

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200580037014. X

[51] Int. Cl.

C12N 15/82 (2006.01)

C12N 15/85 (2006.01)

C12N 15/63 (2006.01)

[43] 公开日 2007 年 10 月 3 日

[11] 公开号 CN 101048509A

[22] 申请日 2005.9.1

[21] 申请号 200580037014. X

[30] 优先权

[32] 2004.9.2 [33] US [31] 60/606,439

[86] 国际申请 PCT/US2005/031081 2005.9.1

[87] 国际公布 WO2006/028889 英 2006.3.16

[85] 进入国家阶段日期 2007.4.27

[71] 申请人 惠氏公司

地址 美国新泽西州

[72] 发明人 郇轶洁 N·M·皮什 耿 梅

S·H·埃尔曼 钟晓天

R·克里茨 吕志坚

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 林柏楠

权利要求书 3 页 说明书 26 页 序列表 17 页
附图 9 页

[54] 发明名称

用于产生蛋白质的体系和方法

[57] 摘要

本发明涉及用于产生的目的蛋白质的体系和方法。本发明使用具有改良的蛋白质折叠或加工能力的基因工程化的动物细胞或植物细胞。在一个方面,本发明描述了基因工程化细胞,其包含编码(1)目的蛋白质和(2)在细胞解折叠蛋白应答(UPR)途径中有功能的多肽的一种或多种重组表达盒。所述多肽的共表达显著地提高基因工程化细胞中目的蛋白质产率。在一个例子中,基因工程化细胞是动物细胞,并且共表达的多肽是 XBP1 介导性或 ATF6 介导性 UPR 途径的组分或调节物。

1. 动物细胞或植物细胞, 其包含一种或多种重组表达盒, 所述重组表达盒编码目的蛋白质和解折叠蛋白应答(UPR)途径的组分或非 IRE1 调节物。

2. 根据权利要求 1 所述的细胞, 其中所述细胞是动物细胞, 并且所述的组分或非 IRE1 调节物是 XBP1 或 ATF6 蛋白质。

3. 如权利要求 1 或 2 所述的细胞, 其中所述的一种或多种重组表达盒包括: 编码所述目的蛋白质的至少一种重组表达盒; 和编码所述的 XBP1 或 ATF6 蛋白的至少另一种重组表达盒,

并且其中所述至少一种重组表达盒总数对所述至少另一种重组表达盒总数的比率在所述细胞中是从 0.5:1 至 10:1。

4. 根据权利要求 3 所述的细胞, 其中所述至少一种重组表达盒总数对所述至少另一种重组表达盒总数的比率在所述细胞中是至少 3:1。

5. 如权利要求 1-4 任一项中所述的细胞, 其中所述细胞是哺乳动物细胞。

6. 转基因的非人动物, 其包含如权利要求 1-5 任一项中所述的细胞。

7. 如权利要求 1-4 任一项中所述的细胞, 其中所述细胞是通过动物细胞与癌细胞融合所产生的杂交细胞。

8. 动物细胞或植物细胞, 其包含一种或多种重组表达盒, 所述重组表达盒编码目的蛋白质和能够结合至所述细胞的 UPR 元件(UPRE)或 ER-应激反应元件(ERSE)的多肽。

9. 根据权利要求 8 所述的细胞, 其中所述细胞是动物细胞, 并且所述多肽包括转录激活结构域。

10. 根据权利要求 9 所述的细胞, 其中所述的转录激活结构域包含 XBP1 或 ATF6 蛋白的转录激活结构域。

11. 动物细胞培养物或植物细胞培养物, 其以一种或多种表达载体转染或转导, 所述表达载体编码: 目的蛋白质; 和 UPR 途径的组分或非 IRE1

调节物。

12. 根据权利要求 11 所述的细胞培养物，其中所述的组分或非 IRE1 调节物是 XBP1 或 ATF6 蛋白，并且所述的一种或多种表达载体包括：编码所述目的蛋白质的至少一种重组表达盒；和编码所述 XBP1 或 ATF6 蛋白的至少另一种重组表达盒，并且其中所述至少一种重组表达盒对所述至少另一种重组表达盒的摩尔比率在所述细胞培养物中是从 0.5:1 至 10:1。

13. 根据权利要求 12 所述的细胞培养物，其中所述的细胞培养物是哺乳动物细胞培养物，并且所述至少一种重组表达盒对所述至少另一种重组表达盒的摩尔比率在所述细胞培养物中是至少 3:1。

14. 包括在动物细胞或植物细胞中表达目的蛋白质的方法，所述细胞包括一种或多种重组表达盒，所述重组表达盒编码：所述目的蛋白质；和 UPR 途径的组分或非 IRE1 调节物。

15. 根据权利要求 14 所述的方法，其中所述细胞是动物细胞，并且所述的组分或非 IRE1 调节物是 XBP1 或 ATF6 蛋白。

16. 包括在动物细胞或植物细胞中表达目的蛋白质的方法，所述细胞包括一种或多种重组表达盒，所述重组表达盒编码：所述目的蛋白质；和能够结合至所述细胞的 UPRE 或 ERSE 的多肽。

17. 根据权利要求 16 所述的方法，其中所述细胞是动物细胞，并且所述多肽包含 XBP1 或 ATF6 蛋白的转录激活结构域。

18. 包括在动物细胞培养物或植物细胞培养物中表达目的蛋白质的方法，所述的细胞培养物以一种或多种表达载体转染或转导，所述表达载体包括：编码所述目的蛋白质的至少一种重组表达盒；和编码 UPR 途径组分或非 IRE1 调节物的至少另一种重组表达盒。

19. 根据权利要求 18 所述的方法，其中所述细胞培养物是哺乳动物细胞培养物，并且所述的组分或非 IRE1 调节物是 XBP1 或 ATF6 蛋白，并且其中所述至少一种重组表达盒对所述至少另一种重组表达盒的摩尔比率在所述细胞培养物中是从 0.5:1 至 10:1。

20. 表达载体，其编码：治疗性或预防性目的蛋白质，和 UPR 途径的

组分或非 IRE1 调节物。

用于产生蛋白质的体系和方法

本申请要求2004年9月2日所提交美国临时申请第60/606,439的权益。

技术领域

本发明涉及用于产生目的蛋白质的表达体系和使用相同体系用于产生目的蛋白质的方法。

背景

分泌性蛋白质和膜蛋白质在内质网(ER)-高尔基体系中进行折叠和其它翻译后修饰。破坏该体系的稳态引起可以导致凋亡的细胞应激。ER 稳态可以因 Ca^{2+} 浓度或氧化还原状态、糖基化作用的变化、或 ER 腔内未折叠或错误折叠蛋白质堆积而改变。为克服应激，分泌体系已经演化了称为解折叠蛋白应答(UPR)的适应性应激反应机制。哺乳动物 UPR 的活化引起至少三种应答：(1) 转运至 ER 腔内的新蛋白质的量因翻译减弱而减少；(2) 堆积在 ER 腔内的蛋白质反转运至胞质并受到降解；以及确实(3)重建 ER-高尔基分泌体系以便增强体系中的蛋白质折叠和加工能力。

增强响应应激的 ER-高尔基体系的能力涉及上调起折叠作用和加工作用的酶。这些酶包括 ER 蛋白伴侣、参与糖基化作用和二硫键形成的酶和参与小泡转运的酶。在哺乳动物细胞内，IRE1 和 ATF6 蛋白是 UPR 途径中此支路的主要转导蛋白。IRE1 蛋白是在其 C 端胞质结构域具有激酶活性和核酸内切酶活性的 ER 跨膜糖蛋白。在小鼠中已经至少鉴定了两个 IRE1 基因-IRE1 α 和 IRE1 β 。IRE1 α 对活力重要并且表达广泛。仅在胃肠粘膜内检测到 IRE1 β 。ER 应激导致 IRE1 蛋白寡聚化和其胞质结构域反式自体磷酸化(trans-autophosphorylation)。IRE1 的磷酸化激活其核酸内

切酶活性,该活性自转录因子 X 盒结合蛋白 1 (XBP1) 的 mRNA 切除内含子。此剪接事件导致将转录失活性 XBP1 同种型 (即 XBP1u)转换为转录激活性 XBP1 同种型 (即 XBP1s 或 XBP1p)。随后 XBP1p 进入细胞核,在那里它结合至 ER-高尔基蛋白伴侣/酶基因调节区中的靶序列,包括 ER 应激反应元件 (ERSE) 和 UPR 元件(UPRE),以诱导它们转录。众多 UPR 靶基因在其启动子区内具有 ERSE 或 UPRE 序列的一个或多个拷贝。

ATF6 (激活性转录因子 6)是另一种 ER 跨膜蛋白。ER 应激导致 ATF6 蛋白转移至高尔基区室,在那里其胞质结构域由 Site 1 蛋白酶和 Site2 蛋白酶切割。切下的胞质结构域进入细胞核并且通过结合至 ERSE 序列充当转录因子,随后上调分泌途径中的多种蛋白伴侣和加工酶。

活化哺乳动物细胞中的 UPR 途径还导致经过 PERK 信号途径短暂抑制蛋白质翻译。PERK 是能够磷酸化真核性翻译起始因子 eIF2 α 作为响应 ER 应激的 ER 跨膜激酶。eIF2 α 的磷酸化阻止 43S 核糖体起始前复合体装配并且因此导致翻译减弱。自相矛盾的是 eIF2 α 的磷酸化还导致转录因子 ATF4 迅速合成,这反过来增强了促进凋亡的转录因子 CHOP 表达。CHOP 启动细胞死亡,此时 ER 应激的有害效应不再可以克服。

分泌性重组蛋白在哺乳动物细胞中过量表达经常导致低产率。为提高蛋白质产生常用的方法是提高转录率,如使用较强启动子,或提高基因拷贝数。然而,提高的转录率可以加剧 ER 应激,并且因此经常不能显著地提高产率。在某些情况下,这可以进一步降低产率。

发明简述

本发明提供具有改进的分泌性蛋白或膜蛋白产率的表达体系。该体系使用具有改良的并且在许多情况下增强的蛋白质折叠或加工能力的基因工程化动物细胞或植物细胞。

在一个方面,本发明描述了基因工程化动物细胞或植物细胞,其包含编码(1)目的蛋白质和(2)UPR 途径的组分或调节物的一种或多种重组表达盒。合适的 UPR 组分或调节物包括,但不限于,在 UPR 途径中发挥作用

的非 IRE1 分子。它们可以是宿主细胞的内源性 UPR 组分，或其变体或功能等同物。它们还可以是直接或间接调节内源性 UPR 组分的活性或表达的天然存在或非天然存在分子。在许多情况下，UPR 组分/调节物如此加以选择以致它们的表达或活化提高宿主细胞的蛋白质折叠或加工能力。

示例性 UPR 组分/调节物包括，但不限于，转录因子如 XBP1 或 ATF6 或其生物活性片段或变体。也可以使用 XBP1 或 ATF6 介导的 UPR 途径中的其它组分。此外，可以使用停留在 ER 中的(ER-resident)蛋白伴侣或加工酶。

可以根据本发明产生的蛋白质包括，但不限于，促红细胞生成素、生长激素、胰岛素、干扰素、生长因子、膜蛋白或其它治疗性、预防性或诊断性蛋白质。在众多实施方案中，目的蛋白质由宿主细胞表达为分泌性蛋白或膜蛋白质。

本发明的基因工程化细胞可以衍生自细胞系、原代培养物或者其它分离的或培养的细胞。基因工程化细胞还可以是杂交细胞。在许多情况下，杂交细胞通过动物细胞和癌细胞(如骨髓瘤细胞)融合产生。编码目的蛋白质或 UPR 组分/调节物的重组表达盒可以先于或后于融合事件引入或导入杂交细胞。此外，本发明的基因工程化细胞可以是转基因动物或转基因植物的细胞。在一个实施方案中，基因工程化细胞是哺乳动物细胞。

重组表达盒可以通过多种方式引入宿主细胞。例如，表达盒可以稳定整合至宿主细胞的染色体或基因组。整合可以为随机或定向(例如通过使用噬菌体 P1 的 Cre-/ox 重组体系)。表达盒还可以通过非整合性表达载体导入宿主细胞。

重组表达盒可以由组成型启动子或诱导型启动子控制。它还可以由组织特异性或发育调节性启动子控制。其他类型的启动子也可以用于本发明。

在另一方面，本发明描述了基因工程化动物细胞或植物细胞，其包含编码(1)目的蛋白质和(2)能够结合至宿主细胞 UPRE 或 ERSE 的多肽的一种或多种重组表达盒。在一个实施方案中，结合至 UPRE 或 ERSE 的多肽是转录因子，如 XBP1 或 ATF6。在另一个实施方案中，结合至 UPRE 或

ERSE 的多肽可以将另一种蛋白质召集至 UPR 基因的启动子区,并且后一种蛋白质包含能够活化 UPR 基因转录的反式激活结构域(例如 XBP1 或 ATF6 的转录激活结构域)。

本发明的基因工程化细胞可以从相同或不同的重组表达盒上表达目的蛋白质和 UPR 组分/调节物。目的蛋白质和 UPR 组分/调节物可以由相同或不同启动子控制。

在一个实施方案中,本发明的基因工程化细胞包含(1)编码目的蛋白质的第一重组表达盒和(2) 编码 UPR 组分或调节物或者 UPR 结合蛋白或 ERSE 结合蛋白 (例如 XBP1 或 ATF6)的第二重组表达盒。第一重组表达盒总数对第二重组表达盒总数的比率范围在细胞中可以是但不限于从不超过 0.1:1 至至少 10:1。在许多情况下,第一重组盒所用的启动子可以具有与第二重组盒所用启动子的相同或相似强度,并且第一重组盒总数对第二重组盒总数的比率范围是从 0.5:1 至 10:1 (如至少 1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1 或 9:1)。第一重组盒和第二重组盒内的启动子还可以具有不同强度。例如,第一重组盒内的启动子可以比第二重组盒内的启动子更强,如强至少 2、3、4、5、6、7、8、9、10 倍或更多倍。

仍在另一方面,本发明描述了包含本发明基因工程化细胞的动物或植物。用于将遗传修饰的细胞引入动物或植物的方法在本领域内众所周知。在众多实施方案中,动物或植物是转基因动物或转基因植物。

仍在另一方面,本发明描述了以编码(1)目的蛋白质和(2) UPR 途径的组分或非 IRE1 调节物的一种或多种表达载体所转染或转导的细胞培养物。在众多实施方案中,表达载体包含编码目的蛋白质的第一重组表达盒和编码 UPR 组分或调节物 (例如 XBP1 或 ATF6)的第二重组表达盒。第一表达盒和第二表达盒可以由相同或不同表达载体携带。第一重组表达盒对第二重组表达盒的摩尔比范围在细胞培养物中可以例如是从不超过 0.1:1 至至少 10:1,如至少 1:1、2:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1 或更多。细胞培养物可以是适用于产生目的蛋白质的哺乳动物细胞培养物、昆虫细胞培养物、植物细胞培养物或其它培养物。

本发明还描述了使用基因工程化的细胞、动物、植物或细胞培养物用于产生目的蛋白质的方法。

此外，本发明描述了编码(1)目的蛋白质和(2) UPR 途径的组分或非IRE1 调节物的表达载体。

本发明的其它特征、目的和优点在后续的详细说明中显而易见。然而，应当理解的是详细的说明虽然显示了本发明优选的实施方案，但仅仅是以示例方式给出，并不构成限制。本发明范围内的多种改动和修改对本领域技术人员而言从详细的说明中是显而易见的。

附图简述

为说明提供附图，但并不构成限制。

图 1 表明在 BMP6 在 PA DUKX 细胞中的过量表达引起 ER 应激。

图 2 显示外源 XBP1 在稳定转染的细胞系中在非应激条件或应激条件下的表达。

图 3 显示 BMP6 在数个 XBP1 细胞系中比亲代 CHO DUKX 细胞中提高的分泌。

图 4 显示 IL11RFc 在数个 XBP1 细胞系中比亲代 CHO DUKX 细胞中提高的分泌。

图 5 显示 GFP 转染效率在所选择的 XBP1 细胞系和亲代 CHO DUKX 细胞系中相似。

图 6 显示 GFP 转录效率和翻译效率在所选择的 XBP1 细胞系和亲代 CHO DUKX 细胞系中相似。

图 7 显示可诱导的 ATF6 蛋白在 COS-1 细胞中的成功表达。

图 8 描述与靶蛋白 cDNA 呈不同比例的 XBP1 或 ATF6 对靶蛋白产生的影响。

图 9 显示 XBP1 或 ATF6 对不同靶蛋白表达的影响。

发明详述

本发明提供了用于产生目的蛋白质的体系和方法。本发明的表达体系使用具有改良或改进蛋白质折叠或加工能力的基因工程化动物细胞或植物细胞。在许多实施方案中，本发明的基因工程化细胞包含编码目的蛋白质和 UPR 途径的组分或调节物的一种或多种重组表达盒。共表达 UPR 组分/调节物显著地提高目的蛋白质产率。适用于本发明的 UPR 组分/调节物包括，但不限于，转录因子如 XBP1 或 ATF6 或其生物活性片段或变体。还可以使用 ER 相关的蛋白伴侣或酶。

本发明的多个方面在如下分项中更详细地描述。分项的用途并非意在限制本发明。每一分项可以用于本发明的任意方面。如本文中所用，术语“或”意指“和/或”，除非另外申明。

A. 目的蛋白质

可以根据本发明产生的蛋白质包括，但不限于治疗性、预防性或诊断性蛋白质如促红细胞生成素、生长激素、胰岛素、白细胞介素、生长因子、干扰素、集落刺激因子、血液因子、疫苗、胶原、纤维蛋白原、人血清白蛋白、组织纤溶酶原激活物、葡萄糖苷酶、阿糖苷酶(alglucerases)、髓鞘碱性蛋白、次黄嘌呤-鸟嘌呤磷酸核糖转移酶、酪氨酸羟化酶、多巴脱羧酶或抗体。服从于本发明的示例性抗体包括，但不限于单克隆抗体、单特异性抗体、多特异性抗体、非特异性抗体、人源化抗体、人抗体、单链抗体、嵌合抗体、合成性抗体、重组抗体、杂合抗体、Fab、F(ab')₂、Fv、scFv、Fd、dAb 或其功能性片段。使用体外 (*in vitro*) 呈现技术或演化策略所选择的高亲和性结合物也可以根据本发明产生。这些高亲和性结合物包括，但不限于肽、抗体和抗体模拟物。见例如 Binz 等, NAT BIOTECHNOL., 22:575-582 (2004) 以及 Lipovsek 和 Pluckthun, J IMMUNOL METHODS, 290:51-67 (2004)。其它目的蛋白质如激酶、磷酸酶、G 蛋白偶联受体、生长因子受体、细胞因子受体、趋化因子受体、细胞表面抗体(膜结合免疫球蛋白)、BMP/GDF-受体、神经元受体、离子通道、蛋白酶、转录因子或聚合酶也可以由本发明产生。

在众多实施方案中，由本发明产生的蛋白质是重组蛋白。如本文中所

用, 重组蛋白指的是使用重组 DNA 技术构建或产生的蛋白质。重组蛋白可以具有天然存在的序列或基因工程化的序列。它可以例如从重组载体或从对宿主细胞为内源但具有基因工程化调节序列的基因表达。例如, 重组蛋白可以自具有基因工程化的病毒启动子的内源性基因表达。

在许多情况下, 重组蛋白是融和蛋白, 包括促进所表达产物的分离、纯化、检测、固定、稳定、折叠或定向的多肽标签。

在众多其它情况下, 重组蛋白包括信号肽。信号肽对所产生的蛋白质可以是内源性或异源性的。信号肽经常决定蛋白质是在粗面 ER 还是在游离核糖体上产生。信号肽可以与信号识别颗粒相互作用并且指导核糖体至发生共翻译性插入的 ER 上。众多信号肽高度疏水, 具有正电荷残基。信号肽可以自生长的肽链由位于 ER 潴泡面的特异性蛋白酶-信号肽酶切下。

由信号序列定向至 ER 的蛋白质可以作为分泌性蛋白质释放至细胞外空间。例如, 含有分泌性蛋白质的小泡可以与细胞膜融合并将其内容物释放至细胞外空间-一种称作胞吐的过程。胞吐可以组成型地发生或在接受触发信号后发生。在后一情况下, 蛋白质贮藏在分泌泡(或分泌颗粒)内直至触发胞吐。类似地, 定居在细胞膜上的蛋白质可以通过蛋白水解性切割将该蛋白质固定至膜上的“接头”而分泌至胞外空间。

目的蛋白质可以自表达体系通过多种方法分离。用于蛋白质分离的初始材料实例包括, 但不限于, 培养基或细胞裂解物。合适的分离方法包括, 但不限于, 亲和层析(包括免疫亲和层析)、离子交换层析、疏水相互作用层析、大小排阻层析、HPLC、蛋白质沉淀(包括免疫沉淀)、差异性溶解、电泳、离心、结晶或其组合。多肽标签如链亲和素标签、FLAG 标签、多组氨酸标签、谷胱甘肽 S 转移酶或 Fc 片段可以与目的蛋白质融合以促进目的蛋白质的分离或纯化。在一个情况下, 多肽标签可从目的蛋白质通过蛋白酶切除。

在众多实施方案中, 根据本发明分离的目的蛋白质基本上没有其它蛋白质或杂质。例如, 分离的蛋白质可以是至少 50%、60%、70%、80%、

90%、95%或99%无其它蛋白质的纯度。在一个实例中，分离的蛋白质含有不超过非显著量的会干扰该蛋白质预期用途的杂质。

根据本发明分离的目的蛋白质可以通过使用标准技术如 SDS-PAGE 或免疫测定法进行验证。SDS-PAGE 可以用考马斯蓝试剂、银试剂或其它合适试剂染色以显影分离的蛋白质。合适的免疫测定法包括，但不限于，Western 印迹、ELISA、RIA、夹心或免疫度量测定法、乳胶或其它颗粒凝集试验或蛋白质组芯片。也可以使用蛋白质测序和质谱法以验证或分析分离的蛋白质。

B. UPR 组分/调节物

服从于本发明的 UPR 组分/调节物包括在 UPR 途径中发挥作用的分子。它们可以是天然存在分子或非天然存在分子。可以对它们进行遗传工程化、化学合成或生物分离。UPR 组分/调节物可以源自与宿主细胞相同或不同的物种。UPR 组分/调节物的表达改进宿主细胞的分泌作用或 ER 加工能力。在许多情况下，UPR 组分/调节物与目的蛋白质在宿主细胞内的共表达改进目的蛋白质的产率至少 2、3、4、5、10 倍或更多。

适合本发明的 UPR 组分/调节物的实例包括，但不限于，转录因子如 XBP1 或 ATF6。也可以使用这些转录因子的功能等同物如 XBP1 或 ATF6 的片段。这些片段至少保留了转录激活的 XBP1 或 ATF6 蛋白的实质部分转录活性。也可以使用 XBP1 或 ATF6 的下游效应物如参与蛋白质糖基化作用或小泡转运的 ER 蛋白伴侣或酶。

此外，也可以使用可活化 XBP1 或 ATF6 介导的 UPR 途径的组分表达或生物学功能的非 IRE1 调节物。该调节物可以通过不同于自我过量表达的机制调节 UPR 途径。例如，该调节物可通过直接结合至 UPR 组分而活化此组分的功能，或通过结合至编码此组分的基因中的调节序列而调节此组分表达。

此外，也可以使用可抑制 PERK 信号途径的组分表达或生物学功能的调节物。这些调节物包括但不限于抗体、反义 RNA 或 RNAi 序列。此外，可以使用 PERK 途径组分的显性失活突变体。此类显性失活突变体的实例

是在第 51 位置(鼠序列)具有丝氨酸替换为丙氨酸的 eIF2a S51A 突变体。该替换取消了蛋白质被磷酸化的能力并且因此终止该蛋白质在 ER 应激期间对蛋白质翻译速率的抑制效应。类似地,可以向 PERK 的激酶结构域导入突变以消除或减少其磷酸化 eIF2a 的激酶活性,因此阻止 ER 应激期间翻译减弱或凋亡的诱导。

在一个实施方案中,使用 XBP1 蛋白质或其生物学活性片段以提高宿主细胞中目的蛋白质产率。XBP1 蛋白质包括在转录因子中经常找到的赋予 DNA 结合和二聚化作用能力的两个结构域。已知 XBP1 是通过结合至称作 X 盒的启动子元件对 MHC II 类基因调节的转录因子。XBP1 还结合至 T 细胞白血病病毒 1 型的启动子中的增强子。

UPR 途径的活化在秀丽隐杆线虫(*Caenorhabditis elegans*)和哺乳动物模型体系中均导致小内含子自 XBP1 mRNA 的 IRE1 依赖性剪接。得到的外显子由 tRNA 连接酶连接。该剪接事件在 XBP1 mRNA 内产生移码,这产生了具有原有 N 末端 DNA 结合结构域但具有新的 C 端反式激活结构域的蛋白质。在鼠细胞中,该剪接事件将 267 个氨基酸 XBP1 同种型转换为 371 个氨基酸 XBP1 同种型(XBP1s 或 XBP1p)。见 Calton 等, NATURE, 415: 92-96 (2002)。

XBP1p 蛋白质结合至众多 ER 蛋白伴侣或 UPR 基因启动子区的 ERSE 或 UPRE 序列上,激活这些基因转录。在哺乳动物中,已经鉴定至少两个 ERSE 序列 ERSE-I 和 ERSE-II。ERSE-I 具有如 SEQ ID NO:1 (CC AATNNNNNNNNCCACG)中所示的保守序列。ERSE-II 具有如 SEQ ID NO:2 (ATTGGNCCACG)中所示的保守序列。此外,已经鉴定至少两个哺乳动物的 UPRE 序列,一个序列具有如 SEQ ID NO:3 (TGACGTGG)中所示的保守序列,并且另一个序列具有如 SEQ ID NO:4 (TGACGTGA)中所示的保守序列。

XBP1 蛋白质编码序列可以自多种来源获得。例如,用于人、小鼠、大鼠、鸡、果蝇和斑马鱼 XBP1 蛋白的编码序列可自国立生物信息中心 (NCBI) (Bethesda, MD) 的 Entrez 核苷酸数据库得到。这些序列分别具有

Entrez 登录号 NM_005080、NM_013842、NM_001004210、NM_001006192、NM_079983 或 NM_131874。

XBP1 蛋白的生物学活性片段至少保留 XBP1p 蛋白的实质部分转录激活活性。例如，在本发明中使用的 XBP1 片段可以保留 XBP1p 至少 50%、60%、70%、80%、90% 或更高的转录激活活性。可通过多种方法选择具有转录活性的 XBP1 片段。在一个实例中，具有转录活性的 XBP1 片段基于其激活 ERSE 或 UPRE 序列下游的基因转录的能力加以选择。

在另一个实施方案中，使用 ATF6 蛋白质或其生物学活性片段以改进宿主细胞中目的蛋白质产率。ATF6 是包括位于 EER 腔内的“感知”结构域和胞质转录激活结构域的跨膜蛋白。ER 应激后，ATF6 的胞质结构域被切下并转运至细胞核，在那里它结合至 ERSE 序列并且因此激活下游的 UPR 基因。已经鉴定了至少两种 ATF6 蛋白，即 ATF6 α 和 ATF6 β 。ATF6 α 和 ATF6 β 在结构上相关并且在其 b-zip 结构域内具有显著的相似性。用于人、小鼠、绵羊和鸡 ATF6 蛋白的示例性编码序列分别具有 Entrez 登录号 NM_007348、XM_129579、AY942654 和 XM_422208。

与本发明中所用 XBP1 片段相似，ATF6 蛋白的生物学活性片段至少保留了活化的 ATF6 蛋白或其胞质结构域的实质部分转录激活活性。生物学活性 ATF6 片段可以通过监测它结合至 ERSE 序列和该片段激活 ERSE 序列下游的基因转录的能力加以选择。

仍在另一个实施方案中，使用停留在 ER 中的加工酶或蛋白伴侣以提高宿主细胞中目的蛋白质产率。位于 ER 的合适酶/蛋白伴侣实例包括，但不限于，GRP78、GRP94、GRP58、蛋白质二硫化物异构酶、钙联结蛋白和肌钙网蛋白(calreticulin)。在一个实例中，本发明中所用 ER 酶(或蛋白伴侣)的内源对应物具有包含一个或多个 ERSE-I 或 ERSE-II 序列的启动子区。在另一个实例中，本发明中所用 ER 酶(或蛋白伴侣)的内源对应物具有包含一个或多个 UPRE 序列的启动子区。

本发明还描述了作为内源性蛋白质变体的 UPR 组分的用途。内源 UPR 组分的变体可以是天然存在的如通过等位变异或多态性，或是为有意设计。

变体的UPR活性与原始蛋白质(例如XBP1p、转录性激活的ATF6蛋白或其生物学活性片段)相比没有显著下降。在众多实施方案中,本发明中所用变体至少保留相应原始蛋白质UPR活性的50%。例如,变体可以至少保留原始蛋白质UPR活性的60%、70%、80%、90%、95%、99%或100%。在一个实施方案中,本发明中所用变体显示与原始蛋白质相比提高的UPR活性。UPR组分的所需要变体可以如此加以选择以致该变体的表达或活化增强宿主细胞中共转染蛋白质的分泌。

变体的氨基酸序列与原始蛋白质的氨基酸序列基本相同。在很多情况下,变体的氨基酸序列具有与原始蛋白质至少80%、85%、90%、95%或99%的总体序列同一性或相似性。序列同一性或相似性可以通过多种方法确定。在一个实施方案中,序列同一性或相似性通过使用序列比对算法确定。用于此目的合适算法包括,但不限于,Altschul等,J. MOL. BIOL., 215:403-410 (1990)中所述的基本局部比对工具(BLAST)、Needleman等,J. MOL. BIOL., 48:444-453 (1970)的算法、Myers和Miller, COMPUTE. APPLE. BIOSCL, 4:11-17 (1988)算法和点矩阵分析(dot matrix analysis)。用于此目的合适计算机程序包括,但不限于,NCBI提供的BLAST程序、DNASTAR (Madison, WI)提供的MegAlign和Genetics Computer Group (GCG) GAP程序(Needleman-Wench算法)。对于GAP程序,可以使用默认值(例如在序列之一中打开缺口的罚值是11并且对延伸缺口的罚值是8)。相似氨基酸可以通过BLOSSOM替换矩阵进行定义。

众多方法可用于制备UPR组分的所需要变体。例如,变体可以通过至少1、2、3、4、5、10、20个或更多个氨基酸的替换、缺失、插入或其它修饰衍生自原始蛋白质。替换可以是保守性的或非保守性的。在多数情况下,保守性氨基酸替换可以导入蛋白质序列而不显著改变该蛋白质的结构或生物学活性。保守性氨基酸转换可以基于残基的极性、电荷、可溶性、疏水性、亲水性或两亲性质实施。例如,保守性氨基酸替换可以在具有碱性侧链的氨基酸如赖氨酸(Lys或K)、精氨酸(Arg或R)和组氨酸(His或H);具有酸性侧链的氨基酸如天冬氨酸(Asp或D)和谷氨酸(Glu或E);具

有不带电荷极性侧链的氨基酸如天冬酰胺 (Asn 或 N)、谷氨酰胺(Gln 或 Q)、丝氨酸(Ser 或 S)、苏氨酸(Thr 或 T)和酪氨酸 (Tyr 或 Y)或具有非极性侧链的氨基酸如丙氨酸(Ala 或 A)、甘氨酸 (Gly 或 G)、缬氨酸(Val 或 V)、亮氨酸 (Leu 或 L)、异亮氨酸(Ile 或 I)、脯氨酸 (Pro 或 P)、苯丙氨酸 (Phe 或 F)、甲硫氨酸 (Met 或 M)、色氨酸 (Trp 或 W)或半胱氨酸 (Cys 或 C)中进行。常用氨基酸替换实例在表 1 中说明。

表 1 氨基酸替换的实例

原始残基	示例性替换	较保守的替换
Ala (A)	Val, Leu, Ile	Val
Arg (R)	Lys, Gln, Asn	Lys
Asn (N)	Gln	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser, Ala	Ser
Gln (Q)	Asn	Asn
Gly (G)	Pro, Ala	Ala
His (H)	Asn, Gln, Lys, Arg	Arg
Ile (I)	Leu, Val, Met, Ala, Phe, 正亮氨酸	Leu
Leu (L)	正亮氨酸, Ile, Val, Met, Ala, Phe	Ile
Lys (K)	Arg, 1, 4 - 二氨基丁酸, Gln, Asn	Arg
Met (M)	Leu, Phe, Ile	Leu
Phe (F)	Leu, Val, Ile, Ala, Tyr	Leu
Pro (P)	Ala	Gly
Ser (S)	Thr, Ala, Cys	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr, Phe	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser	Phe
Val (V)	Ile, Met, Leu, Phe, Ala, 正亮氨酸	Leu

也可以将其它想要的氨基酸替换导入 UPR 组分。例如, 可以将氨基酸替换导入 UPR 组分以提高其稳定性。对另一个实例, 可以将氨基酸替换导入以提高或降低 UPR 组分的 UPR 活性。

此外, 本发明描述了可以结合至宿主细胞 UPRE 或 ERSE 序列的多肽的用途。这些多肽, 无论是单独或与其它蛋白质组合, 均可以作为转录因

子发挥作用以激活具有 UPRE 或 ERSE 启动子区的基因转录。

C. 重组表达盒和宿主细胞

本发明中所用的一般重组表达盒包含有效地连接至 5'非翻译调节区和 3'非翻译调节区的蛋白质编码序列。蛋白质编码序列可以是基因组序列、cDNA 序列、其组合，或者其它可表达的序列。

在一个实施方案中，重组表达盒包括指导编码蛋白质表达所必需的全部调节元件。合适的 5'非翻译调节元件实例包括启动子、增强子或 Kozak 序列。合适的 3'非翻译调节元件实例包括聚腺苷酸化序列或其它转录/翻译终止序列。为表达盒选择合适的启动子、增强子或其它调节元件是本领域内技术人员水平范围内的常规设计。众多此类元件在文献中描述并且可通过供应商得到。

适合于本发明的启动子包括组成型或诱导型启动子。这些启动子可以对宿主细胞为内源性或异源性。在一个实施方案中，使用组织特异性启动子。合适的组织特异性启动子包括，但不限于，肺特异性启动子、淋巴样特异性启动子、T 细胞特异性启动子、神经元特异性启动子、胰腺特异性启动子或乳腺特异性启动子。也可以使用发育调节性启动子。可以使用具有组织特异性启动子或发育调节性启动子的重组表达盒以制备转基因动物或转基因植物。此种重组表达盒所编码的蛋白质可以在转基因动物或转基因植物的特定组织内或特定发育阶段产生。

在另一个实施方案中，使用诱导型表达体系以产生目的蛋白质或 UPR 组分/调节物。适合于此目的的体系包括，但不限于，Tet-开/关体系、蜕皮素体系和雷帕霉素体系。Tet-开/关体系基于衍生自大肠杆菌(*E. coli*) Tn10 转座子中四环素抗性操纵子的两个调节元件。该体系包括两种组分，调节盒和报告盒。在一个形式的 Tet-关闭体系中，调节盒编码了包含与单纯疱疹病毒(HSV) VP 16 激活结构域融合的 Tet 阻遏物(tetR)的杂合蛋白。报告盒包括有效地连接至报告基因的 tet 响应元件(TRE)。报告基因可编码例如目的蛋白质或 UPR 组分/调节物。在缺乏诱导物(例如四环素或强力霉素)时，tetR-VP 16 融和蛋白结合至 TRE，并因此活化报告基因的转录。

在 Tet-on 体系的另一个形式中, 调节盒编码了包含与 HSV VP 16 激活结构域融合的突变 Tet 阻遏物(rtetR)的杂合蛋白。rtetR 是反 Tet 阻遏物, 其在诱导物(例如四环素或强力霉素)存在时结合至并且激活 TRE。

蜕皮素体系基于果蝇 (*Drosophila*) 中的蜕皮诱导体系。在一个形式中, 该体系包括编码功能性蜕皮素受体的调节盒以及包含有效地连接至报告基因的蜕皮素响应启动子的报告盒。在诱导物(如松甾酮 A 或 muristerone A)存在时, 蜕皮素受体结合至蜕皮素响应启动子, 激活报告基因转录。

雷帕霉素体系, 也称作 CID 体系 (“二聚化的化学诱导物”), 使用两种嵌合蛋白。第一种嵌合蛋白包括与结合至 DNA 响应元件的 DNA 结合结构域融合的 FKBP 12。第二种嵌合蛋白包括与转录激活结构域融合的 FRAP。添加雷帕霉素引起两种嵌合蛋白二聚化, 因此激活受 DNA 响应元件控制的基因表达。

在一个实例中, 本发明中所用的重组表达盒包含 SEQ ID NO:5。SEQ ID NO:5 的转录产生非剪接的人 XBP1 mRNA。ER 应激活化 IRE1, 这将内含子从非剪接的 mRNA 中切下。经剪接 mRNA 的翻译产生成熟且具有功能性的 XBP1 蛋白, 其氨基酸序列在 SEQ ID NO:6 中描述。

在另一个实例中, 本发明中所用重组表达盒包含 SEQ ID NO:7。SEQ ID NO:7 不包括任何可切割的内含子序列。SEQ ID NO:7 的表达产生成熟和具有功能性的人 XBP1 蛋白(SEQ ID NO:6)。还可以使用编码 SEQ ID NO:6 或其功能等同物的其它核酸序列以便产生本发明的重组表达盒。这些核酸序列可以包含或不包含内含子或其它可移去的序列。

仍在另一个实例中, 本发明中所用重组表达盒包含 SEQ ID NO:8。SEQ ID NO:8 的转录和翻译产生人 ATF6 蛋白, 其氨基酸序列在 SEQ ID NO:9 中显示。

在又一个实例中, 本发明中所用重组表达盒包含了编码 SEQ ID NO:9 的氨基酸残基 1-366 的核酸序列。此核酸序列的实例是 SEQ ID NO:8 的核苷酸 1-1098。SEQ ID NO:9 的氨基酸残基 1-366 包括人 ATF6 蛋白的完整

碱性区域和大部分亮氨酸拉链区域。已经证实此 ATF6 片段能够活化内源 GRP78 基因。见 Haze 等, *MOL. BIOL. CELL*, 10:3787- 3799 (1999)。

还可以在本发明中使用编码自非人物种衍生的 XBP1 或 ATF6 蛋白的重组盒。例如, 可以使用啮齿类动物或其它动物物种的 XBP1 或 ATF6 蛋白。这些 XBP1 或 ATF6 蛋白可以如此加以选择以致这些蛋白质与目的蛋白质的共表达改进了宿主细胞中的目的蛋白质产率。

重组表达盒可以通过多种方法引入宿主细胞。在一个实施方案中, 重组表达盒通过使用转染载体或转导载体导入真核宿主细胞。适合于此目的的载体包括, 但不限于, 昆虫细胞表达载体 (例如杆状病毒表达载体) 或哺乳动物表达载体。这些载体可以自多种来源如附加体、粘粒、病毒或其组合衍生。在多数情况下, 这些载体包括可选择标记以便促进它们引入宿主细胞。

在另一个实施方案, 本发明中所用重组表达盒通过修饰宿主细胞中的内源基因构建。内源性基因可以编码目的蛋白质或 UPR 组分/调节物。可以修饰内源性基因中的许多部分以实现所需要的表达或调节效果。例如, 内源性基因的原始启动子可以由病毒启动子替换以提高基因表达水平。

可以将重组表达盒以多种形式引入宿主细胞。例如, 可以将重组表达盒整合至宿主细胞的染色体或基因组。重组表达盒还可以由宿主细胞中的非整合性表达载体携带。用于将表达载体或表达盒稳定地或瞬时导入宿主细胞的方法在本领域内已知。在一个实例中, 将表达载体或表达盒通过定向整合导入宿主细胞染色体。适合于此目的的方法包括, 但不限于, Cre-/ox 重组体系和在美国专利号 6,656,727、6,537,542 和 6,541,231 中所描述的那些重组体系。

在众多实施方案中, 本发明的基因工程化细胞包括(1)编码目的蛋白质的第一重组表达盒和(2) 编码 UPR 组分/调节物 (例如 XBP1 或 ATF6)第二重组表达盒。第一重组表达盒总数对第二重组表达盒总数的比率范围在细胞中可以是, 不限于, 例如从不超过 0.1:1 至至少 10:1。合适比例的非限制性实例包括从 0.2:1 至 5:1、从 0.5:1 至 5:1、从 1:1 至 2:1、从 1:1 至 3:1、

从 1:1 至 4:1、从 1:1 至 5:1、从 2:1 至 3:1、从 2:1 至 4:1 以及从 2:1 至 5:1。第一重组表达盒和第二重组表达盒可以由相同或不同载体携带并且由具有相同或不同强度的相同或不同启动子驱动。在一个例子中, 第一重组表达盒中的启动子具有与第二重组表达盒中的启动子相同或相似的强度, 并且第一重组盒总数对第二重组盒总数的比率范围在细胞中是至少 1:1、2:1、3:1、4:1、5:1 或更多。

适合本发明的宿主细胞包括动物细胞或植物细胞。宿主细胞可以是培养的细胞如细胞系或者原代培养物。它们还可以是转基因动物或转基因植物的细胞。选择合适宿主细胞和方法用于培养、转染/转导、扩大、筛选、产生和纯化产物在本领域内是已知的。

在一个实施方案中, 本发明中所用的宿主细胞是哺乳动物细胞。合适的哺乳动物细胞实例包括, 但不限于, 中国仓鼠卵巢 (CHO) 细胞、HeLa 细胞、COS 细胞、293 细胞、CV-1 细胞和其它由美国典型培养物保藏中心 (Manassas, Virginia) 收集的哺乳动物细胞系。在某些情况下, 希望在常常为表达蛋白质提供所需要的翻译后修饰的人细胞中产生治疗性或预防性蛋白质。

本发明中所用的宿主细胞还可以是通过两种或多种细胞融合所产生的杂交细胞。在多种情况下, 本发明中所用的杂交细胞通过融合动物细胞(例如哺乳动物细胞)和癌细胞/永生细胞(例如骨髓瘤细胞或胚细胞瘤细胞)产生。动物细胞和癌细胞/永生细胞可以源自相同物种。它们还可以源自不同物种。可以使用本领域内已知的产生杂合细胞的任意方法。这些方法包括, 但不限于, 电融合或化学融合(例如聚乙二醇融合)。

重组表达盒可以先于或后于融合事件导入或引入杂交细胞。例如, 编码目的蛋白质的重组表达盒可以先于哺乳动物细胞与表达外源 UPR 组分或调节物的癌细胞融合之前引入哺乳动物细胞。对于另一种情况, 哺乳动物细胞可以首先用编码目的蛋白质以及 UPR 组分或调节物的重组表达载体转染或转导并且随后与癌细胞融合。还可以使用其它方法以制备本发明的杂交细胞。

在众多实施方案中，用于制备杂交细胞的癌细胞/永生细胞对一种或多种选择剂敏感。例如，癌细胞/永生细胞可以对含有次黄嘌呤、氨基蝶呤和胸苷的已知为“HAT”培养基的培养基敏感。将这些 HAT 敏感细胞与对 HAT 培养基不敏感的细胞融合。将如此产生的杂交细胞进行杀死非融合细胞的 HAT 选择。随后对所需要的特征筛选融合细胞。

本发明还描述了包含本发明真核宿主细胞的动物或植物。用于将重组细胞引入动物或植物的方法在本领域内众所周知。在众多实施方案中，动物或植物是包含编码目的蛋白质和 UPR 组分/调节物的一个或多个转基因的转基因动物或转基因植物。转基因动物或转基因植物可以使用标准技术产生。在一个实施方案中，转基因动物是非人动物。

本发明还描述了以编码(1)目的蛋白质和(2) UPR 组分/调节物的一种或多种表达载体所转染或所转导的动物细胞培养物或植物细胞培养物。细胞培养物可以是哺乳动物细胞培养物、昆虫细胞培养物、植物细胞培养物或适用于产生目的蛋白质的其它培养物。可以瞬时或稳定转染或转导表达载体。在一个实施方案中，所用表达载体包含编码目的蛋白质的第一重组表达盒和编码 UPR 组分/调节物 (例如 XBP1 或 ATF6)的第二重组表达盒。第一表达盒和第二表达盒可以由相同或不同载体携带。它们可以由相同或不同启动子驱动。第一重组表达盒对第二重组表达盒的摩尔比范围可以是例如从不超过 0.1:1 至至少 10:1。

在一个实例中，第一重组盒所用的启动子可以具有与第二重组盒所用启动子相同或相似的强度，并且第一重组盒对第二重组盒的摩尔比范围在细胞培养物中可以是 0.5:1 至 10:1，如至少 1:1、2:1、3:1、4:1、5:1 或更高。

D. 药物组合物

由本发明产生的治疗性蛋白质或预防性蛋白质可以用于制备药物组合物以治疗需要它的患者或动物。本发明的药物组合物通常包含有效量的治疗性或预防性蛋白质以及可药用载体。合适的可药用载体包括与药物施用兼容的溶剂、增溶剂、填料、稳定剂、粘合剂、吸收剂、基质、缓冲剂、

润滑剂、受控释放载体、稀释剂、乳化剂、湿润剂、润滑剂、分散介质、包衣、抗菌剂或抗真菌剂、等渗剂和延迟吸收剂等。此类介质和试剂用于药物活性物质的用途在本领域内众所周知。辅助剂也可以引入组合物。

本发明的药物组合物可以制剂为与其预期的施用途径兼容。施用途径的实例包括肠胃外、静脉内、皮内、皮下、口服、吸入、经皮、直肠、经粘膜、局部和全身性施用。在一个实例中，施用通过植入物进行。

用于肠胃外、皮内或皮下应用的溶液或混悬液可以包含如下组分：无菌稀释剂如注射用水、盐溶液、固定油、聚乙二醇、甘油；丙二醇或其它合成性溶剂；抗菌剂如苜醇或对羟基苯甲酸甲酯；抗氧化剂如抗坏血酸或亚硫酸氢钠；螯合剂如乙二胺四乙酸；缓冲剂如乙酸盐、柠檬酸盐或磷酸盐以及用于调整张力的试剂如氯化钠或葡萄糖。pH 可以用酸或碱如盐酸或氢氧化钠调节。肠胃外制剂可以在安瓿内、一次性注射器或由玻璃或塑料制成的多剂量小瓶内封装。

本发明的药物组合物还可以向患者以需要的剂量施用。合适剂量的范围可以例如是从 5 mg 至 100 mg、从 15 mg 至 85 mg、从 30 mg 至 70 mg 或者从 40 mg 至 60 mg。也可以使用低于 5 mg 或高于 100 mg 的剂量。药物组合物可以以一个剂量或多个剂量施用。剂量可以间隔性施用，如每日一次、每周一次或每月一次。

治疗性蛋白质的毒性和治疗效果可以通过标准药学方法在细胞培养物或实验动物模型中确定。例如，LD₅₀ (对 50 % 群体致死的剂量) 和 ED₅₀ (50 % 群体中治疗有效的剂量) 可以得到测定。毒性效应和治疗效应之间的剂量比率是治疗指数，并且可以表述为 LD₅₀/ED₅₀ 比率。在多数情况下，选择显示出大治疗指数的治疗性蛋白质。

自细胞培养物分析和动物研究得到的数据可以用于制订在人中使用的剂量范围。在一个实施方案中，剂量是在至少 50 % 群体内表现治疗有效性并具有少量毒性或无毒性的范围内。剂量可以在该范围内根据所采用剂量形式和所使用的施用途径变化。用于施用本发明所产生的治疗性蛋白质的剂量方案可以由住院医师基于多种因素确定，如蛋白质的作用、病理部位、

疾病严重度、患者年龄、性别和饮食、任意感染的严重度、施用时间和其它临床因素。在一个例子中，全身施用或注射施用以最小有效的剂量开始，并且该剂量随着事先选择的时间过程提高直至观察到阳性效果。随后，使剂量的递增量限于在效果上产生相应增加的水平，同时考虑可能出现的任何有害效应。

治疗进展可以通过定期评估疾病进展进行监测。治疗进展可以例如通过 X 射线、MRI 或其它成像方法、滑液分析或临床检验进行监测。

治疗性或预防性目的蛋白质还可以通过使用基因送递载体引入人或动物。适合于此目的的载体包括，但不限于，病毒载体如逆转录病毒、慢病毒、腺病毒、腺相关病毒(AAV)、疱疹病毒、甲病毒、星状病毒、冠状病毒、正粘病毒、乳多空病毒、副粘病毒、细小病毒、痘病毒或披膜病毒载体。脂质体包裹的表达载体也可以用于基因送递。在众多实施方案，基因送递载体既编码目的蛋白质又编码 UPR 调节物。UPR 调节物的共表达增强目的蛋白质在靶细胞 (例如肿瘤细胞或其它功能紊乱的细胞)内的产生。目的蛋白质和 UPR 调节物还可以通过使用不同载体送递至靶细胞。基因送递可以在体内 (in vivo) 或在体外(ex vivo)实施。

在一个实施方案中，细胞特异性基因送递方法可用于导入治疗性/预防性目的蛋白质或 UPR 调节物至靶细胞内。本领域内已知的众多细胞特异性基因送递方法可用于本发明。例如，细胞特异性配体(例如对靶细胞表面抗原特异的抗体)可以引入或缀合于编码治疗性/预防性目的蛋白质或 UPR 调节物的病毒载体的包膜。该配体可以介导病毒载体进入特定细胞类型。抗体缀合的脂质体也可以用于送递基因治疗载体至特定靶细胞。

应当理解如上所述的实施方案和如下的实施例仅以说明方式给出，并不用于限制。本发明范围内的多种改变和修改对于本领域技术人员而言从本描述中将是显而易见的。

E. 实施例

实施例 1. COS-1 和 DUKX 细胞系

COS-1 细胞自美国典型培养物保藏中心 (ATCC) Manassas, VA 得到, ATCC 编号是 CRL-1650。CHO DUKX 细胞和 PA DUKX 细胞均衍生自作为中国仓鼠卵巢 (CHO) 细胞衍生物的 CHO-K1 (ATCC 编号 CCL-61)。

CHO DUKX 细胞, 也称作 DUKX B11 或 DXB11 细胞, 在参与 DNA 复制的重要酶二氢叶酸还原酶(dhfr) 的产生上缺陷。为选择两个 dhfr 等位基因均突变的细胞, Urlaub 和 Chasin, PROC. NATL. ACAD. SCI. U.S.A., 77:4216-4220 (1980) 实施了两步骤的诱变和选择。所用抗 dhfr⁺细胞的选择剂是氟化脱氧尿苷。氟化脱氧尿苷因其掺入 DNA 并且随后发生放射性降解而毒害细胞。脱氧尿苷对 DNA 的掺入需要将脱氧尿苷转换为 thymidilic acid, 这是必需 DHFR 的过程。Dhfr-突变细胞不能将脱氧尿苷掺入其 DNA 并且因此能够在氟化脱氧尿苷存在下存活。某些 dhfr⁺细胞以及在某些对脱氧尿苷掺入 DNA 所需要的酶缺陷的突变体也可以存活。Dhfr-突变体因其不能进行重新合成甘氨酸、次黄嘌呤和胸苷而不同, 并且它们需要用于生长的外源核苷。

在如 Urlaub 和 Chasin 等所用的第一步选择中, 将野生型细胞进行甲磺酸乙酯 (EMS) 诱变并且在 dhfr 抑制性氮甲喋呤 (MTX) 存在时用氟化脱氧尿苷进行选择以便分离推定的杂合子(d⁺/d⁻)。通过使用足以使杂合子(d⁺/d⁻)而非纯合子(d⁺/d⁺)中全部 DHFR 失活的 MTX 浓度, 纯合子具有残余 DHFR 活性并且掺入氟化脱氧尿苷。通过这种掺入, 对(d⁺/d⁺)细胞进行选择并且仅杂合子能够存活。三轮选择、汇集和扩充仍存活的细胞后, 分离了推定的杂合子细胞系 UKB25。UKB25 细胞进一步以 γ 照射诱变并在缺乏 MTX 的氟化脱氧尿苷中选择。存活的集落表现对甘氨酸、次黄嘌呤和胸苷的三重辅源营养性, 显示出 dhfr⁻表型。将显示此三重辅源营养性的集落加以克隆并证实 dhfr 活性缺陷。通过 Southern 印迹杂交对这样一个克隆的分析表明 dhfr 基因没有发生任何总体重排。将此克隆命名为 DXB11。

对如此产生的 DXB11 细胞检查以证实细胞系预期特征。发现 DUKX B11 细胞在遗传上与衍生它们的 CHO-K1 细胞系相似。它们是亚二倍体

CHO 细胞，具有细胞遗传学上已经广泛研究的 20 条染色体。分裂中期 DUKX B11 染色体的姬姆萨带表明 DUKX B11 细胞是 CHO-K1 衍生物。DUKX B11 细胞具有 DHFR 缺陷并且因此对甘氨酸、嘌呤核苷和胸苷的三重辅源营养性。DUKX B11 细胞的这种 DHFR 缺陷表型是对转移重组异源蛋白质表达质粒至细胞的遗传选择基础。

在一些实验中，用于培养 DUKX B11 细胞的培养基不含次黄嘌呤或胸苷。由于没有二氢叶酸还原酶活性，细胞存活和复制的唯一手段是通过在生长培养基中补充核苷以弥补细胞不能产生它们的能力。因此，将腺苷、脱氧腺苷和胸苷添加至用于 DUKX B11 细胞的生长培养基。这些核苷均以 10 μ g/ml 的浓度添加。该浓度超过在常规小规模生长条件下细胞的需要。

DUKX B11 细胞用于导入含有用于所需要蛋白质的 cDNA 的表达质粒。因为 DUKX B11 细胞缺乏内源 DHFR 活性，故 DHFR 活性可以在生成旨在产生异源蛋白质的细胞系时作为选择标记和可扩增标记使用。通过共转染含有用于 dhfr 的 cDNA 的另一种表达载体，或通过将 dhfr cDNA 在相同表达载体上置于目的 cDNA 附近，可以利用 dhfr 活性作为细胞已经摄取含有所需要基因的表达质粒的标志，并且可以产生所需要的蛋白质。转染后通过维持外源核苷，仅引入含有 dhfr 基因的细胞才能够产生 dhfr 并因此存活。存活的细胞现在可以产生目的蛋白质。当目的 cDNA 和 dhfr 在相同质粒上时，这可以借助双顺反子信使实现，或当基因位于不同质粒上时，这可以通过各自信使实现，质粒在转染事件期间通常共定位在染色体上。当使用两个独立的质粒时，改变含有 dhfr 的质粒对含有 cDNA 的质粒的比率可以增强人们选择含有两种质粒的细胞的能力。

DUKX B11 细胞由于 dhfr 基因中的点突变具有 dhfr 缺陷，并且因此可以回复突变为 dhfr⁺ 表型。在进行无血清悬浮适应期间观察到此回复突变和在无外源核苷时生长的能力。培养中的 DUKX 细胞群体对从悬浮培养开始大约 154 次累计群体加倍(CPD)时仍保持 dhfr⁻。然而，当在 190 CPD 再检查群体对外源核苷的依赖性时，表型回复突变明显。与 dhfr⁺表型同时出现的是细胞平均生长速度的显著增加。由于 dhfr⁻表型是转染和基因扩

增策略所需要的,故在 153.8 CPD 后,进行已适应的 DUKX 细胞无血清悬浮培养。因为这些细胞适应于在导入表达载体前生长在无血清悬浮培养物内,故将它们称作“预适应的 (pre-adapted)”,并且叫做“PA DUKX”。

FBS 依赖的非适应性 DUKX 单层可以用于转染表达载体。在获得异源基因和 dhfr 的表达后,每一新细胞系可以适应无 FBS 的悬浮生长。使用单层细胞转染后适应期往往冗长。“预适应的” DUKX 细胞还可以作为用于转染的宿主细胞使用。这些 PA DUKX 细胞经常提供来自时间和工作量方面的优势,因为转染后再适应无血清悬浮生长的时间通常较短。见 Sinacore 等, BIOTECHNOLOGY AND BIOENGINEERING, 52:518-528 (1996)。

实施例 2. 通过 BMP 6 的过量表达在 PA DUKX 细胞中诱导 ER 应激

将 pSMED2/XBP1 和 pSMED2/BMP6 (+)或空的 pSMED2 载体(-)共转染入 PA DUKX 细胞。pSMED2/XBP1 和 pSMED2/BMP6 表达载体分别编码 XBP1 和 BMP6。两种载体均受 CMV 启动子驱动。转染使用 Eugene6(Roche, Indianapolis, DSf)在 6 孔平板中进行。细胞的生长培养基是补加核苷和(热灭活且透析的)10%FBS 以及青霉素/链霉素/谷氨酰胺(Gibco)的 Alpha 培养基(Gibco)。

细胞在添加 400 mM NaCl 和 1 Complete Mini (来自 Roche, Indianapolis, IN 的蛋白酶抑制剂混合物片剂)的细胞裂解缓冲液(Cell Signaling Technology, Beverly, MA)中裂解。裂解物在转染后 7、24、31 和 48 小时获取并在 10 % tricine 凝胶上电泳,随后是 Western 印迹分析(图 1)。Western 印迹分析通过使用 TBS 中 4%脱脂乳、1% BSA 和 0.1% Tween20 的封闭缓冲液和 TBS 中的 0.1% Tween20 的洗涤缓冲液进行。将抗体在封闭缓冲液内稀释。用于 Western 的滴定是 1:1000 抗 XBP1 (Santa Cruz Biotechnology), 随后 1:5000 辣根过氧化物酶缀合的山羊抗兔抗体 (The Jackson Laboratory)。图 1 显示 BMP6 在 PA DUKX 细胞中的过量表达引起 ER 应激,如测量到 XBP1p 蛋白(大约 54 kD)增加。

实施例 3. 以 XBP1 稳定转染的细胞系

将 pSMED2/XBP1 载体转化至 CHO DUKX 细胞以产生稳定的细胞系。在 pSMED2 载体上的 DFHR 基因允许氨甲喋呤 (MTX) 抗性。将转染的细胞在 5、10 或 20 nM MTX 浓度中接种。分离了三个 5nM MTX 集落 (5-2、5-4、5-5)、一个 10nM 集落 (10-3) 和一个 20nM 集落 (20-10)。来自每一集落的细胞用 (+) 或不用 (-) 已知导致 ER 应激的化学品 tunacymicin (Tu) 处理。将裂解物在 10% tricine 凝胶上电泳, 随后使用兔抗-XBP1 多克隆抗体 (Santa Cruz Biotechnology) 进行 Western 印迹分析 (图 2)。图 2 显示当细胞因 Tu 处理应激时更多成熟的 XBP1 在 XBP1 稳定性细胞系中产生。

将 pSMED2/BMP6 瞬时转染入 XBP1 稳定性细胞系 (5-2, 5-4 和 20-10) 和亲代 (CHO DUKX) 细胞系。将转染后 48 小时所收集的条件培养基在 10% tricine 凝胶上电泳, 随后进行 Western 印迹分析。Western 印迹膜 (图 3) 用与 BMP6 强烈交叉反应的小鼠抗 BMP5 单克隆抗体 (1:2000) 探测。二级抗体是辣根过氧化物酶缀合的山羊抗小鼠抗体 (1:5000) (The Jackson Laboratory)。图 3 中每一泳道代表对各自细胞系的独立实验。如图 3 中所示, 以 5nM 和 10nM MTX 选择的 XBP1 稳定性细胞系中比亲代细胞中分泌更多 BMP6。

还将编码 IL11RFc 蛋白的 pSMED2/IL11RFc 瞬时转化至 XBP1 稳定性细胞系和亲代细胞系。将转染后 48 小时所收集的条件培养基在 10% tricine 凝胶上电泳, 随后进行 Western 印迹分析。Western 印迹膜 (图 4) 用辣根过氧化物酶缀合的山羊抗人 Fc 抗体 (The Jackson Laboratory)。在大部分的用 5nM MTX 所选择的 XBP1 细胞系中比 CHO DUKX 细胞中分泌更多 IL11RFc。

实施例 4. 比较 XBP1 稳定性细胞系和它们的亲代 CHO 细胞间的转染、转录效率和翻译效率。

将 5-2、20-10 和 CHO DUKX 细胞计数并且将等量细胞以编码绿色荧光蛋白 (GFP) 的构建体瞬时转染。细胞在 10×10 放大倍率下观察并且通过每一细胞系在三个视野内比较 GFP 荧光细胞与总细胞而测定转染效率 (% 表达 GFP 的细胞)。比较结果显示 GFP 转染效率在所测试的全部 XBP1 和

CHO DUKX 细胞系中相似(图 5)。

在又一个实验中,将编码 GFP (+)或不编码 GFP (-)的构建体瞬时转染至 XBP1 稳定性细胞系和亲代细胞系。将转染后 48 小时所收集的细胞裂解物在 10% tricine 凝胶上电泳,随后进行 Western 印迹分析(图 6)。图 6 中的每一泳道代表独立实验。Western 印迹膜用兔抗 GFP 多克隆抗体以及用于载量对照的小鼠抗肌动蛋白单克隆抗体探测。如图 6 显示,对 GFP 的转录效率和翻译效率总量在所研究的全部细胞系中相似。

实施例 5. 用 ATF6 瞬时转染的细胞

将加 Flag 标签的来自 ATF6 可溶性活化结构域的 cDNA 克隆至 Tet/关闭诱导型表达载体 ptTATOP6 并且瞬时转染入 COS-1 细胞。ptTATOP6 载体包括了控制包含 ATF6 和 Flag 标签的融和蛋白表达的诱导型启动子。将转染后 18、48 和 60 小时所收集的细胞裂解物在 10% tricine 凝胶上电泳,随后进行 Western 印迹分析。Western 印迹膜(图 7)用抗 flag 抗体探测。“V”表示空的 ptTATOP6 载体,并且“P”代表 flag 阳性对照。如图 7 显示,ATF6 蛋白在缺乏强力霉素时在 COS-1 细胞内成功表达。

实施例 6. 正确比率的靶基因与 XBP1 或 ATF 6 在 HEK293 细胞中的共表达增强靶基因分泌

HEK293-FT 和 HEK293-EBNA 在具有 5% CO₂ 的加湿培养箱中在 37°C 于补加 5%胎牛血清的 free-style293 培养基(Invitrogen, Carlsbad, CA)内生长和维持。

瞬时表达在 50ml 旋转器或 24 孔平板或 1L 旋转器内实施。对 50 ml (或 1L)的培养体积,将 25µg (或对 1L 用 0.5mg)质粒 DNA 与 400µg (对 1L 用 8mg)聚乙烯亚胺 (PEI, 25 kDa, 线性、用 HCl 中和至 pH7.0, 1mg/ml, Polysciences, Warrington, PA)在 2.5 ml (对 1L 用 50ml)无血清 293 培养基中混合。收获前旋转器在 37°C 于 P2005 Stirrer (Bellco)上以每分钟 170 转的旋转速率温育 72-144 小时。对 24 孔平板,将 1µg DNA 与 8µg PEI 在 0.5ml 无血清 293 培养基中混合。随后将混合物与在具有 10% FBS 的 293 培养基内的细胞密度 0.5×10^6 个细胞/ml 的 0.5ml HEK293 细胞混合。收获前平板

在 37°C 在定轨摇床(BeIICo) 以每分钟 300 的转旋转速率温育 72 小时。

pSMED2 和 pSMEDA 用于 DNA 构建。将 C 端加 His6 标签的分泌型卷曲相关蛋白 1 (sFRP-1) 和 C 端加 Flag 标签的 aggrecanase-2 (Agg-2) 亚克隆至 pSMEDA。将 C 端加 His6 标签的神经营养酪氨酸激酶受体 2 型 (TrkB) 亚克隆至 pSMED2。这些亚克隆的基因没有任何跨膜结构域或胞浆结构域，因此允许所表达产物分泌。

将 C 端加 His6 标签的 Prop1 和 Prop34-LBD 亚克隆至 pcDNA3.1 (Invitrogen, Carlsbad, CA)。Prop1 和 Prop34-LBD 衍生自低密度脂蛋白受体相关蛋白 5 (LRP5)，具有缺失的跨膜结构域和胞浆结构域。Prop1 和 Prop34-LBD 的氨基酸序列分别在 SEQ ID NO: 10 和 11 中描述。SEQ ID NO: 10 包括在氨基酸 342-347 的 His6 标签和在氨基酸 348-356 的 Flag 标签。SEQ ID NO: 11 括在氨基酸 795-800 的 His6 标签和在氨基酸 778-794 的 V5 标签。

将 1 μ g 的 pcDNA3.1 中的 Prop1-his6-Flag 与 0.3 μ g (1:3) 或 1 μ g (1:1) 的 pSMED2 载体中的 XBP1p 共转染至 HEK293T 细胞。pcDNA3.1 和 pSMED2 均由 CMV 启动子驱动。条件培养基在 DNA 转染后 72 小时收获。样品通过 SDS-PAGE 分离并且用抗 His4 抗体进行免疫印迹。进行重复实验(组 # 1 和组 # 2)。如图 8A 中所示，XBP1 与 Prop1 以 1:1 或 1:3 比率共转染显著地提高 Prop1 表达。

在另一个实验中，将 1 μ g 的 pcDNA3.1 中的 Prop34-LBD-V5-his6 与 0.3 μ g (1:3) 或 1 μ g (1:1) 的 ptTATOP6 载体中的 ATF6 共转染至 HEK293T 细胞。如同 pSMED2，ptTATOP6 也由 CMV 启动子驱动。条件培养基在 DNA 转染后 72 小时通过 SDS-PAGE 分析并且用抗 His4 抗体进行免疫印迹。图 8B 中显示 ATF6 与 Prop34-LBD-V5-his6 以 1:1 或 1:3 比率共转染显著地增强 Prop34-LBD 表达。

图 9 显示 XBP1p 或 ATF6 对不同靶蛋白表达的影响。将 pcDNA3.1 中的 Prop1-his6-Flag 或 Prop34-LBD-V5-His6 与 pSMED2 载体中的 XBP1p 或 ptTATOP6 载体中的 ATF-6 共转染至 HEK293T 细胞。条件培养基在

DNA 转染后 72 小时通过 SDS-PAGE 分析并且用抗 His4 抗体进行免疫印迹。信号通过密度计定量如图 9 中所示。结果表明 XBP1p 和 ATF-6 对 Prop1 和 Prop34-LBD 的表达具有不同影响。还观察到当 TrkB、sFRP-1 和 Agg-2 与 XBP1p 和 ATF-6 共表达时, XBP1p 和 ATF-6 对这些蛋白质不同的增强效果。例如, 与 XBP1p 的共表达提高 TrkB 产率大约 5 倍, 与此同时当 TrkB 与 ATF6 共表达时仅观察到提高大约两倍。

本发明的上述描述提供示例和描述, 但不打算排他性的或将本发明限于所公开的具体之一。符合以上教导的修改和变化可能或可以从本发明的实践中获得。因此应当指出的是本发明范围由权利要求书和它们的等同物定义。

<110> 惠氏公司

<120> 用于产生蛋白质的体系和方法

<130> 031896-085001 (AM101741)

<150> US 60/606,439

<151> 2004-09-02

<160> 11

<170> PatentIn 版本 3.3

<210> 1

<211> 19

<212> DNA

<213> 哺乳动物

<220>

<221> misc_feature

<222> (6)..(14)

<223> n 为 a, c, g, 或 t

<400> 1
ccaatnnnnn nnnnccacg 19

<210> 2

<211> 11

<212> DNA

<213> 哺乳动物

<220>

<221> misc_feature

<222> (6)..(6)

<223> n 为 a, c, g, 或 t

<400> 2
attggnccac g 11

<210> 3

<211> 8

<212> DNA

<213> 哺乳动物

<400> 3	
tgacgtgg	8
<210> 4	
<211> 8	
<212> DNA	
<213> 哺乳动物	
<400> 4	
tgacgtga	8
<210> 5	
<211> 1154	
<212> DNA	
<213> 人 (Homo sapiens)	
<400> 5	
atggtggtgg tggcagccgc gccgaacccg gccgacggga ccctaaagt tctgcttctg	60
tcggggcagc ccgcctccgc cgccggagcc ccggccggcc aggccctgcc gctcatggtg	120
ccagcccaga gaggggccag cccggaggca gcgagcgggg ggctgcccc ggcgcgcaag	180
cgacagcgcc tcacgcacct gagccccgag gagaaggcgc tgaggaggaa actgaaaaac	240
agagtagcag ctcagactgc cagagatcga aagaaggctc gaatgagtga gctggaacag	300
caagtggtag atttagaaga agagaaccaa aaacttttgc tagaaaatca gcttttacga	360
gagaaaactc atggccttgt agttgagaac caggagttaa gacagcgctt ggggatggat	420
gccctggttg ctgaagagga ggcggaagcc aaggggaaatg aagtgaggcc agtggccggg	480
tctgctgagt ccgcagcact cagactacgt gcacctctgc agcaggtgca ggcccagttg	540
tcaccctcc agaacatctc cccatggatt ctggcggtat tgactcttca gattcagagt	600
ctgatatcct gttgggcatt ctggacaact tggaccagct catgttcttc aaatgcctt	660
ccccagagcc tgccagcctg gaggagctcc cagaggtcta ccagaagga cccagttcct	720
taccagctc cttttctctg tcagtgggga cgatcatcagc caagctggaa gccattaatg	780
aactaattcg ttttgaccac atatatacca agcccctagt cttagagata ccctctgaga	840
cagagagcca agctaattgt gtagtgaaaa tcgaggaagc acctctcagc ccctcagaga	900
atgatcacc cgaattcatt gtctcagtga aggaagaacc tgtagaagat gacctcgttc	960

cggagctggg tatctcaaat ctgctttcat ccagccactg cccaaagcca tcttcctgcc 1020
 tactggatgc ttacagtgc tgtggatacg ggggttcctt ttccccattc agtgacatgt 1080
 cctctctgct tgggtgaaac cattcttggg aggacacttt tgccaatgaa ctctttcccc 1140
 agctgattag tgtc 1154

<210> 6
 <211> 376
 <212> PRT
 <213> 人

<400> 6

Met Val Val Val Ala Ala Ala Pro Asn Pro Ala Asp Gly Thr Pro Lys
 1 5 10 15

Val Leu Leu Leu Ser Gly Gln Pro Ala Ser Ala Ala Gly Ala Pro Ala
 20 25 30

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

Glu Ala Ala Ser Gly Gly Leu Pro Gln Ala Arg Lys Arg Gln Arg Leu
 50 55 60

Thr His Leu Ser Pro Glu Glu Lys Ala Leu Arg Arg Lys Leu Lys Asn
 65 70 75 80

Arg Val Ala Ala Gln Thr Ala Arg Asp Arg Lys Lys Ala Arg Met Ser
 85 90 95

Glu Leu Glu Gln Gln Val Val Asp Leu Glu Glu Glu Asn Gln Lys Leu
 100 105 110

Leu Leu Glu Asn Gln Leu Leu Arg Glu Lys Thr His Gly Leu Val Val
 115 120 125

Glu Asn Gln Glu Leu Arg Gln Arg Leu Gly Met Asp Ala Leu Val Ala

130	135	140	
Glu Glu Glu Ala Glu Ala Lys Gly Asn Glu Val Arg Pro Val Ala Gly			
145	150	155	160
Ser Ala Glu Ser Ala Ala Gly Ala Gly Pro Val Val Thr Pro Pro Glu			
	165	170	175
His Leu Pro Met Asp Ser Gly Gly Ile Asp Ser Ser Asp Ser Glu Ser			
	180	185	190
Asp Ile Leu Leu Gly Ile Leu Asp Asn Leu Asp Pro Val Met Phe Phe			
	195	200	205
Lys Cys Pro Ser Pro Glu Pro Ala Ser Leu Glu Glu Leu Pro Glu Val			
	210	215	220
Tyr Pro Glu Gly Pro Ser Ser Leu Pro Ala Ser Leu Ser Leu Ser Val			
	225	230	235
Gly Thr Ser Ser Ala Lys Leu Glu Ala Ile Asn Glu Leu Ile Arg Phe			
	245	250	255
Asp His Ile Tyr Thr Lys Pro Leu Val Leu Glu Ile Pro Ser Glu Thr			
	260	265	270
Glu Ser Gln Ala Asn Val Val Val Lys Ile Glu Glu Ala Pro Leu Ser			
	275	280	285
Pro Ser Glu Asn Asp His Pro Glu Phe Ile Val Ser Val Lys Glu Glu			
	290	295	300
Pro Val Glu Asp Asp Leu Val Pro Glu Leu Gly Ile Ser Asn Leu Leu			
	305	310	315
Ser Ser Ser His Cys Pro Lys Pro Ser Ser Cys Leu Leu Asp Ala Tyr			
	325	330	335

Ser Asp Cys Gly Tyr Gly Gly Ser Leu Ser Pro Phe Ser Asp Met Ser
 340 345 350

Ser Leu Leu Gly Val Asn His Ser Trp Glu Asp Thr Phe Ala Asn Glu
 355 360 365

Leu Phe Pro Gln Leu Ile Ser Val
 370 375

<210> 7

<211> 1128

<212> DNA

<213> 人

<400> 7

atggtggtgg tggcagccgc gccgaacccg gccgacggga cccctaaagt tctgcttctg	60
tcggggcagc ccgcctccgc cgccggagcc ccggccggcc aggccttgcc gctcatggtg	120
ccagcccaga gaggggcccag cccggaggca gcgagcgggg ggctgcccc a ggcgcgcaag	180
cgacagcgcc tcacgcacct gagccccgag gagaaggcgc tgaggaggaa actgaaaaac	240
agagtagcag ctccagactgc cagagatcga aagaaggctc gaatgagtga gctggaacag	300
caagtggtag atttagaaga agagaaccaa aaacttttgc tagaaaatca gcttttacga	360
gagaaaactc atggccttgt agttgagaac caggagttaa gacagcgctt ggggatggat	420
gccctggttg ctgaagagga ggcggaagcc aaggggaatg aagtgaggcc agtggccggg	480
tctgctgagt ccgcagcagg tgcaggcccc gttgtcacc ctccagaaca tctccccatg	540
gattctggcg gtattgactc ttcagattca gagtctgata tcctgttggg cattctggac	600
aacttgacc cagtcattgtt cttcaaatgc cttccccag agcctgccag cctggaggag	660
ctcccagagg tctaccaga aggaccagc tccttaccag cctcccttcc tctgtcagtg	720
gggacgtcat cagccaagct ggaagccatt aatgaactaa ttcgttttga ccacatatat	780
accaagcccc tagtcttaga gataccctct gagacagaga gccaaagctaa tgtggtagtg	840
aaaatcgagg aagcacctct cagcccctca gagaatgatc accctgaatt cattgtctca	900
gtgaaggaag aacctgtaga agatgacctc gttccggagc tgggtatctc aaatctgctt	960

```

tcatccagcc actgccccaa gccatcttcc tgcctactgg atgcttacag tgactgtgga 1020
tacgggggtt ccctttcccc attcagtgac atgtcctctc tgcttggtgt aaaccattct 1080
tgggaggaca cttttgccaa tgaactcttt ccccagctga ttagtgtc 1128

<210> 8
<211> 2010
<212> DNA
<213> 人

<400> 8
atgggggagc cggctggggt tgccggcacc atggagtcac ctttagccc gggactcttt 60
cacaggctgg atgaagattg ggattctgct ctctttgctg aacttggtta ttcacagac 120
actgatgagc tgcaattgga agcagcaa at gagacgtatg aaaacaattt tgataatctt 180
gattttgatt tggatttggt accttgggag tcagacattt gggacatcaa caaccaa atc 240
tgtacagtta aagatattaa ggcagaaccc cagccacttt ctccagcctc ctcaagttat 300
tcagtctcat ctctcggtc agtggactct tattcttcaa ctccagcatgt tcctgaggag 360
ttggatttgt cttctagttc tcagatgtct cccctttcct tatatggtga aaactcta at 420
agtctctctt caccggagcc actgaaggaa gataagcctg tcaactggtc taggaacaag 480
actgaaaatg gactgactcc aaagaaaaaa attcaggtga attcaaaacc ttcaattcag 540
cccaagcctt tattgcttcc agcagcacc aagactcaaa caaactccag tgttccagca 600
aaaaccatca ttattcagac agtaccaacg cttatgccat tggcaaagca gcaaccaatt 660
atcagtttac aacctgcacc cactaaaggc cagacggttt tgctgtctca gcctactgtg 720
gtacaacttc aagcacctgg agttctgccc tctgtcagc cagtccttgc tgttgctggg 780
ggagtcacac agctccctaa tcacgtggtg aatgtggtac cagccccctc agcgaatagc 840
ccagtgaatg gaaaactttc cgtgactaaa cctgtcctac aaagtacat gagaaatgtc 900
ggttcagata ttgctgtgct aaggagacag caacgtatga taaaaaatcg agaatccgct 960
tgtcagtcct gcaagaagaa gaaagaatat atgctagggt tagaggcgag attaaaggct 1020
gccctctcag aaaacgagca actgaagaaa gaaaatggaa cactgaagcg gcagctggat 1080

```

gaagttgtgt cagagaacca gaggcttaaa gtcctagtc caaagcgaag agttgtctgt 1140
 gtgatgatag tattggcatt tataatactg aactatggac ctatgagcat gttggaacag 1200
 gattccagga gaatgaacct tagtgtggga cctgcaaatac aaaggaggca ctttctagga 1260
 ttttctgcta aagaggcaca ggacacatca gatggtatta tccagaaaaa cagctacaga 1320
 tatgatcatt ctgtttcaaa tgacaaagcc ctgatgggtc taactgaaga accattgctt 1380
 tacattcccc cacctccttg tcagccccta attaatacaa cagagtctct caggttaaat 1440
 catgaacttc gaggatgggt tcatagacat gaagtagaaa ggaccaagtc tagaagaatg 1500
 acaaataatc aacagaaaac ccgtattctt cagggtgttg tggaacaggg ctcaaattct 1560
 cagctgatgg ctgttcaata cacagaaacc actagtagta tcagcaggaa ctcagggagt 1620
 gagctacaag tgtattatgc ttcaccaga agttatcaag acttttttga agccatccgc 1680
 agaaggggag acacatttta tgttgtgtca tttcgaaggg atcacctgct gttaccagct 1740
 accaccata acaagaccac aagacaaaa atgtcaattg tgttaccagc aataaacata 1800
 aatgagaatg tgatcaatgg gcaggactac gaagtgatga tgcagattga ctgtcagggt 1860
 atggacacca ggatectcca tatcaaaagt tcgtcgggtc ctccttacct ccgagatcag 1920
 cagaggaatc aaaccaacac cttctttggc tcccctcccg cagccacaga ggcaaccac 1980
 gttgtcagca ccatccctga gtcattacaa 2010

<210> 9

<211> 670

<212> PRT

<213> 人

<400> 9

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

Pro Gly Leu Phe His Arg Leu Asp Glu Asp Trp Asp Ser Ala Leu Phe
20 25 30

Ala Glu Leu Gly Tyr Phe Thr Asp Thr Asp Glu Leu Gln Leu Glu Ala
35 40 45

Ala Asn Glu Thr Tyr Glu Asn Asn Phe Asp Asn Leu Asp Phe Asp Leu
50 55 60

Asp Leu Leu Pro Trp Glu Ser Asp Ile Trp Asp Ile Asn Asn Gln Ile
65 70 75 80

Cys Thr Val Lys Asp Ile Lys Ala Glu Pro Gln Pro Leu Ser Pro Ala
85 90 95

Ser Ser Ser Tyr Ser Val Ser Ser Pro Arg Ser Val Asp Ser Tyr Ser
100 105 110

Ser Thr Gln His Val Pro Glu Glu Leu Asp Leu Ser Ser Ser Ser Gln
115 120 125

Met Ser Pro Leu Ser Leu Tyr Gly Glu Asn Ser Asn Ser Leu Ser Ser
130 135 140

Pro Glu Pro Leu Lys Glu Asp Lys Pro Val Thr Gly Ser Arg Asn Lys
145 150 155 160

Thr Glu Asn Gly Leu Thr Pro Lys Lys Lys Ile Gln Val Asn Ser Lys
165 170 175

Pro Ser Ile Gln Pro Lys Pro Leu Leu Leu Pro Ala Ala Pro Lys Thr
180 185 190

Gln Thr Asn Ser Ser Val Pro Ala Lys Thr Ile Ile Ile Gln Thr Val
195 200 205

Pro Thr Leu Met Pro Leu Ala Lys Gln Gln Pro Ile Ile Ser Leu Gln
210 215 220

Pro Ala Pro Thr Lys Gly Gln Thr Val Leu Leu Ser Gln Pro Thr Val
225 230 235 240

Val Gln Leu Gln Ala Pro Gly Val Leu Pro Ser Ala Gln Pro Val Leu
245 250 255

Ala Val Ala Gly Gly Val Thr Gln Leu Pro Asn His Val Val Asn Val
260 265 270

Val Pro Ala Pro Ser Ala Asn Ser Pro Val Asn Gly Lys Leu Ser Val
275 280 285

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

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

Cys Gln Ser Arg Lys Lys Lys Lys Glu Tyr Met Leu Gly Leu Glu Ala
325 330 335

Arg Leu Lys Ala Ala Leu Ser Glu Asn Glu Gln Leu Lys Lys Glu Asn
340 345 350

Gly Thr Leu Lys Arg Gln Leu Asp Glu Val Val Ser Glu Asn Gln Arg
355 360 365

Leu Lys Val Pro Ser Pro Lys Arg Arg Val Val Cys Val Met Ile Val
370 375 380

Leu Ala Phe Ile Ile Leu Asn Tyr Gly Pro Met Ser Met Leu Glu Gln
385 390 395 400

Asp Ser Arg Arg Met Asn Pro Ser Val Gly Pro Ala Asn Gln Arg Arg
405 410 415

His Leu Leu Gly Phe Ser Ala Lys Glu Ala Gln Asp Thr Ser Asp Gly
420 425 430

Ile Ile Gln Lys Asn Ser Tyr Arg Tyr Asp His Ser Val Ser Asn Asp
435 440 445

Lys Ala Leu Met Val Leu Thr Glu Glu Pro Leu Leu Tyr Ile Pro Pro
450 455 460

Pro Pro Cys Gln Pro Leu Ile Asn Thr Thr Glu Ser Leu Arg Leu Asn
465 470 475 480

His Glu Leu Arg Gly Trp Val His Arg His Glu Val Glu Arg Thr Lys
485 490 495

Ser Arg Arg Met Thr Asn Asn Gln Gln Lys Thr Arg Ile Leu Gln Gly
500 505 510

Val Val Glu Gln Gly Ser Asn Ser Gln Leu Met Ala Val Gln Tyr Thr
515 520 525

Glu Thr Thr Ser Ser Ile Ser Arg Asn Ser Gly Ser Glu Leu Gln Val
530 535 540

Tyr Tyr Ala Ser Pro Arg Ser Tyr Gln Asp Phe Phe Glu Ala Ile Arg
545 550 555 560

Arg Arg Gly Asp Thr Phe Tyr Val Val Ser Phe Arg Arg Asp His Leu
565 570 575

Leu Leu Pro Ala Thr Thr His Asn Lys Thr Thr Arg Pro Lys Met Ser
580 585 590

Ile Val Leu Pro Ala Ile Asn Ile Asn Glu Asn Val Ile Asn Gly Gln
595 600 605

Asp Tyr Glu Val Met Met Gln Ile Asp Cys Gln Val Met Asp Thr Arg
610 615 620

Ile Leu His Ile Lys Ser Ser Ser Val Pro Pro Tyr Leu Arg Asp Gln
625 630 635 640

Gln Arg Asn Gln Thr Asn Thr Phe Phe Gly Ser Pro Pro Ala Ala Thr
 645 650 655

Glu Ala Thr His Val Val Ser Thr Ile Pro Glu Ser Leu Gln
 660 665 670

<210> 10
 <211> 356
 <212> PRT
 <213> 人

<400> 10

Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
 1 5 10 15

Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
 20 25 30

Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
 35 40 45

Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
 50 55 60

Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
 65 70 75 80

Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
 85 90 95

Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
 100 105 110

Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
 115 120 125

Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
 130 135 140

Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala Leu Asp Pro
145 150 155 160

Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
165 170 175

Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
180 185 190

Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
195 200 205

Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
210 215 220

Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
225 230 235 240

Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
245 250 255

Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
260 265 270

Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
275 280 285

Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
290 295 300

Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
305 310 315 320

Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly Arg Thr Cys
325 330 335

Lys Ala Gly Ala Glu His His His His His His Asp Tyr Lys Asp Asp

	340		345		350
Asp Asp Lys Ile					
355					
<210>	11				
<211>	800				
<212>	PRT				
<213>	人				
<400>	11				
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro					
1	5		10		15
Gly Ser Thr Gly Asp Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Val					
	20		25		30
Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His Arg Ile					
	35		40		45
Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr Gly Val					
	50		55		60
Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His Ile Tyr					
65		70		75	80
Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg Ala Phe Met Asn Gly					
	85		90		95
Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro Glu Gly					
	100		105		110
Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp Thr Gly					
	115		120		125
Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg Gln Val					
	130		135		140

Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu Asp Pro
145 150 155 160

Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro Arg Ile
165 170 175

Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val Asp Lys
180 185 190

Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln Arg Leu
195 200 205

Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn Met Leu
210 215 220

Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro Phe Gly
225 230 235 240

Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn Leu His
245 250 255

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

Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His Ser Ser
275 280 285

Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln Cys Gly
290 295 300

Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys Ala Ser
305 310 315 320

His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro Thr Thr
325 330 335

Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile Pro Asp
340 345 350

Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu Arg Asn
355 360 365

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

Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr Gln Pro
385 390 395 400

Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg Gln Pro
405 410 415

His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp Thr Cys
420 425 430

Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu Ala Met
435 440 445

Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile Val Val
450 455 460

Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp Arg Ala
465 470 475 480

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

Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp Asn Thr
500 505 510

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

Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala Asn Ile
530 535 540

Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr Trp Ile
545 550 555 560

Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr Gly Asp
565 570 575

Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly Ile His
580 585 590

Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro Cys Ala
595 600 605

Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly Asp Gly
610 615 620

Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln Asn Leu
625 630 635 640

Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe Ala Cys
645 650 655

Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys Asp Gly
660 665 670

Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro Val Cys
675 680 685

Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp Leu Arg
690 695 700

Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp Glu Ala
705 710 715 720

Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala Ser Gly
725 730 735

Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp Cys Ile
740 745 750

Asp	Gly	Ser	Asp	Glu	Leu	Met	Cys	Glu	Ile	Thr	Lys	Pro	Pro	Ser	Arg
		755					760					765			
Pro	Leu	Glu	Ser	Arg	Gly	Pro	Phe	Glu	Gly	Lys	Pro	Ile	Pro	Asn	Pro
		770					775					780			
Leu	Leu	Gly	Leu	Asp	Ser	Thr	Arg	Thr	Gly	His	His	His	His	His	His
					790					795					800

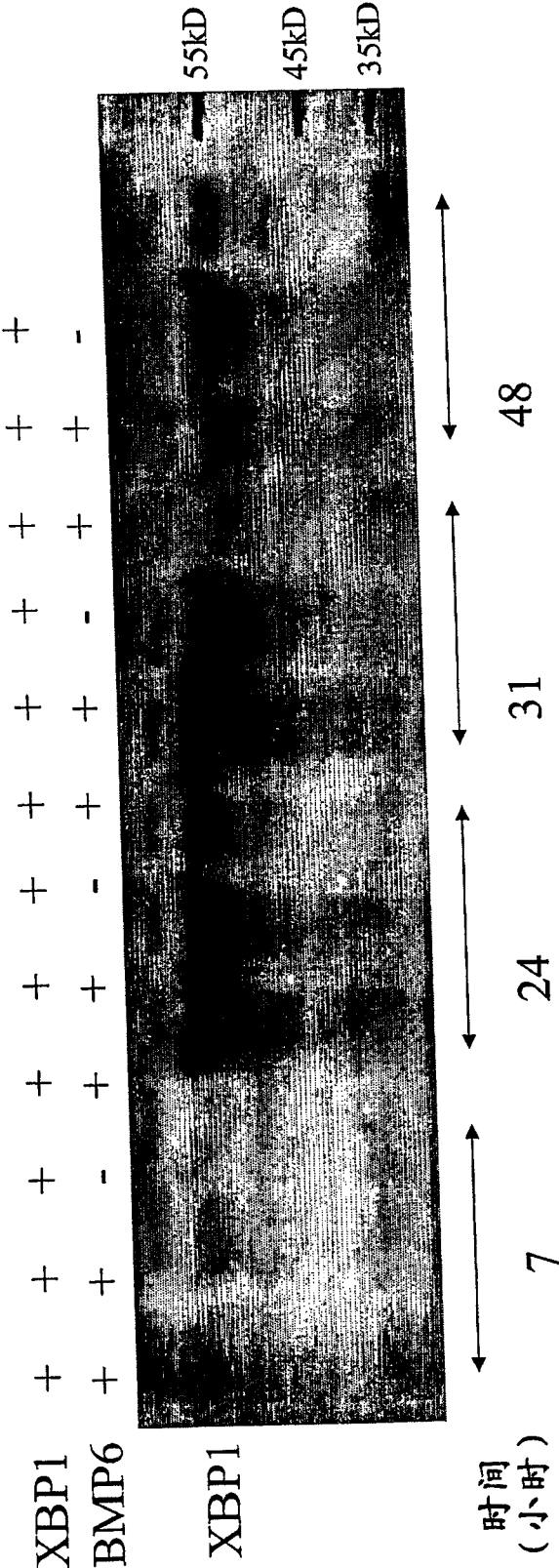


图 1

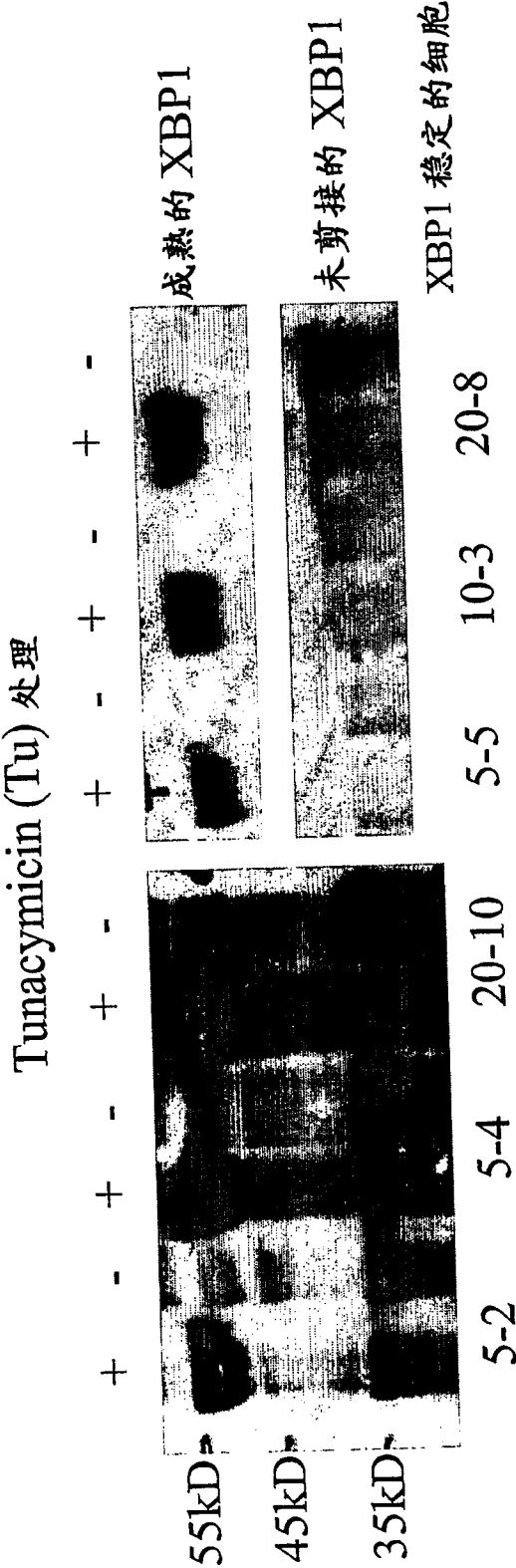


图 2

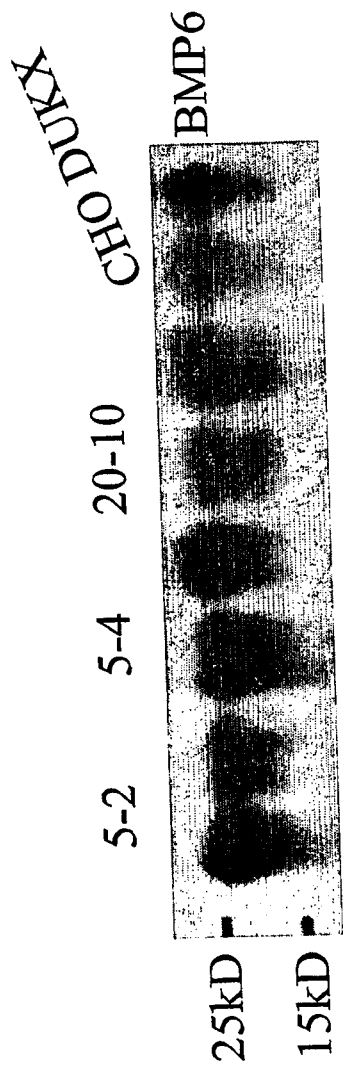


图 3



图 4

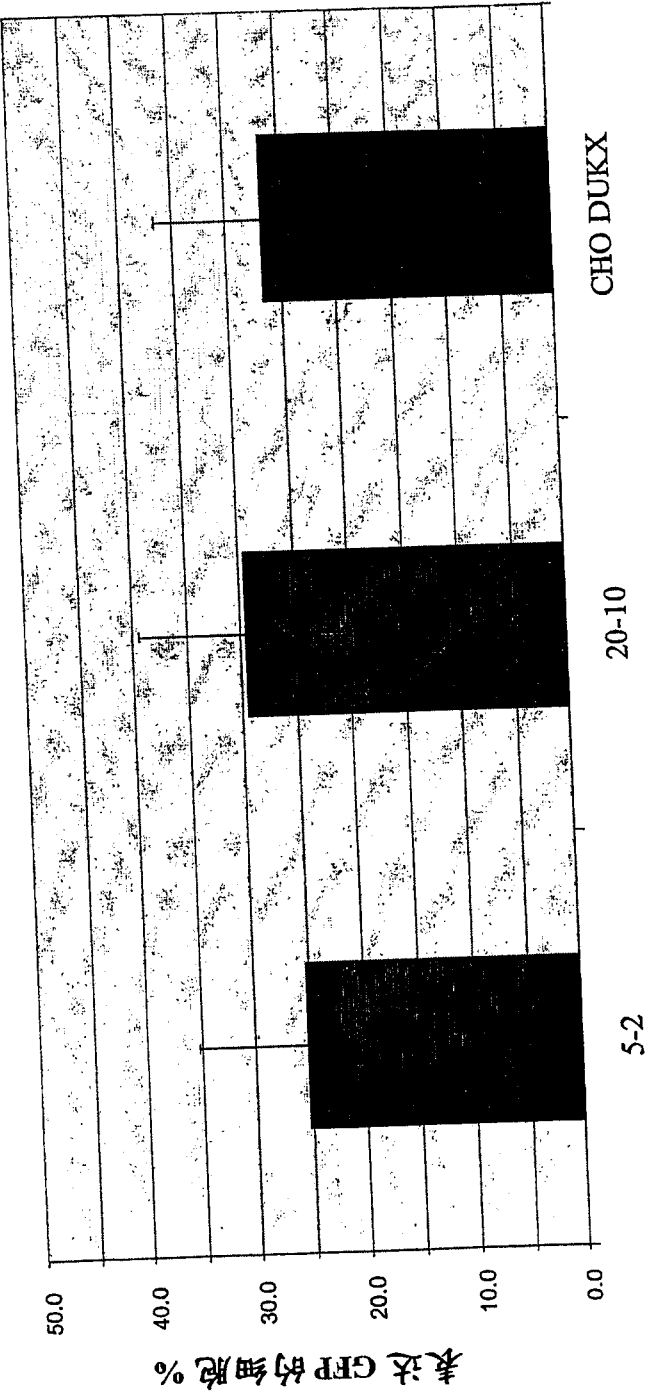


图 5

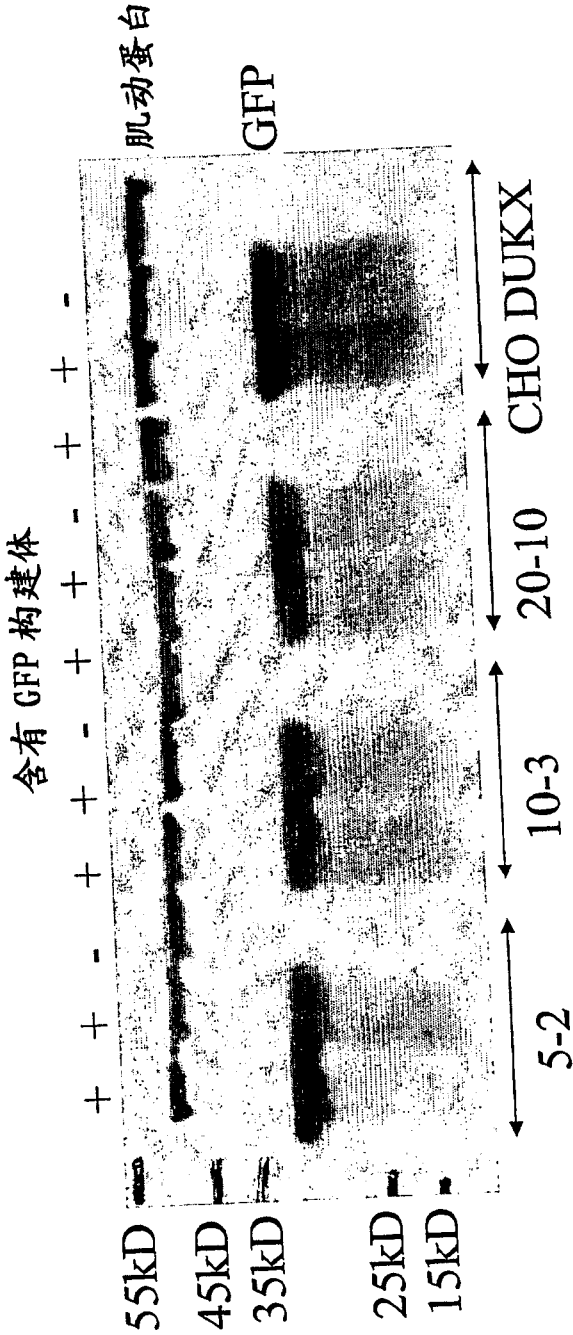


图 6

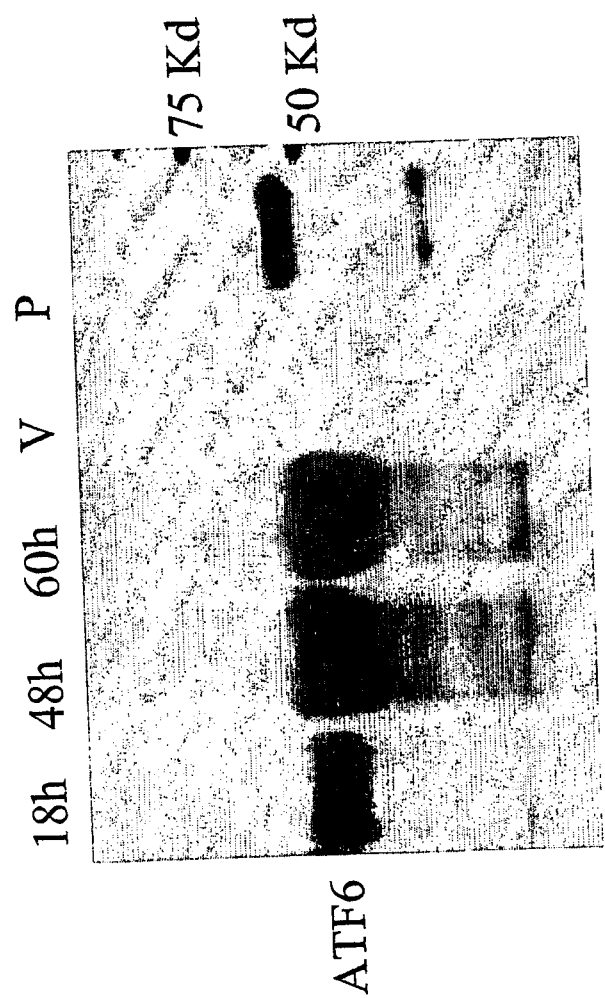


图 7

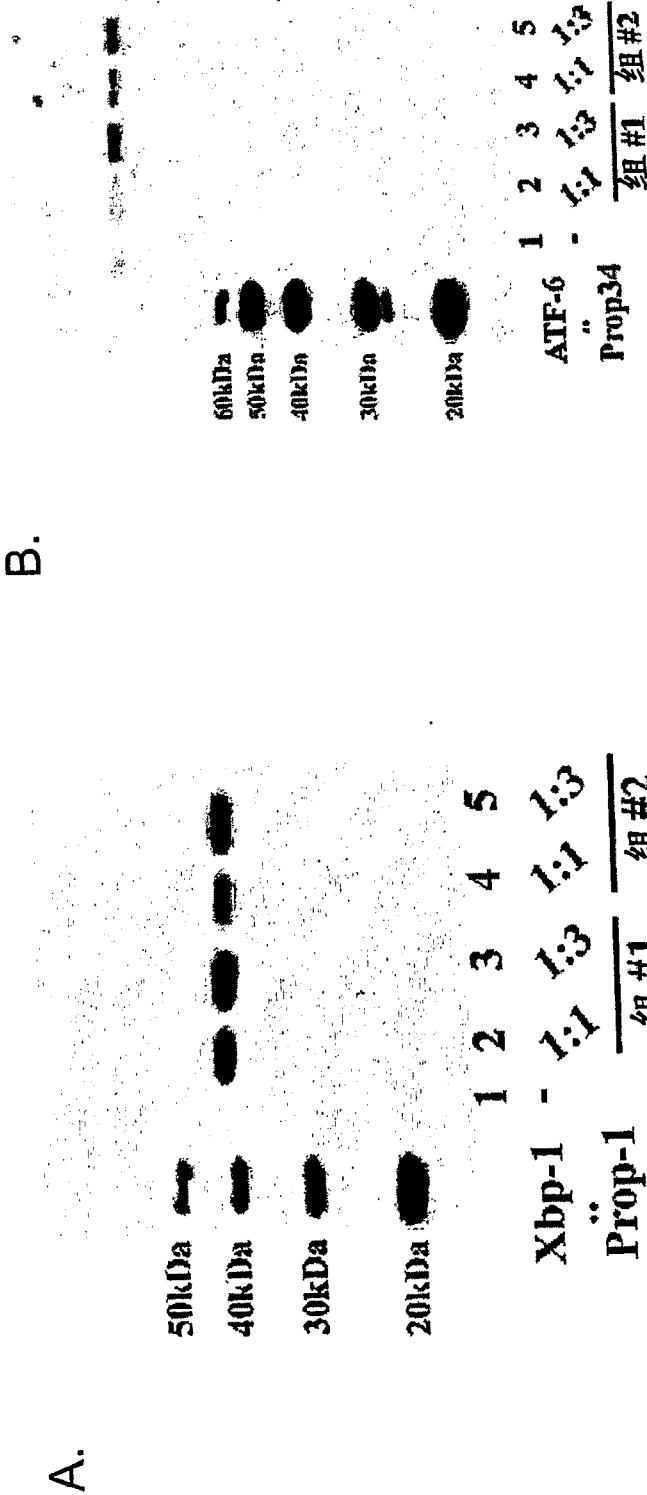


图 8

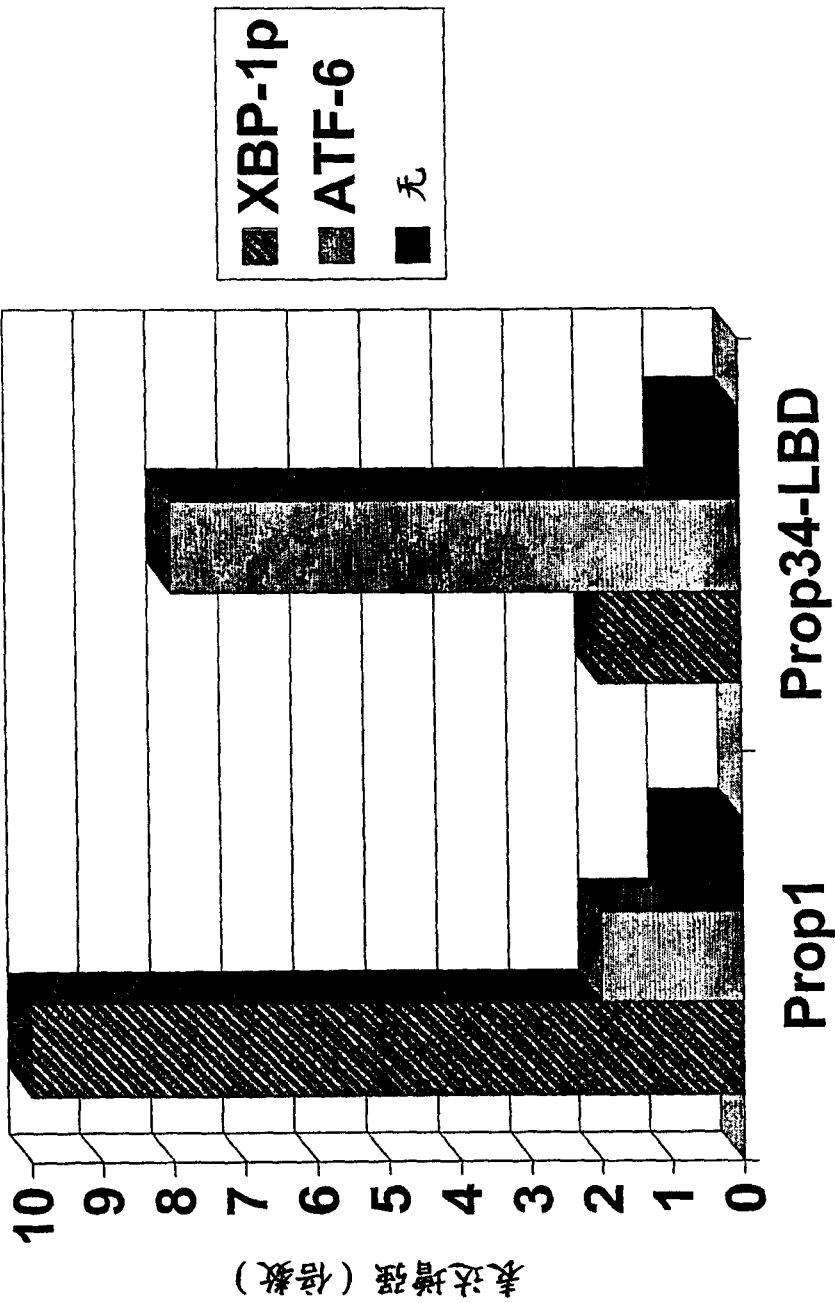


图 9