

[21] 申请号 200910147020.0

C12N 15/12 (2006.01)
C07K 14/47 (2006.01)
C07K 16/18 (2006.01)
G01N 33/53 (2006.01)
G01N 33/68 (2006.01)
G01N 33/574 (2006.01)

[11] 公开号 CN 101643733A

权利要求书 2 页 说明书 127 页 附图 38 页

本发明涉及新的 STRA6 多肽。 本发明涉及具有类似于 Stra6(一种鼠视黄酸应答蛋白)的序列的新多肽, 涉及编码所述多肽的核酸分子。 本发明也提供包含这些核酸序列的载体和宿主细胞, 含有本发明多肽与异源多肽序列融合的嵌合多肽分子, 与本发明多肽结合的抗体, 以及涉及本发明多肽的生产方法。

[illegible]

蛋白的重要特性

4

增補的証

54-63

192-119
148-146

301-320

432-453

560-535

基序排列:

9-3

養濟名稱：

50-5

276-2
262-2

342-3

627-6

540-6

1168

354-3

陈德武 著

112-

1. 一种分离的核酸分子, 其包含与如下 DNA 分子有至少约 80% 序列同一性的 DNA, 所述如下 DNA 分子是指(a)编码 PRO10282 多肽的 DNA 分子, 该多肽包含图 2 (SEQ ID NO:2)中从约 1 到约 667 位氨基酸残基的序列, 或(b)编码 PRO19578 多肽的 DNA 分子, 该多肽包含图 7 (SEQ ID NO:5)中从约 1 到约 658 位氨基酸残基的序列, 或(c) (a)或(b)所述 DNA 分子的互补体。

2. 一种分离的核酸分子, 其包含与如下 DNA 分子有至少约 80% 序列同一性的 DNA: (a)编码同一种成熟多肽的 DNA 分子, 所述多肽由 2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的人蛋白 cDNA (DNA148380-2827)编码, (b)编码同一种成熟多肽的 DNA 分子, 所述多肽由 2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的人蛋白 cDNA (DNA148389-2827-1)编码, 或(c)(a)或(b)所述 DNA 分子的互补体。

3. 包含下述氨基酸序列的分离的 Stra6 多肽, 所述氨基酸序列与(a)图 2 (SEQ ID NO:2)中从约 1 到约 667 位的氨基酸残基的序列, 或(b)图 7 (SEQ ID NO:5)中从约 1 到约 658 位的氨基酸残基的序列有至少约 80% 的序列同一性。

4. 一种分离的 Stra6 多肽, 其与(a)由 2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的载体(DNA148380-2827)的 cDNA 插入子编码的多肽, 或(b)由 2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的载体 (DNA148389-2827-1)的 cDNA 插入子编码的多肽有至少大约 80% 的序列同一性。

5. 特异性结合 Stra6 多肽的抗体。

6. 一种确定怀疑有 Stra6 多肽的样品中所述多肽的存在的方法, 该方法包含使该样品暴露于抗 Stra6 抗体, 并检测所述抗体与所述样品中所述多肽的结合。

7. 一种诊断哺乳动物肿瘤的方法, 该方法包含检测如下样品中编码 Stra6 多肽的基因的表达水平: (a)来自哺乳动物组织细胞的待测样品, 和(b)来自相同细胞类型的已知正常组织细胞的对照样品, 其中, 与对照样品相比较, 待测样品中较高的表达水平提示, 该待测组织细胞的来源哺乳动物中肿瘤的存在。

8. 一种诊断哺乳动物肿瘤的方法, 该方法包含(a)使抗 Stra6 抗体与来

源于哺乳动物组织细胞的待测样品接触,和(b)检测所述抗体与所述待测样品中 Stra6 多肽之间复合物的形成,其中复合物的形成提示所述哺乳动物中肿瘤的存在。

9. 一种抑制肿瘤细胞生长的方法,该方法包含使表达 Stra6 多肽的肿瘤细胞暴露于有效量的一种制剂,该制剂抑制所述多肽的生物活性,从而抑制所述肿瘤细胞的生长。

10. 一种制品,其包含:

一种容器

容器上的标签; 和

一种组合物,所述组合物包含容纳在所述容器中的活性制剂,其中所述组合物可有效抑制肿瘤细胞的生长,并且所述容器上的标签提示该组合物可有效治疗以所述肿瘤细胞相对于相同组织类型的正常细胞而言过表达 Stra6 多肽为特征的情况。

新的 STRA6 多肽

本申请是申请日为 2001 年 1 月 11 日、中国申请号为 01806496.5、发明名称为“新的 STRA6 多肽”的发明申请的分案申请。

发明领域

本发明主要涉及鉴定和分离新的 DNA，涉及重组产生具有类似于鼠 Stra6(一种鼠视黄酸易感蛋白)序列的新多肽。早些时候曾经将这些分子中的一些称为"PRO10282"，但在下文中也将之称为"Stra6"多肽。

发明背景

膜结合蛋白

膜结合蛋白和受体在多细胞生物的构造、分化和维持方面发挥重要作用。许多单个细胞的命运，如增殖、迁移、分化或其它细胞相互作用，通常被接收自其它细胞和/或即时环境的信息所控制。这种信息常常由分泌多肽进行传输(例如，致有丝分裂因子，存活因子，细胞毒因子，分化因子，神经肽和激素)，而分泌多肽反过来被各种细胞受体或膜结合蛋白接收并翻译。此类膜结合蛋白和细胞受体包括但不限于：细胞因子受体、受体激酶、受体磷酸酶、细胞-细胞相互作用中所涉及的受体，以及细胞粘附分子如选择蛋白和整联蛋白。举例来说，调节细胞生长和分化的信号转导部分程度上受各种细胞蛋白质的磷酸化作用的调节。蛋白酪氨酸激酶(催化磷酸化过程的酶)，也起着生长因子受体的作用。实例包括成纤维细胞生长因子受体和神经生长因子受体。

膜结合蛋白和受体分子具有多种工业用途，包括作为药剂和诊断剂。比如，受体免疫粘附素可用作阻断受体-配体相互作用的治疗剂。膜结合蛋白也可用于筛选相关受体/配体相互作用的潜在的肽或小分子抑制剂。

工业和学术界都在进行着鉴定新的、天然受体或膜结合蛋白的不懈努力。许多努力集中在筛选哺乳动物重组 DNA 文库上以确定新受体或膜结合蛋白的编码序列。

本发明 Stra6 多肽与鼠蛋白 Stra6 具有序列同源性(73%相同和 81%相似),所述鼠蛋白 Stra6 的表达是由视黄酸诱导的。(Bouillet 等, Dev. Biol. 170: 420-433 [1995]; Bouillet 等, Mech. Dev. 63:173-186[1997]; Chazaud 等, Dev. Genet. 19:66-73[1996])。既然视黄酸是脊椎动物发育过程中一重要的信号分子,那么,认为应答视黄酸而诱导的基因在胚胎发育的生长和分化中发挥着重要作用。使用设计的为鉴定和分离视黄酸诱导基因的扣除杂交法,从 P19 鼠胚胎癌性细胞中分离出鼠 Stra6 cDNA,后者并未显示出与前面描述的蛋白质具有相似性。它含有与多个跨膜结构域相应的高度疏水伸展(stretches)的氨基酸残基,此是膜内在蛋白的一个特征。基于其表达模式, Stra6 被认为在胚胎发育过程的早期对于背腹侧肢(dorsoventral limb)的结构形式和后期对软骨内成骨的控制方面起着重要作用(Chazaud 等, Dev. Genet. 19:66-73[1996])。

本文公开的 Stra6 多肽含有多个可能是构成跨膜结构域的高度疏水区,其表明 Stra6 多肽是膜内在蛋白。它们可以作为未知配体的受体发挥作用,并可能是影响细胞生长、发育或分化的信号转导通路的一部分。

肿瘤细胞中的基因扩增

在美国,恶性肿瘤(癌)是导致死亡的居心脏疾病之后的第二位主要原因(Boring 等, CA Cancer J. Clin., 43:7[1993])。

恶性肿瘤的特征是,异常细胞数量增加或衍生自正常组织的肿瘤细胞增生形成肿瘤块,这些赘生的肿瘤细胞侵袭毗邻组织,并且增生的恶变细胞逐渐经血液或淋巴系统扩散到局部淋巴结和远端部位(转移)。在癌性状态下,细胞在正常细胞不生长的环境下增生。恶性肿瘤本身表现出多种形态,均以不同程度的侵袭性和侵犯性为特征。

基因表达的改变与不受抑制的细胞生长和去分化密切相关,而不经抑制的细胞生长和去分化是所有恶性肿瘤的共同特征。业已发现,一些进行过较好研究的肿瘤的基因组显示出隐性基因的表达减弱,隐性基因通常指肿瘤抑制基因,正常功能是抑制恶变细胞生长和/或抑制那些促进恶性肿瘤生长的某些优势基因如致癌基因的过度表达。这些中的每一种遗传学变化显得是肿瘤具有某些特性的原因,集合起来,则表现出全部的肿瘤表型(Hunter, Cell, 64:1129[1991]和 Bishop, Cell, 64:235-248[1991])。

众所周知,肿瘤细胞中基因(例如,致癌基因)过度表达机制是基因扩增。这是在祖细胞染色体中产生特定基因多拷贝的过程。该过程涉及含有特定基

因的染色体区域的异常复制,然后将复制的片段再重组回染色体中(Alitalo 等, *Adv. Cancer Res.*, 47:235-281[1986])。据信,基因过度表达与基因扩增相平行,即与生产的拷贝数成正比。

已经确定,编码生长因子和生长因子受体的原癌基因在各种人类恶性肿瘤(包括乳腺癌)的病因学中起着重要作用。例如,业已发现编码与表皮生长因子受体(EGFR)有关的 185 kd 跨膜糖蛋白受体(p185^{HER2}, HER2)的人 ErbB2 基因(*erbB2*, 也称为 *her2* 或 *c-erbB-2*),在约 25%~30%的人乳腺癌中过度表达(Slamon 等, *Science* 235:177-182 [1987]; Slamon 等, *Science* 244:707-712 [1989])。

据报道,原癌基因的基因扩增是更大恶性程度的恶性肿瘤中通常涉及的事件,并且可以起到临床发病预报器的作用(Schwab 等, *Genes Chromosomes Cancer*, 1:181-193[1990]; Alitalo 等, 同上)。因此, *erbB2* 过度表达一般被认为是预后差的预示,尤其对累及腋窝淋巴结的原发性疾病的患者而言(Slamon 等, [1987] 和 [1989], 同上; Ravdin 和 Chamness, *Gene*, 159:19-27[1995]; 以及 Hynes and Stern, *Biochim. Biophys. Acta*, 1198:165-184[1994]),而且与对激素治疗和化学疗法的敏感性和/或抗性相关,所述化学疗法包括 CMF(环磷酰胺,甲氨蝶呤,和氟尿嘧啶)和蒽环霉素(Baselga 等, *Oncology*, 11 (3 Suppl 1): 43-48 [1997])。然而,尽管 *erbB2* 过度表达与预后差有关,但是 HER2-阳性患者对用 taxanes 治疗的临床反应几率却比 HER2-阴性患者的反应几率大 3 倍以上(同上)。重组人源化抗 ErbB2(抗-HER2)单克隆抗体(鼠抗 ErbB2 抗体 4D5 的人源化版本,称作 rhuMAb HER2 或 HERCEPTINTM,对已经接受过广泛抗癌治疗的 ErbB2 过度表达型转移性乳腺癌患者进行的临床治疗有效 (Baselga 等, *J. Clin. Oncol.* 14: 737-744 [1996])。

据上所述,显然存在着对于鉴定在诊断和治疗与基因扩增有关的肿瘤中有用的新的分子、方法和组合物的极大兴趣。

发明概述

已经鉴定了与编码鼠视黄酸易感蛋白 Stra6 并编码 667 个氨基酸的新多肽(在本申请中称为全长天然序列人"PRO10282")的核酸具有同源性的 cDNA 克隆(本文中指定为 DNA148380-2827)。已经鉴定了编码 658 个氨基酸

的新多肽 (PRO 19578) 的另一人 cDNA 克隆 (本文指定为 DNA148389-2827-1), 显示出与鼠 Stra6 和天然序列人 PRO10282 均具有显著序列同源性, 它被认为是天然序列人 PRO10282 的剪切变体。正如后面将要讨论的那样, 根据与鼠 Stra6 的同源性, 也将本发明 PRO10282 多肽(包括天然序列和变体分子) 称为 "Stra6" 多肽。

在一实施方案中, 本发明提供分离的含有编码 PRO10282 多肽的核苷酸序列的核酸分子。

一方面, 分离的核酸分子含有与(a)编码具有图 2(SEQ ID NO: 2)所示从约 1 至约 667 位氨基酸的序列(包括两端残基在内)的 PRO10282 多肽的 DNA 分子, 或者(b)编码具有图 7(SEQ ID NO: 5)所示从约 1 至约 658 位氨基酸的序列(包括两端残基在内)的 PRO19578 多肽的 DNA 分子, 或者(c) (a)或(b)所述 DNA 分子的互补链, 具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98% 或约 99% 核酸序列同源的核苷酸序列。

另一方面, 该分离的核酸分子含有(a) 编码具有图 2(SEQ ID NO: 2)所示从约 1 至约 667 位氨基酸的序列(包括两端残基在内)的 PRO10282 多肽的核苷酸序列, 或者(b)编码具有图 7(SEQ ID NO: 5)所示从约 1 至约 658 位氨基酸的序列(包括两端残基在内)的 PRO19578 多肽的核苷酸序列, 或者(b) (a) 或(b)所述核苷酸序列的互补序列的核酸分子。

再一方面, 该分离的核酸分子含有与(a) 图 1 (SEQ ID NO: 1)所示由约 49 至约 2049 位核苷酸的序列(包括两端残基在内)的 DNA 分子, 或者(b) 图 6 (SEQ ID NO: 4)所示由约 186 至约 2159 位核苷酸的序列(包括两端残基在内)的 DNA 分子, 或者(a)或(b)所述 DNA 分子的互补序列, 具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98% 或约

99%核酸序列同源的核苷酸序列。

另一方面,该分离的核酸分子含有图1 (SEQ ID NO: 1)所示由约49至约2049位核苷酸的序列(包括两端残基在内),或者(b)图6 (SEQ ID NO: 4)所示由约186至约2159位核苷酸的序列(包括两端残基在内),或者(c) (a)或(b)核苷酸序列的互补序列。

此外,本发明涉及分离的核酸分子,其含有与(a)编码由2000年1月11日在ATCC保藏的ATCC保藏号为PTA-1181 (DNA 148380-2827)的人蛋白cDNA编码的相同成熟多肽的DNA分子,或者(b)编码由2000年2月23日在ATCC保藏的ATCC保藏号为PTA-1405 (DNA 148389-2827-1)的人蛋白cDNA编码的相同成熟多肽的DNA分子,或者(c) (a)或(b)所述DNA分子的互补序列,具有至少约80%、优选至少约81%、更优选至少约82%、更优选至少约83%、更优选至少约84%、更优选至少约85%、更优选至少约86%、更优选至少约87%、更优选至少约88%、更优选至少约89%、更优选至少约90%、更优选至少约91%、更优选至少约92%、更优选至少约93%、更优选至少约94%、更优选至少约95%、更优选至少约96%、更优选至少约97%、更优选至少约98%或约99%核酸序列同源的核苷酸序列。在一优选的实施方案中,分离的核酸分子含有(a)编码与2000年1月11日在ATCC保藏的ATCC保藏号为PTA-1181 (DNA 148380-2827)的人蛋白cDNA编码的相同成熟多肽的核苷酸序列,或者(b)编码由2000年2月23日在ATCC保藏的ATCC保藏号为PTA-1405 (DNA 148389-2827-1)的人蛋白cDNA编码的相同成熟多肽的核苷酸序列,或者(c) (a)或(b)所述核苷酸序列的互补序列。

另一方面,发明涉及分离的核酸分子,其含有与(a)2000年1月11日在ATCC保藏的ATCC保藏号为PT-1181 (DNA148380-2827)的人蛋白cDNA的全长多肽编码序列,或者(b)2000年2月23日在ATCC保藏的ATCC保藏号为PTA-1405 (DNA 148389-2827-1)的人蛋白cDNA的全长多肽编码序列,或者(c) (a)或(b)所述核苷酸序列的互补序列,具有至少约80%、优选至少约81%、更优选至少约82%、更优选至少约83%、更优选至少约84%、更优选至少约85%、更优选至少约86%、更优选至少约87%、更优选至少约88%、更优选至少约89%、更优选至少约90%、更优选至少约91%、更优选至少约92%、更优选至少约93%、更优选至少约94%、更优选至少约95%、更优选至少约96%、更优选至少约97%、更优选至少约98%或约99%核酸序

列同源的核苷酸序列。

在一优选的实施方案中,分离的核酸分子含有(a) 2000 年 1 月 11 日在 ATCC 保藏的 ATCC 保藏号为 PT-1181 (DNA148380-2827)的人蛋白 cDNA 的全长多肽编码序列,或者(b) 2000 年 2 月 23 日在 ATCC 保藏的 ATCC 保藏号为 PTA-1405 (DNA 148389-2827-1)的人蛋白 cDNA 的全长多肽编码序列,或者(c) (a)或(b)所述核苷酸序列的互补序列。

还一方面,发明涉及编码下面定义的活性 PRO10282 多肽的分离的核酸分子,其包括能与编码图 2(SEQ ID NO: 2)中 1~约 667 位氨基酸(包括两个端点氨基酸残基在内)的核酸序列的互补序列,或者与编码图 7(SEQ ID NO: 5)中 1~约 658 位氨基酸(包括两个端点氨基酸残基在内)的核酸序列的互补序列杂交的核苷酸序列,其中该分离的核酸分子不是编码鼠 Stra6 的 DNA。优选地,杂交在严格杂交条件和洗涤条件下发生。

另一方面,发明涉及编码下面定义的活性 PRO10282 多肽的分离的核酸分子,其包括与图 1(SEQ ID NO: 1)所示核苷酸约 49 和约 2049 之间(包括两个端点残基在内)核酸序列的互补序列,或者与图 6(SEQ ID NO: 4)所示核苷酸约 186 和约 2159 之间(包括两个端点残基在内)核酸序列的互补序列杂交的核苷酸序列,其中该分离的核酸分子不是编码鼠 Stra6 的 DNA。优选地,杂交在严格杂交条件和洗涤条件下发生。

再一方面,发明涉及具有至少约 765 个核苷酸的分离的核酸分子,其是在严格条件下使测试 DNA 分子与(a) 编码含图 2(SEQ ID NO: 2)所示从约 1 至约 667 的氨基酸序列(包括两端残基在内)的 PRO10282 多肽的 DNA 分子,或者(b) 编码含图 7(SEQ ID NO: 5)所示从约 1 至约 658 的氨基酸序列(包括两端残基在内)的 PRO19578 多肽的 DNA 分子,或者(c) (a)或(b)所述 DNA 分子的互补序列进行杂交而产生的,而且,如果该测试 DNA 分子与(a)或(b)具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%、更优选至少约 99%的核酸序列同一性,则分离该测试 DNA 分子。

另一方面,发明涉及分离的核酸分子,其含有与(a) 编码含图 2(SEQ ID NO: 2)所示从约 1 至约 667 的氨基酸序列(包括两端残基在内),或者(b) 编码含图 7(SEQ ID NO: 5)所示从约 1 至约 658 的氨基酸序列(包括两端残基在内)相比,具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%、更优选至少约 99% 评分阳性的多肽的核苷酸序列,或者含有(c) (a)或(b)所述核苷酸序列的互补序列。

本发明的另一特定方面,本发明提供包含编码 PRO10282 的核苷酸序列的分离的核酸分子,该核苷酸序列要么编码跨膜结构域缺失或者跨膜结构域失活,要么与此编码核苷酸序列互补,其中跨膜结构域已被试验性地确定为:在图 2 (SEQ ID NO: 2)所示序列中,从约氨基酸 54 位向约氨基酸 69 位延伸、从约氨基酸 102 位向约氨基酸 119 位延伸、从约氨基酸 148 位向约氨基酸 166 位延伸、从约氨基酸 207 位向约氨基酸 222 位延伸、从约氨基酸 301 位向约氨基酸 320 位延伸、从约氨基酸 364 位向约氨基酸 380 位延伸、从约氨基酸 431 位向约氨基酸 451 位延伸、从约氨基酸 474 位向约氨基酸 489 位延伸、和从约氨基酸 512 位向约氨基酸 531 位延伸;而在图 7 (SEQ ID NO: 5)所示序列中,从约氨基酸 54 位向约氨基酸 71 位延伸、从约氨基酸 93 位向约氨基酸 111 位延伸、从约氨基酸 140 位向约氨基酸 157 位延伸、从约氨基酸 197 位向约氨基酸 214 位延伸、从约氨基酸 291 位向约氨基酸 312 位延伸、从约氨基酸 356 位向约氨基酸 371 位延伸、从约氨基酸 425 位向约氨基酸 481 位延伸、从约氨基酸 505 位向约氨基酸 522 位延伸。因此,本发明所述 PRO10282 多肽的可溶性细胞外结构域也在考虑之内。

在这点上,本发明的另一方面涉及分离的核酸分子,其含有与(a) 编码图 2 (SEQ ID NO: 2) 1 至 X 氨基酸的 DNA 分子,其中 X 是图 2 (SEQ ID NO: 2)中自 49 到 59 的任何氨基酸,或(b) (a)中 DNA 分子的互补序列具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少

约 91 %、更优选至少约 92 %、更优选至少约 93 %、更优选至少约 94 %、更优选至少约 95 %、更优选至少约 96 %、更优选至少约 97 %、更优选至少约 98 %、更优选至少约 99 % 同一性的核苷酸序列。在一特定方面, 该分离的核酸分子含有(a) 编码图 2 (SEQ ID NO : 2)或图 7 (SEQ ID NO : 5)中 1 至 X 氨基酸的核苷酸序列, 其中 X 是图 2 (SEQ ID NO : 2) 中自 49 至 59 的任何氨基酸, 或者(b) (a)中 DNA 分子的互补序列。一特定方面, 该分离的核酸分子含有 (a)编码图 2 ISEQ ID NO : 2), 或图 7 (SEQ ID NO : 5)中 1 至 X 的氨基酸, 其中 X 是图 2 (SE ID NO : 2), 或图 7 (SEQ ID NO : 5)自 49 至 59 的任何氨基酸(包括两端氨基酸残基在内), 或者(b) (a)中 DNA 分子的互补序列的核苷酸序列。

本发明还有一方面, 分离的核酸分子 (a) 编码与图 2 (SEQ ID NO : 2), 或图 7 (SEQ ID NO : 5) 中约 1 至 X 残基氨基酸序列, 其中 X 是图 2 (SEQ ID NO : 2)中自 49 至 59 的任何氨基酸, 或者(b) (a)中 DNA 分子的互补序列相比, 具有至少约 80 %、优选至少约 81 %、更优选至少约 82 %、更优选至少约 83 %、更优选至少约 84 %、更优选至少约 85 %、更优选至少约 86 %、更优选至少约 87 %、更优选至少约 88 %、更优选至少约 89 %、更优选至少约 90 %、更优选至少约 91 %、更优选至少约 92 %、更优选至少约 93 %、更优选至少约 94 %、更优选至少约 95 %、更优选至少约 96 %、更优选至少约 97 %、更优选至少约 98 %、更优选至少约 99 % 阳性评分的多肽。

另一实施方案, 涉及 PRO10282 多肽编码序列片段, 这些片段例如可用作杂交探针, 或者可任选地编码含抗-PRO10282 抗体结合位点的 PRO10282 多肽的编码片段。此类核酸片段长度通常至少约 20 个核苷酸, 优选至少约 30 个核苷酸, 更优选至少约 40 个核苷酸, 更优选至少约 50 个核苷酸, 更优选至少约 60 个核苷酸, 更优选至少约 70 个核苷酸, 更优选至少约 80 个核苷酸, 更优选至少约 90 个核苷酸, 更优选至少约 100 个核苷酸, 或者至少约 110 个核苷酸, 更优选至少约 120 个核苷酸, 更优选至少约 130 个核苷酸, 更优选至少约 140 个核苷酸, 或者至少约 150 个核苷酸, 更优选至少约 160 个核苷酸, 更优选至少约 170 个核苷酸, 更优选至少约 180 个核苷酸, 或者至少约 190 个核苷酸, 更优选至少约 200 个核苷酸, 更优选至少约 250 个核苷酸, 更优选至少约 300 个核苷酸, 更优选至少约 350 个核苷酸, 更优选至少约 400 个核苷酸, 更优选至少约 450 个核苷酸, 更优选至少约 500 个核苷

酸,更优选至少约 600 个核苷酸,更优选至少约 700 个核苷酸,更优选至少约 800 个核苷酸,更优选至少约 900 个核苷酸,更优选至少约 1000 个核苷酸,其中本文术语“约”是指核苷酸序列有关长度加或减 10%。本发明核酸片段不同于鼠 Stra6 天然编码序列的片段。在一优选的实施方案中,该核苷酸序列片段衍生自图 1 (SEQ ID NO: 1) 或图 6 (SEQ ID NO: 4) 所示核苷酸序列的任何编码区。应当指出,可用常规方法测定编码 PRO10282 多肽的核苷酸序列的新片段,所述方法是:通过使用众多公知的序列对比程序中任何一个,将编码 PRO10282 多肽的核苷酸序列与其它已知核苷酸序列对比,测定出其中哪个或哪几个编码 PRO10282 多肽的核苷酸序列片段是新的。本文考虑到了所有这类编码 PRO10282 多肽的核苷酸序列,并且可以用不繁琐的实验予以测定。也考虑到由这些核苷酸分子片段编码的 PRO10282 多肽片段,优选是含有抗-PRO10282 抗体结合位点的那些 PRO10282 多肽片段。

在另一实施方案中,本发明提供含有编码 PRO10282 或其变体的核苷酸序列的载体。该载体可含有上述鉴定的任何分离的核酸分子。

也提供含有此类载体的宿主细胞。例如,宿主细胞可以是 CHO 细胞、大肠杆菌或酵母。进一步提供生产 PRO10282 多肽的方法,包括在适宜表达 PRO10282 的条件下培养宿主细胞并从细胞培养物中回收 PRO10282。

在另一实施方案中,发明提供由上面鉴定的任一分离的核酸序列所编码的分离的 PRO10282 多肽。

一特定方面中,本发明提供分离的天然序列的 PRO10282 多肽,它在某些实施方案中包括含有图 2 (SEQ ID NO: 2) 中自约 1 至约 667 残基,或者图 7 (SEQ ID NO: 5) 自约 1 至约 658 残基的氨基酸序列。

另一方面,本发明涉及分离的 PRO10282 多肽,其包括与图 2 (SEQ ID NO: 2) 中约 1 到约 667 残基所示氨基酸序列(包括两端残基在内),或与图 7 (SEQ ID NO: 5) 中从约 1 至约 658 的残基氨基酸序列(包括两端氨基酸残基在内)具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%、或更优选至少约 99% 同一性的氨基酸序列。

再一方面, 本发明涉及分离的 PRO10282 多肽, 其含有与 2000 年 1 月 11 日在 ATCC 保藏的 ATCC 保藏号为 PTA-1191 (DNA148380-2827) 或 2000 年 2 月 23 日在 ATCC 保藏的 ATCC 保藏号为 PTA-1405 (DNA148389-2827-1) 的人蛋白 cDNA 编码的氨基酸序列具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98% 或至少约 99% 同一性的氨基酸序列。在一优选的实施方案中, 分离的 PRO10282 多肽含有 2000 年 1 月 11 日在 ATCC 保藏的 ATCC 保藏号为 PTA-1191 (DNA148380-2827) 或 2000 年 2 月 23 日在 ATCC 保藏的 ATCC 保藏号为 PTA-1405 (DNA148389-2827-1) 的人蛋白 cDNA 编码的氨基酸序列。

另一方面, 发明涉及分离的 PRO10282 多肽, 其含有与图 2 (SEQ ID NO: 2) 中约 1 到约 667 位残基的氨基酸序列(包括两端残基在内)或图 7 (SEQ ID NO: 5) 中约 1 到约 659 位残基的氨基酸序列(包括两端残基在内)相比, 具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%、或更优选至少约 99% 评分阳性的氨基酸序列。

本发明另一方面提供分离的 PRO10282 多肽, 其要么跨膜结构域缺失要么跨膜结构域失活。本文也描述产生该多肽的方法, 该方法包括在适宜 PRO10282 多肽表达的条件下培养含有适合编码核酸分子的载体的宿主细胞, 并从细胞培养物中回收 PRO10282 多肽。

照这样, 本发明一方面涉及分离的可溶性 PRO10282 多肽, 其含有与图 2 (SEQ ID NO: 2), 或图 7 (SEQ ID NO: 5) 中 1 至 X 氨基酸(所述 X 是图 2 (SEQ ID NO: 2), 或图 7 (SEQ ID NO: 5) 中自 49 至 59 的任何氨基酸) 具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少

约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%、或更优选至少约 99% 序列同一性的氨基酸序列。

在本发明的又一方面,该分离的可溶性 PRO10282 多肽含有与图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)中约 1 至 X 的残基氨基酸序列相比时, (其中 X 是图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5) 中自 49 至 59 的任何氨基酸)具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%、或更优选至少约 99% 评分阳性的氨基酸序列。

还有一方面,本发明涉及分离的 PRO10282 多肽,其含有图 2 (SEQ ID NO: 2)中从 1 约至约 667 残基的氨基酸序列(包括两端氨基酸残基在内), 或图 7 (SEQ ID NO: 5)中从 1 约至约 658 残基的氨基酸序列(包括两端氨基酸残基在内), 或其生物活性片段,或足以提供与抗-PRO1282 抗体结合位点的片段,其中具有生物活性或提供抗-PRO10282 抗体结合位点的 PRO10282 多肽片段的鉴定,可以用本领域公知的常规技术来完成。这里的片段不是鼠 Stra6 多肽的片段。优选地, PRO10282 片段在性质上保留天然 PRO10282 多肽的生物活性。

再一方面,本发明提供如下制备的多肽: (i)在严格条件下,使测试 DNA 分子与(a)编码具有图 2 (SEQ ID NO: 2)中约 1 至约 667 残基的氨基酸序列(包括两端残基在内)的 PRO10282 多肽的 DNA 分子,或者(b) 编码具有图 7 (SEQ ID NO: 5)中约 1 至约 658 残基氨基酸序列(包括两端残基在内)的 PRO19578 多肽的 DNA 分子, (c) (a)或(b)所述 DNA 分子的互补序列,进行杂交,且如果该测试 DNA 分子与(a)或(b)或(c)具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少

约 96%、更优选至少约 97%、更优选至少约 98% 或至少约 99% 的核酸序列同一性, 则(ii)在适宜表达该多肽的条件下培养含有该测试 DNA 分子的宿主细胞, 和(iii)从细胞培养物中回收该多肽。

在另一实施方案中, 本发明提供含有与异源多肽或氨基酸序列融合的 PRO10282 多肽的嵌合分子, 其中 PRO10282 多肽可含有前述任何 PRO10282 多肽或其变体或片段。此嵌合分子的一个实例是含有与免疫球蛋白表位标签序列或 Fc 区融合的 PRO10282 多肽的嵌合分子。

在另一实施方案中, 发明提供下述定义的与前述 PRO10282 多肽特异结合的抗体。任选地, 抗体是单克隆抗体、抗体片段或单链抗体。

在另一实施方案中, 本发明涉及下述定义天然 PRO10282 多肽的激动剂和拮抗剂。在一具体实施方案中, 激动剂或拮抗剂是抗-PRO10282 抗体或小分子。

在另一实施方案中, 发明涉及鉴定 PRO10282 多肽的激动剂或拮抗剂的方法, 方法包括使 PRO10282 多肽与候选分子接触, 并检测由所述 PRO10282 多肽介导的生物活性。优选地, PRO10282 多肽是天然 PRO10282 多肽。

在另外一个实施方案中, 本发明涉及含有本文所述 PRO10282 多肽、PRO10282 多肽的激动剂或拮抗剂、或抗-PRO10282 抗体、以及载体的组合物。任选地, 载体是可药用载体。

本发明另一实施方案涉及本文所述 PRO10282 多肽或其激动剂或拮抗剂或抗-PRO10282 抗体在制备用于治疗对 PRO10282 多肽、其激动剂或拮抗剂、或抗-PRO10282 抗体敏感的病症的药物中的应用。

再一方面, 本发明涉及与 PRO10282 多肽特异结合的抗体。优选地, 该抗体诱导表达 PRO10282 多肽的细胞死亡。在一具体优选的实施方案中, 细胞是与相同组织类型的正常细胞相比过度表达 PRO10282 多肽的癌细胞。抗体优选是人源化的或人抗体, 并且包括抗体片段单链抗体。

另一方面, 本发明涉及含有与 PRO10282 多肽特异结合的抗体的组合物。

还有一方面, 本发明涉及编码与 PRO10282 多肽特异结合的抗体的分离的核酸分子、含有该核酸的载体以及含有该载体的宿主细胞。本发明也涉及生产抗-PRO10282 抗体的方法。

在另外一个实施方案中, 本发明涉及检测在怀疑含有 PRO10282 多肽的

样品中是否存在该多肽的方法，方法如此进行：将样品暴露于抗-PRO10282 抗体(或 PRO10282 的另一拮抗剂)，并测定样品中该抗体(或拮抗剂)与 PRO10282 多肽的结合。在一具体重要的实施方案中，样品取自癌细胞。

在一不同的实施方案中，本发明涉及诊断哺乳动物肿瘤的方法，包括检测在(a) 取自该哺乳动物的测试组织样品中，和 (b) 取自相同细胞类型的已知正常组织细胞的对照样品中，编码 PRO10282 多肽的基因的表达水平。与对照样品相比，测试样品中表达水平较高，表明测试组织所取自的哺乳动物存在肿瘤。

根据本发明的另一方法，如下步骤诊断哺乳动物中的肿瘤：(a) 将抗 PRO10282 抗体与取自该哺乳动物组织细胞的测试样品进行接触，和(b) 检测测试样品中抗体与 PRO10282 多肽之间复合物的形成。复合物形成，表明哺乳动物中存在肿瘤。

另一方面，本发明涉及癌诊断试剂盒，其含有抗-PRO10282 抗体(或 PRO10282 的不同拮抗剂)和在适宜包装中的载体。任选地，试剂盒也含有指导使用抗体或其它拮抗剂来检测样品中存在 PRO10282 多肽的说明书。

还有一方面，本发明涉及抑制肿瘤细胞生长的方法，通过使表达 PRO10282 多肽的肿瘤细胞暴露于抑制 PRO10282 多肽生物活性的有效量的制剂中，从而抑制肿瘤细胞的生长。制剂可以例如是抗-PRO10282 抗体或 PRO10282 的另一拮抗剂。肿瘤细胞可进一步暴露于常规的肿瘤治疗中，如放射治疗、用细胞毒性剂和/或化疗剂治疗等。

另一不同方面，本发明涉及抑制肿瘤细胞生长的方法，通过使表达 PRO10282 多肽的肿瘤细胞暴露于抑制该多肽表达的有效量的制剂中，从而抑制肿瘤生长。

本发明进一步涉及阻止或减缓(减弱)以 PRO10282 多肽过度表达为特征的肿瘤发展、生长或扩散的类似方法。该治疗可直接减少肿瘤细胞的病变，或者使得肿瘤细胞对于其它治疗剂的治疗(例如放疗和/或化疗)更为敏感。本发明特别包括控制异常或转移，控制对毗邻组织正常功能的干扰、控制细胞因子或其它分泌产物非正常水平释放、控制炎症或免疫应答的抑制或加重等的方法。

本发明进一步涉及一种制品，包括：一容器，贴在容器上的标签，和装在容器中的含有活性剂的组合物；其中组合物对于抑制肿瘤生长是有效的，

而贴在容器上的标签则注明该组合物用于治疗以肿瘤细胞中过度表达 PRO10282 多肽为特征的病症是有效的。

再有一方面，本发明涉及鉴定抑制 PRO10282 多肽生物学或免疫学活性的化合物的方法：在控制条件下，使候选化合物与 PRO10282 多肽接触足够长时间，以使两种组分相互作用，并检测 PRO10282 多肽的生物学或免疫学活性是否被抑制。

另一方面，本发明涉及鉴定抑制 PRO10282 多肽活性的化合物的方法，包括下列步骤：(a) 在存在 PRO10282 多肽、适宜由 PRO10282 多肽正常地诱导细胞应答的条件下，使细胞与被筛选的候选化合物接触，和 (b) 检测所述细胞应答的诱导情况，以决定测试化合物是否是有效的拮抗剂。细胞应答诱导缺乏，表明测试化合物是有效的拮抗剂。

本发明还涉及鉴定抑制在表达 PRO10282 多肽的细胞中表达该肽的化合物的方法，通过使细胞与候选化合物接触，并测定表达是否被抑制。

本发明特别涉及通过本文筛选方法鉴定的 PRO10282 激动剂和拮抗剂。此类激动剂和拮抗剂包括但不限于：抗-PRO10282 抗体、多肽、肽、和具有激动剂或拮抗剂特性的小的有机分子。

本发明涉及以下各项：

1. 一种分离的核酸分子，其包含与如下 DNA 分子有至少约 80% 序列同一性的 DNA，所述如下 DNA 分子是指(a)编码 PRO10282 多肽的 DNA 分子，该多肽包含图 2 (SEQ ID NO:2)中从约 1 到约 667 位氨基酸残基的序列，或(b)编码 PRO19578 多肽的 DNA 分子，该多肽包含图 7 (SEQ ID NO:5)中从约 1 到约 658 位氨基酸残基的序列，或(c) (a)或(b)所述 DNA 分子的互补体。
2. 项 1 的分离的核酸分子，其包含(a)图 1(SEQ ID NO:1)中从约 49 到约 2049 位的核苷酸序列或(b)图 6(SEQ ID NO:4)中从约 186 到约 2159 位的核苷酸序列，或(c)(a)或(b)所述核苷酸序列的互补体。
3. 项 1 的分离的核酸分子，其包含图 1(SEQ ID NO:1)或图 6(SEQ ID NO:4)的核苷酸序列。
4. 项 1 的分离的核酸分子，其包含编码以下氨基酸序列的核苷酸序列：(a)图 2(SEQ ID NO:2)中从约 1 到约 667 位氨基酸残基的序列，或(b)图 7(SEQ ID NO:5)中从约 1 到约 658 位氨基酸残基的序列。
5. 一种分离的核酸分子，其包含与如下 DNA 分子有至少约 80% 序列

同一性的 DNA: (a)编码同一种成熟多肽的 DNA 分子, 所述多肽由 2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的人蛋白 cDNA (DNA148380-2827)编码, (b)编码同一种成熟多肽的 DNA 分子, 所述多肽由 2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的人蛋白 cDNA (DNA148389-2827-1)编码, 或(c)(a)或(b)所述 DNA 分子的互补体。

6. 项 5 的分离的核酸分子, 其包含编码同一种成熟多肽的 DNA, 所述多肽由(a) 2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的人蛋白 cDNA (DNA148380-2827)编码, 或由(b) 2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的人蛋白 cDNA (DNA148389-2827-1)编码。

7. 一种分离的核酸分子, 其包含与如下序列有至少约 80 % 序列同一性的 DNA: (a)2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的人蛋白 cDNA (DNA148380-2827)的全长多肽编码序列, 或(b)2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的人蛋白 cDNA (DNA148389-2827-1)的全长多肽编码序列, 或(c)(a)或(b)的编码序列的互补体。

8. 项 7 的分离的核酸分子, 其包含(a) 2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的人蛋白 cDNA(DNA148380-2827)的全长多肽编码序列, 或(b) 2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的人蛋白 cDNA(DNA148389-2827-1)的全长多肽编码序列。

9. 一种分离的编码 PRO10282 多肽的核酸分子, 其包含与以下核酸序列的互补体杂交的 DNA, 所述以下核酸序列编码(a)图 2(SEQ ID NO:2)中 1 到约 667 位氨基酸或(b)图 7(SEQ ID NO:5)中 1 到约 658 位氨基酸。

10. 项 9 的分离的核酸分子, 其中(a)编码图 2(SEQ ID NO:2)中 1 到约 667 位氨基酸的核酸包含图 1(SEQ ID NO:1)中 49 到约 2049 位的核苷酸, 和 (b)编码图 7(SEQ ID NO:5)中 1 到约 658 位氨基酸的核酸包含图 6(SEQ ID NO:4)中 186 到约 2159 位的核苷酸。

11. 项 9 的分离的核酸分子, 其中所述杂交在严谨的杂交和洗涤条件下进行。

12. 一种分离的核酸分子, 其包含编码一种多肽的 DNA, 所述多肽与如下氨基酸序列相比时阳性分值至少是 80 %: (a)图 2(SEQ ID NO:2)中从 1 到约 667 位的氨基酸残基的序列, 或(b)图 7(SEQ ID NO:5)中从 1 到约 658 位的氨基酸残基的序列, 或(c)(a)或(b)的 DNA 的互补体。

13. 一种分离的核酸分子,其包含至少约 765 个核苷酸,并且通过使待测 DNA 分子在严谨杂交条件下与如下 DNA 分子杂交再分离该待测 DNA 分子来制备:(a)编码包含图 2(SEQ ID NO:2)中从 1 到约 667 位氨基酸残基的序列的 PRO10282 多肽的 DNA 分子,或(b)编码包含图 7(SEQ ID NO:5)中从 1 到约 658 位氨基酸残基的序列的 PRO19578 多肽的 DNA 分子,或(c)(a)或(b)的 DNA 分子的互补体。

14. 项 13 的分离的核酸分子,其具有与(a), (b)或(c)至少约 80% 的序列同一性。

15. 包含项 1 至 14 任一项的核酸分子的载体。

16. 项 15 的载体,其中所述核酸分子可操作地连接到被该载体转化的宿主细胞能识别的调控序列上。

17. 保藏在 ATCC、保藏号为 PTA-1181 的核酸分子(DNA148380-2827)或 PTA-1405 的核酸分子(DNA148389-2827-1)。

18. 包含项 15 的载体的宿主细胞。

19. 项 18 的宿主细胞,其中所述细胞是 CHO 细胞。

20. 项 18 的宿主细胞,其中所述细胞是大肠杆菌。

21. 项 18 的宿主细胞,其中所述细胞是酵母细胞。

22. 制备 Stra6 多肽的方法,其包含在适合所述 Stra6 多肽表达的条件下培养项 18 的宿主细胞并从细胞培养物中回收所述 Stra6 多肽。

23. 包含下述氨基酸序列的分离的 Stra6 多肽,所述氨基酸序列与(a)图 2(SEQ ID NO:2)中从约 1 到约 667 位的氨基酸残基的序列,或(b)图 7(SEQ ID NO:5)中从约 1 到约 658 位的氨基酸残基的序列有至少约 80% 的序列同一性。

24. 项 23 的分离的 Stra6 多肽,其包含图 2(SEQ ID NO:2)中从 1 到约 667 位的氨基酸残基,或图 7(SEQ ID NO:5)中从 1 到约 658 位的氨基酸残基。

25. 一种分离的 Stra6 多肽,其与(a)由 2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的载体(DNA148380-2827)的 cDNA 插入子编码的多肽,或(b)由 2000 年 2 月 23 日保藏在 ATCC、保藏号为 PTA-1405 的载体(DNA148389-2827-1)的 cDNA 插入子编码的多肽有至少大约 80% 的序列同一性。

26. 项 25 的分离的 Stra6 多肽,其由(a)2000 年 1 月 11 日保藏在 ATCC、保藏号为 PTA-1181 的载体(DNA148380-2827)的 cDNA 插入子,或(b)2000

年2月23日保藏在ATCC、保藏号为PTA-1405的载体(DNA148389-2827-1)的cDNA插入子编码。

27. 一种分离的Stra6多肽, 其与(a)图2(SEQ ID NO:2)中从1到约667位的氨基酸残基的序列, 或(b)图7(SEQ ID NO:5)中从1到约658位氨基酸残基的序列比较时, 阳性分值至少是80%。

28. 一种分离的Stra6多肽, 其包含(a)图2(SEQ ID NO:2)中从1到约667位氨基酸残基的序列, 或(b)图7(SEQ ID NO:5)中从1到约658位氨基酸残基的序列, 或(c)足以提供抗Stra6抗体结合位点的(a)或(b)的片段。

29. 一种分离的多肽, 其通过如下步骤制备: (i)在严谨条件下使待测DNA分子与编码Stra6多肽的DNA分子杂交, 所述多肽包含(a)图2(SEQ ID NO:2)中1到约667位氨基酸残基的序列, 或(b)图7(SEQ ID NO:5)中1到约658位氨基酸残基的序列, 或(c)(a)的DNA分子的互补体, (ii)在适合所述多肽表达的条件下培养包含所述待测DNA分子的宿主细胞, 和(iii)从所述细胞培养物中回收所述多肽。

30. 项29的分离的多肽, 其中所述待测DNA与(a)或(b)有至少约80%的序列同一性。

31. 一种嵌合分子, 其包含与异源氨基酸序列融合的Stra6多肽。

32. 项31的嵌合分子, 其中所述异源氨基酸序列是表位标签序列。

33. 项31的嵌合分子, 其中所述异源氨基酸序列是免疫球蛋白的Fc区。

34. 特异性结合Stra6多肽的抗体。

35. 项34的抗体, 其诱导表达所述多肽的细胞死亡。

36. 项35的抗体, 其中所述细胞是与同一组织类型的正常细胞相比所述多肽过表达的癌症细胞。

37. 项34的抗体, 其是单克隆抗体。

38. 项37的抗体, 其包含非人的互补决定区(CDR)或人的框架区(FR)。

39. 项34的抗体, 其是标记的抗体。

40. 项34的抗体, 其是抗体片段或单链抗体。

41. 项34的抗体, 其是人源化的抗体。

42. 一种物质组合物, 其包含项34的抗体, 所述抗体与可药用载体混合。

43. 项42的物质组合物, 其包含治疗有效量的所述抗体。

44. 项 42 的物质组合物, 其进一步包含细胞毒制剂或化疗制剂。
45. 一种分离的核酸分子, 其编码项 34 的抗体。
46. 包含项 45 的核酸分子的载体。
47. 包含项 46 的载体的宿主细胞。
48. 一种制备与 Stra6 多肽结合的抗体的方法, 该方法包含在足以允许所述抗体表达的条件下培养项 47 的宿主细胞, 并从所述细胞培养物中回收所述抗体。
49. 一种 Stra6 多肽的激动剂。
50. 一种 Stra6 多肽的拮抗剂。
51. 项 50 的拮抗剂, 其中所述拮抗剂抑制肿瘤细胞的生长。
52. 一种物质组合物, 其包含与可药用载体混合的(a) Stra6 多肽, (b) Stra6 多肽的激动剂, (c) Stra6 多肽的拮抗剂, 或(d)抗 Stra6 抗体。
53. 一种分离的核酸分子, 其包含一种核苷酸序列, 该核苷酸序列与如下 DNA 分子有至少大约 80% 的序列同一性: (a)编码图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中 1 到 X 位氨基酸的 DNA 分子, 其中 X 是图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中从 49 到 59 位的任何氨基酸。或(b)(a)的 DNA 分子的互补体。
54. 项 53 的分离的核酸, 其包含(a)编码图 2(SEQ ID NO:2)中 1 到 X 位氨基酸的核苷酸序列, 其中 X 是图 2(SEQ ID NO:2)中 49 到 59 位的任何氨基酸, 或(b)(a)的核苷酸序列的互补体。
55. 一种分离的核酸分子, 其包含(a)编码一种多肽的核苷酸序列, 该多肽与图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中从大约 1 到 X 位残基的氨基酸序列相比时, 阳性分值至少是大约 80%, 其中 X 是图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中从 49 到 59 位的任何氨基酸, 或(b)(a)的核苷酸序列的互补体。
56. 一种分离的可溶性 Stra6 多肽, 其包含一种氨基酸序列, 该氨基酸序列与图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中 1 到 X 位的氨基酸有至少大约 80% 的序列同一性, 其中 X 是图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中从 49 到 59 位的任何氨基酸。
57. 项 56 的可溶性 Stra6 多肽, 其包含图 2(SEQ ID NO:2)中 1 到 X 位的氨基酸, 其中 X 是图 2(SEQ ID NO:2)中从 49 到 59 位的任何氨基酸。

58. 一种分离的可溶性 Stra6 多肽, 其包含一种氨基酸序列, 该氨基酸序列与图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中从 1 到 X 位氨基酸的氨基酸序列相比时, 阳性分值至少是大约 80%, 其中 X 是图 2(SEQ ID NO:2)或图 7(SEQ ID NO:5)中从 49 到 59 位的任何氨基酸。

59. 一种确定怀疑有 Stra6 多肽的样品中所述多肽的存在的方法, 该方法包含使该样品暴露于抗 Stra6 抗体, 并检测所述抗体与所述样品中所述多肽的结合。

60. 项 59 的方法, 其中所述样品包含怀疑含有 Stra6 多肽的细胞。

61. 项 60 的方法, 其中所述细胞是癌症细胞。

62. 一种诊断哺乳动物肿瘤的方法, 该方法包含检测如下样品中编码 Stra6 多肽的基因的表达水平: (a)来自哺乳动物组织细胞的待测样品, 和(b)来自相同细胞类型的已知正常组织细胞的对照样品, 其中, 与对照样品相比较, 待测样品中较高的表达水平提示, 该待测组织细胞的来源哺乳动物中肿瘤的存在。

63. 一种诊断哺乳动物肿瘤的方法, 该方法包含(a)使抗 Stra6 抗体与来源于哺乳动物组织细胞的待测样品接触, 和(b)检测所述抗体与所述待测样品中 Stra6 多肽之间复合物的形成, 其中复合物的形成提示所述哺乳动物中肿瘤的存在。

64. 项 63 的方法, 其中所述抗体带有可检测的标记。

65. 项 63 的方法, 其中所述组织细胞的待测样品是从怀疑有肿瘤细胞生长或增生的个体获得的。

66. 一种癌症诊断试剂盒, 其在适当包装中包含抗 Stra6 抗体和一种载体。

67. 项 66 的试剂盒, 其进一步包含使用所述抗体检测怀疑有 Stra6 多肽的样品中该多肽的存在的说明。

68. 一种抑制肿瘤细胞生长的方法, 该方法包含使表达 Stra6 多肽的肿瘤细胞暴露于有效量的一种制剂, 该制剂抑制所述多肽的生物活性, 从而抑制所述肿瘤细胞的生长。

69. 项 68 的方法, 其中所述肿瘤细胞与相同组织类型的正常细胞相比, 所述 Stra6 多肽是过表达的。

70. 项 68 的方法, 其中所述制剂是抗 Stra6 抗体。

71. 项 70 的方法, 其中所述抗 Stra6 抗体诱导细胞死亡。

72. 项 68 的方法, 其中所述肿瘤细胞进一步暴露于放射治疗, 细胞毒制剂, 或化疗制剂。

73. 项 68 的方法, 其中所述肿瘤细胞进一步暴露于刺激编码 Stra6 的基因表达的制剂。

74. 项 73 的方法, 其中所述制剂是视黄酸或其类似物。

75. 项 68 的方法, 其中所述制剂是一种与编码 Stra6 多肽的核酸杂交的反义寡核苷酸或其互补体。

76. 项 75 的方法, 其中所述肿瘤细胞进一步暴露于放射治疗, 细胞毒制剂, 或化疗制剂。

77. 项 75 的方法, 其中所述肿瘤细胞进一步暴露于刺激编码 Stra6 的基因表达的制剂。

78. 项 77 的方法, 其中所述制剂是视黄酸或其类似物。

79. 一种制品, 其包含:

一种容器

容器上的标签; 和

一种组合物, 所述组合物包含容纳在所述容器中的活性制剂, 其中所述组合物可有效抑制肿瘤细胞的生长, 并且所述容器上的标签提示该组合物可有效治疗以所述肿瘤细胞相对于相同组织类型的正常细胞而言过表达 Stra6 多肽为特征的情况。

80. 项 79 的制品, 其中所述活性制剂抑制所述 Stra6 多肽的生物活性和/或表达。

81. 项 80 的制品, 其中所述活性制剂是抗 Stra6 抗体。

82. 项 81 的制品, 其中所述活性制剂是一种反义寡核苷酸。

83. 一种鉴定抑制 Stra6 多肽的生物学或免疫学活性的化合物的方法, 该方法包含使候选化合物与所述多肽在足以使这两种成分相互作用的条件下和时间内接触, 并确定所述多肽的生物学或免疫学活性是否被抑制。

84. 项 83 的方法, 其中所述候选化合物或所述 Stra6 多肽被固定在一固体支持物上。

85. 项 84 的方法, 其中的非固定成分带有可检测的标记。

86. 一种鉴定抑制 Stra6 多肽活性的化合物的方法, 该方法包含如下步

骤: (a)在适于通常由所述 Stra6 多肽诱导的细胞应答被诱导的条件下, 在所述 Stra6 多肽存在下, 使细胞与待筛选的候选化合物接触, 并(b)测定所述细胞应答的诱导, 以确定该待测化合物是否是 Stra6 活性的有效抑制剂, 其中所述细胞应答的诱导的缺失提示所述化合物是有效的抑制剂。

87. 鉴定抑制表达 Stra6 多肽的细胞中所述多肽表达的化合物的方法, 其中所述方法包含使所述细胞与候选化合物接触, 确定所述 Stra6 多肽的表达是否受到抑制。

88. 项 87 的方法, 其中所述候选化合物是一种抗体。

89. 项 87 的方法, 其中所述候选化合物是一种反义寡核苷酸。

90. 一种抑制 Stra6 多肽的生物学或免疫学活性的化合物, 其通过如下方法鉴定, 该方法包括使候选化合物与所述 Stra6 多肽在足以使这两种成分相互作用的条件下和时间内接触, 确定所述 Stra6 多肽的生物学或免疫学活性是否被抑制。

91. 项 90 的化合物, 其是一种抗体。

92. 项 90 的化合物, 其是一种小分子。

93. 抑制 Stra6 多肽活性的 Stra6 拮抗剂化合物, 其通过包括如下步骤的方法鉴定: (a)在适于通常由所述 Stra6 多肽诱导的细胞应答被诱导的条件下, 在所述 Stra6 多肽存在下, 使细胞与待筛选的候选化合物接触, (b)测定所述细胞应答的诱导, 以确定该待测化合物是否是有效的拮抗剂, 其中所述细胞应答的诱导的缺失提示所述化合物是有效的 Stra6 拮抗剂。

94. 项 93 的化合物, 其是一种抗体。

95. 项 93 的化合物, 其是一种小分子。

附图简述

图 1 显示含有编码天然序列 PRO10282 的核苷酸序列(核苷酸 1-2732)cDNA 的核苷酸序列(SEQ ID NO: 1), 其中该核苷酸序列(SEQ ID NO: 1)是本文称为"DNA 148380-2827"的克隆。也以粗字体和下划线标出各自的起始和终止密码子的位置。

图 2 显示衍生自 SEQ ID NO: 1 中编码序列的天然序列 PRO10282 多肽的氨基酸序列(SEQ ID NO: 2)。也显示各种其它重要多肽结构域的近似位置。

图 3A-3D 显示使用 ALIGN-2 序列比较计算机程序, 用下述方法测定氨

氨基酸序列同一性% (表 3 (A-B))和核酸序列同一性% (表 3 (C-D))的假设的例证, 其中"PRO"表示假设的目标 PRO10282 多肽的氨基酸序列, "对照蛋白质"表示使目标"PRO"多肽与其进行比较的多肽的氨基酸序列, "PRO-DNA"表示假设的编码 PRO10282 的目标核酸序列, "对照 DNA"表示使目标核酸分子"PRO-DNA"与其进行比较的核酸分子的核苷酸序列, "X"、"Y"和"Z"各自代表不同的假定氨基酸残基, 而"N"、"L"和"V"各自代表不同的假定核苷酸。

图 4A-4Q 提供 ALIGN 2 序列对比计算机程序的全部源序列码。为在 UNIX 操作系统上使用, 可以对该源序列码进行常规编制, 以提供 ALIGN 2 序列对比计算机程序。

图 5 显示本文指定为 DNA100038 (SEQ ID NO: 3) 的核苷酸序列。

图 6 显示含有编码天然序列人 Stra6 多肽变体的核苷酸序列 (核苷酸 1-2778)的 cDNA 的核苷酸序列 (SEQ ID NO: 4), 其中核苷酸序列 (SEQ ID NO: 4) 是本文指定为"DNA148389-2827-1"的克隆。也以粗体和下划线标出各自起始和终止密码子的位置。

图 7 显示衍生自 SEQ ID NO: 4 编码序列的天然序列人 Stra6 多肽变体的氨基酸序列(SEQ ID NO: 5)。也显示各种其它重要多肽结构域的近似位置。

图 8 是图式表示 DNA 148380-2827 (天然的人 PRO10282)编码的鼠 Stra6 和人 Stra6 蛋白。

图 9 显示 DNA 148380-2827 (天然的人 PRO10282)编码的天然序列人 Stra6 蛋白的疏水性。

图 10 显示 DNA148380-2827 编码的天然序列人 Stra 6 蛋白在各种正常人组织中的相对 RNA 表达情况。

图 11 显示用定量PCR 测定, DNA 148380-2827 编码的天然序列人 Stra6 蛋白在人结肠肿瘤组织中相对于在同一患者正常粘膜中 RNA 表达的 RNA 折叠(fold)表达。数据来自同一试验中所作的一式三份实验。使用不同套 PCR 引物, 至少重复试验 2 次。

图 12A 显示使用看家基因 GAPDH 作对照, DNA 148380-2827 编码的天然序列人 Stra6 蛋白在人结肠组织中相对于在同一患者正常粘膜中 RNA 表达的 RNA 表达。

图 12B 显示通过原位杂交, Stra6 在结肠腺癌的上皮肿瘤细胞上的定位。

图 13 显示 DNA148380-2827 编码的天然人 Stra6 蛋白在人乳腺、肾、结肠和肺肿瘤细胞系中相对于在相应正常细胞系中的 RNA 表达。

图 14 图例说明衍生自 DNA148380-2827 编码的天然序列人 Stra 6 蛋白的肽片段在大肠杆菌中的表达。

图 15 描述人结肠癌细胞中, 分别在存在和不存在全反式视黄酸(ATRA)和 9-顺式-视黄酸(9cRA)的条件下, Stra6 RNA 的表达。

图 16 是 Stra6 在肿瘤切片上的原位杂交情况。显示出证明银粒 (A, C, E, G)的暗视野图像和对应的苏木精-伊红染色的明亮视野图像(B, D, F, H)。在恶性黑素瘤中, 中等密度的银粒压在肿瘤细胞上而不压在血管上(A, B)。在子宫内膜腺癌中, 肿瘤上皮被适度地标记上而肿瘤基质呈阴性(C, D)。在 Wilm's 肿瘤的发育不全的区域显示高表达水平, 而肿瘤基质呈阴性(E, F)。嗜铬细胞瘤显示非常高的 Stra6 mRNA 表达, 而相邻的正常肾上腺皮质则呈阴性(G, H)。刻度棒 = 100 微米。

图 17 (A) 对 9-顺式-RA 或全反-RA 应答, 在 C57MG/ Parent 和 C57MG/Wnt-1 细胞中 Stra6 mRNA 表达的诱导。(B) 对条件介质中的 Wnt-3A 和 9-顺式-RA 应答, 在 C57MG/Parent 细胞中 Stra6 mRNA 表达的诱导。(C) 视黄酸处理后, 在 HCT116 和 WiDr 结肠腺癌细胞中 Stra6 mRNA 表达的诱导。(D) 用视黄酸处理前(上组)和处理后(下组), 经在 HCT116 细胞原位杂交证明 Stra6 表达的暗野图像。(E) 视黄酸治疗前(-RA)和治疗后(+RA), 通过用直接针对人 Stra6 肽 B 的单克隆抗体进行的 Western 印迹, 观察到 WiDr 细胞中 Stra6 蛋白的表达。(F) 在未用视黄酸处理(左组)或用视黄酸处理(右组)的 WiDr 细胞中 Stra6 的膜定位。用抗人 Stra6 肽 B 单克隆杂交瘤培养物上清液(克隆 12F4. 2H9. 1D5)进行免疫组织化学检测。试验用 A-F 表示, 用视黄酸处理细胞 48 小时。完成定量 PCR 反应(每个进行 40 循环)后得到的 Stra6 产物示于下面每一图的 A-C 中。

图 18 (A) Wnt-1 诱导 RAR γ -1 表达。在不含四环素 0、24、48 或 72 小时的条件下, 得自四环素可抑制细胞 C57MG/Wnt 1 细胞的蛋白-等量的全细胞溶解产物, 接受 SDS-PAGE 以及 RAR γ -1 和 ERK2 的免疫印迹分析。(B) 在 Wnt-1 转基因小鼠增生的乳腺和乳腺肿瘤中 RAR γ -1 mRNA 的表达。用定量 RT-PCR 测定 mRNA 的表达, 相对于野生型乳腺中 mRNA 的表达, 数据用表达倍数进行表述。

优选实施方案详述

I. 定义

术语"PRO10282 多肽"、"PRO10282 蛋白"、"PRO10282"、"Stra6 多肽"、"Stra6 蛋白" 和"Stra6"互换使用, 并涵盖天然序列 PRO10282 (Stra6) 和 PRO10282 (Stra6)多肽变体 (本文将作进一步定义)。PRO10282 (Stra6)多肽可以是各种来源(如人组织或其它来源)分离得到的, 或者用重组和/或合成方法制得。

"天然序列 PRO10282"或"天然序列 Stra6", 包括具有与得自天然 PRO10282 相同的氨基酸序列的多肽。此天然序列 PRO10282 (Stra6)可以是自然界分离得到的, 或者用重组和/或合成方法制得的。术语"天然序列 PRO10282"或"天然序列 Stra6", 特别包括 PRO10282 的天然截短型或分泌型(例如, 细胞外结构域序列)、天然变体型(例如, 可替换式剪切型)和天然等位变体。在本发明一实施方案中, 天然序列 PRO10282 是包括图 2(SEQ ID NO: 2)中 1 至 667 氨基酸的成熟或全长天然序列 PRO10282。在本发明另一实施方案中, 天然序列 PRO10282 多肽是包括 SEQ ID NO: 5 中 1 至 658 氨基酸的成熟或全长天然序列 PRO19578 多肽, 后者据信是 SEQ ID NO: 2 的天然序列 PRO10282 的可替换式剪切型。并且, 尽管图 2 (SEQ ID NO: 2)和图 7 (SEQ ID NO: 5)公开的 PRO10282 多肽显示由本文定义为位置 1 的蛋氨酸残基开始, 但是图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)中氨基酸 1 位上游或下游的另一蛋氨酸残基作为 PRO10282 多肽的起始氨基酸残基是可以想象到的和可能的。

PRO10282 多肽(Stra6) "细胞外结构域"或 "ECD", 是指 PRO10282 多肽基本上不含跨膜和胞质结构域的一种形式。通常, PRO10282 多肽 ECD 具有小于约 1 % 的此类跨膜和/或胞质结构域, 优选具有小于约 0.5 % 的此类结构域。应当懂得, 所确定的本发明 PRO10282 多肽的任何跨膜结构域(一个或多个) 是依据本领域在鉴定疏水结构域类型中常规使用的标准而确定的。跨膜结构域的精确的边界会有不同, 但在最初确定的结构域的两末端之一最多不超过约 5 个氨基酸。因此, 在本发明的一实施方案中, 人 PRO10282 多肽的细胞外结构域包括氨基酸 1 至 X; 其中 X 是图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)中从氨基酸 49 到 59 的任何氨基酸。

"PRO10282 变体多肽"或"Stra6 变体多肽", 这些术语是可互换使用的, 它们指如下定义的与下列氨基酸序列具有至少约 80%氨基酸序列同一性的活性 PRO10282 多肽(Stra6): (a)图 2 (SEQ ID NO: 2)所示 PRO10282 多肽的 1 至 667 残基或图 7 (SEQ ID NO: 5)所示 PRO10282 多肽的 1 至 658 残基的氨基酸序列, (b)图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)中 1 至 X 残基的氨基酸序列, 其中 X 是图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)中氨基酸 49 至 59 的任何氨基酸, 或者(c)衍生自图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)所示氨基酸序列的另一特定片段。此类 PRO10282(Stra6)变体多肽包括, 例如, 在图 2 (SEQ ID NO: 2)或图(SEQ ID NO: 5)中所示序列的 N-和/或 C-末端以及在一个或多个内部结构域内增加或删减一个或多个氨基酸残基的 PRO10282 多肽。通常, PRO10282 变体多肽与(a)图 2 (SEQ ID NO: 2)所示 PRO10282 多肽的 1 至 667 残基的氨基酸序列或图 7 (SEQ ID NO: 5)所示 PRO10282 多肽的 1 至 658 残基的氨基酸序列, (b)图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)中 1 至 X 残基的氨基酸序列, 其中 X 是图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)中氨基酸 49 至 59 的任何氨基酸, 或者(c) 衍生自图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)所示氨基酸序列的另一特定片段, 具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98%或至少约 99%的氨基酸序列同一性。PRO10282 变体多肽不包括天然 PRO10282 多肽序列。一般地, PRO10282 变体多肽的长度是至少约 10 个氨基酸, 经常是至少约 20 个氨基酸, 更经常是至少约 30 个氨基酸, 更经常是至少约 40 个氨基酸, 更经常是至少约 50 个氨基酸, 更经常是至少约 60 个氨基酸, 或者至少约 70 个氨基酸, 更经常是至少约 80 个氨基酸, 更经常是至少约 90 个氨基酸, 更经常是至少约 100 个氨基酸, 或者至少约 150 个氨基酸, 更经常是至少约 200 个氨基酸, 更经常是至少约 250 个氨基酸, 更经常是至少约 300 或更多个氨基酸, 并且不同于鼠 Stra6 序列的片段, 见 Bouillet 等, Mechanisms of Development 63, 173-186 (1997)。

关于本文鉴定的 PRO10282 (Stra6)多肽序列的"氨基酸序列同一性百分

比(%)"，是指候选序列的氨基酸残基，在进行序列对比并在必要时导入空隙以获取最大百分比的序列同一性，而不将任何保守取代视为序列同一性时，与 PRO10282 序列的氨基酸残基相同的百分数。可使用本领域各种方法进行序列对比以便测定氨基酸序列同一性百分比，例如，使用公众可得到的计算机软件如 BLAST、BLAST-2、ALIGN、ALIGN-2 或 Megalign (DNASTAR) 软件。本领域技术人员可以决定测量对比的适宜参数，包括针对所比较的序列全长获得最大对比所需的任何算法。然而，为此目的，氨基酸序列同一性 % 值是使用下述序列对比计算机程序 ALIGN-2 而获得的，其中 ALIGN-2 程序的全部源代码见表 4A-Q。ALIGN-2 序列对比计算机程序的作者是 Genentech, Inc.，表 4A-Q 所示源代码和用户数据一起已经提交地处 Washington D. C.，20559 的美国版权局，其美国版权注册登记号为 TXU510087。公众通过 Genentech, Inc., South San Francisco, California 可以得到 ALIGN-2 程序，或者可从表 4A-Q 提供的源代码进行编制。ALIGN2 程序应当为在 UNIX 操作系统，优选在数码 UNIX V4.0D 使用而进行编制。ALIGN-2 程序设定了所有序列对比参数并且不变。

为了本发明目的，给定氨基酸序列 A 相对于给定氨基酸序列 B 的氨基酸序列同一性 % (或者说：给定氨基酸序列 A 具有或含有给定氨基酸序列 B 相同氨基酸序列的 %) 如下计算：

X/Y 比值乘以 100

其中 X 是用序列对比程序 ALIGN-2 比较 A 和 B 的氨基酸残基后计算出的相同氨基酸数，其中 Y 是 B 的氨基酸残基总数。可以理解，当氨基酸序列 A 与氨基酸序列 B 的长度不相等时，A 相对于 B 的氨基酸序列同一性 % 将不等于 B 相对于 A 的氨基酸序列同一性 %。作为计算氨基酸序列同一性 % 的一个实例，表 3A-B 说明如何计算指定为"对照蛋白"的氨基酸序列与指定为"PRO"的氨基酸序列的序列同一性 %。

除非另外特别声明，否则本文所用所有氨基酸序列同一性 % 值是按照上述使用 ALIGN-2 序列对比计算机程序获得的。然而，氨基酸序列同一性 % 也可使用序列对比程序 NCBI-BLAST2 (Altschul 等, Nucleic acids Res. 25: 3389-3402 (1997)) 进行测定。NCBI-BLAST2 序列对比程序可从 <http://www.ncbi.nlm.nih.gov> 下载。NCBI-BLAST2 使用数种检索参数，其中所有这些参数设定为默认值，包括例如，非掩蔽=是(yes)，链(strand) =全部(all)，预

期的出现=10, 最低复杂长度=15/5, 多-程 e-值 = 0.01, 多-程常数= 25, 最终缺口化对比的陡降(dropoff)= 25, 和评分矩阵= BLOSUM62。

当使用 NCBI-BLAST2 进行氨基酸序列比较时, 给定氨基酸序列 A 针对给定氨基酸序列 B 的氨基酸序列同一性%(或者说, 给定氨基酸序列 A 具有或含有与给定氨基酸序列 B 相同的氨基酸序列的%)如下计算:

X/Y 比值乘以 100

其中 X 是用序列对比程序 NCBI-BLAST2 比较 A 和 B 氨基酸残基后计算出的相同氨基酸数量, 其中 Y 是 B 的氨基酸残基总数量。可以理解, 当氨基酸序列 A 与氨基酸序列 B 的长度不相等时, A 针对 B 的氨基酸序列同一性%将不等于 B 针对 A 的氨基酸序列同一性%。

"PRO10282 变体多核苷酸"或"PRO10282 变体核酸序列"是指编码下述定义的活性 PRO10282 多肽的多核苷酸分子, 它与(a)编码图 2 (SEQ ID NO: 2)所示 PRO10282 多肽的 1 至 667 残基的氨基酸序列的核苷酸序列或编码图 7 (SEQ ID NO: 5)中 1 至 658 残基的氨基酸序列的核苷酸序列, (b) 编码图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)中 1 至 X 氨基酸的核苷酸序列, 其中 X 是图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)中 49 至 59 的任何氨基酸, 或者(c) 编码衍生自图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)所示氨基酸序列的另一特定片段的核苷酸序列, 具有至少约 80% 的氨基酸序列同一性。通常, PRO10282 变体多核苷酸与(a)编码图 2 (SEQ ID NO: 2)所示 PRO10282 多肽的 1 至 667 残基的氨基酸序列的核苷酸序列或编码图 7 (SEQ ID NO: 5)中 1 至 658 残基的氨基酸序列的核苷酸序列, (b) 编码图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)中 1 至 X 氨基酸的核苷酸序列, 其中 X 是图 2 (SEQ ID NO: 2) 或图 7 (SEQ ID NO: 5)中 49 至 59 的任何氨基酸, 或者(c) 编码衍生自图 2 (SEQ ID NO: 2)或图 7 (SEQ ID NO: 5)所示氨基酸序列的另一特定片段的核苷酸序列, 具有至少约 80%、优选至少约 81%、更优选至少约 82%、更优选至少约 83%、更优选至少约 84%、更优选至少约 85%、更优选至少约 86%、更优选至少约 87%、更优选至少约 88%、更优选至少约 89%、更优选至少约 90%、更优选至少约 91%、更优选至少约 92%、更优选至少约 93%、更优选至少约 94%、更优选至少约 95%、更优选至少约 96%、更优选至少约 97%、更优选至少约 98% 或至少约 99% 的核酸序列同一性。PRO10282 多核苷酸变体不包括天然 PRO10282 核苷酸序列或天然 PRO19578 核苷酸序列。

一般地, PRO10282(Strat6)变体多核苷酸的长度是至少约 30 个核苷酸, 经常是至少约 60 个核苷酸, 更经常是至少约 90 个核苷酸, 更经常是至少约 120 个核苷酸, 或者至少约 150 个核苷酸, 更经常是至少约 180 个核苷酸, 更经常是至少约 210 个核苷酸, 更经常是至少约 240 个核苷酸, 或者至少约 270 个核苷酸, 更经常是至少约 300 个核苷酸, 更经常是至少约 450 个核苷酸, 或至少 600 个核苷酸, 或至少 900 或更多个核苷酸, 而且特别排除包含在鼠 Strat6 核苷酸序列内的多核苷酸, 参见 Bouillet 等, Mechanisms of Development 63, 173-186 (1997)所公开的内容。

有关本文鉴定的编码 PRO10282(Strat6)多肽的核酸序列的"核酸序列同一性百分比(%)",是指候选序列中与编码PRO10282多肽的核酸序列相比时, 具有相同核苷酸的百分数, 这是在对比序列, 和如果必要则引入空隙, 以达到最大序列同一性百分比的前提下获得的。为测定核苷酸序列同一性百分比而进行的序列对比, 可使用本领域各种方法, 例如, 使用公众可得到的计算机软件如 BLAST、BLAST-2、ALIGN、ALIGN-2 或 Megalign(DNASTAR)软件。本领域技术人员可以决定用于测量对比的适宜参数, 包括针对所比较序列的全长获得最大对比所需的任何算法。然而, 为此目的, 核苷酸序列同一性 % 值是如下述使用序列对比计算机程序 ALIGN-2 而获得的, 其中 ALIGN-2 程序的全部源代码见表 4A-Q。ALIGN-2 序列对比计算机程序的作者是 Genentech, Inc., 表 4A-Q 所示源代码和用户数据一起已经提交地处 Washington D. C., 20559 的美国版权局, 其美国版权注册登记号为 TXU510087。公众通过 Genentech, Inc., South San Francisco, California 可以得到 ALIGN-2 程序, 或者从表 1 提供的源代码进行编制。ALIGN2 程序应当为在 UNIX 操作系统, 优选在数码 UNIX V4.0D 使用而进行编制。ALIGN-2 程序设定了所有序列对比参数并且不变。

为了本发明目的, 给定核酸序列 C 相对于给定核酸序列 D 的核酸序列同一性%(或者说: 给定核酸序列 C 具有或含有与给定核酸序列 D 相同的核酸序列的%)如下计算:

W/Z 比值乘以 100

其中 W 是用序列对比程序 ALIGN-2 比较 C 和 D 后计算出的相同核苷酸的数量, 和其中 Z 是 D 的核苷酸总数量。可以理解, 当核酸序列 C 与核酸序列 D 的长度不相等时, C 针对 D 的核酸序列同一性%将不等于 D 针对

C 的核酸序列同一性%。作为计算核酸序列同一性%的一个实例,表 3C-D 说明如何计算指定为"对照 DNA"的核酸序列与指定为"PRO-DNA"的核酸序列的同一性%。

除非另外特别声明,否则本文所用所有核酸序列同一性%值是按照上述使用 ALIGN-2 序列对比计算机程序获得的。然而,核酸序列同一性%也可使用序列对比程序 NCBI-BLAST2 (Altschul 等, Nucleic Acids Res. 25: 3389-3402 (1997))进行测定。NCBI-BLAST2 序列对比程序可从 <http://www.ncbi.nlm.nih.gov> 下载。NCBI-BLAST2 使用数种检索参数,其中所有这些参数设定为默认值,包括例如,非掩蔽=yes,链=全部,预期的出现=10,最低复杂长度=15/5,多-程 e-值=0.01,多-程常数=25,最终缺口化对比的陡降=25,和评分矩阵=BLOSUM62。

当使用 NCBI-BLAST2 进行序列比较时,给定核酸序列 C 与给定核酸序列 D 的核酸序列同一性%(或者说,给定核酸序列 C 具有或含有与给定核酸序列 D 相同的核酸序列的%)如下计算:

W/Z 比值乘以 100

其中 W 是用序列对比程序 NCBI-BLAST2 比较 C 和 D 后计算出的相同核苷酸数量,和其中 Z 是 D 的核苷酸总数量。可以理解,当核酸序列 C 与核酸序列 D 的长度不相等时,C 针对 D 的核酸序列同一性%将不等于 D 针对 C 的核酸序列同一性%。

在其它实施方案中,PRO10282(Strat6)变体多核苷酸是编码活性 PRO10282 多肽的核酸分子,它能够与编码图 2(SEQ ID NO: 2)或图 7(SEQ ID NO: 5)所示全长 PRO10282 多肽的核苷酸序列杂交(优选地在严格杂交和洗涤条件下杂交)。PRO10282 变体多肽可以是由 PRO10282 变体多核苷酸编码的那些多肽。

在如上述进行氨基酸序列同一性比较时使用的术语"阳性",不仅包括所比较的序列中相同的氨基酸残基,也包括其中具有相似特性的氨基酸残基。针对目的氨基酸残基评分为阳性值的氨基酸残基是那些与目的氨基酸残基相同或是目的氨基酸残基的首选取代的氨基酸残基。(见下表 1 中定义)。

为了本发明目的,给定氨基酸序列 A 相对于给定氨基酸序列 B 的氨基酸序列阳性值%(或者这样说:给定氨基酸序列 A 具有或含有与给定氨基酸序列 B 相同的氨基酸序列的阳性%)如下计算:

X/Y 比值乘以 100

其中 X 是用上述序列对比程序 ALIGN-2 比较 A 和 B 后得出的如上定义的评分阳性氨基酸数量，其中 Y 是 B 的氨基酸残基总数量。可以理解，当氨基酸序列 A 与氨基酸序列 B 的长度不相等时，A 针对 B 的阳性 % 将不等于 B 针对 A 的阳性 %。

当用“分离的”来描述本文的各种多肽时，是指已从天然环境的组成中鉴定和分离和/或回收的多肽。优选地，分离的多肽与其天然结合的所有组分没有任何结合。多肽的天然环境中的污染成分是那些将干扰使用多肽进行诊断和治疗的物质，可包括酶、激素和其它蛋白性或非-蛋白性溶质。在优选的实施方案中，多肽被纯化为：(1)使用转杯式蛋白测序仪，达到足以获得 N-末端或内部氨基酸序列至少 15 个残基的程度，或者(2)达到在非-还原或还原条件下经 SDS-PAGE 电泳和考马斯蓝或优选银染色证实为同质性的程度。分离的多肽包括在重组细胞内的原位多肽，因为 PRO10282(Strat6)自然环境中的至少一个组分是不存在的。然而，通常用至少一个纯化步骤来制备分离的多肽。

编码 PRO10282(Strat6)多肽的“分离的”核酸分子，是从 PRO10282 编码核酸的天然来源中通常与其结合的至少一种杂质核酸分子中鉴定并分离的核酸分子。优选地，所述分离的核酸与其天然结合的所有组分没有任何结合。一分离的编码 PRO10282 的核酸分子不同于其天然形式或组合(setting)。因此，分离的编码 PRO10282 的核酸分子区别于其在天然细胞中的存在形式。但是，编码 PRO10282(Strat6)多肽的分离的核酸分子，包括通常表达 PRO10282 的细胞中所含的 PRO10282-编码核酸分子，例如，所述核酸分子与天然细胞中相应核酸分子所在的染色体位置不同。

术语“控制序列”，指在特定宿主生物中使可操作连接的编码序列表达所需的 DNA 序列。适于原核生物的调控序列可包括如启动子，任选操纵子序列，和核糖体结合位点。已知真核生物细胞利用启动子、聚腺苷酸化信号和增强子。

当一种核酸与另一种核酸序列处于功能相关的位置时，该核酸与该另一种核酸“可操作相连”。例如如果一种多肽表达为参与该多肽的分泌的前体蛋白，则将前序列或分泌前导序列的 DNA 与该多肽的 DNA 可操作相连；启动子或增强子当能影响编码序列的转录时可与该编码序列可操作相连；或

核糖体结合位点当其位置能促进翻译时可与编码序列可操作相连。通常“可操作相连”指被连接的 DNA 序列是相邻的，而且，如果是分泌前导序列，则相邻且为阅读状态。然而，增强子不必是相邻的。连接可通过在方便的限制性位点相连而实现。如果不存在这样的位点，可根据常规实践使用合成的寡核苷酸衔接子或接头。

本文所用术语“抗体”是指最广义上的抗体，并特别包括例如单链抗-PRO10282 单克隆抗体(包括激动剂、拮抗剂和中和抗体)、具有多表位特异性的抗-PRO10282 抗体组合物、单链抗-PRO10282 抗体和抗-PRO10282 抗体的片段(见下述)。本文中术语“单克隆抗体”，是指来自基本上同质的抗体群的抗体，即除了可能少量存在的天然突变以外，该抗体群中的各个抗体均相同。

杂交反应的“严格度”由本领域技术人员很容易地确定，通常根据经验在考虑探针长度、洗涤温度和盐浓度的基础上计算得出。一般而言，长的探针需要较高的温度以便正确退火，而短的探针需要较低温度。杂交通常取决于变性的 DNA 在低于解链温度的环境中存在互补序列时的重退火能力。探针和杂交序列之间所需同源程度越高，则可使用的相对温度越高。结果是，相对温度较高将使反应条件趋于更加严格，而较低温度则使反应的严谨度较小。有关杂交反应的严格度的其它细节和解释，见 Ausubel 等，Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995)。

本文的“严格条件”或“高度严格条件”，可以定义为：(1) 使用低离子强度和高温洗涤，例如，50℃ 0.015M 氯化钠/0.0015M 柠檬酸钠/0.1% 十二烷基硫酸钠；(2) 在杂交过程中 42℃使用变性剂如甲酰胺，譬如，含 0.1% 牛血清白蛋白的 50% (v/v) 甲酰胺/0.1% Ficoll/0.1% 聚乙烯吡咯烷酮/50mM 磷酸钠缓冲液 pH 6.5 及 750 mM 氯化钠，75 mM 柠檬酸钠；或者(3) 42℃使用 50% 甲酰胺、5 x SSC (0.75M NaCl, 0.075 M 柠檬酸钠)、50 mM 磷酸钠(pH 6.8)、0.1% 焦磷酸钠、5 x Denhardt's 溶液、超声处理的鲑精 DNA(50 µg/ml)、0.1% SDS 和 10% 硫酸葡聚糖，于 42℃在 0.2 x SSC(氯化钠/柠檬酸钠)中和于 55℃在 50% 甲酰胺中洗涤，接着于 55℃用含 EDTA 的 0.1 x SSC 进行高度严格洗涤。

“中等严格条件”可以按照 Sambrook 等，Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989 的定义，包括使用严格度

低于前述的洗涤溶液和杂交条件(例如, 温度, 离子强度和 SDS %)。中等严格条件的一个实例是, 于 37℃ 在如下组成的溶液中过夜温育: 20 % 甲酰胺, 5 x SSC (150 mM NaCl, 15 mM 柠檬酸三钠), 50 mM 磷酸钠(pH 7.6), 5 x Denhardt's 溶液, 10 % 硫酸葡聚糖和 20 mg/ml 变性剪切的鲑精 DNA, 接着于 37-50℃ 在 1 x SSC 中洗涤杂交膜。本领域熟练技术人员懂得如何必要地调整温度、离子强度等条件以适应诸如探针长度等因素。

本文使用的术语"表位标签", 是指含有与"标签多肽"融合了的 PRO10282 多肽的嵌合多肽。该标记多肽具有足够多的残基以提供针对所制备抗体的表位, 而且要足够短以便它不干预与其融合的多肽的活性。标记多肽也优选确实是独一无二的, 以便该抗体与其它表位基本上不发生交叉反应。通常, 适宜的标记多肽至少有 6 个氨基酸残基, 一般为约 8~50 个氨基酸残基(优选为约 10~20 个氨基酸残基)。

本文所用术语"抗体"是指最广义上的抗体, 并特别包括例如单链抗-PRO10282(抗-Stra6)单克隆抗体(包括激动剂、拮抗剂和中和抗体)、具有多表位特异性的抗-PRO10282 抗体组合物、单链抗-PRO10282 抗体和抗-PRO10282 抗体的片段(见下述)。

本文中术语"单克隆抗体", 是指来自基本上同质的抗体群的抗体, 即除了可能少量存在的天然突变以外, 该抗体群中的各个抗体均相同。

"抗体片段"包含完整抗体的一部分, 优选是完整抗体的抗原结合或可变区。抗体片段的实例包括 Fab、Fab'、F(ab')₂ 和 Fv 片段; 二价抗体; 线形抗体(Zapata 等, Protein Eng. 8 (10): 1057-1062 [1995]); 单链抗体分子; 和由抗体片段形成的多特异抗体。

木瓜蛋白酶消化抗体可产生两个相同的各带有单个抗原结合位点的抗原结合片段(称为"Fab"片段)和残余的"Fc"片段, Fc 段的名称反应了其易于结晶的能力。经胃蛋白酶处理可产生具有两个抗原结合位点并仍然能交联抗原的 F(ab')₂ 片段。

"Fv"是含有完整的识别和结合位点的最小抗体片段。此区由一个重链可变区与一个轻链可变区紧密地非共价连接形成的二聚体组成。在这个构象中每个可变区的三个 CDR 相互作用, 在 V_H-V_L 二聚体表面限定一个抗原结合位点。这六个 CDR 共同赋予抗体以抗原结合特异性。然而, 即使是单个可变区(或 Fv 的一半仅含有三个抗原特异性 CDR)也具有识别和结合抗原的

能力, 尽管其比完整的结合位点具有较低的亲和力。

Fab 段还包括轻链恒定区和重链的第一个恒定区(CH1)。Fab' 与 Fab 的差别在于 Fab 在重链 CH1 的羧基末端多出几个残基, 其包括抗体铰链区的一个或多个半胱氨酸。Fab'-SH 在本文中是指恒定区半胱氨酸残基中至少有一个游离巯基的 Fab'。F(ab')₂ 抗体片段在最初产生为 Fab' 片段对, 在它们之间具有铰链区半胱氨酸。抗体片段的其它化学偶联是众所周知的。

脊椎动物任何物种的抗体(免疫球蛋白)的“轻链”, 可依据其恒定区氨基酸序列而归为完全不同的两型(称为 k 和 λ)中的一型。

根据其重链恒定区的氨基酸序列, 可将免疫球蛋白分为不同类。主要有 5 类免疫球蛋白: IgA、IgD、IgE、IgG 和 IgM, 其中一些还可进一步分成“亚类”(同种型), 例如 IgG1、IgG2、IgG3、IgG4、IgA1 和 IgA2。对应于不同类抗体的重链恒定区, 分别称为 α 、 δ 、 ϵ 、 γ 和 μ 。不同类免疫球蛋白的亚单位结构和三维构象是众所周知的。

术语“二价抗体”是指具有两个抗原结合位点的小分子抗体片段, 这些片段在一条多肽链(V_H-V_L)上含有相连的一个重链可变区(V_H)和一个轻链可变区(V_L)。利用一种非常短的接头, 其使得同一条链上的两个结构域无法配对, 不得不与另一条链上的互补结构域配对, 从而形成两个抗原结合位点。例如, 在 EP 404, 097; WO93/11161; 和 Hollinger 等, Proc. Natl. Acad. Sci. USA, 90: 6444-6448(1993)中有对二价抗体的更详细描述。

“分离的”抗体是已从其天然环境的组成中鉴定、分离和/或回收的抗体。其天然环境的污染成分是干扰抗体的诊断或治疗应用的物质, 可包括酶、激素和其它蛋白性或非蛋白性溶质。在优选的实施方案中, 该抗体的纯度应达到: (1)经 Lowery 法确定的抗体重量的 95% 以上, 最优选所述重量的 99% 以上, (2)足以获得用旋转杯状(spinning cup)序列分析仪所测至少 15 个残基的 N 端或内部氨基酸序列, (3)通过还原或非还原条件下的 SDS-PAGE 以及考马斯亮蓝染色、优选银染色所证实的同质性。分离的抗体包括重组细胞内的原位抗体, 因为该抗体的自然环境中的至少一种组分已不存在。然而, 一般情况下分离的抗体可通过至少一个纯化步骤来制备。

术语“可变”是指可变区某些部分的序列在抗体之间有很大差异, 它们可在各个抗体对其特定抗原的结合和特异性方面发挥作用。然而, 该变异性不是均匀的分布于整个抗体的可变区。它集中于轻链和重链可变区中三个称

为“互补决定区”(CDR)或超变区的节段中。可变区中保守性较高的区域称为框架区(FR)。天然重链和轻链的可变区各包括4个FR区,主要采取 β -片层构象,由形成环状连接的三个CDR相连接,而在某些情况下可形成部分 β 片层结构。每条链的CDR通过FR区紧密的靠近在一起,并且与其它链的CDR一起形成抗体的抗原结合位点(见Kabat等, NIH Publ. No.91-3242, 第I卷, 第647-669页(1991))。恒定区不直接参与抗体与抗原的结合,但是表现出各种效应功能,例如参与抗体的抗体依赖性细胞毒性作用。

术语“超变区”在本文中是指抗体上负责结合抗原的氨基酸残基。所述超变区通常包括“互补决定区”或“CDR”的氨基酸残基(即轻链可变区的残基24-34(L1)、50-56(L2)和89-97(L3)和重链可变区的残基31-35(H1)、50-65(H2)和95-102(H3); Kabat等, Sequences of Proteins of Immunological Interest, 第5版, Public Health Service, National Institutes of Health, Bethesda, MD(1991))和/或那些“超变环”的氨基酸残基(即轻链可变区的残基26-32(L1)、50-52(L2)和91-96(L3)和重链可变区的残基26-32(H1)、53-55(H2)和96-101(H3); Chothia和Lesk, J. Mol. Biol. 196: 901-917(1987))。“框架区”或“FR”残基是那些除了本文所定义的超变区残基以外的可变区残基。

“人源化”型非人(例如鼠)抗体是嵌合的免疫球蛋白、免疫球蛋白链或其片段(如Fv、Fab、Fab'、F(ab')₂或抗体的其它抗原结合亚序列),它们包含衍生自非人免疫球蛋白的最小序列。在很大程度上,人源化抗体是人免疫球蛋白(受者抗体),其中受者互补决定区(CDR)残基被具有所需特异性、亲和力和性能的小鼠、大鼠、家兔或非人灵长类等非人源物种抗体(供体抗体)的CDR残基所取代。在一些实例中,人免疫球蛋白的Fv FR残基由相应的非人类残基所取代。而且,人源化抗体可包括在受者抗体或供体CDR或框架序列中均不存在的残基。这些修饰旨在进一步改善和最大化抗体的性能。通常,人源化抗体基本上包括至少一个(通常包括两个)可变区的全部,其中CDR区的全部或基本上全部对应于非人免疫球蛋白的相应部分,而FR序列的全部或基本上全部是人免疫球蛋白序列。人源化抗体还任选包括免疫球蛋白恒定区(Fc),通常为人免疫球蛋白的恒定区的至少一部分。详见Jones等, Nature, 321: 522-525(1986); Riechmann等, Nature 332: 323-329(1988); 和Presta, Curr. Op. Struct. Biol. 2: 593-596(1992)。任选地,人源化抗体也包括PRIMATIZEDTM抗体,所述抗体的抗原结合区可从用目标抗原免疫接

种猕猴制得的抗体中衍生获得。

本文中所用术语“免疫粘附素”定义为抗体样分子，其将异源蛋白(一种“粘附素”)的结合特异性与免疫球蛋白恒定区的效应器功能组合。结构上，免疫粘附素包括具有所需结合特异性之氨基酸序列与免疫球蛋白恒定区序列的融合，所述氨基酸序列不同于抗体的抗原识别和结合位点(即为“异源性”)。免疫粘附分子的粘附素部分通常是至少包括一个受体或一个配体的结合位点的邻接氨基酸序列。免疫粘附素中的免疫球蛋白恒定区序列，可以得自任何免疫球蛋白，如 IgG-1、IgG-2、IgG-3 或 IgG-4 亚型、IgA (包括 IgA-1 和 IgA-2)、IgE、IgD 或 IgM。

本文中“活性”是指保留完整或天然 PRO10282 的生物学和/或免疫学活性的 PRO10282 形式，其中“生物学”活性指完整或天然 PRO10282 引起的生物学功能(抑制或兴奋)，并且不包括诱导产生针对完整或天然 PRO10282 所拥有的表位的抗体的能力；而“免疫学”活性则是指诱导产生针对完整或天然 PRO10282 所携表位的抗体的能力。优选的生物活性包括例如下述任何一种或多种活性：应答视黄酸而在生长、发育或分化中的作用。基于与 Stra6(一种应答视黄酸而诱导的小鼠蛋白)的密切相似性，本文公开的 PRO10282 多肽，可能在例如早期的背腹侧四肢结构模式和后期的软骨内成骨中起着重要的作用。另一生物学活性涉及在肿瘤发展和生长中的作用。

本文所用“免疫学活性”，优选是“免疫交叉-反应性”，是指候选多肽能够与用已知活性 PRO10282 (Stra6)多肽制备(raised)的多克隆抗血清竞争性抑制具有此活性的 PRO10282 (Stra6)多肽的定性生物学活性。以常规方式制备此类抗血清，例如给山羊或兔子皮下注射在完全 Freund's 佐剂中的已知活性类似物，接着经腹膜内或皮下注射在不完全 Freund's 佐剂中的已知活性类似物以加强。免疫交叉反应性优选是“特异的”，这意味着所鉴定的免疫交叉-反应分子(例如抗体)与相应 Stra6 多肽的结合亲和力显著高于(优选高至少约 2-倍，更优选至少约 4-倍，更加优选至少约 8 倍，最优选高至少约 10-倍)其与任何其它已知天然多肽的结合亲和力。

本文中术语“拮抗剂”，使用其广义含义，包括部分或完全阻断、抑制或中和本文所述天然 PRO10282 多肽生物活性的任何分子。类似地，术语“激动剂”也使用其广义含义，包括模拟本文所述天然 PRO10282 多肽生物活性的任何分子。适宜的激动剂或拮抗剂分子特别包括：激动剂或拮抗剂抗体或

抗体片段、天然 PRO10282 多肽的片段或氨基酸序列变体、肽、小的有机分子等。鉴定 PRO10282 多肽的激动剂或拮抗剂的方法可以是,使 PRO10282 多肽与待测激动剂或拮抗剂分子接触,并测定正常情况下与 PRO10282 多肽有关的一个或多个生物活性的可测得变化。

本文中"肿瘤"是指所有赘生细胞生长和增殖(无论是恶性还是良性肿瘤),和所有癌前和癌性细胞及组织。

术语"癌"和"癌性的"指或描述哺乳动物中通常以无控细胞生长为特征的生理状况。癌的实例包括但不限于:癌、淋巴瘤、胚细胞瘤、肉瘤和白血病。此类癌的更具体的实例包括 乳腺癌、前列腺癌、结肠癌、鳞状细胞癌、小细胞肺癌、非-小细胞肺癌、胃肠癌、胰腺癌、成胶质细胞瘤、宫颈癌、卵巢癌、肝癌、膀胱癌、肝细胞瘤、结肠直肠癌、子宫内膜癌、唾液腺癌、肾癌、肝癌、外阴癌、甲状腺癌、肝脏癌和各种类型头颈部癌。

"治疗",既指治疗性治疗也指预防性措施,其目的是阻止或延缓(减轻)所针对的病理性病症。那些需要治疗的个体包括,已经患有病症或有患病倾向或者为了预防患上病症的个体。在肿瘤(例如癌)治疗中,治疗剂可直接减少肿瘤细胞的病状,或使得肿瘤细胞对其它治疗剂的治疗(例如,放疗和/和化疗)更为敏感。

癌的"病理学"包括危害患者健康的所有现象。包括但不限于:异常或失去控制的细胞生长、转移、干扰毗邻细胞的正常功能、异常水平释放细胞因子或其它分泌产物、抑制或加重炎症或免疫应答等。

"慢性"给药,指与快速给药模式相反,将试剂以连续给药模式施用,以便在长时间内保持初始治疗效果(活性)。“间断”给药是指治疗不是没有间歇地连续进行,而是以周期性为特征。

意欲治疗的"哺乳动物"是指归为哺乳动物的任何动物,包括人、家畜和农场动物,以及动物园里的动物、参与运动项目的动物或宠物如狗、猫、牛、绵羊、猪、山羊、兔、马等。优选所述哺乳动物为人类。

与一种或多种其它治疗剂"联合"给药,包括同时(并行)和以任何顺序联贯给药。

本文所用"载体"包括在所用剂量和浓度下对细胞或哺乳动物无毒性的可药用载体、赋形剂、稳定剂。可药用载体常常是含水的 pH 缓冲的溶液。可药用载体的实例包括缓冲剂如磷酸盐、柠檬酸盐和其它有机酸;抗氧化剂

包括抗坏血酸；低分子量(小于 10 个残基)多肽；蛋白，如血清白蛋白、明胶或免疫球蛋白；亲水聚合物如聚乙烯吡咯烷酮；氨基酸如甘氨酸、谷氨酰胺、天冬酰胺、精氨酸或赖氨酸；单糖，二糖和其它碳水化合物，包括葡萄糖、甘露糖、或糊精；螯合剂如 EDTA；糖醇如甘露醇或山梨醇；成盐离子如钠；和/或非离子表面活性剂如 TWEENTM、聚乙二醇(PEG)和 PLURONICSTM。

本文中所用"标记"一词，是指直接或间接地与抗体偶联从而产生"标记"抗体的可检测的化合物或组分。标记可以是其本身可被测得(例如放射性同位素标记或荧光标记)，或者如果是酶标记时，可催化底物化合物或组分发生化学变化，这种变化是可检测的。

"固相"是指本发明抗体可粘附的非水性基质。本文中固相的实例包括部分或全部由玻璃(如控制孔径的玻璃)、多糖(如琼脂糖)、聚丙烯酰胺、聚苯乙烯、聚乙烯醇和硅氧烷制成的那些固相。在某些实例中，根据上下文的内容，固相可包括测试板的孔；在其它的例子中，它可以是纯化柱(如亲和层析柱)。该术语也包括美国专利 4275149 中公开的分散颗粒的不连续固相。

"脂质体"是由能向哺乳动物有效运送药物(如本文公开的 PRO10282 多肽和/或其抗体)的各类脂质、磷脂和/或表面活性剂组成的小分子囊泡。脂质体的组分通常排列为双层形式，与生物膜的脂质排列相似。

本文中"小分子"定义为分子量低于约 500 道尔顿的分子。

短语"基因扩增"和"基因复制"可交替使用，它们是指使多拷贝基因或基因片段在具体细胞或细胞系内形成的过程。重复的区域(扩增 DNA 的延伸)一般常指"扩增子"。通常，信使 RNA (mRNA)的产生量，即基因表达水平，也与表达的具体基因所产生的拷贝数量成比例增加。

本文所用术语"细胞毒性剂"，是指抑制或阻止细胞功能和/或引起细胞破坏的物质。该术语意在包括放射性同位素(例如 I¹³¹、I¹²⁵、Y⁹⁰ 和 Re¹⁸⁶)，化疗剂，毒素如细菌、真菌、植物或动物来源的酶活性毒素或其片段。

"化疗剂"是治疗癌有效的化合物。化疗剂的实例包括：阿霉素，阿霉素(doxorubicin)，表阿霉素(epirubicin)，5-氟尿嘧啶，胞嘧啶阿拉伯糖苷(arabinoside)("Ara-C")，环磷酰胺，噻替派，白消安，细胞毒素(cytosine)，紫杉烷例如紫杉醇(Taxol, Bristol-Myers Squibb Oncology, Princeton, NJ)和泰索帝(doxetaxel)(泰索帝(Taxotere), Rhne Poulenc Rorer, Antony, France), taxotere, 甲氨蝶呤，顺铂，苯丙氨酸氮芥，长春花碱，博来霉素，依托泊苷，

异环磷酰胺, 丝裂霉素 C, 米托蒽醌, 长春新碱, 长春瑞宾, 卡铂, 替尼泊苷, 道诺霉素, 洋红霉素, 氨基蝶呤, 放线菌素 D, 丝裂霉素, esperamicins (见美国专利 4, 675, 187), 5-FU, 6-硫代鸟嘌呤, 6-巯基嘌呤, 放线菌素 D, VP-16, 苯丁酸氮芥, 苯丙氨酸氮芥及其它有关的氮芥子。对肿瘤起调节或抑制激素作用的激素制剂如他莫昔芬和奥那司酮也包括在此定义中。

本文所用"生长抑制剂", 是指抑制细胞生长的化合物或组合物, 特别是抑制过度表达本文鉴定的任一种基因的癌细胞, 无论是在体内还是在体外。因此, 生长抑制剂是显著降低过度表达此基因的 S 期细胞百分比的化合物或组合物。生长抑制剂的实例包括阻断细胞周期进程的试剂 (在 S 期之外的环节), 譬如, 诱导 G1 期和 M-期停滞的试剂。传统的 M-期阻断剂包括长春花碱(长春新碱和长春碱)、紫杉酚和拓扑酶 II 抑制剂如阿霉素、表阿霉素、柔红霉素、依托泊苷和博莱霉素。那些使 G1 期停滞的试剂也使 S-期停滞, 例如, DNA 烷化剂如他莫昔芬、泼尼松、达卡巴嗪、氮芥、顺铂、氨基蝶呤、5-氟尿嘧啶和阿糖胞嘧啶-C。进一步的信息可在下列文献中找到: The Molecular Basis of Cancer, Mendelsohn 和 Israel, eds., Chapter 1, entitled "Cell cycle regulation, oncogens, 和 antineoplastic drugs" by Murakami 等 (WB Saunders: Philadelphia, 1995), 特别是第 13 页。

"阿霉素"是蒽环霉素类抗生素。阿霉素的化学全名为(8S-顺式)-10-[(3-氨基-2, 3, 6-三去氧- α -L-来苏-吡喃糖苷)]-7, 8, 9, 10-四氢-6, 8, 11-三羟-8-(羟乙酰)-甲氧基-5, 12-naphthacenedion。

术语"细胞因子"是一般性术语, 指由一个细胞群释放的作用于另一个细胞群起细胞间调节子作用的蛋白。这些细胞因子的实例是淋巴因子、单核因子和传统的多肽类激素。细胞因子包括生长激素, 如人生长激素, N-甲二磺酰人生长激素和牛生长激素; 甲状旁腺素; 甲状腺素; 胰岛素; 前胰岛素; 松弛素; 前松弛素; 糖蛋白激素如卵泡刺激素(FSH), 甲状腺刺激素(TSH), 促黄体(生成)激素(LH); 肝细胞生长因子; 成纤维细胞生长因子; 催乳素; 胎盘催乳素; 肿瘤坏死因子- α 和 β ; 缪氏抑制物质(mullerian-inhibiting substance); 小鼠促性腺激素相关肽; 抑制素; 活化素; 血管内皮细胞生长因子; 整联蛋白; 血小板生成素(TPO); 神经生长因子如 NGF- β ; 血小板生长因子; 转化生长因子(TGF)如 TGF- α 和 TGF- β ; 胰岛素样生长因子-I 和-II; 促红细胞生成素(EPO); 骨诱导因子(osteoinductive factors); 干扰素如干扰素

- α , - β , - γ ; 集落刺激因子(CSF)如巨噬细胞-CSF(M-CSF); 粒细胞-巨噬细胞-CSF(GM-CSF); 粒细胞-CSF(G-CSF); 白细胞介素(IL)如 IL-1, IL-1 α , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12; 肿瘤坏死因子如 TNF- α 或 TNF- β ; 和其它多肽因子包括 LIF 和 kit 配体(KL)。本文中术语细胞因子包括天然蛋白或来自重组细胞培养物的蛋白以及天然序列细胞因子的生物活性等效物。

本申请所用术语"前体药物", 是指药物活性物质的前体或衍生物形式, 其相对于亲本药物对肿瘤细胞的细胞毒作用较小且可酶促活化或被转换成更具活性的亲本形式。例见 Wilman, "Prodrugs in Cancer Chemotherapy", Biochemical Society Transactions, 14 : 375-382, 615th Meeting, Belfast (1986), 和 Stella 等, "Prodrugs : A Chemical Approach to Targeted Drug Delivery", Directed Drug Delivery Borchardt 等, (ed.), pp. 147-267, Humana press(1985)。本发明的前体药物包括但不限于: 含有磷酸盐的前体药物, 含有硫代磷酸盐的前体药物, 含有硫酸盐的前体药物, 含有肽的前体药物, D-氨基酸修饰的前体药物, 糖基化的前体药物, 含有 β -内酰胺的前体药物, 任选含有取代的苯氧乙酰胺的前体药物或任选含有取代的苯基乙酰胺的前体药物, 5-氟胞嘧啶和可转化为更具细胞毒活性的游离药物的其它 5-氟尿嘧啶前体药物。可衍生为本发明所用前体药物形式的细胞毒药物, 包括但不限于上述化疗药物。

本文公开的多肽、其激动剂和/或拮抗剂对于抑制赘生细胞生长、肿瘤生长或癌细胞生长的"有效量", 是指能够将靶细胞生长抑制到一定程度的量。该术语包括可引发对靶细胞的生长抑制、细胞静息和/或细胞毒性作用和/或编程性细胞死亡的量。

而 PRO10282(包括 PRO19578) 多肽或其拮抗剂, 用于抑制赘生细胞生长、肿瘤细胞或癌细胞生长目的时的"有效量", 可以完全根据经验和以常规方式予以确定。

有关治疗肿瘤的"治疗有效量", 是指能够达到下列一种或多种效果的量: (1)将肿瘤生长抑制到一定程度, 包括减缓生长和完全阻止生长; (2)减少肿瘤细胞数量; (3)缩小肿瘤尺寸; (4)抑制(即减少、减缓或完全阻止)肿瘤细胞浸润周边器官; (5)抑制(即减少、减缓或完全阻止)转移; (6)提高抗-肿瘤免疫应答, 这可能但并非必须导致肿瘤消退或者排斥(rejection); 和/

或 (7) 与病症有关的一种或多种症状缓解到一定程度。用于治疗肿瘤为目的的 PRO10282 多肽拮抗剂的"治疗有效量", 可根据经验和按常规方式予以确定。

PRO10282 拮抗剂的"生长抑制量", 是能够抑制细胞尤其是肿瘤细胞(例如癌细胞)在体内或体外生长的量。用于抑制赘生细胞生长目的的 PRO10282 拮抗剂的"细胞毒性量", 可根据经验并以常规方式予以确定。

PRO10282 拮抗剂的"细胞毒性量", 是能够引起体内或体外细胞尤其是肿瘤细胞(例如癌细胞)破坏的量。用于抑制赘生细胞生长的 PRO10282 拮抗剂的"细胞毒性量", 可根据经验和以常规方式予以确定。

"反义寡脱氧核苷酸"或"反义寡核苷酸" (术语之间可互用), 定义为可以按照序列-特异性方式抑制靶基因进行转录和/或翻译的核酸分子。术语"反义"是指核酸与靶基因的编码("正义")遗传序列互补的事实。反义寡核苷酸与新生的 mRNA 通过 Watson-Crick 碱基配对而反向平行杂交。通过与靶 mRNA 模板结合, 反义寡核苷酸阻断 mRNA 模板所编码蛋白质的成功转录。术语特别包括称为"核酶(ribozymes)"的反义因子, 已经证明所述核酶通过加入具有天然自-剪切活性的序列而诱导靶 RNA 催化裂解 (Warzocha 和 Wotowiec, "Antisense strategy" biological utility and prospects in the treatment of hematological malignancies."Leuk. Lymphom 24 : 267-281 [1997])。

II. 本发明组合物和方法

A. 全长天然序列人 PRO10282 和 PRO19578 (Stra6) 多肽

本发明提供编码本申请分别称为天然序列人 PRO10282 (或者也称为 UNQ3126) 和 PRO19578 多肽的新鉴定和分离的核苷酸序列。更具体而言, 如下面的实施例进一步详细公开的那样, 已经鉴定和分离出编码天然人 PRO10282 和 PRO19578 多肽的 cDNA。应注意到在不同表达循环中产生的蛋白可能具有不同的 PRO 数, 但是, 任何特定 DNA 和其编码的蛋白的 UNQ 数是独一无二的, 而且将不会变化。然而, 为了简明起见, 在本说明书中将 DNA148380-2827 编码的蛋白以及前述定义的 PRO10282 的所有其它天然同系物和变体均称为"PRO10282", 而不论其来源或制备模式如何。而"PRO10282"的天然变体 - 由 DNA148389-2827-1 编码的多肽, 则特别称之为"PRO19578"。

如下面的实施例所公开的那样, 本文的 DNA148380-2827 和

DNA148389-2827-1 cDNA 克隆已经在 ATCC 保藏。该克隆的实际核苷酸序列,可由本领域熟练技术人员使用本领域常规方法通过对保藏克隆进行序列分析而很容易地测得。使用本领域常规技术可以预测该核苷酸序列编码的氨基酸序列。关于本文所述天然人 PRO10282 和 PRO19578 多肽以及本文所述其编码核酸,申请人已经确定此时在阅读框上什么是与现有序列信息最一致的。

使用上述 ALIGN-2 序列对比计算机程序,业已发现:全长天然 PRO10282 序列(见图 2 和 SEQ ID NO: 2)与 Stra6(应答视黄酸中诱导的小鼠蛋白(AFO62476))具有一定的氨基酸序列同一性。因此,目前据信本申请公开的 PRO10282 多肽是 Stra6 蛋白家族的新鉴定的成员,而且可能具有该蛋白家族通常具有的一种或多种生物学和/或免疫学活性或者特性。目前,据信 PRO19578 是天然全长人 PRO10282 的其它剪切变体,它在天然人 PRO10282 氨基酸序列的 89-97 位缺失 9 个氨基酸(SPVDFLAGD),并且在 518 位处由 Ile (I))替代了 Met (M),所述 518 位与 PRO10282 的 527 位相对应。这表明 Stra6 可以是多形态的。

B. PRO10282 (Stra6)变体

本文公开了两种具体天然序列人 Stra6 多肽(天然序列人 PRO10282 和 PRO19578)。如上所述,据信 PRO19578 是天然序列人 PRO10282 的其它剪切型天然变体,因而是一种"PRO10282 变体"。除了本文描述的全长天然序列 PRO10282 和 PRO19578 多肽外,也考虑了制备 PRO10282 和 PRO19578 变体。通过在 PRO10282 或 PRO19578 DNA 中引入适当的核苷酸改变,和/或通过合成所需 PRO10282 变体多肽,可以制备 PRO10282 和 PRO19578 变体。本领域技术人员懂得,氨基酸变化将会改变 PRO10282 或 PRO19578 的翻译后加工,例如改变糖基化位点的数量和位置或者改变膜锚着特性。

使用例如在美国专利 5364934 号中阐述的保守和非-保守突变的任一技术和指导路线,可以在天然全长序列 PRO10282、PRO19578 或者在本文所述 PRO10282 或 PRO19578 的各种结构域中制造变化。变化可以是编码 PRO10282 或 PRO19578 的一个或多个密码子的取代、缺失或插入,其导致相对于天然序列 PRO10282 或 PRO19578 的 PRO10282 或 PRO19578 氨基酸序列变化。任选地,变化是 PRO10282 或 PRO19578 的一个或多个结构域中的至少一个氨基酸被任何其它氨基酸取代。将 PRO10282 或 PRO19578 序列

与已知同源蛋白分子进行比较，并使高度同源区氨基酸序列数量变化最小，可以建立决定哪个氨基酸残基可以被插入、取代或缺失而对所需活性无负面影响的指导原则。氨基酸取代可以是一个氨基酸被另一个具有类似结构和/或类似化学特性的氨基酸代替的结果，如丝氨酸置换亮氨酸，即保守氨基酸置换。插入或缺失可任选地在约 1~5 个氨基酸的范围内。通过系统地在序列中进行氨基酸插入、缺失或取代并检测所得变体显示全长或成熟天然序列的活性的情况，可以确定能允许的变化。

本文提供 PRO10282 多肽片段。该类片段可以是，例如，与全长天然蛋白相比，在 N-末端或 C-末端截短的或者是缺乏内部残基的片段。某些片段缺乏那些对于 PRO10282 或 PRO19578 多肽的预期生物活性非必要的氨基酸残基。

PRO10282 或 PRO19578 片段可以按照大量常规技术中的任何一个来制备。可以化学合成所需肽片段。另一方法涉及通过酶消化产生 PRO10282 或 PRO19578 片段，例如，用已知在特定氨基酸残基所限定的位点裂解蛋白的酶处理蛋白，或用适宜的限制性内切酶消化 DNA 并分离所需片段。另一适宜的技术包括分离并用聚合酶链反应(PCR)扩增编码所需多肽片段的 DNA 片段。将限定 DNA 片段所需末端的寡核苷酸用作 PCR 的 5'和 3'端引物。优选地，PRO10282 或 PRO19578 多肽片段具有图 2 (SEQ ID NO: 2)所示天然 PRO10282 或 PRO19578 多肽的至少一种生物学和/或免疫学活性。

在特定实施方案中，目标保守取代冠以优选取代的名称示于下表 1 中。如果这类取代导致生物活性变化，则可引入起更多实质性改变并筛选产物，所述改变在表 1 中命名为示范性取代，或者如下文有关氨基酸种类的详述。

表 1

原始残基	示范性取代	优选取代
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn

Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

对 PRO10282 多肽的功能和免疫学特性的实质性修改可通过选择性取代来完成, 所述取代的效应在维持(a)取代区多肽骨架的结构, 例如片层结构或螺旋构象, (b)该分子靶位点的电荷或疏水性, (c)侧链的大小这几方面有显著差异。天然残基根据共有的侧链特性可分为:

- (1) 疏水性: 正亮氨酸, met, ala, val, leu, ile;
- (2) 中性亲水性: cys, ser, thr;
- (3) 酸性: asp, glu;
- (4) 碱性: asn, gln, his, lys, arg;
- (5) 影响链取向的残基: gly, pro; 和
- (6) 芳香族: trp, tyr, phe。

非保守取代将限定上述某一类的成员被另一类取代。也可将这种取代的残基引入到保守取代位点, 或者更优选地引入其余(非-保守)位点。

可以使用本领域已知方法如寡核苷酸-介导的(定点)诱变、丙氨酸扫描和PCR 诱变技术来产生变异。可以对克隆的 DNA 进行定点诱变[Carter 等, Nucl.

Acids Res., 13: 4331 (1986); Zoller 等, Nucl. Acids Res., 10: 6487 (1987)]、盒式诱变[Wells 等, Gene, 34: 315 (1985)]、限制选择诱变[Wells 等, Philos. Trans. R. Soc. London SerA, 317: 415 (1986)]或其它已知技术以产生 PRO10282 变体 DNA。

扫描氨基酸分析也可沿邻接序列确定一个或多个氨基酸。优选的扫描氨基酸是相对小的中性氨基酸。此类氨基酸包括丙氨酸, 甘氨酸、丝氨酸和半胱氨酸。通常, 丙氨酸是此组中优选的扫描氨基酸, 因为它的 β -碳上无侧链并且不太可能改变变体的主链构象[Cunningham 和 Wells, Science, 244: 1081-1085 (1989)]。优选丙氨酸的另一原因是因为它是最常用氨基酸。而且, 它常常既出现在掩蔽位置也出现在暴露位置[Creighton, The Proteins, (W. H. Freeman & Co., N. Y.); Chothia, J. Mol. Biol., 150: 1 (1976)]。如果丙氨酸取代不能产生足够量的变体, 则可使用同配氨基酸。

C. 天然 PRO10282 (Stra6)和变体 PRO10282 (Stra6)的修饰

本发明包括 PRO1028 和 PRO10282 变体(包括 PRO19578)的共价修饰。一种共价修饰包括使 PRO10282 多肽的靶氨基酸残基和能够与 PRO10282 的选定侧链或 N-或 C-末端残基发生反应的有机衍生化剂进行反应。具有双官能剂的衍生化是非常有用的, 例如, PRO10282 与纯化抗-PRO10282 抗体方法中所用的非水溶性载体基质或表面交联, 反之亦然。通常使用的交联剂包括例如, 1, 1-双(二偶氮乙酰基)-2-苯乙烷, 戊二醛, N-羧琥珀酰亚胺酯如与 4-叠氮基水杨酸的酯, 高双官能亚胺基酯, 包括二琥珀酰亚胺酯如 3, 3'-连二硫酸双(琥珀酰亚胺基丙酸酯)、双官能马来酰亚胺酯如双-N-马来酰亚胺-1, 8-辛酯和化学试剂如甲基-3-[(对-叠氮苯基)二硫]丙炔亚胺酯。

其它修饰包括谷氨酰胺和天冬酰胺残基分别成为相应的谷氨酰胺和天冬酰胺残基的脱酰胺, 脯氨酸和赖氨酸的羟基化, 丝氨酸或苏氨酸残基的羟基磷酸化, 赖氨酸、精氨酸和组氨酸侧链的 α -氨基甲基化[T. E. Creighton, Proteins: Structure and Molecular Properties, W. H. Freeman & Co., San Francisco, pp. 79-86 (1983)], 胺的 N-末端乙酰化, 和任何 C-末端羧基的酰胺化。

本发明还包括 PRO10282 多肽的另一类共价修饰, 所述共价修饰包括改变多肽的天然糖基化模式。天然序列 PRO10282 和 PRO19578 多肽含有位于其氨基酸序列 8-12 位的 N-连接的糖基化位点。"改变天然糖基化模式", 在这里的目的是为了删除天然序列 PRO10282 或 PRO19578 中一个或多个碳水

化合物部分(通过除去发生糖基化的位点或者通过化学和/或酶手段而删除糖基化),和/或在天然序列 PRO10282 或 PRO19578 中加入一个或多个本不存在的糖基化位点。此外,该短语包括天然蛋白糖基化的性质变化,包括不同碳水化合物部分在性质和比例上的变化。

通过改变氨基酸序列可以实现在天然或变体 PRO10282 多肽中加入糖基化位点。此变更可以例如,通过向天然序列 PRO10282 中加入、或者取代一个或多个丝氨酸或苏氨酸残基(针对 O-连接的糖基化位点)而完成。任选地,可以通过 DNA 水平的变化,特别是通过使编码 PRO10282 多肽的 DNA 中预先选定的碱基发生突变从而产生将会翻译成期望氨基酸的密码子,来改变 PRO10282 氨基酸序列。

在 PRO10282 多肽上增加碳水化合物部分数量的另一方法是,将糖苷与多肽化学性或酶性地偶联。现有技术有此类方法的描述,例如 1987 年 9 月 11 日公布的 WO 87/05330,以及 Aplin 和 Wriston, CRC Crit. Rev. Biochem. 第 259-306 页(1981)。

用化学或酶方法,或者通过突变性取代编码作为糖基化靶点的氨基酸残基的密码子,来实现去除 PRO10282 多肽中的碳水化合物部分。化学性去糖基化技术为本领域所公知,并描述在例如下述文献中: Hakimuddin 等, Arch. Biochem. Biophys., 259: 52 (1987)和 Edge 等, Anal. Biochem., 118:131 (1981)。使用 Thotakura 等在 Meth. Enzymol., 138: 350 (1987)中描述的各种内-和外-糖苷酶,可以完成多肽中碳水化合物部分的酶裂解。

PRO10282 的另一类共价修饰,包括以美国专利 4640835、4496689、4301144、4670417、4791192 或 4179337 中所述方式,将 PRO10282 多肽与各种非蛋白类聚合物之一连接,例如聚乙二醇(PEG)、聚丙烯二醇或聚氧化亚烷基。

本发明 PRO10282 多肽也可被修饰成嵌合分子,包括将 PRO10282 与另一异源多肽或氨基酸序列融合。

在一实施方案中,嵌合分子含有 PRO10282 与标记多肽的融合体,其中的标记多肽具有抗-标记抗体可选择性结合的表位。表位标记一般位于 PRO10282 的氨基-或羧基-末端。使用抗标记多肽的抗体,可以检测到此类表位-标记型 PRO10282 的存在。另外,表位标记的提供,使得利用抗-标记抗体或另一类与表位标记结合的亲和基质来亲和纯化 PRO10282 变得非常容

易。各种标记多肽及其各自抗体是本领域众所周知的。实例包括聚-组氨酸(poly-his)或聚-组氨酸-甘氨酸(poly-his-gly)标记; flu HA 标记多肽及其抗体 12CA5 [Field 等, Mol. Cell. Biol., 8: 2159-2165 (1988)]; c-myc 标记和它的 8F9、3C7、6E10、G4、B7 及 9E10 抗体[Evan 等, Molecular and Cellular Biology, 5: 3610-3616 (1985)]; 和单纯疱疹病毒糖蛋白 D (gD)标记及其抗体[Paborsky 等, Protein Engineering, 3 (6): 547-553 (1990)]。其它标记多肽包括 Flag-肽 [Hopp 等, BioTechnology, 6: 1204-1210 (1988)]; KT3 表位肽[Martin 等, Science, 255: 192-194 (1992)]; α -微管蛋白表位肽[Skinner 等, J. Biol. Chem., 266: 15163-15166 (1991)]; 和 T7 基因 10 蛋白肽标记[Lutz-Freyermuth 等, Proc. Natl. Acad. Sci. USA, 87: 6393-6397 (1990)]。

在另一实施方案中,嵌合分子可包括 PRO10282 与免疫球蛋白或免疫球蛋白特定区域的融合。对于二价形式的嵌合分子(也指"免疫粘附素"),融合可以是与 IgG 分子 Fc 区的融合。Ig 融合优选包括 PRO10282 多肽的可溶形式(跨膜结构域缺失或失活)取代 Ig 分子内至少一个可变区。在一特别优选的实施方案中,免疫球蛋白融合包括 IgG1 分子的铰链区、CH2 和 CH3 区,或铰链区、CH1、CH2 和 CH3 区。免疫球蛋白融合体的制备也参见 1995 年 6 月 27 日发布的美国专利 5428130。

D. 天然或变体 PRO10282 (Stra6)多肽的制备

下面的描述主要涉及通过培养转化或转染了含 PRO10282 核酸的载体的细胞来制备天然序列 PRO10282(Stra6)多肽和变体 PRO10282(Stra6)多肽(特别包括 PRO19578)。因此,当然考虑到本领域熟知的供选择方法可用于制备 PRO10282。例如,通过利用固相合成技术的直接合成肽的方法来生产 PRO10282 序列或其一部分[参见例如, Stewart 等, Solid-Phase Peptide Synthesis, W. H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85: 2149-2154 (1963)]。使用手工或自动操作技术,进行蛋白体外合成。例如,按照制造商的指示,使用实用生物系统肽合成仪(Foster City, CA),可以完成自动合成。可以分别化学合成 PRO10282 的各部分,然后利用化学或酶学方法制得全长 PRO10282。

1. 编码 PRO10282 (Stra6) DNA 的分离

可以从 cDNA 文库获得编码 PRO10282 的 DNA,所述 cDNA 文库是从据信具有 PRO10282 mRNA 并且以可测得水平表达 PRO10282 的组织制备

的。相应地,如实施例中所描述的那样,从人组织制备的 cDNA 文库中可以很方便地得到人的 PRO10282 DNA。也可以从基因组文库或者通过已知合成方法(例如自动核酸合成)获得 PRO10282-编码基因。

使用为识别目标基因或其编码的蛋白而设计的探针(如 PRO10282 抗体或至少约 20-80 个碱基的寡核苷酸),可以筛选文库。使用标准操作,如 Sambrook 等在 *Molecular Cloning: A Laboratory Manual*(New York: Cold Spring Harbor Laboratory, 1989 出版)中所描述的方法,用选择的探针进行 cDNA 或基因组文库的筛选。分离 PRO10282 编码基因的另一个可选用方法是使用 PCR 方法[Sambrook 等, 出处同上; Dieffenbach 等, *PCR Primer: A Laboratory Manual*(Cold Spring Harbor Laboratory, 1995 年出版)]。

下述实施例描述 cDNA 文库筛选技术。选作探针的寡核苷酸序列应当具有足够的长度并且足够清晰,以便使假阳性出现机率最小。寡核苷酸优选是标记的,这样通过与正在筛选的文库中的 DNA 杂交而可被检测到。标记方法为本领域所公知,包括使用放射性标记物如 ^{32}P -标记的 ATP、生物素或酶标记。包括中等严格和高度严格的杂交条件,见 Sambrook 等, 出处同上。

可将在筛选文库方法中鉴定的序列与其它已知的保藏序列或公共数据库如 GenBank 或其它私有序列数据库中的序列进行比较。使用本领域已知的和本文描述的方法,可以测定分子中限定区域或全长序列的序列同一性(无论在氨基酸水平还是在核苷酸水平)。

使用本文首次公开的推定氨基酸序列,和如果必要,使用 Sambrook 等(出处同上)描述的常规引物延伸程序,以检测没有被逆转录为 cDNA 的 mRNA 的前体和加工中间体,可以从选定的 cDNA 或基因组文库获得具有蛋白编码序列的核酸。

2. 宿主细胞的选择和转化

用生产本文所述 PRO10282 的表达或克隆载体转染或转化宿主细胞,并在常规营养培养基中培养宿主细胞,对所述培养基进行改良使之适宜于诱导启动子、选择性转化体或扩增编码所需序列的基因。无需繁琐的试验,本领域熟练技术人员就可以选择培养条件如培养基、温度、pH 等。通常,使细胞培养物产量最大化的原则、方案和实用技术见于 *Mammalian Cell Biotechnology: a Practical Approach*, M. Butler 编著(IRL 出版, 1991)和 Sambrook 等, 出处同上。

转染真核细胞和转化原核细胞的方法为本领域熟练技术人员所已知,例如, CaCl_2 、 CaPO_4 、脂质体-介导的方法和电穿孔法。根据所用的宿主细胞,使用适合于该宿主细胞的标准技术进行转化。如 Sambrook 等出处同上所述的使用氯化钙的钙处理,或电穿孔法一般用于原核生物。如 Shaw 等在 *Gene*, 23: 315(1983)和 1989 年 6 月 29 日公布的 WO 89/05859 中描述的那样,根癌农杆菌的感染用于转化某些植物细胞。对于没有细胞壁的哺乳动物来说,可以使用 Graham 和 van der Eb 在 *Virology*, 52: 456-457(1978)中所述的磷酸钙沉淀法。哺乳动物细胞宿主系统转化的总的情况在美国专利 4399216 中已有描述。根据 Van Solingen 等, *J. Bact.*, 130: 946(1977)和 Hsiao 等, *Proc. Natl. Acad. Sci.(USA)*, 76: 3829(1979)所述方法,通常可以完成向酵母的转化。然而,也可使用将 DNA 引入细胞的其它方法,例如通过核的显微注射、电穿孔、细菌原生质体与完整细胞的融合,或者聚阳离子如聚凝胺(polybrene)、聚鸟氨酸。转化哺乳动物细胞的各种方法见 Keown 等, *Methods in Enzymology*, 185: 527-537(1990)和 Mansour 等, *Nature*, 336: 348-352(1988)。

克隆或表达本文所述载体中 DNA 的适宜的宿主细胞,包括原核生物、酵母或高等真核细胞。适宜的原核生物包括但不限于革兰氏阴性或革兰氏阳性细菌等真细菌,譬如肠杆菌科(*Enterobacteriaceae*)如大肠杆菌。各种大肠杆菌株均是公众可以得到的,如大肠杆菌 K12 株 MM294(ATCC 31446); 大肠杆菌 X1776(ATCC 31537); 大肠杆菌株 W3110(ATCC 27325)和 K5 772(ATCC 53635)。其它适宜的原核宿主细胞包括肠杆菌科如埃希氏杆菌属,例如,大肠杆菌,肠杆菌属,欧文菌属,克雷白菌属,变形菌杆属,沙门菌属(如鼠伤寒沙门菌),沙雷菌属(如粘质沙雷菌)和志贺菌属等,以及芽孢杆菌属如枯草芽孢杆菌和地衣芽孢杆菌(例如 1989 年 4 月 12 日出版的 DD 266710 中所述地衣芽孢杆菌 41P)等,假单胞菌属的铜绿菌假单胞菌,及链霉菌。这些实例是用于说明而并非限制于此。W3110 株是一个特别优选的宿主或母宿主,因为它是重组 DNA 产物发酵的常用宿主株。优选地,宿主细胞分泌少量蛋白水解酶。例如,修饰 W3110 株以使编码宿主内源性蛋白的基因发生遗传突变,此类宿主的实例包括大肠杆菌 W3110 株 1 A2,该株具有完整基因型 *tonA*; 大肠杆菌 W3110 株 9E4,该株具有完整基因型 *tonA ptr3*; 大肠杆菌 W3110 株 27C7(ATCC 55244),该株具有完整基因型 *tonA ptr3 phoA E15(argF-lac)169 degP ompT kan'*; 大肠杆菌 W3110 株 37D6,该株具有完整

基因型 *tonA ptr3 phoA E15(argF-lac)169 degP ompT rbs7 ilvG kan^r*; 大肠杆菌 W3110 株 40B4, 它是 37D6 发生非-卡那霉素耐药 *degP* 缺失突变的株; 和具有 1990 年 8 月 7 日发布的美国专利 4946783 中披露的突变型周质蛋白酶的大肠杆菌株。或者, 克隆的体外方法, 例如 PCR 或其它核酸聚合酶反应, 是适宜的。

除了原核生物, 真核微生物如丝状真菌或酵母也是适于使编码 PRO10282 的载体克隆或表达的宿主。酿酒酵母是常用的低等真核宿主微生物。其它真核微生物包括: 粟酒裂殖酵母(*Schizosaccharomyces pombe*)(Beach 和 Nurse, Nature, 290: 140 [1981]; 1985 年 5 月 2 日公布的 EP 139383); 克鲁维酵母属(*Kluyveromyces*)宿主(美国专利 4943529; Fleer 等, Bio/Technology, 9: 968-975(1991)), 例如乳酸克鲁维酵母(*K.lactis*)(MW98-8C, CBS683, CBS4574; Louvencourt 等, J. Bacteriol., 154(2): 737-742 [1983])、脆壁克鲁维酵母(*K.fragilis*)(ATCC 12424)、保加利亚克鲁维酵母(*K.bulgaricus*)(ATCC 16045)、威克曼氏克鲁维酵母(*K. wickerhamii*)(ATCC 24178)、*K.waltii*(ATCC 56500)、果蝇克鲁维酵母(*K. drosophilum*)(ATCC 36906; Van den Berg 等, Bio/Technology, 8:135(1990))、耐热克鲁维酵母(*K. thermotolerans*)和克鲁维氏酵母属(*K. marxianus*)等; *yarrowia*(EP 402226); 巴斯德毕赤酵母(*pichia pastoris*)(EP 183070; Sreekrishna 等, J. Basic Microbiol., 28: 265-278[1988]); 念珠菌属; *Trichoderma reesia*(EP 244234); 粗糙链孢霉(Case 等, Proc. Natl. Acad. Sci. USA, 76: 5259-5263 [1979]); 许旺氏酵母属(*schwanniomyces*)如西方许旺氏酵母(*schwanniomyces occidentalis*)(1990 年 10 月 31 日公布的 EP 394538)等; 和丝状真菌, 例如链孢霉属、青霉属、*Tolypocladium*(1991 年 1 月 10 日公布的 WO 91/00357)以及曲霉属宿主如构巢曲霉(Ballance 等, Biochem. Biophys. Res. Commun., 112: 284-289 [1983]; Tilburn 等, Gene, 26: 205-221 [1983]; Yelton 等, Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984])和黑曲霉等(Kelly 和 Hynes, EMBO J., 4: 475-479 [1985])。在本文中嗜甲基(Methylotropic)酵母是适宜的, 包括但不限于选自下述属的能够在甲醇中生长的酵母: 汉逊氏酵母属(*Hansenula*), 念珠菌属(*Candida*), 克勒克酵母属(*Kloeckera*), 毕赤酵母属(*Pichia*), 酵母属, 球拟酵母属和红酵母属。此类酵母的例证性特定酵母属的清单见 C. Anthony, The Biochemistry of Methylotrophs, 269(1982)。

用于表达糖基化 PRO10282 的适合宿主细胞来自多细胞生物。无脊椎动物细胞的实例包括昆虫细胞(例如果蝇属 S2 和草地夜蛾 Sf9)和植物细胞。有用的哺乳动物宿主细胞系包括中国仓鼠卵巢(CHO)细胞和 COS 细胞。更具体的实例包括 SV40(COS-7, ATCC CRL 1651)转化的猴肾 CV1 细胞系; 人胚肾细胞系(293 细胞或亚克隆培养成在悬浮培养液中生长的 293 细胞, Graham 等, J. Gen Virol. 36: 59(1977)); 中国仓鼠卵巢细胞/-DHFR(CHO, Urlaub 和 Chasin, Proc. Natl. Acad. Sci. USA, 77:4216(1980)); 鼠赛尔托利细胞(TM4, Mather, Biol. Reprod., 23: 243-251(1980)); 人肺细胞(W138, ATCC CCL 75); 人肝脏细胞(Hep G2, HB 8065); 和鼠乳腺癌细胞(MMT 060562, ATCC CCL51)。选择合适宿主细胞是本领域常识。

3. 复制型载体的选择和应用

可将编码 PRO10282 的核酸(例如 cDNA 或基因组 DNA)插入复制型载体中以进行克隆(DNA 扩增)或表达。各种载体是公众可以获得的。载体可以是例如质粒、粘粒、病毒颗粒或噬菌体的形式。使用各种方法可将合适的核酸序列插入载体。通常, 利用本领域已知技术将 DNA 插入到合适的限制性核酸内切酶位点。载体组成部分一般包括但不限于: 一个或多个信号序列、复制起点、一个或多个标记基因、增强子元件、启动子和转录终止序列。使用本领域熟练技术人员熟知的标准连接技术, 构建包括这些组成部分中一个或多个元件的适宜载体。

PRO10282 不仅可直接重组制备, 而且还可制备成与异源多肽的融合多肽, 所述异源多肽为信号序列或在成熟蛋白或多肽的 N 末端上具有特异性裂解位点的其它多肽。通常, 信号序列是载体的一个组分或者是被插入载体的 PRO10282-编码 DNA 的一部分。信号序列可以是选自例如碱性磷酸酶、青霉素酶、lpp 或热稳定肠毒素 II 前导区的原核信号序列。对于酵母分泌, 可以是例如酵母转化酶前导区、 α 因子前导区(包括糖酵母属和克鲁维酵母属的 α 因子前导区, 后者在美国专利 5010182 中有述), 或酸性磷酸酶前导区、白色念珠菌葡萄糖淀粉酶前导区(1990 年 4 月 4 日公布的 EP 362179), 或者 1990 年 11 月 15 日公布的 WO90/13646 中所述信号序列。在哺乳动物细胞表达时, 可使用哺乳动物信号序列以直接分泌蛋白, 如得自相同物种或相关物种分泌型多肽的信号序列以及病毒分泌前导区。

表达载体和克隆载体均含有使载体在一个或多个选定宿主细胞中复制

的核酸序列。在多种细菌、酵母和病毒中，这样的序列众所周知。来自质粒 pBR322 的复制起点适宜于大多数革兰氏阴性细菌，2 μ 质粒复制起始点适于酵母，各种病毒复制起始点(SV40，多瘤病毒，腺病毒，VSV 或 BPV)适宜于哺乳动物细胞中的克隆载体。

表达载体和克隆载体通常包含筛选基因，也称为可筛选标记。典型的筛选基因编码蛋白，该蛋白(a)提供对抗生素或其它毒素，如氨苄青霉素、新霉素、氨基蝶呤或四环素等的抗性，(b)弥补营养缺陷，或(c)提供从复合培养基中不能得到的关键营养物质，例如编码芽孢杆菌属 D-丙氨酸消旋酶的基因。

适于哺乳动物细胞的筛选标记实例，是那些使能接纳编码 PRO10282 核酸的细胞得到鉴定的标记，例如 DHFR 或胸苷激酶。当使用野生型 DHFR 时，适当的宿主细胞是按照 Urlaub 等在 Proc. Natl. Acad. Sci. USA 77:4216(1980)中描述的方法制备和扩增的 DHFR 活性缺陷型中国仓鼠卵巢(CHO)细胞系。适用于酵母的合适筛选基因是存在于酵母质粒 YRp7 中的 trp1 基因[Stinchcomb 等, Nature, 282: 39(1979); Kingsman 等, Gene, 7:141(1979); Tschemper 等, Gene, 10:157(1980)]。trp1 基因为不能在色氨酸中生长的酵母突变株(例如 ATCC 44076 或 PEP4-1)提供了筛选标记[Jones, Genetics, 85: 12(1977)]。

表达和克隆载体通常含有可操作地连接于 PRO10282-编码核酸序列的启动子，以指导 mRNA 合成。被各种潜在宿主细胞识别的启动子是众所周知的。适用于原核宿主的启动子，包括 β -内酰胺酶和乳糖启动子系统[Chang 等, Nature, 275: 615(1978); Goeddel 等, Nature, 281: 544(1979)]，碱性磷酸酶，色氨酸(trp)启动子系统[Goeddel, Nucleic acids Res., 8: 4057(1980); EP 36776]，和杂合启动子如 tac 启动子[deBoer 等, Proc. Natl. Acad. Sci. USA, 80: 21-25(1983)]。适用于细菌系统的启动子，也含有可操作地连接于编码 PRO10282 的 DNA 的 Shine-Dalgarno(S. D.)序列。

适用于酵母宿主的启动序列的实例，包括 3-磷酸甘油酸激酶[Hitzeman 等, J. Biol. Chem., 255: 2073(1980)]或其它糖酵解酶 [Hess 等, J. Adv. Enzyme Reg., 7: 149(1968); Holl and, Biochemistry, 17: 4900(1978)]的启动子，所述其它糖酵解酶如烯醇化酶，甘油醛-3-磷酸脱氢酶，己糖激酶，丙酮酸脱羧酶，磷酸果糖激酶，葡萄糖-6-磷酸异构酶，3-磷酸甘油变位酶，丙酮酸激酶，磷酸丙糖异构酶，磷酸葡萄糖异构酶和葡萄糖激酶。

其它的酵母启动子,即那些具有由生长条件控制转录的优点的诱导型启动子,是下述基因的启动子区,即乙醇脱氢酶 2、异细胞色素 C、酸性磷酸酶、与氮代谢相关的降解酶、金属硫蛋白、甘油醛-3-磷酸脱氢酶和负责麦芽糖和半乳糖利用的酶。在 EP 73657 中进一步描述了用于酵母表达系统的适当载体和启动子。

在哺乳动物宿主细胞中,由载体转录 PRO10282 可受下述启动子调控,所述启动子来自病毒基因组如多瘤病毒、鸡痘病毒(1989 年 7 月 5 日公布的 UK 2211504)、腺病毒(如腺病毒 2)、牛乳头瘤病毒、禽肉瘤病毒、巨细胞病毒、逆转录病毒、乙型肝炎病毒和猴病毒 40(SV40)的启动子,或者来自异源哺乳动物的启动子,如肌动蛋白启动子或免疫球蛋白启动子等,和来自热休克启动子,前提是这些启动子与宿主细胞系统相容。

在载体中插入增强子序列,可以增加编码 PRO10282 的 DNA 在高等真核生物中的转录。增强子是 DNA 的顺式作用元件,作用于启动子以增加所述 DNA 的转录,一般约 10~300 bp。目前已知很多哺乳动物基因(珠蛋白,弹性蛋白酶、白蛋白、 α 胎蛋白和胰岛素)的增强子序列。然而,人们通常使用真核细胞病毒的增强子。实例包括在其复制起始点晚期侧(late side)的 SV40 增强子(bp 100-270),巨细胞病毒早期启动子增强子,在其复制起始点晚期侧的多形瘤增强子,和腺病毒增强子。所述增强子可以剪接入 PRO10282 编码序列的 5'或 3'端,但优选位于启动子的 5'端。

用于真核宿主细胞(酵母、真菌、昆虫、植物、动物、人或来自其它多细胞生物的有核细胞)的表达载体,还包括对转录终止和 mRNA 结构稳定所必须的序列。这些序列通常来自真核或病毒 DNA 或 cDNA 的 5'(偶尔为 3')非翻译区。这些区域包含转录为编码 PRO10282 的 mRNA 的非翻译区中聚腺苷酸化片段的核苷酸区段。

在重组脊椎动物细胞培养中适应 PRO10282 合成的其它适宜方法、载体和宿主细胞,见 Gething 等, *Nature*, 293: 620-625(1981); Mantei 等, *Nature*, 281: 40-46(1979); EP 117060; 和 EP 117058 中的描述。

4. 检测基因扩增/表达

使用基于本文提供的序列的合适标记的探针,利用常规技术如 Southern 印迹、测定 mRNA 转录量的 Northern 印迹[Thomas, *Proc. Natl. Acad. Sci. USA*, 77: 5201-5205(1980)]、点印迹(DNA 分析)或原位杂交,可以直接在样

品中测定基因扩增和/或表达。或者,使用能够识别特异双链体的抗体,所述双链体包括 DNA 双链体、RNA 双链体和 DNA-RNA 杂合双链体或者 DNA-蛋白双链体。然后标记该抗体,进行试验,试验中,双链体结合于表面,使得通过在该表面双链体的形成,来检测与该双链体结合的抗体的存在。

或者,为直接定量测定表达的基因产物,用免疫学方法测定基因表达,所述方法如细胞或组织切片的免疫组化染色和细胞培养物或体液分析。适用于免疫组化染色和/或样品液分析的抗体,可以是单克隆抗体或多克隆抗体,并且可以在任何哺乳动物中制备。可以方便地制备抗天然序列 PRO10282 多肽的抗体,或抗基于本文提供的 DNA 序列的合成肽的抗体,或者抗与 PRO10282 DNA 融合并编码特异抗体表位的外源序列的抗体。

5. 多肽的纯化

可从培养基或宿主细胞裂解物中回收 PRO10282。如果 PRO10282 与膜结合,那么使用适宜去垢剂溶液(例如 Triton-X 100)或通过酶裂解使之自细胞膜释放。使用各种物理或化学手段,如反复冻融、超声、机械破碎或细胞溶解剂,使表达 PRO10282 所用的细胞破裂。

可以期望从重组细胞蛋白或多肽中纯化 PRO10282。下述程序是适宜纯化步骤的实例:在离子-交换柱上分级分离;乙醇沉淀;反相 HPLC;在硅或阳离子-交换树脂如 DEAE 上层析;聚焦层析;SDS-PAGE;硫酸铵沉淀;使用例如交联葡聚糖(Sephadex)G-75 进行凝胶过滤;过蛋白 A 琼脂糖柱以除去污染物如 IgG;和结合表位-标记形式 PRO10282 的金属螯合剂柱。可以使用蛋白纯化的各种方法,这些方法是本领域熟知的,并在例如 Deutscher, Methods in Enzymology, 182(1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York(1982)中作了描述。纯化步骤的选择,取决于例如,所使用的生产方法和所产生的特定 PRO10282 的性质。

E. 编码 PRO10282 (Stra6)多肽的基因在肿瘤组织和细胞系中的扩增

一方面,本发明建立在对在某些癌细胞中扩增的基因的鉴定和描述上。

原核和真核生物的基因组表面上看来存在两个相矛盾的需求。一方面是作为遗传信息的 DNA 以其原始形式保存和扩增,以确保经过多代的稳定遗传。另一方面,细胞或生物体必须能够适应持久的环境变化。适应机制可以包括遗传物质在性质或数量上的修饰。性质上的修饰包括 DNA 突变,其中

编码序列被改变从而产生结构和/或功能不同的蛋白质。基因扩增是数量上的修饰，由此全编码序列即基因的实际数量增加，导致可利用的转录模板的数量增加，可翻译的转录物的数量增加，并最终导致由扩增基因编码的蛋白质的大量增加。

基因扩增的现象及其发生机制，已在数种原核和真核培养系统做过体外研究。基因扩增最具特征的实例，涉及在含有不同浓度细胞毒药物甲氨蝶呤(MTX)的介质中培养真核细胞。MTX是一种叶酸类似物，它通过阻断二氢叶酸还原酶(DHFR)而干扰DNA合成。在最初暴露于低浓度MTX时，大多数细胞(>99.9%)会死亡。少量细胞存活，并且通过产生大量DHFR-RNA和蛋白质而能够在MTX浓度升高的介质中生长。过度产生的基础是单一DHFR基因的扩增。发现该基因增加的拷贝呈小的、超数染色体(双微体)形式的非染色体拷贝或者呈集成染色体拷贝。

在发生对细胞毒性剂(细菌的抗体和真核细胞的化疗剂)抗性和致瘤性转化时最常遇到基因扩增。作为自发事件或由于病毒或化学/环境攻击而致的真核细胞的转化，通常与所述细胞中遗传物质发生变化有关。在人恶性肿瘤中观察到的最常见变化之一是p53蛋白突变。p53控制细胞从停滞期(G1)向复制期(S)转变并且在有DNA损害时阻止这种转变。换言之，p53突变体丧失功能的主要后果之一是DNA损害即遗传变化的积聚和扩增。除点突变外，肿瘤细胞中遗传变化的通常类型是扩增和严重的结构改变，如易位。

正如在DHFR实验系统所阐明的那样，DNA序列扩增可能表明特别的功能需要。因此，恶性肿瘤中一定致癌基因的扩增表明这些基因在恶性变和维持转变表型的过程中的成因性作用。该假说已经得到最近研究的支持。例如，发现bcl-2蛋白在某些类型的非何杰金氏淋巴瘤中增多。该蛋白抑制细胞凋亡并导致肿瘤细胞进行性积聚。

业已发现生长因子受体基因家族的成员在各种型癌中扩增，提示这些受体的过度表达可能使得赘生细胞对于有限量可利用的生长因子更加不敏感。实例包括：在雄激素剥夺治疗期间复发性前列腺癌中雄激素受体扩增，和乳腺癌中生长因子受体同系物ERB2扩增。最终，涉及细胞内信号和控制细胞周期进展的基因在恶性变过程中可以发生扩增。这被bel-1和ras基因在各种上皮和淋巴肿瘤中的扩增证实。

这些早期研究提示，在肿瘤中鉴定扩增的DNA序列的可行性，因为这

种方法可以鉴定对于恶性变而言非常重要的基因。由于改变蛋白可能代表肿瘤治疗的新而特异的靶, 因此, ERB2 的情况从治疗的角度也已经证明了这种可行性。

数种不同的技术可用于证明扩增的基因组序列。对制备自癌细胞的染色体伸展(spreads)的经典细胞遗传学分析足以胜任鉴定大体的结构改变, 如易位、缺失和倒位。如果有大范围的高拷贝数量或者以染色体外物质存在的话, 则仅有扩增的基因组区域可以看得见。虽然细胞遗传学是证明具体染色体改变与特定肿瘤具有一致关联性的首选技术, 但是对于鉴定和分离可控制的 DNA 序列来说是不充分的。最近开发的比较基因组杂交(CGH)技术已经证明肿瘤中广泛存在的基因组扩增现象。肿瘤 DNA 和正常 DNA 在正常细胞分裂中期同时杂交, 通过对肿瘤中以增强的频率出现的 DNA 序列进行图形分析, 可以筛选整个基因组 (WO 93118186; Gray 等, Radiation Res., 137: 275-289 [1994])。作为筛选方法, 该类型分析法已经揭示了各种人肿瘤中大量的复发(recurring)扩增子 (一段扩增的 DNA)。尽管 CGH 较经典的细胞遗传学分析法在鉴定扩增的 DNA 段(stretches)方面更为敏感, 但是使用标准分子遗传学技术 CGH 不能够快速鉴定和分离扩增子内的编码序列。

检测基因扩增最敏感的方法是基于聚合酶链反应(PCR)的分析法。这些分析法精巧灵敏, 使用极小量的肿瘤 DNA 作为起始物料, 提供用于后续分析(如序列分析)所需的 DNA, 并且该方法适于进行高容量通过量的分析。

上面提到的分析法并不相互排斥, 而是经常被联合应用以鉴定肿瘤中的扩增。尽管细胞遗传学分析和 CGH 代表检查整个基因组扩增区域的筛选方法, 但是基于 PCR 的分析法最适于编码序列即扩增区域的基因的最终鉴定。

根据本发明, 使用定量 PCR 方法, 通过将来源于各种原发性肿瘤或肿瘤细胞系的 RNA 与来源于健康供体的库存(pooled) DNA 进行比较, 已经鉴定了这些基因 (S. Gelmini 等, Clin. Chem. 43 : 752 [1997]), 所述肿瘤包括乳腺癌、肺癌、结肠癌、前列腺癌、脑瘤、肝癌、肾癌、胰腺癌、脾脏肿瘤、甲状腺癌、睾丸癌、卵巢癌、子宫癌等。使用 TaqMan 仪(ABI)完成定量 PCR 操作。根据 DNA 的编码序列设计基因-特异性引物和荧光探针。

人肺癌细胞系, 包括 A549 (SRCC768), Calu-1 (SRCC769), Calu-6 (SRCC770), H157 (SRCC771), H441 (SRCC772), H460 (SRCC773), SKMES-1 (SRCC774), SW900 (SRCC775), H522 (SRCC832)和 H810 (SRCC833), 所

有这些均可从 ATCC 获得。原发性人肺肿瘤细胞通常衍生自腺癌, 鳞状细胞癌, 大细胞癌, 非小细胞癌, 小细胞癌和支气管肺泡癌, 并且包括例如 SRCC724 (腺癌, 简称为"AdenoCa") (LT1), SRCC725 (鳞状细胞癌, 简称为"SqCCa") (LT1a), SRCC726 (腺癌) (LT2), SRCC727 (腺癌) (LT3), SRCC728 (腺癌) (LT4), SRCC729 (鳞状细胞癌) (LT6), SRCC730 (腺癌/鳞状细胞癌) (LT7), SRCC731 (腺癌) (LT9), SRCC732 (鳞状细胞癌) (LT10), SRCC733 (鳞状细胞癌) (LT11), SRCC734 (腺癌) (LT12), SRCC735 (腺鳞状细胞癌) (LT13), SRCC736 (鳞状细胞癌) (LT15), SRCC737 (鳞状细胞癌) (LT16), SRCC738 (鳞状细胞癌) (LT17), SRCC739 (鳞状细胞癌) (LT18), SRCC740 (鳞状细胞癌) (LT19), SRCC741 (肺细胞癌, 简称为"LCCa") (LT21), SRCC811 (腺癌) (LT22), SRCC825 (腺癌) (LT8), SRCC886 (腺癌) (LT25), SRCC887 (鳞状细胞癌) (LT26), SRCC888 (adeno-BAC 癌) (LT27), SRCC889 (鳞状细胞癌) (LT28), SRCC890 (鳞状细胞癌) (LT29), SRCC891 (腺癌) (LT30), SRCC892 (鳞状细胞癌) (LT31), SRCC894 (腺癌) (LT33)。也包括指定为 SRCC1125 [HF-000631], SRCC1127 [HF-000641], SRCC1129 [HF-000643], SRCC1133 [HF-000840], SRCC1135 [HF-000842], SRCC1227 [HF-001291], SRCC1229 [HF-001293], SRCC1230 [HF-001294], SRCC1231 [HF001295], SRCC1232 [HF-001296], SRCC1233 [HF-001297], SRCC1235 [HF-001299], and SRCC1236 [HF-001300]的人肺肿瘤。

结肠癌细胞系, 包括例如 ATCC 细胞系 SW480 (腺癌, SRCC776), SW620 (结肠腺癌的淋巴结转移, SRCC777), Colo320 (癌, SRCC778), HT29 (腺癌, SRCC779), HM7 (结肠腺癌细胞系 ATCC 的高粘蛋白产生的变体, SRCC780, 得自 Dr. Robert Warren, UCSF), CaWiDr (腺癌, SRCC781), HCT116 (癌, SRCC782), SKC01 (腺癌, SRCC783), SW403 (腺癌, SRCC784), LS174T (癌, SRCC785), Colo205 (癌, SRCC828), HCT15 (癌, SRCC829), HCC2998 (癌, SRCC830), 和 KM12 (癌, SRCC831)。原发性结肠肿瘤包括指定为 CT2 (SRCC742), CT3 (SRCC743), CT8 (SRCC744), CT10 (SRCC745), CT12 (SRCC746), CT14 (SRCC747), CT15 (SRCC748), CT16 (SRCC749), CT17 (SRCC750), CT1 (SRCC751), CT4 (SRCC752), CT5 (SRCC753), CT6 (SRCC754), CT7 (SRCC755), CT9 (SRCC756), CT11 (SRCC757), CT18 (SRCC758), CT19 (腺癌, SRCC906), CT20 (腺癌, SRCC907), CT21 (腺癌,

SRCC908), CT22 (腺癌, SRCC909), CT23 (腺癌, SRCC910), CT24 (腺癌, SRCC911), CT25 (腺癌, SRCC912), CT26 (腺癌, SRCC913), CT27 (腺癌, SRCC914), CT28 (腺癌, SRCC915), CT29 (腺癌, SRCC916), CT30 (腺癌, SRCC917), CT31 (腺癌, SRCC918), CT32 (腺癌, SRCC919), CT33 (腺癌, SRCC920), CT35 (腺癌, SRCC921), and CT36 (腺癌, SRCC922) 的结肠腺癌。还包括指定为 SRCC1051 [HF-000499], SRCC1052 [HF-000539], SRCC1053 [HF-000575], SRCC1054 [HF-000698], SRCC1142 [HF-000762], SRCC1144 [HF-000789], SRCC1146 [HF-000795] 和 SRCC1148 [HF-000811] 的人结肠肿瘤。

人乳腺癌细胞系, 包括例如 HBL100 (SRCC759), MB435s (SRCC760), T47D (SRCC761), MB468 (SRCC762), MB175 (SRCC763), MB361 (SRCC764), BT20 (SRCC765), MCF7 (SRCC766) 和 SKBR3 (SRCC767), 以及指定为 SRCC1057 [HF-000545] 的人乳腺肿瘤中心。也包括指定为 SRCC1094, SRCC1095, SRCC1096, SRCC1097, SRCC1098, SRCC1099, SRCC1100, SRCC1101 的人乳腺肿瘤, 和指定为 SRCC893 [LT 32] 的人乳腺-转移-肺-NS 肿瘤。

人肾肿瘤中心包括 SRCC989 [HF-000611] 和 SRCC1014 [HF-000613]。

人睾丸肿瘤中心包括 SRCC1001 [HF-000733] 和睾丸肿瘤边缘(margin) SRCC999 [HF-000716]。

人甲状腺旁腺肿瘤包括 SRCC1002 [HF-000831] 和 SRCC1003 [HF-000832]。

F. 组织分布

本文基因扩增分析的结果, 可通过进一步的研究如通过测定各种人组织中 mRNA 表达来加以证实。

如前面指出的那样, 使用基于本文提供的序列的合适标记的探针, 利用常规技术如 Southern 印迹、测定 mRNA 转录量的 Northern 印迹[Thomas, Proc. Natl. Acad. Sci. USA, 77: 5201-5205(1980)]、点印迹(DNA 分析)或原位杂交, 可以直接在样品中测定基因扩增和/或表达。或者, 使用能够识别特异双链体的抗体, 所述双链体包括 DNA 双链体、RNA 双链体和 DNA-RNA 杂化双链体或者 DNA-蛋白双链体。

或者, 为直接定量测定表达的基因产物, 用免疫学方法测定基因在各种

组织的表达,所述免疫方法如细胞或组织切片的免疫组化染色和细胞培养物或体液分析。适用于免疫组化染色和/或样品液分析的抗体,可以是单克隆抗体或多克隆抗体,并且可以在任何哺乳动物中制备抗体。可以方便地制备抗天然序列 PRO10282(StrA6)多肽的抗体,或抗基于本文提供的 DNA 序列和编码特异抗体表位的合成肽的抗体。下文提供生产抗体所用的一般技术和 Northern 印迹与原位杂交所用的特殊方案。

G. 染色体图谱

如果一指定基因的扩增是功能相关性的,那么该基因应当较对于肿瘤存活不重要的邻近基因组区域被更多地扩增。为证实这种观点,可通过例如放射-杂交分析法对特定染色体进行基因作图。然后在所确定的位置和在邻近基因组区域测定扩增水平。对选择性或优先性扩增的基因组区域的基因图谱表明,基因组区域选择性或优先性扩增与观察到的基因扩增促进肿瘤生长或存活的可能性是一致的。染色体图谱既包括框架绘图也包括中心绘图。更进一步的细节请参见例如,Stewart 等,Genome Research, 7: 422-433 (1997)。

H. 抗体结合的研究

基因扩增研究的结果,可通过抗体结合研究而得到进一步证实。在该抗体结合研究试验中,检测抗-PRO10282 (抗-STRA6)抗体抑制肿瘤(癌)细胞上 STRA6 多肽表达的能力。例证性抗体包括多克隆抗体、单克隆抗体、人源化抗体、双特异性抗体和异源偶联抗体,它们的制备将在下面予以描述。

可使用本领域已知的方法进行抗体结合研究,所述方法如竞争性结合试验、直接和间接夹心试验,及免疫沉淀试验。Zola, 单克隆抗体: 技术指南, 147-158 页(CRC Press, Inc. 1987)。

竞争性结合试验依赖于标记的标准品在同限定量抗体结合中与测试样品分析物竞争的能力。测试样品中靶蛋白(由肿瘤细胞中扩增的基因编码)的量与同抗体结合的标准品的量成反比。为便于测定结合形式的标准品的量,优选抗体在竞争之前或之后是不溶性的,这样可以很方便地把与抗体结合的标准品和测试样品从保持未结合状态的标准品和测试样品中区分开来。

夹心试验涉及使用被检测蛋白质的两种抗体,每一种抗体能够结合待测蛋白不同的免疫原性部分或表位。在一夹心试验中,首先使测试样品分析物与固定在固体载体上的第一抗体结合,然后再使第二抗体与分析物结合,从而形成不溶性的三组分复合物。参见例如美国专利 4376110。第二抗体本身

可标记有可检测部分(直接夹心试验),或者通过使用标记有可检测部分的抗-免疫球蛋白抗体来测定第二抗体(间接夹心试验)。例如,一种类型的夹心试验是 ELISA 分析法,其中可检测部分是酶。

对于免疫组织化学而言,组织样品可以是新鲜的或冷冻的,或者是包埋在石蜡中并用防腐剂如福尔马林固定。

I. 基于细胞的肿瘤分析

肿瘤(例如癌)的基于细胞的分析和动物模型,可用于证实基因扩增分析的发现,并进一步了解本文鉴定的基因与肿瘤细胞生长的进展、发病机理之间的关系。使用已经确定扩增本文所述基因的原发性肿瘤细胞或细胞系,可以检测本文鉴定的基因产物在肿瘤或癌发病和病理学中的作用。此类细胞包括例如乳腺、结肠和肺的癌细胞以及上文列举过的细胞系。

在不同的研究中,用本文描述的 cDNA 转染已知参与特定肿瘤的细胞类型的细胞,并分析这些 cDNA 诱导过度生长的能力。适宜的细胞,包括例如稳定的肿瘤细胞系如 B104-1-1 细胞系(neu 原癌基因转染的稳定的 NIH-3T3 细胞系)和 ras-转染的 NIH-3T3 细胞,可用所需的基因转染所述细胞系,并监测肿瘤发生的发展(tumorigenic growth)。然后,可将此转染的细胞系用于检测多克隆或单克隆抗体或抗体组合物通过对转化细胞生长发挥细胞抑制或细胞毒活性,或者通过介导抗体-依赖性细胞毒作用(ADCC)从而抑制肿瘤发生细胞生长的能力。用本文鉴定的基因编码序列转染的细胞,还可用于确定治疗肿瘤的候选药物。

此外,尽管优选稳定的细胞系,但是得自转基因动物(见下述)的原代培养物,可用于本文基于细胞的分析中。从转基因动物获取连续细胞系的技术为本领域所熟知(参见例如 Small 等, Mol. Cell. Biol. 5, 642-648 [1985])。

J. 动物模型

各种众所周知的动物模型,可用于进一步了解本文鉴定的基因在肿瘤发展和发病机理中的作用,并可用于检测候选治疗剂的有效性,所述候选治疗剂包括抗体和天然多肽的其它拮抗剂,后者包括小分子拮抗剂。此类模型的在体(in vivo)特性,使得它们可以具体预见在人类患者的反应。肿瘤和癌(如乳腺癌、结肠癌、前列腺癌、肺癌等)的动物模型包括非-重组动物和重组(转基因)动物。非重组动物模型包括例如啮齿动物如小鼠模型,此类模型可通过利用标准技术将肿瘤细胞引入同系基因鼠体内而制得,譬如,把结肠癌细

胞植入结肠组织中,所述标准技术如皮下注射、尾静脉注射、脾埋入法、腹膜内植入、肾囊植入法或 orthopin 植入。(参见例如,1997年9月18日公布的公开号为 WO 97/33551 的 PCT 申请)。

或许,肿瘤学研究中最常用的动物种系是免疫缺陷小鼠,更具体而言是裸鼠。观察到发育不良裸鼠可以成功地作为人肿瘤异种移植物宿主这一事实,使得裸鼠已经广泛用于此目的。常染色体隐性 nu 基因已被引入到极大量不同同类系裸鼠中,所述同类系裸鼠包括例如 ASW, A/He, AKR, BALB/c, B10.LP, C17, C3H, C57BL, C57, CBA, DBA, DDD, 1/st, NC, NFR, NFS, NFS/N, NZB, NZC, NZW, P, RIII 和 SJL。另外,已经培育了裸鼠之外的具有遗传性免疫缺陷的大量各种其它动物,并且将之用作肿瘤异种移植物的受者。进一步细节请参见例如, *The Nude Mouse in Oncology Research*, E. Boven and B. Winograd, eds., CRC Press, Inc., 1991。

引入此类动物的细胞可来自于瘤癌细胞系(tumoricancer),例如上述列举的任何肿瘤细胞,以及例如 B 104-1-1 细胞系(neu 原癌基因转染的稳定的 NIH-3T3 细胞系); ras-转染的 NIH-3T3 细胞; Caco-2 (ATCC HTB-37); II 级中度分化的人结肠腺癌细胞系, HT-29 (ATCC HTB-38), 或者来自肿瘤或癌细胞。使用标准条件,包括在液氮冷冻和贮存,可从正在接受手术的患者取得肿瘤或癌细胞样品 (Karmali 等, *Br. J.Cancer*, 48 : 689-696 [1983])。

通过多种方法可将瘤细胞引入动物如裸鼠中。小鼠的皮下(s.c.) 间隙非常适宜肿瘤植入。可以将肿瘤以如下方式移植到皮下: 固体块形式,通过使用套针而以针刺组织活检方式,或者以细胞悬液方式。对于固体块或套针方式植入,将合适尺寸的肿瘤组织片段植入到皮下间隙。对于细胞悬液方式植入,则要从原发肿瘤或稳定肿瘤细胞系新鲜配制细胞悬液并进行皮下注射。也可将肿瘤细胞作为皮下植入物进行注射。在这种场合,接种物储存在真皮结缔组织下部和皮下组织之间。Boven and Winograd (1991), 出处同上。

基本上按照 Drebin 等在 *PNAS USA*, 83 : 9129-9133 (1986)中描述的方法,通过例如把大鼠神经母细胞瘤细胞(neu 原癌基因最初就是从神经母细胞瘤中分离出来的)或 neu-转化的 NIH-3T3 细胞植入到裸鼠中,可以制得乳腺癌动物模型。

类似地,通过把结肠癌细胞植入到动物如裸鼠中,致使动物出现肿瘤,从而制得结肠癌动物模型。例如, Wang 等在 *Cancer Research*, 54 : 4726-4728

(1994) 和 Too 等在 *Cancer Research*, 55 : 681-684 (1995) 中已描述了人结肠癌在裸鼠的同位移植模型。该模型是建立在 AntiCancer, Inc. (San Diego, California) 出售的所谓 "METAMOUSE" 基础之上的。

可以把在动物体内长大的肿瘤取出并进行体外 (in vitro) 培养。然后, 再将体外培养的细胞传代给动物。此类肿瘤可以作为进一步检测或筛选药物的靶向目标。或者, 分离由传代细胞生发出的肿瘤, 并从传代前细胞和经一轮或多轮传代后的细胞中分离 RNA, 对目的基因的差异表达进行分析。可以对任何已知的肿瘤或癌细胞系实施此类传代技术。

例如, Meth A、CMS4、CMS5、CMS21 和 WEHI-164 是化学诱导的 BALB/c 雌性小鼠纤维肉瘤 (DeLeo 等, *J. Exp. Med.*, 146 : 720 [1977]), 这为研究各种制剂的抗-肿瘤活性提供了具有高度可控性的模型系统 (Palladino 等, *J. Immunol.*, 138 : 4023-4032 [1987])。简言之, 在体外通过细胞培养来繁殖肿瘤细胞, 将肿瘤细胞注射到动物体内之前, 在缓冲液中洗涤并悬浮细胞系, 细胞密度约 $10 \times 10^6 \sim 10 \times 10^7$ 细胞/ml, 然后经皮下注射 $10 \sim 100 \mu\text{l}$ 细胞悬液而感染动物, 使在 1~3 周出现肿瘤。

另外, 小鼠的 Lewis 肺(3LL)癌(此是研究最彻底的实验用肿瘤之一)可用作实验肿瘤模型。对此肿瘤模型的效力与治疗诊断患有小细胞肺癌(SCCL)的人类患者的有益效果具有良好相关性。通过注射取自感染有肿瘤的小鼠的肿瘤片段或者注射培养在培养物中的肿瘤细胞, 可以将肿瘤引入到正常小鼠体内。(Zupi 等, *Br. J. 癌*, 41 : suppl. 4 : 309 [1980]), 而且证据表明: 甚至注射单个细胞即可引发肿瘤, 和极高比例的感染肿瘤细胞存活。关于此肿瘤模型的更多信息, 请参见 Zacharski, *Haemostasis*, 16 : 300-320 [1986]。

在动物模型中评估受试化合物对植入肿瘤效力的一种方法是测量治疗前、后肿瘤的大小。通常, 用滑动卡规测量植入肿瘤的二维或三维尺寸。由于只进行二维测量并不能精确反映肿瘤的大小, 因此, 常常使用一个数学公式而将二维测量结果转换为相应的体积。然而, 肿瘤大小的测量是极不准确的。治疗引发的生长延缓和具体的生长迟滞能更好描述候选药物的治疗效果。描述肿瘤生长的另一重要变量是肿瘤体积倍增时间。也可利用计算和描述肿瘤生长的计算机程序, 如 Rygaard 和 Spang-Thomsen, *Proc. 6th Int 报道的程序*, Workshop on Immune-Deficient Animals, Wu 和 Sheng 编著, Basel, 1989, 301。但是, 应当指出, 事实上治疗后的坏死和炎症反应致使肿瘤增

大,至少在初始时是这样的。因此,需要通过形态学方法与流式细胞分析法联合,而对这些变化予以小心监控。

使用生产转基因动物的标准技术,通过向感兴趣动物的基因组引入本文鉴定的基因的编码部分,可以对重组(转基因)动物模型进行工程改造。可作为转基因靶的动物包括但不限于小鼠、大鼠、兔、豚鼠、绵羊、山羊、猪和非-人灵长类动物如狒狒、黑猩猩和猴子。本领域已知的将转基因引入动物的技术,包括原核显微注射(Hoppe 和 Wanger, 美国专利 4873191); 逆转录病毒-介导的基因转引入干细胞系(germ lines) (例如, Van der Putten 等, Proc. Natl. Acad. Sci. USA 82: 6148-615 [1985]); 基因靶向进入胚胎干细胞(Thompson 等, Cell 56: 313-321 [1989]); 胚胎的电穿孔(Lo, Mol. Cell. Biol. 3, 1803-1814 [1983]); 精子-介导的基因转移(Lavitrano et al, Cell 57, 717-73 [1989])。有关评述参见例如美国专利 4736866。

为了本发明的目的,转基因动物包括那些仅在其部分细胞携带转基因的动物("镶嵌型动物")。该转基因可以单个转基因形式或者以多联体形式(如头-头或头-尾串联体)进行整合。根据例如 Lasko 等在 Proc. Natl. Acad. Sci. USA 89, 6232-636 (1992)中公开的技术,将转基因选择性引入特定类型细胞也是可能的。

使用标准技术,可以监测转基因动物中转基因的表达。例如, Southern 印迹分析或 PCR 扩增技术可用于证实转基因的整合。然后,使用诸如原位杂交、Northern 印迹分析、PCR 或免疫细胞化学法等技术,可分析 mRNA 的表达水平。进一步检测动物的肿瘤或癌发展体征。

或者,将编码 PRO10282(Strat6)多肽的内源性基因与编码相同多肽的改变了的基因组 DNA 的同源重组物引入动物胚胎干细胞,而构建其编码本文鉴定的多肽的基因缺陷或改变的"基因敲除"动物。例如,按照已经建立的技术,编码 PRO10282 多肽的 cDNA 可用于克隆编码该多肽的基因组 DNA。编码特定 PRO10282 多肽的基因组 DNA 的一部分,可被删除或用其它基因替换,如被用于监控整合的选择性标志的编码基因替换。通常,将数千碱基对未被改变的旁侧 DNA (5'或 3'末端)引入到载体中[参见例如 Thomas 和 Capecchi, Cell, 51: 503 (1987) 对同源重组载体的描述]。将载体引入胚胎干细胞系(例如通过电穿孔),并选出被引入的 DNA 已经与内源性 DNA 发生同源重组的细胞 [参见例如, Li 等, Cell, 69: 915 (1992)]。接着,将选出的细

胞注入动物(例如小鼠或大鼠)的胚胎中,以形成聚集嵌合体[参见例如,Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson 编著(IRL, Oxford, 1987), 113-152 页]。将嵌合胚胎植入适宜的假妊娠雌性养育动物体内,胚胎生长到足月即成为"基因敲除"动物。使用标准技术,鉴定出在其干细胞中携带有同源重组 DNA 的子代动物,用这些子代动物繁殖出在所有细胞中均含有同源重组 DNA 的动物。基因敲除动物的特征在于,例如,抵御某些病理病症的能力和由于动物缺乏 PRO10282 多肽而发生病理病症。

特异结合于本文鉴定的多肽的抗体和其它候选药物的效力,也可在自发性动物肿瘤的治疗中得到测试。此类研究的合适靶是猫口腔鳞状细胞癌(SCC)。猫口腔 SCC 是高度侵袭性的恶性肿瘤,后者是猫最常见的口腔恶性肿瘤,占报道过的猫科口腔肿瘤的 60% 以上。尽管低转移发生率可能仅仅是猫患上此肿瘤后生存时间短的反映,但是该肿瘤极少转移到远端部位。主要由于猫口腔解剖学的原因,使得这些肿瘤通常不适宜接受手术。目前,对此肿瘤尚无有效的治疗办法。在接受研究之前,对每一只猫进行全身临床检查,活检并进行计算机断层摄影扫描(CT)。诊断为舌下口腔鳞状细胞肿瘤的猫排除在此研究之外。即使治疗杀灭肿瘤,舌也会因肿瘤的缘故而变得麻痹,动物不能自己进食。在一较长期间,对每只猫重复治疗。治疗期间的每一天进行肿瘤拍片(Photographs),并与后来的每一张片进行核对。治疗后,每只猫再次接受 CT 扫描。治疗后每 8 周评价 CT 扫描和 X 线胸片。数据用于评价与对照组相比在存活、反应和毒性方面的差异。阳性反应需要肿瘤退缩的证据,优选伴有生活质量改善和/或寿命增加。

另外,也可检测其它自发性动物肿瘤,如犬、猫和狒狒的纤维肉瘤、腺癌、淋巴瘤、chondroma,平滑肌肉瘤。其中,犬和猫的乳腺癌是优选的模型,因为其外观和行为与在人所见到的极为相似。然而,该模型的使用受此类型肿瘤在动物罕有发生的限制。

K. 候选药物筛选试验

候选药物筛选试验,目的是鉴定与由本文确定的基因编码的多肽结合或复合的化合物,或者相反,鉴定那些干扰基因编码的多肽与其它细胞蛋白质相互作用的化合物。此类筛选试验包括使用高通量化学数据库筛选分析,使之特别适于鉴定小分子候选药物。考虑到的小分子包括合成的有机或无机

化合物, 包括肽, 特别是可溶性肽、(多)肽-免疫球蛋白融合体, 和尤其是抗体, 包括但不限于多克隆抗体和单克隆抗体以及抗体片段、单链抗体、抗-独特型抗体和嵌合抗体或人源化抗体或片段, 以及人抗体和抗体片段。试验可按照多种方式进行, 包括蛋白质-蛋白质结合分析、生化筛选分析、免疫测定和基于细胞的分析, 这些分析方法在本领域已被很好描述。

本文中所有的候选药物筛选试验均具有共同的特性, 那就是: 在控制条件下将候选药物与由本文鉴定的核酸编码的多肽接触足够时间, 以使两种分子相互作用。

在结合分析中, 相互作用是结合, 从反应混合物中分离或检测形成的复合物。在一特定实施方案中, 通过共价或非-共价附着方式, 把由本文鉴定的基因编码的多肽或候选药物固定在固相上, 如固定到微量滴定板上。一般通过用多肽溶液包被固体表面并进行干燥而完成非-共价附着。或者, 固定抗体(例如对要被固定的多肽具有特异性的单克隆抗体)可用于使多肽固定到固体表面。试验这样进行: 向固定组分中加入非固定组分, 所述非固定组分可以标记上可检测的标记物, 所述固定组分譬如含有使多肽固着的组分的被包被表面。当反应完成时, 除去未反应的组分(例如, 通过洗涤方式), 并检测固着在固体表面的复合物。当原来非固定组分带有可检测标记物时, 检测到固定在表面的标记物, 就表明复合物形成。如果原来非-固定组分没有携带标记物, 则可以检测复合, 例如, 通过使用与固定复合物特异结合的标记抗体来进行检测。

如果候选化合物与由本文确定的基因编码的特定 PRO10282 多肽相互作用但是又不与之结合, 那么, 使用已知的检测蛋白质-蛋白质相互作用的方法来分析 PRO10282 多肽的相互作用。此类分析方法包括惯常的方法, 例如, 交联、共-免疫沉淀和通过梯度或层析柱的共-纯化。此外, 按照 Fields 与其同事(Fields 和 Song, *Nature* (340: 245-246 (1989); Chien 等, *Proc. Natl. Acad. Sci. USA*, 88: 9578-9582 (1991)) 和 Chevray 和 Nathans, *Proc. Natl. Acad. Sci. USA*, 89: 5789-5793 (1991) 公开的方法, 使用酵母-基础上的遗传系统可以检测蛋白质-蛋白质相互作用。许多转录激活剂, 如酵母 GAL4, 它由两个完全分散的调节结构域构成, 其一作为 DNA-结合结构域, 另一个作为激活转录的结构域。前述出版物中描述的酵母表达系统(一般称为"双杂种体系(two-hybrid system)")利用此特性, 并使用两个杂合蛋白, 其一的靶蛋白质

与 GAL4 的 DNA-结合区融合,另一个的候选激活蛋白质与激活结构域融合。在 GAL4-激活的启动子控制下 GAL1-lacZ 报告基因的表达,依赖于通过蛋白质-蛋白质相互作用重建的 GAL4 活性。用 β -半乳糖苷酶的显色底物检测含有相互作用多肽的菌丛。从 Clontech 可以商购利用双杂种技术鉴定两个特异蛋白质间蛋白质-蛋白质相互作用的全套试剂盒(MATCHMAKER™)。该系统也可被扩展到描绘参与特异蛋白质相互作用的蛋白质结构域,并查明对于这些相互作用起关键作用的氨基酸残基。

为了寻找干扰本文确定的编码 PRO10282 的基因与其它细胞内或细胞外组分相互作用的化合物,通常在控制条件下,制备含有扩增基因产物和细胞内或细胞外组分的反应混合物,使两种产物相互作用和结合一定时间。为了测试候选化合物抑制上述结合的能力,在含有和不含有受试化合物的条件下进行反应。另外,可把安慰剂加入到第三反应混合物中,作为阳性对照。按照上述方法,监测混合物中受试化合物与细胞内或细胞外组分之间的结合(复合物形成)。在对照反应中有复合物形成,而在含有受试化合物的反应混合物中没有复合物形成,则表明该受试化合物干预受试化合物与其配伍化合物之间的相互作用。

为分析拮抗剂,可将 PRO10282 多肽与被筛选化合物一起加入到细胞中,一定活性和能力的化合物抑制共存的 PRO10282 多肽的有关活性,则表明该化合物是 PRO10282 多肽的拮抗剂。或者,在适当条件下使 PRO10282 多肽和潜在拮抗剂与结合在膜上的 PRO10282 多肽受体或重组受体结合,通过竞争性抑制实验来检测拮抗剂。可标记 PRO10282 多肽,如放射性标记,这样,结合在受体上的 PRO10282 多肽分子数量,就可用于测定潜在拮抗剂的效力。使用本领域已知的众多方法,例如,配体淘选和 FACS 分选,可以鉴定编码受体的基因。Coligan 等, *Current Protocols in Immun.*, 1(2): Chapter 5(1991)。优选使用表达克隆,其中多聚腺苷酸化 RNA 是从对 PRO10282 多肽应答细胞中制备的,由该 RNA 建立的 cDNA 文库被分成数个库,用于转染 COS 细胞或其它对 PRO10282 多肽不应答的细胞。使玻璃载玻片上生长的转染细胞与标记的 PRO10282 多肽接触。可用各种方法标记 PRO10282 多肽,包括碘化或位点特异蛋白激酶的识别位点的包埋。固定和温育后,载玻片接受放射自显影分析。鉴定阳性库,制备亚-库,并使用交互作用的亚-库进行再转染,重新进行筛选过程,直至得到编码推定受体的单个克隆。

作为受体鉴定的另一途径,可使标记的 PRO10282 多肽与细胞膜或表达受体分子的提取制品进行光亲和连接。用 PAGE 除去交联材料,暴露于 X-线胶片。可切除含受体的标记复合物,分解为肽片段,接受蛋白微测序。从微测序得到的氨基酸序列,用于设计一套筛选 cDNA 文库的简并寡核苷酸探针,以鉴定编码推定受体的基因。

在拮抗剂另一分析方法中,在候选化合物存在的条件下,哺乳动物细胞或表达受体的膜制品与标记的 PRO10282 多肽温育。然后测定化合物增强或阻断相互作用的能力。

潜在拮抗剂的更特异实例包括与免疫球蛋白和 PRO10282 多肽的融合物结合的寡核苷酸,特别是抗体包括但不限于多克隆和单克隆抗体及抗体片段、单链抗体、抗-独特型抗体,和嵌合抗体或人源化抗体或抗体片段,以及人抗体和抗体片段。或者,潜在拮抗剂是密切相关的蛋白,例如,PRO10282 多肽的突变形式,其识别受体但不引起效应,从而竞争性抑制 PRO10282 多肽的作用。

另一潜在 PRO10282 多肽拮抗剂是使用反义技术制备的反义 RNA 或 DNA 构建体,其中例如,反义 RNA 或 DNA 分子通过与靶 mRNA 杂交和阻止蛋白翻译,起着直接阻断 mRNA 翻译的作用。反义技术可用于通过形成三体螺旋结构或通过反义 DNA 或 RNA 控制基因表达,这两种方法均建立在多核苷酸与 DNA 或 RNA 结合的基础上。例如,编码本文所述成熟 PRO10282 多肽的多核苷酸序列的 5'编码部分,被用于设计约 10~40 个碱基对长度的反义 RNA 寡核苷酸。DNA 寡核苷酸被设计成参与转录基因区的互补序列(三链螺旋参见 Lee 等, Nucl. Acids Res., 6: 3073(1979); Cooney 等, Science, 241: 456(1988); Dervan 等, Science, 251: 1360(1991)),由此阻止转录和 PRO10282 多肽产生。反义 RNA 寡核苷酸在体内与 mRNA 杂交,阻断 mRNA 分子翻译成 PRO10282 多肽(反义,见 Okano, Neurochem., 56: 560(1991); 寡脱氧核苷酸作为基因表达的反义抑制剂(CRC Press: Boca Raton, FL, 1988)。也可把上述寡核苷酸释放入细胞中,这样反义 RNA 或 DNA 可在体内表达,从而抑制 PRO10282 多肽的产生。当使用反义 DNA 时,优选衍生自翻译起始位点寡脱氧核苷酸,例如,靶基因核苷酸序列的约 -10 和 +10 位。

通常,反义 RNA 或 DNA 分子的长度至少约 5 个碱基、约 10 个碱基、

约 15 个碱基、约 20 个碱基、约 25 个碱基、约 30 个碱基、约 35 个碱基、约 40 个碱基、约 45 个碱基、约 50 个碱基、约 55 个碱基、约 60 个碱基、约 65 个碱基、约 70 个碱基、约 75 个碱基、约 80 个碱基、约 85 个碱基、约 90 个碱基、约 95 个碱基、约 100 个碱基或更多。

潜在的拮抗剂包括与活性位点、受体结合位点、或生长因子或 PRO10282 多肽其它相关结合位点结合的小分子，从而阻断 PRO10282 多肽的正常生物活性。小分子的实例包括但不限于，小肽或类-肽分子，优选溶解的肽和合成的非-肽有机或无机化合物。

核酶是能够催化 RNA 特异性裂解的具有酶性质的 RNA 分子。核酶通过序列-特异性杂交于互补的靶 RNA，接着产生核酸内裂解。使用已知技术，可以鉴定位于潜在 RNA 靶内的特异核酶裂解位点。进一步的细节请参见例如 Rossi, Current Biology, 4: 469-471(1994), 和 PCT 出版物 WO 97/33551(1997 年 9 月 18 日公布)。

用于抑制转录的三链螺旋构型的核酸分子，应当是单链并由脱氧核苷酸构成。这些寡核苷酸碱基组成是设计好的，这样它通过 Hoogsteen 碱基配对原则促进三链螺旋的形成，这一般需要双螺旋中一条链上的嘌呤或嘧啶相当大的延伸。更详细情况请参见例如，PCT 出版物 WO 97/33551，出处同上。

用上述讨论的筛选试验的任何一个或多个分析方法，和/或使用本领域技术人员熟知的任何其它筛选技术，可以鉴定这些小分子。

L. PRO10282 (Stra6)的应用

编码 PRO10282(Stra6)多肽的核苷酸序列(或其互补序列)在分子生物学领域具有多种用途，包括用作杂交探针、在染色体和基因图谱以及在制备反义 RNA 和 DNA 中的用途。PRO10282 核酸对于通过本文所述重组技术来制备 PRO10282 多肽，也是非常有用的。

全长天然序列 PRO10282 基因(SEQ ID NO: 1)或 PRO19578 基因(SEQ ID NO:4)或它们的一部分，可用作 cDNA 文库的杂交探针以分离出全长 PRO10282 或 PRO19578 cDNA，或分离出与图 1(SEQ ID NO: 1)中公开的 PRO10282 序列或图 6(SEQ ID NO: 4)中公开的 PRO19578 序列具有期望序列同一性的其它 cDNA(例如，那些编码其它天然发生的 PRO10282 变体或来自其它种属的 PRO10282 的 cDNA)。任选地，探针长度是约 20 ~ 约 50 个碱基。杂交探针可源自于 SEQ ID NO: 1 核苷酸序列的至少部分新区，其中所述新

区是那些不需要十分繁琐实验就可确定的区域,或从包括启动子、增强子元件和内含子的天然 PRO10282 或 PRO19578 基因的基因组序列中确定的区域。举例来说,筛选方法包括:使用已知 DNA 序列合成约 40 个碱基的选定探针,来分离 PRO10282 或 PRO19578 基因的编码区。杂交探针可标记上各种标志,包括放射性核苷酸如 ^{32}P 或 ^{35}S 或酶标记,所述酶标记如通过亲和素/生物素偶联系统将碱性磷酸酶与探针耦合。具有与本发明 PRO10282 或 PRO19578 基因互补序列的标记探针,可用于筛选 cDNA 文库、基因组 DNA 或 mRNA,以测定探针与哪一文库发生杂交。下面的实施例将进一步详细描述杂交技术。

使用本文公开的方法,本申请公开的任何 EST 序列可类似地用作探针。

PRO10282 (Stra6)核酸的其它有用片段,包括含有能与靶 PRO10282 (Stra6) mRNA(正义)或 PRO10282(Stra6) DNA(反义)序列结合的单链核酸序列(RNA 或 DNA)的反义或正义寡核苷酸。根据本发明,反义或正义寡核苷酸包括 PRO10282(Stra6) DNA 编码区的片段。该片段一般含有至少约 14 个核苷酸,优选地由约 14~30 个核苷酸。基于编码一给定蛋白的 cDNA 序列得到反义或正义寡核苷酸的能力,在例如 Stein 和 Cohen (Cancer Res. 48: 2659, 1988)及 van der Krol 等(BioTechniques 6: 958, 1988)的文章中有描述。

反义或正义寡核苷酸与靶核酸序列的结合,导致通过一种或多种方式阻断靶序列转录或翻译的双螺旋形成,所述阻断方式包括增加双螺旋降解、转录或翻译的成熟前终止或者其它方式。因此,反义寡核苷酸可用于阻断 PRO10282(Stra6)蛋白的表达。反义或正义寡核苷酸进一步包括具有修饰过的糖磷酸二酯主链的寡核苷酸(或其它糖链接,如 WO 91/06629 中描述的那些),其中此类糖链接是耐内源性核酸酶的。具有抗性糖链接的寡核苷酸在体内是稳定的(即,能够抗酶降解),同时保留序列与靶核苷酸序列结合的序列特异性。

正义或反义寡核苷酸的其它实例包括那些与有机分子共价键合的寡核苷酸,如 WO 90/10048 中描述的那些寡核苷酸,和其它增加寡核苷酸与靶核酸序列亲和力的分子,如聚-(L-赖氨酸)。另外,嵌入剂(如椭圆玫瑰树碱)和烷基化剂或金属络合物可挂接在正义或反义寡核苷酸上,以修饰反义或正义寡核苷酸对靶核苷酸序列的结合特异性。

使用任何基因转移方法可将反义或正义寡核苷酸引入到含有靶核酸序

列的细胞中,所述方法包括例如, CaPO_4 -介导的 DNA 转染、电穿孔、或使用基因转移载体如 Epstein-Barr 病毒。在一优选的方法中,将反义或正义寡核苷酸插入到一适宜的逆转录病毒载体中,在体内或体外使含有靶核酸序列的细胞与重组逆转录病毒载体接触。适宜的逆转录病毒载体包括但不限于,得自小鼠逆转录病毒 M-MuLV、N2(得自 M-MuLV 的逆转录病毒),或称为 DCT5A、DCT5B 和 DCT5C 的双倍复制载体(见 WO 90/13641)。

也可以通过使正义或反义寡核苷酸与配体结合分子形成偶联物的形式,而将正义或反义寡核苷酸引入到含有靶核苷酸序列的细胞中,见 WO 91/04753 中的描述。适宜的配体结合分子包括但不限于:细胞表面受体、生长因子、其它细胞因子或能与细胞表面受体结合的其它配体。优选地,配体结合分子的偶联基本上不干扰配体结合分子与其相应分子或受体结合的能力,或阻断正义或反义寡核苷酸或其偶联型进入细胞。

或者,按照 WO 90/10448 描述的方法,通过形成寡核苷酸-脂复合物,而将正义或反义寡核苷酸引入到含有靶核酸序列的细胞中。优选地,正义或反义寡核苷酸-脂复合物是在细胞内使用内源性脂酶分离的。

也可在 PCR 技术中使用探针,以产生用于鉴定最接近 PRO10282 编码序列的一组序列。

编码天然序列 PRO10282(Stra6)多肽的核苷酸序列,还可用于构建在描绘编码 PRO10282(Stra6)的基因图谱和对患有遗传病个体进行基因分析中使用的杂交探针。使用已知技术,如原位杂交、抗已知染色体标记物的连锁分析和与文库的杂交筛选技术,可将本文提供的核苷酸序列作图于染色体和染色体的特定区域。

当 PRO10282 的编码序列编码能与另一蛋白结合的蛋白时(例如,PRO10282 是受体时),PRO10282 可用在鉴定参与结合反应的其它蛋白或分子的分析中。使用此类方法,可以鉴定受体/配体结合反应的抑制剂。还可将参与此结合反应的蛋白用于筛选肽或结合反应的小分子抑制剂或激动剂。而且,受体 PRO10282(Stra6)可用于分离相关配体。可设计筛选分析以寻找模拟天然 PRO10282(Stra6)或与 PRO10282 结合的天然多肽(“PRO10282 结合蛋白”)生物活性的主要化合物。此类筛选分析包括使用化学用数据库进行高-通过量筛选分析,使之特别适宜于鉴定小分子候选药物。小分子包括合成的有机或无机化合物。可使用各种形式的分析法,包括蛋白-蛋白结合分

析,生化筛选分析,免疫测定和以细胞为基础的分析,这些分析法均是本领域公知的。

编码 PRO10282 或其修饰形式的核酸,也可用于生产转基因动物或"敲除"动物,所得动物反过来对于开发和筛选适合的治疗剂是非常有用的。转基因动物(例如小鼠或大鼠)是含有转基因细胞的动物,所述转基因被引入到动物或孕期,例如胚胎期祖代动物中。转基因是整合入细胞基因组的 DNA,由所述细胞发展为转基因动物。在一实施方案中,使用已经建立的技术和用于生产含有表达编码 PRO10282 DNA 的细胞的转基因动物的基因组序列,编码 PRO10282 的 cDNA 可用于克隆编码 PRO10282 的基因组 DNA。生产转基因动物的方法,特别是诸如小鼠或大鼠这样的动物,已经成为本领域常规技术,并在例如美国专利 4736866 和 4870009 中有描述。通常,带有组织-特异增强子的 PRO10282 转基因被整合入特定细胞中。包含在胚胎期被整合入动物干细胞一个拷贝编码 PRO10282 转基因的转基因动物,可用于检测增加编码 PRO10282 DNA 表达的效果。此类动物可用作检测认为有可能对与过度表达有关的病理病症具有保护作用的药剂的试验动物。根据本发明的一个方面,与未经治疗的携带转基因的动物相比,用药剂治疗的动物,其病理病症的发生率降低,这表明该药剂对于这种病理病症具有潜在的治疗干预作用。

或者,PRO10282 的非-人源同系物用于构建 PRO10282 的"敲除"动物,该动物具有缺陷的或改变的 PRO10282 编码基因,这是将内源性 PRO10282 编码基因和改变的编码 PRO10282 基因组 DNA 之间的同源重组物引入动物胚胎干细胞的结果。例如,按照已经建立的技术,可将编码 PRO10282 的 cDNA 用于克隆编码 PRO10282 的基因组 DNA。编码 PRO10282 基因组 DNA 的一部分,可被删除或用其它基因替换,如被用于监控整合的选择性标志的基因编码替换。通常,将数千碱基对未被改变的侧面 DNA(5'或 3'末端)引入到载体中[参见例如 Thomas 和 Capecchi, Cell, 51: 503(1987)对同源重组载体的描述]。将载体引入胚胎干细胞系(例如通过电穿孔),并选出引入了已经与内源性 DNA 同源重组的 DNA 的细胞[参见例如, Li 等, Cell, 69: 915(1992)]。接着,将选出的细胞注入动物(例如小鼠或大鼠)的胚泡中,以形成聚集嵌合体[参见例如, Bradley, in Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson 编著(IRL, Oxford, 1987), 113-152

页]。将嵌合胚胎植入适宜的假妊娠雌性养育动物体内,胚胎生长到足月即成为"敲除"动物。使用标准技术,鉴定出在其干细胞中携带有同源重组 DNA 的子代动物,用这些子代动物繁殖出在所有动物细胞中均含有同源重组 DNA 的动物。敲除动物的特征在于,例如,抵御某些病理病症的能力和由于动物缺乏 PRO10282 多肽而发生病理病症。

也可将编码 PRO10282 多肽的核酸用于基因治疗。在基因治疗应用中,将基因引入细胞以便达到在体内合成治疗有效的基因产物,例如置换缺陷基因。"基因治疗",既包括经单次治疗就可获得永久效果的常规基因治疗,也包括基因治疗剂的给药,后者涉及一次或重复多次施用治疗有效量的 DNA 或 mRNA。反义 RNA 和 DNA 可用作阻断体内某些基因表达的治疗剂。业已证明,短链反义寡核苷酸可整合入细胞中,在所述细胞中,短链反义寡核苷酸起抑制剂的作用,而不论由于细胞膜的限制性摄取所导致其低细胞内浓度如何。(Zamecnik 等, *Proc. Natl. Acad. Sci. USA* 83: 4143-4146 [1986])。可修饰寡核苷酸以增强其摄取,例如无电荷的基团取代带阴性电荷的磷酸二酯。

有各种可利用的技术来把核酸引入到活细胞中。所述技术依是否将核酸转移进体外培养细胞或目的宿主体内细胞中而异。将核酸转移入哺乳动物体外细胞的适宜技术包括使用脂质体、电穿孔、微注射、细胞融合、DEAE-葡聚糖、磷酸钙沉淀法等。目前优选的体内基因转移技术包括用病毒(通常为逆转录病毒)载体转染和用病毒包衣蛋白-脂质体介导的转染(Dzau 等, *Trends in Biotechnology* 11, 205-210 [1993])。有些情况下,希望提供带有针对靶细胞的物质的核酸源,所述物质如对细胞表面的膜蛋白具有特异性的抗体,或靶细胞上受体的配体等。使用脂质体时,将与参与胞吞的细胞表面膜蛋白结合的蛋白用作靶向和/或促进摄取的物质,例如对特定型细胞具有定向性的病毒衣壳蛋白或其片段、在循环中发生内化的蛋白的抗体、靶向于细胞内定位(localization)并增加细胞内半衰期的蛋白。有关受体-介导的胞吞的技术在例如 Wu 等, *J. Biol. Chem.* 262, 4429-4432(1987); 和 Wagner 等, *Proc. Natl. Acad. Sci. USA* 87, 3410-3414(1990)中有描述。关于基因标志和基因治疗方案的综述,请参见 Anderson 等, *Science* 256, 808-813(1992)。

本文所述 PRO10282 多肽也可用作电泳中蛋白的分子量标志。

本文所述编码 PRO10282 多肽或其片段的核酸分子在染色体鉴定中非

常有用。在此方面,由于基于实际序列数据只有相对少数几个可利用的染色体标志剂,因而一直存在着对于鉴定新染色体标志的需求。本发明每一个 PRO10282 核酸分子均可用作染色体标志。

本发明 PRO10282 多肽和核酸分子也可用于组织分型,其中本发明 PRO10282 多肽在不同组织间有差别地表达。PRO10282 核酸分子还可用于制备 PCR、Northern 印迹、Southern 印迹和 Western 印迹分析用的探针。

本文所述 PRO10282 多肽也可用作治疗剂。按照已知制备有益药物组合物的方法,可制得本发明 PRO10282 多肽,所制得的 PRO10282 产品与可药用载体呈混合物结合。为了所制备的治疗剂的贮存,将具有所需纯度的活性组分与任选的可药用载体、赋形剂或稳定剂混合(Remington's Pharmaceutical Sciences,第16版,Osol, A.编著(1980)),制备成冷冻干燥制剂或水溶液形式。可接受的载体、赋形剂或稳定剂,在所用剂量和浓度范围对接受者没有毒性,其包括缓冲剂,如磷酸盐、柠檬酸盐和其它有机酸;抗氧化剂,包括抗坏血酸;低分子量多肽(小于约10个残基);蛋白,如血清白蛋白、明胶或免疫球蛋白;亲水性聚合物,如聚乙烯吡咯烷酮,氨基酸如甘氨酸、谷氨酰胺、天冬酰胺、精氨酸或赖氨酸;单糖、二糖和其它碳水化合物,如葡萄糖、甘露糖或糊精;螯合剂,如 EDTA;糖醇如甘露醇或山梨醇;成盐离子,如钠;和/或非离子表面活性剂,如 TWEENTM、PLURONICSTM 或 PEG。

用于体内给药的制剂必须是无菌的。这可以在冷冻干燥和重新配制之前或之后通过除菌膜过滤而轻易实现。

本文的治疗组合物一般盛放在具有无菌存取口的容器中,例如该容器是静脉注射袋或带有能通过皮下注射针穿刺的塞子的小瓶。

给药途径与已知方法一致,例如注射或静脉输注、腹膜内、颅内、肌内、眼内、动脉内或损伤部位内给药,局部用药或使用持续释放系统给药。

本发明药物组合物的剂量和所需药物浓度,将依预想的特定用途而异。决定适宜剂量或给药途径是普通内科医师完全能够很好胜任的事情。动物试验为决定治疗人类患者所用剂量提供可靠指导。可按照 Yacobi 等编著的 Toxicokinetics and New Drug Development(Pergamon Press, New York 1989)一书中, Mordenti, J.和 Chappell 在 W."The use of interspecies scaling in toxicokinetics"第42-96页所给出的方法换算种属间有效剂量的比例。

体内施用 PRO10282 多肽或其激动剂或拮抗剂时,取决于给药途径,每

日标准剂量范围为约 10 ng/kg 至高达 100 mg/kg 哺乳动物体重或更多, 优选地约 $1 \mu\text{g/kg/日}$ ~ 10mg/kg/日。文献提供了有关特定剂量的指导和给药方法, 参见例如美国专利 4657760; 5206344; 或 5225212。对于不同治疗化合物和不同病症而言, 可以预期不同制剂能够奏效, 例如, 施用于一种器官或组织的制剂必须以不同于施用于其它器官或组织的方式进行给药。

当期望 PRO10282 多肽在具有适宜于治疗那些需要施用 PRO10282 多肽的任何疾病或病症的释放特性的制剂中持续释放给药时, 考虑使用微囊化。持续释放用的重组蛋白微囊化已被成功地用于人生长激素(rhGH)、干扰素-(rhIFN-)、白介素-2 和 MN rgpl20 的持续释放。Johnson 等, Nat. Med., 2: 795-799(1996); Yasuda, Biomed. Ther., 27: 1221-1223(1993); Hora 等, Bio/Technology 8: 755-758(1990); 在 Powell 和 Newman 编著的 Vaccine Design: The Subunit and Adjuvant Approach(Plenum Press: New York, 1995)一书中第 439-462 页, Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems"; WO 97/03692, WO 96/40072, WO 96/07399; 和美国专利 5654010。

由于聚乳酸-共乙醇酸 (PLGA) 聚合物具有生物相容性很宽范围的生物降解特性, 所以使用该聚合物来制备蛋白的持续释放制剂。PLGA 的降解产物, 即乳酸和羟基乙酸, 在体内被快速清除。而且, 可以根据该聚合物的分子量和构成来调节其降解特性, 可以是数月 to 数年不等。在 M. Chasin 和 R. Langer(编著), Biodegradable Polymers as Drug Delivery Systems(Marcel Dekker: New York, 1990)一书中第 1-41 页, Lewis, "Controlled release of active agents from lactide/glycolide polymer"。

本发明包括筛选化合物的方法, 以鉴定那些模拟 PRO10282 多肽(激动剂)或者阻止或抑制 PRO10282 多肽(拮抗剂)作用的化合物。拮抗剂候选药物筛选试验, 目的是鉴定与由本文确定的基因编码的 PRO10282 多肽结合或复合的化合物, 或者相反, 鉴定那些干扰基因编码的多肽与其它细胞蛋白相互作用的化合物。此类筛选试验包括使用高通量化学数据库筛选分析, 使之特别适于鉴定小分子候选药物。

试验可按照多种方式进行, 包括蛋白-蛋白结合分析、生化筛选分析、免疫测定和基于细胞的分析, 这些分析方法在本领域已被很好描述。

所有拮抗剂试验共同的方面是, 它们使候选药物与由本文确定的核酸编

码的 PRO10282 多肽接触足够时间以使得这两种组分相互作用。

在结合试验中，相互作用是结合，形成的复合物可被分离或在反应混合物中被检测。在一特定实施方案中，通过共价或非-共价地附着，而把由本文确定的基因编码的 PRO10282 多肽或候选药物固定在固相上，例如，固定在微量滴定盘上。一般通过用 PRO10282 多肽溶液包被固体表面并进行干燥而实现非-共价附着。或者，固定抗体，例如，固定的对 PRO10282 多肽具有特异性的单克隆抗体，可用于使 PRO10282 多肽固定到固体表面。试验这样进行：向固定组分中加入非固定组分，所述非固定组分可以标记上可检测的标记物，譬如，包被的表面含有使 PRO10282 多肽固定的组分。当反应完成时，除去未反应的组分(例如，通过洗涤方式)，并检测固定在固体表面的复合物。当原来非固定组分带有可检测标记物时，检测到固定在表面的标记物，就表明复合物形成。如果原来非-固定组分没有携带标记物，则可以检测到复合，例如，通过使用与固定复合物特异结合的标记的抗体。

如果候选化合物与本文确定的基因编码的特定 PRO10282 多肽相互作用但是又不与之结合，那么，使用已知的检测蛋白-蛋白相互作用的方法来分析那个多肽的相互作用。此类分析方法包括惯常的方法，例如，交联、共-免疫沉淀和通过梯度或层析柱的共-纯化。此外，按照 Fields 与其同事 (Fields 和 Song, Nature(340: 245-246(1989); Chien 等, Proc. Natl. Acad. Sci. USA, 88: 9578-9582(1991))和 Chevray 和 Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793(1991)公开的方法，使用酵母-基础上的遗传系统可以检测蛋白-蛋白相互作用。许多转录激活剂，如酵母 GAL4，它由两个完全分散的调节结构域构成，其一作为 DNA-结合结构域，另一个作为激活转录的结构域。前述出版物中描述的酵母表达系统(一般指"双杂种体系")利用此特性，并使用两个杂合蛋白，其一的靶蛋白与 GAL4 的 DNA-结合域融合，另一个的候选激活蛋白与激活结构域融合。在 GAL4-激活的启动子控制下 GAL1-lacZ 报告基因的表达，依赖于通过蛋白-蛋白相互作用重建的 GAL4 活性。用对-半乳糖苷酶的显色底物检测含有相互作用多肽的菌丛。从 Clontech 可以商购利用双杂种技术鉴定两个特异蛋白间蛋白-蛋白相互作用的全套试剂盒 (MATCHMAKER™)。该系统也可被扩展到描绘参与特异蛋白相互作用的蛋白结构域，并查明对于这些相互作用起关键作用的氨基酸残基。

干预本文鉴定的基因编码的 PRO10282 多肽与其它细胞内或细胞外组

分相互作用的化合物可如下测定：通常在控制条件下，制备含有基因产物和细胞内或细胞外组分的反应混合物，使两种产物相互作用和结合一定时间。为了测试候选化合物抑制结合的能力，在含有和不含有受试化合物的条件下进行反应。另外，可把安慰剂加入到第三反应混合物中，作为阳性对照。按照上述方法，监测混合物中受试化合物与细胞内或细胞外组分之间的结合(复合物形成)。在对照反应中有复合物形成，而在含有受试化合物的反应混合物中没有复合物形成，表明该受试化合物干预受试化合物与其配伍化合物之间的相互作用。

为分析拮抗剂，可将 PRO10282 多肽与被筛选化合物一起加入到细胞中，一定活性和能力的化合物抑制共存的 PRO10282 多肽的有关活性，则表明该化合物是 PRO10282 多肽的拮抗剂。或者，在适当条件下使 PRO10282 多肽和潜在拮抗剂与结合在膜上的 PRO10282 多肽受体或重组受体结合，通过竞争性抑制实验来检测拮抗剂。可标记 PRO10282 多肽，如放射性标记，这样，结合在受体上的 PRO10282 多肽分子数量，就可用于测定潜在拮抗剂的效力。使用本领域已知的众多方法，例如，配体淘选和 FACS 分选，可以鉴定编码受体的基因。Coligan 等，*Current Protocols in Immun.*, 1(2): Chapter 5(1991)。优选使用表达克隆，其中多聚腺苷酸化 RNA 是从对 PRO10282 多肽应答细胞中制备的，由该 RNA 建立的 cDNA 文库被分成数个库，用于转染 COS 细胞或其它对 PRO10282 多肽不反应的细胞。使玻璃载玻片上生长的转染细胞与标记的 PRO10282 多肽接触。可用各种方法标记 PRO10282 多肽，包括碘化或位点特异蛋白激酶的识别位点的包埋。固定和温育后，载玻片接受放射自显影分析。鉴定阳性库，制备亚-库，并使用交互作用的亚-库进行再转染，重新进行筛选过程，直至得到编码推定受体的单个克隆。

作为受体鉴定的另一途径，可使标记的 PRO10282 多肽与细胞膜或表达受体分子的提取制品进行光亲和连接。用 PAGE 除去交联材料，暴露于 X-线胶片。可切除含受体的标记复合物，分解为肽片段，接受蛋白微测序。从微测序得到的氨基酸序列，用于设计一套筛选 cDNA 文库的简并寡核苷酸探针，以鉴定编码推定受体的基因。

在拮抗剂另一分析方法中，在候选化合物存在的条件下，哺乳动物细胞或表达受体的膜制品与标记的 PRO10282 多肽温育。然后测定化合物增强或阻断相互作用的能力。

潜在拮抗剂的更特异实例包括与免疫球蛋白和 PRO10282 多肽的融合物结合的寡核苷酸,特别是抗体,包括但不限于多克隆和单克隆抗体及抗体片段、单链抗体、抗-独特型抗体,和嵌合抗体或人源化抗体或抗体片段,以及人抗体和抗体片段。或者,潜在拮抗剂是密切相关的蛋白,例如,PRO10282 多肽的突变形式,其识别受体但不引起效应,从而竞争性抑制 PRO10282 多肽的作用。

另一潜在 PRO10282 多肽拮抗剂是使用反义技术制备的反义 RNA 或 DNA 构建体,其中例如,反义 RNA 或 DNA 分子通过与靶 mRNA 杂交和阻止蛋白翻译,起着直接阻断 mRNA 翻译的作用。反义技术可用于通过形成三体螺旋或通过反义 DNA 或 RNA 控制基因表达,两种方法均建立在多核苷酸与 DNA 或 RNA 结合的基础上。例如,编码本文所述成熟 PRO10282 多肽的多核苷酸序列的 5'编码部分,,被用于设计约 10~40 个碱基对长度的反义 RNA 寡核苷酸。DNA 寡核苷酸被设计成参与转录基因区的互补序列(三链螺旋,参见 Lee 等, Nucl. Acids Res., 6: 3073(1979); Cooney 等, Science, 241: 456(1988); Dervan 等, Science, 251: 1360(1991)),由此阻止转录和 PRO10282 多肽产生。反义 RNA 寡核苷酸在体内与 mRNA 杂交,阻断 mRNA 分子翻译成 PRO10282 多肽(反义,见 Okano, Neurochem., 56: 560(1991); 寡脱氧核苷酸作为基因表达的反义抑制剂(CRC Press: Boca Raton, FL, 1988)。也可把上述寡核苷酸释放入细胞中,这样反义 RNA 或 DNA 可在体内表达,从而抑制 PRO10282 多肽的产生。当使用反义 DNA 时,优选衍生自翻译起始位点寡脱氧核苷酸,例如,靶基因核苷酸序列的约 -10 和 +10 位。

潜在的拮抗剂包括与活性位点、受体结合位点、或生长因子或 PRO10282 多肽其它相关结合位点结合的小分子,从而阻断 PRO10282 多肽的正常生物活性。小分子的实例包括但不限于,小肽或类-肽分子,优选溶解的肽和合成的非-肽有机或无机化合物。

核酶是能够催化 RNA 特异性裂解的具有酶性质的 RNA 分子。核酶通过序列-特异性杂交于互补的靶 RNA,接着产生核酸内裂解。使用已知技术,可以鉴定位于潜在 RNA 靶内的特异核酶裂解位点。进一步的细节请参见例如 Rossi, Current Biology, 4: 469-471(1994),和 PCT 出版物 WO 97/33551(1997 年 9 月 18 日公布)。

用于抑制转录的三链螺旋构型的核酸分子,应当是单链并由脱氧核苷酸

构成。这些寡核苷酸碱基组成是设计好的，这样它通过 Hoogsteen 碱基对原则促进三链螺旋的形成，这一般需要双螺旋中一条链上的嘌呤或嘧啶相当大的延伸。更详细情况请参见例如，PCT 出版物 WO 97/33551，出处同上。

用上述讨论的筛选试验的任何一个或多个分析方法，和/或使用本领域技术人员熟知的任何其它筛选技术，可以鉴定这些小分子。

Stra6 - 小鼠视黄酸诱导基因，在睾丸的塞尔托利氏(Sertoli)细胞中有生精周期-依赖性表达。基于与 Stra6 的序列同源性，本文公开的 PRO10282 多肽在精子发生中具有重要作用，因而对于生育能力治疗具有潜在用途。目前认为 PRO10282 多肽的表达正如其小鼠同源物 Stra6 一样，是对视黄酸应答而被诱导的。所以，通过杂交测定 PRO10282 在 mRNA 水平的表达，或通过免疫分析法测定 PRO10282 在蛋白水平的表达，可用于决定受试化合物对于治疗广泛的各种维生素 A 类应答性疾病是否具有潜在的治疗作用。考虑到的作为治疗靶的疾病类型包括牛皮癣，痤疮，发育异常，癌和自身免疫性疾病。考虑到的发育异常的种类包括上皮组织的癌前病变如口腔粘膜白斑病，子宫颈发育异常，喉和支气管发育异常。考虑到的癌的种类包括结肠癌(腺癌)，肺癌(癌和腺癌)，皮肤癌，头颈部癌，子宫颈癌，子宫癌，乳腺癌和前列腺癌。考虑到的自身免疫性疾病的种类包括类风湿性关节炎，骨关节炎，系统性红斑狼疮，寻常性天疱疮和落叶状天疱疮。治疗这些疾病或病症的潜在治疗剂是 PRO10282 的抗体和拮抗剂(包括小分子)。

M. 治疗肿瘤的组合物和方法

适宜于治疗与本文鉴定的基因相关的肿瘤的组合物，包括但不限于抑制目标基因产物表达和/或活性的抗体、小的有机分子、肽、磷肽、反义分子和核酶分子、三链螺旋分子等。

举例来说，反义 RNA 和 RNA 分子通过与目标 mRNA 杂交和阻止蛋白质翻译而直接阻断 mRNA 翻译。当使用反义 DNA 时，优选得自翻译起始位点(例如靶基因核苷酸序列约-10~+10 位)的寡脱氧核糖核苷酸。

核酶是能够催化 RNA 特异性裂解的具有酶性质的 RNA 分子。核酶通过序列-特异性杂交于互补的目标 RNA，接着进行核酸内裂解。使用已知技术，可以鉴定位于潜在 RNA 目标内的特异核酶裂解位点。进一步的细节请参见例如 Rossi, Current Biology, 4: 469-471 (1994)，和 PCT 出版物 WO 97/33551 (1997 年 9 月 18 日公布)。

用于抑制转录的三链螺旋构型的核酸分子,应当是单链并由脱氧核苷酸构成。这些寡核苷酸的碱基组成是设计好的,这样它通过 Hoogsteen 碱基对原则促进三链螺旋的形成,这一般需要双螺旋中一条链上的嘌呤或嘧啶相当大的延伸。更详细情况请参见例如, PCT 出版物 WO 97/33551, 出处同上。

用上述讨论的筛选试验中的任何一个或其任何联合,和/或使用本领域技术人员熟知的任何其它筛选技术,可以鉴定这些分子。

N. 抗-PRO10282 (抗-Stra6) 抗体

本发明进一步提供抗-PRO10282(抗-Stra6)抗体。抗体的实例包括多克隆抗体、单克隆抗体、人源化抗体、双特异性抗体和异源偶联抗体。

1. 多克隆抗体

抗-PRO10282 抗体包括多克隆抗体。制备多克隆抗体的方法为本领域熟练技术人员所已知。多克隆抗体可在哺乳动物中产生,例如,通过一次或多次注射致免疫剂,如果需要,可加入佐剂。通常,对哺乳动物皮下或腹膜内多次注射致免疫剂和/或佐剂。致免疫剂可包括 PRO10282 多肽或其融合蛋白。将致免疫剂与针对所免疫的哺乳动物具有免疫原性的蛋白进行偶联是有效的。所述免疫原性蛋白的实例包括但不限于:匙孔血蓝蛋白、血清白蛋白、牛甲状腺球蛋白和大豆胰蛋白酶抑制剂。可以使用的佐剂的实例包括弗氏完全佐剂和 MPL-TDM 佐剂(单磷脂酰脂质 A, 合成的海藻糖双白喉菌酸酯(dicorynomycolate))。无需繁琐的试验,本领域技术人员就可选择免疫方案。

2. 单克隆抗体

或者,抗-PRO10282 抗体是单克隆抗体。可以使用杂交瘤方法制备单克隆抗体,如 Kohler 和 Milstein 在 Nature, 256:495(1975)中描述的方法。在杂交瘤方法中,通常用致免疫剂来免疫小鼠、仓鼠或其它适宜宿主动物,以引发产生或能够产生将与致免疫剂特异性结合的抗体的淋巴细胞。或者,在体外致敏淋巴细胞。

致免疫剂通常包括 PRO10282 多肽或其融合蛋白。一般而言,如果需要来源于人的细胞,则使用外周血淋巴细胞("PBLs"),如果需要非人哺乳动物来源的细胞,那么就使用脾细胞或淋巴结细胞。然后,使用适宜的融合试剂如聚乙二醇,将淋巴细胞与无限增殖细胞系融合,从而形成杂交瘤细胞[Goding, 单克隆抗体: Principles 和 Practice, Academic Press, (1986)59-103 页]。无限增殖细胞系通常是转化的哺乳动物细胞,特别是啮齿类、牛和人

类起源的骨髓瘤细胞。通常使用大鼠或小鼠骨髓瘤细胞系。可在适宜培养基中培养杂交瘤细胞,所述培养基优选含有一种或多种抑制未融合的无限增殖细胞生长或生存的物质。例如,亲代细胞缺乏酶次黄嘌呤鸟嘌呤磷酸核糖转移酶(HGPRT 或 HPRT),杂交瘤的培养基通常包括次黄嘌呤、氨基蝶呤和胸苷("HAT 培养基"),这些物质抑制 HGPRT-缺陷细胞生长。

优选的无限增殖细胞系是那些有效融合、支持筛选出的抗体-生成细胞稳定地高水平表达抗体,并且对培养基如 HAT 敏感的细胞系。更优选的无限增殖细胞系是鼠骨髓瘤系,后者可得自例如 Salk Institute Cell Distribution Center, San Diego, California 和美国典型培养物保藏中心, Manassas, Virginia。人骨髓瘤和鼠-人杂交骨髓瘤细胞系也已用于产生人单克隆抗体[Kozbor, J.Immunol., 133: 3001(1984); Brodeur 等, Monoclonal antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63]。

接下来,可以检定培养杂交瘤细胞的培养基中是否存在直接抗 PRO10282 的单克隆抗体。优选地,用免疫沉淀或体外结合试验如放免分析(RIA)或酶联免疫吸附分析(ELISA),来测定杂交瘤细胞产生的单克隆抗体的结合特异性。这些技术和试验分析方法是本领域所已知的。使用例如 Scatchard analysis of Munson 和 Pollard, Anal. Biochem., 107: 220(1980)公开的方法,可以测定单克隆抗体的结合亲和力。

鉴定出所需的杂交瘤细胞之后,经有限稀释法将克隆进行亚克隆并采用标准方法使之生长[Goding, 出处同上]。用于此目的适宜的培养基包括例如, Dulbecco's 改进的 Eagle's 培养基和 RPMI-1640 培养基。或者,杂交瘤细胞可以在哺乳动物体内以腹水形式生长。

使用常规免疫球蛋白纯化方法,例如蛋白 A-琼脂糖凝胶、羟基磷灰石层析法、凝胶电泳、透析或亲和层析,可以从培养基或腹水中分离或纯化由亚克隆分泌的单克隆抗体。

也可使用重组 DNA 方法生产单克隆抗体,例如在美国专利 4816567 中描述的方法。使用常规技术(例如,使用能够与编码鼠抗体重链和轻链的基因特异结合的寡核苷酸探针),很容易地分离编码本发明单克隆抗体的 DNA 并进行序列测定。本发明的杂交瘤细胞作为此 DNA 的优选来源。一旦分离出该 DNA,就可将其放入表达载体中,然后用该载体转染宿主细胞,所述宿主细胞如猴 COS 细胞、中国仓鼠卵巢(CHO)细胞或不产生免疫球蛋白的骨

髓瘤细胞,以便在重组宿主细胞中合成单克隆抗体。也可对 DNA 进行修饰,例如人重链和轻链恒定区的编码序列置换同源的鼠序列[美国专利 4816567; Morrison 等,出处同上],或者将非-免疫球蛋白多肽的全部或部分编码序列共价连接到免疫球蛋白的编码序列上。此类非-免疫球蛋白多肽可取代本发明抗体的恒定区,或可取代本发明抗体一个抗原结合部位的可变区,以产生嵌合二价抗体。

抗体可以是单价抗体。制备单价抗体的方法为本领域所已知。例如,有一个方法涉及免疫球蛋白轻链和修饰后重链的重组表达。通常在 Fc 区的任何位点将重链截短,从而防止重链交联。或者,用其它氨基酸残基代替相关的半胱氨酸残基或者使半胱氨酸残基缺失以防止交联。

体外方法也适于制备单价抗体。使用本领域常规技术可以完成抗体消化而产生抗体片段,特别是 Fab 片段。

3. 人抗体和人源化抗体

本发明抗-PRO10282 抗体进一步包括人源化抗体或人抗体。非-人(例如,鼠)的人源化抗体是嵌合免疫球蛋白、免疫球蛋白链或其片段(如 Fv, Fab, Fab', F(ab')₂ 或抗体的其他抗原结合亚序列),它们包含极少的非人免疫球蛋白序列。人源化抗体包括人免疫球蛋白(受者抗体)中受者互补决定区(CDR)残基被具有所需特异性、亲和力和性能的小鼠、大鼠、家兔等非人源物种抗体(供体抗体)的 CDR 残基所取代。在一些实例中,人免疫球蛋白的 Fv 框架区残基由相应的非人类残基所取代。人源化抗体也可包含受者抗体或供体 CDR 或框架序列中均不存在的残基。通常,人源化抗体基本上包括至少一个、通常两个可变区的全部,其中 CDR 的全部或基本上全部对应于非人免疫球蛋白的相应部分,而 FR 序列的全部或基本上全部是人免疫球蛋白共有序列。人源化抗体还任选包括免疫球蛋白恒定区(Fc),通常为人免疫球蛋白的恒定区的至少一部分。详见 Jones 等,自然,321: 522-525(1986); Riechmann 等,自然 332: 323-329(1988); 和 Presta, Curr. Op. Struct. Biol. 2: 593-596(1992)。

人源化非-人抗体的方法是本领域技术人员所已知的。通常,人源化抗体中已导入一或多个源自非人类的氨基酸残基。这些非人类氨基酸残基常称为“引进的”残基,它们通常来自“引进的”可变区。人源化过程基本如 Winter 及其同事(Jones 等, Nature, 321:522-525(1986); Riechmann 等, Nature , 332:323-327(1988); Verhoeyen 等, Science , 239:1534-1536(1988))所述,用

一个或多个 CDR 序列取代人类抗体的相应序列来进行。因此，这样的“人源化”抗体是嵌合抗体(美国专利 4816567)，其中完整人类可变区的很大一部分被非人类物种的相应序列取代。实践中，人源化抗体通常是人的抗体，其中一些 CDR 残基且可能有一些 FR 残基被啮齿类抗体中类似位点的残基取代。

使用本领域已知的各种技术可以制备人抗体，所述方法包括噬菌体展示文库[Hoogenboom 和 Winter, *J. Mol. Biol.*, 227: 381(1991); Marks 等, *J. Mol. Biol.*, 222: 581(1991)]。也可使用 Cole 等和 Boerner 等描述的技术制备人单克隆抗体[Cole 等, *Monoclonal Antibody and Cancer Therapy*, Alan R. Liss, (1985)第 77 页, 和 Boerner 等, *J. Immunol.*, 147(1): 86-95(1991)]。类似地，通过把人免疫球蛋白的基因座导入转基因动物例如鼠来生产人抗体，所述宿主动物的内源性免疫球蛋白基因已被部分或全部灭活。攻击后，观察到人抗体产生，这在各个方面均与在人体所见的极为相似，包括基因重排、装配和抗体所有组成成分。有关进展在例如下列专利和科学出版物中作了描述：美国专利 5545807; 5545806; 5569825; 5625126; 5633425; 5661016; Marks 等, *Bio/Technology* 10, 779-783(1992); Lonberg 等, *Nature* 368 856-859(1994); Morrison, *Nature* 368, 812-13(1994); Fishwild 等, *Nature Biotechnology* 14, 845-51(1996); Neuberger, *Nature Biotechnology* 14, 826(1996); Lonberg 和 Huszar, *Intern. Rev. Immunol.* 13 65-93(1995)。

4. 双特异性抗体

双特异性抗体是单克隆抗体，优选具有针对至少两种不同抗原的结合特异性的人抗体或人源化抗体。在本发明中，一种结合特异性针对 PRO10282，另一种针对任何其它抗原，优选针对细胞表面蛋白或受体或受体亚单位。

制备双特异性抗体的方法为本领域所已知。通常，双特异性抗体的重组产生方法，是基于两条免疫球蛋白重链/轻链对的共表达，所述这两条链具有不同特异性 (Millstein 等, *Nature*, 305: 537-539 (1983))。由于免疫球蛋白重链和轻链的随机分配，这些杂交瘤(细胞杂交瘤 (quadromas))产生 10 种不同抗体分子的潜在混合物，其中只有一种具有正确的双特异性结构。对所述正确分子的纯化通常通过亲和层析步骤来完成。类似的方法参见于 1993 年 5 月 13 日公开的 WO 93/08829 和 Traunecker 等, *EMBO J.*, 10: 3655-3659 (1991)。

可将具有所需结合特异性(抗体-抗原结合位点)的抗体可变区与免疫球蛋白恒定区序列融合。该融合优选与包含铰链区的至少一部分、CH2 及 CH3 区的免疫球蛋白重链恒定区融合。优选使含有轻链结合所需位点的第一重链恒定区(CH1)出现在至少一种融合中。可将编码免疫球蛋白重链融合体,以及必要时,编码免疫球蛋白轻链的 DNA 插入不同表达载体,共转染至适当宿主生物。产生双特异性抗体的进一步细节内容,参见例如 Suresh 等, *Methods in Enzymology*, 121: 210(1986)。

根据 WO 96/27011 所述的另一种方法,可改造一对抗体分子之间的界面,使得从重组细胞培养中获得的异源二聚体的百分比最大。优选的界面包括抗体恒定区 CH3 结构域的至少一部分。在该方法中,源于第一抗体分子界面上的一条或多条氨基酸小侧链被较大侧链(如酪氨酸或色氨酸)取代。与前述大侧链大小相同或相近的互补“沟”可通过将氨基酸大侧链用小侧链(如丙氨酸或苏氨酸)取代而在第二抗体分子的界面上形成。这使得异二聚体的产量比不想要的终产物如同源二聚体的高。

双特异性抗体可以制备全长抗体或抗体片段(如 $F(ab')_2$ 双特异性抗体)。从抗体片段制备双特异性抗体的技术已有文献描述。例如,双特异性抗体可利用化学连接制备。Brennan 等, *Science* 229: 81(1985)中描述了将完整抗体经蛋白水解制备 $F(ab')_2$ 片段的方法。这些片段在二硫醇复合剂亚砷酸钠存在时被还原,从而稳定相邻的巯基,并阻止分子间二硫键的形成。生成的 Fab' 片段被转化为硫硝基苯甲酸盐(TNB)衍生物。其中一种 Fab' -TNB 衍生物经巯基乙胺还原成 Fab' -硫醇,再与等分子量的其它 Fab' -TNB 衍生物混合形成双特异性抗体。如此产生的双特异性抗体可作为酶的选择性固相化中所用的试剂。

Fab' -片段可从大肠杆菌中直接回收并经化学偶联形成双特异性抗体。Shalaby 等, *J. Exp. Med.* 175: 217-225(1992)中描述了完全人源化双特异性抗体 $F(ab')_2$ 分子的产生。每一 Fab' 片段分别从大肠杆菌中分泌出来,体外直接化学偶联形成双特异性抗体。如此制备的双特异性抗体能与过度表达 ErbB2 受体的细胞和正常人 T 细胞结合,还能引发人类细胞毒性淋巴细胞对人乳腺癌靶的裂解活性。

直接从重组细胞培养中制备并分离双特异性抗体片段的多种技术也已有描述。例如,可用亮氨酸拉链制备双特异性抗体。Kostelny 等, *J. Immunol.*,

148(5):1547-1553(1992))。将来自 Fos 和 Jun 蛋白的亮氨酸拉链肽与两种不同抗体的 Fab' 部分通过基因融合而连接。使抗体的同型二聚体在铰链区被还原成单体,然后被再氧化形成抗体的异二聚体。该方法也可用于制备抗体同型二聚体。由 Hollinger 等, Proc. Natl. Acad. Sci. USA, 90:6444-6448(1993)) 描述的“二价抗体”技术提供了另一种制备双特异性抗体片段的方法。所述片段中含有重链可变区(V_H),其通过接头与轻链可变区(V_L)相连,该接头非常短,使得同一链的两个结构域之间无法配对。因此,同一片段上的 V_H 和 V_L 结构域被迫与另一片段上的互补 V_L 和 V_H 结构域配对,从而形成两个抗原结合位点。此外还报道了另一种用单链 Fv (sFv)二聚体来制备双特异性抗体的策略。见 Gruber 等, J. Immunol., 152: 5368(1994)。还考虑了二价以上的抗体。如可制备三特异性抗体。Tutt 等, J. Immunol., 147: 60(1991)。

例证性双特异性抗体可与本文指定的 PRO10282 多肽上的两种不同表位结合。或者,可将抗-PRO10282 多肽臂与白细胞上触发分子的臂结合,从而强化针对表达特定 PRO10282 多肽细胞的细胞防御机制,所述引发分子如 T 细胞受体分子(CD2、CD3、CD28 或 B7),或 IgG Fc 受体($Fc\gamma R$)如 $Fc\gamma R$ I (CD64)、 $Fc\gamma R$ II (CD32)和 $Fc\gamma R$ III (CD16)。双特异性抗体还可用于将细胞毒制剂定位至表达特定 PRO10282 多肽的细胞上。这些抗体具有 PRO10282-结合臂和结合细胞毒制剂或放射性核素螯合剂的臂,如 EOTUBE、DPTA、DOTA 或 TETA。另一种感兴趣的双特异性抗体结合 PRO10282 多肽并且进一步结合组织因子(TF)。

5. 异源偶联抗体

异源偶联抗体也包括在本发明范围之内。异源偶联的抗体由两个共价连接的抗体组成。例如,有观点认为,这类抗体可用于将免疫细胞导向不想要的细胞(美国专利 4676980),并可用于治疗 HIV 感染[WO91/00360, WO92/200373, EP03089]。可考虑使用蛋白化学合成中已知方法,包括涉及使用交联剂,在体外制备异源偶联抗体。举例来说,使用二硫化物交换反应或通过形成硫醚键,可以构建免疫毒素。用于此目的适当试剂的实例,包括亚氨基硫醇盐和甲基-4-巯基,巯基 butyrimidate 以及例如在美国专利 4676980 中描述的那些。

6. 效应物功能的工程改造(Effector Function Engineering)

在效应物功能方面可能希望改进本发明的抗体,以增强在治疗(例如治

疗癌症)时的效果。例如,可在Fc区域中引入半胱氨酸残基,从而在该区域中形成链间二硫键。这样制得的同型二聚体可能有改进的内化能力和/或提高的补体介导细胞杀伤能力及依赖于抗体的细胞毒作用(ADCC)。参见 Caron 等, J. Exp Med. 176:1191-1195(1992) 和 Shopes, J. Immunol. 148:2918-2922(1992)。也可如 Wolff 等, Cancer Research 53:2560-2565(1993) 描述的那样,用异二聚功能性交联剂来制备抗肿瘤活性增强的同型二聚抗体。或者,可将抗体经工程化改造成具有双Fc区,从而可能增强补体裂解和ADCC能力。见 Stevenson 等, Anti-Cancer Drug Design 3:219-230(1989)。

7. 免疫偶联物

本发明还涉及免疫偶联物,其中含有与细胞毒制剂偶联的抗体,所述细胞毒药物如化疗剂、毒素(例如小分子毒素或细菌、真菌、植物或动物来源的酶活性毒素,或其片段)或放射性同位素(即放射偶联物)。

适用于制备这类免疫偶联物的化疗剂在上文已作描述。可以应用的酶活性毒素及其片段包括:白喉毒素A链、白喉毒素的非结合活性片段、外毒素A链(来自铜绿假单胞菌)、蓖麻毒蛋白A链、相思豆毒蛋白A链、蒴莲根毒素A链、 α -帚曲毒素、油桐(*Aleutites fordii*)蛋白、石竹素蛋白、美洲商陆(*Phytolacca Americana*)蛋白(PAPI, PAPII, PAP-S)、苦瓜(*Momordica charantia*)抑制因子、麻疯树毒蛋白、巴豆毒蛋白、肥皂草(*Saponaire officinalis*)抑制剂,白树毒素,米托菌素(mitogellin)、局限曲菌素、酚霉素、依诺霉素和单端孢菌毒素(tricothecenes)。多种放射性同位素可用于制备放射性偶联的抗体,实例包括 Bi^{212} 、 I^{131} 、 In^{131} 、 Y^{90} 和 Re^{186} 。

抗体与细胞毒制剂的偶联物可通过多种双功能蛋白偶联剂来连接,所述双功能蛋白偶联剂如:N-琥珀酰亚氨基-3-(2-吡啶基二巯基)丙酸酯(SPDP),琥珀酰亚氨基-4-(N-马来酰亚氨基甲基)环己烷-1-羧酸酯,亚氨基硫烷(iminothiolane)(IT),亚氨酸酯的双功能衍生物(如亚氨基己二酸二甲酯盐酸盐),活性酯类(如二琥珀酰亚胺基辛二酸酯),醛类(如戊二醛(glutareldehyde)),双-叠氮化合物(如双(对-叠氮基苯甲酰基)己二胺),双-重氮衍生物(如双-(对-重氮苯甲酰基)-乙二胺),二异氰酸酯(如亚甲代苯基2,6-二异氰酸酯),和双-活性氟化合物(如1,5-二氟-2,4-二硝基苯)。例如,蓖麻毒蛋白免疫毒素可如 Vitetta 等, Science, 238: 1098(1987)所述制备。 C^{14} 标记的1-异硫氰酸苯甲基-3-甲基二亚乙基三氨基五乙酸酯(MX-DTPA)是将放射性核苷酸偶联至抗

体的偶联剂之一。见 WO94/11026。

在另一实施方案中，抗体可与肿瘤预靶向中应用的“受体”(如链霉亲和素)偶联，将该抗体-受体偶联物施用于患者，之后用清除剂除去循环中未结合的偶联物，再给予已偶联了细胞毒制剂(如放射性核苷酸)的“配体”(如抗生物素蛋白)。

8. 免疫脂质体

本文公开的抗体还可制成免疫脂质体。含抗体的脂质体可通过本领域已知方法制备，如 Epstein 等，Proc. Natl. Acad. Sci. USA, 82: 3688(1985); Hwang 等，Proc. Natl. Acad. Sci. USA, 77: 4030(1980); 美国专利 4485045 和 4544545 及 1997 年 10 月 23 日公开的 WO97/38731。在美国专利 5013566 中公开了循环时间已增加了的脂质体。

特别有用的脂质体可利用包含磷脂酰胆碱、胆固醇和 PEG 衍生的磷脂酰乙醇胺(PEG-PE)的脂质组合物经反相蒸发法而制得。脂质体挤压通过限定孔径的膜，获得具有所需直径的脂质体。本发明抗体的 Fab' 片段可如 Martin 等，生物学化学杂志，257: 286-288(1982)所述的那样，经二硫化物交换反应与脂质体偶联。可任选在所述脂质体中包含一种化疗剂(如阿霉素)。见 Gabizon 等，J. National Cancer Inst, 81(19)1484(1989)。

9. 抗体-依赖性酶介导前体药物治疗法 (ADEPT)

通过将抗体与使前体药物(例如，肽基化疗剂，见 WO 81101145)转化为活性抗癌药物的药物前体-激活酶偶联，本发明抗体也可用于 ADEPT 至。参见例如，WO 88/07378 和美国专利 4 975 278。

对于 ADEPT 免疫偶联有用的酶组分，包括任何能够以使前体药物转化为其更具活性、细胞毒性形式的方式作用于前体药物的酶。

在本发明方法中有用的酶，包括但不限于糖苷酶，葡萄糖氧化酶，人溶菌酶，人葡萄糖醛酸糖苷酶，适于使含磷酸盐的前体药物转化为游离药物的碱性磷酸酶；适于使含有硫酸盐的前体药物转化为游离药物的酰基硫酸酯酶(arylsulfatase)；适于使无毒 5-氟胞嘧啶转化为抗癌药 5-氟尿嘧啶的胞嘧啶脱氨酶；适于使含肽前体药物转化为游离药物的蛋白酶，如沙雷氏菌属蛋白酶(serratia protease)，嗜热菌蛋白酶，枯草杆菌蛋白酶，羧肽酶(例如，羧肽酶 G2 和羧肽酶 A) 以及组织蛋白酶(如组织蛋白酶 B 和 L)；适于使含有 D-氨基酸取代基的前体药物转化的 D-丙氨酰羧肽酶；适于使糖基化前体药物转化为

游离药物的碳水化合物-裂解酶如半乳糖苷酶和神经氨酸苷酶；适于使内酰胺化衍生的药物转化为游离药物的内酰胺酶；和适于使在其胺氮位上苯氧基乙酰基或苯乙酰基衍生的药物分别转化成游离药物的盘尼西林酰胺酶，如盘尼西林 V 酰胺酶或盘尼西林 G 酰胺酶。或者，被本领域称为"抗体酶"具有酶活性的抗体也可用于使本发明前体药物转化为游离活性药物(参见例如，Massey, Nature, 328 : 457-458 (1987))。可按照本发明的描述，制备用于将抗体酶递送到肿瘤细胞群的抗体-抗体酶偶联物。

使用本领域熟知的技术，如使用前面讨论过的异源双功能交联剂，可将本发明酶与抗-PRO10282 (抗-Stra6)抗体共价结合。或者，使用本领域众所周知的重组 DNA 技术，可构建包括本发明抗体的至少一种抗原结合区与本发明酶的至少一个功能活性部分连接的融合蛋白(参见例如，Neuberger 等，Nature, 312 : 604-608 (1984))。

10. 抗体的药物组合物

与本文鉴定的 PRO10282 多肽特异性结合的抗体，以及用前述公开的筛选试验鉴定的其它分子，可以药学组合物的形式施用来治疗各种病症。

如果 PRO10282 多肽是在细胞内并且使用完整抗体作为抑制剂，则优选内化抗体。然而，也可用脂质转染物或脂质体将抗体或抗体片段递送至细胞内。使用抗体片段时，优选与靶蛋白结合结构域特异结合的最小抑制剂片段。例如，基于抗体可变区序列，可把肽分子设计为保留与靶蛋白序列结合的能力。此类肽可以化学合成和/或用重组 DNA 技术产生。参见例如 Marasco 等，Proc. Natl. Acad. Sci. USA, 90: 7889-7893(1993)。这里所述制剂也包含一种以上对于所治特定适应症而言必要的活性化合物，优选具有彼此之间无不利影响的互补活性的那些化合物。或者或另外，组合物可含有增强其功能的试剂，诸如细胞毒制剂、细胞因子、化疗剂或生长抑制剂。这些分子适宜以对所治疗目的而言有效的量联合。

也可将活性组分包被在按照例如凝聚技术或界面聚合技术制备的微胶囊中，例如，分别在胶态药物递送系统(例如脂质体，白蛋白微球体，微乳剂，纳米颗粒和纳米胶囊)或粗滴乳化剂中的羟甲基纤维素或明胶微胶囊和聚-(甲基丙烯酸甲酯)微胶囊中。这些技术公开在 Remington's Pharmaceutical Sciences，出处同上中。

用于体内给药的制剂必须是无菌的。这可以通过除菌膜过滤而轻易实

现。

也可制备控释制剂。控释制剂的适当实例包括含有抗体的固态疏水聚合物的半通透性基质，所述基质为具有一定形状的制品，如膜或微胶囊。控释制剂实例包括聚酯、水凝胶(如聚(2-羟基乙基-异丁烯酸酯)或聚(乙烯醇)，聚交酯(美国专利 3773919)，L-谷氨酸与 γ 乙基-L-谷氨酸酯的共聚物，不可降解的乙烯乙酸乙酯，可降解的乳酸-羟基乙酸共聚物如 LUPRON DEPOTTM(由乳酸-羟基乙酸共聚物和酸酐亮丙瑞林(leuprolide)组成的可注射的微球体)，以及聚 D-(-)-3-羟丁酸。尽管聚合物如乙烯-乙酸乙烯酯和乳酸-羟基乙酸能持续释放分子 100 天以上，但是某些水凝胶释放蛋白的时间却较短。当胶囊化的抗体长时间停留在体内时，它们会由于暴露在 37℃ 水分下而变性或凝聚，从而导致生物活性损失，且免疫原性可能会改变。可以根据涉及的机理来设计稳定化的合理策略。例如，如果发现凝聚的机理是通过硫代二硫键互换而形成了分子间 S-S 键，则可通过修饰巯基残基、自酸性溶液中冻干、控制湿度、采用合适的添加剂和开发特殊的聚合物基质组合物来达到稳定化。

O. 抗-PRO10282 (抗-Stra6)抗体的用途

本发明抗-PRO10282 抗体具有多种用途。例如，抗-PRO10282 抗体可用于 PRO10282 的诊断分析，譬如，检测 PRO10282 在特定细胞、组织或血清的表达。可使用本领域已知的各种诊断分析技术，如以异源或同源形式(phase)进行的竞争性结合试验、直接和间接夹心试验，及免疫沉淀试验。Zola, 单克隆抗体：技术指南，147-158 页(CRC Press, Inc. 1987)。还可将所述抗体标记上可检测部分。可检测部分应当能够直接或间接产生可测得信号。例如，可测得部分是放射性同位素如 ³H、¹⁴C、³²P、³⁵S 或 ¹²⁵I，荧光或化学发光化合物如异硫氰酸荧光素、罗丹明或萤光素，或者酶如碱性磷酸酶、 β -半乳糖苷酶或辣根过氧化物酶。本领域已知的将抗体与可检测部分偶联的任何方法均可使用，包括在 Hunter 等, Nature, 144: 945(1962); David 等, Biochemistry, 13: 1014(1974); Pain 等, J. Immunol. Meth., 40: 219(1981); 和 Nygren, J. Histochem.和 Cytochem., 30: 407(1982)中描述的方法。

抗-PRO10282 抗体也适用于从重组细胞培养物或天然来源中亲和纯化 PRO10282。在该过程中，使用本领域已知方法将抗 PRO10282 抗体固定在适当载体如 Sephadex 树脂或滤纸上。然后，使固定的抗体与含有欲纯化的 PRO10282 的样品接触，接下来用能够基本上将样品中与固定载体结合的

PRO10282 之外的所有物质去除的适宜溶剂洗涤载体。最后, 用能从抗体中释出 PRO10282 的另一适宜溶剂洗涤载体。

抗-PRO10282 抗体可用于监测 PRO10282 多肽的表达, 所述多肽具有作为筛选抗广泛多种维生素 A 类应答性疾病的 治疗有效化合物的敏感方法的潜在用途。在这些类型疾病中, 考虑到作为治疗靶的疾病包括牛皮癣, 痤疮, 发育异常, 癌和自身免疫性疾病。考虑到的发育异常的种类包括上皮组织的癌前病变如口腔粘膜白斑病, 子宫颈发育异常, 喉和支气管发育异常。考虑到的癌的种类包括结肠癌(腺癌), 肺癌(癌和腺癌), 皮肤癌, 头颈部癌, 子宫颈癌, 子宫癌, 乳腺癌和前列腺癌。考虑到的自身免疫性疾病的种类包括类风湿病性关节炎, 骨关节炎, 系统性红斑狼疮, 寻常性天疱疮和落叶状天疱疮。在治疗此类疾病和病症中, 抗-PRO10282 抗体具有潜在的治疗用途。

P. 使用抗-PRO10282 (Stra6) 抗体和其它 Stra6 拮抗剂治疗的方法

考虑到本发明抗体和其它抗-肿瘤化合物可用于治疗多种病症, 包括那些以本文确定的扩增基因的过度表达和/或激活为特征的病症。用此抗体和其它化合物(包括但不限于: 小的有机和无机分子, 肽, 反义分子等)治疗的例证性病症包括良性或恶性肿瘤(例如, 肾的癌, 肝癌, 肾癌, 膀胱癌, 乳腺癌, 胃癌, 卵巢癌, 结肠直肠癌, 前列腺癌, 胰腺癌, 肺癌, 外阴癌, 甲状腺癌, 肝脏癌; 肉瘤; 成胶质细胞瘤; 和各种头颈肿瘤); 白血病和恶性淋巴瘤; 其它如神经细胞瘤, 神经胶质瘤, 星形胶质细胞瘤, 下丘脑和其它腺癌, 巨噬细胞癌, 上皮细胞癌, 基质癌和囊胚腔病症; 以及炎症, 血管生成和免疫紊乱。用本发明抗-Stra6 抗体和其它拮抗剂治疗的特别优选的靶, 是在 Wnt-1 通路中有基因缺陷的肿瘤(例如导致 β -连环蛋白信号异常活化)和/或过度表达 Stra6 的肿瘤。业已发现, 在 Wnt-1 通路中有基因缺陷的人类癌症通常也表现出 Stra6 过度表达, 但并未知道所有 Stra6 过度表达肿瘤均与 Wnt-1 通路上的突变有关。本发明 Stra6 拮抗剂具有治疗 Stra6 过度表达肿瘤的巨大潜力。优选此组肿瘤包括结肠直肠肿瘤, 卵巢癌, 子宫内膜癌, 和 Wilm's 肾脏肿瘤, 黑素瘤, 以及嗜铬细胞瘤(一种衍发自肾上腺髓质的肿瘤)。

按照已知方法, 如经静脉内快速浓注或连续输注一段时间, 或者经肌肉注射、腹膜内注射、脑脊柱内注射(intracerebrospinal)、皮下注射、关节内注射、滑膜腔内注射、鞘内注射、口服、局部施用或吸入途径给药, 将本发明抗-肿瘤剂如抗体施用于哺乳动物, 特别是人, 抗体的静脉给药为优选。

其它治疗方案可与抗癌剂(如本发明抗体)的给药进行联合。例如,接受本发明抗癌剂治疗的患者,也可接受放射治疗。或者,将其它化疗剂施用于患者。此类化疗剂的配制和给药方案,按照制造商的指示进行或者由熟练的临床医师根据经验决定。化疗剂的配制和给药方案在 Chemotherapy Service Ed., M. C. Perry, Williams & Wilkins, Baltimore, MD (1992) 中也有描述。化疗剂可以在抗肿瘤剂如抗体之前或之后给药,或者二者同时给药。此外,抗体可与抗-雌激素化合物如他莫昔芬或抗-孕激素如奥那司酮(见 EP 616812)按照这些分子的已知用量联合给药。

也期望施用抗其它肿瘤相关抗原的抗体,如将会结合 CD20、Cd11a、CD18、ErbB2、EGFR、ErbB3、ErbB4 或血管内皮因子(VEGF)的抗体。或者,或另外,本文公开的结合同一个或二个或多个不同抗原的二个或多个抗体,可联合施用于患者。有时,将一种或多种细胞因子也施用于患者是非常有益的。一个实施方案中,本发明多肽与生长抑制剂联合给药。例如,首先施用生长抑制剂,接下来施用本发明多肽。然而,也考虑了同时施用或首先施用本发明多肽。生长抑制剂的适宜剂量是目前所用剂量,同时由于生长抑制剂和本发明多肽的联合作用(协同作用)而可以降低用量。

为了预防或治疗疾病,本发明抗肿瘤剂如抗体的适宜剂量取决于所治疾病的类型(如上所述)、疾病的严重程度和病程(无论给药目的是预防还是治疗)、先前的治疗、患者的临床病史和对抗肿瘤剂的敏感度以及主治医生的判断。抗肿瘤剂适宜于一次性地或在一系列治疗中施用于患者。

例如,取决于受治疾病的类型和严重程度,施用于患者的抗体的初始候选剂量是 $1\text{ }\mu\text{g/kg}$ ~ 15 mg/kg (如 $0.1\text{--}20\text{ mg/kg}$), 无论是例如通过一次或分多次给药,还是通过连续输注给药。依据上述因素,代表性的每日剂量为约 $1\text{ }\mu\text{g/kg}$ ~ 100 mg/kg 或更多。重复给药数天或更长时间时,视具体情况而将治疗持续至出现对疾病症状的所需抑制为止。然而,其它剂量方案也可能是有用的。治疗的进展很容易通过常规技术和试验分析加以监测。

Q. 制品

在本发明的另一个实施方案中,提供了含有用于诊断或治疗上述疾病的物质的制品。所述制品包括一个容器和一个标签。适当的容器包括瓶子,小瓶,注射器等。容器可由各种材料如玻璃或塑料制成。该容器内盛放能有效诊断或治疗病症的组合物,并具有无菌存取口(例如,该容器可以是静脉注

射袋或带有能通过皮下注射针刺穿的塞子的小瓶)。组合物中的活性剂通常是能够干扰本文确定的基因产物之活性的抗肿瘤剂,如抗体。容器上的或附带的标签说明该组合物用于诊断或治疗所选出的病症。制品还包括第二个容器,其中含可药用的缓冲液,例如磷酸盐缓冲液, Ringer's 溶液和葡萄糖溶液。制品还可进一步包括具有商业需要以及符合用户需要的其它材料,如其它缓冲液、稀释剂,滤器,针头和注射器。

R. 肿瘤的诊断和预后

细胞表面蛋白,如在一定肿瘤中过度表达的生长受体,是肿瘤(如癌)治疗候选药物的极好的靶,相同蛋白质与由肿瘤细胞中扩增的基因编码的分泌蛋白质一起,具有在肿瘤的诊断和预后中的另一用途。譬如,直接针对在肿瘤中增强的基因的蛋白质产物的抗体,可用作肿瘤诊断剂或预后剂。

举例来说,抗体(包括抗体片段),可用于定性或定量测定由扩增的基因或过度表达基因("标记基因产物")编码的蛋白质的表达。抗体优选连接有可检测标记物,例如荧光素标记,并且可用光学显微镜、流式细胞仪、荧光测定仪或本领域已知的其它技术监测抗体结合。如果扩增的基因编码细胞表面蛋白,如生长因子,那么这些技术是特别适合的。此结合分析基本上按照前面第5部分的描述进行。

可通过例如免疫荧光或免疫电子显微镜检查法,完成在原位检测抗体与标志基因产物的结合。为此目的,自患者采出组织学样品,将标记抗体用在样品上,优选将抗体覆盖在生物样品上。此过程也可以测定标志基因产物在被测组织的分布。关于原位检测,大量的各种组织学方法是可以利用的,此点对于本领域技术人员而言是显而易见的。

下述实施例只是为了描述的目的而提供,并非旨在以任何方式限定本发明的保护范围。

实施例

本说明书中引证的所有专利和参考文献,以其全部在此引入作为参考。

实施例中涉及的市售制剂按照生产商的使用说明应用,除非另有说明。以下实施例中,及说明书全文中,用 ATCC 保藏号所指定的那些细胞来源于美国典型培养物保藏中心, Manassas, VA。在以下实施例中,除非另有说明, "Sra6"指天然序列 PRO10282 多肽。

实施例 1

编码人 PRO10282 和 PRO19578 多肽的 cDNA 克隆的分离

采用酵母筛选来鉴定人类胎脑 cDNA 文库中编码天然人 PRO10282 和 PRO19578 多肽的 cDNA (DNA148380-2827 和 DNA148389-2827-1), 所述文库优先代表最初 cDNA 克隆的 5' 末端。

DNA148380-2827 克隆包含单个开放读框, 49-51 位核苷酸是一个明显的(apparent)翻译起始位点, 2050-2052 位核苷酸是中止密码子, 该开放读框中止于此(图 1)。所预测的多肽前体的长度是 667 个氨基酸(图 2)。估测图 2 所示 PRO10282 蛋白全长的分子量约 73502 道尔顿且 pI 约为 9.26。对图 2(SEQ ID NO : 2)所示 PRO10282 序列全长的分析证实其存在多种重要的多肽结构域, 如图 2 所示, 其中, 这些重要的多肽结构域的位置基本如上所述。DNA148380-2827 克隆于 2000 年 1 月 11 日保藏在 ATCC, ATCC 保藏号为 PTA-1181。

DNA148389-2827-1 克隆包含单个开放读框, 186-188 位核苷酸是一个明显的翻译起始位点, 2160-2162 位核苷酸是中止密码子, 该开放读框中止于此(图 6, SEQ ID NO : 4)。预测的多肽前体的长度为 658 个氨基酸(图 7, SEQ ID NO : 5)。图 7 所示 PRO19578 蛋白的全长的估计分子量约为 72583 道尔顿, pI 约为 9.36。对图 7 所示的 PRO19578 序列全长(SEQ ID NO : 5)的分析证实其存在多种重要的多肽结构域, 如图 7 所示, 其中, 这些重要的多肽结构域的位置基本如上所述。值得注意的是存在 9 个潜在的跨膜结构域以及在人和相应的小鼠序列之间保守的 14 个半胱氨酸残基。小鼠 Stra6 有 3 个潜在的 N 连接的糖基化位点, 而人 PRO19578(天然的人 Stra6)多肽有一个糖基化位点。148389-2827-1 克隆在 2000 年 2 月 23 日保藏在 ATCC, ATCC 保藏号为 PTA-1405。

利用对图 2(SEQ ID NO : 2)所示全长序列和图 7 (SEQ ID NO : 5)所示全长序列的 ALIGN 2 序列比对来分析 Dayhoff 数据库 (版本 35.45 SwissProt 35), 结果表明 PRO10282 氨基酸序列和以下 Dayhoff 序列: AF062476, P-W88559 和 HGS-RE259 之间存在序列同一性。

如图 8 所示, 对人 PRO10282 和 PRO19578 多肽全长的比较显示 PRO19578 在 PRO10282 氨基酸序列的 89-97 位包含 9 个氨基酸 (SPVDFLAGD; SEQ ID NO : 13)的缺失。另外, PRO19578 在 518 位包含异

亮氨酸(I), 其取代了 PRO10282 相应位置(527 位)上的蛋氨酸(M), 这是由该位置的 G/A 多态性所致。PRO10282 和天然序列 PRO19578 多肽两者都被认为是小鼠 Stra6 多肽的人同系物, 因而它们也被称为"Stra6"。小鼠 Stra6 和由 DNA148340-2827 编码的天然序列人类全长 PRO10282 多肽有约 74% 的氨基酸序列同一性。

图 9 显示由 DNA148340-2827 编码的天然序列人类全长 PRO10282 多肽(简称"人 Stra6")的疏水图谱。如图 9 所示, 在该长度为 667 个氨基酸的多肽中 50% 的氨基酸残基是疏水性的。

通过 UNIGENE 确定人 Stra6 基因位于染色体 15q23。初步精细作图显示, Stra6 位于 STS 间隔 D15 S124-D15 S160 中, GeneMap 的第 98 位对应于 G3 图谱中的 244.52。

实施例 2

PRO10282 和 PRO19578 作为杂交探针的应用

以下方法描述编码 PRO10282 和 PRO19578 的核苷酸序列作为杂交探针的应用

包含全长或成熟 PRO10282 或 PRO19578 编码序列的 DNA 用作探针在人类组织 cDNA 文库或人类组织基因组文库中筛选同源 DNA(如编码天然存在的 PRO10282 或 PRO19578 变体的那些 DNA)。

对包含上述任一文库 DNA 的膜的杂交和洗涤是在以下高严谨条件下进行的。放射性标记的 PRO10282 来源的探针与膜的杂交在 42℃, 50% 甲酰胺, 5x SSC, 0.1% SDS, 0.1% 焦磷酸钠, 50 mM 磷酸钠, pH 6.8, 2x Denhardt's 溶液, 和 10% 葡聚糖硫酸酯中进行 20 小时。膜的洗地在 0.1xSSC 和 0.1% SDS 的水溶液中 42℃ 下进行。

采用本领域已知的标准技术对与编码全长天然序列 PRO10282 或 PRO19578 的 DNA 有所需同一性的 DNA 进行鉴定。

实施例 3

PRO10282 和 PRO19578 在大肠杆菌中的表达

该实施例描述了 PRO10282 或 PRO19578d 的非糖基化形式通过大肠杆菌中的重组表达所进行的制备。

首先采用所选择的 PCR 引物对编码 PRO10282 或 PRO19578 的 DNA 序列进行扩增。所述引物应该包含与所选择的表达载体上限制性酶位点相应

的限制性酶位点。可以使用各种表达载体。适合载体的一个例子是 pBR322 (来源于大肠杆菌; 参见 Bolivar 等, Gene, 2 : 95 (1977)), 其包含氨苄青霉素和四环素抗性基因。将所述载体用限制性酶消化并且去磷酸化。然后将上述 PCR 扩增序列与所述载体连接。所述载体优选包括编码抗生素抗性基因, trp 启动子, 多组氨酸前导序列(包括最开始的 6 个 STII 密码子, 多组氨酸序列, 和肠激酶切割位点), PRO10282 或 PRO19578 编码区, λ 转录中止子, 和 argU 基因的序列。

然后采用 Sambrook 等, 出处同上所描述的方法用上述所得的连接混合物转化所选择的大肠杆菌菌株。通过转化子在 LB 平板上的生长能力对其进行鉴定且将抗生素抗性菌落选择出来。对质粒 DNA 进行分离, 并通过限制酶切分析和 DNA 测序鉴定。

在补充有抗生素的液体培养基如 LB 肉汤培养基上过夜培养所选择的克隆。随后将此过夜培养物用于接种较大规模的培养。然后使细胞生长到所需光密度, 在此期间表达启动子被激发。

再将细胞培养几个小时后, 离心收获细胞。可用本领域已知的各种制剂溶解离心所获得的细胞团, 然后, 被溶解的 PRO10282 蛋白在允许该蛋白紧密结合的条件下采用金属螯合柱进行纯化。

PRO10282 或 PRO19578 可以以带有多组氨酸标签的形式在大肠杆菌中表达, 表达步骤如下。首先采用所选择的 PCR 引物对编码 PRO10282 或 PRO19578 的 DNA 进行扩增。所述引物包含对应于所选表达载体上限制性酶切位点的限制性酶切位点, 以及提供有效和可靠的翻译起始, 在金属螯合柱上的快速纯化和肠激酶蛋白水解去除所需的其它有用序列。然后将 PCR 扩增的多组氨酸标记的序列连接到表达载体上, 将该载体用于转化基于菌株 52 (W3110 fuhA(tonA)lon galE rpoHts (htpRts)clpP(lacIq)的大肠杆菌宿主。使转化子首先生长在包含 50mg/ml 羧苄青霉素的 LB 中, 30℃振荡培养直至 O.D.600 达到 3-5。用 CRAP 培养基(3.57g (NH₄)₂SO₄, 0.71g 柠檬酸钠 2H₂O, 1.07g KCl, 5.36g Difco 酵母提取物, 5.36g Sheffield hycase SF 溶于 500 mL 水中, 和 110mM MPOS, pH7.3, 0.55 % (w/v) 葡萄糖, 7mM MgSO₄ 混合而成)将培养物稀释 50-100 倍, 然后 30℃下振荡培养大约 20-30 小时。取出样品通过 SDS-PAGE 验证表达, 将大部分培养物离心以沉淀细胞。将所得细胞团冷冻直至纯化和重折叠(refolding)。

将 0.5 到 1 L 发酵(6-10g 细胞团)所得大肠杆菌以 10 倍体积(W/V)重新悬浮在 7M 胍, 20 mM Tris, pH 8 缓冲液中。向该溶液加入固体亚硫酸钠和连四硫酸钠至各自终浓度为 0.1 M 和 0.02 M, 4℃搅拌过夜。该步骤导致蛋白变性, 其中所有的半胱氨酸残基通过亚硫酸化作用被封闭。将上述溶液以 40, 000rpm 在 Beckman 超速离心机中离心 30 分钟。所得上清液用 3-5 个体积的金属螯合柱缓冲液(6M 胍, 20 mM Tris, pH 7.4)稀释, 并通过 0.22 微米膜过滤澄清。将澄清的提取物上样至已在金属螯合柱缓冲液中平衡过的 5 ml Qiagen Ni-NTA 金属螯合柱上。用包含 50 mM 咪唑(Calbiochem, Utrol grade), pH7.4 的另一种缓冲液洗涤该层析柱。用包含 250 mM 咪唑的缓冲液洗脱蛋白。汇集包含所需蛋白的级分并 4℃储存。用根据所述蛋白的氨基酸序列所计算的消光系数, 通过 280 nm 吸光度估算蛋白浓度。

将样品缓慢加入新鲜配制的重折叠缓冲液中进行稀释, 以使蛋白进行重折叠, 所述缓冲液由以下组成: 20mM Tris, pH8.6, 0.3M NaCl, 2.5M 尿素, 5mM 半胱氨酸, 20mM 甘氨酸和 1mM EDTA。选择重折叠的体积使最终的蛋白浓度在 50-100 μ g/ml 之间。4℃下轻轻搅拌重折叠溶液 12-36 小时。加入 TFA 至终浓度为 0.4% (pH 大约是 3)来中止重折叠反应。对蛋白进一步纯化之前, 用 0.22 微米膜过滤上述溶液并加入乙腈使其终浓度为 2%-10%。将重折叠蛋白在 Poros R1/H 反向层析柱上层析, 所用流动缓冲液为 0.1% TFA, 用从 10% 到 80% 的乙腈梯度进行洗脱。对具有 A280 吸光度的各级分取等份试样在 SDS 聚丙烯酰胺凝胶上进行分析, 汇集包含均一重折叠蛋白的级分。通常大多数蛋白中正确重折叠的蛋白在最低乙腈浓度时被洗脱, 因为它们最为紧密, 且它们疏水的内部受到保护而不与反向树脂相互作用。聚集的蛋白通常可用较高浓度的乙腈洗脱。除了可以区分蛋白的所需形式和错误折叠形式外, 反向步骤也可以从样品中除去内毒素。

汇集包含所需折叠的 PRO10282 或 PRO 19578 多肽的级分并用直接引入溶液的柔和氮流除去乙腈。将所述蛋白配制在包含 0.14M 氯化钠和 4% 甘露醇的 20mM Hepes, pH6.8 的溶液中透析或采用在配制缓冲液中平衡的 G25 Superfine (Pharmacia)树脂进行凝胶过滤, 并进行过滤灭菌。

具体地, 天然人 Stra6 蛋白 PRO10282 的两个胞外结构域(ECD)肽 A (氨基酸 229-295)和肽 B(氨基酸 532-667)分别在大肠杆菌细胞质中表达为具有 N 末端多组氨酸前导序列的肽, 其具有氨基酸序列

MKHQHQHQHQHQHQM HQ (SEQ ID NO : 12)。该前导序列提供最佳翻译起始，在镍螯合柱上的纯化，以及必要时用 TAGZyme 系统(Unizyme Laboratories)进行的有效去除。转录通过大肠杆菌碱性磷酸酶启动子(Kikuchi 等, Nucleic Acids Res. 9 : 5671-5678 [1981]) 控制，提供 trp 操纵子核糖体结合位点(Yanofsky 等, Nucleic Acids Res.9:6647-6668[1981]) 用于翻译。翻译中止密码子下游是 λ t0 转录中止子(Scholtissek 和 Grosse, Nucleic Acids Res.15:3185[1987])，其后是稀有密码子 tRNA 基因 pro2, argU, 和 glyT (Komine 等, J. Mol. Biol. 212: 579-598 [1990]; Fournier 和 Ozeki, Microbiol. Rev. 49: 379397 [1985])。

通过 PCR 从全长 cDNA 克隆中制备 Stra6 的两个 ECD 编码序列 DNA 片段，并将其插入到上述称为 pST239 的表达载体中。验证了 DNA 序列后，用称为 PE148380A 和 PE148380B 的新的 Stra6 表达质粒转化大肠杆菌菌株 58F3 ((fhuA Δ (tonA Δ) lon Δ galE rpoHts (htpRts) Δ clpP lacIq Δ ompT Δ (nmpc-fepE) Δ slyD)。这些转化子的 Luria 肉汤培养物首先 30℃ 生长过夜，然后用磷酸盐限制性培养基将其稀释 100 倍以诱导碱性磷酸酶启动子。30℃ 振荡培养 24 小时后，离心所述培养物，并将所得细胞团冰冻直至开始肽的纯化。

进行纯化时需将大肠杆菌菌团(6-10 gm 菌团)在 10 倍(w/v)体积的 7 M 盐酸胍，20mM Tris, pH8 缓冲液中重新悬浮。加入固体的亚硫酸钠和连四硫酸钠使终浓度分别为 0.1M 和 0.02M，4℃ 下搅拌该溶液过夜。通过离心使该溶液澄清，然后将澄清的溶液上样至已在 6M 盐酸胍，20mM Tris, pH 7 中平衡的 Qiagen Ni-NTA 金属螯合柱上。用另一种包含 50mM 咪唑 (Calbiochem, Utrol grade)的缓冲液洗柱。用包含 250mM 咪唑的缓冲液洗脱所述蛋白。汇集包含所需蛋白的级分，对 1mM HCl 透析并 4℃ 储存。

实施例 4

PRO10282 和 PRO19578 在哺乳动物细胞中的表达

该实施例描述了通过哺乳动物细胞中的重组表达制备潜在糖基化形式的 PRO10282 和 PRO19578

载体 pRK5 (参见 EP 307, 247, 1989 年 3 月 15 日公开)用作表达载体。

采用如 Sambrook 等, 出处同上所述的连接方法, 用允许 PRO10282 DNA 任选用所选限制性酶将 PRO10282 或 PRO19578 DNA 连接到 pRK5 中, 以便

用诸如 Sambrook 等, 出处同上所述连接方法插入 PRO 10282 DNA。所产生的载体各自称为 pRK5-PRO10282 和 pRK5-PRO19578。

在一个实施方案中, 所选择的宿主细胞可以是 293 细胞。人 293 细胞 (ATCC CCL 1573) 在组织培养板中生长至汇合, 所用培养基如补充有胎牛血清和任选, 营养成分和/或抗生素的 DMEM。将大约 10g pRK5-PRO10282 或 pRK5-PRO19578 DNA 与约 1g 编码 VA RNA 基因的 DNA [Thimmappaya 等, Cell, 31:543(1982)] 混合并溶解在 500L 1mM Tris-HCl, 0.1mM EDTA, 0.227M CaCl_2 中。向该混合物中逐滴加入 500L 50mM HEPES (pH 7.35), 280mM NaCl, 1.5mM NaPO_4 , 25℃ 10 分钟使沉淀形成。悬浮沉淀物并将其加入到 293 细胞中, 然后 37℃ 放置约 4 小时。吸出培养基并在 30 秒钟内加入 2ml 20% 甘油的 PBS 溶液。然后用无血清的培养基洗涤 293 细胞, 再加入新鲜培养基温育细胞约 5 天。

转染后大约 24 小时, 移走培养基并更换为培养基(单独)或包含 200Ci/ml³⁵S-半胱氨酸和 200Ci/ml³⁵S-蛋氨酸的培养基。12 小时温育后, 收集该条件(conditioned)培养基, 在旋转过滤器上浓缩并上样至 15% SDS 凝胶上。可以将处理后的凝胶进行干燥, 并暴露于胶片一段选定的时间, 以显示 PRO10282 多肽的存在。可将包含转染细胞的培养物进一步温育(在无血清的培养基中)并用所选择的生物检测方法对该培养基进行检测。

在一种备选技术中, 采用 Sompayrac 等, Proc. Natl. Acad. Sci., 12: 7575 (1981) 描述的硫酸葡聚糖方法可以将 PRO10282 或 PRO19578 瞬时导入 293 细胞中。293 细胞在旋转瓶中生长到最大密度后加入 700 g pRK5-PRO10282 或 pRK5-PRO19578 DNA。首先通过离心从旋转瓶中浓缩细胞并用 PBS 洗细胞。将 DNA-葡聚糖沉淀物在细胞团上温育 4 小时。所述细胞用 20% 甘油处理 90 秒, 然后用组织培养基洗涤, 再加入包含组织培养基, 5g/ml 牛胰岛素和 0.1g/ml 牛运铁蛋白的旋转瓶中。大约 4 天之后, 离心该条件培养基并过滤以除去细胞和碎片。浓缩包含表达的 PRO10282 或 PRO19578 的样品并通过任何选定方法如透析和/或柱层析纯化。

在另一实施方案中, PRO10282 或 PRO19578 可在 CHO 细胞中表达。用公知试剂如 CaPO_4 或 DEAE-葡聚糖可将 pRK5-PRO10282 或 pRK5PRO19578 转染到 CHO 细胞中。如上所述, 温育细胞培养物, 将培养基更换为用培养基(单独的)或包含放射性标记物如 ³⁵S-蛋氨酸的培养基。确

定存在 PRO10282 或 PRO19578 多肽之后, 可以将培养基更换为无血清培养基。优选, 将培养物温育大约 6 天, 然后收获该条件培养基。包含表达的 PRO10282 或 PRO19578 的培养基通过任何选定方法浓缩并纯化。

带有表位标签的 PRO10282 或 PRO19578 也可以在宿主 CHO 细胞中表达。PRO10282 或 PRO19578 也可以从 pRK5 载体中亚克隆出来。亚克隆插入子经 PCR 与所选表位标签如多组氨酸标签融合在杆状病毒表达载体的同一阅读框中。然后可将带有多组氨酸标签的 PRO10282 或 PRO19578 插入子亚克隆到 SV40 驱动载体中, 所述驱动载体包含用于选择稳定克隆的选择标志如 DHFR。最后, 用该 SV40 驱动载体转染 CHO 细胞(如上所述)。可如上所述进行标记来验证表达。然后浓缩包含表达的多组氨酸标记型 PRO10282 或 PRO19578 的培养基并通过任何选定方法如通过 Ni^{2+} -螯合亲和层析进行纯化。

PRO10282 或 PRO19578 也可以通过瞬时表达方法在 CHO 和/或 COS 细胞中表达或通过另一稳定表达方法在 CHO 细胞中表达。

采用以下步骤进行 CHO 细胞中的稳定表达。所述各种蛋白表达为 IgG 构建体(免疫粘附素), 其中每种蛋白的可溶性形式(例如细胞外结构域)的编码序列与 IgG1 恒定区序列融合, 所述恒定区序列包含铰链区, CH2 和 CH2 结构域, 和/或所述蛋白是带有多组氨酸标签的形式。

PCR 扩增之后, 采用 Ausubel 等, Current Protocols of Molecular Biology, Unit 3. 16, John Wiley 和 Sons (1997)所述的标准技术将每种 DNA 亚克隆到 CHO 表达载体中。构建 CHO 表达载体, 使目的 DNA 的 5' 和 3' 端含有相容的限制性酶切位点, 以准许 cDNA 便利地穿梭。用于在 CHO 细胞中表达的载体在 Lucas 等, Nucl. Acids Res. 24 : 9 (1774-1779 (1996)中有描述, 且使用 SV40 早期启动子/增强子驱动目的 cDNA 和二氢叶酸还原酶(DHFR)的表达。DHFR 的表达可以用来选择转染后稳定维持的质粒。

采用市售转染试剂 Superfect® (Quiagen), Dospere® 或 Fugene® (Boehringer Mannheim)将 12 μg 的所需质粒 DNA 导入到大约 1×10^7 个 CHO 细胞中。所述细胞如 lucas 等, 出处同上所述生长。将大约 3×10^7 个细胞冷冻在安瓶瓶中以备进一步培养和如下所述的制备。

将包含质粒 DNA 的安瓶瓶放入水浴中解冻, 通过涡旋将其混合。将内容物移到包含 10ml 培养基的离心管中, 1000rpm, 离心 5 分钟。吸出上清, 并

把细胞重悬浮在 10ml 的选择培养基中(0.2 μm 过滤的 PS20, 其含有 5% 0.2 μm 已渗滤的胎牛血清)。将细胞等分到 100ml 的含有 90ml 选择培养基的旋转器(spinner)中。1-2 天后, 将细胞转移到 250ml 的装有 150ml 选择培养基的旋转器中, 37°C 温育。再过 2-3 天后, 以 3×10^5 细胞/ml 将细胞接种到 250ml, 500ml 和 2000ml 的旋转器中。通过在制备培养基中离心和重悬浮将细胞培养基更换为新鲜培养基。尽管可以使用任何适当的 CHO 培养基, 但实际使用的是 1992 年 6 月 16 日授予的美国专利 5122469 中描述的制备培养基。以 1.2×10^6 细胞/ml 接种 3L 的制备旋转器。第 0 天时, 测定细胞数和 pH 值。第一天时, 从旋转器中取样并开始用过滤的空气喷雾。第二天, 从旋转器中取样, 将温度转变到 33°C, 加入 30ml 500g/l 葡萄糖和 0.6ml 10% 消泡剂(例如 35% 的聚二甲基硅氧烷乳剂, Dow Corning 365 医疗级乳剂)。整个生产过程中, 依需要调整 pH 使其保持在 7.2 左右。10 天后, 或存活率降低到 70% 以下时, 通过离心收获细胞培养物并经过 0.22 μm 滤膜过滤。所得过滤物或者 4°C 储存或者立即上样至纯化柱上。

对于多组氨酸标记的构建体, 用 Ni-NTA 层析柱(Qiagen)纯化蛋白。在纯化之前, 向条件培养基中加入咪唑至 5mM 浓度。4°C 下将条件培养基以 4-5 ml/min 的流速泵到 6ml Ni-NTA 层析柱上, 所述层析柱已用含 0.3M NaCl 和 5mM 咪唑的 20mM Hepes pH 7.4 缓冲液平衡。上样后, 用另外的平衡缓冲液洗柱, 用包含 0.25M 咪唑的平衡缓冲液洗脱蛋白。然后用 25 ml G25 Superfine (Pharmacia) 层析柱使高度纯化的蛋白脱盐至含 10 mM Hepes, 0.14M NaCl 和 4% 甘露醇, pH6.8 的储存缓冲液中, -80°C 储存。

免疫粘附素(包含 Fc)构建体从条件培养基中如下进行纯化。将条件培养基泵到 5ml 的蛋白 A 层析柱上(Pharmacia)上, 所述层析柱已经 20mM 磷酸钠缓冲液 pH6.8 平衡。上样后, 用平衡缓冲液充分洗柱, 之后用 100mM 柠檬酸 pH3.5 洗脱。将 1ml 的级分收集到包含 275 μl 1M Tris 缓冲液 pH9 的管中, 使洗脱的蛋白立即得到中和。然后将高度纯化的蛋白脱盐至上述用于多组氨酸标记蛋白的储存缓冲液中。通过 SDS 聚丙烯酰胺凝胶以及利用 Edman 降解法进行的 N 末端氨基酸测序来评估均一性。

实施例 5

PRO10282 和 PRO19578 在酵母中的表达

以下方法描述 PRO10282 和 PRO19578 在酵母中的重组表达。

首先, 构建用于细胞内生产或分泌由 ADH2/GAPDH 启动子控制的 PRO10282 多肽的酵母表达载体。将编码 PRO10282 多肽(包括 PRO19578)的 DNA 和该启动子插入到选定质粒的适当限制性酶切位点上, 以指导 PRO10282 的细胞内表达。对于分泌的情况, 可将编码 PRO10282 的 DNA, 连同编码 ADH2/GAPDH 启动子, 天然 PRO10282 信号肽或其它哺乳动物信号肽, 或例如酵母 α 因子或转化酶分泌信号/前导序列, 或(必要时)用于 PRO10282 表达的接头序列克隆到该选定质粒中。

使用上述表达质粒转化酵母细胞, 如酵母菌株 AB110, 并在选定的发酵培养基中培养。采用 10% 三氯醋酸沉淀对转化的酵母上清液进行分析, 经过 SDS-PAGE 及之后的考马氏兰染色凝胶进行分离。

然后通过离心去除发酵培养基中的酵母细胞, 来实现对重组 PRO10282(包括 PRO19578)的分离和纯化, 并用选定的柱式过滤器浓缩该培养基。采用选定的柱层析树脂对包含 PRO10282(例如 PRO19578)的浓缩物进一步纯化。

实施例 6

PRO10282 和 PRO19578 在杆状病毒感染的昆虫细胞中的表达

以下方法描述 PRO10282 和 PRO19578 在杆状病毒感染的昆虫细胞中的重组表达。

编码 PRO10282 或 PRO19578 的序列被融合在杆状病毒表达载体中表位标记的上游。所述表位标记物包括多组氨酸标签和免疫粘附素标签(如 IgG 的 Fc 区)。多种质粒都可以应用, 包括由市售质粒如 pVL1393 (Novagen)衍生的质粒。简言之, 编码 PRO10282 或 PRO19578 的序列或 PRO10282 或 PRO19578 编码序列的所需部分, 如编码跨膜蛋白的细胞外结构域的序列, 或者, 如果所述蛋白是细胞外蛋白, 编码成熟蛋白的序列, 用与 5' 和 3' 区互补的引物进行 PCR 扩增。所述 5' 引物的旁侧可掺入(选定的)限制性酶切位点。然后将所得产物用那些选定的限制性酶消化并亚克隆到表达载体中。

用 lipofectin(购自 GIBCO-BRL)将上述质粒和 BaculoGold™ 病毒 DNA(PharMingen)共转染到草地夜蛾("Sf9")细胞(ATCC CRL 1711)中, 产生重组杆状病毒。28℃温育 4 到 5 天后, 收获所释放的病毒用于进一步扩增。

如 O' Reilley 等, Baculovirus expression vectors: A laboratory Manual, Oxford: Oxford University Press(1994)所述进行病毒的感染和蛋白的表达。

通过例如 Ni^{2+} 螯合物亲和层析如下对表达的多组氨酸标记的 PRO10282 或 PRO19578 进行纯化。按 Rupert 等, Nature, 362:175-179(1993)描述的从重组病毒感染的 Sf9 细胞制备提取物。简言之, 洗涤 Sf9 细胞, 然后将其重悬浮在超声处理缓冲液中(25ml Hepes, pH7.9; 12mM MgCl_2 ; 0.1mM EDTA; 10% 甘油; 0.1% NP-40; 0.4M KCL), 在冰上超声处理 2 次 20 秒钟。离心澄清超声处理物, 然后将所得上清液在上样缓冲液(50mM 磷酸盐, 300mM NaCl, 10% 甘油, pH7.8)中稀释 50 倍, 并用 $0.45\ \mu\text{m}$ 的膜过滤。制备 5ml 柱床体积的 Ni^{2+} -NTA 琼脂糖层析柱(购自 Qiagen), 用 25ml 的水洗柱, 并用 25ml 的上样缓冲液平衡。将过滤后的细胞提取物以每分钟 0.5ml 上样至层析柱上。用上样缓冲液洗所述层析柱至 A_{280} 基线, 此时开始收集级分。然后, 用第二种洗柱缓冲液(50mM 磷酸盐, 300mM NaCl, 10% 甘油, pH6.0)洗柱, 该缓冲液洗脱非特异性结合的蛋白。重新回到 A_{280} 基线后, 用所述第二种缓冲液中 0 到 500mM 的咪唑梯度过柱。收集 1ml 级分, 通过 SDS-PAGE 和银染或用 Ni^{2+} -NTA-偶联的碱性磷酸酶(Qiagen)经 Western 印迹分析。汇集包含洗脱的 His_{10} -标记型 PRO10282 或 PRO19578 的级分并使其对上样缓冲液进行透析。

备选地, 用已知的层析技术, 包括例如, 蛋白 A 或蛋白 G 柱层析对 IgG 标记的(Fc 标记的) PRO10282 或 PRO19578 进行纯化。

实施例 7

结合 PRO10282 或 PRO19578 的抗体的制备

本实施例描述了能特异性结合 PRO10282 或 PRO19578 的单克隆抗体的制备。

单克隆抗体的制备技术在本领域是公知的并在, 例如, Goding, 出处同上中有描述。可应用的免疫原包括纯化的 PRO10282 和 PRO19578, 包含 PRO10282 或 PRO19578 的融合蛋白, 以及在细胞表面表达重组 PRO10282 或 PRO19578 的细胞。本领域的技术人员无需实验即可对免疫原作出选择。

用乳化在完全弗氏佐剂中的 PRO10282 或 PRO19578 以 1-100 μg 的量皮下或皮内注射来免疫小鼠, 如 Balb/c 小鼠。或者, 用 MPL-TDM 佐剂(Ribi Immunochemical Research, Hamilton, MT)乳化免疫原并注射到所述动物的

后足垫中。10 到 12 天后用乳化在选定佐剂中的免疫原对免疫过的小鼠进行加强免疫。此后几周，也可以再进行免疫注射以加强免疫。定期通过眼眶后取血获得小鼠的血清样品，用 ELISA 检测法检测抗 PRO010282 抗体或抗 PRO019578 抗体。

检测到适当的抗体滴度后，对抗体“阳性”的动物最后静脉内注射 PRO10282 或 PRO19578。3 到 4 天后，处死小鼠，取其脾细胞。然后将所得脾细胞与选定的鼠骨髓瘤细胞系如 P3X63AgU.1(ATCC, No. CRL, 1597 融合(用 35 % 聚乙二醇)。然后将融合后产生的杂交瘤细胞平铺到包含 HAT(次黄嘌呤，氨基蝶呤和胸苷)培养基的 96 孔组织培养板中，所述培养基用于抑制未融合细胞，骨髓瘤杂合体 and 脾细胞杂合体增殖。

用 ELISA 筛选杂交瘤细胞抗 PRO10282 或 PRO19578 的反应性。对分泌所需抗 PRO10282 或 PRO19578 的单克隆抗体的“阳性”杂交瘤细胞的确定方法在本领域是公知的。

将阳性杂交瘤细胞腹膜内注射到同系 Balb/c 小鼠中，产生包含抗 PRO10282 或 PRO19578 单克隆抗体的腹水。或者，使杂交瘤细胞在组织培养瓶或旋转瓶中生长。用硫酸铵沉淀和随后的凝胶排阻层析对产生在腹水中的单克隆抗体进行纯化。或者应用基于抗体与蛋白 A 或蛋白 G 结合的亲和层析。

具体地，用纯化的 Unizyme 偶联型氨基酸肽加 Ribi 佐剂(Ribi Immunochem Reseach, Inc., Hamilton, MO)超免疫 5 只 Balb/c 小鼠(Charles River Laboratories, Wilmington, DE)，所述氨基酸肽相当于人 Stra6 的 532-667 位氨基酸。如前所述(Hongo 等, Hybridoma 14: 253-260 [1995])使来自腭淋巴结的 B 细胞与小鼠骨髓瘤细胞(X63.Ag8.653; 美国典型培养物保藏中心, Rockville, MD)融合。10-14 天后，收集上清，用直接酶联免疫吸附试验(ELISA)对抗体的产生进行筛选。通过有限稀释进行两轮亚克隆后，发现 8 个经直接 ELISA 和免疫组化检测免疫结合性最高的阳性克隆，将这些克隆注射给 Pristine 致敏的(primed)小鼠以体内制备 mAb(Freud 和 Blair, J. Immunol. 129: 2826-2830 [1982])。汇集腹水液并如前所述(Hongo 等, 出处同上)通过蛋白 A 亲和层析(Pharmacia 快速蛋白质液相层析(FPLC); Pharmacia, Uppsala, Sweden)对其进行纯化。纯化的抗体制品过滤(0.2 μ m 孔径; Nalgene, Rochester, NY)除菌，并在磷酸盐缓冲液(PBS)中 4℃ 储存。

实施例 8

用特异性抗体纯化 PRO10282 和 PRO19578 多肽

可通过蛋白纯化领域中各种标准技术对天然或重组的 PRO10282 和 PRO19578 多肽进行纯化。例如，用特异于目标 PRO10282 多肽的抗体通过免疫亲和层析对原 PRO10282 或原 PRO19578 多肽，成熟的 PRO10282 或 PRO19578 多肽，或前 PRO10282 或前 PRO19578 多肽进行纯化。通常，通过使抗 PRO10282 或抗 PRO19578 多肽抗体与一种活化的层析树脂共价偶联来构建免疫亲和层析柱。

通过硫酸铵沉淀或固相蛋白 A(Pharmacia LKB Biotechnology, Piscataway, N.J.)纯化从免疫血清制备多克隆免疫球蛋白。类似的，通过硫酸铵沉淀或固相蛋白 A 层析从小鼠腹水液制备单克隆抗体。

将部分纯化的免疫球蛋白共价偶联到层析树脂如 CnBr-活化的 SEPHAROSE™(Pharmacia LKB Biotechnology)上。使所述抗体与树脂偶联，所述树脂被封闭，并依照生产商的使用说明洗涤所得的衍生树脂。

利用所述免疫亲和层析柱可纯化 PRO10282 或 PRO19578 多肽，其通过从细胞制备包含可溶形式的 PRO10282 或 PRO19578 多肽的组分。该制备可通过将完整细胞或亚细胞组分溶解来实现，所述亚细胞组分可通过加入去污剂然后差速离心而获得或通过本领域公知的其它方法获得。或者，包含信号序列的可溶性 PRO10282 或 PRO19578 多肽可以以有效量被分泌到细胞生长的培养基中。

使包含可溶性 PRO10282 或 PRO19578 多肽的制品流经免疫亲和层析柱，并在允许 PRO10282 或 PRO19578 多肽优先吸附的条件下(例如，在有去污剂的存在的高离子强度缓冲液)洗柱。然后在破坏抗体/PRO10282 或 PRO19578 多肽结合的条件下(例如，低 pH 值如 pH 值约 2-3 的缓冲液或高浓度的离液剂如尿素或硫氰酸离子)洗脱层析柱，并收集 PRO10282 或 PRO19578 多肽。

实施例 9

药物筛选

本发明尤其用于应用 PRO10282(Strat6)多肽(包括 PRO19578)或其结合片段以多种药物筛选技术中的任何一种筛选化合物。在所述检测中应用的 PRO10282 多肽或片段可以是在溶液中游离的，附着在固相支持物上的，细

胞表面携带的，或者位于细胞内。药物筛选的一个方法利用已被表达 PRO10282 多肽或片段的重组核酸稳定转化的真核或原核宿主细胞。在竞争结合试验中筛选针对这种转化细胞的药物。这些细胞，无论是存活的还是固定形式的都可以用于标准结合试验。可以检测，例如，PRO10282 多肽或其片段与待测制剂之间复合物的形成。或者，可以检测由待测制剂所导致的 PRO10282 多肽与其靶细胞或靶受体之间复合物形成的减少。

所以，本发明提供筛选能影响 PRO10282(Stra6)多肽相关疾病或紊乱的药物或其它任何制剂的方法。这些方法包含使所述制剂与 Stra6 多肽或其片段接触，并通过本领域公知的方法检测(i)所述制剂和 Stra6 多肽或其片段之间复合物的存在，或(ii) Stra6 多肽或其片段和细胞之间复合物的存在。在这样的竞争结合试验中，Stra6 多肽或其片段通常被标记。适当温育后，游离的 Stra6 多肽或其片段与其结合形式分离，这样，游离或未形成复合物的标记物的量便成为特定制剂结合 Stra6 多肽或干扰 Stra6 多肽/细胞复合物的能力的一个指标。

药物筛选的另一技术可用于高通量筛选对多肽具有适当结合亲和力的化合物，该技术可参见 1984 年 9 月 13 日公开的 WO 84/03564。简言之，在固体基质上，如塑料针或其它一些表面上合成大量不同的小肽待测化合物。如应用于 PRO10282(Stra6)多肽的，使多肽待测化合物与 Stra6 多肽反应并进行洗涤。结合的 Stra6 多肽通过本领域公知的方法检测。也可将纯化的 Stra6 多肽直接包被在用于前述药物筛选技术的平板上。另外，非中和性抗体能用来捕获肽并将其固定到固相支持物上。

本发明也涉及竞争性药物筛选试验的应用，其中能结合 Stra6 多肽的中和抗体与检测化合物特异性竞争结合 Stra6 多肽或其片段。这样，抗体可用于检测与 Stra6 多肽有一个或多个共同抗原决定簇的任何肽的存在。

实施例 10

合理的药物设计

合理的药物设计的目标是生产所需生物活性多肽(即 PRO10282 (Stra6)多肽)的结构类似物或与所述多肽相互作用的小分子，例如，激动剂，拮抗剂或抑制物。这些实例中的任何一个都能用于设计天然序列 PRO10282 或 PRO19578 多肽更具活性或更稳定形式的药物，或加强或干扰天然序列 PRO10282 或 PRO19578 多肽的体内功能的药物(c.f., Hodgson,

Bio/Technology, 9:19-21(1991))。

通过 X 线晶体衍射, 计算机模型设计或通常, 通过这两种方法的组合, 可以确定天然序列 PRO10282 或 PRO19578 多肽, 或 PRO10282 或 PRO19578 多肽-抑制物复合物的三维结构。必须确定天然 PRO10282 或 PRO19578 多肽的形状和电荷, 以阐明该分子的结构并确定其活性位点。有时, 根据同源蛋白的结构设计模型可获得关于 PRO10282 或 PRO19578 多肽结构的有用信息。在上述两种情况下, 相关的结构信息可用于设计类似的 PRO10282 多肽样分子或确定其有效的抑制物。合理药物设计的有效实例包括活性或稳定性已经改善的那些分子, 如, Braxton 和 Wells, Biochemistry, 31:7796-7801(1992) 所述, 或作为天然肽的抑制物, 激动剂或拮抗剂起作用的那些分子, 如 Athauda 等, J. Biochem., 113: 742-746 (1993) 所描述。

也可以对如上所述功能分析选定的靶特异性抗体进行分离, 然后揭示其晶体结构。该方法原则上产生一种药核(pharmacore), 在此基础上可以设计其后续的药物。通过针对功能性药物活性抗体产生抗独特型抗体(抗-ids), 有可能避开蛋白的结晶学分析。作为镜像的镜像, 可推测出抗-ids 的结合位点是最初受体的类似物。那么抗-id 可用于从化学或生物学方法制备的肽库中鉴定和分离肽。那么这种分离的肽可用作药核。

根据本发明, 可提供足量的 PRO10282 多肽(包括 PRO19578), 用来进行这种分析研究, 如 X 线晶体学研究。另外, 本文提供的 PRO19578 多肽氨基酸序列的知识可对应用计算机模拟技术替代 X 线晶体学或除 X 线晶体学外还应用计算机模拟技术的那些人提供指导。

实施例 11

组织表达分布

从编码 PRO10282 多肽的核苷酸序列构建用于定量 PCR 扩增反应的寡核苷酸探针, 所述核苷酸序列见附图。选择寡核苷酸探针, 使得在标准 PCR 反应中从与其结合的模板 3' 端产生约 200-600 个碱基对的扩增片段。所述寡核苷酸探针用于在标准定量 PCR 反应中, 与分离自不同成人组织的 Clontech RNA 反应, 并通过琼脂糖凝胶电泳进行分析, 以定量 PRO10282 多肽-编码核酸在各种待测组织中的表达水平。对 PRO10282 多肽-编码核酸在各种不同人体组织中的表达模式或差异表达的了解提供了一种诊断标志, 该标志在有或无其它组织特异性标志物的情况下都可以进行组织分型, 用于判

断转移性肿瘤的原始组织来源, 等等。这些分析的结果(显示在图 10)表明, DNA148380-2827 分子在成人肾, 睾丸和子宫中高表达; 在乳腺, 前列腺和气管中明显表达; 在脑, 心脏, 肺和甲状腺中微弱表达; 在肝脏, 骨髓, 结肠, 骨骼肌, 小肠, 脾和胃中不表达。

总 RNA 购自 Clontech (Palo Alto, California) 并采用以下引物/探针进行 PCR 扩增分析:

hStra6.tmf3:5' CACACTCGAGAGCCAGATATTTT (SEQ ID NO:6)

hStra6.tmr4:5' AACAAAGTTTATTGCAGGGAACAC (SEQ ID NO:7)

hStra6.tmp4:5' TGTAGTTTTTATGCCTTTGGCTATTATGAAAGAGGT
(SEQ ID NO:8)

tmf=正向引物

tmr=反向引物

tmp=探针

如以下实施例 12 描述的原位杂交确定了 Stra6 在肾小管上皮细胞, 子宫肌层, 乳腺导管和乳腺小叶周围的间质细胞中的表达, 但在脑, 肝脏, 脾, 胰腺, 心脏, 肺, 胃, 小肠, 结肠, 前列腺, 脾和肾上腺皮质的切片中 Stra6 极少表达或无表达(数据未显示)。原位杂交检测的正常组织中, 胎盘和肾上腺髓质可见最高水平的表达, 在 PCR 分析中并非如此。

实施例 12

人肿瘤中天然人 PRO10282(Stra6)转录物的过表达

该实施例显示编码天然人全长 PRO10282(Stra6)的基因在某些人结肠肿瘤中明显过表达, 并在来源于各种人体肿瘤如结肠, 肺, 肾和乳腺肿瘤的细胞系中也是如此。

由于筛选的起始物质是分离自人结肠肿瘤或人结肠, 肾, 乳腺或肺的各种肿瘤细胞系的总 RNA。在结肠肿瘤组织中, 测定了 Stra6 RNA 相对于同一患者正常结肠组织(粘膜)的 RNA 的表达。Stra6 RNA 在各种肿瘤细胞系中的表达与各种正常细胞系(即, 正常结肠, 肾和肺细胞系)相对比确定。

用 Tagman 分析试剂, 通过实时(real-time)定量 PCR(RT-PCR, 例如, TAQMAN ABI PRIZM 7700™ Sequence Detection System™[Perkin Elmer, Applied Biosystems Division, Foster City, California]), 监测正常细胞中和来源于某些癌或癌细胞系的细胞中 PRO10282(Stra6)编码基因(相当于

DNA148380-2827)表达量的差异。50 μ l RT-PCR 反应的组成为 5 μ l 10 $\times$ Taqman 缓冲液 A, 每种 dNTP 各 300 μ M, 5mM MgCl₂, 10 个单位的 RNA 酶抑制物, 12.5 个单位的 MnLV 反转录酶, 1.25 个单位的 AmpliTaq Gold DNA 聚合酶, 200nM 探针, 500nM 引物和 100ng RNA。反应条件是 48 $^{\circ}$ C 反转录 30 分钟, 95 $^{\circ}$ C 变性 25 秒和 65 $^{\circ}$ C 1 分钟。反应产物在 4-20% 的聚丙烯酰胺凝胶(Novex)上分析。

对每个所分析的样品用标准曲线确定每个目的基因和甘油醛-3-磷酸脱氢酶(GAPDH)管家基因的相对表达水平。用目的基因 mRNA 水平除以 GAPDH mRNA 水平获得相对校正单位。对比试验样品和对照样品之间的相对校正单位以确定诱导倍数。

上述结果用于确定编码 PRO10282 的 mRNA 在任一种原发性结肠癌或被筛选的结肠, 肾脏, 乳腺或肺癌细胞系中是否过表达。用于该分析的一些匹配的人正常和结肠肿瘤样品的组织学情况显示如下 (参见图 11 和 12)。

编号	组织学
850	侵袭性腺癌; 无坏死; 保存完好。
851	侵袭性腺癌; 极少量坏死; 状态完好。
892	侵袭性腺癌; 极少量坏死; 状态完好。
869	侵袭性腺癌; 极少量坏死; 状态完好。
893	正常粘膜-发育不良-侵袭性腺癌; 少量坏死; 状态完好。
870	腺癌-严重发育不良; 极少量坏死; 状态完好。
871	腺癌-发育不良-正常粘膜; 无坏死; 状态完好。
848	腺癌-似乎起源于绒毛状腺瘤; 粘膜/粘膜下正常; 无坏死, 状态完好。
872	侵袭性腺癌; 大约 70% 的肿瘤坏死; 整体保存完好。
778	腺癌-稀有乳头状形态; 粘膜正常/增生; 极少量坏死; 可接受的保存。
17	中度良好分化的腺癌。
18	中度良好分化的腺癌。
64	中度良好分化的腺癌。
76	中度良好分化的腺癌。

人肺细胞系包括正常人肺成纤维细胞系 MRCS (CCL-171)和 MRC5 (CCL-186), 人肺癌上皮细胞系 A549 (SRCC768, CCL-185), 人上皮样肺癌细胞系 Calu-1 (SRCC769; HTB-54), 人未分化的癌细胞系 Calu-6 (SRCC770, HTB-56 可能来自肺), 人上皮细胞系 NCI-H441 (SRCC772; HTB-174)其来源于肺乳头状腺癌患者的心包液, 和人肺鳞状细胞癌细胞系 SW900 (SRCC775; HTB-59), 这些都可从 ATCC 获得。

结肠细胞系包括, 例如正常结肠成纤维细胞系 CCD112Co (CRL-1541), 人结肠直肠腺癌细胞系 CaCo-2 (HTB-37), 人结肠直肠腺癌细胞系 WiDr (CCL-218), 人结肠直肠腺癌细胞系 HCT116 (CCL-247), 人结肠直肠腺癌细胞系 SK-Co1 (HTB-39), 人结肠直肠腺癌细胞系 COLO320 (SRCC778; CCL-220), 人结肠直肠腺癌细胞系 HT29 (SRCC779; HTB-38), 人结肠直肠腺癌细胞系 SW403 (CCL-230), 人结肠癌细胞系 NCI/HCC2998, 和人结肠直肠腺癌细胞系 Colo320DM (CCL-220), 所有这些都来自 ATCC 或其它的公共渠道。

人乳腺癌细胞系包括人乳腺癌细胞系 MCF7 (SRCC766; HTB-22), 和人乳腺癌细胞系 NCI/ADR-RES, 上述两种细胞系都可以从公共途径购买到。

肾细胞系包括 293 细胞系(CRL-1573), 其是用腺病毒 5 DNA 转化的。两种 Wilm' s 肿瘤细胞系, G401 (CRL-1441)和 SK-NEP-1 (CRL-1573)也包括在该分析中。

RNA 的制备:

从前述培养的细胞系制备 RNA。应用 Qiagen 的纯化试剂盒, 缓冲液组和蛋白酶, 依照生产商的使用说明和以下描述进行分离。更具体的, 用 Qiagen RNeasy 微型柱从培养的细胞中分离总 RNA。用 RNA Stat-60 (Tel-Test)分离来自组织样品的总 RNA。通过氯化铯密度梯度离心分离制备从肿瘤制备的 RNA。

固体的人肿瘤样品的制备和裂解:

对肿瘤样品称重, 并放入 50ml 的锥形管中, 冰上静置。限制每次制备仅处理不超过 250mg 组织(1 tip/一次制备)。蛋白酶溶液新鲜配制, 即将蛋白酶放入 6.25ml 冷双蒸水中稀释至终浓度为 20 μ g/ml, 并 4 $^{\circ}$ C 储存。将 DNase A 稀释至终浓度 200 μ g/ml(从 100mg/ml 原液)来制备 G2 缓冲液(20ml)。为避

免气溶胶的吸入，在层流 TC 罩超净台中应用 polytron 的大 tip 对 10ml G2 缓冲液中的肿瘤组织匀浆 60 秒，室温静置。每次取样之间，洗净 polytron，方法是每次在 2L 双蒸水，然后在 G2 缓冲液中(50ml)中以 2×30 秒旋转。如果制备器(generator) tip 仍然有组织存留，则将该设备拆开清洗。

加入 Qiagen 蛋白酶(如上制备的，1.0ml)，然后进行涡旋，并在 50℃ 温育 3 小时。重复温育和离心直至裂解物澄清(例如，另外温育 30-60 分钟，4℃ 以 3000xg，10 分钟离心沉淀)。

定量

从 RNA 的实时 PCR 分析中获得的结果首先表示为 ΔCT 单位。一个单位对应于一个 PCR 循环或相对于正常大约 2 倍的扩增，二个单位相当于 4 倍，三个单位相当于 8 倍扩增等等。将数据转化成倍数差并如此表述出来。首先，用逆转录酶并用寡(dT)作引物从 100ng 总 RNA 或 polyA+RNA 合成 cDNA。所产生的 cDNA 用作 PCR 模板。用来源于 PRO10282 编码序列 3' 非翻译区的引物和相当于各自间插序列的 TAQMAN™ 荧光探针进行定量。应用 3' 区倾向于避免基因组 DNA 中内含子-外显子边界的交叉，这是应用此方法准确评估 RNA 表达的一个必要条件。用于 PRO10282 编码基因扩增的引物和探针的序列(正向，反向，和探针)如下：

一组包括以上实施例 11 中描述的正向和反向引物和探针，其分别显示在 SEQ ID NO: 6, 7 和 8 中。

另一组包括：

hStra6.tmf1:5'AGACCAGGTCCCACACTGA (SEQ ID NO:9)

hStra6.tmr1:5'TTCATAATAGCCAAAGGCATAAAA (SEQ ID NO:10)

hStra6.tmp1:5'CTGCCCACACTCGAGAGCCAGAT3'(SEQ ID NO:11)

人 GAPDH:

正向引物: 5'GAAGATGGTGATGGGATTTC-3' (SEQ ID NO:14)

反向引物: 5'GAAGGTGAAGGTCGGAGTC-3' (SEQ ID NO:15)

探针: 5'CAAGCTTCCCGTTCTCAGCC-3' (SEQ ID NO:16)

5' 核酸酶分析反应是荧光 PCR 基础上的技术，该技术利用 TaqDNA 聚合酶的 5' 外切核酸酶活性实时监测扩增。用两种寡核苷酸引物产生 PCR 反应的典型扩增子。设计第三种寡核苷酸或探针以检测位于两个 PCR 引物之间的核苷酸序列。所述探针是不能通过 TaqDNA 聚合酶延伸的，用报告分子

荧光染料和淬灭剂荧光染料标记。当两种染料在探针上的位置彼此靠近时,任何激光诱导的报告分子染料的发射都被淬灭染料所淬灭。在扩增反应期间, Taq DNA 聚合酶以模板依赖的方式切割探针。所得探针片段在溶液中相互分散, 来自所释放的报告分子染料的信号不受第二种荧光团淬灭作用的影响。每合成一新分子, 便释放一分子的报告分子染料, 对未淬灭的报告分子染料的检测为数据的定量解释打下了基础。

如上所述, 5' 核酸酶在实时定量 PCR 仪如 ABI PRISM 7700™ Sequence Detection System™ 上起作用。所述系统由热循环仪, 激光, 电荷耦合装置 (CCD), 照相机和计算机组成。该系统在热循环仪上扩增 96 孔板中的样品。在扩增期间, 通过所有 96 孔的纤维光缆实时收集激光诱导的荧光信号, 并在 CCD 上进行检测。该系统包括运行该仪器的软件和分析所得数据的软件。

5' 核酸酶分析数据最初表示为 Ct, 或阈循环。将其定义为报告分子信号累积大于荧光的背景水平的循环。当癌细胞中 RNA 的表达与正常细胞 RNA 的表达相比较时, Ct 值被用作核酸样品中特定靶序列起始拷贝相对数的定量值。

结果

当与相应的正常人结肠组织比较时, 人结肠肿瘤组织中 Stra6 RNA(相当于 DNA 148380-2827)显示强烈过表达。如图 11 所示, 在所检查的所有 14 份人肿瘤结肠组织中发现 Stra6 RNA(相当于 DNA 148380-2827)相对于来自同一患者相匹配的正常结肠直肠粘膜 RNA 是过表达的。过表达在 2 到 170 倍之间, 并且 14 个肿瘤组织样品中 7 个过表达至少大约 10 倍。通过定量 PCR 获得的循环阈值显示, 在许多正常的粘膜样品中 Stra6 mRNA 的水平极低或可能不存在。

作为第二种检测方法, 定量 PCR 反应(每次 40 个循环)完成后所获得的产物进行聚丙烯酰胺凝胶电泳并通过溴乙啶染色显影。如图 12A 所示, 用管家基因甘油醛-3-磷酸脱氢酶(GAPDH)作标准, 肿瘤样品与它们的正常对应样品相比, 从 Stra6 mRNA 产生的 PCR 产物的量要大得多。为进行比较, 从所有样品得到内部对照 GAPDH mRNA 的产物的可比较水平。

通过如下进行的原位杂交(ISH)发现结肠腺癌中 Stra6 的表达位于上皮肿瘤细胞中。³²P 标记的正义和反义核酸探针转录自 874bp 的 PCR 产物, 该产物相当于由 DNA 148380-2827 编码的人 Stra6 多肽的编码序列中 432-1247

位核苷酸。如前所述(Pennica 等, Proc.Natl.Acad.Sci.USA 95:14717-22(1998))对甲醛固定且石蜡包埋的组织切片进行处理。所得结果显示在图 12B 中。

如图 13 所示, 在乳腺, 肾脏, 结肠和肺的各种肿瘤细胞系中人 Stra6 RNA(相当于 DNA 148380-2827)也是明显过表达的。“相对的 RNA 表达”指正常和肿瘤组织中 RNA 的表达显示为相对于任意选定的标准细胞系 SW480 中的表达。

从各种肿瘤组织切片获得的原位杂交结果也显示在图 16 中。除结肠腺癌外几种类型的肿瘤也显示高水平的 Stra6 表达。它们包括所有 3 例黑素瘤(图 16A 和 B), 4 例子宫内膜腺癌中的 3 例(图 16C 和 D), 3 例卵巢腺癌中的 2 例, 和肾脏的 Wilm' s 瘤(图 16E 和 F)。在这些不同的肿瘤中 Stra6 的原位杂交信号比结肠腺癌中的大得多, 这与所显示的在正常肾脏和子宫中相对较高表达水平和在正常结肠中的低水平一致。因为在正常肾上腺髓质中检测到 Stra6, 所以我们也检测了二例来源于该组织的肿瘤嗜铬细胞瘤。在这些肿瘤中, Stra6 的表达超过本研究中所检测的其它任何肿瘤或组织中的表达(图 16G 和 H)。尽管 Stra6 在正常肾脏中检测到且在 Wilm' s 肿瘤中强表达, 但在肾细胞癌中未检测到。在肾脏移行(transitional)细胞癌中, 肿瘤相关间质细胞而不是肿瘤上皮细胞表达 Stra6(数据未显示)。

因为在各种肿瘤和许多肿瘤来源的细胞系中, 编码 PRO10282 的 mRNA DNA148380-2827 过表达, 所以其很可能与肿瘤的形成和/或生长有关。因此, 预计直接针对由 DNA148380-2827(PRO10282)编码的蛋白或该蛋白其它天然产生的变体, 如由 DNA148389-2827-1 编码的 PRO19578 的拮抗剂(例如抗体, 有机和无机小分子, 肽和多肽, 如 Stra6 变体, 反义寡核苷酸)可用于诊断, 预防和/或治疗癌症尤其, 不予限制, 结肠, 肺, 乳腺和/或肾脏的癌症。

所述拮抗剂如直接针对由 DNA148380-2827(PRO10282)编码的蛋白的治疗性抗体的有效性可通过某些制剂来加强, 所述制剂可刺激编码 PRO10282 的基因的表达。例如, 用 9-顺-视黄酸或全-反视黄酸处理人结肠直肠癌细胞系导致 Stra6 mRNA 表达的明显增强(图 15)。因此, 预计用直接针对 PRO10282 的治疗性抗体与适当的维生素 A 类组合对癌症患者进行的治疗会提高所述抗体对肿瘤的杀伤。对直接针对在多种肿瘤中过表达的其它天然 Stra6 多肽, 如由 DNA148389-2827-1 编码的人剪接变体的抗体来说也是如此。

实施例 13

Wnt-1 和维生素 A 类对 Stra6 的协同诱导

材料与方法

细胞培养

C57MG 和 C57MG/Wnt-1 细胞生长在 Dulbecco' s 改良型 Eagle 培养基中, 该培养基补充有 10% 胎牛血清, 2mM L-谷氨酰胺和 2.5 μ g/ml 嘌呤霉素 (Edge Biosystems)。C57MG 细胞(该细胞中 Wnt-1 的表达可被四环素阻抑)生长在完全培养基中, 该培养基无嘌呤霉素, 补充有 400 μ g/ml G418 (Gibco BRL), 100 μ g/ml 潮霉素 B(Gibco BRL)和 50ng/ml 四环素(Korinek 等, Mol.Cell Biol.18:1248-56(1998))。对 Wnt-1 的诱导进行研究时, 所述细胞用磷酸盐缓冲液洗涤, 在无四环素的培养基中培养 10, 24, 48, 72 和 96 小时, 然后收获。0 小时的对照培养皿一直维持在含四环素的培养基中。同时收获所有培养皿中的细胞并提取每一时间点的总 RNA。用 Wnt-1, Stra6 和 GAPDH 特异性引物和探针对来自每一样品的 100ng 总 RNA 进行 RT-PCR。

从美国典型培养物保藏中心获得人结肠腺癌细胞系 HCT116 和 WiDr 细胞。HCT116 细胞维持在补充有 10% 胎牛血清(FBS)的 McCoy' s 5A 培养基中。WiDr 细胞维持在补充有 10% FBS 的 Dulbecco' s 极限必需培养基(DMEM)中。对视黄酸诱导进行研究时, 将细胞以 10^5 细胞/60mm 平皿平铺在包含 2.5ml 培养基的培养皿上并使其生长 24 小时。用维生素 D3, 全-反-RA(Spectrum Laboratory Products), 或 9-顺-RA(Toronto Research Chemicals Inc.)(在 DMSO 中的最终浓度为 1 μ M)处理细胞指定的时间。对照细胞用等量的 DMSO 处理。

对 C57MG/亲代细胞用 Wnt-3a 条件培养基处理时, 将细胞在有或无 1 μ M 9-顺-RA 的情况下在常规培养基或已有 L-细胞或 L-W3A 细胞生长的条件培养基中。条件培养基如前所述制备 (Willert 等, Genes Dev.13:1768-73(1999))。到 48 小时的时候, 刮取细胞放入包含氟化钠和钒酸盐的 PBS 中并在液氮中快速冷冻。用 RNAeasy 试剂盒(Qiagen)制备 RNA, 包括依照生产商的使用说明, 在层析柱上进行另外的 DNase I 处理。

Western 印迹

对于 C57MG/Wnt-1 TET 细胞系, 在无四环素存在的情况下培养所述细胞 0, 24, 48 或 72 小时来诱导 Wnt-1 的表达。在 Triton X-100 裂解缓冲液

[20mM tris-HCl(pH8.0), 137mM NaCl, 1 % Triton X-100, 1mM EGTA, 10 % 甘油, 1.5mM MgCl₂, 1mM 二硫苏糖醇, 1mM 钒酸钠, 50mM 氟化钠和完全蛋白酶抑制剂混合物(Boehringer Mannheim)]中裂解细胞并对蛋白等价物进行 SDS-PAGE 和免疫印迹。所得印迹与 0.2 μg/ml 亲和层析纯化的兔多克隆抗 β-连环蛋白抗体(Rubinfeld 等, Science 262:1731-1734[1993]), 0.1μg/ml 抗-ERK2 单克隆抗体(Transduction Laboratories)温育, 或者抗 RAR γ-1 的 1:2000 兔多克隆抗血清(Affinity Bioreagents)温育。对于 WiDr 细胞系, 用 1 μM 全-反-RA 处理细胞 48 小时, 然后在 Triton X-100 裂解缓冲液中将其裂解并如上所述进行处理。使所得印迹与 1:50 抗-Stra6 肽 B 单克隆杂交瘤培养上清(克隆 12F4.2H9.1D5)温育。

免疫组化

WiDr 细胞用 9-顺-视黄酸或 DMSO 处理, 然后脱壁并通过低速离心而沉淀。将细胞团在 10 % 中性缓冲的福尔马林中固定过夜, 脱水并用石蜡包埋。用抗 Stra6 肽 B 单克隆杂交瘤培养上清(克隆 12F4.2H9.1D5)或非特异性小鼠同种型 IgG2A 作为第一抗体进行免疫组化, 然后用亲和素-生物素复合物方法用二氨基联苯胺作为显色剂(Bectastain Elite 试剂盒, Vector 实验室)如前所述(Eberhard 等, Am. J. Pathol. 145:640-9[1994])进行检测。用苏木精复染切片。

Wnt-1 转基因小鼠

在乳腺中表达 Wnt-1 原癌基因的转基因小鼠通常表现过度增生性损伤并在该组织中出现肿瘤(Tsukamoto 等, Cell 55:619-625[1988])。在以下试验中使用这种小鼠。

结果

Easwaran 等以前报导, 当用突变的 β-连环蛋白共转染并用维生素 A 类处理 MCF-7 细胞时, 合成的视黄酸应答性报告基因的活化增强(Easwaran 等, Curr. Biol. 9:1415-1418(1999))。考虑到这一点, 以及 Stra6 最初被鉴定为视黄酸可诱导基因(Bouillet 等, Mech Dev. 63:173-186[1997]), 我们想知道视黄酸是否与 Wnt-1 协同增加 C57MG 细胞系中 Stra6 的水平。用 9-顺-RA 或者全-反-RA(ATRA)处理亲代 C57MG 细胞 48 小时, 发现 Stra6 mRNA 的水平明显增加, 而 DMSO 和维生素 D3 无此作用(图 17A)。如所预期的, 用 DMSO 或维生素 D3 处理 C57MG/Wnt-1 细胞显示 Stra6 mRNA 的水平相对于亲代细

胞系是增加的。Wnt-1 对 Stra6 的诱导水平与用 9-顺-RA 刺激亲代 C57MG 细胞观察到的相当。然而，C57MG/Wnt-1 细胞系的 9-顺-RA 处理诱导的 Stra6 mRNA 相对于未经处理的 C57MG/Wnt-1 或 9-顺-RA 处理的 C57MG 亲代再增加了 10 倍。用全-反-RA 进行的试验也获得了类似的结果。

C57MG/Wnt-1 细胞系中可能存在与 Wnt-1 表达无关的潜在克隆变体，因为这些细胞与亲代对照细胞相比对视黄酸的应答不同。为阐明这一点，我们检测了亲代 C57MG 细胞在有或无 9-顺-RA 存在的情况对可溶性 Wnt-3a 刺激的应答。Wnt-3a 是 Wnt-1 的类似物，其显示与 Wnt-1 类似的转化特性，在小鼠 L 细胞中为可溶性配体形式。表达 Wnt-3a 的培养的 L 细胞而非对照 L 细胞生长过的条件培养基诱导 C57MG 细胞中 Stra6 的表达(图 17B)。用 9-顺-RA 对 C57MG 细胞进行处理可观察到诱导水平略微升高。然而，9-顺-RA 和 Wnt-3a 的组合导致 Stra6 转录子的水平大大超过单独用任一制剂所观察到的水平。

如果 β 连环蛋白水平的增加可加强对视黄酸处理过的 C57MG/Wnt-1 细胞中 Stra6 的诱导，那么人们可以预计在人结肠癌细胞(包含 β 连环蛋白突变或 APC 突变)中对视黄酸应答时对 Stra6 有类似诱导。为确定该现象是否发生，在视黄酸处理 HCT116 细胞(β 连环蛋白中有激活突变)和 WiDr 细胞(野生型 APC 的两个拷贝都失去了)之前和之后检测 Stra6 mRNA 水平。在这两种细胞系中，可观察到用 ATRA 或 9-顺-RA 处理比用 DMSO 或维生素 D3 处理，Stra6 mRNA 的水平明显增加(图 17C)。在 HCT116 细胞系中，ATRA 所致的 Stra6 基因活化通过原位杂交被确定(图 17D)。用 Stra6 特异性的单克隆抗体经 Western 印迹检测到用 ATRA 处理的 WiDr 细胞中预计的 73kDa Stra6 蛋白带的诱导(图 17E)。对 WiDr 细胞的免疫组化分析揭示，对视黄酸应答时 Stra6 蛋白的增加定位在浆膜(图 17F)。

已提出在非洲爪蟾属胚胎中 RAR γ 基因为 Wnt 信号传递的靶，并且显示维生素 A 类对 Stra6 的诱导依赖于该视黄酸受体亚型的存在(McGrew 等, *Mech. Dev.* 87:21-32 [1999]; Taneja 等, *Proc. Natl. Acad. Sci. USA* 92:7854-81[1995])。同时，这些发现提示 Wnt 和维生素 A 类对 Stra6 的协同活化可能是由于 Wnt 信号传递对 RAR γ 表达的活化。为确定 Wnt-1 信号传递对哺乳动物细胞中 RAR γ 的水平是否有任何影响，我们对裂解物上的受体进行了 Western 印迹，所述裂解物制备自有条件地表达 Wnt-1 的 C57MG

细胞。通过 Wnt-1 的表达,在 24 小时时诱导出一种与 RAR γ -1 特异性抗体反应且以 64kDa 的表观分子量迁移的蛋白,在 48 小时时其表达增加(图 18A)。我们也分析了过度增生的乳腺和来源于 Wnt-1 转基因小鼠的乳腺肿瘤,并检测了 RAR γ mRNA 在这些组织中相对于正常乳腺中升高的水平(图 18B)。用定量 PCR 评估从 19 例腺癌中分离到的等量 RNA 中 RAR γ 转录子的水平。很明显,在所检查的 19 例人结肠肿瘤中 14 例(74%)的 RAR γ mRNA 的表达比正常人结肠组织增加大约 2-4 倍(图 18C)。这些结果表明, Wnt-1 信号传递促进 RAR γ 的表达,这种受体受 Wnt-1 途径的驱动,在小鼠和人肿瘤中增加。

实验结果的讨论

基因表达作图法是基于对 mRNA 转录子的无偏检测,因此对基因活化发生的机制有意外的发现。此处显示 Wnt-1 促进视黄酸应答基因 Stra6 的诱导,提示 Wnt 和视黄酸激发的信号传递途径之间有联系。这种联系进一步通过证实 Wnt 和视黄酸的组合对 Stra6 产生协同诱导而得到支持。对这种协同作用至少有三种不同的解释: i)直接应答 Wnt 的转录因子如 TCF/LEFs 可能结合和活化 Stra6 基因中的启动子元件; ii)Wnt 途径中的信号传递成分可能直接与适当的视黄酸受体(RAR)相互作用并增强由 RAR 介导的基因活化;或 iii)通过 Wnt-1 传递的信号能活化相应 RAR 的表达,然后 RAR 致使细胞对视黄酸敏感。尽管赞成这一最终提议,因为本文所呈现的数据显示 Wnt-1 信号传递快速诱导 RAR γ 受体的表达,但本发明没有受到任何特定理论或作用机制的限制。

以前的工作已显示,对 RAR γ 基因特异性的破坏可大大减低维生素 A 类对 Stra6 基因的诱导,而这种诱导在该受体再表达时恢复(Taneja 等,出处同上)。然而,除 Wnt-1 诱导 RAR γ 表达之外仍有可能存在其它机制也对所观察到的协同起作用。 β 连环蛋白结合视黄酸受体的提议是引人注意的,但在我们的细胞系中没有观察到内源性 RAR γ 和 β 连环蛋白之间任何特异性联系。然而,体外显示 β 连环蛋白与 RAR γ 结合(Easwanran 等, 1999, 出处同上),这可能与 Wnt-1 对 Stra6 的活化有关。在无 RAR γ 的转基因动物中 Stra6 的这种表达方式比在野生型同窝动物中观察到的更普遍,并且在裸(null)动物的细胞中视黄酸对 Stra6 的诱导与对照动物相比被加强(Bouillet 等, 1997, 出处同上; Taneja 等, 1995, 出处同上)。所以 RAR γ 可能抑制 Stra6

表达, 且如果 β 连环蛋白抑制 RAR γ 的功能, Wnt-1 对 β 连环蛋白的活化可能会有效解除这种抑制。

Wnt-1 信号传递对视黄酸应答性基因的协同诱导提示那些在 Wnt-1 途径中有基因缺陷的人类癌症会出现 Stra6 的过表达。绝大多数的结肠直肠癌肿瘤在编码 APC 肿瘤抑制物或 β 连环蛋白的基因中含有突变(Polakes, Curr.Opin. Genet. Dev. 9:15-21 [1999]), 因此我们检测了 Stra6 在全部 14 例结肠直肠癌中相对于所匹配的正常组织中的过表达(参见实施例 12)。在卵巢癌和子宫内膜癌, Wilm' s 肾肿瘤和黑素瘤中也鉴定了 β 连环蛋白的活化突变, 这表明 Wnt-1 信号传递中的缺陷对这些癌症的进展是起作用的(Kobayashi 等, Jpn. J. Cancer Res.90:55-9[1999]; Koesters 等, Cancer Res. 59: 3880-2[1999]; Palacinós and Hamallo, Cancer Res. 58:1344-7[1998]; Rimm 等, Am. J. Pathol. 154: 325-9 [1999]; Rubinfeld 等 Science 262: 1731-1734 [1993]; Wright 等 Int. J. Cancer 82:625-9[1999])。通过原位杂交确定所有这四种人类肿瘤都呈现 Stra6 mRNA 的过表达 (参见实施例 12)。嗜铬细胞瘤这种来源于肾上腺髓质的肿瘤显示极高水平的 Stra6 mRNA, 然而, 有关这些肿瘤在 Wnt-1 信号传递途径中的突变的状态尚无报道。产生黑素瘤和嗜铬细胞瘤的细胞类型在胚胎学上都来源于神经管嵴, 这一点是令人非常感兴趣的。值得注意的是, 肾脏的 Wilm' s 肿瘤, 嗜铬细胞瘤和子宫内膜癌都起源于 Stra6 正常表达的器官。这提示肿瘤发生型信号传递, 如由 Wnt-1 途径驱动的那些, 可能与负责分化的信号相互作用, 从而过度活化 Stra6 的表达。

人类 Stra6 蛋白位于细胞表面, 可用人 Stra6 蛋白特异性单克隆抗体对结肠直肠癌细胞进行染色确定。而且, 用视黄酸对细胞进行处理后在细胞膜上观察到染色强度增加。因此, Stra6 可能编码定位于细胞表面的多次跨膜的蛋白。

材料的保藏

以下材料交美国典型培养物保藏中心, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC)保藏:

材料	ATCC 保藏号	保藏日
DNA148380-2827	PTA-1181	2000 年 1 月 11 日
DNA148389-2827-1	PTA-1405	2000 年 2 月 23 日

该保藏是按照国际承认的用于专利程序的微生物保藏的布达佩斯条约

及其章程(布达佩斯条约)进行的。这确保自保藏日起 30 年的时间内保藏培养物保持存活。这使得通过 ATCC 可以获得在布达佩斯条约期限内的保藏物,并服从 Genentech, Inc.与 ATCC 之间的协议,后者确保公众基于颁布的相关美国专利或者基于对任何美国或外国专利申请向公众的开放,无论哪一个先来,而可以永久性和无限制利用保藏培养物的传代物,并确保向美国专利商标委员会根据 35 U.S.C. §122 和所依据的委员会规则(包括 37 C.F.R. §1.14, 特别参见 886 OG 638)确定的个人发放传代物。

本申请的委托人已经同意,如果保藏培养物材料在适宜条件下培养过程中死亡或丢失或遭到破坏,应当在接到通知后及时用相同培养物的另一份材料替换保藏材料。保藏材料的获得不得解释为,违背任何政府当局根据其专利法的规定所赋予的权利,而实施本发明的许可。

前面撰写的说明书已经足以能使本领域技术人员实施本发明。本发明不受保藏构建体范围的限制,因为保藏方案旨在作为本发明一个特定方面的说明,功能等价的任何构建体也包括在本发明范围之内。本文的保藏材料既不构成准许本书面说明书所包含的不足以使本发明任何方面(包括其最佳模式)实施的内容的加入,也不解释为将权利要求的保护范围限定为特定的例证说明。事实上,除了本文所示和所述之外,对前面所描述的本发明进行各种变改,对于本领域技术人员而言是显而易见的,并落入本发明的保护范围。

序列表

<110> 杰南技术公司 (Genentech, Inc.)
 黛安娜·彭尼卡 (Pennica, Diane)
 维多利亚·史密斯 (Smith, Victoria)
 威廉·I·伍德 (Wood, William I.)

<120> 新的 STRA6 多肽

<130> GENENT. 2827VPC

<150> 60/197089

<151> 2000-04-14

<150> 60/175849

<151> 2000-01-13

<160> 5

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2732

<212> DNA

<213> 人 (Homo Sapiens)

<220>

<221> CDS

<222> (49)... (2052)

<400> 1

```

agtcccagac gggcttttcc cagagagcta aaagagaagg gccagaga atg tcg tcc 57
                                     Met Ser Ser
                                     1

```

```

cag cca gca ggg aac cag acc tcc ccc ggg gcc aca gag gac tac tcc 105
Gln Pro Ala Gly Asn Gln Thr Ser Pro Gly Ala Thr Glu Asp Tyr Ser
      5              10              15

```

```

tat ggc agc tgg tac atc gat gag ccc cag ggg ggc gag gag ctg cag 153
Tyr Gly Ser Trp Tyr Ile Asp Glu Pro Gln Gly Gly Glu Glu Leu Gln
      20              25              30              35

```

```

cca gag ggg gaa gtg ccc tcc tgc cac acc agc ata cca ccc ggc ctg 201
Pro Glu Gly Glu Val Pro Ser Cys His Thr Ser Ile Pro Pro Gly Leu
              40              45              50

```

```

tac cac gcc tgc ctg gcc tcg ctg tca atc ctt gtg ctg ctg ctg ctg 249
Tyr His Ala Cys Leu Ala Ser Leu Ser Ile Leu Val Leu Leu Leu
              55              60              65

```

```

gcc atg ctg gtg agg cgc cgc cag ctg tgg cct gac tgt gtg cgt ggc 297
Ala Met Leu Val Arg Arg Arg Gln Leu Trp Pro Asp Cys Val Arg Gly

```

70	75	80	
agg ccc ggc ctg ccc agc cct gtg gat ttc ttg gct ggg gac agg ccc			345
Arg Pro Gly Leu Pro Ser Pro Val Asp Phe Leu Ala Gly Asp Arg Pro			
85	90	95	
cgg gca gtg cct gct gct gtt ttc atg gtc ctc ctg agc tcc ctg tgt			393
Arg Ala Val Pro Ala Ala Val Phe Met Val Leu Leu Ser Ser Leu Cys			
100	105	110	115
ttg ctg ctc ccc gac gag gac gca ttg ccc ttc ctg act ctc gcc tca			441
Leu Leu Leu Pro Asp Glu Asp Ala Leu Pro Phe Leu Thr Leu Ala Ser			
120	125	130	
gca ccc agc caa gat ggg aaa act gag gct cca aga ggg gcc tgg aag			489
Ala Pro Ser Gln Asp Gly Lys Thr Glu Ala Pro Arg Gly Ala Trp Lys			
135	140	145	
ata ctg gga ctg ttc tat tat gct gcc ctc tac tac cct ctg gct gcc			537
Ile Leu Gly Leu Phe Tyr Tyr Ala Ala Leu Tyr Tyr Pro Leu Ala Ala			
150	155	160	
tgt gcc acg gct ggc cac aca gct gca cac ctg ctc ggc agc acg ctg			585
Cys Ala Thr Ala Gly His Thr Ala Ala His Leu Leu Gly Ser Thr Leu			
165	170	175	
tcc tgg gcc cac ctt ggg gtc cag gtc tgg cag agg gca gag tgt ccc			633
Ser Trp Ala His Leu Gly Val Gln Val Trp Gln Arg Ala Glu Cys Pro			
180	185	190	195
cag gtg ccc aag atc tac aag tac tac tcc ctg ctg gcc tcc ctg cct			681
Gln Val Pro Lys Ile Tyr Lys Tyr Tyr Ser Leu Leu Ala Ser Leu Pro			
200	205	210	
ctc ctg ctg ggc ctc gga ttc ctg agc ctt tgg tac cct gtg cag ctg			729
Leu Leu Leu Gly Leu Gly Phe Leu Ser Leu Trp Tyr Pro Val Gln Leu			
215	220	225	
gtg aga agc ttc agc cgt agg aca gga gca ggc tcc aag ggg ctg cag			777
Val Arg Ser Phe Ser Arg Arg Thr Gly Ala Gly Ser Lys Gly Leu Gln			
230	235	240	
agc agc tac tct gag gaa tat ctg agg aac ctc ctt tgc agg aag aag			825
Ser Ser Tyr Ser Glu Glu Tyr Leu Arg Asn Leu Leu Cys Arg Lys Lys			
245	250	255	
ctg gga agc agc tac cac acc tcc aag cat ggc ttc ctg tcc tgg gcc			873
Leu Gly Ser Ser Tyr His Thr Ser Lys His Gly Phe Leu Ser Trp Ala			
260	265	270	275
cgc gtc tgc ttg aga cac tgc atc tac act cca cag cca gga ttc cat			921
Arg Val Cys Leu Arg His Cys Ile Tyr Thr Pro Gln Pro Gly Phe His			
280	285	290	
ctc ccg ctg aag ctg gtg ctt tca gct aca ctg aca ggg acg gcc att			969
Leu Pro Leu Lys Leu Val Leu Ser Ala Thr Leu Thr Gly Thr Ala Ile			
295	300	305	

tac cag gtg gcc ctg ctg ctg ctg gtg ggc gtg gta ccc act atc cag	1017
Tyr Gln Val Ala Leu Leu Leu Leu Val Gly Val Val Pro Thr Ile Gln	
310 315 320	
aag gtg agg gca ggg gtc acc acg gat gtc tcc tac ctg ctg gcc ggc	1065
Lys Val Arg Ala Gly Val Thr Thr Asp Val Ser Tyr Leu Leu Ala Gly	
325 330 335	
ttt gga atc gtg ctc tcc gag gac aag cag gag gtg gtg gag ctg gtg	1113
Phe Gly Ile Val Leu Ser Glu Asp Lys Gln Glu Val Val Glu Leu Val	
340 345 350 355	
aag cac cat ctg tgg gct ctg gaa gtg tgc tac atc tca gcc ttg gtc	1161
Lys His His Leu Trp Ala Leu Glu Val Cys Tyr Ile Ser Ala Leu Val	
360 365 370	
tig tcc tgc tta ctc acc ttc ctg gtc ctg atg cgc tca ctg gtg aca	1209
Leu Ser Cys Leu Leu Thr Phe Leu Val Leu Met Arg Ser Leu Val Thr	
375 380 385	
cac agg acc aac ctt cga gct ctg cac cga gga gct gcc ctg gac ttg	1257
His Arg Thr Asn Leu Arg Ala Leu His Arg Gly Ala Ala Leu Asp Leu	
390 395 400	
agt ccc ttg cat cgg agt ccc cat ccc tcc cgc caa gcc ata ttc tgt	1305
Ser Pro Leu His Arg Ser Pro His Pro Ser Arg Gln Ala Ile Phe Cys	
405 410 415	
igg atg agc ttc agt gcc tac cag aca gcc ttt atc tgc ctt ggg ctc	1353
Trp Met Ser Phe Ser Ala Tyr Gln Thr Ala Phe Ile Cys Leu Gly Leu	
420 425 430 435	
ctg gtg cag cag atc atc ttc ttc ctg gga acc acg gcc ctg gcc ttc	1401
Leu Val Gln Gln Ile Ile Phe Phe Leu Gly Thr Thr Ala Leu Ala Phe	
440 445 450	
ctg gtg ctc atg cct gtg ctc cat ggc agg aac ctc ctg ctc ttc cgt	1449
Leu Val Leu Met Pro Val Leu His Gly Arg Asn Leu Leu Leu Phe Arg	
455 460 465	
tcc ctg gag tcc tgc tgg ccc ttc tgg ctg act ttg gcc ctg gct gtg	1497
Ser Leu Glu Ser Ser Trp Pro Phe Trp Leu Thr Leu Ala Leu Ala Val	
470 475 480	
atc ctg cag aac atg gca gcc cat tgg gtc ttc ctg gag act cat gat	1545
Ile Leu Gln Asn Met Ala Ala His Trp Val Phe Leu Glu Thr His Asp	
485 490 495	
gga cac cca cag ctg acc aac cgg cga gtg ctc tat gca gcc acc ttt	1593
Gly His Pro Gln Leu Thr Asn Arg Arg Val Leu Tyr Ala Ala Thr Phe	
500 505 510 515	
ctt ctc ttc ccc ctc aat gtg ctg gtg ggt gcc atg gtg gcc acc tgg	1641
Leu Leu Phe Pro Leu Asn Val Leu Val Gly Ala Met Val Ala Thr Trp	
520 525 530	
cga gtg ctc ctc tct gcc ctc tac aac gcc atc cac ctt ggc cag atg	1689
Arg Val Leu Leu Ser Ala Leu Tyr Asn Ala Ile His Leu Gly Gln Met	

```

535          540          545
gac ctc agc ctg ctg cca ccg aga gcc gcc act ctc gac ccc ggc tac 1737
Asp Leu Ser Leu Leu Pro Pro Arg Ala Ala Thr Leu Asp Pro Gly Tyr
550          555          560

tac acg tac cga aac ttc ttg aag att gaa gtc agc cag tgc cat cca 1785
Tyr Thr Tyr Arg Asn Phe Leu Lys Ile Glu Val Ser Gln Ser His Pro
565          570          575

gcc atg aca gcc ttc tgc tcc ctg ctc ctg caa gcg cag agc ctc cta 1833
Ala Met Thr Ala Phe Cys Ser Leu Leu Leu Gln Ala Gln Ser Leu Leu
580          585          590          595

ccc agg acc atg gca gcc ccc cag gac agc ctc aga cca ggg gag gaa 1881
Pro Arg Thr Met Ala Ala Pro Gln Asp Ser Leu Arg Pro Gly Glu Glu
600          605          610

gac gaa ggg atg cag ctg cta cag aca aag gac tcc atg gcc aag gga 1929
Asp Glu Gly Met Gln Leu Leu Gln Thr Lys Asp Ser Met Ala Lys Gly
615          620          625

gct agg ccc ggg gcc agc cgc ggc agg gct cgc tgg ggt ctg gcc tac 1977
Ala Arg Pro Gly Ala Ser Arg Gly Arg Ala Arg Trp Gly Leu Ala Tyr
630          635          640

acg ctg ctg cac aac cca acc ctg cag gtc ttc cgc aag acg gcc ctg 2025
Thr Leu Leu His Asn Pro Thr Leu Gln Val Phe Arg Lys Thr Ala Leu
645          650          655

ttg ggt gcc aat ggt gcc cag ccc tga gggcagggaa ggtcaaccca 2072
Leu Gly Ala Asn Gly Ala Gln Pro *
660          665

cctgcccatc tgtgctgagg catgttcctg cctaccatcc tcctccctcc ccggctctcc 2132
tcccagcatc acaccagcca tgcagccagc aggtcctccg gatcactgtg gttgggtgga 2192
ggtctgtctg cactgggagc ctccaggagg ctctgctcca ccacttggc tatgggagag 2252
ccagcagggg ttctggagaa aaaaacttgt gggttagggc ctgtgtccag gagccagtig 2312
agccagggga gccacatcca ggcgtctccc taccctggct ctgccatcag ccttgaagg 2372
ccctgatgaa gccttctctg gaaccactcc agcccagctc caccctagcc ttggcctica 2432
cgctgtggaa gcagccaagg cacttctcca cccctcagc gccacggacc tccttgggga 2492
gtggccggaa agctcccggt cctctggcct gcagggcagc ccaagtcag actcagacca 2552
ggtccacac tgagctgccc acactcgaga gccagatatt ttgtagtgt ttatgccttt 2612
ggctattatg aaagagggtg gtgtgttccc tgcaataaac ttgttcctga gaaaaaaaaa 2672
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2732

```

<210> 2

<211> 667

<212> PRT

<213> 人 (Homo Sapiens)

<400> 2

```

Met Ser Ser Gln Pro Ala Gly Asn Gln Thr Ser Pro Gly Ala Thr Glu
1          5          10          15
Asp Tyr Ser Tyr Gly Ser Trp Tyr Ile Asp Glu Pro Gln Gly Gly Glu
20          25          30
Glu Leu Gln Pro Glu Gly Glu Val Pro Ser Cys His Thr Ser Ile Pro

```

35	40	45
Pro Gly Leu Tyr His Ala Cys Leu Ala Ser Leu Ser Ile Leu Val Leu		
50	55	60
Leu Leu Leu Ala Met Leu Val Arg Arg Arg Gln Leu Trp Pro Asp Cys		
65	70	75
Val Arg Gly Arg Pro Gly Leu Pro Ser Pro Val Asp Phe Leu Ala Gly		
85	90	95
Asp Arg Pro Arg Ala Val Pro Ala Ala Val Phe Met Val Leu Leu Ser		
100	105	110
Ser Leu Cys Leu Leu Leu Pro Asp Glu Asp Ala Leu Pro Phe Leu Thr		
115	120	125
Leu Ala Ser Ala Pro Ser Gln Asp Gly Lys Thr Glu Ala Pro Arg Gly		
130	135	140
Ala Trp Lys Ile Leu Gly Leu Phe Tyr Tyr Ala Ala Leu Tyr Tyr Pro		
145	150	155
Leu Ala Ala Cys Ala Thr Ala Gly His Thr Ala Ala His Leu Leu Gly		
165	170	175
Ser Thr Leu Ser Trp Ala His Leu Gly Val Gln Val Trp Gln Arg Ala		
180	185	190
Glu Cys Pro Gln Val Pro Lys Ile Tyr Lys Tyr Tyr Ser Leu Leu Ala		
195	200	205
Ser Leu Pro Leu Leu Leu Gly Leu Gly Phe Leu Ser Leu Trp Tyr Pro		
210	215	220
Val Gln Leu Val Arg Ser Phe Ser Arg Arg Thr Gly Ala Gly Ser Lys		
225	230	235
Gly Leu Gln Ser Ser Tyr Ser Glu Glu Tyr Leu Arg Asn Leu Leu Cys		
245	250	255
Arg Lys Lys Leu Gly Ser Ser Tyr His Thr Ser Lys His Gly Phe Leu		
260	265	270
Ser Trp Ala Arg Val Cys Leu Arg His Cys Ile Tyr Thr Pro Gln Pro		
275	280	285
Gly Phe His Leu Pro Leu Lys Leu Val Leu Ser Ala Thr Leu Thr Gly		
290	295	300
Thr Ala Ile Tyr Gln Val Ala Leu Leu Leu Leu Val Gly Val Val Pro		
305	310	315
Thr Ile Gln Lys Val Arg Ala Gly Val Thr Thr Asp Val Ser Tyr Leu		
325	330	335
Leu Ala Gly Phe Gly Ile Val Leu Ser Glu Asp Lys Gln Glu Val Val		
340	345	350
Glu Leu Val Lys His His Leu Trp Ala Leu Glu Val Cys Tyr Ile Ser		
355	360	365
Ala Leu Val Leu Ser Cys Leu Leu Thr Phe Leu Val Leu Met Arg Ser		
370	375	380
Leu Val Thr His Arg Thr Asn Leu Arg Ala Leu His Arg Gly Ala Ala		
385	390	395
Leu Asp Leu Ser Pro Leu His Arg Ser Pro His Pro Ser Arg Gln Ala		
405	410	415
Ile Phe Cys Trp Met Ser Phe Ser Ala Tyr Gln Thr Ala Phe Ile Cys		
420	425	430
Leu Gly Leu Leu Val Gln Gln Ile Ile Phe Phe Leu Gly Thr Thr Ala		
435	440	445
Leu Ala Phe Leu Val Leu Met Pro Val Leu His Gly Arg Asn Leu Leu		
450	455	460
Leu Phe Arg Ser Leu Glu Ser Ser Trp Pro Phe Trp Leu Thr Leu Ala		
465	470	475
Leu Ala Val Ile Leu Gln Asn Met Ala Ala His Trp Val Phe Leu Glu		
485	490	495
Thr His Asp Gly His Pro Gln Leu Thr Asn Arg Arg Val Leu Tyr Ala		

```

          500          505          510
Ala Thr Phe Leu Leu Phe Pro Leu Asn Val Leu Val Gly Ala Met Val
          515          520          525
Ala Thr Trp Arg Val Leu Leu Ser Ala Leu Tyr Asn Ala Ile His Leu
          530          535          540
Gly Gln Met Asp Leu Ser Leu Leu Pro Pro Arg Ala Ala Thr Leu Asp
545          550          555          560
Pro Gly Tyr Tyr Thr Tyr Arg Asn Phe Leu Lys Ile Glu Val Ser Gln
          565          570          575
Ser His Pro Ala Met Thr Ala Phe Cys Ser Leu Leu Leu Gln Ala Gln
          580          585          590
Ser Leu Leu Pro Arg Thr Met Ala Ala Pro Gln Asp Ser Leu Arg Pro
          595          600          605
Gly Glu Glu Asp Glu Gly Met Gln Leu Leu Gln Thr Lys Asp Ser Met
          610          615          620
Ala Lys Gly Ala Arg Pro Gly Ala Ser Arg Gly Arg Ala Arg Trp Gly
625          630          635          640
Leu Ala Tyr Thr Leu Leu His Asn Pro Thr Leu Gln Val Phe Arg Lys
          645          650          655
Thr Ala Leu Leu Gly Ala Asn Gly Ala Gln Pro
          660          665

```

<210> 3
 <211> 676
 <212> DNA
 <213> 人 (Homo Sapiens)

<220>
 <221> misc-feature
 <222> (26)... (26)
 <223> n = A, T, C 或 G

```

<400> 3
gtgctctccg aggacaagca ggaggnggtg gagctgggtg agcaccatct gtgggctctg 60
gaaglggtgc acatctcagc ctgggtcttg tctgtcttac tcaccttcct ggtcctgatg 120
cgctcactgg tgacacacag gaccaacctt cgagctctgc accgaggagc tgccctggac 180
ttgagtcctt tgcctcggag tcccatccc tccgccaag ccatactctg ttggatgagc 240
ttcagtgctt accagacagc ctttatctgc ctgggctcc tgggtgcagc gatcatcttc 300
ttcctgggaa ccacggccct ggcttctctg gtgctcatgc ctgtgctcca tggcaggaac 360
ctcctgctct tccgttccct ggagtcctcg tggcccttct ggctgacttt ggccctggct 420
gtgatcctgc agaacaaggc agccatttgg gtcttcttgg agactcatga tggacacca 480
cagctgacca accggcgagt gctctatgca gccaccttct ttctcttccc cctcaatgtg 540
ctgggtgggtg ccatgggtggc cacttggcga gtgtcctctt ctgccctcta caacgccatc 600
caccttggcc agatggacct cagcctgctg ccaccgagag ccgccactct cgaccccggc 660
tactacacgt accgaa 676

```

<210> 4
 <211> 2777
 <212> DNA
 <213> 人 (Homo Sapiens)

<220>
 <221> CDS
 <222> (186)... (2160)

```

<400> 4
cacaaccagc caccctctta ggatcccagc ccagctgggtg ctgggctcag aggagaaggc 60

```

```

cccgigtgg gagcaccctg ctgacctgga gggacaagtt tccgggagag atcaataaag 120
gaaaggaaag agacaaggaa gggagaggtc aggagagcgc ttgattggag gagaagggcc 180
agaga atg tcg tcc cag cca gca ggg aac cag acc tcc ccc ggg gcc aca 230
    Met Ser Ser Gln Pro Ala Gly Asn Gln Thr Ser Pro Gly Ala Thr
        1             5             10             15

gag gac tac tcc tat ggc agc tgg tac atc gat gag ccc cag ggg ggc 278
Glu Asp Tyr Ser Tyr Gly Ser Trp Tyr Ile Asp Glu Pro Gln Gly Gly
        20             25             30

gag gag ctc cag cca gag ggg gaa gtg ccc tcc tgc cac acc agc ata 326
Glu Glu Leu Gln Pro Glu Gly Glu Val Pro Ser Cys His Thr Ser Ile
        35             40             45

cca ccc ggc ctg tac cac gcc tgc ctg gcc tcg ctg tca atc ctt gtg 374
Pro Pro Gly Leu Tyr His Ala Cys Leu Ala Ser Leu Ser Ile Leu Val
        50             55             60

ctg ctg ctc ctg gcc atg ctg gtg agg cgc cgc cag ctc tgg cct gac 422
Leu Leu Leu Leu Ala Met Leu Val Arg Arg Arg Gln Leu Trp Pro Asp
        65             70             75

tgt gtg cgt ggc agg ccc ggc ctg ccc agg ccc cgg gca gtg cct gct 470
Cys Val Arg Gly Arg Pro Gly Leu Pro Arg Pro Arg Ala Val Pro Ala
        80             85             90             95

gct gtt ttc atg gtc ctc ctg agc tcc ctg tgt ttg ctg ctc ccc gac 518
Ala Val Phe Met Val Leu Leu Ser Ser Leu Cys Leu Leu Leu Pro Asp
        100             105             110

gag gac gca ttg ccc ttc ctg act ctc gcc tca gca ccc agc caa gat 566
Glu Asp Ala Leu Pro Phe Leu Thr Leu Ala Ser Ala Pro Ser Gln Asp
        115             120             125

ggg aaa act gag gct cca aga ggg gcc tgg aag ata ctg gga ctg ttc 614
Gly Lys Thr Glu Ala Pro Arg Gly Ala Trp Lys Ile Leu Gly Leu Phe
        130             135             140

tat tat gct gcc ctc tac tac cct ctg gct gcc tgt gcc acg gct ggc 662
Tyr Tyr Ala Ala Leu Tyr Tyr Pro Leu Ala Ala Cys Ala Thr Ala Gly
        145             150             155

cac aca gct gca cac ctg ctc ggc agc acg ctg tcc tgg gcc cac ctt 710
His Thr Ala Ala His Leu Leu Gly Ser Thr Leu Ser Trp Ala His Leu
        160             165             170             175

ggg gtc cag gtc tgg cag agg gca gag tgt ccc cag gtg ccc aag atc 758
Gly Val Gln Val Trp Gln Arg Ala Glu Cys Pro Gln Val Pro Lys Ile
        180             185             190

tac aag tac tac tcc ctg ctg gcc tcc ctg cct ctc ctg ctg ggc ctc 806
Tyr Lys Tyr Tyr Ser Leu Leu Ala Ser Leu Pro Leu Leu Leu Gly Leu
        195             200             205

gga ttc ctg agc ctt tgg tac cct gtg cag ctg gtg aga agc ttc agc 854
Gly Phe Leu Ser Leu Trp Tyr Pro Val Gln Leu Val Arg Ser Phe Ser
        210             215             220

```

cg ₁ ag _g ac _a gg _a gc _a gg _c tc _c aa _g gg _g ct _g ca _g ag _c ag _c ta _c tc _t ga _g	902
Arg Arg Thr Gly Ala Gly Ser Lys Gly Leu Gln Ser Ser Tyr Ser Glu	
225 230 235	
gaa tat ctg agg aac ctc ctt tgc agg aag aag ctg gga agc agc tac	950
Glu Tyr Leu Arg Asn Leu Leu Cys Arg Lys Lys Leu Gly Ser Ser Tyr	
240 245 250 255	
cac acc tcc aag cat gg _c tt _c ct _g tc _c tgg gcc cgc gtc tgc ttg aga	998
His Thr Ser Lys His Gly Phe Leu Ser Trp Ala Arg Val Cys Leu Arg	
260 265 270	
cac tgc atc tac act cca cag cca gga ttc cat ctc ccg ctg aag ctg	1046
His Cys Ile Tyr Thr Pro Gln Pro Gly Phe His Leu Pro Leu Lys Leu	
275 280 285	
gtg ctt tca gct aca ctg aca gg _g acg gcc att tac cag gtg gcc ctg	1094
Val Leu Ser Ala Thr Leu Thr Gly Thr Ala Ile Tyr Gln Val Ala Leu	
290 295 300	
ctg ctg ctg gtg gg _c gtg gta ccc act atc cag aag gtg agg gca gg _g	1142
Leu Leu Leu Val Gly Val Val Pro Thr Ile Gln Lys Val Arg Ala Gly	
305 310 315	
gtc acc acg gat gtc tcc tac ctg ctg gcc gg _c ttt gga atc gtg ctc	1190
Val Thr Thr Asp Val Ser Tyr Leu Leu Ala Gly Phe Gly Ile Val Leu	
320 325 330 335	
tcc gag gac aag cag gag gtg gtg gag ctg gtg aag cac cat ctg tgg	1238
Ser Glu Asp Lys Gln Glu Val Val Glu Leu Val Lys His His Leu Trp	
340 345 350	
gct ctg gaa gtg tgc tac atc tca gcc ttg gtc ttg tcc tgc tta ctc	1286
Ala Leu Glu Val Cys Tyr Ile Ser Ala Leu Val Leu Ser Cys Leu Leu	
355 360 365	
acc ttc ctg gtc ctg atg cgc tca ctg gtg aca cac agg acc aac ctt	1334
Thr Phe Leu Val Leu Met Arg Ser Leu Val Thr His Arg Thr Asn Leu	
370 375 380	
cga gct ctg cac cga gga gct gcc ctg gac ttg agt ccc ttg cat cgg	1382
Arg Ala Leu His Arg Gly Ala Ala Leu Asp Leu Ser Pro Leu His Arg	
385 390 395	
agt ccc cat ccc tcc cgc caa gcc ata ttc tgt tgg atg agc ttc agt	1430
Ser Pro His Pro Ser Arg Gln Ala Ile Phe Cys Trp Met Ser Phe Ser	
400 405 410 415	
gcc tac cag aca gcc ttt atc tgc ctt gg _g ctc ctg gtg cag cag atc	1478
Ala Tyr Gln Thr Ala Phe Ile Cys Leu Gly Leu Leu Val Gln Gln Ile	
420 425 430	
atc ttc ttc ctg gga acc acg gcc ctg gcc ttc ctg gtg ctc atg cct	1526
Ile Phe Phe Leu Gly Thr Thr Ala Leu Ala Phe Leu Val Leu Met Pro	
435 440 445	
gtg ctc cat gg _c agg aac ctc ctg ctc ttc cgt tcc ctg gag tcc tcg	1574
Val Leu His Gly Arg Asn Leu Leu Leu Phe Arg Ser Leu Glu Ser Ser	

450	455	460	
tgg ccc ttc tgg ctg act ttg gcc ctg gct gtg atc ctg cag aac atg			1622
Trp Pro Phe Trp Leu Thr Leu Ala Leu Ala Val Ile Leu Gln Asn Met			
465	470	475	
gca gcc cat tgg gtc ttc ctg gag act cat gat gga cac cca cag ctg			1670
Ala Ala His Trp Val Phe Leu Glu Thr His Asp Gly His Pro Gln Leu			
480	485	490	495
acc aac cgg cga gtg ctc tat gca gcc acc ttt ctt ctc ttc ccc ctc			1718
Thr Asn Arg Arg Val Leu Tyr Ala Ala Thr Phe Leu Leu Phe Pro Leu			
	500	505	510
aat gtg ctg gtg ggt gcc ata gtg gcc acc tgg cga gtg ctc ctc tct			1766
Asn Val Leu Val Gly Ala Ile Val Ala Thr Trp Arg Val Leu Leu Ser			
	515	520	525
gcc ctc tac aac gcc atc cac ctt ggc cag atg gac ctc agc ctg ctg			1814
Ala Leu Tyr Asn Ala Ile His Leu Gly Gln Met Asp Leu Ser Leu Leu			
	530	535	540
cca ccg aga gcc gcc act ctc gac ccc ggc tac tac acg tac cga aac			1862
Pro Pro Arg Ala Ala Thr Leu Asp Pro Gly Tyr Tyr Thr Tyr Arg Asn			
	545	550	555
ttc ttg aag att gaa gtc agc cag tcg cat cca gcc atg aca gcc ttc			1910
Phe Leu Lys Ile Glu Val Ser Gln Ser His Pro Ala Met Thr Ala Phe			
560	565	570	575
tgc tcc ctg ctc ctg caa gcg cag agc ctc cta ccc agg acc atg gca			1958
Cys Ser Leu Leu Leu Gln Ala Gln Ser Leu Leu Pro Arg Thr Met Ala			
	580	585	590
gcc ccc cag gac agc ctc aga cca ggg gag gaa gac gaa ggg atg cag			2006
Ala Pro Gln Asp Ser Leu Arg Pro Gly Glu Glu Asp Glu Gly Met Gln			
	595	600	605
ctg cta cag aca aag gac tcc atg gcc aag gga gct agg ccc ggg gcc			2054
Leu Leu Gln Thr Lys Asp Ser Met Ala Lys Gly Ala Arg Pro Gly Ala			
	610	615	620
agc cgc ggc agg gct cgc tgg ggt ctg gcc tac acg ctg ctg cac aac			2102
Ser Arg Gly Arg Ala Arg Trp Gly Leu Ala Tyr Thr Leu Leu His Asn			
	625	630	635
cca acc ctg cag gtc ttc cgc aag acg gcc ctg ttg ggt gcc aat ggt			2150
Pro Thr Leu Gln Val Phe Arg Lys Thr Ala Leu Leu Gly Ala Asn Gly			
640	645	650	655
gcc cag ccc t gagggcaggg aaggtcaacc cacctgccca tctgtgctga			2200
Ala Gln Pro			
ggcatgttcc tgccctaccac ctctctccctc cccggctctc ctcccagcat cacaccagcc			2260
atgcagccag caggctctcc ggaicactgt ggttgggtgg aggtctgtct gcactgggag			2320
ccicaggagg gctctgctcc acccacttgg ctatgggaga gccagcaggg gtctctggaga			2380
aagaaactgg tgggttaggg ccttgggtcca ggagccagtt gagccagggc agccacatcc			2440

```

aggcgctctcc ctaccctggc tctgccatca gccttgaagg gcctcgatga agccttctct 2500
ggaaccactc cagccagct ccaccacagc ctggccttc acgctgtgga agcagccaag 2560
gcacttcctc accccctcag cgccacggac ctctctgggg agtggccgga aagctcccgg 2620
gcctctggcc tgcagggcag cccaagtcac gactcagacc aggtcccaca ctgagctgcc 2680
cacactcgag agccagatat tttgtagtt tttatgcctt tggctattat gaaagagggt 2740
agtgtgttcc ctgcaataaa ctgttcctg agaaaaa 2777

```

<210> 5

<211> 658

<212> PRT

<213> 人 (Homo Sapiens)

<400> 5

```

Met Ser Ser Gln Pro Ala Gly Asn Gln Thr Ser Pro Gly Ala Thr Glu
1          5          10          15
Asp Tyr Ser Tyr Gly Ser Trp Tyr Ile Asp Glu Pro Gln Gly Gly Glu
20        25        30
Glu Leu Gln Pro Glu Gly Glu Val Pro Ser Cys His Thr Ser Ile Pro
35        40        45
Pro Gly Leu Tyr His Ala Cys Leu Ala Ser Leu Ser Ile Leu Val Leu
50        55        60
Leu Leu Leu Ala Met Leu Val Arg Arg Arg Gln Leu Trp Pro Asp Cys
65        70        75        80
Val Arg Gly Arg Pro Gly Leu Pro Arg Pro Arg Ala Val Pro Ala Ala
85        90        95
Val Phe Met Val Leu Leu Ser Ser Leu Cys Leu Leu Leu Pro Asp Glu
100       105       110
Asp Ala Leu Pro Phe Leu Thr Leu Ala Ser Ala Pro Ser Gln Asp Gly
115       120       125
Lys Thr Glu Ala Pro Arg Gly Ala Trp Lys Ile Leu Gly Leu Phe Tyr
130       135       140
Tyr Ala Ala Leu Tyr Tyr Pro Leu Ala Ala Cys Ala Thr Ala Gly His
145       150       155       160
Thr Ala Ala His Leu Leu Gly Ser Thr Leu Ser Trp Ala His Leu Gly
165       170       175
Val Gln Val Trp Gln Arg Ala Glu Cys Pro Gln Val Pro Lys Ile Tyr
180       185       190
Lys Tyr Tyr Ser Leu Leu Ala Ser Leu Pro Leu Leu Leu Gly Leu Gly
195       200       205
Phe Leu Ser Leu Trp Tyr Pro Val Gln Leu Val Arg Ser Phe Ser Arg
210       215       220
Arg Thr Gly Ala Gly Ser Lys Gly Leu Gln Ser Ser Tyr Ser Glu Glu
225       230       235       240
Tyr Leu Arg Asn Leu Leu Cys Arg Lys Lys Leu Gly Ser Ser Tyr His
245       250       255
Thr Ser Lys His Gly Phe Leu Ser Trp Ala Arg Val Cys Leu Arg His
260       265       270
Cys Ile Tyr Thr Pro Gln Pro Gly Phe His Leu Pro Leu Lys Leu Val
275       280       285
Leu Ser Ala Thr Leu Thr Gly Thr Ala Ile Tyr Gln Val Ala Leu Leu
290       295       300
Leu Leu Val Gly Val Val Pro Thr Ile Gln Lys Val Arg Ala Gly Val
305       310       315       320
Thr Thr Asp Val Ser Tyr Leu Leu Ala Gly Phe Gly Ile Val Leu Ser
325       330       335
Glu Asp Lys Gln Glu Val Val Glu Leu Val Lys His His Leu Trp Ala
340       345       350
Leu Glu Val Cys Tyr Ile Ser Ala Leu Val Leu Ser Cys Leu Leu Thr

```

```

      355              360              365
Phe Leu Val Leu Met Arg Ser Leu Val Thr His Arg Thr Asn Leu Arg
      370              375              380
Ala Leu His Arg Gly Ala Ala Leu Asp Leu Ser Pro Leu His Arg Ser
385              390              395              400
Pro His Pro Ser Arg Gln Ala Ile Phe Cys Trp Met Ser Phe Ser Ala
      405              410              415
Tyr Gln Thr Ala Phe Ile Cys Leu Gly Leu Leu Val Gln Gln Ile Ile
      420              425              430
Phe Phe Leu Gly Thr Thr Ala Leu Ala Phe Leu Val Leu Met Pro Val
      435              440              445
Leu His Gly Arg Asn Leu Leu Leu Phe Arg Ser Leu Glu Ser Ser Trp
      450              455              460
Pro Phe Trp Leu Thr Leu Ala Leu Ala Val Ile Leu Gln Asn Met Ala
465              470              475              480
Ala His Trp Val Phe Leu Glu Thr His Asp Gly His Pro Gln Leu Thr
      485              490              495
Asn Arg Arg Val Leu Tyr Ala Ala Thr Phe Leu Leu Phe Pro Leu Asn
      500              505              510
Val Leu Val Gly Ala Ile Val Ala Thr Trp Arg Val Leu Leu Ser Ala
      515              520              525
Leu Tyr Asn Ala Ile His Leu Gly Gln Met Asp Leu Ser Leu Leu Pro
      530              535              540
Pro Arg Ala Ala Thr Leu Asp Pro Gly Tyr Tyr Thr Tyr Arg Asn Phe
545              550              555              560
Leu Lys Ile Glu Val Ser Gln Ser His Pro Ala Met Thr Ala Phe Cys
      565              570              575
Ser Leu Leu Leu Gln Ala Gln Ser Leu Leu Pro Arg Thr Met Ala Ala
      580              585              590
Pro Gln Asp Ser Leu Arg Pro Gly Glu Glu Asp Glu Gly Met Gln Leu
      595              600              605
Leu Gln Thr Lys Asp Ser Met Ala Lys Gly Ala Arg Pro Gly Ala Ser
      610              615              620
Arg Gly Arg Ala Arg Trp Gly Leu Ala Tyr Thr Leu Leu His Asn Pro
625              630              635              640
Thr Leu Gln Val Phe Arg Lys Thr Ala Leu Leu Gly Ala Asn Gly Ala
      645              650              655
Gln Pro

```

AGTCCCAGACGGGCTTTTCCCAGAGAGCTAAAAGAGAAGGGCCAGAGAATGTCGTCCCAG
 CCAGCAGGGAAACCAGACCTCCCCGGGGCCACAGAGGACTACTCCTATGGCAGCTGGTAC
 ATCGATGAGCCCCAGGGGGCGAGGAGCTCCAGCCAGAGGGGGAAGTGCCCTCCTGCCAC
 ACCAGCATAACCACCGGCTGTACCACGCTGCTGGCCTCGCTGTCAATCCTTGTGCTG
 CTGCTCCTGGCCATGCTGGTGAGGCGCCGCGAGCTCTGGCCTGACTGTGTGCGTGGCAGG
 CCGGCTGCCCAGCCCTGTGGATTTCTTGGCTGGGGACAGGCCCCGGGCAGTGCCCTGCT
 GCTGTTTTTCATGGTCTCCTGAGCTCCCTGTGTTTGCTGCTCCCCGACGAGGACGCAATTG
 CCCTTCCTGACTCTCGCCTCAGCACCCAGCCAAGATGGGAAAACCTGAGGCTCCAAGAGGG
 GCCTGGAAGATACTGGGACTGTTCTATTATGCTGCCCTCTACTACCTCTGGCTGCCTGT
 GCCACGGCTGGCCACACAGCTGCACACCTGCTCGGCAGCACGCTGTCTGGGCCACCTT
 GGGGTCCAGGTCTGGCAGAGGGCAGAGTGTCCCAGGTGCCAAGATCTACAAGTACTAC
 TCCCTGCTGGCCCTCCCTGCCTCTCCTGCTGGGCTCGGATTCCTGAGCCTTTGGTACCT
 GTGCAGCTGGTGAGAAGCTTCAGCCGTAGGACAGGAGCAGGCTCCAAGGGGCTGCAGAGC
 AGCTACTCTGAGGAATATCTGAGGAACCTCCTTTGCAGGAAGAAGCTGGGAAGCAGCTAC
 CACACCTCCAAGCATGGCTTCTGTCTGGGCCCGCGTCTGCTTGAGACACTGCATCTAC
 ACTCCACAGCCAGGATTCCATCTCCTGCTGAAGCTGGTGCCTTCAGCTACACTGACAGGG
 ACGGCCATTTACCAGGTGGCCCTGCTGCTGCTGGTGGGCGTGGTACCCACTATCCAGAAG
 GTGAGGGCAGGGGTCAACACGGATGTCTCTACCTGCTGGCCGGCTTTGGAAATCGTGCTC
 TCCGAGGACAAGCAGGAGGTGGTGAGCTGGTGAAGCACCATCTGTGGGCTCTGGAAGTG
 TGCTACATCTCAGCCTTGGTCTTGTCTGCTTACTCACCCTCCTGGTCTGATGCGCTCA
 CTGGTGACACACAGGACCAACCTTCGAGCTCTGCACCGAGGAGCTGCCCTGGACTTGAGT
 CCTTGCATCGGAGTCCCCATCCCTCCCGCCAAGCCATATTCTGTTGGATGAGCTTCAGT
 GCCTACCAGACAGCCTTTATCTGCCTTGGGCTCCTGGTGCAGCAGATCATCTTCTTCTG
 GGAACACGGCCCTGGCCTTCTGCTGCTCATGCTGTGCTCCATGGCAGGAACCTCCTG
 CTCTTCCGTTCCCTCGAGTCTCTGCTGGCCCTTCTGGCTGACTTTGGCCCTGGCTGTGATC
 CTGCAGAACATGGCAGCCCATTTGGGTCTTCTGGAGATCATGATGGACACCCACAGCTG
 ACCAACCCGGCGAGTGCTCTATGCAGCCACCTTTCTTCTTCTTCCCCCTCAATGTGCTGGT
 GGTGCCATGGTGCCACCTGGCGAGTGCTCCTCTCTGCCCCCTACAACGCCATCCACCTT
 GGCCAGATGGACCTCAGCCTGCTGCCACCGAGAGCCGCCACTCTCGACCCCGGCTACTAC
 ACGTACCGAAACTTCTGAAGATTGAAGTCAGCCAGTCGCATCCAGCCATGACAGCCTTC
 TGCTCCCTGCTCCTGCAAGCGCAGAGCCTCCTACCCAGGACCATGGCAGCCCCCAGGAC
 AGCCTCAGACCAGGGGAGGAAGACGAAGGGATGCAGCTGCTACAGACAAAGGACTCCATG
 GCCAAGGGAGCTAGGCCCGGGGCCAGCCGCGGCAGGGCTCGCTGGGGTCTGGCCTACACG
 CTGCTGCACAACCCAACCTGCAGGTCTTCCGCAAGACGGCCCTGTTGGGTGCCAATGGT
 GCCCAGCCCTGAGGGCAGGGGAAGGTCAACCCACCTGCCCATCTGTGCTGAGGCATGTTCC
 TGCTTACCATCCTCCTCCCTCCCGGCTCTCCTCCAGCATCACACCAGCCATGCAGCCA
 GCAGGTCTCCTCGGATCACTGTGGTTGGGTGGAGGTCTGTCTGCACTGGGAGCCTCAGGAG
 GGCTCTGCTCCACCCACTTGGCTATGGGAGAGCCAGCAGGGGTCTGGAGAAAAAACTG
 GTGGGTAGGGCCTTGGTCCAGGAGCCAGTTGAGCCAGGGCAGCCACATCCAGGCGTCTC
 CTAACCTGGCTCTGCCATCAGCCTTGAAGGGCTCGATGAAGCCTTCTCTGGAACCACT
 CCAGCCCAGCTCCACCTCAGCCTTGGCCTTACGCTGTGGAAGCAGCCAAGGCACTTCCT
 CACCCCTCAGCGCCACGGACCTCTCTGGGAGTGGCCGGAAAGCTCCCGGTCTCTGCG
 CTGCAGGGCAGCCCAAGTCATGACTCAGACCAGGTCCACACTGAGCTGCCACACTCGA
 GAGCCAGATATTTTGTAGTTTTTATGCTTTTGGCTATTATGAAAGAGGTTAGTGTGTTT
 CTTGCAATAAACTTGTTCCTGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA
 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA

图 1

MSSQPAGNQTS PGATEDYSYGSWYIDEPQGGEELQPEGEVPSCHTSIPPGLYHACLASLS
 ILVLLLLAMLVRRRQLWPDCVRGRPGLSPVDFLAGDRPRAVPAAVFMVLLSSLCLLLPD
 EDALPFLTLASAPSQDGKTEAPRGAWKILGLFYAALYYPLAACATAGHTAAHLLGSTLS
 WAHLGVQVWQRAECPQVPKIYKYSSLLASLPLLLGLGLSLWYPVQLVRSFSRRTGAGSK
 GLQSSYSEEYLRNLLCRKKLGSSYHTSKHGFLSWARVCLRHCIYTPQPGFHLPLKLVLSA
 TLTGTAIYQVALLLLVGVVPTIQKVRAGVTTDVSYLLAGFGIVLSEDKQEVVELVKHHLW
 ALEVCYISALVLSCLLTFLVLMRSLVTHRTNLRALHRGAALDLSPLHRSPHPSRQAIFCW
 MSFSAYQTAFICLGLLVQQIIFFLGTTALAFVLMPVLHGRNLLFRSLESSWPFWLTLA
 LAVILQNMAAHWVFLETHDHPQLTNRRLVLYAATFLLFPLNVLVGAMVATWRVLLSALYN
 AIHLGQMDLSLLPPRAATLDPGYTYRNFLKIEVSQSHPAMTAFCSLLLQAQSLLPRTMA
 APQDSL RPGEDEGMQLLQTKDSMAKGARPGASRGARWGLAYTLLHNPTLQVFRKTALL
 GANGAQP

蛋白的重要特征:

信号肽:

无

跨膜结构域:

54-69
 102-119
 148-166
 207-222
 301-320
 364-380
 431-451
 474-489
 560-535

基序排列:

基序名称: N-糖基化位点

8-12

基序名称: N-豆蔻酰化位点

50-56
 176-182
 241-247
 317-323
 341-347
 525-531
 627-633
 631-637
 640-646
 661-667

基序名称: 原核膜脂蛋白脂质附着位点

364-375

基序名称: ATP/GTP-结合位点基序A (P-环)

132-140

图 2

PRO	XXXXXXXXXXXXXXXXXX	(长度 = 15个氨基酸)
对比蛋白	XXXXXXXXYYYYYYY	(长度 = 12个氨基酸)

%氨基酸序列同一性 =

(通过ALIGN-2确定的两个肽序列之间相同或匹配的氨基酸残基数) 除以
(PRO多肽的氨基酸残基总数) =

5除以15 = 33.3 %

图 3A

PRO XXXXXXXXXXXX (长度 = 10个氨基酸)
对比蛋白 XXXXXYYYYYYZZYZ (长度 = 15个氨基酸)

%氨基酸序列同一性 =

(通过ALIGN-2确定的两个肽序列之间相同或匹配的氨基酸残基数) 除以
(PRO多肽的氨基酸残基总数) =

5除以10 = 50 %

图 3B

PRO-DNA 核苷酸)	NNNNNNNNNNNNNNNN	(长度 = 14
对比 DNA 核苷酸)	NNNNNNLLLLLLLLLLLL	(长度 = 16

%核酸序列同一性 =

(通过ALIGN-2确定的两个核酸序列之间相同式匹配的核苷酸数) 除以
(PRO-DNA核酸序列的核苷酸总数) =

6除以14 = 42.9 %

图 3C

PRO-DNA	NNNNNNNNNNNNNN	(长度 = 12核苷酸)
对比 DNA	NNNNLLLLVV	(长度=9核苷酸)

%核酸序列同一性 =

(如ALIGN-2确定的两个核酸序列之间相同或匹配的核苷酸数) 除以
(PRO-DNA核酸序列的核苷酸总数) =

4除以12 = 33.3 %

图 3D

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int  _day[26][26] = {
/*   A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, 2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
0, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

图 4A

```

/*
 */
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX          4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT         3      /* value of matching bases */
#define DMIS         0      /* penalty for mismatched bases */
#define DINSO        8      /* penalty for a gap */
#define DINS1         1      /* penalty per base */
#define PINSO         8      /* penalty for a gap */
#define PINS1         4      /* penalty per residue */

struct jmp {
    short          n[MAXJMP];    /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP];    /* base no. of jmp in seq x */
};
/* limits seq to 2^16 -1 */

struct diag {
    int            score;        /* score at last jmp */
    long           offset;       /* offset of prev block */
    short          ijmp;        /* current jmp index */
    struct jmp     jmp;         /* list of jmps */
};

struct path {
    int            spc;          /* number of leading spaces */
    short          n[JMPS];      /* size of jmp (gap) */
    int            x[JMPS];      /* loc of jmp (last elem before gap) */
};

char              *ofile;        /* output file name */
char              *namex[2];     /* seq names: getseqs() */
char              *prog;         /* prog name for err msgs */
char              *seqx[2];      /* seqs: getseqs() */
int              dmax;           /* best diag: nw() */
int              dmax0;          /* final diag */
int              dna;            /* set if dna: main() */
int              endgaps;        /* set if penalizing end gaps */
int              gapx, gapy;      /* total gaps in seqs */
int              len0, len1;      /* seq lens */
int              ngapx, ngapy;    /* total size of gaps */
int              smax;           /* max score: nw() */
int              *xbm;           /* bitmap for matching */
long             offset;         /* current offset in jmp file */
struct diag      *dx;            /* holds diagonals */
struct path      pp[2];          /* holds paths for seqs */

char              *calloc(), *malloc(), *index(), *strcpy();
char              *getseq(), *g_calloc();

```

Page 1 of nw.h

图 4B

```

/* Needleman-Wunsch alignment program
 *
 * usage: prog file1 file2
 * where file1 and file2 are two dna or two protein sequences.
 * The sequences can be in upper- or lower-case and may contain ambiguity
 * Any lines beginning with ';', '>' or '<' are ignored
 * Max file length is 65535 (limited by unsigned short x in the jmp struct)
 * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 * Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
    int    ac;
    char   *av[];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                /* 1 to penalize endgaps */
    ofile = "align.out";        /* output file */

    nw();                       /* fill in the matrix, get the possible jmps */
    readjmps();                 /* get the actual jmps */
    print();                    /* print stats, alignment */

    cleanup(0);                 /* unlink any tmp files */
}

```

main

Page 1 of nw.c

图 4C

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw0
{
    char          *px, *py;          /* seqs and ptrs */
    int            *ndely, *dely;     /* keep track of dely */
    int            ndelx, delx;       /* keep track of delx */
    int            *tmp;              /* for swapping row0, row1 */
    int            mis;               /* score for each type */
    int            ins0, ins1;        /* insertion penalties */
    register       id;                /* diagonal index */
    register       ij;               /* jmp index */
    register       *col0, *col1;      /* score for curr, last row */
    register       xx, yy;            /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = 'yy';
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
    */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

Page 2 of nw.c

图 4D

```

...nw
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongoing del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongoing del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */
}

```

图 4E

Page 3 of nw.c

```

...DW
    id = xx - yy + len1 - 1;
    if (mis >= delx && mis >= dely[yy])
        coll[yy] = mis;
    else if (delx >= dely[yy]) {
        coll[yy] = delx;
        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINSO)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writejumps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
                offset += sizeof(struct jmp) + sizeof(offset);
            }
            dx[id].jp.n[ij] = ndelx;
            dx[id].jp.x[ij] = xx;
            dx[id].score = delx;
        }
        else {
            coll[yy] = dely[yy];
            ij = dx[id].ijmp;
        }
        if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINSO)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writejumps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
                offset += sizeof(struct jmp) + sizeof(offset);
            }
            dx[id].jp.n[ij] = -ndely[yy];
            dx[id].jp.x[ij] = xx;
            dx[id].score = dely[yy];
        }
        if (xx == len0 && yy < len1) {
            /* last col
            */
            if (endgaps)
                coll[yy] -= ins0+ins1*(len1-yy);
            if (coll[yy] > smax) {
                smax = coll[yy];
                dmax = id;
            }
        }
    }
    if (endgaps && xx < len0)
        coll[yy-1] -= ins0+ins1*(len0-xx);
    if (coll[yy-1] > smax) {
        smax = coll[yy-1];
        dmax = id;
    }
    tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);(void) free((char *)col1);}

```

Page 4 of nw.c

图 4F

```

/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq. [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC      3
#define P_LINE  256      /* maximum output line */
#define P_SPC    3        /* space between name or num and seq */

extern  _day[26][26];
int     olen;             /* set output line length */
FILE    *fx;              /* output file */

print()
{
    int     lx, ly, firstgap, lastgap;      /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                getmat
int    lx, ly;          /* "core" (minus endgaps) */
int    firstgap, lastgap; /* leading trailing overlap */
{
    int    nm, i0, i1, siz0, siz1;
    char    outx[32];
    double    pct;
    register    n0, n1;
    register char    *p0, *p1;

    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbm[*p0-'A'] & xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100. * (double)nm / (double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
}

```

Page 2 of nwprint.c

图 4H

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base": "residue", (ngapx == 1)? "" : "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base": "residue", (ngapy == 1)? "" : "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINSO, DINSI);
    else
        fprintf(fx,
            "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINSO, PINSI);
    if (endgaps)
        fprintf(fx,
            "< endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
            lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
    else
        fprintf(fx, "< endgaps not penalized\n");
}

static      nn;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char  *ps[2];      /* ptr to current element */
static char  *po[2];      /* ptr to next output char slot */
static char  out[2][P_LINE]; /* output line */
static char  star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nn;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(name[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr_align

Page 3 of nwprint.c

图 4I

```

for (nn = nm = 0, more = 1; more;) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        else { /* we're putting a seq element
            */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;

            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i] + 1];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i] + 1];
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nm) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nm = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
}

```

...pr_align

dumpblock

Page 4 of nwprint.c

图 4J

```

...dumpblock

(void)putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[i])
            stars();
        putline(i);
        if (i == 0 && *out[i])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)                                nums
{
    int      ix;      /* index in out[] holding seq line */

    char      nline[P_LINE];
    register  i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void)putc(*pn, fx);
    (void)putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)                                putline
{
    int      ix;

```

图 4K

```

...putline

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(p0[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(p0[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

```

图 4L

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)                                stripname
char *pn; /* file name (may be path) */
{
    register char *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
```

图 4M

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_alloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char    *jname = "/tmp/homgXXXXXX";          /* tmp file for jumps */
FILE    *fj;

int      cleanup();                          /* cleanup tmp file */
long     lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)                                  cleanup
int      i;
{
    if (fj)
        (void) unlink(jname);
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                          getseq
char    *file; /* file name */
int      *len; /* seq len */
{
    char    line[1024], *pseq;
    register char    *px, *py;
    int      natgc, tlen;
    FILE    *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
}

```

Page 1 of nwsubr.c

图 4N

```

...getseq

py = pseq + 4;
*len = llen;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            naigc++;
    }
    *py++ = '\0';
    *py = '\0';
    (void) fclose(fp);
    dna = naigc > (llen/3);
    return(pseq+4);
}

char *
g_alloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;          /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
        }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = lcn0; ; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)

```

Page 2 of nwsubr.c

```

...readjumps

    if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
    fprintf(stderr, "%s: too many gaps in alignment\n", prog);
    cleanup(1);
}
if (j >= 0) {
    siz = dx[dmax].jp.n[j];
    xx = dx[dmax].jp.x[j];
    dmax += siz;
    if (siz < 0) { /* gap in second seq */
        pp[1].n[i1] = -siz;
        xx += siz;

        /* id = xx - yy + len1 - 1
        */
        pp[1].x[i1] = xx - dmax + len1 - 1;
        gapy++;
        ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
/* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
}
else
    break;
}

/* reverse the order of jumps
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;}}

```

Page 3 of nwsubr.c

图 4P

```
/*  
 * write a filled jmp struct offset of the prev one (if any): nw()  
 */  
writejmps(ix)                                writejmps  
{  
    int ix;  
    char *mktemp();  
    if (!fj) {  
        if (mktemp(jname) < 0) {  
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);  
            cleanup(1);  
        }  
        if ((fj = fopen(jname, "w")) == 0) {  
            fprintf(stderr, "%s: can't write %s\n", prog, jname);  
            exit(1);  
        }  
    }  
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);  
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);  
}
```

图 4Q

GTGCTCTCCGAGGACAAGCAGGAGGNGGTGGAGCTGGTGAAGCACCATCTGTGGGCTCTG
GAAGTGTGCTACATCTCAGCCTTGGTCTTGTCTGCTTACTCACCTTCCTGGTCCTGATG
CGCTCACTGGTGACACACAGGACCAACCTTCGAGCTCTGCACCGAGGAGCTGCCCTGGAC
TTGAGTCCCTTGCAATCGGAGTCCCCATCCCTCCCGCCAAGCCATATTCTGTTGGATGAGC
TTCAGTGCCTACCAGACAGCCTTTATCTGCCTTGGGCTCCTGGTGCAGCAGATCATCTTC
TTCCTGGGAACCACGGCCCTGGCCTTCCTGGTGCTCATGCCTGTGCTCCATGGCAGGAAC
CTCCTGCTCTTCCGTTCCTGGAGTCCTCGTGGCCCTTCTGGCTGACTTTGGCCCTGGCT
GTGATCCTGCAGAACATGGCAGCCCATTGGGTCTTCCTGGAGACTCATGATGGACACCCA
CAGCTGACCAACCGGCGAGTGCTCTATGCAGCCACCTTCTTCTTCCCCCTCAATGTG
CTGGTGGGTGCCATGGTGGCCACCTGGCGAGTGCTCCTCTCTGCCCTCTACAACGCCATC
CACCTTGGCCAGATGGACCTCAGCCTGCTGCCACCGAGAGCCGCCACTCTCGACCCCGGC
TACTACACGTACCGAA

图 5

CACAACCAGCCACCCCTCTAGGATCC¹AGCCAGCTGGTGCTGGGCTCAGAGGAGAAGGC
CCCGTGTGGGAGCACCCCTGCTTGCCTGGAGGGACAAGTTTCCGGGAGAGATCAATAAAG
GAAAGGAAGAGACAAGGAAGGGAGAGGTCAGGAGAGCGCTTGATTGGAGGAGAAGGGCC
AGAGAATGTCGTCCCAGCCAGCAGGGAACCAGACCTCCCCGGGGCCACAGAGGACTACT
CCTATGGCAGCTGGTACATCGATGAGCCCCAGGGGGCGAGGAGCTCCAGCCAGAGGGGG
AAGTGCCCTCCTGCCACACCAGCATACCACCCGGCCTGTACCACGCCTGCCTGGCCTCGC
TGTCAATCCTTGTGCTGCTGCTCCTGGCCATGCTGGTGAGGCGCCGCCAGCTCTGGCCTG
ACTGTGTGCGTGGCAGGCCCCGGCCTGCCAGGCCCCGGGCAGTGCTGCTGCTGTTTTCA
TGGTCCTCCTGAGCTCCCTGTGTTTGTGCTCCCCGACGAGGACGCATTGCCCTTCCTGA
CTCTCGCCTCAGCACCCAGCCAAGATGGGAAACTGAGGCTCCAAGAGGGGGCCTGGAAGA
TACTGGGACTGTTCTATTATGCTGCCCTCTACTACCTCTGGCTGCCTGTGCCACGGCTG
GCCACACAGCTGCACACCTGCTCGGCAGCACGCTGTCTGGGCCACCTTGGGGTCCAGG
TCTGGCAGAGGGCAGAGTGTCCCCAGGTGCCAAGATCTACAAGTACTACTCCCTGCTGG
CCTCCCTGCCTCTCCTGCTGGGCCTCGGATTCTTGAGCCTTTGGTACCCTGTGCAGCTGG
TGAGAAGCTTCAGCCGTAGGACAGGAGCAGGCTCCAAGGGGCTGCAGAGCAGCTACTCTG
AGGAATATCTGAGGAACCTCCTTTGCAGGAAGAAGCTGGGAAGCAGCTACCACACCTCCA
AGCATGGCTTCCTGTCTGGGCCCGCGTCTGCTTGAGACACTGCATCTACACTCCACAGC
CAGGATTCCATCTCCCGCTGAAGCTGGTGCTTTCAGCTACACTGACAGGGACGGCCATTT
ACCAGGTGGCCCTGCTGCTGCTGGTGGGCGTGGTACCCACTATCCAGAAGGTGAGGGCAG
GGGTACCACGGATGTCTCCTACCTGCTGGCCGGCTTTGGAATCGTGCTCTCCGAGGACA
AGCAGGAGGTGGTGGAGCTGGTGAAGCACCATCTGTGGGCTCTGGAAGTGTGCTACATCT
CAGCCTTGGTCTTGTCTGCTTACTCACCTTCCTGGTCCTGATGCGCTCACTGGTGACAC
ACAGGACCAACCTTCGAGCTCTGCACCGAGGAGCTGCCCTGGACTTGAGTCCCTTGATC
GGAGTCCCCATCCCTCCCGCCAAGCCATATTCTGTTGGATGAGCTTCAGTGCTTACCAGA
CAGCCTTTATCTGCCTTGGGCTCCTGGTGCAGCAGATCATCTTCTTCCCTGGGAACACGG
CCCTGGCCTTCCTGGTGTCTATGCCTGTGCTCCATGGCAGGAACCTCCTGCTCTTCCGT
CCCTGGAGTCCCTCGTGGCCCTTCTGGCTGACTTTGGCCCTGGCTGTGATCCTGCAGACA
TGGCAGCCCATTGGGTCTTCCCTGGAGACTCATGATGGACACCCACAGCTGACCAACCGGC
GAGTGCTCTATGCAGCCACCTTTCTTCTTCCCCCTCAATGTGCTGGTGGGTGCCATAG
TGGCCACCTGGCGAGTGCTCCTCTCTGCCCTCTACAACGCCATCCACCTTGGCCAGATGG
ACCTCAGCCTGCTGCCACCGAGAGCCGCCACTCTCGACCCCGGCTACTACACGTACCGAA
ACTTCTTGAAGATTGAAGTCAGCCAGTCGCATCCAGCCATGACAGCCTTCTGCTCCCTGC
TCCTGCAAGCGCAGAGCCTCCTACCCAGGACCATGGCAGCCCCCAGGACAGCCTCAGAC
CAGGGGAGGAAGACGAAGGGATGCAGCTGCTACAGACAAAGGACTCCATGGCCAAGGGAG
CTAGGCCCCGGGGCCAGCCGCGGCAGGGCTCGCTGGGGTCTGGCCTACACGCTGCTGCACA
ACCCAACCCTGCAGGTCTTCCGCAAGACGGCCCTGTTGGGTGCCAATGGTGCCAGCCCT
GAGGGCAGGGAAGGTCAACCCACCTGCCCATCTGTGCTGAGGCATGTTCTTGCCTACCAC
CTCCTCCCTCCCCGGCTCTCCTCCCAGCATCACACCAGCCATGCAGCCAGCAGGTCTCTC
GGATCACTGTGTTGGGTGGAGGTCTGTCTGCACTGGGAGCCTCAGGAGGGCTCTGCTCC
ACCCACTTGGCTATGGGAGAGCCAGCAGGGGTCTGGAGAAAGAACTGGTGGGTAGGG
CCTTGGTCCAGGAGCCAGTTGAGCCAGGGCAGCCACATCCAGGCGTCTCCCTACCCTGGC
TCTGCCATCAGCCTTGAAGGGCCTCGATGAAGCCTTCTCTGGAACCACTCCAGCCCAGCT
CCACCTCAGCCTTGGCCTTCACGCTGTGGAAGCAGCCAAGGCACTTCCTACCCCCCTCAG
CGCCACGGACCTCTCTGGGGAGTGGCCGGAAGCTCCCGGGCCTCTGGCCTGCAGGGCAG
CCCAAGTCATGACTCAGACCAGGTCCCACACTGAGCTGCCACACTCGAGAGCCAGATAT
TTTTGTAGTTTTTATGCCTTTGGCTATTATGAAAGAGGTTAGTGTGTTCCCTGCAATAAA
CTTGTTCTTGAGAAAA

图 6

MSSQPAGNQTSPGATEDYSYGSWYIDEPQGGEELQPEGEVPSCHTSIPPGLYHACLASL
 SILVLLLLLAMLVRRRQLWPDVCVRGRPGGLPRPRAVPAAVFMVLLSSLCLLLPDEDALPFL
 TLASAPSQDGKTEAPRGAWKILGLFYAAALYYPLAACATAGHTAAHLLGSTLSWAHLGV
 QVWQRAECPQVPKIYKYSSLLASLPLLLGLGFLSLWYPVQLVRSFSRRTGAGSKGLQSS
 YSEEYLRNLLCRKKLGSSYHTSKHGFLSWARVCLRHCIYTPQPGFHLPLKLVLSTLTG
 TAIYQVALLLLVGVVPTIQKVRAGVTTDVSYLLAGFGIVLSEDKQEVVELVKHHLWALE
 VCYISALVLSCLLTFLVLMRSLVTHRTNLRALHRGAALDLSPLHRSPHPSRQAIFCWMS
 FSAYQTAFICLGLLVQQIIFFLGTTALAFLVLMFVLHGRNLLLFPSLESSWPFWLTAL
 AVILQNMAAHWVFLETHDGHPLTNRRVLYAATFLLFPLNVLVGAIVATWRVLLSALYN
 AIHLGQMDLSLLPPRAATLDPGYTYRNFLEKIEVSQSHPAMTAFCSLLLQAQSLLPRTM
 AAPQDSLRLPGEEDEGMQLLQTKDSMAKGPASRGARWGLAYTLLHNPTLQVFRKTA
 LLGANGAQP

蛋白的重要特征:

信号肽:

无

跨膜结构域:

54-71

93-111

140-157

197-214

291-312

356-371

425-444

464-481

505-522

基序名称: N-糖基化位点

8-12

基序名称: N-豆蔻酰化位点

50-56

167-173

232-238

308-314

332-338

516-522

618-624

622-628

631-637

652-658

基序名称: 原核膜脂蛋白脂质附着位点

355-366

基序名称: ATP/GTP-结合位点基序A (P-环)

123-131

图 7

Stra6 变体克隆

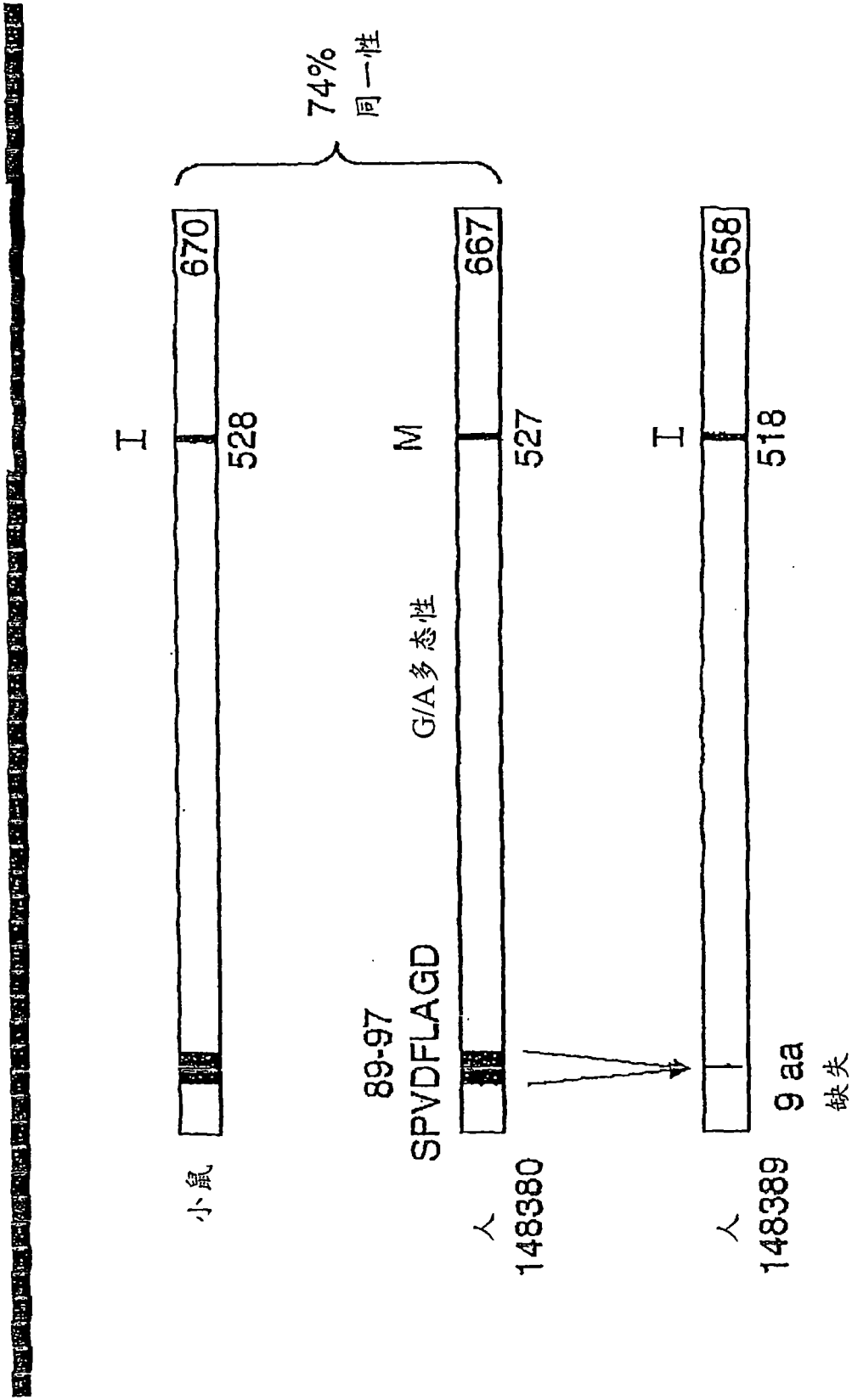
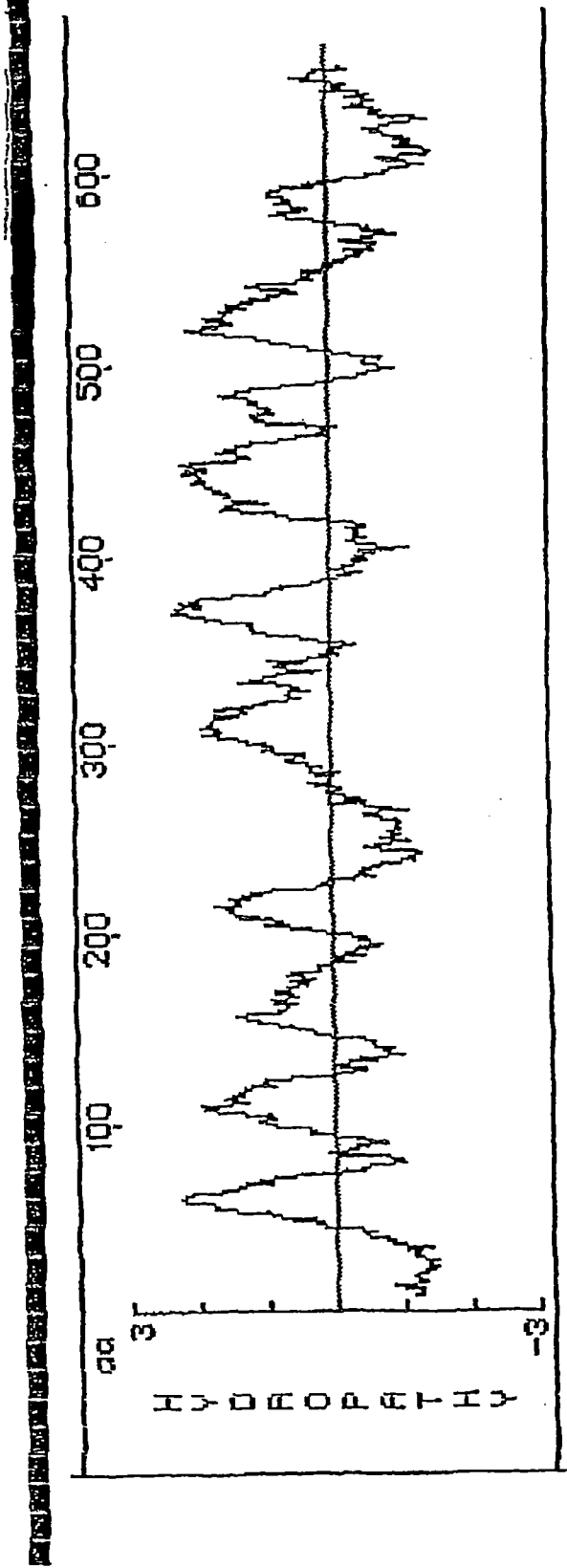


图 8

人 Stra6 的疏水图谱



- 3 kb mRNA
- 667 个氨基酸→50% 疏水性残基
- 73.5 kDa 蛋白
- 9 个潜在的跨膜结构域

图 9

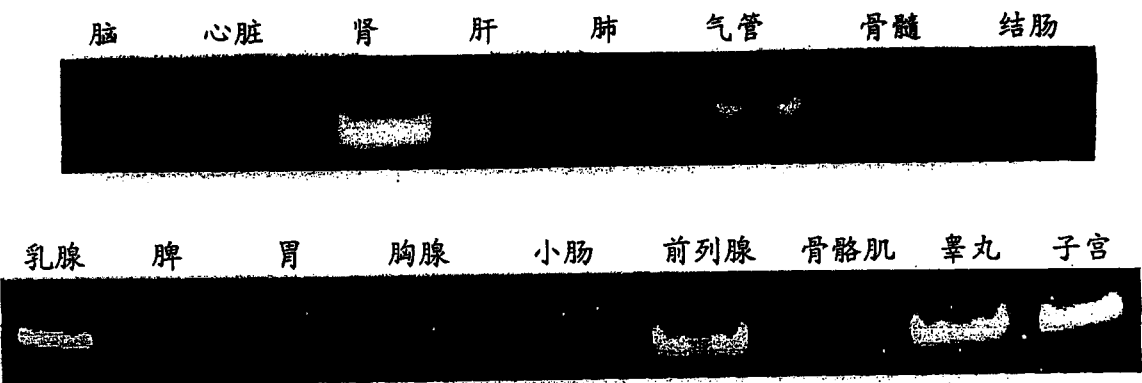


图 10

人结肠肿瘤组织中Stra6 RNA的表达

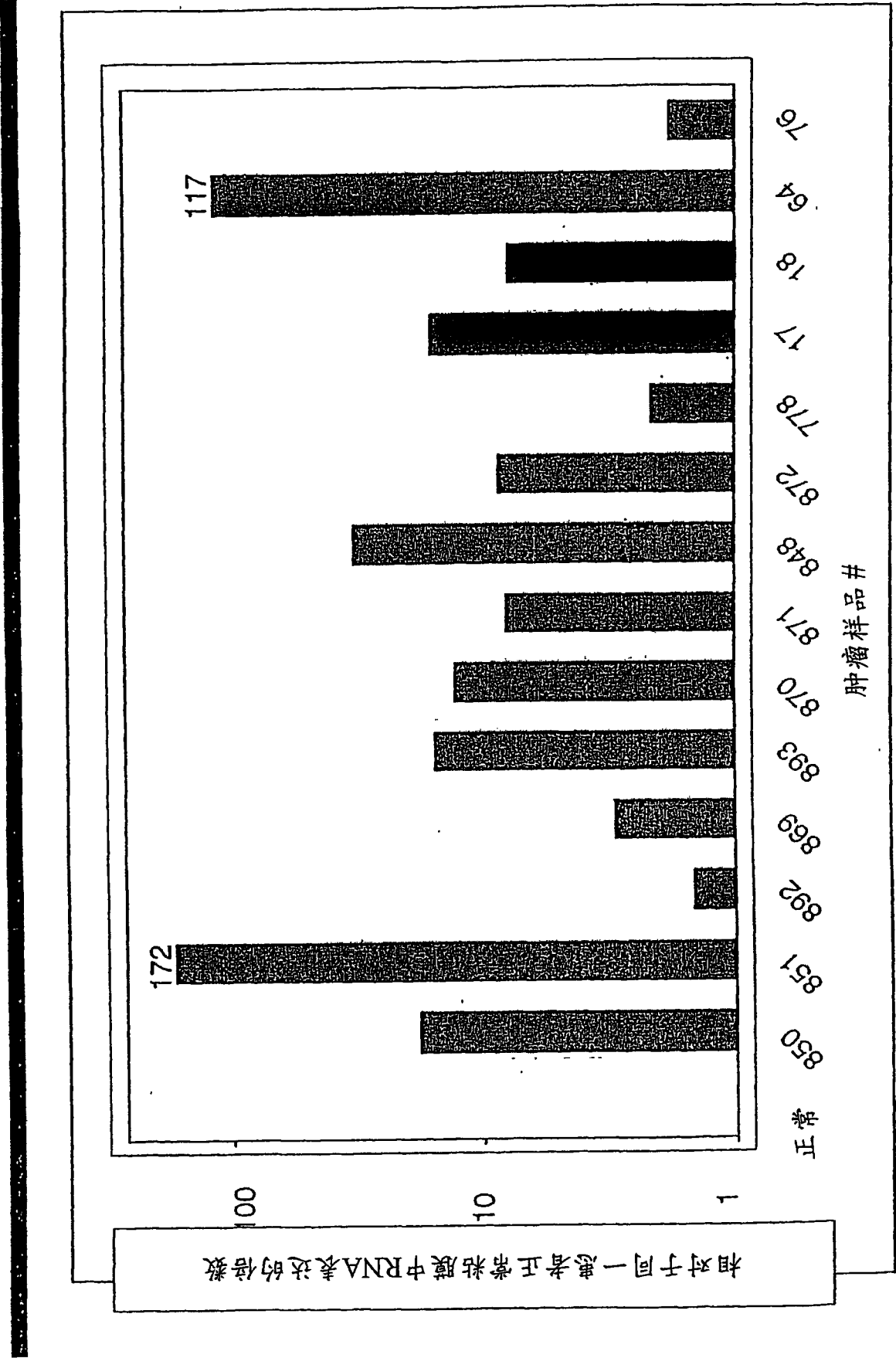
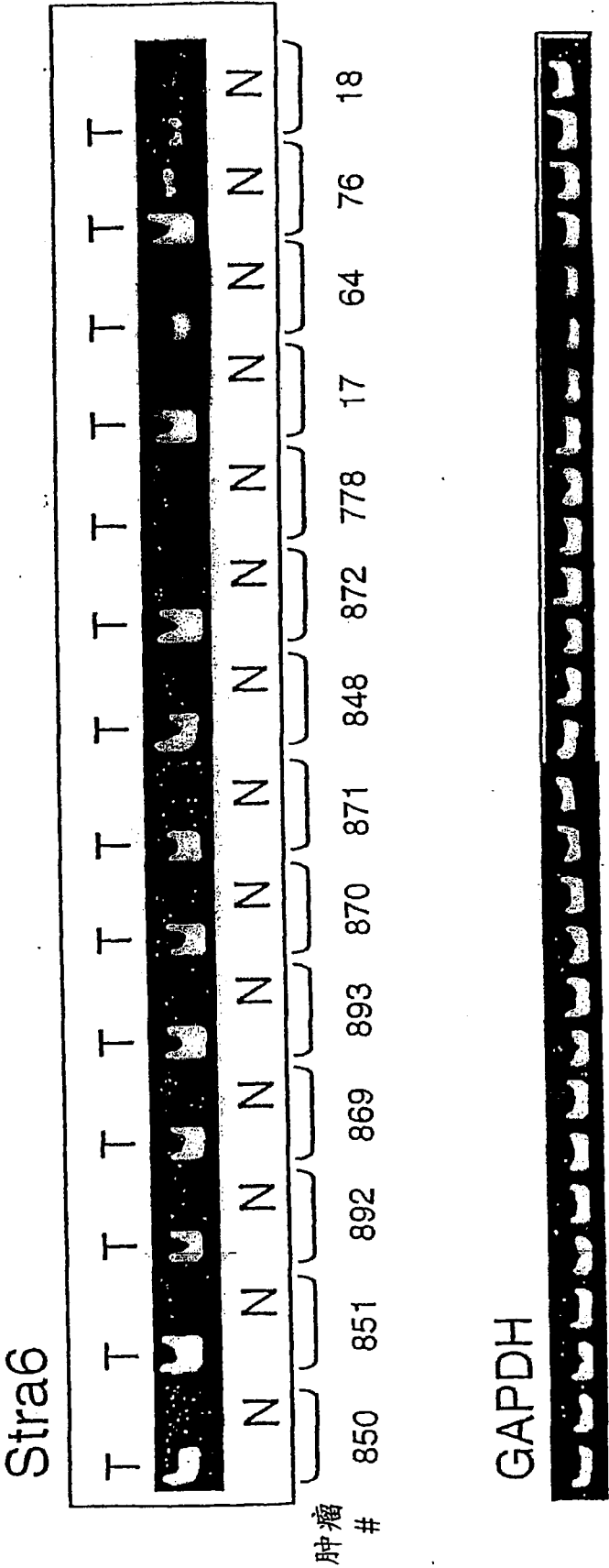


图 11

人结肠肿瘤组织中相对于同一患者正常粘膜中的Stra 6 RNA 的表达

40个循环后的Taqman 产品分析



WJ TM #84 h.Stra6 引物组 #4 1/4/00

图 12A

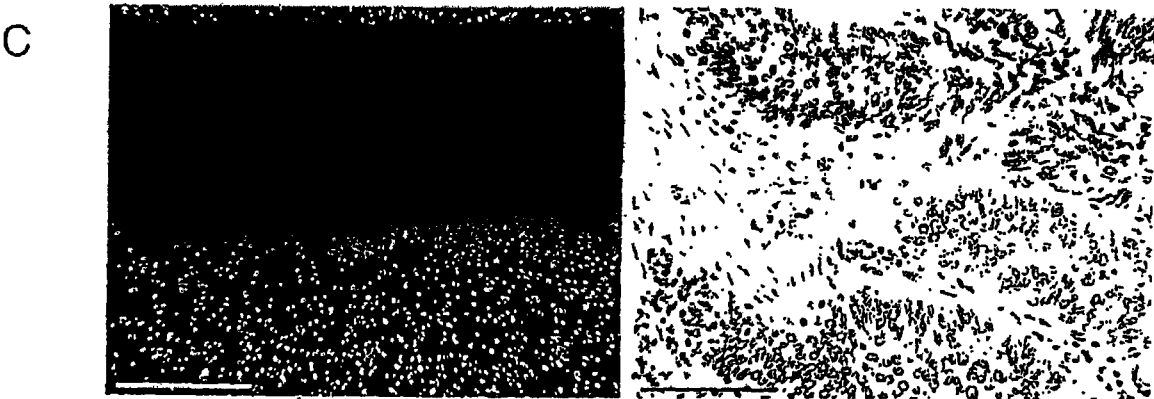


图 12B

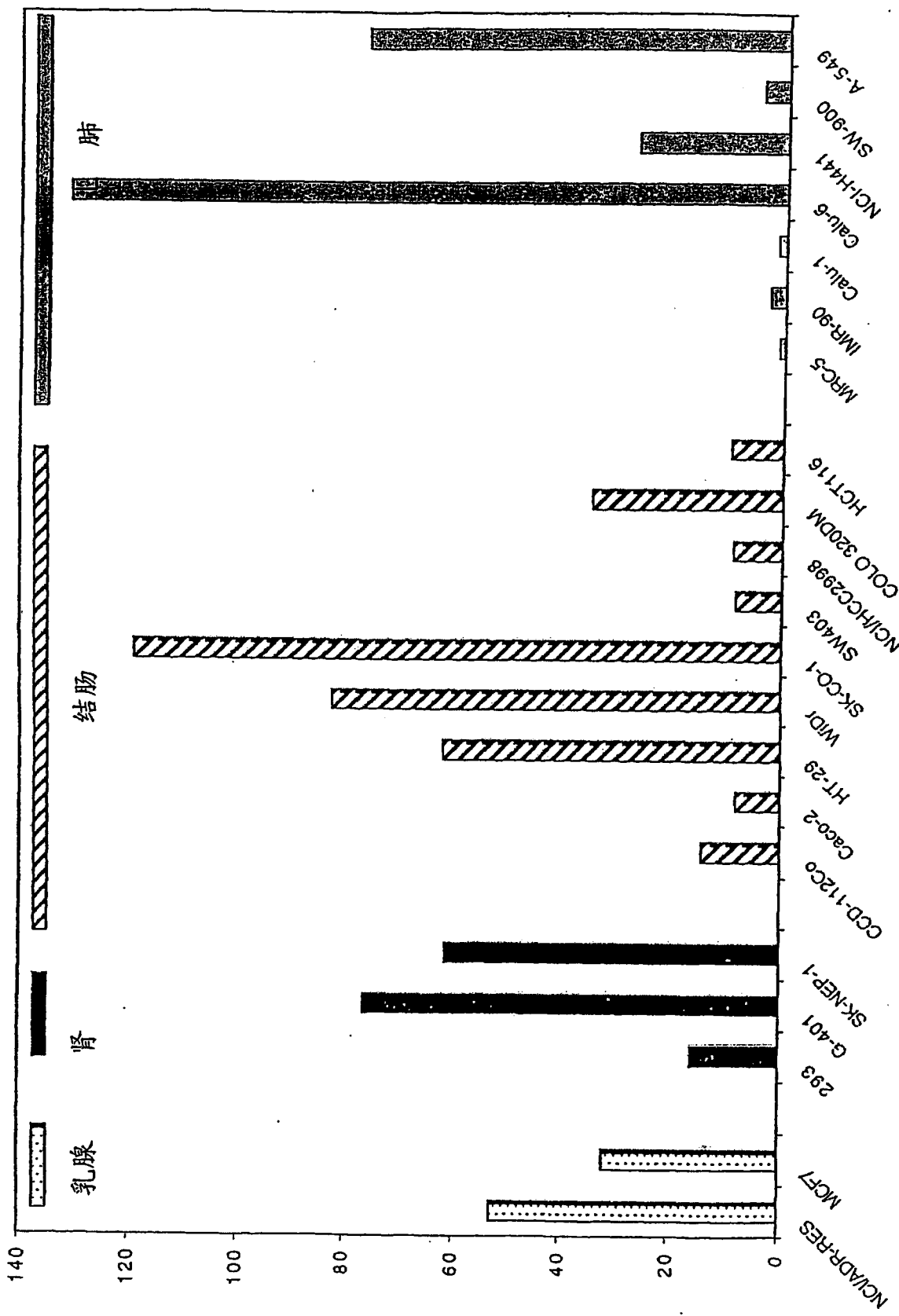


图 13

大肠杆菌中Stra6 肽的表达

N末端Poly-His可裂解性前导序列

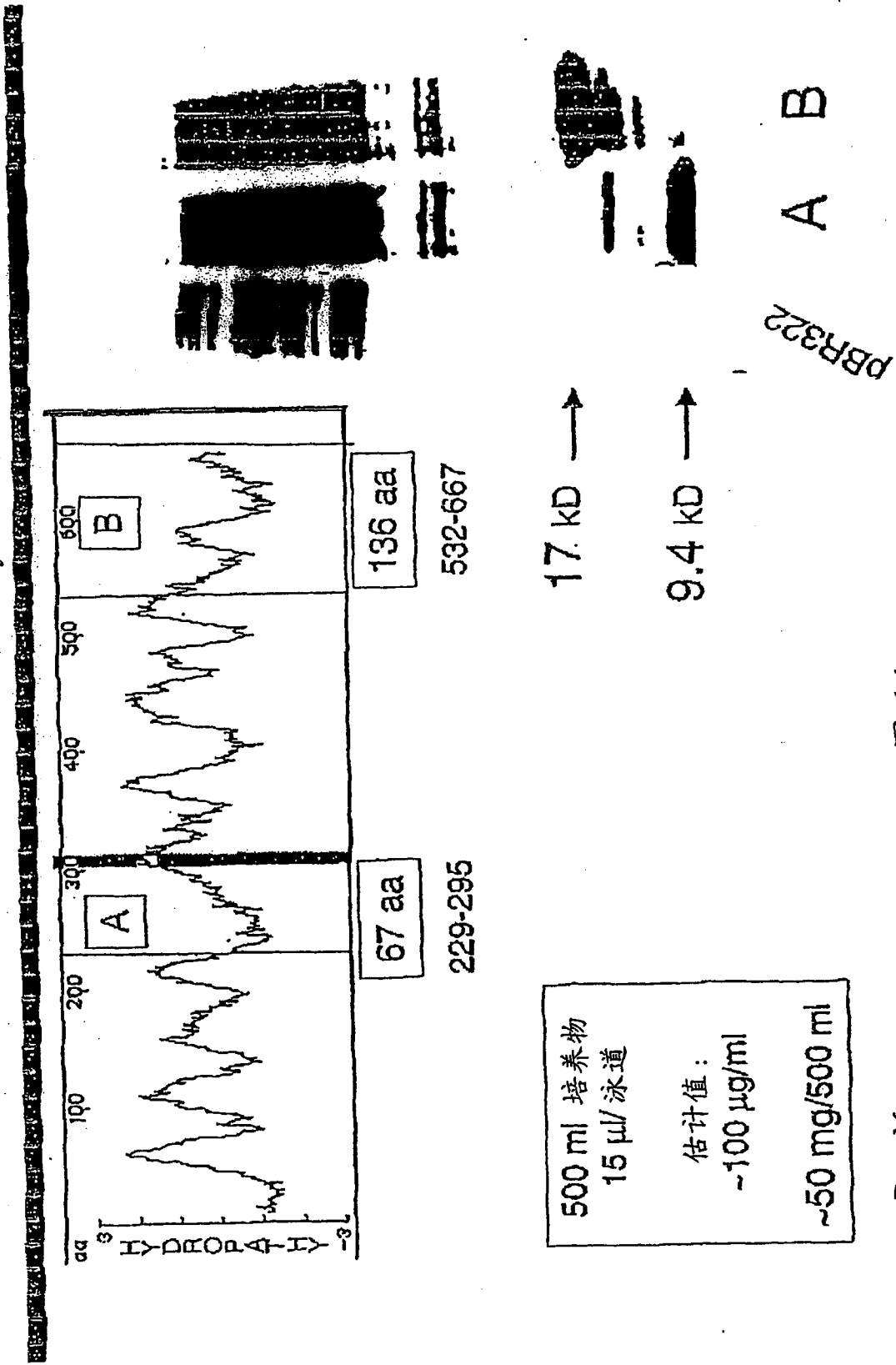


图 14

Dan Yansura

加或不加视黄酸的人结肠癌细胞中Stra6 RNA的表达

TM#75 (2/28/00)
VD3 - 维生素 D3 (1 μ M) ; ATRA - 全-反-视黄酸 (1 μ M)
9cRA - 9-顺-视黄酸

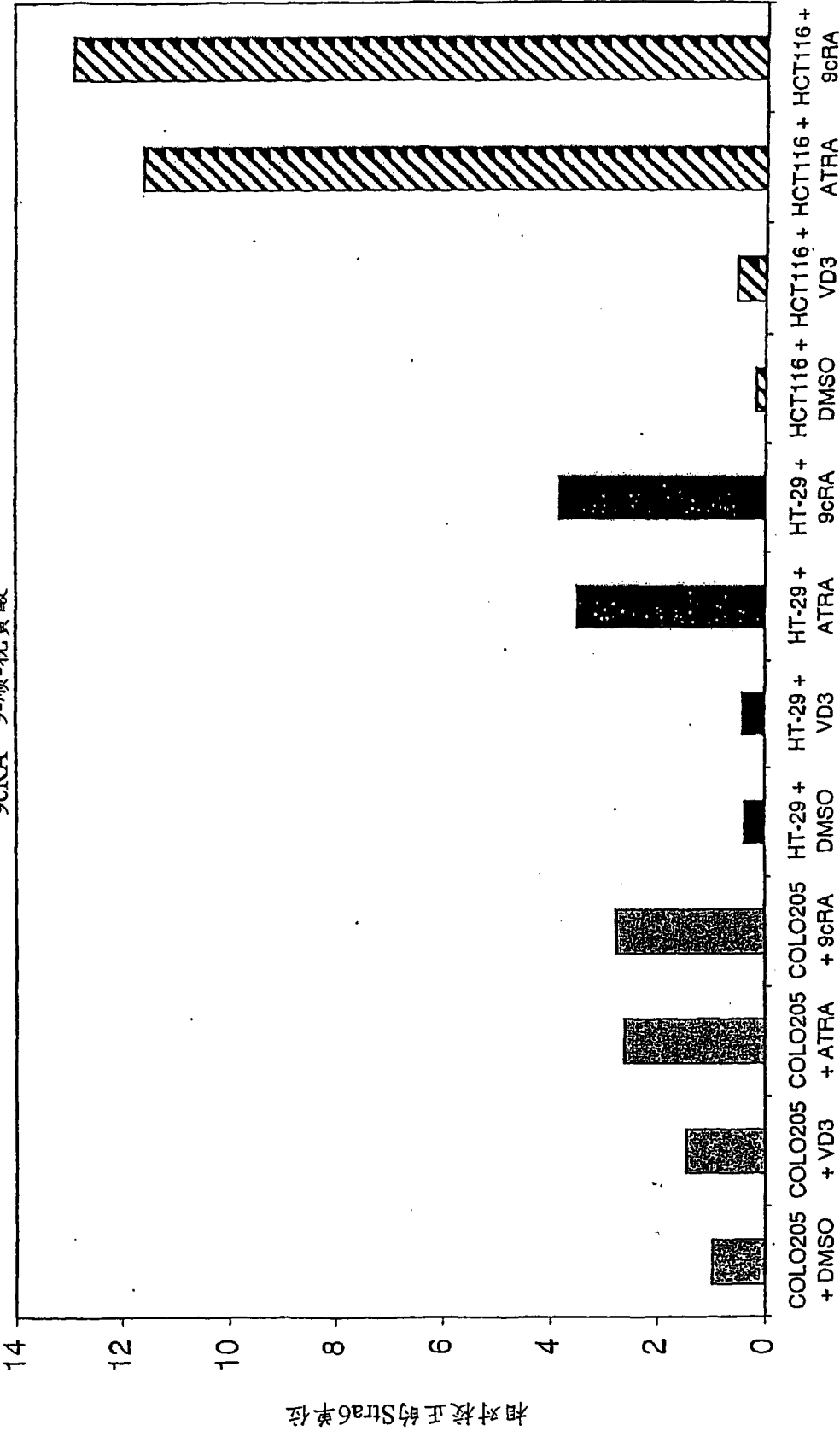


图 15

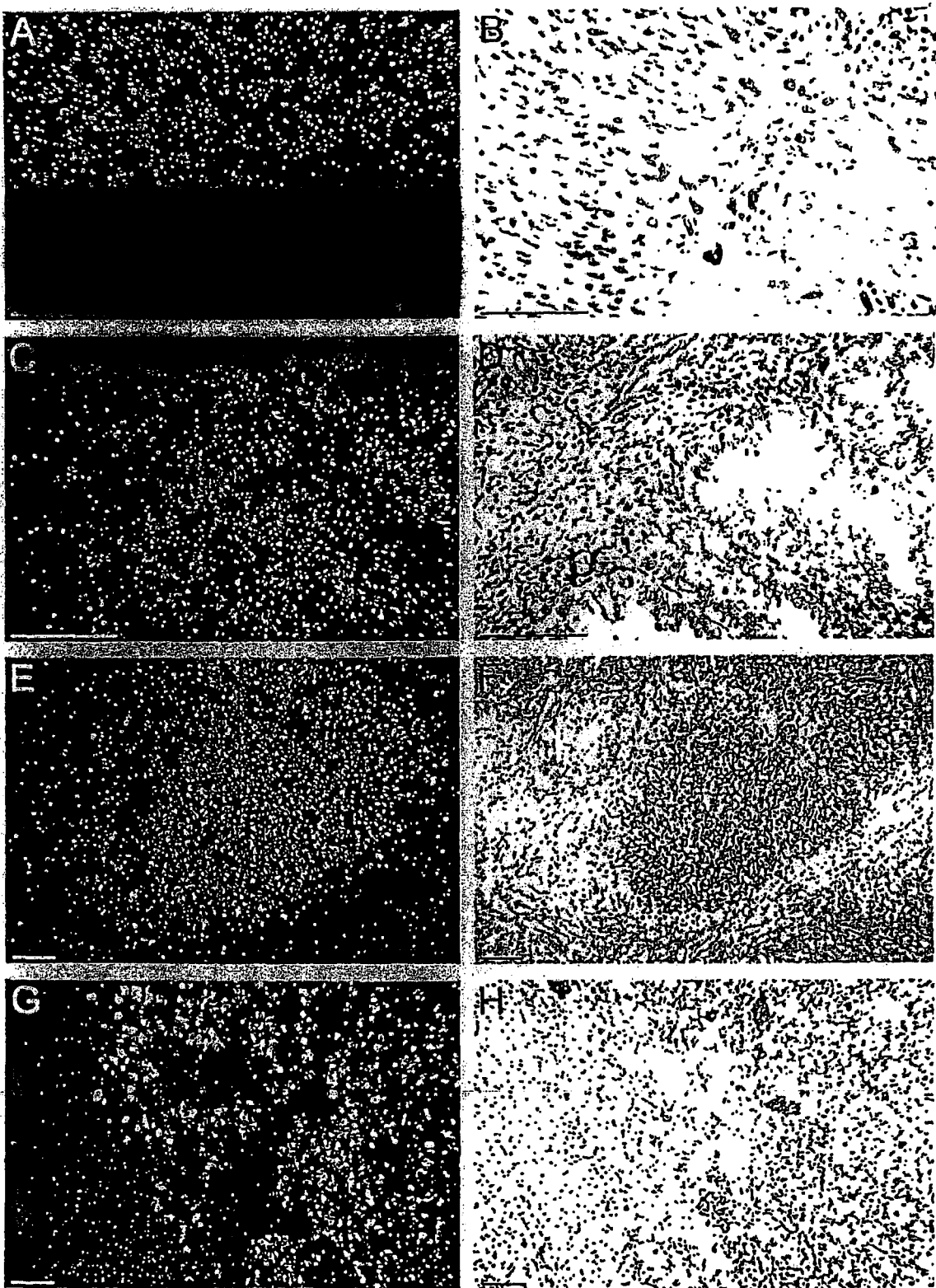


图 16

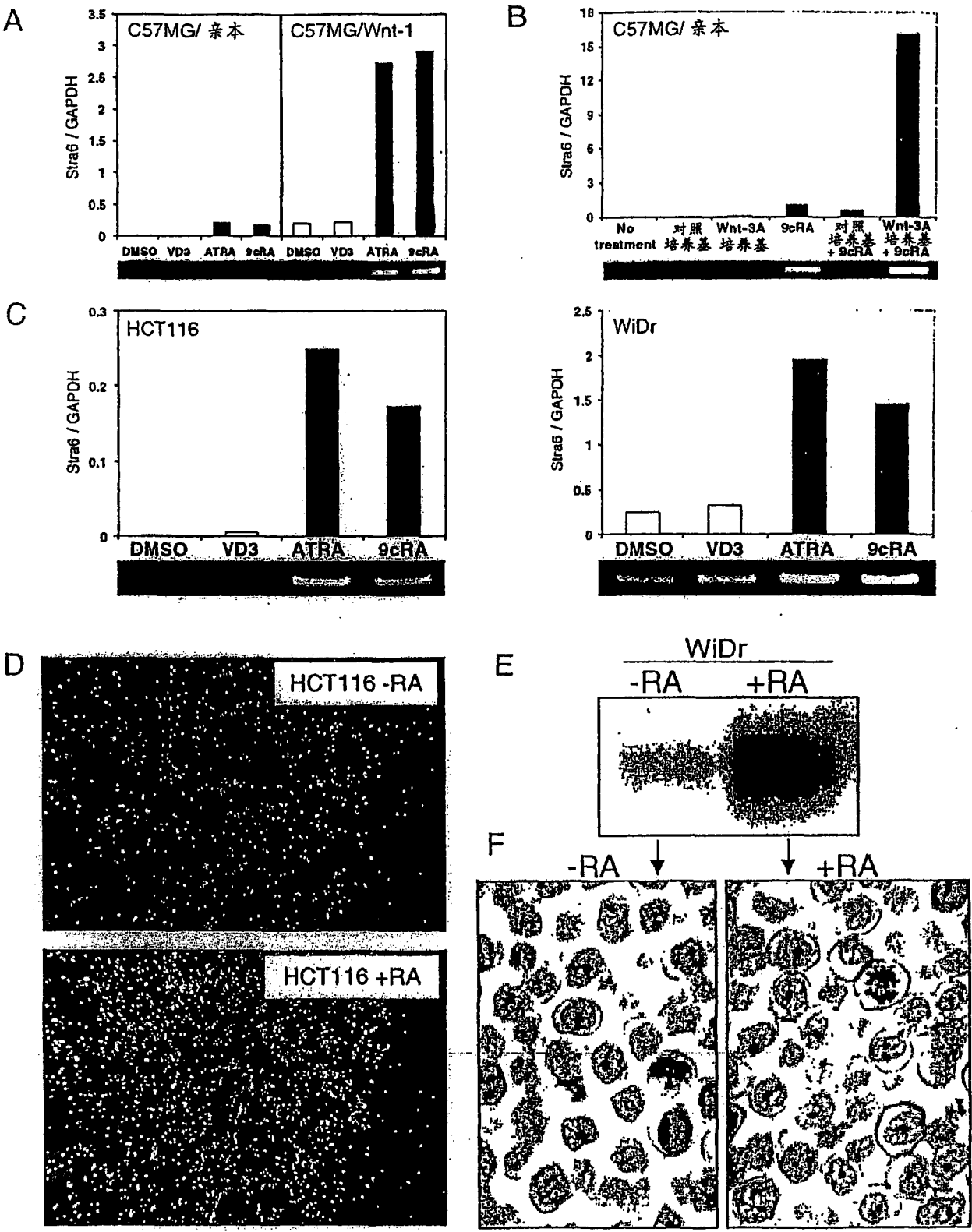


图 17

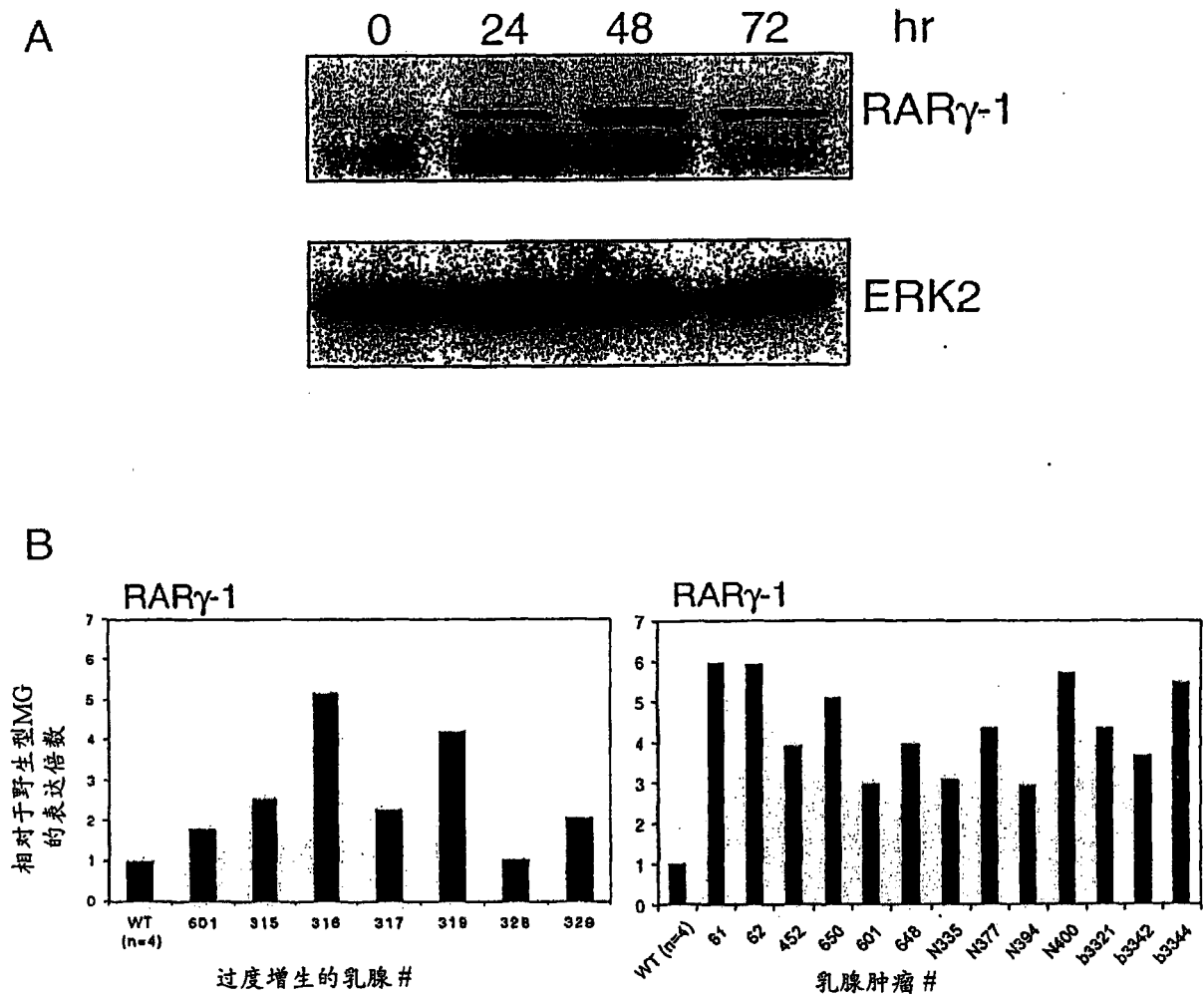


图 18