



[12] 发明专利申请公开说明书

[21] 申请号 88103166.6

[51] Int.Cl⁴
G01N 33/53

[43] 公开日 1989年1月25日

[22] 申请日 88.5.28

[30] 优先权

[32] 87.5.29 [33] US [31] 056,553

[71] 申请人 史密丝克莱恩贝曼公司

地址 美国宾夕法尼亚州

[72] 发明人 克里丝汀·玛丽·迪宝克
马丁·罗森堡

[74] 专利代理机构 中国专利代理有限公司
代理人 林玉贞

G01N 33/563

G01N 33/569 C12N 15/00

说明书页数: 21 附图页数:

[54] 发明名称 使用嵌合抗原的免疫诊断检测法

[57] 摘要

本发明涉及用于对感染或其他病态进行诊断和预后检测的嵌合抗原。本发明提供一种含有嵌合抗原的药盒、一种制备嵌合 HIV 抗原的方法、一种检测样品中抗 HIV 抗体的方法及一种表达载体。

权 利 要 求 书

1. 一种用于检测抗某种特定抗原的抗体的免疫诊断药盒，包括
 - a) 一种固相载体，
 - b) 一种结合到固相载体上的抗体，
 - c) 一种具有第一和第二抗原区的嵌合抗原，其中第一抗原区可与结合的抗体反应，但不与识别第二抗原区的抗体产生交叉反应。

2. 权利要求 1 的药盒，其中第二抗原区代表病原微生物、蛋白质、病毒或细胞。

3. 权利要求 2 的药盒，其中第一抗原区代表不能识别动物体液中结合位点的蛋白质或蛋白质片段。

4. 权利要求 3 的药盒，其中第二抗原区是一种衍生自与特定肿瘤、遗传异常及细菌、真菌、病毒或寄生虫性疾病有关的蛋白质的抗原。

5. 一种制备嵌合 H I V 抗原的方法，包括使 g a l K 基因的一部分与 H I V 基因片段融合，将所得的融合基因片段引入某种生物体内，培养该生物体使之表达嵌合 H I V 抗原。

6. 权利要求 5 的方法，其中嵌合 H I V 抗原在大肠杆菌中表达时以高水平积聚且稳定。

7. 权利要求 5 的方法，其中嵌合抗原是由与 H I V 的 p24 g a g 编码顺序融合的 g a l K 基因编码顺序编码的。

8. 权利要求 5 的方法，其中嵌合抗原是由与 H I V 的 p 5 5 g a g 编码顺序融合的 g a l K 基因编码顺序编码的。

9. 权利要求 5 的方法，其中嵌合抗原是由与 H I V 的反转录酶

—核酸内切酶 (RTendo) 编码顺序融合的 *gal K* 基因编码顺序编码的。

10. 权利要求5的方法, 其中嵌合抗原是由与 HIV 的 p14 *env* 编码顺序融合的 *gal K* 基因编码顺序编码的。

11. 权利要求5的方法, 其中嵌合抗原是由与 HIV 的 p120 *env* 氨基端一半 (N) 编码顺序融合的 *gal K* 基因编码顺序编码的。

12. 权利要求5的方法, 其中嵌合抗原是由与 HIV 的 p120 *env* 羧基端一半 (C) 编码顺序融合的 *gal K* 基因编码顺序编码的。

13. 一种检测样品中抗 HIV 抗体的方法, 包括使样品与结合到固相载体上的 *gal K*—HIV 嵌合抗原接触, 从而只有含 HIV 特异性免疫球蛋白的样品结合到 HIV 嵌合抗原上。

14. 权利要求13的方法, 其中嵌合抗原借助于 *gal K* 区连接到固相载体上, 从而使存在于该分子上的所有 HIV 抗原决定基均可被样品识别。

15. 一种载体, 包括一个启动子和 *gal K* 基因的一部分, 后者邻接有不与 *gal K* 起交叉反应的抗原的编码顺序。

16. 权利要求15的 *gal K* 翻译融合载体, 称为 *gal K*—p24 *gag*。

17. 权利要求15的 *gal K* 翻译融合载体, 称为 *gal K*—p55 *gag*。

18. 权利要求15的 *gal K* 翻译融合载体, 称为 *gal K*—RTendo。

19. 权利要求15的galK翻译融合载体,称为galK-p41 env。

20. 权利要求15的galK翻译融合载体,称为galK-p120 env(N)。

21. 权利要求15的galK翻译融合载体,称为galK-p120 env(C)。

22. 用权利要求16、17、18、19、20或21的galK翻译融合载体转化的大肠杆菌。

23. 一种嵌合HIV抗原,包括一个HIV抗原区和一个第二抗原区,第二抗原区不与识别HIV抗原区的抗体起交叉反应。

24. 按照权利要求5-12中任一项的方法制备的嵌合HIV抗原。

使用嵌合抗原的免疫诊断检测法

本发明涉及用于对感染或其他病态进行诊断和预后检测的嵌合抗原。

各种免疫诊断检测法都是基于抗原与其相应抗体之间的结合亲和力。当形成抗原—抗体复合物时，可用各种检测系统，如放射性同位素、荧光标记或酶如辣根过氧化物酶来检测复合物。

使用免疫测定技术多年来已成为标准实验室常规。所有各种检测法的基本原则都是使标记的抗原与预选的抗体结合，然后测定结合的复合物的量。近年来，使用单克隆抗体代替多克隆抗体，提高了测定的灵敏度和特异性。然而，使用单克隆抗体代替多克隆抗体，并不能完全解决特定检测的特异性或灵敏度问题。在抗体抗原复合物形成之后，仍然继续竞争分子上的特定结合位点。

许多免疫检测法依赖于结合复合物对最初用以结合到固相载体上的同一抗体的识别。因为竞争可利用的结合位点，所以在检测的两个关键点上依赖同一抗体，会进一步降低灵敏度。另外，缺乏抗体的特异性可导致出现“假阳性”，尤其是在这样一些免疫检测中，其中正在对疾病进展的各阶段进行诊断，并且，与特定疾病阶段有关的抗原决定基的识别具有重要意义。为了解决这个问题，已设计了“两位点”或“夹心”免疫测定法。在这种类型的免疫检测法中，有两个抗体结合到抗原表面上，其中一个是放射标记的或已有某些其他检测分子结合到该抗体上，而另一抗体则结合到固相载体上。这样，减少了对分子上特定结合位点的竞争。夹心法则因使用高度纯化的抗体而可进一步

提高特异性。

对于某些疾病，早期诊断以及鉴定疾病的阶段可能是十分重要的，其中一个例子就是获得性免疫缺陷综合症(爱滋病)H I V——爱滋病及相关疾病的致病因子，是还原病毒科的一个成员。存在有几种包括人类Ⅲ型嗜T淋巴细胞病毒(T - Lymphotropic Virus , HTLV - Ⅲ)、淋巴结病病毒(L A V)和爱滋病相关还原病毒(A R V)在内的H I V的分离物。最近发现定名为2型H I V的相关还原病毒与西非的爱滋病有关(Guyader et al . , Nature 326:662, 1987)。

H I V基因组的分子特征表明，该病毒具有与其他还原病毒一样的整个g a g - p o l - e n v结构(Ratner et al . , Nature 313 : 277 , 1985 ; Wain - Hobson et al . , Cell 40 : 9 , 1985)。此外，它至少含有5个未见于较常见的还原病毒的基因：S o r 、t a t 3 、a r t / t r s 、3' o r f和R。

鉴于爱滋病的严重性，早期并准确地诊断H I V感染，以及由血库中检出并除掉H I V污染的样品，是极其重要的。Gallo等人在美国专利4,520,113号中公开了使用得自HTLV - Ⅲ的抗原检测血清中存在的抗HTLV - Ⅲ抗体。Montagnier等人在E P - A - 138,667号中公开了使用特定的L A V抗原检测病毒感染。

Papas等人在美国专利申请6,664,972号(德温特登记号85-110268/18)中公开了HTLV - I被膜蛋白编码顺序在大肠杆菌内的表达，以及所表达的蛋白质在检测HTLV - I感染中的应用。

Crowl等人(Cell , 41 : 979 , 1985)报导了HTLV - Ⅲ被膜蛋白基因(e n v)的某些部分在大肠杆菌内的表达，以及这些蛋

白质在检测 HTLV-III 感染中的应用。Casey 等人 (J. Virol., 55 : 417 , 1985) 报导了一种 gag 基因产物 (一种定名为 p24 的 HTLV-III 内部结构蛋白) 的纯化方法, 以及该蛋白在检测 HTLV-III 感染中的应用。

本领域的许多研究者从事在细菌中表达爱滋病还原病毒的各种抗原的研究。最近构建了具有 HIV 特定片段顺序的一些新肽, 用作接触爱滋病病毒的检测试剂 (Cosand , 美国专利 4 , 629 , 783 号; Capon 等人, EP-A-187041 号)。Steimer 等人 (J. Virology , 58 : 9 , 1986) 描述了使用来自爱滋病还原病毒聚合酶基因 (pol) 的核酸内切酶区域的重组多肽, 检测被感染者血清抗体的方法。

Chang 在 EP-A-185,444 中描述了一种用含有编码 env 基因顺序的 HTLV-III cDNA 的重组载体转化的细胞所表达的多肽, 以及这种免疫活性多肽在检测 HTLV-III 中的应用。Dowbenko 等人描述了爱滋病还原病毒 p24 gag 蛋白在大肠杆菌内的表达, 及其作为诊断试剂的应用。Ghrayeb 等人 (DNA , 5 : 93 , 1986) 和 Steimer 等人 (Virology , 150 : 283 , 1986) 还报导了 HTLV-III 核心抗原在大肠杆菌内的合成及其与人血清的免疫反应活性。

本发明的一个实施方案中描述了用于检测针对特定抗原的抗体的免疫诊断药盒 (device)。该药盒包括一种固相载体, 与固相载体相结合的抗体及具有第一和第二抗原区的嵌合抗原, 其中第一抗原区与已结合的抗体反应, 但不与识别第二抗原区的抗体产生交叉反应。

本发明的另一实施方案是一种制备嵌合 HIV 抗原的方法, 包括

使大肠杆菌半乳糖激酶 (gal K) 基因的一部分与 HIV 基因片段相融合，并将所得的融合基因片段引入一种生物体内，然后培养该生物以使之表达嵌合 HIV 抗原。

本发明的另一实施方案描述了一种表达载体，包括一个启动子及一个 gal K 基因部分，后者邻接一个不与 gal K 起交叉反应的抗原的编码顺序。

在本发明的再一个实施方案中，描述了一种检测样品中抗 HIV 抗体的方法。使样品与结合到固相载体上的 gal K-HIV 嵌合抗原接触，使得只有含 HIV 特异性免疫球蛋白的样品与嵌合 HIV 抗原结合，从而完成抗体的检测。

本发明的嵌合抗原是一种融合蛋白，编码该融合蛋白的顺序包含编码不能与目的蛋白产生交叉反应的基因产物的一部分大肠杆菌基因，以及与其融合的编码目的蛋白的基因片段。目的基因片段可以是编码与特定疾病或疾病状态（如爱滋病、淋病、梅毒等性传播的疾病、衣原体病或乙型肝炎）有关的蛋白质的基因片段。反映特定疾病或疾病状态的其他特异性基因产物，如与特异性肿瘤、遗传异常、其他细菌、真菌、病毒或寄生虫性疾病（如由疟原虫、锥虫、假丝酵母、毕赤酵母、沙门氏菌、假单孢菌等引起的疾病）有关的蛋白质，亦可作为融合蛋白而表达。

因为嵌合抗原的一部分特异地针对特定疾病或疾病状态，所以用嵌合抗原来诊断或预后疾病是有利的。使用融合系统如大肠杆菌 gal K 系统还可提供一种高水平表达嵌合抗原的方法，特别是当嵌合抗原不能由其未融合的基因片段来表达或表达很差的情况下。嵌合抗原的 gal K 部分可起到连接其他特异性蛋白质的“柄”的作用。

由于其特异性，嵌合抗原还提供了一种纯化蛋白质的有效方法。可以使用的其他galK样“柄”包括流感NS-1蛋白质的某些部分、噬菌体 λ O蛋白、大肠杆菌 β -半乳糖苷酶蛋白、或任何可在大肠杆菌或其他细菌、酵母、高等真核细胞等重要重组宿主内很好表达的蛋白质。但在选择“柄”时，必须使融合蛋白的柄部分具有免疫识别能力，但又不被样品中的结合位点识别。

嵌合抗原在免疫诊断药盒中是有用的，它不仅提供了对目的蛋白的高度特异性，而且提高了检测方法的灵敏度。因为嵌合抗原的一部分不被样品（如人或动物血液、血浆、血清、唾液、痰、排泄物或其他体液）所识别，所以目的蛋白的所有抗原决定基均可用来结合样品。这一特性大大提高了特异性，并且是对使用特异性单克隆抗体的一种改进。使用单克隆抗体，有可能出现两种单克隆抗体识别目的抗原上交叠的抗原决定基并因而竞争结合位点的情况。

由于galK柄部分所独具的特异性，所以将嵌合抗原用于免疫诊断测定时，可以迅速地、高度特异地而且灵敏地检查特定疾病，许多情况下还可能确定疾病的特定阶段。当用于HIV抗体检测时尤其如此，因为最近已有证据提示有或没有抗特异性HIV抗原的抗体与疾病进展之间有某种相关性。

本发明的免疫诊断系统是一种俘获免疫检测法，其中结合到固相载体上的抗体具有特异于嵌合抗原一部分上的某些位点的结合位点。嵌合抗原能够结合到已与固相载体结合的抗体上，同时仍留有尚未结合的所有来自特定疾病或疾病状态致病因子的抗原部分的所有抗原决定基。与固相载体结合所必需的位点不同于用来结合样品的位点。具体地说，嵌合抗原是这样构建的：即galK位点或所用的其他融合

对象应使嵌合抗原不被待检样品识别。因此，衍生于疾病或疾病状态的致病因子的嵌合抗原上的所有可利用的位点均可用于结合。

本发明的一个实施方案是使用该俘获免疫诊断系统来检测抗个别病毒蛋白的抗体和与HIV感染有关的由病毒诱导的蛋白质。在酶联免疫吸附试验(ELISA)、“夹心免疫测定法”、竞争结合法和放射免疫法(RIA)等俘获免疫检测法中，使用了galK特异性抗体(多克隆或单克隆的)来检测HIV的血清转化作用(HIV Seroconversion)。俘获检测法的设计使得病人样品最大限度地结合到与疾病的特定阶段有关的特异性嵌合HIV抗原上，从而提供了一种快速而高度灵敏的诊断试验。

在这一俘获检测法中，首先将galK特异性抗体连接到固相载体(如聚苯乙烯96孔平板或小球)上并洗去杂质。然后使纯化或半纯化形式的galK-HIV嵌合抗原与galK抗体相互作用。洗掉所有未结合的杂质。最后，使样品(如人血清或其他体液)与通过galK抗体结合到载体上的嵌合的HIV抗原反应，并洗掉所有未结合的材料。可用各种方法，如荧光分光光度法、使用酶如辣根过氧化物酶的比色法、用 ^{14}C 或 ^{125}I 标记的放射标记同位素或本领域内已知的其他检测系统，来检测与结合的嵌合HIV抗原反应的人免疫球蛋白。

上述俘获系统具有三重优点：(1)不必将galK-HIV抗原嵌合体纯化至均一，因为galK特异性抗体将特异地连接在嵌合抗原的氨基末端上，并在洗涤后仍能使之固着在固相载体上，这样即可除掉存在于嵌合蛋白制剂中的所有杂质；(2)因为嵌合蛋白通过其galK“柄”部分连在固相载体上，所以存在于分子上的所有

H I V 抗原决定基均可被病人样品识别。与传统的俘获 ELISA 法（其中用于俘获和检测的单克隆抗体都是针对目的抗原的）相比较，上述特征提高了检测法的灵敏度；（3）如果样品中的抗体特异地针对特定抗原，便将只识别嵌合抗原，从而提高了特异性。

目的 H I V 抗原可作为翻译融合体而得以表达，其中 H I V 抗原与大肠杆菌 *gal K* 基因所编码的蛋白的氨基末端区域相融合。选择大肠杆菌表达系统旨在使 *gal K* 基因在提高基因片段之表达水平中起到最大作用，特别是当未融合的基因片段表达很差或不能表达时，这一点尤为重要。但亦可使用其他表达系统，如酵母（啤酒酵母）、链霉菌、昆虫细胞（如果蝇属昆虫细胞）及高等真核细胞（哺乳动物细胞）。

可用来在大肠杆菌内表达本发明嵌合抗原的表达系统的例子有：流感病毒 NS 1（Young et al., Proc. Natl. Acad. Sci. USA 80 : 6105, 1983）、 λ 噬菌体 O 蛋白（Rosenberg et al., Meth. Enzymol. 101 : 123, 1983）以及大肠杆菌 β 半乳糖苷酶蛋白（Bartus et al., Antim. Ag. Ch. 25 : 622, 1984）。在链霉菌中表达包括使用大肠杆菌 *gal K* 基因（Brawner et al., Gene 40 : 191, 1985）、来自链霉菌 *gal* 操纵子的基因（Fornwald et al., Proc. Natl. Acad. Sci. USA 84 : 2130, 1987），以及链霉菌 β 半乳糖苷酶基因（Burnett et al., Microbiology, ed. Levin, Amer. Soc. Microbiol., Washington, D. C., 1985）。在酵母中可使用大肠杆菌 *gal K*（Rosenberg et al., Genetic

Engineering, Vol. 8, ed. Setlow, Plenum Press, 1986) 和 3-磷酸甘油酸激酶 (PGK) 基因 (Chen et al., Nucleic Acids Res. 12: 8951, 1984) 实现表达。已在哺乳动物细胞中表达了大肠杆菌 galK (Schumperli et al., Proc. Natl. Acad. Sci. USA 79: 257, 1982); 这种系统亦可用作翻译融合表达系统。

大肠杆菌是生产用于诊断目的的嵌合 HIV 抗原所选用的生物体。在大肠杆菌中确实可迅速地得以克隆和表达,而且迄今已发现在某些情况下,与其他生物体如哺乳动物系统(中国仓鼠卵细胞)或酵母相比较,在大肠杆菌中可获得更高水平的基因表达。此外,用大肠杆菌生产易于进行大规模发酵和蛋白质纯化。

使用 galK 融合系统可以产生大量特异性 HIV 抗原。如此表达的特异性嵌合 HIV 抗原可用于任何已被人们掌握的免疫诊断检测法如 Western 印迹法、点/槽印迹法、直接 ELISA 法或其他免疫检测法中,以检测 HIV 血清转化。

实施例 1 中具体描述了构建大肠杆菌 galK 翻译融合载体即质粒 pOTSKF33 的方法。可以使用实施例中所述的技术制备其他融合载体。

pOTSKF33 融合载体已被用来在大肠杆菌中将几种外来基因产物过量表达为嵌合蛋白,如人 β 球蛋白、组织血纤维蛋白溶酶原激活因子、糖皮质激素受体及金属巯基组氨酸三甲内盐 (metallothionein)。如果没有融合,这些蛋白质则很难表达。为此,该质粒可用于过量表达各种 HIV 的基因产物。

最初的实验所研究的来自 HIV 的基因有 gag、pol 和 env

基因，因为由这些基因编码的蛋白最容易被病人血清识别。首先将 HIV gag 区翻译成长约 55 千道尔顿 (KD) 的多聚蛋白质前体，然后将该前体加工成成熟的 p17、p24 和 p15 gag 结构蛋白。

据信 gag-pol 区可通过交叠的 gag 和 pol 读码之间的翻译移码而翻译成大的前体 (Ratner et al., Nature 313 : 277, 1985)。gag-pol 前体经翻译后加工，产生成熟的 gag 蛋白和包括反转录酶及核酸内切酶在内的 pol 区产物。

env 区被翻译成 gp160 前体多肽，并在两个或两个以上位点被加工。第一加工过程除去长约 30 个氨基酸的分泌信号肽；第二加工过程则产生膜外 gp120 区及 gp41 区，后者系由跨膜的疏水区 and 接在其后的亲水性胞质固着粒 (anchor) 组成。

pOTSKEF33 融合载体具有强的并可调节的 PL 启动子，它邻接 galK 基因的 56 个密码子长的部分，后者则直接连有适于基因融合 (如与 HIV 基因的目的片段融合) 的限制位点。业已发现，当插入表达质粒 (融合载体 pOTSKEF66) 中时，galK 基因的 253 个密码子长的部分亦可与目的基因融合，并导致高水平表达。

当使用 pAS1 系统在大肠杆菌内表达时，全长 galK 基因产物高水平积聚且稳定 (Rosenberg et al., 同上)。另外，发现异源基因产物与 galK 基因产物的氨基末端融合可产生稳定的融合产物，即该产物不被降解、改变或失活。可以获得高水平表达，即 HIV 衍生的抗原可占细胞总蛋白量的 10% 或更多，并可望通过大规模发酵及各种纯化过程来大规模生产这些抗原。

由大肠杆菌产生的特异性 HIV 抗原可用于刺激与 HIV 蛋白上特异位点反应的抗血清产生。因此，所构建并表达的每种嵌合抗原均

可用作产生定位抗体的因子。产生多克隆抗体的能力使它也有可能用最初由Köhler 和 Milstein(Nature 256 : 495 , 1975)所述的技术或其他细胞融合或转化技术产生单克隆抗体。如此产生的多克隆或单克隆抗体可用于俘获ELISA 检测法中(如Western印迹分析、点/槽印迹分析中所述)或用于亲和纯化。这些抗体亦可用于对病人样品中的HIV 抗原(游离抗原、游离病毒、感染的细胞)进行直接免疫检测。

实施例

下列实施例旨在进一步阐述而不是限定本发明。

实施例 1

g a l K 翻译融合载体 pOTSKF33 的构建

pOTSKF33 衍生自 p A S I 表达载体 (Rosenberg et al., 同上), 包含邻接有多连接子的大肠杆菌 g a l K 基因的一部分(56 个密码子), 其中三个翻译读码内均有适于融合的限制位点, 另外包含各阶段的终止密码子和某些附加的克隆位点。其构建可分几个步骤进行:

步骤 1: 将 p U C 1 9 的多连接子区插入 pOTS34 载体中:

用限制性核酸内切酶 B g l II 消化 p A S I (Devare et al., Cell 36 : 43 , 1984) 的衍生物 pOTS34, 用 D N A 聚合酶 I (Klenow 片段) 处理以产生平末端, 用限制性核酸内切酶 SphI 再消化。纯化 5251 碱基对 (b p) 的大 B g l II - S p h I 片段, 并连接到由 p U C 1 9 (Yanisch-Perron et al., Gene 33 : 103 , 1985) 分离纯化的 135bp PvuII - SphI 限制片段上。所得质粒称为 pOTS-UC。

步骤2：将得自pASK的部分galK表达单位插入pOTS-UC中：

用XmaI消化pOTS-UC，用Klenow 片段处理以产生平末端，用HindIII 再消化。纯化大片段并连接到由pASK分离纯化的2007bp BclI 填平的HindIII 片段上。所得质粒称为pOTSKF45。

pASK质粒载体是PAS1的衍生物，其中整个galK编码顺序被插入到ATG起始密码子之后。该galK顺序的5'端与已确定的顺序(Debouck et al., Nucleic Acids Res. 13:1841, 1985)不同：

—pASK顺序： ATG.GAT.CCG.GAA.TTC.CAA.GAA.AAA...

—已确定的顺序： ATG.AGT.CTG.AAA.GAA.AAA...

pASK顺序在ATG处含有一个BamHI 位点(GGATCC)，而已确定的顺序则没有。

步骤3：将携带翻译终止信号的连接子引入pOTSKF45中：

用XmaI和SalI消化pOTSKF45，然后用Klenow 处理产生平末端。然后经用T₄ DNA连接酶连接，在XmaI填平和SalI填平的位点之间引入5'—CTA.GGC.CTA.GTT.AAC.TAG—3'连接子，得到最终载体pOTSKF33。

实施例 2

galK-HIV蛋白融合基因的构建：

galK-p24 gag 融合的构建：

将567bp含p17 gag 后13个密码子和p24 gag 1—177号密码子的PvuII—HindIII 填入片段插入到pOTSKF33

的 *Xma*I 填入位点处，以在读码内与 *gal*K 的前 56 个密码子融合。该片段的端点为 *Pvu*II (689) 和 *Hind* (1256) (使用 Ratner 等 (Nature 313 : 277, 1985) 的编号)。使用标准克隆技术由 HTLV-III (b) 分离物 (Hahn et al., Nature 312 : 166, 1984) 的 BH10 克隆中分离该片段。*gal*K - p24 gag 融合蛋白的氨基酸顺序，使用氨基酸名称的一字母代码列示如下：

长度：253 个氨基酸残基

```

1  MDPEFQEKTO SLFANAFGYF ATHTIQAPGR VNLIGEHTDY NDGFVLPCAI
51  DYQTVIPADT GHSSQVSONY PIVQNIQGQM VHQAISPRTL NAWVKVVEEK
101 AFSPEVIPMF SALSEGATPQ DLNTMLNTVG GHQAAMQMLK ETINEEAAEW
151 DRVHPVHAGP IAPGQMREPR GSDIAGTTST LQEQIGWMTN NPPIPVGEIY
201 KRWIILGLNK IVRMYSPTSI LDIRQGPKEP FRDYVDRFYK TLRAEQAVGL
251 VN*

```

*gal*K - p55 gag 融合的构建

将含有 P17 gag (减去其前 14 个密码子)、整个 p24 gag 和 p15 gag 前 60 个密码子的 1266bp *Cla*I - *Bgl*II 填入片段插入到 pOTSKEF33 的 *Stu*I 位点处，以与 *gal*K 的前 56 个密码子产生读码内融合。该片段的 Ratner 编号为：*Cla*I (374) 和 *Bgl*II (1640)。 *gal*K - p55 gag 融合蛋白的氨基酸顺序如下：

长度：484 个氨基酸残基

```

1 MDPEFQEKTO SLFANAFGYP ATHTIQAPGR VNLIGEHTDY NDGFVLPCAI
51 DYQTVIPGRR WEKIRLRPGG KKKYKLKHIV WASRELERFA VNPGLLETSE
101 GCRQILGQLQ PSLQTGSEEL RSLYNTVATL YCVHQRIEIK DTKEALDKIE
151 EEQNKSKKKA QQAAADTGHS SQVSONYPIV QNIQGQMVHQ AISPRTLNAW
201 VKVVEEKAFS PEVIPMFSAL SEGATPQDLN TMLNTVGGHQ AAMQMLKETI
251 NEEAAEWDRV HPVHAGPIAP GQMPREPRGSD IAGTTSTLQE QIGWMTNNPP
301 IPVGEIYKRW IILGLNKIVR MYSPTSILDI RQGPKEPFRD YVDRFYKTLR
351 AEQASQEVKN WMTETLLVQN ANPDCKTILK ALGPAATLEE MMTACQGVGG
401 PGHKARVLAE AMSQVTNTAT IMMQRGNFRN QRMVKCFNC GKEGHTARNC
451 RAPRKKGCWK CGKEGHQMKD CTERQANFLG KIL*

```

gal K — RTendo 融合的构建：

将含有蛋白酶后 9 个密码子、完整反转录酶基因及核酸内切酶前半基因（即 1—141 号密码子）的 2129bp HincII — EcoRI 填入片段插入到 pOTSKEF33 的 SmaI 位点处，以与 gal K 的前 56 个密码子产生读码内融合。此片段的 Ratner 编号为：HincII (2099) 和 EcoRI (4228)。

gal K — RTendo 融合蛋白的氨基酸顺序如下：

长度：769 个氨基酸残基

```

1 MDPEFQEKTO SLFANAFGYP ATHTIQAPGR VNLIGEHTDY NDGFVLPCAI
51 DYQTVIPQI GCTLNFPISP IETVPVKLKP GMDGPKVKQW PLTEEKIKAL
101 VEICTEMEKE GKISKIGPEN PYNTPVFAIK KKDSTKWRKL VDFRELNKRT
151 QDFWEVQLGI PHPAGLKKKK SVTVLDVGDA YFSVPLDEDF RKYTAFTIPS
201 INNETPGIRY QYNVLPQGWK GSPAIFQSSM TKILEPFFKQ NPDIVIIQYM
251 DDLYVGSDLE IGQHRTKIEE LRQHLLRWGL TTPDKKHQKE PPFLWMGYEL
301 HPDKWTVQPI VLPEKDSWTV NDIQKLVGKL NWASQIYPGI KVRQLCKLLR
351 GTKALTEVIP LTEEALELEA ENREILKEPV HGVYYDPSKD LIAEIQKQGQ
401 GQWTYQIYQE PFKNLKTGKY ARMGAHTND VKQLTEAVQK ITTESIVIWG

```

```

451 KTPKFKLPIQ KETWETWWTE YWQATWIPEW EFVNTPLPLVK LWYQLEKEPI
501 VGAETFYVDG AANRETKLGK AGYVTNKGRQ KVVPLTNTTN QKTELQAIYL
551 ALQDSGLEVN IVTDSQYALG IIQAQPDKSE SELVNQIIEQ LIKKEKVYLA
601 WVPAHKGIGG NEQVDKLVSA GIRKILFLDG IDKAQDEHEK YHSNWRAMAS
651 DFNLPPVVAE EIVASCDKCQ LKGEAMHGQV DCSFGIWQLD CTHLEGKVIL
701 VAVHVASGYI EAEVIPAETG QETAYFLLKL AGRWPVKTIH TDNGSNFTSA
751 TVKAACWWAG IKQEFGLG*

```

gal K — p41 env融合的构建

将含有 p 41 env 基因 29—129 号密码子的 301 bp

RsaI — HindIII 填入片段插入到 pOTSKEF33 的 SmaI 位点处，以与 gal K 的前 56 个密码子形成读码内融合。此片段的 Ratner 编号为：RsaI (7417) 和 HindIII (7718)。

gal K — p41 env 融合蛋白的氨基酸顺序如下：

长度：160 个氨基酸残基

```

1 MDPEFQEKTO SLFANAFGYP ATHTIQAPGR VNLIGEHTDY NDGFVLPCAI
51 DYQTVIPDAR QLLSGIVQQQ NNLLRAIEAQ QHLLQLTVWG IKQLQARILA
101 VERYLKDOOL LGIWGCSGKL ICTTAVPWNA SWSNKSLEQI WNNMTWMEWD
151 REINNYTSWV

```

gal — p120 env(N) 融合的构建：

将含有大部分 p160 env 的 1790bp Asp718I — HindIII 填入片段插入到 pOTSKEF33 的 StuI 位点处，以与 gal K 的前 56 个密码子形成读码内融合。用 StuI 消化所得质粒并连接到所有三个读码中含翻译终止信号的连接子 (5' CTA . GTT . AAC . TAG 3') 上。该构建体表达 12—174 号密码子中 p12 env 基因的氨基末端一半 (N)。该片段的 Ratner 编号为：Asp718I

(5928) 和 *Stu*I(6411) 。

gal K — p120 env (N) 融合的氨基酸顺序如下:

长度: 224 个氨基酸残基

```
1 MDPEFQEKTO SLFANAFGYP ATHTIQAPGR VNLIGEHTDY NDGFVLPCAI
51 DYQTVIPGRV P120 PVWKEATTTL FCASDAKAYD TEVHNVWATH ACVPTDPNPQ
101 EVVLNVNVTEN FNMWKNDMVE QMHEDIISLW DQSLKPCVKL TPLCVSLKCT
151 DLKNDTNTNS SSGRMIMEKG EIKNCSFNIS TSIRGKVQKE YAFFYKLDII
201 PIDNDTTSYT LTSCNTSVIT P120 QAS*
```

gal K — p120 env (C) 融合的构建

将含有 p120 env 的末端部分和 p41 env 的起始部分的

*Pvu*II — *Hind*III 填入片段插入到 pOTSKEF33 的 *Stu*I 位点处, 以与 *gal* K 的前 56 个密码子进行读码内融合。用 *Bgl*II 消化所得质粒 (部分限制), 并连接到所有三个读码内含有翻译终止信号的连接子 (5' CTA . GTT . AAC . TAG 3') 上。该质粒表达来自 258 — 437 号密码子的 p120 env 的羧基末端 (C) 一半。该片段的 Ratner 编号为: *Pvu*II (6662) 和 *Bgl*II (7198) 。

gal K — p120 env (C) 融合的氨基酸顺序如下:

长度: 243 个氨基酸残基

```
1 MDPEFQEKTO SLFANAFGYP ATHTIQAPGR VNLIGEHTDY NDGFVLPCAI
51 DYQTVIPGRV P120 NQSVEINCTR PNNNTRKSIR IQRGPGRAV TIGKIGNMRQ
101 AHCNISRAKW NNTLKQIDSK LREQFGNNKT IIFKQSSGGD PEIVTHSFNC
151 GGEFFYCNST QLFNSTWFNS TWSTKGSNNT EGSDTITLPC RIKQIINMWQ
201 EVGKAMYAPP ISGQIRCSSN ITGLLLTRDG GNSNNESEIL P120 VN*
```

gal K — HIV 嵌合抗原的表达

将含有 *gal* K — HIV 抗原融合基因的各种 pOTSKEF33 衍生物引入大肠杆菌 AR120 菌株内, 并按 Mott 等 (Proc. Natl.

Acad. Sci. USA 82 : 88 , 1985) 所述方法用萘啶酮酸处理进行诱导。亦可通过温度转换 ($32^{\circ}\text{C} \rightarrow 42^{\circ}\text{C}$) , 在那些对可热诱导的缺陷性原噬菌体具有溶源性的大肠杆菌菌株中, 诱导这些表达载体 (Young 等, 同上)。

实施例 3

gal K - HIV 嵌合抗原的纯化

p55 gag

将 100 g gal K - p55 gag 诱导的大肠杆菌细胞沉降物重新悬浮于 500 ml 含 0.5 mg/ml 溶菌酶和各 0.1 mg/ml DNA 酶和 RNA 酶的 p55 溶解缓冲液 ($50 \text{ mM Tris} - \text{HCl}$ ($\text{pH} 8.5$)、 5 mM EDTA 、 1 mM DTT 、 $5\% (\text{V/V})$ 甘油及 0.1% 硫二甘醇) 中。持续搅拌下将混合物于室温保温 15 分钟。以下各步骤均于 4°C 下进行。

使混合物通过压力为 4500 psi 的 Menton - Gaulin 压力匀浆器 2 - 3 次制成匀浆。以 $10,000 \times g$ 将匀浆离心 30 分钟。所得沉淀重新悬浮于 500 ml 含 0.2% 脱氧胆酸钠 (NaDOC) 和 $1\% \text{ Triton X} - 100 (\text{V/V})$ 的 p55 溶解缓冲液中。将该混合物搅拌 15 分钟后以 $10,000 \times g$ 离心 30 分钟。

将所得沉淀重新悬浮于 500 ml 含 $1\% \text{ Triton X} - 100 (\text{V/V})$ 和 0.5 M KCl 的 p55 溶解缓冲液中, 搅拌 15 分钟后以 $10,000 \times g$ 离心 30 分钟。

将所得沉淀重新悬浮于 250 ml 含 8 M 尿素的 p55 溶解缓冲液中并搅拌 30 分钟。经最后离心 ($10,000 \times g$, 30 分钟) 即得到留在上清液中的溶解的产物。

在 Superose 12 (Pharmacia Fine Chemicals , Uppsala , Sweden) 柱上, 于 2 M 氯化钕存在下层析, 得到经凝胶电泳及考马斯亮蓝染色测定其纯度大于 90 % 的制品。均一物质的产率为大约 150 mg / 100 g 冷冻的大肠杆菌沉淀。当改用磷酸盐缓冲盐水 (PBS) 溶解时, p55 gag 融合蛋白仍然是可溶的 (约 1 - 3 mg / ml)。

p41 env

按上述提纯 p55 gag 融合蛋白的方法由 gal K - p41 env 诱导的细胞中纯化 p41 env 融合蛋白, 不同的是在第一次离心时以低速 (1,000×g, 20 分钟) 离心混合物, 小心除去上清并以 27,000×g 再次离心 20 分钟。

在 2 M 氯化钕存在下于 Superose - 12 柱上层析, 结果表明, env 融合蛋白在这些层析条件下表现有高分子量聚集物的行为。在 8 M 尿素存在下于 Mono S (Pharmacia Fine Chemicals , Uppsala , Sweden) 柱上层析, 得到纯度约为 90 % 的未结合部分。产率可达到约 300 mg / 100 g 冷冻的大肠杆菌沉淀。当溶剂直接换成 PBS 时, p41 env 融合蛋白不能溶解。

p24 gag

上述两种提取 / 纯化方法, 都得到纯度约为 70 % 的 p24 gag 融合蛋白。p24 gag 融合蛋白的产率约为 300 mg / 100 g 冷冻的、经诱导的大肠杆菌沉淀。

因 p120 env (N) 和 (C)、以及 RTendo 融合蛋白也见于粗制细菌提取物的不溶部分中, 所以可望使用大致相同的方法纯化之。

实施例 4

g a l K 翻译融合载体 pOTSKF66 的构建

pOTSKF66 衍生于 p A S 1 表达载体 (Rosenberg 等, 同上), 并包含邻接有多连接子的一大部分大肠杆菌 g a l K 基因 (253 个密码子), 在其所有三个翻译读码中均有适于融合的限制位点、各期的终止密码子以及某些附加克隆位点。

将 p A S K 的一个 761bp BamHI — FspI 片段插入 pOTSKF33 的 BamHI 位点与 X m a I 填充位点之间, 以构建该质粒。将 pOTSKF66 衍生物引入同一大肠杆菌菌株内, 并按诱导 pOTSKF33 衍生物的同样方法 (用萘啶酮酸、热处理) 诱导之。

pOTSKF66 中与 g a l K 的 253 个密码子于读码内融合的基因片段可获得高水平表达, 与用 pOTSKF33 融合载体观察到的表达水平相同。

实施例 5

酵母 g a l K 翻译融合载体 p C D 1 9 2 — S K F 3 3 的构建

p C D 1 9 2:

p C D 1 9 2 衍生于 pYSK105 (Gorman et al., Gene 48 : 13, 1986)。它可按下述两个步骤构建:

步骤 1: 用限制性核酸内切酶 N r u I 和 P v u II 消化 pYSK105, 并用 T4 D N A 连接酶使之自身连接, 以缺失 1094bp NruI — PvuII 片段。这一处理去掉了载体的唯一 P v u II 位点, 得到 pYSK105 — LPP。

步骤 2: 切下携带 g a l K 编码顺序的 pYSK105 的 1.2 千碱基 (K b) EcoRI 片段, 并用 D N A 聚合酶 I (Klenow 片段) 处理以填平 EcoRI 末端。将平末端连接到具有顺序

5'CCAGCTGG 3'的PvuII连接子上。所得质粒称为pCD192，它将使插入唯一PvuII位点的任何基因（含有ATG起始密码子）得到CUP1定向表达。

pCD192-SKF33:

如所述，pCD192-SKF33衍生自pCD192酵母表达载体。pCD192-SKF33是一个以pBR322为基础的穿梭载体，包含部分2微米复制起点、trp1选择性标记、可用铜调节的CUP1启动子，其后是大肠杆菌galK基因的一部分（56个密码子），后者邻接有一个多连接子并在三个翻译读码的任一个中均有适于融合的限制位点、各期的终止密码子以及某些附加的克隆位点。按下述两个步骤构建该载体：

步骤1：pCD192的BamHI、SmaI/XmaI、SphI和SalI位点的缺失

用限制性核酸内切酶BamHI和SalI消化pCD192，用DNA聚合酶I（Klenow片段）处理，产生平末端，并用T4DNA连接酶使之自身连接。这一处理缺失了位于pCD192中CUP1启动子上游的286bp片段，从而除去了BamHI、SmaI/XmaI、SphI和SalI位点，得到质粒pCD192-DBS。

步骤2：将pOTSKF33的部分galK表达单位插入pCD192-DBS中：

用限制性核酸内切酶PvuII消化pCD192-DBS，在紧接CUP1启动子下游处切断之。将其连接到由pOTSKF33分离并纯化的313bpNdeI-XhoI填入片段上。该片段包含ATG起始密码子和大肠杆菌galK的56个氨基端密码子，后者邻接适于在所有三个翻译读码内进行融合

的限制位点及各期终止密码子。

实验结果表明，pCD192 - SKF33 可用来提高那些在未融合时表达很差或不能表达的基因片段的表达水平，正如对大肠杆菌中的同一翻译融合系统（pOTSKF33，实施例1）所表明的那样。

实施例 6

galK - β 球蛋白融合基因在酵母内的构建和表达

由pJK06 质粒（O'Donnell，ph. D. thesis，Temple University，1987）中分离含有接在因子Xa蛋白酶识别顺序（Ile - Glu - Gly - Arg）后面的人 β 球蛋白链全长编码顺序的1381bp NcoI - XbaI 填入片段，并插入到pCD192 - SKF33 的SmaI位点处，使之与galK的前56个密码子产生读码内融合。

用氯化锂（LiCl）方法（Ito et al.，Bacteriol. 153 : 163，1983）将这个含galK - β 球蛋白融合的pCD192 - SKF33 衍生物引入啤酒酵母BR10菌株。按下述方法加入CuSO₄ 以诱导CUP1启动子：在含葡萄糖但不含色氨酸的基本培养基中，于30℃下培养含pCD192 - SKF33 衍生物（携带galK - β 球蛋白融合基因）的BR10细胞至对数期。然后加入CuSO₄ 至终浓度为0.1 mM。于30℃下继续通气1小时后，离心收集细胞并用水洗一次。用玻璃珠破碎细胞以制备提取物，然后以SDS - PAGE及免疫印迹法（用 β 球蛋白特异性抗体）进行分析。结果表明，与galK的56个密码子融合后， β 球蛋白基因产物的表达水平比未融合时高得多。

上面的描述和实施例充分公开了本发明及其优选实施方案。对于

分子遗传学及相关领域内的普通技术人员来说，如在下述权利要求的范围内对所述方法作某些改变将是显而易见的。