

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷

C12N 1/21

C12N 15/09

C12P 25/00

[12] 发明专利说明书

[21] ZL 专利号 90104934.4

[45] 授权公告日 2001 年 5 月 30 日

[11] 授权公告号 CN 1066486C

[22] 申请日 1990.6.21 [24] 颁证日 2001.2.3

[21] 申请号 90104934.4

[30] 优先权

[32] 1989.6.22 [33] US [31] 370,378

[73] 专利权人 霍夫曼-拉罗奇有限公司

地址 瑞士巴塞尔

[72] 发明人 约翰·B·珀金斯

贾尼斯·G·佩罗 阿伦·施罗马

[56] 参考文献

FR2546907A 1984.12.7 C12N15/00

审查员 潘爱群

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 曹恒兴

权利要求书 3 页 说明书 76 页 附图页数 36 页

[54] 发明名称 高产核黄素的细菌菌株

[57] 摘要

本发明涉及能高产核黄素的细菌,本发明涉及 rib 操纵子的核苷酸序列和它的开放译读码,以及含有该 rib 操纵子的重组细菌。具体地说,本发明涉及这样一些细菌,这些细菌已被突变,从而使它们产生核黄素和/或嘌呤被失调,而且这些细菌中具有一些被插入在它们染色体 DNA 中、并被扩增的该 rib 操纵子的拷贝。在一个具体的实施例中,该 rib 操纵子本身可以通过用一些能产生组成型或失调表达的序列取代它的控制区域而被失调,通过发酵方法,可以将本发明的细菌、操纵子及序列用于超产大量的核黄素。

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

权 利 要 求 书

1. 一种制备重组细菌的方法，其中，通过一种包括一个编码了一种或多种核黄素生物合成蛋白的细菌来源或酵母来源的核酸序列组成的载体，转化一种细菌，从而至少一个所述核酸序列的拷贝被导入到它的染色体中的一个或多个位点上，而且是可遗传的，并能通过该细菌被表达，使得利用该细菌的核黄素生物合成与没有这种核酸序列的细菌相比，获得了提高。

2. 如权利要求1所述的方法，其中，所述载体还包括一种或多种不与所述细菌或酵母来源的核酸序列天然关联的转录元，并由此将至少一个所述的包括前述转录元的核酸序列的拷贝导入其染色体中的一个或多个位点上，所述的包含所述转录元的核酸序列是可以遗传的，并能在由该细菌表达时使得利用该细菌的核黄素生物合成与没有这种核酸序列的细菌相比，可以产生更多的核黄素。

3. 如权利要求1或2所述的方法，其中所述细菌来源的核酸序列是由芽孢杆菌属，特别是枯草杆菌或大肠杆菌获得的。

4. 如权利要求1—3中任何一个权利要求所述的方法，其中，所述的核酸序列编码了至少5种核黄素生物合成蛋白。

5. 如权利要求2—4中任何一个权利要求所述的方法，其中，所述的载体包含有一种或两种这样的转录元。

6. 如权利要求2—5中任何一个权利要求所述的方法，其中所述的转录元是一种启动子。

7. 如权利要求6所述的方法，其中所述的启动子选自于与一种来自 SPO1 噬菌体的基因天然相关的启动子；与 Veg, amy 天然关联的启动子以及 apr 和其它一些被 sacQ 基因产物活化的转录调节

元。

8. 如权利要求7所述的方法, 其中所述载体是 p R F 5 0, p R F 6 9, p R F 7 0, p R F 7 1, p R F 7 8, p R F 8 1 或 p R F 8 9。

9. 如权利要求1—8中任何一个权利要求所述的方法, 其中所述的核酸序列或包含所述转录元的核酸序列以多个拷贝的形式存在于这样一些位点上。

10. 如权利要求1—9中任何一个权利要求所述的方法, 其中所述的核酸序列或含有所述转录元的核酸序列已被导入所述染色体中的一个或两个这样的位点上。

11. 如权利要求1—10中任何一个权利要求所述的方法, 其中所述的细菌对核黄素基因表达是失调的。

12. 如权利要求1—11中任何一个权利要求所述的方法, 其中所述的细菌是大肠杆菌或芽孢杆菌、特别是枯草杆菌。

13. 如权利要求12所述的方法, 其中所述的枯草杆菌菌株是 R B 5 0 或 R B 5 8。

14. 一种生产核黄素的方法, 其特征在于将由权利要求1—13中任何一个权利要求所述的方法获得的重组细菌在合适的条件下生长。

15. 如权利要求14所述的方法, 其中, 所述的合适条件的特征在于限制该生长介质中的一种组分的存在量, 从而维持所述重组细菌生长所需的需氧条件。

16. 如权利要求15所述的方法, 其中所述的组分选自于碳源、氮源或一种所述重组细菌所需的一个组分。

17. 如权利要求16所述的方法,其中所述的组分是葡萄糖,或一种碳酸。

18. 如权利要求15所述的方法,其中所述的限制步骤包括限制所述组分在补充介质中的加入量。

19. 如权利要求18所述的方法,其中所述的组分是葡萄糖或一种碳酸。

20. 如权利要求14—19中任何一个权利要求所述的方法制备的核黄素。

高产核黄素的细菌菌株

核黄素(维生素B₂)能为所有植物和许多微生物所合成,但不能由高等动物产生。因为它是诸如黄素腺嘌呤二核苷酸和黄素单核苷酸这样的辅酶的前体,它们在碳水化合物的酶氧化作用中被需要,所以核黄素对基础代谢是必不可少的。在高等动物中,缺乏核黄素会引起脱发、皮肤发炎、视觉衰退和发育不良。

核黄素既可以从核糖开始通过完全化学合成,也可以用真菌假囊酵母菌属(*Eremothecium ashbyii*)或 *Ashbya gossypii* 的发酵来大量生产(The Merck Index, Windholz 等人, eds., Merck & Co., P. 1183, 1983)。已报导了通过与嘌呤类似物氮鸟嘌呤和氮杂黄嘌呤接触而选出的 *Bacillus Subtilis* (枯草杆菌)突变种以可回收的量产生核黄素(美国专利3,900,368, Enei 等人, 1975)。通常,与纯的核黄素或核黄素类似物的接触能选出显示出核黄素生物合成增加的失调突变体,因为这种突变使得这种微生物能够通过增加产量来“对抗超过”该类似物(Matsui 等人, Agric. Biol. Chem. 46:2003, 1982)。酿酒酵母(*Saccharomyces Cerevisiae*)的一个能产生核黄素的需嘌呤突变体也已被报导(美国专利4,794,081, Kawai 等人, 1988)。Rabinovich 等人(Genetika 14:1696(1978))报导了枯草杆菌的核黄素操纵子(rib 操纵子)被包含在一个7百万道尔顿(Md)的EcoRI片段内(后来在Chikindas 等人的Mol. Genet. Mik. Virusol. no.

2:20 (1987) 中被称为 6.3 Md 片段)。据报导,可以在 *E. Coli* 中放大 *rib* 操纵子,采用的办法是先将该操纵子克隆到对 α -氨基苄青霉素具有抗性的质粒中,再将含所述质粒的细菌暴露在大量的这种抗生素中。*rib* 操纵子扩增作用的唯一证据是在该介质中有绿色荧光物质存在时的一个相符合的增加,除了操纵子被真正扩增 (amplification) 外,作者还给出了许多别的可供选择的可能性来解释所观察到的现象。

Stepanov 等人的法国专利申请 2,546,907 (1984 年 12 月 7 日公开) 公开了一种生产核黄素的方法,该方法利用了与氮鸟嘌呤和玫瑰黄色素接触的枯草杆菌的突变体菌株,并且用含 *rib* 操纵子拷贝的质粒对它转化。

Morozov 等人 (Mol. Genet. Mik. Virusol. no. 7:42 (1984)) 通过检验枯草杆菌 *rib* 片段对 *E. Coli* 核黄素营养缺陷型的补充能力或者标记—补救 (rescue) 枯草杆菌营养缺陷型的能力来描述枯草杆菌 *rib* 操纵子的基因图。以 *E. Coli rib* 基因已知的功能为基础,对枯草杆菌提出了以下的模型: *ribG* (编码一个脱氨基酶) — *ribO* (控制元) — *ribB* (一个合成酶) — *ribF* — *ribA* (一个 GTP—环化水解酶) — *ribT* / *D* (分别为一个还原酶和异构酶) — *ribH* (一个合成酶)。

Morozov 等人 (Mol. Genet. Mik. Virusol. no 11:11 (1984)) 描述利用含具有野生型 (*ribO*⁺) 或结构性 (*ribO* 335) 操纵基因区域的枯草杆菌 *rib* 操纵子的质粒来检验其对枯草杆菌核黄素营养缺陷型的互补能力。根据此结果,提出了 *rib* 操纵子的一个修改模型。*ribO* 现在定位于所有结构基

因包括 ribG 的上游, 并且有另一个假设的操纵基因存在, 很可能就定位于 ribA 上游。

Morozov 等人 (Mol. Genet. Mik. Virusol. no 12:14 (1985)) 报导, 枯草杆菌 rib 操纵子含有全部三种不同的启动子 (除了一个只能在 E. Coli 中是活性的第四“启动子”外) 据报导操纵子的主要启动子定位于 ribO 区域中, 而两个辅助启动子据报导定位于 ribB 和 ribF 基因之间, 并分别在 ribTD 和 ribH 基因区域内。

Chikindas 等人 (Mol. Genet. Mik. Virusol. no 2:20 (1987)) 提出了一个含有枯草杆菌 rib 操纵子的 6.3 Md DNA 片段的限制酶图。指出了酶 EcoRI、PstI、SalI、EcoRV、PvuII 和 HindIII 的位点。

Chikindas 等人 (Mol. Genet. Mik. Virusol. no 4:22 (1987)) 报导, 枯草杆菌 rib 操纵子所有的结构基因都定位于 2.8 Md BglII — HindIII 片段上, BglII 位点位于操纵子的主要启动子和它的第一结构基因的核蛋白体结合位点之间。如下文所描述的, 申请人证明了该 BglII 位点实际上定位于 rib 操纵子 most -5' 开放译读码内使得所述 2.8 Md 片段不包含所有的 rib 结构基因。因此, 与 Chikindas 等人的报告不同, 1.3 Md BglII 片段并不包含在该第一结构基因的核糖核蛋白体结合位点内, 在该位点的插入导致一个核黄素阴性表型。因此, 为了提高表达能力, 使用该 BglII 位点来构建 rib 操纵子的任何努力, 例如通过用更强的启动子代替 5' 调节区, 实际上将破坏该第

一结构基因的完整性，因此也就破坏了操纵子的完整性。

Chikindas 等人 (Dokl . Akad . Nauk . 5 SSSR
298 : 997 (1988)) 公开了枯草杆菌 rib 操纵子的另
一个含有主要启动子 P_1 和两个次要启动子 P_2 和 P_3 的模型：

ribO(P_1) — ribG — ribB — P_2 — ribF — ribA — ribT
— ribD — P_3 — ribH。如前一样，1.3 Md BglII 片段含
有该操纵子整个的第一结构基因、以及在该主要调节区内的这个近端
的 BglII 部位的基因图都是不正确的报导。

本发明的目的是能大量生产核黄素的细菌。本发明涉及 rib 操
纵子及它的开放译读码的核苷酸序列，以及含有 rib 操纵子的重组
体细菌。具体地说，本发明针对的是已经突变使其核黄素和/或嘌呤
的生产失去调节的细菌，以及将该 rib 操纵子的拷贝插入并扩增到
其染色体 DNA 内的细菌。在一个实施例中，通过用一些允许构成性
表达或失调表达的序列代替它的控制区使 rib 操纵子本身失调。本
发明的细菌、操纵子和序列能用于通过发酵大量产生核黄素。

本发明通过下文中详细描述的一些具体实例的方式被说明。例如，
产生一个核黄素和嘌呤的生产失调、并使 rib 操纵子在其染色体内
被扩增的枯草杆菌 1A382 的突变体，RB50:: (pRF8)_{eo}
(Ade⁺)。该突变体在一个 14 升的容器内发酵 48 小时后，能产
生大于 5 g/l 的核黄素。还描述了别的一些细菌，利用这些细菌，
核黄素的产量在类似条件下提高到 10 g/l 以上。

本发明总的特征是通过构建各种细菌菌株并在适于产生核黄素的
条件下、在培养基内生长这些细菌菌株来大量 (超过 10 g/l) 生
产核黄素。在一个主要方面，本发明的特征是其染色体内至少含有一

种外源导入的核酸拷贝的重组体细菌。该核酸编码了一种或多种核黄素生物合成蛋白质，该核酸是可遗传的，并能用细菌来表达，使得相对于没有这个序列的细菌，由这种细菌生物合成的核黄素得到了提高。

所谓“重组体细菌”是指这样一种细菌，该细菌在其一个位点上含有来自同一种或另一种生物的一个或多个核酸序列，在该位点上，这些序列不会自然出现，和/或按照一个它们不会以此自然出现的拷贝数。因此，该术语包括了在通常只包括一个序列的位点上具有该核酸序列的两个拷贝、例如一个基因或一个操纵子的那些细菌。它还包括了这样一些细菌，其中已将一个核酸序列的一个或多个拷贝导入到通常不包括这种序列的位点上。这样一些重组体细菌是用标准的重组DNA技术构建的。

所谓“外源导入”是指通过任意一种标准技术，包括重组DNA技术、转化和转染，将核酸从染色体外的一个来源导入到该染色体中，它还包括这样一些细菌的子代，例如，通过原始构建、转化或转染细菌的细菌分裂产生的那些细菌。

所谓“核黄素生物合成蛋白质”是指直接参与从三磷酸鸟苷合成核黄素的那些肽、多肽或蛋白质。这些蛋白质可以与细菌内自然出现的、并参与细菌内核黄素合成的那些蛋白质是一样的，或者，它们可能是这样一些蛋白质的修饰物，例如，它们可以包含并不显著影响该蛋白质生物活性的修饰。例如，该天然蛋白质可由导入或替代一个或多个氨基酸来修饰，最好通过保守的氨基酸的替代，或通过去除该蛋白质非必需的区域。这样一些修饰很容易通过一些标准技术来完成。

在一些实施例中，这种细菌包含有该核酸序列的两个或多个拷贝；在该细菌染色体内的至少两个部位上存在编码了一种或多种核黄素生

物合成蛋白质的该核酸。

所谓“位点”是指一个相对于野生型细菌而言的不同的染色体位置，编码了生物合成蛋白的核酸定位于该位点上。例如，这种核酸可能定位于天然存在的编码了这样一些蛋白质（即在 rib 座位上）的基因的位点上，或者它可能定位于远离这一位置的一个位点上，这样一些远距离的位点最好从对重组体细菌不是必不可少的染色体核酸区中选取，例如像编码了对产生核黄素不是必需的蛋白质的区域，这些区的实例包括编码了某些胞外酶如蛋白酶的区域，在这样一些位点上的插入不会干扰所要求的特性或品质，只要细菌在核黄素生产方面的功能未受显著影响，任何位点都是适合的。

在另一些实施例中，该核酸在一个或多个位点上以多个拷贝的形式存在，并且核酸在染色体内的至少三个位点上存在。通过在不同的位点上导入该核酸，就可以提高该染色体该核酸拷贝的总数，提高了拷贝数，就可以提高核黄素产生的数量。

通常，核黄素生物合成蛋白质是通过一种或多种 rib 基因（例如，它的失活会产生核黄素营养缺陷型）、最好通过从图 3 提供的核苷酸序列中可辨识的至少 5 种不同的 rib 基因而被编码的。最好至少提供这样一些基因的 5 个拷贝。所谓“rib 基因”是指编码了天然存在于一个生物内的蛋白质或编码了执行与这些蛋白质类似功能蛋白质的那些基因或基因部分，这些基因或基因部分都涉及细菌内由三磷酸鸟苷向核黄素生物合成的转化。

在一个有关的方面，本发明的特征是重组体细菌，该细菌包括有编码了一种或多种核黄素生物合成蛋白质、例如图 4 中用 ORF 1 和 ORF 6 标记的基因产物的核酸，至少一种这样的蛋白质的表达是受

一个与该核酸不是天然关联的转录元控制的。或者说，该重组体细菌包含一种或多种 rib 基因或转录单元，它们的表达是受一个与该 rib 基因不是天然关联的转录元控制的。

所谓“转录元”是包括能实现（即开动）该转录元下游的核酸转录的任意一种核酸。这样一些元件的例子包括启动子和操纵子，这些转录元并不是与该核酸天然关联的，例如它们可能是异种的转录元。即它们可能由其它细菌种或属或者其它的生物体中分离出来的。换句话说，该转录元可能是一种天然存在于细菌中的、但通常不与现在转录有关的 rib 基因相关联的。它不包括天然与 rib 基因相关联的转录元。

在另一些实施例中，重组体细菌包括至少三种（或至少五种）rib 基因，并且所有这三种 rib 基因的表达是受一个与这些 rib 基因不是天然关联的转录元控制的，至少提供了两种转录元；这些 rib 基因被提供于重组体细菌的染色体内；该重组体细菌对核黄素基因的表达是失调的，而该转录元是一个启动子。例如，该启动子是一个组成性的、生长调节的、或可诱导的启动子，例如像与 S P O 1 噬菌体和/或 veg . amy 相关联的启动子，以及 sacQ 一敏感的启动子如 apr 。

所谓“失调”意指核黄素生产水平大于具有天然核黄素调节系统（即野生型细菌）的细菌中所观察到的生产水平。这样一些失调细菌的实例包括对各种嘌呤类似物或拮抗物有抗性的那些细菌，为此这些类似物或拮抗物可以例如选自于 8-氮鸟嘌呤、狭霉素 C、德夸菌素、8-氮杂黄嘌呤、磺胺胍（Sulfaguanine）、6-硫鸟嘌呤以及甲硫氨酸亚砷、和/或核黄素类似物如玫瑰黄色素。优选的失调细菌

对至少一种下列的嘌呤类似物或拮抗物 8-氮鸟嘌呤或德夸菌素和玫瑰黄色素具有抗性。

在另一些具体的实施例中，至少有一种 rib 基因包括一个不与 rib 基因天然关联的核蛋白体结合位点；该 rib 基因在该染色体的两个位点上存在；而且，该 rib 基因以多个拷贝的形式存在于该染色体。在一些更优选的实施例中，这种 rib 基因是杆菌属的 rib 基因，例如，图 4 中所示的 ORF 3 和 ORF 4，而转录元定位于 ORF 3 或 ORF 5 的 5'—上游区域；该 rib 基因是从 β -核黄素合成酶编码基因、ORF 2、ORF 3、ORF 4 和 ORF 5 中选取的；该细菌属于埃希氏菌属 (*Escherichia*) 的种例如 *E. Coli*，芽孢杆菌属例如枯草杆菌，克雷伯氏菌属 (*Klebsiella*) 或棒杆菌属 (*Corynebacterium*)。

在另一个有关的方面，本发明的特征是这样的核酸，它包含 5 种或更多种的、其表达受一个与其不是自然关联的转录元控制的 rib 基因。

在另外一个方面，本发明涉及这样一个载体，它首先包含细菌，特别是以枯草杆菌 (*B. subtilis*) 为最好的杆菌 (*Bacillus*) 属细菌、或者是 *E. Coli* 的核酸序列，或者编码了一种或多种 (最好 5 种) 核黄素生物合成蛋白质的酵母源的核酸序列。其次包括是一种或多种 (最好一种或两种) 不与该核酸序列天然关联的转录元。具有前面指出的转录元的这样一些载体是最好的，在实例中指出的那些载体例如 pRF 50、pRF 69、pRF 70、pRF 71，pRF 78，pRF 81 和/或 pRF 89 是更优选的载体。

本发明的另一个方面涉及这样一种重组体细菌，它包含的细菌已

由上面规定的载体转化，因此包括所说转录元的所说的核酸序列的至少一个拷贝，最好多个拷贝被引入该染色体中的一个或多个位点上。包含的所说转录元的所说的核酸序列是可以遗传的并能用该细菌表达，使得由该细菌进行的核黄素合成与缺少包含所述转录元的这种核酸序列的细菌相比，得以提高。在一个或两个这样的部位上插入最好，而且对于核黄素基因表达是失调的被转化细菌是更优选的重组体细菌，在这样一些失去调节的细菌中，*E. Coli* 或 *Bacillus* 属的，特别是 *B. Subtilis* 菌株是优选的，其中 *B. Subtilis* 菌株 RB 50 和 RB 58 是最优选的。

本发明还涉及一种制备这种重组体细菌的方法，其中尤其是前面已指出的一种细菌由一种载体转化，该载体包含有编码了一个或多个核黄素生物合成蛋白质的细菌或酵母源的核酸序列，或如前面具体规定的载体，其中所说的包括或不包括所说转录元的核酸序列的至少一个拷贝、最好多个拷贝被导入到该染色体内的一个或多个，最好是一个或两个位点上。并且，包括或不包括所述转录元的所述核酸序列是可遗传的并能用该细菌表达。使得由该细菌进行的核黄素合成与缺少包含或不包含所述转录元的这种核酸顺序的细菌相比得到了提高。

本发明还涉及如本文中具体规定的一些细菌，在这些细菌中，一个或多个与在这样一些细菌的染色体中的 *rib* 生物合成基因天然相关联的转录元已由一些具体规定的转录元代替，根据本发明的描述和采用该技术中已知方法能作出这种置换。

在另一方面，本发明的特征是生产核黄素的方法或工艺过程，该法包括在适当的生长条件下培养细胞，特别是本文特别指明的那些重组体细菌。这样一些适当的生长条件的特征在于以一定的方式限制生

长介质和/或进料介质一种成分的利用率，使得能够维持适合于所述重组体细菌生长的需氧条件。这样一些条件也可以表征为例如将溶解氧的水平维持在大约5%—30%的浓度上。所属技术领域的专业人员知道这样的溶解氧水平会随用于培养所述重组体细菌和用于测量所述溶解氧浓度的特定的技术装备而改变。在厌氧条件下，核黄素的合成要下降，在一些实施例中，该限制成份是从碳源、氮源或一个细胞所需要的成分（例如在进料介质）中选取的。假如这些细胞是营养缺陷型的例如，对于甲硫氨酸（methionine），可以在该生长介质中提供一个限制的甲硫氨酸含量。在另一实例中，这样一个成份可以是葡萄糖或碳酸，例如柠檬酸循环酸，如柠檬酸或琥珀酸，或氨基酸。

在一个有关的方面，本发明的特征是另一个提高细菌生产核黄素的方法。在该方法中，所用的细菌菌株对于核黄素生成是失去调节的。一种以上的编码了一种或多种核黄素生物合成蛋白质的核酸序列的拷贝被引入这种细菌的染色体DNA中。用于本方法的细菌最好是选自前面描述过的细菌中的一种。本发明还涉及一种通过在如上具体指出的条件下培养如上面规定的一种方法获得的重组体细菌来生成核黄素的方法。

在本发明的另一方面，提供了纯化的核酸和这样一些核酸的重组体多肽产物。通常，纯化的核酸基本上是由该rib操纵子所有的或一部分，例如图3所示的具体的开放译读码组成的。这种纯化的核酸可以在诸如质粒、噬菌体或cosmid这样的一种载体内被提供，或者被整合在一种细菌的染色体内。将这种核酸从与它天然相联结的核酸中分离出来，例如，6.5 kb的编码了整个rib操纵子的核酸可被插入枯草杆菌染色体的一个位点上，该位点远离通常存在的6.5

kb DNA位点, 所谓重组体多肽是指具有与这样一个天然产生的多肽的酶活性相同的无外源多肽(即未融合成一个同种多肽)的生物活性蛋白质。

本发明的其他一些特征和优点根据下面对其最佳实施例的描述和权利要求书将变得更加清楚。

首先简单地描述附图:

图1. 核黄素生物合成的途径, 根据Keller 等人, Biochem 27:1117 (1988) 修改的。所表示的相应的中间产物是由 E. Coli 产生的中间产物(它们可能与 B. Subtilis 产生的相同): 结构1, 三磷酸鸟苷(GTP); 结构2, 2, 5-二氨基-6-(核糖基氨基)-4(3H)-嘧啶酮-5'-磷酸盐; 结构3, 5-氨基-6-(核糖基氨基)-2, 4(1H, 3H)-嘧啶二酮-5'-磷酸盐; 结构4, 5-氨基-6-(核糖基氨基)-2, 4(1H, 3H)-嘧啶二酮-5'-磷酸盐; 结构5, 6, 7-二甲基-8-核糖醇 2, 4-二氧四氢喋啶; 结构6, 核黄素。所指出的生物合成酶是由 B. subtilis 编码的那些酶(GTP环化水解酶、核黄素合成酶的 α 和 β 亚单位), 或者是被提议的由 B. subtilis 编码的那些酶(rib-特异脱氨酶、和rib-特异还原酶)。

图2, 示意表示了嘌呤的生物合成, 描绘了嘌呤的生物合成途径。包括决定核黄素生物合成的部份。途径中的各个酶是由它们的基因符号标示的(E. Coli命名法)。这些缩写如下: PRPP; 磷酸核糖焦磷酸; GAR: 甘氨酸核糖核苷酸; pur: GAR甲酰转移酶; PRA: 磷酸核糖基胺; purA: 腺苷酸琥珀酸合成酶; purB: 腺苷酸琥珀酸合成酶; FGAR: 甲酰基甘氨酸核糖核苷酸

SAICAR: 氨基咪唑琥珀羧基酰胺核糖核苷酸, purC: SAICAR合成酶, FGAM: 甲酰甘氨酸核糖核苷酸, purD: GAR合成酶, AIR, 氨基咪唑核糖核苷酸, purE: AIR羧基酶, CAIR: 羧基氨基咪唑核糖核苷酸, purF: PRPP酰胺转移酶, AICAR: 氨基咪唑羧基酰胺核糖核苷酸, purH: AICAR甲酰转移酶, purJ: 次黄苷单磷酸(IMP)环化水解酶, FAICAR: 甲酰氨基甲酰咪唑核糖核苷酸, purL: FGAR酰胺转移酶, guaA: 鸟苷单磷酸(GMP)合成酶, purM: AIR合成酶, guaB: IMP脱氢酶。

图3, *B. subtilis* rib基因完整的核苷酸和推断出的氨基酸序列。该核苷酸序列由M13克隆的二脱氢定序法测定的, 该推断出的氨基酸序列由一个字母的编码(Lehninger, Biochemistry, 第二版, Worth Publishers, Inc. New York, P. 72)指示。

图4. rib基因簇的一个示意表示, 上图是在质粒pRF2中克隆的10 kb EcoRI DNA片段的限制性核酸内切酶的图, 含有*B. subtilis* rib操纵子。对核黄素合成酶基因特异的54-mer探针的同系物由实短线描绘。画阴影线的盒子部分描绘了Rib⁺克隆的DNA, 而细黑线则表示了pBR322 DNA, 下图以6.0 kb片段的全部核苷酸顺序为基础, rib操纵子被定位于该片段上。开放译读码由开口的盒子描述, 箭头指示转录的方向。而封闭的盒子指示假想的核蛋白体结合位点。画出了可能的 σ^A 启动子区域。用“发夹”符号指示由试验鉴定的rho独立转录的终止部位。

图5. RB50的菌株谱系。描绘出*B. subtilis*的核黄素

高产菌株 RB 50 的谱系。各种亲代菌株与核黄素和嘌呤类似物不以使选出合适的突变。

图 6. rib^+ 重组质粒的起原。给出了产生含该 rib 操纵子的重组体质粒 p R F 1, p R F 2, p R F 3, p R F 6 和 p R F 7 的示意, 枯草杆菌 DNA 的大小被选择在 9—11 kb 的片段的基因库被用于在 E. Coli: 质粒载体内产生一个基因库, 通过对核黄素合成酶基因的 β 亚单位特异的 54-mer 探针的杂交选择克隆。

图 7. 枯草杆菌 RB 53 :: (p R F 8)₉₀ 的菌株谱系。质粒 p R F 8 被整合到中间产物菌株 RB 52 的染色体中并被放大; 所得到的菌株与嘌呤类似物氮鸟嘌呤接触。

图 8. 使用插入和缺失鉴别对核黄素生物合成必须的那些区域, 图上表示了具有被标示的对核黄素生物合成必须区域的 10 kb 克隆的 EcoRI DNA 片段, 在指示出的限制性位点上的插入和缺失能实现 rib 操纵子的定位, 未标示出所有的限制性位点。

图 9. 可能的 rho 独立转录终止位点的发夹式的回路结构, 它们在图 3 的核苷酸序列中的位置被表示在每一结构的下边。还给出了它们的按照 Tinoco 等人的测定 (Nature (London) New Biology 246:40 (1973)) 的形成自由能。

图 10. 用于 S—30 体外偶联转录/转译反应中的各种质粒衍生物的结构。示意图画出了包含在用于 S—30 反应中的质粒衍生物中的 rib 操纵子区域的示意图, 以及预计要被表达的开放译读码。

图 11. 核黄素生产的对比曲线, 画出了各种发酵方案的核黄素生产曲线。空的正方形: RB F—14 利用 RB 50 :: (p R F 8)₉₀ (Ade⁻)。黑的方形: RB F—22 利用 RB50

$:: \{ pRF8 \}_{00} (Ad e^{-})$

空的园形: R B F-23 利用 R B 50 $:: \{ pRF8 \}_{00} (Ad e^{-})$,

黑的园形: R B F-29 利用 R B 50 $:: \{ pRF9 \}_{00} (Ad e^{-})$,

图 12. p R F 4 0 的构建.

图 13. p R F 5 0 的构建

图 14, 15 和 16. 各种载体的结构.

图 17. 用于质粒构建的 55-mer

图 18. 用于载体构建的各种寡核苷酸.

在本发明的实施中, 衍生出的宿主细菌株在核黄素生物合成途径的基因中或在各种嘌呤的生物合成途径中含有一种或多种突变, 这些突变导致了核黄素的高产。在一个实施方案中这样一些突变在核黄素生物合成途径中通过失调措施导致核黄素的高产。在另一个实施方案中, 这些突变通过抑制核黄素生物合成中一个前体的一个可能的代谢途径而提高核黄素的产量。

在一个具体的实施例中, 可以通过与嘌呤类似物或核黄素的接触来诱导宿主细菌遗传背景的所希望的突变, 这些突变的宿主细菌会与它们真正的对手竞争。在这种接触中存活的细菌将具有这样一些突变, 这些突变能使它们在产生嘌呤或核黄素类似物上超过它的亲本, 于是, 就“对抗超出”了否则对它是致命的嘌呤或核黄素类似物。核黄素在 *B. subtilis* 中的生物合成以三磷酸鸟苷开始 (图 1, 结构 1), 三磷酸鸟苷 (GTP), 借助鸟苷单磷酸 (GMP), 是嘌呤生物合成途径的一种产物 (图 2)。在一最佳实施例中, 为了获得能高产核黄素的宿主菌株, 人们能运用经典的遗传学方法来增加细胞产生的 GTP 量, 同时使核黄核途径失调, 通过获得对嘌呤类似物或拮抗物

具有抗性的突变体，能够实现枯草杆菌中嘌呤的高产。能够使用的一些嘌呤类似物实例包含但不限于 8-氮鸟嘌呤 (Ishii and Shio, Agric. Biol. Chem. 36:1511, 1972; Konishi and Shiro, Agric. Biol. Chem. 32:396, 1968)、狭霉素 C 和德夸菌素 (Matsui 等人, Agric. Biol. Chem. 43:1739, 1979; Matsui 等人, Agric. Biol. Chem. 43:393, 1979)、8-氮杂黄嘌呤、磺胺鸟嘌呤 (sulfaguanine)、6-硫代鸟嘌呤 (Debabov. V. G. in The Molecular Biology of the Bacilli vol. 1 Bacillus subtilis, D. A. Dubnau, ed. (Academic Press, New York), pp. 331—370, 1982) 和其它, 和/或拮抗物甲硫氨酸亚砷 (Matsui 等人, App. Env. Microbiol. 34:337, 1977), 以及它们任意组合。

核黄素途径由于获得的对核黄素类似物具有抗性的突变而失调。一个可应用的核黄素类似物的实例是玫瑰黄色素 (Matsui 等人, Agric. Biol. Chem. 46:2003, 1982)。

在本发明一个具体的实施例中, 可以使用因突变而对氮鸟嘌呤、德夸菌素和玫瑰黄色素这些类似物有抗性的细菌, 对这些化合物中的每一个具有抗性的具体突变株将在下文予以描述。也可以利用具有使之对其它类似物有抗性的突变细菌, 利用具有使其对这些相同的类似物具有抗性的不同突变的细菌, 或者利用这些突变的不同组合, 任何一种组合方式, 无论是在有或没有对其它类似物的各种突变下的组合

都被认为是在本发明的范围内。

假若只与该类似物的接触不能以足够高的发生率产生抗性突变体时，可以利用各种诱变因素来提高总的突变发生率，因此增加类似物——抗性突变株的数目。作为一个实例，可以使用乙基甲基磺酸盐，但是也可以利用包括但不限于亚硝基胍或紫外线照射的其它的突变因素。

合适的细菌宿主包括所有的杆菌菌种（包括在一最佳实施例中的枯草杆菌 *B. subtilis*）、*E. coli* 和许多其它的革兰氏阳性和革兰氏阴性细菌。能够识别要插入其基因组内的克隆 *rib* 操纵子的启动子顺序的菌种也都适合使用，以下描述的一些质粒可用于通过标准程序，例如转代将 *rib* 基因导入其它的细菌内。插入的 *rib* 基因的表达可以通过以下描述的光谱学方法，或者通过如下所述的在紫外光下观察细菌来测定。

除了通过与嘌呤或核黄素类似物接触产生突变外，可以利用已含有对其嘌呤或核黄素生物合成途径有影响的突变的细菌菌株。例如，本发明能利用但不局限于枯草杆菌 *B. Subtilis* 菌株 1 A 382，该菌株包含有 *pur-60* 的突变，使它对腺嘌呤成为营养缺陷型。因这种突变阻断了在一个与核黄素产生不同的代谢途径中利用核黄素母体次黄苷单磷酸（*IMP*），所以有更多量的 *IMP* 可供核黄素生物合成利用，因此提高了核黄素的产量。有许多其它突变可用于提高核黄素的产量，它们包括但不限于 *guaC3*、*his⁻* 和其它，它们也被包括在本发明范围内。*guaC3* 突变能防止 *GMP* 反向转化成 *IMP*（参见图 2）因此提高了可利用的核黄素生物合成的母体量。

影响核黄素高产的各种合适的突变可采用多种该领域内已知的方法

法进行绘制。在一个具体的实施例中，突变可用营养缺陷型突变体的互补作用来绘制。

由各种细菌得到的核黄素生物合成基因可被克隆以用于本发明中，选自多种菌种的酵母或细菌细胞，包括但不限于杆菌属 *Bacillus*、*E. Coli* 属和许多其它的革兰氏阳性和革兰氏阴性细菌，有可能用作 *rib* 操纵小分子克隆的核酸源，含有 *rib* 操纵子的 DNA 可以例如由一个 DNA 基因库通过在该领域中已知的标准程序得到，该基因库通过将由所需的菌细胞中提纯的染色体 DNA 或其片断克隆到一个适当的用于增殖该基因的载体中制备的。（参见例如 Maniatis et al., 1982, *Molecular Cloning*, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York; Glover, D. M. (ed), 1982, *DNA Cloning: A Practical Approach*, MRL Press, Ltd., Oxford, U. K., Vol. 1, II)。

在分子克隆染色体 DNA 的基因时，产生出一些片段，其中有些编码了所希望的 *rib* 操纵子。利用各种限制性酶可以在特异的位点上切开。作为另一种方法，人们可以在有锰存在的情况下用 DNA 酶来使 DNA 成为片段，或者 DNA 可以例如通过声处理被物理切割，然后可以通过一些标准技术包括但不限于琼脂糖和聚丙烯酰胺凝胶电泳和密度梯度离心法使线性 DNA 片段按大小被分离。

一旦产生了这些 DNA 片段，利用适当的克隆和/或表达载体来制备 DNA 基因库。可以利用该技术领域中已知的大量载体—宿主体系。可能的载体包括但不限于质粒或改性过的病毒。然而该载体系统

必须与所用的宿主细胞是相容的。对于 *E. Coli*, 这样的载体包括但不限于噬菌体如 λ 衍生物, 像 pBR322 或 pUC 质粒这样的高一拷贝质粒, 或者由假单胞菌属质粒 RK2 衍生的低一拷贝质粒。对于 *Bacillus* 杆菌属这样的载体包括但不限于噬菌体如 ϕ 11 (Dean et al., *J. Virol.* 20: 339, 1976; Kawamura et al., *Gene* 5: 87, 1979) 或 Δ 105 衍生物 (Iijima et al., *Gene* 9: 115, 1980; Errington, J. *Gen. Microbiology* 130: 2615, 1984; Dhaese et al., *Gene* 32: 181, 1984; Errington, J. in *Bacillus Molecular Biology and Biotechnology Applications*, A. T. Ganesan and J. A. Hoch, eds (Academic Press, New York) P. 217 1986), 高一拷贝质粒如 pUB 110 (Ehrlich, *Proc. Natl. Acad. Sci (USA)* 74: 1680, 1977) 或 pBD64, 或低一拷贝质粒如 pE194 衍生物 (Gryozan, T. J. in the *Molecular Biology of the Bacilli*, D. A. Dubnau, ed. (Academic Press, New York), pp 307-329, 1982; Horinouchi, and Weisblum, *J. Bacteriol.* 150: 804, 1982)。重组体分子可通过转化、转染、原生质体化, 感染和电击穿等导入宿主细胞内。

一旦产生了该 DNA 基因库, 鉴别包藏了含 rib 操纵系的重组体 DNA 的特异克隆可用许多方法来完成 (例如按 Maniatis 等人同文中所述的)。例如, 假若由另外的细胞源 (例如从 *E. Coli*)

可获得一定量的操纵子或其片段并且它们与杆菌属 (*Bacillus*) 的核黄素生物合成基因同系到足以与其杂交, 则该 DNA 可以被纯化和标记, 并且可通过将核酸与该标记的探针杂交来筛选所产生的 DNA 片段的基因组 (Benton , M . and Davis , R . , 1977 , Science 196 : 180 , Grunstein . M . and Hogness , D . , 1975 , Proc , Natl , Acad . Sci . U . S . A . 72 : 3961) 。作为另一种选择, 包含内源性 rib 操纵子的开放译读码的序列, 或其包含约 10 个, 较好为 15 或更多核苷酸的序列。可被用作杂交探针, 这样一些探针能用合成法制作, 根据该核酸的一部份或一个基因产物的已知被该操纵子逆转遗传编码的氨基酸序列 (其实例下文中给出) 。假若得不到一个纯的 rib 操纵子—特异探针, 限制性片段 (例如由部份 Sau 3 A—消化得到的) 的克隆基因库就可在细菌, 特别是枯草杆菌或大肠杆菌中制备, 并且, 含有 rib 操纵子的重组克隆可以由一些已知 rib 突变的标记—补救或互补作用来鉴别。

在一最佳实施例中, 可从枯草杆菌 DNA 的大肠杆菌质粒基因库分离出枯草杆菌的 rib 操纵子供使用。具体地如下所述, 枯草杆菌的 rib 操纵子借助它与放射标记的合成核苷酸探针是同系的被分离, 该探针是由一个已知被枯草杆菌操纵子编码的基因产物的一个内部区衍生得到的。虽然 β -核黄素合成酶 (Ludwig 等人, J. Biol. Chem . 262 : 1016 , 1987) 氨基酸序列的一部份可以是这种探针的基础, 根据密码子的使用频率估计是每个密码子的第三个核苷酸, 也可以利用一个以该 rib 操纵子的这种蛋白质或另一种蛋白质的别的区域为基础的探针。这也属于本发明的范围。本发明还通

过使用由图 3 所示核酸序列衍生的合成探查针来实现筛选。

可以使用与本文详细说明的方法相类似的方法来分离其它细菌，尤其是其它的 *Bacilli* 属或 *E. Coli* 属的 *rib* 操纵子。在一个具体的实施例中，这样一些克隆能够通过检验与标记的枯草杆菌 *rib* 操纵子或其一个可杂交部份杂交的能力被挑选出来，在 *cDNA* 能够从适当的 *mRNA* 制剂开始制备是本技术领域中所周知的。这样的 *cDNA* 也可以按本发明来制备用于转化合适的核黄素高产细菌的载体。

一旦鉴别出具有包含被分离的 *rib* 操纵子或其一部份的重组 *DNA* 分子的宿主细胞，就可大量获得该 *DNA*。这样就可以采用该领域专业人员熟悉的各种克隆方法和定序方法来处理该 *rib* 操纵子和确定它的核苷酸序列。

例如，插入致突变能用于定位和表征在一个 *DNA* 克隆片段中的 *rib* 操纵子和它的基因，在一个具体的实施例中鉴别含有 *rib* 生物合成区域的方法是将一个小的含 *cat*（氯霉素乙酰转移酶）的限制性片段插入该克隆 *DNA* 的一些不同的限制性酶位点上，并在一个合适的宿主中检验每一衍生物的核黄素生物合成的插入失活作用（参见下文）。

相应于 *rib* 操纵子的克隆 *DNA* 可以用一些方法来分析，这些方法包括但不限于 Southern 杂交法（Southern, E. M., 1975. *J. Mol. Biol.* 98:503-517），Northern 杂交法（例如参见 Freeman 等人., 1983, *Proc. Natl. Acad. Sci. U. S. A.* 80:4094-4098），限制性核酸内切酶作图法（Maniatis 等, 1982,

Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) 和 DNA 定序分析, 限制性核酸内切酶作图法可用于粗略地确定 rib 操纵子的基因结构, 由限制性核酸内切酶切割而得到的限制图可以由 DNA 定序分析来证实。

DNA 定序分析可由该技术领域中已知的任意一个方法来完成, 包括但不限于 Maxam 和 Gilbert 的方法 (1980, Meth. Enzymol. 65 : 499—560). Sanger 的二脱氧法 (Sanger. F. 等人, 1977. Proc. Natl. Acad. Sci. U. S. A. 74 : 5463), 或利用一个自动 DNA 定序仪 (例如, Applied Biosystems, Foster City, CA), 作为一个实例, 图 3 给出了枯草杆菌 rib 操纵子的 DNA 序列。

一旦确定了 rib 操纵子的核苷酸序列, 那么假定的开放译读码 (ORFs) 则可与编码产物的被推定的氨基酸序列一道被鉴定。编码产物的实际鉴定例如可以通过用各种 ORFs 作为模板进行体外 S-30 偶联转录/转化反应来完成。也可测试 ORFs 的各种突变衍生物在 S-30 反应产物功能检定中的活性, 以测试该编码产物的功能。

在本发明涉及枯草杆菌 rib 操纵子的一个在下文中将以实例详述的具体实例中利用了上面这些方法来确定枯草杆菌的核黄素生物合成是受单独一个含有以下 5 种生物合成基因的接近 4.2 kb 的操纵子控制的, 这 5 种生物合成基因是核黄素合成酶的 β 亚单位和命名为 2, 3, 4 和 5 的 ORFs (见图 4)。ORFs 2, 3, 4 和 5 依

次被表明分别编码了分子量为约 15 kd、47 kd、26 kd 和 44 kd 的蛋白质，如上所述，ORF5 被表明编码了一个假定的 rib- 特异脱氨基酶，该酶在核黄素生物合成的最初期阶段中催化脱氨基嘧啶还原成一个核糖基氨基链。我们的数据还指出 ORF4 编码了核黄素合成酶的 α 亚单位以及 ORF3 编码了一个 GTP 环化水解酶，而 ORF2 有可能编码了一个 rib- 特异还原酶，发现 ORF1 和 ORF6 是在 rib 操纵子的基本转录单位的外边。rib 操纵子启动转录的基本部位被确定，是可能定位于自该操纵子的第一个基因 ORF5 起上游 290 bp 的表面 σ^A 启动子 (图 4, p₁)。枯草杆菌 rib 操纵子的编码区。启动子和转录终止部位被表示于下面的表 VI 中。

本发明包括 rib 操纵子基因的核苷酸和氨基酸序列，及它的编码了功能活性肽的亚顺序和基本上与这样一些序列一样的序列。本文所用的功能活性肽是指能催化一个导致核黄素生物合成反应的蛋白质或肽。功能活性核酸序列是指一个能够调节核黄素生物合成的序列。一个基本上与另一个序列相同的序列是指一个能与其互补序列杂交的序列。此外，一个核酸顺序不能自然控制一个第二核酸序列表达的核酸序列是指一个不能控制该第二序列在一种分离出该第二序列的细菌中表达的序列。

一旦 rib 操纵子的遗传结构被了解，就可能对该控制结构进行处理以便最佳地用于本发明中。例如，能构建出核黄素产量最高的 rib 操纵子。

根据所使用的宿主载体体系，可以应用许多合适的转录和转译元中的任何一个由重组体 DNA 或合成技术产生的启动子也可用于保证该插入序列的转录，当在细菌中繁殖时，可以使用该 rib 操纵子本

身的调节序列。在希望以多顺反子信使形式表达完整的rib操纵子或多于它的一个基因的实施例中需要一个原核生物宿主。在一个用真核生物宿主的实施例中，必须在重组DNA每个希望被表达的基因/ORF的上游放置还当的调节序列，例如启动子。

需要特定的起始信号用于有效的转译插入的蛋白编码序列。这些信号包括起始密码子(ATG、GTG或TTG)和邻近的序列，例如核蛋白体结合部位(RBS)。应当指出，可以处理一个给定编码序列的RBS以实现该编码序列以转译水平更加有效地表达。在rib操纵子的整个开放译读码、包括其自己的起始密码子和邻近的调节序列被插入合适的表达载体中的情况下，不需要另外的转译控制信号。然而，在只插入该编码序列的一部份、或者其天然调节信号不能为宿主细胞识别的情况下，必须提供外源转译控制信号，包括起始密码子。该起始密码子还必须与该蛋白编码序列的译读码同相，以保证整个插入物的转译。这些外源转译控制信号和起始密码子可以是各种来源的，天然和合成来源的。

另外，选择的宿主细胞菌株能调整一种或多种rib操纵子基因的表达或以所希望的具体型式来修饰和处理它的一种或多种基因产物，在有某些诱导剂存在的情况下可以提高由一些启动子产生的表达；因此可以控制用遗传方法建造的rib操纵子蛋白质的表达。在一个实施例中，操纵子的一些调节区如启动子和终止/抗终止调节序列，能初处理，或用构成性启动子或生长调节启动子操纵或代替，使rib操纵子失调，因此提高了核黄素产量。此外，可以选择合适的细胞系或宿主体系以保证被表达蛋白质的所希望的修饰和处理。可以进行许多的处理并且它们是在本发明的范围内。

本发明一个具体的实施例中，枯草杆菌 *rib* 操纵子的 5' 调节序列可被去除并用几个枯草杆菌启动子中的一个或多个代替，这样一个构建将引起 *rib* 生物合成基因的高水平表达。这一方法涉及将一些新的限制性位点导入到转录终止子终端和操纵子 O R F 5 的第一个基因的 R B S 序列之间 20—30 bp 区。导入这样一些限制部位既可通过位点定向致突变，也可借助于缺失自定位于 O R F 5 第一个 30 bp 内最右的 Bgl II (Bgl II_R) 位点上游的所有调节序列 (见图 3 和 4) 并在该位点插入能结束 O R F 5 的 5' 终端 (包括核蛋白体—结合位点) 并含有新的上游限制位点的合成寡核苷酸。一旦作出这些构建，就可导入具有与新的限制性位点相容的终端包含了启动子的限制性片段。引起 *rib* 基因在新的启动子的控制下的表达。既可以使用组成性的也可以使用生长调节性的枯草杆菌启动子。包括但不限于由溶解噬菌体 S P O 1 基因、Veg、amy (淀粉酶) 和 apr (枯草溶菌素) 得到的强启动子。

本发明的另一方面，可以用具有转录调节活性的 *rib* 操纵子 D N A 片段 (例如，启动子) 的 *rib* 操纵子 D N A 片段 (例如，启动子) 来调节异种基因产物的表达。

按照本发明，可以将 *rib* 操纵子导入表达它的细菌中，包括例如芽孢杆菌和大肠杆菌。在一个优选的实施例中，细菌宿主是上述突变体宿主中的一种。在一个具体的实施例中，克隆的 *rib* 操纵子被整合到宿主染色体的 D N A 中，它在那里与宿主基因组 D N A 一道被复制和表达。在一个最优选的实施例中，*rib* 操纵子的许多拷贝被整合到宿主染色体 D N A 中，因此在该失调的宿主中提供了 *rib* 操纵子的放大表达。能完成这一点的方法是先在染色体中插入含 Cat

的 rib 操纵子，接着进行操纵子的氯霉素放大，在下文实例部份将被详细描述。采用这种相同的技术，人们还可使用 tet^r 基因或能在芽孢杆菌中被表达的某些其它药品的抗性基因。

在一些具体的实施例中，能够构建出这样一些含 rib 操纵子片段的整合载体，以便可以在最小的可能 DNA 片段上包含该 rib 操纵子，企图在宿主染色体内获得更大的这种载体的放大。例如，可以缺失一些载体 DNA 顺序和/或在 rib 操纵子侧面的非必需的 DNA。

通常，对核黄素是原始营养的细菌能在没有核黄素存在的最低培养基内存活，可用各种方法检测和定量测定核黄素的产生。在一优选的实施例中，当高产细菌如下文所述被暴露于 366 nm 的紫外光下时产生了一种可见的黄色荧光，很容易观察到高产的核黄素。例如，许多本发明构建的质粒在 *E. Coli* 中产生。对于这样一些质粒，核黄素的高产量已为本方法所证实，所产生的核黄素的量可以例如，用反相高效液相色谱 (HPLC) 来定量测定。如下文所述，由细菌得到的无细胞的上清液可在 HPLC 柱上分成几个组份，并在 254 nm 下监测核黄素。根据标准曲线外推，可由色谱上峰的面积来确定核黄素的浓度。

核黄素也可用荧光分光光度测定法来定量测定。例如，含核黄素的样品可在发射波长被设定在 525 nm 和激发波长为 450 nm 的荧光分光光度计读出数据。

此外，还有其它一些在该领域中已知的方法可用于检测和定量测定核黄素，这些方法都是以核黄素的物理和生物性质为基础的。

核黄素高产细菌可以采用已知方法（见下文）在范围从振荡烧瓶到大批量的发酵罐的容器内生长。在一个优选的实例中，可以控制营

养物的供给通过变更培养基中的营养物，以最低的成本达到最大的核黄素产量。

在一个具体的实施例中，放大的含rib 的基因通过在接种菌株中（但不一定在发酵桶中）加入约 $60 \mu\text{g} / \text{ml}$ 的氯霉素。以高拷贝数被保持在细菌染色体中。可使用 1000 rpm ，压头为 0.6 大气压的 Chemap 14 升发酵桶。为了达到高的细胞密度，可以使用各种碳源，例如葡萄糖、蔗糖、柠檬酸循环酸，麦芽糖或淀粉和各种氮源如酵母膏，玉米浆液，氨和/或蛋白水解产物。然而，必须以这样的方式加入这些培养基成分的组成部分。使得对核黄素生产产生不利的限制（例如由过高碳源浓度引起的氧不足）被避免，适用的发酵培养基和条件在以下被详细描述。

例 1：核黄素高产枯草杆菌突变株

在本文实例中我们描述了一些高产核黄素的枯草杆菌菌株的产生。为了完成这一点，我们应用经典遗传学、遗传工程和发酵。选择利用嘌呤和核黄素类似物的经典遗传学可用于使嘌呤（核黄素前体）和核黄素生物合成的途径失调。通过克隆和建造核黄素生物合成途径的基因（rib 操纵子）可进一步提高核黄素的生物合成。为速率有限制的生物合成酶的组成性的，高水平的生产创造条件。

核黄素在枯草杆菌中的生物合成以 GTP 为起源（图 1）。为了获得高产核黄素的宿主。我们用经典遗传学一方面来提高细胞产生的 GTP 的量另一方面使核黄素途径失调。可以通过获得对嘌呤类似物如氮鸟嘌呤和德夸菌素以及其它拮抗物如甲硫氨酸亚砷具有抗性的突变体（参见例如 Ishii 和 Shio, Agric. Biol. Chem. 36 (9): 1511-1522, 1972; Matsui 等人,

Agric. Biol. Chem. 43 (8) : 1739—1744, 1979) 来实现嘌呤在枯草杆菌中的高产, 可以通过获得对核黄素类似物玫瑰黄色素具有抗性的突变株使核黄素途径失调 (Matsui 等人, Agric. Biol. Chem. 46 (8) : 2003—2008, 1982). 玫瑰黄色素—抗性菌株是从一些预先已致突变并对一些嘌呤类似物有抗性的菌株中挑选出来的, 以下描述了一些用于产生一种多产核黄素菌株 (RB50) 的方法。

8-氮鸟嘌呤—抗性突变株

浓度为 $500 \mu\text{g}/\text{ml}$ 的嘌呤类似物 8-氮鸟嘌呤 (Sigma Chemical Co. St. Louis, MO) 能有效地杀死枯草杆菌, 而抗性突变株以低于 10^{-8} 的发生率自发地出现。 $30 \mu\text{g}/\text{ml}$ 的乙基甲基磺酸盐 (EMS: Sigma) 用作突变剂以提高 8-氮鸟嘌呤—抗性 (Ag^r) 突变的发生率。致突变在枯草杆菌菌株 168 得到的细胞上通过标准程序完成的 (Miller, 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) 在含 $500 \mu\text{g}/\text{ml}$ 8-氮鸟嘌呤的最低培养基 (Sloma 等人, J. Bact. 170: 5557, 1988) 里平皿培养 4×10^6 已发生突变的细胞后, 再对单一的菌落划线接种, 获得 35 Ag^r 菌落, 一种突变株, RB11 (Ag^r —11), 用于 RB50 的构建中。

德夸菌素—抗性突变株

可以通过在含 $100 \mu\text{g}/\text{ml}$ 德夸菌素 (Upjohn Co., Kalamazoo MI) 的最低培养基上平皿培养细菌可以 10^{-6} 的发生率自发地获得德夸菌素抗性 (De^r) 突变, 或者在加入 EMS 之后以 10^{-5} 的发生率致突变, RB11 的一个 De^r 突变株如上所

述是通过用 EMS 致突变获得的, 一个 Dc^r 菌落, RB15 (Ag^r-11 , Dc^r-15), 被用于 RB50 的构建中。

Ag 和 Dc 突变的转移

这些嘌呤类似物抗性突变被转移至一个不同的菌株背景中, 以便从任何不需要的 EMS 诱导突变中将它们分离出来, 并核实 Ag^r 和 Dc^r 突变是单一位点引起的。由于来自单磷酸次黄苷 (IMP), 一种核黄素前体的部份“碳流”可被用于腺嘌呤核苷酸的生物合成中, 所以选择一个宿主菌株, 它的单磷酸腺苷 (AMP) 途径由于 $pur-60$ 突变而被阻断, 允许更多的碳物质从 IMP “流”到核黄素的鸟嘌呤核苷酸前体 (图 2)。使枯草杆菌菌株 1A382 ($hisH2$, $trpG2$, $pur-60$) 成为易感性的 (Sloma 等人, J. Bact. 170:5557 (1988)) 并用由 Ag^r/Oc^r 突变株 RB15 制备的完整的 DNA 进行转化 (用 Gryczan 等人, J. Bact. 134:318 (1978) 的方法) 挑选出 Trp^+ (色氨酸) 回复突变型菌落, 其中 3.3% ($10/300$) 是 Dc^r 而 2.3% ($7/300$) 是 Ag^r 。该结果不是预计不到的, 因为由于“中板集合” (一个辅助的未连结标记物的转化作用) 一些 Trp^+ 克隆也应当对德夸菌系或氮鸟嘌呤是有抗性的。

一种 Dc^r 菌落, RB36 ($hisH2$, $pur-60$, Dc^r-15)、一种 Ag^r 菌落, RB40 ($hisH2$, $pur-60$, Ag^r-11) 和一种 Dc^r/Ag^r 菌落 (也发现是 his^+), RB39 ($pur-60$, Ag^r-11 , Dc^r-15) 被一起选出用于进一步的研究。

甲硫氨酸亚砷—抗性突变株

使用高水平的甲硫氨酸亚砷 (MS : 10 mg/ml, Sigma) 进行选择产生自发的突变株, 显示出足够高的发生率以致不必用 EMS 来致突变, 该 Ag^r/Dc^r 突变株, RB 39, 被划线接种到含 10 μ g/ml MS 的最低培养基上。获得一些抗性克隆并对单一的抗性克隆再划线接种, 选择出一菌株 RB 46 (pur-60, Ag^r -11, Dc^r -15, Ms^r -46) 供进一步研究。

玫瑰黄色素抗性突变株

尽管这些 Ag^r 、 Dc^r 和 Ms^r 突变株中的许多似乎是高产 GTP 的, 但是它们中没有一个在平皿上能产生可检测到的核黄素含量, 为了使核黄素生物合成途径失调, 确定了选择对核黄素类似物玫瑰黄色素 (Toronto Research Chemical) 产生抗性的条件。在最低或完全培养基中含 100 μ g/ml 的玫瑰黄色素下, 出现细胞最大致死, 提高浓度并不产生任何额外的致死。对玫瑰黄色素呈抗性的 (RoF^r) 突变以足够高速率 (接近于 5×10^{-5}) 自发地出现, 以致不必用 EMS 或其它化学品致突变。

从上述每一菌株 1A382, RB36, RB39, RB40 和 RB46 获得接近 1000 个 RoF^r 克隆。由所有这些克隆得到的 RoF^r 突变株, 当暴露于长波紫外光 (366 nm) 下时, 在最低培养基平皿上显示的荧光水平很低, 表明产生了一些核黄素。由 RB46, RB46Y (pur-60, Ag^r -11, Dc^r -15, Ms^r -46, RoF^r -46) 获得的一些 RoF^r 克隆中的一个, 当它在最低培养基上培养时, 由 HPLC (如上述) 法测定产生 14 mg/l 的核黄素,

在所有被处理的菌株中, 只有 RB39 和 RB46 在选择 RoF^r

克隆时产生一个明显不同的表型，不管是RB39还是RB46，它们的RoF^r克隆中的大约0.5%—1.0%产生一个具有强烈荧光的黄色克隆。所有这些克隆中，由RB39产生的RB51 (pur-60, Ag^r-11, Dc^r-15, RoF^r-51)，由RB46产生RB50 (pur-60, Ag^r-11, Dc^r-15, Ms^r-46, RoF^r-50) 的稳定的，荧光黄色表型根据HLP C测定，这样的荧光黄色表型与较多核黄素生产水平有关。在最低培养基中进行培养时，RB50和RB51较其它RoF^r菌株在其上清液中能产生更高水平的核黄素，分别约为40 mg/l 或30 mg/l。RB50的系属关系被画在图5上。

由于在非Ms^r菌株如RB51中能获得强的荧光(因此是核黄素高产)克隆，似乎这种突变对更高的生长表型通常不大可能有重要的贡献，其它一些突变中的二种，Ag^r和Dc^r (RB39中的Ag^r-11和Dc^r-15)，似乎对产生高水平的核黄素是必须的，因为在只含Ag^r-11(由RB40产生)或只有Dc^r-15(产自RB36)突变的菌株中没所发现强荧光的RoF^r克隆。

gua C 突变

一种对GTP高产因而核黄素也高产的另外可能是重要的突变是gua C3，它能防止GMP反向转化成IMP(见图2)。为了构建一个含有能高产核黄素的gua C3的菌株，用RB50 DNA转化能够胜任的枯草杆菌菌株62121细胞(gua C3、trp C2、met C7)(Endo等人, J. Bact. 15:169, 1983)，并在含100 μg/ml 德夸菌素的平皿上选择Dc^r。产生了数千的Dc^r克隆。在Dc^r平板上被检验的200个克隆中，

发现一个显示出核黄素高产的表型(根据紫外荧光),并且是 RoF^R , 这一菌落被命名为RB52(*gua* C3、 *trp* C2、 *met* C7、 Dc^R-15 、 RoF^R-50)并被保存供随后研究用。

其它的类似物抗性突变株

最后,由于对一些另外的嘌呤类似物具有抗性的突变株也已被报导在嘌呤代谢中被改变,对这样一些突变株进行了鉴别,以研究它们对核黄素高产菌株的影响。测定到500 g/ml 的8-氮杂黄嘌呤、1 mg / ml 的6-硫代鸟嘌呤或2 mg / ml 的磺胺胍(Sigma)能有效地杀死野生型枯草杆菌。氮鸟嘌呤抗性的、核黄素高产菌株RB50 :: (*pRF8*)₉₀ 和RB53 :: (*pRF8*)₉₀ (见下文)被发现已对氮杂黄嘌呤具有抗性。虽然单独具有不同性质的氮鸟嘌呤和氮杂黄嘌呤抗性突变在先前已被描述,但在这种情况下 Ag^R-11 和 Ag^R-53 突变似乎还赋予了了对氮杂黄嘌呤的抗性。

HPLC分析枯草杆菌粗上清液中的核黄素

核黄素在枯草杆菌培养物中的累积已由反相HPLC定量测定。待试验的一些核黄素标准(Sigma Chemical Co., st. Louis Mo)或由菌株产生的无细胞上清液在用1%醋酸铵(PH 6.0)平衡过的4.6 mm × 250 mm Vydac C18 柱上分离。注入时,层析柱用线性梯度的甲醇展开,并在254 nm 下监测核黄素。核准了的核黄素(即核黄素“标准”)在该梯度的中点提纯。

例2:克隆枯草杆菌 *rib* 操纵子

我们分离含 *rib* 操纵子限制性片段总的策略是通过与合成的寡核苷酸探针杂交的方法来筛选枯草杆菌的一个“微型” *E. Coli* 质粒基因库,该探针的DNA序列部份地是由核黄素合成酶 β 亚单位

被公开的氨基酸序列得到的 (Ludwig 等人, J. Biol. Chem. 262: 1016, 1987), 图6给出了这一方案的梗概。

根据 Ludwig 等人定序 (J. Biol. Chem. 262: 1016-1021, 1987) 的240个氨基酸的核黄素合成酶蛋白中的第84-102氨基酸, 用一个合成的, 54-碱基 “guess-a-mer” 寡核苷酸探针进行筛选。按照基于例如在 GenBank (注册商标) (Los Alamos Nat. Lab, Los Alamos NM) 可得到的一些序列对枯草杆菌最频繁使用的密码子的估计来选取该探针中每个密码子的第三核苷酸。探针由下列序列组成:

5' - GGAGCTACAACACATTATGATTATGTTTGCAATGAAGCTGCTAAAGGAATTGCT - 3'。为测试该探针的特异性, ^{32}P 标记的54-mer DNA被杂交到含有从野生型和枯草杆菌突变株分离的被 EcoRI 消化的染色体 DNA (Southern, J. Mol. Biol. 98: 503, 1975) 的尼龙滤器上。该探针强烈地杂交到 EcoRI-消化的枯草杆菌 ($\text{rib}^+ \text{met}^-$) DNA 的单独一个9-10 kb 片段上, 它与预计的含 rib 片段的大小极为相符 (Osina 等人, FEBS. Lett. 196: 75, 1986)。当探针杂交到两种突变体菌株, RB46 (pur^-60 , Ag^r-11 , Dc^r-15 , MS^r-46) 和 RB50 (pur^-60 , Ag^r-11 , Dc^r-15 , MS^r-46 , $\text{RoFr}-50$) 时, (后者是核黄素高产菌株) 检测到一个同样大小的标记了的片段, 这些杂交试验用 Hind III 切割的染色体 DNA 重复时, 结果使该探针鉴定出一个更小的接近1.8 kb 的一个单独片段。这后一结果在确

定 *rib* 生物合成操纵子在克隆 DNA 内的一般定位是有用的。

分离含野生型 *rib* 生物合成基因的

质粒 p R F 1, p R F 2 和 p R F 3,

用由假单胞菌属复制子 R K 2 衍生的 p R K 2 9 0、一个低拷贝数的载体来制备 (Ditta 等人, Plasmid 13 : 149, 1985) 一个由枯草杆菌 168 (*rib*⁺) DNA 得到的 9-11 Kb EcoRI 片段的“微型”基因库。枯草杆菌 (*rib*⁺*met*⁻) DNA 的 EcoRI 片段 (大小 9-11 kb) 可由蔗糖 (10-40%) 速率分区离心法进行分离, 过量 4 倍的这些片段 (0.22 μg) 与 EcoRI 切割的 p R K 2 9 0 (0.26 μg) 连结, 后者已用小牛肠碱性磷酸酶 (CIAP) 脱磷酸化, 总的 DNA 浓度为 10 μg/ml, 大约 10 ng 的连接了的 DNA 被转化到 E. Coli DH5 (F⁻, end A1, hsd R11, (γ_k^- , m_k^+), SupE₄₄, thi-1, λ^- , rec A1, gyr A96, rel A1) 中, 产生一些四环素抗性 (Tc^r) 的克隆, 其出现率为 7.7×10^4 / μg DNA。为了测定含 9-11 kb DNA 插入物的转化株部份, 由一些 Tc^r 转化株制备质粒的微型溶胞产物, 用限制性酶消化法分析它们的 DNA, 发现大约有 40% 的 Tc^r 转化株含有能产生 9-11 kb 插入物的单一的 EcoRI。

采用对核黄素合成酶基因特异的 ³²P-标记的 54-mer 探针, 筛选出大约 1140 个 Tc^r 克隆。一种克隆给出了阳性信号, 质粒的 DNA 被命名为 p R F 1, 从该克隆中被分离出来。并通过将该 DNA 转化到含有核黄素一缺乏突变 *rib*-2 的枯草杆菌 1A210 中, 并选择 Rib⁺ 原始营养型克隆来检验 Rib⁺ 标记物补救活性。

p R F 1 以高出现率将 1 A 2 1 0 转化成为 Rib⁺ 原生型。由任意选择的 Tc^r 转化株得到的质粒 D N A 不能补救这种标记物。

限制性酶分析揭示, p R F 1 实际含有 1 0 k b 和 1 1 k b 的两个 EcoR I 片段插入极。为了测定那个含有 rib 操纵子的片段, 用 ³²P 标记的 5' 4-mer 核黄素合成酶探针探查 EcoR I 消化的 p R F 1。结果指出只有较小的、1 0 k b 片段与该探针交链反应。而且, 当 1 0 k b 的 EcoR I 片段被再克隆到 p B R 3 2 2 的 EcoR I 位点中时, 可得到重组体质粒 p R F 2 和 p R F 3, 表示了有两个可能的插入定向。发现这两个质粒以高的出现率将枯草杆菌 1 A 2 1 0 的 rib - 2 突变补救成为原始营养型。

从 R o F^r - 枯草杆菌菌株 R B 5 0 中分离出含有 rib

生物合成基因的质粒 p R F 6 和 p R F 7

R B 5 0 是枯草杆菌的一种 R o F^r 突变株, 其制备方法如前所述。它对核黄素的生物合成是失调的。据报导, 大约 8 0 % 的 R O F^r 突变位于 rib 操纵子中的 rib O 座位上 (Stepanov 等人, Genetika (U S S R) 1 3 : 4 9 0, 1 9 7 7), 像野生型 rib 操纵子一样, R B 5 0 中的 rib 基因也是包含在一个 9 - 1 0 k b 的 EcoR I 片段上; 因此, 用图 6 所示的方法, 用 p B R 3 2 2 作为克隆载体将该片段进行克隆。按前面所述的方法, 由 R B 5 0 制备大小选择在 9 - 1 1 k b 的 EcoR I 片段 (0. 1 微克), 并将这些片段与两倍过量的 EcoR I 切口的脱磷酸化的 p B R 3 2 2 D N A (0. 3 4 μ g) 的末端连接, D N A 的总浓度为 2 2 μ g / m l . 将大约 9 纳克 (n g) 被连接了的 D N A 转化到大肠杆菌 D H 5 中, 结果以 3.5×10^6 / 微克 D N A 的发生率产生抗 α - 氨基苄青霉

素的 (AP^r) 克隆。

对由 12 个 Ap^r 克隆的一个试样分离的质粒 DNA 进行限制性内切酶分析表明, 50% 包含具有 9—11 kb $EcoRI$ 插段的质粒。用对核黄素合成酶基因特异的 ^{32}P 标记的 54-mer 探针, 通过克隆杂交法对大约 1140 个 Ap^r 克隆体进行筛选, 有 6 个克隆体显示阳性信号, 通过限制性内切酶分析, 由这 6 个克隆体中的两个分离出的质粒 pRF6 和 pRF7 被识别为分别含有取向与 pRF2 和 pRF3 相同的插段。此外, 这两种质粒均能以高的发生率标志获救 $rib-2$ 突变。

例 3: 将 rib^+ 导入枯草杆菌中

如前所述, 可以将来自野生型枯草杆菌和枯草杆菌 $R_o F^r$ 突变株的 rib 操纵子以同样的 10 kb $EcoRI$ 片段形式克隆到大肠杆菌复制子 pBR322 的 $EcoRI$ 位点中; 图 6 示意地表示这些重组质粒的衍生过程, 为了将含有 rib 操纵子的 10 kb $EcoRI$ 片段以多重拷贝形式导入枯草杆菌中, 并由此进一步提高核黄素产量, 我们构建了一种质粒载体, 该载体能被整合到枯草杆菌染色体中, 通过选择那些能生长在高浓度氯霉素条件下的克隆体, 可以扩增被整合的 DNA。

整合 rib 质粒 pRF4 和 pRF8 的构建以及用整合

rib 质粒 pRF4 和 pRF8 进行转化

为了构建该整合载体, 将抗药基因的氯霉素乙酰转移酶 (Cat) (是在枯草杆菌中可选择的) 导入到 pRF2 和 pRF6, 即具有野生型或 $R_o F^r$ 枯草杆菌的 10 kb 片段的 pBR322 载体中, 用 $BamHI$ 消化质粒 pRF2 和 pRF6, $BamHI$ 专门在 pBR322

序列中切割这些质粒，并用 $CIA P$ 使这两种质粒脱去磷酸，将裂解的 DNA 与含有 Cat 基因的 1.3 kb $BamHI$ 片段连接 (Youngman 等人, *Plasmid* 12:1—9, 1984)，然后将被连接好的 DNA 转化到大肠杆菌 DH5 细胞中 (Hanahand, *J. Mol. Biol.*, 166:557, 1983)。 Ap^r 转化株中约有 80—90% 是氯霉素抗性的 (Cm^r)；对分离的质粒进行限制性分析 (Maniatis 等人) 证实了由 Cm^r 克隆体得到的质粒 DNA 中含有该 1.3 kb 片段。含有野生型核黄素片段和 Cat 基因的质粒被命名为 pRF4；含有由 $Ro F^r$ 菌株得到的克隆核黄素片段的质粒称为 pRF8 (因为该 $Ro F^r$ 突变随后被表明是在 rib 操纵子之外，故可以推测这些质粒是相同的)。

将质粒 pRF4 和 pRF8 转化到四种不同的枯草杆菌菌株中，这四种菌株是：核黄素高产菌株 RB50 (Ag^r — 11, Dc^r — 15, Ms^r — 46, $Ro F^r$ — 50)，RB50 亲本 RB46 (Ag^r — 11, Dc^r — 15, Ms^r — 46)，RB50 亲本 1A382，和 IS75，一种普通的实验菌株。用 pRF4 或 pRF8 转化感受性的 IS75 和 1A382 细胞，通过原在质体的转化作用将这些相同的质粒导入到 RB46 和 RB50 中 (Chang 和 Cohen, *Mol. Gen. Genet* 168:111—115, 1979)。通过选择那些能在高氯霉素浓度条件下生长的克隆体，可以将整合到这四种菌株的一个菌株中的 pRF4 或 pRF8 DNA 扩增，在每一种菌株中，我们均能获得一些能在氯霉素浓度高达 $60 \mu g / ml$ 条件下生长的克隆体。

此外，对于通过用来自 RB50 的 DNA 转化 $gua C 3$ 枯草杆

菌 6 2 1 2 1 而制得的 RB 5 2 ($guaC3$, $trpC2$, $metC7$, $Dcr-15$, $Rofr-50$), 也是感受性的, 且用 pRF8 被转化。用 $90 \mu g / ml$ 氯霉素将整合到多种所得到的 Gm^r 克隆体之一的质粒扩增。将所产生的细胞 RB 5 2 :: (pRF8)₉₀ 生长到对数生长状态中期, 并在含有 $500 \mu g / ml$ 氮鸟嘌呤的最低培养基上平皿培养, 结果产生了大约 20 种 Ag^r 克隆体 一个这样的克隆体似乎产生更强的荧光。图 7 中给出了这种菌株 RB 5 3 :: (pRF8)₉₀ 的谱系。

例 4: 用含有 pRF4 或 pRF8 的菌株高产制备核黄素

含有 pRF4 或 pRF8 的 RB 5 0 显示出核黄素高产表现型 (黄色的和紫外荧光的克隆体), 将 rib^+ DNA 在野生型菌株或 RB 5 0 的亲本菌株中进行扩增, 结果不会产生黄色的或紫外荧光的克隆体, 这说明 $Rofr$ 突变 (它使核黄素的生物合成失调) 是使野生型 DNA 发生染色体扩增从而导致核黄素高产所必需的。在 300 毫升的挡板摇瓶 (Bellco) 中, 在其中的 25 毫升核黄素最低培养基 (RMM, 表 1) 中进行一系列摇瓶发酵实验, 以测定由含有整合和扩增的 rib 操纵子的 RB 5 0 产生的核黄素。

表1 培养基的成分

R M M	克/升
谷氨酸钠	2. 0
酪蛋白氨基酸 (Difco)	0. 2
酵母膏 (Difco)	0. 2
$K H_2 P O_4$	6. 0
$K_2 H P O_4$	14. 0
$(N H_4)_2 S O_4$	2. 0
柠檬酸钠	1. 0
$Mg S O_4 \cdot 7 H_2 O$	0. 2
腺苷	0. 05
(调整到 P H 7. 0 并经热压处理)	
麦芽糖	15. 0

(热压处理后添加成灭菌的 20 % 溶液)

用菌株 R B 4 6。 R B 5 0 以及含有 p R F 4 的 R B 5 0 进行发酵, 该 p R F 4 是由对 $30 \mu g / ml$ 氯霉素有抗性的 (R B 5 0 : : [p R F 4]₃₀) 和对 $90 \mu g / ml$ 氯霉素有抗性的 (R B 5 0 : : [p R F 4]₉₀) 选出物扩增的。在 24 小时和 48 小时, 取出上清液试样, 并用逆相 H P L C 测定核黄素的含量。

如表 2 所示, 在 48 小时内, R B 5 0 : : [p R F 4]₃₀ 产生 0. 3 克/升核黄素, R B 5 0 : : [p R F 4]₉₀ 产生 0. 7 克/升核黄素, 该产量比由那些无 rib 扩增的菌株如 R B 4 6 和 R B 5 0 获得的要高得多。

表 2

对枯草杆菌的含核黄素的上清液进行的定量分析

菌株	培养时间 (小时)	核黄素* (克/升)
RB 4 6	2 4	0. 0 0 9
RB 5 0	2 4	0. 0 2
RB 5 0 : : (p R F 4) ₃₀	2 4	0. 1
RB 5 0 : : (p R F 4) ₉₀	2 4	0. 4
RB 4 6	4 8	0. 0 0 7
RB 5 0	4 8	0. 0 5
RB 5 0 : : (p R F 4) ₃₀	4 8	0. 3
RB 5 0 : : (p R F 4) ₉₀	4 8	0. 7

*核黄素含量是用 H P L C 测定法测定。

由于 rib 基因在失调宿主中的扩增作用,使核黄素产量剧烈增加,这证明了由克隆 D N A 所编码的信息控制着核黄素生物合成的速度。

例 5: R₀ F^r—5 0 突变的定位图

RB 5 0 中的 R₀ F^r—5 0 突变似乎是核黄素高产表现型的关键,为了对该突变进行鉴别并尽可能将它移入不同菌株的本底中,必须绘制 R₀ F^r—5 0 突变在枯草杆菌染色体上的定位图。由于 p R F 4 和 p R F 8 在所有的菌株本底中产生非常相似的核黄素产量,因此, R₀ F^r—5 0 突变似乎不大可能定位于该被克隆的 1 0 kb EcoR I、含 rib 片段上,更大可能量, R₀ F^r—5 0 突变是一种未连接的阻遏型突变,它可能量在 rib C 中,该突变是一种已被报

导为定位于枯草杆菌染色体的 *lys - aro* D 区域内的阻遏突变 (Chernik 等人, Genetika (U S S R) 15 : 1569, 1979)。为了确定该 $RoF^r - 50$ 突变是否与核黄素操纵子相连, 可以用 *RB50 DNA* 将感受性的枯草杆菌 *1A210* (*rib - 2*) 细胞转化, 选出 rib^+ , 结果获得几千个 rib^+ 克隆体, 将 200 个克隆体接种到含有 100 克/毫升玫瑰黄色素的胰蛋白胨血琼脂基上, 结果不产生出 RoF^r 克隆体, 且没有一个克隆体呈现出核黄素高产表现型, 这证实了 $RoF^r - 50$ 突变不是定位于 *rib* 操纵子中。

例 6: 利用 *Cat* 插入诱变法定位 rib^+ 的生物合成基因

图 4 中包括对 *pRF2* 的含 *rib* 的 10 kb *EcoRI* 片段的限制性定位图, 该片段是按标准方法制备的, *XbaI*, *BglII*, *SstI*, *HpaI* 和 *NcoI* 的限制性内切酶位点是该插段 *DNA* 特有的, 而 *SalI* 和 *PstI* 在该插段中切割一次, 且在该载体中切割一次; 该插段中不含任何 *BamHI*, *XhoI* 或 *NheI* 限制性位点; 限制性内切酶 *Hind III* 在多重位点上裂解该插段; 对核黄素合成酶基因特异的 54-mer 探针被杂交到一个约 1.8 kb 的 *Hind III* 片段上, 这说明该 *rib* 操纵子必定也位于围绕 *SalI* 和最左边的 *BglII* (*BglII_L*) 位点的总区域中。

通常, 为了确定 *rib* 操纵子的边界, 需要采用一些较小的、含 *Cat* 的限制性片段。在 *pRF2* 的 rib^+ 被克隆的 *DNA* 片段中构建一些插入和缺失, 大肠杆菌质粒 *PEcc1* 起携带 *Cat* 基因的限制性片段主要来源的作用, *Cat* 基因使大肠杆菌和枯草杆菌对氯霉素具有抗性。这种质粒是 *pMI 1101* 的一种衍生物 (Youngman 等人, Plasmid 12, 1-9, 1984), 在该衍生物中, 该质

粒的非必要的区域已用标准重组DNA技术除去，包含了一个1.3 Kb 的含Cat 片段，该片段 M13 mp7 的“多接头” (polylinker) 侧面相连，因此，这种质粒能产生一些具有 SmaI, EcoRI, SalI 或 BamHI 末端的Cat 盒。为了产生以SstI 或 XbaI 为末端的、含Cat 基因的片段，分离出 pEcc1 的1.3 Kb 含Cat 的BamHI 片段，将其末端用 HindIII 接头修饰，并将修饰后的片段克隆到pIC20R 多接头区域内的HindIII 位点中，产生出质粒pEcc4。

首先，在大肠杆菌中构建出整合质粒衍生物，然后通过DNA转化将它们转化到枯草杆菌的rib 染色体座位中，这一过程可以通过用该克隆DNA插段外面的限制性内切酶切口使该质粒线性化、用该切口DNA转化感受性的枯草杆菌菌株1A382或pY79 (β^c , rib^+) 细胞、并选出 Cm^r 来实现。由于pBR322复制子不能在枯草杆菌中复制，而且该Cat 基因由与 rib^+ 座位同源的序列连于两侧，故只能通过双交叉重组过程将含有Cat 的插入或缺失插入该染色体中，以产生 Cm^r 转化株。为了确定该插入或缺失是否使核黄素生物合成失活，可以在有核黄素(Rib 表现型)存在或无核黄素存在时，测定 Cm^r 克隆在最低培养基琼脂皿上的生长情况。

如图8所示，可以按照标准方法，将含有Cat 的限制性片段通过连接法插入到pRF2上的各个XbaI, SstI, SalI 和 BglII 限制性位点中，插入到一对BglII或NcoI 位点之间(产生去除2.0 Kb BglII 片段或0.8 Kb NcoI 片段的缺失) 或者插入到被杂交到rib 特异的DNA探针上的约1.8 Kb HindIII 片段上的单个HaeIII和EcoRV 位点中。其结果示于表

Ⅲ 中。

表 Ⅲ

rib[±] D N A 的插入和缺失衍生物的特性

<u>插入衍生物^a</u>	<u>枯草杆菌^b 核黄素表现型</u>
A(<u>Xba</u> I)	
r	+
l	N D
B(<u>Sst</u> I ₁)	
r	+
l	N D
C(<u>Sst</u> I _R)	
r	--
l	--
D(<u>Bgl</u> II _L)	
r	--
l	--
E(<u>Sal</u> I)	
r	--
l	--
F(<u>Bgl</u> II _R)	
r	--
l	--
G(<u>Hae</u> III)	
r	N D

1	+
H (<u>E c o R V</u>)	
r	+
1	N D
缺失衍生物	
B g l	
r	— —
1	— —
N c o	
r	+
1	+

a “r” (右) 和“1” (左) 表示插入的 *Cat* 基因相对于图 8 中限制性定位图的转录取向。

b 枯草杆菌菌株 1 A 3 8 2 (*rib*⁺, *trpC* 2, *pur* — 6 0, *hisH* 2) 或 P Y 7 9 (*S P β*^c, *rib*⁺)

由图 8 及表 III 可以看出, 插入到 *Sal* I 中、不论是 *B g l* II 或“最右边”的 *Sst* I (*Sst* I_R) 位点, 或者缺失 2. 0 kb *B g l* II 片段, 都会形成不能产生核黄素 (*Rib*⁻) 的 *Cm*^r 克隆体, 这说明 *rib* 操纵子集中位于克隆的 DNA 中, 有意义的是, 去掉 0. 8 kb *Nco* I 片段对核黄素的产生 (*Rib*⁺) 明显没有影响, 这说明该 *rib* 基因群的一端位于“最左边”的 *Nco* I (*Nco* I_L) 位点的左边, 该 *rib* 操纵子的另一端起初被认为是定位于约 1. 8 kb 的 *Hind* III 片段中, 这是因为在该片段中一些位点如 *Eco* RV 和 *Hae* III 以及该片段远端的一些位点 *Xba* I 和 *Sst* I_L 上的两个插入都产生了能产生核黄素

的Cm^r克隆体。

例7: rib 操纵子的核苷酸序列

根据在被克隆的10 kb DNA片段上进行的Cat 插入式诱变, 整个rib 操纵子定位于以SstI_L和NcoI_L位点为界的一个6.0 kb 区域内。

用Sanger 等人的双脱氧法对pRF2上含有rib 操纵子及其侧面区域的这个6.0 kb 区域进行定序(*proc. Natl.*

Acad. Sci. USA 74:5463, 1977), 简而言之, 先制备定序用的M13克隆体, 其制备方法可以是: 将特定的限制片段克隆到M13中; 利用T4 DNA聚合酶的核酸外切酶活性产生一系列重叠缺失(Dale 等人 *Plasmid* 13:31, 1985); 或者通过“射击枪”将经过声处理过的限制性片段得到的随机的片段克隆到M13中。有时, 也可以用引物延伸定序法来确定穿越相邻片段限制性接合位点的核苷酸序列。在两条链上大约排列着5500个碱基对, 分析其中与具有格兰氏阳性细菌核糖体结合位点、格兰氏阳性启动子和rho 独立转录终止位点的典型的开放式解读码相似的序列。

分析结果表明有6种完整的、非重叠的开放式解读码(图3): ORF2(124个氨基酸), 编码了核黄素合成酶的 β 亚单位的基因(154个氨基酸), ORF3(398个氨基酸), ORF4(215个氨基酸), ORF5(361个氨基酸)和ORF6(105个氨基酸), 每一种ORF前面均有一个强的芽孢杆菌核糖体结合位点(RBS), 经计算的热稳定性范围为 $\Delta G = -16 \sim -22$ 千卡/摩尔, 而且所有ORF的取向均与转录方向相同。此外,

在ORF 5的编码区内,还识别出一个第二RBS位点及ATG起始密码子,它可能编码了一种较小的具有248个氨基酸的蛋白质,但是根据S-30体外偶联转录/转译反应(参见下文),ORF 5似乎仅编码了一种具有361个氨基酸的蛋白。最后,还识别出另一编码区,ORF 1的一部分,它编码了一种蛋白质的最后170个氨基酸,且取向相反。

根据下面的观察结果,杆菌中的核黄素合成是受单独一种操纵子控制的,该操纵子含有5种基因:β核黄素合成酶基因,ORF 2, ORF 3, ORF 4和ORF 5,其中的至少四个基因:β-核黄素合成基因,ORF 3, ORF 4和ORF 5明显地编码了生物合成酶,而余下的一种基因ORF 2则可能编码了一种生物合成酶。

1. ORF 3, ORF 4和ORF 5与插入含Cat 限制性片段会使枯草杆菌中核黄素产生失活的限制性酶位点重叠(图4和图8)。

2. ORF 1与这样一种限制性酶位点重叠,在该位点中插入含Cat 限制性片段不会在 rib^+ 枯草杆菌菌株中的核黄素产生失活(表Ⅲ和图8),也不会使失调的 RoF^r 枯草杆菌菌株RB 5 2中的核黄素产量减少。

3. ORF 2也与这样一种限制性酶位点EcoRV 重叠,在该位点中插入含Cat 限制性片段不会在 rib^+ 枯草杆菌菌株中的核黄素产生失活(表Ⅲ和图8),但是这种插入将使在失调的 RoF^r 枯草杆菌菌株RB 5 2中的核黄素产量可观地减少,说明突变的ORF 2基因产物对于核黄素的产生是部分失活的,这一结果表明,ORF 2确实编码了一种 rib^- 特异酶。

4. 在ORF 1和ORF 2之间以及ORF 5和ORF 6之间的

内顺反子间隙中，识别出两个DNA序列，该DNA序列能形成一些表征rho-独立转录终止位点的杆一环(Stem-loop)结构(图4和图9)，将ORF5和ORF6之间的这些结构去除会提高核黄素的表达。该结构使lacZ-融合结构具有核黄素敏感性，因此，可以用这些结构使任何一种与之在该启动子5'端上游区融合的有关基因具有这种敏感性。

5. 在离ORF5约290bp的上游，还识别出一种DNA序列，TTGCGT-(17bp)-TATAAT，它与由枯草杆菌RNA聚合酶的 σ^A (营养型)识别的启动子相类似，其取向与ORF5的转录方向相同(图4)。将这种启动子(P_1)，在1.1kb BglII-NcoI 限制性片段上)转录融合到一种无启动子的大肠杆菌lacZ基因(P_1 -lacZ)上，显示出在一个rib⁺枯草杆菌菌株(62121)中由 β -半乳糖苷酶活性产生的核黄素的调节表达，以及只有在该启动子的取向与该基因的转录方向一样时显示出在一个rib⁺，RoF^r 枯草杆菌菌株(RB52)中的 β -半乳糖苷酶活性产生高水平的构成型(失调的)表达，如表N所示。可用引物延伸分析来证实该起始位点，用转录及Northern分析法可以得知，一个4.2kb的多顺反子RNA包围了整个rib操纵子。

表 IV
P₁ — LacZ 转录融合物的核
黄素—被调节的表达

菌株 (整合质粒)	β — 半乳糖苷酶特异活性 (米勒单位)	
	+核黄素	—核黄素
	(2 μg / ml)	
枯草杆菌 62121(P ₁ — <u>lacZ</u> ^a)	1. 3	4. 2
枯草杆菌 RB52(P ₁ — <u>lacZ</u> ^a)	3 1	3 8
枯草杆菌 62121(P ₁ — <u>lacZ</u> ^b)	< 0. 1	< 0. 1
枯草杆菌 6 2 1 2 1	< 0. 1	< 0. 1

a. P₁ 和 lacZ 取向相同

b. P₁ 和 lacZ 取向相反

根据这些结果, 可以认为该 σ^A 启动子 P₁ 是 ORF 5, ORF 4, ORF 3, β — 核黄素合成酶基因及 ORF 2 转录过程的基本启动子。

6. 在 ORF 4 的 3' — 末端中, 还识别出一个第二 DNA 序列, TTGAAG — (17 bp) — TACTAT, 它与用枯草杆菌 RNA 聚合酶的 σ^A (营养型) 识别的启动子相类似, 它位于 ORF 3 上游的 295 bp 处, 其取向与 ORF 3 的转录方向相同 (图 4), 通过 Campbell 型重组过程, 将含有这种启动子序列 (在 0. 7 kb SalI — BglII 限制性片段上) 的大肠杆菌质粒整合到枯草杆菌中不会使枯草杆菌中核黄素的产生失活, 结果说明该第二序列 (P₂) 具有启动子活性, 因此可以实际控制 (除 σ^A P₁ 启动子外) ORF 3,

β 亚单位核黄素合成基因和 O R F 2 的转录过程。LacZ 融合及 Northern 分析证实了这种启动子的存在。

7. 在核黄素合成酶基因的 β 亚单位和 O R F 2 之间的内顺反子区域内还识别出一个第三 D N A 序列, T T G A A T - (18bp) - T A A A A A, 该序列可能与用枯草杆菌 R N A 聚合酶的 σ^A 营养型识别的启动子相类似, 它位于 O R F 2 上游约 83 bp 处, 且其取向与转录方向相同 (图 4), 除了 P_1 和 P_2 之外, 该 σ^A 启动子, P_3 也控制着 O R F 2 的转录过程。

8. 对该克隆 D N A 的 S-30 反应进行体外偶联转录/转译分析, 证实了 O R F 2, O R F 3, O R F 4 和 O R F 5 实际上都编码了具有由它们各自序列预测尺寸的蛋白质。

9. 利用 GenBank (注册商标) 或已知的蛋白质尺寸, 通过将预测的一些氨基酸序列或其产物的分子量与公开的蛋白质序列进行比较, 可以将核黄素生物合成过程中 5 个假设的酶步骤中的 3 个分派给特定的编码区。

a. 由 O R F 2 和 O R F 3 之间的开放式解读码编码的该推定的蛋白几乎完全相同地与公开的核黄素合成酶 β 亚单位的 154 个氨基酸序列 (Ludwig 等人, J. Biol. Chem. 262: 1016, 1987) 相一致, 只发现一种氨基酸不同: 残基 65 上的赖氨酸被甘氨酸取代。据报导, 这种酶能催化由 5-氨基-6-ribitylamino-2,4(1H, 3H)-嘧啶二酮-5'-磷酸盐 (图 1, 分别是结构 5 和 4) 和 3,4-二羟基丁酮-4-磷酸盐形成 6,7-二甲基-8-ribityllumazine。

b. 经识别, 在 O R F 5 的推定产物及脱氧胞苷酸脱氨酶 (一种

由大肠杆菌噬菌体 T₂ 编码的 188 氨基酸蛋白质, (Maley 等人, J. Biol. Chem. 258:8290, 1983) 之间, 一个 88 一氨基酸的重叠中有 39% 的等同性。根据这一结果, ORF5 很可能编码了 rib 特异脱氨酶, 该脱氨酶能催化由 2, 5-二氨基-6-(核糖氨基)-4(3H)-嘧啶二酮-5-磷酸盐形成 5-氨基-6-(核糖氨基)-2, 4(1H, 3H)-嘧啶二酮-5'-磷酸盐 (图 1 中, 分别是结构 3 和 2)。

c. 预测的 ORF4 基因产物的分子量 (26000 道尔顿) 与核黄素合成酶 α -亚单位的分子量 (23000 道尔顿, Bacher 等人, J. Biol. Chem. 255:632, 1980) 十分相符, 根据这一结果, ORF4 编码了核黄素合成酶的 α -亚单位, 它能催化该生物合成途径的最后步骤: 6, 7-二甲基-8-ribityl-umazine 歧化成核黄素 (图 1, 分别是结构 5 和 6) 和 5-氨基-6-ribitylamino-2, 4(1H, 3H)-嘧啶二酮。

10. 通过将 ORF 位置与该操纵子中 rib 突变的物理定位图 (Morozov 等人, Mol. Genet. Mik. Virusol. no. 7:42 (1984)) 对准, 可以将核黄素合成中其余的酶催化步骤暂时归结为编码区域, 据报导, 缺陷 GTP 环化水解酶的突变定位于 0.5 kb Hind III 片段上, 由于 ORF3 包围该限制性片段, 所以我们的结论是 ORF3 至少部分地编码了该酶催化的功能部分, 该功能部分能催化由 GTP 形成 2, 5-二氨基-6-(核糖氨基)-4(3H)-嘧啶二酮-5'-磷酸盐 (图 1, 分别为结构 2 和结构 1)。此外, 据报导, 编码了 rib 一特异还原酶的生物合成基因被完全包含在约 1.8 kb Hind III 片段中, 由于该片段仅含有两个完整的编

码区,即核黄素合成酶基因的 β 亚单位及ORF 2,故我们推测ORF 2编码了该还原酶,它能催化由5-氨基-6-(核糖氨基)-2,4(1H,3H)-嘧啶二酮-5'-磷酸盐形成5-氨基-6-(ribitylamino)-2,4(1H,3H)-嘧啶二酮-5'-磷酸盐(图1,分别为结构4和结构3)。

此外,在该操纵子明显的前导区中,可以检测到一种类似的rho-独立转录终止位点,它位于推定的 $\sigma^A P_1$ 启动子的下游,又正好在该操纵子ORF 5第一编码区的上游(图4和图9),这种可能的终止区结构可能通过终止/抗终止作用与rib操纵子的调节作用有关。此外,在ORF 3 0.7 kb的SalI-BglII限制性片段上存在一个与玫瑰黄色素抗性(RoF^R)有关的调节区。

将rib ORF分配到蛋白质产物中

用于证实rib特异ORF是否编码了蛋白质的一种方法是用pRF 2和它的各种衍生物作为模板,对S-30体外偶联转录/转译反应中由该克隆DNA合成的蛋白质的大小及数目进行目测,S-30组分药盒(New England Nuclear:按制造商说明使用)在转译枯草杆菌基因方面是特别有效的,这是因为它们存在强的核糖体结合位点。

用pRF 2或pRF 4克隆的1.0 kb EcoRI片段作模板,预期检测到5种推定的rib特异蛋白质: β -核黄素合成酶,14.7千道尔顿(kd)(Ludwig等人,J. Biol. Chem. 262:1016,1987);来自ORF 2的蛋白质,13.6 kd;ORF 3,43.7 kd;ORF 4,23 kd;和ORF 5,39.7 kd。预计还能检测到至少两种其他的蛋白质,这两种蛋白

质是由ORF6 (11. 6 kd) 和ORF1 (至少18. 7 kd) 编码的, 以及检测到由该10 kb 克隆DNA片段上未被定序区域内的基因编码的任何一种另外的蛋白质, 此外, 预期还有与载体有关的蛋白质, 包括bla 和 Cat 抗生素抗性的基因产物 (tet 基因在S-30反应中不被强烈表达)。

除去bla - 和 Cat - 特异蛋白 (分别是32和18 kd) 及其它与载体有关的蛋白质外, 在用pRF2或pRF4的S-30反应的一个15 %-SDS聚丙烯酰胺凝胶上检测到共6种主要的³⁵S 标记蛋白质, 其分子量分别是47 kd , 44 kd , 38 kd , 26 kd , 20 kd 和15 kd 。为了将这些蛋白质产物分派到与它们相应的rib - 特异ORF上, 用各种可得到的缺失衍生物, Cat 插入衍生物以及该10 kb EcoRI 克隆DNA的亚克隆片段 (图10) 重复S-30反应, 其结果示于表V中。

表 V

在S-30反应中观察到的rib - 特异蛋白质

	47,000	44,000	26,000	15,000
	道尔顿	道尔顿	道尔顿	道尔顿
质粒	(ORF3)	(ORF5)	(ORF4)	(ORF2)
pRF2	+	+	+	+
pRF4	+	+	+	+
pRF21	-	-	+	-
pRF5	-	-	-	+
pRF29	-	-	-	-
pRF12	+	-	+	+

p R F 1 0	—	—	—	—
p R F 3 8	—	—	—	—
pRF24 / pRF20	—	+	+	+
p R F 2 3	+	—	+	+

根据这些结果，可以将蛋白质产物分派到 O R F 3 (4 7 k d)，O R F 5 (4 4 k d)；O R F 4 (2 6 k d) O R F 2 (1 5 k d)。这些分子量与预测的大小极其相符。

将产物分派到 O R F 2 和 β 核黄素合成酶基因上不如分派到其它 O R F 上那么直截了当，因为 p R F 2 的 S-3 0 反应产生一个 1 5 k d 蛋白质，这和由两种基因任一编码的蛋白质的预测大小接近，所以我们首先假定该蛋白质段实际上同时含有两种蛋白质。但是，将 Cat 插入到质粒 p R F 3 8 中的 O R F 2 中就完全除去了这种蛋白段，而代之以一种更小的 6 k d 的蛋白质，该蛋白质分子量与截短的 O R F 2 的预测尺寸极相符合。根据这些结果，该 1 5 k d 蛋白质似乎仅仅是由 O R F 2 产生的，但尚不清楚为什么在 S-3 0 反应的凝胶上看不到 β -核黄素合成酶蛋白，然而，总的来说，该结果证实了存在 5 种 rib-特异编码区：O R F 5，O R F 4，O R F 3，O R F 2 和 β 核黄素合成酶基因。

此外，ORF1 似乎编码了一种 3 8 k d 的蛋白质，而没有识别出 O R F 6 的产物。

rib 操纵子的调节机理

在枯草杆菌中，几位研究者已经观察到了用于生物合成通道的基因组成和调节的循环方式，枯草杆菌的色氨酸生物合成途径 (Henner

等人, Gene 34 : 169, 1984) 及整个嘌呤核苷酸合成途径 (Ebbole 和 Zalkin , J . Biol . Chem . 262 : 8274, 1987) 的核苷酸序列中都含有成群的重叠基因, 这些基因以一种多顺反子信息被转录, 并且至少一部分是由一种新的转录终止/抗终止机理调节的, 在所述的机理中, 涉及到一种阻遏蛋白, 该阻遏蛋白可以由一种未接合到该生物合成操纵子上的基因进行编码 (Zalkin 和 Ebbole , J . Biol . Chem 263 : 1595, 1988)。由于我们发现该 rib 生物合成及调节基因的组成与枯草杆菌 trp 和 pur 途径的组成明显相似, 因此, 我们假设该 rib 操纵子至少有一部分可能以一种类似的方式被调节。

总之, 该转录终止/抗终止模型的几个关键特征包括 (Shimotsu 等人, J. Bacteriol, 166 : 461, 1986): (i) 在该操纵子第一基因之前有一个长的 5' 前导序列; (ii) 在该 RNA 前导序列中存在两种或两种以上重叠的二元对称体, 该对称体能够相互形成一些专有的 RNA 杆-环结构, 一个结构起 rho-独立转录终止子的作用, 而另一个则起“抗终止子”的作用 (阻止该 rho-独立转录终止子的形成); (iii) 在受阻遏条件下, 被途径最后产物活化的阻遏蛋白与阻止抗终止子形成位点上的新生 mRNA 结合, 从而能够形成终止转录的终止子; (iv) 在去除阻遏的条件下, 未活化的阻遏物的结合被防止, 结果形成一个使通读转录进入该操纵子编码区中的抗终止子。

如上所述, rib 操纵子中最有可能的起始转录位点是 σ^A 启动子 P_1 , 定位于该操纵子中第一基因上游的 290 bp 处, 对该 RNA 前导序列的初步分析表明, 它含有通过该终止/抗终止模型进行调节

所必需的绝大部分(如果不是全部的话)结构,在该区域中,在 ORF 5 上游区的 50 bp 处识别出一种杆-环结构,该结构之后有一种类似 rho-独立转录终止区的一串胸苷,这种序列具有形成一种 ΔG 为 -26 千卡/摩尔的发夹的位能(图 9),此外,尚有几种 ΔG 范围在 -13 至 -16 千卡/摩尔之间的可能的杆-环结构定位于可能看作为抗-终止子序列的 rib 5' 前导序列中。

除了通常在操纵子第一基因上游的该主要起始转录位点外,在一些生物合成途径中,还存在一些定位于该启动子内部区域中的辅助启动子位点,以前,人们通过对该 rib 操纵子进行 R-环异源双链的研究,也提出了该 rib 座位内可能存在一些辅助的启动子位点 (Osina 等人, FEBS Letters 196: 75-78, 1986), 显示出两个或两个以上可以引发 mRNA 合成的位点。我们对该 rib 操纵子的内顺反子间隙的初步分析没有探测到这样一些辅助的启动子位点,但当将该分析扩展到该操纵子中所有序列时,在 ORF 4 的 3' 一端中识别出另一个 σ^A 启动子 P_2 , 它刚好在 Sal I 限制性位点的下游(图 4), 因此, ORF 2, ORF 3 和核黄素合成酶 β -亚单位的表达也有可能是在这个辅助启动子的控制下。此外,就在 ORF 2 的上游,识别出了一个可能的第三 σ^A 启动子 P_3 , 因此, ORF 2 可能也是在该附加启动子的控制之下。

表 6 表示推定的编码区,启动子及转录终止位点在 5, 5 kb 枯草杆菌 rib-特异区的 DNA 序列上的定位情况。

表 VI
枯草杆菌 Rib 操纵子的编码区
域、启动子及转录终止位点

		<u>碱基对数目^a</u>
编码区	O R F 6	3 6 4—6 7 8
	O R F 5	1 1 0 1—2 1 8 3
	O R F 4	2 1 9 7—2 8 4 1
	O R F 3	2 8 5 9—4 0 5 2
	β—核黄素合成酶基因	4 0 8 8—4 5 4 9
	O R F 2	4 6 6 5—5 0 3 6
	O R F 1	5 5 6 7—5 0 5 7 ^b
σ A 启动子	P ₁	7 7 1—7 9 9
	P ₂	2 5 2 8—2 5 5 6
	P ₃	4 5 4 5—4 5 7 4
<u>rho</u> 独立终止位点	在 5' 启动子上游	7 0 8—7 4 8
	在 5' 前导 RNA 中	1 0 3 4—1 0 6 7
	在 rib 操纵子的 3' 端	5 0 3 8—5 0 9 0

a. 图 3 上的

b. 取向相反的编码区

例 8: 构建含有一个改性 rib 操纵子的载体

前面对枯草杆菌 rib 操纵子的功能分析用于第一次确定核苷酸序列的调节区和开放解读码, 也使我们可以在提高核黄素产量有用的新的载体, 我们对这些区域中有关核黄素生物合成所必需的特定

基因的定位知识, 有关转录控制区域定位的知识以及这些基因中的其它有关区域(如RBS)的定位知识使我们对这样一些区域能作出一些改变, 下面就是这样一些处理的一些实例。

用一个较小的DNA片段上的

rib操纵子构建一个整合质粒

用来构建核黄素高产菌株RB50:: (pRF8)的整合载体含有包括该rib操纵子的10 kb EcoRI 片段, 由于该rib操纵子似乎仅占有不到6 kb 的DNA, 因此, 构建了一个新的包含了在较小DNA片段上的rib操纵子的整合载体(pRF40), 这种更小尺寸的克隆体可以更高地扩增该rib 基因, 获得更高的核黄素产量。

参见图12, pRF40可以由pRF36构建得到, pRF36是pRF2的一个0.8 kb NcoI片段被一种Cat基因取代了的质粒, 该rib操纵子被包含在一个6.5 kb 的XbaI — EcoRI 片段上, 分离这种片段, 并将它连接到用XbaI 和 EcoRI 消化的pUC 19上(Yanisch — Perron 等人, Gene 33, 103 1985; 可由New England Biolabs Boston, MA, USA和Bethesda Research Laboratories, Maryland, USA获得)。将连接好的DNA转化到DH5 α 大肠杆菌中, 并在含有40 μ g / ml X-gal 和50 μ g / ml α -氨基苄青霉素的LB平皿上培养, 对由白色克隆体制得的DNA微样本(miniprep DNA)进行分析, 结果表明pRF39含有该6.5 kb 的XbaI — EcoRI片段。

用EcoRI消化pRF39, 并用CIA P处理, 然后将它连接到含有Cat 基因的1.6 kb EcoRI 片段上, 然后, 将连接好的

DNA 转化到 DH5 α 大肠杆菌中并选择适合的克隆体用于放在 LB + 10 $\mu\text{g} / \text{ml}$ 氯霉素平皿上进行培养，有两个克隆体是氯霉素抗性的，对由这些克隆体制备的 DNA 微样本进行分析，结果证实了 *Cat* 基因的存在，这些质粒中的一种是 pRF40 (图 14)。

构建含有转录改性 *rib* 操纵子的质粒

如前所述，用能够使核黄素生物合成基因构成性表达的启动子取代核黄素操纵子上的启动子及操纵区将是有用的，含有这样一些结构的质粒然后可以用于产生能使核黄素生产水平提高的细菌菌株，下面将给出一些实例，但本发明不限于这些例子。

参照图 13，将核黄素启动子及调节区除去，并代之以 SPO1 启动子，我们采用的是位于 ORF3 起始端 1130 位上的 BglII 位点，合成寡核苷酸 (RB5 和 RB6，参见图 18)，该寡核苷酸在直至 1058 位置的 BglII 位点 (ORF5 和 SD 序列开始几个不多的氨基酸) 上重新建立该 DNA 序列 5'，该操纵子 5' 一端的新构建过程在任何一种被提议的 DNA 调节结构前停止 (图 13)，在它们的 5' 端，该寡核苷酸含有 BamHI, NsiI 和 EcoRI 限制性位点，从而可以对该 *rib* 操纵子放置各种各样的启动子 5'，由于在该 *rib* 操纵子中存在这多种的限制性位点，因此必须如下地以几个步骤用一些新的启动子来构建该操纵子。

从 pRF36 中分离出一个 1.4 kb SalI - BglII 片段 (图 13)，将该片段与两种寡核苷酸和用 EcoRI - SalI 消化的 pUC19 连接，然后将连接好的混合物转化到大肠杆菌 DH5 α 细胞中，并在含有 50 $\mu\text{g} / \text{ml}$ α -氨基苄青霉素和 40 $\mu\text{g} / \text{ml}$ X-gal 的 LB 上平皿培养，由 AP^{r} 的白色克隆制备出微样

本；一个具有所需结构的质粒是 p R F 4 6 (图 1 3)。

用 BamHI 和 SalI 消化 p R F 4 6，并从中分离出该 1. 4 kb 片段，然后将该片段与 p N H 2 0 2 (含有 SPO1-15 启动子的 p U C 8，Lee 和 Pero，J. Mol. Biol.，152: 247-265，1981) 上的 400 bp EcoRI - BamHI 片段，以及用 SalI 和 EcoRI 切开的 p U C 1 9 连接。将连接好的 DNA 转化到 D H 5 α 大肠杆菌中，并将它放在 LB + α -氨基苄青霉素 + X-gal 上平皿培养，由白色克隆体制备 DNA 微样本；pRF48 具有所需的结构 (图 1 3)。

用 EcoRI 和 SalI 消化 p R F 4 8，并从中分离出 1. 8 kb 片段，将该片段与来自 p R F 2 和 XbaI，EcoRI - 切割 pUC19 的该 4. 0 kb XbaI - SalI 片段 (含有 rib 操纵子的其余部分) 连接，然后将连接好的混合物转化的大肠杆菌 D H 5 α 细胞中，将该细胞放在 LB + α -氨基苄青霉素 + X-gal 上平皿培养。由白色克隆制备 DNA 微样本；p R F 4 9 具有所需的结构，而且由含有该质粒的培养物得到的上清液呈黄色，说明有核黄素产生 (图 1 3)。

为了将 Cat 基因放入 p R F 4 9 中以允许对枯草杆菌进行选择，用 XbaI 消化该质粒，并将它与来自 p Ecc4 的一个 1. 3 kb 的含 Cat XbaI 片段连接，将连接好的 DNA 转化到大肠杆菌 D H 5 细胞中，产生了数百个 A P^r 克隆体，将这些克隆体接种到含有 LB + 10 μ g / ml 氯霉素的平皿上，这些克隆体中大约有 10 % 可以在该氯霉素平皿上生长，这说明该 Cat 基因的存在，含有 Cat 的质粒称为 p R F 5 0 (图 1 4)。

上述例子表明在 O R F 5 上游安放了一种新的启动子，我们发现，

在O R F 3和O R F 4之间、P₂之后放置一个启动子也会是有益的，它可以进一步提高核黄素产量，下面就是这种构建的一个实例。

参见图14和15，为了在O R F 3上游放置SPO1-15启动子的一个拷贝，我们利用了靠近O R F 4-O R F 3接点的一些限制性位点，在2767位的C1a I位点位于O R F 4的末端，该位点是rib操纵子中特有的，另一个有用的靠近O R F 3起始端的限制性位点是在2892位的Dra I位点，合成了一些重新建立从上述Dra I位点直到通过O R F 3起始端以后序列的寡核苷酸，并在O R F 3起始端（接头P2-A和P2-B，图18）之前放置了一个独特的BamHI位点。另一组寡核苷酸重新建立了从C1a I位点直到通过O R F 4末端以后的序，并在该位置（接头P2-CII和P2-DII，图18）上放置一个EcoRI位点。然后将位于EcoRI-BamHI片段上的SPO1-15启动子放在由该寡核苷酸产生的BamHI和EcoRI位点之间。以下列方式将该附加的SPO1-15启动子与整个操纵子放在一起。

参见图15，将含有O R F 4，O R F 3功能的750bp Sal I-Bgl II片段亚克隆到p I C 20 R中（Marsh 等人，Gene 32，481-485，1984），然后，用Dra I和Bgl II消化所产生的质粒，p R F 57，从中分离出该预计的270bp Dra I-Bgl II片段，将该片段与接头P2-A和P2-B一起连接到用Sal I和Bgl II切开的p I C 20 R上，二个接头将BamHI和Sal I位点放置在O R F 3 5'端的上游，（由于Bgl II和BamHI位点是相容的，并在以后被除去，所以Sal I位点应选择简易方便）。将该连接物转化到大肠杆菌DH5α细胞中，并在LB介

质 + Amp 和 X-gal 上平皿培养, 结果产生了一些白色克隆体; p R F 5 8 具有所需的结构。从 p R F 5 8 中分离出 3 3 0 b p . BglII — salI 片段, 并将它与来自 p R F 3 6 的含有 rib 操纵子 3' 一端的 3. 3 k b BglII — XbaI 片段 (图 1 2) 以及与用 XbaI 和 SalI 切割的 p U C 1 9 连接, 然后将连接好的 D N A 转化到大肠杆菌 D H 5 α 细胞中, 结果产生白色克隆体; p R F 6 2 (图 1 5) 具有所需的结构。为方便起见, 从 p R F 6 2 中分离出该 3. 6 k b BamHI — XbaI 片段, 并将它亚克隆到 BamHI — XbaI 切开的 p U C 1 9 中 (p R F 6 4, 图 1 5), 此时, 该质粒含有该 3. 6 k b rib 操纵子的 3' 端, 该操纵子在 O R F 3 之前带有一个被建造的 BamHI 位点。

为了在含有最后三种开放解读码的 rib 操纵子 3' 一半分处之前放置 SPO1 — 1 5 启动子, 我们用 EcoRI 和 BamHI 消化 p R F 6 4, 并将它连接到一个含有 SPO1 — 1 5 启动子的 4 0 0 b p EcoRI — BamHI 片段上, 将连接好的 D N A 转化到大肠杆菌 D H 5 细胞中, 并制备 D N A 微样本; p R F 6 5 具有所需的结构。

将 SPO1 — 1 5 启动子建造成可以在该启动子之前放置一个 ClaI 位点的形式, 从而重新构建 O R F 4 的末端。将来自 p N H 2 0 2 的含有 SPO1 — 1 5 启动子的 EcoRI — BamHI 片段与接头 P 2 — C I I 和 P 2 — D I I 以及用 BamHI 和 ClaI 消化的 p C I 2 0 R 连接。然后将连接好的 D N A 转化到大肠杆菌 D H 5 α 细胞中, 结果产生白色克隆体, 对微样本进行分析, 结果表明 p R F 6 3 具有所需的结构, 然后从 p R F 6 3 中分离出 4 7 0 b p ClaI — BamHI 片段, 并将它与来自 p R F 4 9 的含有 SPO1 — 1 5 启动

子及 rib 操纵子的 5' 端的 2 kb EcoRI — ClaI 片段以及用 EcoRI 和 BamHI 消化的含有 SPO1 — 15 启动子和该操纵子 3' 一端的 p R F 6 4 连接 (图 15)。将连接好的 DNA 转化到大肠杆菌 DH5 α 细胞中, 制备微样本; p R F 6 6 具有所需的结构。此外, 含有 p R F 6 6 的大肠杆菌在 LB 介质 + α -氨基苄青霉素平皿上产生少量的核黄素, 这肯定了该操纵子仍然是完整的。

最后一步是将 Cat 基因连接到上面所述的 p R F 6 6 特有的 XbaI 位点中, 所产生的质粒 p R F 6 9 (图 15) 含有取向和 rib 操纵子相同的 Cat 基因。

为了构建一种包含整个操纵子 (该操纵子在 rib P₂ 之后具有天然的或野生型 rib P₁ 启动子和 SPO1 — 15 启动子) 的质粒, 需要将 p R F 6 4 的 6.3 kb EcoRI — BamHI 片段、p R F 3 6 的 2.75 kb EcoRI — ClaI 片段和 p R F 6 3 的 470 bp ClaI — BamHI 片段连接起来, 并转化到大肠杆菌 DH5 α 细胞中, 约有 50% 的 A P^r 克隆是黄色的, 这说明产生了核黄素, 由黄色克隆体制备 DNA 微样本, p R F 6 8 具有所需的结构 (图 16)。如前所述, 在 XbaI 位点上将一个 Cat 基因加入到 p R F 6 8 中, 从而产生了 p R F 7 1 (图 16), 该质粒含有取向与 rib 操纵子相同的 Cat 基因。

另一种构建本发明中 useful 质粒的例子如下文所述, 在该实例中, 无需事先将现有的 DNA 序列去除, 即可在核黄素操纵子中导入一种或多种启动子。

举一例来说, 可以在 p R F 7 8 中构建一个原型的被改性的操纵子, 所述的 p R F 7 8 含有: (1) 单独一个被插入位于 rib P₁

和推定的 ρ ho — 独立转录终止位点之间 (图 14) 的 30 bp 必需区域的该 SPO1 — 15 启动子的拷贝; (2) 一种被失活的 $\text{rib } P_1$ 启动子, 以防止 SPO1 — 15 启动子可能的转录干扰; (3) 一种活性的 $\text{rib } P_2$ 启动子; (4) 编码 rib 生物合成酶的 5 种结构基因, 以及 (5) 位于核黄素操纵子末端下游的约 1.5 kb 侧面分生的 DNA 核苷酸序列。

参见图 14, 首先将 pRF2 的 1.7 kb NcoI — PstI 片段亚克隆到 mp19 中, 该 NcoI — PstI 片段含有该 rib 操纵子的 5' 启动子区域及侧面分生区域的片段, mp19 是大肠杆菌噬菌体载体 M13 的一种衍生物 (United States Biochemical Catalog, 60—61, 1987; 可由 New England Biolabs, Massachusetts, USA 获得)。回收一种重组噬菌体 M1.7, 并对其启动子区进行标准 DNA 顺序分析, 结果发现在该 $\text{rib } P_1$ 启动子的 -10 区域内有一种自发突变, 即一个由 TA 到 GT 的变化, 该突变可以使该启动子失活。制备一种单股 DNA, 并将它退火成一种可合成制得的 55 bp DNA 寡聚物 (参见图 17), 该寡聚物含有多个限制性酶位点的组合, 即 5' — EcoRI — SmaI — BamHI 3', 其两侧均有与 $\text{rib } P_1$ 上游的 DNA 区域同源的附加序列。用标准的位点导引诱变 (SDM) 方法合成双股 DNA 分子, 通过转染, 将这些 DNA 分子导入大肠杆菌宿主 TG-1 (Amersham Corp., Illinois, USA) 中, 从而产生重组噬菌体噬菌斑, 通过用标准的 DNA 顺序分析测定, 发现一种含有所需改性 DNA 序列的重组噬菌体。

然后, 利用距离 $\text{rib } P_1$ 侧面及周围序列 750 bp 的一对特

有的 Nsi I 限制性内切酶位点, 将该改性 rib 启动子区域重新连接到该操纵子的 rib 结构基因中, 制备该噬菌重组体的双股 DNA 分子, 并用 Nsi I 消化, 从中分离出 750 bp 片段, 然后将该片段与 pRF39 Δ R1 (由 pRF39 衍生的一种质粒, 图 12, 它含有野生型 rib 操纵子) 脱去磷酸的、8.7 kb Nsi I 片段连接。通过转化, 将连接好的 DNA 分子导入大肠杆菌 DH5 α 细胞中, 选出其中抗 α -氨基苄青霉素的克隆, 结果获得一个带有所需重组质粒 pRF75 的 AP^r 克隆。

下一步, 通过用 EcoRI 和 BamHI 酶的结合消化 pRF75, 将该被切开的 DNA 与经过提纯的 400 bp 含 EcoRI - BamHI SPO1-15 的限制性片段连接、并将连接好的 DNA 通过转化导入大肠杆菌 DH5 α 细胞中, 选出其中抗 α -氨基苄青霉素克隆等步骤可以将 SPO1-15 启动子插入 rib P₁ 的上游。发现有一种包含该重组质粒 pRF77 的 AP^r 克隆体。它含有所需的 SPO1-15 一改性的 rib 操纵子。在此之后, 在该特有的 XbaI 位点将位于 1.6 kb XbaI 限制性片段上的氯霉素抗性基因 Cat 导入 pRF77 中, 结果产生质粒 pRF78 (图 14)。

如前所述, 还可以将这种原生型操纵子进一步改性成含有一种活性 rib P₁ 启动子, 和/或一个被引入到 rib 编码区 ORF3 和 ORF4 之间的内顺反子区域内 rib P₂ 下游的该 SPO1-15 启动子的第二拷贝。举例来说, 利用含有野生型 rib P₁ 启动子的重组噬菌体、采用和前面相同的方法, 可以构建质粒 pRF88, 该质粒含有 pRF78 (带有一种活性 rib P₁ 启动子) 中修饰 rib 操纵子的一种衍生物 (图 14)。在其它一些例子中, 通过在任何一

种质粒DNA上除去2.0 kb BglII片段,并插入pRF66的2.4 kb BglII片段,可以将位于rib P₂下游的SPO1 — 15启动子的一个第二拷贝插入到现有的含有改性rib操纵子的质粒pRF78和pRF88中,从而分别产生质粒pRF81和pRF89(图14)。

构建Ade⁺RB50菌株

使用那些只需向发酵介质中添加尽可能少的几种组分的细菌菌株是十分重要的,这样一些菌株可以较廉价地发酵产生核黄素。为此,构建了含有扩增的改性rib操纵子的腺嘌呤回复突变株,这些回复突变株并不一定是真正的pur — 60回复突变株,而在另一个抑制腺嘌呤需求的位点上包括突变,如下文所述,它们比非回复菌株多产生约25%的核黄素,现在来描述这样一些构建过程的实例。

将质粒pRF8, pRF40、pRF50, pRF69, pRF71, pRF78, pRF81, pRF88和pRF89中的每一种都转化到RB50(一种R^oFr⁻、失调的枯草杆菌菌株)中,挑选对氯霉素有抗性(Cm^r)的菌株。对每一种菌株选出一个抗性克隆,通过在含有10 μg / ml腺苷的RMM1培养基中培养细菌,并将培养试样放在最低琼脂平皿上培养,可以分离出每种菌株的Ade⁺回复突变株,从每一种Ade⁺菌株中选出一克隆体,并通过选择那些能在氯霉素含量不断增加(最高到60 μg / ml)的条件下生长的克隆体,扩增该载体DNA。

第二位点整合

如前所述,将被建造的rib操纵子在枯草杆菌染色体中扩增,使核黄素达到高的滴定度是十分重要的。同时确保rib操纵子在染

染色体中的DNA拷贝数不限制核黄素的产生也是十分重要的。可以通过在枯草杆菌染色体一个以上的位点上整合并扩增该rib操纵子的这些拷贝来实现rib操纵子进一步的扩增，从而进一步提高核黄素的产量，下面将描述是如何实现这种第二位点的整合的。

上述这些载体全部依赖于该Cat基因，可以在rib操纵子的位点上整合。为了在一个第二位点上将该rib基因插入，最好在该第二位点上采用一个不同抗生素抗性的基因。例如，可以采用枯草杆菌的一个抗四环素基因(tet) (Perkins 和 Youngman, J. Bacteriol. 155: 607—615, 1983), 这样一些tet基因对本领域的专业人员来说是众所周知的，而且对这类人员来说是容易得到的。在一个这样的构建过程中，例如，可以用Xba I将质粒pRF78 (图14) (含有一个该rib操纵子的修饰形式)，切开，并将它与含有该tet基因的2.4kb Xba I片段连接。将产生的质粒在Xba I位点上含有该tet基因，并把该质粒叫做pRF85。

需要一个缺失整个rib操纵子、并具有一个被整合在一个第二位点上的tet基因的菌株来使pRF85整合在该位点上。一个这样的位点是编码了杆菌肽酶F的bpr基因，杆菌肽酶F是一种较小的，非必需的胞外蛋白酶，用EcoRV消化含有该bpr基因的大肠杆菌质粒、pKT2 (Sloma等人, J. Bacteriol., 172: 1470—1477, 1990)。该EcoRV位点是在bpr的编码区内。然后将该DNA与端部已钝化的含有tet基因的2.4kb的EcoRI片段连接，所产生的质粒(在bpr的EcoRV位点上含有该tet基因)被称为pKT2—tet。用EcoRI将该DNA线

性化，然后转化到RB52中，RB52是一种核黄素合成失调的菌株。结果产生了 tet^r 克隆，一个这样的克隆被叫做RB54，被整合在bpr上的tet基因将起一个整合pRF85的同源序列的作用。

为了确保pRF85的被克隆的核黄素操纵子被插在含有四环素抗性基因的该第二染色体位点上，需要用体外方法从RB54的染色体中缺失掉一个含有原始核黄素操纵子及侧面DNA的区域（等于包含在pRF85中的区域）。概括地说，这包括先要制备一种大肠杆菌重组质粒，在该重组质粒中，在NcoI和XbaI限制性部位之间的克隆核黄素操纵子及侧面区域被去除，并代之以一个在枯草杆菌细菌中被表达的抗氯霉素基因，即Cat基因。然后，通过用线性化的质粒分子转化RB54并选出抗氯霉素（ Cm^r ）的细菌，将该质粒用于使其染色体核黄素操纵子缺失。 Cm^r 细菌是一种用含有Cat基因的被缺失拷贝取代野生型rib基因的重组过程（标志取代）产生的。

具体地说，可以用pRF34（参见例6）产生一种含有一种体外产生的核黄素操纵子缺失的大肠杆菌质粒，该质粒由pRF2衍生的，在pRF2中，该核黄素操纵子的两端均侧接两个特有的XbaI位点（其中一个位点位于缺失的0.8 kb NcoI片段之后的rib操纵子5'一端的上游，该第二位点则位于该操纵子末端下游约1.6 kb处）。而且在该区域的外面还插入了一个Cat基因。通过用XbaI消化pRF34，并在稀释的DNA浓度下，连接该切开的DNA分子，可以获得一种重组质粒pRF82，在该质粒中，一个含有核黄素操纵子的7.2 kb区域被除去，并基本上由Cat基因所取代。通过限制性内切酶消化，将质粒pRF82线性化，并用该切开的

DNA通过转化除去RB54的染色体核黄素操纵子，选取Cm^r细菌，结果导致标记取代，从Cm^r克隆中筛选出核黄素营养缺陷型，获得一种Rib⁻ Cm^r克隆，RB55，供进一步研究用。

将质粒PRF85转化到菌株RB55中，以选出Rib⁺。选择一种Rib⁺转化株，称为RB58，该菌株具有rib操纵子，该操纵子通过在该质粒中的tet^r基因和染色体之间的同源重组而被整合在bpr处，由RB58制备染色体DNA，并用它来转化RB50:: (pRF69)，从中选出Tet^r。这些抗性克隆体就具有被整合在该rib操纵子的该位点上和bpr上的改性rib操纵子。如上所述，通过选择那些能够在具有愈来愈增加的氯霉素含量下生长的克隆体，可以将整合在rib上的rib操纵子扩增，通过选择那些能在愈来愈增加的四环素含量下生长的克隆体，可以将该rib操纵子的第二拷贝扩增。

例9：核黄素的发酵生产

在Chemap 14升容器中，对发酵过程（其配料中碳有限）中的核黄素高产菌株进行评价，用HPLC法测定核黄素含量。由于反应速度是受由核黄素合成基因编码的酶的限制，因此，我们将通过在接种株中（而不是在发酵罐中）包括60 μg/ml的氯霉素将被扩增的rib基因保持在高拷贝数下。

将枯草杆菌RB50:: (pRF69)的培养物在含60微克/毫升氯霉素(CAM)的胰蛋白胍血琼脂培养基(TBAB Difco)上培养，将克隆体转移到300毫升的挡板摇瓶中，摇瓶中含有25毫升具有60 μg/ml CAM的核黄素最低培养基(RMM：含有谷氨酸钠2.0克/升，酪蛋白氨基酸(Difco) 0.2克/升，

酵母膏 (Difco) 0.2 克/升, KH_2PO_4 6.0 克/升,
 K_2HPO_4 14.0 克/升, $(\text{NH}_4)_2\text{SO}_4$ 2.0 克/升, 柠檬酸钠
1.0 克/升, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0.2 克/升, 腺苷 0.05 克/
升, 葡萄糖 15.0 克/升, pH 7.0)。将接种后的摇瓶在 37
℃下, 以 250 转/分摇动培养, 8 小时以后, 加入灭菌甘油, 使最
终浓度达到 15%, 并将 1 毫升的等分试样贮藏在 -80℃ F。

为了引起发酵, 将 RB50 :: (PRF69) 的冰冻小瓶在
37℃下解冻, 并将它转移到 300 毫升的挡板摇瓶中, 该摇瓶含有
25 毫升 RMM (含 60 $\mu\text{g}/\text{ml}$ CAM), 并将摇瓶在 37℃下
以 250 转/分摇动, 8 小时后, 用 6 毫升这种生长培养物接种在一
个 2 升转移烧瓶中的 300 毫升发酵介质 (见下面表 VII) 中。该烧瓶
中含有 300 毫升发酵介质, 发酵介质中已被加入 90 毫升由 15%
葡萄糖和 30% 麦芽糖组成的混合物。加入氯霉素, 使最终浓度达到
60 $\mu\text{g}/\text{ml}$ 。在 37℃下在一个 2" 直径轨道的搅拌器上以 200
转/分搅拌培养 12 小时以后, 将该烧瓶中的物质转移到在, 一个
14 升发酵容器中的 7 升发酵介质中。

在发酵过程中, 要连续监视培养基的 pH 值及溶解的氧气的量
(DO_2)。通过四极质谱分析法连续分析排出的气体, 并记录二氧化
碳的放出量 (CER) 和氧气吸入速度。

将几种发酵过程进行比较, 可以证实该控制系统的可重复性, 原
始的碳水化合物在生长 4 小时以后, 由于 RB50 :: (PRF8)₀₀
的发酵过程而被耗尽, 使 pH 上升, CER 下降。此时, 开始加进碳
水化合物, 对数生长重新开始, 一直到 DO_2 在 6 小时时达到极限为
止, 用计算机控制碳水化合物的供应速度, 以使 DO_2 在剩下的发酵

时间内保持在饱和的10—20%之间。

发酵罐中碳水化合物过量时，肯定会导致氧缺乏现象，并降低核黄素产量，氧气传送的限制决定着对数生长的持续时间、最后的细胞密度，以及核黄素的生产速度，为了增加氧气传送速度，Chcm_{ap}发酵罐应以1000转/分速度运行，其压头为0.6大气压。

用酵母膏向介质中添加碳水化合物将使核黄素产量高于没有添加碳水化合物的介质（图11，空的方块：RBF-14；表Ⅶ）。但是由于酵母膏成本较高，故通过代之以低廉的无机成分，有意地将酵母膏用量降低。用氢氧化钠取代氨水来控制PH值，可以降低酵母膏在加料中的用量，且增加了细胞质量及核黄素的滴定度（图11，黑的方块：RBF-22；表Ⅶ），发酵时间也缩短了。此外，在其它一些发酵过程中，酵母膏可被完全从加料中除去，并代之以一个无机铵盐及磷酸盐的组合，引起核黄素产量进一步增加，且缩短了发酵时间（图11，空的圆：RBF-23；表Ⅶ）。

原始的RB50::[pRF8]_{eo} 由于它的Pur-60突变而对腺嘌呤是有缺陷的，当进行一些实验以确定该菌株所需的最低腺苷量（以便它对包含在形成核黄素前体IMP（图2）过程中的早期生物合成酶的抑制作用达到最小）时，发现RB50::[pRF8]_{eo}（一般还有其染色体中具有被扩增的rib操纵子的RB50菌株）对腺苷的需求量是不稳定的，而且以很高的速率产生原养型回复突变株（Ade⁺）。在摇瓶中，Ade⁺回复突变株似乎至少与其RB50::[pRF8]_{eo}亲本一样地生长和产生核黄素，当在发酵罐中进行评估时，RB50::[pRF8]_{eo}(Ade⁺) 在介质配料中不需要腺苷，尤为重要的一点是，该回复突变株与其亲代菌株相比，

在较短的时间内生长得更快，并且多产生 25 % 的核黄素，在 49 小时内，产生滴定度为 5.4 克/升的核黄素（图 11，密实的园：RBF-29；表 VII），在别的一些发酵过程中，HySoyT 被从原始配料或介质中除去，并代之以玉米浸泡液，结果使 48 小时内的核黄素产量进一步增加到 6.3 克/升（RBF-42，表 VII）。

在这些发酵条件下，利用含有被建造的核黄素操纵子 DNA 的细菌菌株，可以使核黄素产量获得进一步明显的增加，在 6.5 Kb EcoRI - XbaI 限制性片段上含有野生型核黄素操纵子的菌株 RB50 :: (pRF40)_{eo} (Ade⁺) 在 48 小时内产生 7.4 克/升的核黄素。而且，含有转录改性 rib 操纵子（其中，ribP₁ 启动子和调节区被组成型 spo1-15 启动子取代）的菌株、RB50 :: (pRF50)_{eo} (Ade⁺) 在 48 小时内产生 9.0 克/升的核黄素。这些结果证明，通过除去调节区和/或导入更强的组成型外源启动子对该核黄素操纵子的改性导致核黄素滴定度的提高。

表 VII

发酵组分及条件

组 分	R B F-1 4	R B F-2 2	R B F-2 3	R B F-2 9	R B F-4 2
原始配料 (g / l)					
葡萄糖	1 0. 0 0	1 5. 0 0	1 5. 0 0	1 5. 0 0	1 5. 0 0
玉米浆	—	—	—	—	1 0. 0 0
Hy Soy T	1 5. 0 0	1 5. 0 0	1 5. 0 0	1 0. 0 0	—
谷氨酸钠	—	—	—	5. 0 0	5. 0 0
Amberex 5 0 0	1 5. 0 0	1 5. 0 0	2 0. 0 0	2 0. 0 0	2 0. 0 0
KH ₂ PO ₄	5. 0 0	5. 0 0	7. 5 0	7. 5 0	7. 5 0
MgCl ₂ · 6H ₂ O	0. 5	0. 5	1. 5 0	1. 5 0	1. 5 0
MnSO ₄	0. 0 5	0. 0 5	0. 0 5	0. 0 5	0. 0 5
Adenosine	0. 0 5	0. 0 5	0. 0 5	—	—
MAZU DF37C	2. 5 0	2. 5 0	2. 5 0	2. 5 0	2. 5 0
FeCl ₃	—	—	0. 0 2 5	0. 0 2	0. 0 2
CaCl ₂	—	—	0. 5 0	0. 5 0	0. 5 0

表 VII (续 I)

发酵组分及条件

组 分	RBF-14	RBF-22	RBF-23	RBF-29	RBF-42
原始配料 (g/l)					
ZnSO ₄	--	--	0. 0005	--	--
CuCl ₂	--	--	0. 0013	--	--
CoCl ₂	--	--	0. 0013	--	--
进料营养 (g/l)					
Amberex500	160. 00	120. 00	--	--	--
NH ₄ Cl	--	--	7. 50	7. 50	7. 50
(NH ₄) ₂ SO ₄	--	--	7. 50	7. 50	7. 50
KH ₂ PO ₄	--	--	15. 00	15. 00	15. 50
MgSO ₄ · 7H ₂ O	--	--	2. 50	2. 50	2. 50
DL-70 糖浆 (as DS)	600. 00	600. 00	600. 00	660. 00	600. 00

表 VII (续2)

发酵组分及条件

组 分	R B F-14	R B F-22	R B F-23	R B F-29	R B F-42
p H 控制范围					
6. 6	H ₂ SO ₄	H ₂ SO ₄	H ₂ SO ₄	H ₂ SO ₄	H ₂ SO ₄
6. 5	NaOH	NH ₄ OH	NH ₄ OH	NH ₄ OH	NH ₃
条 件					
空气 (vvm)	1. 0	1. 5	1. 5—2. 0	1. 5	1. 5 0
转/分钟	1 0 0 0. 0	1 0 0 0. 0	1 0 0 0. 0	1 0 0 0. 0	1 0 0 0. 0
温度 (°C)	3 7. 0	3 7. 0	3 7. 0	3 7. 0	3 7. 0
压力 (bar)	0. 5	0. 5	0. 5—0. 7 5	0. 6	0. 6
核黄素 (g/l)	3. 4 (6 4 小时)	4. 1 (5 6 小时)	4. 3 (5 3 小时)	5. 4 (4 9 小时)	6. 3 (4 8 小时)
干重 (g/l)	3 3. 6	3 6. 0	3 6. 8	N D	4 4. 6

用 Luedeking - Piret 模型分析各种发酵过程中的核黄素形成动力学, 在所有的情况下, 该特定生产率从指数生产阶段结束时一直下降到发酵过程结束。而且, 很显然, 该核黄素的生产与所用的发酵条件下的生长情况有关。

我们发现, 通过改变发酵组分及条件也能提高核黄素产量, 与前面所述的条件相比, 采用如表 VIII 所示的发酵组分和条件, 可以提高核黄素产量。

表 VIII

原始配料	(克/升)
酵母膏	2 0
葡萄糖	2 5
KH_2PO_4	7. 5
$\text{MgCl}_2 \cdot \text{H}_2\text{O}$	1. 5
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	1. 0
MnSO_4	0. 0 5
$\text{FeCl}_3 \cdot 6 \text{H}_2\text{O}$	0. 0 2 5
MaZU DF 3 7 C	2. 5
玉米浸泡液	1 0
谷氨酸钠	5
<u>加入介质 (总共用 3 升)</u>	
葡萄糖	5 8 3. 3
柠檬酸钠	6. 6 7
KH_2PO_4	1 5
琥珀酸	1. 6 7
$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	1. 6 7

总而言之, 在一个这样的发酵过程中, 起始物料是 6. 65 升配料介质和 0. 35 升细菌 (RB50::[pRF50]₈₀ Ade⁺) 接种物。用一个 chemap 极谱溶氧电极监测氧气含量。通过用计算机调节供入介质的添加量, 将溶解的氧气保持在 15%±5%。在 48—56 小时内共添加约 3. 0 升补充介质, 发酵 pH 值保持在 6. 5±0. 1 (用 1 N H₂SO₄ 和 NH₃ 气体), 发酵罐压力保持在 0. 6 巴, 空气流速为 10. 5 升/分。在这些条件下, 菌株 RB50::[pRF50]₈₀ (Ade⁺) 在 48 小时内产生 11. 0 克/升核黄素, 这表明它比以前的发酵条件多生产约 20% 的核黄素, 最后, 采用含有一种转录改性核黄素操纵子的细菌菌株 RB50::[pRF69]₈₀ (Ade⁺) 可以进一步提高核黄素产量, 所述转录改性核黄素操纵子含有两个 SPO1-15 启动子, 其中一个启动子取代 rib P₁ 和调节区, 另一个启动子插入到 ORF3 和 ORF4 之间。该菌株在 48 小时内产生 13. 0—14. 0 克/升核黄素, 在 56 小时内产生 15 克/升核黄素, 这说明用两个强的外源启动子提高了核黄素操纵子的转录作用, 可以使核黄素增产。

微生物寄存

质粒 pRF69 已根据布达佩斯条约, 于 1990 年 6 月 6 日寄存在美国典型培养物收藏中心 (American Type Culture Collection), 其收藏号为 ATCC 68338。

含有质粒 pRF50 的大肠杆菌菌株 DH5 已根据布达佩斯条约, 于 1990 年 5 月 30 日寄存在美国典型培养物收藏中心, 其收藏号为 ATCC 68332。

含有质粒 pRF78 的大肠杆菌菌株 DH5 α 已根据布达佩斯条

约，于1990年5月30日寄存在美国典型培养物收藏中心，其收藏号为ATCC 68333。枯草杆菌菌株RB58也于同日寄存在该中心，其收藏号为ATCC 55053。

枯草杆菌菌株RB50已于1989年5月23日根据布达佩斯条约寄存在农业研究培养物收藏中心(Agricultural Research Culture Collection , NRRL, Peoria, Illinois)，其收藏号为B18502。

说明书附图

图 1

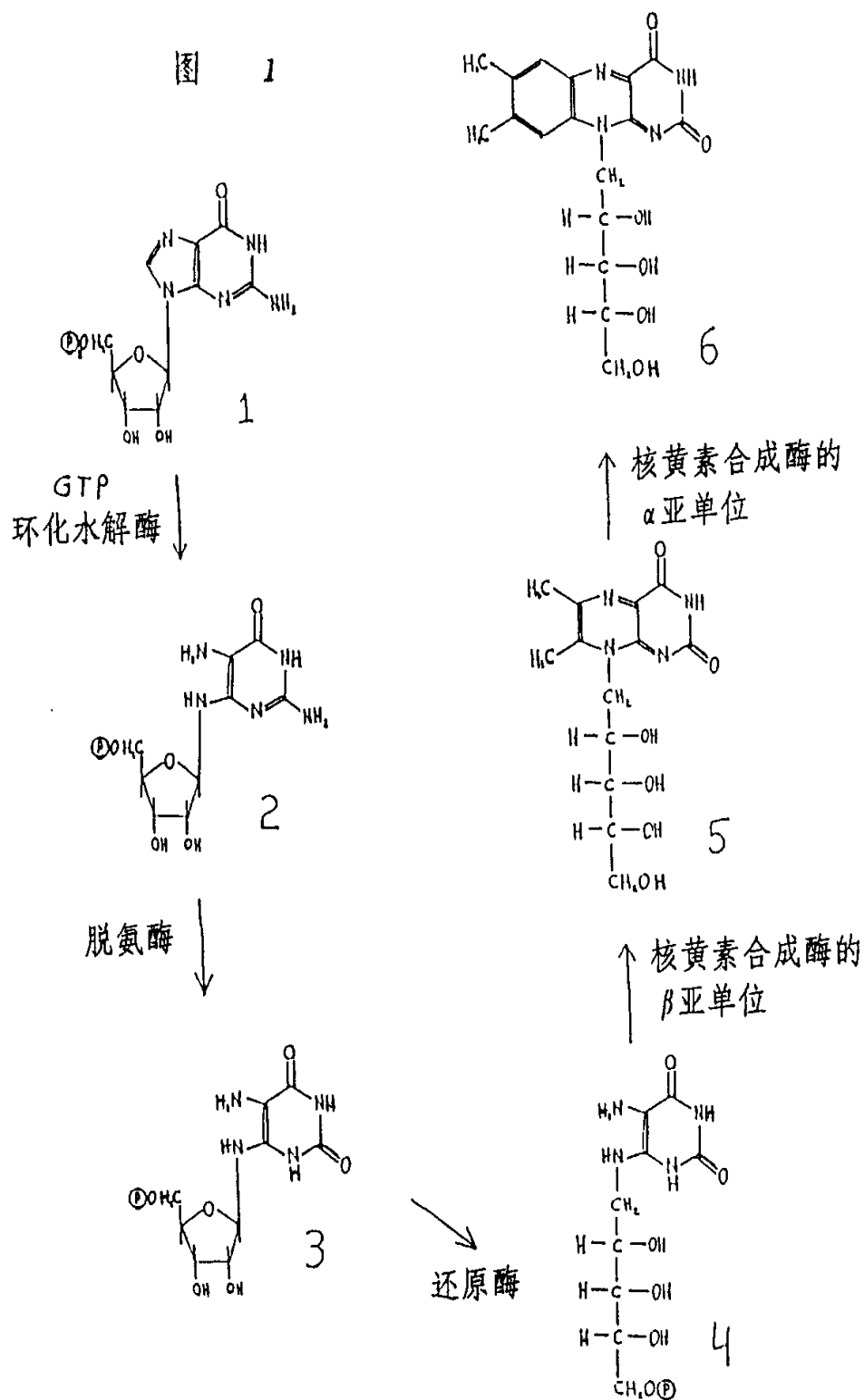
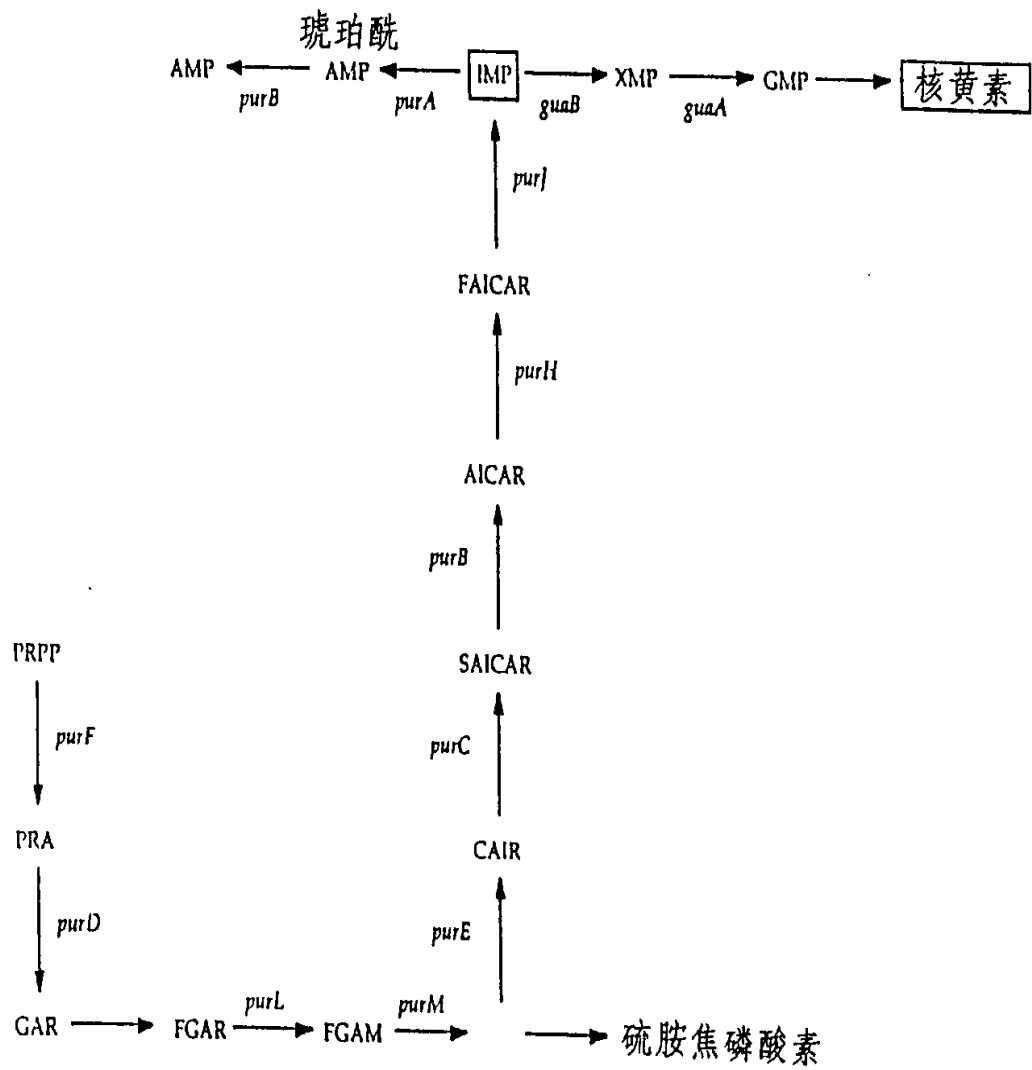


图 2



M13多接头顺序

PPSAHTTHKH	BXBHDBDSNOSBBH	H B HSSBBMH	KE H		
SLACIANNBA	IHALBPITCSESSL	A G ATESSBAA	SA M		RH
TELCHQFLAE	NOMAONWAYOACAAA	E L EYCAAOEE	PR L		SA
1111211111	12141111111JJ3	3 1 311JJ213	21 1		AE
/	///	///	///	/	12

CTGCAGGTCGACTCTAGAGGATCCCCATGGACAGCCGTAACGGCTTGGCCCTTCACGAAAAACAATTCGGGTACGTCAAAGTTTGTTCIACC

GACGTCAGCTGAGATCTCTAGGGGTACCTGTCCGCATTGCCGGAACCGGAGAGTCTTTTGTGTTAACGCCCATGCAGTTTCAAAACAAAAGATGG

100

l q v d s r g e p m d a r n g l o l f t k k k q l a g t s k f v f y p
 c r s t l e d p p w t a v t a l a s s r k n k l r v r q a l f s t
 a g r l . r l p h e q p . r p w p l h e k t n e g y v k v c f l p
 c t s e l p d g m a l r l p r p r k v f f c l a p v d f n t k .
 q l d v r s a g g h v a t v a k s e e r f f l n r t r . l k n e v
 a p r s . l l g w p c g y r g q g r . a f v f q p y t l t q k r g

M	HFN	BM D	F AB	B F	F
S	ISH	IB P	N LB	B H	O
E	NPA	HO N	U UV	V U	K
I	P11	11 I	11 I	I H	I

CGTTTAAAGAAATCGGCAAAACAAATAGGATCAAGCAGCTTCCCATGGGGCTGCTTTTTATATCTTTTACGGTCATCCCTAAAAACAGAACAT

GCAAAATGCTTTACCGGTTTTGTTAATCTAGTTCGTGCAAGGGTAACCCCGACGAAAAAATATAGAAAAATGCCAGTAGGGGATTTTGTCTTGTGA

200

f n e m r k t n . d q a a s h w g c f f y l f f t v l p . k q n l
 r l t k c a k q l r i k q l p l g a e f f l a f l r s e p k n r t .
 v . r n a q n k l g s s s f p l g l l f l y l f y g h p l k t e h
 g n l s l r l v f . a . a a e w q p q k k . l k k v t m g . f c f m
 r k v f h e f c i l l l c a g m p a e k k l d k k r d d g l l l v
 t . r f a c f l n p d i l k g n p a e k k y r k . p . g r f v s c

H		HA	S AM
N		SV	F SS
F		1A	A EE
3		13	N 11

AAATTCGTATATCTATAGAAAAGAAATTTTGCAGAAATGTGAACATATTCGGTATGCATCGTTATTAATAATTTACGAGAAATTCAGGTTTTTT

TTTAAGCATATAGATATCTTTTCTTAAAAACGCTTTACACTTTGTATAAGGGCAATACGTAGCAATATAATTTAAATGCTCTTAAATGCCAAAAA

300

n a y l y r k e l f a e m . n l f p l e l v l l l l y e n l r f l
 l r l s l e k k f l q k e o t y s r y a s l y . . f t r l y g f l
 k f v y l . k r n f c r n v k h l p v m h r y l n n l r e f t v f y
 f a y l . l f s l k a s l h f m n g n h m t l n l l . s f k r n k
 y l r i d l s f f n k c f h a v y e r . a d n y . y n v l l . p k k
 l n t y r y l l f k q l f e f c l g t l c r . l l l k r s n v t k

图 3 A

3	H		H	XH	H		S	H
S	L		D	MN	N		P	B
P	A		E	HF	L		O	O
H	3		1	11	1		1	2

RBS ORF 6

ATTCATGAAAAAAGGAATAACTCATATGAATGAATAGATTTCATATTGGCTGGAGGTTTACAAATGGGAAGAATAAAAAACCAAGATTACCATTCGTITAG
 400
 TAAGTACITTTTTTCCTATTGAGTATACCTTACTTACTAAGTATAACCGACCTCCAAATCTTTACCCCTCTTATTTTGGTTCTAATGGTAAGACAAIC

i h e k k e . l l . m n r f l l a g g l e m g r i k t k l t l l l v
 f m k k r n n a y e . l d s y w l e v . k w e e . k p r l p f c .
 s . k k g l t h m n e . i h l g w r f r n g k n k n q d y h s v s

 l . s f f s y s m h l f l n m n a p p k s l p l l f v l i v m r n
 n m f l l f l e y s h l s e y q s s t . f h s s y f g l n g n q .
 . e h f f p l v . l f s y l . l p q l n l f p f f l f w s . w e t l

R	A	F	BSB H	KEF H
S	L	O	SES H	SAO H
A	U	X	ACA L	PRK L
1	1	1	J1J 1	211 1

/ /

TGCTTTTGCCTTTACITGCAGGCGGTATATGTACATAAATGATATTGAGCTGAAGGATGTTCCGACAGCAATTGGCAAAACCTTGTCTCGGAAGAAGA
 500
 ACGAAACGAAAAATGAACGTCGCGCAATATACATGTATTACTATAACTCGACTTCTACAGGCTGTCTTAACCTGTTTGGACAGGAGCCTTCTTCT

i l l l l a g g y m y i n d i e l k d v p t a l s q t l s s e e
 c f c f y l q a v l c t . m l l a . r m f r q q l d k p c p r k k r
 a f a f t c r r l y v h k . y . a e g e s d s n w t n l v l g r r

 t a k s k s a p p . l y m f s l s s f s t g v a l p c v k d e s s s
 h k q k . k c a t l h v y l l n l q l l n r c c n s l g q g r f f
 a k a k v q l r n y t c l h y q a s p h e s l l q v f r t r p l l

M	H	E	AS	H	D	N	H
B	B	C	PC	A	O	L	B
O	O	R	YR	E	E	A	O
2	2	2	11	3	1	3	2

/

GGAAATACCCATCCAGGAATATAGTGACGAAAAATTGACGGCTCAGAGTATCATGGAGTAGCAGAAAAACGGAACGAAAAATCATCTTCAACGGAAAAAAA
 600
 CCTTATGTGGTAGGTCCTTATTTTCACTGCTTTAAGTCCGAGTCTCATAGTACCTCATCGTCTTTTGGCTTGCCTTTTAGTAGAAGTGGCTTTT

e y t l q e y k v t k l d g s e y h g v a e n g t k l l f n g k k
 n t p s r n l k . r k l t a q s l m e . q k t e r k s s s t e k n
 g l h h p g l . s d e n . r l r v s w s s r k r n e n h l q r k k l

 s y v m v s y l t v f l s p e s y . p t a s l p v f l m k l p l f
 l l v g d l f l f h r f n v a . l l m s y c f v a r f d d e v s f f
 p l c w g p l y l a s f q r s l t d h l l l f r f a f . r . r f f

图 3B

H					
S					
E					
I					

H	HH			
A	SP			
E	EH			
J	II			

BH OT	H	
IB PA	S	
NO NO	E ORF 6	
II II	I	

翻译终止

700

TTAAATCAGGATTATCTGATATAAAGAAGGTGACAAGATTAAAGGCTTACITTCAGCAATCAAAGCGGATCGACGGTTAATCAAGGTTGCAAAAGTGAA
 AATTTAGTCTAAATAGACTATATTTCTTCCACTGTTCTAATTCGCAATGAAGTCGTTTAGTTTCGCTAGCTGCCAATTAGTTCACGTTTTCACCTT
 l n q d l s d l k e g d k l k a y f s k e k r l d g . s r l q k . m
 . l r l y l l . k k v t r l r l t s a n q a g s t v n q g c k s e
 k s g f l . y k r r . q d . g l l q q l k a d r r l i k v a k v n
 n f . a k d s l f s p s l l l a . k l l d f r l s p . d l n c f h
 . l l l . r l y f f t v l n l s v e e f . l p d v t l . p q l l s
 l l d p n l q y l l l h c a . p k s . c l l a s r r n i l t a f t f

HH					
SP					
EH					
II					

BH OT				
IB PA				
NO NO				
II II				

RHD 独立
转录终止区

R	HO	BH D
S	SR	IB P
A	EA	HO N
I	II	II I

P1 启动子

800

TCATTAAAAACATCACCTTCGGATCGAAGGGTGATGTTTGTGTTTTCTCAAATTGTAAGTTTATTCATTGCGTACTTTAAAAAGGATCGCTATAATA
 ACTAATTTTTGTAGTGGAAAGCTAGCTTCCCACTACAAAAAAGAGGTTTAAACATTCAAATAAAGTAACGCATGAAATTTTTCTAGCGATATTAT
 l k k h h l s d r r v m f c f s q l v a l f h e v l . k g a l .
 . l k n i t f r l e g . c f v f l k l . v y f l a y f k k d r y n n
 d . k t s p f g s k g d v l f f a n c k f l s l r e l k r i a l l
 l i l l c . r e e r l t l n q k e . l t l k n . q t s . f p d s y y
 h n f f m v k r i s p h h k t k r l n y t . k m e y k l f a r . l
 s . f v d g k p d f p a t k n k e f q l n l e n r v k f l l a l l

E			
C			
O			
D			

CH		
FP		
RA		
22		

BHM	
ASA	
NPE	
223	

//

900

ACCAATAAGGACAAATGAATAAGATTGTATCCTTCGGGGCAGGGTGGAAATCCCGACCGCGGTAGTAAAGCAATTTGCTTTAGAGCCCGTGACCCGT
 TGGTTATCTCTGTTACTTATTTCTAACATAGGAAGCCCGTCCCACTTATAGGGCTGGCGCCCATCATTTGCGTAAACGAAATCTCGGGCACTGGGCA
 p l r t n e . r l y p a g q g g n p d r r . . s t f a l e p v e r
 q . g q m n k d c i l r g r v e l p t g g s k a h l l . s p . p v
 t n k d k . l k l v s f g a g w k s r p a v v k h l e f r a r d p c
 g l l v l s y l n y g e p c p p f g s r r y y l v n a k s g t v r
 l w y p c i l l s q l r r p l t a l g v p p l t a c k s . l g h g t
 v l l s l h i l l t d k p a p h f d r g a t t f c m q k l a r a g

I	D	H		H	A
H	R	H		S	L
A	A	F		E	U
1	3	1		1	1

F	F	F	MD
O	N	O	SR
K	U	K	EA
1	H	1	11

GTGCATAAGCACGGGTGGATTCAAGTTAAGCTGAAGCCGACAGTGAAGTCTGGATGGGAGAGGATGATGAGCCGCTATGCAAAATGTTAAAAATGC
 CACGTATTCGTGGCCACCTAAGTCAAAATTCGACTTCGGCTGTACATTCAGACCTACCTCTTCTACTACTCGGGATACGTTTACAAAATTTTACG

1000

v h k h a v d e v . a e s d e e s t d e r r m m s r y a k c i k m h
 c l s t r v l q f k l k p t v k v u m g e g . . a a m q n v . k c
 a . a r g g f a l s . s r q . k s g w e k d d e p l c k m f k n a
 t c l c a t s e t . s s a s l a l r s p l l l l l r . a f h k f l
 h m l v r h l . n l s f g v t f t q l p s p h h a l c f t . f h
 h a y e r p p n l k l q l r c h i d p h a f s a s g a h l i n l l a

NA
SV
IA
13
/

H H
H H
F RHO 独立转
3 3
(录终止区)

H H
H H
L L
1 1

RBS

ATAGTGTATTTCCTATTCGGTAAATACCTAAAGCCCGAATTTTATAAATTCGGGGCTTTTTGACGGTAAATAACAAAGAGGGGAGGGAACAA
 TATCAATAAAGGATAACGCATTTTATGGATTTCGGGGCTTAAAAATATTTAAGCCCGCAAAAAACIGCCATTTATTGTTTCTCCCTCCCTTTGTT

1100

s v l a y c v k y l k p r i f y k l g a f l t v n n k r g e g n k
 l v l f p l a . n t . a p e f l l n s g l f . r . l t k e g r e t n
 . c y f l l r k l p k a p n l . l r g l f d g k . q k r g g k q
 c l t l e . q t f y r f g r i k . l n p a k k v t f l l l p s p l l
 m t n n g l a y l v . l g a n k l l e p a k q r y l v f s p l s v
 y h . k r n r l f g l a g f k k y l r p k k s p l y c f l p p f c

KE	H	A	HHD	XBMD
SA	3	L	AAD	HGBP
PR	O	U	EEE	OLON
21	2	1	131	2211

/ ORF 5 / ///

HEH	H HH
HCH	I HA
FPF	N AE
311	P 12

ATGGAAGAGTATTATATGAAGCTGGCCTTAGATCTTCCGAAGCAGGGCGAAGGACAGACCGAATCCAATCCGCTCGTGGCGCTGTTGCTGTAAGGAGG
 TACCTTCTCATAATACTTCGACCGGAATCTAGAAGCGCTTCGTCGGCTTCTGCTCGCTTAGGTTAGGGAGCAGCGCGGACAAACAGCATTTCTCGC

1200

w k s l l . s w p . l l r s r a k d r p n p l r s s a l l s . r t
 g r v l y e a g l r s c e a g r r t d r l q a a r r r c c r k g r
 m e e y y m k l a l d l a k q g e g q t e s n p l v g a v v v k d g
 h f l l l h l q g . l k r l l a f s l g f g l r e d a s n d y l v
 l p l t n y s a p r l d q s a p r l v e r l w d a r r r q q r l p r
 l s s y . l f s a k s a r a f c p a p e v a d l g e t p a t t t l s

图 3D

BHANHBNH MD	A	HHF	N	HB	B	HHN	NN	B N
AAHLABHI SR	L	PLO	L	DS	S	HGS	SL	A L
HRAAEEAN EA	U	HAK	A	ET	T	LIP	PA	N A
1124211P 11	1	131	3	1X	X	1A2	H3	1 4

// /// // /

GACAAATGTGGGAATGGGCGCCCATTTAAATATGGTGAAGCTCATGCAGAAGTTCATGCCATCCATATGGCTGGAGCACATGCAGAGGGTGGCGACAT
 CTGTTTACAGCCTTACCCCGGGTAAATTTTATACCACTTCGAGTACGCTCTCAAGTACGGTACGTATACCGACCTCGTGTACGCTCCCGCGGTGTA 1300

d k l s e w a p l . n m v k l m q k f m p a l w l e h m q r v p t f
 t n c r n g r p f k l w . s s c r s s c h p y g w s t c r g c r h
 q l v g m g a h l k y g e a h a e v h a l h m a g a h a e g a d l
 s l n d s h a g m . f l t f s m c f n m g d m h s s c m c l t g v
 v f q r f p r e n l i h h l e h t l e h w g y p q l v h l p h r c
 p c l t p l p a w k f y p s a . a a t . a m w l a p a c a s p a s m

HBH	T	F	B	N	
ASA	A	H	B	L	XH1
EAE	O	U	V	A	HBA
213	1	H	1	3	OO1

/ 211 //

TTACGTTACACTCGAACCGTGCAGCCATTACGGAACACCGCCATGTGCAGAATTGATTCAACTCTGGTATCAAAAGAGTGTTCGTGGCGATGAGA
 AATGCAATGTGAGCTTGGCAGCTCGGTAATGCCITTTTGTGGCGGTACAGCTCTTAAGTAATAGTTGAGACCATAGTTTCTCACAAGCACCCTACTCT 1400

t l h s n r a a l t e k h r h v q n . l e t l v s k e c s w r . e
 l r y t r t v q p l r k n t a m c r l d y q l w y q k a v r g d e r
 y v t l e p c s h y g k t p p c a e l l i n s g l k r v f v a m r
 n v n c e f r a a m v e f e r w t c f q n d v r t d f s h e h r h s
 k r . v r v t e c n r f f v a m h l l s . s q y . f l t r p s s
 . t v s s g h l w . p f v g g h e a n l l l e p l l l t n e a i l

D BH	M KE	BH D	H N	AM	F	E AS	S E EASH
P 1A	H SA	IB P	B L	LN	O	C PC	f C CPCII
N HM	L PR	HO M	O A	UL	X	R YR	A R PYRL
1 11	1 21	11 1	2 3	11	1	2 11	N 2 1111

/ RBS / ///

GATCCTAATCCGCTTGTGGCTCGAAGAGGATCAGCATGATGAAAGAAGCTGGCATTGAGGTAAGGCAAGGCATCCTGGCAGACCAAGGCGAGAGGCTGA
 CTAGGATTAGGCGAACCCGACCTTCTCCCTAGTCGTACTACTTCTTCGACCGTAAGTCCATTCCTTCCGTAGGACCGTCTGGTCCGCTCTCCGACT 1500

l l l r l w l e e g s a . . k k l a l r . g k a s w q t r r r g .
 s . s a c g w k r d q h d e r s w h . g k g r h p g r p g g e a e
 d p n p l v a g r g l s m m k e a g l e v r e g l l a d q a e r l n
 l r l r k h a s s p d a h h f f s a n t y p l e d q e v l r l p q
 l d . d a q p q f l a . c a a l l q c q p l p l c g p l g p p a s
 s g l g s t a p l p l m l f s a p m a t l a p m r a s w a s l s

X	H	SHH	SS	RHM	B	F	F	B	AH	H
H	N	1AA	PP	SAA	B	N	N	G	LG	A
N	L	UEE	LL	AEE	V	U	U	L	UE	E
1	1	113	11	123	1	H	H	1	12	3

//

ATGAAAAATTTCGCACITTTATGAGGACAGGCTTCCTACGTACGCTAAAAGCGGCTGCCAGCCTTGACGGCAAGATAGCTACCAGCAGGGTGACAG
 IACTTTTTAAAGAGGTGAAATACTCTGTCCGGAAGGATGCAGTGGGATTTTCCCGACGGTCGGAACCTGCCGTTCTATCGATGGTGGTCCCACTGTC

1500

m k n f c e l . g q a f i t s r . k r l p a l t a r . l p e r v t a
 . k i s a l y e d r p s v r h a k s g c q p . r q d s y q h g . q
 e k f l h f m r e g l p y v t l k a e a e l d g k l a t s t g d s
 l f f k q v k h p c a k r v d r . f r s g a k v a l y s g e r t v
 h f l e a s . s a l g e t r . a l l p q w g q r c s l . w e p h c
 f s f n r c k l l v p r g y t v a f a a a l r s p l f a v l v p s l

1	HBHMBM	AFS	ED	F
1	PIBPNBA	LNF	SD	O
H	HNONLVE	WJA	PE	K
2	111112	NNH	11	1

// /

CAAAATGGATCACGTGAGGGCTGCAAGACAGGATGCTCAGCAATACAGGAAAAACACCAAGCAATTTAGTCCGAGTTGGCACAGTGAAAGCCGACAAAT
 GTTACCTAGTGCAGTCTCCGACGTTCTGTCTACGAGTCGTTATGTCTTTTGTGTGGTTTCGTAAATCAGGCTCAACCGTGTGACATTCGGCTGTTA

1700

n g s r q r l q d r m l a n e g k h e k a f . s e l d q . k p e i
 q m d h v r g c k t g c s a l q e n t p k h f a r s w h s e s r q s
 k w l t s e a a r q d a q q y r k t h q s l l v g v g t v k a d n
 a l p d r . l s c s l l s l l v p f e v l a n . d s n a c h f g v i
 c i s . t l p q l v p h e a l c s f v g f e k l r l q c l s l r c
 l h l v d s a e l c s a . c y l l v e w l m k t p t p v t f a s l

AM	PBH	H	CH	R	THH	HD	H	H
LS	SSN	A	FP	S	ANN	SD	G	B
UE	1PF	E	RA	A	QLF	TE	A	O
11	113	3	22	1	111	21	1	1

/

CCGAGCTTAACCTGCAGACTGCCGAATGTAAACAAAACAGCCGGTTCGGGTACACTTGATACCGTACTCTCGATTCTTGAGGACGCTAAAGTGATTTGCG
 GGCTCGAATTCGACGTCTGACGGCTTACATTTGTTTGTCCGCAAGCCAGTATGAACATATGGCATGAGAGCTAAGGACTCCTGCGATTTCACATAAACGC

1800

r a . p a d e r m . q n s r f g s y l l p y s r f l r t t k . f a
 e l n l q t a e c n k t a g s g h t . y r t l d s . g r . s d l r
 p a l t e r l p n v t k q p v r v l d t v l s l p e d a k v l e d
 r a . g a s q r l y c f l r n p d y k l g y e r n r l v s l h n a
 d a s l r e v a s h l l v a p a p . v q y r v r s e q p r . l s k r
 g l k v q l s g f t v f c g t r t m a s v t s e i g s s a l t i q

图 3F

D	H HH	N		THHM		A
P	I HA	L		H1HM		L
H	N AE	A		ANAL		U
I	P 12	3		IP11		I

//

ATCAAAATAGCGCCGACATGGATTTTACGACGGCACGGCAGAGGAAAAAGAAAAAGGGCTTTCAGCTTTCGGAGTGAACATATTTACACTTGAAGC
 TAGTTTATCGCGCTGTACCTAAAAATGCTGCGGTGGCGGTCTGCTCCITTTCTTTTTCGGGAAAGTCGAAAGGCTCAGTTGTATATAATGTGAACITTG

1900

l k . r r h g f l r r h a q t r k r k n g f q l s e . t y l h l k p
 s n s a d m d f y d g t r r r g k e k t s f s f r s e h l y t . n
 q l a p t w i f t t a r a d e e k k k r l s a f g v n l f t l e t
 l l y r r c p n k r r c a c v l l l l f p k . s e s h v y k c k f
 d f l a s m s k . s p v r l r p f s f v a k l k r l s c l . v q f
 s . l a g v h l k v v a r a s s s f f f r s e a k p t f m n v s s v

H HB	XHD BH H	MWS	PHA	HA
I HS	HBP IA B	BLF	VSL	IL
4 AM	ODN HE O	OAA	UPU	HU
P 11	211 11 2	23W	2B1	31

//

CGAGCGCAITCAAATTCCTGATGTTTGAAGATCCTAGCGGAAGGCAATCATGTCGGTGTATGTGGAAGCGGTTTCAGCTGTTCACGGAAGCTTTGTC
 GCTCGGTAAGTTTAAGGACTACAAAACITCTAGGATCGCCTTCTTCGGTAGTACAGCCACATACAGCTTCGGCAAGTCGACAAGTCCCTTCGAAACAG

2000

s a f k f l m f . r s . r k k a s c r e m u k a v q l f t e a l s
 r a h s n s . c f e d p s g r r h h v g v c g r r f a c s r k l c q
 e r l q i p d v l k l l a e e g i m a v y v e g g s a v h g s f v
 g l a n l n r i n q l d . r f f a d h r h l h f e t . s n v s a k d
 r a c e f e q h k s a g l p l l c . t p t h p l r n l q e r f s q
 s r m . l g s t k f l r a s s p m m d t y t a p p e a t . p l k t

H	H	SNW	AH	H	X	H
B	H	PSL	LS	P	H	P
O	L	HPA	UE	A	N	H
2	I	1H3	11	2	1	1

//

AAAGAAGGCTGTTTCAAGAAATCATCTTCTATTTTGGCCCTAACTAATCGGAGGAACGCATGCTCCAGCTTAATCTCCGGTGAAGGTTTCAATCAA
 ITTCTTCCGACAAAAGTCTTTAGTAGAAGATAAAACGGGGATTTGATTAGCCTCCTTCCGTACGAGGTCGAATTAGAGGCCACTTCCAAAAGITAGTT

2100

k k a v f k k s s i l p l n . s e e r m l p a . s p v k v f n q
 r r l f a r n h l l f e p . t n r r n a c s q l n l r . r f s i n
 k e g c f q e l l f y f a p k l l g g t h a p s l l a g e g f q s m
 f f a t k l f d d e i k g r f . d s s r m s g a . d g e f t k l .
 l l s n e l l . r n n q g . v l r l f a h e w a l r r h l n e l l
 l s p q k . s i m k . k a g l s l p p v c a g l k l e p s p k . d

图 3 G

F	XGGCH	E	BH	FN
I	HDDFA	C	SA	OL
N	A11RE	R	1E	KA
1	32213	V	E3	13

ORF 5 翻译终止

////

TGAAGATGTCCTTATTACAATTCAGTCATATAACCCAAATCGGCGTGATAICAACTGACGGCAAAACCGACAAGGAATAGGATGGTGACCATGT

 ACTTCTACAGGGGAATAATGTTAAGTGACTATATTGGGTTTAGCCGGCACTATAGTTGACTGCCGTTTGGCTGTTTCCTTATCTACCACIGGTACA

2200

. k m s p y y n s l l . p k s a v l a n . r q n r q r n r m v t m f
 e r c p l l t i h . y n p n r p . y q t d g k t d k g l g w . p c
 k d v p l l q f t d l t q l g r d i k l t a k p t k e . d g d h v

 h f l d g . . l e s l y g l d a t i d f q r c f r c l f l l t v m
 s l h g r l v l . q y l g f r g h y . v s p l v s l p l p h h g h
 l f s t g k n c n v s l v w l p r a l l a v e f g v f s y s p s w t

H	T	M	TNH	N	SWN	HH	M	M	H	H	H
P	A	B	ANW	L	PSL	AA	S	S	N	A	A
H	O	O	OFF	A	HPA	EE	E	E	L	H	H
1	1	2	131	3	1H3	13	1	1	1	1	1

/ // /

TTACAGGAATTATCGAAGAAACAGGCACAATCGAATCCATGAAAAAGCAGGGCATGCAATGGCCTTAACTATTTAAATGCTCAAAGATTTTACAGGATGT

 AATGCTCTTAATAGCTTCTTTGTCCGTGTAGCTTAGGTACTTTTTTCGTCCCGTACGTTACCGGAATGATAATTTACGAGTTTCTAAATCTCCTACA

2300

t g l l e e t g t l e s m k k a g h a m a l t i k c s k l l e d v
 l q e l s k k q a q s n p . k k q g m q w p . l l n a q r f . r m f
 y r n y r r n r h n r l h e k a r a c n g l n y . m l k d f r g c

 n v p l i s s v p v l a d m f f a p c s l a k v i l h e f l k s s t
 k c s n d f f c a c d f g h l f e p m e h g . s n f a . l n . l l
 . l f . r l f l c l r i w s f l l a h l p r l . . l a l a k l p h

F	H	F
O	A	O
K	E	K
1	3	1

TCATCTTGGCGACAGCATTGCAGTGAACGGCATTTGTCTGACTGTCACTGATTTACAAAAATCAATTCACAGTGGATGTTATGCCCTGAACAGTCAAA

 AGTAGAACCGCTGTGCTAACGTCACTTGGCGTAAACAGACTGACAGTGACTAAAAATGTTTTTAGTTAAGTGTACCTACAATACGGACTTTGTGACGTT

2400

h l g d s l a v n g l c l t v t d f t k n q l t v d v m p e t v k
 l l a t a t q . t a f v . l s l l l q k l n s q w m l c l k q s k
 s s w r q h c s e r h l s d c h . f y k k s l h a g c y e . n a q s

 . r p s l m a t f p m q r v t v a k v f f . n v t s t i g s v t l
 n m k a v a n c h v a n t q s d a l k c f l l e c h l n h r f e d f
 e d q r c c q l s r c k d s q . q n . l f d l . l p h . a q f l .

图 3H

A	MM		H		F	XGGCIN	HH	O	B	HH	HS
L	AA		S		N	MODFAN	AA	D	S	AA	LA
U	EE		E		U	A11REL	EE	E	H	EE	AL
1	23		1		N	322131	13	1	2	13	31

GCTACGTCACCTGAATGATTTAACAAAAGGAAGCAAAGTAAATCTGGAAGAGCGATGGCGGCAACGGCCGTTTCGGAGGCCATTTGCTCTCAGGCCATG
 CGATGCAGTGACITTAATAATGTTTCTTCGTTTCATTTAGACCTTTTCGCTACCGCCGTTTTCGGGCAAGCCTCGGTAAAGCAGAGTCGGGTAC

a t s l n d l t k g s k v n l e r a m a n g r f g g h f v s g h v
 l r h . m i . q k e e k . l u k e r w r q t a v s e a l s a q a m
 y v t e . f n k r k q s k s g k s d g g k r p f r r p f r l r p e
 a v d s f s k v f p l l t f r s l a l a s f p r k p p w k t e p w
 s r . q l l . c f a a f y l q f s r h r e v a t e s a m e d . a m
 l . t v s h n l l l f c l l d p f l a p p l r g n r l g n r r l g h

AHI		II	KE		SH		MD	AAN		H	E		HD
CIA		H	SA		PB		SR	VSL		S	C		SR
CNQ		F	PR		OO		EA	AUA		E	O		EA
121		3	21		12		11	214		1	D		11

TCGACGGAAGTGGCGAAATCACACGAATTGAAGAGAAAAGCAACGCAGTTTACTATGATTTAAAAATGGACCGGTCAATTAACAAAACATTTGGTTTAAA
 AGCTGCCITGACGCCCTTAGTGTCCTTAACCTTCCTTTTCGTTGCGTCAATGATACIAAAATTTTACCTGGGCAGTAATTTGTTTGTAAACAAAATTT

d g t a e l t r l e e k s n a v y y d l k m d p s l t k t l v l k
 s t e l r k s h e l k r k a t q f t m l . k w t r h . q k h w f . r
 r r n c g n h t n . r e k q r a l l . f k n g p v l n k n l g f k
 t s p v a s l v r l s s f l l a t . . s k f l s g d n v f v n t k l
 d v s s r f d c a n f l f a v c n v l l . f h v r . . c f c q n .
 r r f q p f . v f q l s f c r l k s h n l f p g t m l l f m p k l

BH D		A	FM		HH	B	H	HH D	H		MD		MD
IB P		L	OS		NA	B	A	BB P	S		B P		B P
MO H		U	KE		FE	V	E	CO H	E		ON		ON
11 1		1	11		33	2	3	21 1	1		1 1		1 1

GGGATCAATTACTGTGGATGGCGTGAGCTTAACCATATTEGGCCTGACAGAAGACACAGTGACGATCTCTTAATACCGCATACGATCAGCGAAACGATC
 CCTAGTTAATGACACCTACCGCACTCGAATTCGTATAAGCGGAGTGTCTTCTGTGTACGTGATAGGGAATATGGCGTATGCTAGTCGCTTTGCTAG

g s l t v d g v s l t l f g l t e d t v t l s l l p h t l s e t l
 d q l l u m a . a . p y s a . q k t q . r s p . y r l r s a k r s
 g i n y c g w r e l n h l r p d r r h s d d l l n t a y d q r n d l
 p d l v t a p t l k v m n p r v s s v t v l e k l g c v l l s v l
 l s . n a h l a h a . g y e a q c f v c h r d g . y r m r d a f r d
 p l l . q p h r a s l w i r g a l l c l a s r r l v a y s . r f s

图 3 I

H DP	TH	B	H D	CT	
B PV	AM	S	B P	LA	E
O NU	OF	M	O N	AQ	C
I 11	13	I	I 1	11	P
	/			/	I

TTTCAGAAAAACGATCGGCTCTAAAGTGAATATCGAATGCGATATGATCGAAAAATATATGTATCGATTTTGCATAAGCAATGAAAAAAGACCC
 AAAAGTCTTTTTCGTAGCCGAGATTTCACTTATAGCTTACGCTATAGCTTTTATATACATAGCTAAAAACGTATTTGGGTACTTTTATTTCTGGG

fsaktlgakvnlcedmiggkymyrflhkanenkktq
 fqkkrsatk.lsnai.senlcldfclikpmklrp
 frkndrl.seyrmrydrklyvsifasq.k.dp
 kesfvlpeltflshslilplylyrnkclalsflv
 k.lfrdarfhldfahhdeflhlakqmfglfltg
 rklifars.lsyrlryarfllytdfkaylw hfyssg

I	AM	H	F	MHN	H	HO H
I	FS	ORF 4 翻	O	BSL	B	SR G
II	LE	译终止	K	OPA	O	EA A
2	21		1	2H3	2	111

AACAAACCAATACAAAAGCCTTCTTAAGCGAAAAACGGCTTTAGAGAGGAAGATTTGCATGTTTCATCGGATAGAAGCACTGGACGCTTTAAAAAA
 TTGTTTGGTAATGTTTTCGGAAGAAATTCGCTTTTGGCGAAAACTCTCTCTTCTAAACGTACAAGTAGGCTATCTCTTCGTGACCTGCGAAATTTTTT

qtltkaflesngf.rgrfacflr.kkhwtil.kk
 nkpqlqkps.aktafreedlhvssdrstgrfkkr
 tnhyksllkrkrllerklcmfhp leedal k k
 wevmvfakklslpk.lplnahkmryffccqvs.fl
 lfgncfge.afvaklssaketedsl llvprklf
 vfw.lirrlrrfrakslflqmn.glsassakff

B	H	B	SHN	H	H
B	B	S	PSL	P	S
V	O	H	HPA	A	E
2	2	2	1H3	2	1

GGCGAAGTCATCATCGTTGTAGATGATGAAGACAGAGAAAAATGAAGGAGACTTTGTGGCTCTTGGCGAGCATGCAACGCCGGAAGTCATTAACTTTAIGG
 CCGCTTCAGTAGAGCAACATCTACTTTCTCTCTCTTTACTTCTCTGAAACACCGAGAACGGCTCGTACGTTGGGGCTTCAGTAATTGAAATACC

aksssl.mmktekmketllwllpasmqrrkslltlw
 rshhrcr.rqrk.rricgscracnagsh.lyg
 gevillvddedrenegdfvelaehatpevl n f m a
 afdddnlyllf.sflfsvkhskglmcrrfdnvkh
 lrl.rqlhht.lfhllsqpeqrahlapl.s.p
 pstmmctssss.sfspektarascavgatmkl

H H	HD	FHD		H HMD		P R H
H L	BP	HN D		B P B P		L P H
L A	ON	UL E		O H O H		E H F
1 3	1 1	H 1 1		2 1 1 1		1 1 1

CGACACATGGGAGGAGCTGATCTGCACGCGGCTCAGTGAGGAAATCGCAGACAGGCTTGATCTTCACCCATGGTTGAGCATAATACAGACTCTCACCA
 GCCTGTACCCCTCTCTGACTAGACGTGGGGGAGTCACCTCCTTTAGCGTCTGTCGAACTAGAGTGGGATACCAACTCGTATTATGTCTGAGAGTGGT

r h m g e d . s a r r s v r k s q t g l i f t i w l s i l q t l t t
 d t w e r t d l h a a q . g n r r q a . s s p y g . a . y r l s p
 t h g r g l i c t p t s e e l a d r l d l h p m v e h n t d s h h
 r e m p s s q d a r r e t l f d c v p k l k v r h n i m i c v r v
 s v h s l v s r c a a . h p f r i c a q d e g . p q a y y l s e g
 a v c p l p s l q v g s l s s i a s l s a r . g i t s c l v s e . u

E	B	H HEHH	XBMD	T
C	B	B ICIA	HGBP	T
P	V	O N4AE	OLON	H
1	2	2 P712	2211	2

CACITGCATTTACCGTAAGCATAGACCATCGTGAACGAAGACAGGTATCAGCGCTCAAGAAAGATCTTTTACCGTTCAGGATTGCTGGACAGAAAIEC
 GTGACGTAAATGGCATTGGTATCTGGTAGCACTTTGCTTCTGTCCATAGTCGGGAGTCTTTCTAGAAAATGGCAAGTTCGTAACGACCTGTCGTTTAGG

l h l p . a . t l v k r r q v s a l k k d l l p f k h c w t a n p
 h c l y r k h r p s . n e d r y q r s r k l f y r s a l a g q q l r
 t a f t v s i d h r e t k t g l s a q e r a f t v q a l l d s k s
 v s c k g y a y v m t f r l c t d a s l f s r k g n l c q q v a f g
 c q m . r l c l g d h f s s l y . r e l f l k . r e l m a p c c i
 v a n v t l m a w r s v f v p l l a . s l d k v t . a n s s l l d

H	BIHSSB N	H		KH HTABH	AAF	E H D
G	SPCCES S	N		SI IHNSBA	LLH	C B P
A	AAIRCA P	L		PM AAUVE	MAH	IO H
1	J2111J 2	1		IP 11113	N111	S 1 1

GTGCCATCTGATTTACGGGTCGGGGCACATTTTCCACTGATTGCGAAAAAGGAGGTGCTGAAAAGCGGGCCATACAGAAGCTGCTGTTGATC
 CACGGTAGACTAAAAGTCGAGGCGCCGCTGTAAAAAGGTGACTAACGCTTTTTCCTCCACAGGACTTTTCGGCGCCGGTATGTCTTCGAGCACTAG

c h l i f s v r g t f l h . l r k k e v s . k a r a l q k l l l i
 a l . f s a s g a h f s t d c e k r r c p e k r g p y r s c c . s
 v p s d f q r p g h l f p l l a k k g g v l k s a g h t e a v d i
 h w r l k l t r p v n k w q n r f l s t d q f a r a m c f s s n i
 r a m q n e a d p a c k e v s q s f l l h g s f r p g y l l q q q d
 t g d s k . r g p c m k g s l a f f p p t r f l a p w v s a a t s

图 3 K

H A	BXMD	EHAS	NCHN	A	B	H	D	A	SHBH
I L	IHBP	CGPC	LFPA	H	B	B	D	L	SGAS
H U	NOON	RAYR	ARAE	A	V	O	E	U	TJHP
3 1	1211	2111	4221	2	2	2	1	1	1A22

TTGCTGAAGCTTCCGGATCTCCAGGAGCCGGCGTCATTTGTGAATTATGAATGAAGACGGAACGATGGCGAGAGTGCCTGAGCTCATTGAAATTCGGA
 ACGGACTTGAACGCTAGAGGTCTCGGCCGAGTAAACACTTTAATACTTACTTCTGCCCTTGTACCGCTCTCACGGACTCGAGTAACCTTAACGCTT
 l k l a d l q e p a s f v k l . m k t e r w r e c l s s l k l r k
 c . s l r i s r s r r h l . n y e . r r n d g e s a . a h . n c e
 a e a c g s p g a g v l c a l m n e d g t m a r v p e l l e l a k
 k s f s a s r w a g a d n t f n h l f v a r h r s h r l e n f n r
 q q l k r l e l l r r . k h f . s h l r f s p e l a q e . q f q s
 r a s a q p d g p e p t m q s l i f s a p v i a l t g s s m s l a f

H SH	BMD	H	H	H	T	B	H	F
S FP	CBP	S	N	A	A	B	I	N
E AH	LOW	E	F	E	Q	V	N	U
1 11	111	1	1	3	1	1	2	H

AAAGCATCAATTAATAATGATCACCATTAAAGGATTGATTCAAFACCGTTACAATCTGACAACACTTGTGAGCGGTGAAGITGACATTACGCTGCCACT
 TTTCTAGTTAAATTTTACTAGTGGTAATTCCTAACTAAGTTATGGCAATGTTAGACTGTTGTGAACAGCTCGCACTTCAACTGTAATGCCGCGATGA
 s l n . k . s p l r l . f n t v t l . q h l s s v k l t l r c l l
 k a s l k n d h h . g f d s l p l q e d n t c r a . s . h y e a y .
 k h q l k m l t l k d l l q y r y n l t t l v e r e v d l t l p t
 f l m l . f h d g n l l q n l v t v l q e c k d l t f n v n r q r s
 f a d l l f s . w . p n e l g n e d a l v q r a h l q c . a a .
 f c . n f l l v m l s k l . y r . l r v v s t a r s t s m v s g v

F H	H	NH	NCH
I S	N	SL	GFP
H E	L	PA	ERA
1 1	1	H3	222

GATTTTGGGACATTTAAGGTTTATGGATACACAAATGAGGTAGATGGAAAGAGCATGTCCGATTTGTGATGGGAGATGTCCGTTCCGAGAAGAACCGG
 CTAAAACCGCTGTAATTCCTAAATACCTATGTGTTTACTTCCATCTACCTTTTCTGCTACAGCGTAAACACTACCCCTACACGCGCAAGCCTCTTCTTGGCC
 l l g h l r f m d t q m r . m e k s m s h l . w e m c r s e k n r
 f v d l . g l w l h k . g r w k r a c r l c d g r c a v r r r t g
 d f g t f k v y g y t n e v d g k e h v a f v m g d v p f g e e p v
 l k p c k l n l s v c l l y l a f l m d e k h h s l h r e s f f r
 q n q s m . p k h l c l h p l h f l a h r m q s p l h a t r l l v p
 s k p v n l t . p y v l a t s p f s e t a n t l p s t g n p s s g

图 3L

H AA HHS B	B H MA H	B RAA F PNAF FBTKH	F
B VS PCC S	S A AF P	B SVS H VSLH NBSHI	H
U AU AIR H	H E EL H	V RAU U UPUU UVAPAH	U
2 21 211 1	2 3 23 1	1 221 H 2B1H H111P	H
/ /	/	// ///	/

TATGGTCCGGTGCATTGAGATGTCACAGGTGACGTGTTGGCTCTCATCGCTGTGATTGCGGACCGCAGCTGCACGGCGCGCTGAACCAAAATTC
 ATAACGAGGCCACGTAAGTCTTACAGAGTGTCCACTGCCAAACCGAGAGTAGGACACTAACGCGTGGCGTCGACGTGCGGCGCGACITGGTTTAAACG

y w s g c i q n v s q v t c l a l l a v l a d r s c t p r . t k l p
 i g p g a f r m s h r . r v w l s s l . l r t a s e r r a e p n c
 l v r v h s e c i t g d v f g s h r c d c g p q l h a a l n q l a
 y q d p h m . f t e c t v h k a r m a t l a s r l q v g r q v l n
 l i p g p a n l i d . l h r t q s e d a h n r v a a e r r a s g f q
 t n t r t c e s h r v p a t n p e . r q s q p g c s c a a s f w l a

HODSRB	HN R H N M	HS	M H A H A
ASSESS	GS S I H N	SF	S I L I L
EAACAA	IP A N A L	EA	E H U H U
3111JJ	A2 1 P 1 1	1H	1 31 31
////	/		

CGCAGAAGGCCGTGCAGTCTCTGTACTTGGCCCAAGAGGACGAGGCATCGGTTTAACTCAATAAAATAAAGCTTATAAGCTTCAGGAACAAGGCTAT
 GCGTCTTCCGGCACCTCAGGAGACATGAACGGGTCTCTCGTCCGTAGCCAAATTAGTTATTTAATTTTCGAATATTCGAAGTCTTGTTCGGATA

q k a v e c s e t c a k k d e a s v . s i n . k l i s f r n k a m
 r r r p w s a p v l a p r r t r h r f n q . l k s l . a s g t r l .
 a e g r g v l l y l r q e g r g l g l i n k l k a y k l q e q g y
 g c f a t s h e q v q a l f a s a d t . d l f . f a l l k l l t a l
 r l l g h l a g t s a g l l v l c r n l . y l l l k y a e p v l s
 a s p r p t s r y k r u a p r p m p k l l l n f a . l s . s c p .

M	B H HH H	M HBXMD	SHH	F BSB	B11H	H
H	G I HA H	B PIHBP	FGS	H SES	SPH	I
L	L H AE F	O ANOON	A1P	A ACA	PAL	H
1	1 P 12	2 21211	HA2	1 J1J	221	3
/		/		/		

GACACCGTAGAAGCCAAATGAGGGCGCTTGGATTCTTGGCGCATCTTGGCAACTATGGCATCGGAGCACAAATTTACGGGACCTCGGTGTCCGGAATAIGA
 CTGTGGCATCTTGGTTACTCCGGGAACCTAAGAACGGCCTAGAAGCGTTGATACCGTAGCCTCGTGTTTAAATGCGGTGGAGCCACAGGCGTTATACT

t p . k p m r r l d a c r l f a t m e a s h k f y a t s v s g l .
 h r r s q . g a w l l a g s e q l u h r s t n f t r p r c p e y e
 d t v a a n e a l g f l p d l r n y g l g a q l l r d l g v r n m k
 v g y f g l l r k s e q r i k a v l a d a c l n . a v e t d p l h
 h c r l l w h p a q l r a p d e c s h c r l v f k v r g r h g s y s
 s v t a a l s a s p n k g s r r l . p m p a c l k r s r p t r f l

3 M

A	H	I	SHH	B	P	H D	F	H
L	H	H	TAA	G	L	H D	H	H
U	F	A	UEE	L	E	F E	U	L
I	3	I	113	I	I	I I	H	I

//

AGCTTTGACGAATAATCCGCGAAAAATCCGAGGCTTGAAGGCTACGGACTCAGTATTTTCAGAAAGAGTCCCGCTTCAATGGAGGCGAAAGAACACAA
 4000
 TCGAAAACGTCTATTAGGCGCTTTTAGCGTCCGGAAGTTCGATGCGTGAGTCATAAAGTCTTCTCAGGCGAAGTTACCTCCGCTTCTTGIGIT

s f . r i l r e k s q a l k a t d a v f q k e c r f k w r r k n t i
 a f d e . s a k n r r p . r l r t q y f r k s a a s n g g e r t q
 l l t n n p r k l a g l e g y g l s l a e r v p l q m e a k e h n

 l k q r l l r a f d c a k f a v s e t n . f s h r k l h l r f f v
 a k s y d a f f r l g q l s r v . y k l f l e a a f p p s l v c
 f s k v f l g r f l a p r s p . p a l l e s l t g s . l a e f s c l

S	AM	H	MH	H F
S	LA	A	AM	D O
P	UE	H	HF	E K
I	11	I	11	11

ORF 3 翻译终止

TAATAAATATTTGCAAAACCAAAATGAACAAGCTAGGTCATTTACTTCTAATCACAATATCACAAAAAGGATGGGAATCATATGAATATCATAC
 4100
 ATTTTTATAACGTITGGTTTTACTTGTTCGATCCAGTAAATGAAGTAAAGATTAGTGTTTATAGTGTTTTTCTACCCTTAGTACTTATAGTATG

k n l c k p k . t s . v l y f l a n h k y h k k g w e s y o y h t
 . k l f a n q n e q a r s f t a f l l t n l t k k d g n h m n l i q
 k k y l g t k m n k l g h l l h f . s q l s q k r m g l l . l s y

 i f f l q l g f h v l . t m . k m e l . l y . l f p h s d y s y . v
 y f l n a f w f s c a l d n v e n r l v f l v f f s p f . l f l m
 l f y k e v l i f l a p . k s . k . d c l d c f l l p l m h l d y

R	E	H	H	H	H B	AF	EDE	S B	H HF
S	C	S	H	S	B B	LN	SOC	F B	I HN
A	P	E	F	E	O V	UJ	PEI	A V	H AU
I	I	I	I	I	2 I	1H	115	H I	P 1H

/

AAGGAAATTTAGTTGGTACAGGCTTTAAAAATCGGAATCGTAGTAGGAAGATTTAAATGATTTTATTACGAGCAAGCTGCTGAGCGGAGCAGAAGATGCCGT
 4200
 TCCCTTAAATCAACCATGTCAGCAATTTAGCCTTAGCATCATCTCTAAATTACTAAAAAATGCTCGTTCGAGCACTCGGCTCGTCTCTACGCGA

r k f s w y r s . n e n r a r k l . . f y y e q a a e r s r r c a
 g n l v g t g l k l g l v v g r f n d f l t s k l l s g a e d a l
 k e l . l v q v l k s e s . . e d l m l l l r a s c . a e q k m r c

 l l n l q y l d . f r l r l l l l . h n . . s c a a e r l l l h a
 c p f k t p v p r l l p l t t p l n l s k l v l l s a l p a s e a s
 l s l . n t e t k f d a d y y s a k l i k n r a l q q a s c f l r

图 3N

MD	N A	HE AS H H	B	F	S S
BD	L C	LC PC I H	B	N	O F
OE	A C	AR YR W A	V	U	O A
21	3 1	42 11 P 1	1	H	1 N

/

GCTCAGACATGGCGTAGACACAAAIGACATTGATGTGGCTTGGGTCCAGGCGCATTTGAAATACCGTTTGGTCGCAAAAAATGGCGAAACAAAAAA 4300

CGAGTCTGTACCGCATCTGTGTTTACTGTAAGTACACCGAAACCAAGGTCCGCGTAACTTTATGGCAACGACGCTTTTTTACCGCCTTTGTTTTTT

a q t w r r h k . h . c g l g s r r l . n t v c c e k n g g n k k l

l r h g v d t n d l d v a w v p g a f e l p f a a k k m a e t k k

s d m a . t q m t l m w l g f q a h l k y r l l r k k w r k q k n

a . v h r l c l h c q h p k p e i r m q f v t q q s f f p p l l

s l c p t s v f s m a t a q t g p a n s l g n a e f f l a e v f f

q e s m a y v c l v n l h a p n w a c k f y r k s r l f h r f c l l

NH	NH	E	B	AF	11H S F
SN	I H	C	B	LN	11H F H
PL	N A	O	V	UU	AHA A U
21	P 1	O	1	1H	1P1 H H

/

TATGATGCTATTATCACATTGGGCACGTGCATCAGAGGGCGCAACGACACATTACGATTATGTCTGCAATGAAGTGCAAAAGGCATCGCGCAAGCAGCAA 4400

ATACTACGATAATAGTGAACCCGTGACAGTAGTCTCCCGGTTGCTGTGTAATGCTAATACAGACGTTACTTCGACGTTTTCGATAGCGGTTCGTGCTT

. c y y h l g h c h q r r n d t l r l c l q . s c k r h r a s k

y d a l l t l g t v l r g a t t h y d y v e n e a a k g l a q a a n

m m l l s h u a l s s e a q r h l t l m a a m k l q k a s r k q q

l h h . . m p e q . . l r l a v n r n h r c h i q l l c r a l l l

y s a l i v n p v t m l p a v v c . s . t q l s a a f p m a c a a

l l s n d c q a s d d s a e r c m v l l d a i f a c f a d r l c c

B	R	H	T	T	CHN
B	S	A	A	A	FPA
V	A	E	O	O	RAE
1	1	3	1	1	221

ACACTACTGGTGTACCTGTATCTTTTGGAAITGTAACAACIGAAAACATCGAACAGGCTATCGAGCGTGCGGGCACA AAAAGCGGGCAACAAAGGTGTAGA 4500

TGTGATGACCACATGGACAGTAGAAAACCTTAACATTGTTGACTTTTGTAGCTTGTCCGATAGCTCGACGGCGGTGTTTTGCGCCGTTGTTTCCACATCT

h y w c t c h l u n e n n . k h r t g y r a e r h k s g q q r c r

t t g v p v l f g l v t t e n l e q a l e r a g t k a g n k g v d

t l l v y l s a l e l . q l k e s n r l s s v p a q k r a t k v . l

c . q h v q . r q f q l l q f c r v p . r a h r c l l p c c l h l

l v v p t g t m k p i t v v s f m s c a l s r a p v f a p l l p t s

v a s t y r d d k s n y c s f v d f l s d l t g a c f r a v f t y

图 30

B-核黄素合成
酶翻译终止 P3 启动子

B HD SR EA 11
 S SR EA 11
 T EA 11
 X 11

HD SR EA 11
 H P H 1

TTGTGCTGTTTCTGCCATGAAATGGCAAAITTAACCGCTCATTTGAATAATTTGCTGAAACAGTTTAAAAATATGGCAAAATGATATAATGTGAGA
 AACACGACAAAGACCGTAACTTACCGTTTAAATTTGGCGAGTAAACTTATTAACGACCTTTTGTCAAATTTTATACCGCTTTTACTATATTACACICT

l c c f c h . n g k f k p l l . l l c . k q f k n m a k m i . c e k
 c a v s a l e m a n l n r s f e . f a e n s l k l v r k . y n v r
 v l f l p l k w q l . t a h l n n l l k t v . k y g e n d l m . e
 n h q k q w q f p l n l g s m q l l q q f c n l f l a f l l y h s
 q a t e a m s l a f k f r e n s y n a s f l k f l h r f h y l t l
 l t s n r g n f h c l . v a . k f l k s f v t . l y p s f a l l h s

BM D H H M H H
 18 P H S H S H
 HO H F E L E F
 11 1 3 1 1 3

AAACGGATCACCATTCTGATCCGTTAATAGCAGACTGGACATTTTGGATATAGAGGGGTTTTATGTTAATTCGTTATAAAAATCGTTTGAAAAAGAIT
 ITTGCTAGTGGATAAGCATAGGCAATTATCGTCTGACCTGTAAAACTATATCTCCCAAAAAATACAAATTAAGCAATATTTTTAGCAAACTTTTCTAA

t d h l f v s v n s r l d l l d l e g f l c . f v i k n r l k r t
 k r l t y s y p l l a d w t f w l . r g f y v n a l . k l v . k d c
 n g e p l r l r . . q t g h f g y r g v f m l l r y k k s f e k l
 f v s . r n t d t l l l s m k s l a p n k h . n t l f f r k f l n
 f r l v . e y g n l a s q v n q i y l p k . t l e n y f l t q f s
 f p d g l r l r . y c v p c k p y l p t k i n l r . l f d n s f l

H EAM F A E B H E A
 H CFS NL C B S C L
 F PLE UU 1 V E P U
 3 121 H 1 S 1 1 1 1

GCGATGGGGCTTCTTTCTTTATGCCGAATGAAAAAGACCTTAAGCAGCTTCAGCAGACAATTAAGGACTACGAAACGGATACAGACCGCCAGCTCTTTC
 CGCTACCCCGAAGAAAGCAAAATACCGCTTACTTTTCTGGAATTCGTCGAAGTCGTCGTTAATTCCTGATGCITTCCTAIGTCGGCGTCGAGAAAG

r u g l f r l c r m k k t l s s f s r q l r t t k r l q t a e s f
 d g a s f v y a e . k r p . a a s e d n . g l r n g y r p p a l s
 a m g l l s l m p n e k d l k q l q q t i k d y e t d t d r q l f l
 r h p k k r k h r l f f v k l l k l l e n l v v f r l e v a l e k
 q s p a e k t . a a h f l g . a a e a s l . p a r f p y l g g a r e
 a i p s r e n i g f s f s r l c s . c v i l s . s v s v a r w s k

图 3P

H	H	E	H	T	P	H	H	O	XMD	HB	H
H	H	C	H	A	L	H	H	O	HBP	PI	H
L	L	R	F	Q	E	L	F	E	ODM	AM	F
I	I	V	I	I	I	I	I	I	211	21	I

/

TTTGGAAAGAGGACGAGGATATCGTCGGAGCAATCGGAGTCGAAAAAAGGATTCTGAGGTTGAGATCCGGCATATCAGTGTGAATCCCTTCATCGCCA
 AAACCTTTCTCCTGCTCCTATAGCAGCCTCGTTAGCCTCAGCTTTTTTCTAAGACTCCAACCTAGGCGGTATAGTCACACTAGGAAGAGTAGCGGT

4900

l g k r t r l s s e q e e k k r l l r l r a g l a v . i l l l a i
 l e r g r g y r r r n r s r k k g f . g . d p a y q c e a f s s p
 w k e d e d l v g a l g v e k k d s e v e l r h l s v n p s h r h
 k p f l v l l d d s c d s d f f l l r l n l d p m d t h l r r m a
 k s l p r p y r r l l r l r l f p n q p q a g a y . h s d k e d g
 r q f s a s s i t p a l p t s f f s e a t s l r c l l t f g e . r w

H	S	HO	F	SR	H	H	A	AT	H
H	F	SR	O	CS	L	S	L	SA	A
F	A	EA	K	AA	A	E	U	UQ	E
I	H	II	I	II	I	I	I	21	2

/

TCAAGGAATCGGAAACAGATGATGGATGCTTTAAAGCATTTATTCAAACGCAAGTACTGGTTCCAAATGAATTAACGCAGAGCTTTTCGAACGTTGT
 AGTTCCTTAGCCTTTGTCTACTACCTACGAAATTCGTAATTAAGTTTTGCGTTCATGACCAAGGTTACTTAATTCGGTCTCGAAAAAGCTTCCAACA

5000

k e s e n r . w m l . s l y s k r k y w f q m n . r r e f s n v v
 s r n r k t d d g c f k a f l q n a e t g s k . l n a e l f r t l s
 q g l g k q m m d a l k h l f k t q v l v p n e l t q s f f e r c
 m l s d a f l h h l s . l m . e f r l y q n u l f . r l a k e f t t
 d l f l v s s p h k l a n l . f a l v p e l h i l a s s k r v n
 . p l o f c l l s a k f c k n l v c t s t g f a n v c l k k s r q

E	BHD	HM	A	BHD	A	B	E	HB
C	IBP	HS	L	CBP	L	S	C	AS
I	HOH	LE	W	LOH	W	M	O	EA
S	III	CAF	2	II	N	III	N	2

/

CAAGGTCAGCAGGATCAAGACATTTATACAAATAATTAAAGCAGAGGCTGTGATCAGTCTCTGCTTTTTTCTGCGTCTATTTCTTTTCACGTTACAG
 GTTCCAGTCGTCTAGTTCGTAAAGTATGTTTATTAATTCGTCTCCGACACTAGTCAGAGACGAAAAAAGACGCAAGATAAAGAAAAAGTCAAGTGC

5100

ORF 1 TRANSLATION STOP

k v s r l k t f h t l l k q r l . s v s a f f s a f y l l f t l t
 r s a g s r h f l q . l s r g c d q s l l f f l r a l s f s r s r
 q g q q d q d l s y n n . a e a v l s l c l l f c v l l f l h v h g
 l t l l l i v n . v l l l c l s h d t e a k k e a n . k k k v n v
 d l d a p d l c k m c y n l l p q s . d r s k k r r e l e k e r
 . p . c s . s m e y l l . a s a t l l r q k k k q t r n r k . t .

图 3 Q

AM A	FH D	B	S T	DB A	HW	F	SNA	B	IMH
HA A	OB P	I	P A	DB P	GS	N	NLC	S	ABH
AE T	KO W	N	O Q	EY L	IP	U	AAC	H	QOF
22 2	11 1	1	1 1	11 1	A2	H	131	1	113

/ /

GATGACGTCAGTCCGATCCCGCAAACGGTGTGTCGATAAGAAATATGTTGCTGAGTGCACTGGGCTGCCCCATGTATACATTTTTTTCCTGCATTG

CTACTGCAGTCAGGCTAGGGCGTTGCCACAACAGCTATTCTTTATACAACGACTCAGTGACCGGACGGGGTACATATGAAAAAAGGACGTAAGC

5200

d d v s p l p q t v f v d k k y v a e c t g l p p c i l f f p a f d

m t s v r s r k r c l s i r n m l l s e l g c p h v y f f f l h s

r q s d p a n g v e r . a l e c . v h w a a p m y t f f a c l r

s s t l g l g c v e n t a l f y t a s h v p a g g h l s k k g a n

i v d t r d r l r h k d l l f i n a l e s p q g w t y k k k r c e

p h r . d s g a f p t q r y s i h q q t c q a a g m y v k k e q m r

D B	SNH	B	H	B	ST	T
P I	PSL	S	N	S	PA	A
N N	HPA	R	L	H	OQ	O
1 1	1H3	1	1	2	11	1

//

ATCCTGCATGCTTCTCCAGTTTCTCATCTTTGATTGGCAGTATAATGCTTTTATAGCGAGAGCGGTTTCGATTGTTTCGTAACCGATTGCATAAGTT

TAGGACGTACGAAGGAGGTCAAGAGTAGAACTAACCGTCATATTACGAAATATCCGTCTCTGCCAAGCTAAACAAGCAITTCGCTAACGTTATCAA

5300

p a c f l i q f l l f d w q y n a f l g r d g f d l f v n r l h k f

i t h a s s f a s l l g s l m l l . a e t v a l e s . t d c l a s

s c m l p p v s h t . l a v . c f y r q r r f r f v r k p l a . v

s g a h k r u n r m k s q c y l a k l p l s p k a k n t f r n c l n

i r c a e e l k a d k l p l l l s k y a s v t e l q e y v s q m l

d q m s g g t e . r q n a t y h k . l e l r n r n t r l g i a y c

GCH	BHMO	H D	TB	NSH	S H BH	H	F	BSBNH
DFA	CLBP	B D	AB	CCP	F H SP	B	O	SESCCP
IRE	LAOH	O E	OV	IRA	A F PA	O	K	ACATRA
213	1311	2 1	12	112	N 1 22	2	1	J1J112

/ //

CGAGCAACGGCCATGATCAAGCCCTAAGTCTTCGACTGCCCGGTGTTCTGCTTGAAGAATCCGGATGCTGTTGCCATCAGTCTTTTTGCCCGGGCTGT

GCTCGTTGCCGCTACTAGTTCGGGATTGAGAAGCTGACGGGCCACAAGCGAAGCTTCTTAGGCTACGACAAGGGTAGTCAGAAAAAGGGGGCCGACA

5400

e q t a m l k p . v f d c p v f c l k n p d a v r h q a f c p g c

s k r p . s s p k s a s t a r c s a . r l r m l f a l s l f a p a v

r a n g h d q e l a l r l p g v l l e e s g c c s p a v f l p r l y

s c v a m l l g . t k e q g t n q k f f g s a t r w . d k q g p q

e l l r g h d l g l d e v a r h e a q l i r l s n a m l r k a g a t

r a f p w s . a r l r r s g p t r s s s d p h q e g d t k k g r s

图 3R

A		H	BXMD	H	S
L		B	IHP	B	S
U		O	NOON	O	P
I		Z	1211	Z	I

/

ATTCGCCCTTCTGTGATGATATAAAGCCACTGCAAGCTGAATAAAGCCCAATAGCGTTTTCGTTTCTTTGGCGGATCTTCCTTCCAATATTCTTCT
 5500
 TAAGACGGAAGACACTACTATATTTCCGTGACGTTCCACTTATTTCCGGTGGGTTATCCGCAAAAGCAAGAAACCCGCTAGAAGGAAGGTTATAAGAAGA

l l p a v m i . s h c k l n k a h p l a f s f l w r l f l p l f f .
 f c l l . . y k a t a s . l k p t q . r f r f f g g s f q y s s
 s a f c d d l k p l q a e . s p p n e v f v a l a d l p s n l l l

 l r g e t l i y l w q l e f l a u g l a n e n r q r l k r g l n k
 n q r r h h y l a v a l q l f e v w y r k r k k p p d e k w y e e
 y e a k q s s l f g s c a s y l g g l l t k t e k a a r g e l l r r

S	HB	N	HA
S	LS	L	IL
P	AN	A	HU
I	31	3	31

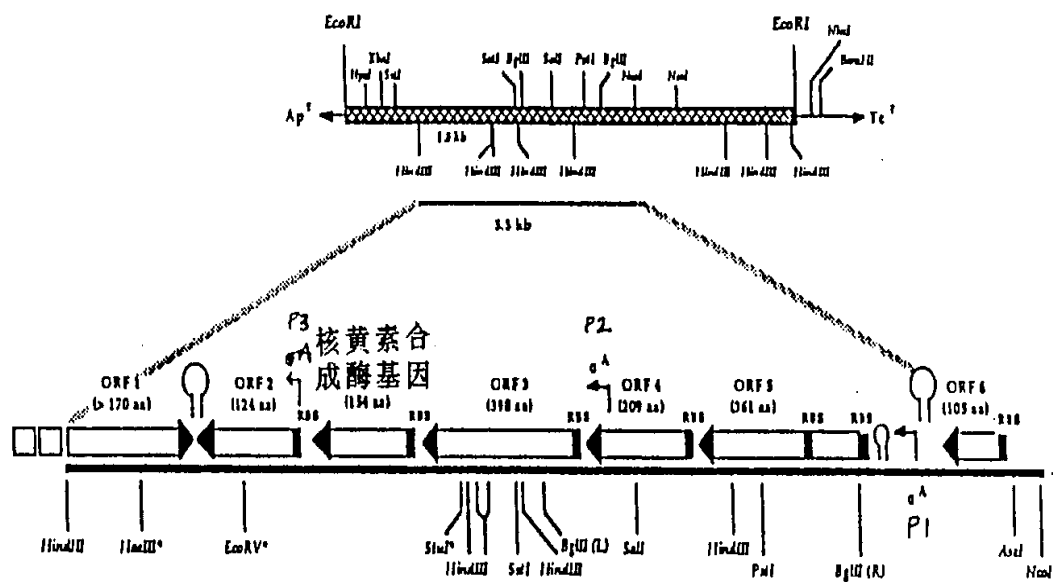
AATATTTTCATGGCATTCAAATAATCCCGTGTCCGATGAACTCAACGAGATAATCTATATAAGCTT
 5567
 TTATAAGTACCGTAAGTTTTATTAGGGCACAGCGTACITTCAGTTGCTCTATTAGATATATTCGAA

y l m a f k l i p e r m k l n e l l y l s f
 n l s w h s k . s r v a . n a t r . s l . a s
 f f h g l q n n p v s h e t q r d n l y k l

 . y k m a n l l i g h r m f a l s l l . l l k
 l i s h c e f y d r t a h f e v l y d l y a q
 l n . p m . f l g t d e a v . r a l r y l s a

图 3 S

图 4



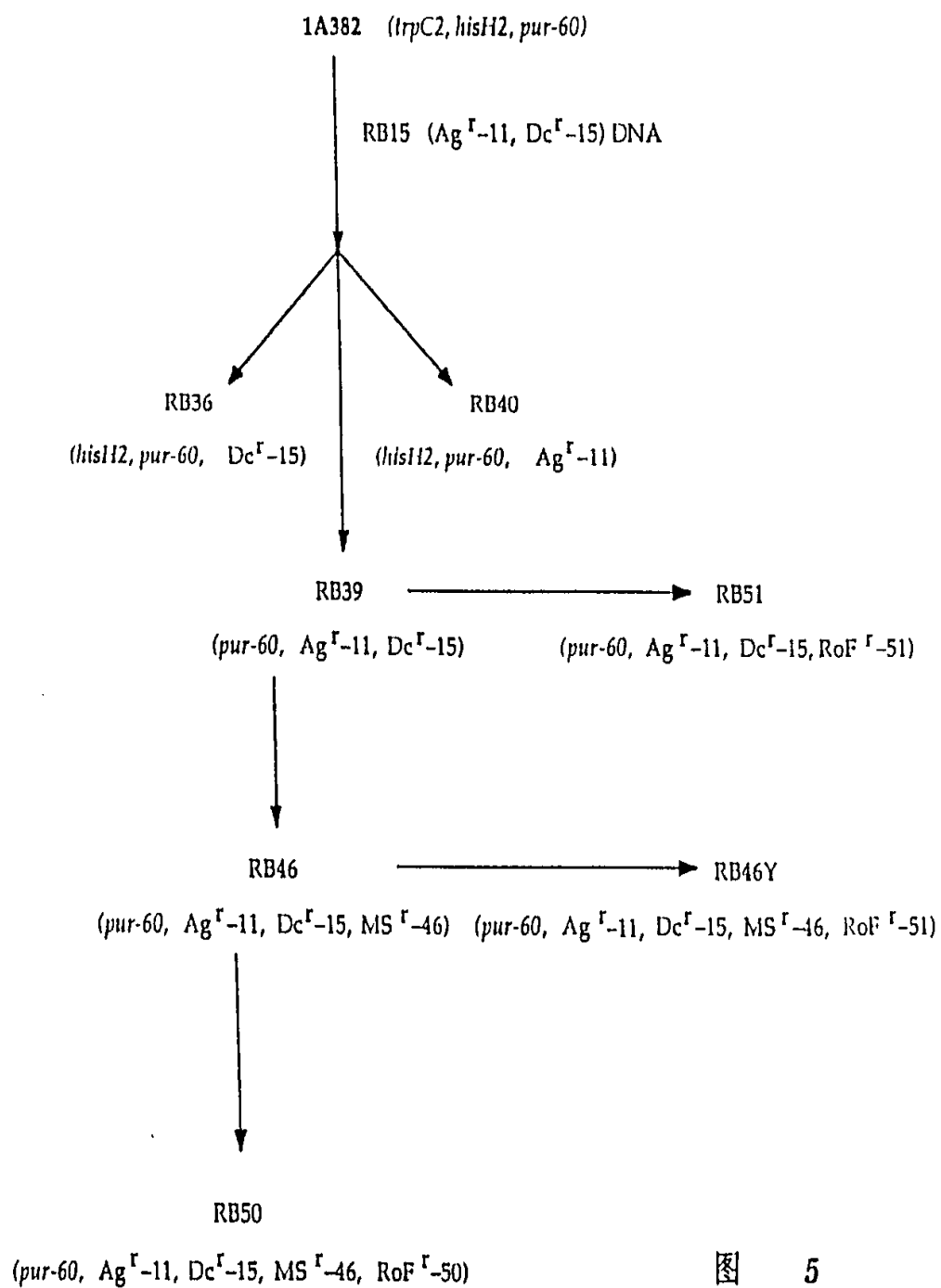


图 5

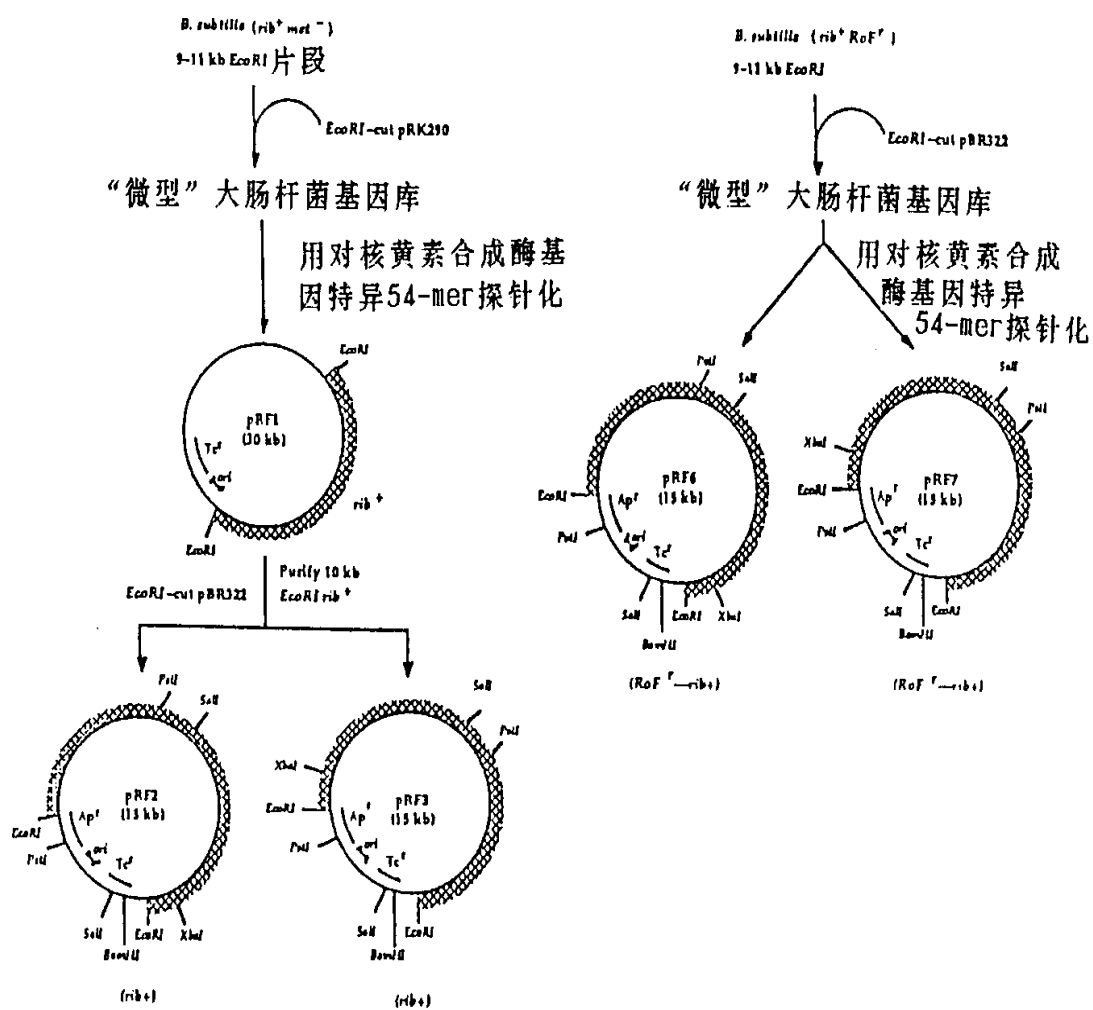


图 6

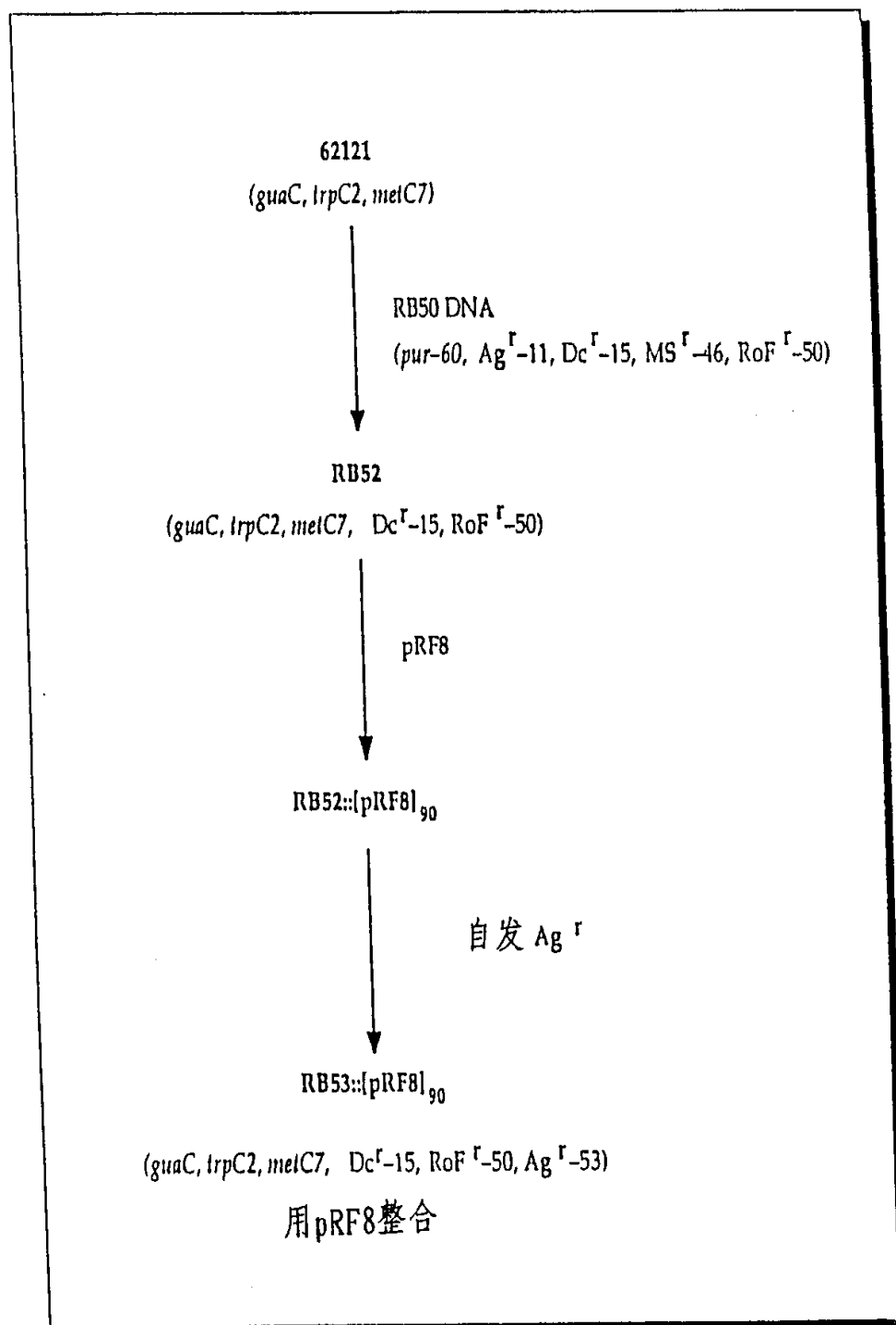


图 7

图 8

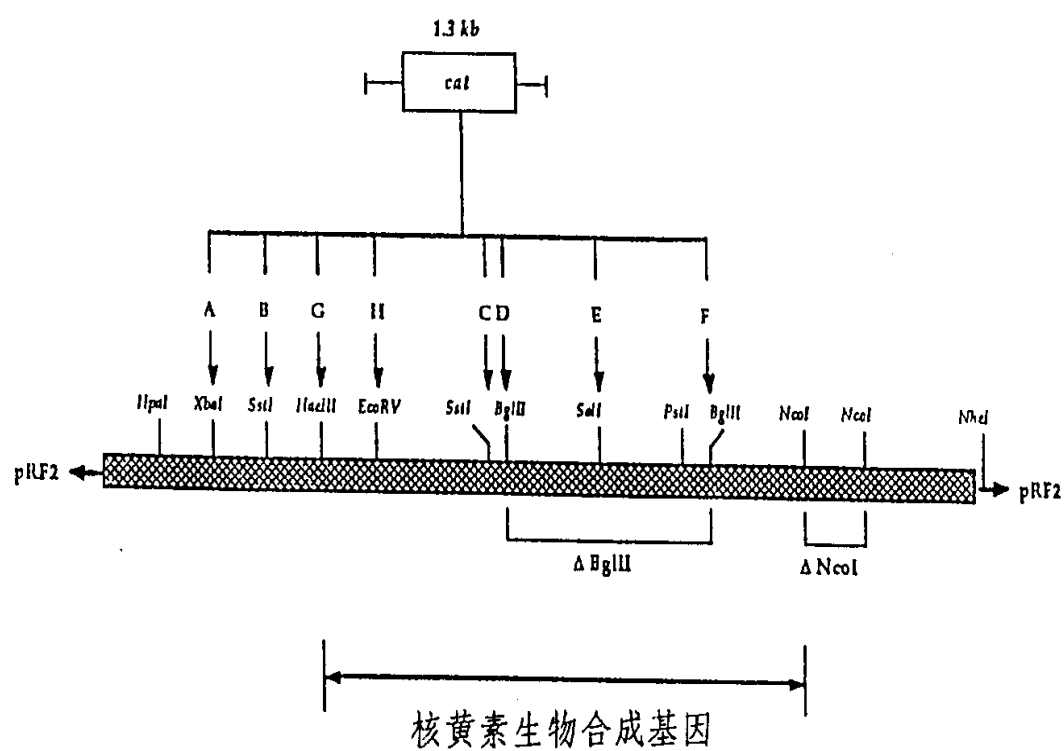
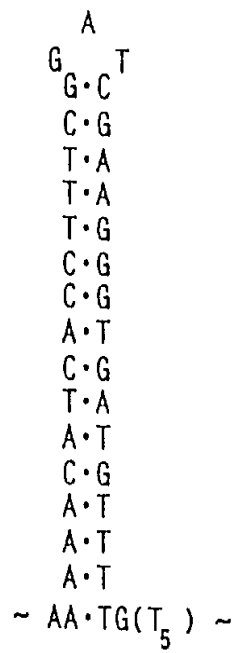


图 9

rib P1的上游

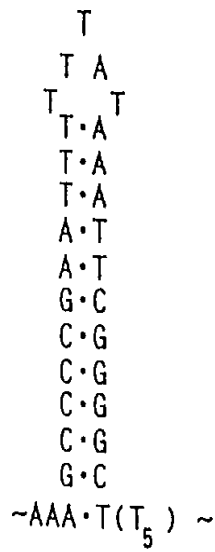
在5'前导mRNA中

在rib 操纵子的3'端



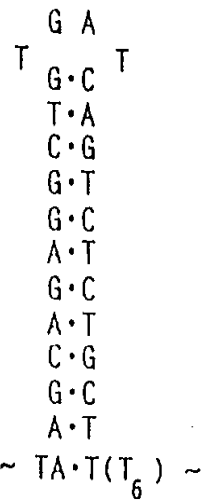
#708

#742



#1034

#1062



#5037

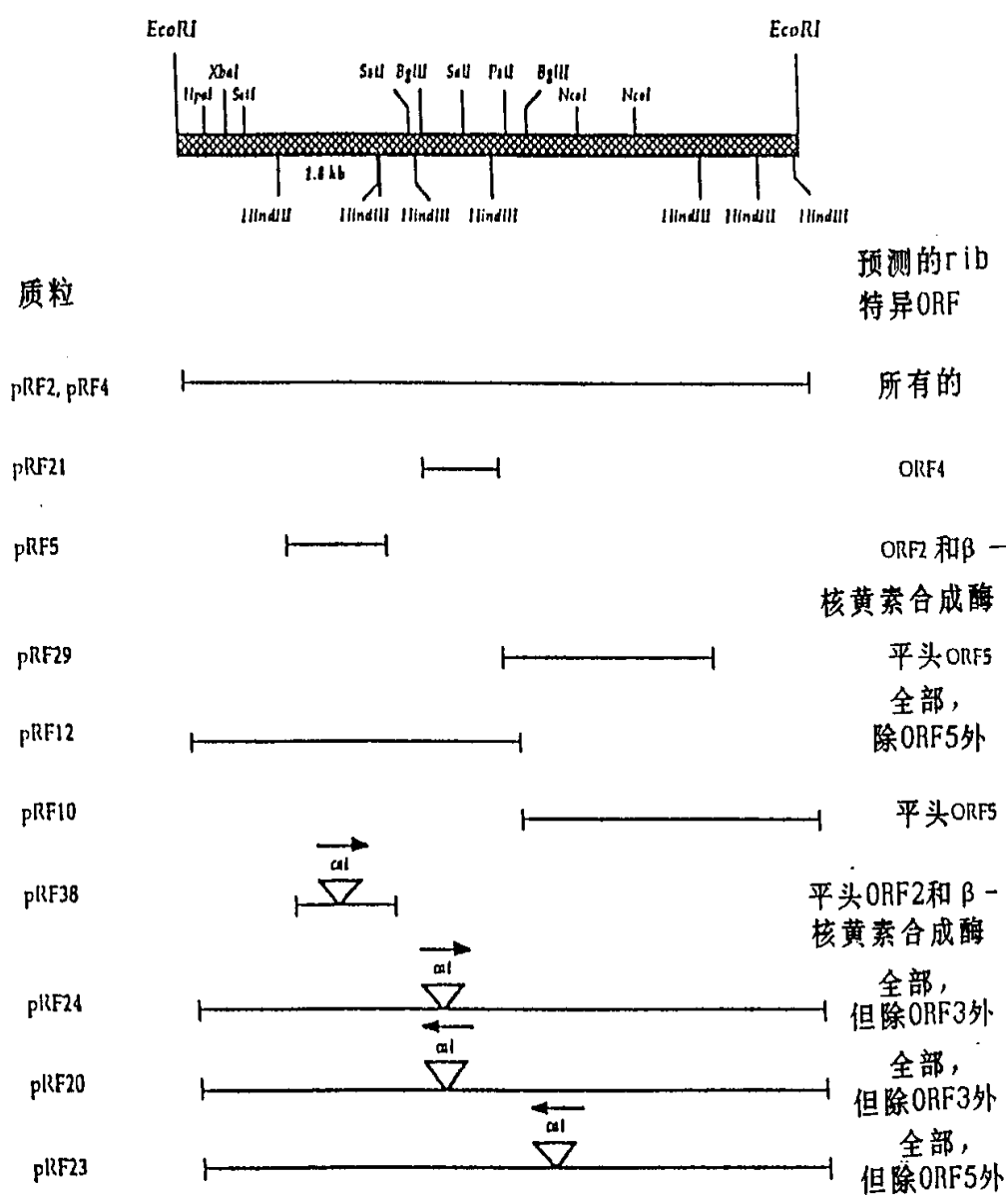
#5064

$\Delta G = -20$ 千卡 / 摩尔

$\Delta G = -26$ 千卡 / 摩尔

$\Delta G = -16.5$ 千卡 / 摩尔

图 10



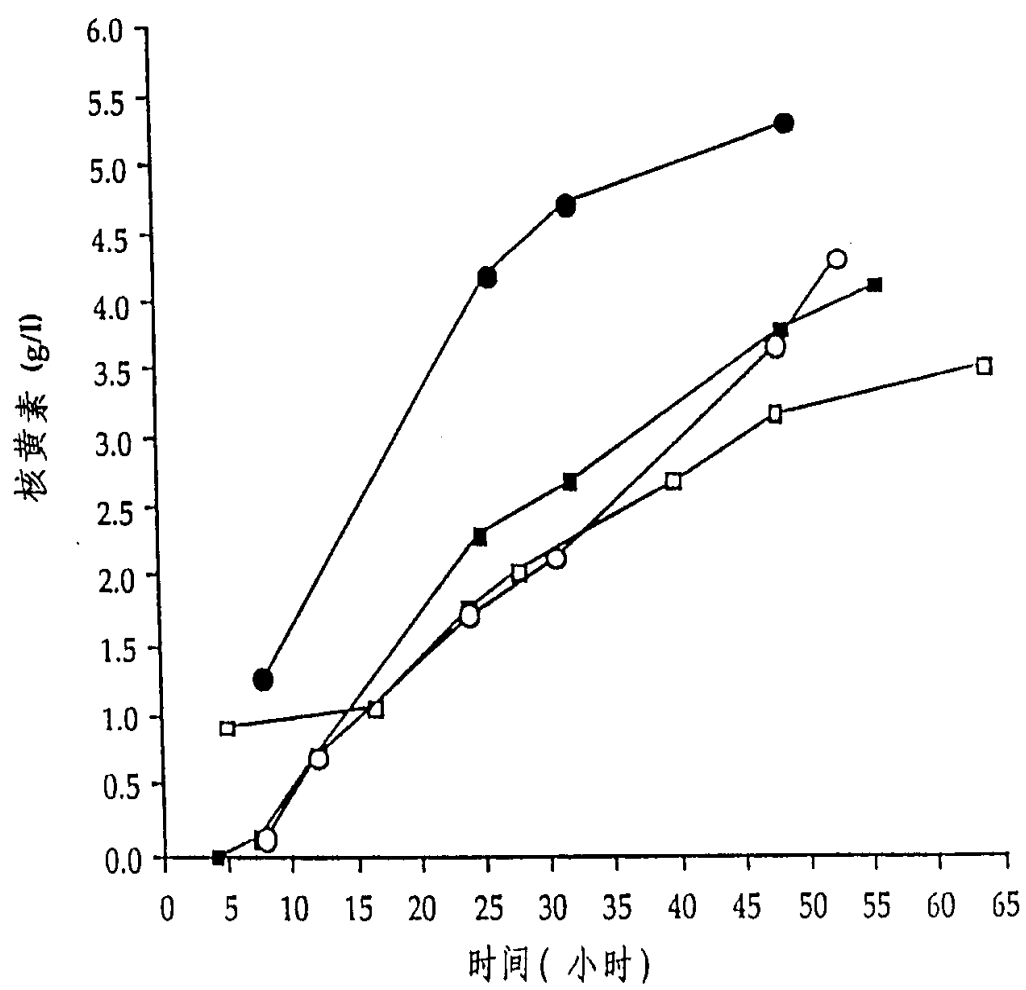
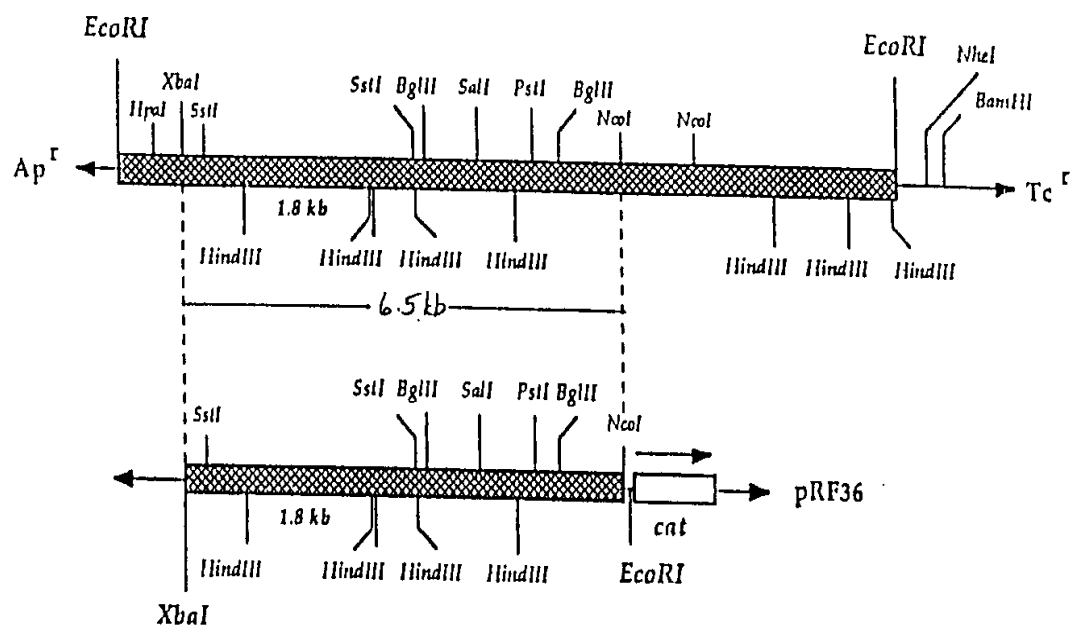


图 1 1



分离 *XbaI*-*EcoRI* 片段

与 pUC19 连接

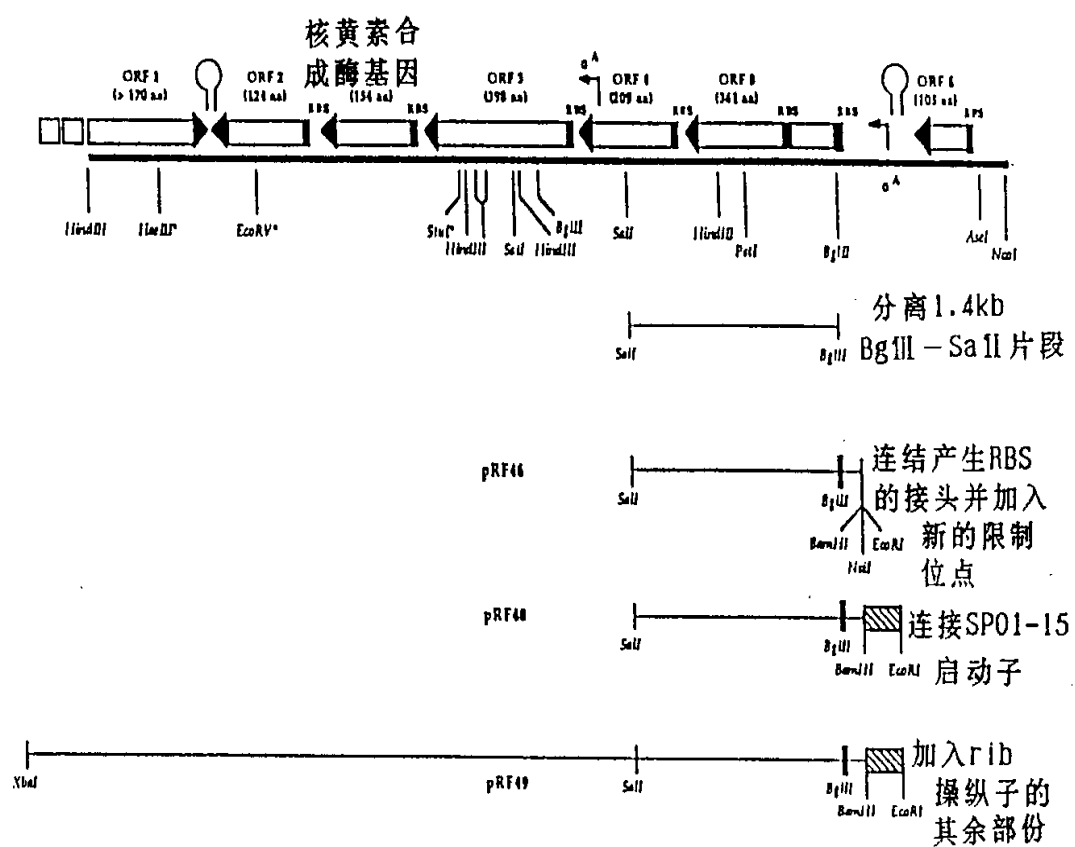
pRF39

在 *EcoRI* 位点加上 *cat* 基因

pRF40

图 1 2

图 1 3



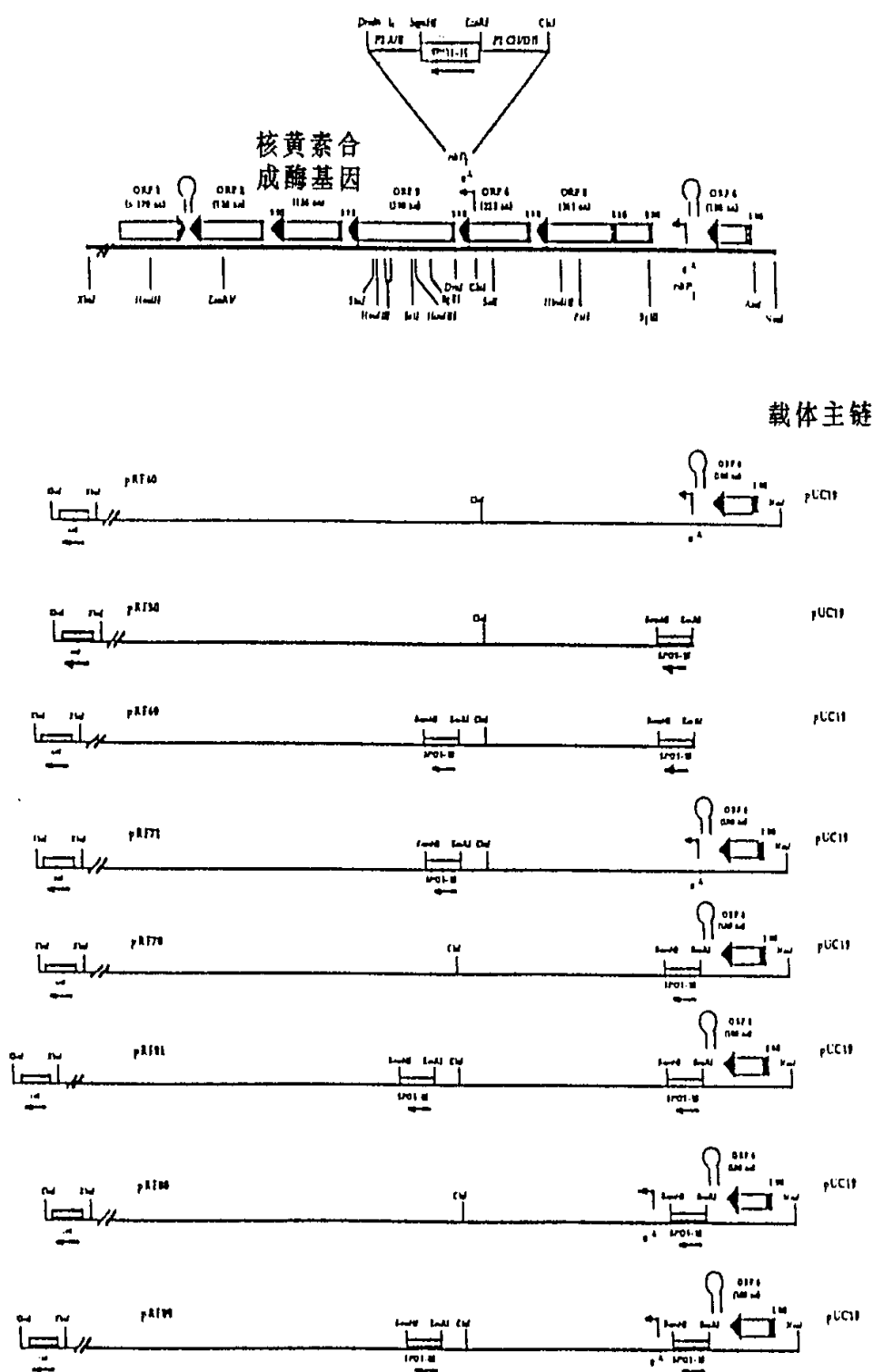


图 14

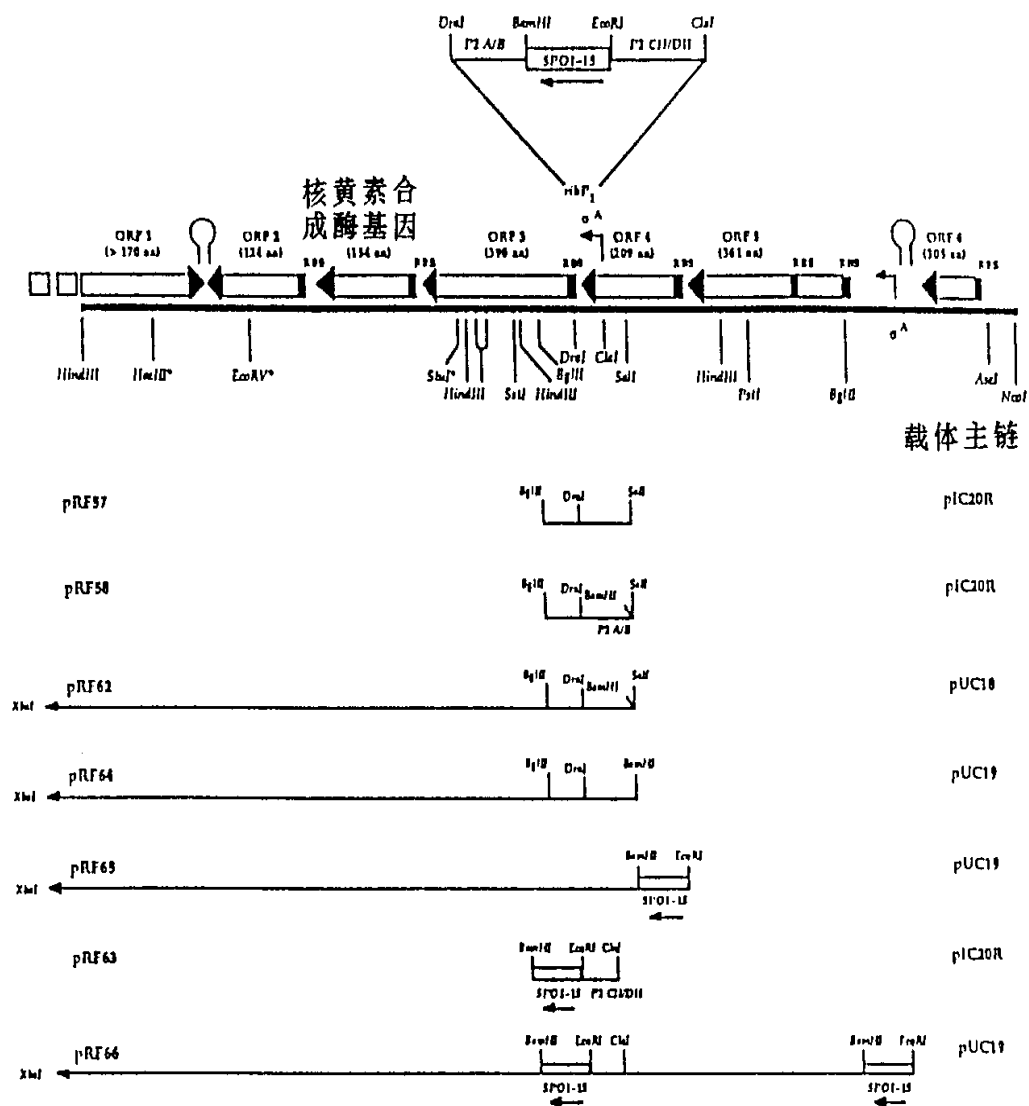
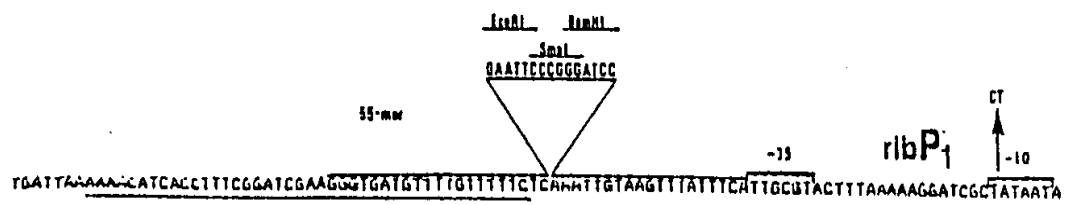


图 15

图 17



RB-5	AATTCATGCATGGATCCGACGGTAAATAAC AAAAGAGGGGAGGGAAACAAATGGAAGAGT ATTATATGAAGCTGGCCTTA
RB-6	GATCTAAGGCCAGCTTCATATAATACTCTT CCATTGTTTCCCTCCCCTCTTTGTTATT TACCGTCGGATCCATGCATG
P2-A	TCGACGGATCCTTTTAGAGAGGAAGATTG CATGTTTCATCCGATAGAAGAAGCACTGGA CGCTTT
P2-B	AAAGCGTCCAGTGCTTCTTCTATCGGATGA AACATGCAAATCTTCCTCTCTAAAAGGATC CG
P2-CII	CGATTTTTGCATAAAGCCAATGAAAATAAG ACCCAACAAACCATTACAAAAGCCTTCTTA AGCGAAAACGGCTTTTAG
P2-DII	AATTCTAAAAGCCGTTTTTCGCTTAAGAAGG CTTTTGTAATGGTTTGTGGGTCTTATTTT CATTGGCTTTATGCAAAAAT