在某些实施方式中，所述标签是包含 6 个、7 个、8 个、9 个或 10 个组氨酸的 HIS 标签。在某些实施方式中，所述第二多肽位于所述第一多肽的 C 末端，所述标签位于所述第二多肽的 C 末端。

在某些实施方式中，本申请提供的融合蛋白包括如 SEQ ID NO: 26、SEQ ID NO: 27、SEQ ID NO: 28、SEQ ID NO: 29、SEQ ID NO: 30、SEQ ID NO: 31、SEQ ID NO: 81、SEQ ID NO: 82、SEQ ID NO: 83、SEQ ID NO: 84 或 SEQ ID NO: 85 所示的氨基酸序列。在某些实施方式中，本申请提供的融合蛋白由选自下组的氨基酸序列组成：SEQ ID NO: 26、SEQ ID NO: 27、SEQ ID NO: 28、SEQ ID NO: 29、SEQ ID NO: 30、SEQ ID NO: 31、SEQ ID NO: 81、SEQ ID NO: 82、SEQ ID NO: 83、SEQ ID NO: 84 或 SEQ ID NO: 85，或者与其具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 的序列同一性。在某些实施方式中，与 SEQ ID NO: 26、SEQ ID NO: 27、SEQ ID NO: 28、SEQ ID NO: 29、SEQ ID NO: 30、SEQ ID NO: 31、SEQ ID NO: 81、SEQ ID NO: 82、SEQ ID NO: 83、SEQ ID NO: 84 或 SEQ ID NO: 85 具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 的序列同一性的融合蛋白仍然保持 IgA 蛋白酶的功能或活性（例如，蛋白水解活性、特异性切割 IgA 的酶活性等）。

SEQ ID NO	氨基酸序列
26	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWL CGNPPKDG YVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ RSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD

	MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQ DAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYE PYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNIISKDNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLGGGGSEPKSCDKTHTCP PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK FNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKE YKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSL TCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTV DKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK
27	MYRMQLLSICIALSLALVTNSGTASKPDIKVG DYVKMGVYNNASIL WRCVSIDNNGPLMLADKIVDTLAYDAKTNDNSNSKSHRSYKRDD YGSNYWKDSNMRSWLNSTAAEGKVDWLCGNPPKDG YVSGVGAY NEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANSDLL YYTDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYVA YNNDGMAWPYWL RTPVTDCNHDMRYISSSGQVGRYAPWYSDLGV RPAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDANPD WNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVV VCGEGYTKSQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYALC TASESTFDNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPE FIEKIHDAHIKKKCDPNTIPSGSEYEPYYVHDYIAQFAMVVNTKSD FGGAYNNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNG YLLDDKELKSLNLSSVEDPEKIKRQLLGFRNTYTCRNAYGSKMLVS SYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEYT GAYSKPSDFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTNMNGKKI ELRTVIQNIISKDNARQLKFKMWIKHSDGSVATDSSGNPLQTVQTFDI

	PVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT VAFQVLGGGGSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTL MISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ PREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHN HYTQKSLSLSPGK
28	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWLCGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ RSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQ GKFIN DVKRLWQ DAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYE PYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGVGGGGSHHHHH HHHHHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDV SHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQ DWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDE LTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGS FFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
29	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWLCGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ

	RSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQ DAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYE PYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTETQG GGGSHHHHHHHHHHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTP EVT CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYR VVS VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT TPPVLDSDGSFFLYSKLTVDKSRWQQGNV FSCSVMHEALHNHYTQ KSLSLSPGK
30	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWLCGNPPKDG YVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ RSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQ DAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYE PYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK

	RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTETQR YTTVSIQYGGGGSHHHHHHHHTCPPCPAPELLGGPSVFLFPPKPK DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPRE EQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCFSVMHE ALHNHYTQKSLSLSPGK
31	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSRSYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWL CGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ RSLVSHPEYNKGIVDGDANSDLLYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQ GK FINDVKRLWQ DAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPEFIEKIHD AHIKKKCDPNTIPSGSEYE PYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTETQR YTTVSIQYKFEDGSEIPNTAGGTFTGGGGSHHHHHHHHHHTCPPCP APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC KVS NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCL VKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK

	SRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK
81	EYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQ SIQLTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYT KSQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYALCTASESTFD NGGSTFFDVIVDKYNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDA HIKKKCDPNTIPSGSEYEPYVVHDYIAQFAMVVNTKSDFGGAYNN REYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKEL KSLNLSSVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMR DTNYQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEYTGAYSKPS DFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQ NISDKNARQLKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDK ANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAFQVLD ENGNVLADDNTETQRYTTVSIQYGGGGGSTCPPCPAPELLGGPSVFL FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI EKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV EWESNGQPENNYKTTPVLDSGDSFFLYSKLTVDKSRWQQGNVFSC SVMHEALHNHYTQKSLSLSPGK
82	TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG KEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQV SLTCLVKGFYPSDIAVEWESNGQPENNYKTTPVLDSGDSFFLYSKL TVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGSEY FVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQ LTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQ QGKFINDVKRLWQDAMKYEPYRSYADRFNVYALCTASESTFDNGG STFFDVIVDKYNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIK KKCDPNTIPSGSEYEPYVVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKS LNLSSVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDT NYQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDF TDLETSSYYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNIS

	DKNARQLKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKAN FWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAFQVLLEN GNVLADDNTETQRYTTVSIQY
83	EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV VDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTV LHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDS DGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSP GKGGGGGGGSASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPL MLADKIVDTLAYDAKTNDNSNSKSHSRYSYKRDDYGSNYWKDSNM RSWLNSTAAEGKVDWLGCNPPKDGYSVSGVGAYNEKAGFLNAFSK SEIAAMKTVTQRSLVSHPEYNGKIVDGDANSDLLYYTDISEAVANY DSSYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMAWPY WLRTPVTD CNHDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYF VTTSGSGSQSSPYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQ LTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQ QGKFINDVKRLWQDAMKYEPYRSYADRFNVYALCTASESTFDNGG STFFDVIVDKYNSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIK KKCDPNTIPSGSEYEPYVYVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKS LNLSSVEDPEKIKWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDT NYQFCEVCRLQGFKRMSQLVKDVLVYVATPEVKEYTGAYSKPSDF TDLETSSYYNYTYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNIS DKNARQLKFKMWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKAN FWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAFQVLLEN GNVLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYGTCLDL TPAKTLYDYEFIKVDGLNKPIVSDGTVVYTYKKNKNEETHNLTLV AAKAATCTTAGNSAYYTCDGCDKWFADATGSVEITDKTSVKIPAP GHTAGTEWKSDDTNHWHECTVAGCGVIIESTKSAHTAGEWIVDTP ATATTAGTKHKECTVCHRLETQPIPISTGTELKIIAGDNQIYNKASG SDVTITCNGDFAKFTGIKVDGSVVDSSNYTAVSGSTVLTALKASYLGT LTDGSHTITFVYTDGEANANLTVRTAGSGHIIHDYGTWKSADNH WHECNCGDKKDEAAHSFKWVVDKEATATKKGSKHEECKICGYKR

	SAVEIPATGTSTAPTDTTKPNDTTKPGNTNGSEKSPQTGDNS
84	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWL CGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ RSLVSHPEYNKGIVDGDANS DLLYYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQ DAMKYEPYRSYADRFNVYALCTASESTFDNNGGSTFFDVIVDKYNP VISNNLHGSQWK NHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYE PYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTETQR YTTVSIQYGGGGGSESKYGPPCPSCPAPEFLGGPSVFLFPPKPKDTL MISRTPEVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPREEQF NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQ PREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE NNYKTTTPVLDS DGSFFLYSRLTVDKSRWQEGNVFSCSVMHEALH NHYTQKSLSLSLGK
85	ASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVDTL AYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTAAE GKVDWL CGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTVTQ RSLVSHPEYNKGIVDGDANS DLLYYTDISEAVANYDSSYFETTTEKV FLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCNHD MRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSSPYI GSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDGKY SNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQ

	DAMKYEPYRSYADRFNVYALCTASESTFDNNGGSTFFDVIVDKYNP VISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGSEYE PYYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSYRAS KTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKIKWR QLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQGFK RMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNYTYN RNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWI KHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHIKSDF NSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNTETQR YTTVSIQYGGGGGSDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQ QCPFEDHVKL VNEVTEFAKTCVADESAENCDSLHTLFGDKLCTVA TLRETYGEMADCCAKQEPERNECFLQHKDDNP NLPRLVRPEVDVM CTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTECCQ AADKAACLLPKLDEL RDEGKASSAKQRLK CASLQKFGERAFAKAWA VARLSQRFPKAEFAEVSKLVTDLT KVHTECCHGDLLECADDRADL AKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMPADLPSLAA DFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSV VLLLRLAKT YETTLKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCLEFEQL GEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEA KRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPCFS ALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHK PKATKEQLKAVMDDFAAFVEKCKADDKETCFAEEGKKLV AASQ AALGL
--	---

- 在某些实施方式中，本申请提供的融合蛋白包括如 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12 所示的氨基酸序列。在某些实施方式中，本申请提供的融合蛋白由选自下组的氨基酸序列组成：
- 5 SEQ ID NO: 2、SEQ ID NO: 4、SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12，或者与其具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 的序列同一性。

SEQ ID NO	氨基酸序列
2	HMASKPDIKVG DYVKMGVYNNASILWRCV SIDNNGPLMLADKIVD TLAYDAKTNDNSNSKSHSR SYKRDDYGSNYWKDSNMRSWLNSTA AEGKVDWL CGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTV TQRSLVSHPEYNKGIVDGDANS DLLYYTDISEAVANYDSSYFETTTE KVFLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCN HDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTS GSGSQSS PYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDG KYSNPTIPVYTIQKTRSDTENM VVVVCGEGYTKSQQ GK FINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFD VIVDKY NSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGS EYEPY VHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSY RASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKI KWRQLLGF RNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQ GFKRMSQLVKD VDLVATPEVKEYTGAYSKPSDFTDLETSSYYNY TYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFK MWIKHSDG SVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHI KSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQVLGGGGSEPKSCDKT HTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDP EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLN GKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSK LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKHHHHHHH AA
4	KLMYRMQLLSICIALSLALVTNSGTASKPDIKVG DYVKMGVYNNASI LWRCV SIDNNGPLMLADKIVDTLAYDAKTNDNSNSKSHSR SYKRD DYGSNYWKDSNMRSWLNSTA AEGKVDWL CGNPPK DGYVSGVGA YNEKAGFLNAFSKSEIAAMKTVTQRSLVSHPEYNKGIVDGDANS DL LYYTDISEAVANYDSSYFETTTEKVFLLDVKQANAVWKNLKGYYV AYNNDGMAWPYWL RTPVTDCNHDMRYISSSGQVGRYAPWYSDLG VRPAFYLDSEYFVTTS GSGSQSSPYIGSAPNKQEDDYTISEPAEDANP

	DWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQKTRSDTENMVV VVCGEGETKSQQGKFINDVKRLWQDAMKYEPYRSYADRFNVYAL CTASESTFDNGGSTFFDVIVDKYNSPVISNNLHGSQWKNHIFERCIGP EFIEKIHDAHIKKKCDPNTIPSGSEYEPYYYVHDYIAQFAMVVNTKS DFGGAYNNREYGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSN GYLLDDKELKSLNLSSVEDPEKIKRQLLGFRNTYTCRNAYGSKMLV SSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLYVATPEVKEY TGAYSKPSDFTDLETSSYYNYTYNRNDRLLSGNSKSRFNTNMNGKK IELRTVIQNISDKNARQLKFKMWIKHSDGSVATDSSGNPLQTVQTFD IPVWNDKANFWPLGALDHIKSDFNGLKSCSLIYQIPSDAQLKSGDT VAFQVLGGGGSEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTL MISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ PREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHN HYTQKSLSLSPGKHHHHHHHHHAA
6	HMASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVD TLAYDAKTNDNSNSKSHSRSYKRDDYGSNYWKDSNMRSWLNSTA AEGKVDWL CGNPPKDG YVSGVGAYNEKAGFLNAFSKSEIAAMKTV TQRSLVSHPEYNKGIVDGDANDLLYYTDISEAVANYDSSYFETTTE KVFLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCN HDMRYISSSGQVGRYAPWYS DLGVRPAFYLDSEYFVTTS GSGSQSS PYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDG KYSNPTIPVYTIQKTRSDTENMVVVVCGEGETKSQQGKFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKY NSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGS EYEPYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSY RASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKI KWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQ GFKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNY TYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFK MWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHI KSDFNGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG NVGGGGSH

	HHHHHHHHHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCV VVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLT VLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPP SRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GKAA
8	HMASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVD TLAYDAKTNDNSNSKSHSRSYKRDDYGSNYWKDSNMRSWLNSTA AEGKVDWL CGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTV TQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTE KVFLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCN HDMRYISSSGQVGRYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSS PYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDG KYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKY NSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGS EYEPYYYVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSY RASKTFAHEFGHGLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKI KWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQ GFKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNY TYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFK MWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHI KSDFN SGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG NVLADDNT ETQGGGGSHHHHHHHHHHTCPPCPAPELLGGPSVFLFPPKPKDTLM ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYN STYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN NYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHN HYTQKSLSLSPGKAA
10	HMASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVD TLAYDAKTNDNSNSKSHSRSYKRDDYGSNYWKDSNMRSWLNSTA AEGKVDWL CGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKTV

	TQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTE KVFLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCN HDMRYISSSGQVG RYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSS PYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDG KYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKY NSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGS EYEPYVVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSY RASKTFAHEFGHLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKI KWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQ GFKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNY TYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFK MWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHI KSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENGNVLADDNT ETQRYTTVSIQYGGGGSHHHHHHHHHHTCPPCPAPPELLGGPSVFLFP PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKT KPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK TISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW ESNGQPENNYKTTTPVLDS DGSFFLYSKLTVDKSRWQQGNVFSCSV MHEALHNHYTQKSLSLSPGKAA
12	HMASKPDIKVG DYVKMGVYNNASILWRCVSIDNNGPLMLADKIVD TLAYDAKTNDNSNSKSHSRSYKRDDYGSNYWKDSNMRSWLNSTA AEGKVDWL CGNPPKDG YVSGVGAYNEKAGFLNAFSKSEIAAMKTV TQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSSYFETTTE KVFLLDVKQANAVWKNLKGYYVAYNNDGMAWPYWL RTPVTDCN HDMRYISSSGQVG RYAPWYSDLGVRPAFYLDSEYFVTTSGSGSQSS PYIGSAPNKQEDDYTISEPAEDANPDWNVSTEQSIQLTLGPWYSNDG KYSNPTIPVYTIQKTRSDTENMVVVVCGEGYTKSQQGKFINDVKRL WQDAMKYEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKY NSPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPNTIPSGS EYEPYVVHDYIAQFAMVVNTKSDFGGAYNNREYGFHYFISPSDSY RASKTFAHEFGHLLGLGDEYSNGYLLDDKELKSLNLSSVEDPEKI KWRQLLGFRNTYTCRNAYGSKMLVSSYECIMRDTNYQFCEVCRLQ

	GFKRMSQLVKDVDLYVATPEVKEYTGAYSKPSDFTDLETSSYYNY TYNRNDRLLSGNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFK MWIKHSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALDHI KSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG NVLADDNT ETQRYTTVSIQYKFEDGSEIPNTAGGTFTGGGGSHHHHHHHHHTC PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV KFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGK EYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVS LTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGKAA
--	---

在某些实施方式中，本申请提供的融合蛋白在受试者体内的血循环中的半衰期为至少 1 天、至少 2 天、至少 3 天、至少 4 天、至少 5 天、至少 6 天、至少 7 天、至少 8 天、至少 9 天、至少 10 天、至少 11 天、至少 12 天、至少 13 天、至少 14 天。

核酸

在另一方面，本申请提供了一种分离的核酸，其包含编码本申请所述的 IgA 蛋白酶截短体的核苷酸序列或包含编码本申请所述的融合蛋白的核苷酸序列。

本申请所用的术语“核酸”或“核苷酸”是指单链或双链形式的脱氧核糖核酸 (DNA) 或核糖核酸 (RNA) 及其聚合物。除非另有说明，否则特定的核苷酸序列还隐含地涵盖其保守修饰的变体（例如简并的密码子取代）、等位基因、直向同源物、SNP 和互补序列以及明确指出的序列。具体而言，简并的密码子取代可通过产生这样的序列来实现：其中一个或多个选定的（或全部）密码子的第三位置被混合碱基和/或脱氧肌苷残基取代（参见 Batzer *et al.*, *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* 260:2605-2608 (1985)以及 Rossolini *et al.*, *Mol. Cell. Probes* 8:91-98 (1994)）。

使用传统的步骤，可以容易地对编码本申请所述的 IgA 蛋白酶截短体的 DNA 或者编码本申请所述的融合蛋白的 DNA 进行分离和测序（例如通过使用能够与编码所述 IgA 蛋白酶截短体或融合蛋白的基因特异性结合的寡核苷酸探针）。编码

DNA 也可以通过合成方法获得。

在某些实施方式中，本申请提供的核酸包括如 SEQ ID NO: 33、SEQ ID NO: 34、SEQ ID NO: 35、SEQ ID NO: 36、SEQ ID NO: 37、SEQ ID NO: 38 所示的核酸序列。在某些实施方式中，本申请提供的核酸由选自下组的核苷酸序列组成：SEQ ID NO: 33、SEQ ID NO: 34、SEQ ID NO: 35、SEQ ID NO: 36、SEQ ID NO: 37、SEQ ID NO: 38，或者与其具有至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性。

SEQ ID NO	核苷酸序列
33	GCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAAAATGGGTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGAGCATCGACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTTGACACGCTGGCGTATGATGCTAAAACCAACGACAATTCGAACAGCAAATCTCATAGTCGTTCCCTACAAACGCGATGACTACGGCAGCAACTATTGGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACCGCGGCCGAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAAAGACGGTTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTGGTTTTCTGAACGCGTTCTCAAAATCGGAAATTGCAGCTATGAAAACGGTGACCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAAGGCATTGTTGATGGTGACGCGAACTCGGATCTGCTGTATTACACCGACATCAGCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGAAACCACGACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGAACGCCGTCTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAACAATGATGGTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGACGGATTGTAATCATGACATGCGCTATATTAGTTCCCTCAGGCCAGGTTGGTCGTTACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGCGTTTTACC TGGACAGTGAATATTTTCGTGACGACCAGCGGCTCTGGTAGTCAGTCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAGAAGATGACTATACCATCTCAGAACCGGCGGAAGATGCCAACCCGGACTGGAATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGGCCCCGTGG

	TACTCGAATGATGGTAAATATAGCAACCCGACCATTCCGGTGTAT ACCATCCAGAAAACGCGCTCGGATACCGAAAACATGGTGGTTGT CGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGGCAAATTTA TCAATGATGTTAAACGTCTGTGGCAGGACGCTATGAAATATGAA CCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATGCACTGTGT ACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTACCTTTTTTC GATGTGATCGTTGACAAATACAACCTCTCCGGTTATCAGTAACAAT CTGCATGGCAGTCAGTGGAAAAATCACATTTTTGAACGCTGCATC GGTCCGGAATTCATTGAAAAAATCCATGATGCCACATTAAGAA AAAATGTGACCCGAACACCATCCCGTCGGGTAGCGAATACGAAC CGTATTACTATGTGCATGATTATATTGCACAGTTTGCTATGGTTG TCAATACCAAATCCGACTTCGGCGGTGCATATAACAATCGCGAA TACGGCTTTCCTACTATTTTCATCTCTCCGAGTGATTCCCTACCGTGCCT CTAAAACCTTTGCACATGAATTCGGCCACGGTCTGCTGGGCCTGG GTGATGAATACTCGAATGGTTATCTGCTGGATGACAAAGAACTG AAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGAAAAAATTAA ATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACACCTGCCGTAA CGCGTATGGTTCTAAAATGCTGGTTTCCTCATAACGAATGTATCAT GCGCGATACCAACTATCAATTTTGCGAAGTCTGTGCGCTGCAGGG CTTCAAACGTATGAGCCAACCTGGTTAAAGATGTCGACCTGTATGT GGCCACGCCGGAAGTTAAAGAATACACCGGTGCATATAGTAAAC CGTCCGATTTTACGGACCTGGAAACCTCGAGCTACTACAACCTACA CCTACAACCGTAACGATCGCCTGCTGAGTGGCAACTCAAAATCG CGTTTCAATACGAACATGAATGGCAAGAAAATTGAACTGCGCAC CGTTATTCAGAACATCAGCGATAAAAACGCCCGTCAACTGAAAT TCAAAATGTGGATCAAACATTCAGATGGCTCGGTGGCAACCGAC TCTAGTGGTAACCCGCTGCAGACCGTCCAAACGTTTGATATTCCG GTGTGGAACGACAAAGCCAATTTCTGGCCGCTGGGCGCACTGGA TCACATCAAATCCGACTTTAATTCAGGTCTGAAAAGCTGCTCTCT GATTTATCAGATCCCGTCTGATGCTCAACTGAAAAGTGGCGACA CCGTGGCGTTCCAGGTTCTG
34	GCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAAAATGG GTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGAGCATC

GACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTTGACAC GCTGGCGTATGATGCTAAAACCAACGACAATTGGAACAGCAAAT CTCATAGTCGTTCCCTACAAACGCGATGACTACGGCAGCAACTATT GGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACCGCGGCC GAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAAAGACGG TTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTGGTTTTCT GAACGCGTTCTCAAAATCGGAAATTGCAGCTATGAAAACGGTGA CCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAAGGCATTG TTGATGGTGACGCGAACTCGGATCTGCTGTATTACACCGACATCA GCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGAAACCACG ACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGAACGCCGT CTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAACAATGATG GTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGACGGATTGTA ATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGTTGGTCGTT ACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGCGTTTTACC TGGACAGTGAATATTTCTGACGACCAGCGGCTCTGGTAGTCAG TCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAGAAGATGA CTATACCATCTCAGAACCGGCGGAAGATGCCAACCCGGACTGGA ATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGGCCCCGTGG TACTCGAATGATGGTAAATATAGCAACCCGACCATTCCGGTGTAT ACCATCCAGAAAACGCGCTCGGATACCGAAAACATGGTGGTTGT CGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGGCAAATTTA TCAATGATGTTAAACGTCTGTGGCAGGACGCTATGAAATATGAA CCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATGCACTGTGT ACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTACCTTTTTC GATGTGATCGTTGACAAATACAACTCTCCGTTATCAGTAACAAT CTGCATGGCAGTCAGTGGAAAAATCACATTTTTGAACGCTGCATC GGTCCGGAATTCATTGAAAAAATCCATGATGCCACATTAAGAA AAAATGTGACCCGAACACCATCCCGTCGGGTAGCGAATACGAAC CGTATTACTATGTGCATGATTATATTGCACAGTTTGCTATGGTTG TCAATACCAAATCCGACTTCGGCGGTGCATATAACAATCGCGAA TACGGCTTTCCTACTTTTCATCTCTCCGAGTGATTCTACCGTGCCT CTAAAACCTTTGCACATGAATTCGGCCACGGTCTGCTGGGCCTGG
--

	GTGATGAATACTCGAATGGTTATCTGCTGGATGACAAAGAACTG AAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGAAAAAATTAA ATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACACCTGCCGTAA CGCGTATGGTTCTAAAATGCTGGTTTCCTCATAACGAATGTATCAT GCGCGATACCAACTATCAATTTTGCGAAGTCTGTGCGCTGCAGGG CTTCAAACGTATGAGCCAACCTGGTTAAAGATGTCGACCTGTATGT GGCCACGCCGGAAGTTAAAGAATACACCGGTGCATATAGTAAAC CGTCCGATTTTACGGACCTGGAAACCTCGAGCTACTACAACCTACA CCTACAACCGTAACGATCGCCTGCTGAGTGGCAACTCAAAATCG CGTTTCAATACGAACATGAATGGCAAGAAAATTGAACTGCGCAC CGTTATTCAGAACATCAGCGATAAAAACGCCCGTCAACTGAAAT TCAAAATGTGGATCAAACATTCAGATGGCTCGGTGGCAACCGAC TCTAGTGGTAACCCGCTGCAGACCGTCCAAACGTTTGATATTCCG GTGTGGAACGACAAAGCCAATTTCTGGCCGCTGGGCGCACTGGA TCACATCAAATCCGACTTTAATTCAGGTCTGAAAAGCTGCTCTCT GATTTATCAGATCCCGTCTGATGCTCAACTGAAAAGTGGCGACA CCGTGGCGTTCCAGGTTCTG
35	GCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAAAATGG GTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGAGCATC GACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTTGACAC GCTGGCGTATGATGCTAAAACCAACGACAATTCGAACAGCAAAT CTCATAGTCGTTCCCTACAAACGCGATGACTACGGCAGCAACTATT GGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACCGCGGCC GAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAAAGACGG TTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTGGTTTTCT GAACGCGTTCTCAAAATCGGAAATTGCAGCTATGAAAACGGTGA CCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAAGGCATTG TTGATGGTGACGCGAACTCGGATCTGCTGTATTACACCGACATCA GCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGAAACCACG ACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGAACGCCGT CTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAACAATGATG GTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGACGGATTGTA ATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGTTGGTCGTT

ACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGCGTTTTACC
TGGACAGTGAATATTTTCGTGACGACCAGCGGCTCTGGTAGTCAG
TCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAGAAGATGA
CTATACCATCTCAGAACCGGCGGAAGATGCCAACCCGGACTGGA
ATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGGCCCCGTGG
TACTCGAATGATGGTAAATATAGCAACCCGACCATTCCGGTGTAT
ACCATCCAGAAAACGCGCTCGGATACCGAAAACATGGTGGTTGT
CGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGGCAAATTTA
TCAATGATGTTAAACGTCTGTGGCAGGACGCTATGAAATATGAA
CCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATGCACTGTGT
ACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTACCTTTTTC
GATGTGATCGTTGACAAATACAACCTCTCCGGTTATCAGTAACAAT
CTGCATGGCAGTCAGTGGAAAAATCACATTTTGAACGCTGCATC
GGTCCGGAATTCATTGAAAAAATCCATGATGCCACATTAAGAA
AAAATGTGACCCGAACACCATCCCGTCGGGTAGCGAATACGAAC
CGTATTACTATGTGCATGATTATATTGCACAGTTTGCTATGGTTG
TCAATACCAAATCCGACTTCGGCGGTGCATATAACAATCGCGAA
TACGGCTTTCACTATTTTCATCTCTCCGAGTGATTCTACCGTGCCT
CTAAAACCTTTGCACATGAATTCGGCCACGGTCTGCTGGGCCTGG
GTGATGAATACTCGAATGGTTATCTGCTGGATGACAAAGAAGT
AAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGAAAAAATTAA
ATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACACCTGCCGTAA
CGCGTATGGTTCTAAAATGCTGGTTTCCTCATACGAATGTATCAT
GCGCGATACCAACTATCAATTTTGCGAAGTCTGTGCGCTGCAGGG
CTTCAAACGTATGAGCCAACTGGTTAAAGATGTCGACCTGTATGT
GGCCACGCCGGAAGTTAAAGAATACACCGGTGCATATAGTAAAC
CGTCCGATTTTACGGACCTGGAAACCTCGAGCTACTACAACCTACA
CCTACAACCGTAACGATCGCCTGCTGAGTGGCAACTCAAAATCG
CGTTTCAATACGAACATGAATGGCAAGAAAATTGAACTGCGCAC
CGTTATTCAGAACATCAGCGATAAAAACGCCCGTCAACTGAAAT
TCAAAATGTGGATCAAACATTCAGATGGCTCGGTGGCAACCGAC
TCTAGTGGTAACCCGCTGCAGACCGTCCAAACGTTTGATATTCCG
GTGTGGAACGACAAAGCCAATTTCTGGCCGCTGGGCGCACTGGA

	TCACATCAAATCCGACTTTAATTCAGGTCTGAAAAGCTGCTCTCT GATTTATCAGATCCCGTCTGATGCTCAACTGAAAAGTGCGGACA CCGTGGCGTTCCAGGTTCTGGATGAAAACGGTAATGTG
36	GCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAAAATGG GTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGAGCATC GACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTTGACAC GCTGGCGTATGATGCTAAAACCAACGACAATTCGAACAGCAAAT CTCATAGTCGTTCTTACAAACGCGATGACTACGGCAGCAACTATT GGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACCGCGGCC GAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAAAGACGG TTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTGGTTTTCT GAACGCGTTCTCAAAATCGGAAATTGCAGCTATGAAAACGGTGA CCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAAGGCATTG TTGATGGTGACGCGAACTCGGATCTGCTGTATTACACCGACATCA GCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGAAACCACG ACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGAACGCCGT CTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAACAATGATG GTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGACGGATTGTA ATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGTTGGTCGTT ACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGCGTTTTACC TGGACAGTGAATATTTCTGTGACGACCAGCGGCTCTGGTAGTCAG TCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAGAAGATGA CTATACCATCTCAGAACCGGCGGAAGATGCCAACCCGGACTGGA ATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGGCCCCGTGG TACTCGAATGATGGTAAATATAGCAACCCGACCATTCCGGTGTAT ACCATCCAGAAAACGCGCTCGGATACCGAAAACATGGTGGTTGT CGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGGCAAATTTA TCAATGATGTTAAACGTCTGTGGCAGGACGCTATGAAATATGAA CCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATGCACTGTGT ACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTACCTTTTC GATGTGATCGTTGACAAATACAACTCTCCGGTTATCAGTAACAAT CTGCATGGCAGTCAGTGGAAAAATCACATTTTGAACGCTGCATC GGTCCGGAATTCATTGAAAAAATCCATGATGCCACATTAAGAA

	AAAATGTGACCCGAACACCATCCCGTCGGGTAGCGAATACGAAC CGTATTACTATGTGCATGATTATATTGCACAGTTTGCTATGGTTG TCAATACCAAATCCGACTTCGGCGGTGCATATAACAATCGCGAA TACGGCTTTCACTATTTTCATCTCTCCGAGTGATTCCCTACCGTGCCT CTAACCTTTGCACATGAATTCGGCCACGGTCTGCTGGGCCTGG GTGATGAATACTCGAATGGTTATCTGCTGGATGACAAAGAACTG AAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGAAAAAATTAA ATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACACCTGCCGTAA CGCGTATGGTTCTAAAATGCTGGTTTCCTCATACGAATGTATCAT GCGCGATACCAACTATCAATTTTGCGAAGTCTGTGCGCTGCAGGG CTTCAAACGTATGAGCCAACTGGTTAAAGATGTCGACCTGTATGT GGCCACGCCGGAAGTTAAAGAATACACCGGTGCATATAGTAAAC CGTCCGATTTTACGGACCTGGAAACCTCGAGCTACTACAACTACA CCTACAACCGTAACGATCGCCTGCTGAGTGGCAACTCAAATCG CGTTTCAATACGAACATGAATGGCAAGAAAATTGAACTGCGCAC CGTTATTCAGAACATCAGCGATAAAAACGCCCGTCAACTGAAAT TCAAATGTGGATCAAACATTCAGATGGCTCGGTGGCAACCGAC TCTAGTGGTAACCCGCTGCAGACCGTCCAAACGTTTGATATTCCG GTGTGGAACGACAAAGCCAATTTCTGGCCGCTGGGCGCACTGGA TCACATCAAATCCGACTTTAATTCAGGTCTGAAAAGCTGCTCTCT GATTTATCAGATCCCGTCTGATGCTCAACTGAAAAGTGGCGACA CCGTGGCGTTCCAGGTTCTGGATGAAAACGGTAATGTGCTGGCG GATGACAACACGGAAACCCAG
37	GCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAAAATGG GTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGAGCATC GACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTTGACAC GCTGGCGTATGATGCTAAAACCAACGACAATTCGAACAGCAAAT CTCATAGTCGTTTCTACAAACGCGATGACTACGGCAGCAACTATT GGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACCGCGGCC GAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAAAGACGG TTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTGGTTTTCT GAACGCGTTCTCAAATCGGAAATTGCAGCTATGAAAACGGTGA CCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAAGGCATTG

TTGATGGTGACGCGAACTCGGATCTGCTGTATTACACCGACATCA GCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGAACACG ACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGAACGCCGT CTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAACAATGATG GTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGACGGATTGTA ATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGTTGGTCGTT ACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGCGTTTTACC TGGACAGTGAATATTTCTGTGACGACCAGCGGCTCTGGTAGTCAG TCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAGAAGATGA CTATACCATCTCAGAACCGGCGGAAGATGCCAACCCGGACTGGA ATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGGCCCCGTGG TACTCGAATGATGGTAAATATAGCAACCCGACCATTCCGGTGTAT ACCATCCAGAAAACGCGCTCGGATACCGAAAACATGGTGGTTGT CGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGGCAAATTTA TCAATGATGTTAAACGTCTGTGGCAGGACGCTATGAAATATGAA CCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATGCACTGTGT ACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTACCTTTTTTC GATGTGATCGTTGACAAATACAACTCTCCGGTTATCAGTAACAAT CTGCATGGCAGTCAGTGGAAAAATCACATTTTTGAACGCTGCATC GGTCCGGAATTCATTGAAAAAATCCATGATGCCACATTAAGAA AAAATGTGACCCGAACACCATCCCGTCGGGTAGCGAATACGAAC CGTATTACTATGTGCATGATTATATTGCACAGTTTGCTATGGTTG TCAATACCAAATCCGACTTCGGCGGTGCATATAACAATCGCGAA TACGGCTTTCACTATTTTCATCTCTCCGAGTGATTCTACCGTGCCT CTAAAACCTTTGCACATGAATTCGGCCACGGTCTGCTGGGCCTGG GTGATGAATACTCGAATGGTTATCTGCTGGATGACAAAGAACTG AAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGAAAAAATTAA ATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACACCTGCCGTAA CGCGTATGGTTCTAAAATGCTGGTTTCCTCATACGAATGTATCAT GCGCGATACCAACTATCAATTTTGCGAAGTCTGTGCGCTGCAGGG CTTCAAACGTATGAGCCAACTGGTTAAAGATGTCGACCTGTATGT GGCCACGCCGGAAGTTAAAGAATACACCGGTGCATATAGTAAAC CGTCCGATTTTACGGACCTGGAAACCTCGAGCTACTACAACCTACA CCTACAACCGTAACGATCGCCTGCTGAGTGGCAACTCAAAATCG
--

	CGTTTCAATACGAACATGAATGGCAAGAAAATTGAACTGCGCAC CGTTATTCAGAACATCAGCGATAAAAACGCCCGTCAACTGAAAT TCAAAATGTGGATCAAACATTCAGATGGCTCGGTGGCAACCGAC TCTAGTGGTAACCCGCTGCAGACCGTCCAAACGTTTGATATTCCG GTGTGGAACGACAAAGCCAATTTCTGGCCGCTGGGCGCACTGGA TCACATCAAATCCGACTTTAATTCAGGTCTGAAAAGCTGCTCTCT GATTTATCAGATCCCGTCTGATGCTCAACTGAAAAGTGGCGACA CCGTGGCGTTCCAGGTTCTGGATGAAAACGGTAATGTGCTGGCG GATGACAACACGGAAACCCAGCGCTACACGACCGTTTCTATCCA ATAC
38	GCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAAAATGG GTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGAGCATC GACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTTGACAC GCTGGCGTATGATGCTAAAACCAACGACAATTCGAACAGCAAAT CTCATAGTCGTTCCCTACAAACGCGATGACTACGGCAGCAACTATT GGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACCGCGGCC GAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAAAGACGG TTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTGGTTTTCT GAACGCGTTCTCAAATCGGAAATTGCAGCTATGAAAACGGTGA CCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAAGGCATTG TTGATGGTGACGCGAACTCGGATCTGCTGTATTACACCGACATCA GCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGAAACCACG ACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGAACGCCGT CTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAACAATGATG GTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGACGGATTGTA ATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGTTGGTCGTT ACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGCGTTTTACC TGGACAGTGAATATTTTCGTGACGACCAGCGGCTCTGGTAGTCAG TCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAGAAGATGA CTATACCATCTCAGAACCGGCGGAAGATGCCAACCCGGACTGGA ATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGGCCCCGTGG TACTCGAATGATGGTAAATATAGCAACCCGACCATTCCGGTGTAT ACCATCCAGAAAACGCGCTCGGATACCGAAAACATGGTGGTTGT

CGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGGCAAATTTA
TCAATGATGTTAAACGTCTGTGGCAGGACGCTATGAAATATGAA
CCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATGCACTGTGT
ACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTACCTTTTTC
GATGTGATCGTTGACAAATACAACCTCTCCGGTTATCAGTAACAAT
CTGCATGGCAGTCAGTGGAAAAATCACATTTTTGAACGCTGCATC
GGTCCGGAATTCATTGAAAAAATCCATGATGCCCACATTAAGAA
AAAATGTGACCCGAACACCATCCCGTCGGGTAGCGAATACGAAC
CGTATTACTATGTGCATGATTATATTGCACAGTTTGCTATGGTTG
TCAATACCAAATCCGACTTCGGCGGTGCATATAACAATCGCGAA
TACGGCTTTCCTACTATTTTCATCTCTCCGAGTGATTCCCTACCGTGCCT
CTAAAACCTTTGCACATGAATTCGGCCACGGTCTGCTGGGCCTGG
GTGATGAATACTCGAATGGTTATCTGCTGGATGACAAAGAACTG
AAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGAAAAAATTA
ATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACACCTGCCGTAA
CGCGTATGGTTCTAAAATGCTGGTTTCCTCATAACGAATGTATCAT
GCGCGATACCAACTATCAATTTTGCGAAGTCTGTGCGCTGCAGGG
CTTCAAACGTATGAGCCAACCTGGTTAAAGATGTCGACCTGTATGT
GGCCACGCCGGAAGTTAAAGAATACACCGGTGCATATAGTAAAC
CGTCCGATTTTACGGACCTGGAAACCTCGAGCTACTACAACCTACA
CCTACAACCGTAACGATCGCCTGCTGAGTGGCAACTCAAAATCG
CGTTTCAATACGAACATGAATGGCAAGAAAATTGAACTGCGCAC
CGTTATTCAGAACATCAGCGATAAAAACGCCCCGTCAACTGAAAT
TCAAAATGTGGATCAAACATTCAGATGGCTCGGTGGCAACCGAC
TCTAGTGGTAACCCGCTGCAGACCGTCCAAACGTTTGATATTCCG
GTGTGGAACGACAAAGCCAATTTCTGGCCGCTGGGCGCACTGGA
TCACATCAAATCCGACTTTAATTCAGGTCTGAAAAGCTGCTCTCT
GATTTATCAGATCCCGTCTGATGCTCAACTGAAAAGTGGCGACA
CCGTGGCGTTCCAGGTTCTGGATGAAAACGGTAATGTGCTGGCG
GATGACAACACGGAAACCCAGCGCTACACGACCGTTTCTATCCA
ATACAAATTCGAAGATGGCAGTGAAATCCCGAATACGGCGGGCG
GTACCTTCACC

在某些实施方式中，本申请提供的核酸包括如 SEQ ID NO: 3、SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13 所示的核酸序列。在某些实施方式中，本申请提供的核酸由选自下组的核苷酸序列组成：SEQ ID NO: 3、SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13，
5 或者与其具有至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性。

SEQ ID NO	核苷酸序列
3	CATATGGCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAA AATGGGTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGA GCATCGACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTT GACACGCTGGCGTATGATGCTAAAACCAACGACAATTTCGAACAG CAAATCTCATAGTCGTTCTTCTACAAACGCGATGACTACGGCAGCA ACTATTGGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACC GCGGCCGAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAA AGACGGTTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTG GTTTTCTGAACGCGTTCTCAAAATCGGAAATTGCAGCTATGAAAA CGGTGACCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAA GGCATTGTTGATGGTGACGCGAACTCGGATCTGCTGTATTACACC GACATCAGCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGA AACCACGACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGA ACGCCGTCTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAAC AATGATGGTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGAC GGATTGTAATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGT TGGTCGTTACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGC GTTTTACCTGGACAGTGAATATTTCTGACGACCAGCGGCTCTGG TAGTCAGTCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAG AAGATGACTATAACCATCTCAGAACCGGCGGAAGATGCCAACCCG GACTGGAATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGG CCCGTGGTACTCGAATGATGGTAAATATAGCAACCCGACCATTC CGGTGTATACCATCCAGAAAACGCGCTCGGATACCGAAAACATG

GTGGTTGTCGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGG CAAATTTATCAATGATGTAAACGTCTGTGGCAGGACGCTATGA AATATGAACCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATG CACTGTGTACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTA CCTTTTTTCGATGTGATCGTTGACAAATACAACCTCTCCGGTTATCA GTAACAATCTGCATGGCAGTCAGTGGAAAAATCACATTTTTGAA CGCTGCATCGGTCCGGAATTCATTGAAAAAATCCATGATGCCCA CATTAAGAAAAAATGTGACCCGAACACCATCCCGTCGGGTAGCG AATACGAACCGTATTACTATGTGCATGATTATATTGCACAGTTTG CTATGGTTGTCAATACCAAATCCGACTTCGGCGGTGCATATAACA ATCGCGAATACGGCTTTCCTACTATTTTCATCTCTCCGAGTGATTCT ACCGTGCCTCTAAAACCTTTGCACATGAATTCGGCCACGGTCTGC TGGGCCTGGGTGATGAATACTCGAATGGTTATCTGCTGGATGAC AAAGAACTGAAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGA AAAAATTAAATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACA CCTGCCGTAACGCGTATGGTTCTAAAATGCTGGTTTTCTCATACG AATGTATCATGCGCGATACCAACTATCAATTTTGCGAAGTCTGTC GCCTGCAGGGCTTCAAACGTATGAGCCAACTGGTTAAAGATGTC GACCTGTATGTGGCCACGCCGGAAGTTAAAGAATACACCGGTGC ATATAGTAAACCGTCCGATTTTACGGACCTGGAAACCTCGAGCT ACTACAACCTACACCTACAACCGTAACGATCGCCTGCTGAGTGGC AACTCAAAATCGCGTTTCAATACGAACATGAATGGCAAGAAAAT TGAACCTGCGCACCGTTATTCAGAACATCAGCGATAAAAACGCCC GTCAACTGAAATTCAAAATGTGGATCAAACATTCAGATGGCTCG GTGGCAACCGACTCTAGTGGTAACCCGCTGCAGACCGTCCAAAC GTTTGATATTCCGGTGTGGAACGACAAAGCCAATTTCTGGCCGCT GGGCGCACTGGATCACATCAAATCCGACTTTAATTCAGGTCTGA AAAGCTGCTCTCTGATTTATCAGATCCCGTCTGATGCTCAACTGA AAAGTGGCGACACCGTGGCGTTCCAGGTTCTGGGCGGCGGTGGA TCCGAGCCCAAATCTTGTGACAAAACCTCACACATGCCACCGTG CCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCC CCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGG TCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTC

	AAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAA GACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTG GTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAA GGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCA TCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCA CAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAA CCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGA CATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAAC TACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTC CTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGG GAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCA CTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAACACCATC ATCATCATCATTAAGCGGCCGC
5	AAGCTTATGTATAGAATGCAGCTGCTGTCCTGTATTGCTCTGAGC CTGGCACTGGTTACAAACAGCGGTACCGCGAGCAAACCGGACAT CAAAGTGGGCGACTACGTGAAAATGGGTGTGTATAATAACGCAA GCATCCTGTGGCGCTGTGTGAGCATCGACAACAATGGCCCGCTG ATGCTGGCCGATAAAATTGTTGACACGCTGGCGTATGATGCTAA AACCAACGACAATTCGAACAGCAAATCTCATAGTCGTTCTTACA AACGCGATGACTACGGCAGCAACTATTGGAAAGATAGTAATATG CGCTCCTGGCTGAACTCAACCGCGGCCGAGGGTAAAGTGGATTG GCTGTGCGGCAATCCGCCGAAAGACGGTTACGTCAGCGGCGTGG GTGCATATAATGAAAAAGCTGGTTTTCTGAACGCGTTCTCAAAAT CGGAAATTGCAGCTATGAAAACGGTGACCCAGCGTAGCCTGGTT TCTCATCCGGAATATAATAAAGGCATTGTTGATGGTGACGCGAA CTCGGATCTGCTGTATTACACCGACATCAGCGAAGCAGTGGCTA ACTACGATAGCTCTTATTTTGAAACCACGACCGAAAAAGTTTTCC TGCTGGATGTCAAACAGGCGAACGCCGTCTGGAAAAATCTGAAA GGCTATTACGTGGCTTACAACAATGATGGTATGGCATGGCCGTAT TGGCTGCGTACCCCGGTGACGGATTGTAATCATGACATGCGCTAT ATTAGTTCCTCAGGCCAGGTTGGTCGTTACGCTCCGTGGTATTCT GATCTGGGCGTCCGTCCGGCGTTTTACCTGGACAGTGAATATTTT GTGACGACCAGCGGCTCTGGTAGTCAGTCGAGCCCGTACATTGG

TTCCGCGCCGAACAAACAAGAAGATGACTATACCATCTCAGAAC
CGGCGGAAGATGCCAACCCGGACTGGAATGTTTCGACGGAACAG
AGCATTCAACTGACCCTGGGCCCCGTGGTACTCGAATGATGGTAA
ATATAGCAACCCGACCATTCCGGTGTATACCATCCAGAAAACGC
GCTCGGATACCGAAAACATGGTGGTTGTCGTGTGCGGCGAAGGT
TATACCAAATCACAGCAAGGCAAATTTATCAATGATGTTAAACG
TCTGTGGCAGGACGCTATGAAATATGAACCGTACCGTAGCTATG
CGGATCGCTTTAATGTGTATGCACTGTGTACGGCTTCCGAATCAA
CCTTCGATAACGGCGGTTCTACCTTTTTTCGATGTGATCGTTGACA
AATACAACCTCCGGTTATCAGTAACAATCTGCATGGCAGTCAGT
GGAAAAATCACATTTTTGAACGCTGCATCGGTCCGGAATTCATTG
AAAAAATCCATGATGCCACATTAAGAAAAAATGTGACCCGAAC
ACCATCCCGTCGGGTAGCGAATACGAACCGTATTACTATGTGCAT
GATTATATTGCACAGTTTGCTATGGTTGTCAATACCAAATCCGAC
TTCGGCGGTGCATATAACAATCGCGAATACGGCTTTCACTATTTT
ATCTCTCCGAGTGATTCTACCGTGCCTCTAAAACCTTTGCACAT
GAATTCGGCCACGGTCTGCTGGGCCTGGGTGATGAATACTCGAA
TGGTTATCTGCTGGATGACAAAGAACTGAAAAGCCTGAACCTGT
CTAGTGTGGAAGATCCGGAAAAAATTAAATGGCGTCAGCTGCTG
GGCTTTCGCAATACGTACACCTGCCGTAACGCGTATGGTTCTAAA
ATGCTGGTTTCCTCATACGAATGTATCATGCGCGATACCAACTAT
CAATTTTGCGAAGTCTGTGCGCTGCAGGGCTTCAAACGTATGAGC
CAACTGGTTAAAGATGTCGACCTGTATGTGGCCACGCCGGAAGT
TAAAGAATACACCGGTGCATATAGTAAACCGTCCGATTTTACGG
ACCTGGAAACCTCGAGCTACTACAACTACACCTACAACCGTAAC
GATCGCCTGCTGAGTGGCAACTCAAAATCGCGTTTCAATACGAA
CATGAATGGCAAGAAAATTGAACTGCGCACCGTTATTTCAGAACAA
TCAGCGATAAAAACGCCCGTCAACTGAAATTCAAAATGTGGATC
AAACATTCAGATGGCTCGGTGGCAACCGACTCTAGTGGTAACCC
GCTGCAGACCGTCCAAACGTTTGATATTCCGGTGTGGAACGACA
AAGCCAATTTCTGGCCGCTGGGCGCACTGGATCACATCAAATCC
GACTTTAATTCAGGTCTGAAAAGCTGCTCTCTGATTTATCAGATC
CCGTCTGATGCTCAACTGAAAAGTGGCGACACCGTGGCGTTCCA

	GGTTCTGGGCGGCGGTGGATCCGAACCTAAGAGTTGCGATAAAA CCCACACTTGCCCTCCCTGTCCGGCCCCCGAACTGCTCGGCGGAC CCTCAGTCTTCCTGTTCCCCCAAAGCCAAAGGACACATTGATGA TCAGCAGGACTCCTGAAGTGACATGCGTGGTTCGTAGACGTGTCA CACGAGGACCCGGAGGTGAAGTTCAACTGGTACGTGGACGGAGT GGAGGTGCATAATGCCAAAACAAAGCCCAGAGAAGAGCAGTAT AACAGTACCTACAGAGTGGTGTGAGTGCTGACCGTGCTTCATCA GGATTGGCTGAACGGGAAGGAGTACAAGTGTAAGGTGAGTAATA AGGCTCTGCCTGCCCCAATTGAGAAGACAATCTCTAAAGCCAAG GGGCAGCCCCGGGAACCCCAAGTGTATACACTCCCACCGTCCCG CGATGAACTGACAAAAAACCAGGTATCACTCACTTGTCTGGTAA AGGGCTTCTATCCATCTGACATTGCCGTGGAGTGGGAATCAAAC GGCCAACCCGAGAATAATTATAAGACAACCCCGCCCGTGCTGGA TTCCGACGGATCTTTTTTTCCTGTATAGCAAATTGACTGTGACAA AAGTCGGTGGCAGCAGGGCAATGTGTTTTCTTGCAGCGTCATGC ATGAGGCGCTGCACAACCACTATACTCAGAAGTCATTGAGCTTG AGCCCTGGTAAGCACCATCATCACCATCACCATCATTAGGCGGC CGC
7	CATATGGCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAA AATGGGTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGA GCATCGACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTT GACACGCTGGCGTATGATGCTAAAACCAACGACAATTGGAACAG CAAATCTCATAGTCGTTCTTACAAACGCGATGACTACGGCAGCA ACTATTGGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACC GCGGCCGAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAA AGACGGTTACGTGAGCGGCGTGGGTGCATATAATGAAAAAGCTG GTTTTCTGAACGCGTTCTCAAATCGGAAATTGCAGCTATGAAAA CGGTGACCCAGCGTAGCCTGGTTTTCTCATCCGGAATATAATAAA GGCATTGTTGATGGTGACGCGAACTCGGATCTGCTGTATTACACC GACATCAGCGAAGCAGTGGCTAACTACGATAGCTCTTATTTGA AACCACGACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGA ACGCCGTCTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAAC AATGATGGTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGAC

GGATTGTAATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGT
TGGTCGTTACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGC
GTTTTACCTGGACAGTGAATATTTTCGTGACGACCAGCGGCTCTGG
TAGTCAGTCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAG
AAGATGACTATAACCATCTCAGAACCGGCGGAAGATGCCAACCCG
GACTGGAATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGG
CCCGTGGTACTCGAATGATGGTAAATATAGCAACCCGACCATTCT
CGGTGTATACCATCCAGAAAACGCGCTCGGATACCGAAAACATG
GTGGTTGTCGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGG
CAAATTTATCAATGATGTTAAACGTCTGTGGCAGGACGCTATGA
AATATGAACCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATG
CACTGTGTACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTA
CCTTTTTTCGATGTGATCGTTGACAAATACAACCTCTCCGGTTATCA
GTAACAATCTGCATGGCAGTCAGTGGAAAAATCACATTTTTGAA
CGCTGCATCGGTCCGGAATTCATTGAAAAAATCCATGATGCCCA
CATTAAAGAAAAAATGTGACCCGAACACCATCCCGTCGGGTAGCG
AATACGAACCGTATTACTATGTGCATGATTATATTGCACAGTTTG
CTATGGTTGTCAATACCAAATCCGACTTCGGCGGTGCATATAACA
ATCGCGAATACGGCTTTCCTACTATTTTCATCTCTCCGAGTGATTCT
ACCGTGCCTCTAAAACCTTTGCACATGAATTCGGCCACGGTCTGC
TGGGCCTGGGTGATGAATACTCGAATGGTTATCTGCTGGATGAC
AAAGAACTGAAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGA
AAAAATTAAATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACA
CCTGCCGTAACGCGTATGGTTCTAAAATGCTGGTTTTCTCATACG
AATGTATCATGCGCGATACCAACTATCAATTTTGCGAAGTCTGTC
GCCTGCAGGGCTTCAAACGTATGAGCCAACTGGTTAAAGATGTC
GACCTGTATGTGGCCACGCCGGAAGTTAAAGAATACACCGGTGC
ATATAGTAAACCGTCCGATTTTACGGACCTGGAAACCTCGAGCT
ACTACAACTACACCTACAACCGTAACGATCGCCTGCTGAGTGGC
AACTCAAAATCGCGTTTCAATACGAACATGAATGGCAAGAAAAT
TGAACCTGCGCACCGTTATTCAGAACATCAGCGATAAAAACGCCC
GTCAACTGAAATTCAAAATGTGGATCAAACATTCAGATGGCTCG
GTGGCAACCGACTCTAGTGGTAACCCGCTGCAGACCGTCCAAAC

	GTTTGATATTCCGGTGTGGAACGACAAAGCCAATTTCTGGCCGCT GGGCGCACTGGATCACATCAAATCCGACTTTAATTCAGGTCTGA AAAGCTGCTCTCTGATTTATCAGATCCCGTCTGATGCTCAACTGA AAAGTGGCGACACCGTGGCGTTCCAGGTTCTGGATGAAAACGGT AATGTGGGCGGCGGTGGATCCCACCATCATCACCACCATCATCA TCACCACACATGCCCACCGTGCCCAGCACCTGAACTCCTGGGGG GACCGTCAGTCTTCCTCTTCCCCCAAAACCCAAGGACACCCTCA TGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTG AGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGG CGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAG TACAACAGCACGTACCGGGTGGTCAGCGTCCTCACCGTCCTGCA CCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCA ACAAAGCCCTCCCAGCCCCATCGAGAAAACCATCTCCAAAGCC AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATC CCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGG TCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGC AATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCT GGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGA CAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGA TGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCC CTGTCTCCGGGTAAATAAGCGGCCGC
9	CATATGGCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAA AATGGGTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGA GCATCGACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTT GACACGCTGGCGTATGATGCTAAAACCAACGACAATTCGAACAG CAAATCTCATAGTCGTTTCCTACAAACGCGATGACTACGGCAGCA ACTATTGGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACC GCGGCCGAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAA AGACGGTTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTG GTTTTCTGAACGCGTTCTCAAAATCGGAAATTGCAGCTATGAAAA CGGTGACCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAA GGCATTGTTGATGGTGACGCGAACTCGGATCTGCTGTATTACACC GACATCAGCGAAGCAGTGGCTAACTACGATAGCTCTTATTTGA

AACCACGACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGA ACGCCGTCTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAAC AATGATGGTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGAC GGATTGTAATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGT TGGTCGTTACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGC GTTTTACCTGGACAGTGAATATTTCTGTGACGACCAGCGGCTCTGG TAGTCAGTCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAG AAGATGACTATAACCATCTCAGAACCGGCGGAAGATGCCAACCCG GACTGGAATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGG CCCGTGGTACTCGAATGATGGTAAATATAGCAACCCGACCATTCTC CGGTGTATACCATCCAGAAAACGCGCTCGGATACCGAAAACATG GTGGTTGTCGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGG CAAATTTATCAATGATGTTAAACGTCTGTGGCAGGACGCTATGA AATATGAACCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATG CACTGTGTACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTA CCTTTTTTCGATGTGATCGTTGACAAATACAACCTCTCCGGTTATCA GTAACAATCTGCATGGCAGTCAGTGGAAAAATCACATTTTTTGAA CGCTGCATCGGTCCGGAATTCATTGAAAAAATCCATGATGCCCA CATTAAAGAAAAAATGTGACCCGAACACCATCCCGTCGGGTAGCG AATACGAACCGTATTACTATGTGCATGATTATATTGCACAGTTTG CTATGGTTGTCAATACCAAATCCGACTTCGGCGGTGCATATAACA ATCGCGAATACGGCTTTCCTACTATTTTCATCTCTCCGAGTGATTCTT ACCGTGCCTCTAAAACCTTTGCACATGAATTCGGCCACGGTCTGC TGGGCCTGGGTGATGAATACTCGAATGGTTATCTGCTGGATGAC AAAGAACTGAAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGA AAAAATTAAATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACA CCTGCCGTAACGCGTATGGTTCTAAAATGCTGGTTTTCTCATACG AATGTATCATGCGCGATACCAACTATCAATTTTGCGAAGTCTGTC GCCTGCAGGGCTTCAAACGTATGAGCCAACTGGTTAAAGATGTC GACCTGTATGTGGCCACGCCGGAAGTTAAAGAATACACCGGTGC ATATAGTAAACCGTCCGATTTTACGGACCTGGAAACCTCGAGCT ACTACAACCTACACCTACAACCGTAACGATCGCCTGCTGAGTGGC AACTCAAAATCGCGTTTCAATACGAACATGAATGGCAAGAAAAT

	TGAAGTGGCGACCGTTATTCAGAACATCAGCGATAAAAACGCCC GTCAACTGAAATTCAAAATGTGGATCAAACATTCAGATGGCTCG GTGGCAACCGACTCTAGTGGTAACCCGCTGCAGACCGTCCAAAC GTTTGATATTCCGGTGTGGAACGACAAAGCCAATTTCTGGCCGCT GGGCGCACTGGATCACATCAAATCCGACTTTAATTCAGGTCTGA AAAGCTGCTCTCTGATTTATCAGATCCCGTCTGATGCTCAACTGA AAAGTGGCGACACCGTGGCGTTCCAGGTTCTGGATGAAAACGGT AATGTGCTGGCGGATGACAACACGGAAACCCAGGGCGGCGGTG GATCCCACCATCATCACCACCATCATCATCACCACACATGCCCAC CGTGGCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCT TCCCCC AAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTG AGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAG GTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGC CAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGG GTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGG CAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCC CCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAA CCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAA GAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCA GCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAA CAACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTT CTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGC AGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACA ACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAA GCGGCCGC
11	CATATGGCGAGCAAACCGGACATCAAAGTGGGCGACTACGT GAAAATGGGTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTG TGAGCATCGACAACAATGGCCCGCTGATGCTGGCCGATAAAATT GTTGACACGCTGGCGTATGATGCTAAAACCAACGACAATTCGAA CAGCAAATCTCATAGTCGTTCCCTACAAACGCGATGACTACGGCA GCAACTATTGGAAAGATAGTAATATGCGCTCCTGGCTGAACTCA ACCGCGGCCGAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCC GAAAGACGGTTACGTCAGCGGCGTGGGTGCATATAATGAAAAAG

CTGGTTTTCTGAACGCGTTCTCAAAATCGGAAATTGCAGCTATGA AAACGGTGACCCAGCGTAGCCTGGTTTTCTCATCCGGAATATAAT AAAGGCATTGTTGATGGTGACGCGAACTCGGATCTGCTGTATTAC ACCGACATCAGCGAAGCAGTGGCTAACTACGATAGCTCTTATTTT GAAACCACGACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGC GAACGCCGTCTGGAAAAATCTGAAAGGCTATTACGTGGCTTACA ACAATGATGGTATGGCATGGCCGTATTGGCTGCGTACCCCGGTG ACGGATTGTAATCATGACATGCGCTATATTAGTTCCTCAGGCCAG GTTGGTCGTTACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCG GCGTTTTACCTGGACAGTGAATATTTCTGACGACCGAGCGGCTCT GGTAGTCAGTCGAGCCCGTACATTGGTTCCGCGCCGAACAAACA AGAAGATGACTATACCATCTCAGAACCGGCGGAAGATGCCAACC CGGACTGGAATGTTTCGACGGAACAGAGCATTCAACTGACCCTG GGCCCGTGGTACTCGAATGATGGTAAATATAGCAACCCGACCAT TCCGGTGTATACCATCCAGAAAACGCGCTCGGATACCGAAAACA TGGTGGTTGTCGTGTGCGGCGAAGGTTATACCAAATCACAGCAA GGCAAATTTATCAATGATGTTAAACGTCTGTGGCAGGACGCTAT GAAATATGAACCGTACCGTAGCTATGCGGATCGCTTTAATGTGTA TGCACTGTGTACGGCTTCCGAATCAACCTTCGATAACGGCGGTTTC TACCTTTTTTCGATGTGATCGTTGACAAATACAACTCTCCGGTTAT CAGTAACAATCTGCATGGCAGTCAGTGGA AAAAATCACATTTTTG AACGCTGCATCGGTCCGGAATTCATTGAAAAAATCCATGATGCC CACATTAAGAAAAAATGTGACCCGAACACCATCCCGTCGGGTAG CGAATACGAACCGTATTACTATGTGCATGATTATATTGCACAGTT TGCTATGGTTGTCAATACCAAATCCGACTTCGGCGGTGCATATAA CAATCGCGAATACGGCTTTCACTATTTTCATCTCTCCGAGTGATTC CTACCGTGCCTCTAAAACCTTTGCACATGAATTCGGCCACGGTCT GCTGGGCCTGGGTGATGAATACTCGAATGGTTATCTGCTGGATG ACAAAGAACTGAAAAGCCTGAACCTGTCTAGTGTGGAAGATCCG GAAAAAATTAAATGGCGTCAGCTGCTGGGCTTTCGCAATACGTA CACCTGCCGTAACGCGTATGGTTCTAAAATGCTGGTTTCCTCATA CGAATGTATCATGCGCGATACCAACTATCAATTTTGCGAAGTCTG TCGCCTGCAGGGCTTCAAACGTATGAGCCAACTGGTTAAAGATG TCGACCTGTATGTGGCCACGCCGGAAGTTAAAGAATACACCGGT
--

	GCATATAGTAAACCGTCCGATTTTACGGACCTGGAAACCTCGAG CTACTACAACTACACCTACAACCGTAACGATCGCCTGCTGAGTG GCAACTCAAAATCGCGTTTCAATACGAACATGAATGGCAAGAAA ATTGAACTGCGCACCGTTATTCAGAACATCAGCGATAAAAACGC CCGTCAACTGAAATTCAAAATGTGGATCAAACATTCAGATGGCT CGGTGGCAACCGACTCTAGTGGTAACCCGCTGCAGACCGTCCAA ACGTTTGATATTCCGGTGTGGAACGACAAAGCCAATTTCTGGCCG CTGGGCGCACTGGATCACATCAAATCCGACTTTAATTCAGGTCTG AAAAGCTGCTCTCTGATTTATCAGATCCCGTCTGATGCTCAACTG AAAAGTGGCGACACCGTGGCGTTCCAGGTTCTGGATGAAAACGG TAATGTGCTGGCGGATGACAACACGGAAACCCAGCGCTACACGA CCGTTTCTATCCAATACGGCGGCGGTGGATCCCACCATCATCACC ACCATCATCATCACACACATGCCACCGTGCCAGCACCTGAA CTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCAAACCCAAG GACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTG GTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTA CGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGG GAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCCTCAC CGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCA AGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAACCATC TCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCT GCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGA CCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGT GGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCC TCCCGTGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCT CACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCAT GCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAG AGCCTCTCCCTGTCTCCGGGTAAATAAGCGGCCGC
13	CATATGGCGAGCAAACCGGACATCAAAGTGGGCGACTACGTGAA AATGGGTGTGTATAATAACGCAAGCATCCTGTGGCGCTGTGTGA GCATCGACAACAATGGCCCGCTGATGCTGGCCGATAAAATTGTT GACACGCTGGCGTATGATGCTAAAACCAACGACAATTCTGAACAG CAAATCTCATAGTCGTTTCTACAAACGCGATGACTACGGCAGCA ACTATTGGAAAGATAGTAATATGCGCTCCTGGCTGAACTCAACC

GCGGCCGAGGGTAAAGTGGATTGGCTGTGCGGCAATCCGCCGAA
AGACGGTTACGTCAGCGGCGTGGGTGCATATAATGAAAAAGCTG
GTTTTCTGAACGCGTTCTCAAATCGGAAATTGCAGCTATGAAAA
CGGTGACCCAGCGTAGCCTGGTTTCTCATCCGGAATATAATAAA
GGCATTGTTGATGGTGACGCGAACTCGGATCTGCTGTATTACACC
GACATCAGCGAAGCAGTGGCTAACTACGATAGCTCTTATTTTGA
AACCACGACCGAAAAAGTTTTCTGCTGGATGTCAAACAGGCGA
ACGCCGTCTGGAAAAATCTGAAAGGCTATTACGTGGCTTACAAC
AATGATGGTATGGCATGGCCGTATTGGCTGCGTACCCCGGTGAC
GGATTGTAATCATGACATGCGCTATATTAGTTCCTCAGGCCAGGT
TGGTCGTTACGCTCCGTGGTATTCTGATCTGGGCGTCCGTCCGGC
GTTTTACCTGGACAGTGAATATTTCTGACGACCAGCGGCTCTGG
TAGTCAGTCGAGCCCGTACATTGGTTCCGCGCCGAACAAACAAG
AAGATGACTATAACCATCTCAGAACCGGCGGAAGATGCCAACCCG
GACTGGAATGTTTCGACGGAACAGAGCATTCAACTGACCCTGGG
CCCGTGGTACTCGAATGATGGTAAATATAGCAACCCGACCATTCT
CGGTGTATACCATCCAGAAAACGCGCTCGGATACCGAAAACATG
GTGGTTGTCGTGTGCGGCGAAGGTTATACCAAATCACAGCAAGG
CAAATTTATCAATGATGTAAACGTCTGTGGCAGGACGCTATGA
AATATGAACCGTACCGTAGCTATGCGGATCGCTTTAATGTGTATG
CACTGTGTACGGCTTCCGAATCAACCTTCGATAACGGCGGTTCTA
CCTTTTTTCGATGTGATCGTTGACAAATACAACCTCTCCGGTTATCA
GTAACAATCTGCATGGCAGTCAGTGGAAAAATCACATTTTTTGAA
CGCTGCATCGGTCCGGAATTCATTGAAAAAATCCATGATGCCCA
CATTAGAAAAAATGTGACCCGAACACCATCCCGTCGGGTAGCG
AATACGAACCGTATTACTATGTGCATGATTATATTGCACAGTTTG
CTATGGTTGTCAATACCAAATCCGACTTCGGCGGTGCATATAACA
ATCGCGAATACGGCTTTCCTACTATTTTCATCTCTCCGAGTGATTCTT
ACCGTGCCTCTAAAACCTTTGCACATGAATTCCGGCCACGGTCTGC
TGGGCCTGGGTGATGAATACTCGAATGGTTATCTGCTGGATGAC
AAAGAACTGAAAAGCCTGAACCTGTCTAGTGTGGAAGATCCGGA
AAAAATTAAATGGCGTCAGCTGCTGGGCTTTCGCAATACGTACA
CCTGCCGTAACGCGTATGGTTCTAAAATGCTGGTTTTCTCATACG

AATGTATCATGCGCGATACCAACTATCAATTTTGCGAAGTCTGTC
GCCTGCAGGGCTTCAAACGTATGAGCCAACTGGTTAAAGATGTC
GACCTGTATGTGGCCACGCCGGAAGTTAAAGAATACACCGGTGC
ATATAGTAAACCGTCCGATTTTACGGACCTGGAAACCTCGAGCT
ACTACAACCTACACCTACAACCGTAACGATCGCCTGCTGAGTGGC
AACTCAAAATCGCGTTTCAATACGAACATGAATGGCAAGAAAAT
TGAAGTGCGCACCGTTATTCAGAACATCAGCGATAAAAACGCCC
GTCAACTGAAATTCAAAATGTGGATCAAACATTCAGATGGCTCG
GTGGCAACCGACTCTAGTGGTAACCCGCTGCAGACCGTCCAAAC
GTTTGATATTCCGGTGTGGAACGACAAAGCCAATTTCTGGCCGCT
GGGCGCACTGGATCACATCAAATCCGACTTTAATTCAGGTCTGA
AAAGCTGCTCTCTGATTTATCAGATCCCGTCTGATGCTCAACTGA
AAAGTGGCGACACCGTGGCGTTCCAGGTTCTGGATGAAAACGGT
AATGTGCTGGCGGATGACAACACGGAAACCCAGCGCTACACGAC
CGTTTCTATCCAATACAAATTCGAAGATGGCAGTGAAATCCCGA
ATACGGCGGGCGGTACCTTCACCGGCGGCGGTGGATCCCACCAT
CATCACCACCATCATCATCACCACACATGCCCACCGTGCCCAGCA
CCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAAAA
CCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATG
CGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCA
ACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAG
CCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGT
CCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACA
AGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAA
ACCATCTCAAAGGCCAAAGGGCAGCCCCGAGAACCACAGGTGTA
CACCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCA
GCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCG
TGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACCTACAAGAC
CACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCTCTACAG
CAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCT
TTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGC
AGAAGAGCCTCTCCCTGTCTCCGGGTAAATAAGCGGCCGC

载体和细胞

在另一方面，本申请提供了一种载体，其包含编码本申请所述的 IgA 蛋白酶截短体的核酸或包含编码本申请所述的融合蛋白的核酸。

使用本领域公知的重组技术，可以将编码所述 IgA 蛋白酶截短体或融合蛋白的分离的多核苷酸插入载体，用于进一步的克隆（DNA 的扩增）或用于表达。有多种载体可供选择。载体组分通常包括但不限于下列的一种或多种：信号序列、复制起始点、一种或多种标记基因、增强子元件、启动子（例如 SV40、CMV、EF-1 α ）和转录终止序列。

在某些实施方式中，本申请提供的核酸编码 IgA 蛋白酶截短体或融合蛋白、与核酸序列可操作连接的至少一种启动子（例如 SV40、CMV、EF-1 α ）和至少一种选择标签。载体的实例包括但不限于：逆转录病毒（包括慢病毒）、腺病毒、腺相关病毒、疱疹病毒（例如单纯疱疹病毒）、痘病毒、杆状病毒、乳头瘤病毒、乳多空病毒（例如 SV40）、 λ 噬菌体和 M13 噬菌体、质粒 pcDNA3.3、pMD18-T、pOptivec、pCMV、pEGFP、pIRES、pQD-Hyg-GSeu、pALTER、pBAD、pcDNA、pCal、pL、pET、pGEMEX、pGEX、pCI、pEGFT、pSV2、pFUSE、pVITRO、pVIVO、pMAL、pMONO、pSELECT、pUNO、pDUO、Psg5L、pBABE、pWPXL、pBI、p15TV-L、pPro18、pTD、pRS10、pLexA、pACT2.2、pCMV-SCRIPT.RTM.、pCDM8、pCDNA1.1/amp、pcDNA3.1、pRc/RSV、PCR 2.1、pEF-1、pFB、pSG5、pXT1、pCDEF3、pSVSPORT、pEF-Bos 等。

可以将包含编码所述 IgA 蛋白酶截短体或融合蛋白的核酸序列的载体引入宿主细胞，用于克隆或基因表达。适用于克隆或表达本申请中所述的载体中的 DNA 的宿主细胞为上述的原核、酵母或高等真核细胞。适用于本申请用途的原核细胞包括真细菌，如革兰氏阴性菌或革兰氏阳性菌，例如，肠杆菌科（*Enterobacteriaceae*），例如，埃希氏菌属（*Escherichia*）（例如，大肠杆菌（*E. coli*））、肠杆菌属（*Enterobacter*）、欧文氏菌属（*Erwinia*）、克雷白氏杆菌属（*Klebsiella*）、变形杆菌属（*Proteus*）、沙门氏菌属（*Salmonella*）（例如，鼠伤寒沙门氏杆菌（*Salmonella typhimurium*））、沙雷氏菌属（*Serratia*）（例如，粘质沙雷氏菌（*Serratia marcescans*））、志贺氏菌属（*Shigella*）、杆菌属（*Bacilli*）（例如，枯草芽孢杆菌（*B. subtilis*）和地衣芽

孢杆菌 (*B. licheniformis*)、假单胞菌属 (*Pseudomonas*) (例如, 绿脓杆菌 (*P. aeruginosa*)、以及链霉菌属 (*Streptomyces*)。在某些实施方式中, 所述细胞是大肠杆菌细胞。

除了原核细胞以外, 真核生物细胞, 例如真核微生物如丝状真菌或酵母也可用作编码 IgA 蛋白酶截短体或融合蛋白的载体的合适的克隆或表达宿主。酿酒酵母 (*Saccharomyces cerevisiae*), 或面包酵母是最常用的低等真核宿主微生物。但是, 许多其他属、种和株都比较常用且在本申请中适用, 例如, 粟酒裂殖酵母 (*Schizosaccharomyces pombe*); 克鲁维酵母属 (*Kluyveromyces*) 宿主, 例如, 乳酸克鲁维酵母 (*K. lactis*)、脆壁克鲁维酵母 (*K. fragilis*) (ATCC 12,424)、保加利亚克鲁维酵母 (*K. bulgaricus*) (ATCC 16,045)、魏氏克鲁维酵母 (*K. wickerhamii*) (ATCC 24,178)、克鲁维酵母 (*K. waltii*) (ATCC 56,500)、果蝇克鲁维酵母 (*K. drosophilae*) (ATCC 36,906)、耐热克鲁维酵母 (*K. thermotolerans*) 和马克斯克鲁维酵母 (*K. marxianus*); 解脂耶氏酵母 (*Yarrowia*) (EP 402,226); 巴斯德毕赤酵母 (*Pichia pastoris*) (EP 183,070); 假丝酵母 (*Candida*); 里氏木霉 (*Trichoderma reesei*) (EP 244,234); 链孢霉 (*Neurospora crassa*); 西方许旺酵母 (*Schwanniomyces*), 例如, 西方许旺酵母 (*Schwanniomyces occidentalis*); 和丝状真菌 (*filamentous fungi*), 例如, 脉孢菌 (*Neurospora*)、青霉菌 (*Penicillium*)、弯颈霉 (*Tolypocladium*) 和曲霉菌 (*Aspergillus*) (例如, 钩巢曲霉 (*A. nidulans*) 和黑曲霉 (*A. niger*))。在某些实施方式中, 所述真核生物细胞是哺乳动物细胞。

在某些实施方式中, 所述哺乳动物细胞是人细胞或中国仓鼠卵巢 (CHO) 细胞。在某些实施方式中, 所述哺乳动物细胞是人胚胎肾细胞 293 (HEK293 细胞)。

药物组合物

在另一方面, 本申请提供了一种药物组合物, 其包含本申请所述的 IgA 蛋白酶截短体、包含本申请所述的融合蛋白、包含本申请所述的核酸、包含本申请所述的载体或者包含本申请所述的细胞, 以及药学上可接受的载体。

用于本申请中公开的药物组合物的药学上可接受的载体可包括, 例如, 药学上可接受的液体、凝胶或固体载剂、水相溶媒、非水相溶媒、抗微生物物质、等渗物质、缓冲液、抗氧化剂、麻醉剂、悬浮剂/分散剂、螯合剂、稀释剂、佐剂、辅料或

无毒辅助物质，其他本领域公知的组分或以上的多种组合。

适用的组分可包括，例如，抗氧化剂、填充剂、粘合剂、崩解剂、缓冲液、防腐剂、润滑剂、搅味剂、增稠剂、着色剂、乳化剂或稳定剂例如糖和环糊精。适用的抗氧化剂可包括，例如，甲硫氨酸、抗坏血酸、EDTA、硫代硫酸钠、铂、过氧化氢酶、柠檬酸、半胱氨酸、巯基甘油、巯基乙酸、巯基山梨醇、丁基甲基茴香醚、丁基化羟基甲苯和/或没食子酸丙酯。如本申请所公开，在包含本申请公开的 IgA 蛋白酶截短体或融合蛋白的组合物中包括一种或多种抗氧化剂如甲硫氨酸，可降低所述 IgA 蛋白酶截短体或融合蛋白的氧化。本申请进一步提供多种防止所述 IgA 蛋白酶截短体或融合蛋白氧化、延长其保质期和/或提高其活性的方法，例如，通过将本申请中提供的 IgA 蛋白酶截短体或融合蛋白与一种或多种抗氧化剂（例如，甲硫氨酸）混合来实现。

进一步地说，药学上可接受的载体可包括，例如，水相介质如氯化钠注射液、林格氏液注射液、等渗葡萄糖注射液、无菌水注射液、或葡萄糖和乳酸林格注射液、非水介质例如：植物来源的不挥发性油、棉花子油、玉米油、芝麻油、或者花生油、细菌抑制或真菌抑制浓度下的抗菌物质、等渗剂如：氯化钠或葡萄糖、缓冲液如：磷酸盐或枸橼酸盐缓冲液，抗氧化剂如：硫酸氢钠，局部麻醉剂如：盐酸普鲁卡因，助悬剂和分散剂如：羧甲基纤维素钠、羟丙基甲基纤维素或聚乙烯吡咯烷酮，乳化剂如：聚山梨醇酯 80（吐温-80）、螯合试剂如 EDTA（乙二胺四乙酸）或 EGTA（乙二醇双（2-氨基乙基醚）四乙酸）、乙醇、聚乙二醇、丙二醇、氢氧化钠、盐酸、柠檬酸或乳酸。作为运载体的抗菌剂可加入多剂量容器中的药物组合物中，其包括酚类或甲酚、汞制剂、苯甲醇、氯代丁醇、甲基和丙基对羟基苯甲酸酯、噻汞撒、氯苯甲烷铵和氯苯乙铵。适用的辅料可包括，例如，水、盐、葡萄糖、甘油或乙醇。适用的无毒辅助物质可包括，例如，润湿剂、乳化剂、pH 缓冲剂、稳定剂、增溶剂，或者醋酸钠、去水山梨糖醇月桂酸酯、三乙醇胺油酸酯或者环糊精之类的物质。

所述药物组合物可以是液体溶液、悬浮液、乳剂、丸剂、胶囊、片剂、持续释放制剂或粉末。口服制剂可以包括标准运载体如药物级的甘露醇、乳糖、淀粉、硬脂酸镁、聚乙烯吡咯烷酮、糖精钠、纤维素、碳酸镁等。

在某些实施方式中，所述药物组合物被制剂成可注射的组合物。可注射的药物组合物可以任何常规的形式制备，例如，液体溶剂、悬浮剂、乳化剂或适用于产生液体溶剂、悬浮剂或乳化剂的固体形式。注射制剂可包括现用的无菌和/或无热原溶液、使用前现与溶剂结合的无菌干燥的可溶物，如冻干粉，包括皮下片、注射即用的无菌悬浮剂、使用前现与介质结合的无菌干燥不溶产品，和无菌和/或无热原的乳剂。溶剂可以为水相或非水相。

在某些实施方式中，单位剂量的注射制剂包装在一个安瓿、一支管或一支带有针的针筒中。本领域习知，所有注射给药的制剂应为无菌无热原。

在某些实施方式中，通过将本申请中公开的 IgA 蛋白酶截短体或融合蛋白溶解于某适当的溶剂中可制备无菌冻干的粉末。所述溶剂可含有一种可提高粉末或由粉末制得的重组溶液的稳定性，或改善粉末或重组溶液的其他药理组分。适用的辅料包括，但不限于，水、葡萄糖、三梨糖醇、果糖、玉米糖浆、木糖醇、甘油、葡萄糖、蔗糖或其他适用的物质。溶剂可含有缓冲液，如枸橼酸缓冲液、磷酸钠或磷酸钾缓冲液或其他本技术人员公知的缓冲液，在一种实施方式中，缓冲液的 pH 为中性。在本领域公知的标准条件下进行对所述溶解进行随后的过滤除菌，然后冻干制得理想的制剂。在一种实施方式中，将所得的溶剂分装至小管中冻干。每支小管可容纳单次剂量或多次剂量的所述 IgA 蛋白酶截短体或融合蛋白或其组合物。每支小管中的装入量可略微高于每次剂量所需或多次剂量所需（例如 10%过量），从而保证取样精确和给药精确。冻干粉可在适当的条件下储存，如在约 4℃ 到室温范围。

用注射用水将冻干粉重溶得到用于注射给药的制剂。在一种实施方式中，可将冻干粉加至无菌无热原水或其他适用的液体载体中重溶。精确的量由选择的疗法决定，可根据经验值决定。

治疗或预防疾病的方法

在另一方面，本申请提供了一种治疗或预防 IgA 沉积相关疾病的方法，其包括向需要治疗或预防的受试者施用本申请所述的 IgA 蛋白酶截短体、本申请所述的融合蛋白或者本申请所述的药物组合物。

在另一方面，本申请提供了一种治疗或预防 IgA 沉积相关疾病的方法，其包括向需要治疗或预防的受试者施用 IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物组合物，其中，所述 IgA 蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75、SEQ ID NO: 76 或其组合。在某些实施方式中，所述 IgA 蛋白酶的氨基酸序列是如 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的氨基酸序列去掉信号肽序列之后形成的氨基酸序列。在某些实施方式中，所述 IgA 蛋白酶截短体与如 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的多肽具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性）。在某些实施方式中，所述 IgA 蛋白酶截短体与如 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的多肽具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性），并且仍然保持 IgA 蛋白酶的功能或活性（例如，蛋白水解活性、特异性切割 IgA 的酶活性等）。

SEQ ID NO	NCBI 登 录号	氨基酸序列
61	WP_2488 35846.1	MTKKITAIFLALCMAISVLPITIQAAASKPDIKVG DYVKMGA YNNASILWRCV SIDNNGPLMLADKIVDTLAYDAKTNDNS NSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKVD WLCGNPPKDG YVSGVGAYNEKAGFLNAFSKSEIAAMKTV TQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDSS YFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMA WPYWL RTPVTDCNHDMRYISSSGQVGRYAPWYSDLGVR PAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAE DANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQK TRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQDAMK YEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYN SPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPN TIPSGSEYEPYVVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLL DDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTTCRNAYGSK MLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLY VATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALD HIKSDFN SGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG VLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYG TKLDLTPAKTLYDYEFIKVDGLNKPIVSDGT VVTY YYYKNK NEETHNLTLVAAKAATCITAGNSAYYTC DGC DKWFAD ATGSVE
62	WP_0068 58468.1	MTKKITAIFLALCMAISVLPMTIQAAASKPDIKVG DYVKMG AYNNASILWRCV SIDNNGPLMLADKIVDTLAYDAKTNDN SNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKV DWLCGNPPKDG YVSGVGAYNEKAGFLNAFSKSEIAAMKT VTQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDS

		SYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMA WPLYWLRTPVTD CNHDMRYISSSGQVGRYAPWYSDLGVR PAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAE DANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQK TRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQDAMK YEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKHN SPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPN TIPSGSEYEPYYYVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYS DGYLL DDKELKSLNLSSVEDPEKIKWRQLLGRNTYTCRNAYGSK MLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLY VATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALD HIKSDFN SGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG VLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYG TKLDLTPAKTLYDYEFIKVDGLNKPIVSDGTVVYTYYYKNK NEETHNLTLVAAKAATCTTAGNSAYYTC DGC DKWFAD ATGSVEITDKTSVKIPAPGHTAGTEWKSDDTNHWHECSRC HDKKDEAAHDY GSDNVC DTCGY YKTVPHTHNLTLVAAK AATCTESGKEAYYKCEGCGKFYEDVLGTKEITDLASWGN AKIAHTTKQTVTKASSIKLKATSLTYNGKVRTPKVIVKDR TGKTLVKNTDYTVSYAKGRKYVGKYAVKITFKGKYS GTK TLYFTIKPKATSISSLKAGSKKFTVKWKKQATQTTGYQVQ YSASSKFSKAKTVTVGKN TTVSKKISKLSGKKKYVVRVT YKTVKINGKSIRIYSGWSKAKTVTTKK
63	WP_0053 63310.1	MTKKITAIFLALCMAISVLPMTIQAASKPDIKVG DYVKMG AYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTNDN SNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKV DWLCGNPPK DGYVSGVGAYNEKAGFLNAFSKSEIAAMKT VTQRSLVSHPEYNKGIVDGDANS D LLYTDISEAVANYDS SYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMA

		WPYWL RTPVTDCN HDMRYISSSGQVGRYAPWYSDLGVR PAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAE DANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQK TRSDTENMVVVVCGEGYTKSQQGKFINDVKRLWQDAMK YEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYN SPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPN TIPSGSEYEPYYYVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLL DDKELKSLNLSSVEDPEKIKWRQFLGFRNTYTCRNAYGSK MLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLY VATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALD HIKSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG VLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYG TKLDLTPAKTLYDYEFIKVDGLNKPIVSDGTVVTTYYYKNK NEEHTHNLTVA AKAATCTTAGNSAYYTCDGCDKWFAD ATGSIEITDKTSVKIPAPGHTAGTEWKSDDTNHWHECSRC HDKKDEAAHSASEWIIDTAATETAEGAKHKECTVCKKVL ETATIPATGSSHTHSYGVYVGMTYTAGNLIYQITSIDTATL GQSKVIGVVAACKNKITKITIPDRADCKGYRLNVTTIGNN AFAGCKALKKLTIGNKVTVIGKNAFKKCSKLKTVVIGKAV KTISSKAFIGDNKIKKITFKGKKLKTVNKNAFSKKAKKNIK SKKTKLKGNKKAIKLFFKKLKI
64	WP_0700 97494.1	MTKKITAIFLALCMAISVLPMTIQAASKPDIKVGDYVKMG AYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTNDN SNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKV DWLCGNPPKDGYSVSGVGAYNEKAGFLNAFSKSEIAAMKT VTQRSLVSHPEYNKGIVDGDANSDLLYTDISEAVANYDS SYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMA WPYWL RTPVTDCN HDMRYISSSGQVGRYAPWYSDLGVR PAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAE

		DANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQK TRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMK YEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYN SPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPN TIPSGSEYEPYYYVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLL DDKELKSLNLSSVEDPEKIKWRQLLGRNTYTTCRNAYGSK MLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLY VATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS GNSKSRFNTNMNGKKIELRTVIQNI SDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALD HIKSDFN SGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG VLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYG TKLDLTPAKTLYDYEFIKVDGLNKPIVSDGTVVTTYYYKNK NEETHNLTLVAAKAATCTEGGKEAAYKCEGCGKFYED VLGTKEITDLASWGNIAKIAHTTKQTVTKATPTANGKIVN YCSVCKKTLSTTVIPKASSIKLKATSLTYNGKVRTPKVIVK DRTGKTLVKNTDYTVSYAKGRKYVGKYAVKITFKGKYS GKTLYFTIKPKATSISSLKAGSKKFTVKWKKQATQTTGY QVQYSASSKFSKAKTVTVGKNTTVSKKISLKGKKKYV RVRTYKTVKINGKSIRIYSGWSKAKTVTTKK
65	WP_1603 40763.1	MTKKITAIFLALYMAISVLPMTIQAASKPDIVGDYVKMG VYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTNDN SNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKV DWLCGNPPKDGYSVSGVGAYNEKAGFLNAFSKSEIAAMKT VTQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDS SYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMA WPYWLRTPVTD CNHDMRYISSSGQVG RYAPWYSDLGVR PAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAE DANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQK TRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMK YEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYN

		SPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPN TIPSGSEYEPYVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLL DDKELKSLNLSSVEDPEKIKWRQLLGFRTYTCRNAYGSK MLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLY VATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS GNSKSRFNTNMNGKKIELRTVIQNISDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALD HIKSDFNSGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG VLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYG TKLDLTPAKTLYDYEFIKVDGLNKPIVSDGTVVTTYKKNK NEEHNLTLVAAKAATCTTAGNSAYYTCDGCDKWFAD ATGSVEITDKTSVKIPAPGHTAGTEWKSDDTNHWHECSRC HDKKDEAAHDYGSNDVCDTCGYKTVPHNLTVA AATCTTAGNSAYYTCDGCDKWFADATGSVEITDKTSVKIP APGHTAGTEWKSDDTNHWHECTVAGCGVHIESTKSAHTA GEWIVDTPATATTAGTKHKECTVCHRVLETQPIPSTGTEL KIIAGDNQIYNKASGSDVTITCNGDFAKFTGKIVDGSVDS SNYTAVSGSTVLTLKASYLGTLTDGSHTITFVYTDGEANA NLVTRTAGSGHIHDYGTWKSADNHWHECNCGDKKDE AAHSFKWVVDKEATATKKGSKHEECKICGYKRSAPVEIPA TGTSTAPTDTTKPNDTTKPGNTNGSEKSPQTGDNSNIFLW FALLFVSAAGVTGITAYNKKKKEHAE
66	MCJ7966 723.1	MTKKITAIFLALCMAISVLPMTIQAASKPDIVGDYVKMG AYNNASILWRCVSIDNNGPLMLADKIVDTLAYDAKTNDN SNSKSHRSYKRDDYGSNYWKDSNMRSWLNSTAAEGKV DWLCGNPPKDGYSVSGVAYNEKAGFLNAFSKSEIAAMKT VTQRSLVSHPEYNKGIVDGDANSDLLYYTDISEAVANYDS SYFETTTEKVFLLDVKQANAVWKNLKGYYVAYNNDGMA WPYWLRTPTDCNHDMMRYISSSGQVGRYAPWYSDLGVR PAFYLDSEYFVTTSGSGSQSSPYIGSAPNKQEDDYTISEPAE DANPDWNVSTEQSIQLTLGPWYSNDGKYSNPTIPVYTIQK

		TRSDTENMVVVVCGEGYTKSQQGFINDVKRLWQDAMK YEPYRSYADRFNVYALCTASESTFDNGGSTFFDVIVDKYN SPVISNNLHGSQWKNHIFERCIGPEFIEKIHDAHIKKKCDPN TIPSGSEYEPYYYVHDYIAQFAMVVNTKSDFGGAYNNRE YGFHYFISPSDSYRASKTFAHEFGHGLLGLGDEYSNGYLL DDKELKSLNLSSVEDPEKIKWRQLLGFRNTYTTCRNAYGSK MLVSSYECIMRDTNYQFCEVCRLQGFKRMSQLVKDVDLY VATPEVKEYTGAYSKPSDFTDLETSSYYNYTYNRNDRLLS GNSKSRFNTNMNGKKIELRTVIQNIISKDKNARQLKFKMWIK HSDGSVATDSSGNPLQTVQTFDIPVWNDKANFWPLGALD HIKSDFNGLKSCSLIYQIPSDAQLKSGDTVAFQVLDENG VLADDNTETQRYTTVSIQYKFEDGSEIPNTAGGTFTVPYG TKLDLTPAKTLYDYEFIKVDGLNKPIVSDGTVVTTYKKNK NEETHNLTLVAAKAATCTTAGNSAYYTCDGCDKWFAD ATGLVEITDKTSVKIPALGHTAGTEWKSDDTNHWHECSR CHDKKDEAAHDYGSNDVCDTCGYKTVPHTHNLTLVAA KAATCTTAGNSAYYTCDGCDKWFADATGLVEITDKTSVK IPALGHTAGTEWKSDDTNHWHECSRCHDKNDEAAHSTSE WIIDTAATETAEGAKHKECTVCKKVLETATIPATGSSHTN SYGVYVGMTYTAGNLIYQITSIDTATVGQSKVIGVVAACK NKKIKVTIPDRADCKGYRLNVTIGNNAFAGCKALEKLT GNKVTVIGKNAFKNCSKLETVVIGKAVKTISSKAFIGDNKI KKITFKGDKLKT VKKNAFSKKAKKNIKSKKTKLKGNKKA IKLFKKKLKIK
67	HAC1090 2.1	MKKYFEKTSIALIHAMMFI AIFGGEAMKTH TIDDITKYKM VVNAQGVKTENGTRTTTQVELGNYISLGKYN GKEILWRC VGEDENGALMLADNIIDTLPYDAKTNDNNRSKSHSRNYK RDTYGSNYWKDSNMRSWLNSTAVAGEVKWLCGNPPRED SVNGNAYDQKAGFLNDFSKAEIAAMKNVTQRSIVSHPEY NLGFHDGEGRSDELELNFDIENVASNFDSAYGENSTEKVFL LDVKQVNTVWKNFGNYIIGRNEQGM AWPYWL RTPVTD CNHDMRYVHSNGSVGREWPNTDYIGVRPAFYLDSDYYA

		TTSGDGSASNPHYVGSAPDKIEDDYTVAEPEEDPNQEWDIS LDQQLRLTLGPYYSSDGKYATPTIPVYTIQKTRSDTENMVI LICGEGYTKSQQQKFIDDVKKVWEGAMQYEPYRSYADRF NVYALCTASESNFNSGGSTFFDVVIDKKSGPMISVNSAW KNHIFERCIGPTFLEQIHDAHIPNKTDPDFTIWDDDKMYPP FYYVHKYINQFAVLVNTTQDFGGSHRNYKRGHILITPAD SPRAQKTFTHELGHGLLELGDEYMTTAAESTDYTSLNVA YTHDPEKVKKWQMLGFRKFTCNTSPSYTAYNSSWECLM RDTTYQFCEVCKLQGSKRMSQLIDGKSLYVADPEVKKYT VQYSKPSDFADTTYNGYYYFENYRNNVLLSGVDKNKFNT SMAGEKIQLRITVQNLSDTTQRYVTMKLWIKHAGGSVAT TTGGQRLEATQFTIPVWSEKSKFWPKGALSYESNMNSG LENCELIYQIPLDADLKHGDYVAFEVTDESGNILANDNTE TQTYANINIEYKFEDGTPMPNANKATIPVAVGSQLNWTAP STMFGHTFVRAEGHDQMVNGSGQTVTYYYTKQSNVHIH DWGEVKYTWTSNTCKAERVCKHDSHIESETVTATGTI TAATCKEKGKMKYTATFVNTAFSMQEKEVDIDFAEHTYG AWIEEVPATCIAGGMKAHYKCTVCGKYFDENKNETTEEA LKTPVSPYYGHSFGFWVEEQYATCQAPGRKGYKHCSICN KDYDASNTEITDFVIPINPDGHELGLVAEVPATCKDTGV KEHRDCRLCGKHCDPITRKEIADLTPTTNNHTYGELIPEV SPTTTEFGVKEHKDCTVCGRHFDKDGNEITELRIAKIGTHN VIVNGESKFYAHGESVTVTANEPAGKVFKGWQDASGKI VSTDKSYTFTVNGETTLTAVYEDKSSGGGEITPPAKKDGL SGGQIAGIVTGSAAVAGLGGFAVLWFVVKKKTFADLGAL LKKGFTAIGNFFKTLGEKIKALFTKKK
68	WP_0552 60806.1	MKKQLTALVLCICMVLSVLPFSSTQAAAEETSSVGTSNISI GDYIRLGNYNGQPILWRCVDVDEMGLMLSDQVLATMA YDAKTSENSATRSHSRNLKRGTYGSNQWRDSNMRSWLN SKAEAGKVEWLCGNPPKSGFVGENPYDQAAGFLNGFKED EIAAIKTVTQRSIVSHPEYNAGMIEGQGADLPYDTNIEAAA NGFDQAYYENVTDKVFLMDVKQINKVYQNNSQLGGSYHI

		AYKGGVRWPYWL RTPVTDCNHDMRYVETDGRIDRNAPY LGFYGV RPAFYLD TQYYQVTGGDGTADSPYRGAAVNKPE ENFTVSGDGPTPGQEWDVSLDKSIQLYLGPNYSKSKKYES ATIQVIQKTRSDNENMVVICGEGYTKGEQQKFVAAAK RLWEGAMQYEPYKTYKDRFNVYALCTASDKTYS AIDGY DSTFFDVWGKNISVNGSQWKNHIFERCIGPAFIEKVHDAHI PQQADPNVDWDFEKYKYVHDYISQFVLLVNSANDFGGAF NDL DYGFHYIVSPAYSQRAVETLTHELGHGLLWLGDEYN SGSFMGEASEKTSLNRTGISDPEQVKWRQLLGFRKTYSP HTDYDTDKIYNSSRECMRQTWNGFCEVCKLQGNKRM RQLVTEGPDLYVAEPEVTKRTDAYTKLSDFSDATGWGYTK FDADKKTRLLTGADKITFQPTMKGQKIELRTIVQNLSDT KLSQVTLQVWVNHADGTIATADGQPVAASETFKIPLWTE KGNFRPKGTLEYHGSDFN SGLKNCSLLYTIPSNADLRTGD TVGYAVRDEQGTVLAYEGTLPNKGQDILPAPEPAKSYTVT FCYNDGRENTTKTTGINGKLGDLPAPAREGYVFDGWYTT GGEKVDLTRVYSSNTVLYARWSEYIAPSPNVKKAPVILLA ASPDTVTEGEQVMLSVSETSGFGVDLSGVTYTADPSLPISG TGEAQ TIRLDQAGTYTFTAHYSGDNEKYLAADSNRVTVT VTKKADVSGGTTSGGGSSSGGGSTAGGGSSAGGGSSAGG GSSSGGGAAGGATAGGGAANGNASATTPDIKDSDGTTV AIVNGKKGMITAEVQLSEKAIAANAEKSGEAVKLPVEVKA GKNIKAASTVTINLPEGAGKTKVKIPVKNMTAGTVAVLV NADGTEKIVKKSVA AKDGVQLIVDEDTTVKLV DKA NFK DTKKHWA KDSIDFVSARGLMNGKSSTAF APEAKITRRL WTILARWEDVDLTGGKKWYSKARAWAKNQGISDGSRPN AAITRAEAITMLWRAQGKPAAEQETA FKDVSSDEYYAQA VAWAKEKGIAQANSKGRFNPDA ACTRAEIAAFLYRMSLS E
69	CDE2481 1.1	MKKNFGKASIALIHAMMFI LAFFGGDVMKAHTIDGSASNL YVGSALDKVEDDYTVAEPEKAPNQELDISIEQSLNLTGE WYSSDGKYANPTIPVYTIQKTRSDTENMVILICGEGYTKN

		QQQKFVNDVKKVWEGAMQHEPYRSYADRFNVYALCTVS ESSFNSSGGSTFFDVVIDKNSGPMIAISKSICKNHIFERCIGPA FLEQIHDVHIPKKVDPNSSYWVGNNSPLESEYEPFYVHEYI NQFAILVNTTQDFGGSHRNYERGIHYLVTPADSDRAQKTF THELGHGLLELGDEYMSSTTQQTDLTSLNVAHTHDPNNV KWKQLLGFRKTYTCNALGYGNAYNSSYECLMRDTAYQF CEVCKLQGSKRMSQLIDGKSLEYAVPEVKKYTGQYSKPS DFIDTTYNGYYYFENYRKGVLLSGTDKNKFNTSMAGETIQ LRTIVQNLSDTKQRYVTMKLWIKHADGSVATTTGGQRLE ATQFTTIPVWSEKSKFWPKGALSYEGSHMNSGLENCELIY QIPSDAVLNNGDTVAFEVIDENGNILANDNTETQAYANINI EYKFEDGTPMPNVNQAMIPVAVGSQLNWTAPSTMFGHKF VRAEGHDQMVNGSGQTVTYYYKKQSNVHIHDWGEVIYT WISDNICKAERVCKHDSAHESETITATGTVIKASTCTEKG KVKYTARFTNTAFGVQYREVDIDLVEHKFGEWIDEIPATT ENFGTKGHKDC TVCKKHFDKDGNEITDLRIAKISTYTVTV KDGADETNTHYKSGDTVTIKVTIPTGKHFKWSAVTGISL SASQLTQEEITFTMPDNDVTLIAELEDILYRVTVIGGTTTLN EAKYQENVTVTANTPEVGKEFDKWIVSGITLSNTNLTRST LTFTMPENDVTFTATYKDIVYRVMVEEGKATPEMAIYQTE VTMANEPaidMYFDKWEVMGLDTTGMDLTKTQIKFQM PAGNVTFKATYLPHIKYGILVVDGTDKDKSPVMAGEIVTIT ANPAKPGKVFDKWT CETVGV TIEFASATSKQTTFVMPAQ DIKIKAHFKDIEVAPSVEIKVEGGTGAGTYKPGDSVTITAN EPAEGKVFKCWKDEKGEIVSTDRSYAFIVNGETTTLTAVYE DKASGGAIAGIVIGSILGVGIIGFVIFWFAVKKKDVf
70	HCS2457 7.1	MRWNIMKKYFGKANIALIHAMMFILAFFGGEAMKTHIDG SASNLYVGSALDKVEDDYTVAEPEKAPNQELDISIEQSLN LTLGEWYSSDGKYANPTIPVYTIQKTRSDTENMVILICGEG YTKNQQQKFVNDVKKVWEGAMQHEPYRSYADRFNVYA LCTVSESSFNSSGGSTFFDVVIDKNSGPMIAISKSICKNHIFE RCIGPAFLEQIHDVHIPKKVDPNSSYWVGNNSPLESEYEPFY

		YVHEYINQFAILVNTNQDFGGSHRNYERGIHYLVTPADSD RAQKTFTHELGHGELLELGDEYMSSTTQQTDLTSLNVAHT HDPNNVKWKQLLGFRKTYTCNALGYGNAYNSSYECLMR DTAYQFCEVCKLQGSKRMSQLIDGKSLYVAVPEVKKYTG QYSKPSDFIDTTYNGYFFFENYRKGVLSSGTDKNKFNTS MAGETIQLRTIVQNLSDTKQRYVTMKLWIKHADGSAVATT TGGQRLEATQAFTIPVWSEKSKFWPKGALSYEGSHMNSG LENCELIYQIPSDAVLNNGDVAFAFEVIDENGNILANDNTET QAYANINIEYKFEDGTLMPNVNQAMIPVAVGSRLNWTAP STMFGHKFVRAEGHDQMVNGSGQTVTYYYKKQSNVHHI DWGEAIYTWTSNICKAERVCKHDSAHESETITATGTVIK AATCTEKGVKYTARFTNTAFGVQYREVDIDLVEHKFGE WIDEIPATTENFGTKGHKDCVCKKHFDKDGNEITDLRIA KISTYTVTVKDGADETNTHYKSGDVTIKVTIPTGKHFVK WSAVTGISLSASQLTQEEITFTMPDNDVTLIAELEDILYRV TVIGGTTTTLNEAKYQENVTVTANTSEVGKEFDKWIVSGIT LSNTNLTRSTLTFTMSENDVTFTATYKDIVYRVMVEEGKA TPEMAIYQTEVTVMANEPAYDMYFDKWEVMGLDTTGMD LTKTQIKFQMPAGNVTFKATYLPHIKYGILVVDGTDKSP VMAGEIVTITANPAKPGKVFDKWTCTETVGTIEFASATSK QTTFVMPAQDIKIKAHFKDIEVAPSVEIKVEGGTGAGTYK PGDSVTITANEPAGKVFKCWKDEKGEIVSTDRSYAFTVN GETTLTAVYEDKASGGAIAGIVIGSILGVGIIGFVIFWFAVK KKDVF
71	MBD902 5975.1	MILLQIYYTKEGVKMKNKQINRTLSLLSVVMVLSLCPLI AKAEGTKPNIKIGDYIKLGTYENEPILWRCVDIDDNGPLM LMDKVLGSMPLYDAKTSSENSATRSHSRNSFRSSYGSNHWR DSNMRSWLNSDADAGKVDWLCCGNPPKSDYVGYGSEYDK KAGFLNGFSKAEIAAIKTVTQRSIVSHPEYSAGYIAGPGAD LPYNTDIASVAYGFEEKAYYENIIDKVFLPDVKQLNTIYNNS NILGNYYLAKNKDGIRWSYWL RTPITDCNNDHMRYVETDG NIYRVAPYFGHIGVRPAFYLDTDYIYVSEGNGEVNSPYVG

		DAADKPGDDISISGPDEEGGDGDWDIDTDQSIQLNLGPWY SSDGEYANSTIPVQVIQKTRSDLENMVIVICGEGYTKDQQ QKFINDVKRIWAGVLKHEPYRSMADRFNVYALCTASKTS GFASENTFFDITMSTTSRSPMISLYKSVLKNQILTRCIGPAFI EKIHDAHIKEKTNPNEITIGDEYAPYYYVNEYISQFVVLVN SGQYGGASMNNLDVGLHYVTATVDNIQSEYTLAHELGHG LLHLGDEYNAYGGAYTMPEQQDKQSLNIAGLRESPITIKW KDMLGFRKTYTCRDSNTSNSSNMVNSSWQCMMRTQNQE LCDVCQLQGFKVMSQLIKDTDDIYIAIPEVKLYTGNYKNP FEDYSAYTEAEYYGYLAYASDRAQRLLSGTSKNKFTKDM KGQEVELRTIAQNLSGIEEQEITLQLWVEHEDGTRAVTEN GEEILKEQTFTVPVWDEKENFYIKGMRNYSGTEFDSGLMN CSLIYKIPENADLKDGDITIKFSVIDKMGKTLADDNTETQN YANVTISYQLEDSSNAVPNTQTAVIPVPIGTKMDIEPPGELY GYKFVKAEGLGKIVGDDGLNIICYEDPSGKLPVEYKVEY DWGTDFTDTTLPTDNTKYDSIENAKESVKNQKYDENSTS TVKKNDKDGTWTFSGWTATVEGTTVKFTGAWTFTATPII TYTVTYDWGTDFTGEMPLPVDSKTYKSEEDAKAAMDGK YTSLSSTAEKDGKSGTWKFSGWIATLIGTTVKFNGMWTF TPDAPVVDADTPTNIKLVSDEYKIGDKATALDGKATVSDN GVLSYQWYKSDKADNFNGTAIDGQNGETFPDTSKEGTY YYYVIATNTKADATGKKTASVTSSMAIIHVKESVKYTVVY DWGSDYPTDVTVPKDDTKYENIEKAKEAVKNQKYDENST STAEKNSKSGKWSFSGWTTAVEGTTVKFTGVWTFTEANAIP VVTRKPSSGGSGSGSSTYNIKVSPEITNGSLSVNPSRASN GKKVSVIVKPNNGYVLNSVIVKDSNDKEIAVTKQSDGTYT FIMPSSNVTVSASFDTLAKDVVTEIEKSIEFKDVKKGDW YFDAVQWAVKNNITEGSGKDTFSPDVICTRTQMVTFLWR VAGSPEPKITKCDFRDVDNSAYYYKAVLWAVEKGVTVGT SDTTFSPNENVTRGQTVAFLYREAGSPFETGEDVFNVDNS NDYYFKAVSWATKNGITVGTGNGKFEPDMDCNRAQIVA MLYRTQR
--	--	--

72	CDE1602 7.1	MKKHLKKT S I A L T I V M M F I P A I F G G K A I T I H T N A D N T N Y K T A V N A Q G V K T E K E T K A T T Q V E L G N Y I S L G K Y N G N E V I W R C V S I D E K G A L M L A D N I I D T L P Y D A K T N D N N H S K S H S R N N N R D N H G S N Y W K D S N M R S W L N S T A V A G E V T W L C G N P P R A G Y L N E N A Y D Q K A G F L N D F S K A E I A L M K N V T Q R S L V S H P E Y N H G F H D G D G H S D L E F N E N I E N V S S N F N S A Y G E N S T E K V F L L D V Q Q V N K V W E N F D N Y Y I G R K E G V A W P Y W L R T P L S S C N H L M R Y V G S N G L V G K D Y P T N A I G V R P A F Y L D S D Y Y V T T S G N G S A S N P Y V G S A P D K I E G D Y I I A E P E E D P N Q E W D V S L D Q Q L R L A L G P Y Y S S D G K Y S T P T I P V Y T I Q K T R S D T E N M V I L I C G E G Y T K S Q Q Q K F I N D A K K V W E G A I Q Y E P Y R S Y A D R F N V Y A L C T A S E S S F N N S G S T F F D V V I D K K V G P M I S V N K S S W K N H I F E R C I G P A F I E Q I H D A H I P N K T D P D T F I W D D D K M Y S P F Y Y V H K Y I N Q F A L L V N T S Q D F G G S H R N Y K R G I H Y F I T P A D N N R A V K S F A H E L G H G L L E L G D E Y M T V A A E S T D Y T S L N V A Y N H N P E Q V K W K L L G F R K T F T C N T Y P F Y T A Y N S S W E C L M R D T N Y Q F C E V C K L Q G Y K R M S Q L I D G K N L Y V A D P E V K K Y T D R Y T N P S D F A E T N Y N G Y I N F T N Y R D E I L L S G W N K N K F N T G M V G E K I Q L R T I V Q N L S D T T E R Q V T M K L W I K H A D G S I A T T T N S Q R L E A T Q T F T I P V W S E K S K F W P K G A L E Y N G S N L N S G L E N C E L I Y Q I P S D A V L N N G D T V A F E V T D E N G N V L A H D N T E T Q P Y A N V N I E Y K F E D G S P M P N A N K A V I P L A V G S Y I N W T A A P S L Y G Y A L S R V E G L N Q I V S G S D Q T V T Y Y Y T K K I G T H I H D W G D W V S N G D G T H T R T C T K D S S H T E T E N C S G G T A T C T T K A V C S V C G F T Y G E K L G H N W G E V K Y T W M S D N I C K A E R V C K H D S T H I E S E M V T A T G T V I T K A T C K E K G K M K Y T A T F V N T A F T V Q E K E I D T D F A P H T F G A W K D E I P A T T E E F G T K G H K D C M D C G R H F D K D G K E I T E L R I A K I G T Y N V V I N G E S K F Y A D G E S V T V K A E D K E G K I F K G W Q D E S G E I V S T E K S Y T F T V T G D R S L T A V Y E N V L A T K K G L S D R Q I A G I I G S V V A A G L G G F A I F W F V I K K K G L R L
73	WP_1182	M N I I K H K Y G K R T V S L L L A V I L V L C P L P V R A A D N K P T I E I G D Y I Q M G T Y G G V P I V W R C V A K D S N G P L M L S D R V L C D Y M P Y

	65594.1	DAKTNKNAETGSHRRNSWRDNFGSNHWRDSNIRSWLNS NAEAGKVKWLCGNPPTEDSVYPKTAAYDQKEGFLRSFRS DELGAIRTVKQRSIVSHPEYTAGYIDAAGVDLLYNTTIDTV ADGYDSAHYEYIWDRVFLLDVQQLKTVNDNLNGYHIAK NRSGVAWNYWL RTPITTCNHDMRFVTPQGNILRDAPYKG YYGVRPAFYLN TENYTVSSGTGQSAQDPYVVSAPDATDD SIGISGAVREDVNGDWNVNTDEYLQLEMSTLYTEDPAYA NVTVPVYTIQKPRSDKENMVIVYCAEGYTKSQKQFVED VKKLWGEVLQIEPYRSMADRFNVYALCTASVDGYGGTST FFAATAKGGISTNKG NWRNVHLERIIGPAFIEKIHDAHIPN ETHPNENTMDHNYRQYDYVYENINQFVVLANSGEYFGGS HDNKQYGIHYIVASARNAYS AFTQRHELGHGLFHLGDEY NYSTVPVDEWNYTTSLNMTATKDPTKVWKQLLGRNT YTCPHLDYYPYTYNSSRDCLMRETFQND FCDVCKLQGIK VMSQLITNPPALYVAVPEVKKYIGGYRNPTKDPSAFEAAAN SSAYASYQNDRNSRLLSGGSKNSFDYSSMKGQQVELRTII QNLSNTQAKTVTLRLWVEHSNGEKAVTTDGEQVFTTQEF DIPVWKEKSKFWTKGALDYEGSDFNSGLVNC SLVYTIPEN AILQSGDTIGFEIVDHATGEVLADDDTEQQRVNVVTIQYQ LEDGTDVPNTMPTTFTVPVGKKVDWQPPQELHG YTFVKA EGMENAVPNSGMTIRYIYKRSEERPEPPVTKNYTVQYNW GSVFPTGATLPLNSSSYSSVQAKAAVDKKYTSTTRIQAQ KDGKNGTWAFSGWDAGNLNGTTVVFRG SWSFTADTAPIT PPSGTASYKITATAGIGGTISPGGTTTVSAAGQLTYTIKANE GYYIADV KVDGSEVTATTSYTFSDVNTDHTIEVTFKQESQ TPDVPDVIAPSITTQPGNATVKVGETASFTIAASGTDLT YQ WQIDRNDGKGWVNIDGATATS YTTSTVNISCNGFKYKCV VSNSAGNVESNSATLTVQDAGGSDNPDPNNTYQIHDGAN SSWTHDS DGNITIRGNGDFS KFTGVKVDGNLIDKSNYTAK EGSTIITLKASYLN TLSAGNHTVEILWTDGSASTFTTKANI SDNSNNNQNDNNNSNSSDDKPSSGTDKKDVTAPKTGDNT PSVWLFILSILSGTGLIITVKKRRENLS
--	---------	--

74	WP_2431 21302.1	MNVIKHRYGKRTVSLLLAVILVLCPLRVRAADNKPTIGIG DYIKSAQDPYVVSAPDAPDYSIGISGAVRDDVNGDWNVN TDEYLQLKMSTYYTEDITAYANVTVPVYTIQKPRSDKENM VIVYCAEGYTKSQKQFIEDVKKLWGEVLQIEPYRSMAD RFNVYALCTASVDSFGGTSTFFNATKKGISNSKGAWRNHI LERIIGPAFIEKIHDAHIPNKTHPNENPGDHDYRQYDYVYE NINQFVLLANSGEYFGGSHDNKEHGIHYIIASARSQYSFT QRHELGHGLFHLGDEYNYSTVPVAEANYTTSLNMTATKD PTKVWKQLLGFRNTYTCPHDDRYPTYTNSSRDCLMRET FQNDFCDVCKLQGIKVMSQLITNPPALYVAVPEVKKYIGD YRNPTEKPSAFEAAANSSAYASYQYDRNSRLLSGGSKNSFD YGSMKGQQVELRTIIQNLSDTQAKTVTLRLWVEHSNGEK AVTTEGQQVFATKDFAI PAWSEKSKFWPKGALDYKGSDF DSGLVNCSLVYTIPENATLQYGDTIGFDIVDRATGEVLAH DDTEKQPYADVTIQYQLEDGTDVPNTMPTTFTVPVGKKV DWQPPQELNGYTFVQAEGMEETVPSNGMTIRYIYKRTEER PEPPVTKNYTVKYDWGSVFPTGVTLPQNSNSYSSEQQAK AAVDKKYTTSTRIKAQKDGNKGTWAFSGWDSGSLNGTT VVFRGSWSFTADTAPITPPSGGGGGGGAATTASYKITATA GIGGTISPGGTTTVSAAGQLTYTIKANEGYYIADV KVDGK SVGAVSSFAFEKITASHTIEASF AKKDSATVKDPIKRLPIHL PIKMKVKKM
75	UK14907 4.1	MKKTYLFKSYKRKSQKEVEHNEKNILKKTIIALIIAMMFIL AIFFGGEAMKTHTIDDVAKYKMVVNAQGVKMGNTRTQ TQVDLGNYISLDQQLKLTLPYYSSDGKYSTPTIPVYTIQK TRSDTENMIILICGEGYTKSQKQFIEDVKRVWDGAIQHEP YRSYADRFNVYALCTASESSFNSSGGSTFFDVVIDKKSGPRI SGNKS AWKNHIFERCIGPTFLEQIHDAHIPNKTDPDFTIWD DDKMYPPFYVHKYINQFAVLVNTEQDFGGSHRNYKSGI HYLITPADSPRAQKTFTHELGHGLLELGDEYMTSATESTD YTSLNVAYTHDPEKVKWEKMLGFRKFTTCHTNSSYTAYN SSWECLMRDTTYQFCEVCKLQGSKRMSQLIDGKS LYVAD

		PEVKKYTGQYSQPSDFADTTYNGYANFSYYRSGVLLSGW DKNKFNTDMAGEKIQLRITIVQNLSDDTQRYVTMKLWIKH ADGSVATTTGGQRLEATQFTTIPVWSEKSKFWPKGALSIE GSNMNSGLENCIELYQIPLDAVLNKGDTVAFEVTDENGIV LANDNTETQTYANINIEYKFEDGTPMPNVNKATIPAVGSK LNWTAPSTMFGHTFVRAEGHDQMVNGSDQTVTYYYAKQ SNVHIHDWEEWVSNGNGTHTRTCRTDNSHSETANCVGGT ATCTHKPVCEVCHGEYGGQAKSHDWGKATYTWTDTVCKA ERVCKHDSAHTESETRATGTVIKAATCKEKGKMKYTAT FENTAFTKQEKKVDINFAGHTFGKWQDEIPATTEAFGTKG HKDCSVCGRHFDKDGNEITELRIKIVTHNVIVNGESKFY AHGESVTVTANEPAEKGKVFKGWQDASGKIVSTKKSFTT VNGETNLTAVYEDKTSGGEIVPPAKKDGLSGGQVAGVVI GSAAVAGIGGFAIFWFTVKKKTFADLIAAIKSLFTKKKTK
76	WP_0056 04305.1	MNIKHKHYGKRTVSLLLAVILVLCPLQVRAADNKPTIEIGD YIQMGTYGGVPIVWRCVAKDSNGPLMLSDRVLCDYMPY DAKTNKNAETGSHRRNSWRDNFGSNHWRDSNIRSWLNS NAEAGKVKWLCGNPPTEDSVYPKTAAYDQKEGFLRSFRS DELGAIRTVKQRSIVSHPEYTAGYIDVAGVDLPYNTTIDTV ADGYDSAHYEYIWDRVFLLDVQQLKTVNDNLNGYHIAK NRSQVAVNYWLRTPITTCNHDMRFVTPQGNILRDAPYKG YYGVRPAFYLNTEYTVSSGTGQSAQDPYVVSAPDAPDD SIGISGAVREDVNGDWNVNTDEYLQLEMSTLYTEDPAYA NVTVPVYTIQKPRSDKENMVIVYCAEGYTKSQKQFVED VKKLWGEVLQIEPYRSMADRFNVYALCTASVDGYGGTST FFAATAKGGISTNKGWNRNHVLERIIGPAFIEKIHDAHIPN ETHPNENTMDHNYRQYDYVYENINQFVVLANSGEYFGGS HDNKQYGIHYIVASARNAYSFTQRHELGHGLFHLGDEY NYSTVPVDEWNYTTSLNMTATKDPTKVWKQLLGRNT YTCPLDYYPYTYNSSRDCLMRETFQNDFCDVCKLQGIK VMSQLITNPPALYVAVPEVKKYIGGYRNPTKDPSAFEAN SSAYASYQNDRNSRLLSGGSKNSFDYSSMKGQQVELRTII

		QNLSTQAKTVTLRLWVEHSNGEKAVTTDGEQVFTTQEF DIPVWKEKSKFWTKGALDYEGSDFNSGLVNCSLVYTIPEN AILQSGDTIGFEIVDHATGEVLADDDTEQQRVNVNTIQYQ LEDGTDVPNTMPTTFTVPVGKKVDWQPPQELHGYTFVKA EGMENAVPNSGMTIRYIYKRSEERPEPPVTKNYTVQYNW GSVFPTGATLPLNSSSYSSVQQAQAAVDKKYTSTTRIAQ KDGKNGTWAFSGWDAGNLNGTTVVFRGSWSFTADTAPIT PPSGTASYKITATAGIGGTISPGGTTTVSAAGQLTYTIKANE GYIADV KVDGSEVTATTSYTFSDVNTDHTIEVTFKQESQ TPDNTYQIIDGANSSWTPSDGNITIRGNDFSKFTGVKVD GNLIDKSNYTAKEGSTIITLKASYLNTLSAGTHTVEILWTD GSASTTFTIKANTSDDKPSSGTDKKDDAPKTGDNTPSVWL FILSILSGTGLIITVKKRRENLS
--	--	---

在另一方面，本申请提供了本申请所述的 IgA 蛋白酶截短体、本申请所述的融合蛋白或者本申请所述的药物组合物在制备用于治疗或预防 IgA 沉积相关疾病的药物中的用途。

5 在另一方面，本申请提供了 IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物组合物在制备用于治疗或预防 IgA 沉积相关疾病的药物中的用途，其中，所述 IgA 蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、
10 SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75、SEQ ID NO: 76 或其组合。在某些实施方式中，所述 IgA 蛋白酶的氨基酸序列是如 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的氨基酸序列去掉信号肽序列之后形成的氨基酸序列。在某些实施方式中，所述 IgA 蛋白酶截短体与如
15 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、

SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、
SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75
或 SEQ ID NO: 76 所示的多肽具有至少 70% 的序列同一性（例如，具有至少 75%、
至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、
5 至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、
至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性）。
在某些实施方式中，所述 IgA 蛋白酶截短体与如 SEQ ID NO: 61、SEQ ID NO: 62、
SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、
SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、
10 SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的多肽具
有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、
至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、
至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、
至少 97%、至少 98%、至少 99% 的序列同一性），并且仍然保持 IgA 蛋白酶的功能或活性（例如，蛋白水解活性、特异性切割 IgA 的酶活性等）。
15

在另一方面，本申请提供了用于治疗或预防 IgA 沉积相关疾病的如本申请所述的 IgA 蛋白酶截短体、融合蛋白或者药物组合物。

在另一方面，本申请提供了用于治疗或预防 IgA 沉积相关疾病的 IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物组合物，其中，所述 IgA 蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、
20 SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、
SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、
SEQ ID NO: 75、SEQ ID NO: 76 或其组合。在某些实施方式中，所述 IgA 蛋白酶的
25 氨基酸序列是如 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、
SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、
SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、
SEQ ID NO: 75 或 SEQ ID NO: 76 所示的氨基酸序列去掉信号肽序列之后形成的氨基酸序列。在某些实施方式中，所述 IgA 蛋白酶截短体与如 SEQ ID NO: 61、SEQ

ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的多肽具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性）。在某些实施方式中，所述 IgA 蛋白酶截短体与如 SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75 或 SEQ ID NO: 76 所示的多肽具有至少 70% 的序列同一性（例如，具有至少 75%、至少 80%、至少 81%、至少 82%、至少 83%、至少 84%、至少 85%、至少 86%、至少 87%、至少 88%、至少 89%、至少 90%、至少 91%、至少 92%、至少 93%、至少 94%、至少 95%、至少 96%、至少 97%、至少 98%、至少 99% 的序列同一性），并且仍然保持 IgA 蛋白酶的功能或活性（例如，蛋白水解活性、特异性切割 IgA 的酶活性等）。

在某些实施方式中，本申请所述的 IgA 沉积相关疾病包括 IgA 肾病、疱疹样皮炎、亨诺赫-舍恩莱因紫癜（又称 IgA 血管炎）、川崎病、紫癜性肾炎、IgA 血管炎肾损害、IgA 类风湿因子阳性的类风湿性关节炎、IgA 型抗 GBM 病或 IgA 型 ANCA 相关血管炎。在某些实施方式中，本申请所述的 IgA 沉积相关疾病为 IgA 肾病。在某些实施方式中，本申请所述的 IgA 沉积相关疾病为 IgA1 肾病。在某些实施方式中，本申请所述的 IgA 沉积相关疾病为 IgA 血管炎。在某些实施方式中，本申请所述的 IgA 沉积相关疾病为川崎病。

具体实施方式

所有实施例中的涉及的生物学材料如大肠杆菌菌株、各种克隆与表达质粒、培养基、工具酶、缓冲液，和各种培养方法、蛋白提取和纯化方法、其它的分子生物学操作方法，均为该领域技术人员所熟悉，可以参考 Sambrook 等人编著的“分子克隆”（实验室手册，冷泉港，1989）及“精编分子生物学实验指南”（美/F. 奥斯伯等著，颜子颖等译，北京，科学出版社，1998）。

实施例 1: AK183 IgA 蛋白酶的最短活性位点研究

发明人将来自多枝梭菌 (*Clostridium ramosum*) AK183 株的野生型 IgA 蛋白酶 (其氨基酸序列如 SEQ ID NO: 1 所示) N 末端的信号肽 (即, SEQ ID NO: 1 的第 1 至 30 位氨基酸) 及 C 末端的跨膜区加胞内区 (即, SEQ ID NO: 1 的第 1205 至 1234 位氨基酸) 去掉, 然后将人 IgG1 的 Fc 序列 (HR-CH2-CH3, 其氨基酸序列如 SEQ ID NO: 24 所示) 加在了去掉了信号肽、跨膜区和胞内区的 IgA 蛋白酶的氨基酸序列 (即, 由 SEQ ID NO: 1 的第 31 位至 1204 位氨基酸组成的 IgA 蛋白酶截短体) 的 N 末端, 构建了 PET30a-Fc-AK183 质粒。

然后, 发明人以 PET30a-Fc-AK183 质粒为模板, 进行终止突变, 构建了一系列 Fc-AK183 截短体, 以研究 AK183 IgA 蛋白酶的 C 末端最短活性位点。基于前期研究结果, 发明人认为在 AK183 IgA 蛋白酶的第 730 位至第 840 位氨基酸之间存在自切位点。因此, 发明人进行了第一轮终止突变, 终止突变位点为 AK183 IgA 蛋白酶的第 738 位、第 769 位、第 799 位、第 834 位这四个氨基酸位点, 结果如图 1 所示。结果显示, 第 738 位、第 769 位的氨基酸终止突变后得到的 AK183(31-737)、AK183(31-768) IgA 蛋白酶截短片段无体外酶切活性, 而第 799 位或第 834 位的氨基酸终止突变后得到的 AK183(31-798)、AK183(31-833) IgA 蛋白酶截短片段有活性。因此, 第一轮终止突变结论为: AK183 IgA 蛋白酶的 C 末端最短活性位点位于第 768 位至第 798 位氨基酸之间; 接着进行了第二轮终止突变, 突变位点为 AK183 IgA 蛋白酶的第 774 位、第 779 位、第 783 位、第 788 位或者第 793 位这五个氨基酸位点, 结果如图 2 所示, 第 774 位、第 779 位、第 783 位或者第 788 位的氨基酸终止突变后得到的 AK183(31-773)、AK183(31-778)、AK183(31-782)、AK183(31-787) IgA 蛋白酶截短片段无体外酶切活性, 而第 793 位的氨基酸终止突变后得到的 AK183(31-792) IgA 蛋白酶截短片段仍有活性。因此, 第二轮终止突变结论为: AK183 IgA 蛋白酶的 C 末端最短活性位点位于第 787 位至第 792 位氨基酸之间; 然后, 发明人进行了第三轮终止突变, 突变位点为 AK183 IgA 蛋白酶的第 789 位、第 790 位、第 791 位或者第 792 位这四个氨基酸位点, 结果如图 3 所示, 第 789 位和第 790 位的氨基酸终止突变后得到的 AK183(31-788)、AK183(31-789) IgA 蛋白酶截短片段无体外酶切活性, 而第 791 位或第 792 位的氨基酸终止突变后得到的 AK183(31-790)、AK183(31-791) IgA 蛋白酶截短片段仍有活性 (其中第 791

位可能由于蛋白酶构像问题，活性不完全，仅表现出轻微酶切作用）。因此，第三轮终止突变结论为：AK183 IgA 蛋白酶的 C 末端最短活性片段为 AK183(31-790)。

类似地，发明人进行了三轮截短突变，以研究 AK183 IgA 蛋白酶的 N 末端最短活性位点。首先，发明人进行了第一轮截短突变，在 AK183(31-792)的基础上去掉 AK183 的 N 末端的一段功能未知的结构域（domain of unknown, DUF），C 末端的氨基酸位点固定在第 792 位。例如，在 AK183(31-792)的基础上去掉 N 末端的对应于由 SEQ ID NO: 1 的第 31 位至 284 位氨基酸组成的 DUF，得到 AK183(285-792) IgA 蛋白酶截短片段。采取类似的方法，分别得到 AK183(330-792)、AK183(380-792)、AK183(430-792)、AK183(480-792)、AK183(530-792)、AK183(580-792) IgA 蛋白酶截短片段。得到的 IgA 蛋白酶截短片段对 IgA1 的体外酶切活性实验结果如图 10 所示。如图 10 所示，AK183(285-792)、AK183(330-792) IgA 蛋白酶截短片段仍有体外酶切活性，而 AK183(380-792)、AK183(430-792)、AK183(480-792)、AK183(530-792)、AK183(580-792) IgA 蛋白酶截短片段没有体外酶切活性。因此，第一轮截短突变的结论为：AK183 IgA 蛋白酶的 N 末端最短活性位点位于第 330 位至第 380 位氨基酸之间。接着进行了第二轮截短突变，在第 330 位至第 380 位氨基酸之间每隔 5 个氨基酸进行一个截短体构建，分别得到 AK183(335-792)、AK183(340-792)、AK183(345-792)、AK183(350-792)、AK183(355-792)、AK183(360-792)、AK183(365-792)、AK183(370-792)、AK183(375-792) IgA 蛋白酶截短片段。得到的 IgA 蛋白酶截短片段对 IgA1 的体外酶切活性实验结果如图 11 所示。如图 11 所示，AK183(335-792) IgA 蛋白酶截短片段仍有体外酶切活性，而 AK183(340-792)、AK183(345-792)、AK183(350-792)、AK183(355-792)、AK183(360-792)、AK183(365-792)、AK183(370-792)、AK183(375-792) IgA 蛋白酶截短片段没有体外酶切活性。因此，第二轮截短突变的结论为：AK183 IgA 蛋白酶的 N 末端最短活性位点位于第 335 位至第 340 位氨基酸之间。然后，发明人进行了第三轮截短突变，在第 335 位至第 340 位氨基酸之间逐个氨基酸进行截短体构建，分别得到 AK183(336-792)、AK183(337-792)、AK183(338-792)、AK183(339-792) IgA 蛋白酶截短片段。得到的 IgA 蛋白酶截短片段对 IgA1 的体外酶切活性实验结果如图 12 所示。如图 12 所示，AK183(336-792)、AK183(337-792)、AK183(338-792)、AK183(339-792) IgA 蛋白酶截短片段均没有体外酶切活性。因此，第三轮截短突变的结论为：

AK183 IgA 蛋白酶的 N 末端最短活性位点位于第 335 位氨基酸。最后，发明人对前三轮的结果再次进行了验证，重新同时表达 AK183(285-792)、AK183(330-792)、AK183(335-792)、AK183(336-792)、AK183(337-792)、AK183(338-792)、AK183(339-792)、AK183(340-792)、AK183(345-792)、AK183(350-792) IgA 蛋白酶截短片段。

5 得到的 IgA 蛋白酶截短片段对 IgA1 的体外酶切活性实验结果如图 13 所示。如图 13 所示，AK183(285-792)、AK183(330-792)、AK183(335-792) IgA 蛋白酶截短片段仍有体外酶切活性，而 AK183(336-792)、AK183(337-792)、AK183(338-792)、AK183(339-792)、AK183(340-792)、AK183(345-792)、AK183(350-792) IgA 蛋白酶截短片段没有体外酶切活性，与前述三轮截短突变的结论一致，即，AK183 IgA 蛋白酶

10 蛋白酶的 N 末端最短活性位点位于第 335 位氨基酸。

综上，AK183 IgA 蛋白酶的最短活性片段为 AK183(335-790)。

实施例 2：制备包含 AK183 IgA 蛋白酶截短体或 AK183 IgA 蛋白酶全长的融合蛋白

2.1 质粒构建

15 在确定了 AK183 IgA 蛋白酶的 C 末端最短活性片段 AK183(31-790)后，发明人将 Fc 结构域放在 AK183 IgA 蛋白酶的第 790 位氨基酸的 C 端，中间加 GGGGS 连接，Fc 的 C 端加 6XHis 标签用于蛋白纯化，构建了 PET30a-AK183(31-790)-Fc 质粒，构建流程如图 4 所示。然后，发明人以 PET30a-AK183(31-790)-Fc 质粒为模板，通过 PCR 在 AK183(31-790)截短体后面加了第 791 个及第 792 个氨基酸，构建

20 了 PET30a-AK183(31-792)-Fc 质粒。

申请人同时委托北京六合华大基因科技有限公司构建了 PET30a-AK183(31-798)-Fc、PET30a-AK183(31-807)-Fc、PET30a-AK183(31-816)-Fc、PET30a-AK183(31-833)-Fc 四个备选亚克隆，备选亚克隆的 Fc (CH2-CH3) 除去了铰链区，其氨基酸序列如 SEQ ID NO: 6 所示 (SEQ ID NO: 6 与 SEQ ID NO: 2 相比，缺少了 SEQ ID

25 NO: 2 的前 9 个氨基酸 EPKSCDKTH)，且在 IgA 蛋白酶截短体与 Fc 中间添加了 10 个 His (位于连接子 GGGGS 后面，Fc 前面)。四个备选亚克隆用于后期蛋白酶产量及纯度筛选的备选方案。

为了考察 AK183 IgA 蛋白酶截短片段与 Fc 区的连接方式是否会影响其对 IgA 的酶切活性，发明人还构建了 PET30a-AK183(285-816)-Fc、PET30a-Fc-AK183(285-816)这两个备选亚克隆，其中 Fc 的氨基酸序列如 SEQ ID NO: 25 所示。

5 为了比较 AK183 IgA 蛋白酶截短片段与 Fc 形成的融合蛋白、AK183 IgA 蛋白酶全长与 Fc 形成的融合蛋白对 IgA 的酶切活性，发明人还构建了备选亚克隆 PET30a-Fc-AK183(31-1203)，其中 Fc 的氨基酸序列如 SEQ ID NO: 24 所示。

为了考察 IgG1 Fc、IgG4 Fc 和白蛋白对包含 AK183 IgA 蛋白酶截短体的融合蛋白的 IgA 酶切活性的影响，发明人还构建了 PET30a-AK183(31-816)-IgG4 Fc、PET30a-AK183(31-816)-白蛋白这两个备选亚克隆，其中 IgG4 Fc 的氨基酸序列如
10 SEQ ID NO: 77 所示，白蛋白的氨基酸序列如 SEQ ID NO: 60 所示。

为了考察不同的连接子对包含 AK183 IgA 蛋白酶截短体的融合蛋白的 IgA 酶切活性的影响，发明人还构建了六个备选亚克隆 PET30a-AK183(285-816)-linker-Fc，这六个备选亚克隆表达的融合蛋白序列中，除了连接子不同之外，AK183(285-816)和 Fc 的氨基酸序列均相同，其中 AK183(285-816)的氨基酸序列如 SEQ ID NO: 46
15 所示，Fc 的氨基酸序列如 SEQ ID NO: 25 所示，连接子的氨基酸序列分别为 HHHHHHHHHH (SEQ ID NO: 59，也称为“10xHis”)、EEKKKEKEKEEQEERETK (SEQ ID NO: 58，也称为“IgD linker”)、GGGGS (SEQ ID NO: 22，也称为“1xlinker”)、GGGSGGGGS (SEQ ID NO: 78，也称为“2xlinker”)、GGGSGGGSGGGGS (SEQ ID NO: 79，也称为“3xlinker”)和 GGGSGGGSGGGSGGGGS (SEQ ID NO: 80，
20 也称为“4xlinker”)。

2.2 融合蛋白制备方法

将表达载体转染入大肠杆菌 (BL21-DE3) 感受态细胞，经含 50ug/ml 的卡那霉素的 LB 琼脂培养基抗性选择，然后挑取单克隆菌落至含相应抗生素的 LB 培养液中摇菌至指数生长期 (OD600: 0.6-0.8)，至指数生长期后，加入 0.1-0.5 mM 的异丙基-β-D-硫代半乳糖苷 (IPTG) 诱导，16℃低温诱导表达 24h。完成表达后按照常规方法处理大肠杆菌胞体超声碎裂后高速离心并保留上清，然后采用亲和层析及分子筛纯化得到所述重组融合蛋白。
25

PET30a-AK183(31-792)-Fc 质粒表达的 AK183(31-792)-Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 2 所示, 其编码核酸序列如 SEQ ID NO: 3 所示;

PET30a-AK183(31-798)-Fc 质粒表达的 AK183(31-798)-Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 6 所示, 其编码核酸序列如 SEQ ID NO: 7 所示;

5 PET30a-AK183(31-807)-Fc 质粒表达的 AK183(31-807)-Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 8 所示, 其编码核酸序列如 SEQ ID NO: 9 所示;

PET30a-AK183(31-816)-Fc 质粒表达的 AK183(31-816)-Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 10 所示, 其编码核酸序列如 SEQ ID NO: 11 所示;

10 PET30a-AK183(31-833)-Fc 质粒表达的 AK183(31-833)-Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 12 所示, 其编码核酸序列如 SEQ ID NO: 13 所示;

PET30a-AK183(285-816)-Fc 质粒表达的 AK183(285-816)-Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 81 所示;

PET30a-Fc-AK183(285-816)质粒表达的 Fc-AK183(285-816)融合蛋白的氨基酸序列如 SEQ ID NO: 82 所示;

15 PET30a-Fc-AK183(31-1203)质粒表达的 Fc-AK183(31-1203)融合蛋白的氨基酸序列如 SEQ ID NO: 83 所示;

PET30a-AK183(31-816)-IgG4 Fc 质粒表达的 AK183(31-816)-IgG4 Fc 融合蛋白的氨基酸序列如 SEQ ID NO: 84 所示;

20 PET30a-AK183(31-816)-白蛋白质粒表达的 AK183(31-816)-白蛋白融合蛋白的氨基酸序列如 SEQ ID NO: 85 所示。

2.3 体外活性测试方法

将获得的包含 AK183 IgA 蛋白酶截短体的融合蛋白与从 IgA 肾病患者血浆中提纯的底物 IgA1 在体外混合, 37°C下反应 2-12h, 然后进行 Western blot, 验证其对底物 IgA1 的酶切活性。

25 2.4 体内活性测试方法

将获得的包含 AK183 IgA 蛋白酶截短体的融合蛋白通过尾静脉注射到人源化 IgA1 alpha 链敲入(α 1KI-Tg)C57BL/6 小鼠体内,分别在收集注射前,注射后 5min, 2h, 4h, 24h 的血样, 然后进行 Western blot 验证。

2.5 结果

5 实验表明, PET30a-AK183(31-790)-Fc 质粒成功表达 AK183(31-790)-Fc 融合蛋白(如图 5 所示)。同时, AK183(31-792)-Fc 融合蛋白有预期的全长蛋白表达(如图 6a 所示), 也具有体外针对 IgA1 的酶切活性(如图 6b 所示)。

备选四个亚克隆 PET30a-AK183(31-798)-Fc、PET30a-AK183(31-807)-Fc、PET30a-AK183(31-816)-Fc、PET30a-AK183(31-833)-Fc 均表达融合蛋白, 并且均具有体外针对 IgA1 的酶切活性(如图 7 所示)。

另外, 亚克隆 PET30a-AK183(285-816)-Fc、PET30a-Fc-AK183(285-816)均表达了融合蛋白(如图 14 所示), 并且均具有体外酶切活性(如图 15 所示)。

发明人还验证了亚克隆 PET30a-AK183(31-807)-Fc 表达的 AK183(31-807)-Fc 融合蛋白、亚克隆 PET30a-Fc-AK183(285-816)表达的 Fc-AK183(285-816)融合蛋白的体内活性, 结果如图 8(AK183(31-807)-Fc, 还原条件下)和图 17(Fc-AK183(285-816), 非还原条件下)所示。如图 8 所示, 人源化 IgA1 小鼠(α 1KI-Tg) C57BL/6 在接受单针尾静脉注射 AK183(31-807)-Fc 融合蛋白后, 血液中完整的 IgA1 重链(H)全部消失, 并且持续到至少 24h。如图 17 所示, 人源化 IgA1 小鼠(α 1KI-Tg) C57BL/6 在接受单针尾静脉注射 Fc-AK183(285-816)融合蛋白后, 血液中完整的 IgA1 重链(H)全部消失, 并且持续到至少 2 周。

发明人还比较了 Fc-AK183(285-816)融合蛋白、AK183(285-816)-Fc 融合蛋白、AK183(285-816) IgA 蛋白酶截短体等三者对 IgA1 的酶切活性, 结果如图 16 所示。如图 16 所示, 这三种蛋白对 IgA1 均具有酶切活性。

发明人还比较了 AK183(285-816)-Fc 融合蛋白、Fc-AK183(31-1203)融合蛋白对 IgA1 的酶切活性, 结果如图 18 所示。如图 18 所示, AK183(285-816)-Fc 融合蛋白、Fc-AK183(31-1203)融合蛋白对 IgA1 均具有酶切活性。

发明人还比较了 AK183(31-816)-IgG1 Fc 融合蛋白、AK183(31-816)-IgG4 Fc 融

合蛋白和 AK183(31-816)-白蛋白融合蛋白对 IgA1 的酶切活性，结果如图 19 所示。如图 19 所示，这三种融合蛋白对 IgA1 均具有酶切活性。

发明人还比较了具有不同连接子(10xHis、IgD linker、1xlinker、2xlinker、3xlinker 或 4xlinker) 的 AK183(285-816)-Fc 融合蛋白对 IgA1 的酶切活性，结果如图 20 所示。如图 20 所示，这六种融合蛋白对 IgA1 均具有酶切活性。

2.6 真核表达系统

以上实验均在大肠杆菌 (BL21-DE3) 感受态细胞 (即，原核表达系统) 中进行。接下来，发明人将 AK183(31-792)-Fc 融合 cDNA 序列克隆入 pcDNA3.1/hygro(+) 表达载体中，融合蛋白的 N 端加上
 10 ATGTACAGGATGCAACTCCTGTCTTGCATTGCACTAAGTCTTGCACTTGTCAC
 GAATTCG (SEQ ID NO: 41) 编码表达人 IL-2 的信号肽序列，构建了
 pcDNA3.1/hygro(+)-IL2-AK183(31-792)-Fc 质粒，用于转染真核表达系统 HEK293
 细胞。其中对 Fc 序列进行了针对真核表达系统的密码子优化。pcDNA3.1/hygro(+)-
 IL2-AK183(31-792)-Fc 表达的 IL2-AK183(31-792)-Fc 融合蛋白的氨基酸序列如 SEQ
 15 ID NO: 4 所示，其编码核酸序列如 SEQ ID NO: 5 所示。

AK183(31-792)-Fc 融合蛋白在 HEK293 细胞中的表达结果如图 9 所示。结果表明，AK183(31-792)-Fc 融合蛋白有预期的全长表达，且在真核系统中表达的融合蛋白存在二聚体 (dimer) 形式。

实施例 3: AK183 IgA 蛋白酶突变体的制备及活性测试

20 发明人在 AK183(31-1203) IgA 蛋白酶截短体的基础上，分别对第 844 位、第 862 位、第 931 位和第 933 位、第 978 位、第 1002 位和第 1004 位 (上述位点均为相对于 SEQ ID NO: 1 的位点) 的脯氨酸 (P) 进行定点突变，突变为甘氨酸 (G)，得到 AK183(31-1173) IgA 蛋白酶截短体的五种突变体，其氨基酸序列分别如 SEQ
 ID NO: 53 (也称为“PA-GA Mut”)、SEQ ID NO: 54 (也称为“PI-GI Mut”)、SEQ
 25 ID NO: 55 (也称为“PAP-GAG Mut”)、SEQ ID NO: 56 (也称为“PAT-GAT Mut”) 和 SEQ ID NO: 57 (也称为“PIP-GIG Mut”) 所示。

发明人分别测试了这五种突变体对 IgA1 的酶切活性，结果如图 21 所示。如图

21 所示，这五种突变体对 IgA1 均具有酶切活性。

另外，发明人还在实施例 2 中制备的 AK183(31-816)-Fc 融合蛋白的氨基酸序列（即，SEQ ID NO: 10）的基础上，对其 Fc 区第 7 位（该位点是相对于 SEQ ID NO: 25 而言）的丙氨酸（A）进行定点突变，分别突变为缬氨酸（V）、甘氨酸（G）、
5 丝氨酸（S）、亮氨酸（L），得到 AK183(31-816)-Fc 融合蛋白的四种突变体，分别称为 A-V Mut、A-G Mut、A-S Mut 和 A-L Mut。

发明人分别测试了这四种突变体对 IgA1 的酶切活性，结果如图 22 所示。如图 22 所示，这四种突变体对 IgA1 均具有酶切活性。

实施例 4：探索其他 IgA 蛋白酶

10 发明人从宏基因组数据库里筛选了若干种与 AK183 的野生型 IgA 酶具有一定同源性的氨基酸序列，并合成了 16 种 AK183 同源酶。它们的氨基酸序列分别如 SEQ ID NO: 61 ~ SEQ ID NO: 76 所示。发明人按照实施例 2.3 所述的体外活性测试方法，分别测试了这些 AK183 同源酶对 IgA1 的酶切活性。结果如图 23a 和图 23b 所示。图 23a 中的“1 + IgA1”表示如 SEQ ID NO: 61 所示的多肽与底物 IgA1 在体外
15 混合，“2 + IgA1”表示如 SEQ ID NO: 62 所示的多肽与底物 IgA1 在体外混合，以此类推，图 23b 中的“16 + IgA1”表示如 SEQ ID NO: 76 所示的多肽与底物 IgA1 在体外混合。

由图 23a 和图 23b 可知，如 SEQ ID NO: 61~76 所示的多肽均对 IgA1 具有酶切活性。

20 虽然本申请通过引用特定实施例的方式对发明进行特定的展现和描述，但是本领域技术人员都应该理解，在不脱离本申请披露的主旨和保护范围的情况下，上述内容还可以进行各种形式和细节上的变化。

权 利 要 求 书

1. 一种分离的 IgA 蛋白酶截短体，其包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的非天然截短片段，或者与所述非天然截短片段具有至少 70% 的序列同一性。
2. 如权利要求 1 所述的 IgA 蛋白酶截短体，其中所述非天然截短片段在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的基础上有氨基酸取代、缺失、插入或修饰，使得所述 IgA 蛋白酶截短体丧失或降低自酶切功能。
3. 如权利要求 2 所述的 IgA 蛋白酶截短体，其中所述氨基酸取代、缺失、插入或修饰发生在所述多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的天然自酶切位点、所述天然自酶切位点的上游 5 个位点以内和/或下游 5 个位点以内。
4. 如权利要求 1-3 中任一项所述的 IgA 蛋白酶截短体，其中所述非天然截短片段为获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的 N 末端截短片段或 C 末端截短片段。
5. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体，其中所述多枝梭菌（*Clostridium ramosum*）为 *Clostridium ramosum* AK183 株。
6. 如权利要求 4 或 5 所述的 IgA 蛋白酶截短体，其中所述 N 末端截短片段包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的 N 末端第 31 位起的至少 760 个连续氨基酸的多肽片段，或者与所述多肽片段具有至少 70% 的序列同一性。
7. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体，其中所述非天然截短片段包含获得自或衍生自多枝梭菌（*Clostridium ramosum*）的野生型 IgA 蛋白酶的 N 末端第 335 位起的至少 456 个连续氨基酸的多肽片段，或者与所述多肽片段具有至少 90% 或至少 95% 的序列同一性。

8. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体, 其中所述多枝梭菌 (*Clostridium ramosum*) 的野生型 IgA 蛋白酶的氨基酸序列如 SEQ ID NO: 1 所示。
9. 如权利要求 3 至 8 中任一项所述的 IgA 蛋白酶截短体, 其中所述天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 730 位至第 840 位之间 (例如, 第 792 位至第 797 位之间)。
10. 如权利要求 9 所述的 IgA 蛋白酶截短体, 其中所述天然自酶切位点在如 SEQ ID NO: 1 所示的氨基酸序列的第 790 位、第 791 位、第 792 位、第 793 位、第 794 位、第 795 位、第 796 位、第 797 位、第 798 位、第 799 位或第 800 位。
11. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体, 其包含如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位起的至少 760 个 (例如, 至少 761 个、至少 762 个、至少 763 个、至少 764 个、至少 765 个、至少 766 个、至少 767 个、至少 768 个、至少 769 个、至少 770 个、至少 771 个、至少 772 个、至少 773 个、至少 774 个、至少 775 个、至少 776 个、至少 777 个、至少 778 个、至少 779 个、至少 780 个、至少 781 个、至少 782 个、至少 783 个、至少 784 个、至少 785 个、至少 786 个、至少 787 个、至少 788 个、至少 789 个、至少 790 个、至少 791 个、至少 792 个、至少 793 个、至少 794 个、至少 795 个、至少 796 个、至少 797 个、至少 798 个、至少 799 个、至少 800 个、至少 801 个、至少 802 个、至少 803 个、至少 804 个、至少 805 个、至少 806 个、至少 807 个、至少 808 个、至少 809 个、至少 810 个、至少 900 个、至少 950 个、至少 1000 个、至少 1100 个、至少 1150 个或者至少 1200 个) 连续氨基酸的多肽片段。
12. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体, 其包含选自下组的多肽片段: 如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 798 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 807 位氨基酸、如 SEQ ID NO:

1 所示的氨基酸序列的第 31 位至第 816 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 31 位至第 833 位氨基酸以及与其具有至少 70% 的序列同一性的多肽片段。

13. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体, 其包含如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位起的至少 456 个 (例如, 至少 457 个、至少 458 个、至少 459 个、至少 460 个、至少 461 个、至少 462 个、至少 463 个、至少 464 个、至少 465 个、至少 466 个、至少 467 个、至少 468 个、至少 469 个、至少 470 个、至少 471 个、至少 472 个、至少 473 个、至少 474 个、至少 475 个、至少 476 个、至少 477 个、至少 478 个、至少 479 个、至少 480 个、至少 481 个、至少 482 个、至少 483 个、至少 484 个、至少 485 个、至少 486 个、至少 487 个、至少 488 个、至少 489 个、至少 490 个、至少 491 个、至少 492 个、至少 493 个、至少 494 个、至少 495 个、至少 496 个、至少 497 个、至少 498 个、至少 499 个、至少 500 个、至少 550 个、至少 600 个、至少 650 个、至少 700 个、至少 750 个、至少 800 个、至少 850 个或者至少 900 个) 连续氨基酸的多肽片段。
14. 如权利要求 13 所述的 IgA 蛋白酶截短体, 其包含选自下组的多肽片段: 如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 335 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 790 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 791 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 330 位至第 792 位氨基酸、如 SEQ ID NO: 1 所示的氨基酸序列的第 285 位至第 816 位氨基酸以及与其具有至少 90% 或至少 95% 的序列同一性的多肽片段。

15. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体，其在所述多肽片段的氨基酸序列基础上，在一个或多个位点具有氨基酸的保守替换。
16. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体，其中所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置有氨基酸突变。
17. 如权利要求 16 所述的 IgA 蛋白酶截短体，其中所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位、第 862 位、第 931 位、第 933 位、第 978 位、第 1002 位、第 1004 位中的一个或多个位置突变为甘氨酸。
18. 如权利要求 16 或 17 所述的 IgA 蛋白酶截短体，其中所述多肽片段在对应于 SEQ ID NO: 1 的第 844 位有氨基酸突变、第 862 位有氨基酸突变、第 931 位和第 933 位有氨基酸突变、第 978 位有氨基酸突变、或者第 1002 位和第 1004 位有氨基酸突变。
19. 如权利要求 16 至 18 中任一项所述的 IgA 蛋白酶截短体，其中所述多肽片段的氨基酸序列如 SEQ ID NO: 53（也称为“PA-GA Mut”）、SEQ ID NO: 54（也称为“PI-GI Mut”）、SEQ ID NO: 55（也称为“PAP-GAG Mut”）、SEQ ID NO: 56（也称为“PAT-GAT Mut”）或者 SEQ ID NO: 57（也称为“PIP-GIG Mut”）所示。
20. 如前述权利要求中任一项所述的 IgA 蛋白酶截短体，其具有特异性切割人 IgA 的酶活性。
21. 如权利要求 20 所述的 IgA 蛋白酶截短体，其具有特异性切割人 IgA 重链的酶活性。
22. 如权利要求 21 所述的 IgA 蛋白酶截短体，其具有特异性切割人 IgA 重链 CH1 与铰链区交汇处的酶活性。
23. 如权利要求 20 至 22 中任一项所述的 IgA 蛋白酶截短体，其具有特异性切割人 IgA1 的酶活性。

24. 一种融合蛋白，其包含第一多肽和第二多肽，其中所述第一多肽包含获得自或衍生自多枝梭菌(*Clostridium ramosum*)的野生型 IgA 蛋白酶的全长、获得自或衍生自多枝梭菌(*Clostridium ramosum*)的野生型 IgA 蛋白酶去掉信号肽之后形成的多肽或者如前述权利要求中任一项所述的 IgA 蛋白酶截短体，所述第二多肽包含用于延长所述第一多肽在受试者体内半衰期的氨基酸序列。
25. 根据权利要求 24 所述的融合蛋白，其中所述第一多肽包含如 SEQ ID NO: 1 或者 SEQ ID NO: 42 所示的序列。
26. 如权利要求 24 所述的融合蛋白，其中所述第二多肽位于所述第一多肽的 N 末端或者 C 末端。
27. 如权利要求 24 至 26 中任一项所述的融合蛋白，其中所述第一多肽和所述第二多肽之间通过连接子连接。
28. 如权利要求 24 至 26 中任一项所述的融合蛋白，其中所述第一多肽和所述第二多肽之间直接连接。
29. 如权利要求 27 所述的融合蛋白，其中所述连接子选自下组：可切割连接子、不可切割连接子、肽连接子、柔性连接子、刚性连接子、螺旋连接子和非螺旋连接子。
30. 如权利要求 29 所述的融合蛋白，其中所述连接子包括肽连接子。
31. 如权利要求 30 所述的融合蛋白，所述肽连接子包括含有甘氨酸和丝氨酸的连接子。
32. 如权利要求 31 所述的融合蛋白，其中所述含有甘氨酸和丝氨酸的连接子包括如 SEQ ID NO: 21 (GGGS)、SEQ ID NO: 22 (GGGGS)、SEQ ID NO: 86 (GGGGGS)或 SEQ ID NO: 87 (GGGGGGGS)所示的一个、两个、三个、四个或更多个重复。
33. 如权利要求 30 所述的融合蛋白，其中所述连接子包括如 SEQ ID NO: 23 (GGCGGCGGTGGATCC) 所示、如 SEQ ID NO: 58

(EEKKKEKEKEEQEERETK) 所示或者如 SEQ ID NO: 59 (HHHHHHHHHH) 所示的氨基酸序列。

34. 如权利要求 24 至 33 中任一项所述的融合蛋白，其中所述第二多肽选自：Fc 结构域和白蛋白。

35. 如权利要求 34 所述的融合蛋白，其中所述 Fc 结构域包含铰链区。

36. 如权利要求 35 所述的融合蛋白，其中所述 Fc 结构域来源于人 IgG Fc 结构域。

37. 如权利要求 36 所述的融合蛋白，其中所述 Fc 结构域来源于人 IgG1 Fc 结构域、人 IgG2 Fc 结构域、人 IgG3 Fc 结构域或人 IgG4 Fc 结构域。

38. 如权利要求 34 至 37 中任一项所述的融合蛋白，其中所述 Fc 结构域包含与 SEQ ID NO: 24、SEQ ID NO: 25、SEQ ID NO: 32 或 SEQ ID NO: 77 具有至少 70%、至少 75%、至少 80%、至少 85%、至少 90%或至少 95%的序列同一性的氨基酸序列。

39. 如权利要求 38 所述的融合蛋白，其中所述 Fc 结构域包含如 SEQ ID NO: 24、SEQ ID NO: 25、SEQ ID NO: 32 或 SEQ ID NO: 77 所示的氨基酸序列。

40. 如权利要求 39 所述的融合蛋白，其中所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位有氨基酸突变。

41. 如权利要求 40 所述的融合蛋白，其中所述 Fc 结构域在对应于 SEQ ID NO: 25 的第 7 位的氨基酸（例如，丙氨酸）突变为缬氨酸、甘氨酸、丝氨酸或者亮氨酸。

42. 如权利要求 34 至 41 中任一项所述的融合蛋白，其中所述 Fc 结构域包含一个或多个延长所述融合蛋白的半衰期的突变。

43. 如权利要求 34 至 42 中任一项所述的融合蛋白，其中所述 Fc 结构域与所述第一多肽的 C 末端或者 N 末端连接。

44. 如权利要求 34 所述的融合蛋白，其中所述白蛋白包含人血清白蛋白的一个或多个结构域。
45. 如权利要求 44 所述的融合蛋白，其中所述白蛋白包含人血清白蛋白的 D3 结构域。
46. 如权利要求 24 至 45 中任一项所述的融合蛋白，其进一步包含标签。
47. 如权利要求 46 所述的融合蛋白，其中所述标签选自下组：荧光标签、发光标签、纯化标签和生色标签。
48. 如权利要求 46 或 47 所述的融合蛋白，其中所述标签选自下组：c-Myc 标签、HA 标签、VSV-G 标签、FLAG 标签、V5 标签和 HIS 标签。
49. 如权利要求 48 所述的融合蛋白，其中所述标签是包含 6 个、7 个、8 个、9 个或 10 个组氨酸的 HIS 标签。
50. 如权利要求 46 至 49 中任一项所述的融合蛋白，其中所述第二多肽位于所述第一多肽的 C 末端，所述标签位于所述第二多肽的 C 末端。
51. 如权利要求 46 至 50 中任一项所述的融合蛋白，其中所述融合蛋白在受试者体内的血循环中的半衰期为至少 1 天、至少 2 天、至少 3 天、至少 4 天、至少 5 天、至少 6 天、至少 7 天、至少 8 天、至少 9 天、至少 10 天、至少 11 天、至少 12 天、至少 13 天、至少 14 天。
52. 一种分离的核酸，其包含编码如权利要求 1 至 23 中任一项所述的 IgA 蛋白酶截短体的核苷酸序列或包含编码如权利要求 24 至 51 中任一项所述的融合蛋白的核苷酸序列。
53. 如权利要求 52 所述的核酸，其包含选自下组的核苷酸序列：SEQ ID NO: 3、SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 33、SEQ ID NO: 34、SEQ ID NO: 35、SEQ ID NO: 36、SEQ ID NO: 37、SEQ ID NO: 38 以及与其具有至少 70% 的序列同一性的核苷酸序列。
54. 一种载体，其包含如权利要求 52 或 53 所述的核酸。

55. 一种细胞，其包含如权利要求 52 或 53 所述的核酸或如权利要求 54 所述的载体。
56. 如权利要求 55 所述的细胞，其中所述细胞是原核生物细胞或真核生物细胞。
57. 如权利要求 56 所述的细胞，其中所述原核生物细胞是大肠杆菌细胞。
58. 如权利要求 56 所述的细胞，其中所述真核生物细胞是哺乳动物细胞。
59. 如权利要求 58 所述的细胞，其中所述哺乳动物细胞是人细胞或中国仓鼠卵巢（CHO）细胞。
60. 如权利要求 58 所述的细胞，其中所述哺乳动物细胞是人胚胎肾细胞 293（HEK293 细胞）。
61. 一种药物组合物，其包含如权利要求 1 至 23 中任一项所述的 IgA 蛋白酶截短体、包含如权利要求 24 至 51 中任一项所述的融合蛋白、包含如权利要求 52 或 53 所述的核酸、包含如权利要求 54 所述的载体或者包含如权利要求 55 至 60 中任一项所述的细胞，以及药学上可接受的载体。
62. 一种产生融合蛋白的方法，其包括培养如权利要求 55 至 60 中任一项所述的细胞的步骤。
63. 一种治疗或预防 IgA 沉积相关疾病的方法，其包括向需要治疗或预防的受试者施用如权利要求 1 至 23 中任一项所述的 IgA 蛋白酶截短体、如权利要求 24 至 51 中任一项所述的融合蛋白或者如权利要求 61 所述的药物组合物。
64. 一种治疗或预防 IgA 沉积相关疾病的方法，其包括向需要治疗或预防的受试者施用 IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物组合物，其中，所述 IgA 蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO:

70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75、SEQ ID NO: 76 或其组合。

65. 如权利要求 1 至 23 中任一项所述的 IgA 蛋白酶截短体、如权利要求 24 至 51 中任一项所述的融合蛋白或者如权利要求 61 所述的药物组合物在制备用于治疗或预防 IgA 沉积相关疾病的药物中的用途。

66. IgA 蛋白酶或其截短体、包含所述 IgA 蛋白酶或其截短体的融合蛋白或者包含所述 IgA 蛋白酶或其截短体或者所述融合蛋白的药物组合物在制备用于治疗或预防 IgA 沉积相关疾病的药物中的用途，其中，所述 IgA 蛋白酶的氨基酸序列选自下组：SEQ ID NO: 61、SEQ ID NO: 62、SEQ ID NO: 63、SEQ ID NO: 64、SEQ ID NO: 65、SEQ ID NO: 66、SEQ ID NO: 67、SEQ ID NO: 68、SEQ ID NO: 69、SEQ ID NO: 70、SEQ ID NO: 71、SEQ ID NO: 72、SEQ ID NO: 73、SEQ ID NO: 74、SEQ ID NO: 75、SEQ ID NO: 76 或其组合。

67. 如权利要求 63 或 64 所述的方法或者如权利要求 65 或 66 所述的用途，其中所述 IgA 沉积相关疾病包括 IgA 肾病、疱疹样皮炎、亨诺赫-舍恩莱因紫癜（又称 IgA 血管炎）、川崎病、紫癜性肾炎、IgA 血管炎肾损害、IgA 类风湿因子阳性的类风湿性关节炎、IgA 型抗 GBM 病或 IgA 型 ANCA 相关血管炎。

68. 如权利要求 63 或 64 所述的方法或者如权利要求 65 或 66 所述的用途，其中所述 IgA 沉积相关疾病为 IgA 肾病、IgA 血管炎或川崎病。