



(12) 发明专利申请

(10) 申请公布号 CN 113383008 A

(43) 申请公布日 2021.09.10

(21) 申请号 201980069663.X

(22) 申请日 2019.10.18

(30) 优先权数据

62/752,546 2018.10.30 US

(85) PCT国际申请进入国家阶段日

2021.04.22

(86) PCT国际申请的申请数据

PCT/US2019/056872 2019.10.18

(87) PCT国际申请的公布数据

W02020/092017 EN 2020.05.07

(71) 申请人 先正达参股股份有限公司

地址 瑞士巴塞尔

(72) 发明人 C·M·雷诺

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

代理人 张小勇

(51) Int.Cl.

C07K 14/32 (2006.01)

C07K 14/325 (2006.01)

C07K 14/24 (2006.01)

C12N 15/63 (2006.01)

C12N 15/70 (2006.01)

C12N 15/75 (2006.01)

权利要求书4页 说明书47页
序列表66页

(54) 发明名称

杀昆虫蛋白

(57) 摘要

披露了用于控制昆虫有害生物的组合物和方法。特别是,提供了对至少鞘翅目昆虫有害生物具有毒性的新颖杀昆虫蛋白。还提供了编码所述新颖杀昆虫蛋白的核酸分子。还披露了制备所述杀昆虫蛋白的方法和使用所述杀昆虫蛋白以及编码本发明的杀昆虫蛋白的核酸的方法,例如,在转基因植物中以赋予免于昆虫损害的保护。

1. 一种核酸分子,所述核酸分子包含编码对昆虫有害生物有毒的蛋白质的核苷酸序列,其中所述核苷酸序列(a) 编码蛋白质,所述蛋白质包含与SEQ ID NO:1-22中的任一个或其毒素片段具有至少80%到至少99%序列同一性的氨基酸序列;(b) 与SEQ ID NO:23-54中的任一个或其毒素编码片段具有至少80%到至少99%序列同一性;或者(c) 是(a) 或(b) 的合成序列,所述合成序列已经进行密码子优化用于在转基因生物中表达。

2. 如权利要求1所述的核酸分子,其中所述杀昆虫蛋白包含SEQ ID NO:1-22中的任一个或其毒性片段的氨基酸序列。

3. 如权利要求1所述的核酸分子,其中所述核苷酸序列包含SEQ ID NO:23-54中的任一个或其毒素编码片段。

4. 如权利要求1所述的核酸分子,其中所述合成核苷酸序列包含SEQ ID NO:26-54中的任一个或其毒素编码片段。

5. 一种嵌合基因,所述嵌合基因包含可操作地连接至如权利要求1-4中任一项所述的核酸分子的异源启动子。

6. 如权利要求5所述的嵌合基因,其中所述异源启动子是植物可表达型启动子。

7. 如权利要求6所述的嵌合基因,其中所述植物可表达型启动子选自下组启动子,该组由以下组成:泛素、夜香树属黄病毒、玉米TrpA、OsMADS 6、玉蜀黍H3组蛋白、噬菌体T3基因9 5'UTR、玉米蔗糖合成酶1、玉米醇脱氢酶1、玉米捕光复合物、玉米热休克蛋白、玉蜀黍mt1、豌豆小亚基RuBP羧化酶、稻肌动蛋白、稻亲环蛋白、Ti质粒甘露碱合酶、Ti质粒胭脂碱合酶、矮牵牛查尔酮异构酶、豆类富甘氨酸蛋白1、马铃薯糖蛋白、凝集素、CaMV 35S以及S-E9小亚基RuBP羧化酶启动子。

8. 如权利要求5所述的嵌合基因,其中所述昆虫有害生物是鞘翅目昆虫有害生物。

9. 如权利要求8所述的嵌合基因,其中所述鞘翅目昆虫有害生物是根萤叶甲属昆虫有害生物。

10. 如权利要求9所述的嵌合基因,其中所述根萤叶甲属昆虫有害生物选自由以下组成的组:玉米根萤叶甲、巴氏根萤叶甲、黄瓜十一星叶甲食根亚种和玉蜀黍根萤叶甲。

11. 如权利要求5所述的嵌合基因,其中所述转基因生物是细菌或植物。

12. 一种对昆虫有害生物有毒的分离的蛋白质,其中所述蛋白质包含(a) 与SEQ ID NO:1-22中的任一个或其毒素片段的氨基酸序列具有至少80%到至少99%的序列同一性的氨基酸序列;(b) 包含SEQ ID NO:1-22中的任一个或其毒素片段的氨基酸序列;(c) 由与SEQ ID NO:23-54中的任一个或其毒素编码片段的核苷酸序列具有至少80%到至少99%序列同一性的核苷酸序列编码的氨基酸序列;或(d) 由包含SEQ ID NO:22-54中的任一个或其毒素编码片段的核苷酸序列编码的氨基酸序列。

13. 如权利要求12所述的分离的蛋白质,其中所述昆虫有害生物是鞘翅目昆虫有害生物。

14. 如权利要求13所述的分离的蛋白质,其中所述鞘翅目昆虫有害生物是根萤叶甲属昆虫有害生物。

15. 如权利要求14所述的分离的蛋白质,其中所述根萤叶甲属昆虫有害生物选自由以下组成的组:玉米根萤叶甲、巴氏根萤叶甲、黄瓜十一星叶甲食根亚种和玉蜀黍根萤叶甲。

16. 一种重组载体,所述重组载体包含如权利要求5所述的嵌合基因。

17. 一种宿主细胞, 所述宿主细胞包含如权利要求16所述的重组载体, 其中所述宿主细胞是细菌细胞或植物细胞。

18. 如权利要求17所述的转基因细菌细胞, 其中所述细菌细胞在芽孢杆菌属、梭菌属、致病杆菌属、发光杆菌属、巴斯德氏芽菌属、埃希氏菌属、假单胞菌属、欧文氏菌属、沙雷氏菌属、克雷伯菌属、沙门氏菌属、巴氏杆菌属、黄单胞菌属、链霉菌属、根瘤菌属、红假单胞菌属、嗜甲基菌属、农杆菌属、醋杆菌属、乳杆菌属、节杆菌属、固氮菌属、明串珠菌属或产碱杆菌属中。

19. 如权利要求18所述的转基因芽孢杆菌属细胞, 其中所述芽孢杆菌属细胞是苏云金芽孢杆菌细胞。

20. 如权利要求17所述的转基因植物细胞, 其中所述植物细胞是双子叶植物细胞或单子叶植物细胞。

21. 如权利要求20所述的转基因植物细胞, 其中 (a) 所述双子叶植物细胞选自由以下组成的组: 大豆细胞、向日葵细胞、番茄细胞、芸苔属作物细胞、棉花细胞、甜菜细胞以及烟草细胞; 或 (b) 所述单子叶植物细胞选自由以下组成的组: 大麦细胞、玉蜀黍细胞、燕麦细胞、稻细胞、高粱细胞、甘蔗细胞以及小麦细胞。

22. 一种转基因植物或植物部分, 所述转基因植物或植物部分包含如权利要求21所述的转基因植物细胞。

23. 如权利要求22所述的转基因植物或植物部分, 所述转基因植物或植物部分是转基因玉蜀黍植物或植物部分。

24. 一种工程化杀昆虫蛋白, 所述工程化杀昆虫蛋白包含与SEQ ID NO:1-3中的任一个具有至少80%到至少90%序列同一性的氨基酸序列并且进一步在对应于以下的位置处包含至少一个突变: (a) SEQ ID NO:1的氨基酸位置1-489; 或 (b) SEQ ID NO:2的氨基酸位置1-488; 或 (c) SEQ ID NO:3的氨基酸位置1-489。

25. 如权利要求24所述的工程化杀昆虫蛋白, 其中所述突变位于与SEQ ID NO:1的氨基酸位置8、11、188、382、397、398、413、428、430或482或其任意组合对应的氨基酸位置处。

26. 如权利要求25所述的工程化杀昆虫蛋白, 其中所述突变位于SEQ ID NO:1的位置8、11、188、382、397、398、413、428、430或482处。

27. 如权利要求25所述的工程化杀昆虫蛋白, 其中所述突变位于与SEQ ID NO:1的氨基酸8和11或396和397对应的位置处。

28. 如权利要求27所述的工程化杀昆虫蛋白, 其中所述突变位于SEQ ID NO:1的氨基酸位置8和11或396和397处。

29. 如权利要求26所述的工程化杀昆虫蛋白, 其中位置8的突变是V8A, 位置11的突变是L11I, 位置188的突变是Y188F, 位置382的突变是C382A、C382S或C382T, 位置397的突变是I397L, 位置398的突变是V398L, 位置413的突变是Y413W, 位置428的突变是F428W, 位置430的突变是Y430W, 或者位置482的突变是C482L。

30. 如权利要求27所述的工程化杀昆虫蛋白, 其中所述蛋白质包含SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:17的氨基酸序列。

31. 如权利要求24-31中任一项所述的工程化杀昆虫蛋白, 其中所述蛋白质对至少鞘翅

目昆虫有害生物具有活性。

32. 如权利要求31所述的工程化杀昆虫蛋白,其中所述鞘翅目昆虫有害生物是根萤叶甲属昆虫有害生物。

33. 如权利要求32所述的工程化杀昆虫蛋白,其中所述根萤叶甲属昆虫有害生物选自由以下组成的组:玉米根萤叶甲、巴氏根萤叶甲、黄瓜十一星叶甲食根亚种和玉蜀黍根萤叶甲。

34. 一种杀昆虫组合物,所述杀昆虫组合物包含如权利要求12所述的蛋白质和农业上可接受的载剂。

35. 如权利要求34所述的组合物,其中所述农业上可接受的载剂选自由以下组成的组:粉剂、尘剂、丸剂、颗粒剂、喷雾剂、乳剂、胶体以及溶液。

36. 如权利要求34所述的组合物,其中所述组合物通过脱水、冷冻干燥、均化、提取、过滤、离心、沉降或浓缩苏云金芽孢杆菌菌株的培养物而制备。

37. 如权利要求34所述的组合物,所述组合物包含按重量计从约1%至约99%的所述杀昆虫蛋白。

38. 如权利要求34-37中任一项所述的组合物,其中所述组合物还包含第二杀有害生物剂。

39. 如权利要求38所述的组合物,其中所述第二杀有害生物剂是生物剂或化学剂。

40. 如权利要求39所述的组合物,其中 (a) 所述生物剂是或衍生自苏云金芽孢杆菌杀昆虫蛋白、蜡样芽孢杆菌杀昆虫蛋白、致病杆菌属杀昆虫蛋白、光杆菌属杀昆虫蛋白、侧孢短芽孢杆菌杀昆虫蛋白、球形赖氨酸芽孢杆菌杀昆虫蛋白、色杆菌属杀昆虫蛋白、噬虫霉耶尔森菌杀昆虫蛋白、乳状芽孢杆菌杀昆虫蛋白、或梭菌属杀昆虫蛋白; (b) 所述生物剂是或衍生自dsRNA、Cry蛋白、Vip蛋白、马铃薯糖蛋白、蛋白酶、蛋白酶抑制剂、脲酶、 α -淀粉酶抑制剂、成孔蛋白、凝集素、工程化抗体或抗体片段、或几丁质酶; (c) 所述化学剂是氨基甲酸酯、拟除虫菊酯、有机磷酸酯、fiprole、新烟碱、有机氯化物、沙蚕毒素或其组合; 或者 (d) 所述化学剂包含选自下组的活性成分,该组由以下组成:克百威、胺甲萘、灭多虫、联苯菊酯、七氟菊酯、氯菊酯、氟氯氰菊酯、 λ -氯氟氰菊酯、氯氰菊酯、溴氰菊酯、毒死蜱、氯氧磷、乐果、灭线磷、马拉硫磷、甲基对硫磷、甲拌磷、特丁磷、叔丁嘧啶磷、氟虫腈、啉虫脒、吡虫啉、噻虫啉、噻虫嗪、硫丹、杀虫磺及其组合。

41. 一种用于产生杀昆虫蛋白的方法,所述方法包括在如权利要求17所述的宿主细胞产生所述杀昆虫蛋白的条件下培养所述宿主细胞或包含所述宿主细胞的生物。

42. 一种产生与对照植物或植物部分相比具有增强的昆虫抗性的转基因植物或植物部分的方法,所述方法包括: (a) 向植物或植物部分中引入如权利要求5所述的嵌合基因,其中所述杀昆虫蛋白在所述植物或植物部分中表达,从而产生具有增强的昆虫抗性的植物或植物部分。

43. 如权利要求42所述的方法,其中所述引入步骤通过以下实现: (a) 转化所述植物或植物部分; 或者 (b) 使包含所述嵌合基因的第一植物与不同的第二植物杂交。

44. 如权利要求43所述的方法,其中所述嵌合基因编码具有SEQ ID NO:4的氨基酸序列的杀昆虫蛋白。

45. 一种控制昆虫有害生物的方法,所述方法包括将有效量的如权利要求12所述的杀

昆虫蛋白递送至所述昆虫有害生物或其环境。

46. 如权利要求45所述的方法, 其中所述杀昆虫蛋白通过转基因植物或通过局部施用包含所述杀昆虫蛋白的杀昆虫组合物递送。

47. 如权利要求46所述的方法, 其中所述转基因植物或所述杀昆虫组合物包含不同于所述杀昆虫蛋白的第二杀昆虫剂。

48. 如权利要求47所述的方法, 其中所述第二杀昆虫剂是蛋白质、dsRNA或化学品。

49. 如权利要求48所述的方法, 其中 (a) 所述蛋白质选自由以下组成的组: Cry蛋白、Vip蛋白、马铃薯糖蛋白、蛋白酶、蛋白酶抑制剂、脲酶、 α -淀粉酶抑制剂、成孔蛋白、凝集素、工程化抗体或抗体片段、或几丁质酶; (b) 所述化学品是氨基甲酸酯、拟除虫菊酯、有机磷酸酯、fiprole、新烟碱、有机氯化物、沙蚕毒素或其组合; 或 (c) 所述化学品包含选自下组的活性成分, 该组由以下组成: 克百威、胺甲萘、灭多虫、联苯菊酯、七氟菊酯、氯菊酯、氟氯氰菊酯、 λ -氯氟氰菊酯、氯氰菊酯、溴氰菊酯、毒死蜱、氯氧磷、乐果、灭线磷、马拉硫磷、甲基对硫磷、甲拌磷、特丁磷、叔丁嘧啶磷、氟虫腈、啉虫脒、吡虫啉、噻虫啉、噻虫嗪、硫丹、杀虫磺及其组合。

50. 如权利要求45-49中任一项所述的方法, 其中所述昆虫有害生物是鞘翅目昆虫有害生物。

51. 如权利要求50所述的方法, 其中所述鞘翅目昆虫有害生物是根萤叶甲属物种。

52. 如权利要求51所述的方法, 其中所述根萤叶甲属物种选自由以下组成的组: 玉米根萤叶甲、巴氏根萤叶甲、黄瓜十一星叶甲食根亚种和玉蜀黍根萤叶甲。

53. 一种降低根萤叶甲属昆虫群体对如权利要求5所述的杀昆虫蛋白的抗性发展的方法, 所述方法包括在由所述根萤叶甲属昆虫群体摄食的转基因植物中表达所述杀昆虫蛋白和干扰RNA分子, 所述干扰RNA分子抑制幼虫和成虫根萤叶甲属昆虫中靶基因的表达, 从而使所述根萤叶甲属昆虫群体中的抗性发展与仅暴露于所述杀昆虫蛋白的根萤叶甲属昆虫群体相比降低。

54. 一种向玉米种植者提供控制玉米作物中根萤叶甲属昆虫有害生物群体的手段的方法, 所述方法包括 (a) 向所述种植者出售或提供包含如权利要求1所述的核酸分子的转基因玉米种子; 和 (b) 向所述种植者宣传所述转基因玉米种子产生控制根萤叶甲属有害生物群体的转基因玉米植物。

55. 一种鉴定杀昆虫蛋白的方法, 所述杀昆虫蛋白包含与SEQ ID NO:1-3中的任一个具有至少80%到至少99%序列同一性的核苷酸序列, 所述方法包括以下步骤: (a) 产生引物对, 所述引物对将从核酸样品或其互补序列扩增选自组SEQ ID NO:1-3的多核苷酸, (b) 从所述核酸样品中扩增直系同源基因, (c) 鉴定直系同源基因的多核苷酸序列, (d) 产生由所述直系同源基因编码的蛋白质, 以及 (e) 确定步骤 (d) 的蛋白质针对昆虫有害生物具有杀昆虫活性。

杀昆虫蛋白

技术领域

[0001] 本发明涉及蛋白工程、植物分子生物学以及有害生物控制的领域。更具体地,本发明涉及具有杀昆虫活性的新颖蛋白质及其变体,核酸(其表达产生杀昆虫蛋白),以及制造这些杀昆虫蛋白和相应的核酸的方法和使用这些杀昆虫蛋白和相应的核酸来控制昆虫的方法。

背景技术

[0002] 昆虫有害生物是引起作物损失的一个主要原因。仅在美国,由于各个属的昆虫的侵染每年就损失数十亿美元。除了大田作物的损失之外,昆虫有害生物对于菜农和果农、对于观赏性花卉的生产商而言也是一种负担,并且对于园丁和房屋所有者而言是一种令人讨厌的东西。

[0003] 玉米根虫的多个物种被认为是最具破坏性的玉米有害生物。单在美国,三个物种,玉米根萤叶甲(*Diabrotica virgifera virgifera*, 又称西方玉米根虫)、长角巴氏根萤叶甲(*D. longicornis barberi*, 又称北方玉米根虫)、以及黄瓜十一星叶甲食根亚种(*D. undecimpunctata howardi*, 又称南方玉米根虫),在美国玉米带中,每年对玉米造成的损失超过十亿美元。美国南部的一种重要的玉米根虫有害生物是墨西哥玉米根虫(墨西哥玉米根萤叶甲(*Diabrotica virgifera zea*))。在南美洲,南美叶甲(*Diabrotica speciosa*)被认为是玉米的重要有害生物。西方玉米根虫于1992年传播到欧洲,自2008年以来一直在整个主要玉米种植地区造成经济损失。玉米根虫幼虫通过几乎专一摄食玉米根而造成最为实质性的植物损害。已显示这种损伤增加了植物倒伏、减少了谷物产量以及营养产量连同改变了谷物的营养物含量。幼虫摄食还通过为导致根腐病和茎腐病的细菌和真菌感染打开了进入根部的途径,而对玉米造成了间接的影响。成年玉米根虫在晚夏时活跃在玉米地中,在此它们摄食穗、穗丝以及花粉,由此干扰了正常的授粉。

[0004] 玉米根虫主要是通过密集施用化学杀有害生物剂而得到控制的,这些化学杀有害生物剂通过抑制昆虫生长、预防昆虫摄食或繁殖、或者导致死亡而具有活性。由此可以达到良好的玉米根虫控制,但这些化学品有时也能影响其他有益的生物。由化学杀有害生物剂的广泛使用产生的另一个问题是出现了抗性昆虫品种。又另一个问题是由于下述事实:玉米根虫幼虫在地下摄食因此使得施用杀昆虫剂的救护处理变得困难。因此,大多数杀昆虫剂的施用是在种植时预防性地进行的。这种实践导致了巨大的环境负担。通过各种农田管理实践已部分地改善了这种状况,但对替代性有害生物控制的机制存在着不断增加的需求。

[0005] 生物性有害生物控制剂,如表达杀有害生物毒素像 δ -内毒素(Δ -内毒素;也被称为结晶毒素或Cry蛋白)的苏云金芽孢杆菌(Bt)菌株,也已经施用至作物植物中,产生了对昆虫有害生物的令人满意的结果。这些 δ -内毒素是在结晶基质内所容纳的蛋白质,已知它们当被某些昆虫摄取时拥有杀昆虫活性。来自苏云金芽孢杆菌的若干天然Cry蛋白或工程化的Cry蛋白已经在转基因作物植物中表达并且在商业上被开发以控制某些鳞翅目和鞘

翅目昆虫有害生物。例如,起始于2003年,通过表达Cry3Bb1、Cry34Ab1/Cry35Ab1、或修饰的Cry3A (mCry3A) 或eCry3.1Ab蛋白而控制玉米根虫的转基因玉米杂交种在美国已经是可商购的。

[0006] 尽管已经显示使用表达Cry蛋白的转基因植物非常有效,但现在针对某些转基因植物中表达的Cry蛋白具有抗性的昆虫有害生物是已知的。因此,仍需要鉴定新型并有效的有害生物控制剂,这些有害生物控制剂为农民提供经济利益并且是环境可接受的。特别需要的是对根萤叶甲属 (*Diabrotica*) 物种 (一种主要的玉米有害生物) 有毒的蛋白质,与现有昆虫控制产品相比这些蛋白质具有不同的作用方式、以缓和抗性发展。此外,通过这些使环境负担最小化的产品 (如通过转基因植物) 递送昆虫控制剂是令人希望的。

发明内容

[0007] 鉴于这些需要,本发明提供了从沙雷氏菌属细菌和相关细菌中分离的新颖杀昆虫蛋白,即SproCRW、SplyCRW和SquiCRW,统称为沙雷氏菌属杀昆虫蛋白 (SIP)。本发明还提供本发明的沙雷氏菌属杀昆虫蛋白的变体,以及与本发明的SIP及其变体基本相同的蛋白质。本发明的蛋白质对昆虫有害生物具有毒性,特别是玉米根虫 (根萤叶甲属物种) 有害生物。本发明进一步提供编码SIP和/或SIP变体的核酸分子,它们的互补序列,或编码与SIP和/或SIP变体基本相同的蛋白的核酸分子,它们的互补序列。

[0008] 本发明还提供了含有重组 (或与其互补的) 核酸的载体;包括和能够表达这样的核酸的植物或微生物;用这样的核酸转化的植物,例如转基因玉米植物;这样的植物的后代,其含有稳定地掺入和以孟德尔方式遗传的核酸,和/或这样的植物的种子和这样的后代。本发明还提供育种的方法,以将包含本发明的核酸分子的转基因引入子代植物和各种种质中。

[0009] 本发明还提供了包含本发明的SIP或SIP变体的组合物和配制品,其能够抑制昆虫有害生物存活、生长和/或繁殖的能力,或限制与昆虫相关的对作物的损害或损失的能力,例如,将SIP或SIP变体作为组合物或配制品的一部分施用于受昆虫侵染的区域或植物,或预防性处理易受昆虫侵染的区域或植物以赋予对昆虫有害生物的保护。

[0010] 本发明进一步通过制造SIP或SIP变体的方法,以及使用所述核酸的方法,例如,在微生物中控制昆虫或者在转基因植物中赋予免于昆虫损害的保护。

[0011] 本文描述的新颖蛋白针对昆虫具有活性。例如,在一些实施例中,本发明的蛋白质可用于控制经济上重要的昆虫有害生物,包括鞘翅目昆虫,如西方玉米根虫 (WCR; 玉米根萤叶甲)、北方玉米根虫 (NCR; 长角巴氏根萤叶甲)、南方玉米根虫 (SCR; 黄瓜十一星叶甲食根亚种) 和/或墨西哥玉米根虫 (MCR; 墨西哥玉米根萤叶甲 (*D. virgifera zea*))。本发明的杀昆虫蛋白可以单独使用或与其他昆虫控制策略组合使用,来以最小的环境影响赋予针对相同昆虫有害生物的增强的有害生物控制效率和/或增加靶昆虫的谱。

[0012] 通过学习本发明的以下说明书和非限制性实例,本发明的其他方面和优点对于本领域的技术人员而言将变得清楚。

序列表中的序列简述

SEQ ID NO:1是变形斑沙雷氏菌SproCRW氨基酸序列。

SEQ ID NO:2是普利茅斯沙雷氏菌SplyCRW氨基酸序列。

SEQ ID NO:3是奎沃斯沙雷氏菌 (*Serratia Quinivorans*) SquiCRW氨基酸序列。

SEQ ID NO:4是SproCRW V8A L11I氨基酸序列。

SEQ ID NO:5是SproCRW Y188Xaa氨基酸序列。

SEQ ID NO:6是SproCRW Y413W氨基酸序列。

SEQ ID NO:7是SproCRW F428W氨基酸序列。

SEQ ID NO:8是SproCRW Y430W氨基酸序列。

SEQ ID NO:9是SproCRW S190P氨基酸序列。

SEQ ID NO:10是SproCRW V192Y氨基酸序列。

SEQ ID NO:11是SproCRW C312L氨基酸序列。

SEQ ID NO:12是SproCRW C382Xaa氨基酸序列。

SEQ ID NO:13是SproCRW C482L氨基酸序列。

SEQ ID NO:14是SproCRW V398L氨基酸序列。

SEQ ID NO:15是SproCRW P396L I397L氨基酸序列。

SEQ ID NO:16是SproCRW I397L氨基酸序列。

SEQ ID NO:17是SproCRW I397L V398L氨基酸序列。

SEQ ID NO:18是SproCRW Y480L C482L氨基酸序列。

SEQ ID NO:19是SproCRW C482L T483L氨基酸序列。

SEQ ID NO:20是SplyCRW C481L氨基酸序列。

SEQ ID NO:21是SplyCRW Y479L C481L氨基酸序列。

SEQ ID NO:22是SplyCRW C481L T482L氨基酸序列。

SEQ ID NO:23是变形斑沙雷氏菌SproCRW核苷酸序列。

SEQ ID NO:24是普利茅斯沙雷氏菌SplyCRW核苷酸序列。

SEQ ID NO:25是奎沃斯沙雷氏菌SquiCRW核苷酸序列。

SEQ ID NO:26是SproCRW大肠杆菌优化的序列。

SEQ ID NO:27是SplyCRW大肠杆菌优化的序列。

SEQ ID NO:28是SquiCRW大肠杆菌优化的序列。

SEQ ID NO:29是SproCRW V8A L11I大肠杆菌优化的核苷酸序列。

SEQ ID NO:30是SproCRW Y188W大肠杆菌优化的核苷酸序列。

SEQ ID NO:31是SproCRW Y188F大肠杆菌优化的核苷酸序列。

SEQ ID NO:32是SproCRW Y413W大肠杆菌优化的核苷酸序列。

SEQ ID NO:33是SproCRW F428W大肠杆菌优化的核苷酸序列。

SEQ ID NO:34是SproCRW Y430W大肠杆菌优化的核苷酸序列。

SEQ ID NO:35是SproCRW S190P大肠杆菌优化的核苷酸序列。

SEQ ID NO:36是SproCRW V192Y大肠杆菌优化的核苷酸序列。

SEQ ID NO:37是SproCRW C312L大肠杆菌优化的核苷酸序列。

SEQ ID NO:38是SproCRW C382A大肠杆菌优化的核苷酸序列。

SEQ ID NO:39是SproCRW C382S大肠杆菌优化的核苷酸序列。

SEQ ID NO:40是SproCRW C382T大肠杆菌优化的核苷酸序列。

SEQ ID NO:41是SproCRW C382M大肠杆菌优化的核苷酸序列。

SEQ ID NO:42是SproCRW C382G大肠杆菌优化的核苷酸序列。
SEQ ID NO:43是SproCRW C382F大肠杆菌优化的核苷酸序列。
SEQ ID NO:44是SproCRW C482L大肠杆菌优化的核苷酸序列。
SEQ ID NO:45是SproCRW V398L大肠杆菌优化的核苷酸序列。
SEQ ID NO:46是SproCRW P396-L-L-I397大肠杆菌优化的核苷酸序列。
SEQ ID NO:47是SproCRW I397L大肠杆菌优化的核苷酸序列。
SEQ ID NO:48是SproCRW I397L V398L大肠杆菌优化的核苷酸序列。
SEQ ID NO:49是SproCRW Y480L C482L大肠杆菌优化的核苷酸序列。
SEQ ID NO:50是SproCRW C482L T483L大肠杆菌优化的核苷酸序列。
SEQ ID NO:51是SplyCRW C481L大肠杆菌优化的核苷酸序列。
SEQ ID NO:52是SplyCRW Y479LC481L大肠杆菌优化的核苷酸序列。
SEQ ID NO:53是SplyCRW C481L T482L大肠杆菌优化的核苷酸序列。
SEQ ID NO:54是SproCRW V8A L11I玉蜀黍优化的核苷酸序列。

具体实施方式

[0013] 本说明不旨在是可以实施本发明的所有不同方式,或可以添加到本发明中的所有特征的详细目录。例如,关于一个实施例所说明的特征可以并入其他实施例中,并且关于一个特定实施例所说明的特征可以从那个实施例删除。因此,本发明考虑了,在本发明的一些实施例中,可以排除或省略本文阐述的任何特征或特征的组合。此外,鉴于本披露,本文建议的不同实施例的众多变化以及附加对于本领域技术人员是显而易见的,这不脱离本发明。因此,以下说明旨在阐述本发明的一些特定实施例,并且并没有穷尽地叙述其所有排列、组合和变化。

[0014] 除非另外定义,本文所使用的所有技术和科学术语均具有与本发明所属领域的普通技术人员通常所理解的相同的含义。在本文的发明的说明中使用的术语是仅出于描述特定实施例的目的,且并不旨在限制本发明。

[0015] 本文引用的所有的公开、专利申请、专利以及其他参考文件对于引用中提及的有关句子和/或段落的传授内容通过引用以其全文并入本文。

[0016] 在此提供的核苷酸序列以5'至3'方向从左至右表示,并且使用代表核苷酸碱基的标准代码表示,如37CFR§§1.821-1.825和世界知识产权组织(WIPO)标准ST.25中所述,例如:腺嘌呤(A)、胞嘧啶(C)、胸腺嘧啶(T)、以及鸟嘌呤(G)。

[0017] 氨基酸同样也是使用WIPO标准ST.25来指示,例如: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)、谷氨酰胺(Gln;Q)、谷氨酸(Glu;E)、甘氨酸(Gly;G)、组氨酸(His;H)、异亮氨酸(Ile;I)、亮氨酸(Leu;L)、赖氨酸(Lys;K)、甲硫氨酸(Met;M)、苯丙氨酸(Phe;F)、脯氨酸(Pro;P)、丝氨酸(Ser;S)、苏氨酸(Thr;T)、色氨酸(Trp;W)、酪氨酸(Tyr;Y)、以及缬氨酸(Val;V)。

[0018] 除非上下文另外指示,否则明确地预期本文所述的本发明的不同特征能以任何组合使用。此外,本发明还考虑到在本发明的一些实施例中,本文陈述的任何特征或特征的组合可以被排除或省略。举例说明,如果本说明书陈述组合物包含组分A、B和C,明确地预期A、B或C的任何一种或其组合可单一地或以任何组合被省略和放弃。

定义

[0019] 为了清晰起见,定义了在本说明书中所使用的某些术语并且将其呈现如下:

[0020] 当在此和所附权利要求书中使用时,单数形式“一个/种(a/an)”和“所述(the)”包括复数指代物,除非上下文另外明确地指示。因此,例如,提及“一种植物”是提及一种或多种植物并且包括本领域技术人员已知的其等效物等。

[0021] 如在此使用的,词语“和/或”是指并且涵盖一个或多个相关联的列出项的任何及全部可能组合,连同当以可替代性(“或”)解释时组合的缺少。

[0022] 术语“约”本文用于意指大约、大致、约或在……左右。当术语“约”结合数值范围来使用时,它通过将边界延伸至高于以及低于所阐述的数值来限定这个范围。一般而言,术语“约”本文用于将数值限定至以20%的变化,优选地10%上下(更高或更低)地高于以及低于规定值。关于温度,术语“约”意指 $\pm 1^{\circ}\text{C}$,优选 $\pm 0.5^{\circ}\text{C}$ 。当术语“约”被用于本发明的上下文中(例如与温度或分子量值组合)时,确切值(即,无“约”)是优选的。

[0023] 除非上下文另有说明,如本文所使用的,例如“在约X和Y之间”、“在约X和约Y之间”、“从X至Y”和“从约X至约Y”(以及类似短语)的短语应被解释为包括X和Y。

[0024] 如本文使用的,术语“扩增的”意指使用至少一种核酸分子作为模板,构建核酸分子的多个拷贝或与该核酸分子互补的多个拷贝。扩增系统包括聚合酶链式反应(PCR)系统、连接酶链式反应(LCR)系统、基于核酸序列的扩增(NASBA,安大略省密西索加的坎基尼公司(Cangene, Mississauga, Ontario))、Q- β 复制酶系统、基于转录的扩增系统(TAS)、以及链置换扩增(SDA)。参见,例如,Diagnostic Molecular Microbiology: Principles and Applications[诊断分子微生物学:原理与应用],PERSING等人编,华盛顿美国微生物学会(American Society for Microbiology, Washington, D.C.), (1993)。扩增产物被称为“扩增子”。

[0025] 本发明的杀昆虫蛋白的“活性”意指杀昆虫蛋白作为口服活性的昆虫控制剂发挥作用,具有毒性作用、和/或能够干扰或阻止昆虫摄食,这可能引起或者可能不引起昆虫的死亡。当本发明的杀昆虫蛋白被递送至昆虫时,这种结果典型地是该昆虫的死亡,或者该昆虫不以使该杀昆虫蛋白可达该昆虫的来源为食。“杀有害生物的”被定义为有毒的生物活性,其能够控制有害生物(如昆虫、线虫、真菌、细菌或病毒),优选地通过杀死或破坏它们来进行控制。“杀昆虫的”被定义为有毒的生物活性,其能够控制昆虫,优选地通过杀死它们来进行控制。“杀有害生物剂”是具有杀有害生物活性的药剂。“杀昆虫剂”是具有杀昆虫活性的药剂。

[0026] 如在此使用的术语“嵌合构建体”或“嵌合基因”或“嵌合多核苷酸”或“嵌合核酸”(或类似术语)是指如下构建体或分子,该构建体或分子包含被组装进单个核酸分子中的不同来源的两个或更多个多核苷酸。术语“嵌合构建体”、“嵌合基因”、“嵌合多核苷酸”或“嵌合核酸”是指如下任何构建体或分子,该构建体或分子含有但不限于(1)多核苷酸(例如,DNA),包括在自然界中没有被发现在一起的调节多核苷酸和编码多核苷酸(即,构建体中的至少一个多核苷酸相对于它的其他多核苷酸中的至少一个是异源的),或(2)编码不是天然毗邻的蛋白质部分的多核苷酸,或(3)不是天然毗邻的启动子部分。另外,嵌合构建体、嵌合基因、嵌合多核苷酸或嵌合核酸可以包含衍生自不同来源的调节多核苷酸和编码多核苷酸,或包含衍生自相同来源、但以与在自然界中所发现的不同的方式进行布置的调节多核

苷酸和编码多核苷酸。在本发明的一些实施例中,嵌合构建体、嵌合基因、嵌合多核苷酸或嵌合核酸包含表达盒,该表达盒包含在调节多核苷酸的控制下、特别地在植物或细菌中具有功能性的调节多核苷酸的控制下的本发明的多核苷酸。

[0027] “编码序列”(CDS)是转录成RNA(例如mRNA、rRNA、tRNA、snRNA、有义RNA或反义RNA)的核酸序列。优选地,RNA进而在生物中被翻译以产生蛋白质。

[0028] 如在此使用的,“密码子优化的”序列意指如下核苷酸序列,其中这些密码子被选择以反映宿主细胞或生物可以具有的特定的密码子偏好性。这典型地是以这样一种方式来完成,该方式是为了保持由待优化的核苷酸序列所编码的多肽的氨基酸序列。在某些实施例中,重组DNA构建体的DNA序列包括已经针对该构建体有待在其中进行表达的细胞(例如,动物、植物、或真菌细胞)进行了密码子优化的序列。例如,有待在植物细胞中表达的构建体可以使其全部或部分序列(例如,第一基因抑制元件或基因表达元件)进行密码子优化用于在植物中表达。参见例如,美国专利号6,121,014,通过引用并入本文。

[0029] “控制”昆虫意指通过毒性作用抑制了昆虫有害生物存活、生长、摄食、和/或繁殖的能力,或限制与昆虫有关的损害或作物植物中的损失。“控制”昆虫可以是或可以不是意指杀死昆虫,尽管其优选意指杀死昆虫。

[0030] 术语“包含(comprises或comprising)”当用于本说明书中时指示所说明的特征、整数、步骤、操作、要素、或组分的存在,但并不排除一个或多个其他特征、整数、步骤、操作、要素、组分、或其组的存在或添加。

[0031] 如本文使用的,过渡短语“基本上由……组成”(以及语法变体)意指,权利要求书的范围有待被解读为涵盖权利要求书中所列举的指定材料或步骤以及不实质上改变所要求的发明的一个或多个基本和新颖特征的那些。因此,当用于本发明的权利要求中时,术语“基本上由……组成”并不旨在被解释为等同于“包含(comprising)”。

[0032] 在本发明的上下文中,“对应于(corresponding to或corresponds to)”意指当杀昆虫蛋白或其变体或同源物的氨基酸序列与彼此比对时,“对应于”在该变体或同系物蛋白中某些枚举的位置的氨基酸是与参考蛋白中的这些位置比对的那些,但在相对于本发明的特定参考氨基酸序列而言的这些精确的数字位置中是不必要的。例如,如果SEQ ID NO:1是参考序列,并且与SEQ ID NO:2比对,则SEQ ID NO:2的位置420的氨基酸Asn(Asn420)“对应于”SEQ ID NO:1的位置421的Asn(Asn421),或例如SEQ ID NO:2的Asn424“对应于”SEQ ID NO:1的Gly425。

[0033] “递送”组合物或毒性蛋白意指该组合物或毒性蛋白与昆虫接触,这促进该组合物或毒性蛋白的经口摄取,产生对昆虫的毒性作用和控制。可以按照许多公认的方式,包括但不限于转基因植物表达、一种或多种配制的蛋白质组合物、一种或多种可喷洒的蛋白质组合物、饵基(bait matrix)、或任何其他领域公认的蛋白质递送系统来递送该组合物或毒性蛋白。

[0034] 术语“结构域”是指沿着进化相关蛋白的序列的比对在特定位置处保守的一组氨基酸。虽然其他位置上的氨基酸可在同系物之间有所不同,但是在特定位置处高度保守的氨基酸指示在蛋白质的结构、稳定性或功能中很可能是必需的氨基酸。通过其在蛋白质同系物家族的经比对序列中的高度保守性进行鉴别,其可用作鉴别物(identifier),用来确定所讨论的任何多肽是否属于先前鉴别的多肽组。

[0035] “控昆虫有效量”意指杀昆虫蛋白的浓度,其通过毒性作用抑制昆虫存活、生长、摄食和/或繁殖的能力,或限制昆虫相关的损害或作物植物损失。“控昆虫有效量”可以是或可以不是意指杀死昆虫,尽管其优选意指杀死昆虫。

[0036] 如本文使用的“表达盒”意指能够在适当的宿主细胞中指导特定的核苷酸序列的表达的核酸序列,包含可操作地连接至目的核苷酸序列的启动子,该目的核苷酸序列可操作地连接至终止信号。它还典型地包含适当翻译该核苷酸序列所需要的序列。包含该目的核苷酸序列的表达盒可能具有其组分中的至少一种,该组分相对于它的其他组分中的至少一种而言是异源的。该表达盒还可以是天然存在的但已经是以对于异源表达有用的重组形式而获得的表达盒。然而,典型地,该表达盒相对于该宿主而言是异源的,即该表达盒的特定核酸序列不是天然存在于该宿主细胞中的,并且必须已经通过转化事件引入到该宿主细胞或该宿主细胞的祖先中。在该表达盒中核苷酸序列的表达可以是在组成型启动子或诱导型启动子的控制之下,该启动子只有当该宿主细胞暴露于一些特定的外界刺激时才引发转录。在多细胞生物体(例如植物)的情况下,该启动子对于特定组织、或器官、或者发育阶段也可以是特异的。

[0037] 包含目的核苷酸序列的表达盒可以是嵌合的,意味着它的组分中的至少一种相对于它的其他组分中的至少一种而言是异源的。表达盒还可以是包括驱动其天然基因的天然启动子,然而已经以对于异源表达有用的重组形式获得。表达盒的这种用途使它如此在其被引入的细胞中不是天然存在的。

[0038] 表达盒还可以任选地包括在植物中发挥作用的转录和/或翻译终止区(即,终止区)。多种转录终止子是可供用于在表达盒中使用的并且负责在超出目的异源核苷酸序列时的转录终止以及正确的mRNA聚腺苷酸化。终止区对于转录起始区可以是天然的,对于可操作地连接的目的核苷酸序列可以是天然的,对于植物宿主可以是天然的,或者可以是源自另一种来源(即,对于启动子、目的核苷酸序列、植物宿主、或其任何组合而言是外来的或异源的)。适当的转录终止子包括但不限于CAMV 35S终止子、tm1终止子、胭脂碱合酶终止子和/或豌豆rbcS E9终止子。这些终止子可以在单子叶植物和双子叶植物两者中使用。此外,可以使用编码序列的天然转录终止子。任何已知在植物中发挥作用的可供使用的终止子均可以在本发明的上下文中使用。

[0039] 当参考多核苷酸(如植物的基因、ORF或其部分、或转基因)使用时,术语“表达”是指通过基因的“转录”(即经由RNA聚合酶的酶促作用)将基因中编码的遗传信息转化为RNA(例如,mRNA、rRNA、tRNA或snRNA),并在适用的情况下(例如,如果基因编码蛋白)通过mRNA的“翻译”转化为蛋白质的过程。基因表达可以在该过程的许多阶段进行调节。例如,在反义构建体或dsRNA构建体的情况下,各自地表达可仅指该反义RNA或仅指dsRNA的转录。在实施例中,“表达”是指正义(mRNA)或功能性RNA的转录和稳定累积。“表达”还可指蛋白质的产生。

[0040] “基因”是位于基因组内并且包含编码核酸序列的限定区域,并且典型地还包含其他负责控制该编码部分表达(也就是转录和翻译)的主要调节性核酸。基因还可以包含其他5'和3'未翻译序列和终止序列。另外的可能存在的元件是例如内含子。如在自然界中所发现,基因的调节核酸序列在正常情况下可能不与该相关联的核酸序列进行可操作地连接,并因此不会是嵌合基因。

[0041] “目的基因”是指当转移至植物时,在该植物上赋予所希望的性状(如抗生素抗性、病毒抗性、昆虫抗性、疾病抗性、或对其他有害生物的抗性、除草剂耐受性、非生物胁迫耐受性、雄性不育、改性脂肪酸代谢、改性碳水化合物代谢、改善的营养价值、工业过程中改善的性能或改变的繁殖能力)的任何核酸分子。“目的基因”还可以是被转移至植物用于在该植物中产生商业上有价值的酶或代谢物的基因。

[0042] “异源”核酸序列或核酸分子是天然地不与将该核酸序列引入其中的宿主细胞相关联的核酸序列或核酸分子,包括天然存在的核酸序列的非天然存在的多个拷贝。异源核酸序列或核酸分子可以包含嵌合序列,如嵌合表达盒,在该表达盒中,启动子和编码区源自多源生物。启动子序列可以是组成型启动子序列、组织特异性启动子序列、化学诱导型启动子序列、伤口诱导型启动子序列、胁迫诱导型启动子序列、或发育阶段特异性启动子序列。

[0043] “同源”核酸序列是与其被引入的宿主细胞天然相关联的核酸序列。

[0044] “同源重组”是在同源核酸分子之间的核酸片段的相互交换。

[0045] 术语“基序”或“共有序列”或“特征(signature)”是指进化相关蛋白的序列中的短保守区。基序常常是结构域的高度保守部分,但是也可只包括结构域的一部分,或位于保守结构域外部(如果基序的所有氨基酸位于所限定结构域的外部)。

[0046] 在两个核酸或氨基酸序列的上下文中,术语“同一性”或“相同的”或“基本上相同的”是指当针对最大对应性进行比较和比对时具有至少60%、优选至少80%、更优选90%、甚至更优选95%、并且最优选至少99%核苷酸或氨基酸残基同一性的两个或更多个序列或子序列,如使用以下序列比较算法之一或通过目测检查所测量的。优选地,该基本的同一性存在于整个具有长度为至少约50个残基或碱基的序列的区域中,更优选地在整个至少约100个残基或碱基的区域中,并且最优选地这些序列在至少约150个残基或碱基上是基本上相同的。在尤其优选的实施例中,在整个编码区域长度中的序列是基本上相同的。此外,基本上相同的核酸或氨基酸序列基本上执行相同的功能。

[0047] 对于序列比较,典型地,一个序列充当与测试序列进行比较的参考序列。当使用序列比较算法时,将测试序列和参考序列输入到计算机中(如有必要,指定子序列坐标),并且指定序列算法程序的参数。然后,该序列比较算法基于所指定的程序参数来计算测试序列相对于参考序列的序列同一性百分比。

[0048] 用于比较的序列的最佳比对可以按照以下方式进行,例如通过Smith和Waterman, Adv. Appl. Math. [应用数学进展] 2:482 (1981) 的局部同源性算法、通过Needleman和Wunsch, J. Mol. Biol. [分子生物学杂志] 48:443 (1970) 的同源比对算法、通过Pearson和Lipman, Proc. Nat'l. Acad. Sci. USA [美国国家科学院院刊] 85:2444 (1988) 的相似性方法的搜索,通过这些算法的计算机化实施(威斯康星州遗传学分析软件包中的GAP、BESTFIT、FASTA和TFASTA,遗传学计算机组(Genetics Computer Group),科学街575号(575 Science Dr.),麦迪逊,威斯康星州),或通过目测检查(总体上参见Ausubel等人,下文)。

[0049] 适合于确定序列同一性百分比以及序列相似性的算法的一个实例是BLAST算法,其描述于以下文献中:Altschul等人, J. Mol. Biol. [分子生物学杂志] 215:403-410 (1990)。执行BLAST分析的软件是通过国家生物技术信息中心(the National Center for Biotechnology Information, 美国国家医学图书馆(U.S. National Library of Medicine), 洛克维尔大道8600号(8600 Rockville Pike), 贝塞斯达, 马里兰州20894美国)

可供公众使用的。这种算法涉及首先通过鉴定查询序列中具有长度W的短字码而鉴定得分高的序列对(HSP),这些得分高的序列对当与数据库序列中具有相同长度的字码(word)进行比对时匹配或满足一些正值阈值的得分T。T被称为邻近字码得分阈(Altschul等人,1990)。这些初始的邻近字码命中充当种子用于起始搜索以发现含有它们的较长的HSP。然后,将这些字码命中在两个方向上沿着每个序列延伸直到累积的比对得分可以增加。对于核苷酸序列,使用参数M(对于一对匹配残基的奖赏得分;总是>0)和N(对于错配残基的罚分;总是<0)来计算累积的得分。对于氨基酸序列,使用评分矩阵来计算累积得分。当累积的比对得分从它的最大达到值降低了数量X;由于累积一个或多个负得分的残基比对使累积得分趋于0或0以下;或者到达任一序列的末端时,停止字码命中在每个方向上的延伸。BLAST算法的参数W、T、以及X决定了比对的灵敏度与速度。BLASTN程序(对核苷酸序列来说)使用字长(W)为11、期望值(E)为10、截止值(cutoff)为100、M=5、N=-4、以及两条链的比较作为默认值。对于氨基酸序列,BLASTP程序使用字长(W)为3、期望值(E)为10、以及BLOSUM62评分矩阵作为默认值(参见Henikoff和Henikoff,Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]89:10915(1989))。

[0050] 除了计算序列同一性百分比之外,BLAST算法还进行两个序列之间相似性的统计分析(参见例如,Karlin和Altschul,Proc.Nat'l.Acad.Sci.USA[美国国家科学院院刊]90:5873-5787(1993))。由BLAST算法提供的相似性的一种量度是最小概率总和(P(N)),它提供了在两个核苷酸或氨基酸序列之间会偶然发生匹配的概率的指示。例如,若在测试核酸序列与参考核酸序列的比较中最小概率总和小于约0.1、更优选地小于约0.01、并且最优选地小于约0.001,则该测试核酸序列被认为是与该参考序列相似的。

[0051] 两个核酸序列基本上相同的另一个指示是这两种分子在严格条件下彼此杂交。短语“特异性杂交”是指分子在严格条件下仅与特定的核苷酸序列结合、双链化或杂交,这是在所述序列存在于复合混合物(例如,总细胞的)DNA或RNA中时进行的。“基本上结合”是指在探针核酸与靶核酸之间的互补杂交,并且涵盖少量错配,所述错配可以通过降低杂交介质的严格来容纳,以实现靶核酸序列的所希望的检测。

[0052] 在核酸杂交实验(如DNA杂交和RNA杂交)的上下文中,“严格杂交条件”和“严格杂交洗涤条件”是序列依赖性的,并且在不同的环境参数下是不同的。较长的序列在较高的温度下特异性杂交。对核酸杂交的广泛指导见于以下文献中:Tijssen(1993)Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes[生物化学和分子生物学实验室技术-使用核酸探针的杂交]第2章第I部分“Overview of principles of hybridization and the strategy of nucleic acid probe assays[杂交原理和核酸探针测定策略综述]”Elsevier[爱思唯尔集团],纽约。通常,高严格杂交和洗涤条件在限定的离子强度和pH下被选定为比特定序列的热熔点(T_m)低约5℃。典型地,在“严格条件”下,探针将会与它的靶子序列进行杂交,但不会与其他序列杂交。

[0053] T_m 是50%的靶序列与完全匹配的探针进行杂交时的温度(在限定的离子强度和pH下)。非常严格条件被选定为等于特定探针的 T_m 。对于互补核酸(它们在DNA或RNA印迹中在过滤器上具有超过100个互补的残基)的杂交的严格杂交条件的实例是在42℃下、具有1mg肝素的50%甲酰胺、将杂交进行过夜。高严格洗涤条件的实例是0.15M NaCl在72℃下持续约

15分钟。严格洗涤条件的实例是 $0.2 \times \text{SSC}$ 洗涤在 65°C 下持续15分钟(参见,Sambrook,下文,对于SSC缓冲液的说明)。通常,高严格洗涤之前会先进行低严格洗涤,以去除背景探针信号。对于例如超过100个核苷酸的双链体的中等严格洗涤的实例是 $1 \times \text{SSC}$ 在 45°C 下持续15分钟。对于例如超过100个核苷酸的双链体的低严格洗涤的实例是 $4-6 \times \text{SSC}$ 在 40°C 下持续15分钟。对于短探针(例如,约10-50个核苷酸),严格条件典型地涉及小于约 1.0M 的 Na^+ 离子的盐浓度,典型地在 $\text{pH } 7.0-8.3$ 下约 0.01 至 1.0M 的 Na^+ 离子浓度(或其他盐类),并且所述温度典型地是至少约 30°C 。还可以通过添加去稳定剂(如甲酰胺)来达到严格条件。一般而言,相比于不相关的探针,在特定的杂交测定中观察到高出2倍(或更高)的信噪比就表明检测到特异性杂交。如果在严格条件下彼此不杂交的核酸所编码的蛋白质是基本上相同的,则它们仍然是基本上相同的。例如,当使用遗传密码允许的最大程度的密码子简并而创建核酸的拷贝时,则发生这种情况。

[0054] 以下是可以用来克隆同源核苷酸序列(所述序列与本发明的参考核苷酸序列基本上相同)的杂交/洗涤条件的设置的实例:参考核苷酸序列在以下条件下优选地与所述参考核苷酸序列杂交:在 7% 十二烷基硫酸钠(SDS)、 0.5M NaPO_4 、 1mM EDTA 中在 50°C ,并且在 $2 \times \text{SSC}$ 、 0.1% SDS中在 50°C 洗涤;更令人希望的是在 7% 十二烷基硫酸钠(SDS)、 0.5M NaPO_4 、 1mM EDTA 中在 50°C ,并且在 $1 \times \text{SSC}$ 、 0.1% SDS中在 50°C 洗涤;仍更令人希望的是在 7% 十二烷基硫酸钠(SDS)、 0.5M NaPO_4 、 1mM EDTA 中在 50°C ,并且在 $0.5 \times \text{SSC}$ 、 0.1% SDS中在 50°C 洗涤;优选地在 7% 十二烷基硫酸钠(SDS)、 0.5M NaPO_4 、 1mM EDTA 中在 50°C ,并且在 $0.1 \times \text{SSC}$ 、 0.1% SDS中在 50°C 洗涤;更优选地在 7% 十二烷基硫酸钠(SDS)、 0.5M NaPO_4 、 1mM EDTA 中在 50°C ,并且在 $0.1 \times \text{SSC}$ 、 0.1% SDS中在 65°C 洗涤。

[0055] 两个核酸序列或蛋白质基本上相同的另一个指示是由第一核酸编码的蛋白质与由第二核酸编码的蛋白质进行免疫性交联反应或与其特异性结合。因此,蛋白质典型地是与第二蛋白质基本上相同的,例如其中这两种蛋白质仅在保守性取代上不同。

[0056] 术语“分离的”核酸分子、多核苷酸或蛋白质是不再存在于其天然环境中的核酸分子、多核苷酸或蛋白质。本发明的分离的核酸分子、多核苷酸或蛋白质可以按照纯化的形式存在,或者可以存在于重组宿主中,例如转基因细菌或转基因植物中。因此,当核酸分子或多核苷酸包含在转基因植物基因组内或蛋白质在转基因植物中表达时,如本文所列举的对“分离的”核酸分子、多核苷酸或蛋白质的权利要求涵盖了所述核酸分子、多核苷酸或蛋白质。

[0057] “核酸分子”或“核酸序列”或“多核苷酸”是可以从任何来源中分离的单链或双链DNA或RNA的区段。在本发明的上下文中,该核酸分子典型地是DNA的区段。在一些实施例中,本发明的核酸分子是分离的核酸分子。

[0058] “可操作地连接”是指在单一核酸片段上多核苷酸的关联,这样使得一者的功能影响另一者的功能。例如,当启动子能够影响编码多核苷酸或功能RNA的表达时(即,该编码多核苷酸或功能RNA处于该启动子的转录控制之下),则该启动子与该编码多核苷酸或功能RNA是可操作地连接的。正义方向或者反义方向的编码多核苷酸能够与调节多核苷酸可操作地连接。

[0059] 如在此使用的“杀有害生物”、“杀昆虫”等是指本发明的蛋白质控制有害生物的能力或者可以控制如在此所定义的有害生物的蛋白质的量。因此,杀有害生物蛋白可以杀死

或抑制有害生物(例如,昆虫有害生物)存活、生长、摄食、或繁殖的能力。

[0060] 术语“蛋白质”、“肽”和“多肽”在本文可以互换地使用。

[0061] “植物”是在发育的任何阶段的任何植物,特别是种子植物。示例性植物包括但不限于玉米(玉蜀黍(*Zea mays*))、卡罗拉(*canola*) (欧洲油菜(*Brassica napus*))、芜菁亚种(*Brassica rapa ssp.*)、苜蓿(紫花苜蓿(*Medicago saliva*))、水稻(栽培稻(*Oryza sativa*)),包括但不限于籼稻和/或粳稻)、油菜(欧洲油菜)、黑麦(*rye*, *Secale cereale*)、高粱(两色蜀黍(*Sorghum bicolor*))、*Sorghum vulgare*)、向日葵(*sunflower*, *Helianthus annuus*)、小麦(普通小麦(*Triticum aestivum*))、大豆(*soybean*, *Glycine max*)、烟草(普通烟草(*Nicotiana tobacum*))、马铃薯(洋芋(*Solanum tuberosum*))、花生(落花生(*Arachis hypogaea*))、棉花(陆地棉(*Gossypium hirsutum*))、甘薯(番薯(*Ipomoea batatas*))、木薯(树薯(*Manihot esculenta*))、咖啡(咖啡属亚种(*Cofea spp.*))、椰子(可可椰子(*Cocos nucifera*))、菠萝(凤梨(*Ananas comosus*))、柑桔树(柑橘属物种(*Citrus spp.*))、可可(*cocoa*, *Theobroma cacao*)、茶(茶树(*Camellia sinensis*))、香蕉(芭蕉属物种(*Musa spp.*))、鳄梨(*avocado*, *Persea americana*)、无花果(*fig*, *Ficus casica*)、番石榴(*guava*, *Psidium guajava*)、芒果(杧果(*Mangifera indica*))、橄榄(油橄榄(*Olea europaea*))、番木瓜(木瓜(*Carica papaya*))、腰果(*cashew*, *Anacardium occidentale*)、澳洲坚果(*macadamia*, *Macadamia integrifolia*)、扁桃(巴旦杏(*Prunus amygdalus*))、糖用甜菜(甜菜(*Beta vulgaris*))、苹果(*apple*, *Malus pumila*)、黑莓(悬钩子属(*Rubus*))、草莓(草莓属(*Fragaria*))、胡桃(普通胡桃(*Juglans regia*))、葡萄(酿酒葡萄(*Vitis vinifera*))、杏(*apricot*, *Prunus armeniaca*)、樱桃(李属(*Prunus*))、桃(碧桃(*Prunus persica*))、李(欧洲李(*Prunus domestica*))、梨(西洋梨(*Pyrus communis*))、西瓜(*watermelon*, *Citrullus vulgaris*)、浮萍(浮萍属(*Lemna spp.*))、燕麦(*oats*, *Avena sativa*)、大麦(*barley*, *Hordium vulgare*)、蔬菜、观赏植物、针叶树和草坪草(例如观赏性、娱乐或饲料用)、以及生物物质禾草(例如柳枝稷和芒草)。

[0062] 蔬菜包括但不限于茄属物种(例如西红柿、番茄(*Lycopersicon esculentum*))、莴苣(例如莴苣(*Lactuea sativa*))、胡萝卜(*Caucus carota*)、花椰菜(甘蓝(*Brassica oleracea*))、芹菜(旱芹(*Apium graveolens*))、茄子(*eggplant*, *Solanum melongena*)、芦笋(石刁柏(*Asparagus officinalis*))、秋葵(黄秋葵(*Abelmoschus esculentus*))、四季豆(菜豆(*Phaseolus vulgaris*))、青豆(*Phaseolus limensis*)、豌豆(香豌豆属(*Lathyrus spp.*))、南瓜属(*Cucurbita*)的成员如古巴瓜(*C.hubbard*)、冬南瓜(南瓜(*C.moschata*))、密生西葫芦(*zucchini*) (西葫芦(*C.pepo*))、曲颈南瓜(*C.crookneck*)、*C.argyrosperma*、*C.argyrosperma ssp sororia*、*C.digitata*、*C.ecuadorensis*、旱地油瓜(*C.foetidissima*)、*C.lundelliana*、以及*C.martinezii*、以及黄瓜属(*Cucumis*)的成员如黄瓜(*cucumber*, *Cucumis sativus*)、哈密瓜(*C.cantalupensis*)、以及甜瓜(香瓜(*C.melo*))。

[0063] 观赏植物包括但不限于杜鹃花(杜鹃属(*Rhododendron spp.*))、绣球花(*hydrangea*, *Macrophylla hydrangea*)、木槿(*Hibiscus rosasanensis*)、玫瑰(玫瑰属(*Rosa spp.*))、郁金香(郁金香属(*Tulipa spp.*))、水仙花(水仙属(*Narcissus spp.*))、矮牵牛(*petunias*, *Petunia hybrida*)、康乃馨(*carnation*, *Dianthus caryophyllus*)、猩猩木(一品红(*Euphorbia pulcherima*))、以及菊花。

[0064] 可以用于实践本发明的针叶树包括例如,松树如火炬松(loblolly pine,Pinus taeda)、湿地松(slash pine,Pinus elliotii)、杰克松(西黄松(Pinus ponderosa))、黑松(小干松(Pinus contorta))以及蒙特利松树(辐射松(Pinus radiata));花旗松(Douglas-fir,Pseudotsuga menziesii);美国西部铁杉(加拿大铁杉(Tsuga canadensis));美国西加云杉(白云杉(Picea glauca));红杉(北美红杉(Sequoia sempervirens));冷杉如银杉(温哥华冷杉(Abies amabilis))和香脂冷杉(加拿大胶杉(Abies balsamea));以及雪松,如美国西部侧柏(北美香柏(Thuja plicata))以及阿拉斯加黄杉(Alaska yellow-cedar)(黄扁柏(Chamaecyparis nootkatensis))。

[0065] 草坪草包括但不限于结缕草、剪股颖、羊茅草、早熟禾、奥古斯丁草、百慕达草、butfallograsses、黑麦草、以及鸭茅。

[0066] 还包括的是主要充当实验室模式的植物,例如拟南芥。

[0067] “植物细胞”是植物的结构和生理单位,包含原生质体和细胞壁。植物细胞可以处于分离的单个细胞或培养细胞的形式,或者是作为较高级的组织单位(诸如像植物组织、植物器官、或全株植物)的一部分。

[0068] “植物细胞培养物”意指植物单元(如例如,原生质体、细胞培养物细胞、植物组织中的细胞、花粉、花粉管、胚珠、胚囊、接合子以及处于不同发育阶段的胚)的培养物。

[0069] “植物材料”是指叶、茎、根、花或花的部分、果实、花粉、卵细胞、接合子、种子、切条、细胞或组织培养物、或植物的任何其他部分或产物。

[0070] “植物器官”是植物的独特而明显的已结构化并且分化的部分,如根、茎、叶、花蕾或胚。

[0071] 如本文所使用的,“植物组织”意指组织化成结构和功能单元的一组植物细胞。包括植物中或培养物中的任何植物组织。这个术语包括但不限于全株植物、植物器官、植物种子、组织培养物以及被组织化成结构和/或功能单元的任何植物细胞群组。这个术语与如以上列出的或由该定义以其他方式涵盖的任何具体类型的植物组织的联合应用或单独应用并不旨在排除任何其他类型的植物组织。

[0072] “多核苷酸”是指由共价键合于链中的许多核苷酸单体构成的聚合物。此类“多核苷酸”包括DNA、RNA、经修饰的寡核苷酸(例如,包含对于生物RNA或DNA不典型的碱基的寡核苷酸,如2'-O-甲基化寡核苷酸)等。在一些实施例中,核酸或多核苷酸可以是单链的、双链的、多链的或其组合。除非另外指示,否则本发明的具体核酸或多核苷酸任选地包含或编码除明确指示的任何多核苷酸之外的互补多核苷酸。

[0073] “目的多核苷酸”是指任何多核苷酸,当其转移至生物(例如,植物)中时赋予该生物所希望的特征,如昆虫抗性、疾病抗性、除草剂耐受性、抗生素抗性、改进的营养价值、工业过程中改进的性能、商业上有价值的酶或代谢物的生产、或者改变的繁殖能力。

[0074] “启动子”是编码区的不被翻译的DNA序列上游,其包含RNA聚合酶结合位点并且起始DNA的转录。启动子区还可以包括充当基因表达的调节物的其他元件。

[0075] 如在此使用的,术语“重组”是指核酸分子(例如,DNA或RNA)或蛋白质或生物的如下形式,该形式通常不会在自然界中发现并且正因为如此通过人类干预来产生。如在此使用的,“重组核酸分子”是包括多核苷酸组合的核酸分子,这些多核苷酸不会天然地一起存在并且是人类干预的结果,例如,由至少两种彼此异源的多核苷酸的组合组成的核酸分子,

或人工合成的(例如,使用组装的核苷酸序列合成的多核苷酸)并且包含偏离通常存在于自然界中的多核苷酸的多核苷酸的核酸分子,或包含人工掺入至宿主细胞的基因组DNA中和该宿主细胞基因组相关侧翼DNA中的转基因的核酸分子。重组核酸分子的另一个实例是由将转基因插入至植物的基因组DNA中产生的DNA分子,其可以最终导致该生物中的重组RNA/或蛋白质分子的表达。如在此使用的,“重组植物”是通常不会在自然界中存在的植物,是人类干预的结果,并且含有掺入至其基因组中的转基因和/或异源核酸分子。由于此类基因组改变,该重组植物明显不同于相关的野生型植物。

[0076] “调节元件”是指参与控制核苷酸序列的表达的序列。调节元件包含可操作地连接至目的核苷酸序列的启动子以及终止信号。它们还典型地涵盖适当翻译该核苷酸序列所需的序列。

[0077] “转化”是用于将异源核酸引入到宿主细胞或生物的方法。在具体实施例中,“转化”意指DNA分子稳定地整合到目的生物的基因组(核或质体)中。

[0078] “转化的/转基因的/重组的”是指异源核酸分子已经引入其中的宿主生物例如细菌或植物。核酸分子可以稳定整合到宿主基因组中,或者,核酸分子也可以作为染色体外分子存在。这种染色体外分子能够自主复制。转化的细胞、组织或植物应当理解为不仅涵盖转化过程的终产物,而且涵盖其转基因子代。“非转化的”、“非转基因的”、或“非重组的”宿主是指不含该异源的核酸分子的野生型生物,例如细菌或植物。

[0079] 本发明提供了用于控制有害的昆虫有害生物的组合物以及方法。特别地,本发明涉及新颖杀昆蛋白,特别是沙雷氏菌属杀昆虫蛋白(SIP),其至少对鞘翅目昆虫具有活性,例如玉米根萤叶甲(西方玉米根虫;WCR)、巴氏根萤叶甲(*Diabrotica barberi*) (北方玉米根虫;NCR)和/或黄瓜十一星叶甲食根亚种(南方玉米根虫;SCR)和/或其他根萤叶甲属物种(包括墨西哥玉米根萤叶甲(墨西哥玉米根虫)。天然SIP是由沙雷氏菌属的革兰氏阴性细菌产生的。本文所述的这样的SIP的实例包括由以下产生的那些:变形斑沙雷氏菌(其产生SproCRW杀昆虫蛋白)、普利茅斯沙雷氏菌(其产生SpIyCRW杀昆虫蛋白)和奎沃斯沙雷氏菌(其产生SquiCRW杀昆虫蛋白)。在一些实施例中,本发明的新颖杀昆虫蛋白可能对鳞翅目昆虫有害生物具有活性,包括但不限于小地老虎(黑色地老虎)、小蔗螟(甘蔗蛀虫;SCB)和/或西南玉米螟虫(西南玉米螟;SWCB))。本发明还涉及核酸(它们的表达产生了本发明的SIP),以及涉及制造和使用这些SIP以控制昆虫有害生物的方法。在某些实施例中,这些核酸的表达产生杀昆虫蛋白,其可用于至少控制鞘翅目昆虫(如西方玉米根虫、北方玉米根虫和/或南方玉米根虫),特别是当表达于转基因植物(如转基因玉米植物)中时。

[0080] 在一些实施例中,本发明涵盖核酸分子,所述核酸分子包含以下、基本上由以下组成或由以下组成:编码对昆虫有害生物有毒的蛋白质(即杀昆虫蛋白)的核苷酸序列,其中所述核苷酸序列(a) 编码包含氨基酸序列的蛋白质,所述氨基酸序列与SEQ ID NO:1-22中的任一个或其毒素片段具有至少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%或具有100%序列同一性;(b) 与SEQ ID NO:23-54中的任一个或其毒素编码片段具有至少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%

或具有100%序列同一性;或者(c)是(a)或(b)的合成序列,所述合成序列已经进行密码子优化用于在转基因生物中表达。在其他实施例中,该杀昆虫蛋白包含以下、基本上由以下组成、或由以下组成:SEQ ID NO:1-22中的任一个的氨基酸序列或其毒性片段。在其他实施例中,该核苷酸序列包含以下、基本由以下组成、或由以下组成:SEQ ID NO:23-54中的任一个或其毒素编码片段。

[0081] 在一些实施例中,本发明涵盖嵌合基因,所述嵌合基因包含可操作地连接到核酸分子的异源启动子,所述核酸分子包含以下、基本上由以下组成或由以下组成:编码对昆虫有害生物有毒的蛋白质的核苷酸序列,其中所述核苷酸序列(a)编码包含氨基酸序列的蛋白质,所述氨基酸序列与SEQ ID NO:1-22中的任一个或其毒素片段具有至少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%或具有100%序列同一性;(b)与SEQ ID NO:23-54中的任一个或其毒素编码片段具有至少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%或具有100%序列同一性;或者(c)是(a)或(b)的合成序列,所述合成序列已经进行密码子优化用于在转基因生物中表达。在其他实施例中,杀昆虫蛋白包含SEQ ID NO:1-22中的任一个或其毒性片段的氨基酸序列。在其他实施例中,核苷酸序列包含SEQ ID NO:23-54中的任一个或其毒素编码片段。在这些实施例的一些方面中,嵌合基因是表达盒。

[0082] 在其他实施例中,包含在本发明嵌合基因或表达盒中的启动子是植物可表达型启动子。在这些实施例的方面,所述植物可表达型启动子选自下组启动子,该组由以下组成:泛素、夜香树属黄病毒、玉米TrpA、OsMADS 6、玉蜀黍H3组蛋白、噬菌体T3基因9'5'UTR、玉米蔗糖合成酶1、玉米醇脱氢酶1、玉米捕光复合物、玉米热休克蛋白、玉蜀黍mt1、豌豆小亚基RuBP羧化酶、稻肌动蛋白、稻亲环蛋白、Ti质粒甘露碱合酶、Ti质粒胭脂碱合酶、矮牵牛查尔酮异构酶、豆类富甘氨酸蛋白1、马铃薯糖蛋白、凝集素、CaMV 35S以及S-E9小亚基RuBP羧化酶启动子。

[0083] 在一些实施例中,由本发明的核酸分子或本发明的嵌合基因或本发明的表达盒编码的杀昆虫蛋白对鞘翅目昆虫有害生物具有活性。在这些实施例的一些方面中,鞘翅目昆虫有害生物在根萤叶甲属中。在其他方面,根萤叶甲属昆虫有害生物是玉米根萤叶甲(西方玉米根虫;WCR)、巴氏根萤叶甲(北方玉米根虫;NCR)和/或黄瓜十一星叶甲食根亚种(南方玉米根虫;SCR)和/或其他根萤叶甲属物种(包括墨西哥玉米根萤叶甲(墨西哥玉米根虫))。

[0084] 在一些实施例中,本发明的嵌合基因或表达盒包括编码本发明的杀昆虫蛋白的核苷酸序列,其中所述核苷酸序列经密码子优化以用于在转基因生物中表达。在这些实施例的一些方面中,转基因生物是细菌或植物。

[0085] 在一些方面,转基因细菌在芽孢杆菌属、梭菌属、致病杆菌属、发光杆菌属、巴斯德氏芽菌属、埃希氏菌属、假单胞菌属、欧文氏菌属、沙雷氏菌属、克雷伯菌属、沙门氏菌属、巴氏杆菌属、黄单胞菌属、链霉菌属、根瘤菌属、红假单胞菌属、嗜甲基菌属、农杆菌属、醋杆菌属、乳杆菌属、节杆菌属、固氮菌属、明串珠菌属或产碱杆菌属中。在其他方面,转基因细菌是大肠杆菌。在其他方面,该核苷酸序列包含以下、基本由以下组成、或由以下组成:SEQ ID

N0:26-53中的任一个。

[0086] 在其他方面,转基因植物是单子叶植物或双子叶植物。在仍其他方面,双子叶植物选自由以下组成的组:大豆、向日葵、番茄、芸苔属作物、棉花、甜菜和烟草。在另外的方面中,所述单子叶植物选自由以下组成的组:大麦、玉蜀黍、燕麦、稻、高粱、甘蔗和小麦。在一些方面,转基因植物是玉蜀黍植物。在其他方面,该核苷酸序列包含以下、基本由以下组成、或由以下组成:SEQ ID N0:54。

[0087] 在一些实施例中,本发明涵盖对昆虫有害生物有毒的蛋白质,以及任选地分离的蛋白质,即杀昆虫蛋白,其中所述蛋白质或分离的蛋白质包含以下、基本上由以下组成或由以下组成:(a) 氨基酸序列,所述氨基酸序列与SEQ ID N0:1-22中的任一个或其毒素片段具有至少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%或具有100%序列同一性;(b) 氨基酸序列,所述氨基酸序列包含以下、基本上由以下组成、或由以下组成:SEQ ID N0:1-22中的任一个或其毒素片段;(c) 由核苷酸序列编码的氨基酸序列,所述核苷酸序列与SEQ ID N0:23-54中的任一个或其毒素编码片段具有至少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%或具有100%序列同一性;或(d) 由核苷酸序列编码的氨基酸序列,所述核苷酸序列包含以下、基本上由以下组成、或由以下组成:SEQ ID N0:22-54中的任一个或其毒素编码片段。本领域技术人员将认识到,可以对本发明所涵盖的示例性杀昆虫蛋白进行修饰。此类修饰和基本上一致的核酸或氨基酸分子涵盖于本发明中。

[0088] 本发明还涵盖工程化的沙雷氏菌属杀昆虫蛋白(eSIP),其是本发明的突变SIP或变体SIP或经修饰的SIP。在一些实施例中,修饰可包括任何天然存在和/或非天然存在的氨基酸中的一个或多个的取代和/或插入。在一些实施例中,修饰包含以下、基本上由以下组成、或由以下组成:在SIP氨基酸序列的氨基酸位置处丙氨酸、精氨酸、天冬酰胺、天冬氨酸、半胱氨酸、谷氨酰胺、谷氨酸、甘氨酸、组氨酸、异亮氨酸、亮氨酸、赖氨酸、甲硫氨酸、苯丙氨酸、脯氨酸、丝氨酸、苏氨酸、色氨酸、酪氨酸和/或缬氨酸中的一个或多个的插入和/或取代。这种取代和/或插入可以通过改变编码SIP的核苷酸序列中的密码子来完成,从而得到编码eSIP的经修饰的SIP核苷酸序列。

[0089] 在一些实施例中,通过以下的取代和/或插入来修饰SIP:(a) 具有脂肪族疏水侧链的一个或多个氨基酸(例如丙氨酸、异亮氨酸、甲硫氨酸和/或缬氨酸;在实施例中,氨基酸不是丙氨酸);(b) 具有芳香族疏水侧链的一个或多个氨基酸(例如,苯丙氨酸、色氨酸和/或酪氨酸);(c) 具有极性中性侧链的一个或多个氨基酸(例如,天冬酰胺、半胱氨酸、谷氨酰胺、丝氨酸和/或苏氨酸);(d) 具有酸性侧链的一个或多个氨基酸(例如,天冬氨酸和/或谷氨酸);具有碱性侧链的一个或多个氨基酸(例如,精氨酸、组氨酸和/或赖氨酸);(e) 一个或多个甘氨酸残基;(f) 一个或多个脯氨酸残基;或(g) (a)至(f)的任何组合。

[0090] 在一些实施例中,本发明涵盖工程化杀昆虫蛋白,所述工程化杀昆虫蛋白包含以下、基本上由以下组成或由以下组成:与SEQ ID N0:1-3中的任一个具有至少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%或具有100%序列同一性的氨基酸序列并且进一步在对应于以下的位置处包含至少一个突变：(a) SEQ ID NO:1的氨基酸位置1-489；或(b) SEQ ID NO:2的氨基酸位置1-488；或(c) SEQ ID NO:3的氨基酸位置1-489。在这些实施例的一些方面中，突变位于与SEQ ID NO:1的氨基酸位置8、11、188、382、397、398、413、428、430或482或其任意组合对应的氨基酸位置处。在其他方面，突变位于SEQ ID NO:1的位置8、11、188、382、397、398、413、428、430或482处。在仍其他方面，突变位于与SEQ ID NO:1的氨基酸8和11或396和397对应的位置处。在另外的方面中，突变位于SEQ ID NO:1的氨基酸位置8和11或396和397处。在仍进一步的方面，位置8的突变是V8A，位置11的突变是L11I，位置188的突变是Y188F，位置382的突变是C382A、C382S或C382T，位置397的突变是I397L，位置398的突变是V398L，位置413的突变是Y413W，位置428的突变是F428W，位置430的突变是Y430W，或者位置482的突变是C482L。在这些实施例的一些方面中，蛋白质包含以下、基本上由以下组成或由以下组成：SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:17的氨基酸序列。

[0091] 在一些方面，本发明的工程化杀昆虫蛋白包含在与SEQ ID NO:2的氨基酸位置479、481和/或482对应的氨基酸位置处的突变。在其他方面，突变位于SEQ ID NO:2的位置479、481和/或482处。在仍其他方面，突变位于与SEQ ID NO:2的氨基酸479和481对应的位置处。在其他方面，突变位于SEQ ID NO:2的氨基酸位置481和482处。在仍进一步的方面，位置479的突变是Y479L，位置481的突变是C481L或位置482的突变是T482L。在这些实施例的一些方面中，蛋白质包含以下、基本上由以下组成或由以下组成：SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22的氨基酸序列。

[0092] 在一些实施例中，当在相同条件（例如酶浓度、蛋白质浓度、pH、温度和/或时间）下测试时，与合适的对照和/或不含本发明的修饰的亲本分子相比，本发明的突变体或变体SIP具有增强的被哺乳动物消化蛋白酶（例如，胃蛋白酶）的消化。评估胃蛋白酶和其他消化蛋白酶的蛋白质消化的方法是本领域众所周知的，例如，实例14中所述的模拟的胃液（SGF）测定。例如，使用胃蛋白酶的消化可以在以下条件进行：约37℃和约pH 1.2，任选地具有约10单位（U）胃蛋白酶/微克蛋白质的酶浓度。在这些实施例的一些方面中，通过取代与SEQ ID NO:1的氨基酸位置398或482对应的位置处的氨基酸来修饰SIP。在这些实施例的其他方面，通过取代SEQ ID NO:1的位置398或482处的氨基酸来修饰SIP。在这些实施例的其他方面，氨基酸位置398处的取代是V398L或氨基酸位置482处的取代是C482L。在仍进一步的方面，在SGF测定中具有增强的消化性的突变蛋白或变体蛋白包含以下、基本上由以下组成或由以下组成：SEQ ID NO:13或SEQ ID NO:14。

[0093] 在一些实施例中，本发明的杀昆虫蛋白，包括本发明的工程化杀昆虫蛋白，对鞘翅目昆虫有害生物具有活性。鞘翅目的昆虫包括但不限于目前已知的或之后鉴定的任何鞘翅目昆虫，包括原鞘亚目、粘食亚目、肉食亚目和多食亚目、及其任何组合中的那些。

[0094] 在这些实施例的一些方面中，本发明的杀昆虫蛋白针对根萤叶甲属物种具有活性。根萤叶甲属是鞘翅目的一种甲虫属，通常被称为“玉米根虫”或“黄瓜甲虫”。示例性根萤叶甲属物种包括但不限于：长角巴氏根萤叶甲（*Diabrotica longicornis barberi*）（北方玉米根虫）、玉米根萤叶甲指名亚种（西方玉米根虫）、黄瓜十一星叶甲食根亚种（南方玉米

根虫)、黄瓜条根萤叶甲 (*D.balteata*) (带状黄瓜甲虫 (banded cucumber beetle))、黄瓜十一星叶甲 (*D.undecimpunctata undecimpunctata*) (西方斑点黄瓜甲虫 (western spotted cucumber beetle))、斯格尼根萤叶甲 (*D.significata*) (3斑叶甲 (3-spotted leaf beetle))、南美叶甲 (*D.speciosa*) (菊花甲虫 (chrysanthemum beetle))、墨西哥玉米根萤叶甲 (墨西哥玉米根虫)、班尼根萤叶甲 (*D.beniensis*)、克里斯塔根萤叶甲 (*D.cristata*)、科威根萤叶甲 (*D.curvipustulata*)、双斑根萤叶甲 (*D.dissimilis*)、华丽根萤叶甲 (*D.elegantula*)、伊墨根萤叶甲 (*D.emorsitans*)、禾本科根萤叶甲 (*D.graminea*)、伊斯帕根根萤叶甲 (*D.hispanolae*)、莱米妮根萤叶甲 (*D.lemniscata*)、赭腿根萤叶甲 (*D.linsleyi*)、米勒根萤叶甲 (*D.milleri*)、钱币形根萤叶甲 (*D.nummularis*)、扇叶根萤叶甲 (*D.occlusa*)、普拉根萤叶甲 (*D.porracea*)、蜗牛根萤叶甲 (*D.scutellata*)、胫骨根萤叶甲 (*D.tibialis*)、三线根萤叶甲 (*D.trifasciata*) 以及微绿根萤叶甲 (*D.viridula*) ; 及其任何组合。

[0095] 根据本发明的鞘翅目昆虫有害生物的其他非限制性实例包括瘦跗叶甲属物种 (*Leptinotarsa* spp.), 如马铃薯叶甲 (科罗拉多马铃薯甲虫); 叶甲虫属物种 (*Chrysomela* spp.), 如黑杨叶甲 (*C.scripta*) (黑杨叶甲虫 (cottonwood leaf beetle)); 咪小蠹属物种 (*Hypothenemus* spp.), 如咖啡果小蠹 (*H.hampei*) (咖啡豆钻孔虫 (coffee berry borer)); 米象属物种 (*Sitophilus* spp.), 如玉米象 (*S.zeamais*) (玉蜀黍象 (maize weevil)); 毛跳甲属物种 (*Epitrix* spp.), 如烟草跳甲 (*E.hirtipennis*) (烟草跳甲 (tobacco flea beetle)) 和黄瓜跳甲 (*E.cucumeris*) (马铃薯跳甲 (potato flea beetle)); 条跳甲属物种 (*Phyllotreta* spp.), 如十字花科跳甲 (*P.cruciferae*) (十字花科植物跳甲 (crucifer flea beetle)) 和西方黑跳甲 (*P.pusilla*) (西方黑色跳甲 (western black flea beetle)); 花象属物种 (*Anthonomus* spp.), 如胡椒花象 (*A.eugenii*, pepper weevil); 金针虫属物种 (*Hemicrepidus* spp.), 如金针虫 (*H.memnonius*) (铁线虫 (wireworms)); 梳爪叩头虫属物种 (*Melanotus* spp.), 如普通梳爪叩头甲 (*M.communis*) (铁线虫 (wireworms)); 莢象甲属物种 (*Ceutorhynchus* spp.), 如甘蓝莢象甲 (*C.assimilis*) (甘蓝心皮象鼻虫 (cabbage seedpod weevil)); 条跳甲属物种, 如十字花科跳甲 (十字花科植物跳甲); *Aeolus* 属物种 (*Aeolus* spp.), 如 *A.mellillus* (铁线虫); *Aeolus* 属物种, 如 *A.mancus* (小麦铁线虫 (wheat wireworm)); 砂铁线属物种 (*Horistonotus* spp.), 如砂铁线虫 (*H.uhlerii*) (沙子铁线虫 (sand wireworm)); 尖隐喙象属物种 (*Sphenophorus* spp.), 如玉米谷象 (*S.maidis*) (玉蜀黍谷象 (maize billbug))、梯牧草谷象 (*S.zeae*) (梯牧草谷象虫 (timothy billbug))、牧草长喙象 (*S.parvulus*) (早熟禾谷象 (bluegrass billbug)) 和南方玉米长喙象 (*S.callosus*) (南方玉米谷象 (southern corn billbug)); 鳃角金龟属物种 (*Phyllophaga* spp.) (蛴螬 (white grub)); 凹胫跳甲属物种 (*Chaetocnema* spp.), 如玉米铜色跳甲 (*C.pulicaria*) (玉米跳甲 (corn flea beetle)); 弧丽金龟属物种 (*Popillia* spp.), 如日本弧丽金龟 (*P.japonica*) (日本金龟子 (Japanese beetle)); 食植瓢虫属物种 (*Epilachna* spp.), 如墨西哥豆瓢虫 (*E.varivestis*) (墨西哥豆甲虫 (Mexican bean beetle)); 萤叶甲属物种 (*Cerotoma* spp.), 如菜豆莹叶甲 (*C.trifurcate*) (豆叶甲 (Bean leaf beetle)); 豆芫菁属物种 (*Epicauta* spp.), 如边缘豆芫菁 (*E.pestifera*) 和芫菁 (*E.lemniscata*) (斑蝥 (lister beetles)); 以及前述的任何组合。

[0096] 本发明的杀昆虫蛋白也可对鳞翅目昆虫具有活性。这样的鳞翅目昆虫包括但不限于

于目前已知的或之后鉴定的任何被分类为鳞翅目昆虫的昆虫,包括在轭翅亚目、有喙亚目和异石蛾亚目及其任何组合内的那些昆虫物种。示例性鳞翅目昆虫包括但不限于:玉米螟属物种,如欧洲玉米螟(*O.nubilalis*,European corn borer);菜蛾属物种,如小菜蛾(*P.xylostella*,diamondback moth);灰翅夜蛾属物种(*Spodoptera* spp.),如草地贪夜蛾(*S.frugiperda*) (秋夜蛾(fall armyworm))、黄条粘虫(*S.ornithogalli*,yellowstriped armyworm)、西部黄条粘虫(*S.praefica*,western yellowstriped armyworm)、南部粘虫(*S.eridania*,southern armyworm)和甜菜夜蛾(*S.exigua*,beet armyworm);地老虎属物种,如小地老虎(*A.ipsilon*) (黑色地老虎)、普通地老虎(*A.segetum*,common cutworm)、泥背地老虎(*A.gladiaria*,claybacked cutworm)和西部灰地老虎(*A.orthogonia*,pale western cutworm);切根虫属物种(*Striacosta* spp.),如豆白缘切根虫(*S.albicosta*) (西部豆切根虫(western bean cutworm));铃夜蛾属物种(*Helicoverpa* spp.),如玉米穗虫(*H.zea*) (玉米穗蛾(corn earworm))、茶色铃夜蛾(*H.punctigera*) (原生夜蛾(native budworm))、海灰翅夜蛾(*S.littoralis*) (埃及棉树叶虫(Egyptian cotton leafworm)) and 棉铃虫(*H.armigera*,cotton bollworm);实夜蛾属物种(*Heliothis* spp.),如烟芽夜蛾(*H.virescens*) (烟夜蛾(tobacco budworm));蔗螟属物种(*Diatraea* spp.),如西南玉米螟(*D.grandiosella*,southwestern corn borer)和小蔗螟(*D.saccharalis*) (甘蔗蛀虫));粉纹夜蛾属物种(*Trichoplusia* spp.),如粉纹夜蛾(*T.ni*,cabbage looper);蛀茎夜蛾属物种(*Sesamia* spp.),如地中海玉米螟(*S.nonagroides*,Mediterranean corn borer);红铃虫属物种(*Pectinophora* spp.),如棉红铃虫(*P.gossypiella*,pink bollworm);纹卷蛾属物种(*Cochylis* spp.),如向日葵细卷叶蛾(*C.hospes*,banded sunflower moth);天蛾属物种(*Manduca* spp.),如烟草天蛾(*M.sexata*,tobacco hornworm)和番茄天蛾(*M.quinquemaculata*,tomato hornworm);玉米苗斑螟属物种(*Elasmopalpus* spp.),如南美玉米苗斑螟(*E.Lignosellus*) (小玉米茎蛀虫(lesser cornstalk borer));尺夜蛾属物种(*Pseudoplusia* spp.),如大豆尺蠖(*P.includens*) (大豆夜蛾);干煞夜蛾属物种(*Anticarsia* spp.),如黎豆夜蛾(*A.gemmatalis*) (绒毛豆毛虫(velvetbean caterpillar));绿夜蛾属物种(*Plathypena* spp),如苜蓿绿夜蛾(*P.scabra*,green cloverworm);马醉木属物种(*Pieris* spp.),如大菜粉蝶(*P.brassicae*) (纹白蝶(cabbage butterfly));夜蛾属物种(*Papaipema* spp.),如蛀茎夜蛾(*P.nebris*,stalk borer);黏虫属物种(*Pseudaletia* spp.),如一星黏虫(*P.unipuncta*) (普通黏虫);疆夜蛾属物种(*Peridroma* spp.),如杂色地老虎(*P.saucia*,variegated cutworm);茄茎麦蛾属物种(*Keiferia* spp.),如番茄蠹蛾(*K.lycopersicella*,tomato pinworm);菜粉蝶属物种(*Artogeia* spp.),如菜粉蝶(*A.rapae*,imported cabbageworm);茄麦蛾属物种(*Phthorimaea* spp.),如马铃薯麦蛾(*P.operculella*,potato tuberworm);透翅缓夜蛾属物种(*Crymodes* spp.),如*C.devastator*,glassy cutworm;脏切叶蛾属物种(*Feltia* spp.),如脏切夜蛾(*F.ducens*,dingy cutworm);以及前述的任何组合。

[0097] 本发明的杀昆虫蛋白还可以针对半翅类、双翅类、草盲蝽属物种和/或其他刺吸式昆虫(例如直翅目或缨翅目的刺吸式昆虫)具有活性。双翅目的昆虫包括但不限于目前已知的或之后鉴定的任何双翅目昆虫,包括但不限于斑潜蝇属物种(*Liriomyza* spp.),如三叶斑潜蝇(*L.trifolii*) (潜叶虫(leafminer)) and 美洲斑潜蝇(*L.sativae*) (蔬菜潜叶虫

(vegetable leafminer)); *Scrobipalpula* 属物种, 如番茄潜叶虫 (*S. absoluta*, tomato leafminer); 地种蝇属物种 (*Delia* spp.), 如玉米蝇蛆 (*D. platura*) (玉米种蝇 (seedcorn maggot))、甘蓝种蝇蛆 (*D. brassicae*) (甘蓝种蝇 (cabbage maggot)) 和甘蓝根花蝇 (*D. radicum*, cabbage root fly); 锈蝇属物种 (*Psilia* spp.), 如胡萝卜锈蝇 (*P. rosae*, carrot rust fly); 根斑蝇属物种 (*Tetanops* spp.), 如甜菜根蛆 (*T. myopaeformis*) (甜菜根斑蝇 (sugarbeet root maggot)); 以及前述的任何组合。

[0098] 直翅目的昆虫包括但不限于目前已知的或之后鉴定的任何直翅目昆虫, 包括但不限于黑蝗属物种 (*Melanoplus* spp.), 如异黑蝗 (*M. differentialis*) (长额负蝗 (*Differential grasshopper*))、赤胫黑蝗 (*M. femurrubrum*, Redlegged grasshopper)、双带黑蝗 (*M. bivittatus*, Twostriped grasshopper); 及其任何组合。

[0099] 缨翅目的昆虫包括但不限于目前已知的或之后鉴定的任何缨翅目昆虫, 包括但不限于花蓟马属物种 (*Frankliniella* spp.), 如西花蓟马 (*F. occidentalis*) (西方花蓟马 (western flower thrips)) 和烟褐花蓟马 (*F. fusca*) (烟草蓟马 (tobacco thrips)); 和蓟马属物种 (*Thrips* spp.), 如烟蓟马 (*T. tabaci*) (葱蓟马 (onion thrips))、瓜蓟马 (*T. palmi*, melon thrips); 以及前述的任何组合。

[0100] 本发明的杀昆虫蛋白还可以针对线虫具有活性。如本文使用的术语“线虫”涵盖目前已知的或之后鉴定的任何被分类为动物界线虫门的生物, 包括但不限于有腺纲 (包括例如: 嘴刺目、等咽目、单齿目、矛线目、毛首目、索虫目、姆斯帕目、薄咽目、色矛目、带线虫目、链环目和单宫目) 和/或胞管肾纲 (包括例如: 小杆目、圆线虫目、蛔虫目、旋尾目、驼形目、双胃目、垫刃目和滑刃目) 中的线虫。

[0101] 线虫包括但不限于寄生线虫, 如根结线虫、胞囊线虫和/或腐线虫。根据本发明的线虫的示例性属包括但不限于: 根结线虫属 (根结线虫)、异皮线虫属 (胞囊线虫)、球异皮线虫属 (*Globodera*) (胞囊线虫)、穿孔线虫属 (穿孔线虫)、肾状线虫属 (肾形肾状线虫)、短体线虫属 (腐线虫)、滑刃线虫属 (叶面线虫)、螺旋线虫属 (螺旋线虫)、纽带线虫属 (矛线虫)、拟毛刺属 (短粗根线虫)、长针线虫属、珍珠线虫属 (假根结线虫)、亚粒线虫属 (*Subanguina*)、刺线虫属 (刺线虫)、小环线虫属 (小环线虫)、环线虫属 (环线虫)、茎线虫属 (茎线虫)、锥线虫属 (锥线虫)、半轮线虫属 (半轮线虫)、鞘线虫属 (鞘线虫)、潜根线虫属 (潜根线虫)、根结线虫属 (*Hypsoperine*)、大茎线虫属 (大茎线虫)、*Melinius* 属、刻点胞囊属 (*Punctodera*)、五沟线虫属 (*Quinisulcius*)、盾线虫属 (盾线虫)、剑线虫属 (匕首线虫)、矮化线虫属 (矮化线虫)、穿刺线虫属 (穿刺线虫)、伞滑刃属 (蛔虫)、及其任何组合。

[0102] 根据本发明的示例性植物寄生线虫包括但不限于: 细小刺线虫 (*Belonolaimus gracilis*)、长尾刺线虫 (*Belonolaimus longicaudatus*)、松材线虫 (*Bursaphelenchus xylophilus*, pine wood nematode)、龙胆轮线虫属 (*Criconemoides ornata*)、马铃薯腐烂线虫 (*Ditylenchus destructor*, potato rot nematode)、鳞球茎茎线虫 (*Ditylenchus dipsaci*, stem and bulb nematode)、马铃薯胞囊线虫 (*Globodera pallida*, potato cyst nematode)、马铃薯金线虫 (*Globodera rostochiensis*) (金黄线虫 (golden nematode))、大豆胞囊线虫 (*Heterodera glycines*, soybean cyst nematode)、甜菜胞囊线虫 (*Heterodera schachtii*, sugar beet cyst nematode); 玉米胞囊线虫 (*Heterodera zeae*, corn cyst nematode)、禾谷胞囊线虫 (*Heterodera avenae*, cereal cyst nematode)、胡萝卜异皮线虫

(*Heterodera carotae*)、*Heterodera trifolii* (三叶异皮线虫)、哥伦布纽带线虫 (*Hoplolaimus columbus*)、帽状纽带线虫 (*Hoplolaimus galeatus*)、*Hoplolaimus magnistylus*、短环长针线虫 (*Longidorus breviannulatus*)、花生根结线虫 (*Meloidogyne arenaria*)、哥伦比亚根结线虫 (*Meloidogyne chitwoodi*)、北方根节线虫 (*Meloidogyne hapla*)、南方根结线虫 (*Meloidogyne incognita*)、爪哇根结线虫 (*Meloidogyne javanica*)、异盘中环线虫 (*Mesocriconema xenoplax*)、异常珍珠线虫 (*Nacobbus aberrans*)、*Nacobbus dorsalis*、*Paratrichodorus christiei*、微小拟毛刺线虫 (*Paratrichodorus minor*)、最短尾短体线虫 (*Pratylenchus brachyurus*)、刻痕短体线虫 (*Pratylenchus crenatus*)、*Pratylenchus hexincisus*、落选短体线虫 (*Pratylenchus neglectus*)、穿刺短体线虫 (*Pratylenchus penetrans*)、*Pratylenchus projectus*、斯克里布纳短体线虫 (*Pratylenchus scribneri*)、*Pratylenchus tenuicaudatus*、*Pratylenchus thornei*、玉米短体线虫 (*Pratylenchus zeae*)、*Punctodera chaccoensis*、*Quinisulcius acutus*、香蕉穿孔线虫 (*Radopholus similis*)、肾形轮线虫 (*Rotylenchulus reniformis*)、顺逆矮化线虫 (*Tylenchorhynchus dubius*)、柑桔半穿刺线虫 (*Tylenchulus semipenetrans*)、美洲剑线虫 (*Siphinema americanum*)、*X.Mediterraneum*、以及前述的任何组合。

[0103] 本发明还涵盖重组载体和/或重组构建体,这些重组载体或构建体也可被称为载体或构建体,其包含本发明的表达盒和/或核酸分子。在此类载体中,这些核酸优选地处于表达盒中,这些表达盒包含用于在能够表达核苷酸分子的宿主细胞中表达核苷酸分子的调节元件。此类调节元件通常包含启动子和终止信号并且优选地还包含元件,这些元件允许由本发明的核酸所编码的多肽的有效的翻译。包含核酸的载体能够在特定的宿主细胞中复制(优选是作为染色体外分子)并因此可用来在这些宿主细胞中扩增本发明的核酸。

[0104] 本发明还涵盖宿主细胞,所述宿主细胞包含本发明的重组载体、表达盒或核酸分子。在其他实施例中,此类载体是病毒载体并被用于在特定的宿主细胞(例如昆虫细胞或植物细胞)中复制核苷酸序列。重组载体也用于将本发明的核酸分子转化到宿主细胞中,由此这些核酸分子被稳定整合到转基因宿主的DNA中。在一些实施例中,宿主细胞是细菌细胞或植物细胞。在这些实施例的一些方面中,细菌细胞在芽孢杆菌属、梭菌属、致病杆菌属、发光杆菌属、巴斯德氏芽菌属、埃希氏菌属、假单胞菌属、欧文氏菌属、沙雷氏菌属、克雷伯菌属、沙门氏菌属、巴氏杆菌属、黄单胞菌属、链霉菌属、根瘤菌属、红假单胞菌属、嗜甲基菌属、农杆菌属、醋杆菌属、乳杆菌属、节杆菌属、固氮菌属、明串珠菌属或产碱杆菌属中。在这些实施例的其他方面,用于此类重组载体的宿主细胞是内生菌或附生菌。在这些实施例的一些其他方面,宿主细胞是植物细胞,例如双子叶植物细胞或单子叶植物细胞。在其他方面,双子叶植物细胞选自下组,该组由以下组成:大豆细胞、向日葵细胞、番茄细胞、芸苔属作物细胞、棉花细胞、甜菜细胞以及烟草细胞。在仍其他方面,单子叶植物细胞选自下组,该组由以下组成:大麦细胞、玉蜀黍细胞、燕麦细胞、水稻细胞、高粱细胞、甘蔗细胞以及小麦细胞。

[0105] 在本发明的一些实施例中,将本发明的核酸分子中的至少一个插入到适当的表达盒(包含启动子以及终止信号)中。核酸的表达可以是组成型的,或者可以使用响应于各种类型的刺激以起始转录的诱导型启动子。在另一个实施例中,其中表达了本发明的杀昆虫蛋白的细胞是微生物,如病毒、细菌或真菌。在又一个实施例中,病毒(如杆状病毒)在其

基因组中含有本发明的核酸并且在感染了适当的真核细胞(适合于病毒复制以及该核酸的表达)之后表达了大量相应的杀昆虫蛋白。将由此产生的杀昆虫蛋白用作杀昆虫剂。可替代地,被工程化以包括该核酸的杆状病毒被用来体内感染昆虫并通过该杀昆虫毒素的表达或通过病毒感染与该杀昆虫毒素的表达的组合而将其杀死。在一个另外的实施例中,本发明还涵盖了用于产生具有杀昆虫活性的多肽的方法,该方法包括在编码该多肽的核酸分子被表达的条件下培养宿主细胞。

[0106] 细菌细胞也是用于表达本发明的核酸的宿主。在一个实施例中,使用了能够在植物组织内生活并复制的非致病的共生细菌(所谓的内生菌),或能够定居在叶际或根际的非致病的共生细菌(所谓的附生菌)。此类细菌包括以下属的细菌:农杆菌属、产碱杆菌属、固氮螺菌属、固氮菌属、芽孢杆菌属、棒形杆菌属、肠杆菌属、欧文氏菌属、黄杆菌属、克雷伯菌属、假单胞菌属、根瘤菌属、沙雷氏菌属、链霉菌属以及黄单胞菌属。共生的真菌(如木霉属以及胶枝霉属)也是为了相同目的表达本发明的核酸的可能的宿主。

[0107] 这些基因操作技术对于不同的可供使用的宿主而言是特异性的并且在本领域中是已知的。例如,表达载体pKK223-3以及pKK223-2可以用来在大肠杆菌中(在转录或翻译融合中)在tac或trc启动子之后表达异源基因。为了表达编码多个ORF的操纵子,最简单的方法是在转录融合中将该操纵子插入载体(如pKK223-3)中,允许对该异源基因的同源核糖体结合位点进行利用。在革兰氏阳性物种(如芽孢杆菌属)中的过表达技术在本领域中也是已知的,并且可在本发明的上下文中使用(Quax等人,在:Industrial Microorganisms:Basic and Applied Molecular Genetics[工业微生物:基础和应用分子遗传学]中,编辑Baltz等人,American Society for Microbiology[美国微生物学会],华盛顿(1993))。用于过表达的替代性系统依赖于例如酵母载体,并包括毕赤酵母属、酵母属、以及克鲁维酵母属的使用(Sreekrishna,在:Industrial microorganisms:basic and applied molecular genetics[工业微生物:基础和应用分子遗传学]中,Baltz,Hegeman,和Skatrud编辑,American Society for Microbiology[美国微生物学会],华盛顿(1993);Dequin和Barre,Biotechnology[生物科技]L2:173-177(1994);van den Berg等人,Biotechnology[生物科技]8:135-139(1990))。

[0108] 在又其他实施例中,本发明涵盖了控制昆虫有害生物的方法,该方法包括将本发明的控昆虫有效量的杀昆虫蛋白递送至昆虫有害生物。在这些实施例的一些方面中,所述杀昆虫蛋白通过转基因植物或通过局部施用包含所述杀昆虫蛋白的杀昆虫组合物递送。在其他方面,所述转基因植物或所述杀昆虫组合物包含不同于所述杀昆虫蛋白的第二杀昆虫剂。在仍其他方面,第二杀昆虫剂是蛋白质、dsRNA或化学品。在仍其他方面,蛋白质选自以下组成的组:Cry蛋白、Vip蛋白、马铃薯糖蛋白、蛋白酶、蛋白酶抑制剂、脲酶、 α -淀粉酶抑制剂、成孔蛋白、凝集素、工程化抗体或抗体片段、或几丁质酶;或者化学品是氨基甲酸酯、拟除虫菊酯、有机磷酸酯、fiprole、新烟碱、有机氯化物、沙蚕毒素或其组合;或者化学品包含选自下组的活性成分,该组由以下组成:克百威、胺甲萘、灭多虫、联苯菊酯、七氟菊酯、氯菊酯、氟氯氰菊酯、 λ -氯氟氰菊酯、氯氰菊酯、溴氰菊酯、毒死蜱、氯氧磷、乐果、灭线磷、马拉硫磷、甲基对硫磷、甲拌磷、特丁磷、叔丁噻啉磷、氟虫腈、啉虫脒、吡虫啉、噻虫啉、噻虫嗪、硫丹、杀虫磺及其组合。

[0109] 在本发明的一些实施例中,本发明的杀昆虫蛋白中的至少一种在较高级的生物

(如植物)中表达。表达控昆虫有效量的杀昆虫蛋白的转基因植物保护其本身免受昆虫有害生物损害。在昆虫有害生物开始以这种转基因植物为食时,它也摄取了这种已表达的杀昆虫蛋白。这可以阻止昆虫进一步咬食植物组织和/或甚至可以伤害或杀死昆虫。将本发明的核酸分子插入到表达盒中,然后该核酸可被稳定地整合到该植物的基因组中。在其他实施例中,核酸分子包括在非致病的自我复制的病毒中。根据本发明所转化的植物可以是单子叶植物或双子叶植物,并且包括但不限于:玉米、小麦、燕麦、草坪草、牧场草、亚麻、大麦、黑麦、甘薯、豆、豌豆、菊苣、莴苣、卷心菜、花椰菜、西兰花、芜菁、小萝卜、菠菜、芦笋、洋葱、大蒜、胡椒、芹菜、小南瓜(squash)、南瓜(pumpkin)、大麻、绿皮西葫芦、苹果、梨、温柏、甜瓜、李子、樱桃、桃、油桃、杏、草莓、葡萄、覆盆子、黑莓、菠萝、鳄梨、番木瓜、芒果、香蕉、大豆、番茄、高粱、甘蔗、甜菜、向日葵、油菜籽、三叶草、烟草、胡萝卜、棉花、苜蓿、稻、马铃薯、茄子、黄瓜、拟南芥属物种,以及木本植物,如针叶树以及落叶树。

[0110] 在一些实施例中,本发明涵盖生产对昆虫有害生物有毒的蛋白质,即杀昆虫蛋白的方法,所述方法包括:(a)获得包含基因的宿主细胞,该基因自身包含本发明的表达盒和/或核酸分子;和(b)在所述转基因宿主细胞产生所述对昆虫有害生物有毒的蛋白质的条件下使所述宿主细胞或包含所述宿主细胞的转基因宿主生长。

[0111] 在其他实施例中,本发明涵盖产生与对照植物或植物部分相比具有增强的昆虫抗性的转基因植物或植物部分的方法,所述方法包括:(a)将包含编码本发明的杀昆虫蛋白的核酸分子的嵌合基因或表达盒或载体引入植物或植物部分,其中所述杀昆虫蛋白在所述植物或植物部分中表达,从而产生具有增强的昆虫抗性的植物或植物部分。在这些实施例的一些方面中,嵌合基因、表达盒或载体可编码蛋白,所述蛋白质包含以下、基本上由以下组成或由以下组成:(b)与SEQ ID NO:1-22中的任一个至少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%或100%相同或相似的氨基酸序列。在其他方面,表达盒可以编码包含以下氨基酸序列的多肽,该氨基酸序列与SEQ ID NO:4具有至少80%同一性。“增强的”昆虫抗性可以用转基因植物对以转基因植物为食的昆虫有害生物的任何毒性作用来衡量。增强的昆虫抗性与不表达所述杀昆虫蛋白的对照植物相比可高出0%、至少1%、至少2%、至少3%、至少4%、至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少100%、至少125%、至少150%、至少200%、至少300%、至少400%、至少500%、至少600%、至少700%、至少800%、至少900%、或至少1000%的杀昆虫活性。可通过植物转化、植物组织培养或育种的方法产生与对照植物或植物部分相比具有增强的昆虫抗性的植物或植物部分。可通过有性 or 无性繁殖的方法来产生植物或植物部分。可以使用任何合适的对照植物或植物部分,例如在相同环境下生长的、具有相同或相似遗传背景的植物。在实施例中,对照植物或植物部分与所描述的植物具有相同的遗传背景并在相同的环境中生长,但其不包含本发明的分子,而所描述的植物包含本发明的核酸分子。

[0112] 在其他实施例中,本发明涵盖了使植物或植物部分的昆虫抗性与对照植物或植物部分相比增强的方法,该方法包括在该植物或植物部分中表达本发明的核酸分子或表达盒,其中该表达盒的异源核酸的表达导致植物或植物部分与对照植物或植物部分相比具有增强的昆虫抗性。在实施例中,表达盒或核酸分子包含可操作地连接至包含核苷酸序列的

异源核酸分子的启动子,所述核苷酸序列包含以下、基本上由以下组成或由以下组成:(a) SEQ ID NO:23-54中的任一个的核苷酸序列;(b)与SEQ ID NO:23-54中的任一个的核苷酸序列至少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%或100%相同的核苷酸序列;(c)编码蛋白质的核苷酸序列,其中所述蛋白质的氨基酸序列包含以下、基本上由以下组成或由以下组成:SEQ ID NO:1-22中的任一个,并且具有昆虫控制活性;(d)编码蛋白质的核苷酸序列,其中所述蛋白质的氨基酸序列与SEQ ID NO:1-22中的任一个的氨基酸序列至少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%或100%相同;(e)上述(a)至(d)中任一个的核苷酸序列,其经密码子优化以在转基因宿主生物中表达;或(f)与上述(a)至(e)中的任一个的核苷酸序列互补的核苷酸序列。可将核酸分子或表达盒引入植物。在一些实施例中,可将核酸分子或表达盒引入植物部分中,并且可从该植物部分产生包含核酸分子或表达盒的植物。

[0113] 在一些实施例中,本发明涵盖了产生与对照植物相比具有增强的昆虫抗性的植物的方法,该方法包括在植物部分中检测包含本发明的核酸分子或表达盒的异源核酸,以及从该植物部分产生植物,从而产生与对照植物相比具有增强的昆虫抗性的植物。在一个另外的实施例中,本发明涵盖了鉴定与对照植物或植物部分相比具有增强的昆虫抗性的植物或植物部分的方法,该方法包括在该植物或植物部分中检测本发明的核酸分子或表达盒,从而鉴定具有增强的昆虫抗性的植物或植物部分。在一个另外的实施例中,将该表达盒或其诊断片段在来自该植物或植物部分的核酸样品中的扩增产物中进行检测。诊断片段可以是至少10个连续核苷酸长的核酸分子,其对于本发明的表达盒是独特的。

[0114] 在其他实施例中,本发明涵盖了产生与对照植物或植物部分相比具有增强的昆虫抗性的植物的方法,该方法包括将第一亲本植物与第二亲本植物进行杂交,其中至少该第一亲本植物在其基因组中包含含有本发明的核酸分子或表达盒的异源核酸;以及产生子代后代,其中该子代后代包含至少一种在其基因组中具有所述异源核酸并且与对照植物相比展现出增强的昆虫抗性的植物。

[0115] 在上述实施例的一些方面中,本发明的方法赋予植物或植物部分针对鞘翅目昆虫有害生物的增强的昆虫抗性。在实例中证明了对鞘翅目昆虫有害生物的昆虫控制。在另外的方面中,本发明的方法赋予植物或植物部分针对根萤叶甲属物种(包括玉米根萤叶甲、巴氏根萤叶甲、黄瓜十一星叶甲食根亚种、墨西哥玉米根萤叶甲和/或南美叶甲(*Diabrotica speciosa*))和/或相关物种的增强的昆虫抗性。在另外的实施例中,本发明的方法赋予植物或植物部分针对西方玉米根萤叶甲、巴氏根萤叶甲和/或黄瓜十一星叶甲食根亚种的增强的昆虫抗性。

[0116] 在一些实施例中,本发明涵盖了转基因植物,所述转基因植物包含本发明的异源核酸分子或表达盒,所述异源核酸分子或表达盒在被转录或被翻译时向所述转基因植物赋予增强的昆虫抗性。在这些实施例的一些方面中,异源核酸分子或表达盒包含与SEQ ID NO:23-54中的任一个至少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%或100%相同的核苷酸序列。在其他方面,转基因植物包含异源核酸分子或表达盒,所述异源核酸分子或表达盒包含与SEQ ID NO:4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%或100%相同的核苷酸序列。在这些实施例的其他方面,转基因植物是双子叶植物或单子叶植物。在另外的方面中,转基因植物是苜蓿、小茴香、苹果、杏、朝鲜蓟、芝麻菜、芦笋、鳄梨、香蕉、豆类、甜菜、黑莓、蓝莓、西蓝花、抱子甘蓝、卷心菜、卡诺拉、罗马甜瓜、胡萝卜、木薯、花椰菜、芹菜、樱桃、芫荽、柑橘、克莱门氏小柑橘、咖啡豆、玉米、棉花、黄瓜、花旗松、茄子、苦苣、茅菜、桉树、茴香、无花果、葫芦、葡萄、葡萄柚、白兰瓜、豆薯、奇异果、莴苣、韭菜、柠檬、酸橙、火炬松、芒果、甜瓜、蘑菇、坚果、秋葵、洋葱、橙子、观赏植物、木瓜、欧芹、豌豆、桃、花生、梨、胡椒、柿子、松树、菠萝、车前草、李、石榴、杨树、马铃薯、南瓜、温柏、辐射松、红菊苣、小萝卜、覆盆子、稻、黑麦、高粱、南方松、大豆、菠菜、小南瓜、草莓、甜菜、向日葵、甘薯、枫香、柑桔、茶、烟草、番茄、草皮、藤、西瓜、洋芋或绿皮西葫芦。在仍其他方面,转基因植物是粟、柳枝稷、玉蜀黍、高粱、小麦、燕麦、草坪草、牧场草、亚麻、稻、甘蔗、油菜或大麦。

[0117] 在一些实施例中,本发明涵盖编码本发明的杀昆虫蛋白的核酸分子,所述杀昆虫蛋白经修饰和优化以在转基因植物中表达。尽管在许多情况中来自微生物有机体的基因可以在没有修饰的情况下在植物中以高水平表达,但在转基因植物中的低表达可能是由于具有在植物中不优选的密码子的微生物核酸所造成的。在本领域中已知的是,所有生物针对密码子使用具有特定偏好,并且可以对本发明中所描述的核酸的密码子进行改变以符合植物的偏好,同时保持由此所编码的氨基酸,或对编码的杀昆虫蛋白进行某些氨基酸改变。此外,在植物中高表达最好是由以下编码序列实现的,这些编码序列具有至少大约35%、优选大于约45%、更优选大于约50%、并且最优选大于约60%的GC含量。具有低GC含量的微生物核酸可能在植物中表达欠佳,这是由于存在可使信息不稳定的ATTGA基序,以及可能引起不适当的聚腺苷酸化作用的AATAAA基序。在实施例中,可以对序列进行修饰以便迎合单子叶植物或双子叶植物的特定密码子偏好以及GC含量偏好,因为这些偏好已经被证明是不同的(Murray等人Nucl. Acids Res. [核酸研究]17:477-498(1989))。此外,针对可能引起信息缩短的不合理剪接位点的存在对这些核酸筛选。可以使用熟知的位点定向诱变、PCR、以及合成基因的构建,例如,使用公开的专利申请EP 0 385 962、EP 0 359 472、以及WO 93/07278中所述的方法,对在这些核酸之内所有需要做出的变化(如以上所描述的那些变化)进行改变。

[0118] 在本发明的一些实施例中根据在美国专利5,625,136(通过引用并入本文)中所披露的程序来制造本发明的杀昆虫蛋白的编码序列。在这个程序中,使用了玉蜀黍优选的密码子,即最频繁地编码玉蜀黍中的氨基酸的单个密码子。针对特定的氨基酸的玉蜀黍优选的密码子可源自(例如)来自玉蜀黍的已知基因序列。针对来自玉蜀黍植物的28个基因的玉蜀黍密码子使用发现于Murray等人,Nucleic Acids Research[核酸序列]17:477-498(1989)中,其披露内容通过引用并入本文。以这种方式,这些核苷酸序列可以进行优化用于在任何植物中表达。认识到的是该基因序列的所有或任何部分可以是优化的或合成的。即,还可以使用合成的或部分优化的序列。

[0119] 为了更加有效的翻译起始,可修饰与起始甲硫氨酸相邻的序列。例如,它们可以通过包含已知在植物中有效的序列而被修饰。Joshi已经提出了针对植物的适当的共有序列(NAR 15:6643-6653(1987)),并且Clontech提出了另一种共有翻译起始子(1993/1994目录,第210页)。这些共有序列适合与本发明的核酸一起使用。在实施例中,将这些序列并入

包含核酸的构建体中,达到并且包括ATG(而未对第二个氨基酸进行修饰),或者可替代地达到并且包括在ATG后的GTC(具有修饰该转基因的第二氨基酸的可能性)。

[0120] 在转基因植物中这些核酸的表达是由在植物中发挥作用的启动子所驱动的。启动子的选择将根据表达的时间和空间需要而变化,并且还根据靶物种而变化。因此,本发明的核酸在叶、柄(stalk)或茎(stem)、穗、花序(例如穗状花序、圆锥花序、穗轴等等)、根、和/或籽苗中的表达是优选的。然而在许多情况下,寻求针对多于一种类型昆虫有害生物的保护,并且因此在多个组织中的表达是令人希望的。尽管已经显示来自双子叶植物的很多启动子在单子叶植物中是可操作的并且反之亦然,但理想的是选择双子叶植物启动子用于在双子叶植物中表达,并且选择单子叶植物启动子用于在单子叶植物中表达。不过,对于所选的启动子的起源没有限制;只要它们可有效驱动核酸在所希望的细胞中表达就足够了。

[0121] 在一些实施例中,使用了以组成型表达的启动子,包括肌动蛋白或泛素或cnp启动子,或CaMV 35S和19S启动子。本发明的核酸也可以在用化学方法进行调节的启动子的调节下表达。用于基因表达的化学诱导的优选的技术详述于公开申请EP 0 332 104(汽巴-嘉基公司(Ciba-Geigy))以及美国专利5,614,395中。用于化学诱导的优选的启动子是烟草PR-1a启动子。

[0122] 在其他实施例中,可使用一类伤口诱导型启动子。已经描述了数量众多的在创伤部位并且还在植物病原菌感染的部位表达的启动子。理想的是,这种启动子应该仅在感染的部位有局部活性,并且以这种方式,本发明的杀昆虫蛋白仅在需要合成这些蛋白质的细胞中累积以杀死入侵的昆虫有害生物。这类优选的启动子包括由以下文献所描述的那些:Stanford等人Mol.Gen.Genet.[分子遗传学]215:200-208(1989),Xu等人Plant Molec.Biol.[植物分子生物学]22:573-588(1993),Logemann等人Plant Cell[植物细胞]1:151-158(1989),Rohrmeier和Lehle,Plant Molec.Biol.[植物分子生物学]22:783-792(1993),Firek等人Plant Molec.Biol.[植物分子生物学]22:129-142(1993),以及Warner等人Plant J.[植物杂志]3:191-201(1993)。

[0123] 用于在植物(特别是玉米)中表达对本发明的杀昆虫蛋白进行编码的基因的组织特异性的或组织优选的启动子是那些直接在根、髓、叶或花粉(特别是根)中表达的启动子。此类启动子,例如从PEPC或trpA中分离的那些披露于美国专利号5,625,136中,或从MTL中分离的那些披露于美国专利号5,466,785中。这两篇美国专利以其全文通过引用并入本文。

[0124] 此外,可以使用在质体中发挥作用的启动子。此类启动子的非限制性实例包括噬菌体T3基因9'5'UTR以及其他的披露于美国专利号7,579,516中的启动子。适用于本发明的其他启动子包括但不限于S-E9小亚基RuBP羧化酶启动子和Kunitz胰蛋白酶抑制剂基因启动子(Kti3)。

[0125] 在本发明的一些实施例中,可以使用诱导型启动子。因此,例如,可以使用化学调节型启动子以通过应用外源化学调节物来调节本发明的核苷酸序列的表达。本发明的核苷酸序列的表达经由化学调节过的启动子进行的调节使得本发明的多肽仅当用诱导的化学物处理作物植物时能被合成。取决于目的,在应用化学品诱导本发明的核苷酸序列的表达时,该启动子可以是化学诱导型启动子,或者在应用化学品抑制本发明的核苷酸序列表达时该启动子可以是化学阻抑型启动子。

[0126] 化学诱导型启动子在本领域是已知的,并且包括但不限于玉蜀黍In2-2启动子(其

由苯磺酰胺除草剂安全剂激活)、玉蜀黍GST启动子(其由用作发芽前除草剂的疏水亲电子化合物激活)、以及烟草PR-1a启动子(其由水杨酸激活)(例如,PR1a系统)、类固醇反应性启动子(参见例如,Schena等人(1991)Proc.Natl.Acad.Sci.[美国国家科学院院刊]USA 88, 10421-10425和McNellis等人(1998)Plant J.[植物杂志]14,247-257中的糖皮质激素诱导型启动子)以及四环素诱导型启动子和四环素-阻抑型启动子(参见例如,Gatz等人(1991)Mol.Gen.Genet.[分子遗传学]227,229-237和美国专利号5,814,618和5,789,156)、Lac阻遏物系统启动子、铜诱导型系统启动子、水杨酸诱导型系统启动子(例如,PR1a系统)、糖皮质激素诱导型启动子(Aoyama等人(1997)Plant J.[植物杂志]11:605-612)以及蜕皮激素诱导型系统启动子。

[0127] 诱导型启动子的其他非限制性实例包括ABA诱导型和细胞膨胀诱导型启动子、植物生长素结合蛋白基因启动子(Schwob等人(1993)Plant J.[植物杂志]4:423-432)、UDP葡萄糖类黄酮糖基转移酶启动子(Ralston等人(1988)Genetics[遗传]119:185-197)、MPI蛋白酶抑制剂启动子(Cordero等人(1994)Plant J.[植物杂志]6:141-150)以及甘油醛-3-磷酸脱氢酶启动子(Kohler等人(1995)Plant Mol.Biol.[植物分子生物学]29:1293-1298; Martinez等人(1989)J.Mol.Biol.[分子生物学杂志]208:551-565;和Quigley等人(1989)J.Mol.Evol.[分子进化杂志]29:412-421)。还包括苯磺酰胺诱导型(美国专利号5,364,780)和乙醇诱导型(国际专利申请公开号WO 97/06269和WO 97/06268)系统和谷胱甘肽S-转移酶启动子。同样地,可以使用描述于以下文献中的诱导型启动子中的任一种:Gatz(1996)Current Opinion Biotechnol.[生物技术新见]7:168-172和Gatz(1997)Annu.Rev.Plant Physiol.Plant Mol.Biol.[植物生理学与植物分子生物学年度综述]48:89-108。适用于指导本发明的核苷酸序列在植物中的表达的其他化学诱导型启动子披露于美国专利5,614,395中,该专利通过引用以其全文并入本文。基因表达的化学诱导还详述于公开申请EP 0 332 104(授予汽巴-嘉基公司(Ciba-Geigy))和美国专利5,614,395中。在一些实施例中,用于化学诱导的启动子可以是烟草PR-1a启动子。

[0128] 在进一步的实施例中,本发明的核苷酸序列可以与启动子可操作地相关联,该启动子是创伤诱导型或由有害生物或病原体侵染(例如,昆虫或线虫植物有害生物)进行的诱导型。已经描述了在创伤部位和/或在有害生物攻击(例如,昆虫/线虫摄食)或植物病原菌侵染的部位处表达的数量众多的启动子。理想的是,这种启动子应仅在攻击部位处或邻近该部位具有局部活性,并且以这种方式,本发明的核苷酸序列的表达将集中在所侵入或摄食的细胞中。此类启动子包括但不限于由以下描述的那些启动子:Stanford等人,Mol.Gen.Genet.[分子与普通遗传学]215:200-208(1989);Xu等人,Plant Molec.Biol.[植物分子生物学]22:573-588(1993);Logemann等人,Plant Cell[植物细胞]1:151-158(1989);Rohrmeier和Lehle,Plant Molec.Biol.[植物分子生物学]22:783-792(1993);Firek等人,Plant Molec.Biol.[植物分子生物学]22:129-142(1993);Warner等人,Plant J.[植物杂志]3:191-201(1993);美国专利号5,750,386;美国专利号5,955,646;美国专利号6,262,344;美国专利号6,395,963;美国专利号6,703,541;美国专利号7,078,589;美国专利号7,196,247;美国专利号7,223,901;以及美国专利申请公开2010043102。

[0129] 在本发明的一些实施例中,使用“最小启动子”或“基本启动子”。最小启动子能够募集并结合RNA聚合酶II复合体及其辅助蛋白,以允许转录起始和延伸。在一些实施例中,

最小启动子被构建成仅包含来自转录因子的结合和目的核苷酸序列的转录所必需的选定启动子的核苷酸/核苷酸序列的启动子,这一目的核苷酸序列可操作地与包括但不限于TATA盒序列的最小启动子相关联。在其他实施例中,最小启动子缺少募集并结合转录因子的cis序列,这些转录因子调节(例如,增强、阻抑、赋予组织特异性,赋予诱导性或阻抑性)转录。最小启动子通常被放置在待表达的核苷酸序列的上游(即,5')。因此,可以选择来自与本发明一起可用的任何启动子的核苷酸/核苷酸序列用作最小启动子。

[0130] 可以将许多其他序列掺入本发明所描述的表达盒中。这些序列包括已经显示出增强表达的序列,如内含子序列(例如,来自Adh1和bronze1)以及病毒的前导序列(例如,来自TMV、MCMV、和AMV)。

[0131] 本发明的核酸在植物中针对不同的细胞定位的靶向表达可能是更优选的。在一些情况下,在胞质溶胶中的定位可能是令人希望的,而在其他情况下,在某个亚细胞器中的定位可能是优选的。使用本领域内熟知的技术进行编码酶类的转基因的亚细胞定位。典型地,对编码来自已知细胞器靶向的基因产物的靶肽的DNA进行操作并将其融合在该核酸的上游。针对叶绿体的许多此类靶序列是已知的并且已经证明了它们在异源构建体中的功能。还将本发明的核酸的表达靶向至宿主细胞的内质网或液泡。实现其的技术在本领域中是熟知的。

[0132] 适合于植物转化的载体描述于在本说明书的其他地方。对于农杆菌介导的转化,二元载体或携带至少一种T-DNA边界序列的载体是合适的,而对于直接基因转移,任何载体都是合适的,并且仅含有目的构建体的线性DNA也许是优选的。在直接基因转移的情况下,可以使用以单个DNA种类的转化或共转化(Schocher等人,Biotechnology[生物技术]4:1093-1096(1986))。对于直接基因转移以及农杆菌介导的转化这两者,转化通常(但不是必需的)用选择性标记进行,这种选择性标记可以提供针对抗生素(卡那霉素、潮霉素、或甲氨蝶呤)或除草剂(basta)的抗性。包含本发明的核酸分子的植物转化载体还可包含以下基因(例如磷酸甘露糖异构酶;PMI),这些基因提供转基因植物的阳性选择,如美国专利5,767,378和5,994,629(通过引用并入本文)中所披露的。然而,选择性标记的选择对于本发明并不是至关重要的。

[0133] 在一些实施例中,核酸可以转化到核基因组中。在另一个实施例中,本发明的核酸被直接转化到质体基因组中。质体转化的主要优点在于质体通常能够表达细菌基因而无需实质性的密码子优化,而且质体能够在单一启动子的控制下表达多个开放阅读框。在美国专利号5,451,513、5,545,817和5,545,818中,在PCT申请号WO 95/16783中,以及在McBride等人,(1994),Proc.Nati.Acad.Sci.USA[美国国家科学院院刊]91,7301-7305中广泛描述了质体转化技术。基本的叶绿体转化技术涉及例如使用生物射弹(biolistic)或原生质体转化(例如,氯化钙或PEG介导的转化),将位于选择性标记侧翼的经克隆的质体DNA区连同目的基因一起引入合适的靶组织中。这些1至1.5kb的侧翼区(被命名为靶向序列)促进了与质体基因组的同源重组,并且因而允许置换或修饰原质体(plastome)的特定区域。最初,将叶绿体16S rRNA和rps12基因的点突变(赋予对壮观霉素和/或链霉素抗性)用作用于转化的选择性标记(Svab,Z.,Hajdukiewicz,P.,和Maliga,P.(1990)Proc.Nati.Acad.Sci.USA[美国国家科学院院刊]87,8526-8530;Staub,J.M.,和Maliga,P.(1992)Plant Cell[植物细胞]4,39-45)。这以大约每100次靶叶轰击大约1个的频率产生稳定的同质转化株。在这些

标记之间克隆位点的存在允许建立质体靶向载体用于外来基因的引入 (Staub, J.M. 和 Maliga, P. (1993) EMBO J. [欧洲分子生物学杂志] 12, 601-606)。转化频率的实质性增加是通过用显性选择性标记 (细菌 *aadA* 基因, 其编码了壮观霉素-解毒酶氨基糖苷-3'-腺苷酰转移酶) 置换隐性 *rRNA* 或 *r*-蛋白抗生素抗性基因来获得的 (Svab, Z., 和 Maliga, P. (1993) Proc. Natl. Acad. Sci. USA [美国国家科学院院刊] 90, 913-917)。先前, 这种标记已经被成功地用于莱茵衣藻这种绿藻的质体基因组的高频率转化 (Goldschmidt-Clermont, M. (1991) Nucl. Acids Res. [核酸研究] 19:4083-4089)。有用于质体转化的其他选择性标记在本领域是已知的, 并且被涵盖在本发明的范围之内。典型地, 转化之后需要大约 15-20 个细胞分裂循环以便达到同质状态。质体表达 (其中基因通过同源重组被插入到在每个植物细胞中存在的所有数千个环状质体基因组的拷贝中) 利用了超过核表达的基因的庞大的拷贝数目的优点, 以便允许能够很容易超过总的可溶性植物蛋白的 10% 的表达水平。在一个优选的实施例中, 将本发明的核酸插入到质体靶向的载体中并且转化到所希望的植物宿主的质体基因组中。获得了对于包含本发明的核酸的质体基因组同型的植物, 并且这些植物能够优先地高表达该核酸。

[0134] 在又其他实施例中, 本发明的转基因植物可包含异源核酸分子, 该异源核酸分子编码至少一种另外的所希望的性状。另外的性状可以在与本发明的核酸分子相同的异源核酸分子上编码, 或者可以在第二异源核酸分子上编码。另外的所希望的性状可赋予对第二昆虫有害生物的昆虫抗性、对相同昆虫有害生物的昆虫抗性、非生物胁迫耐受性、雄性不育、除草剂抗性、细菌疾病抗性、真菌疾病抗性、病毒疾病抗性、线虫抗性、改性脂肪酸代谢、改性碳水化合物代谢、改善的营养价值、工业过程中改善的性能或改变的繁殖能力。另外的所希望的性状还可诱导商业上有价值的酶或代谢物在植物内产生。

[0135] 在一些实施例中, 所希望的另外的性状是第二杀有害生物剂。第二杀有害生物剂可对任何植物有害生物 (包括昆虫、线虫、真菌、病毒或细菌) 具有活性。昆虫植物有害生物的实例包括但不限于: 褐飞虱属物种 (*Nilaparvata* spp.) (例如褐飞虱 (*N. lugens*) (褐色飞虱 (brown planthopper))) ; 灰飞虱属物种 (*Laodelphax* spp.) (例如灰飞虱 (*L. striatellus*) (褐色小飞虱 (small brown planthopper))) ; 黑尾叶蝉属物种 (*Nephotettix* spp.) (例如二点黑尾叶蝉 (*N. virescens*) 或黑尾叶蝉 (*N. cincticeps*) (青叶蝉 (green leafhopper))) , 或二条黑尾叶蝉 (*N. nigropictus*) (稻叶蝉 (rice leafhopper))) ; 白背飞虱属物种 (*Sogatella* spp.) (例如白背飞虱 (*S. furcifera*, white-backed planthopper)) ; 土长蝽属物种 (*Blissus* spp.) (例如麦长蝽 (*B. leucopterus leucopterus*) (麦小蝽 (chinch bug))) ; 黑蝽属物种 (*Scotinophora* spp.) (例如稻黑蝽 (*S. vermidulate*, rice blackbug)) ; 拟缘蝽属物种 (*Acrosternum* spp.) (例如拟绿蝽 (*A. hilare*) (绿椿象 (green stink bug))) ; 稻弄蝶属物种 (*Parnara* spp.) (例如直纹稻弄蝶 (*P. guttata*, rice skipper)) ; 禾草螟属物种 (*Chilo* spp.) (例如二化螟 (*C. suppressalis*) (水稻二化螟 (rice striped stem borer))、台湾稻螟 (*C. auricilius*) (金线二化螟 (gold-fringed stem borer))、或稻多丽螟 (*C. polychrysus*) (黑头茎螟 (dark-headed stem borer))) ; 稻螟属物种 (*Chilo traea* spp.) (例如稻多丽螟 (*C. polychrysa*) (稻茎螟 (rice stalk borer))) ; 蛀茎夜蛾属物种 (*Sesamia* spp.) (例如大螟 (*S. inferens*) (粉红色稻螟 (pink rice borer))) ; 蛀禾螟属物种 (*Tryporyza* spp.) (例

如稻白螟(*T.innotata*) (白稻螟(*white rice borer*))、或三化螟(*T.incertulas*) (黄稻螟(*yellow rice borer*))) ;稻纵卷叶螟属物种(*Cnaphalocrocis spp.*) (例如稻纵卷叶螟(*C.medinalis, rice leafroller*))) ;潜蝇属物种(*Agromyza spp.*) (例如日本稻潜蝇(*A.oryzae*) (潜叶虫(*leafminer*))、或美洲黍潜叶蝇(*A.parvicornis*) (玉米污点斑潜蝇(*corn blot leafminer*))) ;蔗螟属物种(例如小蔗螟(甘蔗蛀虫)、或西南玉米螟(*D.grandiosella, southwestern corn borer*)) ;*Narnaga*属物种(例如稻螟蛉(*N.aenescens, green rice caterpillar*)) ;黄夜蛾属物种(*Xanthodes spp.*) (例如犁纹黄夜蛾(*X.transversa, green caterpillar*)) ;灰翅夜蛾属物种(例如草地贪夜蛾(*S.frugiperda*) (秋夜蛾(*fall armyworm*))、甜菜夜蛾(*S.exigua, beet armyworm*)、海灰翅夜蛾(*S.littoralis, climbing cutworm*)或西部黄条粘虫(*S.praefica, western yellowstriped armyworm*)) ;光腹夜蛾属物种(*Mythimna spp.*) (例如粘虫(*Mythimna seperata, Pseudaletia seperata*)) ;铃夜蛾属物种(例如玉米穗虫(玉米穗蛾)) ;*Colaspis*属物种(例如葡萄肖叶甲(*C.brunnea, grape colaspis*)) ;稻水象属物种(*Lissorhoptrus spp.*) (例如稻水象虫(*L.oryzophilus*) (稻水象鼻虫(*rice water weevil*)) ;稻象甲属物种(*Echinocnemus spp.*) (例如稻鳞象虫(*E.squamos*) (稻象鼻虫(*rice plant weevil*))) ;稻铁甲属物种(*Diclodispa spp.*) (例如水稻铁甲(*D.armigera, rice hispa*)) ;负泥虫属物种(*Oulema spp.*) (例如水稻负泥虫(*O.oryzae* (稻叶泥虫(*leaf beetle*))) ;米象属物种(*Sitophilus spp.*) (例如米象(*S.oryzae, rice weevil*)) ;稻瘿蚊属物种(*Pachydiplosis spp.*) (例如稻瘿蚊(*P.oryzae, rice gall midge*)) ;毛眼水蝇属物种(*Hydrellia spp.*) (例如麦叶毛眼水蝇(*H.griseola*) (小米潜叶虫(*small rice leafminer*))、或稻茎毛眼水蝇(*H.sasakii, rice stem maggot*)) ;黄潜蝇属物种(*Chlorops spp.*) (例如秆蝇(*C.oryzae, stem maggot*)) ;根萤叶甲属物种(例如玉米根萤叶(西方玉米根虫)、巴氏根萤叶甲(北方玉米根虫)、黄瓜十一星叶甲食根亚种(南方玉米根虫)、墨西哥玉米根萤叶甲(墨西哥玉米根虫) ;黄瓜条根萤叶甲(带状黄瓜甲虫) ;秆野螟属物种(*Ostrinia spp.*) (例如欧洲玉米螟(*O.nubilalis, European corn borer*)) ;地老虎属物种(例如小地老虎(黑色地老虎)) ;玉米苗斑螟属物种(例如南美玉米苗斑螟(*E.Lignosellus*)) ;梳爪叩头虫属物种(铁线虫) ;方头甲属物种(*Cyclocephala spp.*) (例如北方圆头犀金龟(*C.borealis, northern masked chafer*)或南方圆头犀金龟(*C.immaculata, southern masked chafer*)) ;弧丽金龟属物种(*Popillia spp.*) (例如日本甲虫(*P.japonica, Japanese beetle*)) ;凹胫跳甲属物种(*Chaetocnema spp.*) (例如玉米跳甲(*C.pulicaria, corn flea beetle*)) ;尖隐喙象属物种(*Sphenophorus spp.*) (例如玉米谷象(*S.maidis, maize billbug*)) ;缢管蚜属物种(*Rhopalosiphum spp.*) (例如玉米蚜(*R.maidis*) (玉米叶蚜虫(*corn leaf aphid*))) ;圆尾蚜属物种(*Anuraphis spp.*) (例如玉米根蚜(*A.maidiradicis, corn root aphid*)) ;黑蝗属物种(*Melanoplus spp.*) (例如赤胫黑蝗(*M.femurrubrum*) (红胫蝗虫(*redlegged grasshopper*))、异黑蝗(*M.differentialis, differential grasshopper*)或血黑蝗(*M.sanguinipes*) (迁徙蝗虫(*migratory grasshopper*)) ;种蝇属物种(*Hylemya spp.*) (例如玉米种蝇(*H.platura, seedcorn maggot*)) ;呆蓟马属物种(*Anaphothrips spp.*) (例如草蓟马(*A.obscrurus, grass thrips*)) ;水蚁属物种(*Solenopsis spp.*) (例如窃叶蚁(*S.milesta, thief ant*)) ;四爪螨属物种(*Tetranychus spp.*) (例如二斑叶螨(*T.urticae,*

twospotted spider mite)、朱砂叶螨(*T.cinnabarinus*, carmine spider mite);铃夜蛾属物种(例如玉米穗虫(棉铃虫(*cotton bollworm*))、或棉铃虫(*H.armigera*) (美国螟蛉虫(*American bollworm*)));红铃虫属物种(例如棉红铃虫(*P.gossypiella*, pink bollworm);钻夜蛾属物种(*Earias* spp.) (例如翠纹钻夜蛾(*E.vittella*, spotted bollworm));实夜蛾属物种(*Heliothis* spp.) (例如烟芽夜蛾(*H.virescens*) (烟夜蛾(*tobacco budworm*)));花象属物种(例如棉铃象甲(*A.grandis*) (棉铃象虫(*boll weevil*)));*Pseudatomoscelis*属物种(例如棉跳盲蝽(*P.seriatum*, cotton fleahopper));粉虱属物种(*Trialeurodes* spp.) (例如纹翅粉虱(*T.abutiloneus*) (结翅粉虱(*banded-winged whitefly*))、温室白粉虱(*T.vaporariorum*) (温室粉虱(*greenhouse whitefly*)));小粉虱属物种(*Bemisia* spp.) (例如银叶粉虱(*B.argentifolii*) (银叶白粉虱(*silverleaf whitefly*)));蚜虫属物种(*Aphis* spp.) (例如棉蚜(*A.gossypii*) (棉花蚜虫(*cotton aphid*)));草盲蝽属物种(*Lygus* spp.) (例如牧草盲蝽(*L.lineolaris*) (无泽盲蝽(*tarnished plant bug*))或豆荚草盲蝽(*L.hesperus*) (西方无泽盲蝽(*western tarnished plant bug*)));美洲蝽属物种(*Euschistus* spp.) (例如斑点美洲蝽(*E.conspersus*) (具散点的椿象(*conspersus stink bug*)));*Chlorochroa*属物种(例如佐井氏椿象(*C.sayi*) (佐井氏椿象(*Say stinkbug*)));绿椿象属物种(*Nezara* spp.) (例如稻绿椿象(*N.viridula*) (南方绿椿象(*southern green stinkbug*)));蓟马属物种(例如烟蓟马(*T.tabaci*) (葱蓟马(*onion thrips*)));花蓟马属物种(*Frankliniella* spp.) (例如烟褐花蓟马(*F.fusca*) (烟草蓟马(*tobacco thrips*))或西花蓟马(*F.occidentalis*) (西方花蓟马(*western flower thrips*)));叶甲属物种(*Leptinotarsa* spp.) (例如马铃薯叶甲(*L.decemlineata*) (科罗拉多马铃薯甲虫(*Colorado potato beetle*))、伪马铃薯叶甲(*L.juncta*) (伪马铃薯甲虫(*false potato beetle*))或胡颓子叶茄叶甲(*L.texana*) (德克萨斯伪马铃薯甲虫(*Texan false potato beetle*)));*Lema*属物种(例如三线马铃薯甲虫(*L.trilineata*, three-lined potato beetle));毛跳甲属物种(*Epitrix* spp.) (例如马铃薯跳甲(*E.cucumeris*, potato flea beetle), 烟草跳甲(*E.hirtipennis*) (跳甲(*flea beetle*)), 或块茎跳甲(*E.tuberis*, tuber flea beetle));豆芫菁属物种(例如横带芫菁(*E.vittata*, striped blister beetle));猿叶甲属物种(*Phaedon* spp.) (例如辣根猿叶甲(*P.cochleariae*) (芥菜叶甲虫(*mustard leaf beetle*)));食植瓢虫属物种(*Epilachna* spp.) (例如墨西哥豆瓢虫(*E.variverticis*) (墨西哥豆甲虫(*mexican bean beetle*)));蟋蟀属物种(*Acheta* spp.) (例如家蟋蟀(*A.domesticus*) (室内蟋蟀(*house cricket*)));叶蝉属物种(*Empoasca* spp.) (例如马铃薯微叶蝉(*E.fabae*) (马铃薯微叶蝉(*potato leafhopper*)));瘤蚜属物种(*Myzus* spp.) (例如桃蚜(*M.persicae*) (绿桃蚜(*green peach aphid*)));木虱属物种(*Paratrioza* spp.) (例如考氏木虱(*P.cockerelli*) (木虱(*psyllid*)));宽胸叩头虫属物种(*Conoderus* spp.) (例如南方马铃薯金针虫(*C.fallii*, southern potato wireworm) 或马铃薯金针虫(*C.vespertinus*, tobacco wireworm));茄麦蛾属物种(例如马铃薯麦蛾(*P.operculella*, potato tuberworm));长管蚜属物种(*Macrosiphum* spp.) (例如马铃薯长管蚜(*M.euphorbiae*) (马铃薯蚜(*potato aphid*)));Thyanta属物种(例如T.*Pallidovirens* (红肩椿象(*redshouldered stinkbug*)));茄麦蛾属物种(例如马铃薯麦蛾(*P.operculella*, potato tuberworm));铃夜蛾属物种(例如玉米穗虫(柿铃虫(*tomato fruitworm*)));茄茎

麦蛾属物种(例如番茄蠹蛾(番茄蛴虫));*Limonius*属物种(*Limonius* spp.) (铁线虫);天蛾属物种(*Manduca* spp.),例如烟草天蛾(*M.sexta*, tobacco hornworm)和番茄天蛾(*M.quinquemaculata*, tomato hornworm);斑潜蝇属物种(*Liriomyza* spp.) (例如美洲斑潜蝇(*L.sativae*)、三叶斑潜蝇(*L.trifolli*)或南美斑潜蝇(*L.huidobrensis*) (潜叶虫));果蝇属物种(*Drosophilla* spp.) (例如黑腹果蝇(*Drosophila melanogaster*)、*D.yakuba*、拟暗果蝇(*D.Pseudoobscura*)或拟果蝇(*D.simulans*));林大步甲属物种(*Carabus* spp.) (例如粒步甲(*C.granulatus*));摇蚊属物种(*Chironomus* spp.) (例如破伤风摇蚊(*C.tentanus*));栉头蚤属物种(猫蚤(*C.felis*, cat flea));*Diaprepes*属物种(例如根象甲(*D.abbreviatus*, root weevil));*Ips*属物种(例如*I.pini* (pine engraver));拟谷盗属物种(*Tribolium* spp.) (例如赤拟谷盗(*T.innotata*, red floor beetle));舌蝇属物种(*Glossina* spp.) (例如采采蝇(*G.Morsitans*, tsetse fly));疟蚊属物种(*Anopheles* spp.) (例如冈比亚疟蚊(*A.gambiae*) (疟蚊(malaria mosquito)));棉铃虫属物种(例如棉铃虫(*H.armigera*) (非洲螟蛉虫(African Bollworm));无网管蚜属物种(*Acyrtosiphon* spp.) (例如豌豆蚜(*A.pisum*, pea aphid));蚜属物种(*Apis* spp.) (例如意大利蜂(*A.melifera*) (蜜蜂));叶蝉属物种(*Homalodisca* spp.) (例如琉璃叶蝉(*H.coagulate*) (玻璃翅叶蝉(glassy-winged sharpshooter)));伊蚊属物种(*Aedes* spp.) (例如埃及伊蚊(*Ae.aegypti*) (黄热病伊蚊(yellow fever mosquito)));蚕蛾属物种(*Bombyx* spp.) (例如家蚕(*B.mori*) (桑蚕(silkworm)));飞蝗属物种(*Locusta* spp.) (例如非洲飞蝗(*L.migratoria*) (飞蝗(migratory locust)));牛蜱属物种(*Boophilus* spp.) (例如微小牛蜱(*B.microplus*) (牛蜱(cattle tick)));*Acanthoscurria*属物种(例如*A.Gomesiana* (红发巧克力色食鸟蛛(red-haired chocolate bird eater)));折翅属物种(*Diploptera* spp.) (例如太平洋折翅蠊(*D.punctata*, pacific beetle cockroach));袖蝶属物种(*Heliconius* spp.) (例如*H.erato* (红西番莲蛱蝶(red passion flower butterfly))或红带袖蝶(*H.melpomene*) (邮差蝴蝶(postman butterfly)));象虫属物种(*Curculio* spp.) (例如栎实象(*C.glandium*, acorn weevil));菜蛾属物种(例如小菜蛾(*P.xylostella*, diamondback moth));钝眼蜱属物种(*Amblyomma* spp.) (例如牛蜱(*A.variegatum*, cattle tick));*Anteraea*属物种(例如天蚕(*A.Yamamai*) (柞蚕(silkmoth)));和阿蚊属物种(*Armigeres* spp.) (例如骚扰阿蚊(*A.subalbatus*))。

[0136] 本发明的杀昆虫蛋白可以与其他杀有害生物剂(例如Bt Cry蛋白)组合使用以增加有害生物的靶范围。此外,本发明的杀昆虫蛋白与杀昆虫剂(其具有不同的作用方式或靶向昆虫肠中不同的受体)相组合进行使用对于预防和/或管理昆虫抗性而言具有特定的功用。

[0137] 第二杀有害生物剂可以是源自苏云金芽孢杆菌的杀昆虫蛋白。苏云金芽孢杆菌杀昆虫蛋白可以是许多杀昆虫蛋白中的任一种,包括但不限于:Cry1蛋白、Cry3蛋白、Cry7蛋白、Cry8蛋白、Cry11蛋白、Cry22蛋白、Cry23蛋白、Cry36蛋白、Cry37蛋白、Cry34蛋白连同Cry35蛋白、二元杀昆虫蛋白CryET33和CryET34、二元杀昆虫蛋白TIC100和TIC101、二元杀昆虫蛋白PS149B1、VIP(营养期杀昆虫蛋白,披露于美国专利5849870和5877012中,通过引用并入本文)、TIC900或相关蛋白、TIC901、TIC1201、TIC407、TIC417、修饰的Cry3A蛋白、或由前述杀昆虫蛋白中的任一个制成的杂合蛋白或嵌合体。在其他实施例中,苏云金芽孢杆

菌杀昆虫蛋白选自下组,该组由以下组成:Cry3Bb1、Cry34Ab1连同Cry35Ab1一起、mCry3A(美国专利号7276583,通过引用并入本文)、eCry3.1Ab(美国专利号8309516,通过引用并入本文)、和Vip3A蛋白(包括Vip3Aa(美国专利号6137033,通过引用并入本文))。

[0138] 在其他实施例中,本发明的转基因植物可包含第二杀有害生物剂(其可源自除苏云金芽孢杆菌以外的来源)。第二杀昆虫剂可以是选自下组的药剂,该组由以下组成: α 淀粉酶、过氧化物酶、胆固醇氧化酶、马铃薯糖蛋白、蛋白酶、蛋白酶抑制剂、脲酶、 α -淀粉酶抑制剂、成孔蛋白、几丁质酶、凝集素、工程化抗体或抗体片段、蜡样芽孢杆菌杀昆虫蛋白、致病杆菌属物种(如嗜线虫致病杆菌(*X.nematophila*)或伯氏致病杆菌(*X.bovienii*))杀昆虫蛋白、发光杆菌属物种(如发光杆菌(*P.luminescens*)或*P.asymbiotica*)杀昆虫蛋白、短芽孢杆菌属物种(如侧孢短芽孢杆菌(*B.laterosporous*))杀昆虫蛋白、赖氨酸芽孢杆菌属物种(*Lysinibacillus* spp.) (如球形赖氨酸芽孢杆菌(*L.Sphearicus*))杀昆虫蛋白、色杆菌属物种(如*C.subtsugae*或*C.piscinae*)杀昆虫蛋白、耶尔森菌属物种(如耶尔森噬虫霉(*Y.entomophaga*))杀昆虫蛋白、类芽孢杆菌属物种(如丙型类芽孢杆菌属(*P.propylaea*))杀昆虫蛋白、梭菌属物种(如双酶梭菌(*C.bifermentans*))杀昆虫蛋白和木质素。在其他实施例中,该第二药剂可以是至少一种源自杀昆虫毒素复合物(Tc)(该复合物来自发光杆菌属、异短杆菌属(*Xenorhabus*)、沙雷氏菌属或耶尔森菌属)的杀昆虫蛋白。在其他实施例中,杀昆虫蛋白可以是源自杀昆虫细菌(如发光杆菌属物种)的ADP-核糖基转移酶。在仍其他实施例中,杀昆虫蛋白可以是Axmi205或源自Axmi205(美国专利号8575425和9394345,其各通过引用并入本文)。在其他实施例中,杀昆虫蛋白可以是VIP蛋白(如来自蜡样芽孢杆菌的VIP1和/或VIP2)。在仍其他实施例中,杀昆虫蛋白可以是源自杀昆虫细菌(如来自侧孢短芽孢杆菌的ISP1A和ISP2A或来自球形芽孢杆菌的BinA和BinB)的二元毒素。在仍其他实施例中,杀昆虫蛋白可以被工程化或可以是前述杀昆虫蛋白中的任一个的杂交体或嵌合体。

[0139] 在一些实施例中,本发明的转基因植物可包含至少一种非蛋白质的第二杀有害生物剂。在优选的实施例中,该第二杀有害生物剂是干扰RNA分子。干扰RNA分子典型地包含针对靶基因的至少一种RNA片段、间隔序列、和与第一RNA片段互补的第二RNA片段,从而可以形成双链RNA结构。当生物识别双链RNA(dsRNA)分子并水解它们时,RNA干扰(RNAi)发生。所得的水解产物是约19-24个核苷酸长度的小RNA片段,这些小RNA片段被称为小干扰RNA(siRNA)。然后这些siRNA扩散或被携带至整个生物,包括越过细胞膜,在那里它们与mRNA(或其他的RNA)杂交并且导致RNA的水解。干扰RNA由RNA干扰沉默复合体(RISC)识别,RNA的效应链(或“引导链”)位于该复合体中。此引导链充当双链体序列的识别和破坏模板。每次siRNA与其互补RNA靶杂交,此过程都被重复,有效防止那些mRNA被翻译,并且因此“沉默”mRNA自其中转录的特异性基因的表达。本领域已知干扰RNA可用于昆虫控制(参见例如,公开物W0 2013/192256,其通过引用并入本文)。设计用于昆虫控制的干扰RNA产生非天然存在的双链RNA,其利用昆虫中的天然RNAi途径来触发靶基因的下调,这可能导致停止摄食和/或生长并可能导致昆虫有害生物死亡。干扰RNA分子可赋予针对与本发明的蛋白质相同的靶有害生物的昆虫抗性,或可靶向不同的有害生物。靶昆虫植物有害生物可以通过咀嚼、吸吮或刺穿来摄食。本领域已知干扰RNA可用于昆虫控制。在其他实施例中,干扰RNA可赋予针对非昆虫植物有害生物(如线虫有害生物或病毒有害生物)的抗性。

[0140] 多于一种的杀有害生物剂在相同转基因植物中的共表达可以通过制造单一的重

组载体(在一种所谓的分子堆积(molecular stack)中包含多于一种的杀有害生物剂的编码序列)并对植物进行遗传工程化以便在该转基因植物中含有并表达所有的杀有害生物剂而实现。此类分子堆积还可以通过使用微型染色体进行制造,如在(例如)美国专利7,235,716中所说明。可替代地,包含对第一杀有害生物剂进行编码的核酸的转基因植物可以用对第二有害生物剂等进行编码的不同的核酸进行再转化。可替代地,可以将植物(亲本1)遗传工程化,用于本发明的基因的表达。可以将第二植物(亲本2)遗传工程化,用于第二杀有害生物剂的表达。通过将亲本1与亲本2杂交,获得了表达被引入至亲本1和亲本2中的所有基因的子代植物。

[0141] 包括本发明的杀昆虫蛋白的转基因植物或种子还可以用杀昆虫剂或杀昆虫种子包衣进行处理,如美国专利号5,849,320和5,876,739(通过引用并入本文)中所述的。在本发明的杀昆虫剂或杀昆虫种子包衣以及转基因植物或种子针对相同靶昆虫(例如鞘翅目有害生物或根萤叶甲属靶有害生物)具有活性的情况下,该组合(i)在用于进一步增强本发明的组合物针对该靶昆虫的活性的方法中以及(ii)在用于通过提供针对该靶昆虫的又另一种作用机制而防止对本发明的组合物产生抗性的方法中是有用的。因此,本发明提供了增强根萤叶甲属昆虫群体的控制的方法,该方法包括提供本发明的转基因植物或种子并且向该植物或该种子施用本发明的杀昆虫剂或杀昆虫种子包衣。

[0142] 即使在杀昆虫种子包衣针对不同昆虫具有活性的情况下,该杀昆虫种子包衣对于扩展昆虫控制的范围是有用的,例如通过将针对鳞翅目昆虫具有活性的杀昆虫种子包衣添加到本发明的转基因种子(在一些实施例中针对鞘翅目和一些鳞翅目具有活性)中,所产生的包被的转基因种子控制鳞翅目和鞘翅目有害生物两者。

[0143] 此类杀昆虫剂和/或杀昆虫种子包衣的实例包括但不限于,氨基甲酸酯、拟除虫菊酯、有机磷酸酯、fiprole、新烟碱、有机氯化物、沙蚕毒素或其组合。在另一个实施例中,该杀昆虫剂或杀昆虫种子包衣选自下组,该组由以下组成:克百威、胺甲萘、灭多虫、联苯菊酯、七氟菊酯、氯菊酯、氟氯氰菊酯、 λ -三氟氯氰菊酯、氯氰菊酯、溴氰菊酯、毒死蜱、氯氧磷、乐果、灭线磷、马拉硫磷、甲基对硫磷、甲拌磷、特丁磷、叔丁嘧啶磷(tebupirimiphos)、氟虫腈、啉虫脒、吡虫啉、噻虫啉、噻虫嗪、硫丹、杀虫磺、及其组合。包含此类杀昆虫剂和杀昆虫种子包衣的商业产品包括但不限于以下:**Furadan®**(克百威)、**Lanate®**(灭多虫、灭多威、纳乃得)、**Sevin®**(胺甲萘)、**Talstar®**(联苯菊酯)、**Force®**(七氟菊酯)、**Ammo®**(氯氰菊酯)、**Cymbush®**(氯氧菊酯)、**Delta Gold®**(溴氰菊酯)、**Karate®**(λ -氯氟氰菊酯)、**Ambush®**(氯菊酯)、**Pounce®**(氯菊酯)、**Brigade®**(联苯菊酯)、**Capture®**(联苯菊酯)、**ProShield®**(七氟菊酯)、**Warrior®**(λ 氯氟氰菊酯)、**Dursban®**(毒死蜱)、**Fortress®**(氯氧磷)、**Mocap®**(灭克磷)、**Thimet®**(甲拌磷)、**AAstar®**(甲拌磷、氟氰戊菊酯(flucythinate))、**Rampart®**(甲拌磷)、**Counter®**(特丁磷)、**Cygon®**(乐果)、异氯磷、**Regent®**(氟虫腈)、**Cruiser®**(噻虫嗪)、**Gaucha®**(吡虫啉)、**Prescribe®**(吡虫啉)、**Poncho®**(噻虫胺)和**Aztec®**(氟氯氰菊酯、噻丙磷)。

[0144] 在一些实施例中,本发明还涵盖一种组合物,其包含控昆虫有效量的本发明杀昆虫蛋白。在进一步的实施例中,组合物包含合适的农业载剂和本发明的杀昆虫蛋白。该农业载剂可包括有益于活性成分(如本发明的蛋白质,包括包含以下、基本上由以下组成或由以下组成的蛋白质:与SEQ ID NO:1-22中的任一个至少80%、至少85%、至少90%、至少95%或100%相同的氨基酸序列)施用的佐剂、混合剂、增强剂等。合适的载剂不应应对有价值的作物具有植物毒性(特别是在作物的存在下施用组合物时所用的浓度),并且不应与本文的活性成分的化合物(即本发明的多肽或其他组合物成分)进行化学反应。此类混合物可以设计用于直接施用于作物,或者可以是浓缩物或配制品,其通常在施用前用另外的载剂和佐剂进行稀释。它们可以包括惰性或活性组分,并且可以是固体(如例如尘剂、粉剂、颗粒剂、水分散性颗粒剂或可湿性粉剂)或液体(如例如可乳化浓缩物、溶液、乳剂或悬浮液)。合适的农业载剂可包括液体载剂,例如水、甲苯、二甲苯、石脑油、作物油、丙酮、甲基乙基酮、环己酮、三氯乙烯、全氯乙烯、乙酸乙酯、乙酸戊酯、乙酸丁酯、丙二醇单甲醚和二乙二醇甲醚、甲醇、乙醇、异丙醇、戊醇、乙二醇、丙二醇、甘油等。水通常是用以稀释浓缩物的选用载剂。合适的固体载剂可包括滑石、叶腊石粘土、硅石、凹凸棒石粘土、砂藻土(kieselguhr)、白垩、硅藻土(diatomaceous earth)、石灰、碳酸钙、膨润土、漂白土、棉子壳、小麦粉、大豆粉、浮石、木粉、核桃壳粉、木质素等。在其他实施例中,本发明的蛋白质可以包封在合成基质(如聚合物)中,并施用于宿主(如植物)的表面。昆虫摄取宿主细胞允许将昆虫控制剂递送至昆虫并导致对昆虫有害生物的毒性作用。

[0145] 在另外的实施例中,本发明的组合物可以是粉剂、尘剂、丸剂、颗粒剂、喷雾剂、乳剂、胶体或溶液。本发明的组合物可以通过脱水、冷冻干燥、均化、提取、过滤、离心、沉降或浓缩细菌细胞的培养物来制备。本发明的组合物可包含按重量计至少1%,约5%、至少10%、至少20%、至少25%、至少30%、至少35%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%、至少95%、至少97%或至少99%的本发明的多肽。本发明的组合物可包含至少第二杀有害生物剂(其可以是杀昆虫的、杀线虫的、杀真菌的、杀细菌的)。至少第二杀有害生物剂可以针对与本发明的多肽相同的昆虫或不同的昆虫具有杀昆虫作用。第二杀有害生物剂可以是多肽。杀有害生物剂可以是干扰RNA。第二杀有害生物剂可以是微生物(如细菌),其包含编码杀有害生物剂的核酸分子和/或包含杀有害生物剂(如多肽或干扰RNA)。可对该微生物进行减毒、热灭活或冷冻干燥。微生物可能死亡或无法繁殖。第二杀有害生物剂可以是杀昆虫剂、例如克百威、胺甲萘、灭多虫、联苯菊酯、七氟菊酯、氯菊酯、氟氯氰菊酯、 λ -氯氟氰菊酯、氯氰菊酯、溴氰菊酯、毒死蜱、氯氧磷、乐果、灭线磷、马拉硫磷、甲基对硫磷、甲拌磷、特丁磷、叔丁嘧啶磷(tebupirimiphos)、氟虫腈、啉虫脒、吡虫啉、噻虫啉、噻虫嗪、硫丹、杀虫磺、及其组合,或含有如上所述的杀虫剂和杀虫种子包衣的商业产品。

[0146] 本发明的组合物(例如包含本发明的蛋白质和农业上可接受的载剂的组合物)可用于传统农业方法中。农业上可接受的载剂是可用于将包含本发明的多肽的组合物施用于植物或种子的配制品。例如,本发明的组合物可以与水和/或肥料混合,并且可以通过任何手段在出苗前和/或出苗后施用至所希望的场所,如飞机喷雾桶、灌溉设备、直接注射喷雾设备、背负式喷雾桶、家畜浸渍槽、在地面喷雾中使用的农场设备(例如,喷管式喷雾器、手动喷雾器)等。所希望的场所可以是土壤、植物等。

[0147] 本发明的组合物可以在以下时间施用至处于任何生理状态下的种子或植物繁殖

体：在种子收获和播种之间的任何时间；或在播种期间或播种后；和/或发芽后。优选种子或植物繁殖体处于足够耐用的状态，使得在处理过程期间不会造成损害或造成最小的损害，包括物理损害或生物损害。配制品可以使用常规的包衣技术以及机器（如流化床技术、滚筒研磨方法、静态转动 (rotostatic) 种子处理器和转鼓包衣器）施用至种子或植物繁殖体上。

[0148] 在一些实施例中，本发明还包括控制鞘翅目有害生物群体的方法，所述方法包括使所述有害生物群体与控昆虫有效量的本发明的杀昆虫蛋白接触，其中所述蛋白质包含以下、基本上由以下组成或由以下组成：与SEQ ID NO:1-16中的任一个至少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%或100%相同的氨基酸序列。接触包括有害生物群体的成员摄食或摄取杀昆虫蛋白。所述杀昆虫蛋白可以并入昆虫膳食食物中，或者可以在昆虫群体然后摄取的植物组织中表达或存在。在另外的实施例中，控制鞘翅目有害生物群体包括通过使昆虫与控昆虫有效量的本发明的杀昆虫蛋白接触来杀死昆虫。

[0149] 本发明进一步包含用于增加植物产量的方法，该方法包括在田地中生长已经在其基因组中稳定地并入本发明的表达盒的核酸分子的植物或其种子，并且其中所述田地被害生物侵染，所述多肽针对该有害生物具有杀昆虫活性。

[0150] 一旦将所希望的核酸转化至特定的植物物种中，它可以在那个物种中进行增殖或者使用传统的育种技术将其移至相同物种的其他品种中（特别是包括商品化的品种）。

[0151] 在一些实施例中，本发明涵盖向玉米种植者提供控制玉米作物中根萤叶甲属昆虫有害生物群体的手段的方法，所述方法包括 (a) 向所述种植者出售或提供包含本发明核酸分子、表达盒、载体或嵌合基因的转基因玉米种子；和 (b) 向所述种植者宣传所述转基因玉米种子产生控制根萤叶甲属有害生物群体的转基因玉米植物。

[0152] 在一些实施例中，本发明还涵盖一种鉴定杀昆虫蛋白的方法，所述杀昆虫蛋白包含以下、基本上由以下组成或由以下组成：与SEQ ID NO:1-22中的任一个或其毒素片段具有至少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%或具有100%序列同一性的核苷酸序列，所述方法包括以下步骤：(a) 产生引物对，所述引物对将从核酸样品或其互补序列扩增SEQ ID NO:1-3中任一个的多核苷酸，(b) 从所述核酸样品中扩增直系同源多核苷酸，(c) 鉴定直系同源多核苷酸的核苷酸序列，(d) 产生由所述直系同源多核苷酸编码的蛋白质，以及 (e) 确定步骤 (d) 的蛋白质针对昆虫有害生物具有杀昆虫活性。

实例

[0153] 本发明的实施例可以通过参考以下实例而被更好地理解。前述的和以下的本发明的实施例以及各种实施例的描述不是旨在限制权利要求书，而是对其具有说明性。因此，应理解的是权利要求书不旨在受限于这些实例的具体细节。本领域技术人员应理解的是本发明的其他实施例可以在不偏离本披露的精神和范围的情况下进行实践，本披露的范围是由所附权利要求书限定的。本领域公认的重组DNA和分子克隆技术可以在以下中找到，例如 J.Sambrook等人, Molecular Cloning: A Laboratory Manual [分子克隆：实验室手册]，第3版, Cold Spring Harbor [冷泉港], NY [纽约]: Cold Spring Harbor Laboratory Press [冷泉港实验室出版社] (2001) ; T.J.Silhavy, M.L.Berman, 和 L.W.Enquist, Experiments with

Gene Fusions[基因融合实验], Cold Spring Harbor Laboratory[冷泉港实验室], Cold Spring Harbor[冷泉港], NY[纽约] (1984) 和 Ausubel, F.M. 等人, Current Protocols in Molecular Biology[分子生物学实验手册], New York[纽约], John Wiley and Sons Inc. [约翰威利父子出版公司], (1988), Reiter 等人, Methods in Arabidopsis Research[拟南芥属研究方法], World Scientific Press[世界科学出版社] (1992), 以及 Schultz 等人, Plant Molecular Biology Manual[植物分子生物学手册], Kluwer Academic Publishers [克鲁沃学术出版社] (1998)。

实例1: SproCRW 杀昆虫蛋白的鉴定。

[0154] 属于肠杆菌科的变形斑沙雷氏菌 568 是本领域已知的, 最初作为根内寄生菌从毛果杨分离。已发现变形斑沙雷氏菌促进植物生长, 并已被假设和证明这是通过产生特定的化合物发生的, 例如脂质-几丁质低聚糖, 这些化合物被用作细菌到植物的特定信号。

[0155] 出人意料的是, 从变形斑沙雷氏菌 568 中分离的包含 SEQ ID NO:1 氨基酸序列的蛋白质被发现对西方玉米根虫 (WCR; 玉米根萤叶甲) 具有杀昆虫活性。编码所述蛋白质的天然 DNA 序列由 SEQ ID NO:23 表示。产生变形斑沙雷氏菌序列的大肠杆菌密码子优化版本 (SEQ ID NO:26) 并将其引入 pET29a 细菌表达载体以产生称为 pET-spro 的蛋白质。将构建体转化进入大肠杆菌 BL21* (DE3) 并且从异丙基 β -D-1-硫代吡喃半乳糖苷 (IPTG) 诱导的培养物 (在 18°C 下过夜生产蛋白质) 制得裂解物。在饲料掺入生物测定实验中, 检测了裂解物对 WCR 的杀昆虫活性。简言之, 将大肠杆菌裂解物与等体积的加热的人造昆虫饲料 (Bioserv 公司, 弗伦奇敦 (Frenchtown), 新泽西州 (NJ)) 在 1.5mL 离心管中混合, 然后施用于小的培养皿中。在膳食-样品混合物冷却并固化后, 向每个板中添加 12 只 WCR 幼虫。将板密封并且保持在就温度、光照以及相对湿度而言的环境实验室条件下。不含裂解物的缓冲液, 来自大肠杆菌 BL21* (DE3) 培养物 (含有空 pET29a 载体) 的裂解物和单独的人工昆虫饲料作为阴性对照。结果示出在表 1 中。获取侵染后 3 天和 6 天时的死亡率百分数和生长抑制观察, 其中 s = 小型幼虫, m = 中型幼虫, 并且 l = 大型幼虫。

表1: SproCRW 针对 WCR 的杀昆虫活性。

处理	第 3 天		第 6 天	
	死亡率%	生长	死亡率%	生长
50 mM Tris 8.5, 50 mM NaCl	0	b	0	b
BL21*/pET29a-空	8	b	17	b
BL21*/pET-Spro	92	s	100	s
仅饲料	0	b	0	b

[0156] 如表 1 所示, 来自表达变形斑沙雷氏菌蛋白的大肠杆菌培养物的裂解物对西方玉米根虫 (玉米根萤叶甲) 是杀昆虫活性的。变形斑沙雷氏菌杀昆虫蛋白被命名为 SproCRW。SproCRW 具有 489 个氨基酸, 长 53.4kDa 并且包含来自 SEQ ID NO:1 的约氨基酸 117-323 的膜攻击复合物 (MACPF) 区域。

实例2. SplyCRW 杀昆虫蛋白的鉴定。

[0157] 普利茅斯沙雷氏菌是腐生发酵性的非运动性革兰氏阴性杆菌, 其产生红色素 (灵菌红素)。目前所描述的大多数菌株都是从淡水和鱼类中分离出来的。

[0158] 出人意料的是,从普利茅斯沙雷氏菌分离的包含SEQ ID NO:2氨基酸序列的蛋白质也被发现对西方玉米根虫具有杀昆虫活性。该蛋白质的天然DNA序列为SEQ ID NO:27。产生普利茅斯沙雷氏菌序列的大肠杆菌密码子优化版本(SEQ ID NO:28)并将其引入pET29a细菌表达载体以产生称为pET-Sply的蛋白质。将构建体转化到大肠杆菌BL21*(DE3)中,并使用与上述针对pET-Spro所述的方法相似的方法制备裂解物。在与上述针对pET-Spro相似的饮食掺入生物测定实验中,对裂解物进行了针对WCR的杀昆虫活性测试。结果示出在表2中。获取侵染后4天和6天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。

表2.SplyCRW针对WCR的杀昆虫活性。

处理	第 4 天		第 6 天	
	死亡率%	生长	死亡率%	生长
pET29a-空	8%	m/l	25%	l
SplyCRW	67%	s/m	100%	-

[0159] 如表2所示,来自表达普利茅斯沙雷氏菌蛋白的大肠杆菌培养物的裂解物对西方玉米根虫(玉米根萤叶甲)是杀昆虫活性的。普利茅斯沙雷氏菌杀昆虫蛋白被命名为SplyCRW。SplyCRW具有488个氨基酸,长53.5kDa并且包含来自SEQ ID NO:2的约氨基酸117-323的膜攻击复合物区域。

实例3.SquiCRW杀昆虫蛋白的鉴定。

[0160] 奎沃斯沙雷氏菌是本领域已知的,并且最初是从土壤中分离出来的。已经发现奎沃斯沙雷氏菌与植物和无脊椎动物的消化道有关。

[0161] 出人意料的是,来自奎沃斯沙雷氏菌的包含SEQ ID NO:3氨基酸序列的蛋白质也被发现对西方玉米根虫具有杀昆虫活性。所述蛋白质的天然DNA序列由SEQ ID NO:25表示。产生奎沃斯沙雷氏菌序列的大肠杆菌密码子优化版本(SEQ ID NO:28)并将其引入pET29a细菌表达载体以产生称为pET-Squi的蛋白质。将构建体转化到大肠杆菌BL21*(DE3)中,并使用与上述针对pET-Spro所述的方法相似的方法制备裂解物。在与上述针对pET-Spro和pET-Sply相似的饮食掺入生物测定实验中,对裂解物进行了针对WCR的杀昆虫活性测试。结果示出在表3中。获取侵染后4天和6天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。

表3.SquiCRW针对WCR的杀昆虫活性。

处理	第 4 天		第 6 天	
	死亡率%	生长	死亡率%	生长
pET29a-空	0%	l	8%	l
SquiCRW	33%	m/l	75%	m/l

[0162] 如表3所示,来自表达奎沃斯沙雷氏菌蛋白的大肠杆菌培养物的裂解物对西方玉米根虫(玉米根萤叶甲)是杀昆虫活性的。奎沃斯沙雷氏菌杀昆虫蛋白被命名为SquiCRW。SquiCRW具有489个氨基酸,长53.2kDa并且包含来自SEQ ID NO:3的约氨基酸117-322的膜攻击复合物区域。

实例4.SproCRW、SpIyCRW、SquiCRW之间的关系。

[0163] 使用BLOSUM 62评分矩阵对分别为SEQ ID NO:1-3的SproCRW、SpIyCRW和SquiCRW氨基酸序列进行比对。如表4所示,SproCRW (SEQ ID NO:1) 和SquiCRW (SEQ ID NO:3) 杀昆虫蛋白的氨基酸序列99%相同,并且SproCRW (SEQ ID NO:1) 和SpIyCRW (SEQ ID NO:2) 杀昆虫蛋白的氨基酸序列82%相同。SpIyCRW氨基酸序列 (SEQ ID NO:2) 与SquiCRW氨基酸序列 (SEQ ID NO:3) 具有83%同一性。表4中的粗体文字表示膜攻击复合体 (MAC) 区域。

表4.CRW活性SIP的比对。

位置	序列	起始	结束	长度	匹配	% 同一性
参考 1	SproCRW (SEQ ID NO:1)	1	489	489 aa		
2	SquiCRW (SEQ ID NO:2)	1	489	489 aa	485	99
3	SpIyCRW (SEQ ID NO:3)	1	488	488 aa	405	82
(SEQ ID NO:1)	1 MKIESSKVEGLFSSSFTRVNAVPLPSDTLPGVGIIGCGYNPFLAYADASAVLHPILDWSK					
(SEQ ID NO:3)	1					
(SEQ ID NO:2)	1 ..F..L...N....L..I..I...N.A.....S.....V.....					
(SEQ ID NO:1)	61 SQFNEITMNGQQYQLPDVLQAVWLSNQSYASVTGKSLQSYLTELANSIKVSGNYGFFSAS					
(SEQ ID NO:3)	61 ...H.....					
(SEQ ID NO:2)	61 ...HTV....T....EI.N.....T.S..S.....S.....					
(SEQ ID NO:1)	121 ATNEFSDSSLRKSENEFSRCQQSFDLWSISIPADIARLQNYVSDDFIKLINAINPESKDS					
(SEQ ID NO:3)	121					
(SEQ ID NO:2)	121 T.....A.....I....K...SS.D.NNQQT					
(SEQ ID NO:1)	181 IATVFNVYGVSHVLMGVMGGKAHVSAASANKLTLTQKFEMSTIVQAKYEQLTSQSLVEDKL					
(SEQ ID NO:3)	181 L.....					
(SEQ ID NO:2)	181 L..I.....I.....A.....					
(SEQ ID NO:1)	241 KYSEAFDSFSESGSYTYDILGGSPSLGALVFKNNSQGSSDDNLKNWISISSMPVLTKFI					
(SEQ ID NO:3)	241 A.....S.....					
(SEQ ID NO:2)	241 E.....V.N.-D....RK..D...T.....					
(SEQ ID NO:1)	301 DQTSLMPVWLLCEDKTKADALKKYDNTWSKSQMAVASLRANYIDELTFVLGDNSDIPAP					
(SEQ ID NO:3)	301					
(SEQ ID NO:2)	300 L...T....QV....N..N...QL....R....V.....					
(SEQ ID NO:1)	361 VGYTKVPIDLNSDAGGKYVYLCYHEAQFTFVNGKQPIVDIQVLYGSQMPAPGYIKIDVDL					
(SEQ ID NO:3)	361					
(SEQ ID NO:2)	360 A.....V....G....FI.....S..A..GL....K.E...D.SR.NI..					
(SEQ ID NO:1)	421 NSGAGGEFVYLSYKKGEPTSSDVINKITAVYKNEYVPTPYGYKQISGDLNAGAGGDFVY					
(SEQ ID NO:3)	421					
(SEQ ID NO:2)	420 ...N.DD.....DA..KE.....DQ.....P...S.....					
(SEQ ID NO:1)	481 LCTYQGGTE					
(SEQ ID NO:3)	481					
(SEQ ID NO:2)	480 F.....					

实例5.SproCRW针对西方玉米根虫的活性。

[0164] 为了纯化SproCRW,从使用弗氏细胞压碎器 (French pressure cell) 的2升B121*培养物产生的细胞颗粒中,在20mM pH 7.0的磷酸钾、10%甘油和5mM 2-巯基乙醇 (BME) 中制备裂解物。裂解液离心除去细胞碎片和膜。然后收集上清液,添加洗涤剂N-十二烷基β-D-麦芽糖苷 (DDM) 至3μM的浓度。然后将裂解液施用至HiPrepQ阴离子交换FPLC柱。洗涤柱以除去污染物质,然后使用高盐缓冲液以线性梯度洗脱结合蛋白超过20柱体积。SproCRW的预期分子量为53.4kDa。如实例1所述,在饲料结合生物测定中测试纯化的蛋白质针对WCR的

功效,不同之处在于将已知浓度($\mu\text{g}/\text{mL}$)的纯化蛋白(“剂量”)与加热的人工饲料而不是细菌裂解物混合。使用1X PBS作为阴性对照。结果示出在表5中。获取侵染后3天和6天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。

[0165] 如表5所示,纯化的SproCRW在测试浓度范围(63-500 $\mu\text{g}/\text{mL}$)下显示出对WCR的强杀昆虫活性。

表5.纯化的SproCRW针对WCR的杀昆虫活性。

处理	浓度 ($\mu\text{g}/\text{ml}$)	第 3 天		第 6 天	
		死亡率%	生长	死亡率%	生长
1X PBS	0	8	l	21	l
SproCRW	500	75	s, m	100	s
SproCRW	250	42	s, m	100	s
SproCRW	125	33	s, m	100	s
SproCRW	63	25	s, m	93	s, m

实例6.SproCRW变体针对西方玉米根虫的活性。

[0166] 将SproCRW的Y188残基改变为色氨酸(W)或苯丙氨酸(F),以制备SproCRW-Y188W和SproCRW-Y188F突变变体(SEQ ID NO:5)。在类似于实例1所述的饮食掺入生物测定实验中测试这些变体的裂解物对WCR的活性。结果示出在表4中。获取侵染后3天和6天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。

[0167] SproCRW-Y188W变体针对WCR不再具有杀昆虫活性,而SproCRW-Y188F变体保留了杀昆虫活性。这些结果表明,SproCRW蛋白中的Y188残基和/或位置188对杀昆虫活性是重要的。

表6:SproCRW变体针对WCR的杀昆虫活性

处理	第 3 天		第 6 天	
	死亡率%	生长	死亡率%	生长
pET29a 空	13	l	31	l
SproCRW	33	s, m	84	s, m
SproCRW-Y188W	0	l	0	l
SproCRW-Y188F	21	s, m	85	s, m
仅饲料	0	l	0	l

[0168] 在表面覆盖生物测定中比较了野生型SproCRW和SproCRW-Y188F变体针对WCR在不同浓度范围内的杀昆虫活性。简单地说,在24孔塑料板的每个孔中添加1ml WCR饲料。每个孔施用每个样品(缓冲液阴性对照、空pET29-a载体阴性对照的细菌裂解物、或纯化蛋白)60 μl ,并将孔风干至少1小时。接下来,在饲料板上每孔侵染12只幼虫,然后用橡胶塞盖住孔。板放置在黑暗中并保持在室温下。获取侵染后4天和5天时的幼虫死亡率和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。SproCRW-Y188F变体针对WCR的杀昆虫活性与野生型SproCRW相似。

表7:SproCRW和SproCRW Y188F针对WCR的杀昆虫活性

处理	剂量 ($\mu\text{g}/\text{cm}^2$)	第 4 天		第 5 天	
		死亡率%	生长	死亡率%	生长
只用缓冲剂	0	0	l	0	l
pET29a-空	0	33	l	33	l
SproCRW	31	80	s, m	100	s
SproCRW-Y188F	31	75	s, m	100	s
SproCRW	16	67	s, m	100	s
SproCRW-Y188F	16	64	s, m	77	s, m
SproCRW	8	62	s, m	73	s, m
SproCRW-Y188F	8	40	s, m	40	s, m

实例7. 另外的SproCRW变体针对WCR的杀昆虫活性。

[0169] 通过诱变改变SproCRW的三个C末端残基,产生SproCRW-Y413W (SEQ ID NO:6)、SproCRW-F428W (SEQ ID NO:7) 和SproCRW-Y430W (SEQ ID NO:8) 变体。包含这些突变变体每者中的一个的裂解物与野生型SproCRW一起使用前面描述的饮食掺入生物测定来测试针对WCR的杀昆虫活性。所述测定以每个样品10只WCR幼虫进行,阴性对照是包含50mM磷酸钾、pH 7.0和50mM NaCl的缓冲液。生长抑制仅在第7天确定。结果示出在表8中,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。所有3个变体均保持针对玉米根虫的杀昆虫活性。

表8.SproCRW变体针对WCR的杀昆虫活性

处理	第 5 天	第 7 天	
	死亡率%	死亡率%	生长
只用缓冲剂	10	10	l
pET29a-空	0	20	l
SproCRW	80	100	s
SproCRW-Y413W	80	100	s
SproCRW-F428W	70	100	s
SproCRW-Y430W	90	100	s

[0170] 此外,改变SproCRW的两个残基以产生SproCRW-V8A/L11I变体 (SEQ ID NO:4)。还产生了SproCRW-S190P (SEQ ID NO:9) 和SproCRW-V192Y (SEQ ID NO:10) 突变变体。表达这些SproCRW突变变体的细菌的裂解物与野生型SproCRW裂解物一样具有针对WCR的杀昆虫活性。结果示出在表9中。获取侵染后2、5和8天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。SproCRW-V8A/L11I变体显示出与野生型SproCRW相当的杀昆虫活性。有趣的是,突变体SproCRW-S190P和SproCRW-V192Y在测定条件下不具有杀昆虫活性,这表明这两个氨基酸和/或位置对SproCRW蛋白的杀昆虫活性也是重要的。

表9:SproCRW变体针对WCR的杀昆虫活性。

处理	第 2 天		第 5 天		第 8 天	
	死亡率%	生长	死亡率%	生长	死亡率%	生长
只用缓冲剂	0	l	0	l	8	l
pET29a-空	0	l	8	l	8	l
SproCRW	8	m	46	s, m	67	s, m
SproCRW-V8A/L11I	0	m	50	s, m	50	s, m
SproCRW-S190P	0	l	0	l	0	l
SproCRW-V192Y	0	l	0	l	0	l

[0171] 产生了SproCRW的其他突变变体,包括SproCRW-C382M、SproCRW-C382G和SproCRW-C382F (SEQ ID NO:12),但没有测试对WCR的生物活性,因为没有产生这些突变体的可溶性蛋白。产生SproCRW突变变体SproCRW-C382A、SproCRW-C382S和SproCRW-C382T (SEQ ID NO:12),并产生包含这些变体中每一个的细菌裂解物。三个变体均产生可溶性突变SproCRW蛋白。使用饮食掺入生物测定,按照与实例1相似的方法测定裂解物。结果示出在表10中。获取侵染后4天和6天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。所有3个变体均保持针对CRW的杀昆虫活性。这些结果表明位置382在确定SproCRW杀昆虫蛋白的溶解度中可能是重要的。

表10:SproCRW变体针对WCR的杀昆虫活性

处理	第 4 天		第 6 天	
	死亡率%	生长	死亡率%	生长
pET29a-空	8	m/l	17	m/l
SproCRW-C382A	42	m	100	
SproCRW-C382S	8	m	50	m/l
SproCRW-C382T	17	m/l	67	m/l

[0172] SproCRW-C482L变体 (SEQ ID NO:13) 产生高可溶性蛋白,使用与上述类似的饲料掺入生物测定,测试包含SproCRW-C482L蛋白的裂解物与野生型SplyCRW蛋白相比针对西方玉米根虫的杀昆虫活性。结果示出在表11中。获取侵染后4天和6天时的死亡率百分数和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。SproCRW-C482L变体保留了杀昆虫活性。

表11.SproCRW-C482L变体针对WCR的杀昆虫活性。

处理	第 4 天		第 6 天	
	死亡率%	生长	死亡率%	生长
pET29a-空	8	m/l	25	l
SproCRW-C482L	58	m/l	75	m/l
SproCRW	67	s/m	100	

[0173] 生产了另外的SproCRW变体,并对其溶解度和对西方玉米根虫的杀昆虫活性进行了测试,基本上如上文所述。制备和测试的变体包括SproCRW-V398L (SEQ ID NO:14)、SproCRW-P396L/I397L (SEQ ID NO:15)、SproCRW-I397L (SEQ ID NO:16)、SproCRW-I397L/V398L (SEQ ID NO:17)、SproCRW-Y480L/C482L (SEQ ID NO:18) 和SproCRW-C482L/T483L

(SEQ ID NO:19)。结果示出在表12中。溶解度被评估为H=高溶解度、M=中溶解度和L=低溶解度。WCR活性记录如下

表12.SproCRW变体的表征。

处理	溶解度	WCR活性
pET29a-空	N/A	-
SproCRW	H	+++
SproCRW-V398L	H	++
SproCRW-P396L/I397L	L	-
SproCRW-I397L	M	+++
SproCRW-I397L/V398L	L	++
SproCRW-Y480L/C482L	L	-
SproCRW-C482L/T483L	L	-

实例8.SproCRW针对Cry抗性WCR的杀昆虫活性。

[0174] 为了确定SproCRW的毒性是否通过不同于Cry3相关蛋白的作用模式,如实例2中所述纯化SproCRW,并测试其针对抗经修饰的Cry3A (mCry3A) 毒素的WCR菌株 (mCry3A-R) 和抗eCry3.1Ab毒素的WCR菌株 (eCry3.1Ab-R) 的功效。基本上如实例2所述进行饲料掺入测定,并且获取侵染后第4和6天时的死亡率和生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。在0.6mg/ml和0.3mg/ml两种不同浓度下对SproCRW进行测试。阴性对照是1xPBS。还测定了对mCry3A或eCry3.1Ab(即敏感株(sus))无抗性的WCR菌株。如表13所示,SproCRW针对Cry抗性WCR菌株具有杀昆虫活性,这表明该蛋白质与来自苏云金芽孢杆菌的Cry蛋白相比具有独特的作用模式。因此,本发明的SIP,特别是SproCRW,和Cry蛋白的组合将有效地减轻对Cry蛋白或SIP,特别是SproCRW的抗性的发展。

表13.SproCRW针对Cry-R WCR的杀昆虫活性

处理	昆虫株系	第4天		第6天	
		死亡率%	生长	死亡率%	生长
SproCRW 0.6 mg/mL	WCRW-sus	58	m	67	m
SproCRW 0.3 mg/mL	WCRW-sus	50	m	75	m
1 x PBS	WCRW-sus	0	m/l	8	m/l
SproCRW 0.6 mg/mL	WCRW-mCry3A-R	100		100	
SproCRW 0.3 mg/mL	WCRW-mCry3A-R	25	s/m	75	m
1 x PBS	WCRW-mCry3A-R	0	m/l	25	m/l
SproCRW 0.6 mg/mL	WCRW-eCry3.1Ab-R	58	s/m	92	m
SproCRW 0.3 mg/mL	WCRW-eCry3.1Ab-R	58	m	100	
1 x PBS	WCRW-eCry3.1Ab-R	0	m/l	0	m/l

实例9.SproCRW针对北方玉米根虫的杀昆虫活性。

[0175] 使用如实例2所述的饲料掺入生物测定,对纯化的SproCRW进行针对北方玉米根虫

(NCR;长角根萤叶甲 (*Diabrotica longicornis*)) 的测试,不同的是对每种浓度测试12只NCR幼虫。结果示出在表14中。获取侵染后第1、2和6天时的死亡率百分比,并获取侵染后第6天时的生长抑制观察,其中s=小型幼虫,m=中型幼虫,并且l=大型幼虫。结果表明,SproCRW针对NCR具有杀昆虫活性。

表14.SproCRW针对NCR的杀昆虫活性

处理	第 1 天	第 2 天	第 6 天	
	死亡率%	死亡率%	死亡率%	生长
1X PBS	0	0	0	l
SproCRW 0.2 mg/mL	0	0	83	s
SproCRW 0.1 mg/mL	0	0	42	m

实例10.SproCRW针对秋夜蛾的杀昆虫活性。

[0176] 为了确定SproCRW是否具有针对秋夜蛾 (FAW; 草地贪夜蛾) 的杀昆虫活性,使用与实例1相似的方法,在饲料掺入生物测定中测试包含SproCRW蛋白的细菌裂解物。WCR和FAW的测定并行进行,使用12只L1 FAW幼虫或12只CRW幼虫。包含仅50mM Tris (pH 8.5) 和50mM NaCl的缓冲液、来自携带空pET-29a载体的细菌的细菌裂解物和不添加细菌裂解物的单独饲料各自被包括作为阴性对照。表达苏云金芽孢杆菌Vip3 FAW活性杀昆虫蛋白的大肠杆菌B121* (DE3) 衍生的裂解物作为FAW活性测定的阳性对照。在侵染后第4天评估死亡率、生长抑制(对于WCR而言) (其中s=小型幼虫、m=中型幼虫和l=大型幼虫) 和摄食行为(对于FAW而言) (其中f=摄食和nf=无摄食)。在这些实验条件下,SproCRW针对FAW没有活性(表15)。

表15.SproCRW针对秋夜蛾的杀昆虫活性

处理	WCRW, %死亡率		FAW, %死亡率	
	第 4 天	生长	第 4 天	摄食
缓冲液	0	l	0	f
pET29a-空	8	l	0	f
SproCRW	100	s	0	f
Vip3D	0	l	100	nf
饲料	0	l	0	f

实例11.SplyCRW变体的表征。

[0177] 构建了三个SplyCRW突变体,并对其溶解度和对西方玉米根虫的杀昆虫活性进行了测试,基本上如上文所述。所制备的变体包括SplyCRW-C482L/T483L (SEQ ID NO:20)、SplyCRW-C481L (SEQ ID NO:21) 和SplyCRW-C481L/T482L (SEQ ID NO:22)。结果示出在表16中。溶解度被评估为H=高溶解度、M=中溶解度和L=低溶解度。WCR活性记录如下

表16.SplyCRW变体的表征。

处理	溶解度	WCR活性
pET29a-空	N/A	-
SplyCRW	H	+++
SplyCRW-Y479L/C481L	L	-
SplyCRW-C481L	H	+++

SplyCRW-C481L/T482L	M	-
---------------------	---	---

实例12.用SproCRW V8A/L11I转化玉蜀黍。

[0178] 如美国专利号6,051,760 (通过引用并入本文) 中所述,产生了编码SproCRW突变蛋白SproCRW-V8A/L11I (SEQ ID NO:4) 的玉蜀黍优化的核苷酸序列 (SproCRWV8A-L11I)。SproCRWV8A-L11I) 编码序列在SEQ ID NO:58中列出。

[0179] 构建了两个植物表达盒,以将SproCRW-V8A/L11I编码序列引入玉蜀黍。第一盒包含可操作地连接到SproCRW-V8A/L11I编码序列的玉蜀黍泛素1 (Ubi1) 启动子,所述SproCRW-V8A/L11I编码序列可操作地连接到玉蜀黍Ubi361终止子。第二盒包含可操作地连接到编码选择性标记磷酸甘露糖异构酶 (PMI) 的pmi编码序列的玉蜀黍Ubi1启动子,所述pmi编码序列可操作地连接到玉蜀黍Ubi1终止子。生成包含所述两个表达盒的重组植物转化二元载体,用于玉蜀黍转化实验。

[0180] 使用标准分子生物学技术将所述二元载体转化到根癌农杆菌中。为了制备用于转化的农杆菌,将细胞在28℃和220rpm下在液体YPC培养基中培养过夜。

[0181] 基本如描述于Negrotto等人,2000,Plant Cell Reports[植物细胞报告]19:798-803中的进行未成熟的玉蜀黍胚的农杆菌转化。对于这个实例,所有的培养基组分基本上如在以上Negrotto等人所说明。然而,本领域内已知的各种培养基组分可以被代替。

[0182] 简言之,使包含二元载体植物转化载体的农杆菌菌株LBA4404 (pSB1) 在28℃下在YEP (酵母提取物 (5g/L)、蛋白胨 (10g/L)、NaCl (5g/L)、15g/l琼脂,pH 6.8) 固体培养基上生长2-4天。将大约 0.8×10^9 个农杆菌悬浮于补充有100μM As的LS-inf培养基中 (Negrotto等人,同上)。在这个培养基中对细菌预诱导30至60分钟。

[0183] 将来自适合的基因型的未成熟胚从8-12天大的穗中切除到液体LS-inf+100μM As中。用新鲜的感染培养基漂洗这些胚。然后添加农杆菌溶液,并且将这些胚涡旋30秒并且允许其与细菌一起沉降5分钟。然后将这些胚盾片向上地转移到LSA培养基中,并且在暗处培养两到三天。随后,将每皮氏板 (petri plate) 20与25个之间的胚转移到补充有头孢噻肟 (250mg/l) 和硝酸银 (1.6mg/l) 的LSDc培养基中,并且在28℃在黑暗中培养10天。

[0184] 将产生胚性愈伤组织的未成熟胚转移至LSD1M0.5S培养基中。在这种培养基上对培养物进行持续大约6周的选择,具有大约3周的传代培养步骤。将存活的愈伤组织转移至补充有甘露糖的Reg1培养基中。在光照中 (16小时光照/8小时黑暗方案) 培养之后,然后将绿色组织转移至没有生长调节剂的Reg2培养基中并且孵育约1-2周。将这些小植株转移至含有Reg3培养基的Magenta GA-7盒 (马真塔公司 (Magenta Corp), 芝加哥, 伊利诺伊州) 中并使其在光照中生长。

[0185] 在转化、选择和再生后,使用**TaqMan®**分析来测定植物中是否存在pmi基因和SproCRW-V8A/L11I玉蜀黍密码子优化的编码序列 (SEQ ID NO:54)。还测试了植物中是否存在载体骨架。将对载体骨架阴性并且包含来自二元载体的转基因的一个拷贝的十四个植物转移到温室中,并测试其针对WCR的杀昆虫活性。

实例12:转基因玉蜀黍植物的杀昆虫活性

[0186] 对于每个转基因玉蜀黍事件,通过ELISA检测SproCRW-V8A/L11I蛋白的表达并将其定量为ng SproCRW/mg叶和根组织中的总可溶性蛋白 (TSP)。当表达SproCRW-V8A/L11I的玉蜀黍事件达到V3-V4阶段时,采集来自每次事件的玉蜀黍根组织的样品。将玉蜀黍根组织

从植物上切除,置于皮氏培养皿中,然后用10-12只新生WCR幼虫侵染。在第4天评估了两个根组织样品(Rep1和Rep2)的摄食损伤,并根据摄食孔的数量和根组织瘢痕形成将其评分为L=低,M=中和H=高。将来自未转化的玉蜀黍(无效)的根组织用作阴性对照。与无效样品根组织相比,SproCRW-V8A/L11I在转基因玉蜀黍中的表达在大多数SproCRW转基因根组织中保护了其免受WCR(表14)。

表14. 表达转基因SproCRW-V8A/L11I玉蜀黍的杀昆虫活性。

植物 ID	SproCRW-V8A/L11I 浓度 (ng/mg TSP)	根 损伤	WCR 活性
1	39.27	L	++
2	41.19	L/M	+++
3	70.41	M/H	+
4	47.56	L	+++
5	34.59	M/H	+
6	39.9	L	+++
7	63.31	L/M	++
8	39.78	L/M	++
9	30.87	M/H	+
10	36.3	L	+++
11	31.57	M/H	+
12	36.82	M/H	+
13	41.86	L/M	++
14	65.76	M/H	+
对照	-	H	-

实例13. 与第二杀昆虫剂组合的SproCRW。

[0187] 如实例2所述纯化SproCRW。制备针对必需靶标并已知具有杀昆虫活性的dsRNA。在非限制性实例中,dsRNA可靶向编码液泡ATP合酶、 β -微管蛋白、26S蛋白体亚基p28蛋白、EF1 α 48D、肌钙蛋白I、跨膜四蛋白、网格蛋白重链、 γ -外被体、 β -外被体和/或保幼激素环氧化物水解酶的基因(PCT专利申请号PCT/US17/044825;PCT/US 17/044831;PCT/US 17/044832;美国专利号7,812,219;各自通过引用并入本文)。在基本上如实例1所述进行的膳食掺入测定中测试dsRNA和纯化的蛋白质针对WCR的功效。结果将表明dsRNA和SproCRW蛋白针对WCR具有活性。

实例14. 模拟胃液测定中SIP的命运

[0188] 先前的实验已经确定,包含SEQ ID NO:1氨基酸序列的天然SproCRW杀昆虫蛋白在标准的模拟胃液(SGF)测定中显示出延迟的消化,其中在30分钟的时间段中仍存在未消化的条带。SGF测定用于估算哺乳动物肠道中蛋白质的消化率,是评估任何新的杀昆虫蛋白用于监管部门批准的标准组成。因此,SproCRW突变体在SGF测定中被测试以确定其消化率。

[0189] 模拟胃液(SGF)测定在代表哺乳动物上消化道的严格控制的条件下测量测试蛋白

的体外消化率。简言之,在37℃下经一小时的时间段,将细菌产生的测试SproCRW蛋白和突变体(在0.5-5mg/ml的浓度下)以10单位的胃蛋白酶活性/μg测试蛋白的比率暴露于胃蛋白酶(来自猪胃粘膜,溶解于2mg/ml NaCl中,pH 1.2)。将样品在1、2、5、10、30和60分钟时取出,并通过加入预热的(95℃-2分钟)终止缓冲液(65%0.5M碳酸氢钠(pH 11)、35%Tricine加样缓冲液)立即淬灭,以使胃蛋白酶立即失活,并返回加热再持续5分钟。一旦测定完成,便通过SDS-PAGE在10%-20%Tris-Tricine凝胶(肽可见低至1kDa)上检查时间点样品和对照(只有测试蛋白,只有胃蛋白酶))以跟踪由胃蛋白酶进行的消化的动力学和水平。

[0190] SGF测定的结果表明,SproCRW突变体SproCRW-C482L(SEQ ID NO:13)和SproCRW-V398L(SEQ ID NO:14)降解非常迅速。这些结果提供了证据,证明在标准SGF测定中,可以突变SproCRW以提高消化率。

实例15.在植物细胞中原位编辑基因组以生成经修饰的SIP。

[0191] 以下实例说明了利用植物细胞基因组的原位基因组编辑以将本文所述的突变(包括但不限于实例6、7和11中所述的突变)掺入天然SIP(包括SproCRW(SEQ ID NO:1)、SplyCRW(SEQ ID NO:2)和/或SquiCRW(SEQ ID NO:3))的编码序列或已修饰的SproCRW、SplyCRW和/或SquiCRW蛋白的编码序列。

[0192] 定向基因组修饰(也称为基因组编辑)可用于在特定DNA序列中引入突变。这些基因组编辑技术,其包括锌指核酸酶(ZNF)、转录激活子样效应子核酸酶(TALEN)、大范围核酸酶和成簇规律间隔短回文重复(CRISPR),已成功施用至包括作物植物在内的50多种不同生物。参见例如,Belhaj,K.等人,Plant Methods[植物方法]9,39(2013);Jiang,W.等人,Nucleic Acids Res[核酸研究]41,e188(2013)。用于基因组编辑的CRISPR/Cas系统基于Cas9核酸酶和指定靶多核苷酸序列的工程化的单引导RNA(sgRNA)的瞬时表达。

[0193] Cas9是一种大的单体DNA核酸酶,其借助两个20-核苷酸(nt)非编码RNA的复合物被引导至DNA靶序列:CRIPSR RNA(crRNA)和反式激活crRNA(tracrRNA),其功能上可作为单合成RNA嵌合体获得。Cas9蛋白含有两个与RuvC和HNH核酸酶同源的核酸酶结构域。HNH核酸酶结构域切割互补DNA链,而RuvC样结构域切割非互补链,因此,在靶DNA中引入钝切(blunt cut)。

[0194] 当Cas9和sgRNA在活的玉蜀黍细胞中瞬时表达时,在转基因玉蜀黍细胞中产生特异性靶DNA中的双链断裂(DSB)。通过非同源末端连接和同源定向DNA修复途径引入断裂位点处的突变。

[0195] 通过使用在转基因玉蜀黍中表达Cas9核酸酶和sgRNA靶(针对SproCRW或经修饰的SproCRW序列进行了玉蜀黍密码子优化)的重组质粒,将特定突变(例如上述实例6、7和11中所述的突变)引入天然SproCRW杀昆虫蛋白(SEQ ID NO:1)或经修饰的SproCRW蛋白的编码序列。该方法的实施是通过根癌农杆菌的农杆菌浸润方法,该根癌农杆菌携带含有指定的目的靶序列的二元质粒.sgRNA与靶SproCRW或经修饰的SproCRW编码序列结合后,Cas9核酸酶对编码序列进行特异性切割,并在DNA修复过程中引入所需的一个或多个突变。因此,现在突变的SproCRW编码序列将编码经修饰的SproCRW变体蛋白,例如表7中描述的变体,例如,其中位置413的突变用色氨酸(W)替代酪氨酸(Y),或位置428的突变用色氨酸(W)替代苯丙氨酸(F);或例如表9中所述的变体,例如其中位置8的突变用丙氨酸(A)替代缬氨酸(V)与位置11的突变(其用异亮氨酸(I)替代亮氨酸(L)的突变)组合。

[0196] 通过PCR和测序筛选包含基因组编辑的SproCRW编码序列的植物细胞。诱导在SproCRW或经修饰的SproCRW编码序列中具有基因组编辑的突变的愈伤组织,以再生植物进行表型评估,评估表达的SproCRW蛋白对以下的杀昆虫活性:WCRW,北方玉米根虫(巴氏根萤叶甲)、南方玉米根虫(黄瓜十一星叶甲食根亚种)和/或墨西哥玉米根虫(墨西哥玉米根萤叶甲)。

[0197] 应当理解的是,本文所述的实例和实施例仅是出于说明性的目的,并且根据其说明书的不同修改或变化将提示本领域的技术人员并且将会包含在本申请的精神和范围内以及所附权利要求书的范围内。

[0198] 在本说明书中提到的所有公开物和专利申请显示了本发明所属领域中的技术人员的技术水平。所有这些公开物和专利申请均通过引用并入本文,其程度如同每个单独的公开物或专利申请被明确地并单独地指示通过引用而被并入。

序列表

<110> 先正达基团(Syngenta Participations AG)

REYNOLDS, Clarence Michael

<120> 杀昆虫蛋白

<130> 81599-W0-REG-ORG-P-1

<150> 62/752,546

<151> 2018-10-30

<160> 54

<170> PatentIn 版本3.5

<210> 1

<211> 489

<212> PRT

<213> 变形斑沙雷氏菌

<400> 1

```

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

```

Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala		
195	200	205
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln		
210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Val Glu Asp Lys Leu		
225	230	235
Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr		
	245	250
		255
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys		
	260	265
		270
Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln		
	275	280
		285
Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser		
290	295	300
Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala		
305	310	315
		320
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val		
	325	330
		335
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly		
	340	345
		350
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile		
	355	360
		365
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His		
370	375	380
Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile		
385	390	395
		400
Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile		
	405	410
		415
Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser		
	420	425
		430
Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr		
	435	440
		445
Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys		
450	455	460
Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr		
465	470	475
		480
Leu Cys Thr Tyr Gln Gly Gly Thr Glu		
	485	

<210> 2

<211> 488

<212> PRT

<213> 普利茅斯沙雷氏菌

<400> 2

```

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

```

275	280	285
Ile Ser Thr Met Pro Val	Leu Thr Lys Phe Ile Asp	Gln Thr Ser Leu
290	295	300
Leu Pro Val Trp Thr Leu	Cys Glu Asp Gln Val Lys	Ala Asp Ala Leu
305	310	315
Lys Asn Tyr Tyr Asn Asn	Thr Trp Ser Gln Leu Gln	Met Ala Val Ala
325	330	335
Arg Leu Arg Ala Asn Tyr	Ile Asp Glu Val Thr Phe	Val Leu Gly Asp
340	345	350
Asn Ser Asp Ile Pro Ala	Pro Ala Gly Tyr Thr Lys	Val Pro Val Asp
355	360	365
Leu Asn Ser Gly Ala Gly	Gly Lys Phe Ile Tyr Leu	Cys Tyr His Glu
370	375	380
Ala Gln Phe Thr Pro Val	Asn Ser Lys Gln Ala Ile	Val Gly Leu Gln
385	390	395
Val Leu Tyr Gly Lys Gln	Glu Pro Ala Pro Asp Tyr	Ser Arg Ile Asn
405	410	415
Ile Asp Leu Asn Ser Gly	Ala Asn Gly Asp Asp Val	Tyr Leu Ser Tyr
420	425	430
Lys Lys Gly Asp Ala Thr	Ser Lys Glu Val Ile Asn	Lys Ile Thr Ala
435	440	445
Val Tyr Gly Lys Asp Gln	Tyr Val Pro Thr Pro Tyr	Gly Tyr Lys Gln
450	455	460
Ile Pro Gly Asp Leu Asn	Ser Gly Ala Gly Gly Asp	Phe Val Tyr Phe
465	470	475
Cys Thr Tyr Gln Gly Gly	Thr Glu	
485		

<210> 3

<211> 489

<212> PRT

<213> 奎沃斯沙雷氏菌

<400> 3

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

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

Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His
 370 375 380
 Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile
 385 390 395 400
 Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile
 405 410 415
 Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser
 420 425 430
 Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr
 435 440 445
 Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys
 450 455 460
 Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr
 465 470 475 480
 Leu Cys Thr Tyr Gln Gly Gly Thr Glu
 485
 <210> 4
 <211> 489
 <212> PRT
 <213> 人工序列
 <220>
 <223> SproCRW V8A L11I
 <400> 4
 Met Lys Ile Glu Ser Ser Lys Ala Glu Gly Ile Phe Ser Ser Ser Phe
 1 5 10 15
 Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val
 20 25 30
 Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala
 35 40 45
 Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn
 50 55 60
 Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln
 65 70 75 80
 Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser
 85 90 95
 Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly
 100 105 110
 Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser
 115 120 125

Ser	Leu	Arg	Lys	Ser	Glu	Asn	Glu	Phe	Ser	Arg	Cys	Gln	Gln	Ser	Phe
130						135					140				
Asp	Leu	Trp	Ser	Ile	Ser	Ile	Pro	Ala	Asp	Ile	Ala	Arg	Leu	Gln	Asn
145					150					155					160
Tyr	Val	Ser	Asp	Asp	Phe	Ile	Lys	Leu	Ile	Asn	Ala	Ile	Asn	Pro	Glu
				165					170					175	
Ser	Lys	Asp	Ser	Ile	Ala	Thr	Val	Phe	Asn	Val	Tyr	Gly	Ser	His	Val
			180					185						190	
Leu	Met	Ser	Gly	Val	Met	Gly	Gly	Lys	Ala	His	Val	Ser	Ala	Ser	Ala
		195					200					205			
Asn	Lys	Leu	Thr	Leu	Thr	Gln	Lys	Phe	Glu	Met	Ser	Thr	Ile	Val	Gln
		210				215					220				
Ala	Lys	Tyr	Glu	Gln	Leu	Thr	Ser	Gln	Leu	Ser	Val	Glu	Asp	Lys	Leu
225					230					235					240
Lys	Tyr	Ser	Glu	Ala	Phe	Asp	Ser	Phe	Ser	Glu	Ser	Gly	Ser	Tyr	Thr
				245					250					255	
Tyr	Asp	Ile	Leu	Gly	Gly	Ser	Pro	Ser	Leu	Gly	Ala	Leu	Val	Phe	Lys
			260					265						270	
Asn	Asn	Ser	Gln	Gly	Ser	Ser	Asp	Asp	Asn	Leu	Lys	Asn	Trp	Ile	Gln
		275					280					285			
Ser	Ile	Ser	Ser	Met	Pro	Val	Leu	Thr	Lys	Phe	Ile	Asp	Gln	Thr	Ser
		290				295					300				
Leu	Met	Pro	Val	Trp	Leu	Leu	Cys	Glu	Asp	Lys	Thr	Lys	Ala	Asp	Ala
305					310					315					320
Leu	Lys	Lys	Tyr	Tyr	Asp	Asn	Thr	Trp	Ser	Lys	Ser	Gln	Met	Ala	Val
				325					330					335	
Ala	Ser	Leu	Arg	Ala	Asn	Tyr	Ile	Asp	Glu	Leu	Thr	Phe	Val	Leu	Gly
			340					345						350	
Asp	Asn	Ser	Asp	Ile	Pro	Ala	Pro	Val	Gly	Tyr	Thr	Lys	Val	Pro	Ile
		355					360					365			
Asp	Leu	Asn	Ser	Asp	Ala	Gly	Gly	Lys	Tyr	Val	Tyr	Leu	Cys	Tyr	His
		370				375					380				
Glu	Ala	Gln	Phe	Thr	Pro	Val	Asn	Gly	Lys	Gln	Pro	Ile	Val	Asp	Ile
385					390					395					400
Gln	Val	Leu	Tyr	Gly	Ser	Gln	Met	Pro	Ala	Pro	Gly	Tyr	Ile	Lys	Ile
				405					410					415	
Asp	Val	Asp	Leu	Asn	Ser	Gly	Ala	Gly	Gly	Glu	Phe	Val	Tyr	Leu	Ser
			420					425					430		
Tyr	Lys	Lys	Gly	Glu	Pro	Thr	Ser	Ser	Asp	Val	Ile	Asn	Lys	Ile	Thr

435	440	445
Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys		
450	455	460
Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr		
465	470	475
Leu Cys Thr Tyr Gln Gly Gly Thr Glu		480
485		
<210> 5		
<211> 489		
<212> PRT		
<213> 人工序列		
<220>		
<223> SproCRW Y188W		
<220>		
<221> 尚未归类的特征		
<222> (188) .. (188)		
<223> Xaa = W或F		
<400> 5		
Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe		
1	5	10
Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val		15
20	25	30
Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala		
35	40	45
Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn		
50	55	60
Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln		
65	70	75
Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser		
85	90	95
Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly		
100	105	110
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser		
115	120	125
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe		
130	135	140
Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn		
145	150	155
Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu		160

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

Leu Cys Thr Tyr Gln Gly Gly Thr Glu

485

<210> 6

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW Y413W

<400> 6

Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe

1 5 10 15

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

20 25 30

Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala

35 40 45

Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn

50 55 60

Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln

65 70 75 80

Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser

85 90 95

Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly

100 105 110

Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser

115 120 125

Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe

130 135 140

Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn

145 150 155 160

Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu

165 170 175

Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Val

180 185 190

Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala

195 200 205

Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln

210 215 220

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

225 230 235 240

Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr
 245 250 255
 Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys
 260 265 270
 Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln
 275 280 285
 Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser
 290 295 300
 Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala
 305 310 315 320
 Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val
 325 330 335
 Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly
 340 345 350
 Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile
 355 360 365
 Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His
 370 375 380
 Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile
 385 390 395 400
 Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Trp Ile Lys Ile
 405 410 415
 Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser
 420 425 430
 Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr
 435 440 445
 Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys
 450 455 460
 Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr
 465 470 475 480
 Leu Cys Thr Tyr Gln Gly Gly Thr Glu
 485

<210> 7

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW F428W

<400> 7

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

305	310	315	320
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val			
	325	330	335
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly			
	340	345	350
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile			
	355	360	365
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His			
	370	375	380
Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile			
385	390	395	400
Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile			
	405	410	415
Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Trp Val Tyr Leu Ser			
	420	425	430
Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr			
	435	440	445
Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys			
	450	455	460
Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr			
465	470	475	480
Leu Cys Thr Tyr Gln Gly Gly Thr Glu			
	485		
<210> 8			
<211> 489			
<212> PRT			
<213> 人工序列			
<220>			
<223> SproCRW Y430W			
<400> 8			
Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe			
1	5	10	15
Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val			
	20	25	30
Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala			
	35	40	45
Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn			
	50	55	60
Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln			

65	70	75	80
Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser			
	85	90	95
Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly			
	100	105	110
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser			
	115	120	125
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe			
	130	135	140
Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn			
145	150	155	160
Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu			
	165	170	175
Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Val			
	180	185	190
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala			
	195	200	205
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln			
	210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Val Glu Asp Lys Leu			
225	230	235	240
Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr			
	245	250	255
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys			
	260	265	270
Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln			
	275	280	285
Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser			
	290	295	300
Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala			
305	310	315	320
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val			
	325	330	335
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly			
	340	345	350
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile			
	355	360	365
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His			
	370	375	380

Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile
 385 390 395 400
 Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile
 405 410 415
 Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Trp Leu Ser
 420 425 430
 Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr
 435 440 445
 Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys
 450 455 460
 Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr
 465 470 475 480
 Leu Cys Thr Tyr Gln Gly Gly Thr Glu
 485

<210> 9

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW S190P

<400> 9

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

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

450	455	460
Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr		
465	470	475
Leu Cys Thr Tyr Gln Gly Gly Thr Glu		480
	485	
<210> 10		
<211> 489		
<212> PRT		
<213> 人工序列		
<220>		
<223> SproCRW V192Y		
<400> 10		
Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe		
1	5	10
Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val		15
	20	25
Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala		30
	35	40
Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn		45
	50	55
Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln		60
65	70	75
Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser		80
	85	90
Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly		95
	100	105
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser		110
	115	120
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe		125
	130	135
Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn		140
145	150	155
Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu		160
	165	170
Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Tyr		175
	180	185
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala		190
	195	200
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln		205

210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Val Glu Asp Lys Leu		
225	230	235
Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr		240
	245	250
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys		255
	260	265
Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln		270
	275	280
Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser		285
290	295	300
Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala		
305	310	315
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val		320
	325	330
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly		335
	340	345
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile		350
	355	360
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His		365
370	375	380
Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile		
385	390	395
Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile		400
	405	410
Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser		415
	420	425
Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr		430
	435	440
Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys		445
	450	455
Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr		460
465	470	475
Leu Cys Thr Tyr Gln Gly Gly Thr Glu		480
	485	

<210> 11

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW C312L

<400> 11

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

Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser
 290 295 300
 Leu Met Pro Val Trp Leu Leu Glu Asp Lys Thr Lys Ala Asp Ala
 305 310 315 320
 Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val
 325 330 335
 Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly
 340 345 350
 Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile
 355 360 365
 Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His
 370 375 380
 Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile
 385 390 395 400
 Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile
 405 410 415
 Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser
 420 425 430
 Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr
 435 440 445
 Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys
 450 455 460
 Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr
 465 470 475 480
 Leu Cys Thr Tyr Gln Gly Gly Thr Glu
 485

<210> 12

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW C382A

<220>

<221> 尚未归类的特征

<222> (382) .. (382)

<223> Xaa = A, S, T, M, G, F或L.

<400> 12

Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe
 1 5 10 15

Thr	Arg	Val	Asn	Ala	Val	Pro	Leu	Pro	Ser	Asp	Thr	Leu	Pro	Gly	Val
20				25				30							
Gly	Ile	Ile	Gly	Cys	Gly	Tyr	Asn	Pro	Phe	Leu	Ala	Tyr	Ala	Asp	Ala
35				40				45							
Ser	Ala	Val	Leu	His	Pro	Ile	Leu	Asp	Trp	Ser	Lys	Ser	Gln	Phe	Asn
50				55				60							
Glu	Ile	Thr	Met	Asn	Gly	Gln	Gln	Tyr	Gln	Leu	Pro	Asp	Val	Leu	Gln
65				70				75				80			
Ala	Val	Trp	Leu	Ser	Asn	Gln	Ser	Tyr	Ala	Ser	Val	Thr	Gly	Lys	Ser
85				90				95							
Leu	Gln	Ser	Tyr	Leu	Thr	Glu	Leu	Ala	Asn	Ser	Ile	Lys	Val	Ser	Gly
100				105				110							
Asn	Tyr	Gly	Phe	Phe	Ser	Ala	Ser	Ala	Thr	Asn	Glu	Phe	Ser	Asp	Ser
115				120				125							
Ser	Leu	Arg	Lys	Ser	Glu	Asn	Glu	Phe	Ser	Arg	Cys	Gln	Gln	Ser	Phe
130				135				140							
Asp	Leu	Trp	Ser	Ile	Ser	Ile	Pro	Ala	Asp	Ile	Ala	Arg	Leu	Gln	Asn
145				150				155				160			
Tyr	Val	Ser	Asp	Asp	Phe	Ile	Lys	Leu	Ile	Asn	Ala	Ile	Asn	Pro	Glu
165				170				175							
Ser	Lys	Asp	Ser	Ile	Ala	Thr	Val	Phe	Asn	Val	Tyr	Gly	Ser	His	Val
180				185				190							
Leu	Met	Ser	Gly	Val	Met	Gly	Gly	Lys	Ala	His	Val	Ser	Ala	Ser	Ala
195				200				205							
Asn	Lys	Leu	Thr	Leu	Thr	Gln	Lys	Phe	Glu	Met	Ser	Thr	Ile	Val	Gln
210				215				220							
Ala	Lys	Tyr	Glu	Gln	Leu	Thr	Ser	Gln	Leu	Ser	Val	Glu	Asp	Lys	Leu
225				230				235				240			
Lys	Tyr	Ser	Glu	Ala	Phe	Asp	Ser	Phe	Ser	Glu	Ser	Gly	Ser	Tyr	Thr
245				250				255							
Tyr	Asp	Ile	Leu	Gly	Gly	Ser	Pro	Ser	Leu	Gly	Ala	Leu	Val	Phe	Lys
260				265				270							
Asn	Asn	Ser	Gln	Gly	Ser	Ser	Asp	Asp	Asn	Leu	Lys	Asn	Trp	Ile	Gln
275				280				285							
Ser	Ile	Ser	Ser	Met	Pro	Val	Leu	Thr	Lys	Phe	Ile	Asp	Gln	Thr	Ser
290				295				300							
Leu	Met	Pro	Val	Trp	Leu	Leu	Cys	Glu	Asp	Lys	Thr	Lys	Ala	Asp	Ala
305				310				315				320			
Leu	Lys	Lys	Tyr	Tyr	Asp	Asn	Thr	Trp	Ser	Lys	Ser	Gln	Met	Ala	Val

	325		330		335
Ala Ser Leu Arg	Ala Asn Tyr Ile	Asp Glu Leu Thr	Phe Val Leu Gly		
	340		345		350
Asp Asn Ser Asp	Ile Pro Ala Pro	Val Gly Tyr Thr	Lys Val Pro Ile		
	355		360		365
Asp Leu Asn Ser	Asp Ala Gly Gly	Lys Tyr Val Tyr	Leu Xaa Tyr His		
	370		375		380
Glu Ala Gln Phe	Thr Pro Val Asn	Gly Lys Gln Pro	Ile Val Asp Ile		
385		390		395	400
Gln Val Leu Tyr	Gly Ser Gln Met	Pro Ala Pro Gly	Tyr Ile Lys Ile		
	405		410		415
Asp Val Asp Leu	Asn Ser Gly Ala	Gly Gly Glu Phe	Val Tyr Leu Ser		
	420		425		430
Tyr Lys Lys Gly	Glu Pro Thr Ser	Ser Asp Val Ile	Asn Lys Ile Thr		
	435		440		445
Ala Val Tyr Gly	Lys Asn Glu Tyr	Val Pro Thr Pro	Tyr Gly Tyr Lys		
	450		455		460
Gln Ile Ser Gly	Asp Leu Asn Ala	Gly Ala Gly Gly	Asp Phe Val Tyr		
465		470		475	480
Leu Cys Thr Tyr	Gln Gly Gly Thr	Glu			
	485				
<210>	13				
<211>	489				
<212>	PRT				
<213>	人工序列				
<220>					
<223>	SproCRW C482L				
<400>	13				
Met Lys Ile Glu	Ser Ser Lys Val	Glu Gly Leu Phe	Ser Ser Ser Phe		
1		5		10	15
Thr Arg Val Asn	Ala Val Pro Leu	Pro Ser Asp Thr	Leu Pro Gly Val		
	20		25		30
Gly Ile Ile Gly	Cys Gly Tyr Asn	Pro Phe Leu Ala	Tyr Ala Asp Ala		
	35		40		45
Ser Ala Val Leu	His Pro Ile Leu	Asp Trp Ser Lys	Ser Gln Phe Asn		
	50		55		60
Glu Ile Thr Met	Asn Gly Gln Gln	Tyr Gln Leu Pro	Asp Val Leu Gln		
65		70		75	80
Ala Val Trp Leu	Ser Asn Gln Ser	Tyr Ala Ser Val	Thr Gly Lys Ser		

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

Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile
 405 410 415
 Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser
 420 425 430
 Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr
 435 440 445
 Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys
 450 455 460
 Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr
 465 470 475 480
 Leu Leu Thr Tyr Gln Gly Gly Thr Glu
 485

<210> 14

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW V398L

<400> 14

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

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

465	470	475	480
Leu Cys Thr Tyr Gln Gly Gly Thr Glu			
	485		
<210> 15			
<211> 491			
<212> PRT			
<213> 人工序列			
<220>			
<223> SproCRW P396-L-L-I397			
<400> 15			
Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe			
1	5	10	15
Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val			
	20	25	30
Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala			
	35	40	45
Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn			
	50	55	60
Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln			
65	70	75	80
Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser			
	85	90	95
Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly			
	100	105	110
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser			
	115	120	125
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe			
	130	135	140
Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn			
145	150	155	160
Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu			
	165	170	175
Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Val			
	180	185	190
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala			
	195	200	205
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln			
	210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Val Glu Asp Lys Leu			

225	230	235	240
Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr			
	245	250	255
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys			
	260	265	270
Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln			
	275	280	285
Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser			
	290	295	300
Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala			
305	310	315	320
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val			
	325	330	335
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly			
	340	345	350
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile			
	355	360	365
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His			
	370	375	380
Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Leu Leu Ile Val			
385	390	395	400
Asp Ile Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile			
	405	410	415
Lys Ile Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr			
	420	425	430
Leu Ser Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys			
	435	440	445
Ile Thr Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly			
	450	455	460
Tyr Lys Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe			
465	470	475	480
Val Tyr Leu Cys Thr Tyr Gln Gly Gly Thr Glu			
	485	490	

<210> 16

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW I397L

<400> 16

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

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

<210> 17

<211> 489

<212> PRT

<213> 人工序列

<220>

<223> SproCRW I397L V398L

<400> 17

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

Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln			
65	70	75	80
Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser			
	85	90	95
Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly			
	100	105	110
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser			
	115	120	125
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe			
	130	135	140
Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn			
145	150	155	160
Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu			
	165	170	175
Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Val			
	180	185	190
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala			
	195	200	205
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln			
	210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Val Glu Asp Lys Leu			
225	230	235	240
Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr			
	245	250	255
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys			
	260	265	270
Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln			
	275	280	285
Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser			
	290	295	300
Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala			
305	310	315	320
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val			
	325	330	335
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly			
	340	345	350
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile			
	355	360	365
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His			

370	375	380
Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Leu Leu Asp Ile		
385	390	395
Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile		400
	405	410
Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser		415
	420	425
Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr		430
	435	440
Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys		445
	450	455
Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Tyr		460
465	470	475
Leu Cys Thr Tyr Gln Gly Gly Thr Glu		480
	485	
<210>	18	
<211>	489	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	SproCRW Y480L C482L	
<400>	18	
Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe		
1	5	10
Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val		15
	20	25
Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala		30
	35	40
Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn		45
	50	55
Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln		60
65	70	75
Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser		80
	85	90
Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly		95
	100	105
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser		110
	115	120
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe		125

130	135	140
Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn		
145	150	155
Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu		
	165	170
Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Val		
	180	185
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala		
	195	200
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln		
210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Val Glu Asp Lys Leu		
225	230	235
Lys Tyr Ser Glu Ala Phe Asp Ser Phe Ser Glu Ser Gly Ser Tyr Thr		
	245	250
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys		
	260	265
Asn Asn Ser Gln Gly Ser Ser Asp Asp Asn Leu Lys Asn Trp Ile Gln		
	275	280
Ser Ile Ser Ser Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser		
290	295	300
Leu Met Pro Val Trp Leu Leu Cys Glu Asp Lys Thr Lys Ala Asp Ala		
305	310	315
Leu Lys Lys Tyr Tyr Asp Asn Thr Trp Ser Lys Ser Gln Met Ala Val		
	325	330
Ala Ser Leu Arg Ala Asn Tyr Ile Asp Glu Leu Thr Phe Val Leu Gly		
	340	345
Asp Asn Ser Asp Ile Pro Ala Pro Val Gly Tyr Thr Lys Val Pro Ile		
	355	360
Asp Leu Asn Ser Asp Ala Gly Gly Lys Tyr Val Tyr Leu Cys Tyr His		
370	375	380
Glu Ala Gln Phe Thr Pro Val Asn Gly Lys Gln Pro Ile Val Asp Ile		
385	390	395
Gln Val Leu Tyr Gly Ser Gln Met Pro Ala Pro Gly Tyr Ile Lys Ile		
	405	410
Asp Val Asp Leu Asn Ser Gly Ala Gly Gly Glu Phe Val Tyr Leu Ser		
	420	425
Tyr Lys Lys Gly Glu Pro Thr Ser Ser Asp Val Ile Asn Lys Ile Thr		
435	440	445

Ala Val Tyr Gly Lys Asn Glu Tyr Val Pro Thr Pro Tyr Gly Tyr Lys
 450 455 460
 Gln Ile Ser Gly Asp Leu Asn Ala Gly Ala Gly Gly Asp Phe Val Leu
 465 470 475 480
 Leu Leu Thr Tyr Gln Gly Gly Thr Glu
 485
 <210> 19
 <211> 489
 <212> PRT
 <213> 人工序列
 <220>
 <223> SproCRW C482L T483L
 <400> 19
 Met Lys Ile Glu Ser Ser Lys Val Glu Gly Leu Phe Ser Ser Ser Phe
 1 5 10 15
 Thr Arg Val Asn Ala Val Pro Leu Pro Ser Asp Thr Leu Pro Gly Val
 20 25 30
 Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ala
 35 40 45
 Ser Ala Val Leu His Pro Ile Leu Asp Trp Ser Lys Ser Gln Phe Asn
 50 55 60
 Glu Ile Thr Met Asn Gly Gln Gln Tyr Gln Leu Pro Asp Val Leu Gln
 65 70 75 80
 Ala Val Trp Leu Ser Asn Gln Ser Tyr Ala Ser Val Thr Gly Lys Ser
 85 90 95
 Leu Gln Ser Tyr Leu Thr Glu Leu Ala Asn Ser Ile Lys Val Ser Gly
 100 105 110
 Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Ser Asp Ser
 115 120 125
 Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe
 130 135 140
 Asp Leu Trp Ser Ile Ser Ile Pro Ala Asp Ile Ala Arg Leu Gln Asn
 145 150 155 160
 Tyr Val Ser Asp Asp Phe Ile Lys Leu Ile Asn Ala Ile Asn Pro Glu
 165 170 175
 Ser Lys Asp Ser Ile Ala Thr Val Phe Asn Val Tyr Gly Ser His Val
 180 185 190
 Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala
 195 200 205

Asn	Lys	Leu	Thr	Leu	Thr	Gln	Lys	Phe	Glu	Met	Ser	Thr	Ile	Val	Gln
210						215					220				
Ala	Lys	Tyr	Glu	Gln	Leu	Thr	Ser	Gln	Leu	Ser	Val	Glu	Asp	Lys	Leu
225					230					235					240
Lys	Tyr	Ser	Glu	Ala	Phe	Asp	Ser	Phe	Ser	Glu	Ser	Gly	Ser	Tyr	Thr
				245					250					255	
Tyr	Asp	Ile	Leu	Gly	Gly	Ser	Pro	Ser	Leu	Gly	Ala	Leu	Val	Phe	Lys
			260					265					270		
Asn	Asn	Ser	Gln	Gly	Ser	Ser	Asp	Asp	Asn	Leu	Lys	Asn	Trp	Ile	Gln
		275					280					285			
Ser	Ile	Ser	Ser	Met	Pro	Val	Leu	Thr	Lys	Phe	Ile	Asp	Gln	Thr	Ser
290						295					300				
Leu	Met	Pro	Val	Trp	Leu	Leu	Cys	Glu	Asp	Lys	Thr	Lys	Ala	Asp	Ala
305					310					315					320
Leu	Lys	Lys	Tyr	Tyr	Asp	Asn	Thr	Trp	Ser	Lys	Ser	Gln	Met	Ala	Val
				325					330					335	
Ala	Ser	Leu	Arg	Ala	Asn	Tyr	Ile	Asp	Glu	Leu	Thr	Phe	Val	Leu	Gly
			340					345					350		
Asp	Asn	Ser	Asp	Ile	Pro	Ala	Pro	Val	Gly	Tyr	Thr	Lys	Val	Pro	Ile
		355					360					365			
Asp	Leu	Asn	Ser	Asp	Ala	Gly	Gly	Lys	Tyr	Val	Tyr	Leu	Cys	Tyr	His
370						375					380				
Glu	Ala	Gln	Phe	Thr	Pro	Val	Asn	Gly	Lys	Gln	Pro	Ile	Val	Asp	Ile
385					390					395					400
Gln	Val	Leu	Tyr	Gly	Ser	Gln	Met	Pro	Ala	Pro	Gly	Tyr	Ile	Lys	Ile
				405					410					415	
Asp	Val	Asp	Leu	Asn	Ser	Gly	Ala	Gly	Gly	Glu	Phe	Val	Tyr	Leu	Ser
			420					425					430		
Tyr	Lys	Lys	Gly	Glu	Pro	Thr	Ser	Ser	Asp	Val	Ile	Asn	Lys	Ile	Thr
		435					440					445			
Ala	Val	Tyr	Gly	Lys	Asn	Glu	Tyr	Val	Pro	Thr	Pro	Tyr	Gly	Tyr	Lys
		450					455				460				
Gln	Ile	Ser	Gly	Asp	Leu	Asn	Ala	Gly	Ala	Gly	Gly	Asp	Phe	Val	Tyr
465					470					475					480
Leu	Leu	Leu	Tyr	Gln	Gly	Gly	Thr	Glu							
				485											

<210> 20

<211> 488

<212> PRT

<213> 人工序列

<220>

<223> SplyCRW C481L

<400> 20

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

[illegible]

21

488 $\langle 211 \rangle$

<212> PRT

〈213〉 人工序列

 $\langle 220 \rangle$

<223> Sply Y479L C481L

<400> 21

Met Lys Phe Glu Ser Leu Lys Val Glu Asn Leu Phe Ser Ser Ser Leu
1 5 10 15

Thr Arg Ile Asn Ala Ile Pro Leu Pro Asn Asp Ala Leu Pro Gly Val
20 25 30

Gly Ile Ile Gly Cys Gly Tyr Asn Pro Phe Leu Ala Tyr Ala Asp Ser

35	40	45
Ser Ala Val Leu His Pro Val Leu Asp Trp Ser Lys Ser Gln Phe His		
50	55	60
Thr Val Thr Met Asn Gly Gln Thr Tyr Gln Leu Pro Glu Ile Leu Asn		
65	70	75
Ala Val Trp Leu Ser Asn Gln Thr Tyr Ser Ser Val Ser Gly Lys Ser		
85	90	95
Leu Gln Ser Tyr Leu Thr Glu Leu Ser Asn Ser Ile Lys Val Ser Gly		
100	105	110
Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Thr Asp Ser		
115	120	125
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe		
130	135	140
Asp Leu Trp Ser Ile Ser Ala Pro Ala Asp Ile Ala Arg Leu Gln Asn		
145	150	155
Tyr Ile Ser Asp Asp Phe Lys Lys Leu Ile Ser Ser Ile Asp Pro Asn		
165	170	175
Asn Gln Gln Thr Leu Ala Thr Ile Phe Asn Val Tyr Gly Ser His Ile		
180	185	190
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala		
195	200	205
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln		
210	215	220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Ala Glu Asp Lys Leu		
225	230	235
Lys Tyr Ser Glu Ala Phe Glu Ser Phe Ser Glu Ser Gly Ser Tyr Thr		
245	250	255
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys		
260	265	270
Val Asn Asn Gln Asp Ser Asp Asp Asn Leu Arg Lys Trp Ile Asp Ser		
275	280	285
Ile Ser Thr Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser Leu		
290	295	300
Leu Pro Val Trp Thr Leu Cys Glu Asp Gln Val Lys Ala Asp Ala Leu		
305	310	315
Lys Asn Tyr Tyr Asn Asn Thr Trp Ser Gln Leu Gln Met Ala Val Ala		
325	330	335
Arg Leu Arg Ala Asn Tyr Ile Asp Glu Val Thr Phe Val Leu Gly Asp		
340	345	350

Asn Ser Asp Ile Pro Ala Pro Ala Gly Tyr Thr Lys Val Pro Val Asp
 355 360 365
 Leu Asn Ser Gly Ala Gly Gly Lys Phe Ile Tyr Leu Cys Tyr His Glu
 370 375 380
 Ala Gln Phe Thr Pro Val Asn Ser Lys Gln Ala Ile Val Gly Leu Gln
 385 390 395 400
 Val Leu Tyr Gly Lys Gln Glu Pro Ala Pro Asp Tyr Ser Arg Ile Asn
 405 410 415
 Ile Asp Leu Asn Ser Gly Ala Asn Gly Asp Asp Val Tyr Leu Ser Tyr
 420 425 430
 Lys Lys Gly Asp Ala Thr Ser Lys Glu Val Ile Asn Lys Ile Thr Ala
 435 440 445
 Val Tyr Gly Lys Asp Gln Tyr Val Pro Thr Pro Tyr Gly Tyr Lys Gln
 450 455 460
 Ile Pro Gly Asp Leu Asn Ser Gly Ala Gly Gly Asp Phe Val Leu Phe
 465 470 475 480
 Leu Thr Tyr Gln Gly Gly Thr Glu
 485

<210> 22

<211> 488

<212> PRT

<213> 人工序列

<220>

<223> Sply C481L T482L

<400> 22

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

Asn Tyr Gly Phe Phe Ser Ala Ser Ala Thr Asn Glu Phe Thr Asp Ser		
115	120	125
Ser Leu Arg Lys Ser Glu Asn Glu Phe Ser Arg Cys Gln Gln Ser Phe		
130	135	140
Asp Leu Trp Ser Ile Ser Ala Pro Ala Asp Ile Ala Arg Leu Gln Asn		
145	150	155
Tyr Ile Ser Asp Asp Phe Lys Lys Leu Ile Ser Ser Ile Asp Pro Asn		
	165	170
		175
Asn Gln Gln Thr Leu Ala Thr Ile Phe Asn Val Tyr Gly Ser His Ile		
	180	185
		190
Leu Met Ser Gly Val Met Gly Gly Lys Ala His Val Ser Ala Ser Ala		
	195	200
		205
Asn Lys Leu Thr Leu Thr Gln Lys Phe Glu Met Ser Thr Ile Val Gln		
	210	215
		220
Ala Lys Tyr Glu Gln Leu Thr Ser Gln Leu Ser Ala Glu Asp Lys Leu		
225	230	235
		240
Lys Tyr Ser Glu Ala Phe Glu Ser Phe Ser Glu Ser Gly Ser Tyr Thr		
	245	250
		255
Tyr Asp Ile Leu Gly Gly Ser Pro Ser Leu Gly Ala Leu Val Phe Lys		
	260	265
		270
Val Asn Asn Gln Asp Ser Asp Asp Asn Leu Arg Lys Trp Ile Asp Ser		
	275	280
		285
Ile Ser Thr Met Pro Val Leu Thr Lys Phe Ile Asp Gln Thr Ser Leu		
	290	295
		300
Leu Pro Val Trp Thr Leu Cys Glu Asp Gln Val Lys Ala Asp Ala Leu		
305	310	315
		320
Lys Asn Tyr Tyr Asn Asn Thr Trp Ser Gln Leu Gln Met Ala Val Ala		
	325	330
		335
Arg Leu Arg Ala Asn Tyr Ile Asp Glu Val Thr Phe Val Leu Gly Asp		
	340	345
		350
Asn Ser Asp Ile Pro Ala Pro Ala Gly Tyr Thr Lys Val Pro Val Asp		
	355	360
		365
Leu Asn Ser Gly Ala Gly Gly Lys Phe Ile Tyr Leu Cys Tyr His Glu		
	370	375
		380
Ala Gln Phe Thr Pro Val Asn Ser Lys Gln Ala Ile Val Gly Leu Gln		
385	390	395
		400
Val Leu Tyr Gly Lys Gln Glu Pro Ala Pro Asp Tyr Ser Arg Ile Asn		
	405	410
		415
Ile Asp Leu Asn Ser Gly Ala Asn Gly Asp Asp Val Tyr Leu Ser Tyr		

420	425	430
Lys Lys Gly Asp Ala Thr Ser	Lys Glu Val Ile Asn Lys	Ile Thr Ala
435	440	445
Val Tyr Gly Lys Asp Gln Tyr	Val Pro Thr Pro Tyr Gly	Tyr Lys Gln
450	455	460
Ile Pro Gly Asp Leu Asn Ser	Gly Ala Gly Gly Asp Phe	Val Tyr Phe
465	470	475
Leu Leu Tyr Gln Gly Gly Thr	Glu	
485		
<210>	23	
<211>	1470	
<212>	DNA	
<213>	变形斑沙雷氏菌	
<400>	23	
atgaaaattg aatcatccaa ggttgaaggt ttgttttagct catctttttac tcgtgttaat		60
gctgtccctc tcccgagcga tactttaccc ggcgtaggta ttattggttg tggctacaat		120
ccatttttag cctatgcaga tgcttctgct gtccctgcac ctattctcga ctgggtcaaaa		180
tctcagttta acgaaattac tatgaacggc cagcaatata aattacctga tgtgcttcag		240
gcggtatggc tcagtaacca gtcttacgca tcggtaacgg ggaaatcgct ccaaagttat		300
ttaactgaat tggctaacag tattaagta agtggtaact atggcttttt tagtgcctca		360
gccacaaatg aatttagcga cagttcattg agaaagtcag aaaatgagtt tagtcgatgt		420
cagcagtcac ttgatttgtg gtcaattagt atcccagccg atattgcaag gttgcaaaac		480
tatgtttcag atgattttat aaaactgatt aatgcaataa atccagaaaag taaagattca		540
atagccactg tattcaacgt ttatggcagc catgtttctaa tgagtgggtg tatgggcggg		600
aaagctcacg tatccgcctc agcaaacaag ttaacactca cccaaaaatt tgagatgagt		660
acgattgttc aggctaaata cgaacaacta acatcacaac ttagcgtcga agacaaactt		720
aagtattcag aagccttcga tagtttcagc gaaagcggta gttacacata tgatattcta		780
ggcggtagcc cttcttttagg ggcgcttggt tttaaaaata actcgcaagg tagtagtgac		840
gataatttaa aaaattggat acagtctatt tccagtatgc cgggtgtgac caagttcata		900
gaccacacct cgtaaatgcc tgtatggcta ttgtgtgaag acaagacaaa agccgatgca		960
ttaaaaaaat attacgataa tacatgggtca aaatcgacga tggccgttgc cagtctcaga		1020
gcaaaactaca tcgacgagct gacatttgct ctgggtgaca actcagatat ccctgctcca		1080
gtcggataca ctaaagttcc gattgatctt aattctgatg ctgggggtaa atatgtctat		1140
ctttgctatc atgaagcgca gttcacgcca gtaaattggca aacagccgat tgttgatata		1200
caagtgtctc atggtagcca gatgccagcg cctggatata taaaaattga tgttgatttg		1260
aatagcgggtg ctggcgggga atttgtctac ctttcataca aaaagggaga accaacgtcg		1320
tccgatgtaa ttaataaaaat caccgctgta tatggaaaaa atgagtacgt cccaacccca		1380
tatggatata aacaaaattag cggtgacctt aatgctggag caggcgggtga ctttgtttac		1440
ctctgtactt accaaggtgg tacagaataa		1470

<210> 24

<211> 1467

<212> DNA

<213> 普利茅斯沙雷氏菌

<400> 24

```

atgaaatttg aatcattaac agttgaaaac ttattcagct catcattaac tcgcatcaat      60
gccattccac ttcctaacga cgccttacca ggcgtgggta ttattggctg tggttacaat      120
ccattcttgg cttatgcaga ttcgtcagct gtgctacacc ctgtactcga ttgggtctaag      180
tcacagttcc aactgttaac catgaatggg cagacttata agttgccgga aatacttaat      240
gcagtatggc taagcaacca gacgtactct tcagtatcag gtaaactcact tcagagttat      300
ttaactgaac tttctaatag tattaaggtg agtgggaatt atggtttttt tagtgcttca      360
gcaacaaatg aattcactga tagctcgttg agaaaatctg agaatgaatt tagtcgatgt      420
cagcagtcgt ttgatttgtg gtcaataagt gccccggcgg atatcgcaag gttacagaat      480
tacatttcag atgattttta aaaactcatt agttcaattg atccaaataa tcaacaaaca      540
ttggcaacca ttttcaacgt atacggaagt catattttta tgagtgggtg tatggggggg      600
aaagctcacg tttcagcatc agcgaacaag ctaactctca cgcaaaaatt cgaaatgagc      660
acgattgtac aggccaaata tgagcaactg acatcgagc tcagcgctga agataaatta      720
aaatactctg aagcatttga aagtttcagt gagagtggta gctatactta cgatattttta      780
ggtggtagtc catcgctagg agcattggta ttttaaggta acaaccaaga cagtgcacac      840
aacctaagaa aatggataga ttcgatttca accatgcccc tacttactaa gtttatcgac      900
cagacatcgc tactaccggt gtggacgctg tgtgaagatc aggtaaaggc tgatgcgttg      960
aaaaattatt acaacaatac atggtctcaa ttgcaaatgg ctgtagcccc tctgagagcg      1020
aattacattg atgaagtgc ctttgttctt ggtgacaatt cagatatccc cgcacccgct      1080
ggttatacta aagtaccagt tgatctcaat tcaggcgagc ggggtaaatt tatttacctt      1140
tgctaccatg aagcacaatt taccctgtt aactcaaac aggctattgt tggctctccag      1200
gttctctatg gtaagcagga accggctcct gattattcga gaataaatat tgatttgaat      1260
agcggcgcaa atggagatga cgtctacctt tcatataaga aaggatgatc aacttctaag      1320
gaggttaatta ataaaatcac cgctgtttat ggtaaagatc agtacgttcc gactccatat      1380
ggctataaac aaattccccg cgaccttaac tctggtgcgg gaggtgattt tgtttatttc      1440
tgcacatacc aggttgaac cgaatag

```

<210> 25

<211> 1470

<212> DNA

<213> 奎沃斯沙雷氏菌

<400> 25

```

atgaaaattg aatcatccaa ggttgaaggt ttgttttagct catcttttac tcgtgttaat      60
gctgtccctc tcccgagcga tactttaccc ggtgtaggta ttattggttg tggctacaat      120
ccatttttag cctatgcaga tgcttctgct gtctctgcat ctattctcga ctgggtcaaaa      180
tctcagtttc acgaaattac tatgaacggc cagcaatatc aattacctga tgtgtcttcag      240

```

gcggtatggc tcagcaacca gtcttacgca tcggtaacgg ggaaatcgct ccaaagctat	300
ttaactgaat tggctaacag tattaaagta agtggtaact atggcttttt tagtgcctca	360
gccacaaatg aatttagcga cagttcattg agaaagtcag aaaatgagtt tagtcgatgt	420
cagcagtcac ttgactttgt gtcaattagt atccctgccg atattgcaag gttgcaaaac	480
tatgtttcag atgattttat aaaactgatt aatgcaataa atccagaaaag taaagattca	540
ttagccactg tattcaacgt ttatggcagc catgtttctaa tgagtgggtgt tatgggcggg	600
aaagctcacg tatccgcctc agcaaacaag ttaacactca cccaaaaatt tgagatgagt	660
acgattgttc aggctaaata cgaacaacta acatcacaac ttagcgtcga agacaaactt	720
aagtattcag aagccttcga tagtttcagc gaaagcggta gttacacata tgatattcta	780
ggcggtagcc cttcttttagg ggcgctagtt tttaaagcta actcgcaggg tagtagtgac	840
gataatttaa aaagttggat acagtctatt tccagtatgc cgggtgttgac caagttcata	900
gaccaaacct cgtaaatgcc tgtatggcta ttgtgtgaag acaagacaaa agccgatgca	960
ttaaaaaaat attacgataa tacatggtea aaatcgcaga tggccgttgc cagtctcaga	1020
gcaaaactaca tcgacgagct gacattttgtc ctgggtgaca actcagatat ccctgctcca	1080
gtcggataca ctaaagttcc gattgatctt aattctgatg ctgggggtaa atatgtctat	1140
ctttgctatc atgaagcgca gttcacgcca gtaaattggca aacagccgat tgttgatatc	1200
caagtgtctc atggtagcca gatgccagcg cctggatata taaaaattga tgttgatttg	1260
aatagcggtg ctggcgggga atttgtctac ctttcataca aaaagggaga accaacgtcg	1320
tccgatgtaa ttaataaaat caccgctgta tatggaaaaa atgagtacgt cccaacccca	1380
tatggatata aacaaattag cggtgacctt aatgctggag caggcgggtga ctttgtttac	1440
ctctgtactt atcaaggtgg gacagaataa	1470

<210> 26

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> 大肠杆菌优化的SproCRW序列.

<400> 26

atgaaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtagccga cgtagcgcc gttctgcacc cgattctcga ttgggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctcgtcactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctgggtgt aatggggggg	600
aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc	660

accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg 720
 aaatatagtg aagcggtcga ttctgttttcg gagtcctggct cgtataccta cgatatcctg 780
 ggcggttcgc cgtccttagg ggcttttagtt ttcaaaaaca atagtcaggg ctcgtctgac 840
 gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc 900
 gaccagacaa gcctgatgcc cgtttggctg ttgtgctgagg ataaaaccaa agccgatgcg 960
 ctgaaaaaat attacgataa cacctgggtcc aagagccaaa tggccgtcgc aagcctgcgc 1020
 gcgaactaca tcgatgaact gaccttcgtc ctcgggtgata attctgacat tcctgcgccg 1080
 gtaggttata cgaaaagtcct gatcgatctt aattcggatg ccggcggtta atacgtttat 1140
 ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt 1200
 caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg 1260
 aacagcgggg ccggtgggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg 1320
 tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg 1380
 tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga ctctgtttat 1440
 ttgtgtactt atcaaggggg gaccgaataa 1470

<210> 27

<211> 1467

<212> DNA

<213> 人工序列

<220>

<223> 大肠杆菌-优化的Sp1y序列

<400> 27

atgaagttcg aaagccttaa ggtggagaac ctgttttcga gttcactgac ccgatcaat 60
 gccattcccc ttcccaatga cgcattacct ggtgttggga ttattgggtg tggatacaac 120
 ccttttctgg cctatgcgga ctcatctgct gtcttacatc ctgtgttga ctggtcaaag 180
 agtcaatttc atacggttac tatgaacgga caaacatata aattgccgga aatccttaat 240
 gccgtgtggt taagtaacca aacttactca tcggtctctg gaaaaagtct gcaaagttat 300
 ctgaccgagc tttctaattc aatcaaagtt agtggaaatt atggattttt ttccgcgtct 360
 gccacgaacg aattcaccga ttcattccta agaaaatccg aaaatgaatt cagccgttgt 420
 cagcagtcct tcgacttatg gtctataagc gccctgcag atatagcacg cctgcagaat 480
 tatataagtg acgattttta gaagcttatt tcttcgatag atcccaacaa ccagcaaaca 540
 ttggctacaa tattcaacgt gtatggatct catatactga tgagcggcgt catgggcgga 600
 aaagcccacg tctcggcgag tgcgaataag cttaccttaa ctcagaaatt tgagatgtcc 660
 acgattgttc aggcaaaaata cgagcagctt acatcacac tgctcgccga agacaagtta 720
 aagtacagcg aggcgttcga atcattctcc gagtcgggtt cctacaccta tgatatctta 780
 ggaggttcac ccagtctggg agcattgggtc ttcaagggtta acaaccagga ttctgacgat 840
 aatttacgga agtggatcga ttcgatctcc accatgcctg tgttgacgaa atttatcgat 900
 cagaccagtt tattgcctgt ctggacgctt tgcgaagatc aggtgaaagc cgacgccctg 960
 aagaactact ataacaacac atggtcgcag ttgcagatgg ccgttgctag attacgggcc 1020
 aactacattg acgaggtaac gtttgttctt ggcgataact ctgacatccc tgcgccagca 1080

gggtatacta aggtaccagt ggacctgaat tctggggcag gtggtaaatt catttatctt	1140
tgtctatcatg aagcgcagtt tactcccgtg aactctaagc aagccatagt tggcttgcag	1200
gtcttgtacg ggaaacaaga accagcacct gattattcac gtattaacat agatctgaat	1260
tctggagcaa acggggatga tgtctacctt tcttataaaa agggggacgc cacttcgaaa	1320
gaagtcatca acaaaatcac tgcggtgtac gggaaggatc agtacgttcc cacaccgtat	1380
ggctacaagc aaattccggg ggacctgaat tcaggtgcgg gtggcgactt tgtctatctt	1440
tgtacttatc agggaggaac tgaataa	1467
<210> 28	
<211> 1467	
<212> DNA	
<213> 人工序列	
<220>	
<223> 大肠杆菌优化的SquiCRW序列.	
<400> 28	
atgaaaattg aatcttcaaa ggtagaaggg ttattttcat catcctttac gcgtgtcaat	60
gcggttccgc tgccttcgga tactttgccc ggtgtaggca ttataggttg cgggtacaat	120
ccattcctgg cttatgcgga tgcgagcgcc gtgttgccac ccattcctgga ttggtccaaa	180
agccagtttc atgagatcac gatgaatgga cagcaatacc agttaccaga tgttctgcaa	240
gctgtctggt tatcaaatca gtcctacgct tcggtaacgg gcaaatcatt gcaatcttac	300
cttacggaat tagctaactc aatcaaagta tcggggaact atggcttctt cagtgtctagt	360
gccacaaacg agttctcaga ctcttcggtg cggaaaagcg aaaatgaatt tagccgctgc	420
caacaatctt tcgacctgtg gagcatatcc ataccggctg acattgcccc ctttcaaaat	480
tacgtttccg atgatttcat aaagttgatc aatgcgatta atcccagag caaggactcc	540
ttagcgaccg tttttaatgt ctatggttcc cacgtcttaa tgagcggagt catgggaggc	600
aaagcccatg tcagtgcgaag cgctaacaaa ttaactctga cccaaaaatt cgagatgtcc	660
acgatcgtgc aggcgaagta tgaacaactg acctctcaac tgagtgttga agacaagctg	720
aaatacagtg aggcctttga cagcttttcg gagagcggat cgtacactta tgatatactg	780
ggagggtcgc catctctggg agctcttgta ttcaaagcaa attcacaggg gagttccgac	840
gacaatctga agagttggat acagtccatt tcgtctatgc cagtcttaac gaaatttata	900
gatcaaaacta gcctgatgcc cgtgtgggta ctttgtgaag ataaaacaaa ggctgatgcc	960
ttaaaaaaat attatgataa cacttggagc aaatcgcaga tggcagttgc gagcttgcgt	1020
gcaaaattaca ttgatgagtt gacatttgtt ttgggggaca atagcgacat accggcacca	1080
gtaggctaca caaaagtacc gattgatctg aacagtgatg ctggtgggaa gtacgtttat	1140
ctttgttatc atgaagccca atttaccggt gtgaacggta aacagccaat cgtggatatc	1200
caggctcctgt atggctctca aatgccccga cctgggtata tcaagataga tgtggatttg	1260
aactctgggg caggaggtga gttcgtatac ttatcttaca agaaagggga accaacgagc	1320
tcagacgtta ttaacaaaaat cactgcagtc tatggcaaaa atgagtatgt acccacgcct	1380
tacgggtata aacaaaatttc tggcgatctg aacgctggcg caggcgggga tttcgtgtat	1440
ctttgcactt atcaaggtgg gacagag	1467

<210>	29	
<211>	1470	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SproCRW V8A L11I 大肠杆菌优化的序列	
<400>	29	
atgaaaattg agagcagcaa ggctgaaggc attttctcct ctagcttcac acgggtgaat	60	
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120	
ccgtttcttg cgtagccga cgctagcgcc gttctgcacc cgattctcga ttgggtcaaag	180	
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240	
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300	
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttccgctctc	360	
gccacgaacg aatttagcga ctcgctactg cgcaaatcag aaaacgaatt ttcacgtgc	420	
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac	480	
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540	
atgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg	600	
aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc	660	
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720	
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780	
ggcggttcgc cgtccttagg ggctttagtt ttcaaaaaca atagtcaggg ctcgtctgac	840	
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900	
gaccagacaa gcctgatgcc cgtttggtg ttgtgcgagg ataaaaccaa agccgatgcg	960	
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc	1020	
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080	
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat	1140	
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt	1200	
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260	
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320	
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380	
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440	
ttgtgtactt atcaaggggg gaccgaataa	1470	
<210>	30	
<211>	1470	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	SproCRW Y188W 大肠杆菌优化的序列	
<400>	30	

```

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat      60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat      120
ccgtttctgg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttgggtcaaag      180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa      240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatcgct ccagtcatac      300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct      360
gccacgaacg aatttagcga ctcgtcactg cgcaaatcag aaaacgaatt ttcacgctgc      420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac      480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct      540
atcgccaccg tgtttaacgt ctgggggtcc catgttttaa tgtctggtgt aatggggggg      600
aaagcgcacg ttagcgctc ggctaataag cttacctga cccagaaatt tgaaatgtcc      660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg      720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg      780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac      840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc      900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg      960
ctgaaaaaat attacgataa cacctggctc aagagccaaa tggccgtcgc aagcctgcgc      1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg      1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat      1140
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt      1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg      1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaaggga gccgacatcg      1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg      1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat      1440
ttgtgtactt atcaaggggg gaccgaataa      1470

```

<210> 31

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW Y188F 大肠杆菌优化的序列

<400> 31

```

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat      60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat      120
ccgtttctgg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttgggtcaaag      180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa      240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatcgct ccagtcatac      300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct      360
gccacgaacg aatttagcga ctcgtcactg cgcaaatcag aaaacgaatt ttcacgctgc      420

```

cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctttgggtcc catgttttaa tgtctggtgt aatggggggg	600
aaagcgcacg ttagcgccctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctggctc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tctgcgcgcg	1080
gtaggttata cgaaaagtcg gatcgatctt aattcgatg ccggcggtta atacgtttat	1140
ctgtgctatc acgaagccca gtttaccggt gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470

<210> 32

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW Y413W 大肠杆菌优化的序列

<400> 32

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatcgct ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctcgtcactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg	600
aaagcgcacg ttagcgccctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac	840

gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggtg ttgtgagagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat	1140
ctgtgctatc acgaagccca gtttaccggt gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtga ttaagattga tgtagatctg	1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470
<210> 33	
<211> 1470	
<212> DNA	
<213> 人工序列	
<220>	
<223> SproCRW F428W 大肠杆菌优化的序列	
<400> 33	
atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtagccga cgtagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggatTTTT ttcggcctct	360
gccacgaacg aatttagcga ctgcgtactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg	600
aaagcgcacg ttagcgctc ggctaataag cttaccctga ccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctgctctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggtg ttgtgagagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat	1140
ctgtgctatc acgaagccca gtttaccggt gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260

aacagcgggg	ccggtggtga	atgggtgtac	ctgtcctata	aaaaagggga	gccgacatcg	1320
tccgatgtta	ttaacaaaat	caccgcggtt	tatggaaaga	atgaatatgt	accgaccccg	1380
tatggctata	aacagatcag	cggcgacctc	aacgctgggg	cggggggtga	cttcgtttat	1440
ttgtgtactt	atcaaggggg	gaccgaataa				1470

<210> 34

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW Y430W 大肠杆菌优化的序列

<400> 34

atgaaaattg	agagcagcaa	ggtcgaaggc	cttttctcct	ctagcttcac	acgggtgaat	60
gctgtgccgc	ttccgagtga	cacctgccg	ggtgtaggta	tcattggttg	cggttataat	120
ccgtttctgg	cgtatgccga	cgctagcgcc	gttctgcacc	cgattctcga	ttggtcaaag	180
tcacaattta	atgaaattac	catgaatggc	cagcagtacc	aacttcctga	tgctcctcaa	240
gcagtgtggt	taagcaacca	gtcgtacgcc	tctgtgacag	gcaaatcgct	ccagtcatac	300
cttactgaac	tcgcgaactc	tatcaaagtc	tctggcaatt	atggattttt	ttcggcctct	360
gccacgaacg	aatttagcga	ctcgtcactg	cgcaaacag	aaaacgaatt	ttcacgctgc	420
cagcaaagtt	ttgatattatg	gtctatttcg	atccctgctg	atattgcccc	cctccagaac	480
tacgtgtcgg	atgactttat	taaactcatc	aatgcaatta	atccagagtc	caaagattct	540
atcgccaccg	tgtttaacgt	ctacgggtcc	catgttttaa	tgtctggtgt	aatggggggg	600
aaagcgcacg	ttagcgcctc	ggctaataag	cttaccctga	cccagaaatt	tgaaatgtcc	660
accattgtac	aagctaaata	cgaacagctt	accagtcagc	tgagcgtaga	agacaagctg	720
aaatatagtg	aagcgttcga	ttcgttttcg	gagtctggct	cgtataccta	cgatatcctg	780
ggcggttcgc	cgtccttagg	ggcttttagt	ttcaaaaaca	atagtcaggg	ctcgtctgac	840
gataacctta	agaactggat	ccaatctatt	agctcgatgc	cagtacttac	caaattcatc	900
gaccagacaa	gcctgatgcc	cgtttggtg	ttgtgcgagg	ataaaaccaa	agccgatgcg	960
ctgaaaaaat	attacgataa	cacctgggtc	aagagccaaa	tggccgtcgc	aagcctgcgc	1020
gcgaactaca	tcgatgaact	gaccttcgtc	ctcggtgata	attctgacat	tcctgcgccg	1080
gtaggttata	cgaaggtccc	gatcgatctt	aattcggatg	ccggcggtaa	atacgtttat	1140
ctgtgctatc	acgaagccca	gtttaccccg	gtgaacggca	aacagcccat	cgtggatatt	1200
caggttctgt	acggttctca	gatgcctgcg	ccgggggtata	ttaagattga	tgtagatctg	1260
aacagcgggg	ccggtggtga	atgtgtgtgg	ctgtcctata	aaaaagggga	gccgacatcg	1320
tccgatgtta	ttaacaaaat	caccgcggtt	tatggaaaga	atgaatatgt	accgaccccg	1380
tatggctata	aacagatcag	cggcgacctc	aacgctgggg	cggggggtga	cttcgtttat	1440
ttgtgtactt	atcaaggggg	gaccgaataa				1470

<210> 35

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW S190P 大肠杆菌优化的序列

<400> 35

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtagtccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaactcgct ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctctgtactg cgcaaactcg aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggccc catgttttaa tgtctgggtg aatggggggg	600
aaagcgcacg ttagcgccct ggctaataag cttaccctga ccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctctgtctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctggctc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat	1140
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaaggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470

<210> 36

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW V192Y 大肠杆菌优化的序列

<400> 36

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtagtccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180

tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctctgactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc cattacttaa tgtctggtgt aatggggggg	600
aaagcgcacg ttagcgctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctctgtctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctgggtcc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtta atacgtttat	1140
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470

<210> 37

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C312L 大肠杆菌优化的序列

<400> 37

atgaaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctctgactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg	600

```

aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc 660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg 720
aaatatagtg aagcggttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg 780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctctgtctgac 840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc 900
gaccagacaa gcctgatgcc cgtttggtg ttgtcgcagg ataaaaccaa agccgatgcg 960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc 1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgcgc 1080
gtaggttata cgaaagtcgc gatcgatctt aattcggatg ccggcggtaa atacgtttat 1140
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt 1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg 1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg 1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg 1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga ctctgtttat 1440
ttgtgtactt atcaaggggg gaccgaataa 1470

```

<210> 38

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C382A 大肠杆菌优化的序列

<400> 38

```

atgaaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat 60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat 120
ccgtttcttg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag 180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa 240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac 300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct 360
gccacgaacg aatttagcga ctctgactg cgcaaatecgt aaaacgaatt ttcacgtgc 420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac 480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct 540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg 600
aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc 660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg 720
aaatatagtg aagcggttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg 780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctctgtctgac 840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc 900
gaccagacaa gcctgatgcc cgtttggtg ttgtcgcagg ataaaaccaa agccgatgcg 960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc 1020

```

gcgaactaca	tcgatgaact	gaccttcgtc	ctcggtgata	attctgacat	tcctgcgccg	1080
gtaggttata	cgaaagtccc	gatcgatctt	aattcggatg	ccggcggtaa	atacgtttat	1140
ctggcctatc	acgaagccca	gtttaccccc	gtgaacggca	aacagcccat	cgtggatatt	1200
caggttctgt	acggttctca	gatgcctgcg	ccgggggtata	ttaagattga	tgtagatctg	1260
aacagcgggg	ccggtggtga	atttgtgtac	ctgtcctata	aaaaagggga	gccgacatcg	1320
tccgatgtta	ttaacaaaat	caccgcggtt	tatggaaaga	atgaatatgt	accgaccccc	1380
tatggctata	aacagatcag	cggcgacctc	aacgctgggg	cgggggggtga	cttcgtttat	1440
ttgtgtactt	atcaaggggg	gaccgaataa				1470

<210> 39

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C382S 大肠杆菌优化的序列

<400> 39

atgaaaattg	agagcagcaa	ggtcgaaggc	cttttctcct	ctagcttcac	acgggtgaat	60
gctgtgccgc	ttccgagtga	cacctgccc	ggtgtaggta	tcattggttg	cggttataat	120
ccgtttctgg	cgtatgccga	cgctagcgcc	gttctgcacc	cgattctcga	ttggtcaaag	180
tcacaattta	atgaaattac	catgaatggc	cagcagtacc	aacttcctga	tgctcctcaa	240
gcagtgtggt	taagcaacca	gtcgtacgcc	tctgtgacag	gcaaatecgt	ccagtcatac	300
cttactgaac	tcgcgaactc	tatcaaagtc	tctggcaatt	atggattttt	ttcggcctct	360
gccacgaacg	aatttagcga	ctcgtcactg	cgcaaatcag	aaaacgaatt	ttcacgctgc	420
cagcaaagtt	ttgatattatg	gtctatttcg	atccctgctg	atattgcccc	cctccagaac	480
tacgtgtcgg	atgactttat	taaactcatc	aatgcaatta	atccagagtc	caaagattct	540
atcgccaccg	tgtttaacgt	ctacgggtcc	catgttttaa	tgtctggtgt	aatggggggg	600
aaagcgcacg	ttagcgcctc	ggctaataag	cttacctga	cccagaaatt	tgaaatgtcc	660
accattgtac	aagctaaata	cgaacagctt	accagtcagc	tgagcgtaga	agacaagctg	720
aaatatagtg	aagcgttcga	ttcgttttcg	gagtctggct	cgtataccta	cgatatcctg	780
ggcggttcgc	cgtccttagg	ggcttttagt	ttcaaaaaca	atagtcaggg	ctcgtctgac	840
gataacctta	agaactggat	ccaatctatt	agctcgatgc	cagtacttac	caaattcatc	900
gaccagacaa	gcctgatgcc	cgtttggtcg	ttgtgcgagg	ataaaaccaa	agccgatgcg	960
ctgaaaaaat	attacgataa	cacctgggtc	aagagccaaa	tggccgtcgc	aagcctgcgc	1020
gcgaactaca	tcgatgaact	gaccttcgtc	ctcggtgata	attctgacat	tcctgcgccg	1080
gtaggttata	cgaaagtccc	gatcgatctt	aattcggatg	ccggcggtaa	atacgtttat	1140
ctgagttatc	acgaagccca	gtttaccccc	gtgaacggca	aacagcccat	cgtggatatt	1200
caggttctgt	acggttctca	gatgcctgcg	ccgggggtata	ttaagattga	tgtagatctg	1260
aacagcgggg	ccggtggtga	atttgtgtac	ctgtcctata	aaaaagggga	gccgacatcg	1320
tccgatgtta	ttaacaaaat	caccgcggtt	tatggaaaga	atgaatatgt	accgaccccc	1380
tatggctata	aacagatcag	cggcgacctc	aacgctgggg	cgggggggtga	cttcgtttat	1440

ttgtgtactt atcaaggggg gaccgaataa	1470
<210> 40	
<211> 1470	
<212> DNA	
<213> 人工序列	
<220>	
<223> SproCRW C382T 大肠杆菌优化的序列	
<400> 40	
atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtatgccga cgtagcgcc gttctgcacc cgattctcga ttgggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctctgtcactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctgggtg aatggggggg	600
aaagcgcacg ttagcgcctc ggctaataag cttaccctga ccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggctttagtt ttcaaaaaca atagtcaggg ctcgtctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctggctc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat	1140
ctgacgtatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtgggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgtgggg cggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470
<210> 41	
<211> 1470	
<212> DNA	
<213> 人工序列	
<220>	
<223> SproCRW C382M 大肠杆菌优化的序列	

<400> 41

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctcgctactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctgggtg aatggggggg	600
aaagcgcacg ttagcgcctc ggctaataag cttacctga cccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatacctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctgggtcc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtta atacgtttat	1140
ctgatgtatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtgggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470

<210> 42

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C382G 大肠杆菌优化的序列

<400> 42

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttcttg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360

```

gccacgaacg aatttagcga ctctgactg cgcaaatcag aaaacgaatt ttcacgtgc 420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac 480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct 540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg 600
aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc 660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg 720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatacctg 780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctctgtctgac 840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc 900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg 960
ctgaaaaaat attacgataa cacttggtcc aagagccaaa tggccgtcgc aagcctgcgc 1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcttgcgccg 1080
gtaggttata cgaaagtcgc gatcgatctt aattcggatg ccggcggtaa atacgtttat 1140
ctgggctatc acgaagccca gtttaccctg gtgaacggca aacagcccat cgtggatatt 1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg 1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaaggga gccgacatcg 1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg 1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat 1440
ttgtgtactt atcaaggggg gaccgaataa 1470

```

<210> 43

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C382F 大肠杆菌优化的序列

<400> 43

```

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat 60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat 120
ccgtttctgg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag 180
tcacaattta atgaaattac catgaatggc cagcagtagc aacttcctga tgtccttcaa 240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatacgt ccagtcatac 300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct 360
gccacgaacg aatttagcga ctctgactg cgcaaatcag aaaacgaatt ttcacgtgc 420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac 480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct 540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg 600
aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc 660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg 720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatacctg 780

```

ggcggttcgc	cgtccttagg	ggcttttagtt	ttcaaaaaca	atagtcaggg	ctcgtctgac	840
gataacctta	agaactggat	ccaatctatt	agctcgatgc	cagtacttac	caaattcatc	900
gaccagacaa	gcctgatgcc	cgtttggctg	ttgtgcgagg	ataaaaccaa	agccgatgcg	960
ctgaaaaaat	attacgataa	cacctgggtcc	aagagccaaa	tggccgctgc	aagcctgcgc	1020
gcgaactaca	tcgatgaact	gaccttcgtc	ctcggtgata	attctgacat	tcctgcgccg	1080
gtaggttata	cgaaagtccc	gatcgatctt	aattcggatg	ccggcggtaa	atacgtttat	1140
ctgtttttatc	acgaagccca	gtttaccccc	gtgaacggca	aacagcccat	cgtggatatt	1200
caggtttctgt	acggttctca	gatgcctgcg	ccgggggtata	ttaagattga	tgtagatctg	1260
aacagcgggg	ccggtgggtga	atttgtgtac	ctgtcctata	aaaaagggga	gccgacatcg	1320
tccgatgtta	ttaacaaaaat	caccgcggtt	tatggaaaga	atgaatatgt	accgaccccc	1380
tatggctata	aacagatcag	cggcgacctc	aacgctgggg	cggggggtga	cttcgtttat	1440
ttgtgtactt	atcaaggggg	gaccgaataa				1470

<210> 44

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C482L 大肠杆菌优化的序列

<400> 44

atgaaaattg	agagcagcaa	ggtcgaaggc	cttttctcct	ctagcttcac	acgggtgaat	60
gctgtgccgc	ttccgagtga	cacctgccc	gggtgtaggt	tcattgggtg	cggttataat	120
ccgtttctgg	cgtagtccga	cgctagcgcc	gttctgcacc	cgattctcga	ttgggtcaaag	180
tcacaattta	atgaaattac	catgaatggc	cagcagtacc	aacttcctga	tgctccttcaa	240
gcagtgtggt	taagcaacca	gtcgtacgcc	tctgtgacag	gcaaatcgct	ccagtcatac	300
cttactgaac	tcgcgaactc	tatcaaagtc	tctggcaatt	atggattttt	ttcggcctct	360
gccacgaacg	aatttagcga	ctcgtcactg	cgcaaatcag	aaaacgaatt	ttcacgctgc	420
cagcaaagtt	ttgatttatg	gtctatttcg	atccctgctg	atattgcccc	cctccagaac	480
tacgtgtcgg	atgactttat	taaactcatc	aatgcaatta	atccagagtc	caaagattct	540
atcgccaccg	tgtttaacgt	ctacgggtcc	catgttttaa	tgtctgggtg	aatggggggg	600
aaagcgcacg	ttagcgcctc	ggctaataag	cttacctga	cccagaaatt	tgaaatgtcc	660
accattgtac	aagctaaata	cgaacagctt	accagtcagc	tgagcgtaga	agacaagctg	720
aaatatagtg	aagcgttcga	ttcgttttcg	gagtctggct	cgtataccta	cgatatcctg	780
ggcggttcgc	cgtccttagg	ggcttttagtt	ttcaaaaaca	atagtcaggg	ctcgtctgac	840
gataacctta	agaactggat	ccaatctatt	agctcgatgc	cagtacttac	caaattcatc	900
gaccagacaa	gcctgatgcc	cgtttggctg	ttgtgcgagg	ataaaaccaa	agccgatgcg	960
ctgaaaaaat	attacgataa	cacctgggtcc	aagagccaaa	tggccgctgc	aagcctgcgc	1020
gcgaactaca	tcgatgaact	gaccttcgtc	ctcggtgata	attctgacat	tcctgcgccg	1080
gtaggttata	cgaaagtccc	gatcgatctt	aattcggatg	ccggcggtaa	atacgtttat	1140
ctgtgctatc	acgaagccca	gtttaccccc	gtgaacggca	aacagcccat	cgtggatatt	1200

caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cgggggggtga cttcgtttat	1440
ttgcttactt atcaaggggg gaccgaataa	1470
<210> 45	
<211> 1470	
<212> DNA	
<213> 人工序列	
<220>	
<223> SproCRW V398L 大肠杆菌优化的序列	
<400> 45	
atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat	60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat	120
ccgtttctgg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag	180
tcacaattta atgaaaattac catgaatggc cagcagtacc aacttcctga tgccttcaa	240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatcgct ccagtcatac	300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct	360
gccacgaacg aatttagcga ctcgctactg cgcaaatcag aaaacgaatt ttcacgtgc	420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccc cctccagaac	480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct	540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg	600
aaagcgcacg ttagcgctc ggctaataag cttaccctga ccagaaatt tgaaatgtcc	660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg	720
aaatatagtg aagcggtcga ttcgttttcg gagtctggct cgtataccta cgatatcctg	780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac	840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc	900
gaccagacaa gcctgatgcc cgtttggtg ttgtgcgagg ataaaaccaa agccgatgcg	960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc	1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg	1080
gtaggttata cgaaaagtc gatcgatctt aattcggatg ccggcggtaa atacgtttat	1140
ctgtgctatc acgaagccca gtttaccggt gtgaacggca aacagcccat cctggatatt	1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg	1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg	1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg	1380
tatggctata aacagatcag cggcgacctc aacgctgggg cgggggggtga cttcgtttat	1440
ttgtgtactt atcaaggggg gaccgaataa	1470
<210> 46	
<211> 1476	

<212> DNA

<213> 人工序列

<220>

<223> SproCRW P396-L-L-I397 大肠杆菌优化的序列

<400> 46

```

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat      60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat      120
ccgtttcttg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttgggtcaaag      180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa      240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatcgct ccagtcatac      300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct      360
gccacgaacg aatttagcga ctcgtaactg cgcaaatcag aaaacgaatt ttcacgctgc      420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac      480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct      540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg      600
aaagcgcacg ttagcgcctc ggctaataag cttaccctga cccagaaatt tgaaatgtcc      660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg      720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg      780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac      840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc      900
gaccagacaa gcctgatgcc cgtttggtg ttgtgcgagg ataaaaccaa agccgatgcg      960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc     1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg     1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat     1140
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccct gttgatcgtg     1200
gatattcagg ttctgtacgg ttctcagatg cctgcgccgg ggtatattaa gattgatgta     1260
gatctgaaca gcggggccgg tgggtgaatt gtgtacctgt cctataaaaa aggggagccg     1320
acatcgtccg atgttattaa caaaatcacc gcggtttatg gaaagaatga atatgtaccg     1380
accccgtatg gctataaaca gatcagcggc gacctcaacg ctggggcggg gggtgacttc     1440
gtttatttgt gtacttatca aggggggacc gaataa                                1476

```

<210> 47

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW I397L 大肠杆菌优化的序列

<400> 47

```

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat      60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat      120

```

```

ccgtttctgg cgatatgccga cgctagcgcc gttctgcacc cgattctcga ttgggtcaaag 180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa 240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac 300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct 360
gccacgaacg aatttagcga ctcgtcactg cgcaaatcag aaaacgaatt ttcacgtgc 420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac 480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct 540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctgggtg aatggggggg 600
aaagcgcacg ttagcgccct ggctaataag cttaccctga ccagaaatt tgaaatgtcc 660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg 720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg 780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctcgtctgac 840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc 900
gaccagacaa gcctgatgcc cgtttggctg ttgtgcgagg ataaaaccaa agccgatgcg 960
ctgaaaaaat attacgataa cacctgggtc aagagccaaa tggccgtcgc aagcctgcgc 1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg 1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat 1140
ctgtgctatc acgaagccca gtttaccctg gtgaacggca aacagcccct ggtggatatt 1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg 1260
aacagcgggg ccggtgggtg atttgtgtac ctgtcctata aaaaagggga gccgacatcg 1320
tccgatgtta ttaacaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccg 1380
tatggctata aacagatcag cggcgacctc aacgctgggg cgggggggtg cttegtttat 1440
ttgtgtactt atcaaggggg gaccgaataa 1470

```

<210> 48

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW I397L V398L 大肠杆菌优化的序列

<400> 48

```

atgaaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat 60
gctgtgccgc ttccgagtga cacctgccg ggtgtaggta tcattggttg cggttataat 120
ccgtttctgg cgatatgccga cgctagcgcc gttctgcacc cgattctcga ttgggtcaaag 180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgtccttcaa 240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac 300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggattttt ttcggcctct 360
gccacgaacg aatttagcga ctcgtcactg cgcaaatcag aaaacgaatt ttcacgtgc 420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccg cctccagaac 480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct 540

```

atcgccaccg	tgtttaacgt	ctacgggtcc	catgttttaa	tgtctggtgt	aatggggggg	600
aaagcgcacg	ttagcgctc	ggctaataag	cttaccctga	cccagaaatt	tgaaatgtcc	660
accattgtac	aagctaaata	cgaacagctt	accagtcagc	tgagcgtaga	agacaagctg	720
aaatatagtg	aagcgttcga	ttcgttttcg	gagtcctggc	cgtataccta	cgatatcctg	780
ggcggttcgc	cgtccttagg	ggcttttagt	ttcaaaaaca	atagtcaggg	ctcgtctgac	840
gataacctta	agaactggat	ccaatctatt	agctcgatgc	cagtacttac	caaattcatc	900
gaccagacaa	gcctgatgcc	cgtttggtcg	ttgtgcgagg	ataaaaccaa	agccgatgcg	960
ctgaaaaaat	attacgataa	cacctgggtc	aagagccaaa	tggccgtcgc	aagcctgcgc	1020
gcgaactaca	tcgatgaact	gaccttcgtc	ctcggtgata	attctgacat	tcctgcgccg	1080
gtaggttata	cgaaagtccc	gatcgatctt	aattcggatg	ccggcggtaa	atacgtttat	1140
ctgtgctatc	acgaagccca	gtttaccccg	gtgaacggca	aacagcccct	gttgatattt	1200
caggttctgt	acggttctca	gatgcctgcg	ccgggggtata	ttaagattga	tgtagatctg	1260
aacagcgggg	ccggtggtga	atttgtgtac	ctgtcctata	aaaaagggga	gccgacatcg	1320
tccgatgtta	ttaacaaaat	caccgcggtt	tatggaaaga	atgaatatgt	accgaccccg	1380
tatggctata	aacagatcag	cggcgacctc	aacgctgggg	cggggggtga	cttcgtttat	1440
ttgtgtactt	atcaaggggg	gaccgaataa				1470

<210> 49

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW Y480L C482L 大肠杆菌优化的序列:

<400> 49

atgaaaattg	agagcagcaa	ggtcgaaggc	cttttctcct	ctagcttcac	acgggtgaat	60
gctgtgccgc	ttccgagtga	cacctgccc	ggtgtaggta	tcattggttg	cggttataat	120
ccgtttcttg	cgtagtccga	cgtagcgcc	gttctgcacc	cgattctcga	ttggtcaaag	180
tcacaattta	atgaaattac	catgaatggc	cagcagtacc	aacttcctga	tgtccttcaa	240
gcagtgtggt	taagcaacca	gtcgtacgcc	tctgtgacag	gcaaatcgct	ccagtcatac	300
cttactgaac	tcgcgaactc	tatcaaagtc	tctggcaatt	atggattttt	ttcggcctct	360
gccacgaacg	aatttagcga	ctcgtcactg	cgcaaatacg	aaaacgaatt	ttcacgtgc	420
cagcaaaagt	ttgatttatg	gtctatttcg	atccctgctg	atattgccc	cctccagaac	480
tacgtgtcgg	atgactttat	taaactcatc	aatgcaatta	atccagagtc	caaagattct	540
atcgccaccg	tgtttaacgt	ctacgggtcc	catgttttaa	tgtctggtgt	aatggggggg	600
aaagcgcacg	ttagcgctc	ggctaataag	cttaccctga	cccagaaatt	tgaaatgtcc	660
accattgtac	aagctaaata	cgaacagctt	accagtcagc	tgagcgtaga	agacaagctg	720
aaatatagtg	aagcgttcga	ttcgttttcg	gagtcctggc	cgtataccta	cgatatcctg	780
ggcggttcgc	cgtccttagg	ggcttttagt	ttcaaaaaca	atagtcaggg	ctcgtctgac	840
gataacctta	agaactggat	ccaatctatt	agctcgatgc	cagtacttac	caaattcatc	900
gaccagacaa	gcctgatgcc	cgtttggtcg	ttgtgcgagg	ataaaaccaa	agccgatgcg	960

```

ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc      1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg      1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat      1140
ctgtgctatc acgaagccca gtttaccceg gtgaacggca aacagcccat cgtggatatt      1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg      1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg      1320
tccgatgtta ttaacaaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccc      1380
tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgttctg      1440
ttgcttactt atcaaggggg gaccgaataa                                     1470

```

<210> 50

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW C482L T483L 大肠杆菌优化的序列:

<400> 50

```

atgaaaattg agagcagcaa ggtcgaaggc cttttctcct ctagcttcac acgggtgaat      60
gctgtgccgc ttccgagtga caccctgccg ggtgtaggta tcattggttg cggttataat      120
ccgtttctgg cgtatgccga cgctagcgcc gttctgcacc cgattctcga ttggtcaaag      180
tcacaattta atgaaattac catgaatggc cagcagtacc aacttcctga tgctcttcaa      240
gcagtgtggt taagcaacca gtcgtacgcc tctgtgacag gcaaatecgt ccagtcatac      300
cttactgaac tcgcgaactc tatcaaagtc tctggcaatt atggatTTTT ttcggcctct      360
gccacgaacg aatttagcga ctctgctact cgcaaactcg aaaacgaatt ttcacgtgc      420
cagcaaagtt ttgatttatg gtctatttcg atccctgctg atattgcccc cctccagaac      480
tacgtgtcgg atgactttat taaactcatc aatgcaatta atccagagtc caaagattct      540
atcgccaccg tgtttaacgt ctacgggtcc catgttttaa tgtctggtgt aatggggggg      600
aaagcgcacg ttagcgccct ggctaataag ctaccctga ccagaaatt tgaaatgtcc      660
accattgtac aagctaaata cgaacagctt accagtcagc tgagcgtaga agacaagctg      720
aaatatagtg aagcgttcga ttcgttttcg gagtctggct cgtataccta cgatatcctg      780
ggcggttcgc cgtccttagg ggcttttagt ttcaaaaaca atagtcaggg ctctgtctgac      840
gataacctta agaactggat ccaatctatt agctcgatgc cagtacttac caaattcatc      900
gaccagacaa gcctgatgcc cgtttggtg ttgtgcgagg ataaaaccaa agccgatgcg      960
ctgaaaaaat attacgataa cacctggtcc aagagccaaa tggccgtcgc aagcctgcgc      1020
gcgaactaca tcgatgaact gaccttcgtc ctcggtgata attctgacat tcctgcgccg      1080
gtaggttata cgaaagtccc gatcgatctt aattcggatg ccggcggtaa atacgtttat      1140
ctgtgctatc acgaagccca gtttaccceg gtgaacggca aacagcccat cgtggatatt      1200
caggttctgt acggttctca gatgcctgcg ccgggggtata ttaagattga tgtagatctg      1260
aacagcgggg ccggtggtga atttgtgtac ctgtcctata aaaaagggga gccgacatcg      1320
tccgatgtta ttaacaaaaat caccgcggtt tatggaaaga atgaatatgt accgaccccc      1380

```

tatggctata aacagatcag cggcgacctc aacgctgggg cggggggtga cttcgtttat	1440
ttgcttctgt atcaaggggg gaccgaataa	1470
<210> 51	
<211> 1467	
<212> DNA	
<213> 人工序列	
<220>	
<223> SplyCRW C481L 大肠杆菌优化的序列:	
<400> 51	
atgaagttcg aaagccctaa ggtggagaac ctgttttcga gttcactgac ccggatcaat	60
gccattcccc ttcccaatga cgcattacct ggtgttggga ttattgggtg tggatacaac	120
ccttttcttg cctatgcgga ctcatctgct gtcttacatc ctgtgttggga ctgggtcaaag	180
agtcaatttc atacggttac tatgaacgga caaacatata aattgccgga aatccctaat	240
gccgtgtggt taagtaacca aacttactca tcggtctctg gaaaaagtct gcaaagttat	300
ctgaccgagc tttctaattc aatcaaagtt agtggaaatt atggattttt ttccgcgtct	360
gccacgaacg aattcaccga ttcatcctta agaaaatccg aaaatgaatt cagccgttgt	420
cagcagtctt tcgacttatg gtctataagc gccctgcag atatagcacg cctgcagaat	480
tatataagtg acgattttta gaagcttatt tcttcgatag atcccaacaa ccagcaaaca	540
ttggctacaa tattcaacgt gtatggatct catatactga tgagcggcgt catgggcgga	600
aaagcccacg tctcggcgag tgcgaataag cttaccttaa ctcagaaatt tgagatgtcc	660
acgattgttc aggcaaaaata cgagcagctt acatcacaaac tgtcggccga agacaagtta	720
aagtacagcg aggcgttcga atcattctcc gagtccggtt cctacaccta tgatatctta	780
ggaggttcac ccagtctggg agcattgggtc ttcaaggta acaaccagga ttctgacgat	840
aatttacgga agtggatcga ttcgatctcc accatgcctg tgttgacgaa atttatcgat	900
cagaccagtt tattgcctgt ctggacgctt tgccaagatc aggtgaaagc cgacgccctg	960
aagaactact ataacaacac atggtcgcag ttgcagatgg ccgttgctag attacgggcc	1020
aactacattg acgaggtaac gtttgttctt ggcgataact ctgacatccc tgcgccagca	1080
gggtatacta aggtaccagt ggacctgaat tctggggcag gtggtaaatt catttatctt	1140
tgctatcatg aagcgcagtt tactcccgtg aactctaagc aagccatagt tggcttgcag	1200
gtcttgtagc ggaaacaaga accagcacct gattattcac gtattaacat agatctgaat	1260
tctggagcaa acggggatga tgtctacctt tcttataaaa agggggacgc cacttcgaaa	1320
gaagtcatca acaaaaatcac tgcggtgtac gggaaggatc agtacgttcc cacaccgtat	1380
ggctacaagc aaattccggg ggacctgaat tcaggtgcgg gtggcgactt tgtctatttt	1440
ttgacttatc agggaggaac tgaataa	1467
<210> 52	
<211> 1467	
<212> DNA	
<213> 人工序列	
<220>	

<223> SplyCRW Y479L C481L 大肠杆菌优化的序列:

<400> 52

```

atgaagttcg aaagccttaa ggtggagaac ctgttttcga gttcactgac ccgatcaat      60
gccattcccc ttcccaatga cgcattacct ggtgttggga ttattgggtg tggatacaac      120
ccttttctgg cctatgcgga ctcatctgct gtcttacatc ctgtgttga ctggtcaaag      180
agtcaatttc atacggttac tatgaacgga caaacatatac aattgccgga aatccttaat      240
gccgtgtggt taagtaacca aacttactca tcggtctctg gaaaaagtct gcaaagttat      300
ctgaccgagc tttctaattc aatcaaagtt agtggaaatt atggattttt ttccgcgtct      360
gccacgaacg aattcaccga ttcactctta agaaaatccg aaaatgaatt cagccgttgt      420
cagcagtctt tcgacttatg gtctataagc gccctgcag atatagcacg cctgcagaat      480
tatataagtg acgattttta gaagcttatt tcttcgatag atcccaacaa ccagcaaaca      540
ttggctacaa tattcaacgt gtatggatct catatactga tgagcggcgt catgggcgga      600
aaagcccacg tctcggcgag tgcgaataag cttaccttaa ctcaaaaatt tgagatgtcc      660
acgattgttc aggcaaaata cgagcagctt acatcacac tgctcgccga agacaagtta      720
aagtacagcg aggcgttcga atcattctcc gagtccggtt cctacaccta tgatatctta      780
ggaggttcac ccagtctggg agcattggtc ttcaaggtta acaaccagga ttctgacgat      840
aatttacgga agtggatcga ttcgatctcc accatgcctg tgttgacgaa atttatcgat      900
cagaccagtt tattgcctgt ctggacgctt tgcgaagatc aggtgaaagc cgacgccctg      960
aagaactact ataacaacac atggtcgcag ttgcagatgg ccgttgctag attacgggcc      1020
aactacattg acgaggtaac gtttgttctt ggcgataact ctgacatccc tgcgccagca      1080
gggtatacta aggtaccagt ggacctgaat tctggggcag gtggtaaatt catttatctt      1140
tgctatcatg aagcgcagtt tactcccgtg aactctaagc aagccatagt tggcttgcag      1200
gtcttgtacg ggaaacaaga accagcacct gattattcac gtattaacat agatctgaat      1260
tctggagcaa acggggatga tgtctacctt tcttataaaa agggggacgc cacttcgaaa      1320
gaagtcatca acaaaatcac tgcggtgtac gggaaggatc agtacgttcc cacaccgtat      1380
ggctacaagc aaattccggg ggacctgaat tcaggtgcgg gtggcgactt tgtccttttt      1440
ttgacttatc agggaggaac tgaataa      1467

```

<210> 53

<211> 1467

<212> DNA

<213> 人工序列

<220>

<223> SplyCRW C481L T482L 大肠杆菌优化的序列:

<400> 53

```

atgaagttcg aaagccttaa ggtggagaac ctgttttcga gttcactgac ccgatcaat      60
gccattcccc ttcccaatga cgcattacct ggtgttggga ttattgggtg tggatacaac      120
ccttttctgg cctatgcgga ctcatctgct gtcttacatc ctgtgttga ctggtcaaag      180
agtcaatttc atacggttac tatgaacgga caaacatatac aattgccgga aatccttaat      240
gccgtgtggt taagtaacca aacttactca tcggtctctg gaaaaagtct gcaaagttat      300

```

ctgaccgagc	tttctaattc	aatcaaagtt	agtggaaatt	atggattttt	ttccgcgtct	360
gccacgaacg	aattcaccga	ttcatcctta	agaaaatccg	aaaatgaatt	cagccgttgt	420
cagcagtcctt	tcgacttatg	gtctataagc	gccctgcag	atatagcacg	cctgcagaat	480
tatataagtg	acgattttta	gaagcttatt	tcttcgatag	atcccaacaa	ccagcaaaca	540
ttggctacaa	tattcaacgt	gtatggatct	catatactga	tgagcggcgt	catgggcgga	600
aaagcccacg	tctcggcgag	tgcaataaag	cttaccttaa	ctcagaaatt	tgagatgtcc	660
acgattgttc	aggcaaaaata	cgagcagctt	acatcacaaac	tgctcgccga	agacaagtta	720
aagtacacgcg	aggcggttcga	atcattctcc	gagtcgggtt	cctacaccta	tgatatctta	780
ggaggttcac	ccagtctggg	agcattggtc	ttcaagggtta	acaaccagga	ttctgacgat	840
aattttacgga	agtggatcga	ttcgatctcc	accatgcctg	tgttgacgaa	atttatcgat	900
cagaccagtt	tattgcctgt	ctggacgctt	tgcaagatc	aggtgaaagc	cgacgccttg	960
aagaactact	ataacaacac	atggtcgcag	ttgcagatgg	ccgttgctag	attacgggcc	1020
aactacattg	acgaggtaac	gtttgttctt	ggcgataact	ctgacatccc	tgcccgagca	1080
gggtatacta	aggtaccagt	ggacctgaat	tctggggcag	gtggtaaatt	catttatctt	1140
tgctatcatg	aagcgcagtt	tactcccggtg	aactctaagc	aagccatagt	tggttcgcag	1200
gtcttgtacg	ggaaacaaga	accagcacct	gattattcac	gtattaacat	agatctgaat	1260
tctggagcaa	acggggatga	tgtctacctt	tcttataaaa	agggggacgc	cacttcgaaa	1320
gaagtcatca	acaaaatcac	tgcggtgtac	gggaaggatc	agtacgttcc	cacaccgtat	1380
ggctacaagc	aaattccggg	ggacctgaat	tcaggtgcgg	gtggcgactt	tgtctatctt	1440
ttgctttatc	agggaggaac	tgaataa				1467

<210> 54

<211> 1470

<212> DNA

<213> 人工序列

<220>

<223> SproCRW V8A L11I 玉蜀黍优化的序列

<400> 54

atgaagattg	agtccagcaa	ggccgagggc	atcttctcgt	cgctggtcac	ccgcgttaat	60
gccgttccac	tcccatcgga	tacactgccg	ggcgtcggga	tcattggctg	cgggtacaac	120
cccttcctcg	cttacgccga	cgcttcggcg	gtgctccacc	cgattctgga	ttgggtctaag	180
tcacagttca	acgagatcac	catgaacggc	cagcagtacc	agctccccga	cgttctgcag	240
gctgtgtggc	tctcaaaacca	gtcctacgcc	agcgtaaccg	gcaagtcgct	gcagtcttac	300
ctcacggagc	tggccaactc	tatcaagggtg	tcaggcaatt	acgggttctt	ctccgcgagc	360
gctactaacg	agttcagcga	ctccagctc	cgcaagtccg	agaatgagtt	cagcaggtgc	420
cagcagtcgt	tcgacctgtg	gtcgatctct	attccagccg	atattgcgcg	gctccagaac	480
tacgtctccg	acgatttcat	caagctgatt	aacgcgatca	atcctgagtc	aaaggactcc	540
atcgtaccg	tgttcaacgt	ctacggctct	cacgtgctca	tgtcaggggt	catgggcggg	600
aaggctcatg	tgtcagcttc	cgcaataaag	ctcacctga	cgcagaagtt	cgagatgtcc	660
accatcgttc	aggccaagta	cgagcagctc	acgagccagc	tgctcggtgga	ggacaagctc	720

aagtactccg aggcgttcga tagcttctcg gagtctggct catacaccta cgacatcctc	780
ggcggctccc cgagcctggg cgcgctgggtg ttcaagaaca attcgcaggg gtcgtctgac	840
gataacctga agaattggat tcagtctatc tcatccatgc cagtcctcac aaagttcatc	900
gatcagactt ccctgatgcc tgtttggctc ctgtgcgagg acaaaaccaa ggctgatgcc	960
ctcaagaaat actacgacaa cacttggctg aagtctcaga tggcggctgc ttccctgcgc	1020
gcgaactaca tcgatgagct gaccttcgtg ctgggcgaca atagcgatat tccggctccc	1080
gttggctaca cgaaggtgcc gatcgacctc aactcggatg ccggcggcaa gtatgtgtac	1140
ctgtgctacc atgaggcgca gttcacgcca gtcaatggca agcagcctat tgtcgacatc	1200
caggttctct acggctccca gatgccagct cctggataca taaagattga cgtggatctc	1260
aactcgggcg ccggcgggga gttcgtttac ctgagctaca agaaggcgga gcccacaagc	1320
tcggacgtca taaacaagat cactgcggtt tacggcaaga acgagtatgt gccgacaccc	1380
tacggctaca agcagatcag cggggatctg aatgcggggg cgggcggcga tttcgtttac	1440
ctctgcacct accagggcgg gaccgagtga	1470