



(12) 发明专利

(10) 授权公告号 CN 111100197 B

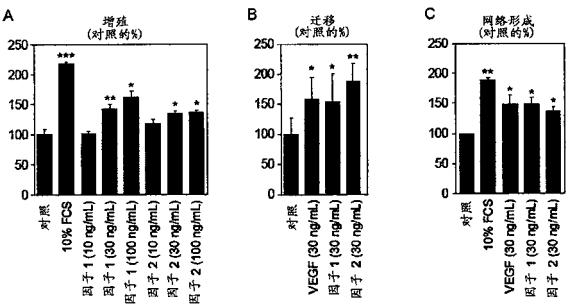
(45) 授权公告日 2024. 11. 15

(21) 申请号 201911137853.9  
(22) 申请日 2014.01.16  
(65) 同一申请的已公布的文献号  
    申请公布号 CN 111100197 A  
(43) 申请公布日 2020.05.05  
(30) 优先权数据  
    13151593.4 2013.01.17 EP  
(62) 分案原申请数据  
    201480005079.5 2014.01.16  
(73) 专利权人 汉诺威医学院  
    地址 德国汉诺威  
(72) 发明人 卡伊·克里斯托夫·沃勒特  
    莫蒂默·科尔夫-克林格比尔  
(74) 专利代理机构 北京集佳知识产权代理有限公司 11227  
    专利代理师 顾晋伟 尹玉峰  
(51) Int.Cl.  
    C07K 14/47 (2006.01)

C07K 16/18 (2006.01)  
C12N 15/12 (2006.01)  
C12N 15/13 (2006.01)  
A61K 38/17 (2006.01)  
A61K 39/395 (2006.01)  
A61P 35/00 (2006.01)  
A61P 19/02 (2006.01)  
A61P 29/00 (2006.01)  
A61P 17/06 (2006.01)  
A61P 27/02 (2006.01)  
A61P 9/14 (2006.01)  
A61P 7/04 (2006.01)  
A61P 17/02 (2006.01)  
A61P 9/10 (2006.01)  
A61P 1/00 (2006.01)  
A61P 17/00 (2006.01)  
(56) 对比文件  
    US 2008004232 A1, 2008.01.03  
    审查员 管冰  
    权利要求书1页 说明书34页  
    序列表30页 附图11页

(54) 发明名称  
    用于治疗或预防疾病的因子1蛋白、因子2蛋白及其抑制剂

(57) 摘要  
    本发明涉及用于治疗或预防疾病的因子1蛋白、因子2蛋白及其抑制剂。本发明涉及包含由来源于人染色体区C19orf10和/或C19orf63的核酸编码的氨基酸序列的蛋白质(称为因子1和/或因子2)或者其抑制剂,所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。本发明还涉及因子1和因子2的抑制剂用于医学用途的用途,优选用于治疗或预防其中血管发生有助于疾病发生或进展之疾病的用途。



1. 由根据SEQ ID NO:2的氨基酸序列组成的蛋白质或SEQ ID NO:2所示氨基酸序列的N端1-30位氨基酸缺失后形成的蛋白质用于制造药物的用途,其中所述药物用于治疗心肌梗死或再灌注损伤。

2. 权利要求1所述的用途,其中所述药物配制成通过经口、静脉内、粘膜内、动脉内、肌内或冠状动脉内途径施用。

3. 权利要求1或2所述的用途,其中所述药物配制成在再灌注疗法之前、与其同时或在其之后施用。

4. 权利要求3所述的用途,其中所述施用通过一次或更多次快速推注和/或输注进行。

5. 核酸用于制造用于治疗心肌梗死或再灌注损伤之药物的用途,所述核酸编码权利要求1中所述的蛋白质。

6. 权利要求5所述的用途,其中所述药物配制成通过经口、静脉内、粘膜内、动脉内、肌内或冠状动脉内途径施用。

7. 权利要求5或6所述的用途,其中所述药物配制成在再灌注疗法之前、与其同时或在其之后施用。

8. 权利要求7所述的用途,其中所述施用通过一次或更多次快速推注和/或输注进行。

9. 载体用于制造用于治疗心肌梗死或再灌注损伤之药物的用途,所述载体包含权利要求5中所述的核酸。

10. 根据权利要求9所述的用途,其中所述载体选自:质粒载体;粘粒载体;噬菌体载体;病毒载体;以及细菌芽孢。

11. 根据权利要求10所述的用途,其中所述噬菌体载体选自 $\lambda$ 噬菌体载体和丝状噬菌体载体。

12. 根据权利要求10所述的用途,其中所述病毒载体选自腺病毒载体、腺相关病毒(AAV)载体、甲病毒载体、疱疹病毒载体、麻疹病毒载体、痘病毒载体、疱疹性口炎病毒载体、逆转录病毒载体和慢病毒载体。

13. 根据权利要求9所述的用途,其中所述药物配制成通过经口、静脉内、粘膜内、动脉内、肌内或冠状动脉内途径施用。

14. 根据权利要求9或13所述的用途,其中所述药物配制成在再灌注疗法之前、与其同时或在其之后施用。

15. 根据权利要求14所述的用途,其中所述施用通过一次或更多次快速推注和/或输注进行。

## 用于治疗或预防疾病的因子1蛋白、因子2蛋白及其抑制剂

[0001] 本申请是申请号为201480005079.5的中国专利申请的分案申请,原申请是2014年1月16日提交的PCT国际申请PCT/EP2014/050788于2015年7月16日进入中国国家阶段的申请。

### 技术领域

[0002] 本发明涉及包含由来源于人染色体区C19orf10和/或C19orf63的核酸编码之氨基酸序列的蛋白质(称为因子1和/或因子2),其用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。还提供了因子1和因子2的抑制剂,其用于医学用途,特别是用于治疗或预防其中血管发生有助于疾病发生或进展的疾病。

### 背景技术

[0003] 急性心肌梗死(AMI)是全世界发病率和死亡率的一个主要原因。仅在德国,每年发生约280,000例。对患有AMI的患者的治疗包括通过再灌注疗法打开闭塞的冠状动脉,并且联合施用血小板聚集抑制剂和凝固抑制剂以防止血管再闭合。另外,可通过施用 $\beta$ 阻滞剂和ACE抑制剂降低脉搏和血压。还常见的是使用斯特汀类(statins)以降低胆固醇水平。用于直接修复心肌的医学方法目前只限于实验性地使用身体自身的骨髓细胞。对于具有类似效果但是不使用骨髓细胞的药物有巨大需求。

[0004] AMI期间持续的组织坏死引起伤口愈合应答,其导致坏死区域被肉芽组织替换并最终被富含胶原蛋白的疤痕替换。在AMI之后,单核细胞被从骨髓中募集到梗死心肌中,并且在伤口愈合中发挥重要作用。心肌中的单核细胞响应是短暂双相的。促炎性单核细胞出现较早并且促进梗死组织的消化和坏死碎片的移除,而修复性单核细胞在较晚占据优势并且传播血管发生和修复。趋化因子受体CXCR4的细胞表面表达鉴定小鼠和人中的修复性单核细胞亚群。CXCR4+骨髓细胞的促血管发生和促愈合效果被认为通过以旁分泌方式作用的分泌蛋白介导,但是对于这些因子的鉴定在很大程度上尚未可知。因此,本发明人在人CXCR4+骨髓细胞中进行了生物信息学分泌蛋白质组(secretome)分析以鉴定在AMI之后控制梗死愈合以及表现出治疗潜力的新的分泌蛋白。

[0005] 这些研究鉴定了两种表现出促血管发生和/或细胞保护作用的不同多肽,本发明人将其命名为因子1(Factor 1)和因子2(Factor 2)蛋白。

[0006] 这两种因子均在多种关于生物背景的科学出版物中描述,其不包括这些因子在非转化细胞或非转化组织中的促血管发生和/或细胞保护作用。没有研究公开证据或甚至暗示这些因子的功能,或者一种因子与仅和非转化细胞或非转化组织相关的疾病或病症之间的显著相关性。

[0007] 人因子1的氨基酸序列在人染色体19上的开放阅读框10(C19orf10)中被编码。在2007年对作为滑膜中的新分泌因子的所谓的成纤维细胞样滑膜细胞(FLS细胞)的蛋白质组分析中描述了所述蛋白质。猜测了所述蛋白质的分泌和关节的炎性疾病之间的相关性,但是没有任何试验或统计学证据(Weiler等,Arthritis Research and Therapy 2007,The

identification and characterization of a novel protein, c19orf10, in the synovium)。一个相应专利申请要求保护所述蛋白质作为用于治疗关节或用于诊断经历了改变的生长的组织以及监测组织中变化的治疗剂 (US 2008/0004232 A1, Characterization of c19orf10, a novel synovial protein)。另一个科学出版物描述了肝细胞癌细胞中所述蛋白质的增强的表达 (Sunagozaka等, International Journal of Cancer, 2010, Identification of a secretory protein c19orf10 activated in hepatocellular carcinoma)。重组产生的蛋白质表现出对培养的肝细胞癌细胞的增殖增强的效果。应注意的是, C19orf10也被称为IL-25、IL-27和IL-27W, 因为其最初被认为是白细胞介素。但是, 术语“IL-25”和“IL-27”已经在本领域中不一致地使用, 并且用于指多种不同的蛋白质。例如, US 2004/0185049指出一种蛋白质为IL-27, 并且公开了其在调控免疫应答中的用途。这种蛋白质在结构上不同于因子1 (根据SEQ ID NO:1的因子1氨基酸序列与根据UniProt:Q8NEV9的“IL-27”的氨基酸序列相比)。类似地, EP 2130547A1指出一种蛋白质为IL-25, 并且公开了其在治疗炎症中的用途。这种蛋白质在本领域中也称为IL-17E, 并且其在结构上不同于因子1 (根据SEQ ID NO:1的因子1氨基酸序列与根据UniProt:Q9H293的“IL-25”的氨基酸序列相比)。

[0008] 人因子2的氨基酸序列在人染色体19上的开放阅读框63 (C19orf63) 中被编码。所述蛋白质在2009年被作为新的分泌因子INM02描述 (Wang等, Journal of Endocrinology 2009, Molecular cloning of a novel secreted peptide, INM02, and regulation of its expression by glucose)。在具有多克隆抗体的人血清中示出了所述蛋白质的存在。此外, 示出了培养的MIN6 ( $\beta$ 细胞) 以及分离的胰腺大鼠胰岛中所述蛋白质的表达和培养基中的葡萄糖浓度之间的相关性。对于糖尿病和INM2的表达之间的相关性的分析没有提供显著性结果。一个相应专利申请要求保护抗所述蛋白质的多克隆抗体的产生及其用于治疗糖尿病的用途 (CN 200910055490, Novel polyclonal antibody of secretive peptide INM02 and preparation method thereof)。

[0009] 另一个科学出版物描述了作为新的分泌因子hHSS1 (人造血含信号肽分泌1) 的所述蛋白质 (Junes-Gill等, J Neurooncol, 2011, hHSS1: a novel secreted factor and suppressor of glioma growth located at chromosome 19q13.33)。公开的数据示出了造血系统中干细胞中hHSS1的表达, 并提出了作为特定脑肿瘤 (神经胶质瘤) 的产生中之肿瘤抑制剂的功能。一个相应专利申请要求保护hHSS1在治疗脑肿瘤中的用途 (W02011/094446A1, A method for treating brain cancer using a novel tumor suppressor gene and secreted factor)。

## 发明内容

[0010] 在第一个方面, 本发明提供了一种蛋白质, 所述蛋白质包含根据SEQ ID NO:1的氨基酸序列或者其片段或变体, 所述片段或变体与SEQ ID NO:1具有至少80%序列同一性, 所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0011] 在第二个方面, 本发明提供了一种蛋白质, 所述蛋白质包含根据SEQ ID NO:3的氨基酸序列或者其片段或变体, 所述片段或变体与SEQ ID NO:3具有至少80%序列同一性, 所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0012] 在第三方面,本发明提供了编码根据所述第一和第二方面的蛋白质的核酸,所述核酸用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0013] 在第四方面,本发明提供了包含根据所述第三方面的核酸的载体,所述载体用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0014] 在第五方面,本发明提供了药物组合物,所述药物组合物包含所述第一和/或第二方面的蛋白质和/或所述第三方面的核酸/或所述第四方面的载体并且任选地包含合适的药物赋形剂,所述药物组合物用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0015] 在第六方面,本发明提供了因子1和因子2分别的抑制剂,所述抑制剂用于医学用途,优选是治疗或预防其中血管发生有助于疾病发生或发展的疾病。

[0016] 在第七方面,本发明提供了编码这种抑制剂的核酸,所述核酸用于治疗或预防其中血管发生有助于疾病发生或发展的疾病。

[0017] 在第八方面,本发明提供了包含所述第七方面的编码这种抑制剂的核酸的载体,所述载体用于治疗或预防其中血管发生有助于疾病发生或发展的疾病。

[0018] 在第九方面,本发明提供了药物组合物,其包含所述第六方面的抑制剂和/或所述第七方面的核酸和/或所述第八方面的载体并且任选地包含合适的药物赋形剂,所述药物组合物用于治疗或预防其中血管发生有助于疾病发生或发展的疾病。以上发明内容未必描述了本发明的所有方面。

## 附图说明

[0019] 图1:如指出的,在不存在(对照)或存在10%FCS、人重组VEGF-A(R&D系统)或不同浓度的重组人因子1(根据SEQ ID NO:2的氨基酸序列)或因子2(根据SEQ ID NO:4的氨基酸序列)的情况下,将人冠状动脉内皮细胞(HCAEC)和人脐带静脉内皮细胞(HUVEC)在基本培养基中培养24小时。(A)通过溴脱氧尿苷渗入测量HCAEC增殖。(B)在用移液管尖端对汇合的内皮细胞单层制造创伤后,评估HCAEC迁移。(C)在培养在生长因子减少的基质胶(Matrigel)上的细胞中评估HUVEC网络形成。N=每种条件3-5个独立试验;\* $P<0.05$ ,\*\* $P<0.01$ ,\*\*\* $P<0.001$ 相对于对照。

[0020] 图2:通过Percoll密度梯度离心从1至3天龄的Sprague-Dawley大鼠中分离心室心肌细胞。如指出的,在不存在(对照)或存在重组人GDF-15或不同浓度的重组小鼠因子1(根据SEQ ID NO:13的氨基酸序列)的情况下,使心肌细胞暴露于模拟的局部缺血180分钟(5%CO<sub>2</sub>/95%N<sub>2</sub>气氛下的含2-脱氧葡萄糖的不含葡萄糖培养基),然后进行60分钟的模拟再灌注(返回5%CO<sub>2</sub>/95%室内空气下的含葡萄糖培养基)。通过原位TdT介导的dUTP缺口末端标记(TUNEL)评估细胞死亡。N=3个独立试验每种条件;\* $P<0.05$ 相对于对照。

[0021] 图3:将小鼠因子1或因子2cDNA(根据SEQ ID NO:7和10的核酸)克隆到复制缺陷型腺病毒中。使用编码半乳糖苷酶(lacZ)的复制缺陷型腺病毒作为对照。将10至12周龄的雄性C57BL/6小鼠用异氟烷麻醉和通气,并且进行永久的冠状动脉左前降支(LAD)结扎。在LAD结扎后立即将病毒注入到左心室(LV)腔中。(A)LAD结扎后28天通过经胸超声心动图(transthoracic echocardiography)评估左心室收缩功能(切面面积变化,FAC)。(B)LAD结扎后28天通过荧光显微术对梗死边界区中的同工凝集素阳性毛细血管密度进行定量。\* $P$

$<0.05$ ,  $**P<0.01$ 相对于Ad.lacZ对照。

[0022] 图4:将10至12周龄的雄性C57BL/6小鼠用异氟烷麻醉和通气,并且进行短暂的冠状动脉左前降支结扎一小时,然后再灌注28天。在再灌注时,小鼠接受重组小鼠因子1(根据SEQ ID NO:13的氨基酸序列)或因子2(根据SEQ ID NO:24的氨基酸序列)的单次皮下(s.c.)注射。这之后进行7天持续的重组因子1或因子2的皮下输注。对照小鼠输注PBS。(A)再灌注后28天通过经胸超声心动图评估左心室收缩功能(部分面积变化,FAC)。(B)再灌注后28天通过荧光显微术对梗死边界区中的同工凝集素阳性毛细血管密度进行定量。 $*P<0.05$ ,  $**P<0.01$ 相对于PBS对照。

[0023] 图5:将10至12周龄的雄性C57BL/6小鼠用异氟烷(1-2%)麻醉和通气,并且进行短暂的冠状动脉左前降支结扎一小时,然后再灌注28天。在再灌注时,小鼠接受重组小鼠因子1或因子2(分别为SEQ ID NO:13和24)的单次皮下注射(对照小鼠注射PBS)。这之后进行7天持续的重组因子1或因子2的皮下输注。对照小鼠输注PBS。28天中每天检查小鼠以评估梗死后存活。

[0024] 图6:用BLASTP算法检索人C19orf10编码的蛋白质的序列同源物。从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物、鸟类和鱼类中选择一个实例。使用CLUSTALW2算法对所选择的氨基酸序列进行比对。所有比对物种之间的同一性标记为“\*”,最高保守氨基酸位置(即仅示出保守替换的位置)标记为“:”,高度保守氨基酸位置标记为“.”。

[0025] 图7:用BLASTP算法检索人C19orf10编码的蛋白质的序列同源物。从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物、鸟类和鱼类中选择一个实例。使用CLUSTALW2算法对所选择的哺乳动物氨基酸序列进行比对。所有比对物种之间的同一性标记为“\*”,最高保守氨基酸位置(即仅示出保守替换的位置)标记为“:”,高度保守氨基酸位置标记为“.”。

[0026] 图8:用BLASTP算法检索人C19orf63剪接变体HSS1编码的蛋白质的序列同源物。从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物和鱼类中选择一个序列。使用CLUSTALW2算法对所选择的氨基酸序列进行比对。所有比对物种之间的同一性标记为“\*”,最高保守氨基酸位置(即仅示出保守替换的位置)标记为“:”,高度保守氨基酸位置标记为“.”。

[0027] 图9:用BLASTP算法检索人C19orf63剪接变体HSS1编码的蛋白质的序列同源物。从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物和鱼类物种中选择一个序列。使用CLUSTALW2算法对所选择的氨基酸序列进行比对。所有比对的物种之间的同一性标记为“\*”,最高保守氨基酸位置(即仅示出保守替换的位置)标记为“:”,高度保守氨基酸位置标记为“.”。

[0028] 图10:用BLASTP算法检索人C19orf63剪接变体HSM1编码的蛋白质的序列同源物。从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物和鱼类中选择一个序列。使用CLUSTALW2算法对所选择的氨基酸序列进行比对。所有比对的物种之间的同一性标记为“\*”,最高保守氨基酸位置(即仅示出保守替换的位置)标记为“:”,高度保守氨基酸位置标记为“.”。

[0029] 图11:用BLASTP算法检索人C19orf63剪接变体HSM1编码的蛋白质的序列同源物。

从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物和鱼类中选择一个序列。使用CLUSTALW2算法对所选择的哺乳动物氨基酸序列进行比对。所有比对的物种之间的同一性标记为“\*”,最高保守氨基酸位置(即仅示出保守替换的位置)标记为“:”,高度保守氨基酸位置标记为“.”。

[0030] 图12:示出了因子1和2特异性抗体对用重组因子1(图A)和重组因子2(图B)刺激的HCAEC的增殖的影响。数据为来自3-6次实验的平均值 $\pm$ SEM。图A: #P<0.05, ##P<0.01相对于未刺激的对照(左列) \*P<0.05, \*\*P<0.01相对于无抗体的因子1;图B: ##P<0.01相对于未刺激的对照(最左列) \*P<0.05, \*\*P<0.01相对于无抗体的因子2。

### 具体实施方式

[0031] 在下文详细描述本发明之前,应理解的是,本发明不限于本文描述的特定方法、方案和试剂,因为这些可以改变。还应理解的是,本文使用的术语仅仅是为了描述特定实施方案,而不是旨在限制本发明的范围,本发明的范围仅由所附权利要求限定。除非另有限定,否则本文使用的技术和科学术语具有与本领域普通技术人员的通常理解相同的含义。

[0032] 定义

[0033] 优选地,本文使用的术语如以下文献所述定义:“*A multilingual glossary of biotechnological terms: (IUPAC Recommendations)*”, H.G.W. Leuenberger, B. Nagel, 和 H. Kölbl, 编辑, Helvetica Chimica Acta, CH-4010 Basel, Switzerland, (1995)。

[0034] 为了实践本发明,除非另外指出,否则使用本领域文献中解释的化学、生物化学、细胞生物学和重组DNA技术的常规方法(参考,例如Molecular Cloning: A Laboratory Manual, 第2版, J. Sambrook等编辑, Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989)。此外,还使用本领域文献中解释的常规临床心脏病学方法(参考,例如Braunwald's Heart Disease. A Textbook of Cardiovascular Medicine, 第9版, P. Libby等编辑, Saunders Elsevier Philadelphia, 2011)。

[0035] 在整个说明书和所附权利要求书中,除非上下文另有要求,否则词语“包括”及类似用语如“包含”和“含有”将被理解为意指包括所述整体或步骤或者整体或步骤的组,但是不排除其他整体或步骤或者整体或步骤的组。如本说明书和所附权利要求中使用的,未用数量词限定的名词包括复数,除非内容清楚地另外规定。

[0036] 核酸分子应理解为由核苷酸单体形成的聚合物大分子。核苷酸单体由核碱基、五碳糖(例如但不限于核糖或2'-脱氧核糖)和一至三个磷酸基团构成。通常,多核苷酸通过单个核苷酸单体之间的磷酸二酯键形成。在本发明上下文中,提到的核酸分子包括但不限于核糖核酸(RNA)和脱氧核糖核酸(DNA)。术语“多核苷酸”和“核酸”在本文中可互换使用。

[0037] 术语“开放阅读框”(ORF)是指可翻译成氨基酸的核苷酸序列。通常,这样的ORF包含起始密码子、随后通常具有三核苷酸的倍数之长度的区域,但是在给定的阅读框中不包含终止密码子(TAG、TAA、TGA、UAG、UAA或UGA)。通常,ORF天然存在或被人工构建,即通过基因技术手段。ORF编码其中的氨基酸可翻译形成肽连接链的蛋白质。

[0038] 术语“蛋白质”和“多肽”在本文中可互换使用,是指任何氨基酸的肽键连接链,而不论长度或翻译后修饰。本发明中的可用蛋白质(包括蛋白质衍生物、蛋白质变体、蛋白质片段、蛋白质区段、蛋白质表位和蛋白质结构域)可进一步通过化学修饰进行修饰。这意味

着这种化学修饰的多肽包含除20种天然存在的氨基酸之外的化学基团。这样的其他化学基团的实例包括但不限于糖基化氨基酸和磷酸化氨基酸。多肽的化学修饰可提供与亲本多肽相比有利的性质,例如,一种或更多种增强的稳定性,延长的生物半衰期或增加的水溶性。适用于本发明中的可用变体的化学修饰包括但不限于:PEG化、非糖基化亲本多肽的糖基化、与治疗性小分子(如胰高血糖素样肽1激动剂包括艾塞那肽(exenatide)、阿必鲁肽(albiglutide)、他司鲁肽(taspoglutide)、DPP4抑制剂、肠降血糖素和利拉鲁肽(liraglutide))共价偶联,或者亲本多肽中存在的糖基化模式的改变。适用于本发明的可用变体的这些化学修饰可共翻译或在翻译后发生。

[0039] 术语“氨基酸”涵盖天然存在的氨基酸以及氨基酸衍生物。本发明上下文中的疏水非芳香族氨基酸优选具有大于0.5,更优选大于1.0,甚至更优选大于1.5的Kyte-Doolittle亲水指数并且不是芳香族的任何氨基酸。优选地,本发明上下文中的疏水非芳香族氨基酸选自氨基酸丙氨酸(Kyte Doolittle亲水指数1.8)、甲硫氨酸(Kyte Doolittle亲水指数1.9)、异亮氨酸(Kyte Doolittle亲水指数4.5)亮氨酸(Kyte Doolittle亲水指数3.8)和缬氨酸(Kyte Doolittle亲水指数4.2)或其具有上文限定之Kyte Doolittle亲水指数的变体。

[0040] 本文使用的术语“翻译后”是指发生在核苷酸三联体翻译成氨基酸并与序列中行进(proceeding)的氨基酸形成肽键之后的事件。这样的翻译后事件可在整个多肽形成之后发生或者在翻译过程期间在已经被翻译的多肽的那些部分上已经发生。翻译后事件通常改变或修改所得多肽的化学或结构性质。翻译后事件的实例包括但不限于这样的事件如氨基酸的糖基化或磷酸化,或者例如通过内肽酶对肽链的切割。

[0041] 本文使用的术语“共翻译(co-translational)”是指在核苷酸三联体翻译成氨基酸链的过程期间发生的事件。这些事件通常改变或修改所得氨基酸链的化学或结构性质。共翻译事件的实例包括但不限于可停止整个翻译过程或干扰肽键形成而导致两个分立(discreet)的翻译产物的事件。

[0042] 本文使用的术语“变体”是指与其来源的多肽或其片段相比差别在于氨基酸序列中的一个或更多个变化的多肽。蛋白质变体来源的多肽也称为亲本多肽。同样地,蛋白质片段变体来源的片段也称为亲本片段。通常,变体是人工构建的,优选地通过基因技术手段。通常,亲本多肽是野生型蛋白质或野生型蛋白质结构域。另外,本发明中的可用变体还可来源于亲本多肽的同源物、直系同源物或旁系同源物或来源于人工构建的变体,只要所述变体表现出亲本多肽的至少一种生物学活性即可。氨基酸序列中的变化可以是氨基酸交换、插入、缺失、N端截短或C端截短,或者这些变化的任意组合,其可发生在一个或数个位点处。在一些优选实施方案中,本发明中的可用变体表现出在氨基酸序列中总数至多100(至多1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95或100)个变化(即,交换、插入、缺失、N端截短和/或C端截短)。氨基酸交换可以是保守的和域半保守的和/或非保守的。在一些优选实施方案中,本发明中的可用变体与其来源的蛋白质或结构域相差至多1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45或50个氨基酸交换,优选保守氨基酸变化。

[0043] 通常替换在脂肪族氨基酸之间,具有脂肪族羟基侧链的氨基酸之间,具有酸性残基的氨基酸之间,酰胺衍生物之间、具有碱性残基的氨基酸之间或具有芳香族残基的氨基



酸之间。通常半保守和保守替换为：

	氨基酸	保守替换	半保守
	A	G; S; T	N; V; C
	C	A; V; L	M; I; F; G
	D	E; N; Q	A; S; T; K; R; H
	E	D; Q; N	A; S; T; K; R; H
[0044]	F	W; Y; L; M; H	I; V; A
	G	A	S; N; T; D; E; N; Q
	H	Y; F; K; R	L; M; A
	I	V; L; M; A	F; Y; W; G
	K	R; H	D; E; N; Q; S; T; A
	L	M; I; V; A	F; Y; W; H; C
	M	L; I; V; A	F; Y; W; C;
	N	Q	D; E; S; T; A; G; K; R
	P	V; I	L; A; M; W; Y; S; T; C; F
	Q	N	D; E; A; S; T; L; M; K; R
[0045]	R	K; H	N; Q; S; T; D; E; A
	S	A; T; G; N	D; E; R; K
	T	A; S; G; N; V	D; E; R; K; I
	V	A; L; I	M; T; C; N
	W	F; Y; H	L; M; I; V; C
	Y	F; W; H	L; M; I; V; C

[0046] 如果新半胱氨酸保留游离巯基,从A、F、H、I、L、M、P、V、W或Y到C的变化是半保守的。此外,技术人员将理解,空间需求位置的甘氨酸不应被替换,并且P不应被引入到具有 $\alpha$ 螺旋或 $\beta$ 片层结构的蛋白质部分中。

[0047] 作为替代或补充,本文使用的“变体”可通过与其来源的亲本多肽或亲本多核苷酸的一定程度的序列同一性来表征。更确切地说,本发明上下文中的蛋白质变体表现出与其亲本多肽至少80%的序列同一性。优选地,所讨论的多肽和参照多肽在20、30、40、45、50、60、70、80、90、100或更多个氨基酸的连续段上或在参照多肽的整个长度上表现出指定的序列同一性。优选地,所讨论多核苷酸和参照多核苷酸在60、90、120、135、150、180、210、240、270、300或更多个核苷酸的连续段上或在参照多肽的整个长度上表现出指定的序列同一性。

[0048] 整个说明书关于多肽和多核苷酸序列比较使用术语“至少80%序列同一性”。这种表述优选地是指与相应参照多肽或相应参照多核苷酸具有至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%或至少99%序列同一性。

[0049] 蛋白质片段包含氨基酸的缺失,其可以是N端截短、C端截短或内部缺失,或者这些的任意组合。包含N端截短、C端截短和/或内部缺失的这些变体在本申请上下文中被称为“片段”。片段可以是天然存在的(例如,剪接变体)或可以是人工构建的,优选地通过基因技术手段。优选地,与亲本多肽相比,片段(或缺失变体)在其N端和/或其C端和/或内部具有至

多1、2、3、4、5、6、7、8、9、10、15、20、25、30、35、40、45或50个氨基酸的缺失,优选地在其N端、在其N和C端或在其C端。

[0050] 如果没有明确地另外指出,在两个序列相比较并且没有指明计算序列同一性百分比所比较的参照序列的情况下,参考两个待比较序列中较长的计算序列同一性。

[0051] 可通过序列比对来确定核苷酸和氨基酸序列的相似性,即,序列同一性百分比。这样的比对可利用多种本领域已知算法进行,优选Karlin和Altschul的数学算法(Karlin&Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5877)、hmmalign (HMMER包, <http://hmmmer.wustl.edu/>) 或者CLUSTAL算法(Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673-80) 或CLUSTALW2算法(Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. (2007) . Clustal W and Clustal X 2.0版. Bioinformatics, 23, 2947-2948), 其可获自例如[http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_clustalw.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html)或<http://www.ebi.ac.uk/Tools/clustalw2/index.html>。优选地,使用<http://www.ebi.ac.uk/Tools/clustalw2/index.html>的CLUSTALW2算法,其中所使用的参数是<http://www.ebi.ac.uk/Tools/clustalw2/index.html>上设置的默认参数:比对类型=慢,对于慢配对比对选择,蛋白质权重矩阵=Gonnet,空位开放=10,空位延伸=0.1,以及蛋白质权重矩阵=Gonnet,空位开放=10,空位延伸=0.20,空位距离=5,无结束空位=无,输出选项:格式=Aln w/数字,顺序=对齐。

[0052] 可使用例如BLAST、BLAT或BlastZ (或BlastX) 计算序列同一性 (序列匹配) 级别。类似算法被整合在了BLASTN和BLASTP程序中:Altschul等(1990) J. Mol. Biol. 215:403-410。可利用BLASTP程序进行BLAST蛋白搜索,所述程序可获自例如[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST\\_PROGRAMS=blastp&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)。优选地,使用的算法参数是[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST\\_PROGRAMS=blastp&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)上设置的默认参数:预期阈值=10,字长=3,查询范围中的最大匹配=0,矩阵=BLOSUM62,空位罚分=存在:11延伸:1,组成调整=条件组成得分矩阵调节以及非冗余蛋白序列(nr)的数据库,以获得与因子1和因子2多肽同源的氨基酸序列。

[0053] 为了获得比较目的的空位比对,使用如Altschul等(1997) Nucleic Acids Res. 25:3389-3402中所述的空位BLAST。当使用BLAST和空位BLAST程序时,使用各程序的默认参数。可通过建立的同源性作图技术如Shuffle-LAGAN (Brudno M., Bioinformatics 2003b, 19 Suppl 1: I54-I62) 或马尔科夫随机域(Markov random fields)对序列匹配分析进行补充。当在本申请中提到序列同一性百分比时,这些百分比相对较长序列的全长计算,如果没有另外明确地说明的话。

[0054] 本文使用的术语“宿主细胞”是指具有本发明之核酸的细胞(例如,质粒或病毒)。这种宿主细胞可以是原核细胞(例如,细菌细胞)或真核细胞(例如,真菌、植物或动物细胞)。细胞可以是转化的或非转化的。细胞可以是例如细胞培养物中的分离的细胞或组织的一部分,其本身可以是分离的更复杂的组织结构(如器官或个体)或其一部分。

[0055] 术语“因子1”、“因子1蛋白”或“因子1多肽”可互换使用,并且是指NCBI参照序列NM\_019107.3(人同源物)指示的蛋白质以及其哺乳动物同源物,特别是来自小鼠或大鼠。人同源物的氨基酸序列在人染色体19上的开放阅读框10(C19orf10)中被编码。优选地,因子1蛋白是指包含具有根据SEQ ID NO:1的氨基酸序列的人因子1核心区段或基本由其组成或由其组成的蛋白质。在一个更优选的实施方案中,因子1蛋白具有根据SEQ ID NO:2的氨基酸序列。

[0056] 术语“因子2”、“因子2蛋白”或“因子2多肽”可互换使用,并且是指NCBI参照序列NM\_175063.4(人同源物)指示的蛋白质以及其哺乳动物同源物,特别是来自小鼠或大鼠。人因子2的氨基酸序列在人染色体19上的开放阅读框63(C19orf63)中被编码。优选地,因子2蛋白是指包含具有根据SEQ ID NO:3的氨基酸序列的人因子2核心区段或基本由其组成或由其组成的蛋白质。在一个更优选的实施方案中,因子2蛋白具有分别根据SEQ ID NO:4和5的氨基酸序列。在最优选的实施方案中,因子2蛋白是分泌形式,优选地具有根据SEQ ID NO:4的氨基酸序列。

[0057] 术语“非转化组织”或“非转化细胞”是指表现出可比较的非癌或肿瘤发生细胞或组织之生理参数的组织和细胞。这种参数是例如但不限于细胞周期控制、细胞分裂速率、接触抑制、非锚定依赖性生长或代谢。可比较的非癌或肿瘤发生细胞或组织可以是健康的、受损的或患病的。

[0058] 术语“愈合”包括活细胞、组织、器官和生物系统整体的再生和修复,以及正常功能的部分或完全恢复。在组织、器官或生物系统整体的情况下,其包括通过体内细胞再生和修复以减小受损或坏死区域的尺寸或使其被新的活组织替代的过程。替代可例如通过用形成与原组织类似的组织的新细胞替代坏死细胞的再生或通过其中用疤痕组织替代损伤组织的修复来发生。因为再生是导致正常功能部分或完全恢复的过程,其为愈合过程中的优选变量。因此,本发明上下文中的术语愈合包括本领域技术人员可与该术语组合的全部过程,但是优选地应促进再生过程,而不是导致产生非功能性组织(如疤痕组织)的过程。本发明上下文中的术语愈合还优选地指促进增殖、迁移、网络形成和血管发生。

[0059] 术语“增强增殖”是指如果与不用本发明的蛋白质、核酸、载体或药物组合物处理的细胞或细胞组相比,细胞或细胞组的细胞分裂速率提高。本领域中公知如何测量细胞的细胞分裂速率,例如通过使用FACS对有丝分裂细胞进行计数。

[0060] 术语“抑制凋亡”是指在对照细胞或细胞组进入凋亡的情况下,本发明的蛋白质、核酸、载体或药物组合物阻止细胞或细胞组进入凋亡的能力。本领域技术人员公知如何测量细胞是否经历凋亡,例如通过TUNEL测定。

[0061] 实施方案的描述包括对贯穿本申请使用的术语的进一步限定和解释。这些描述和限定对于整个申请有效,除非另有说明。

[0062] 实施方案

[0063] 在下文中,将描述本发明的要素。这些要素与具体的实施方案一起列出,但是,应当理解,它们可以以任何方式以及任何数量进行组合以产生另外的实施方案。多种描述的实施例和优选的实施方案不应被解释为将本发明限制于仅明确地描述的实施方案。这种描述应被理解为支持和涵盖将明确描述的实施方案与任何数量的公开和/或优选的要素相组合的实施方案。此外,除非上下文另有说明,否则本申请中所有描述的要素的任何排列和组

合应被视为被本申请的说明书所公开。

[0064] 在第一方面,本发明提供了蛋白质,所述蛋白质包含因子1蛋白、基本上由因子1蛋白组成或由因子1蛋白组成,优选地具有根据SEQ ID NO:1的氨基酸序列或者其片段或变体,所述片段或变体与SEQ ID NO:1具有至少80%序列同一性,所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡,特别地增强非转化组织的增殖、增强非转化细胞的增殖、抑制非转化组织的凋亡或抑制非转化细胞的凋亡。在一个特别优选的本发明的实施方案中,所述蛋白质包含氨基酸序列SEQ ID NO:1或其片段。优选地,所述蛋白质与SEQ ID NO:1具有至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的同一性。

[0065] 在本发明这个方面的一个优选实施方案中,所述蛋白质包含氨基酸序列SEQ ID NO:2,与SEQ ID NO:2具有至少80%序列同一性的其片段或变体。优选的SEQ ID NO:2片段缺乏N-端信号序列MAAPSGGWNGVGASLWAALLLGAVALRPAEA (SEQ ID NO:35)。本领域技术人员能够在没有不适当的负担下决定亲本多肽的哪些位置可以突变至何种程度以及哪些位置必须保持以保留所述多肽的功能。例如,这样的信息可以从通过本领域公知的生物信息学的方法鉴定、比对和分析的同源物序列获得。在实施例7中示例性地描述了这种分析,且将结果示于图6和7中。突变优选引入到物种(优选哺乳动物)之间不充分保守的蛋白质的那些区域中,即,被突变的一个或更多个那些氨基酸位置未标有“\*”。在一个更优选的实施方案中,仅改变不完全保守的(用“\*”表示)或保守程度较低(用“:”或“.”表示)的氨基酸。在一个特别优选的本发明的实施方案中,因子1蛋白包含氨基酸序列SEQ ID NO:2或其片段,基本上由氨基酸序列SEQ ID NO:2或其片段组成或者由氨基酸序列SEQ ID NO:2或其片段组成。优选地,所述蛋白质与SEQ ID NO:2具有至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的序列同一性。

[0066] 这些突变可存在于根据SEQ ID NO:2的全长蛋白质中或存在于根据SEQ ID NO:1缺乏N-端信号序列的蛋白质中。

[0067] 除了N-端信号,N-端缺失变体可能缺少一个或更多个来自32至55位氨基酸(根据SEQ ID NO:2),即来自N端保守区的氨基酸。因此,缺失的因子1蛋白的N端可以在位置32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55或56处,另外地或可替代地,缺失的因子1蛋白可能缺少一个或更多个来自146至173位氨基酸(基于SEQ ID NO:2),即来自C端保守区的氨基酸。因此,缺失的因子1蛋白的C端可以在位置145、146、147、148、149、150、151、152、153、154、155、156、157、158、159、160、161、162、163、164、165、166、167、168、169、170、171或172处。

[0068] 本发明第一方面的蛋白质还可包含例如用于稳定或纯化所得蛋白质的另外的氨基酸序列。这样的氨基酸的实例是His<sub>6</sub>-标签(SEQ ID NO:36)、myc-标签或FLAG-标签。

[0069] 在一些实施方案中,优选在本发明第一方面的蛋白质内突变蛋白酶切割位点来稳定蛋白质(参见Segers等,Circulation 2007,2011)。技术人员知晓如何确定蛋白质内潜在的蛋白水解切割位点。例如,可将蛋白质序列提交到提供这种分析的网站上,例如[http://web.expasy.org/peptide\\_cutter/](http://web.expasy.org/peptide_cutter/)或<http://pmap.burnham.org/proteases>。如果根据SEQ ID NO:2的蛋白质序列被提交到[http://web.expasy.org/peptide\\_cutter/](http://web.expasy.org/peptide_cutter/)上,则确定了以下较低频率(小于10)的切割位点:

[0070] 表1

	蛋白酶	频率	位置 (参照 SEQ ID NO: 2)
	<b>Arg-C 蛋白酶</b>	<b>7</b>	<b>27 43 96 113 130 151 170</b>
	<b>Asp-N 内肽酶</b>	<b>4</b>	<b>40 58 85 132</b>
[0071]	<b>梭菌蛋白酶</b>	<b>7</b>	<b>27 43 96 113 130 151 170</b>
	<b>LysN</b>	<b>9</b>	<b>59 99 108 124 136 144 155 160 166</b>
	<b>脯氨酸-内肽酶</b>	<b>4</b>	<b>28 44 97 152</b>

[0072] 可改变这些位点以去除各自鉴定的蛋白酶的识别/切割序列从而增加蛋白质的血清半衰期。

[0073] 因子1和因子2显示了增殖增强活性,特别是增强血管发生。在一个优选的实施方案中,根据第一方面的蛋白质用于增强非转化组织或非转化细胞的增殖,优选血管发生。因此,优选的是因子1或因子2用于治疗可从增强血管发生中受益的疾病。下面进一步举例说明这些疾病的一些实例。

[0074] 本发明的上下文中增殖的增强包括与未施用本发明蛋白质的对照细胞或对照组织相比,细胞或组织每级的增殖增强。例如,可通过如实施例2中所描述的掺入溴脱氧尿苷来测量细胞的增殖。例如,可通过测量各组织重量或尺寸的增加以及用组织学方法来确定组织的增强增殖。这样的方法在本领域中是公知的,因为它们中许多是用于临床应用的标准方法。

[0075] 在另一个优选的实施方案中,根据第一方面的蛋白质用于使非转化组织或非转化细胞愈合。在另一个优选的实施方案中,根据第一方面的蛋白质用于使非转化组织或非转化细胞愈合并且增强非转化组织或非转化细胞的增殖。在一个特别优选的实施方案中,根据第一方面的蛋白质用于使非转化组织或非转化细胞愈合并且增强非转化组织或非转化细胞的增殖并且抑制非转化组织或非转化细胞的凋亡。

[0076] 在另一个优选的实施方案中,根据第一方面的蛋白质用于抑制非转化组织或非转化细胞的凋亡。因此,本发明的蛋白质表现出抗凋亡的潜能并保护细胞或组织免于凋亡细胞死亡。在此上下文中“保护”或“细胞保护作用”意指,与对照相比,在用根据本发明的因子1蛋白处理的细胞中凋亡细胞死亡的程度降低至少20%,优选至少30%,更优选至少40%,甚至更优选至少50%,并且最优选至少60%。技术人员能够通过如实施例3中所描述的原位TdT介导的dUTP缺口末端标记(TUNEL)来评估细胞死亡。其他凋亡的指标是,例如可通过DNA梯度检查的片段化的基因组(Liu等,2005,Circulation 111:90-96)、细胞色素-c的释放或胱天蛋白酶-3的活性(Most等,2003,J.Biol.Chem.278:48404-48412)。可体内在实验的心力衰竭的动物模型中来评估肽的抗凋亡作用。例如,可评估蛋白质处理的局部缺血后收缩功能障碍的小鼠和经处理过的心脏组织以及对照小鼠之心肌细胞凋亡的程度。肽优选可经肠胃外,例如腹膜内、静脉内或皮下施用。在一个特别优选的实施方案中,本发明的蛋白质显示具有两种上述功能且优选具有全部上述功能,即,增强非转化组织或非转化细胞的增殖以及使非转化组织或非转化细胞愈合和抑制非转化组织或非转化细胞的凋亡。包括在本发明中的因子1蛋白的片段和变体表现出抗凋亡的潜能并保护细胞或组织免于凋亡细胞死亡,即,与根据SEQ ID NO:1或SEQ ID NO:2,更优选SEQ ID NO:1的氨基酸序列具有至少50%,优选60%,优选70%,优选80%,优选90%且更优选至少100%的蛋白质。

[0077] 在第二方面,本发明提供了蛋白质,所述蛋白质包含因子2蛋白、基本上由因子2蛋白组成或由因子2蛋白组成,具有根据SEQ ID NO:3的氨基酸序列或者其片段或变体,所述片段或变体与SEQ ID NO:3具有至少80%序列同一性,所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈合。在本发明的一个优选的实施方案中,因子2蛋白包含氨基酸序列SEQ ID NO:3,与SEQ ID NO:3具有至少80%同一性的其片段或变体。本领域技术人员能够在没有不适当的负担下决定多肽的哪些位置可以突变为何种程度以及哪些位置必须保持以保留所述多肽的功能。例如,这样的信息可以例如从通过本领域公知的生物信息学方法鉴定、比对和分析的同源物序列获得。在一个特别优选的本发明的实施方案中,本发明的第二方面的蛋白质包含氨基酸序列SEQ ID NO:3或其片段。优选地,所述蛋白质与SEQ ID NO:3具有至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的序列同一性。

[0078] 本发明第二方面的蛋白质还可包含例如用于稳定或纯化所得蛋白质的另外的氨基酸序列。

[0079] 在一些实施方案中,优选突变在本发明第一方面的蛋白质内的蛋白酶切割位点以稳定蛋白质。可如以上所述鉴定合适的蛋白水解切割位点。

[0080] 在本发明的另一个优选实施方案中,本发明的第二方面的蛋白质包含氨基酸序列SEQ ID NO:4或其片段或变体、基本上由氨基酸序列SEQ ID NO:4或其片段或变体组成或者由氨基酸序列SEQ ID NO:4或其片段或变体组成,所述片段或变体与SEQ ID NO:4具有至少80%的同一性。优选的片段缺乏N-端信号序列MAAASAGATRLLLLLLMAVAAPSRARG' (SEQ ID NO:37)。本领域技术人员能够在没有不适当的负担下决定多肽的哪些位置可以突变为何种程度以及哪些位置必须保持以保留所述多肽的功能。例如,这样的信息可以从通过本领域公知的生物信息学方法鉴定、比对和分析的同源物序列获得。在实施例8中示例性地描述这种分析,且将结果示于图8和9中。突变优选引入到物种(优选哺乳动物)之间不充分保守的蛋白质的那些区域中,即,被突变的一个或更多个那些氨基酸位置未标有“\*”。在一个更优选的实施方案中,仅改变不完全保守的(用“\*”表示)或保守程度较低(用“:”或“.”表示)的氨基酸。在一个本发明特别优选的实施方案中,因子2蛋白包含氨基酸序列SEQ ID:4或其片段。优选地,所述蛋白质与SEQ ID NO:4具有至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的同一性。

[0081] 这样的突变可存在于根据SEQ ID NO:4的全长蛋白质中或存在于缺乏N端信号序列的蛋白质中。

[0082] 除了N-端信号,N-端缺失变体可能缺少一个或更多个来自27至73位氨基酸(根据SEQ ID NO:4),即来自N端保守区的氨基酸。因此,缺失的因子2蛋白的N端可以在位置27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55或56处。另外地或可替代地,缺失的因子1蛋白可能缺少一个或更多个来自190至254位氨基酸(根据SEQ ID NO:4)即来自C端保守区的氨基酸。因此,缺失的因子2蛋白的C端可以在位置189、190、191、192、193、194、195、196、197、198、199、200、201、202、203、204、205、206、207、208、209、210、211、212、213、214、215、216、217、218、219、220、221、222、223、224、225、226、227、228、229、230、231、232、233、234、235、236、237、238、239、240、241、242、243、244、245、246、247、248、249、250、251、252或253处。

[0083] 在本发明的另一个优选实施方案中,本发明的第二方面的蛋白质包含SEQ ID NO:5的氨基酸序列或其片段或变体、基本上由SEQ ID NO:5的氨基酸序列或其片段或变体组成或者由SEQ ID NO:5的氨基酸序列或其片段或变体组成,所述片段或变体与SEQ ID NO:5具有至少80%的同一性。优选的片段缺乏N-端信号序列MAAASAGATRLLLLLLMAVAAPSRARG (SEQ ID NO:37)。本领域技术人员能够在没有不适当的负担下决定多肽的哪些位置可以突变为何种程度以及哪些位置必须保持以保留所述多肽的功能。例如,这样的信息可以从通过本领域公知的生物信息学的方法鉴定、比对和分析的同源物序列获得。实施例9中示例性地描述了这种分析,且将结果示于图10和11中。突变优选引入到物种(优选哺乳动物)之间不充分保守的蛋白质的那些区域,即,被突变的一个或更多个那些氨基酸位置未标有“\*”。在一个更优选的实施方案中,仅改变不完全保守的(用“\*”表示)或保守程度较低(用“:”或“.”表示)的氨基酸。在一个本发明的特别优选的实施方案中,因子2蛋白包含氨基酸序列SEQ ID:5或其片段。优选地,所述蛋白质与SEQ ID NO:5具有至少85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%的同一性。

[0084] 这样的突变可能存在于根据SEQ ID NO:5的全长蛋白质中或存在于根据SEQ ID NO:5缺乏N端信号序列的蛋白质中。

[0085] 除了N-端信号,N-端缺失变体可能缺少一个或更多个来自27至73位氨基酸(根据SEQ ID NO:4)即来自N端保守区的氨基酸。因此,缺失的因子2蛋白的N端可以在位置27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55或56处。另外地或可替代地,缺失的因子1蛋白可能缺少一个或更多个来自190至262位氨基酸(根据SEQ ID NO:4)即来自C端保守区的氨基酸。因此,缺失的因子2蛋白的C端可以在位置189、190、191、192、193、194、195、196、197、198、199、200、201、202、203、204、205、206、207、208、209、210、211、212、213、214、215、216、217、218、219、220、221、222、223、224、225、226、227、228、229、230、231、232、233、234、235、236、237、238、239、240、241、242、243、244、245、246、247、248、249、250、251、252、253、254、255、256、257、258、259、260或261处。

[0086] 本发明第二方面的蛋白质的片段和变体表现出抗凋亡的潜能并保护细胞或组织免于凋亡细胞死亡。其与根据SEQ ID NO:3,SEQ ID NO:4或SEQ ID NO:5,最优选SEQ ID NO:3的氨基酸序列具有至少50%,优选60%,优选70%,优选80%,优选90%,且更优选至少100%的蛋白质。

[0087] 在一个优选的实施方案中,根据第二方面的蛋白质用于增强非转化组织或非转化细胞的增殖。与未施用蛋白质的对照细胞或组织相比,增殖的增强包括每级的增强。上文描述了用于测量增殖的实验方法。在实施例2中也描述了通过掺入溴脱氧尿苷来评估根据本发明第二方面的蛋白质之各自的增殖测量值。

[0088] 在另一个优选的实施方案中,根据第二方面的蛋白质用于使非转化组织或非转化细胞愈合。在一个特别优选的实施方案中,根据本发明第二方面的蛋白质表现出上述两种功能,即,非转化组织或非转化细胞的增强增殖和愈合。

[0089] 在本发明的另一个实施方案中,在体内、离体或体外,优选体内施用第一方面和/或第二方面的蛋白质。一个示例性实施方案,当离体或体外施用本发明第一方面的蛋白质和/或本发明第二方面的蛋白质时,是为了组织工程的目的增强增殖和/或愈合和/或抑制细胞凋亡,其中产生组织以用于移植到个体中。用于组织工程的细胞可以来自相同物种的

相同个体或来自其他物种,但也可来自相同物种的不同个体或来自其他物种。本发明不包括除去所述细胞或组织以及将新的组织移植到个体中的方法。

[0090] 在本发明的一个优选的实施方案中,非转化细胞是干细胞。这样的干细胞可以是胚胎干细胞或成体干细胞以及祖细胞。本发明包括全能干细胞以及多能干细胞。

[0091] 在一个优选的实施方案中,非转化细胞或非转化组织是患病的。在另一个优选的实施方案中,非转化细胞或非转化组织是受损的。在另一个优选的实施方案中,非转化细胞或非转化组织是受损且患病的。

[0092] 在一个优选的实施方案中,非转化细胞或非转化组织是肌肉细胞或肌肉组织。肌肉包括由本领域的技术人员已知的所有类型的肌肉。这样的肌肉是例如,骨骼肌、平滑肌或心肌。在一个更具体的优选的实施方案中,肌肉是心肌。在另一个优选的实施方案中,非转化细胞或非转化组织是上皮细胞或上皮组织。在另一个优选的实施方案中,非转化细胞或非转化组织是神经细胞或神经组织。

[0093] 在一个优选的实施方案中,非转化细胞或非转化组织属于个体的循环系统。

[0094] 在另一个优选的实施方案中,非转化细胞或所述非转化组织属于或来自个体身体的限定系统,所述系统选自:消化系统、内分泌系统、排泄系统、免疫系统、皮肤(integumentary)系统、肌肉系统、神经系统、生殖系统、呼吸系统或骨骼系统。在本发明的另一个优选的实施方案中,所述细胞属于所列出个体系统中的两个、三个、四个、五个、六个、七个、八个、九个、十个或全部。

[0095] 在另一个优选的实施方案中,非转化细胞或非转化组织属于或来自个体身体的限定部分或器官,所述限定部分或器官选自:皮肤、骨、心脏、软骨、血管、食道、胃、肠、腺、肝、肾、肺、脑和脾。在一个特别优选的实施方案中,非转化细胞或非转化组织属于或来自心脏。

[0096] 更加优选地,在受损的非转化细胞或非转化组织的情况下,受损通过遗传性/继承性(inherited)疾病或通过例如局部缺血、再灌注损伤、炎症、感染、外伤、机械性超负荷(mechanical overload)、中毒或手术导致的获得性疾病而引起。在一个特别优选的实施方案中,所述受损是通过局部缺血引起的。在另一个特别优选的实施方案中,所述受损是通过再灌注损伤引起的。

[0097] 在本发明的上下文中,优选的是在患病的或受损的细胞或组织的情况下,其中所述受损是通过与萎缩、发育不全、炎症、损伤或创伤有关的疾病引起的。特别优选的是与损伤相关有关的疾病。还特别优选的是与创伤相关有关的疾病。

[0098] 在本发明的一个优选的实施方案中,所述疾病是骨骼肌病症,所述骨骼肌病症选自:肌营养不良、肌无力、肌肉萎缩、肌炎、中央轴空病、线状体(杆)肌病、中央核肌病(centronuclear myopathy)、肌管性肌病、中央核肌管性肌病、眼的眼肌麻痹以及线粒体肌病。肌营养不良可选自:贝克肌营养不良、先天性肌营养不良、杜氏肌营养不良、远端肌营养不良、艾-德肌营养不良(Emery-Dreifuss muscular dystrophy)、面肩肱型(facioscapulohumeral)肌营养不良、肢带型肌营养不良、强直性肌营养不良和眼咽型肌营养不良。肌炎可选自骨化性肌炎、纤维肌炎、特发性炎性肌病(例如皮肌炎、多肌炎以及包涵体肌炎)和脓性肌炎。

[0099] 在本发明的另一个优选的实施方案中,疾病是原发性或获得性心肌病。原发性心肌病是选自继承性心肌病和由自发突变引起的心肌病。心肌病为但不限制于:例如,肥厚性



心肌病 (HCM或HOCM)、致心律失常性右心室心肌病 (ARVC)、孤立心室非致密线粒体肌病 (isolated ventricular non-compaction mitochondrial myopathy)、扩张性心肌病 (DCM)、限制性心肌病 (RCM)、Takotsubo心肌病、吕弗勒氏心内膜炎 (Loeffler endocarditis)、糖尿病性心肌病、酒精性心肌病、肥胖相关的心肌病。

[0100] 在本发明的上下文中,获得性心肌病优选地选自动脉粥样硬化或其他冠状动脉疾病引起的缺血性心肌病,由心肌感染或中毒引起的心肌病,由肺动脉高血压和/或动脉高血压引起的高血压心脏疾病以及心脏瓣膜的疾病,其中特别优选由动脉粥样硬化或其他冠状动脉疾病引起的缺血性心肌病。

[0101] 在本发明的一个最优选的实施方案中,通过局部缺血或再灌注损伤而受损的非转化细胞或非转化组织属于心脏。因此,优选的是,所得到的待治疗的疾病选自心肌梗死、心绞痛和心力衰竭,其中特别优选心肌梗死。如在本发明的上下文中使用的术语“心肌梗死”包括急性心肌梗死 (AMI)。

[0102] 在一个特别优选的实施方案中,根据本发明的第一方面的蛋白质或根据本发明第二方面的蛋白质的用途包括在心肌梗死后向个体施加并且愈合包括改善左心室收缩功能且可能与梗死边界区毛细血管密度的增加相关。此外,根据本发明的第一方面的蛋白质或根据本发明的第二方面的蛋白可降低心肌梗死后的死亡率。可用于确定如左心室收缩功能的改善,梗死边界区中毛细血管密度的增加以及心肌梗死后死亡率的降低之参数的方法是本领域公知的并在实施例4-6中示例性地进行了描述。

[0103] 上述的因子1和因子2蛋白、片段或变体用于治疗或改善萎缩;发育不全;炎症、损伤;创伤、局部缺血;再灌注损伤;炎症;感染;外伤;机械性超负荷;中毒;原发性或获得性心肌病,优选继承性心肌病和由自发突变引起的心肌病。心肌病例如但不限于:肥厚性心肌病 (HCM或HOCM)、致心律失常性右心室心肌病 (ARVC)、孤立心室非致密线粒体肌病、扩张性心肌病 (DCM)、限制性心肌病 (RCM)、Takotsubo心肌病、吕弗勒氏心内膜炎、糖尿病性心肌病、酒精性心肌病、肥胖相关的心肌病;心肌梗死或改善左心室收缩功能。

[0104] 上述的因子1和因子2蛋白、片段或变体也可在治疗分别所示的病症和疾病的方法中使用。

[0105] 在第三方面,本发明提供了编码根据第一方面和第二方面的蛋白质的核酸,其用于增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈合和/或抑制非转化组织或非转化细胞的凋亡。

[0106] 术语“增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈合和/或抑制非转化组织或非转化细胞的凋亡”具有上述定义的含义,以及优选的含义。

[0107] 可优化核酸序列来努力增强在宿主细胞中的表达。要考虑的参数包括:C:G含量、优选的密码子和避免抑制性二级结构。可以以不同的方式组合这些因素来试图获得在特定宿主中具有增强表达的核酸序列 (参见,例如Donnelly等,国际公开号W097/47358)。特定序列在特定宿主中具有增强表达的能力涉及到一些经验性实验。这种实验包括测量预期的核酸序列的表达以及 (如果需要的话) 改变所述序列。从特定的氨基酸序列和遗传密码已知的简并性开始,可以得到大量不同的编码核酸序列。遗传密码的简并性的产生是因为几乎所有的氨基酸都是由不同组合的核苷酸三联体或“密码子”进行编码的。特定的密码子翻译成特定的氨基酸是本领域公知的 (参见,例如,Lewin GENES IV,第119页,Oxford University

Press, 1990)。

[0108] 在本发明的一个优选实施方案中,所述核酸还包括定位成控制蛋白质表达的转录调控元件或表达控制序列。这样的核酸与控制元件经常一起被称为表达系统。如本文所用的术语“表达系统”指的是设计以产生一个或更多个目的基因产物的系统。通常,这样的系统是“人工”设计的,即通过在体内、体外或离体可用于产生目的基因产物的基因技术手段。术语“表达系统”还涵盖目的基因产物的表达,所述表达包括多核苷酸的转录、mRNA剪接、翻译成多肽、多肽或蛋白质的共翻译修饰和翻译后修饰以及使蛋白质靶向至细胞内的一个或更多个区室、从细胞分泌以及在同一个细胞或另一个细胞中蛋白质的摄取。这种一般的描述指用于真核细胞、组织或生物体中的表达系统。用于原核系统的表达系统可能会不同,其中本领域公知如何构建原核细胞的表达系统。

[0109] 存在于基因表达盒中的调控元件一般包括:(a)与编码多肽的核苷酸序列转录性偶联的启动子,(b)与所述核苷酸序列功能性偶联的5'核糖体结合位点(c)与所述核苷酸序列的3'末端连接的终止子;和(d)与所述核苷酸序列功能偶联的3'聚腺苷酸化信号。还可存在用于增强或调控基因表达或多肽加工的另外的调控元件。启动子是由RNA聚合酶识别并介导下游区域转录的遗传元件。优选的启动子是提供提高的转录水平的强启动子。强启动子的实例是人巨细胞病毒立即早期启动子(CMV)和具有内含子A的CMV(Chapman等, Nucl. Acids Res. 19:3979-3986, 1991)。启动子的另外的实例包括天然存在的启动子,例如EF1 $\alpha$ 启动子、鼠CMV启动子、劳斯肉瘤病毒启动子和SV40早期/晚期启动子以及[ $\beta$ ]-肌动蛋白启动子;以及人工启动子,例如合成肌肉特异性启动子和嵌合肌肉特异性/CMV启动子(Li等, Nat. Biotechnol. 17:241-245, 1999, Hagstrom等, Blood 95:2536-2542, 2000)。

[0110] 核糖体结合位点位于或靠近起始密码子。优选的核糖体结合位点的实例包括CCACCAUGG、CCGCCAUGG和ACCAUGG,其中AUG是起始密码子(Kozak, Cell 44:283-292, 1986)。聚腺苷酸化信号负责切割转录的RNA并向RNA添加poly(A)尾。在高等真核生物中聚腺苷酸化信号包含AAUAAA序列的来自聚腺苷酸化添加位点的约11-30个核苷酸。AAUAAA序列包含在信号传导RNA切割中(Lewin, Genes IV, Oxford University Press, NY, 1990)。poly(A)尾对于mRNA的加工、从细胞核的输出、翻译和稳定性是重要的。

[0111] 可用作基因表达盒的一部分的聚腺苷酸化信号包括最小的兔[ $\beta$ ]-珠蛋白聚腺苷酸化信号和牛生长激素聚腺苷酸化(BGH)(Xu等, Gene 272:149-156, 2001, Post等, 美国专利 US 5,122,458)。

[0112] 可能存在的用于增强或调控基因表达或多肽加工之另外的调控元件的实例包括增强子、前导序列和操纵子。增强子区提高了转录。增强子区的实例包括CMV增强子和SV40增强子(Hitt等, Methods in Molecular Genetics 7:13-30, 1995, Xu等, Gene 272:149-156, 2001)。增强子区可与启动子相关联。

[0113] 可以调控根据本发明的第二方面的蛋白质或根据本发明的第二方面的蛋白质的表达。可在基因表达的许多步骤中实现这种调控。可能的调控步骤为例如但不限制于:转录的起始、启动子清除、转录的延伸、剪接、从细胞核的输出、mRNA稳定性、翻译的起始、翻译效率、翻译的延伸和蛋白质的折叠。影响细胞内部因子1或因子2多肽浓度的其他调控步骤影响蛋白质的半衰期。这样的调控步骤是例如,蛋白质的经调控变性。因为本发明的蛋白包含分泌蛋白,所述蛋白可被引导到宿主细胞的分泌途径。分泌的效率连同调控步骤一起调控

表达和蛋白质稳定性,细胞外部各蛋白质的浓度。细胞外部可指,例如但不限制于:培养基、组织、细胞内基质或空间或者体液例如血液或淋巴。

[0114] 上面提到的调控步骤的控制可以,例如不依赖于细胞类型或组织类型或者是细胞类型或组织类型特异性的。在一个本发明的特别优选的实施方案中,调控步骤的控制是细胞类型或组织类型特异性的。优选通过有关核酸转录的调控步骤实现这样的细胞类型或组织类型特异性调控。可通过使用细胞类型或组织类型特异性的启动子序列来完成这种转录调控。这种细胞类型或组织类型特异性调控的结果可以有不同等级的特异性。这意味着,相比于其他细胞类型或组织类型或表达限制于相应的细胞或组织类型,增强了相应的细胞或组织中相应多肽的表达。细胞或组织类型特异性启动子序列是本领域公知的并且可用于范围广泛的细胞或组织类型。

[0115] 在另一个优选的实施方案中,表达不是细胞类型或组织类型特异性的而是取决于生理条件。这样的条件是,例如炎症或创伤。也可通过调控所有上面提到的调控步骤来完成这种生理条件特异性的表达。对于生理条件特异性表达的调控的优选方式是转录调控。用于此目的,可使用创伤或炎症特异性启动子。各自的启动子为例如天然存在的序列(其可以例如来自基因,所述基因在免疫反应和/或创伤组织的再生期间特异性表达)。另一种可能性是使用人工启动子序列,其是,例如通过两个或更多个天然存在的序列的组合而构建。

[0116] 在另一个优选的实施方案中,调控是细胞类型或组织类型特异性的并且是生理条件特异性的。在一个特别优选的实施方案中,表达是心脏特异性表达。在另一个特别的实施方案中,表达是心脏特异性和创伤特异性的。

[0117] 用于调控根据本发明的第二方面的蛋白质或根据本发明的第二方面的蛋白质的表达的另一种可能性是基因表达的条件调控。为了实现条件调控,可以使用操纵子序列。例如,Tet操纵子序列可用于抑制基因表达。通过Tet操纵子连同Tet阻抑物手段进行基因表达的条件调控是本领域公知的并且已经建立了用于广泛的原核和真核生物体的许多相应的系统。本领域技术人员知晓如何选择合适的系统并使其适应于相应应用的特殊需求。

[0118] 在一个特别优选的实施方案中,根据本发明的核酸的用途包括在心肌梗死后向个体施加并且所述愈合包括左心室收缩功能的改善并与在梗死边界区的毛细血管密度的增加相关。此外,它可以在心肌梗死后降低死亡率。可用于确定如左心室收缩功能的改善,梗死边界区中毛细血管密度的增加以及心肌梗死后死亡率的降低之参数的方法是本领域公知的并在实施例4-6中示例性地进行了描述。

[0119] 上述的编码因子1和因子2蛋白、片段或变体的核酸优选用于治疗或改善萎缩;发育不全;炎症、损伤;创伤、局部缺血;再灌注损伤;炎症;感染;外伤;机械性超负荷;中毒;原发性或获得性心肌病,优选继承性心肌病和自发突变引起的心肌病。心肌病为例如但不限制于:肥厚性心肌病(HCM或HOCM)、致心律失常性右室心肌病(ARVC)、孤立心室非致密线粒体肌病、扩张性心肌病(DCM)、限制性心肌病(RCM)、Takotsubo心肌病、吕弗勒氏心内膜炎、糖尿病性心肌病、酒精性心肌病、肥胖相关的心肌病;心肌梗死或改善左心室收缩功能。

[0120] 在第四方面,本发明提供了包含第三方面的核酸或表达系统的载体,其用于增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈合和/或抑制非转化组织或非转化细胞的凋亡。

[0121] 术语“增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈

合和/或抑制非转化组织或非转化细胞的凋亡”具有上述定义的含义,以及优选的含义。

[0122] 如本文所使用的术语“载体”指的是蛋白质或多核苷酸或其混合物,其能够被引入到细胞中或将包含在其中的蛋白质和/或核酸引入到细胞中。在本发明的上下文中,优选在引入载体后在宿主细胞内表达由引入的多核苷酸编码的目的基因。合适的载体的实例包括但不限于:质粒载体、粘粒载体、噬菌体载体(例如 $\lambda$ 噬菌体载体、丝状噬菌体载体)、病毒载体、病毒样颗粒和细菌芽孢。

[0123] 在本发明的一个优选的实施方案中,载体是病毒载体。合适的病毒载体包括但不限于:腺病毒载体、腺相关病毒(AAV)载体、甲病毒载体、疱疹病毒载体、麻疹病毒载体、痘病毒载体、疱疹性口炎病毒载体、逆转录病毒载体和慢病毒载体。

[0124] 在一个本发明的特别优选的实施方案中,载体是腺病毒载体或腺相关病毒(AAV)载体。

[0125] 使用适于治疗施用的载体可将编码本发明第一方面的一种或更多种蛋白质和/或本发明第二方面的一种或更多种蛋白质的核酸引入到宿主细胞、组织或个体中。合适的载体可优选将核酸递送至靶细胞中,而不引起不可接受的副作用。

[0126] 在一个特别优选的实施方案中,根据本发明的载体的用途包括在心肌梗死后向个体施加并且愈合包括左心室收缩功能的改善且与梗死边界区中毛细血管密度的增加相关。此外,它可以在心肌梗死后降低死亡率。可用于确定如左心室收缩功能的改善、梗死边界区毛细血管密度的增加以及心肌梗死后死亡率的降低之参数的方法是本领域公知的并在实施例4-6中示例性地进行了描述。

[0127] 上述包含编码因子1和因子2蛋白、片段或变体之核酸的载体优选用于治疗或改善萎缩;发育不全;炎症、损伤;创伤、局部缺血;再灌注损伤;炎症;感染;外伤;机械性超负荷;中毒;原发性或获得性心肌病,优选继承性心肌病和自发突变引起的心肌病。心肌病为例如但不限制于:肥厚性心肌病(HCM或HOCM)、致心律失常性右心室心肌病(ARVC)、孤立心室非致密线粒体肌病、扩张性心肌病(DCM)、限制性心肌病(RCM)、Takotsubo心肌病、吕弗勒氏心内膜炎、糖尿病性心肌病、酒精性心肌病或肥胖相关的心肌病;心肌梗死或改善左心室收缩功能。

[0128] 在第五方面,本发明提供了包含第一方面和/或第二方面的蛋白质和/或第三方面的核酸和/或第四方面的载体的药物组合物,且任选地包含载体,其用于增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈合和/或抑制非转化组织或非转化细胞的凋亡。

[0129] 术语“增强非转化组织或非转化细胞的增殖和/或使非转化组织或非转化细胞愈合和/或抑制非转化组织或非转化细胞的凋亡”具有上述定义的含义,以及优选的含义。

[0130] 本文所使用的术语“载体”指的是药理学上无活性的物质,例如但不限制于:与治疗活性成分一起施用的稀释剂、赋形剂、表面活性剂、稳定剂、生理缓冲溶液或载剂。这样的药物载体可以是液体或固体。液体载体包括但不限于:无菌液体,例如在水和油中的盐水溶液,包括但不限于石油、动物油、植物油或合成来源的油例如花生油、大豆油、矿物油、芝麻油等。盐溶于水液以及水性葡萄糖和甘油溶液也可用作液体载体,特别是用于可注射的溶液。当药物组合物经静脉内施用,盐水溶液是优选的载体。合适的药物载体的实例由E.W.Martin在“Remington's Pharmaceutical Sciences”中进行了描述。在本发明的一个

优选实施方案中,载体是合适的药物赋形剂。合适的药物“赋形剂”包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、大米、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石、氯化钠、脱脂奶粉、甘油、丙烯、乙二醇、水、乙醇等。这种合适的药物赋形剂是优选的可药用的赋形剂。

[0131] “可药用的”意指由联邦或州政府管理机构批准的或在美国药典或其他一般公认的用于动物,且更特别地用于人的药典中列出的。

[0132] 术语“组合物”意在包括作为载体的封装材料与活性化合物的制剂,提供这样一种胶囊,其中具有或不具有其他载体的活性组分被载体包围,所述载体因此与所述活性化合物结合。

[0133] 术语“活性成分”指的是在药物组合物或制剂中具有生物学活性,即提供了药用价值的物质。在本发明的上下文中,活性成分是第一方面和/或第二方面的蛋白质和/或第三方面的核酸和/或第四方面的载体。药物组合物可包括可以相互结合起作用或彼此独立起作用的一种或更多种活性成分。所述活性成分可配制成中性或盐的形式。盐的形式优选可药用盐。

[0134] 术语“可药用盐”指的是例如但不限制于本发明多肽的盐。合适的可药用盐包括,例如,其可通过混合本发明多肽的溶液与可药用酸的溶液形成的酸加成盐,所述可药用酸例如盐酸、硫酸、富马酸、马来酸、琥珀酸、乙酸、苯甲酸、柠檬酸、酒石酸、碳酸或磷酸。此外,当肽携带酸性部分时,其合适的可药用盐可包括碱金属盐(例如,钠盐或钾盐);碱土金属盐(例如,钙盐或镁盐);以及与合适的有机配体(例如,铵、季铵以及使用抗衡阴离子例如卤化物、氢氧化物、羧酸盐、硫酸盐、磷酸盐、硝酸盐、烷基磺酸盐和芳基磺酸盐形成的胺阳离子)形成的盐。可药用盐的示例性实例包括但不限于:乙酸盐、己二酸盐、藻酸盐、抗坏血酸盐、天冬氨酸盐、苯磺酸盐、苯甲酸盐、碳酸氢盐、硫酸氢盐、酒石酸氢盐、硼酸盐、溴化物、丁酸盐、乙二胺四乙酸钙、樟脑酸盐(camphorate)、樟脑磺酸盐、右旋樟脑磺酸盐(camsylate)、碳酸盐、氯化物、柠檬酸盐、克拉维酸盐(clavulanate)、环戊烷丙酸盐、二葡萄糖酸盐、二盐酸盐、十二烷基硫酸盐、乙二胺四乙酸盐、乙二磺酸盐(edisylate)、依托酸盐(estolate)、乙磺酸盐(esylate)、乙磺酸盐、甲酸盐、富马酸盐、葡庚糖酸盐(gluceptate)、葡庚糖酸盐(glucoheptonate)、葡萄糖酸盐、谷氨酸盐、甘油磷酸盐、乙醇酰阿散酸盐、半硫酸盐(hemisulfate)、庚酸盐、己酸盐、己基间苯二酚盐、海巴明盐(hydrabamine)、氢溴酸盐、盐酸盐、氢碘酸盐、2-羟基-乙磺酸盐、羟基萘甲酸盐、碘化物、异硫代硫酸盐、乳酸盐、乳糖酸盐(lactobionate)、月桂酸盐、十二烷基硫酸盐、苹果酸盐、马来酸盐、丙二酸盐、扁桃酸盐、甲磺酸盐、甲基磺酸盐、甲基硫酸盐、粘酸盐(muicate)、2-萘磺酸盐、萘磺酸盐(napsylate)、烟酸盐、硝酸盐、N-甲基葡萄糖胺铵盐、油酸盐、草酸盐、扑酸盐(双羟萘酸盐)、棕榈酸盐、泛酸盐、果胶酸盐、过硫酸盐、3-苯丙酸盐、磷酸盐/二磷酸盐、苦味酸盐、新戊酸盐、聚半乳糖醛酸盐、丙酸盐、水杨酸盐、硬脂酸盐、硫酸盐、碱式乙酸盐、琥珀酸盐、单宁酸盐、酒石酸盐、茶氯酸盐(teoclolate)、甲苯磺酸盐、三乙基碘化物(triethiodide)、十一烷酸盐(undecanoate)、戊酸盐等(参见,例如,S.M.Berge等,“Pharmaceutical Salts” J.Pharm.Sci.66,第1-19页(1977))。

[0135] 将活性成分以有效量施用于细胞、组织或个体。“有效量”是足以实现预期目的的有效成分的量。活性成分可以是治疗剂。给定的活性成分的有效量将随参数,例如成分的性质、施用途径、接受活性成分的个体的尺寸和物种以及施用的目的而不同。在每个单独的情

况下,可通过技术人员根据本领域已建立的方法凭经验来确定有效量。如本发明上下文中的使用的“施用”包括在体内向个体施用,以及在体外或离体直接向细胞或组织施用。

[0136] 在本发明的一个优选实施方案中,药物组合物被定制为治疗疾病或病症。如本文所使用的疾病或病症的“治疗”意指实现以下的一个或多个:(a)降低病症的严重程度;(b)限制或阻止被治疗病症的症状的特征发展;(c)抑制被治疗病症的症状特征恶化;(d)限制或防止之前具有该病症的患者中该病症的复发;(e)限制或防止之前具有该病症之症状的患者中症状的复发;(f)降低疾病或病症发生后的死亡率;(g)愈合;及(h)预防疾病。如本文所使用的疾病或病症的“防止”或“预防”意指防止这样的疾病或病症在患者体内发生。

[0137] 在一个本发明的特别优选的实施方案中,用根据本发明的药物组合物的治疗,愈合包括在心肌梗死后个体的治疗且愈合包括改善左心室的收缩功能,并且可与梗死边界区毛细血管密度的增加相关。此外,它可在心肌梗死后降低死亡率。可用于确定如左心室收缩功能的改善、梗死边界区毛细血管密度的增加以及心肌梗死后死亡率的降低之参数的方法是本领域公知的,并在实施例4-6中进行了示例性地描述。

[0138] 可将本发明考虑的药物组合物配制成本领域技术人员公知的多种形式。例如,本发明的药物组合物可以是液体的形式例如以溶液、乳液或悬浮液的形式。优选地,将本发明的药物组合物配制成用于肠胃外施用,优选地用于静脉内、动脉内、肌内、皮下、经皮、肺内、腹膜内、冠状动脉内、心内施用或通过粘膜施用,优选地用于静脉内、皮下或腹膜内施用。用于经口服或经肛门施用的制剂也是可能的。优选地,本发明的药物组合物是无菌水溶液的形式,其可含有另外的物质例如以使溶液与血液等张的足够的盐或葡萄糖。如果必要的话,应将水溶液适当缓冲(优选pH为3至9,更优选pH为5至7)。药物组合物优选为单位剂型。在这样的形式中药物组合物被分为含有适量活性成分的单位剂量。单位剂型可以是包装的制剂,含有离散量的药物组合物的包装例如小瓶或安瓿。

[0139] 药物组合物的施用优选通过静脉内、动脉内、肌内、皮下、经皮、肺内、腹膜内、冠状动脉内或心内途径施用,其中还包括本领域已知的其他施用途径。

[0140] 在这种情况下,药物组合物用于个体的治疗,使用药物组合物可代替用于相应疾病或病症的标准治疗或可额外地施用于标准治疗。在额外地使用药物组合物的情况下,可在标准治疗之前、同时或之后施用药物组合物。在一个优选的实施方案中,标准治疗是再灌注治疗且在再灌注治疗之前、同时或之后施用药物组合物。

[0141] 更加优选地,所述药物组合物施用一次或多次。这包括2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45或50次。所述药物组合物施用的时间跨度没有限制。优选地,施用不超过1、2、3、4、5、6、7或8周。

[0142] 药物组合物的单次剂量可独立于施用剂量的总量或施用的相应时间跨度,作为一次或多次快速推注(bolus injection)和/或输注施用。

[0143] 本发明的第一方面至第五方面是基于本发明人观察到因子1和因子2在体外和体内是有效的促血管发生分子。因此,上述的蛋白质、核酸和载体也可预期用于离体用途,其可以是例如用于离体刺激上述细胞的或用于细胞培养应用中的治疗剂。

[0144] 然而,因子1和2的促血管发生活性的观察导致本发明人研究了因子1和因子是否2分别是抗血管发生治疗的靶标。本发明人成功地证明因子1的抑制或因子2的抑制可用于抑制血管发生。

[0145] 抗血管发生的策略被用于治疗癌性病症和其他疾病,例如年龄相关性黄斑变性,其中血管发生有助于疾病进展(参见,例如,Ferrara N和Kerbel RS (2005) *Nature*:438:967-974或Potente M等,(2011) *Cell*.146(6):873-887)。因此,在另一些方面,本发明涉及因子1和因子2之抑制剂的抗血管发生特性。在这些方面中,以上在定义部分提供的定义同样适用。此外,除非上下文对其用途另有明确说明,否则在第一方面至第五方面的描述中以及优选的实施方案中提供的具体定义,例如术语“载体”以及优选的载体也适用于本发明的以下方面。

[0146] 在第六方面,本发明提供了因子1和/或2蛋白质的抑制剂,其用于医疗用途,优选用于治疗或预防其中血管发生有助于疾病的发生或进展的疾病。广泛使用的术语“抑制剂”指干扰因子1或因子2之促血管发生活性的化合物。抑制剂可能作用于编码因子1或因子2的mRNA的转录和/或翻译,从而防止其细胞产生以及进而防止分泌到循环中和/或疾病发生或进展的位点。抑制剂也可通过与因子1或因子2蛋白或者与因子1或因子2蛋白特异性结合的细胞蛋白质特异性地结合而发挥作用,优选地与其细胞受体地结合。这种结合可防止或干扰因子1或因子2与其他细胞蛋白质,优选与其相应细胞受体的自然相互作用。技术人员非常清楚如何干扰受体与其激动剂之间的结合,并且可利用这一知识来设计因子1和因子2的合适的抑制剂。此外,也可通过分别缺失或突变因子1和因子2的蛋白质赋予发挥因子1和因子2抗血管发生功能的那些部分,从因子1或因子2蛋白本身衍生出抑制剂。这种非活性的突变的或缺失的因子1或因子2将与野生型因子1和因子2竞争其天然结合配体。干扰因子1或因子2的促血管发生活性的化合物降低至少20%,优选至少30%,更优选至少40%的活性。在抑制剂分别地特异性地结合因子1或因子2或者包含因子1和因子2的突变体或片段、基本上由因子1和因子2的突变体或片段组成或由因子1和因子2的突变体或片段组成的情况下,优选的是它们在等摩尔浓度发挥了对因子1蛋白或因子2蛋白的这种抑制水平。在本文的实施例10中描述了可用于确定抑制促血管发生活性的优选的测定法。为了分别确定给定的抑制剂在等摩尔量的因子1和2摩尔量下是否具有这种活性,必须确定各自的抑制剂。根据SEQ ID NO:1的因子1具有15.84kD的分子量和根据SEQ ID NO:3的因子2具有21.57kD的分子量以及IgG具有约150kD的分子量。因此,100ng因子1与947ng因子1特异性IgG是约等摩尔的,100ng因子2和695ng因子2特异性IgG是约等摩尔的。分别从图12的A和B明显看出,在此意义上,本文提供的因子1和因子2特异性抗体是因子1蛋白或因子2蛋白的抑制剂。如果在抑制剂干扰编码因子1或因子2之mRNA的转录和/或翻译的情况下,优选在基于细胞中产生的蛋白质(自然产生的因子1或因子3)的基础上测量抑制水平。技术人员熟知许多测量编码因子1或因子3之mRNA的量以及因子1或因子2蛋白质的量的方法,所述方法可用于评估化合物干扰编码因子1或因子2的mRNA的转录和/或翻译的能力。优选地,因子1蛋白包含根据SEQ ID NO:1的氨基酸序列或其变体、基本上由根据SEQ ID NO:1的氨基酸序列或其变体组成或由根据SEQ ID NO:1的氨基酸序列或其变体组成,因子2包含根据SEQ ID NO:3的氨基酸序列或其变体、基本上由根据SEQ ID NO:3的氨基酸序列或其变体组成或由根据SEQ ID NO:3的氨基酸序列或其变体组成,所述变体与SEQ ID NO:1或3具有至少80%序列同一性。

[0147] 优选地,在本发明的上下文中所使用的抑制剂是蛋白质,所述蛋白质包含根据SEQ ID NO:1或3的氨基酸序列或其变体的抑制性片段或抑制性突变体、基本上由根据SEQ ID NO:1或3的氨基酸序列或其变体的抑制性片段或抑制性突变体组成或由根据SEQ ID NO:1

或3的氨基酸序列或其变体的抑制性片段或抑制性突变体组成,所述变体与SEQ ID NO:1或3的氨基酸序列具有至少80%序列同一性。本领域公知通过受体由蛋白质-蛋白质相互作用发挥其功能的蛋白质包括结合到受体所必需的结构域,以及引发受体发送信号到细胞中的结构域。因此,技术人员清楚地知晓如何产生此类受体结合蛋白质的抑制片段或突变体。例如,可生成一系列N-端截短的和/或C-端截短的因子1或因子2蛋白并按照实施例2中所概述的试样测试其促血管发生活性。然后,按照实施例10中概述的试验对不再显示促血管发生活性的那些片段进行测试,测试其抑制因子1或2蛋白的促血管发生活性的能力。如本领域已知的,可通过丙氨酸扫描诱变生成因子1和2的突变体。在丙氨酸扫描中产生了一系列的突变体,其中每个突变体包含1、2、3或更多个已被突变为丙氨酸的氨基酸(所谓的盒),并且其中突变体的不同在于因子1和因子2内盒的位置。可按照实施例2中所概述的再次鉴定已丧失促血管发生活性的因子1或2的突变体。随后可通过使用如实施例10中概述的测定法来鉴定抑制性突变体。

[0148] 在另一个优选的实施方案中,抑制剂是与根据SEQ ID NO:1或3的氨基酸序列或其变体特异性结合的配体,所述变体与根据SEQ ID NO:1或3的氨基酸序列具有至少80%序列同一性,或者与根据SEQ ID NO:1或3的氨基酸序列或其变体自然相互作用之受体特异性结合的配体,所述变体与根据SEQ ID NO:1或3的氨基酸序列具有至少80%序列同一性。在本发明中使用的术语“配体”指的是与指定的抗原特异性结合的化学部分。优选的配体是氨基酸基配体,例如免疫球蛋白,优选抗体或其抗原结合片段和抗体样蛋白。或者,配体可以是肽模拟物。

[0149] 本文中所使用的术语“免疫球蛋白(Ig)”指的是免疫球蛋白超家族的赋予免疫性的糖蛋白。“表面免疫球蛋白”通过其跨膜区与效应细胞的膜连接并且涵盖分子例如但不限制于:B细胞受体、T细胞受体、I类和II类主要组织相容性复合体(MHC)蛋白、 $\beta$ -2微球蛋白( $\beta$ 2M)、CD3、CD4和CD8。一般,如本文所使用的术语“抗体”指的是缺少跨膜区的分泌的免疫球蛋白,因此,可被释放到血液和体腔中。基于抗体具有的重链,将其分为不同的同种型。有五种类型的人Ig重链,用希腊字母: $\alpha$ 、 $\delta$ 、 $\epsilon$ 、 $\gamma$ 和 $\mu$ 表示。存在的重链类型定义了抗体的种类,即这些链分别在IgA、IgD、IgE、IgG和IgM抗体中被发现,并各自执行不同的作用,并指导针对不同类型抗原的适当的免疫应答。不同重链的大小和组成不同; $\alpha$ 和 $\gamma$ 包含约450个氨基酸,而 $\mu$ 和 $\epsilon$ 具有约550个氨基酸(Janeway等,(2001) Immunobiology, Garland Science)。抗体包含四条多肽链,即通过二硫键相互连接的两条重链(H)和两条轻链(L)。每条重链由重链可变区(本文缩写为HCVR或VH)和重链恒定区组成。重链恒定区由三个结构域CH1、CH2和CH3组成。每条轻链由轻链可变区(本文缩写为LCVR或VL)和轻链恒定区组成。轻链恒定区由结构域CL组成。VH和VL区可以进一步细分为称为互补决定区(CDR)的高变区,其散布于更保守的区域(称为框架区(FR))中。每个VH和VL由三个CDR和四个FR组成,从氨基端至羧基端按以下顺序排列:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4。可按照本领域已知的来确定重链和轻链的CDR。例如,可用以下这组规则来寻找分别在抗体轻链和重链序列内的CDR:

[0150] 轻链CDR-1:开始:约第24位残基,在CDR-1之前的残基总是Cys,在CDR1之后的残基总是Trp。一般是Trp-Tyr-Gln,但也可可是Trp-Leu-Gln、Trp-Phe-Gln、Trp-Tyr-Leu;长度:10至17个残基。

[0151] 轻链CDR-2:开始:总是在L1末端后的16个残基,之前的残基一般是Ile-Tyr,但也



可是Val-Tyr-Ile-Lys-Ile-Phe,长度总是7个残基;

[0152] 轻链CDR-3:开始:总是在CDR-2末端后的33个残基;之前的残基总是Cys,之后的残基总是Phe-Gly-XXX-Gly,长度:7至11个残基。

[0153] 重链的CDR-1:开始:约第26位残基,总是在Cys之后的4个残基(根据Chothia AbM定义,Kabat定义5个残基之后开始);之前的残基总是Cys-XXX-XXX-XXX;之后的残基总是Trp。一般,Trp-Val,但是也可以是Trp-Ile、Trp-Ala,长度:10至12个残基(AbM定义,Chothia定义排除了最后4个残基);

[0154] 重链CDR-2:开始:总是在Kabat/AbM定义的重链CDR-1末端之后的15个残基;之前的残基一般是Leu-Glu-Trp-Ile-Gly,但是有多种变化,之后的残基是Lys/Arg-Leu/Ile/Val/Phe/Thr/Ala-Thr/Ser/Ile/Ala,长度Kabat定义为16至19个残基(根据AbM定义;较早Chothia定义末端7个残基)。

[0155] 重链CDR-3:开始:总是重链CDR-2末端之后的33个残基(总是在半胱氨酸之后的2个氨基酸残基);之前的残基总是Cys-XXX-XXX(典型的Cys-Ala-Arg);之后的残基总是Trp-Gly-XXX-Gly;长度:3至25个残基。这组规则是本领域技术人员已知的并且也可在<http://www.bioinf.org.uk/abs/#cdrid>上找到。

[0156] 如本文所使用的术语“人抗体”旨在包括源自人种系免疫球蛋白序列的具有可变区和恒定区的抗体。人mAb可包括不由人种系免疫球蛋白序列编码的氨基酸残基(例如,通过体外随机或位点特异性的诱变或通过体内体细胞突变引入的突变),例如在CDR中。然而,如本文所使用的术语“人抗体”,并非意在包括“人源化抗体”,其中来自另一哺乳动物物种(如小鼠)的种系来源的CDR序列已被移植到人FR序列上。人抗体还包括从人免疫球蛋白文库中分离的抗体或从转基因有一个或更多个人免疫球蛋白但不表达内源性免疫球蛋白的动物中分离的抗体。

[0157] 如本文所使用的术语“单克隆抗体”指的是单一分子组成的抗体分子的制备物。单克隆抗体显示出对特定表位的单一的结合特异性和亲和力。在一个实施方案中,单克隆抗体由杂交瘤产生,所述杂交瘤包括与无限增殖化细胞融合的来自非人类动物例如小鼠中所获得的B细胞。如本文所使用的术语“重组抗体”,包括通过重组方法制备、表达、产生或分离的所有人抗体,例如(a)从具有人免疫球蛋白基因的转基因或转染色体动物(例如小鼠)或由所述动物制备的杂交瘤中分离的抗体,(b)从经转化表达所述抗体的宿主细胞(例如从转染瘤)中分离的抗体,(c)从重组体、组合人抗体文库中分离的抗体,以及(d)通过任何包括将人免疫球蛋白基因序列剪接至其他DNA序列上的其他方法制备、表达、产生或分离的抗体。如本文所使用的“异源抗体”是相对于产生此类抗体的转基因生物而定义的。该术语指的是指这样一种抗体,它具有与存在于不包括转基因生物体的生物体中的氨基酸序列或编码核酸序列相对应,一般是来源于转基因生物体以外的物种获得的氨基酸序列或编码核酸序列。如本文所用的“异源杂种抗体”指的是具有来源于不同生物体的轻链和重链的抗体。例如,具有与鼠轻链结合的人重链的抗体是异源杂种抗体。

[0158] 术语“抗原结合片段”指的是保留特异性结合抗原或抗原蛋白质之功能的抗体的片段,但其缺乏抗体或包含抗体部分的人工构建体之一些或全部的其他结构特征。抗原结合片段的优选实例包括但不限于以下的Fab片段、Fc片段、Fab'片段、F(ab')<sub>2</sub>、单结构域抗体(sdAb)、纳米抗体、单链Fv、二价单链可变片段(di-scFv)、串联的scFv、双抗体、单链双

抗体(scDB)、三抗体、双特异性T细胞接合器(Bi-specific T-cell engager, BiTE)或双亲和性重定向分子(dual affinity retargeting molecule, DART分子)。

[0159] “Fab片段”(也称为“Fab部分”或“Fab区”),各自具有单个抗原结合位点以及残余的“Fc片段”(也称为“Fc部分”或“Fc区”),它的名称反映了其易于结晶的能力。“Fab’片段”指的是还包括Ig分子的铰链区的Fab片段,而“F(ab’)<sub>2</sub>片段”被理解为包括通过化学连接的或通过二硫键连接的两个Fab’片段。虽然sdAb(Desmyter等,1996)和“纳米抗体”只包括单个V<sub>H</sub>结构域,但是“单链Fv(scFv)”片段包含通过短接头肽连接至轻链重链可变结构域的重链可变结构域(Huston等,1988)。可通过连接两个scFv(scFvA-scFvB)改造成di-scFv。这可通过制造具有两个V<sub>H</sub>和两个VL区的单一肽链得到“串联scFv”(V<sub>H</sub>A-V<sub>L</sub>A-V<sub>H</sub>B-V<sub>L</sub>B)来完成。另一种可能性是产生具有接头的scFv,所述接头过短以至于两个可变区不能折叠在一起,迫使scFv二聚化。通常使用具有5个残基长度的接头来产生这些二聚体。这种类型被称为“双抗体”。仍然是V<sub>H</sub>和V<sub>L</sub>结构域之间较短的接头(一个或两个氨基酸)导致形成单特异性三聚体,所谓的“三抗体”或“三体”。通过分别表达具有排列V<sub>H</sub>A-V<sub>L</sub>B和V<sub>H</sub>B-V<sub>L</sub>A或者V<sub>L</sub>A-V<sub>H</sub>B和V<sub>L</sub>B-V<sub>H</sub>A的链形成双特异性双抗体。单链双抗体(scDb)包含由12至20个氨基酸,优选14个氨基酸的接头肽(P)连接的V<sub>H</sub>A-V<sub>L</sub>B和V<sub>H</sub>B-V<sub>L</sub>A片段(V<sub>H</sub>A-V<sub>L</sub>B-P-V<sub>H</sub>B-V<sub>L</sub>A)。“双特异性T细胞接合器(BiTE)”是由不同抗体的2个scFv组成的融合蛋白,其中一个scFv通过CD3受体与T细胞结合,另一个通过肿瘤特异性分子与肿瘤细胞结合(Kufer等,2004)。双重亲和性重定向分子(“DART”分子)是另外通过C端二硫桥稳定的双抗体。

[0160] 术语“抗体样蛋白”指的是在与抗体在结合抗原或抗原性蛋白方面具有相似特性但无需具有抗体的结构特征的蛋白质。抗体样蛋白可天然存在或可被人工设计(例如通过生物技术)。天然存在的抗体样蛋白的实例包括但不限于:抗原-结合蛋白例如脂质运载蛋白家族,它们通常服务于生理上重要的化合物的储存或运输的多种蛋白质家族。它们共享八个反平行β链的保守桶作为其中央折叠基序,并在此桶结构末端包括六个连接至每个β链对的高变环。这些环形成结合口袋的入口。脂质运载蛋白家族的成员之间的结构多样性反映了它们的结合伴侣的不同形状和化学性质。因此,虽然它们由单链多肽所组成且比免疫球蛋白小得多,但是它们表现出结合不同特异性的抗原的巨大潜力。人工设计的抗体样蛋白的实例包括:基于骨架的蛋白质(scaffold-based protein),其通过将已知亲和力的肽与某一靶标融合或通过将所述肽插入骨架蛋白中来产生,使肽的结合特性与所需骨架载体的有利的特性相结合。本领域技术人员知晓许多具有这样基于骨架的蛋白质。如本文所使用的术语“骨架蛋白”指的是具有结构刚性,即折叠成稳定的三级结构的蛋白质。骨架蛋白的氨基酸可能占据骨架蛋白中限定的三维位置。因此,如果骨架蛋白的一个或更多个氨基酸被适当长度的多肽取代,多肽会占据与取代那些位置类似的位置。这使得给定的多肽定位在骨架蛋白中限定的三维位置和/或方向。相应地,骨架蛋白可作为抗体的替代用于分子识别(参见,例如,Skerra A: (2007) Curr. Opin. Biotechnol. 2007, 18: 295-304或Skerra A (2000) J. Mol. Recognit. 2000, 13: 167-187)。这样的骨架蛋白的一个实例是Fyn SH3结构域,所述Fyn SH3结构域包括可被突变以转移新的结合特异性至SH3结构域的两个结构域。在例如W0 2000/072742或W0 2008/022759中公开了选择与给定抗原特异性地结合的Fyn SH3的方法。

[0161] 在本发明的上下文中,术语“肽模拟物(peptidomimetics)”是用来指其基本要素

(药效团)模拟了在3D空间中的天然肽或蛋白质并保留了与生物学靶标相互作用的能力以及产生相同生物学效应的任何分子。肽模拟物包括设计为模拟肽的小的蛋白样链,其一般可通过修饰现有的肽或通过设计成模拟肽的类似的系统(例如拟肽和 $\beta$ -肽)得到。不论方法如何,经改变的化学结构设计成有利地调节分子性质使得增加或减少稳定性或生物活性。因此,修饰包括不会自然发生的对肽的改变,包括但不限于改变主链以及掺入非天然的氨基酸。

[0162] 术语对抗原例如因子1或因子2的“特异性结合”或“特异性地结合”指的是配体高亲和性地与抗原的抗原决定簇结合的能力。在这种情况下,“高亲和力”指的相互作用的 $K_d$ 低于 $1 \times 10^{-5}M$ ,优选低于 $1 \times 10^{-6}M$ ,更优选低于 $1 \times 10^{-7}$ ,甚至更优选低于 $1 \times 10^{-8}M$ 且最优选低于 $1 \times 10^{-9}M$ 。

[0163] 在本发明的第六方面的上下文中使用的优选的抗体是单克隆抗体,优选人或人源化抗体。

[0164] 在另一个优选的实施方案中,抑制剂是抑制或防止mRNA转录和/或翻译的核酸,所述mRNA编码的蛋白质包含根据SEQ ID NO:1或3的氨基酸序列或其变体,其与根据SEQ ID NO:1或3的氨基酸序列具有至少80%序列同一性。技术人员清楚地知道如何基于编码包含根据SEQ ID NO:1或3的氨基酸序列的蛋白质之基因组序列确定此类核酸的序列。这种抑制剂核酸的一个实例是对编码因子1或因子2的mRNA特异的siRNA。

[0165] 在其中本发明第六方面的抑制剂是能够由核酸编码的蛋白质的这种情况下,可以设想通过提供编码抑制剂的核酸施用抑制剂。因此,在第七方面,本发明提供了编码本发明第六方面之抑制剂的核酸,用于治疗或预防其中血管发生有助于疾病的发生或进展的疾病。所述核酸还可包括本发明第三方面的上下文中描述的任何元素。

[0166] 因此,在第八方面,本发明提供了包含本发明第六方面的核酸的载体以用于治疗或预防其中血管发生有助于疾病的发生或进展的疾病。在第七方面的上下文中的术语“载体”与如在本发明的第四方面的上下文中描述的具有相同的含义。

[0167] 在本发明这个方面的一个优选实施方案中,所述载体是病毒载体。合适的病毒载体包括但不限于腺病毒载体、腺相关病毒(AAV)载体、甲病毒载体、疱疹病毒载体、麻疹病毒载体、痘病毒载体、疱疹性口炎病毒载体、逆转录病毒载体和慢病毒载体。此外,在第九方面,本发明提供了包括第六方面的抑制剂、第七方面的核酸或第八方面的载体以及任选地合适的药物赋形剂的药物组合物,用于治疗或预防其中血管发生有助于疾病的发生或进展的疾病。该药物组合物还可包含任何上述的关于本发明的第五方面教导的成分。

[0168] 如在本发明第六方面至第九方面使用的术语“其中血管发生有助于疾病的发生或进展的疾病”指的是其中形成血管的细胞在疾病的发作、发生过程和/或进展中发生增殖的疾病。形成血管的细胞以及其可增殖的包括衬在血管内壁上的内皮细胞以及形成血管壁的平滑肌细胞。在健康的血管中内皮细胞或平滑肌细胞均不增殖。这些细胞在例如响应损伤或化学信号(例如VEGF)时才增殖。血管发生也被称为新血管形成并表征从预先存在的血管形成新血管的过程。在某些疾病例如眼部血管发生疾病中,血管的异常形成是疾病的起因,以及在像良性肿瘤或恶性肿瘤的一些疾病中,血管发生在疾病进展期间发生并用氧和营养物质支持生长的肿瘤块。在这些疾病中,血管发生不是疾病的起因,但促进了疾病的进展。恶性肿瘤的新血管形成通过向肿瘤细胞提供从肿瘤块的逃脱途径,从而有助于肿瘤转移而

有助于疾病进展。

[0169] 优选其中血管发生有助于疾病的发生或进展的疾病是增殖性疾病。优选的增殖性疾病选自：良性肿瘤、恶性肿瘤、类风湿性关节炎、银屑病、眼部血管发生疾病、奥斯勒-韦伯综合征 (Osier-Webber Syndrome)、斑块新血管形成 (plaque neovascularization)、移植物和血管成形术后狭窄 (graft and post-angioplasty stenosis)、毛细血管扩张 (telangiectasia)、血友病性关节、血管纤维瘤、伤口肉芽、肠粘连、动脉粥样硬化、硬皮病、肥大性瘢痕、猫抓病、和溃疡、特别是黄斑变性、肾上腺皮质癌、膀胱癌、骨癌、脑癌、乳腺癌、宫颈癌、结肠癌、结直肠癌、子宫内膜癌、食道癌、眼癌、胆囊癌、胃癌、头颈癌、喉癌、肝癌、肺癌、黑色素瘤、骨髓增生性疾病、颈癌、非黑色素瘤皮肤癌 (nonmelanoma skin cancer)、卵巢癌、前列腺癌、良性前列腺增生、胰腺癌、直肠癌和睾丸癌。待治疗的优选的疾病为良性肿瘤、恶性肿瘤和眼部血管发生疾病。

[0170] 本发明人已鉴定了针对因子1和因子2的某些表位,分别干扰了因子1和因子2的促血管发生功能的抗体,即是拮抗性抗体。因此,在第十方面,本发明分别涉及因子1和因子2的配体,所述配体分别抑制因子1和因子2的促血管发生活性。特别优选的是抑制因子1和因子2的促血管发生活性的抗体或其片段。本发明人已通过产生与因子1和2的表面暴露之结构域特异性结合的抗体,成功地提供了这样的抑制剂抗体的实例。在本发明这一方面的一个优选实施方案中,它涉及分别能够与因子1和2的这些片段特异性结合的配体。在本发明这个方面的上下文中,在一般定义部分中提供的定义和在本发明第六方面的上下文中的具体定义同样适用。

[0171] 因此,在一个优选实施方案中,本发明涉及配体,优选与人因子1蛋白的表位或对应于该表位的来自另一个因子1蛋白的区域特异性结合的抗体或其片段或者抗体样蛋白,所述表位包含在SEQ ID NO:1的61至76位氨基酸中或由SEQ ID NO:1的61至76位氨基酸组成。

[0172] 在另一个优选的实施方案中,本发明涉及配体,优选与人因子2蛋白的表位或对应于该表位的来自另一个因子2蛋白的区域特异性结合的抗体或其片段或者抗体样蛋白,所述表位包含在SEQ ID NO:3的181至195位氨基酸中或由SEQ ID NO:3的181至195位氨基酸组成。

[0173] 在本发明上述优选方面的上下文中,术语“对应于该表位的来自另一个因子1或2蛋白的区域”指的是来自另一个因子1或2蛋白的氨基酸序列,当使用标准比对工具例如像 ClustalW 以及如上概述的标准参数时,其与指定的因子1或2的氨基酸序列比对。图6至11显示了因子1和2蛋白几个这样的比对。技术人员能够容易地鉴定具有氨基酸序列 CTIWRPQGKSYLYFTQ (SEQ ID NO:38) 的SEQ ID NO:2的片段并确定另一个因子1蛋白的相应氨基酸。

[0174] 实施例

[0175] 设计实施例是为了进一步说明本发明并有助于更好地理解。他们不应被解释为以任何方式限制本发明的范围。

[0176] 实施例1

[0177] 在多中心,安慰剂对照的临床试验中,本发明人之一测试了在患有AMI的患者中,自体骨髓细胞的冠状动脉内输注的作用 (BOOST-2,对照试验识别号 ISRCTN17457407)。在试

验中,从AMI患者中获得骨髓穿刺液用于研究的目的。通过磁性细胞分离(MiniMACS™, Miltenyi Biotec)分离CXCR4+骨髓细胞。经过随后的两次纯化步骤后,获得了CXCR4+富集的细胞群体(通过流式细胞术证实纯度>95%)。下一步从这些细胞中分离了RNA且将RNA用于微阵列分析(Affymetrix GeneChip HG\_U133Plus 2.0)中。在随后的生物信息学分析中,检测了CXCR4+骨髓细胞在微阵列中表达最强烈的4000个表达序列标签(EST)。使用一系列的生物信息学工具来鉴定那些EST中以N端信号肽缺失线粒体或核信号肽、缺失内质滞留序列以及缺失跨膜结构域为特征的推定的分泌因子。总计鉴定了283个推定的分泌因子;发现其中117个在NCBI Blast中具有鼠同源物。将人同源物的cDNA克隆到表达质粒中,然后将质粒单独地转染到人胚胎肾(HEK)细胞中。在无血清培养基中培养转染的HEK细胞,30小时后获得条件培养上清液。条件HEK细胞上清液单独用于在小型化的血管发生测定中测试促血管发生作用以及在心肌细胞死亡测定中测试细胞保护作用。这种筛选导致鉴定了2个分泌蛋白;在上述测定中“因子1”显示了促血管发生作用和细胞保护作用;在上述测定中“因子2”显示了促血管发生作用。

[0178] 在筛选中鉴定的序列以及各自的小鼠或人同源物是:

[0179] 因子1:

[0180] 在筛选中鉴定了人因子1并用于实施例2和图1中:

[0181] 人19号染色体开放阅读框10(C19orf10)

[0182] 根据NCBI参照序列:NM\_019107.3(SEQ ID NO:6)获得编码人因子1的核酸序列。在图6中描述了人因子1的氨基酸(SEQ ID NO:2)。

[0183] 小鼠同源物用于实施例3-6和图2-5中:

[0184] 小鼠的DNA区段,Chr17,Wayne State University 104,表达的(D17Wsu104e)

[0185] 根据NCBI参照序列NM\_080837.2(SEQ ID NO:7)获得编码小鼠因子1的核酸序列。在图6中描述了小鼠因子1的氨基酸序列(SEQ ID NO:13)。

[0186] 因子2:

[0187] 在筛选中鉴定了分泌形式的人因子2并用于实施例2和图1中:

[0188] 人19号染色体开放阅读框63(C19orf63),转录变异体HSS1。根据NCBI参照序列:NM\_175063.4(SEQ ID NO:8)获得编码人因子2的核酸序列。在图8中描述了分泌形式的人因子2的氨基酸序列(SEQ ID NO:4)。

[0189] 跨膜形式的人因子2:

[0190] 人19号染色体开放阅读框63(C19orf63),转录变体HSM1。根据NCBI参照序列:NM\_206538.2(SEQ ID NO:9)获得编码跨膜形式的人因子2的核酸序列。在图10中描述了跨膜形式的人因子2的氨基酸序列(SEQ ID NO:5)。

[0191] 在GenBank:AY358710.1中获得跨膜形式的人因子2的变体的氨基酸序列(SEQ ID NO:34)。

[0192] 分泌形式的小鼠同源物用于实施例4-6和图3-5中:

[0193] 含有小鼠造血信号肽的分泌1(2310044H10Rik)mRNA、全长cds、选择性剪接。根据GenBank:AY761096.1获得编码跨膜形式的小鼠因子2的核酸序列(SEQ ID NO:10)。在图8中描述了分泌形式的小鼠因子2的氨基酸序列(SEQ ID NO:24)。

[0194] 跨膜形式的小鼠同源物:

[0195] 小鼠RIKEN cDNA 2310044H10基因(2310044H10Rik)。根据NCBI参照序列:NM\_197991.2获得编码跨膜形式的小鼠因子2的核酸序列(SEQ ID NO:11)。在图10中描述了跨膜形式的小鼠因子2的氨基酸序列(SEQ ID NO:29)。

[0196] 实施例2:

[0197] 为了证实在筛选中观察到的促血管发生活性,在COS7细胞中产生两种因子(由在SEQ ID NO:6和8所示的核酸序列编码的人类同源物)都作为His标记的重组蛋白。如图1中所示,在培养的人内皮细胞中,重组因子1和重组因子2促进剂量依赖性促血管发生作用。

[0198] 人冠状动脉内皮细胞(HCAEC)和人脐静脉内皮细胞(HUVEC)购自Provitro (Berlin, Germany)。如所示的,在不存在(对照)或存在10%胎牛血清(FCS)、人重组VEGF-A (R&D系统)或不同浓度的重组人因子1(SEQ ID NO:2)或因子2(SEQ ID NO:4)的情况下,将细胞于基本培养基中培养24小时。(A) 溴脱氧尿苷掺入法测量HCAEC增殖。(B) 在用移液管尖端对汇合的内皮细胞单层制造创伤后评估HCAEC迁移。(C) 在于生长因子减少的基质胶上培养的细胞中评估HUVEC网络形成。N=每个条件下3至5次独立的实验; \* $P<0.05$ , \*\* $P<0.01$ , \*\*\* $P<0.001$ 相对于对照(参见图1)。

[0199] 实施例3:

[0200] 为了证实在筛选中观察到的因子1的心肌细胞保护作用,在重组因子1存在或不存在的条件下,将新生大鼠心室心肌细胞进行模拟的局部缺血再灌注损伤。在COS7细胞中产生因子1(由在SEQ ID NO:7中所示的核酸序列编码的小鼠同源物)作为His标记的重组蛋白。如图2中所示,在培养的心肌细胞中重组因子1促进剂量依赖性抗凋亡作用。

[0201] 通过Percoll密度梯度离心从1至3天龄的Sprague-Dawley大鼠中分离心室心肌细胞。将心肌细胞暴露于模拟的局部缺血条件下180分钟(5%CO<sub>2</sub>/95%N<sub>2</sub>气氛下的含2-脱氧葡萄糖的不含葡萄糖培养基),随后如所示的,在不存在(对照)或存在重组人GDF-15 (R&D系统,已知的抗凋亡细胞因子)或不同浓度的重组小鼠因子1的情况下,进行60分钟模拟的再灌注(返回5%CO<sub>2</sub>/95%N<sub>2</sub>室内空气下的含葡萄糖培养基)。通过原位TdT-介导的dUTP缺口末端标记(TUNEL)来评估细胞死亡。N=每个条件下3次独立的实验; \* $P<0.05$ 相对于对照。

[0202] 实施例4:

[0203] 为了探索因子1和因子2在AMI的条件下的治疗潜力,在AMI的小鼠模型中产生并测试了编码因子1或因子2(核酸序列示于SEQ ID NO:7和10中)的鼠同源物的腺病毒。在梗死后28天,两种因子的腺病毒表达导致左心室收缩功能的改善。这是与梗死边界区毛细血管密度的增加相关(图3)。

[0204] 使用AdEasy XL载体系统(Stratagene)将小鼠因子1或因子2的cDNA克隆到复制缺陷型腺病毒中。使用编码半乳糖苷酶(lacZ)的复制缺陷型腺病毒作为对照。使用Adeno-X病毒纯化试剂盒(BD Biosciences)纯化病毒。将10至12周龄的雄性C57BL/6小鼠用异氟烷(1-2%)麻醉并通气,并进行永久的冠状动脉左前降支(LAD)结扎。LAD结扎后,立即将病毒( $5 \times 10^9$  p.f.u)注入左心室(LV)腔中。(A) 在LAD结扎28天后,通过经胸超声心动图(Visualsonics)评估左室收缩功能(切面面积变化,FAC) (N=每组10至12只小鼠)。(B) 在LAD结扎28天后,通过荧光显微术对在梗死边界区同工凝集素阳性的毛细血管密度进行定量(每组N=3只小鼠)。\* $P<0.05$ , \*\* $P<0.01$ 相对于Ad.lacZ对照。

[0205] 实施例5:

[0206] 为了研究当在再灌注的AMI的条件下(模拟AMI患者接收再灌注治疗的临床情况)作为重组蛋白质施用的因子1和因子2的治疗潜力,将10至12周龄的雄性C57BL/6小鼠经历1小时的冠状动脉结扎(局部缺血),随后再灌注28天。在灌注后的前7天,用两种因子经皮下注射治疗小鼠。在HEK293细胞中产生因子1和因子2(小鼠同源物;根据SEQ ID NOS:13和24的氨基酸序列)作为带His标签的重组蛋白。在梗死后28天,用重组因子1或重组因子2的治疗导致左心室收缩功能的显著改善。这与梗死边界区毛细血管密度的显著增加有关(图4)。

[0207] 将10至12周龄的雄性C57BL/6小鼠用异氟烷(1-2%)麻醉并通气,并经历1小时的短暂的左冠状动脉左前降支结扎,随后再灌注28天。在再灌注时(将PBS注射到对照小鼠中),使小鼠接受重组因子1或因子2(各10 $\mu$ g)的单次皮下注射。随后,使用Alzet微型泵(10 $\mu$ g/天)将重组因子1或因子2进行7天连续的皮下输注。对照小鼠用PBS输注。(A)在再灌注后28天,经胸超声心动图评估左心室的收缩功能(切面面积变化,FAC)(每组N=10至13只小鼠)。(B)在再灌注后28天,通过荧光显微术对在梗死边界区同工凝集素阳性的毛细血管密度进行定量(N=每组6只小鼠)。 $*P<0.05$ , $**P<0.01$ 相对于PBS对照。

#### [0208] 实施例6:

[0209] 为了研究因子1和因子2在作为重组蛋白质应用时是否能够增强AMI再灌注后的存活,使小鼠经历1小时的冠状动脉结扎,随后再灌注28天。在灌注后的前7天期间,用两种因子经皮下注射治疗小鼠。在HEK293细胞中产生因子1和因子2(小鼠同源物;SEQ ID NO:7和10)作为带His标签的重组蛋白。用重组因子1或重组因子2的治疗导致梗死前28天期间显著的存活提高(图5)。

[0210] 将10至12周龄的雄性C57BL/6小鼠用异氟烷(1-2%)麻醉并通气,并经受1小时的短暂的冠状动脉左前降支结扎,随后再灌注28天。在再灌注时(将PBS注射到对照小鼠),使小鼠接受单次重组因子1或因子2(各10 $\mu$ g)的皮下注射。随后使用Alzet微型泵(10 $\mu$ g/天)进行重组因子1或因子2的7天连续的皮下输注。用PBS输注对照小鼠。28天内每天检查小鼠以评估梗死后的存活。N=29PBS处理的小鼠;N=25因子1处理的小鼠;N=15因子2处理的小鼠。

#### [0211] 实施例7:

[0212] 用[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=Blastp&BLAST\\_PROGRAMS=Blastp&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=Blastp&BLAST_PROGRAMS=Blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)上的BLASTP算法鉴定由人C190rf10编码的蛋白质的同源物序列。SEQ ID NO:2作为矩阵。

[0213] 所用的参数为默认参数:11延伸:1,组成调整=条件组成得分矩阵与非冗余蛋白质序列(nr)的数据库一起调整。

[0214] 从鉴定的序列中选择来自不同脊椎动物的实例,所述脊椎动物主要是哺乳动物,但是也从两栖动物、鸟类和鱼类中选择一个序列。各序列列于表1中。

[0215] 使用<http://www.ebi.ac.uk/Tools/clustalw2/index.html>上的CLUSTALW2算法对所选择的氨基酸序列进行比对。所使用的各参数为默认参数:对齐类型=慢,蛋白质权重矩阵=Gonnet,空位开放=10,空位延伸=0.20,空位距离=5,无结束空位=无,输出选项:格式=Alnw/数字,顺序=对齐。在图6中显示所得到的所有序列的多重比对并在图7中显示哺乳动物序列的多重比对。

[0216] 表1:用BLASTP算法鉴定的人因子1的同源氨基酸序列。人序列作为矩阵。

[0217]	登录号	来源	总分	查询覆盖	最大鉴定率	SEQ ID NO
	NP_061980.1	智人 <i>Homo sapiens</i>	359	100%	100%	2
	EGH29500.1	猕猴 <i>Macaca mulatta</i>	357	100%	99%	12
	NP_543027.1	小鼠 <i>Mus musculus</i>	280	100%	84%	13
	NP_001001164.1	牛 <i>Bos taurus</i>	305	100%	89%	14
	XP_003421710.1	非洲象 <i>Loxodonta africana</i>	296	100%	87%	15
	EBH15128.1	裸鼯鼠 <i>Heterocephalus glaber</i>	294	100%	87%	16
	AES10565.1	雪貂 <i>Mustela putorius furo</i>	289	100%	91%	17
	NP_001006342.1	原鸡 <i>Gallus gallus</i>	209	100%	60%	18
	XP_003440079.1	罗非鱼 <i>Oreochromis niloticus</i>	185	100%	53%	19
	NP_001093679.1	热带爪蟾 <i>Xenopus (Silurana) tropicalis</i>	182	91%	55%	20

[0218] 实施例8:

[0219] 用[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST\\_PROGRAMS=blastp&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)上的BLASTP算法检索了人C19orf63剪接变体HSS1编码的蛋白质的同源序列。SEQ ID NO:4作为矩阵。

[0220] 所用的参数为默认参数:11延伸:1,组成调整=条件组成得分矩阵与非冗余蛋白质序列(nr)的数据库一起调整。

[0221] 从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物、鸟类和鱼类中选择一个序列。各序列列于表2中。

[0222] 使用<http://www.ebi.ac.uk/Tools/clustalw2/index.html>上的CLUSTALW2算法对所选择的氨基酸序列进行比对。所使用的各参数为默认参数:对齐类型=慢,蛋白质权重矩阵=Gonnet,空位开放=10,空位延伸=0.20,空位距离=5,无结束空位=无,输出选项:格式=Alnw/数字,顺序=对齐。在图8中显示所得到的所有序列的多重比对并在图9中显示哺乳动物序列的多重比对。

[0223] 表2:BLASTP算法鉴定的分泌形式的人因子2的同源氨基酸序列。人序列作为矩阵。



[0224]

登记号	来源	总分	查询覆盖	最大鉴定率	SEQ ID NO:
NP_778233.4	智人 <i>Homo sapiens</i>	511	100%	100%	4
AFE66256.1	猕猴 <i>Macaca mulatta</i>	503	100%	98%	21
XP_003465531.1	豚鼠 <i>Cavia porcellus</i>	434	100%	91%	22
XP_863321.1	狗 <i>Canis lupus familiaris</i>	427	100%	91%	23
AAV30544.1	小鼠 <i>Mus musculus</i>	401	93%	86%	24
XP_001494158.1	马 <i>Equus caballus</i>	429	100%	95%	25
AAI41719.1	非洲爪蟾 <i>Xenopus laevis</i>	275	89%	61%	26
NP_001157390.1	斑马鱼 <i>Danio rerio</i>	233	80%	55%	27

[0225] 实施例9:

[0226] 用[http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST\\_PROGRAMS=blastp&PAGE\\_TYPE=BlastSearch&SHOW\\_DEFAULTS=on&LINK\\_LOC=blasthome](http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome)上的BLASTP算法检索了由人C19orf63剪接变体HSMI编码的蛋白质的同源序列。SEQ ID NO: 5作为矩阵。

[0227] 所用的参数为默认参数:11延伸:1,组成调整=条件组成得分矩阵与非冗余蛋白质序列(nr)的数据库一起调整。

[0228] 从鉴定的序列中选择来自不同脊椎动物物种的实例,所述脊椎动物物种主要是哺乳动物,但是也从两栖动物、鸟类和鱼类中选择一个序列。各序列列于表3中。

[0229] 使用<http://www.ebi.ac.uk/Tools/clustalw2/index.html>上的CLUSTALW2算法对所选择的氨基酸序列进行比对。所使用的各参数为默认参数:对齐类型=慢,蛋白质权重矩阵=Gonnet,空位开放=10,空位延伸=0.20,空位距离=5,无结束空位=无,输出选项:格式=Alnw/数字,顺序=对齐。在图10中显示所得到的所有序列的多重比对并在图11中显示哺乳动物序列的多重比对。

[0230] 表3:用BLASTP算法鉴定的跨膜形式的人因子2的同源氨基酸序列。人序列作为矩阵。

登记号	来源	总	查询	最大	SEQ ID
-----	----	---	----	----	--------

[0231]

		分	覆盖	鉴定率	NO:
NP_996261.1	智人 <i>Homo sapiens</i>	527	100%	100%	5
XP_001173798.1	黑猩猩 <i>Pan troglodytes</i>	520	100%	99%	28
AAH20179.1	小鼠 <i>Mus musculus</i>	421	93%	87%	29
EH05689.1	裸鼯鼠 <i>Heterocephalus glaber</i>	426	93%	92%	30
XP_003406885.1	非洲象 <i>Loxodonta africana</i>	380	93%	92%	31
NP_988902.1	热带爪蟾 <i>Xenopus (Silurana) tropicalis</i>	310	89%	64%	32
NP_001157390.1	斑马鱼 <i>Danio rerio</i>	223	77%	54%	33

[0232] 实施例10:

[0233] 使人冠状动脉内皮细胞 (HCAEC, 来自Provitro) 于T75烧瓶中在补充了10% FCS (Biochrom) 的EGM-2培养基 (Lonza) 上生长。使用来自通道3-6的细胞。在用多种试剂刺激之前, 将细胞在含有2% FCS的MCDB131 (LifeTechnologies) 中培养过夜。将HCAEC接种在96孔板 (每孔 $5 \times 10^3$ 细胞) 上, 然后在存在或不存在不同浓度的兔抗因子1抗体、兔抗因子2抗体或对照IgG的情况下, 用重组人因子1、重组人因子2或VEGF (阳性对照) 刺激16小时。通过Eurogentec产生抗体并针对包含在人因子1 (CTIWRPQGKSYLYFTQ, SEQ ID NO:38, 即SEQ ID NO:1的氨基酸61至76) 或因子2 (CEQAQKAKNPQEQKSF; SEQ ID NO:39, 即SEQ ID NO:3的氨基酸181至195, 加上N-端Cys) 中的多肽而增加。用比色BrdU掺入免疫测定法 (Roche) 测量细胞增殖。数据示于图12中 (图A因子1; 图B因子2; 数据为来自3-6次实验的平均值 $\pm$ SEM)。

[0234] 以下内容对应于母案申请中的原始权利要求书, 现作为说明书的一部分并入此处:

[0235] 1. 蛋白质, 其包含根据SEQ ID NO:1的氨基酸序列或者其片段或变体, 所述片段或变体与SEQ ID NO:1具有至少80%序列同一性, 所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0236] 2. 蛋白质, 其包含根据SEQ ID NO:3的氨基酸序列或者其片段或变体, 所述片段或变体与SEQ ID NO:3具有至少80%序列同一性, 所述蛋白质用于增强非转化组织或非转化细胞的增殖和/或愈合。

[0237] 3. 根据项2所述的蛋白质, 其中所述氨基酸序列选自SEQ ID NO:4或SEQ ID NO:5。

[0238] 4. 根据项1至4中任一项所述的蛋白质, 其中所述非转化组织或所述非转化细胞是健康的、患病的或受损的。

[0239] 5. 根据项1至5中任一项所述的蛋白质, 其中所述细胞或所述组织

[0240] (i) 是肌肉组织细胞、结缔组织细胞、上皮组织细胞或神经组织细胞或者肌肉组织、结缔组织、上皮组织或神经组织; 和/或

[0241] (ii) 属于或来源于循环系统、消化系统、内分泌系统、排泄系统、免疫系统、皮肤系统、肌肉系统、神经系统、生殖系统、呼吸系统或骨骼系统; 和/或

[0242] (iii) 属于或来源于心脏、皮肤、骨、软骨、血管、食道、胃、肠、腺、肝、肾、肺、脑和脾。

[0243] 6. 根据项5中任一项所述的蛋白质,其中所述受损由遗传性/继承性疾病或者起因于局部缺血、再灌注损伤、炎症、感染、外伤、机械性超负荷、中毒或手术造成的获得性疾病引起,优选地,引起的疾病选自心肌梗死、心绞痛和心力衰竭。

[0244] 7. 根据项5或6中任一项所述的蛋白质,其中所述疾病与萎缩、发育不全、炎症、损伤或创伤有关。

[0245] 8. 根据项6或7所述的蛋白质,其中所述疾病是原发性心肌病,优选继承性心肌病和由自发突变引起的心肌病;或者获得性心肌病,优选由动脉粥样硬化或其他冠状动脉疾病引起的缺血性心肌病、由心肌的感染或中毒引起的心肌病、由肺动脉高血压和/或动脉高血压引起的高血压性心脏病以及心脏瓣膜的疾病。

[0246] 9. 核酸,其编码根据项1至8所述的蛋白质,所述核酸用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0247] 10. 载体,其包含根据项9所述的核酸,所述载体用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0248] 11. 根据项10所述的载体,其中所述载体选自:质粒载体;粘粒载体;噬菌体载体,如λ噬菌体载体、丝状噬菌体载体;病毒载体,优选腺病毒载体、腺相关病毒(AAV)载体、甲病毒载体、疱疹病毒载体、麻疹病毒载体、痘病毒载体、疱疹性口炎病毒载体、逆转录病毒载体和慢病毒载体;病毒样颗粒;以及细菌芽孢。

[0249] 12. 药物组合物,其包含根据项1至8中任一项所述的蛋白质、根据项9所述的核酸或根据项10或11所述的载体,并且任选地包含合适的药物赋形剂,所述药物组合物用于增强非转化组织或非转化细胞的增殖和/或使其愈合和/或抑制其凋亡。

[0250] 13. 根据项12所述的药物组合物,其中所述药物组合物通过经口、静脉内、粘膜内、动脉内、肌内或冠状动脉内途径施用。

[0251] 14. 根据项12或13所述的药物组合物,其中所述药物组合物在再灌注疗法之前、与其同时或在其之后施用。

[0252] 15. 根据项14所述的药物组合物,其中所述施用通过一次或更多次快速推注和/或输注进行。

[0253] 16. 因子1和/或因子2蛋白的抑制剂,其用于医学用途。

[0254] 17. 根据项16所述的因子1和/或因子2的抑制剂,其中所述医学用途是治疗或预防其中血管发生有助于疾病发生或进展的疾病。

[0255] 18. 根据项17所述的抑制剂,其中因子1蛋白包含根据SEQ ID NO:1的氨基酸序列或其变体,因子2蛋白包含根据SEQ ID NO:3的氨基酸序列或其变体,所述变体与SEQ ID NO:1或3具有至少80%序列同一性。

[0256] 19. 根据项17或18所述的抑制剂,其选自:包含SEQ ID NO:1或3或其变体的抑制性片段的蛋白质,所述变体与SEQ ID NO:1或3具有至少80%序列同一性;与根据SEQ ID NO:1或3的氨基酸序列或其变体特异性结合的配体,所述变体与SEQ ID NO:1或3具有至少80%序列同一性;以及抑制或阻止编码蛋白质的mRNA转录和/或翻译的核酸,所述蛋白质包含根据SEQ ID NO:1或3的氨基酸序列或其变体,所述变体与SEQ ID NO:1或3具有至少80%序列

同一性。

[0257] 20. 根据项19所述的抑制剂,其中所述配体选自抗体或其片段、抗体样蛋白或肽模拟物。

[0258] 21. 核酸,其编码根据项17至20所述的抑制剂,所述核酸用于治疗或预防其中血管发生有助于疾病发生或进展的疾病。

[0259] 22. 载体,其包含根据项22所述的核酸,所述载体用于治疗或预防其中血管发生有助于疾病发生或进展的疾病。

[0260] 23. 药物组合物,其包含根据项17至20中任一项所述的抑制剂、根据项21所述的核酸或根据项22所述的载体,并且任选地包含合适的药物赋形剂,所述药物组合物用于治疗或预防其中血管发生有助于疾病发生或进展的疾病。

[0261] 24. 根据项17至20中任一项所述的抑制剂、根据项21所述的核酸、根据项22所述的载体或根据项23所述的药物组合物,其中所述其中血管发生有助于疾病发生或进展的疾病是增殖性疾病。

[0262] 25. 根据项17至20中任一项所述的抑制剂、根据项21所述的核酸、根据项22所述的载体或根据项23所述的药物组合物,其中所述增殖性疾病选自:良性肿瘤、恶性肿瘤、类风湿性关节炎、银屑病、眼部血管发生疾病、奥斯勒-韦伯综合征、斑块新血管形成、移植物和血管成形术后狭窄、毛细血管扩张、血友病性关节、血管纤维瘤、伤口肉芽、肠粘连、动脉粥样硬化、硬皮病、肥大性瘢痕、猫抓病和溃疡,特别是黄斑变性、肾上腺皮质癌、膀胱癌、骨癌、脑癌、乳腺癌、宫颈癌、结肠癌、结直肠癌、子宫内膜癌、食道癌、眼癌、胆囊癌、胃癌、头颈癌、喉癌、肝癌、肺癌、黑素瘤、骨髓增生性病症、颈癌、非黑素瘤皮肤癌、卵巢癌、前列腺癌、良性前列腺增生、胰腺癌、直肠癌和睾丸癌。

[0263] 26. 与因子1或因子2特异性结合的配体,其抑制因子1或因子2的促血管发生活性。

[0264] 27. 根据项26所述的配体,所述配体特异性结合包含在SEQ ID NO:1的61至76位氨基酸中或由SEQ ID NO:1的61至76位氨基酸组成的人因子1蛋白的表位,或者特异性结合对应于该表位的来自另一因子1蛋白的区域。

[0265] 28. 根据项26所述的配体,所述配体特异性结合包含在SEQ ID NO:3的181至195位氨基酸中或由SEQ ID NO:3的181至195位氨基酸组成的人因子2蛋白的表位,或者特异性结合对应于该表位的来自另一因子2蛋白的区域。

## 序列表

<110> Medizinische Hochschule Hannover

<120> 用于治疗或预防疾病的因子1蛋白、因子2蛋白及其抑制剂

<130> 750-2 PCT

<150> EP13151593

<151> 2013-01-17

<160> 39

<170> PatentIn version 3.5

<210> 1

<211> 142

<212> PRT

<213> Homo sapiens

<400> 1

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

Val His Ser Phe Ser His Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys  
20 25 30

Met Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met  
35 40 45

Ser Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp  
50 55 60

Arg Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu  
65 70 75 80

Val Arg Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala Ala  
85 90 95

Phe Glu Arg Glu Ser Asp Val Pro Leu Lys Thr Glu Glu Phe Glu Val  
100 105 110

Thr Lys Thr Ala Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu  
115 120 125

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

<210> 2

<211> 173

<212> PRT

<213> Homo sapiens

<400> 2

Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Gly Ala Ser Leu Trp  
1 5 10 15

Ala Ala Leu Leu Leu Gly Ala Val Ala Leu Arg Pro Ala Glu Ala Val  
20 25 30

Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val Val

[0001]

	35	40	45
	His Ser Phe Ser His Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys Met 50 55 60		
	Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met Ser 65 70 75 80		
	Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp Arg 85 90 95		
	Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu Val 100 105 110		
	Arg Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala Ala Phe 115 120 125		
	Glu Arg Glu Ser Asp Val Pro Leu Lys Thr Glu Glu Phe Glu Val Thr 130 135 140		
	Lys Thr Ala Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu Ser 145 150 155 160		
	Lys Leu Val Ile Val Ala Lys Ala Ser Arg Thr Glu Leu 165 170		
[0002]	<210> 3		
	<211> 200		
	<212> PRT		
	<213> Homo sapiens		
	<400> 3		
	Ser Gly Cys Arg Ala Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly 1 5 10 15		
	Arg Glu Gly Glu Ala Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser 20 25 30		
	Phe Glu Ile Asp Asp Ser Ala Asn Phe Arg Lys Arg Gly Ser Leu Leu 35 40 45		
	Trp Asn Gln Gln Asp Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser 50 55 60		
	Glu Glu Glu Arg Gly Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu 65 70 75 80		
	Tyr Arg Val Arg Ile Pro Arg Arg Pro Gly Ala Leu Asp Gly Leu Glu 85 90 95		
	Ala Gly Gly Tyr Val Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu 100 105 110		
	Ser His Leu Ser Asp Gln Leu Thr Leu His Val Asp Val Ala Gly Asn 115 120 125		
	Val Val Gly Val Ser Val Val Thr His Pro Gly Gly Cys Arg Gly His		

130	135	140
Glu Val Glu Asp Val Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu 145 150 155 160		
Gln Pro Pro Thr Thr Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu 165 170 175		
Arg Leu Glu Met Glu Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln 180 185 190		
Lys Ser Phe Phe Ala Lys Tyr Trp 195 200		
<210> 4 <211> 254 <212> PRT <213> Homo sapiens		
<400> 4		
Met Ala Ala Ala Ser Ala Gly Ala Thr Arg Leu Leu Leu Leu Leu Leu 1 5 10 15		
Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Ala 20 25 30		
Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Ala 35 40 45		
[0003] Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp 50 55 60		
Ser Ala Asn Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp 65 70 75 80		
Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly 85 90 95		
Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Ile 100 105 110		
Pro Arg Arg Pro Gly Ala Leu Asp Gly Leu Glu Ala Gly Gly Tyr Val 115 120 125		
Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp 130 135 140		
Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser 145 150 155 160		
Val Val Thr His Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val 165 170 175		
Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Gln Pro Pro Thr Thr 180 185 190		
Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu		

	195	200	205
	Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala 210	215	220
	Lys Tyr Trp His Ile Ile Leu Gly Gly Ala Val Leu Leu Thr Ala Leu 225	230	235 240
	Arg Pro Ala Ala Pro Gly Pro Ala Pro Pro Pro Gln Glu Ala 245	250	
<210>	5		
<211>	262		
<212>	PRT		
<213>	Homo sapiens		
<400>	5		
	Met Ala Ala Ala Ser Ala Gly Ala Thr Arg Leu Leu Leu Leu Leu Leu 1	5 10 15	
	Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Ala 20	25 30	
	Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Ala 35	40 45	
	Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp 50	55 60	
[0004]	Ser Ala Asn Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp 65	70 75 80	
	Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly 85	90 95	
	Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Ile 100	105 110	
	Pro Arg Arg Pro Gly Ala Leu Asp Gly Leu Glu Ala Gly Gly Tyr Val 115	120 125	
	Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp 130	135 140	
	Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser 145	150 155 160	
	Val Val Thr His Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val 165	170 175	
	Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Gln Pro Pro Thr Thr 180	185 190	
	Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu 195	200 205	
	Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala		



	210	215	220	
	Lys Tyr Trp Met Tyr Ile Ile Pro Val Val Leu Phe Leu Met Met Ser			
	225	230	235	240
	Gly Ala Pro Asp Thr Gly Gly Gln Gly Gly Gly Gly Gly Gly Gly			
		245	250	255
	Gly Gly Gly Ser Gly Arg			
		260		
	<210> 6			
	<211> 1067			
	<212> DNA			
	<213> Homo sapiens			
	<400> 6			
	ggcggacgct ccacgtgtcc ctcgccgcgc cccgtctacc cgcccctgcc ctgaggaccc		60	
	tagtccaaca tggcggcgcc cagcggaggg tggaacggcg tcggcgcgag cttgtgggccc		120	
	gcgctgtccc taggggcccgt ggcgctgagg ccggcggagg cgggtgtccga gccacacacg		180	
	gtggcggttg acgtgcggcc cggcggcgtc gtgcattcct tctcccataa cgtgggccc		240	
	ggggacaaat atacgtgtat gttcacttac gcctctcaag gagggaccaa tgagcaatgg		300	
	cagatgagtc tggggaccag cgaagaccac cagcacttca cctgcacat ctggaggccc		360	
	cagggggaagt cctatctgta cttcacacag ttcaaggcag aggtgcgggg cgctgagatt		420	
	gagtacgcca tggcctactc taaagccgca ttgaaaggg aaagtgatgt cccctctgaaa		480	
[0005]	actgaggaat ttgaagtgac caaaacagca gtggctcaca ggcccggggc attcaaagct		540	
	gagctgtcca agctgggtgat tgtggccaag gcacgcgcga ctgagctgtg accagcagcc		600	
	ctgttgccgg tggcaccttc tcctctccgg tgaagctgaa ggggcctgtg tccctgaaag		660	
	ggccagcaca tcaactggtt tctaggaggg actcttaagt ttctacctg ggctgacgtt		720	
	gcctgtgccg gaggggcttg cagggtggct gaagccctgg ggccagagaac agagggtcca		780	
	ggccctctct ggctcccaac agcttctcag ttccacttc ctgctgagct cttctggact		840	
	caggatcgca gatccggggc acaagagggg tggggaacat gggggctatg ctggggaaaag		900	
	cagccatgct cccccgacc tccagccgag catccttcat gagcctgcag aactgtttc		960	
	ctatgtttac ccaggggacc tcctttcaga tgaactggga agagatgaaa tgttttttca		1020	
	tatttaaata aataagaaca ttaaaaagca aaaaaaaaaa aaaaaaa		1067	
	<210> 7			
	<211> 913			
	<212> DNA			
	<213> Mus musculus			
	<400> 7			
	gagcgccctgc gcattgcccc ggaagcaaga tggcagcccc cagcggaggc ttctggactg		60	
	cgggtggctct ggcggccgca gcctgaaat tggcgcgcgc tgtgtccgag cccaccaccg		120	
	tgccatttga cgtgaggccc ggaggggicg tgcattcgtt ctcccaggac gtaggacccg		180	
	ggaacaagt ttacatgtaca ttacctacg cttccaagg agggaccaac gagcaatggc		240	
	agatgagcct ggggacaagt gaagacagcc agcactttac ctgtaccatc tggaggcccc		300	
	aggggaaatc ctacctctac ttcacacagt tcaaggctga gttagcaggt gctgagatcg		360	

	agtatgccat ggcctactcc aaagccgcat ttgagagaga gagtgaatgc cccctgaaaa	420
	gtgaggagtt tgaagtgacc aagacagcag tgtctcacag gcctggggcc ttcaaagctg	480
	agctctccaa gctgggtgac gtagccaagg cggcacgctc ggagctgtga ccctcgctg	540
	tcaagggcct tcatgtccac gttcctcagg cacactgacc gggactactt gtctagggca	600
	ctggttccca taggagctgc cctgccctgc acaggtcaca ctgtgtcact ccgcagaact	660
	ctctgagccc ggtcacctgt ttgccaggg aagatgcagg gcatgtgcgg ggggtgggatg	720
	gaaggacttc ctggctttcc tgaagtcagg atgtggtgtg gtttccctc tgagccacag	780
	atgagtgtcc ccatcccagg accactttct aaccccatcc agggcagctc cactcagaag	840
	gatgggaaag gatagaaaaa ataaataaat aagtagccac cttagtggtg gctctgtggg	900
	gtcaggactc aga	913
	<210> 8	
	<211> 2048	
	<212> DNA	
	<213> Homo sapiens	
	<400> 8	
	ttcctcccgg cgtgctccgc ggctcttggc tcacagccgt cccttcgctg gtgggaagaa	60
	gccgagatgg cggcagccag cgctggggca acccggtcgc tctgtctctt gctgatggcg	120
	gtagcagcgc ccagtcgagc ccggggcagc ggctgccggg ccgggactgg tgcgcgaggg	180
	gctggggcgg aaggctcaga gggcgaggcc tgtggcacgg tggggctgct gctggagcac	240
	tcatittaga tcgatgacag tgccaacttc cgggaagcggg gctcactgct ctggaaccag	300
	caggatggta ccttgtccct gtcacagcgg cagctcagcg aggaggagcg gggccgactc	360
	cgggatgtgg cagccctgaa tggcctgtac cgggtccgga tcccaaggcg acccggggcc	420
	ctggatggcc tggaaagtgg tggctatgtc tctctctttg tccctgcgtg ctccctgggtg	480
	gagtcgcacc tgcggacca gctgacctg cactgtgatg tggccggcaa cgtgggtggc	540
	gtgtcggtag tgacgcacc cgggggctgc cggggccatg aggtggagga cgtggacctg	600
	gagctgttca acacctcgtt gcagctgcag ccggccacca cagccccagg ccctgagacg	660
	gccccttca ttgagccct ggagatggaa caggcccaga agggccaagaa cccccaggag	720
	cagaagtctt tcttcgcaa atactggcac atcatcctgg ggggggcccgt gttgtcaca	780
	gccctgcgtc ctgctgcgc agggcccgcg ccaccgccac aggaggcctg atggatgtac	840
	atcatccccg tcttctgtt cctcatgatg tcaggagcgc cagacaccgg gggccagggt	900
	gggggtgggg gtgggggtgg tgggtgggggt agtggccggg gagggcccag gctggtcagc	960
	gtcccgtctt gcacaccag gggcctccct ttctgctgga gtcccctgtg tctcagcca	1020
	tcccaagaag ggtttgcctg tccctctttt cccccctcc cagcaggcca cctgggccag	1080
	ccccctgtcc tctgccttct gctggcagag gacagctgg actggggcct ttggcacagc	1140
	agccggtgtc tctgcgcgc gccctcccca tggcccatg cagccccagg ggcttccccc	1200
	ctgccatgg agtagagccc gagatcctgg ccactatgcc agttctgacc tcgcatcccc	1260
	ctaccccag cccatgcagt ctgggaacat gccgccttct ctccagctc tgtgcctttg	1320
	ttccaggtag tctcaccctc ctgtccctgg ctgggctagg tggctctgtc caggctcctg	1380
	cagcgcctcc ctactttga cactggacta ggatgcagcc tcccttctgt gtccccttga	1440

[0006]

gggtaccctg ggccccctca tcaggggcag aggcataaaa gagtcggggc tggatggccg	1500
ggggcttctg ggcccgacgc ctatgacagc ccttggggtc gtggtttgac atttgtctgc	1560
ctgggtgcaa caaggaatcc ttgcctttaa ggtgacaggc cctccacagg ctccagact	1620
tgaaggaaaa ggtttaagaa agaaaacaaa accaacagtt agtggagtca aagcccagac	1680
actgtaaata gaacccccct caccaccccc cgccgcccag catcctacct ggactgcggt	1740
gctacgaggg cctgegggee tttgtgtgt gccacccctc ctgtaagtct atttaaaaac	1800
atcgacgata cattgaaatg tgtgaacgtt ttgaaaagct acagcttcca gcagccaaaa	1860
gcaactgttg ttttggcaag acggctctga tgtacaagct tgattgaaat tcaactgtca	1920
cttgatacgt tattcagaaa cccaaggaat ggctgtcccc atcctcatgt ggctgtgtgg	1980
agctcagctg tgttgtgtgg cagtttatta aactgtcccc cagatcgaca cgcaaaaaaa	2040
aaaaaaaa	2048
<210> 9	
<211> 1960	
<212> DNA	
<213> Homo sapiens	
<400> 9	
ttcctcccgg cgtgctccgc ggctcttggc tcacagccgt cccttcgctg gtgggaagaa	60
gccgagatgg cggcagccag cgctggggca acccggtcgc tcctgctctt gctgatggcg	120
gtagcagcgc ccagtcgagc ccggggcagc ggctgccggg ccgggactgg tgcgcgaggg	180
gctggggcgg aaggctgaga gggcgaggcc tgtggcacgg tggggctgct gctggagcac	240
tcatttgaga tcgatgacag tgccaacttc cggaagcggg gctcactgct ctggaaccag	300
caggatggta ccttgtccct gtcacagcgg cagctcagcg agggaggagcg gggccgactc	360
cgggatgtgg cagccctgaa tggcctgtac cgggtccgga tcccaaggcg acccggggcc	420
ctggatggcc tggaagctgg tggctatgtc tcctcctttg tccttcgctg ctccctggtg	480
gagtcgcacc tgtcggacca gctgacctg cacgtggatg tggccggcaa cgtggtgggc	540
gtgtcggtag tgacgcacc cgggggctgc cggggccatg aggtggagga cgtggacctg	600
gagctgttca acacctcggg gcagctgcag ccgcccacca cagcccagg cctgagacg	660
ggcgcttca ttgagcgcct ggagatggaa caggcccaga agggcaagaa cccccaggag	720
cagaagtccct tcttcgcaa atactggatg tacatcattc ccgtcgtcct gttcctcatg	780
atgcaggag cgcagacac cgggggcccag ggtgggggtg ggggtggggg tgggtgtggg	840
ggtagtggcc ggtgagggcc caggctggtc agcgtcccgt ctgacacac caggggcctc	900
cctttctgct ggagtcacct gtgtcctcag ccatcccaag aagggtttgc tggtccttc	960
tttccccccg tcccacagg ccacctgggc cagccccttg tcctctgcct tctgctggca	1020
gaggagcagc tggactgggg cctttggcac agcagccggt gtctcctgcg cccgcctccc	1080
ccatggcccc atgcagcccc aggggcttcc cccctgccca tggagtagag cccgagatcc	1140
tggccactat gccagttctg acctgcacac cccctacccc gageccatgc agtctgggaa	1200
catgccgcct tctctcagc ctctgtgcct ttgttcagg tggctcacc ctctgtccc	1260
tggctgggct aggtggctct gtccaggctc ctgcagcgcc cccctcactt tgacactgga	1320
ctaggatgca gcctccctc tgtgtccctc tgagggtacc ctgggtcccc tcatcagggg	1380

[0007]

cagaggcatg aaagagtcgg ggctggatgg ccgggggctt ctgggcccga cgcctagtgc	1440
agccctctggg gtcgtgggtt gacatttgtc tgcttggtgc aaacaaggaa tccttgccctt	1500
taagggtgaca ggccctccac aggcctccag acttgaagga aaagggttaa gaaagaaaac	1560
aaaaccaaca gttagtggag tcaaaagcca gacactgtaa atagaacccc ctccaccacc	1620
ccccgccgcc cagcatccta cctggactgc ggtgctacga gggcctgcgg gcctttgctg	1680
tgtgccaccc tccttgtaag tctattttaa aacatcgacg atacattgaa atgtgtgaac	1740
gttttgaaaa gctacagctt ccagcagcca aaagcaactg ttgttttggc aagacggctc	1800
tgatgtacaa gcttgattga aattcactgc tcaattgata cgttattcag aaaccaagg	1860
aatggctgtc cccatctca tgtggctgtg tggagctcag ctgtgtgtg tggcagttta	1920
ttaaactgtc cccagatcg acacgcaaaa aaaaaaaaaa	1960
<210> 10	
<211> 1627	
<212> DNA	
<213> Mus musculus	
<400> 10	
gggctgtatg gctctcggtt tttctcaacg ctcccgtatg gtggccgcgg gtgccgggggt	60
gacccggctg ctagtgcctt tgctgatggg agccgcggct cctagcagag cccgaggcag	120
cggctgcggg gtcggggcct ccgcgcgtgg gaccggggcc gatggccgtg aagctgaggg	180
ctgtggcacc gtggccttgc tgctggagca ttcatattgag ctcggtgatg gagecaactt	240
ccagaagcga ggcttgctgc tctggaacca gcaggatggc accctgtcgg caacacagcg	300
acagctcagt gaggaggagc gtggccgact ccgggatgtg gctgctgtca atggcctcta	360
cagggtccgg gtcceagagg ggcttgggac acttgatggg tcagaagctg gcggccatgt	420
gtcttccttc gtccacagct gctccctggg ggagtcgcac ctttcggacc agctgacctt	480
gcacgtggat gtggctggca acgtgggtgg cctgtctgtg gtgggtgtacc ctgggggctg	540
ccggggctcc gagggtggaag atgaggacct ggagctgttc aatacatctg tgcagctgcg	600
gcctcccagc actgctccag gccccgagac tgcagccttc attagagccc tggagatgga	660
gcaggcccag aaggccaaga acccacagga gcagaagtct ttctttgcca aatactggca	720
cctcatcctg gggggggcgg tggtgtcac agccctacgt cctgctgccc cagggcctgc	780
accagcgcca acggaggcct aatggatgta catcattcca gtgtgtgtgt tcctcatgat	840
gtcgggagcg ccggacgtg ggggccaggg cggcgggtgg ggcgggggca gcagccggtg	900
agcagctgtg ccacctagag cccccccag agccagccca agaaggagtt cctgtcccca	960
catttccta ttgcatgaat atggaaggct gtcccttcag tgagccctct ggccttctctg	1020
taagcccttc tttctgtccc tgagcctctc tctcatcctg ttgactgaga gcttgggtgg	1080
acctccctgt agccagctca ctgcaactgt gtcccacat gtggcactgt gctcctctgt	1140
ctgctaaaca cccaccagcc tgccccaccc caccacacca tacactttgg gaacttgcca	1200
agctctctcc agcctctgtg cctttgcctt gcaggccccg tgcgcccctc actgtcactc	1260
tccagccctt tgccaaggat ctgtggccca gaggcctctg ctcttagtgg ctaggtcagc	1320
ctccagccca ctgtccaggt ggcatgtgtt cttctttgcc cccctctctg gtgcccaga	1380
ataccatggg gacctaccac taccctttct gcctttggat gtcatagcct ggatctgtca	1440

[0008]

	ccaggagagg attgtgggcc tccacgttag tctgtgaatg cacacttcga gtgacttgtg	1500
	tgcaggtttt gagagccggt tttagactag ctgctcgaca gctgctggca tggccgtgct	1560
	cttgacatg cgcgcgtgtg ggcatgggga ttgctgtgca gcctcagctg tgttgtgtgg	1620
	ctgctga	1627
	<210> 11	
	<211> 1855	
	<212> DNA	
	<213> Mus musculus	
	<400> 11	
	gagaaatggc cgcctgact gcgggagcag gcggacgggc gctagtgcgc aggcgcggcg	60
	tgcggcgacg gcgcgtgagc ctgagatga acctgtgtt tccatagcggg ctgtatggct	120
	ctcggttttt ctcaacgctc ccgtatggtg gccgcgggtg ccgggggtgac ccggctgcta	180
	gtgtctttgc tgaatgtagc cgcggctcct agcagagccc gaggcagcgg ctgccgggtc	240
	ggggcctccg cgcgtgggac cggggccgat ggccgtgaag ctgagggctg tggcaccgtg	300
	gctttgtctg tggagcattc atttagctc ggtgatggag ccaacttcca gaagcaggc	360
	ttgtgtctct ggaaccagca ggatggcacc ctgtcgga caccagcaca gctcagttag	420
	gaggagcgtg gccgactccg ggatgtggct gctgtcaatg gcctctacag ggtccgggtc	480
	ccgagggcgc ctgggacact tgatggttca gaagctggcg gccatgtgtc ttccctctgc	540
	ccagcgtgct ccttggttga gtgcacattc tcggaccagc tgaccttgca cgtggatgtg	600
	gttggaacg tgggtggcct gtctgtggtg gtgtacctg ggggtgccc gggctccgag	660
[0009]	gtggaagatg aggacctgga gctgttcaat acatctgtgc agctcgggcc tcccagcact	720
	gctccaggcc ccgagctgc agccttcatt gagcgctgg agatggagca ggcccagaag	780
	gccaagaacc caccaggaca gaagtcttcc ttgccaat actggatgta catcattcca	840
	gttgtgtgt tctcatgat gtcgggagcg ccggacgtg ggggccaggc cggcgggtgg	900
	ggcgggggca gcagccgggt agcagctgtg ccacctagag cccccccag agccagccca	960
	agaaggagtt cctgtcccca catttcccta ttgcatgaat atggaaggct gtcccttcag	1020
	tgagccctct ggcttccctg taagccctc ttctgtccc tgagcctctc tctcatcctg	1080
	ttgactgaga gcttgggtgg acctccctgt agccagctca ctgcaactgt gtcccacct	1140
	gtggcactgt gctcctctgt ctgtataaca cccaccagcc tgccccacc caccaccaca	1200
	tacactttgg gaacttgcca agctctctcc agcctctgtg cctttgccct gcaggccccg	1260
	tgcgccccct actgtcactc tccagccctt tgccaaggat ctgtggccca gaggcctctg	1320
	ctcttagtgg ctaggtcagc ctccagccca ctgtccaggt ggcatgctgt cttctttgcc	1380
	ccctctctg gtgccccaga ataccatggt gacctaccac tatectttct gcctttggat	1440
	gtcatagcct ggatctgtca ccaggagagg attgtgggcc tccacgttag tctgtgaatg	1500
	cacacttcga gtgacttgtg tgcaggtttt gagagccggt tttagactag ctgctcgaca	1560
	gctgctggca tggccgtgct cttgcacatg cgcgcgtgtg ggcatgggga ttgctgtgca	1620
	gcctcagctg tgttgtgtgg ctgctgatta aactgtcccc taaacagcca ctcttcagct	1680
	cacttccctg cttctgtgct tgtgaatagt cctgagttgc cgtgtgtgtt tgccgtgttt	1740
	atgtttgaat ggctttctta gggatatgta cagaggggtg cctgagcaga ttaaagttgc	1800

tgtgagcaag gacgccttcc gaactctggg aggaggtgg ttctgaccc tecta 1855

<210> 12  
<211> 173  
<212> PRT  
<213> Macaca mulatta

<400> 12

Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Gly Ala Ser Leu Trp  
1 5 10 15

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

Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val Val  
35 40 45

His Ser Phe Ser His Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys Met  
50 55 60

Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met Ser  
65 70 75 80

Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp Arg  
85 90 95

Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu Val  
100 105 110

[0010]

Arg Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala Ala Phe  
115 120 125

Glu Arg Glu Ser Asp Val Pro Leu Lys Thr Glu Glu Phe Glu Val Thr  
130 135 140

Lys Thr Ala Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu Ser  
145 150 155 160

Lys Leu Val Ile Val Ala Lys Ala Ser Arg Thr Glu Leu  
165 170

<210> 13  
<211> 166  
<212> PRT  
<213> Mus musculus

<400> 13

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

Ala Ala Leu Lys Leu Ala Ala Ala Val Ser Glu Pro Thr Thr Val Pro  
20 25 30

Phe Asp Val Arg Pro Gly Gly Val Val His Ser Phe Ser Gln Asp Val  
35 40 45

Gly Pro Gly Asn Lys Phe Thr Cys Thr Phe Thr Tyr Ala Ser Gln Gly

50	55	60
Gly Thr Asn Glu Gln Trp Gln Met Ser Leu Gly Thr Ser Glu Asp Ser 65 70 75 80		
Gln His Phe Thr Cys Thr Ile Trp Arg Pro Gln Gly Lys Ser Tyr Leu 85 90 95		
Tyr Phe Thr Gln Phe Lys Ala Glu Leu Arg Gly Ala Glu Ile Glu Tyr 100 105 110		
Ala Met Ala Tyr Ser Lys Ala Ala Phe Glu Arg Glu Ser Asp Val Pro 115 120 125		
Leu Lys Ser Glu Glu Phe Glu Val Thr Lys Thr Ala Val Ser His Arg 130 135 140		
Pro Gly Ala Phe Lys Ala Glu Leu Ser Lys Leu Val Ile Val Ala Lys 145 150 155 160		
Ala Ala Arg Ser Glu Leu 165		
<210> 14 <211> 174 <212> PRT <213> Bos taurus		
<400> 14		
[0011] Met Ala Ala Pro Ser Gly Arg Arg Asn Gly Ser Gly Gly Ala Asn Leu 1 5 10 15		
Trp Val Ser Leu Leu Leu Ala Ala Ala Ala Leu Arg Pro Val Glu Thr 20 25 30		
Val Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val 35 40 45		
Val His Ser Phe Ser Gln Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys 50 55 60		
Val Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Lys Trp Gln Met 65 70 75 80		
Ser Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp 85 90 95		
Arg Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu 100 105 110		
Val Arg Gly Ala Glu Ile Glu Tyr Gly Met Ala Tyr Ser Lys Ala Ala 115 120 125		
Phe Glu Lys Glu Ser Asp Val Pro Leu Lys Asn Glu Glu Phe Glu Val 130 135 140		
Thr Lys Thr Ala Val Phe His Arg Pro Gly Ala Phe Lys Ala Glu Leu		

145	150	155	160
Ser Lys Leu Val Ile Val Ala Lys Ala Thr Arg Ser Glu Leu			
	165	170	
<210> 15 <211> 174 <212> PRT <213> <i>Loxodonta africana</i> <400> 15			
Met Ala Ala Pro Arg Gly Arg Arg Asn Gly Ser Ala Gly Ala Ser Met			
1	5	10	15
Trp Gly Ala Leu Leu Leu Ala Ala Val Ala Leu Arg Ser Val Glu Ala			
	20	25	30
Val Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val			
	35	40	45
Val His Ser Phe Ser His Ser Ala Gly Pro Gly Asp Arg Phe Thr Cys			
	50	55	60
Thr Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met			
65	70	75	80
Ser Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp			
	85	90	95
Arg Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu			
	100	105	110
Val Arg Gly Ala Gln Ile Glu Tyr Gly Met Ala Tyr Ser Lys Ala Ala			
	115	120	125
Ser Glu Arg Glu Ser Asp Val Pro Leu Lys Asn Glu Glu Phe Glu Val			
	130	135	140
Thr Lys Thr Thr Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu			
145	150	155	160
Ser Lys Leu Val Ile Val Ala Lys Ala Ser His Ser Glu Leu			
	165	170	
<210> 16 <211> 173 <212> PRT <213> <i>Heterocephalus glaber</i> <400> 16			
Met Ala Ala Pro Arg Gly Asn Ser Asp Gly Cys Gly Gly Ala Trp Phe			
1	5	10	15
Ala Ala Leu Leu Leu Ala Ala Val Ala Leu Arg Pro Ala Glu Ala Val			
	20	25	30
Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val Val			
	35	40	45

[0012]



His Ser Phe Ser Gln Asn Val Gly Pro Gly Asp Lys Phe Thr Cys Thr  
 50 55 60  
 Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Gln Trp Gln Met Ser  
 65 70 75 80  
 Leu Gly Thr Ser Glu Asp His Gln His Phe Thr Cys Ile Ile Trp Arg  
 85 90 95  
 Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Lys Ala Glu Val  
 100 105 110  
 His Gly Ala Glu Ile Glu Tyr Ala Met Ala Tyr Ser Lys Ala Ala Phe  
 115 120 125  
 Glu Arg Glu Ser Asp Val Pro Leu Lys Asn Glu Glu Phe Glu Val Thr  
 130 135 140  
 Lys Ala Ala Val Ala His Arg Pro Gly Ala Phe Arg Ala Glu Leu Ser  
 145 150 155 160  
 Lys Leu Val Ile Val Ala Lys Glu Ala His Ser Glu Leu  
 165 170

<210> 17  
 <211> 173  
 <212> PRT  
 <213> *Mustela putorius furo*  
 <400> 17

[0013]

Met Ala Ala Pro Ser Glu Arg Arg Asn Gly Gly Gly Ala Ser Leu Trp  
 1 5 10 15  
 Ala Ala Leu Leu Leu Ala Ala Ala Ala Leu Arg Pro Ala Glu Ala Val  
 20 25 30  
 Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly Val Val  
 35 40 45  
 His Ser Phe Ser Gln Asn Val Gly Pro Gly Asp Lys Tyr Thr Cys Ala  
 50 55 60  
 Phe Thr Tyr Ala Ser Gln Gly Gly Thr Asn Glu Lys Trp Gln Met Ser  
 65 70 75 80  
 Leu Gly Ile Ser Glu Asp His Gln His Phe Thr Cys Thr Ile Trp Arg  
 85 90 95  
 Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln Phe Arg Ala Glu Val  
 100 105 110  
 Arg Gly Ala Glu Ile Glu Tyr Gly Met Ala Tyr Ser Lys Ala Ala Phe  
 115 120 125  
 Glu Arg Glu Ser Asp Val Pro Leu Lys Ser Glu Glu Phe Glu Val Thr  
 130 135 140

Lys Thr Ala Val Ser His Arg Pro Gly Ala Phe Lys Ala Glu Leu Ser  
145 150 155 160

Lys Leu Val Ile Val Ala Lys Ala Ser Arg Ser Glu Leu  
165 170

<210> 18  
<211> 167  
<212> PRT  
<213> Gallus gallus

<400> 18

Met Ala Ala Pro Cys Gly Arg Ser Ser Arg Trp Leu Trp Ala Ala Val  
1 5 10 15

Val Pro Ala Ala Val Leu Cys Leu Ala Val Arg Ala Ala Glu Glu Ala  
20 25 30

Ser Thr Ala Glu Phe Asp Val Arg Pro Gly Gly Glu Val His Phe Phe  
35 40 45

Ser Arg Ser Leu Gly Asp Tyr Thr Cys Thr Phe Thr Tyr Ser Ala Gln  
50 55 60

Gly Gly Thr Asn Glu Gln Trp Gln Met Asn Ile Gly Val Ser Glu Asp  
65 70 75 80

[0014]

Asn Leu Leu Phe Ser Cys Ser Val Trp Arg Pro Gln Gly Lys Ser Tyr  
85 90 95

Leu Phe Phe Thr Gln Phe Lys Ala Glu Val Lys Gly Ala Lys Ile Glu  
100 105 110

Tyr Ala Met Ala Tyr Ser Gln Ala Ala Val Gly Ala Gln Ser Asp Ile  
115 120 125

Pro Leu Lys Gln Glu Glu Phe Glu Ile Thr Glu Thr Thr Val Ser His  
130 135 140

Arg Glu Gly Lys Phe Arg Phe Glu Leu Ser Lys Leu Met Ile Val Ala  
145 150 155 160

Lys Thr Pro His Asp Glu Leu  
165

<210> 19  
<211> 169  
<212> PRT  
<213> Oreochromis niloticus

<400> 19

Met Ala Arg Gln Ser Asn Thr Cys Ala Gly Asn Leu Ala Phe Leu Phe  
1 5 10 15

Ala Leu Ala Leu Ile Ala Ala Arg Val Pro Ala Glu Ala Ser Glu Glu  
20 25 30

Gln Ala Lys Thr Val Glu Phe Asn Val Lys Pro Gly Gly Val Val His  
 35 40 45  
 Thr Phe Ser Glu Gly Ile Gly Glu Tyr Glu Cys Ser Phe Thr Tyr Ala  
 50 55 60  
 Ser Gln Gly Gly Thr Asn Glu Gln Trp Leu Met Ser Val Gly Leu Thr  
 65 70 75 80  
 Asp Asp Asn Arg Leu Phe Ser Cys Ser Val Trp Arg Pro Gln Gly Lys  
 85 90 95  
 Ser Tyr Leu Phe Phe Thr Gln Phe Lys Ala Glu Leu Lys Gly Thr Lys  
 100 105 110  
 Ile Glu Tyr Ala Asn Ala Tyr Ser Gln Ser Ala Ala Gly Gly Gln Ser  
 115 120 125  
 Asp Val Pro Leu Lys Pro Glu Glu Phe Thr Ile Gly Glu Ser Thr Val  
 130 135 140  
 Thr His Lys Asp Gly Lys Phe Ser Ala Gln Leu Ser Lys Leu Thr Val  
 145 150 155 160  
 Ile Gly Arg Thr Gln Lys Asp Glu Leu  
 165

[0015]

<210> 20  
 <211> 158  
 <212> PRT  
 <213> *Xenopus tropicalis*  
 <400> 20  
 Met Ala Thr Tyr Gly Ile Ile Cys Ala Phe Leu Leu Leu Leu Ala Val  
 1 5 10 15  
 Cys Ser Ala Gln Glu Lys Ser Ser Thr Glu Glu Phe Asp Val Arg Pro  
 20 25 30  
 Gly Gly Leu Gln His Ser Phe Thr Ser Lys Leu Gly Asp Tyr Ala Cys  
 35 40 45  
 Thr Phe Thr Tyr Ala Ala Gln Gly Gly Thr Asn Glu Lys Trp His Met  
 50 55 60  
 Ser Val Gly Leu Ser Asp Asp Asn Gln His Phe Ser Cys Ser Ile Trp  
 65 70 75 80  
 Arg Pro Gln Gly Lys Ser Tyr Leu Phe Phe Thr Gly Phe Lys Ala Glu  
 85 90 95  
 Val Thr Gly Gly Lys Ile Glu Phe Ser Glu Ala Tyr Ser Gln Ala Ser  
 100 105 110  
 Ser Asp Gly Ser Ser Asp Val Lys Leu Lys Ser Ser Glu Tyr Asp Val  
 115 120 125

Thr Asp Asn Val Val Ser His Arg Pro Gly Ser Phe Ser Ser Ser Leu  
 130 135 140

Cys Lys Leu Val Leu Val Ala Arg Ser Glu His Asp Glu Leu  
 145 150 155

<210> 21  
 <211> 254  
 <212> PRT  
 <213> *Macaca mulatta*

<400> 21

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

Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Ala  
 20 25 30

Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Ala  
 35 40 45

Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
 50 55 60

Ser Ala Asn Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
 65 70 75 80

[0016] Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
 85 90 95

Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Val  
 100 105 110

Pro Arg Arg Pro Gly Ala Leu Asp Gly Leu Glu Ala Ser Gly Tyr Val  
 115 120 125

Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140

Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser  
 145 150 155 160

Val Val Thr His Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val  
 165 170 175

Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Gln Pro Pro Ala Thr  
 180 185 190

Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205

Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220

Lys Tyr Trp His Ile Ile Leu Gly Gly Ala Val Leu Leu Thr Ala Leu  
 225 230 235 240

Arg Pro Ala Ala Pro Gly Pro Ala Pro Pro Pro Gln Glu Ala  
 245 250

<210> 22  
 <211> 254  
 <212> PRT  
 <213> *Cavia porcellus*

<400> 22

Met Ala Ala Ala Gly Ala Gly Ala Pro Arg Leu Leu Leu Leu Leu Leu  
 1 5 10 15

Ile Val Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Ser Cys Arg Ala  
 20 25 30

Gly Ala Ala Thr Arg Gly Val Gly Ala Glu Gly Arg Glu Gly Glu Ser  
 35 40 45

Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
 50 55 60

Thr Ala Gln Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
 65 70 75 80

Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Asn Glu Glu Glu Arg Gly  
 85 90 95

[0017] Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Val  
 100 105 110

Pro Arg Arg Pro Gly Ala Leu Asp Ser Ala Glu Ala Gly Gly Tyr Val  
 115 120 125

Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140

Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser  
 145 150 155 160

Val Val Thr Tyr Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val  
 165 170 175

Asp Leu Glu Leu Phe Asn Thr Ser Val Arg Leu Arg Pro Pro Gly Thr  
 180 185 190

Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205

Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220

Lys Tyr Trp His Leu Ile Leu Gly Gly Ala Val Leu Leu Thr Ala Leu  
 225 230 235 240

Arg Pro Ala Ala Pro Gly Pro Thr Pro Pro Pro Gln Glu Ala  
 245 250

<210> 23  
 <211> 254  
 <212> PRT  
 <213> *Canis lupus familiaris*  
 <400> 23  
 Met Ala Ala Ala Gly Ala Val Val Thr Arg Leu Phe Leu Leu Leu Leu  
 1 5 10 15  
 Met Ala Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Ser  
 20 25 30  
 Gly Ala Ala Leu Arg Gly Ala Gly Ala Glu Gly Arg Glu Ser Glu Gly  
 35 40 45  
 Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
 50 55 60  
 Ser Ala His Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
 65 70 75 80  
 Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Asn Glu Glu Glu Arg Gly  
 85 90 95  
 Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Val  
 100 105 110  
 Pro Gln Arg Pro Gly Val Pro Asp Gly Ala Glu Ala Gly Gly Tyr Val  
 115 120 125  
 Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140  
 Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser  
 145 150 155 160  
 Val Val Thr His Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val  
 165 170 175  
 Asp Leu Glu Leu Phe Asn Thr Ser Val His Leu Gln Pro Pro Ala Thr  
 180 185 190  
 Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205  
 Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220  
 Lys Tyr Trp His Leu Val Leu Gly Gly Ala Val Leu Leu Thr Ala Leu  
 225 230 235 240  
 Arg Pro Ala Ala Pro Gly Pro Thr Pro Pro Pro Gln Glu Ala  
 245 250  
 <210> 24  
 <211> 254  
 <212> PRT

[0018]

<213> Mus musculus  
 <400> 24  
 Met Val Ala Ala Gly Ala Gly Val Thr Arg Leu Leu Val Leu Leu Leu  
 1 5 10 15  
 Met Val Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Val  
 20 25 30  
 Gly Ala Ser Ala Arg Gly Thr Gly Ala Asp Gly Arg Glu Ala Glu Gly  
 35 40 45  
 Cys Gly Thr Val Ala Leu Leu Leu Glu His Ser Phe Glu Leu Gly Asp  
 50 55 60  
 Gly Ala Asn Phe Gln Lys Arg Gly Leu Leu Leu Trp Asn Gln Gln Asp  
 65 70 75 80  
 Gly Thr Leu Ser Ala Thr Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
 85 90 95  
 Arg Leu Arg Asp Val Ala Ala Val Asn Gly Leu Tyr Arg Val Arg Val  
 100 105 110  
 Pro Arg Arg Pro Gly Thr Leu Asp Gly Ser Glu Ala Gly Gly His Val  
 115 120 125  
 [0019] Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140  
 Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Leu Ser  
 145 150 155 160  
 Val Val Val Tyr Pro Gly Gly Cys Arg Gly Ser Glu Val Glu Asp Glu  
 165 170 175  
 Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Arg Pro Pro Ser Thr  
 180 185 190  
 Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205  
 Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220  
 Lys Tyr Trp His Leu Ile Leu Gly Gly Ala Val Leu Leu Thr Ala Leu  
 225 230 235 240  
 Arg Pro Ala Ala Pro Gly Pro Ala Pro Ala Pro Thr Glu Ala  
 245 250  
 <210> 25  
 <211> 254  
 <212> PRT  
 <213> Equus caballus  
 <400> 25

Met Ala Ala Ala Gly Ala Gly Ala Thr Arg Leu Leu Leu Leu Leu Leu  
1 5 10 15

Met Ala Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Pro  
20 25 30

Gly Thr Ala Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Gly  
35 40 45

Cys Gly Pro Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
50 55 60

Asn Ala His Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
65 70 75 80

Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
85 90 95

Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Val  
100 105 110

Pro Arg Arg Pro Gly Thr Pro Asp Gly Leu Glu Ala Gly Gly Tyr Val  
115 120 125

Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
130 135 140

[0020] Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser  
145 150 155 160

Val Val Thr His Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val  
165 170 175

Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Gln Pro Pro Val Thr  
180 185 190

Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
195 200 205

Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
210 215 220

Lys Tyr Trp His Leu Ile Leu Gly Gly Ala Val Leu Leu Thr Ala Leu  
225 230 235 240

Arg Pro Ala Ala Pro Gly Pro Ala Pro Pro Pro Gln Glu Ala  
245 250

<210> 26  
<211> 258  
<212> PRT  
<213> Xenopus laevis

<400> 26

Gly Gly Gln Arg Ala Gly Pro Leu Ser Ser Ile Val Thr Gly Asn Cys  
1 5 10 15



Trp	Val	Trp	Leu	Ile	Ala	Leu	Pro	Phe	Leu	Ala	Val	Thr	Ala	Gln	Gly	
			20					25					30			
Ser	Val	Cys	Arg	Leu	Lys	Thr	Gly	Asp	Gly	Arg	Glu	Ser	Glu	Ser	Cys	
		35					40					45				
Gly	Thr	Asn	Leu	Glu	Leu	Glu	His	Ser	Phe	Glu	Leu	Asp	Asp	Ser	Ile	
	50					55					60					
His	Phe	Thr	Lys	Arg	Gly	Ser	Leu	Phe	Trp	Ser	Gly	Thr	Ala	Glu	Gln	
65					70					75					80	
Ser	Ile	Ser	Ile	Leu	Gln	Lys	Gln	Leu	Thr	Glu	Asp	Glu	Arg	Asn	Lys	
				85					90					95		
Leu	Arg	Asp	Ile	Ala	Asn	Leu	Asn	Gly	Leu	Tyr	Arg	Ile	Arg	Ile	Pro	
			100					105					110			
Arg	Lys	Leu	Gly	Ile	Thr	Glu	Glu	Ala	Asn	Glu	Tyr	Val	Thr	Ser	Phe	
		115					120					125				
Val	Arg	Ala	Cys	Ser	Met	Val	Glu	Ser	His	Leu	Ser	Asp	Glu	Ile	Thr	
	130					135					140					
Val	His	Thr	Asp	Leu	Ser	Gly	Asn	Val	Ile	Gly	Val	Ser	Ile	Val	Thr	
145					150					155					160	
[0021]	Phe	Pro	Gly	Ser	Cys	Asn	Gly	Ala	Glu	Val	Glu	Asp	Val	Asp	Leu	Glu
					165					170					175	
Met	Phe	Asn	Thr	Thr	Val	His	Met	Gln	Gln	Pro	Ile	Pro	Ala	Ala	Val	
			180					185					190			
Pro	Glu	Thr	Ala	Ala	Phe	Ile	Glu	Arg	Leu	Glu	Met	Glu	Gln	Ala	Gln	
		195					200					205				
Lys	Ala	Lys	Asn	Pro	Gln	Glu	Gln	Lys	Ser	Phe	Phe	Ala	Lys	Tyr	Trp	
	210					215					220					
Met	Tyr	Ile	Ile	Pro	Val	Val	Leu	Phe	Leu	Met	Met	Ser	Gly	Ala	Ser	
225					230					235					240	
Asp	Ala	Gly	Asn	Gln	Gly	Gly	Asn	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Gly	
			245					250					255			
Gly	Arg															
<210>	27															
<211>	251															
<212>	PRT															
<213>	Danio rerio															
<400>	27															
Met	Ala	Pro	Ile	Arg	Val	Leu	Ser	Leu	Val	Leu	Pro	Ile	Leu	Ser	Thr	
1					5				10					15		

Val Pro Leu Leu Leu Thr Gln Phe Gly Glu Cys Asn Asn Gly Arg Arg  
 20 25 30  
 Ser Gly Asp Ala Val Asp Thr Asp Phe Ser Gly Phe Ser Val Pro Leu  
 35 40 45  
 Glu His Ser Phe Glu Val Asp Asp Val Pro Arg Phe Arg Leu Arg Gly  
 50 55 60  
 Ala Leu Gln Phe Arg Gly Gly Arg Glu Asn Ser Val Tyr Leu Ser Gln  
 65 70 75 80  
 Asn Gln Leu Ser Glu Lys Asp Arg Asn Thr Leu Lys Asp Val Ala Ala  
 85 90 95  
 Val Asp Gly Leu Tyr Arg Ile Arg Val Pro Arg Val Ser Leu Gln Val  
 100 105 110  
 Asp Arg Gln Thr Glu Arg Gln Tyr Glu Gly Tyr Leu Thr Ala Phe Val  
 115 120 125  
 Arg Ala Cys Ala Leu Val Glu Ser His Leu Ser Asp Val Ile Thr Leu  
 130 135 140  
 His Thr Asp Val Ser Gly Tyr Val Ile Gly Ile Ser Ile Val Thr Ile  
 145 150 155 160  
 [0022] Pro Gly Ser Cys Arg Gly Ile Glu Val Glu Asp Glu Val Asp Leu Glu  
 165 170 175  
 Val Phe Asn Thr Thr Ile Ser Val Met Ala Pro Val Thr Ala Pro Val  
 180 185 190  
 Pro Glu Thr Ala Pro Tyr Ile Glu Arg Met Glu Met Glu Met Glu Lys  
 195 200 205  
 Lys Gly Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala Lys Tyr Trp  
 210 215 220  
 Tyr Leu Ile Leu Gly Gly Ala Val Phe Leu Met Ala Thr Ser Ser Ala  
 225 230 235 240  
 Gln Thr Pro Pro Gly Gly Ala Arg Glu Gln Ser  
 245 250  
 <210> 28  
 <211> 262  
 <212> PRT  
 <213> Pan troglodytes  
 <400> 28  
 Met Ala Ala Ala Ser Ala Gly Ala Thr Arg Leu Leu Leu Leu Leu Leu  
 1 5 10 15  
 Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly Ser Ser Cys Arg Ala  
 20 25 30

Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Ala  
 35 40 45  
 Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
 50 55 60  
 Ser Ala Asn Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
 65 70 75 80  
 Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
 85 90 95  
 Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Ile  
 100 105 110  
 Pro Arg Arg Pro Gly Ala Leu Asp Gly Leu Glu Ala Gly Gly Tyr Val  
 115 120 125  
 Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140  
 Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser  
 145 150 155 160  
 Val Val Thr Gln Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val  
 165 170 175  
 [0023] Asp Leu Glu Leu Phe Thr Thr Ser Val Gln Leu Gln Pro Pro Thr Thr  
 180 185 190  
 Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205  
 Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220  
 Lys Tyr Trp Met Tyr Ile Ile Pro Val Val Leu Phe Leu Met Met Ser  
 225 230 235 240  
 Gly Ala Pro Asp Thr Gly Gly Gln Gly Gly Gly Gly Gly Gly Gly Gly  
 245 250 255  
 Gly Gly Gly Ser Gly Arg  
 260  
 <210> 29  
 <211> 258  
 <212> PRT  
 <213> Mus musculus  
 <400> 29  
 Met Val Ala Ala Gly Ala Gly Val Thr Arg Leu Leu Val Leu Leu Leu  
 1 5 10 15  
 Met Val Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Val  
 20 25 30

Gly Ala Ser Ala Arg Gly Thr Gly Ala Asp Gly Arg Glu Ala Glu Gly  
 35 40 45  
 Cys Gly Thr Val Ala Leu Leu Leu Glu His Ser Phe Glu Leu Gly Asp  
 50 55 60  
 Gly Ala Asn Phe Gln Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
 65 70 75 80  
 Gly Thr Leu Ser Ala Thr Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
 85 90 95  
 Arg Leu Arg Asp Val Ala Ala Val Asn Gly Leu Tyr Arg Val Arg Val  
 100 105 110  
 Pro Arg Arg Pro Gly Thr Leu Asp Gly Ser Glu Ala Gly Gly His Val  
 115 120 125  
 Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140  
 Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Leu Ser  
 145 150 155 160  
 Val Val Val Tyr Pro Gly Gly Cys Arg Gly Ser Glu Val Glu Asp Glu  
 165 170 175  
 [0024] Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Arg Pro Pro Ser Thr  
 180 185 190  
 Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205  
 Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220  
 Lys Tyr Trp Met Tyr Ile Ile Pro Val Val Leu Phe Leu Met Met Ser  
 225 230 235 240  
 Gly Ala Pro Asp Ala Gly Gly Gln Gly Gly Gly Gly Gly Gly Ser  
 245 250 255

Ser Arg

<210> 30  
 <211> 210  
 <212> PRT  
 <213> *Heterocephalus glaber*

<400> 30

Met Val Met Ala Ala Ser Gly Ala Ser Ala Ser Arg Leu Leu Leu Leu  
 1 5 10 15  
 Leu Leu Ile Val Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys  
 20 25 30

Arg Ala Gly Ala Ala Ala Arg Gly Val Gly Ala Glu Gly Arg Glu Gly  
35 40 45

Glu Ser Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile  
50 55 60

Asp Asp Ala Ala His Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln  
65 70 75 80

Gln Asp Gly Thr Leu Ser Pro Ser Gln Arg Gln Leu Ser Glu Glu Glu  
85 90 95

Arg Gly Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val  
100 105 110

Arg Val Pro Arg Arg Pro Gly Ala Leu Asp Ser Ser Glu Ala Gly Gly  
115 120 125

Tyr Val Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu  
130 135 140

Ser Asp Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly  
145 150 155 160

Val Ser Val Val Thr Asp Pro Gly Gly Cys Arg Gly His Glu Val Glu  
165 170 175

[0025] Asp Val Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Gln Pro Pro  
180 185 190

Gly Thr Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu  
195 200 205

Met Glu  
210

<210> 31  
<211> 262  
<212> PRT  
<213> *Loxodonta africana*

<400> 31

Met Ala Ala Ala Gly Ala Gly Ala Thr Arg Leu Leu Leu Leu Leu  
1 5 10 15

Met Ala Ala Ala Ala Pro Ser Arg Ala Arg Gly Ser Ser Cys Arg Ala  
20 25 30

Gly Ala Ala Thr Arg Gly Ala Gly Ala Glu Gly Arg Glu Asn Glu Gly  
35 40 45

Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
50 55 60

Ala Met His Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
65 70 75 80

Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
 85 90 95  
 Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Val  
 100 105 110  
 Pro Arg Arg Pro Gly Ala Pro Glu Gly Pro Glu Ala Gly Gly Tyr Val  
 115 120 125  
 Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
 130 135 140  
 Gln Leu Thr Leu His Val Asp Val Val Gly Asn Val Val Gly Val Ser  
 145 150 155 160  
 Val Val Thr Leu Pro Gly Gly Cys Arg Gly Tyr Glu Val Glu Asp Val  
 165 170 175  
 Asp Leu Glu Leu Phe Asn Thr Thr Val Gln Leu Gln Pro Pro Thr Thr  
 180 185 190  
 Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
 195 200 205  
 Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
 210 215 220  
 [0026] Lys Tyr Trp Met Tyr Ile Ile Pro Val Val Leu Phe Leu Met Met Ser  
 225 230 235 240  
 Gly Ala Pro Asp Thr Gly Gly Gln Gly Gly Gly Gly Gly Gly Gly Gly  
 245 250 255  
 Gly Gly Gly Ser Gly Arg  
 260  
 <210> 32  
 <211> 263  
 <212> PRT  
 <213> *Xenopus tropicalis*  
 <400> 32  
 Met Ala Ala Gly Cys Leu Val Gly Gln Arg Ala Gly Pro Leu Ser Asp  
 1 5 10 15  
 Lys Leu Ser Gly Tyr Cys Trp Val Leu Leu Pro Leu Leu Val Ala  
 20 25 30  
 Thr Ala Gln Ala Ser Val Cys Arg Leu Lys Thr Gly Asp Gly Arg Asp  
 35 40 45  
 Ser Glu Ser Cys Gly Thr Asn Leu Glu Leu Glu His Ser Phe Glu Leu  
 50 55 60  
 Asp Asp Ser Ile His Phe Lys Lys Arg Gly Ser Leu Ile Trp Ser Gly  
 65 70 75 80

Thr Ala Glu Gln Ser Ile Ser Ile Leu Gln Lys Gln Leu Thr Glu Asp  
 85 90 95  
 Glu Arg Asn Lys Leu Arg Asp Ile Ala Asn Leu Asn Gly Leu Tyr Arg  
 100 105 110  
 Ile Arg Val Pro Arg Lys Leu Gly Ile Thr Glu Glu Ala Asn Glu Tyr  
 115 120 125  
 Val Thr Ser Phe Val Arg Ala Cys Ser Met Val Glu Ser His Leu Ser  
 130 135 140  
 Asp Gln Ile Ser Val His Thr Asp Ile Ser Gly Asn Val Val Gly Ile  
 145 150 155 160  
 Ser Ile Val Thr Phe Pro Gly Ser Cys Asn Gly Ala Glu Val Glu Asp  
 165 170 175  
 Val Asp Leu Glu Met Phe Asn Thr Thr Val Tyr Ile Gln Gln Pro Ile  
 180 185 190  
 Ala Ala Ala Val Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met  
 195 200 205  
 Glu Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe  
 210 215 220  
 [0027] Ala Lys Tyr Trp Met Tyr Ile Ile Pro Val Val Leu Phe Leu Met Met  
 225 230 235 240  
 Ser Gly Ala Ser Asp Ala Gly Asn Gln Gly Gly Asn Gly Gly Gly Gly  
 245 250 255  
 Gly Gly Ser Gly Gly Gly Arg  
 260  
 <210> 33  
 <211> 251  
 <212> PRT  
 <213> Danio rerio  
 <400> 33  
 Met Ala Pro Ile Arg Val Leu Ser Leu Val Leu Pro Ile Leu Ser Thr  
 1 5 10 15  
 Val Pro Leu Leu Leu Thr Gln Phe Gly Glu Cys Asn Asn Gly Arg Arg  
 20 25 30  
 Ser Gly Asp Ala Val Asp Thr Asp Phe Ser Gly Phe Ser Val Pro Leu  
 35 40 45  
 Glu His Ser Phe Glu Val Asp Asp Val Pro Arg Phe Arg Leu Arg Gly  
 50 55 60  
 Ala Leu Gln Phe Arg Gly Gly Arg Glu Asn Ser Val Tyr Leu Ser Gln  
 65 70 75 80

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

Val Asp Gly Leu Tyr Arg Ile Arg Val Pro Arg Val Ser Leu Gln Val  
100 105 110

Asp Arg Gln Thr Glu Arg Gln Tyr Glu Gly Tyr Leu Thr Ala Phe Val  
115 120 125

Arg Ala Cys Ala Leu Val Glu Ser His Leu Ser Asp Val Ile Thr Leu  
130 135 140

His Thr Asp Val Ser Gly Tyr Val Ile Gly Ile Ser Ile Val Thr Ile  
145 150 155 160

Pro Gly Ser Cys Arg Gly Ile Glu Val Glu Asp Glu Val Asp Leu Glu  
165 170 175

Val Phe Asn Thr Thr Ile Ser Val Met Ala Pro Val Thr Ala Pro Val  
180 185 190

Pro Glu Thr Ala Pro Tyr Ile Glu Arg Met Glu Met Glu Met Glu Lys  
195 200 205

Lys Gly Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala Lys Tyr Trp  
210 215 220

[0028] Tyr Leu Ile Leu Gly Gly Ala Val Phe Leu Met Ala Thr Ser Ser Ala  
225 230 235 240

Gln Thr Pro Pro Gly Gly Ala Arg Glu Gln Ser  
245 250

<210> 34  
<211> 269  
<212> PRT  
<213> Homo sapiens

<400> 34

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

Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly Ser Gly Cys Arg Ala  
20 25 30

Gly Thr Gly Ala Arg Gly Ala Gly Ala Glu Gly Arg Glu Gly Glu Ala  
35 40 45

Cys Gly Thr Val Gly Leu Leu Leu Glu His Ser Phe Glu Ile Asp Asp  
50 55 60

Ser Ala Asn Phe Arg Lys Arg Gly Ser Leu Leu Trp Asn Gln Gln Asp  
65 70 75 80

Gly Thr Leu Ser Leu Ser Gln Arg Gln Leu Ser Glu Glu Glu Arg Gly  
85 90 95



Arg Leu Arg Asp Val Ala Ala Leu Asn Gly Leu Tyr Arg Val Arg Ile  
100 105 110

Pro Arg Arg Pro Gly Ala Leu Asp Gly Leu Glu Ala Gly Gly Tyr Val  
115 120 125

Ser Ser Phe Val Pro Ala Cys Ser Leu Val Glu Ser His Leu Ser Asp  
130 135 140

Gln Leu Thr Leu His Val Asp Val Ala Gly Asn Val Val Gly Val Ser  
145 150 155 160

Val Val Thr His Pro Gly Gly Cys Arg Gly His Glu Val Glu Asp Val  
165 170 175

Asp Leu Glu Leu Phe Asn Thr Ser Val Gln Leu Gln Pro Pro Thr Thr  
180 185 190

Ala Pro Gly Pro Glu Thr Ala Ala Phe Ile Glu Arg Leu Glu Met Glu  
195 200 205

Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe Phe Ala  
210 215 220

Lys Tyr Trp Met Tyr Ile Ile Pro Val Val Leu Phe Leu Met Met Ser  
225 230 235 240

[0029]

Gly Ala Pro Asp Thr Gly Gly Gln Gly Gly Gly Gly Gly Gly Gly Gly  
245 250 255

Gly Gly Gly Ser Gly Leu Cys Cys Val Pro Pro Ser Leu  
260 265

<210> 35  
<211> 31  
<212> PRT  
<213> Homo sapiens

<400> 35

Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Gly Ala Ser Leu Trp  
1 5 10 15

Ala Ala Leu Leu Leu Gly Ala Val Ala Leu Arg Pro Ala Glu Ala  
20 25 30

<210> 36  
<211> 6  
<212> PRT  
<213> 人工序列

<220>  
<223> 用于纯化蛋白质的六组氨酸标签

<400> 36

His His His His His His  
1 5

<210> 37  
<211> 27

[0030]

<212> PRT  
<213> Homo sapiens

<400> 37

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

Met Ala Val Ala Ala Pro Ser Arg Ala Arg Gly  
20 25

<210> 38  
<211> 16  
<212> PRT  
<213> Homo sapiens

<400> 38

Cys Thr Ile Trp Arg Pro Gln Gly Lys Ser Tyr Leu Tyr Phe Thr Gln  
1 5 10 15

<210> 39  
<211> 16  
<212> PRT  
<213> Homo sapiens

<400> 39

Cys Glu Gln Ala Gln Lys Ala Lys Asn Pro Gln Glu Gln Lys Ser Phe  
1 5 10 15

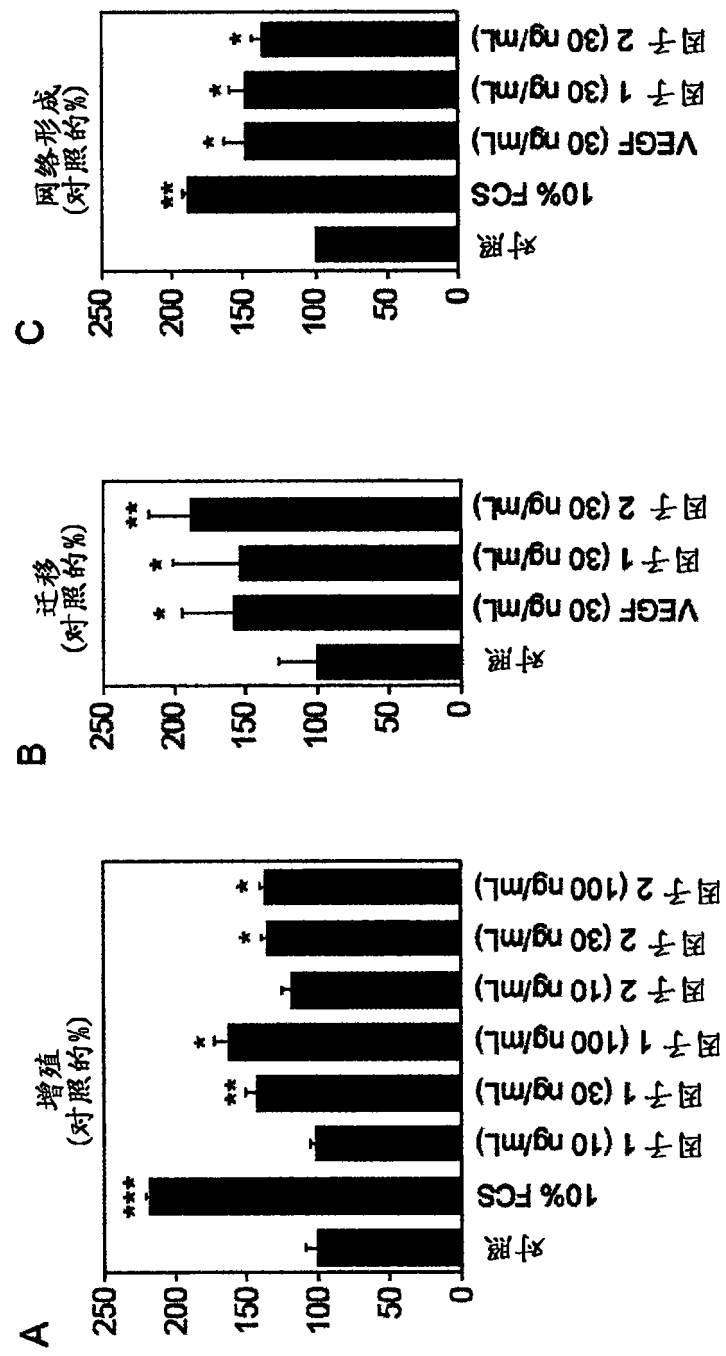


图1

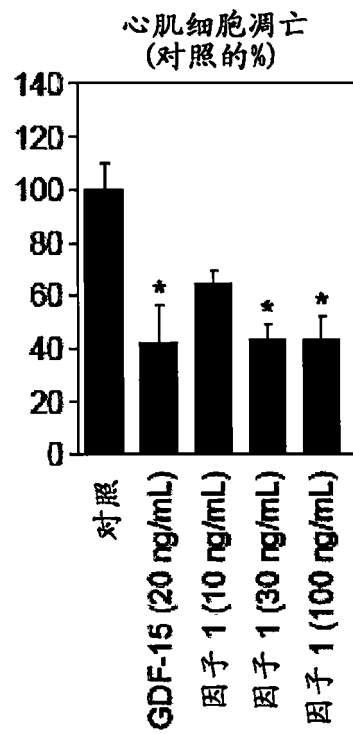


图2

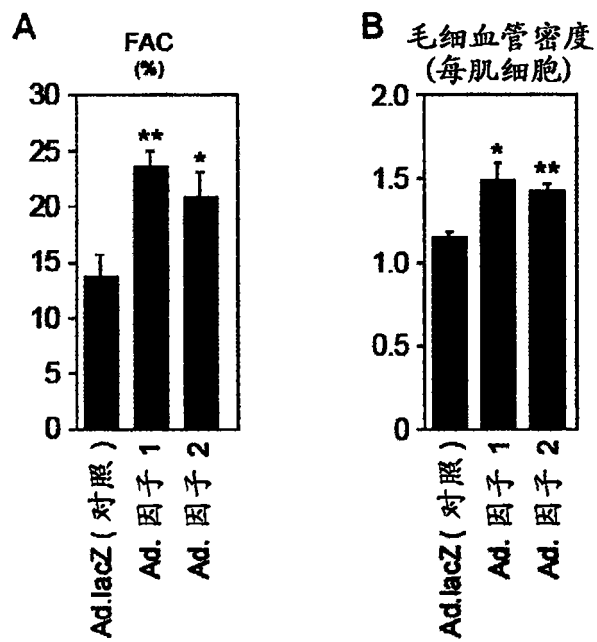


图3

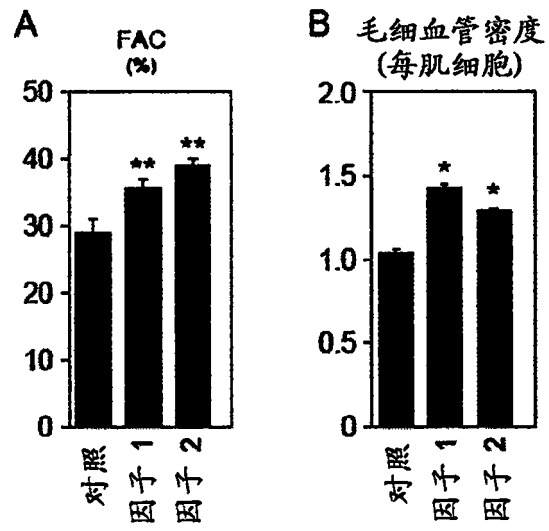


图4

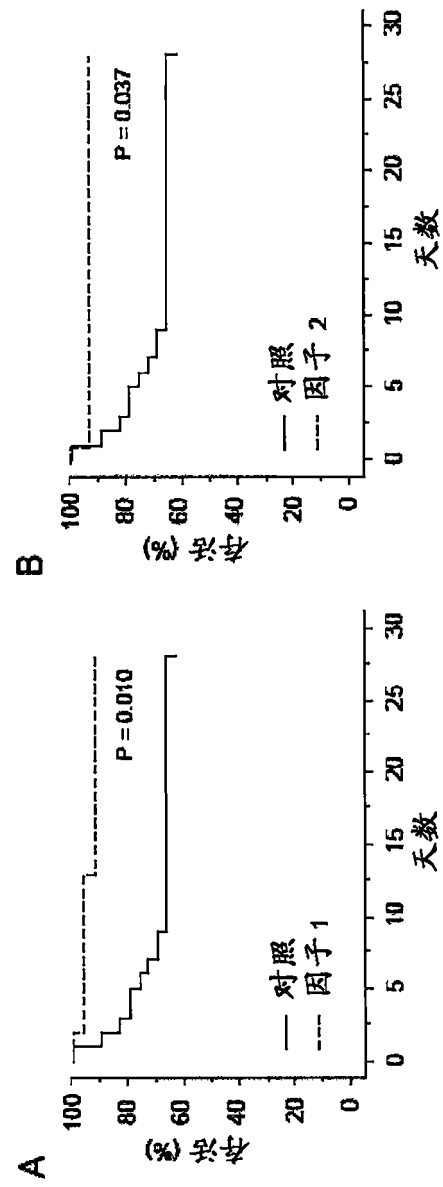


图5

## CLUSTAL 2.1 多序列比对

```

SEQ_ID_NO_2      HAAPSGGUNG-VGASLWAAILLGAVALRPAEAVSEPTTVAFDVRPGGVVHVSFHNVPD 59
SEQ_ID_NO_12     HAAPSGGUNG-VGASLWAAILLGAVALSFAEAVSEPTTVAFDVRPGGVVHVSFHNVPD 59
SEQ_ID_NO_14     HAAPSGRRNGSGGANTLWVSLILAAAAALRPVETVSEPTTVAFDVRPGGVVHVSFQNVGPD 60
SEQ_ID_NO_17     HAAPSERRNG-GGASLWAAILLAAAAALRPAEAVSEPTTVAFDVRPGGVVHVSFQNVGPD 59
SEQ_ID_NO_15     HAAPRGRRNGSAGASHWGAILLAAVALRSVEAVSEPTTVAFDVRPGGVVHVSFHSAGPD 60
SEQ_ID_NO_16     HAAPRGNSDG-CGGAWFAAILLAAVALRPAEAVSEPTTVAFDVRPGGVVHVSFQNVGPD 59
SEQ_ID_NO_13     HAAPS-----CGFOTAVVLAAILKLAAYSEPTTVPFDVRPGGVVHVSQDVGPD 52
SEQ_ID_NO_18     HAAPCGRSSR----TLWAAVYPAAVLCLAVRAAEASTAEFDVRPGGVVHVSFQNVGPD 53
SEQ_ID_NO_20     NATYG-----IICAFLLLLAVCS----AQEKSTEEFDVRPGGVVHVSFQNVGPD 44
SEQ_ID_NO_19     HARQSNICAG--NLAFLLALALAAVPAEASEEQAKTVEFVVKPGGVVHVSFQNVGPD 55
                **          :  *          :  . . . *  *:***  *:  *

SEQ_ID_NO_2      KYTCEFIYASQGGTNEQQWMSLGTSEDHQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 119
SEQ_ID_NO_12     KYTCEFIYASQGGTNEQQWMSLGTSEDHQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 119
SEQ_ID_NO_14     KYTCVFIYASQGGTNEQQWMSLGTSEDHQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 120
SEQ_ID_NO_17     KYTCAFIYASQGGTNEQQWMSLGTSEDHQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 119
SEQ_ID_NO_15     RFTCTFIYASQGGTNEQQWMSLGTSEDHQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 120
SEQ_ID_NO_16     KFTCTFIYASQGGTNEQQWMSLGTSEDHQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 119
SEQ_ID_NO_13     KFTCTFIYASQGGTNEQQWMSLGTSEDSQHFTCTIWRPQCKSYLYFTQFKAIEVRGAEIEY 112
SEQ_ID_NO_18     DYTCTFIYASQGGTNEQQWMSLGTSEDELLFSCSVWRPQCKSYLYFTQFKAIEVRGAEIEY 113
SEQ_ID_NO_20     DYACTFIYASQGGTNEQQWMSLGTSEDDNQHFSCSVWRPQCKSYLYFTQFKAIEVRGAEIEY 104
SEQ_ID_NO_19     EYECSEFIYASQGGTNEQQWMSLGTSEDDNQHFSCSVWRPQCKSYLYFTQFKAIEVRGAEIEY 115
                : * ***:*****:* *. * :.*  *: * :*****:*** *:*** : **

SEQ_ID_NO_2      ANAYSKAAFERE SDVPLKTEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 173
SEQ_ID_NO_12     ANAYSKAAFERE SDVPLKTEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 173
SEQ_ID_NO_14     GHAYSKAAFEKE SDVPLKNEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 174
SEQ_ID_NO_17     GHAYSKAAFERE SDVPLKSEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 173
SEQ_ID_NO_15     GHAYSKAASERE SDVPLKNEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 174
SEQ_ID_NO_16     ANAYSKAAFERE SDVPLKNEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 173
SEQ_ID_NO_13     ANAYSKAAFERE SDVPLKSEE FEVTKTAVHRPGAFAKAE LSKLVIKASRTEL 166
SEQ_ID_NO_18     ANAYSQAAGVGAQSD IPLKQEE FEITETTAVSHREGKFRFELS KLMIKARTPHDEL 167
SEQ_ID_NO_20     SEATYQASSDGS SDVPLKSEE FEITETTAVSHREGKFRFELS KLMIKARTPHDEL 158
SEQ_ID_NO_19     ANAYSQAAGGQSDVPLKPEE FTIGESTVTHKDKFSAQLSKLTVIGRTQKDEL 169
                . ***:::  .**:* *: : . . * *: * * .*.** :.: : **

```

图6

## CLUSTAL 2.1 多序列比对

```

SEQ_ID_NO_2      MAAPSGGQNG-VGASLWAALLLCAVALRPAEAVSEPTTVAFDVVRPGGVVHSFSSHVVGPGD 59
SEQ_ID_NO_12     MAAPSGGQNG-VGASLWAALLLCAVALSPAFAVSEPTTVAFDVVRPGGVVHSFSSHVVGPGD 59
SEQ_ID_NO_15     MAAPRGRRNGSAGASMMGALLLAAVALRSVEAVSEPTTVAFDVVRPGGVVHSFSSHAGPGD 60
SEQ_ID_NO_16     MAAPRGNSDC-CGGAWFAALLLAAVALRPAEAVSEPTTVAFDVVRPGGVVHSFSONVVGPGD 59
SEQ_ID_NO_14     MAAPSGRRNGSGGANLWVSLLLAAALRPVETVSEPTTVAFDVVRPGGVVHSFSONVVGPGD 60
SEQ_ID_NO_17     MAAPSERRNG-GGASLWAALLLAAALRPAEAVSEPTTVAFDVVRPGGVVHSFSONVVGPGD 59
SEQ_ID_NO_13     MAAPS-----GGFOTAVVLAALAAALKLAAYSEPTTVFDFVRPGGVVHSFSQDVGPGD 52
                ****          . : ::*.*.** . :*****.*****:..***:

SEQ_ID_NO_2      KYTCMFTYASQGGTNEQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAEVGAEIEY 119
SEQ_ID_NO_12     KYTCMFTYASQGGTNEQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAEVGAEIEY 119
SEQ_ID_NO_15     RFTCTFTYASQGGTNEQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAEVGAQIEY 120
SEQ_ID_NO_16     KFTCTFTYASQGGTNEQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAEVHGAIEY 119
SEQ_ID_NO_14     KYTCVFTYASQGGTNEKQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAEVGAEIEY 120
SEQ_ID_NO_17     KYTCAFTYASQGGTNEKQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAEVGAEIEY 119
SEQ_ID_NO_13     KFTCTFTYASQGGTNEQWQMSLGTSEDHQHFTCTIWRPQGKSYLYFTQFKAELRGAEIEY 112
                :; ** *****:***** *** ***** *****:***:***:***

SEQ_ID_NO_2      ANAYSKAAFERESDVPLKTEEFVTKTAVAHRRPGAFKAELSKLVIVAKASRTEL 173
SEQ_ID_NO_12     ANAYSKAAFERESDVPLKTEEFVTKTAVAHRRPGAFKAELSKLVIVAKASRTEL 173
SEQ_ID_NO_15     GMAYSKAASERESDVPLKNEEFVTKTAVAHRRPGAFKAELSKLVIVAKASHSEL 174
SEQ_ID_NO_16     ANAYSKAAFERESDVPLKNEEFVTKAAVAHRRPGAFRAELSKLVIVAKEANSEL 173
SEQ_ID_NO_14     GMAYSKAAFEKESDVPLKNEEFVTKTAVFHRPGAFKAELSKLVIVAKATRSEL 174
SEQ_ID_NO_17     GMAYSKAAFERESDVPLKSEEFVTKTAVSHRRPGAFKAELSKLVIVAKASRSEL 173
SEQ_ID_NO_13     ANAYSKAAFERESDVPLKSEEFVTKTAVSHRRPGAFKAELSKLVIVAKAARSEL 166
                .***** *:*****.*****:;* *****:*****:***:***

```

图7



## CLUSTAL 2.1 多序列比对

```

SEQ_ID_NO_4      ----MAAASAGAT-RLLLLLLMAVAAPSRARGSGCCRAFTGARGAGAGAEGRGEACGTGVL 55
SEQ_ID_NO_21     ----HAATSGGAT-RLLLLLLMAVAAPSRARGSGCCRAFTGARGAGAGAEGRGEACGTGVL 55
SEQ_ID_NO_25     ----MAAAGAGAT-RLLLLLLMAAAPSARGSGCCRPCTAARGAGAEGRGEAGCGPVGL 55
SEQ_ID_NO_23     ----MAAAGAVYT-RLFLLLLMAAAPSARGSGCCRSAGALRGAGAEGRGECGCTVGL 55
SEQ_ID_NO_22     ----MAAAGACAP-RLLLLLLIVAAAPSARGSSCRAGAATRGVGAEGREGESCCTVGL 55
SEQ_ID_NO_24     ----MYAAGAVYT-RLLVLLLLHYAAAPSARGSGCCRVGASARGTGADGREAGCGTVALL 55
SEQ_ID_NO_26     GGQRAGPLSSIVTGNCWUOLIALFLFVAVTAQGSVCRLKIG-----DGRESESCGTN-LE 53
SEQ_ID_NO_27     ----MAPIRVLS---LVLPILSTVPLLLTQFGECCNNGRRS----GDAVDTDFSGFS-VP 47
      .      .      :      :      :      :      :      :      :      :      :

SEQ_ID_NO_4      LEHSFEIDDSANFRKRGSLLWN-QQDGTLSLSQRQLSEERGRLRDVAALNGLYRVRI 114
SEQ_ID_NO_21     LEHSFEIDDSANFRKRGSLLWN-QQDGTLSLSQRQLSEERGRLRDVAALNGLYRVRV 114
SEQ_ID_NO_25     LEHSFEIDDSANFRKRGSLLWN-QQDGTLSLSQRQLSEERGRLRDVAALNGLYRVRV 114
SEQ_ID_NO_23     LEHSFEIDDSANFRKRGSLLWN-QQDGTLSLSQRQLNEEERGRLRDVAALNGLYRVVPQ 114
SEQ_ID_NO_22     LEHSFEIDDTAQFRKRGSLLWN-QQDGTLSLSQRQLNEEERGRLRDVAALNGLYRVVP 114
SEQ_ID_NO_24     LEHSFELGGGANFQKRGLLLWN-QQDGTLSATQRQLSEERGRLRDVAALNGLYRVVP 114
SEQ_ID_NO_26     LEHSFEIDDSIHFTKRGSLLFWSGTAEQSISILQKQLTEDERMKLRDIANLNGLYRIRI 113
SEQ_ID_NO_27     LEHSFEVDDVPFRILRGALQFRGCRENSVLSQKLTSEKDRNTLKDVAALNGLYRVVP 107
      *****;.* ** * * : : : *;.*;.*. *;.* :;****;.*;

SEQ_ID_NO_4      R---PGALDGLAEGGYVSSFYPACSLVESHSLSQTLTHVDVAGNVVGVSVVTHPGGCRGH 171
SEQ_ID_NO_21     R---PGALDGLAEGGYVSSFYPACSLVESHSLSQTLTHVDVAGNVVGVSVVTHPGGCRGH 171
SEQ_ID_NO_25     R---PGTDPGLAEGGYVSSFYPACSLVESHSLSQTLTHVDVAGNVVGVSVVTHPGGCRGH 171
SEQ_ID_NO_23     R---PGVPDCAEAGGYVSSFYPACSLVESHSLSQTLTHVDVAGNVVGVSVVTHPGGCRGH 171
SEQ_ID_NO_22     R---PGALDSEAGGYVSSFYPACSLVESHSLSQTLTHVDVAGNVVGVSVVTHPGGCRGH 171
SEQ_ID_NO_24     R---PGTLDGSEAGGHVSSFYPACSLVESHSLSQTLTHVDVAGNVVGVSVVTHPGGCRGH 171
SEQ_ID_NO_26     K---LGITE--EANEYVTSFVRACSNVESHSLSDEITVHDLSCGVIGSVITVTPGSCNGA 168
SEQ_ID_NO_27     VSLQVDRQTERQYEGYLTAFFRACALVESHSLSVDTLTHDVSQYVIGISIVTTPGSCRG 167
      .      :      :;*** **;***** :*;*;.* *;.*;.*. **;.*

SEQ_ID_NO_4      EVED-VDLELFNTSVQLQPPPTAPGPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHII 230
SEQ_ID_NO_21     EVED-VDLELFNTSVQLQPPPTAPGPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHII 230
SEQ_ID_NO_25     EVED-VDLELFNTSVQLQPPPTAPGPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHII 230
SEQ_ID_NO_23     EVED-VDLELFNTSVHLQPPPTAPGPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHIV 230
SEQ_ID_NO_22     EVED-VDLELFNTSVRLRPPTAPGPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHII 230
SEQ_ID_NO_24     EVED-VDLELFNTSVQLRPPSTAPGPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHII 230
SEQ_ID_NO_26     EVED-VDLEHFNNTVHMQQPIPAAVPETAATFIERLEMEQAQKAKNPQE QKSFFAKYWHYI 227
SEQ_ID_NO_27     EVEDEVLDLVFNNTTISVMAPIVTAAPVPTAPVPTIERNEHEMEKKGNPQE QKSFFAKYWHII 227
      **** **;***; : : *.*. ****;***;*** *;***** :

SEQ_ID_NO_4      LGGAVLLTALRPAAPGPAPPPQEA----- 254
SEQ_ID_NO_21     LGGAVLLTALRPAAPGPAPPPQEA----- 254
SEQ_ID_NO_25     LGGAVLLTALRPAAPGPAPPPQEA----- 254
SEQ_ID_NO_23     LGGAVLLTALRPAAPGPTTTPPQEA----- 254
SEQ_ID_NO_22     LGGAVLLTALRPAAPGPTTTPPQEA----- 254
SEQ_ID_NO_24     LGGAVLLTALRPAAPGPAPAPTEA----- 254
SEQ_ID_NO_26     IPVVLFLMHS GASDAGNQGGGNGGGGGGGGG 258
SEQ_ID_NO_27     LGGAVFLMATSSAQTPPGGAREQS----- 251
      :      :;***

```

图8

## CLUSTAL 2.1 多序列比对

```

SEQ_ID_NO_4      MAASAGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGRGEACGTIVGLLLEHSF 60
SEQ_ID_NO_21     MAATSGGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGRGEACGTIVGLLLEHSF 60
SEQ_ID_NO_25     MAAGAGATRLLLLLLMAVAAPSRARGSGCRPGTAARGAGAEGRGEACGPPVGLLLEHSF 60
SEQ_ID_NO_23     MAAGAVVTRFLLLLLMAVAAPSRARGSGCRSGAALRGAGAEGRGEACGCTIVGLLLEHSF 60
SEQ_ID_NO_22     MAAGAGAPRLLLLLLIVAAAPSRARGSSCRAGAAATRGVGAEGREGEACGTIVGLLLEHSF 60
SEQ_ID_NO_24     MVAAGAGVTRLLVLLLMVAAPSRARGSGCRVGAASARGTGADGREAGCGTIVALLLEHSF 60
                  *:.. .*:;***:..*****:** *:. **.**:***.*.*.*:*****

SEQ_ID_NO_4      EIDDSANFRKRGSLLWNQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVVPRRPGALD 120
SEQ_ID_NO_21     EIDDSANFRKRGSLLWNQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVVPRRPGALD 120
SEQ_ID_NO_25     EIDDMAHFRKRGSLLWNQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVVPRRPGTPD 120
SEQ_ID_NO_23     EIDDSAHFRKRGSLLWNQDGTLSLSQRQLNEEERGRLRDVAALNGLYRVVPRRPGVPD 120
SEQ_ID_NO_22     EIDDTAQFRKRGSLLWNQDGTLSLSQRQLNEEERGRLRDVAALNGLYRVVPRRPGALD 120
SEQ_ID_NO_24     ELGDGANFQKRGLLLWNQDGTLSATQRQLSEEERGRLRDVAALNGLYRVVPRRPGTLD 120
                  *:.* *:;*:*** *****:****.*****:*****:*:***.*

SEQ_ID_NO_4      GLEAGGYVSSFVPACSLVESHSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLEL 180
SEQ_ID_NO_21     GLEAGGYVSSFVPACSLVESHSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLEL 180
SEQ_ID_NO_25     GLEAGGYVSSFVPACSLVESHSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLEL 180
SEQ_ID_NO_23     GAEAGGYVSSFVPACSLVESHSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLEL 180
SEQ_ID_NO_22     SAEAGGYVSSFVPACSLVESHSDDLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLEL 180
SEQ_ID_NO_24     GSEAGGHVSSFVPACSLVESHSDDLTLHVDVAGNVVGLSVVTHPGGCRGSEVEDDLEL 180
                  . **.*:*****:***.*:***** ****

SEQ_ID_NO_4      FNTSVQLQPPPTAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWHILGGAVLITAL 240
SEQ_ID_NO_21     FNTSVQLQPPATAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWHILGGAVLITAL 240
SEQ_ID_NO_25     FNTSVQLQPPPTAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWHILGGAVLITAL 240
SEQ_ID_NO_23     FNTSVHLQPPATAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWHILGGAVLITAL 240
SEQ_ID_NO_22     FNTSVRLRPPPTAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWHILGGAVLITAL 240
SEQ_ID_NO_24     FNTSVQLRPPPTAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYWHILGGAVLITAL 240
                  *****:*:*** *****:*****:*****:*****

SEQ_ID_NO_4      RPAAPGPAPPPQEA 254
SEQ_ID_NO_21     RPAAPGPAPPPQEA 254
SEQ_ID_NO_25     RPAAPGPAPPPQEA 254
SEQ_ID_NO_23     RPAAPGPTPPPQEA 254
SEQ_ID_NO_22     RPAAPGPTPPPQEA 254
SEQ_ID_NO_24     RPAAPGPAAPTEA 254
                  *****:*.* **

```

图9

## CLUSTAL 2.1 多序列比对

```

SEQ_ID_NO_5      --MAA-----ASAGATRL--LLLLLMAVAAPSRARGSGCRACTGARGAGAEGREGEAC 49
SEQ_ID_NO_28     --MAA-----ASAGATRL--LLLLLMAVAAPSRARGSSCRAGTGARGAGAEGREGEAC 49
SEQ_ID_NO_31     --MAA-----AGAGATRL--LLLLLMAAAAPSRARGSSCRAGAAATRGAGAEGRENEGC 49
SEQ_ID_NO_30     EVMAA-----SCASATRL--LLLLLTVAAPSRARGSGCRAGAAARGVGAEGREGESG 51
SEQ_ID_NO_29     --MVA-----AGAGVTRL--LVLLLLMVAAPSRARGSGCRVGAASARGTGADGREAEGC 49
SEQ_ID_NO_32     --MAAGCLVGQRAGPLSDKLSGYCWVLP LLLLVATAQASVCRLKTG-----DGRDSESC 52
SEQ_ID_NO_33     --MAP-----IRVLSLVLPILSTVPLLLTQFGECNNGRS----GDAVDIDFS 42
                  *..          ::      ::      :: . *.          :. : . .

SEQ_ID_NO_5      GTVGLLLEHSFEIDDSANFRKRGSLLWN-QQDGTLSLSQRQLSEEERGLRDVAALNGIY 108
SEQ_ID_NO_28     GTVGLLLEHSFEIDDSANFRKRGSLLWN-QQDGTLSLSQRQLSEEERGLRDVAALNGIY 108
SEQ_ID_NO_31     GTVGLLLEHSFEIDDAHHFRKRGSLLWN-QQDGTLSLSQRQLSEEERGLRDVAALNGIY 108
SEQ_ID_NO_30     GTVGLLLEHSFEIDDAHHFRKRGSLLWN-QQDGTLSPSQRQLSEEERGLRDVAALNGIY 110
SEQ_ID_NO_29     GTVALLLEHSFELGDGANFQKRGSLLWN-QQDGTLSATQRQLSEEERGLRDVAALNGIY 108
SEQ_ID_NO_32     GTW-LELEHSFELDDSIHFRKRGSLLWSGTAEQSSISILQKQLTEDERNKLELDIANLNGIY 111
SEQ_ID_NO_33     GFS-VPLEHSFEVDDVPRFRLRGALQFRGGRENSVYLSQNQLSEKDRNTLKDVAADVGLY 101
                  * ; *****:* :*: **:*: : : : : * :*:*:*:*. :*:*: :*:**

SEQ_ID_NO_5      RVRIPRR---PGALDGLEAGGYVSSFYPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHP 165
SEQ_ID_NO_28     RVRIPRR---PGALDGLEAGGYVSSFYPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHP 165
SEQ_ID_NO_31     RVRVPRR---PGAPEGPEAGGYVSSFYPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHP 165
SEQ_ID_NO_30     RVRVPRR---PGALDSSEAGGYVSSFYPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHP 167
SEQ_ID_NO_29     RVRVPRR---PGTLDGSEAGGHVSSFYPACSLVESHLSDQLTLHVDVAGNVVGLSVVYYP 165
SEQ_ID_NO_32     RIRVPRK---LGITE--EAMEYVTSFVRACSNVESHLSDQISVHTDISGNVVGISIVTTP 166
SEQ_ID_NO_33     RIRVPRVSLQVDRQTERQYEGYLTAFFRACALVESHLSDVITLHTDVSgyvIGISIVTTP 161
                  *:*:** . : : : : ** :*:***** :*:*: * :*:*:*. *

SEQ_ID_NO_5      GGCRGHEVED-VDLELFNTSVQLQPPPTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFA 224
SEQ_ID_NO_28     GGCRGHEVED-VDLELFNTSVQLQPPPTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFA 224
SEQ_ID_NO_31     GGCRGYEVED-VDLELFNTIVQLQPPPTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFA 224
SEQ_ID_NO_30     GGCRGHEVED-VDLELFNTSVQLQPPPTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFA 226
SEQ_ID_NO_29     GGCRGSEVED-EDLELFNTSVQLRPPSTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFA 224
SEQ_ID_NO_32     GSCNGAEVED-VDLEMFNTIVYIQPIAAAVPETAAFIERLEMEQAQKAKNPQEQKSFFA 225
SEQ_ID_NO_33     GSCRGIEVEDEVLDLVFNTTISVHAPVTPVPETAPYIERMEMEMEKKCKNPQEQKSFFA 221
                  *.*. * **** *:*:*:*: : * :*. **** :*:*:** :*.*****

SEQ_ID_NO_5      KYWNYIIPVVVLFNMSCAPDTGGQGGGGGGGGGGGSGR 262
SEQ_ID_NO_28     KYWNYIIPVVVLFNMSCAPDTGGQGGGGGGGGGGGSGR 262
SEQ_ID_NO_31     KYWNYIIPVVVLFNMSCAPDTGGQGGGGGGGGGGGSGR 262
SEQ_ID_NO_30     KYWNYIIPVVVLFNMSCAPDAGGQGGGGGGG----- 258
SEQ_ID_NO_29     KYWNYIIPVVVLFNMSCAPDAGGQGGGGGGGSGR---- 258
SEQ_ID_NO_32     KYWNYIIPVVVLFNMSCADAGNQGNGGGGGGGGSGR 263
SEQ_ID_NO_33     KYWYLILGCAVFLMATSSAQTP-PGCAREQS----- 251
                  *** *: .:*** :.: : : ** .

```

图10

```

SEQ_ID_NO_5      --MAAASAGATRLLLLLLLMAYAAPSRARGSSCRAGTGARGAGAGAEGREGEACGTVGLLLEH 58
SEQ_ID_NO_28     --MAAASAGATRLLLLLLLMAYAAPSRARGSSCRAGTGARGAGAGAEGREGEACGTVGLLLEH 58
SEQ_ID_NO_31     --MAAAGAGATRLLLLLLLMAAAPSRARGSSCRAGAAATRGAGAE GRENEGCGTGTGLLLEH 58
SEQ_ID_NO_30     MYMAASGASASRLLLLLLLIVAAAPSRARGSSCRAGAAARGVGAGYEGREGESCGTGTGLLLEH 60
SEQ_ID_NO_29     --MYAAGAGVTRLVLVLLMVAAPSRARGSSCRVGASARGTGADGREAGCGGTVALLEH 58
                  *.*.*.*;***;***;..*****.***.*.*;*.**;*** *.****.*****

SEQ_ID_NO_5      SFEIDDSANFRKRKGSLLUNQODGTL$LSQRQLSEEEERGLRDVAALNGLYRVRIPRRPGA 118
SEQ_ID_NO_28     SFEIDDSANFRKRKGSLLUNQODGTL$LSQRQLSEEEERGLRDVAALNGLYRVRIPRRPGA 118
SEQ_ID_NO_31     SFEIDANHFYKRKGSLLUNQODGTL$LSQRQLSEEEERGLRDVAALNGLYRVRVPRRPGA 118
SEQ_ID_NO_30     SFEIDDAHFYKRKGSLLUNQODGTL$PSQRQLSEEEERGLRDVAALNGLYRVRVPRRPGA 120
SEQ_ID_NO_29     SFEILOGANFPQKRKGSLLUNQODGTL$ATQRQLSEEEERGLRDVAALNGLYRVRVPRRPGT 118
                  ***.*.*.*;*****.*****;*****;*****;

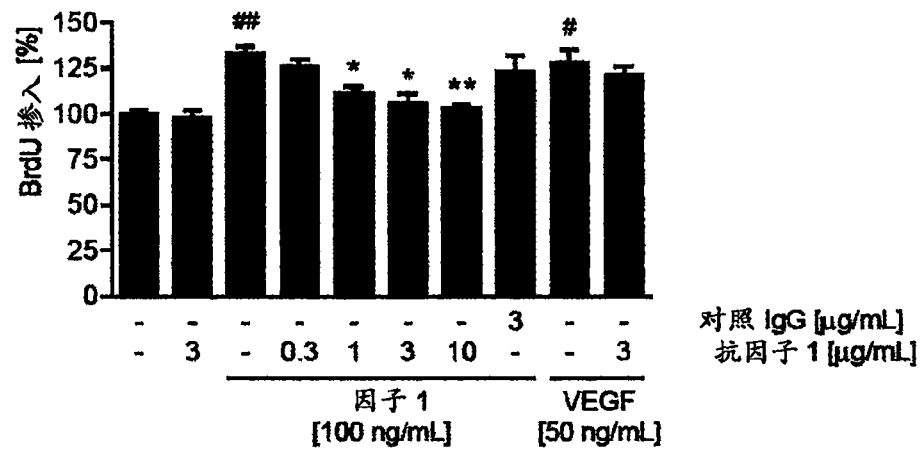
SEQ_ID_NO_5      LDGLEAGGYVSSFFVACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDL 178
SEQ_ID_NO_28     LDGLEAGGYVSSFFVACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDL 178
SEQ_ID_NO_31     PEGPEAGGYVSSFFVACSLVESHLSDQLTLHVDVGVGVVGVSVVTHPGGCRGHEVEDVDL 178
SEQ_ID_NO_30     LDSSEAGGYVSSFFVACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDL 180
SEQ_ID_NO_29     LDGSEAGGHVSSFFVACSLVESHLSDQLTLHVDVAGNVVGLSVVVTHPGGCRGHEVEDDL 178
                  :. ****;*****.*****;***. ***** **

SEQ_ID_NO_5      ELFNITSVQLQPPITAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYUMYIIPVVLFLM 238
SEQ_ID_NO_28     ELFNITSVQLQPPITAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYUMYIIPVVLFLM 238
SEQ_ID_NO_31     ELFNITVQLQPPITAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYUMYIIPVVLFLM 238
SEQ_ID_NO_30     ELFNITSVQLQPPITAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYUMYIIPVVLFLM 240
SEQ_ID_NO_29     ELFNITSVQLRPPSTAPGPETAAFIERLEMEQAQKAKNPQEOKSFFAKYUMYIIPVVLFLM 238
                  ***.*.*;***;.. *****.*****.*****.*****.*****

SEQ_ID_NO_5      MSGAPDTGGQGGGGGGGGGGGSSR 262
SEQ_ID_NO_28     MSGAPDTGGQGGGGGGGGGGGSSR 262
SEQ_ID_NO_31     MSGAPDTGGQGGGGGGGGGGGSSR 262
SEQ_ID_NO_30     MSGAPDAGGQGGGGGGGGG----- 258
SEQ_ID_NO_29     MSGAPDAGGQGGGGGGGGGSSR---- 258
                  *****.*****

```

A



B

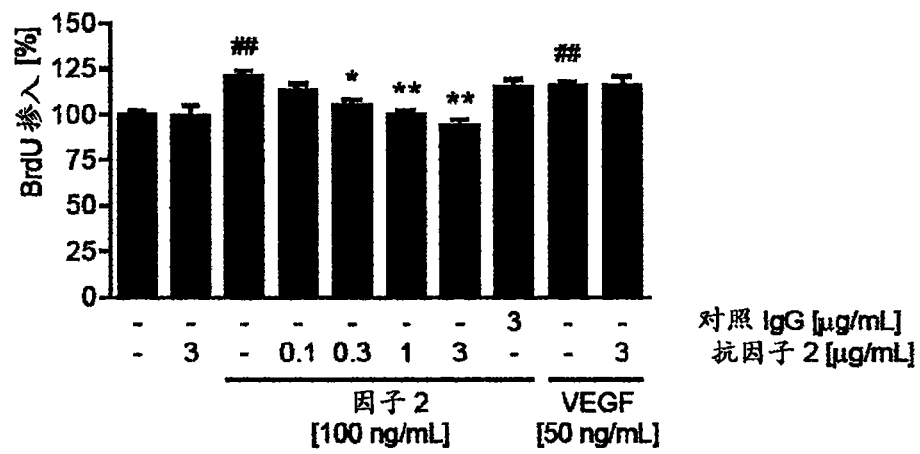


图12