

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷
C07K 14/62



[12] 发明专利说明书

[21] ZL 专利号 98117909.6

[45] 授权公告日 2003 年 12 月 31 日

[11] 授权公告号 CN 1132845C

[22] 申请日 1998.8.17 [21] 申请号 98117909.6

[30] 优先权

[32] 1997.8.18 [33] DE [31] 19735711.3

[71] 专利权人 阿文蒂斯药物德国有限公司

地址 联邦德国法兰克福

[72] 发明人 F-J·鲁伯洛德 R·凯勒

审查员 王亦然

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 郭建新

权利要求书 4 页 说明书 20 页

[54] 发明名称 获得具有正确键合的胱氨酸键的胰
岛素前体的改进的方法

[57] 摘要

本发明涉及一种在半胱氨酸或盐酸半胱氨酸和
离液助剂存在下, 获得具有正确键合的胱氨酸键的
胰岛素或胰岛素衍生物的前体的改进的方法。

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

1. 一种在半胱氨酸或盐酸半胱氨酸和离液助剂存在下，获得具有正确键合的胱氨酸键的胰岛素或胰岛素衍生物的前体的方法，该方法包括依次实施以下步骤：

(a) 将胰岛素或胰岛素衍生物的前体的水悬浮液与一定量的半胱氨酸或盐酸半胱氨酸混合，该量导致1~15个半胱氨酸或盐酸半胱氨酸的SH残基/前体的半胱氨酸残基；

(b) 在pH值8~11.5、温度15~55℃下，将含半胱氨酸或盐酸半胱氨酸的前体悬浮液加入到浓度为4~9M的离液助剂溶液中，使所得混合物在所述温度下保持10~60分钟；和

(c) 在pH值8~11.5、温度5~30℃下，将混合物加入到一定量的水中，使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到1~5mM，离液助剂的浓度稀释到0.2~1.0M。

2. 如权利要求1所要求的方法，其中

在步骤(a)中，半胱氨酸或盐酸半胱氨酸的量相当于导致1~6个半胱氨酸或盐酸半胱氨酸的SH残基/前体的半胱氨酸残基的量；

在步骤(b)中，在pH值8~11、温度30~45℃下，将含半胱氨酸或盐酸半胱氨酸的前体悬浮液加入到浓度为4~9M的离液助剂溶液中，使所得混合物在所述温度下保持20~40分钟；而且

在步骤(c)中，在pH值8~11、温度15~20℃下，将混合物加入到一定量的水中，使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到1~5mM，离液助剂的浓度稀释到0.2~1.0M。

3. 如权利要求1所要求的方法，其中所说的离液助剂为胍或盐酸胍。

4. 如权利要求2所要求的方法，其中所说的离液助剂为胍或盐酸胍。

5. 如权利要求1所要求的方法, 其中所说的离液助剂为尿素。
6. 如权利要求2所要求的方法, 其中所说的离液助剂为尿素。
7. 如权利要求1~6任一项所要求的方法, 其中步骤(b)中的离液助剂浓度为7.0~9M。
8. 如权利要求1~6任一项所要求的方法, 其中步骤(b)中的温度为40℃。
9. 如权利要求1~6任一项所要求的方法, 其中步骤(b)中的pH值为10~11。
10. 如权利要求1~6任一项所要求的方法, 其中步骤(c)中的pH值为10~11。
11. 如权利要求1~6任一项所要求的方法, 其中步骤(c)中水的量使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到2.5~3mM, 离液助剂的浓度稀释到0.5M。
12. 如权利要求2~6任一项所要求的方法, 其中步骤(b)中的离液助剂浓度为8M, 步骤(b)中的温度为40℃, 步骤(b)中的pH值为10.6, 步骤(c)中的pH值为10.6, 步骤(c)中水的量使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到2.5~3mM, 离液助剂的浓度稀释到0.5M。
13. 如权利要求1~6任一项所要求的方法, 其中所说的胰岛素或胰岛素衍生物的前体具有如式II的序列:



其中

R^2 为a) 一个氢原子,

b) 选自赖氨酸 (Lys) 和精氨酸 (Arg) 的一个氨基酸残基,

或者

c) 具有2~45个氨基酸残基的肽, 该肽的羧基端含有赖氨酸 (Lys) 或精氨酸 (Arg) 残基

R^1 为苯丙氨酸残基 (Phe) 或共价键,

(B2-B29) 为人胰岛素、动物胰岛素或胰岛素衍生物的B链上B2到B29位置上的氨基酸残基, 所述衍生物在一个或多个这些位置上任选改变,

Y 为一个可基因编码的氨基酸残基,

X 为a) 选自赖氨酸 (Lys) 和精氨酸 (Arg) 的一个氨基酸残基,

b) 具有2~35个氨基酸残基的肽, 该肽的N-端和羧基端含有氨基酸残基赖氨酸 (Lys) 或精氨酸 (Arg), 或者

c) 具有2~35个可基因编码的氨基酸的肽, 其中含有1~5个组氨酸残基,

(A2-A20) 为人胰岛素、动物胰岛素或胰岛素衍生物的B链上A2到A20位置上的氨基酸残基, 所述衍生物在一个或多个这些位置上任选改变, 而且

R^3 为一个可基因编码的氨基酸残基。

14. 如权利要求13所要求的方法, 其中式II中

R^2 为a) 一个氢原子或

b) 具有2~25个氨基酸残基的肽, 该肽的羧基端含有精氨酸 (Arg) 残基,

R^1 为一个苯丙氨酸残基 (Phe),

(B2-B29) 是人胰岛素B链B2到B29位置上的氨基酸残基,

Y 为选自丙氨酸 (Ala)、苏氨酸 (Thr) 和丝氨酸 (Ser) 的一个氨基酸残基,

X 为氨基酸残基精氨酸 (Arg) 或具有人胰岛素C链氨基酸序列的肽,

(A2- A20) 是人胰岛素B链A2到A20位置上的氨基酸残基, 而且

R³ 为选自天冬酰胺 (Asn)、丝氨酸 (Ser) 和甘氨酸 (Gly) 的一个氨基酸残基。

15. 如权利要求13所要求的方法, 其中式II中

R² 为a) 一个氢原子或

b) 具有2~15个氨基酸残基的肽, 该肽的羧基端存在精氨酸残基 (Arg),

R¹ 为一个苯丙氨酸残基 (Phe),

(B2-B29) 是人胰岛素B链B2到B29位置上的氨基酸残基,

Y 为苏氨酸残基 (Thr),

X 为氨基酸残基精氨酸 (Arg) 或具有2~35个氨基酸残基的肽, 该肽的首端和末端具有两个碱性氨基酸残基,

(A2- A20) 是人胰岛素B链A2到A20位置上的氨基酸残基, 而且

R³ 为氨基酸残基天冬酰胺 (Asn) 或甘氨酸 (Gly)。

16. 如权利要求15所要求的方法, 其中所述碱性氨基酸残基是精氨酸 (Arg) 和/或赖氨酸 (Lys)。

获得具有正确键合的胱氨酸键的胰岛素前体的改进的方法

本发明涉及一种在半胱氨酸或盐酸半胱氨酸和一种离液助剂存在下,获得具有正确键合的胱氨酸键的胰岛素或胰岛素衍生物的前体的改进的方法。

人胰岛素蛋白具有两条氨基酸链,总共具有51个氨基酸残基。两条氨基酸链中含有6个半胱氨酸残基,每两个半胱氨酸残基通过二硫键彼此连接。具有生物活性的人胰岛素,其A和B两条链间通过两个胱氨酸键进行连接,且A链中含有另一个胱氨酸键。从统计学角度看,人胰岛素分子中的二硫键形成方式具有15种可能性。然而,只有一种二硫键形成方式的人胰岛素具有生物活性。人胰岛素分子中下列半胱氨酸残基彼此连接:

A6-A11

A7-B7

A20-B19

字母A和B分别代表胰岛素的两条氨基酸链,数字代表氨基酸残基的位置,分别从每条链的氨基端到羧基端数起。两个人胰岛素分子间也可以形成二硫键,因此很容易形成许多不同的二硫键。

已知可从人胰岛素原制备人胰岛素的方法。人胰岛素原是有一条具有86个氨基酸残基的线性氨基酸链的蛋白,人胰岛素的B和A两条链是通过具有35个氨基酸残基的C肽彼此连接的。人胰岛素中二硫键是通过一个中间物-人胰岛素中的半胱氨酸残基形成的,该半胱氨酸带有硫保护基团,如S-磺酸根(-S-SO₃⁻)基团(EP0037255)。另外,已有方法可以制备具有正确键合的胱氨酸键的胰岛素原(Biochemistry, 60, (1968),p622-629),该方法是从猪胰腺中制备的胰岛素原开始,其中半胱氨酸残基是以SH残基形式存在的。“正确键合的胱氨酸键”这一术

语指的是具有生物活性的哺乳动物胰岛素中发现的二硫键。

DNA重组技术能够使胰岛素或胰岛素衍生物的前体，尤其能够使人胰岛素原或者氨基酸序列和/或氨基酸链长度区别于人胰岛素的胰岛素原在微生物中制备。由基因修饰大肠杆菌细胞制备的胰岛素原不具有正确键合的胱氨酸键。下面是利用大肠杆菌（EP 0 055 945）制备人胰岛素的方法步骤：

微生物发酵 - 细胞破碎 - 融合蛋白的分离 - 卤化氰裂解融合蛋白 - 含有胰岛素原序列的裂解产物的分离 - S-磺酸基对胰岛素原中胱氨酸残基的保护 - S-磺酸基的色谱提纯 - 正确键合的胱氨酸键的形成 - 胰岛素原的脱盐 - 含有正确键合的胱氨酸键的胰岛素原的色谱提纯 - 胰岛素原溶液的浓缩 - 胰岛素原浓缩溶液的色谱提纯 - 酶裂解胰岛素原以获得人胰岛素 - 所获得的人胰岛素的色谱提纯。

该方法的不足之处为步骤较多且在每一提纯步骤中都有产物损失，结果胰岛素的产率很低。由于操作多步骤，造成大量产物的损失。从以卤化氰裂解融合蛋白到亚硫酸盐化(sulfitolysis)和胰岛素原的提纯，预计损失40%的胰岛素原（EP 0 055 945）。在以后获取终产物的纯化过程中，同样会有很大的损失。

如果能大大简化操作步骤，通过DNA重组技术制备人胰岛素或其衍生物的产率将会得到提高。

EP 0 600 372 A1（或US 5, 473, 049）和EP 0 668 292 A2公开了一种制备胰岛素或胰岛素衍生物的改进方法，在该方法中，在有硫醇，如半胱氨酸，以及至少一种离液助剂，如尿素或盐酸胍的情况下，使胰岛素或胰岛素衍生物的前体发生反应，结果使胰岛素或胰岛素衍生物的前体中非正确键合的胱氨酸键转变为正确键合的胱氨酸键。在已知方法中，这些蛋白首先以很低浓度溶解在含有一种离液助剂或多种离液助剂的混合物的水溶液中，然后，再将蛋白混合物与硫醇水溶液进行混合。

惊奇的是，发现胰岛素或胰岛素衍生物的正确折叠的前体的产率可以通过以下方法提高并且还可以减少折叠反应所需要的时间，首先将硫醇，即半胱氨酸或盐酸半胱氨酸加到前体的水悬浮液中，而不是在第一

步中将前体借助离液助剂加入溶液中；然后加入离液助剂水溶液使前体发生溶解；最后用适量的水稀释混合物，使其中的半胱氨酸或盐酸半胱氨酸的浓度达到一个优选的浓度，从而使得前体发生正确的折叠反应。

因此，本发明涉及一种在半胱氨酸或盐酸半胱氨酸和离液助剂存在下，获得具有正确键合的胱氨酸键的胰岛素或胰岛素衍生物的前体的改进的方法，该方法包括依次实施以下步骤：

(a) 将胰岛素或胰岛素衍生物的前体的水悬浮液与一定量的半胱氨酸或盐酸半胱氨酸混合，使得每一个前体半胱氨酸残基1~15个半胱氨酸或盐酸半胱氨酸的SH残基；

(b) 在pH值约8~约11.5、温度约15~约55℃下，将含半胱氨酸或盐酸半胱氨酸的前体悬浮液加入到浓度为4~9M的离液助剂溶液中，使所得混合物在所述温度下保持约10~60分钟；

(c) 在pH值约8~约11.5、温度约5~约30℃下，将混合物加入到一定量的水中，使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到大约1~5mM，离液助剂的浓度稀释到0.2~1.0M。

优选地，所说的方法是这样一种方法，其中

在步骤(a)中，半胱氨酸或盐酸半胱氨酸的量相当于使得每一个前体半胱氨酸残基1~6个半胱氨酸或盐酸半胱氨酸的SH残基的量；

在步骤(b)中，在pH值8~11、温度30~45℃下，将含半胱氨酸或盐酸半胱氨酸的前体悬浮液加入到浓度为4~9M的离液助剂溶液中，使所得混合物在所述温度下保持20~40分钟；

在步骤(c)中，在pH值8~11、温度15~20℃下，将混合物加入到一定量的水中，使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到大约1~5mM，离液助剂的浓度稀释到0.2~1.0M。

离液助剂为在水溶液中能够打开氢键的化合物，如硫酸铵、盐酸胍、碳酸亚乙酯、硫氰酸盐、二甲亚砜和尿素。

根据本发明，优选采用的离液助剂为胍、盐酸胍，或特别优选尿素。

根据本发明的方法，步骤(b)中，离液助剂的优选浓度为7.0~9M；优选的温度为40℃；优选的pH值为10~11。

根据本发明的方法，步骤(c)中，优选的pH值为10~11；混合物进入其中的水的优选量为使得混合物中半胱氨酸或盐酸半胱氨酸的浓度稀释到2.5~3mM，离液助剂的浓度为0.5M的量。

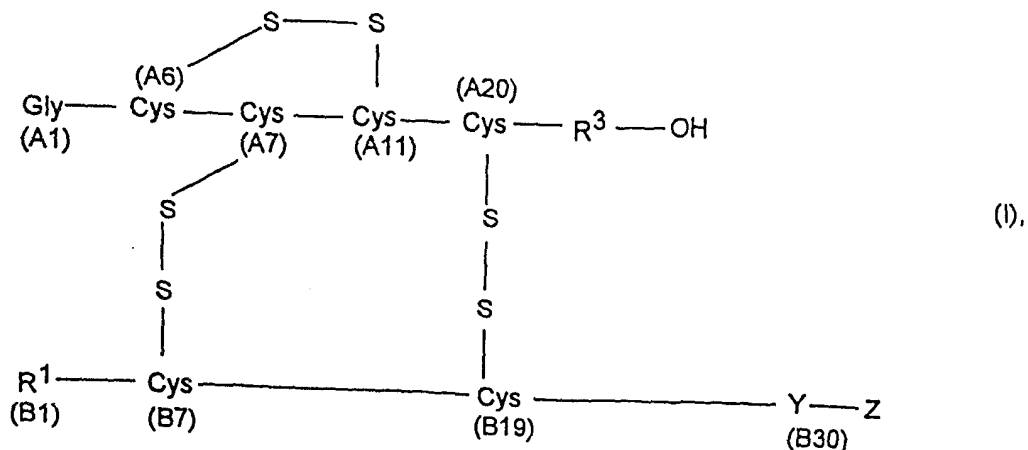
更优选的是，根据本发明的方法是这样一种方法，其中步骤(b)中的离液助剂的浓度大约为8M，温度大约为40℃，pH值大约为10.6；步骤(c)中的pH值大约为10.6，水的量是使得混合物中半胱氨酸或盐酸半胱氨酸的浓度大约稀释到2.5~3mM，离液助剂的浓度稀释到0.5M的量。

根据本发明的方法所制备的胰岛素或胰岛素衍生物的前体，特别是胰岛素原，其胱氨酸键是正确键合的。

胰岛素衍生物是天然胰岛素即人胰岛素（见SEQ ID NO.:1=人胰岛素的A链；见SEQ ID NO.:2=人胰岛素的B链，序列表）或动物胰岛素的衍生物，其与天然胰岛素不同之处在于天然胰岛素中的至少一个天然氨基酸残基被替换和/或添加了至少一个氨基酸残基和/或相应的、而不是相同的天然胰岛素的有机残基。

根据本发明的方法制得的含有正确键合的胱氨酸键的胰岛素或胰岛素衍生物的前体，可以通过EP 0 600 372 A1（或US5, 473, 049）或EP 0 668 292 A2中所描述的方法来制备含有正确键合的胱氨酸键的胰岛素或胰岛素衍生物，其中所描述的方法是采用胰蛋白酶或类胰蛋白酶进行酶裂解，另外，合适的话还可采用羧肽酶B，以及在吸附树脂上纯化的后续步骤。

可以从前体制备的胰岛素或胰岛素衍生物可以更好地采用式I来描述：



其中:

Y 为可基因编码的氨基酸残基;

Z 为 a)选自组氨酸、精氨酸和赖氨酸的一种氨基酸残基,

b)具有2或3个氨基酸残基的肽,该肽的羧基端含有精氨酸或赖氨酸残基,

c)具有2-35个可基因编码的氨基酸的肽,含有1-5个组氨酸残基,或,

d)OH,

R¹ 为苯丙氨酸残基(Phe)或一共价键,

R³ 为可基因编码的氨基酸残基,

为了简化对应于动物的胰岛素或胰岛素衍生物的式I,没有具体给出人胰岛素A链氨基酸序列上的基因A2-A20,同样也没有具体给出人胰岛素B链氨基酸序列上的基因B2-B29.

肽和蛋白的氨基酸序列是从氨基酸链的N-端开始排列的。式I中括号中的内容如A6、A20、B1、B7或B19对应于胰岛素A链或B链中氨基酸残基的位置。

“可基因编码的氨基酸残基”一词是指甘氨酸、丙氨酸、丝氨酸、苏氨酸、缬氨酸、亮氨酸、异亮氨酸、天冬氨酸、天冬酰胺、谷氨酸、谷氨酰胺、半胱氨酸、蛋氨酸、精氨酸、赖氨酸、组氨酸、酪氨酸、苯丙氨酸、色氨酸、脯氨酸和硒代半胱氨酸。

动物胰岛素的“A2-A20残基”和“B2-B29残基”可以理解为,如牛、猪或鸡胰岛素的氨基酸序列。胰岛素衍生物的“A2-A20残基”和“B2-B29残基”指的是通过用其它可基因编码的氨基酸替换氨基酸形成的、人胰岛素的相应氨基酸序列。

如人胰岛素的A链具有以下序列(SEQ ID NO.: 1):

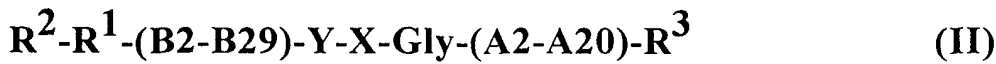
```
Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln
Leu
Glu Asn Tyr Cys Asn.
```

人胰岛素的B链具有以下序列 (SEQ ID NO.: 2) :

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu
Tyr
Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr.

在这种情况下, 式I中的R³为天冬酰胺 (Asn), R¹为苯丙氨酸 (Phe), Y为苏氨酸 (Thr), Z为OH.

本发明方法特别适用于制备式II的胰岛素或胰岛素衍生物的前体, 其中胱氨酸键折叠正确 (式II中未显示),



其中:

R² 为a) 一个氢原子,

b) 选自赖氨酸 (Lys) 和精氨酸 (Arg) 的一个氨基酸残基, 或者

c) 具有2~45个氨基酸残基的肽, 该肽的羧基端含有赖氨酸 (Lys) 或精氨酸 (Arg)

R¹ 为苯丙氨酸残基 (Phe) 或共价键,

(B2-B29) 为人胰岛素、动物胰岛素或胰岛素衍生物的B链上B2到B29位置上的氨基酸残基, 所述衍生物在一个或多个这些位置上任选改变,

Y 为一个可基因编码的氨基酸残基,

X 为a) 选自赖氨酸 (Lys) 和精氨酸 (Arg) 的一个氨基酸残基,

b) 具有2~35个氨基酸残基的肽, 该肽的N-端和羧基端含有赖氨酸 (Lys) 或精氨酸 (Arg) 残基, 或者

c) 具有2~35个可基因编码的氨基酸的肽, 其中含有1~5个组氨酸残基,

(A2- A20) 为人胰岛素、动物胰岛素或胰岛素衍生物的B链上A2到A20位置上的氨基酸残基, 所述衍生物在一个或多个这些位置上任选改变,

R^3 为一个可基因编码的氨基酸残基。

1. 优选的是，在式II中：

R^2 为a) 一个氢原子或

b) 具有2~25个氨基酸残基的肽，该肽的羧基端含有精氨酸 (Arg) 残基，

R^1 为一个苯丙氨酸残基 (Phe)，

(B2-B29) 是人胰岛素B链B2到B29位置上的氨基酸残基，

Y 为选自丙氨酸 (Ala)、苏氨酸 (Thr) 和丝氨酸 (Ser) 的一个氨基酸残基，

X 为氨基酸残基精氨酸 (Arg) 或具有人胰岛素C链氨基酸序列的肽，

(A2-A20) 是人胰岛素B链A2到A20位置上的氨基酸残基，

R^3 为选自天冬酰胺 (Asn)、丝氨酸 (Ser) 和甘氨酸 (Gly) 的一个氨基酸残基，

人胰岛素C链具有以下氨基酸序列 (SEQ ID NO.: 3)：

```
Arg Arg Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly
Gly
Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser
Leu
Gln Lys Arg.
```

2. 优选的是，在式II中：

R^2 为a) 一个氢原子或

b) 具有2~15个氨基酸残基的肽，该肽的羧基端含有精氨酸残基 (Arg)，

R^1 为一个苯丙氨酸残基 (Phe)，

(B2-B29) 是人胰岛素B链B2到B29位置上的氨基酸残基，

Y 为苏氨酸残基 (Thr)，

X 为氨基酸残基精氨酸 (Arg) 或具有2~35个氨基酸残基的肽, 该肽的首端和末端有两个碱性氨基酸残基, 尤其是精氨酸 (Arg) 和/或赖氨酸 (Lys),

(A2- A20) 是人胰岛素B链A2到A20位置上的氨基酸残基,

R³ 为氨基酸残基天冬酰胺 (Asn) 或甘氨酸 (Gly)。

式I的胰岛素或胰岛素衍生物的残基Z是指式II前体的X的部分氨基酸序列, 残基Z是通过蛋白酶, 如胰蛋白酶、类胰蛋白酶或羧肽酶B的活性作用而产生的。基团R³是胰岛素A链位置A21上的氨基酸残基。基团Y是胰岛素B链位置B30上的氨基酸残基。

胰蛋白酶、类胰蛋白酶是在精氨酸或赖氨酸残基的位置上切割氨基酸链的蛋白酶。

羧肽酶B是除去氨基酸链羧基端上的碱性氨基酸残基如Arg 或Lys的外切蛋白酶 (Kemmler等, 生物化学杂志 246, p6786-6791)。

从1中提到的前体, 例如, 可以制备式I所示的含有正确键合的脱氨酸键的胰岛素或胰岛素衍生物, 其中Y, R¹, R², R³, A2-A20和B2-B29的含义如1中提到的, 并且Z是精氨酸残基(Arg)、肽的残基Arg- Arg或-OH。

从2中提到的前体, 例如, 可以制备式I所示的含有正确键合的脱氨酸键的胰岛素或胰岛素衍生物, 其中Y, R¹, R², R³, A2-A20和B2-B29的含义如2中提到的, 并且Z是精氨酸残基(Arg)、肽的残基Arg- Arg或Lys- Lys或-OH。

借助于大量的基因构建体, 可以在微生物中形成式II中的前体 (EP 0 489 780, EP 0 347 781, EP 0 453 969)。通过发酵, 基因构建体在诸如大肠杆菌或链霉菌中获得表达。所产生的蛋白贮存于微生物体内 (EP 0 489 780) 或者分泌到发酵液中。

本发明的方法可以直接采用细胞破碎后的式II的胰岛素或胰岛素衍生物的前体, 该前体仍然含有大量的来自发酵液和微生物的杂蛋白。然而, 也可以采用预纯化了的式II前体, 例如经过沉淀的或色谱纯化的前体。

实施例1（对照例，现有技术）

遗传修饰的大肠杆菌细胞经发酵（EP 0 489 780），制备出具有下列氨基酸序列的融合蛋白。

胰岛素原序列1（SEQ ID NO.: 4）：

Ala Thr Thr Ser Thr Gly Asn Ser Ala Arg Phe Val Asn Gln His
Leu
Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu
Arg
Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu
Gln
Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu
Gln
Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu
Gln
Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
Asn

胰岛素原序列1对应于式II，式中

X 为人胰岛素的C-肽（SEQ ID NO.: 3）

Y 为苏氨酸（B30），

R¹ 为苯丙氨酸（B1），

R² 为具有10个氨基酸残基的肽，

R³ 为天冬酰胺（A21）和

A2-A20是人胰岛素A链的氨基酸序列（氨基酸残基2~20），B2- B29是人胰岛素B链的氨基酸序列（氨基酸残基2~29）。

具有胰岛素原的序列1的表达的融合蛋白聚集在大肠杆菌的细胞内，并形成包涵体。发酵后，离心分离细胞，采用普通的高压匀浆法破碎细胞，释放出来的融合蛋白包涵体通过离心进行分离。

将分离到的20kg具有胰岛素原序列1的融合蛋白包涵体（基于冻干后的干物质；通过HPLC测定含胰岛素的融合蛋白的含量，结果为50%）溶解于550L、pH值为10.6的8M尿素溶液中。合适的话，离心掉引起混浊的少量物质后，将清液加入9000L，pH值为10.6和温度为4℃

的半胱氨酸水溶液（5kg 盐酸半胱氨酸水合物）。

经过大约24小时的折叠反应，通过分析用HPLC，测定出这批反应中具有正确键合的胱氨酸键的胰岛素原序列1的含量为3.0 kg，相当于30%的转化率。

用1N的HCl将9500L溶液的pH值调节到5.0并进行分离。然后，用1N的NaOH将pH值调节到9。往溶液中加入3g的胰蛋白酶。根据HPLC测定的结果，具有2个羧基端精氨酸残基的胰岛素含量为1.25 kg。

经羧肽酶B裂解后，再进行色谱纯化，便可得到人胰岛素。

人胰岛素对应于式I，式中

Y 为苏氨酸（B30），

Z 为OH，

R¹ 为苯丙氨酸（B1），

R³ 为天冬酰胺（A21）和

A2-A20是人胰岛素A链的氨基酸序列（氨基酸残基2~20），B2- B29是人胰岛素B链的氨基酸序列（氨基酸残基2~29）。

人胰岛素2由SEQ ID NO.: 1和2组成，其中序列1和2是通过正确键合的胱氨酸键彼此连接的。

如EP 0 668 292中所描述的，溶液通过吸附树脂进行浓缩和纯化。含有胰岛素2的洗脱物用水稀释并调节pH值后，可以直接采用色谱柱进一步进行纯化。

HPLC分析

于95℃下，将0.5g的蛋白在40ml下述溶液中溶解2分钟，该溶液含6M的盐酸胍、50mM的tris(pH值为8.5)、5 mM的乙二胺四乙酸（EDTA）、1%的2-巯基乙醇、10 mM的二硫苏糖醇，然后，在14000 g下离心20分钟。取0.02 ml的上清液加到高压液相色谱柱上。

柱子：®Nucleogel RP 300-5/46 (Macherey & Nagel, Aachen, 德国)

梯度：缓冲液A：0.1%三氟乙酸（TFA）

缓冲液B：0.09%TFA/乙腈

温度：55℃

运行总时间：40分钟

梯度以对应于运行时间的缓冲液B的浓度来表示：

10分钟25%，12分钟60%，13分钟90%，15分钟100%

流速：1 ml/分钟

检测波长：215nm

胰岛素的保留时间：约19分钟

实施例2（本发明的方法）

通过对遗传修饰的大肠杆菌细胞进行发酵培养（EP 0 489 780），制备出具有实施例1中所示的氨基酸序列的融合蛋白（胰岛素原序列1. SEQ ID NO.: 4）。

具有胰岛素原序列1的表达的融合蛋白聚集在大肠杆菌细胞内，形成包涵体。发酵培养结束后，离心分离细胞，常规高压匀浆破碎细胞，最后通过离心分离到释放出的融合蛋白的包涵体。

将5kg的盐酸半胱氨酸水合物加到含有40kg融合蛋白的融合蛋白水悬浮液中（通过冻干样品测定的）。

将含有胰岛素原序列1的悬浮液（经HPLC测定含胰岛素的融合蛋白的浓度为50%）于40℃溶解到550L、pH值为10.2、浓度为8M的尿素溶液中。在pH值为10.6和温度为15℃的条件下，将所述清液加入9000L的水。

经大约5小时完成折叠反应后，通过分析用HPLC测出反应液中具有正确键合的胱氨酸键的胰岛素原序列1的含量为10.0kg，相当于50%的转化率。

用1N HCl 将9500L溶液的pH值调节到5.0并进行分离，再添加1N NaOH 将溶液的pH值调节到9。往溶液中加入10g的胰蛋白酶。结果得到4kg具有2个羧基端精氨酸残基的胰岛素。用羧肽酶B进行裂解后，得到人胰岛素（SEQ ID NO.: 1和2，有正确键合的胱氨酸键）。

采用吸附树脂对溶液进行浓缩和纯化。

用水稀释并进行pH值调节后，可直接采用色谱柱进一步纯化含人胰岛素的洗脱物。

实施例3 (对照实施例, 现有技术)

通过对遗传修饰的大肠杆菌细胞进行发酵培养 (EP 0 489 780), 制备出具有下列氨基酸序列的融合蛋白

胰岛素原序列2 (SEQ ID NO.: 5):

```

Ala Thr Thr Ser Thr Gly Asn Ser Ala Arg Phe Val Asn Gln His
Leu
Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu
Arg
Gly Phe Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu
Gln
Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu
Gln
Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu
Gln
Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys
Gly

```

胰岛素原序列2对应于式II, 式中

X 为人胰岛素的C-肽 (SEQ ID NO.: 3)

Y 为苏氨酸 (B30),

R¹ 为苯丙氨酸 (B1),

R² 为具有10个氨基酸残基的肽,

R³ 为甘氨酸 (A21) 和

A2-A20是人胰岛素A链的氨基酸序列 (氨基酸残基2~20), B2- B29是人胰岛素B链的氨基酸序列 (氨基酸残基2~29)。

具有胰岛素原序列2的表达融合蛋白聚集在大肠杆菌的细胞内, 并形成包涵体。发酵后, 离心分离细胞, 采用普通的高压匀浆法破碎细胞, 释放出来的融合蛋白包涵体通过离心进行分离。

将分离到的20kg具有胰岛素原序列2的融合蛋白包涵体 (基于冻干后的干物质; 通过HPLC测定含胰岛素的融合蛋白的含量, 结果为50%) 于20℃条件下, 溶解于550L、pH值为10.6的8M尿素溶液中。将清液加入9000L, pH值为10.6和温度为4℃的半胱氨酸水溶液 (5kg盐酸

半胱氨酸水合物)。

经过大约24小时的折叠反应后，通过分析用HPLC测出反应液中具有正确键合的胱氨酸键的胰岛素原序列2的含量为3.0 kg，相当于30%的转化率。

用1N的HCl将9500L溶液的pH值调节到5.0并进行分离。然后，添加1N的NaOH将pH值调节到9。往溶液中加入3g的胰蛋白酶。根据HPLC测定，具有2个羧基端精氨酸残基的胰岛素衍生物含量为0.98kg。该胰岛素衍生物对应于式I，其中，

Y 为苏氨酸 (B30)，

Z 为精氨酸-精氨酸，

R¹ 为苯丙氨酸 (B1)，

R³ 为甘氨酸 (A21) 和

A2-A20是人胰岛素A链的氨基酸序列(氨基酸残基2~20)，B2-B29是人胰岛素B链的氨基酸序列(氨基酸残基2~29)，由SEQ ID NO.: 6和7组成，该两个序列通过正确的胱氨酸键彼此连接。

溶液通过吸附树脂进行浓缩和纯化。

含有胰岛素衍生物的洗脱物用水稀释并调节pH值后，可以直接采用色谱柱进一步进行纯化。

实施例4(本发明的方法)

通过对遗传修饰的大肠杆菌细胞进行发酵培养(EP 0 489 780)，按照实施例3制备含有胰岛素原序列2的融合蛋白(SEQ ID NO.: 5)。

具有胰岛素原序列2的表达的融合蛋白聚集在大肠杆菌细胞内，并形成包涵体。发酵培养结束后，离心分离细胞，常规高压匀浆破碎细胞。通过离心分离到释放的融合蛋白包涵体。

将5kg的盐酸半胱氨酸水合物加到含有40kg融合蛋白的融合蛋白水悬浮液中(通过冻干样品测定的)。

将含有胰岛素原序列2的悬浮液(经高压液相色谱分析，含胰岛素的融合蛋白的浓度为50%)于40℃溶解到550L、pH值为10.2、浓度为8M的尿素溶液中。在pH值为10.6和温度为15℃的条件下，将所述清液加

入9000L的水。

经大约5小时完成折叠反应后，通过分析用HPLC测定出具有正确键合的胱氨酸键的胰岛素原序列1的含量为10.0kg，相当于50%的转化率。

用1N HCl 将9500L溶液的pH值调节到5.0并进行分离，再添加1N NaOH 溶液将pH值调节到9。往溶液中加入10g的胰蛋白酶，结果得到2.8kg的胰岛素衍生物（HPLC测定），胰岛素衍生物由SEQ ID NO.: 6和7的序列组成，该两个序列通过正确键合的胱氨酸键彼此连接。

采用吸附树脂对溶液进行浓缩并进行纯化。

用水稀释并进行pH值调节后，可直接采用色谱柱进一步纯化含人胰岛素衍生物的洗脱物。

序列表

(1) 一般信息:

(i) 申请人:

(A) 名称: Hoechst Marion Roussel Deutschland GmbH

(B) 街: -

(D) 城市: Frankfurt am Main

(E) 国家: 联邦德国

(F) 邮编: 65926

(G) 电话: 069-305-5307

(H) 传真: 069-35-7175

(I) 电报: 041234-700 hod

(ii) 发明名称: 获得具有正确键合的胱氨酸键的胰岛素前体的改进的方法

(iii) 序列数目: 7

(iv) 计算机可读形式:

(A) 介质类型: 软盘

(B) 计算机: IBM PC 兼容机

(C) 操作系统: PC-DOS/MS-DOS

(D) 软件: PatentIn Release #1.0, 版本 #1.25 (EPO)

(2) SEQ ID NO: 1的信息:

(i) 序列特征:

(A) 长度: 21个氨基酸

(B) 类型: 氨基酸

(C) 链型: 单链

(D) 拓扑学: 线性

(ii) 分子类型: 蛋白质

(vi) 初始来源:

(A) 生物体: 大肠杆菌

(ix) 特征:**(A) 名称/关键词: 蛋白质****(B)位置: 1..21****(xi) 序列描述: SEQ ID NO: 1:**

```

Leu      Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln
1              5              10              15
      Glu Asn Tyr Cys Asn
                20

```

(2)SEQ ID NO: 2的信息:**(i) 序列特征:****(A) 长度: 30个氨基酸****(B) 类型: 氨基酸****(C) 链型: 单链****(D) 拓扑学: 线性****(ii) 分子类型: 蛋白质****(vi) 初始来源:****(A) 生物体: 大肠杆菌****(ix) 特征:****(A) 名称/关键词: 蛋白质****(B)位置: 1..30****(xi) 序列描述: SEQ ID NO: 2:**

```

Tyr      Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu
1              5              10              15
      Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr
                20              25              30

```

(2)SEQ ID NO: 3的信息:**(i) 序列特征:**

- (A) 长度: 35个氨基酸
- (B) 类型: 氨基酸
- (C) 链型: 单链
- (D) 拓扑学: 线性
- (ii) 分子类型: 蛋白质
- (vi) 初始来源:
 - (A) 生物体: 大肠杆菌
- (ix) 特征:
 - (A) 名称/关键词: 蛋白质
 - (B) 位置: 1..35
- (xi) 序列描述: SEQ ID NO: 3:

```

Arg Arg Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly
Gly      1           5           10           15
Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser
Leu      20           25           30
Gln Lys Arg
          35

```

(2)SEQ ID NO: 4的信息:

- (i) 序列特征:
 - (A) 长度: 96个氨基酸
 - (B) 类型: 氨基酸
 - (C) 链型: 单链
 - (D) 拓扑学: 线性
- (ii) 分子类型: 蛋白质
- (vi) 初始来源:
 - (A) 生物体: 大肠杆菌
- (ix) 特征:
 - (A) 名称/关键词: 蛋白质

(B)位置: 1..96**(xi) 序列描述: SEQ ID NO: 4:**

```

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

```

(2)SEQ ID NO: 5的信息:**(i) 序列特征:****(A) 长度: 96个氨基酸****(B) 类型: 氨基酸****(C) 链型: 单链****(D) 拓扑学: 线性****(ii) 分子类型: 蛋白质****(vi) 初始来源:****(A) 生物体: 大肠杆菌****(ix) 特征:****(A) 名称/关键词: 蛋白质****(B)位置: 1..96**

(xi) 序列描述: SEQ ID NO: 5:

```

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

```

(2)SEQ ID NO: 6的信息:**(i) 序列特征:****(A) 长度: 32个氨基酸****(B) 类型: 氨基酸****(C) 链型: 单链****(D) 拓扑学: 线性****(ii) 分子类型: 蛋白质****(vi) 初始来源:****(A) 生物体: 大肠杆菌****(ix) 特征:****(A) 名称/关键词: 蛋白质****(B)位置: 1..32****(xi) 序列描述: SEQ ID NO: 6:**


```

Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu
Tyr
1           5           10           15
Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr Arg
Arg
           20           25           30

```

(2)SEQ ID NO: 7的信息:**(i) 序列特征:****(A) 长度: 21个氨基酸****(B) 类型: 氨基酸****(C) 链型: 单链****(D) 拓扑学: 线性****(ii) 分子类型: 蛋白质****(vi) 初始来源:****(M) 生物体: 大肠杆菌****(ix) 特征:****(A) 名称/关键词: 蛋白质****(B)位置: 1..21****(xi) 序列描述: SEQ ID NO: 7:**

```

Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln
Leu
1           5           10           15
Glu Asn Tyr Cys Gly
           20

```