



(12)发明专利申请

(10)申请公布号 CN 109844105 A

(43)申请公布日 2019.06.04

(21)申请号 201780048121.5

(74)专利代理机构 北京集佳知识产权代理有限

(22)申请日 2017.07.10

公司 11227

(30)优先权数据

代理人 郑斌 刘振佳

62/360,745 2016.07.11 US

(51)Int.Cl.

C12N 9/00(2006.01)

(85)PCT国际申请进入国家阶段日

C12N 9/86(2006.01)

2019.01.31

C12N 15/11(2006.01)

(86)PCT国际申请的申请数据

C12N 15/70(2006.01)

PCT/US2017/041373 2017.07.10

C12P 21/02(2006.01)

(87)PCT国际申请的公布数据

W02018/013481 EN 2018.01.18

(71)申请人 得克萨斯州大学系统董事会

地址 美国得克萨斯州

(72)发明人 安德鲁·D·埃林顿

罗西·塞耶尔

权利要求书8页 说明书20页

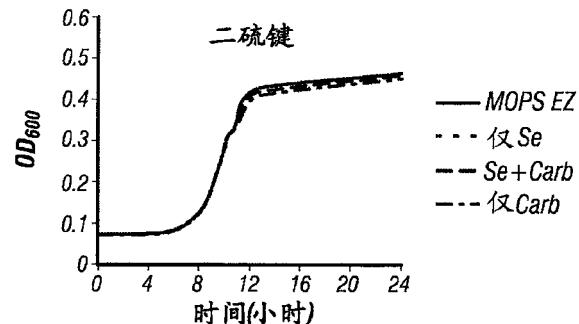
序列表52页 附图9页

(54)发明名称

包含硒代半胱氨酸的重组多肽及其产生方法

(57)摘要

提供了包含纯化的重组硒蛋白，如抗体和酶的组合物。同样提供了产生此类重组多肽的方法和用于产生此类重组多肽的细菌菌株。



1. 一种包含纯化的重组多肽的组合物，所述多肽在所述多肽的野生型型式中未发现的选定位置包含至少一个硒代半胱氨酸残基，其中所述组合物中至少80%的所述重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

2. 如权利要求1所述的组合物，其中所述组合物中80%-99.9%的所述重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

3. 如权利要求1所述的组合物，其中所述组合物中90%-99%的所述重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

4. 如权利要求1所述的组合物，其中所述组合物中至少95%的所述重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

5. 如权利要求1所述的组合物，其中所述组合物中至少99%的所述重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

6. 如权利要求1-5中任一项所述的组合物，其中所述重组多肽与人多肽至少90%相同。

7. 如权利要求6所述的组合物，其中所述人多肽是疾病中涉及的多肽。

8. 如权利要求1-5中任一项所述的组合物，其中所述重组多肽包含抗体或酶。

9. 如权利要求1-5中任一项所述的组合物，其中所述多肽在选定位置包含至少两个硒代半胱氨酸残基。

10. 如权利要求9所述的组合物，其中所述选定位置处的所述两个硒代半胱氨酸残基形成二硒键。

11. 如权利要求10所述的组合物，其中所述多肽包含酶或抗体。

12. 如权利要求11所述的组合物，其中所述抗体是无糖基化抗体。

13. 如权利要求1-5中任一项所述的组合物，其包含至少50μg所述纯化的重组多肽。

14. 如权利要求1-5中任一项所述的组合物，其中纯化的重组多肽是97%至99.9%纯的。

15. 如权利要求1-11中任一项所述的组合物，其中所述多肽在选定位置包含2、3、4、5、6、7、8、9或10个硒代半胱氨酸残基。

16. 一种药物组合物，其包含根据权利要求1-11中任一项所述的组合物。

17. 一种治疗受试者的方法，所述方法包括向所述受试者施用有效量的根据权利要求16所述的药物组合物。

18. 一种编码多肽的核酸分子，所述多肽：

(a) 包含与SEQ ID NO:1至少90%相同的氨基酸序列并且在对应于SEQ ID NO:1的位置344的位置处具有氨基酸取代或缺失；

(b) 包含与SEQ ID NO:2至少90%相同的氨基酸序列并且在对应于SEQ ID NO:2的位置702的位置处具有氨基酸取代或缺失；

(c) 包含与SEQ ID NO:3至少90%相同的氨基酸序列并且在对应于SEQ ID NO:3的位置655的位置处具有氨基酸取代或缺失；

(d) 包含与SEQ ID NO:4至少90%相同的氨基酸序列并且在对应于SEQ ID NO:4的位置73的位置处具有氨基酸取代或缺失；

(e) 包含与SEQ ID NO:5至少90%相同的氨基酸序列并且在对应于SEQ ID NO:5的位置781的位置处具有氨基酸取代或缺失；

- (f) 包含与SEQ ID NO:6至少90%相同的氨基酸序列并且在对应于SEQ ID NO:6的位置136的位置处具有氨基酸取代或缺失；
- (g) 包含与SEQ ID NO:7至少90%相同的氨基酸序列并且在对应于SEQ ID NO:7的位置183的位置处具有氨基酸取代或缺失；
- (h) 包含与SEQ ID NO:8至少90%相同的氨基酸序列并且在对应于SEQ ID NO:8的位置1的位置处具有氨基酸取代或缺失；
- (i) 包含与SEQ ID NO:9至少90%相同的氨基酸序列并且在对应于SEQ ID NO:9的位置102的位置处具有氨基酸取代或缺失；
- (j) 包含与SEQ ID NO:10至少90%相同的氨基酸序列并且在对应于SEQ ID NO:10的位置105的位置处具有氨基酸取代或缺失；
- (k) 包含与SEQ ID NO:11至少90%相同的氨基酸序列并且在对应于SEQ ID NO:11的位置673的位置处具有氨基酸取代或缺失；
- (l) 包含与SEQ ID NO:12至少90%相同的氨基酸序列并且在对应于SEQ ID NO:12的位置69的位置处具有氨基酸取代或缺失；
- (m) 包含与SEQ ID NO:13至少90%相同的氨基酸序列并且在对应于SEQ ID NO:13的位置107的位置处具有氨基酸取代或缺失；
- (n) 包含与SEQ ID NO:17至少90%相同的氨基酸序列并且在对应于SEQ ID NO:17的位置246的位置处具有氨基酸取代或缺失；
- (o) 包含与SEQ ID NO:18至少90%相同的氨基酸序列并且在对应于SEQ ID NO:18的位置545的位置处具有氨基酸取代或缺失；和/或
- (p) 包含与SEQ ID NO:19至少90%相同的氨基酸序列并且在对应于SEQ ID NO:19的位置124的位置处具有氨基酸取代或缺失。

19. 一种重组多肽，所述重组多肽：

- (a) 包含与SEQ ID NO:1至少90%相同的氨基酸序列并且在对应于SEQ ID NO:1的位置344的位置处具有氨基酸取代或缺失；
- (b) 包含与SEQ ID NO:2至少90%相同的氨基酸序列并且在对应于SEQ ID NO:2的位置702的位置处具有氨基酸取代或缺失；
- (c) 包含与SEQ ID NO:3至少90%相同的氨基酸序列并且在对应于SEQ ID NO:3的位置655的位置处具有氨基酸取代或缺失；
- (d) 包含与SEQ ID NO:4至少90%相同的氨基酸序列并且在对应于SEQ ID NO:4的位置73的位置处具有氨基酸取代或缺失；
- (e) 包含与SEQ ID NO:5至少90%相同的氨基酸序列并且在对应于SEQ ID NO:5的位置781的位置处具有氨基酸取代或缺失；
- (f) 包含与SEQ ID NO:6至少90%相同的氨基酸序列并且在对应于SEQ ID NO:6的位置136的位置处具有氨基酸取代或缺失；
- (g) 包含与SEQ ID NO:7至少90%相同的氨基酸序列并且在对应于SEQ ID NO:7的位置183的位置处具有氨基酸取代或缺失；
- (h) 包含与SEQ ID NO:8至少90%相同的氨基酸序列并且在对应于SEQ ID NO:8的位置1的位置处具有氨基酸取代或缺失；

(i) 包含与SEQ ID NO:9至少90%相同的氨基酸序列并且在对应于SEQ ID NO:9的位置102的位置处具有氨基酸取代或缺失；

(j) 包含与SEQ ID NO:10至少90%相同的氨基酸序列并且在对应于SEQ ID NO:10的位置105的位置处具有氨基酸取代或缺失；

(k) 包含与SEQ ID NO:11至少90%相同的氨基酸序列并且在对应于SEQ ID NO:11的位置673的位置处具有氨基酸取代或缺失；

(l) 包含与SEQ ID NO:12至少90%相同的氨基酸序列并且在对应于SEQ ID NO:12的位置69的位置处具有氨基酸取代或缺失；

(m) 包含与SEQ ID NO:13至少90%相同的氨基酸序列并且在对应于SEQ ID NO:13的位置107的位置处具有氨基酸取代或缺失；

(n) 包含与SEQ ID NO:17至少90%相同的氨基酸序列并且在对应于SEQ ID NO:17的位置246的位置处具有氨基酸取代或缺失；

(o) 包含与SEQ ID NO:18至少90%相同的氨基酸序列并且在对应于SEQ ID NO:18的位置545的位置处具有氨基酸取代或缺失；或

(p) 包含与SEQ ID NO:19至少90%相同的氨基酸序列并且在对应于SEQ ID NO:19的位置124的位置处具有氨基酸取代或缺失。

20. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:1的位置344的位置处的Pro取代。

21. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:2的位置702的位置处的His取代。

22. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:3的位置655的位置处的Ala取代。

23. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:4的位置73的位置处的Ala取代。

24. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:5的位置781的位置处的Gly取代。

25. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:6的位置136的位置处的Val取代。

26. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:7的位置183的位置处的Ala取代。

27. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:8的位置1的位置处的Arg取代。

28. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:9的位置102的位置处的Cys取代。

29. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:10的位置105的位置处的Cys取代。

30. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:11的位置673的位置处的Leu取代。

31. 如权利要求19所述的重组多肽，其中所述多肽包含在对应于SEQ ID NO:12的位置

69的位置处的Gly取代。

32. 如权利要求19所述的重组多肽,其中所述多肽包含在对应于SEQ ID NO:13的位置107的位置处的Ser取代。

33. 如权利要求19所述的重组多肽,其中所述多肽包含在对应于SEQ ID NO:17的位置246的位置处的Ala取代。

34. 如权利要求19所述的重组多肽,其中所述多肽包含在对应于SEQ ID NO:18的位置545的位置处的Ile取代。

35. 如权利要求19所述的重组多肽,其中所述多肽包含在对应于SEQ ID NO:19的位置124的位置处的Pro取代。

36. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID NO:1的位置999的位置处的Thr取代、在对应于SEQ ID NO:1的位置457的位置处的Thr取代、在对应于SEQ ID NO:1的位置591的位置处的Pro取代、在对应于SEQ ID NO:1的位置183的位置处的Thr取代、在对应于SEQ ID NO:1的位置358的位置处的Leu取代、在对应于SEQ ID NO:1的位置23的位置处的Arg取代、在对应于SEQ ID NO:1的位置902的位置处的Ile取代、在对应于SEQ ID NO:1的位置889的位置处的Val取代、在对应于SEQ ID NO:1的位置620的位置处的Cys取代和/或在对应于SEQ ID NO:1的位置174的位置处的Gly取代。

37. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID NO:3的位置398的位置处的Cys取代、在对应于SEQ ID NO:3的位置652的位置处的Ala取代、在对应于SEQ ID NO:3的位置264的位置处的Cys取代和/或在对应于SEQ ID NO:3的位置21的位置处的Ala取代。

38. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID NO:4的位置45的位置处的Asn取代、在对应于SEQ ID NO:4的位置290的位置处的Leu取代、在对应于SEQ ID NO:4的位置271的位置处的Asp取代、在对应于SEQ ID NO:4的位置153的位置处的Tyr取代、在对应于SEQ ID NO:4的位置45的位置处的Val取代、在对应于SEQ ID NO:4的位置284的位置处的Pro取代、在对应于SEQ ID NO:4的位置73的位置处的Ile取代、在对应于SEQ ID NO:4的位置68的位置处的Leu取代、在对应于SEQ ID NO:4的位置69的位置处的Ile取代、在对应于SEQ ID NO:4的位置305的位置处的Cys取代、在对应于SEQ ID NO:4的位置144的位置处的Ser取代和/或在对应于SEQ ID NO:4的位置281的位置处的Val取代。

39. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID NO:5的位置916的位置处的Gly取代、在对应于SEQ ID NO:5的位置938的位置处的His取代、在对应于SEQ ID NO:5的位置860的位置处的His取代、在对应于SEQ ID NO:5的位置925的位置处的Asp取代和/或在对应于SEQ ID NO:5的位置470的位置处的Met取代。

40. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID NO:6的位置115的位置处的Ala取代、在对应于SEQ ID NO:6的位置386的位置处的Arg取代、在对应于SEQ ID NO:6的位置155的位置处的Arg取代、在对应于SEQ ID NO:6的位置98的位置处的Ser取代、在对应于SEQ ID NO:6的位置201的位置处的Ala取代、在对应于SEQ ID NO:6的位置294的位置处的Thr取代、在对应于SEQ ID NO:6的位置159的位置处的Tyr取代和/或在对应于SEQ ID NO:6的位置112的位置处的Ile取代。

41. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID NO:18的位

置76的位置处的Gly取代、在对应于SEQ ID N0:18的位置293的位置处的Val取代、在对应于SEQ ID N0:18的位置637的位置处的Thr取代、在对应于SEQ ID N0:18的位置3的位置处的Val取代、在对应于SEQ ID N0:18的位置311的位置处的Ser取代、在对应于SEQ ID N0:18的位置471的位置处的Thr取代、在对应于SEQ ID N0:18的位置228的位置处的Val取代、在对应于SEQ ID N0:18的位置311的位置处的Ser取代和/或在对应于SEQ ID N0:18的位置257的位置处的Thr取代。

42. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID N0:7的位置173的位置处的Val取代、在对应于SEQ ID N0:7的位置196的位置处的Leu取代、在对应于SEQ ID N0:7的位置180的位置处的Phe取代、在对应于SEQ ID N0:7的位置249的位置处的Ala取代、在对应于SEQ ID N0:7的位置5的位置处的Val取代、在对应于SEQ ID N0:7的位置273的位置处的Leu取代和/或在对应于SEQ ID N0:7的位置176的位置处的Asn取代。

43. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID N0:9的位置15的位置处的Cys取代和/或在对应于SEQ ID N0:9的位置30的位置处的取代。

44. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID N0:19的位置193的位置处的Ile取代、在对应于SEQ ID N0:19的位置233的位置处的Thr取代、在对应于SEQ ID N0:19的位置300的位置处的Ala取代和/或在对应于SEQ ID N0:19的位置199的位置处的Arg取代。

45. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID N0:17的位置246的位置处的Ala取代。

46. 如权利要求19所述的重组多肽,其中所述多肽还包含在对应于SEQ ID N0:11的位置119的位置处的Val取代、在对应于SEQ ID N0:11的位置535的位置处的Pro取代、在对应于SEQ ID N0:11的位置373的位置处的Arg取代、在对应于SEQ ID N0:11的位置535的位置处的Ser取代、在对应于SEQ ID N0:11的位置119的位置处的Thr取代、在对应于SEQ ID N0:11的位置601的位置处的Ala取代、在对应于SEQ ID N0:11的位置103的位置处的Lys取代、在对应于SEQ ID N0:11的位置31的位置处的Asp取代、在对应于SEQ ID N0:11的位置662的位置处的Ile取代、在对应于SEQ ID N0:11的位置359的位置处的Lys取代和/或在对应于SEQ ID N0:11的位置519的位置处的Asp取代。

47. 一种编码多肽的核酸分子,所述多肽包含:(i)与SEQ ID N0:14至少90%相同的氨基酸序列,以及(ii)在对应于位置212的位置处的Ile取代、在对应于位置162的位置处的Asn取代、在对应于位置299的位置处的Ala取代和/或在对应于位置220的位置处的Arg取代。

48. 一种重组多肽,其包含:(i)与SEQ ID N0:14至少90%相同的氨基酸序列,以及(ii)在对应于位置212的位置处的Ile取代、在对应于位置162的位置处的Asn取代、在对应于位置299的位置处的Ala取代和/或在对应于位置220的位置处的Arg取代。

49. 一种编码多肽的核酸分子,所述多肽包含:(i)与SEQ ID N0:15至少90%相同的氨基酸序列,以及(ii)在对应于位置184的位置处的Ala取代、在对应于位置1730的位置处的Asn取代、在对应于位置1888的位置处的Asp取代、在对应于位置352的位置处的Thr取代、在对应于位置374的位置处的Ser取代、在对应于位置1423的位置处的Arg取代、在对应于位置1502的位置处的Glu取代、在对应于位置285的位置处的Gln取代、在对应于位置470的位置

处的Lys取代、在对应于位置939的位置处的Thr取代、在对应于位置1669的位置处的Phe取代、在对应于位置2034的位置处的Ile取代、在对应于位置1713的位置处的Asn取代、在对应于位置704的位置处的Thr取代和/或在对应于位置1084的位置处的Ile取代。

50. 一种重组多肽,其包含:(i)与SEQ ID NO:15至少90%相同的氨基酸序列,以及(ii)在对应于位置184的位置处的Ala取代、在对应于位置1730的位置处的Asn取代、在对应于位置1888的位置处的Asp取代、在对应于位置352的位置处的Thr取代、在对应于位置374的位置处的Ser取代、在对应于位置1423的位置处的Arg取代、在对应于位置1502的位置处的Glu取代、在对应于位置285的位置处的Gln取代、在对应于位置470的位置处的Lys取代、在对应于位置939的位置处的Thr取代、在对应于位置1669的位置处的Phe取代、在对应于位置2034的位置处的Ile取代、在对应于位置1713的位置处的Asn取代、在对应于位置704的位置处的Thr取代和/或在对应于位置1084的位置处的Ile取代。

51. 一种编码多肽的核酸分子,所述多肽包含:(i)与SEQ ID NO:16至少90%相同的氨基酸序列,以及(ii)在对应于位置112的位置处的Phe取代、在对应于位置126的位置处的Ala取代、在对应于位置978的位置处的Leu取代、在对应于位置199的位置处的Gly取代、在对应于位置476的位置处的Thr取代和/或在对应于位置735的位置处的Val取代。

52. 一种重组多肽,其包含:(i)与SEQ ID NO:16至少90%相同的氨基酸序列,以及(ii)在对应于位置112的位置处的Phe取代、在对应于位置126的位置处的Ala取代、在对应于位置978的位置处的Leu取代、在对应于位置199的位置处的Gly取代、在对应于位置476的位置处的Thr取代和/或在对应于位置735的位置处的Val取代。

53. 一种细菌菌株,其包含选自如权利要求18、47、49或51所述的核酸分子的至少一种核酸分子。

54. 如权利要求53所述的细菌菌株,其中所述细菌菌株表达由选自如权利要求18、47、49或51所述的核酸分子的至少一种核酸分子编码的多肽。

55. 如权利要求53所述的细菌菌株,其包含至少2、3、4、5、6、7、8、9或10种所述核酸分子。

56. 如权利要求55所述的细菌菌株,其中所述细菌菌株表达由至少2、3、4、5、6、7、8、9或10种所述核酸分子编码的多肽。

57. 如权利要求55所述的细菌菌株,其中所述细菌菌株是大肠杆菌菌株。

58. 如权利要求53所述的细菌菌株,其包含编码tRNA的核酸,所述tRNA与SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22至少90%相同,并且包括一个或多个以下特征:

- (i) 在对应于位置7的位置处的G或C;
- (ii) 在对应于位置49的位置处的T;
- (iii) 在对应于位置50的位置处的A或C;
- (iv) 在对应于位置64的位置处的T;
- (v) 在对应于位置65的位置处的G或A;和/或
- (vi) 在对应于位置66的位置处的G、T或C。

59. 如权利要求58所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含与SEQ ID NO:20至少约90%相同的序列;并且包括一个或多个以下特征:

- (i) 在对应于位置7的位置处的G;

- (ii) 在对应于位置49的位置处的T;
- (iii) 在对应于位置50的位置处的C;
- (iv) 在对应于位置64的位置处的T;
- (v) 在对应于位置65的位置处的G; 和/或
- (vi) 在对应于位置66的位置处的C。

60. 如权利要求59所述的细菌菌株,其中所述分子包含所述特征(i)-(vi)中的2、3、4、5或6个。

61. 如权利要求58所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含与SEQ ID NO:20至少95%相同的序列。

62. 如权利要求61所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含SEQ ID NO:20的序列。

63. 如权利要求58所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含与SEQ ID NO:21至少约90%相同的序列。

64. 如权利要求63所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含SEQ ID NO:21的序列。

65. 如权利要求58所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含与SEQ ID NO:22至少约90%相同的序列。

66. 如权利要求65所述的细菌菌株,其中所述分子编码tRNA,所述tRNA包含SEQ ID NO:22的序列。

67. 如权利要求53所述的细菌菌株,其还包含编码目标多肽的可表达核酸序列,所述目标多肽在编码序列中具有至少一个位置,所述至少一个位置具有用于硒代半胱氨酸并入的TAG密码子。

68. 如权利要求67所述的细菌菌株,其中所述可表达核酸序列编码人多肽。

69. 如权利要求67所述的细菌菌株,其中所述可表达核酸序列编码酶或抗体。

70. 如权利要求67所述的细菌菌株,其中所述可表达核酸序列包含T7RNA聚合酶启动子。

71. 如权利要求53所述的细菌菌株,其还包含编码T7RNA聚合酶的核酸序列。

72. 一种大肠杆菌细菌菌株,其在NCIMB以保藏号42595保藏。

73. 一种细菌培养物,其包含根据权利要求53-72中任一项所述的细菌菌株。

74. 一种细菌培养物,其包含所述培养物,所述培养物包含1至100mg/L所述培养物的量的表达的重组多肽,所述表达的重组多肽包含至少一个硒代半胱氨酸残基。

75. 如权利要求74所述的培养物,其包含10至40mg/L所述培养物的量的表达的重组多肽,所述表达的重组多肽包含至少一个硒代半胱氨酸残基。

76. 如权利要求74或75所述的培养物,其中所述表达的重组多肽是根据权利要求中任一项所述的多肽。

77. 如权利要求74或75所述的培养物,其中所述表达的重组多肽在所述多肽的野生型型式中未发现的选定位置包含至少一个硒代半胱氨酸残基。

78. 如权利要求77所述的培养物,其中所述培养物中至少80%的所述表达的重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

79. 如权利要求77所述的培养物,其中所述培养物中80%-99.9%的所述重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

80. 如权利要求77所述的培养物,其中所述培养物中90%-99%的所述表达的重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

81. 如权利要求77所述的培养物,其中所述培养物中至少95%的所述表达的重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

82. 如权利要求77所述的培养物,其中所述培养物中至少99%的所述表达的重组多肽在所述选定位置包含所述硒代半胱氨酸残基。

83. 如权利要求74或75所述的培养物,其中所述表达的重组多肽与人多肽至少90%相同。

84. 如权利要求83所述的培养物,其中所述人多肽是疾病中涉及的多肽。

85. 如权利要求74或75所述的培养物,其中所述表达的重组多肽包含抗体或酶。

86. 如权利要求74或75所述的培养物,其中所述多肽在选定位置包含至少两个硒代半胱氨酸残基。

87. 如权利要求86所述的培养物,其中所述选定位置处的所述两个硒代半胱氨酸残基形成二硒键。

88. 如权利要求86所述的培养物,其中所述多肽在选定位置包含2、3、4、5、6、7、8、9或10个硒代半胱氨酸残基。

89. 一种多肽,其从根据权利要求74-88中任一项所述的培养物中纯化,所述多肽包含至少第一硒代半胱氨酸残基。

90. 一种表达多肽的方法,所述多肽包含至少一个硒代半胱氨酸残基,所述方法包括:

(a) 在根据权利要求53-72中任一项所述的细菌菌株中且在硒源存在下表达编码所述多肽的核酸;以及

(b) 从所述细菌中纯化所述重组多肽。

91. 一种通过以下方法产生的重组多肽,所述重组多肽在选定位置包含至少一个硒代半胱氨酸残基,所述方法包括:

(a) 在根据权利要求53-72中任一项所述的细菌菌株中且在硒源存在下表达编码所述多肽的核酸;以及

(b) 从所述细菌中纯化所述重组多肽。

包含硒代半胱氨酸的重组多肽及其产生方法

[0001] 本申请要求2016年7月11日提交的美国临时专利申请号62/360,745的权益，所述临时专利申请的全部内容以引用的方式并入本文。

[0002] 本发明是在政府的支持下根据国家科学基金会授予的资助号CHE1402753进行的。政府拥有本发明的某些权利。

[0003] 发明背景

1. 技术领域

[0004] 本发明总体上涉及分子生物学领域。更具体地，本发明涉及包含非规范氨基酸的多肽。

2. 相关技术说明

[0006] 某些生物体使用硒来表达硒蛋白。硒蛋白是一组独特的多肽，其存在于原核生物和真核生物中并且含有非规范氨基酸硒代半胱氨酸。硒代半胱氨酸具有比半胱氨酸显著更低的pK_a（对于游离氨基酸5.2对8.5）和更强的亲核性质，从而使其成为改变蛋白质化学和功能的有吸引力的靶标。不幸的是，大多数硒蛋白已被证明难以在大肠杆菌中产生，大肠杆菌是用于重组蛋白产生的标准宿主。这是由于细菌中硒代半胱氨酸并入的固有低效率（4%-5%对蛋白质合成的终止）。此外，要求SECIS元件紧跟UGA密码子，从而形成编码序列的一部分，极大地限制了哪些蛋白质适合于硒代半胱氨酸插入。

[0007] 最近，还开发了一种进化的大肠杆菌tRNA^{Sec}，其与经典翻译机制相容并且可遏制琥珀终止密码子以高效率地并入硒代半胱氨酸。然而，对于允许有效并入硒代半胱氨酸以产生商业上相关量的重组硒蛋白的系统存在未满足的需求。

发明内容

[0008] 本公开的第一实施方案提供一种包含纯化的重组多肽的组合物，所述多肽在所述多肽的野生型型式中未发现的选定位置包含至少一个硒代半胱氨酸残基，其中所述组合物中至少80%的重组多肽在所述选定位置包含硒代半胱氨酸残基。在某些方面，所述重组多肽包含抗体或酶。在某些方面，所述组合物包含至少10μg、50μg、100μg、500μg或1mg的纯化的重组多肽。在一些方面，所述纯化的重组多肽是95%-99.9%纯的，例如至少约95%、96%、97%、98%、99%或99.5%纯的。

[0009] 在一些方面，所述组合物中80%-99.9%的重组多肽在所述选定位置包含硒代半胱氨酸残基。在某些方面，所述组合物中90%-99%的重组多肽在所述选定位置包含硒代半胱氨酸残基。在一些方面，所述组合物中至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的重组多肽在所述选定位置包含硒代半胱氨酸残基。在其它方面，组合物包含在选定位置具有至少两个硒代半胱氨酸残基的重组多肽，并且所述组合物中80%-99.9%（例如，至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%）的重组多肽在所述选定位置包含两个硒代半胱氨酸残基。在其它方面，组合物包含在形成二硒键的选定位置具有至少两个硒代半胱氨酸残基的重组多肽，并且所述组合物中80%-99.9%（例

如,至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%)的重组多肽包含在所述选定位置的硒代半胱氨酸残基之间的二硒键。

[0010] 在一些方面,所述重组多肽与人多肽至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同。在某些方面,所述人多肽是疾病中涉及的多肽。在一些方面,所述人多肽是酶、趋化因子、细胞因子、抗体或T细胞受体。在一些方面,所述抗体是无糖基化抗体。在其它方面,所述人多肽是包含二硫键的多肽,并且所述重组多肽包含代替所述二硫键的二硒键。

[0011] 在某些方面,所述多肽在选定位置包含2、3、4、5、6、7、8、9或10个硒代半胱氨酸残基。在其它方面,所述组合物中至少约80%-99.9% (例如,至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%) 的重组多肽在所述选定位置中的每一个处包含硒代半胱氨酸残基。在其它方面,所述多肽在选定位置包含至少两个硒代半胱氨酸残基。在某些方面,所述选定位置处的两个硒代半胱氨酸残基形成二硒键。

[0012] 另一实施方案提供一种药物组合物,所述药物组合物包含实施方案的组合物,所述组合物包含纯化的重组多肽,所述多肽在所述多肽的野生型型式中未发现的选定位置包含至少一个硒代半胱氨酸残基,其中所述组合物中至少80%-99.9% (例如,至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%) 的重组多肽在所述选定位置包含硒代半胱氨酸残基。

[0013] 其它实施方案提供一种治疗受试者的方法,所述方法包括施用有效量的实施方案的药物组合物,所述药物组合物包含纯化的重组多肽,所述多肽在所述多肽的野生型型式中未发现的选定位置包含至少一个硒代半胱氨酸残基,其中所述组合物中至少80%-99.9% (例如,至少90%、91%、92%、93%、94%、95%、96%、97%、98%或99%) 的重组多肽在所述选定位置包含硒代半胱氨酸残基。

[0014] 在另一实施方案中,提供了一种编码多肽的核酸分子,所述多肽: (a) 包含与SEQ ID NO:1至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:1的位置344的位置处具有氨基酸取代或缺失; (b) 包含与SEQ ID NO:2至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:2的位置702的位置处具有氨基酸取代或缺失; (c) 包含与SEQ ID NO:3至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:3的位置655的位置处具有氨基酸取代或缺失; (d) 包含与SEQ ID NO:4至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:4的位置73的位置处具有氨基酸取代或缺失; (e) 包含与SEQ ID NO:5至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:5的位置781的位置处具有氨基酸取代或缺失; (f) 包含与SEQ ID NO:6至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:6的位置136的位置处具有氨基酸取代或缺失; (g) 包含与SEQ ID NO:7至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:7的位置183的位置处具有氨基酸取代或缺失; (h) 包含与SEQ ID NO:8至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:8的位置1的位置处具

有氨基酸取代或缺失；(i) 包含与SEQ ID NO:9至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:9的位置102的位置处具有氨基酸取代或缺失；(j) 包含与SEQ ID NO:10至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:10的位置105的位置处具有氨基酸取代或缺失；(k) 包含与SEQ ID NO:11至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:11的位置673的位置处具有氨基酸取代或缺失；(l) 包含与SEQ ID NO:12至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:12的位置69的位置处具有氨基酸取代或缺失；(m) 包含与SEQ ID NO:13至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:13的位置107的位置处具有氨基酸取代或缺失；(n) 包含与SEQ ID NO:17至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:17的位置246的位置处具有氨基酸取代或缺失；(o) 包含与SEQ ID NO:18至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:18的位置545的位置处具有氨基酸取代或缺失；和/或(p) 包含与SEQ ID NO:19至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:19的位置124的位置处具有氨基酸取代或缺失。

[0015] 另一实施方案提供一种重组多肽，所述重组多肽：(a) 包含与SEQ ID NO:1至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:1的位置344的位置处具有氨基酸取代或缺失；(b) 包含与SEQ ID NO:2至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:2的位置702的位置处具有氨基酸取代或缺失；(c) 包含与SEQ ID NO:3至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:3的位置655的位置处具有氨基酸取代或缺失；(d) 包含与SEQ ID NO:4至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:4的位置73的位置处具有氨基酸取代或缺失；(e) 包含与SEQ ID NO:5至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:5的位置781的位置处具有氨基酸取代或缺失；(f) 包含与SEQ ID NO:6至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:6的位置136的位置处具有氨基酸取代或缺失；(g) 包含与SEQ ID NO:7至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:7的位置183的位置处具有氨基酸取代或缺失；(h) 包含与SEQ ID NO:8至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:8的位置1的位置处具有氨基酸取代或缺失；(i) 包含与SEQ ID NO:9至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:9的位置102的位置处具有氨基酸取代或缺失；(j) 包含与SEQ ID NO:10至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:10的位置105的位置处具有氨基酸取代或缺失；(k) 包含与SEQ ID NO:11至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、

98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:11的位置673的位置处具有氨基酸取代或缺失; (l) 包含与SEQ ID NO:12至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:12的位置69的位置处具有氨基酸取代或缺失; (m) 包含与SEQ ID NO:13至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:13的位置107的位置处具有氨基酸取代或缺失; (n) 包含与SEQ ID NO:17至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:17的位置246的位置处具有氨基酸取代或缺失; (o) 包含与SEQ ID NO:18至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:18的位置545的位置处具有氨基酸取代或缺失; 或 (p) 包含与SEQ ID NO:19至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列并且在对应于SEQ ID NO:19的位置124的位置处具有氨基酸取代或缺失。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0016] 在一些方面,根据前一段落的多肽包含在对应于SEQ ID NO:1的位置344的位置处的Pro取代。在某些方面,所述多肽包含在对应于SEQ ID NO:2的位置702的位置处的His取代。在一些方面,所述多肽包含在对应于SEQ ID NO:3的位置655的位置处的Ala取代。在一些方面,所述多肽包含在对应于SEQ ID NO:4的位置73的位置处的Ala取代。在某些方面,所述多肽包含在对应于SEQ ID NO:5的位置781的位置处的Gly取代。在一些方面,所述多肽包含在对应于SEQ ID NO:6的位置136的位置处的Val取代。在一些方面,所述多肽包含在对应于SEQ ID NO:7的位置183的位置处的Ala取代。在某些方面,所述多肽包含在对应于SEQ ID NO:8的位置1的位置处的Arg取代。在一些方面,所述多肽包含在对应于SEQ ID NO:9的位置102的位置处的Cys取代。在一些方面,所述多肽包含在对应于SEQ ID NO:10的位置105的位置处的Cys取代。在某些方面,所述多肽包含在对应于SEQ ID NO:11的位置673的位置处的Leu取代。在一些方面,所述多肽包含在对应于SEQ ID NO:12的位置69的位置处的Gly取代。在某些方面,所述多肽包含在对应于SEQ ID NO:13的位置107的位置处的Ser取代。在一些方面,所述多肽包含在对应于SEQ ID NO:17的位置246的位置处的Ala取代。在某些方面,所述多肽包含在对应于SEQ ID NO:18的位置545的位置处的Ile取代。在一些方面,所述多肽包含在对应于SEQ ID NO:19的位置124的位置处的Pro取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0017] 在一些方面,所述多肽还包含与SEQ ID NO:1至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:1的位置999的位置处的Thr取代、在对应于SEQ ID NO:1的位置457的位置处的Thr取代、在对应于SEQ ID NO:1的位置591的位置处的Pro取代、在对应于SEQ ID NO:1的位置183的位置处的Thr取代、在对应于SEQ ID NO:1的位置358的位置处的Leu取代、在对应于SEQ ID NO:1的位置23的位置处的Arg取代、在对应于SEQ ID NO:1的位置902的位置处的Ile取代、在对应于SEQ ID NO:1的位置889的位置处的Val取代、在对应于SEQ ID NO:1的位置620的位置处的Cys取代和/或在对应于SEQ ID NO:1的位置174的位置处的Gly取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0018] 在某些方面,所述多肽还包含与SEQ ID NO:3至少85%、90%、91%、92%、93%、

94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:3的位置398的位置处的Cys取代、在对应于SEQ ID NO:3的位置652的位置处的Ala取代、在对应于SEQ ID NO:3的位置264的位置处的Cys取代和/或在对应于SEQ ID NO:3的位置21的位置处的Ala取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0019] 在一些方面,所述多肽还包含与SEQ ID NO:4至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:4的位置45的位置处的Asn取代、在对应于SEQ ID NO:4的位置290的位置处的Leu取代、在对应于SEQ ID NO:4的位置271的位置处的Asp取代、在对应于SEQ ID NO:4的位置153的位置处的Tyr取代、在对应于SEQ ID NO:4的位置45的位置处的Val取代、在对应于SEQ ID NO:4的位置284的位置处的Pro取代、在对应于SEQ ID NO:4的位置73的位置处的Ile取代、在对应于SEQ ID NO:4的位置68的位置处的Leu取代、在对应于SEQ ID NO:4的位置69的位置处的Ile取代、在对应于SEQ ID NO:4的位置305的位置处的Cys取代、在对应于SEQ ID NO:4的位置144的位置处的Ser取代和/或在对应于SEQ ID NO:4的位置281的位置处的Val取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0020] 在某些方面,所述多肽还包含与SEQ ID NO:5至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:5的位置916的位置处的Gly取代、在对应于SEQ ID NO:5的位置938的位置处的His取代、在对应于SEQ ID NO:5的位置860的位置处的His取代、在对应于SEQ ID NO:5的位置925的位置处的Asp取代和/或在对应于SEQ ID NO:5的位置470的位置处的Met取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0021] 在一些方面,所述多肽还包含与SEQ ID NO:6至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:6的位置115的位置处的Ala取代、在对应于SEQ ID NO:6的位置386的位置处的Arg取代、在对应于SEQ ID NO:6的位置155的位置处的Arg取代、在对应于SEQ ID NO:6的位置98的位置处的Ser取代、在对应于SEQ ID NO:6的位置201的位置处的Ala取代、在对应于SEQ ID NO:6的位置294的位置处的Thr取代、在对应于SEQ ID NO:6的位置159的位置处的Tyr取代和/或在对应于SEQ ID NO:6的位置112的位置处的Ile取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0022] 在一些方面,所述多肽还包含与SEQ ID NO:18至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:18的位置76的位置处的Gly取代、在对应于SEQ ID NO:18的位置293的位置处的Val取代、在对应于SEQ ID NO:18的位置637的位置处的Thr取代、在对应于SEQ ID NO:18的位置3的位置处的Val取代、在对应于SEQ ID NO:18的位置311的位置处的Ser取代、在对应于SEQ ID NO:18的位置471的位置处的Thr取代、在对应于SEQ ID NO:18的位置228的位置处的Val取代、在对应于SEQ ID NO:18的位置311的位置处的Ser取代和/或在对应于SEQ ID NO:18的位置257的位置处的Thr取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0023] 在一些方面,所述多肽还包含与SEQ ID NO:7至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:7的位置173的位置处的Val取代、在对应于SEQ ID NO:7的位置196的位置处的Leu取代、在对应于

SEQ ID NO:7的位置180的位置处的Phe取代、在对应于SEQ ID NO:7的位置249的位置处的Ala取代、在对应于SEQ ID NO:7的位置5的位置处的Val取代、在对应于SEQ ID NO:7的位置273的位置处的Leu取代和/或在对应于SEQ ID NO:7的位置176的位置处的Asn取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0024] 在一些方面,所述多肽还包含与SEQ ID NO:9至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:9的位置15的位置处的Cys取代和/或在对应于SEQ ID NO:9的位置30的位置处的取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0025] 在某些方面,所述多肽还包含与SEQ ID NO:19至少90%相同的氨基酸序列以及在对应于SEQ ID NO:19的位置193的位置处的Ile取代、在对应于SEQ ID NO:19的位置233的位置处的Thr取代、在对应于SEQ ID NO:19的位置300的位置处的Ala取代和/或在对应于SEQ ID NO:19的位置199的位置处的Arg取代。

[0026] 在一些方面,所述多肽还包含与SEQ ID NO:17至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:17的位置246的位置处的Ala取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0027] 在一些方面,所述多肽还包含与SEQ ID NO:11至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列以及在对应于SEQ ID NO:11的位置119的位置处的Val取代、在对应于SEQ ID NO:11的位置535的位置处的Pro取代、在对应于SEQ ID NO:11的位置373的位置处的Arg取代、在对应于SEQ ID NO:11的位置535的位置处的Ser取代、在对应于SEQ ID NO:11的位置119的位置处的Thr取代、在对应于SEQ ID NO:11的位置601的位置处的Ala取代、在对应于SEQ ID NO:11的位置103的位置处的Lys取代、在对应于SEQ ID NO:11的位置31的位置处的Asp取代、在对应于SEQ ID NO:11的位置662的位置处的Ile取代、在对应于SEQ ID NO:11的位置359的位置处的Lys取代和/或在对应于SEQ ID NO:11的位置519的位置处的Asp取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0028] 另一个实施方案提供了一种编码多肽的核酸分子,所述多肽包含:(i)与SEQ ID NO:14至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列,以及(ii)在对应于位置212的位置处的Ile取代、在对应于位置162的位置处的Asn取代、在对应于位置299的位置处的Ala取代和/或在对应于位置220的位置处的Arg取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0029] 在又一个实施方案中,提供了一种重组多肽,所述重组多肽包含:(i)与SEQ ID NO:14至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列,以及(ii)在对应于位置212的位置处的Ile取代、在对应于位置162的位置处的Asn取代、在对应于位置299的位置处的Ala取代和/或在对应于位置220的位置处的Arg取代。在其它方面,提供了一种核酸分子,所述核酸分子编码前述多肽之一。

[0030] 在另一实施方案中,提供了一种编码多肽的核酸分子,所述多肽包含:(i)与SEQ ID NO:15至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列,以及(ii)在对应于位置184的位置处的Ala取代、在对应于位置1730的位置处的Asn

取代、在对应于位置1888的位置处的Asp取代、在对应于位置352的位置处的Thr取代、在对应于位置374的位置处的Ser取代、在对应于位置1423的位置处的Arg取代、在对应于位置1502的位置处的Glu取代、在对应于位置285的位置处的Gln取代、在对应于位置470的位置处的Lys取代、在对应于位置939的位置处的Thr取代、在对应于位置1669的位置处的Phe取代、在对应于位置2034的位置处的Ile取代、在对应于位置1713的位置处的Asn取代、在对应于位置704的位置处的Thr取代和/或在对应于位置1084的位置处的Ile取代。

[0031] 在另一实施方案中,提供了一种重组多肽,所述重组多肽包含: (i) 与SEQ ID NO: 15至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列,以及 (ii) 在对应于位置184的位置处的Ala取代、在对应于位置1730的位置处的Asn取代、在对应于位置1888的位置处的Asp取代、在对应于位置352的位置处的Thr取代、在对应于位置374的位置处的Ser取代、在对应于位置1423的位置处的Arg取代、在对应于位置1502的位置处的Glu取代、在对应于位置285的位置处的Gln取代、在对应于位置470的位置处的Lys取代、在对应于位置939的位置处的Thr取代、在对应于位置1669的位置处的Phe取代、在对应于位置2034的位置处的Ile取代、在对应于位置1713的位置处的Asn取代、在对应于位置704的位置处的Thr取代和/或在对应于位置1084的位置处的Ile取代。

[0032] 在一个实施方案中,提供了一种编码多肽的核酸分子,所述多肽包含: (i) 与SEQ ID NO:16至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列,以及 (ii) 在对应于位置112的位置处的Phe取代、在对应于位置126的位置处的Ala取代、在对应于位置978的位置处的Leu取代、在对应于位置199的位置处的Gly取代、在对应于位置476的位置处的Thr取代和/或在对应于位置735的位置处的Val取代。

[0033] 另一实施方案提供了一种重组多肽,所述重组多肽包含: (i) 与SEQ ID NO:16至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的氨基酸序列,以及 (ii) 在对应于位置112的位置处的Phe取代、在对应于位置126的位置处的Ala取代、在对应于位置978的位置处的Leu取代、在对应于位置199的位置处的Gly取代、在对应于位置476的位置处的Thr取代和/或在对应于位置735的位置处的Val取代。

[0034] 一个实施方案还提供了一种细菌菌株,所述细菌菌株包含实施方案的至少一种核酸分子或表达实施方案的至少一种多肽(例如,与SEQ ID NO:1的多肽之一至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同并且包含表1中列出的取代之一的多肽)。在一些方面,所述细菌菌株包含编码表1的多肽的至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18或19种核酸分子。在一些方面,所述细菌菌株表达由至少2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18或19种核酸分子编码的多肽。

[0035] 在其它方面,所述细菌菌株包含编码tRNA和用于并入硒代半胱氨酸的氨酰基-tRNA合成酶的核酸。在一些方面,所述tRNA识别UAG密码子。在一些方面,所述细菌菌株包含编码tRNA的核酸,所述tRNA与SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22至少90%相同。在一些方面,所述细菌菌株还包括一个或多个以下特征: (i) 在对应于位置7的位置处的G或C; (ii) 在对应于位置49的位置处的T; (iii) 在对应于位置50的位置处的A或C; (iv) 在对应于位置64的位置处的T; (v) 在对应于位置65的位置处的G或A; 和/或 (vi) 在对应于位置66的位置处的G、T或C。在一些方面,所述分子编码tRNA,所述tRNA包含与SEQ ID NO:20至少约90%相同的序列; 并且包括一个或多个以下特征: (i) 在对应于位置7的位置处的G; (ii) 在

对应于位置49的位置处的T; (iii) 在对应于位置50的位置处的C; (iv) 在对应于位置64的位置处的T; (v) 在对应于位置65的位置处的G; 和/或 (vi) 在对应于位置66的位置处的C。在其它方面,所述分子编码tRNA,所述tRNA包含与SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20或SEQ ID NO:4至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的序列。在具体方面,所述分子编码tRNA,所述tRNA包含SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20或SEQ ID NO:4。

[0036] 在一些方面,所述细菌菌株还包含编码目标多肽的可表达核酸序列,所述目标多肽在编码序列中具有至少一个位置,所述位置具有用于硒代半胱氨酸并入的TAG密码子。在某些方面,所述可表达的核酸序列编码人多肽。在一些方面,所述可表达的核酸序列编码酶或抗体。在某些方面,所述可表达的核酸序列包含T7RNA聚合酶启动子。在一些方面,所述细菌菌株还包含编码T7RNA聚合酶的核酸序列。

[0037] 在其它方面,所述细菌菌株是革兰氏阴性细菌,如大肠杆菌菌株。在另一实施方案中,提供了在NCIMB以登录号42595保藏的大肠杆菌细菌菌株。在另一实施方案中,提供了一种重组多肽,所述重组多肽包含在选定位置处的至少一个硒代半胱氨酸残基,所述重组多肽是通过在根据实施方案的细菌菌株中且在硒源存在下表达编码多肽的核酸并且从所述细菌中纯化所述重组多肽而产生。

[0038] 另一实施方案提供了一种细菌培养物,所述细菌培养物包含5至100mg/L所述培养物的量的表达的重组多肽。在某些方面,所述表达的重组多肽包含至少一个硒代半胱氨酸残基。在一些方面,所述细菌培养物包含10至40mg/L所述培养物的量的表达的重组多肽,所述表达的重组多肽包含至少一个硒代半胱氨酸残基。在某些方面,所述表达的重组多肽以5至50mg/L、10至80mg/L、15至60mg/L、10至30mg/L、20至80mg/L、30至90mg/L、40至80mg/L、50至70mg/L、60至90mg/L、70至80mg/L或90至100mg/L所述培养物的量存在。在某些方面,所述表达的重组多肽以5至10mg/L、7至15mg/L、10至20mg/L、15至30mg/L、20至35mg/L、30至40mg/L、35至45mg/L、40至50mg/L、45至55mg/L、50至50mg/L、55至65mg/L、50至60mg/L、65至70mg/L、75至85mg/L、85至90mg/L、80至95mg/L、85至98mg/L或95至100mg/L所述培养物的量存在。在某些方面,所述表达的重组多肽在所述培养物中的量是至少1mg/L、5mg/L、10mg/L、15mg/L、20mg/L、25mg/L、30mg/L、35mg/L、40mg/L、45mg/L、50mg/L、55mg/L、60mg/L、65mg/L、70mg/L、75mg/L、80mg/L、85mg/L、90mg/L、95mg/L、100mg/L或更高。

[0039] 在一些方面,所述表达的重组多肽是实施方案的多肽(例如与SEQ ID NO:1的多肽之一至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同并且包含表1中列出的取代之一的多肽)。在某些方面,所述表达的重组多肽在所述多肽的野生型型式中未发现的选定位置包含至少一个硒代半胱氨酸残基。在一些方面,所述培养物中至少80%的所述表达的重组多肽在所述选定位置包含硒代半胱氨酸残基。在某些方面,所述培养物中80%-99.9%的重组多肽在所述选定位置包含硒代半胱氨酸残基。在一些方面,所述培养物中90%-99%的所述表达的重组多肽在所述选定位置包含硒代半胱氨酸残基。在一些方面,所述培养物中至少95%的所述表达的重组多肽在所述选定位置包含硒代半胱氨酸残基。在一些方面,所述培养物中至少99%的所述表达的重组多肽在所述选定位置包含硒代半胱氨酸残基。在一些方面,所述表达的重组多肽与人多肽至少90%相同。在某些方面,所述人多肽是疾病中涉及的多肽。在一些方面,所述表达的重组多肽包含抗体或酶。在某些

方面,所述多肽在选定位置包含至少两个硒代半胱氨酸残基。在一些方面,所述选定位置处的两个硒代半胱氨酸残基形成二硒键。在具体方面,所述多肽在选定位置包含2、3、4、5、6、7、8、9或10个硒代半胱氨酸残基。另一实施方案提供了一种从实施方案的培养物中纯化的多肽,所述多肽包含至少第一硒代半胱氨酸残基。在另一个实施方案中,提供了一种表达多肽的方法,所述多肽包含至少一个硒代半胱氨酸残基,所述方法包括:(a)在实施方案的细菌菌株中且在硒源存在下表达编码所述多肽的核酸;以及(b)从所述细菌中纯化所述重组多肽。在另一个实施方案中,提供了一种通过以下方法产生的重组多肽,所述重组多肽在选定位置包含至少一个硒代半胱氨酸残基,所述方法包括:(a)在实施方案的细菌菌株中且在硒源存在下表达编码所述多肽的核酸;以及(b)从所述细菌中纯化所述重组多肽。

[0040] 在另一实施方案中,提供了实施方案的细菌菌株作为宿主用于产生包含至少一个硒代半胱氨酸残基的多肽的用途。在其它方面,所述细菌菌株在包含硒源的培养基中培养。

[0041] 在另一实施方案中,提供了一种包含异源核酸的转基因细菌菌株,所述异源核酸编码用于并入至少第一非规范氨基酸的翻译组分和可筛选或选择性标记多肽,当所述标记多肽的至少一个位置是所述第一非规范氨基酸时,所述标记多肽表现出增强的活性。在具体方面,所述可筛选标记是荧光或发光多肽。在其它方面,所述细菌菌株包含实施方案的至少一种核酸分子或表达实施方案的至少一种多肽(例如,与SEQ ID NO:1的多肽之一至少90%相同并且包含表1中列出的取代之一的多肽)。

[0042] 在其它方面,所述细菌菌株包含编码选择性标记的异源核酸,当标记多肽的至少一个位置是所述第一非规范氨基酸时,所述选择性标记表现出增强的活性。在其它方面,所述选择性标记是提供抗生素抗性的多肽。在特定方面,所述选择性标记是 β -内酰胺酶。

[0043] 在一些方面,所述细菌菌株是革兰氏阳性或革兰氏阴性细菌细胞。在一些具体方面,所述细菌细胞是大肠杆菌细胞。在其它特定方面,所述细菌细胞是肠杆菌属或沙雷氏菌属细菌。在其它方面,所述细菌细胞是阴沟肠杆菌(Enterobacter cloacae)或粘质沙雷氏菌(Serratia marcescens)细菌细胞。

[0044] 在某些方面,用于并入第一非规范氨基酸的翻译组分包含编码tRNA和用于第一非规范氨基酸的氨酰基-tRNA合成酶的核酸。在一些方面,所述tRNA识别UAG密码子。在其它方面,所述tRNA与SEQ ID NO:20、SEQ ID NO:21、SEQ ID NO:22或SEQ ID NO:23至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同。在具体方面,所述tRNA包含SEQ ID NO:20、SEQ ID NO:21、SEQ ID NO:22或SEQ ID NO:23。

[0045] 在其它方面,用于并入第一非规范氨基酸的翻译组分还包含编码用于合成第一非规范氨基酸的酶的核酸。在一个特定方面,所述非规范氨基酸是硒代半胱氨酸。在其它方面,所述细胞包含编码selA、selB和/或selC的核酸。在一个具体方面,所述细胞包含编码selA的核酸。在一些方面,所述细菌细胞包含失活或缺失的prfA基因。在某些方面,所述细胞已被工程改造为缺乏内源性琥珀(TAG)密码子。在一些特定方面,所述细胞是或源自大肠杆菌C321.ΔA。

[0046] 在另一实施方案中,本发明提供了一种根据上述实施方案和方面的细菌细胞群体。在某些方面,所述群体包含 1×10^3 至 1×10^{12} 个细菌细胞。

[0047] 在另一实施方案中,提供了一种产生商业多肽的方法,所述商业多肽包含至少第一非规范氨基酸,所述方法包括(i)获得根据实施方案的细菌菌株和编码所述商业多肽的

表达盒；以及(ii)在允许所述商业多肽表达的条件下孵育所述细菌菌株。在一些方面，编码所述商业多肽的表达盒在诱导型启动子的控制之下。在某些方面，所述方法还包括分离表达的商业多肽。

[0048] 在另一实施方案中，本发明提供了一种筛选具有所需活性的多肽的方法，所述方法包括(i)获得根据实施方案的细菌细胞群体，所述细胞编码候选多肽文库，所述多肽包含至少第一非规范氨基酸位置；以及(ii)筛选所述细菌群体以鉴别具有所需生物活性的候选多肽。在具体方面，所述细菌细胞群体包含编码100至10,000,000种不同候选多肽的核酸构建体。

[0049] 在另一实施方案中，提供了一种重组核酸分子，其中所述分子编码tRNA，所述tRNA与SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同；并且包括一个或多个以下特征：在对应于位置7的位置处的G或C；在对应于位置49的位置处的T；在对应于位置50的位置处的A或C；在对应于位置64的位置处的T；在对应于位置65的位置处的G或A；和/或在对应于位置66的位置处的G、T或C。在一些特定方面，所述分子编码tRNA，所述tRNA包含与SEQ ID NO:20至少约90%相同的序列；并且包括一个或多个以上列出的特征。在其它具体方面，所述分子包括以上列出的特征中的2、3、4、5或6个。在某些方面，所述分子编码tRNA，所述tRNA包含与SEQ ID NO:21或SEQ ID NO:19至少约90%相同的序列。在特定方面，所述分子编码tRNA，所述tRNA包含SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22的序列。

[0050] 在另一实施方案中，提供了一种重组多肽，所述重组多肽编码 β -内酰胺酶，所述酶包含为所述酶的活性所需的两个半胱氨酸位置之间的二硫键，其中所述两个半胱氨酸位置中的至少一个被硒代半胱氨酸取代。在一些方面，两个所述半胱氨酸残基都被硒代半胱氨酸取代。在其它方面，所述 β -内酰胺酶是SME型 β -内酰胺酶或NMC-A β -内酰胺酶。例如， β -内酰胺酶可包含与SEQ ID NO:24至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同的序列，并且其中对应于C69和/或C238的位置是硒代半胱氨酸。在某些方面，对应于C69和C238的位置是硒代半胱氨酸。

[0051] 本发明的其它实施方案提供了一种重组核酸分子，所述重组核酸分子编码根据上述实施方案和方面的多肽。在一些方面，对应于一个或多个硒代半胱氨酸位置的密码子是UAG密码子。在其它方面，所述序列与SEQ ID NO:25至少约85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%相同。

[0052] 如本文所用，就指定组分而言“基本上不含”在本文中用于表示指定组分未被有目的地配制到组合物中和/或仅以污染物或痕量存在。因此，由组合物的任何意外污染产生的指定组分的总量远低于0.01%。最优选的是用标准分析方法检测不到指定组分量的组合物。

[0053] 如本文在说明书中所用，“一个/种(a/an)”可指一个(种)或多个(种)。如本文在权利要求中所用，当与词语“包括/包含(comprising)”结合使用时，词语“一个/种”可指一个(种)或多于一个(种)。

[0054] 除非明确指明仅仅指代替物，或代替物相互排斥，否则权利要求书中所用的术语“或”用于指“和/或”，尽管本公开支持仅仅指代替物和“和/或”的定义。如本文所用，“另一”可指至少第二个(种)或更多个(种)。

[0055] 在整个本申请中,术语“约”用于指示值包括装置、用以测定所述值的方法的误差的固有变化,或研究受试者中存在的变化。

[0056] 通过以下详细描述,本发明的其它目的、特征和优点将变得显而易见。然而,应理解的是,尽管指示本发明的优选实施方案,但是详细描述和特定实施例仅通过说明的方式给出,因为从此详细描述中,本发明的精神和范围内的各种改变和修改对于本领域技术人员来说将变得显而易见。

附图说明

[0057] 以下附图形成本说明书的一部分,并且被包括以进一步说明本发明的某些方面。本发明可通过参考与本文提出的具体实施方案的详细描述结合的这些附图中的一个或多个来更好地理解。

[0058] 图1A-1C:用于产生硒蛋白的大肠杆菌菌株条件性地依赖于硒代半胱氨酸来生长和存活。含有整合的 β -内酰胺酶与天然二硫键的大肠杆菌菌株(图1A)不需要硒补充来在 β -内酰胺抗生素羧苄青霉素存在下存活。具有一个(图1B)或两个(图1C)必需硒代半胱氨酸残基(其分别形成氢硒基-巯基或二硒键)的大肠杆菌菌株需要硒来获得对 β -内酰胺抗生素的抗性。对硒代半胱氨酸并入的条件依赖性防止另外有毒的硒代半胱氨酸生物合成和并入机制的丧失或减弱。在图1B和1C中,在时间16小时处从顶部至底部的线表示MOPS EZ、仅Se、Se+Carb和仅Carb。

[0059] 图2A-2B:用于产生硒蛋白的大肠杆菌菌株含有一组突变,所述突变经由硒代半胱氨酸依赖性 β -内酰胺酶显著增强细胞生长和对 β -内酰胺抗生素的抗性。亲本大肠杆菌菌株(图2A)在丰富培养基中生长期间对经由硒代半胱氨酸依赖性 β -内酰胺酶介导的 β -内酰胺抗生素具有中等抗性。用于产生含有一系列突变的重组硒蛋白(图2B)的大肠杆菌菌株具有较高的生长速率、最终细胞密度和对羧苄青霉素的抗性。这种改善的生长使突变菌株成为用于产生重组硒蛋白的优良宿主。在图2A中,在时间16小时处从顶部至底部的线表示LB、100Carb、1000Carb和10000Carb,其中1000Carb和10000Carb的线重叠。在图2B中,在时间16小时处从顶部至底部的线表示LB、100Carb、1000Carb和10000Carb,其中100Carb和1000Carb的线重叠。

[0060] 图3A-3B:用于产生硒蛋白的大肠杆菌菌株含有一组突变,所述突变显著增强限定培养基中的细胞生长。亲本大肠杆菌菌株(图3A)具有延长的迟滞期并且在限定生长培养基中显示出较差生长。用于产生含有一系列突变的重组硒蛋白(图3B)的大肠杆菌菌株具有较高的生长速率、最终细胞密度和对 β -内酰胺抗生素的硒代半胱氨酸依赖性抗性。为了再现性和质量保证,限定生长培养基通常用于重组蛋白的商业生产。在图3A和3B中,在时间20小时处从顶部至底部的线表示MOPS EZ、仅Se、Se+Carb和仅Carb。

[0061] 图4A-4B:含有二硒键的大肠杆菌二氢叶酸还原酶的完整质谱和UVPD片段化图谱。大肠杆菌DHFR的质谱(图4A,所示的平均质量)证实以大约100%效率并入两个硒代半胱氨酸残基。未检测到对应于并入一个或两个丝氨酸残基的质量。UVPD片段化作图(图4B,SEQ ID NO:26)证实在位置39和85处并入硒代半胱氨酸(U)和形成二硒键。二硒键形成由对应于两个硒代半胱氨酸残基之间的片段化事件的离子的缺乏来指示。DHFR的产量是8mg/L。含有二硒键的其它蛋白质已经以超过40mg/L的产量表达。

[0062] 图5A-5C:含有两个必需二硒键的抗-MS2scFv的质谱和UVPD片段化图谱。被开发用于表达重组硒蛋白的大肠杆菌菌株能够在细菌细胞质中产生二硒化物稳定的抗体片段。与被开发用于表达含有二硫键的蛋白质的大肠杆菌菌株不同,这种菌株不需要升高的细胞质氧化还原电位。实验确定的单同位素质量(图5A)与含有四个硒代半胱氨酸残基的重组抗-MS2scFv一致,所述硒代半胱氨酸残基已经形成了两个二硒键。二硒键形成由四个质子的损失指示(图5A,第3行对第2行)。含有四个硒代半胱氨酸残基的抗-MS2scFv的完整质谱(图5B)显示为平均质量。未检测到对应于丝氨酸残基并入的质量。UVPD片段化作图(图5C,SEQ ID NO:27)证实在位置42、116、179和249并入硒代半胱氨酸(U)以及还有二硒键形成。二硒键形成由对应于硒代半胱氨酸残基对(U42:U116和U179:U249)之间的片段化事件的离子的缺乏来指示。

[0063] 图6A-6D:野生型抗蓖麻毒素A链scFv(分别图6A和6B)和硒抗蓖麻毒素A链scFv(分别图6C和6D)的UVPD片段图谱和ELISA。对于193nm UVPD序列信息,共价键合(或潜在键合)的半胱氨酸残基(图6A,SEQ ID NO:28)和共价键合(或潜在键合)的硒代半胱氨酸残基(图6C,SEQ ID NO:29)以灰色阴影表示。序列覆盖中的间隙表明在半胱氨酸残基之间形成共价二硫键。硒代半胱氨酸残基之间不存在片段化证实两个二硒键的形成。键连接性与野生型抗蓖麻毒素A链scFv相同。用DTT处理野生型抗蓖麻毒素A链scFv导致显著活性丧失(图6B)。含有硒代半胱氨酸的scFv对还原条件具有强烈抗性(图6D)。用50mM DTT处理仅导致轻微亲和力损失(EC_{50} 7.95nM至 EC_{50} 11.4nM)。在图6B和6D中,在scFv浓度10nM处从顶部至底部的线代表0mM DTT、1mM DTT、10mM DTT和50mM DTT。

[0064] 图7A-7C:在大肠杆菌菌株BL21DE3(图7A,SEQ ID NO:30)、T7Shuffle Express(图7B,SEQ ID NO:30)和RT Δ A-2X310K(图7C,SEQ ID NO:31)中产生的曲妥珠单抗(赫赛汀)scFv的UVPD片段图谱。对于上文所示的193nm UVPD序列信息,共价键合(或潜在键合)的半胱氨酸残基(图7B)和共价键合的硒代半胱氨酸残基(图7C)以灰色阴影表示。(图7A)整个蛋白质序列中的均匀片段化表明未形成二硫键。(图7B)序列的后半部分缺乏片段化证实仅在VH区域中形成二硫键。尽管在氧化条件下表达,但在VL区中未形成二硫键。(图7C) VL区和VH区两者中缺乏片段化证实形成两个二硒键。只有二硒化物scFv采用了天然和预期的共价结构。

具体实施方式

[0065] 在蛋白质中使用非规范氨基酸提供了具有极大扩展的功能的多肽的可能性,所述功能可用于广泛应用。例如,通过将硒代半胱氨酸并入多肽中,可能有可能开发具有增强的稳定性或活性水平的酶并产生高活性治疗性多肽。然而,迄今为止,这些方法因不能产生稳定地保留翻译途径的生物体并可预测且可靠地将硒代半胱氨酸并入编码的多肽中而受到阻碍。因此,本公开通过提供能够有效产生硒代半胱氨酸蛋白的进化的大肠杆菌菌株和先进的大肠杆菌基因克服了与当前技术相关的挑战。特别地,首次可实现商业上相关量的硒蛋白,所述硒蛋白显示基本上完全并入编码的硒代半胱氨酸残基。还提供了能够实现这种产生的细菌菌株和突变基因。例如,在此提出的研究已经鉴别了基因中的若干点突变,包括acrB、adhE、arcB、cysK、dnaE、ftsA和ftsL,所述点突变优化富含硒的培养基中的存活并且含有确保维持对硒代半胱氨酸的依赖性的途径。一种示例性细菌菌株是突变体RT Δ A

2X310K大肠杆菌菌株，所述菌株含有这些核心突变，从而提供增强的生长速率、更大的最终细胞密度和使用这种菌株产生的多肽中大约100%并入硒代半胱氨酸。因此，本文提供了用于有效并入硒代半胱氨酸和用于产生并入硒代半胱氨酸位置的多肽的方法和组合物。

[0066] I. 用于硒代半胱氨酸并入的系统

[0067] 本公开的实施方案提供了一种用于并入硒代半胱氨酸，如用于产生硒蛋白的系统。在一方面，所述系统是细菌菌株，如大肠杆菌菌株，所述菌株包含一种或多种增强细胞生长或硒代半胱氨酸并入的基因内的一个或多个氨基酸取代或缺失。细菌菌株可以是大肠杆菌菌株K012 MG1655 (GenBank登录号U00096.3)、大肠杆菌菌株C321 Δ A (US2016/0060301) 或大肠杆菌菌株RT Δ A。在某些实施方案中，所述细菌菌株是RT Δ A 2X310K菌株。

[0068] 在特定方面，所述系统包含条件性地依赖于硒代半胱氨酸的进化的细菌菌株。在一个实例中，所述进化的细菌菌株源自菌株，如RT Δ A菌株，所述菌株包含基因组selA、selB 和 selC基因 (分别编码SelA、SelB和tRNA^{Sec}) 的缺失并且缺少编码释放因子1 (RF1) 的prfA基因，从而允许有效并入一系列非天然氨基酸。这些基因的缺失消除了UAG遏制tRNA (例如，tRNA^{SecUx}) 和内源性硒代半胱氨酸并入机制之间的任何串扰。此外，RT Δ A菌株含有硒代半胱氨酸的报告蛋白，其是来自阴沟肠杆菌的具有必需氢硒基-巯基或二硒键 (Δ aphC::nmca C69U C238或C69U C238U) 的NMC-Aβ-内酰胺酶，从而使得条件性依赖于硒代半胱氨酸并入来获得对β-内酰胺抗生素的抗性。对硒代半胱氨酸并入的条件依赖性防止另外有毒的硒代半胱氨酸生物合成和并入机制的丧失或减弱。硒代半胱氨酸并入的效率可通过表达用于位点特异性并入硒代半胱氨酸的进化的tRNA来增强 (例如，例如tRNA^{SecUX}；美国专利申请公布2017/0166945，以引用的方式并入本文)。进一步增强可包括各种基因的表达，包括但不限于selD、selA和/或pstK基因。例如，细菌菌株可表达大肠杆菌selD、大肠杆菌selA和詹氏甲烷球菌pstK。

[0069] 在某些实施方案中，所述硒代半胱氨酸并入系统编码具有一个或多个氨基酸缺失或取代的一种或多种大肠杆菌基因。所述基因可包括acrB (SEQ ID NO:1)、adhE (SEQ ID NO:2)、arcB (SEQ ID NO:3)、cysK (SEQ ID NO:4)、dnaE (SEQ ID NO:5)、ftsA (SEQ ID NO:6)、ftsL (SEQ ID NO:18)、hemA (SEQ ID NO:7)、mdfA (SEQ ID NO:8)、ompR (SEQ ID NO:9)、oxyR (SEQ ID NO:19)、pcnB (SEQ ID NO:10)、prfB (SEQ ID NO:17)、pta (SEQ ID NO:11)、queE (SEQ ID NO:69) 和/或ydiL (SEQ ID NO:13)。另外的基因包括pdxB (SEQ ID NO:14)、yeeJ (SEQ ID NO:15) 和yphG (SEQ ID NO:16)。例如，并入硒代半胱氨酸的细菌菌株可包含prfB基因中的T246A突变与至少一种hemA突变的组合。作为另一个实例，并入硒代半胱氨酸的细菌菌株可包含prfB基因中的T246A突变与至少一种hemA突变和至少一种yeeJ突变的组合。表1中描绘了这些基因的示例性氨基酸取代。

[0070] 表1：进化的菌株中富含的突变。

大肠杆菌基因	突变
AcrB (SEQ ID NO:1)	A999T、A457T、 L344P 、 L591P:A183T、F358L、G23R、M902I、 L230R、A889V:R620C、D174G
AdhE (SEQ ID NO:2)	R702H
ArcB (SEQ ID NO:3)	T655A 、R398C、T652A、R264C、 V21A
CysK (SEQ ID NO:4)	I45N、F290L、G271D、 T73A 、H153Y、 I45V、L284P、T73I、P68L、T69I、 R305C、F144S:A281V
DnaE (SEQ ID NO:5)	E916G、Y938H、R860H、 E781G 、 G925D、T470M
FtsA (SEQ ID NO:6)	T115A、G386R、 A136V 、Q155R、 P98S:E201A、A294T、H159Y、V112I
HemA (SEQ ID NO:7)	A173V、 V183A 、P196L、L180F、 D127A:S249A、A5V:P273L、A5V、 I176N
MdfA (SEQ ID NO:8)	M1R
OmpR (SEQ ID NO:9)	Y102C 、R15C、Q30R
PcnB (SEQ ID NO:10)	R105C
Pta (SEQ ID NO:11)	A119V、S535P、W373R、 P673L 、 P673S、S535P、A119T、D601A、 A119T:P673L、E103K、G31D、T662I、

[0071]

	E359K:N519D
QueE (SEQ ID NO:12)	A69G
YdiK (SEQ ID NO:13)	L107S
PdxB (SEQ ID NO:14)	V212I、D162N、T299A、C220R
YeeJ (SEQ ID NO:15)	S1467P、V1184A、S1730N、G1888D、 A352T、G374S、S1423R、D1502E、 R285Q、E470K、A939T、L1669F、 V2034I、S1713N、D1233N、A704T、 M1084I
YphG (SEQ ID NO:16)	S112F、T126A、P978L、E199G、 A476T、A735V
PrfB (SEQ ID NO:17)	T246A
FtsI (SEQ ID NO:18)	V545I 、D76G:A293V、A537T、 A3V:P311S、M471T、A228V:P311S、 A257T
OxyR (SEQ ID NO:19)	M193I、A233T、 L124P 、V300A、 C199R

[0072]

[0073] 呈粗体的突变是大肠杆菌RT Δ A 2X310K中的核心突变。

[0074] 在另外的方面,大肠杆菌基因可通过一个或多个其它氨基酸取代进一步修饰。例如,氨基酸取代可在一或多个位置进行,其中所述取代是用于具有相似亲水性的氨基酸。本领域中通常了解亲水氨基酸指数在对蛋白质赋予相互作用性生物功能方面的重要性(Kyte和Doolittle,1982)。可接受的是,氨基酸的相对亲水特性有助于所得蛋白质的次级结构,这进而界定蛋白质与其它分子(例如酶、底物、受体、DNA、抗体、抗原等)的相互作用。因此,这种保守取代可在含有硒代半胱氨酸的多肽中进行并且可能仅对其活性具有轻微影响。如美国专利4,554,101中所详述,以下亲水性值已指定给氨基酸残基:精氨酸(+3.0);赖氨酸(+3.0);天冬氨酸(+3.0±1);谷氨酸(+3.0±1);丝氨酸(+0.3);天冬酰胺(+0.2);谷氨酰胺(+0.2);甘氨酸(0);苏氨酸(-0.4);脯氨酸(-0.5±1);丙氨酸(0.5);组氨酸(-0.5);半胱氨酸(-1.0);甲硫氨酸(-1.3);缬氨酸(-1.5);亮氨酸(-1.8);异亮氨酸(-1.8);酪氨酸(-2.3);苯丙氨酸(-2.5);色氨酸(-3.4)。这些值可用作指导并且因此优选亲水性值在±2内的氨基酸的取代,特别优选亲水性值在±1内的那些,并且甚至更特别优选亲水性值在±0.5内的那些。因此,对于具有相似亲水性值的不同但同源的氨基酸,可通过氨基酸的取代来修饰本文所述的任何含硒代半胱氨酸的多肽。亲水性在+/-1.0或+/-0.5点内的氨基酸被认为是同源的。

[0075] II. 包含硒代半胱氨酸的重组多肽的产生

[0076] 本文提供的硒代半胱氨酸并入系统可用于产生包含硒代半胱氨酸的多肽,如抗体或二硫键键合的蛋白质。在一些情况下,硒代半胱氨酸残基可被天然存在的半胱氨酸取代,或在其中硒代半胱氨酸残基的取代不会改变多肽的结构和功能的任何位点(例如在丝氨酸残基处)被取代。在其它方面,根据实施方案的方法可用于产生天然地包含一个或多个硒代半胱氨酸残基的多肽,如人硒蛋白(例如,人TrxR)。

[0077] 在某些实施方案中,编码多肽的核酸(例如质粒)可被引入所述系统并在含硒的生长培养基中在细胞将至少一个硒代半胱氨酸残基并入多肽的条件下培养。然后可分离所得含硒代半胱氨酸的多肽并且如通过质谱法或x射线晶体学进行分析。

[0078] 可将核酸引入并维持在重组载体中的细胞中,所述重组载体能够根据标准技术在细胞中自主复制。转化方法和表达载体的选择将取决于待表达的多肽的性质和所选择的宿主系统。转化方法描述于例如Ausubel等人(编辑)Current Protocols in Molecular Biology(John Wiley&Sohs,New York,1994);表达媒介物可从本领域中熟知的那些中选择,例如Cloning Vectors:A Laboratory Manual(P.H.Pouwels等人,1985,增刊1987)中。为了诱导异源多肽的表达,细胞培养基通常含有1与50ng/ml之间、优选2至40ng/ml且最优选5至25ng/ml的硒。硒可以亚硒酸钠或另一种可溶的氧化形式的硒存在(例如,可使用0.1至50μM、优选5至25μM Na₂SeO₃)。“异源”核酸是指对其所引入的细胞或动物来说部分或完全外来的核酸,或与所述细胞或动物的内源性基因同源的核酸,除了异源蛋白质含有被至少一个氨基酸取代的硒代半胱氨酸。

[0079] 为了获得核酸序列的表达,可将所述序列并入具有一个或多个与核酸可操作连接的控制序列的载体中以控制其表达。所述载体可包括其它序列,如驱动插入的核酸、核酸序列的表达的启动子,以使得所述多肽或肽作为编码分泌信号的融合体和/或核酸产生,以使得在宿主细胞中产生的多肽从细胞中分泌。然后可通过以下方式来获得多肽:将载体转化

到载体有功能的宿主细胞中、培养所述宿主细胞以使得产生多肽并从所述宿主细胞或周围培养基中回收所述多肽。可用于本发明的实施方案中的原核细胞包括大肠杆菌，并且为了易于过表达，可优选具有T7RNA聚合酶的菌株。

[0080] 在一些情况下，硒代半胱氨酸残基可被天然存在的半胱氨酸取代，或在其中硒代半胱氨酸残基的取代不会改变多肽的结构和功能的任何位点(例如在丝氨酸残基处)被取代。此类氨基酸可通过本领域技术人员熟知的方法鉴别，并且通常将在不参与蛋白质的催化或结合活性的位置(如例如通过突变分析所确定)或在被认为对多肽的结构完整性至关重要的位置处发生。在特定多肽中，可将硒代半胱氨酸残基并入通常将被形成二硫桥的两个半胱氨酸残基占据的位置中；因此，所述硒代半胱氨酸残基将形成二硒键，所述二硒键可通过质谱法鉴别。使用标准技术，还可修饰硒代半胱氨酸以形成硒化物、硒亚砜、亚硒酸、硒酸、硒砜或硒代-硫基团。

[0081] 在另一个实施方案中，将硒代半胱氨酸残基并入多肽中的位点，在所述位点中已知(或预测)所述取代改变多肽的结构和/或功能。这可用于改进多肽的功能或特征，以帮助确定多肽的生物学功能、确定多肽的结构或用于纯化多肽，例如，以帮助鉴别多肽结构域、活性位点或用于药物或另一多肽的结合位点。

[0082] A. 具有硒代半胱氨酸的抗体

[0083] 包含硒代半胱氨酸的重组多肽可编码抗体，如单克隆抗体。在一些方面，所述硒代半胱氨酸残基可在于野生型抗体中将是半胱氨酸的位置处被取代。抗体可以是IgG、IgM、IgA或其抗原结合片段。在某些方面，抗体是Fab'、F(ab')₂、F(ab')₃、单价scFv、二价scFv或单结构域抗体。抗体可以是非人抗体、鼠抗体、人抗体、人源化抗体或脱免疫抗体。在一些情况下，抗体可与成像剂、化学治疗剂、毒素或放射性核素缀合。本文还提供了一种组合物，所述组合物包含在药学上可接受的载体中的本文所述的实施方案和方面的抗体。

[0084] 如本文所用，术语“抗体”意图广泛地指任何免疫结合剂，如IgG、IgM、IgA、IgD、IgE和遗传修饰的IgG以及包含保留抗原结合活性的抗体CDR结构域的多肽。抗体可选自由以下组成的组：嵌合抗体、亲和力成熟抗体、多克隆抗体、单克隆抗体、人源化抗体、人抗体或抗原结合抗体片段或者天然或合成配体。因此，通过已知的方式并且如本文所述，可产生含有至少一个硒代半胱氨酸残基的多克隆或单克隆抗体、抗体片段以及结合结构域和CDR(包括前述中的任一种的工程改造形式)。在一些方面，所述抗体包含可形成二硒键的至少两个硒代半胱氨酸残基。

[0085] 单克隆抗体是单一种类的抗体，其中每个抗体分子识别相同的表位，因为所有产生抗体的细胞都源自单一B淋巴细胞细胞系。用于产生单克隆抗体(MAb)的方法通常以与用于制备多克隆抗体的那些方式相同的方式开始。在一些实施方案中，啮齿动物如小鼠和大鼠用于产生单克隆抗体。在一些实施方案中，兔、绵羊或青蛙细胞用于产生单克隆抗体。大鼠的使用是众所周知的并且可提供某些优点。小鼠(例如，BALB/c小鼠)常规地使用并且通常给出高百分比的稳定融合体。

[0086] 在一个实施方案中，所述抗体是嵌合抗体，例如，包含来自非人供体的抗原结合序列的抗体，所述非人供体被移植至异源非人、人或人源化序列(例如，框架和/或恒定结构域序列)。已经开发了方法来用人起源的类似结构域置换单克隆抗体的轻链和重链恒定结构域，从而使外源抗体的可变区保持完整。或者，在针对人免疫球蛋白基因转基因的小鼠中产

生“完全人”单克隆抗体。还开发了方法来通过重组构建具有啮齿动物(例如小鼠)和人氨基酸序列的抗体可变结构域来将单克隆抗体的可变结构域转化为更具人形式。在“人源化”单克隆抗体中,仅高变CDR源自小鼠单克隆抗体,并且框架和恒定区源自人氨基酸序列(参见美国专利号5,091,513和6,881,557)。据认为,用在人抗体的相应位置中发现的氨基酸序列置换抗体中为啮齿动物所特有的氨基酸序列将降低治疗使用期间不良免疫反应的可能性。杂交瘤或产生抗体的其它细胞也可经受基因突变或其它变化,所述基因突变或其它变化可以或可以不改变由所述杂交瘤产生的抗体的结合特异性。

[0087] 适用于本发明实施方案的抗体片段的实例包括但不限于:(i)由V_L、V_H、C_L和C_{H1}结构域组成的Fab片段;(ii)由V_H和C_{H1}结构域组成的“Fd”片段;(iii)由单一抗体的V_L和V_H结构域组成的“Fv”片段;(iv)“dAb”片段,其由V_H结构域组成;(v)分离的CDR区;(vi)F(ab')₂片段,其是包含两个连接的Fab片段的二价片段;(vii)单链Fv分子(“scFv”),其中V_H结构域和V_L结构域通过肽接头连接,所述肽接头允许两个结构域缔合以形成结合结构域;(viii)双特异性单链Fv二聚体(参见美国专利号5,091,513);以及(ix)通过基因融合构建的双抗体、多价或多特异性片段(美国专利申请公布20050214860)。可通过并入连接V_H和V_L结构域的二硫桥来稳定Fv、scFv或双抗体分子。还可制备包含与CH3结构域连接的scFv的微型抗体(Hu等人,1996)。

[0088] 还在实施方案中考虑了抗体样结合肽模拟物。Liu等人(2003)描述“抗体样结合肽模拟物”(ABiP),其是充当简化抗体并且具有较长血清半衰期以及较不麻烦的合成方法的某些优点的肽。

[0089] 抗体可从任何动物来源(包括鸟类和哺乳动物)产生。优选地,抗体是牛、绵羊、鼠、大鼠、兔、山羊、豚鼠、骆驼、马或鸡。此外,较新的技术允许从人组合抗体文库开发和筛选人抗体。例如,噬菌体抗体表达技术允许在没有动物免疫的情况下产生特异性抗体,如美国专利号6,946,546中所描述,所述专利以引用的方式并入本文。这些技术进一步描述于:Marks(1992);Stemmer(1994);Gram等人(1992);Barbas等人(1994);以及Schier等人(1996)。

[0090] 在一些方面,具有硒代半胱氨酸残基的抗体可以是用作治疗剂的抗体。例如,抗体可包含商业抗体治疗剂如西妥昔单抗的CDR序列。

[0091] III. 保藏信息

[0092] 大肠杆菌菌株RTΔA 2X310K的代表性冷冻保藏物已在2016年6月22日保藏于国家工业、食品和海洋细菌保藏中心(NCIMB),23 St. Machar Drive, Aberdeen AB21RY, Scotland, United Kingdom。那些沉积的细胞已被指定登录号NCIMB 42595。

[0093] 前述保藏是根据布达佩斯条约关于微生物保藏的条款和规定进行的,并且在保藏所接收供给保藏物的样品的最近一次请求后至少三十(30)年和至少五(05)年的保藏期或持续专利的有效期限(以较长者为准)收取,并且如果在此期间变得不可存活,则将予以更换。

[0094] IV. 实施例

[0095] 包括以下实施例以展示本发明的优选实施方案。本领域的普通技术人员应理解的是,在以下实施例中公开的技术代表由本发明人发现的在本发明的实践中起良好作用的技术,并且因此可被认为构成本发明实践的优选模式。然而,根据本公开,本领域的技术人员应理解,在不脱离本发明的精神和范围的情况下可在已公开并仍获得类似或相似结果的特

定实施方案中做出许多改变。

[0096] 实施例1-RT Δ A 2X310K的产生和表征

[0097] 产生若干细菌菌株用于硒代半胱氨酸并入且然后进化。先前描述的大肠杆菌RT Δ A菌株(Thyer等人,2015;以引用的方式并入本文)用于衍生更有效的硒代半胱氨酸并入菌株。简言之,RT Δ A菌株包含缺少编码释放因子1(RF1)的prfA基因的菌株C321. Δ A(Lajole等人,2013)中se1A、se1B和se1C基因(分别编码Se1A、Se1B和tRNA^{Sec})的缺失,从而允许有效并入一系列非天然氨基酸。这些基因的缺失消除了新tRNA^{Sec}文库与内源性硒代半胱氨酸并入机制之间的任何串扰。

[0098] 此外,RT Δ A菌株含有硒代半胱氨酸依赖性报告蛋白,其是来自阴沟肠杆菌的具有必需氢硒基-巯基或二硒键(Δ aphC::nmcA C69U C238或C69U C238U)的NMC-AB-内酰胺酶,从而产生对硒代半胱氨酸的条件性依赖。这种酶与来自粘质沙雷氏菌的SME-1 β -内酰胺酶具有高度序列相似性,所述SME-1 β -内酰胺酶是先前已被证明需要与活性位点丝氨酸残基相邻的二硫键来获得活性、但赋予大肠杆菌显著适合度成本的酶。观察到,含有整合的 β -内酰胺酶与天然二硫键的大肠杆菌菌株不需要硒补充来在 β -内酰胺抗生素羧苄青霉素存在下存活。具有一个或两个必需硒代半胱氨酸残基(其分别形成氢硒基-巯基或二硒键)的大肠杆菌菌株需要硒来获得对 β -内酰胺抗生素的抗性(图1A-1C)。对硒代半胱氨酸并入的条件依赖性防止另外有毒的硒代半胱氨酸生物合成和并入机制的丧失或减弱。

[0099] 引入了两种质粒,所述质粒提供硒代半胱氨酸并入所必需的机制。第一质粒含有CIoDF13复制起点并表达大肠杆菌se1D基因和用于位点特异性并入硒代半胱氨酸的进化tRNA(tRNA^{SecUX};美国专利申请公布2017/0166945)。第二质粒(含有RSF1030复制起点)表达大肠杆菌se1A基因和詹氏甲烷球菌pstK基因。在用这两种质粒转化后,含有零个、一个或两个硒代半胱氨酸残基的NMC-AB内酰胺酶代替形成半胱氨酸残基的必需二硫键被整合到aphC基因座处的基因组中。这赋予对一些 β -内酰胺抗生素的硒代半胱氨酸依赖性抗性,并且用于设定细胞中硒代半胱氨酸并入的最小阈值。

[0100] 将所得菌株(一式三份,与对照菌株一起)在两种不同条件下进化超过2500代(205次传代至汇合):增加抗生素浓度或增加温度。在进化后,对所有菌株进行全基因组测序。菌株在编码区内含有一至三百个非同义突变。将高度富集的这些突变的子集(存在于多个独立品系中)引入亲本菌株以用于表征。这些突变(参见以上表1)赋予增加的生长速率、活力、亚硒酸盐抗性或其它有益特征。与各种不同条件下的亲本菌株相比,所有进化的菌株显示出显著改善的生长。

[0101] 分离来自其中一个品系的单克隆(命名为2X310K)并用于比较这些进化的菌株用于重组硒蛋白产生的潜力。这种菌株能够以显著产率产生含有二硒化物的蛋白质。虽然含有整合的NMC-AB内酰胺酶与两个硒代半胱氨酸残基的亲本RT Δ A菌株显示出对 β -内酰胺抗生素羧苄青霉素的中等抗性,但突变体2X310K菌株显示出对羧苄青霉素的最大抗性(图2A-2B)。这种改善的生长使突变菌株成为用于产生重组硒蛋白的优良宿主。此外,虽然亲本菌株显示出延长的迟滞期并在限定培养基中生长不良,但突变体2X310K菌株显示出更高的生长速率和最终的细胞密度(图3A-3B)。

[0102] 为了监测硒代半胱氨酸并入的效率并证明蛋白质工程改造的可能性,产生了含有工程改造的非必需氢硒基-巯基键的大肠杆菌二氢叶酸还原酶(DHFR)。自上而下质谱法显

示接近100%的硒代半胱氨酸并入而无对应于含丝氨酸的DHFR的可检测的背景。观察到两个硒代半胱氨酸残基分别在位置39和85以大约100%的效率并入(图4A-4B)。所述分析还证实存在二硒键。

[0103] 突变体2X310K菌株用于在细菌细胞质中产生二硒化物稳定的抗-MS2抗体片段。与被开发用于表达含有二硫键的蛋白质的大肠杆菌菌株不同,这种菌株不需要升高的细胞质氧化还原电位。实验确定的单同位素质量与含有形成两个二硒键的四个硒代半胱氨酸残基的重组抗-MS2scFv一致(图5A-5C)。因此,所述菌株可有效地产生稳定的硒蛋白。此外,将菌株工程改造为含有T7RNA聚合酶,这将进一步增加重组硒蛋白的产量。

[0104] 突变体2X310K菌株用于产生硒代抗蓖麻毒素A链scFv。在UVPD序列信息中硒代半胱氨酸残基之间不存在片段化(如图6C中所示)证实两个二硒键的形成。硒代抗蓖麻毒素A链scFv的键连接性与野生型抗蓖麻毒素A链scFv相同,如图6A中所示。然而,用DTT处理野生型抗蓖麻毒素A链scFv导致显著活性丧失(图6B),而含有硒代半胱氨酸的scFv对还原条件具有强烈抗性(图6D)。

[0105] 大肠杆菌菌株BL21DE3、T7Shuffle Express和RT Δ A-2X310K用于产生曲妥珠单抗(赫赛汀)scFv。在BL21DE3菌株中,未形成二硫键,如可通过图7A中蛋白质序列的均匀片段化所观察到。在T7Shuffle Express菌株中,仅在VH区域中形成二硫键,如可通过图7B中序列的后半部分中缺乏片段化所观察到。然而,在RT Δ A-2X310K菌株中,形成两个二硒键,如可通过图7C中VL区和VH区两者中缺乏片段化所观察到。因此,只有二硒化物scFv采用了天然和预期的共价结构。

[0106] ***

[0107] 根据本公开,本文公开且要求保护的所有方法可在无需过度实验的情况下进行和实施。尽管本发明的组合物和方法已经根据优选实施方案进行了描述,但对本领域技术人员显而易见的是可使本文所述的方法和本文所述方法的步骤或步骤的顺序发生变化,而不偏离本发明的概念、精神和范围。更具体地说,显而易见的是在化学上和生理学上相关的某些试剂可取代本文所述的试剂,同时达到相同或相似结果。对于本领域技术人员显而易见的所有此类相似的替换和修改均被认为是在所附权利要求书所限定的本发明的精神、范围和概念内。

[0108] 参考文献

[0109] 以下参考文献以引用方式特别并入本文,在某种程度上,它们提供示例性程序或对本文所阐述的那些进行补充的其它细节。

[0110] 美国专利号4,554,101

[0111] 美国专利号5,091,513

[0112] 美国专利号6,946,546

[0113] 美国专利号6,881,557

[0114] 美国专利申请公布2005/0214860

[0115] 美国专利申请公布2017/0166945

[0116] Ausubel等人(编辑)Current Protocols in Molecular Biology(John Wiley& Sons, New York, 1994)。

[0117] Barbas等人, Proc. Natl. Acad. Sci., USA, 91:3809-13, 1994.

- [0118] Cloning Vectors:A Laboratory Manual (P.H.Pouwels等人,1985,增刊1987) .
- [0119] Gram等人,Proc.Natl.Acad.Sci.,USA,89:3576-80,1992.
- [0120] Hu等人,Cancer Res.,56:3055-61,1996.
- [0121] Kyte和Doolittle,J.Mol.Biol.,157 (1) :105-32,1982.
- [0122] Liu等人,Cell Mol.Biol.,49 (2) :209-16,2003.
- [0123] Marks等人,J.Biol.Chem.,267:16007-101992.
- [0124] Schier等人,J.Mol.Biol.,263:551-67,1996.
- [0125] Stemmer,Nature,370:389-91,1994.
- [0126] Thyer等人,J.Am.Chem.Soc.,137 (1) :46-49,2015.

<110> BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM

<120> 包含硒代半胱氨酸的重组多肽及其产生方法

<130> UTRB.P1120WO

<150> US 62/360,745
<151> 2016-07-11

<160> 31

<170> PatentIn 3.5版

<210> 1
<211> 1049

<212> PRT

<213> 大肠杆菌

<400> 1

Met Pro Asn Phe Phe Ile Asp Arg Pro Ile Phe Ala Trp Val Ile Ala
1 5 10 15

Ile Ile Ile Met Leu Ala Gly Gly Leu Ala Ile Leu Lys Leu Pro Val
20 25 30

Ala Gln Tyr Pro Thr Ile Ala Pro Pro Ala Val Thr Ile Ser Ala Ser
35 40 45

Tyr Pro Gly Ala Asp Ala Lys Thr Val Gln Asp Thr Val Thr Gln Val
50 55 60

[0001] Ile Glu Gln Asn Met Asn Gly Ile Asp Asn Leu Met Tyr Met Ser Ser
65 70 75 80

Asn Ser Asp Ser Thr Gly Thr Val Gln Ile Thr Leu Thr Phe Glu Ser
85 90 95

Gly Thr Asp Ala Asp Ile Ala Gln Val Gln Val Gln Asn Lys Leu Gln
100 105 110

Leu Ala Met Pro Leu Leu Pro Gln Glu Val Gln Gln Gln Gly Val Ser
115 120 125

Val Glu Lys Ser Ser Ser Phe Leu Met Val Val Gly Val Ile Asn
130 135 140

Thr Asp Gly Thr Met Thr Gln Glu Asp Ile Ser Asp Tyr Val Ala Ala
145 150 155 160

Asn Met Lys Asp Ala Ile Ser Arg Thr Ser Gly Val Gly Asp Val Gln
165 170 175

Leu Phe Gly Ser Gln Tyr Ala Met Arg Ile Trp Met Asn Pro Asn Glu
180 185 190

Leu Asn Lys Phe Gln Leu Thr Pro Val Asp Val Ile Thr Ala Ile Lys
195 200 205

Ala Gln Asn Ala Gln Val Ala Ala Gly Gln Leu Gly Gly Thr Pro Pro
210 215 220

Val Lys Gly Gln Gln Leu Asn Ala Ser Ile Ile Ala Gln Thr Arg Leu
 225 230 235 240

Thr Ser Thr Glu Glu Phe Gly Lys Ile Leu Leu Lys Val Asn Gln Asp
 245 250 255

Gly Ser Arg Val Leu Leu Arg Asp Val Ala Lys Ile Glu Leu Gly Gly
 260 265 270

Glu Asn Tyr Asp Ile Ile Ala Glu Phe Asn Gly Gln Pro Ala Ser Gly
 275 280 285

Leu Gly Ile Lys Leu Ala Thr Gly Ala Asn Ala Leu Asp Thr Ala Ala
 290 295 300

Ala Ile Arg Ala Glu Leu Ala Lys Met Glu Pro Phe Phe Pro Ser Gly
 305 310 315 320

Leu Lys Ile Val Tyr Pro Tyr Asp Thr Thr Pro Phe Val Lys Ile Ser
 325 330 335

Ile His Glu Val Val Lys Thr Leu Val Glu Ala Ile Ile Leu Val Phe
 340 345 350

Leu Val Met Tyr Leu Phe Leu Gln Asn Phe Arg Ala Thr Leu Ile Pro
 355 360 365

[0002] Thr Ile Ala Val Pro Val Val Leu Leu Gly Thr Phe Ala Val Leu Ala
 370 375 380

Ala Phe Gly Phe Ser Ile Asn Thr Leu Thr Met Phe Gly Met Val Leu
 385 390 395 400

Ala Ile Gly Leu Leu Val Asp Asp Ala Ile Val Val Val Glu Asn Val
 405 410 415

Glu Arg Val Met Ala Glu Glu Gly Leu Pro Pro Lys Glu Ala Thr Arg
 420 425 430

Lys Ser Met Gly Gln Ile Gln Gly Ala Leu Val Gly Ile Ala Met Val
 435 440 445

Leu Ser Ala Val Phe Val Pro Met Ala Phe Phe Gly Gly Ser Thr Gly
 450 455 460

Ala Ile Tyr Arg Gln Phe Ser Ile Thr Ile Val Ser Ala Met Ala Leu
 465 470 475 480

Ser Val Leu Val Ala Leu Ile Leu Thr Pro Ala Leu Cys Ala Thr Met
 485 490 495

Leu Lys Pro Ile Ala Lys Gly Asp His Gly Glu Gly Lys Lys Gly Phe
 500 505 510

Phe Gly Trp Phe Asn Arg Met Phe Glu Lys Ser Thr His His Tyr Thr
 515 520 525

Asp Ser Val Gly Gly Ile Leu Arg Ser Thr Gly Arg Tyr Leu Val Leu
 530 535 540

Tyr Leu Ile Ile Val Val Gly Met Ala Tyr Leu Phe Val Arg Leu Pro
 545 550 555 560

Ser Ser Phe Leu Pro Asp Glu Asp Gln Gly Val Phe Met Thr Met Val
 565 570 575

Gln Leu Pro Ala Gly Ala Thr Gln Glu Arg Thr Gln Lys Val Leu Asn
 580 585 590

Glu Val Thr His Tyr Tyr Leu Thr Lys Glu Lys Asn Asn Val Glu Ser
 595 600 605

Val Phe Ala Val Asn Gly Phe Gly Phe Ala Gly Arg Gly Gln Asn Thr
 610 615 620

Gly Ile Ala Phe Val Ser Leu Lys Asp Trp Ala Asp Arg Pro Gly Glu
 625 630 635 640

Glu Asn Lys Val Glu Ala Ile Thr Met Arg Ala Thr Arg Ala Phe Ser
 645 650 655

Gln Ile Lys Asp Ala Met Val Phe Ala Phe Asn Leu Pro Ala Ile Val
 660 665 670

[0003] Glu Leu Gly Thr Ala Thr Gly Phe Asp Phe Glu Leu Ile Asp Gln Ala
 675 680 685

Gly Leu Gly His Glu Lys Leu Thr Gln Ala Arg Asn Gln Leu Leu Ala
 690 695 700

Glu Ala Ala Lys His Pro Asp Met Leu Thr Ser Val Arg Pro Asn Gly
 705 710 715 720

Leu Glu Asp Thr Pro Gln Phe Lys Ile Asp Ile Asp Gln Glu Lys Ala
 725 730 735

Gln Ala Leu Gly Val Ser Ile Asn Asp Ile Asn Thr Thr Leu Gly Ala
 740 745 750

Ala Trp Gly Gly Ser Tyr Val Asn Asp Phe Ile Asp Arg Gly Arg Val
 755 760 765

Lys Lys Val Tyr Val Met Ser Glu Ala Lys Tyr Arg Met Leu Pro Asp
 770 775 780

Asp Ile Gly Asp Trp Tyr Val Arg Ala Ala Asp Gly Gln Met Val Pro
 785 790 795 800

Phe Ser Ala Phe Ser Ser Ser Arg Trp Glu Tyr Gly Ser Pro Arg Leu
 805 810 815

Glu Arg Tyr Asn Gly Leu Pro Ser Met Glu Ile Leu Gly Gln Ala Ala
 820 825 830

Pro Gly Lys Ser Thr Gly Glu Ala Met Glu Leu Met Glu Gln Leu Ala
835 840 845

Ser Lys Leu Pro Thr Gly Val Gly Tyr Asp Trp Thr Gly Met Ser Tyr
850 855 860

Gln Glu Arg Leu Ser Gly Asn Gln Ala Pro Ser Leu Tyr Ala Ile Ser
865 870 875 880

Leu Ile Val Val Phe Leu Cys Leu Ala Ala Leu Tyr Glu Ser Trp Ser
885 890 895

Ile Pro Phe Ser Val Met Leu Val Val Pro Leu Gly Val Ile Gly Ala
900 905 910

Leu Leu Ala Ala Thr Phe Arg Gly Leu Thr Asn Asp Val Tyr Phe Gln
915 920 925

Val Gly Leu Leu Thr Thr Ile Gly Leu Ser Ala Lys Asn Ala Ile Leu
930 935 940

Ile Val Glu Phe Ala Lys Asp Leu Met Asp Lys Glu Gly Lys Gly Leu
945 950 955 960

Ile Glu Ala Thr Leu Asp Ala Val Arg Met Arg Leu Arg Pro Ile Leu
965 970 975

[0004] Met Thr Ser Leu Ala Phe Ile Leu Gly Val Met Pro Leu Val Ile Ser
980 985 990

Thr Gly Ala Gly Ser Gly Ala Gln Asn Ala Val Gly Thr Gly Val Met
995 1000 1005

Gly Gly Met Val Thr Ala Thr Val Leu Ala Ile Phe Phe Val Pro
1010 1015 1020

Val Phe Phe Val Val Val Arg Arg Arg Phe Ser Arg Lys Asn Glu
1025 1030 1035

Asp Ile Glu His Ser His Thr Val Asp His His
1040 1045

<210> 2
<211> 891
<212> PRT
<213> 大肠杆菌

<400> 2

Met Ala Val Thr Asn Val Ala Glu Leu Asn Ala Leu Val Glu Arg Val
1 5 10 15

Lys Lys Ala Gln Arg Glu Tyr Ala Ser Phe Thr Gln Glu Gln Val Asp
20 25 30

Lys Ile Phe Arg Ala Ala Leu Ala Ala Asp Ala Arg Ile Pro
35 40 45

Leu Ala Lys Met Ala Val Ala Glu Ser Gly Met Gly Ile Val Glu Asp
 50 55 60

Lys Val Ile Lys Asn His Phe Ala Ser Glu Tyr Ile Tyr Asn Ala Tyr
 65 70 75 80

Lys Asp Glu Lys Thr Cys Gly Val Leu Ser Glu Asp Asp Thr Phe Gly
 85 90 95

Thr Ile Thr Ile Ala Glu Pro Ile Gly Ile Ile Cys Gly Ile Val Pro
 100 105 110

Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu
 115 120 125

Lys Thr Arg Asn Ala Ile Ile Phe Ser Pro His Pro Arg Ala Lys Asp
 130 135 140

Ala Thr Asn Lys Ala Ala Asp Ile Val Leu Gln Ala Ala Ile Ala Ala
 145 150 155 160

Gly Ala Pro Lys Asp Leu Ile Gly Trp Ile Asp Gln Pro Ser Val Glu
 165 170 175

Leu Ser Asn Ala Leu Met His His Pro Asp Ile Asn Leu Ile Leu Ala
 180 185 190

[0005] Thr Gly Gly Pro Gly Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro
 195 200 205

Ala Ile Gly Val Gly Ala Gly Asn Thr Pro Val Val Ile Asp Glu Thr
 210 215 220

Ala Asp Ile Lys Arg Ala Val Ala Ser Val Leu Met Ser Lys Thr Phe
 225 230 235 240

Asp Asn Gly Val Ile Cys Ala Ser Glu Gln Ser Val Val Val Asp
 245 250 255

Ser Val Tyr Asp Ala Val Arg Glu Arg Phe Ala Thr His Gly Gly Tyr
 260 265 270

Leu Leu Gln Gly Lys Glu Leu Lys Ala Val Gln Asp Val Ile Leu Lys
 275 280 285

Asn Gly Ala Leu Asn Ala Ala Ile Val Gly Gln Pro Ala Tyr Lys Ile
 290 295 300

Ala Glu Leu Ala Gly Phe Ser Val Pro Glu Asn Thr Lys Ile Leu Ile
 305 310 315 320

Gly Glu Val Thr Val Val Asp Glu Ser Glu Pro Phe Ala His Glu Lys
 325 330 335

Leu Ser Pro Thr Leu Ala Met Tyr Arg Ala Lys Asp Phe Glu Asp Ala
 340 345 350

Val Glu Lys Ala Glu Lys Leu Val Ala Met Gly Gly Ile Gly His Thr
 355 360 365

Ser Cys Leu Tyr Thr Asp Gln Asp Asn Gln Pro Ala Arg Val Ser Tyr
 370 375 380

Phe Gly Gln Lys Met Lys Thr Ala Arg Ile Leu Ile Asn Thr Pro Ala
 385 390 395 400

Ser Gln Gly Gly Ile Gly Asp Leu Tyr Asn Phe Lys Leu Ala Pro Ser
 405 410 415

Leu Thr Leu Gly Cys Gly Ser Trp Gly Gly Asn Ser Ile Ser Glu Asn
 420 425 430

Val Gly Pro Lys His Leu Ile Asn Lys Lys Thr Val Ala Lys Arg Ala
 435 440 445

Glu Asn Met Leu Trp His Lys Leu Pro Lys Ser Ile Tyr Phe Arg Arg
 450 455 460

Gly Ser Leu Pro Ile Ala Leu Asp Glu Val Ile Thr Asp Gly His Lys
 465 470 475 480

Arg Ala Leu Ile Val Thr Asp Arg Phe Leu Phe Asn Asn Gly Tyr Ala
 485 490 495

[0006] Asp Gln Ile Thr Ser Val Leu Lys Ala Ala Gly Val Glu Thr Glu Val
 500 505 510

Phe Phe Glu Val Glu Ala Asp Pro Thr Leu Ser Ile Val Arg Lys Gly
 515 520 525

Ala Glu Leu Ala Asn Ser Phe Lys Pro Asp Val Ile Ile Ala Leu Gly
 530 535 540

Gly Gly Ser Pro Met Asp Ala Ala Lys Ile Met Trp Val Met Tyr Glu
 545 550 555 560

His Pro Glu Thr His Phe Glu Glu Leu Ala Leu Arg Phe Met Asp Ile
 565 570 575

Arg Lys Arg Ile Tyr Lys Phe Pro Lys Met Gly Val Lys Ala Lys Met
 580 585 590

Ile Ala Val Thr Thr Ser Gly Thr Gly Ser Glu Val Thr Pro Phe
 595 600 605

Ala Val Val Thr Asp Asp Ala Thr Gly Gln Lys Tyr Pro Leu Ala Asp
 610 615 620

Tyr Ala Leu Thr Pro Asp Met Ala Ile Val Asp Ala Asn Leu Val Met
 625 630 635 640

Asp Met Pro Lys Ser Leu Cys Ala Phe Gly Gly Leu Asp Ala Val Thr
 645 650 655

His Ala Met Glu Ala Tyr Val Ser Val Leu Ala Ser Glu Phe Ser Asp
660 665 670

Gly Gln Ala Leu Gln Ala Leu Lys Leu Leu Lys Glu Tyr Leu Pro Ala
675 680 685

Ser Tyr His Glu Gly Ser Lys Asn Pro Val Ala Arg Glu Arg Val His
690 695 700

Ser Ala Ala Thr Ile Ala Gly Ile Ala Phe Ala Asn Ala Phe Leu Gly
705 710 715 720

Val Cys His Ser Met Ala His Lys Leu Gly Ser Gln Phe His Ile Pro
725 730 735

His Gly Leu Ala Asn Ala Leu Leu Ile Cys Asn Val Ile Arg Tyr Asn
740 745 750

Ala Asn Asp Asn Pro Thr Lys Gln Thr Ala Phe Ser Gln Tyr Asp Arg
755 760 765

Pro Gln Ala Arg Arg Arg Tyr Ala Glu Ile Ala Asp His Leu Gly Leu
770 775 780

Ser Ala Pro Gly Asp Arg Thr Ala Ala Lys Ile Glu Lys Leu Leu Ala
785 790 795 800

[0007] Trp Leu Glu Thr Leu Lys Ala Glu Leu Gly Ile Pro Lys Ser Ile Arg
805 810 815

Glu Ala Gly Val Gln Glu Ala Asp Phe Leu Ala Asn Val Asp Lys Leu
820 825 830

Ser Glu Asp Ala Phe Asp Asp Gln Cys Thr Gly Ala Asn Pro Arg Tyr
835 840 845

Pro Leu Ile Ser Glu Leu Lys Gln Ile Leu Leu Asp Thr Tyr Tyr Gly
850 855 860

Arg Asp Tyr Val Glu Gly Glu Thr Ala Ala Lys Lys Glu Ala Ala Pro
865 870 875 880

Ala Lys Ala Glu Lys Lys Ala Lys Lys Ser Ala
885 890

<210> 3

<211> 778

<212> PRT

<213> 大肠杆菌

<400> 3

Met Lys Gln Ile Arg Leu Leu Ala Gln Tyr Tyr Val Asp Leu Met Met
1 5 10 15

Lys Leu Gly Leu Val Arg Phe Ser Met Leu Leu Ala Leu Leu Val
20 25 30

Val Leu Ala Ile Val Val Gln Met Ala Val Thr Met Val Leu His Gly
 35 40 45

Gln Val Glu Ser Ile Asp Val Ile Arg Ser Ile Phe Phe Gly Leu Leu
 50 55 60

Ile Thr Pro Trp Ala Val Tyr Phe Leu Ser Val Val Val Glu Gln Leu
 65 70 75 80

Glu Glu Ser Arg Gln Arg Leu Ser Arg Leu Val Gln Lys Leu Glu Glu
 85 90 95

Met Arg Glu Arg Asp Leu Ser Leu Asn Val Gln Leu Lys Asp Asn Ile
 100 105 110

Ala Gln Leu Asn Gln Glu Ile Ala Val Arg Glu Lys Ala Glu Ala Glu
 115 120 125

Leu Gln Glu Thr Phe Gly Gln Leu Lys Ile Glu Ile Lys Glu Arg Glu
 130 135 140

Glu Thr Gln Ile Gln Leu Glu Gln Gln Ser Ser Phe Leu Arg Ser Phe
 145 150 155 160

Leu Asp Ala Ser Pro Asp Leu Val Phe Tyr Arg Asn Glu Asp Lys Glu
 165 170 175

[0008] Phe Ser Gly Cys Asn Arg Ala Met Glu Leu Leu Thr Gly Lys Ser Glu
 180 185 190

Lys Gln Leu Val His Leu Lys Pro Ala Asp Val Tyr Ser Pro Glu Ala
 195 200 205

Ala Ala Lys Val Ile Glu Thr Asp Glu Lys Val Phe Arg His Asn Val
 210 215 220

Ser Leu Thr Tyr Glu Gln Trp Leu Asp Tyr Pro Asp Gly Arg Lys Ala
 225 230 235 240

Cys Phe Glu Ile Arg Lys Val Pro Tyr Tyr Asp Arg Val Gly Lys Arg
 245 250 255

His Gly Leu Met Gly Phe Gly Arg Asp Ile Thr Glu Arg Lys Arg Tyr
 260 265 270

Gln Asp Ala Leu Glu Arg Ala Ser Arg Asp Lys Thr Thr Phe Ile Ser
 275 280 285

Thr Ile Ser His Glu Leu Arg Thr Pro Leu Asn Gly Ile Val Gly Leu
 290 295 300

Ser Arg Ile Leu Leu Asp Thr Glu Leu Thr Ala Glu Gln Glu Lys Tyr
 305 310 315 320

Leu Lys Thr Ile His Val Ser Ala Val Thr Leu Gly Asn Ile Phe Asn
 325 330 335

Asp Ile Ile Asp Met Asp Lys Met Glu Arg Arg Lys Val Gln Leu Asp
 340 345 350

Asn Gln Pro Val Asp Phe Thr Ser Phe Leu Ala Asp Leu Glu Asn Leu
 355 360 365

Ser Ala Leu Gln Ala Gln Gln Lys Gly Leu Arg Phe Asn Leu Glu Pro
 370 375 380

Thr Leu Pro Leu Pro His Gln Val Ile Thr Asp Gly Thr Arg Leu Arg
 385 390 395 400

Gln Ile Leu Trp Asn Leu Ile Ser Asn Ala Val Lys Phe Thr Gln Gln
 405 410 415

Gly Gln Val Thr Val Arg Val Arg Tyr Asp Glu Gly Asp Met Leu His
 420 425 430

Phe Glu Val Glu Asp Ser Gly Ile Gly Ile Pro Gln Asp Glu Leu Asp
 435 440 445

Lys Ile Phe Ala Met Tyr Tyr Gln Val Lys Asp Ser His Gly Gly Lys
 450 455 460

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

[0009] Lys Asn Met Gly Gly Asp Ile Thr Val Thr Ser Glu Gln Gly Lys Gly
 485 490 495

Ser Thr Phe Thr Leu Thr Ile His Ala Pro Ser Val Ala Glu Glu Val
 500 505 510

Asp Asp Ala Phe Asp Glu Asp Asp Met Pro Leu Pro Ala Leu Asn Val
 515 520 525

Leu Leu Val Glu Asp Ile Glu Leu Asn Val Ile Val Ala Arg Ser Val
 530 535 540

Leu Glu Lys Leu Gly Asn Ser Val Asp Val Ala Met Thr Gly Lys Ala
 545 550 555 560

Ala Leu Glu Met Phe Lys Pro Gly Glu Tyr Asp Leu Val Leu Leu Asp
 565 570 575

Ile Gln Leu Pro Asp Met Thr Gly Leu Asp Ile Ser Arg Glu Leu Thr
 580 585 590

Lys Arg Tyr Pro Arg Glu Asp Leu Pro Pro Leu Val Ala Leu Thr Ala
 595 600 605

Asn Val Leu Lys Asp Lys Gln Glu Tyr Leu Asn Ala Gly Met Asp Asp
 610 615 620

Val Leu Ser Lys Pro Leu Ser Val Pro Ala Leu Thr Ala Met Ile Lys
 625 630 635 640

Lys Phe Trp Asp Thr Gln Asp Asp Glu Glu Ser Thr Val Thr Thr Glu
645 650 655

Glu Asn Ser Lys Ser Glu Ala Leu Leu Asp Ile Pro Met Leu Glu Gln
660 665 670

Tyr Leu Glu Leu Val Gly Pro Lys Leu Ile Thr Asp Gly Leu Ala Val
675 680 685

Phe Glu Lys Met Met Pro Gly Tyr Val Ser Val Leu Glu Ser Asn Leu
690 695 700

Thr Ala Gln Asp Lys Lys Gly Ile Val Glu Glu Gly His Lys Ile Lys
705 710 715 720

Gly Ala Ala Gly Ser Val Gly Leu Arg His Leu Gln Gln Leu Gly Gln
725 730 735

Gln Ile Gln Ser Pro Asp Leu Pro Ala Trp Glu Asp Asn Val Gly Glu
740 745 750

Trp Ile Glu Glu Met Lys Glu Glu Trp Arg His Asp Val Glu Val Leu
755 760 765

Lys Ala Trp Val Ala Lys Ala Thr Lys Lys
770 775

[0010] <210> 4
<211> 323
<212> PRT
<213> 大肠杆菌
<400> 4

Met Ser Lys Ile Phe Glu Asp Asn Ser Leu Thr Ile Gly His Thr Pro
1 5 10 15

Leu Val Arg Leu Asn Arg Ile Gly Asn Gly Arg Ile Leu Ala Lys Val
20 25 30

Glu Ser Arg Asn Pro Ser Phe Ser Val Lys Cys Arg Ile Gly Ala Asn
35 40 45

Met Ile Trp Asp Ala Glu Lys Arg Gly Val Leu Lys Pro Gly Val Glu
50 55 60

Leu Val Glu Pro Thr Ser Gly Asn Thr Gly Ile Ala Leu Ala Tyr Val
65 70 75 80

Ala Ala Ala Arg Gly Tyr Lys Leu Thr Leu Thr Met Pro Glu Thr Met
85 90 95

Ser Ile Glu Arg Arg Lys Leu Leu Lys Ala Leu Gly Ala Asn Leu Val
100 105 110

Leu Thr Glu Gly Ala Lys Gly Met Lys Gly Ala Ile Gln Lys Ala Glu
115 120 125

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

130

135

140

Ser Asn Pro Ala Asn Pro Glu Ile His Glu Lys Thr Thr Gly Pro Glu
 145 150 155 160

Ile Trp Glu Asp Thr Asp Gly Gln Val Asp Val Phe Ile Ala Gly Val
 165 170 175

Gly Thr Gly Gly Thr Leu Thr Gly Val Ser Arg Tyr Ile Lys Gly Thr
 180 185 190

Lys Gly Lys Thr Asp Leu Ile Ser Val Ala Val Glu Pro Thr Asp Ser
 195 200 205

Pro Val Ile Ala Gln Ala Leu Ala Gly Glu Glu Ile Lys Pro Gly Pro
 210 215 220

His Lys Ile Gln Gly Ile Gly Ala Gly Phe Ile Pro Ala Asn Leu Asp
 225 230 235 240

Leu Lys Leu Val Asp Lys Val Ile Gly Ile Thr Asn Glu Glu Ala Ile
 245 250 255

Ser Thr Ala Arg Arg Leu Met Glu Glu Gly Ile Leu Ala Gly Ile
 260 265 270

Ser Ser Gly Ala Ala Val Ala Ala Leu Lys Leu Gln Glu Asp Glu
 275 280 285

[0011]

Ser Phe Thr Asn Lys Asn Ile Val Val Ile Leu Pro Ser Ser Gly Glu
 290 295 300

Arg Tyr Leu Ser Thr Ala Leu Phe Ala Asp Leu Phe Thr Glu Lys Glu
 305 310 315 320

Leu Gln Gln

<210> 5

<211> 1160

<212> PRT

<213> 大肠杆菌

<400> 5

Met Ser Glu Pro Arg Phe Val His Leu Arg Val His Ser Asp Tyr Ser
 1 5 10 15

Met Ile Asp Gly Leu Ala Lys Thr Ala Pro Leu Val Lys Ala Ala
 20 25 30

Ala Leu Gly Met Pro Ala Leu Ala Ile Thr Asp Phe Thr Asn Leu Cys
 35 40 45

Gly Leu Val Lys Phe Tyr Gly Ala Gly His Gly Ala Gly Ile Lys Pro
 50 55 60

Ile Val Gly Ala Asp Phe Asn Val Gln Cys Asp Leu Leu Gly Asp Glu
 65 70 75 80

Leu Thr His Leu Thr Val Leu Ala Ala Asn Asn Thr Gly Tyr Gln Asn
 85 90 95

Leu Thr Leu Leu Ile Ser Lys Ala Tyr Gln Arg Gly Tyr Gly Ala Ala
 100 105 110

Gly Pro Ile Ile Asp Arg Asp Trp Leu Ile Glu Leu Asn Glu Gly Leu
 115 120 125

Ile Leu Leu Ser Gly Gly Arg Met Gly Asp Val Gly Arg Ser Leu Leu
 130 135 140

Arg Gly Asn Ser Ala Leu Val Asp Glu Cys Val Ala Phe Tyr Glu Glu
 145 150 155 160

His Phe Pro Asp Arg Tyr Phe Leu Glu Leu Ile Arg Thr Gly Arg Pro
 165 170 175

Asp Glu Glu Ser Tyr Leu His Ala Ala Val Glu Leu Ala Glu Ala Arg
 180 185 190

Gly Leu Pro Val Val Ala Thr Asn Asp Val Arg Phe Ile Asp Ser Ser
 195 200 205

Asp Phe Asp Ala His Glu Ile Arg Val Ala Ile His Asp Gly Phe Thr
 210 215 220

[0012] Leu Asp Asp Pro Lys Arg Pro Arg Asn Tyr Ser Pro Gln Gln Tyr Met
 225 230 235 240

Arg Ser Glu Glu Glu Met Cys Glu Leu Phe Ala Asp Ile Pro Glu Ala
 245 250 255

Leu Ala Asn Thr Val Glu Ile Ala Lys Arg Cys Asn Val Thr Val Arg
 260 265 270

Leu Gly Glu Tyr Phe Leu Pro Gln Phe Pro Thr Gly Asp Met Ser Thr
 275 280 285

Glu Asp Tyr Leu Val Lys Arg Ala Lys Glu Gly Leu Glu Glu Arg Leu
 290 295 300

Ala Phe Leu Phe Pro Asp Glu Glu Glu Arg Leu Lys Arg Arg Pro Glu
 305 310 315 320

Tyr Asp Glu Arg Leu Glu Thr Glu Leu Gln Val Ile Asn Gln Met Gly
 325 330 335

Phe Pro Gly Tyr Phe Leu Ile Val Met Glu Phe Ile Gln Trp Ser Lys
 340 345 350

Asp Asn Gly Val Pro Val Gly Pro Gly Arg Gly Ser Gly Ala Gly Ser
 355 360 365

Leu Val Ala Tyr Ala Leu Lys Ile Thr Asp Leu Asp Pro Leu Glu Phe
 370 375 380

Asp Leu Leu Phe Glu Arg Phe Leu Asn Pro Glu Arg Val Ser Met Pro
 385 390 395 400

Asp Phe Asp Val Asp Phe Cys Met Glu Lys Arg Asp Gln Val Ile Glu
 405 410 415

His Val Ala Asp Met Tyr Gly Arg Asp Ala Val Ser Gln Ile Ile Thr
 420 425 430

Phe Gly Thr Met Ala Ala Lys Ala Val Ile Arg Asp Val Gly Arg Val
 435 440 445

Leu Gly His Pro Tyr Gly Phe Val Asp Arg Ile Ser Lys Leu Ile Pro
 450 455 460

Pro Asp Pro Gly Met Thr Leu Ala Lys Ala Phe Glu Ala Glu Pro Gln
 465 470 475 480

Leu Pro Glu Ile Tyr Glu Ala Asp Glu Glu Val Lys Ala Leu Ile Asp
 485 490 495

Met Ala Arg Lys Leu Glu Gly Val Thr Arg Asn Ala Gly Lys His Ala
 500 505 510

Gly Gly Val Val Ile Ala Pro Thr Lys Ile Thr Asp Phe Ala Pro Leu
 515 520 525

[0013] Tyr Cys Asp Glu Glu Gly Lys His Pro Val Thr Gln Phe Asp Lys Ser
 530 535 540

Asp Val Glu Tyr Ala Gly Leu Val Lys Phe Asp Phe Leu Gly Leu Arg
 545 550 555 560

Thr Leu Thr Ile Ile Asn Trp Ala Leu Glu Met Ile Asn Lys Arg Arg
 565 570 575

Ala Lys Asn Gly Glu Pro Pro Leu Asp Ile Ala Ala Ile Pro Leu Asp
 580 585 590

Asp Lys Lys Ser Phe Asp Met Leu Gln Arg Ser Glu Thr Thr Ala Val
 595 600 605

Phe Gln Leu Glu Ser Arg Gly Met Lys Asp Leu Ile Lys Arg Leu Gln
 610 615 620

Pro Asp Cys Phe Glu Asp Met Ile Ala Leu Val Ala Leu Phe Arg Pro
 625 630 635 640

Gly Pro Leu Gln Ser Gly Met Val Asp Asn Phe Ile Asp Arg Lys His
 645 650 655

Gly Arg Glu Glu Ile Ser Tyr Pro Asp Val Gln Trp Gln His Glu Ser
 660 665 670

Leu Lys Pro Val Leu Glu Pro Thr Tyr Gly Ile Ile Leu Tyr Gln Glu
 675 680 685

Gln Val Met Gln Ile Ala Gln Val Leu Ser Gly Tyr Thr Leu Gly Gly
 690 695 700

Ala Asp Met Leu Arg Arg Ala Met Gly Lys Lys Pro Glu Glu Met
 705 710 715 720

Ala Lys Gln Arg Ser Val Phe Ala Glu Gly Ala Glu Lys Asn Gly Ile
 725 730 735

Asn Ala Glu Leu Ala Met Lys Ile Phe Asp Leu Val Glu Lys Phe Ala
 740 745 750

Gly Tyr Gly Phe Asn Lys Ser His Ser Ala Ala Tyr Ala Leu Val Ser
 755 760 765

Tyr Gln Thr Leu Trp Leu Lys Ala His Tyr Pro Ala Glu Phe Met Ala
 770 775 780

Ala Val Met Thr Ala Asp Met Asp Asn Thr Glu Lys Val Val Gly Leu
 785 790 795 800

Val Asp Glu Cys Trp Arg Met Gly Leu Lys Ile Leu Pro Pro Asp Ile
 805 810 815

Asn Ser Gly Leu Tyr His Phe His Val Asn Asp Asp Gly Glu Ile Val
 820 825 830

[0014] Tyr Gly Ile Gly Ala Ile Lys Gly Val Gly Glu Gly Pro Ile Glu Ala
 835 840 845

Ile Ile Glu Ala Arg Asn Lys Gly Gly Tyr Phe Arg Glu Leu Phe Asp
 850 855 860

Leu Cys Ala Arg Thr Asp Thr Lys Lys Leu Asn Arg Arg Val Leu Glu
 865 870 875 880

Lys Leu Ile Met Ser Gly Ala Phe Asp Arg Leu Gly Pro His Arg Ala
 885 890 895

Ala Leu Met Asn Ser Leu Gly Asp Ala Leu Lys Ala Ala Asp Gln His
 900 905 910

Ala Lys Ala Glu Ala Ile Gly Gln Ala Asp Met Phe Gly Val Leu Ala
 915 920 925

Glu Glu Pro Glu Gln Ile Glu Gln Ser Tyr Ala Ser Cys Gln Pro Trp
 930 935 940

Pro Glu Gln Val Val Leu Asp Gly Glu Arg Glu Thr Leu Gly Leu Tyr
 945 950 955 960

Leu Thr Gly His Pro Ile Asn Gln Tyr Leu Lys Glu Ile Glu Arg Tyr
 965 970 975

Val Gly Gly Val Arg Leu Lys Asp Met His Pro Thr Glu Arg Gly Lys
 980 985 990

Val Ile Thr Ala Ala Gly Leu Val Val Ala Ala Arg Val Met Val Thr
 995 1000 1005

Lys Arg Gly Asn Arg Ile Gly Ile Cys Thr Leu Asp Asp Arg Ser
 1010 1015 1020

Gly Arg Leu Glu Val Met Leu Phe Thr Asp Ala Leu Asp Lys Tyr
 1025 1030 1035

Gln Gln Leu Leu Glu Lys Asp Arg Ile Leu Ile Val Ser Gly Gln
 1040 1045 1050

Val Ser Phe Asp Asp Phe Ser Gly Gly Leu Lys Met Thr Ala Arg
 1055 1060 1065

Glu Val Met Asp Ile Asp Glu Ala Arg Glu Lys Tyr Ala Arg Gly
 1070 1075 1080

Leu Ala Ile Ser Leu Thr Asp Arg Gln Ile Asp Asp Gln Leu Leu
 1085 1090 1095

Asn Arg Leu Arg Gln Ser Leu Glu Pro His Arg Ser Gly Thr Ile
 1100 1105 1110

Pro Val His Leu Tyr Tyr Gln Arg Ala Asp Ala Arg Ala Arg Leu
 1115 1120 1125

[0015] Arg Phe Gly Ala Thr Trp Arg Val Ser Pro Ser Asp Arg Leu Leu
 1130 1135 1140

Asn Asp Leu Arg Gly Leu Ile Gly Ser Glu Gln Val Glu Leu Glu
 1145 1150 1155

Phe Asp
 1160

<210> 6
 <211> 420
 <212> PRT
 <213> 大肠杆菌

<400> 6

Met Ile Lys Ala Thr Asp Arg Lys Leu Val Val Gly Leu Glu Ile Gly
 1 5 10 15

Thr Ala Lys Val Ala Ala Leu Val Gly Glu Val Leu Pro Asp Gly Met
 20 25 30

Val Asn Ile Ile Gly Val Gly Ser Cys Pro Ser Arg Gly Met Asp Lys
 35 40 45

Gly Gly Val Asn Asp Leu Glu Ser Val Val Lys Cys Val Gln Arg Ala
 50 55 60

Ile Asp Gln Ala Glu Leu Met Ala Asp Cys Gln Ile Ser Ser Val Tyr
 65 70 75 80

Leu Ala Leu Ser Gly Lys His Ile Ser Cys Gln Asn Glu Ile Gly Met
 85 90 95

Val Pro Ile Ser Glu Glu Glu Val Thr Gln Glu Asp Val Glu Asn Val
 100 105 110

Val His Thr Ala Lys Ser Val Arg Val Arg Asp Glu His Arg Val Leu
 115 120 125

His Val Ile Pro Gln Glu Tyr Ala Ile Asp Tyr Gln Glu Gly Ile Lys
 130 135 140

Asn Pro Val Gly Leu Ser Gly Val Arg Met Gln Ala Lys Val His Leu
 145 150 155 160

Ile Thr Cys His Asn Asp Met Ala Lys Asn Ile Val Lys Ala Val Glu
 165 170 175

Arg Cys Gly Leu Lys Val Asp Gln Leu Ile Phe Ala Gly Leu Ala Ser
 180 185 190

Ser Tyr Ser Val Leu Thr Glu Asp Glu Arg Glu Leu Gly Val Cys Val
 195 200 205

Val Asp Ile Gly Gly Thr Met Asp Ile Ala Val Tyr Thr Gly Gly
 210 215 220

[0016] Ala Leu Arg His Thr Lys Val Ile Pro Tyr Ala Gly Asn Val Val Thr
 225 230 235 240

Ser Asp Ile Ala Tyr Ala Phe Gly Thr Pro Pro Ser Asp Ala Glu Ala
 245 250 255

Ile Lys Val Arg His Gly Cys Ala Leu Gly Ser Ile Val Gly Lys Asp
 260 265 270

Glu Ser Val Glu Val Pro Ser Val Gly Gly Arg Pro Pro Arg Ser Leu
 275 280 285

Gln Arg Gln Thr Leu Ala Glu Val Ile Glu Pro Arg Tyr Thr Glu Leu
 290 295 300

Leu Asn Leu Val Asn Glu Glu Ile Leu Gln Leu Gln Glu Lys Leu Arg
 305 310 315 320

Gln Gln Gly Val Lys His His Leu Ala Ala Gly Ile Val Leu Thr Gly
 325 330 335

Gly Ala Ala Gln Ile Glu Gly Leu Ala Ala Cys Ala Gln Arg Val Phe
 340 345 350

His Thr Gln Val Arg Ile Gly Ala Pro Leu Asn Ile Thr Gly Leu Thr
 355 360 365

Asp Tyr Ala Gln Glu Pro Tyr Tyr Ser Thr Ala Val Gly Leu Leu His
 370 375 380

Tyr Gly Lys Glu Ser His Leu Asn Gly Glu Ala Glu Val Glu Lys Arg
385 390 395 400

Val Thr Ala Ser Val Gly Ser Trp Ile Lys Arg Leu Asn Ser Trp Leu
405 410 415

Arg Lys Glu Phe
420

<210> 7
<211> 418
<212> PRT
<213> 大肠杆菌

<400> 7

Met Thr Leu Leu Ala Leu Gly Ile Asn His Lys Thr Ala Pro Val Ser
1 5 10 15

Leu Arg Glu Arg Val Ser Phe Ser Pro Asp Lys Leu Asp Gln Ala Leu
20 25 30

Asp Ser Leu Leu Ala Gln Pro Met Val Gln Gly Gly Val Val Leu Ser
35 40 45

Thr Cys Asn Arg Thr Glu Leu Tyr Leu Ser Val Glu Glu Gln Asp Asn
50 55 60

[0017] Leu Gln Glu Ala Leu Ile Arg Trp Leu Cys Asp Tyr His Asn Leu Asn
65 70 75 80

Glu Glu Asp Leu Arg Lys Ser Leu Tyr Trp His Gln Asp Asn Asp Ala
85 90 95

Val Ser His Leu Met Arg Val Ala Ser Gly Leu Asp Ser Leu Val Leu
100 105 110

Gly Glu Pro Gln Ile Leu Gly Gln Val Lys Lys Ala Phe Ala Asp Ser
115 120 125

Gln Lys Gly His Met Lys Ala Ser Glu Leu Glu Arg Met Phe Gln Lys
130 135 140

Ser Phe Ser Val Ala Lys Arg Val Arg Thr Glu Thr Asp Ile Gly Ala
145 150 155 160

Ser Ala Val Ser Val Ala Phe Ala Ala Cys Thr Leu Ala Arg Gln Ile
165 170 175

Phe Glu Ser Leu Ser Thr Val Thr Val Leu Leu Val Gly Ala Gly Glu
180 185 190

Thr Ile Glu Leu Val Ala Arg His Leu Arg Glu His Lys Val Gln Lys
195 200 205

Met Ile Ile Ala Asn Arg Thr Arg Glu Arg Ala Gln Ile Leu Ala Asp
210 215 220

Glu Val Gly Ala Glu Val Ile Ala Leu Ser Asp Ile Asp Glu Arg Leu
225 230 235 240

Arg Glu Ala Asp Ile Ile Ile Ser Ser Thr Ala Ser Pro Leu Pro Ile
245 250 255

Ile Gly Lys Gly Met Val Glu Arg Ala Leu Lys Ser Arg Arg Asn Gln
260 265 270

Pro Met Leu Leu Val Asp Ile Ala Val Pro Arg Asp Val Glu Pro Glu
275 280 285

Val Gly Lys Leu Ala Asn Ala Tyr Leu Tyr Ser Val Asp Asp Leu Gln
290 295 300

Ser Ile Ile Ser His Asn Leu Ala Gln Arg Lys Ala Ala Ala Val Glu
305 310 315 320

Ala Glu Thr Ile Val Ala Gln Glu Thr Ser Glu Phe Met Ala Trp Leu
325 330 335

Arg Ala Gln Ser Ala Ser Glu Thr Ile Arg Glu Tyr Arg Ser Gln Ala
340 345 350

Glu Gln Val Arg Asp Glu Leu Thr Ala Lys Ala Leu Ala Ala Leu Glu
355 360 365

Gln Gly Gly Asp Ala Gln Ala Ile Met Gln Asp Leu Ala Trp Lys Leu
370 375 380

[0018]

Thr Asn Arg Leu Ile His Ala Pro Thr Lys Ser Leu Gln Gln Ala Ala
385 390 395 400

Arg Asp Gly Asp Asn Glu Arg Leu Asn Ile Leu Arg Asp Ser Leu Gly
405 410 415

Leu Glu

<210> 8

<211> 410

<212> PRT

<213> 大肠杆菌

<400> 8

Met Gln Asn Lys Leu Ala Ser Gly Ala Arg Leu Gly Arg Gln Ala Leu
1 5 10 15

Leu Phe Pro Leu Cys Leu Val Leu Tyr Glu Phe Ser Thr Tyr Ile Gly
20 25 30

Asn Asp Met Ile Gln Pro Gly Met Leu Ala Val Val Glu Gln Tyr Gln
35 40 45

Ala Gly Ile Asp Trp Val Pro Thr Ser Met Thr Ala Tyr Leu Ala Gly
50 55 60

Gly Met Phe Leu Gln Trp Leu Leu Gly Pro Leu Ser Asp Arg Ile Gly

65	70	75	80
Arg Arg Pro Val Met Leu Ala Gly Val Val Trp Phe Ile Val Thr Cys			
85	90	95	
Leu Ala Ile Leu Leu Ala Gln Asn Ile Glu Gln Phe Thr Leu Leu Arg			
100	105	110	
Phe Leu Gln Gly Ile Ser Leu Cys Phe Ile Gly Ala Val Gly Tyr Ala			
115	120	125	
Ala Ile Gln Glu Ser Phe Glu Glu Ala Val Cys Ile Lys Ile Thr Ala			
130	135	140	
Leu Met Ala Asn Val Ala Leu Ile Ala Pro Leu Leu Gly Pro Leu Val			
145	150	155	160
Gly Ala Ala Trp Ile His Val Leu Pro Trp Glu Gly Met Phe Val Leu			
165	170	175	
Phe Ala Ala Leu Ala Ala Ile Ser Phe Phe Gly Leu Gln Arg Ala Met			
180	185	190	
Pro Glu Thr Ala Thr Arg Ile Gly Glu Lys Leu Ser Leu Lys Glu Leu			
195	200	205	
Gly Arg Asp Tyr Lys Leu Val Leu Lys Asn Gly Arg Phe Val Ala Gly			
210	215	220	
[0019]			
Ala Leu Ala Leu Gly Phe Val Ser Leu Pro Leu Leu Ala Trp Ile Ala			
225	230	235	240
Gln Ser Pro Ile Ile Ile Thr Gly Glu Gln Leu Ser Ser Tyr Glu			
245	250	255	
Tyr Gly Leu Leu Gln Val Pro Ile Phe Gly Ala Leu Ile Ala Gly Asn			
260	265	270	
Leu Leu Leu Ala Arg Leu Thr Ser Arg Arg Thr Val Arg Ser Leu Ile			
275	280	285	
Ile Met Gly Gly Trp Pro Ile Met Ile Gly Leu Leu Val Ala Ala Ala			
290	295	300	
Ala Thr Val Ile Ser Ser His Ala Tyr Leu Trp Met Thr Ala Gly Leu			
305	310	315	320
Ser Ile Tyr Ala Phe Gly Ile Gly Leu Ala Asn Ala Gly Leu Val Arg			
325	330	335	
Leu Thr Leu Phe Ala Ser Asp Met Ser Lys Gly Thr Val Ser Ala Ala			
340	345	350	
Met Gly Met Leu Gln Met Leu Ile Phe Thr Val Gly Ile Glu Ile Ser			
355	360	365	
Lys His Ala Trp Leu Asn Gly Gly Asn Gly Leu Phe Asn Leu Phe Asn			

370

375

380

Leu Val Asn Gly Ile Leu Trp Leu Ser Leu Met Val Ile Phe Leu Lys
 385 390 395 400

Asp Lys Gln Met Gly Asn Ser His Glu Gly
 405 410

<210> 9
 <211> 239
 <212> PRT
 <213> 大肠杆菌

<400> 9

Met Gln Glu Asn Tyr Lys Ile Leu Val Val Asp Asp Asp Met Arg Leu
 1 5 10 15

Arg Ala Leu Leu Glu Arg Tyr Leu Thr Glu Gln Gly Phe Gln Val Arg
 20 25 30

Ser Val Ala Asn Ala Glu Gln Met Asp Arg Leu Leu Thr Arg Glu Ser
 35 40 45

Phe His Leu Met Val Leu Asp Leu Met Leu Pro Gly Glu Asp Gly Leu
 50 55 60

Ser Ile Cys Arg Arg Leu Arg Ser Gln Ser Asn Pro Met Pro Ile Ile
 65 70 75 80

[0020] Met Val Thr Ala Lys Gly Glu Val Asp Arg Ile Val Gly Leu Glu
 85 90 95

Ile Gly Ala Asp Asp Tyr Ile Pro Lys Pro Phe Asn Pro Arg Glu Leu
 100 105 110

Leu Ala Arg Ile Arg Ala Val Leu Arg Arg Gln Ala Asn Glu Leu Pro
 115 120 125

Gly Ala Pro Ser Gln Glu Ala Val Ile Ala Phe Gly Lys Phe Lys
 130 135 140

Leu Asn Leu Gly Thr Arg Glu Met Phe Arg Glu Asp Glu Pro Met Pro
 145 150 155 160

Leu Thr Ser Gly Glu Phe Ala Val Leu Lys Ala Leu Val Ser His Pro
 165 170 175

Arg Glu Pro Leu Ser Arg Asp Lys Leu Met Asn Leu Ala Arg Gly Arg
 180 185 190

Glu Tyr Ser Ala Met Glu Arg Ser Ile Asp Val Gln Ile Ser Arg Leu
 195 200 205

Arg Arg Met Val Glu Glu Asp Pro Ala His Pro Arg Tyr Ile Gln Thr
 210 215 220

Val Trp Gly Leu Gly Tyr Val Phe Val Pro Asp Gly Ser Lys Ala
 225 230 235

<210> 10
 <211> 465
 <212> PRT
 <213> 大肠杆菌

<400> 10

Met	Phe	Thr	Arg	Val	Ala	Asn	Phe	Cys	Arg	Lys	Val	Leu	Ser	Arg	Glu
1				5				10			15				

Glu Ser Glu Ala Glu Gln Ala Val Ala Arg Pro Gln Val Thr Val Ile
 20 25 30

Pro Arg Glu Gln His Ala Ile Ser Arg Lys Asp Ile Ser Glu Asn Ala
 35 40 45

Leu Lys Val Met Tyr Arg Leu Asn Lys Ala Gly Tyr Glu Ala Trp Leu
 50 55 60

Val Gly Gly Val Arg Asp Leu Leu Leu Gly Lys Lys Pro Lys Asp
 65 70 75 80

Phe Asp Val Thr Thr Asn Ala Thr Pro Glu Gln Val Arg Lys Leu Phe
 85 90 95

Arg Asn Cys Arg Leu Val Gly Arg Arg Phe Arg Leu Ala His Val Met
 100 105 110

Phe Gly Pro Glu Ile Ile Glu Val Ala Thr Phe Arg Gly His His Glu
 [0021] 115 120 125

Gly Asn Val Ser Asp Arg Thr Thr Ser Gln Arg Gly Gln Asn Gly Met
 130 135 140

Leu Leu Arg Asp Asn Ile Phe Gly Ser Ile Glu Glu Asp Ala Gln Arg
 145 150 155 160

Arg Asp Phe Thr Ile Asn Ser Leu Tyr Tyr Ser Val Ala Asp Phe Thr
 165 170 175

Val Arg Asp Tyr Val Gly Gly Met Lys Asp Leu Lys Asp Gly Val Ile
 180 185 190

Arg Leu Ile Gly Asn Pro Glu Thr Arg Tyr Arg Glu Asp Pro Val Arg
 195 200 205

Met Leu Arg Ala Val Arg Phe Ala Ala Lys Leu Gly Met Arg Ile Ser
 210 215 220

Pro Glu Thr Ala Glu Pro Ile Pro Arg Leu Ala Thr Leu Leu Asn Asp
 225 230 235 240

Ile Pro Pro Ala Arg Leu Phe Glu Glu Ser Leu Lys Leu Leu Gln Ala
 245 250 255

Gly Tyr Gly Tyr Glu Thr Tyr Lys Leu Leu Cys Glu Tyr His Leu Phe
 260 265 270

Gln Pro Leu Phe Pro Thr Ile Thr Arg Tyr Phe Thr Glu Asn Gly Asp
275 280 285

Ser Pro Met Glu Arg Ile Ile Glu Gln Val Leu Lys Asn Thr Asp Thr
290 295 300

Arg Ile His Asn Asp Met Arg Val Asn Pro Ala Phe Leu Phe Ala Ala
305 310 315 320

Met Phe Trp Tyr Pro Leu Leu Glu Thr Ala Gln Lys Ile Ala Gln Glu
325 330 335

Ser Gly Leu Thr Tyr His Asp Ala Phe Ala Leu Ala Met Asn Asp Val
340 345 350

Leu Asp Glu Ala Cys Arg Ser Leu Ala Ile Pro Lys Arg Leu Thr Thr
355 360 365

Leu Thr Arg Asp Ile Trp Gln Leu Gln Leu Arg Met Ser Arg Arg Gln
370 375 380

Gly Lys Arg Ala Trp Lys Leu Leu Glu His Pro Lys Phe Arg Ala Ala
385 390 395 400

Tyr Asp Leu Leu Ala Leu Arg Ala Glu Val Glu Arg Asn Ala Glu Leu
405 410 415

[0022] Gln Arg Leu Val Lys Trp Trp Gly Glu Phe Gln Val Ser Ala Pro Pro
420 425 430

Asp Gln Lys Gly Met Leu Asn Glu Leu Asp Glu Glu Pro Ser Pro Arg
435 440 445

Arg Arg Thr Arg Arg Pro Arg Lys Arg Ala Pro Arg Arg Glu Gly Thr
450 455 460

Ala
465

<210> 11
<211> 714
<212> PRT
<213> 大肠杆菌

<400> 11

Met Ser Arg Ile Ile Met Leu Ile Pro Thr Gly Thr Ser Val Gly Leu
1 5 10 15

Thr Ser Val Ser Leu Gly Val Ile Arg Ala Met Glu Arg Lys Gly Val
20 25 30

Arg Leu Ser Val Phe Lys Pro Ile Ala Gln Pro Arg Thr Gly Gly Asp
35 40 45

Ala Pro Asp Gln Thr Thr Ile Val Arg Ala Asn Ser Ser Thr Thr
50 55 60

Thr Ala Ala Glu Pro Leu Lys Met Ser Tyr Val Glu Gly Leu Leu Ser
 65 70 75 80

Ser Asn Gln Lys Asp Val Leu Met Glu Glu Ile Val Ala Asn Tyr His
 85 90 95

Ala Asn Thr Lys Asp Ala Glu Val Val Leu Val Glu Gly Leu Val Pro
 100 105 110

Thr Arg Lys His Gln Phe Ala Gln Ser Leu Asn Tyr Glu Ile Ala Lys
 115 120 125

Thr Leu Asn Ala Glu Ile Val Phe Val Met Ser Gln Gly Thr Asp Thr
 130 135 140

Pro Glu Gln Leu Lys Glu Arg Ile Glu Leu Thr Arg Asn Ser Phe Gly
 145 150 155 160

Gly Ala Lys Asn Thr Asn Ile Thr Gly Val Ile Val Asn Lys Leu Asn
 165 170 175

Ala Pro Val Asp Glu Gln Gly Arg Thr Arg Pro Asp Leu Ser Glu Ile
 180 185 190

Phe Asp Asp Ser Ser Lys Ala Lys Val Asn Asn Val Asp Pro Ala Lys
 195 200 205

[0023] Leu Gln Glu Ser Ser Pro Leu Pro Val Leu Gly Ala Val Pro Trp Ser
 210 215 220

Phe Asp Leu Ile Ala Thr Arg Ala Ile Asp Met Ala Arg His Leu Asn
 225 230 235 240

Ala Thr Ile Ile Asn Glu Gly Asp Ile Asn Thr Arg Arg Val Lys Ser
 245 250 255

Val Thr Phe Cys Ala Arg Ser Ile Pro His Met Leu Glu His Phe Arg
 260 265 270

Ala Gly Ser Leu Leu Val Thr Ser Ala Asp Arg Pro Asp Val Leu Val
 275 280 285

Ala Ala Cys Leu Ala Ala Met Asn Gly Val Glu Ile Gly Ala Leu Leu
 290 295 300

Leu Thr Gly Gly Tyr Glu Met Asp Ala Arg Ile Ser Lys Leu Cys Glu
 305 310 315 320

Arg Ala Phe Ala Thr Gly Leu Pro Val Phe Met Val Asn Thr Asn Thr
 325 330 335

Trp Gln Thr Ser Leu Ser Leu Gln Ser Phe Asn Leu Glu Val Pro Val
 340 345 350

Asp Asp His Glu Arg Ile Glu Lys Val Gln Glu Tyr Val Ala Asn Tyr
 355 360 365

Ile Asn Ala Asp Trp Ile Glu Ser Leu Thr Ala Thr Ser Glu Arg Ser
 370 375 380

Arg Arg Leu Ser Pro Pro Ala Phe Arg Tyr Gln Leu Thr Glu Leu Ala
 385 390 395 400

Arg Lys Ala Gly Lys Arg Ile Val Leu Pro Glu Gly Asp Glu Pro Arg
 405 410 415

Thr Val Lys Ala Ala Ala Cys Ala Glu Arg Gly Ile Ala Thr Cys
 420 425 430

Val Leu Leu Gly Asn Pro Ala Glu Ile Asn Arg Val Ala Ala Ser Gln
 435 440 445

Gly Val Glu Leu Gly Ala Gly Ile Glu Ile Val Asp Pro Glu Val Val
 450 455 460

Arg Glu Ser Tyr Val Gly Arg Leu Val Glu Leu Arg Lys Asn Lys Gly
 465 470 475 480

Met Thr Glu Thr Val Ala Arg Glu Gln Leu Glu Asp Asn Val Val Leu
 485 490 495

Gly Thr Leu Met Leu Glu Gln Asp Glu Val Asp Gly Leu Val Ser Gly
 500 505 510

[0024] Ala Val His Thr Thr Ala Asn Thr Ile Arg Pro Pro Leu Gln Leu Ile
 515 520 525

Lys Thr Ala Pro Gly Ser Ser Leu Val Ser Ser Val Phe Phe Met Leu
 530 535 540

Leu Pro Glu Gln Val Tyr Val Tyr Gly Asp Cys Ala Ile Asn Pro Asp
 545 550 555 560

Pro Thr Ala Glu Gln Leu Ala Glu Ile Ala Ile Gln Ser Ala Asp Ser
 565 570 575

Ala Ala Ala Phe Gly Ile Glu Pro Arg Val Ala Met Leu Ser Tyr Ser
 580 585 590

Thr Gly Thr Ser Gly Ala Gly Ser Asp Val Glu Lys Val Arg Glu Ala
 595 600 605

Thr Arg Leu Ala Gln Glu Lys Arg Pro Asp Leu Met Ile Asp Gly Pro
 610 615 620

Leu Gln Tyr Asp Ala Ala Val Met Ala Asp Val Ala Lys Ser Lys Ala
 625 630 635 640

Pro Asn Ser Pro Val Ala Gly Arg Ala Thr Val Phe Ile Phe Pro Asp
 645 650 655

Leu Asn Thr Gly Asn Thr Thr Tyr Lys Ala Val Gln Arg Ser Ala Asp
 660 665 670

Leu Ile Ser Ile Gly Pro Met Leu Gln Gly Met Arg Lys Pro Val Asn
675 680 685

Asp Leu Ser Arg Gly Ala Leu Val Asp Asp Ile Val Tyr Thr Ile Ala
690 695 700

Leu Thr Ala Ile Gln Ser Ala Gln Gln Gln
705 710

<210> 12
<211> 223
<212> PRT
<213> 大肠杆菌

<400> 12

Met Gln Tyr Pro Ile Asn Glu Met Phe Gln Thr Leu Gln Gly Glu Gly
1 5 10 15

Tyr Phe Thr Gly Val Pro Ala Ile Phe Ile Arg Leu Gln Gly Cys Pro
20 25 30

Val Gly Cys Ala Trp Cys Asp Thr Lys His Thr Trp Glu Lys Leu Glu
35 40 45

Asp Arg Glu Val Ser Leu Phe Ser Ile Leu Ala Lys Thr Lys Glu Ser
50 55 60

Asp Lys Trp Gly Ala Ala Ser Ser Glu Asp Leu Leu Ala Val Ile Gly
65 70 75 80

[0025]

Arg Gln Gly Tyr Thr Ala Arg His Val Val Ile Thr Gly Gly Glu Pro
85 90 95

Cys Ile His Asp Leu Leu Pro Leu Thr Asp Leu Leu Glu Lys Asn Gly
100 105 110

Phe Ser Cys Gln Ile Glu Thr Ser Gly Thr His Glu Val Arg Cys Thr
115 120 125

Pro Asn Thr Trp Val Thr Val Ser Pro Lys Leu Asn Met Arg Gly Gly
130 135 140

Tyr Glu Val Leu Ser Gln Ala Leu Glu Arg Ala Asn Glu Ile Lys His
145 150 155 160

Pro Val Gly Arg Val Arg Asp Ile Glu Ala Leu Asp Glu Leu Leu Ala
165 170 175

Thr Leu Thr Asp Asp Lys Pro Arg Val Ile Ala Leu Gln Pro Ile Ser
180 185 190

Gln Lys Asp Asp Ala Thr Arg Leu Cys Ile Glu Thr Cys Ile Ala Arg
195 200 205

Asn Trp Arg Leu Ser Met Gln Thr His Lys Tyr Leu Asn Ile Ala
210 215 220

<210> 13

<211> 370

<212> PRT

<213> 大肠杆菌

<400> 13

Met Val Asn Val Arg Gln Pro Arg Asp Val Ala Gln Ile Leu Leu Ser			
1	5	10	15

Val Leu Phe Leu Ala Ile Met Ile Val Ala Cys Leu Trp Ile Val Gln		
20	25	30

Pro Phe Ile Leu Gly Phe Ala Trp Ala Gly Thr Val Val Ile Ala Thr		
35	40	45

Trp Pro Val Leu Leu Arg Leu Gln Lys Ile Met Phe Gly Arg Arg Ser		
50	55	60

Leu Ala Val Leu Val Met Thr Leu Leu Leu Val Met Val Phe Ile Ile			
65	70	75	80

Pro Ile Ala Leu Leu Val Asn Ser Ile Val Asp Gly Ser Gly Pro Leu		
85	90	95

Ile Lys Ala Ile Ser Ser Gly Asp Met Thr Leu Pro Asp Leu Ala Trp		
100	105	110

Leu Asn Thr Ile Pro Val Ile Gly Ala Lys Leu Tyr Ala Gly Trp His		
115	120	125

[0026] Asn Leu Leu Asp Met Gly Gly Thr Ala Ile Met Ala Lys Val Arg Pro

130	135	140
-----	-----	-----

Tyr Ile Gly Thr Thr Thr Trp Phe Val Gly Gln Ala Ala His Ile			
145	150	155	160

Gly Arg Phe Met Val His Cys Ala Leu Met Leu Leu Phe Ser Ala Leu		
165	170	175

Leu Tyr Trp Arg Gly Glu Gln Val Ala Gln Gly Ile Arg His Phe Ala		
180	185	190

Thr Arg Leu Ala Gly Val Arg Gly Asp Ala Ala Val Leu Leu Ala Ala		
195	200	205

Gln Ala Ile Arg Ala Val Ala Leu Gly Val Val Val Thr Ala Leu Val		
210	215	220

Gln Ala Val Leu Gly Gly Ile Gly Leu Ala Val Ser Gly Val Pro Tyr			
225	230	235	240

Ala Thr Leu Leu Thr Val Leu Met Ile Leu Ser Cys Leu Val Gln Leu		
245	250	255

Gly Pro Leu Pro Val Leu Ile Pro Ala Ile Ile Trp Leu Tyr Trp Thr		
260	265	270

Gly Asp Thr Thr Trp Gly Thr Val Leu Leu Val Trp Ser Gly Val Val		
275	280	285

Gly Thr Leu Asp Asn Val Ile Arg Pro Met Leu Ile Arg Met Gly Ala

290 295 300

Asp Leu Pro Leu Ile Leu Ile Leu Ser Gly Val Ile Gly Gly Leu Ile

305 310 315 320

Ala Phe Gly Met Ile Gly Leu Phe Ile Gly Pro Val Leu Leu Ala Val

325 330 335

Ser Trp Arg Leu Phe Ala Ala Trp Val Glu Glu Val Pro Pro Pro Thr

340 345 350

Asp Gln Pro Glu Glu Ile Leu Glu Glu Leu Gly Glu Ile Glu Lys Pro

355 360 365

Asn Lys

370

<210> 14

<211> 378

<212> PRT

<213> 大肠杆菌

<400> 14

Met Lys Ile Leu Val Asp Glu Asn Met Pro Tyr Ala Arg Asp Leu Phe

1 5 10 15

Ser Arg Leu Gly Glu Val Thr Ala Val Pro Gly Arg Pro Ile Pro Val
[0027]

35 40 45

Val Asn Glu Ser Leu Leu Ala Gly Lys Pro Ile Lys Phe Val Gly Thr

50 55 60

Ala Thr Ala Gly Thr Asp His Val Asp Glu Ala Trp Leu Lys Gln Ala

65 70 75 80

Gly Ile Gly Phe Ser Ala Ala Pro Gly Cys Asn Ala Ile Ala Val Val

85 90 95

Glu Tyr Val Phe Ser Ser Leu Leu Met Leu Ala Glu Arg Asp Gly Phe

100 105 110

Ser Leu Tyr Asp Arg Thr Val Gly Ile Val Gly Val Gly Asn Val Gly

115 120 125

Arg Arg Leu Gln Ala Arg Leu Glu Ala Leu Gly Ile Lys Thr Leu Leu

130 135 140

Cys Asp Pro Pro Arg Ala Asp Arg Gly Asp Glu Gly Asp Phe Arg Ser

145 150 155 160

Leu Asp Glu Leu Val Gln Arg Ala Asp Ile Leu Thr Phe His Thr Pro

165 170 175

Leu Phe Lys Asp Gly Pro Tyr Lys Thr Leu His Leu Ala Asp Glu Lys
180 185 190

Leu Ile Arg Ser Leu Lys Pro Gly Ala Ile Leu Ile Asn Ala Cys Arg
195 200 205

Gly Ala Val Val Asp Asn Thr Ala Leu Leu Thr Cys Leu Asn Glu Gly
210 215 220

Gln Lys Leu Ser Val Val Leu Asp Val Trp Glu Gly Glu Pro Glu Leu
225 230 235 240

Asn Val Glu Leu Leu Lys Lys Val Asp Ile Gly Thr Ser His Ile Ala
245 250 255

Gly Tyr Thr Leu Glu Gly Lys Ala Arg Gly Thr Thr Gln Val Phe Glu
260 265 270

Ala Tyr Ser Lys Phe Ile Gly His Glu Gln His Val Ala Leu Asp Thr
275 280 285

Leu Leu Pro Ala Pro Glu Phe Gly Arg Ile Thr Leu His Gly Pro Leu
290 295 300

Asp Gln Pro Thr Leu Lys Arg Leu Val His Leu Val Tyr Asp Val Arg
305 310 315 320

[0028] Arg Asp Asp Ala Pro Leu Arg Lys Val Ala Gly Ile Pro Gly Glu Phe
325 330 335

Asp Lys Leu Arg Lys Asn Tyr Leu Glu Arg Arg Glu Trp Ser Ser Leu
340 345 350

Tyr Val Ile Cys Asp Asp Ala Ser Ala Ala Ser Leu Leu Cys Lys Leu
355 360 365

Gly Phe Asn Ala Val His His Pro Ala Arg
370 375

<210> 15
<211> 2367
<212> PRT
<213> 大肠杆菌

<400> 15

Met Leu Ala Arg Ser Gly Lys Val Ser Met Ala Thr Lys Lys Arg Ser
1 5 10 15

Gly Glu Glu Ile Asn Asp Arg Gln Ile Leu Cys Gly Met Gly Ile Lys
20 25 30

Leu Arg Arg Leu Thr Ala Gly Ile Cys Leu Ile Thr Gln Leu Ala Phe
35 40 45

Pro Met Ala Ala Ala Ala Gln Gly Val Val Asn Ala Ala Thr Gln Gln
50 55 60

Pro Val Pro Ala Gln Ile Ala Ile Ala Asn Ala Asn Thr Val Pro Tyr
 65 70 75 80

Thr Leu Gly Ala Leu Glu Ser Ala Gln Ser Val Ala Glu Arg Phe Gly
 85 90 95

Ile Ser Val Ala Glu Leu Arg Lys Leu Asn Gln Phe Arg Thr Phe Ala
 100 105 110

Arg Gly Phe Asp Asn Val Arg Gln Gly Asp Glu Leu Asp Val Pro Ala
 115 120 125

Gln Val Ser Glu Lys Lys Leu Thr Pro Pro Pro Gly Asn Ser Ser Asp
 130 135 140

Asn Leu Glu Gln Gln Ile Ala Ser Thr Ser Gln Gln Ile Gly Ser Leu
 145 150 155 160

Leu Ala Glu Asp Met Asn Ser Glu Gln Ala Ala Asn Met Ala Arg Gly
 165 170 175

Trp Ala Ser Ser Gln Ala Ser Gly Ala Met Thr Asp Trp Leu Ser Arg
 180 185 190

Phe Gly Thr Ala Arg Ile Thr Leu Gly Val Asp Glu Asp Phe Ser Leu
 195 200 205

[0029] Lys Asn Ser Gln Phe Asp Phe Leu His Pro Trp Tyr Glu Thr Pro Asp
 210 215 220

Asn Leu Phe Phe Ser Gln His Thr Leu His Arg Thr Asp Glu Arg Thr
 225 230 235 240

Gln Ile Asn Asn Gly Leu Gly Trp Arg His Phe Thr Pro Thr Trp Met
 245 250 255

Ser Gly Ile Asn Phe Phe Asp His Asp Leu Ser Arg Tyr His Ser
 260 265 270

Arg Ala Gly Ile Gly Ala Glu Tyr Trp Arg Asp Tyr Leu Lys Leu Ser
 275 280 285

Ser Asn Gly Tyr Leu Arg Leu Thr Asn Trp Arg Ser Ala Pro Glu Leu
 290 295 300

Asp Asn Asp Tyr Glu Ala Arg Pro Ala Asn Gln Trp Asp Val Arg Ala
 305 310 315 320

Glu Ser Trp Leu Pro Ala Trp Pro His Leu Gly Gly Lys Leu Val Tyr
 325 330 335

Glu Gln Tyr Tyr Gly Asp Glu Val Ala Leu Phe Asp Lys Asp Asp Arg
 340 345 350

Gln Ser Asn Pro His Ala Ile Thr Ala Gly Leu Asn Tyr Thr Pro Phe
 355 360 365

Pro Leu Met Thr Phe Ser Ala Glu Gln Arg Gln Gly Lys Gln Gly Glu
 370 375 380

Asn Asp Thr Arg Phe Ala Val Asp Phe Thr Trp Gln Pro Gly Ser Ala
 385 390 395 400

Met Gln Lys Gln Leu Asp Pro Asn Glu Val Ala Ala Arg Arg Ser Leu
 405 410 415

Ala Gly Ser Arg Tyr Asp Leu Val Asp Arg Asn Asn Asn Ile Val Leu
 420 425 430

Glu Tyr Arg Lys Lys Glu Leu Val Arg Leu Thr Leu Thr Asp Pro Val
 435 440 445

Thr Gly Lys Ser Gly Glu Val Lys Ser Leu Val Ser Ser Leu Gln Thr
 450 455 460

Lys Tyr Ala Leu Lys Gly Tyr Asn Val Glu Ala Thr Ala Leu Glu Ala
 465 470 475 480

Ala Gly Gly Lys Val Val Thr Thr Gly Lys Asp Ile Leu Val Thr Leu
 485 490 495

Pro Ala Tyr Arg Phe Thr Ser Thr Pro Glu Thr Asp Asn Thr Trp Pro
 500 505 510

[0030] Ile Glu Val Thr Ala Glu Asp Val Lys Gly Asn Leu Ser Asn Arg Glu
 515 520 525

Gln Ser Met Val Val Val Gln Ala Pro Thr Leu Ser Gln Lys Asp Ser
 530 535 540

Ser Val Ser Leu Ser Thr Gln Thr Leu Asn Ala Asp Ser His Ser Thr
 545 550 555 560

Ala Thr Leu Thr Phe Ile Ala His Asp Ala Ala Gly Asn Pro Val Val
 565 570 575

Gly Leu Val Leu Ser Thr Arg His Glu Gly Val Gln Asp Ile Thr Leu
 580 585 590

Ser Asp Trp Lys Asp Asn Gly Asp Gly Ser Tyr Thr Gln Ile Leu Thr
 595 600 605

Thr Gly Ala Met Ser Gly Thr Leu Thr Leu Met Pro Gln Leu Asn Gly
 610 615 620

Val Asp Ala Ala Lys Ala Pro Ala Val Val Asn Ile Ile Ser Val Ser
 625 630 635 640

Ser Ser Arg Thr His Ser Ser Ile Lys Ile Asp Lys Asp Arg Tyr Leu
 645 650 655

Ser Gly Asn Pro Ile Glu Val Thr Val Glu Leu Arg Asp Glu Asn Asp
 660 665 670

Lys Pro Val Lys Glu Gln Lys Gln Gln Leu Asn Asn Ala Val Ser Ile
675 680 685

Asp Asn Val Lys Pro Gly Val Thr Thr Asp Trp Lys Glu Thr Ala Asp
690 695 700

Gly Val Tyr Lys Ala Thr Tyr Thr Ala Tyr Thr Lys Gly Ser Gly Leu
705 710 715 720

Thr Ala Lys Leu Leu Met Gln Asn Trp Asn Glu Asp Leu His Thr Ala
725 730 735

Gly Phe Ile Ile Asp Ala Asn Pro Gln Ser Ala Lys Ile Ala Thr Leu
740 745 750

Ser Ala Ser Asn Asn Gly Val Leu Ala Asn Glu Asn Ala Ala Asn Thr
755 760 765

Val Ser Val Asn Val Ala Asp Glu Gly Ser Asn Pro Ile Asn Asp His
770 775 780

Thr Val Thr Phe Ala Val Leu Ser Gly Ser Ala Thr Ser Phe Asn Asn
785 790 795 800

Gln Asn Thr Ala Lys Thr Asp Val Asn Gly Leu Ala Thr Phe Asp Leu
805 810 815

[0031] Lys Ser Ser Lys Gln Glu Asp Asn Thr Val Glu Val Thr Leu Glu Asn
820 825 830

Gly Val Lys Gln Thr Leu Ile Val Ser Phe Val Gly Asp Ser Ser Thr
835 840 845

Ala Gln Val Asp Leu Gln Lys Ser Lys Asn Glu Val Val Ala Asp Gly
850 855 860

Asn Asp Ser Val Thr Met Thr Ala Thr Val Arg Asp Ala Lys Gly Asn
865 870 875 880

Leu Leu Asn Asp Val Met Val Thr Phe Asn Val Asn Ser Ala Glu Ala
885 890 895

Lys Leu Ser Gln Thr Glu Val Asn Ser His Asp Gly Ile Ala Thr Ala
900 905 910

Thr Leu Thr Ser Leu Lys Asn Gly Asp Tyr Arg Val Thr Ala Ser Val
915 920 925

Ser Ser Gly Ser Gln Ala Asn Gln Gln Val Asn Phe Ile Gly Asp Gln
930 935 940

Ser Thr Ala Ala Leu Thr Leu Ser Val Pro Ser Gly Asp Ile Thr Val
945 950 955 960

Thr Asn Thr Ala Pro Gln Tyr Met Thr Ala Thr Leu Gln Asp Lys Asn
965 970 975

Gly Asn Pro Leu Lys Asp Lys Glu Ile Thr Phe Ser Val Pro Asn Asp
 980 985 990

Val Ala Ser Lys Phe Ser Ile Ser Asn Gly Gly Lys Gly Met Thr Asp
 995 1000 1005

Ser Asn Gly Val Ala Ile Ala Ser Leu Thr Gly Thr Leu Ala Gly
 1010 1015 1020

Thr His Met Ile Met Ala Arg Leu Ala Asn Ser Asn Val Ser Asp
 1025 1030 1035

Ala Gln Pro Met Thr Phe Val Ala Asp Lys Asp Arg Ala Val Val
 1040 1045 1050

Val Leu Gln Thr Ser Lys Ala Glu Ile Ile Gly Asn Gly Val Asp
 1055 1060 1065

Glu Thr Thr Leu Thr Ala Thr Val Lys Asp Pro Ser Asn His Pro
 1070 1075 1080

Val Ala Gly Ile Thr Val Asn Phe Thr Met Pro Gln Asp Val Ala
 1085 1090 1095

Ala Asn Phe Thr Leu Glu Asn Asn Gly Ile Ala Ile Thr Gln Ala
 1100 1105 1110

[0032] Asn Gly Glu Ala His Val Thr Leu Lys Gly Lys Lys Ala Gly Thr
 1115 1120 1125

His Thr Val Thr Ala Thr Leu Gly Asn Asn Asn Thr Ser Asp Ser
 1130 1135 1140

Gln Pro Val Thr Phe Val Ala Asp Lys Ala Ser Ala Gln Val Val
 1145 1150 1155

Leu Gln Ile Ser Lys Asp Glu Ile Thr Gly Asn Gly Val Asp Ser
 1160 1165 1170

Ala Thr Leu Thr Ala Thr Val Lys Asp Gln Phe Asp Asn Glu Val
 1175 1180 1185

Asn Asn Leu Pro Val Thr Phe Ser Ser Ala Ser Ser Gly Leu Thr
 1190 1195 1200

Leu Thr Pro Gly Val Ser Asn Thr Asn Glu Ser Gly Ile Ala Gln
 1205 1210 1215

Ala Thr Leu Ala Gly Val Ala Phe Gly Glu Lys Thr Val Thr Ala
 1220 1225 1230

Ser Leu Ala Asn Asn Gly Ala Ser Asp Asn Lys Thr Val His Phe
 1235 1240 1245

Ile Gly Asp Thr Ala Ala Ala Lys Ile Ile Glu Leu Ala Pro Val
 1250 1255 1260

Pro	Asp	Ser	Ile	Ile	Ala	Gly	Thr	Pro	Gln	Asn	Ser	Ser	Gly	Ser	
1265					1270						1275				
Val	Ile	Thr	Ala	Thr	Val	Val	Asp	Asn	Asn	Gly	Phe	Pro	Val	Lys	
1280					1285						1290				
Gly	Val	Thr	Val	Asn	Phe	Thr	Ser	Asn	Ala	Ala	Thr	Ala	Glu	Met	
1295					1300						1305				
Thr	Asn	Gly	Gly	Gln	Ala	Val	Thr	Asn	Glu	Gln	Gly	Lys	Ala	Thr	
1310					1315						1320				
Val	Thr	Tyr	Thr	Asn	Thr	Arg	Ser	Ser	Ile	Glu	Ser	Gly	Ala	Arg	
1325					1330						1335				
Pro	Asp	Thr	Val	Glu	Ala	Ser	Leu	Glu	Asn	Gly	Ser	Ser	Thr	Leu	
1340					1345						1350				
Ser	Thr	Ser	Ile	Asn	Val	Asn	Ala	Asp	Ala	Ser	Thr	Ala	His	Leu	
1355					1360						1365				
Thr	Leu	Leu	Gln	Ala	Leu	Phe	Asp	Thr	Val	Ser	Ala	Gly	Glu	Thr	
1370					1375						1380				
Thr	Ser	Leu	Tyr	Ile	Glu	Val	Lys	Asp	Asn	Tyr	Gly	Asn	Gly	Val	
1385					1390						1395				
[0033]	Pro	Gln	Gln	Glu	Val	Thr	Leu	Ser	Val	Ser	Pro	Ser	Glu	Gly	Val
	1400				1405						1410				
Thr	Pro	Ser	Asn	Asn	Ala	Ile	Tyr	Thr	Thr	Asn	His	Asp	Gly	Asn	
1415					1420						1425				
Phe	Tyr	Ala	Ser	Phe	Thr	Ala	Thr	Lys	Ala	Gly	Val	Tyr	Gln	Leu	
1430					1435						1440				
Thr	Ala	Thr	Leu	Glu	Asn	Gly	Asp	Ser	Met	Gln	Gln	Thr	Val	Thr	
1445					1450						1455				
Tyr	Val	Pro	Asn	Val	Ala	Asn	Ala	Glu	Ile	Thr	Leu	Ala	Ala	Ser	
1460					1465						1470				
Lys	Asp	Pro	Val	Ile	Ala	Asp	Asn	Asn	Asp	Leu	Thr	Thr	Leu	Thr	
1475					1480						1485				
Ala	Thr	Val	Ala	Asp	Thr	Glu	Gly	Asn	Ala	Ile	Ala	Asn	Thr	Glu	
1490					1495						1500				
Val	Thr	Phe	Thr	Leu	Pro	Glu	Asp	Val	Lys	Ala	Asn	Phe	Thr	Leu	
1505					1510						1515				
Ser	Asp	Gly	Gly	Lys	Val	Ile	Thr	Asp	Ala	Glu	Gly	Lys	Ala	Lys	
1520					1525						1530				
Val	Thr	Leu	Lys	Gly	Thr	Lys	Ala	Gly	Ala	His	Thr	Val	Thr	Ala	
1535					1540						1545				

Ser Met Thr Gly Gly Lys Ser Glu Gln Leu Val Val Asn Phe Ile
 1550 1555 1560

 Ala Asp Thr Leu Thr Ala Gln Val Asn Leu Asn Val Thr Glu Asp
 1565 1570 1575

 Asn Phe Ile Ala Asn Asn Val Gly Met Thr Arg Leu Gln Ala Thr
 1580 1585 1590

 Val Thr Asp Gly Asn Gly Asn Pro Leu Ala Asn Glu Ala Val Thr
 1595 1600 1605

 Phe Thr Leu Pro Ala Asp Val Ser Ala Ser Phe Thr Leu Gly Gln
 1610 1615 1620

 Gly Gly Ser Ala Ile Thr Asp Ile Asn Gly Lys Ala Glu Val Thr
 1625 1630 1635

 Leu Ser Gly Thr Lys Ser Gly Thr Tyr Pro Val Thr Val Ser Val
 1640 1645 1650

 Asn Asn Tyr Gly Val Ser Asp Thr Lys Gln Val Thr Leu Ile Ala
 1655 1660 1665

 Asp Ala Gly Thr Ala Lys Leu Ala Ser Leu Thr Ser Val Tyr Ser
 1670 1675 1680

 [0034] Phe Val Val Ser Thr Thr Glu Gly Ala Thr Met Thr Ala Ser Val
 1685 1690 1695

 Thr Asp Ala Asn Gly Asn Pro Val Glu Gly Ile Lys Val Asn Phe
 1700 1705 1710

 Arg Gly Thr Ser Val Thr Leu Ser Ser Thr Ser Val Glu Thr Asp
 1715 1720 1725

 Asp Arg Gly Phe Ala Glu Ile Leu Val Thr Ser Thr Glu Val Gly
 1730 1735 1740

 Leu Lys Thr Val Ser Ala Ser Leu Ala Asp Lys Pro Thr Glu Val
 1745 1750 1755

 Ile Ser Arg Leu Leu Asn Ala Ser Ala Asp Val Asn Ser Ala Thr
 1760 1765 1770

 Ile Thr Ser Leu Glu Ile Pro Glu Gly Gln Val Met Val Ala Gln
 1775 1780 1785

 Asp Val Ala Val Lys Ala His Val Asn Asp Gln Phe Gly Asn Pro
 1790 1795 1800

 Val Ala His Gln Pro Val Thr Phe Ser Ala Glu Pro Ser Ser Gln
 1805 1810 1815

 Met Ile Ile Ser Gln Asn Thr Val Ser Thr Asn Thr Gln Gly Val
 1820 1825 1830

Ala Glu Val Thr Met Thr Pro Glu Arg Asn Gly Ser Tyr Met Val
 1835 1840 1845

Lys Ala Ser Leu Pro Asn Gly Ala Ser Leu Glu Lys Gln Leu Glu
 1850 1855 1860

Ala Ile Asp Glu Lys Leu Thr Leu Thr Ala Ser Ser Pro Leu Ile
 1865 1870 1875

Gly Val Tyr Ala Pro Thr Gly Ala Thr Leu Thr Ala Thr Leu Thr
 1880 1885 1890

Ser Ala Asn Gly Thr Pro Val Glu Gly Gln Val Ile Asn Phe Ser
 1895 1900 1905

Val Thr Pro Glu Gly Ala Thr Leu Ser Gly Gly Lys Val Arg Thr
 1910 1915 1920

Asn Ser Ser Gly Gln Ala Pro Val Val Leu Thr Ser Asn Lys Val
 1925 1930 1935

Gly Thr Tyr Thr Val Thr Ala Ser Phe His Asn Gly Val Thr Ile
 1940 1945 1950

Gln Thr Gln Thr Thr Val Lys Val Thr Gly Asn Ser Ser Thr Ala
 1955 1960 1965

[0035] His Val Ala Ser Phe Ile Ala Asp Pro Ser Thr Ile Ala Ala Thr
 1970 1975 1980

Asn Thr Asp Leu Ser Thr Leu Lys Ala Thr Val Glu Asp Gly Ser
 1985 1990 1995

Gly Asn Leu Ile Glu Gly Leu Thr Val Tyr Phe Ala Leu Lys Ser
 2000 2005 2010

Gly Ser Ala Thr Leu Thr Ser Leu Thr Ala Val Thr Asp Gln Asn
 2015 2020 2025

Gly Ile Ala Thr Thr Ser Val Lys Gly Ala Met Thr Gly Ser Val
 2030 2035 2040

Thr Val Ser Ala Val Thr Thr Ala Gly Gly Met Gln Thr Val Asp
 2045 2050 2055

Ile Thr Leu Val Ala Gly Pro Ala Asp Thr Ser Gln Ser Val Leu
 2060 2065 2070

Lys Ser Asn Arg Ser Ser Leu Lys Gly Asp Tyr Thr Asp Ser Ala
 2075 2080 2085

Glu Leu Arg Leu Val Leu His Asp Ile Ser Gly Asn Pro Ile Lys
 2090 2095 2100

Val Ser Glu Gly Met Glu Phe Val Gln Ser Gly Thr Asn Val Pro
 2105 2110 2115

Tyr Ile Lys Ile Ser Ala Ile Asp Tyr Ser Leu Asn Ile Asn Gly
2120 2125 2130 2135

Asp Tyr Lys Ala Thr Val Thr Gly Gly Gly Glu Gly Ile Ala Thr
2135 2140 2145

Leu Ile Pro Val Leu Asn Gly Val His Gln Ala Gly Leu Ser Thr
2150 2155 2160

Thr Ile Gln Phe Thr Arg Ala Glu Asp Lys Ile Met Ser Gly Thr
2165 2170 2175

Val Ser Val Asn Gly Thr Asp Leu Pro Thr Thr Thr Phe Pro Ser
2180 2185 2190

Gln Gly Phe Thr Gly Ala Tyr Tyr Gln Leu Asn Asn Asp Asn Phe
2195 2200 2205

Ala Pro Gly Lys Thr Ala Ala Asp Tyr Glu Phe Ser Ser Ser Ala
2210 2215 2220

Ser Trp Val Asp Val Asp Ala Thr Gly Lys Val Thr Phe Lys Asn
2225 2230 2235

Val Gly Ser Asn Ser Glu Arg Ile Thr Ala Thr Pro Lys Ser Gly
2240 2245 2250

[0036] Gly Pro Ser Tyr Val Tyr Glu Ile Arg Val Lys Ser Trp Trp Val
2255 2260 2265

Asn Ala Gly Glu Ala Phe Met Ile Tyr Ser Leu Ala Glu Asn Phe
2270 2275 2280

Cys Ser Ser Asn Gly Tyr Thr Leu Pro Arg Ala Asn Tyr Leu Asn
2285 2290 2295

His Cys Ser Ser Arg Gly Ile Gly Ser Leu Tyr Ser Glu Trp Gly
2300 2305 2310

Asp Met Gly His Tyr Thr Thr Asp Ala Gly Phe Gln Ser Asn Met
2315 2320 2325

Tyr Trp Ser Ser Ser Pro Ala Asn Ser Ser Glu Gln Tyr Val Val
2330 2335 2340

Ser Leu Ala Thr Gly Asp Gln Ser Val Phe Glu Lys Leu Gly Phe
2345 2350 2355

Ala Tyr Ala Thr Cys Tyr Lys Asn Leu
2360 2365

<210> 16

<211> 1093

<212> PRT

<213> 大肠杆菌

<400> 16

Met Thr Pro Val Lys Val Trp Gln Glu Arg Val Glu Ile Pro Thr Tyr

1	5	10	15														
Glu	Thr	Gly	Pro														
20		25	Gln														
			Asp														
			Ile														
			His														
			Pro														
			Met														
			Phe														
			Leu														
			Glu														
			Asn														
			Arg														
			Val														
Tyr	Gln	Gly	Ser	Ser	Gly	Ala	Val	Tyr	Pro	Tyr	Gly	Val	Thr	Asp	Thr		
35																	
Leu	Ser	Glu	Gln	Lys	Thr	Leu	Lys	Ser	Trp	Gln	Ala	Val	Trp	Leu	Glu		
50																	
Asn	Asp	Tyr	Ile	Lys	Val	Met	Ile	Leu	Pro	Glu	Leu	Gly	Gly	Arg	Val		
65																	
His	Arg	Ala	Trp	Asp	Lys	Val	Lys	Gln	Arg	Asp	Phe	Val	Tyr	His	Asn		
85																	
Glu	Val	Ile	Lys	Pro	Ala	Leu	Val	Gly	Leu	Leu	Gly	Pro	Trp	Ile	Ser		
100																	
Gly	Gly	Ile	Glu	Phe	Asn	Trp	Pro	Gln	His	His	Arg	Pro	Thr	Thr	Phe		
115																	
Met	Pro	Val	Asp	Phe	Thr	Leu	Glu	Ala	His	Glu	Asp	Gly	Ala	Gln	Thr		
130																	
Val	Trp	Val	Gly	Glu	Thr	Glu	Pro	Met	His	Gly	Leu	Gln	Val	Met	Thr		
145																	
[0037]	Gly	Phe	Thr	Leu	Arg	Pro	Asp	Arg	Ala	Ala	Leu	Glu	Ile	Ala	Ser	Arg	
	165																
Val	Tyr	Asn	Gly	Asn	Ala	Thr	Pro	Arg	His	Phe	Leu	Trp	Trp	Ala	Asn		
180																	
Pro	Ala	Val	Lys	Gly	Gly	Glu	Gly	His	Gln	Ser	Val	Phe	Pro	Pro	Asp		
195																	
Val	Thr	Ala	Val	Phe	Asp	His	Gly	Lys	Arg	Ala	Val	Ser	Ala	Phe	Pro		
210																	
Ile	Ala	Thr	Gly	Thr	Tyr	Tyr	Lys	Val	Asp	Tyr	Ser	Ala	Gly	Val	Asp		
225																	
Ile	Ser	Arg	Tyr	Lys	Asn	Val	Pro	Val	Pro	Thr	Ser	Tyr	Met	Ala	Glu		
245																	
Lys	Ser	Gln	Tyr	Asp	Phe	Val	Gly	Ala	Trp	Cys	His	Asp	Glu	Asp	Gly		
260																	
Gly	Leu	Leu	His	Val	Ala	Asn	His	His	Ile	Ala	Pro	Gly	Lys	Lys	Gln		
275																	
Trp	Ser	Trp	Gly	His	Ser	Glu	Phe	Gly	Gln	Ala	Trp	Asp	Lys	Ser	Leu		
290																	
Thr	Asp	Asn	Asn	Gly	Pro	Tyr	Ile	Glu	Leu	Met	Thr	Gly	Ile	Phe	Ala		

305	310	315	320
-----	-----	-----	-----

Asp Asn Gln Pro Asp Phe Thr Trp Leu Asp Ala Tyr Glu Glu Lys Arg	325	330	335
---	-----	-----	-----

Phe Glu Gln Tyr Phe Leu Pro Tyr His Ser Leu Gly Met Val Gln Asn	340	345	350
---	-----	-----	-----

Ala Ser Arg Asp Ala Val Ile Lys Leu Gln Arg Ser Lys Arg Gly Ile	355	360	365
---	-----	-----	-----

Glu Trp Gly Leu Tyr Ala Ile Ser Pro Leu Asn Gly Tyr Arg Leu Ala	370	375	380
---	-----	-----	-----

Ile Arg Glu Ile Gly Lys Cys Asn Ala Leu Leu Asp Asp Ala Val Ala	385	390	395	400
---	-----	-----	-----	-----

Leu Met Pro Ala Thr Ala Ile Gln Gly Val Leu His Gly Ile Asn Pro	405	410	415
---	-----	-----	-----

Glu Arg Leu Thr Ile Glu Leu Ser Asp Ala Asp Gly Asn Ile Val Leu	420	425	430
---	-----	-----	-----

Ser Tyr Gln Glu His Gln Pro Gln Ala Leu Pro Leu Pro Asp Val Ala	435	440	445
---	-----	-----	-----

Lys Ala Pro Leu Ala Ala Gln Asp Ile Thr Ser Thr Asp Glu Ala Trp	450	455	460
---	-----	-----	-----

[0038]

Phe Ile Gly Gln His Leu Glu Gln Tyr His His Ala Ser Arg Ser Pro	465	470	475	480
---	-----	-----	-----	-----

Phe Asp Tyr Tyr Leu Arg Gly Val Ala Leu Asp Pro Leu Asp Tyr Arg	485	490	495
---	-----	-----	-----

Cys Asn Leu Ala Leu Ala Met Leu Glu Tyr Asn Arg Ala Asp Phe Pro	500	505	510
---	-----	-----	-----

Gln Ala Val Ala Tyr Ala Thr Gln Ala Leu Lys Arg Ala His Ala Leu	515	520	525
---	-----	-----	-----

Asn Lys Asn Pro Gln Cys Gly Gln Ala Ser Leu Ile Arg Ala Ser Ala	530	535	540
---	-----	-----	-----

Tyr Glu Arg Gln Gly Gln Tyr Gln Gln Ala Glu Glu Asp Phe Trp Arg	545	550	555	560
---	-----	-----	-----	-----

Ala Val Trp Ser Gly Asn Ser Lys Ala Gly Gly Tyr Tyr Gly Leu Ala	565	570	575
---	-----	-----	-----

Arg Leu Ala Ala Arg Asn Gly Asn Phe Asp Ala Gly Leu Asp Phe Cys	580	585	590
---	-----	-----	-----

Gin Gln Ser Leu Arg Ala Cys Pro Thr Asn Gln Glu Val Leu Cys Leu	595	600	605
---	-----	-----	-----

His Asn Leu Leu Leu Val Leu Ser Gly Arg Gln Asp Asn Ala Arg Val			
---	--	--	--

610

615

620

Gln Arg Glu Lys Leu Leu Arg Asp Tyr Pro Leu Asn Ala Thr Leu Trp
 625 630 635 640

Trp Leu Asn Trp Phe Asp Gly Arg Ser Glu Ser Ala Leu Ala Gln Trp
 645 650 655

Arg Gly Leu Cys Gln Gly Arg Asp Val Asn Ala Leu Met Thr Ala Gly
 660 665 670

Gln Leu Ile Asn Trp Gly Met Pro Thr Leu Ala Ala Glu Met Leu Asn
 675 680 685

Ala Leu Asp Cys Gln Arg Thr Leu Pro Leu Tyr Leu Gln Ala Ser Leu
 690 695 700

Leu Pro Lys Ala Glu Arg Gly Glu Leu Val Ala Lys Ala Ile Asp Val
 705 710 715 720

Phe Pro Gln Phe Val Arg Phe Pro Asn Thr Leu Glu Glu Val Ala Ala
 725 730 735

Leu Glu Ser Ile Glu Glu Cys Trp Phe Ala Arg His Leu Leu Ala Cys
 740 745 750

Phe Tyr Tyr Asn Lys Arg Ser Tyr Asn Lys Ala Ile Ala Phe Trp Gln
 755 760 765

[0039]

Arg Cys Val Glu Met Ser Pro Glu Phe Ala Asp Gly Trp Arg Gly Leu
 770 775 780

Ala Ile His Ala Trp Asn Lys Gln His Asp Tyr Glu Leu Ala Ala Arg
 785 790 795 800

Tyr Leu Asp Asn Ala Tyr Gln Leu Ala Pro Gln Asp Ala Arg Leu Leu
 805 810 815

Phe Glu Arg Asp Leu Leu Asp Lys Leu Ser Gly Ala Thr Pro Glu Lys
 820 825 830

Arg Leu Ala Arg Leu Glu Asn Asn Leu Glu Ile Ala Leu Lys Arg Asp
 835 840 845

Asp Met Thr Ala Glu Leu Leu Asn Leu Trp His Leu Thr Gly Gln Ala
 850 855 860

Asp Lys Ala Ala Asp Ile Leu Ala Thr Arg Lys Phe His Pro Trp Glu
 865 870 875 880

Gly Gly Glu Gly Lys Val Thr Ser Gln Phe Ile Leu Asn Gln Leu Leu
 885 890 895

Arg Ala Trp Gln His Leu Asp Ala Arg Gln Pro Gln Gln Ala Cys Glu
 900 905 910

Leu Leu His Ala Ala Leu His Tyr Pro Glu Asn Leu Ser Glu Gly Arg

915	920	925
-----	-----	-----

Leu Pro Gly Gln Thr Asp Asn Asp Ile Trp Phe Trp Gln Ala Ile Cys
930 935 940

Ala Asn Ala Gln Gly Asp Glu Thr Glu Ala Thr Arg Cys Leu Arg Leu
945 950 955 960

Ala Ala Thr Gly Asp Arg Thr Ile Asn Ile His Ser Tyr Tyr Asn Asp
965 970 975

Gln Pro Val Asp Tyr Leu Phe Trp Gln Gly Met Ala Leu Arg Leu Leu
980 985 990

Gly Glu Gln Gln Thr Ala Gln Gln Leu Phe Ser Glu Met Lys Gln Trp
995 1000 1005

Ala Gln Glu Met Ala Lys Thr Ser Ile Glu Ala Asp Phe Phe Ala
1010 1015 1020

Val Ser Gln Pro Asp Leu Leu Ser Leu Tyr Gly Asp Leu Gln Gln
1025 1030 1035

Gln His Lys Glu Lys Cys Leu Met Val Ala Met Leu Ala Ser Ala
1040 1045 1050

Gly Leu Gly Glu Val Ala Gln Tyr Glu Ser Ala Arg Ala Glu Leu
1055 1060 1065

[0040]

Thr Ala Ile Asn Pro Ala Trp Pro Lys Ala Ala Leu Phe Thr Thr
1070 1075 1080

Val Met Pro Phe Ile Phe Asn Arg Val His
1085 1090

<210> 17

<211> 365

<212> PRT

<213> 大肠杆菌

<400> 17

Met Phe Glu Ile Asn Pro Val Asn Asn Arg Ile Gln Asp Leu Thr Glu
1 5 10 15

Arg Ser Asp Val Leu Arg Gly Tyr Leu Asp Tyr Asp Ala Lys Lys Glu
20 25 30

Arg Leu Glu Glu Val Asn Ala Glu Leu Glu Gln Pro Asp Val Trp Asn
35 40 45

Glu Pro Glu Arg Ala Gln Ala Leu Gly Lys Glu Arg Ser Ser Leu Glu
50 55 60

Ala Val Val Asp Thr Leu Asp Gln Met Lys Gln Gly Leu Glu Asp Val
65 70 75 80

Ser Gly Leu Leu Glu Leu Ala Val Glu Ala Asp Asp Glu Glu Thr Phe
85 90 95

序 列 表

CN 109844105 A

41/52 页

Asn Glu Ala Val Ala Glu Leu Asp Ala Leu Glu Glu Lys Leu Ala Gln
 100 105 110

Leu Glu Phe Arg Arg Met Phe Ser Gly Glu Tyr Asp Ser Ala Asp Cys
 115 120 125

Tyr Leu Asp Ile Gln Ala Gly Ser Gly Gly Thr Glu Ala Gln Asp Trp
 130 135 140

Ala Ser Met Leu Glu Arg Met Tyr Leu Arg Trp Ala Glu Ser Arg Gly
 145 150 155 160

Phe Lys Thr Glu Ile Ile Glu Glu Ser Glu Gly Glu Val Ala Gly Ile
 165 170 175

Lys Ser Val Thr Ile Lys Ile Ser Gly Asp Tyr Ala Tyr Gly Trp Leu
 180 185 190

Arg Thr Glu Thr Gly Val His Arg Leu Val Arg Lys Ser Pro Phe Asp
 195 200 205

Ser Gly Gly Arg Arg His Thr Ser Phe Ser Ser Ala Phe Val Tyr Pro
 210 215 220

Glu Val Asp Asp Asp Ile Asp Ile Glu Ile Asn Pro Ala Asp Leu Arg
 225 230 235 240

[0041] Ile Asp Val Tyr Arg Thr Ser Gly Ala Gly Gly Gln His Val Asn Arg
 245 250 255

Thr Glu Ser Ala Val Arg Ile Thr His Ile Pro Thr Gly Ile Val Thr
 260 265 270

Gln Cys Gln Asn Asp Arg Ser Gln His Lys Asn Lys Asp Gln Ala Met
 275 280 285

Lys Gln Met Lys Ala Lys Leu Tyr Glu Leu Glu Met Gln Lys Lys Asn
 290 295 300

Ala Glu Lys Gln Ala Met Glu Asp Asn Lys Ser Asp Ile Gly Trp Gly
 305 310 315 320

Ser Gln Ile Arg Ser Tyr Val Leu Asp Asp Ser Arg Ile Lys Asp Leu
 325 330 335

Arg Thr Gly Val Glu Thr Arg Asn Thr Gln Ala Val Leu Asp Gly Ser
 340 345 350

Leu Asp Gln Phe Ile Glu Ala Ser Leu Lys Ala Gly Leu
 355 360 365

<210> 18
<211> 121
<212> PRT
<213> 大肠杆菌

<400> 18

Met Ile Ser Arg Val Thr Glu Ala Leu Ser Lys Val Lys Gly Ser Met
 1 5 10 15

Gly Ser His Glu Arg His Ala Leu Pro Gly Val Ile Gly Asp Asp Leu
 20 25 30

Leu Arg Phe Gly Lys Leu Pro Leu Cys Leu Phe Ile Cys Ile Ile Leu
 35 40 45

Thr Ala Val Thr Val Val Thr Thr Ala His His Thr Arg Leu Leu Thr
 50 55 60

Ala Gln Arg Glu Gln Leu Val Leu Glu Arg Asp Ala Leu Asp Ile Glu
 65 70 75 80

Trp Arg Asn Leu Ile Leu Glu Glu Asn Ala Leu Gly Asp His Ser Arg
 85 90 95

Val Glu Arg Ile Ala Thr Glu Lys Leu Gln Met Gln His Val Asp Pro
 100 105 110

Ser Gln Glu Asn Ile Val Val Gln Lys
 115 120

<210> 19

<211> 305

<212> PRT

<213> 大肠杆菌

[0042]

<400> 19

Met Asn Ile Arg Asp Leu Glu Tyr Leu Val Ala Leu Ala Glu His Arg
 1 5 10 15

His Phe Arg Arg Ala Ala Asp Ser Cys His Val Ser Gln Pro Thr Leu
 20 25 30

Ser Gly Gln Ile Arg Lys Leu Glu Asp Glu Leu Gly Val Met Leu Leu
 35 40 45

Glu Arg Thr Ser Arg Lys Val Leu Phe Thr Gln Ala Gly Met Leu Leu
 50 55 60

Val Asp Gln Ala Arg Thr Val Leu Arg Glu Val Lys Val Leu Lys Glu
 65 70 75 80

Met Ala Ser Gln Gln Gly Glu Thr Met Ser Gly Pro Leu His Ile Gly
 85 90 95

Leu Ile Pro Thr Val Gly Pro Tyr Leu Leu Pro His Ile Ile Pro Met
 100 105 110

Leu His Gln Thr Phe Pro Lys Leu Glu Met Tyr Leu His Glu Ala Gln
 115 120 125

Thr His Gln Leu Leu Ala Gln Leu Asp Ser Gly Lys Leu Asp Cys Val
 130 135 140

Ile Leu Ala Leu Val Lys Glu Ser Glu Ala Phe Ile Glu Val Pro Leu
145 150 155 160

Phe Asp Glu Pro Met Leu Leu Ala Ile Tyr Glu Asp His Pro Trp Ala
165 170 175

Asn Arg Glu Cys Val Pro Met Ala Asp Leu Ala Gly Glu Lys Leu Leu
180 185 190

Met Leu Glu Asp Gly His Cys Leu Arg Asp Gln Ala Met Gly Phe Cys
195 200 205

Phe Glu Ala Gly Ala Asp Glu Asp Thr His Phe Arg Ala Thr Ser Leu
210 215 220

Glu Thr Leu Arg Asn Met Val Ala Ala Gly Ser Gly Ile Thr Leu Leu
225 230 235 240

Pro Ala Leu Ala Val Pro Pro Glu Arg Lys Arg Asp Gly Val Val Tyr
245 250 255

Leu Pro Cys Ile Lys Pro Glu Pro Arg Arg Thr Ile Gly Leu Val Tyr
260 265 270

Arg Pro Gly Ser Pro Leu Arg Ser Arg Tyr Glu Gln Leu Ala Glu Ala
275 280 285

Ile Arg Ala Arg Met Asp Gly His Phe Asp Lys Val Leu Lys Gln Ala
290 295 300

[0043]

Val
305

<210> 20
<211> 95
<212> DNA
<213> 人工序列

<220>
<223> 合成寡核苷酸

<400> 20
ggaagatgt cgtctccgtt gaggcggtg gactctaaat ccagttgggg ccgccacgg 60
tcccggttag gttcgactcc ttgcataatc cgcca 95

<210> 21
<211> 95
<212> DNA
<213> 人工序列

<220>
<223> 合成寡核苷酸

<400> 21
ggaagatgt cgtctccgtt gaggcggtg gactctaaat ccagttgggg ccgccacgg 60
tcccggttag gttcgactcc ttgcataatc cgcca 95

<210> 22
<211> 94
<212> DNA
<213> 人工序列

<220>
 <223> 合成寡核苷酸
 <400> 22
 ggaagatcg t cgtctccggt gaggcggctg gactctaaat ccagttgggg ccgccagcgg 60
 tcccgtaag gttcgactcc tatatcttcc gcca 94

<210> 23
 <211> 90
 <212> DNA
 <213> 人工序列

<220>
 <223> 合成寡核苷酸
 <400> 23
 ggaagatgt gccgagcggt tgaaggcacc ggtctctaaa accggcgacc cgaaagggtt 60
 ccagagttcg aatctctgca tcttccgc 90

<210> 24
 <211> 292
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成多肽
 <400> 24

Met	Ser	Leu	Asn	Val	Lys	Gln	Ser	Arg	Ile	Ala	Ile	Leu	Phe	Ser	Ser
1				5				10						15	

[0044] Cys Leu Ile Ser Ile Ser Phe Phe Ser Gln Ala Asn Thr Lys Gly Ile
 20 25 30

Asp Glu Ile Lys Asn Leu Glu Thr Asp Phe Asn Gly Arg Ile Gly Val
 35 40 45

Tyr Ala Leu Asp Thr Gly Ser Gly Lys Ser Phe Ser Tyr Arg Ala Asn
 50 55 60

Glu Arg Phe Pro Leu Cys Ser Ser Phe Lys Gly Phe Leu Ala Ala Ala
 65 70 75 80

Val Leu Lys Gly Ser Gln Asp Asn Arg Leu Asn Leu Asn Gln Ile Val
 85 90 95

Asn Tyr Asn Thr Arg Ser Leu Glu Phe His Ser Pro Ile Thr Thr Lys
 100 105 110

Tyr Lys Asp Asn Gly Met Ser Leu Gly Asp Met Ala Ala Ala Leu
 115 120 125

Gln Tyr Ser Asp Asn Gly Ala Thr Asn Ile Ile Leu Glu Arg Tyr Ile
 130 135 140

Gly Gly Pro Glu Gly Met Thr Lys Phe Met Arg Ser Ile Gly Asp Glu
 145 150 155 160

Asp Phe Arg Leu Asp Arg Trp Glu Leu Asp Leu Asn Thr Ala Ile Pro
 165 170 175

Gly Asp Glu Arg Asp Thr Ser Thr Pro Ala Ala Val Ala Lys Ser Leu
180 185 190

Lys Thr Leu Ala Leu Gly Asn Ile Leu Ser Glu His Glu Lys Glu Thr
195 200 205

Tyr Gln Thr Trp Leu Lys Gly Asn Thr Thr Gly Ala Ala Arg Ile Arg
210 215 220

Ala Ser Val Pro Ser Asp Trp Val Val Gly Asp Lys Thr Gly Ser Cys
225 230 235 240

Gly Ala Tyr Gly Thr Ala Asn Asp Tyr Ala Val Val Trp Pro Lys Asn
245 250 255

Arg Ala Pro Leu Ile Ile Ser Val Tyr Thr Thr Lys Asn Glu Lys Glu
260 265 270

Ala Lys His Glu Asp Lys Val Ile Ala Glu Ala Ser Arg Ile Ala Ile
275 280 285

Asp Asn Leu Lys
290

<210> 25
<211> 879
<212> DNA
<213> 人工序列

[0045]

<220>
<223> 合成寡核苷酸

<400> 25		
atgtcacita atgtaaagca aagttagataa gccatcttgt ttagcttttg tttaatttca	60	
atatcatttt tctcacaggc caatacgaag ggcattgatg agattaaaaaa ctttggaaaca	120	
gatttcaatg gcaggattgg tgtctacgct ttagacactg gctcggttaa atcatttcg	180	
tacagagcaa atgaacgatt tccattatag agttttttta aaggtttttt agctgtgct	240	
gtattaaaag gctctcaaga taatcgactt aatcttaatc agattgtgaa ttataataca	300	
agaagtttag agttccatttc acccatcaca actaaatata aagataatgg aatgtcatta	360	
ggtgatatgg ctgtgtgtgc tttacaatat agcgacaatg gtgtactaa tattatttt	420	
gaacgttata tcgggtgtcc agagggtatg actaaatcca tgccgtcgat tggagatgaa	480	
gattttagac tcgatcggtt ggagtttagat ctaaacacag ctattccagg cgatgagcgt	510	
gacacatcta cacatgcgc agtagccaag agtctgaaaa cccttgcctt gggtaacata	600	
ctttagtgaac atgaaaagga aacctatcg acatggtaa aggtaaacac aaccgggtca	660	
gcgtgttacc gtgttagcgt accaagcgat tggtagtttgcgtataaaac tggtagtt	720	
ggaggcatacg gtacggcaaa tgattatgc gtagtctggc caaagaaccc ggctcttctt	780	
ataatttctg tatacacaac aaaaaacgaa aaagaagcca agcatgagga taaagtaatc	840	
gcagaagctt caagaattgc aattgataac cttaaataaa	879	

<210> 26
<211> 167

<212> PRT
 <213> 人工序列

<220>
 <223> 合成多肽

<220>
 <221> MISC_特征
 <222> (39)..(39)
 <223> Xaa = 硒代半胱氨酸

<220>
 <221> MISC_特征
 <222> (85)..(85)
 <223> Xaa = 硒代半胱氨酸

<400> 26

Met Ile Ser Leu Ile Ala Ala Leu Ala Val Asp Arg Val Ile Gly Asn
 1 5 10 15

Glu Met Ala Met Pro Trp Asn Leu Pro Ala Asp Leu Ala Trp Phe Lys
 20 25 30

Arg Asn Thr Leu Asn Lys Xaa Val Ile Met Gly Arg His Thr Trp Glu
 35 40 45

Ser Ile Gly Arg Pro Leu Pro Gly Arg Lys Asn Ile Ile Leu Ser Ser
 50 55 60

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

[0046]

Ala Ile Ala Ala Xaa Gly Asp Val Pro Glu Ile Met Val Ile Gly Gly
 85 90 95

Gly Arg Val Tyr Glu Gln Phe Leu Pro Lys Ala Gln Lys Leu Tyr Leu
 100 105 110

Thr His Ile Asp Ala Glu Val Glu Gly Asp Thr His Phe Pro Asp Tyr
 115 120 125

Glu Pro Asp Asp Trp Glu Ser Val Phe Ser Glu Phe His Asp Ala Asp
 130 135 140

Ala Gln Asn Ser His Ser Tyr Cys Phe Glu Ile Leu Glu Arg Arg Gly
 145 150 155 160

Ser His His His His His
 165

<210> 27
 <211> 294
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成多肽

<220>
 <221> MISC_特征
 <222> (42)..(42)
 <223> Xaa = 硒代半胱氨酸

<220>
 <221> MISC 特征
 <222> (116)..(116)
 <223> Xaa = 硒代半胱氨酸

<220>
 <221> MISC 特征
 <222> (179)..(179)
 <223> Xaa = 硒代半胱氨酸

<220>
 <221> MISC 特征
 <222> (249)..(249)
 <223> Xaa = 硒代半胱氨酸

<400> 27

Met	His	His	His	His	His	Leu	Val	Pro	Arg	Gly	Ser	Gly	Lys	Ser
1					5			10					15	

Ala Met Ala Ala Glu Val Lys Leu Val Glu Ser Gly Gly Gly Leu Val
 20 25 30

Lys Pro Gly Gly Ser Leu Lys Leu Ser Xaa Ala Ala Ser Gly Phe Thr
 35 40 45

Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln Thr Pro Glu Lys Arg
 50 55 60

Leu Glu Trp Val Ala Thr Ile Ser Thr Gly Gly Tyr Thr Tyr Phe
 65 70 75 80

[0047] Pro Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys
 85 90 95

Asn Ala Leu Tyr Leu Gln Met Lys Ser Leu Arg Ser Glu Asp Thr Ala
 100 105 110

Met Tyr Tyr Xaa Ala Arg Gln Gly Asp Phe Gly Asp Trp Tyr Phe Asp
 115 120 125

Val Trp Gly Ala Gly Thr Thr Val Thr Val Ser Ser Gly Gly Gly
 130 135 140

Ser Gly Gly Gly Ser Gly Gly Ser Thr Asp Val Leu Met
 145 150 155 160

Thr Gln Thr Pro Leu Ser Leu Pro Val Ser Leu Gly Asp Gln Ala Ser
 165 170 175

Ile Ser Xaa Arg Ser Ser Gln Ser Leu Val His Ser Asn Gly Asn Thr
 180 185 190

Tyr Leu His Trp Tyr Leu Gln Lys Pro Gly Gln Ser Pro Lys Leu Leu
 195 200 205

Ile Tyr Lys Val Ser Asn Arg Phe Ser Gly Val Pro Asp Arg Phe Ser
 210 215 220

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile Ser Arg Val Glu
 225 230 235 240

Ala Glu Asp Leu Gly Val Tyr Phe Xaa Ser Gln Ser Thr His Val Pro
245 250 255

Trp Thr Phe Gly Gly Thr Lys Leu Glu Ile Lys Arg Ala Ala Ala
260 265 270

Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr
275 280 285

Lys Asp Asp Asp Asp Lys
290

<210> 28

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 合成多肽

<400> 28

Met Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Ala Lys Pro Gly
1 5 10 15

Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Thr Asp
20 25 30

Tyr Tyr Met Asp Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
35 40 45

[0048]

Val Ser Arg Ile Ser Pro Gly Gly Asp Val Thr Trp Tyr Ala Asp Ser
50 55 60

Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Gln Asn Thr Leu
65 70 75 80

Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe
85 90 95

Cys Ala Arg Asp Asp Ile Val Val Ser Arg Ile Phe Asp Asp Trp Gly
100 105 110

Gln Gly Val Leu Leu Thr Val Ser Ser Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Gly Gly Ser Glu Leu Gln Met Thr Gln Ser Pro
130 135 140

Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg
145 150 155 160

Ala Ser Gln Ser Ile Arg Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro
165 170 175

Gly Lys Ala Pro Lys Leu Leu Ile Tyr Asp Ala Ala His Leu Gln Ser
180 185 190

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe Ser
195 200 205

Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Val Tyr Tyr Cys
210 215 220

Gln Gln Arg Asn Ser Tyr Pro Leu Thr Phe Gly Gly Gly Thr Lys Val
225 230 235 240

Glu Ile Lys Arg Gly Ser His His His His His His
245 250

<210> 29

<211> 252

<212> PRT

<213> 人工序列

<220>

<223> 合成多肽

<220>

<221> MISC_特征

<222> (23)..(23)

<223> Xaa = 硒代半胱氨酸

<220>

<221> MISC_特征

<222> (97)..(97)

<223> Xaa = 硒代半胱氨酸

<220>

<221> MISC_特征

<222> (159)..(159)

[0049] <223> Xaa = 硒代半胱氨酸

<220>

<221> MISC_特征

<222> (224)..(224)

<223> Xaa = 硒代半胱氨酸

<400> 29

Met Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Ala Lys Pro Gly
1 5 10 15

Gly Ser Leu Arg Leu Ser Xaa Ala Ala Ser Gly Phe Thr Phe Thr Asp
20 25 30

Tyr Tyr Met Asp Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
35 40 45

Val Ser Arg Ile Ser Pro Gly Gly Asp Val Thr Trp Tyr Ala Asp Ser
50 55 60

Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Gln Asn Thr Leu
65 70 75 80

Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Phe
85 90 95

Xaa Ala Arg Asp Asp Ile Val Val Ser Arg Ile Phe Asp Asp Trp Gly
100 105 110

Gln Gly Val Leu Leu Thr Val Ser Ser Gly Gly Gly Ser Gly Gly

115	120	125
-----	-----	-----

Gly Gly Ser Gly Gly Gly Ser Glu Leu Gln Met Thr Gln Ser Pro
 130 135 140

Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr Xaa Arg
 145 150 155 160

Ala Ser Gln Ser Ile Arg Ser Tyr Leu Ala Trp Tyr Gln Gln Lys Pro
 165 170 175

Gly Lys Ala Pro Lys Leu Leu Ile Tyr Asp Ala Ala His Leu Gln Ser
 180 185 190

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Glu Phe Ser
 195 200 205

Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Val Tyr Tyr Xaa
 210 215 220

Gln Gln Arg Asn Ser Tyr Pro Leu Thr Phe Gly Gly Gly Thr Lys Val
 225 230 235 240

Glu Ile Lys Arg Gly Ser His His His His His His
 245 250

[0050] <210> 30
 <211> 255
 <212> PRT
 <213> 人工序列

<220>
 <223> 合成多肽
 <400> 30

Met Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val
 1 5 10 15

Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Val Asn Thr
 20 25 30

Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu
 35 40 45

Ile Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser
 50 55 60

Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu Tyr Ile Ser Ser Leu Gln
 65 70 75 80

Pro Glu Asp Phe Ala Thr Tyr Cys Gln Gln His Tyr Thr Thr Pro
 85 90 95

Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Gly Ala Ser Gly
 100 105 110

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Ser Glu
 115 120 125

Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly Ser
 130 135 140

Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Asn Ile Lys Asp Thr Tyr
 145 150 155 160

Ile His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ala
 165 170 175

Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr Ala Asp Ser Val Lys
 180 185 190

Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr Leu
 195 200 205

Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ser
 210 215 220

Arg Trp Gly Gly Asp Gly Thr Tyr Ala Met Asp Tyr Trp Gly Gln Gly
 225 230 235 240

Thr Leu Val Thr Val Ser Ser Gly Ser His His His His His His
 245 250 255

<210> 31

<211> 255

<212> PRT

<213> 人工序列

[0051]

<220>

<223> 合成多肽

<220>

<221> MISC_特征

<222> (24)..(24)

<223> Xaa = 硒代半胱氨酸

<220>

<221> MISC_特征

<222> (89)..(89)

<223> Xaa = 硒代半胱氨酸

<220>

<221> MISC_特征

<222> (149)..(149)

<223> Xaa = 硒代半胱氨酸

<220>

<221> MISC_特征

<222> (223)..(223)

<223> Xaa = 硒代半胱氨酸

<400> 31

Met Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val
 1 5 10 15

Gly Asp Arg Val Thr Ile Thr Xaa Arg Ala Ser Gln Asp Val Asn Thr
 20 25 30

Ala Val Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu
 35 40 45

Ile Tyr Ser Ala Ser Phe Leu Tyr Ser Gly Val Pro Ser Arg Phe Ser
 50 55 60

Gly Ser Arg Ser Gly Thr Asp Phe Thr Leu Tyr Ile Ser Ser Leu Gln
 65 70 75 80

Pro Glu Asp Phe Ala Thr Tyr Tyr Xaa Gln Gln His Tyr Thr Thr Pro
 85 90 95

Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Gly Ala Ser Gly
 100 105 110

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Ser Glu
 115 120 125

Val Gln Leu Val Glu Ser Gly Gly Leu Val Gln Pro Gly Gly Ser
 130 135 140

[0052] Leu Arg Leu Ser Xaa Ala Ala Ser Gly Phe Asn Ile Lys Asp Thr Tyr
 145 150 155 160

Ile His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ala
 165 170 175

Arg Ile Tyr Pro Thr Asn Gly Tyr Thr Arg Tyr Ala Asp Ser Val Lys
 180 185 190

Gly Arg Phe Thr Ile Ser Ala Asp Thr Ser Lys Asn Thr Ala Tyr Leu
 195 200 205

Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Xaa Ser
 210 215 220

Arg Trp Gly Gly Asp Gly Thr Tyr Ala Met Asp Tyr Trp Gly Gln Gly
 225 230 235 240

Thr Leu Val Thr Val Ser Ser Gly Ser His His His His His His
 245 250 255

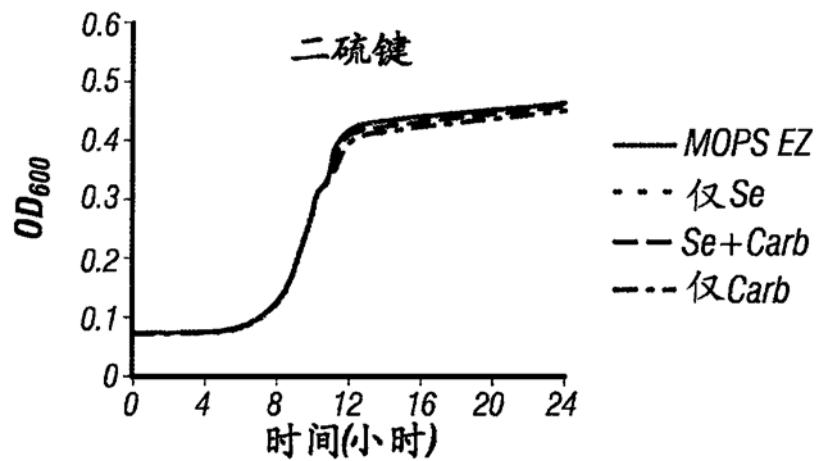


图1A

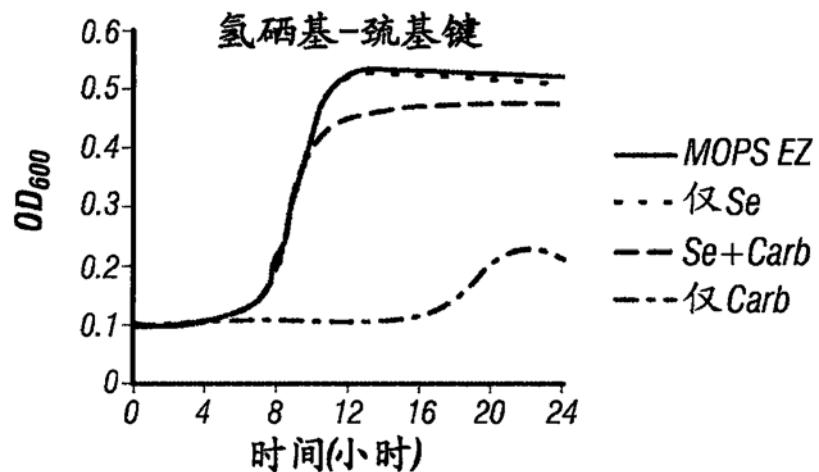


图1B

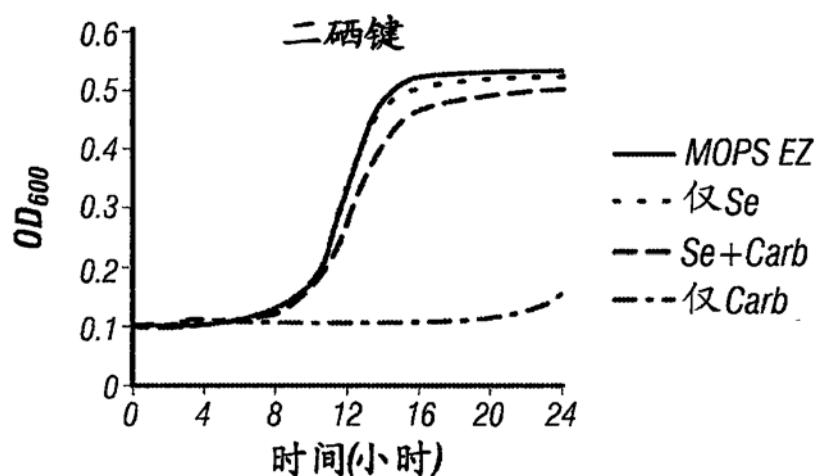


图1C

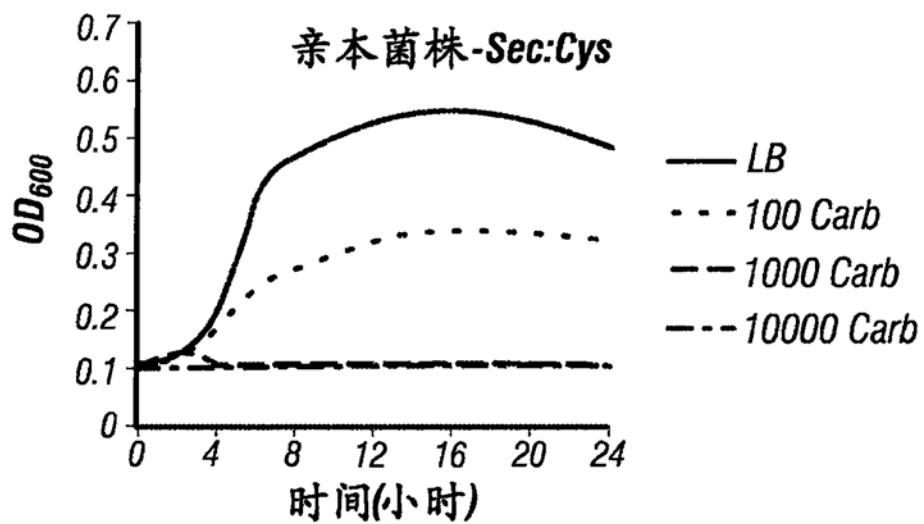


图2A

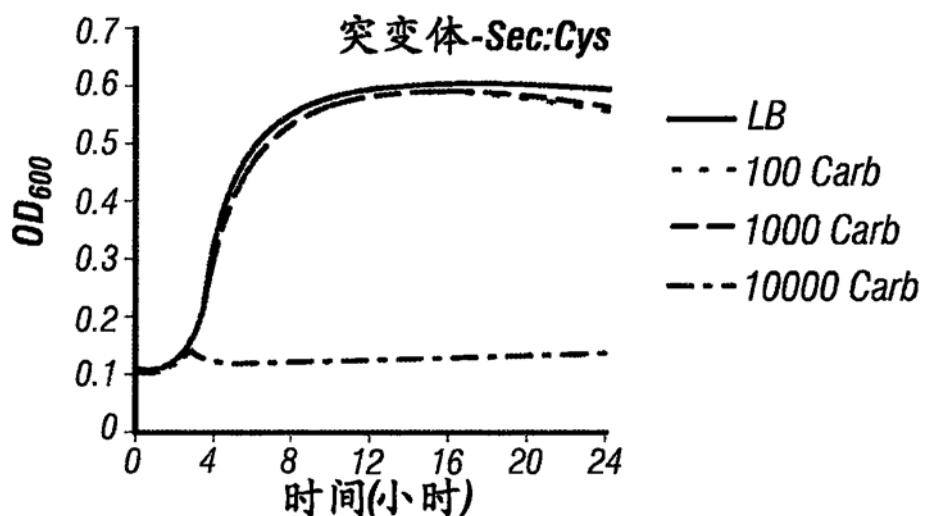


图2B

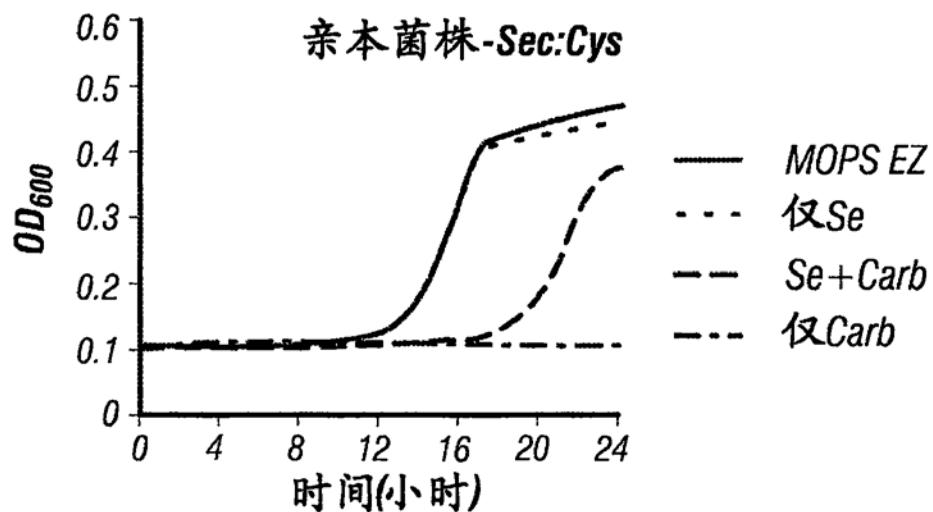


图3A

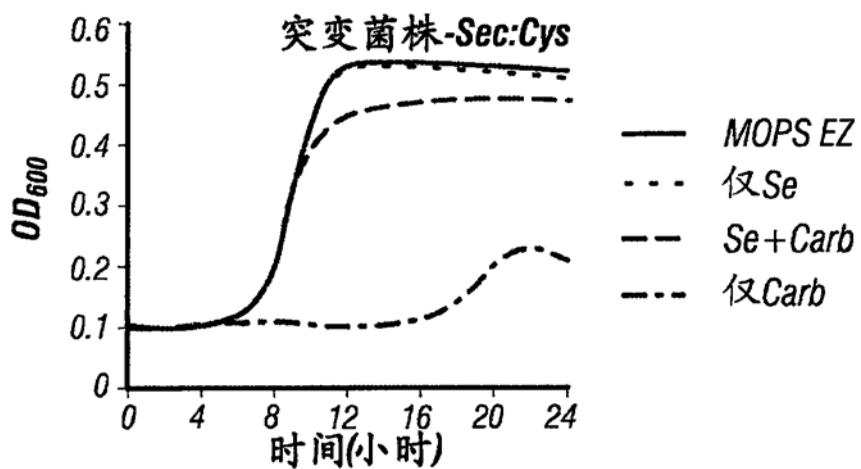


图3B

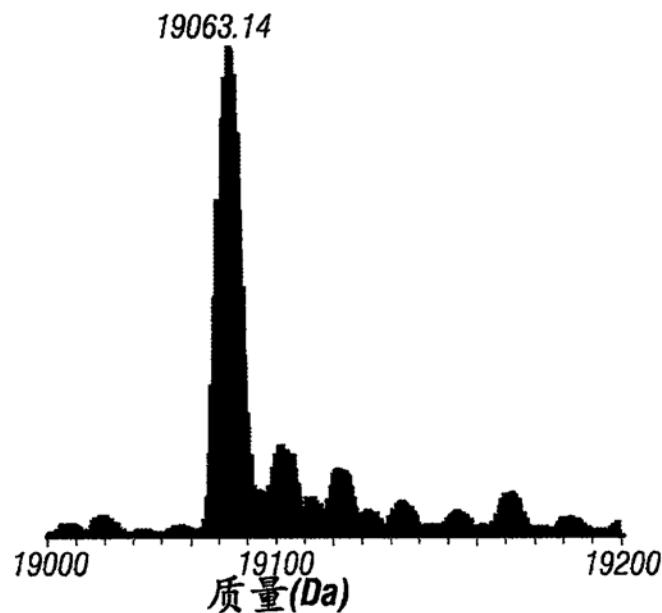


图4A

N M|S | | | A| A| L| A| V| D R| V| | G|M| E| N| A| M| P| W| | | P 25
 26 | A| D| | A| W| F| K|R| N| T| | | N| K| | | V| I| M| G| R| H| T| W| E| S| I| 50
 51 G|R| P| L| P| G| R| K| N| I| I| L| S| S| Q| P| G| T| D| D| R| V| T| W| V| 75
 76 K| S| V| D| | E| A| | | A| A| | | G| D| V| P| E| | | M| V| | | G| G| G| O| R| V| Y| 100
 101 | | Q| F| | | P| | K| | A| Q| | K| | Y| | | T| H| | | D| A| | | E| V| | | G| D| T| H| F| 125
 126 | | P| D| Y| | E| P| | D| W| E| S| V| F| S| E| F| H| D| A| Q| N| S| H| S| 150
 151 | | Y| C| F| | E| | | L| E| R| | R| G| S| H| | H| H| H| H| H| C

图4B

	理论单同位素质量 (Da)	实验单同位素质量 (Da)	质量精度 (ppm)
抗-MS2 scFv (Ser)	31603.40		
抗-MS2 scFv (2xU,无交联)	31855.10	31851.08	-125.57
抗-MS2 scFv (2xU,交联)	31851.06	31851.08	+0.6

图5A

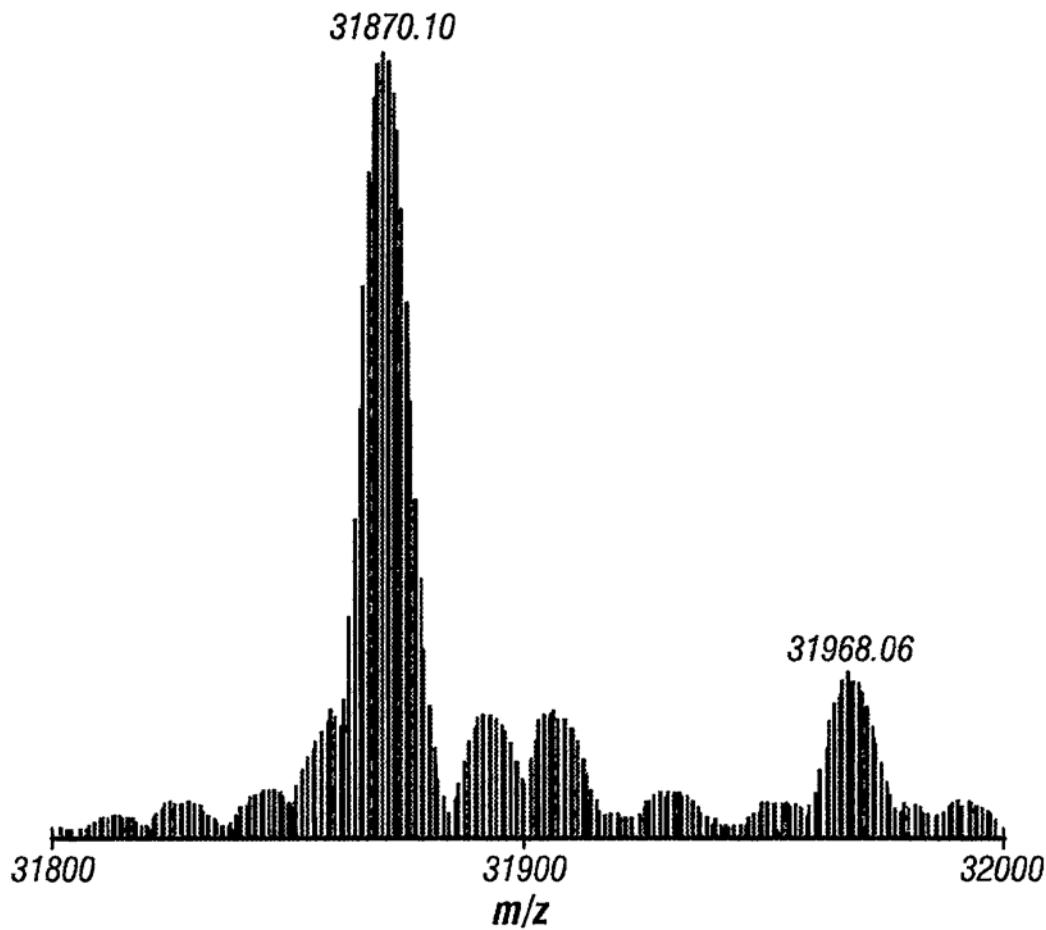


图5B

N M H̄H̄ H̄H̄H̄H̄H̄ L V̄ P̄ R̄ Ḡ S̄ L̄ Ḡ K̄ S Ā M̄ Ā Ā Ē V̄ K̄ L V 25
26 Ē S̄ Ḡ Ḡ Ḡ L̄ V̄ K̄ P̄ Ḡ S̄ L̄ K̄ L̄ S̄ U A A S G F T̄ F S 50
51 S Y A M S W V R Q T P E K R L E W V A T I S T G G 75
76 G Y T Y F̄ P D S V K G R F̄ T I S R D N A K N A L Y 100
101 L Q M K S L R S E D T Ā M Y Y U Ā R Q̄ Ḡ D̄ F̄ Ḡ D W 125
126 Ȳ F̄ D V W̄ Ḡ L Ā Ḡ T̄ T̄ V T̄ V S̄ S̄ Ḡ Ḡ Ḡ S̄ Ḡ Ḡ Ḡ S 150
151 Ḡ Ḡ Ḡ G S̄ L T̄ D L V L M T Q T̄ P L L S L P L V L S L G D Q A 175
176 S I S U R S S Q S L V H S N G N T Y L H W Y L Q K 200
201 P G Q S P K L L I Y K V S N R F S G V P D R F S G 225
226 S G S G T D F T L K I S R V I E A I E D L L G V L Y F U S 250
251 Q̄ S̄ L T̄ H I V L P L W L T̄ F L G I Ḡ G L T I K I L E L I I K I R I A L A L A L D Y K 275
276 D I H D I G D I Y I K D I H D I D I Y I K D D D D K C

图5C

N M E>V Q>L>L>E>S>G G>G>L>A K>P>G G>S L R>L S>C A >I 25
 26 S G F T>F T D>Y Y>M D W V R Q A P G K G L E W >V S 50
 51 R I S P G G D V T W Y A D S V K G R F T I S R D N 75
 76 A Q N T L Y L Q M N S L R>A E D T A V Y F>C A >R>D 100
 101 D> I >N>V>S>R> I >F> D> D>W>G>Q>G>V>L> L L T>V>S> S>G G>G G 125
 126 S G G>G>G S G>G>G>G S L E L L Q M L T Q L S L P L S S L L S I A S 150
 151 V G D R>V T I T>C R A S Q S I R S Y L A W Y Q Q K 175
 176 P G K A P K L L I Y D A A H L Q S G V P S R F S G 200
 201 S G S G T E F S L T I S S L Q Q P L E D F A L V Y L Y>C Q 225
 226 Q L R>N S L Y P L L T F G G G L T K V L E L I K L R L G I S I H I H I H I H 250
 251 H H C

图 6A

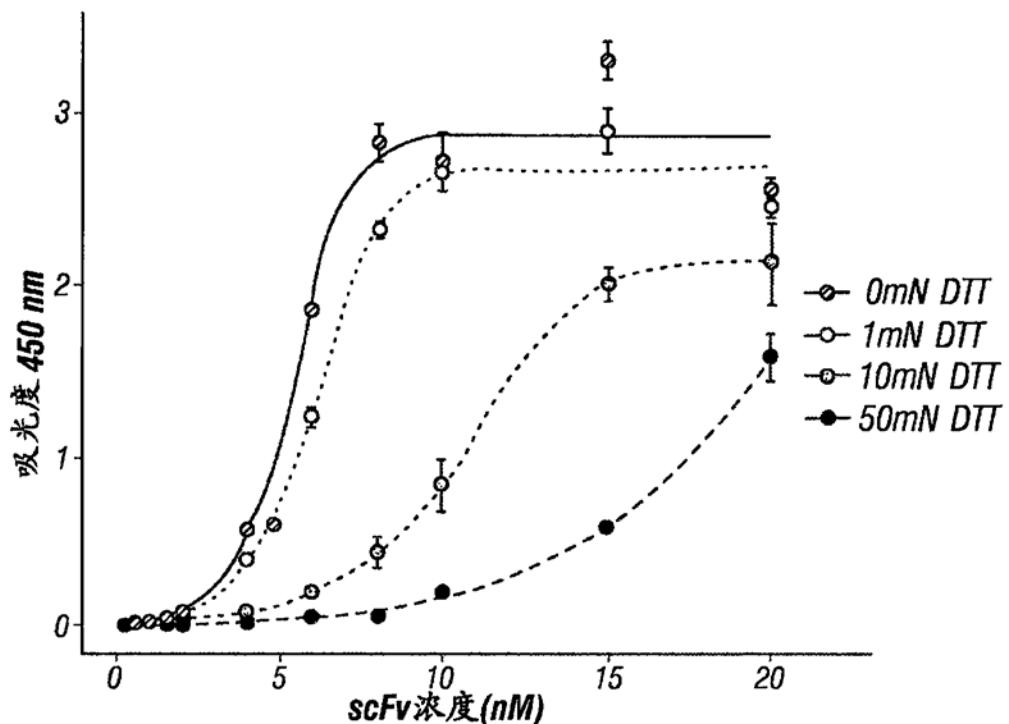


图6B

NM E[V Q L]L[E]S G G G L[A]K[P G G S L R]L[S]U AA 25
 26 S G F T F T D Y Y M D W V R Q A P G K G L E W V S 50
 51 R I S P G G D V T W Y A D S V K G R F T I S R D N 75
 76 A Q N T L Y L Q M N S L[R]A E D T A[V Y R]U A R[D 100
 101 D[I]V[V]S[R] I F[D D]W[G]Q G[V]L L T[V]S[S]G[G]G[G] 125
 126 S[G]G[G]G S G[G]G[G]SLE L[Q]MLT Q S P[S]S L[S]A S 150
 151 V G D R[V T[I]T]U R A S Q S I R S Y L A W Y Q Q K 175
 176 P G K A P K L L I Y D A A H L Q S G V P S R F S G 200
 201 S G S G T E F S L T I S S L[Q]P[L]E D F[L]A[V Y Y U]Q 225
 226 [Q]R[N]S[Y]P[L]T[F]G[G]G[T]K[V]E[L]I[K]R[G S]H[H]H[H]H 250
 251 H C

图6C

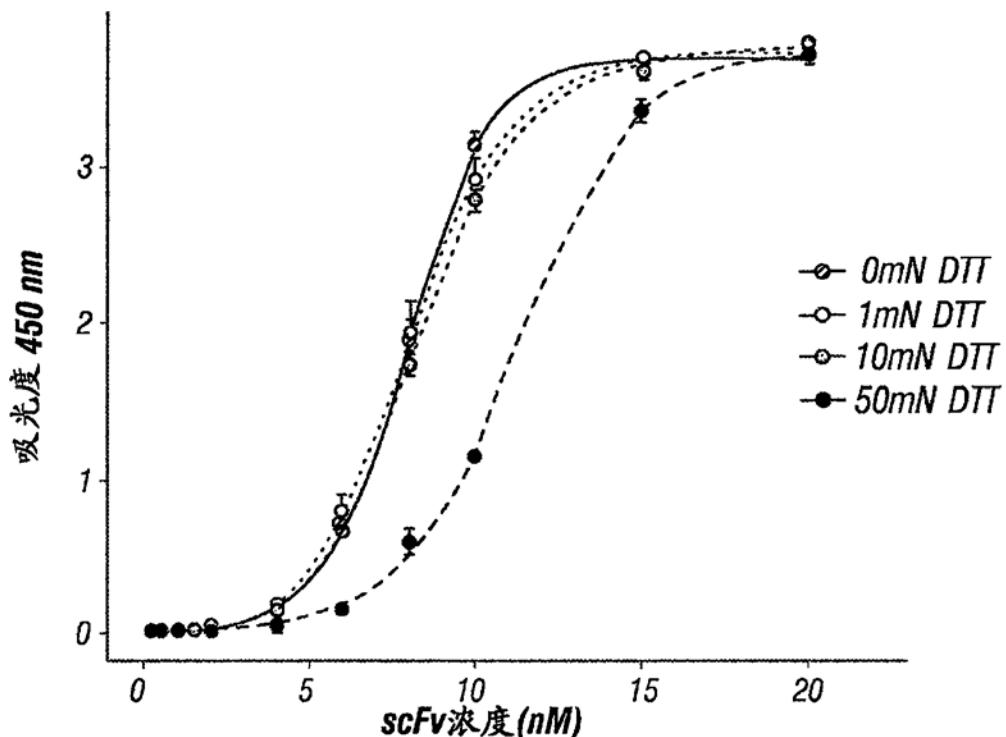


图6D

N M|D|I Q|M T Q|S P S S L S A S V G D|R V T I T C R 25
 26|A S Q D V|N|T|A V A|W Y Q Q|K P|G|K A|P|K| | | | Y 50
 51 S|A|S|F| |Y|S|G V|P|S|R|F|S|G S|R|S|G|T|D| F|T|L|T 75
 76 | |S|S | |Q P|E|D| F|A|T Y|Y|C Q Q|H Y|T|T|P| P T F|G 100
 101|Q G T K V E I K|G|A S G G|G G S G|G|G|G|S|G|G|G|G 125
 126|S|S E|V|Q L|V E|S G G|G|L|V|Q|P G G S|L|R L S|C A 150
 151 A S G F|N I |K D T Y I H|W V R|Q|A|P|G|K|G|L E|W V 175
 176 A|R|I Y|P|T N G|Y T R|Y|A|D|S V|K|G R|F T I S A D 200
 201 T S K|N T A Y|L Q M N|S L R A|E|D|T A V|Y|Y|C|S R 225
 226 W|G G|D G|F Y|A|M D|Y|W|G Q|G|T|L|L|V|T V|S|S G|S H 250
 251 H H H H H C

图7A

N M D I₁Q₁M₁T₁Q₁S₁P₁S S L S A S V G D R V T I₁T₁C R 25
 26 A I₁S₁Q₁D₁V₁N₁T₁A₁V₁A₁W₁Y₁Q Q₁K₁P₁G₁K₁A₁P₁K₁ I₁I₁ I₁Y 50
 51 S A I₁S₁F₁ I₁Y₁S₁G₁ V₁P S₁R₁ F₁S₁G₁R₁S₁G₁ T₁D₁ F₁T₁L₁T₁ 75
 76 I₁I₁S₁S L₁Q P₁E₁D₁ F₁A₁T₁Y₁Y₁C Q Q H Y₁T₁T₁P₁P₁T F₁G 100
 101 Q G T K V E I K₁G₁A S G₁G₁G G S G G G G₁S₁G₁G₁G₁G 125
 126 S S₁E₁V₁Q₁L₁V₁E₁S₁G₁G₁G L₁V Q P G G S L R₁L S₁C A 150
 151 A S G F₁N I K D T Y I H W V R Q A P G K G L E W V 175
 176 A R I Y P T N G Y T R Y A D S V K G R F T I S A D 200
 201 T S K N T A Y L Q M N S L R A E D T A V Y Y C S R 225
 226 W G G D G F₁Y₁A₁M₁D₁Y₁W₁G₁Q₁G₁T₁L₁V₁T₁V₁S₁S₁G₁S₁H 250
 251 H₁H₁H₁H H H C

图7B

N M₁D I₁Q₁M₁T₁Q₁S₁P₁S S L S A S V₁G₁D R₁V₁T I₁T₁U R 25
 26 A S Q D V N T₁A V A W Y Q Q K P G K A₁P K L L I₁Y 50
 51 S A S F₁I₁Y S G V P S R F S G S R S G T D₁F₁T L T 75
 76 I S S L Q₁P E D F A T Y Y U Q Q H Y₁T T₁P₁P₁T F₁G 100
 101 Q G₁T K₁V E I K₁G₁A S G₁G₁G₁G₁G₁G₁G₁S G₁G₁G₁G 125
 126 S S₁E₁V₁Q L₁V E S₁G₁G₁G L₁V₁Q₁P₁G₁G₁S L₁R₁L S₁U A 150
 151 A S G F N I K D T Y I H W V R Q A P G K G L E W V 175
 176 A R I Y P T N G Y T R Y A D S V K G R F T I S A D 200
 201 T S K N T A Y L Q M N S L R A E D T A V Y Y U S R 225
 226 W G G D G F₁Y₁A₁M₁D₁Y₁W₁G₁Q₁G₁T₁L₁V₁T₁V₁S₁S₁G₁S₁H 250
 251 H₁H₁H₁H H H C

图7C