

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 96197233.5

[51] Int. Cl.

C07K 14/00 (2006.01)

A61K 38/00 (2006.01)

[45] 授权公告日 2006 年 11 月 8 日

[11] 授权公告号 CN 1283659C

[22] 申请日 1996.9.16 [21] 申请号 96197233.5

[30] 优先权

[32] 1995. 9. 28 [33] US [31] 08/535,681

[86] 国际申请 PCT/US1996/014915 1996.9.16

[87] 国际公布 WO1997/011964 英 1997.4.3

[85] 进入国家阶段日期 1998.3.26

[71] 专利权人 安姆根有限公司

地址 美国加利福尼亚州

[72] 发明人 胡小芬

审查员 丁 海

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 卢新华 姜建成

权利要求书 4 页 说明书 71 页 附图 9 页

[54] 发明名称

截短的神经胶质细胞系来源的神经营养因子

[57] 摘要

本发明公开了新的蛋白质，称为截短的神经胶质细胞系来源的神经营养因子(截短的 GDNF)蛋白，它促进多巴胺被多巴胺能细胞吸收且促进神经细胞的存活。还公开了经重组遗传工程技术获得截短的 GDNF 蛋白质的方法。

1. 由氨基酸序列 X-[Cys⁴¹-Cys¹³³]-Y 组成的截短的神经胶质细胞系来源的神经营养因子 (GDNF) 蛋白质产物,

其中: [Cys⁴¹-Cys¹³³]由 SEQ ID NO: 2 的 Cys⁴¹ 到 Cys¹³³ 构成;

Y 代表 Cys¹³³ 的羧基末端基团或 Ile¹³⁴ 的羧基末端氨基酸残基;

X 代表 Cys⁴¹ 的甲硫氨酰化的或非甲硫氨酰化的氨基或选自下列基团的氨基末端氨基酸残基:

NRG
 KNRG (SEQ ID NO:3)
 GKNRG (SEQ ID NO:4)
 RGKNRG (SEQ ID NO:5)
 QRGKNRG (SEQ ID NO:6)
 GQRGKNRG (SEQ ID NO:7)
 RGQRGKNRG (SEQ ID NO:8)
 RRGQRGKNRG (SEQ ID NO:9)
 G RRGQRGKNRG (SEQ ID NO:10)
 KG RRGQRGKNRG (SEQ ID NO:11)
 GKKG RRGQRGKNRG (SEQ ID NO:12)
 RGKG RRGQRGKNRG (SEQ ID NO:13); 和
 SRGKG RRGQRGKNRG (SEQ ID NO:14)

并且其中所说的 GDNF 蛋白质产物具有提供对多巴胺能细胞的神经营养效应的能力。

2. 根据权利要求 1 的截短的 GDNF 蛋白质产物,其中 X = SRGKG RRGQRGKNRG (SEQ ID NO: 14)。

3. 根据权利要求 1 的截短的 GDNF 蛋白质产物,其中 X = RGQRGKNRG (SEQ ID NO: 8)。

4. 根据权利要求 1 的截短的 GDNF 蛋白质产物,其中 X = GQRGKNRG (SEQ ID NO: 7)。

5. 根据权利要求 1 的截短的 GDNF 蛋白质产物,其中 X = KNRG(SEQ ID NO: 3)。

6. 根据权利要求 1 的截短的 GDNF 蛋白质产物,其中 X = NRG。

7. 根据权利要求 1-6 中任一项的截短的 GDNF 蛋白质产物,其中氨基酸序列与至少一种水溶性多聚体结合。

8. 编码根据权利要求1的截短的 GDNF 蛋白质的多核苷酸。
9. 一种载体, 包含可操作地连接到表达控制序列上的权利要求8的多核苷酸。
10. 用权利要求8的多核苷酸转化或转染的原核或真核宿主细胞。
11. 制备截短的 GDNF 蛋白质的方法, 包含在合适的营养培养基中培养权利要求10的宿主细胞并任选从所说的细胞或所说的营养培养基中分离所说的截短的 GDNF。
12. 根据权利要求11的制备截短的 GDNF 蛋白质的方法, 其中所说的宿主细胞是大肠杆菌。
13. 根据权利要求11的制备截短的 GDNF 蛋白质的方法, 其中所说的宿主细胞是中国仓鼠卵巢细胞。
14. 制备截短的神经胶质细胞系来源的神经营养因子 (GDNF) 蛋白质的方法, 包含下列步骤:
 - (a) 培养用权利要求9的载体转化或转染的原核或真核宿主细胞;
 - (b) 在允许所说的宿主细胞表达截短的 GDNF 蛋白质的条件下维持所说的宿主细胞; 和
 - (c) 任选分离由所说的宿主细胞表达的截短的 GDNF 蛋白质。
15. 一种截短的 GDNF 蛋白质, 它是含权利要求8的外源性多核苷酸的原核或真核宿主细胞的重组表达产物。
16. 一种药物组合物, 包含与药学上可接受的载体结合的根据权利要求1的截短的 GDNF 蛋白质产物。
17. 一种药物组合物, 包含与药学上可接受的载体结合的根据权利要求11的方法生产的截短的 GDNF 蛋白质。
18. 一种药物组合物, 包含与药学上可接受的载体结合的、根据权利要求14的方法产生的截短的 GDNF 蛋白质。
19. 权利要求16的药用组合物在制备用于治疗帕金森氏病的药物中的应用。
20. 权利要求8的多核苷酸序列在制备用于治疗帕金森氏病的方法的药物中的应用, 所述方法包含给病人施用权利要求8的多核苷酸序列以提供体内生产所说的截短的 GDNF 蛋白质。
21. 权利要求8的多核苷酸序列在制备用于治疗帕金森氏病的方法的药物中的应用, 所述方法包含给病人植入用权利要求8的多核苷酸序列转

化的细胞以提供体内生产所说的截短的 GDNF 蛋白质。

22. 权利要求 1 的神经营养因子 (GDNF) 蛋白质产物, 它从重组修饰的细菌或哺乳动物细胞表达的成熟 GDNF 蛋白质产生。

23. 根据权利要求 22 的神经营养因子 (GDNF) 蛋白质产物, 其中 X 选自下列序列:

G
RG
NRG
KNRG (SEQ ID NO:3)
GKNRG (SEQ ID NO:4)
RGKNRG (SEQ ID NO:5)
QRGKNRG (SEQ ID NO:6)
GQRGKNRG (SEQ ID NO:7)
RGQRGKNRG (SEQ ID NO:8) 和
RRGQRGKNRG (SEQ ID NO:9)

24. 根据权利要求 22 的神经营养因子 (GDNF) 蛋白质产物, 其中成熟的 GDNF 蛋白质由重组修饰的细菌细胞表达且所说的截短的 GDNF 蛋白质在体外或体内生产。

25. 一种制备药物组合物的方法, 其中治疗上有效量的根据权利要求 1 的截短的 GDNF 蛋白质产物与一种或多种药学上可接受的载体混合。

26. 根据权利要求 1 的截短的 GDNF 蛋白质产物在治疗由疾病或损伤引起的神经系统的损伤中的应用。

27. 根据权利要求 26 的截短的 GDNF 蛋白质产物的应用, 用于治疗帕金森氏病。

28. 根据权利要求 1 的截短的 GDNF 蛋白质产物在制备治疗由疾病或损伤引起的神经系统的损伤的药物组合物中的应用。

29. 根据权利要求 1 的截短的 GDNF 蛋白质产物, 其中所说的蛋白质产物从重组修饰的宿主细胞表达的 SEQ ID NO:2 的成熟 GDNF 氨基酸序列产生。

30. 神经胶质细胞系来源的神经营养因子 (GDNF) 蛋白质产物, 包含 SEQ ID NO:2 的成熟 GDNF 氨基酸序列的二聚体和权利要求 1 的截短的 GDNF 氨基酸序列, 其中所说的二聚体对多巴胺能神经细胞有神经营养

效应。

31. 神经胶质细胞系来源的神经营养因子（GDNF）蛋白质产物，包含两个权利要求 1 的截短的 GDNF 氨基酸序列的二聚体，其中所说的二聚体对多巴胺能神经细胞有神经营养效应。

32. 根据权利要求 31 的神经胶质细胞系来源的神经营养因子（GDNF）蛋白质产物，其中所说的二聚体是同二聚体。

33. 根据权利要求 31 的神经胶质细胞系来源的神经营养因子（GDNF）蛋白质产物，其中所说的二聚体是异二聚体。

34. 根据权利要求 1-6 中任一项的截短的 GDNF 蛋白质产物，其中所说的氨基酸序列是糖基化的。

截短的神神经胶质细胞系来源 的神经营养因子

发明领域

一般来说,本发明涉及本文称为神经胶质细胞系来源的神经营养因子(也称为胶质来源的神经营养因子或 GDNF)的蛋白质,其特征在于促进多巴胺被多巴胺神经元的吸收及支持在帕金森氏病中死亡的神经元的存活的能力。更具体地说,本发明涉及新的截短的 GDNF 蛋白质。

发明背景

神经营养因子是在神经系统中或由神经系统支配的非神经组织中发现的蛋白质,其功能是促进存活和维持神经和/或神经胶质细胞的表型分化(Varon 等,神经科学年鉴 1:327, 1979; Thoenen 等,科学, 229: 238, 1985)。由于该生理学作用,神经营养因子在治疗各种神经退化性疾病中出现的神经细胞的退化和分化功能的丧失中 useful。

为了便于将具体的神经营养因子有效地用于治疗神经损伤,损伤的神经细胞的类型或种类应对该因子有效应。不同的神经营养因子典型地影响明显不同类型的神经细胞。因此,拥有各种不同的神经营养因子来治疗在不同形式的疾病或损伤中出现的各种类型的损伤的神经元具有优势。

神经营养因子可保护效应神经元抵抗各类无关损害。例如,神经生长因子(NGF)从经过切断其轴突突起引起的死亡中(Rich 等,神经细胞学杂志 16: 261 1987; Otto 等,神经科学杂志: 83: 156, 1987),从胚胎发育期间的个体发育死亡中(Hamburger 等,神经科学杂志: 4:767, 1984)和从施用紫杉醇或顺式铂氨引起的损伤中(Apfel 等,神经学年鉴, 29: 87, 1991)挽救有效部分的感觉神经元。保护的明显普遍性产生了这样的观点:如果一种神经营养因子保护效应神经元抵抗实验性损伤,那么它可用于治疗病人中涉及这些神经元损伤的疾病,即使该病原学可能是未知的。

一种给定的神经营养因子除了具有正确的神经元特异性外,还必须能获得足够量以用作药物治疗。由于神经营养因子在组织中典型地以少量存在(例如, Hofer 和 Barde, 自然: 331: 261, 1988; Lin 等,科学 246: 1023 1989)可能不便于直接从动物组织制备药用量的神经营养因

子。作为选择，需要使用重组表达系统生产所需的蛋白质。

Lin 等以前描述了筛选对灰质多巴胺能神经元的胚胎前体具有神经营养活性的生物学样品的方法（参见 1994 年 5 月 23 日申请的美国专利申请号 08/182,183 及其专利申请；1992 年 9 月 17 日申请的 PCT/US92/07888（WO 93/06116）和欧洲专利申请号 92921022.7（公开号 EP610254），本文引用其说明书以供参考）。该生物测定在鉴定可用于治疗帕金森氏病的神经营养因子中（Friedman 等，神经科学通讯，79: 65-72, 1987），因为该疾病的特征在于支配纹状体的中脑中的多巴胺能神经元的退化。

Lin 等进一步描述了从一种该来源，成神经胶质瘤细胞系 B49 的条件培养基中纯化的新神经营养因子的鉴定（Schubert 等，自然 249:224-27, 1974）。以前报道了该细胞系的条件培养基含多巴胺能神经营养活性（Bohn 等，神经科学协会文摘，15:277, 1989）。在 Lin 等的文献前，神经胶质细胞系来源的神经营养因子（GDNF）尚未作为分离的生物学活性物质被鉴定或作为基本上纯的蛋白质被分离。另外，Lin 等描述了克隆编码 GDNF 的人基因的方法，编码 GDNF 的人基因的核酸序列和 GDNF 蛋白质的氨基酸序列。GDNF 基因被亚克隆进表达载体，且该载体用于表达生物学活性 GDNF。GDNF 蛋白质是由 2 个以二硫键连接的 134 氨基酸，22kDa 的亚基组成的同源二聚体。该描述进一步包括将 GDNF 用于防止和治疗神经损伤和诸如帕金森氏病的神经相关疾病。

GDNF 疗法有助于治疗由损伤一种或多种类型的神经细胞的存活和/或合适功能的情况引起的神经损伤。该神经损伤可从许多种类的不同原因引起。神经损伤可对一种或多种类型的神经细胞产生，它经过：

- （1）物理损伤：引起靠近损伤位点的轴突突起和/或神经细胞体退化；
- （2）部分神经系统的临时或永久性血流停止，如中风；
- （3）故意或碰巧接触神经毒素，如用于治疗癌症的化学治疗剂（例如，顺式铂氨）或用于治疗 AIDS 的双脱氧胞苷（ddC）；
- （4）慢性代谢疾病，如糖尿病或肾机能障碍；或
- （5）神经退化疾病，如帕金森氏病，Alzheimer 氏疾病和肌萎缩性侧索硬化（ALS），它们由特定神经元群体的退化引起。

GDNF 治疗特别有助于治疗涉及灰质多巴胺能神经元退化的神经

退化性症状, 如帕金森氏病。对于帕金森氏病的唯一流行疗法是治标, 针对纹状体中增加的多巴胺水平。GDNF 疗法预期的效果不是简单地在纹状体多巴胺能神经末端产生多巴胺能神经传递的增加(这将导致症状的减轻), 而且还减缓了或甚至终止了退化性突起的进展并修复损伤的黑质纹状体途径并恢复其功能。GDNF 还可用于治疗病人的其它形式的多巴胺能神经细胞的损伤或不适当的功能。这种损伤或机能失常可能在精神分裂症和其它形式的精神病中出现。对这类症状的唯一常规疗法是针对有症状的且需要作用于多巴胺吸收位点的多巴胺受体, 这与该疾病过程可能涉及支配具有这些受体的神经元群体的多巴胺能神经元具有不合适的功能相一致。

发明概述

在一个方面, 本发明提供了新的截短的神经胶质细胞系来源的神经营养因子 (GDNF) 蛋白产物。在一个实施方案中, 以重组遗传工程技术生产截短的 GDNF 蛋白质。在一个另选的实施方案中, 以化学技术, 或重组与化学技术的结合合成该截短的 GDNF 蛋白质。

本发明截短的 GDNF 蛋白质产物包括由氨基酸序列 X-[Cys⁴¹-Cys¹³³]-Y 代表的蛋白质。图 1 (SEQ ID NO: 2) 中氨基酸残基的编号方式用于帮助比较成熟 GDNF 蛋白质。[Cys⁴¹-Cys¹³³]代表图 1 (SEQ ID NO: 2) 所示的 Cys⁴¹ 到 Cys¹³³ 的氨基酸序列。Y 代表 Cys¹³³ 的羧基末端基团或 Ile¹³⁴ 的羧基末端氨基酸残基。X 代表 Cys⁴¹ 的甲硫氨酰化或非甲硫氨酰化的氨基或选自下列基团的氨基末端氨基酸残基。

G
RG
NRG
KNRG (SEQ ID NO:3)
GKNRG (SEQ ID NO:4)
RGKNRG (SEQ ID NO:5)
QRGKNRG (SEQ ID NO:6)
GQRGKNRG (SEQ ID NO:7)
RGQRGKNRG (SEQ ID NO:8)
RRGQRGKNRG (SEQ ID NO:9)
G RRGQRGKNRG (SEQ ID NO:10)

	KG	RRGQRGKNRG	(SEQ ID NO:11)
	GKG	RRGQRGKNRG	(SEQ ID NO:12)
	RGKG	RRGQRGKNRG	(SEQ ID NO:13)
	SRGKG	RRGQRGKNRG	(SEQ ID NO:14)
	NSRGKG	RRGQRGKNRG	(SEQ ID NO:15)
	ENSRGKG	RRGQRGKNRG	(SEQ ID NO:16)
	PENSRGKG	RRGQRGKNRG	(SEQ ID NO:17)
	NPENSRGKG	RRGQRGKNRG	(SEQ ID NO:18)
	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:19)
A	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:20)
AA	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:21)
AAA	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:22)
QAAA	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:23)
RQAAA	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:24)
NRQAAA	ANPENSRGKG	RRGQRGKNRG	(SEQ ID NO:25)
	RNRQAAA	ANPENSRGKG	(SEQ ID NO:26)
	ERNRQAAA	ANPENSRGKG	(SEQ ID NO:27)
	RERNRQAAA	ANPENSRGKG	(SEQ ID NO:28)
	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:29)
P	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:30)
LP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:31)
VLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:32)
AVLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:33)
MAVLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:34)
QMAVLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:35)
KQMAVLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:36)
DKQMAVLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:37)
PDKQMAVLP	RRERNRQAAA	ANPENSRGKG	(SEQ ID NO:38)

预期该截短的 GDNF 蛋白产物包括具有由 X-[Cys⁴¹-Cys¹³³]-Y 所代表的氨基酸序列的截短的 GDNF 蛋白质和其变异体和衍生物。因此, 本发明的截短的 GDNF 蛋白质产物还包括由 X-[Cys⁴¹-Cys¹³³]-Y 所代表的氨基酸序列的添加, 取代和内部缺失变异体及衍生物。截短的 GDNF 蛋白质产物进一步包括甲硫氨酰化或非甲硫氨酰化的形式以及截短的 GDNF 蛋白的糖基化或非糖基化形式。

本发明的举例性的截短的 GDNF 蛋白质包括[Arg¹⁶-Ile¹³⁴], [Asn²²-Ile¹³⁴], [Pro²³-Ile¹³⁴], [Ser²⁶-Ile¹³⁴], [Arg³²-Ile¹³⁴], [Gly³³-Ile¹³⁴], [Lys³⁷-Ile¹³⁴]

和[Asn³⁸-Ile¹³⁴] 截短的 GDNF 蛋白质, 甲硫氨酰化或非甲硫氨酰化形式及其变异体或衍生物。目前本发明优选的截短的 GDNF 蛋白质包括 [Lys³⁷-Ile¹³⁴]和[Asn³⁸-Ile¹³⁴]截短的 GDNF 蛋白质, 甲硫氨酰化或非甲硫氨酰化形式及其变异体或衍生物。举例性的取代变异体是[Asn²²ΔSer²²-Ile¹³⁴]和[Pro²³-Lys³⁷ΔAsn³⁷-Ile¹³⁴]截短的 GDNF 蛋白质。举例性的添加变异体是 Ser-[Por²³-Ile¹³⁴]截短的 GDNF 蛋白质。

在本发明的另一方面, 可以以糖基化的或非糖基化的形式制备截短的 GDNF 蛋白质。截短的 GDNF 蛋白质衍生物典型地包括将截短的 GDNF 蛋白质附着于水溶性多聚体上。例如, 截短的 GDNF 蛋白质可结合到一个或多个聚乙二醇分子上以减小截短的 GDNF 蛋白质产物在水环境中的沉淀。

本发明还有另一个方面包括编码截短 GDNF 蛋白质的各种多核苷酸。这些核酸序列一般用于在真核或原核宿主细胞中表达截短的 GDNF, 其中该表达产物或其衍生物的特征在于减少多巴胺被多巴胺能细胞吸收的能力。该多核苷酸也可用于细胞治疗或基因治疗的应用中。合适的核酸序列包括在附图中被具体描述的序列以及其它的简并序列和天然存在的等位基因变异体。

本发明的另一方面包括含有可操作地连接到扩增和/或表达控制序列上的编码截短的 GDNF 蛋白质的多核苷酸的载体。用该载体稳定转化或转染原核和真核宿主细胞以表达截短的神经胶质细胞来源的神经营养因子。本发明进一步包括截短的 GDNF 蛋白质的重组生产, 其中该转化的或转染的宿主细胞在合适的营养培养基中生长, 且任选从宿主细胞和/或营养培养基中分离由该细胞表达的截短的 GDNF。本发明进一步包括编码截短的 GDNF 的多核苷酸和含有该多核苷酸的载体在基因治疗或细胞治疗中的应用。

另一方面, 本发明涉及含有成熟 GDNF 蛋白质和一种或多种从其得来的截短的 GDNF 蛋白质的混合物的重组产生的 GDNF 组合物, 其中成熟 GDNF 蛋白质分子量为大约 44kDa, 且其中截短的 GDNF 蛋白质分子量大约为 36 到 40kDa。GDNF 组合物可含至少两种截短的 GDNF 种类, 其中第一类分子量大约为 36kDa, 第二类分子量大约为 40kDa。分子量为大约 40kDa 的截短的 GDNF 种类是具有分子量大约为 22kDa 的 GDNF 单体和具有分子量大约为 18kDa 的截短的 GDNF 单体的异型

二聚体。还预期可从该混合物中分离一种或多种截短的 GDNF 种类用于治疗用途。

本发明的另一方面包括含有截短的 GDNF 蛋白产物的药用组合物。典型的是，与药学上可接受的载体一起配制截短的 GDNF 蛋白质产物。可使用各种其它的制剂物质以利于生产，贮存，处理，传递和/或效能。在本发明的另一方面，截短的 GDNF 蛋白质产物增加多巴胺吸收和多巴胺能神经元的存活。因此，截短的 GDNF 蛋白质产物特别适用于治疗由损伤或疾病，如帕金森氏病，引起的神经系统损伤。

根据下面对本发明的实施的详细描述，本发明的其它方面和优点对本领域的技术人员而言将是显而易见的。

附图简述

参考附图，本发明的许多特征和优点将变得显而易见，其中：

图 1 描述了编码成熟的人神经胶质细胞系来源的神经营养因子（hGDNF）的核苷酸序列（SEQ ID NO:1）。还描述了成熟人 GDNF 蛋白的氨基酸序列（SEQ ID NO:2）。

图 2 描述了为表达重组截短 GDNF 蛋白而制备的质粒构建体的图示。

图 3 描述了编码 GDNF 的另一核酸序列（SEQ ID NO:39）和截短的 GDNF 多核苷酸的限制性图谱。

图 4 描述了编码 GDNF 的另一个核酸序列（SEQ ID NO:40）和截短的 GDNF 多核苷酸的限制性图谱。

图 5 描述了编码[Pro²³-Lys³⁷ΔAsn³⁷-Ile¹³⁴]截短的 GDNF 蛋白质取代变异体（SEQ ID NO:42）的核酸序列（SEQ ID NO:41）。该蛋白质也可描述成 Met-Ser-[Pro²³-Lys³⁷ΔAsn³⁷-Ile¹³⁴]截短的 GDNF 蛋白质添加/取代变异体。

图 6 描述了编码[Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质（SEQ ID NO:44）的核酸序列（SEQ ID NO:43）。

图 7 描述了编码[Gly³³-Ile¹³⁴]截短的 GDNF 蛋白质（SEQ ID NO:46）的核酸序列（SEQ ID NO:45）。

图 8 描述了成熟 hGDNF 的氨基酸序列（SEQ ID NO:47）与一些举例性的截短 GDNF 蛋白质：Met-[Arg³²-Ile¹³⁴]（SEQ ID NO:48）。Met-[Gly³³-Ile¹³⁴]（SEQ ID NO:49）和 Met-Ser-[Pro²³-Lys³⁷ΔAsn³⁷-Ile¹³⁴]

(SEQ ID NO:50) 的比较。

发明详述

人神经胶质细胞系来源的神经营养因子 (hGDNF) 以前体合成, 该前体被加工并分泌为 134 个氨基酸的成熟蛋白质。以前测定了成熟人 GDNF 具有图 1 所示的氨基酸序列 (SEQ ID NO:2) 。

本发明基于这样一种意想不到的发现: 成熟 GDNF 蛋白质大小可减小 (本文也称为“剪短”或“截短”蛋白质或截短 GDNF 蛋白质), 但仍保留其生物学活性。剪短蛋白在中国仓鼠卵巢 (CHO) 细胞重组生产 GDNF 期间首次发现。简单地说, 重组人 GDNF (rhGDNF) 制备如下。将编码成熟人 GDNF 蛋白的完整开放阅读框的核酸序列克隆进表达质粒。证实该核酸序列的正确性 (DNA 测序为相当于 GeneBank 中 hGDNF 序列) 并翻译成与成熟人 GDNF 公开序列相同的氨基酸序列 (Lin 等, 科学 260, 1130-1132, 1993)。将质粒 DNA 线性化并使用磷酸钙沉淀法转染入二氢叶酸还原酶缺陷型 CHO 细胞 (CHO⁻细胞) 中。在选择培养基中培养转染的细胞, 选择在选择过程中存活的那些菌落用于分别分析 hGDNF 表达。

收集来自单个克隆的无血清条件培养基并使用对 hGDNF 特异性的抗血清进行 Western 印迹分析。该抗血清包含从用在大肠杆菌中表达的重组 hGDNF 免疫的兔中得到的兔多克隆抗体。在还原条件下, 存在于这些样品中的 hGDNF 分辨出具有表观分子量为大约 22kDa 和 18kDa 的两条主要的带。每条带分别由大约 22 + 22.5kDa 和 18 + 17.5kDa 的位置接近的双联体组成(为简单起见, 这些双联体称为 22kDa 和 18kDa 带或种类)。

以前报道了 GDNF 以由分子量为大约 20 到 22kDa 的成熟 GDNF 蛋白的两个相同亚基组成的二硫键键合的同型二聚体存在。在非还原条件下分析 GDNF 时, 据报道已鉴定出 32 到 42kDa (Lin 等, 科学, 260, 1130-1132, 1993) 或 33 到 45kDa (Lin 等, 神经化学杂志, 63(2), 758-768, 1994) 的宽带。该范围的存在解释为由于在成熟单体上存在糖基化的不均一性, 且进一步用去糖基化实验证实。

尽管该 22kDa 带相应于文献报道的成熟 GDNF 蛋白质, 但以前未报道 18kDa 的带。在从不同克隆收集的样品中 22kDa 和 18kDa 蛋白质的相对量有变化。另外, 发现从同一克隆的多次收获物中显示出两种带

的比例可变化。而且,发现 CHO 表达的 GDNF 蛋白质的贮存通常导致 18kDa 带的存在增加,同时 22kDa 带减少。

以 Western 印迹在非还原条件分析转化的 CHO^d 细胞的条件培养基时,观察到具有表观分子量为 36, 40 和 44kDa 的 3 种高分辨带。该发现也与以前的报道相反。这些带的相对强度可变,但它们与存在于各样品中的 22 和 18kDa 单体带的比例非常相关。用单克隆抗血清进一步分析时,确定了在非还原凝胶中的 3 条带相应于由两个单体组成的 3 种可能的二聚体。最大的 44kDa 蛋白质是以前所报道的 22kDa 成熟 GDNF 蛋白质的二聚体。中间的 40kDa 蛋白质由二聚体组成,其中一个成熟蛋白质分子量减少为 18kDa 的形式。最小的 36kDa 二聚体似乎含 2 个 18kDa 蛋白质,即, 2 个 22kDa 形式分子量减小。这些数据首次证实在二聚体构型中不仅存在 GDNF 单体的新形式,而且还存在剪短的 GDNF 蛋白质。贮存时还发现,样品的单体组成向剪短形式转变,且可见相应的二聚体种类,即 36kDa 蛋白质的量增加。

然后进行研究以鉴定与以前报道的成熟 GDNF 蛋白质相比哪一部分蛋白质被消除或改变,从而引起了分子量的减小。首先确定了分子量的减小不是由于糖基化的改变而引起。

GDNF 含 2 个潜在的 N - 连接的糖基化位点并已报道被糖基化。然而,剪短蛋白并不简单地是成熟 GDNF 的非糖基化的或糖基化不足的形式。在去糖基化实验中证实了这一点,其中用 N - 聚糖酶, O - 聚糖酶和神经氨酸苷酶处理样品。在还原凝胶上,以 N - 聚糖酶消化使 18kDa 蛋白质减小成 13.5kDa 的带,这表明存在相当于 4.5kDa 的 N - 连接的糖。用神经氨酸苷酶和 O - 聚糖酶处理导致 18kDa 带略微改变成 17kDa。这表明在蛋白质上存在 O - 连接的糖。已报道成熟的 22kDa 带被糖基化且也被 N - 葡聚糖酶减小到 18kDa (即,也有 4.5kDa)。通过使用对凝胶上的 22kDa 带特异性的单克隆抗体进一步证实了这点。未还原的二聚体的聚糖酶消化方式更复杂,但能解释且与 3 种形式的最初说明一致。

结果,认为蛋白质分子量减小 4.5kDa 是由于缺失大约 30 - 35 个氨基酸残基而不是由糖基化改变引起。由于下列原因预期缺失很可能在成熟 GDNF 蛋白质的氨基末端发生。成熟 GDNF 含总共 7 个胱氨酸。如果缺失从羧基末端,将失去 7 个胱氨酸中的 2 到 4 个,且这可能导致

蛋白质失活。然而，当对主要由剪短形式组成的试验样品进行生物测定以测量其多巴胺能神经元神经营养活性时，该样品表现出的活性比得上在比例上含更多 GDNF 成熟形式的样品。

然后经过对纯化蛋白的氨基酸序列分析测定裂解位点。根据厂家说明书，使用应用生物系统 494A 蛋白质测序仪经 10 个循环测序样品。尽管氨基酸序列分析技术和程序对本领域的技术人员是熟知的，但在 Fausset 等，电泳 12:22-27, 1991 和 1990 年 8 月 24 日申请的美国专利申请号 576316（欧洲专利申请号 90310899，公开号 EP 423980，申请日 1990 年 10 月 4 日，题目“干细胞因子”）中提供了蛋白质测序的进一步的描述，本文引用这些文献作为参考文献。经过分析，测定了剪短蛋白的氨基末端是“RGQRGK”或 Arg-Gly-Gln-Arg-Gly-Lys。因此在条件培养基中从成熟蛋白质去掉了前 31 个氨基酸。此外，以氨基酸 Arg³² 开始的剪短蛋白质的剩余氨基酸序列与图 1（SEQ ID NO:2）所示的成熟 GDNF 氨基酸序列一致。

在多巴胺能神经元试验中，发现[Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质在定性基础上有活性。多巴胺能神经营养活性试验用于鉴定可能在治疗帕金森氏病中有益的神经营养因子。该试验以以前描述的试验（Friedman 等，神经科学通讯，79:65-72, 1987，本文引用该文献以供参考）为基础且可包括在 Lin 等（见 1994 年 5 月 23 日申请的美国专利申请号 08/182,183 和其母申请；1992 年 9 月 17 日申请的 PCT/US92/07888（WO 93/06116）；和欧洲专利申请号 92921022.7）（公开号 EP 610254）所述的修改。该试验的详细描述在下面实施例 5 中提供。

随后的纯化程序中经氨基酸测序后发现另一种蛋白质，其中成熟 GDNF 的 N 端被去掉了前 36 个氨基酸残基，得到[Lys³⁷-Ile¹³⁴]截短的 GDNF 蛋白质，N 端序列为 KNRG(C)VL--。此外剪短蛋白的剩余氨基酸残基与成熟人 GDNF 氨基酸序列一致。在多巴胺吸收生物测定中还分析了[Lys³⁷-Ile¹³⁴]截短的 GDNF 蛋白质。发现截短的 GDNF 蛋白质具有活性，ED₅₀ 为大约 50 pg/ml，相似于纯化的重组大肠杆菌表达的成熟 GDNF。

进一步发现细菌表达的成熟 GDNF 可变成截短形式。在转化的大肠杆菌中表达的成熟 GDNF（如在 Lin 等，美国专利申请号 08/182,183，出

处同上)用 CHO - 产生的条件培养基培养。重组大肠杆菌 GDNF 在还原凝胶上的表观分子量为 17kDa。当该物质与 CHO 细胞条件培养基混合并在 4℃ 培养 5 天时,该蛋白质完全剪短为 12.5kDa。在培养 1 小时或 24 小时时这种裂解不太完全,表明在该条件下是一种时间依赖性过程。还发现用含 0.1%胎牛血清的培养基简单培养重组大肠杆菌 GDNF 过夜不产生剪短形式。因此,培养基中活细胞的存在似乎对剪切加工的发现是必需的。因此可能在体内某些组织内也发生剪短事件。

另外,发现成熟大肠杆菌表达的 hGDNF 的衍生物,如 PEG 化的 GDNF (也在 Lin 等,美国专利申请号 08/182,183,出处同上中描述)在 CHO-产生的条件培养基存在下可被加工成截短形式。成熟 GDNF 可在氨基末端被 PEG 化以延长其在循环中的清除时间。PEG 化增加了蛋白质的大小,在还原条件下修饰的成熟 GDNF 迁移到大约 45kDa。与未 PEG 化的成熟形式一样,PEG 化的大肠杆菌 GDNF 与 CHO 细胞(未转染的)条件培养基培养产生 12.5kDa 的带。在两种情况下,12.5kDa 种类以二硫键结合的二聚体存在,如在非还原凝胶上所示。从 N 端 PEG 化的成熟蛋白产生该剪短形式进一步证实剪短事件在该蛋白质的 N 端发生,因为在剪切过程中 PEG 化的残基丢失。

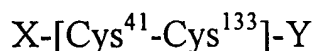
根据这些发现,且由于剪短事件也可在体内发生,因此 GDNF 蛋白质的截短形式可能是在生理条件下 hGDNF 的最终自然加工的形式。因此,认为产生截短的 GDNF 蛋白质或其衍生物用于治疗应用具有优势。例如,直接表达或合成截短的 GDNF 蛋白质,如 [Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质预期可对上述蛋白水解活性产生抗性。而且,如果需要产生截短的 GDNF 衍生物,如 PEG 化的 [Arg³²-Ile¹³⁴]截短 GDNF 蛋白质,预期所得的衍生物具有对成熟 GDNF 衍生物所观察到的特异性剪切不敏感的优势。

也可预期截短的 GDNF 蛋白质产物的其它优势。首先,截短的蛋白质,如 [Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质的 pI 从大约 10 减小到大约 8.0-8.5。这使得蛋白质碱性明显变小,从而提供了有益的效果,包括更好的受体结合和在用药物位点,如膜内注射位点的细胞毒性下降。其次,在成熟 GDNF 的前 26 个氨基酸内,氨基酸序列有两个脱酰胺位点: Arg-Asn-Arg (氨基酸 14 - 16) 和 Glu-Asn-Ser (氨基酸 24 - 26)。在截短 GDNF 蛋白质中缺失一个或两个这类位点预期增加了该蛋白质

的稳定性。

截短的 GDNF 蛋白质产物:

在基本实施方案中可用下面氨基酸序列代表本发明的截短的 GDNF 蛋白质, 其中使用图 1 的氨基酸残基编号方式以利于与成熟 GDNF 蛋白质相比较:



其中:

[Cys⁴¹-Cys¹³³]代表图 1 (SEQ ID NO:2) 所示的从 Cys⁴¹ 到 Cys¹³³ 的氨基酸序列;

Y 代表 Cys¹³³ 的羧基端基团或 Ile¹³⁴ 的羧基端氨基酸残基;

X 代表 Cys⁴¹ 的甲硫氨酰化或非甲硫氨酰化的氨基或选自下列基团的氨基末端氨基酸残基:

	G
	RG
	NRG
	KNRG (SEQ ID NO:3)
	GKNRG (SEQ ID NO:4)
	RGKNRG (SEQ ID NO:5)
	QRGKNRG (SEQ ID NO:6)
	GQRGKNRG (SEQ ID NO:7)
	RGQRGKNRG (SEQ ID NO:8)
	RRGQRGKNRG (SEQ ID NO:9)
G	RRGQRGKNRG (SEQ ID NO:10)
KG	RRGQRGKNRG (SEQ ID NO:11)
GKG	RRGQRGKNRG (SEQ ID NO:12)
RGKG	RRGQRGKNRG (SEQ ID NO:13)
SRGKG	RRGQRGKNRG (SEQ ID NO:14)
NSRGKG	RRGQRGKNRG (SEQ ID NO:15)
ENSRGKG	RRGQRGKNRG (SEQ ID NO:16)
PENSRGKG	RRGQRGKNRG (SEQ ID NO:17)
NPENSRGKG	RRGQRGKNRG (SEQ ID NO:18)
ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:19)
A	ANPENSRGKG RRGQRGKNRG (SEQ ID NO:20)
AA	ANPENSRGKG RRGQRGKNRG (SEQ ID NO:21)
AAA	ANPENSRGKG RRGQRGKNRG (SEQ ID NO:22)

	QAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:23)
	RQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:24)
	NRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:25)
	RNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:26)
	ERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:27)
	RERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:28)
	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:29)
P	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:30)
LP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:31)
VLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:32)
AVLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:33)
MAVLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:34)
QMAVLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:35)
KQMAVLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:36)
DKQMAVLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:37)
PDKQMAVLP	RRERNRQAAA	ANPENSRGKG	RRGQRGKNRG (SEQ ID NO:38)

本文使用的术语“截短的 GDNF 蛋白质产物”包括有生物学活性的合成的或重组的截短 GDNF 蛋白质,从成熟 GDNF 生产的截短的 GDNF 蛋白质,有生物学活性的截短的 GDNF 变异体(包括插入,取代和缺失变异体)及其化学修饰的衍生物。还包括与具有 SEQ ID NO:2 所述氨基酸序列的人 GDNF 蛋白质基本上同源的截短的 GDNF 蛋白质。

本文所用的术语“生物学活性”指截短的 GDNF 蛋白质表现出与具有 SEQ ID NO:2 所述氨基酸序列的 GDNF 蛋白质相似的神经营养特性,但不必是全部相同的特性,且不必是相同的程度。感兴趣的具体神经营养特性的选择取决于施用截短的 GDNF 蛋白质产品的用途。截短的 GDNF 蛋白质产物具有生物学活性,且正如下面实施例中所讨论的,使用评价多巴胺吸收和酪氨酸羟化酶(TH)表达作为举例性生物测定时表现出的多巴胺能神经元存活特征相似于成熟 GDNF 蛋白质所表现出的特征。

本文所用的术语“基本上同源的”指与具有 SEQ ID NO:2 所示的氨基酸序列的人 GDNF 的同源性程度优选超过 70%,更优选超过 80%,甚至更优选超过 90%或甚至 95%。本文所述的同源性百分数计算为,在 100 个氨基酸长度中引入 4 个残基变化以帮助序列比较时在两个序列

的较小序列中发现的待比较的序列中对比为相同氨基酸残基的氨基酸残基的百分数，如 Dayhoff 在蛋白质序列和结构图表集 Vol.5, p.124 (1972)，国家生化研究基金会，华盛顿区中所述，本文引用该文献以供参考。基本上同源还包括用与抗 SEQ ID NO:2 的 GDNF 的抗体的交叉反应性特征可分离的任何截短的 GDNF 蛋白质或其基因可通过与 SEQ ID NO:1 的编码 GDNF 的基因或其片段杂交而分离的任何截短的 GDNF 蛋白质。

通过阅读本说明书对本领域的技术人员而言显而易见的是，基本上同源的蛋白质包括对 X-[Cys⁴¹-Cys¹³³]-Y 所代表的截短的 GDNF 蛋白质的氨基酸残基进行的一个或多个的缺失，添加或取代。该变异体的产生在下文进一步详细描述。还可预期由于本发明清楚地阐述了“截短”的 GDNF 蛋白质，可预期氨基末端添加变异体包括添加甲硫氨酸残基或非 - GDNF 氨基酸残基或序列，但不包括添加可导致成熟 GDNF 蛋白质重构的氨基酸残基。以天然存在的等位基因突变体或变异体为基础的截短的 GDNF 蛋白质也在本发明的范围内。变异的截短 GDNF 蛋白质在下面进一步详细地描述。

Lin 等（美国专利申请号 08/182,183，出处同上）描述了经蛋白水解加工 Lys-Arg 残基在羧基末端截短成熟 GDNF，其中该 Lys-Arg 分别从成熟 GDNF 羧基末端为第 6 位和第 5 位残基（即、根据图 1（也同 SEQ ID NO:1 或 SEQ ID NO:2）的氨基酸残基编号为 Lys¹²⁹-Arg¹³⁰）。这种截短从成熟 GDNF 蛋白质上去掉了两个半胱氨酸残基。这很可能导致蛋白质的折叠不合适，从而导致形成失活的蛋白质。相反，本发明的 X-[Cys⁴¹-Cys¹³³]-Y 截短的 GDNF 蛋白质产物含 Cys¹³¹ 和 Cys¹³³ 残基且经多巴胺吸收试验测定为活性蛋白质。

在本发明的一个实施方案中，优选的截短的 GDNF 蛋白质产物缺乏一个或多个脱酰胺化位点。这种缺乏脱酰胺化位点可能导致纯化蛋白质的生化稳定性增强及减少可能的降解产物，从而产生贮存更稳定的蛋白质。举例性的截短的 GDNF 蛋白质产物是 [Ser²⁶-Ile¹³⁴] 截短的 GDNF 蛋白质，它缺乏可能导致成熟蛋白质脱酰胺化的位点。另外，[Arg¹⁶-Ile¹³⁴] 截短的 GDNF 蛋白质缺乏至少存在于成熟蛋白质中的第一脱酰胺化位点。

目前优选的截短的 GDNF 蛋白质产物是 [Arg³²-Ile¹³⁴] 截短的 GDNF

蛋白质。该截短的 GDNF 蛋白质缺乏发生成熟蛋白质的蛋白水解剪切的位点或靠近它的位点。因此,预期该截短的 GDNF 蛋白质对也可能发生于体内的加工事件有抗性。目前另一优选的截短的 GDNF 蛋白质产物是 [Lys³⁷-Ile¹³⁴]截短的 GDNF 蛋白质。这种截短可能进一步降低了截短蛋白质的 pI, 正如分别从 N 端和 C 端去掉达到且包括 Gly⁴⁰ 和 Ile¹³⁴ 的残基的其它截短一样。目前最优选的截短 GDNF 蛋白质产物保留在成熟 GDNF 蛋白质中发现的全部半胱氨酸残基, 但缺乏在表达和生产期间或体内用药后截短的 GDNF 蛋白质迅速蛋白水解加工的任何可识别的位点。这些优选的蛋白质包括 [Arg³²-Ile¹³⁴], [Gly³³-Ile¹³⁴], [Gln³⁴-Ile¹³⁴], [Arg³⁵-Ile¹³⁴], [Gly³⁶-Ile¹³⁴], [Lys³⁷-Ile¹³⁴], [Asn³⁸-Ile¹³⁴]和 [Arg³⁹-Ile¹³⁴]截短的 GDNF 蛋白产物。

与以前所述的 Lin 等(美国专利申请号 07/855,413 出处同上)的结果相似,已证实本发明的截短的 GDNF 蛋白质具有增强灰质多巴胺能神经元胚胎前体对多巴胺的吸收的能力。截短的 GDNF 蛋白质的生物测定在下面实施例 4 中进一步描述。

典型地分离新截短的 GDNF 蛋白质并纯化形成基本上不含其它(非 GDNF)蛋白物质的截短的 GDNF 蛋白质。优选的是,截短的 GDNF 蛋白质产物为大约 80% 不含由于用于截短的 GDNF 蛋白质产物的制备的生产技术而存在的其它蛋白质。更优选的是,截短的 GDNF 蛋白质产物为大约 90% 不含其它蛋白质,特别优选的,大约 95% 不含其它蛋白质,最优选大约 >98% 不含其它蛋白质。另外,本发明具有提供用于生产同源截短的 GDNF 蛋白质的多核苷酸序列的独特优势。例如,使用编码 [Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质的多核苷酸序列允许在大肠杆菌和其它合适的表达系统中重组生产截短的 GDNF 蛋白质。换句话说,新的多核苷酸允许生产对蛋白水解加工不敏感的或对该加工及上面所述的其它生化加工效应的敏感性降低的截短的 GDNF 蛋白质。因此,新的多核苷酸使其易于制备和/或分离单个种类的截短 GDNF 蛋白质,因此,截短的 GDNF 蛋白质和/或其产物不含异型和同型二聚体的上述混合物或其含量下降。然而可预期最终截短的 GDNF 蛋白质产物可在用药前与其它因子,化学组合物和/或合适的药用组合物物质结合,如下面所进行的详细描述。

在本发明的一个方面,截短的 GDNF 蛋白质经重组技术生产具有优

势，因为它们能以更大的纯度获得比较起来含量更高的蛋白质。重组截短的 GDNF 蛋白质形式包括蛋白质的糖基化和非糖基化形式及在细菌，哺乳动物或昆虫细胞系统中表达的蛋白质。另外，截短的 GDNF 蛋白质可经化学合成。目前优选的生产方法在下面进行更详细的描述。

截短的 GDNF 变异体和衍生物

A. 截短的 GDNF 变异体

本发明的另一方面包括截短的 GDNF 蛋白质的变异体。本文所用的术语“截短的 GDNF 蛋白质产物”包括变异的蛋白质，其中在天然存在的 GDNF 氨基酸序列内氨基酸残基被缺失(“缺失变异体”)，插入(“添加变异体”)或取代(“取代变异体”)。经过将合适的核苷酸变化导入编码该蛋白质的 DNA 中或经体外化学合成所需的蛋白质来制备该变异体。本领域的技术人员可预期可制备许多缺失，插入和取代的组合，只要最终的蛋白质具有 GDNF 生物学活性。

一个或多个选定的氨基酸残基的取代，插入或缺失的诱变技术对本领域的技术人员来说是熟知的(例如，美国专利号 4518584，本文引用其说明书以供参考)。在氨基酸序列变异体的构建中有 2 个原则上的变化：诱变位点的位置和诱变性质。在截短的 GDNF 变异体的设计中，诱变位点的位置和诱变性质取决于待修饰的生化特征。可单独或系列修饰突变位点，例如，经过(1)首先用保守的氨基酸选择，然后根据所得的结果用更激进的选择进行取代，(2)缺失靶氨基酸残基，或(3)相邻于选定位点插入氨基酸残基。

氨基酸序列缺乏范围一般从大约 1 到 30 个氨基酸残基，更常见的是从大约 1 到 10 个残基，典型地从大约 1 到 5 个残基。例如，位于 Cys⁴¹N 端的氨基酸残基“X”部分的缺失范围从大约 1 到 30 个残基，而 [Cys⁴¹-Cys¹³³]半胱氨酸残基之间的缺失根据其位置典型地从大约 1 到 5 个残基，以便不破坏蛋白质折叠。可在与转化生长因子- β (TGF- β) 家族成员同源性较低的区域对截短的 GDNF 蛋白质进行缺失。在与其它 TGF- β 家族序列基本上同源的区域截短 GDNF 蛋白质的缺失很可能会更显著地改变生物学活性。选择总缺失数和/或连续缺失以保护受影响的区域的截短 GDNF 蛋白质的三级结构，例如，半胱氨酸交联。

氨基酸序列添加包括长度范围从 1 个残基到 100 或更多个残基的氨基和/或羧基端融合，以及单个或多个氨基酸残基的内部序列内插入。

内部添加范围一般从大约 1 到 10 个氨基酸残基，更典型地是从大约 1 到 5 个氨基酸残基，通常是从大约 1 到 3 个氨基酸残基。如上所述，预期本发明的氨基末端添加变异体包括添加甲硫氨酸（例如，在细菌重组细胞培养物中作为直接表达 GDNF 的假像）或非 GDNF 氨基酸残基或序列。氨基末端添加变异体不包括添加可导致成熟 GDNF 蛋白质重构的氨基酸残基。末端插入的另一个例子包括异源性 N - 端信号序列与 N 端的融合以利于从重组宿主细胞分泌蛋白质。该信号序列一般可从目的宿主细胞种类获得且因此与之同源。插入或添加还可包括来自其它神经营养因子的序列的氨基酸序列。

另一组变异体是氨基酸取代变异体。这些变异体在截短的 GDNF 蛋白质至少去掉了一个氨基酸残基且在其位置插入了不同的残基。例如，见图 5，其中天然存在的 Asn²² 被变成 Ser 以利于进一步去掉 Met 残基。使用 X-[Cys⁴¹-Cys¹³³]-Y 氨基酸序列表达式和该截短的 GDNF 蛋白产物的该定义，这种截短的 GDNF 蛋白质可称为取代变异体 Met-[Asn²²ΔSer²²-Ile¹³⁴]截短的 GDNF 蛋白质或添加变异体 Met-Ser-[Pro²³-Ile¹³⁴]截短的 GDNF 蛋白质。取代变异体包括等位基因变异体，其特征在于在种群中存在可导致氨基酸改变或不变的天然存在的核苷酸序列改变。

截短的 GDNF 蛋白质的序列的特异性突变可包括糖基化位点（例如，丝氨酸，苏氨酸或天冬酰胺）的修饰。缺乏糖基化或仅部分糖基化可由在任意天冬酰胺连接的糖基化识别位点或在经添加 O - 连接的糖类修饰的蛋白质的任意位点的氨基酸取代或缺失引起。天冬酰胺连接的糖基化识别位点包括被合适的细胞糖基化酶特异性识别的三肽序列。这些三肽序列是 Asn-Xaa-Thr 或 Asn-Xaa-Ser，其中 Xaa 可以是非 Pro 的任意氨基酸。在糖基化识别位点的一个或第一及第三氨基酸位置的各种氨基酸取代或缺失（和/或在第二位置的氨基酸缺失）导致在修饰的三肽序列上非糖基化。因此，适当改变的核苷酸序列的表达产生在该位点未糖基化的变异体。另外，该序列可修饰成在截短的 GDNF 蛋白质上添加糖基化位点。

鉴定用于诱变的截短 GDNF 氨基酸残基或区域的一个方法称为“丙氨酸扫描诱变”，如 Cuninghame 和 Wells 所述（科学，244:1081-1085, 1989）。在该方法中，鉴定了氨基酸残基或靶残基的基团（例如，带

电残基，如 Arg, Asp, His, Lys 和 Glu) 并用中性或带负电的氨基酸 (最优选丙氨酸或多聚丙氨酸) 取代以影响氨基酸与细胞内或细胞外的周围水环境的相互作用。经过在该取代位点导入其它或任选的残基精选对取代表现出功能性敏感的区域。从而预定用于导入氨基酸序列修饰的位点并选择在给定位点的最佳突变效果，进行丙氨酸扫描或随机诱变并筛选所需活性和活性程度的最佳组合的变异体。

用于取代诱变的最感兴趣的位点包括在各种类型的 GDNF 蛋白质中发现的氨基酸在侧链大小，电荷和/或疏水性方面基本上不同的位点。感兴趣的其它位点包括从各种类型获得的 GDNF 样蛋白质的特定残基相同的那些位点。这些位置一般来说对蛋白质的生物学活性是重要的。开始，以相当保守的方式经取代修饰这些位点。这些保守的取代在表 1 中优选的取代下显示。如果这些取代导致生物学活性改变，那么导入更大的改变 (举例性取代) 和/或进行增加/缺失，筛选所得的产品。

表 1
氨基酸取代

原始残基	优选的取代	举例性取代
Ala (A)	Val	Val; Leu; Ile
Arg (R)	Lys	Lys; Gln; Asn
Asn (N)	Gln	Gln; His; Lys; Arg
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro	Pro
His (H)	Arg	Asn; Gln; Lys; Arg
Ile (I)	Leu	Leu; Val; Met; Ala; Phe; 正亮氨酸
Leu (L)	Ile	正亮氨酸 · Ile; Val; Met; Ala; Phe
Lys (K)	Arg	Arg; Gln; Asn
Met (M)	Leu	Leu; Phe; Ile
Phe (F)	Leu	Leu; Val; Ile; Ala
Pro (P)	Gly	Gly
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr	Tyr
Tyr (Y)	Phe	Trp; Phe; Thr; Ser
Val (V)	Leu	Ile; Leu; Met; Phe; Ala; 正亮氨酸

预期对氨基酸序列的保守修饰(和对编码核酸序列的相应修饰)产生功能和化学特征相似于下面实施例中所述的截短 GDNF 蛋白质的截短 GDNF 蛋白质。相反,经过选择对维持(a)取代区域的多肽骨架结构,例如,片层或螺旋构象,(b)靶位点的蛋白质电荷或疏水性,或(c)侧链大小其影响明显不同的取代可实现截短 GDNF 蛋白质的功能和/或化学特征的明显改变。天然存在的残基根据其共同的侧链特征分组:

- 1)疏水的: 正亮氨酸, Met, Ala, Val, Leu, Ile;
- 2)中性亲水的: Cys, Ser, Thr;
- 3)酸性: Asp, Glu;
- 4)碱性: Asn, Gln, His, Lys, Arg;
- 5)影响链取向的残基: Gly, Pro; 和
- 6)芳香族: Trp, Tyr, Phe

非保守性取代包括一种类型中的成员与另一类型的交换。这些取代的残基可导入与其它 TGF- β 蛋白质同源的截短 GDNF 蛋白质区域或该蛋白的非同源区域。

B. 截短的 GDNF 衍生物

按本文说明书本领域的技术人员可制备截短 GDNF 或截短 GDNF 变异体的化学修饰的衍生物。最适于衍变截短 GDNF 蛋白质的化学基团包括水溶性多聚体。需要水溶性多聚体是因为附着于其上的蛋白质在水环境,如生理环境中不沉淀。优选的是,该多聚体对于制备治疗性产品或组合物是药学上可接受的。本领域的技术人员根据多聚体/蛋白质结合物是否能在治疗上使用,及如果能,所需的剂量,循环时间,对蛋白水解的抗性及其它因素的这种考虑可选择所需的多聚体。经过以所需形式(即,经渗透压泵,或更优选的是,经注射或输注或进一步配制成口服、肺或其它传递途径)施用该衍生物并测定其效果可确定衍变的效果。

合适的水溶性多聚体包括,但不限于,聚乙二醇(PEG),乙二醇/丙二醇共聚体,单甲氧基聚乙二醇,羧甲基纤维素,葡聚糖,聚乙烯醇,聚乙烯吡咯烷酮,聚-1,3-二氧戊烷,聚-1,3,6-三噁烷,乙烯/马来酸酐共聚体,多聚氨基酸(同型多聚体或随机共聚体),聚(正乙烯吡咯烷酮)聚乙二醇,丙二醇同型多聚体,脯氨酸基氧化丙烯/环氧乙烷共聚体,聚氧乙烯化多元醇(例如,甘油),聚乙二醇丙醛

和其混合物。本文使用的聚乙二醇指包含用于衍变其它蛋白质的 PEG 的任何形式，如单 - (C1 - C10) 烷氧基或芳氧基聚乙二醇。聚乙二醇丙醛由于其在水中的稳定性而在生产中具有优势。该多聚体可以是任意分子量，且可以是分枝的或不分枝的。

本发明特别涉及包括截短的 GDNF 蛋白质连接到至少一个 PEG 分子上的截短的 GDNF 蛋白质产物。另一方面，本发明涉及经酰基或烷基键附着于至少一个 PEG 分子上的截短的 GDNF 蛋白质。

以本领域已知的任一 PEG 化反应可进行 PEG 化。例如，参见：Focus on Growth Factors, 3(2): 4-10 (1992); EP 0154316; EP 0401384; 和 Malik 等，实验血液学; 20:1028-1035 (1992) (报道了使用 tresyl chloride 将 GM-CSF PEG 化)。优选的是，用活性水溶性多聚体经酰基化反应或烷基化反应进行 PEG 化。这些用于衍变的优选的方法在下面进行更详细的讨论。对于酰化反应，优选选择的多聚体具有单个活性酯基。对于还原性烷基化反应，优选选择的多聚体具有单个活性醛基。另外，选择的多聚体可修饰成具有单个反应基团，如用于酰基化的活性酯或用于烷基化的醛，以便控制聚合程度。一般来说，水溶性多聚体不能选自天然存在的糖基残基，因为这些残基通常可经哺乳动物重组表达系统更方便地制备。

酰基化

在本发明中，经酰基化的 PEG 化通常涉及将聚乙二醇的活性酯衍生物与截短的 GDNF 蛋白质反应。任何已知的或随后发现的活性 PEG 分子可用于进行 PEG 化过程。优选的活化的 PEG 酯是 PEG 酯化到 N - 羟基琥珀酰亚胺 (“ NHS ”) 上。本文使用的“酰基化”预期包括但不限于下列类型的在截短的 GDNF 蛋白质和诸如 PEG 的水溶性多聚体间形成的键：酰胺，氨基甲酸，尿烷等。见生物结合物化学， 5: 133-140 (1994)。反应条件可选自 PEG 化领域已知的或随后建立的任何条件，但应避免或限制接触能使待修饰的截短的 GDNF 蛋白质失活的诸如温度、溶剂和 pH 水平的反应条件。

经酰基化反应来 PEG 化一般产生多个 PEG 化的截短 GDNF 蛋白质，其中经酰基连接基团 PEG 化赖氨酸 ϵ - 氨基。优选的是，连接链可以是酰胺。还优选的是，所得的产物基本上仅 (例如 $\geq 95\%$) 单个，二个或三个 PEG 化。然而，根据所使用的具体反应条件可形成在总量上

具有更高的 PEG 化程度的一些结合物。如果需要，经过标准纯化技术，其中包括透析、盐析、超滤、离子交换色谱，凝胶过滤色谱和电泳，可从混合物中制备更纯化的 PEG 化的结合物。

烷基化

在本发明中，以烷基化进行的 PEG 化一般包括在还原剂存在的条件下将 PEG 的末端醛基衍生物与截短的 GDNF 蛋白质反应。经烷基化进行的 PEG 化也可产生多个 PEG 化的截短的 GDNF 蛋白质。另外，可控制反应条件以帮助基本上仅在蛋白质 N 端的 α 氨基上 PEG 化（即，单个 PEG 化的种类）。在单个 PEG 化或多个 PEG 化的情况下，PEG 基团优选经 $-\text{CH}_2-\text{NH}-$ 基团附着于蛋白质上。特别对于 $-\text{CH}_2-$ 基团，这类连接本文称为“烷基”链。

经还原性烷基化可实现选择性 N 端化学修饰，这利用了对于特定蛋白质的衍变可用的不同类型的伯胺基的不同反应性（赖氨酸与 N 端）。在合适的反应条件下，用含羧基的多聚体在 N 端进行对蛋白质的基本上选择性的衍化。例如，在允许利用赖氨酸残基 ϵ -氨基和蛋白质 N 端残基 α -氨基之间的 pKa 差异的 pH 下进行反应可选择性地在 N 端 PEG 化该蛋白质。经过这种选择性衍变，可控制水溶性多聚体与蛋白质的附着：与多聚体的连接主要发生在蛋白质的 N 端，且在其它反应基团，如赖氨酸侧链氨基基团上没有明显的修饰出现。使用还原性烷基化时，水溶性多聚体优选具有用于偶联蛋白质的单个反应性醛。可使用含单个反应性醛的聚乙二醇丙醛。

本发明包括 PEG 化截短的 GDNF 蛋白质，其中 PEG 基团经酰基或烷基基团附着。正如上面所讨论的，这种截短的 GDNF 蛋白质产物可被单个 PEG 化或多个 PEG 化（例如，含 2 - 6 个，优选 2 - 5 个 PEG 基团）。这种 PEG 基团一般在氨基酸的 α -或 ϵ -氨基基团上附着于蛋白质上，但还预期 PEG 基团可附着于蛋白质的任何反应基团上。在合适的反应条件下，该反应基团具有足够的反应性以附着于 PEG 基团上。因此，聚乙二醇可经过诸如游离氨基或羧基的反应基团共价结合于蛋白质上。反应性基团是那些激活的 PEG 分子可结合上的基团。具有游离氨基的氨基酸残基可包括赖氨酸残基和 N-端氨基酸残基。具有游离羧基的那些残基包括天冬氨酸残基，谷氨酸残基，和 C-端氨基酸残基。硫氢基基团也可用作附着 PEG 分子的反应基团。为达到治疗目的，

典型地优选在氨基上附着，如在 N 端或赖氨酸基团上附着。如果需要受体结合，应避免附着在对受体结合重要的残基上。

在一个方面，本发明提供了单个多聚体/截短的 GDNF 蛋白质结合物的基本上均一的制品，其中多聚体分子基本上仅（即 $\geq 95\%$ ）在单个位置附着。更具体地说，如果使用 PEG，本发明还提供了缺乏可能的抗原性连接基团，且 PEG 分子直接偶联到截短的 GDNF 蛋白质上的 PEG 化的截短的 GDNF 蛋白质。

另外，使用糖基化的，非糖基化的或去糖基化的截短 GDNF 蛋白质可制备衍生物。典型地是，使用非糖基化的截短的 GDNF 蛋白质。例如，原核表达的[Arg³²-Ile¹³⁴]截短 GDNF 蛋白质可化学衍变成包括经酰基或烷基基团附着的单个或多个，例如 2 - 4 个 PEG 部分：

一般来说，在用于将生物学活性物质与活化的多聚体分子反应的任何合适的条件下可进行化学衍变。制备 PEG 化的截短的 GDNF 蛋白质的方法一般包括下列步骤：（a）在截短的 GDNF 蛋白质附着到一个或多个 PEG 基团的条件下将截短的 GDNF 蛋白质与聚乙二醇（如 PEG 的反应性酯或醛衍生物）反应，（b）获得反应产物。一般来说，根据已知的参数和所需的结果逐例测定酰基化反应的最佳反应条件。例如，PEG:蛋白质的比例越大，多个 PEG 化的产物百分数越大。以诸如衍化的所需程度（例如，单个，2 个，3 个等），选择的多聚体的分子量，多聚体是分枝的还是不分枝的及所用的反应条件的因素可测定最佳比例（反应的效率在于无剩余的未反应的蛋白质或多聚体）。

产生基本上均一的单个多聚体/截短的 GDNF 蛋白质结合物的群体的还原性烷基化一般包括下列步骤：（a）在还原性烷基化条件下，在适于使截短的 GDNF 蛋白质的氨基末端 α -氨基选择性修饰的 pH 下将截短的 GDNF 蛋白质与反应性 PEG 分子反应，（b）获得反应产物。

为得到单个多聚体/截短的 GDNF 蛋白质连接物的基本上均一的群体，还原性烷基化反应条件是允许水溶性多聚体部分选择性附着于截短的 GDNF 蛋白质 N 端的那些条件。该反应条件一般提供了赖氨酸氨基和 N 端 α -氨基间的 pKa 差异（pKa 是 50 % 氨基质子化而 50 % 非质子化的 pH）。pH 还影响多聚体与所用的蛋白质之间的比率。一般来说，如果 pH 较低，需要超出蛋白质的多聚体量越大（即，N 端 α -氨基的反应性越小，达到最佳条件需要的多聚体越多）。如果 pH 更高，

多聚体：蛋白质比率不必太大（即，可获得反应基团越多，所需的多聚体分子越小）。为达到本发明的目的，pH一般落在3 - 9的范围内，优选3 - 6。

另一考虑是多聚体的分子量。一般来说，多聚体的分子量越高，可附着于蛋白质上的多聚体分子数越少。同样，在优选这些参数时应考虑多聚体的分枝。一般来说，分子量越高（或分枝越多），多聚体：蛋白质的比率越高。一般情况下，为进行本文预期的PEG化反应，优选的平均分子量是大约2kDa到大约100kDa（术语“大约”指在聚乙二醇制品中，有些分子比所述的分子量更大，有些更小）。优选的平均分子量是大约5kDa到大约50kDa，特别优选大约12kDa到大约25kDa。水溶性多聚体与截短的GDNF蛋白质的比率一般范围为从1:1到100:1，优选（对于多个PEG化）1:1到20:1，和（对于单个PEG化）1:1到5:1。

使用上面所述条件，还原性烷基化可使多聚体选择性附着到在氨基末端具有 α -氨基的任意截短的GDNF蛋白质，且产生单个多聚体/截短的GDNF蛋白质结合物的基本上均一的制品。本文使用的术语“单个多聚体/截短的GDNF蛋白质结合物”指含附着于截短的GDNF蛋白质上的单个多聚体分子的衍生物。单个多聚体/截短的GDNF蛋白质结合物优选具有位于N端，而不在赖氨酸氨基上的多聚体分子。该制品优选含有大于90%的单个多聚体/截短的GDNF蛋白质结合物，更优选含有大于95%的单个多聚体/截短的GDNF蛋白质结合物，余下的值得注意的蛋白质是未反应的（即，缺乏多聚体部分的蛋白质）。

为进行还原性烷基化，还原剂应在水溶液中稳定，且优选仅能还原在还原性烷基化的起始过程中形成的西佛碱。举例性的还原剂可选自氢硼化钠，氰基氢硼化钠，二甲胺甲硼烷，三甲胺甲硼烷和吡啶甲硼烷。特别优选的还原剂是氰基氢硼化钠。其它反应参数，如溶剂，反应时间，温度等，和产品的纯化方式，可根据按常规可获得的涉及用水溶性多聚体衍变蛋白质的信息逐一测定。

经酰基化和/或烷基化方法可选择制备多聚体/蛋白质结合物的混合物，本文提供的优势在于可选择单个多聚体/蛋白质连接物的比例以包括进混合物中。因此，如果需要，可制备各种数目的多聚体分子附着于其上的蛋白质的混合物（即，二个，三个，四个等）并与使用本方法制

备的单个多聚体/蛋白质结合物组合及具有预定比例的单个多聚体/蛋白质连接物的混合物。

编码截短的 GDNF 蛋白质的多核苷酸

本发明进一步提供了编码截短的 GDNF 蛋白质的新的多核苷酸。当用作杂交探针或扩增引物时，该核酸序列基本上不含所有其它核酸序列。为在重组蛋白质表达中使用，该核酸序列一般基本上不含编码其它蛋白质的核酸序列，除非需要一种融合蛋白质。基于本说明书并使用通用密码子表，本领域的普通技术人员易于确定编码截短的 GDNF 蛋白质的氨基酸序列的所有核酸序列。目前优选的核酸序列包括编码[Arg¹⁶-Ile¹³⁴], [Ser²⁶-Ile¹³⁴], [Arg³²-Ile¹³⁴]及[Lys³⁷-Ile¹³⁴]截短的 GDNF 蛋白质的那些多核苷酸。各种多核苷酸的例子在图 5, 6 和 7 及编码截短的 GDNF 蛋白质的图 1, 3 和 4 的那些部分中描述。本领域的技术人员还可预期到编码截短的 GDNF 蛋白质的新多核苷酸包括编码变异的截短的 GDNF 蛋白质的那些核酸序列，不管是人造的还是天然存在的核酸序列。

根据下面所提供的描述可实施重组表达技术以生产这些多核苷酸并表达各种截短的 GDNF 蛋白质。例如，经过将编码截短的 GDNF 蛋白质的核酸序列插入合适的载体中，本领域的技术人员易于生产大量的所需核苷酸序列。然后该序列可用于生产检测探针或扩增引物。另外，可将编码截短的 GDNF 蛋白质的多核苷酸插入表达载体中。经过将表达载体导入合适的宿主中，可大量生产所需的截短的 GDNF 蛋白质。

如本文进一步所述，有许多宿主/载体系统可用于扩增核酸序列和/或生产截短的 GDNF 蛋白质。这些包括，但不限于，质粒，病毒和插入载体及原核和真核宿主。本领域的技术人员可改进能扩增或表达异源 DNA 以生产或表达本发明的序列的宿主/载体系统。

借助于该重组技术，容易以商用量生产本发明的截短的 GDNF 蛋白质。而且，根据本说明书，本领域的技术人员可预料到新核酸序列包括编码在附图中特别提到的截短的 GDNF 蛋白质，该截短的 GDNF 蛋白质的变异体的简并核酸序列，及那些优选在严紧杂交条件下与这些核酸序列的互补序列杂交的核酸序列（见，Maniatis 等，分子克隆（实验手册）；冷泉港实验室，387 到 389 页，1982）。举例性的严紧杂交条件是在 4×SSC 中 62 - 67 °C 杂交，接着在 0.1×SSC 中 62 - 67 °C 洗涤大约 1 小时。另外，举例性的严紧杂交条件是在 45 - 55 % 甲酰胺，

4×SSC 中 40 - 45 °C 杂交, 在松驰型杂交条件下与编码截短的 GDNF 蛋白质的互补序列杂交的且编码本发明的截短的 GDNF 蛋白质的 DNA 序列也在本文中包括。这种松驰型严紧杂交条件的例子是 4×SSC, 45 - 55 °C 或用 30 - 40 % 甲酰胺在 40 - 45 °C 杂交。

本发明还提供了包括载体 DNA 与编码截短 GDNF 蛋白质的 DNA 序列的重组 DNA 构建体。在该 DNA 构建体中, 编码截短的 GDNF 蛋白质(具有或不含信号肽)的核酸序列与能指导截短 GDNF 蛋白质在选定的宿主中复制和/或表达的合适的表达控制或调节序列以可操作的方式相连。

截短的 GDNF 蛋白质的重组表达.

编码截短的 GDNF 的多核苷酸的制备:

编码截短的 GDNF 或成熟 GDNF 起始材料的核酸序列可用包括但不限于化学合成, cDNA 或基因组文库筛选, 表达文库筛选和/或 cDNA 的 PCR 扩增的各种方法易于获得。这些方法和用于分离这类核酸序列的其它方法, 例如由 Sambrook 等(分子克隆; 实验手册, 冷泉港实验室出版社, 冷泉港, NY, 1989), Ausubel 等编辑(分子生物学常规方法, 常规方法出版社, (1994)和 Berger 和 Kimmel(酶学方法: 分子克隆技术指南, Vol. 152, 学院出版公司, San Diego, CA, 1987)进行了阐述。优选的编码 GDNF 的核酸序列是哺乳动物序列。

编码截短的 GDNF 蛋白质的核酸序列的化学合成也可使用本领域已知的方法实现, 如由 Engels 等所述的方法(Angew, Chem, Intl. Ed. 28:716-734, 1989)。这些方法特别包括核酸序列分析的磷酸三酯, 亚磷酰胺和 H - 磷酸酯方法。编码截短的 GDNF 蛋白质的核酸序列具有几百个碱基对(bp)或核苷酸的长度。大型的核酸序列, 例如长度大于大约 100 个核苷酸的那些核酸序列可作为一些片段来合成。然后将这些片段连接起来形成编码截短的 GDNF 蛋白质的核酸序列。优选的方法是使用标准亚磷酰胺化学进行多聚体支持的合成。

另外, 经过筛选合适的 cDNA 文库(即从据信表达该蛋白质的一个或多个组织来源制备的文库)或基因组文库(从总基因组 DNA 制备的文库)可获得合适的核酸序列。这些 cDNA 文库的来源典型地是据信以合理量表达 GDNF 的任何种类的组织。基因组文库的来源可以是任意组织或据信含有编码 GDNF 或 GDNF 同源物的基因的任何哺乳动物或其

它种类的组织。可使用选择性地与存在于文库中的 GDNF 或 GDNF 同源性 cDNA 或基因杂交的一个或多个核酸探针（与待克隆的 GDNF 或 GDNF 同源性 cDNA 或基因具有可接受水平的同源性的寡核苷酸 cDNA 或基因组 DNA 片段）筛选存在 GDNF cDNA/基因的文库。典型地用于该文库筛选的探针通常编码与制备文库的种类相同或相似的种类的 GDNF DNA 序列的小区域。另外，该探针可以是本文所讨论的简并序列。

典型地是，在防止非特异性结合但允许与探针或引物具有有意义的同源水平的那些克隆结合的严紧条件下经过退火寡核苷酸探针或 cDNA 与文库中的克隆实现文库的筛选。典型的杂交和洗涤严紧条件部分取决于 cDNA 或寡核苷酸探针的大小（即，核苷酸数目的长度）以及该探针是否具简并性。获得克隆的可能性也考虑到杂交溶液的设计（即，待筛选的是 cDNA 还是基因组文库，如果是 cDNA 文库，以高水平存在感兴趣的 cDNA 的可能性）。

如果 DNA 片段（如 cDNA）用作探针，典型的杂交条件包括在 Ausubel 等，编辑，出处同上中阐述的那些杂交条件。杂交后，含该文库的印迹根据诸如探针大小，探针与克隆的预期同源性，待筛选的文库类型，待筛选的克隆数等一些因素以合适的严紧性洗涤。严紧洗涤溶液的例子（通常为低离子强度和在相当高温下使用）如下。一种这类严紧洗涤为 0.015M NaCl, 0.005M 柠檬酸 Na 和 0.1% SDS, 55-65 °C。另一种该严紧缓冲液是 1mM Na₂EDTA, 40mM NaHOP₄, pH 7.2 和 1 % SDS, 大约 40 - 50 °C。一种其它严紧洗涤为 0.2×SSC 和 0.1%SDS, 大约 50 - 65 °C。

对于将寡核苷酸探针用于选择 cDNA 或基因组文库的严紧洗涤条件还有一些举例性的方案。例如，第一种方案使用具有 0.05%焦磷酸钠的 6×SSC，根据探针长度温度为大约 35 到 62 °C 之间。例如，14 个碱基探针在 35 - 40 °C 洗涤，17 个碱基探针在 45 - 50 °C，20 个碱基探针在 52 - 57 °C，23 个碱基探针在 57 - 63 °C。当出现较高的背景非特异性结合时，温度可增加 2 - 3 °C。第二方案使用氯化四甲铵（TMAC）用于洗涤。一个这种严紧洗涤溶液是 3M TMAC, 50mM Tris-HCl, pH8.0 和 0.2%SDS。

获得编码 GDNF 蛋白质的核酸序列的另一合适的方法是聚合酶链

式反应 (PCR)。在该方法中, poly(A)+RNA 或总 RNA 从表达 GDNF 的组织中提取。然后使用逆转录酶从 RNA 制备 cDNA。然后向 cDNA 与诸如 Taq 聚合酶的聚合酶中加入典型地与 GDNF cDNA 的两个分离的区域互补的二种引物 (寡核苷酸), 该聚合酶扩增两种引物间的 cDNA 区域。

如果选择制备编码所需截短的 GDNF 蛋白质的核酸序列的方法需要使用寡核苷酸引物或探针 (例如, PCR, cDNA 或基因组文库筛选), 选择作为探针或引物的寡核苷酸序列应足够长和足够清楚以减小在文库筛选或 PCR 扩增期间出现的非特异性结合的量。探针或引物的实际序列通常以另一生物体的相同或相似基因的保守的或高度同源的序列或区域为基础。另外, 探针或引物可以是完全或部分简并的, 即, 含探针/引物的混合物, 均编码相同氨基酸序列, 但使用不同的密码子编码。制备简并探针的一种选择是在不同种类中有变化的一些或全部密码子中放入肌苷。用上面所述的 DNA 的化学合成方法可制备寡核苷酸探针或引物。

以编码 GDNF 的这些核酸序列及其突变体或变异序列为基础的截短的 GDNF 蛋白质也包含在本发明范围内。如上所述, 突变体或变异体序列是与野生型序列相比含一个或多个核苷酸取代, 缺失和/或插入的序列及与野生型氨基酸序列相引起表达氨基酸序列变异体的序列。在某些情况下, 由于存在天然等位基因变异体, 可存在天然产生的 GDNF 氨基酸突变体或变异体。以这些天然产生的突变体或变异体为基础的截短的 GDNF 蛋白质也在本发明的范围内。合成突变体序列的制备也是本领域熟知的, 且在例如 Wells 等 (基因, 34: 315, 1985), 和 Sambrook 等, 出处同上中进行了描述。

载体

将编码截短 GDNF 蛋白质的 cDNA 或基因组 DNA 插入载体中以便进一步克隆 (DNA 的扩增) 或表达合适的载体是商业上可获得的, 或该载体可特定构建。合适载体的选择或构建取决于: 1). 它是用于 DNA 扩增还是用于 DNA 表达, 2). 将要插入载体中的 DNA 的大小, 3). 将要使用该载体转化的宿主细胞 (例如, 哺乳动物, 昆虫, 酵母, 真菌, 植物或细菌细胞)。各载体根据其功能 (DNA 扩增或 DNA 表达) 及其与目的宿主细胞的相容性可含有各种成分。载体成分一般包括, 但不限于

下列的一种或多种：信号序列，复制起点，一个或多个选择或标记基因，增强子元件，启动子，转录终止序列，等。这些成分可从天然来源获得或用已知方法合成。本发明的载体包括选择性连接到一个或多个能指导，控制或否则影响选定的宿主细胞表达截短的 GDNF 蛋白质的下列表达控制或调节序列上的编码感兴趣的截短的 GDNF 蛋白质的核酸序列。

信号序列

信号序列可以是载体的一种成分，或它可以是插入载体中的部分 GDNF DNA。天然 GDNF，DNA 编码在蛋白质的翻译后加工期间被裂解以形成成熟 GDNF 蛋白质的位于蛋白质的氨基末端的信号序列。在本发明的范围内包括具有天然信号序列和其它蛋白原前体序列的截短的 GDNF 多核苷酸以及天然信号序列被删除并用异源性信号序列取代的截短的 GDNF 多核苷酸。选定的异源性信号序列应是能被宿主细胞识别并加工（即，经过信号肽酶裂解）的序列。对于不能识别和加工天然 GDNF 信号序列的原核宿主细胞，用选自，例如碱性磷酸酶，青霉素酶或热不稳定性肠毒素 II 前导区的原核信号序列取代信号序列。对于酵母分泌，可用酵母转化酶， α 因子或酸性磷酸酶前导区取代天然 GDNF 信号序列。在哺乳动物细胞表达中，天然信号序列是合适的，尽管其它哺乳动物信号序列可能是合适的。

复制起点

表达和克隆载体一般包括能使载体在一种或多种选定的宿主细胞中复制的核酸序列。在克隆载体中，该序列典型地是能使载体独立于宿主染色体 DNA 复制的序列且包括复制起点或自主复制序列。该序列对于各种细菌，酵母和病毒来说是熟知的。质粒 pBR322 的复制起点适合于大多数革兰氏阴性细菌，且各种起点（例如，SV40，多形瘤，腺病毒，VSV 或 BPV）对于哺乳动物细胞中的克隆载体有用。一般来说，复制成分的起点对于哺乳动物表达载体是不需要的（例如，经常使用 SV40 起点仅仅因为它含有早期启动子）。

选择基因

表达和克隆载体典型地含有一个选择基因。该基因编码生长于选择性培养基中时对转化宿主细胞的存活或生长必需的“标记”蛋白质。未被载体转化的宿主细胞不含选择基因，因此，它们不能在培养基中存

活。典型的选择基因编码 (a) 提供对抗生素或其它毒素, 例如氨苄青霉素, 新霉素, 氨甲蝶呤或四环素的抗性; (b) 补充营养缺陷; 或 (c) 提供不能从培养基中获得的关键营养成分的蛋白质。

其它选择基因可用于扩增待表达的基因。扩增是为生产对生长关键性的蛋白质大量需要的基因在重组细胞的连续传代染色体内串联重复的过程。哺乳动物细胞的合适的选择标记的例子包括二氢叶酸还原酶 (DHFR) 和胸苷激酶。哺乳动物细胞转化子置于选择压力下, 由于在载体中存在的标记的特征, 在该选择压力下仅专一性地适于转化子存活。在培养基中选择剂的浓度连续改变, 从而导致选择基因和编码截短 GDNF 的 DNA 扩增的条件下培养转化细胞来加强选择压力。结果, 从扩增的 DNA 合成增加量的截短的 GDNF。

例如, 经过在含氨甲蝶呤 (DHFR 的竞争性拮抗剂) 的培养基中培养所有的转化子首先鉴定用 DHFR 选择基因转化的细胞。当使用野生型 DHFR 时合适的宿主细胞是 DHFR 活性缺陷的中国仓鼠卵巢细胞系 (例如, 参见 Urlaub 和 Chasin, 美国国家科学院学报, 77(7): 4216-4220 (1980))。然后, 将转化的细胞系与水平增加的氨甲蝶呤接触。这导致合成多拷贝的 DHFR 基因, 并伴随着合成多拷贝的存在于表达载体中的其它 DNA, 如编码截短的 GDNF 蛋白质的 DNA。

启动子

本发明的表达和克隆载体典型地含有被宿主生物识别的, 可操作地连接到编码截短的 GDNF 蛋白质的核酸序列上的启动子。启动子是位于结构基因 (一般大约为 100 到 1000bp) 起始密码子上游 (5') 的不翻译序列, 它控制特定核酸序列, 如编码截短的 GDNF 的序列的转录和翻译。启动子按常规分成 2 类, 诱导性启动子和组成性启动子。诱导性启动子在其控制下与培养条件中的某些改变, 如存在或不存在一种营养成分或温度改变反应而启动增加从 DNA 的转录水平。被各种潜在的宿主细胞识别的许多启动子是熟知的。经过用限制性酶消化从来源 DNA 取出启动子并将所需的启动子序列插入载体中将这启动子可操作地连接到编码截短 GDNF 的 DNA 上。可使用天然 GDNF 启动子序列指导截短的 GDNF DNA 的扩增和/或表达。然而如果异源性启动子与天然启动子相比允许表达的蛋白质转录更多和产量更高, 且如果它与选择使用的宿主细胞系统相配伍, 那么优选该异源性启动子。

适用于原核宿主的启动子包括 β -内酰胺酶和乳糖启动子系统；碱性磷酸酶，色氨酸（trp）启动子系统和诸如tac启动子的杂合启动子。其它已知的细菌启动子也是合适的。已公开了其核苷酸序列，从而使本领域的技术人员能按需要使用接头或衔接子以提供任意所需的限制性位点将它们连接到所需DNA序列上。

用于酵母宿主的合适的启动子序列也是本领域所熟知的。酵母增强子与酵母启动子一起使用具有优势。用于哺乳动物宿主细胞的合适的启动子是熟知的且包括从诸如多瘤病毒，禽痘病毒，腺病毒（如腺病毒2），牛乳头瘤病毒，禽肉瘤病毒，巨细胞病毒，逆转录病毒，乙肝病毒且最优选猴病毒40（SV40）的病毒基因组获得的那些启动子。其它合适的哺乳动物启动子包括异源性哺乳动物启动子，例如，热休克启动子和肌动蛋白启动子。在CHO细胞中生产GDNF蛋白质的常用启动子是SR α 。见Takebe等，分子细胞生物学，8(1): 466-472 (1988)。合适的表达载体是pDSR α 2，在下面进一步描述。

增强子元件

可将增强子序列插入载体中以增强高等真核生物转录编码本发明的截短GDNF蛋白质的DNA序列。增强子是DNA的顺式作用元件，通常大约10-300 bp长，它作用于启动子上以增强其转录。增强子相对来说对取向和位置是非依赖性。已发现它们位于转录单位的5'和3'端。可从哺乳动物基因获得的一些增强子序列是已知的（例如，珠蛋白，弹性蛋白酶，白蛋白， α -甲胎蛋白和胰岛素）。然而，典型的是，可使用来自病毒的增强子。SV40增强子，巨细胞病毒早期启动子增强子，多瘤增强子和腺病毒增强子是激活真核启动子的举例性增强元件。尽管增强子可剪切进载体中的截短的GDNF DNA的5'或3'位置，但典型的是它位于启动子5'位置。

转录终止

用于真核宿主细胞（酵母，真菌，昆虫，植物，动物，人类或来自其它多细胞生物的真核细胞）的表达载体还含有为终止转录和稳定mRNA所必需的序列。这些序列通常从真核DNA或cDNA的5'及偶尔从3'非翻译区获得。这些区域含有转录为编码截短的GDNF的mRNA非翻译部分的多聚腺苷酸化片段的核苷酸片段。

含一个或多个上述成分及所需截短的GDNF编码序列的合适的载

体的构建以标准连接技术实现。分离的质粒或 DNA 片段被裂解，加尾并以所需的顺序再连接以产生所需的质粒。为证实已构建了正确的序列，可使用连接混合物转化大肠杆菌，用已知技术，如上面所述的氨苄青霉素或四环素抗性可选择成功的转化子。然后制备来自转化子的质粒，以限制性内切酶消化分析，和/或测序以证实所需构建体的存在。

也可使用在哺乳动物细胞中提供瞬时表达编码截短 GDNF 的 DNA 的载体。一般来说，瞬时表达涉及使用能在宿主细胞中有效复制的表达载体，使宿主细胞积累许多拷贝的表达载体，然后合成高水平的该表达载体编码的所需蛋白质。包含合适的表达载体和宿主细胞的瞬时表达系统允许用于对克隆 DNA 编码的蛋白质进行常规阳性鉴定，以及用于迅速筛选所需生物学或生理学特性的该蛋白质。因此，瞬时表达系统在鉴定蛋白质的变体中特别有用。

宿主细胞的选择和转化

本发明还提供了用核酸序列转化的宿主细胞（例如，细菌，哺乳动物，昆虫，酵母，或植物细胞）用于表达重组截短的 GDNF 蛋白质。该转化的宿主细胞在允许核酸序列表达的合适条件下培养转化的宿主细胞。合适的宿主细胞的选择及用于转化，培养，扩增，筛选和产品生产及纯化的方法在本领域中是熟知的。例如见，Gething 和 Sambrook，自然 293:620-625 (1981)，或 Kaufman 等，分子细胞生物学，5 (7):1750-1759 (1985) 或 Howley 等，美国专利号 4419446。根据 Lin 等（美国专利申请号 07/855,413；申请号 PCT/US 92/07888, WO 93/06116）的描述可在大肠杆菌中表达截短的 GDNF，该文献中涉及成熟 GDNF 的表达。其它举例性的材料和方法在下文进一步详细讨论。在合适的培养基中培养转化的宿主细胞，然后以本领域的技术人员已知的合适方法选择性地从培养基（或从细胞，如果在细胞内表达）中回收，分离和纯化表达因子。

用于克隆或表达本文的载体的合适的宿主细胞是如上所述的原核、酵母或高等真核细胞。原核宿主细胞包括但不限于真细菌，如革兰氏阴性或革兰氏阳性生物，例如大肠杆菌，诸如枯草杆菌的杆菌属，诸如铜绿假单胞菌的假单胞菌属种类，伤寒沙门氏菌或粘质沙雷氏菌。另外，体外克隆方法，例如，PCR 或其它核酸聚合酶反应，也是合适的。

除了原核宿主细胞外，诸如丝状真菌或酵母的真核微生物可以是表

达截短的 GDNF 蛋白质的合适的宿主。在低等真核宿主微生物中最常使用的是啤酒糖酵母或常见面包酵母，但许多其它的属，种类和菌株也是熟知的和通常可得到的。

表达糖基化的截短的 GDNF 蛋白质的合适的宿主细胞来自多细胞生物。该宿主细胞具有复杂的加工和糖基化活性。原则上，可使用任何高等真核细胞培养物，不管该培养物是脊椎的还是非脊椎动物的细胞，包括植物和昆虫细胞。脊椎动物细胞一般用于在培养物（组织培养物）中繁殖脊椎动物细胞，这是一种熟知的方法。有用的哺乳动物宿主细胞系的例子包括、但不限于以 SV 40 转化的猴肾 CV1 细胞系（COS-7），人胚胎肾细胞系（293 或亚克隆用于在悬浮培养物中生长的 293 细胞），幼仓鼠肾细胞和中国仓鼠卵巢细胞。其它合适的哺乳动物细胞系包括但不限于 Hela，小鼠 L-929 细胞，来自 Swiss，Balb-c 或 NIH 小鼠的 3T3 细胞系，BHK 或 HaK 仓鼠细胞系。

作为适用于本发明的宿主细胞同样有用的有细菌细胞。例如，大肠杆菌的各种菌株（例如 HB101，DH5 α ，DH10 和 MC1061）在生物技术领域中作为宿主细胞是熟知的。也可采用链霉菌属种类的各种菌株和类似菌株。目前优选的用于生产截短的 GDNF 蛋白质的宿主细胞是细菌细胞（例如，大肠杆菌）和哺乳动物细胞（如中国仓鼠卵巢细胞，COS 细胞等）。

宿主细胞用上面所述的表达或克隆载体转染并优选转化并在常规营养培养基中培养。培养基可改变为适于诱导启动子，选择转化子或扩增编码所需序列的基因。使用本领域的技术人员熟知的且选择为适于所涉及的宿主细胞的标准技术进行转染和转化。例如，对于无细胞壁的哺乳动物细胞，可使用磷酸钙沉淀方法。也可使用电穿孔，显微注射和其它已知技术。

培养宿主细胞

用于生产本发明的截短的 GDNF 蛋白质的转化细胞在合适的培养基中培养。该培养基可按需要补充激素和/或其它生长因子（如胰岛素，转铁蛋白或表皮生长因子），盐（如氯化钠，钙，镁和磷酸盐），缓冲液（如 HEPES），核苷（如腺苷和胸苷），抗生素（如庆大霉素），痕量元素（定义为无机化合物，通常以微摩尔范围的终浓度存在），葡萄糖或其它能源。也可包括适当浓度的其它补充物，正如本领域的技术

人员可预料到的。对于选定的宿主细胞使用的合适的培养条件，如温度，pH等也是本领域的技术人员所熟知的。

还可能经过同源重组或使用将控制元件导入已含编码 GDNF 的 DNA 的细胞中的重组生产方法来生产截短的 GDNF 蛋白质。同源重组是最初为靶向基因以诱导或修正转录活性基因中的突变而建立的一种技术（Kucherlapati，核酸研究和分子生物学进展，36:301 (1989)）。该基本技术作为将特异性突变导入哺乳动物基因组特定区域（Thomas 等，细胞，44:419-428, 1986; Thomas 和 Capecchi，细胞，51:503-512, 1987; Doetschman 等，美国国家科学院学报，85:8583-8587, 1988）或校正缺陷基因内的特异性突变（Doetschman 等，自然，330:576-578, 1987）的方法而建立。举例性的同源重组技术在 U.S. 5272071（EP 91903051，欧洲专利公开号 505 500; PCT/US 90/07642，国际公开号 WO 91/09955）中描述，本文引用其说明书以供参考。

通过同源重组，将要插入基因组的 DNA 序列经过将其附着于靶向 DNA 上而导入目的基因的特定区域。靶向 DNA 是与基因组 DNA 区域互补（同源）的 DNA。互补于基因组特定区域的靶向 DNA 的小片段在 DNA 复制过程中与母链接触。插入细胞后通过共享同源区与内源性 DNA 的其它片段杂交和重组是 DNA 的一般特征。如果该互补链附着于含突变或 DNA 的不同序列的寡核苷酸上，作为重组的结果它也掺入新合成的链上。作为校正功能的结果，DNA 的新序列可能用作模板。因此，将转移的 DNA 掺入基因组。

如果已知特定基因的序列，如 GDNF 的核酸序列，蛋白原前体序列或表达控制序列，可合成或经过如在感兴趣的特定识别位点结合区合适的限制性酶切消化天然 DNA 获得与基因的选定区域互补的 DNA 片段。该片段在插入细胞时用作靶向序列且在基因组内与其同源区杂交。如果在 DNA 复制期间发生杂交，该 DNA 片段及附着于其上的任何其它序列可充当冈奇片段并反向填补（backstitched）进新合成的 DNA 子链中。

在本发明中，附着到靶向 DNA 的这些片段上的是可与 GDNF 蛋白质的表达相互作用的 DNA 区域。例如，启动子/增强子元件，抑制子或外源性转录调节元件以足以影响编码所需截短的 GDNF 的 DNA 转录的邻近距离和取向插入目的宿主细胞的基因组中。控制元件不编码截短的

GDNF，但控制宿主细胞基因组中存在的部分 DNA。因此，不用编码截短的 GDNF 基因本身的 DNA 转染而使用与 DNA 调节片段偶联的靶向 DNA（含与感兴趣的内源性基因同源的区域）可实现截短的 GDNF 蛋白质的表达，其中该 DNA 调节片段含内源性基因序列和转录截短的 GDNF 蛋白质的可识别的信号。

根据本发明，也可使用同源重组方法以修饰含正常情况下转录静止的 GDNF 基因的细胞以产生表达 GDNF 的细胞。然后可加工 GDNF 蛋白质以形成截短的 GDNF 蛋白质。

截短的 GDNF 药用组合物

截短的 GDNF 蛋白质产物的药用组合物典型地包括与一种或多种药学上和生理学上可接受的制剂材料混合的治疗上有效量的截短的 GDNF 蛋白质产物。合适的制剂材料包括但不限于抗氧化剂，防腐剂，着色剂，调味剂和稀释剂，乳化剂，悬浮剂、溶剂，填充剂，混合剂，缓冲剂，传递载体，稀释剂，赋形剂和/或药用佐剂。例如，合适的载体可以是注射用水，生理盐水溶液或人工脑脊液（CSF），很可能补充在肠胃外用药的组合物中常用的其它物质。中性缓冲盐水或与血清白蛋白混合的盐水是进一步的举例性载体。

载体中的第一溶剂可以是含水或不含水性质的溶剂。另外，载体可含有用于修饰或维持组合物的 pH，渗透压，粘度，澄清性，颜色，无菌性，稳定性，溶解率或气味的其它药用上可接受的赋形剂。同样，载体还可含有用于修饰或维持截短的 GDNF 蛋白产物的稳定性，溶解率或释放率或用于促进截短的 GDNF 蛋白质产物吸收或穿透脑血屏障的其它药学上可接受的赋形剂。该赋形剂是通常和习惯上用于配制以单位剂量或多剂量形式用于肠胃外用药或以从埋植泵连续或间歇性输注用于直接输注入 CSF 的剂型的那些物质。

一旦配制了治疗性组合物，可作为溶液，悬液，凝胶，乳剂，固体或脱水或冻干的粉末在无菌小瓶中贮存。该组合物可以以易于使用的形式或以例如冻干，需在用药前重配的形式贮存。

最佳药物制剂可由本领域的技术人员根据用药途径和所需剂量来测定。例如，参见《Remington 的药用科学》，第 18 版（1990，Mack 出版公司，Easton, PA, 18042），1435-1712 页，本文引用其公开内容以供参考。组合物也可包括诸如聚乳酸，聚乙醇酸等的多聚体化合物或

在脂质体中的特定制品。也可使用 Hylauronic acid, 它具有延长在循环中的保留时间的效力。该组合物可影响该蛋白质及衍生物的物理状态, 稳定性体内释放速率和体内清除速率。

其它有效的用药形式, 如肠胃外缓释制剂, 吸入雾剂, 口服活性制剂或栓剂也是可想象的。通常的截短的 GDNF 蛋白质产物药用组合物被配制成肠胃外用, 例如脑静脉内注射。该肠胃外用治疗性组合物典型地以无致热原的包含在药用上可接受的载体中的截短 GDNF 蛋白质产物的肠胃外可接受的水溶液的形式。一个优选的载体是生理盐水。

还预期某些含截短的 GDNF 蛋白质产物的组合物以口服方式用药。以该方式用药的截短的 GDNF 蛋白质产物可制成胶囊且可用或不用通常在固体剂量形式的制剂中使用的那些载体来配制。该胶囊可设计成当生物可获得性最大且系统前降解最小时在胃肠道的该位点释放该制剂的活性部分。可包括其它赋形剂以利于截短的 GDNF 蛋白质产物的吸收。也可采用稀释剂, 调味剂, 低熔点蜡, 植物油, 润滑剂, 悬浮剂, 片剂崩解剂和粘合剂。

截短的 GDNF 蛋白质产物的用药:

截短的 GDNF 蛋白质产物可经过皮下, 肌肉内, 静脉内, 经肺, 经皮肤, 鞘内或脑内途径进行肠胃外用。不穿过血脑屏障的蛋白质生长因子可直接经脑内给药或与将它们运输通过屏障的其它成分相联给药。优选的是截短的 GDNF 蛋白质产物经脑静脉内给药或施用到脑或脊索蛛网膜下隙。截短的 GDNF 蛋白质产物也可经脑内直接施用到脑实质中。含包埋于生物可降解的多聚体基质中的神经营养因子的脑内缓释型埋植体也能传递截短的 GDNF 蛋白质产物。截短的 GDNF 蛋白质产物可以化学修饰或包装的形式经脑外用以便其穿过血脑屏障, 或可与一种或多种能促进截短的 GDNF 蛋白质产物穿过屏障的试剂相联用药。例如, NGF 与单克隆抗运铁蛋白受体抗体的连接物已显示经过结合运铁蛋白受体而运输到大脑。为了达到截短的 GDNF 蛋白质产物所需剂量, 可每天重复或偶尔注射来用药, 或从恒速或可控制流速的植入泵连续或间歇性地输注截短的 GDNF 蛋白质产物。剂量频率取决于所配制的截短的 GDNF 蛋白质产物的药物动力学参数和用药途径。

不管何种用药方式, 典型地根据体重或体表面积计算具体剂量。对于与大脑有关的疾病, 典型地根据病人的近似大脑重计算具体剂量, 也

可根据体重或体表面积估计。测定包括上述各种配方的合适治疗剂量所需的计算的进一步精选可由本领域的普通技术人员按常规,特别是根据本文公开的剂量信息和试验来作出。通过使用为测定所用的剂量与合适的剂量反应数相联系而建立的测定来确定合适的剂量。对于治疗具体症状的方法的最终剂量方案由出诊医生考虑改变药品作用的各种因素,如病人的年龄,状态,体重,性别和饮食,感染的严重性,用药时间和其它临床因素来确定。

本发明的截短的 GDNF 蛋白质产物在神经性疾病的治疗中也可单独或与其它生长因子结合使用。例如,截短的 GDNF 蛋白质产物可与神经生长因子结合在治疗某些形式的神经性疾病中使用。另外,其它因子或其它分子,包括化学组合物可与截短的 GDNF 蛋白质产物一起使用。在帕金森氏病的治疗中,可预期使用截短的 GDNF 蛋白质产物本身或与施用左旋多巴结合,其中截短的 GDNF 可增加内源性多巴胺的生产和神经元对浓度增加的多巴胺的吸收。

如上所述,也期望其它的神经营养或神经元营养因子对于治疗某些神经元细胞群体或某些类型的损伤或疾病有用或是必需的。可用于与截短的 GDNF 结合的其它因子包括但不限于:诸如胰岛素,胰岛素样生长因子,表皮生长因子,血管活性生长因子,垂体腺苷酸环化酶激活多肽,干扰素和促生长素抑制素的促细胞分裂原;诸如大脑产生的神经营养因子的神经营养因子,神经营养蛋白-3,神经营养蛋白4/5,神经营养蛋白-6,胰岛素样生长因子,睫状神经营养因子,酸性和碱性成纤维细胞生长因子,成纤维细胞生长因子-5,转化生长因子- β ,可卡因苯异丙胺调节的转录本(CART)和成熟 GDNF;其它的生长因子,如表皮生长因子,白血病抑制因子,白细胞介素,干扰素和集落刺激因子;以及与这些因子功能相当的分子和物质。

预期连续用药或持久传递截短的 GDNF 蛋白质产物对于给定治疗具有优势。尽管经机械装置,如用输注泵可实现连续用药,预期可实施连续或几乎连续用药的其它方式。例如,化学衍变可产生蛋白质的持续释放形式,它具有根据确定的剂量方案以可预计量在血流中连续存在的效果。因此,截短的 GDNF 蛋白质产物包括衍变成实现该连续用药的截短的 GDNF 蛋白质。

本发明还包括截短的 GDNF 蛋白质细胞疗法,例如,脑内植入生产

截短的 GDNF 蛋白质的细胞。本发明的该实施方案可包括将能合成并分泌截短的 GDNF 蛋白质的生物活性形式的细胞植入病人体内。这种生产截短的 GDNF 蛋白质的细胞可以是在正常情况下不生产神经营养因子但被修饰成生产截短的 GDNF 的细胞,或者它们可以是经过用适于表达和分泌截短的 GDNF 蛋白质的多核苷酸转化而增强了其生产 GDNF 的能力的细胞。为了减小因施用外源种类的 GDNF 而在病人中的潜在的免疫学反应,优选该细胞是人类来源的且产生截短的人 GDNF 蛋白质。

植入的细胞可包入胶囊以避免细胞渗入脑组织。人类或非人类动物细胞可以在生物相容性,半通透性聚合外套或膜中植入病人以允许释放截短的 GDNF 蛋白质产物,但防止病人免疫系统或来自周围组织的其它有害因素破坏该细胞。另外,来自体内的转化成生产截短的 GDNF 的病人自身的细胞可不经这类包被直接植入病人。

活细胞的膜包被的方法学对本领域的普通技术人员来说是熟悉的,可实现包被细胞的制备并将其植入病人。例如,见美国专利号 4892538; 5011472; 和 5106627, 本文引用其说明书以供参考。用于包装活细胞的系统也在 Aebischer 等的 PCT 申请 WO 91/10425 中进行了描述, 本文特别引用以供参考。也参见 Aebischer 等的 PCT 申请 WO 91/10470; Winn 等, 实验神经学, 113:322-329, 1991; Aebischer 等, 实验神经学, 111:269-275, 1991; Tresco 等, ASAIO, 38:17-23, 1992, 此处引用这些文献以供参考。

也可想象体内截短的 GDNF 蛋白质基因治疗。其中将编码截短的 GDNF 蛋白质的核酸序列直接导入病人。例如, 经过局部注射含或不含合适的传递载体, 如腺伴随病毒载体的核酸构建体可将编码截短的 GDNF 蛋白质的核酸序列导入靶细胞。任选的病毒载体包括, 但不限于逆转录病毒, 腺病毒, 单纯疱疹病毒和乳头瘤病毒载体。经过局部注射所需的核酸构建体或含所需核酸序列的其它合适的传递载体, 直接注射(裸露 DNA), 受体介导的转移(配体-DNA 复合物), 或微粒子轰击(基因枪)可在体内实现物理转移。

应注意本文描述的截短的 GDNF 蛋白质制剂可用于脊椎动物及人类的应用, 且术语“病人”不应以限制性方式解释。在脊椎动物应用的情况下, 剂量范围应与上面限定的相同。

作为进一步鉴定本发明的截短的 GDNF 蛋白质的一种工具, 可形成

结合截短的 GDNF 蛋白质，如 X-[Cys⁴¹-Cys¹³³]-Y 氨基酸序列内的抗原决定簇的抗体。本领域的普通技术人员可使用熟知的，公开的程序获得特异性地识别和结合由本发明的氨基酸序列所编码的各种蛋白质的单克隆和多克隆抗体或重组抗体。然后，该抗体可用于纯化和鉴定截短的 GDNF 蛋白质。另外，该抗体可用作它们所针对的蛋白质的治疗性抑制剂。

考虑到下列举例性的实施例将会明白本发明的其它方面和优势。实施例 1 阐述了成熟 GDNF 在哺乳动物细胞系统中的表达及截短的 GDNF 蛋白质的制备。实施例 2 阐述了成熟 GDNF 在细菌细胞系统中的表达。实施例 3 阐述了各种截短的 GDNF 蛋白质在细菌细胞系统中的表达。实施例 4 在多巴胺能神经元神经营养活性试验中比较了成熟 GDNF 蛋白质与截短的 GDNF 蛋白质的生物学活性。

实施例

实施例 1

成熟的人 GDNF 在 CHO 细胞中的表达及 CHO 细胞产生的截短的 GDNF 蛋白质的纯化

材料

下列材料用于在二氢叶酸还原酶缺陷型 CHO 细胞（CHOd⁻细胞，例如，如 Urlaub 和 Chasin，美国国家科学院学报 77(7): 4216-4220 (1980) 所述）中表达人 GDNF。

CHOd⁻培养基含：Dulbecco 改良的 Eagle 培养基（DMEM）- 高葡萄糖（Gibco/BRL）；5% 胎牛血清（HyClone）；MEM 非必需氨基酸（1%）（Gibco/BRL）；次黄嘌呤/胸苷（1%）（Gibco/BRL）；和谷氨酰胺/青霉素/链霉素（1%）（Irvine Scientific）。

选择性培养基含：DMEM（高葡萄糖）；5% 透析的胎牛血清（HyClone）；MEM 非必需氨基酸；和谷氨酰胺/青霉素/链霉素。

2X HEPES - 缓冲盐水（HBS）含：280mM NaCl；10mM KCl；1.5mM Na₂HPO₄；12mM 葡聚糖；和 50mM HEPES。

Tris-缓冲盐水加 Tween（TBST）含：137mM NaCl；20mM Tris/HCl pH 7.5；和 0.1% Tween-20。

方法

转染和选择

将 CHOd⁻细胞（第 20 代）以 8×10^5 个细胞/皿的密度接种入 60mm 组织培养皿（Falcon）的 CHOd⁻生长培养基中。在第二天，转染前大约 3 小时，用新鲜培养基换掉细胞中的培养基。

使用熟知技术制备含合适的 GDNF cDNA 的质粒构建体。例如，基本上按 1990 年 3 月 29 日申请的共同拥有和共同未决的美国专利申请系列号 501904（也见 1990 年 3 月 18 日申请的欧洲专利申请号 90305433，公开号 EP 398753 和 WO 90/14363 (1990)，本文引用其说明书以供参考）所述的方法制备质粒构建体 pDSR α 2。说明载体的结构组织的举例性质粒图谱在图 2 中描述。本领域的技术人员可预料到也可使用各种编码成熟 GDNF 蛋白质的核酸序列，如在图 1, 3 和 4 中所述的序列。

经限制性酶消化从 pcDNA3 为基础的表达载体（Invitrogen, San Diego, CA）中回收含人 GDNF 编码序列和共有 Kozak 序列，CCACC(ATG)的 Hind III-XbaI DNA 片段。将 DNA 片段直接克隆进 Hind III/XbaI 切割 pDSR α 2 中。所得的质粒称为 pSW5。pSW5 的质粒 DNA 在转染前在 PvuI 位点线性化。

pDSR α 2（图 2）是质粒 pCD（Okayama 和 Berg, 分子细胞生物学, 3: 280-289, 1983）的衍生物，具有 3 个主要的改变：（i）SV40 多聚腺苷化信号被来自牛卵泡刺激激素的 α 亚基， α -bFSH 的信号（Goodwin 等，核酸研究, 11: 6873-6882, (1983)取代，（ii）将小鼠二氢叶酸还原酶小基因（Gasser 等，美国国家科学院学报, 79: 6522-6526, 1982）插入表达盒下游以允许选择和扩增转化子；和（iii）按以前所述（Takebe 等，分子细胞生物学, 8: 466-472, 1988）克隆含：“R-元件”和部分人 T-细胞白血病病毒 1 型（HTLV-1）的长末端重复（LTR）的“U5”序列的 267 bp 片段并插入 SV40 启动子和剪接信号之间。

制备含终浓度为 $3.0 \mu\text{g}/\text{皿}$ GDNF - 质粒 DNA， $7.0 \mu\text{g}/\text{皿}$ 小鼠肾基因组载体 DNA（Clontech）， $25 \mu\text{l}/\text{皿}$ 2.5M CaCl_2 和无菌蒸馏水至终体积为 $250 \mu\text{l}/\text{皿}$ 的 DNA 溶液。同样制备含 pDSR α 2 载体 DNA 的 DNA 溶液或单独的载体 DNA 分别作为阳性和阴性对照。向等体积的 $2 \times$ HEPES - 缓冲盐水中逐滴加入 DNA 溶液，同时向溶液通入气泡。将

DNA/HBS 溶液在室温下培养 30 分钟。

从 CHO^d 细胞培养物中去掉培养基，在每皿中加入 500 μ l DNA 溶液。培养皿在室温下培养 30 分钟，此后向每皿中加入 CHO^d 培养基（5.0ml）。然后将皿在 37℃ 培养过夜。

在第二天用新鲜的 CHO^d 培养基换掉培养基。在第二天当细胞达到汇合时，培养物用胰酶消化并以 1 $\times$ 60mm 皿到 8 $\times$ 100mm 皿的比率再涂布到 100mm 皿（Falcon）上。在选择性培养基中再涂布细胞。每 2 到 3 天用新鲜培养基再培养培养物。

15 天后，使用玻璃克隆筒分离转染的细胞集落，胰蛋白酶消化并再涂布到 24 孔盘（Falcon）上。从 GDNF/pSW5 - 转染的细胞分离总共 40 个集落。剩余的在盘上的细胞经胰酶消化，合并并且再涂布进 2 个 100mm 皿中（每个 DNA 构建体一个细胞库）。

转染细胞的筛选

24 孔及合并培养物生长到汇合，此时去掉生长培养基并用无血清培养基取代（400 μ l/孔或 4ml/皿）。将细胞培养 48 小时并收获条件培养基。以 Western 印迹分析条件培养基样品的 GDNF 蛋白质表达。条件培养基的等分试样（20 μ l 或 40 μ l）用电泳样品缓冲液（有或无 β -巯基乙醇）稀释。含 β -巯基乙醇的样品煮沸 3 分钟（还原条件）。将还原的和非还原的样品走 16% Tris - 甘氨酸凝胶（Novex）。将凝胶电印迹到硝酸纤维素滤膜（Schleicher 和 Schuell BA-83, 0.2 μ ）上。用 TBST 漂洗印迹，然后在 TBST 中的 5% 干牛牛奶（Carnation）的封闭溶液中室温下培养 30 分钟。然后将印迹用 GDNF 抗血清（抗大肠杆菌产生的 GDNF 的兔多克隆抗血清；1:1000 溶于 5% 牛奶/TBST 中）在室温下处理 1 小时。然后用 TBST 漂洗印迹并用 1% 牛奶/TBST 洗涤 1 $\times$ 10 分钟和 2 $\times$ 5 分钟。然后用抗兔 Ig 辣根过氧化物酶连接的第二抗体（1:15,000, 在 1% 牛奶/TBST 中）处理 20 分钟。漂洗印迹并用 TBST 洗涤 1 $\times$ 20 分钟和 2 $\times$ 10 分钟，接着用 ECL 试剂（Amersham）处理 1 分钟并对 Hyperfilm-ECL（Amersham）曝光。

下面过程描述了从 1 升条件培养基中纯化 CHO-表达的 GDNF 和 CHO - 产生的剪短的 GDNF 同型二聚体。由于在 CHO 培养基中明显的蛋白酶作用在残基 31 处剪断该链，该过程可包括在纯化期间使用蛋白酶抑制剂。

步骤 1, 玻珠色谱:

经过加入 1/5 体积的 1M MES, pH 6.0, 使无血清条件培养基含 20mM 2-[N - 吗啉代]乙烷磺酸 (MES), pH6.0。加入 25ml 用 20mM MES, pH 6.0 平衡的 SP 琼脂糖大颗粒树脂 (Pharmacia) 并在 4 °C 搅拌 1 小时。经过使其沉降并倒掉条件培养基来收集树脂。倒掉的培养基通过多孔圆片滤纸过滤以回收未沉降的树脂。沉降的树脂及经过滤回收的树脂重悬并倒入 2.5 cm 直径的柱中, 用 3 倍柱体积的 0.15 M NaCl, 20mM MES, pH 6.0 (A 缓冲液) 洗涤。用 300ml 从 A 缓冲液到 1.0M NaCl, 20mM MES, pH 6.0 (B 缓冲液) 的梯度以 0.2 柱体积/分的流速洗脱蛋白质, 在 280 nm 处监测吸光率。收集含 1.1 柱体积的级份。以 Western 印迹分析检测级份中 GDNF 的存在。合并含 GDNF 的级份用于进一步纯化。GDNF 在 0.3 和 0.6M NaCl 之间洗脱。

步骤 2, HPLC C4 色谱:

使步骤 1 的合并液含 0.1% (v/v) 三氯乙酸 (TFA), 通过 0.45 微米滤膜真空过滤, 使之通过用含水的 10 % 的乙腈, 0.1% TFA (A 缓冲液) 条件化的 Vydac C4 柱 (0.46×25 cm)。用 2 %/分的线性梯度在 50 分钟内从 A 缓冲液到含水的 90 % 乙腈, 0.1% TFA (B 缓冲液) 洗脱蛋白质, 在 280 nm 处测吸光率。收集 1 毫升级分, 以 Western 印迹分析检测 GDNF 的存在。在 45 % 和 55 % 乙腈之间洗脱 GDNF。级分在真空中干燥。

步骤 3, 高效 S 色谱:

来自步骤 2 的含 GDNF 的级分重溶于 1ml 0.15M NaCl, 10mM Tris, pH 8.0 中并使之过 0.75×7.5 cm TSK-Gel 5WP 高效 S 柱 (Toso Haas)。以 1ml/分的流速在 50 分钟内从 0.15M NaCl, 10mM Tris, pH 8.0 (A 缓冲液) 到 1.0 M NaCl, 10mM Tris, pH 8.0 (B 缓冲液) 流过 0.4%/分的线型梯度。在 280 nm 处测量吸光率收集 1 分钟级分。在 35 % B 缓冲液时, 将梯度变成在 10 分钟内 6.5%/分。级分的 Western 印迹分析显示出四个主要的 GDNF 成分。在 0.4%/分钟梯度期间洗脱出其中的 3 种成分, 在 6.5%/分梯度期间洗脱出第 4 种成分。制备具有相似成分的合适的合并液并进行测序。测序分析将大约 29 - 36 kD 的合并液鉴定为 [Arg³²-Ile¹³⁴] 截短的 GDNF 蛋白质。大约 38-40 kD 的成分鉴定为 [Arg³²-Ile¹³⁴] 截短的 GDNF/成熟 GDNF 异型二聚体。最后, 在后一部分梯度期间分

离的大约 41-44kD 的成分经测序鉴定为成熟 GDNF 同型二聚体。

实施例 2

在大肠杆菌中生产的成熟人 GDNF

根据 Lin 等所述的方法（1994 年 5 月 23 日申请的美国专利申请号 08/182,183 和其优先权申请；1992 年 9 月 17 日申请的 PCT/US 92/07888（WO 93/06116）；和欧洲专利申请号 92921022.7（公开号 EP 610254）；本文引用其说明书以供参考）实现成熟的人 GDNF 在细菌中的表达。根据本发明的描述，本领域的普通技术人员可预料到各种材料和方法容易用于或适应于在大肠杆菌和其它细菌中进行合适的表达。例如，选择的多核苷酸，如在图 1, 3 和 4 中所述的那些多核苷酸可用于表达方法。表达的成熟 GDNF 的再折叠和纯化

按下面所述在 5 °C（除非另有说明）处理转化的细胞：使用含 5mM EDTA 的 25mM Tris, pH8.5 将细胞沉淀（paste）（30gm）悬浮于 200 ml 的终体积中以产生 15 %（w/v）的最终细胞悬液。使用 Biospec 手握式低剪切匀浆器彻底分散细胞。将悬液以 14,500 psi 通过微流仪 2 次以破碎细胞并释放包含体。然后将所得的匀浆产物在 16,000xg 下离心 30 分钟。离心所得的包含体沉淀经过使用如前所述的 Biospec 匀浆器重悬于冷水中达到终体积为 240ml 以形成悬液进行洗涤。保留该悬液样品用于 GDNF 表达水平的 HPLC 分析。所得的悬液在 16,000xg 下离心 30 分钟。弃去上清，向离心瓶中加入少量冷水并轻轻旋动以去掉在包含体沉淀顶部疏松形成的膜层。用 Biospec 匀浆器使用足够体积的冷水重悬沉淀以产生 2 mg/ml 的 GDNF 浓度。经过将最终包含体悬液（25ml）与含 180mM 半胱氨酸 HCl，和 50mM Tris HCl, pH 8.7 的 8M 盐酸胍（25ml）混合溶解包含体。溶解混合物在 25 °C 搅拌 60 至 90 分钟，然后伴随混合将其注入含 20mM Tris HCl, pH 8.75 和 0.2M 盐酸胍的 2M 尿素（450ml, 5 °C）中。在 5 °C 缓慢搅拌重折叠混合物 72 小时。

按如下方法部分纯化重折叠的 GDNF：伴随着迅速搅拌向重折叠混合物中加入 20mM 乙酸钠缓冲液（250ml, pH 5），5 °C。用冰乙酸将 pH 调到 5。经过以 13,600xg 在 5 °C 离心 45 分钟去掉所得的沉淀。离心的上清用作上样溶液，用于涉及用 SP - 大颗粒树脂（Pharmacia）的阳离子交换色谱的下一步纯化。使用 20mM 乙酸钠（pH5）作为平衡，漂洗和洗脱缓冲系统在 5 °C 操作柱。用 5 倍柱体积（CV）的 0.2N NaOH

卫生处理树脂床 (5ml) , 然后用乙酸缓冲液 (5CV) 平衡。将上样溶液 (190ml) 以 0.5 CV/分过柱, 接着用乙酸缓冲液以相同的流速进行 10 CV 洗涤。然后用在乙酸缓冲液中的从 0.3M 到 0.9M 的 20 CV 的 NaCl, 线性梯度以 0.1 CV/分的流速从树脂中洗脱出 GDNF。以 280 nm 的吸光率监测柱的洗脱并收集级分进行 SDS-PAGE 试验。从前面 GDNF 峰 10 % 的峰高到峰后 10 % 的峰高合并含 GDNF 的级分。在该合并液中的蛋白质完全是 GDNF, 根据所使用的生产菌株, 含有 32 % 到 12 % 修饰的 GDNF 形式的污染物。然后将合并液用 PBS 或其它配方形式的缓冲液透析, 在某些情况下, 经超滤浓缩到 25mg/ml。为了比较涉及相应的生产菌株的制品的纯度, 用反相 HPLC, 阳离子交换 HPLC, 质谱法、和内毒素水平鉴定以该方法纯化的 GDNF 野生型和类似物形式。

实施例 3

截短的 GDNF 在大肠杆菌中的重组生产

基本上按照 Lin 等所述的技术 (1994 年 5 月 23 日申请的美国专利申请号 08/182,183, 出处同上) 生产举例性的截短的 GDNF 蛋白质。也可使用如上所述的另选的细菌表达材料和方法。大肠杆菌表达的截短的 GDNF 蛋白质包括分别在图 5, 6 和 7 中所述的 $[\text{Pro}^{23}\text{-Ile}^{134}]$, $[\text{Arg}^{32}\text{-Ile}^{134}]$ 和 $[\text{Gly}^{33}\text{-Ile}^{134}]$ 截短的 GDNF 蛋白质。编码这些举例性的截短的 GDNF 蛋白质的多核苷酸按图 5, 6 和 7 所述构建, 但也可使用诸如在图 1, 3 和 4 中所述的那些多核苷酸的相应多核苷酸。按在 PCR 技术, DNA 扩增的原则和应用, Henry A, Erlich 编辑, Stockton Press, NY, 1989 (第 6 章, 使用 PCR 改造 DNA) 中所述的标准 PCR 程序构建多核苷酸, 本文引用该文献以供参考。

实施例 4

多巴胺能神经元神经营养活性的生物测定

在定量基础上评价实施例 3 的大肠杆菌表达的 $[\text{Pro}^{23}\text{-Ile}^{134}]$, $[\text{Arg}^{32}\text{-Ile}^{134}]$, $[\text{Gly}^{33}\text{-Ile}^{134}]$ 和 $[\text{Lys}^{37}\text{-Ile}^{134}]$ 截短的 GDNF 蛋白质及实施例 1 的 CHO 产生的 $[\text{Arg}^{32}\text{-Ile}^{134}]$ 截短的 GDNF 蛋白质促进灰质多巴胺能神经元的多巴胺吸收的能力。

材料

下列材料用于在截短的 GDNF 蛋白质存在下评价多巴胺能神经元存活的试验:

细胞培养基

高葡萄糖 Dulbecco's 修饰的 Eagle's 培养基 (DMEM , 目录号 11965-092) , Ham's F12 培养基 (F12, #11765-021) , 无碳酸氢钠的 Leibovitz's L15 培养基 (#41300-039) , B27 培养基补充剂 (#17504-010) , 青霉素/链霉素 (#15070-014) , L - 谷氨酰胺 (#25030-016) , Dulbecco's 磷酸盐缓冲盐水 (D-PBS, #14190-052) , 具有钙和镁盐的 Hank 平衡盐溶液 (HBSS; #24020-026) , N - 2 - 羟乙基哌嗪 - N' - 2 - 乙磺酸 (HEPES; #15630-015) , 小鼠层粘连蛋白 (#23017-015) , 和牛血清白蛋白级分 V (#110-18-017) 均来自 GIBCO , Grand Island, NY。热灭活的马血清来自 HyClone, Logan, Utah。伴白蛋白 (C-7786) , 多聚 - L - 鸟氨酸溴氢化物 (P-3655) , 牛胰岛素 (I-5500) , 人运铁蛋白 (T-2252) , 腐胺 (P-6024) , 孕酮 (P-6149) , 亚硒酸钠 (S-9133) , 三碘苯甲酰氨基葡萄糖 (M-3383) 均来自 Sigma 化学公司 , Saint-Louis, MO。木瓜蛋白酶, 脱氧核糖核酸酶 I (DNAase) 和卵清蛋白 (木瓜蛋白酶解离系统) 来自 Worthington Biochemicals, Freehold, NJ。

Falcon 无菌 96-孔微量平板 (#3072) , 组织培养塑料用品和聚丙烯离心管来自 Becton-Dickinson, Oxnard, CA。Nunc Lab-Tek 组织培养室盖玻片 (#136439) 来自 Baxter, Irvine, CA; 20 μ m (#460) 尼龙网来自 Tetko, Elmsford, NY; 4[#]解剖镊和 4"解剖剪来自 Roboz Surgical, Washington, DC。

抗体和相关试剂

多克隆兔抗酪氨酸羟化酶抗体 (TE101) 来自 Eugene Tech, Ridgefield Park, NJ; 多克隆兔抗 - 神经元特异性烯醇化酶抗体 (NSE, AB951) 来自 Chemicon, Temecula, CA; 生物素标记的羊抗兔 IgG 和过氧化物酶连接的亲和素/生物素复合物 (ABC Elite; Vectastain 试剂盒 PK-6100) 来自载体实验室, Burlingame, CA。3', 3' - 二氨基联苯胺来自 Cappel Laboratories, West Chester, PA。PBS 中的 Superblock 封闭缓冲液来自 Pierce 化学公司, Rockford, IL。Triton X-100 (X100) , Nonidet P-40 (N6507) 和过氧化氢 (30 % , v/v, H1009) 来自 Sigma。GBR-12909 多巴胺吸收抑制剂 (D-052) 来自 RBI, Natick, MA。³H-多巴胺 (滴定多巴胺, NE-131; 21 Ci/mmol) 来自 New England Nuclear,

Boston, MA。Optiphase Supermix 闪烁混合物来自 Wallac, Turku, Finland。白色 ViewPlate-96 孔微滴定板 (#6005182) 来自 Packard 仪器公司, Meriden, CT。所有其它试剂可从 Sigma 化学公司获得, 除非另有说明。

培养基的制备

以 DMEM 和 F12 培养基的 1:1 混合物制备基础培养基, 以加入 50 倍浓缩的贮液补充 B27 培养基。以大约 2mM 的终浓度加入 L-谷氨酰胺, 加入青霉素至大约 100 IU/l, 链霉素至大约 100mg/l。加入热灭活的马血清至大约 15 % 的终浓度。混合后, 将 pH 调到大约 7.3, 培养基在 4 °C 保存。为了减小实验内变化, 用前新制备培养基。整个过程中使用塑料吸管和容器以减少蛋白质吸附。

培养基质

为了使神经元和神经突的外延最佳附着于基质上, 按下面所述用多聚-L-鸟氨酸和层粘连蛋白顺序覆盖修饰微滴度板表面(培养基质)。用在 0.1M 硼酸 (pH 8.4) 中的 0.1mg/ml 无菌多聚-L-鸟氨酸溶液在室温下完全覆盖平板表面至少 1 小时, 接着用 Super-Q 水无菌洗涤。然后吸出水洗液, 加入在 PBS 中的 1.0µg/ml 小鼠层粘连蛋白溶液并在 37 °C 下培养 2 小时。为了保证结果的可重复性, 刚好在使用平板前进行这些程序。

胚胎大鼠灰质培养物的制备

胚胎大鼠脑用作灰质多巴胺能神经元的来源。使用 15 日胚期的定时受孕 Sprague-Dawley 大鼠。每个实验最多处理 36 个胚胎(大约 3 升)。经过接触 CO₂ 杀死妊娠大鼠, 用解剖剪打开腹腔, 从子宫取出胎儿。然后解剖胚胎的大脑, 消除血液和脑膜, 使用精确的解剖界标解剖含灰质的下被盖区 (Altman 和 Bayer, 产前大鼠脑发育图谱, CRC 出版社, Boca Raton, FL. 1995)。在冰冷的 D-PBS 中收集组织, 转移进 10ml 解离培养基 (120 单位木瓜蛋白酶和 2000 单位的 DNAase 溶于 HBSS) 中并在大约 200rpm 的摇床上大约 37 °C 培养 45 分钟。然后经过用火磨光的巴斯德吸管研磨分散细胞, 筛过 20µm Nitex 网以弃去未解离的组织, 使用 IEC 临床离心管以 200xg 离心 5 分钟。所得的细胞沉淀重悬入含卵清蛋白和大约 500 单位的 DNA 酶的 HBSS 中, 顶部覆盖 4 % 的卵清蛋白溶液 (HBSS), 在 500xg 下离心大约 10 分钟。最终的沉淀重

悬于完全培养基（见上）中，调到大约 28,000 个细胞/ml，以等份试样（90 μ l）接种入预先用多聚鸟氨酸和层粘连蛋白覆盖的 6mm 孔的 96 孔微滴度平板中。迅速发生细胞的附着且涂板效率为大约 75 %。

多巴胺能神经元的免疫组织化学

使用如下略有修改的 Louis 等（药理学实验及治疗杂志，262:1274-1283, 1992；科学，259: 689-692, 1993）所述的间接免疫过氧化物酶方法鉴定灰质培养物中的多巴胺能神经元。在室温下用 D-PBS, pH7.4 中的 4 % 低聚甲醛固定培养物大约 30 分钟，接着在 D-PBS（每 6mm 孔 200 μ l）中洗涤 3 次。然后在含 1 % NP-40 的 PBS 中的 Superblock 封闭缓冲液中培养固定的培养物以增加抗体的穿透。然后应用在相同缓冲液中以大约 1:2000 稀释的第一兔抗酪氨酸羟化酶抗体并在摇床中 37 °C 培养 1 小时。用 D-PBS 洗涤 3 次后，使用大约 1:500 稀释的羊抗兔生物素标记的 IgG 检测结合的抗体；将这些第二抗体与细胞在 37 °C 下培养大约 1 小时。然后用 D-PBS 洗涤细胞 3 次。用 1:500 稀释的抗生物素蛋白 - 生物素 - 过氧化物酶复合物检测第二抗体，将细胞在 37 °C 下培养大约 45 分钟。用 D-PBS 再洗涤 3 次后，培养物在含 0.04%, 3', 3' - 二氨基联苯胺(HCl)4, 0.06 % NiCl_2 和 0.02%过氧化氢的 0.1M Tris-HCl pH7.4 的溶液中反应 5 - 20 分钟。

测定神经元的存活

如上所述固定和处理灰质培养物用于免疫染色，然后用亮视野显微镜以 200 $\times$ 放大倍数检查。在 96 孔微滴度平板的整个 6mm 孔中计数酪氨酸羟化酶染色的神经元数目。活神经元鉴定为具有规则形状的细胞体，具有主要轴状突起和一些树状突起。从计数中排除显示出退化特征的神经元，如具有不规则的，具泡的核周质或片段化的神经突（然而，大多数退化神经元从培养基质上脱离）。多巴胺能神经元细胞数表达为 TH 阳性神经元/6-mm 孔或相对于对照多巴胺能神经元密度的倍性改变。

测定多巴胺吸收

在白色 View Plate-96 孔微滴度平板上形成的 15 日龄胚胎大鼠灰质神经元培养物中测定多巴胺吸收。用由含大约 120mM NaCl, 4.7mM KCl, 1.8mM CaCl_2 , 1.2mM MgSO_4 , 32mM NaHPO_4 , 1.3mM EDTA, 和 5.6mM D - 葡萄糖的修饰的 Krebs-Ringer 溶液 pH7.4 组成的预温的吸收缓冲液

(大约 100 μ l) 洗涤培养物。吸收缓冲液也含 1mM 抗坏血酸和 50 μ M pargyline 以防止多巴胺的氧化。然后细胞在吸收缓冲液中 37 $^{\circ}$ C 预培养大约 10 分钟。然后以在 75 μ l 吸收缓冲液中大约 50nM 的浓度向灰质培养物中加入氘标记的多巴胺 (3 H-DA, 21Ci/mmol), 培养物在 37 $^{\circ}$ C 培养大约 60 分钟。经过用含多巴胺吸收抑制剂 GBR-12909 (1 μ M) 的吸收缓冲液培养培养物测定非特异性多巴胺吸收。非特异性的吸收代表不超过总吸收的大约 1%。经过吸出培养基接着用冰冻的吸收缓冲液 (大约 120 μ l) 迅速洗涤 3 次停止吸收试验。然后经过加入 Optiphas Supermix 闪烁混合物 (200 μ l) 裂解细胞, 使用 Wallac MicrobetaPlus 96-孔微滴度平板计数器经闪烁谱计测定放射活性 (即, 经闪烁计数培养物中保留的氘分析多巴胺的吸收)。结果表达为 dpm/6-mm 孔或相对于对照培养物的倍数改变。

分析

多巴胺能神经元存活和形态学发育

富含多巴胺能神经元的 15 日胚 (E15) 大鼠灰质培养物用于证实截短的 GDNF 蛋白质对多巴胺能神经元存活的效力。培养物在多聚鸟氨酸和层粘连蛋白覆盖的 96-孔微滴度平板上单独或在各种浓度 (变化范围从大约 1pg/ml 到大约 10ng/ml) 的下列蛋白质: 大肠杆菌表达的成熟 hGDNF; 大肠杆菌表达的 [Pro²³-Ile¹³⁴], [Arg³²-Ile¹³⁴], [Gly³³-Ile¹³⁴] 和 [Lys³⁷-Ile¹³⁴] 截短的 GDNF 蛋白质; CHO 细胞表达的成熟人 GDNF; 和 CHO 细胞产生的 [Arg³²-Ile¹³⁴] 截短的 GDNF 蛋白质, 存在的条件下生长达 6 天。培养基由补充 15% 热灭活的马血清 (E15 培养物) 或 2.5% 热灭活的马血清, D-葡萄糖, HEPES, 胰岛素和运铁蛋白 (P6 培养物) 的 DMEM/F12 组成。对酪氨酸羟化酶 (TH), 多巴胺生物合成中的限速酶的免疫染色用作多巴胺能神经元的标记。由于菱脑中的去甲肾上腺素能神经元也对 TH 染色呈阳性, 所以应极小心地解剖局限于中脑下被盖的区域以避免含去甲肾上腺素能细胞体的更向后的区域。6 天后, E15 培养物典型地由以神经元特异性烯醇化酶免疫染色 (上述) 鉴定的大约 70% 的神经元和 30% 非神经元细胞 (具有扁平的, 颜色较暗的外观) 组成; 多巴胺能神经元代表大约 10 - 15% 的神经元群体。

6 天后, 用低聚甲醛固定培养物并免疫染色鉴定培养物中多巴胺能神经元的标记, 酪氨酸氧化酶。在亮视野显微镜下计数存在于 6-mm 孔中

的所有酪氨酸羟化酶阳性神经元。在各实验条件下分析 3 到 6 个不同的孔。结果表达为在对照培养物中发现的酪氨酸羟化酶阳性神经元数的百分数。

用 1.0ng/ml GDNF, CHO 细胞表达的 GDNF 或大肠杆菌表达的 GDNF 处理的 E15 灰质培养物比未处理的对照培养物分别多含大约 38 % 和 27 % 的 TH - 免疫活性神经元, 表明两种 GDNF 种类促进多巴胺能神经元的存活。用 1.0ng/ml 截短的 GDNF 蛋白质处理的 E15 灰质培养物体外培养 6 天后培养物中 TH-阳性神经元数同样增加; CHO 细胞产生的[Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质为 42 %; 大肠杆菌表达的[Arg³²-Ile¹³⁴]和[Gly³³-Ile¹³⁴]截短的 GDNF 蛋白质分别为 26 % 和 17 %。

对照培养物与用成熟和截短的 GDNF 蛋白质处理的培养物的比较也揭示了所有 GDNF 蛋白质对多巴胺能神经元形态分化的显著影响。[Arg³²-Ile¹³⁴]和[Gly³³-Ile¹³⁴]截短的 GDNF 蛋白质的效力与它们各自的成熟 GDNF 蛋白质对应物的效力相同。在所有 GDNF 处理的培养物中 TH 免疫反应神经元比对照培养物中的 TH - 阳性神经元具有明显更复杂和广泛的神经突树状分枝以及更高层次的神经突分枝和总体上更大的细胞体大小。

多巴胺吸收

多巴胺吸收测量了高亲和力多巴胺再吸收运载蛋白位点的数目和活性并反应了多巴胺能神经元的功能分化。在含有或不含成熟 GDNF 或截短的 GDNF 蛋白质的 E15 大鼠灰质培养物中体外培养 6 天后测量多巴胺的吸收。在这些培养物中, 多巴胺吸收具有多巴胺能神经元的药理学形态特征, 即被 1.0 μ M GBR-12909 一种对多巴胺能神经元特异性的多巴胺运载蛋白抑制剂 (ID₅₀ = 20nM) 几乎完全抑制 (超过 98 %)。这表明多巴胺吸收测量不反应污染的去甲肾上腺素能神经元的存在, 该神经元能通过去甲肾上腺素吸收多巴胺, 但对 GBR-12909 抑制不敏感。CHO - 细胞表达的成熟 GDNF 和 CHO 产生的[Arg³²-Ile¹³⁴]截短的 GDNF 蛋白质的效力相同, 大约增加 65 %, ED₅₀ 为大约 20pg/ml。图 5 所示的大肠杆菌表达的[Pro²³-Lys³⁷ Δ Asn³⁷-Ile¹³⁴]截短的 GDNF 蛋白质表现出增加 65 %, ED₅₀ 为大约 40pg/ml。大肠杆菌表达的成熟蛋白质和大肠杆菌表达的[Arg³²-Ile¹³⁴], [Gly³³-Ile¹³⁴]和[Lys³⁷-Ile¹³⁴]截短的 GDNF 蛋白质对多巴胺吸收的影响相同: 大约增加 50 %, ED₅₀ 为大约 50pg/ml。

这些结果表明,截短的 GDNF 蛋白质对灰质多巴胺能神经元充当有效的促进存活和诱导分化的因子。因此,可想象它们在帕金森氏病,以情绪敏感度下降,随意和不随意肌肉运动变慢,肌肉僵硬和震颤为特征的神经学紊乱的治疗中特别有用。这些症状由位于灰质中的产多巴胺神经元的逐步退化引起。这些神经元(“多巴胺能神经元”)的退化导致称为纹状体的大脑相近区域的多巴胺减少。

序列描述

(1) 一般信息:

- (i) 申请人: Hu, Sylvia
- (ii) 发明题目: 截短的神经胶质细胞系来源的神经营养因子
- 5 (iii) 序列数: 50
- (iv) 联系地址:
- (A) 联系人: 安姆根有限公司
- (B) 街道: 德哈维兰路 1840 号
- (C) 城市: 千橡树
- 10 (D) 州: 加利福尼亚
- (E) 国家: 美国
- (F) 邮政编码: 91320
- (v) 计算机可读形式:
- (A) 媒体类型: 软盘
- 15 (B) 计算机: IBM PC 兼容机
- (C) 操作系统: PC-DOS/MS-DOS
- (D) 软件: PatentIn Release #1.0, Version #1.25
- (vi) 目前申请资料:
- (A) 申请号:
- 20 (B) 申请日:
- (C) 分类:
- (viii) 律师/代理人信息:
- (A) 姓名: Curry, Daniel R.
- (B) 登记号: 32, 727
- 25 (C) 参考/摘要号: A-357
- (ix) 电信信息:
- (A) 电话: 805-447-8102
- (B) 电传: 805-499-8011
- (C) 电报:
- 30

(2) SEQ ID NO:1 信息:

- (i) 序列特征:
- (A) 长度: 402 个碱基对
- (B) 类型: 核酸
- 35 (C) 链型: 单链
- (D) 拓扑学: 线性
- (ii) 分子类型: 蛋白质
- (ix) 特征:

(A)名称/关键词: CDS

(B)位置: 1 ... 402

(xi) 序列描述: SEQ ID NO: 1:

5	TCA CCA GAT AAA CAA ATG GCA GTG CTT CCT AGA AGA GAG CGG AAT CGG	48
	Ser Pro Asp Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg	
	1 5 10 15	
	CAG GCT GCA GCT GCC AAC CCA GAG AAT TCC AGA GGA AAA GGT CGG AGA	96
	Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg	
	20 25 30	
10	GGC CAG AGG GGC AAA AAC CGG GGT TGT GTC TTA ACT GCA ATA CAT TTA	144
	Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu	
	35 40 45	
	AAT GTC ACT GAC TTG GGT CTG GGC TAT GAA ACC AAG GAG GAA CTG ATT	192
	Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile	
	50 55 60	
15	TTT AGG TAC TGC AGC GGC TCT TGC GAT GCA GCT GAG ACA ACG TAC GAC	240
	Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr Asp	
	65 70 75 80	
	AAA ATA TTG AAA AAC TTA TCC AGA AAT AGA AGG CTG GTG AGT GAC AAA	288
	Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser Asp Lys	
	85 90 95	
20	GTA GGG CAG GCA TGT TGC AGA CCC ATC GCC TTT GAT GAT GAC CTG TCG	336
	Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu Ser	
	100 105 110	
	TTT TTA GAT GAT AAC CTG GTT TAC CAT ATT CTA AGA AAG CAT TCC GCT	384
	Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser Ala	
	115 120 125	
	AAA AGG TGT GGA TGT ATC	402
	Lys Arg Cys Gly Cys Ile	
	130	

25 (2) SEQ ID NO:2 信息:

(i) 序列特征:

(A) 长度: 134 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

30 (ii)分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 2:

35	Ser Pro Asp Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg	
	1 5 10 15	
	Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg	
	20 25 30	
	Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu	
	35 40 45	

[illegible]

(2) SEQ ID NO:3 信息:

(i) 序列特征:

- 15 (A) 长度: 4 个氨基酸
(B) 类型: 氨基酸
(C) 链型: 单链
(D) 拓扑学: 线性

(ii) 分子类型: 肽

20 (xi) 序列描述: SEQ ID NO: 3:

Lys Asn Arg Gly
1

(2) SEQ ID NO:4 信息:

(i) 序列特征:

- (A) 长度: 5 个氨基酸
(B) 类型: 氨基酸
(C) 链型: 单链
(D) 拓扑学: 线性

30 (ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 4:

Gly Lys Asn Arg Gly
1 5

35 (2) SEQ ID NO:5 信息:

(i) 序列特征:

(A) 长度: 6 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

5 (ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 5:

Arg Gly Lys Asn Arg Gly
1 5

10 (2) SEQ ID NO:6 信息:

(i) 序列特征:

(A) 长度: 7 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

15 (D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 6:

Gln Arg Gly Lys Asn Arg Gly
1 5

20

(2) SEQ ID NO:7 信息:

(i) 序列特征:

(A) 长度: 8 个氨基酸

(B) 类型: 氨基酸

25 (C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 7:

Gly Gln Arg Gly Lys Asn Arg Gly
1 5

30

(2) SEQ ID NO:8 信息:

(i) 序列特征:

(A) 长度: 9 个氨基酸

35 (B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 8:

5 Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5

(2) SEQ ID NO:9 信息:

(i) 序列特征:

(A) 长度: 10 个氨基酸

10 (B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 9:

15 Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5 10

(2) SEQ ID NO:10 信息:

(i) 序列特征:

20 (A) 长度: 11 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

25 (xi) 序列描述: SEQ ID NO: 10:

Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5 10

(2) SEQ ID NO:11 信息:

30 (i) 序列特征:

(A) 长度: 12 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

35 (ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 11:

Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5 10

(2) SEQ ID NO:12 信息:

5 (i) 序列特征:

(A) 长度: 13 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

10 (ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 12:

Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5 10

15 (2) SEQ ID NO:13 信息:

(i) 序列特征:

(A) 长度: 14 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

20 (D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 13:

Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5 10

25

(2) SEQ ID NO:14 信息:

(i) 序列特征:

(A) 长度: 15 个氨基酸

(B) 类型: 氨基酸

30 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 14:

Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
1 5 10 15

35

(2) SEQ ID NO:15 信息:

(i) 序列特征:

- 5 (A) 长度: 16 个氨基酸
 (B) 类型: 氨基酸
 (C) 链型: 单链
 (D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 15:

10 Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
 1 5 10 15

(2) SEQ ID NO:16 信息:

(i) 序列特征:

- 15 (A) 长度: 17 个氨基酸
 (B) 类型: 氨基酸
 (C) 链型: 单链
 (D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 16:

20 Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg
 1 5 10 15
 Gly

(2) SEQ ID NO:17 信息:

25 (i) 序列特征:

- (A) 长度: 18 个氨基酸
 (B) 类型: 氨基酸
 (C) 链型: 单链
 (D) 拓扑学: 线性

30 (ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 17:

Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn
 1 5 10 15
 Arg Gly

35

(2) SEQ ID NO:18 信息:

(i) 序列特征:

(A) 长度: 19 个氨基酸

(B) 类型: 氨基酸

5 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 18:

10 Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys
 1 5 10 15
 Asn Arg Gly

(2) SEQ ID NO:19 信息:

15 (i) 序列特征:

(A) 长度: 20 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

20 (ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 19:

Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly
 1 5 10 15
 Lys Asn Arg Gly
 20

25

(2) SEQ ID NO:20 信息:

(i) 序列特征:

(A) 长度: 21 个氨基酸

30 (B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 20:

35

Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg
1 5 10 15

Gly Lys Asn Arg Gly
20

5

(2) SEQ ID NO:21 信息:

(i) 序列特征:

(A) 长度: 22 个氨基酸

(B) 类型: 氨基酸

10 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 21:

15 Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln
1 5 10 15

Arg Gly Lys Asn Arg Gly
20

(2) SEQ ID NO:22 信息:

20 (i) 序列特征:

(A) 长度: 23 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

25 (ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 22:

Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly
1 5 10 15

Gln Arg Gly Lys Asn Arg Gly
20

30

(2) SEQ ID NO:23 信息:

(i) 序列特征:

(A) 长度: 24 个氨基酸

35 (B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 23 :

```

5      Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg
      1              5              10              15
      Gly Gln Arg Gly Lys Asn Arg Gly
                20

```

(2) SEQ ID NO:24 信息:

10 (i) 序列特征:

(A) 长度: 25 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

15 (ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 24 :

```

      Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg
      1              5              10              15
      Arg Gly Gln Arg Gly Lys Asn Arg Gly
                20              25
20

```

(2) SEQ ID NO:25 信息:

(i) 序列特征:

(A) 长度: 26 个氨基酸

25 (B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 25 :

```

30      Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly
      1              5              10              15
      Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
                20              25

```

35 (2) SEQ ID NO:26 信息:

(i) 序列特征:

(A) 长度: 27 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

5 (ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 26:

```

Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys
1          5          10          15
Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
20          25

```

10

(2) SEQ ID NO:27 信息:

(i) 序列特征:

(A) 长度: 28 个氨基酸

15 (B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 27:

```

Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly
1          5          10          15
Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
20          25

```

20

25 (2) SEQ ID NO:28 信息:

(i) 序列特征:

(A) 长度: 29 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

30 (D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO:28:

```

Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg
1          5          10          15
Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
20          25

```

35

(2) SEQ ID NO:29 信息:

(i) 序列特征:

- (A) 长度: 30 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 29 :

Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser
1 5 10 15
Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
20 25 30

(2) SEQ ID NO:30 信息:

(i) 序列特征:

- (A) 长度: 31 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 30 :

Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu Asn
1 5 10 15
Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
20 25 30

(2) SEQ ID NO:31 信息:

(i) 序列特征:

- (A) 长度: 32 个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑学: 线性

(ii)分子类型: 肽

(xi) 序列描述: SEQ ID NO: 31 :

```

Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro Glu
1          5          10          15
Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg Gly
20          25          30

```

5

(2) SEQ ID NO:32 信息:

(i) 序列特征:

(A) 长度: 33 个氨基酸

(B) 类型: 氨基酸

10 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO:32:

```

Val Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn Pro
1          5          10          15
Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn Arg
20          25          30
Gly

```

20 (2) SEQ ID NO:33 信息:

(i) 序列特征:

(A) 长度: 34 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

25 (D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 33:

```

Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala Asn
1          5          10          15
Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys Asn
20          25          30
Arg Gly

```

(2) SEQ ID NO:34 信息:

35 (i) 序列特征:

(A) 长度: 35 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

5 (xi) 序列描述: SEQ ID NO: 34:

Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Ala
1 5 10 15

Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys
20 25 30

10 Asn Arg Gly
35

(2) SEQ ID NO:35 信息:

(i) 序列特征:

15 (A) 长度: 36 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

20 (xi) 序列描述: SEQ ID NO: 35:

Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala
1 5 10 15

Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly
20 25 30

25 Lys Asn Arg Gly
35

(2) SEQ ID NO:36 信息:

(i) 序列特征:

30 (A) 长度: 37 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 肽

35 (xi) 序列描述: SEQ ID NO: 36:

Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala
 1 5 10 15
 Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg
 20 25 30
 5 Gly Lys Asn Arg Gly
 35

(2) SEQ ID NO:37 信息:

(i) 序列特征:

- (A) 长度: 38 个氨基酸
 (B) 类型: 氨基酸
 (C) 链型: 单链
 (D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 37:

Asp Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala
 1 5 10 15
 Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln
 20 25 30
 Arg Gly Lys Asn Arg Gly
 35

(2) SEQ ID NO:38 信息:

(i) 序列特征:

- (A) 长度: 39 个氨基酸
 (B) 类型: 氨基酸
 (C) 链型: 单链
 (D) 拓扑学: 线性

(ii) 分子类型: 肽

(xi) 序列描述: SEQ ID NO: 38:

Pro Asp Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln
 1 5 10 15
 Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly
 20 25 30
 Gln Arg Gly Lys Asn Arg Gly
 35

(2) SEQ ID NO:39 信息:

(i) 序列特征:

(A) 长度: 417 个碱基对

(B) 类型: 核酸

5 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: DNA

(xi) 序列描述: SEQ ID NO: 39:

10 CATATGTCTC CGGATAAACA AATGGCTGTT CTTCCACGTC GTGAACGTAA CCGTCAGGCG 60
 GCCGCTGCTA ACCCGGAGAA TTCCCGTGGT AAAGGTCGTC GTGGTCAGCG TGGTAAAAAC 120
 CGCGGTTGCG TTCTGACCGC TATCCACCTG AACGTTACCG ACCTGGGTCT CGGTTACGAA 180
 ACCAAAGAAG AATTAATCTT CCGTTACTGC TCCGGTTCCT GCGACGCTGC TGAAACCACG 240
 TACGACAAAA TCCTGAAAAA CCTGTCCCGT AACCGTCGTC TGGTTTCCGA CAAAGTTGGT 300
 15 CAAGCTTGCT GCCGTCCGAT CGCTTTCGAC GACGACCTGT CCTTCCTGGA CGACAACCTG 360
 GTTTACCACA TCCTGCGTAA ACACTCCGCT AAGCGTTGCG GTTGCATCTA AGGATCC 417

20 (2) SEQ ID NO:40 信息:

(i) 序列特征:

(A) 长度: 417 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

25 (D) 拓扑学: 线性

(ii) 分子类型: DNA

(xi) 序列描述: SEQ ID NO: 40:

30 CATATGAGCC CGGACAAACA GATGGCAGTA CTTCCACGTC GTGAACGTAA TCGCCAGGCA 60
 GCAGCTGCAA ACCCGGAAAA CTCCCGTGGT AAAGGTCGCC GTGGCCAGCG CGGCAAAAAC 120
 CGTGGTTGTG TTCTGACTGC AATCCACCTG AACGTTACTG ACCTGGGTCT GGGCTACGAA 180
 ACCAAAGAAG AACTGATCTT CCGCTACTGC AGCGGCTCTT GCGACGCAGC TGAAACCACT 240
 TACGACAAAA TCCTGAAAAA CCTGTCCCGT AACCGCCGTC TGGTAAGCGA CAAAGTAGGT 300
 CAGGCATGCT GCCGTCCGAT CGCATTCGAC GATGACCTGA GCTTCCTGGA TGACAACCTG 360
 35 GTTTACCACA TCCTGCGTAA ACACTCCGCT AAACGCTGCG GTTGCATCTA AGGATCC 417

(2) SEQ ID NO:41 信息:

(i) 序列特征:

(A) 长度: 345 个碱基对

(B) 类型: 核酸

5 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(ix) 特征:

(A) 名称/关键词: CDS

10 (B) 位置: 1 ... 342

(xi) 序列描述: SEQ ID NO: 41:

	ATG TCC CCA GAA AAT TCT CGT GGT AAA GGT CGT CGT GGT CAG CGT GGT	48
	Met Ser Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly	
	135 140 145 150	
15	AAT AAC CGC GGT TGC GTT CTG ACC GCT ATC CAC CTG AAC GTT ACC GAC	96
	Asn Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu Asn Val Thr Asp	
	155 160 165	
	CTG GGT CTC GGT TAC GAA ACC AAA GAA GAA TTA ATC TTC CGT TAC TGC	144
	Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile Phe Arg Tyr Cys	
	170 175 180	
20	TCC GGT TCC TGC GAC GCT GCT GAA ACC ACG TAC GAC AAA ATC CTG AAA	192
	Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr Asp Lys Ile Leu Lys	
	185 190 195	
	AAC CTG TCC CGT AAC CGT CGT CTG GTT TCC GAC AAA GTT GGT CAA GCT	240
	Asn Leu Ser Arg Asn Arg Arg Leu Val Ser Asp Lys Val Gly Gln Ala	
	200 205 210	
25	TGC TGC CGT CCG ATC GCT TTC GAC GAC GAC CTG TCC TTC CTG GAC GAC	288
	Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu Ser Phe Leu Asp Asp	
	215 220 225 230	
	AAC CTG GTT TAC CAC ATC CTG CGT AAA CAC TCC GCT AAG CGT TGC GGT	336
	Asn Leu Val Tyr His Ile Leu Arg Lys His Ser Ala Lys Arg Cys Gly	
	235 240 245	
	TGC ATC TAA	345
	Cys Ile	

30

(2) SEQ ID NO:42 信息:

(i) 序列特征:

(A) 长度: 114 个氨基酸

35 (B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii)分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 42:

```

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

```

(2) SEQ ID NO:43 信息:

(i) 序列特征:

(A) 长度: 315 个碱基对

(B) 类型: 核酸

20 (C) 链型: 单链

(D) 拓扑学: 线性

(ii)分子类型: 蛋白质

(ix)特征:

(A)名称/关键词: CDS

25 (B)位置: 1 ... 312

(xi) 序列描述: SEQ ID NO: 43:

```

ATG CGT GGT CAA CGT GGT AAA AAC CGC GGT TGC GTT CTG ACT GCA ATC 48
Met Arg Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile
115           120           125           130
30 CAC CTG AAC GTT ACT GAC CTG GGT CTG GGC TAC GAA ACC AAA GAA GAA 96
His Leu Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu
    135           140           145
CTG ATC TTC CGC TAC TGC AGC GGC TCT TGC GAC GCA GCT GAA ACC ACT 144
Leu Ile Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr
    150           155           160
35 TAC GAC AAA ATC CTG AAA AAC CTG TCC CGT AAC CGC CGT CTG GTA AGC 192
Tyr Asp Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser
    165           170           175

```

```

GAC AAA GTA GGT CAG GCA TGC TGC CGT CCG ATC GCA TTC GAC GAT GAC 240
Asp Lys Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp
180 185 190

CTG AGC TTC CTG GAT GAC AAC CTG GTT TAC CAC ATC CTG CGT AAA CAC 288
Leu Ser Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His
195 200 205 210

TCC GCT AAA CGC TGC GGT TGC ATC TAA 315
Ser Ala Lys Arg Cys Gly Cys Ile
215

```

(2) SEQ ID NO:44 信息:

10 (i) 序列特征:

(A) 长度: 104 个氨基酸

(B) 类型: 氨基酸

(D) 拓扑学: 线性

(ii)分子类型: 蛋白质

15 (xi) 序列描述: SEQ ID NO: 44 :

```

Met Arg Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile
1 5 10 15

His Leu Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu
20 25 30

Leu Ile Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr
35 40 45

Tyr Asp Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser
50 55 60

Asp Lys Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp
65 70 75 80

Leu Ser Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His
85 90 95

Ser Ala Lys Arg Cys Gly Cys Ile
100

```

30 (2) SEQ ID NO:45 信息:

(i) 序列特征:

(A) 长度: 312 个碱基对

(B) 类型: 核酸

(C) 链型: 单链

35 (D) 拓扑学: 线性

(ii)分子类型: 蛋白质

(ix)特征:

(A)名称/关键词: CDS

(B)位置: 1 ... 309

(xi) 序列描述: SEQ ID NO: 45:

```

5      ATG GGT CAA CGT GGT AAA AAC CGT GGT TGT GTT CTG ACT GCA ATC CAC      48
      Met Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His
      105                      110                      115                      120

      CTG AAC GTT ACT GAC CTG GGT CTG GGC TAC GAA ACC AAA GAA GAA CTG      96
      Leu Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu
                      125                      130                      135

10     ATC TTC CGC TAC TGC AGC GGC TCT TGC GAC GCA GCT GAA ACC ACT TAC      144
      Ile Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr
                      140                      145                      150

      GAC AAA ATC CTG AAA AAC CTG TCC CGT AAC CGC CGT CTG GTA AGC GAC      192
      Asp Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser Asp
                      155                      160                      165

15     AAA GTA GGT CAG GCA TGC TGC CGT CCG ATC GCA TTC GAC GAT GAC CTG      240
      Lys Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu
      170                      175                      180

      AGC TTC CTG GAT GAC AAC CTG GTT TAC CAC ATC CTG CGT AAA CAC TCC      288
      Ser Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser
      185                      190                      195                      200

      GCT AAA CGC TGC GGT TGC ATC TAA
      Ala Lys Arg Cys Gly Cys Ile
                      205
20

```

(2) SEQ ID NO:46 信息:

(i) 序列特征:

(A) 长度: 103 个氨基酸

(B) 类型: 氨基酸

25 (D) 拓扑学: 线性

(ii)分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 46:

```

      Met Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His
      1                      5                      10                      15

30     Leu Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu
      20                      25                      30

      Ile Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr
      35                      40                      45

      Asp Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser Asp
      50                      55                      60

35     Lys Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu
      65                      70                      75                      80

```

Ser Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser
 85 90 95
 Ala Lys Arg Cys Gly Cys Ile
 100

5 (2) SEQ ID NO:47 信息:

(i) 序列特征:

(A) 长度: 135 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

10 (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 47:

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

30 (2) SEQ ID NO:48 信息:

(i) 序列特征:

(A) 长度: 104 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

35 (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 48:

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

15 (2) SEQ ID NO:49 信息:

(i) 序列特征:

(A) 长度: 103 个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

20 (D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 49:

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

35

(2) SEQ ID NO: 50 信息:

(i) 序列特征:

(A) 长度: 114 个氨基酸

(B) 类型: 氨基酸

5 (C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(xi) 序列描述: SEQ ID NO: 50:

```

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

```

成熟人 GDNF

TCA	CCA	GAT	AAA	CAA	ATG	GCA	GTG	CTT	CCT	AGA	AGA	GAG	CGG	AAT	
Ser	Pro	Asp	Lys	Gln	Met	Ala	Val	Leu	Pro	Arg	Arg	Glu	Arg	Asn	
				5					10					15	
CGG	CAG	GCT	GCA	GCT	GCC	AAC	CCA	GAG	AAT	TCC	AGA	GGA	AAA	GGT	
Arg	Gln	Ala	Ala	Ala	Ala	Asn	Pro	Glu	Asn	Ser	Arg	Gly	Lys	Gly	
				20					25					30	
CGG	AGA	GGC	CAG	AGG	GGC	AAA	AAC	CGG	GGT	TGT	GTC	TTA	ACT	GCA	
Arg	Arg	Gly	Gln	Arg	Gly	Lys	Asn	Arg	Gly	Cys	Val	Leu	Thr	Ala	
				35					40					45	
ATA	CAT	TTA	AAT	GTC	ACT	GAC	TTG	GGT	CTG	GGC	TAT	GAA	ACC	AAG	
Ile	His	Leu	Asn	Val	Thr	Asp	Leu	Gly	Leu	Gly	Tyr	Glu	Thr	Lys	
				50					55					60	
GAG	GAA	CTG	ATT	TTT	AGG	TAC	TGC	AGC	GGC	TCT	TGC	GAT	GCA	GCT	
Glu	Glu	Leu	Ile	Phe	Arg	Tyr	Cys	Ser	Gly	Ser	Cys	Asp	Ala	Ala	
				65					70					75	
GAG	ACA	ACG	TAC	GAC	AAA	ATA	TTG	AAA	AAC	TTA	TCC	AGA	AAT	AGA	
Glu	Thr	Thr	Tyr	Asp	Lys	Ile	Leu	Lys	Asn	Leu	Ser	Arg	Asn	Arg	
				80					85					90	
AGG	CTG	GTG	AGT	GAC	AAA	GTA	GGG	CAG	GCA	TGT	TGC	AGA	CCC	ATC	
Arg	Leu	Val	Ser	Asp	Lys	Val	Gly	Gln	Ala	Cys	Cys	Arg	Pro	Ile	
				95					100					105	
GCC	TTT	GAT	GAT	GAC	CTG	TCG	TTT	TTA	GAT	GAT	AAC	CTG	GTT	TAC	
Ala	Phe	Asp	Asp	Asp	Leu	Ser	Phe	Leu	Asp	Asp	Asn	Leu	Val	Tyr	
				110					115					120	
CAT	ATT	CTA	AGA	AAG	CAT	TCC	GCT	AAA	AGG	TGT	GGA	TGT	ATC		
His	Ile	Leu	Arg	Lys	His	Ser	Ala	Lys	Arg	Cys	Gly	Cys	Ile		
				125					130						

图 1

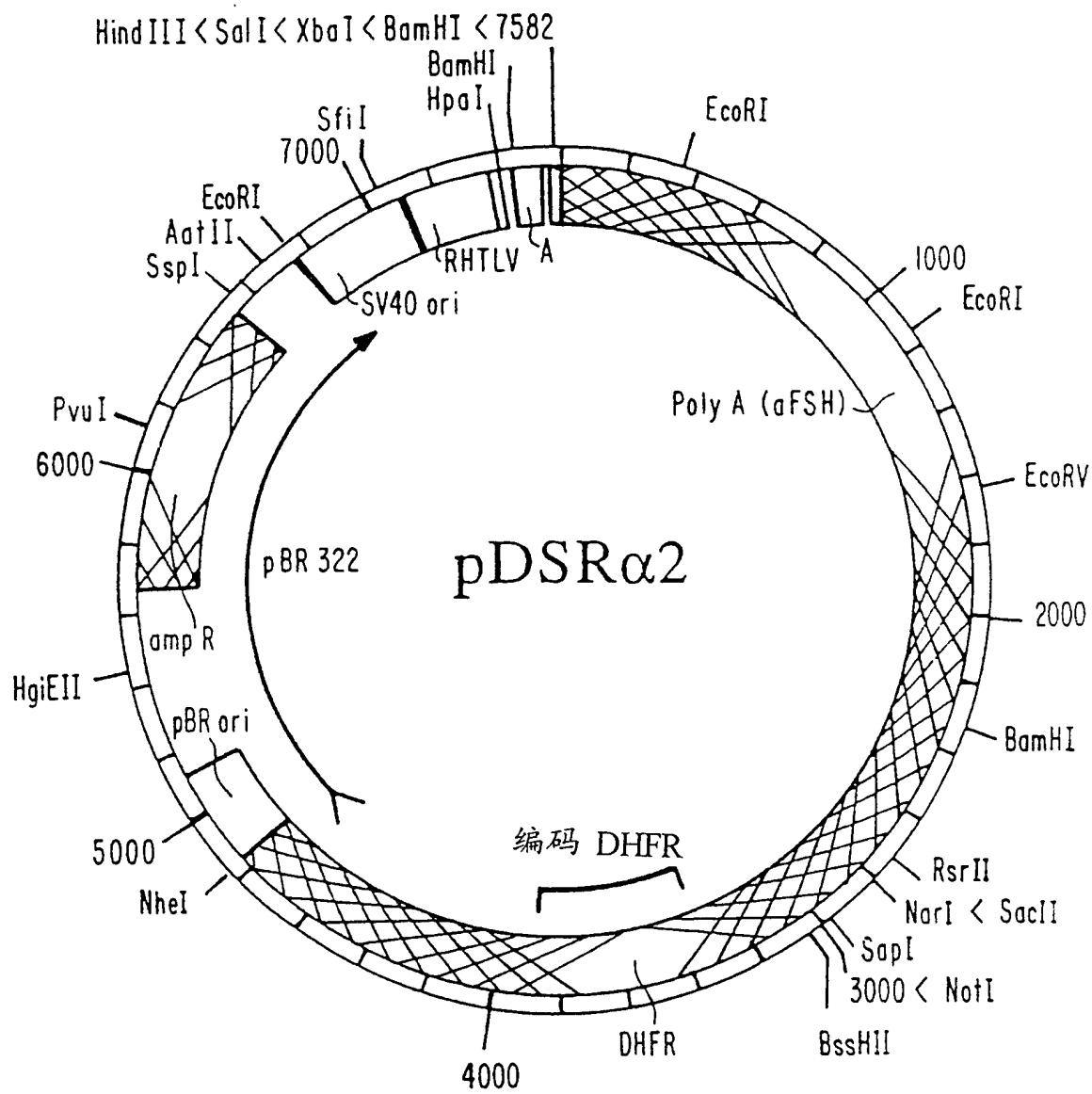


图 2

metGDNF 简并 DNA 序列

```

                                B
                                S
                                P
                                E
                                I
                                I
                                CATATGTCTCCGGATAAACAAATGGCTGTTCTTCCAC
1  -----+-----+-----+-----+-----+-----+-----+ 60
                                MetSer

                                N
                                o
                                t
                                I
                                GTCGTGAACGTAACCGTCAGGCGGCCGCTGCTAACCCGGAGAATTCCCGTGGTAAAGGTC
61 -----+-----+-----+-----+-----+-----+-----+ 120

                                S
                                a
                                c
                                I
                                I
                                GTCGTGGTCAGCGTGGTAAAAACCGCGGTTGCGTTCTGACCGCTATCCACCTGAACGTTA
121 -----+-----+-----+-----+-----+-----+-----+ 180

                                P
                                s
                                h
                                A
                                I
                                CCGACCTGGGTCTCGGTTACGAAACCAAAGAAGAAATTAATCTTCCGTTACTGCTCCGTT
181 -----+-----+-----+-----+-----+-----+-----+ 240

```

图 3 A

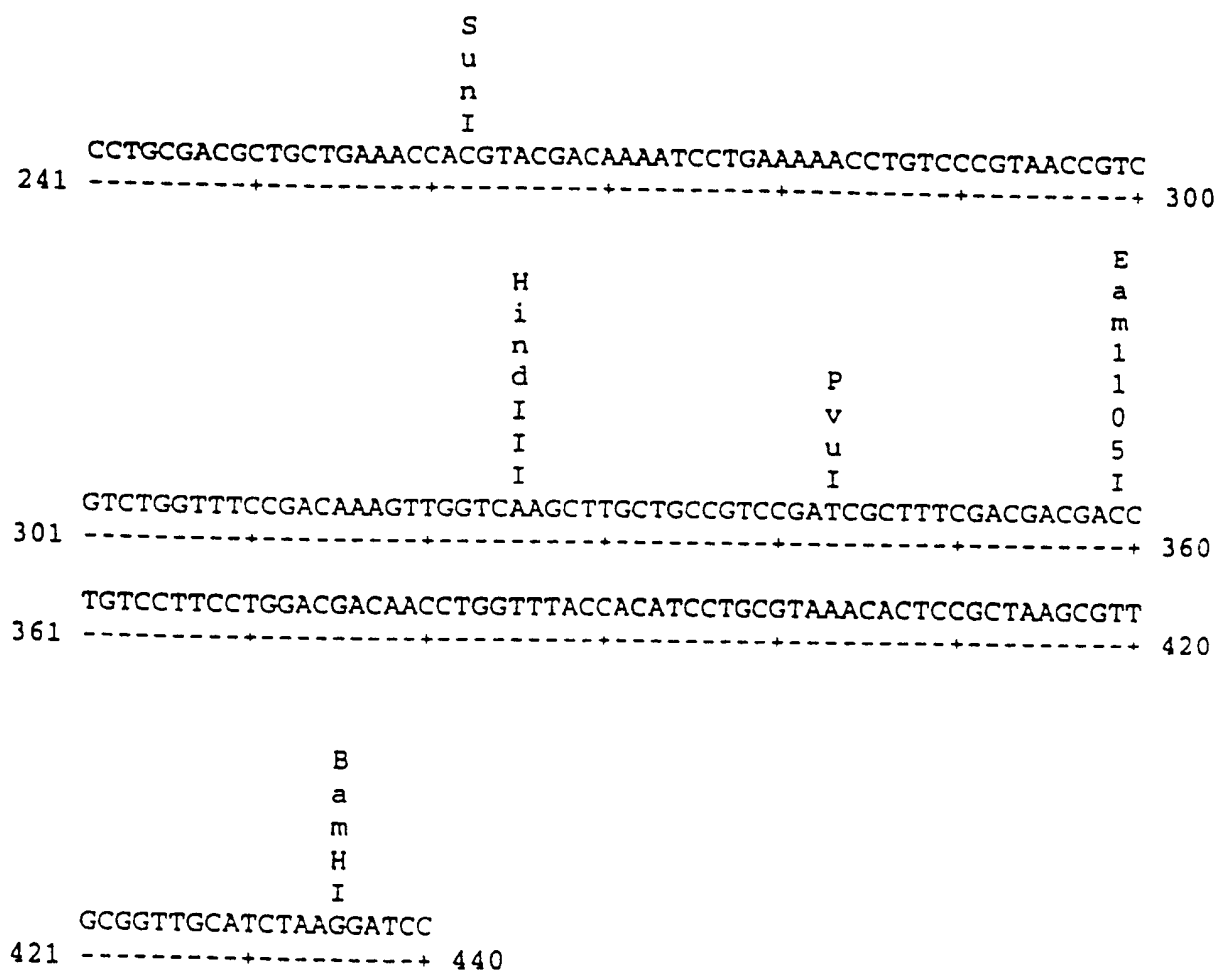


图 3 B

metGDNF 简并 DNA 序列

```

                                N
                                d
                                e
                                I
                                CATATGAGCCCGGACAAACAG
1  -----+-----+-----+-----+-----+-----+ 60
                                MetSer

ATGGCAGTACTTCCACGTCGTGAACGTAATCGCCAGGCAGCAGCTGCAAACCCGGAAAAC
61 -----+-----+-----+-----+-----+-----+ 120

TCCCGTGGTAAAGGTCGCCGTGGCCAGCGCGGCAAAAACCGTGGTTGTGTTCTGACTGCA
121 -----+-----+-----+-----+-----+-----+ 180

                                P
                                s
                                h
                                A
                                I
181 -----+-----+-----+-----+-----+-----+ 240

                                P
                                s
                                t
                                I
241 -----+-----+-----+-----+-----+-----+ 300

                                P
                                v
                                u
                                I
301 -----+-----+-----+-----+-----+-----+ 360

B
s
m
I
361 -----+-----+-----+-----+-----+-----+ 420

                                B
                                a
                                m
                                H
                                I
421 -----+-----+-----+-----+-----+-----+ 456

```

图 4

[Pro²³-Lys³⁷ΔAsn³⁷-Ile¹³⁴] 截短的 GDNF 蛋白质

```

ATGTCCCCAGAAAATTCTCGTGGTAAAGGTCGTCGTGGTCAGCGTGGTAATAACCGCGGT
21  -----+-----+-----+-----+-----+-----+ 80
   M S P E N S R G K G R R G Q R G N N R G

TGCGTTCTGACCGCTATCCACCTGAACGTTACCGACCTGGGTCTCGGTTACGAAACCAAA
81  -----+-----+-----+-----+-----+-----+ 140
   C V L T A I H L N V T D L G L G Y E T K

GAAGAATTAATCTTCCGTTACTGCTCCGGTTCCTGCGACGCTGCTGAAACCACGTACGAC
141 -----+-----+-----+-----+-----+-----+ 200
   E E L I F R Y C S G S C D A A E T T Y D

AAAATCCTGAAAAACCTGTCCCGTAACCGTCGTCTGGTTTCCGACAAAGTTGGTCAAGCT
201 -----+-----+-----+-----+-----+-----+ 260
   K I L K N L S R N R R L V S D K V G Q A

TGCTGCCGTCGATCGCTTTCGACGACGACCTGTCCTTCCTGGACGACAACCTGGTTTAC
261 -----+-----+-----+-----+-----+-----+ 320
   C C R P I A F D D D L S F L D D N L V Y

CACATCCTGCGTAAACACTCCGCTAAGCGTTGCGGTTGCATCTAA
321 -----+-----+-----+-----+-----+-----+
   H I L R K H S A K R C G C I *

```

图 5

[Arg³²-Ile¹³⁴] 截短的 GDNF 蛋白质

```

ATGCGTGGTCAACGTGGTAAAAACCGCGGTTGCGTTCTGACTGCAATCCACCTGAACGTT
41  -----+-----+-----+-----+-----+ 100
   M R G Q R G K N R G C V L T A I H L N V

ACTGACCTGGGTCTGGGCTACGAAACCAAAGAAGAACTGATCTTCCGCTACTGCAGCGGC
101 -----+-----+-----+-----+-----+ 160
   T D L G L G Y E T K E E L I F R Y C S G

TCTTGCGACGCAGCTGAAACCACTTACGACAAAATCCTGAAAAACCTGTCCCGTAACCGC
161 -----+-----+-----+-----+-----+ 220
   S C D A A E T T Y D K I L K N L S R N R

CGTCTGGTAAGCGACAAAGTAGGTCAGGCATGCTGCCGTCCGATCGCATTTCGACGATGAC
221 -----+-----+-----+-----+-----+ 280
   R L V S D K V G Q A C C R P I A F D D D

CTGAGCTTCCTGGATGACAACCTGGTTTACCACATCCTGCGTAAACACTCCGCTAAACGC
281 -----+-----+-----+-----+-----+ 340
   L S F L D D N L V Y H I L R K H S A K R

TGCGGTTGCATCTAA
341 -----+----- 355
   C G C I *

```

图 6

[Gly³³-Ile¹³⁴] 截短的 GDNF 蛋白质

ATGGGTCAACGTGGTAAAAACCGTGGTTGTGTTCTGACTGCAATCCACCTGAACGTTACT
 41 -----+----- 100
 M G Q R G K N R G C V L T A I H L N V T

 GACCTGGGTCTGGGCTACGAAACCAAAGAAGAACTGATCTTCCGCTACTGCAGCGGCTCT
 101 -----+-----+-----+-----+-----+-----+ 160
 D L G L G Y E T K E E L I F R Y C S G S

 TGCGACGCAGCTGAAACCACTTACGACAAAATCCTGAAAAACCTGTCCCGTAACCGCCGT
 161 -----+-----+-----+-----+-----+-----+ 220
 C D A A E T T Y D K I L K N L S R N R R

 CTGGTAAGCGACAAAGTAGGTCAGGCATGCTGCCGTCCGATCGCATTTCGACGATGACCTG
 221 -----+-----+-----+-----+-----+-----+ 280
 L V S D K V G Q A C C R P I A F D D D L

 AGCTTCCTGGATGACAACCTGGTTTACCACATCCTGCGTAAACACTCCGCTAAACGCTGC
 281 -----+-----+-----+-----+-----+-----+ 340
 S F L D D N L V Y H I L R K H S A K R C

 GGTTCATCTAA
 341 -----+--- 352
 G C I *

图 7

蛋白质序列比较

					50
	GDNF	MSPDKQMAVL	PRRERNRQAA	AANPENSAGK	GRRGQRGKNR GCVLTAIHLN
-31	GDNF				.MRGQRGKNR GCVLTAIHLN
-32	GDNF				..MGQRGKNR GCVLTAIHLN
-22	GDNF			.MSPENSAGK	GRRGQRGNR GCVLTAIHLN
		51			100
	GDNF	VTDLGLGYET	KEELIFRYCS	GSCDAAETTY	DKILKNLSRN RRLVSDKVGQ
-31	GDNF	VTDLGLGYET	KEELIFRYCS	GSCDAAETTY	DKILKNLSRN RRLVSDKVGQ
-32	GDNF	VTDLGLGYET	KEELIFRYCS	GSCDAAETTY	DKILKNLSRN RRLVSDKVGQ
-22	GDNF	VTDLGLGYET	KEELIFRYCS	GSCDAAETTY	DKILKNLSRN RRLVSDKVGQ
		101		135	
	GDNF	ACCRPIAFDD	DLSFLDDNLV	YHILRKHS AK	RCGCI
-31	GDNF	ACCRPIAFDD	DLSFLDDNLV	YHILRKHS AK	RCGCI
-32	GDNF	ACCRPIAFDD	DLSFLDDNLV	YHILRKHS AK	RCGCI
-22	GDNF	ACCRPIAFDD	DLSFLDDNLV	YHILRKHS AK	RCGCI

图 8