



## (12) 发明专利申请

(10) 申请公布号 CN 104364259 A

(43) 申请公布日 2015. 02. 18

(21) 申请号 201380027365. 7

R · 奇拉库鲁 M · J · 哈米尔

(22) 申请日 2013. 03. 15

(74) 专利代理机构 北京市金杜律师事务所

11256

(30) 优先权数据

代理人 陈文平 马慧

61/615, 819 2012. 03. 26 US

(85) PCT国际申请进入国家阶段日

(51) Int. Cl.

2014. 11. 25

*C07K 14/435* (2006. 01)

(86) PCT国际申请的申请数据

*A23J 3/04* (2006. 01)

PCT/US2013/032218 2013. 03. 15

*A23L 1/305* (2006. 01)

(87) PCT国际申请的公布数据

*C12N 15/63* (2006. 01)

W02013/148330 EN 2013. 10. 03

*C12N 1/21* (2006. 01)

(71) 申请人 普罗努塔利亚公司

地址 美国马萨诸塞州

(72) 发明人 D · A · 贝利 B · A · 伯希吉安

权利要求书7页 说明书117页

N · W · 希尔弗 G · 冯马尔特扎恩

序列表 (电子公布) 附图1页

按照条约第19条修改的权利要求书7页

(54) 发明名称

营养性片段、蛋白质和方法

(57) 摘要

本发明提供营养性蛋白质。在一些实施方案中,所述营养性蛋白质包含含有天然存在的营养性蛋白质的片段的第一多肽序列。在一些实施方案中,所述片段包含以下至少一者:a) 所述营养性蛋白质中存在的支链氨基酸残基与总氨基酸残基的比率增大;b) 所述营养性蛋白质中存在的亮氨酸残基与总氨基酸残基的比率增大;和c) 所述营养性蛋白质中存在的必需氨基酸残基与总氨基酸残基的比率增大。在一些实施方案中,所述片段包含以下至少一者:a) 所述营养性蛋白质中存在的支链氨基酸残基与总氨基酸残基的比率;b) 所述营养性蛋白质中存在的亮氨酸残基与总氨基酸残基的比率;和c) 所述营养性蛋白质中存在的必需氨基酸残基与总氨基酸残基的比率等于或大于如乳清、蛋或大豆蛋白质的基准蛋白质中存在的相应比率。也尤其提供编码所述蛋白质的核酸、制备所述蛋白质的重组微生物、使用重组微生物制备所述蛋白质的方法、包含所述蛋白质的组合物和使用所述蛋白质的方法。

1. 一种包含与天然存在的蛋白质的片段同源的第一多肽序列的分离的营养性蛋白质，其中所述第一多肽序列包含以下至少一者：

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%；
- b. Leu 残基与总氨基酸残基的比率为至少 11%；和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。

2. 如权利要求 1 所述的分离的营养性蛋白质，其中所述第一多肽序列进一步包含各必需氨基酸的至少一者。

3. 如权利要求 1 或 2 所述的分离的营养性蛋白质，其中所述第一多肽序列包含与天然存在的蛋白质的所述片段的至少 70% 同源性。

4. 如权利要求 1 或 2 所述的分离的营养性蛋白质，其中所述第一多肽序列包含与天然存在的蛋白质的所述片段的至少 95% 同源性。

5. 如权利要求 1 至 4 中任一项所述的分离的营养性蛋白质，其中天然存在的蛋白质的所述片段包含至少 25 个氨基酸残基。

6. 如权利要求 1 至 4 中任一项所述的分离的营养性蛋白质，其中天然存在的蛋白质的所述片段包含至少 50 个氨基酸残基。

7. 如权利要求 1 至 6 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列包含：

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%；
- b. Leu 残基与总氨基酸残基的比率为至少 11%；和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。

8. 如权利要求 1 至 7 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列包含天然存在的营养性蛋白质的所述片段。

9. 如权利要求 1 至 7 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列由天然存在的营养性蛋白质的所述片段组成。

10. 如权利要求 1 至 9 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列不是过敏原。

11. 如权利要求 10 所述的分离的营养性蛋白质，其中所述第一多肽序列与已知过敏原具有小于 50% 整体同源性。

12. 如权利要求 1 至 11 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列不是毒素。

13. 如权利要求 12 所述的分离的营养性蛋白质，其中所述第一多肽序列与已知毒素具有小于 50% 整体同源性。

14. 如权利要求 1 至 13 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列具有模拟胃消化半衰期为小于 60 分钟。

15. 如权利要求 14 所述的分离的营养性蛋白质，其中所述第一多肽序列具有模拟胃消化半衰期为小于 30 分钟。

16. 如权利要求 15 所述的分离的营养性蛋白质，其中所述第一多肽序列具有模拟胃消化半衰期为小于 10 分钟。

17. 如权利要求 1 至 16 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列

在模拟胃液中完全消化。

18. 如权利要求 1 至 17 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。

19. 如权利要求 1 至 18 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含半胱氨酸残基。

20. 如权利要求 1 至 19 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含二硫键。

21. 如权利要求 1 至 20 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含 N 连接的糖基化。

22. 如权利要求 1 至 21 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含 O 连接的糖基化。

23. 如权利要求 1 至 22 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列对聚集具有抗性。

24. 如权利要求 1 至 23 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列在 pH 7 下是阴离子型的。

25. 如权利要求 1 至 24 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列在 pH 7 下具有水溶解度为至少 12.5g/L。

26. 如权利要求 1 至 25 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有计算溶合计分为 -20 或更低。

27. 如权利要求 1 至 26 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有计算聚集计分为 0.75 或更低。

28. 如权利要求 1 至 27 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有计算聚集计分为 0.5 或更低。

29. 如权利要求 1 至 28 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含选自以下的氨基酸序列:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

30. 如权利要求 1 至 29 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列由选自以下的氨基酸序列组成:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

31. 如权利要求 1 至 30 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列与至少一个选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列至少 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 99.5% 同源。

32. 如权利要求 1 至 31 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列

是所述天然存在的营养性蛋白质的至少 25 个氨基酸的片段。

33. 如权利要求 1 至 32 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列是所述天然存在的营养性蛋白质的至少 50 个氨基酸的片段。

34. 如权利要求 1 至 33 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质由所述第一多肽序列组成。

35. 如权利要求 1 至 34 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质进一步包含用于亲和纯化的多肽标签。

36. 如权利要求 35 所述的分离的营养性蛋白质,其中用于亲和纯化的所述标签是聚组氨酸标签。

37. 一种包含与天然存在的蛋白质的片段同源的第一多肽序列的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含以下至少一者:

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%;
- b. Leu 残基与总氨基酸残基的比率为至少 11%;和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。

38. 如权利要求 37 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质进一步包含各必需氨基酸的至少一者。

39. 如权利要求 37 或 38 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含与天然存在的蛋白质的所述片段的至少 70% 同源性。

40. 如权利要求 37 或 38 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含与天然存在的蛋白质的所述片段的至少 95% 同源性。

41. 如权利要求 37 至 40 中任一项所述的分离的营养性蛋白质,其中天然存在的蛋白质的所述片段包含至少 25 个氨基酸残基。

42. 如权利要求 37 至 40 中任一项所述的分离的营养性蛋白质,其中天然存在的蛋白质的所述片段包含至少 50 个氨基酸残基。

43. 如权利要求 37 至 42 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含:

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%;
- b. Leu 残基与总氨基酸残基的比率为至少 11%;和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。

44. 如权利要求 37 至 43 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含天然存在的营养性蛋白质的片段。

45. 如权利要求 37 至 43 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质由天然存在的营养性蛋白质的片段组成。

46. 如权利要求 37 至 45 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质不是过敏原。

47. 如权利要求 46 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质与已知过敏原具有小于 50% 整体同源性。

48. 如权利要求 37 至 47 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质不是毒素。



49. 如权利要求 48 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质与已知毒素具有小于 50%整体同源性。

50. 如权利要求 37 至 49 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有模拟胃消化半衰期为小于 60 分钟。

51. 如权利要求 50 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有模拟胃消化半衰期为小于 30 分钟。

52. 如权利要求 51 所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有模拟胃消化半衰期为小于 10 分钟。

53. 如权利要求 37 至 52 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质在模拟胃液中完全消化。

54. 如权利要求 37 至 53 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。

55. 如权利要求 37 至 54 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质不包含半胱氨酸残基。

56. 如权利要求 37 至 55 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质不包含二硫键。

57. 如权利要求 37 至 56 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质不包含 N 连接的糖基化。

58. 如权利要求 37 至 57 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质不包含 O 连接的糖基化。

59. 如权利要求 37 至 58 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质对聚集具有抗性。

60. 如权利要求 37 至 59 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质在 pH 7 下是阴离子型的。

61. 如权利要求 37 至 60 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质在 pH 7 下具有水溶解度为至少 12.5g/L。

62. 如权利要求 37 至 61 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有计算溶合计分为 -20 或更低。

63. 如权利要求 37 至 62 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有计算聚集计分为 0.75 或更低。

64. 如权利要求 37 至 63 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有计算聚集计分为 0.5 或更低。

65. 如权利要求 37 至 64 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含选自以下的氨基酸序列:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

66. 如权利要求 37 至 65 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性

蛋白质由选自以下的氨基酸序列组成：

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列；
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物；和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

67. 如权利要求 37 至 66 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质与至少一个选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列至少 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 99.5% 同源。

68. 如权利要求 37 至 67 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质进一步包含用于亲和纯化的多肽标签。

69. 如权利要求 68 所述的分离的营养性蛋白质，其中用于亲和纯化的所述标签是聚组氨酸标签。

70. 一种分离的核酸，其包含编码根据权利要求 1 至 69 中任一项所述的营养性蛋白质的核酸序列。

71. 如权利要求 70 所述的分离的核酸，其中所述分离的核酸选自基因组 DNA、cDNA、正义 RNA 和反义 RNA。

72. 如权利要求 71 所述的分离的核酸，其中所述分离的核酸是基因组 DNA。

73. 如权利要求 71 所述的分离的核酸，其中所述分离的核酸是 cDNA。

74. 如权利要求 70 至 73 中任一项所述的分离的核酸，其进一步包含可操作地连接于编码所述营养性蛋白质的所述核酸序列的表达控制序列。

75. 一种载体，其包含编码如权利要求 1 至 69 中任一项所述的营养性蛋白质的核酸序列。

76. 如权利要求 75 所述的载体，其进一步包含可操作地连接于编码所述营养性蛋白质的所述核酸序列的表达控制序列。

77. 一种重组微生物，其包含如权利要求 70 至 74 中任一项所述的核酸和如权利要求 75 和 76 中任一项所述的载体的至少一者。

78. 根据权利要求 77 所述的重组微生物，其中所述重组微生物是原核生物。

79. 根据权利要求 78 所述的重组原核生物，其中所述原核生物是异养性的。

80. 根据权利要求 78 所述的重组原核生物，其中所述原核生物是自养性的。

81. 根据权利要求 78 所述的重组原核生物，其中所述原核生物是细菌。

82. 一种制备根据权利要求 1 至 69 中任一项所述的营养性蛋白质的方法，所述方法包括在足以由根据权利要求 77 至 81 中任一项所述的重组微生物产生所述营养性蛋白质的条件下培养所述重组微生物。

83. 如权利要求 82 所述的方法，其进一步包括自所述培养物分离所述营养性蛋白质。

84. 一种营养性组合物，其包含如权利要求 1 至 69 中任一项所述的营养性蛋白质和至少一种第二组分。

85. 如权利要求 84 所述的营养性组合物，其中所述至少一种第二组分选自蛋白质、多肽、肽、游离氨基酸、碳水化合物、脂肪、矿物质或矿物质来源、维生素、补充剂、生物体、医药和赋形剂。

86. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是蛋白质。
87. 如权利要求 86 所述的方法,其中所述至少一种第二组分是营养性蛋白质。
88. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是选自必需氨基酸、非必需氨基酸、支链氨基酸、非标准氨基酸和修饰的氨基酸的游离氨基酸。
89. 如权利要求 88 所述的营养性组合物,其中所述至少一种第二组分是选自必需氨基酸的游离氨基酸。
90. 如权利要求 88 所述的营养性组合物,其中所述至少一种第二组分是选自支链氨基酸的游离氨基酸。
91. 如权利要求 88 所述的营养性组合物,其中所述至少一种第二组分是 Leu。
92. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是脂质。
93. 如权利要求 92 所述的营养性组合物,其中所述脂质选自脂肪、油、甘油三酯、胆固醇、磷脂和脂肪酸。
94. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分选自矿物质和维生素。
95. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是补充剂。
96. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是生物体。
97. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是医药。
98. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是赋形剂。
99. 如权利要求 98 所述的营养性组合物,其中所述至少一种赋形剂选自缓冲剂、防腐剂、稳定剂、粘合剂、压实剂、润滑剂、分散增强剂、崩解剂、调味剂、甜味剂、着色剂。
100. 如权利要求 84 至 99 中任一项所述的营养性组合物,其中所述营养性组合物被配制成液体溶液、浆液、混悬液、凝胶剂、糊剂、粉末或固体。
101. 一种制备如权利要求 84 至 100 中任一项所述的营养性组合物的方法,其包括提供根据权利要求 1 至 69 中任一项的营养性蛋白质以及组合所述营养性蛋白质与所述至少一种第二组分。
102. 一种维持或增加受试者的肌肉质量、肌肉强度和功能性表现的至少一者的方法,所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。
103. 一种维持或达成受试者的合乎需要的身体质量指数的方法,所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。
104. 如权利要求 102 或 103 所述的方法,其中所述受试者处于以下至少一种状态下: 年长、患有医学上危急疾病、以及罹患蛋白质-能量营养不良症。
105. 一种向患有蛋白质-能量营养不良症的受试者提供蛋白质的方法,所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。
106. 一种增加受试者的产热的方法,所述方法包括向所述受试者提供足量的如权利要

求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

107. 一种诱导受试者的饱足反应和饱腹感反应的至少一者的方法,所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

108. 如权利要求 106 或 107 所述的方法,其中所述受试者是肥胖的。

109. 如权利要求 102 至 108 中任一项所述的方法,其中如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物是由所述受试者与进行运动配合加以消耗。

110. 一种治疗受试者的恶病质、肌肉减少症和虚弱的至少一者的方法,所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

111. 如权利要求 102 至 110 中任一项所述的方法,其中如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物是由所述受试者通过经口、肠或胃肠外途径加以消耗。

## 营养性片段、蛋白质和方法

[0001] 相关申请的引用：

[0002] 本申请要求 2012 年 3 月 26 日提交的美国临时专利申请号 61/615,819 的优先权，所述专利申请据此以引用的方式整体并入本文。

[0003] 序列表：

[0004] 本申请含有已通过 EFS-Web 以 ASCII 格式提交且据此以引用的方式整体并入本文的序列表。2013 年 3 月 12 日创建的所述 ASCII 拷贝名为 1005.005-PCT\_SL.txt 且大小为 3,234,600 字节。

[0005] 引言：

[0006] 膳食蛋白质是为人类健康和生长所必需之营养物。世界卫生组织推荐当处于能量平衡和重量稳定时，膳食蛋白质应占能量摄取量的约 10 至 15%。各个国家的平均每日蛋白质摄取量指示这些推荐与全世界消耗的蛋白质的量一致。当以能量平衡消耗时，平均 20 至 30% 的能量来自蛋白质的餐食代表高蛋白质膳食。

[0007] 身体不能合成为健康和生长所必需的某些氨基酸，且代之以必须自食物获得它们。称为“必需氨基酸”的这些氨基酸是组氨酸 (H)、异亮氨酸 (I)、亮氨酸 (L)、赖氨酸 (K)、甲硫氨酸 (M)、苯丙氨酸 (F)、苏氨酸 (T)、色氨酸 (W) 和缬氨酸 (V)。提供所有必需氨基酸的膳食蛋白质称为“高品质”蛋白质。如肉、鱼、家禽、蛋和乳制品的动物性食物通常被认为是提供必需氨基酸的良好平衡的高品质蛋白质来源。酪蛋白（一种通常见于哺乳动物奶中，占牛奶中的蛋白质的 80% 的蛋白质）和乳清（在奶已凝固和应变之后剩余的呈液体状态的蛋白质）是高品质膳食蛋白质的主要来源。不提供必需氨基酸的良好平衡的食物称为“低品质”蛋白质。大多数水果和蔬菜是蛋白质的不良来源。包括菜豆、豌豆、小扁豆、坚果和谷粒（如小麦）的一些植物性食物是蛋白质的较好来源。大豆，作为一种由黄豆制得的植物性蛋白质，被一些人视为高品质蛋白质。

[0008] 对消耗大量蛋白质在人中的急性效应的研究已显示在膳食中包括以及在一些情况下增加蛋白质含量可具有有益作用。举例来说，研究已显示摄取蛋白质可在人受试者中诱导餐后饱腹感（包括通过抑制饥饿感），诱导产热以及降低血糖反应。

[0009] 对高蛋白质膳食有关重量减轻的研究已显示蛋白质正面影响能量消耗和瘦体重。进一步研究已显示就含有的能量的至少 5% 来自蛋白质的膳食而言，过度进食会产生显著较小重量增加，且高蛋白质膳食降低能量摄取。

[0010] 临床研究提供证据表明蛋白质阻止归因于衰老或卧床休息的肌肉损失。具体来说，研究已显示补充蛋白质增加长期卧床休息期间的肌肉分数合成率 (FSR)，维持长期卧床休息期间的腿部质量和强度，增加瘦体重，改进步态和平衡的功能性量度，且可充当用于由于不活动或长期卧床休息而处于肌肉减少症的风险下的个体的可行性介入手段。

[0011] 在运动员中进行的关于增加肌肉蛋白质合成代谢的研究已显示在运动之后提供的蛋白质促进肌肉肥大的程度大于由单独运动达成的肌肉肥大。也已显示在运动之后提供的蛋白质支持蛋白质合成而不使蛋白质分解有任何增加，从而导致净正性蛋白质平衡和肌肉质量增大。尽管肌肉蛋白质合成似乎以剂量 - 反应方式对补充必需氨基酸起反应，但并

非所有蛋白质在构建肌肉方面都等同。举例来说,在用阻力训练支持肌肉质量增大方面,奶蛋白似乎优于大豆,而两者均优于单独碳水化合物。氨基酸亮氨酸是刺激肌肉蛋白质合成的重要因素。

[0012] 通常见于食物中的全部蛋白质未必以高效方式提供满足如人的哺乳动物的氨基酸需求的氨基酸组成。结果是为达到各必需氨基酸的最小需求,在膳食中消耗的总蛋白质的量必须大于如果膳食蛋白质的品质较高所需需要的量。通过增加膳食中的蛋白质品质,有可能相较于包括品质较低的蛋白质的膳食,降低必须消耗的蛋白质总量。

[0013] 一般来说,在哺乳动物膳食中,具有较高蛋白质品质的蛋白质被视为比不具有较高蛋白质品质的其它蛋白质更有益。所述蛋白质例如适用作哺乳动物膳食的组分。在某些情况下,所述蛋白质尤其促进维持肌肉质量、健康身体质量指数和血糖平衡。因此,对具有高蛋白质品质的蛋白质的来源存在需要。

[0014] 在传统上,氨基酸的合乎需要混合物(如包含必需氨基酸的混合物)已通过水解必需氨基酸的水平相对较高的蛋白质(如乳清蛋白质)和/或通过在任选也包括如乳清的水解蛋白质的混合物中组合游离氨基酸来提供。这个类型的混合物可具有苦味且可被视为不适合某些用途或不合乎某些用途需要。因此,所述混合物有时包括调味剂来掩蔽游离氨基酸和/或水解蛋白质的味道。在一些情况下,发现其中一定比例的氨基酸内含物是由多肽或蛋白质提供的组合物具有的味道好于较高比例的总氨基酸是以游离氨基酸和/或某些水解蛋白质形式提供的组合物。然而,所述组合物的可用性已受到限制,因为营养性制剂在传统上已由自天然食物产品分离的蛋白质(如自奶分离的乳清或自大豆分离的大豆蛋白质)制得。那些蛋白质的氨基酸分布未必满足哺乳动物的氨基酸需求。此外,商品蛋白质通常由蛋白质和/或蛋白质水解物的混合物组成,所述混合物在它们的蛋白质组成方面可变化,由此导致关于它们的营养价值的不可预测性。此外,所述高品质蛋白质的来源的数目有限已意味着仅某些氨基酸组合可以蛋白质形式以大规模用于摄取。

[0015] 为供应如酪蛋白和乳清、蛋及肉的高品质动物蛋白质来源、以及如大豆的植物蛋白质所需的农业方法也需要大量能量输入且具有潜在有害环境影响。因此,将在某些情况下适用的是具有供应供哺乳动物消耗的蛋白质的替代性来源和方法。

[0016] 在理论上,可在实验室环境下设计和产生包含所需氨基酸混合物的合成多肽序列。然而,这个方法可产生各种顾虑,且因此并非始终适用。首先,熟练技术人员了解的是获得所述合成序列的高产生水平可极具挑战性。其次,即使所述合成蛋白质被合成,它用于营养性产品中的适合性也将是不确定的。举例来说,所述非天然存在的多肽可为过敏原或毒素。因此,在一些实施方案中,本公开提供天然蛋白质或多肽序列或其变体。

[0017] 在本公开中,发明者提供包含与天然存在的蛋白质的片段同源的第一多肽序列的营养性蛋白质。营养性蛋白质和主要组成它们的片段由适用的氨基酸的组合构成。蛋白质可通过不仅仅依赖于传统农业的生产方法来产生。举例来说,发明者已发现且本公开提供包含天然存在的蛋白质的片段且主要由氨基酸的组合组成的营养性蛋白质,其含有适用水平的以下至少一者:支链氨基酸与总氨基酸的比率、氨基酸亮氨酸与总氨基酸的比率、和必需氨基酸与总氨基酸的比率。在一些实施方案中,营养性蛋白质包含一定水平的以下至少一者:a) 营养性蛋白质中存在的支链氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质中存在的支链氨基酸残基与总氨基酸残基的比率;b) 营养性蛋白质中存在的亮氨

酸残基与总氨基酸残基的比率等于或大于乳清蛋白质中存在的亮氨酸残基与总氨基酸残基的比率 ;和 c) 营养性蛋白质中存在的必需氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质中存在的必需氨基酸残基与总氨基酸残基的比率。

[0018] 本公开也尤其提供编码蛋白质的核酸、制备蛋白质的重组微生物、使用重组微生物制备蛋白质的方法、包含蛋白质的组合物和使用蛋白质的方法。

[0019] 概述 :

[0020] 在第一方面,本公开提供包含与天然存在的蛋白质的片段同源的第一多肽序列的分离的营养性蛋白质,其中所述第一多肽序列包含以下至少一者 :a. 支链氨基酸残基与总氨基酸残基的比率为至少 24% ;b. Leu 残基与总氨基酸残基的比率为至少 11% ;和 c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。在一些实施方案中,第一多肽序列进一步包含各必需氨基酸的至少一者。在一些实施方案中,第一多肽序列包含与天然存在的蛋白质的片段的至少 70% 同源性。在一些实施方案中,第一多肽序列包含与天然存在的蛋白质的片段的至少 95% 同源性。在一些实施方案中,天然存在的蛋白质的片段包含至少 25 个氨基酸残基。在一些实施方案中,天然存在的蛋白质的片段包含至少 50 个氨基酸残基。在一些实施方案中,第一多肽序列包含 :a. 支链氨基酸残基与总氨基酸残基的比率为至少 24% ;b. Leu 残基与总氨基酸残基的比率为至少 11% ;和 c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。在一些实施方案中,第一多肽序列包含天然存在的营养性蛋白质的片段。在一些实施方案中,第一多肽序列由天然存在的营养性蛋白质的片段组成。

[0021] 在一些实施方案中,第一多肽序列不是过敏原。在一些实施方案中,第一多肽序列与已知过敏原具有小于 50% 整体同源性。

[0022] 在一些实施方案中,第一多肽序列不是毒素。在一些实施方案中,第一多肽序列与已知毒素具有小于 50% 整体同源性。

[0023] 在一些实施方案中,第一多肽序列具有模拟胃消化半衰期为小于 60 分钟。在一些实施方案中,第一多肽序列具有模拟胃消化半衰期为小于 30 分钟。在一些实施方案中,第一多肽序列具有模拟胃消化半衰期为小于 10 分钟。在一些实施方案中,第一多肽序列在模拟胃液中完全消化。在一些实施方案中,第一多肽序列包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。在一些实施方案中,第一多肽序列不包含半胱氨酸残基。在一些实施方案中,第一多肽序列不包含二硫键。在一些实施方案中,第一多肽序列不包含 N 连接的糖基化。在一些实施方案中,第一多肽序列不包含 O 连接的糖基化。

[0024] 在一些实施方案中,第一多肽序列对聚集具有抗性。在一些实施方案中,第一多肽序列在 pH 7 下是阴离子型的。在一些实施方案中,第一多肽序列在 pH 7 下具有水溶解度为至少 12.5g/L。在一些实施方案中,第一多肽序列具有计算溶合计分为 -20 或更低。在一些实施方案中,第一多肽序列具有计算聚集计分为 0.75 或更低。在一些实施方案中,第一多肽序列具有计算聚集计分为 0.5 或更低。

[0025] 在一些实施方案中,第一多肽序列包含选自以下的氨基酸序列 :i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列 ;ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物 ;和 iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。在一些实施方案中,第一多肽序列由选自以下的氨基酸序列组成 :i. 选自 SEQ ID

NO:1 至 SEQ ID NO:2609 的氨基酸序列;ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物,和 iii. 选自 v 的氨基酸序列的突变蛋白。在一些实施方案中,第一多肽序列与至少一个选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列至少 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 99.5% 同源。

[0026] 在一些实施方案中,第一多肽序列是天然存在的营养性蛋白质的至少 25 个氨基酸的片段。在一些实施方案中,第一多肽序列是天然存在的营养性蛋白质的至少 50 个氨基酸的片段。在一些实施方案中,分离的营养性蛋白质由第一多肽序列组成。在一些实施方案中,分离的营养性蛋白质进一步包含用于亲和纯化的多肽标签。在一些实施方案中,用于亲和纯化的标签是聚组氨酸标签。

[0027] 在另一方面,本公开提供包含与天然存在的蛋白质的片段同源的第一多肽序列的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含以下至少一者:a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%;b. Leu 残基与总氨基酸残基的比率为至少 11%;和 c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。在一些实施方案中,分离的营养性蛋白质进一步包含各必需氨基酸的至少一者。在一些实施方案中,分离的营养性蛋白质包含与天然存在的蛋白质的片段的至少 70% 同源性。在一些实施方案中,分离的营养性蛋白质包含与天然存在的蛋白质的片段的至少 95% 同源性。在一些实施方案中,天然存在的蛋白质的片段包含至少 25 个氨基酸残基。在一些实施方案中,天然存在的蛋白质的片段包含至少 50 个氨基酸残基。在一些实施方案中,分离的营养性蛋白质包含:a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%;b. Leu 残基与总氨基酸残基的比率为至少 11%;和 c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。在一些实施方案中,分离的营养性蛋白质包含天然存在的营养性蛋白质的片段。在一些实施方案中,分离的营养性蛋白质由天然存在的营养性蛋白质的片段组成。

[0028] 在一些实施方案中,分离的营养性蛋白质不是过敏原。在一些实施方案中,分离的营养性蛋白质与已知过敏原具有小于 50% 整体同源性。

[0029] 在一些实施方案中,分离的营养性蛋白质不是毒素。在一些实施方案中,分离的营养性蛋白质与已知毒素具有小于 50% 整体同源性。

[0030] 在一些实施方案中,分离的营养性蛋白质具有模拟胃消化半衰期为小于 60 分钟。在一些实施方案中,分离的营养性蛋白质具有模拟胃消化半衰期为小于 30 分钟。在一些实施方案中,分离的营养性蛋白质具有模拟胃消化半衰期为小于 10 分钟。在一些实施方案中,分离的营养性蛋白质在模拟胃液中完全消化。在一些实施方案中,分离的营养性蛋白质包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。在一些实施方案中,分离的营养性蛋白质不包含半胱氨酸残基。在一些实施方案中,分离的营养性蛋白质不包含二硫键。在一些实施方案中,分离的营养性蛋白质不包含 N 连接的糖基化。在一些实施方案中,分离的营养性蛋白质不包含 O 连接的糖基化。

[0031] 在一些实施方案中,分离的营养性蛋白质对聚集具有抗性。在一些实施方案中,分离的营养性蛋白质在 pH 7 下是阴离子型的。在一些实施方案中,分离的营养性蛋白质在 pH 7 下具有水溶解度为至少 12.5g/L。在一些实施方案中,分离的营养性蛋白质具有计算溶合计分为 -20 或更低。在一些实施方案中,分离的营养性蛋白质具有计算聚集计分为 0.75 或



更低。在一些实施方案中,分离的营养性蛋白质具有计算聚集计分为 0.5 或更低。

[0032] 在一些实施方案中,分离的营养性蛋白质包含选自以下的氨基酸序列:i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和 iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。在一些实施方案中,分离的营养性蛋白质由选自以下的氨基酸序列组成:i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和 iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。在一些实施方案中,分离的营养性蛋白质与至少一个选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列至少 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 99.5% 同源。在一些实施方案中,分离的营养性蛋白质进一步包含用于亲和纯化的多肽标签。在一些实施方案中,用于亲和纯化的标签是聚组氨酸标签。

[0033] 在另一方面,本公开提供包含编码根据本公开的营养性蛋白质的核酸序列的分离的核酸。在一些实施方案中,分离的核酸选自基因组 DNA、cDNA、正义 RNA 和反义 RNA。在一些实施方案中,分离的核酸是基因组 DNA。在一些实施方案中,分离的核酸是 cDNA。在一些实施方案中,分离的核酸进一步包含可操作地连接于编码营养性蛋白质的核酸序列的表达控制序列。在一些实施方案中,编码根据本公开的营养性蛋白质的核酸序列存在于载体中。在一些实施方案中,载体进一步包含可操作地连接于编码营养性蛋白质的核酸序列的表达控制序列。

[0034] 在另一方面,本公开提供包含根据本公开的核酸和载体的至少一者的重组微生物。在一些实施方案中,重组微生物是原核生物。在一些实施方案中,原核生物是异养性的。在一些实施方案中,原核生物是自养性的。在一些实施方案中,原核生物是细菌。

[0035] 在另一方面,本公开提供制备根据本公开的营养性蛋白质的方法,所述方法包括在足以由重组微生物产生所述营养性蛋白质的条件下培养根据本公开的重组微生物。在一些实施方案中,方法进一步包括自培养物分离营养性蛋白质。

[0036] 在另一方面,本公开提供包含本公开的分离的营养性蛋白质和至少一种第二组分的营养性组合物。在一些实施方案中,至少一种第二组分选自蛋白质、多肽、肽、游离氨基酸、碳水化合物、脂肪、矿物质或矿物质来源、维生素、补充剂、生物体、医药和赋形剂。在一些实施方案中,至少一种第二组分是蛋白质。在一些实施方案中,至少一种第二组分是营养性蛋白质。在一些实施方案中,至少一种第二组分是选自必需氨基酸、非必需氨基酸、支链氨基酸、非标准氨基酸和修饰的氨基酸的游离氨基酸。在一些实施方案中,至少一种第二组分是选自必需氨基酸的游离氨基酸。在一些实施方案中,至少一种第二组分是选自支链氨基酸的游离氨基酸。在一些实施方案中,至少一种第二组分是 Leu。在一些实施方案中,至少一种第二组分是脂质。在一些实施方案中,脂质选自脂肪、油、甘油三酯、胆固醇、磷脂和脂肪酸。在一些实施方案中,至少一种第二组分选自矿物质和维生素。在一些实施方案中,至少一种第二组分是补充剂。在一些实施方案中,至少一种第二组分是生物体。在一些实施方案中,至少一种第二组分是医药。在一些实施方案中,至少一种第二组分是赋形剂。在一些实施方案中,至少一种赋形剂选自缓冲剂、防腐剂、稳定剂、粘合剂、压实剂、润滑剂、分散增强剂、崩解剂、调味剂、甜味剂、着色剂。在一些实施方案中,将营养性组合物配制成液

体溶液、浆液、混悬液、凝胶剂、糊剂、粉末或固体。

[0037] 在另一方面，本公开提供制备本公开的营养性组合物的方法，其包括提供根据本公开的营养性蛋白质以及组合所述营养性蛋白质与至少一种第二组分。

[0038] 在另一方面，本公开提供维持或增加受试者的肌肉质量、肌肉强度和功能性表现的至少一者的方法，所述方法包括向所述受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中，受试者处于以下至少一种状态下：年长、患有医学上危急疾病、以及罹患蛋白质-能量营养不良症。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0039] 在另一方面，本公开提供维持或达成受试者的合乎需要的身体质量指数的方法，所述方法包括向所述受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中，受试者处于以下至少一种状态下：年长、患有医学上危急疾病、以及罹患蛋白质-能量营养不良症。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0040] 在另一方面，本公开提供向患有蛋白质-能量营养不良症的受试者提供蛋白质的方法，所述方法包括向所述受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中，受试者是肥胖的。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0041] 在另一方面，本公开提供增加受试者的产热的方法，所述方法包括向所述受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中，受试者是肥胖的。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0042] 在另一方面，本公开提供诱导受试者的饱足反应和饱腹感反应的至少一者的方法，所述方法包括向所述受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中，受试者是肥胖的。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0043] 在另一方面,本公开提供治疗受试者的恶病质、肌肉减少症和虚弱的至少一者的方法,所述方法包括向所述受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中,本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0044] 附图简述:

[0045] 图 1 显示二维直方图,其指示随溶合积分(y 轴)和聚集积分(x 轴)而变的蛋白质在大肠杆菌表达筛选中表达的相对可能性(在对数尺度上)。

[0046] 图 2 显示二维直方图,其指示随溶合积分(y 轴)和聚集积分(x 轴)而变的蛋白质在大肠杆菌表达筛选中可溶性表达的相对可能性(在对数尺度上)

[0047] 详述:

[0048] 除非本文另外定义,否则与本公开关联使用的科学和技术术语将具有由本领域普通技术人员通常所理解的含义。此外,除非另外为上下文所需,否则单数术语将包括复数且复数术语将包括单数。通常,与本文所述的生物化学、酶学、分子和细胞生物学、微生物学、遗传学、及蛋白质和核酸化学、以及杂交关联使用的命名法以及本文所述的生物化学、酶学、分子和细胞生物学、微生物学、遗传学、及蛋白质和核酸化学、以及杂交的技术是本领域中熟知且通常使用的命名法以及技术。本文引用的某些参考文献和其它文件以引用的方式明确并入本文。另外,本文引用的所有 UniProt/SwissProt 记录都据此以引用的方式并入本文。在起冲突的情况下,以包括定义的本说明书为准。材料、方法和实施例仅具有说明性而非意图具有限制性。

[0049] 除非另外指示,否则本公开的方法和技术通常是根据本领域中熟知及如本说明书整篇引用和论述的各种一般性和更特定参考文献中所述的常规方法来执行。参见例如 Sambrook 等, *Molecular Cloning: A Laboratory Manual*, 第 3 版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (2001); Ausubel 等, *Current Protocols in Molecular Biology*, Greene Publishing Associates (1992 及截至 2002 的补编); Taylor 和 Drickamer, *Introduction to Glycobiology*, Oxford Univ. Press (2003); Worthington Enzyme Manual, Worthington Biochemical Corp., Freehold, N. J.; *Handbook of Biochemistry: Section A Proteins*, 第 I 卷, CRC Press (1976); *Handbook of Biochemistry: Section A Proteins*, 第 II 卷, CRC Press (1976); *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press (1999)。可适用于蓝细菌的许多分子生物学和遗传技术描述于 Heidorn 等, "Synthetic Biology in Cyanobacteria: Engineering and Analyzing Novel Functions," *Methods in Enzymology*, 第 497 卷, 第 24 章 (2011) 中,其据此以引用的方式并入本文。

[0050] 本公开涉及在因特网上公开的针对某些蛋白质和基因序列的序列数据库条目(例如, UniProt/SwissProt)以及因特网上的其它信息。熟练技术人员了解的是因特网上的信息(包括序列数据库条目)时时更新,且例如用于指代特定序列的参考编号可变化。当提及序列信息的公开数据库或因特网上的其它信息时,应了解的是可发生所述变化且因特网上的信息的特定实施方案可时有时无。因为熟练技术人员可通过在因特网上搜索来得到等效信息,所以提及因特网网页地址或序列数据库条目证实所论述信息的可用性和公开传

播性。

[0051] 在公开和描述本发明蛋白质、组合物、方法和其它实施方案之前,应了解的是本文使用的术语是仅出于描述特定实施方案的目的且不意图具有限制性。必须注意的是,除非上下文另外明确规定,否则如说明书和随附权利要求中所用,单数形式“一”和“所述”包括复数个指示物。

[0052] 如本文所用的术语“包含”与“包括”或“含有”同义,且是包括性的或开放的,且不排除其它未叙述成员、要素或方法步骤。

[0053] 本公开提及氨基酸。氨基酸的全称可与各自的标准三字母和单字母缩写互换使用。为避免疑问,这些氨基酸是:丙氨酸 (Ala、A)、精氨酸 (Arg、R)、天冬酰胺 (Asn、N)、天冬氨酸 (Asp、D)、半胱氨酸 (Cys、C)、谷氨酸 (Glu、E)、谷氨酰胺 (Gln、Q)、甘氨酸 (Gly、G)、组氨酸 (His、H)、异亮氨酸 (Ile、I)、亮氨酸 (Leu、L)、赖氨酸 (Lys、K)、甲硫氨酸 (Met、M)、苯丙氨酸 (Phe、F)、脯氨酸 (Pro、P)、丝氨酸 (Ser、S)、苏氨酸 (Thr、T)、色氨酸 (Trp、W)、酪氨酸 (Tyr、Y)、缬氨酸 (Val、V)。

[0054] 如本文所用,术语“体外”是指事件发生在人工环境中,例如在试管或反应容器中、在细胞培养物中、在皮氏培养皿 (Petri dish) 中等,而非在生物体 (例如,动物、植物或微生物) 内。

[0055] 如本文所用,术语“体内”是指事件发生在生物体 (例如,动物、植物或微生物) 内。

[0056] 如本文所用,术语“分离的”是指物质或实体已 (1) 与它在最初产生时 (无论在自然界中或在实验环境中) 与之相伴的至少一些组分分离,和 / 或 (2) 通过人手产生、制备和 / 或制造。分离的物质和 / 或实体可与至少约 10%、约 20%、约 30%、约 40%、约 50%、约 60%、约 70%、约 80%、约 90%、或更多它们最初与之相伴的其它组分分离。在一些实施方案中,分离的试剂的纯度大于约 80%、约 85%、约 90%、约 91%、约 92%、约 93%、约 94%、约 95%、约 96%、约 97%、约 98%、约 99% 或大于约 99%。如本文所用,如果物质大致上不含其它组分,那么它是“纯的”。

[0057] 如本文所用,“支链氨基酸”是选自亮氨酸、异亮氨酸和缬氨酸的氨基酸。

[0058] 如本文所用,“必需氨基酸”是选自组氨酸、异亮氨酸、亮氨酸、赖氨酸、甲硫氨酸、苯丙氨酸、苏氨酸、色氨酸和缬氨酸的氨基酸。

[0059] 如本文所用的术语“肽”是指短多肽,例如通常含有小于约 50 个氨基酸且更通常小于约 30 个氨基酸的多肽。如本文所用的所述术语涵盖模拟结构功能以及因此生物功能的类似物和模拟物。

[0060] 术语“多肽”涵盖天然存在的蛋白质与非天然存在的蛋白质两者及其片段、突变体、衍生物和类似物。多肽可为单体或聚合物。此外,多肽可包含各自具有一种或多种不同活性的许多不同结构域。为避免疑问,“多肽”可具有大于两个氨基酸的任何长度。

[0061] 术语“分离的蛋白质”或“分离的多肽”是就它的起源或来源而言, (1) 不与在它的天然状态下伴随它的天然相伴组分相伴, (2) 以不见于自然界中的纯度存在,其中纯度可关于其它细胞物质的存在加以断定 (例如,不含来自相同物种的其它蛋白质) (3) 由来自不同物种的细胞表达,或 (4) 不存在于自然界中 (例如,它是见于自然界中的多肽的片段或它包括不见于自然界中的氨基酸类似物或衍生物或除标准肽键以外的键联) 的蛋白质或多肽。因此,化学合成或在不同于它所天然源于的细胞的细胞系统中合成的多肽将与它的天然相

伴组分“分离”。也可通过使用本领域中熟知的蛋白质纯化技术进行分离来致使多肽或蛋白质大致上不含天然相伴组分。如因此所定义,“分离”未必需要如此描述的蛋白质、多肽、肽或寡肽已自于其中合成它的细胞物理移除。

[0062] 如本文所用的术语“多肽片段”是指相较于全长多肽(如天然存在的蛋白质),具有缺失(例如氨基末端和/或羧基末端缺失)的多肽。在一实施方案中,多肽片段是连续序列,其中所述片段的氨基酸序列与天然存在的序列中的相应位置一致。片段的长度通常是至少 5、6、7、8、9 或 10 个氨基酸、或至少 12、14、16 或 18 个氨基酸、或至少 20 个氨基酸、或至少 25、30、35、40 或 45 个氨基酸、或至少 50 或 60 个氨基酸、或至少 70 个氨基酸。

[0063] 术语“融合蛋白”是指包含偶联于异源氨基酸序列的多肽或片段的多肽。融合蛋白是适用的,因为它们可被构建来含有两个或更多个可来自两种或更多种不同蛋白质的所需功能性元件。融合蛋白包含至少 10 个来自目标多肽的连续氨基酸、或至少 20 或 30 个氨基酸、或至少 40、50 或 60 个氨基酸、或至少 75、100 或 125 个氨基酸。融合蛋白内包括的异源多肽的长度通常是至少 6 个氨基酸、或至少 8 个氨基酸、或至少 15、20 或 25 个氨基酸。包括较大多肽(如 IgG Fc 区域)以及甚至整个蛋白质(如含有绿色荧光蛋白(“GFP”)发色团的蛋白质)的融合物具有特定效用。融合蛋白可通过与编码不同蛋白质或肽的核酸序列同框构建编码多肽或其片段的核酸序列,且接着表达融合蛋白来重组产生。或者,可通过使多肽或其片段交联于另一蛋白质来化学产生融合蛋白。

[0064] 如本文所用,如果编码蛋白质的核酸序列与编码第二蛋白质的核酸序列具有类似序列,那么所述蛋白质与所述第二蛋白质具有“同源性”或与所述第二蛋白质“同源”。或者,如果两个蛋白质具有类似氨基酸序列,那么某一蛋白质与第二蛋白质具有同源性。(因此,术语“同源蛋白质”被定义来指两个蛋白质具有类似氨基酸序列。)如本文所用,氨基酸序列的两个区域之间的同源性(尤其关于预测的结构类似性)解释为暗示在功能方面的类似性。

[0065] 当“同源”关于蛋白质或肽使用时,应认识到不相同的残基位置常因保守性氨基酸取代而不同。“保守性氨基酸取代”是其中某一氨基酸残基被具有类似化学性质(例如,电荷或疏水性)侧链(R 基团)的另一氨基酸残基取代的取代。一般来说,保守性氨基酸取代将不实质上改变蛋白质的功能性质。在两个或更多个氨基酸序列因保守性取代而不同于彼此的情况下,序列同一性百分比或同源性程度可向上调整以校正取代的保守性性质。用于进行这个调整的手段为本领域技术人员所熟知。参见例如 Pearson, 1994, *Methods Mol. Biol.* 24:307-31 和 25:365-89。

[0066] 以下六组各自含有是彼此的保守性取代的氨基酸:1) 丝氨酸、苏氨酸;2) 天冬氨酸、谷氨酸;3) 天冬酰胺、谷氨酰胺;4) 精氨酸、赖氨酸;5) 异亮氨酸、亮氨酸、甲硫氨酸、丙氨酸、缬氨酸和 6) 苯丙氨酸、酪氨酸、色氨酸。

[0067] 也称为序列同一性百分比的多肽的序列同源性通常是使用序列分析软件测量。参见例如威斯康星大学生物技术中心的遗传学计算机组(GCG)的序列分析软件包(910 University Avenue, Madison, Wis. 53705)。蛋白质分析软件使用对各种取代、缺失和其它修饰(包括保守性氨基酸取代)指定的同源性量度匹配类似序列。举例来说,GCG 含有可以缺省参数用于确定密切相关多肽(如来自生物体的不同物种的同源多肽)之间或野生型蛋白质与其突变蛋白之间的序列同源性或序列同一性的程序,如“Gap”和“Bestfit”。参

见例如 GCG6.1 版。

[0068] 当将特定多肽序列与含有来自不同生物体的许多序列的数据库比较时，一示例性算法是计算机程序 BLAST(Altschul 等, J. Mol. Biol. 215:403-410(1990); Gish 和 States, Nature Genet. 3:266-272(1993); Madden 等, Meth. Enzymol. 266:131-141(1996); Altschul 等, Nucleic Acids Res. 25:3389-3402(1997); Zhang 和 Madden, Genome Res. 7:649-656(1997)), 尤其 blastp 或 tblastn(Altschul 等, Nucleic Acids Res. 25:3389-3402(1997))。

[0069] 用于 BLASTp 的示例性参数是：期望值：10(缺省值)；过滤：seg(缺省值)；打开缺口的代价：11(缺省值)；延伸缺口的代价：1(缺省值)；最大比对：100(缺省值)；字长：11(缺省值)；描述数目：100(缺省值)；罚分矩阵：BLOWSUM62。针对同源性比较的多肽序列的长度将通常是至少约 16 个氨基酸残基、或至少约 20 个残基、或至少约 24 个残基、或至少约 28 个残基、或大于约 35 个残基。当搜索含有来自许多不同生物体的序列的数据库时，可适用的是比较氨基酸序列。使用氨基酸序列进行的数据库搜索可通过本领域中已知的除 blastp 以外的算法来测量。举例来说，可使用 GCG 6.1 版中的程序 FASTA 比较多肽序列。FASTA 提供在查询序列与搜索序列之间最佳重叠的区域的比对和序列同一性百分比。Pearson, Methods Enzymol. 183:63-98(1990)。举例来说，如以引用的方式并入本文的 GCG 6.1 版中所提供，可使用 FASTA 以它的缺省参数（字长 2 和 PAM250 计分矩阵）确定氨基酸序列之间的序列同一性百分比。

[0070] 在一些实施方案中，如果聚合分子（例如多肽序列或核酸序列）的序列至少 25%、至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 或至少 99% 同一，那么它们被视为彼此“同源”。在一些实施方案中，如果聚合分子的序列至少 25%、至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%、至少 60%、至少 65%、至少 70%、至少 75%、至少 80%、至少 85%、至少 90%、至少 95% 或至少 99% 类似，那么它们被视为彼此“同源”。术语“同源”必须是指至少两个序列（核苷酸序列或氨基酸序列）之间的比较。在一些实施方案中，如果两个核苷酸序列编码的多肽就至少一个具有至少约 20 个氨基酸的链段而言是至少约 50% 同一、至少约 60% 同一、至少约 70% 同一、至少约 80% 同一或至少约 90% 同一，那么所述两个核苷酸序列被视为同源。在一些实施方案中，同源核苷酸序列的特征在于能够编码具有至少 4-5 个独特指定的氨基酸的链段。这些氨基酸相对于彼此的同一性与大致间隔两者均必须为将核苷酸序列考虑为同源时所考虑。在长度小于 60 个核苷酸的核苷酸序列的一些实施方案中，同源性通过能够编码具有至少 4-5 个独特指定的氨基酸的链段来确定。在一些实施方案中，如果蛋白质就至少一个具有至少约 20 个氨基酸的链段而言是至少约 50% 同一、至少约 60% 同一、至少约 70% 同一、至少约 80% 同一或至少约 90% 同一，那么两个蛋白质序列被视为同源。

[0071] 如本文所用，“修饰的衍生物”是指在一级结构序列方面与参照多肽序列大致上同源，但包括例如体内或体外化学和生物化学修饰或并有不见于参照多肽中的氨基酸的多肽或其片段。所述修饰包括例如乙酰化、羧化、磷酸化、糖基化、泛素化、标记（例如用放射性核素标记）和各种酶修饰，如将易于为本领域技术人员所了解。多种用于标记多肽的方法和适用于所述目的取代基或标记在本领域中是熟知的，且包括放射性同位素

(如  $^{125}\text{I}$ 、 $^{32}\text{P}$ 、 $^{35}\text{S}$  和  $^3\text{H}$ )、结合标记的抗配体(例如,抗体)的配体、荧光团、化学发光剂、酶和可充当标记的配体的特异性结合对成员的抗配体。对标记的选择取决于所需灵敏性、与引物缀合的简易性、稳定性要求和可用仪器。用于标记多肽的方法在本领域中是熟知的。参见例如 Ausubel 等, *Current Protocols in Molecular Biology*, Greene Publishing Associates(1992 及截至 2002 的补编)。

[0072] 如本文所用,“多肽突变体”或“突变蛋白”是指相较于参照蛋白质或多肽(如天然或野生型蛋白质)的氨基酸序列,序列含有一个或多个氨基酸的插入、重复、缺失、重排或取代的多肽。突变蛋白可具有一个或多个氨基酸点取代,其中在某一位置处的单一氨基酸已变为另一氨基酸;一个或多个插入和/或缺失,其中一个或多个氨基酸分别被插入在参照蛋白质的序列中或在参照蛋白质的序列中缺失;和/或氨基酸序列在氨基或羧基末端的任一者或两者处的截短。突变蛋白相较于参照蛋白质可具有相同或不同生物活性。

[0073] 在一些实施方案中,突变蛋白与它的对应参照蛋白质具有例如至少 85% 总体序列同源性。在一些实施方案中,突变蛋白与野生型蛋白质具有至少 90% 总体序列同源性。在其它实施方案中,突变蛋白展现至少 95% 序列同一性、或 98% 或 99% 或 99.5% 或 99.9% 总体序列同一性。

[0074] 如本文所用,“用于亲和纯化的多肽标签”是具有可用于分离或纯化融合于第一“标签”多肽的目标第二蛋白质或多肽序列的结合配偶体的任何多肽。若干实例在本领域中是熟知的且包括 His-6 标签、FLAG 表位、c-myc 表位、Strep-TAGII、生物素标签、谷胱甘肽 5- 转移酶 (GST)、几丁质结合蛋白 (CBP)、麦芽糖结合蛋白 (MBP) 或金属亲和标签。

[0075] 如本文所用,“重组”是指例如基因或蛋白质的生物分子 (1) 已自它的天然存在的环境移除, (2) 不与在自然界中发现所述基因所处的多核苷酸的全部或部分相伴, (3) 可操作地连接于它在自然界中不与之连接的多核苷酸,或 (4) 不存在于自然界中。术语“重组”可关于克隆的 DNA 分离物、化学合成的多核苷酸类似物或由异源系统生物合成的多核苷酸类似物、以及由所述核酸编码的蛋白质和/或 mRNA 加以使用。因此,举例来说,例如如果由微生物合成的蛋白质是自由细胞中存在的重组基因合成的 mRNA 合成,那么它是重组的。

[0076] 术语“多核苷酸”、“核酸分子”、“核酸”或“核酸序列”是指长度是至少 10 个碱基的聚合形式的核苷酸。所述术语包括 DNA 分子(例如, cDNA 或基因组或合成 DNA) 和 RNA 分子(例如, mRNA 或合成 RNA)、以及 DNA 或 RNA 的含有非天然核苷酸类似物、非天然核苷间键或两者的类似物。核酸可呈任何拓扑构象。举例来说,核酸可为单链、双链、三链、四链体、部分双链、分支式、发夹式、环状或呈挂锁 (padlocked) 构象。

[0077] “合成”RNA、DNA 或混合聚合物是在细胞外部产生的 RNA、DNA 或混合聚合物,例如化学合成的 RNA、DNA 或混合聚合物。

[0078] 如本文所用的术语“核酸片段”是指相较于全长参照核苷酸序列,具有缺失(例如 5' 末端或 3' 末端缺失)的核酸序列。在一实施方案中,核酸片段是连续序列,其中所述片段的核苷酸序列与天然存在的序列中的相应位置一致。在一些实施方案中,片段的长度是至少 10、15、20 或 25 个核苷酸、或至少 20、30、40、50、60、70、80、90、100、110、120、130、140 或 150 个核苷酸。在一些实施方案中,核酸序列的片段是开放阅读框序列的片段。在一些实施方案中,所述片段编码由开放阅读框核苷酸序列编码的蛋白质的多肽片段(如本文所定义)。

[0079] 如本文所用,如果邻近于生物体的基因组中的内源性核酸序列(或那个序列编码的蛋白质产物)放置异源序列,以使这个内源性核酸序列的表达被改变,那么所述内源性核酸序列在本文中被视为“重组的”。在这个情形下,异源序列是不天然邻近于内源性核酸序列的序列,无论所述异源序列是否本身是内源性的(源于相同宿主细胞或其子代)或外源性的(源于不同宿主细胞或其子代)。举例来说,启动子序列可取代(例如,通过同源重组)宿主细胞的基因组中的基因的天然启动子,以使这个基因具有改变的表达样式。这个基因现将变为“重组的”,因为它与天然侧接它的至少一些序列分离。

[0080] 如果核酸含有不天然存在于基因组中的相应核酸中的任何修饰,那么它也被视为“重组的”。举例来说,如果内源性编码序列含有例如通过人为干预来人工引入的插入、缺失或点突变,那么它被视为“重组的”。“重组核酸”也包括在异源位点处整合至宿主细胞染色体中的核酸和以游离体形式存在的核酸构建体。

[0081] 如本文所用,短语参照核酸序列的“简并变体”涵盖可根据标准遗传密码翻译以提供与自所述参照核酸序列翻译的氨基酸序列一致的氨基酸序列的核酸序列。术语“简并寡核苷酸”或“简并引物”用于表示能够与在序列方面未必一致,但在一个或多个特定区段内彼此同源的靶标核酸序列杂交的寡核苷酸。

[0082] 术语“序列同一性百分比”或“同一”在核酸序列的情形下是指当比对以达成最大对应时,两个序列中的残基是相同的。序列同一性比较的长度可为历经至少约 9 个核苷酸、通常至少约 20 个核苷酸、更通常至少约 24 个核苷酸、通常至少约 28 个核苷酸、更通常至少约 32 个、且甚至更通常至少约 36 个或更多个核苷酸的链段。存在本领域中已知的可用于测量核苷酸序列同一性的许多不同算法。举例来说,可使用威斯康星程序包 10.0 版 (Genetics Computer Group (GCG), Madison, Wis) 中的程序 FASTA、Gap 或 Bestfit 比较多核苷酸序列。FASTA 提供在查询序列与搜索序列之间最佳重叠的区域的比对和序列同一性百分比。Pearson, *Methods Enzymol.* 183:63-98 (1990)。举例来说,如以引用的方式并入本文的 GCG 6.1 版中所提供,可使用 FASTA 以它的缺省参数(字长 6 和用于计分矩阵的 NOPAM 因子)或使用 Gap 以它的缺省参数确定核酸序列之间的序列同一性百分比。或者,可使用计算机程序 BLAST (Altschul 等, *J. Mol. Biol.* 215:403-410 (1990); Gish 和 States, *Nature Genet.* 3:266-272 (1993); Madden 等, *Meth. Enzymol.* 266:131-141 (1996); Altschul 等, *Nucleic Acids Res.* 25:3389-3402 (1997); Zhang 和 Madden, *Genome Res.* 7:649-656 (1997)), 尤其 blastp 或 tblastn (Altschul 等, *Nucleic Acids Res.* 25:3389-3402 (1997)) 比较序列。

[0083] 术语“实质性同源性”或“实质性类似性”在关于核酸或其片段时指示当在适当核苷酸插入或缺失下与另一核酸(或它的互补链)最优比对时,在至少约 76%、80%、85%、或至少约 90%、或至少约 95%、96%、97%、98% 或 99% 的核苷酸碱基中存在核苷酸序列同一性,如通过序列同一性的任何熟知算法,如如上所论述的 FASTA、BLAST 或 Gap 所测量。

[0084] 或者,当核酸或其片段与另一核酸、另一核酸的一条链、或其互补链在严格杂交条件下杂交时,存在实质性同源性或类似性。“严格杂交条件”和“严格洗涤条件”在核酸杂交实验的情形下取决于许多不同物理参数。核酸杂交将受如以下的条件影响:盐浓度、温度、溶剂、杂交物质的碱基组成、互补区域的长度和杂交核酸之间的核苷酸碱基错配数目,如将易于为本领域技术人员所了解。本领域普通技术人员了解如何改变这些参数来达成特定杂



交严格性。

[0085] 一般来说,特定 DNA 杂交物的“严格杂交”是在约 25℃下在低于热熔点 ( $T_m$ ) 下在一组特定条件下进行。特定 DNA 杂交物的“严格洗涤”是在温度约 5℃下在低于  $T_m$  下在一组特定条件下进行。 $T_m$  是 50%靶标序列与完全匹配探针杂交所处的温度。参见 Sambrook 等, *Molecular Cloning: A Laboratory Manual*, 第 2 版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989), 第 9.51 页, 据此以引用的方式并入本文。出于本文目的,溶液相杂交的“严格条件”被定义为在 65℃下在 6xSSC(其中 20xSSC 含有 3.0M NaCl 和 0.3M 柠檬酸钠)、1% SDS 中进行水性杂交(即不含甲酰胺)8-12 小时,随后在 65℃下在 0.2xSSC、0.1% SDS 中进行 2 次 20 分钟洗涤。熟练工作者应了解的是在 65℃下杂交将取决于许多因素而在不同速率下发生,所述因素包括杂交的序列的长度和同一性百分比。

[0086] 如本文所用,“表达控制序列”是指为实现它们所可操作地连接的编码序列的表达所必需的多核苷酸序列。表达控制序列是控制核酸序列的转录、转录后事件和翻译的序列。表达控制序列包括适当转录起始、终止、启动子和增强子序列;高效 RNA 加工信号,如剪接和聚腺苷酸化信号;使细胞质 mRNA 稳定的序列;增强翻译效率的序列(例如,核糖体结合位点);增强蛋白质稳定性的序列;和当需要时,增强蛋白质分泌的序列。所述控制序列的性质视宿主生物体而不同;在原核生物中,所述控制序列通常包括启动子、核糖体结合位点和转录终止序列。术语“控制序列”意图至少涵盖其存在为表达所必需的任何组分,且也可涵盖其存在是有利的其它组分,例如前导序列和融合配偶体序列。

[0087] 如本文所用,“可操作地连接的”表达控制序列是指以下连接:表达控制序列与目标基因邻接以控制所述目标基因,以及表达控制序列反式作用或在一定距离下控制所述目标基因的作用。

[0088] 如本文所用,“载体”意图指代能够转运它已连接的另一核酸的核酸分子。一种载体类型是“质粒”,其通常是指可向其中连接其它 DNA 区段的环状双链 DNA 环,但也包括线性双链分子,如通过聚合酶链式反应 (PCR) 由扩增或由用限制酶处理环状质粒产生的线性双链分子。其它载体包括粘粒、细菌人工染色体 (BAC) 和酵母人工染色体 (YAC)。另一载体类型是病毒载体,其中可将其它 DNA 区段连接至病毒基因组中(以下更详细论述)。某些载体能够在它们所引入的宿主细胞中自主复制(例如,具有在宿主细胞中起作用的复制起点的载体)。其它载体可在引入宿主细胞中后整合至所述宿主细胞的基因组中,且由此与宿主基因组一起复制。此外,某些载体能够指导它们所可操作地连接的基因的表达。所述载体在本文中称为“重组表达载体”(或简称为“表达载体”)。

[0089] 如本文所用的术语“重组宿主细胞”(或简称为“重组细胞”或“宿主细胞”)意图指代已引入有如重组载体的重组核酸的细胞。在一些情况下,用词“细胞”由指定细胞类型的名称替代。举例来说,“重组微生物”是为微生物宿主细胞的重组宿主细胞且“重组蓝细菌”是作为蓝细菌宿主细胞的重组宿主细胞。。应了解的是所述术语意图不仅指代特定受试细胞,而且指代所述细胞的子代。因为某些修饰可由于突变或环境影响而发生在继代中,所以所述子代可能实际上不与亲本细胞一致,但仍然包括在如本文所用的术语“重组宿主细胞”、“重组细胞”和“宿主细胞”的范围内。重组宿主细胞可为生长在培养物中的分离的细胞或细胞系或可为存在于活组织或生物体中的细胞。

[0090] 如本文所用,术语“异养性”是指生物体不能固定碳且使用有机碳来生长。

[0091] 如本文所用,术语“自养性”是指生物体使用来自光(通过光合作用)或无机化学反应(化学合成)的能量自简单无机分子产生复杂有机化合物(如碳水化合物、脂肪和蛋白质)。

[0092] 如本文所用,“肌肉质量”是指受试者的身体中的肌肉重量。肌肉质量包括骨骼肌、平滑肌(如心脏和消化道肌肉)和这些肌肉中含有的水。特定肌肉的肌肉质量可使用双能x射线吸光测定法(DEXA)(Padden-Jones等,2004)来测定。总瘦体重(减去脂肪)、总身体质量和骨矿物质含量也可通过DEXA来测量。在一些实施方案中,受试者的特定肌肉的肌肉质量的变化是例如通过DEXA来测定,且所述变化用作所述受试者的肌肉质量的总变化的代表。因此,举例来说,如果受试者消耗如本文公开的营养性蛋白质且历经一段时期经历特定肌肉或肌肉群中的肌肉质量增加,那么可推断所述受试者已经历肌肉质量增加。

[0093] 如本文所用,“肌肉强度”是指在单次最大出力下,肌肉可产生的力的量。存在两种类型的肌肉强度,即静态强度和动态强度。静态强度是指肌肉的等长收缩,其中肌肉产生力量,而肌肉长度保持恒定和/或此时不存在关节运动。实例包括固持或携带物件或相对于墙壁推进。动态强度是指肌肉产生导致运动的力量。动态强度可为等张收缩,其中肌肉在恒定负荷或等动力收缩下缩短,其中肌肉在恒定速度下收缩且缩短。动态强度也可包括等惯性强度。

[0094] 除非指定,否则“肌肉强度”是指最大动态肌肉强度。最大强度称为“最大重复一次”(1RM)。这是可在不失败或损伤下完全移动(提升、推进或牵拉)最大负荷(以千克计)一次的测量结果。这个值可直接测量,但这样做需要增加重量直至受试者未能将活动进行完毕。或者,1RM是通过使用小于受试者可移动的最大量的负荷,对受试者可达成的最大运动重复次数进行计数来估计。常在临床试验中测量腿部伸展和腿部弯曲(Borsheim等,“Effect of amino acid supplementation on muscle mass, strength and physical function in elderly,”Clin Nutr 2008 ;27:189-195 ;Paddon-Jones,等,“Essential amino acid and carbohydrate supplementation ameliorates muscle protein loss in humans during 28days bed rest,”J Clin Endocrinol Metab 2004 ;89:4351-4358)。

[0095] 如本文所用,“功能性表现”是指模拟每日活动的功能性测试。“功能性表现”是通过任何适合接受测试来测量,所述测试包括定时举步测试(自4英寸台阶尽可能快地上下举步5次)、定时平台传送测试(在平台上尽可能快地自立位变为仰卧位且此后再次向上变为立位,重复一次)、和身体表现连串测试(静态平衡测试、坐椅测试和行走测试)(Borsheim等,“Effect of amino acid supplementation on muscle mass, strength and physical function in elderly,”Clin Nutr 2008 ;27:189-195)。

[0096] 如本文所用,“身体质量指数”或“BMI”或“克托莱指数(Quetelet index)”是以千克计的受试者的重量除以以米计的受试者的身高的平方( $\text{kg}/\text{m}^2$ )。

[0097] 对于成人,常使用BMI来评估个体的体重偏离具有他的或她的身高的人士的正常或合乎需要体重的程度。重量过度或不足可部分地用身体脂肪来解释,但如肌肉发达性的其它因素也显著影响BMI。世界卫生组织将BMI小于18.5视为重量不足且可指示营养不良、饮食病症或其它健康问题,而BMI大于25被视为超重且高于30被视为肥胖。(世界卫生组织。BMI分类。2012年3月19日访问[http://apps.who.int/bmi/index.jsp?introPage=intro\\_3.html](http://apps.who.int/bmi/index.jsp?introPage=intro_3.html)。)如本文所用,“合乎需要的身体质量指数”是身体质量指数自约18.5至

约 25。因此,如果受试者的 BMI 低于约 18.5,那么所述受试者的 BMI 增加是所述受试者的 BMI 的合乎需要增加。如果相反,受试者的 BMI 高于约 25,那么所述受试者的 BMI 降低是所述受试者的 BMI 的合乎需要增加。

[0098] 如本文所用,“年长”哺乳动物是经历身体质量指数和肌肉质量的至少一者的年龄相关变化(例如,年龄相关肌肉减少症)的哺乳动物。在一些实施方案中,“年长”人是至少 50 岁、至少 60 岁、至少 65 岁、至少 70 岁、至少 75 岁、至少 80 岁、至少 85 岁、至少 90 岁、至少 95 岁或至少 100 岁。在一些实施方案中,年长动物、哺乳动物或人是已经历肌肉质量自峰值寿命肌肉质量损失至少 5%、至少 10%、至少 15%、至少 20%、至少 25%、至少 30%、至少 35%、至少 40%、至少 45%、至少 50%、至少 55%或至少 60%的人。因为已知身体质量指数和肌肉质量的至少一者的年龄相关变化与年龄增加相关,所以在一些实施方案中,仅基于年龄鉴定或确定年长哺乳动物。因此,在一些实施方案中,“年长”人是通过他们的年龄是至少 60 岁、至少 65 岁、至少 70 岁、至少 75 岁、至少 80 岁、至少 85 岁、至少 90 岁、至少 95 岁或至少 100 岁的事实,而不依靠身体质量指数和肌肉质量的至少一者的测量结果来简单地鉴定或确定。

[0099] 如本文所用,如果患者由于医学疾病而经历身体质量指数和肌肉质量的至少一者的变化(例如,肌肉减少症),那么所述患者患有“医学上危急疾病”。在一些实施方案中,患者持续他们醒来时间的至少 25%、至少 50%、至少 60%、至少 70%、至少 80%、至少 90%、至少 95%或 100%被限制卧床。在一些实施方案中,患者是无意识的。在一些实施方案中,患者已如这个段落中所述被限制卧床至少 1 天、2 天、3 天、4 天、5 天、10 天、2 周、3 周、4 周、5 周、10 周或更长。

[0100] 如本文所用,“蛋白质-能量营养不良症”是指一种形式的营养不良,其中存在蛋白质摄取不充分。类型包括夸休可尔症(Kwashiorkor)(突出性蛋白质营养不良)、消瘦(缺乏卡路里与蛋白质营养物两者)和消瘦性夸休可尔症(存在显著蛋白质缺乏和显著卡路里不足征象,有时称为最重度形式的营养不良)。

[0101] 如本文所用,“恶病质”是指导致损耗和重量减轻的多层面临床综合征。它是一种复杂病状,其中蛋白质分解代谢超过蛋白质合成代谢,此使得肌肉损耗成为所述病状的主要特征。除蛋白质代谢的代谢紊乱之外,它的特征也在于厌食和发炎。这些紊乱加上蛋白质代谢受损在不同程度上对营养物疗法起反应。

[0102] 如本文所用,“产热”是哺乳动物中的热产生过程。产热伴随有能量消耗增加。产热具体来说是在食物组分(如蛋白质)代谢之后燃烧能量。这也可称为食物的热效应。由个体消耗的总能量等于静息能量消耗(在空腹状态下在静息时消耗来支撑基础代谢的能量)、食物的热效应和与身体活动相关的能量消耗的总和。静息能量消耗占人中总能量消耗的约 65-75%。肌肉质量的量和活性是静息能量消耗的一种影响因素。足以支撑肌肉的蛋白质消耗也影响静息能量消耗。蛋白质的摄取倾向于增加餐后能量消耗;这是食物的热效应。食物的热效应占人中总能量消耗的约 10%。尽管这是总能量消耗的较小比例,但这个值的较小增加可影响体重。蛋白质具有的热效应高于脂肪或碳水化合物;这个效应连同蛋白质的其它代谢影响一起使得蛋白质成为适用于重量控制、糖尿病管理和其它情况的基质。

[0103] 如本文所用,“饱足”是当进食时变饱或对进食的需求降低的行为。这会停止或减

少进食。

[0104] 如本文所用,“饱腹感”是餐后保持饱满的行为,其表示为餐后不进食的时期。

[0105] 如本文所用,最广泛而言,“运动”是增强或维持身体康健以及总体健康和安康的任何身体活动。运动是出于各种原因来进行,包括强化肌肉和心血管系统、磨练运动技能、减轻或维持重量以及出于享乐的目的。

[0106] 如本文所用,“足量”是本文公开的蛋白质或多肽足以产生所需作用的量。举例来说,如果需要肌肉质量增加,那么足量是历经一段时期导致肌肉质量增加的量。足量蛋白质或多肽片段可直接(即通过向受试者施用所述蛋白质或多肽片段)提供,或它可作为包含所述蛋白质或多肽片段的组合物的一部分提供。在本文中其它地方论述施用模式。

[0107] 如本文所用,术语“哺乳动物”是指分类类别哺乳纲的任何成员,包括胎盘哺乳动物和有袋哺乳动物。因此,“哺乳动物”包括人、灵长类动物、家畜和实验室哺乳动物。示例性哺乳动物包括啮齿动物、小鼠、大鼠、兔、狗、猫、绵羊、马、山羊、美洲驼、牛、灵长类动物、猪和任何其它哺乳动物。在一些实施方案中,哺乳动物是转基因哺乳动物、遗传工程化哺乳动物和克隆哺乳动物中的至少一者。

[0108] A. 营养性蛋白质

[0109] 出于本公开的目的,“营养性蛋白质”是含有合乎需要量的必需氨基酸的蛋白质。在一些实施方案中,以重量计,营养性蛋白质包含至少 30%必需氨基酸。在一些实施方案中,以重量计,营养性蛋白质包含至少 40%必需氨基酸。在一些实施方案中,以重量计,营养性蛋白质包含至少 50%必需氨基酸。在一些实施方案中,营养性蛋白质包含以下或由以下组成:天然存在于可食用物质中的蛋白质或蛋白质的片段。“可食用物质”以它的最广泛意义涵盖已知由至少一种类型的哺乳动物进食而无有害作用的任何物质。有害作用包括有毒作用和毒性作用。在一些实施方案中,可食用物质是已知由人进食而无有害作用的物质。一些可食用物质是处于有限地理位置的仅一小组某一类型的哺乳动物的膳食的不常见但已知的组分,而其它可食用物质是遍及世界上大部分区域的膳食主要成分。在其它实施方案中,可食用物质是不知先前由任何哺乳动物进食,但在测试后被证明可食用的物质。可食用物质包括但不限于特伦纳棉 (*Gossypium turneri*)、白黄侧耳 (*Pleurotus cornucopiae*)、大豆 (*Glycine max*)、水稻 (*Oryza sativa*)、大眼金枪鱼 (*Thunnus obesus*)、苞片冷杉 (*Abies bracteata*)、红腹刺鼠 (*Acomys ignitus*)、叶轴香豌豆 (*Lathyrus aphaca*)、印度野牛 (*Bos gaurus*)、黑耳岩羚 (*Raphicerus melanotis*)、格陵兰海豹 (*Phocagroenlandica*)、中华鲟 (*Acipenser sinensis*)、马来灵猫 (*Viverra zangalunga*)、漏斗状侧耳 (*Pleurotus sajor-caju*)、苦荞麦 (*Fagopyrum tataricum*)、美国五针松 (*Pinus strobus*)、大花牵牛 (*Ipomoea nil*)、东北红豆杉 (*Taxus cuspidata*)、槭叶小牵牛 (*Ipomoea wrightii*)、砂海螂 (*Mya arenaria*)、美味猕猴桃 (*Actinidia deliciosa*)、戈氏瞪羚 (*Gazella granti*)、欧洲山杨 (*Populus tremula*)、西洋李 (*Prunus domestica*)、银鸥 (*Larus argentatus*)、毛野豌豆 (*Vicia villosa*)、乳斑棘鳞鱼 (*Sargocentron punctatissimum*)、叉枝蝇子草 (*Silene latifolia*)、霍氏海豚 (*Lagenodelphis hosei*)、大西洋浪蛤 (*Spisula solidissima*)、暗长毛獐 (*Crossarchus obscurus*)、赤豆 (*Phaseolus angularis*)、*Lathyrus vestitus*、细鳞大麻哈鱼 (*Oncorhynchus gorbuscha*)、密西西比短吻鳄 (*Alligator mississippiensis*)、哈列朋松 (*Pinus halepensis*)、海鸥 (*Larus canus*)、芜菁甘蓝 (*Brassica napus*)、麦瓶草

(*Silene cucubalus*)、带纹海豹 (*Phoca fasciata*)、印度瞪羚 (*Gazella bennettii*)、火炬松 (*Pinus taeda*)、加拿大红豆杉 (*Taxus canadensis*)、鳞秕泽米 (*Zamia furfuracea*)、云南松 (*Pinus yunnanensis*)、乔松 (*Pinus wallichiana*)、石刁柏 (*Asparagus officinalis*)、风铃辣椒 (*Capsicum baccatum*)、大盆地刺果松 (*Pinus longaeva*)、浆果紫杉 (*Taxus baccata*)、西伯利亚红松 (*Pinus sibirica*)、甜橙 (*Citrus sinensis*)、条纹金鳞鱼 (*Sargocentron xantherythrum*)、美洲野牛 (*Bison bison*)、汤姆森瞪羚 (*Gazella thomsonii*)、救荒野豌豆 (*Vicia sativa*)、加拿大黑雁 (*Branta canadensis*)、芹菜 (*Apium graveolens*)、栓皮槭 (*Acer campestre*)、芫荽 (*Coriandrum sativum*)、圆锥麦瓶草 (*Silene conica*)、莴苣 (*Lactuca sativa*)、黄灯笼辣椒 (*Capsicum chinense*)、白桧 (*Abies veitchii*)、家山羊 (*Capra hircus*)、斯氏瞪羚 (*Gazella spekei*)、大麻哈鱼 (*Oncorhynchus keta*)、小心叶薯 (*Ipomoea obscura*)、菜瓜 (*Cucumis melo* var. *conomon*)、环斑海豹 (*Phoca hispida*)、火狐 (*Vulpes vulpes*)、羽叶茛苣 (*Ipomoea quamoclit*)、多毛番茄 (*Solanum habrochaites*)、杨属的种 (*Populus* sp.)、刚松 (*Pinus rigida*)、提琴叶栎 (*Quercus lyrata*)、多花菜豆 (*Phaseolus coccineus*)、红嘴鸥 (*Larus ridibundus*)、尖吻棘鳞鱼 (*Sargocentron spiniferum*)、金枪鱼 (*Thunnus thynnus*)、北极狐 (*Vulpes lagopus*)、大额印度野牛 (*Bos gaurus frontalis*)、意大利槭 (*Acer opalus*)、鸡爪槭 (*Acer palmatum*)、石栎 (*Quercus ilex*)、中欧山松 (*Pinus mugo*)、赤颈鹤 (*Grus antigone*)、山地松 (*Pinus uncinata*)、乌梅 (*Prunus mume*)、大鳞大马哈鱼 (*Oncorhynchus tshawytscha*)、鹅喉羚 (*Gazella subgutturosa*)、耳廓狐 (*Vulpes zerda*)、大果松 (*Pinus coulteri*)、海岛棉 (*Gossypium barbadense*)、欧亚槭 (*Acer pseudoplatanus*)、红大麻哈鱼 (*Oncorhynchus nerka*)、须野猪 (*Sus barbatus*)、荞麦祖先亚种 (*Fagopyrum esculentum* subsp. *Ancestrale*)、刺苞菜蓟 (*Cynara cardunculus*)、绿豆 (*Phaseolus aureus*)、黑杨 (*Populus nigra*)、施温迪茫棉 (*Gossypium schwendimanii*)、野生马铃薯 (*Solanum chacoense*)、赤栎 (*Quercus rubra*)、黄瓜 (*Cucumis sativus*)、普通斑马 (*Equus burchelli*)、银士麻哈鱼 (*Oncorhynchus kisutch*)、辐射松 (*Pinus radiata*)、港海豹 (*Phoca vitulina richardsi*)、黑颈鹤 (*Grus nigricollis*)、大冷杉 (*Abies grandis*)、孟苏大麻哈鱼 (*Oncorhynchus masou*)、菠菜 (*Spinacia oleracea*)、智利番茄 (*Solanum chilense*)、旋角羚 (*Addax nasomaculatus*)、甘薯 (*Ipomoea batatas*)、格列威斑马 (*Equus grevyi*)、库页冷杉 (*Abies sachalinensis*)、意大利石松 (*Pinus pinea*)、康氏蹄蝠 (*Hipposideros commersoni*)、走茎番红花 (*Crocus nudiflorus*)、柚 (*Citrus maxima*)、美洲鲟 (*Acipenser transmontanus*)、拟似棉 (*Gossypium gossypoides*)、麝香猫 (*Viverra zibetha*)、西里栎 (*Quercus cerris*)、斑头雁 (*Anser indicus*)、狐尾松 (*Pinus balfouriana*)、黄雪轮 (*Silene otites*)、大麻哈鱼属的种 (*Oncorhynchus* sp.)、大斑灵猫 (*Viverra megaspila*)、野牦牛 (*Bos mutus grunniens*)、湿地松 (*Pinus elliottii*)、土库曼野驴 (*Equus hemionus kulan*)、羴羊 (*Capra ibex ibex*)、大蒜 (*Allium sativum*)、萝卜 (*Raphanus sativus*)、萌芽松 (*Pinus echinata*)、野黑樱 (*Prunus serotina*)、白纹鳃 (*Sargocentron diadema*)、匙叶麦瓶草 (*Silene gallica*)、甘蓝 (*Brassica oleracea*)、胡萝卜 (*Daucus carota*)、虹鳟 (*Oncorhynchus mykiss*)、芥蓝 (*Brassica oleracea* var. *alboglabra*)、陆地棉 (*Gossypium hirsutum*)、欧洲冷杉 (*Abies alba*)、柑桔 (*Citrus*

reticulata)、菊苣 (*Cichorium intybus*)、柬埔寨野牛 (*Bos sauveli*)、马驼 (*Lama glama*)、玉米 (*Zea mays*)、石菖蒲 (*Acorus gramineus*)、赤狐 (*Vulpes macrotis*)、盘羊蒙古亚种 (*Ovis ammon darwini*)、沙氏石羚 (*Raphicerus sharpei*)、扭叶松 (*Pinus contorta*)、瘤牛 (*Bos indicus*)、西伯利亚北山羊 (*Capra sibirica*)、西黄松 (*Pinus ponderosa*)、苦杏仁 (*Prunus dulcis*)、马铃薯 (*Solanum soganandinum*)、蕹菜 (*Ipomoea aquatica*)、白喙斑纹海豚 (*Lagenorhynchus albirostris*)、加拿大盘羊 (*Ovis canadensis*)、欧洲甜樱桃 (*Prunus avium*)、鹿瞪羚 (*Gazella dama*)、长鳍金枪鱼 (*Thunnus alalunga*)、白花蝇子草 (*Silene pratensis*)、瑞士松 (*Pinus cembra*)、番红花 (*Crocus sativus*)、西瓜 (*Citrullus lanatus*)、红额瞪羚 (*Gazella rufifrons*)、亚洲芥 (*Brassica tournefortii*)、捻角山羊 (*Capra falconeri*)、民都洛水牛 (*Bubalus mindorensis*)、长叶松 (*Pinus palustris*)、月桂樱 (*Prunus laurocerasus*)、白枕鹤 (*Grus vipio*)、牵牛花 (*Ipomoea purpurea*)、平滑叶松 (*Pinus leiophylla*)、暗色斑纹海豚 (*Lagenorhynchus obscurus*)、小岩羚 (*Raphicerus campestris*)、大白菜 (*Brassica rapa* subsp. *Pekinensis*)、*Acmella radicans*、三裂叶薯 (*Ipomoea triloba*)、展叶松 (*Pinus patula*)、甜瓜 (*Cucumis melo*)、矮松 (*Pinus virginiana*)、番茄 (*Solanum lycopersicum*)、赤松 (*Pinus densiflora*)、大叶松 (*Pinus engelmannii*)、欧洲白栎 (*Quercus robur*)、刺毛月光花 (*Ipomoea setosa*)、红平菇 (*Pleurotus djamor*)、冕蹄蝠 (*Hipposideros diadema*)、绵羊 (*Ovis aries*)、长棘鳞鱼 (*Sargocentron microstoma*)、花茎甘蓝 (*Brassica oleracea* var. *italica*)、东高加索羴羊 (*Capra cylindricornis*)、杂交白杨 (*Populus kitakamiensis*)、编织葱 (*Allium textile*)、蚕豆 (*Vicia faba*)、荞麦 (*Fagopyrum esculentum*)、草原野牛 (*Bison priscus*)、栓皮栎 (*Quercus suber*)、*Lagophylla ramosissima*、马达加斯加地蚶 (*Acrantophis madagascariensis*)、西伯利亚鲟 (*Acipenser baerii*)、樱桃辣椒 (*Capsicum annum*)、小麦 (*Triticum aestivum*)、爪蟾 (*Xenopus laevis*)、西伯利亚海豹 (*Phoca sibirica*)、纳氏鲟 (*Acipenser naccarii*)、中华猕猴桃 (*Actinidia chinensis*)、白大角羊 (*Ovis dalli*)、马铃薯 (*Solanum tuberosum*)、菲律宾水牛 (*Bubalus carabanensis*)、粗柠檬 (*Citrus jambhiri*)、欧洲野牛 (*Bison bonasus*)、非洲野驴 (*Equus asinus*)、低地倭水牛 (*Bubalus depressicornis*)、刺芹侧耳 (*Pleurotus eryngii*)、矮茄 (*Solanum demissum*)、赤羊 (*Ovis vignei*)、野生玉米 (*Zea mays* subsp. *Parviglumis*)、丹吉尔山黧豆 (*Lathyrus tingitanus*)、百岁兰 (*Welwitschia mirabilis*)、澳洲鹤 (*Grus rubicunda*)、橙红茑萝 (*Ipomoea coccinea*)、洋葱 (*Allium cepa*)、索曼莫格齐瞪羚 (*Gazella soemmerringii*)、芜菁 (*Brassica rapa*)、瘦驼 (*Lama vicugna*)、秘鲁番茄 (*Solanum peruvianum*)、北欧爪蟾 (*Xenopus borealis*)、高加索山羊 (*Capra caucasica*)、黄鳍鲂 (*Thunnus albacares*)、山斑马 (*Equus zebra*)、原鸡 (*Gallus gallus*)、野生墨西哥土豆 (*Solanum bulbocastanum*)、台湾叶鼻蝠 (*Hipposideros terasensis*)、白腰斑纹海豚 (*Lagenorhynchus acutus*)、河马 (*Hippopotamus amphibius*)、红松 (*Pinus koraiensis*)、蒙皮利埃槭 (*Acer monspessulanum*)、三角叶杨 (*Populus deltoides*)、毛果杨 (*Populus trichocarpa*)、俄罗斯鲟 (*Acipenser guldenstadti*)、黑松 (*Pinus thunbergii*)、结球甘蓝 (*Brassica oleracea* var. *capitata*)、柯氏渊杜父鱼 (*Abyssocottus korotneffi*)、居维叶瞪羚 (*Gazella cuvieri*)、日光冷杉 (*Abies*

homolepis)、辽东冷杉 (*Abies holophylla*)、山瞪羚 (*Gazella gazella*)、五针松 (*Pinus parviflora*)、羽衣甘蓝 (*Brassica oleracea* var. *acephala*)、西葫芦 (*Cucurbita pepo*)、华山松 (*Pinus armandii*)、大白叶冷杉 (*Abies mariesii*)、太平洋黑鲔 (*Thunnus thynnus orientalis*)、温州蜜柑 (*Citrus unshiu*)、契斯曼尼番茄 (*Solanum cheesmanii*)、太平洋短吻海豚 (*Lagenorhynchus obliquidens*)、挪威槭 (*Acer platanoides*)、柠檬 (*Citrus limon*)、杜氏蚬 (*Acrantophis dumerili*)、科民茄 (*Solanum commersonii*)、木本棉 (*Gossypium arboreum*)、桃 (*Prunus persica*)、糙皮侧耳 (*Pleurotus ostreatus*)、日本冷杉 (*Abies firma*)、细角羚 (*Gazella leptoceros*)、盐湖鲑鱼 (*Salmo salar*)、美洲海螯虾 (*Homarus americanus*)、紫果冷杉 (*Abies magnifica*)、爪哇野牛 (*Bos javanicus*)、环斑海豹 (*Phoca largha*)、卷毛野猪 (*Sus cebifrons*)、茄子 (*Solanum melongena*)、海豹 (*Phoca vitulina*)、欧洲赤松 (*Pinus sylvestris*)、佛州泽米 (*Zamia floridana*)、沙狐 (*Vulpes corsac*)、韭葱 (*Allium porrum*)、里海海豹 (*Phoca caspica*)、南非狐 (*Vulpes chama*)、中国红豆杉 (*Taxus chinensis*)、花椰菜 (*Brassica oleracea* var. *botrytis*)、灰雁 (*Anser anser anser*)、金甲豆 (*Phaseolus lunatus*)、油菜 (*Brassica campestris*)、糖槭 (*Acer saccharum*)、偃松 (*Pinus pumila*)、潘那利番茄 (*Solanum pennellii*)、可食松 (*Pinus edulis*)、南方牵牛 (*Ipomoea cordatotriloba*)、银白杨 (*Populus alba*)、克拉克大麻哈鱼 (*Oncorhynchus clarki*)、无梗花栎 (*Quercus petraea*)、多疣野猪 (*Sus verrucosus*)、野生蒙古野马 (*Equus caballus przewalskii*)、胡杨 (*Populus euphratica*)、热带爪蟾 (*Xenopus tropicalis*)、短叶紫杉 (*Taxus brevifolia*)、原驼 (*Lama guanicoe*)、北美短叶松 (*Pinus banksiana*)、龙葵 (*Solanum nigrum*)、印尼野猪 (*Sus celebensis*)、芥菜 (*Brassica juncea*)、大西洋斑纹海豚 (*Lagenorhynchus cruciger*)、颤杨 (*Populus tremuloides*)、刺针松 (*Pinus pungens*)、西里伯斯野水牛 (*Bubalus quarlesi*)、*Quercus gamelliflora*、欧洲盘羊 (*Ovis orientalis musimon*)、印度水牛 (*Bubalus bubalis*)、琉球松 (*Pinus luchuensis*)、菲律宾野猪 (*Sus philippensis*)、菜豆 (*Phaseolus vulgaris*)、海鲱 (*Salmo trutta*)、波斯鲟 (*Acipenser persicus*)、野生种马铃薯 (*Solanum brevidens*)、多脂松 (*Pinus resinosa*)、黑马羚 (*Hippotragus niger*)、努比亚獐羊 (*Capra nubiana*)、*Asparagus scaber*、旋转牵牛 (*Ipomoea platensis*)、野猪 (*Sus scrofa*)、野山羊 (*Capra aegagrus*)、山魃豆 (*Lathyrus sativus*)、赤鳍棘鳞鱼 (*Sargocentron tiere*)、大西洋庸鲽 (*Hippoglossus hippoglossus*)、美国菖蒲 (*Acorus americanus*)、马 (*Equus caballus*)、欧洲牛 (*Bos taurus*)、山芥 (*Barbarea vulgaris*)、小羊驼 (*Lama guanicoe pacos*)、海岸松 (*Pinus pinaster*)、普通章鱼 (*Octopus vulgaris*)、星花茄 (*Solanum crispum*)、褐马羚 (*Hippotragus equinus*)、普通斑马达马亚种 (*Equus burchellii antiquorum*)、亚氏长毛獐 (*Crossarchus alexandri*)、月光花 (*Ipomoea alba*)、一粒小麦 (*Triticum monococcum*)、长叶杨 (*Populus jackii*)、皮氏斑纹海豚 (*Lagenorhynchus australis*)、多卡瞪羚 (*Gazella dorcas*)、胭脂虫栎 (*Quercus coccifera*)、雪雁 (*Anser caerulescens*)、白菖蒲 (*Acorus calamus*)、西藏长叶松 (*Pinus roxburghii*)、油松 (*Pinus tabulaeformis*)、费切尔泽米 (*Zamia fischeri*)、肉垂鹤 (*Grus carunculatus*)、棘小鼠 (*Acomys cahirinus*)、网纹甜瓜 (*Cucumis melo* var. *reticulatus*)、锡兰野鸡 (*Gallus lafayettei*)、豌豆 (*Pisum sativum*)、狭松 (*Pinus*

attenuata)、美国沙松 (*Pinus clausa*)、沙特瞪羚 (*Gazella saudiya*)、羴羊 (*Capra ibex*)、三浅裂野牵牛 (*Ipomoea trifida*)、繁茂大刍草 (*Zea luxurians*)、越南松 (*Pinus krempfii*)、威氏刺毛鼠 (*Acomys wilsoni*)、绉叶石蛇床 (*Petroselinum crispum*)、针栎 (*Quercus palustris*)、提莫菲维小麦 (*Triticum timopheevi*)、火鸡 (*Meleagris gallopavo*)、甘蓝 (*Brassica oleracea*)、甘蓝 (*Brassica oleracea*)、甜菜 (*Beta vulgaris*)、番茄 (*Solanum lycopersicum*)、菜豆 (*Phaseolus vulgaris*)、剑鱼 (*Xiphias gladius*)、美洲条纹狼鲈 (*Morone saxatilis*)、大口黑鲈 (*Micropterus salmoides*)、巨扇贝 (*Placopecten magellanicus*)、黍鲱 (*Sprattus sprattus*)、大西洋鲱 (*Clupea harengus*)、欧洲鳀 (*Engraulis encrasicolus*)、笋瓜 (*Cucurbita maxima*)、双孢子蘑菇 (*Agaricus bisporus*)、香蕉 (*Musa acuminata x balbisiana*)、苹果 (*Malus domestica*)、鹰嘴豆 (*Cicer arietinum*)、绿头鸭 (*Anas platyrhynchos*)、大果酸果蔓 (*Vaccinium macrocarpum*)、北美红木莓 (*Rubus idaeus x strigosus*)、矮丛蓝莓 (*Vaccinium angustifolium*)、草莓 (*Fragaria ananassa*)、欧洲黑莓 (*Rubus fruticosus*)、甜瓜 (*Cucumis melo*)、菠萝 (*Ananas comosus*)、西葫芦 (*Cucurbita pepo*)、南瓜 (*Cucurbita moschata*)、家猪 (*Sus scrofa domesticus*)、罗勒 (*Ocimum basilicum*)、迷迭香 (*Rosmarinus officinalis*)、茴香 (*Foeniculum vulgare*)、波叶大黄 (*Rheum rhabarbarum*)、番木瓜 (*Carica papaya*)、芒果 (*Mangifera indica*)、美味猕猴桃 (*Actinidia deliciosa*)、杏 (*Prunus armeniaca*)、欧洲甜樱桃 (*Prunus avium*)、椰子 (*Cocos nucifera*)、油橄榄 (*Olea europaea*)、西洋梨 (*Pyrus communis*)、无花果 (*Ficus carica*)、鸡蛋果 (*Passiflora edulis*)、水稻亚种粳稻 (*Oryza sativa* subsp. *Japonica*)、籼稻 (*Oryza sativa* subsp. *Indica*)、鹌鹑 (*Coturnix coturnix*)、酿酒酵母 (*Saccharomyces cerevisiae*)。

[0110] 在一些实施方案中,营养性蛋白质包含以下或由以下组成:天然存在于可食用物质中的蛋白质或蛋白质的片段的衍生物或突变蛋白。所述营养性蛋白质可称为“工程化营养性蛋白质”。在所述实施方案中,天然蛋白质或其片段是“参照”蛋白质或多肽且工程化营养性蛋白质或其第一多肽序列相对于所述参照蛋白质或多肽的氨基酸序列包含至少一个序列修饰。举例来说,在一些实施方案中,工程化营养性蛋白质或其第一多肽序列与至少一个参照营养性蛋白质氨基酸序列至少 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 99.5% 同源。通常,工程化营养性蛋白质或其第一多肽序列中存在的支链氨基酸残基与总氨基酸残基的比率、必需氨基酸残基与总氨基酸残基的比率、以及亮氨酸残基与总氨基酸残基的比率的至少一者大于参照营养性蛋白质或多肽序列中存在的支链氨基酸残基与总氨基酸残基的相应比率、必需氨基酸残基与总氨基酸残基的相应比率、以及亮氨酸残基与总氨基酸残基的相应比率的至少一者。

[0111] 在一些实施方案中,营养性蛋白质是食物中的丰富蛋白质或其衍生物或突变蛋白,或是食物中的丰富蛋白质或其衍生物或突变蛋白的片段。丰富蛋白质是相对于食物中存在的其它蛋白质,以较高浓度存在于所述食物中的蛋白质。食物可为处于有限地理位置的仅一小组某一类型的哺乳动物的膳食的已知组分,或为遍及世界上大部分区域的膳食主要成分。在一些实施方案中,食物中的丰富蛋白质选自鸡蛋蛋白质,如卵清蛋白、卵铁传递



蛋白和卵类粘蛋白；肉蛋白质，如肌球蛋白、肌动蛋白、原肌凝蛋白、胶原蛋白和肌钙蛋白；谷类蛋白质，如酪蛋白、 $\alpha 1$  酪蛋白、 $\alpha 2$  酪蛋白、 $\beta$  酪蛋白、 $\kappa$  酪蛋白、 $\beta$ -乳球蛋白、 $\alpha$ -乳白蛋白、大豆球蛋白、 $\beta$ -伴大豆球蛋白、谷蛋白、谷醇溶蛋白、麦醇溶蛋白、麦谷蛋白、白蛋白、球蛋白；鸡肌肉蛋白质，如白蛋白、烯醇酶、肌酸激酶、磷酸甘油酸变位酶、磷酸丙糖异构酶、载脂蛋白、卵铁传递蛋白、葡萄糖磷酸变位酶、磷酸甘油酸激酶、甘油-3-磷酸脱氢酶、3-磷酸甘油醛脱氢酶、血红素、丝切蛋白、糖原磷酸化酶、果糖-1,6-二磷酸酶、肌动蛋白、肌球蛋白、原肌球蛋白  $\alpha$  链、酪蛋白激酶、糖原磷酸化酶、果糖-1,6-二磷酸酶、醛缩酶、微管蛋白、波形蛋白、内质蛋白 (endoplasmic)、乳酸脱氢酶、肌动蛋白解聚因子 (destain)、转甲状腺素蛋白、双磷酸果糖醛缩酶、碳酸酐酶、醛脱氢酶、膜联蛋白、腺苷高半胱氨酸酶 (adenosyl homocysteinase)；猪肉肌肉蛋白质，如肌动蛋白、肌球蛋白、烯醇酶、肌联蛋白 (titin)、丝切蛋白、磷酸甘油酸激酶、烯醇酶、丙酮酸脱氢酶、糖原磷酸化酶、磷酸丙糖异构酶、肌激酶；和鱼蛋白质，如小白蛋白、丙酮酸脱氢酶、肌间线蛋白 (desmin) 和磷酸丙糖异构酶。

[0112] 通常被认为是高品质氨基酸的良好来源的三种蛋白质天然来源是乳清蛋白质、蛋蛋白质和大豆蛋白质。各来源包含多种蛋白质。表 1 呈现各氨基酸在蛋白质来源中的表示为百分比的重量比例表示 (gAA/g 蛋白质)。

[0113] 表 1

[0114]

| 氨基酸  | 乳清    | 蛋     | 大豆    |
|------|-------|-------|-------|
| 异亮氨酸 | 6.5%  | 5.5%  | 5.0%  |
| 亮氨酸  | 11.0% | 8.6%  | 8.0%  |
| 赖氨酸  | 9.1%  | 7.2%  | 6.3%  |
| 甲硫氨酸 | 2.1%  | 3.1%  | 1.3%  |
| 苯丙氨酸 | 3.4%  | 5.3%  | 1.2%  |
| 苏氨酸  | 7.0%  | 4.8%  | 3.7%  |
| 色氨酸  | 1.7%  | 1.2%  | 1.3%  |
| 缬氨酸  | 6.2%  | 6.1%  | 4.9%  |
| 组氨酸  | 2.0%  | 2.4%  | 2.7%  |
| 其它   | 51.7% | 49.5% | 60.4% |

[0115] 表 2 呈现必需氨基酸、支链氨基酸 (L、I 和 V) 和亮氨酸 (L) (单独) 在各蛋白质来源中的重量比例。

[0116] 表 2

[0117]

| 蛋白质来源 | 必需氨基酸 | 支链氨基酸 | 亮氨酸   |
|-------|-------|-------|-------|
| 乳清    | 49.0% | 23.7% | 11.0% |
| 蛋     | 50.5% | 20.1% | 8.6%  |
| 大豆    | 39.6% | 17.9% | 8.0%  |

[0118] 被依赖来确定乳清的氨基酸含量的出处是:Belitz HD., Grosch W. 和 Schieberle P. Food Chemistry (第4版). Springer-Verlag, Berlin Heidelberg 2009; <http://www.gnc.com/product/index.jsp?productId=2986027>; [http://www.nutrabio.com/Products/whey\\_protein\\_concentrate.htm](http://www.nutrabio.com/Products/whey_protein_concentrate.htm); 和 [http://www.nutrabio.com/Products/whey\\_protein\\_isolate.htm](http://www.nutrabio.com/Products/whey_protein_isolate.htm)。将来自那些出处的氨基酸含量值平均化以得到表1和2中呈现的数值。用于大豆蛋白质的出处是 Egg, National Nutrient Database for Standard Reference, 发布版 24 (<http://ndb.nal.usda.gov/ndb/foods/list>)。用于大豆蛋白质的出处是 Self Nutrition Data (<http://nutritiondata.self.com/facts/legumes-and-legume-products/4389/2>)。

[0119] 如本文所用,“乳清蛋白质”或“乳清”是指包含根据表1和2的氨基酸组成的蛋白质混合物。如本文所用,以重量计,乳清蛋白质包含49%必需氨基酸、24%支链氨基酸和11%亮氨酸。

[0120] 如本文所用,“蛋蛋白质”或“蛋”是指包含根据表1和2的氨基酸组成的蛋白质混合物。如本文所用,以重量计,蛋蛋白质包含51%必需氨基酸、20%支链氨基酸和9%亮氨酸。

[0121] 如本文所用,“大豆蛋白质”或“大豆”是指包含根据表1和2的氨基酸组成的蛋白质混合物。如本文所用,以重量计,大豆蛋白质包含40%必需氨基酸、18%支链氨基酸和8%亮氨酸。

[0122] 在本文的一些情况下,特定类型的氨基酸在多肽、蛋白质或组合物内所占的部分是基于所述类型的氨基酸与所论述多肽、蛋白质或组合物中存在的氨基酸的总重量的重量比率来定量。这个值是通过用多肽、蛋白质或组合物中的特定氨基酸的重量除以多肽、蛋白质或组合物中存在的所有氨基酸的重量来计算。

[0123] 在其它情况下,使用多肽或蛋白质中存在的特定类型的氨基酸残基与所论述多肽或蛋白质中存在的氨基酸的总数的比率。这个值是通过用各多肽或蛋白质分子中存在的所论述氨基酸的数目除以各多肽或蛋白质分子中存在的氨基酸残基的总数来计算。熟练技术人员了解的是这两种方法是可互换的且多肽或蛋白质中存在的某一类型的氨基酸的重量比例可换算成特定类型的氨基酸残基的比率,且反过来也是一样。

[0124] 在本文的某些实施方案中,乳清、蛋或大豆中的支链氨基酸、亮氨酸和/或必需氨基酸的重量比例用作测量多肽、蛋白质或包含多肽和蛋白质的至少一者的组合物的氨基酸组成的基准。在那些实施方案中,应了解的是两种量度标准不完全等效,但也应了解的是所述量度标准产生的测量结果就用于这个目的而言足够类似。举例来说,当目标蛋白质被表征为包含的支链氨基酸残基与总氨基酸残基的比率等于或大于24%(乳清中存在的支链

氨基酸残基的重量比例)时,那就是对蛋白质的支链氨基酸含量的准确描述。同时,那种蛋白质中存在的支链氨基酸残基的重量比例不必恰好等于 24%。即使如此,熟练技术人员也应了解的是这是一种适用比较。如果提供目标蛋白质中存在的氨基酸残基的总数,那么熟练技术人员也可确定目标蛋白质中的支链氨基酸残基的重量比例。

[0125] 在本文一些实施方案中,选择且任选分离天然存在的营养性蛋白质的片段。在一些实施方案中,片段包含至少 25 个氨基酸。在一些实施方案中,片段包含至少 50 个氨基酸。在一些实施方案中,片段由至少 25 个氨基酸组成。在一些实施方案中,片段由至少 50 个氨基酸组成。在一些实施方案中,提供分离的重组营养性蛋白质。在一些实施方案中,营养性蛋白质包含第一多肽序列,且所述第一多肽序列包含天然存在的营养性蛋白质的至少 25 或至少 50 个氨基酸的片段。在一些实施方案中,营养性蛋白质是分离的营养性蛋白质。在一些实施方案中,营养性蛋白质是重组营养性蛋白质。在一些实施方案中,营养性蛋白质包含含有天然存在的营养性蛋白质的至少 50 个氨基酸的片段的第一多肽序列。在一些实施方案中,营养性蛋白质是分离的重组营养性蛋白质。在一些实施方案中,本文公开的分离的重组营养性蛋白质是以非分离和 / 或非重组形式提供。

[0126] 在一些实施方案中,根据本公开的营养性蛋白质包含含有天然存在的营养性蛋白质的片段的第一多肽序列。在营养性蛋白质的一些实施方案中,营养性蛋白质由第一多肽序列组成。在营养性蛋白质的一些实施方案中,营养性蛋白质由天然存在的营养性蛋白质的片段组成。

[0127] 在一些实施方案中,根据本公开的营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的支链氨基酸残基与总氨基酸残基的比率。因此,在所述实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于选自 24%、20% 和 18% 的比率。

[0128] 在一些实施方案中,根据本公开的营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的 L 残基与总氨基酸残基的比率。因此,在所述实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于选自 11%、9% 和 8% 的比率。

[0129] 在一些实施方案中,根据本公开的营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的必需氨基酸残基与总氨基酸残基的比率。因此,在所述实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于选自 49%、51% 和 40% 的比率。

[0130] 在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的支链氨基酸残基与总氨基酸残基的比率;且包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的 L 残基与总氨基酸残基的比率。在一些所述实施方案中,营养性蛋白质进一步包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸

残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的必需氨基酸残基与总氨基酸残基的比率。

[0131] 在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24% 且 L 残基与总氨基酸残基的比率等于或大于 11%。在一些所述实施方案中,营养性蛋白质进一步包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于 49%。

[0132] 在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 20% 且 L 残基与总氨基酸残基的比率等于或大于 9%。在一些所述实施方案中,营养性蛋白质进一步包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于 51%。

[0133] 在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 18% 且 L 残基与总氨基酸残基的比率等于或大于 8%。在一些所述实施方案中,营养性蛋白质进一步包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于 40%。

[0134] 在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的支链氨基酸残基与总氨基酸残基的比率;且包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的必需氨基酸残基与总氨基酸残基的比率。在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 49%。在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 20% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 51%。在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 18% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 40%。

[0135] 在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的 L 残基与总氨基酸残基的比率;且包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的必需氨基酸残基与总氨基酸残基的比率。在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于 11% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 49%。在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的 L 氨基酸残基与总氨基酸残基的比率等于或大于 9% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 51%。在一些实施方案中,营养性蛋白质包含第一多肽序列,所述第一多肽序列包含的 L 氨基酸残基与总氨基酸残基的比率等于或大于 8% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 40%。

[0136] 在营养性蛋白质的一些实施方案中,第一多肽序列包含第一多肽序列,所述第一

多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24%、L 残基与总氨基酸残基的比率等于或大于 11%、必需氨基酸残基与总氨基酸残基的比率等于或大于 49%，且包含每种必需氨基酸的至少一者，且所述第一多肽序列选自 SEQ ID NO:1 至 SEQ ID NO:1000。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1 至 SEQ ID NO:1000 的修饰的衍生物。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1 至 SEQ ID NO:1000 的突变蛋白。

[0137] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24%、L 残基与总氨基酸残基的比率等于或大于 11% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 49%，且所述第一多肽序列选自 SEQ ID NO:1001 至 SEQ ID NO:1414。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1001 至 SEQ ID NO:1414 的修饰的衍生物。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1001 至 SEQ ID NO:1414 的突变蛋白。

[0138] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24%、必需氨基酸残基与总氨基酸残基的比率等于或大于 49%，且包含每种必需氨基酸的至少一者，且所述第一多肽序列选自 SEQ ID NO:1415 至 SEQ ID NO:1899。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1415 至 SEQ ID NO:1899 的修饰的衍生物。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1415 至 SEQ ID NO:1899 的突变蛋白。

[0139] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于 49% 且 L 残基与总氨基酸残基的比率等于或大于 11%，且包含每种必需氨基酸的至少一者，且所述第一多肽序列选自 SEQ ID NO:1900 至 SEQ ID NO:2102。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1900 至 SEQ ID NO:2102 的修饰的衍生物。在一些实施方案中，第一多肽序列选自 SEQ ID NO:1900 至 SEQ ID NO:2102 的突变蛋白。

[0140] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24% 且 L 残基与总氨基酸残基的比率等于或大于 11%，且包含每种必需氨基酸的至少一者，且所述第一多肽序列选自 SEQ ID NO:2103 至 SEQ ID NO:2518。在一些实施方案中，第一多肽序列选自 SEQ ID NO:2103 至 SEQ ID NO:2518 的修饰的衍生物。在一些实施方案中，第一多肽序列选自 SEQ ID NO:2103 至 SEQ ID NO:2518 的突变蛋白。

[0141] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24%、L 残基与总氨基酸残基的比率等于或大于 11%，且包含每种必需氨基酸的至少一者。

[0142] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 49%。

[0143] 在营养性蛋白质的一些实施方案中，第一多肽序列包含第一多肽序列，所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于 11% 且必需氨基酸残基与总氨基酸残基的比率等于或大于 49%。

[0144] 在营养性蛋白质的一些实施方案中,第一多肽序列包含第一多肽序列,所述第一多肽序列包含的必需氨基酸残基与总氨基酸残基的比率等于或大于 49% 且包含每种必需氨基酸的至少一者。

[0145] 在营养性蛋白质的一些实施方案中,第一多肽序列包含第一多肽序列,所述第一多肽序列包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24% 且包含每种必需氨基酸的至少一者。

[0146] 在营养性蛋白质的一些实施方案中,第一多肽序列包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于 11% 且包含每种必需氨基酸的至少一者。

[0147] 在营养性蛋白质的一些实施方案中,第一多肽序列包含第一多肽序列,所述第一多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于 11%。

[0148] 在营养性蛋白质的一些实施方案中,营养性蛋白质进一步包含第二多肽序列。在一些实施方案中,第二多肽序列由 3 至 10、5 至 20、10 至 30、20 至 50、25 至 75、50 至 100、或 100 至 200 个氨基酸组成。在一些实施方案中,第二多肽序列不源于天然存在的营养性蛋白质。在一些实施方案中,第二多肽序列选自用于亲和纯化的标签、蛋白质结构域接头和蛋白酶识别位点。在一些实施方案中,用于亲和纯化的标签是聚组氨酸标签。在一些实施方案中,蛋白质结构域接头包含序列 GGSG 的至少一个拷贝。在一些实施方案中,蛋白酶选自胃蛋白酶、胰蛋白酶和胰凝乳蛋白酶。

[0149] 在一些实施方案中,重组营养性蛋白质由第一多肽序列和第二多肽序列组成。

[0150] 在一些实施方案中,营养性蛋白质进一步包含第三多肽序列,所述第三多肽序列包含天然存在的营养性蛋白质的至少 25 或至少 50 个氨基酸的片段。在一些实施方案中,第一多肽序列和第三多肽序列是相同的。在一些实施方案中,第一多肽序列和第三多肽序列是不同的。在一些实施方案中,第一多肽序列和第三多肽序列源于相同天然存在的营养性蛋白质。在一些实施方案中,第一多肽序列和第三多肽序列在分离的重组营养性蛋白质中的顺序与第一多肽序列和第三多肽序列在天然存在的营养性蛋白质中的顺序相同。在一些实施方案中,第一多肽序列和第三多肽序列在营养性蛋白质中的顺序不同于第一多肽序列和第三多肽序列在天然存在的营养性蛋白质中的顺序。在一些实施方案中,第一多肽序列和第三多肽序列源于不同天然存在的营养性蛋白质。在一些实施方案中,第二多肽序列由第一多肽序列和第三多肽序列侧接。

[0151] 在营养性蛋白质的一些实施方案中,第一和 / 或第三多肽序列包含天然存在的营养性蛋白质的片段且第一和 / 或第三多肽序列中存在的支链氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的支链氨基酸残基与总氨基酸残基的比率。

[0152] 在营养性蛋白质的一些实施方案中,第一和 / 或第三多肽序列包含含有天然存在的营养性蛋白质的片段的第一和 / 或第三多肽序列且第一和 / 或第三多肽序列中存在的 L 残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的 L 残基与总氨基酸残基的比率。

[0153] 在营养性蛋白质的一些实施方案中,第一和 / 或第三多肽序列包含含有天然存在的营养性蛋白质的片段的第一和 / 或第三多肽序列且第一和 / 或第三多肽序列中存在的必

需氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的必需氨基酸残基与总氨基酸残基的比率。

[0154] 在营养性蛋白质的一些实施方案中,第一和 / 或第三多肽序列包含含有天然存在的营养性蛋白质的片段的第一和 / 或第三多肽序列且第一和 / 或第三多肽序列中存在的支链氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的支链氨基酸残基与总氨基酸残基的比率;且第一和 / 或第三多肽序列包含的 L 残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的 L 残基与总氨基酸残基的比率。在一些所述实施方案中,第一和 / 或第三多肽序列进一步包含的必需氨基酸残基与总氨基酸残基的比率等于或大于乳清蛋白质、蛋蛋白质和大豆蛋白质的至少一者中存在的必需氨基酸残基与总氨基酸残基的比率。

[0155] 在营养性蛋白质的一些实施方案中,第一和 / 或第三多肽序包含第一和 / 或第三多肽序,所述第一和 / 或第三多肽序包含的支链氨基酸残基与总氨基酸残基的比率等于或大于 24% 且 L 残基与总氨基酸残基的比率等于或大于 11%。在一些所述实施方案中,第一和 / 或第三多肽序列进一步包含的必需氨基酸残基与总氨基酸残基的比率等于或大于 49%。

[0156] 在一些实施方案中,片段所源于的天然存在的营养性蛋白质是除至少一种选自以下的营养性蛋白质以外的营养性蛋白质:蛋蛋白质,如卵清蛋白、卵铁传递蛋白和卵类粘蛋白;肉蛋白质,如肌球蛋白、肌动蛋白、原肌凝蛋白、胶原蛋白和肌钙蛋白;奶蛋白质,如乳清和酪蛋白;谷类蛋白质,如酪蛋白、 $\alpha 1$  酪蛋白、 $\alpha 2$  酪蛋白、 $\beta$  酪蛋白、 $\kappa$  酪蛋白、 $\beta$ -乳球蛋白、 $\alpha$ -乳白蛋白、大豆球蛋白、 $\beta$ -伴大豆球蛋白、谷蛋白、谷醇溶蛋白、麦醇溶蛋白、麦谷蛋白、白蛋白、球蛋白;鸡肌肉蛋白质,如白蛋白、烯醇酶、肌酸激酶、磷酸甘油酸变位酶、磷酸丙糖异构酶、载脂蛋白、卵铁传递蛋白、葡糖磷酸变位酶、磷酸甘油酸激酶、甘油-3-磷酸脱氢酶、3-磷酸甘油醛脱氢酶、血红素、丝切蛋白、糖原磷酸化酶、果糖-1,6-二磷酸酶、肌动蛋白、肌球蛋白、原肌球蛋白  $\alpha$ -链、酪蛋白激酶、糖原磷酸化酶、果糖-1,6-二磷酸酶、醛缩酶、微管蛋白、波形蛋白、内质蛋白、乳酸脱氢酶、肌动蛋白解聚因子、转甲状腺素蛋白、双磷酸果糖醛缩酶、碳酸酐酶、醛脱氢酶、膜联蛋白、腺苷高半胱氨酸酶;猪肉肌肉蛋白质,如肌动蛋白、肌球蛋白、烯醇酶、肌联蛋白、丝切蛋白、磷酸甘油酸激酶、烯醇酶、丙酮酸脱氢酶、糖原磷酸化酶、磷酸丙糖异构酶、肌激酶;和鱼蛋白质,如小白蛋白、丙酮酸脱氢酶、肌间线蛋白和磷酸丙糖异构酶。

[0157] 苯丙酮酸尿症 (PKU) 是一种常染色体隐性代谢遗传病症,其特征在于肝酶苯丙氨酸羟化酶 (PAH) 的基因突变,从而致使它无功能。这种酶为将苯丙氨酸代谢成酪氨酸所必需。当 PAH 活性降低时,苯丙氨酸累积且转化成在尿中检测到的苯丙酮酸 (也称为苯基甲酮)。未治疗的儿童在出生时是正常的,但未能达到早期发育里程碑,显现小头畸形,且显示进行性脑功能损害。活动过度、EEG 异常和癫痫发作以及重度学习失能是后来生活中的主要临床问题。也观察到皮肤、毛发、汗和尿的特征性气味 (归因于乙酸苯酯累积);以及色素沉着不足和湿疹的倾向。所有 PKU 患者都必须遵循低 Phe 特殊膳食。因此,包含低数量 Phe 残基或不包含 Phe 残基的营养性蛋白质合乎 PKU 患者需要。所述蛋白质可通过选择本文提供的具有少许 Phe 残基或不具有 Phe 残基的营养性蛋白质来获得。因此,在一些实施方案中,营养性蛋白质包含的 Phe 残基与总氨基酸残基的比率等于或低于 5%、4%、3%、

2%或1%。在一些实施方案中,营养性蛋白质包含10个或更少Phe残基、9个或更少Phe残基、8个或更少Phe残基、7个或更少Phe残基、6个或更少Phe残基、5个或更少Phe残基、4个或更少Phe残基、3个或更少Phe残基、2个或更少Phe残基、1个Phe残基,或不包含Phe残基。在一些实施方案中,营养性蛋白质不包含Phe残基。

[0158] 精氨酸是一种条件性非必需氨基酸,这意味在大多数时间,它可由人体制造,且不需要直接通过膳食获得。营养不良的个体、老年人或患有某些身体病状(例如,败血症)的人可能不产生足量精氨酸且因此需要增加他们对含有精氨酸的食物的摄取。据信精氨酸具有有益健康性质,包括降低损伤(特别是骨损伤)愈合时间,以及降低血压,特别是在高风险妊娠期间(惊厥前期)的高血压。此外,研究已显示膳食性补充L-精氨酸有益于增强具有天然存在的子宫内生长迟缓的猪的繁殖性能,增强奶喂食的仔猪的蛋白质沉积和出生后生长,使链脲佐菌素诱发性糖尿病大鼠的血浆葡萄糖水平正常,降低肥胖 Zucker 糖尿病多脂肪(ZDF)大鼠的脂肪质量,以及改进糖尿病大鼠的血管功能。为将这些益处与本文公开的营养性蛋白质的至少一种效用组合,在本文公开的营养性蛋白质的一些实施方案中,营养性蛋白质包含的营养性蛋白质中的精氨酸残基与总氨基酸残基的比率等于或大于3%、等于或大于4%、等于或大于5%、等于或大于6%、等于或大于7%、等于或大于8%、等于或大于9%、等于或大于10%、等于或大于11%、或等于或大于12%。

[0159] 可消化性是营养性蛋白质的营养性益处和效用相关的参数。与相对消化完全性相关的信息可充当肽生物可用度的预测因子(Daniel, H., 2003. *Molecular and Integrative Physiology of Intestinal Peptide Transport*. Annual Review of Physiology, 第66卷, 第361-384页)。在一些实施方案中,筛选本文公开的营养性蛋白质以评估它们的可消化性。可通过本领域中已知的任何适合方法评估营养性蛋白质的可消化性。在一些实施方案中,可消化性是通过包括一个或两个蛋白质消化阶段(即模拟胃消化和模拟肠消化)的生理相关体外消化反应来评估(参见例如, Moreno, 等, 2005. Stability of the major allergen Brazil nut 2S albumin(Ber e 1) to physiologically relevant in vitro gastrointestinal digestion. FEBS Journal, 第341-352页; Martos, G., Contreras, P., Molina, E. 和 Lopez-Fandino, R., 2010. Egg White Ovalbumin Digestion Mimicking Physiological Conditions. Journal of Agricultural and food chemistry, 第5640-5648页; Moreno, F. J., Mackie, A. R. 和 Clare Mills, E. N., 2005. Phospholipid interactions protect the milk allergen  $\alpha$ -Lactalbumin from proteolysis during in vitro digestion. Journal of agricultural and food chemistry, 第9810-9816页)。简而言之,使测试蛋白质依序暴露于模拟胃液(SGF)120分钟(使90%液体餐食自胃传至小肠所花费的时长;参见 Kong, F. 和 Singh, R. P., 2008. Disintegration of Solid Foods in Human Stomach. Journal of Food Science, 第67-80页),且接着转移至模拟十二指肠液(SDF)中以再消化120分钟。通过电泳(例如,芯片电泳或SDS-PAGE)分析不同消化阶段(例如,2、5、15、30、60和120分钟)的样本以监测完整蛋白质以及任何大型消化片段(例如,大于4kDa)的大小和量。蛋白质随时间消失指示蛋白质在测定中被消化的速率。通过监测随时间观察到的完整蛋白质的量,计算SGF的消化半衰期( $\tau_{1/2}$ ),且如果在用SGF处理之后检测到完整蛋白质,那么计算SIF的消化 $\tau_{1/2}$ 。这个测定可用于评估比较可消化性(即,相对于如乳清的基准蛋白质)或评估绝对可消化



性。在一些实施方案中,营养性蛋白质的可消化性高于(即,SGF  $\tau$  1/2 和 / 或 SIF  $\tau$  1/2 较短)乳清蛋白质。在一些实施方案中,营养性蛋白质的 SGF  $\tau$  1/2 是 30 分钟或更短、20 分钟或更短、15 分钟或更短、10 分钟或更短、5 分钟或更短、4 分钟或更短、3 分钟或更短、2 分钟或更短、或 1 分钟或更短。在一些实施方案中,营养性蛋白质的 SIF  $\tau$  1/2 是 30 分钟或更短、20 分钟或更短、15 分钟或更短、10 分钟或更短、5 分钟或更短、4 分钟或更短、3 分钟或更短、2 分钟或更短、或 1 分钟或更短。在一些实施方案中,截至 2 分钟、5 分钟、15 分钟、30 分钟、60 分钟或 120 分钟,不可在 SGF 和 SIF 测定的一者或两者中检测到营养性蛋白质。在一些实施方案中,在 SGF 和 SIF 的一者或两者中,营养性蛋白质是在恒定速率下和 / 或在控制速率下被消化。在所述实施方案中,可不将营养性蛋白质的消化速率最优化成最高可能消化速率。在所述实施方案中,相较于在 SGF 和 SIF 的一者或两者中在较快初始速率下消化的具有类似氨基酸组成的营养性蛋白质,蛋白质在由哺乳动物摄取之后的吸收速率可较慢且在摄取之后进行吸收所历经的总时期可较长。在一些实施方案中,营养性蛋白质在 SGF 中被完全或大致上完全消化。在一些实施方案中,营养性蛋白质大致上未由 SGF 消化或未由 SGF 消化;在大多数所述实施方案中,蛋白质在 SIF 中被消化。

[0160] 评估蛋白质可消化性也可提供对蛋白质的潜在过敏原性的了解,因为对消化性蛋白酶具有抗性的蛋白质或蛋白质的大片段可具有导致过敏原性反应的较高风险 (Goodman, R. E. 等, 2008. Allergenicity assessment of genetically modified crops-what makes sense? *Nature Biotechnology*, 第 73-81 页)。为检测和鉴定太小以致不能进行芯片电泳分析的肽,可使用液相色谱和质谱分析。在 SGF 样本中,可通过 LC/MS 直接检测和鉴定肽。SIF 蛋白质消化可需要在通过 LC/MS 检测和鉴定之前进行纯化以移除胆汁酸。

[0161] 在一些实施方案中,营养性蛋白质的可消化性是通过鉴定和定量蛋白质氨基酸序列中的消化性蛋白酶识别位点来评估。在一些实施方案中,营养性蛋白质包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。

[0162] 如本文所用,“胃蛋白酶识别位点”是多肽序列中的实验显示由胃蛋白酶裂解的任何位点。在一些实施方案中,所述位点是在选自 Phe、Trp、Tyr、Leu、Ala、Glu 和 Gln 的氨基酸残基之后(即,下游)的肽键,前提是随后残基不是选自 Ala、Gly 和 Val 的氨基酸残基。

[0163] 如本文所用,“胰蛋白酶识别位点”是多肽序列中的实验显示由胰蛋白酶裂解的任何位点。在一些实施方案中,所述位点是在选自 Lys 或 Arg 的氨基酸残基之后的肽键,前提是随后残基不是脯氨酸。

[0164] 如本文所用,“胰凝乳蛋白酶识别位点”是多肽序列中的实验显示由胰凝乳蛋白酶裂解的任何位点。在一些实施方案中,所述位点是在选自 Phe、Trp、Tyr 和 Leu 的氨基酸残基之后的肽键。

[0165] 相较于在不存在二硫键下将存在的消化速率,蛋白质中的二硫键键结的半胱氨酸残基倾向于降低蛋白质的消化速率。举例来说,已显示当蛋白质 b- 乳球蛋白的二硫桥被裂解时,它的消化速率增加 (I. M. Reddy, N. K. D. Kella 和 J. E. Kinsella. “Structural and Conformational Basis of the Resistance of B-Lactoglobulin to Peptic and Chymotryptic Digestion”. *J. Agric. Food Chem.* 1988, 36, 737-741)。因此,二硫键较少的营养性蛋白质的可消化性倾向于高于二硫键的数目较大的类似营养性蛋白质。在一些实施

方案中,筛选本文公开的营养性蛋白质以鉴定各营养性蛋白质中存在的半胱氨酸残基的数目且具体来说以允许选择包含相对较低数目的半胱氨酸残基的营养性蛋白质。举例来说,可鉴定不包含 Cys 残基或包含相对较低数目的 Cys 残基,如 10 个或更少 Cys 残基、9 个或更少 Cys 残基、8 个或更少 Cys 残基、7 个或更少 Cys 残基、6 个或更少 Cys 残基、5 个或更少 Cys 残基、4 个或更少 Cys 残基、3 个或更少 Cys 残基、2 个或更少 Cys 残基、1 个 Cys 残基,或不包含 Cys 残基的天然存在的营养性蛋白质或片段。在一些实施方案中,通过缺失和 / 或通过用另一氨基酸取代来移除天然存在的营养性蛋白质或其片段中的一个或多个 Cys 残基。在一些实施方案中,缺失或置换 1 个 Cys 残基,缺失或置换 1 个或多个 Cys 残基,缺失或置换 2 个或多个 Cys 残基,缺失或置换 3 个或多个 Cys 残基,缺失或置换 4 个或多个 Cys 残基,缺失或置换 5 个或多个 Cys 残基,缺失或置换 6 个或多个 Cys 残基,缺失或置换 7 个或多个 Cys 残基,缺失或置换 8 个或多个 Cys 残基,缺失或置换 9 个或多个 Cys 残基,或缺失或置换 10 个或多个 Cys 残基。在一些实施方案中,本公开的营养性蛋白质包含的 Cys 残基与总氨基酸残基的比率等于或低于 5%、4%、3%、2% 或 1%。在一些实施方案中,营养性蛋白质包含 10 个或更少 Cys 残基、9 个或更少 Cys 残基、8 个或更少 Cys 残基、7 个或更少 Cys 残基、6 个或更少 Cys 残基、5 个或更少 Cys 残基、4 个或更少 Cys 残基、3 个或更少 Cys 残基、2 个或更少 Cys 残基、1 个 Cys 残基,或不包含 Cys 残基。在一些实施方案中,营养性蛋白质包含 1 个或更少 Cys 残基。在一些实施方案中,营养性蛋白质不包含 Cys 残基。

[0166] 或者或此外,可移除存在或可能存在于营养性蛋白质中的二硫键。可使用化学方法,通过用如  $\beta$ -巯基乙醇、二硫苏糖醇 (DTT) 或三 (2-羧乙基) 膦 (TCEP) 的还原剂将二硫化物还原成两个硫醇基团来移除二硫化物。硫醇可接着用如碘乙酰胺、N-乙基马来酰亚胺或亚硫酸钠的试剂共价修饰或“加帽”(参见例如, Crankshaw, M. W. 和 Grant, G. A. 2001. Modification of Cysteine. Current Protocols in Protein Science. 15. 1. 1 - 15. 1. 18)。

[0167] 真核蛋白质常被糖基化,且连接于蛋白质的碳水化合物链起各种作用。N 连接和 O 连接的糖基化是存在于蛋白质中的两种最常见形式的糖基化。N 连接的糖基化是糖分子连接于蛋白质中的氨基酸残基中的氮原子。N 连接的糖基化存在于天冬酰胺和精氨酸残基处。O 连接的糖基化是糖分子连接于蛋白质中的氨基酸残基中的氧原子。O 连接的糖基化存在于苏氨酸和丝氨酸残基处。

[0168] 糖基化蛋白质常比它们的未糖基化形式更可溶。就蛋白质药物而言,适当糖基化通常赋予高活性、适当抗原结合性、在血液中的更好稳定性等。然而,糖基化必须是指蛋白质“携带”糖部分。所述糖部分可降低包括重组营养性蛋白质的本公开的营养性蛋白质的适用性。举例来说,如实施例中所显示,将相同蛋白质的糖基化形式和非糖基化形式的消化进行比较显示非糖基化形式比糖基化形式更快速被消化。出于这些原因,在一些实施方案中,根据本公开的营养性蛋白质包含较低糖基化或不包含糖基化。举例来说,在一些实施方案中,营养性蛋白质包含的非糖基化残基与总氨基酸残基的比率是至少 80%、至少 85%、至少 90%、至少 95%、至少 96%、至少 97%、至少 98% 或至少 99%。在一些实施方案中,营养性蛋白质不包含任何糖基化。

[0169] 在一些实施方案中,本公开的营养性蛋白质在它产生之后或在它分离之后被去糖基化。低糖基化或无糖基化的营养性蛋白质可通过本领域中已知的任何方法制备。举例来说,可使用酶方法和 / 或化学方法 (Biochem. J. (2003) 376, 第 339-350 页.)。酶是在研究

规模上商业产生以移除 N 连接和 O 连接的寡糖。化学方法包括使用三氟甲烷磺酸来选择性断裂 N 连接和 O 连接的肽 - 糖键。这个方法产生的去糖基化常比使用酶方法达成的去糖基化更完全。

[0170] 在其它实施方案中,根据本公开的营养性蛋白质是由宿主生物体在低糖基化或无糖基化下产生。大多数细菌和其它原核生物具有极其有限的使蛋白质(尤其异源蛋白质)糖基化的能力。因此,在本公开的一些实施方案中,营养性蛋白质是在微生物中重组制备以使重组蛋白的糖基化水平较低或无糖基化。在一些实施方案中,重组营养性蛋白质的糖基化水平低于如蛋白质所源于的生物体中存在的蛋白质的糖基化水平。

[0171] 在一些实施方案中,根据本公开的营养性蛋白质或多肽包含的选自 Asn、Arg、Ser 和 Thr 的氨基酸与总氨基酸的比率是 20%或更低、19%或更低、18%或更低、17%或更低、16%或更低、15%或更低、14%或更低、13%或更低、12%或更低、11%或更低、10%或更低、9%或更低、8%或更低、7%或更低、6%或更低、5%或更低、4%或更低、3%或更低、2%或更低、1%或更低。在一些实施方案中,根据本公开的营养性蛋白质或多肽不包含选自 Asn、Arg、Ser 和 Thr 的氨基酸。在一些实施方案中,根据本公开的营养性蛋白质或多肽包含小于 20 个、小于 19 个、小于 18 个、小于 17 个、小于 16 个、小于 15 个、小于 14 个、小于 13 个、小于 12 个、小于 11 个、小于 10 个、小于 9 个、小于 8 个、小于 7 个、小于 6 个、小于 5 个、小于 4 个、小于 3 个、小于 2 个、小于 1 个或不包含选自 Asn、Arg、Ser 和 Thr 的氨基酸。

[0172] 在本文公开的营养性蛋白质的一些实施方案中,营养性蛋白质是可溶性的。溶解度可通过本领域中已知的任何方法来测量。在一些实施方案中,溶解度是通过离心浓缩,随后进行蛋白质浓度测定来考查。根据方案使用两种方法测试营养性蛋白质在 20mM HEPES(pH 7.5) 中的样本的蛋白质浓度,所述方法是考马斯加强型(Coomassie Plus)(Bradford) 蛋白质测定(Thermo Scientific)和二喹啉甲酸(BCA) 蛋白质测定(Sigma-Aldrich)。基于这些测量结果,添加 10mg 蛋白质至 Amicon Ultra 3kDa 离心过滤器(Millipore) 中。样本是通过在 10,000Xg 下离心 30 分钟来浓缩。考查最终此刻浓缩的样本的蛋白质沉淀且接着如上使用两种方法(Bradford 和 BCA) 测试蛋白质浓度。

[0173] 在一些实施方案中,营养性蛋白质在生理 pH 下具有的最终溶解度限度是至少 5g/L、10g/L、20g/L、30g/L、40g/L、50g/L 或 100g/L。在一些实施方案中,在生理 pH 下在浓度大于 5g/L、或 10g/L、或 20g/L、或 30g/L、或 40g/L、或 50g/L、或 100g/L 下,大于 50%、大于 60%、大于 70%、大于 80%、大于 90%、大于 95%、大于 96%、大于 97%、大于 98%、大于 99%或大于 99.5%的营养性蛋白质可溶,其中未观察到蛋白质沉淀。在一些实施方案中,营养性蛋白质的溶解度高于通常在考查乳清(12.5g/L;Pelegrine 等,Lebensm.-Wiss. U.-Technol. 38(2005)77-80) 和大豆(10g/L;Lee 等,JAOC80(1)(2003)85-90) 的溶解度限度的研究中报道的溶解度。

[0174] 如本文所用,“稳定”蛋白质是抵抗改变目标蛋白质的生物物理学(例如,溶解度)、生物(例如,可消化性)或组成(例如,亮氨酸氨基酸的比例)特性的变化(例如,解折叠、氧化、聚集、水解等)的蛋白质。

[0175] 可使用本领域中已知的各种测定测量蛋白质稳定性且可选择本文公开且具有高于阈值的稳定性的营养性蛋白质。在一些实施方案中,选择显示的热稳定性类似于或好于乳清蛋白质的热稳定性的蛋白质。热稳定性是可帮助预测营养性蛋白质的储存期限的性

质。在测定的一些实施方案中,营养性蛋白质样本的稳定性是通过在暴露于极端温度之后使用尺寸排阻色谱法 (SEC) 监测聚集形成来确定。将待测试的蛋白质的水性样本置放在 90℃ 下的加热块中且在 0、1、5、10、30 和 60 分钟之后取样以进行 SEC 分析。蛋白质是通过监测 214nm 下的吸光度来检测,且聚集体被表征为洗脱快于目标蛋白质的峰。峰面积无总体变化指示在热处理期间无蛋白质沉淀。当在所述测定中暴露于 90℃ 时,乳清蛋白质已显示快速形成约 80% 聚集体。

[0176] 在一些实施方案中,营养性蛋白质的热稳定性是通过在结合随温度增加而作为蛋白质变性物形成的聚集蛋白质的疏水性染料 (例如, **ProteoStat®** 热漂移稳定性测定试剂盒, Enzo Life Sciences) 存在下将样本自 25℃ 缓慢加热至 95℃ 来确定 (Niesen, F. H., Berglund, H. 和 Vadadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, 第 2 卷, 第 2212-2221 页)。在结合后,染料的荧光显著增加,此由 rtPCR 仪器记录且表示为蛋白质的熔解曲线 (Lavinder, J. J., Hari, S. B., Sullivan, B. J. 和 Magilery, T. J., 2009. High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. *Journal of the American Chemical Society*, 第 3794-3795 页)。在完成热漂移之后,考查样本的不溶性沉淀且由分析型尺寸排阻色谱法 (SEC) 加以进一步分析。

[0177] 在一些实施方案中,本公开的营养性蛋白质显示抗聚集性,展现例如在高温 (例如, 50℃、60℃、70℃、80℃、85℃、90℃ 或 95℃) 下小于 80% 聚集、10% 聚集,或不展现可检测聚集。

[0178] 如本文公开的稳定营养性蛋白质的一个益处是它们在使用之前可能够储存延长时期,在一些情况下无需冷藏或冷却。在一些实施方案中,营养性蛋白质被加工成干燥形式 (例如,通过冻干)。在一些实施方案中,营养性蛋白质在冻干后是稳定的。在一些实施方案中,所述冻干的营养性蛋白质在复水 (例如,液体制剂) 后维持它们的稳定性。

[0179] 对于大多数实施方案,优选的是营养性蛋白质不展现不适当较高过敏原性。因此,在一些实施方案中,评估营养性蛋白质的潜在过敏原性。这可通过本领域中已知的任何适合方法进行。在一些实施方案中,计算过敏原性计分。过敏原性计分是一种基于 WHO 推荐 (<http://www.fao.org/ag/agn/food/pdf/allergygm.pdf>) 的以一级序列为基础的用于评估蛋白质与任何已知过敏原的类似程度的量度,其中主要假设是靶标与已知过敏原之间的同一性百分比较高可能指示交叉反应性。对于给定蛋白质,可通过一互补的基于序列同源性的测试的一者或两者来评估引发过敏反应的可能性。第一测试通过使用采用 BLOSUM50 取代矩阵、缺口开放罚分 10 和缺口延伸罚分 2 的 FASTA 算法,与某一数据库的已知过敏原进行整体-整体序列比对来确定蛋白质的跨越整个序列的同一性百分比。已表明整体同源性小于 50% 的蛋白质不可能具有过敏原性 (Goodman R. E. 等 Allergenicity assessment of genetically modified crops—what makes sense? *Nat. Biotech.* 26, 73-81 (2008); Aalberse R. C. Structural biology of allergens. *J. Allergy Clin. Immunol.* 106, 228-238 (2000))。

[0180] 在营养性蛋白质的一些实施方案中,营养性蛋白质与用于分析的数据库中的任何已知过敏原具有小于 50% 整体同源性。在一些实施方案中,使用小于 40% 同源性的截断

值。在一些实施方案中,使用小于 30%同源性的截断值。在一些实施方案中,使用小于 20%同源性的截断值。在一些实施方案中,使用小于 10%同源性的截断值。在一些实施方案中,使用 40%至 50%的截断值。在一些实施方案中,使用 30%至 50%的截断值。在一些实施方案中,使用 20%至 50%的截断值。在一些实施方案中,使用 10%至 50%的截断值。在一些实施方案中,使用 5%至 50%的截断值。在一些实施方案中,使用 0%至 50%的截断值。在一些实施方案中,使用与用于分析的数据库中的任何已知过敏原具有大于 50%整体同源性的截断值。在一些实施方案中,使用 50%至 60%的截断值。在一些实施方案中,使用 50%至 70%的截断值。在一些实施方案中,使用 50%至 80%的截断值。在一些实施方案中,使用 50%至 90%的截断值。在一些实施方案中,使用 55%至 60%的截断值。在一些实施方案中,使用 65%至 70%的截断值。在一些实施方案中,使用 70%至 75%的截断值。在一些实施方案中,使用 75%至 80%的截断值。

[0181] 第二测试通过使用采用 BLOSUM50 取代矩阵、缺口开放罚分 10 和缺口延伸罚分 2 的 FASTA 算法,将各片段与某一数据库的已知过敏原进行整体-局部序列比对以确定所有可能的连续 80 个氨基酸的片段的局部过敏原性来沿蛋白质序列评估局部过敏原性。任何 80 个氨基酸的窗口与任何过敏原的最高同一性百分比被取作目标蛋白质的最终计分。WHO 指导方针关于这个片段测试建议使用 35%同一性截断值。在营养性蛋白质的一些实施方案中,营养性蛋白质的所有可能的片段都与用于使用这个测试进行分析的数据库中的任何已知过敏原具有小于 35%局部同源性。在一些实施方案中,使用小于 30%同源性的截断值。在一些实施方案中,使用 30%至 35%同源性的截断值。在一些实施方案中,使用 25%至 30%同源性的截断值。在一些实施方案中,使用 20%至 25%同源性的截断值。在一些实施方案中,使用 15%至 20%同源性的截断值。在一些实施方案中,使用 10%至 15%同源性的截断值。在一些实施方案中,使用 5%至 10%同源性的截断值。在一些实施方案中,使用 0%至 5%同源性的截断值。在一些实施方案中,使用大于 35%同源性的截断值。在一些实施方案中,使用 35%至 40%同源性的截断值。在一些实施方案中,使用 40%至 45%同源性的截断值。在一些实施方案中,使用 45%至 50%同源性的截断值。在一些实施方案中,使用 50%至 55%同源性的截断值。在一些实施方案中,使用 55%至 60%同源性的截断值。在一些实施方案中,使用 65%至 70%同源性的截断值。在一些实施方案中,使用 70%至 75%同源性的截断值。在一些实施方案中,使用 75%至 80%同源性的截断值。

[0182] 熟练技术人员能够出于这个目的鉴定和使用已知过敏原的适合数据库。在一些实施方案中,数据库是通过自一个以上数据库来源选择蛋白质加以定制。在一些实施方案中,定制数据库包含由食物过敏研究和资源程序 (Food Allergy Research and Resource Program) (<http://www.allergenonline.org/>)、UNIPROT 注释 (UNIPROT annotations) (<http://www.uniprot.org/docs/allergen>) 和过敏原性蛋白质的结构数据库 (Structural Database of Allergenic Proteins) (SDAP, [http://fermi.utmb.edu/SDAP/sdap\\_lnk.html](http://fermi.utmb.edu/SDAP/sdap_lnk.html)) 收集的汇合过敏原清单。这个数据库包括由国际免疫学联合会 (International Union of Immunological Societies) (IUIS, <http://www.allergen.org/>) 当前认可的所有过敏原以及尚未官方命名的许多其它过敏原。在一些实施方案中,数据库包含可在已知数据库中获得的一小组已知过敏原蛋白质;也就是说数据库是一小组定制选择的已知过敏原蛋白质。在一些实施方案中,已知过敏原的数据库包含至少 10 种蛋白质、至少 20 种蛋白

质、至少 30 种蛋白质、至少 40 种蛋白质、至少 50 种蛋白质、至少 100, 种蛋白质、至少 200 种蛋白质、至少 300 种蛋白质、至少 400 种蛋白质、至少 500 种蛋白质、至少 600 种蛋白质、至少 700 种蛋白质、至少 800 种蛋白质、至少 900 种蛋白质、至少 1,000 种蛋白质、至少 1,100 种蛋白质、至少 1,200 种蛋白质、至少 1,300 种蛋白质、至少 1,400 种蛋白质、至少 1,500 种蛋白质、至少 1,600 种蛋白质、至少 1,700 种蛋白质、至少 1,800 种蛋白质、至少 1,900 种蛋白质、或至少 2,000 种蛋白质。在一些实施方案中,已知过敏原的数据库包含 100 至 500 种蛋白质、200 至 1,000 种蛋白质、500 至 1,000 种蛋白质、500 至 1,000 种蛋白质、或 1,000 至 2,000 种蛋白质。

[0183] 在一些实施方案中,相对于过敏原数据库测试营养性蛋白质的所有(或一小组选择的)具有不同长度的连续氨基酸窗口(例如 70、60、50、40、30、20、10、8 或 6 个氨基酸的窗口),且鉴定具有 100%同一性、95%或更高同一性、90%或更高同一性、85%或更高同一性、80%或更高同一性、75%或更高同一性、70%或更高同一性、65%或更高同一性、60%或更高同一性、55%或更高同一性、或 50%或更高同一性匹配的肽序列以进一步考查潜在过敏原性。

[0184] 预测蛋白质的过敏原性的另一方法是评估所述蛋白质与人来源的蛋白质的同源性。人免疫系统被定期暴露于众多可能的过敏原性蛋白质且具有在宿主身体的蛋白质与外源性蛋白质之间进行区分的固有能力。这个能力的精确本质并非总是明确的,且存在由于身体未能区分自身与非自身而产生的许多疾病(例如关节炎)。尽管如此,基本假设是与人蛋白质共有一定程度的序列同源性的蛋白质引发免疫反应的可能性较小。具体来说,已显示对于一些具有已知过敏原性成员的蛋白质家族(原肌凝蛋白、小白蛋白、酪蛋白),相对于已知过敏原性蛋白质,携带与它们的人对应物的更大序列同源性的那些蛋白质不被认为具有过敏原性(Jenkins J.A. 等 Evolutionary distance from human homologs reflects allergenicity of animal food proteins. J. Allergy Clin Immunol. 120(2007):1399-1405)。对于给定蛋白质,人同源性计是通过由使用采用 BLOSUM50 取代矩阵、缺口开放罚分 10 和缺口延伸罚分 2 的 FASTA 算法进行整体-局部比对以确定所述蛋白质与某一数据库(例如,UNIPROT 数据库)的人蛋白质的最大同一性百分比来测量。根据 Jenkins 等(Jenkins J.A. 等 Evolutionary distance from human homologs reflects allergenicity of animal food proteins. J. Allergy Clin Immunol. 120(2007):1399-1405),与人蛋白质的序列同一性高于约 62%的蛋白质具有过敏原性的可能性较小。熟练技术人员能够出于这个目的鉴定和使用已知人蛋白质的适合数据库,例如通过搜索 UNIPROT 数据库(<http://www.uniprot.org>)。在一些实施方案中,数据库是通过自一个以上数据库来源选择蛋白质加以定制。当然,数据库可为但无需为全面的。在一些实施方案中,数据库包含一小组人蛋白质;也就是说数据库是一小组定制选择的人蛋白质。在一些实施方案中,人蛋白质的数据库包含至少 10 种蛋白质、至少 20 种蛋白质、至少 30 种蛋白质、至少 40 种蛋白质、至少 50 种蛋白质、至少 100, 种蛋白质、至少 200 种蛋白质、至少 300 种蛋白质、至少 400 种蛋白质、至少 500 种蛋白质、至少 600 种蛋白质、至少 700 种蛋白质、至少 800 种蛋白质、至少 900 种蛋白质、至少 1,000 种蛋白质、至少 2,000 种蛋白质、至少 3,000 种蛋白质、至少 4,000 种蛋白质、至少 5,000 种蛋白质、至少 6,000 种蛋白质、至少 7,000 种蛋白质、至少 8,000 种蛋白质、至少 9,000 种蛋白质、或至少 10,000 种

蛋白质。在一些实施方案中,数据库包含 100 至 500 种蛋白质、200 至 1,000 种蛋白质、500 至 1,000 种蛋白质、500 至 1,000 种蛋白质、1,000 至 2,000 种蛋白质、1,000 至 5,000 种蛋白质、或 5,000 至 10,000 种蛋白质。在一些实施方案中,数据库包含所有已知人蛋白质的至少 90%、至少 95%或至少 99%。

[0185] 在营养性蛋白质的一些实施方案中,营养性蛋白质与人蛋白质至少 20%同源。在一些实施方案中,使用至少 30%同源性的截断值。在一些实施方案中,使用至少 40%同源性的截断值。在一些实施方案中,使用至少 50%同源性的截断值。在一些实施方案中,使用至少 60%同源性的截断值。在一些实施方案中,使用至少 70%同源性的截断值。在一些实施方案中,使用至少 80%同源性的截断值。在一些实施方案中,使用至少 62%同源性的截断值。在一些实施方案中,使用至少 20%同源性至至少 30%同源性的截断值。在一些实施方案中,使用至少 30%同源性至至少 40%同源性的截断值。在一些实施方案中,使用至少 50%同源性至至少 60%同源性的截断值。在一些实施方案中,使用至少 60%同源性至至少 70%同源性的截断值。在一些实施方案中,使用至少 70%同源性至至少 80%同源性的截断值。

[0186] 对于大多数实施方案,优选的是营养性蛋白质不展现不适当较高毒性。因此,在一些实施方案中,评估营养性蛋白质的潜在毒性。这可通过本领域中已知的任何适合方法进行。在一些实施方案中,毒性计分是通过确定蛋白质与的已知毒性蛋白质的数据库(例如自 UNIPROT 数据库鉴定的毒性蛋白质)的同一性百分比来计算。目标蛋白质相对于已知毒素的数据库的整体-整体比对是使用采用 BLOSUM50 取代矩阵、缺口开放罚分 10 和缺口延伸罚分 2 的 FASTA 算法来进行。在营养性蛋白质的一些实施方案中,营养性蛋白质与已知毒素小于 35%同源。在一些实施方案中,使用小于 35%同源性的截断值。在一些实施方案中,使用 30%至 35%同源性的截断值。在一些实施方案中,使用 25%至 35%同源性的截断值。在一些实施方案中,使用 20%至 35%同源性的截断值。在一些实施方案中,使用 15%至 35%同源性的截断值。在一些实施方案中,使用 10%至 35%同源性的截断值。在一些实施方案中,使用 5%至 35%同源性的截断值。在一些实施方案中,使用 0%至 35%同源性的截断值。在一些实施方案中,使用大于 35%同源性的截断值。在一些实施方案中,使用 35%至 40%同源性的截断值。在一些实施方案中,使用 35%至 45%同源性的截断值。在一些实施方案中,使用 35%至 50%同源性的截断值。在一些实施方案中,使用 35%至 55%同源性的截断值。在一些实施方案中,使用 35%至 60%同源性的截断值。在一些实施方案中,使用 35%至 70%同源性的截断值。在一些实施方案中,使用 35%至 75%同源性的截断值。在一些实施方案中,使用 35%至 80%同源性的截断值。熟练技术人员能够出于这个目的鉴定和使用已知毒素的适合数据库,例如通过搜索 UNIPROT 数据库(<http://www.uniprot.org>)。在一些实施方案中,数据库是通过自一个以上数据库来源选择鉴定为毒素的蛋白质加以定制。在一些实施方案中,数据库包含一小组已知毒性蛋白质;也就是说数据库是一小组定制选择的已知毒性蛋白质。在一些实施方案中,毒性蛋白质的数据库包含至少 10 种蛋白质、至少 20 种蛋白质、至少 30 种蛋白质、至少 40 种蛋白质、至少 50 种蛋白质、至少 100, 种蛋白质、至少 200 种蛋白质、至少 300 种蛋白质、至少 400 种蛋白质、至少 500 种蛋白质、至少 600 种蛋白质、至少 700 种蛋白质、至少 800 种蛋白质、至少 900 种蛋白质、至少 1,000 种蛋白质、至少 2,000 种蛋白质、至少 3,000 种蛋白质、至少 4,000 种蛋白质、至少 5,000 种

蛋白质、至少 6,000 种蛋白质、至少 7,000 种蛋白质、至少 8,000 种蛋白质、至少 9,000 种蛋白质、或至少 10,000 种蛋白质。在一些实施方案中,数据库包含 100 至 500 种蛋白质、200 至 1,000 种蛋白质、500 至 1,000 种蛋白质、500 至 1,000 种蛋白质、1,000 至 2,000 种蛋白质、1,000 至 5,000 种蛋白质或 5,000 至 10,000 种蛋白质。

[0187] 对于一些实施方案,优选的是营养性蛋白质不展现抗营养活性(“抗营养性”),即蛋白质具有阻止自食物吸收营养物的潜力。抗营养性因子的实例包括蛋白酶抑制剂,其抑制肠中胰蛋白酶、胃蛋白酶和其它蛋白酶的作用,从而阻止蛋白质的消化和随后吸收。因此,在一些实施方案中,评估营养性蛋白质的潜在抗营养性。这可通过本领域中已知的任何适合方法进行。在一些实施方案中,抗营养性计分是通过确定蛋白质与已知蛋白酶抑制剂的数据库(例如自 UNIPROT 数据库鉴定的蛋白酶抑制剂)的同一性百分比来计算。目标蛋白质相对于已知蛋白酶抑制剂的数据库的整体-整体比对是使用采用 BLOSUM50 取代矩阵、缺口开放罚分 10 和缺口延伸罚分 2 的 FASTA 算法来进行,以鉴定营养性蛋白质是否与已知抗营养性蛋白质同源。在营养性蛋白质的一些实施方案中,营养性蛋白质与用于分析的数据库中的任何已知抗营养性蛋白质(例如,任何已知蛋白酶抑制剂)具有小于 35% 整体同源性。在一些实施方案中,使用小于 35% 同一性的截断值。在一些实施方案中,使用 30% 至 35% 的截断值。在一些实施方案中,使用 25% 至 35% 的截断值。在一些实施方案中,使用 20% 至 35% 的截断值。在一些实施方案中,使用 15% 至 35% 的截断值。在一些实施方案中,使用 10% 至 35% 的截断值。在一些实施方案中,使用 5% 至 35% 的截断值。在一些实施方案中,使用 0% 至 35% 的截断值。在一些实施方案中,使用大于 35% 同一性的截断值。在一些实施方案中,使用 35% 至 40% 的截断值。在一些实施方案中,使用 35% 至 45% 的截断值。在一些实施方案中,使用 35% 至 50% 的截断值。在一些实施方案中,使用 35% 至 55% 的截断值。在一些实施方案中,使用 35% 至 60% 的截断值。在一些实施方案中,使用 35% 至 70% 的截断值。在一些实施方案中,使用 35% 至 75% 的截断值。在一些实施方案中,使用 35% 至 80% 的截断值。熟练技术人员能够出于这个目的鉴定和使用已知蛋白酶抑制剂的适合数据库,例如通过搜索 UNIPROT 数据库(<http://www.uniprot.org>)。在一些实施方案中,数据库是通过自一个以上数据库来源选择鉴定为蛋白酶抑制剂的蛋白质加以定制。在一些实施方案中,数据库包含可在数据库中获得的一小组已知蛋白酶抑制剂;也就是说数据库是一小组定制选择的已知蛋白酶抑制剂蛋白质。在一些实施方案中,已知蛋白酶抑制剂蛋白质的数据库包含至少 10 种蛋白质、至少 20 种蛋白质、至少 30 种蛋白质、至少 40 种蛋白质、至少 50 种蛋白质、至少 100 种蛋白质、至少 200 种蛋白质、至少 300 种蛋白质、至少 400 种蛋白质、至少 500 种蛋白质、至少 600 种蛋白质、至少 700 种蛋白质、至少 800 种蛋白质、至少 900 种蛋白质、至少 1,000 种蛋白质、至少 1,100 种蛋白质、至少 1,200 种蛋白质、至少 1,300 种蛋白质、至少 1,400 种蛋白质、至少 1,500 种蛋白质、至少 1,600 种蛋白质、至少 1,700 种蛋白质、至少 1,800 种蛋白质、至少 1,900 种蛋白质、或至少 2,000 种蛋白质。在一些实施方案中,已知蛋白酶抑制剂蛋白质的数据库包含 100 至 500 种蛋白质、200 至 1,000 种蛋白质、500 至 1,000 种蛋白质、500 至 1,000 种蛋白质、或 1,000 至 2,000 种蛋白质、或 2,000 至 3,000 种蛋白质。

[0188] 在其它实施方案中,使用展现一定程度的蛋白酶抑制剂活性的营养性蛋白质。举例来说,在一些实施方案中,所述蛋白质可为适用的,因为当营养性蛋白质消耗时,它延迟



蛋白酶消化以使营养性蛋白质在被消化之前在胃肠 (GI) 道内经过较大距离,由此延迟吸收。举例来说,在一些实施方案中,营养性蛋白质抑制胃消化而非肠消化。

[0189] Delaney B. 等 (Evaluation of protein safety in the context of agricultural biotechnology. Food. Chem. Toxicol. 46(2008:S71-S97)) 建议当评估可能的食物蛋白质的安全性时,应避免已知毒性蛋白质与抗营养性蛋白质两者。在营养性蛋白质的一些实施方案中,营养性蛋白质与某一数据库的已知毒性蛋白质具有有利较低水平的整体同源性和/或与某一数据库的已知抗营养性蛋白质(例如,蛋白酶抑制剂)具有有利较低水平的整体同源性,如本文所定义。

[0190] 可增强营养性蛋白质的效用的一个特征是它的电荷(或每个氨基酸电荷)。在一些实施方案中,电荷较高的营养性蛋白质可展现合乎需要的特征,如增加的溶解度、增加的稳定性、抗聚集性和合乎需要的味道分布。举例来说,展现溶解度增强的带电荷营养性蛋白质可被配制成呈相对较低体积的溶液形式的包括高浓度营养性蛋白质的饮料或液体制剂,由此递送每单位体积较大剂量的蛋白质营养物。展现溶解度增强的带电荷营养性蛋白质可例如适用于其中使用者(例如运动员)想要在身体活动之前、期间或之后摄取营养性蛋白质的运动饮料或恢复饮料中。展现溶解度增强的带电荷营养性蛋白质也可特别适用于其中受试者(例如患者或年长人士)需要蛋白质营养物,但不能摄取固体食物或大体积的液体的临床环境中。

[0191] 举例来说,多肽在 pH 7 下的净电荷(电荷 p)可使用下式来计算:

[0192] 电 荷  $P = -0.002 - (C)(0.045) - (D)(0.999) - (E)(0.998) + (H)(0.091) + (K)(1.0) + (R)(1.0) - (Y)(-0.001)$

[0193] 其中 C 是所述多肽中的半胱氨酸残基的数目, D 是天冬氨酸残基的数目, E 是谷氨酸残基的数目, H 是组氨酸残基的数目, K 是赖氨酸残基的数目, R 是精氨酸残基的数目且 Y 是酪氨酸残基的数目。多肽的每个氨基酸电荷(电荷 A)可通过用净电荷(电荷 P)除以氨基酸残基的数目(N)来计算,即电荷  $A = \text{电荷 } P / N$ 。(参见 Bassi S(2007), "A Primer on Python for Life Science Researchers." PLoS Comput Biol 3(11):e199. doi:10.1371/journal.pcbi.0030199)。

[0194] 一种用于评估给定蛋白质的亲水性和潜在溶解度的量度是溶合计分。如果各残基都独立地溶合,那么溶合计分定义为所有氨基酸侧链的用序列中的残基总数加以标准化的总溶合自由能(即与自气相转换至稀释溶液相关的自由能变化)。侧链溶合自由能是通过计算真空介电常数 1 与水介电常数 80 之间的静电能量差异(通过解析泊松-玻尔兹曼方程 (Poisson-Boltzmann equation)), 以及使用线性溶剂可及表面积模型计算非极性范德华 (Van der Waals) 能量来计算得到 (D. Sitkoff, K. A. Sharp, B. Honig. "Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models". J. Phys. Chem. 98, 1994)。对于具有可电离侧链的氨基酸 (Arg、Asp、Cys、Glu、His、Lys 和 Tyr), 平均溶合自由能是基于在指定 pH 下, 各离子化状态的相对概率加以使用。溶合计分开始于 0 且继续以变为负值, 且溶合计分负性越大, 预测蛋白质的亲水性和潜在可溶性越大。在营养性蛋白质的一些实施方案中, 营养性蛋白质在 pH 7 下具有溶合计分为 -10 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在 pH 7 下具有溶合计分为 -15 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在 pH 7 下具有溶合计分为 -20 或更

低。在营养性蛋白质的一些实施方案中,营养性蛋白质在 pH 7 下具有溶合计分为 -25 或更低。在营养性蛋白质的一些实施方案中,营养性蛋白质在 pH 7 下具有溶合计分为 -30 或更低。在营养性蛋白质的一些实施方案中,营养性蛋白质在 pH 7 下具有溶合计分为 -35 或更低。在营养性蛋白质的一些实施方案中,营养性蛋白质在 pH 7 下具有溶合计分为 -40 或更低。

[0195] 由于如由汉德森-海森巴赫 (Henderson-Hasselbalch) 方程所定义, pH 依赖于未解离弱酸 ([HA]) 与共轭碱 ([A<sup>-</sup>]) 的摩尔比, 所以溶合计分随 pH 而变:

[0196] 
$$pH = pKa + \log \left( \frac{[A^-]}{[HA]} \right)$$

[0197] 所有弱酸相较于它们的共轭碱都具有不同溶合自由能, 且当计算给定 pH 下的溶合计时用于给定残基的溶合自由能是那两个值的加权平均值。

[0198] 因此, 在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -10 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -15 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -20 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -25 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -30 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -35 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在酸性 pH 下具有溶合计分为 -40 或更低。

[0199] 因此, 在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -10 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -15 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -20 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -25 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -30 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -35 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在碱性 pH 下具有溶合计分为 -40 或更低。

[0200] 因此, 在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶合计分为 -10 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶合计分为 -15 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶合计分为 -20 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶合计分为 -25 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶合计分为 -30 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶合计分为 -35 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质在选自 2-3、3-4、4-5、5-6、6-7、7-8、8-9、9-10、10-11、和 11-12 的 pH 范围下具有溶

合计分为 -40 或更低。

[0201] 聚集计分是用于评估给定蛋白质的疏水性和聚集可能性的基于一级序列的量度。使用赋予疏水性残基正值且赋予亲水性残基负值的 Kyte 和 Doolittle 疏水性量表 (Kyte J, Doolittle RF (1982 年 5 月) “A simple method for displaying the hydropathic character of a protein”. J. Mol. Biol. 157(1):105-32), 利用五个残基的移动平均值计算蛋白质序列的平均疏水性。聚集计分是通过确定大于零的数值的曲线下面积且用蛋白质的总长度加以标准化来自所得曲线获取。潜伏假设是聚集是两个或更多个疏水性片块汇集来一起以排除水且降低表面暴露的结果, 且蛋白质将聚集的可能性随它的疏水性 (即聚集倾向性) 残基密集压紧程度而变。聚集计分开始于 0 且继续以变为正值, 且聚集计分越小, 预测蛋白质的疏水性越小且潜在聚集倾向越小。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 2 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 1.5 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 1 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.9 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.8 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.7 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.6 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.5 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.4 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.3 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.2 或更低。在营养性蛋白质的一些实施方案中, 营养性蛋白质具有聚集计分为 0.1 或更低。

[0202] 在一些情况下, 可溶性表达是合乎需要的, 因为它可增加营养性蛋白质的量和 / 或产率且有助于营养性蛋白质的分离和纯化的一者或多者。在一些实施方案中, 本公开的营养性蛋白质在宿主生物体中可溶性表达。溶合计分和聚集计分可用于预测重组营养性蛋白质在宿主生物体中的可溶性表达。如实施例 8 中所示, 本公开提供证据表明溶合计分  $\leq -20$  且聚集计分  $\leq 0.75$  的营养性蛋白质更可能在特定大肠杆菌表达系统中重组表达。此外, 数据也表明溶合计分  $\leq -20$  且聚集计分  $\leq 0.5$  的营养性蛋白质更可能在这个系统中可溶性表达。因此, 在一些实施方案中, 本公开的营养性蛋白质具有溶合计分为 -20 或更低。在一些实施方案中, 营养性蛋白质具有聚集计分为 0.75 或更低。在一些实施方案中, 营养性蛋白质具有聚集计分为 0.5 或更低。在一些实施方案中, 营养性蛋白质具有溶合计分为 -20 或更低和聚集计分为 0.75 或更低。在一些实施方案中, 营养性蛋白质具有溶合计分为 -20 或更低和聚集计分为 0.5 或更低。

[0203] 已知某些游离氨基酸和游离氨基酸的混合物具有苦味或另外不合意味道。此外, 常见蛋白质 (例如, 乳清和大豆) 的水解产物常具有苦味或不合意味道。在一些实施方案中, 本文公开和所述的营养性蛋白质不具有苦味或另外不合意味道。在一些实施方案中, 本文公开和所述的营养性蛋白质相较于游离氨基酸、游离氨基酸的混合物和 / 或蛋白质水解产物的至少一者具有更可接受的味道。在一些实施方案中, 本文公开和所述的营养性蛋白质具有的味道等于或超过至少一种乳清蛋白质。

[0204] 已知蛋白质具有涵盖以下五种确定味道形式的味道: 甜味、酸味、苦味、咸味和鲜

味。特定蛋白质的味道（或所述蛋白质缺乏所述味道）可归因于若干因素，包括一级结构、存在带电荷侧链和蛋白质的电子和构象特征。在一些实施方案中，本文公开和所述的营养性蛋白质被设计来具有所需味道（例如，甜味、咸味、鲜味）和 / 或不具有非所需味道（例如，苦味、酸味）。在这个情形下，“设计”包括例如选择体现达成所需味道性质的特征天然存在的蛋白质，以及产生天然存在的蛋白质的具有所需味道性质的突变蛋白。举例来说，营养性蛋白质可被设计来与特定味道受体，如甜味受体（T1R2-T1R3 杂二聚体）或鲜味受体（T1R1-T1R3 杂二聚体、mGluR4 和 / 或 mGluR1）相互作用。此外，营养性蛋白质可被设计来不与其它味道受体（如苦味受体（T2R 受体）相互作用或与其它味道受体具有减弱的相互作用。

[0205] 本文公开和所述的营养性蛋白质在被摄取时也可引发口中不同体感，有时称为“口感”。营养性蛋白质的口感可归因于一个或多个因素，包括一级结构、存在带电荷侧链、以及蛋白质的电子和构象特征。在一些实施方案中，营养性蛋白质在被摄取时引发黄油或脂肪样口感。

[0206] 在一些实施方案中，营养性蛋白质包含 20 至 5,000 个氨基酸、20-2,000 个氨基酸、20-1,000 个氨基酸、20-500 个氨基酸、20-250 个氨基酸、20-200 个氨基酸、20-150 个氨基酸、20-100 个氨基酸、20-40 个氨基酸、30-50 个氨基酸、40-60 个氨基酸、50-70 个氨基酸、60-80 个氨基酸、70-90 个氨基酸、80-100 个氨基酸、至少 25 个氨基酸、至少 30 个氨基酸、至少 35 个氨基酸、至少 40 个氨基酸、至少 45 个氨基酸、至少 50 个氨基酸、至少 55 个氨基酸、至少 60 个氨基酸、至少 65 个氨基酸、至少 70 个氨基酸、至少 75 个氨基酸、至少 80 个氨基酸、至少 85 个氨基酸、至少 90 个氨基酸、至少 95 个氨基酸、至少 100 个氨基酸、至少 105 个氨基酸、至少 110 个氨基酸、至少 115 个氨基酸、至少 120 个氨基酸、至少 125 个氨基酸、至少 130 个氨基酸、至少 135 个氨基酸、至少 140 个氨基酸、至少 145 个氨基酸、至少 150 个氨基酸、至少 155 个氨基酸、至少 160 个氨基酸、至少 165 个氨基酸、至少 170 个氨基酸、至少 175 个氨基酸、至少 180 个氨基酸、至少 185 个氨基酸、至少 190 个氨基酸、至少 195 个氨基酸、至少 200 个氨基酸、至少 205 个氨基酸、至少 210 个氨基酸、至少 215 个氨基酸、至少 220 个氨基酸、至少 225 个氨基酸、至少 230 个氨基酸、至少 235 个氨基酸、至少 240 个氨基酸、至少 245 个氨基酸、或至少 250 个氨基酸。在一些实施方案中，营养性蛋白质由 20 至 5,000 个氨基酸、20-2,000 个氨基酸、20-1,000 个氨基酸、20-500 个氨基酸、20-250 个氨基酸、20-200 个氨基酸、20-150 个氨基酸、20-100 个氨基酸、20-40 个氨基酸、30-50 个氨基酸、40-60 个氨基酸、50-70 个氨基酸、60-80 个氨基酸、70-90 个氨基酸、80-100 个氨基酸、至少 25 个氨基酸、至少 30 个氨基酸、至少 35 个氨基酸、至少 40 个氨基酸、至少 45 个氨基酸、至少 50 个氨基酸、至少 55 个氨基酸、至少 60 个氨基酸、至少 65 个氨基酸、至少 70 个氨基酸、至少 75 个氨基酸、至少 80 个氨基酸、至少 85 个氨基酸、至少 90 个氨基酸、至少 95 个氨基酸、至少 100 个氨基酸、至少 105 个氨基酸、至少 110 个氨基酸、至少 115 个氨基酸、至少 120 个氨基酸、至少 125 个氨基酸、至少 130 个氨基酸、至少 135 个氨基酸、至少 140 个氨基酸、至少 145 个氨基酸、至少 150 个氨基酸、至少 155 个氨基酸、至少 160 个氨基酸、至少 165 个氨基酸、至少 170 个氨基酸、至少 175 个氨基酸、至少 180 个氨基酸、至少 185 个氨基酸、至少 190 个氨基酸、至少 195 个氨基酸、至少 200 个氨基酸、至少 205 个氨基酸、至少 210 个氨基酸、至少 215 个氨基酸、至少 220 个氨基酸、至少 225 个氨基酸、至少 230 个氨基酸、至

少 235 个氨基酸、至少 240 个氨基酸、至少 245 个氨基酸、或至少 250 个氨基酸组成。

#### [0207] B. 核酸

[0208] 本文也提供编码营养性多肽或蛋白质的核酸。在一些实施方案中,核酸是分离的核酸。在一些实施方案中,核酸是纯化核酸。

[0209] 在核酸的一些实施方案中,核酸包含编码以上章节 A 中公开的第一多肽序列的核酸序列。在核酸的一些实施方案中,核酸由编码以上章节 A 中公开的第一多肽序列的核酸序列组成。在核酸的一些实施方案中,核酸包含编码以上章节 A 中公开的营养性蛋白质的核酸序列。在核酸的一些实施方案中,核酸由编码以上章节 A 中公开的营养性蛋白质的核酸序列组成。在核酸的一些实施方案中,编码第一多肽序列的核酸序列可操作地连接于至少一种表达控制序列。举例来说,在核酸的一些实施方案中,编码第一多肽序列的核酸序列可操作地连接于启动子,如以下章节 D 中所述的启动子。

[0210] 因此,在一些实施方案中,本公开的核酸分子编码自身是营养性多肽或蛋白质的多肽或蛋白质。所述核酸分子可称为“营养性核酸”。在一些实施方案中,营养性核酸编码自身包含以下至少一者的多肽或蛋白质:a) 支链氨基酸残基与总氨基酸残基的比率为至少 24%;b) Leu 残基与总氨基酸残基的比率为至少 11%;和 c) 必需氨基酸残基与总氨基酸残基的比率为至少 49%。在一些实施方案中,营养性核酸包含至少 10 个核苷酸、至少 20 个核苷酸、至少 30 个核苷酸、至少 40 个核苷酸、至少 50 个核苷酸、至少 60 个核苷酸、至少 70 个核苷酸、至少 80 个核苷酸、至少 90 个核苷酸、至少 100 个核苷酸、至少 200 个核苷酸、至少 300 个核苷酸、至少 400 个核苷酸、至少 500 个核苷酸、至少 600 个核苷酸、至少 700 个核苷酸、至少 800 个核苷酸、至少 900 个核苷酸、至少 1,000 个核苷酸。在一些实施方案中,营养性核酸包含 10 至 100 个核苷酸、20 至 100 个核苷酸、10 至 50 个核苷酸、或 20 至 40 个核苷酸。在一些实施方案中,营养性核酸包含编码天然存在的营养性多肽或蛋白质的开放阅读框的全部或一部分。在一些实施方案中,营养性核酸由编码天然存在的营养性蛋白质的片段的开放阅读框组成,其中所述开放阅读框不编码完整天然存在的营养性蛋白质。

[0211] 在一些实施方案中,营养性核酸是 cDNA。

[0212] 在一些实施方案中,提供包含与天然存在的营养性核酸具有至少 50%、60%、70%、80%、85%、90%、95%、96%、97%、98%、99% 或 99.9% 同一性的序列的核酸分子。在一些实施方案中,提供在严格杂交条件下与至少一个参照营养性核酸杂交的核酸。

[0213] 本公开中提供的营养性核酸及其片段显示在多种系统和方法中具有效用。举例来说,片段可在各种杂交技术中用作探针。视方法而定,靶标核酸序列可为 DNA 或 RNA。靶标核酸序列可在杂交之前加以分级分离(例如,通过凝胶电泳),或可对样本进行原位杂交。本领域技术人员应了解的是具有已知序列的核酸探针在确定染色体结构(例如,通过 DNA 印迹)和测量基因表达(例如,通过 RNA 印迹)方面具有效用。在所述实验中,序列片段优选被可检测标记,以使可检测以及任选定量它们与靶标序列的特异性杂交。本领域技术人员应了解的是本公开的核酸片段可用于本文未明确描述的广泛多种印迹技术中。

[0214] 也应了解的是当固定在微阵列上时,本文公开的核酸序列片段也适用作探针。用于通过将核酸沉积和固定于支撑基板上产生微阵列的方法在本领域中是熟知的。综述于以下中:DNA Microarrays:A Practical Approach (Practical Approach Series), Schena(编), Oxford University Press(1999) (ISBN:0199637768);

Nature Genet. 21(1) (增刊):1-60(1999); Microarray Biochip: Tools and Technology, Schena(编), Eaton Publishing Company/BioTechniques Books Division(2000) (ISBN:1881299376), 其公开内容以引用的方式整体并入本文。使用包含核酸序列片段(如本文公开的核酸序列片段)的微阵列分析例如基因表达在细胞和分子生物学领域中是序列片段的明确确定效用。固定在微阵列上的序列片段的其它用途描述于以下中: Gerhold 等, Trends Biochem. Sci. 24:168-173(1999) 和 Zweiger, Trends Biotechnol. 17:429-436(1999); DNA Microarrays: A Practical Approach (Practical Approach Series), Schena(编), Oxford University Press(1999) (ISBN:0199637768); Nature Genet. 21(1) (增刊):1-60(1999); Microarray Biochip: Tools and Technology, Schena(编), Eaton Publishing Company/BioTechniques Books Division(2000) (ISBN:1881299376)。

#### [0215] C. 载体

[0216] 也提供包含至少一种本文公开的核酸分子的载体, 包括表达载体, 如本文进一步所述。在一些实施方案中, 载体包含至少一种编码如本文公开的营养性蛋白质的分离的核酸分子。在替代性实施方案中, 载体包含可操作地连接于一种或多种表达控制序列的所述核酸分子。载体可因此用于在重组微生物宿主细胞中表达至少一种重组蛋白。

[0217] 适用于在微生物中表达核酸的载体为本领域技术人员所熟知。适用于蓝细菌中的载体例如描述于 Heidorn 等, “Synthetic Biology in Cyanobacteria: Engineering and Analyzing Novel Functions,” Methods in Enzymology, 第 497 卷, 第 24 章 (2011) 中。可用于使如本文公开的蓝细菌工程化的示例性复制型载体包括 pPMQAK1、pSL1211、pFC1、pSB2A、pSCR119/202、pSUN119/202、pRL2697、pRL25C、pRL1050、pSG111M 和 pPBH201。

[0218] 也可使用能够接受本文公开的核酸序列的其它载体, 如 pJB161。如 pJB161 的载体包含与某些光合微生物内源性质粒(例如, 某些聚球藻属种 (*Synechococcus* species) 的质粒 pAQ1、pAQ3 和 pAQ4) 中存在的序列同源的序列。所述载体的实例以及如何使用它们在本领域中是已知的且例如提供于 Xu 等, “Expression of Genes in Cyanobacteria: Adaptation of Endogenous Plasmids as Platforms for High-Level Gene Expression in *Synechococcus* sp. PCC 7002,” 第 21 章, Robert Carpentier(编), “Photosynthesis Research Protocols,” Methods in Molecular Biology, 第 684 卷, 2011 中, 所述文献据此以引用的方式并入本文。在 pJB161 与内源性质粒之间进行体内重组产生自它们的内源性质粒表达目标基因的工程化微生物。或者, 载体可被工程化来与宿主细胞染色体重组, 或载体可被工程化来独立于宿主细胞染色体或任何宿主细胞的内源性质粒来复制和表达目标基因。

[0219] 适于重组蛋白产生的载体的另一实例是 pET 系统(Novagen®)。这个系统已被广泛表征以用于大肠杆菌和其它微生物中。在这个系统中, 将靶标基因在强噬菌体 T7 转录和(任选)翻译信号控制下克隆在 pET 质粒中; 通过在宿主细胞中提供 T7 RNA 聚合酶来源来诱导表达。T7 RNA 聚合酶具有如此选择性和活性以致当被充分诱导时, 几乎所有微生物的资源都被转向靶标基因表达; 在诱导之后数小时, 所需产物可占总细胞蛋白质的 50% 以上。也有可能仅通过降低诱导剂的浓度来减弱表达水平。降低表达水平可增强一些靶标蛋白质的可溶性产率。在一些实施方案中, 这个系统也允许维持靶标基因处于转录沉默未诱

导状态。

[0220] 在使用这个系统的一些实施方案中,靶标基因是使用不含有 T7 RNA 聚合酶基因的宿主来克隆,由此减轻与归因于产生对宿主细胞具有潜在毒性的蛋白质的质粒不稳定性相关的潜在问题。一旦在非表达宿主中建立,靶标蛋白质表达即可通过用  $\lambda$  CE6 (一种携带在  $\lambda$  pL 和 pI 启动子控制下的 T7 RNA 聚合酶基因的噬菌体) 感染宿主,或通过将质粒转移至含有在 lacUV5 控制下的 T7 RNA 聚合酶基因的染色体拷贝的表达宿主中来起始。在第二种情况下,表达是通过向细菌培养物中添加 IPTG 或乳糖或使用自身诱导培养基来诱导。由 lac 操纵子控制,但不需要 T7 RNA 聚合酶基因且依赖于大肠杆菌的天然 RNA 聚合酶的其它质粒系统包括 pTrc 质粒套件 (Invitrogen) 或 pQE 质粒套件 (QIAGEN)。

[0221] 在其它实施方案中,有可能直接克隆至表达宿主中。在它们抑制基础表达水平的严格性方面不同的两种类型的 T7 启动子和若干宿主是可用的,从而提供使广泛多种靶标基因的表达优化的巨大灵活性和能力。

[0222] 适用于在哺乳动物细胞中表达核酸的载体通常包含由病毒调控元件提供的控制功能。举例来说,通常使用的启动子是源于多形瘤病毒、腺病毒 2、巨细胞病毒或猿猴病毒 40。

#### [0223] D. 启动子

[0224] 适用于表达本文所述的重组基因的启动子包括组成型启动子与诱导型 / 可抑制性启动子两者。诱导型 / 可抑制性启动子的实例包括镍诱导型启动子 (例如, PnrsA、PnrsB; 参见例如 Lopez-Mauy 等, Cell (2002) 第 43 卷: 247-256) 和脲可抑制性启动子, 如 PnirA (例如描述于 Qi 等, Applied and Environmental Microbiology (2005) 第 71 卷: 5678-5684 中)。诱导型 / 可抑制性启动子的其它实例包括 PnirA (驱动 nirA 基因的表达的启动子, 由硝酸盐诱导且由脲阻遏) 和 Psuf (驱动 sufB 基因的表达的启动子, 由铁应激诱导)。组成型启动子的实例包括 Pcpc (驱动 cpc 操纵子的表达的启动子)、Prbc (驱动二磷酸核酮糖羧化酶的表达的启动子)、PpsbAII (驱动 PpsbAII 的表达的启动子)、Pcro (驱动 cro 的表达的  $\lambda$  噬菌体启动子)。在其它实施方案中, PaphII 和 / 或 lacIq-Ptrc 启动子可用于控制表达。当在工程化微生物中表达多种重组基因时, 不同基因可由不同启动子或由单独操纵子中的相同启动子控制, 或两种或更多种基因的表达可由作为操纵子的一部分的单一启动子控制。

[0225] 诱导型启动子的其它非限制性实例可包括但不限于通过表达外源性蛋白质 (例如, T7 RNA 聚合酶、SP6RNA 聚合酶), 通过存在小分子 (例如, IPTG、半乳糖、四环素、类固醇激素、脱落酸 (abscisic acid)), 通过不存在小分子 (例如, CO<sub>2</sub>、铁、氮), 由金属或金属离子 (例如, 铜、锌、镉、镍), 以及由环境因素 (例如, 热、冷、应激、光、黑暗) 和由生长期加以诱导的启动子。在一些实施方案中, 诱导型启动子受严密调控以使在不存在诱导下, 大致上无转录通过启动子被起始。在一些实施方案中, 诱导启动子不实质上改变通过其它启动子进行的转录。此外, 一般说来, 诱导型启动子的化合物或条件不天然存在于寻求表达所处的生物体或环境中。

[0226] 在一些实施方案中, 诱导型启动子是通过限制向蓝细菌培养物供应 CO<sub>2</sub> 来诱导。作为非限制性实例, 诱导型启动子可为在 CO<sub>2</sub> 限制条件下上调的集胞藻属 (Synechocystis) PCC 6803 (如 cmp 基因、ntp 基因、ndh 基因、sbt 基因、chp 基因和 rbc 基因或其变体或片

段)的启动子序列。

[0227] 在一些实施方案中,诱导型启动子是通过铁饥饿或通过进入稳定生长期来诱导。在一些实施方案中,诱导型启动子可为在Fe饥饿条件下上调的蓝细菌基因(如*isiA*)或当培养物进入稳定生长期时上调的蓝细菌基因(如*isiA*、*phrA*、*sigC*、*sigB*和*sigH*基因或其变体或片段)的启动子序列的变体序列。

[0228] 在一些实施方案中,诱导型启动子由金属或金属离子诱导。作为非限制性实例,诱导型启动子可由铜、锌、镉、汞、镍、金、银、钴和铋或其离子诱导。在一些实施方案中,诱导型启动子由镍或镍离子诱导。在一些实施方案中,诱导型启动子由如Ni<sup>2+</sup>的镍离子诱导。在另一示例性实施方案中,诱导型启动子是来自集胞藻属PCC 6803的镍诱导性型启动子。在另一实施方案中,诱导型启动子可由铜或铜离子诱导。在另一实施方案中,诱导型启动子可由锌或锌离子诱导。在另一实施方案中,诱导型启动子可由镉或镉离子诱导。在另一实施方案中,诱导型启动子可由汞或汞离子诱导。在一替代性实施方案中,诱导型启动子可由金或金离子诱导。在另一替代性实施方案中,诱导型启动子可由银或银离子诱导。在另一替代性实施方案中,诱导型启动子可由钴或钴离子诱导。在另一替代性实施方案中,诱导型启动子可由铋或铋离子诱导。

[0229] 在一些实施方案中,启动子通过使包含诱导型启动子的细胞暴露于金属或金属离子来诱导。可通过向微生物生长培养基中添加金属来使细胞暴露于金属或金属离子。在某些实施方案中,添加至微生物生长培养基中的金属或金属离子可自培养基高效回收。在其它实施方案中,在回收之后剩余在培养基中的金属或金属离子不实质上阻碍对培养基或细菌基因产物的下游加工。

[0230] 组成型启动子的其它非限制性实例包括来自革兰氏阴性细菌(Gram-negative bacteria)或在革兰氏阴性细菌中增殖的噬菌体的组成型启动子。举例来说,可使用用于编码高度表达的革兰氏阴性基因产物的基因的启动子,如*Lpp*、*OmpA*、*rRNA*和核糖体蛋白质的启动子。或者,可调控启动子可用于缺乏那个启动子的调控性蛋白质的菌株中。举例来说,*P<sub>lac</sub>*、*P<sub>tac</sub>*和*P<sub>trc</sub>*可用作缺乏*LacI*的菌株中的组成型启动子。类似地,*P<sub>22P<sub>R</sub></sub>*和*P<sub>L</sub>*可用于缺乏 $\lambda$ C2阻遏蛋白的菌株中,且 $\lambda$ *P<sub>R</sub>*和*P<sub>L</sub>*可用于缺乏 $\lambda$ C1阻遏蛋白的菌株中。在一个实施方案中,组成型启动子来自噬菌体。在另一实施方案中,组成型启动子来自沙门氏菌属(*Salmonella*)噬菌体。在另一实施方案中,组成型启动子来自蓝绿藻噬菌体。在一些实施方案中,组成型启动子是集胞藻属启动子。举例来说,组成型启动子可为*PpsbA11*启动子或它的变体序列、*Prbc*启动子或它的变体序列、*P<sub>cp</sub>*启动子或它的变体序列、和*PrnpB*启动子或它的变体序列。

#### [0231] E. 宿主细胞

[0232] 也提供用本文公开的核酸分子或载体转化的宿主细胞及其后代。在一些实施方案中,宿主细胞是微生物细胞。在一些实施方案中,宿主细胞携带在可为但无需为自由复制载体的载体上的核酸序列。在其它实施方案中,核酸已整合至宿主细胞的基因组中和/或宿主细胞的内源性质粒中。转化的宿主细胞例如适用于产生本文公开的重组营养性蛋白质。

[0233] “微生物”包括来自古细菌域、细菌域和真核(Eucarya)域的原核和真核微生物物种,真核域包括酵母和丝状真菌、原虫、藻类或高等原生生物。术语“微生物细胞”和“微生物(microbe)”可与术语微生物(microorganism)互换使用。



[0234] 多种宿主微生物可用本文公开的核酸序列转化,且在一些实施方案中,可用于产生本文公开的重组营养性蛋白质。适合的宿主微生物包括自养性微生物与异养性微生物两者。在一些应用中,自养性微生物允许降低为制备由引入宿主微生物中的重组核酸序列编码的营养性蛋白质所需的矿物燃料和/或电输入量。在一些应用中,这又会降低产生营养性蛋白质的成本和/或环境影响,和/或相较于制造替代性营养性蛋白质(如乳清、蛋和大豆)的成本和/或环境影响,降低成本和/或环境影响。举例来说,在一些实施方案中,使用如本文公开的宿主微生物制备本文公开的营养性蛋白质的成本和/或环境影响低于通过加工牛奶以适于人消耗的形式制备乳清蛋白质的成本和/或环境影响。

[0235] 异养生物的非限制性实例包括大肠杆菌 (*Escherichia coli*)、鼠伤寒沙门氏菌 (*Salmonella typhimurium*)、枯草芽孢杆菌 (*Bacillus subtilis*)、巨大芽孢杆菌 (*Bacillus megaterium*)、谷氨酸棒杆菌 (*Corynebacterium glutamicum*)、天蓝色链霉菌 (*Streptomyces coelicolor*)、变铅青链霉菌 (*Streptomyces lividans*)、委内瑞拉链霉菌 (*Streptomyces venezuelae*)、玫瑰孢链霉菌 (*Streptomyces roseosporus*)、弗雷德氏链霉菌 (*Streptomyces fradiae*)、灰色链霉菌 (*Streptomyces griseus*)、棒状链霉菌 (*Streptomyces calvuligerus*)、吸水链霉菌 (*Streptomyces hygroscopicus*)、普拉特链霉菌 (*Streptomyces platensis*)、糖多孢红霉菌 (*Saccharopolyspora erythraea*)、谷氨酸棒杆菌、黑曲霉 (*Aspergillus niger*)、构巢曲霉 (*Aspergillus nidulans*)、米曲霉 (*Aspergillus oryzae*)、土曲霉 (*Aspergillus terreus*)、酱油曲霉 (*Aspergillus sojae*)、产黄青霉 (*Penicillium chrysogenum*)、里氏木霉 (*Trichoderma reesei*)、丙酮丁醇梭菌 (*Clostridium acetobutylicum*)、贝季尔林斯基梭菌 (*Clostridium beijerinckii*)、热纤梭菌 (*Clostridium thermocellum*)、*Fusibacter paucivorans*、酿酒酵母 (*Saccharomyces cerevisiae*)、布拉迪酵母 (*Saccharomyces boulardii*)、巴斯德毕赤酵母 (*Pichia pastoris*) 和树干毕赤酵母 (*Pichia stipitis*)。

[0236] 光合自养性微生物包括真核藻类以及原核蓝细菌、绿色硫细菌、绿色非硫细菌、紫色硫细菌和紫色非硫细菌。

[0237] 也涵盖极端生物作为适合的生物体。所述生物体经受得起各种环境参数,如温度、辐射、压力、重力、真空、干燥、盐度、pH、氧张力和化学品。它们包括在 80°C 或超过 80°C 下生长的极端嗜热菌,如延胡索酸火叶菌 (*Pyrolobus fumarii*);在 60–80°C 之间生长的嗜热菌,如灰篮聚球藻 (*Synechococcus lividis*);在 15–60°C 之间生长的嗜温菌;和在 15°C 或低于 15°C 下生长的嗜冷菌,如嗜冷杆菌 (*Psychrobacter*) 和一些昆虫。辐射耐受性生物体包括耐辐射异常球菌 (*Deinococcus radiodurans*)。压力耐受性生物体包括耐受 130MPa 压力的嗜压生物。重量耐受性生物体包括嗜压生物。也涵盖超重(例如 >1g)、低重(例如 <1g)耐受性生物体。真空耐受性生物体包括节肢动物、昆虫、微生物和种子。干燥耐受性和脱水生活生物体包括嗜旱生物,如盐水卤虫 (*Artemia salina*);线虫、微生物、真菌和地衣。盐耐受性生物体包括嗜盐生物(例如,2–5M NaCl)盐杆菌科 (*Halobacteriaceae*) 和盐生杜氏藻 (*Dunaliella salina*)。pH 耐受性生物体包括嗜碱菌,如嗜盐碱杆菌属 (*Natronobacterium*)、坚强芽孢杆菌 (*Bacillus firmus*) OF4、螺旋藻属的种 (*Spirulina spp.*) (例如 pH>9);以及嗜酸菌,如温泉红藻 (*Cyanidium caldarium*)、铁原体属某种 (*Ferroplasma sp.*) (例如低 pH)。也涵盖如詹氏甲烷球菌 (*Methanococcus*

jannaschii) 的不能耐受  $O_2$  的厌氧菌 ;如需要  $O_2$  的梭菌属 (*Clostridium*) 和需氧菌的耐受一定  $O_2$  的微需氧菌。耐受纯  $CO_2$  的气体耐受性生物体包括温泉红藻且金属耐受性生物体包括金属耐受性生物,如嗜酸铁原体 (*Ferroplasma acidarmanus*) (例如 Cu、As、Cd、Zn)、雷尔氏菌属 (*Ralstonia*) 的种 CH34 (例如 Zn、Co、Cd、Hg、Pb)。Gross, Michael. Life on the Edge: Amazing Creatures Thriving in Extreme Environments. New York: Plenum (1998) 和 Seckbach, J. "Search for Life in the Universe with Terrestrial Microbes Which Thrive Under Extreme Conditions.", Cristiano Batalli Cosmovici, Stuart Bowyer 和 Dan Wertheimer 编, Astronomical and Biochemical Origins and the Search for Life in the Universe, 第 511 页. Milan: Editrice Compositori (1997)。

[0238] 混合营养性生物体也是适合的生物体。混合营养性生物体可利用混合不同来源的能量和碳,例如光营养和化学营养、无机和有机营养、自体营养和异体营养及其组合或它的组合。混合营养生物可为真核的或原核的。另外,混合营养生物可为专性的或兼性的。适合的混合营养性生物体包括混合营养性藻类和混合营养性细菌。

[0239] 藻类和蓝细菌包括但不限于以下属:刺菊石属 (*Acanthoceras*)、*Acanthococcus*、*Acaryochloris*、弯杆藻属 (*Achnanthes*)、翼硅藻属 (*Achnanthidium*)、星形藻属 (*Actinastrum*)、*Actinochloris*、辐环藻属 (*Actinocyclus*)、辐射鼓藻属 (*Actinotaenium*)、双金藻属 (*Amphichrysis*)、前沟藻属 (*Amphidinium*)、*Amphikrikos*、双肋藻属 (*Amphipleura*)、茧形藻属 (*Amphiprora*)、分须藻属 (*Amphithrix*)、双眉藻属 (*Amphora*)、鱼腥藻属 (*Anabaena*)、项圈藻属 (*Anabaenopsis*)、暗额藻属 (*Aneumastus*)、针连藻属 (*Ankistrodesmus*)、锚藻属 (*Ankyra*)、异菱藻属 (*Anomoeoneis*)、虚幻球藻属 (*Apatococcus*)、束丝藻属 (*Aphanizomenon*)、隐球藻属 (*Aphanocapsa*)、隐毛藻属 (*Aphanochaete*)、隐杆藻属 (*Aphanothece*)、梨囊藻属 (*Apiocystis*)、顶丝藻属 (*Apistonema*)、四棘鼓藻属 (*Arthrodesmus*)、节旋藻属 (*Artherospira*)、*Ascochloris*、星杆藻属 (*Asterionella*)、星球藻属 (*Asterococcus*)、奥氏藻属 (*Audouinella*)、浮生直链藻属 (*Aulacoseira*)、杆状藻属 (*Bacillaria*)、巴尔比亚藻属 (*Balbiana*)、似竹鼓藻属 (*Bambusina*)、红毛菜属 (*Bangia*)、*Basichlamys*、串珠藻属 (*Batrachospermum*)、骈胞藻属 (*Binuclearia*)、角藻属 (*Bitrichia*)、盘苔属 (*Blidingia*)、*Botridiopsis*、气球藻属 (*Botrydium*)、葡萄藻属 (*Botryococcus*)、球葡萄藻属 (*Botryosphaerella*)、咸胞藻属 (*Brachiomonas*)、短螺旋体属 (*Brachysira*)、海雹菜属 (*Brachytrichia*)、*Brebissonia*、毛鞘藻属 (*Bulbochaete*)、柱杆藻属 (*Bumilleria*)、拟杆藻属 (*Bumilleriopsis*)、美壁藻属 (*Caloneis*)、眉藻属 (*Calothrix*)、马鞍藻属 (*Campylodiscus*)、盒管藻属 (*Capsosiphon*)、四鞭藻属 (*Carteria*)、*Catena*、*Cavinula*、顶刺藻属 (*Centritractus*)、中脊贝属 (*Centronella*)、角藻属 (*Ceratium*)、角毛藻属 (*Chaetoceros*)、*Chaetochloris*、硬毛藻属 (*Chaetomorpha*)、*Chaetonella*、毛丝藻属 (*Chaetonema*)、盾毛藻属 (*Chaetopeltis*)、胶毛藻属 (*Chaetophora*)、毛球藻属 (*Chaetosphaeridium*)、管胞藻属 (*Chamaesiphon*)、轮藻属 (*Chara*)、*Characiochloris*、拟小桩藻属 (*Characiopsis*)、小桩藻属 (*Characium*)、轮藻目 (*Charales*)、缘胞藻属 (*Chilomonas*)、厚胞藻属 (*Chlainomonas*)、*Chlamydonepharis*、*Chlamydocapsa*、衣藻属 (*Chlamydomonas*)、*Chlamydomonopsis*、衣粘藻属 (*Chlamydomyxa*)、*Chlamydonephris*、*Chlorangiella*、拟绿囊藻属 (*Chlorangiopsis*)、小球藻属 (*Chlorella*)、

绿葡萄藻属 (Chlorobotrys)、绿幅藻属 (Chlorobrachis)、绿点藻属 (Chlorochytrium)、绿球藻属 (Chlorococcum)、绿胶藻属 (Chlorogloea)、拟绿胶藻属 (Chlorogloeopsis)、绿梭藻属 (Chlorogonium)、绿带藻属 (Chlorolobion)、拟衣藻属 (Chloromonas)、Chlorophysema、绿藻门 (Chlorophyta)、绿胶囊藻属 (Chlorosaccus)、Chlorosarcina、Choricystis、色植藻属 (Chromophyton)、单鞭金藻属 (Chromulina)、拟色球藻属 (Chroococcidiopsis)、色球藻属 (Chroococcus)、色指藻属 (Chroodactylon)、蓝隐藻属 (Chroomonas)、Chroothece、金变形藻属 (Chrysamoeba)、金网藻属 (Chrysapsis)、金星藻属 (Chrysidiastrium)、金囊藻属 (Chrysocapsa)、Chrysocapsella、Chrysochaete、金色藻属 (Chrysochromulina)、金颗藻属 (Chrysococcus)、Chrysocrinus、Chrysolepidomonas、Chrysolykos、Chrysonebula、金藻门 (Chrysophyta)、金中藻属 (Chrysopyxis)、Chrysosaccus、Chrysophaerella、金环藻属 (Chrysostephanosphaera)、刚毛藻属 (Clodophora)、链孢藻属 (Clastidium)、拟新月藻属 (Closteriopsis)、新月藻属 (Closterium)、胶球藻属 (Coccomyxa)、卵形藻属 (Cocconeis)、Coelastrella、空星藻属 (Coelastrum)、腔球藻属 (Coelosphaerium)、Coenochloris、无胶集球藻属 (Coenococcus)、聚囊藻属 (Coenocystis)、柄裸藻属 (Colacium)、鞘毛藻属 (Coleochaete)、Collodictyon、Compsogonopsis、弯枝藻属 (Compsopogon)、Conjugatophyta、Conochaete、Coronastrum、鼓藻属 (Cosmarium)、Cosmioneis、胶球鼓藻属 (Cosmocladium)、Crateriportula、Craticula、发毛针藻属 (Crinalium)、十字藻属 (Crucigenia)、卵胞藻属 (Crucigeniella)、Cryptoaulax、隐藻属 (Cryptomonas)、隐藻门 (Cryptophyta)、栉水母门 (Ctenophora)、Cymodictyon、Cyanonephron、Cyanophora、蓝藻门 (Cyanophyta)、蓝杆藻门 (Cyanothece)、Cyanothomonas、环胞藻属 (Cyclonexis)、环冠藻属 (Cyclostephanos)、小环藻属 (Cyclotella)、筒藻属 (Cylindrocapsa)、柱胞鼓藻属 (Cylindrocystis)、柱孢藻属 (Cylindrospermum)、细柱藻属 (Cylindrotheca)、波缘藻属 (Cymatopleura)、桥弯藻属 (Cymbella)、Cymbellonitzschia、胞甲藻属 (Cystodinium)、蓝纤维藻属 (Dactylococcopsis)、单板藻属 (Debarya)、细齿藻属 (Denticula)、Dermatochrysis、皮果藻属 (Dermocarpa)、果菌属 (Dermocarpella)、缢带藻属 (Desmatractum)、角丝鼓藻属 (Desmidium)、Desmococcus、带线藻属 (Desmonema)、Desmosiphon、长刺藻属 (Diacanthos)、Diacronema、Diadesmis、等片藻属 (Diatoma)、Diatomella、双细胞藻属 (Dicellula)、双须藻属 (Dichothrix)、叉球藻属 (Dichotomococcus)、Dicranochaete、网绿藻属 (Dictyochloris)、球网藻属 (Dictyococcus)、胶网藻属 (Dictyosphaerium)、泡双吸虫属 (Didymocystis)、Didymogenes、异形藻属 (Didymosphenia)、丝藻属 (Dilabifilum)、双形藻属 (Dimorphococcus)、锥囊藻属 (Dinobryon)、球甲藻属 (Dinococcus)、双绿藻属 (Diplochloris)、双壁藻属 (Diploneis)、Diplostauron、Distrionella、基纹鼓藻属 (Docidium)、竹枝藻属 (Draparnaldia)、杜氏藻属 (Dunaliella)、孔壳藻属 (Dysmorphococcus)、Eckillocystis、纺锤藻属 (Elakatothrix)、Ellerbeckia、双眉藻属 (Encyonema)、浒苔属 (Enteromorpha)、内枝藻属 (Entocladia)、Entomoneis、石囊藻属 (Entophysalis)、附金藻属 (Epichrysis)、附钟藻属 (Epipyxis)、窗纹藻属 (Epithemia)、独球藻属 (Eremosphaera)、Euastropsis、凹顶鼓藻属 (Euastrum)、立方藻属 (Eucapsis)、真卵形藻属 (Eucocconeis)、空球藻属 (Eudorina)、裸藻属 (Euglena)、裸藻门

(Euglenophyta)、短缝藻属 (Eunotia)、Eustigmatophyta、双鞭藻属 (Eutreptia)、Fallacia、费氏藻属 (Fischerella)、脆杆藻属 (Fragilaria)、Fragilariforma、披刺藻属 (Franceia)、肋缝藻属 (Frustulia)、Curcilla、双胞藻属 (Geminella)、短水绵属 (Genicularia)、灰胞藻属 (Glaucocystis)、灰色藻门 (Glaucophyta)、Glenodiniopsis、薄甲藻属 (Glenodinium)、粘球藻属 (Gloeocapsa)、Gloeochaete、Gloeochrysis、圆球藻属 (Gloeococcus)、胶囊藻属 (Gloeocystis)、Gloeodendron、胶胞藻属 (Gloeomonas)、Gloeoplax、粘杆藻属 (Gloeotheca)、胶丝藻属 (Gloeotila)、胶刺藻属 (Gloeotrichia)、Gloiodictyon、多芒藻属 (Golenkinia)、拟多芒藻属 (Golenkiniopsis)、抱根藻属 (Gomontia)、楔桥弯藻属 (Gomphocymbella)、异极藻属 (Gomphonema)、束球藻属 (Gomphosphaeria)、棒形鼓藻属 (Gonatozygon)、Gongrosia、链瘤藻属 (Gongrosira)、角绿藻属 (Goniochloris)、盘藻属 (Gonium)、膝口藻属 (Gonyostomum)、粒绿藻属 (Granulochloris)、Granulocystopsis、Groenbladia、裸甲藻属 (Gymnodinium)、缢丝鼓藻属 (Gymnozyga)、布纹藻属 (Gyrosigma)、红球藻属 (Haematococcus)、Hafniomonas、Hallassia、双尖藻属 (Hammatoidea)、Hannaea、菱板藻属 (Hantzschia)、软管藻属 (Hapalosiphon)、Haplotaenium、定鞭藻门 (Haptophyta)、Haslea、半沟藻属 (Hemidinium)、Hemitonia、茸壳藻属 (Heribaudiella)、异毛藻属 (Heteromastix)、异线藻属 (Heterothrix)、Hibberdia、胭脂藻属 (Hildenbrandia)、隐鞭藻属 (Hillea)、Holopedium、须藻属 (Homoeothrix)、管鞘藻属 (Hormanthionema)、皮襟藻属 (Hormotila)、Hyalobrachion、Hyalocardium、明盘藻属 (Hyalodiscus)、透明棱藻属 (Hyalogonium)、圆丝鼓藻属 (Hyalotheca)、Hydrianum、Hydrococcus、水鞘藻属 (Hydrocoleum)、水条藻属 (Hydrocoryne)、水网藻属 (Hydrodictyon)、水涟藻属 (Hydrosera)、水树藻属 (Hydrurus)、蓝枝藻属 (Hyella)、膜胞藻属 (Hymenomonas)、细绿藻属 (Isthmochloron)、Johannesbaptistia、肾粒藻属 (Juranyiella)、Karayevia、尖眼藻属 (Kathablepharis)、Katodinium、金杯藻属 (Kephyrion)、角球藻属 (Keratococcus)、蹄形藻属 (Kirchneriella)、克里藻属 (Klebsormidium)、Kolbesia、Koliella、Komarekia、Korshikoviella、Kraskella、拉氏藻属 (Lagerheimia)、烧瓶藻属 (Lagynion)、丽枝藻属 (Lamprothamnium)、鱼子菜属 (Lemanea)、鳞孔藻属 (Lepocinclis)、钩端螺旋体属 (Leptosira)、Lobococcus、obocystis、叶衣藻属 (Lobomonas)、Luticola、鞘丝藻属 (Lyngbya)、Malleochloris、鱼鳞藻属 (Mallomonas)、Mantoniella、射星藻属 (Marssoniella)、Martyana、鞭鞘藻属 (Mastigocoleus)、胸隔藻属 (Gastogloia)、直链藻属 (Melosira)、平列藻属 (Merismopedia)、Mesostigma、中带鼓藻属 (Mesotaenium)、微芒藻属 (Micractinium)、小星藻属 (Micrasterias)、微毛藻属 (Microchaete)、微鞘藻属 (Microcoleus)、微胞藻属 (Microcystis)、软壳藻属 (Microglana)、微单胞菌属 (Micromonas)、微孢藻属 (Microspora)、小丛藻属 (Microthamnion)、柄球藻属 (Mischococcus)、单鞭金藻属 (Monochrysis)、蒜头藻属 (Monodus)、Monomastix、单针藻属 (Monoraphidium)、礁膜属 (Monostroma)、转板藻属 (Mougeotia)、拟转板藻属 (Mougeotiopsis)、喙藻属 (Myochloris)、Myromecia、粘囊藻属 (Myxosarcina)、瓶丝藻属 (Naegeliella)、微绿球藻属 (Nannochloris)、类球藻属 (Nautococcus)、舟形藻属 (Navicula)、Neglectella、长篦藻属 (Neidium)、Nephroclamys、肾形藻属 (Nephrocytium)、

Nephrodiella、肾藻属 (Nephroselmis)、梭形鼓藻属 (Netrium)、丽藻属 (Nitella)、拟丽藻属 (Nitellopsis)、菱形藻属 (Nitzschia)、节球藻属 (Nodularia)、念珠藻属 (Nostoc)、棕鞭藻属 (Ochromonas)、鞘藻属 (Oedogonium)、Oligochaetophora、棘接鼓藻属 (Onychonema)、Oocardium、卵囊藻属 (Oocystis)、具隙藻属 (Opephora)、黄管藻属 (Ophiocytium)、Orthoseira、颤藻属 (Oscillatoria)、Oxyneis、厚枝藻属 (Pachycladella)、四集藻属 (Palmella)、掌网藻属 (Palmodictyon)、实球藻属 (Pnadorina)、Pannus、Paralia、巴喧氏藻属 (Pascherina)、Paulschulzia、盘星藻属 (Pediastrum)、柄钟藻属 (Pedinella)、平藻属 (Pedinomonas)、指藻属 (Pedinopera)、Pelagodictyon、柱形鼓藻属 (Penium)、袋鞭藻属 (Peranema)、拟多甲藻属 (Peridiniopsis)、多甲藻属 (Peridinium)、Peronia、石藻属 (Petroneis)、壳衣藻属 (Phacotus)、扁裸藻属 (Phacus)、Phaeaster、褐皮藻属 (Phaeodermatium)、褐藻门 (Phaeophyta)、Phaeosphaera、金枝藻属 (Phaeothamnion)、席藻属 (Phormidium)、叶楯藻属 (Phycopeltis)、叶绿藻属 (Phyllariochloris)、Phyllocardium、Phyllomitas、羽纹藻属 (Pinnularia)、Pitophora、Placoneis、游丝藻属 (Planctonema)、浮球藻属 (Planktosphaeria)、Planothidium、织线藻属 (Plectonema)、杂球藻属 (Pleodorina)、Pleurastrum、厚皮藻属 (Pleurocapsa)、侧枝藻属 (Pleurocladia)、双盘藻属 (Pleurodiscus)、斜纹藻属 (Pleurosigma)、侧链藻属 (Pleurosira)、宽带鼓藻属 (Pleurotaenium)、Pocillomonas、Podohedra、多鞭藻属 (Polyblepharides)、Polychaetophora、多角藻属 (Polyedriella)、多突藻属 (Polyedriopsis)、Polygoniochloris、Polyepidomonas、Polytaenia、素衣藻属 (Polytoma)、Polytomella、紫球藻属 (Porphyridium)、Posteriochromonas、Prasiniochloris、绿枝藻属 (Prasinocladus)、Prasinophyta、溪菜属 (Prasiola)、Prochlorophyta、原绿发藻属 (Prochlorothrix)、原皮藻属 (Protoderma)、原生管属 (Protosiphon)、朴罗藻属 (Provasoliella)、三毛金藻属 (Prymnesium)、Psammodictyon、Psammothidium、伪项圈藻属 (Pseudanabaena)、Pseudenoclonium、Psuedocarteria、Pseudochate、Pseudocharacium、Pseudococcomyxa、Pseudodictyosphaerium、假金杯藻属 (Pseudokephyrion)、伪瘤皮藻属 (Pseudoncobyrza)、Pseudoquadrigula、Pseudosphaerocystis、Pseudostaurastrum、Pseudostaurosira、Pseudotetrastrum、翼膜藻属 (Pteromonas)、Punctastruata、Pyramichlamys、塔胞藻属 (Pyramimonas)、甲藻门 (Pyrrophyta)、四毛藻属 (Quadrichloris)、Quadricoccus、并联藻属 (Quadrigula)、芒球藻属 (Radiococcus)、辐丝藻属 (Radiofilum)、尖头藻属 (Raphidiopsis)、Raphidocelis、针丝藻属 (Raphidonema)、Raphidophyta、Peimeria、棒条藻属 (Rhabdoderma)、杆胞藻属 (Rhabdomonas)、根枝藻属 (Rhizoclonium)、红细胞藻属 (Rhodomonas)、红藻门 (Rhodophyta)、弯楔藻属 (Rhoicosphenia)、棒杆藻属 (Rhopalodia)、胶须藻属 (Rivularia)、Rosenvingiella、Rossithidium、Roya、栅藻属 (Scenedesmus)、Scherffelia、Schizochlamydella、裂壁藻属 (Schizochlamys)、裂线藻属 (Schizomeris)、裂须藻属 (Schizothrix)、弓形藻属 (Schroederia)、Scolioneis、螺翼藻属 (Scotiella)、Scotiellopsis、Scourfieldia、双歧藻属 (Scytonema)、月牙藻属 (Selenastrum)、月绿藻属 (Selenochloris)、Sellaphora、Semiorbis、褐胞藻属 (Siderocelis)、拟铁囊藻属 (Diderocystopsis)、Dimonsenia、管线

藻属、Sirocladium、链膝藻属 (Sirogonium)、骨条藻属 (Skeletonema)、群星藻属 (Sorastrum)、Spermatozopsis、Sphaerello cystis、拟球藻属 (Sphaerellopsis)、Sphaerodinium、环藻属 (Sphaeroplea)、瘤接鼓藻属 (Sphaerosoma)、刺胞藻属 (Spiniferomonas)、水绵属 (Spirogyra)、螺旋藻属 (Spirulina)、椎槲藻属 (Spondylomorom)、顶接鼓藻属 (Spondylosium)、Sporotetras、Spumella、角星鼓藻属 (Staurostrum)、叉链藻属 (Staurerodesmus)、幅节藻属 (Stauroneis)、十字脆杆藻亚属 (Staurosira)、Staurosirella、长羽藻属 (Stenopterobia)、Stephanocostis、冠盘藻属 (Stephanodiscus)、冠孔藻属 (Stephanoporos)、Stephanosphaera、裂丝藻属 (Stichococcus)、粘胶藻属 (Stichogloea)、毛枝藻属 (Stigeoclonium)、真枝藻属 (Stigonema)、柄球藻属 (Stipitococcus)、斯特克藻属 (Stokesiella)、陀螺藻属 (Strombomonas)、柄胞藻属 (Stylochrysalis)、Stylodinium、柱钟藻属 (Styloxyis)、绿柄球藻属 (Stylosphaeridium)、双菱藻属 (Surirella)、Sykidion、束藻属 (Symploca)、聚球藻属 (Synechococcus)、集胞藻属 (Synechocystis)、针杆藻属 (Synedra)、聚赭胞藻属 (Synochromonas)、黄群藻属 (Synura)、平板藻属 (Tabellaria)、Tabularia、Teilingia、切孢藻属 (Temnogametum)、裂顶鼓藻属 (Tetmemorus)、四球藻属 (Tetrachlorella)、四环藻属 (Tetrademus)、四链藻属 (Tetrademus)、Tetraedriella、四角藻属 (Tetraedron)、绿色鞭毛藻属 (Tetraselmis)、四孢藻属 (Tetraspora)、四星藻属 (Tetrastrum)、海链藻属 (Thalassiosira)、丛毛藻属 (Thamniochaete)、Thorakochloris、红索藻属 (Thorea)、鸟巢藻属 (Tolypella)、单歧藻属 (Tolypothrix)、囊裸藻属 (Trachelomonas)、Trachydiscus、共球藻属 (Trebouxia)、橘色藻属 (Trentepholia)、四棘藻属 (Treubaria)、黄丝藻属 (Tribonema)、束毛藻属 (Trichodesmium)、Trichodiscus、小箍藻属 (Trochiscia)、盘杆藻属 (Tryblionella)、丝藻属 (Ulothrix)、辐尾藻属 (Uroglena)、尾丝藻属 (Uronema)、尾管藻属 (Urosolenia)、尾孢藻属 (Urospora)、Uva、周泡藻属 (Vacuolaria)、无节藻属 (Vaucheria)、团藻属 (Volvox)、Volvulina、韦斯藻属 (Westella)、网甲藻属 (Woloszynskia)、多棘鼓藻属 (Xanthidium)、黄藻门 (Xanthophyta)、异球藻属 (Xenococcus)、双星藻属 (Zygnema)、拟双星藻属 (Zygnemopsis) 和 Zygonium。

[0240] 其它蓝细菌包括以下属的成员：管胞藻属 (Chamaesiphon)、色球藻属 (Chroococcus)、蓝细菌属 (Cyanobacterium)、蓝菌属 (Cyanobium)、蓝丝菌属 (Cyanotheca)、蓝纤维藻属 (Dactylococcopsis)、粘杆菌属 (Gloeobacter)、粘球藻属 (Gloeocapsa)、粘杆藻属 (Gloeotheca)、微孢藻属 (Microcystis)、原绿球藻属 (Prochlorococcus)、原绿藻属 (Prochloron)、聚球藻属 (Synechococcus)、集胞藻属 (Synechocystis)、蓝囊胞菌属 (Cyanocystis)、果菌属 (Dermocarpella)、斯塔尼尔氏菌属 (Stanieria)、异球藻属 (Xenococcus)、拟色球藻属 (Chroococcidiopsis)、粘囊藻属 (Myxosarcina)、节旋藻属 (Arthrospira)、搏氏藻属 (Borzia)、发毛针藻属 (Crinalium)、Geitlerinema、瘦鞘丝藻属 (Leptolyngbya)、湖丝藻属 (Limnothrix)、鞘丝藻属 (Lyngbya)、微鞘藻属 (Microcoleus)、颤藻属 (Oscillatoria)、浮丝藻属 (Planktothrix)、Prochlorothrix、伪项圈藻属 (Pseudanabaena)、螺旋藻属 (Spirulina)、斯塔氏蓝细菌属 (Starria)、束藻属 (Symploca)、束毛藻属 (Trichodesmium)、常丝藻属 (Tychonema)、

鱼腥藻属 (*Anabaena*)、项圈藻属 (*Anabaenopsis*)、束丝藻属 (*Aphanizomenon*)、蓝螺菌属 (*Cyanospira*)、拟柱孢藻属 (*Cylindrospermopsis*)、柱孢藻属 (*Cylindrospermum*)、节球藻属 (*Nodularia*)、念珠藻属 (*Nostoc*)、伪枝藻属 (*Scytonema*)、眉藻属 (*Calothrix*)、胶须藻属 (*Rivularia*)、单歧藻属 (*Tolypothrix*)、拟绿胶藻属 (*Chlorogloeopsis*)、费氏藻属 (*Fischerella*)、吉特勒氏蓝细菌属 (*Geitieria*)、*Iyengariella*、拟珠藻属 (*Nostochopsis*)、真枝藻属 (*Stigonema*) 和嗜热聚球藻属 (*Thermosynechococcus*)。

[0241] 绿色非硫细菌包括但不限于以下属：绿曲挠菌属 (*Chloroflexus*)、绿丝菌属 (*Chloronema*)、颤绿菌属 (*Oscillochloris*)、螺丝菌属 (*Heliothrix*)、滑柱菌属 (*Herpetosiphon*)、玫瑰弯菌属 (*Roseiflexus*) 和热微菌属 (*Thermomicrobium*)。

[0242] 绿色硫细菌包括但不限于以下属：绿菌属 (*Chlorobium*)、格状绿菌属 (*Clathrochloris*) 和突柄绿菌属 (*Prosthecochloris*)。

[0243] 紫色硫细菌包括但不限于以下属：*Allochromatium*、着色菌属 (*Chromatium*)、盐着色菌属 (*Halochromatium*)、等着色菌属 (*Isochromatium*)、*Marichromatium*、小红卵菌属 (*Rhodovulum*)、热着色菌属 (*Thermochromatium*)、荚硫菌属 (*Thiocapsa*)、硫红球菌属 (*Thiorhodococcus*) 和囊硫菌属 (*Thiocystis*)。

[0244] 紫色非硫细菌包括但不限于以下属：褐螺菌属 (*Phaeospirillum*)、*Rhodobaca*、红细菌属 (*Rhodobacter*)、红微菌属 (*Rhodomicrobium*)、红球形菌属 (*Rhodopila*)、红假单胞菌属 (*Rhodopseudomonas*)、玫瑰菌属 (*Rhodothalassium*)、红螺菌属 (*Rhodospirillum*)、*Rhodovibrio* 和玫瑰螺旋菌属 (*Roseospira*)。

[0245] 好氧化能无机营养性细菌包括但不限于硝化细菌，如硝化杆菌科 (*Nitrobacteraceae*) 的种、硝化细菌属 (*Nitrobacter*) 的种、硝化刺菌属 (*Nitrospina*) 的种、硝化球菌属 (*Nitrococcus*) 的种、硝化螺菌属 (*Nitrospira*) 的种、亚硝化单胞菌属 (*Nitrosomonas*) 的种、亚硝化球菌属 (*Nitrosococcus*) 的种、亚硝化螺菌属 (*Nitrosospira*) 的种、亚硝化叶菌属 (*Nitrosolobus*) 的种、亚硝化弧菌属 (*Nitrosovibrio*) 的种；无色硫细菌，如卵硫菌属 (*Thiovulum*) 的种、硫杆菌属 (*Thiobacillus*) 的种、硫微螺菌属 (*Thiomicrospira*) 的种、硫球形菌属 (*Thiosphaera*) 的种、高温毛发菌属 (*Thermothrix*) 的种；专性化能无机营养性氢细菌，如氢杆菌属 (*Hydrogenobacter*) 的种、铁和锰氧化和 / 或沉积细菌 (如铁球菌属 (*Siderococcus*) 的种)、和磁性细菌 (*magnetotactic bacteria*) (如水螺菌属 (*Aquaspirillum*) 的种)。

[0246] 古细菌包括但不限于产甲烷的古细菌，如甲烷杆菌属 (*Methanobacterium*) 的种、甲烷短杆菌属 (*Methanobrevibacter*) 的种、甲烷嗜热菌属 (*Methanothermus*) 的种、甲烷球菌属 (*Methanococcus*) 的种、甲烷微菌属 (*Methanomicrobium*) 的种、甲烷螺菌属 (*Methanospirillum*) 的种、产甲烷菌属 (*Methanogenium*) 的种、甲烷八叠球菌属 (*Methanosarcina*) 的种、甲烷叶菌属 (*Methanolobus*) 的种、甲烷丝菌属 (*Methanothrix*) 的种、拟甲烷球菌属 (*Methanococcoides*) 的种、甲烷盘菌属 (*Methanoplanus*) 的种；极端嗜热硫代谢菌，如热变形菌属 (*Thermoproteus*) 的种、热网菌属 (*Pyrodictium*) 的种、硫化裂片菌属 (*Sulfolobus*) 的种、嗜酸菌属 (*Acidianus*) 的种，和其它微生物，如枯草芽孢杆菌、酿酒酵母、链霉菌属 (*Streptomyces*) 的种、雷尔氏菌属 (*Ralstonia*) 的种、红球菌属 (*Rhodococcus*) 的种、棒状杆菌属 (*Corynebacteria*) 的种、短杆菌属 (*Brevibacteria*) 的

种、分枝杆菌属 (*Mycobacteria*) 的种和油脂酵母。

[0247] 其它适合生物体包括如 Venter 等美国专利公布号 2007/0264688 中所述的合成细胞或由合成基因组产生的细胞、以及如 Glass 等美国专利公布号 2007/0269862 中所述的细胞样系统或合成细胞。

[0248] 其它适合的生物体包括大肠杆菌、醋酸杆菌 (*Acetobacter aceti*)、枯草芽孢杆菌、酵母和真菌, 如扬氏梭菌 (*Clostridium ljungdahlii*)、热纤梭菌 (*Clostridium thermocellum*)、产黄青霉菌 (*Penicillium chrysogenum*)、巴斯德毕赤酵母、酿酒酵母、粟酒裂殖酵母 (*Schizosaccharomyces pombe*)、荧光假单胞菌 (*Pseudomonas fluorescens*) 或运动发酵单胞菌 (*Zymomonas mobilis*)。在一些实施方案中, 那些生物体被工程化来固定二氧化碳, 而在其它实施方案中, 它们不被工程化。

[0249] 在一些实施方案中, 如昆虫细胞或哺乳动物细胞 (如人细胞) 的真核细胞用作宿主细胞。熟知用于所述细胞的载体和表达控制序列, 包括启动子和增强子。适用于这个目的的哺乳动物宿主细胞系的实例是由 SV40 转化的猴肾 CV1 株系 (COS-7, ATCC CRL 1651); 人胚肾株系 (293 细胞或被亚克隆来在悬浮培养物中生长的 293 细胞, Graham 等, *J. Gen. Virol.* 36:59(1977)); 幼小仓鼠肾细胞 (BHK, ATCC CCL 10); 中国仓鼠卵巢细胞 /-DHFR (CHO, Urlaub 等, *Proc. Natl. Acad. Sci. USA* 77:4216(1980)); 小鼠塞尔托利细胞 (sertoli cell) (TM4, Mather, *Biol. Reprod.* 23:243-251(1980)); 猴肾细胞 (CV1 ATCC CCL 70); 非洲绿猴肾细胞 (VERO-76, ATCC CRL-1587); 人宫颈癌细胞 (HELA, ATCC CCL 2); 犬肾细胞 (MDCK, ATCC CCL 34); 水牛大鼠肝细胞 (BRL 3A, ATCC CRL 1442); 人肺细胞 (W138, ATCC CCL 75); 人肝细胞 (Hep G2, HB 8065); 小鼠乳腺肿瘤 (MMT 060562, ATCC CCL51); TRI 细胞 (Mather 等, *Annals N. Y. Acad. Sci.* 383:44-68(1982)); MRC 5 细胞; FS4 细胞; 和人肝细胞瘤株系 (Hep G2)。

#### [0250] F. 营养性蛋白质的产生

[0251] 熟练技术人员了解可用于培养重组细胞以产生 (以及任选分泌) 如本文公开的重组营养性蛋白质, 以及用于纯化和 / 或分离表达的重组蛋白的许多适合的方法。选择用于蛋白质纯化的方法取决于许多变量, 包括目标蛋白质的性质、它在细胞内的位置和形式、载体、宿主菌株背景和表达的蛋白质的预定应用。培养条件也可对给定靶标蛋白质的溶解度和定位具有影响。许多方法可用于纯化在如本文公开的重组微生物细胞中表达的靶标蛋白质, 所述方法不加限制地包括离子交换和凝胶过滤。

[0252] 在一些实施方案中, 添加肽融合标签至重组蛋白中, 从而使得有可能使用利用肽融合标签的多种亲和纯化方法。在一些实施方案中, 使用亲和方法使得能够在一步中将靶标蛋白质纯化至接近均质。纯化可包括用例如肠激酶、因子 Xa、凝血酶或 HRV 3C 蛋白酶裂解融合标签的一部分或全部。在一些实施方案中, 在表达的靶标蛋白质的纯化或活性测量之前, 对靶标蛋白质的表达水平、细胞定位和溶解度进行初步分析。靶标蛋白质可见于任何或全部以下部分中: 可溶性或不溶性细胞质部分、周质或培养基。视预定应用而定, 优先定位于包涵体、培养基或周质间隙可为有利的, 在一些实施方案中, 以通过相对简单程序达成快速纯化。

[0253] 尽管大肠杆菌被广泛被认为是用于异源蛋白质表达的强健宿主, 但也广泛已知的是在这个宿主中过度表达许多蛋白质倾向于以不溶性包涵体形式聚集。一种用于挽救包涵



体形成或改进蛋白质自身的效价的最通常使用的方法在于包括融合于目标蛋白质的氨基末端麦芽糖结合蛋白 (MBP) [Austin BP, Nallamsetty S, Waugh DS. Hexahistidine-tagged maltose-binding protein as a fusion partner for the production of soluble recombinant proteins in *Escherichia coli*. *Methods Mol Biol*. 2009 ;498:157-72]、或小泛素相关修饰因子 (SUMO) [Saitoh H, Uwada J, Azusa K. Strategies for the expression of SUMO-modified target proteins in *Escherichia coli*. *Methods Mol Biol*. 2009 ;497:211-21 ;Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Funct Genomics*. 2004 ;5(1-2):75-86 ;Panavas T, Sanders C, Butt TR. SUMO fusion technology for enhanced protein production in prokaryotic and eukaryotic expression systems. *Methods Mol Biol*. 2009 ;497:303-17]。这两种蛋白质在大肠杆菌中极其良好且以可溶性形式表达以使目标蛋白质也以可溶性形式有效产生。可通过在目标蛋白质与融合蛋白之间设计位点特异性蛋白酶识别序列 (如烟草蚀刻病毒 (TEV) 蛋白酶) 来裂解目标蛋白质 [1]。

[0254] 在一些实施方案中,重组蛋白最初不正确折叠或是不溶性的。熟知用于再折叠不溶性蛋白质的多种方法。大多数方案包括依次通过离心和在变性条件下溶解来分离不溶性包涵体。接着透析蛋白质或稀释至非变性缓冲液中,在所述缓冲液中发生再折叠。因为每一种蛋白质都具有独特折叠性质,所以用于任何给定蛋白质的最优再折叠方案可由熟练技术人员凭经验确定。最优再折叠条件可例如通过矩阵方法以小规模快速确定,在所述方法中,测试如蛋白质浓度、还原剂、氧化还原处理、二价阳离子等的变量。一旦得到最优浓度,它们即可被应用于靶标蛋白质的较大规模溶解和再折叠。

[0255] 在一些实施方案中,营养性蛋白质不包含三级结构。在一些实施方案中,营养性蛋白质中小于半数氨基酸参与三级结构。在一些实施方案中,营养性蛋白质不包含二级结构。在一些实施方案中,营养性蛋白质中小于半数氨基酸参与二级结构。重组营养性蛋白质可自以包含这些结构特征的一者或多者的状态表达它们的细胞的培养物分离。在一些实施方案中,在蛋白质自产生它的培养物分离之后,重组营养性蛋白质的三级结构被减少或消除。在一些实施方案中,在蛋白质自产生它的培养物分离之后,重组营养性蛋白质的二级结构被减少或消除。

[0256] 在一些实施方案中,在碱性 pH 下的 CAPS 缓冲液与 N-月桂酰基肌氨酸的组合用于达成包涵体溶解,随后在 DTT 存在下透析以促进再折叠。视靶标蛋白质、表达条件和预定应用而定,自洗涤的包涵体溶解的蛋白质可为 >90% 均质且可不需要进一步纯化。使用 His•Tag® 融合蛋白和 His•Bind® 固定的金属亲和色谱 (Novogen®), 有可能在完全变性条件下 (在再折叠之前) 进行纯化。此外,使用 6M 脲自包涵体溶解的 S•Tag™、T7•Tag® 和 Strep•Tag® II 融合蛋白可通过稀释至 2M 脲 (S•Tag 和 T7•Tag) 或 1M 脲 (Strep•Tag II), 随后在适当树脂上进行色谱分析来在部分变性条件下纯化。再折叠的融合蛋白可使用 His•Tag、S•Tag、Strep•Tag II 和其它适当亲和标签 (例如 GST•Tag™ 和 T7•Tag) (Novogen®), 在天然条件下亲和纯化。

[0257] 在一些实施方案中,重组营养性蛋白质是用于表达它的宿主细胞的内源性蛋白

质。也就是说,宿主细胞的细胞基因组包含编码重组营养性蛋白质的开放阅读框。在一些实施方案中,将足以增加营养性蛋白质的表达的调控序列插入宿主细胞基因组中且可操作地连接于内源性开放阅读框以使所述调控序列驱动自重组核酸过度表达重组营养性蛋白质。在一些实施方案中,使异源核酸序列融合于营养性蛋白质的内源性开放阅读框且导致合成包含异源氨基酸序列的营养性蛋白质,所述异源氨基酸序列改变重组营养性蛋白质的细胞迁移,如将它引导至细胞器或分泌路径中。在一些实施方案中,将编码内源性宿主细胞蛋白质的开放阅读框于质粒上引入宿主细胞中,所述质粒进一步包含可操作地连接于所述开放阅读框的调控序列。在一些实施方案中,重组宿主细胞表达的重组营养性蛋白质比由在类似条件下生长的类似宿主细胞产生的营养性蛋白质的量多至少 2 倍、至少 3 倍、至少 4 倍、至少 5 倍、至少 10 倍、或至少 20 倍、至少 30 倍、至少 40 倍、至少 50 倍或至少 100 倍。

[0258] 在一些实施方案中,本公开的营养性蛋白质是在不使用重组产生系统下化学合成。可使用本领域中已知的技术在液相系统中或在固相系统中进行蛋白质合成(参见例如, Atherton, E., Sheppard, R. C. (1989). *Solid Phase peptide synthesis: a practical approach*. Oxford, England: IRL Press; Stewart, J. M., Young, J. D. (1984). *Solid phase peptide synthesis* (第 2 版). Rockford: Pierce Chemical Company)。

[0259] G. 在植物中重组营养性蛋白质的产生

[0260] 包含编码本公开的营养性蛋白质的核酸序列的核酸分子使得能够产生包含所述核酸序列的转基因植物。因此,本公开也提供包含包括编码本公开的营养性蛋白质的核酸序列的重组核酸分子的植物。植物可为经受转化和再生的任何植物且包括但不限于刺槐、苜蓿草、小茴香、苹果、杏、洋蓟、芝麻菜、芦笋、鳄梨、香蕉、大麦、菜豆、甜菜、黑莓、蓝莓、茎椰菜、抱子甘蓝、卷心菜、油菜、棱瓜、胡萝卜、木薯、花椰菜、芹菜、大白菜、樱桃、芜荑、柑橘、细皮小柑桔、咖啡豆、玉米、棉花、黄瓜、花旗松、茄子、菊苣、苦苣、桉树、茴香、无花果、森林树木、葫芦、葡萄、葡萄柚、蜜瓜、豆薯、猕猴桃、莴苣、韭菜、柠檬、酸橙、火炬松、芒果、甜瓜、蘑菇、坚果、燕麦、秋葵、洋葱、橙、观赏植物、番木瓜、欧芹、豌豆、桃子、花生、梨、胡椒、柿子、松树、菠萝、车前草、李子、石榴、白杨、马铃薯、南瓜 (pumpkin)、槲寄生、辐射松、红菊苣、萝卜、油菜籽、悬钩子、稻米、黑麦、高粱、美国长叶松、黄豆、菠菜、南瓜 (squash)、草莓、糖用甜菜、甘蔗、向日葵、甜玉米、甜薯、枫香、红桔、茶叶、烟草、番茄、草皮、藤、西瓜、小麦、薯蓣和绿皮西葫芦。在优选实施方案中,植物是菜豆、茎椰菜、卷心菜、油菜、胡萝卜、花椰菜、芹菜、大白菜、玉米、棉花、黄瓜、茄子、韭菜、莴苣、甜瓜、豌豆、胡椒、南瓜 (pumpkin)、萝卜、菠菜、黄豆、南瓜 (squash)、甘蔗、甜玉米、番茄、西瓜和小麦植物。在一些实施方案中,植物是玉米植物。在一些实施方案中,植物是黄豆植物。在一些实施方案中,植物是棉花植物。在一些实施方案中,植物是油菜植物。在一些实施方案中,植物是选自以下的属的成员: 拟南芥属 (*Arabidopsis*)、甜菜属 (*Beta*)、大豆属 (*Glycine*)、麻风树属 (*Jatropha*)、芒属 (*Miscanthus*)、黍属 (*Panicum*)、藨草属 (*Phalaris*)、白杨属 (*Populus*)、蔗属 (*Saccharum*)、柳属 (*Salix*)、希蒙得木属 (*Simmondsia*) 和玉蜀黍属 (*Zea*)。

[0261] 在植物细胞中具有活性的众多启动子已描述于文献中。这些包括植物基因组中存在的启动子以及来自其它来源的启动子,包括根癌土壤杆菌 (*Agrobacterium tumefaciens*) 的肿瘤诱导性质粒上携带的胭脂碱 (nopaline) 合成酶 (NOS) 启动子和章鱼碱 (octopine) 合成酶 (OCS) 启动子、花椰菜花叶病毒组启动子 (如花椰菜花叶病毒

(cauliflower mosaic virus))。举例来说,参见美国专利号 5,858,742 和 5,322,938,其公开源于花椰菜花叶病毒 (CaMV35S) 的组成型启动子的各种形式;美国专利号 5,641,876,其公开稻米肌动蛋白启动子;美国专利申请公布 2002/0192813A1,其公开适用于设计有效植物表达载体的 5'、3' 和内含子元件;美国专利申请号 09/757,089,其公开玉米叶绿体醛缩酶启动子;美国专利申请序列号 08/706,946,其公开稻米谷蛋白启动子;美国专利申请序列号 09/757,089,其公开玉米醛缩酶 (FDA) 启动子;和美国专利申请序列号 60/310,370,其公开玉米烟酰胺合成酶启动子。在植物细胞中起作用的这些启动子和众多其它启动子为本领域技术人员所知且可用于重组核酸中以提供在转基因植物中表达营养性蛋白质。

[0262] 对于一些应用,需要在植物绿色组织中优先表达。用于所述用途的目标启动子包括来自拟南芥 (*Arabidopsis thaliana*) 核酮糖 -1,5- 双磷酸羧化酶 (二磷酸核酮糖羧化酶) 小亚单位 (Fischhoff 等 (1992) *Plant Mol. Biol.* 20:81-93)、醛缩酶和丙酮酸正磷酸二激酶 (PPDK) (Taniguchi 等 (2000) *Plant Cell Physiol.* 41(1):42-48) 的基因的启动子。

[0263] 此外,启动子可被改变以含有至少一种增强子序列来有助于提高基因表达。所述增强子在本领域中是已知的。通过在所述构建体内包括增强子序列,营养性蛋白质的表达可得以增强。这些增强子常见于在真核细胞中起作用的启动子中的转录起始位置的 5',但可常常插入在编码序列的上游 (5') 或下游 (3')。在一些情况下,这些 5' 增强元件是内含子。特别适用作增强子的是稻米肌动蛋白 1 (参见美国专利号 5,641,876) 和稻米肌动蛋白 2 基因的 5' 内含子、玉米醇脱氢酶基因内含子、玉米热休克蛋白 70 基因内含子 (美国专利号 5,593,874) 和玉米 *shrunk1* 基因。

[0264] 对于一些应用,需要在植物种子组织中表达以实现改良种子组成。用于种子组成改良的示例性启动子包括来自油菜籽蛋白 (美国专利号 5,420,034)、玉米 L3 油脂蛋白 (美国专利号 6,433,252)、玉米蛋白 Z27 (Russell 等 (1997) *Transgenic Res.* 6(2):157-166)、球蛋白 1 (Belanger 等 (1991) *Genetics* 129:863-872)、谷蛋白 1 (Russell (1997) (上文)) 和抗氧化过氧化物氧化酶 (Per1) (Stacy 等 (1996) *Plant Mol. Biol.* 31(6):1205-1216) 的种子基因的启动子

[0265] 根据本公开制备的重组核酸构建体也将通常包括通常含有聚腺苷酸化信号和位点的 3' 元件。熟知的 3' 元件包括例如美国专利号 6,090,627 中公开的来自如 *nos* 3'、*tml* 3'、*tmr* 3'、*tms* 3'、*ocs* 3'、*tr73* 的根癌土壤杆菌基因的元件;来自如小麦 (普通小麦 (*Triticum aestivum*)) 热休克蛋白 17 (*Hsp17* 3')、小麦泛素基因、小麦果糖 -1,6- 二磷酸酶基因、稻米谷蛋白基因、稻米乳酸脱氢酶基因和稻米  $\beta$ -微管蛋白基因的植物基因的 3' 元件,其全部公开于美国公布专利申请 2002/0192813 A1 中;以及豌豆 (豌豆 (*Pisum sativum*)) 二磷酸核酮糖羧化酶基因 (*rbs* 3') 和来自宿主植物内的基因的 3' 元件。

[0266] 构建体和载体也可包括用于使基因靶标靶向植物细胞器,特别是叶绿体、白色体或其它质体细胞器的转运肽。对于使用叶绿体转运肽的描述,参见美国专利号 5,188,642 和美国专利号 5,728,925。对于拟南芥属 *EPSPS* 基因的转运肽区域的描述,参见 Klee, H. J. 等 (MGG (1987) 210:437-442)。

[0267] 用于用重组 DNA 转化植物细胞的众多方法在本领域中是已知的且可用于本公开中。用于植物转化的两种通常使用的方法是土壤杆菌属介导的转化和微粒轰击。微粒轰击方法说明于美国专利号 5,015,580 (黄豆); 5,550,318 (玉米); 5,538,880 (玉

米);5,914,451(黄豆);6,160,208(玉米);6,399,861(玉米)和6,153,812(小麦)中,且土壤杆菌属介导的转化描述于美国专利号5,159,135(棉花);5,824,877(黄豆);5,591,616(玉米);和6,384,301(黄豆)中。对于基于根癌土壤杆菌的植物转化系统,存在于转化构建体上的其它元件将包括用以促进重组多核苷酸并入植物基因组中的T-DNA左侧和右侧边缘序列。

[0268] 一般来说,适用的是在靶标植物株系的基因组中随机引入重组DNA,即引入在非特定位置处。在特殊情况下,可适用的是进行靶向性重组DNA插入以达成位点特异性整合,例如以置换基因组中的现存基因,使用植物基因组中的现存启动子,或在已知就基因表达而言具有活性的预定位点处插入重组多核苷酸。存在已知在植物中起作用的若干位点特异性重组系统,包括如美国专利号4,959,317中公开的cre-lox和如美国专利号5,527,695中公开的FLP-FRT。

[0269] 转化方法是通常在培养基上以及在受控环境中进行组织培养时实施。“培养基”是指用于使细胞在体外(也就是说在完整活生物体外部)生长的众多营养物混合物。接受者细胞靶标包括但不限于分生组织细胞、愈伤组织、不成熟胚胎和配子细胞,如小孢子、花粉、精子和卵细胞。预期可繁殖植物可自其再生的任何细胞都适用作接受者细胞。愈伤组织可起始于包括但不限于不成熟胚胎、幼苗顶端分生组织、小孢子等的组织来源。能够以愈伤组织形式增殖的细胞也是用于遗传转化的接受者细胞。美国专利号6,194,636和6,232,526中公开用于制备本公开的转基因植物的实际转化方法和材料,例如各种培养基和接受者靶标细胞、转化不成熟胚胎细胞以及随后再生可繁殖转基因植物。

[0270] 转基因植物的种子可自可繁殖转基因植物加以收集且用于使产生本公开的重组营养性蛋白质的转化植物的子代生长。除用重组DNA直接转化植物之外,转基因植物也可通过使具有重组DNA的第一植物与缺乏所述DNA的第二植物杂交来制备。举例来说,可将重组DNA引入可经受转化以产生转基因植物的第一植物株系中,所述第一植物株系可与第二植物株系杂交以使重组DNA渗入第二植物株系中。可使具有编码本公开的营养性蛋白质的重组DNA的转基因植物与具有赋予另一特性(例如除草剂抗性 or 害虫抗性 or 产生如油的第二营养性产物)的其它重组DNA的转基因植物株系杂交以产生具有赋予两种特性的重组DNA的子代植物。通常,在用于组合特性的所述育种中,贡献其它特性的转基因植物是雄性株系且携带基础特性的转基因植物是雌性株系。这个杂交的子代将分离以使一些植物将携带两种亲本特性的DNA而一些将携带一种亲本特性的DNA;所述植物可通过与亲本重组DNA相关的标记来鉴定,例如通过分析重组DNA,或在可选择标记连接于重组体的情况下通过应用选择剂(如供与除草剂耐受性标记一起使用的除草剂),或通过选择增强的特性来进行标记鉴定。携带两种亲本特性的DNA的子代植物可与雌性亲本株系回交多次(例如通常6至8代)以产生除另一转基因亲本株系的重组DNA之外,基因型大致上与一个原始转基因亲本株系相同的子代植物。

[0271] 在实施转化时,通常在任一转化实验中都将DNA引入仅较小百分比的靶标植物细胞中。标记基因用于提供用于鉴定通过接受转基因DNA构建体并将所述构建体整合至它们的基因组中而被稳定转化的那些细胞的高效系统。优选标记基因提供赋予对如抗生素或除草剂的选择剂的抗性的选择性标记。转化植物可对其具有抗性的任何除草剂都是适用于选择性标记的试剂。使潜在转化细胞暴露于选择剂。在存活群体中,细胞将为其中通常赋予

抗性的基因被整合且在足以允许细胞存活的水平下表达的那些细胞。可进一步测试细胞以确认外源性 DNA 的稳定整合。通常使用的选择性标记基因包括赋予对如卡那霉素和巴龙霉素 (nptII)、潮霉素 B(aph IV) 和庆大霉素 (aac3 和 aacC4) 的抗生素的抗性或对如草铵膦 (bar 或 pat) 和草甘膦 (aroA 或 EPSPS) 的除草剂的抗性的标记基因。所述可选择标记的实例说明于美国专利号 5,550,318 ;5,633,435 ;5,780,708 和 6,118,047 中。也可采用提供目视鉴定转化体的能力的可选择标记,例如表达着色或荧光蛋白质(如荧光素酶或绿色荧光蛋白质(GFP))的基因或表达  $\beta$ -葡萄糖醛酸酶的基因或已知其各种显色底物的 uidA 基因(GUS)。

[0272] 可在再生培养基中培养在暴露于选择剂时存活的植物细胞或在筛选测定中已计分为阳性的植物细胞且使所述植物细胞成熟成植物。可将自转化植物细胞再生的发育幼苗转移至植物生长混合物中,且例如在环境受控的腔室中经受耐寒锻炼。视初始组织而定,在鉴定转化体之后约 6 周至 10 个月再生植物。植物可使用本领域技术人员已知的常规植物育种方法加以授粉且例如自花授粉产生的种子通常与转基因玉米一起使用。可测试再生转化植物或它的子代种子或植物的重组 DNA 表达且针对存在异源营养性蛋白质加以选择。

[0273] 使源于本公开的植物细胞的转基因植物生长以产生包含编码本公开的营养性蛋白质的异源核酸的转基因植物且产生包含异源核酸序列的转基因种子和单倍体花粉。具有增强的特性的所述植物是通过选择具有增强的特性的转化植物或子代种子来鉴定。自本文提供的转基因种子生长的转基因植物显示促进产率增加的改良的农艺学特性或提供增加的植物价值的其它特性,包括例如改良的蛋白质品质,如增加必需氨基酸、支链氨基酸或 Leu 的至少一者的含量。

[0274] 转基因植物适用作营养性蛋白质的来源。举例来说,在一些实施方案中,包含本公开的重组营养性蛋白质的转基因植物相较于对照非转基因植物包含总蛋白质的重量分数增加。在一些实施方案中,包含本公开的重组营养性蛋白质的转基因植物相较于对照非转基因植物包含必需氨基酸的重量分数增加。在一些实施方案中,包含本公开的重组营养性蛋白质的转基因植物相较于对照非转基因植物包含支链氨基酸的重量分数增加。在一些实施方案中,包含本公开的重组营养性蛋白质的转基因植物相较于对照非转基因植物包含 Leu 的重量分数增加。在一些实施方案中,包含本公开的重组营养性蛋白质的转基因植物包含以下至少一者:a) 相较于对照非转基因植物,支链氨基酸残基与总氨基酸残基的比率增加;b) 相较于对照非转基因植物,Leu 残基与总氨基酸残基的比率增加;和 c) 相较于对照非转基因植物,必需氨基酸残基与总氨基酸残基的比率增加。在一些实施方案中,包含本公开的重组营养性蛋白质的转基因植物包含:a) 相较于对照非转基因植物,支链氨基酸残基与总氨基酸残基的比率增加;b) 相较于对照非转基因植物,Leu 残基与总氨基酸残基的比率增加;和 c) 相较于对照非转基因植物,必需氨基酸残基与总氨基酸残基的比率增加。

[0275] 因此,转基因植物适用作高品质蛋白质的来源。植物可被收集且在或不在进一步加工下用于哺乳动物膳食中。举例来说,相较于由不包含重组营养性蛋白质的植物制得的类似产品,由转基因小麦制得的面粉、由转基因玉米制得的玉米粉、或由转基因稻获得的稻米或米粉富含蛋白质、必需氨基酸、支链氨基酸和 Leu 的至少一者。在一些实施方案中,重组营养性蛋白质是植物蛋白质或包含植物蛋白质或其衍生物或突变蛋白的多肽序列,如但不必定是相同类型的植物的蛋白质或多肽序列。在其它实施方案中,重组营养性蛋白质不

是植物蛋白质或其衍生物或突变蛋白。

[0276] 在一些实施方案中,重组营养性蛋白质是在它由哺乳动物消耗之前自转基因植物回收或部分回收。

[0277] H. 组合物

[0278] 本文公开的至少一种营养性蛋白质可与至少一种第二组分组合以形成营养性组合物。在一些实施方案中,组合物中的氨基酸的唯一来源是本文公开的至少一种营养性蛋白质。在所述实施方案中,组合物的氨基酸组成将与本文公开的至少一种营养性蛋白质的氨基酸组成相同。在一些实施方案中,组合物包含本文公开的至少一种营养性蛋白质和至少一种第二蛋白质。在一些实施方案中,至少一种第二蛋白质是本文公开的第二营养性蛋白质,而在其它实施方案中,至少一种第二蛋白质不是本文公开的营养性蛋白质。在一些实施方案中,组合物包含 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 种或更多种本文公开的营养性蛋白质。在一些实施方案中,组合物包含 0、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 种或更多种不是本文公开的营养性蛋白质的蛋白质。在一些实施方案中,组合物包含 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 种或更多种营养性蛋白质且组合物包含 0、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 种或更多种不是本文公开的营养性蛋白质的蛋白质。

[0279] 在一些实施方案中,如先前段落中所述的营养性组合物进一步包含至少一种多肽、至少一种肽和至少一种游离氨基酸的至少一者。在一些实施方案中,营养性组合物包含至少一种多肽和至少一种肽。在一些实施方案中,营养性组合物包含至少一种多肽和至少一种游离氨基酸。在一些实施方案中,营养性组合物包含至少一种肽和至少一种游离氨基酸。在一些实施方案中,至少一种多肽、至少一种肽和 / 或至少一种游离氨基酸包含选自 1) 支链氨基酸, 2) 亮氨酸, 和 3) 必需氨基酸的氨基酸。在一些实施方案中,至少一种多肽、至少一种肽和 / 或至少一种游离氨基酸由选自 1) 支链氨基酸, 2) 亮氨酸, 和 3) 必需氨基酸的氨基酸组成。在一些实施方案中,营养性组合物包含至少一种修饰的氨基酸或非标准氨基酸。修饰的氨基酸包括具有对羧基末端、氨基末端和 / 或侧链的一者或多者的修饰的氨基酸。非标准氨基酸可选自通过翻译后修饰蛋白质所形成的氨基酸,例如羧化谷氨酸、羧基脯氨酸或尾下素(hypusine)。其它非标准氨基酸不见于蛋白质中。实例包括羊毛硫氨酸(lanthionine)、2-氨基异丁酸、脱氢丙氨酸、 $\gamma$ -氨基丁酸、鸟氨酸和瓜氨酸。在一些实施方案中,营养性组合物包含一种或多种 D-氨基酸。在一些实施方案中,营养性组合物包含一种或多种 L-氨基酸。在一些实施方案中,营养性组合物包含一种或多种 D-氨基酸和一种或多种 L-氨基酸的混合物。

[0280] 通过添加多肽、肽和游离氨基酸的至少一者至营养性组合物中,组合物中存在的支链氨基酸、亮氨酸和必需氨基酸的至少一者与总氨基酸的比例可得以增加。

[0281] 在一些实施方案中,组合物包含至少一种碳水化合物。“碳水化合物”是指糖或糖的聚合物。术语“糖”、“多糖”、“碳水化合物”和“寡糖”可互换使用。大多数碳水化合物是具有许多羟基的醛或酮,通常在分子各碳原子上具有一个羟基。碳水化合物通常具有分子式  $C_nH_{2n}O_n$ 。碳水化合物可为单糖、二糖、三糖、寡糖或多糖。大多数基础碳水化合物是单糖,如葡萄糖、蔗糖、半乳糖、甘露糖、核糖、阿拉伯糖、木糖和果糖。二糖是两个接合的单糖。示例性二糖包括蔗糖、麦芽糖、纤维二糖和乳糖。通常,寡糖包括三个与六个之间的单糖单

元（例如棉子糖、水苏糖）且多糖包括六个或更多个单糖单元。示例性多糖包括淀粉、糖原和纤维素。碳水化合物可含有修饰的糖单元，如 2'-脱氧核糖（其中羟基被移除）、2'-氟代核糖（其中羟基被氟置换）、或 N-乙酰葡萄糖胺、含氮形式的葡萄糖（例如，2'-氟代核糖、脱氧核糖和己糖）。碳水化合物可以许多不同形式存在，例如构象异构体、环状形式、无环形式、立体异构体、互变异构体、端基异构体和异构体。

[0282] 在一些实施方案中，组合物包含至少一种脂质。如本文所用，“脂质”包括脂肪、油、甘油三酯、胆固醇、磷脂、呈任何形式的脂肪酸，包括游离脂肪酸。脂肪、油和脂肪酸可为饱和的、不饱和的（顺式或反式）或部分不饱和的（顺式或反式）。在一些实施方案中，脂质包含至少一种选自以下的脂肪酸：月桂酸（12:0）、肉豆蔻酸（14:0）、棕榈酸（16:0）、棕榈油酸（16:1）、十七酸（17:0）、十七碳烯酸（17:1）、硬脂酸（18:0）、油酸（18:1）、亚麻油酸（linoleic acid）（18:2）、次亚麻油酸（18:3）、十八碳四烯酸（18:4）、花生酸（20:0）、二十烯酸（20:1）、二十碳二烯酸（20:2）、二十碳四烯酸（20:4）、二十碳五烯酸（20:5）（EPA）、二十二酸（22:0）、二十二烯酸（22:1）、二十二碳五烯酸（22:5）、二十二碳六烯酸（22:6）（DHA）和二十四酸（24:0）。在一些实施方案中，组合物包含至少一种改性脂质，例如已通过烹调来改性的脂质。

[0283] 在一些实施方案中，组合物包含至少一种补充性矿物质或矿物质来源。矿物质的实例不加限制地包括：氯化物、钠、钙、铁、铬、铜、碘、锌、镁、锰、钼、磷、钾和硒。任何前述矿物质的适合形式包括可溶性矿物盐、微溶性矿物盐、不溶性矿物盐、螯合矿物质、矿物质复合物、非反应性矿物质（如羰基矿物质）和还原的矿物质及其组合。

[0284] 在一些实施方案中，组合物包含至少一种补充性维生素。至少一种维生素可为脂溶性或水溶性维生素。适合的维生素包括但不限于维生素 C、维生素 A、维生素 E、维生素 B12、维生素 K、核黄素、烟酸、维生素 D、维生素 B6、叶酸、吡哆醇、硫胺、泛酸和生物素。任何前述维生素的适合形式是维生素的盐、维生素的衍生物、具有维生素的相同或类似活性的化合物、和维生素的代谢物。

[0285] 在一些实施方案中，组合物包含至少一种生物体。适合实例在本领域中是熟知的且包括益生菌（例如，乳酸杆菌属或双歧杆菌属的种）、螺旋藻属、小球藻属和紫菜属。

[0286] 在一些实施方案中，组合物包含至少一种膳食补充剂。适合的实例在本领域中是熟知的且包括草本制剂、植物制剂和某些激素。非限制性实例包括银杏、人参和褪黑素（melatonin）。

[0287] 在一些实施方案中，组合物包含赋形剂。适合的赋形剂的非限制性实例包括缓冲剂、防腐剂、稳定剂、粘合剂、压实剂、润滑剂、分散增强剂、崩解剂、调味剂、甜味剂、着色剂。

[0288] 在一些实施方案中，赋形剂是缓冲剂。适合的缓冲剂的非限制性实例包括柠檬酸钠、碳酸镁、碳酸氢镁、碳酸钙和碳酸氢钙。

[0289] 在一些实施方案中，赋形剂包含防腐剂。适合的防腐剂的非限制性实例包括抗氧化剂，如  $\alpha$ -生育酚和抗坏血酸；以及抗微生物剂，如对羟基苯甲酸酯、氯丁醇和苯酚。

[0290] 在一些实施方案中，组合物包含粘合剂作为赋形剂。适合的粘合剂的非限制性实例包括淀粉、预胶凝淀粉、明胶、聚乙烯吡咯烷酮、纤维素、甲基纤维素、羧甲基纤维素钠、乙基纤维素、聚丙烯酰胺、聚乙烯噁唑烷酮、聚乙烯醇、C12-C18 脂肪酸醇、聚乙二醇、多元醇、糖、寡糖及其组合。

[0291] 在一些实施方案中,组合物包含润滑剂作为赋形剂。适合的润滑剂的非限制性实例包括硬脂酸镁、硬脂酸钙、硬脂酸锌、氢化植物油、sterotex、聚氧乙烯单硬脂酸酯、滑石、聚乙二醇、苯甲酸钠、月桂基硫酸钠、月桂基硫酸镁和轻质矿物油。

[0292] 在一些实施方案中,组合物包含分散增强剂作为赋形剂。适合的分散剂的非限制性实例包括淀粉、海藻酸、聚乙烯吡咯烷酮、瓜尔胶(guar gum)、高岭土、膨润土、纯化木纤维素、淀粉乙醇酸钠、同形硅酸盐和微晶纤维素作为高 HLB 乳化剂表面活性剂。

[0293] 在一些实施方案中,组合物包含崩解剂作为赋形剂。在一些实施方案中,崩解剂是非泡腾性崩解剂。适合的非泡腾性崩解剂的非限制性实例包括淀粉(如玉米淀粉、马铃薯淀粉、其预凝胶化和改性淀粉)、甜味剂、粘土(如膨润土)、微晶纤维素、海藻酸盐、淀粉乙醇酸钠、胶(如琼脂、瓜尔胶、刺槐豆胶、刺梧桐胶、果胶(pectin)和黄蓍胶)。在一些实施方案中,崩解剂是泡腾性崩解剂。适合的泡腾性崩解剂的非限制性实例包括碳酸氢钠与柠檬酸组合、以及碳酸氢钠与酒石酸组合。

[0294] 在一些实施方案中,赋形剂包含调味剂。并入外层中的调味剂可选自合成调味油和调味香料;天然油;来自植物、叶子、花和果实的提取物;及其组合。在一些实施方案中,调味剂选自肉桂油;冬青油;胡椒薄荷油;三叶草油;干草油;茴香油;桉树;香草;柑橘油,如柠檬油、橙油、葡萄和葡萄柚油;以及果实精油,包括苹果、桃、梨、草莓、悬钩子、樱桃、李子、菠萝和杏。

[0295] 在一些实施方案中,赋形剂包含甜味剂。适合的甜味剂的非限制性实例包括葡萄糖(玉米糖浆)、右旋糖、转化糖、果糖及其混合物(当不用作载体时);糖精和它的各种盐,如钠盐;二肽甜味剂,如阿斯巴甜;二氢查耳酮(dihydrochalcone)化合物、甘草甜(glycyrrhizin);甜菊(*Stevia Rebaudiana*)(甜菊苷(Stevioside));蔗糖的氯代衍生物,如蔗糖素(sucralose);以及糖醇,如山梨糖醇、甘露糖醇、木糖醇(sylitol)等。也涵盖氢化淀粉水解产物和合成甜味剂 3,6-二氢-6-甲基-1,2,3-噁嗪酮-2,2-二氧化物,特别是其钾盐(安赛蜜(acesulfame-K))以及钠盐和钙盐。

[0296] 在一些实施方案中,组合物包含着色剂。适合的着色剂的非限制性实例包括食物、药物和化妆品着色剂(FD和C)、药物和化妆品着色剂(D和C)、以及外部药物和化妆品着色剂(Ext.D和C)。着色剂可用作染料或它们的相应色淀。

[0297] 制剂中的赋形剂或赋形剂的组合的重量分数通常是组合物中的氨基酸的总重量的约 50%或更少、约 45%或更少、约 40%或更少、约 35%或更少、约 30%或更少、约 25%或更少、约 20%或更少、约 15%或更少、约 10%或更少、约 5%或更少、约 2%或更少、或约 1%或更少。

[0298] 本文公开的营养性蛋白质和营养性组合物可被配制成多种形式且通过许多不同手段来施用。组合物可以根据需要含有常规可接受的载体、佐剂和媒介物的制剂形式经口、经直肠或胃肠外施用。如本文所用的术语“胃肠外”包括皮下、静脉内、肌肉内或胸骨内注射和输注技术。在一示例性实施方案中,营养性蛋白质或组合物是经口施用。

[0299] 用于经口施用的固体剂型包括胶囊、片剂、囊片、丸剂、锭剂、糖锭、粉末和颗粒。胶囊通常包含包括营养性蛋白质或组合物的核心材料和囊封所述核心材料的外壳壁。在一些实施方案中,核心材料包含固体、液体和乳剂的至少一者。在一些实施方案中,外壳壁材料包含软质明胶、硬质明胶和聚合物的至少一者。适合的聚合物包括但不限于:纤维素聚合



物,如羟丙基纤维素、羟乙基纤维素、羟丙基甲基纤维素 (HPMC)、甲基纤维素、乙基纤维素、乙酸纤维素、乙酸邻苯二甲酸纤维素、乙酸偏苯三甲酸纤维素、邻苯二甲酸羟丙基甲基纤维素、琥珀酸羟丙基甲基纤维素和羧甲基纤维素钠;丙烯酸聚合物和共聚物,如由丙烯酸、甲基丙烯酸、丙烯酸甲酯、甲基丙烯酸铵、丙烯酸乙酯、甲基丙烯酸甲酯和 / 或甲基丙烯酸乙酯形成的那些(例如,在商标名“Eudragit”下销售的那些共聚物);乙烯基聚合物和共聚物,如聚乙烯吡咯烷酮、聚乙酸乙烯酯、聚乙酸邻苯二甲酸乙酯、乙酸乙酯巴豆酸共聚物、和乙烯-乙酸乙酯共聚物;以及虫胶(纯化紫胶)。在一些实施方案中,至少一种聚合物充当味道掩蔽剂。

[0300] 片剂、丸剂等可被压制、多重压制、多重层化、和 / 或包覆。包覆可为单一的或多重的。在一个实施方案中,包覆材料包含自植物、真菌和微生物的至少一者提取的糖、多糖和糖蛋白的至少一者。非限制性实例包括玉米淀粉、小麦淀粉、马铃薯淀粉、木薯淀粉、纤维素、半纤维素、葡聚糖、麦芽糖糊精、环糊精、菊糖、果胶、甘露聚糖、阿拉伯胶、刺槐豆胶、牧豆树胶、瓜尔胶、刺梧桐胶、印度胶、黄蓍胶、海萝 (funori)、卡拉胶、琼脂、海藻酸盐、壳聚糖或结冷胶。在一些实施方案中,包覆材料包含蛋白质。在一些实施方案中,包覆材料包含脂肪和油的至少一者。在一些实施方案中,脂肪和油的至少一者具有高温熔融性。在一些实施方案中,脂肪和油的至少一者被氢化或部分氢化。在一些实施方案中,脂肪和油的至少一者源于植物。在一些实施方案中,脂肪和油的至少一者包含甘油酯、游离脂肪酸和脂肪酸酯的至少一者。在一些实施方案中,包覆材料包含至少一种可食用蜡。可食用蜡可源于动物、昆虫或植物。非限制性实例包括蜂蜡、羊毛脂、杨梅蜡、棕榈蜡和米糠蜡。可另外制备具有肠溶包衣的片剂和丸剂。

[0301] 或者,可将包含本文公开的营养性蛋白质和营养性组合物的粉末或颗粒并入食品中。在一些实施方案中,食品是用于经口施用的饮料。适合的饮料的非限制性实例包括果汁、果汁饮料、人工调味饮料、人工增甜饮料、碳酸饮料、运动饮料、液体乳制品、搅合饮料、含酒精饮料、含咖啡因饮料、婴儿配方奶粉等。用于经口施用的其它适合手段包括水性和非水性溶液、霜剂、糊剂、乳剂、混悬液和浆液,其各自可任选也含有适合溶剂、防腐剂、乳化剂、混悬剂、稀释剂、甜味剂、着色剂和调味剂的至少一者。

[0302] 在一些实施方案中,食品是固体食物。固体食物的适合实例不加限制地包括成块食物、成块点心、曲奇饼、核仁巧克力饼、松饼、脆饼、饼干、奶油或面团、成块冰淇淋、成块冷冻酸奶等。

[0303] 在一些实施方案中,将本文公开的营养性蛋白质和营养性组合物并入治疗性食物中。在一些实施方案中,治疗性食物是任选含有一些或全部必需大量营养素和微量营养素的备用食物。在一些实施方案中,将本文公开的营养性蛋白质和营养性组合物并入被设计来掺合进现存餐食中的补充性食物中。在一些实施方案中,补充性食物含有一些或全部必需大量营养素和微量营养素。在一些实施方案中,将本文公开的营养性蛋白质和营养性组合物与现存食物掺合或添加至现存食物中以加强所述食物的蛋白质营养。实例包括食物主要成分(谷物、盐、糖、烹调油、人造黄油)、饮料(咖啡、茶、苏打水、啤酒、烈酒、运动饮料)、点心、甜食和其它食物。

[0304] 本文公开的组合物可在方法中用于增加例如肌肉质量、强度和身体功能、产热、代谢消耗、饱腹感、线粒体生物发生、重量或脂肪损失、以及瘦性身体组成的至少一者。

### [0305] I. 使用方法

[0306] 在一些实施方案中,本文公开的营养性蛋白质和营养性组合物是向患者或使用者(有时总称为“受试者”)施用。如本文所用,“施用”涵盖其中一个人指导另一人以某一方式和/或出于某一目的消耗营养性蛋白质或营养性组合物的实施方案,以及其中使用者不依赖于自第二人接收的任何指导或以所述任何指导的变化形式,以某一方式和/或出于某一目的消耗营养性蛋白质或营养性组合物的情况。其中一个人指导另一人以某一方式和/或出于某一目的消耗营养性蛋白质或营养性组合物的实施方案的非限制性实例包括当医师对患者指定执行和/或治疗过程时、当教练员建议使用者(如运动员)遵循特定执行和/或治疗过程时、以及当制造商、销售者或商人向终端使用者推荐使用条件时,例如通过广告或在包装上或在与产品的出售或销售相伴提供的其它材料上进行标记。

[0307] 在一些实施方案中,以某一剂型提供营养性蛋白质或营养性组合物。在一些实施方案中,设计用于施用本文公开的至少一种营养性蛋白质的剂型,其中施用的营养性蛋白质的总量选自0.1g至1g、1g至5g、2g至10g、5g至15g、10g至20g、15g至25g、20g至40g、25-50g、以及30-60g。在一些实施方案中,设计用于施用本文公开的至少一种营养性蛋白质的剂型,其中施用的营养性蛋白质的总量选自约0.1g、0.1g-1g、1g、2g、3g、4g、5g、6g、7g、8g、9g、10g、15g、20g、25g、30g、35g、40g、45g、50g、55g、60g、65g、70g、75g、80g、85g、90g、95g和100g。

[0308] 在一些实施方案中,设计用于施用本文公开的至少一种营养性蛋白质的剂型,其中施用的必需氨基酸的总量选自0.1g至1g、1g至5g、2g至10g、5g至15g、10g至20g、以及1-30g。在一些实施方案中,设计用于施用本文公开的至少一种营养性蛋白质的剂型,其中施用的营养性蛋白质的总量选自约0.1g、0.1-1g、1g、2g、3g、4g、5g、6g、7g、8g、9g、10g、15g、20g、25g、30g、35g、40g、45g、50g、55g、60g、65g、70g、75g、80g、85g、90g、95g和100g。

[0309] 在一些实施方案中,营养性蛋白质或营养性组合物是在1天0.1g至1g、1天1g至5g、1天2g至10g、1天5g至15g、1天10g至20g、1天15g至30g、1天20g至40g、1天25g至50g、1天40g至80g、1天50g至100g或更大的速率下消耗。

[0310] 在由受试者达成的总蛋白质摄取量的一些实施方案中,由受试者历经膳食时期达成的总蛋白质摄取量的至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%或约100%是由至少一种根据本公开的营养性蛋白质构成。在由受试者达成的总蛋白质摄取量的一些实施方案中,历经膳食时期,由受试者达成的总蛋白质摄取量的5%至100%、由受试者达成的总蛋白质摄取量的5%至90%、由受试者达成的总蛋白质摄取量的5%至80%、由受试者达成的总蛋白质摄取量的5%至70%、由受试者达成的总蛋白质摄取量的5%至60%、由受试者达成的总蛋白质摄取量的5%至50%、由受试者达成的总蛋白质摄取量的5%至40%、由受试者达成的总蛋白质摄取量的5%至30%、由受试者达成的总蛋白质摄取量的5%至20%、由受试者达成的总蛋白质摄取量的5%至10%、由受试者达成的总蛋白质摄取量的10%至100%、由受试者达成的总蛋白质摄取量的10%至100%、由受试者达成的总蛋白质摄取量的20%至100%、由受试者达成的总蛋白质摄取量的30%至100%、由受试者达成的总蛋白质摄取量的40%至100%、由受试者达成的总蛋白质摄取量的50%至100%、由受试者达成的总蛋白质摄取量

的 60% 至 100%、由受试者达成的总蛋白质摄取量的 70% 至 100%、由受试者达成的总蛋白质摄取量的 80% 至 100%、或由受试者达成的总蛋白质摄取量的 90% 至 100% 是由至少一种根据本公开的营养性蛋白质构成。在一些实施方案中,本公开至少一种的营养性蛋白质占受试者历经膳食时期的卡路里摄取量的至少 5%、至少 10%、至少 15%、至少 20%、至少 25%、至少 30%、至少 35%、至少 40%、至少 45% 或至少 50%。

[0311] 在一些实施方案中,至少一种根据本公开的营养性蛋白质包含至少 2 种本公开的营养性蛋白质、至少 3 种本公开的营养性蛋白质、至少 4 种本公开的营养性蛋白质、至少 5 种本公开的营养性蛋白质、至少 6 种本公开的营养性蛋白质、至少 7 种本公开的营养性蛋白质、至少 8 种本公开的营养性蛋白质、至少 9 种本公开的营养性蛋白质、至少 10 种本公开的营养性蛋白质或更多种本公开的营养性蛋白质。

[0312] 在一些实施方案中,膳食时期是 1 餐、2 餐、3 餐、至少 1 天、至少 2 天、至少 3 天、至少 4 天、至少 5 天、至少 6 天、至少 1 周、至少 2 周、至少 3 周、至少 4 周、至少 1 个月、至少 2 个月、至少 3 个月、至少 4 个月、至少 5 个月、至少 6 个月或至少 1 年。在一些实施方案中,膳食时期是 1 天至 1 周、1 周至 4 周、1 个月至 3 个月、3 个月至 6 个月或 6 个月至 1 年。

[0313] 临床研究提供证据表明蛋白质会阻止归因于衰老或卧床休息的肌肉损失。具体来说,研究已显示补充蛋白质会增加长期卧床休息期间的肌肉分数合成率 (FSR),维持长期卧床休息期间的腿部质量和强度,增加瘦体重,改进步态和平衡的功能性量度,且可充当用于由于不活动或长期卧床休息而处于肌肉减少症的风险下的个体的可行性介入手段。(参见例如, Paddon-Jones D, 等 J Clin Endocrinol Metab 2004, 89:4351 - 4358; Ferrando, A 等 Clinical Nutrition 2009:1-6; Katsanos C 等 Am J Physiol Endocrinol Metab. 2006, 291:381-387)。

[0314] 在运动员中进行的关于增加肌肉蛋白质合成代谢的研究已显示在运动之后提供的蛋白质促进肌肉肥大的程度大于由单独运动达成的肌肉肥大。也已显示在运动之后提供的蛋白质支持蛋白质合成而不使蛋白质分解有任何增加,从而导致净正性蛋白质平衡和肌肉质量增大。尽管肌肉蛋白质合成似乎以剂量 - 反应方式对补充必需氨基酸起反应,但并非所有蛋白质在构建肌肉方面都等同。举例来说,氨基酸亮氨酸是刺激肌肉蛋白质合成的重要因素。(参见例如, Borsheim E 等 Am J Physiol Endocrinol Metab 2002, 283:E648-E657; Borsheim E 等 Clin Nutr. 2008, 27:189-95; Esmarck B 等 J Physiol 2001, 535:301-311; Moore D 等 Am J Clin Nutr 2009, 89:161-8)。

[0315] 在另一方面,本公开提供维持或增加受试者的肌肉质量、肌肉强度和功能性表现的至少一者的方法。在一些实施方案中,方法包括向受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中,受试者处于以下至少一种状态下: 年老、患有医学上危急疾病、以及罹患蛋白质 - 能量营养不良症。在一些实施方案中,足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中,本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0316] 在另一方面,本公开提供维持或达成受试者的合乎需要的身体质量指数的方法。在一些实施方案中,方法包括向受试者提供足量的本公开的营养性蛋白质、本公开的营养

性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中,受试者处于以下至少一种状态下:年长、患有医学上危急疾病、以及罹患蛋白质-能量营养不良症。在一些实施方案中,足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中,本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0317] 在另一方面,本公开提供向患有蛋白质-能量营养不良症的受试者提供蛋白质的方法。在一些实施方案中,方法包括向受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中,本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0318] 已建议需要在癌症患者和罹患恶病质的其它患者中补充必需氨基酸。在小鼠中进行的膳食研究已显示通过用必需氨基酸进行膳食性干预,对恶病质癌症携带小鼠具有存活性和功能性益处。除癌症之外,补充必需氨基酸也已在具有运动困难且因此遭受肌肉退化的罹患其它疾病(如慢性阻塞性肺病、慢性心脏衰竭、HIV 和其它疾病病况)的患者中显示各种益处,如肌肉功能和肌肉增加得以改进。

[0319] 研究已显示特定氨基酸有利于管理恶病质。膳食中的 BCAA 和 Leu 的含量相对较高被认为因通过发出使翻译增加的信号、增强胰岛素释放和抑制蛋白质降解来促进总蛋白质合成而在恶病质中具有正面作用。因此,消耗的膳食性 BCAA(一般来说)和/或 Leu(具体来说)增加将正面促进降低或逆转恶病质的影响。因为氮平衡在对抗恶病质的潜伏病因的方面是重要的,所以认为消耗的膳食性谷氨酰胺和/或精氨酸增加将正面促进降低或逆转恶病质的影响。(参见例如,Op den Kamp C,Langen R,Haegens A,Schols A. "Muscle atrophy in cachexia:can dietary protein tip the balance?" *Current Opinion in Clinical Nutrition and Metabolic Care* 2009,12:611-616;Poon RT-P,Yu W-C,Fan S-T,等"Long-term oral branched chain amino acids in patients undergoing chemoembolization for hepatocellular carcinoma:a randomized trial." *Aliment Pharmacol Ther* 2004;19:779-788;Tayek JA,Bistrian BR,Hehir DJ,Martin R,Moldawer LL,Blackburn GL. "Improved protein kinetics and albumin synthesis by branched chain amino acid-enriched total parenteral nutrition in cancer cachexia." *Cancer*.1986;58:147-57;Xi P,Jiang Z,Zheng C,Lin Y,Wu G "Regulation of protein metabolism by glutamine:implications for nutrition and health." *Front Biosci*.2011年1月1日;16:578-97)。

[0320] 因此,本文也提供治疗受试者的恶病质的方法。在一些实施方案中,本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物足以用于患有恶病质的受试者的量是以下量:由人摄取的所述量的本公开的蛋白质满足或超过代谢需要(所述需要常被提高)。每天每千克体重 1.5g 或总热量摄取量的 15-20% 的蛋白质摄取量似乎是患有恶病质的人士的适当目标。在一些实施方案中,由受试者消耗的全部蛋白质都是根据本公开的营养性蛋白质。在一些实施方案中,根据本公开的营养性蛋白质与其它来源的蛋白质和/或游离氨基酸组合以提供受试者的总蛋白质摄取量。在一些实施方案中,

受试者处于以下至少一种状态下：年长、患有医学上危急疾病、以及罹患蛋白质-能量营养不良症。在一些实施方案中，受试者罹患使得运动困难且因此导致肌肉退化的疾病，如慢性阻塞性肺病、慢性心脏衰竭、HIV、癌症和其它疾病病况。在一些实施方案中，根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0321] 肌肉减少症是骨骼肌肉质量（通常在 25 岁之后每年损失 0.5-1%）、品质和强度的与衰老相关的退行性损失。肌肉减少症是虚弱综合征的组成部分。欧洲关于年长人的肌肉减少症的工作组 (European Working Group on Sarcopenia in Older People, EWGSOP) 已建立针对年龄相关的肌肉减少症的实际临床规定和统一诊断准则。对于肌肉减少症的诊断，工作组已提出使用存在低肌肉质量与低肌肉功能（强度或表现）两者。肌肉减少症的特征首先在于肌肉萎缩（肌肉尺寸降低），以及由如肌肉纤维被脂肪置换的因素引起的肌肉组织“品质”降低、纤维化增加、肌肉代谢变化、氧化应激和神经肌肉接点退化。这些变化组合在一起导致肌肉功能进行性损失且最终导致虚弱。虚弱是一种在年长成人中常见的体现灾难性健康和功能衰退的风险提高的老年综合征。虚弱的促成因素可包括肌肉减少症、骨质疏松和肌肉衰弱。肌肉衰弱，也称为肌肉疲劳（或“缺乏强度”），是指不能用骨骼肌施加力量。衰弱常在肌肉萎缩和活动减少之后，如在由于疾病的一段长期卧床休息之后。也存在由于肌肉减少症的肌肉衰弱逐渐发作。

[0322] 本公开的营养性蛋白质适用于一旦肌肉减少症或虚弱在受试者中显现时对它进行治疗或适用于防止作为风险组的成员的受试者的肌肉减少症或虚弱的发作。在一些实施方案中，由受试者消耗的全部蛋白质都是根据本公开的营养性蛋白质。在一些实施方案中，根据本公开的营养性蛋白质与其它来源的蛋白质和 / 或游离氨基酸组合以提供受试者的总蛋白质摄取量。在一些实施方案中，受试者处于以下至少一种状态下：年长、患有医学上危急疾病、以及罹患蛋白质-能量营养不良症。在一些实施方案中，根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中，根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0323] 肥胖症是一种与许多合并症相关的多因性病症，所述合并症包括高血压、2 型糖尿病、血脂异常、冠心病、中风、癌症（例如，子宫内膜癌、乳癌和结肠癌）、骨关节炎、睡眠呼吸暂停和呼吸道问题。在美国，定义为身体质量指数  $>30\text{kg/m}^2$  的肥胖症的发病率已自 15% (1976 - 1980) 显著增加至 33% (2003 - 2004)，且它会继续增长。尽管促成肥胖症的机制是复杂的且涉及行为组成部分与激素、遗传和代谢过程的相互作用，但肥胖症主要被视为一种取决于生活方式的病状，具有 2 种主要病因：过度能量摄取和身体活动不足。关于能量摄取，有证据表明适度增加膳食中的蛋白质的比例，同时控制总能量摄取可改进身体组成、促进脂肪损失、以及改进在重量减轻之后的体重维持。与膳食蛋白质增加相关的正面结果被认为主要归因于与饱腹感增加相关的能量摄取降低、能量效率降低和 / 或产热增加、对身体组成（具体来说瘦性肌肉质量）的正面作用、以及血糖控制增强。

[0324] 膳食蛋白质在增加膳后能量消耗方面比等热量摄取的碳水化合物或脂肪更有效(参见例如, Dauncey M, Bingham S. “Dependence of 24h energy expenditure in man on composition of the nutrient intake.” *Br J Nutr* 1983, 50:1-13; Karst H 等 “Diet-induced thermogenesis in man: thermic effects of single proteins, carbohydrates and fats depending on their energy amount.” *Ann Nutr Metab.* 1984, 28:245-52; Tappy L 等 “Thermic effect of infused amino acids in healthy humans and in subjects with insulin resistance.” *Am J Clin Nutr* 1993, 57(6):912-6)。这个性质以及其它性质(饱腹感诱导; 维持瘦体重)使得蛋白质成为在重量管理时所针对的有吸引力的膳食组分。由所述膳食引起的能量消耗增加可部分归因于消化和代谢蛋白质的能量消耗高于其它卡路里来源的事实。包括蛋白质合成的蛋白质转换是一种能量消耗过程。此外, 高蛋白质膳食也可上调肝和棕色脂肪中的未偶联蛋白质, 此与能量消耗增加正相关。已建立理论的是不同蛋白质可对能量消耗具有独特影响。

[0325] 研究表明摄取蛋白质(特别是EAA和/或BCAA含量较高的蛋白质)会对产热和能量消耗产生不同影响(参见例如, Mikkelsen P. 等 “Effect of fat-reduced diets on 24h energy expenditure: comparisons between animal protein, vegetable protein and carbohydrate.” *Am J Clin Nutr* 2000, 72:1135-41; Acheson K. 等 “Protein choices targeting thermogenesis and metabolism.” *Am J Clin Nutr* 2011, 93:525-34; Alfenas R. 等 “Effects of protein quality on appetite and energy metabolism in normal weight subjects” *Arg Bras Endocrinol Metabol* 2010, 54(1):45-51; Lorenzen J. 等 “The effect of milk proteins on appetite regulation and diet-induced thermogenesis.” *J Clin Nutr* 2012, 66(5):622-7)。另外, L-酪氨酸已被鉴定为在产热中起作用的氨基酸(参见例如, Belza A. 等 “The beta-adrenergic antagonist propranolol partly abolishes thermogenic response to bioactive food ingredients.” *Metabolism* 2009, 58(8):1137-44)。进一步研究表明补充亮氨酸和精氨酸似乎通过将底物引向瘦体重而非脂肪组织来改变能量代谢(Dulloo A. “The search for compounds that stimulate thermogenesis in obesity management: from pharmaceuticals to functional food ingredients.” *Obes Rev* 2011, 12:866-83)。

[0326] 总之, 文献表明不同蛋白质类型对产热产生不同影响。因为富含EAA、BCAA和/或Tyr、Arg和Leu的至少一者的蛋白质或肽据信对产热具有刺激作用, 且因为刺激产热据信会对重量管理产生正面影响, 所以本公开也提供适用于刺激产热和/或大体上对重量管理产生正面影响的产品和方法。

[0327] 更具体来说, 本公开提供增加受试者的产热的方法。在一些实施方案中, 方法包括向受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中, 受试者是肥胖的。在一些实施方案中, 根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中, 根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0328] 在基本层面上, 显现超重病状的原因是归因于能量摄取与能量消耗之间的不平

衡。试图减少在任何特定时机（饱足）以及在进食时机之间（饱腹感）的食物已成为新近研究的主要焦点。由于在进餐期间感觉饱足以及在进餐之后感觉饱满而降低热量摄取是由内部信号和外部信号的复杂相互作用所致。各种营养研究已显示改变食物性质（如能量密度、内含物、质地和味道）会影响饱足与饱腹感两者。

[0329] 存在递送能量的三种大量营养素：脂肪、碳水化合物和蛋白质。1 克蛋白质或碳水化合物提供 4 卡而 1 克脂肪提供 9 卡。蛋白质使饱腹感增加的程度通常大于碳水化合物或脂肪且因此可促进卡路里摄取降低。然而，有可观证据指示蛋白质的类型在诱导饱腹感方面具有重要性（参见例如，W.L.Hall, 等“Casein and whey exert different effects on plasma amino acid profiles, gastrointestinal hormone secretion and appetite.” *Br J Nutr.* 2003 年 2 月, 89(2):239-48 ;R.Abou-Samra, 等“Effect of different protein sources on satiation and short-term satiety when consumed as a starter.” *Nutr J.* 2011 年 12 月 23 日, 10:139 ;T.Akhavan, 等“Effect of premeal consumption of whey protein and its hydrolysate on food intake and postmeal glycemia and insulin responses in young adults.” *Am J Clin Nutr.* 2010 年 4 月, 91(4):966-75, 电子发表 2010 年 2 月 17 日 ;MA Veldhorst“Dose-dependent satiating effect of whey relative to casein or soy” *Physiol Behav.* 2009 年 3 月 23 日, 96(4-5):675-82)。证据指示富含亮氨酸的蛋白质在诱导饱腹感方面特别有效（参见例如，Fromentin G 等“Peripheral and central mechanisms involved in the control of food intake by dietary amino acids and proteins.” *Nutr Res Rev* 201225:29-39)。

[0330] 在一些实施方案中，本公开的营养性蛋白质是与至少一种医药或生物药物产品同时由受试者消耗。在一些实施方案中，营养性蛋白质和至少一种医药或生物药物产品的有益作用具有累加效应，而在一些实施方案中，营养性蛋白质和至少一种医药或生物药物产品的有益作用具有协同效应。可与本公开的营养性蛋白质一起施用的医药或生物药物产品的实例在本领域中是熟知的。举例来说，当本公开的营养性蛋白质用于维持或增加受试者的肌肉质量、肌肉强度和功能性表现的至少一者时，所述营养性蛋白质可与某一治疗剂量方案的至少一种指示用以维持或增加受试者的肌肉质量、肌肉强度和功能性表现的至少一者的医药或生物药物产品（如合成代谢类固醇）同时由受试者消耗。当本公开的营养性蛋白质用于维持或达成受试者的合乎需要的身体质量指数时，所述营养性蛋白质可与某一治疗剂量方案的至少一种指示用以维持或达成受试者的合乎需要的身体质量指数的医药或生物药物产品（如奥利斯他 (orlistat)、氯卡色林 (lorcaserin)、西布曲明 (sibutramine)、利莫那班 (rimonabant)、二甲双胍 (metformin)、艾塞那肽 (exenatide) 或普兰林肽 (pramlintide)) 同时由受试者消耗。当本公开的营养性蛋白质用于诱导受试者的饱足反应和饱腹感反应的至少一者时，所述营养性蛋白质可与某一治疗剂量方案的至少一种指示用以诱导受试者的饱足反应和饱腹感反应的至少一者的医药或生物药物产品（如利莫那班、艾塞那肽或普兰林肽）同时由受试者消耗。当本公开的营养性蛋白质用于治疗受试者的恶病质、肌肉减少症和虚弱的至少一者时，所述营养性蛋白质可与某一治疗剂量方案的至少一种指示用以治疗恶病质、肌肉减少症和虚弱的至少一者的医药或生物药物产品（如  $\omega$ -3 脂肪酸或合成代谢类固醇）同时由受试者消耗。由于膳食蛋白质在诱导饱足和饱腹感方面的作用，本文公开的营养性蛋白质和营养性组合物可用于诱导受试者的饱

足反应和饱腹感反应的至少一者。在一些实施方案中,方法包括向受试者提供足量的本公开的营养性蛋白质、本公开的营养性组合物或通过本公开的方法制备的营养性组合物。在一些实施方案中,受试者是肥胖的。在一些实施方案中,根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者与进行运动配合加以消耗。在一些实施方案中,根据本公开的营养性蛋白质、根据本公开的营养性组合物或通过根据本公开的方法制备的营养性组合物是由受试者通过经口、肠或胃肠外途径加以消耗。

[0331] 在一些实施方案中,将至少一种本公开的营养性蛋白质或营养性组合物并入受试者的膳食中具有至少一种选自以下的作用:诱导餐后饱腹感(包括通过抑制饥饿感)、诱导产热、降低血糖反应、正面影响能量消耗、正面影响瘦体重、降低由过度进食引起的重量增加、以及降低能量摄取。在一些实施方案中,将至少一种本公开的营养性蛋白质或营养性组合物并入受试者的膳食中在所述受试者中具有至少一种选自以下的作用:增加身体脂肪损失、降低瘦性组织损失、改进脂质分布、以及改进葡萄糖耐受性和胰岛素敏感性。

## 实施例

[0332] 以下实施例用于更充分描述使用本发明的方式。这些实施例是出于说明目的而呈现且不应用于限制本发明的真实范围。

[0333] 实施例 1:含有的必需氨基酸、支链氨基酸和亮氨酸的比率大于或等于乳清以及含有所有必需氨基酸的蛋白质片段的鉴定

[0334] UniProtKB/Swiss-Prot(欧洲生物信息学研究所和瑞士生物信息学研究所的合作)是手动策划和复查的蛋白质数据库,且用作蛋白质鉴定的起始点。将来自可食用物质番茄、玉米、粳稻、大豆、绵羊、豌豆、菠菜、籼稻、小麦、野猪、桃、樱桃辣椒、苹果、黄鳍鲷、家山羊、鹰嘴豆、盐湖鲑鱼、火鸡、马铃薯和双孢子蘑菇的具有大于或等于五十(50)个氨基酸的蛋白质视为可行目标。选择大致 8,415 种所述蛋白质供评估且接着使蛋白质断裂成原始蛋白质的大致五十(50)、大致一百(100)、或大致一百五十(150)个氨基酸的部分,所述部分被设计来在两端含有消化酶裂解位点。此外,相对于已知过敏原的数据库筛选蛋白质以确定任何蛋白质是否与已知过敏原具有大于 50%整体同源性。鉴定出总计 1,000 个片段含有大于或等于 51% EAA、大于或等于 25% BCAA 和大于或等于 13% Leu,且含有所有必需氨基酸,并与已知过敏原具有小于 50%整体同源性(SEQ ID NO:1 至 1000)。(应注意对于实施例 1-5,51% EAA/25% BCAA/13% Leu 表示各自比乳清的 EAA/BCAA/Leu(在本文中定义为 49% EAA/24% BCAA/11% Leu)大 1-2% 的值。这些值仅用于出于这个实施例的目的鉴定目标营养性蛋白质片段,以确保鉴定的蛋白质具有的 EAA、BCAA 和 Leu 含量高于乳清)。对于所述组的蛋白质,计算在 pH 7 下的溶合计分(“SolvScore”)、在 pH 7 下的聚集计分(“AggScore”)、过敏原性(即,如本文所述的与已知过敏原的局部同源性百分比)、毒性(即,如本文所述的与已知毒素的同源性百分比)、抗营养性(即,如本文所述的与已知蛋白酶抑制剂的同源性百分比)、和人同源性(即,如本文所述的与已知人蛋白质的同源性百分比),且测定 Cys 残基(“C”)的总数。由此鉴定的 75 种代表性片段的特征呈现于表 3A 和 3B 中。

[0335] 表 3A



[0336]

| Seq ID No | UniProt | FragEnds | EAA  | BCAA | L    | C |
|-----------|---------|----------|------|------|------|---|
| 1         | Q29550  | 101 :301 | 0.53 | 0.27 | 0.13 | 3 |
| 2         | Q29078  | 1 :176   | 0.52 | 0.27 | 0.14 | 5 |
| 3         | Q29550  | 201 :426 | 0.56 | 0.25 | 0.13 | 2 |
| 4         | Q29078  | 1 :76    | 0.60 | 0.34 | 0.20 | 3 |
| 5         | Q9N2D1  | 1 :76    | 0.57 | 0.27 | 0.18 | 3 |
| 6         | Q29079  | 1 :76    | 0.56 | 0.29 | 0.18 | 1 |
| 7         | Q29550  | 101 :176 | 0.53 | 0.31 | 0.15 | 1 |
| 8         | Q70BM6  | 201 :326 | 0.58 | 0.26 | 0.15 | 0 |
| 9         | Q70BM6  | 201 :426 | 0.59 | 0.30 | 0.17 | 3 |
| 10        | Q70BM6  | 101 :326 | 0.56 | 0.28 | 0.14 | 2 |
| 11        | O81122  | 101 :301 | 0.53 | 0.29 | 0.16 | 1 |
| 12        | Q9M7M1  | 101 :201 | 0.55 | 0.28 | 0.19 | 1 |
| 13        | O81122  | 101 :201 | 0.55 | 0.28 | 0.19 | 1 |
| 14        | Q9M7M1  | 101 :301 | 0.53 | 0.28 | 0.16 | 1 |
| 15        | Q8WNV7  | 151 :226 | 0.51 | 0.29 | 0.14 | 1 |
| 16        | P50118  | 176 :276 | 0.52 | 0.26 | 0.15 | 3 |
| 17        | P12263  | 451 :526 | 0.55 | 0.26 | 0.13 | 0 |
| 18        | Q29197  | 26 :126  | 0.52 | 0.29 | 0.18 | 0 |
| 19        | O02799  | 151 :301 | 0.54 | 0.26 | 0.19 | 2 |
| 20        | P80229  | 201 :276 | 0.55 | 0.27 | 0.18 | 1 |
| 21        | Q29624  | 226 :326 | 0.59 | 0.25 | 0.13 | 2 |
| 22        | A5GFQ5  | 26 :151  | 0.52 | 0.25 | 0.15 | 3 |

|    |        |            |       |       |       |   |
|----|--------|------------|-------|-------|-------|---|
| 23 | Q5GN48 | 1601 :1676 | 0. 57 | 0. 26 | 0. 15 | 0 |
| 24 | Q5GN48 | 1026 :1101 | 0. 55 | 0. 26 | 0. 16 | 2 |
| 25 | A6M931 | 176 :276   | 0. 54 | 0. 26 | 0. 13 | 1 |
| 26 | P49039 | 1 :101     | 0. 51 | 0. 27 | 0. 14 | 0 |
| 27 | P79430 | 226 :326   | 0. 62 | 0. 25 | 0. 14 | 3 |
| 28 | P61291 | 1 :151     | 0. 53 | 0. 27 | 0. 14 | 2 |
| 29 | B5X2B8 | 126 :326   | 0. 52 | 0. 25 | 0. 13 | 3 |
| 30 | Q1A730 | 226 :301   | 0. 52 | 0. 28 | 0. 20 | 2 |
| 31 | D2SW95 | 276 :476   | 0. 52 | 0. 31 | 0. 13 | 3 |
| 32 | 002799 | 251 :326   | 0. 57 | 0. 29 | 0. 22 | 1 |

[0337]

|    |        |            |       |       |       |   |
|----|--------|------------|-------|-------|-------|---|
| 33 | B5R154 | 126 :251   | 0. 54 | 0. 26 | 0. 13 | 2 |
| 34 | P21999 | 76 :176    | 0. 51 | 0. 28 | 0. 13 | 5 |
| 35 | Q07511 | 151 :276   | 0. 52 | 0. 27 | 0. 13 | 2 |
| 36 | A5A779 | 351 :476   | 0. 53 | 0. 31 | 0. 24 | 4 |
| 37 | Q61TT4 | 476 :601   | 0. 54 | 0. 30 | 0. 16 | 1 |
| 38 | 046374 | 876 :1076  | 0. 55 | 0. 26 | 0. 13 | 2 |
| 39 | Q4U116 | 901 :1076  | 0. 55 | 0. 25 | 0. 14 | 2 |
| 40 | P04045 | 376 :476   | 0. 53 | 0. 27 | 0. 13 | 1 |
| 41 | Q29121 | 151 :226   | 0. 52 | 0. 30 | 0. 13 | 2 |
| 42 | Q2VEC3 | 1601 :1726 | 0. 54 | 0. 27 | 0. 16 | 1 |
| 43 | B5X737 | 26 :176    | 0. 57 | 0. 26 | 0. 13 | 2 |
| 44 | Q9XSW2 | 201 :401   | 0. 53 | 0. 25 | 0. 15 | 8 |
| 45 | P19756 | 276 :351   | 0. 60 | 0. 32 | 0. 14 | 0 |

|    |        |           |       |       |       |   |
|----|--------|-----------|-------|-------|-------|---|
| 46 | P67931 | 101 :201  | 0. 51 | 0. 26 | 0. 19 | 1 |
| 47 | Q9AWA5 | 401 :476  | 0. 58 | 0. 27 | 0. 13 | 0 |
| 48 | 046374 | 976 :1051 | 0. 56 | 0. 26 | 0. 19 | 2 |
| 49 | Q9XSW2 | 301 :401  | 0. 59 | 0. 29 | 0. 16 | 3 |
| 50 | Q8M1Q9 | 551 :651  | 0. 56 | 0. 28 | 0. 15 | 1 |
| 51 | 002668 | 526 :626  | 0. 53 | 0. 26 | 0. 14 | 0 |
| 52 | Q6UAQ8 | 51 :226   | 0. 51 | 0. 27 | 0. 13 | 3 |
| 53 | P21343 | 326 :401  | 0. 54 | 0. 33 | 0. 17 | 1 |
| 54 | P79430 | 126 :326  | 0. 58 | 0. 26 | 0. 13 | 5 |
| 55 | P10691 | 26 :176   | 0. 52 | 0. 25 | 0. 14 | 0 |
| 56 | 002697 | 626 :701  | 0. 52 | 0. 27 | 0. 19 | 1 |
| 57 | Q5V141 | 676 :751  | 0. 53 | 0. 30 | 0. 13 | 1 |
| 58 | 097716 | 201 :376  | 0. 52 | 0. 25 | 0. 13 | 4 |
| 59 | Q66UW5 | 51 :151   | 0. 52 | 0. 25 | 0. 16 | 2 |
| 60 | D2SW95 | 1 :201    | 0. 54 | 0. 27 | 0. 14 | 5 |
| 61 | 046491 | 201 :351  | 0. 53 | 0. 25 | 0. 15 | 1 |
| 62 | Q52QT8 | 1 :151    | 0. 53 | 0. 26 | 0. 16 | 5 |
| 63 | Q9P4T7 | 1 :76     | 0. 53 | 0. 25 | 0. 17 | 0 |
| 64 | Q2VEE8 | 201 :276  | 0. 60 | 0. 25 | 0. 14 | 0 |
| 65 | P37831 | 251 :376  | 0. 52 | 0. 26 | 0. 16 | 0 |
| 66 | B9EP11 | 1 :126    | 0. 56 | 0. 30 | 0. 16 | 4 |
| 67 | B5X2B8 | 126 :201  | 0. 61 | 0. 30 | 0. 15 | 0 |

[0338]

|    |        |          |       |       |       |   |
|----|--------|----------|-------|-------|-------|---|
| 68 | 002799 | 526 :651 | 0. 57 | 0. 26 | 0. 16 | 1 |
|----|--------|----------|-------|-------|-------|---|

|    |        |            |       |       |       |   |
|----|--------|------------|-------|-------|-------|---|
| 69 | Q99028 | 76 :151    | 0. 53 | 0. 30 | 0. 16 | 3 |
| 70 | Q07511 | 51 :276    | 0. 51 | 0. 26 | 0. 13 | 4 |
| 71 | P08454 | 1 :76      | 0. 57 | 0. 28 | 0. 13 | 1 |
| 72 | Q5GN48 | 2251 :2501 | 0. 51 | 0. 26 | 0. 14 | 0 |
| 73 | 064390 | 276 :426   | 0. 52 | 0. 25 | 0. 15 | 1 |
| 74 | P09954 | 1 :101     | 0. 53 | 0. 31 | 0. 14 | 0 |
| 75 | P29196 | 801 :876   | 0. 57 | 0. 28 | 0. 18 | 0 |

[0339] 表 3B

[0340]

| Seq ID No | 溶合计分   | 聚集计分 | 过敏原性 | 毒性   | 抗营养性 | 人同源性 |
|-----------|--------|------|------|------|------|------|
| 1         | -16.14 | 0.70 | 0.58 | 0.40 | 0.23 | 0.83 |
| 2         | -15.72 | 0.57 | 0.54 | 0.00 | 0.25 | 0.50 |
| 3         | -18.83 | 0.61 | 0.45 | 0.30 | 0.00 | 0.76 |
| 4         | -18.32 | 0.80 | 0.43 | 0.28 | 0.29 | 0.46 |
| 5         | -10.78 | 0.78 | 0.42 | 0.42 | 0.29 | 0.78 |
| 6         | -18.03 | 0.62 | 0.41 | 0.28 | 0.28 | 0.48 |
| 7         | -14.86 | 0.78 | 0.40 | 0.43 | 0.26 | 0.79 |
| 8         | -17.96 | 0.60 | 0.36 | 0.00 | 0.25 | 0.81 |
| 9         | -13.70 | 0.90 | 0.36 | 0.00 | 0.21 | 0.82 |
| 10        | -15.03 | 0.80 | 0.36 | 0.00 | 0.22 | 0.81 |
| 11        | -18.93 | 0.53 | 0.36 | 0.00 | 0.21 | 0.00 |
| 12        | -22.36 | 0.47 | 0.35 | 0.21 | 0.00 | 0.00 |
| 13        | -22.36 | 0.47 | 0.35 | 0.21 | 0.00 | 0.00 |
| 14        | -18.90 | 0.52 | 0.35 | 0.00 | 0.24 | 0.00 |
| 15        | -12.53 | 0.69 | 0.35 | 0.00 | 0.00 | 0.82 |
| 16        | -20.41 | 0.43 | 0.35 | 0.20 | 0.25 | 0.86 |
| 17        | -18.65 | 0.43 | 0.25 | 0.43 | 0.24 | 0.83 |
| 18        | -28.11 | 0.37 | 0.31 | 0.00 | 0.00 | 0.98 |
| 19        | -26.99 | 0.37 | 0.33 | 0.24 | 0.22 | 0.81 |
| 20        | -26.56 | 0.35 | 0.34 | 0.29 | 1.00 | 0.79 |
| 21        | -26.33 | 0.75 | 0.30 | 0.29 | 0.22 | 0.84 |
| 22        | -25.32 | 0.42 | 0.29 | 0.23 | 0.29 | 0.95 |
| 23        | -25.11 | 0.34 | 0.29 | 0.00 | 0.00 | 0.93 |

[0341]

|    |        |      |      |      |      |      |
|----|--------|------|------|------|------|------|
| 24 | -24.91 | 0.25 | 0.28 | 0.00 | 0.26 | 0.88 |
| 25 | -24.78 | 0.47 | 0.31 | 0.00 | 0.20 | 1.00 |
| 26 | -24.60 | 0.44 | 0.29 | 0.25 | 0.31 | 0.25 |
| 27 | -24.55 | 0.79 | 0.28 | 0.24 | 0.27 | 0.77 |
| 28 | -24.41 | 0.38 | 0.33 | 0.00 | 0.25 | 1.00 |
| 29 | -24.09 | 0.41 | 0.31 | 0.00 | 0.25 | 0.50 |
| 30 | -23.74 | 0.38 | 0.27 | 0.26 | 0.26 | 0.99 |
| 31 | -23.68 | 0.62 | 0.31 | 0.00 | 0.23 | 1.00 |
| 32 | -23.68 | 0.38 | 0.27 | 0.00 | 0.24 | 0.80 |
| 33 | -23.57 | 0.39 | 0.32 | 0.00 | 0.00 | 0.89 |
| 34 | -23.51 | 0.64 | 0.33 | 0.17 | 0.27 | 0.99 |
| 35 | -23.41 | 0.45 | 0.34 | 0.21 | 0.00 | 0.29 |
| 36 | -23.38 | 0.50 | 0.32 | 0.23 | 0.26 | 0.98 |
| 37 | -23.11 | 0.63 | 0.30 | 0.22 | 0.24 | 1.00 |
| 38 | -23.11 | 0.41 | 0.33 | 0.22 | 0.00 | 0.99 |
| 39 | -23.02 | 0.72 | 0.29 | 0.00 | 0.23 | 0.99 |
| 40 | -22.98 | 0.36 | 0.30 | 0.00 | 0.28 | 0.57 |
| 41 | -22.88 | 0.59 | 0.30 | 0.27 | 0.29 | 1.00 |
| 42 | -22.87 | 0.49 | 0.32 | 0.00 | 0.23 | 0.00 |
| 43 | -22.79 | 0.42 | 0.31 | 0.21 | 0.23 | 0.59 |
| 44 | -22.74 | 0.56 | 0.31 | 0.00 | 0.21 | 0.88 |
| 45 | -22.73 | 0.80 | 0.26 | 0.00 | 0.30 | 0.91 |
| 46 | -22.72 | 0.50 | 0.29 | 0.00 | 0.30 | 0.73 |
| 47 | -22.62 | 0.38 | 0.23 | 0.23 | 0.00 | 0.24 |
| 48 | -22.55 | 0.42 | 0.24 | 0.27 | 0.24 | 1.00 |
| 49 | -22.53 | 0.71 | 0.31 | 0.00 | 0.26 | 0.96 |
| 50 | -22.43 | 0.59 | 0.30 | 0.00 | 0.00 | 0.90 |
| 51 | -22.38 | 0.48 | 0.30 | 0.00 | 1.00 | 0.76 |
| 52 | -22.32 | 0.47 | 0.33 | 0.00 | 0.00 | 0.96 |
| 53 | -22.21 | 0.49 | 0.26 | 0.00 | 0.26 | 0.30 |
| 54 | -22.18 | 0.64 | 0.31 | 0.20 | 0.21 | 0.78 |
| 55 | -22.18 | 0.45 | 0.32 | 0.20 | 0.24 | 0.00 |
| 56 | -22.16 | 0.52 | 0.24 | 0.00 | 0.23 | 1.00 |
| 57 | -22.06 | 0.82 | 0.27 | 0.24 | 0.27 | 0.79 |
| 58 | -22.05 | 0.59 | 0.30 | 0.00 | 0.22 | 1.00 |

[0342]

|    |        |      |      |      |      |      |
|----|--------|------|------|------|------|------|
| 59 | -22.01 | 0.36 | 0.29 | 0.00 | 0.00 | 0.96 |
| 60 | -21.91 | 0.62 | 0.33 | 0.00 | 0.21 | 1.00 |
| 61 | -21.88 | 0.46 | 0.33 | 0.21 | 0.18 | 0.73 |
| 62 | -21.83 | 0.49 | 0.30 | 0.24 | 0.22 | 0.99 |
| 63 | -21.80 | 0.39 | 0.29 | 0.00 | 0.30 | 0.51 |
| 64 | -21.75 | 0.44 | 0.28 | 0.00 | 0.24 | 0.22 |
| 65 | -21.74 | 0.58 | 0.31 | 0.00 | 0.26 | 0.00 |
| 66 | -21.65 | 0.52 | 0.33 | 0.25 | 0.26 | 0.82 |
| 67 | -21.59 | 0.58 | 0.24 | 0.00 | 0.27 | 0.71 |
| 68 | -21.58 | 0.40 | 0.30 | 0.26 | 0.25 | 0.85 |
| 69 | -21.57 | 0.64 | 0.25 | 0.22 | 0.28 | 0.96 |
| 70 | -21.54 | 0.46 | 0.31 | 0.22 | 0.00 | 0.28 |
| 71 | -21.50 | 0.73 | 0.25 | 0.28 | 1.00 | 0.28 |
| 72 | -21.47 | 0.34 | 0.32 | 0.22 | 0.00 | 0.86 |
| 73 | -21.43 | 0.46 | 0.29 | 0.22 | 0.25 | 0.38 |
| 74 | -21.41 | 0.62 | 0.31 | 0.00 | 0.27 | 0.85 |
| 75 | -21.41 | 0.58 | 0.28 | 0.00 | 0.00 | 0.29 |

[0343] 实施例 2 : 含有的必需氨基酸、支链氨基酸和亮氨酸的比率大于或等于乳清的蛋白质片段的鉴定

[0344] 评估来自实施例 1 中所述的所述组的 8,415 种蛋白质的片段的氨基酸含量、必需氨基酸 (“EAA”) 的百分比、支链氨基酸 (“BCAA”) 的百分比、亮氨酸 (“L”) 的百分比, 且对于各蛋白质计算蛋白质是否含有所有必需氨基酸。此外, 相对于已知过敏原的数据库筛选蛋白质片段以确定任何蛋白质片段是否与已知过敏原具有大于 50% 整体同源性。鉴定出总计 414 个蛋白质片段含有大于或等于 51% EAA、大于或等于 25% BCAA 和大于或等于 13% Leu, 且与已知过敏原具有小于 50% 整体同源性 (SEQ ID NO:1001 至 1414)。对于所述组的蛋白质, 计算在 pH 7 下的溶合计分 (“SolvScore”)、在 pH 7 下的聚集计分 (“AggScore”)、过敏原性 (即, 如本文所述的与已知过敏原的局部同源性百分比)、毒性 (即, 如本文所述的与已知毒素的同源性百分比)、抗营养性 (即, 如本文所述的与已知蛋白酶抑制剂的同源性百分比)、和人同源性 (即, 如本文所述的与已知人蛋白质的同源性百分比), 且测定 Cys 残基 (“C”) 的总数。由此鉴定的 75 种代表性蛋白质的特征呈现于表 4A 和 4B 中。

[0345] 表 4A

[0346]

| Seq ID No | UniProt | FragEnds | EAA  | BCAA | L    | C |
|-----------|---------|----------|------|------|------|---|
| 1001      | P21641  | 1 :126   | 0.52 | 0.30 | 0.19 | 1 |
| 1002      | Q41782  | 151 :276 | 0.52 | 0.25 | 0.14 | 4 |
| 1003      | O24174  | 176 :251 | 0.53 | 0.29 | 0.15 | 3 |

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1004 | Q41784 | 201 :276   | 0. 55 | 0. 26 | 0. 17 | 3 |
| 1005 | P09315 | 276 :351   | 0. 51 | 0. 35 | 0. 14 | 1 |
| 1006 | Q43695 | 201 :276   | 0. 53 | 0. 26 | 0. 17 | 3 |
| 1007 | P18025 | 201 :276   | 0. 54 | 0. 26 | 0. 17 | 3 |
| 1008 | P18026 | 201 :276   | 0. 54 | 0. 26 | 0. 17 | 3 |
| 1009 | Q43697 | 201 :276   | 0. 54 | 0. 26 | 0. 17 | 3 |
| 1010 | Q41783 | 201 :276   | 0. 54 | 0. 26 | 0. 17 | 3 |
| 1011 | Q41785 | 201 :276   | 0. 54 | 0. 26 | 0. 17 | 3 |
| 1012 | 024562 | 126 :276   | 0. 53 | 0. 28 | 0. 14 | 1 |
| 1013 | Q6ZHS4 | 126 :301   | 0. 52 | 0. 27 | 0. 13 | 1 |
| 1014 | Q6ERW7 | 226 :301   | 0. 55 | 0. 29 | 0. 13 | 1 |
| 1015 | Q01595 | 51 :126    | 0. 52 | 0. 36 | 0. 22 | 7 |
| 1016 | Q6ERW9 | 226 :301   | 0. 54 | 0. 29 | 0. 13 | 1 |
| 1017 | Q0J6T3 | 151 :276   | 0. 55 | 0. 28 | 0. 13 | 1 |
| 1018 | P93203 | 201 :276   | 0. 55 | 0. 25 | 0. 18 | 0 |
| 1019 | Q7XWU3 | 151 :276   | 0. 52 | 0. 27 | 0. 14 | 1 |
| 1020 | Q41874 | 126 :251   | 0. 51 | 0. 29 | 0. 16 | 2 |
| 1021 | Q8H859 | 151 :276   | 0. 54 | 0. 25 | 0. 14 | 1 |
| 1022 | POCD58 | 301 :476   | 0. 53 | 0. 32 | 0. 18 | 3 |
| 1023 | POCD59 | 301 :476   | 0. 53 | 0. 32 | 0. 18 | 3 |
| 1024 | 004161 | 251 :376   | 0. 58 | 0. 26 | 0. 13 | 5 |
| 1025 | Q2R114 | 151 :276   | 0. 54 | 0. 26 | 0. 13 | 1 |
| 1026 | Q5Z987 | 1051 :1276 | 0. 54 | 0. 30 | 0. 14 | 3 |
| 1027 | Q41808 | 176 :276   | 0. 51 | 0. 28 | 0. 17 | 2 |

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1028 | Q41807 | 176 :276 | 0. 51 | 0. 29 | 0. 17 | 2 |
| 1029 | P54773 | 101 :276 | 0. 55 | 0. 29 | 0. 15 | 0 |

[0347]

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1030 | Q5Z987 | 1151 :1251 | 0. 60 | 0. 34 | 0. 14 | 2 |
| 1031 | Q5Z987 | 1151 :1351 | 0. 54 | 0. 33 | 0. 16 | 4 |
| 1032 | P31542 | 326 :401   | 0. 53 | 0. 30 | 0. 14 | 0 |
| 1033 | P31541 | 126 :201   | 0. 51 | 0. 32 | 0. 15 | 0 |
| 1034 | P04706 | 1 :76      | 0. 51 | 0. 31 | 0. 17 | 2 |
| 1035 | 064411 | 1 :76      | 0. 53 | 0. 32 | 0. 18 | 0 |
| 1036 | Q0JI49 | 351 :426   | 0. 60 | 0. 26 | 0. 17 | 0 |
| 1037 | Q9XGC9 | 476 :551   | 0. 57 | 0. 26 | 0. 13 | 0 |
| 1038 | P93648 | 326 :401   | 0. 51 | 0. 25 | 0. 17 | 1 |
| 1039 | Q6H7U5 | 376 :451   | 0. 56 | 0. 26 | 0. 14 | 0 |
| 1040 | Q43143 | 76 :176    | 0. 52 | 0. 26 | 0. 14 | 2 |
| 1041 | Q9LGV5 | 51 :126    | 0. 52 | 0. 27 | 0. 14 | 0 |
| 1042 | POC8M8 | 126 :201   | 0. 51 | 0. 26 | 0. 16 | 4 |
| 1043 | Q05761 | 51 :126    | 0. 56 | 0. 32 | 0. 16 | 0 |
| 1044 | Q84NP7 | 326 :426   | 0. 54 | 0. 26 | 0. 16 | 1 |
| 1045 | Q75HC2 | 276 :351   | 0. 51 | 0. 27 | 0. 14 | 2 |
| 1046 | Q6QNU4 | 126 :201   | 0. 52 | 0. 27 | 0. 14 | 2 |
| 1047 | Q5JMF2 | 51 :126    | 0. 55 | 0. 26 | 0. 14 | 4 |
| 1048 | Q7XU38 | 276 :351   | 0. 54 | 0. 25 | 0. 16 | 1 |
| 1049 | Q8GUQ5 | 1076 :1151 | 0. 53 | 0. 27 | 0. 16 | 2 |
| 1050 | P51059 | 76 :176    | 0. 52 | 0. 27 | 0. 14 | 0 |



|      |        |           |       |       |       |   |
|------|--------|-----------|-------|-------|-------|---|
| 1051 | POC8M8 | 26 :176   | 0. 54 | 0. 27 | 0. 14 | 6 |
| 1052 | P49037 | 26 :101   | 0. 51 | 0. 29 | 0. 15 | 0 |
| 1053 | 049187 | 101 :176  | 0. 56 | 0. 30 | 0. 18 | 0 |
| 1054 | Q0J411 | 26 :176   | 0. 52 | 0. 25 | 0. 14 | 2 |
| 1055 | Q6ZAG3 | 51 :176   | 0. 53 | 0. 25 | 0. 13 | 1 |
| 1056 | P29620 | 51 :126   | 0. 54 | 0. 27 | 0. 14 | 0 |
| 1057 | P33544 | 201 :276  | 0. 54 | 0. 26 | 0. 15 | 1 |
| 1058 | 024573 | 1 :101    | 0. 55 | 0. 28 | 0. 14 | 2 |
| 1059 | Q3HRN8 | 126 :201  | 0. 55 | 0. 26 | 0. 15 | 0 |
| 1060 | Q3HRP5 | 1 :101    | 0. 53 | 0. 27 | 0. 15 | 3 |
| 1061 | P93647 | 801 :876  | 0. 53 | 0. 32 | 0. 16 | 0 |
| 1062 | Q9ZS62 | 951 :1026 | 0. 53 | 0. 35 | 0. 16 | 0 |
| 1063 | Q6Z4U2 | 276 :351  | 0. 53 | 0. 29 | 0. 13 | 1 |
| 1064 | P29185 | 201 :301  | 0. 57 | 0. 31 | 0. 13 | 1 |

[0348]

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1065 | Q75LU8 | 1 :76    | 0. 51 | 0. 31 | 0. 19 | 4 |
| 1066 | Q43298 | 201 :301 | 0. 55 | 0. 31 | 0. 13 | 1 |
| 1067 | P49971 | 26 :101  | 0. 52 | 0. 27 | 0. 13 | 2 |
| 1068 | P46641 | 1 :76    | 0. 64 | 0. 26 | 0. 16 | 1 |
| 1069 | Q2QMX9 | 701 :776 | 0. 52 | 0. 25 | 0. 14 | 0 |
| 1070 | Q43270 | 226 :301 | 0. 53 | 0. 25 | 0. 13 | 0 |
| 1071 | P93841 | 76 :176  | 0. 52 | 0. 26 | 0. 15 | 1 |
| 1072 | P10581 | 151 :226 | 0. 52 | 0. 29 | 0. 14 | 1 |
| 1073 | P93647 | 151 :251 | 0. 51 | 0. 26 | 0. 14 | 0 |

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1074 | Q2QQS5 | 126 :201 | 0. 52 | 0. 30 | 0. 13 | 0 |
| 1075 | Q0JNK5 | 301 :426 | 0. 52 | 0. 34 | 0. 15 | 1 |

[0349] 表 4B

[0350]

| Seq ID No | 溶合计分    | 聚集计分  | 过敏原性  | 毒性    | 抗营养性  | 人同源性  |
|-----------|---------|-------|-------|-------|-------|-------|
| 1001      | -7. 76  | 1. 07 | 0. 63 | 0. 00 | 0. 00 | 0. 28 |
| 1002      | -16. 11 | 0. 62 | 0. 49 | 0. 24 | 0. 26 | 0. 87 |
| 1003      | -11. 65 | 0. 84 | 0. 48 | 0. 00 | 0. 29 | 0. 54 |
| 1004      | -13. 93 | 0. 66 | 0. 46 | 0. 00 | 0. 27 | 0. 87 |
| 1005      | -15. 37 | 0. 78 | 0. 45 | 0. 32 | 0. 00 | 0. 47 |
| 1006      | -14. 07 | 0. 63 | 0. 45 | 0. 00 | 0. 27 | 0. 86 |
| 1007      | -14. 01 | 0. 63 | 0. 45 | 0. 00 | 0. 27 | 0. 87 |
| 1008      | -14. 01 | 0. 63 | 0. 45 | 0. 00 | 0. 27 | 0. 87 |
| 1009      | -14. 01 | 0. 63 | 0. 45 | 0. 00 | 0. 27 | 0. 87 |
| 1010      | -14. 01 | 0. 63 | 0. 45 | 0. 00 | 0. 27 | 0. 87 |
| 1011      | -14. 01 | 0. 63 | 0. 45 | 0. 00 | 0. 27 | 0. 87 |
| 1012      | -14. 02 | 0. 70 | 0. 44 | 0. 24 | 0. 27 | 0. 29 |
| 1013      | -13. 33 | 0. 73 | 0. 42 | 0. 00 | 0. 26 | 0. 27 |
| 1014      | -10. 77 | 0. 79 | 0. 42 | 0. 00 | 0. 30 | 0. 34 |
| 1015      | -14. 67 | 0. 92 | 0. 40 | 0. 31 | 0. 29 | 0. 28 |
| 1016      | -13. 59 | 0. 79 | 0. 39 | 0. 00 | 0. 29 | 0. 35 |
| 1017      | -13. 51 | 0. 76 | 0. 39 | 0. 00 | 0. 00 | 0. 28 |
| 1018      | -35. 15 | 0. 19 | 0. 38 | 0. 00 | 0. 00 | 0. 35 |
| 1019      | -16. 25 | 0. 65 | 0. 38 | 0. 22 | 0. 00 | 0. 27 |

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1020 | -13.96 | 0.54 | 0.38 | 0.00 | 0.00 | 0.83 |
|------|--------|------|------|------|------|------|

[0351]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1021 | -15.75 | 0.75 | 0.37 | 0.22 | 0.22 | 0.27 |
| 1022 | -9.84  | 1.10 | 0.36 | 0.00 | 0.25 | 0.26 |
| 1023 | -9.84  | 1.10 | 0.36 | 0.00 | 0.25 | 0.26 |
| 1024 | -9.79  | 0.97 | 0.36 | 0.27 | 0.24 | 0.29 |
| 1025 | -14.86 | 0.77 | 0.36 | 0.21 | 0.00 | 0.27 |
| 1026 | -20.01 | 0.60 | 0.36 | 0.00 | 0.00 | 0.20 |
| 1027 | -11.76 | 0.58 | 0.36 | 0.00 | 0.00 | 0.85 |
| 1028 | -11.75 | 0.58 | 0.36 | 0.00 | 0.00 | 0.85 |
| 1029 | -14.00 | 0.90 | 0.36 | 0.00 | 0.25 | 0.26 |
| 1030 | -19.70 | 0.84 | 0.35 | 0.00 | 0.21 | 0.00 |
| 1031 | -19.59 | 0.71 | 0.35 | 0.00 | 0.17 | 0.24 |
| 1032 | -24.74 | 0.63 | 0.28 | 0.67 | 0.29 | 0.00 |
| 1033 | -20.05 | 0.57 | 0.24 | 0.40 | 0.28 | 0.30 |
| 1034 | -4.01  | 0.67 | 0.31 | 0.40 | 0.38 | 0.43 |
| 1035 | -10.91 | 1.01 | 0.29 | 0.39 | 0.34 | 0.42 |
| 1036 | -30.64 | 0.41 | 0.27 | 0.00 | 0.30 | 0.00 |
| 1037 | -27.88 | 0.25 | 0.23 | 0.00 | 0.00 | 0.31 |
| 1038 | -27.52 | 0.30 | 0.27 | 0.00 | 0.27 | 0.64 |
| 1039 | -27.42 | 0.47 | 0.29 | 0.00 | 0.00 | 0.00 |
| 1040 | -27.42 | 0.29 | 0.32 | 0.00 | 0.27 | 0.00 |
| 1041 | -26.41 | 0.31 | 0.26 | 0.00 | 0.29 | 0.53 |
| 1042 | -26.11 | 0.38 | 0.28 | 0.22 | 0.28 | 0.66 |

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1043 | -25.86 | 0.56 | 0.25 | 0.00 | 0.26 | 0.79 |
| 1044 | -25.75 | 0.31 | 0.32 | 0.00 | 0.28 | 0.58 |
| 1045 | -25.74 | 0.51 | 0.27 | 0.29 | 0.30 | 0.31 |
| 1046 | -25.62 | 0.48 | 0.26 | 0.30 | 0.21 | 0.66 |
| 1047 | -25.19 | 0.41 | 0.27 | 0.32 | 0.29 | 0.26 |
| 1048 | -25.11 | 0.72 | 0.28 | 0.00 | 0.28 | 0.36 |
| 1049 | -24.99 | 0.51 | 0.21 | 0.27 | 0.27 | 0.34 |
| 1050 | -24.92 | 0.42 | 0.28 | 0.00 | 0.00 | 0.00 |
| 1051 | -24.77 | 0.34 | 0.30 | 0.26 | 0.24 | 0.50 |
| 1052 | -24.70 | 0.55 | 0.26 | 0.00 | 0.22 | 0.32 |
| 1053 | -24.58 | 0.52 | 0.25 | 0.00 | 0.00 | 0.00 |
| 1054 | -24.51 | 0.39 | 0.33 | 0.00 | 0.00 | 0.52 |
| 1055 | -24.44 | 0.27 | 0.31 | 0.26 | 0.00 | 0.46 |

[0352]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1056 | -24.32 | 0.51 | 0.27 | 0.00 | 0.00 | 0.56 |
| 1057 | -24.32 | 0.57 | 0.25 | 0.00 | 0.31 | 0.00 |
| 1058 | -24.14 | 0.55 | 0.31 | 0.00 | 0.28 | 0.48 |
| 1059 | -24.14 | 0.48 | 0.29 | 0.00 | 0.30 | 0.41 |
| 1060 | -23.96 | 0.56 | 0.34 | 0.00 | 0.26 | 0.30 |
| 1061 | -23.86 | 0.55 | 0.26 | 0.00 | 0.26 | 0.58 |
| 1062 | -23.81 | 0.67 | 0.23 | 0.00 | 0.24 | 0.00 |
| 1063 | -23.35 | 0.63 | 0.23 | 0.26 | 0.26 | 0.28 |
| 1064 | -23.30 | 0.56 | 0.28 | 0.00 | 0.00 | 0.56 |
| 1065 | -23.28 | 0.63 | 0.33 | 0.26 | 0.30 | 0.27 |

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1066 | -23.25 | 0.56 | 0.29 | 0.00 | 0.27 | 0.55 |
| 1067 | -23.24 | 0.32 | 0.28 | 0.29 | 0.28 | 0.50 |
| 1068 | -23.20 | 0.71 | 0.29 | 0.33 | 0.25 | 0.27 |
| 1069 | -22.91 | 0.44 | 0.27 | 0.00 | 0.30 | 0.53 |
| 1070 | -22.86 | 0.47 | 0.28 | 0.31 | 0.25 | 0.34 |
| 1071 | -22.79 | 0.51 | 0.29 | 0.25 | 0.25 | 0.25 |
| 1072 | -22.76 | 0.62 | 0.25 | 0.00 | 0.00 | 0.00 |
| 1073 | -22.76 | 0.47 | 0.30 | 0.00 | 0.00 | 0.31 |
| 1074 | -22.72 | 0.56 | 0.27 | 0.00 | 0.21 | 0.31 |
| 1075 | -22.70 | 0.75 | 0.30 | 0.00 | 0.00 | 0.51 |

[0353] 实施例 3 :含有的必需氨基酸和支链氨基酸的比率大于或等于乳清以及含有所有必需氨基酸的蛋白质片段的鉴定

[0354] 评估来自实施例 1 中所述的所述组的 8,415 种蛋白质的片段的氨基酸含量、必需氨基酸 (“EAA”) 的百分比、支链氨基酸 (“BCAA”) 的百分比、亮氨酸 (“L”) 的百分比,且对于各蛋白质计算蛋白质是否含有所有必需氨基酸。此外,相对于已知过敏原的数据库筛选蛋白质片段以确定任何蛋白质片段是否与已知过敏原具有大于 50% 整体同源性。鉴定出总计 485 种蛋白质含有大于或等于 51% EAA 和大于或等于 25% BCAA,且含有所有必需氨基酸,并与已知过敏原具有小于 50% 整体同源性 (SEQ ID NO:1415 至 1899)。对于所述组的蛋白质片段,计算在 pH 7 下的溶合计分 (“SolvScore”)、在 pH 7 下的聚集计分 (“AggScore”)、过敏原性 (即,如本文所述的与已知过敏原的局部同源性百分比)、毒性 (即,如本文所述的与已知毒素的同源性百分比)、抗营养性 (即,如本文所述的与已知蛋白酶抑制剂的同源性百分比)、和人同源性 (即,如本文所述的与已知人蛋白质的同源性百分比),且测定 Cys 残基 (“C”) 的总数。由此鉴定的 75 种代表性蛋白质的特征呈现于表 5A 和 5B 中。

[0355] 表 5A

[0356]

| Seq ID No | UniProt | FragEnds  | EAA  | BCAA | L    | C |
|-----------|---------|-----------|------|------|------|---|
| 1415      | Q96576  | 151 :326  | 0.52 | 0.27 | 0.10 | 2 |
| 1416      | P15722  | 176 :276  | 0.53 | 0.28 | 0.04 | 1 |
| 1417      | Q8GUQ5  | 876 :1001 | 0.51 | 0.25 | 0.12 | 1 |

|      |        |           |       |       |       |   |
|------|--------|-----------|-------|-------|-------|---|
| 1418 | P17801 | 476 :751  | 0. 52 | 0. 26 | 0. 12 | 3 |
| 1419 | Q6ERW7 | 201 :351  | 0. 53 | 0. 26 | 0. 10 | 1 |
| 1420 | Q7XWU3 | 51 :276   | 0. 51 | 0. 25 | 0. 09 | 6 |
| 1421 | Q6ER94 | 76 :226   | 0. 52 | 0. 26 | 0. 10 | 1 |
| 1422 | Q43272 | 176 :376  | 0. 52 | 0. 26 | 0. 10 | 6 |
| 1423 | Q43272 | 76 :326   | 0. 52 | 0. 28 | 0. 11 | 7 |
| 1424 | Q6ERW9 | 226 :351  | 0. 55 | 0. 26 | 0. 11 | 1 |
| 1425 | Q10717 | 251 :326  | 0. 51 | 0. 30 | 0. 10 | 1 |
| 1426 | Q6ERW9 | 126 :276  | 0. 51 | 0. 25 | 0. 06 | 5 |
| 1427 | Q6ERW5 | 26 :101   | 0. 52 | 0. 25 | 0. 03 | 4 |
| 1428 | P16243 | 226 :476  | 0. 51 | 0. 27 | 0. 12 | 4 |
| 1429 | Q43272 | 76 :226   | 0. 51 | 0. 28 | 0. 12 | 4 |
| 1430 | 004161 | 151 :376  | 0. 55 | 0. 25 | 0. 11 | 5 |
| 1431 | Q9ATN1 | 101 :251  | 0. 56 | 0. 29 | 0. 11 | 2 |
| 1432 | Q9ATN1 | 1 :251    | 0. 51 | 0. 25 | 0. 10 | 5 |
| 1433 | P29185 | 26 :301   | 0. 52 | 0. 25 | 0. 08 | 3 |
| 1434 | Q08451 | 126 :276  | 0. 53 | 0. 27 | 0. 08 | 2 |
| 1435 | Q96575 | 1 :276    | 0. 52 | 0. 26 | 0. 12 | 4 |
| 1436 | P09607 | 1 :276    | 0. 51 | 0. 26 | 0. 12 | 5 |
| 1437 | Q6ETN3 | 401 :501  | 0. 52 | 0. 25 | 0. 06 | 0 |
| 1438 | Q5BQN5 | 151 :226  | 0. 60 | 0. 26 | 0. 11 | 1 |
| 1439 | Q2MI42 | 976 :1176 | 0. 64 | 0. 25 | 0. 10 | 0 |
| 1440 | P28523 | 26 :176   | 0. 51 | 0. 26 | 0. 09 | 3 |

[0357]

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1441 | Q8H6B1 | 126 :226 | 0. 57 | 0. 26 | 0. 10 | 1 |
| 1442 | P54767 | 176 :251 | 0. 54 | 0. 26 | 0. 11 | 1 |
| 1443 | Q67W82 | 401 :526 | 0. 51 | 0. 26 | 0. 06 | 0 |
| 1444 | POC8M8 | 1 :226   | 0. 53 | 0. 25 | 0. 11 | 7 |
| 1445 | Q9FYT6 | 551 :651 | 0. 53 | 0. 25 | 0. 09 | 0 |
| 1446 | POC8M8 | 151 :226 | 0. 52 | 0. 26 | 0. 12 | 2 |
| 1447 | Q10S72 | 426 :551 | 0. 52 | 0. 26 | 0. 10 | 1 |
| 1448 | Q42982 | 426 :551 | 0. 53 | 0. 26 | 0. 06 | 0 |
| 1449 | Q2MI46 | 26 :101  | 0. 51 | 0. 25 | 0. 09 | 4 |
| 1450 | B4G0F3 | 276 :401 | 0. 52 | 0. 26 | 0. 07 | 1 |
| 1451 | P93648 | 76 :251  | 0. 52 | 0. 26 | 0. 12 | 1 |
| 1452 | C1K5M3 | 551 :751 | 0. 55 | 0. 25 | 0. 12 | 5 |
| 1453 | Q41741 | 151 :276 | 0. 53 | 0. 27 | 0. 12 | 1 |
| 1454 | O22424 | 151 :251 | 0. 60 | 0. 25 | 0. 05 | 0 |
| 1455 | Q9XFH3 | 51 :201  | 0. 54 | 0. 26 | 0. 10 | 3 |
| 1456 | Q7XTS4 | 201 :451 | 0. 51 | 0. 25 | 0. 11 | 3 |
| 1457 | Q9FUZO | 176 :276 | 0. 52 | 0. 25 | 0. 12 | 2 |
| 1458 | P05118 | 1 :76    | 0. 59 | 0. 26 | 0. 10 | 1 |
| 1459 | Q8S7E1 | 51 :126  | 0. 52 | 0. 27 | 0. 12 | 1 |
| 1460 | C1K5M2 | 101 :201 | 0. 55 | 0. 25 | 0. 10 | 3 |
| 1461 | Q6Z4U2 | 151 :351 | 0. 52 | 0. 26 | 0. 12 | 4 |
| 1462 | Q8H6B1 | 26 :176  | 0. 56 | 0. 25 | 0. 11 | 0 |
| 1463 | Q0J7J7 | 176 :351 | 0. 51 | 0. 27 | 0. 11 | 5 |
| 1464 | P10581 | 651 :751 | 0. 51 | 0. 27 | 0. 12 | 1 |

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1465 | Q2RAS0 | 26 :251  | 0. 51 | 0. 25 | 0. 09 | 2 |
| 1466 | 024594 | 351 :426 | 0. 53 | 0. 29 | 0. 08 | 2 |
| 1467 | Q0J7J7 | 276 :351 | 0. 54 | 0. 31 | 0. 12 | 3 |
| 1468 | Q6XZ79 | 176 :276 | 0. 51 | 0. 25 | 0. 10 | 1 |
| 1469 | P19862 | 651 :801 | 0. 53 | 0. 26 | 0. 11 | 3 |
| 1470 | P28554 | 276 :426 | 0. 53 | 0. 26 | 0. 11 | 2 |
| 1471 | C1K5M3 | 451 :651 | 0. 52 | 0. 25 | 0. 12 | 4 |
| 1472 | P31927 | 451 :551 | 0. 51 | 0. 26 | 0. 12 | 1 |
| 1473 | Q65X71 | 451 :526 | 0. 61 | 0. 31 | 0. 12 | 0 |
| 1474 | Q00497 | 76 :226  | 0. 51 | 0. 25 | 0. 11 | 4 |
| 1475 | Q7XKC0 | 76 :176  | 0. 51 | 0. 26 | 0. 06 | 3 |

[0358]

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1476 | P10582 | 276 :451   | 0. 55 | 0. 27 | 0. 12 | 3 |
| 1477 | P29535 | 176 :401   | 0. 51 | 0. 26 | 0. 11 | 8 |
| 1478 | Q43298 | 26 :301    | 0. 51 | 0. 25 | 0. 08 | 3 |
| 1479 | Q7XRA1 | 176 :251   | 0. 61 | 0. 29 | 0. 12 | 0 |
| 1480 | Q852N2 | 826 :1001  | 0. 52 | 0. 27 | 0. 12 | 6 |
| 1481 | Q65X71 | 451 :726   | 0. 52 | 0. 26 | 0. 09 | 5 |
| 1482 | Q852N2 | 826 :901   | 0. 51 | 0. 25 | 0. 10 | 2 |
| 1483 | Q65X71 | 51 :326    | 0. 52 | 0. 27 | 0. 10 | 3 |
| 1484 | Q8H6B1 | 801 :901   | 0. 57 | 0. 29 | 0. 12 | 1 |
| 1485 | P54767 | 176 :451   | 0. 52 | 0. 26 | 0. 10 | 2 |
| 1486 | P61242 | 1626 :1876 | 0. 51 | 0. 27 | 0. 12 | 4 |
| 1487 | Q65X71 | 451 :626   | 0. 54 | 0. 26 | 0. 09 | 3 |



|      |        |          |      |      |      |   |
|------|--------|----------|------|------|------|---|
| 1488 | P04712 | 326 :426 | 0.52 | 0.27 | 0.11 | 1 |
| 1489 | Q75LR2 | 351 :426 | 0.55 | 0.28 | 0.12 | 0 |

[0359] 表 5B

[0360]

| Seq ID No | 溶合计分   | 聚集计分 | 过敏原性 | 毒性   | 抗营养性 | 人同源性 |
|-----------|--------|------|------|------|------|------|
| 1415      | -16.93 | 0.58 | 0.54 | 0.53 | 0.24 | 0.00 |
| 1416      | -19.58 | 0.41 | 0.53 | 0.00 | 0.21 | 0.00 |
| 1417      | -21.91 | 0.46 | 0.47 | 0.27 | 0.27 | 0.37 |
| 1418      | -19.73 | 0.54 | 0.46 | 0.00 | 0.00 | 0.34 |
| 1419      | -11.15 | 0.74 | 0.43 | 0.00 | 0.25 | 0.29 |
| 1420      | -16.57 | 0.61 | 0.42 | 0.22 | 0.00 | 0.27 |
| 1421      | -20.09 | 0.54 | 0.41 | 0.23 | 0.00 | 0.63 |
| 1422      | -16.36 | 0.68 | 0.41 | 0.00 | 0.24 | 0.34 |
| 1423      | -16.55 | 0.67 | 0.40 | 0.00 | 0.21 | 0.32 |
| 1424      | -11.31 | 0.79 | 0.40 | 0.24 | 0.00 | 0.27 |
| 1425      | -17.05 | 0.64 | 0.39 | 0.28 | 0.35 | 0.53 |
| 1426      | -11.49 | 0.73 | 0.38 | 0.00 | 0.26 | 0.35 |
| 1427      | -17.25 | 0.63 | 0.38 | 0.25 | 0.28 | 0.41 |
| 1428      | -19.07 | 0.59 | 0.38 | 0.21 | 0.00 | 0.60 |
| 1429      | -16.46 | 0.67 | 0.36 | 0.25 | 0.23 | 0.33 |
| 1430      | -11.00 | 0.89 | 0.36 | 0.00 | 0.00 | 0.26 |
| 1431      | -6.06  | 1.08 | 0.36 | 0.27 | 0.27 | 0.36 |

[0361]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1432 | -8.76  | 0.87 | 0.36 | 0.23 | 0.23 | 0.28 |
| 1433 | -21.28 | 0.49 | 0.36 | 0.00 | 0.00 | 0.60 |

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1434 | -10.55 | 0.95 | 0.35 | 0.25 | 0.24 | 0.44 |
| 1435 | -17.94 | 0.61 | 0.34 | 0.37 | 0.00 | 0.00 |
| 1436 | -18.54 | 0.63 | 0.33 | 0.36 | 0.19 | 0.00 |
| 1437 | -28.54 | 0.46 | 0.30 | 0.00 | 0.23 | 0.38 |
| 1438 | -25.54 | 0.40 | 0.27 | 0.00 | 0.00 | 0.26 |
| 1439 | -25.12 | 0.37 | 0.33 | 0.24 | 0.00 | 0.19 |
| 1440 | -25.09 | 0.34 | 0.31 | 0.00 | 0.21 | 0.80 |
| 1441 | -24.93 | 0.38 | 0.32 | 0.00 | 0.25 | 0.27 |
| 1442 | -24.55 | 0.45 | 0.22 | 0.00 | 0.00 | 0.32 |
| 1443 | -24.41 | 0.49 | 0.29 | 0.24 | 0.00 | 0.33 |
| 1444 | -24.26 | 0.36 | 0.31 | 0.00 | 0.00 | 0.49 |
| 1445 | -23.95 | 0.53 | 0.28 | 0.00 | 0.00 | 0.00 |
| 1446 | -23.76 | 0.45 | 0.30 | 0.00 | 0.26 | 0.70 |
| 1447 | -23.69 | 0.45 | 0.30 | 0.24 | 0.27 | 0.39 |
| 1448 | -23.58 | 0.48 | 0.31 | 0.00 | 0.24 | 0.36 |
| 1449 | -23.51 | 0.56 | 0.27 | 0.25 | 0.27 | 0.36 |
| 1450 | -23.40 | 0.48 | 0.31 | 0.25 | 0.00 | 0.43 |
| 1451 | -23.21 | 0.41 | 0.30 | 0.00 | 0.22 | 0.27 |
| 1452 | -23.11 | 0.50 | 0.30 | 0.00 | 0.23 | 0.00 |
| 1453 | -22.98 | 0.53 | 0.30 | 0.29 | 0.26 | 0.71 |
| 1454 | -22.62 | 0.40 | 0.27 | 0.00 | 0.00 | 0.70 |
| 1455 | -22.54 | 0.48 | 0.31 | 0.24 | 0.24 | 0.42 |
| 1456 | -22.52 | 0.37 | 0.33 | 0.24 | 0.00 | 0.30 |
| 1457 | -22.50 | 0.44 | 0.30 | 0.00 | 0.00 | 0.44 |

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1458 | -22.49 | 0.61 | 0.27 | 0.24 | 1.00 | 0.26 |
| 1459 | -22.48 | 0.47 | 0.23 | 0.00 | 0.27 | 0.24 |
| 1460 | -22.30 | 0.61 | 0.30 | 0.23 | 0.19 | 0.36 |
| 1461 | -22.29 | 0.45 | 0.30 | 0.21 | 0.24 | 0.00 |
| 1462 | -22.25 | 0.48 | 0.31 | 0.00 | 0.21 | 0.29 |
| 1463 | -22.20 | 0.44 | 0.30 | 0.00 | 0.22 | 0.00 |
| 1464 | -22.05 | 0.53 | 0.28 | 0.00 | 0.00 | 0.00 |
| 1465 | -22.05 | 0.69 | 0.30 | 0.00 | 0.00 | 0.29 |
| 1466 | -22.02 | 0.45 | 0.29 | 0.00 | 0.32 | 0.54 |

[0362]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1467 | -22.00 | 0.48 | 0.27 | 0.00 | 0.30 | 0.32 |
| 1468 | -21.97 | 0.50 | 0.29 | 0.00 | 0.00 | 0.31 |
| 1469 | -21.83 | 0.54 | 0.30 | 0.23 | 0.00 | 0.00 |
| 1470 | -21.80 | 0.58 | 0.31 | 0.25 | 0.00 | 0.23 |
| 1471 | -21.78 | 0.56 | 0.30 | 0.21 | 0.23 | 0.00 |
| 1472 | -21.60 | 0.54 | 0.30 | 0.00 | 0.00 | 0.00 |
| 1473 | -21.53 | 0.69 | 0.24 | 0.00 | 0.26 | 0.33 |
| 1474 | -21.44 | 0.48 | 0.33 | 0.21 | 0.00 | 0.27 |
| 1475 | -21.39 | 0.65 | 0.28 | 0.26 | 0.00 | 0.59 |
| 1476 | -21.27 | 0.57 | 0.30 | 0.00 | 0.24 | 0.00 |
| 1477 | -21.26 | 0.50 | 0.33 | 0.00 | 0.23 | 0.30 |
| 1478 | -21.26 | 0.49 | 0.32 | 0.00 | 0.00 | 0.60 |
| 1479 | -21.13 | 0.55 | 0.28 | 0.00 | 0.00 | 0.28 |
| 1480 | -21.08 | 0.61 | 0.33 | 0.17 | 0.24 | 0.27 |

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1481 | -21.00 | 0.50 | 0.31 | 0.19 | 0.22 | 0.39 |
| 1482 | -20.97 | 0.43 | 0.24 | 0.17 | 0.26 | 0.00 |
| 1483 | -20.94 | 0.68 | 0.33 | 0.00 | 0.23 | 0.34 |
| 1484 | -20.89 | 0.61 | 0.29 | 0.00 | 0.26 | 0.51 |
| 1485 | -20.84 | 0.48 | 0.31 | 0.00 | 0.00 | 0.23 |
| 1486 | -20.74 | 0.51 | 0.30 | 0.00 | 0.22 | 0.22 |
| 1487 | -20.65 | 0.51 | 0.31 | 0.28 | 0.27 | 0.34 |
| 1488 | -20.58 | 0.44 | 0.32 | 0.00 | 0.00 | 0.00 |
| 1489 | -20.54 | 0.44 | 0.28 | 0.00 | 0.00 | 0.21 |

[0363] 实施例 4 :含有的必需氨基酸和亮氨酸的比率大于或等于乳清以及含有所有必需氨基酸的蛋白质片段的鉴定。

[0364] 评估来自实施例 1 中所述的所述组的 8,415 种蛋白质的片段的氨基酸含量、必需氨基酸 (“EAA”) 的百分比、支链氨基酸 (“BCAA”) 的百分比、亮氨酸 (“L”) 的百分比,且对于各蛋白质计算蛋白质是否含有所有必需氨基酸。此外,相对于已知过敏原的数据库筛选蛋白质片段以确定任何蛋白质片段是否与已知过敏原具有大于 50% 整体同源性。鉴定出总计 203 个蛋白质片段含有大于或等于 51% EAA 和大于或等于 13% Leu,且含有所有必需氨基酸,并与已知过敏原具有小于 50% 整体同源性 (SEQ ID NO:1900 至 2102)。对于所述组的蛋白质片段,计算在 pH 7 下的溶合计分 (“SolvScore”)、在 pH 7 下的聚集计分 (“AggScore”)、过敏原性 (即,如本文所述的与已知过敏原的局部同源性百分比)、毒性 (即,如本文所述的与已知毒素的同源性百分比)、抗营养性 (即,如本文所述的与已知蛋白酶抑制剂的同源性百分比)、和人同源性 (即,如本文所述的与已知人蛋白质的同源性百分比),且测定 Cys 残基 (“C”) 的总数。由此鉴定的 75 个代表性蛋白质片的特征呈现于表 6A 和 6B 中。

[0365] 表 6A

[0366]

| Seq ID No | UniProt | FragEnds | EAA  | BCAA | L    | C |
|-----------|---------|----------|------|------|------|---|
| 1900      | Q41784  | 201 ;301 | 0.51 | 0.22 | 0.15 | 4 |
| 1901      | P18025  | 201 ;301 | 0.51 | 0.22 | 0.15 | 4 |
| 1902      | Q43697  | 201 ;301 | 0.51 | 0.22 | 0.15 | 4 |
| 1903      | Q41783  | 201 ;301 | 0.51 | 0.22 | 0.15 | 4 |

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1904 | Q41785 | 201 :301   | 0. 51 | 0. 22 | 0. 15 | 4 |
| 1905 | P18026 | 201 :301   | 0. 51 | 0. 22 | 0. 15 | 4 |
| 1906 | P10708 | 76 :176    | 0. 51 | 0. 24 | 0. 13 | 1 |
| 1907 | P27525 | 51 :151    | 0. 53 | 0. 22 | 0. 14 | 0 |
| 1908 | Q3HRN9 | 51 :201    | 0. 52 | 0. 23 | 0. 14 | 1 |
| 1909 | Q9ZSV1 | 726 :801   | 0. 51 | 0. 23 | 0. 14 | 1 |
| 1910 | Q5BQN5 | 51 :176    | 0. 52 | 0. 20 | 0. 13 | 1 |
| 1911 | P56669 | 1 :76      | 0. 53 | 0. 18 | 0. 14 | 1 |
| 1912 | Q75LU8 | 1 :201     | 0. 52 | 0. 24 | 0. 13 | 4 |
| 1913 | Q8S0F0 | 601 :851   | 0. 51 | 0. 23 | 0. 14 | 3 |
| 1914 | Q9ZSV1 | 626 :801   | 0. 51 | 0. 23 | 0. 13 | 1 |
| 1915 | Q0J149 | 301 :451   | 0. 53 | 0. 23 | 0. 14 | 0 |
| 1916 | P93703 | 201 :401   | 0. 51 | 0. 24 | 0. 14 | 1 |
| 1917 | Q2M142 | 426 :601   | 0. 52 | 0. 23 | 0. 13 | 1 |
| 1918 | Q8S0F0 | 501 :751   | 0. 52 | 0. 23 | 0. 13 | 4 |
| 1919 | Q2M142 | 1601 :1701 | 0. 51 | 0. 23 | 0. 14 | 1 |
| 1920 | Q7XTS3 | 401 :476   | 0. 57 | 0. 21 | 0. 13 | 2 |

[0367]

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1921 | P49037 | 1 :201     | 0. 51 | 0. 24 | 0. 14 | 1 |
| 1922 | Q9FQ11 | 1 :76      | 0. 53 | 0. 24 | 0. 16 | 1 |
| 1923 | Q7X996 | 201 :376   | 0. 55 | 0. 22 | 0. 13 | 1 |
| 1924 | P27061 | 126 :201   | 0. 53 | 0. 24 | 0. 14 | 0 |
| 1925 | P53682 | 326 :401   | 0. 51 | 0. 24 | 0. 13 | 1 |
| 1926 | Q84ZL0 | 1201 :1476 | 0. 51 | 0. 24 | 0. 16 | 4 |

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1927 | Q8S0F0 | 701 :851   | 0. 51 | 0. 24 | 0. 14 | 1 |
| 1928 | Q6ZAG3 | 51 :201    | 0. 54 | 0. 24 | 0. 13 | 1 |
| 1929 | Q69Q02 | 276 :451   | 0. 52 | 0. 22 | 0. 13 | 1 |
| 1930 | P10581 | 276 :526   | 0. 51 | 0. 24 | 0. 13 | 2 |
| 1931 | P49037 | 201 :276   | 0. 52 | 0. 24 | 0. 15 | 2 |
| 1932 | P61242 | 776 :951   | 0. 51 | 0. 23 | 0. 15 | 2 |
| 1933 | Q5Z987 | 1826 :1926 | 0. 51 | 0. 19 | 0. 14 | 1 |
| 1934 | Q5JMM1 | 251 :326   | 0. 52 | 0. 22 | 0. 14 | 0 |
| 1935 | Q84N48 | 426 :526   | 0. 51 | 0. 24 | 0. 14 | 0 |
| 1936 | Q7G6K7 | 826 :1101  | 0. 51 | 0. 23 | 0. 14 | 4 |
| 1937 | Q0J4I1 | 1 :201     | 0. 53 | 0. 24 | 0. 14 | 2 |
| 1938 | Q7XWS7 | 1226 :1476 | 0. 51 | 0. 22 | 0. 13 | 3 |
| 1939 | Q0JF02 | 76 :276    | 0. 51 | 0. 24 | 0. 14 | 4 |
| 1940 | Q2QN26 | 201 :351   | 0. 51 | 0. 24 | 0. 14 | 5 |
| 1941 | P49037 | 101 :276   | 0. 53 | 0. 22 | 0. 14 | 2 |
| 1942 | Q6K5F8 | 351 :626   | 0. 51 | 0. 24 | 0. 14 | 3 |
| 1943 | Q10QA2 | 101 :226   | 0. 51 | 0. 22 | 0. 13 | 6 |
| 1944 | Q9XHR2 | 201 :301   | 0. 52 | 0. 24 | 0. 14 | 0 |
| 1945 | Q0DCT8 | 176 :351   | 0. 52 | 0. 24 | 0. 15 | 5 |
| 1946 | Q6AU53 | 476 :576   | 0. 52 | 0. 24 | 0. 16 | 0 |
| 1947 | A7LFZ6 | 276 :451   | 0. 52 | 0. 23 | 0. 15 | 6 |
| 1948 | Q6ZKB2 | 451 :726   | 0. 52 | 0. 22 | 0. 13 | 2 |
| 1949 | Q8S983 | 176 :251   | 0. 51 | 0. 22 | 0. 14 | 0 |
| 1950 | Q3LXA7 | 526 :626   | 0. 53 | 0. 24 | 0. 14 | 2 |

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1951 | P93703 | 201 :276 | 0. 55 | 0. 24 | 0. 14 | 1 |
| 1952 | A3B9A0 | 76 :226  | 0. 51 | 0. 21 | 0. 13 | 2 |
| 1953 | Q9XGD6 | 51 :126  | 0. 52 | 0. 22 | 0. 13 | 0 |
| 1954 | P18485 | 301 :376 | 0. 54 | 0. 20 | 0. 14 | 2 |
| 1955 | Q32SG6 | 876 :951 | 0. 52 | 0. 23 | 0. 16 | 0 |

[0368]

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 1956 | Q5VS72 | 501 :576   | 0. 54 | 0. 24 | 0. 18 | 2 |
| 1957 | Q43503 | 401 :476   | 0. 54 | 0. 21 | 0. 15 | 1 |
| 1958 | Q42896 | 101 :326   | 0. 52 | 0. 24 | 0. 13 | 4 |
| 1959 | Q9MAX6 | 151 :251   | 0. 52 | 0. 24 | 0. 14 | 3 |
| 1960 | Q9LXX9 | 1 :201     | 0. 52 | 0. 22 | 0. 14 | 5 |
| 1961 | P09562 | 176 :301   | 0. 52 | 0. 23 | 0. 15 | 3 |
| 1962 | Q7XWS7 | 1226 :1376 | 0. 51 | 0. 22 | 0. 13 | 1 |
| 1963 | Q7XU38 | 151 :351   | 0. 51 | 0. 24 | 0. 14 | 4 |
| 1964 | Q69LX2 | 926 :1151  | 0. 51 | 0. 24 | 0. 13 | 5 |
| 1965 | Q7XD96 | 601 :726   | 0. 51 | 0. 22 | 0. 14 | 5 |
| 1966 | Q7XI73 | 26 :101    | 0. 52 | 0. 17 | 0. 13 | 5 |
| 1967 | Q42881 | 301 :376   | 0. 53 | 0. 20 | 0. 14 | 3 |
| 1968 | C1K5M3 | 176 :301   | 0. 53 | 0. 23 | 0. 14 | 4 |
| 1969 | Q6L4F8 | 226 :301   | 0. 53 | 0. 22 | 0. 15 | 2 |
| 1970 | P48022 | 1 :101     | 0. 51 | 0. 23 | 0. 14 | 0 |
| 1971 | Q8LHA8 | 51 :201    | 0. 51 | 0. 24 | 0. 14 | 3 |
| 1972 | Q6ZAG3 | 151 :301   | 0. 52 | 0. 24 | 0. 14 | 3 |
| 1973 | P04711 | 751 :901   | 0. 53 | 0. 24 | 0. 13 | 0 |

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 1974 | Q5N870 | 576 ;801 | 0. 53 | 0. 23 | 0. 13 | 8 |
|------|--------|----------|-------|-------|-------|---|

[0369] 表 6B

[0370]

| Seq ID No | 溶合计分   | 聚集计分 | 过敏原性 | 毒性   | 抗营养性 | 人同源性 |
|-----------|--------|------|------|------|------|------|
| 1900      | -14.31 | 0.56 | 0.48 | 0.00 | 0.25 | 0.88 |
| 1901      | -14.31 | 0.54 | 0.46 | 0.00 | 0.25 | 0.88 |
| 1902      | -14.31 | 0.54 | 0.46 | 0.00 | 0.25 | 0.88 |
| 1903      | -14.31 | 0.54 | 0.46 | 0.00 | 0.25 | 0.88 |
| 1904      | -14.31 | 0.54 | 0.46 | 0.00 | 0.25 | 0.88 |
| 1905      | -14.42 | 0.53 | 0.46 | 0.00 | 0.24 | 0.87 |
| 1906      | -16.21 | 0.63 | 0.41 | 0.00 | 0.00 | 0.28 |
| 1907      | -14.10 | 0.61 | 0.39 | 0.00 | 0.25 | 0.00 |
| 1908      | -24.00 | 0.41 | 0.36 | 0.22 | 0.26 | 0.31 |
| 1909      | -27.98 | 0.27 | 0.25 | 0.29 | 0.33 | 0.49 |
| 1910      | -27.07 | 0.17 | 0.33 | 0.23 | 0.24 | 0.00 |
| 1911      | -27.03 | 0.25 | 0.22 | 0.00 | 0.27 | 0.41 |

[0371]



|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1912 | -25.61 | 0.41 | 0.35 | 0.23 | 0.25 | 0.32 |
| 1913 | -24.50 | 0.36 | 0.33 | 0.00 | 0.00 | 0.30 |
| 1914 | -24.43 | 0.35 | 0.35 | 0.00 | 0.23 | 0.39 |
| 1915 | -24.41 | 0.37 | 0.32 | 0.24 | 0.23 | 0.00 |
| 1916 | -24.40 | 0.42 | 0.32 | 0.00 | 0.24 | 0.30 |
| 1917 | -24.34 | 0.23 | 0.33 | 0.00 | 0.23 | 0.00 |
| 1918 | -23.77 | 0.39 | 0.32 | 0.00 | 0.00 | 0.34 |
| 1919 | -23.65 | 0.44 | 0.31 | 0.00 | 0.22 | 0.00 |
| 1920 | -23.49 | 0.37 | 0.28 | 0.00 | 0.00 | 0.33 |
| 1921 | -23.32 | 0.40 | 0.33 | 0.22 | 0.25 | 0.00 |
| 1922 | -23.31 | 0.48 | 0.25 | 0.26 | 0.24 | 0.00 |
| 1923 | -23.23 | 0.39 | 0.30 | 0.00 | 0.21 | 0.27 |
| 1924 | -23.18 | 0.42 | 0.24 | 0.00 | 0.00 | 0.00 |
| 1925 | -23.13 | 0.29 | 0.28 | 0.25 | 0.26 | 0.27 |
| 1926 | -23.12 | 0.37 | 0.34 | 0.00 | 0.00 | 0.32 |
| 1927 | -23.00 | 0.35 | 0.34 | 0.00 | 0.25 | 0.29 |
| 1928 | -22.83 | 0.31 | 0.31 | 0.00 | 0.22 | 0.49 |
| 1929 | -22.74 | 0.36 | 0.29 | 0.00 | 0.23 | 0.00 |
| 1930 | -22.67 | 0.39 | 0.31 | 0.00 | 0.00 | 0.00 |
| 1931 | -22.62 | 0.50 | 0.26 | 0.25 | 0.27 | 0.26 |
| 1932 | -22.44 | 0.44 | 0.33 | 0.24 | 0.22 | 0.00 |
| 1933 | -22.39 | 0.41 | 0.31 | 0.22 | 0.25 | 0.34 |
| 1934 | -22.25 | 0.36 | 0.22 | 0.00 | 0.33 | 0.33 |
| 1935 | -22.24 | 0.40 | 0.30 | 0.00 | 0.25 | 0.00 |
| 1936 | -22.14 | 0.35 | 0.32 | 0.00 | 0.00 | 0.31 |
| 1937 | -22.12 | 0.41 | 0.34 | 0.00 | 0.00 | 0.48 |
| 1938 | -21.99 | 0.35 | 0.31 | 0.00 | 0.00 | 0.29 |
| 1939 | -21.98 | 0.49 | 0.33 | 0.00 | 0.23 | 0.00 |
| 1940 | -21.79 | 0.55 | 0.32 | 0.26 | 0.24 | 0.36 |
| 1941 | -21.63 | 0.40 | 0.29 | 0.00 | 0.21 | 0.00 |
| 1942 | -21.60 | 0.42 | 0.33 | 0.00 | 0.24 | 0.65 |
| 1943 | -21.54 | 0.62 | 0.30 | 0.25 | 0.24 | 0.33 |
| 1944 | -21.23 | 0.37 | 0.28 | 0.00 | 0.00 | 0.42 |
| 1945 | -21.20 | 0.48 | 0.32 | 0.00 | 0.25 | 0.40 |
| 1946 | -21.19 | 0.55 | 0.30 | 0.00 | 0.25 | 0.00 |

[0372]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 1947 | -21.18 | 0.43 | 0.30 | 0.00 | 0.25 | 0.00 |
| 1948 | -21.10 | 0.36 | 0.32 | 0.00 | 0.00 | 0.31 |
| 1949 | -20.98 | 0.41 | 0.24 | 0.00 | 0.26 | 0.24 |
| 1950 | -20.85 | 0.56 | 0.30 | 0.00 | 0.24 | 0.37 |
| 1951 | -20.78 | 0.47 | 0.26 | 0.00 | 0.23 | 0.31 |
| 1952 | -20.67 | 0.40 | 0.30 | 0.23 | 0.22 | 0.00 |
| 1953 | -20.65 | 0.55 | 0.26 | 0.00 | 0.00 | 0.37 |
| 1954 | -20.62 | 0.51 | 0.28 | 0.28 | 0.26 | 0.30 |
| 1955 | -20.55 | 0.33 | 0.26 | 0.00 | 0.29 | 0.00 |
| 1956 | -20.51 | 0.45 | 0.26 | 0.28 | 0.00 | 0.46 |
| 1957 | -20.51 | 0.66 | 0.26 | 0.00 | 0.25 | 0.26 |
| 1958 | -20.48 | 0.48 | 0.32 | 0.00 | 0.20 | 0.25 |
| 1959 | -20.40 | 0.55 | 0.32 | 0.00 | 0.26 | 0.56 |
| 1960 | -20.37 | 0.48 | 0.29 | 0.24 | 0.22 | 0.00 |
| 1961 | -20.36 | 0.45 | 0.29 | 0.25 | 0.23 | 0.00 |
| 1962 | -20.34 | 0.33 | 0.30 | 0.23 | 0.24 | 0.30 |
| 1963 | -20.33 | 0.59 | 0.30 | 0.00 | 0.22 | 0.25 |
| 1964 | -20.31 | 0.50 | 0.30 | 0.21 | 0.20 | 0.25 |
| 1965 | -20.18 | 0.36 | 0.29 | 0.30 | 0.21 | 0.32 |
| 1966 | -20.15 | 0.38 | 0.28 | 0.31 | 0.29 | 0.27 |
| 1967 | -20.14 | 0.50 | 0.31 | 0.00 | 0.27 | 0.36 |
| 1968 | -20.11 | 0.36 | 0.27 | 0.00 | 0.00 | 0.00 |
| 1969 | -19.97 | 0.40 | 0.23 | 0.23 | 0.28 | 0.63 |
| 1970 | -19.89 | 0.73 | 0.30 | 0.00 | 0.00 | 0.29 |
| 1971 | -19.81 | 0.55 | 0.32 | 0.22 | 0.22 | 0.31 |
| 1972 | -19.49 | 0.46 | 0.33 | 0.00 | 0.25 | 0.51 |
| 1973 | -19.43 | 0.48 | 0.34 | 0.26 | 0.20 | 0.00 |
| 1974 | -19.42 | 0.46 | 0.32 | 0.00 | 0.00 | 0.26 |

[0373] 实施例 5 : 含有的支链氨基酸和亮氨酸的比率大于或等于乳清以及含有所有必需氨基酸的蛋白质的鉴定。

[0374] 评估来自实施例 1 中所述的所述组的 8,415 种蛋白质的片段的氨基酸含量、必需氨基酸 (“EAA”) 的百分比、支链氨基酸 (“BCAA”) 的百分比、亮氨酸 (“L”) 的百分比, 且对于各蛋白质计算蛋白质是否含有所有必需氨基酸。此外, 相对于已知过敏原的数据库筛选蛋白质以确定任何蛋白质是否与已知过敏原具有大于 50% 整体同源性。鉴定出总计 415 个蛋白质片段含有大于或等于 25% BCAA 和大于或等于 13% Leu, 且含有所有必需氨基酸, 并与已知过敏原具有小于 50% 整体同源性 (SEQ ID NO: 2103 至 2518)。对于所述组的蛋白质片段, 计算在 pH 7 下的溶合计分 (“SolvScore”)、在 pH 7 下的聚集计分 (“AggScore”)、过敏原性 (即, 如本文所述的与已知过敏原的局部同源性百分比)、毒性 (即, 如本文所述的与已知毒素的同源性百分比)、抗营养性 (即, 如本文所述的与已知蛋白酶抑制剂的同源性百分比)、和人同源性 (即, 如本文所述的与已知人蛋白质的同源性百分比), 且测定 Cys 残

基（“C”）的总数。由此鉴定的 75 个代表性蛋白质片段的特征呈现于表 7A 和 7B 中。

[0375] 表 7A

[0376]

| Seq ID No | UniProt | FragEnds | EAA  | BCAA | L    | C |
|-----------|---------|----------|------|------|------|---|
| 2103      | O24174  | 76 :251  | 0.49 | 0.26 | 0.14 | 4 |
| 2104      | P24396  | 1 :126   | 0.48 | 0.27 | 0.13 | 2 |
| 2105      | P17801  | 501 :776 | 0.49 | 0.26 | 0.13 | 3 |
| 2106      | O24585  | 501 :751 | 0.47 | 0.26 | 0.14 | 3 |
| 2107      | O24585  | 626 :751 | 0.46 | 0.29 | 0.14 | 1 |
| 2108      | P42211  | 1 :101   | 0.50 | 0.27 | 0.20 | 2 |
| 2109      | Q75L42  | 51 :226  | 0.48 | 0.28 | 0.13 | 3 |
| 2110      | Q9LGV5  | 1 :251   | 0.48 | 0.26 | 0.13 | 4 |
| 2111      | Q9LGV5  | 101 :226 | 0.45 | 0.26 | 0.14 | 3 |
| 2112      | Q10PS6  | 101 :276 | 0.47 | 0.26 | 0.13 | 5 |
| 2113      | Q5JLQ9  | 76 :226  | 0.47 | 0.26 | 0.15 | 3 |
| 2114      | Q2QM10  | 26 :226  | 0.43 | 0.26 | 0.15 | 3 |
| 2115      | Q0D4B2  | 201 :301 | 0.47 | 0.27 | 0.14 | 3 |
| 2116      | Q8GUQ5  | 351 :626 | 0.49 | 0.30 | 0.20 | 6 |
| 2117      | Q8GUQ5  | 451 :626 | 0.47 | 0.31 | 0.22 | 3 |
| 2118      | Q8GUQ5  | 451 :726 | 0.46 | 0.28 | 0.18 | 4 |
| 2119      | Q2QY53  | 126 :226 | 0.46 | 0.25 | 0.17 | 3 |
| 2120      | Q2RAX3  | 126 :226 | 0.46 | 0.25 | 0.17 | 3 |
| 2121      | Q75GK4  | 1 :226   | 0.46 | 0.25 | 0.16 | 3 |

[0377]

|      |        |          |      |      |      |   |
|------|--------|----------|------|------|------|---|
| 2122 | P16243 | 176 :376 | 0.49 | 0.28 | 0.13 | 4 |
|------|--------|----------|------|------|------|---|

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 2123 | Q41807 | 151 :401   | 0. 49 | 0. 26 | 0. 13 | 3 |
| 2124 | Q69Q47 | 51 :226    | 0. 49 | 0. 27 | 0. 14 | 3 |
| 2125 | P46420 | 1 :201     | 0. 48 | 0. 25 | 0. 13 | 3 |
| 2126 | Q6Z9F4 | 51 :226    | 0. 47 | 0. 27 | 0. 13 | 3 |
| 2127 | P16025 | 1026 :1226 | 0. 50 | 0. 26 | 0. 14 | 3 |
| 2128 | P16025 | 926 :1151  | 0. 45 | 0. 26 | 0. 13 | 2 |
| 2129 | P49174 | 1 :76      | 0. 47 | 0. 29 | 0. 14 | 0 |
| 2130 | P48186 | 1 :176     | 0. 49 | 0. 25 | 0. 13 | 1 |
| 2131 | P93648 | 276 :401   | 0. 50 | 0. 26 | 0. 16 | 2 |
| 2132 | P49037 | 551 :626   | 0. 49 | 0. 26 | 0. 15 | 1 |
| 2133 | Q2MIA9 | 201 :276   | 0. 48 | 0. 28 | 0. 15 | 1 |
| 2134 | P16024 | 201 :276   | 0. 48 | 0. 28 | 0. 15 | 1 |
| 2135 | P16024 | 201 :376   | 0. 44 | 0. 26 | 0. 13 | 2 |
| 2136 | P49037 | 1 :101     | 0. 49 | 0. 27 | 0. 14 | 1 |
| 2137 | Q84N48 | 376 :451   | 0. 49 | 0. 27 | 0. 13 | 2 |
| 2138 | Q941W5 | 251 :351   | 0. 45 | 0. 27 | 0. 14 | 1 |
| 2139 | Q75LD5 | 576 :726   | 0. 44 | 0. 25 | 0. 13 | 1 |
| 2140 | Q84ZX8 | 1 :76      | 0. 47 | 0. 25 | 0. 17 | 0 |
| 2141 | Q336R3 | 26 :101    | 0. 43 | 0. 26 | 0. 20 | 3 |
| 2142 | P93231 | 576 :851   | 0. 48 | 0. 26 | 0. 14 | 6 |
| 2143 | P93648 | 276 :551   | 0. 48 | 0. 25 | 0. 13 | 4 |
| 2144 | P16024 | 201 :476   | 0. 47 | 0. 28 | 0. 13 | 3 |
| 2145 | Q2NM15 | 26 :251    | 0. 47 | 0. 26 | 0. 16 | 3 |
| 2146 | P93231 | 576 :751   | 0. 50 | 0. 25 | 0. 14 | 3 |

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 2147 | C7E5V8 | 26 :251    | 0. 47 | 0. 26 | 0. 16 | 3 |
| 2148 | C7E5V7 | 26 :251    | 0. 48 | 0. 27 | 0. 17 | 3 |
| 2149 | P30792 | 126 :226   | 0. 49 | 0. 28 | 0. 15 | 0 |
| 2150 | Q5VS72 | 151 :276   | 0. 46 | 0. 25 | 0. 14 | 1 |
| 2151 | Q9XGC9 | 426 :501   | 0. 45 | 0. 27 | 0. 16 | 0 |
| 2152 | P61242 | 1626 :1776 | 0. 47 | 0. 28 | 0. 14 | 3 |
| 2153 | Q9XGD6 | 126 :251   | 0. 46 | 0. 27 | 0. 14 | 2 |
| 2154 | B6SJQ0 | 1 :176     | 0. 40 | 0. 26 | 0. 16 | 4 |
| 2155 | P93207 | 1 :126     | 0. 42 | 0. 25 | 0. 14 | 2 |
| 2156 | Q653U3 | 801 :876   | 0. 44 | 0. 26 | 0. 14 | 1 |

[0378]

|      |        |           |       |       |       |   |
|------|--------|-----------|-------|-------|-------|---|
| 2157 | P51059 | 351 :576  | 0. 44 | 0. 25 | 0. 15 | 4 |
| 2158 | P07920 | 76 :226   | 0. 48 | 0. 25 | 0. 15 | 2 |
| 2159 | P16024 | 301 :451  | 0. 48 | 0. 28 | 0. 13 | 2 |
| 2160 | P29390 | 176 :251  | 0. 50 | 0. 25 | 0. 14 | 1 |
| 2161 | Q5VQ78 | 576 :676  | 0. 47 | 0. 25 | 0. 14 | 0 |
| 2162 | P16025 | 926 :1026 | 0. 46 | 0. 27 | 0. 13 | 1 |
| 2163 | P51059 | 476 :576  | 0. 46 | 0. 25 | 0. 15 | 1 |
| 2164 | P17847 | 326 :451  | 0. 38 | 0. 25 | 0. 16 | 1 |
| 2165 | P05116 | 76 :226   | 0. 47 | 0. 25 | 0. 15 | 2 |
| 2166 | P24157 | 76 :226   | 0. 47 | 0. 25 | 0. 15 | 2 |
| 2167 | Q9XGC9 | 301 :401  | 0. 48 | 0. 25 | 0. 15 | 1 |
| 2168 | P04711 | 351 :551  | 0. 45 | 0. 25 | 0. 13 | 3 |
| 2169 | Q84ZW8 | 376 :551  | 0. 50 | 0. 26 | 0. 14 | 5 |

|      |        |            |      |      |      |   |
|------|--------|------------|------|------|------|---|
| 2170 | Q6H6V4 | 801 :876   | 0.45 | 0.26 | 0.13 | 1 |
| 2171 | Q9XGD6 | 26 :251    | 0.48 | 0.25 | 0.14 | 2 |
| 2172 | Q41342 | 301 :401   | 0.49 | 0.27 | 0.15 | 0 |
| 2173 | Q42883 | 451 :626   | 0.45 | 0.25 | 0.14 | 6 |
| 2174 | Q2MI98 | 26 :201    | 0.45 | 0.25 | 0.13 | 1 |
| 2175 | Q7XTS4 | 251 :426   | 0.49 | 0.26 | 0.13 | 3 |
| 2176 | Q5Z987 | 1276 :1451 | 0.47 | 0.25 | 0.14 | 8 |
| 2177 | Q6AVM3 | 276 :351   | 0.47 | 0.26 | 0.16 | 2 |

[0379] 表 7B

[0380]

| Seq ID<br>No | 溶合计分   | 聚集计分 | 过敏原性 | 毒性   | 抗营养性 | 人同源性 |
|--------------|--------|------|------|------|------|------|
| 2103         | -17.26 | 0.66 | 0.55 | 0.00 | 0.27 | 0.46 |
| 2104         | -14.60 | 0.68 | 0.49 | 0.25 | 0.26 | 0.23 |
| 2105         | -21.22 | 0.41 | 0.46 | 0.00 | 0.00 | 0.36 |
| 2106         | -19.98 | 0.48 | 0.44 | 0.20 | 0.23 | 0.35 |
| 2107         | -18.81 | 0.59 | 0.44 | 0.00 | 0.00 | 0.40 |
| 2108         | -15.83 | 0.61 | 0.43 | 0.29 | 0.27 | 0.40 |
| 2109         | -20.27 | 0.45 | 0.42 | 0.00 | 0.25 | 0.55 |
| 2110         | -21.01 | 0.39 | 0.42 | 0.00 | 0.00 | 0.51 |

[0381]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 2111 | -19.90 | 0.44 | 0.42 | 0.25 | 0.00 | 0.59 |
| 2112 | -13.77 | 0.63 | 0.41 | 0.00 | 0.19 | 0.28 |
| 2113 | -18.33 | 0.59 | 0.40 | 0.27 | 0.23 | 0.53 |
| 2114 | -19.91 | 0.56 | 0.40 | 0.25 | 0.25 | 0.44 |
| 2115 | -17.67 | 0.47 | 0.39 | 0.00 | 0.24 | 0.59 |
| 2116 | -12.97 | 0.46 | 0.39 | 0.00 | 0.24 | 0.32 |
| 2117 | -12.09 | 0.48 | 0.39 | 0.00 | 0.24 | 0.35 |
| 2118 | -13.41 | 0.44 | 0.39 | 0.00 | 0.22 | 0.31 |
| 2119 | -17.69 | 0.50 | 0.38 | 0.00 | 0.00 | 0.55 |
| 2120 | -17.69 | 0.50 | 0.38 | 0.00 | 0.00 | 0.55 |
| 2121 | -18.70 | 0.55 | 0.38 | 0.21 | 0.00 | 0.42 |
| 2122 | -18.11 | 0.64 | 0.38 | 0.00 | 0.24 | 0.61 |
| 2123 | -16.10 | 0.45 | 0.38 | 0.00 | 0.00 | 0.78 |
| 2124 | -19.89 | 0.47 | 0.37 | 0.24 | 0.23 | 0.48 |
| 2125 | -18.94 | 0.60 | 0.37 | 0.24 | 0.25 | 0.28 |
| 2126 | -21.69 | 0.54 | 0.37 | 0.19 | 0.25 | 0.52 |
| 2127 | -17.96 | 0.48 | 0.36 | 0.24 | 0.00 | 0.00 |
| 2128 | -19.80 | 0.48 | 0.36 | 0.20 | 0.00 | 0.00 |
| 2129 | -10.98 | 0.80 | 0.36 | 0.31 | 0.28 | 0.34 |
| 2130 | -20.67 | 0.47 | 0.35 | 0.00 | 0.23 | 0.00 |
| 2131 | -29.02 | 0.26 | 0.33 | 0.26 | 0.24 | 0.60 |
| 2132 | -28.96 | 0.53 | 0.26 | 0.00 | 0.27 | 0.26 |
| 2133 | -28.54 | 0.45 | 0.28 | 0.23 | 0.26 | 0.30 |
| 2134 | -28.35 | 0.49 | 0.25 | 0.27 | 0.00 | 0.23 |
| 2135 | -27.16 | 0.36 | 0.31 | 0.00 | 0.24 | 0.28 |
| 2136 | -25.78 | 0.45 | 0.33 | 0.00 | 0.25 | 0.30 |
| 2137 | -25.22 | 0.53 | 0.27 | 0.19 | 0.00 | 0.33 |
| 2138 | -24.85 | 0.53 | 0.30 | 0.00 | 0.26 | 0.29 |
| 2139 | -24.69 | 0.52 | 0.33 | 0.20 | 0.00 | 0.00 |
| 2140 | -24.65 | 0.46 | 0.24 | 0.00 | 0.30 | 0.27 |
| 2141 | -24.35 | 0.46 | 0.24 | 0.29 | 0.27 | 0.27 |
| 2142 | -24.29 | 0.47 | 0.30 | 0.00 | 0.00 | 0.40 |
| 2143 | -24.22 | 0.34 | 0.33 | 0.22 | 0.00 | 0.67 |
| 2144 | -23.92 | 0.43 | 0.31 | 0.21 | 0.25 | 0.27 |
| 2145 | -23.69 | 0.40 | 0.32 | 0.00 | 0.00 | 0.00 |

[0382]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 2146 | -23.69 | 0.50 | 0.30 | 0.00 | 0.20 | 0.43 |
| 2147 | -23.61 | 0.41 | 0.32 | 0.00 | 0.00 | 0.23 |
| 2148 | -23.59 | 0.42 | 0.33 | 0.00 | 0.00 | 0.00 |
| 2149 | -23.57 | 0.39 | 0.31 | 0.00 | 0.27 | 0.00 |
| 2150 | -23.48 | 0.33 | 0.30 | 0.00 | 0.24 | 0.34 |
| 2151 | -23.44 | 0.29 | 0.30 | 0.00 | 0.25 | 0.26 |
| 2152 | -23.42 | 0.51 | 0.30 | 0.00 | 0.23 | 0.27 |
| 2153 | -23.41 | 0.52 | 0.29 | 0.27 | 0.00 | 0.40 |
| 2154 | -23.34 | 0.47 | 0.32 | 0.00 | 0.26 | 0.00 |
| 2155 | -23.10 | 0.47 | 0.34 | 0.20 | 0.22 | 0.54 |
| 2156 | -23.09 | 0.45 | 0.28 | 0.25 | 0.28 | 0.00 |
| 2157 | -23.07 | 0.47 | 0.32 | 0.00 | 0.21 | 0.00 |
| 2158 | -23.01 | 0.33 | 0.31 | 0.00 | 0.25 | 0.24 |
| 2159 | -22.95 | 0.47 | 0.31 | 0.24 | 0.21 | 0.32 |
| 2160 | -22.81 | 0.26 | 0.26 | 0.00 | 0.00 | 0.42 |
| 2161 | -22.75 | 0.46 | 0.33 | 0.00 | 0.21 | 0.53 |
| 2162 | -22.74 | 0.44 | 0.27 | 0.27 | 0.28 | 0.00 |
| 2163 | -22.72 | 0.54 | 0.30 | 0.00 | 0.28 | 0.24 |
| 2164 | -22.69 | 0.36 | 0.30 | 0.24 | 0.00 | 0.32 |
| 2165 | -22.64 | 0.34 | 0.31 | 0.25 | 0.00 | 0.00 |
| 2166 | -22.64 | 0.34 | 0.31 | 0.25 | 0.00 | 0.00 |
| 2167 | -22.52 | 0.32 | 0.29 | 0.00 | 0.00 | 0.47 |
| 2168 | -22.29 | 0.48 | 0.31 | 0.00 | 0.25 | 0.00 |
| 2169 | -22.23 | 0.51 | 0.32 | 0.19 | 0.22 | 0.00 |
| 2170 | -22.16 | 0.48 | 0.29 | 0.25 | 0.27 | 0.22 |
| 2171 | -22.13 | 0.49 | 0.32 | 0.00 | 0.00 | 0.38 |
| 2172 | -22.05 | 0.63 | 0.34 | 0.00 | 0.00 | 0.00 |
| 2173 | -21.94 | 0.45 | 0.34 | 0.00 | 0.26 | 0.37 |
| 2174 | -21.89 | 0.34 | 0.29 | 0.00 | 0.21 | 0.00 |
| 2175 | -21.84 | 0.41 | 0.33 | 0.00 | 0.25 | 0.28 |
| 2176 | -21.74 | 0.50 | 0.32 | 0.00 | 0.25 | 0.30 |
| 2177 | -21.74 | 0.50 | 0.32 | 0.30 | 0.28 | 0.38 |

[0383] 实施例 6 :蛋白质表达

[0384] 对编码本文公开的营养性蛋白质的基因针对在大肠杆菌中表达进行密码子最优化且由 LifeTechnologies/GeneArt 或 DNA 2.0 合成。基因被设计来含有两种氨基末端标签之一以有助于纯化：

[0385] MGSHHHHHHHHH(SEQ ID NO:2615)

[0386] MGSSHHHHHHSSGLVPRGSH(SEQ ID NO:2614)

[0387] 使用 NcoI-BamHI 限制位点（在第一标签的情况下）或使用 NdeI-BamHI 限制位点（在第二标签的情况下），将这些基因构建体插入 pET15b 质粒载体（Novagen）中。所有限



制酶都自 New England Biolabs 购买。将质粒转化至大肠杆菌 T7 表达系统 (New England Biolabs) 中且在含有 100mg/l 羧苄青霉素的溶原性培养基 (LB) 板上进行选择。挑选单一菌落,使其在具有 100mg/l 羧苄青霉素的 LB 中生长至  $OD_{600nm} \approx 0.6$ , 且以甘油储备物形式 (于具有 10% 甘油 (v/v) 的 LB 中) 储存在  $-80^{\circ}\text{C}$  下以充当基础细胞储备物 (master cell stock)。

[0388] 2ml 具有 100mg/l 羧苄青霉素的 LB (在 14mm×100mm 培养管中) 用刺针自甘油储备物接种且在  $37^{\circ}\text{C}$  和 250rpm 下生长过夜。次日, 2ml 具有 100mg/l 羧苄青霉素的 LB (在 14mm×100mm 培养管中) 用过夜培养物接种以达到  $OD_{600nm} = 0.05$  且在  $30^{\circ}\text{C}$  或  $37^{\circ}\text{C}$  和 250rpm 下生长。在  $OD_{600nm} \approx 0.8$  下, 用 1mM 异丙基  $\beta$ -D-1-硫代半乳糖吡喃糖苷 (IPTG) 引发异源基因表达且再生长 2 小时 (当在  $37^{\circ}\text{C}$  下生长时) 或 4 小时 (当在  $30^{\circ}\text{C}$  下生长时) 直至收集。在收集后, 测量  $OD_{600nm}$ , 离心 1ml 等分试样, 且倾析上清液。将细胞再混悬以达到  $OD_{600nm} = 1.50$  来进行 SDS-PAGE 分析以评估表达水平。将 10  $\mu$ l 再悬浮的培养物装载于以下之上: 1) **Novex® NuPAGE®** 12% Bis-Tris 凝胶 (Life Technologies), 或 2) **Novex®** 16% Tricine 凝胶 (Life Technologies), 且使用标准制造商方案进行操作。使用标准制造商方案, 利用 SimplyBlue™ SafeStain (Life Technologies) 将凝胶染色, 且使用 Molecular **Imager®** Gel Doc™ XR+ 系统 (Bio-Rad) 进行成像。通过相对于分子量标记和对照培养物进行比较来鉴定过度表达的异源蛋白质。

[0389] 使用这个方法, 观察到表 8A 和 8B 中所列的蛋白质的重组表达。

[0390] 表 8A

[0391]

| Seq ID No | UniProt | FragEnds    | EAA  | BCAA | L    | C |
|-----------|---------|-------------|------|------|------|---|
| 2519      | P15989  | 388 ; 487   | 0.52 | 0.36 | 0.15 | 0 |
| 2520      | P15989  | 341 ; 442   | 0.48 | 0.36 | 0.16 | 0 |
| 2521      | —       | —           | 0.50 | 0.41 | 0.16 | 1 |
| 2522      | P10587  | 1287 ; 1386 | 0.49 | 0.26 | 0.17 | 0 |
| 2523      | Q27991  | 1353 ; 1452 | 0.46 | 0.23 | 0.17 | 0 |
| 2524      | —       | —           | 0.64 | 0.33 | 0.21 | 0 |
| 2525      | P15989  | 398 ; 447   | 0.57 | 0.41 | 0.18 | 0 |
| 2526      | P02662  | 2 ; 51      | 0.59 | 0.31 | 0.18 | 1 |
| 2527      | P02662  | 3 ; 52      | 0.59 | 0.33 | 0.18 | 1 |
| 2528      | P02662  | 4 ; 53      | 0.57 | 0.31 | 0.16 | 1 |

|      |        |            |       |       |       |    |
|------|--------|------------|-------|-------|-------|----|
| 2529 | P02662 | 5 :54      | 0. 55 | 0. 29 | 0. 14 | 1  |
| 2530 | P02662 | 6 :55      | 0. 55 | 0. 29 | 0. 16 | 1  |
| 2531 | P02662 | 79 :128    | 0. 41 | 0. 28 | 0. 12 | 0  |
| 2532 | P02662 | 86 :135    | 0. 41 | 0. 27 | 0. 13 | 0  |
| 2533 | Q9FRT9 | 1 :90      | 0. 32 | 0. 11 | 0. 07 | 10 |
| 2534 | P09643 | 1 :322     | 0. 44 | 0. 20 | 0. 08 | 10 |
| 2535 | Q9HD67 | 472 :521   | 0. 49 | 0. 29 | 0. 20 | 1  |
| 2536 | Q02440 | 461 :511   | 0. 50 | 0. 23 | 0. 13 | 1  |
| 2537 | Q13402 | 452 :509   | 0. 52 | 0. 24 | 0. 11 | 1  |
| 2538 | Q02440 | 413 :512   | 0. 52 | 0. 22 | 0. 09 | 2  |
| 2539 | Q28970 | 427 :527   | 0. 51 | 0. 21 | 0. 10 | 1  |
| 2540 | Q90584 | 436 :486   | 0. 60 | 0. 41 | 0. 32 | 0  |
| 2541 | P22281 | 82 :134    | 0. 58 | 0. 46 | 0. 24 | 1  |
| 2542 | Q17R14 | 270 :321   | 0. 64 | 0. 39 | 0. 20 | 0  |
| 2543 | Q27991 | 1396 :1446 | 0. 49 | 0. 31 | 0. 23 | 0  |
| 2544 | P12106 | 170 :220   | 0. 70 | 0. 32 | 0. 14 | 1  |
| 2545 | P32191 | 112 :163   | 0. 58 | 0. 31 | 0. 15 | 1  |
| 2546 | P19524 | 1398 :1449 | 0. 56 | 0. 25 | 0. 18 | 2  |
| 2547 | Q03262 | 225 :277   | 0. 65 | 0. 30 | 0. 13 | 1  |
| 2548 | P12863 | 82 :132    | 0. 51 | 0. 33 | 0. 17 | 1  |
| 2549 | P12106 | 159 :209   | 0. 67 | 0. 29 | 0. 14 | 1  |
| 2550 | A6QR56 | 589 :691   | 0. 47 | 0. 34 | 0. 21 | 2  |

[0392]

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 2551 | P32492 | 1218 :1319 | 0. 67 | 0. 25 | 0. 14 | 1 |
|------|--------|------------|-------|-------|-------|---|

|      |        |            |       |       |       |   |
|------|--------|------------|-------|-------|-------|---|
| 2552 | P79114 | 688 :788   | 0. 48 | 0. 26 | 0. 15 | 2 |
| 2553 | Q5M1B5 | 492 :592   | 0. 53 | 0. 27 | 0. 14 | 2 |
| 2554 | P04119 | 28 :129    | 0. 51 | 0. 26 | 0. 17 | 2 |
| 2555 | P06642 | 14 :115    | 0. 55 | 0. 27 | 0. 14 | 2 |
| 2556 | Q2XQV4 | 108 :213   | 0. 49 | 0. 26 | 0. 15 | 2 |
| 2557 | P32492 | 1086 :1287 | 0. 58 | 0. 29 | 0. 15 | 1 |
| 2558 | Q76FS2 | 59 :260    | 0. 46 | 0. 23 | 0. 12 | 7 |
| 2559 | Q36967 | 1 :105     | 0. 65 | 0. 46 | 0. 26 | 0 |
| 2560 | Q31721 | 1 :58      | 0. 63 | 0. 38 | 0. 19 | 2 |
| 2561 | Q5ZMNO | 44 :93     | 0. 53 | 0. 36 | 0. 26 | 0 |
| 2562 | Q5ZMNO | 46 :95     | 0. 54 | 0. 36 | 0. 26 | 0 |
| 2563 | Q5ZMNO | 47 :96     | 0. 52 | 0. 36 | 0. 26 | 0 |
| 2564 | Q5ZMNO | 1 :146     | 0. 53 | 0. 30 | 0. 21 | 2 |
| 2565 | Q5ZMNO | 1 :147     | 0. 52 | 0. 30 | 0. 21 | 2 |
| 2566 | Q5ZMNO | 1 :148     | 0. 53 | 0. 30 | 0. 21 | 2 |
| 2567 | Q5ZMNO | 1 :149     | 0. 52 | 0. 29 | 0. 21 | 2 |
| 2568 | Q5ZMNO | 1 :150     | 0. 52 | 0. 29 | 0. 21 | 2 |
| 2569 | Q5ZMNO | 1 :151     | 0. 52 | 0. 29 | 0. 20 | 2 |
| 2570 | Q5ZMNO | 1 :152     | 0. 51 | 0. 29 | 0. 20 | 2 |
| 2571 | Q5ZMNO | 1 :153     | 0. 51 | 0. 29 | 0. 20 | 2 |
| 2572 | Q5ZMNO | 1 :154     | 0. 51 | 0. 28 | 0. 20 | 2 |
| 2573 | Q5ZMNO | 1 :161     | 0. 49 | 0. 27 | 0. 19 | 2 |
| 2574 | Q9JLT0 | 127 :231   | 0. 48 | 0. 21 | 0. 15 | 0 |
| 2575 | Q27991 | 141 :190   | 0. 51 | 0. 28 | 0. 23 | 0 |

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 2576 | Q27991 | 136 :185 | 0. 51 | 0. 28 | 0. 21 | 0 |
| 2577 | Q27991 | 116 :185 | 0. 51 | 0. 27 | 0. 21 | 0 |
| 2578 | Q27991 | 146 :200 | 0. 51 | 0. 29 | 0. 21 | 0 |
| 2579 | Q27991 | 146 :210 | 0. 52 | 0. 27 | 0. 20 | 0 |
| 2580 | Q27991 | 136 :190 | 0. 50 | 0. 27 | 0. 21 | 0 |
| 2581 | Q27991 | 146 :195 | 0. 50 | 0. 30 | 0. 21 | 0 |
| 2582 | Q27991 | 126 :190 | 0. 52 | 0. 26 | 0. 19 | 0 |
| 2583 | Q27991 | 141 :200 | 0. 51 | 0. 28 | 0. 21 | 0 |
| 2584 | Q27991 | 161 :210 | 0. 50 | 0. 24 | 0. 19 | 0 |
| 2585 | Q27991 | 126 :185 | 0. 53 | 0. 27 | 0. 19 | 0 |

[0393]

|      |        |          |       |       |       |   |
|------|--------|----------|-------|-------|-------|---|
| 2586 | Q27991 | 111 :210 | 0. 50 | 0. 26 | 0. 19 | 0 |
| 2587 | Q27991 | 116 :215 | 0. 50 | 0. 26 | 0. 18 | 0 |
| 2588 | Q27991 | 126 :225 | 0. 46 | 0. 22 | 0. 15 | 0 |
| 2589 | Q27991 | 126 :237 | 0. 45 | 0. 21 | 0. 15 | 0 |
| 2590 | Q9JLT0 | 194 :244 | 0. 45 | 0. 20 | 0. 14 | 0 |
| 2591 | Q61879 | 154 :252 | 0. 41 | 0. 19 | 0. 14 | 0 |
| 2592 | P15989 | 99 :208  | 0. 52 | 0. 37 | 0. 17 | 0 |
| 2593 | P15989 | 99 :203  | 0. 51 | 0. 37 | 0. 16 | 0 |
| 2594 | P15989 | 49 :166  | 0. 44 | 0. 28 | 0. 15 | 0 |
| 2595 | Q90339 | 201 :265 | 0. 45 | 0. 21 | 0. 18 | 0 |
| 2596 | Q9BE41 | 216 :265 | 0. 48 | 0. 21 | 0. 18 | 0 |
| 2597 | Q9TV62 | 241 :290 | 0. 48 | 0. 17 | 0. 15 | 0 |
| 2598 | Q5SX39 | 236 :285 | 0. 45 | 0. 19 | 0. 17 | 0 |

|      |        |            |      |      |      |   |
|------|--------|------------|------|------|------|---|
| 2599 | Q9TV61 | 151 :265   | 0.52 | 0.21 | 0.16 | 0 |
| 2600 | Q27991 | 20 :289    | 0.38 | 0.20 | 0.15 | 0 |
| 2601 | Q0WVK7 | 53 :105    | 0.57 | 0.27 | 0.12 | 2 |
| 2602 | Q94A52 | 149 :261   | 0.54 | 0.26 | 0.12 | 2 |
| 2603 | P38111 | 1093 :1165 | 0.59 | 0.30 | 0.13 | 3 |
| 2604 | P38111 | 1093 :1182 | 0.57 | 0.27 | 0.13 | 3 |
| 2605 | P38111 | 1093 :1162 | 0.57 | 0.27 | 0.13 | 3 |
| 2606 | P38111 | 1092 :1166 | 0.59 | 0.29 | 0.13 | 3 |
| 2607 | P38111 | 1093 :1168 | 0.58 | 0.29 | 0.13 | 3 |
| 2608 | P38111 | 1091 :1164 | 0.59 | 0.30 | 0.13 | 3 |
| 2609 | P38111 | 1089 :1164 | 0.58 | 0.29 | 0.13 | 3 |

[0394] 表 8B

[0395]

| Seq ID<br>No | 溶合计分   | 聚集计分 | 过敏原性 | 毒性   | 抗营养性 | 人同源性 |
|--------------|--------|------|------|------|------|------|
| 2519         | -15.79 | 0.80 | 0.31 | 0.00 | 1.00 | 0.67 |
| 2520         | -16.50 | 0.84 | 0.29 | 0.00 | 1.00 | 0.62 |
| 2521         | -21.97 | 1.00 | 0.30 | 0.00 | 0.25 | 0.39 |
| 2522         | -28.44 | 0.16 | 0.36 | 0.00 | 0.00 | 0.84 |
| 2523         | -36.33 | 0.16 | 0.33 | 0.00 | 0.00 | 0.87 |

[0396]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 2524 | -27.36 | 0.49 | 0.34 | 0.00 | 0.30 | 0.41 |
| 2525 | -17.59 | 0.99 | 0.20 | 0.00 | 1.00 | 0.64 |
| 2526 | -17.98 | 0.90 | 0.63 | 0.00 | 0.00 | 0.44 |
| 2527 | -16.59 | 0.95 | 0.63 | 0.00 | 0.32 | 0.44 |
| 2528 | -16.79 | 0.87 | 0.63 | 0.00 | 0.33 | 0.44 |
| 2529 | -18.37 | 0.77 | 0.63 | 0.00 | 0.30 | 0.42 |
| 2530 | -18.37 | 0.71 | 0.63 | 0.00 | 0.32 | 0.40 |
| 2531 | -24.84 | 0.25 | 0.63 | 0.26 | 0.00 | 0.35 |
| 2532 | -26.10 | 0.22 | 0.63 | 0.26 | 0.00 | 0.33 |
| 2533 | -15.45 | 0.39 | 0.32 | 1.00 | 0.34 | 0.34 |
| 2534 | -19.21 | 0.44 | 0.99 | 0.00 | 0.00 | 0.96 |
| 2535 | -25.30 | 0.46 | 0.22 | 0.33 | 0.34 | 1.00 |
| 2536 | -22.49 | 0.43 | 0.21 | 0.31 | 0.35 | 0.94 |
| 2537 | -22.06 | 0.51 | 0.21 | 0.28 | 0.27 | 1.00 |
| 2538 | -21.01 | 0.41 | 0.31 | 0.00 | 0.28 | 0.94 |
| 2539 | -23.16 | 0.43 | 0.28 | 0.00 | 0.00 | 0.98 |
| 2540 | -15.02 | 1.17 | 0.22 | 0.34 | 0.35 | 0.75 |
| 2541 | -10.60 | 1.38 | 0.22 | 0.00 | 0.34 | 0.28 |
| 2542 | -18.91 | 0.89 | 0.20 | 0.00 | 0.32 | 0.94 |
| 2543 | -31.57 | 0.29 | 0.21 | 0.00 | 0.31 | 0.88 |
| 2544 | -17.13 | 0.77 | 0.20 | 0.31 | 0.28 | 0.61 |
| 2545 | -16.98 | 0.76 | 0.19 | 0.25 | 0.17 | 0.52 |
| 2546 | -20.09 | 0.24 | 0.21 | 0.29 | 0.30 | 0.43 |
| 2547 | -19.93 | 0.52 | 0.22 | 0.32 | 0.32 | 0.33 |
| 2548 | -16.57 | 0.79 | 0.58 | 0.33 | 0.24 | 0.59 |
| 2549 | -16.05 | 0.65 | 0.20 | 0.28 | 0.31 | 0.65 |
| 2550 | -17.15 | 0.83 | 0.30 | 0.30 | 0.00 | 0.84 |
| 2551 | -21.64 | 0.55 | 0.26 | 0.26 | 0.26 | 0.00 |
| 2552 | -24.45 | 0.54 | 0.32 | 0.24 | 0.25 | 0.91 |
| 2553 | -23.71 | 0.37 | 0.30 | 0.27 | 0.29 | 0.93 |
| 2554 | -22.23 | 0.52 | 0.74 | 0.00 | 0.19 | 0.44 |
| 2555 | -18.41 | 0.58 | 0.29 | 0.29 | 0.25 | 0.78 |
| 2556 | -17.59 | 0.58 | 0.60 | 0.00 | 0.23 | 0.98 |
| 2557 | -19.98 | 0.57 | 0.31 | 0.00 | 0.24 | 0.00 |
| 2558 | -17.02 | 0.48 | 0.53 | 0.00 | 0.22 | 0.88 |

[0397]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 2559 | -5.92  | 1.63 | 0.30 | 0.00 | 0.27 | 0.61 |
| 2560 | -6.10  | 1.53 | 0.21 | 0.00 | 0.26 | 0.40 |
| 2561 | -18.69 | 0.54 | 0.18 | 0.24 | 0.35 | 0.84 |
| 2562 | -17.31 | 0.54 | 0.20 | 0.24 | 0.33 | 0.84 |
| 2563 | -17.40 | 0.47 | 0.20 | 0.22 | 0.34 | 0.84 |
| 2564 | -22.65 | 0.39 | 0.32 | 0.25 | 0.25 | 0.76 |
| 2565 | -22.50 | 0.39 | 0.32 | 0.25 | 0.25 | 0.76 |
| 2566 | -22.35 | 0.39 | 0.32 | 0.25 | 0.26 | 0.76 |
| 2567 | -22.74 | 0.39 | 0.32 | 0.00 | 0.25 | 0.76 |
| 2568 | -22.59 | 0.39 | 0.32 | 0.00 | 0.25 | 0.75 |
| 2569 | -22.97 | 0.38 | 0.32 | 0.24 | 0.25 | 0.76 |
| 2570 | -23.34 | 0.38 | 0.32 | 0.24 | 0.25 | 0.75 |
| 2571 | -23.25 | 0.38 | 0.32 | 0.24 | 0.25 | 0.75 |
| 2572 | -23.61 | 0.37 | 0.32 | 0.00 | 0.25 | 0.75 |
| 2573 | -24.11 | 0.36 | 0.32 | 0.23 | 0.26 | 0.75 |
| 2574 | -33.59 | 0.15 | 0.34 | 0.25 | 0.00 | 1.00 |
| 2575 | -36.26 | 0.21 | 0.21 | 0.29 | 0.33 | 0.92 |
| 2576 | -34.27 | 0.27 | 0.22 | 0.25 | 0.00 | 0.90 |
| 2577 | -32.73 | 0.22 | 0.25 | 0.31 | 0.00 | 0.87 |
| 2578 | -34.50 | 0.23 | 0.21 | 0.00 | 0.00 | 0.89 |
| 2579 | -34.08 | 0.25 | 0.26 | 0.00 | 0.00 | 0.91 |
| 2580 | -34.67 | 0.22 | 0.23 | 0.25 | 0.00 | 0.91 |
| 2581 | -33.39 | 0.29 | 0.20 | 0.00 | 0.31 | 0.88 |
| 2582 | -37.57 | 0.19 | 0.25 | 0.31 | 0.00 | 0.92 |
| 2583 | -35.42 | 0.21 | 0.24 | 0.21 | 0.00 | 0.90 |
| 2584 | -33.68 | 0.26 | 0.23 | 0.00 | 0.00 | 0.92 |
| 2585 | -37.47 | 0.23 | 0.25 | 0.00 | 0.00 | 0.92 |
| 2586 | -33.38 | 0.20 | 0.31 | 0.00 | 0.00 | 0.87 |
| 2587 | -32.78 | 0.18 | 0.33 | 0.27 | 0.29 | 0.88 |
| 2588 | -36.89 | 0.16 | 0.35 | 0.00 | 0.00 | 0.92 |
| 2589 | -37.66 | 0.15 | 0.39 | 0.00 | 0.00 | 0.93 |
| 2590 | -29.21 | 0.16 | 0.25 | 0.00 | 0.27 | 1.00 |
| 2591 | -30.24 | 0.12 | 0.33 | 0.00 | 0.00 | 0.99 |
| 2592 | -16.02 | 0.94 | 0.32 | 0.00 | 1.00 | 0.63 |
| 2593 | -16.69 | 0.94 | 0.30 | 0.00 | 1.00 | 0.63 |

[0398]

|      |        |      |      |      |      |      |
|------|--------|------|------|------|------|------|
| 2594 | -17.69 | 0.65 | 0.31 | 0.00 | 1.00 | 0.62 |
| 2595 | -34.37 | 0.07 | 0.29 | 0.00 | 0.00 | 0.98 |
| 2596 | -38.56 | 0.08 | 0.24 | 0.00 | 0.00 | 0.98 |
| 2597 | -41.59 | 0.04 | 0.26 | 0.00 | 0.00 | 0.98 |
| 2598 | -40.29 | 0.04 | 0.23 | 0.28 | 0.00 | 0.98 |
| 2599 | -35.70 | 0.11 | 0.36 | 0.26 | 0.22 | 0.99 |
| 2600 | -34.97 | 0.13 | 0.40 | 0.23 | 0.22 | 0.93 |
| 2601 | -26.91 | 0.45 | 0.18 | 0.30 | 0.22 | 0.31 |
| 2602 | -25.52 | 0.46 | 0.29 | 0.00 | 0.25 | 0.81 |
| 2603 | -23.00 | 0.80 | 0.27 | 0.00 | 0.24 | 0.26 |
| 2604 | -20.82 | 0.80 | 0.27 | 0.00 | 0.25 | 0.27 |
| 2605 | -23.98 | 0.70 | 0.25 | 0.27 | 0.25 | 0.27 |
| 2606 | -22.45 | 0.79 | 0.27 | 0.27 | 0.25 | 0.25 |
| 2607 | -22.10 | 0.80 | 0.27 | 0.28 | 0.24 | 0.25 |
| 2608 | -22.75 | 0.77 | 0.28 | 0.31 | 0.25 | 0.25 |
| 2609 | -23.28 | 0.71 | 0.28 | 0.28 | 0.25 | 0.24 |

[0399] 实施例 7 :按比例扩大产生重组营养性蛋白质

[0400] 用于产生大量如本公开中所述的营养性蛋白质的代表性方案如下

[0401] 5ml 具有 100mg/l 羧苄青霉素的 LB(在 50ml 有挡板 Pyrex 振荡烧瓶中)用刺针自包含编码营养性蛋白质的重组基因的重组大肠杆菌菌株的甘油储备物接种且在 37℃ 和 250rpm 下生长直至指数期晚期 ( $OD_{600nm} \approx 2$ )。2.5l Ultra Yield 烧瓶 (Thomson Instrument Company) 用 500ml 无菌水和足够 EnBase EnPresso™ 片 (BioSilta) 接种以配制 500ml 生长培养基。这个培养基补充以 100mg/l 羧苄青霉素、0.001% Industrol 204 消泡剂和 0.6U/l EnzI'm (BioSilta)。对振荡烧瓶进行接种以达到  $OD_{600nm} = 0.05$  且在 30℃ 和 250rpm 下生长 16 小时。此时,已达到  $OD_{600nm} \approx 10$ , 生长培养基进一步补充以 EnPresso™ 增效片 (BioSilta)、1.2U/l EnzI'm 和 1mM IPTG 以诱导异源蛋白质产生。在 30℃ 和 250rpm 下再振荡 8-24 小时之后,通过在 10,000rpm 下离心收集烧瓶,倾析上清液,且测量湿细胞重量。此时,回收约 20gWCW(克湿细胞重量)/1 培养基。

[0402] 将由各振荡烧瓶发酵收集的细胞混悬于 25mL IMAC 平衡溶液 (30mM 咪唑、50mM 磷酸盐、0.5M NaCl, pH 7.5) 中。接着通过在冰上在 90 瓦特下超声处理 90 秒来溶解混悬的细胞。将溶解的细胞在 15,500RCF 下离心 60 分钟且倾析。丢弃细胞碎片,且使上清液经受 0.2  $\mu m$  过滤。过滤器接着再用 10mL IMAC 平衡溶液冲洗。这些过滤的蛋白质溶液接着通过固定金属亲和色谱 (IMAC) 纯化。

[0403] IMAC 树脂 (GE Healthcare, 速流 IMAC 琼脂糖 6) 以镍装料且加以平衡。将 30mL 各蛋白质溶液装载于 5mL IMAC 柱上,且再用平衡溶液洗涤以移除未结合的杂质。目标蛋白质接着用 15mL 0.5M NaCl、0.2M 咪唑 (pH 7.5) 洗脱。根据 SDS-PAGE, 纯化蛋白质显示纯度 90%。在 IMAC 洗脱份中回收约 20 至 60mg 各蛋白质。通过透析将各 IMAC 洗脱份的缓冲液交换成配制溶液 (20mM HEPES, pH 7.5)。在缓冲液交换之后,回收这些蛋白质溶液以进行所有下游处理。



[0404] 实施例 8 :营养性蛋白质的可溶性表达的预测

[0405] 克隆编码一组 292 种营养性蛋白质的开放阅读框且将其引入大肠杆菌中来使用实施例 6 的方法评估重组蛋白表达。在所用系统中,163 种蛋白质被鉴定为表达,而 129 种蛋白质并非如此。在表达的 163 种蛋白质中,测试 125 种蛋白质的可溶性表达。发现 75 种蛋白质可溶性表达,而 50 种蛋白质并非如此。

[0406] 图 1 显示大肠杆菌表达筛选中的蛋白质表达的二维直方图。图 1 显示随溶合计分(y 轴)和聚集计分(x 轴)而变的蛋白质表达的相对可能性(在对数尺度上)。直方图上的深色标记指示较高数目的蛋白质表达,而浅色标记指示较少数目的蛋白质表达。图 1 显示成功表达的那些蛋白质倾向于群集在图的顶部左侧区域中,其中溶合计分负性更大( $\leq -20$ )且聚集计分较小( $\leq 0.75$ )。存在少数几例成功表达的蛋白质具有负性较小的溶合计分( $\geq -15$ )和大聚集计分( $\geq 1$ )。这个结果表明溶合计分是 -20 或更低且聚集计分是 0.75 或更低的营养性蛋白质更可能在这个系统中表达。

[0407] 图 2 显示大肠杆菌表达筛选中的可溶性蛋白质表达的数目的二维直方图。图 2 显示随溶合计分(y 轴)和聚集计分(x 轴)而变的蛋白质可溶性表达的相对可能性(在对数尺度上)。再次,直方图上的深色标记指示较高数目的蛋白质表达,而浅色标记指示较少数目的蛋白质表达。图 1 显示可溶性表达的那些蛋白质倾向于群集在图的顶部左侧区域中,其中溶合计分负性更大( $\leq -20$ )且聚集计分较小( $\leq 0.5$ )。存在少数几例可溶性表达的蛋白质具有负性较小的溶合计分( $\geq -15$ )和大聚集计分( $\geq 0.75$ )。这个结果表明溶合计分是 -20 或更低且聚集计分是 0.5 或更低的营养性蛋白质更可能在这个系统中可溶性表达。

[0408] 实施例 9 :溶解度筛选

[0409] 通过离心浓缩,随后进行蛋白质浓度测定来考查如实施例 6 和 7 中所述产生的九种营养性蛋白质的溶解度。根据用于考马斯加强型(Bradford)蛋白质测定(Thermo Scientific)的方案测试样本在 20mM HEPES(pH 7.5)中的蛋白质浓度以及在 280nm 下的吸光度(如果适用)。基于这些测量结果,添加 10mg 蛋白质至 Amicon Ultra 3kDa 离心过滤器(Millipore)中。通过在 10,000Xg 下离心 30 分钟来浓缩样本。考查最终浓缩样本的蛋白质沉淀和颜色,且接着如上所述测试蛋白质浓度。结果显示于表 9 中。

[0410] 表 9

[0411]

| Seq ID No | 外观    | 浓度 (g/L) |
|-----------|-------|----------|
| 2580      | 澄清淡黄色 | 44       |
| 2582      | 澄清淡黄色 | 166      |
| 2587      | 澄清无色  | 107      |
| 2595      | 澄清无色  | 60       |
| 2596      | 澄清无色  | 29       |

|      |      |     |
|------|------|-----|
| 2598 | 澄清无色 | 151 |
| 2599 | 澄清无色 | 207 |
| 2600 | 澄清黄色 | 95  |
| 2603 | 澄清黄色 | 191 |

[0412] 发现这些营养性蛋白质的溶解度显著高于乳清 (12.5g/L) 和大豆 (10g/L) 通常所见的浓度 (Pelegri, D. H. G. 和 Gasparetto, C. A., 2005. Whey proteins solubility as function of temperature and pH. LWT-Food Science and Technology, 第 77 - 80 页; Lee, K. H., Ryu, H. S. 和 Rhee, K. C., 2003. Protein solubility characteristics of commercial soy protein products. Journal of the American Oil Chemists' Society, 第 85-90 页)。这说明本文公开的营养性蛋白质的适用性。举例来说, 营养性蛋白质的溶解度可改进高品质蛋白质以尽可能小的体积递送的顺应性, 同时避免常为以这个方式递送的蛋白质的特征的“粉笔气味 (chalkyness)”。这可例如适用于向年长受试者或其它受试者递送蛋白质。

#### [0413] 实施例 10 : 稳定性筛选

[0414] 营养性蛋白质的热稳定性提供关于蛋白质是否可能具有适用储存期限的见地。使用快速热稳定性筛选方法平行筛选如实施例 6 和 7 中所述产生的蛋白质的样本。在这个方法中, 在结合随温度增加而作为蛋白质变性物形成的聚集蛋白质的疏水性染料 (Enzo Life Sciences, **ProteoStat®** 热漂移稳定性测定试剂盒) 存在下, 将蛋白质以两种代表性制剂形式自 25°C 缓慢加热至 95°C (Niesen, F. H., Berglund, H. 和 Vadadi, M., 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols, 第 2 卷, 第 2212-2221 页。)。在结合后, 染料的荧光显著增加, 此接着由 rtPCR 仪器记录且表示为蛋白质的熔融曲线 (Lavinder, J. J., Hari, S. B., Sullivan, B. J. 和 Magilery, T. J., 2009. High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. Journal of the American Chemical Society, 第 3794-3795 页。)。在完成热漂移之后, 考查样本的不溶性沉淀且由分析型尺寸排阻色谱法 (SEC) 加以进一步分析。

[0415] 在各自含有 1X ProteoStat TS 检测试剂的 PBS 与 20mM HEPES (pH 7.7) 缓冲液两者中制备蛋白质溶液 (12.5mg/ml)。使用实时 PCR (rtPCR) 热循环仪, 在监测染料的荧光的同时, 将各溶液的样本以 0.5°C / 30 秒自 25°C 缓慢加热至 95°C。根据这个热扫描, 如果观察到荧光增加, 那么由斜率最强的温度确定聚集温度 ( $T_{agg}$ )。为补充所述测定, 在热漂移之前和之后获取样本且通过可检测大型可溶性聚集体的 SEC (GE Healthcare-Superdex 75 5/150) 来分析。七种本公开的营养性蛋白质和乳清标准物的结果呈现于表 10 中。如果观察到, 那么由 SEC 检测到存在可溶性聚集体被用“是”注释, 如果未观察到, 那么用“否”注释, 且乳清标准物的“n/a”条目指示产生不溶性沉淀以致未进行 SEC 分析。

[0416] 表 10

[0417]

| Seq ID<br>No | HEPES-T <sub>agg</sub> | PBS-T <sub>agg</sub> | HEPES-SEC<br>Agg? | PBS-SEC<br>Agg? |
|--------------|------------------------|----------------------|-------------------|-----------------|
| 2582         | 95                     | 95                   | 否                 | 否               |
| 2587         | 95                     | 95                   | 否                 | 否               |
| 2596         | 45                     | 95                   | 否                 | 否               |
| 2598         | 95                     | 95                   | 否                 | 否               |
| 2599         | 95                     | 95                   | 否                 | 否               |
| 2600         | 95                     | 95                   | 是                 | 是               |
| 2603         | 95                     | 95                   | 是                 | 是               |
| 乳清           | 79                     | 81.5                 | n/a               | n/a             |

[0418] 如表 10 中所示, SEQ ID NO:2582、2587、2596、2598 和 2599 具有的 HEPES T<sub>agg</sub> 比乳清高得多(实际上,蛋白质在测定上限 95℃下不形成任何聚集体),且因此预期比乳清更稳定。SEQ ID NO:2600 和 2603 蛋白质具有的 HEPES T<sub>agg</sub> 和 PBS T<sub>agg</sub> 高于乳清,但两者均在溶液中形成可溶性聚集体,而乳清形成自溶液沉淀的不溶性聚集体。因此,即使在聚集状态下,也将预期这两种营养性蛋白质保持溶解的时期大于乳清。

[0419] 实施例 11:可消化性筛选 - 消化半衰期的测定

[0420] 筛选蛋白质可消化性的目标在于消除潜在不安全过敏原性蛋白质且将相对消化完全性确定为肽生物可用度的预测因子。这个筛选方法利用包括两个蛋白质消化阶段(即模拟胃消化和模拟肠消化)的生理相关体外消化反应(Moreno, J. F. 等, 2005. Stability of the major allergen Brazil nut 2S albumin(Ber e l)to physiologically relevant in vitro gastrointestinal digestion. FEBS Journal, 第 341-352 页。)。可在整个反应期间获取样本且使用芯片电泳和 LC-QTOF-MS 分析完整蛋白质和肽片段。可通过鉴定对消化性蛋白酶具有抗性且因此具有导致过敏原性反应的较高风险的蛋白质或蛋白质的大片段来评估具有过敏原性性质的蛋白质(Goodman, R. E. 等, 2008. Allergenicity assessment of genetically modified crops-what makes sense? .Nature Biotechnology, 第 73-81 页。)。可消化性是通过测定蛋白质被分解成肽的高效性程度来测量(Daniel, H., 2003. Molecular and Integrative Physiology of Intestinal Peptide Transport. Annual Review of Physiology, 第 66 卷, 第 361-384 页。))。

[0421] 方法使用针对蛋白质的体外消化的自动测定,其中测定条件和蛋白酶浓度是生理相关的(Moreno, F. J., Mackie, A. R. 和 Clare Mills, E. N., 2005. Phospholipid interactions protect the milk allergen  $\alpha$ -Lactalbumin from proteolysis during in vitro digestion. Journal of agricultural and food chemistry, 第 9810-9816 页; Martos, G., Contreras, P., Molina, E. 和 Lopez-Fandino, R., 2010. Egg White Ovalbumin Digestion Mimicking Physiological Conditions. Journal of Agricultural and food chemistry, 第 5640-5648 页; Moreno, J. F. 等, 2005. Stability of the major

allergen Brazil nut 2S albumin(Ber e 1)to physiologically relevant in vitro gastrointestinal digestion.FEBS Journal, 第 341-352 页。)。第一消化阶段在模拟胃液 (SGF) 中且在 pH 1.5 下配制并具有胃蛋白酶:底物比率 (1:10w/w)。第二消化阶段在模拟肠液 (SIF) 中,与胆汁盐一起在 pH 6.5 下配制并具有胰蛋白酶:胰凝乳蛋白酶:底物比率 (1:4:400w/w)。在模拟胃液中处理蛋白质 120 分钟,这是 90%液体餐食自胃传至小肠花费的时间 (Kong, F. 和 Singh, R. P., 2008. Disintegration of Solid Foods in Human Stomach. Journal of Food Science, 第 67-80 页),且接着用模拟肠液处理 120 分钟。在整个两个反应期间的时间点获取样本且淬灭以进行分析。易于由胃蛋白酶消化的牛血清白蛋白是 SGF 溶液的阳性对照,且天然对胃蛋白酶具有抗性但在 SIF 中消化的  $\beta$ -乳球蛋白是 SIF 溶液的阳性对照。使用电泳检测完整蛋白质和大片段。对于芯片电泳,配备有 HT 低分子量蛋白质测定试剂盒的 Caliper Labchip GXII 用于监测完整蛋白质以及大于 4kDa 的任何消化片段的尺寸和量。通过监测随时间观察到的完整蛋白质的量,计算 SGF 消化半衰期 ( $\tau_{1/2}$ ),且如果在 SGF 消化之后检测到完整蛋白质,那么计算在 SIF 中的消化半衰期。

[0422] 这个方法用于分析十三种如实施例 6 和 7 中所述产生的本公开的营养性蛋白质 (SEQ ID NO:2547、2575、2578、2580、2582、2587、2595、2596、2598、2599、2600、2602 和 2603) 以及天然和重组卵清蛋白 (分别是 OVA 和 rOVA;SEQ ID NO:2610) 和  $\beta$ -乳球蛋白 (分别是 BLG 和 rBLG;SEQ ID NO:2611) 蛋白质和乳清标准物的消化半衰期。这些实验的结果概述于表 11 中。模拟肠液区域中的“n/a”条目指示在 SGF 消化之后未检测到完整蛋白质。

[0423] 表 11

[0424]

| Seq ID No.  | 消化 $\tau_{1/2}$ (分钟) |      |
|-------------|----------------------|------|
|             | 模拟胃液                 | 模拟胃液 |
| 2547        | 0.3                  | n/a  |
| 2575        | 0.3                  | n/a  |
| 2578        | 2                    | n/a  |
| 2580        | 6                    | n/a  |
| 2582        | 0.5                  | n/a  |
| 2587        | 0.7                  | n/a  |
| 2595        | 10                   | n/a  |
| 2596        | 0.6                  | n/a  |
| 2598        | 0.7                  | n/a  |
| 2599        | 0.3                  | n/a  |
| 2600        | 29                   | 1    |
| 2602        | 1                    | n/a  |
| 2603        | 6                    | n/a  |
| BLG (2611)  | 77                   | 4    |
| rBLG (2611) | 50                   | 0.7  |
| OVA (2610)  | 18                   | 1    |
| rOVA (2610) | 5                    | n/a  |
| 乳清          | 99                   | 4    |

[0425] 表 11 中所示的结果指示十三种本公开的营养性蛋白质中的十二者都由 SGF 完全消化且具有 SGF 半衰期十分钟或更短。比较而言,乳清不由 SGF 完全消化且具有 SGF 半衰期 99 分钟和 SIF 半衰期 4 分钟。这个研究表明本公开的营养性蛋白质可能易于消化且当摄取时不可能引发过敏反应。

[0426] 表 11 中的结果也显示根据本公开产生的重组  $\beta$ -乳球蛋白和卵清蛋白均比它们的天然存在的对应物更易于被消化。可通过针对改进或限制胃肠蛋白酶的可及性的性质加以选择来控制蛋白质被分解的速度。可证明两种典型蛋白质性质(即糖基化和二硫交联)具有这个能力。与许多天然存在的蛋白质类似,天然存在的 OVA 和 BLG 是由它们的宿主生物体加以糖基化。相反,根据本公开产生的重组蛋白不被糖基化,因为宿主生物体(在这个情况下是大肠杆菌)不进行糖基化。根据本公开的营养性蛋白质中缺乏糖基化可产生更易于被消化的蛋白质。此外,BLG 具有已知会减缓或干扰消化的四个二硫键。当这些二硫键被破坏时,消化速率增加 (Reddy, I. M., Kella, N. K. D. 和 Kinsella, J. E., 1988. Structural and conformational Basis of the Resistance of  $\beta$ -Lactoglobulin to Peptic and Chymotryptic Digestion. J. Agric. Food Chem., 第 36 卷, 第 737-741 页)。缺乏或破坏根据本公开的营养性蛋白质中的二硫键形成可产生更易于被消化的蛋白质。

[0427] 实施例 12:可消化性筛选 - 消化产物的分析

[0428] 使如实施例 6 和 7 中所述产生的两种营养性蛋白质 (SEQ ID NO:2612 和 2613) 经受如实施例 11 中所述的 SGF 和 SIF 消化。两种蛋白质均在 SGF 中完全消化,且 SGF 半衰期显示于表 12 中。

[0429] 表 12

[0430]

| Seq ID<br>No | 消化 $\tau_{1/2}$ (分钟) |      |
|--------------|----------------------|------|
|              | 模拟胃液                 | 模拟胃液 |
| 2612         | 0.7                  | n/a  |
| 2613         | 6                    | n/a  |

[0431] 为检测和鉴定在 SGF 和 SIF 消化之后存在的肽,通过 LC/Q-TOF MS/MS 分析 SGF 和 SIF 消化液的样本。通过 LC/Q-TOF MS/MS 直接分析来自 SGF 消化液的样本,而 SIF 蛋白质消化需要在通过 LC/Q-TOF MS/MS 检测和鉴定之前通过 SCX 进行纯化以移除胆汁酸。自色谱图摘取肽且使用 Bioconfirm 软件 (Aglient) 加以鉴定。基于准确质量匹配 ( $\pm 10\text{ppm}$ ) 对肽进行序列指定且通过 MS/MS 断裂加以进一步确认。结果显示于下表 13 和 14 中。

[0432] 表 13

[0433] (SEQ ID NO :2612)

[0434]

| SGF 肽 120 分钟 | SEQ ID NO | SIF 肽 120 分钟 | SEQ ID NO |
|--------------|-----------|--------------|-----------|
| LL           |           | SE           |           |
| LAL          |           | PSE          |           |
| HVL          |           | HVL          |           |
| LEL          |           | FKV          |           |
| LALA         | 2616      | HQI          |           |
| LLLD         | 2617      | PSEA         | 2618      |
| IAEF         | 2619      | REV          |           |
| IQQF         | 2620      | FDK          |           |
| YDKL         | 2621      | AEFK         | 2622      |
| SNLTE        | 2623      | LKHV         | 2624      |

|       |      |        |      |
|-------|------|--------|------|
| ELLEA | 2625 | SSSEL  | 2626 |
| EELAL | 2627 | FKVF   | 2628 |
| DDLLE | 2629 | AELKH  | 2630 |
| LAYDK | 2631 | LKHVL  | 2632 |
| TKTRL | 2633 | FKEAF  | 2634 |
| DLDHQ | 2635 | NGSISS | 2636 |

[0435]

|            |      |           |      |
|------------|------|-----------|------|
| GTLENL     | 2637 | SLGLSPS   | 2638 |
| EKLTD      | 2639 | GEKLT     | 2640 |
| AEVDDM     | 2641 | ELATVM    | 2642 |
| LDDLL      | 2643 | GSGEINI   | 2644 |
| LDLDHQ     | 2645 | RSLGLSP   | 2646 |
| DLKKKL     | 2647 | ELKHVL    | 2648 |
| KTKTRL     | 2649 | KLDAEV    | 2650 |
| SQRLEE     | 2651 | FKVFDK    | 2652 |
| AEVDDML    | 2653 | AEVDDML   | 2654 |
| QQELDDL    | 2655 | DAEVDDM   | 2656 |
| LEKTKTR    | 2657 | AELKHVL   | 2658 |
| LEKTKTR    | 2659 | HVLTSGE   | 2660 |
| GTLENLEE   | 2661 | SIGEKLT   | 2662 |
| QQELDDL    | 2663 | RSLGLSPSE | 2664 |
| KLEKTKTRLQ | 2665 | VLTSIGEL  | 2666 |
| SRQLKSNSEQ | 2667 | TSIGEKLT  | 2668 |
| EKTKTRLQEL | 2669 | HVLTSGEK  | 2670 |

|                        |      |                       |      |
|------------------------|------|-----------------------|------|
| YDKLEKTKTRL            | 2671 | AELKHVLTSL            | 2672 |
| LAYDKLEKTKTRL          | 2673 | RSLG LSPSEA           | 2674 |
| EELKKKLLKDLEL          | 2675 | SSNLTEEQIA            | 2676 |
| EELKKKLLKDLELL         | 2677 | RSLGLSPSEAE           | 2678 |
| AELKHVLTSLGEKLT        | 2679 | KVFDKNGDGLISA         | 2680 |
| MGSHHHHHHHSSNL         | 2681 | FDKDNNGSISSSEL        | 2682 |
| LREVSDGSGEINIQQF       | 2683 | REVSDGSGEINIQQ        | 2684 |
| AAELKHVLTSLGEKLT       | 2685 | DVDGNHQIEFSEF         | 2686 |
| AELKHVLTSLGEKLTDAE     | 2687 | LREVSDGSGEINIQQ       | 2688 |
| AYDKLEKTKTRLQQEL       | 2689 | REVSDGSGEINIQQF       | 2690 |
| AAELKHVLTSLGEKLTDAE    | 2691 | LREVSDGSGEINIQQFAALLS | 2692 |
| LENLEELKKKLLKDLEL      | 2693 |                       |      |
| KLEKTKTRLQQELDDL       | 2694 |                       |      |
| DKLEKTKTRLQQELDDL      | 2695 |                       |      |
| LAYDKLEKTKTRLQQELDDL   | 2696 |                       |      |
| LALAYDKLEKTKTRLQQELDDL | 2697 |                       |      |

[0436] 表 14

[0437] (SEQ ID NO :2613)

[0438]

| SGF 肽 120 分钟 | SEQ ID NO | SIF 肽 120 分钟 | SEQ ID NO |
|--------------|-----------|--------------|-----------|
| GVL          |           | TKH          |           |
| ALL          |           | INDI         | 2698      |
| LVL          |           | HLVL         | 2699      |
| IGVL         | 2700      | TIKF         | 2701      |
| TIKF         | 2702      | IGVLD        | 2703      |



|                |      |               |      |
|----------------|------|---------------|------|
| TIKF           | 2704 | RNLD          | 2705 |
| EVYDL          | 2706 | IGVLDV        | 2707 |
| LNDSVQ         | 2708 | QTIKF         | 2709 |
| IWVIND         | 2710 | VQTIKF        | 2711 |
| DLNDSVQ        | 2712 | SVQTIKF       | 2713 |
| SVQTIKF        | 2714 | KCAKCISMIGVL  | 2715 |
| HHLVLGALLD     | 2716 | EKCAKCISMIGV  | 2717 |
| HHHHHHLVL      | 2718 | HEFKRTTYSE    | 2719 |
| HHHHHHHHLVL    | 2720 | SHKFRNLDKDL   | 2721 |
| DVTKEFKRTTY    | 2722 | ISMIGVLDVTKHE | 2723 |
| SHHHHHHHHLVL   | 2724 |               |      |
| TKHEFKRTTYSEN  | 2725 |               |      |
| GSHHHHHHHHLVLG | 2726 |               |      |
| MGSHHHHHHHHLVL | 2727 |               |      |
| KRTTYSENEVYDLN | 2728 |               |      |

[0439] 如可在表 13 和 14 中所见,各蛋白质被消化成大小在 2 至 22 个氨基酸 (SEQ ID NO : 2612) 或 2 至 13 个氨基酸 (SEQ ID NO :2613) 的范围内的多个较小肽片段。未发现这些肽片段与任何已知过敏原同源。

[0440] 尽管本发明已关于其特定实施方案加以描述,但本领域技术人员应了解的是可在不脱离本发明的真实精神和范围下进行各种变化且可用等效物进行替代。此外,可进行许多修改以使特定情况、材料、物质组成、方法、一个或多个方法步骤适于本发明的目标、精神和范围。所有所述修改都意图在随附于此的权利要求的范围内。

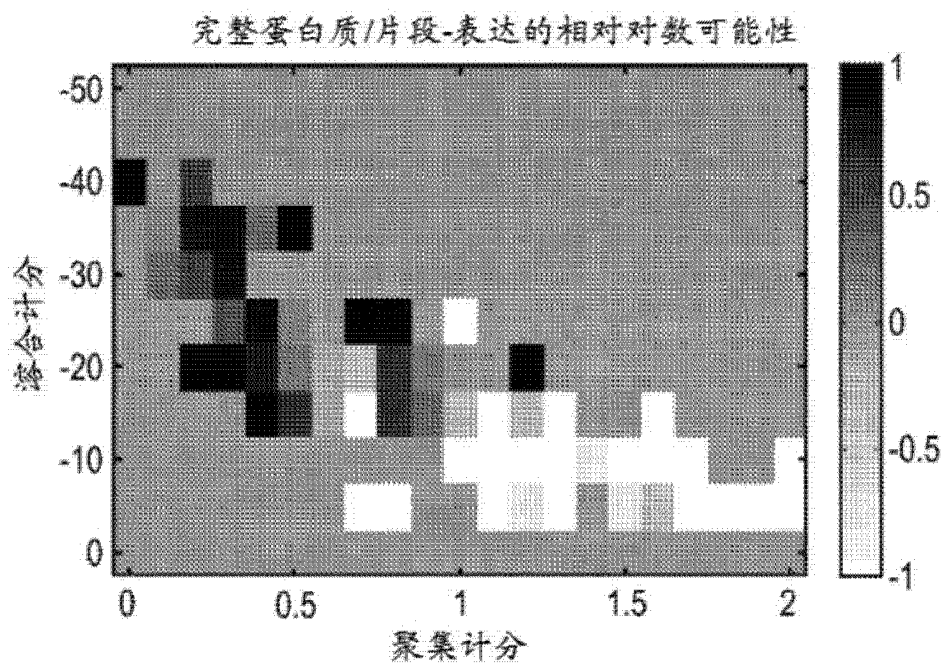


图 1

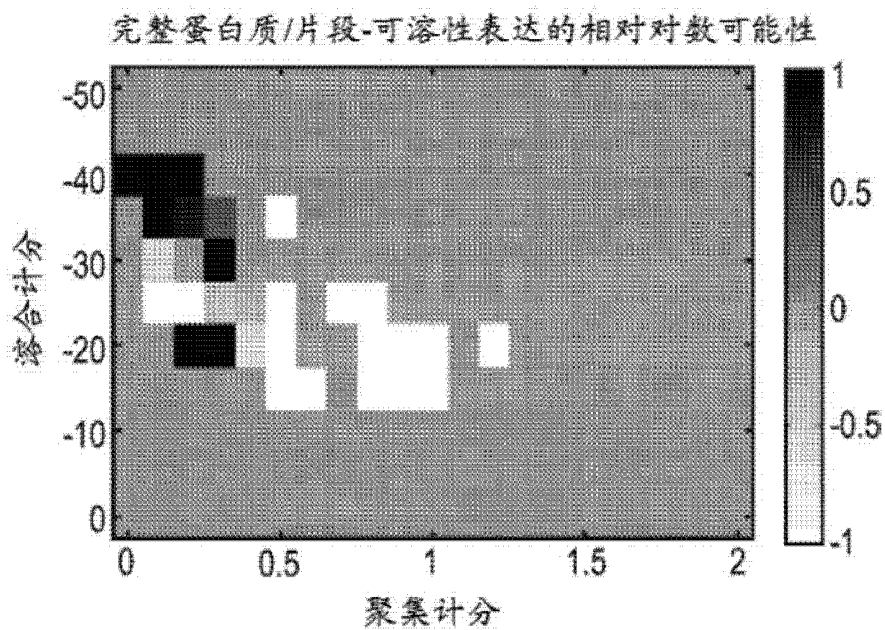


图 2

1. 一种包含含有与天然存在的蛋白质的片段同源的第一多肽序列的第一多肽的分离的营养性蛋白质,其中所述片段的长度是至少 25 个氨基酸,其中所述第一多肽序列包含以下至少一者:

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24% ;
  - b. Leu 残基与总氨基酸残基的比率为至少 11% ;和
  - c. 必需氨基酸残基与总氨基酸残基的比率为至少 49% ,
- 并且其中所述第一多肽序列与已知过敏原具有小于 90%整体同源性。

2. 如权利要求 1 所述的分离的营养性蛋白质,其中所述第一多肽序列进一步包含各必需氨基酸的至少一者。

3. 如权利要求 1 或 2 所述的分离的营养性蛋白质,其中所述第一多肽序列包含与天然存在的蛋白质的所述片段的至少 70%同源性。

4. 如权利要求 1 或 2 所述的分离的营养性蛋白质,其中所述第一多肽序列包含与天然存在的蛋白质的所述片段的至少 95%同源性。

5. 如权利要求 1 至 4 中任一项所述的分离的营养性蛋白质,其中天然存在的蛋白质的所述片段包含至少 25 个氨基酸残基。

6. 如权利要求 1 至 4 中任一项所述的分离的营养性蛋白质,其中天然存在的蛋白质的所述片段包含至少 50 个氨基酸残基。

7. 如权利要求 1 至 6 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含:

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24% ;
- b. Leu 残基与总氨基酸残基的比率为至少 11% ;和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少 49% 。

8. 如权利要求 1 至 7 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含天然存在的营养性蛋白质的所述片段。

9. 如权利要求 1 至 7 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列由天然存在的营养性蛋白质的所述片段组成。

10. 如权利要求 1 至 9 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不是过敏原。

11. 如权利要求 1 至 10 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列与已知过敏原具有小于 5、10、15、20、25、30、35、40、45、50、55、60、70、75、80、85 或 90%整体同源性。

12. 如权利要求 1 至 11 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不是毒素。

13. 如权利要求 12 所述的分离的营养性蛋白质,其中所述第一多肽序列与已知毒素具有小于 50%整体同源性。

14. 如权利要求 1 至 13 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有的模拟胃消化半衰期短于乳清。

15. 如权利要求 14 所述的分离的营养性蛋白质,其中所述第一多肽序列具有模拟胃消化半衰期为小于 60 分钟或小于 30 分钟。

16. 如权利要求 15 所述的分离的营养性蛋白质,其中所述第一多肽序列具有模拟胃消化半衰期为小于 10 分钟。

17. 如权利要求 1 至 16 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列在模拟胃液中完全消化。

18. 如权利要求 1 至 17 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。

19. 如权利要求 1 至 18 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含 0 或 1 个半胱氨酸残基。

20. 如权利要求 1 至 19 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含二硫键。

21. 如权利要求 1 至 20 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含 N 连接的糖基化。

22. 如权利要求 1 至 21 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列不包含 O 连接的糖基化。

23. 如权利要求 1 至 22 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列对聚集具有抗性。

24. 如权利要求 1 至 23 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列在 pH 7 下是阴离子型的。

25. 如权利要求 1 至 24 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列在 pH 7 下具有水溶解度为至少 12.5g/L。

26. 如权利要求 1 至 25 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有计算溶合计分为 -20 或更低。

27. 如权利要求 1 至 26 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有计算聚集计分为 0.75 或更低。

28. 如权利要求 1 至 27 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列具有计算聚集计分为 0.5 或更低。

29. 如权利要求 1 至 28 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列包含选自以下的氨基酸序列:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

30. 如权利要求 1 至 29 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列由选自以下的氨基酸序列组成:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

31. 如权利要求 1 至 30 中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列与至少一个选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列至少 40%、45%、50%、55%、

60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或99.5%同源。

32. 如权利要求1至31中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列是所述天然存在的营养性蛋白质的至少25个氨基酸的片段。

33. 如权利要求1至32中任一项所述的分离的营养性蛋白质,其中所述第一多肽序列是所述天然存在的营养性蛋白质的至少50个氨基酸的片段。

34. 如权利要求1至33中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质由所述第一多肽序列组成和/或其中所述营养性蛋白质是除卵清蛋白、卵铁传递蛋白、卵类粘蛋白、肌球蛋白、肌动蛋白、原肌凝蛋白、胶原蛋白、肌钙蛋白、乳清、酪蛋白、 $\alpha 1$ 酪蛋白、 $\alpha 2$ 酪蛋白、 $\beta$ 酪蛋白、 $\kappa$ 酪蛋白、 $\beta$ -乳球蛋白、 $\alpha$ -乳白蛋白、大豆球蛋白、 $\beta$ -伴大豆球蛋白、谷蛋白、谷醇溶蛋白、麦醇溶蛋白、麦谷蛋白、白蛋白、球蛋白、白蛋白、烯醇酶、肌酸激酶、磷酸甘油酸变位酶、磷酸丙糖异构酶、载脂蛋白、葡糖磷酸变位酶、5磷酸甘油酸激酶、甘油-3-磷酸脱氢酶、3-磷酸甘油醛脱氢酶、血红素、丝切蛋白、糖原磷酸化酶、果糖-1,6-二磷酸酶、原肌球蛋白 $\alpha$ 链、酪蛋白激酶、醛缩酶、微管蛋白、波形蛋白、内质蛋白、乳酸脱氢酶、肌动蛋白解聚因子、转甲状腺素蛋白、双磷酸果糖醛缩酶、碳酸酐酶、醛脱氢酶、膜联蛋白、腺苷高半胱氨酸酶、肌联蛋白、丝切蛋白、磷酸甘油酸激酶、丙酮酸脱氢酶、糖原磷酸化酶、肌激酶、小白蛋白和/或肌间线蛋白以外的营养性蛋白质。

35. 如权利要求1至34中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质进一步包含用于亲和纯化的多肽标签。

36. 如权利要求35所述的分离的营养性蛋白质,其中用于亲和纯化的所述标签是聚组氨酸标签。

37. 一种包含含有与天然存在的蛋白质的片段同源的第一多肽序列的第一多肽的分离的营养性蛋白质,其中所述片段的长度是至少25个氨基酸,其中所述分离的营养性蛋白质包含以下至少一者:

- a. 支链氨基酸残基与总氨基酸残基的比率为至少24%;
- b. Leu残基与总氨基酸残基的比率为至少11%;和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少49%,

且其中所述第一多肽序列与已知过敏原具有小于90%整体同源性。

38. 如权利要求37所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质进一步包含各必需氨基酸的至少一者。

39. 如权利要求37或38所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含与天然存在的蛋白质的所述片段的至少70%同源性。

40. 如权利要求37或38所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含与天然存在的蛋白质的所述片段的至少95%同源性。

41. 如权利要求37至40中任一项所述的分离的营养性蛋白质,其中天然存在的蛋白质的所述片段包含至少25个氨基酸残基。

42. 如权利要求37至40中任一项所述的分离的营养性蛋白质,其中天然存在的蛋白质的所述片段包含至少50个氨基酸残基。

43. 如权利要求37至42中任一项所述的分离的营养性蛋白质,其中所述分离的营养性

蛋白质包含：

- a. 支链氨基酸残基与总氨基酸残基的比率为至少 24%；
- b. Leu 残基与总氨基酸残基的比率为至少 11%；和
- c. 必需氨基酸残基与总氨基酸残基的比率为至少 49%。

44. 如权利要求 37 至 43 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质包含天然存在的营养性蛋白质的片段。

45. 如权利要求 37 至 43 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质由天然存在的营养性蛋白质的片段组成。

46. 如权利要求 37 至 45 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质不是过敏原。

47. 如权利要求 37 至 46 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质与已知过敏原具有小于 5、10、15、20、25、30、35、40、45、50、55、60、70、75、80、85 或 90% 整体同源性。

48. 如权利要求 37 至 47 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质不是毒素。

49. 如权利要求 48 所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质与已知毒素具有小于 50% 整体同源性。

50. 如权利要求 37 至 49 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质具有的模拟胃消化半衰期短于乳清。

51. 如权利要求 50 所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质具有模拟胃消化半衰期为小于 60 分钟或小于 30 分钟。

52. 如权利要求 51 所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质具有模拟胃消化半衰期为小于 10 分钟。

53. 如权利要求 37 至 52 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质在模拟胃液中完全消化。

54. 如权利要求 37 至 53 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质包含至少一个选自胃蛋白酶识别位点、胰蛋白酶识别位点和胰凝乳蛋白酶识别位点的蛋白酶识别位点。

55. 如权利要求 37 至 54 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质包含 0 或 1 个半胱氨酸残基。

56. 如权利要求 37 至 55 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质不包含二硫键。

57. 如权利要求 37 至 56 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质不包含 N 连接的糖基化。

58. 如权利要求 37 至 57 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质不包含 O 连接的糖基化。

59. 如权利要求 37 至 58 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性蛋白质对聚集具有抗性。

60. 如权利要求 37 至 59 中任一项所述的分离的营养性蛋白质，其中所述分离的营养性

蛋白质在 pH 7 下是阴离子型的。

61. 如权利要求 37 至 60 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质在 pH 7 下具有水溶解度为至少 12.5g/L。

62. 如权利要求 37 至 61 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有计算溶合计分为 -20 或更低。

63. 如权利要求 37 至 62 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有计算聚集计分为 0.75 或更低。

64. 如权利要求 37 至 63 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质具有计算聚集计分为 0.5 或更低。

65. 如权利要求 37 至 64 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质包含选自以下的氨基酸序列:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

66. 如权利要求 37 至 65 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质由选自以下的氨基酸序列组成:

- i. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列;
- ii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的修饰的衍生物;和
- iii. 选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列的突变蛋白。

67. 如权利要求 37 至 66 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质与至少一个选自 SEQ ID NO:1 至 SEQ ID NO:2609 的氨基酸序列至少 40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99% 或 99.5% 同源。

68. 如权利要求 37 至 67 中任一项所述的分离的营养性蛋白质,其中所述分离的营养性蛋白质进一步包含用于亲和纯化的多肽标签。

69. 如权利要求 68 所述的分离的营养性蛋白质,其中用于亲和纯化的所述标签是聚组氨酸标签。

70. 一种分离的核酸,其包含编码根据权利要求 1 至 69 中任一项所述的营养性蛋白质的核酸序列。

71. 如权利要求 70 所述的分离的核酸,其中所述分离的核酸选自基因组 DNA、cDNA、正义 RNA 和反义 RNA。

72. 如权利要求 71 所述的分离的核酸,其中所述分离的核酸是基因组 DNA。

73. 如权利要求 71 所述的分离的核酸,其中所述分离的核酸是 cDNA。

74. 如权利要求 70 至 73 中任一项所述的分离的核酸,其进一步包含可操作地连接于编码所述营养性蛋白质的所述核酸序列的表达控制序列。

75. 一种载体,其包含编码如权利要求 1 至 69 中任一项所述的营养性蛋白质的核酸序列。

76. 如权利要求 75 所述的载体,其进一步包含可操作地连接于编码所述营养性蛋白质的所述核酸序列的表达控制序列。

77. 一种重组微生物,其包含如权利要求 70 至 74 中任一项所述的核酸和如权利要求 75 和 76 中任一项所述的载体的至少一者。

78. 根据权利要求 77 所述的重组微生物,其中所述重组微生物是原核生物。

79. 根据权利要求 78 所述的重组原核生物,其中所述原核生物是异养性的。

80. 根据权利要求 78 所述的重组原核生物,其中所述原核生物是自养性的。

81. 根据权利要求 78 所述的重组原核生物,其中所述原核生物是细菌。

82. 一种制备根据权利要求 1 至 69 中任一项所述的营养性蛋白质的方法,所述方法包括在足以由根据权利要求 77 至 81 中任一项所述的重组微生物产生所述营养性蛋白质的条件下培养所述重组微生物。

83. 如权利要求 82 所述的方法,其进一步包括自所述培养物分离所述营养性蛋白质。

84. 一种营养性组合物,其包含如权利要求 1 至 69 中任一项所述的营养性蛋白质和至少一种第二组分。

85. 如权利要求 84 所述的营养性组合物,其中所述至少一种第二组分选自蛋白质、多肽、肽、游离氨基酸、碳水化合物、脂肪、矿物质或矿物质来源、维生素、补充剂、生物体、医药和赋形剂。

86. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是蛋白质。

87. 如权利要求 86 所述的方法,其中所述至少一种第二组分是营养性蛋白质。

88. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是选自必需氨基酸、非必需氨基酸、支链氨基酸、非标准氨基酸和修饰的氨基酸的游离氨基酸。

89. 如权利要求 88 所述的营养性组合物,其中所述至少一种第二组分是选自必需氨基酸的游离氨基酸。

90. 如权利要求 88 所述的营养性组合物,其中所述至少一种第二组分是选自支链氨基酸的游离氨基酸。

91. 如权利要求 88 所述的营养性组合物,其中所述至少一种第二组分是 Leu。

92. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是脂质。

93. 如权利要求 92 所述的营养性组合物,其中所述脂质选自脂肪、油、甘油三酯、胆固醇、磷脂和脂肪酸。

94. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分选自矿物质和维生素。

95. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是补充剂。

96. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是生物体。

97. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是医药。

98. 如权利要求 85 所述的营养性组合物,其中所述至少一种第二组分是赋形剂。

99. 如权利要求 98 所述的营养性组合物,其中所述至少一种赋形剂选自缓冲剂、防腐剂、稳定剂、粘合剂、压实剂、润滑剂、分散增强剂、崩解剂、调味剂、甜味剂、着色剂。

100. 如权利要求 84 至 99 中任一项所述的营养性组合物,其中所述营养性组合物被配制成液体溶液、浆液、混悬液、凝胶剂、糊剂、粉末或固体。

101. 一种制备如权利要求 84 至 100 中任一项所述的营养性组合物的方法,其包括提供根据权利要求 1 至 69 中任一项的营养性蛋白质以及组合所述营养性蛋白质与所述至少一



种第二组分。

102. 一种维持或增加受试者的肌肉质量、肌肉强度和功能性表现的至少一者的方法，所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

103. 一种维持或达成受试者的合乎需要的身体质量指数的方法，所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

104. 如权利要求 102 或 103 所述的方法，其中所述受试者处于以下至少一种状态下：年长、患有医学上危急疾病、以及罹患蛋白质 - 能量营养不良症。

105. 一种向患有蛋白质 - 能量营养不良症的受试者提供蛋白质的方法，所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

106. 一种增加受试者的产热的方法，所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

107. 一种诱导受试者的饱足反应和饱腹感反应的至少一者的方法，所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

108. 如权利要求 106 或 107 所述的方法，其中所述受试者是肥胖的。

109. 如权利要求 102 至 108 中任一项所述的方法，其中如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物是由所述受试者与进行运动配合加以消耗。

110. 一种治疗受试者的恶病质、肌肉减少症和虚弱的至少一者的方法，所述方法包括向所述受试者提供足量的如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物。

111. 如权利要求 102 至 110 中任一项所述的方法，其中如权利要求 1 至 69 中任一项所述的营养性蛋白质、如权利要求 84 至 100 中任一项所述的营养性组合物或通过如权利要求 101 所述的方法制备的营养性组合物是由所述受试者通过经口、肠或胃肠外途径加以消耗。

112. 一种制备如权利要求 1 至 69 中任一项所述的营养性蛋白质的方法，所述方法包括化学合成所述蛋白质。

113. 一种制备如权利要求 1 至 69 中任一项所述的营养性蛋白质的方法，所述方法包括分离所述蛋白质。

114. 如权利要求 1 至 69 中任一项所述的分离的营养性蛋白质，其中所述第一多肽序列的长度是至少 25 个氨基酸。