

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200710186675.X

[51] Int. Cl.

C07K 16/24 (2006.01)
A61K 39/395 (2006.01)
C07H 15/12 (2006.01)
C12N 5/10 (2006.01)
C12N 5/20 (2006.01)
C12N 15/02 (2006.01)

[43] 公开日 2009年1月28日

[11] 公开号 CN 101353381A

[51] Int. Cl. (续)

C12N 15/03 (2006.01)
C12P 21/08 (2006.01)
G01N 33/53 (2006.01)

[22] 申请日 1995.12.22

[21] 申请号 200710186675.X

分案原申请号 95197701.6

[30] 优先权

[32] 1994.12.23 [33] US [31] 08/363131

[32] 1995.6.6 [33] US [31] 08/467420

[32] 1995.6.6 [33] US [31] 08/470110

[71] 申请人 史密丝克莱恩比彻姆公司

地址 美国宾夕法尼亚州

共同申请人 史密丝克莱恩比彻姆有限公司

[72] 发明人 R·S·阿米斯

E·R·阿佩尔鲍姆 I·M·柴肯

R·M·库克 M·S·格罗斯

S·D·霍尔米斯

L·J·麦克米兰 T·W·泰森

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 刘 玥

权利要求书1页 说明书78页 附图13页

[54] 发明名称

用于治疗 IL-5 介导的疾病的重组 IL-5 拮抗剂

[57] 摘要

本发明涉及用于治疗 IL-5 介导的疾病的重组 IL-5 拮抗剂。本发明提供得自高亲和性中和单克隆抗体的, 嵌合的、人源化的和其他的 IL-5 单克隆抗体, 含有这些单克隆抗体的药物组合物, 治疗和诊断方法。

1. 一种人源化中和单克隆抗体，它对人类白介素 5 有特异性，并包含以下互补决定区：

SEQ ID NO: 7 的 CDRH1

SEQ ID NO: 8 的 CDRH2

SEQ ID NO: 9 的 CDRH3

SEQ ID NO: 10 的 CDRL1

SEQ ID NO: 11 的 CDRL2

SEQ ID NO: 12 的 CDRL3。

2. 权利要求 1 的人源化中和单克隆抗体，它对人类白介素 5 有特异性，包含 SEQ ID NO: 19 的重链可变区和 SEQ ID NO: 21 的轻链可变区。

3. 一种药物组合物，它包含权利要求 1 或 2 的人源化抗体和药用载体。

4. 权利要求 1 或 2 的人源化抗体在制备用于治疗与嗜曙红细胞过量产生有关的人类病症的药物中的用途，所述病症选自气喘、过敏性鼻炎、特应性皮炎、慢性嗜曙红性肺炎、过敏性支气管曲霉病、腹腔病、嗜曙红性胃肠炎、Churg-Strauss 综合症、嗜曙红肌痛综合症、嗜曙红细胞过多综合症和水肿反应。

用于治疗 IL-5 介导的疾病的重组 IL-5 拮抗剂

本申请是申请日为 1995 年 12 月 22 日, 申请号为 95197701.6, 发明名称为“用于治疗 IL-5 介导的疾病的重组 IL-5 拮抗剂”的发明专利申请的分案申请。

相关申请

本申请是 1995 年 6 月 6 日提交的美国专利 08/470,110 和 08/467,420 号的部分继续, 这两个专利又是 1994 年 12 月 23 日提交的美国专利 08/363,131 号的部分继续。

发明领域

本发明一般地涉及抗体和改变的抗体(altered antibodies)领域, 用于治疗 and 诊断 IL-5 介导的病症和与嗜曙红细胞过量产生有关的病症, 更具体地说, 涉及 mAb、Fab、嵌合的和人源化的抗体。

发明背景

嗜曙红细胞被认为与许多炎症疾病状态的病因有关, 这些疾病包括与肺组织的过敏反应有关的变态疾病(Butterfield 等, Immunopharmacology of Eosinophils, H. Smith and R. Cook, Eds., p.151-192, Academic Press, London(1993))。一个著名的实例是气喘, 该疾病的特征是空气通路的可逆性阻塞, 导致非特异性的支气管过敏。而这又依赖于支气管粘液水平的慢性炎症反应的生成, 以及巨噬细胞、淋巴细胞和嗜曙红细胞的特征性侵入。看来嗜曙红细胞在起始疾病的典型粘液损伤中起到关键作用(Corrigan 等, Immunol. Today, 13:501-507(1992))。据报道, 在患有慢性气喘的病人的循环系统、支气管分泌物和肺实质中, 活化的嗜曙红细胞的数目增加, 并且通过对不同肺功能的测试测量出, 疾病的严重程度与血液嗜曙红细胞数相关(Griffen 等, J. Aller. Clin. Immunol., 67:548-557(1991))。嗜曙红细胞数目增加, 通常在去粒过程中, 也在进行晚期气喘反应的患者的支气管肺泡洗出液(bronchoalveolar lavage, BAL)中收集到; 嗜曙红细胞减少, 通常是类固醇治疗的结果, 与临床症状的改善有关(Bousquet 等, N. Eng. J. Med., 323:1033-1039(1990))。

白介素 5 (IL-5) 是由活化的 CD4+ T 淋巴细胞产生的同二聚体糖蛋白。在人体中, IL-5 主要负责控制嗜曙红细胞的生长和分化。在气喘患者的支气管泡洗出液中检测到 IL-5 水平升高(Motojima 等, Allergy, 48:98(1993))。在没有抗原的刺激下, IL-5 转基因小鼠的外周血和组织显示出显著的嗜曙红细胞水平(Dent 等, J. Exp. Med.,

172:1425(1990)), 而且抗鼠 IL - 5 单克隆抗体具有在小鼠的血液和组织中减少嗜曙红细胞(Hitoshi 等, Int. Immunol., 3:135(1991)), 以及在实验动物中减少与寄生物感染和过敏原攻击有关的嗜曙红细胞的作用(Coffman 等, Science, 245:308-310(1989), Sher 等, Proc. Natl. Acad. Sci., 83:61-65(1990), Chand 等, Eur. J. Pharmacol., 211:121-123(1992)).

虽然皮质类固醇对抑制嗜曙红细胞数量和气喘的其他炎症成分有显著的作用, 但他们对严重气喘, 以及近来发现的对温和至中等气喘有副作用。仅有的另一种主要抗炎药物治疗 - 色甘酸(色甘酸钠和 nedocromil) - 被认为具有比皮质类固醇少的副作用。它们的精确作用机制仍然未知。

近来的开发工作集中于新的吸入的类固醇、长作用支气管舒张剂和作用于新的生化或药理学靶的药剂(例如钾通道激活剂、白三烯拮抗剂、5 - 脂加氧酶(5-LO)抑制剂等)。一种理想的药物应当将类固醇的有效性和色甘酸钠的安全性结合起来, 并具有较高的选择性和快速作用的能力。中和 IL - 5 抗体可能对治疗人的与嗜曙红细胞有关的疾病有用。

因此, 本领域中需要一种高亲和性的 IL - 5 拮抗剂, 如人 IL - 5 的中和单克隆抗体, 它能减少嗜曙红细胞的分化和增殖(即嗜曙红细胞的积累), 由此减少嗜曙红细胞炎症。

发明概述

在第一方面, 本发明提供啮齿动物(例如大鼠和鼠)中和单克隆抗体, 该抗体对人 IL - 5 有特异性, 并具有一定的结合亲和性, 该亲和性的特征在于, 解离常数等于或小于大约 3.5×10^{-11} M, 如发明详述中所述。这些单克隆抗体的实例是鼠单克隆抗体 2B6、2E6 和 2F2, 以及大鼠单克隆抗体, 如 4A6。本发明的另一方面是一些杂交瘤, 例如 SK119-2B6.206.75(1)、SK119-2E3.39.40.2、SK119-2F2.37.80.12、4A6(1)G1F7 和 5D3(1)F5D6。

在相关方面, 本发明提供中和 Fab 片段或通过去除本发明的啮齿动物中和单克隆抗体的 Fc 区得到的, 对人 IL - 5 有特异性的 F(ab')₂ 片段。

在另一个相关方面, 本发明提供中和 Fab 片段或通过链改组(shuffling)技术得到的, 对人 IL - 5 有特异性的 F(ab')₂ 片段, 由此, 分离自本发明啮齿动物中和单克隆抗体的重(或轻)链免疫球蛋白, 被轻链(或相应的重链)免疫球蛋白文库表达, 该文库用 IL - 5 免疫的啮齿动物, 存在与纤毛噬菌体 Fab 显示文库中。

在另一个相关方面，本发明提供对人 IL - 5 有特异性的改变的抗体，它包括来源于非人类中和单克隆抗体(mAb)的互补决定区(CDR)，非人类中和单克隆抗体的特征在于，人类 IL - 5 和编码它的核酸分子的解离常数等于或小于大约 $3.5 \times 10^{-11}M$ 。当改变的抗体是人源化的抗体时，来源于非人类免疫球蛋白的，编码互补决定区的序列被插入第一免疫球蛋白配对物(partner)，在该配对物中，第一免疫球蛋白配对物的至少一个，优选全部的互补决定区(CDR)被来源于非人类单克隆抗体的 CDR 置换。优选地，第一免疫球蛋白配对物也与第二免疫球蛋白配对物可操作地连接(operatively linked)，第二免疫球蛋白配对物包括免疫球蛋白恒定区的全部或部分。

在相关方面，本发明提供来源于非人类中和单克隆抗体(mAb)的 CDR，其特征在于，人类 IL - 5 和编码所述 CDR 的核酸分子的解离常数等于或小于大约 $3.5 \times 10^{-11}M$ 。

在另一个相关方面，本发明提供一种嵌合抗体，该抗体含有来源于非人类中和单克隆抗体的人类重链和轻链恒定区以及重链和轻链可变区，其特征在于，对人类 IL - 5 的解离常数等于或小于大约 $3.5 \times 10^{-11}M$ 。

在另一个相关方面，本发明提供一种药物组合物，它含有一种(或多种)上面描述的改变的抗体和药用载体。

在另一方面，本发明提供一种治疗与嗜曙红细胞过量产生有关的人类疾病的方法，该方法是提供对人施用有效量的本发明药物组合物来完成的。

在另一方面，本发明提供改变的抗体(例如工程化的抗体、CDR、Fab 或 F(ab)₂ 片段，或其类似物)的重组制备方法和有用的成分，该抗体来源于非人类中和单克隆抗体(mAb)，其特征在于，对人类 IL - 5 的解离常数等于或小于大约 $3.5 \times 10^{-11}M$ 。这些成分包括编码它们的分离的核酸序列，含有在特定调节序列控制下的核酸序列的重组质粒，所述调节序列能够指导核酸序列在被重组质粒转染的宿主细胞(优选哺乳动物)中表达。该制备方法包括在改变的抗体(优选人源化抗体)可以在所述细胞中表达的条件下培养本发明的被转染的宿主细胞系，并从中分离表达产物。

在另一方面，本发明提供诊断与嗜曙红细胞在人体中过量产生有关的疾病的方法，其包括从患者体内获得生物液样本，使本发明的抗体和改变的抗体与该样本接触，所用条件使得 IL - 5/(单克隆或改变的)抗体复合物得以形成，检测 IL - 5/抗体复合物存在与否。

本发明的其他方面和优点在发明详述和优选实施方案中进一步描述。

附图简述

图 1[SEQ ID NO: 1 和 15]表明鼠抗体 2B6 和鼠/人抗体 2B6 嵌合抗体的重链可变区。方框区指示 CDR。

图 2[SEQ ID NO: 2 和 16]表明鼠和鼠/人抗体 2B6 嵌合抗体的轻链可变区。方框区指示 CDR。

图 3[SEQ ID NO: 3]表明鼠抗体 2F2 的重链可变区。方框区指示 CDR。

图 4[SEQ ID NO: 4]表明鼠抗体 2F2 的轻链可变区。方框区指示 CDR。

图 5[SEQ ID NO: 5]表明鼠抗体 2E3 的重链可变区。方框区指示 CDR。

图 6[SEQ ID NO: 6]表明鼠抗体 2E3 的轻链可变区。方框区指示 CDR。

图 7[SEQ ID NO: 7-14]表明鼠抗体 2B6、2F2 和 2E3 的重链和轻链可变区。

图 8[SEQ ID NO: 18,19]表明人源化的抗体 2B6 的重链可变区。方框区指示 CDR。

图 9[SEQ ID NO: 20,21]表明人源化的抗体 2B6 的轻链可变区。方框区指示 CDR。

图 10 是质粒 pCDIL5HZHC1.0 的示意图, 该质粒用于在哺乳动物细胞中表达人源化的重链基因。该质粒含有 β 乳糖酶基因 (BETA LAC)、SV - 40 复制起始区 (SV40)、巨细胞病毒启动子序列 (CMV)、一段信号序列、人源化的重链、来自牛生长激素 (BGH) 的 poly A 信号、 β 珠蛋白启动子 (beta glopro)、二氢叶酸还原酶基因 (DHFR) 和另一种 BGH 序列 poly A 信号, 它们位于 pUC19 背景中。

图 11 是质粒 pCNIL5HZHC1.0 的示意图, 该质粒用于在哺乳动物细胞中表达人源化的轻链基因。

图 12[SEQ ID NO:61,62]表明人源化的抗体 2B6 的 NewM 重链可变区。方框区指示 CDR。

图 13[SEQ ID NO:69,70]表明人源化的抗体 2B6 的 REI 轻链可变区。方框区指示 CDR。

发明详述

本发明提供多种抗体、改变的抗体及其片段, 其特征在于人类 IL - 5 结合的特异性、中和活性和对人类 IL - 5 的高亲和性, 这些特征由鼠单克隆抗体 2B6 例示。本发明的抗体可通过常规杂交瘤技术、噬菌体显示组合文库、免疫球蛋白链改组以及人源化技术制备, 以得到新的中和抗体。这

些产物可用于治疗 IL - 5 介导的疾病（如气喘）的治疗和药物组合物。这些产物也可用于通过测量（例如通过酶联免疫检测(ELISA)）人体内源 IL - 5 水平或由激活的细胞间接体内(ex vivo)释放的 IL - 5，来诊断 IL - 5 介导的病症。

1. 定义

“改变的抗体”指由改变的免疫球蛋白编码区编码的蛋白，它可以通过在特定宿主细胞中表达得到。这些改变的抗体是工程化的抗体（例如嵌合抗体或人源化的抗体）或抗体片段，该抗体片段缺少免疫球蛋白恒定区（例如 Fv、Fab 或 F(ab)₂ 等）的全部或部分。

“改变的免疫球蛋白编码区”指编码本发明的改变的抗体的核酸序列。当改变的抗体是移植的 CDR 或人源化的抗体时，来源于非人类免疫球蛋白的互补编码决定区（CDR）的序列被插入第一免疫球蛋白配对物中，该配对物含有人类可变框架序列。任选地，第一免疫球蛋白配对物与第二免疫球蛋白配对物可操作地连接。

“第一免疫球蛋白配对物”指编码人类框架或人类免疫球蛋白可变区的核酸序列，其中天然（或天然存在的）CDR 编码区被供体抗体的 CDR 编码区置换。人类可变区可以是免疫球蛋白重链、轻链（或两种链同时存在）、其类似物或功能性片段。这些 CDR 区位于抗体（免疫球蛋白）的可变区内，可以用本领域已知的方法确定。例如，Kabat 等(Sequences of Proteins of Immunological Interest, 4th Ed., U.S.Department of Health and Human Services, National Institutes of Health(1987))公开了 CDR 的定位方法。此外，可用于鉴定 CDR 区/结构的计算机程序也是已知的。

“中和”指一种抗体通过阻止人类 IL - 5 与其特异性受体结合，或抑制 IL - 5 通过其受体的信号来抑制 IL - 5 活性，应使结合发生。如果在 B13 细胞生物检测中（IL - 5 中和检测，见实施例 2C），一种 mAb 对抑制 IL - 5 活性是 90 % 有效的，优选 95 % 有效的，最有效 100 % 有效的，则该 mAb 是中和性的。

术语“高亲和性”指抗体具有一种结合亲和性，其特征在于，在光学生物传感器分析中（见实施例 2D），对人类 IL - 5 的 Kd 等于或小于 3.5×10^{-11} M。

“对人类 IL - 5 结合的特异性”指对人而不是鼠的 IL - 5 的高亲和性。

“第二免疫球蛋白配对物”指另一编码蛋白或肽的核酸序列，第一免疫

球蛋白配对物在框架内或以任意合适的连接序列（即可操作连接）与之融合。优选地，它是免疫球蛋白基因。第二免疫球蛋白配对物可包括一段核酸序列，该序列编码同样（即同源的 - 第一和第二改变的可以来自同一来源）或另一（即异源的）感兴趣的抗体的全部恒定区。它可以是免疫球蛋白的重链或轻链（或同时具有两种链，它们作为单一多肽的一部分）。第二免疫球蛋白配对物不限于具体的免疫球蛋白类型或同种型。此外，第二免疫球蛋白配对物可包含免疫球蛋白的恒定区的一部分，例如 Fab 或 F(ab)₂ 的一部分（即合适的人类恒定区或框架区的离散部分）。这样的第二免疫球蛋白配对物可包含编码整合膜蛋白的序列，该蛋白暴露于宿主细胞的外表面，例如噬菌体显示文库的一部分，或编码用于分析性或诊断性检测的蛋白（例如辣根过氧化物酶、β 半乳糖苷酶等）。

术语 Fv、Fc、Fd、Fab 或 F(ab)₂ 以它们的标准含义使用（见例如，Harlow 等，Antibodies A Laboratory Manual, Cold Spring Harbor Laboratory(1988)）。

本文所用的术语“工程化的抗体”描述一种改变的抗体，即全长合成的抗体（例如针对一个抗体片段的嵌合的或人源化的抗体），其中特定受体抗体的轻链和/或重链可变结构域的一部分被来自一个或多个供体抗体的类似部分替换，所述供体抗体对特定表位具有特异性。例如，这类分子可包括这样一些抗体，其特征在于，人源化的重链与未修饰的轻链（或嵌合的轻链）相连，反之亦然。工程化的抗体的特征也可以是编码受体抗体轻链和/或重链可变结构域框架区(framework regions)的核酸序列发生改变，使之保留供体抗体结合特异性。这些抗体可包含用来自本文描述的供体抗体的 CDR 替换来自受体抗体的一个或多个 CDR。

“嵌合抗体”指一类工程化的抗体，其包含来源于供体抗体的天然存在的可变区（轻链和重链），以及来源于受体抗体的轻链和重链恒定区。

“人源化抗体”指一类工程化的抗体，其 CDR 来源于非人类供体免疫球蛋白，分子中残余的来源于免疫球蛋白的部分来源于一种（或多种）人类免疫球蛋白。此外，框架支持残基可被改变以维持结合亲和性（见例如，Queen 等，Proc. Natl Acad Sci USA, 86:10029-10032(1989), Hodgson 等，Bio/Technology, 9:421(1991)）。

术语“供体抗体”指这样一种（单克隆或多克隆）抗体，它将其可变区、CDR 或其他功能片段或其类似物提供给第一免疫球蛋白，以提供改变的免

疫球蛋白编码区和所得的表达的抗体，该抗体具有供体抗体的抗原特异性或中和活性特征。一种适用于本发明的供体抗体是命名为 2B6 的非人类（即鼠的）中和单克隆抗体。抗体 2B6 被限定为高亲和性的、人类 IL - 5 特异性的（即不识别鼠 IL - 5）同种型 IgG₁ 的中和抗体，它在合适的鼠 IgG 恒定区内，分别具有 SEQ ID NO:2 和 16 的可变轻链 DNA 和氨基酸序列，以及 SEQ ID NO:1 和 15 的可变重链 DNA 和氨基酸序列。

术语“受体抗体”指与供体抗体异源的（单克隆或多克隆）抗体，它将编码其重链和/或轻链框架区和/或其重链和/或轻链恒定区全部（或任意部分，但优选全部）的核酸序列贡献给第一免疫球蛋白配对物。优选地，人类抗体是该受体抗体。

“CDR”被限定为一种抗体的互补决定区氨基酸序列，它是免疫球蛋白重链和轻链的超可变区。见例如，Kabat 等，Sequences of Proteins of Immunological Interest, 4th Ed., U.S. Department of Health and Human Services, National Institutes of Health(1987)。免疫球蛋白的可变部分中有三种重链 CDR 和三种轻链 CDR（或 CDR 区）。因此，本文所用“CDR”指全部三种重链 CDR，或全部三种轻链 CDR（或全部重链和轻链 CDR，若合适的话）。

在抗体与抗原或表位结合时，CDR 提供主要的接触残基。本发明感兴趣的 CDR 来源于供体抗体可变重链和轻链序列，包括天然存在的 CDR 的类似物，该类似物分享或维持与和它来源相同的供体抗体同样的抗原结合特异性和/或中和能力。

“分享抗原结合特异性或中和能力”指，例如，虽然 mAb 可能具有某种程度的抗原亲和性，但被 2B6 核酸序列编码的 CDR 在适当的结构环境中可能具有较低的或较高的亲和性。人们希望 2B6 的 CDR 在这样的环境中与 2B6 识别相同的表位。2B6 重链 CDR 的实例包括 SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9，2B6 轻链 CDR 的实例包括 SEQ ID NO:10、SEQ ID NO:11 和 SEQ ID NO:12。

“功能性片段”是重链或轻链可变序列的一部分（例如免疫球蛋白可变区的氨基或羧基末端的小的缺失），它保留了与产生该片段的抗体相同的抗原结合特异性和/或中和能力。

“类似物”是至少一个氨基酸被修饰的氨基酸序列，其中所述修饰可以是化学性的，或是少量氨基酸（即不超过 10 个）的替换或重排，该修饰使

得氨基酸序列保留未修饰的序列的生物学特性，例如抗原特异性和高亲和性。例如，当在 CDR 编码区内或周围产生了某些内切核酸酶限制性位点时，可以通过替换构建（不活动）突变。

类似物也可由等位变异产生。“等位变异或修饰”是编码本发明的氨基酸或肽序列的核酸序列中的改变。这种变异或修饰可来自于遗传密码的简并性，或被精心地工程化，以提供所需的特性。这些变异或修饰可以或不导致任何所编码的氨基酸序列发生改变。

术语“效应子试剂”指非蛋白载体分子，改变的抗体和/或供体抗体的天然或合成的轻链或重链，或供体抗体的其他片段可通过常规方式与之相连。这样的非蛋白载体可包括诊断领域所用的常规载体，例如聚苯乙烯或其他塑料珠，多糖，如 BIAcore[Pharmacia]系统中所用的，或其他可用于医药领域的非蛋白物质，它们对人和动物施用是安全的。其他效应子试剂可包括用于整合重金属原子或放射性同位素的大环。这些效应子试剂也可用于提高改变的抗体的半衰期，例如聚乙二醇。

II. 高亲和性 IL - 5 单克隆抗体

为了用于构建本发明的抗体、改变的抗体和片段，可用非人类的物种（例如牛、羊、猴、鸡、啮齿动物（例如鼠和大鼠）等）来产生所需的免疫球蛋白，该免疫球蛋白呈递天然的人类 IL - 5 或其它肽表位。常规的杂交瘤技术可用于提供分泌非人类抗 IL - 5 单克隆抗体的杂交瘤细胞系。然后用 IL - 5 包被的 96 孔板筛选结合的杂交瘤，如实施例部分所述，或可选地用与链亲和素包被的平板结合的生物素化的 IL - 5。

本发明的一个示例性的，高亲和性的，中和 mAb 是 mAb 2B6，它是一种鼠类抗体，可用于开发嵌合的、人源化的抗体，在下面的实施例 1 中有进一步的描述。2B6 单克隆抗体的特征在于特异性结合人类 IL - 5 的抗原，对 IL - 5 的 Kd 值小于 $3.5 \times 10^{-11} \text{M}$ （大约 $2.2 \times 10^{-11} \text{M}$ ）。通过光学生物传感器确定的 2B6 的 Fab 片段对 IL - 5 的 Kd 值是大约 $9 \times 10^{-11} \text{M}$ 。mAb 2B6 看来阻遏人类 IL - 5 和人类 IL - 5 受体 α 链之间的结合反应。

另一种所需的供体抗体是鼠类 mAb - 2E3。这种 mAb 是同种型 IgG2a，对 hIL - 5 的解离常数小于 $3.5 \times 10^{-11} \text{M}$ （大约 $2.0 \times 10^{-11} \text{M}$ ）。

另一种所需的供体抗体是大鼠 mAb - 4A6。这种 mAb 对 hIL - 5 的解离常数小于 $3.5 \times 10^{-11} \text{M}$ （大约 $1.8 \times 10^{-11} \text{M}$ ）。此外，mAb 4A6 看

来阻遏人类 IL - 5 和人类 IL - 5 受体 β 链之间的结合反应。

本发明不限于应用 2B6 mAb、2E3 mAb 或其超可变（即 CDR）序列。任何其他对人类 IL - 5 的解离常数小于或等于 3.5×10^{-11} M 的高亲和性 IL - 5 抗体和相应的抗 IL - 5 CDR

可替换它们。下面说明书中的供体抗体被限定为 2B6 或 2E3，这个名称仅是用于说明并使说明书简化。

III. 抗体片段

本发明也包括 Fab 片段或 F(ab')₂ 片段，这些片段来源于抗人类 IL - 5 的 mAb。这些片段可用作体内抗 IL - 5 和嗜曙红细胞介导的病症的药剂，或作为 IL - 5 体外诊断的一部分。Fab 片段包含轻链的全部和重链的氨基末端部分；F(ab')₂ 片段是由两个 Fab 片段通过二硫键结合形成的片段。单克隆抗体 2B6、2E3 和其他类似的高亲和性的 IL - 5 结合抗体，提供了 Fab 片段和 F(ab')₂ 片段来源，这些片段可以通过常规方式得到，例如用合适的蛋白水解酶（木瓜蛋白酶或胃蛋白酶）切割 mAb，或用重组方法得到。这些 Fab 和 F(ab')₂ 片段本身可用作治疗、预防或诊断药剂，以及用作序列供体，所述序列包括可变区和 CDR 序列，它们可用于形成本文描述的重组或人源化的抗体。

Fab 和 F(ab')₂ 片段可通过重组噬菌体文库构建（见例如 Winter 等，Ann. Rev. Immunol., 12:433-455(1994)）或免疫球蛋白链改组技术（见例如 Marks 等，Bio/Technology, 10:779-783(1992)，两篇文献均完整地在此引入作为参考）构建，其中来自选定抗体的 Fd 或 v_H 免疫球蛋白（例如 2B6）被允许与轻链球蛋白 v_L（或 v_K）的所有组成部分结合，以形成新的 Fab。相反，来自选定抗体的轻链可被允许与重链免疫球蛋白 v_H（或 Fd）的所有组成部分结合，以形成新的 Fab。当 mAb 2B6 的 Fd 被允许与轻链免疫球蛋白的所有组成部分结合时，得到中和 IL - 5 Fab，如实施例部分所详述的。因此，人们可由链改组技术用单一序列（核苷酸或氨基酸）回收中和 Fab。

IV. 感兴趣的抗 IL - 5 氨基酸和核苷酸序列

上述 mAb 2B6 或其他抗体可提供序列，如可变重链和/或轻链肽序列、框架序列、CDR 序列、功能片段及其类似物，以及编码它们的核酸序列，可用于设计和得到不同的改变的抗体，这些抗体对供体抗体具有抗原特异性。

作为一个实例，本发明提供可变轻链和可变重链序列，这些序列来自 IL

- 5 鼠抗体 2B6 及其衍生的序列。2B6 的重链可变区由图 1 给出。CDR 编码区用方框区域标出,并由 SEQ ID NO:7; SEQ ID NO:8 和 SEQ ID NO:9 给出。2B6 的轻链克隆可变区由图 2 给出。CDR 编码区由 SEQ ID NO:10; SEQ ID NO:11 和 SEQ ID NO:12 给出。

人源化的重链可变区由图 8 给出 [SEQ ID NO:18 和 19]。信号序列也由 SEQ ID NO:17 给出。本领域技术人员已知的其他合适的信号序列,可以替换本实例中的信号序列。此构建体的 CDR 氨基酸序列与天然鼠和嵌合重链 CDR 的相同,由 SEQ ID NO:7、SEQ ID NO:8 和 SEQ ID NO:9 给出。人源化的轻链可变序列的一个实例(合成的)由图 9 给出 [SEQ ID NO:20 和 21]。

本发明的编码可变轻链和重链肽序列的核酸序列或其片段,也可用于在编码 CDR 或框架区的核酸序列内诱变产生特定的变化,并将所得的被修饰的或融合的核酸序列掺入表达质粒中。例如,框架和核酸序列和 CDR 编码区内的不活动替换被用于形成限制酶位点,这些位点有助于诱变的 CDR (和/或框架)区的插入。这些 CDR 编码区被用于构建本发明的人源化的抗体。

考虑到遗传密码的简并性,可以构建不同的编码序列,这些序列编码可变重链和轻链氨基酸序列,本发明的 CDR 序列及其功能性片段或类似物,它们也具有供体抗体的抗原特异性。本发明的编码可变轻链肽序列或 CDR 的分离的核酸序列或其片段,当与第二免疫球蛋白配对物可操作地结合时,可用于产生改变的抗体,例如嵌合的或人源化的抗体,或其他工程化的抗体。

应当注意到,除了本文描述的编码改变的抗体部分的分离的核酸序列外,本发明也包括其他这些的核酸序列,例如与天然 CDR 编码序列互补的,或与位于 CDR 编码区周围的被修饰的人类框架区互补的序列。有用的 DNA 序列包括这样一些序列,它们在严谨杂交条件下与 DNA 序列杂交(见 T. Maniatis 等,分子克隆(实验指南),冷泉港实验室(1982),387-389 页)。严谨杂交条件的一个实例是用 $4 \times \text{SSC}$ 在 65°C 下杂交,然后用 $0.1 \times \text{SSC}$ 在 65°C 下清洗一个小时。另一严谨杂交条件是 50% 甲酰胺, $4 \times \text{SSC}$, 42°C 。优选地,这些杂交的 DNA 序列的长度至少是大约 18 个核苷酸,即大约 CDR 的大小。

V. 改变的免疫球蛋白分子和改变的抗体

改变的免疫球蛋白分子可编码改变的抗体，该抗体包括工程化的抗体，如嵌合抗体和人源化的抗体。所需的改变的免疫球蛋白编码区包含 CDR 编码区，CDR 编码区编码对 IL - 5 抗体有抗原特异性的肽，优选的是本发明提供的高亲和性抗体，它被插入第一免疫球蛋白配对物（人类框架或人类免疫球蛋白可变区）中。

优选地，第一免疫球蛋白配对物与第二免疫球蛋白配对物可操作地相连。第二免疫球蛋白配对物由上文限定，可包括编码感兴趣的第二抗体区域（例如 Fc 区）的序列。第二免疫球蛋白配对物也可包括编码另一免疫球蛋白的序列，所述另一免疫球蛋白与轻链或重链恒定区在框架内或以连接物序列的方式融合。针对 IL - 5 的功能片段或类似物的工程化抗体可被设计得与同一抗体有更强的结合能力。

第二免疫球蛋白配对物也可与上文限定的效应物试剂（包括非蛋白载体分子）结合，第二免疫球蛋白配对物可以以常规方式与之可操作地连接。

在第二免疫球蛋白配对物（例如抗体序列）和效应物试剂之间的融合和连接可通过任何合适的方式进行，例如通过常规的共价或离子键，蛋白融合，或异双功能交联，如碳化二亚胺，戊二醛等。这些技术是本领域已知的，在常规的化学和生物化学课本中已有描述。

此外，常规的连接序列也可被构建到改变的免疫球蛋白编码区内，这些连接序列仅在第二免疫球蛋白配对物和效应物试剂之间提供所需数量的空间。这样的连接物的设计是本领域技术人员充分了解的。

此外，本发明的分子的信号序列可被修饰，以增强表达。作为一个实例，具有信号序列和来源于鼠重链序列的 CDR 的 2B6 人源化抗体，其初始的信号肽被另一信号序列[SEQ ID NO:17]替换。

一种示例性的改变的抗体含有可变重链和/或轻链肽，或含有对 mAb 2B6 具有抗原特异性的蛋白序列，例如 V_H 和 V_L 链。本发明另一种所需的改变的抗体的特征在于这样的氨基酸序列或其功能片段或类似物，它包含鼠抗体分子 2B6 的重链和/或轻链可变区的至少一个，优选全部的 CDR，残余序列来源于人。见例如人源化的 V_H 和 V_L 区（图 8 和 9）。

在另一个实施方案中，本发明的工程化的抗体可与另一试剂相连。例如，重组 DNA 技术的方法可用于产生本发明的工程化抗体，其中完整抗体分子的 Fc 片段或 CH₂CH₃ 结构域被一种酶或其他可检测分子（即多肽效应物或报道分子）替换。

第二免疫球蛋白配对物也可与非免疫球蛋白肽、蛋白或其片段可操作地连接，这些非免疫球蛋白肽、蛋白或其片段是与对鼠 2B6 有抗原特异性的含 CDR 的序列异源的。表达后所得的蛋白同时具有抗 IL - 5 抗原特异性和非免疫球蛋白的特性。融合配对物特性可以是例如功能特性，如其他结合或受体结构域；或治疗特性，如果融合配对物本身是治疗性蛋白；或其他抗原特性。

本发明另一种所需的蛋白可包括具有全长重链和轻链的完整抗体分子，或其任何离散的片段，如 Fab 或 F(ab')₂ 片段，重链二聚体，或其任何最小重组片段，例如 Fv 或单链抗体（SCA）或任何其他具有与选定供体单克隆抗体（例如 mAb 2B6 或 2E3）相同的特异性的分子。这样的蛋白可以以改变的抗体形式应用，或以其非融合形式应用。

不论第二免疫球蛋白配对物是否来源于不同于供体抗体的抗体（例如免疫球蛋白框架或恒定区的任何同种型或类型），都可得到工程化抗体。工程化抗体可包含来自例如受体抗体的免疫球蛋白(Ig)恒定区和可变框架区，和来自供体抗体（如本文描述的抗 IL - 5 抗体）的一个或多个（优选所有）CDR。此外，可对受体 mAb 轻链和/或重链可变结构域框架区在核酸或氨基酸水平上进行改变，例如缺失、替换或添加，或改变供体 CDR 区，以便保留供体抗体的抗原结合特异性。

这些工程化的抗体被设计为具有 IL - 5 mAb（如上述任意地修饰）可变重链和/或轻链中的一个（或两个），或下面确定的重链或轻链 CDR（见图 7）中的一个或多个。本发明的工程化抗体被中和，即它们按需要阻遏与 IL - 5 蛋白受体的结合，它们也阻遏或防止 IL - 5 依赖型细胞的增殖。

这些工程化抗体可包括人源化抗体，该人源化抗体含有选定的人类免疫球蛋白或亚型的框架区，或是包含与 IL - 5 抗体功能片段融合的人类重链和轻链恒定区的嵌合抗体。合适的人类（或其他动物）受体抗体可根据与供体抗体的核苷酸和氨基酸序列的同源性选自常规的数据库，例如 KABAT®数据库、Los Alamos 数据库和 Swiss 蛋白数据库。与供体抗体的框架区同源的（基于氨基酸）的人类抗体可能适于提供用于插入供体 CDR 的重链恒定区和/或重链可变框架区。可以用类似方式选择能提供轻链恒定区或可变框架区的合适的受体抗体。应注意到受体抗体的重链和轻链对由同一受体抗体起始而言是不需要的。

异源框架和恒定区最好选自人类免疫球蛋白及其同种型，如 IgG（亚型

1至4), IgM, IgA 和 IgE。但是, 受体抗体不必仅包含人类免疫球蛋白序列。例如可以这样构建基因, 使得其中编码部分人类免疫球蛋白链的 DNA 序列与编码非免疫球蛋白氨基酸序列(例如多肽效应子或受体分子)的 DNA 分子融合。

特定的所需人源化抗体的一个实例包含 2B6 的 CDR, 它被插入选定的人类抗体序列的框架区。对于中和人类抗体, 一个, 两个或优选三个来自 IL - 5 抗体重链和/或轻链可变区的 CDR 被插入选定的人类抗体序列的框架区, 替换了人类抗体的天然 CDR。

优选地, 在人源化的抗体中, 人类重链和轻链的可变结构域通过一个或多个 CDR 替换而被工程化。可以使用全部六个 CDR, 或使用少于六个的 CDR 的不同组合。优选全部六个 CDR 被替换。可以仅在人类重链中替换 CDR, 用轻链作为来自人类受体抗体的未修饰的轻链。可选地, 可相容的轻链可通过常规抗体数据库选自另一人类抗体。工程化抗体的剩余物可来源于任何合适的受体人类免疫球蛋白。

工程化的人源化抗体优选地具有天然人类抗体或其片段的结构, 并具有有效治疗应用或诊断应用所要求的特性的组合, 所述治疗应用包括治疗 IL - 5 介导的人类炎症。

作为另一实例, 工程化抗体可包含三个 2E3 可变轻链区的 CDR[SEQ ID NO:10、11 和 13]和三个 2B6 可变重链区的 CDR[SEQ ID NO:7、8 和 9]。所得人源化抗体应具有与 mAb 2B6

相同的抗原结合特异性和高亲和性。

本领域技术人员应当理解, 工程化抗体可通过改变可变结构域氨基酸而被进一步修饰, 同时不会显著影响供体抗体(即类似物)的特异性和高亲和性。预期重链和轻链氨基酸可被其他氨基酸取代, 取代位于可变结构域框架和/或 CDR 中。

此外, 可改变恒定区以提高或降低本发明分子的选择特性。例如二聚化, 与 Fc 受体结合或结合并激活补体的能力(见例如 Angal 等, 分子免疫学, 30:105-108(1993), Xu 等, 生物化学杂志, 269:3469-3474(1994), Winter 等, EP 307, 434-B)。

作为嵌合抗体的改变的抗体与上面描述的人源化抗体不同, 它提供完整的非人类供体抗体重链和轻链可变区(包括框架区), 它们均与两条链的人类免疫球蛋白恒定区相连。预期相对人源化抗体而言保留了附加的非人

类序列的嵌合抗体可在人体中引发显著的免疫应答。

这样的抗体在本发明和 IL - 5 介导的疾病的治疗中是有用的，如下所述。

VI. 改变的抗体和工程化抗体的制备

优选地，mAb 2B6 或其他合适的供体单克隆抗体（例如 2E3、2F2、4A6 等）的可变轻链和/或重链序列以及 CDR，和它们的编码核酸序列，被用于通过如下方法构建本发明的改变的抗体，优选人源化抗体。相同或类似的技术也可用于产生本发明的其他实施方案。

用常规方法克隆产生选定供体 mAb（例如鼠抗体 2B6）的杂交瘤，该抗体的重链和轻链可变区的 DNA 可通过本领域技术人员已知的技术得到，例如 Sambrook 等（分子克隆实验指南，第二版，冷泉港实验室（1989））描述的技术。用多核苷酸引物和逆转录酶得到 2B6 的可变重链和轻链，它们至少包含 CDR 编码区和受体 mAb 轻链和/或重链可变结构域框架区的某些部分（这些部分是为保留供体 mAb 结合特异性所需的），以及残余的免疫球蛋白来源的部分，所述免疫球蛋白来源于人类免疫球蛋白的抗体链。用已知的数据库并通过与其他抗体进行比较确定出 CDR 编码区。

可制备鼠/人嵌合抗体并检测结合活性。这种嵌合抗体包含完整的非人类供体抗体 V_H 和 V_L 区，它们在两条链中均与人类 Ig 恒定区相连。用计算机化的数据库（例如 KABAT®）来确定来自人类抗体的重链可变区的同源框架区，并且选择与 2B6 有源性的人类抗体作为受体抗体。设计在人类抗体框架内含有 2B6 CDR 编码区的合成的重链可变区的序列，它在框架区内具有任选的核苷酸替换，以引入限制性位点。随后用长的合成寡聚物合成出所设计的序列。可选地，所设计的序列是通过重叠寡核苷酸，用聚合酶链式反应（PCR）扩增和修正错误而被合成的。

用类似的方式可以设计出合适的轻链可变框架区。

人源化的抗体可来源于嵌合抗体，或优选地通过合成制备，合成是通过在选定的重链和轻链框架内适当地插入来自重链和轻链的供体 mAb CDR 编码区而完成的。可选地，本发明的人源化的抗体可用标准诱变技术制备。因此，所得的人源化抗体包含人类框架区和供体 mAb CDR 编码区。随后可对框架残基进行进一步操作。所得人源化抗体可在重组宿主细胞（例如 COS、CHO 或骨髓瘤细胞）中表达。其他人源化抗体可用这一技术在其他合适的 IL - 5 特异性的、中和的、高亲和性的非人类抗体上制备。

常规的表达载体或重组质粒可这样制备：将改变的抗体的这些编码序列与常规的调节控制序列可操作地连接，所述调节控制序列能控制在宿主细胞中的复制和表达，和/或从宿主细胞中分泌。调节序列包括启动子序列，例如 CMV 启动子，和信号序列，它可来源于其他已知抗体。类似地可制备第二表达载体，该载体具有编码互补抗体轻链或重链的 DNA 序列。优选地，第二表达载体与第一表达载体相同，只是涉及编码序列和可选择的标记物，以确保各多肽链被功能性地表达。可选地，改变的抗体的重链和轻链编码序列可存在于单一载体上。

通过常规技术将第一和第二载体共转染选定的宿主细胞（或只用单一载体转染），以产生本发明的被转染的细胞，该细胞同时包含重组的或合成的轻链和重链。随后用常规技术培养被转染的细胞，以制备本发明的工程化的抗体。通过适当的检测技术，如 ELISA 或 RIA，从培养物中筛选出与重组重链和/或轻链相连的人源化抗体。类似的常规技术可用于构建本发明其他改变的抗体和分子。

本领域的技术人员可筛选出克隆和亚克隆步骤所用的合适的载体，所述的克隆和亚克隆步骤被用于本发明的方法和组合物的构建中。例如可使用常规的 pUC 系列克隆载体。所用的一种载体是 pUC19，它可以从例如 Amersham(Buckinghamshire, 英国)或 Pharmacia(Uppsala, 瑞典)购得。此外，任何容易复制，具有大量克隆位点和可选择的基因（例如抗生素抗性），容易操作的载体可用于克隆。因此，对克隆载体的选择在本发明中不是限制性因素。

类似地，本领域的技术人员可从任意常规载体中选择出表达本发明的工程化抗体所需的载体。这些载体也包含选定的调节序列（如 CMV 启动子），它指导异源 DNA 序列在选定的宿主细胞中的复制和表达。这些载体含有上面描述的 DNA 序列，所述序列编码工程化抗体或改变的免疫球蛋白编码区。此外，载体中可掺入选定的免疫球蛋白序列，该序列通过所需的限制性位点的插入而被修饰，以利于操作。

表达载体的特征也可在于这样一些基因，它们适于扩大异源 DNA 序列（例如哺乳动物二氢叶酸还原酶基因（DHFR））的表达。其他优选的载体序列包含一段 poly A 信号序列，该序列来自例如牛生长激素（BGH）和 β 珠蛋白启动子序列（betaglopro）。本文所用的表达载体可用本领域技术人员充分了解的技术合成。

这些载体的组分，例如复制子、选择基因、增强子、启动子、信号序列等，可从商业或天然来源得到，或用已知的方法合成，它们被用来在宿主细胞中指导重组 DNA 的产物的表达和/或分泌。可为此目的选择其他合适的表达载体，它们之中用于哺乳动物、细菌、昆虫、酵母和真菌表达的多类型是本领域已知的。

本发明也包括用重组质粒转染的细胞系，所述重组质粒包含工程化抗体或其改变的免疫球蛋白分子的编码序列。适于这些克隆载体克隆和其他操作的宿主细胞也是常规的。但是，最优选地，来自大肠杆菌的不同菌株被用于克隆载体的复制和本发明的改变的抗体的构建中的其他步骤。

用于表达本发明的工程化抗体或改变的抗体的适当的宿主细胞或细胞系优选地是哺乳动物细胞，例如 CHO、COS、成纤维细胞（例如 3T3）和骨髓细胞，更优选的是 CHO 或骨髓细胞。可用人类细胞，这使得分子可以以人类糖基化模式被修饰。可选地，可以使用其他真核细胞系。合适的哺乳动物宿主细胞的选择，转化、培养、扩增、筛选和产物制备和纯化的方法是本领域已知的。见例如 Sambrook 等，引文同上。

细菌细胞可用作适于表达本发明的重组 Fab 的宿主细胞（见例如 Pluckthum, A., 免疫学综述, 130:151-188(1992)）。但是，由于细菌细胞表达的蛋白有处于非折叠或不正确折叠形式或处于非糖基化形式的倾向，应对任何细菌细胞产生的重组 Fab 筛选其抗原结合能力的保留。如果细菌细胞表达的分子以正确折叠的形式产生，该细菌细胞将是所需的宿主。例如，用于表达的不同大肠杆菌菌株是生物技术领域众所周知的宿主细胞。枯草芽孢杆菌、链霉菌和其他杆菌等的不同菌株也可用于本方法。

如果需要，本领域技术人员已知的酵母菌株也可作为宿主细胞，昆虫细胞（例如果蝇和鳞翅目昆虫）和病毒表达系统也是如此。见例如 Miller 等，遗传工程, 8:277-298, Plenum Press(1986)以及其中所引用的文献。

构建本发明的载体的一般方法，产生本发明的宿主细胞所需的转染方法，和从该宿主细胞产生本发明的改变的抗体所必需的培养方法都是常规的技术。同样，本发明的改变的抗体一旦生成，即可根据本领域的标准方法从细胞培养物中将其纯化，所述方法包括硫酸铵沉淀、亲和柱、柱层析、凝胶电泳等。这些技术属于本领域之内，并且没有对本发明进行限制。

人源化抗体的另一种表达方法可利用在转基因动物中的表达，如美国专利 4,873,316 号中描述的。这涉及利用动物酪蛋白启动子的表达系统，当其

被转基因性地掺入哺乳动物中时，能使雌性动物在其奶中产生所需的重组蛋白。

一旦用所需的方法表达，即通过合适的检测手段检查工程化抗体的体外活性。用目前常规的 ELISA 检测方案来定量和定性地评估工程化抗体与 IL - 5 的结合。此外，可用其他体外检测手段确认中和效果，然后进行人类临床研究以评价工程化抗体除通常的清除机制外的机体内持续能力。

在由 2B6 制备人源化抗体的方法之后，本领域的技术人员可从本文描述的其他供体 IL - 5 抗体、可变区序列和 CDR 肽构建人源化的抗体。工程化抗体可用可变区框架制备，该框架可能通过接受工程化抗体而识别“自身”。可变区框架的小的改变可使抗原结合产生大的增加，但对接受者的免疫原性没有可觉察的增加。这种工程化的抗体可有效地治疗 IL - 5 介导的人类疾病。这类抗体也用于诊断这些疾病。

VII. 治疗/预防应用

本发明也涉及治疗人类嗜曙红细胞相关症状（如气喘）的方法，包括施用有效剂量的抗体，所述抗体包括本文描述的工程化抗体或改变的抗体或其片段中的一种或多种。

利用本发明的分子诱导的治疗反应可通过结合人类 IL - 5 然后阻遏嗜曙红细胞刺激作用而产生。因此，当处于适于治疗应用的制备物和制剂中时，本发明的分子对这样一些人是非常有用的，他们患有变态和/或特应性反应，或与嗜曙红细胞有关的反应，例如但不限于，过敏性鼻炎、气喘、慢性嗜曙红性肺炎、过敏性支气管曲霉病、腹腔病、嗜曙红性胃肠炎、Churg-Strauss 综合症（结节性动脉外膜炎和特应性病）、嗜曙红肌痛综合症、嗜曙红细胞过多综合症、水肿反应（包括发作性 angiodema）、蠕虫感染（其中嗜曙红细胞起保护性作用）、盘尾丝虫皮炎和特应性皮炎。

本发明的改变的抗体、抗体及其片段也可与其他抗体结合应用，特别是可与其他标记物（表位）反应的人类 mAb，所述标记物与疾病相关，本发明的工程化抗体是针对这些标记物的。

据信，本发明的治疗性药剂可在大约 2 天至 3 周内或根据需要治疗过敏性疾病。例如，当治疗季节性鼻炎等时，可能需要较长的治疗时间。这比目前在治疗 IL - 5 介导的疾病时所用的输注方案要好。治疗的剂量和持续时间与本发明分子在人类循环系统中的相对持续时间有关，本领域的技术人员可根据患者的病情和一般健康状况对其进行调整。

本发明药剂的施用方式可以是任何将药剂输送到宿主中的合适的途径。本发明的改变的抗体、抗体、工程化的抗体及其片段以及药物组合物特别适于胃肠外施用，即皮下、肌内、静脉内或鼻腔内。

本发明的治疗药剂可被制成药物组合物，该组合物中含有有效量的处于药用载体中的本发明的工程化（例如人源化）抗体作为活性成分。在本发明的预防性药剂中，以适于注射的形式存在的含有工程化抗体的水悬液或溶液（优选在生理 pH 下缓冲的）是优选的。用于胃肠外施用的组合物通常含有本发明工程化抗体的溶液或其溶于药用载体形成的混合液，所述载体优选地是水基载体。可用不同的水基载体，例如 0.4% 盐水，0.3% 甘氨酸等。这些溶液是灭菌的，通常没有颗粒状物种。这些溶液可用常规的、众所周知的灭菌技术（例如过滤）灭菌。组合物中可含有接近生理条件所需的药用辅助性物质，例如 pH 调节剂和缓冲剂等。本发明抗体在药物组合物中的浓度可在较宽的范围内变化，即重量比从小于 0.5%，通常等于或小于 1%，到 15% 或 20%，根据所选择的具体施用方式，基于液体的体积、粘度等来选择抗体的浓度。

因此，用于肌肉内注射的本发明的药物组合物可包含 1 毫升无菌缓冲液，和大约 1 纳克至 100 毫克，例如大约 50 纳克至 30 毫克，或更优选大约 5 毫克至 25 毫克的本发明的工程化抗体。类似地，用于静脉输注的本发明的药物组合物可包含 250 毫升无菌的 Ringer's 溶液，和大约 1 至 30，优选 5 毫克至大约 25 毫克的本发明的工程化抗体。制备可胃肠外施用的组合物的精确方法对本领域的技术人员是已知的或是显而易见的，更详细的描述见例如 Remington's Pharmaceutical Science, 15th ed., Mack Publishing Company, Easton, Pennsylvania。

优选地，当处于药物制备物中时，本发明的治疗药剂以单位剂量形式存在。适当的治疗有效剂量可由本领域的技术人员容易地确定。为了有效地治疗人或其他动物的炎症，应当对每 70 公斤体重通过胃肠外（优选静脉内或肌肉内）施用大约 0.1 毫克至 20 毫克剂量的本发明的蛋白或抗体。如果必要的话，在炎症反应期间，医生可选择合适的时间间隔重复这一剂量。

本发明的改变的抗体和工程化抗体也可用于诊断，例如确定 IL - 5 介导的疾病或跟踪这类疾病的治疗进程。作为诊断药剂，这些改变的抗体可被常规地标记以用于 ELISA 和其他常规的测量血清、血浆或其他适当的组织中的，或由培养物中的人类细胞释放的 IL - 5 水平的检测方案。其中利

用了改变的抗体的检测的性质是常规的，并且不限制本发明公开。

因此，本发明的一个实施方案涉及帮助诊断过敏或其他与嗜曙红细胞过量产生有关的病症的方法，它包括以下步骤：确定得自所述患者的样本（血浆或组织）中人类 IL - 5 的量，将该量与普通群体中的人类 IL - 5 平均量进行比较，由此，如果患者样本中 IL - 5 的量显著升高，则表明有与嗜曙红细胞过量产生有关的过敏或其他病症。

本文描述的抗体、改变的抗体或其片段可被冷冻干燥以便贮存，并在使用前在合适的载体中重建。这一技术已被证明对常规的免疫球蛋白有效。本领域已知的冷冻干燥和重建技术可被采用。

下述实施例阐明了本发明的不同实施方案，包括示例性工程化抗体的构建及其在合适的载体和宿主细胞中的表达，并且不应被理解为限制本发明的范围。所有的氨基酸都以常规的三字母或单字母编码表示。除非另外说明，所有必要的限制性酶、质粒和其他药剂和材料都是从商业途径得到的。所有的一般克隆连接和其他重组 DNA 方法都按下面文献中描述的进行：T. Maniatis 等，引文同上，或其第二版（1989），Sambrook 等编辑，出版者相同（“Sambrook 等”）。

实施例 1 - hIL - 5 的 MAb 的制备

在果蝇 Schneider 2 (S₂) 细胞中表达人类 IL - 5，将其纯化至均一。用草地贪夜蛾 21(Sf21) 在 Baculovirus 中表达鼠 IL - 5，将其纯化至均一。单克隆抗体 TRFK - 5（一种中和大鼠抗小鼠 IL - 5 抗体）得自 Genzyme 公司(Cambridge, MA)。

A. 免疫方法：

重组人类 IL - 5 (IL - 5) 被用作一组七只 CAF1 雌性小鼠(Charles River, Wilmington, MA) 的免疫原。动物接受三次皮下注射，注射的是磷酸缓冲液 (PBS) 中的 IL - 5，用 1: 1 的 TiterMAX™(CytoRx Corp., Nocross, GA) 将其乳化四个月。激发抗原的剂量是 50 微克，加强剂量是 25 和 10 微克。加强之后收集血清样本，检测与 IL - 5 的结合，并通过受体结合抑制试验和 B13 增殖试验（或 IL - 5 中和试验（实施例 2C））检测中和活性。所有小鼠产生的血清样本都结合 IL - 5。在无痛苦致死之前 3 天，用 10 微克重组人类 IL - 5 对被选作脾供体的动物进行静脉内加强注射。

B: 杂交瘤发育：

所用的是 Kohler 等(Nature, 256:495(1975)) 首先报道的融合方法，对其进

行了修改以使用细胞单层来完成这一技术(Kennet 等编, “杂交瘤: 生物分析的新方向”, pp.368-377, Plenum Press, New York)。合并来自两只供体小鼠的脾细胞, 用 5 千万脾细胞对 1 千万 SP2/0/Ag14 骨髓瘤细胞的比例进行融合。通过 ELISA 检测融合阳性孔的上清液对 IL - 5 的结合。含有产生 IL - 5 的抗体的细胞的孔被展开, 用 IL - 5 受体结合抑制试验和 B13 (中和) 增殖试验 (下面将要描述) 筛选上清。

分离出 16 个杂交瘤, 它们分泌可与 IL - 5 反应的单克隆抗体。将杂交瘤上清与碘化的 IL - 5 混合, 加入到由表达 IL - 5 受体 (IL - 5R) α 链的果蝇细胞制备的膜提取物中, 检测受体结合的抑制。11 个杂交瘤上清液对碘化 IL - 5 与 IL - 5 受体 α 链结合的抑制大于 60%。三种单克隆抗体 (2B6、2E3 和 2F2) 也对由人而不是鼠 IL - 5 引发的鼠 B13 细胞的增殖的抑制大于 70%。杂交瘤中的 5 个, 其中 4 个阻遏结合和/或增殖 (1C6、2B6、2E3 和 2F2), 1 个是非中和的 (24G9), 在软琼脂中被反复地亚克隆, 以产生稳定的克隆细胞系。用 ELISA 筛选来自克隆细胞系上清的交叉反应, 这些上清不与人类 IL - 1 α 、IL - 1 β 、IL - 4、IL - 8、M - CSF 或 TGF α 结合。这些单克隆抗体被纯化, 其结合亲和性用光学生物传感器(BIAcore)测定为范围是 10 - 100pM。用同种型试剂(PharMingen, San Diego, CA)将来自这些细胞系的上清同种型化。

通过类似的方法, 如对小鼠进行的描述, 用相近的免疫方案和用于融合的大鼠杂交瘤, 从免疫的大鼠中得到大鼠杂交瘤。两个大鼠杂交瘤 (4A6 和 5D3) 被鉴定为产生可与 IL - 5 结合的单克隆抗体。象 mAb 2B6、2E3 和 2F2 一样, 发现 mAb 4A6 和 5D3 在下文描述的 B13 试验中被中和。

C. 杂交瘤保藏:

产生单克隆抗体 2B6 的杂交瘤细胞系 SK119-2B6.206.75(1)被保藏在美国典型培养物保藏中心 (ATCC), Rockville, MD, USA, 保藏号 HB 11783, 并且依照 1977 年的布达佩斯协议被专利保藏处接受, 该协议管理为专利程序目的的微生物保藏。

产生单克隆抗体 2E3 的杂交瘤细胞系 SK119-2E3.39.40.2 被保藏在美国典型培养物保藏中心 (ATCC), Rockville, MD, USA, 保藏号 HB 11782, 并且依照 1977 年的布达佩斯协议被专利保藏处接受, 该协议管理为专利程序目的的微生物保藏。

产生单克隆抗体 2F2 的杂交瘤细胞系 SK119-2F2.37.80.12(1)被保藏在美

国典型培养物保藏中心 (ATCC) , Rockville, MD, USA, 保藏号 HB 11781, 并且依照 1977 年的布达佩斯协议被专利保藏处接受, 该协议管理为专利程序目的的微生物保藏。

产生单克隆抗体 24G9 的杂交瘤细胞系 SK119-24G9.8.20.5(1)被保藏在美国典型培养物保藏中心 (ATCC) , Rockville, MD, USA, 保藏号 HB 11780, 并且依照 1977 年的布达佩斯协议被专利保藏处接受, 该协议管理为专利程序目的的微生物保藏。

产生单克隆抗体 4A6 的杂交瘤细胞系 4A6(1)G1F7 于 1995 年 6 月 8 日被保藏在美国典型培养物保藏中心 (ATCC) , Rockville, MD, USA, 保藏号 HB 11943, 并且依照 1977 年的布达佩斯协议被专利保藏处接受, 该协议管理为专利程序目的的微生物保藏。

产生单克隆抗体 5D3 的杂交瘤细胞系 5D3(1)F5D6 于 1995 年 6 月 8 日被保藏在美国典型培养物保藏中心 (ATCC) , Rockville, MD, USA, 保藏号 HB 11942, 并且依照 1977 年的布达佩斯协议被专利保藏处接受, 该协议管理为专利程序目的的微生物保藏。

实施例 2 - 检测

A. ELISA:

用溶于 0.05M, pH9.6 碳酸缓冲液的 0.2 微克 IL - 5 包被 MaxiSorb™ 免疫平板的单独的孔。4°C 保温过夜后, 用含 0.025%吐温®20 的 PBS 清洗平板, 用溶于含 0.025%吐温®20 的 PBS 的 1%BSA 在室温下封闭 2 小时。未稀释的杂交上清液被加到 IL - 5 包被的孔中, 在室温下保温 2 小时。清洗平板之后, 加入用 PBS 稀释成 1/7500 的过氧化物酶标记的山羊抗小鼠 IgG & IgM(Boehringer Mannheim, Indianapolis, IN), 所述 PBS 中含有 1% BSA 和 0.025%吐温®20。2 小时之后清洗平板, 加入 0.2 毫升 pH4.75 的 0.1M 柠檬酸缓冲液, 该缓冲液中含有 0.1%的尿素过氧化物和 1mg/ml 的邻苯二胺。15 分钟后, 将平板在 VMax™ Microplate Reader(Molecular Devices, Menlo Park, CA)上以 450nm 读数。

B. 受体结合抑制试验:

表达人类 IL - 5 受体 (IL - 5R) α 链的果蝇 S2 细胞的膜提取物被用于测量抗体对 IL - 5 与受体结合的影响。为制备膜, 在 4°C 下用 1000 × g 离心 10 分钟沉淀 10⁹ 个细胞。用干冰/乙醇浴冷冻细胞沉淀 15 分钟。融化沉淀, 用 10 毫升 PBS 在 4°C 下重悬, 用 1000 × g 离心 10 分钟。用 2 ×

PBS 清洗细胞沉淀, 重悬于 13.5 毫升高渗缓冲液(10mM Tris pH7.5, 3mM MgCl₂, 1mM 二硫苏糖醇, 1mM 苯甲磺酰氟, 1uM leupeptin, 1uM pepstatin A), 在冰上保温 5 分钟。细胞悬液在 15 毫升 Dounce 匀浆器中匀浆, 用 2.5M 蔗糖溶液使蔗糖的终浓度达到 0.25M。1000 × g 离心 15 分钟除去细胞碎片。4°C 下 100,000 × g 离心 90 分钟沉淀细胞膜, 用 50 毫升 10mM Tris pH7.5, 3mM MgCl₂, 250mM 蔗糖重悬细胞膜, -70°C 保存。

用含有受体的果蝇膜进行的检测在 MultiscreenGV™ 平板(Millipore Corp., Bedford, MA)上用果蝇果蝇组织培养基 M3(Lindquist 等, *Drosophila Inf. Serv.*, 58:163(1982))完成, M3 中含有 25mM HEPES 缓冲液, pH7.2 和 0.1% BSA(结合缓冲液)。用 0.1 毫升结合缓冲液将孔预封闭。50 微升的待测样本, 每样三份, 被加入孔中, 随后加入 25 微升碘化(¹²⁵I)IL - 5。室温下温育 20 分钟后, 25 微升表达人类 IL5R α 链的果蝇 S2 细胞的膜提取物被加入孔中。进一步培养 1 小时后, 通过真空过滤收集膜, 用结合缓冲液洗 3 次。过滤物被干燥和计数。

C. IL - 5 中和检测:

鼠 IL - 5/IL - 3 依赖型细胞系 LyH7.B13(B13)得自 R. Plalcios, Basel Institute of Immunology, Switzerland。细胞在 RPMI 1640 培养液(GibcoBRL, Renfrewshire, UK)中每周深层培养两次, 培养液中补充了 L - 谷氨酰胺, 非必需氨基酸, 丙酮酸钠, 青霉素 - 链霉素(均为 GibcoBRL), 加 2 - 巯基乙醇(5×10^{-5} M, Sigma), 10%胎牛血清(Globepharm, Surrey, UK)和 1 - 10 单位的鼠 IL - 5。为进行检测, 在存在被适当稀释的待测样本的条件下, 将细胞在 96 孔圆底平板中每样三份地培养 48 小时, 在最后 4 个小时中快速施加 0.5 uCi ³H - 胸苷(Amersham, Bucks, UK)。在 1205 Betaplate(LKB Wallac, Beds, UK)中对其进行闪烁计数。

D. 光学生物传感器:

用 BIAcore 光学传感器(Pharmacia Biosensor, Uppsala, Sweden)测量固定化的 hIL - 5 和抗体的动力学和平衡结合特性。用前面描述的(Karlsson 等, *J. Immunol. Meth.*, 145:229-240(1991))

相关内容评价动力学数据, 该文献完整地在此引入作为参考。

三种中和单克隆抗体, 即 2B6、2E3 和 2F2, 具有非常类似的抑制 ¹²⁵I - IL - 5 与膜受体结合和 B 细胞增殖的中和的能力, 并对 IL - 5 有非常类似的亲和性(见表 I)。来自这三种 mAb (2 种 IgG₁ 和一种 IgG_{2a}) 的

V_H 和 V_L 的核苷酸序列已被确定。所得的序列非常类似，仅有几个残基的差异。

表 I

可与人类 IL - 5 反应的 mAb 的亲亲和性和中和能力

| mAb | Kd(pM) ^a | 中和 | | |
|------|---------------------|---------------------------------|----------------------------------|---------------------|
| | | 结 合 IC50(nM) ^b | 合 增 殖 IC50 ^c | 100%抑制 ^c |
| 2B6 | 22 | 1 | 70 | 200 |
| 2E3 | 20 | 1 | 90 | 600 |
| 2F2 | 13 | 1 | 150 | 340 |
| 1C6 | 86 | 43 | 12,200 | ND |
| 24G9 | ND | >133 | >100,000 | ND |
| 4A6 | 18 | >88 | 28 | 100 |
| 5D3 | ND | ND | 100 | 10,000 |

^a 用光学生物传感器(BIOcore)分析(25°C)确定

^b 对 125I - IL - 5 与来自果蝇膜的 IL - 5R(α 链)结合的抑制

^c 对 8pM 人类 IL - 5 引发的 B13 细胞增殖(以 pM 为单位)的抑制

ND = 没有数据

实施例 3 - 来自组合文库的 IL - 5 Fab 的分离和定性

A. PCR 和组合文库的构建:

用 cDNA 试剂盒(Boehringer Mannheim, Indianapolis, IN)将纯化自三只小鼠的脾的 RNA 逆转录, 使用试剂盒中提供的引物(dT)15' 或 3' Fd(IgG1, IgG2a & IgG3), 以及 kappa 轻链引物, 如 Huse 等(Science, 246:1275(1989))和 Kang, S.A.(Methods:Companion Methods Enzymol., 2:111(1991))所述, 文献在此引入作为参考。用所描述的(Huse 等, 同上)引物和热循环条件通过 PCR 扩增免疫球蛋白 cDNA。对所有反应采用使用了 AmpliWax™ PCR Gem 100(Perkin Elmer Cetus, Norwalk, CT)小珠的“热起始(Hot Start)”技术和制造商的方案。PCR 产物被凝胶纯化、消化并连接到 pMKFabGene3 载体(Ames 等, J. Immunol., 152:4572(1994))上。与 Fd cDNA 连接后的文库滴度为 5.1×10^7 CFU, 与 kappa cDNA 连接后的文库滴度为 1.5×10^6 CFU。

用辅助噬菌体 VCSM13(Stratagene)感染 XL1-Blue 细胞(Stratagene, La Jolla, CA), 该细胞已用噬菌粒文库进行了转化, 所述噬菌体是如 Barbas 和 Lerner(Methods:Companion Methods Enzymol., 2:119(1991))所述制备的。

B. 生物淘选 (Biopanning) :

用溶于 0.1M 碳酸氢盐, pH8.6 的 IL - 5(1ug/孔)在 4°C 下将四微滴定孔 (Immulon II Removawell Strips, Dynatech Laboratories Inc., Chantilly, VA)包被过夜。用水清洗孔, 用含有 3% BSA 的 PBS 在 37°C 下将孔封闭 1 小时。除去封闭溶液, 将文库加入微滴定孔中(50ug/孔), 在 37°C 下保温 2 小时。用 TBS/吐温®(50mM Tris-HCl, pH7.5, 150mM NaCl, 0.5%吐温®20)将孔洗 10 次, 用水洗 1 次, 然后用下述溶液将粘附的噬菌体洗脱: 0.1M HCl, 用甘氨酸调至 pH2.2, 含 1mg/ml BSA。

C. 菌落转移 s:

按所述方法(Barbas 和 Lerner, 同上)处理菌落转移 s, 该菌落转移 s 来自分离自第 3 和第 4 轮生物淘选的克隆。用 0.5-1.0 uCi ¹²⁵I-IL-5 在室温下, 在含有 1%BSA 的 PBS 中温育过滤物 1 小时, ¹²⁵I-IL-5 是按制造商建议的方法用 Bolton-Hunter 试剂(NEN, Billerica, MA)碘化的。然后用 PBS 0.25% 吐温清洗, 在柯达 XAR 胶片上曝光。用放射自显影检测表达 IL - 5 反应性 Fab 的克隆。

D. 可溶 FAB 的制备:

用 NheI 和 SpeI 消化噬菌粒 DNA 以除去基因 III 并自连接。XL1-Blue 细胞被转化, 分离的克隆在含有 1%葡萄糖和 50ug/ml 羧苄青霉素的 5.0 毫升 super broth(SB)培养基 (30 克胰蛋白胨, 20 克酵母提取物, 10 克 3 - [N - 吗啉代]丙磺酸, MOPS, pH 调至 7) 中, 在 37°C 下生长过夜。用 Beckman GS-6G 离心机 3500rpm 离心 10 分钟将 1 毫升培养物中的细胞沉淀, 该细胞被用于接种 5 毫升 SB, SB 中含有 50ug/ml carbenicillin。将培养物在 37°C 下振摇 1 小时, 加入异丙基 - b - D - 硫代半乳糖苷 (IPTG; 1mM), 将培养物移至 28°C 过夜。通过在 4°C 下, 在悬于 30mM Tris pH8.0 的 20%蔗糖中裂解细胞来从胞外质中制备可溶性 Fab, 然后在 Microfuge 中离心 10 分钟。用 western 印迹, 通过与含有已知量的鼠 Fab 的样本进行比较来估计 Fab 的浓度。不同细菌的胞外质提取物含有相似浓度的 Fab, 按 western 印迹分析所估计的, 其浓度范围是 1 - 20ug/ml。

E. FAB 的纯化:

用基因工程方法将一段整合肽加到重链的羧基末端，以利于蛋白纯化。用 NheI 和 SpeI 消化除去 M13 基因 III 编码区之后，编码 6 个组氨酸残基的一对重叠寡核苷酸被亚克隆到 Fab 表达载体中：

[SEQ ID NO:43] 5'-CTAGCCACCACCACCACCACCACTAA-3'

[SEQ ID NO:44] 3'-GGTGGTGGTGGTGGTGGTGGTGATTGATC-5'.

用上面描述的方法诱导 Fab 的表达。在 28°C 下诱导过夜之后，通过在 20%蔗糖，30mM TRIS pH8.0 中，在 4°C 下保温 30 分钟制备细胞沉淀的胞外质裂解物。向澄清的上清中加入尿素和 Brij-35 试剂至终浓度分别为 2M 和 1%。在室温下摇动 1 小时后，处理后的澄清上清液以 0.5 毫升/分钟的速率直接加到 5 毫升镍 - NTA 金属螯合柱 (1.5 × 3 厘米) 上，该柱已用缓冲液 A (100 mM 磷酸钠，10mM Tris, 0.3M NaCl, 2M 尿素, pH8.0) 平衡。在 4 个柱体积 (20 毫升) 的洗柱之后，用 6 个柱体积 (30 毫升) 的从 pH8 到 pH4 的反向 pH 梯度，在与上面相同的缓冲液中洗脱下结合的物质。从柱中洗脱下的纯化的 Fab 在 pH5.5 时呈现尖锐的对称峰。它们的纯度大于 90%，不含 DNA。

F. FAB ELISA :

用悬于 0.1M pH8.6 碳酸氢盐缓冲液的蛋白 (1 毫克/毫升，每孔 50 毫升) 包被 Immulon II 平板 (Dynatech)。稀释和清洗在含有 0.05% 吐温™20 的 PBS 中进行。将平板在室温下用含有 1% BSA 的 PBS 清洗并封闭 1 小时。将含有可溶 Fab 或纯化 Fab 的，不同稀释率的细菌上清加到平板上。温育 1 小时之后清洗平板，加入生物素化的山羊抗小鼠 kappa (Southern Biotechnology Associates, Inc., Birmingham, AL) (1:2000 稀释；50 微升/孔) 1 小时。清洗平板，加入链亲和素标记的辣根过氧化物酶 (1:2000 稀释；50 微升/孔) 1 小时。清洗平板，加入 ABTS 过氧化物酶底物 (100 微升/孔；Kirkegaard & Perry Laboratoris, Gaithersburg, MD)，在 UVmax™ (Molecular Devices) 微平板计数器上读取 405nm 的光密度。

G. 来自组合文库的 Fab 的分离和定性

通过多轮抗包被了 IL - 5 的微滴定孔的生物淘选，从文库中选出带有抗 IL - 5 的 Fab 的噬菌体。筛选 4 轮之后，用菌落转移试验鉴定 IL - 5 反应性 Fab，所述试验利用了 ¹²⁵I - IL - 5。第 3 轮的 34 个克隆和第 4 轮的 4 个克隆被鉴定为结合标记的 IL - 5。与 IL - 5 的结合通过直接结合 ELISA 而被确认，ELISA 中应用了表达 Fab - 基因 III 融合蛋白的培养

物上清。从这些克隆中分离出 DNA，在去除 M13 基因 III 的编码区之后，可溶 Fab 的表达被诱导。制备胞外质级分，用 ELISA 检测其与 IL - 5 的结合。Fab 特异性地结合 IL - 5，没有显示出结合另一种蛋白，rC5a。

未稀释的胞外质提取物（含有 1 - 20 ug/ml Fab）在 IL - 5R 结合抑制试验（实施例 2）中被检测。Fab 中没有一个抑制碘化 IL - 5 与 IL - 5R α 的结合超过 35%。

H. 中和 mAb 转变成 FAB：

用上面描述的条件，通过 PCR 分离 mAb(2B6)的 Fd 和 κ cDNA。凝胶纯化的片段被亚克隆到 pMKFabGene3 载体中，该载体已被修饰为包含基因 III cDNA 的 hexa-His 序列 3'，结果生成质粒 pMKFabGene3H。通过菌落转移试验鉴定出含有 2B6 重链和轻链的功能性 IL - 5 结合 Fab 克隆。用 Nhe I/Spe II 消化除去基因 III 和自连接之后重链在框架内与 hexa-His 融合，使得可以用上述方法纯化。Fab 以剂量依赖方式抑制受体结合，IC₅₀ 大约为 7.5ug/ml，与亲本 mAb（鼠 2B6）类似。

I. 链改组文库的构建和筛选：

编码中和 mAb 2B6 的 Fd 的 cDNA 作为 XhoI/SpeI 片段被亚克隆到 pMKFabGene3H 中，pMKFabGene3H 含有 SstI/XbaI 片段以代替轻链 cDNA。用 SstI 和 XbaI 消化该噬菌粒，然后与 SstI/XbaI 消化的轻链 PCR 产物连接，所述 PCR 产物来源于 IL - 5 免疫的小鼠（上文描述的）。连接后的文库滴度是 4×10^5 CFU。如上所述对组合文库进行生物淘选和菌落转移试验。

通过将编码中和 mAb 2B6 的 Fd 与相同轻链的所有组成部分匹配而构建文库，所述轻链的所有组成部分是从 IL - 5 免疫的小鼠中回收的，被用于产生组合文库。对该链改组文库进行 4 轮生物淘选（对固定化的 IL - 5），用菌落转移试验对所得的克隆进行 IL - 5 反应性检测。用 ELISA 和 IL - 5R α 结合试验对结合碘化 IL - 5 的阳性克隆进一步检测。从链改组文库中回收的两个 Fab（2 和 5）阻遏 IL - 5 与 IL - 5R α 的结合，并在 B13 试验中抑制 IL - 5 依赖性增殖。这两个 V κ 的序列与 2B6 V κ （2B6 V H 的初始轻链配对物）的序列类似。Fab 2 & 15 的轻链序列分别是 SEQ ID NO:45 和 46。对于 Fab 2，CDR 1-3 分别是 SEQ ID NO:10，11 和 47。对于 Fab 15，CDR1-3 分别是 SEQ ID NO:10，11 和 48。

下面实施例 4 和 5 中列出的所有抗体氨基酸序列使用 KABAT 编号系

统，它允许 CDR 和框架长度的变化。即，不论在它们前面的氨基酸的实际编号是多少，关键氨基酸总是被指定相同的编号。例如，所有轻链 CDR 之前的半胱氨酸总是 KABAT 位置 23，CDR1 之后的色氨酸残基总是 KABAT 位置 35，即使 CDR1 可含有多达 17 个氨基酸。

实施例 4 - 人源化的抗体

一种人源化的抗体被设计为在人类抗体框架内包含鼠类 CDR。通过下述操作来制备人源化的 IL - 5 特异性小鼠抗体 2B6。

A. 基因克隆:

用得自 Boehringer Mannheim(Indianapolis, IN)试剂盒分别从 2B6、2F2 和 2E3 杂交瘤细胞系(见实施例 1)中分离 mRNA，然后用 cDNA 试剂盒(Boehringer Mannheim)提供的引物(dT)15 将其逆转录以制备 cDNA。用小鼠免疫球蛋白特异性 PCR 引物扩增 DNA，该 DNA 编码从重链可变区的氨基酸#9(KABAT 编号系统)到铰链区，以及从轻链可变区氨基酸#9(KABAT 编号系统)到恒定区末端的结构域。通过独立的 PCR 反应得到各抗体链的几个克隆。

所用的小鼠 γ 1 铰链区引物是[SEQ ID NO:22]:

5' GTACATATGCAAGGCTTACAACCACAATC 3'

所用的小鼠 γ 2a 铰链区引物是[SEQ ID NO:23]:

5' GGACAGGGCTTACTAGTGGGCCCTCTGGGCTC 3'

所用的小鼠重链可变区引物是[SEQ ID NO:24]:

5' AGGT(C 或 G)(C 或 A)A(G 或 A)CT(G 或 T)TCTCGAGTC(T 或 A)GG 3'

所用的小鼠 kappa 链恒定区引物是[SEQ ID NO:25]:

5' CTAACACTCATTCCCTGTTGAAGCTCTTGACAATGGG 3'

所用的小鼠轻链可变区引物是[SEQ ID NO:26]:

5' CCAGATGTGAGCTCGTGATGACCCAGACTCCA 3'

PCR 片段被克隆到质粒 pGEM7f+(Promega)中，然后转化大肠杆菌 DH5 α (Bethesda Research Labs)。

B. DNA 测序:

对上面 A 部分的重链和轻链鼠 cDNA 克隆进行测序。这些克隆的可变区的测序结果显示在 SEQ ID NO:1 - 6 (图 1 - 6)中。各克隆含有这样一些氨基酸，它们在小鼠重链可变区或轻链可变区中是保守的。CDR 氨基酸序列在下面列出。

2B6 重链的 CDR 区是 SEQ ID NO:7、8 和 9。见图 7。这些序列由 SEQ ID NO:1 编码。轻链的 CDR 区是 SEQ ID NO:10、11 和 12。见图 7。这些序列由 SEQ ID NO:2 编码。

2F2 重链的 CDR 区是 SEQ ID NO:7、8 和 9。见图 7。这些序列由 SEQ ID NO:3 编码。轻链的 CDR 区是 SEQ ID NO:10、11 和 13。见图 7。这些序列由 SEQ ID NO:4 编码。

2E3 重链的 CDR 区是 SEQ ID NO:7、8 和 14。见图 7。这些序列由 SEQ ID NO:5 编码。轻链的 CDR 区是 SEQ ID NO:10、11 和 13。见图 7。这些序列由 SEQ ID NO:6 编码。

C. 人类框架的选择:

在 2B6 的克隆之后, 将重链和轻链的可变区的氨基酸序列(图 1 和 2)(分别是 SEQ ID NO:15 和 16)与 KABAT 和 SWISS - PROT(Nuc. Acids Res., 20:2019-2022(1992))蛋白序列数据库中已知的鼠免疫球蛋白序列进行比较, 以将氨基酸归属至 N 末端残基。然后将推断的 2B6 重链和轻链可变区氨基酸序列与人类免疫球蛋白序列数据库进行比较, 以确定重链和轻链的人类框架, 它应与鼠类序列最为匹配。此外, 用位置数据库(positional database)评价重链和轻链, 以估计由于氨基酸而导致的冲突, 该冲突可能影响 CDR 的呈递, 所述数据库是从 Fab 结构域的结构模型产生的。通过在冲突位置替换相应的小鼠氨基酸, 可以在人源化可变区的合成过程中解决冲突。

得自人类骨髓瘤免疫球蛋白(COR)的抗体的重链框架区被利用(E. M. Press 和 N. M. Hogg, Biochem. J., 117:641-660(1970))。发现人类重链框架氨基酸序列与 2B6 框架有大约 66%的同源性。

对于合适的轻链可变区框架, 利用了 Bence-Jones 蛋白(Len)的轻链可变框架序列(Schneider 等, Hoppe-Seyler's Z. Physiol. Chem., 356:507-557(1975))。在氨基酸水平上, 人类轻链框架区与鼠类 2B6 轻链框架区有大约 82%的同源性。

选定的人类框架被反向翻译以提供 DNA 序列。

D. 人源化 Mab 基因的构建:

提供 2B6 重链 CDR[图 7 和 SEQ ID NO:1 - 2]和人类抗体的框架序列, 制备合成的重链可变区[SEQ ID NO:18]。它是用四个合成的寡核苷酸[SEQ ID NO:27 和 28][SEQ ID NO:29 和 30]制备的, 当这些寡核苷酸连接到一起

时, 编码氨基酸#21-#106 (KABAT 编号)。然后将寡核苷酸连接到基于 pUC18 的质粒的 HpaI-KpnI 限制性位点中, 该质粒含有基于 COR 框架(同上)的来源于另一人源化重链的序列。这个质粒提供单一序列[SEQ ID NO:17]和残余的可变区序列。用诱变引物通过 PCR 或通过合成的连接物添加到已存在的限制性位点中, 对作图的序列中的任何错误进行改正。

信号序列和人源化重链可变区作为 EcoRI-ApaI 片段被从基于 pUC 的质粒中切下, 并被连接到表达载体 pCD 中, pCD 包含 IgG₁ 人类恒定区。合成的重链可变区的核苷酸和氨基酸序列在图 8[SEQ ID NO:18 和 19]中给出。人类框架残基是 SEQ ID NO:19 的氨基酸 1 - 30, 36 - 49, 66 - 97 和 109 - 119。CDR 的氨基酸序列与鼠 2B6 CDR 相同。所得的表达载体 pCDIL5HZHC1.0 在图 10 中给出。

提供 2B6 轻链 CDR[图 7 和 SEQ ID NO:10、11 和 12]和人类抗体的框架序列, 制备合成的轻链可变区[SEQ ID NO:20]。四个合成的寡核苷酸被插入基于 pUC18 的质粒中; 这四个合成的寡核苷酸编码人源化 V_L 氨基酸 #27-#58 (KABAT 编号)[SEQ ID NO:31 和 32]和氨基酸#80-#109[SEQ ID NO:33 和 34], 它们分别带有 SacI-KpnI 和 PstI-HindIII 末端; 所述质粒含有来源于另一人类轻链框架 (B17) (Marsh 等, Nuc. Acids Res., 13:6531-6544(1985))的序列, 它与 LEN 框架有高度的同源性。这个质粒提供残余的可变区序列。用诱变引物通过 PCR 或通过合成的连接物添加到已存在的限制性位点中, 对作图的序列中的任何错误和 LEN 与 B17 之间的单氨基酸差异进行改正。

人源化轻链可变区以 EcoRV-NarI 片段从 pUC 质粒中分离出来, 连接到表达载体 pCN 中, 该载体含有单一序列[SEQ ID NO:17]和 kappa 人类恒定区。合成的轻链可变区核苷酸和氨基酸序列在图 9[SEQ ID NO:20 和 21]中给出。人类框架残基是 SEQ ID NO:21 的氨基酸 1 - 23、41 - 55、63 - 94 和 104 - 113。CDR 的氨基酸序列与鼠 2B6 CDR 相同。但是, 这些 CDR 的编码序列与鼠 2B6 编码序列不同, 以产生限制性酶位点。所得的一个表达载体, pCNIL5HZLC1.0 在图 11 中给出。这些合成的可变轻链和/或质粒序列被用于构建人源化抗体。

E. 人源化 MAB 的表达:

来源于 IgG₁ 同种型的人源化重链利用了 SEQ ID NO:19 提供的合成的重链可变区。按上面描述的方法设计并合成含有 2B6 重链 CDR 的该合成的

V_H。

人源化的轻链，一种人类 kappa 链，利用了 SEQ ID NO:21 提供的合成的轻链可变区。按上面描述的方法设计并合成含有 2B6 轻链 CDR 的该合成的 V_L。编码人源化可变区的 DNA 片段被插入基于 pUC19 的哺乳动物细胞表达质粒，该质粒利用了单一序列，并含有 CMV 启动子，以及嵌合体的人类重链或人类轻链恒定区，用常规方法(Maniatis 等，引文同上)得到质粒 pCDIL5HZHC1.0 (重链)[SEQ ID NO:49，也见图 10]和 pCNIL5HZLC1.0 (轻链)[SEQ ID NO:50，也见图 11]。质粒共转染到 COS 细胞中，分别在 3 和 5 天后用实施例 5 描述的 ELISA 进行上清试验，以确定人类抗体的存在。

上述实施例描述了示例性工程化抗体的制备。利用其他抗 IL - 5 抗体 (例如 2F2、2E3、4A6、5D3、24G9 等)，可用类似的方法开发其他工程化的抗体。

F. 纯化:

CHO 表达的嵌合和人源化 2B6 的纯化可通过常规的蛋白 A (或 G) 亲和层析，离子交换和分子筛层析来完成。类似的方法已被成功地用于其他单克隆抗体 (例如呼吸道合胞病毒、白介素 4 和疟疾环孢子蛋白抗原) 的纯化，纯度 >95%。

G. 其他人源化单克隆抗体和表达质粒:

提供质粒 pCDIL5HZHC1.0[SEQ ID NO:49]，制得表达质粒 pCDIL5HZHC1.1，其中在框架位置 73 的 Asp 被 Thr 替换。这是通过将带有 EcoRV 和 XhoI 末端的合成的连接物[SEQ ID NO:51 和 SEQ ID NO:52]与同样消化的 pCDIL5HZHC1.0 连接而完成的。类似地，表达质粒 pCDIL5HZHC1.2 在框架位置 37 将 Ile 替换成 Val。这是通过将带有 HpaI 和 XbaI 末端的合成的连接物[SEQ ID NO:53 和 SEQ ID NO:54]与同样消化的 pCDIL5HZHC1.0 连接而完成的。表达质粒 pCDIL5HZHC1.3 也是通过将带有 HpaI 和 XbaI 末端的合成的连接物[SEQ ID NO:53 和 SEQ ID NO:54]与同样消化的 pCDIL5HZHC1.1 连接而完成的。

提供前面描述的基于 pUC18 的质粒，制备人源化的轻链可变区，其中框架位置#15 从 Leu 变成 Ala，所述基于 pUC18 的质粒含有四个合成的寡核苷酸序列[SEQ ID NO:31、32、33 和 34]。用 NheI 和 SacI 限制性内切酶消化该质粒，并插入一个合成的连接物[SEQ ID NO:55 和 56]。然后分离

EcoRV-NarI 片段, 将其连接到同样消化的表达载体 pCDIL5HZHC1.0 中, 得到 pCDIL5HZHC1.1。

用得自免疫球蛋白 (NEW) 的重链框架区 (Saul 等, J. Biol. Chem. 253:585-597(1978)) 和 2B6 重链 CDR [图 7 和 SEQ ID NO:1 - 2] 制备合成的可变区。可能影响 CDR 呈递的框架氨基酸被鉴定并用前面描述的方法替换。产生了四个重叠的合成寡核苷酸 [SEQ ID NO:57、58、59 和 60], 当它们被退火并延伸时, 编码单一序列 [SEQ ID NO:17] 和重链可变区的氨基酸。用 PCR 引物 [SEQ ID NO:63 和 64] 扩增该合成的基因, 并作为 BstXI-HindIII 限制性片段连接到基于 pUC18 的质粒中, 该质粒含有基于 COR 框架的来源于另一人源化重链的序列。通过将合成的寡核苷酸连接物 [SEQ ID NO:75 和 76] 插入 SacII 和 KpnI 限制性位点, 在氨基酸位置 91 (Kabat 编号系统) (等价于图 12 的位置 94) 进行苯丙氨酸到酪氨酸框架替换。所得的重链可变区 [图 12 和 SEQ ID NO:61, 62] 被称作 NEWM 人源化的重链。

用诱变引物通过 PCR 或通过将合成的连接物添加到已存在的限制性位点中, 对 mapped 序列中的任何错误进行改正。信号序列和人源化重链可变区作为 EcoRI-ApaI 片段被从基于 pUC 的质粒中切下, 并被连接到表达载体 pCD 中, 得到质粒 pCDIL5NEWM, 其中 pCD 包含 IgG₁ 人类恒定区。CDR 的氨基酸序列与鼠 2B6 重链 CDR 的相同。

用得自免疫球蛋白 (REI) 的轻链框架区 (Palm 等, Hoppe-Seyler's Z. Physiol. Chem. 356:167-191(1975)) 和 2B6 轻链 CDR [图 7 和 SEQ ID NO:10、11 和 12] 制备合成的可变区。可能影响 CDR 呈递的框架氨基酸被鉴定并用前面描述的方法替换。产生了四个重叠的合成寡核苷酸 [SEQ ID NO:65、66、67 和 68], 当它们被退火并延伸时, 编码代表的氨基酸 [图 13 和 SEQ ID NO:69, 70], 该轻链可变区被称为 REI 人源化轻链。用 PCR 引物 [SEQ ID NO:71 和 72] 扩增该合成的基因, 并作为 EcoRI-HindIII 限制性片段连接到 pGEM-7Zf(+)(Promega 公司, Madison, WI) 中。

用诱变引物通过 PCR 或通过将合成的连接物添加到已存在的限制性位点中, 对 mapped 序列中的任何错误进行改正。人源化轻链可变区作为 EcoRV-NarI 片段被从基于 pGEM-7Zf(+) 的质粒中切下, 并被连接到表达载体 pCN 中, 得到质粒 pCNIL5REI, 其中 pGEM-7Zf(+) 包含单一序列 [SEQ ID NO:17] 以及人类 Kappa 恒定区。CDR 的氨基酸序列与鼠 2B6 轻链 CDR 的相同。但是, 这些 CDR 的编码序列与鼠 2B6 的编码序列不同, 以产生限

制性酶位点。这些合成的可变轻链和/或重链序列被用于构建人源化抗体。

提供上述基于 pGEM-7Zf(+) 的质粒, 可制备人源化的轻链可变区, 其中框架位置#15 从缬氨酸变为丙氨酸。可用限制性内切酶 NheI 和 SacI 消化该质粒并插入合成的引物[SEQ ID NO:73 和 74]。然后分离出 EcoRR-NarI 片段, 将其连接到同样消化的表达载体 pCNIL5HZREI 中, 得到质粒 pCNIL5REIV15A。

实施例 5 - 嵌合抗体的构建

编码鼠单克隆抗体 2B6 重链可变区的氨基酸#9-#104 (KABAT 编号) 的 DNA 被作为 *Ava*II-*Sty*I 限制性片段从 cDNA 的基于 pGEM7Zf+ 的 PCR 克隆中分离, 所述 cDNA 由 2B6 杂交瘤细胞系 (见实施例 4) 产生。通过将该片段与四个小的合成寡连接物[SEQ ID NO:35、和 36][SEQ ID NO:37 和 38]结合到用 *Bst*X1-*Hind*III 消化的基于 pUC18 的质粒中, 可以提供侧面 (flanking) 重链可变区序列和一段信号序列[SEQ ID NO:17]。从紧密相关的鼠重链推断的 N 末端氨基酸的共有序列被指认为前 8 个 V_H 残基, 并且在 SEQ ID NO:35 和 36 中被编码。通过对 2B6 重链的前 15 个 N 末端氨基酸进行测序, 确认了重链的推断氨基酸序列。

含有信号和 V_H 区序列的 *Eco*RI-*Ap*aI 片段被分离, 并被连接到质粒 pCD 中, 该质粒已编码人类 IgG₁ 恒定区。

编码鼠单克隆抗体 2B6 轻链可变区的氨基酸#12-#99 (KABAT 编号) 的 DNA 被作为 *Dde*I-*Ava*I 限制性片段从 cDNA 的基于 pGEM7Zf+ 的 PCR 克隆中分离, 所述 cDNA 由 2B6 杂交瘤细胞系 (见实施例 4) 产生。通过将该片段与四个小的合成寡连接物[SEQ ID NO:39、和 40][SEQ ID NO:41 和 42]结合到用 *Eco*RV-*Hind*III 消化的基于 pUC18 的质粒中, 可以提供侧面 (flanking) 轻链可变区序列。从紧密相关的鼠轻链推断的 N 末端氨基酸的共有序列被指认为前 8 个 V_L 残基, 并且在 SEQ ID NO:39 和 40 中被编码。通过对 2B6 轻链的前 15 个 N 末端氨基酸进行测序, 确认了轻链的推断氨基酸序列。该可变区随后作为 *Eco*RV-*Nar*I 片段被分离出来, 并连接到表达载体 pCN 中, 该载体已含有人类 kappa 区和信号序列。

通过 pCD 和基于 pCN 的质粒共转染 COS 细胞来完成嵌合抗体的表达。3 和 5 天后收集培养物上清, 通过下面描述的 ELISA 检测免疫球蛋白的表达: 除最后一步外, 各步骤之后用 PBS 清洗。用 100 纳克/50 微升/孔的山羊抗体将微滴定平板包被过夜, 所述抗体对人类抗体的 Fc 区有特异性。加

入培养物上清，温育 1 小时。然后加入辣根过氧化物酶结合的山羊抗人 IgG 抗体，温育 1 小时。然后加入 ABTS 过氧化物酶底物(Kirkegaard & Perry Laboratories Inc., Gaithersburg, MD)。温育 1 小时后，用微滴定平板计数器(Molecular Devices 公司, Menlo Park, CA)在 405nm 处读取吸收值。嵌合抗体的表达被检测。在类似的 ELISA 中，含有嵌合抗体的 COS 细胞上清与包被了人类 IL - 5 蛋白的微滴定孔特异性地结合。这个结果证实编码抗 IL - 5 抗体的基因已被合成并表达。

上述实施例描述了示例性工程化抗体的制备。利用通过常规方式开发的其他抗 IL - 5 供体抗体（例如 2F2、2E3、4A6、5D3、24G9 等），可以用类似的方法开发其他工程化的抗体。

序列表

(1)一般资料:

(i)申请人: Ames, Robert S.

Appelbaum, Edward R.

Chaiken, Irwin M.

Cook, Richard M.

Gross, Mitchell S.

Holmes, Stephen D.

McMillan, Lynette J.

Theisen, Timothy W.

(ii)发明名称: 用于治疗 IL - 5 介导的疾病的重组 IL5 拮抗剂

(iii)序列数: 76

(iv)相关地址:

(A)收信人: SmithKline Beecham Corp./Corporate

(B)街道: P.O. Box 1539-UW2220

(C)城市: King of Prussia

(D)州: Pennsylvania

(E)国家: USA

(F)邮编: 19406-0936

(v)计算机可读形式:

(A)介质类型: 软盘

(B)计算机: IBM 兼容型

(C)操作系统: PC - DOS/MS - DOS

(D)软件: PatentIn Release #1.0, Version #1.30

(vi)当前申请资料:

(A)申请号:

(B)提交日期:

(C)分类:

(vii)在先申请资料:

(A)申请号: US 08/470110

(B)提交日期: 1995 年 6 月 6 日

(vii)在先申请资料:

- (A)申请号: US 08/467420
 (B)提交日期: 1995年6月6日
- (vii)在先申请资料:
 (A)申请号: US 08/363131
 (B)提交日期: 1994年12月23日
- (viii)律师/代理人资料:
 (A)姓名: Sutton, Jeffrey A.
 (B)注册号: 34,028
 (C)参考/登记号: P50282-2
- (ix)远程通讯资料:
 (A)电话: 610 - 270 - 5024
 (B)电传: 610 - 270 - 5090
- (2)SEQ ID NO:1 的资料:
- (i)序列特征:
 (A)长度: 334 个碱基对
 (B)类型: 核酸
 (C)链型: 单链
 (D)拓扑结构: 线性
- (ii)分子类型: DNA (基因组)
- (iv)特性:
 (A)名称/关键词: misc_feature
 (B)位置: 1..334
 (D)其他资料: /注 = “第一个碱基相应于 Kabat 位置 24”
- (xi)序列描述: SEQ ID NO:1:
- ```

ACCTGGCCTG GTGGCGCCCT CACAGAGCCT GTCCATCACT TGCACTGTCT CTGGGTTTTTC 60
ATTAACCAGC TATAGTGTAC ACTGGGTTTCG CCAGCCTCCA GGAAAGGGTC TGGAGTGGCT 120
GGGAGTAATA TGGGCTAGTG GAGGCACAGA TTATAATTTCG GCTCTCATGT CCAGACTGAG 180
CATCAGCAAA GACAACCTCCA AGAGCCAAGT TTTCTTAAAA CTGAACAGTC TGCAAACCTGA 240
TGACACAGCC ATGTACTACT GTGCCAGAGA TCCCCCTTCT TCCTTACTAC GGCTTGACTA 300

```

CTGGGGCCAA GGCACCACTC TCACAGTCTC CTCA

334

(2)SEQ ID NO:2 的资料:

(i)序列特征:

(A)长度: 315 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(iv)特性:

(A)名称/关键词: misc\_feature

(B)位置: 1..315

(D)其他资料: /注 = “第一个碱基相应于 Kabat 位置 25”

(xi)序列描述: SEQ ID NO:2:

TCCTCCCTGA GTGTGTCAGC AGGAGAGAAG GTCACTATGA GCTGCAAGTC CAGTCAGAGT 60

CTGTTAAACA GTGGAAATCA AAAGA ACTAC TTGGCCTGGT ACCAGCAGAA ACCAGGGCAG 120

CCTCCTAAAC TTTTGATCTA CGGGGCATCC ACTAGGGAAT CTGGGGTCCC TGATCGCTTC 180

ACAGGCAGTG GATCTGGAAC CGATTTCACT CTTTCCATCA GCAGTGTGCA GGCTGAAGAC 240

CTGGCAGTTT ATTACTGTCA GAATGTTTCAT AGTTTTCCAT TCACGTTCGG CTCGGGGACA 300

GAGTTGGA AA TAAAA 315

(2)SEQ ID NO:3 的资料:

(i)序列特征:

(A)长度: 334 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(iv)特性:

(A)名称/关键词: misc\_feature

(B)位置: 1..334

(D)其他资料: /注 = “第一个碱基相应于 Kabat 位置 24”

(xi)序列描述: SEQ ID NO:3:

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| ACCTGGCCTG GTGGCGCCCT CACAGAGCCT GTCCATCACT TGCCTGTCT CTGGGTTTTTC | 60  |
| ATTAACCAGT TATAGTGAC ACTGGGTTTCG CCAGCCTCCA GGAAAGGTC TGGAGTGGCT  | 120 |
| GGGAGTAATA TGGGCTAGTG GAGGCACAGA TTATAATTCG GCTCTCATGT CCAGACTGAG | 180 |
| CATCAGCAA GACAACCTCA AGAGCCAAGT TTTCTTAAAA CTGAACAGTC TCGAACTGA   | 240 |
| TGACACAGCC ATGTACTACT GTGCCAGAGA TCCCCCTTCT TCCTTACTAC GGCTTGACTA | 300 |
| CTGGGGCCAA GGCACCACTC TCACAGTCTC CTCA                             | 334 |

(2)SEQ ID NO:4 的资料:

(i)序列特征:

(A)长度: 315 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(iv)特性:

(A)名称/关键词: misc\_feature

(B)位置: 1..315

(D)其他资料: /注 = “第一个碱基相应于 Kabat 位置 25”

(xi)序列描述: SEQ ID NO:4:

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| TCCTCCCTGA GTGTGTCAGC AGGAGAGAAG GTCCTATGA GCTGCAAGTC CAGTCAGAGT  | 60  |
| CTATTAAACA GTGGAAATCA AAAGAACTAC TTGGCCTGGT ACCAACAGAA ACCAGGGCAG | 120 |
| CCTCCTAAAC TTTTGATCTA CGGGGCATCC ACTAGGGAAT CTGGGGTCCC TGATCGCTTC | 180 |
| ACAGGCAGTG GATCTGGAAC CGATTTCACT CTTACCATCA GCAGTGTGCA GGCTGAAGAC | 240 |
| CTGGCAGTTT ATTACTGTCA GAATGATCAT AGTTTTCCAT TCACGTTCGG CTCGGGGACA | 300 |
| GAGTTGGAAA TAAAA                                                  | 315 |

(2)SEQ ID NO:5 的资料:

(i)序列特征:

(A)长度: 334 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(iv)特性:

(A)名称/关键词: misc\_feature

(B)位置: 1..334

(D)其他资料: /注 = “第一个碱基相应于 Kabat 位置 24”

(xi)序列描述: SEQ ID NO:5:

```

ACCTGGCCTG GTGGCGCCCT CACAGAGCCT GTCCATCACT TGCACGTGTCT CTGGGTTTTTC 60
ATTAACCAGC TATAGTGTAC ACTGGGTTTCG CCAGCCTCCA GGAAAGGGTC TGGAGTGGCT 120
GGGAGTAATC TGGGCTAGTG GAGGCACAGA TTATAATTTCG GCTCTCATGT CCAGACTGAG 180
CATCAGCAAA GACAACCTCA AGAGCCAAGT TTTCTTAAAA CTGAACAGTC TGCAAACCTGA 240
TGACGCAGCC ATGTACTACT GTGCCAGAGA TCCCCCTTTT TCCTTACTAC GGCTTGACTT 300
CTGGGGCCAA GGCACCACTC TCACAGTCTC CTCA 334

```

(2)SEQ ID NO:6 的资料:

(i)序列特征:

(A)长度: 315 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(iv)特性:

(A)名称/关键词: misc\_feature

(B)位置: 1..315

(D)其他资料: /注 = “第一个碱基相应于 Kabat 位置 25”

(xi)序列描述: SEQ ID NO:6:

TCCTCTCTGA GTGTGTCAGC AGGAGAGAAG GTCACTATGA GCTGCAAGTC CAGTCAGAGT 60

CTGTTAAACA GTGGAAATCA AAAAAACTAC TTGGCCTGGT ACCAGCAGAA ACCAGGGCAG 120

CCTCCTAAAC TTTTGATCTA CGGGGCATCC ACTAGGGAAT CTGGGGTCCC TGATCGCTTC 180

ACAGGCAGTG GATCTGGAAC CGATTTCACT CTTACCATCA GCAGTGTGCA GGCTGAAGAC 240

CTGGCAGTTT ATTACTGTCA GAATGATCAT AGTTTTCAT TCACGTTCGG CTCGGGGACA 300

GAGTTGGAAA TAAAA 315

## (2)SEQ ID NO:7 的资料:

## (i)序列特征:

(A)长度: 5 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:1:

Ser Tyr Ser Val His

1 5

## (2)SEQ ID NO:8 的资料:

## (i)序列特征:

(A)长度: 16 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:8:

Val Ile Trp Ala Ser Gly Gly Thr Asp Tyr Asn Ser Ala Leu Met Ser

1 5 10 15

## (2)SEQ ID NO:9 的资料:

## (i)序列特征:

(A)长度: 11 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:9:

```
Asp Pro Pro Ser Ser Leu Leu Arg Leu Asp Tyr
1 5 10
```

## (2)SEQ ID NO:10 的资料:

## (i)序列特征:

(A)长度: 17 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:10:

```
Lys Ser Ser Gln Ser Leu Leu Asn Ser Gly Asn Gln Lys Asn Tyr Leu
1 5 10 15
```

Ala

## (2)SEQ ID NO:11 的资料:

## (i)序列特征:

(A)长度: 7 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:11:

Gly Ala Ser Thr Arg Glu Ser

1 5

(2)SEQ ID NO:12 的资料:

(i)序列特征:

(A)长度: 9 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:12:

Gln Asn Val His Ser Phe Pro Phe Thr

1 5

(2)SEQ ID NO:13 的资料:

(i)序列特征:

(A)长度: 9 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:13:

Gln Asn Asp His Ser Phe Pro Phe Thr

1 5

(2)SEQ ID NO:14 的资料:

(i)序列特征:

(A)长度: 11 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:14:



Gly Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln  
                   35                                  40                                  45

Pro Pro Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg Glu Ser Gly Val  
           50                                  55                                  60

Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Ser  
   65                                  70                                  75                                  80

Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn  
                                   85                                  90                                  95

Val His Ser Phe Pro Phe Thr Phe Gly Ser Gly Thr Glu Leu Glu Ile  
                   100                                  105                                  110

Lys

(2)SEQ ID NO:17 的资料:

(i)序列特征:

- (A)长度: 60 个碱基对
- (B)类型: 核酸
- (C)链型: 单链
- (D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:17:

ATGGTGTTC AGACCCAGGT CTCATTCT CTGTTGCTCT GGATCTCTGG TGCCTACGGG 60

(2)SEQ ID NO:18 的资料:

(i)序列特征:

- (A)长度: 357 个碱基对
- (B)类型: 核酸
- (C)链型: 单链
- (D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:18:

CAGTTACCC TGCGTGAATC CGGTCCGGCA CTAGTTAAAC CGACCCAGAC CCTGACGTTA 60  
 ACCTGCACCG TCTCCGGTTT CTCCTGACG AGCTATAGTG TACACTGGGT CCGTCAGCCG 120

CCGGGTAAAG GTCTAGAATG GCTGGGTGTA ATATGGGCTA GTGGAGGCAC AGATTATAAT 180  
 TCGGCTCTCA TGTCCCGTCT GTCGATATCC AAAGACACCT CCCGTAACCA GGTTGTTCTG 240  
 ACCATGACTA ACATGGACCC GGTTGACACC GCTACCTACT ACTGCGCTCG AGATCCCCCT 300  
 TCTTCCTTAC TACGGCTTGA CTACTGGGGT CGTGGTACCC CAGTTACCGT GAGCTCA 357

## (2)SEQ ID NO:19 的资料:

## (i)序列特征:

(A)长度: 119 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:19:

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

## (2)SEQ ID NO:20 的资料:

## (i)序列特征:

(A)长度: 339 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: DNA (基因组)

## (xi)序列描述: SEQ ID NO:20:

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| GATATCGTGA TGACCCAGTC TCCAGACTCG CTAGCTGTGT CTCTGGGCGA GAGGGCCACC | 60  |
| ATCAACTGCA AGAGCTCTCA GAGTCTGTTA AACAGTGGAA ATCAAAAGAA CTACTTGGCC | 120 |
| TGGTATCAGC AGAAACCCGG GCAGCCTCCT AAGTTGCTCA TTTACGGGGC GTCGACTAGG | 180 |
| GAATCTGGGG TACCTGACCG ATTCAGTGGC AGCGGGTCTG GGACAGATTT CACTCTCACC | 240 |
| ATCAGCAGCC TGCAGGCTGA AGATGTGGCA GTATACTACT GTCAGAATGT TCATAGTTTT | 300 |
| CCATTACCGT TCGGCGGAGG GACCAAGTTG GAGATCAAA                        | 339 |

## (2)SEQ ID NO:21 的资料:

## (i)序列特征:

(A)长度: 113 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:21:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Ile | Val | Met | Thr | Gln | Ser | Pro | Asp | Ser | Leu | Ala | Val | Ser | Leu | Gly |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     | 15  |     |     |

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Asn Ser

20

25

30

Gly Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln

35

40

45

Pro Pro Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg Glu Ser Gly Val

50

55

60

Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr

65

70

75

80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Asn

85

90

95

Val His Ser Phe Pro Phe Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile

100

105

110

Lys

(2)SEQ ID NO:22 的资料:

(i)序列特征:

(A)长度: 29 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:22:

GTACATATGC AAGGCTTACA ACCACAATC

29

(2)SEQ ID NO:23 的资料:

(i)序列特征:

- (A)长度: 32 个碱基对
- (B)类型: 核酸
- (C)链型: 单链
- (D)拓扑结构: 线性
- (ii)分子类型: DNA (基因组)
- (xi)序列描述: SEQ ID NO:23:

GGACAGGGCT TACTAGTGGG CCCTCTGGGC TC

32

(2)SEQ ID NO:24 的资料:

- (i)序列特征:
  - (A)长度: 23 个碱基对
  - (B)类型: 核酸
  - (C)链型: 单链
  - (D)拓扑结构: 线性
- (ii)分子类型: DNA (基因组)
- (xi)序列描述: SEQ ID NO:24:

AGGTSMARCT KTCTCGAGTC WGG

23

(2)SEQ ID NO:25 的资料:

- (i)序列特征:
  - (A)长度: 36 个碱基对
  - (B)类型: 核酸
  - (C)链型: 单链
  - (D)拓扑结构: 线性
- (ii)分子类型: DNA (基因组)
- (xi)序列描述: SEQ ID NO:25:

CTAACACTCA TTCCTGTTGA AGCTCTTGAC AATGGG

36

(2)SEQ ID NO:26 的资料:

- (i)序列特征:
  - (A)长度: 32 个碱基对
  - (B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:26:

CCAGATGTGA GCTCGTGATG ACCCAGACTC CA

32

(2)SEQ ID NO:27 的资料:

(i)序列特征:

(A)长度: 140 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:27:

AACCTGCACC GTCTCCGGTT TCTCCCTGAC GAGCTATAGT GTACTGTTGG TCCGTCAGCC 60

GCCGGGTAAA GGTCTAGAAT GGCTGGGTGT AATATGGGCT AGTGGAGGCA CAGATTATAA 120

TTCGGCTCTC ATGTCCCGTC 140

(2)SEQ ID NO:28 的资料:

(i)序列特征:

(A)长度: 149 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:28:

ATATCGACAG ACGGGACATG AGAGCCGAAT TATAATCTGT GCCTCCACTA GCCCATATTA 60

CACCCAGCCA TTCTAGACCT TTACCCGGCG GCTGACGGAC CCAGTGTACA CTATAGCTCG 120

TCAGGGAGAA ACCGGAGACG GTGCAGGTT 149

(2)SEQ ID NO:29 的资料:

## (i)序列特征:

(A)长度: 139 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: DNA (基因组)

## (xi)序列描述: SEQ ID NO:29:

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| TGTCGATATC CAAAGACACC TCCCGTAACC AGGTTGTTCT GACCATGACT AACATGGACC | 60  |
| CGGTTGACAC CGCTACCTAC TACTGCGCTC GAGATCCCCC TTCTTCCTTA CTACGGCTTG | 120 |
| ACTACTGGGG TCGTGGTAC                                              | 139 |

## (2)SEQ ID NO:30 的资料:

## (i)序列特征:

(A)长度: 126 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: DNA (基因组)

## (xi)序列描述: SEQ ID NO:30:

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| CACGACCCCA GTAGTCAAGC CGTAGTAAGG AAGAAGGGGG ATCTCGAGCG CAGTAGTAGG | 60  |
| TAGCGGTGTC AACCGGGTCC ATGTTAGTCA TGGTCAGAAC AACCTGGTTA CGGGAGGTGT | 120 |
| CTTTGG                                                            | 126 |

## (2)SEQ ID NO:31 的资料:

## (i)序列特征:

(A)长度: 117 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:31:

CTCAGAGTCT GTTAAACAGT GGAAATCAAA AGAACTACTT GGCCTGGTAT CAGCAGAAAC 60

CCGGGCAGCC TCCTAAGTTG CTCATTTACG GGGCGTCGAC TAGGGAATCT GGGGTAC 117

(2)SEQ ID NO:32 的资料:

(i)序列特征:

(A)长度: 117 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:32:

CCCAGATTCC CTAGTCGACG CCCCCTAAAT GAGCAACTTA GGAGGCTGCC CGGGTTTCTG 60

CTGATACCAG GCCAAGTAGT TCTTTTGATT TCCACTGTTT AACAGACTCT GAGAGCT 117

(2)SEQ ID NO:33 的资料:

(i)序列特征:

(A)长度: 102 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:33:

GCTGAAGATG TGGCAGTATA CTA CTGTCAG AATGTTTCATA GTTTTCCATT CACGTTCCGGC 60

GGAGGGACCA AGTTGGAGAT CAAACGTACT GTGGCGGCGC CA 102

(2)SEQ ID NO:34 的资料:

(i)序列特征:

(A)长度: 111 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:34:

AGCTTGGCGC CGCCACAGTA CGTTTGATCT CCAACTTGGT CCCTCCGCCG AACGTGAATG 60

GAAAACATATG AACATTCTGA CAGTAGTATA CTGCCACATC TTCAGCCTGC A 111

(2)SEQ ID NO:35 的资料:

(i)序列特征:

(A)长度: 82 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:35:

ATGGTGTTC AGACCCAGGT CTTCAATTTCT CTGTTGCTCT GGATCTCTGG TGCCTACGGG 60

CAGGTTCAAC TGAAAGAGTC AG 82

(2)SEQ ID NO:36 的资料:

(i)序列特征:

(A)长度: 89 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:36:

GTCCTGACTC TTTCAGTTGA ACCTGCCCCGT AGGCACCAGA GATCCAGAGC AACAGAGAAA 60

TGAAGACCTG GGTCTGCAAC ACCATGTTG 89

(2)SEQ ID NO:37 的资料:

(i)序列特征:

(A)长度: 45 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:37:

CAAGGCACCA CTCTCACAGT CTCCTCAGCT AGTACGAAGG GCCCA

45

(2)SEQ ID NO:38 的资料:

(i)序列特征:

(A)长度: 43 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:38:

AGCTTGGGCC CTTGCTACTA GCTGAGGAGA CTGTGAGTGG TGC

43

(2)SEQ ID NO:39 的资料:

(i)序列特征:

(A)长度: 28 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:39:

ATCGTGATGA CCCAGTCTCC ATCCTCCC

28

## (2)SEQ ID NO:40 的资料:

## (i)序列特征:

(A)长度: 31 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:40:

TCAGGGAGGA TGGAGACTGG GTCATCACGA T

31

## (2)SEQ ID NO:41 的资料:

## (i)序列特征:

(A)长度: 43 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:41:

TCGGGGGACA GAGTTGGAAA TAAAACGTAC TGTGGCGGCG CCA

43

## (2)SEQ ID NO:42 的资料:

## (i)序列特征:

(A)长度: 42 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:42:

AGCTTGGCGC CGCCACAGTA CGTTTTATTT CCAACTCTGT CC

42

## (2)SEQ ID NO:43 的资料:

## (i)序列特征:



Gly Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln  
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg Glu Ser Gly Val  
 50 55 60

Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr  
 65 70 75 80

Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn  
 85 90 95

Asp His Ser Tyr Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile  
 100 105 110

Lys

(2)SEQ ID NO:46 的资料:

(i)序列特征:

(A)长度: 113 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:46:

Glu Leu Val Met Thr Gln Ser Pro Ser Ser Leu Ser Val Ser Ala Gly  
 1 5 10 15

Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser Leu Leu Asn Ser  
 20 25 30

Gly Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln  
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg Glu Ser Gly Val  
 50 55 60

```

Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
65 70 75 80

Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys Gln Asn
 85 90 95

Asp Tyr Ser Tyr Pro Phe Thr Phe Gly Ser Gly Thr Lys Leu Glu Ile
 100 105 110

```

Lys

(2)SEQ ID NO:47 的资料:

(i)序列特征:

(A)长度: 9 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:47:

```

Gln Asn Asp His Ser Tyr Pro Phe Thr
1 5

```

(2)SEQ ID NO:48 的资料:

(i)序列特征:

(A)长度: 9 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:48:

```

Gln Asn Asp Tyr Ser Tyr Pro Phe Thr
1 5

```

(2)SEQ ID NO:49 的资料:

(i)序列特征:

(A)长度: 6285 个碱基对

- (B)类型: 核酸  
(C)链型: 双链  
(D)拓扑结构: 环状  
(ii)分子类型: DNA (基因组)  
(xi)序列描述: SEQ ID NO:49:

```
GACGTCGCGG CCGCTCTAGG CCTCCAAAA AGCCTCCTCA CTA CTCTCTGG AATAGCTCAG 60
AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGCATGGGGC 120
GGAGAATGGG CGGA ACTGGG CGGAGTTAGG GGC GGGATGG GCGGAGTTAG GGGCGGGACT 180
ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG 240
GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT 300
GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAT TAATTCCCGG 360
GGATCGATCC GTCGACGTAC GACTAGTTAT TAATAGTAAT CAATTACGGG GTCATTAGTT 420
CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA 480
CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA 540
ATAGGGACTT TCCATTGACG TCAATGGGTG GACTATTTAC GGTA AACTGC CCACTTGGCA 600
GTACATCAAG TGTATCATAT GCCAAGTACG CCCCTATTG ACGTCAATGA CGGTAAATGG 660
CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC 720
TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT GGCAGTACAT CAATGGGCGT 780
```

|                                                                     |      |
|---------------------------------------------------------------------|------|
| GGATAGCGGT TTGACTCAGG GGGATTTCCTA AGTCTCCACC CCATTGACGT CAATGGGAGT  | 340  |
| TTGTTTTGGC ACCAAAATCA ACGGGACTTT CCAAAAATGTC GTAACAACCTC CGCCCCATTG | 900  |
| ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA TAAGCAGAGC TGGGTACGTG   | 960  |
| AACCGTCAGA TCGCCTGGAG ACGCCATCGA ATTCGAGGAC GCCAGCACA TGGTGTGCA     | 1020 |
| GACCCAGGTC TTCATTTCTC TGTGCTCTG GATCTCTGGT GCCTACGGGC AGGTTACCCT    | 1080 |
| GCGTGAATCC GGTCCGGCAC TAGTTAAACC GACCCAGACC CTGACGTAA CCTGCACCGT    | 1140 |
| CTCCGGTTTC TCCCTGACGA GCTATAGTGT ACACTGGGTC CGTCAGCCGC CGGGTAAAGG   | 1200 |
| TCTAGAATGG CTGGGTGTAA TATGGGCTAG TGGAGGCACA GATTATAATT CGGCTCTCAT   | 1260 |
| GTCCCGTCTG TCGATATCCA AAGACACCTC CCGTAACCAG GTTGTCTGA CCATGACTAA    | 1320 |
| CATGGACCCG GTTGACACCG CTACCTACTA CTGCGCTCGA GATCCCCCTT CTTCCTTACT   | 1380 |
| ACGGCTTGAC TACTGGGGTC GTGGTACCCC AGTTACCGTG AGCTCAGCTA GTACCAAGGG   | 1440 |
| CCCATCGGTC TTCCCCCTGG CACCCTCCTC CAAGAGCACC TCTGGGGGCA CAGCGGCCCT   | 1500 |
| GGGCTGCCTG GTCAGGACT ACTTCCCCGA ACCGGTGACG GTGTCGTGGA ACTCAGGCGC    | 1560 |
| CCTGACCAGC GCGGTGCACA CCTTCCCGGC TGTCTACAG TCCTCAGGAC TCTACTCCCT    | 1620 |
| CAGCAGCGTG GTGACCGTGC CCTCCAGCAG CTGGGGCACC CAGACCTACA TCTGCAACGT   | 1680 |
| GAATCACAAAG CCCAGCAACA CCAAGGTGGA CAAGAGAGTT GAGCCCCAAT CTTGTGACAA  | 1740 |
| AACTCACACA TGCCCACCGT GCCCAGCACC TGAACCTCTG GGGGGACCGT CAGTCTTCCT   | 1800 |
| CTTCCCCCA AAACCAAGG ACACCCTCAT GATCTCCCGG ACCCCTGAGG TCACATGCGT     | 1860 |
| GGTGGTGGAC GTGAGCCACG AAGACCCTGA GGTCAAGTTC AACTGGTACG TGGACGGCGT   | 1920 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| GGAGGTGCAT AATGCCAAGA CAAAGCCCGG GGAGGAGCAG TACAACAGCA CGTACCSTGT | 1980 |
| GGTCAGCCTC CTCACCGTCC TGCACCAGGA CTGGCTGAAT GGCAAGGAGT ACAAGTGCAA | 2040 |
| GGTCTCCAC AAAGCCCTCC CAGCCCCCAT CGAGAAAACC ATCTCCAAAG CCAAAGGGCA  | 2100 |
| GCCCCGAGAA CCACAGGTGT ACACCCTGCC CCCATCCCGG GAGGAGATGA CCAAGAACCA | 2160 |
| GGTCAGCCTG ACCTGCCTGG TCAAAGGCTT CTATCCCAGC GACATCCCGG TGGAGTGGGA | 2220 |
| GAGCAATGGG CAGCCGGAGA ACAACTACAA GACCACGCCT CCGTGCTGG ACTCCGACGG  | 2280 |
| CTCCTTCTTC CTCTATAGCA AGCTCACCGT GGACAAGAGC AGGTGGCAGC AGGGGAACGT | 2340 |
| CTTCTCATGC TCCGTGATGC ATGAGGCTCT GCACAACCAC TACACGCAGA AGAGCCTCTC | 2400 |
| CCTGTCTCCG GGTAAGTGAG TGTAGTCTAG ATCTACGTAT GATCAGCCTC GACTGTGCCT | 2460 |
| TCTAGTTGCC AGCCATCTGT TGFTTGCCCC TCCCCCGTGC CTTCCCTGAC CCTGGAAGGT | 2520 |
| GCCACTCCCA CTGTCCCTTC CTAATAAAAT GAGGAAATTG CATCGCATTG TCTGAGTAGG | 2580 |
| TGTCATTCTA TTCTGGGGGG TGGGGTGGGG CAGGACAGCA AGGGGGAGGA TTGGGAAGAC | 2640 |
| AATAGCAGGC ATGCTGGGGA TGCGGTGGGC TCTATGGAAC CAGCTGGGGC TCGACAGCGC | 2700 |
| TGGATCTCCC GATCCCCAGC TTTGCTTCTC AATTTCTTAT TTGCATAATG AGAAAAAAG  | 2760 |
| GAAAATTAAT TTAAACACCA ATTCAGTAGT TGATTGAGCA AATGCGTTGC CAAAAAGGAT | 2820 |
| GCTTTAGAGA CAGTGTCTC TGCACAGATA AGGACAAACA TTATTCAGAG GGAGTACCCA  | 2880 |
| GAGCTGAGAC TCCTAAGCCA GTGAGTGGCA CAGCATTCTA GGGAGAAATA TGCTTGTCAT | 2940 |
| CACCGAAGCC TGATTCCGTA GAGCCACACC TTGGTAAGGG CCAATCTGCT CACACAGGAT | 3000 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| AGAGAGGGCA GGAGCCAGGG CAGAGCATAT AAGGTGAGGT AGGATCAGTT GCTCCTCACA | 3060 |
| TTTGCTTCTG ACATAGTTGT GTTGGGAGCT TGGATAGCTT GGACAGCTCA GGGCTGCGAT | 3120 |
| TTCCGGCCAA ACTTGACGGC AATCCTAGCG TGAAGGCTGG TAGGATTTTA TCCCCGCTGC | 3180 |
| CATCATGGTT CGACCATTGA ACTGCATCGT CGCCGTGTCC CAAAATATGG GGATTGGCAA | 3240 |
| GAACGGAGAC CTACCCTGSC CTCCGCTCAG GAACGAGTTC AAGTACTTCC AAAGAATGAC | 3300 |
| CACAACCTCT TCAGTGGAA GTAACAGAA TCTGGTGATT ATGGGTAGGA AAACCTGGTT   | 3360 |
| CTCCATTCCCT GAGAAGAATC GACCTTTAA GGACAGAATT AATATAGTTC TCAGTAGAGA | 3420 |
| ACTCAAAGAA CCACCACGAG GAGCTCATT TCTTGCCAAA AGTTTGGATG ATGCCTTAAG  | 3480 |
| ACTTATTGAA CACCCGAAT TGGCAAGTAA AGTAGACATG GTTGGATAG TCGGAGGCAG   | 3540 |
| TTCTGTTTAC CAGGAAGCCA TGAATCAACC AGGCCACCTT AACTCTTTG TGACAAGGAT  | 3600 |
| CATGCAGGAA TTTGAAAGTG ACACGTTTTT CCCAGAAATT GATTTGGGGA AATATAAACT | 3660 |
| TCTCCCAGAA TACCCAGGCG TCCTCTCTGA GGTCCAGGAG GAAAAAGGCA TCAAGTATAA | 3720 |
| GTTTGAAGTC TACGAGAAGA AAGACTAACA GGAAGATGCT TTCAAGTTCT CTGCTCCCCT | 3780 |
| CCTAAAGCTA TGCATTTTTA TAAGACCATG GGACTTTTGC TGGCTTTAGA TCAGCCTCGA | 3840 |
| CTGTGCCTTC TAGTTGCCAG CCATCTGTTG TTTGCCCCCT CCCCCTGCCT TCCTTGACCC | 3900 |
| TGGAAGGTGC CACTCCCCT GTCCTTTCCT AATAAAATGA GGAATTTGCA TCGCATTGTC  | 3960 |
| TGAGTAGGTG TCATCTATT CTGGGGGGTG GGGTGGGGCA GGACAGCAAG GGGGAGGATT  | 4020 |
| GGGAAGACAA TAGCAGGCAT GCTGGGGATG CGGTGGGCTC TATGGAACCA GCTGGGGCTC | 4080 |
| GATCGAGTGT ATGACTGCGG CCGGATCCC GTCGAGAGCT TGGCGTAATC ATGGTCATAG  | 4140 |

---

|                                                                    |      |
|--------------------------------------------------------------------|------|
| CTGTTTCCTG TGTGAAATTG TTATCCGCTC ACAATTCCAC ACAACATACG AGCCGGAAGC  | 4200 |
| ATAAAGTGTA AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT TGCCTTGCGC  | 4260 |
| TCACTGCCCC CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG AATCGGCCAA  | 4320 |
| CGCGCGGGGA GAGGCGGTTT GCGTATTGGG CGCTCTTCCG CTTCCCTCGCT CACTGACTCG | 4380 |
| CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG  | 4440 |
| TTATCCACAG AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG  | 4500 |
| GCCAGGAACC GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC  | 4560 |
| GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA  | 4620 |
| TACCAGGCGT TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT  | 4680 |
| ACCGGATACC TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA ATGCTCACGC  | 4740 |
| TGTAGGTATC TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC  | 4800 |
| CCCGTTCAGC CCGACCGCTG CGCCTTATCC GGTAACTATC GTCTTGAGTC CAACCCGGTA  | 4860 |
| AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT  | 4920 |
| GTAGGCGGTG CTACAGAGTT CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGGACA  | 4980 |
| GTATTTGGTA TCTGCGCTCT GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGSTAGCTCT  | 5040 |
| TGATCCGGCA AACAAACCAC CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT  | 5100 |
| ACGCGCAGAA AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT  | 5160 |
| CAGTGAACG AAAACTCAG TTAAGGGATT TTGGTCATGA GATTATCAA AAGGATCTTC     | 5220 |

|              |            |            |             |            |            |      |
|--------------|------------|------------|-------------|------------|------------|------|
| ACCTAGATCC   | TTTTAAATTA | AAAATGAAGT | TTTAAATCAA  | TCTAAAGTAT | ATATGAGTAA | 5280 |
| ACTTGGTCTG   | ACAGTTACCA | ATGCTTAATC | AGTGAGGCAC  | CTATCTCAGC | GATCTGTCTA | 5340 |
| TTTCGTTTTCAT | CCATAGTTGC | CTGACTCCCC | GTCGTGTAGA  | TAACTACGAT | ACGGGAGGGC | 5400 |
| TTACCATCTG   | GCCCCAGTGC | TGCAATGATA | CCGCGAGACC  | CACGCTCACC | GGCTCCAGAT | 5460 |
| TTATCAGCAA   | TAAACCAGCC | AGCCGGAAGG | GCCGAGCGCA  | GAAGTGGTCC | TGCAACTTTA | 5520 |
| TCCGCCTCCA   | TCCAGTCTAT | TAATTGTTGC | CGGGAGGCTA  | GAGTAAGTAG | TTCGCCAGTT | 5580 |
| AATAGTTTGC   | GCAACGTTGT | TGCCATTGCT | ACAGGCATCG  | TGGTGTACAG | CTCGTCGTTT | 5640 |
| GGTATGGCTT   | CATTCAGCTC | CGGTTCCTCA | CGATCAAGGC  | GAGTTACATG | ATCCCCCATG | 5700 |
| TTGTGCAAAA   | AAGCGGTTAG | CTCCTTCGGT | CCTCCGATCG  | TTGTCAGAAG | TAAGTTGGCC | 5760 |
| GCAGTGTAT    | CACTCATGGT | TATGGCAGCA | CTGCATAATT  | CTCTTACTGT | CATGCCATCC | 5820 |
| GTAAGATGCT   | TTTCTGTGAC | TGGTGAGTAC | TCAACCAAGT  | CATTCTGAGA | ATAGTGTATG | 5880 |
| CGGCGACCGA   | GTTGCTCTTG | CCCGGCGTCA | -ATACGGGATA | ATACCGCGCC | ACTTAGCAGA | 5940 |
| ACTTTAAAG    | TGCTCATCAT | TGGAAAACGT | TCTTCGGGGC  | GAAACTCTC  | AAGGATCTTA | 6000 |
| CCGCTGTTGA   | GATCCAGTTC | GATGTAACCC | ACTCGTGCAC  | CCAAGTATC  | TTCAGCATCT | 6060 |
| TTTACTTTCA   | CCAGCGTTTC | TGGGTGAGCA | AAAACAGGAA  | GGCAAAATGC | CGCAAAAAG  | 6120 |
| GGAATAAGGG   | CGACACGGAA | ATGTTGAATA | CTCATACTCT  | TCCTTTTTC  | ATATTATTGA | 6180 |
| AGCATTTATC   | AGGGTTATTG | TCTCATGAGC | GGATACATAT  | TTGAATGTAT | TTAGAAAAAT | 6240 |
| AAACAAATAG   | GGGTTCGCG  | CACATTTCCC | CGAAAAGTGC  | CACCT      |            | 6285 |

## (2)SEQ ID NO:50 的资料:

## (i)序列特征:

(A)长度: 5703 个碱基对

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 环状

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:50:

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| GACGTCGCGG CCGCTCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG  | 60  |
| AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGCATGGGGC  | 120 |
| GGAGAATGGG CGGAACTGGG CGGAGTTAGG GCGGGGATGG GCGGAGTTAG GGGCGGGACT  | 180 |
| ATGGTTGCTG ACTAATTGAG ATGCATGCTT TGCATACTTC TGCCTGCTGG GGAGCCTGGG  | 240 |
| GACTTTCCAC ACCTGGTTGC TGACTAATTG AGATGCATGC TTTGCATACT TCTGCCTGCT  | 300 |
| GGGGAGCCTG GGGACTTTCC ACACCCTAAC TGACACACAT TCCACAGAAT TAATTCCCGG  | 360 |
| GGATCGATCC GTCGACGTAC GACTAGTTAT TAAATAGTAAT CAATTACGGG GTCATTAGTT | 420 |
| CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG TAAATGGCCC GCCTGGCTGA  | 480 |
| CCGCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT ATGTTCCCAT AGTAACGCCA   | 540 |
| ATAGGGACTT TCCATTGACG TCAATGGGTG GACTATTTAC GGTAAACTGC CCACTTGCCA  | 600 |
| GTACATCAAG TGTATCATAT GCCAAGTACG CCCCTATTG ACGTCAATGA CGGTAAATGG   | 660 |
| CCCGCCTGGC ATTATGCCCA GTACATGACC TTATGGGACT TTCCTACTTG GCAGTACATC  | 720 |

|             |            |             |             |            |             |      |
|-------------|------------|-------------|-------------|------------|-------------|------|
| TACGTATTAG  | TCATCGCTAT | TACCATGGTG  | ATGCGGTTTT  | GGCAGTACAT | CAATGGGCGT  | 780  |
| GGATAGCGGT  | TTGACTCACG | GGGATTTCCA  | AGTCTCCACC  | CCATTGACGT | CAATGGGAGT  | 840  |
| TTGTTTTGGC  | ACCAAAATCA | ACGGGACTTT  | CCAAAATGTC  | GTAACAACTC | CGCCCCATTG  | 900  |
| ACGCAAATGG  | GCGGTAGGCG | TGTACGGTGG  | GAGGTCTATA  | TAAGCAGAGC | TGGGTACGTG  | 960  |
| AACCGTCAGA  | TCGCCTGGAG | ACGCCATCGA  | ATTCATTGAT  | AGGATCCAGC | AAGATGGTGT  | 1020 |
| TGCAGACCCA  | GGTCTTCATT | TCTCTGTTGC  | TCTGGATCTC  | TGCTGCCTAC | GGGGATATCG  | 1080 |
| TGATGACCCA  | GTCTCCAGAC | TCGCTAGCTG  | TGTCTCTGGG  | CGAGAGGGCC | ACCATCAACT  | 1140 |
| GCAAGAGCTC  | TCAGAGTCTG | TTAAACAGTG  | GAAATCAAAA  | GAACTACTTG | GCCTGGTATC  | 1200 |
| AGCAGAAAAC  | CGGGCAGCCT | CCTAAGTTGC  | TCATTTACGG  | GGCGTCGACT | AGGGAATCTG  | 1260 |
| GGGTACCTGA  | CCGATTCAGT | GGCAGCGGGT  | CTGGGACAGA  | TTTCACTCTC | ACCATCAGCA  | 1320 |
| GCCTGCAGGC  | TGAAGATGTG | GCAGTATACT  | ACTGTCAGAA  | TGTTCATAGT | TTTCCATTCA  | 1380 |
| CGTTCCGGCG  | AGGGACCAAG | TTGGAGATCA  | AACGTA CTGT | GGCGGCGCCA | TCTGTCTTCA  | 1440 |
| TCTTCCCGCC  | ATCTGATGAG | CAGTTGAAAT  | CTGGA ACTGC | CTCTGTTGTG | TGCCTGCTGA  | 1500 |
| ATAACTTCTA  | TCCCAGAGAG | GCCAAAGTAC  | AGTGG AAGGT | GGATAACGCC | CTCCAA TCGG | 1560 |
| GTA ACTCCCA | GGAGAGTGTC | ACAGAGCAGG  | ACAGCAAGGA  | CAGCACCTAC | AGCCTCAGCA  | 1620 |
| GCACCC TGAC | GCTGAGCAAA | GCAGACTACG  | AGAAACACAA  | AGTCTACGCC | TGCGAAGTCA  | 1680 |
| CCCATCAGGG  | CCTGAGCTCG | CCCGTCACAA  | AGAGCTTCAA  | CAGGGGAGAG | TGTTAATTCT  | 1740 |
| AGATCCGTTA  | TCTACGTATG | ATCAGCCTCG  | ACTGTGCCTT  | CTAGTTGCCA | GCCATCTGTT  | 1800 |
| GTTTGCCCCCT | CCCCCGTGCC | TTCCCTTGACC | CTGGAAGGTG  | CCACTCCCAC | TGTCCTTTCC  | 1860 |

---

|                                                                   |      |
|-------------------------------------------------------------------|------|
| TAATAAAATG AGGAAATTC ATCGCATTGT CTGAGTAGGT GTCATTCTAT TCTGGGGGGT  | 1920 |
| GGGGTGGGGC AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA TGCTGGGGAT | 1980 |
| GCGGTGGGCT CTATGGAACC AGCTGGGGCT CGACAGCTCG AGCTAGCTTT GCTTCTCAAT | 2040 |
| TTCTTATTTG CATAATGAGA AAAAAAGGAA AATTAATTTT AACACCAATT CAGTAGTTGA | 2100 |
| TTGAGCAAAT GCGTTGCCAA AAAGGATGCT TTAGAGACAG TGTTCTCTGC ACAGNTAAGG | 2160 |
| ACAAACATTA TTCAGAGGGA GTACCCAGAG CTGAGACTCC TAAGCCAGTG AGTGGCACAG | 2220 |
| CATTCTAGGG AGAAATATGC TTGTCATCAC CGAAGCCTGA TTCCGTAGAG CCACACCTTG | 2280 |
| GTAAGGGCCA ATCTGCTCAC ACAGGATAGA GAGGGCAGGA GCCAGGGCAG ACCTATAAG  | 2340 |
| GTGAGGTAGG ATCAGTTGCT CCTCACATTT GCTTCTGACA TAGTTGTGTT GGGAGCTTGG | 2400 |
| ATCGATCCAC CATGGTTGAA CAAGATGGAT TGCACGCAGG TTCTCCGGCC GCTTGGGTGG | 2460 |
| AGAGGCTATT CGGCTATGAC TGGGCACAAC AGACAATCGG CTGCTCTGAT GCCGCCGTGT | 2520 |
| TCCGGCTGTC AGCGCAGGGG CGCCCGGTTT TTTTGTCAA GACCGACCTG TCCGGTGGCC  | 2580 |
| TGAATGAACT GCAGGACGAG GCAGCGCGGC TATCGTGGCT GGCCACGACG GGCCTTCCTT | 2640 |
| GCGCAGCTGT GCTCGACCTT GTCACTGAAG CGGGAAGGGA CTGGCTGCTA TTGGGCGAAG | 2700 |
| TGCCGGGGCA GGATCTCCTG TCATCTCACC TTGCTCCTGC CGAGAAAGTA TCCATCATGG | 2760 |
| CTGATGCAAT GCGGCGGCTG CATACGCTTG ATCCGGCTAC CTGCCCATTC GACCACCAAG | 2820 |
| CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC CGGTCTTGTC GATCAGGATG | 2880 |
| ATCTGGACGA AGAGCATCAG GGGCTCGCGC CAGCCGAACT GTTCGCCAGG CTCAAGGCGC | 2940 |

|             |            |            |             |            |            |      |
|-------------|------------|------------|-------------|------------|------------|------|
| GCATGCCCCGA | CGGCGAGGAT | CTCGTCGTGA | CCCATGGCGA  | TGCCTGCTTG | CCGAATATCA | 3000 |
| TGGTGGAAA   | TGGCCGCTTT | TCTGGATTCA | TCGACTGTGG  | CCGGCTGGGT | GTGGCGGACC | 3060 |
| GCTATCAGGA  | CATAGCGTTG | GCTACCCGTG | ATATTGCTGA  | AGAGCTTGGC | GGCGAATGGG | 3120 |
| CTGACCGCTT  | CCTCGTGCTT | TACGGTATCG | CCGCTCCCGA  | TTCGCAGCGC | ATCGCCTTCT | 3180 |
| ATCGCCTTCT  | TGACGAGTTC | TTCTGAGCGG | GACTCTGGGG  | TTCGAAATGA | CCGACCAAGC | 3240 |
| GACGCCAAC   | CTGCCATCAC | GAGATTTCGA | TTCCACCGCC  | GCCTTCTATG | AAAGSTTGGG | 3300 |
| CTTCGGAAATC | GTTTTCCGGG | ACGCCGGCTG | GATGATCCTC  | CAGCGCGGGG | ATCTCATGCT | 3360 |
| GGAGTTCTTC  | GCCCACCCCA | ACTTGTTTAT | TGCAGCTTAT  | AATGGTTACA | AATAAAGCAA | 3420 |
| TAGCATCACA  | AATTCACAA  | ATAAAGCATT | TTTTTCACTG  | CATTCTAGTT | GTGGTTTGTG | 3480 |
| CAAACTCATC  | AATGTATCTT | ATCATGTCTG | GATCGCGGCC  | GCGATCCCGT | CGAGAGCTTG | 3540 |
| GCGTAATCAT  | GGTCATAGCT | GTTTCCTGTG | TGAAATTGTT  | ATCCGCTCAC | AATCCACAC  | 3600 |
| AACATACGAG  | CCGGAAGCAT | AAAGTGTA   | GCCTGGGGTG  | CCTAATGAGT | GAGCTAACTC | 3660 |
| ACATTAATTG  | CGTTGCGCTC | ACTGCCCGCT | TTCCAGTCGG  | GAAACCTGTC | GTGCCAGCTG | 3720 |
| CATTAATGAA  | TCGGCCAACG | CGCGGGGAGA | GGCGGTTTGC  | GTATTGGGCG | CTCTCCCGCT | 3780 |
| TCCTCGCTCA  | CTGACTCGCT | GCGCTCGGTC | GTTCCGGCTGC | GGCGAGCGGT | ATCAGCTCAC | 3840 |
| TCAAAGGCGG  | TAATACGGTT | ATCCACAGAA | TCAGGGGATA  | ACGCAGGAAA | GAACATGTGA | 3900 |
| GCAAAAGGCC  | AGCAAAGGC  | CAGGAACCGT | AAAAGGCCG   | CGTTGCTGGC | GTTTTCCAT  | 3960 |
| AGGCTCCGCC  | CCCCTGACGA | GCATCACAAA | AATCGACGCT  | CAAGTCAGAG | GTGGCGAAAC | 4020 |
| CCGACAGGAC  | TATAAAGATA | CCAGGCGTTT | CCCCTGGAA   | GCTCCCTCGT | GCGCTCTCCT | 4080 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG  | 4140 |
| CTTTCTCAAT GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG  | 4200 |
| GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCCTGCG COTTATCCGG TAACTATCGT   | 4260 |
| CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG  | 4320 |
| ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC  | 4380 |
| GGCTACACTA GAAGGACAGT ATTTGGTATC TCGCTCTGC TGAAGCCAGT TACCTTCGGG   | 4440 |
| AAAAGAGTTG GTAGCTCTTG ATCCGGCAA CAAACCACCG CTGGTAGCGG TGTTTTTTTT   | 4500 |
| GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT  | 4560 |
| TCTACGGGST CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCAATGAGA | 4620 |
| TTATCAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAT AATGAAGTTT TAAATCAATC   | 4680 |
| TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT  | 4740 |
| ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT GACTCCCCGT CGTGTAGATA  | 4800 |
| ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGCTG CAATGATACC GCGAGACCCA  | 4860 |
| CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC CGAGCGCAGA  | 4920 |
| AGTGGTCCTG CAACTTTATC CGCCTCCATC CAGTCTATTA ATTGTTGCCG GGAAGCTAGA  | 4980 |
| GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG CCATTGCTAC AGGCATCGTG  | 5040 |
| GTGTCACGCT CGTCGTTTGG TATGGCTTCA TTCAGCTCCG GTTCCCAACG ATCAAGGCCA  | 5100 |
| GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT CCTTCGGTCC TCCGATCGTT  | 5160 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| GTCAGAAGTA AGTTGGCCGC AGTGTTATCA CTCATGGTTA TGGCAGCACT GCATAATTCT  | 5220 |
| CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG GTGAGTACTC AACCAAGTCA  | 5280 |
| TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC CGGCGTCAAT ACGGGATAAT  | 5340 |
| ACCGCGCCAC ATAGCAGAAC TTTAAAAGTG CTCATCATTG GAAAACGTTC TTCGGGGCGA  | 5400 |
| AAACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA TGTAACCCAC TCGTGCACCC  | 5460 |
| AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG GGTGAGCAAA AACAGGAAGG  | 5520 |
| CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGGAAAT GTTGAATACT CATACTCTTC  | 5580 |
| CTTTTTCAAT ATTATTGAAG CATTATCAG GGTTATTGTC TCATGAGCGG ATACATATTT   | 5640 |
| GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA CATTTCCTCCG AAAAGTGCCA | 5700 |
| CCT                                                                | 5703 |

## (2)SEQ ID NO:51 的资料:

## (i)序列特征:

(A)长度: 81 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:51:

|                                                                   |    |
|-------------------------------------------------------------------|----|
| ATCCAAAGAC AACTCCCGTA ACCAGGTTGT TCTGACCATG ACTAACATGG ACCCGGTTGA | 60 |
|-------------------------------------------------------------------|----|

|                         |    |
|-------------------------|----|
| CACCGCTACC TACTACTGCG C | 81 |
|-------------------------|----|

## (2)SEQ ID NO:52 的资料:

## (i)序列特征:

(A)长度: 85 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:52:

TCGAGCGCAG TAGTAGGTAG CGGTGTCAAC CGGGTCCATG TTAGTCATGG TCAGAACAAC 60

CTGGTTACGG GAGTTGTCTT TGGAT 85

(2)SEQ ID NO:53 的资料:

(i)序列特征:

(A)长度: 73 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:53:

AACCTGCACC GTCTCCGGTT TCTCCCTGAC GAGCTATAGT GTACTGGA TCCGTCAGCC 60

GCCGGGTAAA GGT 73

(2)SEQ ID NO:54 的资料:

(i)序列特征:

(A)长度: 77 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:54:

CTAGACCTTT ACCCGGCGGC TGACGGATCC AGTGTACT ATAGCTCGTC AGGGAGAAAC 60

CGGAGACGGT GCAGGTT 77

(2)SEQ ID NO:55 的资料:

(i)序列特征:

(A)长度: 46 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:55:

CTAGCTGTGT CAGCTGGCGA GAGGGCCACC ATCAACTGCA AGAGCT

46

(2)SEQ ID NO:56 的资料:

(i)序列特征:

(A)长度: 38 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:56:

CTTGCAGTTG ATGGTGGCCC TCTCGCCAGC TGACACAG

38

(2)SEQ ID NO:57 的资料:

(i)序列特征:

(A)长度: 140 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:57:

TTCGAGGACG CCAGCAACAT GGTGTTGCAG ACCCAGGTCT TCATTTCTCT GTTGCTCTGG

60

ATCTCTGGTG CCTACGGGCA GGTCCAACCTG CAGGAGAGCG GTCCAGGTCT TGTGAGACCT

120

AGCCAGACCC TGAGCCTGAC

140

(2)SEQ ID NO:58 的资料:

(i)序列特征:

(A)长度: 138 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:58:

GTGCCTCCAC TAGCCCATAT TACTCCAAGC CACTCTAGAC CTCGTCCAGG TGGCTGTCTC 60

ACCCAGTGTA CACTATAGCT GGTGAGGGAG AAGCCGAGA CGGTGCAGGT CAGGCTCAGG 120

GTCTGGCTAG GTCTCACA 138

(2)SEQ ID NO:59 的资料:

(i)序列特征:

(A)长度: 143 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:59:

GGCTTGGAGT AATATGGGCT AGTGGAGGCA CAGATTATAA TTCGGCTCTC ATGTCCAGAC 60

TGAGTATACT GAAAGACAAC AGCAAGAACC AGGTCAGCCT GAGACTCAGC AGCGTGACAG 120

CCGCCGACAC CGCGGTCTAT TTC 143

(2)SEQ ID NO:60 的资料:

(i)序列特征:

(A)长度: 136 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:60:

CCAGTGCCAA GCTTGGGCCC TTGGTGGAGG CGCTCGAGAC GGTGACCGTG GTACCTTGTC 60

CCCAGTAGTC AAGCCGTAGT AAGGAAGAAG GGGGATCTCG AGCACAGAAA TAGACCGCGG 120

TGTCGGCGGC TGTCAC 136

(2)SEQ ID NO:61 的资料:

(i)序列特征:

(A)长度: 357 个碱基对

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:61:

CAGGTCCAAC TGCAGGAGAG CGGTCCAGGT CTTGTGAGAC CTAGCCAGAC CCTGAGCCTG 60

ACCTGCACCG TCTCGGGCTT CTCCCTCACC AGCTATAGTG TACTGCGGT GAGACAGCCA 120

CCTGGACGAG GTCTAGAGTG GCTTGGAGTA ATATGGGCTA GTGGAGGCAC AGATTATAAT 180

TCGGCTCTCA TGTCCAGACT GAGTATACTG AAAGACAACA GCAAGAACCA GGTCAGCCTG 240

AGACTCAGCA GCGTGACAGC CGCCGACACC GCGGTCTATT ACTGTGCTCG GGATCCCCCT 300

TCTTCCTTAC TACGGCTTGA CTA CTG GGGGA CAAGGTACCA CGGTCACCGT CTCGAGC 357

(2)SEQ ID NO:62 的资料:

(i)序列特征:

(A)长度: 119 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: 蛋白

(xi)序列描述: SEQ ID NO:62:

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Arg Pro Ser Gln

1

5

10

15

Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Thr Ser Tyr  
 20 25 30

Ser Val His Trp Val Arg Gln Pro Pro Gly Arg Gly Leu Glu Trp Leu  
 35 40 45

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

Ser Arg Leu Ser Ile Leu Lys Asp Asn Ser Lys Asn Gln Val Ser Leu  
 65 70 75 80

Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala  
 85 90 95

Arg Asp Pro Pro Ser Ser Leu Leu Arg Leu Asp Tyr Trp Gly Gln Gly  
 100 105 110

Thr Thr Val Thr Val Ser Ser  
 115

(2)SEQ ID NO:63 的资料:

(i)序列特征:

(A)长度: 28 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:63:

AGGACGCCAG CAACATGGTG TTGCAGAC

28

(2)SEQ ID NO:64 的资料:

(i)序列特征:

(A)长度: 36 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:64:

TGCCAAGCTT GGGCCCTTGG TGGAGGCGCT CGAGAC

36

(2)SEQ ID NO:65 的资料:

(i)序列特征:

(A)长度: 121 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:65:

GACCATGATT ACGAATTCGT AGTCGGATAT CGTGATGACC CAGAGCCCAA GCAGCCTGAG 60

CGCTAGCGTG GGTGACAGAG TGACCATCAC CTGTAAGAGC TCTCAGAGTC TGTAAACAG 120

T 121

(2)SEQ ID NO:66 的资料:

(i)序列特征:

(A)长度: 116 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:66:

AGATTCCCTA GTCGATGCC CGTAGATCAG CAGCTTTGGA GCCTTACCGG GTTTCTGCTG 60

ATACCAGGCC AAGTAGTTCT TTTGATTTCC ACTGTTTAAAC AGACTCTGAG AGCTCT 116

(2)SEQ ID NO:67 的资料:

(i)序列特征:

(A)长度: 116 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:67:

TCTACGGGGC ATCGACTAGG GAATCTGGGG TACCAGATAG ATTCAGCGGT AGCGGTAGCG 60  
GAACCGACTT CACCTTCACC ATCAGCAGCC TGCAGCCAGA GGACATCGCC ACCTAC 116

(2)SEQ ID NO:68 的资料:

(i)序列特征:

(A)长度: 117 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:68:

TCGATGCCAA GCTTGGCGCC GCCACAGTAC GTTTGATCTC CACCTTGGTC CCTTGTC CGA 60  
ACGTGAATGG AAAACTATGA ACATTCTGGC AGTAGTAGGT GGCGATGTCC TCTGGCT 117

(2)SEQ ID NO:69 的资料:

(i)序列特征:

(A)长度: 339 个碱基对

(B)类型: 核酸

(C)链型: 双链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:69:

GATATCGTGA TGACCCAGAG CCCAAGCAGC CTGAGCGCTA GCGTGGGTGA CAGAGTGACC 60  
ATCACCTGTA AGAGCTCTCA GAGTCTGTTA AACAGTGGAA ATCAAAAAGAA CTA CTGAGCC 120  
TGGTATCAGC AGAAACCCGG TAAGGCTCCA AAGCTGCTGA TCTACGGGGC ATCGACTAGG 180

```

GAATCTGGGG TACCAGATAG ATTCAGCGGT AGCGGTAGCG GAACCGACTT CACCTTCACC 240

ATCAGCAGCC TGCAGCCAGA GGACATCGCC ACCTACTACT GCCAGAATGT TCATAGTTTT 300

CCATTACAGT TCGGACAAGG GACCAAGGTG GAGATCAAA 339

```

## (2)SEQ ID NO:70 的资料:

## (i)序列特征:

(A)长度: 113 个氨基酸

(B)类型: 氨基酸

(C)链型: 单链

(D)拓扑结构: 线性

## (ii)分子类型: 蛋白

## (xi)序列描述: SEQ ID NO:70:

```

Asp Ile Val Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15

```

```

Asp Arg Val Thr Ile Thr Cys Lys Ser Ser Gln Ser Leu Leu Asn Ser
 20 25 30

```

```

Gly Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys
 35 40 45

```

```

Ala Pro Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60

```

```

Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr
65 70 75 80

```

```

Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Asn
 85 90 95

```

```

Val His Ser Phe Pro Phe Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110

```

Lys

## (2)SEQ ID NO:71 的资料:

## (i)序列特征:

(A)长度: 24 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:71:

GATTACGAAT TCGTAGTCGG ATAT

24

## (2)SEQ ID NO:72 的资料:

## (i)序列特征:

(A)长度: 24 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:72:

TGCCAAGCTT GCGCCGCCA CAGT

24

## (2)SEQ ID NO:73 的资料:

## (i)序列特征:

(A)长度: 39 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:73:

CTAGTGCGGG TGACCGAGTG ACCATCACCT GTAAGAGCT

39

## (2)SEQ ID NO:74 的资料:

## (i)序列特征:

(A)长度: 31 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:74:

CTTACAGGTG ATGGTCACTC GGTCACCCGC A

31

(2)SEQ ID NO:75 的资料:

(i)序列特征:

(A)长度: 66 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:75:

GGTCTATTAC TGTGCTCGGG ATCCCCCTTC TTCCTTACTA CGGCTTGACT ACTGGGGACA

60

AGGTAC

66

(2)SEQ ID NO:76 的资料:

(i)序列特征:

(A)长度: 64 个碱基对

(B)类型: 核酸

(C)链型: 单链

(D)拓扑结构: 线性

(ii)分子类型: DNA (基因组)

(xi)序列描述: SEQ ID NO:76:

CTTGTCCCCA GTAGTCAAGC CGTAGTAAGG AAGAAGGGGG ATCCCGAGCA CAGTAATAGA

60

CCGC

64

图 1

## 2B6 重链可变区 DNA 序列

|     |     |     |     |     |     |     |     |     |     |         |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------|-----|
| 1   |     |     |     |     |     |     |     |     |     | CCT GGC | 30  |
|     | Gln | Val | Gln | Leu | Lys | Glu | Ser | Gly | Pro | Gly     |     |
| 31  | CTG | GTG | GCG | CCC | TCA | CAG | AGC | CTG | TCC | ATC     | 60  |
|     | Leu | Val | Ala | Pro | Ser | Gln | Ser | Leu | Ser | Ile     |     |
| 61  | ACT | TGC | ACT | GTC | TCT | GGG | TTT | TCA | TTA | ACC     | 90  |
|     | Thr | Cys | Thr | Val | Ser | Gly | Phe | Ser | Leu | Thr     |     |
| 91  | AGC | TAT | AGT | GTA | CAC | TGG | GTT | CGC | CAG | CCT     | 120 |
|     | Ser | Tyr | Ser | Val | His | Trp | Val | Arg | Gln | Pro     |     |
| 121 | CCA | GGA | AAG | GGT | CTG | GAG | TGG | CTG | GGA | GTA     | 150 |
|     | Pro | Gly | Lys | Gly | Leu | Glu | Trp | Leu | Gly | Val     |     |
| 151 | ATA | TGG | GCT | AGT | GGA | GGC | ACA | GAT | TAT | AAT     | 180 |
|     | Ile | Trp | Ala | Ser | Gly | Gly | Thr | Asp | Tyr | Asn     |     |
| 181 | TCG | GCT | CTC | ATG | TCC | AGA | CTG | AGC | ATC | AGC     | 210 |
|     | Ser | Ala | Leu | Met | Ser | Arg | Leu | Ser | Ile | Ser     |     |
| 211 | AAA | GAC | AAC | TCC | AAG | AGC | CAA | GTT | TTC | TTA     | 240 |
|     | Lys | Asp | Asn | Ser | Lys | Ser | Gln | Val | Phe | Leu     |     |
| 241 | AAA | CTG | AAC | AGT | CTG | CAA | ACT | GAT | GAC | ACA     | 270 |
|     | Lys | Leu | Asn | Ser | Leu | Gln | Thr | Asp | Asp | Thr     |     |
| 271 | GCC | ATG | TAC | TAC | TGT | GCC | AGA | GAT | CCC | CCT     | 300 |
|     | Ala | Met | Tyr | Tyr | Cys | Ala | Arg | Asp | Pro | Pro     |     |
| 301 | TCT | TCC | TTA | CTA | CGG | CTT | GAC | TAC | TGG | GGC     | 330 |
|     | Ser | Ser | Leu | Leu | Arg | Leu | Asp | Tyr | Trp | Gly     |     |
| 331 | CAA | GGC | ACC | ACT | CTC | ACA | GTC | TCC | TCA |         | 357 |
|     | Gln | Gly | Thr | Thr | Leu | Thr | Val | Ser | Ser |         |     |

## 图 2

## 2B6 轻链可变区 DNA 序列

|     |                                         |         |    |
|-----|-----------------------------------------|---------|----|
| 1   |                                         | TCC TCC | 30 |
|     | Asp Ile Val Met Thr Gln Ser Pro Ser Ser |         |    |
| 31  | CTG AGT GTG TCA GCA GGA GAG AAG GTC ACT | 60      |    |
|     | Leu Ser Val Ser Ala Gly Glu Lys Val Thr |         |    |
| 61  | ATG AGC TGC AAG TCC AGT CAG AGT CTG TTA | 90      |    |
|     | Met Ser Cys Lys Ser Ser Gln Ser Leu Leu |         |    |
| 91  | AAC AGT GGA AAT CAA AAG AAC TAC TTG GCC | 120     |    |
|     | Asn Ser Gly Asn Gln Lys Asn Tyr Leu Ala |         |    |
| 121 | TGG TAC CAG CAG AAA CCA GGG CAG CCT CCT | 150     |    |
|     | Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro |         |    |
| 151 | AAA CTT TTG ATC TAC GGG GCA TCC ACT AGG | 180     |    |
|     | Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg |         |    |
| 181 | GAA TCT GGG GTC CCT GAT CGC TTC ACA GGC | 210     |    |
|     | Glu Ser Gly Val Pro Asp Arg Phe Thr Gly |         |    |
| 211 | AGT GGA TCT GGA ACC GAT TTC ACT CTT TCC | 240     |    |
|     | Ser Gly Ser Gly Thr Asp Phe Thr Leu Ser |         |    |
| 241 | ATC AGC AGT GTG CAG GCT GAA GAC CTG GCA | 270     |    |
|     | Ile Ser Ser Val Gln Ala Glu Asp Leu Ala |         |    |
| 271 | GTT TAT TAC TGT CAG AAT GTT CAT AGT TTT | 300     |    |
|     | Val Tyr Tyr Cys Gln Asn Val His Ser Phe |         |    |
| 301 | CCA TTC ACG TTC GGC TCG GGG ACA GAG TTG | 330     |    |
|     | Pro Phe Thr Phe Gly Ser Gly Thr Glu Leu |         |    |
| 331 | GAA ATA AAA                             | 339     |    |
|     | Glu Ile Lys                             |         |    |

图 3

## 2F2 重链可变区 DNA 序列

|     |                                                                                                                                     |     |
|-----|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1   | CCT GGC CTG GTG GCG CCC TCA CAG AGC CTG                                                                                             | 30  |
|     | Pro Gly Leu Val Ala Pro Ser Gln Ser Leu                                                                                             |     |
| 31  | TCC ATC ACT TGC ACT GTC TCT GGG TTT TCA                                                                                             | 60  |
|     | Ser Ile Thr Cys Thr Val Ser Gly Phe Ser                                                                                             |     |
| 61  | TTA ACC <span style="border: 1px solid black;">AGT TAT AGT GTA CAC</span> TGG GTT CGC                                               | 90  |
|     | Leu Thr <span style="border: 1px solid black;">Ser Tyr Ser Val His</span> Trp Val Arg                                               |     |
| 91  | CAG CCT CCA GGA AAG GGT CTG GAG TGG CTG                                                                                             | 120 |
|     | Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu                                                                                             |     |
| 121 | GGA <span style="border: 1px solid black;">GTA ATA TGG GCT AGT GGA GGC ACA GAT</span>                                               | 150 |
|     | Gly <span style="border: 1px solid black;">Val Ile Trp Ala Ser Gly Gly Thr Asp</span>                                               |     |
| 151 | <span style="border: 1px solid black;">TAT AAT TCG GCT CTC ATG TCC</span> <span style="border: 1px solid black;">AGA CTG AGC</span> | 180 |
|     | <span style="border: 1px solid black;">Tyr Asn Ser Ala Leu Met Ser</span> <span style="border: 1px solid black;">Arg Leu Ser</span> |     |
| 181 | ATC AGC AAA GAC AAC TCC AAG AGC CAA GTT                                                                                             | 210 |
|     | Ile Ser Lys Asp Asn Ser Lys Ser Gln Val                                                                                             |     |
| 211 | TTC TTA AAA CTG AAC AGT CTG CGA ACT GAT                                                                                             | 240 |
|     | Phe Leu Lys Leu Asn Ser Leu Arg Thr Asp                                                                                             |     |
| 241 | GAC ACA GCC ATG TAC TAC TGT GCC AGA <span style="border: 1px solid black;">GAT</span>                                               | 270 |
|     | Asp Thr Ala Met Tyr Tyr Cys Ala Arg <span style="border: 1px solid black;">Asp</span>                                               |     |
| 271 | <span style="border: 1px solid black;">CCC CCT TCT TCC TTA CTA CGG CTT GAC TAC</span>                                               | 300 |
|     | <span style="border: 1px solid black;">Pro Pro Ser Ser Leu Leu Arg Leu Asp Tyr</span>                                               |     |
| 301 | TGG GGC CAA GGC ACC ACT CTC ACA GTC TCC                                                                                             | 330 |
|     | Trp Gly Gln Gly Thr Thr Leu Thr Val Ser                                                                                             |     |
| 331 | TCA                                                                                                                                 | 333 |
|     | Ser                                                                                                                                 |     |

图 4

## 2F2 轻链可变区 DNA 序列

|     |                                         |     |
|-----|-----------------------------------------|-----|
| 1   | TCC TCC CTG AGT GTG TCA GCA GGA GAG AAG | 30  |
|     | Ser Ser Leu Ser Val Ser Ala Gly Glu Lys |     |
| 31  | GTC ACT ATG AGC TGC AAG TCC AGT CAG AGT | 60  |
|     | Val Thr Met Ser Cys Lys Ser Ser Gln Ser |     |
| 61  | CTA TTA AAC AGT GGA AAT CAA AAG AAC TAC | 90  |
|     | Leu Leu Asn Ser Gly Asn Gln Lys Asn Tyr |     |
| 91  | TTG GCC TGG TAC CAA CAG AAA CCA GGG CAG | 120 |
|     | Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln |     |
| 121 | CCT CCT AAA CTT TTG ATC TAC GGG GCA TCC | 150 |
|     | Pro Pro Lys Leu Leu Ile Tyr Gly Ala Ser |     |
| 151 | ACT AGG GAA TCT GGG GTC CCT GAT CGC TTC | 180 |
|     | Thr Arg Glu Ser Gly Val Pro Asp Arg Phe |     |
| 181 | ACA GGC AGT GGA TCT GGA ACC GAT TTC ACT | 210 |
|     | Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr |     |
| 211 | CTT ACC ATC AGC AGT GTG CAG GCT GAA GAC | 240 |
|     | Leu Thr Ile Ser Ser Val Gln Ala Glu Asp |     |
| 241 | CTG GCA GTT TAT TAC TGT CAG AAT GAT CAT | 270 |
|     | Leu Ala Val Tyr Tyr Cys Gln Asn Asp His |     |
| 271 | AGT TTT CCA TTC ACG TTC GGC TCG GGG ACA | 300 |
|     | Ser Phe Pro Phe Thr Phe Gly Ser Gly Thr |     |
| 301 | GAG TTG GAA ATA AAA                     | 315 |
|     | Glu Leu Glu Ile Lys                     |     |

## 图 5

## 2E3 重链可变区 DNA 序列

|     |                                                                                       |     |
|-----|---------------------------------------------------------------------------------------|-----|
| 1   | CCT GGC CTG GTG GCG CCC TCA CAG AGC CTG                                               | 30  |
|     | Pro Gly Leu Val Ala Pro Ser Gln Ser Leu                                               |     |
| 31  | TCC ATC ACT TGC ACT GTC TCT GGG TTT TCA                                               | 60  |
|     | Ser Ile Thr Cys Thr Val Ser Gly Phe Ser                                               |     |
| 61  | TTA ACC <span style="border: 1px solid black;">AGC TAT AGT GTA CAC</span> TGG GTT CGC | 90  |
|     | Leu Thr <span style="border: 1px solid black;">Ser Tyr Ser Val His</span> Trp Val Arg |     |
| 91  | CAG CCT CCA GGA AAG GGT CTG GAG TGG CTG                                               | 120 |
|     | Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu                                               |     |
| 121 | GGA <span style="border: 1px solid black;">GTA ATC TGG GCT AGT GGA GGC ACA GAT</span> | 150 |
|     | Gly <span style="border: 1px solid black;">Val Ile Trp Ala Ser Gly Gly Thr Asp</span> |     |
| 151 | <span style="border: 1px solid black;">TAT AAT TCG GCT CTC ATG TCC</span> AGA CTG AGC | 180 |
|     | <span style="border: 1px solid black;">Tyr Asn Ser Ala Leu Met Ser</span> Arg Leu Ser |     |
| 181 | ATC AGC AAA GAC AAC TCC AAG AGC CAA GTT                                               | 210 |
|     | Ile Ser Lys Asp Asn Ser Lys Ser Gln Val                                               |     |
| 211 | TTC TTA AAA CTG AAC AGT CTG CAA ACT GAT                                               | 240 |
|     | Phe Leu Lys Leu Asn Ser Leu Gln Thr Asp                                               |     |
| 241 | GAC GCA GCC ATG TAC TAC TGT GCC AGA <span style="border: 1px solid black;">GAT</span> | 270 |
|     | Asp Ala Ala Met Tyr Tyr Cys Ala Arg <span style="border: 1px solid black;">Asp</span> |     |
| 271 | <span style="border: 1px solid black;">CCC CCT TTT TCC TTA CTA CGG CTT GAC TTC</span> | 300 |
|     | <span style="border: 1px solid black;">Pro Pro Phe Ser Leu Leu Arg Leu Asp Phe</span> |     |
| 301 | TGG GGC CAA GGC ACC ACT CTC ACA GTC TCC                                               | 330 |
|     | Trp Gly Gln Gly Thr Thr Leu Thr Val Ser                                               |     |
| 331 | TCA                                                                                   | 333 |
|     | Ser                                                                                   |     |

图 6

## 2E3 轻链可变区 DNA 序列

|     |                                         |     |
|-----|-----------------------------------------|-----|
| 1   | TCC TCT CTG AGT GTG TCA GCA GGA GAG AAG | 30  |
|     | Ser Ser Leu Ser Val Ser Ala Gly Glu Lys |     |
| 31  | GTC ACT ATG AGC TGC AAG TCC AGT CAG AGT | 60  |
|     | Val Thr Met Ser Cys Lys Ser Ser Gln Ser |     |
| 61  | CTG TTA AAC AGT GGA AAT CAA AAA AAC TAC | 90  |
|     | Leu Leu Asn Ser Gly Asn Gln Lys Asn Tyr |     |
| 91  | TTG GCC TGG TAC CAG CAG AAA CCA GGG CAG | 120 |
|     | Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln |     |
| 121 | CCT CCT AAA CTT TTG ATC TAC GGG GCA TCC | 150 |
|     | Pro Pro Lys Leu Leu Ile Tyr Gly Ala Ser |     |
| 151 | ACT AGG GAA TCT GGG GTC CCT GAT CGC TTC | 180 |
|     | Thr Arg Glu Ser Gly Val Pro Asp Arg Phe |     |
| 181 | ACA GGC AGT GGA TCT GGA ACC GAT TTC ACT | 210 |
|     | Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr |     |
| 211 | CTT ACC ATC AGC AGT GTG CAG GCT GAA GAC | 240 |
|     | Leu Thr Ile Ser Ser Val Gln Ala Glu Asp |     |
| 241 | CTG GCA GTT TAT TAC TGT CAG AAT GAT CAT | 270 |
|     | Leu Ala Val Tyr Tyr Cys Gln Asn Asp His |     |
| 271 | AGT TTT CCA TTC ACG TTC GGC TCG GGG ACA | 300 |
|     | Ser Phe Pro Phe Thr Phe Gly Ser Gly Thr |     |
| 301 | GAG TTG GAA ATA AAA                     | 315 |
|     | Glu Leu Glu Ile Lys                     |     |

## 图 7

## 2B6 CDRs

重链 1: SYSV11  
重链 2: VIWASGGTDYNSALMS  
重链 3: DPPSLLRLDY

轻链 1: KSSQSLLNSGNQKNYLA  
轻链 2: GASTRES  
轻链 3: QNVHSFPFT

## 2F2 CDRs

重链 1: SYSVH  
重链 2: VIWASGGTDYNSALMS  
重链 3: DPPSLLRLDY

轻链 1: KSSQSLLNSGNQKNYLA  
轻链 2: GASTRES  
轻链 3: QNDHSFPFT

## 2E3 CDRs

重链 1: SYSVH  
重链 2: VIWASGGTDYNSALMS  
重链 3: DPPSLLRLDF

轻链 1: KSSQSLLNSGNQKNYLA  
轻链 2: GASTRES  
轻链 3: QNDHSFPFT

## 图 8

## IL5 人源化重链可变区

|     |                                         |     |
|-----|-----------------------------------------|-----|
| 1   | CAG GTT ACC CTG CGT GAA TCC GGT CCG GCA | 30  |
|     | Gln Val Thr Leu Arg Glu Ser Gly Pro Ala |     |
| 31  | CTA GTT AAA CCG ACC CAG ACC CTG ACG TTA | 60  |
|     | Leu Val Lys Pro Thr Gln Thr Leu Thr Leu |     |
| 61  | ACC TGC ACC TTC TCC GGT TTC TCC CTG ACG | 90  |
|     | Thr Cys Thr Phe Ser Gly Phe Ser Leu Thr |     |
| 91  | AGC TAT AGT GTA CAC TGG ATC CGT CAG CCG | 120 |
|     | Ser Tyr Ser Val His Trp Ile Arg Gln Pro |     |
| 121 | CCG GGT AAA GGT CTG GAG TGG CTG GGT GTA | 150 |
|     | Pro Gly Lys Gly Leu Glu Trp Leu Gly Val |     |
| 151 | ATA TGG GCT AGT GGA GGC ACA GAT TAT AAT | 180 |
|     | Ile Trp Ala Ser Gly Gly Thr Asp Tyr Asn |     |
| 181 | TCG GCT CTC ATG TCC CGT CTG ACG ATA TCC | 210 |
|     | Ser Ala Leu Met Ser Arg Leu Thr Ile Ser |     |
| 211 | AAA GAC ACC TCC CGT AAC CAG GTT GTT CTG | 240 |
|     | Lys Asp Thr Ser Arg Asn Gln Val Val Leu |     |
| 241 | ACC ATG ACT AAC ATG GAC CCG GTT GAC ACC | 270 |
|     | Thr Met Thr Asn Met Asp Pro Val Asp Thr |     |
| 271 | GCT ACC TAC TAC TGC GCT CGA GAT CCC CCT | 300 |
|     | Ala Thr Tyr Tyr Cys Ala Arg Asp Pro Pro |     |
| 301 | TCT TCC TTA CTA CCG CTT GAC TAC TGG GGT | 330 |
|     | Ser Ser Leu Leu Arg Leu Asp Tyr Trp Gly |     |
| 331 | CGT GGT ACC CCA GTT ACC GTG AGC TCA     | 357 |
|     | Arg Gly Thr Pro Val Thr Val Ser Ser     |     |

图 9

## 人源化轻链可变区

|     |                                         |     |
|-----|-----------------------------------------|-----|
| 1   | GAT ATC GTG ATG ACC CAG TCT CCA GAC TCG | 30  |
|     | Asp Ile Val Met Thr Gln Ser Pro Asp Ser |     |
| 31  | CTA GCT GTG TCT CTG GGC GAG AGG GCC ACC | 60  |
|     | Leu Ala Val Ser Leu Gly Glu Arg Ala Thr |     |
| 61  | ATC AAC TGC                             | 90  |
|     | Ile Asn Cys                             |     |
|     | AAG AGC TCT CAG AGT CTG TTA             |     |
|     | Lys Ser Ser Gln Ser Leu Leu             |     |
| 91  | AAC AGT GGA AAT CAA AAG AAC TAC TTG GCC | 120 |
|     | Asn Ser Gly Asn Gln Lys Asn Tyr Leu Ala |     |
| 121 | TGG TAT CAG CAG AAA CCC GGG CAG CCT CCT | 150 |
|     | Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro |     |
| 151 | AAG TTG CTC ATT TAC                     | 180 |
|     | Lys Leu Leu Ile Tyr                     |     |
|     | GGG GCG TCG ACT AGG                     |     |
|     | Gly Ala Ser Thr Arg                     |     |
| 181 | GAA TCT                                 | 210 |
|     | Glu Ser                                 |     |
|     | GGG GTA CCT GAC CGA TTC AGT GGC         |     |
|     | Gly Val Pro Asp Arg Phe Ser Gly         |     |
| 211 | AGC GGG TCT GGG ACA GAT TTC ACT CTC ACC | 240 |
|     | Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr |     |
| 241 | ATC AGC AGC CTG CAG GCT GAA GAT GTG GCA | 270 |
|     | Ile Ser Ser Leu Gln Ala Glu Asp Val Ala |     |
| 271 | GTA TAC TAC TGT                         | 300 |
|     | Val Tyr Tyr Cys                         |     |
|     | CAG AAT GTT CAT AGT TTT                 |     |
|     | Gln Asn Val His Ser Phe                 |     |
| 301 | CCA TTC ACG                             | 330 |
|     | Pro Phe Thr                             |     |
|     | TTC GGC GGA GGG ACC AAG TTG             |     |
|     | Phe Gly Gly Gly Thr Lys Leu             |     |
| 331 | GAG ATC AAA                             | 339 |
|     | Glu Ile Lys                             |     |

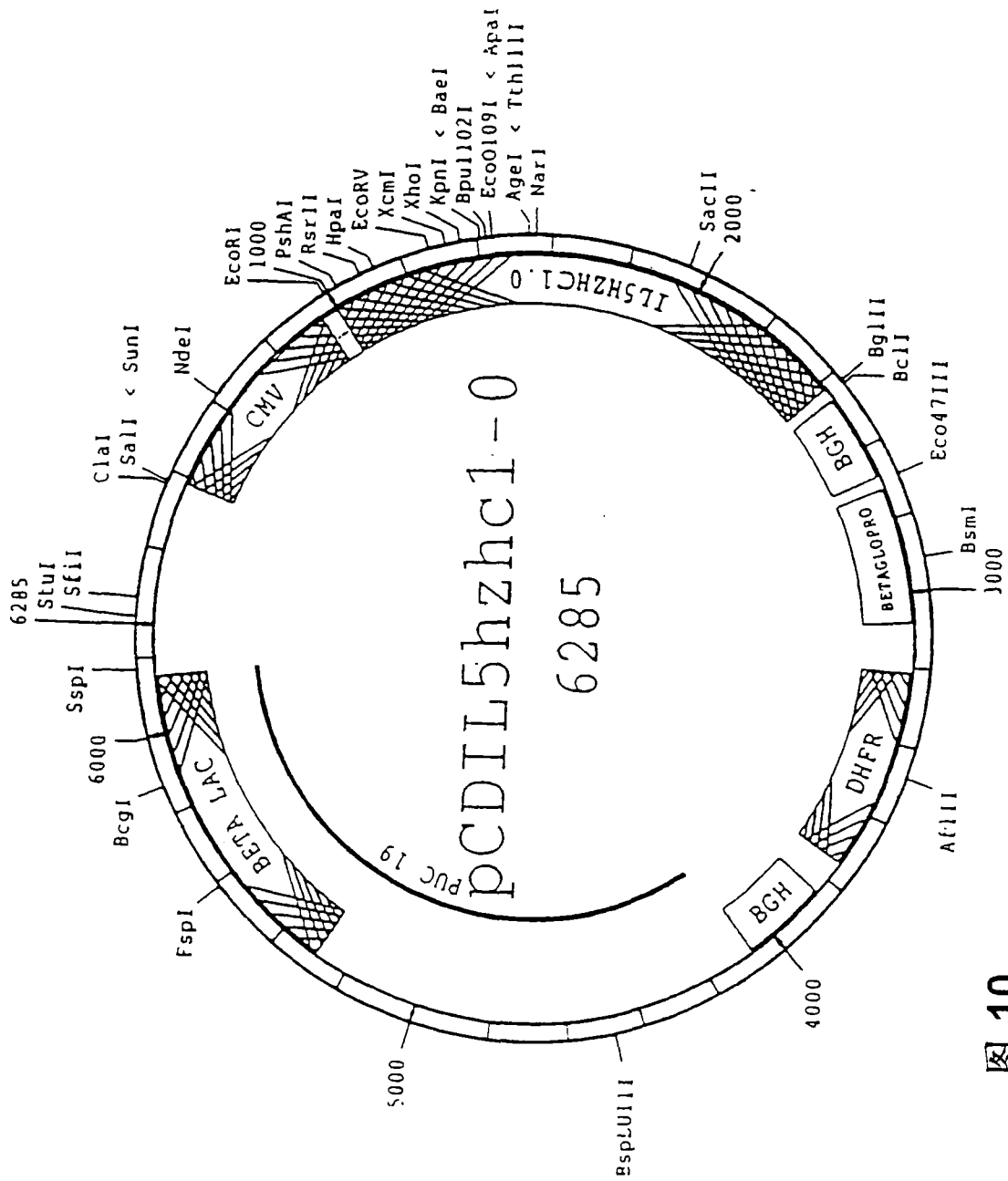


图 10

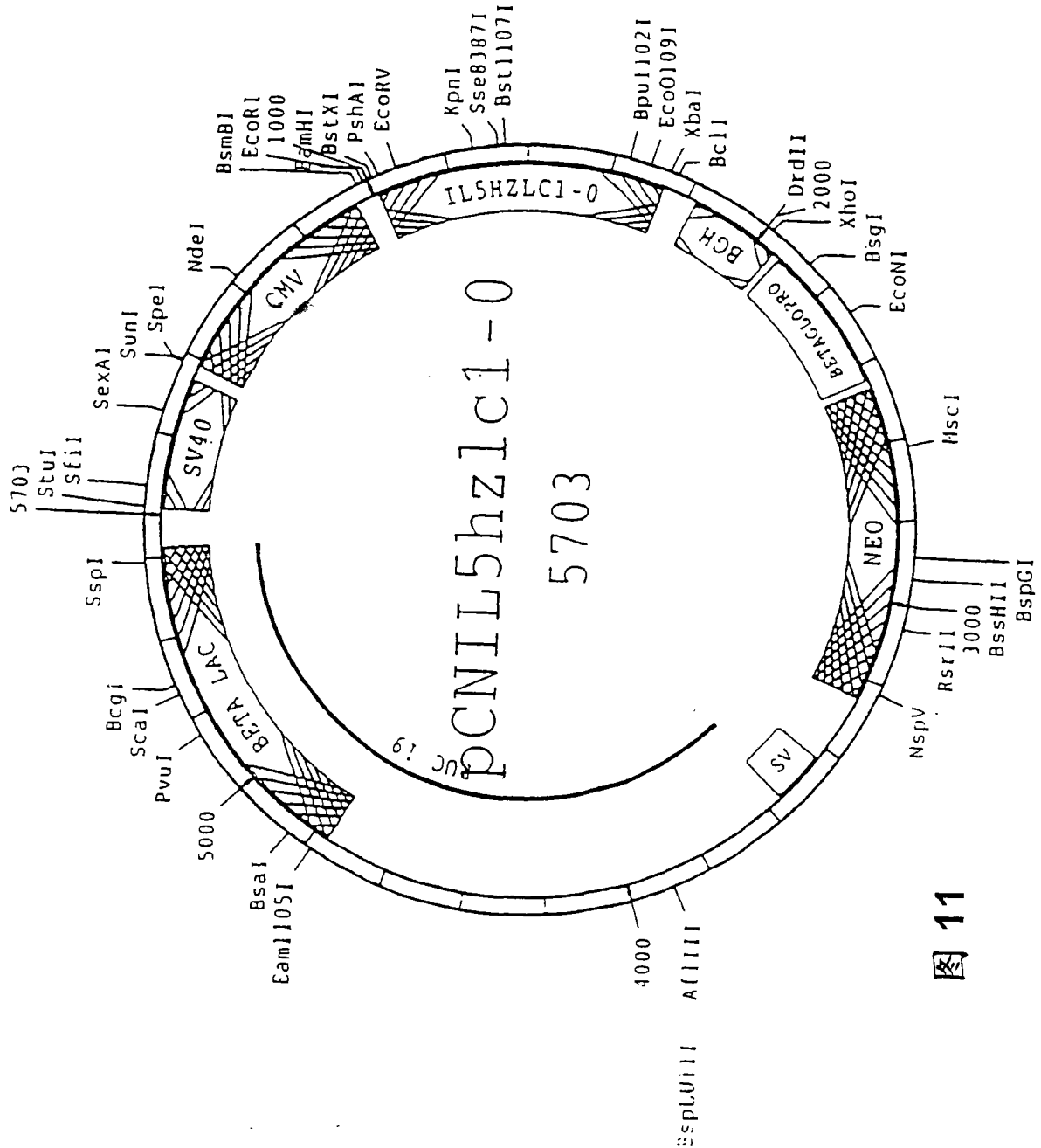


图 11

图 12

## IL5 NEWM 人源化重链可变区

|     |                                         |     |
|-----|-----------------------------------------|-----|
| 1   | CAG GTC CAA CTG CAG GAG AGC GGT CCA GGT | 30  |
|     | Gln Val Gln Leu Gln Glu Ser Gly Pro Gly |     |
| 31  | CTT GTG AGA CCT AGC CAG ACC CTG AGC CTG | 60  |
|     | Leu Val Arg Pro Ser Gln Thr Leu Ser Leu |     |
| 61  | ACC TGC ACC GTC TCG GGC TTC TCC CTC ACC | 90  |
|     | Thr Cys Thr Val Ser Gly Phe Ser Leu Thr |     |
| 91  | AGC TAT AGT GTA CAC TGG GTG AGA CAG CCA | 120 |
|     | Ser Tyr Ser Val His Trp Val Arg Gln Pro |     |
| 121 | CCT GGA CGA GGT CTA GAG TGG CTT GGA GTA | 150 |
|     | Pro Gly Arg Gly Leu Glu Trp Leu Gly Val |     |
| 151 | ATA TGG GCT AGT GGA GGC ACA GAT TAT AAT | 180 |
|     | Ile Trp Ala Ser Gly Gly Thr Asp Tyr Asn |     |
| 181 | TCG GCT CTC ATG TCC AGA CTG TCA ATA CTG | 210 |
|     | Ser Ala Leu Met Ser Arg Leu Ser Ile Leu |     |
| 211 | AAA GAC AAC AGC AAG AAC CAG GTC AGC CTG | 240 |
|     | Lys Asp Asn Ser Lys Asn Gln Val Ser Leu |     |
| 241 | AGA CTC AGC AGC GTG ACA GCC GCC GAC ACC | 270 |
|     | Arg Leu Ser Ser Val Thr Ala Ala Asp Thr |     |
| 271 | GCG GTC TAT TAC TGT GCT CGG GAT CCC CCT | 300 |
|     | Ala Val Tyr Tyr Cys Ala Arg Asp Pro Pro |     |
| 301 | TCT TCC TTA CTA CGG CTT GAC TAC TGG GGA | 330 |
|     | Ser Ser Leu Leu Arg Leu Asp Tyr Trp Gly |     |
| 331 | CAA GGT ACC ACG GTC ACC GTC TCG AGC     | 357 |
|     | Gln Gly Thr Thr Val Thr Val Ser Ser     |     |

## 图 13

## IL5 REI 人源化轻链可变区

|     |                                                                                       |     |
|-----|---------------------------------------------------------------------------------------|-----|
| 1   | GAT ATC GTG ATG ACC CAG AGC CCA AGC AGC                                               | 30  |
|     | Asp Ile Val Met Thr Gln Ser Pro Ser Ser                                               |     |
| 31  | CTG AGC GCT AGC GTG GGT GAC AGA GTG ACC                                               | 60  |
|     | Leu Ser Ala Ser Val Gly Asp Arg Val Thr                                               |     |
| 61  | ATC ACC TGT <span style="border: 1px solid black;">AAG AGC TCT CAG AGT CTG TTA</span> | 90  |
|     | Ile Thr Cys Lys Ser Ser Gln Ser Leu Leu                                               |     |
| 91  | <span style="border: 1px solid black;">AAC AGT GGA AAT CAA AAG AAC TAC TTG GCC</span> | 120 |
|     | Asn Ser Gly Asn Gln Lys Asn Tyr Leu Ala                                               |     |
| 121 | TGG TAT CAG CAG AAA CCC GGT AAG GCT CCA                                               | 150 |
|     | Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro                                               |     |
| 151 | AAG CTG CTG ATC TAC <span style="border: 1px solid black;">GGG GCA TCG ACT AGG</span> | 180 |
|     | Lys Leu Leu Ile Tyr Gly Ala Ser Thr Arg                                               |     |
| 181 | <span style="border: 1px solid black;">GAA TCT</span> GGG GTA CCA GAT AGA TTC AGC GGT | 210 |
|     | Glu Ser Gly Val Pro Asp Arg Phe Ser Gly                                               |     |
| 211 | AGC GGT AGC GGA ACC GAC TTC ACC TTC ACC                                               | 240 |
|     | Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr                                               |     |
| 241 | ATC AGC AGC CTG CAG CCA GAG GAC ATC GCC                                               | 270 |
|     | Ile Ser Ser Leu Gln Pro Glu Asp Ile Ala                                               |     |
| 271 | ACC TAC TAC TGC <span style="border: 1px solid black;">CAG AAT GTT CAT AGT TTT</span> | 300 |
|     | Thr Tyr Tyr Cys Gln Asn Val His Ser Phe                                               |     |
| 301 | <span style="border: 1px solid black;">CCA TTC ACG</span> TTC GGA CAA GGG ACC AAG GTG | 330 |
|     | Pro Phe Thr Phe Gly Gln Gly Thr Lys Val                                               |     |
| 331 | GAG ATC AAA                                                                           | 339 |
|     | Glu Ile Lys                                                                           |     |