

[19]中华人民共和国专利局

[51]Int.Cl⁶

C12P 21/02

A61K 38/00



[12] 发明专利说明书

[21] ZL 专利号 88103482.7

[45]授权公告日 1997 年 1 月 8 日

[11] 授权公告号 CN 1033760C

[22]申请日 88.6.11 [24]颁证日 96.10.26

[21]申请号 88103482.7

[30]优先权

[32]87.6.12 [33]EP[31]87201121.8

[73]专利权人 伊姆诺股份公司

地址 奥地利维也纳

[72]发明人 瓦·奥耶·阿尔伯特·约翰森·约瑟夫

帕尼科克·汉森

弗比特·马蒂尼·菲利浦斯

瓦利·罗伯特·威廉姆

[74]专利代理机构 中国国际贸易促进委员会专利商标

事务所

代理人 辛敏忠

审查员 马昭若

权利要求书 0.5 页 说明书 26.5 页 附图页数 3 页

[54]发明名称 产生具有第八因子活性的第八因子缺失突变体的方法

[57]摘要

本发明提供通过 908—949 氨基酸缺失制备突变第Ⅷ因子 cDNA 的方法。cDNA 在转化的 COS 细胞中表达,得到在赖-338 到谷酰-1638 间氨基酸有相应缺失的第Ⅷ因子蛋白。所得蛋白似乎能纠正血友病 A 病人的凝血缺陷,该蛋白具有完整的链,表现出相对高的表达水平。

权利要求书

1. 产生有第Ⅷ因子活性的第Ⅷ因子缺失突变体的方法，该方法包括在一营养培养基中培养含有表达载体的哺乳动物宿主细胞，所述表达载体在其转录方向上包含：

A) 一转录调节区，

B) 在哺乳动物宿主细胞中具有功能的翻译起始区，

C) 编码一蛋白质的 DNA 序列，该蛋白质具有第Ⅷ因子的生物活性且除以下缺失外具有对应于第Ⅷ因子的氨基酸序列，

i) Ser-741 至 Ser1637 和 Gln-1638 至 Ile-1668 的氨基酸序列（缺失 741—1668，△928；或

ii) Ser-741 至 Ser1637 和 Lys-713 至 Arg-740 的氨基酸序列（缺失 713—1637，△925；或

iii) Ser-741 至 Ser1637 和 Lys-713 至 Arg-740 的氨基酸，并在缺失位点插入氨基酸序列 Pro-Arg-Val-Ala（缺失 713—1637，插入 PRVA，△925；或

iv) Ser-741 至 Ser1637 和 Ser-1637 的氨基酸序列和 Gln-1638 至 Ile-1668 的部分氨基酸序列，所述部分为 Gln-1638 至 Arg-1648（缺失 741—1648，△908；或

v) Ser-741 至 Ser-1637、Lys-713 至 Arg-740 和 Gln-1638 至 Ile-1668 的氨基酸序列（缺失 713—1668，△956）；或

vi) Ser-741 至 Ser-1637、Lys-713 至 Arg-740 和 Gln-1638 至 Ile-1668 的氨基酸序列，在缺失位点插入氨基酸序列 Pro-Arg-Val-Ala（缺失 713—1668，插入 PRVA，△956）；

D) 在所述宿主细胞中具有功能的翻译和转录终止区，其中所述 DNA 序列的表达是由所述起始和终止区调节的，以及分离所述蛋白质。

本发明涉及到一些具有第 8 因子活性的新的蛋白质以及使用遗传工程细胞系或微生物来制备它们的方法。

第 8 因子是一种医学上重要的蛋白质，因为它是血液凝固机制中的一种重要成分。它在健康人血浆中的浓度大约为 100ng/ml。第 8 因子缺乏导致称为血友病 A 的一种出血性疾病，它主要见于男性，因为第 8 因子基因位于 X-染色体上。本病是一种重要的遗传性出血性疾病，男性人群发病率约 0.01%。

血友病 A 可以通过输入自健康人获得的含第 8 因子的血浆来治疗。然而该疗法有几种缺点。首先，第 8 因子的来源有限且价格昂贵，因为它在供者血中的浓度很低。其次，现行的血浆分级分离方法产率较低，第 8 因子的实际含量不足。第三，每一例血友病 A 患者接受了来自许多供血者的血液制品，从而获得传染病的危险性升高，这些传染病包括例如由存在于供血中的非甲、非乙型肝炎病毒，乙型肝炎病毒，艾滋病（AIDS）病毒引起的各种疾病。另一问题是所谓抑制因子的出现，它是接受外源性，同种基因性第 8 因子制品治疗的血友病患者所产生的抗第 8 因子抗体。

这些问题已促使人们去认真研究用重组 DNA 技术的方法来获得第 8 因子，以代替自供者血液中的分离方法。几组学者已报道了由人类 mRNA 所获的第 8 因子 cDNA 分子克隆和在哺乳类，酵母菌，或细菌细胞中生产具有第 8 因子活性的蛋白质（WO85/01961，EP0160457，EP150735，EP0253455）。

欧洲专利申请 0253455 介绍了一种使用转化微生物生产具有第 8 因子活性蛋白质的方法。该种蛋白适于制备具有第 8 因子活性的药学制品并可激发能识别具有第 8 因子活性蛋白的抗体。

另一组研究者的目标是发现具有第 8 因子活性并能纠正血友病 A 患者凝血缺陷的其它种蛋白质。欧洲专利申请 EP150735，EP123945，和 Brinkhous 等（1985）指出，第 8 因子活性不仅取决于第 8 因子蛋白本身，而且还决定于它的某些蛋白水解产物。蛋白水解酶作为凝血链式反应的一环见于血液之中，并部分降解第 8 因子为分子量 92KDa 和 80KDa 的两个亚单位多肽。据认为，当这两个亚单位的复合体或其一部分在血中出现时，第 8 因子活性增强（Fay 等，1986，Eaton 等 1986a）。

与其它血浆蛋白不同，当以已知的重组 DNA

方法制备时,第8因子是一种不尽理想的低产蛋白质。甚至该蛋白并不以完整肽链的形式出现,因此制品很难进一步分离和纯化,费用也高。

Eaton等(1986b)发现,从第8因子蛋白的中心区去除第797-1562之间的766个氨基酸之后,余下的蛋白仍保有第8因子活性。而且,用含有编码该缩短的蛋白质的DNA的载体转化的哺乳类细胞比带有编码全部长度蛋白的密码的易感细胞表现出更高的产率。

PCT专利申请WO86/06101透露,既使将第8因子蛋白删除到仅离中心区880个氨基酸仍显示第8因子活性。最多的缺失延伸达Thr-760到Asn-1639(按表IV计数)。已提及的从Thr-760到Lys-1674的更大的缺失显然对其并无影响。这些蛋白可以用含编码这些蛋白质的cDNA的一个载体转化得到的哺乳类细胞系体外培养来制备。对所有天然的或制备的第8因子蛋白来讲,缺失绝不能延伸到中心区边界以外,因为该分子的余下部分是维持凝血活性的基础。

第8因子是纠正象血友病A那样的凝血缺陷的蛋白质,其特性即第8因子活性。第8因子以与一种蛋白质,Von Willebrand因子(VWF),结合状态循环于血液中,该蛋白防止第8因子过早衰变。

第8因子由表IV所述的氨基酸序列所定性。序列的编号开始于Ala-1,它为含19个氨基酸的信号序列后的第一氨基酸,末端为Tyr-2332。该编号用于本专利说明书。

“具有第8因子活性的蛋白质”“第8因子蛋白”或“类第8因子化合物”包括在天然第8因子之内的各种大小的蛋白质,它们均具有经确定的氨基酸序列的核心部分或与第8因子的一部分氨基酸序列是同质的,当输给血友病A患者时,它们具有第8因子活性以纠正凝血缺陷。该凝血缺陷的纠正可能受到已被破坏的凝血机能中直接干扰的影响或上述抑制因子患者体内的抗第8因子蛋白抗体中和作用的影响从而随后输入的凝血第8因子蛋白能显示其全部活性。

作为激活蛋白对全长第8因子蛋白作用的结果,形成了亚单位蛋白质,它们的分子量分别为92KDa和80KDa。92KDa-亚单位自全长第8因子蛋白的氨基末端始而向下延伸,80KDa-亚单位

自羧基末端始而向上延伸。

为本发明的需要,92KDa-多肽规定为全长第8因子的自Ala-1至Arg-740的氨基酸序列,8KDa-多肽定为自Gln-1649至TYR-2332氨基酸序列,这些序列是经蛋白水解位点标明的。第8因子蛋白的其余部分为中心区,它与所谓的B-区间非常适应,B-区间的边界氨基酸是Ser-741和Arg-1648。

“具有第8因子活性蛋白质”的概念包括92KDa和80KDa多肽及其一部分,它们可能本身不具有活性,但当一起输入或与另一种适当的蛋白质结合时,显示出第8因子的活性,该后一种蛋白可能存在于血友病A患者体内。

“具有第8因子活性的蛋白质”的概念还包括含有既与80KDa又与92KDa-蛋白质或它们的片段相对应的氨基酸序列的蛋白质。据认为它们来自于缺失了在上述92KDa和80KDa亚单位之间的中心区的第8因子蛋白或缺失了92KDa-和80KDa-亚单位界限以外部分。它们一般被称为缺失突变蛋白。

该概念包括在人体外用如重组DNA技术生产的具第8因子活性的蛋白质。它包括在体内显示出天然的第8因子生物学特性以及具有其主要部分已确定或与第8因子的一部分氨基酸序列同质的任何蛋白质。

由遗传工程得到的“具有第8因子活性的蛋白”可能与天然存在的产品一致,它还可以是象80KDa-或92-KDa片段那样的一个片段或是有同样活性但勿需具备同样的氨基酸序列或分子量的一个片段。

“具有第8因子活性的药物制品”意指输入哺乳类体内后能纠正与第8因子缺乏有关的凝血缺陷的任何制品。

为方便起见,名词“蛋白质”常包括蛋白质和更小的多肽链。至于“来自另一序列的一个序列”指的是与另一序列一致或是天然形成、经遗传工程所得变种、或化学合成的一个序列,该序列维持或增强所需功能。

“信号序列”指的是蛋白质氨基端的一部分,在细胞膜运输时往往由成熟蛋白上被清除。

所用缩写:Ug=微克,Kb千碱基对,bp=碱基对,KDa=千道尔顿

已发现, 能从全长第 8 因子蛋白制备它的缺失突变。虽然它具有目前普遍认为有损于其生物活性的超长缺失, 但表现出第 8 因子活性。

这种新型蛋白质显示出未曾预料到的令人满意的特性, 能使它适于有规律地生产。它们在遗传工程的宿主细胞中被生产, 该细胞作为含有编码全长第 8 因子或已知缺失突变体的 DNA 的易感态细胞, 大量分泌该蛋白。该突变体系已在 PCT 专利申请 WO86/06101 所述。

该新蛋白可以用 N-L-C 式表示。N 代表从第 8 因子蛋白氨基端伸展的氨基酸序列, C 代表由羧基端伸展的序列。中间连接体 L 代表一个连接肽键或一个 1-20 个氨基酸的连接序列。

按本发明的蛋白的特征在于 N 区本质上第 8 因子丙-1 到脯-740 氨基酸序列的复制品或在该序列羧基末端每次截掉一个氨基酸直到最短的丙-1-到天冬-712 序列的一个较短序列复制品。C 区本质上是第 8 因子谷酰 1638 到酪-2332 氨基酸序列的复制品或在该序列氨基末端每次截掉一个氨基酸直到最短的丝-1690 到酪-2332C 序列的一个较短序列的复制品。第 8 因子氨基酸顺序见表 IV。

优先的蛋白包括序列 3, 3b, 7, 8 和 9 (如表 1 所示)。序列 1 涉及全部长度的第 8 因子, 序列 2 系缺失 695 个氨基酸活化的第 8 因子蛋白。

本发明还包括两种分离蛋白的混合物, 其一以序列 N 为特征, 另一以序列 C 为特征, 序列 N 和 C 的含义如前文所述。

本发明还提供了由 PSV2 衍生的通用克隆载体, 它被重组到一种表达载体之中, 该表达载体含有至少一种编码上述蛋白质之一的结构基因 CNA。它还能在转化的宿主细胞内表达至少一种上述可分离形式的蛋白质。当使用与编码信号序列的 DNA 结合的载体时, 该蛋白可以被分泌于介质之中。需被表达的 DNA 可被整合到宿主染色体上, 也可使用一个随体复制子而使其在染色体外保存和复制。

按照本发明的表达载体可以特别包括适于在哺乳类细胞中以 5'-3' 的转录方向被表达的下列 DNA 序列: 一个转录和翻译起始调节区, 具有信号序列的为第 8 因子多肽编码的开放读码框架, 和转录及翻译终止调节区。

对 COS 细胞较理想是用可被哺乳类细胞转录

系统识别的 SV40 早期转录单位的转录启动子 (SVEP)。最好是, 将启动子 SVep 与 Rous 肉瘤病毒长末端重复启动子 (RSV-LTR) 按一前一后串接方式结合, 二者按正确方向表达。见图 1。

然而, 在另一些特殊的细胞系中, 更适于使用其它启动子, 如脊椎动物、哺乳类、病毒的启动子, 包括 SV40 后期启动子, 腺病毒启动子, 金属硫蛋白启动子 (metallothionin promoter), β -actin 病毒, 或巨细胞病毒启动子。

成帽位点定位于 cDNA 插入段之前, 并确定了来自表达载体上的转录单位的 mRNA 的开端。

为了适宜的基因表达, 已发现在翻译起始密码 ATG 周围的核苷酸序列在动物细胞中是重要的。例如, KOZaK (1983) 已深入研究了这些区域对于在 COS-细胞中胰岛素的表达的影响。因此, 修饰起始密码 ATG 周围的核苷酸序列可能是重要的。这一工作可经点突变或将外源基因融合于一个高效表达基因的起始区来完成。

象许多可从细胞中正常地分泌到周围培养基中的蛋白质一样, 成熟的第 8 因子蛋白也以信号肽为先导, 一般来说, 普遍认为信号肽具有能使蛋白穿过细胞膜的特性。对于第 8 因子多肽链的分泌, 宁可使用天然形成的第 8 因子先导顺序, 但其它可能是衍生于天然第 8 因子信号序列的信号序列也同样适用。

信号序列之后接着含 cDNA 插入段的片段, 该插入段编码既可以是全长蛋白, 也可以是按本发明提供的一种新的缺失突变蛋白的成熟第 8 因子蛋白质。

能以几种方式完成对结构基因有目的的删除, 如对全长第 8 因子 cDNA 的切割、再修饰和连接, 纯化片段, 或直接点突变, 特别是如 Kramer 等 (1984) (General Methods section, z) 所介绍的圈出突变 (loop-out)。

终止区可来源于获得起始区的同一基因的 3' 末端区, 也可来自于另一不同的基因。迄今已知和已发现的许多终止区均令同一或不同属和种的宿主细胞满意地接受。终止区可以包括将哺乳类细胞 mRNA 适当处理所得序列, 即一个小的内含子和多聚腺苷信号。

在组建表达盒时, 通常将 DNA 的各种片段克隆于一个适合的克隆载体之中, 该载体考虑到了

DNA 的扩展和修饰, 或对顺序的连接和移除等操作。正常情况下, 该载体应能够在大肠杆菌中至少以相当高的拷贝数来复制。许多载体很易于供克隆使用, 包括 PBR322, PML2, PUC7-PUC19, PSP64, PSP65, PSP18, PSP19, PTZ18, PTZ18R, PTZ19, PTZ19R, 等载体。

克隆载体以在大肠杆菌中带有至少一个在功能上是有效的复制起点为特征。该克隆载体将有至少一个特定的限制酶切点, 而通常是很多特定的限制酶切点, 并可包括多种限制点, 特别是两个同样酶的限制点供置换。此外, 克隆载体将有一个以上的标记以供选择转化体时用。这些标记一般表现为抗细胞毒试剂, 如抗菌素, 重金属和毒素等, 也可是对营养缺陷型宿主的补充, 或对噬菌体具有免疫性, 载体或盒通过适当的限制酶作用, 适宜的末端修饰, 反复处理, 或填平突出部分等方法来得到平纯末端。再经过加入连接体, 或加尾等方法。为连接和载体与表达盒或其它成分的结合提供互补末端。

某些情况下, 需使用一个穿梭载体, 它能在需要不同的复制系统的不同宿主体内复制。本发明更愿意使用 Simian, 病毒 (SV40) 复制起点, 它使质粒在 COS-1-细胞内增殖 (Gluzman, 1981)。

牛乳头瘤病毒 (BPV-1) 用于表达载体在 C127 细胞中的随体维持和增殖 (见 Howley 等, 1983)。

为微生物生产选用的已建细胞系或宿主微生物是优选的宿主系统。来源于培养的原始组织的正常初级细胞或细胞系适于选用。宿主细胞最好是哺乳类细胞系, 如猴 COS-细胞, 仓鼠 CHO 细胞或鼠 C127-细胞, 但某些微生物细胞, 如酵母菌, 最好是克鲁维酵母, 或细菌, 最好是枯草杆菌也适用。

为使这些细胞系或带有第 8 因子编码序列的表达载体的宿主稳定地转化, 以及随后能使插入宿主染色体中的载体扩增, CHO (中国仓鼠卵巢) 细胞适于选用。转化必须包括与一种选择标记基因, 如 dhfr (双氢叶酸还原酶) 基因或 G418 (neo-) 抗药基因等, 连在一起的表达载体对宿主细胞的转染, 以及它在染色体上的整合和对稳定转化体的选择。获得稳定的转化体之后, 用标准方法选择第 8 因子活性和第 8 因子抗原性。

使宿主细胞系稳定转化的另一方法使用了含有 BPV-1-顺序的表达载体。C127 细胞是达此目的理想选择。转化包括表达载体对宿主细胞的转染。用中心点 (foci) 的形成来识别转化体。用标准方法确定稳定的转化体和选择第 8 因子活性。反之, 当表达实现于 PSV-2 来源的表达载体向 COS-1 细胞的过渡转化系统时, 没有选择步骤。

按照本发明更进一步的内容, 提供了按本发明使用培养的转化哺乳类细胞或微生物宿主生产蛋白质的方法, 该微生物宿主按照本发明存在于营养培养基中而产生含该蛋白的培养物, 由此分离出蛋白质。

本发明更进一步的优点在于所表达的蛋白能从完整的细胞膜或宿主细胞壁分泌到培养基之中。

使用常规技术从培养基中分离和纯化表达蛋白, 如使用固相抗体的亲和层析, 氨基己基琼脂糖层析, 或混合电解质法。

本发明的主要优点是所需蛋白以未断裂的糖基化单链分子形式分泌。经凝胶电泳分离后, 它们表现为预计大小的完整单链。没发现 80KDa 多肽的降解链, 见表 3 和图 3。该现象在制备缺失后的成熟蛋白中是新出现的, 并对第 8 因子蛋白的成功分离与纯化至关重要。

本发明进一步证实, 当与缺失的蛋白比较, 该蛋白质溶液显示出未曾估计到的很高的第 8 因子活性 (见表 2)。还发现较大的缺失导致较高的蛋白产量。随之获得具有较高第 8 因子活性的溶液。从对抗原决定簇试验的评价 (该试验是用以测定第 8 因子蛋白量) 中可以看出培养基中第 8 因子活性的增加与第 8 因子蛋白量的增加呈正比。结论是新的第 8 因子蛋白有与全部长度的第 8 因子同样的单位蛋白重量活性 (比活性)。

按照本发明, 表达蛋白适于作为药物制品, 当将其输给哺乳类, 包括患有第 8 因子缺乏症的病人和作为对含该因子抑制体病人的治疗时, 它可对抗血友病 A 的凝血缺陷。该蛋白在生产工艺中通常要与非肠道吸收的赋形剂和其它适宜的赋形剂相混合。

为了适于非肠道吸收, 本发明的药物制品可以常规包括该蛋白的灭菌冷冻干燥制品, 使用时加入与受者血液等渗的无菌溶剂而制成溶液。该制剂可以单位容量或大剂量分装, 如密封安瓶或小瓶,

它们的用法与已知的人第 8 因子蛋白类似, 为增效可加入适宜的佐剂。

按本发明得到的具有第 8 因子活性的蛋白质, 比带有相应较少抗原决定簇的全长度天然第 8 因子蛋白更小, 所以预期抗原性更弱, 从而受者更少诱发抑制体。它们也更少被存在于血友病人体内的抑制因子识别。

一般方法:

1. 克隆技术:

常规克隆技术参见 Maniatis 等人的手册 (1982)

限制酶按制造者的推荐使用, 它们可分别从新英格兰生物学试验室 (Biolabs), 贝塞斯达研究实验室 (BRL), 或贝林格, 曼海姆 (Boehringer) 实验室获取。一般来说处理 $1\mu\text{gDNA}$ 需要 1-5 单位的酶。

连接步骤, DNA 片段混于琼脂糖凝胶, 电泳抽提, 过 NACS 柱 (BRL; 见制造者的介绍) 纯化, 乙醇沉淀, 于含 50mM 和 Tris-HCl 缓冲液 pH7.4; 10mM 氯化镁, 10mM 二硫苏糖醇, 1mMATP 和 $T_4\text{DNA}$ 连接酶 (Biolabs) 中温育。

填充 5'末端突出部限制性位点使用 DNA 聚合酶 1 的大的 Klenow 片段 (Biolabs)。在含 10mMTris-HCl 缓冲液 PH7.7; 50mMNaCl, 10mM 二硫苏糖醇和 20mM 未标记核苷酸中温育 30°C , 15 分钟。

3'末端突出部限制性片段末端的降解如 Maniatis 等 (1982) 所述, 用 $T_4\text{DNA}$ 聚合酶来完成。

大肠杆菌的转化可以按经典的氯化钙法 (Maniatis 等, 1982) 或 Hanahan (1983) 法进行。转化体的选择和所需质粒的确定按 Maniatis 等 (1982) 方法实行。

从限制性片段的连接体或自制寡核苷酸连接顺序得到的全部组建体均经序列分析检查, 以便证实读码结构的完整性和在第 8 因子编码序列或表达载体的必须的调节序列中已排斥在外的核苷酸取代结构。序列分析可以用 Maxam 和 Gilbert (1980) 技术或 Sanger 等人 (1977) 的双脱氧核苷酸法进行。

寡核苷酸的人工合成可在一种应用生物系统 308B 型 DNA 人工合成器内完成。该寡核苷酸经

乙醇沉淀法纯化。

2. 圈出突变:

由寡核苷酸指引的圈出突变按双缺口技术 (Kramer 等, 1984) 进行。

M13mp8 和 M13mp9 基因组在主要的 M13 基因上包括两个 amber 突变, 并仅能在大肠杆菌抑制菌株, 如 JM101 (Yanisch-Perron 等, 1985) 中繁殖。取自不含 amber 突变的 M13mp18 或 M13mp19 的重复形式 DNA 被分离和用 EcoRI 和 HindIII 限制酶作用以去除多聚连接体区。用琼脂糖凝胶电泳分离出近 7.2Kb 的限制性 M13mp18 片段。纯化后的 M13mp18 7.2Kb 片段加入含靶片段的单股 M13mp8 (或 M13mp9) DNA 和适量的突变诱导剂 (1.25nmol/ml)。混合液加热至 90°C , 10 分钟, 随后在 55°C 下退火 30 分钟。所获双缺口通过用 Klenow-片段 DNA 聚合酶加三磷酸脱氧核苷酸的温育而被填充, 再用 $T_4\text{DNA}$ 连接酶连接。所获闭环以股 DNA 通常转染非抑制性大肠杆菌菌株 HB2154 (mutI+) 及其 HB2151 (Carter 1985)。由于噬菌体基因中有琥珀突变使得含靶片段的 M13mp18 (+) DNA 链不能增殖。反之, 含突变片段的 M13mp8 (-) 互补 DNA 链能够对噬菌体繁殖起到模板作用, 所致噬斑应该全部或至少富含突变片段。取自认定为突变体的单股 DNA 被分离, 再用适当的引物将其顺接于突变下游的 50—100 个碱基对区。

3. 转染, 过渡表达, 代谢标记和凝胶电泳

以 COS-1 细胞和 PSV2 衍生表达载体为基础的一个类第 8 因子组过渡表达系统分泌于转染细胞培养基之中。按 Lopata 等 (1984) 和 Lulhman 与 Magnusson (1983) 所述, 转染原本使用载体 DNA 的 DEAE-右旋糖 2 小时内沉淀物, 随后 2 小时内经氯喹休克以加强转化。用于 COS-1 细胞生长和维持的培养基是添加了热灭活 10% (v/v) 小牛血清的 Iscove 氏 DMEM 培养基。转染后将培养基吸水保鲜 48 小时, 再过 48 小时后进行收获。

转染细胞中的蛋白质代谢标记过程使用无血清的 RPMI 培养基完成。转染后两天用 $50\mu\text{Ci/ml}$ 的 L- ^{35}S -蛋氨酸 (Amersham, 1985; $370\text{MBq/330}\mu\text{l}$ 特异活性: 1000Ci/mmol 以上) 温育 1 小时, 接着在收获控制条件培养物之前

再用 1mM L-蛋氨酸温育过夜。为了抑制蛋白质降解衰变,通常是在收获后将蛋白酶抑制剂,如 PMSF 加入到控制条件培养基中。当需抑制蛋白糖基化时,向控制条件培养基中加入终浓度 0.001mg/ml 的 tunicamycin。

4. 免疫沉淀分析

为确定制造出的第 8 因子蛋白质分子大小,使用由血浆来源第 8 因子诱发得到的单克隆和多克隆血清抗体,先结合于蛋白 A-琼脂糖(每 10mg 蛋白 A-sepharose 对 10 μ l 血清)。接着与从转化的 COS-1 细胞分泌的类第 8 因子样组分代谢标记物共同温育。按 Laemmli (1972) 法在 5% 聚丙烯酰胺凝胶电泳上用 10% β -巯基乙醇还原后分离沉淀的类第 8 因子分子。

5. 第 8 因子的生物学活性

按制造者 Kabi-vitrum 所述进行产色素活性或 Kabi coatest 试验,除了全部容量分为四份和每次检查维持培养基 25 μ l 以外。第 8 因子类蛋白在 15 分钟内被活化,接着加入产色底物(S2222)。

标准凝集或血凝试验(所谓 activated partial thromboplast intime 试验)按 Veltkamp 等(1968)所述使用血友病人血浆进行。检查前维持培养基被柠檬酸化抗凝处理。

按标准的贝塞斯达方法 Kasper c. k 等, (1975) 测定抑制作用,所用免疫球蛋白经离子交换和蛋白 A-琼脂糖层析纯化获得。

第 8 因子生物学活性试验的标准品是柠檬酸化血浆库(经浓度 0.5mM),假定它每毫升含 1 个单位的第 8 因子活性或抗原性。

6. 抗原性的决定,以便确定第 8 因子的存在

为了确认培养基中的第 8 因子抗原,使用了 Riethorst, w 等人建立的新的 ELISA (酶联免疫吸附试验)。

以适当稀释的特异性单克隆抗体 α CagA (5mg/L) (溶于 0.05M 碳酸盐缓冲液 pH9.8) 200 μ l/每孔包被 ELISA 塑料板孔。封好后在 4 $^{\circ}$ C 温育过夜。每两次温育之间,反应板都应用含 0.05% 吐温 20 的磷酸盐缓冲液冲洗三次,每次至少一分钟。

试验样品(正常血浆)系列稀释后加入反应孔内(200 μ l/每孔)。反应板封好后 37 $^{\circ}$ C 温育 2 小时,勿震荡。稀释抗原所用缓冲液含

50mM Tris- HCl, 2% 牛血清白蛋白和 1.0M NaCl, pH7.4。

α Cag117-辣根过氧化物酶结合物用含 50mM Tris- HCl 缓冲液, 0.2% 吐温 20, 和 1.0M NaCl pH7.4 缓冲液稀释 10000 倍(依据敏感度的需要)。每孔加入该稀释液 200 μ l, 反应板封好后在 37 $^{\circ}$ C 暗处温育 2 小时。

冲洗反应板后,每孔加入 150 μ l 用 0.1M 醋酸/枸橼酸缓冲液 pH5.5 新配制的四甲基联苯胺(TMB)液(0.1g/l)和过氧化氢(0.006%),室温下暗处温育 30 分钟。加入 150 μ l 2N 硫酸终止酶反应,用 ELISA 微量测定仪在光波 450nm 下测定吸光率。

7. 抗第 8 因子单克隆和多克隆抗体的制备以其定性

以纯化的人类第 8 因子-Von Willebrand 因子复合物免疫 Balb/c 品系小鼠来制备单克隆抗体。按 Van Mourik 和 Mochtar (1970) 所述,纯化过程经对结晶沉淀物的琼脂糖凝胶过滤法或用于治疗目的的第 8 因子浓缩制品(由荷兰红十字输血服务中心实验室提供,阿姆斯特丹,荷兰)来完成。按 Galfre'等(1977)所述进行淋巴细胞杂交。用以分离产生抗第 8 因子单克隆抗体克隆的技术细节另有提供(Stel, 1984)。抗第 8 因子单克隆抗体与第 8 因子多肽的反应按英国专利申请 GB8617594 所述鉴定。

要得到抗第 8 因子多肽的多克隆抗体,根据标准步骤用层析法纯化的第 8 因子制剂免疫家兔(见 stel, 1984)。再根据以前的专利申请 GB8617594 所述的免疫吸印法鉴定所得抗体,该方法用纯化的第 8 因子-VonWillebrand 因子或由 SDS-聚丙烯酰胺凝胶电泳分离纯化的多肽进行。然后将该抗体转移到硝酸纤维素纸上作为靶蛋白。分析结果表明产生了三个明显不同的多克隆抗血清(RH63271, RH63272, RH63275)。这些抗血清与 80KDa 双带, 92KDa 多肽及更大的多肽反应,它们也与较小的肽反应。

附图说明

图 1 表示为全长第 8 因子蛋白编码的表达载体(PCLB201)的结构。

环形质粒 PCLB201 显示了在从质粒 PSV2 (Gorman, 1985; Mulligan 和 Berg, 1981) 衍生

的 4217 对碱基的 HindIII-BglII 片段中 EcoRI 位置有断裂。图还指明了早期 SV40 启动子区: SVep。该载体也含有从质粒 PRSV-neo (Gorman, C, 1985) 衍生的 NdeI-HindIII 片段, 质粒 PRSV-neo 带有插入 SalI 点的 Rous 肉瘤病毒长末端重复区 (LTR)。全长第 8 因子是由 7440 对碱基的 SalI-HpaI 片段编码。

该片段用第 8 因子 cDNA 表示 (H=Hind III, Sal=SalI)。组建 PCLB201 时失去的限制性内切酶切点在括号内。质粒大小用千对碱基 (Kb) 表示。

图 2: 过滤表达第 8 因子蛋白缺失突变的载体

a. PSV2 衍生的载体的标志: 两个位置随机的启动子: SV40 早期转录启动子 (SVep) 和 Rous 肉瘤病毒长末端重复区 (RSV-LTR); 带帽位点 (Capsite) 和信息 RNA 的 5' 端; 携带具有起始密码 (ATG), 开链读码和终止密码 (TGA) 的全长第 8 因子编码区的 cDNA 插入体; 具有短内含子和从 SV40DNA 衍生多腺苷信号 (polyA) 的 RNA 3' 端非编码区

b. 表示携带全长第 8 因子编码 cDNA 的 7440 对碱基片段 SalI-HpaI。也指出了读码的起始和终止密码。还显示了在全长 cDNA 内限制内切酶切点涉及组建 PCLB202 和 PCLB203 的突变。

c. 第 8 因子蛋白的氨基酸顺序图。显示了 19 氨基酸长信号肽和血浆第 8 因子的显性或重复结构, A_1 , A_2 和 A_3 是同质氨基酸序列。B 是一个单一区, 而 C_1 和 C_2 又是同质序列 (Vebar 等, 1984)。给出了靠近 A 和 C 重复区的氨基酸位置。图下端的蛋白酶切点, 这些切点剪接第 8 因子, 已用箭头标出了活化的蛋白质 c (APC), 凝血因子 (IIa), 活化的凝血因子 X (Xa) 和类胰蛋白酶 (用 ? 表示)。还说明 Xa 也一直切到 IIa- 和 APC 切点 (Eaton 等人, 1986a; Fay 等人, 1986) 图中都给出了它们的切点, B 区含有多个切点。

d. 活化第 8 因子的亚单位结构

92KDa 和 80KDa 的第 8 因子亚单位分别用 92K 和 80K 表示。还给出了全长顺序内氨基端和羧基端氨基酸的位置。

e. 表示 cDNA 和在 B 中提及 PstI 和 BamHI 点间含有直接融合的 PCLB202 蛋白的结构。显示

所得蛋白在丙-867 和天门-1563 间含有一个肽键。该蛋白全长为 1637 个氨基酸。

f. 表示 cDNA 和含为 4 个额外氨基酸编码的 MluI 连接子序列的 PCLB203 蛋白的结构, 这四个氨基酸桥接着 b 所示的 MaeIII 点和 HgiAI 点。该蛋白全长为 1411 个氨基酸。

e, f, g 给出了重复区边界位置的界碑和特异蛋白酶切点, 其解释见上文。

图 3: 用免疫沉淀法及电泳法测定 n 种缺失突变的第 8 因子蛋白分子量 (见一般方法 3)。

出现的放射自显影表明是载体 PCLB202 (2 泳道) 和 PCLB203 (1 泳道) 产生蛋白的泳道, 还指出了分子量的参照标志物 (见表 III)。

例 1: 组建 PCLB201

以前的申请 GB 8617594 和 GB 8630699 中已叙述内容, 如从人肝脏中提取第 8 因子的 mRNA, 制备, 纯化和鉴定其 cDNA, 以及质粒 PEP121 中组建质粒 PCLB99 均包括在本申请中。

为在细胞培养系统中表达, 已组建和修饰了从 Mulligan 和 Berg (1981) 的 PSV2 载体衍生的一个载体。质粒 PSV2.tPA (Van Zonneveld 等, 1986) 的单一 HindIII 点修饰成 SalI 位点。为此, 将 HindIII 粘性末端变成平钝端, SalI-连接子 (5'-CGTCGACG-3'); 连接到这些末端, 筛选含 SalI 位点的质粒 PCLB91。这个质粒与 PSVtpA 的不同仅在于它含插入 HindIII 点中的一个 SalI 连接子。

纯化来自 PCLB89 含有第 8 因子 cDNA (见表 IV) 的长 7440 碱基对的 SalI-HpaI 片段 (见 Maniatis 等人, 1982), 用 SalI 和 BglII 酶解, 将其插入 PSV2 衍生的载体 PCLB91 中。筛选出的表达载体 PCLB200 含有完整的 SalI 切点, 其 HpaI 点连接到 BglII 点, 恰好吻合。

将一个额外的 Rous 肉瘤病毒长末端重复区序列作为一个附加转录启动子接到质粒 PCLB200 上。用 NdeI 和 HindIII 酶解质粒 PRSVneo (Gorman, 1985), 再与 DNA 聚合酶的 Klenow 片段一起温育, 制造平钝端, 提取 0.6 千对碱基片段, 插入到同样被变成平钝端的 PCLB200 的 SalI 位点中。所得表达载体 PCLB201, 在转录起始点右边含 RSV-LTR, 再提取全长第 8 因子的 cDNA (见图 1)。

例 2: 组建 PCLB202

如专利申请 GB86175-94 和 GB8630699 中所述, 在质粒 PGB860 中组建一个突变体, 该突变体在第 8 因子序列的第 2660 个核苷酸 (见表 IV) 到 4749 个核苷酸一段缺失。PGB860 中第 8 因子的编码 DNA 在中心区含有一段为 695 个氨基酸编码的缺失。表达载体 PCLB202 是从 PGB860 和 PCLB201 衍生的。两种质粒均用 KpnI 和 XbaI 限制酶酶解。在 PGB860 中第 8 因子中心区含有缺失的 3.1 千对碱基 KpnI-XbaI 片段被连接到来自 PelB201 的 7 千对碱基的 KpnI-XbaI 片段上。所得质粒 PCLB202 具有完整的 KpnI 和 XbaI 切点。图 2e 图解了它的结构。

例 3: 组建 PCLB203

将来自 PCLB897440 对碱基的 SalI-HpaI 片段 (图 1) 插入用 SalI 和 SmaI 酶切的质粒 PsP64 中 (见 Melton 等人, 1984)。所得质粒含一个完整 SalI 点, 并将 HpaI 端连接到 SmaI 端。

PCLB101 的组建包括下列步骤 (括号中数字参照表 IV)

步骤 1. 用 KpnI 和 TthIII 酶解质粒 PCLB100, 得到一个 631 对碱基的片段: 从 KpnI (1817) 到 TthIII (2447)。用 MacIII (切点在 2199 位) 切纯化的 631 对碱基长的片段, 填平所得 MacIII 粘性端, 得到 383 对碱基片段。

步骤 2. 用 ApaI 和 BanI 酶解质粒 PCLB100, 提取 669 对碱基的 ApaI (6200) - BanI (5532) 片段。

步骤 3. 用 HgiAI 酶解质粒 PCLB100, 产生 634 对碱基的 HgiAI 片段, 将其与 T₄DNA 聚合酶温育, 使其变成平钝端。再用 BamI 酶解该片段, 得到 565 对碱基的片段。

步骤 4. 用 ApaI 和 KpnI 酶解质粒 PCLB100, 制造一个含 5.9 千对碱基长的片段 ApaI (6200) - KpnI (1817) 的载体序列, 保存并在大肠杆菌中增殖。

步骤 5. 纯化步骤 1-4 所得片段, 将等量的各片段与 10 倍摩尔数的过量 MluI 连接子 (CCACGCGTGG) 连接, 得到质粒 PCLB101。

鉴定在 MaeIII 和 HgiAI 点间 (见图 26) 含 MluI-连接子的质粒 PCLB101。证实了 KpnI, BamI 和 ApaI 点均是完整的。图 2f 图解天冬-712

和谷氨酰-1638 间 (见表 IV) 的四个额外氨基酸 (脯-精-缬-丙)

质粒 PCLB203 是直接来自 PCLB101 和 PCLB201 衍生的。两种质粒均用 KpnI 和 XbaI 酶解。将来自含有缺失的 PCLB101 的 2.4 千对碱基片段 KpnI-XbaI 连接到 PCLB200 衍生的 7.0 千对碱基长 KpnI-XbaI 片段。筛选的质粒 PCLB203 具有完整的 KpnI 和 XbaI 切点, 它们的结构图解在图 2f 中。

例 3a: 组建 PCLB212

用 Kranser 等人所述圈出突变技术 (loop-out mutagenesis) (见一般方法 2) 组建 PCLB212。用下列寡核苷酸使第 8 因子 cDNA 中心区缺失突变:

3'TTC CAA AGA TCA ACA CTG GTT 引物 D

TTG GGT GGT CAG AAC5'

在第 8 因子 cDNA 内发生为第 713 (赖) 到 1674 (丝) 氨基酸编码的序列 (引物 D) 的缺失。

其相应的蛋白质与全长第 8 因子蛋白比较含有 925 个氨基酸的内缺失部分。

为得到圈出突变的靶片段。选用来源于 PCLB203 (见例 3) 的 1.4 千对碱基的 HindIII-Pst I 片段。HindIII 点的核苷酸位置 (见表 IV) 在全长第 8 因子序列的 1025 核苷酸处, 恰好在欲饰区的上游; Pst I 点在 5165 核苷酸处, 恰在该区下游。图 1 指出了这些点。将 1.4 千对碱基片段亚克隆到 M13mp9 中, 然后进行圈出突变。筛选具有准确缺失的片段后再将含所需缺失的 HindIII-Pst I 片段的 KpnI-Pst I 部分插入到合适的表达载体中, 插入是通过制备具粘性单一限制酶末端 (数字参照表 IV) 的片段完成的, 步骤如下:

步骤 1: 为表达突变第 8 因子分子, 组建一个新质粒 PCLB211。因此, 从 PCLB89 衍生有 7440 对碱基 (见图 1) 的 Sal I-Hpa I 片段插入表达载体 PSVL (pharmacia, N, 27-4507-01, Uppsala, Sweden), 使其能在哺乳动物细胞中从晚期 SV40 启动子转录。用 XhoI 和 SmaI 酶解质粒 PSVL, 所得质粒 PCLB211 含连接到 SalI 端的 XhoI 端, 因为这些酶的 5'端突起端是相同的, SmaI 端连接到 HpaI 端。

步骤 2: 用 *ApaI* 和 *SalI* 酶解质粒 PCLB211, 提取长 1.4 千对碱基的 *ApaI* (6200) -*SalI* (该点在 PCLB211 的 PSVL 部分中是唯一的) 片段。

步骤 3: 用 *NdeI*, *PstI* 和 *ApaI* 酶解质粒 PCLB100, 提取 363 对碱基的 *PstI* (5165) -*NdeI* (5527) 片段和 674 对碱基的 *NdeI* (5527) -*ApaI* (6200) 片段。

步骤 4: 用 *KpnI* 和 *SacI* 酶解质粒 PCLB100, 产生一个 1799 对碱基的 *KpnI* (1817) -*SacI* (19) 片段, 提取之。

步骤 5: 用 *SalI* 和 *SacI* 酶解质粒 PCLB211。得到一个 4.5 千对碱基的 *SalI* (在 PSVL 载体中的唯一切点) -*SalI* (19) 片段。

步骤 6: 用 *KpnI* 和 *PstI* 酶解具有含序列分析证实有所需缺失的 *Hind* (1025) -*PstI* (5165) 片段的 M13 噬菌体。得到并纯化 584 对碱基的 *PstI* (5165) -*KpnI* (1819) 片段。

步骤 7: 将步骤 2 到 6 分离所得 6 种片段等量混合及连接。筛选含所有片段的质粒。质粒 PCLB212 表达典型化合物 3b (见表 II), 化合物 3b 不同于化合物 3 (见图 3f), 由于它失去 *MluI* 连接子和因而失去相应的四种氨基酸。

例 4: 组建 PCLB208, PCLB209, 和 PCLB210

用 Kramer 等人 (1984) 所述 (一般方法 2) 圈出突变技术组建 PCLB208, PCLB209, 和 PCLB210。用下列寡核苷酸进行第 8 因子 cDNA 中央区的缺失突变:

3'TTA CGG TAA CTT GGT TCT
CTT
引物 a

TAT TGA GCA TGA TGA5'
3'TTA CGG TAA CTT GGT
TCT AGT

引物 b

CAA CTT TAC TTC TTC5'
3'TTA CGG TAA CTT GGT
TCT TCG

引物 c

AAA GTT TTC TTT TGT5'

由于在第 8 因子 cDNA 产生了为丝-747 到精

1648 (引物 a), 丝-741 到异亮-1668 (引物 b), 及丝-741 到精 1689 (引物 c) 编码序列的内部缺失, 相应的蛋白质分别含有 908, 928, 和 949 个氨基酸残基的内部缺失。

为得到一个圈出突变的靶片段, 由 PCLB101 衍生的 0.8 千对碱基的片段组建如下:

用 *EcoRI* 酶解质粒 PCLB101, 将 *EcoRI* 切点填平, 再用 *KpnI* 进一步酶解。分离 479 个碱基对的由 *KpnI* (在表 IV 的核苷酸号为 1819) 开始到 *EcoRI* (2279) 平端片段。用 *BamHI* 酶解质粒 PCLB101, 填平粘性末端, 用 *PstI* 进一步酶解。第二个长 416 个碱基对的从 *BamHI* (4746) (末端已填平) 到 *PstI* (5165) 的片段被分离出来。将这些片段等摩尔连接, 亚克隆到 M13mp19 载体中。含有从欲修饰区上游的 1819 核苷酸到该区下游 5165 核苷酸的 895 对碱基的 *KpnI*-*PstI* 片段并具有第 8 因子区编码序列大缺失区的 M13 噬菌体被选用于圈出突变。选择在 M13 噬菌体中的引物 a, b 和 c 得到准确缺失的片段后, 将含有所需缺失的 *KpnI*-*PstI* 片段插入来源于 PCLB211 的合适表达载体, 见例 36, 其步骤如下, 制备具有单一限制酶粘性末端的片段 (核苷酸号参见表 IV):

步骤 1: 用 *ApaI* 和 *SalI* 酶解质粒 PCLB211。分离 1.4 千对碱基的 *ApaI* (6200) -*SalI* (PCLB211) 质粒中 PSVL 部分的唯一切点) 片段。

步骤 2: 用 *NdeI*, *PstI* 和 *ApaI* 酶解质粒 PCLB100, 分离 363 对碱基的 *PstI* (5165) -*NdeI* (5527) 片段和另一个 674 对碱基的 *NdeI* (5527) -*ApaI* (6200) 片段。

步骤 3: 用 *KpnI* 和 *SacI* 酶解质粒 PCLB100, 分离产生的 1799 对碱基的 *KpnI* (1817) -*SacI* (19) 片段。

步骤 4: 用 *SalI* 和 *SacI* 和酶解质粒 PCLB211。得到 4.5 千对碱基的 *SalI* (PSVL 载体的唯一切点) -*SacI* (19) 片段。

步骤 5: 用 *KpnI* 和 *PstI* 酶解含有所需突变片段的 M13 噬菌体。分离含有三种所需突变的 *KpnI*-*PstI* 片段。

步骤 6: 分离步骤 1-5 所得 6 种片段, 等摩尔数混合并连接。筛选含有所需片段的质粒。用上面所示的引物 a, b 或 c 组建表达表 I 所示举例化合物

的新表达载体:

1. 用引物 a 组建的表达化合物 7 的 PCLB208。

2. 用引物 b 组建的表达化合物 8 的 PCLB209。

3. 用引物 c 组建的表达化合物 9 的 PCLB210。

用不同启动子和细胞株可提高表达的水平。

例 5: 组建生产具有第 8 因子活性蛋白的稳定细胞株

a. 在 CHO 细胞中表达

用磷酸钙沉淀技术 (Graman Van der Eb, 1973) 将质粒 PCLB203 (10 μ g) 和 PAdD26sv (A)-3 (1 μ g, kaufman 和 sharp, 1982) 引入 DHFR 缺陷型 CHO 细胞中 (chasin 和 Urlaub, 1980)。如 Kaufman 和 Sharp (1982a) 所述, 用逐滴加入增加氨甲喋呤 (MTX) 浓度的方法使第 8 因子和 DHFR 编码序列扩增。将抗 200nM 氨甲喋呤的转化体取出, 建立稳定的在宿主基因组含为第 8 因子编码 DNA 的细胞株。这些细胞株可使每毫升培养基产生约 95 毫单位的第 8 因子活性。通过增加氨甲喋呤浓度进一步扩增第 8 因子编码序列, 从而使分泌到培养基中第 8 因子的量进一步增加。

建立生产具有第 8 因子活性蛋白质的细胞株是合适的。

为大量生产这种凝血因子, 也可用其它启动子 (如巨细胞病毒启动子或腺病毒启动子) 和 / 或细胞株 (如人 293 细胞, 骨髓瘤细胞或诸如肝细胞的特异分化细胞)。

b. 在 C127 细胞株中表达

如上所述, 表达载体可被引入真核细胞中, 它们整合到宿主细胞的基因组中。随后再扩增表达暗盒 (表达序列)。对表达载体整合的补充是作为随体稳定地保存在细胞内, 例如: 用随体 BPV 系统 (Howley 等人 1983) 表达一种突变第 8 因子蛋白。BPV-1 基因组 (BamHI 切的) 首先被引入 PTZ18R (来自 Pharmacia 公司) 用 BamHI 切的衍生物, PTZ18R 在 BamHI 切点的两处含有 XhoI 点。随后用 XhoI 酶切所得 PTZX-BPV 质粒, 得到 2.9 千对碱基的 PTZ 片段和具有 Xho-I 突端的 8 千对碱基的 BPV 片段。后一个片段连接

到质粒 PCLB212, 该质粒在它唯一的 SalI 切点处被切 (即在原始 PSVL 载体的 2040 位置)。所得表达载体 PGB881 含有 SV40 晚期启动子, 为缺失 925 个氨基酸的第 8 因子编码的 cDNA (主要为 B-部分) 和 SV40 晚期聚腺苷信号。由于有 BPV-基因组存在, 这个载体在合适宿主细胞, 如: 小鼠 C127 细胞中可以随体形式保存下来。用磷酸钙沉淀技术 (Graham 和 Van der Eb, Supra) 将质粒 PGB881 (10 μ g) 转染到 C127 细胞中。转染后 14 天分离 Foci, 随后建立稳定的细胞株。一些细胞株可在每毫升培养基中产生 40 个毫单位的第 8 因子活性。

一般在本发明范围内可改变表达载体, 特别是启动子和 / 或启动子 / 增强子组合 (Hurwitz 1987) 以及宿主, 目的是提高本发明第 8 因子活性蛋白的水平。

例 6: 表达重组第 8 因子 DNA 和测试产生的蛋白

将根据上述例子得到的表达载体用 DEAE 转染技术 (见一般方法 3) 引入 Cos-1 细胞。转染后 4 天收集控制条件的培养物。测试产生的蛋白。

根据一般方法部分 (第 5 章) 用 (1) 标准凝血或成血块试验和 (2) 成色活性试验 (Kabi coatest) 测定控制条件培养物中第 8 因子的活性。加入已知是从血浆衍生具有第 8 因子活照泳道出现一个较小的带。

同样从表 III 看出, 似乎第 8 因子是糖蛋白, 与天然第 8 因子相似。因为在有或无已知是天冬氨酸连接的糖化抑不同蛋白的结果均收集在表 II A 和 II B, 表明缺失突变蛋白质甚至影响 92KDa 区, 显示第 8 因子的凝血活性。

用 ELISA 试验 (见一般方法 6) 确定第 8 因子蛋白量。从表 II A 和 II B 可看出培养液中第 8 因子活性是随着其中第 8 因子蛋白的含量增加而增加的。

用一般方法部分第 3 章所述的凝胶电泳确定产生的第 8 因子蛋白的大小。从图 3 及表 III 资料可推导出新的第 8 因子蛋白分子大小与糖蛋白分子相仿。对照泳道出现一个较小的带。

同样从表 III 看出, 似乎第 8 因子是糖蛋白, 与天然第 8 因子相似。因为在有或无已知是天冬氨酸连接的糖化抑制剂时, 形成的蛋白大小有明显的区别。

表 1
缺失突变第八因子蛋白

化合物	N	L	C	de l	表达载体
1	Ala-1→Ser-1637.....	31a-1638 → Tyr-2332	0		pCLB 201
2	Ala-1→Ala-867.....	Asp-1563 → Tyr-2332	695		pCLB 202
3	Ala-1→Asp-712.....Pro.....Arg.....Val...Ala...Gln-1638 → Tyr-2332		925		pCLB 203
3b	Ala-1→Asp-712.....	Gln-1638 → Tyr-2332	925		pCLB 212
7	Ala-1→Arg-740.....	Ser-1649 → Tyr-2332	908		pCLB 208
8	Ala-1→Arg-740.....	Ser-1669 → Tyr-2332	928		pCLB 219
9	Ala-1→Arg-740.....	Ser-1690 → Tyr-2332	949		pCLB 210

— 未断裂的第八因子氨基酸序列

... 一个肽键

与全长第八因子比较缺失氨基酸的数目列在 de l. 栏中, 数字与表 4 相同.

表 2

第八因子活性和蛋白量

化合物 ¹⁾	氨基酸数	(缺失数)	第八因子活性 (mU/ml) ²⁾		加入抗体后剩余活性 ²⁾		抗原测定 ³⁾
			2)	3)	4)	5)	
1	2332	(0)	1.0	- 6)	- 6)	- 6)	<0.5 mU
2	1637	(695)	6.7	2	< 1%	< 1%	7 mU
3	1411	(925)	17.3	5	< 2%	< 1%	20 mU
3b	1607	(925)	15.0	ND ⁹⁾	ND	ND	20 mU
4	1069	(1263)	0	0	0	0	20 mU
7	1424	(908)	15.0	ND	ND	ND	ND
8	1404	(928)	15.0	ND	ND	ND	ND
9	1383	(949)	0	ND	ND	ND	ND
对照 (pSV2)			0	0	ND	ND	0

- 1) 见表 1
- 2) 根据 K a b i C o a t e s t 成色试验
- 3) 标准成血块试验
- 4) a C a l g A, 第八因子的单克隆抗体
- 5) 从病人血清中分离的抑制剂
- 6) 对鉴别可能的失活起始活性太低
- 7) 用 100 n g 第八因子蛋白 (相当 1 个单位第八因子) 在控制下的培养基中温育 48 小时
- 8) 酶联免疫试验 (E L I S A)
- 9) ND = 未测定

表 3
分泌的第八因子的大小

化合物 ¹⁾	表达载体	氨基酸数 ²⁾	缺失 ³⁾	计算的大小 ⁴⁾	测定的大小 ⁵⁾	糖基化抑制 ⁶⁾
1	pCLB 201	2332	0	265 kDa		
2	pCLR 202	1637	695	138 kDa	192 kDa	135 kDa
3	pCLB 203	1411	925	162 kDa	168 kDa	158 kDa
3b	pCLB 212	1407	925	162 kDa	168 kDa	

1) 见表 1

2) 氨基酸数是第八因子氨基酸总数

3) 缺失是从全长第八因子缺失的氨基酸数目

4) 根据已知氨基酸组成推断的大小

5) 放射免疫治疗后电泳测得的大小

6) 从生长在含 tunicamycin (0.001 mg/ml) 培养基中细胞得到蛋白的大小

表 4

第八因子CDNA插入到pCLB89中。第八因子从丙-1到酪-2332的氨基酸顺序和它的信号序列蛋(-19)到丝(-1)。下面对应有其编码的核苷酸序列。苯丙-973是由TTC编码。

10 20 30 40 50 60
 GTCGACATGCAAATASAGCTCTCCACCTGCTTCTTCTGTGCGCTTTTGGGATTCTGCTTT
 -19 M Q I E L S T C F F I C L L R F C F -2
 70 80 90 100 110 120
 AGTGCCACCAGAAAGATACTACCTGGGTGCAGTGGAACTGTGATGGGACTATATGCAAAGT
 S A T R R Y Y L G A V E L S W D Y M Q S 19
 130 140 150 160 170 180
 GATCTCGGTGAGCTGCCTGTGGACGCAAGATTTCCTCTAGAGTGGCAAAATCTTTTCCA
 D L G E L F V D A R F P P R V P K S F P
 190 200 210 220 230 240
 TTCAACACCTCAGTCGTGTACAAAAGACTCTGTTTGTAGGATTTCACGGATCACCCTTTTC
 F N T S V V Y K K T L F V E F T D H L F 59
 250 260 270 280 290 300
 AACATCGCTAAGCCCAAGGCCACCTGGATGGGTCTGCTAGGTCTTACCATCCAGGCCTGAG
 N I A K P R P P W M G L L G P T I Q A E
 310 320 330 340 350 360
 GTTTATGATACAGTGGTCATTACACTTAAGAACATGGCTTCCCATCCTGTCACTCTTCAT
 V Y D T V V I T L K N M A S H P V S L H 99
 370 380 390 400 410 420
 GCTGTTGGTGTATCCTACTGGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCACT
 A V G V S Y W K A S E G A E Y D D Q T S
 430 440 450 460 470 480
 CAAAGGGAGAAAGAAGATGATAAAGTCTTCCCTGGTGGAAAGCCATACATATGTCTGGGAG
 Q R E K E D D K V F P G G S H T Y V W Q 179
 490 500 510 520 530 540
 GTCCTGAAAGAGAAATGGTCCAATGGCCTCTGACCCCACTGTGCTTACCTACTCATATCTT
 V L K E N G P M A S D P L C L T Y S Y L
 550 560 570 580 590 600
 TCTCATGTGGACCTGGTAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGT
 S H V D L V K D L N S G L I G A L L V C 179
 610 620 630 640 650 660
 AGAGAAGGGAGTCTGGCCCAAGGAAAGACACAGACCTTGCACAAATTTATCTACTTTT
 R E G S L A K E K T Q T L H E F I L L F
 670 680 690 700 710 720
 GCTGTATTTGATGAAGGGAAAGTTGGCACCTGAGAAACAGGAACTCTTGTATGACAGGAT
 A V F D E G K S W H S E T E N S L M Q D 219
 730 740 750 760 770 780
 AGGGATGCTGCATCTGCTCGGGCCTGGCCTAAATGCACACAGTCAATGGTTATGTAAC
 R D A A S A R A W P K M H T V N G Y V N
 790 800 810 820 830 840
 AGGTCTCTGCCAGGTCTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGA
 R S L P G L I G C H R K S V Y W H V I G 259
 850 860 870 880 890 900
 ATGGGCACCACTCCTGAAGTGCACCTCAATATTCTCTGAAAGGTCACACATTCTTGTGAGG
 M G T T F E V H S I F L E G H T F L V R

910 920 930 940 950 960
 AACCATCGCCAGGCGTCCCTTGGAATCTCGCCAAATACCTTCTTACTGCTCAACACTC 299
 N H R Q A S L E I S P I T F L T A Q T L

970 980 990 1000 1010 1020
 TTGATGGACCTTGGACAGTTTCTACTGTTTTGTCAATATCTCTTCCCACCAACATGATGGC
 L M D L G Q F L L F C H I S S H Q H D G

1030 1040 1050 1060 1070 1080
 ATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAAT 339
 M E A Y V K V D S C P E E P Q L R M K N

1090 1100 1110 1120 1130 1140
 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGG
 N E E A E D Y D D D L T D S E M D V V R

1150 1160 1170 1180 1190 1200
 TTTGATGATGACAACTCTCCTTCCCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCCT 379
 F D D D N S P S F I Q I R S V A K K H F

1210 1220 1230 1240 1250 1260
 AAACTTGGGTACATTACATTGCTGCTGAAGAGGAGGACAGGACTATGCTCCCTTAGTC
 K T W V H Y I A A E E E D W D Y A P L V

1270 1280 1290 1300 1310 1320
 CTCGCCCCCGATGACAGAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATT 419
 L A P D D R S Y K S Q Y L N N G F Q R I

1330 1340 1350 1360 1370 1380
 GGTAGGAAGTACAAAAAGTCCGATCTATGGCATACAGGATGAACCTTTAAGAGTCGT
 G R K Y K K V R F M A Y I D E T F K T R

1390 1400 1410 1420 1430 1440
 GAAGCTATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAAGTTGGAGAC 459
 E A I Q H E S G I L G P L L Y G E V G D

1450 1460 1470 1480 1490 1500
 ACACTGTTGATTATATTTAAGAATCAAGCAAGCAGACCATATAACATCTACCCCTCAGGGA
 T L L I I F K N Q A S R P Y N I Y P H G

1510 1520 1530 1540 1550 1560
 ATCACTGATGTCCGTCTCTTTGTATTCAAGGAGATTACCAAAAGGTTGTAACACATTTGAAG 499
 I T D V R P L Y S R R L P K G V K H L K

1570 1580 1590 1600 1610 1620
 GATTTTCCAATTCTGCCAGGAGAAATATTCAATATAAATGGACAGTGAAGTGTAGAGAT
 D F P I L P G E I F K Y K W T V T V E D

1630 1640 1650 1660 1670 1680
 GGGCCAACTAAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATATG 539
 G P T K S D P R C L T R Y Y S S F V N M

1690 1700 1710 1720 1730 1740
 GAGAGAGATCTAGCTTCAGGACATCATGGCCCTCTCTCATCTGCTACAAAGAATCTGTA
 E R D L A S G L I G P L I C Y K E S V

1750 1760 1770 1780 1790 1800
 GATCAAGAGGGAACCCAGATAATGTCAGACAAGAGGAATGTCATCCTGTTTTCTGATTT 579
 D Q R G N D I M S D K R N V I L F S V F

1810 1820 1830 1840 1850 1860
 GATGAGAACCGAAGCTGGTACCTCACAGAGGAAATATACCAAGCTTTCTCCCAATCCAGCT
 D E N R S W Y L T E N T O R F L P N P A

1870 1880 1890 1900 1910 1920
 GGAGTGCAGCTTGAGGATCCAGAGTTCCAAGCCACCAACATCATGCACAGCATCAATGGC
 G V Q L E D P E F Q A S N I H H S I N G 619

1930 1940 1950 1960 1970 1980
 TATGTTTTTGATAGTTTGCAGTTGTGAGTTTGTTCATGAGGTGGGATACTGGTACATT
 Y V F D S L O L S V C L H E V A Y W Y I

1990 2000 2010 2020 2030 2040
 CTAAGCATTGGAGCACAGACTGACTTCCTTTCTGCTCTCTCTCTGGATATACCTTCAA
 L S I G A Q T D F L S V F F S G Y T F K 659

2050 2060 2070 2080 2090 2100
 CACAAAATGGTCTATGAAGACACACTCACCTATTCCCATTTCTCAGGAGAAACTGTCTTC
 H K M V Y E D T L T L F P F S G E T V F

2110 2120 2130 2140 2150 2160
 ATGTCGATGGAAAACCCAGGTCTATGGATTCTGGGGTCCCACAACCTCAGACTTTCCGGAAC
 M S M E N P G L W I L G C H N S D F R N 699

2170 2180 2190 2200 2210 2220
 AGAGGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAGAGAACACTGGTGATTATTAC
 R G M T A L L K V S S C D K N T G D Y Y

2230 2240 2250 2260 2270 2280
 GAGGACAGTTATGAAGATATTTCAGCATACTTGGTGAAGTAAAGACAATGCCATTGAACCA
 E D S Y E D I S A Y L L S K N N A I E P 739

2290 2300 2310 2320 2330 2340
 AGAAGCTTCTCCCAAGATTCAAGACACCCTAGCACAGGCAAAAGCAATTTAATGCCACC
 R S F S Q N S R H P S T R Q K Q F N A T

2350 2360 2370 2380 2390 2400
 ACAATTCCAGAAAATGACATAGAGAGAGACTGACCCCTTGGTTTGCACACAGAACACCTATG
 T I P E N D I E K T D F W F A H R T F M 779

2410 2420 2430 2440 2450 2460
 CCTAAAATACAAATGTCTCCTCTAGTGATTGTGTTGATGCTCTTGGCAGAGAGTCTACT
 P K I Q N V S S S D L L M L L R Q S P T

2470 2480 2490 2500 2510 2520
 CCACATGGGCTATCCTTATCTGATCTCCAAGAGGCCAATATGAGACTTTTCTGATGAT
 P H G L S L S D L Q E A K Y E T F S D D 819

2530 2540 2550 2560 2570 2580
 CCATCACCTGGAGCAATAGACAGTAATAACAGCCTGTCTGAAATGACACACTTCAGGACA
 P S P G A I D S N N S L S E M T H F R P

2590 2600 2610 2620 2630 2640
 CAGCTCCATCACAGTGGGGACATGGTATTTACCCCTGAGTCAAGGCTCCAATTAAGATT
 Q L H H S G D M V F T P E S G L Q L R L 859

2650 2660 2670 2680 2690 2700
 AATGAGAAACTGGGGACAACCTGCAGCAACAGAGTTGAAGAAACTTGATTTCAAAGTTTCT
 N E K L G T T A A T E L K K L D F K V S

2710 2720 2730 2740 2750 2760
 AGTACATCAAATAATCTGATTTCAACAATTCATCAGACAATTTGGCAGCAGGTAAGTAT 899
 S T S N N L I S T I P S D N L A A G T D

2770 2780 2790 2800 2810 2820
 AATACAAGTTTCCTTAGGACCCCCAAGTATGCCAGTTCATATGATAGTCAATTAGATACC
 N T S S L G P P S M P V H Y D S Q L D T

2830 2840 2850 2860 2870 2880
 ACTCTATTTGGCAAAAAGTCATCTCCCCTTACTGAGTCTGGTGGACCTCTGAGCTTGAGT 939
 T L F G K K S S P L T E S G G P L S L S

2890 2900 2910 2920 2930 2940
 GAAGAAAATAATGATTCAAAGTTGTTAGGAATCAGGTTTAAATGAATAGCCAAAGAAAGTTCA
 E E N N D S K L L E S G L M N S Q E S S

2950 2960 2970 2980 2990 3000
 TGGGGAAAAAATGTATCGTCAACAGAGAGTGGTAGGTTATTCAAAGGGAAAAAGAGCTCAT 979
 W G K N V S S T E S G R L F K G K R A H

3010 3020 3030 3040 3050 3060
 GGACCTGCTTTGTTGACTAAAGATAATGCCTTATTCAAAGTTAGCATCTCTTTGTTAAAG
 G P A L L T K D N A L F K V S I S L L K

3070 3080 3090 3100 3110 3120
 ACAACAAAACTTCCAATAATTCAGCAACTAATAGAAAGACTCACATTGATGGCCCATCA 1019
 T N K T S N N S A T N R K T H I D G P S

3130 3140 3150 3160 3170 3180
 TTATTAAATTGAGAATAGTCCATCAGTCTGGCAAAATATATTAGAAAGTGCACCTGAGTTT
 L L I E N S P S V W Q N I L E S D T E F

3190 3200 3210 3220 3230 3240
 AAAAAAGTGACACCTTTGATTTCATGACAGAAATGCTTATGGACAAAAATGCTACAGCTTTG 1059
 K K V T P L I H D R M L M D K N A T A L

3250 3260 3270 3280 3290 3300
 AGGCTAAATCATATGTCAAATAAACTACTTCATCAAAAAACATGGAAATGGTCCAACAG
 R L N H M S N K T T S S K N M E M V Q Q

3310 3320 3330 3340 3350 3360
 AAAAAAGAGGGCCCCATTCCACCAGATGCACAAAATCCAGATATGTCGTTCTTTAAGATG 1099
 K K E G P I P P D A Q N P D M S F F K M

3370 3380 3390 3400 3410 3420
 CTATTCTTGCCAGAATCAGCAAGGTGGATACAAAGGACTCATGGAAAGAACTCTCTGAAC
 L F L P E S A R W I Q R T H G K N S L N

3430 3440 3450 3460 3470 3480
 TCTGGGCAAGGCCCCAGTCCAAAGCAATTAGTATCCTTAGGACCAGAAAAATCTGTGGAA 1139
 S G Q G P S P K Q L V S L G P E K S V E

3490 3500 3510 3520 3530 3540
 GGTGAGAATTTCTTGTCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTACAAAG
 G Q N F L S E K N K V V V G R G E F T K

3550 3560 3570 3580 3590 3600
 GACGTAGGACTCAAGAGAGATGGTTTTTCCAAAGCAGCAGAAACCTATTTCTTACTAACTTG 1179
 D V G L K E M V F P S S R N L F L T N L

3610 3620 3630 3640 3650 3660
 GATAATTTACATGAAAAATAATACACACAATCAAGAAAAAAATTCAGGAAGAAATAGAA
 D N L H E N N T H N Q E K K I Q E E I E

3670 3680 3690 3700 3710 3720
 AAGAAGGAAACATTAATCCAAGAGAATGTAGTTTIGCCTCAGATACATACAGTGACTGGC
 K K E T L I Q E N V V L P Q I H T V T G 1219

3730 3740 3750 3760 3770 3780
 ACTAAGAATTTTCATGAAGAACCCTTTTCTTACTGAGCACTAGGCAAAATGTAGAAAGGTTCA
 T K N F M K N L F L L S T R Q N V E G S

3790 3800 3810 3820 3830 3840
 TATGACGGGGCATATGCTCCAGTACTTCAAGATTTTAGGTTCATTAAATGATTCAACAAAT
 Y D G A Y A P V L Q D F R S L N D S T N 1259

3850 3860 3870 3880 3890 3900
 AGAACAAAGAAACACACAGCTCATTCTCAAAAAAAGGGGAGGAAGAAAACCTTGGAAAGGC
 R T K K H T A H F S K K G E E E N L E G

3910 3920 3930 3940 3950 3960
 TTGGGAAATCAAACCAAGCAAATTGTAGAGAAATATGCATGCACCACAAAGGATATCTCCT
 L G N Q T K Q I V E K Y A C T T R I S F 1299

3970 3980 3990 4000 4010 4020
 AATACAAGCCAGCAGAAATTTGTACAGCAACCTAGTAAGAGAGCCTTGAACCAATTTCAGA
 N T S Q Q N F V T Q R S E R A L K Q F R

4030 4040 4050 4060 4070 4080
 CTCCCACTAGAAGAAACAGAACTTGAAAAAAGGATAATTTGTTGATGACACCTCAACCCAG
 L P L E E T E L E K R I I V D D T S T Q 1339

4090 4100 4110 4120 4130 4140
 TGGTCCAAAAACATGAACCATTTGACCCCGAGGCACCTCAGCAGATAGACTACAATGAG
 W S K N M K H L T P S T L T Q I D Y N E

4150 4160 4170 4180 4190 4200
 AAGGAGAAAGGGGCCATTACTCAGTCTCCCTATACAGATIGLCITACGAGGAGTCATAGC
 K E K G A J T Q S P L S D C L T R S H S 1379

4210 4220 4230 4240 4250 4260
 ATCCCTCAAGCAAATAGATCTCCATTACCCATTGCAAGGATATCATCATTTCCATCTATT
 I P Q A N R S P L P I A F V S S F P S I

4270 4280 4290 4300 4310 4320
 AGACCTATATATCTGACCAGGGTCCCTATTCCAAAGACAATCTTCTCATCTTCCAGCAGCA
 R P I Y L T R V L F Q D N S S H L P A A 1419

4330 4340 4350 4360 4370 4380
 TCTTATAGAAAGAAAGATTCTGGGGTCCAAAGAAAGCAGTCATTCTTACAAAGGAGCCAAA
 S Y R K K D S G V Q E S S H F L Q G A K

4390 4400 4410 4420 4430 4440
 AAAAATAACCTTTCTTTAGCCATTCTAACCTTGGAGATGACTGGTGATCAAGAGAGGTTT
 K N N L S L A I L T L E M T G D Q R E V 1459

4450 4460 4470 4480 4490 4500
 GGCTCCCTGGGGACAAGTGCCACAAATTCAGTCACATACAAAGAAAGTTGAGAACACTGTT
 G S L G T S A T N S V T Y L E V E N T V

4510 4520 4530 4540 4550 4560
 CTCCCGAAACCAGACTTGCCTCAAAACATCTGGCAAGTTGAAATTGCTTCCAAAAGTTTCAC
 L P K P D L P P T S G E V E L L P K V H 1499

4570 4580 4590 4600 4610 4620
 ATTTATCAGAAGGACCTATTCCCTACGGAAACTAGCAATGGGTCTCCTGGCCATCTGGAT
 I Y Q K D L F P T E T S N G S P G H L D

4630 4640 4650 4660 4670 4680
 CTCGTGGAAGGGGAGCCTTCTTCAGGGAAACAGAGGGAGCGATTAAAGTGGGAATGAAGCAAC
 L V E G S L L Q G T E G A I K W N E A N 1539

4690 4700 4710 4720 4730 4740
 AGACCTGGAAAAGTTCCCTTTTCTGAGAGTAGCAACAGAAAGCTCTGCAAAAGACTCCCTCC
 R P G K V P F L R V A T E S S A K T P S

4750 4760 4770 4780 4790 4800
 AAGCTATTGGATCCTCTTGTCTTGGGATAACCACTATGGTACTCAGATACCAAAAGAAGAG
 K L L D P L A W D N H Y G I Q I P K E E 1579

4810 4820 4830 4840 4850 4860
 TGGAAATCCCAAGAGAAGTCACCAGAAAAACAGCTTTTAAAGAAAAGGATACCATTIIG
 W K S Q E K S P E K T A F K K K D T I L

4870 4880 4890 4900 4910 4920
 TCCCTGAACGCTTGTGAAAGCAATCATGCAATAGCAGCAATAAATGAGGGACAAAATAAG
 S L N A C E S N H A I A A I N E G Q N K 1619

4930 4940 4950 4960 4970 4980
 CCCGAAATAGAAATCAGCTGGCAAAAGCAAGGTAGGAGTGAAGGCTGTCTCTCAAAAT
 P E I E V T W A K Q G R T E R L C S Q N
 37 38

4990 5000 5010 5020 5030 5040
 CCACCACTCTTGAAACGCCATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCA
 P P V L K R H Q R F I T R I T L Q S D Q 1659

5050 5060 5070 5080 5090 5100
 GAGGAATTGACTATGATGATACCATATCAGTTGAAATGAAGGAAGGAAGATTTTGACATT
 E E I D Y D D I I S V E H K E E D F D I

5110 5120 5130 5140 5150 5160
 TATGATGAGGATGAAAATCAGAGCCCCCGCAGCTTTCAAAAGAAAACACGACACTATTTT
 Y D E D E N Q S P R S F Q K K T R H Y F 1699
 89 90

5170 5180 5190 5200 5210 5220
 ATTGCTGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGA
 I A A V E R L W D Y G M S S S P H V L R

5230 5240 5250 5260 5270 5280
 AACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTTCCAGGAATTTACT
 N R A Q S G S V P Q F K K V V F Q E F T 1739

5290 5300 5310 5320 5330 5340
 GATGGCTCCTTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTG
 D G S F T Q P L Y R G E L N E H L G L L

5350 5360 5370 5380 5390 5400
 GGGCCATATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACTTTCAGAAATCAGGCC
 G P Y I R A E V E D N I M V T F R N Q A 1779

5410 5420 5430 5440 5450 5460
 TCTCGTCCCTATTTCCTTCTATTCTAGCCTTATTTCTATGAGGAAAGATCAGAGGCAAGGA
 S R P Y S F Y S S L I S Y E E D Q R Q G

5470 5480 5490 5500 5510 5520
 GCAGAACCTAGAAAAAAGCTTTGTCAGGCGCTAATGAAAGCAAGATTACTTTTGGAAAGTG
 A E P R K N F V K P N E T K T Y F W K V 1819

5530 5540 5550 5560 5570 5580
 CAACATCATATGGCAGCCCACTAAAGATGAGTCTGAGCTGCAAGGCTGGGCTTATTCTCT
 Q H H M A P T L D F F D C F A W A Y F S

5590 5600 5610 5620 5630 5640
 GATGTTGACCTGGAAAAAGATGTGCACTCAGGCGCTGATTTGGAGCCCTTCTGGTCTGCCAC
 D V D L E K D V H S G L I G P L L V C H 1859

5650 5660 5670 5680 5690 5700
 ACTAACCACTGAACCCCTGCTCATGGAGACAAAGTGACAGTACAGGAATTTGCTCTGTTT
 T N T L N P A H G R Q V T V Q E F A L F

5710 5720 5730 5740 5750 5760
 TTCACCATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAATATGGAAAGAAAGCTGC
 F T I F D E T K S W Y F T E N M E R N C 1899

5770 5780 5790 5800 5810 5820
 AGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCAT
 R A P C N I Q M E D P T F K E N Y R F H

5830 5840 5850 5860 5870 5880
 GCAATCAATGGCTACATAATGGATACACTACCTGGCTTAGTAGTGGCTCAGGATCAAGG
 A I N G Y I M D T L P G L V H A Q D Q R 1939

5890 5900 5910 5920 5930 5940
 ATTCGATGGTATCTGCTCAGCATGGGCAGCAATGAAAACATCCATTCTATTTCATTTAGT
 I R W Y L L S M G S N E N I H S I H F S

5950 5960 5970 5980 5990 6000
 GGACATGTGTTCACTGTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTAT
 G H V F T V R K K E E Y K M A L Y N L Y 1979

6010 6020 6030 6040 6050 6060
 CCAGGTGTTTTTGAGACAGTGGAAATGTTACCATCCAAAGCTGGAATTTGGCGGGTGGAA
 P G V F E T V E M L P S L A G I W R V E

6070 6080 6090 6100 6110 6120
 TGCCTTATTGGCGAGCATCTACATGCTGGGATGAGCAACACTTTTCTGGTGTACAGCAAT
 C L I G E H L H A G M S T L F L V Y S N 2019

6130 6140 6150 6160 6170 6180
 AAGTGTCAAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGATTACAGCT
 K C Q T P L G M A S G H I R D F Q I T A

6190 6200 6210 6220 6230 6240
 TCAGGACAATATGGACAGTGGGCGCCCAAGGCTGGCCAGACTTCATTATCCGGATCAATC
 S G Q Y G Q W A P T L A R L H Y S G S I 2059

6250 6260 6270 6280 6290 6300
 AATGCCTGGAGCACCAAGGAGCCCTTTCTTGGATCAAGGTGGATCTGTTGGCAUCAAAG
 N A W S T K E P F S W I E V D L L A P M

6310 6320 6330 6340 6350 6360
 ATTATTACGGCATCAAGACCCAGGGTGCCCGTCAGAAATTCTCCAGCCTCTACATCTCT
 I I H G I K T Q G A R Q K F S S L Y I S 2099

6370 6380 6390 6400 6410 6420
 CAGTTTATCATCATGTATAGTCTTTGATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCC
 Q F I I M Y S L D G K K W Q T Y R G N S

6430 6440 6450 6460 6470 6480
 ACTGGAACCTTAATGGTCTTCTTTGGCAATGTGCATTTCATCTGGGATAAAACACAATATT
 T G T L M V F F G N V D S S G I K H N I 2139

6490 6500 6510 6520 6530 6540
 TTTAACCCCTCCAATTATTGCTCGATACATCCGTTTGGACCCCAACTCATTATAGCATTCCG
 F N P P I I A R Y I R L H P T H Y S I R

6550 6560 6570 6580 6590 6600
 AGCACTCTTCGCATGGAGTGGATGGGCTGTGATTAAATAGTTGCAGCATGCCATTGGGA
 S T L R M E W M G C D L N S C S M P L G 2179

6610 6620 6630 6640 6650 6660
 ATGGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTCATCCTACTTTACCAATATG
 M E S K A I S D A Q I T A S S Y F T N M

6670 6680 6690 6700 6710 6720
 TTTGCCACCTGGTCTCCTTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATGCCTGG
 F A T W S P S K A R L H L Q Q R S N A W 2219

6730 6740 6750 6760 6770 6780
 AGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGAUACTTCAGAAAGACAATGAAA
 R P Q V N N P K E W L Q V D F Q K T M K

6790 6800 6810 6820 6830 6840
 GTCACAGGAGTAAGTACTCAGGGAGTAAATCTCTGCTTACCAGCATGTATGTGAAGGAG
 V T G V T T Q G V K S L L T S M Y V K E 2259

6850 6860 6870 6880 6890 6900
 TTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTTCAGAAATGGCAA
 F L I S S S Q D G H Q W T L F F Q N G K

6910 6920 6930 6940 6950 6960
 GTAAAGGTTTTTCAGGGAAATCAAGACTCCTTCACACCTGTGGTGAAGTCTCTAGACCCA
 V K V F Q G N Q D S F T P V V N S L D P 2299

6970 6980 6990 7000 7010 7020
 CCGTTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTG
 P L L T R Y L R I H P Q S W V H Q I A L

7030 7040 7050 7060 7070 7080
 AGGATGGAGGTTCTGGGCTGCGAGGCACAGGACCTCTACTGAGGGTGGCCACTGCAGCAC
 R M E V L G C E A Q D L Y

7090 7100 7110 7120 7130 7140
 CTGCCACTGCCGTCACCTCTCCCTCCTCAGCTCCAAGGCCAGTGTCCCTCCCTGGCTTGCC
 2332 stop

7150 7160 7170 7180 7190 7200
 TTCTACCTTTGTGCTAAATCCTAGCAGACACTGCCCTTGAAGCCTCCTGAATTAAGTATCA

7210 7220 7230 7240 7250 7260
TCAGTCCTGCATTTCTTTGGTGGGGGCCAGGAGGGTGCATCCATTTAACTTAACTCTT

7270 7280 7290 7300 7310 7320
ACCTATTTTCTGCAGCTGCTCCCGAGTTACTCCTTCCTTCCAATATACTAGGCAAAAAAG

7330 7340 7350 7360 7370 7380
AAGTGAGGAGAAACCTGCATGGAAGCATTCTTCCTGAAAAGTTAGGCCTCTCAGAGTCA

7390 7400 7410 7420 7430 7440
CCACTTCCTCTGTTGTAGAAAACTATGTGATGAAACTTTGAAAAAGATATTTATGATGT

7450 7460 7470 7480 7490 7500
TAACATTTTCAAGTTAAGCCTCATACGTTTAAATAAAGTCTCAGTTGTTATTATCCTG

7510 7520 7530 7540 7550 7560
ATCAAGCATGGAACAAAGCATGTTTCAGGATCAGATCAATACAATCTTGGAGTCAAAAGG

7570 7580 7590 7600 7610 7620
CAAATCATTTTGGACAATCTGCAAAATGGAGAGAAATACAATACTACTACAGTAAAGTCTG

7630 7640 7650 7660 7670 7680
TTTCTGCTTCCTTACACATAGATATAATTATGTTATTAGTCATATGAGGGGCACATTC

7690 7700 7710 7720 7730 7740
TTATCTCCAAAAGTACGATTCTTAAAGTGAATATAGATGGGGTTCAAGAATCCCTAA

7750 7760 7770 7780 7790 7800
GTCCCCTGAAATTATATAAGGCATTCTGTATAAATGCAAAATGTGCATTTTCTGACGAGT

7810 7820 7830 7840 7850 7860
GTCCATAGATAAAAAGCCATTGGGTCTTAATTCTGACCAATAAAAAAATAAGTCAGGAGG

7870 7880 7890 7900 7910 7920
ATGCAATTGTTGAAAGCTTTGAAATAAAATAACAATGTCCTTCTTGAAATTTGTGATGGCC

7930 7940 7950 7960 7970 7980
AAGAAAGAAAATGATGATGACATTAGGCTTCTAAAGGACATACATTTAATATTTCTGTGG

7990 8000 8010 8020 8030 8040
AAATATGAGGAAAATCCATGGTTATCTGAGATAGGAGATACAAACTTTGTAATTCTAATA

8050 8060 8070 8080 8090 8100
ATGCACTCAGTTTACTCTCTCCTCTACTAATTTCTGCTGAAAATAACACAAACAAAAT

8110 8120 8130 8140 8150 8160
GTAACAGGGGAAATTATATACCGTGAACCTAGGCTCCTACTTACATAGTTGAAAT

8170 8180 8190 8200 8210 8220
ATCAAGGAGGTCAGAGGAAATTGGACTGGTGAAAACAGAAAAACACTCCAGTCTGCCA

8230 8240
TATCACCACACAATAGGATCC

参考资料

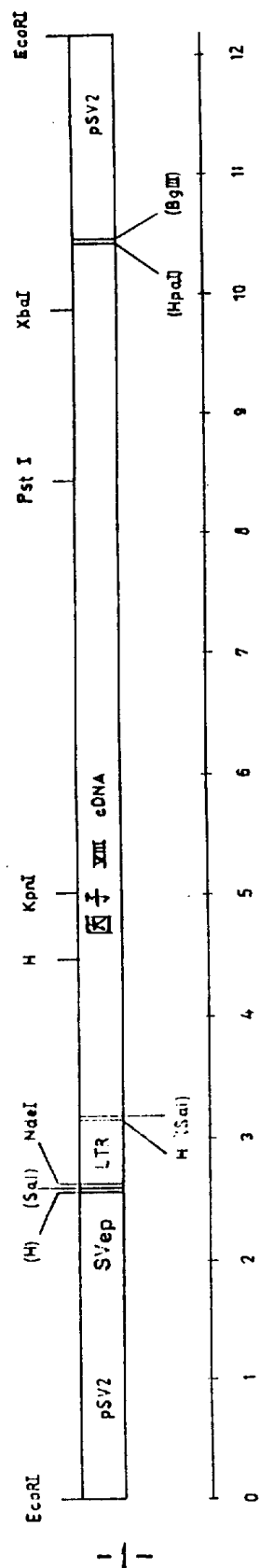
- Brinkhous, K.M. et al., (1985)
Proc. Natl. Acad. Sci. USA 82, 3752-3756
- Burke, R.L. et al., (1986)
J. Biol. Chem. 261, 12574-12578.
- Carter, P. et al., (1985)
Nucleic Acids Res. 13, 4431
- Chasin, L.A. and Urlaub, G. (1980)
Proc. Natl. Acad. Sci. USA 77, 4216-4220
- Eaton, D.L. et al., (1986a)
Biochemistry 25, 505-512
- Eaton, D.L. et al., (1986b)
Biochemistry 25, 8343-8347
- Fay, J.F. et al., (1986)
Biochim. Biophys. Acta 871, 268-278
- Galfré, G. et al., (1977)
Nature 266, 550-552
- Gluzman, Y., (1981)
Cell 23, 175-182
- Gorman, C., (1985)
DNA cloning (Ed.: D. Glover), IRL-press, Washington,
Oxford, pp. 143-169
- Graham, F. and Van der Eb, A. (1973)
Virology 52, 456-467

- Hanahan, D., (1982)
J. Mol. Biol. 166, 557-580
- Howley, P.M., Sarver, N. and Law, M.F. (1983)
Meth. Enzymol. 101, 337-402
- Hurwitz, D.R., Hodges, R., Drohan, W. and Sarver, H. (1987)
Nucl. Acids Res. 15, 7137-7153
- Kasper, C.K. et al., (1975)
Thrombs. Diath. Haemorrh. 34, 869-871
- Kaufman, R.J. and Sharp, P.A. (1982)
Mol. Cell. Biol. 2, 1304-1319
- Kaufman, R.J. and Sharp, P.A. (1982a)
J. Mol. Biol. 159, 601-621
- Kozak, M. (1983)
Microbiol. Rev. 47, 1-45
- Kramer, W. et al., (1984)
Nucleic Acids Res. 12, 9441-9456
- Laemmli, U.K., (1970)
Nature 227, 680-685
- Luthman, H. and Magnusson, G. (1983)
Nucleic Acids Res. 11, 1295
- Maniatis, T. et al., (1982)
Molecular Cloning. Cold Spring Harbor Laboratory

- Maxam, A.M. and Gilbert, W. (1980)
Methods Enzymol. 65, 499-560
- Melton, D.A. et al., (1984)
Nucleic Acids Res. 12, 7035-7056
- Mulligan, R. and Berg, P. (1981)
Proc. Natl. Acad. Sci. USA, 78, 2072
- Sanger, F. et al., (1977)
Proc. Natl. Acad. Sci. USA 74, 5463-5467
- Stel, H.V., (1984)
Ph.D. thesis, University of Amsterdam, The Netherlands
- Toole, J.T., et al., (1986)
Proc. Natl. Acad. Sci. USA 83, 5939-5942
- Van Mourik, J.A. and Mochtar, J.A., (1970)
Biochim. Biophys. Acta 221, 677-679
- Van Zonneveld, A.J., et al., (1986)
Proc. Natl. Acad. Sci USA 83, 4670-4674
- Vehar, G.A., et al., (1984)
Nature 312, 337-342
- Veldkamp, J.J. et al., (1968)
Thromb. Diathes Haemorrh. 19, 279
- Wood, W.I. et al., (1984)
Nature 312, 330-337
- Yanisch-Perron, C. et al., (1985)
Gene 33, 103-119

说明书附图

图.1



2.

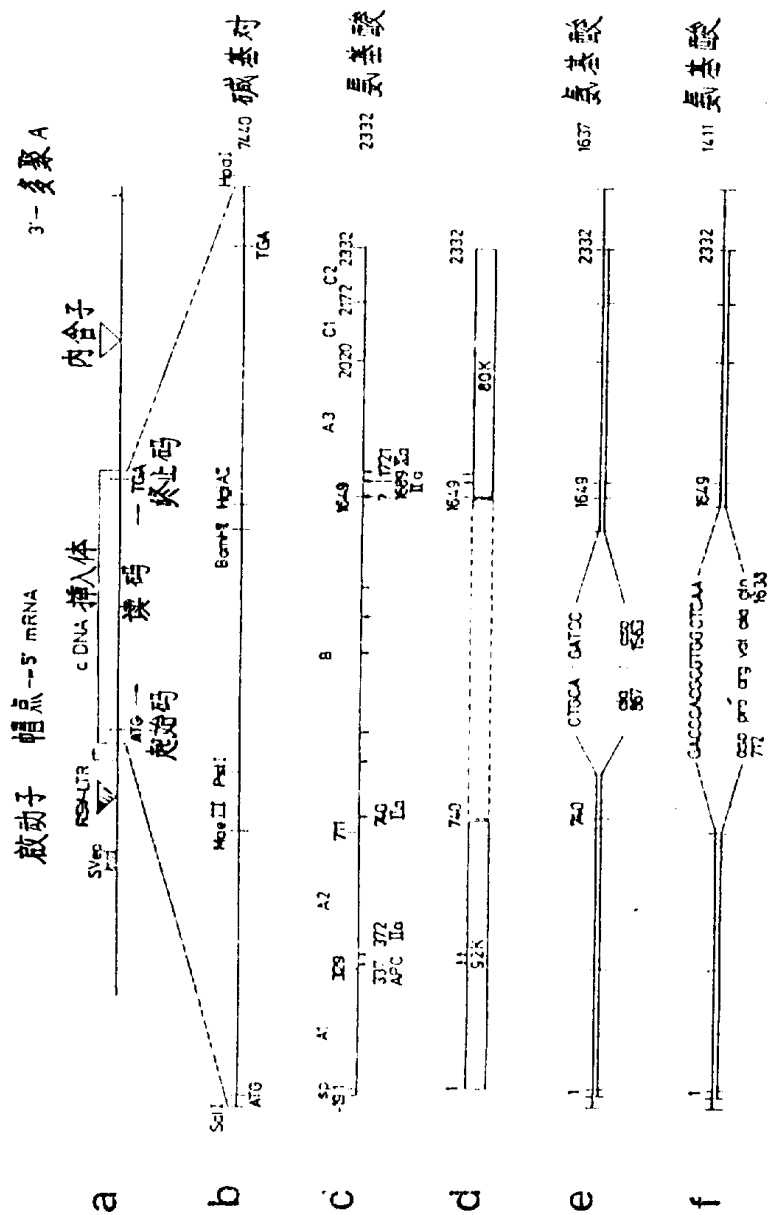


图. 3

1 2

$$M_r \times 10^{-3}$$

—200



—192



—168

—97







—68