

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

A61K 47/48

C07K 14/535 C12N 15/27



[12] 发明专利说明书

[21] ZL 专利号 01803590.6

[45] 授权公告日 2005 年 2 月 9 日

[11] 授权公告号 CN 1188172C

[22] 申请日 2001.1.9 [21] 申请号 01803590.6

[30] 优先权

[32] 2000. 1. 10 [33] DK [31] PA200000024

[32] 2000. 3. 2 [33] DK [31] PA200000341

[32] 2000. 6. 16 [33] DK [31] PA200000943

[86] 国际申请 PCT/DK2001/000011 2001. 1. 9

[87] 国际公布 WO2001/051510 英 2001. 7. 19

[85] 进入国家阶段日期 2002. 7. 10

[71] 专利权人 马克西根控股公司

地址 美国加利福尼亚州

[72] 发明人 托本·L·尼森 金·V·安德森

克里斯琴·K·汉森

简·M·米凯尔森

汉斯·T·沙姆拜伊

审查员 樊婵娟

[74] 专利代理机构 北京市柳沈律师事务所

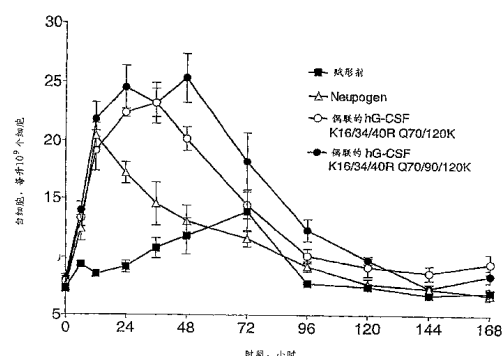
代理人 巫肖南 黄益芬

权利要求书 3 页 说明书 69 页 附图 3 页

[54] 发明名称 G-CSF 偶联物

[57] 摘要

本发明涉及多肽偶联物，所述偶联物含有显示 G-CSF 活性的多肽且该多肽的氨基酸序列不同于人类 G-CSF 序列之处在于至少引入和/或去除一个特定的氨基酸残基，所述残基含有一个能附着非多肽组分的基团，并且该多肽有至少一个偶联到该多肽的附着基团上的非多肽组分。附着基团可以是例如赖氨酸，半胱氨酸，天冬氨酸或谷氨酸残基或糖基化位点，非多肽组分可以是例如聚合物例如聚乙二醇或多糖。偶联物有一个或多个改善的特性例如生物半寿期增加且副作用降低。



1. 一种多肽偶联物, 该偶联物包含:

i) 一种显示 G-CSF 活性的多肽, 该多肽包含与 SEQ ID NO:1 所示氨基酸序列不同的氨基酸序列, 该序列的不同之处在于至少一个选自下组的取代: T1K, P2K, L3K, G4K, P5K, A6K, S7K, S8K, L9K, P10K, Q11K, S12K, F13K, L14K, L15K, E19K, Q20K, V21K, Q25K, G26K, D27K, A29K, A30K, E33K, A37K, T38K, Y39K, L41K, H43K, P44K, E45K, E46K, V48K, L49K, L50K, H52K, S53K, L54K, I56K, P57K, P60K, L61K, S62K, S63K, P65K, S66K, Q67K, A68K, L69K, Q70K, L71K, A72K, G73K, S76K, Q77K, L78K, S80K, F83K, Q86K, G87K, Q90K, E93K, G94K, S96K, P97K, E98K, L99K, G100K, P101K, T102K, D104K, T105K, Q107K, L108K, D109K, A111K, D112K, F113K, T115K, T116K, W118K, Q119K, Q120K, M121K, E122K, E123K, L124K, M126K, A127K, P128K, A129K, L130K, Q131K, P132K, T133K, Q134K, G135K, A136K, M137K, P138K, A139K, A141K, S142K, A143K, F144K, Q145K, S155K, H156K, Q158K, S159K, L161K, E162K, V163K, S164K, Y165K, V167K, L168K, H170K, L171K, A172K, Q173K 和 P174K, 和

ii) 至少一个附着至该多肽的赖氨酸残基上的非多肽组分。

2. 权利要求 1 的偶联物, 其中 SEQ ID NO:1 的氨基酸残基 K16, K23, K34 和 K40 中的至少一个已缺失或用另一个氨基酸残基取代。

3. 权利要求 2 的偶联物, 其中 K16, K23, K34 和 K40 中的至少一个已被 R 或 Q 残基取代。

4. 权利要求 1-3 中任一项的偶联物, 其中所述多肽包含至少一个选自下组的取代: P5K, A6K, S7K, S8K, L9K, P44K, E45K, E46K, V48K, L49K, L50K, H52K, S53K, L54K, I56K, P57K, P60K, L61K, S62K, S63K, P65K, S66K, Q67K, A68K, L69K, Q70K, L71K, A72K, G73K, S76K, Q77K, L78K, S80K, F83K, Q86K, G87K, Q90K, E93K, G94K, S96K, P97K, E98K, L99K, P101K, D104K, T105K, Q107K, L108K, D109K, A111K, D112K, F113K, T115K, T116K, W118K, Q119K, Q120K, M121K, E122K, E123K, L124K, M126K, A127K, P128K, A129K, L130K, Q131K, P132K, T133K, Q134K,

G135K, A136K, M137K, P138K, A139K, A141K, S142K, A143K, F144K, Q145K, S155K, H156K, Q158K, S159K, L161K, E162K, V163K, S164K, Y165K, V167K, L168K, H170K, L171K, A172K, Q173K 和 P174K。

5. 权利要求4的偶联物,其中所述多肽包含至少一个选自 E46K, S66K, Q70K, Q90K, D104K, T105K, Q120K, T133K, S142K, S155K, S159K 和 H170K 的取代。

6. 权利要求4的偶联物,其中所述多肽包含至少一个选自 Q70K, Q90K, T105K, Q120K 和 T133K 的取代。

7. 权利要求1-6中任一项的偶联物,其中所述非多肽组分是聚环氧烷。

10 8. 权利要求7的偶联物,其中所述聚环氧烷选自聚乙二醇和线性或分支的聚乙二醇组成的组。

9. 权利要求8的偶联物,其中所述聚合物是线性或分支的聚乙二醇。

10. 权利要求9的偶联物,其中所述聚乙二醇的分子量约1000-15,000 Da。

15 11. 权利要求9或10的偶联物,其包含2-8个聚乙二醇组分。

12. 权利要求1-11中任一项的偶联物,其中所述多肽与 SEQ ID NO:1 所示氨基酸序列相比有1-15个氨基酸残基不同。

13. 权利要求12的偶联物,其中所述多肽与 SEQ ID NO:1 所示氨基酸序列相比有2-10个氨基酸残基不同。

20 14. 权利要求1-13中任一项的偶联物,其中所述多肽已糖基化。

15. 权利要求14的偶联物,其中所述多肽包含至少一个非天然存在的糖基化位点。

16. 一种显示 G-CSF 活性的多肽,其具有权利要求1-6中任一项限定的氨基酸序列。

25 17. 一种核苷酸序列,其编码权利要求16的多肽。

18. 一种表达载体,其携有权利要求17的核苷酸序列。

19. 一种宿主细胞,其包含权利要求17的核苷酸序列或权利要求18的表达载体。

30 20. 制备权利要求1-15中任一项所述多肽偶联物的方法,包括在可导致所述多肽表达的条件下依照权利要求19的宿主细胞,并收获所述多肽,其中 a)所述多肽包含至少一个 N-或 O-糖基化位点且所述宿主细胞是一种能

进行体内糖基化的真核宿主细胞，和/或 b)使所述多肽在体外与一个非多肽组分偶联。

¹ 21. 一种组合物，其包含权利要求 1-15 中任一项的偶联物和一种可药用载体或赋形剂。

- 5 22. 权利要求 1-15 中任一项的偶联物或权利要求 16 的多肽在制备用于治疗疾病的药物中的用途，所述疾病具体指广泛性造血疾病包括放射治疗或化疗所致的疾病，白细胞减少症，AIDS 或其它免疫缺陷病，以及细菌感染或其它感染。

G-CSF 偶联物

5

发明领域

本发明涉及能显示粒细胞集落刺激因子(G-CSF)活性的新的多肽,显示 G-CSF 活性的多肽与非多肽组分(moiety)的偶联物,制备这种多肽或偶联物的方法以及所述多肽或偶联物在治疗,尤其对白细胞减少症的治疗方面的应用。

10

发明背景

白细胞在骨髓中的生长,分裂和分化的过程称为造血(Dexter and Spooncer, Ann. Rev. Cell. Biol., 3: 423, 1987)。每种类型的血细胞都来自多能干细胞。体内通常能产生三种类型的血细胞:红血细胞(红细胞),血小板和白血细胞(白细胞),后者的大多数参与宿主的免疫防御。造血前体细胞的增殖和分化由一个细胞因子家族调节,这一家族的细胞因子包括集落刺激因子(CSF's)例如 G-CSF 和白细胞介素(Arai et al., Ann.Rev. Biochem., 59: 783-836, 1990)。G-CSF 在体内的主要生物学作用是刺激某些称为嗜中性粒细胞的白细胞的生长和发育(Welte et al., PNAS-USA 82: 1526-1530, 1985, Souza et al., science, 232: 61-65, 1986)。当嗜中性粒细胞被释放到血流时,这些细胞便行使抗御细菌感染的功能。

Nagata et al., nature 319: 415-418, 1986 曾报道过人 G-CSF (hG-CSF)的氨基酸序列。hG-CSF 是一种单体蛋白,它通过形成 2 个 G-CSF 分子和 2 个受体的 2:2 复合物与 G-CSF 受体二聚化(Horan et al., (1996), Biochemistry 35 (15): 4886-96)。Aritomi et al., nature 401: 713-717, 1999 描述了 hG-CSF 和 G-CSF 受体的 BN-BC 功能域之间形成的复合物的 X 线结构。它们确定以下 hG-CSF 的残基为受体结合界面的一部分: G4, P5, A6, S7, S8, L9, P10, Q11, S12, L15, K16, E19, Q20, L108, D109, D112, T115, T116, Q119, E122, E123, 和 L124。rhG-CSF 在大肠杆菌,酿酒酵母和哺乳动物细胞中的表达已有报道(Souza et al., Science 232:61-65, 1986, Nagata et al., Nature 319:415-418, 1986, Robinson and Wittrup, Biotechnol. Prog. 11:171-177, 1985)。

重组人 G-CSF (rhG-CSF)通常用于治疗各种形式的白细胞减少症。因此,

可得到冠以 filgrastim (Gran® 和 Neupogen®), lenograstim (Neutrogin® 和 Granocyte®)和 nartograstim (Neu-up®)名称的 rhG-CSF 商业制品。Gran® 和 Neupogen®是非糖基化的且可以在重组大肠杆菌细胞中制备。Neutrogin® 和 Granocyte®是糖基化的且可在重组 CHO 细胞中制备, Neu-up®是非糖基化的, 其在重组大肠杆菌细胞产生的完整 rhG-CSF 的 N 末端区域有 5 个氨基酸被取代。

已报道了 hG-CSF 的多个蛋白工程变体(US5, 581, 476, US5, 214, 132, US5, 362, 853, US4, 904, 584 和 Riedhaar-Olson et al., Biochemistry 35: 9034-9041, 1996)。为引入比天然多肽至少多一个的糖链, 已有人建议对 hG-CSF 和其它多肽进行修饰(US 5, 218, 092)。所述多肽的氨基酸序列可以通过氨基酸取代, 氨基酸缺失或氨基酸插入进行修饰以添加另外的糖链。除此之外, 对天然 hG-CSF 的聚合物修饰, 包括附着 PEG 基团, 已有报道 (Satake-Ishikawa Cell Structure and Function 17: 157-160, 1992, US 5, 824, 778, US 5, 824, 784, WO 96/11953, WO 95/21629, WO 94/20069)。

Bowen et al., Experimental Hematology 27 (1999), 425-432 公开了 PEG 偶联的 G-CSF 突变蛋白的分子质量及其活性维持时间之间的关系。提示所述蛋白上 PEG 偶联部分的分子量与体外活性之间有明显反比关系, 然而体内活性随分子量的增加而增加。推测偶联物的较低亲和力有增加半寿期的作用, 因为受体介导的内吞作用是一种调节造血生长因子水平的重要机制。

市售 rhG-CSF 有短期的药理学作用, 且在白细胞减少状态的持续期间通常必须一天给药一次以上。具有较长循环半寿期的分子会减少缓解白细胞减少症和预防随后的感染所必需的给药次数。现有 rG-CSF 产品的另一问题是剂量依赖性骨痛的出现。因为骨痛是患者接受 rG-CSF 治疗时体会到的明显副作用, 所以能提供一种不引起骨痛的 rG-CSF 产品是人们所期望的, 这种产品可以本身没有这种作用, 或者是这种产品在足够小的剂量下有效但该剂量不引起骨痛。因此, 很明显有必要改进重组 G-CSF 样分子。

关于半寿期, 增加蛋白的循环半寿期的一个途径是保证蛋白的清除, 尤其是通过肾脏的清除和受体介导的清除的下降。这可以通过将蛋白与能增加表观分子量的化学组分偶联, 从而降低肾脏的清除作用, 增加上述分子的体内半寿期来实现。进一步地, 蛋白附着化学组分可以有效地阻断蛋白裂解性酶与蛋白的物理接触, 由此可以使蛋白免遭非特异性蛋白水解作用的降

解。聚乙二醇(Polyethylene glycol)(PEG)便是这样一种化学组分,它已在治疗性蛋白制品的制备中应用。一种被称为 SD/01 的 N 末端连接了单个 PEG 基团的 G-CSF 分子目前正进行临床试验。尽管这种 PEG 化 G-CSF 分子比非 PEG 化 G-CSF 半寿期增加,但它活性低,所以必须以相当高的剂量给药。

- 5 由此,上面讨论的骨痛问题并没有通过这个分子解决。此外,已知 G-CSF 分子的半寿期的改进仍有相当大的空间。

所以,仍然有必要提供一种既能显示在白细胞减少症的治疗中有用的活性,同时半寿期增加但副作用降低的新型分子。本发明涉及这样的分子。

10

发明简述

- 更具体地,本发明涉及特异性偶联物,其包含一种显示 G-CSF 活性的多肽和一种非多肽组分,本发明还涉及上述偶联物的制备方法和它们在医学治疗和药物制备方面的应用。由此,本发明第一个方面涉及各种特异性偶联物,这些偶联物包含显示 G-CSF 活性的多肽,该多肽具有与 SEQ ID NO: 1 所示人 G-CSF 的氨基酸序列不同在于至少一个特定的包含非多肽组分附着基团的氨基酸残基的氨基酸序列,且至少有一个非多肽组分附着到该多肽的附着基团上。本发明的偶联物与商业提供的 rhG-CSF 相比有一个或多个改进的特性,包括增加了体内的功能半寿期,增加了血清半寿期,降低了副作用,降低了免疫原性和/或增加了生物利用度。所以,应用本发明的偶联物进行
- 15 医学治疗可以表现出优于现有 G-CSF 化合物的许多优点。
- 20

另一方面,本发明涉及显示 G-CSF 活性且形成本发明的偶联物的一部分的多肽。预计本发明的多肽可以用于治疗,诊断或其它的目的,但也可具体作为制备本发明偶联物的中间产物。

- 本发明另一方面涉及一种多肽偶联物,该偶联物包含显示 G-CSF 活性的多肽,该多肽的氨基酸序列与 rhG-CSF 的氨基酸序列(其氨基酸序列在 SEQ ID NO: 1 中列出)不同在于至少一个氨基酸残基是引入或除去的包含非多肽组分附着基团的氨基酸残基,并且足够数量或类型的非多肽组分提供了与已知重组 G-CSF 产品相比半寿期增加的偶联物。
- 25

- 本发明的另一方面还涉及制备本发明偶联物的方法,包括编码本发明多肽的核苷酸序列,包含此核苷酸序列的表达载体,和包含上述核苷酸序列或表达载体的宿主细胞。
- 30

本发明的最后一方面涉及包含本发明偶联物或多肽的组合物,制备药物组合物的方法,本发明的偶联物或组合物作为药物的应用,和用上述组合物治疗哺乳动物的方法。尤其,本发明的多肽,偶联物或组合物可以用于使正在
5 5 进行某些类型的放射治疗,化疗,和骨髓移植的癌症患者预防感染,用于动员祖细胞在外周血祖细胞移植中聚集,用于治疗无论什么原因所致的严重慢性或相对的白细胞减少症,和用于支持对急性髓细胞性白血病(acute myeloid leukaemia)患者的治疗。另外,本发明的多肽,偶联物或组合物可以用于 AIDS 或其它免疫缺陷病以及细菌感染的治疗。

10

发明详述

定义

在本申请和发明的上下文中应用了以下定义:

术语“偶联物”是指由一个或多个多肽,通常为单独一个多肽与一个或多个非多肽组分例如聚合物分子,亲脂性化合物,糖(carbohydrate)组分或有
15 15 机衍生制剂共价连接形成的异源分子。术语共价连接是指多肽和非多肽组分直接彼此共价连接,或者是通过一个或多个间插组分,例如桥,间隔臂,或连接组分彼此间接共价连接。优选,偶联物在相应的浓度和条件下是可溶性的,即在生理性的体液中例如血液是可溶性的。术语“非偶联的多肽”可用于指偶联物的多肽部分。

20

本发明中术语“多肽”可以和术语“蛋白”交换使用。

术语“聚合物分子”是由二个或多个单体共价连接形成的分子,其中,除聚合物是人白蛋白或另一种高丰度血浆蛋白外,所有单体都不是仅含有一种氨基酸残基。术语“聚合物”可以和术语“聚合物分子”交换使用。该术语将包括糖分子,但通常情况下,该术语不包括通过体内 N-或 O-的糖基化(以下进一步描述)连接到所述多肽上的这类糖分子,因为这类分子在下文中被称为“寡糖组分”。除非对聚合物分子的数目有清楚的说明,每次提及本发明多肽中包含的“一个聚合物”,“一个聚合物分子”,“聚合物”或“聚合物分子”或在本发明中应用的上述称谓都涉及一个或多个聚合物分子。
25 25

术语“附着基团”是指多肽中能够偶联到相关非多肽组分上的氨基酸残基基团。例如,在聚合物偶联,尤其与 PEG 偶联的情况中,常用的附着基团是赖氨酸的 ϵ 氨基或 N 末端氨基。其它的聚合物附着基团包括游离羧酸基
30 30

团(例如, C 末端氨基酸残基的游离羧酸基团或天冬氨酸或谷氨酸残基的游离羧酸基团), 适当活化的羰基基团, 氧化的糖组分和巯基基团。有用的附着基团和它们相配对的非肽组分见下表。

附着基团	氨基酸	非多肽组分的实例	偶联方法/-活化的 PEG	参考文献
-NH ₂	N 末端, Lys, His, Arg	聚合物, 例如 PEG, 带酰胺基 或亚胺基	mPEG-SPA Tresylated mPEG	Shearwater Inc. Delgado et al., Critical reviews in Therapeutic Drug Carrier Systems 9 (3, 4) : 249-304 (1992)
-COOH	C 末端, Asp, Glu	聚合物, 例如 PEG, 带酯或酰胺基 寡糖组分	mPEG-Hz 体外偶联	Shearwater Inc.
-SH	Cys	聚合物, 例如 PEG, 带二硫键, 马来酰亚胺或乙 烯砜基团 寡糖组分	PEG-乙烯砜 PEG-马来酰 亚胺 体外偶联	Shearwater Inc. Delgado et al., Critical reviews in Therapeutic Drug Carrier Systems 9 (3, 4) : 249-304 (1992)
-OH	Ser, Thr, -OH, Lys	寡糖组分 具有酯, 乙醚, 氨 基甲酸酯, 碳酸盐 的 PEG	体内 O-连接 的糖基化	
-CONH ₂	Asn 作为 N 糖基化 位点的一 部分	寡糖组分 聚合物, 例如 PEG	体内 N 糖基 化	
芳香族 残基	Phe, Tyr, Trp	寡糖组分	体外偶联	
-CONH ₂	Gln	寡糖组分	体外偶联	Yan and Wold, Biochemistry, 1984, Jul 31;23(16):3759-65
醛酮	氧化的寡 糖	聚合物, 例如 PEG, PEG-胍	PEG 化	Andresz et al., 1978, Makromol. Chem. 179:301, WO92/16555, WO02/23114
胍	Arg	寡糖组分	体外偶联	Lundblad and Noyes, Chemical Reagents for Protein Modification, CRC Press Inc., Florida, USA
咪唑环	His	寡糖组分	体外偶联	与胍相同

- 对于体内的 N 糖基化, 术语“附着基团”以非常规的方式应用以表示
- 5 构成 N 糖基化位点的氨基酸残基 (带有 N-X'-S/T/C-X" 序列, 其中 X' 是除脯氨酸外的任何氨基酸残基, X" 可以是任何氨基酸残基, 其可以是脯氨酸也可

以不是脯氨酸，优选其不是脯氨酸，N 是天冬酰胺，S/T/C 可以是丝氨酸，苏氨酸或半胱氨酸，优选是丝氨酸或苏氨酸，最优选是苏氨酸)。尽管 N-糖基化位点的天冬酰胺残基是糖基化过程中寡糖部分附着的位置，但这种附着不能实现，除非存在 N 糖基化位点的其它氨基酸残基。因此，当非肽组分是寡糖组分且经过 N-糖基化实现偶联时，与目标多肽氨基酸序列的改变连用的术语“包含非肽组分附着基团的氨基酸残基”应理解为构成 N 糖基化位点的一个或多个氨基酸残基以下述方式改变，该是将功能性 N 糖基化位点引入到该氨基酸序列或从所述序列中除去。

在本申请中，氨基酸的名称和原子名称(例如，CA, CB, NZ, N, O, C, 等)依照蛋白数据库(PDB) (www.pdb.org)的定义应用，这些名称是根据 IUPAC 命名法命名的(IUPAC Nomenclature and Symbolism for Amino Acids and Peptides (residue names, atom names etc.), Eur.J.Biochem. 138, 9-37 (1984) 和在 Eur.J.Biochem.中对它们的更正 152, 1 (1985)。术语“氨基酸残基”主要是指 20 个天然氨基酸中的氨基酸，即丙氨酸(Ala 或 A)，半胱氨酸(Cys 或 C)，天冬氨酸(Asp 或 D)，谷氨酸(Glu 或 E)，苯丙氨酸(Phe 或 F)，甘氨酸(Gly 或 G)，组氨酸(His 或 H)，异亮氨酸(Ile 或 I)，赖氨酸(Lys 或 K)，亮氨酸(Leu 或 L)，蛋氨酸(Met 或 M)，天冬酰胺(Asn 或 N)，脯氨酸(Pro 或 P)，谷氨酰胺(Gln 或 Q)，精氨酸(Arg 或 R)，丝氨酸(Ser 或 S)，苏氨酸(Thr 或 T)，缬氨酸(Val 或 V)，色氨酸(Trp 或 W)和酪氨酸(Tyr 或 Y)残基。

用于定义氨基酸的位置/取代的术语举例说明如下：F13 是指参考氨基酸序列中编号 13 的位置由苯丙氨酸残基占据。F13K 是指位置 13 的苯丙氨酸残基已经被赖氨酸残基取代。除非另有说明，本文中氨基酸残基的编号与 SEQ ID NO : 1 所示 hG-CSF 的氨基酸序列相对应。可供选择的取代用"/"表示，例如，Q67D/E 是指其中第 67 位的谷氨酰胺被天冬氨酸或谷氨酸取代的氨基酸序列。多个取代用"+"表示，例如，S53N+G55S/T 是指其中第 53 位的丝氨酸残基被天冬酰胺取代且第 55 位的甘氨酸残基被丝氨酸或苏氨酸残基取代的一个氨基酸序列。

术语“核苷酸序列”是指二个或多个核苷酸分子的连续延伸。核苷酸序列可以是基因组型，cDNA 型，RNA 型，半合成或合成来源的，或它们的任何组合。

术语“聚合酶链反应”或“PCR”通常是指一种在体外扩增目的核苷酸

序列的方法，例如 US 4, 683, 195 中所述。通常，PCR 方法涉及应用能优先与模板核酸杂交的寡核苷酸引物，进行引物延伸合成的反复循环。

“细胞”，“宿主细胞”，“细胞系”和“细胞培养物”在本发明中可以交换使用且所有这些术语都应该包括细胞生长或培养后产生的后代。“转化”

5 和“转染”可以交换使用，指将 DNA 导入细胞的过程。

“可操作地连接(Operably linked)”是指二个或多个核苷酸序列通过酶连接或其它方式共价连接以使该序列的正常功能能够实现。例如，如果前序列或分泌前导序列作为参与多肽的分泌的前蛋白表达，那么将编码前序列或分泌前导序列(secretory leader)的核苷酸序列可操作地连接到多肽的核苷酸序列上：启动子或增强子如果影响该序列的转录则可操作地连接到编码序列上；如果核糖体结合部位位于可促进翻译的位置，那么将核糖体结合部位可操作地连接到编码序列上。通常“可操作地连接”是指被连接的核苷酸序列是邻近的，且在分泌前导序列的情况下，是邻近的同时又是处在阅读状态的。连接可以方便的在限制性酶切位点完成。如果这样的部位不存在，那么需要应用合成的寡核苷酸衔接头或接头，并结合标准的重组 DNA 方法。

术语“引入”是指引入含非多肽组分附着基团的氨基酸残基，尤其通过对现有氨基酸基的取代，或通过插入另外的氨基酸残基。术语“除去”是指将包含非多肽组分附着基团的氨基酸残基除去，尤其通过将该氨基酸残基用另外的氨基酸残基取代的方式除去，或者选择地通过删除(没有取代)预除去的氨基酸残基。

当对亲代多肽进行取代时，这些取代优选是“保守性取代”，换句话说取代是在具有相似特点的氨基酸，例如，小氨基酸，酸性氨基酸，极性氨基酸，碱性氨基酸，疏水性氨基酸和芳香性氨基酸之间进行取代。

本发明中优选的取代尤其选自下表中列出的保守取代组

25 保守取代组：

1	丙氨酸(A)	甘氨酸(G)	丝氨酸(S)	苏氨酸(T)
2	天冬氨酸(D)	谷氨酸(E)		
3	天冬酰胺(N)	谷氨酰胺(Q)		
4	精氨酸(R)	组氨酸(H)	赖氨酸(K)	
5	异亮氨酸(I)	亮氨酸(L)	蛋氨酸(M)	缬氨酸(V)
6	苯丙氨酸(F)	酪氨酸(Y)	色氨酸(W)	

在与给定的物质联用的术语“免疫原性”是指该物质能诱导免疫系统的

反应。免疫反应可以是细胞或抗体介导的反应(对免疫原性的进一步解释参见, 例如 Roitt: Essential Immunology (8th Edition, Blackwell))。正常情况下, 抗体反应性的降低将意味着免疫原性的降低。降低的免疫原性可以通过应用本领域熟知的任何适当的方法, 例如体内的方法或体外的方法确定。

- 5 术语“体内功能半寿期”以其正常的含意使用, 即在体内/靶器官内多肽或偶联物仍然存在 50%生物学活性的时间, 或多肽或偶联物的活性是最初活性的 50%的时间。确定体内功能半寿期的另一种方法可以是确定“血清半寿期”, 即 50%的多肽或偶联物分子在被清除之前在血浆或血流中循环的时间。血清半寿期的其它术语包括“血浆半寿期”, “循环半寿期”, “血清清除
- 10 率”“血浆清除率”, “清除半寿期”。多肽或偶联物由网状内皮系统(RES), 肾脏, 脾脏或肝脏中的一个或多个, 通过受体介导的降解, 或通过特异性或非特异性蛋白水解作用, 尤其通过受体介导的清除和肾脏的清除而被清除。通常, 清除取决于蛋白的大小(相对于肾小球过滤的截留值), 电荷, 附着的糖链, 和该蛋白的细胞受体的存在。被保留的功能性通常选自增殖或受体结
- 15 合活性。体内功能半寿期和血清半寿期可以通过本领域熟知的任何适当的方法确定, 这在以下的材料和方法中进一步讨论。

- 关于体内功能半寿期或血清半寿期所用的术语“增加的”是指偶联物或多肽的相应半寿期经可比条件下的测定, 相对于参考分子, 例如非偶联的 hG-CSF (例如 Neupogen®)的半寿期在统计学意义上是增加的。例如, 相应
- 20 半寿期可能增加至少大约 25%, 例如至少大约 50%, 例如至少大约 100%, 200%, 500% 或 1000%。

- 所用术语“肾脏清除”的正常的含意是指肾脏, 例如通过肾小球的过滤, 肾小管的分泌或肾小管的排泄进行的任何清除。肾脏清除依赖于偶联物的物理特征, 包括大小(直径), 对称性, 形状/刚性和电荷。通过任何适当的检测,
- 25 例如, 一个现有体内检测法可以确定肾脏清除的降低。通常, 肾脏清除可通过给予患者一种标记的(例如, 放射性标记或荧光标记的)多肽偶联物并检测所收集的患者尿中标记物的活性进行确定。在可比较的条件下, 肾脏清除的降低可对比相关的参考多肽例如, 相应的非偶联多肽, 非偶联的相应的野生型多肽或另一偶联多肽(例如并非依照本发明的偶联多肽)确定。优选, 偶联
- 30 物的肾脏清除率与相关的参考多肽相比降低至少 50%, 优选至少降低 75%, 最优选至少降低 90%。

通常,受体的活化是和受体介导的清除(RMC)相关联的,这使得多肽与其受体在非活化的情况下的结合不能导致 RMC,而受体的活化导致 RMC。清除是由于受体结合的多肽的内化并随后被溶酶体降解所致。设计偶联物使其能够结合和活化足够数目的受体,以获得最佳体内生物学反应并避免活化比获得该反应所需数量更多的受体,从而可使 RMC 降低。这可以反映在降低的体外生物学活性和/或增加的分离率方面。

通常,体外生物学活性的降低反映了功效/效率的降低和/或有效性的降低,这可以通过用确定这些特性中任何一种的任何适当方法确定。例如,体外生物学活性可以在基于萤光素酶的试验中确定(参见材料和方法)。经可比条件下的测定,与相应的参考多肽的体外生物学活性比较,偶联物的体外生物学活性可降低至少 30%,例如可降低至少 50%, 60% 或 75%,例如降低至少 90%。换言之,偶联物可能具有相应非偶联多肽或相应野生型多肽的体外生物学活性的大约 1%,通常,至少大约 2%,例如至少大约 5%。例如,经可比条件下的测定,该体外生物学活性可以是参考多肽的大约 1-50%,例如,大约 2-40%,例如大约 5-25%。在需要降低体外生物学活性以降低受体介导的清除的情况下,但很明显为获得所期望的受体活化显然需要维持足够的生物学活性。

优选,多肽偶联物和其受体的分离率的增加将达到一定程度,使得可以导致在受体配体复合物的任何实质性内吞发生之前将多肽偶联物从其受体上释放出来。受体多肽结合亲和力(包括分离率)可以依照本文材料和方法部分的描述确定。体外 RMC 可以如下确定:标记(例如放射性的或荧光的标记)多肽偶联物,刺激包含针对该多肽的受体的细胞,洗涤这些细胞并检测标记物的活性。或者,将偶联物暴露于表达相关受体的细胞。适当时间的温育后将上清液取出转移到包含相似细胞的培养孔中。这些细胞对上清液的生物反应可以与非偶联多肽或另一参考多肽对比确定,这就是 RMC 降低程度的测量值。

通常,偶联物体外生物学活性的降低可以是其被非多肽组分修饰的结果。然而,为进一步降低体外生物学活性或为其它原因,可能还需要对偶联物的多肽部分进行修饰。例如,在一个实施方案中,与相应的野生型的多肽相比,位于多肽的受体结合区或其附近的至少一个氨基酸残基可以用另一氨基酸残基取代,以降低体外生物学活性。通过取代引入的氨基酸残基可以是

能够降低偶联物的体外生物学活性的任何氨基酸残基。便利地,被引入的氨基酸残基包含本文定义的非多肽组分附着基团。尤其,当非多肽组分是聚合物分子例如 PEG 时,被引入的氨基酸残基可以是赖氨酸残基。

术语“显示 G-CSF 活性”是指多肽或偶联物有天然 G-CSF,尤其具有 SEQ ID NO: 1 显示的氨基酸序列的 G-CSF 的一个或多个功能,包括与 G-CSF 受体结合的能力(Pukunaga et al., J.Bio.Chem, 265: 14008, 1990)。G-CSF 活性用下面材料和方法中描述的初步检测方法可以便利地检测。“显示”G-CSF 活性的多肽当显示可检测的功能,例如可检测的增殖活性或受体结合活性(例如那些通过材料和方法中描述的初步检测方法确定的活性)时,被认为具备了这样的活性。本文中显示 G-CSF 活性的多肽也可以称为“G-CSF 分子”。

术语“亲代 G-CSF”或“亲代多肽”是指将根据本发明被修饰的分子。亲代 G-CSF 通常是 hG-CSF 或其变体。“变体”是其一个或多个氨基酸残基与亲代多肽不同的多肽,通常 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 或 15 个氨基酸残基不同。rhG-CSF 的实例包括 filgrastim (Gran® 和 eupogen®), lenograstim (Neutrogen® 和 Granocyte®) 和 nartograstim (Neu-up®)。

本发明的偶联物

如上所述,本发明第一方面涉及包含显示 G-CSF 活性的多肽以及附着至该多肽的附着基团上的至少一个非多肽组分的偶联物,该多肽包含的氨基酸序列与 SEQ ID NO: 1 所示的氨基酸序列不同在于引入或去除至少一个包含非多肽组分附着基团的氨基酸残基。被引入和/或去除的氨基酸残基在以下进一步描述。应理解偶联物本身也显示 G-CSF 活性。

通过引入和/或去除包含非多肽组分附着基团的氨基酸残基,可以定向改造多肽,以使多肽分子更易于与所选定的非多肽组分偶联,使偶联模式最优化(例如保证非多肽组分在 G-CSF 分子表面最优化分布并保证分子中只存在将进行偶联的附着基团),从而获得既具备 G-CSF 活性还具有与现有 G-CSF 分子相比改善了一个或多个特性的新的偶联分子。

多肽可以是任何来源的,尤其可以来源于哺乳动物,但目前优选来源于人。

在本发明的优选实施例中,具备 G-CSF 活性的多肽的一个以上的氨基酸残基被变更,例如这种变更包括去除和引入含所选非多肽组分附着基团的

氨基酸残基。

除本发明公开的旨在去除和/或引入非多肽组分的附着位点的氨基酸残基的变更之外，可以理解本发明的多肽的氨基酸序列在需要时可以包含其它的与引入或去除附着位点无关的变更，即其它的取代，插入或缺失。这些变更可以，例如，包括截去 N 末端和/或 C 末端一个或多个氨基酸残基，或在 N 末端和/或 C 末端添加一个或多个额外的残基，例如在 N 末端添加蛋氨酸残基。

本发明的偶联物与 hG-CSF 相比，尤其与 rhG-CSF (例如 filgrastim, lenograstim 或 artograstim)或已知的 hG-CSF 变体相比有一个或多个以下改善的特性：体内的功能半寿期增加，血清半寿期增加，肾脏清除减少，受体介导的清除减少，副作用例如骨痛减弱，和免疫原性降低。

可以理解，包含非多肽组分的附着基团的氨基酸残基，不管它是被引入还是被去除，都将根据所选择的非多肽组分的性质，且在多数情况下，根据多肽和非多肽组分之间实现偶联的方法选择。例如，非多肽组分是聚合物分子例如聚乙二醇或聚环氧烷衍生的分子时，包含附着基团的氨基酸残基可以从赖氨酸，半胱氨酸，天冬氨酸，谷氨酸，组氨酸和精氨酸中选择。当偶联赖氨酸残基时，适当的活化分子是例如来自 Shearwater Polymers, Inc.的 PEG-SPA，氧基羰基氧-N-二羧酰亚胺-PEG (US 5, 122, 614)，或可从 PolyMASC Pharmaceuticals plc.得到的 PEG。对这些分子中的第一个分子将在以下作进一步描述。

为避免对亲代 hG-CSF 分子的结构和功能的破坏，依照本发明需要变更的氨基酸残基总数，例如如本发明下面描述的，(与 SEQ ID NO: 1 显示的氨基酸序列相比)通常不超过 15 个。引入或去除的实际的氨基酸残基的数目和类型尤其依赖于所期望的偶联的性质和程度(例如，非肽组分的特征，多少非多肽组分期望或可能与多肽偶联，期望偶联发生的部位或应该避免偶联发生的部位，等)。优选，本发明偶联物的多肽部分或本发明的多肽有 1-15 个氨基酸残基，通常 2-10 个氨基酸残基，例如 3-8 个氨基酸残基，例如 4-6 个氨基酸残基，与 SEQ ID NO: 1 显示的氨基酸序列相比不同。所以，通常本发明的偶联物的多肽部分或本发明的多肽有 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 或 15 个氨基酸残基不同于 SEQ ID NO: 1 显示的氨基酸序列。

通常偶联物的多肽部分有至少大约 80%与 SEQ ID NO: 1 相同的氨基酸

序列, 优选至少大约 90%, 例如至少大约 95%。氨基酸序列的同源性/同一性可以便利地, 用例如 ClustalW 程序, 版本 1.8, 1999 年 6 月, 用缺省参数从被比对的序列确定 (Thompson et al., 1994, ClustalW : Improving the sensitivity of progressive multiple sequence alignment through sequence
5 weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22: 4673-4680) 或从 PFAM 家族数据库 版本 4.0 (<http://pfam.wustl.edu/>) (Nucleic Acids Res. 1999 Jan 1; 27 (1) : 260-2) 通过应用 GENEDOC 版本 2.5 (Nicholas, K. B., Nicholas H. B. Jr., 和 Deerfield, D. W. II. 1997 GeneDoc : Analysis and Visualization of Genetic Variation, EMBNEW.
10 NEWS 4: 14; Nicholas, K. B. and Nicholas H. B. Jr. 1997 GeneDoc: Analysis and Visualization of Genetic Variation)进行确定。

在优选实施方案中, 多肽的氨基酸序列和 SEQ ID NO:1 显示的氨基酸序列之间的差异在于引入至少一个且经常是多个, 例如 1-15 个包含非多肽组分附着基团的氨基酸残基, 优选通过替换的方式引入到氨基酸序列中。所以, 多肽部分变更了与所选非多肽组分结合的特定氨基酸残基的含量, 从而
15 获得更有效的, 特异的和/或更大范围的偶联。例如, 当包含所选非多肽的附着基团的氨基酸残基总数变更到较佳水平时, 由于偶联获得的分子的形状, 大小和/或电荷的改变, 偶联物的清除通常明显降低。进一步地, 当包含所选非多肽的附着基团的氨基酸残基总数增加时, 更大比例的多肽分子被
20 所选非多肽组分掩蔽, 导致较低免疫反应。

本申请中所用术语“一种不同”是指允许另外的不同存在。所以, 除了特定的氨基酸的不同之外, 其它的氨基酸残基可以发生突变。

在另一优选实施例, 多肽的氨基酸序列和 SEQ ID NO:1 所示氨基酸序列之间的不同是至少一个, 优选多个, 例如 1-15 个含非多肽组分附着基团的氨基酸残基从氨基酸序列中被去除, 优选被取代。通过去除一个或多个
25 包含非多肽组分附着基团的氨基酸残基, 可以避免非多肽组分偶联在多肽上导致不利偶联的部位, 例如在多肽的功能部位处或其附近的氨基酸残基上 (因为在此类部位的偶联可能由于受体的识别受到影响导致所产生的偶联物失活或 G-CSF 活性下降)。本文术语“功能部位”是指一个或多个为 hG-CSF
30 功能或效应所必不可少或者相关的氨基酸残基。这样的氨基酸残基构成功能部位的一部分。功能部位可以通过本领域已知的方法确定, 优选通过分析多

肽和相关受体, 例如 hG-CSF 受体(参见 Aritomi et al., nature 401: 713-717, 1999)的复合物的结构确定。

在另一优选实施方案中, 多肽的氨基酸序列与 SEQ ID NO: 1 显示的氨基酸序列有如下不同: a)至少一个特定的包含非多肽组分附着基团且存在于 SEQ ID NO: 1 显示的氨基酸序列中的氨基酸残基被去除, 优选通过取代的方式去除, b)至少一个特定的包含非多肽组分附着基团的氨基酸残基被引入到氨基酸序列中, 优选通过取代, 上述特定的氨基酸残基可以是本文以下部分描述的那些氨基酸残基中的任何一种。这个实施方案被认为具有特定的意义, 因为它使得有可能定向设计多肽使之能与选择的非多肽组分获得最佳的偶联。例如, 通过引入和去除以下部分公开的选定氨基酸残基, 有可能保证所选择的非多肽组分的附着基团的最佳分布, 所产生的偶联物的非多肽组分的位置可以使得: a)有效地掩蔽多肽的表位和其它的表面部分和 b)保证偶联物具有最佳的 Stokes 半径, 不引起太多的结构破坏从而减弱多肽的功能。

本发明的偶联物通常包含足够数量和类型的非多肽组分以使偶联物的体内功能半寿期和/或血清半寿期与 hG-CSF, 例如 filgrastim, lenograstim 或 nartograstim 比较, 优选与包含 N 末端附着的单个 20 kDa PEG 组分的 rhG-CSF 比较得到增加。体内功能半寿期的增加可以通过如本文材料和方法中的描述便利地确定。

本发明的偶联物可以包含非多肽组分的至少一个尚未偶联、可进行偶联的附着基团。在本发明的上下文中所用的术语“可偶联的附着基团”是指位于多肽中可进行偶联的部位的附着基团, 且当进行偶联时, 若不存在特殊的预防措施, 上述附着基团便可与相关的非多肽组分偶联。例如, 这类附着基团可以是多肽表现其活性必需或相关的氨基酸残基的一部分。避免另外的可偶联的附着基团的偶联的一个便利的是通过辅助分子, 例如以在题目为“对功能部位的阻断”部分描述的方式对附着基团进行掩蔽。应理解尚未被偶联但可偶联附着基团的数目由特定的 G-SCF 多肽和可偶联的附着基团的位置决定。例如, 多肽偶联物包含一个或二个未被偶联但可偶联的附着基团, 和至少一个, 优选二个或多个已被偶联的附着基团。

本发明的偶联物, 其中非多肽组分被附着到赖氨酸上或 N 末端氨基酸残基上

本发明一方面涉及一种多肽偶联物, 该偶联物包含 i)显示 G-CSF 活性

的多肽，该多肽包含的氨基酸序列与 SEQ ID NO: 1 显示的氨基酸序列不同在于它至少存在选自以下取代中的一个取代：T1K, P2K, L3K, G4K, P5K, A6K, S7K, S8K, L9K, PICK, Q11K, S12K, F13K, L14K, L15K, E19K, Q20K, V21K, Q25K, G26K, D27K, A29K, A30K, E33K, A37K, T38K, Y39K, L41K, H43K, P44K, E45K, E46K, V48K, L49K, L50K, H52K, S53K, L54K, I56K, P57K, P60K, L61K, S62K, S63K, P65K, S66K, Q67K, A68K, L69K, Q70K, L71K, A72K, G73K, S76K, Q77K, L78K, S80K, F83K, Q86K, G87K, Q90K, E93K, G94K, S96K, P97K, E98K, L99K, G100K, P101K, T102K, D104K, T105K, Q107K, L108K, D109K, A111K, D112K, F113K, T115K, T116K, W118K, Q119K, Q120K, M121K, E122K, E123K, L124K, M126K, A127K, P128K, A129K, L130K, Q131K, P132K, T133K, Q134K, G135K, A136K, M137K, P138K, A139K, A141K, S142K, A143K, F144K, Q145K, S155K, H156K, Q158K, S159K, L161K, E162K, V163K, S164K, Y165K, V167K, L168K, H170K, L171K, A172K, Q173K 和 P174K, 和 ii)至少有一个非多肽组分附着到该多肽的赖氨酸残基上。

hG-CSF 包含四个赖氨酸残基，其中 K16 位于受体结合功能域，其它的分别位于距离受体结合功能域相对较近的 23, 34 和 40 位。为了避免这些赖氨酸残基的一个或多个氨基酸残基的偶联反应(因为这种偶联可以使所产生的偶联物的活性丧失或严重下降)需要去除至少一个赖氨酸残基，例如二个，三个或所有这些残基。所以，在另一更优选的方面本发明涉及以上定义的多肽偶联物，其中选自 K16, K23, K34 和 K40 的至少一个氨基酸残基被删除或被另一氨基酸残基取代。优选至少 K16 被另一氨基酸残基取代。

优选的氨基酸取代实例包括 Q70K, Q90K, T105K, Q120K 和 T133K 中的一个或多个，例如这些取代中的二个，三个或四个，例如：Q70K+Q90K, Q70K+T105K, Q70K+120K, Q70K+T133K, Q90K+T105K, Q90K+Q120K, Q90K+T133K, T105K+Q120K, T105K+T133K, Q120K+T133K, Q70K+Q90K+T105K, Q70K+Q90K+Q120K, Q70K+Q90K+T133K, Q70K+T105K+T120K, Q70K+T105K+T133K, Q70K+Q120K+T133K, Q09K+T105K+Q120K, Q90K+T105K+T133K, Q90K+Q120K+T133K, T105K+T120K+T133K, Q90K+T105K+Q120K+T133K, Q70K+T105K+Q120K+T133K, Q70K+Q90K+Q120K+T133K,

Q70K+Q90K+T105K+T133K 或 Q70K+Q90K+T105K+Q120K。

这一部分中描述的第二个方面的偶联物的多肽(即至少存在一个被引入和一个被去除的赖氨酸)优选包含至少一个, 例如一个, 二个, 三个或四个选自 K16R, K16Q, K23R, K23Q, K34R, K34Q, K40R 和 K40Q 的取代, 更优选包含取代 K16R 和 K23R 中的至少一个, 以便避免这些残基的偶联。优选, 上述多肽包含至少选自 K16R+K23R, K16R+K34R, K16R+K40R, K23R+K34R, K23R+K40R, K34R+K40R, K16R+K23R+K34R, K16R+K23R+K40R, K23R+K34R+K40R, K16R+K34R+K40R 和 K16R+K23R+K34R+K40R 的一个取代。这些取代可能产生与天然或亲代的多肽相比最小的结构差异。

除了以上列举的取代, 本发明的多肽偶联物也包括选自 R22K, R146K, R147K, R166K 和 R169K 的一个或多个取代。

依照本发明的这一方面偶联物的非多肽组分可以是任何分子, 当应用给定的偶联方法时, 该分子有赖氨酸作为附着基团例如糖部分的附着基团, 优选非多肽组分是聚合物分子。聚合物分子可以是在标题为“聚合物的偶联”部分中提及的分子中的任何分子, 但上述聚合物分子优选选自线性或分支状的聚乙二醇或另一聚环氧烷。优选的聚合物分子是例如来自 Shearwater Polymers, Inc.的 PEG-SPA 或 oxycarbonyl-oxy-N-dicarboxyimide PEG (US 5, 122, 614)。

可以理解, 在这一部分指定的任何氨基酸的变更, 尤其是取代可与本发明的其它部分指定的任何氨基酸变更, 优选取代, 即公开的特定氨基酸的修饰, 包括糖基化位点的引入和/或去除结合。

本发明的偶联物, 其中非多肽组分是以半胱氨酸作为附着基团的分子

本发明的另一方面涉及一种偶联物, 该偶联物包含 i) 显示 G-CSF 活性的多肽, 该多肽包含的氨基酸序列与 SEQ ID NO: 1 显示的氨基酸序列不同, 它至少存在选自以下取代中的一个取代: T1C, P2C, L3C, G4C, P5C, A6C, S7C, S8C, L9C, P10C, Q11C, S12C, F13C, L14C, L15C, E19C, Q20C, V21C, R22C, Q25C, G26C, D27C, A29C, A30C, E33C, A37C, T38C, Y39C, L41C, H43C, P44C, E45C, E46C, V48C, L49C, L50C, H52C, S53C, L54C, I56C, P57C, P60C, L61C, S62C, S63C, P65C, S66C, Q67C, A68C, L69C, Q70C, L71C, A72C, G73C, S76C, Q77C, L78C, S80C, F83C, Q86C, G87C, Q90C, E93C,

G94C, S96C, P97C, E98C, L99C, G100C, P101C, T102C, D104C, T105C, Q107C, L108C, D109C, A111C, D112C, F113C, T115C, T116C, W118C, Q119C, Q120C, M121C, E122C, E123C, L124C, M126C, A127C, P128C, A129C, L130C, Q131C, P132C, T133C, Q134C, G135C, A136C, M137C, P138C, 5 A139C, A141C, S142C, A143C, F144C, Q145C, R146C, R147C, S155C, H156C, Q158C, S159C, L161C, E162C, V163C, S164C, Y165C, R166C, V167C, L168C, R169C, H170C, L171C, A172C, Q173C 和 P174C, 和 ii)至少一个非多肽组分偶联到多肽的半胱氨酸残基上。

hG-CSF 的受体结合功能域在第 17 位包含一个半胱氨酸残基, 该半胱氨酸不形成胱氨酸, 且为了避免非多肽组分与上述半胱氨酸的偶联将该半胱氨酸去除是有利的。所以, 在本发明另一更优选方面涉及一种偶联物, 该偶联物包含 i) 显示 G-CSF 活性的多肽, 该多肽包含的氨基酸序列与 SEQ ID NO: 1 显示的氨基酸序列不同, 它至少存在选自以下取代中的一个取代: T1C, P2C, L3C, G4C, P5C, A6C, S7C, S8C, L9C, P10C, Q11C, S12C, F13C, L14C, L15C, 15 E19C, Q20C, V21C, R22C, Q25C, G26C, D27C, A29C, A30C, E33C, A37C, T38C, Y39C, L41C, H43C, P44C, E45C, E46C, V48C, L49C, LSOC, H52C, S53C, L54C, I56C, P57C, P60C, L61C, S62C, S63C, P65C, S66C, Q67C, A68C, L69C, Q70C, L71C, A72C, G73C, S76C, Q77C, L78C, S80C, F83C, Q86C, G87C, Q90C, E93C, G94C, S96C, P97C, E98C, L99C, G100C, P101C, T102C, 20 D104C, T105C, Q107C, L108C, D109C, A111C, D112C, F113C, T115C, T116C, W118C, Q119C, Q120C, M121C, E122C, E123C, L124C, M126C, A127C, P128C, A129C, L130C, Q131C, P132C, T133C, Q134C, G135C, A136C, M137C, P138C, A139C, A141C, S142C, A143C, F144C, Q145C, R146C, R147C, S155C, H156C, Q158C, S159C, L161C, E162C, V163C, S164C, Y165C, R166C, 25 V167C, L168C, R169C, H170C, L171C, A172C, Q173C 和 P174C, 上述取代与 C17 的去除结合, 优选 C17 用任何其它的氨基酸残基, 例如用丝氨酸残基取代, 和 ii) 以半胱氨酸残基作为其附着基团的非多肽组分。

依照本发明的这一方面优选的取代是用半胱氨酸取代精氨酸, 例如 R146C, R147C, R166C 和 R169C 取代中的一种或多种。

30 可以理解, 在这一部分指定的任何氨基酸修饰, 尤其是取代可与本发明的其它部分指定的任何氨基酸变更, 优选取代, 即公开的特定氨基酸的修饰,

包括糖基化位点的引入和/或去除结合。

本发明的偶联物，其中的非多肽组分结合到一个酸性基团上或 C 末端的氨基酸残基上

- 本发明进一步涉及一种偶联物，该偶联物包含 i) 显示 G-CSF 活性的多肽，该多肽包含的氨基酸序列与 SEQ ID NO: 1 显示的氨基酸序列不同，它至少存在选自以下取代中的一个取代：T1D, P2D, L3D, G4D, P5D, A6D, S7D, S8D, L9D, P10D, Q11D, S12D, F13D, L14D, L15D, K16D, Q20D, V21D, R22D, K23D, Q25D, G26D, A29D, A30D, K34D, A37D, T38D, Y39D, K40D, L41D, H43D, P44D, V48D, L49D, L50D, H52D, S53D, L54D, I56D, P57D, P60D, L61D, S62D, S63D, P65D, S66D, Q67D, A68D, L69D, Q70D, L71D, A72D, G73D, S76D, Q77D, L78D, S80D, F83D, Q86D, G87D, Q90D, G94D, S96D, P97D, L99D, G100D, P101D, T102D, T105D, Q107D, L108D, A11D, F113D, T115D, T116D, W118D, Q119D, Q120D, M121D, L124D, M126D, A127D, P128D, A129D, L130D, Q131D, P132D, T133D, Q134D, G135D, A136D, M137D, P138D, A139D, A141D, S142D, A143D, F144D, Q145D, R146D, R147D, S155D, H156D, Q158D, S159D, L161D, V163D, S164D, Y165D, R166D, V167D, L168D, R169D, H170D, L171D, A172D, Q173D 和 P174D; 或至少一个取代选自 T1E, P2E, L3E, G4E, P5E, A6E, S7E, S8E, L9E, P10E, Q11E, S12E, F13E, L14E, L15E, K16E, Q20E, V21E, R22E, K23E, Q25E, G26E, A29E, A30E, K34E, A37E, T38E, Y39E, K40E, L41E, H43E, P44E, V48E, L49E, L50E, H52E, S53E, L54E, I56E, P57E, P60E, L61E, S62E, S63E, P65E, S66E, Q67E, A68E, L69E, Q70E, L71E, A72E, G73E, S76E, Q77E, L78E, S80E, F83E, Q86E, G87E, Q90E, G94E, S96E, P97E, L99E, G100E, P101E, T102E, T105E, Q107E, L108E, A11E, F113E, T115E, T116E, W118E, Q119E, Q120E, M121E, L124E, M126E, A127E, P128E, A129E, L130E, Q131E, P132E, T133E, Q134E, G135E, A136E, M137E, P138E, A139E, A141E, S142E, A143E, F144E, Q145E, R146E, R147E, S155E, H156E, Q158E, S159E, L161E, V163E, S164E, Y165E, R166E, V167E, L168E, R169E, H170E, L171E, A172E, Q173E 和 P174E; 和 ii) 含有一个以天冬氨酸或谷氨酸残基作为其附着基团的非多肽组分。
- 依照本发明的这一方面优选取代的实例包括 Q67D/E, Q70D/E, Q77D/E, Q86D/E, Q90D/E, Q120D/E, Q131D/E, Q134D/E, Q145D/E 和 Q173D/E。

- 除了上述列举的取代, 依照上述的任何方面的偶联物的多肽包含选自 D27, D104, D109, D112, E19, E33, E45, E46, E93, E98, E122, E123, 和 E163 的至少一个氨基酸残基的去除, 优选通过取代去除。对任何其它的氨基酸残基可以进行取代, 尤其可以对天冬酰胺或谷氨酰胺残基进行取代, 以避免这些残基的偶联。尤其, 多肽可以包含以下取代中的至少一个取代: D27N, D104N, D109N, D112N, E19Q, E33Q, E45Q, E46Q, E93Q, E98Q, E122Q, E123Q 和 E163Q。优选, 在上述一个或多个位置上的氨基酸取代另外可以与以下取代中的至少一个结合: D109N, D112N, E19Q, E122Q 和 E123Q。用这些氨基酸残基中的任何一个氨基酸的取代可能产生最小的结构差异。
- 10 依照本发明这一方面的以酸性基团作为附着基团的偶联物的非多肽组分可以是具备上述特征的任何非多肽组分, 目前优选非多肽组分是聚合物分子或一种有机的衍生物, 尤其是聚合物分子, 且偶联物通过例如 Sakane and Pardridge, Pharmaceutical Research Vol. 14, No. 8, 1997, pp 1085-1091 所述进行制备。
- 15 可以理解, 在这一部分指定的任何氨基酸的变更, 尤其是取代可与本发明的其它部分指定的任何氨基酸变更, 优选取代, 即公开的特定氨基酸的修饰, 包括糖基化位点的引入和/或去除结合。

本发明的其它偶联物

- 除了上述例如在标题为“本发明的偶联....”的那部分中指定的非多肽组分外, 本发明的偶联物可以进一步包含多肽在可糖基化的宿主细胞中被表达且在糖基化位点或被引入的糖基化位点上获得糖基化时所产生的糖部分。
- 20

本发明的偶联物, 其中非多肽组分是糖组分

- 本发明进一步涉及一种偶联物, 该偶联物包含一个显示 G-CSF 活性的糖基化的多肽, 该多肽包含的氨基酸序列和 SEQ ID NO: 1 显示的氨基酸序列不同, 它至少有一个非天然存在的糖基化位点是通过选自以下取代中的至少一个取代的方式被引入到氨基酸序列中: L3N+P5S/T, P5N, A6N, S8N+P10S/T, P10N, Q11N+F13S/T, S12N+L14S/T, F13N+L15S/T, L14N+K16S/T, K16N+L18S/T, E19N+V21S/T, Q20N+R22S/T, V21N+K23S/T, R22N+I24S/T, K23N+Q25S/T, Q25N+D27S/T, G26N+G28S/T, D27N+A29S/T, A29N+L31S/T, A30N+Q32S/T, E33N+L35S/T, A37N+Y39S/T, T38N+K40S/T, Y39N+L41S/T, P44N+E46S/T, E45N+L47S/T, E46N+V48S/T, V48N+L50S/T,
- 25
- 30

L49N+G51S/T, L50N+H52S/T, H52N+L54S/T, S53N+G55S/T, P60N, L61N, S63N+P65S/T, P65N+Q67S/T, S66N+A68S/T, Q67N+L69S/T, A68N+Q70S/T, L69N+ L71S/T, Q70N+A72S/T, L71N+G73S/T, G73N+L75S/T, S76N+L78S/T, Q77N+H79S/T, L78N, S80N+L82S/T, F83N+Y85S/T, Q86N+L88S/T, 5 G87N+L89S/T, Q90N+L92S/T, E93N+I95S/T, P97N+L99S/T, L99N+P101S/T, P101N+L103S/T, T102N+D104S/T, D104N+L106S/T, T105N+Q107S/T, Q107N+D109S/T, L108N+V110S/T, D109N+A111S/T, A111N+F113S/T, D112N+A114S/T, F113N, T115N+I117S/T, T116N+W118S/T, W118N+Q120S/T, Q119N+M121S/T, Q120N+E122S/T, M121N+E123S/T, E122N+L124S/T, 10 E123N+G125S/T, L124N+M126S/T, M126N+P128S/T, P128N+L130S/T, L130N+P132S/T, P132N+Q134S/T, T133N+G135S/T, Q134N+A136S/T, A136N+P138S/T, P138N+F140S/T, A139N+A141S/T, A141N+A143S/T, S142N+F144S/T, A143N+Q145S/T, F144N+R146S/T, Q145N+R147S/T, R146N+A148S/T, R147N+G149S/T, S155N+L157S/T, H156N+Q158S/T, 15 S159N+L161S/T, L161N+V163S/T, E162N, V163N+Y165S/T, S164N+R166S/T, Y165N+V167S/T, R166N+L168S/T, V167N+R169S/T, L168N+H170S/T, R169N+L171S/T 和 H170N+A172S/T, 其中 S/T 代表一个 S 或 T 残基, 优选 T 残基。

可以理解, 为了制备依照这一方面的偶联物, 其多肽必须在能够将寡糖 20 部分附着到糖基化位点的可进行糖基化的宿主细胞中表达, 或者使多肽在体外得到糖基化。可糖基化的宿主细胞的实例在下面题目为“寡糖部分的偶联”部分给出。

或者, 依照这一方面的偶联物包含显示 G-CSF 活性的多肽, 该多肽包含的氨基酸序列与 SEQ ID NO: 1 显示的氨基酸序列不同, 它存在选自 P5N, 25 A6N, P10N, P60N, L61N, L78N, F113N 和 E162N, 更优选选自 PSN, A6N, PION, P60N, L61N, F113N 和 E162N, 例如选自 P60N, L61N, F113N E162N 中的至少一个取代。

或者, 本发明该方面的偶联物包含显示 G-CSF 活性的多肽, 该多肽含有与 SEQ ID NO: 1 所示氨基酸序列不同在于选自下组的至少一个取代的氨基酸序列: 30 D27N+A29S, D27N+A29T, D104N+L106S, D104N+L106T, D109N+A111S, D109N+A111T, D112N+A114S 和 D112N+A114T, 更优选选

自 D27N+A29S, D27N+A29T, D104N+L106S, D104N+L106T, D112N+A114S, 和 D112N+A114T, 如选自 D27N+A29S, D27N+A29T, D104N+L106S, 和 D104N+L106T.

除了糖分子, 在这部分描述的依照本发明这一方面的偶联物可以包含另
5 外的非多肽组分, 尤其是在本申请中描述的, 与偶联物的多肽部分存在的一个或多个附着基团偶联的聚合物分子。

可以理解, 在这一部分指定的任何氨基酸的变更, 尤其是取代可与本发明的其它部分指定的任何氨基酸变更, 尤其是取代, 即公开的特定氨基酸的修饰结合。

10 环状变换的变体(Circularly permuted variants)

在进一步的实施例中, 本发明的多肽偶联物的多肽部分可以是本发明中另外公开的多肽序列的环状变换的变体形式。在这样的环状变换的多肽中, 当新的 N 末端和 C 末端在两个原来由肽键连接的邻近的氨基酸残基之间形成时, 最初的 N 末端和 C 末端或者直接地通过肽键或者通过一个肽接头间接地连接在一起。因为通常最初的 N 末端和 C 末端彼此之间有一些距离,
15 所以它们通常通过有适当的长度和成分的肽接头的方式连接, 以使偶联物的结构和活性不受负面地影响。很明显, 新的 N 末端和 C 末端在氨基酸残基对之间不应该形成, 因为这会干扰多肽的活性。环状变换的 G-CSF 受体拮抗剂在 US 6, 100, 070 中公开, 关于肽接头和新的 N 末端和 C 末端位置和制备该变体的方法的进一步的参考信息可以从中得到。
20

本发明的偶联物对白细胞和嗜中性粒细胞的形成作用

在进一步的实施例中, 本发明的多肽偶联物的特征是偶联物显示 G-CSF 活性且包含与 SEQ ID NO: 1 显示的氨基酸序列至少有一个氨基酸不同的氨基酸序列, 并且该偶联物至少有一个非多肽组分附着到多肽的附着基团上,
25 多肽的偶联物至少进一步满足下列(A)-(D)标准中的至少一个标准:

(A) 一次皮下给予大鼠每 kg 体重 100mg 药物后(在偶联物多肽部分的重量的基础上):

i) 与给药每 kg 体重 100 μ g 未偶联的 hG-CSF 相比, 给药后 12 小时白细胞形成增加的数率至少与其相同且至少达到同一水平(按每升血的细胞数量
30 计算), 和

ii) 在至少大约 96 小时, 优选至少大约 120 小时期间, 白细胞的水平增

加(按每升血的细胞数量计算)其水平高于给药前的白细胞水平;

(B) 一次皮下给予大鼠每 kg 体重 25 μ g 药物后(在偶联物多肽部分的重量的基础上):

5 i) 与给药每 kg 体重 100 μ g 未偶联的 hG-CSF 相比, 给药后 12 小时白细胞形成增加的数率至少与其相同且至少达到同一水平(按每升血的细胞数量计算), 和

ii) 在至少大约 72 小时, 优选至少大约 96 小时, 更优选至少大约 120 小时期间, 白细胞的水平增加(按每升血的细胞数量计算), 其水平高于给药前的白细胞水平;

10 (C)一次皮下给予大鼠每 kg 体重 100 μ g 药物后(在偶联物多肽部分的重量的基础上):

i) 与给药每 kg 体重 100 μ g 未偶联的 hG-CSF 相比, 给药后 12 小时白细胞形成增加的数率至少与其相同且至少达到同一水平(按每升血的细胞数量计算), 和

15 ii) 在至少大约 96 小时, 优选至少大约 120 小时期间, 白细胞的水平增加(按每升血的细胞数量计算)其水平高于给药前的白细胞水平;

(D)一次皮下给予大鼠每 kg 体重 25 μ g 药物后(在偶联物多肽部分的重量的基础上):

20 i)与给药每 kg 体重 100 μ g 未偶联的 hG-CSF 相比, 给药后 12 小时嗜中性粒细胞形成增加的数率至少与其相同且至少达到同一水平(按每升血的细胞数量计算), 和

ii)在至少大约 72 小时, 优选至少大约 96 小时, 更优选至少大约 120 小时期间, 嗜中性粒细胞的水平增加(按每升血的细胞数量计算), 其水平高于给药前的嗜中性粒细胞水平。

25 本发明的偶联物的非多肽组分

如以上进一步指出的本发明的偶联物的非多肽组分优选选自聚合物分子, 亲脂性化合物, 糖部分(例如, 通过体内糖基化的方式)和有机衍生物。所有这些制品都可以呈现所期望的对偶联物的多肽部分的特性, 尤其是体内功能半寿期和/或血清半寿期增加。偶联物的多肽部分通常仅仅与一种类型的非多肽组分偶联, 但也可以与二种或多种不同类型的非多肽组分, 例如与
30 聚合物分子和一种寡糖组分, 与一种亲脂性的基团和一种寡糖组分, 与一种

有机衍生物和一种寡糖组分，与一种亲脂性的基团和一种聚合物分子等等偶联。与二种或多种不同的非多肽组分的偶联可以同时也可以先后进行。

制备本发明偶联物的方法

- 在以下“与亲脂性化合物的偶联”，“与聚合物分子的偶联”，“与寡糖组分的偶联”和“与有机衍生物偶联”部分，对与特定类型非多肽组分的偶联进行描述。通常，依照本发明的多肽偶联物可以通过在有益于多肽表达的条件下培养适当的宿主细胞制备，然后获得多肽，其中 a)该多肽包含至少一种非 O-糖基化位点且宿主细胞是能够进行体内糖基化的真核宿主细胞，和/或 b)在体外多肽易于和非多肽组分偶联。

10 与亲脂性化合物的偶联

- 多肽和亲脂性化合物可以通过直接的或通过应用接头彼此偶联。亲脂性化合物可以是天然化合物例如一种饱和的或非饱和的脂肪酸，脂肪酸二酮，萜，前列腺素，维生素，类胡萝卜素或类固醇，或合成的化合物例如碳酸，乙醇，胺和具有一个或多个烷基，芳基，链烯基或其它多元不饱和化合物。
- 15 多肽和亲脂性化合物之间的偶联，选择地通过一种接头，可以按照本领域已知的方法，例如按照 Bodanszky Peptide Synthesis John Wiley, New York, 1976 and in WO 96/12505 描述的方法进行。

与聚合物分子的偶联

- 将与多肽偶联的聚合物分子可以是任何一种适当的聚合物分子，例如天然
- 20 然的或合成的均聚物或杂聚物，通常多肽与分子量在大约 300-100,000Da 的范围的分子偶联，例如在大约 500-20,000 Da，更优选在大约 1000-15,000 Da，甚至更优选在大约 2000-12,000 Da，例如在大约 3000-10,000 的范围之间的分子。本发明关于聚合物分子所使用的一词“大约”是指一个大概范围的分子量且反映这样一个事实，即通常在给定的聚合物制品中会有一个确定的
- 25 分子量分布。

均聚物的实例包括多羟基化合物(即聚-OH)，聚胺(即聚-NH₂)和聚羧酸(即聚-COOH)。杂聚物是包含不同偶联基团，例如羟基和胺基的聚合物。

- 适当的聚合物分子的实例包括选自聚环氧烷(PAO)，包括聚亚烷基二醇(PAG)，例如线性的或分支的聚乙二醇(PEG)和聚丙烯乙二醇(PPG)，聚乙烯醇(PVA)，聚羧酸盐(poly-carboxylate)，poly-(vinylpyrrolidone)，聚乙烯-马来酸酐
- 30 酐 (polyethylene-co- maleic acid anhydride)，聚苯乙烯马来酸酐酐

(polystyrene-co- maleic acid anhydride), 右旋糖酐, including 羧甲基右旋糖酐或任何其它的适合降低免疫原性和/或增加体内功能半寿期和/或血清半寿期的生物聚合物的聚合物分子。聚合物分子的另一实例是人白蛋白或另一丰富的血浆蛋白。通常, 聚环氧烷来源的聚合物是适合生物的, 无毒性的, 无抗原性的, 无免疫原性的, 有各种水溶解特性, 且易从活的生物体排泄。

PEG 是优选的聚合物分子, 因为与多肽例如右旋糖酐相比它仅有几个能够进行交叉连接的反应基团。尤其, 单一功能的 PEG, 例如甲氧基聚乙二醇(mPEG), 因为它的偶联化学反应相对简单(仅仅一个反应基团可用于与多肽上的附着基团的偶联)所以受到关注。因而, 交叉连接的危险被排除, 所产生的多肽偶联物是较均一的且聚合物分子与多肽的反应比较容易受到控制。

为了实现聚合物分子和多肽的共价连接, 提供的聚合物分子的羟基结尾的基团是活化的形式, 即是功能性的反应基团。适当的活化的聚合物分子可以购买到, 例如从 Shearwater Polymers, Inc., Huntsville, AL, USA, 或从 PolyMASC Pharmaceuticals plc, UK. 购买到。或者, 聚合物分子可以通过本领域已知的便利的方法, 例如 WO 90/13540 中公开的方法活化。本发明中应用的活化的线性或分支状的聚合物分子的具体的实例在 Shearwater Polymers, Inc. 1997 and 2000 Catalogs(Functionalized Biocompatible Polymers for Research and pharmaceuticals, Polyethylene Glycol and Derivatives, 纳入本发明作参考)中描述。活化的 PEG 聚合物的具体实例包括以下线性 PEGs: NHS-PEG (例如 SPAPEG, SSPA-PEG, SBA-PEG, SS-PEG, SSA-PEG, SC-PEG, SG-PEG, 和 SCM-PEG), 和 NOR-PEG), BTC-PEG, EPOX-PEG, NCO-PEG, NPC-PEG, CDI-PEG, ALD-PEG, TRESPEG, VS-PEG, IODO-PEG, 和 MAL-PEG, 和分支状的 PEGs 例如 PEG2-NHS 和那些在 US 5, 932, 462 and US 5, 643, 575 公开的聚合物, 它们都被纳入本发明作参考。进一步地, 在以下被纳入本发明作参考的出版物中, 公开了有用的聚合物分子和/或 PEG 化合物: US 5, 824, 778, US 5, 476, 653, WO 97/32607, EP 229, 108, EP 402, 378, US 4, 902, 502, US 5, 281, 698, US 5, 122, 614, US 5, 219, 564, WO 92/16555, WO 94/04193, WO 94/14758, WO 94/17039, WO 94/18247, WO 94/28024, WO 95/00162, WO 95/11924, W095/13090, WO 95/33490, WO 96/00080, WO 97/18832, WO 98/41562, WO 98/48837, WO 99/32134, WO 99/32139, WO

99/32140, WO 96/40791, WO 98/32466, WO 95/06058, EP 439 508, WO 97/03106, WO 96/21469, WO 95/13312, EP 921 131, US 5, 736, 625, WO 98/05363, EP 809 996, US 5, 629, 384, WO 96/41813, WO 96/07670, US 5, 473, 034, US 5, 516, 673, EP 605 963, US 5, 382, 657, EP 510 356, EP 400 472, EP 5 183 503 and EP 154 316.

多肽和活化的聚合物分子的偶联可通过应用任何便利的方法,例如按照以下参考文献中描述的方法(这些文献也描述了聚合物分子的适当的活化方法)进行: R. F. Taylor, (1991), "Protein immobilisation. Fundamental and applications", Marcel Dekker, N. Y.; S. S. Wong, (1992), "Chemistry of Protein
10 Conjugation and Crosslinking", CRC Press, Boca Raton; G. T. Hermanson et al., (1993), "Immobilized Affinity Ligand Techniques", Academic Press, N. Y.). 根据多肽的附着基团(实例在上面已给出)和聚合物的功能基团(例如胺, 羟基, 羧基, 醛, sulfydryl, 琥珀酰亚胺(succinimidyl), 马来酰亚胺, viny sulfone 或卤代醋酸盐(haloacetate)技术人员知道所要应用的活化方法和/或偶联化学。
15 PEG 化可以是直接指向所有的多肽上的可利用的附着基团(即被暴露在多肽表面的那些附着基团)的偶联或者可以是直接指向一个或多个特定的附着基团, 例如 N 末端氨基(US 5, 985, 265)的偶联。进一步地, 偶联可以以一步或逐步的方式完成(例如象 WO 99/55377 描述的那样)。

可以理解, 为了制备最佳的分子, 针对附着的 PEG 分子的数量, 附着
20 的 PEG 分子的大小和形状(例如, 这些分子是线性的还是分支的), 这些分子附着在多肽中的位置, 可以对 PEG 化进行设计。对使用的聚合物的分子量的选择要考虑到期望达到的效果。例如如果偶联的主要目的是获得高分子量和较大外形的偶联物(例如降低肾脏的清除), 那么可以选择与一个或几个高分子量的聚合物分子或许多较小分子量的聚合物分子偶联以获得所期望的
25 效果。然而, 优选, 使用几个分子量较低的聚合物分子。当存在多数表位需要掩蔽时, 可以应用足够数量的低分子量的聚合物分子(例如分子量大约在 5, 000 Da)以有效掩蔽所有或大部分多肽表位。例如可以应用 2-8, 例如 3-6 个这样的聚合物。如以下列举的实例, 与应用具有较高分子量的较少的聚合物分子(例如 1-3 个分子量为 12, 000-20, 000 的分子)相比, 应用具有较低分
30 子量的较多的聚合物分子(例如 4-6 个分子量为 5000 的分子)在改善多肽偶联物的体内功能半寿期方面更有优点, 即使在附着的聚合物分子的总分子量相

同的情况下,也是如此。可以相信,较小分子量的较多的聚合物分子的存在比例如单独一个较大的聚合物分子可以为多肽提供较大的半径或外观形状,至少当聚合物分子是相对均一地分布在多肽表面时情况是这样的。

‘只有一个聚合物分子偶联到蛋白的单个附着基团上不是优选的,在这种
5 仅仅一个聚合物分子偶联的情况下,通常该聚合物分子可以是线性的或分支的,且分子量相对高,例如大约 20kDa 时是有利的。

通常,聚合物的偶联是在使尽可能多的可利用的聚合物附着基团与聚合物分子反应的条件下进行的。这可以通过使聚合物的摩尔数适当超过多肽的摩尔数的方式达到。活化的聚合物分子与多肽的典型的摩尔比是达到大约
10 1000-1, 例如达到大约 200-1 或达到大约 100-1。然而,在一些情况下,上述比例可以有所降低,例如达到大约 50-1, 10-1 或 5-1。

本发明对聚合物分子通过一个接头与多肽偶联也进行了研究。适当的接头对技术人员来说是熟知的。优选的实例是氰尿酸氯(Abuchowski et al., (1977), J.Biol.Chem. 252, 3578-3581 ; US 4, 179, 337; Shafer et al., (1986), J.
15 Polym. Sci. Polym. Chem. Ed., 24, 375-378)。

偶联后,按本领域已知的方法如通过将伯胺加入到反应混合物中而封闭残留的活化聚合物分子,并通过合适的方法除去所产生的失活的聚合物分子(参见材料与方法)。

在一个优选实施例中,本发明的多肽偶联物包含 PEG 分子,其附着至
20 可用于 PEG 化多肽中的一些,大多数或优选基本上全部的赖氨酸残基上,所述 PEG 分子具体地是线性或分支的 PEG 分子,例如分子量约 1-15kDa,通常约 2-12kDa,例如约 3-10kDa,例如约 5 或 6kDa 的 PEG 分子。

可以理解,根据多种情况,例如多肽的氨基酸序列,所使用的活化的 PEG 化合物的性质和特定的 PEG 化条件,包括 PEG 与多肽的摩尔比,可获得不同程度的 PEG 化,较高程度的 PEG 化通常可以通过 PEG 与多肽的较高
25 比率获得。然而,由任何给定的 PEG 化过程所产生的 PEG 化多肽通常将包含多肽偶联物的随机分布。

在另一实施例中,本发明的多肽偶联物包括附着至可用于 PEG 化多肽赖氨酸残基上的 PEG 分子,和另外附着至多肽的 N 末端氨基酸残基上的 PEG
30 分子。

与寡糖组分的偶联

- 与寡糖组分的偶联可以在体内或体外进行。为了完成包含一个或多个糖基化位点的 G-CSF 分子的体内糖基化, 必须将编码多肽的核苷酸序列插入到可糖基化的真核表达宿主中。表达宿主细胞可选自真菌(丝状真菌或酵母)、昆虫、动物细胞, 或选自转基因植物细胞。在一个实施方案中, 宿主细胞是
- 5 哺乳动物细胞, 如 CHO 细胞、BHK 细胞或 HEK 细胞, 如 HEK293, 或昆虫细胞, 如 SF9 细胞, 或者酵母细胞如酿酒酵母(*Saccharomyces Cerevisiae*)、毕赤巴斯德酵母(*Pichia pastoris*)或下文描述的任何宿主细胞。也可应用体外糖(右旋糖酐)与多肽的氨基酸残基的共价偶联, 例如, 按照 WO 87/05330 和 Aplin 等, CRC Crit Rev. Biochem., pp. 259-306, 1981 中所描述的方法进行。
- 10 寡糖组分或 PEG 与蛋白-和肽-结合的 Gln 残基的在体外的偶联可通过谷氨酰胺转移酶(TG 酶)来实现。在所谓交联反应中谷氨酰胺转移酶催化供体的胺基转移到蛋白结合型和肽结合型的 Gln 残基上。供体的胺基可以是蛋白结合型或肽结合型的, 例如象赖氨酸残基中的 ϵ 胺基, 或者可以是小的或大的有机分子的一部分。在谷氨酰胺转移酶催化的交联中一个可起胺基供体作用的小的有机分子的实例是腐胺(1, 4 丁二胺)。在谷氨酰胺转移酶催化的交
- 15 联中起胺基供体作用的较大的有机分子的一个实例是包含胺基的 PEG(Sato et al., Biochemistry 35, 13072-13080)。

通常, 谷氨酰胺转移酶是高度特异性的酶, 不是每个暴露在蛋白质表面的 Gln 残基都可以在谷氨酰胺转移酶的催化下交联到包含胺基的底物上。相反, 仅仅几个 Gln 残基可自然起到谷氨酰胺转移酶底物的功能, 但决定哪一个 Gln-残基是谷氨酰胺转移酶的好的底物的准确标准尚未可知。所以, 为了使蛋白易于进行谷氨酰胺转移酶催化的交联反应, 通常在方便的位置添加已知可以很好地行使谷氨酰胺转移酶底物功能的氨基酸序列是首要条件。已知几种氨基酸序列本身是或者包含极好的天然谷氨酰胺转移酶底物例如 P 物

20 质, 快反应型弹性蛋白酶抑制肽, 血纤蛋白原, 纤连蛋白, α_2 纤溶酶抑制物, α 酪蛋白和 β 酪蛋白。

与有机衍生剂的偶联

- 显示 G-CSF 活性的多肽的共价修饰可通过多肽的附着基团与有机衍生剂反应来进行。合适的衍生剂和方法在本领域中是公知的。例如, 半胱氨酰基残基常常与 α -卤代乙酸酯(和相应的胺), 如氯乙酸或氯乙酰胺反应产生羧甲基或羧基酰胺基甲基衍生物。半胱氨酰基残基也可与溴三氟丙酮、 α -溴
- 30

- β -(4-咪唑基)丙酸、氯乙酰基磷酸酯、N-烷基马来酰亚胺、3-硝基-2-吡啶二硫化物、甲基 2-吡啶二硫化物、对氯汞基苯甲酸盐、2-氯汞基-4-硝基酚或氯-7-硝基苯并-2-氧杂-1, 3-二唑反应进行衍生。组氨酸残基通过与二乙基焦碳酸酯在 pH5.5-7.0 下反应进行衍生, 因为该试剂对组氨酸侧链来说是相对特异的。对溴苯甲酰甲基溴化物也可用于衍生。该反应优选在 0.1M 卡可酸钠中于 pH6.0 下进行反应。赖氨酸和氨基末端残基可与琥珀酸酐或其它羧酸酐反应。用这些试剂衍生具有使赖氨酸残基的电荷逆转的作用。衍生含 α -氨基的残基的其它合适的试剂包括亚氨基酯如 methyl picolinimide、磷酸吡多醛、吡多醛、氯代氢硼化物、三硝基苯磺酸、邻甲基异脲、2, 4-戊二酮和转氨酶催化的与乙醛酸盐的反应。精氨酸残基通过与一种或多种常规试剂的反应来修饰, 所述试剂包括苯基乙二醛、2, 3-丁二酮、1, 2-环己二酮和茚三酮。精氨酸残基的衍生要求反应在碱性条件下进行, 因为胍功能基具有高的 pKa。

而且, 这些试剂可与赖氨酸以及精氨酸胍基反应。羧基侧基(天冬氨酸基或谷氨酸基)通过与碳二亚胺($R-N=C=N-R'$)反应来进行选择性地修饰, 其中 R 和 R' 是不同的烷基, 如 1-环己基-3-(2-咪唑基-4-乙基)碳二亚胺或 1-乙基-3-(4-氮杂-4, 4-二甲基戊基)碳二亚胺。而且, 天冬氨酸基和谷氨酸基通过与铵离子反应转化为天冬酰胺基和谷氨酰胺基。

功能性位点的封闭

已经报道聚合物的过度偶联可导致与聚合物偶联的多肽活性丧失。该问题可通过如去除位于功能性位点的附着基团或在偶联前封闭功能性位点而消除。后一种策略构成本发明进一步的技术方案(第一种策略在上文中进行了举例说明, 如通过去除邻近功能性位点的赖氨酸残基)。更具体来说, 按照第二种策略, 多肽和非多肽组分之间的偶联是在多肽功能性位点在能够结合多肽功能性位点的辅助分子封闭的条件下进行的。

优选地, 辅助分子能特异性识别多肽的功能性位点, 如受体, 特别是 G-CSF 受体或 G-CSF 受体的一部分。

另外, 辅助分子可以是抗体, 特别是识别显示 G-CSF 活性的多肽的单克隆抗体。特别是, 辅助分子可以是中和性单克隆抗体。

可使多肽在偶联之前与辅助分子相互作用。这确保多肽的功能性位点被掩盖或保护并因此不能被非多肽组分如聚合物衍生。从辅助分子洗脱后, 非

多肽组分和多肽之间的偶联可用至少部分保留的功能性位点来恢复。

具有被封闭的功能性位点的多肽与聚合物、亲脂性化合物、寡糖组分、有机衍生剂或任何其它化合物的随后偶联以常规方式进行，如按上文“与……偶联”的章节中所述的方法进行。

- 5 不考虑用于掩盖偶联物中多肽功能性位点的辅助分子的性质，理想的是辅助分子不含或仅仅含有少量的与分子中指定非多肽组分结合的基团，其中与这些基团的偶联将阻止偶联的多肽从辅助分子上解除吸附。因此，可与多肽非掩盖部分中存在的附着基团选择性偶联并且辅助分子可反复用于重复偶联。例如，如果非多肽组分是聚合物分子如 PEG 这种具有赖氨酸 ϵ 氨基或
- 10 N-末端氨基酸残基作为附着基团的分子时，理想的是辅助分子基本上不含可偶联的 ϵ 氨基，优选不含任何 ϵ 氨基。因此，在优选的实施方案中，辅助分子是能够与多肽功能性位点结合的蛋白质或肽，该蛋白质或肽不含任何与指定非多肽组分可偶联的附着基团。

- 15 本发明的实施例中多肽偶联物是从多种多样的编码目的多肽的核苷酸序列中制备的，功能基团的阻断是在偶联之前在微滴定板中实现的，例如，将表达出来的多肽变体加入到包含被固定的阻断基团，例如受体，抗体或类似物的微滴定板中。

- 20 在另一实施例中，辅助分子首先与固相，例如充填材料，例如交联葡聚糖或琼脂糖珠，的柱子，或一种表面，例如反应器皿共价连接。随后，将多肽装载到携带辅助分子的柱体材料上，然后依照本领域熟知的方法，例如按照上面题目为“与…的偶联”章节中描述的进行偶联。该程序允许多肽偶联物通过洗脱与辅助分子分离。多肽偶联物通过常规技术在不导致多肽偶联物实质性降解的物理化学条件下洗脱。包含多肽偶联物的液相从固相中被分离，但辅助分子仍然与多肽偶联物共价连接。多肽偶联物与辅助分子的分离
- 25 可以用其它的方法实现：例如，辅助分子可以用一个第二分子(例如生物素)衍生，所述第二分子可以被一个特异性结合物(例如链霉抗生物素)识别。可以将特定的结合物与固相连接，这样使多肽偶联物与辅助分子-第二分子复合物在经过第二个辅助物-固相柱时得到分离，在随后洗脱时，辅助分子-第二分子复合物保留在固相柱上，而多肽偶联物从柱上被洗脱下来。多肽偶联
- 30 物可以任何适当的方式从辅助分子中释放出来。通过提供辅助分子从与其结合的 G-CSF 的功能位点分离的条件可解除保护作用。例如，聚合物偶联的

抗体和抗独特型抗体之间形成的复合物可以通过将 pH 值调整到酸性或碱性水平得到解离。

标记后的多肽的偶联

在另一实施方案中多肽和一段标记物，即通常一段由 1-30，例如 1-20 个氨基酸残基组成的一段氨基酸序列或肽片段作为融合蛋白被表达。除了能使纯化快速和容易外，该标记物还是实现被标记的多肽和非多肽组分之间偶联的便利工具。尤其，标记物可以用来完成微滴定板中的偶联或通过标记物将标记后的多肽固定到其它载体，例如顺磁性珠上。在微滴定板上与标记的多肽偶联具有以下优点，标记的多肽可从培养肉汤直接固定在微滴定板上（原则上不经任何纯化）并进行偶联。因此，可减少操作步骤的总数（从表达到偶联）。而且，标记物可起间隔臂分子的作用以确保更易于使固定的多肽偶联。使用标记多肽进行的偶联可以是与本文中公开的任何非多肽组分的偶联，如与聚合物分子如 PEG 的偶联。

所要应用的特定标记物的特征并不重要，只要该标记物能够和多肽一起被表达且能够被固定到适当的表面或载体物质上。可以市场上购买到许多适当的标记物，例如购自 Unizyme boratories, Denmark.。例如所述标记物可以是任一下列序列：

His-His-His-His-His-His

Met-Lys-His-His-His-His-His

Met-Lys-His-His-Ala-His-His-Gln-His-His

Met-Lys-His-Gln-His-Gln-His-Gln-His-Gln-His-Gln

Met-Lys-His-Gln-His-Gln-His-Gln-His-Gln-His-Gln-Gln

或以下的任何序列：

EQKLISEEDL(一种在 Mol.Cell.Biol.C 5: 3610-16, 1985 中描述的 C 末端标记物)

DYKDDDDK (一种 C 末端或 N 末端标记物)

YPYDVPDYA

也可以从市场上获得抗上述标记物的抗体，例如可购自 ADI, Aves Lab and Research Diagnostics。

应用标记后的多肽进行 PEG 化便利方法在以下材料与方法部分描述。随后可以用从市场上可购到的酶将标记物从多肽中切割下来。

制备本发明的多肽或本发明偶联物的多肽组分的方法

本发明的多肽或偶联物的多肽部分，选择地，以糖基化形式，通过本领域熟知的任何适当的方法制备。所述方法包括构建编码多肽的核苷酸序列并且在适当的转化或转染宿主中表达。然而，本发明的多肽也可以通过化学合成的方法制备，尽管这种方法的效率较低，或者将化学合成和重组 DNA 技术结合应用。

编码本发明的多肽或偶联物的多肽部分的核苷酸序列可以通过分离或合成编码亲代 G-CSF，例如具有 SEQ ID NO : 1 显示的氨基酸序列的 hG-CSF 的核苷酸序列构建，然后改变核苷酸序列以引入(即插入或取代)或删除(即去除或取代)相关的氨基酸残基。

依照常规方法通过定点突变可便利地修饰核苷酸序列。或者，可通过化学合成核苷酸序列，例如通过应用一种寡聚核苷酸合成器制备，其中寡聚核苷酸是在目的多肽的氨基酸序列的基础上设计的，且优选那些有利于在制备重组多肽的宿主细胞中表达的那些密码子。例如可以合成编码目的多肽部分的几个小寡核苷酸，然后通过 PCR，连接反应或连接酶链反应(LCR)将所述寡核苷酸装配起来(Barany, PNAS. 88 : 189-193, 1991)。通常，每个寡核苷酸包含用于补足性装配的 5'或 3'的悬垂部分。

可以应用另外的核苷酸序列修饰方法，这种方法用于制备高产量筛选的多肽变体，例如，在 US 5, 093, 257, 中公开的涉及同源交叉的方法，和涉及基因改组(shuffling)的方法，即二个或多个同源核苷酸序列的重组导致新的核苷酸序列与起初的核苷酸序列比较具有多个核苷酸改变。基因改组(也称为 DNA 改组)涉及一个或多个随机片段的循环和核苷酸序列的组装，然后进行筛选以选择编码具有所需特性的多肽的核苷酸序列。为了进行同源基础的核酸改组，核苷酸序列相关的部分优选至少 50%相同，例如至少 60%相同，更优选至少 70%相同，例如至少 80%相同。重组可以在体外或者体内进行。

适当的体外基因改组方法的实例由 Stemmer et al., (1994), Proc. Natl. Acad. Sci. USA; vol. 91, pp. 10747-10751; Stemmer (1994), nature vol. 370, pp. 389-391; Smith (1994), Nature vol. 370, pp. 324-325; Zhao et al., nature Biotechnology, 1998, Mar; 16 (3): 258-61; Zhao H. and Arnold, FB, 1997, Vol. 25. No. 6 pp. 1307-1308; Shao et al., Nucleic Acids Research 1998, Jan 15; 26 (2): pp. 681-83; and WO 95/17413 公开。适当的体内改组方法的实例在 WO

97/07205 中公开。其它的通过体外或体内重组的核酸序列的突变技术已由例如 WO 97/20078 和 US 5,837,458 公开。特定的改组技术的实例包括“家族改组”，“合成改组”和“计算机模拟(in silico)改组”。家族改组涉及将来自不同种属的同源基因的家族进行一个或多个改组循环，然后进行筛选或选择。家族改组技术已由例如 Crameriet al., (1998), nature vol. 391, pp. 288-291; Christians et al., (1999), nature Biotechnology vol. 17, pp. 259-264; Chang et al., (1999), nature Biotechnology vol. 17, pp. 793-797; and Ness et al., (1999), nature Biotechnology Vol. 17, 893-896 公开。合成改组涉及提供重叠的合成寡核苷酸文库，所述寡核苷酸是以例如令人感兴趣的同源基因序列比对为基础的。将合成的寡核苷酸重组，然后将所形成的重组核酸序列进行筛选，如果需要，可将所形成的重组序列用于进一步的改组循环。合成改组技术在 WO 00/42561 中公开。In silico 改组中涉及 DNA 的改组程序，该改组程序是应用计算机系统进行或模拟的，所以部分或全部避免了核酸的人工操作。WO 00/42560 公开了 In silico 改组的技术。

一旦完成组装(通过合成，定点突变或其它的方法)，即将编码多肽的核苷酸序列插入到一种重组载体中，且可操作地连接到期望转化的宿主细胞中 G-CSF 表达所需的调控序列上。

可以理解，不是所有的载体和表达调控序列都能同样完好的表达编码本发明描述的多肽的核苷酸序列。即使应用同样的表达系统所有的宿主也不是都能同样完好地行使其功能。然而，本领域中有一种技术可以在无不当试验的情况下对这些载体，表达调控序列和宿主进行选择。例如，在选择载体时，必须考虑到宿主，因为载体必须在宿主中复制或能够整合到染色体中。也应该考虑载体的拷贝数量，其调控拷贝数量的能力，和任何其它的由载体编码的蛋白的表达，例如抗生素标志物。在选择表达调控序列时，也应该考虑许多因素。这些因素包括，例如，序列的相对长度，序列的可控制性，和该调控序列与编码多肽的核苷酸序列的相容性，尤其要注意多肽的核苷酸序列的可能的二级结构。选择宿主时应该考虑宿主与所选载体的相容性，由核苷酸序列编码的产物的毒性，宿主的分泌特性，宿主正确折叠多肽的能力，宿主的发酵或培养条件，和由核苷酸序列编码的产物的纯化的难易。

重组载体可以是自主复制的载体，即这种载体可以作为染色体外的实体存在，其复制独立于染色体的复制，例如质粒。或者，载体也可以是这样一

种载体,当导入到宿主细胞时,其被整合到宿主细胞基因组中且同被整合的染色体一起复制。

载体优选是一种表达载体,在该表达载体中编码本发明的多肽的核苷酸序列被可操作性地连接到核苷酸序列转录所需的另外的区段上。通常,载体来源于质粒或病毒 DNA。本发明中提到的在宿主细胞中表达的许多适宜的
5 表达载体可从商业渠道购买或者在文献中有描述。用于真核宿主的表达载体包括,例如来自 SV40,牛乳突淋巴瘤病毒,腺病毒和细胞巨化病毒的包含表达调控序列的载体。特定的载体是,例如 pCDNA3.1 (+) \Hyg (Invitrogen, Carlsbad, CA, USA) 和 pCI-neo (Stratagene, La Jolla, CA, USA)。用于酵母细
10 胞的表达载体包括 2 μ 质粒和它们的衍生物,POT1(US 4, 931, 373), 在 Okkels, Ann. New York Acad. Sci. 782, 202-207, 1996, 和 pPICZ A, B 或 C (Invitrogen) 中描述了 pJS037 载体。用于昆虫细胞的载体包括 pVL941, pBG311 (Cate 等, "Isolation of the Bovine and Human Genes for Mullerian Inhibiting Substance And Expression of the Human Gene In Animal Cells", 细胞 45, pp. 685-98
15 (1986), pBluebac 4.5 和 pMelbac (两者都可以从 Invitrogen 获得)。用于细菌宿主的表达载体包括已知的细菌质粒,例如来自 E. Coli 的质粒,包括 pBR322, pET3a 和 pET12a (两个都来自 Novagen Inc., WI, USA), 较宽宿主范围的质粒,例如 RP4, 噬菌体 DNA, 例如 λ 噬菌体的许多衍生物,例如 NM989, 和其它的 DNA 噬菌体,例如 M13 和 丝状单链 DNA 噬菌体。

20 本发明中应用的其它载体包括允许编码多肽的核苷酸序列以一定的拷贝数扩增的那些载体。上述可扩增的载体在本领域中已熟知。它们包括,例如能够通过 DHFR 扩增(参见,例如, Kaufman, U. S. Pat. No. 4, 470, 461, Kaufman 和 Sharp, "Construction Of A Modular Dihydrafolate Reductase cDNA Gene: Analysis Of Signals Utilized For Efficient Expression", Mol. Cell. Biol., 2,
25 pp. 1304-19 (1982)) 和谷氨酰胺合成酶扩增("GS")的载体(参见,例如, US 5, 122, 464 and EP 338, 841)。

重组载体可进一步包含能够使所述载体在所述宿主细胞中复制的 DNA 序列。此序列的一个例子(宿主细胞是哺乳动物细胞时)是 SV40 的复制起始点。当宿主细胞是酵母细胞时,能够使载体复制的合适序列是酵母质粒 2 μ
30 复制基因 REP1-3 和复制起始点。

载体也可含有能选择的标记,如基因,其产物补充宿主细胞的缺陷,如

编码二氢叶酸还原酶(DHFR)或粟酒裂殖酵母 TPI 的基因(P. R. Russell, Gene 40, 1985, pp. 125-130), 或者赋予对药物如氨苄青霉素、卡那霉素、四环素、氯霉素、新霉素、潮霉素或氨基蝶呤的抗性的基因。对于丝状真菌来说, 能选择的标记包括 amdS、pyrG、arcB、niaD、sC。

5 术语“控制序列”在本文中包括对本发明多肽表达必需的或有利的所有成分。每个控制序列对编码多肽的核苷酸序列来说可以是天然的或外来的。这些控制序列包括但不限于前导区、聚腺苷酸化序列、前肽序列、启动子、增强子或上游活化序列、信号肽序列和转录终止子。控制序列至少包括启动子。

10 本发明可以使用各种各样的表达控制序列。这些可使用的表达控制序列包括与前述表达载体的结构基因相连的表达控制序列以及已知控制原核或真核细胞或其病毒的基因表达的任何序列及其各种组合。

哺乳动物细胞中与转录相关的合适控制序列的实例包括SV40和腺病毒的早期或晚期启动子, 如腺病毒2的主要晚期启动子、MT-1(金属硫蛋白基
15 因)启动子、人巨细胞病毒立即早期基因启动子(CMV)、人延伸因子1 α (EF-1 α)启动子、果蝇最小热休克蛋白70启动子、劳氏肉瘤病毒(RSV)启动子、人遍在蛋白C(UbC)启动子、人生长激素终止子、SV40或腺病毒E1b区聚腺苷酸化信号和Kozak共有序列(Kozak, M. J Mol Biol 1987 Aug 20; 196(4):947-50)。

为了改善在哺乳动物细胞中的表达, 可以将合成的内含子插入编码所述
20 多肽的核苷酸序列的5'未翻译区。合成内含子的实例是来自质粒pCI-Neo的合成内含子(购自Promega Corporation, WI, USA)。

昆虫细胞中指导转录的合适控制序列的实例包括多角体蛋白启动子、P10启动子、苜蓿银纹夜蛾(*Autographa californica*)多角体病毒碱性蛋白启动子、杆状病毒立即早期基因1启动子和杆状病毒39K延迟早期基因启动子和
25 SV40聚腺苷酸化序列。在酵母宿主细胞中使用的合适控制序列的实例包括酵母 α 交配系统的启动子、酵母丙糖磷酸异构酶(TPI)启动子、来自酵母糖酵解基因或乙醇脱氢酶基因的启动子、ADH2-4c启动子和诱导型GAL启动子。在丝状真菌宿主细胞中使用的合适控制序列的实例包括ADH3启动子和终止子、由编码米曲霉TAKA淀粉酶丙糖磷酸异构酶或碱性蛋白酶、黑曲霉 α
30 淀粉酶、黑曲霉或构巢曲霉(*A. Nidulans*)葡糖淀粉酶、构巢曲霉乙酰胺酶、米赫根毛霉(*Rhizomucor miehei*)天冬氨酸蛋白酶或脂肪酶的基因衍生的启动

子、TPI1 终止子和 ADH3 终止子。在细菌宿主细胞中使用的合适控制序列的实例包括 lac 系统、trp 系统、TAC 或 TRC 系统的启动子和噬菌体 λ 的主要启动子区域。

5 信号肽的存在或不存在如取决于用于多肽生产的表达宿主细胞、被表达的蛋白质(是细胞内的还是细胞外的蛋白质)和是不是需要分泌。为了在丝状真菌中使用,信号肽可以由编码曲霉属菌株的淀粉酶或葡糖淀粉酶的基因、编码米赫根毛霉脂肪酶或蛋白酶或 *Humicola lanuginosa* 脂肪酶的基因方便衍生。信号肽优选由编码米曲霉 TAKA 淀粉酶、黑曲霉中性 α -淀粉酶、黑曲霉酸稳定淀粉酶或黑曲霉葡糖淀粉酶的基因衍生。为了在昆虫细胞中使用,10 信号肽可以由昆虫基因方便地衍生(参见 WO90/05783),如鳞翅目 *Manduca sexta* 脂动激素前体(US5023328)、蜜蜂蜂毒素(Invitrogen)、蛻皮甾类 UDP 葡萄糖基转移酶(glucosyltransferase)(egt) (Murphy 等人, Protein Expression and Purification 4, 349-357(1993))或人胰脂酶(hpl)(Methods in Enzymology 284, pp. 262-272, 1997)。哺乳动物细胞中使用的优选信号肽是下文实施例所示人干扰素 β 的或鼠 Ig κ 轻链信号肽(Coloma, M(1992) J. Imm. Methods 152:89-104)。为了在酵母细胞中使用,发现合适的信号肽是酿酒酵母 α -因子信号肽(参见 US4870008)、改性的羧肽酶信号肽(参见 L. A. Valls 等人, Cell 48, 1987, pp.887-897)、酵母 BAR1 信号肽(参见 WO87/02670)和酵母天冬氨酸蛋白酶 3(YAP3)信号肽(参见 M. Egel-Mitani 等人, Yeast 6, 1990, pp. 127-137)和合成的前导序列 TA57(WO98/32867)。已发现适宜于大肠杆菌的信号肽是 ompA。

编码显示 G-CSF 活性的多肽的本发明核苷酸序列,无论是通过定点诱变、合成还是其它方法制备的,都可以包括也可以不包括编码信号肽的核苷酸序列。多肽从表达该多肽的细胞中分泌时,存在信号肽。如果存在,此信号肽应当由选来用于表达多肽的细胞识别。信号肽可以与多肽同源(如,通常25 与 G-CSF 相连)或异源(即,来自于除 hG-CSF 之外的另一来源)或者可以与宿主细胞同源或异源,即是由宿主细胞正常表达的信号肽或是由宿主细胞非正常表达的。因此,信号肽可以是原核的如由细菌如大肠杆菌衍生的,或真核的,如由哺乳动物或昆虫或酵母细胞衍生的。

可以使用任何合适的宿主产生本发明偶联物的多肽或多肽部分,包括细菌、真菌(包括酵母)、植物、昆虫、哺乳动物或其它合适的动物细胞或细胞系以及转基因动物或植物。细菌宿主细胞的实例包括革兰氏阳性细菌如芽孢30

杆菌属,如短芽孢杆菌或枯草芽孢杆菌、假单孢菌属或链霉菌属或革兰氏阴性细菌,如大肠杆菌菌株。载体可引入细菌宿主细胞,例如可通过原生质体转化(参见如 Chang and Cohen, 1979, Molecular General Genetics 168:111-115),使用感受态细胞(参见如 Young and Spizizin, 1961, Journal of Bacteriology 81:823-829, 或 Dubnau and Davidoff-Abelson, 1971, Journal of Molecular Biology 56:209-221)、电穿孔(参见如 Shigekawa and Dower, 1988, Biotechniques 6:742-751)或偶联(参见如 Koehler and Thorne, 1987, Journal of Bacteriology 169:5771-5278)来进行。合适的丝状真菌宿主细胞的实例包括曲霉属菌株,如米曲霉、黑曲霉或构巢曲霉,镰刀菌属或木霉菌属。真菌细胞可以转化,转化过程涉及原生质体形成、原生质体转化和细胞壁通过本身已知的方式再生。在 EP238023 和 US5679543 中描述了转化曲霉属宿主细胞的合适方法。在 Malardier 等人 1989, Gene 78:147-156 和 WO96/00787 中描述了转化镰刀菌属的合适方法。合适的酵母宿主细胞的实例包括糖酵母属的菌株,如酿酒酵母属、裂殖酵母属、克鲁维酵母属、毕赤酵母属,如巴斯德毕赤酵母或 *P. Methanolica*、汉逊酵母属,如多形汉逊酵母或 *Yarrowia*。酵母可使用 Becker 和 Guarente, In Abelson, J. N. And Simon, M. I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology Volume 194, pp 182-187, Academic Press, Inc., New York; Ito 等人, 1983, Journal of Bacteriology 153:163; 和 Hinnen 等人 1978, Proceedings of the National Academy of Sciences USA 75:1920 中所述的方法和 Clontech Laboratories, Inc, Palo Alto, CA, USA(在 Yeastmaker™ Yeast Transformation System Kit 的产品使用说明书)公开的来转化。合适昆虫宿主细胞的实例包括鳞翅目细胞系,如草地夜蛾(Sf9 或 Sf21)或粉纹夜蛾细胞(High Five)(US5077214)。可以按 Invitrogen 所述方法转化昆虫细胞并在其中产生异源多肽。合适哺乳动物宿主细胞的实例包括中国仓鼠卵巢(CHO)细胞系(如, CHO-K1; ATCC CCL-61)、绿猴细胞系(COS)(如, COS1(ATCC CRL-1650)、COS7(ATCC CRL-1651)); 小鼠细胞(如, NS/O)、仓鼠幼鼠肾脏(BHK)细胞系(如, ATCC CRL-1632 或 ATCC CCL-10)和人细胞(如, HEK293(ATCC CRL-1573)),以及组织培养中的植物细胞。另外合适的细胞系在本领域中是已知的并可购自公开的保藏中心如美国典型培养物保藏中心, Rockville, Maryland。将外源性 DNA 引入哺乳动物宿主细胞的方法包括磷酸钙介导的转染、电穿孔、DEAE-葡聚糖介导

的转染、脂质体介导的转染、病毒载体和由 Life Technologies Ltd, Paisley, UK 所述的使用 Lipofectamin2000 的转染方法。这些方法在本领域中是已知的, 例如在 Ausbel 等人(eds.), 1996, Current Protocols in Molecular Biology, John Wiley & Sons, New York, USA 中所述的。按建立的方法进行哺乳动物细胞的

5 培养, 例如在 Animal Cell Biotechnology, Methods and Protocols, Nigel Jenkins 编, 1999, Human Press Inc, Totowa, New Jersey, USA and Harrison MA and Rae IF, General Techniques of Cell Culture, Cambridge University Press 1997 中的公开的。

在本发明生产方法中, 细胞用本领域已知的方法在适于生产多肽的营养

10 培养基中培养。例如, 细胞可以在实验室中通过振荡烧瓶培养、小规模或大规模发酵(包括连续、成批、补料分批或固体状态发酵)或在合适培养基和允许表达和/或分离所述多肽的条件下进行工业发酵来发酵。培养在含有碳和氮源和无机盐的合适营养培养基中使用本领域已知的方法来进行。合适的培养基可商购或可以按公开的组成来制备(例如, 在美国典型培养物保藏中心的

15 目录中所述的)。如果多肽分泌在营养培养基中, 多肽可从培养基中直接回收。如果多肽不被分泌, 可从细胞裂解物中回收。

可通过本领域已知的方法来回收所得多肽。例如, 可通过常规方法包括但不限于离心、过滤、提取、喷雾干燥、蒸发或沉淀从营养培养基中回收多肽。

20 可通过本领域已知的各种方法包括但不限于层析(如, 离子交换、亲和、疏水、层析聚焦和大小排阻)、电泳方法(如, 制备性等电聚焦)、溶解度差异(如, 硫酸铵沉淀)、SDS-PAGE 或提取(参见, 如 Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989)来纯化多肽。在下列文献中公开了纯化具有 G-CSF 活性的多肽的具体方法由 D. Metcalf 和 N. A.

25 Nicola 在 The hemopoietic colony-stimulating factors, p.50-51, Cambridge University Press (1995)中, 由 C. S. Bae 等, 在 Appl. Microbiol. Biotechnol, 52: 338-344 (1999)和在 US 4, 810, 643 中描述。

本发明的药物组合物和其应用

本发明进一步包含本发明描述的多肽或偶联物的组合物和至少一种可

30 药用载体或赋形剂。

用本发明的多肽, 偶联物或药物组合物可以制备治疗某些疾病的药物,

尤其用于预防正接受某些类型的化疗,放疗和骨髓移植治疗的癌症患者的感染,在外周血前体细胞移植时用于动员前体细胞汇聚,用于严重的慢性或相对的白细胞减少症患者的治疗,用于急性髓细胞性白血病患者治疗,用于AIDS病或其它的免疫缺陷病的治疗,和抗真菌治疗,尤其用于对全身性或侵袭性念珠菌病的治疗。

另一方面,依照本发明的多肽,偶联物或药物组合物应用于治疗患广泛的造血疾病,包括放疗或化疗引发的那些造血疾病,尤其是白细胞减少症,AIDS或其它的免疫缺陷病的患者,该方法包含给予患病的哺乳动物上述多肽,偶联物或药物组合物。

10 将本发明的多肽和偶联物以“治疗有效的”剂量,即足以对给药治疗的疾病产生所期望的效果的剂量,给予患者。准确的剂量将根据所要治疗的疾病决定,且可通过本领域技术人员用已知的技术确定。本发明的多肽或偶联物可以例如用类似于rhG-CSF,例如Neupogen®的治疗剂量给药。本发明偶联物的适当剂量预计是在大约5-300 μ g/kg体重的范围内(根据偶联物的蛋白部分的重量),10-200 μ g/kg,例如25-100 μ g/kg。本发明的多肽,偶联物或组合物的有效量尤其依赖于,疾病,剂量和给药的时间表,多肽或偶联物或组合物是否单独给药或与其它的治疗制剂结合给药,组合物的血清半寿期,患者的一般健康状况,和给药的频率,这些对本领域的技术人员来说是显而易见的。优选,本发明的多肽,偶联物,制品或组合物以有效量给予,尤其所给剂量足以使所涉及的患者体内白细胞,尤其是嗜中性粒细胞的数量达到正常。依照惯例通过每隔一定时间间隔简单地计数白细胞来确定白细胞数量是否已达到正常。

25 本发明的多肽或偶联物优选以包含一个或多个可药用载体或赋形剂的组合物形式给药。可把多肽或偶联物以本领域人人尽知的方式配制成药物组合物以产生储存十分稳定且适合对人或动物给药的多肽药物。可以将药物组合物以多种形式配制,包括液态或凝胶体,或冻干的或任何其它的形式。优选形式取决于所要治疗的具体指征,这一点对本领域的技术人员来说是显而易见的。

30 因此,本发明提供治疗各种形式的白细胞减少症的组合物和方法。尤其本发明的多肽,偶联物或组合物可用来预防正接受某些类型的放疗,化疗和骨髓移植的癌症患者的感染,在外周血前体细胞移植时用于动员前体细胞汇

聚,用于治疗严重的慢性或相对的白细胞减少症和用于患急性髓细胞性白血病患者支持治疗。另外,本发明的多肽,偶联物或组合物可用于治疗 AIDS 或其它的免疫缺陷病,用于抗真菌治疗,尤其用于治疗全身性或侵袭性念珠菌病,和用于治疗细菌感染。

5 *药物的形式*

本发明的多肽或偶联物可以“原型”和/或以它们的盐形式应用。适当的盐包括,但不限于,碱金属盐或碱土金属盐,例如钠,钾和镁盐,和例如锌盐。这些盐和复合物可以作为一种水晶型和/或无定形的结构存在。

赋形剂

- 10 “可药用的”是指一种所使用的剂量和浓度在接受药物的患者体内不产生任何不良作用的载体或赋形剂。上述可药用的载体和赋形剂在本领域是人所共知的(参见 Remington's Pharmaceutical Sciences, 18th edition, A. R. Gennaro, Ed., Mack Publishing Company [1990]; Pharmaceutical Formulation Development of Peptides and Proteins, S.Frokjaer 和 L. Hovgaard, Eds., Taylor & Francis [2000]; 和 Handbook of Pharmaceutical Excipients, 3rd edition, A. Kibbe, Ed., Pharmaceutical Press [2000])。

药物的混合

- 20 本发明的药物组合物可以单独给药或与其它的治疗剂结合给药。可以将这些制剂作为同一药物组合物的一部分给药,或与本发明的多肽或偶联物分开,同时给药或按照另一治疗时间表给药。另外,本发明的多肽或偶联物或药物组合物可以用作对其它治疗的一种辅助治疗。

患者

本发明中的“患者”既包括人也包括其它哺乳动物。所以一些对人治疗和对牲畜使用的方法都是可应用的。

25 *给药途径*

本发明制剂可以多种方式给药,包括但不限于,经口服、皮下、静脉内、脑内、鼻内、真皮内、腹膜内、肌肉内、肺内、阴道内、直肠内、眼球内或以任何其它可接受的方式给药。可以连续地灌输给药,尽管大药丸注射剂是可接受的,但应用技术在本领域是熟知的。

30 *非胃肠道给药剂(Parentals)*

药物组合物的实例是设计成非肠道给药的溶液剂。尽管在很多情况下,

药物溶液剂可以以适用于立即使用的液体制剂形式提供,但所述非肠道制剂也可以以冷冻或冻干形式提供。在前者情况下,所述组合物在使用前必须解冻。后一剂型通常用于增强组合物中所包含的活性组分在各种贮存条件下的稳定性,正如本领域技术人员已知的那样,冻干制剂通常比其液体相应物更稳定。所述冻干制剂在使用前通过加入一种或多种适宜的可药用稀释剂如灭菌注射用水或灭菌生理盐水溶液重新配制。

在非胃肠道给药的情况下,制成冻干制剂或水溶液备用,如通过将具有所需纯度的多肽与一种或多种本领域常用的可药用载体、赋形剂或稳定剂(统称为“赋形剂”)适当混合来制备,赋形剂如缓冲剂、稳定剂、防腐剂、等渗剂、非离子去污剂、抗氧化剂和/或其它各种添加剂。

缓冲剂有助于将 pH 保持在接近生理条件的范围中。它们通常以大约 2mM-50mM 的浓度范围存在。本发明使用的合适缓冲剂包括有机和无机酸及其盐如柠檬酸盐缓冲剂(如,柠檬酸一钠-柠檬酸二钠混合物、柠檬酸-柠檬酸三钠混合物、柠檬酸-柠檬酸一钠混合物等)、琥珀酸盐缓冲剂(如,琥珀酸-琥珀酸一钠混合物、琥珀酸-氢氧化钠混合物、琥珀酸-琥珀酸二钠混合物等)、酒石酸盐缓冲剂(如,酒石酸-酒石酸钠混合物、酒石酸-酒石酸钾混合物、酒石酸-氢氧化钠混合物等)、富马酸盐缓冲剂(如,富马酸-富马酸一钠混合物、富马酸-富马酸二钠混合物、富马酸一钠-富马酸二钠混合物等)、葡萄糖酸(gluconate)盐缓冲剂(如,葡萄糖酸-葡萄糖酸钠混合物、葡萄糖酸-氢氧化钠混合物、葡萄糖酸-葡萄糖酸钾混合物等)、草酸盐缓冲剂(如,草酸-草酸钠混合物、草酸-氢氧化钠混合物、草酸-草酸钾混合物等)、乳酸盐缓冲剂(如,乳酸-乳酸钠混合物、乳酸-氢氧化钠混合物、乳酸-乳酸钾混合物等)和乙酸盐缓冲剂(如,乙酸-乙酸钠混合物、乙酸-氢氧化钠混合物等)。其它可能的缓冲剂是磷酸盐缓冲剂、组氨酸缓冲剂和三甲胺盐如 Tris。

加入防腐剂来阻滞微生物生长,加入的量通常约为 0.2%-1%(w/v)。本发明使用的合适防腐剂包括酚、苯甲醇、间甲酚、羟苯甲酸(paraben)甲酯、羟苯甲酸丙酯、十八烷基二甲基苯甲基铵氯化物、苯亚甲基 卤化物(如苯亚甲基 氯化物、溴化物或碘化物)、氯化己烷双胺、羟苯甲酸烷基酯如甲酯或丙酯、儿茶酚、间苯二酚、环己醇和 3-戊醇。

可以加入等渗剂来确保液体组合物的等渗性,等渗剂包括多羟糖醇,优选三羟或更高级糖醇,如甘油、赤藓醇、阿拉伯糖醇、木糖醇、山梨糖醇和

甘露糖醇。多羟基醇的量可为 0.1%-25%重量比,通常为 1%-5%,以其它组分的相对量计算。

- 稳定剂涉及一大类赋形剂,其功能范围从膨胀剂到溶解治疗剂的或有助于防止变性或与容器壁粘合的添加剂。通常的稳定剂可以是多羟基糖醇(上文列举的);氨基酸如精氨酸、赖氨酸、甘氨酸、谷氨酰胺、天冬酰胺、组氨酸、丙氨酸、omithine、L-亮氨酸、2-苯丙氨酸、谷氨酸、苏氨酸等,有机糖或糖醇,如乳糖、海藻糖、水苏糖、甘露糖醇、山梨糖醇、木糖醇、核糖醇、肌醇、半乳糖醇、甘油等,包括环醇如环己六醇;聚乙二醇;氨基酸聚合物;含硫还原剂,如脲,谷胱甘肽、硫辛酸、巯基乙酸钠、硫代甘油、 α -单硫代甘油和硫代硫酸钠;低分子量多肽(即<10 个残基);蛋白质如人血清白蛋白、牛血清白蛋白、明胶或免疫球蛋白;亲水性聚合物如聚乙烯吡咯烷酮;单糖如木糖、甘露糖、果糖和葡萄糖;二糖如乳糖、麦芽糖和蔗糖;三糖如棉子糖,和多糖如葡聚糖。基于活性蛋白质重量计算,稳定剂通常的存在范围为 0.1-10000 重量份。

- 15 可以存在非离子表面活性剂或清洁剂(也称作“湿润剂”)以有助于溶解治疗剂以及保护治疗性多肽来避免搅拌诱导的聚集,其也允许配方剂暴露于剪切表面压力而没有引起多肽变性。合适的非离子表面活性剂包括聚山梨酸酯(polysorbate)(20, 80 等)、polyoxamer(184, 188 等)、Pluronic® 多元醇、聚氧乙烯脱水山梨糖醇单醚(吐温 20、吐温 80 等)。

- 20 其它各种赋形剂包括膨胀剂或填充剂(如淀粉)、螯合剂(如 EDTA)、抗氧化剂(如抗坏血酸、蛋氨酸、维生素 E)和共溶剂。

- 25 活性组分也可以被包裹在制备的微囊中,例如通过 coascervation 技术或通过界面聚合制备的,例如羟甲基纤维素、明胶或聚(甲基丙烯酸酯)微囊,包裹在胶体状药物释放体系(例如脂质体、白蛋白微球、微乳、纳米颗粒和纳米胶囊)中或包裹在大乳剂中。在 Remington's Pharmaceutical Sciences(同上)中公开了这些技术。

体内给药所使用的非胃肠道制剂必须是灭菌的。该过程容易完成,例如,通过无菌滤膜来过滤。

持续释放制剂

- 30 持续释放制剂的合适例子包括含有多肽或偶联物的固体疏水聚合物半渗透材料、具有合适形式如膜或微囊的材料。持续释放材料的例子包括聚酯、

水凝胶(例如, 聚(2-羟基乙基甲丙烯酸酯)或聚(乙烯醇))、聚交酯、L-谷氨酸和 L-谷氨酸乙酯的共聚物、不可降解的乙烯乙酸乙酯、可降解的乳酸-乙醇酸共聚物如 ProLease®技术或 Lupron Depot®(由乳酸-乙醇酸共聚物和醋酸亮丙瑞林构成的可注射微球)和聚-D-(-)-3-羟基丁酸。聚合物如乙烯乙酸乙酯和乳酸-乙醇酸能够长时间释放分子如长达或超过 100 天, 某些水凝胶在较短时间内释放蛋白质。包囊中的多肽长时间保留在体内时, 它们因在 37℃ 潮湿的环境中暴露而变性或聚集, 导致丧失生物学活性并且可能改变免疫原性。合理的稳定策略可根据所涉及的机理设计。例如, 如果发现聚集机理是通过硫代二硫化物相互交换形成分子内 S-S 键, 那么可通过修饰巯基、冻干酸性溶液、控制含湿量、使用适宜的添加剂和开发特异性聚合物型组合物来达到稳定。

肺释放

适于与雾化剂(喷射或超声雾化剂)一起使用的偶联物制剂通常包含以每 ml 溶液大约 0.01-25mg 偶联物, 优选大约 0.1-10mg/ml 的浓度溶于水中的偶联物。所述制剂也可以包含缓冲剂和简单的糖(例如用于稳定蛋白质和调节渗透压), 和/或 0.1-10mg/ml 浓度范围的人血清白蛋白。可使用的缓冲剂的实例为醋酸钠、枸橼酸钠和甘氨酸。优选地, 所述缓冲剂将含有适用于调节溶液的 pH 在 3-9 范围的组成和克分子浓度。通常, 1mM-50mM 的缓冲剂克分子浓度适用于该目的。可使用的糖的实例为乳糖、麦芽糖、甘露糖醇、山梨糖醇、海藻糖和木糖, 通常用量为所述制剂重量的 1%-10% 范围。

所述雾化制剂也可以包含表面活性剂用于减少或防止在形成气雾剂时溶液雾化引起的蛋白质的表面诱导性聚集。可使用各种常规的表面活性剂, 如聚氧乙烯脂肪酸酯和醇, 和聚氧化乙烯山梨糖醇脂肪酸酯。用量通常在所述配方剂重量的 0.001-4% 范围。本发明特别优选的表面活性剂是聚氧乙烯山梨糖醇单油酸酯。

产生适宜的本发明液体微粒分散体的特定制剂和方法描述于 WO 9420069、US 5915378、US 5960792、US5957124、US 5934272、US 5915378、US 5855564、US 5826570 和 US 5522385, 所有文献引入本文供参考。

用于与可计量剂量的吸入装置一起使用的偶联物制剂通常包括细粉。该粉剂可通过冻干, 然后研磨液体偶联物制剂制备并且也可以包含稳定剂如人血清白蛋白(HAS)。通常, 加入超过 0.5%(w/w)的 HAS。因此, 如果需要,

可将一种或多种糖或糖醇加到所述制剂中。其实例包括乳糖、麦芽糖、甘露糖醇、山梨糖醇、山梨糖、海藻糖、木糖醇和木糖，制剂中所加入的量可在本发明偶联物的大约 0.01 %-200 % (w/w) 范围，优选大约为 1-50 %。然后，将所述制剂冻干并研磨至所需粒度。

- 5 然后，将适宜大小的微粒在表面活性剂的辅助下悬浮在推进剂中。所述推进剂可以是适用于本发明目的的任何常规物质，如含氯氟烃、氢氯氟烃、氢氟烃或烃，包括三氯氟甲烷、二氯氟甲烷、二氯四氟乙醇和 1, 1, 1, 2-四氟乙烷或其组合。适宜的表面活性剂包括三油酸山梨糖醇酯和大豆卵磷脂。油酸也可以用作表面活性剂。然后，将混合物装入释放装置中。适用于本发
- 10 的市售可计量剂量的吸入器的实例为由 Glaxo Inc., Research Triangle Park, N. C. 生产的 Ventolin 可计量剂量吸入器。

所述与粉末吸入器一起使用的偶联物制剂将包括含有偶联物的极细干粉末并且也可以包含膨胀剂，如乳糖、山梨糖醇、蔗糖或甘露糖醇，其促进粉末从装置中分散的量，例如为所述制剂重量的 50%-90%。所述粉末微粒

15 在肺中应该具有与直径中度值小于 10 微米，优选在 0.5-5 微米之间，最优选在 1.5-3.5 微米之间，密度大约为 1g/cm^2 的微粒相当的气动性。适用于本说明书的粉末吸入器的实例为 Fisons Corp., Bedford, Mass. 生产的旋转吸入型 (Spinhaler) 粉末吸入器。

应用于这些装置的粉末可以按 US 5, 997, 848, US 5, 993, 783, US 5, 985, 248, US 5, 976574, US 5, 922, 354, US 5, 785, 049 和 US 5, 654, 007 中公开的方法制备和/或传输。

20 应用于这些装置的粉末可以按 US 5, 997, 848, US 5, 993, 783, US 5, 985, 248, US 5, 976574, US 5, 922, 354, US 5, 785, 049 和 US 5, 654, 007 中公开的方法制备和/或传输。

为肺部传输治疗产品设计的机械装置包括但不限于雾化器、可计量剂量吸入器和粉末吸入器，所有这些都是本领域技术人员熟悉的。一些适用于本发

25 的市售装置的特定实例为由 Mallinckrodt, Inc., St. Louis, Missouri 生产的 Ultravent 雾化器，由 Marquest Medical Products, Englewood, Colorado 生产的 Acorn II 雾化器，由 Glaxo Inc., Research Triangle Park, North Carolina 生产的 Ventolin 可计量剂量吸入器；由 Fisons Corp., Bedford, Massachusetts 制造的旋转吸入型粉末吸入器；“持续雾化”装置，Inhale Therapeutic System, Inc., San Carlos, California；由 Alkermes, Cambridge, Massachusetts 制造的 AIR 吸

30 入器；和由 Aradigm Corporation, Hayward, California 制造的 AERx 肺部药物释放系统。

也关注编码本发明多肽的核苷酸序列在基因治疗中的应用。特别感兴趣的是按上文“本发明偶联物中的非多肽组分是糖”章节中所述方法使用编码多肽的核苷酸序列。因此，在基因治疗过程中即核苷酸序列在人体内表达后完成多肽的糖基化。

5 关注的基因治疗的应用包括对那些期待多肽能提供有效治疗的疾病的治疗。所述治疗包括对各种类型的白细胞减少症的治疗，尤其用于预防正接受某些类型的化疗和骨髓移植，外周前体细胞移植时对前体细胞进行动员，汇聚治疗的患者的感染，包括对严重的慢性或相对的白细胞减少症的治疗，对急性髓细胞性白血病患者治疗，对 AIDS 或其它免疫缺陷病的治疗。

10 基因治疗的 G-CSF 的局部传输可使治疗剂到达靶部位而避免了与非特异性给药有关的潜在毒性问题。

本发明还涉及体外和体内基因治疗方法学。

将潜在的治疗基因转移到特定细胞群中的一些方法是已知的。可进一步参见如，Mulligan, “The Basic Science Of Gene Therapy”, Science, 260, 15 pp.926-31(1993)。这些方法包括：

直接的基因转移，如 Wolff 等人在 “Direct Gene Transfer Into Mouse Muscle In vivo”, Science 247, pp.1465-68(1990)中公开的；

脂质体介导的 DNA 转移，如，Caplen 等人在 “Liposome-mediated CFTR Gene Transfer to the Nasal Epithelium Of Patients With Cystic Fibrosis” Nature 20 Med., 3, pp.39-46(1995)中；Crystal 在 “The Gene As A Drug”, Nature Med., 1, pp. 15-17(1995)中和 Gao and Huang 在 “A Novel Cationic Liposome Reagent For Efficient Transfection of Mammalian Cells”, Biochem. Biophys Res. Comm., 179, pp.280-85(1991)中公开的；

25 逆转录病毒介导的 DNA 转移，如，Kay 等人在 “In vivo Gene Therapy of Hemophilia B: Sustained Partial Correction In Factor IX-Deficient Dogs”, Science, 262, pp.117-19(1993)中，Anderson 在 “Human Gene Therapy”, Science, 256, pp.808-13(1992)中公开的；

30 DNA 病毒介导的 DNA 转移。这些 DNA 病毒包括腺病毒(优选基于 Ad-2 或 Ad-5 的载体)、疱疹病毒(优选基于单纯疱疹病毒的载体)和微小病毒(优选基于 “缺损” 或非自主微小病毒的载体，更优选基于腺伴随病毒的载体，最优选基于 AAV-2 的载体)。如参见，Ali 等人 “The Use Of DNA Viruses as

Vectors for Gene Therapy” , Gene Therapy, 1, pp.367-84(1994); US4797368 和 US5139941。

下面在非限制性的实施例中对本发明进一步描述。

5

附图说明

图 1: rhG-CSF (Neupogen®) 和 SPA-PEG 5000-偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 的体内半寿期。

图 2: rhG-CSF (Neupogen®), SPA-PEG 5000-偶联的 hG-CSF K16R K34R K40R Q70K Q120K 和 SPA-PEG 5000-偶联的 hG-CSF K16R K34R K40R
10 Q70K Q90K Q120K 的体内生物学活性。

图 3: rhG-CSF (Neupogen®), SPA-PEG 12000 偶联的 hG-CSF K16R K34R K40R 和不同剂量的 SPA-PEG 5000-偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 的体内生物学活性。

15

序列表

所附序列表包含以下序列:

SEQ ID NO : 1:人 G-CSF 的氨基酸序列。

SEQ ID NO : 2:编码人 G-CSF 的合成的 DNA 序列, 具有利于在大肠杆菌中最佳表达的密码子使用特性。

20

SEQ ID NO : 3:OmpA 信号序列的氨基酸序列。

SEQ ID NO : 4: 编码 OmpA 信号序列的合成的 DNA 序列。

SEQ ID NO : 5: 合成的组氨酸标记物。

SEQ ID NO : 6: 编码 SEQ ID NO : 5 所示组氨酸标记物的合成的 DNA 序列。

25

SEQ ID NO : 7: 人 G-CSF 信号肽的氨基酸序列。

SEQ ID NO : 8: 编码人 G-CSF, 包括 SEQ ID NO : 7 所示信号肽的, 合成的 DNA 序列, 其具有适于在 CHO 细胞中最佳表达的密码子使用特性。

材料与方法

30

确定所要修饰的氨基酸的方法

可接近的表面区域(ASA)

通过 NMR 波谱仪确定的 10 种结构的 3D 全貌(Zink et al., (1994) Biochemistry33: 8453-8463)可从蛋白数据库(Protein Data Bank)(PDB) (www.rcsb.org/pdb/)中获得。可将该信息输入计算机程序 Access(B. Lee 和 F. M. Richards, J. Mol. Biol. 55: 379-400(1971))版本 2(Copyright© 1983 Yale University)并用于计算结构中各个原子的可接近的表面区域(ASA)。该方法通常使用 1.4Å 大小的探针并将可接近的表面区域(ASA)定义为探针中心形成的面积。于该计算前,从坐标中去除所有水分子和所有氢原子,以及其它不与该蛋白质直接相连的原子。

侧链的分部(fractional)ASA

10 侧链原子的分部 ASA 通过侧链中原子的 ASA 总和除以延长的 ALA-x-ALA 三肽中该残基类型侧链原子的 ASA 代表值来计算。参见 Hubbard, Campbell & Thornton(1991) J. Mol. Biol. 220, 507-530。在该例中, CA 原子被看作甘氨酸残基侧链的一部分但不被看作其它残基的一部分。下表表示侧链的 100% ASA 标准:

15	Ala 69.23 Å ²	Leu 140.76 Å ²
	Arg 200.35 Å ²	Lys 162.50 Å ²
	Asn 106.25 Å ²	Met 156.08 Å ²
	Asp 102.06 Å ²	Phe 163.90 Å ²
	Cys 96.69 Å ²	Pro 119.65 Å ²
20	Gln 140.58 Å ²	Ser 78.16 Å ²
	Glu 134.61 Å ²	Thr 101.67 Å ²
	Gly 32.28 Å ²	Trp 210.89 Å ²
	His 147.00 Å ²	Tyr 176.61 Å ²
	Ile 137.91 Å ²	Val 114.14 Å ²

25 该结构中未探测到的残基被定义为 100%暴露,因为可以认为这些残基位于弹性区域。

确定原子之间的距离

用分子制图软件,例如 InsightII v. 98.0, MSI INC 可以非常容易地确定原子之间的距离。

30 对所要修饰的残基的总体考虑事项

如上所述,依照本发明需要修饰的氨基酸残基优选是下述氨基酸残基:

其侧链表面是暴露的,尤其其侧链的 25%以上暴露在分子表面,更优选 50% 以上的侧链暴露在分子表面。另一种考虑是优选将位于受体接触面上的残基排除以免其对受体的结合或活化可能产生干扰或至少使干扰降低到最小。进一步考虑的是应该将距离最近的 lys(Glu, Asp) CB-CB (对 Gly 而言为 CA)不足 10Å 的残基排除。最后,优选的修饰部位具体是具有亲水和/或带电残基,即 Asp, Asn, Glu, Gln, Arg, His, Tyr, Ser 和 Thr 的那些,具有精氨酸残基的部位尤其优选。

对进行修饰的 G-CSF 氨基酸残基的确定

以下例举了确定本发明所要修饰的氨基酸残基时一般应该考虑的因素。

10 对应用 X 线晶体学(Hill et al., (1993) Proc. Natl. Acad. Sci. USA 90: 5167-5171)和 NMR 波谱学方法(Zink et al., (1994) Biochemistry33: 8453-8463)得到的人 G-CSF 三维结构已有报道。如上所述, Aritomi 等(nature 401: 713-717, 1999)已确定以下 hG-CSF 残基为受体结合界面的一部分: G4, P5, A6, S7, S8, L9, P10, Q11, 512, L15, K16, E19, Q20, L108, D109, D112, T115, T116, 15 Q119, E122, E123, 和 L124。因此, 尽管可以对这些残基进行修饰, 但优选不修饰这些残基。

通过将 Zink 等(1994)确定的 G-CSF 的 10 种 NMR 结构用作输入结构, 然后计算侧链的平均 ASA, 而将以下残基确定为具有 25%以上的 ASA: M0, T1, P2, L3, G4, P5, A6, S7, S8, L9, P10, Q11, S12, F13, L14, L15, K16, C17, E19, 20 Q20, V21, R22, K23, Q25, G26, D27, A29, A30, E33, K34, C36, A37, T38, Y39, K40, L41, H43, P44, E45, E46, V48, L49, L50, H52, S53, L54, I56, P57, P60, L61, S62, S63, P65, S66, Q67, A68, L69, Q70, L71, A72, G73, C74, S76, Q77, L78, S80, F83, Q86, G87, Q90, E93, G94, S96, P97, E98, L99, G100, P101, T102, D104, T105, Q107, L108, D109, A111, D112, F113, T115, T116, W118, 25 Q119, Q120, M121, E122, E123, L124, M126, A127, P128, A129, L130, Q131, P132, T133, Q134, G135, A136, M137, P138, A139, A141, S142, A143, F144, Q145, R146, R147, S155, H156, Q158, S159, L161, E162, V163, S164, Y165, R166, V167, L168, R169, H170, L171, A172, Q173, P174。

同样, 以下残基具有 50%以上 ASA: M0, T1, P2, L3, G4, P5, A6, S7, S8, 30 L9, P10, Q11, S12, F13, L14, L15, K16, C17, E19, Q20, R22, K23, G26, D27, A30, E33, K34, T38, K40, L41, H43, P44, E45, E46, L49, L50, S53, P57, P60,

L61, S62, S63, P65, S66, Q67, A68, L69, Q70, L71, A72, G73, S80, F83, Q90, G94, P97, E98, P101, D104, T105, L108, D112, F113, T115, T116, Q119, Q120, E122, E123, L124, M126, P128, A129, L130, Q131, P132, T133, Q134, G135, A136, A139, A141, S142, A143, F144, R147, S155, S159, E162, R166, V167,
5 R169, H170, L171, A172, Q173, P174.

用分子制图程序 InsightII® v. 98.0 确定以下残基, 它们具有距离最近的氨基 15Å 以上的 CB 原子(在甘氨酸的情况下是 CA), 这样的原子被确定为赖氨酸的 NZ 原子和 N 末端残基 T1 的 N 原子。以下所列包括在 10 种 NMR 结构的至少一种中满足这一标准的残基。G4, P5, A6, S7, S8, L9, P10, Q11, L14,
10 L15, L18, V21, R22, Q25, G26, D27, G28, A29, Q32, L35, C36, T38, Y39, C42, H43, P44, E45, E46, L47, V48, L49, L50, G51, H52, S53, L54, G55, I56, P57, W58, A59, P60, L61, S62, S63, C64, P65, S66, Q67, A68, L69, Q70, L71, A72, G73, C74, L75, S76, Q77, L78, H79, S80, G81, L82, F83, L84, Y85, Q86, G87, L88, L89, Q90, A91, L92, E93, G94, I95, S96, P97, E98, L99, G100, P101, T102,
15 L103, D104, T105, L106, Q107, L108, D109, V110, A111, D112, F113, A114, T115, T116, I117, W118, Q119, Q120, M121, E122, E123, L124, G125, M126, A127, P128, A129, L130, Q131, P132, T133, Q134, G135, A136, M137, P138, A139, F140, A141, S142, A143, P144, Q145, R146, R147, A148, G149, G150, V151, L152, V153, A154, S155, H156, L157, Q158, S159, F160, L161, E162,
20 V163, S164, Y165, R166, V167, L168, R169, H170, L171, A172, Q173, P174.

分子制图程序 InsightII® v. 98.0 也用于确定以下残基, 它们具有距离最近的氨基 10Å 以上的 CB 原子(在甘氨酸的情况下是 CA), 这样的原子被确定为天冬氨酸的 CG 原子, 谷氨酸的 CD 原子和 C 末端残基 P174 的 C 原子。以下所列包括在 10 种 NMR 结构的至少一种中满足这一标准的残基。M0, T1,
25 P2, L3, G4, P5, A6, S7, S8, L9, P10, Q11, S12, P13, L14, T38, Y39, K40, M1, C42, L50, G51, H52, S53, L54, G55, I56, P57, W58, A59, P60, L61, S62, S63, C64, P65, S66, Q67, A68, L69, Q70, L71, A72, G73, C74, L75, S76, Q77, L78, H79, S80, G81, L82, F83, L84, Y85, Q86, G87, L88, I117, M126, A127, P128, A129, L130, Q131, P132, T133, Q134, G135, A136, M137, P138, A139, F140, A141,
30 S142, A143, F144, Q145, R146, R147, A148, G149, G150, V151, L152, V153, A154, S155, H156, L157, V167, L168, R169, H170, L171.

综合以上所列并经比较发现,有可能选择单个氨基酸残基进行修饰,从而产生一组的有限数目的氨基酸残基,它们在特定 G-CSF 多肽中的修饰很可能产生所期望的特性。

hG-CSF 和其变体的 PEG 化方法

5 hG-CSF 和其变体在溶液中 PEG 化

人 hG-CSF 和其变体在 50 mM 磷酸钠, 100 mM NaCl, pH 8.5 的溶液中以 250 μ g/ml 的浓度进行 PEG 化。PEG 的摩尔数超过所述蛋白上 PEG 化位点的 100 倍。将反应混合物放入 37°C 热混合器中, 1200 rpm, 30 分钟。30 分钟后, 加入摩尔数过量的甘氨酸终止反应。

- 10 用阳离子交换层析去除反应混合物中过量的 PEG, 甘氨酸和其它副产品。用 20 mM 柠檬酸钠 pH 2.5 稀释 PEG 化反应混合物至离子强度 7 mS/cm 以下。用 5 N HCl 调整 pH 值至 2.5。将上述混合物加到用 20 mM 柠檬酸钠 pH 2.5 平衡过的 SP-sepharose FF 柱上。用 4 个柱体积的平衡缓冲液将未结合的物质从柱上洗下来。用 20 mM 的柠檬酸钠, 750 mM 氯化钠以三个柱体积
- 15 洗脱 PEG 化蛋白。浓缩纯 PEG 化 G-CSF 并用分子量截留值(mwco)为 10kDa 的 VivaSpin 浓缩仪对其进行缓冲液的交换。

有 G-CSF 活性的标记多肽在微量滴定板中的 PEG 化

- 表达携带适当的标记物, 例如上述一般描述中列举的任何一种标记物, 并显示 G-CSF 活性的多肽, 然后, 将培养液转移至能固定标记多肽的微量
- 20 滴定板的一个或多个孔中。当标记物是 Met-Lys-His-Gln-His-Gln-His-Gln-His-Gln-His-Gln-His-Gln 时, 使用可购自 QIAGEN 的镍-氨三乙酸 (Ni-NTA) HisSorb 微量滴定板。

- 将标记多肽固定到微量滴定板, 然后用适于结合和随后 PEG 化的缓冲液洗涤各孔, 然后与选择的活化 PEG 一起温育。例如, 可使用 Shearwater
- 25 Polymers 的 M-SPA-5000。应使活化的 PEG 与多肽的摩尔比优化, 但通常大于 10: 1, 例如大约 100: 1 或更高。室温反应合适的时间, 通常约 1 小时后, 除去活化的 PEG 溶液以便中止反应。偶联蛋白通过与适当的缓冲液温育而从微量滴定板中被洗脱下来。适当的洗脱缓冲液可能包含咪唑, 过量 NTA 或另一种螯合化合物。分析所述偶联蛋白的相应生物学活性和免疫原性。可
- 30 任选使用本领域公知的方法, 如使用二氨基肽酶将该标记物切割下来并用 GCT(谷氨酰环转移酶)将 1-位的 Gln 转化为焦谷氨酰基, 最后用 PGAP(焦-

谷氨酰基-氨肽酶)将其切割下来,产生无标记的蛋白质。该方法包括几步骤金属螯合亲和层析。或者,可以将经过标记的多肽进行偶联。

有一个封闭的受体结合位点的 hG-CSF 活性多肽的 PEG 化

- 为了使 hG-CSF 以一定方式最佳 PEG 化,并保证参与受体识别的赖氨酸不被 PEG 化,已经开发了下列方法:

纯化的 hG-CSF 如实施例 3 中描述的方法获得。hG-CSF 多肽和 G-CSF 受体可溶性功能域以 2:2 化学计量组成的同型二聚体复合物在磷酸盐缓冲液(PBS) pH 7 中形成。hG-CSF 多肽的浓度是大约 20 ug/ml 或 1uM 且受体以相同的摩尔浓度存在。

- 10 将来自 Shearwater Polymers, Inc 的 M-SPA-5000 以相当于比 hG-CSF 多肽过量 5, 20 和 100 摩尔的三种不同浓度加入。反应时间为室温 30 分钟。反应了 30 分钟后,将反应混合物调为 pH 2.0,且将反应混合物上样至 Vydac C18 柱上,然后基本上按已描述的方法(Utsumi 等, J. Biochem., Vol. 101, 1199-1208, (1987)用乙腈梯度液洗脱。或者,可使用异丙醇梯度洗脱。

- 15 用本文描述的初步筛选试验对各组分进行分析,然后将上述方法获得的活性 PEG 化 hG-CSF 多肽于-80℃储存在包含 1 mg/ml 人血清白蛋白(HSA)的 PBS pH7 中。

用于确定偶联和非偶联型 hG-CSF 和其变体的特性的方法

确定 hG-CSF 和其变体的分子大小

- 20 偶联型或非偶联型 hG-CSF 或其变体的分子量可以通过 SDS-PAGE, 凝胶过滤, 基质辅助的激光解吸质谱测定法(matrix assisted laser desorption mass spectrometry)或者平衡离心法确定。

确定多肽浓度

- 25 用 280nm 光密度检测法, 酶联免疫吸附实验(ELISA), 放射免疫分析(RIA), 或本领域熟知的其它此类免疫检测技术检测多肽的浓度。而且, 样品中的多肽浓度也可用 Biacore® 仪器用包被了该多肽特异性抗体的 Biacore®芯片进行检测。

- 30 将上述抗体通过多种化学方法共价偶联到 Biacore®芯片上。或者, 使抗体非共价结合, 例如通过特异性针对抗多肽抗体 Fc 部分的抗体来结合。将 Fc 特异性抗体首先共价偶联到芯片上。然后使抗多肽抗体流过该芯片并被第一抗体直接结合。进一步地, 可用链霉亲和素包被的表面(例如 Biacore

Sensor Chip SA®)固定生物素化抗体 (Real-Time Analysis of Biomolecular Interactions, Nagata and Handa (Eds.), 2000, Springer Verlag, Tokyo ; Biacore 2000 Instrument Handbook, 1999, Biacore AB)。

- 5 当样品流过该芯片时,多肽将结合到包被抗体上且质量的增加能得到检测。可用已知浓度的多肽制品绘出标准曲线,随后确定样品中多肽的浓度。每次注射样品后,通过能去除结合型分析物的适当洗脱液(例如低 pH 值的缓冲液)使传感器芯片再生。

- 10 通常,所应用的抗体是抗野生型多肽的单克隆抗体。在野生型多肽中引入突变或其它操作(额外糖基化或聚合物偶联)可以改变上述抗体的识别。进一步地,能使多肽的分子量增加的操作将使胞质基因共振信号(plasmon resonance signal)增强。因此,有必要对每一种待测分子建立一个标准曲线。

检测偶联和非偶联型 hG-CSF 和其变体的体外和体内活性的方法

初步试验 1 - 体外 G-CSF 活性试验

- 15 鼠细胞系 NFS-60 (从 Dr. J. Ihle, St. Jude Children's, Research Hospital, Tennessee, USA 获得)的增殖依赖于生长培养基中活性 G-CSF 的存在。因此, hG-CSF 和其变体的体外生物学活性可以如下检测,把 G-CSF 样品加到生长培养基中,温育一段时间后检测正在分裂的 NFS-60 细胞的数量。

- 20 将 NFS-60 细胞培养在包含 10% w/w FBS(胎牛血清), 1% w/w Pen/Strep, 每升 10 μ g 的 hG-CSF 和 2 mM Glutamax 的 Iscoves DME 培养基中。在加入样品前,用无 hG-CSF 的生长培养基洗涤细胞两次,然后稀释细胞至 2.2×10^5 细胞/ml 的浓度。加 100 μ l 细胞悬液至 96 孔微滴定板(Corning)的每孔中。

- 25 包含偶联型或非偶联型 G-CSF 或其变体的样品用生长培养基稀释至浓度达到 1.1×10^{-6} M 至 1.1×10^{-13} M。每种样品取 10 μ l 加至包含 NFS-60 细胞的 3 个孔中。由 10 μ l 哺乳动物生长培养基组成的对照加在每个微滴定板的 8 个孔中。细胞培养 48 小时(37 $^{\circ}$ C, 5%CO₂)后用 WST-1 细胞增殖剂(Roche Diagnostics GmbH, Mannheim, Germany)对每个孔中正在分裂的细胞进行定量。各孔加 0.01 ml WST-1,然后在 5% CO₂ 的空气环境下 37 $^{\circ}$ C 温育 150min。活细胞中线粒体脱氢酶分解四唑盐 WST-1 形成甲臞(formazan),该产物可通过 450 nm 下的吸光率定量。从而对每个孔中的活细胞进行定量。

- 30 在上述检测基础上,计算每种偶联和非偶联 G-CSF 分子或其变体的剂量反应曲线,然后确定每种分子的 EC50 值。该值等于获得非偶联的人 G-CSF

的最大增殖活性的 50%所必需的活性 G-CSF 蛋白的量。所以, EC50 值是对给定蛋白的体外活性的直接测量。

初步试验 2 - 体外 G-CSF 活性试验

用质粒转染鼠造血细胞系 BaF3, 该质粒携带人 G-CSF 受体基因, 转录调节子, fos 的启动子和其后的萤光素酶报告基因。用 G-CSF 样品刺激上述细胞系, 多种细胞内反应引发对 fos 表达的刺激, 随后导致萤光素酶的表达。这种刺激可以通过 Steady-Glo™ 萤光素酶检测系统(Promega, Cat. No. E2510)监测, 籍此可以对 G-CSF 样品的体外活性进行检测。

在 37℃潮湿的 5%CO₂ 空气环境下, 在完全培养基(RPMI-1640/HEPES (Gibco/BRL, Cat. No. 22400), 10% FBS (HyClone, 特定的), 1 × 青霉素/链霉素(Gibco/BRL, Cat. No. 15140-122), 1 × L-谷氨酰胺, (Gibco/BRL, Cat. No. 25030-081), 10% WEHI-3 培养液(可产生 muIL-3)中培养 BaF3/hGCSF-R/pfos-lux 细胞, 使其生长到 5 × 10⁵ 细胞/mL 的密度(汇合的)。每 2-3 天以大约 2 × 10⁴ 细胞/mL 进行再接种。

在检测的前一天, 将对数期细胞以 2 × 10⁵ 细胞/mL 悬浮在饥饿培养基中(DMEM/F-12 (Gibco/BRL, Cat. No. 11039), 1% BSA (Sigma, Cat. No. A3675), 1x 青霉素/链霉素(Gibco/BRL, Cat. No. 15140-122), 1x L-谷氨酰胺(Gibco/BRL, Cat. No. 25030-081), 0.1% WEHI-3 培养液), 维持饥饿状态 20 小时。然后用 PBS 洗涤细胞两次并用台盼兰活性染色法检测所述细胞的活性。将细胞以 4 × 10⁶ 细胞/mL 的浓度重新悬浮在试验培养基中(RPMI-1640 (无酚红, Gibco/BRL, Cat. No. 11835), 25 mM HEPES, 1% BSA (Sigma, Cat. No. A3675), 1x 青霉素/链霉素(Gibco/BRL, Cat.No. 15140-122), 1x L-谷氨酰胺(Gibco/BRL, Cat. No. 25030-081), 然后 96 孔微滴定板(Corning) 中每孔加 50μl 所述细胞。用试验培养基将包含偶联型或非偶联型 G-CSF 或其变体的样品稀释到浓度为 1.1 × 10⁻⁷M 至 1.1 × 10⁻¹²M。每种样品 50μl 加入包含 BaF3/hGCSF-R/pfos-lux 细胞的三个孔中。将阴性对照 50μl 培养基加入到每块微滴定板的 8 个孔中。将平板轻轻混合并于 37℃保温 2 小时。依照 Promega Steady-Glo™ 的操作步骤(Promega Steady-Glo 萤光素酶检测系统, Cat. No. E2510)检测萤光素酶的活性。每孔加 100μl 底物, 然后轻轻混匀。在 TopCount 发光仪 (Packard)上以 SPC(单光子计数)的模式检测发光。

根据上述检测, 可以绘出每种偶联和非偶联 G-CSF 分子或其变体的剂

量反应曲线,之后,可确定每种分子的 EC₅₀ 值。

第二步试验 - G-CSF 或其变体与 hG-CSF 受体的结合层析

应用标准的结合实验对 rhG-CSF 或其变体与 hG-CSF 受体的结合进行研究。所述受体可以是纯化的细胞外受体功能区,结合到纯化的细胞浆膜上的受体,或整个细胞,这些细胞可以是天然表达 G-CSF 受体的细胞系(例如 NFS-60)或用编码所述受体的 cDNAs 转染的细胞。rhG-CSF 或其变体与天然 hG-CSF 竞争结合位点的能力可通过其与标记后的 G-CSF-类似物,例如生物素化的 hG-CSF 或放射性碘标记的 hG-CSF 共同温育进行分析。这种检测的实例由 Yamasaki 等(Drugs. Exptl. Clin. Res. 24: 191-196 (1998))描述。

可选择地将 hG-CSF 受体的细胞外功能域偶联到 Fc 上且将其固定到 96 孔板上。随后加入 RhG-CSF 或其变体,它们的结合可以用特异性的抗 hG-CSF 抗体或生物素化的或放射性碘标记的 hG-CSF 进行检测。

对偶联型或非偶联型 rhG-CSF 和其变体的体内半寿期的检测

本发明一个重要方面是延长生物学半寿期,这可以通过构建以下 hG-CSF 实现,其中的多肽与聚合物组分偶联或不偶联。hG-CSF 血清浓度的快速消减使评估对使用偶联型和非偶联型 hG-CSF 及其变体进行的治疗的生物学反应变得重要。优选,在静脉内(i.v.)给药后,本发明的偶联型和非偶联型 hG-CSF 和其变体也具有延长的血清半寿期,使例如 ELISA 方法或初步筛选检测成为可能。体内生物半寿期的检测按照以下描述进行。

应用雄性 Sprague Dawley 大鼠(7 周龄)。在给药的当天,称量动物体重(每个动物 280-310g)。将非偶联型和偶联型 hG-CSF 样品按每 kg 体重 100 μ g 经尾静脉注射给三只大鼠。注射后 1 分钟,30 分钟,1,2,4,6,和 24 小时,在 CO₂ 麻醉下每只大鼠眼球取血 500 μ l。室温储存血样品 1 $\frac{1}{2}$ 小时后,离心分离血清(4 $^{\circ}$ C, 18000xg 离心 5 分钟)。-80 $^{\circ}$ C 储存血样品至分析日。将样品在冰上解冻后通过 G-CSF 体外活性试验(参见初步试验 1 和 2)对血清样品中的 G-CSF 进行定量。

US 5, 824, 778 中描述了测定 G-CSF 或其变体的体内半寿期的另一试验,该文献引入本发明作参考。

偶联型和非偶联型 hG-CSF 和其变体的体内生物学活性的检测

对 SPF Sprague Dawley 大鼠(购自 M & B A/S, 丹麦)中 hG-CSF 的体内生物学活性的检测用于评估偶联型和非偶联型 G-CSF 和其变体的生物学功

效。

新到的大鼠当天随机分为6组。所有动物需要适应7天,其间要去掉状态差或体重极端的个体。顺应阶段开始时的体重范围是250-270g。

给药当天大鼠需禁食16小时,随后按每kg体重100 μ g皮下注射hG-CSF或其变体。每种hG-CSF样品给一组6只随机分配的大鼠注射。在给药前和给药后6, 12, 24, 36, 48, 72, 96, 120 和144小时从大鼠尾静脉抽取300 μ l血样品,用EDTA抗凝。对血样品进行以下血液学参数的分析:血红蛋白,红细胞计数,血细胞比容,平均细胞体积(mean cell volume),平均细胞血红蛋白浓度(mean cell haemoglobin concentration),平均细胞血红蛋白(mean cell haemoglobin),白细胞计数,分化型白细胞计数(嗜中性粒细胞,淋巴细胞,嗜酸性粒细胞,嗜碱性粒细胞,单核细胞)。根据这些检测可以对偶联型和非偶联型hG-CSF和其变体进行评估。

US 5, 681, 720, US 5, 795, 968, US 5, 824, 778, US 5, 985, 265 和 Bowen et al., Experimental Hematology 27: 425-432 (1999)描述了测定hG-CSF或其变体的体内生物学活性的其它试验例。

确定多肽-受体结合亲和力(结合率和解离速度)

用酶联免疫吸附实验(ELISA),放射免疫分析(RIA),或本领域熟知的其它此类免疫检测技术可以检测受体和配体之间的结合强度。使用胞质基因共振检测的Biacore®仪(Zhou et al., Biochemistry 3, 32, 8193-98; Faegerstram 和 O'Sh annessy, 1993, In Handbook of Affinity Chromatography, 229-52, Marcel Dekker, Inc., NY)也可以检测配体-受体结合相互作用。

Biacore®技术允许受体结合到金颗粒表面并使配体流过所述受体。胞质基因共振检测实时对结合到该表面的物质进行直接计量。使用该技术可获得结合速度常数和解离速度常数并因此直接确定配体-受体解离常数和亲和力常数。

hG-CSF 偶联物的体外免疫原性检测

用检测偶联物相对于参考分子或制品的免疫原性的ELISA方法确定本发明偶联物的免疫原性下降。通常,参考分子或制品是重组人G-CSF制品例如Neupogen®或另一种重组人G-CSF制品,例如US 5, 824, 784中描述的N末端PEG化rhG-CSF分子。ELISA方法是以抗体为基础建立的,所述抗体来自用上述重组G-CSF制品中的一种进行治疗的患者。当试验中本发明

偶联物的反应与参考分子或制品相比在统计学意义上明显下降时应认为免疫原性是下降的。

G-CSF 生物检测中对活性的中和

用上述 G-CSF 生物检测方法分析抗 G-CSF 血清对 hG-CSF 偶联物的中和作用。

使用来自 G-CSF 参考分子治疗的患者或来自已被免疫的动物的血清。加入固定浓度(1: 20-1: 500 (患者血清) 或 20-1000 ng/ml (动物血清)稀释)或从 1: 20 (pt 血清) 或 1000 ng/ml (动物血清)开始进行 5 倍连续稀释的上述血清。在总量为 80 μ l DMEM 培养基+ 10% FCS 中以从 10 nM 开始进行 7 倍稀释的浓度或以固定浓度 (1-100 pM) 加入 HG-CSF 偶联物。将血清与 hG-CSF 偶联物在 37°C 温育 1 小时。

然后将样品(0.01 ml)转移到包含 NFS-60 细胞的 0.1 ml DMEM 培养液的 96 孔组织培养板中。培养物在 5%CO₂ 的空气环境下 37°C 温育 48 小时。然后加 0.01 ml WST-1 (WST-1 细胞增殖剂 Roche Diagnostics GmbH, Mannheim, Germany)至培养物中, 在 5%CO₂ 的空气环境下 37°C 温育 150 分钟。活细胞中线粒体脱氢酶对四唑盐 WST-1 的分解导致形成甲臞, 该产物通过检测 450 nm 的吸光度进行定量。

在有固定量血清的条件下对 hG-CSF 偶联物样品进行滴定时, 中和效应被定义为抑制倍数(FI), 计量为 EC50(有血清)/EC50(无血清)。G-CSF 变体蛋白的抗体中和作用的下降被定义为

$$\text{FI 变体-1} \\ (1 - \frac{\text{FI 变体-1}}{\text{FI wt-1}}) \times 100\%$$

25 实施例 1

编码 hG-CSF 的合成基因的构建和克隆

以下 DNA 片段是按照 Stemmer et al., (1995), Gene 164, pp. 49-53 描述的通用步骤合成的:

片段 1, 由以下序列组成: Bam HI 消化位点, 编码 YAP3 信号肽的序列 (WO 98/32867), 编码 TA57 前导序列的序列(WO 98/32867), 编码 KEX2 蛋白酶识别位点的序列(AAAAGA), 编码 hG-CSF 的序列(其密码子使用特点有

利于在大肠杆菌中表达), (SEQ ID NO : 2)和 Xba I 消化位点。

片段 2, 由以下序列组成: Bam HI 消化位点, 编码 YAP3 信号肽的序列 (WO 98/32867), 一个编码 TA 57 前导序列的序列 (WO 98/32867), 编码组氨酸标记物的序列 (SEQ ID NO : 5), 编码 KEX2 蛋白酶识别位点的序列 (AAAAGA), 编码 hG-CSF 的序列 (其密码子使用特点有利于在大肠杆菌中表达), (SEQ ID NO : 2)和 Xba I 消化位点。

片段 3, 由以下序列组成: Nde I 消化位点, 编码 OmpA 信号肽的序列 (SEQ ID NO : 3), 编码 hG-CSF 的序列 (其密码子使用特点有利于在大肠杆菌中表达), (SEQ ID NO : 2)和 Bam HI 消化位点。

片段 4, 由以下序列组成: Bam HI 消化位点, Kozak 共有序列 (Kozak, M.J Mol Biol 1987 Aug 20; 196 (4): 947-50), 编码 hG-CSF 信号肽的序列 (SEQ ID NO : 7)和编码 hG-CSF 的序列 (其密码子使用特点有利于在 CHO 中表达), (SEQ ID NO : 8)和 Xba I 消化位点。

用标准 DNA 技术将 DNA 片段 1 和 2 插入到质粒 pJSO37 (Okkels, Ann. New York Acad. Sci. 782: 202-207, 1996)中的 Bam HI 和 Xba I 消化位点。这样产生质粒 pG-CSFcerevisiae 和 pHISG-CSFcerevisiae。

用标准 DNA 技术将 DNA 片段 3 插入到质粒 pET12a (Invitrogen)中的 Nde I 和 Bam HI 消化位点。这样产生质粒 pG-CSFcoli。

用标准 DNA 技术将 DNA 片段 4 插入到质粒 pcDNA3.1 (+) (Invitrogen)中的 Bam HI 和 Xba I 消化位点。这样产生质粒 pG-CSFCHO。

实施例 2

hG-CSF 在酿酒酵母(S. cerevisiae)和大肠杆菌(E. coli)中的表达

用质粒 pG-CSFcerevisiae 或 pHISG-CSFcerevisiae 转化酿酒酵母 (*Saccharomyces cerevisiae*) YNG318 (可从美国典型培养物保藏中心, VA, USA 以 ATCC 208973 的保藏号获得), 分离出包含这两种质粒中任一种的转化子, 然后在细胞外分别表达携带和不携带 HIS 标记物的 hG-CSF, 以上各个步骤按照文献中描述的标准技术进行。用 pG-CSFcoli 转化大肠杆菌 BL21 (DE3) (Novagen, Cat. No. 69387-3), 分离出包含该质粒的转化子, 随后在上清液和细胞的外周胞质中表达 hG-CSF, 以上步骤按照 Novagen 的 pET System Manual (第 8 版)中描述的进行。

酿酒酵母和大肠杆菌对 hG-CSF 的表达可以用 ImmunoPure

Ultra-Sensitive ABC 兔 IgG 染色试剂盒(Pierce)和多克隆抗 hG-CSF 抗体(Pepro Tech EC Ltd.)通过 Western 印迹分析来检验。结果发现所述蛋白具有正确的大小。

在酿酒酵母和大肠杆菌中的携带和不携带 N 末端组氨酸标记物的 hG-CSF 的表达水平可用市售 G-CSF 特异性 ELISA 试剂盒(Quantikine Human G-CSF Immunoassay, R & D Systems Cat. No. DCS50)进行定量。检测到的数值在以下列出。

表达系统	表达水平(mg G-CSF/ L)
酿酒酵母中的 hG-CSF	30
酿酒酵母中具有组氨酸标记物的 hG-CSF	25
大肠杆菌中的 hG-CSF	0.05

实施例 3

从酿酒酵母培养上清中纯化 hG-CSF 及其变体

hG-CSF 的纯化如下进行:

通过离心去除细胞。然后将去除了细胞的上清液通过 0.22μm 的过滤器过滤除菌。经过滤除菌后的上清液在 10mM, pH 4.5 的醋酸钠中稀释 5 倍。每 5 升稀释后的上清液加入 10 ml 浓缩的醋酸调整 pH。在将该溶液应用于阳离子交换柱之前其离子强度应该在 8mS/cm 以下。

将稀释后的上清液以 90cm/h 的线性流速装载到已用 50mM, pH 4.5 的醋酸钠平衡的 SP-sepharose FF (Pharmacia)柱上,直到该柱的洗出液达到稳定的 UV 和基线导电性水平。用所述平衡缓冲液洗柱直到柱流出液的 UV 吸光度和导电性达到稳定水平,从而去除未结合的物质。用线性梯度; 30 个柱体积; 0-80%的缓冲液 B(50mM NaAc, pH 4.5, 750mM NaCl)以 45cm/h 的流速将结合的 G-CSF 蛋白从柱子上洗脱下来。根据 SDS 聚丙烯酰胺凝胶电泳的结果将包含 G-CSF 的组分汇集。加入氯化钠直到该溶液的离子强度在 80mS/cm 以上。

将上述蛋白溶液装载到已用 50mM NaAc, pH 4.5, 750mM NaCl 平衡的 Phenyl Toyo Pearl 650S 柱上。用该平衡缓冲液洗下未结合的物质。通过应用 MilliQ 水的逐步梯度进行 G-CSF 的洗脱。包含 G-CSF 的组分被汇集。通过应用上述 2 步冲洗(down stream)的策略,能获得 90%以上纯度的 G-CSF。然后纯化的蛋白通过在 280nm 的分光光度检测和/或通过氨基酸分析进行定

量。

包含 G-CSF 的组分被汇集。用 VivaSpin 浓缩器(mwco: 5kDa)进行缓冲液交换和浓缩。

¹ 实施例 4

5 对非偶联的和偶联的 hG-CSF 和其变体的鉴定和定量

SDS 聚丙烯酰胺凝胶电泳

对纯化并浓缩的 G-CSF 进行 SDS-PAGE 分析。结果主要是表观分子量大约 17kDa 的单一一条带。

吸光度

10 通过分光光度测量法评估 G-CSF 浓度。通过检测在 280 nm 的吸光度和应用 0.83 的理论消光系数可以计算出蛋白的浓度。

氨基酸分析

通过氨基酸分析更准确测定蛋白。对纯化的 G-CSF 所进行的氨基酸分析揭示，经实验确定的氨基酸组分与根据 DNA 序列推测的氨基酸组成是一致的。

15 实施例 5

PEG 化 wt G-CSF 和 G-CSF 变体的 MALDI-TOF 质谱测定

用 MALDI-TOF 质谱测定法评估附着到 PEG 化 wt G-CSF 和所选择的 PEG 化 G-CSF 变体上的 PEG 基团数。

20 Wt G-CSF 包含 5 个预计为 SPA-PEG 附着位点的伯氨基(N 末端氨基和 K16, K23, K34 和 K40 上的 ϵ 氨基)。wt G-CSF 用 SPA-PEG 5000 PEG 化后, MALDI-TOF 质谱检测显示主要存在附着有 4, 5 和 6 个 PEG 基团的 wt G-CSF 变体, 此外, 附着 7 个 PEG 基团的 wt G-CSF 变体也清晰可见, 但量较少。

25 具有 K16R, K34R, K40R, Q70K, Q90K, 和 Q120K 取代的 G-CSF 变体也包含 5 个伯氨基(N 末端氨基和 K23, K70, K90 和 K120 上的 ϵ 氨基)。这种 G-CSF 变体被 SPA-PEG5000 PEG 化后, MALDI-TOF 质谱检测显示主要存在附着有 4, 5 和 6 个 PEG 基团的 G-CSF 变体, 此外, 附着 7 个 PEG 基团的 G-CSF 变体也清晰可见, 但量较少。

30 具有 K16R, K34R, 和 K40R 取代的 G-CSF 包含 2 个伯氨基(N 末端氨基和 K23 上的 ϵ 氨基)。该 G-CSF 变体用 SPA-PEG 12000 PEG 化后,

MALDI-TOF 质谱检测显示主要存在附着有 2 和 3 个 PEG 基团的 G-CSF 变体, 此外, 附着 4 个 PEG 基团的 G-CSF 变体也清晰可见, 但量较少。

上述观察清楚地显示除包含氨基的氨基酸残基外, 其它的氨基酸残基在所应用的 PEG 化条件下有时也可以进行 PEG 化。上述观察还显示在引入氨基的部位进行 PEG 化具有一定重要性。这在 wt G-CSF 和 G-CSF 变体的 SDS-PAGE 分析中也已观察到。

如实施例 11 所述, 当应用 SPA-PEG 化学法时组氨酸 170 完全被 PEG 化。此外, K23 和 S159 被部分 PEG 化。这解释了除 hG-CSF 和其变体中已经形成的伯氨基外, 还有 1-2 个另外的 PEG 化位点。

10 实施例 6

PEG 化和非 PEG 化 G-CSF 变体的肽作图

为了绘制 G-CSF 和 G-CSF 变体上另外的 SPA-PEG 附着位点, 应用以下策略。

为了将预期的 PEG 化位点数降至最少, 选择具有少量氨基的 G-CSF 变体。所选择的 G-CSF 变体具有 K16R, K34R, K40R 和 H170Q 取代。除了 K23 上的在先资料显示无任何程度 PEG 化的ε胺基外, 该变体仅仅包含一个 N 末端伯胺。所以, 在该 G-CSF 变体中胺基团上的 PEG 化背景明显降低。应用 SPA-PEG 5000 对 G-CSF 变体进行 PEG 化。PEG 化后, 使 G-CSF 变体变性, 使其二硫键还原, 使所得巯基烷基化, 然后用谷氨酸特异性蛋白酶降解已烷基化并 PEG 化的蛋白。最后, 所形成的肽通过反相 HPLC 分离。

平行地对具有 K16R, K34R, 和 K40R 取代的非 PEG 化 G-CSF 变体进行同样的处理, 以获得一个 HPLC 参考层析图。

对 PEG 化 G-CSF 变体和非 PEG 化 G-CSF 变体的降解的 HPLC 层析作图比较应该揭示出在 PEG 化时哪个肽消失。通过对非 PEG 化 G-CSF 变体的肽进行 N 末端氨基酸测序而对这些肽进行的鉴定可间接指出 PEG 化位置。

原则上, 优选应用具有 K16R, K34R, K40R 和 H170Q 所有这些取代的非 PEG 化 G-CSF 变体, 但实践中不必如此。

更具体地, 将大约 1mg 具有 K16R, K34R, K40R 和 H170Q 取代的 PEG 化 G-CSF 变体和大约 500μg 具有 K16R, K34R, 和 K40R 取代的非 PEG 化 G-CSF 变体在 SpeedVac 浓缩器中干燥。将两种样品分别溶解在 400μl 6 M 胍, 0.3 M Tris HCl, pH 8.3 中, 然后 37℃变性过夜。变性后, 通过加入 50μl 0.1 M

- DTT 的 6 M 胍, 0.3 M TrisHCl, pH 8.3 溶液还原蛋白质中的二硫键。室温下温育 2h 后, 加入 50 μ l 0.6 M 碘乙酰胺的 6 M 胍, 0.3 M Tris-HCl, pH 8.3 溶液使原有巯基烷基化。烷基化在室温下进行 30min, 然后用 NAP5 柱将已还原并烷基化的蛋白的缓冲液换成 50 mM NH_4HCO_3 。在 SpeedVac 浓缩器中样品的
- 5 体积降低到大于 200 μ l 后, 分别加入 20 μ g 和 10 μ g 谷氨酸特异性蛋白酶。谷氨酸特异性蛋白酶的降解在 37 $^\circ\text{C}$ 进行 16h。所产生的肽应用 Phenomenex Jupiter C18 柱(0.2 $\times$ 5 cm)以 0.1%TFA 水溶液中的线性乙腈梯度洗脱, 从而进行反相 HPLC 分离。通过 MALDI-TOF 质谱法分析所收集的组分, 随后对所选择的肽进行 N 末端氨基酸序列分析。
- 10 对 PEG 化 G-CSF 变体和非 PEG 化 G-CSF 变体的降解的 HPLC 层析作图比较揭示, 在 PEG 化时仅二个组分消失。对非 PEG 化 G-CSF 变体的这两个组分进行 N 末端氨基酸序列分析显示, 这两种肽都来源于 G-CSF 的 N 末端。其中一种肽由 Gln11 以后的意外切割产生的氨基酸残基 1-11 组成。另一肽由 Glu19 后的预期切割产生的氨基酸残基 1-19 组成。
- 15 预期 G-CSF 的 N 末端肽可用该方法鉴定, 因为 N 末端氨基容易被 PEG 化。然而, 用该方法没有鉴定出 SPA-PEG 5000 的其它附着位点。
- 间接鉴定 PEG 5000 附着位点的另一种做法是直接鉴定 PEG 化肽中的附着位点。然而, 在对降解的 PEG 化 G-CSF 变体进行 HPLC 分离时, 包含 PEG 化肽的组分彼此之间没有分开且未能与几个包含非 PEG 化肽的组分分开。
- 20 所以, 对这些组分的 N 末端氨基酸序列分析除了提示 K23 能被部分 PEG 化外没获得任何有用的数据。
- 为克服这些问题, 从第一次 HPLC 分离的组分中制备两份 PEG 化肽的汇集物。将这二个汇集物在 SpeedVac 浓缩器中干燥, 溶解在 200 μ l 新鲜配制的 50 mM NH_4HCO_3 中并用 1 μ g 糜蛋白酶进一步降解。所产生的肽应用
- 25 Phenomenex Jupiter C18 柱(0.2 $\times$ 5 cm)以 0.1%TFA 水溶液中的线性乙腈梯度洗脱, 从而经反相 HPLC 分离。通过 MALDI-TOF 质谱法分析所收集的级分, 随后对所选择的肽进行 N 末端氨基酸序列分析。
- 经 N 末端氨基酸序列测定, K23 和 S159 都被部分 PEG 化。无法确定这两个位置上 PEG 化实际程度, 但因为从最初的 HPLC 分离中鉴定出 K23
- 30 和 S159 未被修饰的肽并已测序, 所以可确定 PEG 化仅仅是部分的。

实施例 7

wt G-CSF 和 G-CSF 变体的糖基化

当通过 MALDI-TOF 质谱分析对纯化的 wt G-CSF 和 G-CSF 变体进行分析时总是能观察到比所分析的 G-CSF 分子的质量大了约 324Da 的另一组分。因为具有最小质量的组分始终具有 G-CSF 分子的质量且因为 G-CSF 分子有正确的 N 末端氨基酸序列，所以可以得出这样的结论，上述另外的成分 5 是携带二个己糖残基的经过修饰的 G-CSF 分子。在许多情况下，未经修饰的 G-CSF 分子产生最强的信号但在一些情况下经过修饰的 G-CSF 分子的信号强度是最强的。

在分析为鉴定另外的 PEG 化位点而产生的肽时，从每次降解中鉴定出 10 二个可用于鉴定糖基化位点的肽。

在两次 HPLC 分离中，这两种肽相继被洗脱出且 MALDI-TOF 质谱检测显示这两种肽有约 324Da 的质量差异。质谱检测数据提示该肽包含氨基酸残基 124-162。对所有 4 种肽的 N 末端氨基酸序列分析显示这种分配是正确的且 Thr133 是唯一的修饰位点。在具有未经修饰的肽的质量的肽中，Thr133 15 在序列中清晰可见，但在具有 324 Da 的附着质量的肽中第 133 位无法排布氨基酸残基。因为所有其它的氨基酸残基都能排布在该序列中，所以可以得出这样的结论，即 Thr133 是唯一的修饰位点。这个糖基化位点以前曾报道用于 CHO 细胞中表达的重组 G-CSF 中，但此聚糖与酵母中附着的聚糖不同。

非糖基化 wt G-CSF 利用反相 HPLC 与糖基化 wt G-CSF 分离，具体是 20 应用 Vydac C₁₈ 柱(0.21 × 5cm)以 51%乙腈的 0.1%TFA 溶液经等度洗脱为一个经 MALDI-TOF 质谱分析显示仅含非糖基化 wt G-CSF 的级分。

实施例 8

共价偶联有不同数量的 PEG 分子的 G-CSF 分子的分离

共价偶联 4, 5 或 6 个 PEG 基团的 G-CSF 分子如下分离。用 20mM 柠檬酸钠 pH 2.5 平衡 SPsepharose FF 柱，然后将 20mM 柠檬酸钠, pH 2.5 中的 25 PEG 化蛋白装载到上述 SPsepharose FF 柱上。任何未结合的物质都被从柱子上洗掉。洗脱用 pH 梯度进行。PEG 化 G-CSF 在大约 pH3.8 时开始从柱子中洗脱且继续洗脱在 pH3.8 到 4.5 的级分中。

对这些级分进行 SDS-PAGE 和质谱分析。这些分析表明 PEG 化程度最高的 G-CSF 存在于低 pH 级分中。PEG 化程度较低 G-CSF 洗脱在高 pH 级分 30 中。

对 PEG 化 G-CSF 所进行的氨基酸分析显示, 消光系数的理论值与实验测定值非常一致。

实施例 9

hG-CSF 变体的构建

- 5 将 hG-CSF 中的氨基酸特异性取代为其它氨基酸残基, 例如上面在一般描述中所讨论的特异性取代, 是用本领域已知的标准 DNA 技术引入的。用包含编码无 HIS 标记物之 hG-CSF 的基因的质粒 pG-CSF *cerevisiae* 可制备新的 G-CSF 变体, 所述 hG-CSF 基因在 PCR 反应中作 DNA 模板。使这些变体在酿酒酵母中表达且如实施例 3 所述纯化。所构建的部分 G-CSF 变体在以下列举(参见实施例 9 和 10)
- 10

实施例 10

SPA-PEG 向 hG-CSF 或其变体的共价偶联

- 如上所述(hG-CSF 和其变体在溶液中的 PEG 化)使人 G-CSF 和其变体共价连接 SPA-PEG 5000, SPA-PEG 12000 和 SPA-PEG 20000 (Shearwater)。偶联物的体外活性在实施例 12 中列举。
- 15

实施例 11

通过定点诱变和随后对纯化变体的 PEG 化来鉴定 G-CSF 中的 SPA-PEG 附着位点

- 可以将 SPA-PEG 附着到 G-CSF 中除赖氨酸之外的其它氨基酸残基上。
- 20 为确定 SPA-PEG 是否附着到组氨酸, 丝氨酸, 苏氨酸和精氨酸上, 制备一些变体, 使其中的这些氨基酸被取代为赖氨酸, 丙氨酸或谷氨酸。使这些变体在酿酒酵母中表达, 纯化和 PEG 化, 然后在 SDS-PAGE 上分析附着的 SPA-PEG 分子的数量。给定氨基酸被谷氨酸或丙氨酸取代后附着 SPA-PEG 的数量的下降强有力地提示该氨基酸通过 SPA-PEG 被 PEG 化, 且这种现象
- 25 进一步得到将该氨基酸取代成赖氨酸后 PEG 化程度不变的结果的支持。所分析的变体在以下列举。

	G-CSF 变体	附着的 PEG 基团的数量
5	hG-CSF	10% 4 PEG, 75% 5 PEG, 15% 6 PEG
	K23R	10% 4 PEG, 85% 5 PEG, 5% 6 PEG
	H43Q	10% 4 PEG, 75% 5 PEG, 15% 6 PEG
	H43K	10% 5 PEG, 75% 6 PEG, 15% 7 PEG
	H52Q	10% 4 PEG, 75% 5 PEG, 15% 6 PEG
	H52K	10% 5 PEG, 75% 6 PEG, 15% 7 PEG
	H156Q	10% 4 PEG, 75% 5 PEG, 15% 6 PEG
	H156K	10% 5 PEG, 75% 6 PEG, 15% 7 PEG
	H170Q	10% 3 PEG, 75% 4 PEG, 15% 5 PEG
	H170K	10% 4 PEG, 75% 5 PEG, 15% 6 PEG
10	K16/34R	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34R R22K	10% 3 PEG, 75% 4 PEG, 15% 5 PEG
	K16/34R R22Q	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34R S142A	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R	10% 1 PEG, 75% 2 PEG, 15% 3 PEG
	K16/34/40R S53K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S53A	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S62K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S66K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S80K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
15	K16/34/40R T105K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R T133K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S142K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R R147K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S155K	10% 2 PEG, 75% 3 PEG, 15% 4 PEG
	K16/34/40R S159K	10% 2 PEG, 85% 3 PEG, 5% 4 PEG
	K16/34/40R S170K	10% 1 PEG, 75% 2 PEG, 15% 3 PEG
	K16/34/40R S170Q	5% 0 PEG, 85% 1 PEG, 15% 2 PEG
20		

以上数据显示除 N 末端, K16, K34 和 K40 之外, SPA-PEG 也可共价结合到 H170 上。进一步地, 上述数据显示仅仅 10% 的可利用的 K23 氨基酸残基被 PEG 化, 大约 10% 的 S159 被 PEG 化。

25 实施例 12

非偶联的和偶联的 hG-CSF 和其变体的体外生物学活性

如以上“主要试验 2-体外 hG-CSF 活性检测”所述对非偶联的和偶联的 hG-CSF 和其变体的体外生物学活性进行检测。对 SPA-PEG 5000 与可利用的 PEG 化位点偶联和不偶联所产生的每个变体所检测的 EC50 值在以下列举。这些值已经相对于非偶联的 hG-CSF (Neupogen®) 标准化。每次在同样的检测条件下检测该值以及变体的值。在所述检测中 hG-CSF 的 EC50 值是

30 pM.

G-CSF 变体	EC50 (占 hG-CSF 的%) 未偶联的	EC50 (占 hG-CSF 的%) 与 SPA-PEG5000 偶联
G-CSF, 含 N 末端组氨酸标记	10	未测定
G-CSF, 不含 N 末端组氨酸标记	100	0.1
16R	100	1
16Q	80	1
23Q	80	0.1
23R	100	0.1
34R	100	1
34A	80	1
34Q	70	1
40R	50	1
K16/23R	100	1
K16/23Q	80	1
K34/40R	50	5
K16/34R	100	10
K16/40R	50	5
K16/23/34R	50	10
K16/23/40R	50	5
K16/34/40R	35	30
K16/23/34/40R	20	15
K16/34/40R E45K	低水平表达	未测定
K16/34/40R E46K	10	1
K16/34/40R S53K	5	0.5
K16/34/40R S62K	10	0.5
K16/34/40R S66K	20	2
K16/34/40R Q67K	10	0.2
K16/34/40R Q70K	30	20

5	K16/34/40R S80K	10	0.2
	K16/34/40R Q90K	30	20
	K16/34/40R E98K	低水平表达	未测定
	K16/34/40R D104K	10	0.9
	K16/34/40R T105K	30	10
10	K16/34/40R Q120K	30	20
	K16/34/40R Q131K	低水平表达	未测定
	K16/34/40R T133K	30	10
	K16/34/40R Q134K	10	0.2
	K16/34/40R S142K	20	7
15	K16/34/40R R147K	20	1
	K16/34/40R S155K	20	1
	K16/34/40R S159K	20	3
	K16/34/40R Q70K Q120K	25	15
	K16/34/40R Q90K Q120K	25	15
20	K16/34/40R T105K Q120K	20	8
	K16/34/40R Q120K T133K	20	8
	K16/34/40R Q120K S142K	10	2
	K16/34/40R Q70K Q90K T105K	10	2
	K16/34/40R Q70K Q90K Q120K	20	10
25	K16/34/40R Q70K Q90K T133K	15	5
	K16/34/40R Q70K T105K Q120K	10	2
	K16/34/40R Q70K Q120K T133K	15	5
	K16/34/40R Q70K Q120K S142K	10	1
	K16/34/40R Q90K T105K Q120K	10	2
30	K16/34/40R Q90K T105K T133K	10	2
	K16/34/40R Q90K Q120K T133K	15	5
	K16/34/40R Q90K Q120K S142K	10	1
	K16/34/40R T105K Q120K T133K	10	1
	K16/34/40R Q120K T133K S142K	10	1

以上数据显示 K23 取代成精氨酸不增加偶联蛋白的活性。这是由于这样一个事实，即仅有 10% 的 K23 被 PEG 化，偶联后的 K23R 变体基本与 hG-CSF 具有相同的 PEG 基团附着数和相同的 PEG 化位置。在 K16, K34 和 K40 位置上去除赖氨酸产生了在 PEG 化后具有明显活性的 G-CSF 变体。该变体与 SPA-PEG5000 偶联后与未偶联的变体相比不明显降低活性。所以，N 末端和 H170 被 SPA-PEG 5000 PEG 化不降低 hG-CSF 的活性。由此决定应用 hG-CSF K16R K34R K40R 作为插入新 PEG 化位点的主链。在该主链中，E45 和 H159 残基之间引入了 21 个新的 PEG 化位点。这些残基分布在 hG-CSF 中不与 G-CSF 受体相互作用的部位。在 Q70, Q90, T105, Q120, T133 和 S142 位置引入新的 PEG 化位点得到被 SPA-PEG 5000 PEG 化后仍保持明显活性的

hG-CSF 变体。所以, 具有 2 或 3 个新的 PEG 化位点的 hG-CSF 变体结合了这些新的 PEG 化位点。用 SPA-PEG 5000 PEG 化后这些变体中最具活性的是 hG-CSF K16R K34R K40R Q70K Q120K 和 hG-CSF K16R K34R K40R Q70K Q90K Q120K。

- 5 此外, 可使 SPA-PEG 5000, SPA-PEG 12000 和 SPA-PEG 20000 附着到 hG-CSF K16R K34R K40R 上。它们各自的体外 EC50 值分别是非偶联的 hG-CSF (Neupogen®) 的 35%, 10% 和 1%。

实施例 13

非偶联的和偶联的 hG-CSF 和其变体的体内半寿期

- 10 非偶联的 hG-CSF (Neupogen®) 和 SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 的体内半寿期如以上(非偶联的和偶联的 hG-CSF 和其变体的体内半寿期的检测)所述进行检测。所得结果在图 1 中显示。Neupogen® 和 SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 的体内半寿期经测定分别为 2.01 小时和 12.15 小时。所以, 在
- 15 hG-CSF 中引入新的 PEG 化位点并使这些位点与 SPA-PEG 5000 偶联可使体内半寿期明显增加。在较早的一项试验中, 将 Neupogen® 的体内半寿期与具有较大的 N 末端 PEG 偶联分子(10 kDa)的 hG-CSF 相比较。此时, 体内半寿期分别确定为 1.75 小时和 7.01 小时。所以, 与 Neupogen® 和具有 10 kDa 的 N 末端 PEG 偶联分子的 hG-CSF 相比, SPA-PEG 5000 偶联的 hG-CSF K16R
- 20 K34R K40R Q70K Q90K Q120K 的半寿期明显延长。

实施例 14

非偶联的和偶联的 hG-CSF 和其变体的体内生物学活性

- 非偶联的 hG-CSF (Neupogen®), SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q120K 和 SPA-PEG 5000 偶联的 hG-CSF K16R K34R
- 25 K40R Q70K Q90K Q120K 的体内生物学活性如以上(非偶联的和偶联的 hG-CSF 和其变体的体内生物学活性的检测)所述进行检测。所得结果在图 2 中显示。在 t=0 小时每 kg 体重注射 100μg 后 48 小时未检测出 Neupogen® 的活性。初次注射后, 直到 72 小时都能检测出 SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q120K 的活性, 而直到 96 小时, SPA-PEG 5000 偶
- 30 联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 仍保留体内活性。所以, 可以看出两种偶联的变体与 Neupogen® 比较都有较长时间的体内生物学活

性且 SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 的活性在体内维持的时间是 Neupogen® 的两倍。

此外,对 Neupogen®, SPA-PEG 12000 偶联的 hG-CSF K16R K34R K40R 和不同剂量 SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 的体内生物学活性进行了检测。所得结果显示在图 3。如较早时观察到的,在以每 kg 体重 100 μ g 最初注射后 48 小时未检测出 Neupogen® 的活性。每 kg 体重给药 5 μ g SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 与 Neupogen 相比可使体内生物学活性略微延长,而每 kg 体重给药上述化合物 25 μ g 和 100 μ g 使得 hG-CSF 活性在最初注射后各自到 72 和 96 小时仍维持活性。SPA-PEG 12000 偶联的 hG-CSF K16R K34R K40R 在每 kg 体重给药 100 μ g 后直到 72 小时仍维持体内活性。如实施例 5 所述, SPA-PEG 12000 偶联的 hG-CSF K16R K34R K40R 偶联有 2 或 3 个 SPA-PEG 12000 基团,而 SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 偶联有 5 或 6 个 SPA-PEG 5000 基团。所以,两种化合物的分子量分别是 42-54 kDa 和 43-48 kDa。SPA-PEG 5000 偶联的 hG-CSF K16R K34R K40R Q70K Q90K Q120K 与具有基本相同分子量的 SPA-PEG 12000 偶联的 hG-CSF K16R K34R K40R 相比体内生物学活性能维持较长时间的结果显示,被偶联的 PEG 分子的分布对体内生物学活性的维持时间是至关重要的。

序列表

<110> 马克西根公司 (Maxygen ApS)

<120> G-CSF 偶联物

<130> 8wo02

<150> DK PA 2000 00024

<151> 2000-01-10

<150> DK PA 2000 00341

<151> 2000-03-02

<150> DK PA 2000 00943

<151> 2000-06-16

<160> 8

<170> PatentIn version 3.0

<210> 1

<211> 174

<212> PRT

<213> 人 (Homo sapiens)

<400> 1

Thr	Pro	Leu	Gly	Pro	Ala	Ser	Ser	Leu	Pro	Gln	Ser	Phe	Leu	Leu	Lys
1				5					10					15	

Cys	Leu	Glu	Gln	Val	Arg	Lys	Ile	Gln	Gly	Asp	Gly	Ala	Ala	Leu	Gln
		20						25					30		

Glu	Lys	Leu	Cys	Ala	Thr	Tyr	Lys	Leu	Cys	His	Pro	Glu	Glu	Leu	Val
		35					40					45			

Leu	Leu	Gly	His	Ser	Leu	Gly	Ile	Pro	Trp	Ala	Pro	Leu	Ser	Ser	Cys
		50				55					60				

Pro	Ser	Gln	Ala	Leu	Gln	Leu	Ala	Gly	Cys	Leu	Ser	Gln	Leu	His	Ser
65					70					75					80

Gly	Leu	Phe	Leu	Tyr	Gln	Gly	Leu	Leu	Gln	Ala	Leu	Glu	Gly	Ile	Ser
				85					90					95	

Pro	Glu	Leu	Gly	Pro	Thr	Leu	Asp	Thr	Leu	Gln	Leu	Asp	Val	Ala	Asp
			100					105					110		

Phe	Ala	Thr	Thr	Ile	Trp	Gln	Gln	Met	Glu	Glu	Leu	Gly	Met	Ala	Pro
		115					120					125			

Ala	Leu	Gln	Pro	Thr	Gln	Gly	Ala	Met	Pro	Ala	Phe	Ala	Ser	Ala	Phe
	130					135					140				

Gln	Arg	Arg	Ala	Gly	Gly	Val	Leu	Val	Ala	Ser	His	Leu	Gln	Ser	Phe
145					150					155					160

Leu Glu Val Ser Tyr Arg Val Leu Arg His Leu Ala Gln Pro
165 170

1
<210> 2
<211> 525
<212> DNA
<213> 人工序列

<400> 2
accctctgg gccggccag cagtctgcct cagagttttt tactgaaatg cttagaacag 60
gtgcgtaaaa tccagggcga tggcgggcc ctgcaggaaa aactgtgcgc gacctataaa 120
ctgtgccatc ctgaagaact ggtcctgtta ggccatagct taggcatccc gtgggcgcct 180
ctgagtagct gcccgagtca ggccctgcag ctggccgget gcctgagtca gttacatagt 240
ggcttatttt tatatcaggg cttactgcag gcgttagaag gcattagtcc ggaactgggc 300
ccgacctgg ataccttaca gttagatgtc gcggattttg ccaccaccat ttggcagcag 360
atggaagaat taggcatggc gcctgcgtta cagcctaccc agggcgccat gcctgcgttt 420
gcgagtgcgt ttcagcgtcg cgccgggggc gtgttagtgg ccagccatct gcagagcttt 480
ctggaagtga gttatcgtgt gttacgccat ctggcccagc cttaa 525

<210> 3
<211> 21
<212> PRT
<213> 大肠杆菌 (Escherichia coli)

<400> 3

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala
1 5 10 15

Thr Val Ala Gln Ala
20

<210> 4
<211> 63
<212> DNA
<213> 人工序列

<400> 4
atgaaaaaga cagctatcgc gattgcagtg gcactggctg gtttcgctac cgtagcgcag 60
gcc 63

<210> 5
<211> 15
<212> PRT
<213> 人工序列

<400> 5

Met Lys His Gln His Gln His Gln His Gln His Gln His Gln Gln
 1 5 10 15

<210> 6

<211> 45

<212> DNA

<213> 人工序列

<400> 6

atgaaacacc aacaccaaca tcaacatcaa catcaacatc aacag 45

<210> 7

<211> 30

<212> PRT

<213> 人 (Homo sapiens)

<400> 7

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

Leu Leu Leu Trp His Ser Ala Leu Trp Thr Val Gln Glu Ala
 20 25 30

<210> 8

<211> 615

<212> DNA

<213> 人工序列

<400> 8

atggccggcc ctgccacaca gtcccccattg aagctgatgg ccctgcagct gctgctgtgg 60
 cactccgccc tgtggacagt gcaggaggcc acccctctgg gcccgcagag ctccctgcct 120
 cagtccttcc tgctgaagtg cctggagcag gtgagaaaga tccagggcga cggcgccgcc 180
 ctgcaggaga agctgtgcgc cacatacaag ctgtgccacc ctgaggagct ggtgctgctg 240
 ggccacagcc tgggcatccc ctgggcccct ctgtccagct gccctccca ggccctgcag 300
 ctggccggct gcctgtccca gctgcactcc ggccgtttcc tgtaccaggg cctgctgcag 360
 gccctggagg gcatctcccc cgagctgggc cccacactgg ataccctgca gctggacgtg 420
 gccgatttcg ccaccacaat ctggcagcag atggaggagc tgggcatggc ccctgccctg 480
 cagcctaccc agggcgccat gcctgccttt gcctccgctt ttcagagacg ggccggcggc 540
 gtgctggtgg ccagccacct gcagagcttt ctggagggtgt cctacagagt gctgcggcac 600
 ctggcccagc cttga 615

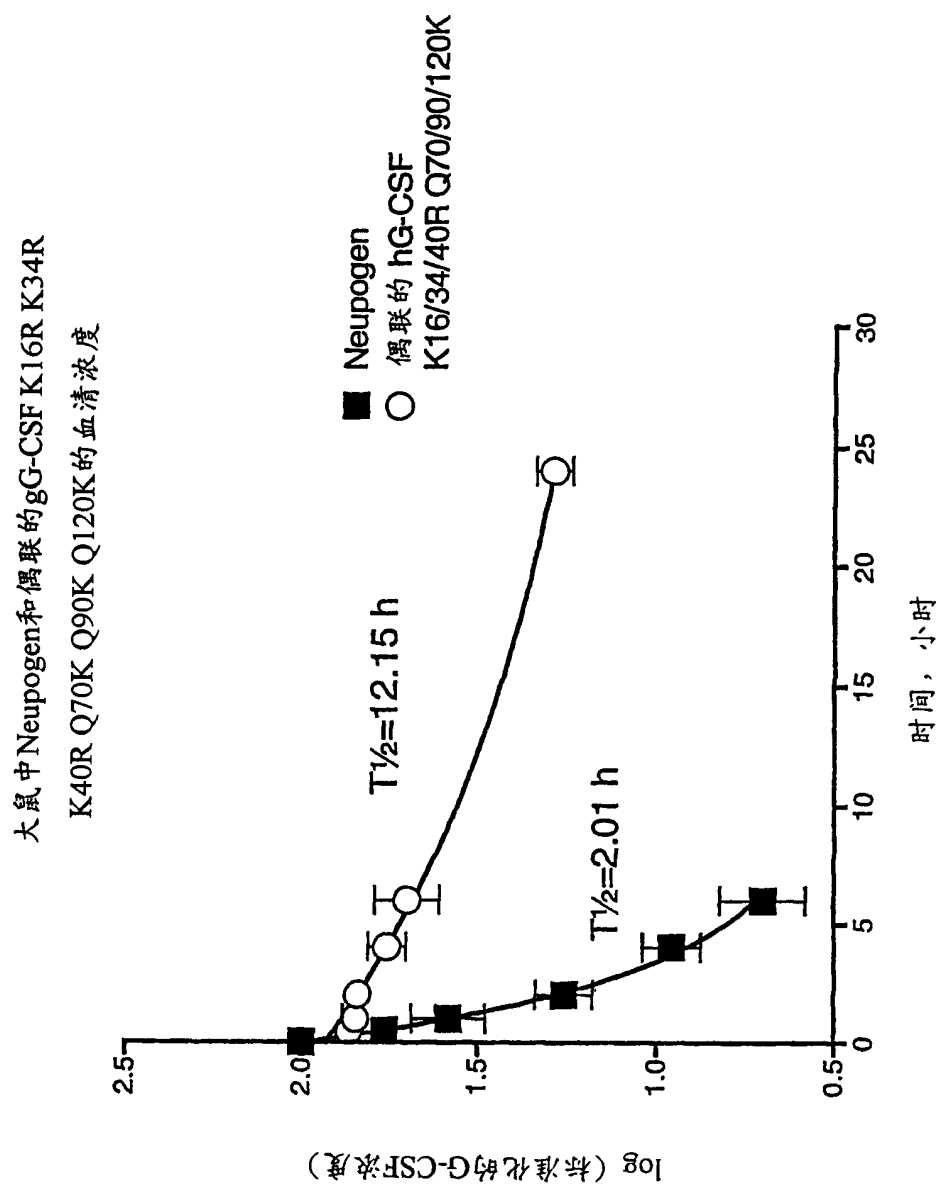


图 1

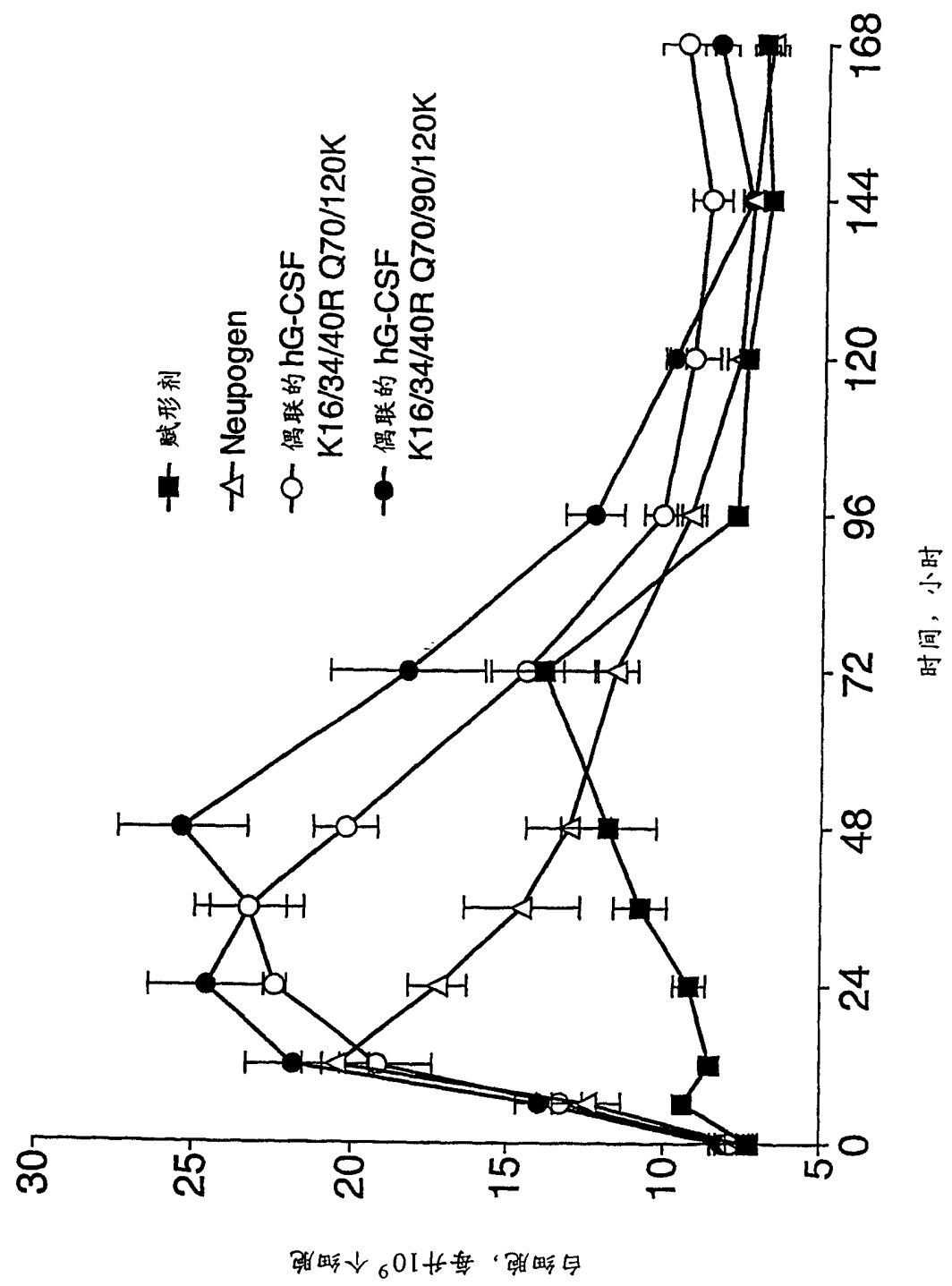


图 2

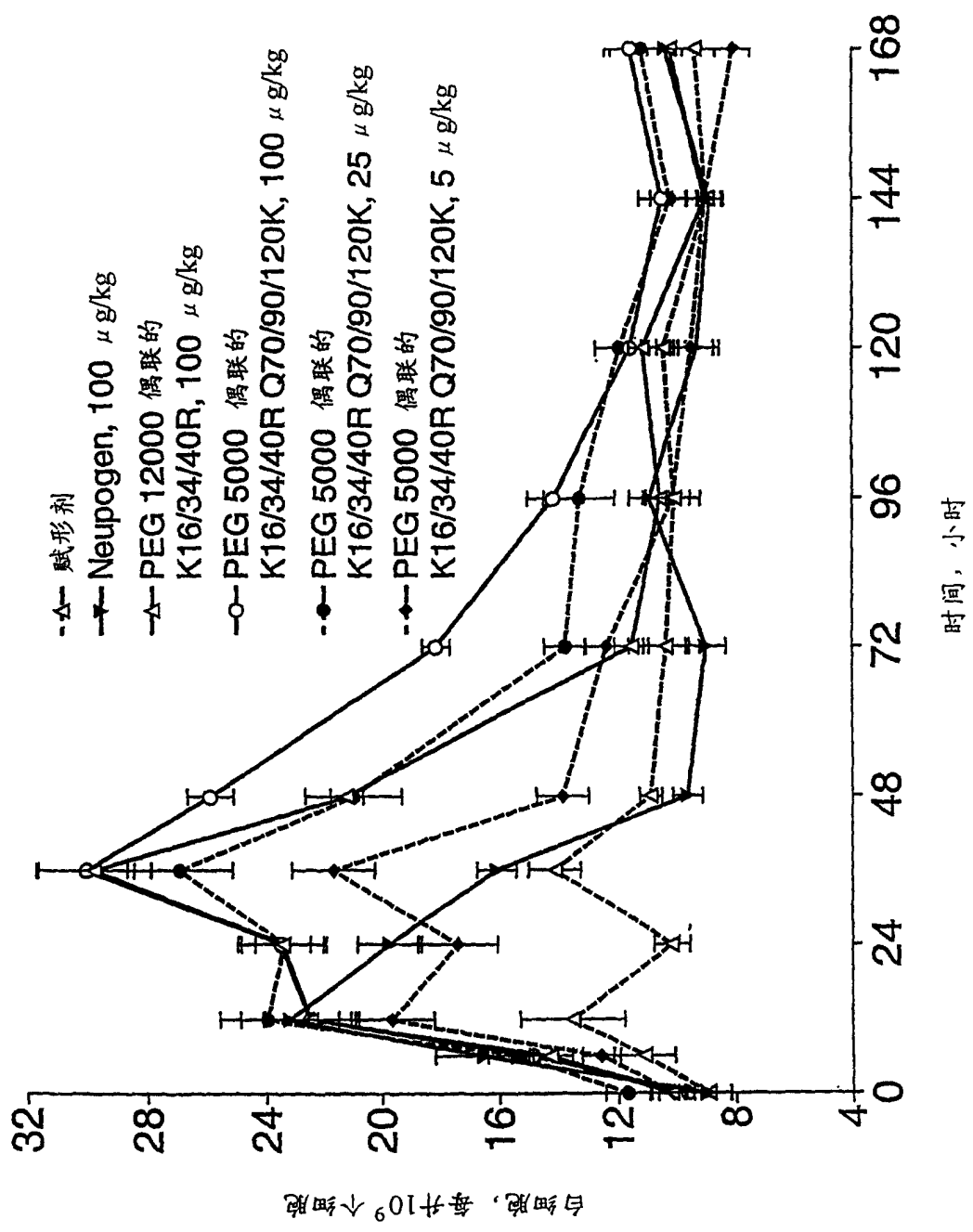


图 3