

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁷

[12] 发明专利申请公开说明书

[21] 申请号 99811157.0

C12N 15/12
C12N 15/20 C07K 14/435
C07K 14/555 A61K 38/17
A61K 38/21

[43]公开日 2001 年 10 月 24 日

[11]公开号 CN 1319137A

[22]申请日 1999.7.21 [21]申请号 99811157.0

[30]优先权

[32]1998.7.21 [33]US [31]60/093,643

[86]国际申请 PCT/US99/16424 1999.7.21

[87]国际公布 WO00/05371 英 2000.2.3

[85]进入国家阶段日期 2001.3.21

[71]申请人 人体基因组科学有限公司

地址 美国马里兰州

[72]发明人 史蒂文·M·鲁宾 保罗·A·穆尔

戴维·W·拉弗勒

[74]专利代理机构 永新专利商标代理有限公司

代理人 林晓红

权利要求书 4 页 说明书 71 页 附图页数 7 页

[54]发明名称 角质形成细胞衍生的干扰素

[57]摘要

本发明涉及是干扰素家族成员的新 KDI 蛋白。具体地,提供了 编码称为“KDI”的人干扰素多肽的分离的核酸分子,还提供了 KDI 多肽,用于产生该多肽的载体,宿主细胞和重组方法。本发明进一步涉及鉴别 KDI 活性的兴奋剂和拮抗剂的筛选方法。本发明还提供了治疗免疫系统相关疾病的治疗方法。

I S S N 1 0 0 8 - 4 2 7 4

权 利 要 求 书

1、一种分离的多核苷酸，其包含至少 95% 相同于选自以下一组的成员的核酸序列：

- (a) 编码 SEQ ID NO: 2 的 1~207 位残基所示多肽的核苷酸序列；
- (b) 编码 SEQ ID NO: 2 的 2~207 位残基所示多肽的核苷酸序列；
- (c) 编码 SEQ ID NO: 2 的 28~207 位残基所示多肽的核苷酸序列；
- (d) 编码 SEQ ID NO: 2 的 30~207 位残基所示多肽的核苷酸序列；
- (e) 编码 SEQ ID NO: 2 的 165~183 位残基所示多肽的核苷酸序列；
- (f) 编码由包含于克隆 HKAPI15 中人 cDNA 编码的完整多肽的核苷酸序列；
- (g) 编码由包含于克隆 HKAPI15 中人 cDNA 编码的除 N 端甲硫氨酸外完整多肽的核苷酸序列；
- (h) 编码由包含于克隆 HKAPI15 中人 cDNA 编码的成熟多肽的核苷酸序列；
- (i) 与以上(a)，(b)，(c)，(d)，(e)，(f)，(g)或(h)中任一核苷酸序列互补的核苷酸序列。

2、一种分离的多核苷酸，其包含选自以下一组的核酸序列：

- (a) 编码 SEQ ID NO: 2 的 1~207 位残基所示多肽的生物学活性片段的核苷酸序列；及
- (b) 与(a)核苷酸序列互补的核苷酸序列。

3、权利要求 1 的核酸分子，其中所述多核苷酸具有编码具有 SEQ ID NO: 2 的 165~183 位氨基酸序列的 KDI 多肽的核苷酸序列，如图 1 (SEQ ID NO: 1) 所示。

4、权利要求 1 的核酸分子，其中所述多核苷酸具有编码具有 SEQ

ID NO: 2 的 28~207 位氨基酸序列的 KDI 多肽的核苷酸序列, 如图 1 (SEQ ID NO: 1) 所示。

5、一种分离的核酸分子, 其包含具有至少 95% 相同于选自以下一组的序列的核苷酸序列的多核苷酸:

(a) 编码包含 SEQ ID NO: 2 的 $n \sim 207$ 位残基的氨基酸序列的多肽的核苷酸序列, 其中 n 是 1~59 范围内的一整数;

(b) 编码包含 SEQ ID NO: 2 的 $1 \sim m$ 位残基的氨基酸序列的多肽的核苷酸序列, 其中 m 是 183~207 范围内的一整数;

(c) 编码具有由 SEQ ID NO: 2 的 $n \sim m$ 位残基组成的氨基酸序列的多肽的核苷酸序列, 其中 n 和 m 是分别在(a)和(b)中限定的整数;

(d) 编码由包含于 HKPI15 中人 cDNA 编码的多肽的核苷酸序列, 其中所述多肽 N 端缺失 1~58 之间的氨基酸;

(e) 编码由包含于 HKAPI15 中人 cDNA 编码的多肽的核苷酸序列, 其中所述多肽 C 端缺失 1~23 之间的氨基酸;

(f) 编码由包含于 HKAPI15 中人 cDNA 编码的多肽的核苷酸序列, 其中所述多肽具有(d)和(e)中所述 N 及 C 端缺失的任意组合。

6、一种分离的核酸分子, 其包含在严格杂交条件下, 与具有相同于权利要求 1(a), (b), (c), (d), (e), (f), (g)或(h)中核苷酸序列的核苷酸序列的多核苷酸杂交的多核苷酸, 其中所述多核苷酸在严格杂交条件下与具有只由 A 残基或 T 残基组成的核苷酸序列的多核苷酸不杂交。

7、一种分离的核酸分子, 其包含编码具有权利要求 1 的(a), (b), (c), (d), (e), (f)或(g)中氨基酸序列的 KDI 多肽携带表位部分的氨基酸序列的多核苷酸。

8、权利要求 7 的分离的核酸分子, 其包含编码选自以下一组的 KDI 多肽携带表位部分的核酸序列:

包含 SEQ ID NO: 2 中大约 Ser49~Ser54 氨基酸残基的多肽;

包含 SEQ ID NO: 2 中大约 Cys59~Ala65 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Pro78~Tyr88 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 His101~Gln113 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Gln120~Glu123 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Cys128~Pro155 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Leu160~Arg168 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Asn171~Asp180 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Val186~Cys193 氨基酸残基的多肽；
包含 SEQ ID NO: 2 中大约 Phe204~Lys207 氨基酸残基的多肽。

9、一种生产重组载体的方法，包括将权利要求 1 的分离的核酸分子插入载体。

10、一种通过权利要求 9 的方法生产的重组载体。

11、一种生产重组宿主细胞的方法，包括将权利要求 10 的重组载体导入宿主细胞。

12、一种通过权利要求 11 的方法生产的重组宿主细胞。

13、一种生产 KDI 多肽的重组方法，包括在所述多肽被表达的条件下培养权利要求 12 的重组宿主细胞，并回收所述多肽。

14、一种分离的 KDI 多肽，其包含至少 95% 相同于选自以下一组的成员的氨基酸序列：

(a)SEQ ID NO: 2 的 1~207 位残基所示多肽；

(b)SEQ ID NO: 2 的 2~207 位残基所示多肽；

(c)SEQ ID NO: 2 的 28~207 位残基所示多肽；

(d)SEQ ID NO: 2 的 165~183 位残基所示多肽；

(e)由包含于克隆 HKAPI15 中人 cDNA 编码的完整多肽；

(f)由包含于克隆 HKAPI15 中人 cDNA 编码的除 N 端甲硫氨酸外的完整多肽；

(g)由包含于克隆 HKAPI15 中人 cDNA 编码的成熟多肽。

15、一种分离的多肽，其包含 KDI 蛋白携带表位部分，其中所述部分选自以下一组：

包含 SEQ ID NO: 2 中大约 Ser49-Ser54 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Cys59-Ala65 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Pro78-Tyr88 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 His101-Gln113 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Gln120-Glu123 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Cys128-Pro155 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Leu160-Arg168 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Asn171-Asp180 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Val186-Cys193 氨基酸残基的多肽；

包含 SEQ ID NO: 2 中大约 Phe204-Lys207 氨基酸残基的多肽。

16、一种分离的抗体，其与权利要求 14 的 KDI 多肽特异结合。

17、一种分离的核酸分子，其包含具有至少 95% 相同于 SEQ ID NO: 1 的至少 50 个连续核苷酸的核苷酸序列的多核苷酸。

18、一种分离的多肽，其包含 SEQ ID NO: 2 的 1~207 位残基所示多肽生物学活性片段的氨基酸序列。

19、一种药物组合物，其包含在药物可接受载体中的权利要求 14 的多肽。

20、一种治疗患者病毒感染的方法，包括给患者施用权利要求 19 的组合物。

说明书

角质形成细胞衍生的干扰素

发明领域

本发明涉及一种编码是干扰素家族成员的多肽的新的人基因，更具体地本发明提供了编码称为“角质形成细胞衍生的干扰素”或“KDI”的人多肽的分离的核酸分子。还提供了 KDI 多肽，生产这种物质的载体、宿主细胞及重组方法。本发明还提供了检测与免疫系统相关疾病的诊断方法，及治疗免疫系统疾病的方法。本发明还涉及鉴别 KDI 的兴奋剂和拮抗剂的筛选方法。

发明背景

干扰素 (IFN) 是由大量真核细胞在各种刺激下分泌的细胞因子家族。干扰素已通过其化学和生物特性分为四类： α -IFN (白细胞)， β -IFN (成纤维细胞) γ -IFN (淋巴细胞)， Ω -IFN (白细胞)。 α 和 β -IFN 已知是 I 型干扰素， γ -干扰素已知是 II 型或免疫干扰素。人基因组中一种单功能基因编码 Ω 干扰素 (Ω -IFN)，其是一种单体糖蛋白，在结构上与 α -IFN 和 β -IFN 远缘相关，但与 γ -IFN 不相关。IFN 呈现具有抗病毒，免疫调节和抗增殖活性。干扰素的临床潜能已被证实，并将在下文总结。

抗病毒：IFN 已用于临床抗病毒治疗，例如用于治疗 AIDS (Lane. Semin. Oncol.18:46-52(Ocl.1991))，病毒性肝炎包括慢性乙肝，丙肝 (Woo, M.H. and Brunakis, T.G., Ann.Parmacother. 31: 330-337(1997 年 3 月); Gibas.A.L.Gastroenterologist,1:129-142(1993.6))，丁肝，乳头瘤病毒 (Levine, L.A.等, Urology 47:553-557(1996.4))，疱疹病毒 (Ho,M.,Annu.Rev.Med.38:51-59(1987))，脑炎病毒 (Wintergerst 等, 感染, 20: 207-212(1992, 7))，及用于预防鼻炎和呼吸道感染

(Ho.M.,Annu.Rev.Med.38:51-59(1987))。

抗寄生虫：IFN 已被建议用于抗寄生虫治疗，例如 γ -IFN 可用于治疗 *Cryptosporidium parvum* 感染 (Rehg,J.E.,J.Infect.Des.174:229-232(1996.7))。

抗细菌：IFN 已临床用于抗菌治疗。例如： γ -IFN 已用于治疗多抗药性肺结核(Condos,R.等, Lancet.349:1513-1515(1997))。

抗癌：干扰素疗法已用于治疗各种癌症（例如，多毛细胞白血病(Hofmann 等, 癌症治疗综述, 12 (增刊.B): 33—37 (1985, 12)) 急性骨髓样白血病 (Stone,R.M.等, Am.J.Clin.Oncol.16:159-163 (1993, 4)), 骨肉瘤 (Strander, H.等, Acta Oncol, 34:877-880(1995)), 基细胞癌 (Dogan, B.等, 癌症信息 91: 215—219(1995, 5)), 神经胶质瘤 (Fetell, M.R.等, 癌症 65: 78—83(1990.1), 肾细胞癌 (Aso.Y.等, Prof.Clin.Biol.Res.303:653 ~ 659(1989)), 多骨髓瘤 (Peest.D.等, Br.J.Haematol. 94:425-432(Sept.1996)), 黑素瘤 (Ikic.D. 等., Int.J.Dermatol.34:872-874(1995.12)) 和 Hodgkin' s 病 (Rybak, M.E.等, J.Biol.Response Mod.9:1-41(1990.2))。 α 干扰素和替莫唑胺组合治疗癌症的增效治疗进展已经报道 (专利公开 WO9712630, Dugan, M.H.)。

免疫疗法：IFN 已临床用于免疫治疗或更特别地用于 (1) 例如预防移植体对宿主的排斥，或缩短自身免疫系统疾病的渐进如关节炎、多发硬化，(2) 糖尿病。 β -IFN 在美国批准销售治疗多发硬化 (即作为免疫抑制剂)。近来有报道多发硬化患者 I 型干扰素和白细胞介素-2 的产生减少 (Wandinger, K.P.等, J.Neurol.Sci.149:87-93(1997))。另外，重组 α -IFN 免疫疗法 (与重组人 IL-2 组合) 已成功用于自体骨髓或造血干细胞移植后的淋巴瘤患者，其可在翻译后促进病情好转 (Nagler,A.等, Blood 89:3951-3959(1997.7))。

抗变态反应：服用 γ -IFN 已用于治疗哺乳动物变态反应 (参

见专利公开 WO8701288, Parkin, J.M.和 Pinching, A.J.) 现也已证实变态反应哮喘患者体内 IL-12 和 IL-12 依赖型 γ -IFN 释放减少(van der Pouw kraan, T.C.等, 免疫杂志 158: 5560-5565 (1997))。因此, IFN 通过抑制体液应答而可用于治疗变态反应。

疫苗佐剂: 干扰素可用作佐剂或辅佐剂以增强或刺激预防或治疗接种情况下的免疫应答。(Heath, A.W.和 Playfair, J.H.L.疫苗 10: 427-434) 1992))。

显然地, 本领域中需要揭示新干扰素蛋白的各种应用, 例如免疫治疗, 以及抗病毒, 抗寄生虫, 抗细菌, 或抗癌治疗中, 或需要提高干扰素活性的症状。

发明概述

本发明提供了分离的核酸分子, 其包括编码具有 SEQ ID NO: 2 所示完整氨基酸序列, 或由以质粒 DNA 在 1998 年 12 月 1 日保藏在 ATCC, 保藏号 203500 的 cDNA 克隆编码的完整氨基酸序列的 KDI 多肽的至少一部分的多核苷酸。图 1 (SEQ ID NO: 1) 所示的通过测序保藏的 KDI 克隆 (HKAPI 15) 测定的核苷酸序列, 含有编码 207 个氨基酸残基的完整多肽的开放读框, 包括在 35-37 位核苷酸的编码 N-末端甲硫氨酸的起始密码子。本发明的核酸分子包括那些编码 SEQ ID NO: 2 所示除 N 末端甲硫氨酸之外的完整氨基酸序列的核酸, 这些分子也能编码融合在 KDI 氨基酸序列 N 末端的额外氨基酸。

编码的多肽具有推导的 27 个氨基酸的前导序列, 在图 1 中下划线处; 推导的成熟 KDI 蛋白质的氨基酸序列在图 1 中如 28~207 位氨基酸所示, 在 SEQ ID NO: 2 中如 1~207 位残基所示。

图 1 (SEQ ID NO: 2) 的 165-183 位残基包括干扰素多肽尤其 KDI 多肽的特征序列。因此, 优选的是包括图 1 的 165-183 位残基的多肽及编码这种多肽的多核苷酸。

因此，本发明一方面提供了分离的核酸分子，其包括含有选自以下一组核苷酸序列的多核苷酸：(a)编码具有 SEQ ID NO: 2 中完整氨基酸序列的 KDI 多肽的核苷酸序列；(b)编码具有除 N 末端甲硫氨酸之外完整 SEQ ID NO: 2 氨基酸序列（即 SEQ ID NO: 2 的 2～207 位残基）的 KDI 多肽的核苷酸；(c)编码 SEQ ID NO: 2 中 28～207 位残基所示成熟 KDI 多肽的核苷酸序列；(d)编码 SEQ ID NO: 2 的 165～183 位残基的核苷酸序列；(e)编码由克隆 HKAPI15 中所含人 cDNA 编码的完整多肽的核苷酸序列；(f)编码由克隆 HKAPI15 中所含的人 cDNA 编码的除 N 末端甲硫氨酸之外完整多肽的核苷酸序列；(g)编码由克隆 HKAPI15 中所含的人 cDNA 编码的成熟多肽的核苷酸序列；和(h)与以上(a),(b), (c), (d), (e), (f)或(g)的任一核苷酸序列互补的核苷酸序列。

本发明的另一实施方案包括分离的核酸分子，其包括具有至少 90%，优选至少 95%，96%，97%，98%或 99%相同于上述(a), (b), (c), (d), (e), (f), (g)或(h)任一核苷酸序列的核苷酸序列的多核苷酸，或具有在严格杂交条件下与(a), (b), (c), (d), (e), (f), (g)或(h) 中多核苷酸杂交的多核苷酸。此杂交的多核苷酸在严格杂交条件下与具有只由 A 或 T 残基组成的核苷酸序列的多核苷酸不杂交。本发明的另一核酸实施方案涉及分离的核酸分子，其包括编码具有上述(a), (b), (c), (d), (e), (f)或(g)氨基酸序列的 KDI 多肽的携带表位部分的氨基酸序列的多核苷酸。

本发明还涉及重组载体，其包括本发明分离的核酸分子，还涉及含有重组载体的宿主细胞，以及生产这种载体和宿主细胞及用它们通过重组技术生产 KDI 多肽或肽的方法。

本发明还提供了分离的 KDI 多肽，其包括选自以下一组氨基酸序列的氨基酸序列：(a)具有 SEQ ID NO: 2 所示完整氨基酸序列的全长 KDI 多肽的氨基酸序列；(b)具有除 N 末端甲硫氨酸之外的 SEQ

ID NO:2 所示完整氨基酸序列的全长 KDI 多肽的氨基酸序列(即 SEQ ID NO: 2 的 2~207 位残基); (c) SEQ ID NO: 2 中 28~207 位残基所示成熟 KDI 多肽的氨基酸序列; (d)SEQ ID NO: 2 的 165~183 位残基所示氨基酸序列; (e)由克隆 HKAPI15 中含有的人 cDNA 编码的全长 KDI 多肽; (f)由克隆 HKAPI15 中含有的人 cDNA 编码的除 N-末端甲硫氨酸之外的全长 KDI 多肽; 和(g)由克隆 HKAPI15 含有的人 cDNA 编码的成熟 KDI 多肽。本发明的多肽还包括具有至少 80%, 优选至少 90%, 更优选 95%, 96%, 97%, 98%或 99%相同于上述(a), (b), (c), (d), (e), (f)或(g)中所述氨基酸序列的氨基酸序列的多肽, 以及与上述那些序列具有至少 90%, 优选至少 95%相似性的氨基酸序列的多肽。

本发明的另一实施方案涉及肽或多肽, 其包括具有以上(a), (b), (c), (d), (e),(f)或(g)所述氨基酸序列的 KDI 多肽的携带表位部分的氨基酸序列。具有本发明 KDI 多肽的携带表位部分氨基酸序列的肽或多肽包括具有至少 6 或 7 个, 优选至少 9 个, 更优选至少 30~50 个氨基酸的这种多肽的一部分, 当然直至并包括本发明多肽的完整氨基酸序列的任何长度的携带表位的多肽也包括在本发明内。

在另一实施方案中, 本发明提供了分离的抗体, 其特异地与具有(a), (b), (c), (d), (e), (f)或(g)所述氨基酸序列的 KDI 多肽结合。本发明还提供了分离与具有本文所述氨基酸序列的 KDI 多肽特异结合的抗体的方法。这种抗体在以下所述治疗中是有效的。

本发明还提供了含 KDI 多肽的药物组合物, 其可用于治疗免疫系统相关疾病如病毒感染, 寄生虫感染, 细菌感染, 癌症, 自身免疫系统疾病, 多发硬化、淋巴瘤及变态反应疾病。本发明还提供了治疗各种需要干扰素多肽的个体的方法。

本发明还提供了含有 KDI 多肽或 KDI 多核苷酸的组合物以施用于体外细胞, 来自体内的细胞及体内细胞, 或多细胞生物体。在本

发明的某些特别优选的实施方案中，组合物含有 KDI 多核苷酸，用于在宿主机体内表达 KDI 多肽以治疗疾病。尤为优选的是在患者体内表达以治疗与干扰素畸变内源活性相关的功能异常。

本发明还提供了鉴别能增强或抑制 KDI 多肽生物活性的化合物的筛选法，其包括将由 KDI 多肽增强的受体在存在 KDI 多肽时与候选化合物相接触，分析例如在候选化合物和 KDI 多肽存在时的抗病毒活性，并将此活性与标准活性水平相对比，标准水平是在受体与 KDI 之间没有候选化合物的情况下相接触分析而得。在此分析中，活性高于标准水平表明候选化合物是 KDI 活性的兴奋剂，活性低于标准水平表明候选化合物是 KDI 活性的拮抗剂。

现已揭示 KDI 在角质形成细胞中表达。因而本发明的核酸用作杂交探针以示差鉴别生物样本中存在的组织或细胞类型。相似地，多肽及其抗体用于提供免疫探针以示差鉴别组织或细胞类型。另外，对上述组织或细胞尤其免疫系统的疾病而言，在取自具有相关的疾病患者的某些组织（例如癌及创作组织）或体液（例如血清、血浆、尿、滑液或脑脊液）中可检测与“标准”KDI 基因表达水平相比明显较高或较低水平 KDI 基因表达，标准 KDI 基因表达水平即在无免疫系统疾病的健康组织中 KDI 的表达水平。因此，本发明提供了在诊断这种疾病中的有效诊断方法，其包括：(a)分析在个体细胞或体液中 KDI 基因表达水平；(b)对比此 KDI 基因表达水平与标准 KDI 基因表达水平，从而根据对比结果高于或低于标准水平来指示免疫系统疾病。

本发明另一方面涉及一种治疗需要机体内干扰素活性水平增高的患者的方法，包括给予这种患者含有治疗有效量的本发明分离的 KDI 多肽或其兴奋剂的组合物。

本发明另一方面还涉及一种治疗需要机体内干扰素活性水平降低的患者的方法，包括给予这种患者含有治疗有效量的 KDI 拮抗剂

的组合物。本发明中使用的优选拮抗剂是 KDI 特异抗体。

附图描述

图 1 示出 KDI 的核苷酸序列(SEQ ID NO: 1)和氨基酸序列(SEQ ID NO: 2)。

图 2 示出 KDI 蛋白质与人 Ω 干扰素 mRNA 的翻译产物 (SEQ ID NO: 3) 的氨基酸序列之间的相同区域, 这通过 Bestfit 计算机程序确定 (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711)。

图 3 示出 KDI 氨基酸序列分析。示出了 α , β , 转角和卷曲区, 亲水性和疏水性; 两亲性区; 柔性区; 抗原指数及表面概率。在“抗原指数—Jameson-Wolf”图中, 正峰值表明 KDI 蛋白的高抗原区的位置, 即从中可获得本发明携带表位的肽。

图 4 示出本发明的 KDI 多肽与干扰素多肽家族一些其它成员的序列对比。所示的是人干扰素 $\beta-1$ (SEQ ID NO: 4), 人胎盘干扰素 (SEQ ID NO: 5), 人 Ω 干扰素 (SEQ ID NO: 6), 人 $\alpha-c$ 干扰素 (SEQ ID NO: 7), 人 $\alpha-F$ 干扰素 (SEQ ID NO: 8), 人干扰素 II-1 (SEQ ID NO: 9), 人 α 干扰素-N (SEQ ID NO: 10), 牛 TP-1 (SEQ ID NO: 11), ovi TP-1 (SEQ ID NO: 12); 猪 TP (SEQ ID NO: 13); 人干扰素 $\beta-2a$ (IL-6) (SEQ ID NO: 14), 牛干扰素 $\beta-2$ (SEQ ID NO: 15), 牛干扰素 $\beta-1$ (SEQ ID NO: 20), 合成干扰素 $\beta-1$ (SEQ ID NO: 21)。此序列对比是通过 Megalign 常规 Clustal 方法用 PAM250 残基重量表产生的。Megalign 包含在 DNA star suite 程序中。

发明的详细描述

本发明提供了分离的核酸分子, 其包括编码具有 SEQ ID NO: 2 所示氨基酸序列的角质形成细胞衍生的干扰素多肽 (KDI) 的多核

苷酸。图 1 所示核苷酸序列 (SEQ ID NO: 1) 是通过对 1998 年 12 月 1 日保藏在美国典型培养物保藏中心 (ATCC), 10801 Boulevard 大学, Manassas, Virginia 20110-2209, 保藏号 ATCC 203500 的人 HKAPI15 cDNA 进行测序而获得的。此保藏的 cDNA 包含在质粒 pCMVSPORT 2.0 中 (Life Technologies, Gaithersburg MD), 并可通过位于人 cDNA 侧翼的 SalI/NotI 限制性酶位点切除。

本发明的 KDI 蛋白质与干扰素家族的一些成员呈序列同源, 特别是人 Ω 干扰素 mRNA 的翻译产物 (图 2) (SEQ ID NO: 3)。 Ω 干扰素已示出抑制各种肿瘤细胞体外增殖, 刺激天然杀伤细胞活性, 增强主要组织相容性 I 类 (非 II 类) 抗原表达, 并抑制用促细胞分裂原或同种异型细胞刺激的淋巴细胞增殖。参见 Adolf, G.R. 人 Ω 干扰素综述, Mult Scler 1995; 1 Suppl 1:S44-S47。已测定 KDI 表达主要在角质形成细胞、树状细胞和单核细胞中发现, 但在角质形成细胞中尤其强。用 TNF- α 或 PolyIC (刺激病毒感染) 刺激角质形成细胞特异并迅速地刺激 KDI 转录物的过表达。基于其结构与 Ω -IFN 相似且在对刺激的病毒感染应答中表达增加, 确认 KDI 呈现 Ω -IFN 和其它干扰素蛋白的许多生物活性, 包括抑制肿瘤增殖, 抗病毒活性, NK 细胞功能及免疫系统增强作用。

核酸分子

除非特别表明, 所有通过对本文 DNA 分子进行测序测定的核苷酸序列均用自动 DNA 测序仪 (如 Model 373 from Applied Biosystems, Inc. Foster City, CA) 测定, 由本文测定的 DNA 分子编码的多肽的所有氨基酸序列通过如上测定的 DNA 序列翻译推导。因而如本领域已知通过自动方法测定的任何 DNA 序列一样, 本文测定的核苷酸序列可有一些误差。自动测定的核苷酸序列与测序 DNA 分子的真实核苷酸序列典型地至少大约 90%, 更典型地至少大约 95%~99.9% 相同。真实序列可通过其它方法更精确测定, 包括本领域熟知的人工 DNA

测序法。本领域也已知测定的核苷酸序列与真实序列相对比有一个插入或缺失将导致核苷酸序列翻译中读框移位，由此由测定的核苷酸序列编码的推导的氨基酸序列从这个插入或缺失点开始将完全不同于由测序的 DNA 分子真实编码的氨基酸序列。

核酸分子或多核苷酸的“核苷酸序列”对 DNA 分子或多核苷酸而言是指脱氧核糖核苷酸的序列，对 RNA 分子或多核苷酸而言是指相应核糖核苷酸（A，G，C 和 U）的序列，其中在特异的脱氧核糖核苷酸序列中的每个胸腺嘧啶脱氧核苷（T）被尿嘧啶核苷（U）取代。

用本文提供的信息，如图 1 中的核苷酸序列，用标准分子生物学方法如用 mRNA 作起始物克隆 cDNA，可获得编码 KDI 多肽的本发明的核酸分子。本发明例证了图 1 所示核酸分子（SEQ ID NO: 1）在衍生自分离的角质形成细胞的 cDNA 文库中发现。

图 1 的 KDI DNA 的核苷酸序列（SEQ ID NO: 1）包括编码 207 个氨基酸残基的蛋白质的一开放读框，起始密码子在图 1 所示核苷酸序列（SEQ ID NO: 1）的 35~37 位核苷酸。SEQ ID NO: 2 所示的 KDI 蛋白质的氨基酸序列大约 35% 相同于 Ω -IFN（图 2）。 Ω -IFN 的序列可在 GenBank 获得，登记号 No.gblA 12140。

熟练技术人员知道由于可能存在上述序列误差，由保藏的 cDNA 编码的包括大约 207 个氨基酸的真正完整 KDI 多肽有时可以较长或较短。一般地，真正开放读框可以在推导自图 1 所示（SEQ ID NO: 1）N 末端甲硫氨酸密码子的 ± 20 个，尤其 ± 10 个氨基酸范围内。

前导和成熟序列

完整 KDI 蛋白的氨基酸序列包括前导序列和成熟蛋白质，如 SEQ ID NO: 2 所示。本发明尤其提供了编码成熟形式 KDI 蛋白质的核酸分子。因此根据信号假说，一旦生长蛋白链经过粗面内质网开始输出，由哺乳动物细胞分泌的蛋白质具有信号或分泌性前导序列，

其裂解自完整多肽以产生分泌的“成熟”形式蛋白质。大多数哺乳动物细胞甚至昆虫细胞以相同特异性裂解分泌的蛋白质。然而在一些情况中，分泌蛋白的裂解不是完全均匀的，则产生两或多种成熟蛋白质。另外，长期以来已知分泌蛋白质的裂解特异性由完整蛋白质的一级结构最终确定，即其在多肽的氨基酸序列中是固有的。因而本发明提供了编码具有由人克隆 HKAPI15（ATCC.保藏号 No.203500）中 cDNA 编码的氨基酸序列的成熟 KDI 多肽的核苷酸序列。“具有由人克隆 HKAPI15 中的 cDNA 编码的氨基酸序列的成熟 KDI 多肽”是指由包含在保藏载体中的克隆的人 DNA 序列编码的完整开放读框在哺乳动物细胞中（如下述 COS 细胞）表达而产生的成熟形式的 KDI 蛋白质。

另外，本发明提供了推测蛋白质是否具有分泌前导序列以及前导序列裂解点的方法。例如 McGeoch 的方法（病毒研究 3: 271—286(1995)）使用来自完整的（未裂解）蛋白质的短 N—末端荷电区及随后的非荷电区的信息。Heinje 的方法（核酸研究 14: 4683～4690(1986)）使用来自围绕裂解位点，典型地—13～+2 残基的残基信息，其中+1 表示成熟蛋白质的氨基末端。这些方法中已知哺乳动物分泌蛋白的推导的裂解位点准确度为 75—80%（Heinje, 前述）。然而，这两种方法对于一给定蛋白质不总是得出相同推导的切割点。

在本发明中，完整 KDI 多肽的推导的氨基酸序列用“PSORT”计算机程序分析，该程序得自京都大学化学研究所的 Kenta Nakai 博士（参见 K.Nakai and M.Kanehisa, Genomics 14:897-911(1992)），其是基于氨基酸序列推测蛋白质的细胞定位的专门系统。作为此计算机推测定位的一部分，可将 McGeoch 和 Heinje 的方法掺入。以上计算机分析推测一个在完整氨基酸序列内的潜在切割位点示于 SEQ ID NO: 2；即在图 1（SEQ ID NO: 2）中 27 和 28 位残基之间。当然，被天然发生的酶使用的切割位点的准确位置与推测的切割位点可有

轻微变化且在物种间可以变化。由此提供了起自大约 20~34 位残基的成熟多肽。尤其地，本发明提供了具有如下 SEQ ID NO: 2 一部分的多肽：SEQ ID NO: 2 中 20~207 位残基，SEQ ID NO: 2 中 21~207 位残基，SEQ ID NO: 2 中 22~207 位残基，SEQ ID NO: 2 中 23~207 位残基，SEQ ID NO: 2 中 24~207 位残基，SEQ ID NO: 2 中的 25~207 位残基，SEQ ID NO: 2 中的 26~207 位残基，SEQ ID NO: 2 中的 27~207 位残基，SEQ ID NO: 2 中的 28~207 位残基，SEQ ID NO: 2 中 29~207 位残基，SEQ ID NO: 2 中 30~207 位残基，SEQ ID NO: 2 中 31~207 位残基，SEQ ID NO: 2 中 32~207 位残基，SEQ ID NO: 2 中 33~207 位残基，SEQ ID NO: 2 中 34~207 位残基。本发明还提供了编码这种多肽的多核苷酸。

本发明的核酸分子可以是 RNA 形式或 DNA 形式。此 DNA 可以是双链或单链的。单链 DNA 或 RNA 可以是编码链，也指作有义链，或其可以是非编码链，也指作反义链。

在特异实施方案中，本发明的多核苷酸长度小于 300kb，200kb，100kb，50kb，15kb，10kb 或 7.5kb。在另一实施方案中，本发明的多核苷酸包括 KDI 编码序列的至少 15 个连续核苷酸，但不包括任何 KDI 内含子的全部或部分。在另一实施方案中，包括 KDI 编码序列的核酸不含有基因组侧翼基因的编码序列（即基因组中 KDI 编码序列的 5' 或 3' ）。

“分离的”核酸分子是指已从其天然环境中移出的核酸分子，DNA 或 RNA。例如，本发明包含于载体中的重组 DNA 分子是分离的。分离的 DNA 分子例如包括保存在异源宿主细胞中的重组 DNA 分子，或溶液中纯化的（部分或基本纯化的）DNA 分子。分离的 RNA 分子包括本发明 DNA 分子的体内或体外 RNA 转录物。然而作为还未与文库（如基因组或 cDNA 文库）的其它成员分离的文库成员（如含有克隆和文库其它成员的均匀溶液形式）的克隆中含有的核酸不

是分离的，或分离或取自细胞或细胞裂解物的染色体（如染色体组型中“染色体涂片”）不是分离的。如本文所述，根据本发明的分离的核酸分子可以是天然、重组或合成产生的。

本发明分离的核酸分子包括具有开放读框（ORF）的 DNA 分子，该读框起始密码子在图 1（SEQ ID NO: 1）所示核苷酸序列 35～37 位。

本发明分离的核酸分子还包括具有缺失 N 末端甲硫氨酸的 SEQ ID NO: 2 的 2～207 位所示 KDI 蛋白编码序列的 DNA 分子。

另外，本发明的分离的核酸分子由于遗传密码的简并包括含有基本不同于上述那些序列，但仍编码 KDI 蛋白的序列的 DNA 分子。当然本领域熟知遗传密码和物种特异密码的优先性。因此，本领域技术人员产生上述简并变体，例如使特殊宿主密码子表达最佳化（例如利用细菌宿主如 E.coli 将人 mRNA 中密码改变为其优选的）是常规技术。

本发明另一方面提供了编码具有由克隆 HKAPI15（ATCC 保藏号 No.203500）中人 cDNA 编码的氨基酸序列的 KDI 多肽的分离的核酸分子。优选地，此核酸分子将编码由上述保藏的人 cDNA 编码的成熟多肽。

本发明还提供了具有图 1（SEQ ID NO: 1）所示核苷酸序列，或具有包含于上述保藏的克隆中的 KDI cDNA 的核苷酸序列的分离的核酸分子，或具有互补于上述序列之一的序列的核酸分子。这种分离的分子，尤其 DNA 分子用于生产本发明的 KDI 多肽，并作为探针检测用载体转染生产 KDI 的细胞中的 mRNA，即作为标记测定宿主细胞中异源基因的表达。

本发明还涉及编码本文所述核苷酸序列一部分的核酸分子，以及所述分离的核酸分子的片段。本发明尤其提供了具有代表 SEQ ID NO: 1 一部分的核苷酸序列的多核苷酸，其由 SEQ ID NO: 1 的 35～

655 位组成。本发明其它特别优选的多核苷酸片段包括或由 SEQ ID

NO: 1 的以下核苷酸残基组成:

38-655, 41-655, 44-655, 47-655, 50-655, 53-655, 56-655, 59-655, 62-655, 65-655, 68-655, 71-655, 74-655, 77-655, 80-655, 83-655, 86-655, 89-655, 92-655, 95-655, 98-655, 101-655, 104-655, 107-655, 110-655, 113-655, 116-655, 119-655, 122-655, 125-655, 128-655, 131-655, 134-655, 137-655, 140-655, 143-655, 146-655, 149-655, 152-655, 155-655, 158-655, 161-655, 164-655, 167-655, 170-655, 173-655, 176-655, 179-655, 182-655, 185-655, 188-655, 191-655, 194-655, 197-655, 200-655, 203-655, 206-655, 209-655, 212-655, 215-655, 218-655, 221-655, 224-655, 227-655, 230-655, 233-655, 236-655, 239-655, 242-655, 245-655, 248-655, 251-655, 254-655, 257-655, 260-655, 263-655, 266-655, 269-655, 272-655, 275-655, 278-655, 281-655, 284-655, 287-655, 290-655, 293-655, 296-655, 299-655, 302-655, 305-655, 308-655, 311-655, 314-655, 317-655, 320-655, 323-655, 326-655, 329-655, 332-655, 335-655, 338-655, 341-655, 344-655, 347-655, 350-655, 353-655, 356-655, 359-655, 362-655, 365-655, 368-655, 371-655, 374-655, 377-655, 380-655, 383-655, 386-655, 389-655, 392-655, 395-655, 398-655, 401-655, 404-655, 407-655, 410-655, 413-655, 416-655, 419-655, 422-655, 425-655, 428-655, 431-655, 434-655, 437-655, 440-655, 443-655, 446-655, 449-655, 452-655, 455-655, 458-655, 461-655, 464-655, 467-655, 470-655, 473-655, 476-655, 479-655, 482-655, 485-655, 488-655, 491-655, 494-655, 497-655, 500-655, 503-655, 506-655, 509-655, 512-655, 515-655, 518-655, 521-655, 524-655, 527-655, 530-655, 533-655, 536-655, 539-655, 542-655, 545-655, 548-655, 551-655, 554-655, 557-655, 560-655, 563-655, 566-655, 569-655, 572-655, 575-655, 578-655, 581-655, 584-655, 587-655, 590-655, 593-655, 596-655, 599-655, 602-655, 605-655, 608-655, 611-655, 614-655, 617-655, 620-655, 623-655, 626-655, 629-655, 632-655, 和 635-655

本发明尤为优选的多核苷酸片段包括或由 SEQ ID NO: 1 的以下核苷酸残基组成:

38-68, 38-71, 38-74, 38-77, 38-80, 38-83, 38-86, 38-89, 38-92, 38-95, 38-98, 38-101, 38-104, 38-107, 38-110, 38-113, 38-116, 38-119, 38-122, 38-125, 38-128, 38-131, 38-134, 38-137, 38-140, 38-143, 38-146, 38-149, 38-152, 38-155, 38-158, 38-161, 38-164, 38-167, 38-170, 38-173, 38-176, 38-179, 38-182, 38-185, 38-188, 38-191, 38-194, 38-197, 38-200, 38-203, 38-206, 38-209, 38-212, 38-215, 38-218, 38-221, 38-224, 38-227, 38-230, 38-233, 38-236, 38-239, 38-242, 38-245, 38-248, 38-251, 38-254, 38-257, 38-260, 38-263, 38-266, 38-269, 38-272, 38-275, 38-278, 38-281, 38-284, 38-287, 38-290, 38-293, 38-296, 38-299, 38-302, 38-305, 38-308, 38-311, 38-314, 38-317, 38-320, 38-323, 38-326, 38-329, 38-335, 38-338, 38-341, 38-344, 38-347, 38-350, 38-353, 38-356, 38-359, 38-362, 38-365, 38-368, 38-371, 38-374, 38-377, 38-380, 38-383, 38-386, 38-389, 38-392, 38-395, 38-398, 38-401, 38-404, 38-407, 38-410, 38-413, 38-416, 38-419, 38-422, 38-425, 38-428, 38-431, 38-434, 38-437, 38-440, 38-443, 38-446, 38-449, 38-452, 38-455, 38-458, 38-461, 38-464, 38-467, 38-470, 38-473, 38-476, 38-479, 38-482, 38-485, 38-488, 38-491, 38-494, 38-497, 38-500, 38-503, 38-506, 38-509, 38-512, 38-515, 38-518, 38-521, 38-524, 38-527, 38-530, 38-

533, 38-536, 38-539, 38-542, 38-545, 38-548, 38-551, 38-554, 38-557, 38-560, 38-563, 38-566, 38-569, 38-572, 38-575, 38-578, 38-581, 38-584, 38-587, 38-590, 38-593, 38-596, 38-599, 38-602, 38-605, 38-608, 38-611, 38-614, 38-617, 38-620, 38-623, 38-626, 38-629, 38-632, 和 38-635

另外，本发明包括含有 SEQ ID NO: 1 的至少大约 30 个，优选至少大约 50 个连续核苷酸的任何部分的多核苷酸。

通常地，具有保藏的 cDNA 的核苷酸序列或图 1 (SEQ ID NO: 1) 所示核苷酸序列的分离的核酸分子的片段是指用作所述诊断探针和引物的长度至少大约 15nt，优选至少大约 20nt，更优选至少大约 30nt，最优选至少大约 40nt 的片段。当然长度为 50~600nt 的较大片段（长度为 400nt, 450nt, 500nt, 550nt 和 600nt 的片段也属于所有长度在 15 和 600nt 之间片段，但是考虑空间特性不特别列举）根据本发明也是有用的，相当于保藏的 cDNA 或图 1 (SEQ ID NO: 1) 所示的大部分（如果不是全部）核苷酸序列的片段也是有用的。例如，长度为至少 20nt 的片段是指包括保藏的 cDNA 的核苷酸序列或图 1 (SEQ ID NO: 1) 所示核苷酸序列的 20 或更多个连续碱基的片段，当然可包括非衍生自 SEQ ID NO: 1（或保藏的 cDNA）的融合在 SEQ ID NO: 1 或保藏的 cDNA 的 20+连续碱基的任一个末端的额外核酸序列。本发明优选的核酸片段包括编码如图 3 中鉴别的 KDI 多肽携带表位部分的核酸分子，并如下详述。

本发明另一方面提供了分离的核酸分子，其包括在严格杂交条件下与上述本发明核酸分子中一部分多核苷酸例如克隆 HKAPI15 (ATCC 保藏号 203500) 中人 cDNA 杂交的多核苷酸。“严格杂交条件”是指在含以下成份的溶液在 42℃ 培养过夜：50% 甲酰胺，5XSSC (150mM NaCl, 15mM 柠檬酸三钠)，50mM 磷酸钠 (pH7.6)，5x Denhardt's 溶液，10% 硫酸葡聚糖，和 20 μg/ml 变性的剪切蛙精 DNA，随后在大约 65℃ 用 0.1×SSC 将滤膜洗涤 2 次。

与多核苷酸“一部分”杂交的多核苷酸是指与相关多核苷酸的

至少大约 15 个核苷酸(nt)，优选至少大约 20nt，更优选至少大约 30 个 nt，最优选至少大约 30—70 个 nt（例如 50 个）杂交的多核苷酸（DNA 或 RNA）。这些多核苷酸如上述用作诊断探针和引物并在下文详述。

例如“长度至少为 20 个 nt”的一部分多核苷酸是指相关多核苷酸的核苷酸序列的 20 或更多个连续核苷酸（如保藏的 cDNA 或图 1（SEQ ID NO: 1）所示核苷酸序列）。当然，只与聚 A 序列（例如图 1（SEQ ID NO: 1）所示 KDI cDNA 的 3' 末端聚（A）段序列）或 T（或 U）残基的互补序列杂交的多核苷酸不包括在本发明用于杂交本发明一部分核酸的多核苷酸内，因为这种多核苷酸将与含有聚（A）序列或其互补序列的任何核酸分子（例如实际上任何双链 cDNA 克隆）杂交。

在特异的实施方案中，本发明的多核苷酸长度小于 100000kb，50000kb，10000kb，1000kb，500kb，400kb，350kb，300kb，250kb，200kb，175kb，150kb，125kb，100kb，75kb，50kb，40kb，30kb，25kb，20kb，15kb，7.5kb，或 5kb。

本发明编码 KDI 多肽的核酸分子可包括但非限于那些通过自身编码完整多肽的氨基酸序列的核酸，及完整多肽的编码序列及其它序列，如那些编码加入的前导分泌序列如前蛋白，蛋白原或前蛋白原序列的序列。

本发明核酸还编码具有额外的，非编码序列的上述蛋白质序列，例如包括但非限于内含子和非编码 5' 和 3' 序列，如转录的，未翻译的在转录，mRNA 加工，包括剪接和聚腺苷酸化信号，例如核糖体结合和 mRNA 的稳定性中起作用的序列；编码额外氨基酸如那些提供其它功能的氨基酸的额外编码序列。

因此，编码多肽的序列可与标记序列，如编码促进融合多肽纯化的肽的序列融合。在本发明此方面的某些优选实施方案中，标记

氨基酸序列是 6-组氨酸肽，如 pQE 载体中提供的标记（QIAGEN, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311），其中许多可商购。如 Gentz 等在 Proc. Natl. Acad. Sci USA 86: 821-824(1989)中所述，6-组氨酸使融合蛋白便于纯化。“HA”标记是用于纯化的另一种肽，其相当于衍生自流感血凝素蛋白的表位，这已由 Wilson 等，细胞 37: 767(1984)中所述。如下所述，其它这种融合蛋白包括在 N-或 C-末端融合 Fc 的 KDI。

变体和突变的多核苷酸

本发明还涉及本发明核酸分子的变体，其编码 KDI 蛋白质的一部分，类似物或衍生物。变体可天然发生如天然等位基因变体。“等位基因变体”是指占据机体染色体上一基因座的基因的几个变化形式之一。参见，基因 II, Lewin, B., ed., John Wiley & Sons, 纽约(1985)。非天然发生的变体可用本领域已知诱变技术产生。

这种变体包括通过核苷酸取代，缺失或加入产生的那些变体。取代，缺失或加入可包括一或多个核苷酸。变体可以在编码区，非编码区或二区均有改变。在编码区内改变可产生保守或非保守氨基酸取代、缺失或加入。其中特别优选的是沉默取代，加入和缺失，其不改变 KDI 蛋白或其部分的性质和活性。保守取代也是特别优选的。

其它实施方案包括一分离的核酸分子，其包括具有至少 90%，优选至少 95%，96%，97%，98%或 99%相同于选自以下一组的多核苷酸的核苷酸序列的多核苷酸：(a)编码具有 SEQ ID NO: 2 中完整氨基酸序列的 KDI 多肽的核苷酸序列；(b)编码具有除 N-末端甲硫氨酸之外，SEQ ID NO: 2 完整氨基酸序列（即 SEQ ID NO: 2 的 2-161 残基）的 KDI 多肽的核苷酸序列；(c)编码具有 SEQ ID NO: 2 的 28-207 位残基所示序列的成熟 KDI 多肽的核苷酸序列；(d)编码 SEQ ID NO: 2 的 165-183 位残基的核苷酸序列；(e)编码由包含

在克隆 HKAPI15 中人 cDNA 编码的完整氨基酸序列的核苷酸序列；
 (f)编码由包含在克隆 HKAPI15 中人 cDNA 编码的除 N-末端甲硫氨酸外完整氨基酸序列的核苷酸序列；(g)编码由包含在克隆 HKAPI15 中人 cDNA 编码的成熟多肽的氨基酸序列的核苷酸序列；和(h)互补于以上(a), (b), (c), (d), (e), (f), (g)和(h)的任何核苷酸序列的核苷酸序列。

本发明的其它实施方案包括分离的核酸分子,其包括具有至少 90 %, 优选至少 95 %, 96 %, 97 %, 98 %或 99 %相同于以上(a) (b) (c) (d) (e) (f) (g)或(h)中任何核苷酸序列的核苷酸序列的多核苷酸, 或在严格杂交条件下与以上(a), (b), (c), (d), (e), (f), (g)或(h)中的多核苷酸杂交的多核苷酸。杂交的多核苷酸在严格杂交条件下, 与具有只由 A 或 T 残基组成的核苷酸序列的多核苷酸不杂交。本发明其它核酸实施方案涉及一分离的核酸分子, 其包括编码具有以上(a), (b), (c), (d), (e), (f)或(g)中氨基酸序列的 KDI 多肽的携带表位部分的氨基酸序列的多核苷酸。

本发明还涉及含有本发明分离的核酸分子的重组载体, 含有重组载体的宿主细胞, 生产这种载体和宿主细胞的方法, 及利用它们通过重组技术生产 KDI 多肽或肽的方法。

具有至少例如 95 % “相同于” 编码 KDI 多肽的相关核苷酸序列的核苷酸序列的多核苷酸是指除了此多核苷酸序列在编码 KDI 多肽的相关核苷酸序列的每 100 个核苷酸中包括至多 5 个点突变外, 此多核苷酸的核苷酸序列相同于相关序列。换言之, 为获得具有至少 95 %相同于相关核苷酸序列的核苷酸序列的多核苷酸, 相关核苷酸序列中至多 5 %的核苷酸可以缺失或用其它核苷酸取代, 或占相关序列中全部核苷酸数的 5 %的核苷酸可以插入相关序列中。相关序列的这些突变可发生在相关核苷酸序列的 5' 或 3' 末端位置, 或发生于末端位置之间的任何地方, 独立散在相关序列的核苷酸之间或在

相关序列内一或多个连续基组中。

事实上，任何特定的核酸分子是否至少 90%，95%，96%，97%，98%或 99%相同于图 1 所示核苷酸序列或保藏的 cDNA 克隆的核苷酸序列，可常规使用已知计算机程序如 Bestfit 程序(Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711)进行测定。Bestfit 利用 Smith 和 Waterman 的局部同源性算法以发现二个序列间的最佳同源区段 (Advances in Applied Mathematics 2:482-489(1981))。当使用 Bestfit 或任何其它序列对比程序以测定一特殊序列是否例如 95%相同于根据本发明的相关序列时，要设立参数，从而可以在全长相关核苷酸序列之上计算相同性百分率，并且允许相关序列中至多全部核苷酸数 5%的同源性缺口。测定查询序列（本发明的序列）和样本序列之间最佳完全匹配，也称作总体序列对比的一种优选方法，可用基本上如 Brutlag 等的公式的 FASTDB 计算机程序测定 (Comp.App.Biosci.(1990)6:237-245)。在一序列对比中，查询和样本序列均是 DNA 序列。RNA 序列可通过 U 转变为 T 加以对比。所述总体序列对比结果以相同性百分率表示。用于 DNA 序列的 FASTDB 序列对比以计算百分率的优选参数是：Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500,或样本核苷酸序列的长度，取两者较短的。

如果样本序列因为 5' 或 3' 缺失而非因为内在缺失而比查询序列短，必须对结果进行手工校正。这是因为 FASTDB 程序当计算相同性百分率时未考虑到样本序列 5' 和 3' 的截短。对相对于查询序列在 5' 或 3' 末端截短的样本序列而言，相同性百分率通过计算样本序列 5' 和 3' 端的未匹配/对比的查询序列的碱基数占查询序列

总碱基数的百分率加以校正。核苷酸是否匹配/对比可通过 FASTDB 序列对比结果确定。然后将此百分数从通过用特殊参数经 FASTDB 程序计算的相同性百分率中减去，以得到最终相同性百分率。此校正结果用于本发明。只计算通过 FASTDB 序列对比展示的不与查询序列匹配/对比的在样本序列 5' 和 3' 碱基之外的碱基以用于人为调节相同性百分率。

例如，90 个碱基的样本序列与 100 个碱基的查询序列相对比测定相同性百分率。缺失发生在样本序列的 5' 端，因而 FASTDB 序列对比不能示出 5' 端头 10 个碱基的匹配/对比。此 10 个未配对的碱基代表序列的 10%（在 5' 和 3' 端末匹配的碱基数/查询序列中碱基总数），所以此 10% 从经 FASTDB 程序计算的相同性百分率中减去。如果所剩 90 个碱基完全匹配则最终相同性百分率为 90%。在另一实施例中，一 90 个碱基的样本序列与 100 个碱基的查询序列相对比。这次缺失是内部缺失，所以在样本序列的 5' 或 3' 末端没有不与查询序列匹配/对比的碱基。在经情况中经 FASTDB 计算的相同性百分率不用手工校正。同样，样本序列只有 5' 和 3' 碱基不与查询序列匹配/对比时，要手工校正。本发明不需要进行其它手工校正。

本发明涉及至少 90%，95%，96%，97%，98% 或 99% 相同于图 1（SEQ ID NO: 1）所示核酸序列的或保藏的 DNA 的核酸序列的核酸分子，不管它们是否编码具有 KDI 活性的多肽。这是因为即使特殊的核酸分子不编码具有 KDI 活性的多肽，本领域技术人员仍知怎样用该核酸分子作为例如杂交探针或聚合酶链反应（PCR）引物。本发明使用的不编码具有 KDI 活性多肽的核酸分子特别包括 cDNA 文库中分离的等位基因变体。

然而优选的是具有至少 90%，95%，96%，97%，98% 或 99% 相同于图 1（SEQ ID NO: 1）所示核酸序列或保藏的 DNA 的核酸

序列的核酸分子，且该分子编码具有 KDI 蛋白质活性的多肽。

“具有 KDI 活性的多肽”是指在特定生物学分析中，活性类似于但非必须相同于本发明 KDI 蛋白活性的多肽。例如本发明的 KDI 蛋白质抑制体外骨髓集落形成并可依据 Tiefenthaler M. 等的方法 (Interferon Cytokine Res, 1997, Jun; 17(6):327-329, 全文并入参考) 分析。另外，根据由 Mire-Sluis A.R. 等报道的分析法 (免疫学方法杂志 1996.4.9; 195 (1-2): 55-61, 并入参考)，KDI 可抑制 GM-CSF 诱导的红白血病细胞系 TF-1 的增殖。或者，通过本领域已知的一些分析方法，例如由 Sugiyama, K 等报道的分析法 (Yakugaku Zasshi 1995 5; 115(5):390-393) 可分析 KDI 经典抗病毒活性。

在上述分析中，本发明的 KDI 蛋白抑制骨髓增殖并示出剂量依赖方式抗病毒活性。因此，“具有 KDI 蛋白活性的多肽”包括在上述分析中也呈剂量依赖方式相同活性的多肽。尽管剂量依赖活性程度不需相同于 KDI 蛋白的剂量依赖程度，但优选地，“具有 KDI 蛋白活性的多肽”与 KDI 蛋白相比给定活性呈基本相似剂量依赖性 (即候选多肽将比相关 KDI 蛋白呈较高活性或其活性降低不超过大约 25 倍且优选不超过 10 倍)。

当然，由于遗传密码的简并，本领域技术人员将立即意识到大量具有至少 90%，95%，96%，97%，98% 或 99% 相同于保藏的 cDNA 的核酸序列或图 1 (SEQ ID NO: 1) 所示核酸序列的核酸分子将编码“具有 KDI 蛋白活性”的多肽。事实上，由于这些核苷酸序列的简并变体都编码相同多肽，即使不进行上述对比分析，本领域技术人员也知此结果。本领域技术人员另外还知晓有适量不是简并变体的这种核酸分子也编码具有 KDI 蛋白活性的多肽。这是因为本领域技术人员充分知道有的氨基酸取代较少或不明显影响蛋白质功能 (例如用另一个脂族氨基酸置换一个脂族氨基酸)。

载体和宿主细胞

本发明还涉及包括本发明分离的 DNA 分子的载体，用重组载体遗传工程化的宿主细胞，及通过重组技术生产 KDI 多肽或其片段。载体例如可以是噬菌体，质粒，病毒或逆转录病毒载体。逆转录病毒可以是可复制型载体或复制缺陷型载体。在后种情况下，病毒增殖通常只在互补宿主细胞中发生。

此多核苷酸可与含选择标志的载体连接以在宿主中增殖。一般地，质粒载体以沉淀物如磷酸钙沉淀物或与荷电脂质复合导入。如果载体是病毒，其可用适当包装细胞系在体外包装，然后转导入宿主细胞中。

DNA 插入片段应有效地与适当启动子连接，如 λ 噬菌体 PL 启动子，E.coli lac, trp, phoA 和 tac 启动子，SV40 早期和晚期启动子和逆转录病毒 LTRs 的启动子。技术人员还知其它适当启动子。表达构建体还包括转录起始位点，终止位点及在转录区内的翻译核糖体结合位点。由构建体表达的转录物的编码部分优选包括恰当地位于被翻译的多肽起始端的翻译起始密码子和多肽末端合适位置的终止密码子（UAA，UGA 或 UAG）。

表达载体优选包括至少一个选择标记。这种标记包括对于真核细胞培养物的二氢叶酸还原酶，G418 或新霉素抗性，对于 E.coli 和其它细菌培养物的四环素，卡那霉素或氨苄青霉素抗性基因。适当的宿主的代表性实例包括但非限于细菌细胞如 E.coli，链霉菌和鼠伤寒沙门氏菌细胞；真菌细胞如酵母细胞；昆虫细胞如果蝇 S2 和 Spodoptera St9 细胞；动物细胞如 CHO，COS，293 和 Bowes 黑素瘤细胞；及植物细胞。本领域已知上述宿主细胞的培养基和培养条件。

优选的细菌载体包括购自 QIAGEN 公司的 pQE70，pQE60 和 pQE09；购自 Stratagene 的 PBS 载体，Phagescript 载体，Bluescript 载体，pNH8A，pNH16a，pNH18A，pNH46A；和购自 Pharmacia 的

ptc99a, pKK223-3, pKK233-3, pDR540, pRIT5。优选的真核载体是购自 Stratagene 的 pWLNEO, pSV2CAT, pOG44, pXT1 和 pSG; 及购自 Pharmacia 的 pSVK3, pBPV, pMSG 和 pSVL。本领域技术人员也知其它适当载体。

构建体导入宿主细胞可通过磷酸钙转染, DEAE-葡聚糖介导的转染, 阳离子脂质介导的转染, 电穿孔, 转导, 感染或其它方法进行。这些方法在许多标准实验手册中均有阐述, 如 Davis 等, 分子生物学基本方法(1986)。

多肽可以修饰形式表达, 如融合蛋白, 且不仅可包括分泌信号, 也可包括其它异源功能区。例如, 额外氨基酸尤其荷电氨基酸区可加入多肽的 N 末端, 以改良其在纯化或随后的操作和贮存期间在宿主细胞中的稳定性及持续性。肽基序也可加入多肽中以促进纯化。这个区域可在最后制备多肽之前除去。将肽基序加入多肽以实现分泌或排泄, 改良稳定性及促进纯化等, 是本领域熟知且常规技术。一优选的融合蛋白包括来自免疫球蛋白的异源区, 其用于稳定和纯化蛋白质。例如, EP-A-0 464 533 (加拿大相应专利 2045869) 揭示了包括各种免疫球蛋白 C 区部分与人体蛋白或其部分的融合蛋白。在许多情况中, 融合蛋白的 Fc 部分用于诊断和治疗是非常有利的, 因此例如产生改良的药物动力学性质 (EP-A023262)。另一方面, 在融合蛋白在以上述有利方式表达, 检测及纯化后, 对于某些应用需缺失 Fc 部分, 这是在当 Fc 部分证实妨碍诊断和治疗时, 例如融合蛋白被用作免疫反应的抗原时的情况中。在药物开发方面, 人体蛋白例如 hIL-5 已经与 Fc 部分融合以通过高通量筛选分析鉴别 hIL-5 的拮抗剂。参见 D.Bennett 等, 分子识别杂志 8: 52-58(1995) 和 K.Johanson 等。生物化学杂志 270: 9459-9471(1995)。

KDI 蛋白通知熟知方法包括硫酸铵或乙醇沉淀, 酸提取, 阳离子或阴离子交换层析, 磷酸纤维素层析, 疏水作用层析, 亲和层析,

羟磷灰石层析和凝集素层析可从重组细胞培养物中回收并纯化。更优选地用高效液相层析（HPLC）进行纯化。本发明的多肽包括：纯化自天然来源包括体液，组织和细胞的产物，其可以是直接分离的或是培养；及通过重组技术产自原核或真核宿主如细菌，酵母，高等植物，昆虫和哺乳动物细胞的产物。根据用于重组生产过程的宿主，本发明的多肽可以是糖基化的或非糖基化的。另外，本发明的多肽也可包括起始的修饰的甲硫氨酸残基。因此，本领域熟知由翻译起始密码子编码的 N-末端甲硫氨酸一般在所有真核细胞中翻译之后，高效地从任何蛋白质中除去。尽管大多数蛋白质上的 N-末端甲硫氨酸也能有效地在大多数原核细胞中除去，但根据 N-末端甲硫氨酸共价键连的氨基酸的性质，对一些蛋白质而言此原核细胞除去过程是无效的。

多肽和片段

本发明还提供了具有由保藏的 DNA 编码的氨基酸序列，或 SEQ ID NO: 2 中氨基酸序列的分离的 KDI 多肽，或含有上述多肽一部分的肽或多肽。

变体和突变多肽

为改良或改变 KDI 多肽的特性，可应用蛋白质工程。本领域技术人员已知的重组 DNA 技术可用于产生新的突变体蛋白或“突变蛋白”，包括单或多个氨基酸取代，缺失，加入或融合蛋白。这种修饰的蛋白可示出例如增强的活性或提高的稳定性。另外，它们可以高产量纯化并比相应的天然多肽至少在一定纯化和贮存条件下呈现更好的溶解性。

N-末端和 C-末端缺失突变体

例如，对一些蛋白质包括膜相关蛋白的胞外区或分泌蛋白的成熟形式而言，本领域已知一或多个氨基酸可从 N-末端或 C-末端缺失而基本不丧失生物功能。例如，Ron 等在生物化学杂志 268:

2984~2988(1993)报道了修饰的 KGF 蛋白即使缺失 3, 8 或 27 个氨基末端的氨基酸残基, 仍具有肝素结合活性。在本文中, 由于本发明的蛋白质是干扰素多肽家族成员, SEQ ID NO: 2 N 末端氨基酸直至 SEQ ID NO: 2 所示 59 位半胱氨酸的缺失仍可保留一些生物活性如抗病毒活性或抑制骨髓增殖。具有包括 SEQ ID NO: 2 中 Cys-59 残基的进一步 N 末端缺失的多肽不能保留这种生物功能, 因为干扰素相关多肽中的此残基在一些 (非全部) 家族成员中是保守的, 紧邻其的亮氨酸残基 (60 位残基) 也是如此。在 59 位的半胱氨酸残基被认为是形成二硫键之所需, 以保证为受体结合及信号转导之所需的结构稳定性。

然而, 即使一或多个氨基酸从蛋白质 N 末端缺失导致蛋白质一或多种生物功能丧失变化, 其它生物活性仍可保留。因此, 缩短的蛋白质诱导和/或结合识别完整蛋白质的抗体的能力, 当少于半数的完整蛋白质残基从 N 末端除去时一般仍被保留。缺失完整蛋白质 N 末端残基的特定多肽是否保留这种免疫活性可通过本文所述常规方法及本领域已知其它方法测定。

因此本发明还提供了具有一或多个缺失自 SEQ ID NO: 2 所示 KDI 氨基酸序列氨基末端直至 Cys-59 的残基的多肽, 及编码这种多肽的多核苷酸。本发明尤其提供了包括 SEQ ID NO: 2 的 n-207 位残基的氨基酸序列的多肽, 其中 n 是 1~59 范围内的整数, 且 Cys-59 是确信为 KDI 蛋白活性所需的起自完整 KDI 多肽 (如 SEQ ID NO: 2 所示) N-末端的第一位残基。

本发明尤其提供了具有 SEQ ID NO: 2 的如下残基的氨基酸序列的多肽: 1~207, 2~207, 3~207, 4~207, 5~207, 6~207, 7~207, 8~207, 9~207, 10~207, 11~207, 12~207, 13~207, 14~207, 15~207, 16~207, 17~207, 18~207, 19~207, 20~207, 21~207, 22~207, 23~207, 24~207, 25~207, 26~207, 27~207,

28~207, 29~207, 30~207, 31~207, 32~207, 33~207, 34~207, 35~207, 36~207, 37~207, 38~207, 39~207, 40~207, 41~207, 42~207, 43~207, 44~207, 45~207, 46~207, 47~207, 48~207, 49~207, 50~207, 51~207, 52~207, 53~207, 54~207, 55~207, 56~207, 57~207, 58~207 和 59~207, 也提供了编码这些多肽的多核苷酸。

相似地, 已知许多 C 末端缺失突变蛋白的生物功能的实例。例如, γ -干扰素通过缺失蛋白质羧基末端 8~10 个氨基酸残基呈现活性提高接近 10 倍(Dobeli 等, 生物技术学杂志 7: 199-216(1988))。在本发明中, 由于本发明的蛋白质是干扰素多肽家族成员, SEQ ID NO: 2 的 C 末端氨基酸直至 183 位的色氨酸残基缺失仍可保留一些生物活性如抗病毒活性或抑制骨髓增殖。具有包括 SEQ ID NO: 2 的 Ile-183 的进一步 C-末端缺失的多肽呈现不能保留这种生物活性, 因为已知干扰素相关多肽中此残基在许多成员中是保守的, 且被认为对受体结合和信号转导是重要的。另外, 在 181 位半胱氨酸是高保守的且已知是干扰素家族成员抗病毒活性所需的。

然而, 即使蛋白质 C 末端的一或多个氨基酸缺失导致一或多种生物活性丧失, 但仍可保留其它生物活性。因此缩短的蛋白质诱导和/或结合识别完整蛋白质的抗体的能力。当少于半数的完整蛋白质的残基从 C 末端除去时一般被保留。缺失完整蛋白质 C 末端残基的特定多肽是否保留这种免疫活性通过本文所述常规方法及本领域已知其它方法可易于测定。

因此本发明还提供了具有缺失自 SEQ ID NO: 2 所示 KDI 氨基酸序列的羧基末端直至 Trp-183 的一或多个氨基酸的多肽及编码这种多肽的多核苷酸。尤其地本发明提供了具有 SEQ ID NO: 2 氨基酸序列 1~m 残基的氨基酸序列的多肽, 其中 m 是 183~207 范围内的任何整数, 且残基 Trp-183 是确认为 KDI 蛋白活性所需的完整 KDI

多肽（如 SEQ ID NO: 2 所示）C 末端的第一个残基位。

更为具体地，本发明提供了具有 SEQ ID NO: 2 的以下残基的氨基酸序列的多肽：1-183, 1-184, 1-185, 1-186, 1-187, 1-188, 1-189, 1-190, 1-191, 1-192, 1-193, 1-194, 1-195, 1-196, 1-197, 1-198, 1-199, 1-200, 1-201, 1-202, 1-203, 1-204, 1-205, 1-206 和 1-207。还提供了编码这些多肽的多核苷酸。

本发明还提供了具有一或多个缺失自氨基和羧基末端的氨基酸的多肽，其通常被描述为具有 SEQ ID NO: 2 的 n-m 残基，其中 n 和 m 是如上所述的整数。另外，本发明提供了任选地具有 N 末端甲硫氨酸的这些突变体多肽。此多肽因而也可用公式 X-n-m 代表，其中 X 是 NH₂ 或 Met, n 和 m 是如上述的整数。当然也提供了编码这些多肽的多核苷酸。

更为具体地，本发明优选提供了具有 SEQ ID NO: 2 的以下残基的氨基酸序列的多肽：

20-183, 21-183, 22-183, 23-183, 24-183, 25-183, 26-183, 27-183, 28-183, 29-183, 30-183, 31-183, 32-183, 33-183, 34-183, 35-183, 36-183, 37-183, 38-183, 39-183, 40-183, 41-183, 42-183, 43-183, 44-183, 45-183, 46-183, 47-183, 48-183, 49-183, 50-183, 51-183, 52-183, 53-183, 54-183, 55-183, 56-183, 57-183, 58-183, 59-183, 20-184, 21-184, 22-184, 23-184, 24-184, 25-184, 26-184, 27-184, 28-184, 29-184, 30-184, 31-184, 32-184, 33-184, 34-184, 35-184, 36-184, 37-184, 38-184, 39-184, 40-184, 41-184, 42-184, 43-184, 44-184, 45-184, 46-184, 47-184, 48-184, 49-184, 50-184, 51-184, 52-184, 53-184, 54-184, 55-184, 56-184, 57-184, 58-184, 59-184, 20-185, 21-185, 22-185, 23-185, 24-185, 25-185, 26-185, 27-185, 28-185, 29-185, 30-185, 31-185, 32-185, 33-185, 34-185, 35-185, 36-185, 37-185, 38-185, 39-185, 40-185, 41-185, 42-185, 43-185, 44-185, 45-185, 46-185, 47-185, 48-185, 49-185, 50-185, 51-185, 52-185, 53-185, 54-185, 55-185, 56-185, 57-185, 58-185, 59-185, 20-186, 21-186, 22-186, 23-186, 24-186, 25-186, 26-186, 27-186, 28-186, 29-186, 30-186, 31-186, 32-186, 33-186, 34-186, 35-186, 36-186, 37-186, 38-186, 39-186, 40-186, 41-186, 42-186, 43-186, 44-186, 45-186, 46-186, 47-186, 48-186, 49-186, 50-186, 51-186, 52-186, 53-186, 54-186, 55-186, 56-186, 57-186, 58-186, 59-186, 20-187, 21-187, 22-187, 23-187, 24-187, 25-187, 26-187, 27-187, 28-187, 29-187, 30-187, 31-187, 32-187, 33-187, 34-187, 35-187, 36-187, 37-187, 38-187, 39-187, 40-187, 41-187, 42-187, 43-187, 44-187, 45-187, 46-187, 47-187, 48-187, 49-187, 50-187, 51-187, 52-187, 53-187, 54-187,

[illegible]

[illegible]

前述每个多肽可另外包括 N-末端甲硫氨酸残基。也提供了编码每个具有或不具有 N-末端甲硫氨酸残基的这些多肽的多核苷酸。

本发明还包括由克隆 HKAPI15 中人 cDNA 编码的完整 KDI 氨基酸序列的一部分组成的多肽，其中此部分不包括克隆 HKAPI15 中人 cDNA 编码的完整氨基酸序列氨基末端 1~大约 58 位氨基酸，或羧基末端 1~大约 23 位氨基酸，或克隆 HKAPI15 中人 cDNA 编码的完整氨基酸序列的上述氨基酸末端和羧基末端缺失的任何组合。本发明还提供了编码所有上述缺失突变体多肽的多核苷酸。

其他突变体

除上述蛋白质末端缺失形式之外，本领域技术也将意识到 KDI 多肽的一些氨基酸序列可不明显影响蛋白质结构或功能而加以改变。如果涉及这种序列中的差异，应记得在蛋白质上有确定活性的关键区。

因此，本发明还包括 KDI 多肽的变体，其呈现 KDI 多肽的活性，或其包括 KDI 蛋白质的如下所述部分的区域。这种突变体包括缺失，插入，倒位，重复及根据本领域已知一般规则选择的取代类型，对活性只有轻微影响。例如，Bowie, J.V.等在“译解蛋白质序列中信息：对氨基酸取代的耐受性”，科学 247: 1306—1310(1990)中提供了关于怎样产生表型沉默氨基酸取代的指导，其中作者指明有两种主要方法研究氨基酸序列对改变的耐受性。第一种方法依赖于进化过程，其中通过自然选择接纳或排斥突变。第二种方法用遗传工程在克隆的基因的特殊位置进行氨基酸调换，并进行选择或筛选以鉴别保留功能的序列。

如作者所述，这些研究揭示蛋白质对氨基酸取代非常耐受。作者还指明氨基酸调换似乎在蛋白质的一定位置是允许的。例如，大多数隐蔽的氨基酸残基要求非极性侧链，而表面侧链的特征一般很

少是保守的。其它这种表型沉默取代见于 Bowie, J.U.等所述, 并在此并入参考。典型的保守取代是脂肪族氨基酸 Ala, Val, Leu 和 Ile 之间的彼此置换; 羟基残基 Ser 和 Thr 的互换, 酸性残基 Asp 和 Glu 的置换, 酰胺残基 Asn 和 Gln 间的取代, 碱性残基 Lys 和 Arg 的置换, 及芳香族残基 Phe, Tyr 间的置换。

因此, SEQ ID NO: 2 的多肽或由保藏的 cDNA 编码的多肽的类似物或衍生物或片段可以是(i)其中一或多个氨基酸残基是用保守或非保守的氨基酸残基(优选保守的氨基酸残基)取代的, 且这种取代的氨基酸残基可以是或不是由遗传密码编码的, 或(ii)其中一或多个氨基酸残基包括取代基团, 或(iii)其中 KDI 多肽是与其它化合物如提高多肽半衰期的化合物(例如聚乙二醇)融合的, 或(iv)其中额外氨基酸与以上形式的多肽融合的, 如 IgG Fc 融合区肽或前导或分泌序列或用于纯化以上形式多肽的序列或蛋白原序列。这种片段, 衍生物及类似物根据本文教导被本领域技术人员认为包括在本发明内。

因此, 本发明的多肽可以包括通过自然突变或人为操作的一或多个氨基酸取代, 缺失或加入。如上所述, 氨基酸改变优选是较小的, 如不明显影响蛋白质折叠或活性的保守氨基酸取代。

本发明 KDI 蛋白质中为功能所必需的氨基酸可以通过本领域已知方法鉴别, 如定点诱变或丙氨酸扫描诱变(Cunningham 和 Wells, 科学 244: 1081~1085(1989)。后一种方法在分子中的每个残基引入单丙氨酸突变。然后测试所得突变体分子的生物活性, 如受体结合或体外增殖活性。

特别感兴趣的是用可生产具有所需明显改良特性如较少聚集性的其它极性或中性氨基酸取代极性氨基酸。聚集在制备药物配方时不仅降低活性还会产生问题, 因为聚集物可以是免疫原性的。

(Pinckard 等, Clin . Exp. Immunol, 2:331-340(1967); Robbins 等,

Diabetes 36:838-845(1987);Cleland 等, Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377(1993)。

氨基酸的置换也可改变配体对细胞表面受体结合的选择性。例如, Ostade 等在自然 361: 266—268 (1993) 中阐述了导致 TNF- α 仅与两种已知 TNF 受体之一选择性结合的一些突变。配体—受体结合的临界位点也可通过结构分析确定, 如结晶, 核磁共振或光亲和标记。(Smith 等, 分子生物学杂志 224: 899—904 (1992) 及 Vos 等, 科学 255: 306—312(1992))。

本文所述的每个 KDI 多肽的特别优选的取代是在 192 位的精氨酸残基用赖氨酸置换(后文有时称为“R192K”), 及在 193 位的半胱氨酸残基用丝氨酸置换(后文有时称为“C193S”)。这些取代可在单独的 KDI 多肽中发现, 或者它们可在相同 KDI 多肽中发生。编码所有前述 KDI 多肽的多核苷酸均含有取代。

本发明的多肽优选是分离形式的且优选是基本纯化的。重组产生的 KDI 多肽可以通过 Smith 与 Johnson 在基因 67: 31—40 (1988) 中所述的一步法而基本纯化。本发明的多肽也可以通过本领域熟知的蛋白质纯化方法, 用本发明的抗 KDI 抗体从天然或重组来源中纯化。

本发明还提供了含有选自以下一组组成的氨基酸序列的分离的 KDI 多肽: (a)具有 SEQ ID NO: 2 所示完整氨基酸序列的全长 KDI 多肽的氨基酸序列; (b)具有 SEQ ID NO: 2 所示完整氨基酸序列除 N—末端甲硫氨酸外的全长 KDI 多肽的氨基酸序列(即 SEQ ID NO: 2 的 2~207 位残基); (c) SEQ ID NO: 2 所示 28~207 位残基所示成熟 KDI 多肽的氨基酸序列; (d) SEQ ID NO: 2 中 165~183 位残基所示氨基酸序列; (e)由克隆 HKAPI15 中包含的人 cDNA 编码的全长 KDI 多肽; (f)由克隆 HKAPI15 中包含的人 cDNA 编码的除 N 末端甲硫氨酸外全长 KDI 多肽; 及(g)由克隆 HKAPI15 中包含的人

cDNA 编码的成熟 KDI 多肽。

本发明的多肽还包括与上述那些多肽至少 90%，优选至少 95%，更优选至少 96%，97%，98%或 99%相似的多肽。本发明的多肽还包括与由保藏的 DNA 编码的多肽或 SEQ ID NO: 2 的多肽至少 80%，优选至少 90%或 95%，更优选至少 96%，97%，98%或 99%相同的多肽，且也包括具有这种多肽至少 10，20 或 30 个氨基酸且更优选至少 50 个氨基酸的一部分。

本发明的另一实施方案涉及一种肽或多肽，其包括具有含有至少一个保守氨基酸取代，但不超过 50 个，优选不超过 40 个，更优选不超过 30 个，特别优选不超过 20 个保守氨基酸取代的氨基酸序列的 KDI 多肽的氨基酸序列。当然，这种肽或多肽最优选的是包括 KDI 多肽氨基酸序列的氨基酸序列，其含有至少一个，但不超过 10，9，8，7，6，5，4，3，2 或 1 个保守氨基酸序列。

两种多肽的“相似百分率”是指用 Bestfit 程序(Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575, Science Drive, Madison, WI 53711)及确定相似性的默认设定，对比两种多肽的氨基酸序列产生的相似性值。Bestfit 用 Smith 与 Waterman 的局部同源性算法 (Advances in Applied Mathematics 2:482-489,1981)，以发现两序列间相似的最佳区段。

具有至少例如 95% “相同于” 参照的 IL-20 多肽的氨基酸序列的氨基酸序列的多肽是指除了此多肽序列在参照的 IL-20 多肽的氨基酸序列的每 100 个氨基酸中包括至多 5 个氨基酸改变，此多肽的氨基酸序列相同于查询序列。换言之，为获得具有至少 95%相同于查询氨基酸序列的氨基酸序列的多肽，查询序列中至多 5%的氨基酸残基可以缺失或用其它氨基酸取代，或占总氨基酸残基数 5%的氨基酸可插入参照序列。参照序列的这些变化可发生在参照序列的

氨基或羧基末端位置或这些末端位置之间的任何地方，独立散在参照序列的残基间或在参照序列间的一或多个连续组中。

事实上，任何特殊多肽是否至少 90%，95%，96%，97%，98% 或 99% 相同于例如 SEQ ID NO: 2 所示氨基酸序列或由保藏的 cDNA 克隆编码的氨基酸序列，通常可用已知计算机程序如 Bestfit 程序（Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711）确定。当用 Bestfit 或其它序列对比程序确定一特殊序列是否例如 95% 相同于本发明的参照序列时，当然要设立参数，如此在全长参照序列之上计算相同性百分率，且允许参照序列中有至多占总氨基酸残基数 5% 的同源缺口。一种优选的确定查询序列（本发明的序列）与样本序列间最佳总体配对的方法，也称作总体序列对比，可以用基本上如 Brutlay 等的算法的 FASTDB 计算机程序确定（Comp. App. Biosci(1990)6;237-245）。在序列对比中，查询和样本序列均是核苷酸序列或均是氨基酸序列。所述总体序列对比的结果以相同性百分率表示。用于 FASTDB 氨基酸序列对比中的优选参数是：Matrix=PAM, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=序列长, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 或样本氨基酸序列的长度，取较短的。

如果样本序列由于 N-或 C-末端缺失而不是因为内部缺失而比查询序列短，必须对结果进行手工校正，这是因为 FASTDB 程序当计算总体相同性百分率时未考虑到样本序列 N-和 C-末端截短。对与查询序列相比在 N 端和 C 端截短的样本序列，相同性百分率通过计算样本序列 N 和 C-端的未匹配/对比的查询序列残基数占查询序列总碱基百分率加以校正。残基是否配对/对比通过 FASTDB 序列对比结果可确定。然后将此百分数从相同性百分率中减去以得到最

终相同性百分率范围，相同性百分率是用特殊参数通过以上 FASTDB 程序计算的。此最终相同性百分率即为本发明所用。只有未与查询序列配对/对比的样本序列的 N 和 C 端残基需要人为调节相同性百分率范围。即只有查询残基位在样本序列最远的 N 和 C 端残基之外。

例如，90 个氨基酸残基的样本序列与 100 个残基的查询序列对比确定相同性百分率。缺失发生在样本序列的 N 端，因而 FASTDB 序列对比未示出 N 端头 10 个残基的配对/对比。这 10 个未配对的残基代表序列的 10%（在 N 和 C 端未配对的残基数/查询序列中总残基数），因此这 10% 从通过 FASTDB 程序计算的相同性百分率中减去。如果剩余的 90 个残基充分配对，则最终相同性百分率为 90%。在另一实施例中，90 个残基的样本序列与 100 个残基的查询序列相对比。这次缺失是内部缺失，因而在样本序列的 N 或 C 端没有不与查询序列配对/对比的残基。在这种情况下，通过 FASTDB 计算的相同性百分率不用手工校正。同样，只有样本序列 N 和 C 端外侧的残基位不与查询序列配对/对比时，通过 FASTDB 计算结果需要手工校正。本发明不需其它手工校正。

本发明的多肽可用本领域技术人员熟知的方法用作 SDS-PAGE 凝胶上或分子筛凝胶过滤柱上的分子量标记。

如下所述，本发明的多肽还可用于产生多克隆和单克隆抗体，其用于如下所述检测 KDI 蛋白质表达的分析中，或作为能增强或抑制 KDI 蛋白质功能的兴奋剂与拮抗剂。另外，这种多肽可用于酵母双重杂交系统以“捕捉”KDI 蛋白结合蛋白，其也是根据本发明的候选兴奋剂和拮抗剂。此酵母双重杂交系统见于 Fields 和 Song，自然 340：245—246（1989）中所述。

携带表位部分

另一方面，本发明提供了包括本发明多肽携带表位部分的肽或多肽。此多肽部分的表位是本发明多肽的免疫原性或抗原性表位。

“免疫原性表位”是指当整个蛋白质是免疫原时激发抗体应答的蛋白质的一部分。另一方面，“抗原性表位”是指抗体能结合的蛋白质分子的一区域。蛋白质免疫原性表位数通常少于抗原性表位数。例如参见于 Geysen 等, Proc.Natl.Acad. Sci. USA 81:3998-4002(1983)。

为选择具有抗原性表位的肽或多肽（即含有蛋白质分子的抗体能结合的区域），本领域熟知模拟蛋白质序列部分的相对短的合成肽通常能激发与部分模拟的蛋白质反应的抗血清。例如见于 Sutcliffe, J.G.Shinnick, T.M., Green, N. and Learner, R.A.(1983) “与蛋白质上预确定位点反应的抗体”，科学 219：660～666。能激发蛋白质反应性血清的肽通常以蛋白质的一级序列代表，能通过一系列简便化学规则定性，并被限定为既不是完整蛋白质的免疫优势区（即免疫原性表位），也不是氨基或羧基端。本发明的具有抗原性表位的肽与多肽因而用于产生抗体，包括与本发明多肽特异结合的单克隆抗体。例如参见 Wilson 等，细胞 37：767—778（1984），在 777 页。

本发明的携带抗原性表位的肽和多肽优选含有至少 7 个，更优选至少 9 个，最优选大约 15～30 个包含于本发明多肽的氨基酸序列中的氨基酸的序列。可用于产生 KDI 特异抗体的抗原性多肽或肽的非限制性实例包括：包含 SEQ ID NO：2 中大约 Ser49～Ser54 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Cys59～Ala65 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Pro78～Tyr88 氨基酸残基的多肽；包含 SEQ ID NO：2 中 His101～Gln113 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Gln120～Glu123 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Cys128～Pro155 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Leu160～Arg168 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Asn171～Asp180 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Val186～Cys193 氨基酸残基的多肽；包含 SEQ ID NO：2 中 Phe 204～Lys207 氨基酸残基的多肽。这些多肽片段通过 Jameson-Wolf 抗原性指数分析已确

定其携带 KDI 蛋白质的抗原性表位，如图 3 所示。

本发明的携带表位的肽和多肽可通过常规方式生产。例如见于 Houghten, R.A.(1985) “快速固相合成大量肽的常用方法：在单独氨基酸水平抗原—抗体相互作用的特异性” Proc. Natl. Acad. Sci. USA 82:5131~5135；此“同时大量肽合成(SMPS)”法还见于 Houghten 等(1986)在美国专利 No.4631211 中所述。

本发明的携带表位的肽与多肽根据本领域熟知方法用于诱导抗体。例如见于 Sutcliffe 等；前述，Wilson 等；前述，Chow, M 等 Proc. Natl. Acad. Sci. USA 82:910-914；及 F.J.等，J.Gen. Virol 66:2347-2354(1985)所述。本发明的携带免疫原性表位的肽，即当整个蛋白质是免疫原时，激发抗体应答的蛋白质的那些部分，根据本领域已知方法鉴别。例如见于 Geysen 等所述。另外，Geysen(1990)的美国专利 No.5194392 阐述了一种检测或确定是表位拓扑学等价物（即“mimotope”）的单体（氨基酸或其它化合物）序列的方法，该单体序列互补于相应抗体的特殊互补位（抗原结合位点）。Geysen(1989)的美国专利 No.443092 阐述了一种检测或确定是配体拓扑等价物的单体的序列的方法，该单体序列互补于相应特定受体的配体结合位点。相似地，Houghten, R.A.等(1996)关于过烷基化的寡肽混合物的美国专利 No.5480971 揭示了线性 C1—C7—烷基过烷基化寡肽及这种肽的系列及文库，以及用这种寡肽系列及文库确定优先与相应受体分子结合的过烷基化的寡肽的序列的方法。因此，本发明携带表位的肽的非肽类似物也可通过这些方法常规生产。

融合蛋白质

本领域技术人员将懂得本发明的 KDI 多肽和上述其携带表位的片段可与免疫球蛋白(IgG)的恒定区部分组合，产生嵌合多肽。这些融合蛋白促进纯化，并呈现体内半衰期延长。例如已经揭示由人 CD4—多肽的头两个结构域和哺乳动物免疫球蛋白的重或轻链恒定

区的各个结构域组成的融合蛋白 (EPA394827: Traunecker 等; 自然 331: 84—86(1988))。由于 IgG 部分而具有二硫键二聚体结构的融合蛋白在结合及中和其它分子方面比单体 KDI 蛋白或蛋白质片段更有效 (Fountoulakis 等, 生物化学杂志 270: 3958—3964(1995))。

抗体

用于本发明的 KDI 蛋白质特异抗体可针对完整 KDI 蛋白或其抗原性多肽片段而产生, 其可与载体蛋白如动物 (如兔或鼠) 的白蛋白一起呈递, 或者如果其足够长 (至少大约 25 个氨基酸), 则不用载体。

本文所用术语 “抗体” (Ab) 或 “单克隆抗体” (Mab) 包括能与 KDI 蛋白特异结合的完整分子以及抗体片段 (例如 Fab 和 F(ab')₂ 片段)。Fab 和 F(ab')₂ 片段缺少完整抗体的 Fc 片段, 从循环中更快速清除, 并可以具有低于完整抗体的非特异组织结合性 (Wahl 等, J.Nucl.Med 24:316-325(1983))。因此这些片段是优选的。

本发明的抗体可通过各种方法制备。例如, 可将表达 KDI 蛋白质或其抗原性片段的细胞施用于动物, 以诱导含有多克隆抗体的血清产生。在一优选方法中, 预先制备 KDI 蛋白质并纯化以使其基本上没有天然污染物。然后将这种制备物导入动物以产生较高特异活性的多克隆抗血清。

在一更优选的方法中, 本发明的抗体是单克隆抗体 (或其 KDI 蛋白质结合片段)。这种单克隆抗体可用杂交瘤技术学制备 (Kohler 等, 自然 256: 495(1975); Kohler 等, 欧洲免疫学杂志 6: 511(1976); Kohler 等, 欧洲免疫学杂志 6: 292(1976); Hammerling 等, 单克隆抗体及 T 细胞杂交瘤, Elsevier, N.Y., (1981) PP 563-681)。一般地, 这种方法包括用 KDI 蛋白质抗原或优选地用表达 KDI 蛋白质的细胞接种动物 (优选鼠)。通过其结合抗 KDI 蛋白质抗体的能力可识别适当的细胞。这种细胞可在任何适当组织培养基中培养; 但优选在

补加 10% 胎牛血清（在大约 56℃ 灭活），并补加大约 10g/l 非必需氨基酸，大约 1000U/ml 青霉素，及大约 100 μ g/ml 四环素的 Earle 改良的 Eagle 培养基中培养细胞。提取这种鼠的脾细胞并与适当骨髓瘤细胞系融合。根据本发明可使用任何适当的骨髓瘤细胞系，然而优选应用购自美国典型培养物保藏中心 Rockville, Maryland 的亲代骨髓瘤细胞系（SP20）。在融合后，所得杂交瘤细胞选择性地保持在 HAT 培养基中，然后如 Wands 等所述（Gastroenterology 80:225-232(1981)）通过限制稀释克隆。然后分析通过这种选择获得的杂交瘤细胞以鉴别分泌能结合 KDI 蛋白质抗原的抗体的克隆。

或者，能结合 KDI 蛋白质抗原的其它抗体可通过用抗独特型抗体以二步方法生产。这种方法利用抗体本身也是抗原这一事实，且因而能获得结合第二种抗体的抗体。根据这种方法，KDI 蛋白质特异抗体用于接种动物，优选鼠。然后将这种动物的脾细胞用于生产杂交瘤细胞，并筛选此杂交瘤细胞以鉴别能产生抗体的克隆，其结合 KDI 蛋白质特异抗体的能力可由 KDI 蛋白质抗原阻断。这种抗体包括 KDI 蛋白质特异抗体的抗独特型抗体，并能用于接种动物以诱导其它 KDI 蛋白特异抗体形成。

应意识到本发明的 Fab 和 F(ab')₂ 及其它抗体片段可根据本文所示方法使用。这种片段典型地用酶如木瓜蛋白酶（以产生 Fab 片段）或胃蛋白酶（以产生 F(ab')₂ 片段），通过蛋白酶解而产生。或者，KDI 蛋白结合片段可通过应用重组 DNA 技术或合成化学而产生。

对体内应用抗 KDI 而言，可优选使用“人化的”嵌合单克隆抗体。这种抗体可用衍生自上述产生单克隆抗体的杂交瘤细胞的遗传构建体产生。本领域已知生产嵌合抗体的方法。参见 Morrison, 科学 229: 1202 (1985); Oi 等, 生物技术 4: 214(1986); Cabilly 等, 美国专利 No.4816567; Taniguchi 等, EP171496; Morrison 等,

EP173494; Neuberger 等, WO 8601533; Robinson 等, WO8702671; Boulianne 等, 自然 312: 643 (1984); Neuberger 等, 自然 314: 268(1985)。

免疫系统相关疾病

治疗

本领域熟练技术人员还意识到, 由于本发明的 KDI 蛋白是干扰素家族的一成员, 当 KDI 加入个体的细胞, 组织或机体时, 此蛋白质将对个体的靶细胞发挥生理作用。因而应意识到由于个体干扰素活性的标准或正常水平的降低导致的疾病, 尤其免疫系统疾病, 可通过施用 KDI 多肽治疗。因此, 本发明还提供了治疗需要提高干扰素水平的个体的方法, 包括给这种个体施用包含一定量的本发明分离的 KDI 多肽的药物组合物, 所述的量能有效提高这种个体干扰素活性。

介导 α -IFN 和 β -IFN 生物活性的人工类 IFN 受体复合物除与 KDI 结合外也与 Ω -IFN 结合。由此, KDI 能临床用于抗病毒治疗, 例如治疗 AIDS, 病毒性肝炎包括慢性乙肝, 丙肝, 乳头瘤病毒, 病毒性脑炎及预防鼻炎及呼吸道感染。

KDI 还用于治疗各种癌症 (例如多毛细胞白血病, 急性骨髓瘤, 骨肉瘤, 基细胞癌, 神经胶质瘤, 肾细胞癌, 多发骨髓瘤, 黑色素细胞瘤, 及 Hodgkin 病)。

KDI 被认为可刺激天然杀伤细胞活性。由此, KDI 可用于治疗寄生虫及细菌感染, 例如治疗 *Cryptosporidium parvum* 感染及耐受多种药物的肺结核。

KDI 还被认为可用作免疫治疗制剂, 尤其用作免疫抑制剂。例如, 据信 KDI 抑制用促细胞分裂原或同种异体细胞, 骨髓样祖细胞及其它骨髓细胞刺激的淋巴细胞增殖。由此, KDI 当在化疗前施用时用作防护剂, 并另外可用于治疗淋巴细胞, 骨髓样祖细胞及骨髓

干细胞的过高增殖，例如治疗慢性多毛细胞白血病。KDI 多肽也可用于防止移植物与宿主的排斥，或缩短自身免疫疾病的进程如关节炎，多发硬化，(2)或糖尿病 (3)。KDI 还用于治疗哺乳动物变态反应，例如通过抑制体液应答进行。

KDI 可用作佐剂或共佐剂增强或刺激预防性或治疗性接种情况下的免疫应答。

另外，本发明还提供了治疗患者感染的方法，包括给需要抗感染治疗的患者施用有效量的本发明多肽。在一优选的实施方案中，感染是病毒，细菌或寄生虫感染。在一尤为优选的实施方案中，感染是病毒感染。

另外，本发明还提供了治疗癌症患者的方法，包括给需要抗癌治疗的患者施用有效量的本发明多肽。

另外，本发明还提供了免疫治疗癌症患者的方法，包括给需要抗癌治疗的患者施用有效量的本发明多肽。

另外，本发明还提供了免疫治疗患者的方法，包括给需要免疫治疗的患者施用有效量的本发明多肽。

配方

KDI 多肽组合物将以公认良好医学实践，考虑到个体患者的临床条件（尤其单用 KDI 多肽治疗的副作用），KDI 多肽组合物的释放位点，施用方法，施用程序及其它已知临床因素而加以配制。文中 KDI 多肽的“有效量”通过这种考虑加以确定。

一般而言，非肠道施用的每剂 KDI 多肽的药物有效量在大约 $1 \mu\text{g}$ /每公斤体重/天 $\sim$ 10mg/每公斤体重/天范围内，当然如上所述这要进行治疗性判断。优选地，剂量至少为 0.01mg/每公斤体重/天，更优选对人而言在大约 0.01 $\sim$ 10mg 激素/每公斤体重/天之间。如果连续提供，KDI 多肽典型地以大约 $1 \mu\text{g}$ /每公斤体重/小时 $\sim$ 50 μg /每公斤体重/小时的速度施用，每天注射 1 $\sim$ 4 次或连续皮下灌注，例

如用微泵。也可使用静脉袋溶液。治疗时间需根据所需效果观测治疗后发生的反应而加以变化。

含有本发明 KDI 的药物组合物可以口服，直肠，非肠道，脑池内，阴道内，腹膜内，局部（以粉末，软膏，滴剂或皮贴），颊部施用或口服或鼻喷。“药物可接受载体”是指非毒性固体，半固体或液体充填剂，稀释剂，胶囊化物或任何类型的配方辅剂。文中术语“非肠道”是指施用模式包括静脉内，肌内，腹膜内，臀部内，皮下及关节内注射与灌注。

KDI 多肽也适于通过持续释放系统施用。适当的持续释放组合物例如包括有形形式的半通透聚合物基质例如薄膜或微胶囊。适当的释放基质包括聚交酯（美国专利 No.3773919,EP58481），L-谷氨酸和 γ -乙基-L-谷氨酸的共聚物（Sidman, U.等，生物聚合物 22: 547-556（1983）），聚(2-羟乙基甲基丙烯酸)（R.Langer 等，J.Biomed Mater Res 15:167-277(1981)，及 R.Langer Chem. Tech. 12:98-105(1982)），乙烯乙烯基醋酸酯（R.Langer 等，Id.）或聚-D-3-羟基丁酸（EP133988）。持续释放 KDI 多肽组合物也包括脂质截留的 KDI 多肽。含有 KDI 多肽的脂质体通过已知方法制备，参见 DE3218121；Epstein 等，Proc.Natl.Acad.Sci(USA)82:3688-3692(1985)；Hwang 等，Proc.Natl.Acad. Sci(USA)77:4030-4034(1980)；EP52322；EP36676；EP88046；EP143949；EP142641；日本专利公开 83-118008；美国专利 4485045 和 4544545；及 EP102324。通常地，脂质体是小的均匀片形的（大约 200-800A(埃)），其中脂质含量高于大约 30mol%胆固醇，选择调整的比例以使 KDI 多肽治疗最佳化。

在一实施方案中，对非肠道施用而言，KDI 多肽通常通过以所需纯度，以单位剂量可注射形式（溶液，悬浮液或乳状液）与药物可接受载体混合而配制，该载体在所用剂量及浓度对受体无毒性，且可与配方其它中配料相容。例如，配方优选不包括氧化剂及其它

已知对多肽有害的化合物。

通常地，通过将 KDI 多肽均匀及密切地与液体载体或细分散的固定载体或二者接触而制备配方。然后如果需要，将产物制成所需形状。优选地此载体是非肠道载体，更优选与受体血液等渗的溶液。这种载体例如包括水，生理盐水，Ringer's 液及葡萄糖液。本发明还使用非水相载体如固定的油及油酸乙酯，以及脂质体。

载体适当地含有微量添加剂如增强等渗性及化学稳定性的物质。这种物质在所用剂量及浓度对受体是无毒的，且包括缓冲剂如磷酸，柠檬酸，琥珀酸，乙酸，及其它有机盐及其盐；抗氧化剂如抗坏血酸（VC）；低分子量（少于 10 个残基）多肽，例如聚精氨酸或三肽；蛋白质如血清白蛋白，明胶或免疫球蛋白；疏水聚合物如聚乙烯吡咯烷酮；氨基酸如甘氨酸，谷氨酸，天冬氨酸及精氨酸；单糖，二糖及其它碳水化合物包括纤维素或其衍生物，葡萄糖，甘露糖，或糊精；螯合剂如 EDTA；糖醇如甘露糖醇或山梨糖醇；平衡离子如钠离子；和/或非离子表面活性剂如 Polysorbates，Poloxamers，或 PEG。

KDI 多肽典型地以大约 0.1mg/ml~100mg/ml，优选 1—10mg/ml 浓度，pH 大约为 3~8 条件配制在这种载体中。应知道使用前述的某些赋形剂，载体或稳定剂将导致 KDI 多肽盐产生。

用于治疗 KDI 多肽必须是无菌的。通过无菌过滤膜（例如 0.2 微米膜）过滤很容易达到无菌。治疗性 KDI 多肽组合物通常置于具有便携无菌设备的容器中，如静脉内溶液包或具有可通过皮下注射针穿刺塞子的小瓶。

KDI 多肽通常以单位剂量或多剂量含量贮存，例如封蜡安瓿或小瓶，以水溶液或冻干配方贮存以重配。冻干配方例如是在 10ml 小瓶中充填 5ml 过滤的无菌 1%（w/v）KDI 多肽水溶液，然后将所得混合物冻干。灌注液通过用无菌注射用水注入冻干的 KDI 多肽而制

备。

本发明还提供了药物包或试剂盒，其包括一或多种本发明药物组合物的配料充填的一或多个容器，这种容器中还附有政府机构对药物或生物制品的生产，使用或销售调节的告示，该告示反映了政府机构对人施用的这些药物或生物制品的生产，使用及销售的认可。另外，本发明的多肽可与其它治疗性化合物联合应用。

兴奋剂及拮抗剂一分析和分子

本发明还提供了筛选化合物的方法，以鉴别那些增强或破坏 KDI 对细胞的作用，如其与 KDI 结合分子如受体分子的相互作用的化合物。兴奋剂是一种提高 KDI 天然生物功能或与 KDI 功能相似的化合物，而拮抗剂降低或消除其功能。

本发明此实施方案另一方面提供了鉴别受体蛋白或其它与 KDI 多肽特异结合的配体结合蛋白的一种方法。例如，细胞区室，如细胞膜或其制备物可从表达结合 KDI 的分子中制备。将此制备物与标记的 KDI 保温。根据本领域已知常规方法分离并定性 KDI 和与受体或其它结合蛋白结合的 KDI 复合物。或者，KDI 多肽可与固体支持物结合，由此从细胞中溶解的结合分子可与柱结合，然后根据常规方法洗脱及定性。

在本发明关于兴奋剂与拮抗剂的分析中，细胞区室如细胞膜或其制备物可从表达结合 KDI 的分子的细胞中制备，该分子如通过 KDI 调节的信号或调节途径的分子。此制备物用标记的 KDI 在没有或有可能是 KDI 拮抗剂或兴奋剂的候选分子情况下进行培养。候选分子与结合分子的结合能力通过标记配体结合降低而反映。无缘由结合分子，即不诱导 KDI 对 KDI 结合分子的结合作用，是良好的拮抗剂。与 KDI 结合良好且激发与 KDI 相同或相近作用的分子是兴奋剂。

测定潜在的兴奋剂与拮抗剂的类 KDI 作用例如可通过在候选分子与细胞或适当的细胞制备物相互反应之后，确定第二信使系统活

性，并对比 KDI 或分子激发同 KDI 相同作用的能力而进行。用于此目的的第二信使系统包括但非限于 AMP 鸟苷酸环化酶，离子通道或磷酸肌苷水解第二信使系统。

另一种对 KDI 拮抗剂的分析例如是竞争分析，其是在适当条件下组合 KDI 和潜在拮抗剂及膜结合的 KDI 受体分子或重组 KDI 受体分子，以进行竞争抑制分析。KDI 可被标记如经放射标记，这样可准确确定与受体分子结合的 KDI 分子数目，以评价潜在拮抗剂的效力。

潜在拮抗剂包括与本发明多肽结合的小的有机分子，肽，多肽及抗体，从而抑制或消除其活性。潜在拮抗剂还可以是与结合分子上相同位点结合的小的有机分子，肽，多肽如密切相关的蛋白质或抗体，而不诱导 KDI 诱导的活性，从而通过使 KDI 不能结合阻止 KDI 功能。

其它潜在拮抗剂包括反义分子。反义技术通过反义 DNA 或 RNA 或通过三螺旋形成用于控制基因表达。反义技术例如由 Okano, J.Neurochem. 56:560(1990) “寡脱氧核苷酸作为基因表达的反义抑制物” CRC 出版社, Boca Raton, FL(1988)所揭示。三螺旋形成例如见于 Lee 等, 核酸研究 6: 3073(1979); Coony 等, 科学 241: 456(1988); 及 Dervan 等, 科学 251: 1360(1991)所揭示。这些方法基于多核苷酸与互补 DNA 或 RNA 的结合。例如，编码本发明成熟多肽的多核苷酸的 5' 编码区可用于设计长度大约为 10~40 个碱基对的反义 RNA 寡核苷酸。设计一与参与转录的基因区互补的 DNA 寡核苷酸，从而阻止转录和 KDI 产生。此反义 RNA 寡核苷酸与 mRNA 体内杂交并破坏 mRNA 分子翻译为 KDI 多肽。上述寡核苷酸也可输入细胞，这样反义 RNA 或 DNA 可在体内表达从而抑制 KDI 蛋白产生。

兴奋剂和拮抗剂可以与如上所述药物可接受载体组合应用。

拮抗剂例如可用于抑制干扰素活性，例如在化疗后刺激骨髓和

造血原细胞增殖，任何上述拮抗剂均可与如后文所述药物可接受载体组合应用。

本发明参考以下例证非限制性的实施例将更易于理解。

实施例 1：在 E.coli 克隆和表达 KDI

新的 pHE4 系列细菌表达载体，尤其 pHE4a 载体在本实施例中用于细菌表达。（QIAGEN, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311）pHE4-5/KDI 载体质粒 DNA 含有插入在单一限制酶位点 NdeI 和 Asp 718 之间的编码图 1 所示多核苷酸的 DNA。此构建体为本领域技术人员的方便于 1998 年 2 月 25 日保藏于 ATCC，保藏号 No.209645。

pHE4a 细菌表达载体包括供选择的新霉素磷酸转移酶基因，E.coli 复制起点，T5 噬菌体启动子序列，2 个 lac 操纵基因序列，Shine-Delgarno 序列，及乳糖操纵子阻抑物基因（lacIq）。这些元件的排列使得编码多肽的插入 DNA 片段表达具有与多肽氨基端共价键连的 6 个 His 残基（即“6xHis 标记”）的多肽。

编码成熟 KDI 蛋白的 DNA 序列用 PCR 寡核苷酸引物扩增，引物与 KDI 蛋白所需部分的氨基端序列及保藏构建体中 cDNA 编码序列 3' 的序列退火。含有促进在 pHE4a 载体中克隆的限制位点的额外核苷酸分别加至 5' 和 3' 引物序列中。

为克隆 KDI 蛋白编码区，5' 引物序列为 5' GGCCGCATATGC TGGACTGTAACTTACTG3'（SEQ ID NO: 6），下划线处为 NdeI 限制位点，本领域技术人员将意识到蛋白质编码序列中 5' 引物起始点可以变化，以扩增编码完整 KDI 蛋白任何所需部分的 DNA 区段。3' 引物序列为 5' GGCCGCGGTACCTTATTCCTCCTGAAT AGAGC3'（SEQ ID NO: 17），划线处为 Asp 718 限制位点。

扩增的 KDI DNA 片段用 NdeI 和 Asp 718 酶切，并然后将与线性化质粒连接在一起。KDI DNA 插入限制的 pHE4a 载体中使 KDI

蛋白编码区在 IPTG 可诱导的启动子的下游及在起始 AUG 和 6 个组氨酸密码子的框架内。

用如 Sambrook 等所述标准方法（分子克隆实验手册，2nd Ed; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)），将连接混合物转化入感受态 *E.coli* 细胞中。本发明实施例中使用的是含有多拷贝的质粒 pREP4 的 *E.coli* 菌株 M15/rep4，其表达 lac 阻抑物并赋予卡那霉素抗性（“Kan^r”）。这个唯一适于表达 KDI 蛋白的菌株可以商购（QIAGEN, Inc, 前述）。通过其在有氨苄青霉素和卡那霉素的 LB 平板上的生长能力鉴别转化体。从抗性克隆中分离质粒 DNA，并通过限制性分析，PCR 及 DNA 测序确定克隆的 DNA 的正确性。

含有所需构建体的克隆在补加氨苄青霉素（100 μ g/ml）和卡那霉素（25 μ g/ml）的液体 LB 培养基中生长过夜（O/N）。此 O/N 培养物以大约 1: 25~1: 250 的稀释度用于接种大培养基。细胞生长至 600nm 光密度（OD 600）为 0.4 和 0.6 之间。然后加入异丙基硫代- β -D-半乳糖苷（IPTG）至终浓度为 1mM，通过灭活 lacI 阻抑物，诱导从 lac 阻抑物敏感启动子转录。接着将细胞另外培养 3~4 小时。然后经离心收集细胞。

细胞然后在 4°C，在 6M 盐酸胍 pH8 中搅拌 3~4 小时。经离心除去细胞碎片，并将含有 KDI 多肽的上清加样于镍-胍-三乙酸（“Ni-NTA”）亲和树脂柱中（QIAGEN, Inc.）。具有 6xHis 标记的蛋白质与 Ni-NTA 树脂高亲和性结合，并可用简便的一步法纯化（详见 The QIAexpressionist, 1995, QIAGEN, Inc）。简而言之，上清在 6M 盐酸胍，pH8 中加样于柱上，该柱首先用 10 体积的 6M 盐酸胍，pH8 洗涤，然后用 10 体积的 6M 盐酸胍 pH6 洗涤，最后用 6M 盐酸胍 pH5 洗脱 KDI。

然后通过用磷酸盐缓冲盐水（PBS）或 50mM 乙酸钠缓冲液 pH6，

加上 200mM NaCl 将纯化的蛋白质透析复性。或者，此蛋白质当固定在 Ni-NTA 柱上时可被充分重折叠。建议条件如下：在含有蛋白酶抑制剂的 500mM NaCl, 20% 甘油, 20mM Tris/HCl pH7.4 中，以线性 6M~1M 尿素梯度复性。复性应进行 1.5 小时或更长。复性后，蛋白质可通过加入 250mM 咪唑洗脱。咪唑通过最后用 PBS 或 50mM 乙酸钠缓冲液 pH6 加上 200mM NaCl 透析而除去，纯化的蛋白质在 4℃ 贮存或在 -80℃ 冷冻。

当其以包涵体形式存在时，以下另一种方法可用于纯化在 E.coli 中表达的 KDI，除非特别指出，以下步骤均在 4~10℃ 进行。

在 E.coli 发酵生产阶段结束时，细胞培养物冷却至 4~10℃，并通过以 15000rpm 持续离心收集细胞（Heraeus Sepatech）。基于每单位重细胞糊的蛋白质的期望产量及所需纯化的蛋白质量，将适量重的细胞糊悬浮于含有 100mM Tris, 50mM EDTA, pH 7.4 的缓冲液中，用高剪切混合仪将细胞分散为均匀悬浮液。

然后将此溶液在 4000~6000 psi 两次流经显微流化仪（Microfluidics, Corp 或 APV Gaulin. Inc.）将细胞裂解。匀浆然后与 NaCl 液混合至终浓度为 0.5M NaCl，随后以 7000xg 离心 15 分钟。所得沉淀再用 0.5M NaCl, 100mM Tris, 50mM EDTA, pH 7.4 洗涤。

所得洗涤的包涵体用 1.5M 盐酸胍（GnHCl）溶解 2~4 小时，在 7000xg 离心 15 分钟后，弃去沉淀并将含 KDI 多肽的上清在 4℃ 培养过夜，以进行下一步的 GuHCl 提取。

在高速离心（30000xg）除去不溶颗粒后，通过快速混合 GnHCl 提取物与 20 体积的含 50mM 乙酸钠, pH4.5, 150mM NaCl, 2mM EDTA 的缓冲液，经强力搅拌，将 GnHCl 溶解的蛋白质重折叠。重折叠的稀释蛋白质溶液在进一步纯化之前，在 4℃ 不混合保持 12 小时。

为澄清重折叠的 KDI 多肽溶液，使用具有适当表面积（如

Filtron) 的 $0.16\ \mu\text{m}$ 膜滤器, 用 40mM 乙酸钠, pH6.0 平衡的提前制备的切向过滤装置。过滤的样品加样于阳离子交换树脂 (如 Poros HS-50, Perseptive Biosystems) 上。该柱用 40mM 乙酸钠, pH6.0 洗涤, 并用含有 250mM, 500mM, 1000mM 和 1500mM NaCl 的相同缓冲液分步洗脱。连续监测流出物在 280nm 的吸光度。收集各级组分并进行进一步 SDS-PAGE 分析。

将含有 KDI 多肽的组分集中并与 4 体积的水混合。稀释的样品然后加样于提前制备的一系列强阴离子 (Poros HQ-50, Perseptive Biosystems) 和弱阴离子 (Poros CM-20, Perseptive Bio systems) 交换树脂串联柱上。该柱用 40mM 乙酸钠 pH6.0 平衡。两个柱均用 40mM 乙酸钠, pH6.0, 200mM NaCl 洗涤。CM-20 柱然后在 0.2M NaCl, 50mM 乙酸钠, pH 6.0 至 1.0M NaCl, 50mM 乙酸钠, pH 6.5 范围, 用 10 柱体积线性梯度洗脱。在 A280 之下恒定监测流出物收集各级组分。然后集中含有 KDI 多肽的组分 (例如通过 16% SDS-PAGE 确定)。

所得 KDI 多肽在上述重折叠及纯化步骤后呈 95% 以上的纯度。当加样 $5\ \mu\text{g}$ 纯化的蛋白质时, 在 Commassie 蓝染色的 16% SDS-PAGE 凝胶中未观测到主要污染带。纯化的蛋白质也进行内毒素/LPS 污染测试, 典型地根据 LAL 分析 LPS 污染物少于 0.1ng/ml。

本发明人已产生了多个 KDI 表达构建体, 以促进各种大小及各种系统中 KDI 多肽的生产。基于 E.coli 的构建体如下: (1) pQE9: KDI.S27-K207 (表达 SEQ ID NO: 2 的 27-207 位氨基酸); (2) pHE4: KDI.S27-K207 (表达 SEQ ID NO: 2 的 27-207 位氨基酸); (3) pHE4: KDI.A23-K207 (表达 SEQ ID NO: 2 的 23-207 位氨基酸); (4) pHE4.KDI.G24-K207 (表达 SEQ ID NO: 2 的 24-207 位氨基酸); (5) pHE4.KDI.C30-K207 (表达 SEQ ID NO: 2 的 30-207 位氨基酸)。

实施例 2：在杆状病毒表达系统中克隆和表达 KDI 蛋白质

在此实施例中，质粒穿梭载体 pA2GP 用于将编码 KDI 的克隆的 DNA 插入杆状病毒中，以表达 KDI 蛋白质，所用的杆状病毒前导序列及标准方法如 Summer 等，杆状病毒载体及昆虫细胞培养方法手册，Texas Agricultural Experimental Station Bulletin No.1555 (1987) 所述。此表达载体含有苜蓿银纹夜蛾多核型多角体病毒 (AcMNPV) 的强多角体蛋白启动子，后接杆状病毒 gp67 蛋白的分泌性信号肽 (前导序列) 及常规限制位点如 BamHI, XbaI 和 Asp718。猴病毒 40 (SV40) 的聚腺苷酸化位点用于有效地聚腺苷酸化。为便于选择重组的病毒，此质粒含有在弱果蝇启动子控制下相同方向的 E.coli β 一半乳糖苷酶基因，后接多角体蛋白基因的聚腺苷酸化信号。插入的基因位于与野生型病毒 DNA 经细胞介导的同源重组的病毒序列两侧翼，以产生表达克隆的多核苷酸的活病毒。

本领域技术人员充分意识到许多其它杆状病毒载体可代替以上载体而使用，如 pAc373, pVL941 和 pAc IM1，只要此构建提供适当的转录，翻译，分泌等的定位信号，包括所需信号肽及框内 AUG。这种载体例如由 Luckow 等，病毒学 170: 31—39 (1989) 所述。

编码保藏的克隆中的成熟 KDI 多肽的缺失 AUG 起始密码子及 SEQ ID NO: 2 所示天然相关的前导序列的 cDNA 序列，用相应于基因 5' 和 3' 序列的 PCR 寡核苷酸引物扩增，5' 引物序列为 5' GGCCGGGATCCGCCATCATGAGCACCAAACCTGATATG3' (SEQ ID NO:18), 划线处为 Bam HI 限制酶位点。3' 引物序列为 5' GGCCGCGGTACCTTATTTCTCCTGAATAGAGC3' (SEQ ID NO:19), 划线处为 Asp718 限制酶位点。

扩增的片段用商购的试剂盒 (“Geneclean” BIO 101 Inc., La Jolla, CA) 从 1% 琼脂糖凝胶中分离出。该片段然后用 BamHI 和 Asp718 酶切，并再在 1% 琼脂糖凝胶上纯化。

此质粒用限制酶 BamHI 和 Asp 718 酶切，并任选地，用本领域已知常规方法，用牛肠磷酸酶去磷酸化。该 DNA 然后用商购试剂盒（“Geneclean” BIO 101 Inc., La Jolla, CA）从 1% 琼脂糖凝胶中分离。

片段和去磷酸化的质粒用 T4DNA 连接酶连在一起。用连接混合物转化 E.coli HB 101 或其它适当 E.coli 宿主如 XL-1 Blue（Statagene Cloning Systems, La Jolla, CA）并涂布于培养平板上。用 BamHI 和 Asp 718 酶切各个菌落的 DNA，然后经凝胶电泳分析酶切产物，鉴定细菌含有具有人 KDI 基因的质粒。克隆的片段的序列通过 DNA 测序确定。此质粒在此称为 pA2GPKDI。

用 Felgner 等, Proc Natl Acad Sci USA 84:7413-7417(1987)所述脂转染法，将 5 μ g 质粒 pA2GPKDI 与 1.0 μ g 商购线性化的杆状病毒 DNA（“Baculo Gold™ baculovirus DNA”, Pharmingen, San Diego, CA）共转染。将 1 μ g BaculoGold™ 病毒 DNA 和 5 μ g 质粒 pA2GPKDI 混合在含 50 μ l 无血清 Grace' s 培养基的微滴定平板的无菌孔中（Life Technologies Inc., Gaithersburg, MD）随后加入 10 μ l Lipofectin 和 90 μ l Grace' s 培养基，混合，并在室温培养 15 分钟。转染混合物然后滴加入 Sf9 昆虫细胞中（ATCC CRL 1711），该细胞种植于具有 1ml 无血清 Grace' s 培养基的 35mm 组织培养平板中。此平板在 27°C 培养 5 小时，然后从平板中除去转染溶液，并加入 1ml 补加 10% 胎牛血清的 Grace' s 昆虫培养基，在 27°C 持续培养 4 天。

四天后，收集上清，并如 Summer 和 Smith（前述）所述进行噬斑分析。具有“Blue Gal”（Life Technologies Inc., Gaithersburg）的琼脂糖凝胶使产生兰染噬斑的 gal 表达克隆易于鉴别及分离。（关于此类“噬斑分析”可见于昆虫细胞培养使用手册及由 Life Technologies Inc., Gaithersburg, P 9-10 所分类的杆状病毒学）。在适当培养后，用微量移液管的尖端（如 Eppendorf）挑取兰染的噬斑。含有重组病毒的琼脂然后重悬浮于含 200 μ l Grace' s 培养基的微离

心管中，并将含有重组杆状病毒的悬浮液施加于种植于 35mm 平皿中的昆虫 Sf9 细胞中。4 天后收集这些培养平皿的上清，然后在 4℃ 贮存。此重组病毒称为 V-KDI。

为检定 KDI 基因的表达，Sf9 细胞在补加 10% 热灭活的 FBS 的 Grace' s 培养基中生长。以感染复数 (MOI) 大约为 2 用重组杆状病毒 V-KDI 感染细胞，如果放射标记的蛋白质是所需的，6 小时后除去培养基，并用减去甲硫氨酸和半胱氨酸的 SF900I II 培养基（可购自 Life Technologies Inc., Rockville, MD）置换。42 小时后，加入 5 μ Ci 的 ^{35}S -甲硫氨酸和 5 μ Ci ^{35}S -半胱氨酸（可购自 Amersham）。将细胞进一步培养 16 小时，然后经离心收集，上清中蛋白质以及细胞内蛋白质通过 SDS-PAGE，随后经自动放射自显影（如果是放射标记的）分析。

纯化的蛋白质氨基端的氨基酸序列的微测序可用于确定 KDI 蛋白质氨基端序列。

其它杆状病毒表达构建体如下构建：（1）pA2:KDI（表达 SEQ ID NO: 2 的 1-207 位残基）；（2）pA2.KDI.M7-K207（表达 SEQ ID NO: 2 的 7-207 位残基）；（3）pA2gp.KDI.L28-K207（表达 SEQ ID NO: 2 的 28-207 位残基）；（4）pA2gp.KDI.C30-K207（表达 SEQ ID NO: 2 的 30-207 位残基）；（5）pA2.KDI.M1-R192（表达 SEQ ID NO: 2 的 1-192 位残基）。

实施例 3：在哺乳动物细胞中克隆和表达 KDI

典型的哺乳动物表达载体包括启动子元件，其引导 mRNA 转录开始，蛋白质编码序列，为转录终止和转录物聚腺苷酸化所需的信号。其它元件包括增强子，Kozak 序列及侧翼于供体和受体 RNA 剪切位点的间插序列。用 SV 40 的早期和晚期启动子，反转录病毒如 RSV, HTLV, HIV 的长末端重复 (LTRs) 和巨细胞病毒 (CMV) 的早期启动子可实现高效转录。然而，也可使用细胞元件（例如人

肌动蛋白启动子)。适用于本发明的表达载体例如包括 pSVL 和 pMSG 载体 (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) 和 pBC12MI (ATCC 67109)。可用的哺乳动物宿主细胞包括人 HeLa, 293, H9 和 Jurkat 细胞, 鼠 NIH3T3 和 C127 细胞, Cos1, Cos7 和 CV1, 鹌鹑 QC1-3 细胞, 鼠 L 细胞和中国仓鼠卵巢 (CHO) 细胞。

或者, 基因可在含有整合入染色体中的基因的的稳定细胞系中表达。用选择标记如 dhfr, gpt, 新霉素, 潮霉素, 可鉴别共转染及分别转染的细胞。

转染的基因也可扩增以表达大量编码的蛋白质, DHFR (二氢叶酸还原酶) 标记用于开发携带相应基因的几百个或者甚至几千个拷贝的细胞系。其它有效选择标记是谷氨酰胺合酶 (GS) (Murphy 等, 生物化学杂志 277: 227-279 (1991); Bebbington 等, 生物/技术学 10: 169-175 (1992))。使用这些标记, 哺乳动物细胞在选择性培养基中生长, 并选择最高抗性的细胞, 这些细胞系含有整合入染色体中的扩增的基因, 中国仓鼠卵巢 (CHO) 细胞和 NSO 细胞通常用于蛋白质的生产。

表达载体 pC1 和 pC4 含有 Rous 肉瘤病毒 (Cullen 等, 分子和细胞生物学, 438-447 (1985, 3)) 的强启动子 (LTR), 加上 CMV 一增强子的片段 (Boshart 等, 细胞 41: 521-530 (1985))。例如具有 BamHI, XbaI 和 Asp 718 限制酶切位点的多克隆位点, 促进相应基因的克隆。除 3' 内含子之外, 载体还含有鼠前胰岛素原基因的聚腺苷酸化及终止信号。

在 CHO 细胞中克隆及表达

载体 pC4-Sig 用于表达 KDI 多肽, 质粒 pC4-Sig 是质粒 pSV2-dhfr (ATCC 登录号 No.37146) 的衍生物。其含有位于多克隆位点上游的趋化因子 β -8 (见 US95/09508) 的分泌前导序列的编

码区，且被设计在插入的异源 DNA 框内。此质粒含有在 SV40 早期启动子控制下的鼠 DHFR 基因。用这些质粒转染的，缺失二氢叶酸活性的 CHO 或其它细胞，可通过将细胞在补加化疗剂氨甲喋呤的选择性培养基（ α 减 MEM, Life Technologies）生长加以选择。抗氨甲喋呤（MTX）细胞中的 DHFR 基因的扩增已被充分论证（参见如 Alt, F.W., Kellems, R.M., Bertino, J.R.和 Schimke, R.T.1978,生物化学杂志 253: 1357—1370, Hamlin, J.L.和 Ma, C 1990, Biochem,et Biophys Acta,1097:107-143,Page,M.J.和 Sydenham,M.A.1991,生物技术学 9: 64—68）。细胞在 MTX 浓度增高下生长，由于 DHFR 基因扩增，导致定向酶 DHFR 的过量产生，引发对药物的抗性。如果第二个基因连于 DHFR 基因，其通常同时被扩增及过表达。本领域已知此方法可用于开发携带 1000 以上拷贝的扩增基因的细胞系。随后，当除去氨甲喋呤时，可获得含有整合入宿主细胞一或多条染色体中的扩增的基因的细胞系。

质粒 pC4 含有表达相应基因的 Rouse 肉瘤病毒（Cullen 等，分子和细胞生物学，1985.3.:438-447）的长末端重复序列（LTR）的强启动子，加上分离自人巨细胞病毒（CMV）（Boshart 等，细胞 41: 521—530（1985））的立即早期基因增强子的片段。启动子的下游是下述使基因整合的单限制酶切位点：BamHI,XbaI 和 Asp 718。除这些克隆位点外，质粒还含有 3' 内含子和鼠前胰岛素原基因的聚腺苷酸化位点。其它高效启动子也可用于表达，例如人 β -肌动蛋白启动子，SV40 早期或晚期启动子，或其它逆转录病毒的长末端重复序列，例如 HIV 和 HTLVII。Clontech 的 Tet-off 和 Tet-on 基因表达系统及相似系统可用于在哺乳动物细胞中以可调节的方式表达 KDI 多肽。（Gossen.M.,& Bujard, H.1992, Proc. Natl Acad Sci USA 89:5547-5551）。为 mRNA 的聚腺苷酸化，其它例如人生长素或珠蛋白基因的信号也可使用。携带整合入染色体的相应基因的稳定细

胞系也可在共转染后用选择标记如 gpt, G418 或潮霉素选择。在开始时使用一个以上选择标记例如 G 418 加上氨甲喋呤是有利的。

质粒 pC4 用限制酶 BamHI 和 Asp 718 酶切, 然后用牛肠磷酸酶经本领域已知方法去磷酸化。载体然后从 1% 琼脂糖凝胶中分离出。

编码 KDI 多肽的 DNA 序列用相当于基因所需部分的 5' 和 3' 序列的 PCR 寡核苷酸引物扩增。5' 引物含有下划线处的 BamHI 位点, Kozak 序列及 AUG 起始密码子, 5' 引物序列为 5' GGCCGGGATCCGCCATCATGAGCACCAAACCTGATATG3' (SEQ ID NO:18)。3' 引物含有下划线处的 Asp 718 限制位点, 其序列如下: 5' GGCCGCGGTACCTTATTTCCTCCCTGAATAGAGC 3' (SEQ ID NO:19)。

扩增的片段用内切核酸酶 BamHI 和 Asp718 酶切, 然后再在 1% 琼脂糖凝胶上纯化。分离的片段与去磷酸化的载体然后用 T4 DNA 连接酶连接。然后转化 E.coli HB101 或 XL-1 Blue 细胞, 并用例如限制酶分析鉴别含有插入质粒 pC4 中的片段的细菌。

缺失活性 DHFR 基因的 CHO 细胞用于转染。将 5 μ g 的表达质粒 pC4 与 0.5 μ g 的质粒 pSVneo 用脂转染法 (Felgner 等, 前述) 共转染。质粒 pSV2-neo 含有显性选择标记, Tn5 的 neo 基因, 其编码赋予包括 G418 的一组抗生素抗性的酶。将细胞种植于补加 1mg/ml G418 的 α 减 MEM 培养基中。2 天后, 细胞用胰蛋白酶消化, 并种植于补加 10, 25 或 50ng/ml 氨甲喋呤加上 1mg/ml G418 的 α 减 MEM 培养基中的杂交瘤克隆平板 (Greiner, 德国) 中。10-14 天后, 将单个克隆用胰蛋白酶消化, 并种植于具有不同氨甲喋呤浓度 (50nM, 100nM, 200nM, 400nM, 800nM) 的 6 孔培养皿或 10ml 烧瓶中。然后将在最高氨甲喋呤浓度生长的克隆移至含有更高浓度氨甲喋呤 (1 μ M, 2 μ M, 5 μ M, 10 μ M, 20 μ M) 的新 6 孔平板上。重复相同步骤直至获得在 100~200 μ M 浓度生长的克隆。所需基因产物的表

达通过例如 SDS-PAGE 及 Western 印迹或反相 HPLC 分析进行分析。

其它哺乳动物表达载体构建如下：(1) pC4: KDI (表达 SEQ ID NO: 2 的 1-207 残基); (2) pC4sp: KDI.C30-K207 (表达后接 SEQ ID NO: 2 的 30-207 位残基的异源信号肽 (趋化因子 β -8 (MPIF-1) 信号肽)); 及(3)pC4sp: KDI.L28-K207 (表达后接 SEQ ID NO: 2 的 28~207 位残基的异源信号肽 (趋化因子 β -8 (MPIF-1) 信号肽))。

应清楚本发明除前述阐述及实施例外还可以其它方式实施。根据以上教导及在权利要求范围内可对本发明做各种修改及变化。

所有引证的出版物 (包括专利, 专利申请, 杂志, 实验手册, 书或其它文件) 均并入参考。

关于微生物保藏的说明
(PCT 细则 13 之二)

A. 对说明书第 <u>8</u> 页, 第 <u>1-3</u> 行所述的微生物的说明。	
B. 保藏事项 其它保藏在补充页中 ☐	
保藏单位名称 美国典型培养物保藏中心	
保藏单位地址 (包括邮政编码和国名) 美国弗吉尼亚州马纳萨斯大学大道 10801 号 (2011-2209)	
保藏日期 1998 年 12 月 1 日	保藏编号 203500
C. 补充说明 (必要时) 本栏有补充页 ☐	
D. 本说明是为下列指定国作的 (如果说明不是为所有指定国而作的)	
E. 补充说明 (必要时)	
下列说明将随后向国际局提供 (写出说明的类别, 例如: “保藏的编号”)	

由受理局填写 ☒ 本页已经和国际申请一起收到 受权官员	由国际局填写 ☐ 国际局收到本页日期 受权官员
--	---

序列表

<110> 人体基因组科学有限公司等

<120> 角质形成细胞衍生的干扰素

<130> PF482

<140> 未给出

<141> 1999-07-21

<150> 60/093,643

<151> 1998-07-21

<160> 21

<170> PatentIn Ver. 2.0

<210> 1

<211> 1170

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (35)..(655)

<400> 1

```

ccacgcgtcc gggatttttt agcttgcaaa aaaa atg agc acc aaa cct gat atg 55
                                   Met Ser Thr Lys Pro Asp Met
                                   1               5

att caa aag tgt ttg tgg ctt gag atc ctt atg ggt ata ttc att gct 103
Ile Gln Lys Cys Leu Trp Leu Glu Ile Leu Met Gly Ile Phe Ile Ala
      10              15              20

ggc acc cta tcc ctg gac tgt aac tta ctg aac gtt cac ctg aga aga 151
Gly Thr Leu Ser Leu Asp Cys Asn Leu Leu Asn Val His Leu Arg Arg
      25              30              35

gtc acc tgg caa aat ctg aga cat ctg agt agt atg agc aat tca ttt 199
Val Thr Trp Gln Asn Leu Arg His Leu Ser Ser Met Ser Asn Ser Phe
      40              45              50              55

cct gta gaa tgt cta cga gaa aac ata gct ttt gag ttg ccc caa gag 247
Pro Val Glu Cys Leu Arg Glu Asn Ile Ala Phe Glu Leu Pro Gln Glu
      60              65              70

ttt ctg caa tac acc caa cct atg aag agg gac atc aag aag gcc ttc 295
Phe Leu Gln Tyr Thr Gln Pro Met Lys Arg Asp Ile Lys Lys Ala Phe
      75              80              85

tat gaa atg tcc cta cag gcc ttc aac atc ttc agc caa cac acc ttc 343
Tyr Glu Met Ser Leu Gln Ala Phe Asn Ile Phe Ser Gln His Thr Phe
      90              95              100

aaa tat tgg aaa gag aga cac ctc aaa caa atc caa ata gga ctt gat 391
Lys Tyr Trp Lys Glu Arg His Leu Lys Gln Ile Gln Ile Gly Leu Asp
      105              110              115

```

cag caa gca gag tac ctg aac caa tgc ttg gag gaa gac gag aat gaa 439
 Gln Gln Ala Glu Tyr Leu Asn Gln Cys Leu Glu Glu Asp Glu Asn Glu
 120 125 130 135

 aat gaa gac atg aaa gaa atg aaa gag aat gag atg aaa ccc tca gaa 487
 Asn Glu Asp Met Lys Glu Met Lys Glu Asn Glu Met Lys Pro Ser Glu
 140 145 150

 gcc agg gtc ccc cag ctg agc agc ctg gaa ctg agg aga tat ttc cac 535
 Ala Arg Val Pro Gln Leu Ser Ser Leu Glu Leu Arg Arg Tyr Phe His
 155 160 165

 agg ata gac aat ttc ctg aaa gaa aag aaa tac agt gac tgt gcc tgg 583
 Arg Ile Asp Asn Phe Leu Lys Glu Lys Lys Tyr Ser Asp Cys Ala Trp
 170 175 180

 gag att gtc cga gtg gaa atc aga aga tgt ttg tat tac ttt tac aaa 631
 Glu Ile Val Arg Val Glu Ile Arg Arg Cys Leu Tyr Tyr Phe Tyr Lys
 185 190 195

 ttt aca gct cta ttc agg agg aaa taagaatcat ctaccttcaa gcaagaatta 685
 Phe Thr Ala Leu Phe Arg Arg Lys
 200 205

 acagagattg tggctacgca aatgcaccaa aaaaggggtga aatatatctg aaatgtacct 745
 ggttctgccc ttggaagcca cttctgtctc atgccactaa cagcatgctg ccaaactgtt 805
 cagattcaag attattccaa gcgcagggcc caaatgttat agccaaagaa agtcttatga 865
 taaaagtgag gcaaatttca gccagaagt tagaagagat gtttaaaaga acaagaacaa 925
 attgtggatc atggtatatg caggctatca gcagaaggat cagacaataa aatgagttag 985
 tgcaaaccat ttagtaaaaa taactatcag cagagttgtt ccagattaaa aatagtacta 1045
 caagcttgta aaggagttag gacatgcaag ctactgagca taaaatatat acttgctatt 1105
 ttcatgact ttctctaata aagtctttga ctgttctctc taataaaaaa aaaaaaaaaa 1165
 aaaaa 1170

<210> 2
 <211> 207
 <212> PRT
 <213> Homo sapiens

<400> 2
 Met Ser Thr Lys Pro Asp Met Ile Gln Lys Cys Leu Trp Leu Glu Ile
 1 5 10 15
 Leu Met Gly Ile Phe Ile Ala Gly Thr Leu Ser Leu Asp Cys Asn Leu
 20 25 30
 Leu Asn Val His Leu Arg Arg Val Thr Trp Gln Asn Leu Arg His Leu
 35 40 45
 Ser Ser Met Ser Asn Ser Phe Pro Val Glu Cys Leu Arg Glu Asn Ile
 50 55 60

Ala Phe Glu Leu Pro Gln Glu Phe Leu Gln Tyr Thr Gln Pro Met Lys
65 70 75 80

Arg Asp Ile Lys Lys Ala Phe Tyr Glu Met Ser Leu Gln Ala Phe Asn
85 90 95

Ile Phe Ser Gln His Thr Phe Lys Tyr Trp Lys Glu Arg His Leu Lys
100 105 110

Gln Ile Gln Ile Gly Leu Asp Gln Gln Ala Glu Tyr Leu Asn Gln Cys
115 120 125

Leu Glu Glu Asp Glu Asn Glu Asn Glu Asp Met Lys Glu Met Lys Glu
130 135 140

Asn Glu Met Lys Pro Ser Glu Ala Arg Val Pro Gln Leu Ser Ser Leu
145 150 155 160

Glu Leu Arg Arg Tyr Phe His Arg Ile Asp Asn Phe Leu Lys Glu Lys
165 170 175

Lys Tyr Ser Asp Cys Ala Trp Glu Ile Val Arg Val Glu Ile Arg Arg
180 185 190

Cys Leu Tyr Tyr Phe Tyr Lys Phe Thr Ala Leu Phe Arg Arg Lys
195 200 205

<210> 3

<211> 238

<212> PRT

<213> Homo sapiens

<400> 3

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

Ser Pro Val Gly Ser Leu Gly Cys Asp Leu Pro Gln Asn His Gly Leu
20 25 30

Leu Ser Arg Asn Thr Leu Val Leu Leu His Gln Met Arg Arg Ile Ser
35 40 45

Pro Phe Leu Cys Leu Lys Asp Arg Arg Asp Phe Arg Phe Pro Gln Glu
50 55 60

Met Val Lys Gly Ser Gln Leu Gln Lys Ala His Val Met Ser Val Leu
65 70 75 80

His Glu Met Leu Gln Gln Ile Phe Ser Leu Phe His Thr Glu Arg Ser
85 90 95

Ser Ala Ala Trp Asn Met Thr Leu Leu Asp Gln Leu His Thr Glu Leu
100 105 110

His Gln Gln Leu Gln His Leu Glu Thr Cys Leu Leu Gln Val Val Gly
115 120 125

Glu Gly Glu Ser Ala Gly Ala Ile Ser Ser Val Pro Gln Leu Ser Ser

130 135 140
 Leu Glu Leu Arg Arg Tyr Phe His Arg Ile Asp Asn Phe Leu Lys Glu
 145 150 155 160
 Lys Lys Tyr Ser Asp Cys Ala Trp Glu Ile Val Arg Val Glu Ile Arg
 165 170 175
 Arg Cys Leu Tyr Tyr Phe Tyr Lys Phe Thr Ala Leu Pro Ala Leu Thr
 180 185 190
 Leu Arg Arg Tyr Phe Gln Gly Ile Arg Val Tyr Leu Lys Glu Lys Lys
 195 200 205
 Tyr Ser Asp Cys Ala Trp Glu Val Val Arg Met Glu Ile Met Lys Ser
 210 215 220
 Leu Phe Leu Ser Thr Asn Met Gln Glu Arg Leu Arg Ser Lys
 225 230 235

 <210> 4
 <211> 187
 <212> PRT
 <213> Homo sapiens

 <400> 4
 Met Thr Asn Lys Cys Leu Leu Gln Ile Ala Leu Leu Leu Cys Phe Ser
 1 5 10 15
 Thr Thr Ala Leu Ser Met Ser Tyr Asn Leu Leu Gly Phe Leu Gln Arg
 20 25 30
 Ser Ser Asn Phe Gln Cys Gln Lys Leu Leu Trp Gln Leu Asn Gly Arg
 35 40 45
 Leu Glu Tyr Cys Leu Lys Asp Arg Met Asn Phe Asp Ile Pro Glu Glu
 50 55 60
 Ile Lys Gln Leu Gln Gln Phe Gln Lys Glu Asp Ala Ala Leu Thr Ile
 65 70 75 80
 Tyr Glu Met Leu Gln Asn Ile Phe Ala Ile Phe Arg Gln Asp Ser Ser
 85 90 95
 Ser Thr Gly Trp Asn Glu Thr Ile Val Glu Asn Leu Leu Ala Asn Val
 100 105 110
 Tyr His Gln Ile Asn His Leu Lys Thr Val Leu Glu Glu Lys Leu Glu
 115 120 125
 Lys Glu Asp Phe Thr Arg Gly Lys Leu Met Ser Ser Leu His Leu Lys
 130 135 140
 Arg Tyr Tyr Gly Arg Ile Leu His Tyr Leu Lys Ala Lys Glu Tyr Ser
 145 150 155 160
 His Cys Ala Trp Thr Ile Val Arg Val Glu Ile Leu Arg Asn Phe Tyr
 165 170 175

Phe Ile Asn Arg Leu Thr Gly Tyr Leu Arg Asn
 180 185

<210> 5
 <211> 194
 <212> PRT
 <213> Homo sapiens

<400> 5
 Met Ala Phe Val Leu Ser Leu Leu Met Ala Leu Val Leu Val Ser Tyr
 1 5 10 15

Gly Pro Phe Gly Ser Leu Gly Cys Asp Leu Ser Gln Asn His Val Leu
 20 25 30

Val Gly Arg Lys Asn Leu Arg Leu Leu Asp Glu Met Arg Arg Leu Ser
 35 40 45

Pro His Phe Cys Leu Gln Asp Arg Lys Asp Phe Ala Leu Pro Gln Glu
 50 55 60

Met Val Glu Gly Gly Gln Leu Gln Glu Ala Gln Ala Ile Ser Val Leu
 65 70 75 80

His Glu Met Leu Gln Gln Ser Phe Asn Leu Phe His Thr Glu His Ser
 85 90 95

Ser Ala Ala Trp Asp Thr Thr Leu Leu Glu Pro Cys Arg Thr Gly Leu
 100 105 110

His Gln Gln Leu Asp Asn Leu Asp Ala Cys Leu Gly Gln Val Met Gly
 115 120 125

Glu Glu Asp Ser Ala Leu Gly Arg Thr Gly Pro Leu Ala Leu Lys Arg
 130 135 140

Tyr Phe Gln Gly Ile His Val Tyr Leu Lys Glu Lys Gly Tyr Ser Asp
 145 150 155 160

Cys Ala Trp Glu Thr Val Arg Leu Glu Ile Met Arg Ser Phe Ser Ser
 165 170 175

Leu Ile Ser Leu Gln Glu Arg Leu Arg Met Met Asp Gly Asp Leu Ser
 180 185 190

Ser Pro

<210> 6
 <211> 245
 <212> PRT
 <213> Homo sapiens

<400> 6
 Met Ala Leu Leu Phe Pro Leu Leu Ala Ala Leu Val Met Thr Ser Tyr
 1 5 10 15

Ser Pro Val Gly Ser Leu Gly Cys Asp Leu Pro Gln Asn His Gly Leu

20 25 30
 Leu Ser Arg Asn Thr Leu Val Leu Leu His Gln Met Arg Arg Ile Ser
 35 40 45
 Pro Phe Leu Cys Leu Lys Asp Arg Arg Asp Phe Arg Phe Pro Gln Glu
 50 55 60
 Met Val Lys Gly Ser Gln Leu Gln Lys Ala His Val Met Ser Val Leu
 65 70 75 80
 His Glu Met Leu Gln Gln Ile Phe Ser Leu Phe His Thr Glu Arg Ser
 85 90 95
 Ser Ala Ala Trp Asn Met Thr Leu Leu Asp Gln Leu His Thr Glu Leu
 100 105 110
 His Gln Gln Leu Gln His Leu Glu Thr Cys Leu Leu Gln Val Val Gly
 115 120 125
 Glu Gly Glu Ser Ala Gly Ala Ile Ser Ser Val Pro Gln Leu Ser Ser
 130 135 140
 Leu Glu Leu Arg Arg Tyr Phe His Arg Ile Asp Asn Phe Leu Lys Glu
 145 150 155 160
 Lys Lys Tyr Ser Asp Cys Ala Trp Glu Ile Val Arg Val Glu Ile Arg
 165 170 175
 Arg Cys Leu Tyr Tyr Phe Tyr Lys Phe Thr Ala Leu Pro Ala Leu Thr
 180 185 190
 Leu Arg Arg Tyr Phe Gln Gly Ile Arg Val Tyr Leu Lys Glu Lys Lys
 195 200 205
 Tyr Ser Asp Cys Ala Trp Glu Val Val Arg Met Glu Ile Met Lys Ser
 210 215 220
 Leu Phe Leu Ser Thr Asn Met Gln Glu Arg Leu Arg Ser Lys Asp Arg
 225 230 235 240
 Asp Leu Gly Ser Ser
 245

<210> 7
 <211> 189
 <212> PRT
 <213> Homo sapiens

<400> 7
 Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
 1 5 10 15
 Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
 20 25 30
 Gly Asn Arg Arg Ala Leu Ile Leu Leu Gly Gln Met Gly Arg Ile Ser
 35 40 45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Arg Ile Pro Gln Glu
 50 55 60
 Glu Phe Asp Gly Asn Gln Phe Gln Asp Ala Gln Ala Ile Ser Val Leu
 65 70 75 80
 His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Glu Asp Ser
 85 90 95
 Ser Ala Ala Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
 100 105 110
 Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
 115 120 125
 Val Glu Glu Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
 130 135 140
 Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Ile Glu Arg Lys Tyr Ser
 145 150 155 160
 Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser
 165 170 175
 Phe Ser Thr Asn Leu Gln Lys Arg Leu Arg Arg Lys Asp
 180 185

<210> 8
 <211> 189
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
 1 5 10 15
 Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
 20 25 30
 Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
 35 40 45
 Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
 50 55 60
 Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala His Val Met Ser Val Leu
 65 70 75 80
 His Glu Met Leu Gln Gln Ile Phe Ser Leu Phe His Thr Glu Arg Ser
 85 90 95
 Ser Ala Ala Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
 100 105 110
 Asn Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
 115 120 125
 Val Glu Glu Thr Pro Leu Met Asn Val Asp Ser Ile Leu Ala Val Lys
 130 135 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145 150 155 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
165 170 175

Leu Ser Lys Ile Phe Gln Glu Arg Leu Arg Arg Lys Glu
180 185

<210> 9

<211> 195

<212> PRT

<213> Homo sapiens

<400> 9

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

Ser Pro Val Gly Ser Leu Gly Cys Asp Leu Pro Gln Asn His Gly Leu
20 25 30

Leu Ser Arg Asn Thr Leu Val Leu Leu His Gln Met Arg Arg Ile Ser
35 40 45

Pro Phe Leu Cys Leu Lys Asp Arg Arg Asp Phe Arg Phe Pro Gln Glu
50 55 60

Met Val Lys Gly Ser Gln Leu Gln Lys Ala His Val Met Ser Val Leu
65 70 75 80

His Glu Met Leu Gln Gln Ile Phe Ser Leu Phe His Thr Glu Arg Ser
85 90 95

Ser Ala Ala Trp Asn Met Thr Leu Leu Asp Gln Leu His Thr Glu Leu
100 105 110

His Gln Gln Leu Gln His Leu Glu Thr Cys Leu Leu Gln Val Val Gly
115 120 125

Glu Gly Glu Ser Ala Gly Ala Ile Ser Ser Pro Ala Leu Thr Leu Arg
130 135 140

Arg Tyr Phe Gln Gly Ile Arg Val Tyr Leu Lys Glu Lys Lys Tyr Ser
145 150 155 160

Asp Cys Ala Trp Glu Val Val Arg Met Glu Ile Met Lys Ser Leu Phe
165 170 175

Leu Ser Thr Asn Met Gln Glu Arg Leu Arg Ser Lys Asp Arg Asp Leu
180 185 190

Gly Ser Ser
195

<210> 10

<211> 378

<212> PRT

<213> Homo sapiens

<400> 10

```

Met Pro Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
 1             5             10             15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
          20             25             30

Gly Asn Arg Arg Ala Trp Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
          35             40             45

His Phe Ser Cys Leu Lys Asp Arg Tyr Asp Phe Gly Phe Pro Gln Glu
 50             55             60

Val Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Ala Phe
 65             70             75             80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
          85             90             95

Ser Ala Ala Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Ile Glu Leu
          100            105            110

Phe Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Thr Gln Glu Val Gly
          115            120            125

Val Glu Glu Ile Ala Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
          130            135            140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Met Gly Lys Lys Tyr Ser
          145            150            155            160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
          165            170            175

Phe Ser Thr Asn Leu Gln Lys Gly Leu Arg Arg Lys Asp Met Pro Leu
          180            185            190

Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr Lys Ser Ile
          195            200            205

Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg
          210            215            220

Arg Ala Trp Ile Leu Leu Ala Gln Met Gly Arg Ile Ser His Phe Ser
          225            230            235            240

Cys Leu Lys Asp Arg Tyr Asp Phe Gly Phe Pro Gln Glu Val Phe Asp
          245            250            255

Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Ala Phe His Glu Met
          260            265            270

Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala
          275            280            285

Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Ile Glu Leu Phe Gln Gln
          290            295            300

```

Leu Asn Asp Leu Glu Ala Cys Val Thr Gln Glu Val Gly Val Glu Glu
305 310 315 320

Ile Ala Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe
325 330 335

Gln Arg Ile Thr Leu Tyr Leu Met Gly Lys Lys Tyr Ser Pro Cys Ala
340 345 350

Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser Phe Ser Thr
355 360 365

Asn Leu Gln Lys Gly Leu Arg Arg Lys Asp
370 375

<210> 11

<211> 195

<212> PRT

<213> Homo sapiens

<400> 11

Met Ala Phe Val Leu Ser Leu Leu Met Ala Leu Val Leu Val Ser Tyr
1 5 10 15

Gly Pro Gly Arg Ser Leu Gly Cys Tyr Leu Ser Glu Asp His Met Leu
20 25 30

Gly Ala Arg Glu Asn Leu Arg Leu Leu Ala Arg Met Asn Arg Leu Ser
35 40 45

Pro His Pro Cys Leu Gln Asp Arg Lys Asp Phe Gly Leu Pro Gln Glu
50 55 60

Met Val Glu Gly Asn Gln Leu Gln Lys Asp Gln Ala Ile Ser Val Leu
65 70 75 80

His Glu Met Leu Gln Gln Cys Phe Asn Leu Phe Tyr Thr Glu His Ser
85 90 95

Ser Ala Ala Trp Asn Thr Thr Leu Leu Glu Gln Leu Cys Thr Gly Leu
100 105 110

Gln Gln Gln Leu Glu Asp Leu Asp Ala Cys Leu Gly Pro Val Met Gly
115 120 125

Glu Lys Asp Ser Asp Met Gly Arg Met Gly Pro Ile Leu Thr Val Lys
130 135 140

Lys Tyr Phe Gln Gly Ile His Val Tyr Leu Lys Glu Lys Glu Tyr Ser
145 150 155 160

Asp Cys Ala Trp Glu Ile Ile Arg Met Glu Met Met Arg Ala Leu Ser
165 170 175

Ser Ser Thr Thr Leu Gln Lys Arg Leu Arg Lys Met Gly Gly Asp Leu
180 185 190

Asn Ser Leu
195

<210> 12
 <211> 196
 <212> PRT
 <213> Homo sapiens

<400> 12
 Met Ala Phe Val Leu Ser Leu Leu Met Ala Leu Val Leu Val Ser Tyr
 1 5 10 15
 Gly Pro Gly Gly Ser Leu Gly Cys Tyr Leu Ser Gln Arg Leu Met Leu
 20 25 30
 Asp Ala Arg Glu Asn Leu Lys Leu Leu Glu Pro Met Asn Arg Leu Ser
 35 40 45
 Pro His Ser Cys Leu Gln Asp Arg Lys Asp Phe Gly Leu Pro Gln Glu
 50 55 60
 Met Val Glu Gly Asp Gln Leu Gln Lys Asp Gln Ala Phe Pro Val Leu
 65 70 75 80
 Tyr Glu Met Leu Gln Gln Thr Phe Asn Leu Phe His Thr Glu His Ser
 85 90 95
 Ser Ala Ala Trp Asp Thr Thr Leu Leu Glu Gln Leu Cys Thr Gly Leu
 100 105 110
 Gln Gln Gln Leu Glu Asp Leu Asp Thr Cys Cys Arg Gly Gln Val Met
 115 120 125
 Gly Glu Glu Asp Ser Glu Leu Gly Asn Met Asp Pro Ile Val Thr Val
 130 135 140
 Lys Lys Tyr Phe Gln Gly Ile Tyr Asp Tyr Leu Gln Glu Lys Gly Tyr
 145 150 155 160
 Ser Asp Cys Ala Trp Glu Ile Val Arg Val Glu Met Met Arg Ala Leu
 165 170 175
 Thr Val Ser Thr Thr Leu Gln Lys Arg Leu Thr Lys Met Gly Gly Asp
 180 185 190
 Leu Asn Ser Pro
 195

<210> 13
 <211> 170
 <212> PRT
 <213> Homo sapiens

<400> 13
 Met Ala Gln Ile Tyr Leu Val Met Ala Gly Val Met Leu Cys Ser Ile
 1 5 10 15
 Ser Val Cys Phe Leu Asp Gln Asn Leu Ser Ala Val His Cys Val Glu
 20 25 30

Lys Arg Glu Ile Phe Lys His Leu Gln Glu Ile Lys Lys Ile Pro Ser
 35 40 45

Gln Leu Cys Leu Lys Asp Arg Ile Asp Phe Lys Phe Pro Trp Lys Arg
 50 55 60

Glu Ser Ile Thr Gln Leu Gln Lys Asp Gln Ala Phe Pro Val Leu Tyr
 65 70 75 80

Glu Met Leu Gln Gln Thr Phe Asn Leu Phe His Thr Glu His Ser Ser
 85 90 95

Ala Ala Trp Asn Thr Thr Leu Leu Asp Gln Leu Leu Ser Ser Leu Asp
 100 105 110

Leu Gly Leu Arg Arg Leu Glu His Met Lys Lys Asp Asn Met Asp Cys
 115 120 125

Pro His Val Gly Ser Ala Leu Arg Lys Tyr Phe Gln Gly Ile Gly Leu
 130 135 140

Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Ile Val Arg
 145 150 155 160

Val Glu Ile Glu Arg Cys Phe Ser Leu Thr
 165 170

<210> 14

<211> 212

<212> PRT

<213> Homo sapiens

<400> 14

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

Gly Leu Leu Leu Val Leu Pro Ala Ala Phe Pro Ala Pro Val Pro Pro
 20 25 30

Gly Glu Asp Ser Lys Asp Val Ala Ala Pro His Arg Gln Pro Leu Thr
 35 40 45

Ser Ser Glu Arg Ile Asp Lys Gln Ile Arg Tyr Ile Leu Asp Gly Ile
 50 55 60

Ser Ala Leu Arg Lys Glu Thr Cys Asn Lys Ser Asn Met Cys Glu Ser
 65 70 75 80

Ser Lys Glu Ala Leu Ala Glu Asn Asn Leu Asn Leu Pro Lys Met Ala
 85 90 95

Lys Glu Asp Gly Cys Phe Gln Ser Gly Phe Asn Glu Glu Thr Cys Leu
 100 105 110

Val Lys Ile Ile Thr Gly Leu Leu Glu Phe Glu Val Tyr Leu Glu Tyr
 115 120 125

Leu Gln Asn Arg Phe Glu Ser Ser Glu Glu Gln Ala Arg Ala Val Gln
 130 135 140

Met Ser Thr Lys Val Leu Ile Gln Phe Leu Gln Lys Lys Ala Lys Asn
145 150 155 160

Leu Asp Ala Ile Thr Thr Pro Asp Pro Thr Thr Asn Ala Ser Leu Leu
165 170 175

Thr Lys Leu Gln Ala Gln Asn Gln Trp Leu Gln Asp Met Thr Thr His
180 185 190

Leu Ile Leu Arg Ser Phe Lys Glu Phe Leu Gln Ser Ser Leu Arg Ala
195 200 205

Leu Arg Gln Met
210

<210> 15

<211> 186

<212> PRT

<213> Homo sapiens

<400> 15

Met Thr His Arg Cys Leu Leu Gln Met Val Leu Leu Leu Cys Phe Ser
1 5 10 15

Thr Thr Ala Leu Ser Arg Ser Tyr Ser Leu Leu Arg Phe Gln Gln Arg
20 25 30

Arg Ser Leu Ala Leu Cys Gln Lys Leu Leu Arg Gln Leu Pro Ser Thr
35 40 45

Pro Gln His Cys Leu Glu Ala Arg Met Asp Phe Gln Met Pro Glu Glu
50 55 60

Met Lys Gln Ala Gln Gln Phe Gln Lys Glu Asp Ala Ile Leu Val Ile
65 70 75 80

Tyr Glu Met Leu Gln Gln Ile Phe Asn Ile Leu Thr Arg Asp Phe Ser
85 90 95

Ser Thr Gly Trp Ser Glu Thr Ile Ile Glu Asp Leu Leu Glu Glu Leu
100 105 110

Tyr Glu Gln Met Asn His Leu Glu Pro Ile Gln Lys Glu Ile Met Gln
115 120 125

Lys Gln Asn Ser Thr Met Gly Asp Thr Thr Val Leu His Leu Arg Lys
130 135 140

Tyr Tyr Phe Asn Leu Val Gln Tyr Leu Lys Ser Lys Glu Tyr Asn Arg
145 150 155 160

Cys Ala Trp Thr Val Val Arg Val Gln Ile Leu Arg Asn Phe Ser Phe
165 170 175

Leu Thr Arg Leu Thr Gly Tyr Leu Arg Glu
180 185

<210> 16
 <211> 29
 <212> DNA
 <213> Homo sapiens

<400> 16
 ggccgcatat gctggactgt aacttactg 29

<210> 17
 <211> 33
 <212> DNA
 <213> Homo sapiens

<400> 17
 ggccgcggtta ccttatttcc tcctgaatag agc 33

<210> 18
 <211> 38
 <212> DNA
 <213> Homo sapiens

<400> 18
 ggccgggatc cgccatcatg agcaccaaac ctgatatg 38

<210> 19
 <211> 33
 <212> DNA
 <213> Homo sapiens

<400> 19
 ggccgcggtta ccttatttcc tcctgaatag agc 33

<210> 20
 <211> 156
 <212> PRT
 <213> Homo sapiens

<400> 20
 Met Thr Tyr Arg Cys Leu Leu Gln Met Val Leu Leu Leu Cys Phe Ser
 1 5 10 15
 Thr Thr Ala Leu Ser Arg Ser Tyr Ser Leu Leu Arg Phe Gln Gln Arg
 20 25 30
 Gln Ser Leu Lys Glu Cys Gln Lys Leu Leu Gly Gln Leu Pro Ser Thr
 35 40 45
 Ser Gln His Cys Leu Glu Ala Arg Met Asp Phe Gln Met Pro Glu Glu
 50 55 60
 Met Lys Gln Glu Gln Gln Phe Gln Lys Glu Asp Ala Ile Leu Val Met
 65 70 75 80
 Tyr Glu Val Leu Gln His Ile Phe Gly Ile Leu Thr Arg Asp Phe Ser
 85 90 95

Ser Thr Gly Trp Asn Ser Thr Thr Glu Asp Thr Ile Val Pro His Leu
 100 105 110

Gly Lys Tyr Tyr Phe Asn Leu Met Gln Tyr Leu Glu Ser Lys Glu Tyr
 115 120 125

Asp Arg Cys Ala Trp Thr Val Val Gln Val Gln Ile Leu Thr Asn Val
 130 135 140

Ser Phe Leu Met Arg Leu Thr Gly Tyr Val Arg Asp
 145 150 155

<210> 21

<211> 166

<212> PRT

<213> Homo sapiens

<400> 21

Met Ser Tyr Asn Leu Leu Gly Phe Leu Gln Arg Ser Ser Asn Phe Gln
 1 5 10 15

Cys Gln Lys Leu Leu Trp Gln Leu Asn Gly Arg Leu Glu Tyr Cys Leu
 20 25 30

Lys Asp Arg Met Asn Phe Asp Ile Pro Glu Glu Ile Lys Gln Leu Gln
 35 40 45

Gln Phe Gln Lys Glu Asp Ala Ala Leu Thr Ile Tyr Glu Met Leu Gln
 50 55 60

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

Glu Thr Ile Val Glu Asn Leu Leu Ala Asn Val Tyr His Gln Ile Asn
 85 90 95

His Leu Lys Thr Val Leu Glu Glu Lys Leu Glu Lys Glu Asp Phe Thr
 100 105 110

Arg Gly Lys Leu Met Ser Ser Leu His Leu Lys Arg Tyr Tyr Gly Arg
 115 120 125

Ile Leu His Tyr Leu Lys Ala Lys Glu Tyr Ser His Cys Ala Trp Thr
 130 135 140

Ile Val Arg Val Glu Ile Leu Arg Asn Phe Tyr Phe Ile Asn Arg Leu
 145 150 155 160

Thr Gly Tyr Leu Gly Asn
 165

说明书附图

角质形成细胞衍生的干扰素

1 CCA CGC GTC CGG GAT TTT TTA GCT TGC AAA AAA AAT GAG CAC CAA ACC TGA TAT GAT TCA 60
 1 M S T K P D M L Q 9

61 AAA GTG TTT GTG GCT TGA GAT CCT TAT GGG TAT ATT CAT TGC TGG CAC CCT ATC CCT GGA 120
 9 K C L W L E I L M G I F I A G T L S L D 29

121 CTG TAA CTT ACT GAA CGT TCA CCT GAG AAG AGT CAC CTG GCA AAA TCT GAG ACA TCT GAG 180
 29 C N L L N V H L R R V T W Q N L R H L S 49

181 TAG TAT GAG CAA TTC ATT TCC TGT AGA ATG TCT ACG AGA AAA CAT AGC TTT TGA GTT GCC 240
 49 S M S N S F P V E C L R E N I A F E L P 69

241 CCA AGA GTT TCT GCA ATA CAC CCA ACC TAT GAA GAG GGA CAT CAA GAA GGC CTT CTA TGA 300
 69 Q E F L Q Y T Q P M K R D I K K A F Y E 89

301 AAT GTC CCT ACA GGC CTT CAA CAT CTT CAG CCA ACA CAC CTT CAA ATA TTG GAA AGA GAG 360
 89 M S L Q A F N I F S Q H T F K Y W K E R 109

361 ACA CCT CAA ACA AAT CCA AAT AGG ACT TGA TCA GCA AGC AGA GTA CCT GAA CCA ATG CTT 420
 109 H L K Q I Q I G L D Q Q A E Y L N Q C L 129

421 GGA GGA AGA CGA GAA TGA AAA TGA AGA CAT GAA AGA AAT GAA AGA GAA TGA GAT GAA ACC 480
 129 E E D E N E N E D M K E M K E N E M K P 149

481 CTC AGA AGC CAG GGT CCC CCA GCT GAG CAG CCT GGA ACT GAG GAG ATA TTT CCA CAG GAT 540
 149 S E A R V P Q L S S L E L R R Y F H R I 169

541 AGA CAA TTT CCT GAA AGA AAA GAA ATA CAG TGA CTG TGC CTG GGA GAT TGT CCG AGT GGA 600
 169 D N F L K E K K Y S D C A W E I V R V E 189

601 AAT CAG AAG ATG TTT GTA TTA CTT TTA CAA ATT TAC AGC TCT ATT CAG GAG GAA ATA AGA 660
 189 I R R C L Y Y F Y K F T A L F R R K * 207

661 ATC ATC TAC CTT CAA GCA AGA ATT AAC AGA GAT TGT GGC TAC GCA AAT GCA CCA AAA AAG 720
 721 GGT GAA ATA TAT CTG AAA TGT ACC TGG TTC TGC CCT TGG AAG CCA CTT CCT GCT CAT GCC 780
 781 ACT AAC AGC ATG CTG CCA AAC TGT TCA GAT TCA AGA TTA TTC CAA GCG CAG GGC CCA AAT 840
 841 GTT ATA GCC AAA GAA AGT CTT ATG ATA AAA GTG AGG CAA ATT TCA GCC AAG AAG TTA GAA 900
 901 GAG ATG TTT AAA AGA ACA AGA ACA AAT TGT GGA TCA TGG TAT ATG CAG GCT ATC AGC AGA 960
 961 AGG ATC AGA CAA TAA AAT GAG TTA GTG CAA ACC ATT TAG TAA AAA TAA CTA TCA GCA GAG 1020
 1021 TTG TTC CAG ATT AAA AAT AGT ACT ACA AGC TTG TAA AGG AGT TAG GAC ATG CAA GCT ACT 1080
 1081 GAG CAT AAA ATA TAT ACT TGC TAT TTT TCA TGA CTT TCT CTA ATA AAG TCT TTG ACT GTT 1140
 1141 CTC TCT AAT AAA AAA AAA AAA AAA AAA AAA 1170

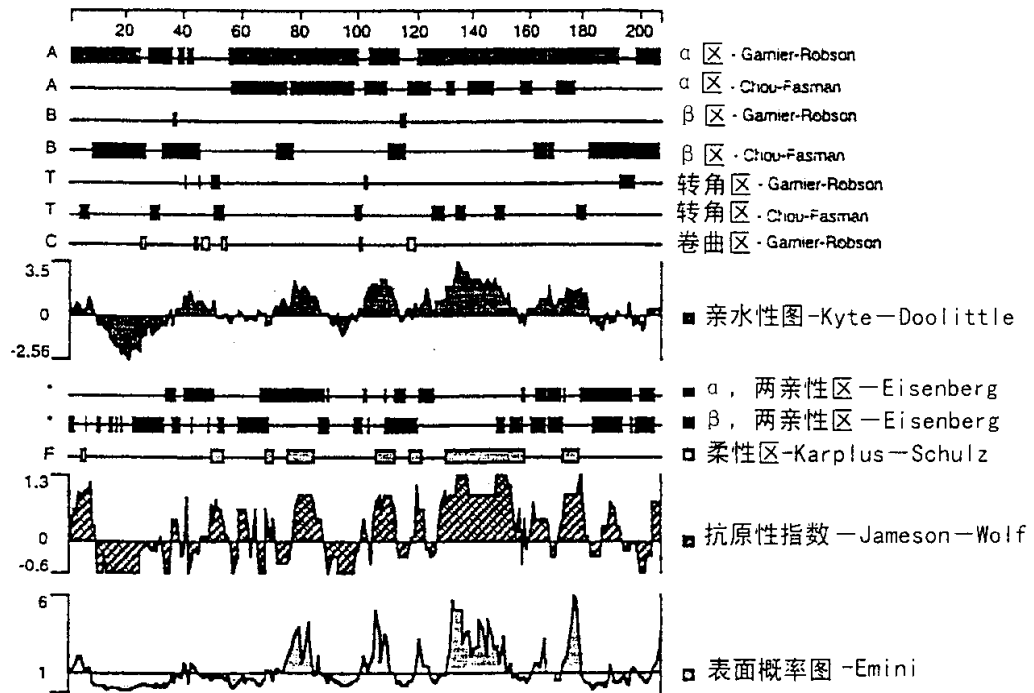


图3

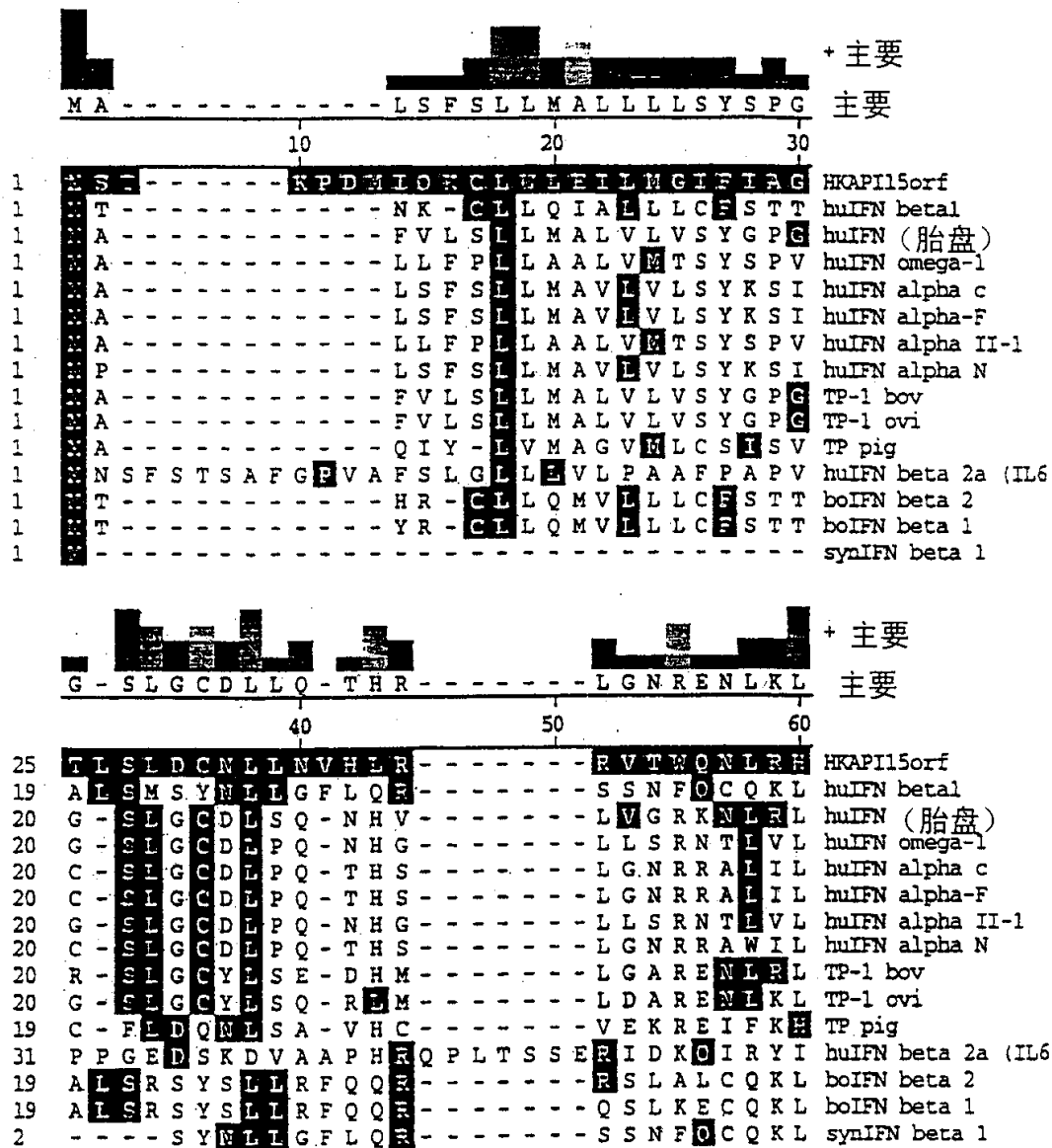


图4A

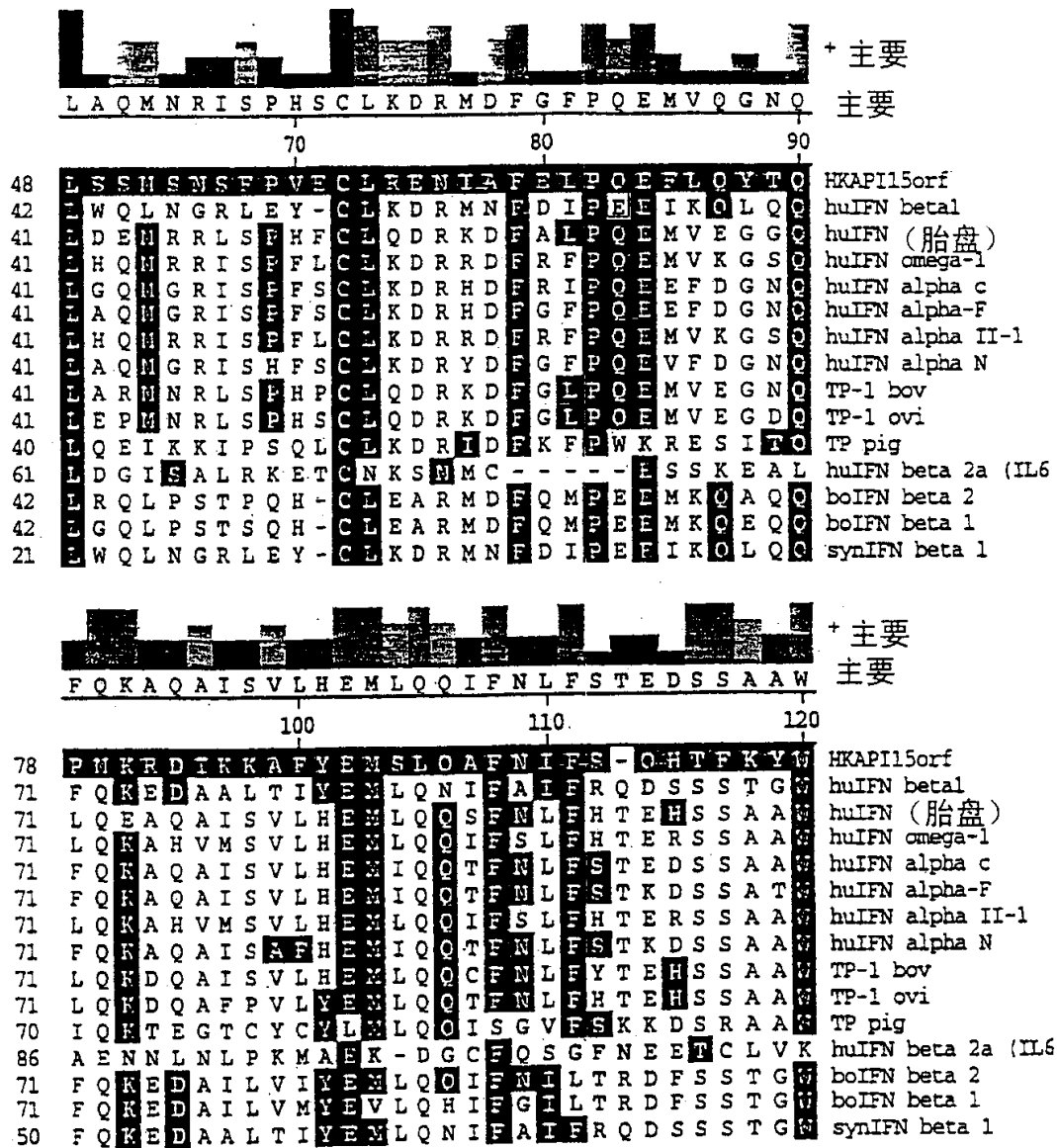


图4B

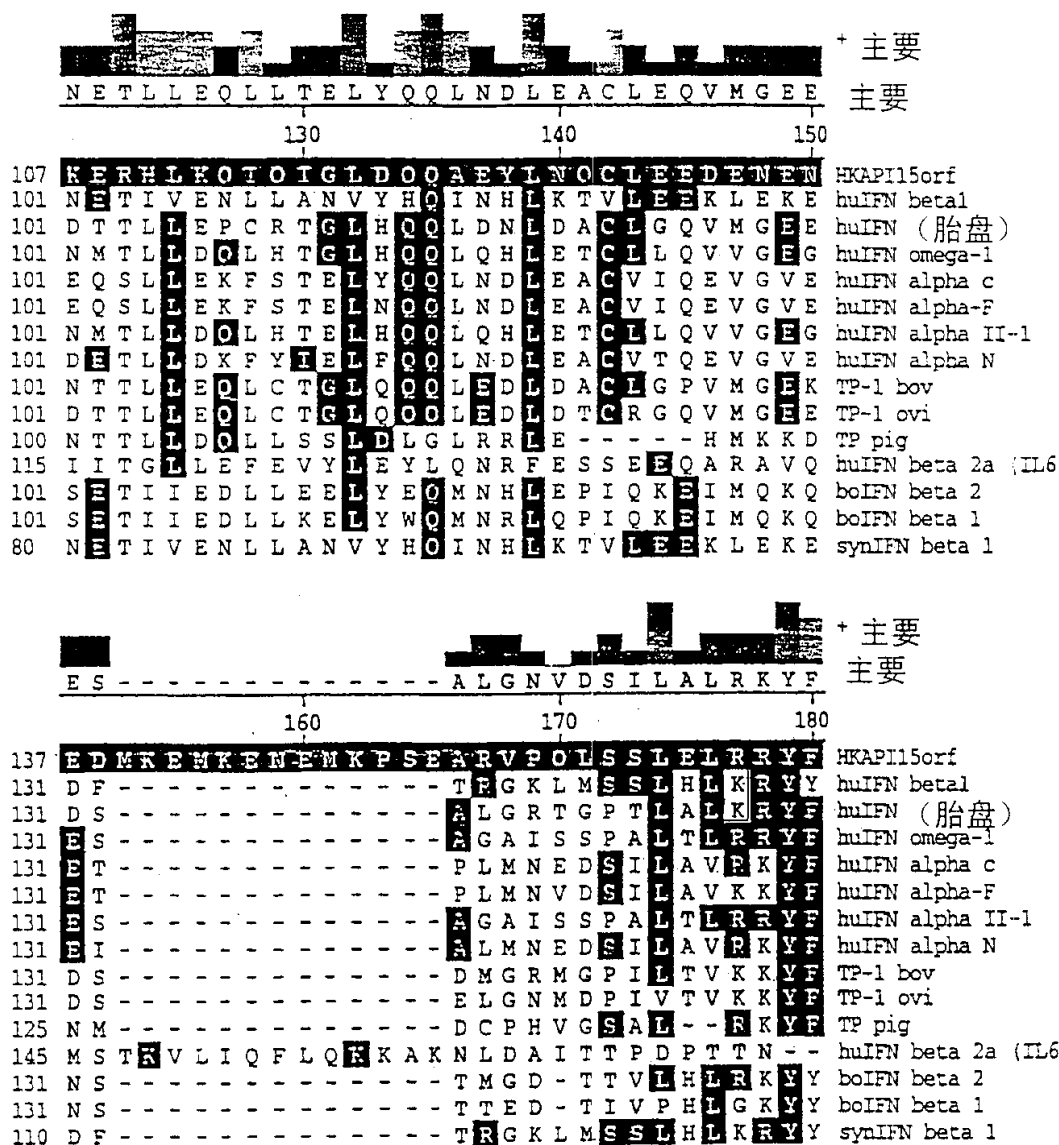


图4C

