

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷
C12N 9/42
C12N 15/56



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 97198010.1

[45] 授权公告日 2004 年 6 月 23 日 [11] 授权公告号 CN 1154721C

[22] 申请日 1997.7.24 [21] 申请号 97198010.1 [30] 优先权 [32] 1996. 7.24 [33] JP [31] 194974/1996 [86] 国际申请 PCT/JP1997/002561 1997.7.24 [87] 国际公布 WO1998/003640 日 1998.1.29 [85] 进入国家阶段日期 1999.3.17 [71] 专利权人 明治制果株式会社 地址 日本东京都 [72] 发明人 村岛弘一郎 浜谷彻 古贺仁一郎 河野敏明 守屋达树 隅田奈绪美 青柳薰 村上健 审查员 丁 海	[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 隗永良
权利要求书 1 页 说明书 31 页	

[54] 发明名称 葡聚糖内切酶以及含有该酶的纤维素酶制剂

[57] 摘要

公开了理想的可用于含纤维素纤维细毛的去除、减量加工和斜纹粗棉布染色中含纤维素纤维的脱色加工的高活性纤维素酶及其基因。从 *Humicola insolens* 分离的新的纤维素酶 NCE4 是高活性的纤维素酶，该酶可以用于各种含纤维素纤维的处理。

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

1. 一种分离的蛋白质，其含有序列 1 所示的氨基酸序列或含有序列 1 中 1~284 位氨基酸所示的序列。
2. 一种分离的 DNA 序列，其编码权利要求 1 所述的蛋白质。
3. 权利要求 2 所述的 DNA 序列，其含有序列 2 所示的碱基序列。
4. 权利要求 3 所述的 DNA 序列，其含有序列 2 所示的碱基序列中 118~1088 位的碱基。
5. 载体，其含有权利要求 2 的 DNA 序列。
6. 宿主细胞，其被权利要求 5 所述的载体所转化。
7. 权利要求 1 所述的蛋白质的制备方法，包括培养权利要求 6 所述的宿主细胞，和从其培养物中提取权利要求 1 所述的蛋白质的工序。
8. 纤维素酶制剂，其含有权利要求 1 所述的蛋白质。
9. 去除含纤维素纤维的细毛的方法，包括使权利要求 1 的蛋白质或权利要求 8 的纤维素酶制剂与含纤维素纤维接触以去除含纤维素纤维的细毛的工序。
10. 斜纹粗棉布染色中含纤维素纤维的脱色加工方法，包括使权利要求 1 的蛋白质或权利要求 8 的纤维素酶制剂与斜纹粗棉布染色中含纤维素纤维接触的工序。
11. 减少含纤维素纤维的方法，包括使权利要求 1 的蛋白质或权利要求 8 的纤维素酶制剂与含纤维素纤维接触以减少含纤维素纤维的工序。
12. 权利要求 1 的蛋白质或权利要求 8 的纤维素酶制剂在去除含纤维素纤维细毛中的应用。
13. 权利要求 1 的蛋白质或权利要求 8 的纤维素酶制剂在斜纹粗棉布染色中含纤维素纤维的脱色加工中的应用。
14. 权利要求 1 的蛋白质或权利要求 8 的纤维素酶制剂在减少含纤维素纤维中的应用。

葡聚糖内切酶以及含有该酶的纤维素酶制剂

本发明涉及葡聚糖内切酶以及含有该酶的纤维素酶制剂，以及利用该制剂去除含纤维素纤维细毛、减量加工和斜纹粗棉布染色中含纤维素纤维的脱色加工。

为了使含纤维素纤维具有所期望的特性，要用纤维素酶对这样的纤维进行处理。例如，在纺织界中为了改善含纤维素纤维与皮肤的触感和其外观，以及为了赋予染色斜纹粗棉布中含纤维素纤维具有磨石水洗（ストーンウォッシュ）的外观，就利用纤维素酶对这样的纤维进行处理。

另外，近年来，作为来自木材纸浆的纤维素用有机溶剂溶解、纺线而制成的再生纤维素系纤维的Tencel由于它的高的强度、吸水度等性质，同时其制造方法又不容易引起环境污染，所以一直受到人们的注目。然而由于Tencel在制造过程中产生细毛，所以直接作为纤维的商品价值低。因此，利用纤维素酶除去制造过程中产生的细毛的方法的提案不断提出。

现在，处理含纤维素纤维使用的纤维素酶主要来自木材不朽菌木霉菌Trichoderma或霉质菌Humicola。然而，为了使纤维达到所期望的效果，目前的状况是必须使用很多的这类纤维素酶。

如果能够通过使用比以前用量少的高活性的纤维素酶改善含纤维素纤维的触感和外观，赋予染色斜纹粗棉布中含纤维素纤维具有磨石水洗的外观，并除去Tencel的细毛的话，就能够进一步降低这些处理中的成本。

来自Humicola菌的纤维素酶在W091/17243号公报（特表平5-509223号公报）中已有报道，43kD的葡聚糖内切酶基因已从Humicola insolens DSM1800菌株中分离出来，并确定了其碱基序列。

本发明人从Humicola insolens菌中分离出了对各种含有纤维素

的纤维的处理非常有用的新的高活性的纤维素酶及其基因。本发明是基于上述认识的产物。

因此，本发明的目的在于提供新的高活性的纤维素酶及其基因。

另外，提供利用新的纤维素酶除去含纤维素纤维细毛的方法，减量加工方法以及斜纹粗棉布染色中含纤维素纤维的脱色加工方法也是本发明的目的。

而本发明的新的纤维素酶是序列 1 记载的氨基酸序列的一部分或含有 1~284 号序列的蛋白质和其变化的蛋白质。

本发明的含纤维素纤维细毛的除去方法，减量加工方法以及斜纹粗棉布染色中含纤维素纤维的脱色加工方法中包括使上述蛋白质或其变化蛋白质与染色的斜纹粗棉布中含纤维素纤维接触的工序。

因此，本发明涉及以下方面：

1. 一种分离的蛋白质，其含有序列 1 所示的氨基酸序列或含有序列 1 中 1~284 位氨基酸所示的序列。
2. 一种分离的 DNA 序列，其编码项目 1 所述的蛋白质。
3. 项目 2 所述的 DNA 序列，其含有序列 2 所示的碱基序列。
4. 项目 3 所述的 DNA 序列，其含有序列 2 所示的碱基序列中 118~1088 位的碱基。
5. 载体，其含有项目 2 的 DNA 序列。
6. 宿主细胞，其被项目 5 所述的载体所转化。
7. 项目 1 所述的蛋白质的制备方法，包括培养项目 6 所述的宿主细胞，和从其培养物中提取项目 1 所述的蛋白质的工序。
8. 纤维素酶制剂，其含有项目 1 所述的蛋白质。
9. 去除含纤维素纤维的细毛的方法，包括使项目 1 的蛋白质或项目 8 的纤维素酶制剂与含纤维素纤维接触以去除含纤维素纤维的细毛的工序。
10. 斜纹粗棉布染色中含纤维素纤维的脱色加工方法，包括使项目 1 记载的蛋白质或项目 8 的纤维素酶制剂与斜纹粗棉布染色中含纤维素纤维接触的工序。

11. 减少含纤维素纤维的方法，包括使项目 1 的蛋白质或项目 8 的纤维素酶制剂与含纤维素纤维接触以减少含纤维素纤维的工序。
12. 项目 1 的蛋白质或项目 8 的纤维素酶制剂在去除含纤维素纤维细毛中的应用。
13. 项目 1 的蛋白质或项目 8 的纤维素酶制剂在斜纹粗棉布染色中含纤维素纤维的脱色加工中的应用。
14. 项目 1 的蛋白质或项目 8 的纤维素酶制剂在减少含纤维素纤维中的应用。

微生物的保藏

用含有本发明纤维素酶基因的质粒 pNCE4Sa1（参见实施例 A5）转化的大肠杆菌 JM109 菌株保藏于通商产业省工业技术院生命工程工业技术研究所（日本国茨城县筑波市东 1-1-3），保藏号为 FERM BP-5976（原保藏号:FERM P-15732，原保藏日:1996 年 7 月 12 日）。

纤维素酶及其基因

本发明的纤维素酶是具有序列 1 记载的氨基酸序列的一部分或含有 1~284 号序列的蛋白质。本发明中所谓序列 1 记载的氨基酸序列的一部分是指可用作探针程度的序列长度，而且即使是其中的一部分序列也仍然具有纤维素酶的活性，尤其是指维持葡聚糖内切酶活性的那部分序列。

本发明中，也包含在上述蛋白质的 N 末端还含有序列 1 的 -22~-1 的氨基酸序列的一部分或全部的蛋白质。由于序列 1 的 -22~-1 的氨基酸序列被认为是信号肽，所以所谓的其中一部分序列也意味着是加在保持信号肽活性的那部分序列上的，留在 N 末端的序列，其结果是因表达宿主的种类不同而在加工的位置上产生了差别。

另外，本发明包括上述蛋白质的变化蛋白质。本发明中所谓的变

化蛋白质是指在上述蛋白质的氨基酸序列中，附加、插入、剔除、缺失或置换一~数个氨基酸后而产生了变化的蛋白质，但仍然保持纤维素酶活性，尤其是保持葡聚糖内切酶活性的蛋白质。

以下序列表1记载的氨基酸序列中1~284的序列称之纤维素酶NCE4，而该酶基因有时称之纤维素酶NCE4基因。

本发明的纤维素酶活性高，只需少量就可以在各种用途中获得所期望的效果。例如，如果按照本发明的优选状态，将NCE4与从*Humicola insolens*菌的培养液制备的未纯化的纤维素酶比较，前者的用量大约为后者的百分之一就可看到同等纤维素系纤维的细毛去除效果，若是二十五分之一的用量就可看到同等的染色斜纹粗棉布中含有的含纤维素纤维的脱色效果，若是五分之一用量就可看到同等纤维素系纤维的减量效果。因此，本发明纤维素酶可以在要效率更高而且经济的纤维素系纤维的处理中使用。

按照本发明的其它实施方案，可以提供编码上述蛋白质的氨基酸序列的DNA序列。该DNA序列的典型序列应当是含有序列2记载的碱基序列的一部分或全部序列。

序列2记载的碱基序列含有开始于118~120的ATG，终止于1089~1091的TAA的开放读框。而171~173的碱基序列对应于由284个残基组成的上述成熟蛋白质。另外在序列2的碱基序列中确认存在着内含子（参照实施例A7）。

如果给出蛋白质的氨基酸序列，很容易确定编码蛋白质的DNA序列，可以选择编码序列1记载的氨基酸序列的全部或部分序列的各种碱基序列。因此，所谓编码本发明序列1记载的氨基酸序列的一部分或全部序列的DNA序列，除序列2记载的一部分或全部的碱基序列上的序列外，也是指编码同一氨基酸序列的但以存在简并关系的密码为碱基序列的序列。

本发明的DNA既可以是来自天然的DNA，也可以全是合成的。而即使是利用一部分来自天然的序列进行合成的也可以。作为获得DNA的典型方法有来自*Humicola insolens*菌的染色体文库或cDNA文库的在

基因工程领域中惯用的方法，有利用例如以部分氨基酸序列信息为基础作成适当DNA探针进行筛选的方法等。另外，从上述保藏的菌株获得DNA也是可能的。

表达载体和转化微生物

另外，根据本发明，使上述的本发明的DNA序列在宿主微生物中复制是可能的，而且可以提供含有能够表达该DNA编码的蛋白质的载体。同时，根据本发明，能够提供通过该载体转化的微生物。该宿主-载体系统并没有特别限定，例如可以使用大肠杆菌、放线菌、酵母、霉菌等系统，以及使用与其它蛋白质融合的融合蛋白质表达系统，本发明的载体的构建过程和方法可以使用基因工程领域中常用的方法。

为了使本发明的载体实际上导入宿主微生物中，表达出所期望的蛋白质，该载体除了含有上述的本发明的DNA以外，最好还含有控制目的DNA表达的DNA序列以及选择微生物的标记基因等。由于序列2的碱基序列已含有这些控制序列，所以直接利用该碱基序列有时随场合不同也是有利的。该表达载体即使含有重复的编码纤维素酶的DNA序列（串联的）也可以。含有这些DNA序列的表达载体可以按照常规方法构建。而通过该表达载体对微生物的转化也可以按照该领域中常用的方法实施。

用适当的培养基培养该转化体，从该培养物中可以分离到上述本发明的蛋白质。此外，本发明还提供上述本发明的新蛋白质的制造方法。转化体的培养和培养条件只要是与使用的微生物所采用的在本质上一样的条件就可以。而从培养液中回收、纯化本发明的新蛋白质也可以按照常规方法进行。

纤维素酶的用途/纤维素酶制剂

本发明的上述纤维素酶可以直接或作成纤维素酶制剂用于各种用途。尤其是为了赋予纤维素系纤维所期望的特性而使用纤维素酶。具体讲，纤维素酶被用于含纤维素纤维细毛的去除、减量加工、以及染色的斜纹粗棉布中含纤维素纤维的脱色加工。

纤维素酶制剂是将本发明的纤维素酶与通常纤维素酶制剂含有的成分，例如和赋形剂（如乳糖、氯化钠、山梨糖）、表面活性剂、防腐剂等混合在一起制备的。

利用本发明的纤维素酶或纤维素酶制剂去除含纤维素纤维细毛、减量加工以及斜纹粗棉布染色中含纤维素纤维的脱色加工可以通过使纤维素酶与含纤维素纤维接触来实施。

接触温度、纤维素酶的用量等条件可以在考虑了其它各种条件之后适当地确定，例如在去除含纤维素纤维细毛时，可以于50~60℃温度下通过使用蛋白质浓度为5~50mg/l的纤维素酶进行处理。而在进行减量加工时，可以于50~60℃温度下通过使用蛋白质浓度为100~300mg/l的纤维素酶进行处理。如果进行斜纹粗棉布染色中含纤维素纤维的脱色加工时，可以于50~60℃温度下通过使用蛋白质浓度为2~10mg/l的纤维素酶进行处理。

以下通过实施例更详细地说明本发明，但本发明并不限于这些例子。

实施例A 1: 从*Humicola insolens*分离纯化具有使Tencel细毛除去的活性成分

将*Humicola insolens* MN 200-1于37℃下在(N)培养基(5.0%微晶纤维素(Avicel)、2.0%酵母提取物、0.1%蛋白胨、0.03%氯化钙、0.03%硫酸镁、pH6.8)中培养。培养七天后，将得到的培养液用7000rpm离心20分钟，除去菌体，培养液的上清液就作为粗纤维素酶液。

将粗纤维素酶液进行疏水层析(Phenyl-Sepharose High Performance 16/100, Phamacia Biotech公司制)，用浓度梯度为1-0 M硫酸铵的50mM磷酸缓冲液(pH7.0)洗脱，分级。其中，由于在浓度梯度为1-0 M硫酸铵洗脱时得到的级分表现出很强的去除Tencel细毛的活性，所以再将收集的该级分进行疏水层析(Phenyl-Sepharose High Performance 16/100)，用浓度梯度为1-0 M硫酸铵的50mM磷酸缓冲液(pH7.0)洗脱，收集活性成分。

将收集的活性级分进行反相层析 (Source15 IS0, Phamacia Biotech公司制), 用浓度梯度为1-0 M硫酸铵的50mM磷酸缓冲液 (pH 7.0) 洗脱, 分级。其中, 由于在0 M硫酸铵的50mM磷酸缓冲液洗脱时得到的级分表现出很强的去除Tencel细毛的活性, 所以将该级分进行反相层析 (Source15 PHE, Phamacia Biotech公司制), 用浓度梯度为1-0 M硫酸铵的50mM磷酸缓冲液 (pH7.0) 洗脱, 将分离出的表现出很强的去除Tencel细毛的活性的级分作为纯化酶NCE4。该NCE4在SDS-PAGE中显示出分子量为43kDa的单一一条带。

实施例A2: 纤维素酶NCE4的部分氨基酸序列

(1) N末端氨基酸残基的确定

为了确定实施例1中纯化的蛋白质的N末端氨基酸序列, 使用FPLC系统 (Phamacia Biotech公司制) 进行柱层析 (柱子: RESOURCE (商品名) RPC 3ml、含有0.1%的TFA的5%~60%乙腈梯度), 分别收集主要的峰。

将收集的成分冷冻干燥后, 溶解于少量的水中, 使用小型电泳装置 (TEFCO公司制造), 进行胶浓度为8%的SDS-PAGE电泳。电泳后, 利用Multiphor II电泳装置 (Phamacia Biotech公司制) 将胶中的蛋白质经电转移到PVDF膜 (Millipore公司) 上, 然后用考马斯亮蓝 R-250 (Nakalai-texque公司) 染色后, 再脱色, 经水洗干净, 最后风干。从PVDF膜上将分子量为43kDa的蛋白质斑剪下来供蛋白质序列仪Model 492 (Perkin Elmer公司) 进行序列分析。确定了N末端一侧15个氨基酸残基序列。得到的氨基酸序列如下所示。

N末端氨基酸序列: Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp-(Cys)-(Cys)-Lys-Pro-Ser (15个残基)

(2) 肽谱峰

将通过上述(1)中FPLC纯化的蛋白质冷冻干燥后, 溶解于100mM的碳酸氢铵缓冲液中 (pH8.0)。然后添加大约相对于蛋白质摩尔量的1/20的胰蛋白酶 (Promega公司制造), 于37℃反应48小时。用Model 172 μ 制备量HPLC系统 (Perkin Elmer公司) 进行柱层析 (柱子: C8

220 × 2.1mm, 0.1%TFA、0% 乙腈 ~ - 0.085%TFA、35% 乙腈梯度), 分别获得3种肽。然后通过上述的蛋白质序列仪确定得到的肽段的氨基酸序列。其结果如下所示。

TP - 1: Tyr-Gly-Gly-Ile-Ser-Ser (6个残基)

TP - 2: Phe-Pro-Asp-Ala-Leu-Lys (6个残基)

TP - 3: Phe-Asp-Trp-Phe-Lys-Asn-Ala-Asp-Asn-Pro-Ser-Phe-Ser-Phe-Arg

* * * * *

(15个残基)

这些N末端氨基酸序列和通过肽段峰得到的氨基酸序列由于与W091/17243号公报记载的由*Humicola insolens* DSM1800得到的43kDa葡聚糖内切酶的氨基酸序列具有同源性, 所以明显地暗示出该蛋白质是纤维素酶的一种。

另外, 将上述序列与Protein Identification Resource(PIR) R 44.0, March, 1995, 或者SWISS-PROT R31.0, March 1995登录的序列比较, 虽然具有表现出同源性的序列, 但并不是同一种蛋白质, 而是一种新的蛋白质已经很清楚。

实施例A3: 基因组DNA文库的制作

基因组DNA的分离按照Horiuchi等人的方法(Horiyuki Horiuchi等, 《细菌学杂志》J. Bacteriol., 170:272-278, 1988)进行。

首先将*Humicola insolens* MN 200-1于37℃下, 在上述(N)培养基中培养。培养2天后, 通过离心分离(3500rpm、10分钟)收集菌体。将得到的菌体进行苯酚处理、蛋白酶K和核糖核酸酶A处理、再通过聚乙二醇(PEG)沉淀, 得到基因组DNA。

接下来, 用Sau3A I消化*Humicola insolens*基因组DNA, 通过琼脂糖凝胶电泳确认在9~23kbp的范围有部分分解带, 然后利用乙醇沉淀回收该电泳带。使用T4连接酶(东洋纺织社制造)将该DNA片段连接在噬菌体载体、EMBL3克隆试剂盒(Stratagene 公司制造)中的BamH I臂上。用乙醇沉淀后, 溶解于TE(10mM Tris盐酸(pH8.0)、1mMEDTA)缓冲液中。

使用就象Hohn, B. 方法 (Hohn, B. 《酶学方法》Methods Enzymol., 68:299-309, 1979) 记载的那样冷冻干燥的包装成分和Giga-Pack II包装试剂盒 (Stratagene 公司制造), 将上述连接后的混合物都包装在 λ 噬菌体头部, 然后用得到的噬菌体感染大肠杆菌LE392菌株。使用通过这种方法得到的 5×10^4 个噬菌体文库进行目的基因的克隆。

实施例A4: 利用PCR方法制作长链探针

作为DNA探针是通过以*Humicola insolens*全长DNA作模板利用PCR方法制作的扩增的长链探针。

各个引物是对应于N末端和多肽TP-3中用*表示的氨基酸合成的DNA。制作的合成寡核糖核酸序列如下所示。

NCE4N1: 5'-GCXGA(CT)GGXAA(AG)TC(AGCT)AC-3' (17mer)

NCE4N2: 5' GCXGA(CT)GGXAA(AG)AG(CT)AC-3' (17mer)

NCE4C: 5'-CXGC(AG)TT(CT)TT(AG)AACCA(AG)TC-3' (19mer)

(X: 次黄嘌呤核苷)

PCR反应的条件如下。首先对于1 μ g *Humicola insolens*基因组DNA, 一组的试管中加入引物NCE4N1、NCE4C各1 μ M, 另一组试管中加入NCE4N2、NCE4C各1 μ M, 在dNTP存在下, 于95 $^{\circ}$ C进行5分钟热变性, 然后加入Taq聚合酶 (重组Taq, 宝酒造公司制造), 通过25轮反复反应进行扩增, 每次的反应条件都是94 $^{\circ}$ C 1分钟, 45 $^{\circ}$ C 2分钟, 72 $^{\circ}$ C 3分钟。扩增的结果是, 只使用引物NCE4N1、NCE4C时, 扩增得到大约750bp的DNA。这些DNA用作以下的筛选探针。

实施例A5: 纤维素酶成分NCE4基因的克隆

(1) 利用噬斑原位杂交进行筛选

首先利用ECL Director DNA/RNA标记检测系统 (Amersham公司制造) 对通过PCR扩增的大约750bp的DNA片段100ng预先进行标记。

将按照实施例2记载的方法制作的噬菌体噬斑转移到High-Bond N + 尼龙转移膜 (Amersham公司制造) 上, 用0.4N的氢氧化钠变性

后, 用5倍浓度的SSC (15mM柠檬酸三钠盐, 150mM氯化钠) 洗净, 干燥, 使DNA固定。按照试剂盒的方法, 进行1小时的预杂交 (42℃) 后, 加入先前标记的探针, 进行4小时 (42℃) 的杂交。按照上述试剂盒的方法洗净标记物。首先用加有0.4%SDS、6M尿素的0.5倍浓度的SSC于42℃反复洗2次, 每次20分钟, 然后再用2倍浓度的SSC于室温反复洗2次, 每次5分钟。

洗过探针的尼龙膜浸入到检测液中1分钟, 然后使Hyper-Film-EC L (Amersham公司制造) 感光, 得到4个阳性克隆。

(2) 噬菌体DNA的制备

使噬菌体感染E. coli LE392, 8小时后收集噬菌体颗粒, 按照Grossberger的方法 (Grossberger, 《核酸研究》D., Nucleic Acids. Res. 15 6737, 1987) 用蛋白激酶K和苯酚处理后, 通过乙醇沉淀, 分离出噬菌体DNA。

(3) 目的基因的亚克隆

4种噬菌体DNA用Sal I切, 然后走琼脂糖电泳。

按照Southern的方法 (Southern, E. M., 《分子生物学杂志》J. Mol. Biol. 98:503-517, 1975) 将DNA转移到尼龙膜上, 在与上述 (1) 的噬斑杂交相同的条件下, 用大约750bp的探针进行杂交, 检测出含有5.2kbp的目的基因的DNA片段。其结果4种噬菌体DNA都有同样大小的Sal I片段。

用Sephaglass Band Prep试剂盒 (Pharmacia Biotech公司制造) 分离5.2kbp DNA片段, 使用E. coli JM109在质粒pUC119的Sal I部位进行亚克隆。得到的质粒定为pNCE4Sal。

实施例A6: 碱基序列的确定

(1) 基因组DNA碱基序列的解析

碱基序列的确定按以下方式实施。

碱基序列解析设备使用的是A. L. F. DNA序列仪II (Pharmacia Biotech公司制造)。作为测序胶使用的是ReadyMix胶 (Pharmacia Biotech公司制造), 或者使用可能作为HydroLink Long Ranger (FMC公

司制造)得到的丙烯酰胺载体。作胶用的各种试剂(N, N, N', N'-四甲基乙二胺、尿素、过硫酸铵)可以使用A. L. F级试剂(Pharmacia Biotech公司制造)。解读碱基序列反应使用的是自动解读测序试剂盒Autoread Sequencing Kit(Pharmacia Biotech公司制造)。凝胶制作条件、反应条件和电泳条件等参照各个说明书设定。

将作为模板DNA的pNCE4Sal用10 μg的2M氢氧化钠进行碱变性后,与自动解读测序试剂盒附带的Universal 引物退火,进行延长反应。反应产物用测序仪解读,判断有546bp碱基序列。从这一结果可以制作MNEG 01的FITC标记的测序引物,与pNCE4Sal反应,可进一步进行解读。由得到的结果又可制作下一个引物,继续解读。其结果就可解读出整个的NCE4。制作的FITC标记测序引物如下所示。

MNEG - 01: 5'-GTGATGAGGGCTGGCGACAGGCC-3' (23mer)

MNEG - 02: 5'-CTGCCACCTCTATTGCCGGCAGC-3' (23mer)

MNEG - 03: 5'-CCCGACGCCCTCAAGCCCGGCTG-3' (23mer)

MNEG - 04: 5'-GGCTGGAGCGGCTGCACCACCTG-3' (23mer)

(2) 碱基序列的确定

以上述(1)的结果为基础,合成MNEG - 05 ~ MNEG - 08的FITC标记测序引物,制作的FITC标记测序引物如下所示。

MNEG - 05: 5'-GACCTGACGGAAGCTGAAGCTCG-3' (23mer)

MNEG - 06: 5'-AGCAGTGCAGCCGCTGGGAGTCG-3' (23mer)

MNEG - 07: 5'-TGGCAGATGAGGACCTGGTGTG-3' (23mer)

MNEG - 08: 5'-CGCAGCCGGACTTGGCGTCGAAG-3' (23mer)

使这些引物与pNCE4Sal通过自动解读测序试剂盒进行反应。首先,对10 μg的质粒进行碱变性,然后与各个引物进行退火,用T7聚合酶使其反应。其结果可以确定Sal I片段内的1257bp碱基序列。该序列如序列3所示。

实施例A7: 内含子的确定

为了确定内含子,首先从*Humicola insolens* MN200-1制备mRNA,通过逆转录酶合成cDNA,将它与基因组的碱基序列比较,判断它

们的相同区域。

(1) 总RNA的制备

将*Humicola insolens* MN200-1于纤维素酶诱导培养基培养, 优选于上述(N)培养基中培养2天, 通过离心分离(3500rpm、10分钟), 收集菌体。取出其中的2g菌体用灭菌水洗净, 于液氮冷冻状态下用搅拌机(日本精机公司制造的匀浆机AM-3)粉碎。然后悬浮于含有4M胍基硫氰酸盐的变性溶液10ml中(4M胍基硫氰酸盐、25mM柠檬酸三钠盐、0.5%N-十二烷基肌氨酸钠、0.1M巯基乙醇)。于室温搅拌数分钟后, 用1ml的2M醋酸钠(pH4.5)中和, 加入10ml的TE饱和的苯酚, 继续搅拌。然后加入2ml的氯仿-异戊醇(24:1), 充分搅拌后, 通过离心分离(3500rpm、10分钟), 除去苯酚变性了的菌体成分。吸取上层液(水相), 用10ml的异丙醇将核酸沉淀。通过离心分离(3500rpm、10分钟)从该沉淀中回收核酸, 用70%乙醇通过再离心分离洗沉淀。

将该沉淀溶解于3.5ml的TE后, 向溶液中加入880 μ l的10M氯化锂溶液, 于5℃冷藏2小时后, 经离心分离(12000rpm、10分钟), 回收沉淀。用70%乙醇洗该沉淀, 得到沉淀即为总RNA级分。回收RNA的量为2.7mg, 收率是0.14%。

(2) PolyA尾巴⁺RNA(=mRNA)的制备

mRNA的制备是利用mRNA纯化试剂盒(Pharmacia Biotech公司制造)进行的。

首先, 从(1)中制备的总RNA中取出1mg溶解于1ml的洗脱缓冲液中, 于65℃进行10分钟的热变性处理。然后于冰中骤冷后, 加入0.2ml的样品缓冲液。将该总RNA溶液加到寡聚物(dT)纤维素柱上, 用高盐缓冲液洗柱3次, 再用低盐缓冲液洗柱3次, 然后用于65℃加温的洗脱缓冲液洗脱。上述对柱子的操作反复进行2次。得到mRNA级分。得到的mRNA量为19.2 μ g, 收率为2%。

(3) cDNA的合成

使用Timesaver cDNA合成试剂盒(Pharmacia Biotech公司制造)

进行cDNA合成。

首先将5 μ g的mRNA溶解于20 μ l的样品缓冲液。于65℃进行10分钟的热变性处理后，与二硫苏糖醇溶液以及寡聚物(dT)引物一起加到第一条链合成混合物中，于37℃反应1小时。然后再都加到第二条链合成混合物中，于12℃反应30分钟，再于22℃反应1小时，得到cDNA。

(4) 通过PCR方法扩增纤维素酶NCE4cDNA

从合成的cDNA取出1 μ g作为模板，通过PCR方法只扩增目的cDNA。制作的N末端和C末端引物的寡核苷酸序列如下所示。

NCE4 - CN: 5' -ATGCGTTCCTCCCTCTCCTCCGCTCCGCC-3' (30mer)

NCE4 - CC: 5' -TACAGGCACTGATGGTACCAGTCATTAATC-3' (30mer)

PCR反应按以下条件进行。首先对于1 μ g *Humicola insolens*的cDNA，各加入1 μ M引物，在dNTP存在下，于94℃进行10分钟热变性，然后加入Taq聚合酶(重组Taq, 宝酒造公司制造)，通过30轮反复反应进行扩增，每次的反应条件都是94℃1分钟，50℃2分钟，72℃3分钟。扩增的片段经琼脂糖电泳确定为0.9bp大小的片段。通过乙醇沉淀浓缩该片段，利用pT7 blue-T vector试剂盒(Novagen公司制造)进行克隆。该质粒定为pCNCE4。

(5) cDNA碱基序列的解析

使用与上述一样的自动解读测序试剂盒进行测序反应。

质粒pCNCE4用2M的氢氧化钠进行碱变性后，用乙醇进行沉淀。该一条链质粒作为模板，用T7聚合酶进行聚合反应。利用前面合成的引物MNEG-01、MNEG-02、MNEG-03、MNEG-04、MNEG-05、MNEG-06、MNEG-07、MNEG-08以及试剂盒附带的Universal引物、Reverse引物进行反应，解读序列。

结果发现存在着一个56bp的内含子。在序列3的序列中，非翻译开始序列以及终止序列、内含子内部的调控序列如下所示(数字为序列3中的序列位置号)。

内含子: 453~458、506~508、491~497

实施例A8: NCE4去除含纤维素纤维细毛活性的评价

使用大型洗衣机, 加有表面活性剂和橡皮球, 使预先染色的Tencel (Courtauls公司制造) 布料生出细毛。将通过处理生出细毛的Tencel按照下面的条件用纤维素酶处理除去细毛, 计算为了使细毛完全除去所需要的纤维素酶的蛋白质浓度。

试验机械: 20kg洗衣机 (SANYO公司制造 全自动洗衣机SCW5101)

浴比: 1: 20

加热: 55℃

时间: 60分钟

pH : 7 (10mM磷酸缓冲液)

橡皮球与纤维素酶液一起加入到处理液中, 总重量大约是布的2倍。

表1

	除去Tencel细毛 所需要的蛋白质量*
NCE4	5mg/L
粗制纤维素酶液	500mg/L

*:蛋白质量是使用蛋白质分析试剂盒 (Biorad公司)、以牛血清白蛋白作为标准来定量的。

结果表明, NCE4的用量为粗制纤维素酶液的百分之一就可以使同等程度的Tencel细毛除去。

实施例A9: NCE4对斜纹粗棉布染色中含纤维素纤维的脱色活性的评价

对已下水脱浆的12-Ounce的蓝色牛仔褲按照下面的条件进行脱色处理。

试验机械: 20kg洗衣机 (SANYO公司制造 全自动洗衣机SCW5101)

浴比: 1: 50

加热: 55℃

时间: 60分钟

pH : 7 (10mM磷酸缓冲液)

橡皮球与纤维素酶液一起加入到处理液中, 总重量大约是布的2倍。

脱色度用色差系COLOR ANSLYZER TOPSCAN MODELTC-1800MK2 (东京电色株式会社制造) 测定, Lab表示系列的L值(明度)。通过L值相对于对照的增加值(白色度的增加) = ΔL 来评价脱色, 就评价脱色的各个试验区测定了5个点的 ΔL 值($n=5$), 去掉最大值和最小值, 采用剩下的3点的平均值。可以算出为了使脱色达到 ΔL 值 = 7所需要的纤维素酶的蛋白质浓度。

表2

	蓝色牛仔裤脱色 所需要的蛋白质量
NCE4	1.8mg/L
粗制纤维素酶液	45.0mg/L

试验结果表明NCE4的用量为粗制纤维素酶液的二十五分之一就可以使同等程度的蓝色牛仔裤脱色。

实施例A10: NCE4对含纤维素纤维的减量活性的评价

按照下面条件对作为预先测定了绝对干重的再生纤维素系纤维的库普拉(铜铵短纤维, 旭化成工业株式会社制: 长15cm×宽10cm)进行酶处理。

试验机械: 洗涤牢固度试验机 型号L-12(株式会社大荣科学精密仪器制作所制造)

浴比: 1: 50

加热: 55℃

时间：60分钟

pH：7（40mM磷酸缓冲液）

不锈钢球与纤维素酶液一起加入到处理液中。酶处理后，使其干燥，测定库普利的绝对干重，测定相对于酶处理前的重量减少率。计算出为了达到8%的重量减少率所需要的纤维素酶的蛋白质浓度。

表3

为使库普拉重量减少8% 所需要的蛋白质量	
NCE4	100mg/L
粗制纤维素酶液	500mg/L

试验结果表明NCE4的用量为粗制纤维素酶液的五分之一就可以实现同等程度的减量加工。

实施例B1：质粒pMKD01的制备

（1）质粒pUC118BN的制备

用BamH I切1 μ g pUC118 DNA，利用苯酚使限制性内切酶失活。然后进行乙醇沉淀，将沉淀溶解于少量的TE（10mM Tris盐酸（pH8.0）、1mM EDTA）缓冲液中。使用DNA Branching kit（宝酒造社制）使上述DNA的末端补平。再利用DNA连接试剂盒使平滑化的DNA连接，使其自身环化。用得到的连接混合物转化E. coli感受态细胞 JM 109（宝酒造社制）。转化菌可以在含有100 μ g/ml的氨苄青霉素、1mM IPTG、0.004% X-gal的 LB琼脂培养基（1%蛋白胨、0.5%酵母提取物、1% NaCl、1.5%琼脂）上繁殖，选择白色的菌斑在含有100 μ g/ml的氨苄青霉素LB培养基（1%蛋白胨、0.5%酵母提取物、1%NaCl）上于37℃培养过夜。从得到的培养基中通过碱-SDS方法回收质粒DNA。将该质粒DNA用BamH I切，走0.8%琼脂糖凝胶电泳，选择pUC118 DNA被BamH I切断的质粒DNA。该质粒DNA被定为pUC118BN。

（2）质粒pUC118BSN的制备

用Sph I切1 μ g pUC118BN DNA, 利用与上述同样的方法获得pUC118BN DNA的Sph I部位被破坏的质粒DNA。该质粒DNA被定为pUC118BSN。

(3) 质粒pM21的制备

(A) 纤维素酶NCE2基因的分离

将按照特开平8-126492号公报记载的方法从*Humicola insolens*得到的含有作为纤维素酶NCE2基因、以及含有作为该基因启动子和终止区的上游1.4kb、下游0.5kb的DNA序列的全长3.4kb的Pst I~Xba I的酶切片段连接到先前的pUC118BSN Pst I~Xba I位点上, 得到的质粒DNA被定为pUC118BSN-PX。

(B) pUC118BSN-PX的指定部位的变异处理

在紧靠NCE2基因的N末端的下游以及终止密码的下游象以下那样通过指定部位变异导入BamH I位点。通过质粒pUC118BSN-PX转化*E. coli* JM109菌株, 使辅助噬菌体M13K07感染后, 在含有氨苄青霉素和卡那霉素各150 μ g/ml、70 μ g/ml的30ml的2 $\times$ YT液体培养基(1.6%细菌胰蛋白酶、0.8%酵母提取物、0.5% NaCl)上于37 $^{\circ}$ C培养16~20小时。从培养上清中回收M13的单链DNA(ssDNA)。利用该ssDNA和两种合成的寡核苷酸, 使用雕纹体外诱变系统(Sculpture In Vitro Mutagenesis System (Amersham公司制造))进行特定部位的变异处理。制备的合成寡核苷酸引物如下所示。

MNC-02 5'-GAGCGCCAGAACTGTGGATCCACTTGGTGAGCAATG-3' (36mer)

MNC-03 5'-TCCGCCGTTCTGAGCGGATCCAGGCGTTTGGCGCG-3' (35mer)

将经指定部位变异处理的DNA混合液导入*E. coli* TG1, 得到的转化菌株在含有100 μ g/ml氨苄青霉素的LB培养基(1% 胨、0.5%酵母提取物、1% NaCl)中培养, 回收质粒DNA。该质粒DNA用BamH I切, 走0.8%琼脂糖凝胶电泳, 选择在pUC118BSN-PX中导入两个BamH I位点的质粒DNA, 该质粒DNA被定为pM21。

(4) 纤维素酶NCE3基因的分离

以来自众所周知的灰腐质霉*Humicola grisea*纤维二糖水解酶基因(de Oliveira Alzevedo, M. 等,《普通分子生物学杂志》J. General Microbiol., 136:2569-2576, 1990)的序列为基础利用PCR方法分离来自*Humicola insolens*纤维二糖水解酶基因(NCE3)。

(A) 基因组DNA的分离

利用上述实施例A3的方法得到*Humicola insolens* MN200-1的基因组。

(B) 利用PCR方法扩增纤维素酶NCE3基因

以来自*Humicola grisea*的纤维二糖水解酶基因的序列为基础利用PCR方法分离*Humicola insolens*的NCE3基因。为了能够使含有该NCE3的PCR产物连接在质粒pM21的BamH I位点上, 预先在各引物中设计了含有BamH I位点, 作为引物合成了如下的寡核苷酸序列。

MKA-05: 5'-GCCGCCAGCAGGCGGGATCCCTCACCACCGAGAGG-3' (36mer)

MKA-06: 5'-TGATCGTCGAGTCAGGGATCCAGAATTTACAGGCAC-3' (36mer)

PCR反应根据LA PCR Kit Ver.2(宝酒造社制), 按照以下方法进行。首先对通过上述方法得到的1 μ g *Humicola insolens*基因组DNA, 各加入1 μ M引物、400 μ M dNTP、LA Taq聚合酶2.5U, 通过30轮重复反应进行扩增, 每次的反应条件都是94 $^{\circ}$ C 1分钟, 55 $^{\circ}$ C 2分钟, 72 $^{\circ}$ C 3分钟。扩增的片段经0.8% 琼脂糖凝胶电泳确认1.6 Kb DNA的扩增。该1.6 Kb DNA的片段用Sephaglass Band Prep 试剂盒(Pharmacia Biotech. 公司制造)回收, 并与pT7 Blue-T-vector kit (Novagen公司制造)连接。该质粒DNA定为pK21。

(5) 质粒pKM04的制备

质粒pK21DNA用BamH I消化, 回收1.6 Kbp片段。然后质粒pM21DNA用BamH I消化, 再于70 $^{\circ}$ C 处理10分钟, 使限制性内切酶失活。再通过小牛碱性磷酸酶脱磷酸后, 经0.8% 琼脂糖凝胶电泳分离, 回收5.2 Kbp DNA片段。最后将来自pK21的1.6 Kbp片段与来自pM21的5.2 Kbp DNA片段连接, 就得到质粒pKM04。

(6) 质粒pMKD01的制备

首先, 使用通过众所周知的方法得到的来自构巢曲霉*Aspergillus nidulans*的trp C基因的启动子和终止子 (Mullaney, E. J. 等, 《分子遗传学与普通遗传学》Mol. Gen. Genet. 199:37-45, 1985), 将特开昭59-175889号公报记载的越霉素抗性基因制备成在*Humicola insolens*中能够表达的基因。将该基因导入质粒pMKD04的Xba I位点, 制备质粒pMKD01。

实施例B2: 利用质粒pMKD01转化*Humicola insolens*

(1) 质粒pMKD01高浓度纯化样品的制备

为了使质粒pMKD01导入*Humicola insolens*, 首先要制备纯化的高浓度质粒pMKD01。将pMKD01导入*E. coli* JM109, 在含有100 $\mu\text{g}/\text{ml}$ 氨苄青霉素的LB培养基中于37℃培养过夜。得到的培养液用Flexiprep kit (Pharmacia Biotech公司制造) 纯化, 得到1 $\mu\text{g}/\mu\text{l}$ 的pMKD01质粒DNA。

(2) 转化*Humicola insolens*

将*Humicola insolens* MN200-1在(S)培养基中于37℃培养24小时后, 通过3000rpm、离心分离10分钟, 收集菌体。其中(S)培养基的组成是上述(N)培养基中加了葡萄糖(3.0%), 并去掉了Avicel后的培养基。得到的菌体用0.5M蔗糖洗净, 悬浮于经0.45 μm 滤器过滤的10ml的原生质体化酶液中(5mg/ml Novozyme 234(NL I公司制造)、5mg/ml Cellulase Onozuka R-10(Yacult公司制造)、0.5M蔗糖)。于30℃振荡60~90分钟, 使菌原生质体化。将该悬浊液过滤后, 经2500rpm、10分钟离心分离, 回收原生质体, 用SUTC缓冲液(0.5M蔗糖、10mM氯化钙、10mM TRIS盐酸(pH7.5))洗净。

将以上制备的原生质体悬浮于1ml的SUTC缓冲液中, 向100 μl 该悬浮液中加入10 μg 的DNA (TE) 溶液(10 μl), 静置于冰中5分钟。然后加入400 μl 的PEG溶液(60% PEG4000、10mM氯化钙、10mM TRIS盐酸(pH7.5)), 静置于冰中20分钟后, 加入10ml的SUTC缓冲液, 经2500rpm、10分钟离心分离, 收集原生质体。将该原生质体悬

浮于1ml的SUTC缓冲液后，经4000rpm、5分钟离心分离，将原生质体最终悬浮于100 μ l的SUTC缓冲液中。

将经以上处理的原生质体与YMG软琼脂一起铺在含有200 μ g/ml潮霉素B的YMG培养基（1%葡萄糖、0.4%酵母提取液、0.2%麦芽提取液、1%琼脂（pH6.8））上，于37℃培养5天后，形成的菌落就是转化体。

（3）pMKD01转化的菌株的培养以及经SDS - PAGE的评价

象前述那样将质粒pMKD01导入*Humicola insolens* MN200 - 1，挑选50个潮霉素抗性菌株。将这些菌株在（N）培养基上于37℃培养5天。获得的培养液上清通过SDS - PAGE解析，pMKD01转化的菌株中的5个菌落推断为NCE3的蛋白质带比亲本菌株增加了3~4倍。

（4）重组NCE3的N末端氨基酸残基的鉴定

为了进一步确认SDS - PAGE结果给出的大量表达的蛋白质带是来自NCE3基因的，所以要确定该蛋白质的N末端氨基酸序列。首先，就从亲本菌株和NCE3高表达菌株得到的培养上清按照前述实施例A2的方法进行FPLC柱层析，比较它们的主要的峰。收集NCE3高表达菌株中特别增加的峰，并冷冻干燥。冷冻干燥的样品溶解于少量的水中，进行胶浓度为8%的微型SDS - PAGE(Difco公司制造)。然后按照前述实施例A2的方法，将胶中蛋白质电转移到PVDF膜上，用考马斯亮蓝R - 250染色后，进行脱色，用水洗净。从膜中剪下印记为分子量66KD的蛋白质部分，将该印记蛋白质按照Podell, D. N等人的方法（Podell, D. N. 等，《生物化学与生物物理学研究通讯》*Biochem. Biophys. Res. Commun.*, 81:176, 1978），除去修饰N末端残基。首先，剪下目的蛋白质，在少量的0.5%聚乙烯吡咯烷酮（分子量40,000 Sigma公司制造）/ 100mM醋酸溶液中于37℃保温30分钟后，用水充分洗净。然后利用Pfu焦谷氨酸氨肽酶（宝酒造公司制造）除去修饰的N末端残基，用水洗净、风干。利用蛋白质测序仪Model 492，确定了N末端的15个氨基酸残基序列。得到的序列如下所示。

N末端氨基酸序列: Asn-Cys-Gly-Ser-Leu-Thr-Thr-Glu-Arg-His-Pro-Ser
-Leu-Ser-Trp (15个氨基酸残基)

该N末端氨基酸序列与由质粒pMKD01的碱基序列推测的纤维素酶NCE2、NCE3融合蛋白质的氨基酸序列一致。

(5) pMKD01转化的菌株的FPLC评价

为了进一步定量象上述那样用SDS - PAGE确认NCE3大量表达的5个菌落的培养上清, 利用FPLC柱进行柱层析。层析条件与上述(4)的一样。分别收集NCE3的峰, 并冷冻干燥, 测定重量, 比较高表达菌株与亲本菌株的产率。比较结果如下表所示。

表4

	NCE3产量*
<i>Humicola insolens</i> MN200-1(亲本菌株)	0.46g
<i>Humicola insolens</i> pMKD01	1.8g

*: 产量是每一升培养液的产量。

实施例B3; 质粒pEGD01的制备

质粒pMKD01经BamH I消化, 然后通过70℃的热处理使限制性内切酶失活, 经脱磷酸处理, 回收8.2Kbp的DNA片段。

然后, 以在上述实施例A1~7得到的来自*Humicola insolens*的NCE4基因的序列为基础, 通过PCR方法扩增NCE4基因。按照在各个引物中预先含有BamH I位点的形式设计该NCE4的PCR产物, 使之能够符合读码框架, 连接于上述质粒pMKD01的8.2Kbp的BamH I片段。

作为引物合成具有以下序列的寡核苷酸。

NCE4 - N: 5'-CCGGTGTTGGCCGATCCGCTGATGGCAAG-3' (30mer)

NCE4 - C: 5'-TAAGGCCCTCAAGGATCCCTGCGTCTACAG-3' (30mer)

PCR反应如下。对于1μg *Humicola insolens*基因组DNA, 各加入1μM引物、400μM dNTP、Pfu DNA聚合酶(Stratagene公司制造)

2.5U, 通过25轮重复反应扩增0.8Kbp的DNA片段, 每次的反应条件都是94℃ 1分钟, 55℃ 2分钟, 72℃ 3分钟。回收该0.8Kbp的DNA片段, 将该片段连接到上述pMKD01的8.2KbpBamH I片段。该质粒DNA定为pEGD01。

实施例B4: 质粒pEGD01的表达

(1) 质粒pEGD01对*Humicola insolens*的转化

质粒pEGD01对*Humicola insolens*MN200-1的转化按照实施例B2的方法进行。首先制备纯化的高浓度的质粒pEGD01, 得到1μg/μl pEGD01质粒DNA。使用10μl该pEGD01溶液, 转化*Humicola insolens* MN200-1, 挑选50个潮霉素抗性菌株。将这些菌株在(N)培养基上于37℃培养5天。获得的培养液上清通过SDS-PAGE解析, pMKD01转化的菌株中的10个菌落中推断为NCE4蛋白质带比亲本菌株增加了10~16倍。

(2) 重组NCE4的N末端氨基酸残基的鉴定

为了进一步确认SDS-PAGE结果给出的确认大量表达的蛋白质带是来自NCE4基因的, 所以要确定该蛋白质的N末端氨基酸序列。首先, 就从亲本菌株和NCE4高表达菌株得到的培养上清进行FPLC柱层析, 比较它们的主要的峰。条件与上述实施例B2一样。收集NCE3高表达菌株中特别增加的峰, 并冷冻干燥。冷冻干燥的样品溶解于少量的水中。按照实施例B2的方法除去修饰N末端残基后, 通过上述蛋白质测序仪确定N末端的氨基酸序列。结果得到比例大约为7:3的如下所示的两种N末端氨基酸序列。而如果不除去修饰N末端残基, 通过上述蛋白质测序仪确定N末端氨基酸序列, 结果只得到如下所示的氨基酸序列1。

N末端氨基酸序列1: Val-Val-Glu-Glu-Arg-Gln-Asn-Cys-Gly-Ser-Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp (20个残基)

N末端氨基酸序列2: Asn-(Cys)-Gly-Ser-Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp-Asp-(Cys)-(Cys)-Lys-Pro-Ser-(Cys) (20个残基)

这些N末端氨基酸序列与由质粒pMKD01的碱基序列推测的纤维素酶NCE2和NCE4融合蛋白质的氨基酸序列一致。因为得到的是两种N末端氨基酸序列，很清楚，当切断该融合蛋白质的信号序列时，会在多个位置进行切断。

(3) pEGD01转化的菌株的FPLC评价

为了进一步定量象上述那样用SDS - PAGE确认NCE4大量表达的5个菌落的培养上清，利用FPLC柱进行柱层析。分别收集NCE4的峰，并冷冻干燥，测定重量，比较高表达菌株与亲本菌株的产率。比较结果如下表所示。

表5

	NCE4产量*
<i>Humicola insolens</i> MN200-1(亲本菌株)	0.28g
<i>Humicola insolens</i> pMKEG1	4.5g

*: 产量是每一升培养液的产量。

实施例B5: 质粒pIED02的制备

(1) 质粒pID01的制备

pEGD01经Hind III ~ BamH I消化，回收7.2Kbp的DNA片段。

然后，以用特开平8-5663号公报中记载的方法得到的来自*Humicola insolens*的NCE1基因的序列为基础，通过PCR方法扩增NCE1基因的启动子和编码信号序列部分的DNA。含有该NCE1启动子和信号序列的PCR产物按照在各个引物中预先含有Hind III 位点、BamH I位点的形式设计，使之能够符合读码框架，连接于上述质粒pEGD01的7.2 Kbp的Hind III ~ BamH I片段。作为引物制备了如下的合成寡核苷酸。

PNCE1 - N: 5'-GTCATGAAGCTTCATTAAGGTACGTATGCAAC-3' (32mer)

PNCE1 - C: 5'-GGTGATGGATCCGGCCTGCTGGGCAGCGACGC-3' (32mer)

PCR反应与实施例3一样。对于1 μ g *Humicola insolens* 基因组DNA, 各加入1 μ M引物、400 μ M dNTP、Pfu DNA聚合酶2.5U, 通过23轮重复反应扩增1.5Kbp的DNA片段, 每轮的反应条件都是94℃1分钟, 55℃2分钟, 72℃4分钟。该PCR产物用Hind III和BamH I消化, 回收1.5Kbp的DNA片段, 将该片段连接到上述pEGD01的7.2Kbp Hind III~BamH I片段。该质粒DNA定为pIED01。

(2) 质粒pIED02的制备

质粒pIED01经BamH I消化, 然后通过70℃的热处理使限制性酶失活, 经脱磷酸处理, 回收8.6Kbp的DNA片段。然后质粒pEGD01经BamH I消化后, 回收含有NCE4基因的0.8Kbp的DNA片段, 将两个片段连接起来, 得到质粒pIED02。

实施例B6: 质粒pIED02的表达

(1) 质粒pIED02对*Humicola insolens*的转化

质粒pIED02对*Humicola insolens* MN200-1的转化按照实施例B2的方法进行。首先制备纯化的高浓度的质粒pIED02, 得到1 μ g/ μ l pIED02质粒DNA。使用10 μ l该pIED02溶液, 转化*Humicola insolens* MN200-1, 挑选50个潮霉素抗性菌株。将这些菌株在(N)培养基上于37℃培养5天。获得的培养液上清通过SDS-PAGE解析, pIED02转化的菌株中的5个菌落中推断为NCE4蛋白质带, 比亲本菌株增加了5~10倍。

(2) 重组NCE4的N末端氨基酸残基的鉴定

为了确认SDS-PAGE的结果给出的大量表达的蛋白质带是来自NCE4基因的, 所以要确定该蛋白质的N末端氨基酸序列。首先, 按照与实施例B2一样的方法, 就从亲本菌株和NCE4高表达菌株得到的培养上清进行FPLC柱层析, 收集NCE3高表达菌株中特别增加的峰, 并冷冻干燥。冷冻干燥的样品溶解于少量的水中。按照实施例B2的方法除去修

饰N末端残基后，通过上述蛋白质测序仪确定N末端的15个氨基酸残基序列。得到的序列如下所示。

N末端氨基酸序列: Gln-Ala-Gly-Ser-Ala-Asp-Gly-Lys-Ser-Thr-Arg-Tyr-Trp
-Asp-(Cys) （15个残基）

该N末端氨基酸序列与由质粒pIED02的碱基序列推测的纤维素酶NCE1、NCE4融合蛋白质的氨基酸序列一致。

（3）pEGD02转化的菌株的FPLC评价

为了进一步定量象上述那样用SDS - PAGE确认的NCE4大量表达的5个克隆的培养上清，利用FPLC柱进行柱层析。分别收集NCE4的峰，并冷冻干燥，测定重量，比较高表达菌株与亲本菌株的产率。比较结果如下表所示。

表6

	NCE4产量*
Humicola insolens MN200-1(亲本菌株)	0.28g
Humicola insolens pIED02	2.9g

*: 产量是每一升培养液的产量。

序列表

序列号: 1

序列的长度: 305

拓扑型: 直链形

序列种类: 蛋白质

序列

```

Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro
-20          -15          -10
Val Leu Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
-5          1          5          10
Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro
15          20          25
Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala
30          35          40
Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln
45          50          55
Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr
60          65          70          75
Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Glu
80          85          90
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln
95          100         105

```

Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn
 110 115 120
 Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe
 125 130 135
 Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu
 140 145 150 155
 Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe
 160 165 170
 Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val
 175 180 185
 Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp
 190 195 200
 Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser
 205 210 215
 Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Thr
 220 225 230 235
 Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu
 240 245 250
 Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys
 255 260 265
 Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys
 270 275 280
 Leu

序列号: 2

序列的长度: 1257

序列类型: 核酸

链数: 二条链

拓扑型: 直链形

序列种类: 基因组DNA

起源

生物名: *humicola insolens*

序列特征

表示特征的记号: intron

存在位置: 453..509

决定特征的方法: E

序列

```

AATGACGGGG CAACCTCCCG CCCGGGCCCCA ACTCTTGGGT TTGGTTTGAC AGGCCGTCTG      60
TCTCTTGCGT CCTCTTACTA CGCCTGCCTG GACCCTACGT CTCAACTCCG ATTCAAG      117
ATG CGT TCC TCC CCT CTC CTC CGC TCC GCC GTT GTG GCC GCC CTG CCG      165
Met Arg Ser Ser Pro Leu Leu Arg Ser Ala Val Val Ala Ala Leu Pro
      -20              -15              -10
GTG TTG GCC CTT GCC GCT GAT GGC AAG TCC ACC CGC TAC TGG GAC TGC      213
Val Leu Ala Leu Ala Ala Asp Gly Lys Ser Thr Arg Tyr Trp Asp Cys
      -5              1              5              10
TGC AAG CCT TCG TGC GGC TGG GCC AAG AAG GCT CCC GTG AAC CAG CCT      261
Cys Lys Pro Ser Cys Gly Trp Ala Lys Lys Ala Pro Val Asn Gln Pro
              15              20              25

```

GTC TTC TCC TGC AAC GCC AAC TTC CAG CGT CTC ACT GAC TTC GAC GCC	309
Val Phe Ser Cys Asn Ala Asn Phe Gln Arg Leu Thr Asp Phe Asp Ala	
30 35 40	
AAG TCC GGC TGC GAG CCG GGC GGT GTC GCC TAC TCG TGC GCC GAC CAG	357
Lys Ser Gly Cys Glu Pro Gly Gly Val Ala Tyr Ser Cys Ala Asp Gln	
45 50 55	
ACC CCA TGG GCT GTG AAC GAC GAC TTC GCG TTC GGT TTT GCT GCC ACC	405
Thr Pro Trp Ala Val Asn Asp Asp Phe Ala Phe Gly Phe Ala Ala Thr	
60 65 70 75	
TCT ATT GCC GGC AGC AAT GAG GCG GGC TGG TGC TGC GCC TGC TAC GA	452
Ser Ile Ala Gly Ser Asn Glu Ala Gly Trp Cys Cys Ala Cys Tyr Gl	
80 85 90	
GTAAGCTTTG GTCGCGTGTG TAACACTGTG CAGGCATAGC ACTAACCACC TCCCAG G	509
	u
CTC ACC TTC ACA TCC GGT CCT GTT GCT GGC AAG AAG ATG GTC GTC CAG	557
Leu Thr Phe Thr Ser Gly Pro Val Ala Gly Lys Lys Met Val Val Gln	
95 100 105	
TCC ACC AGC ACT GGC GGT GAT CTT GGC AGC AAC CAC TTC GAT CTC AAC	605
Ser Thr Ser Thr Gly Gly Asp Leu Gly Ser Asn His Phe Asp Leu Asn	
110 115 120	
ATC CCC GGC GGC GGC GTC GGC ATC TTC GAC GGA TGC ACT CCC CAG TTC	653
Ile Pro Gly Gly Gly Val Gly Ile Phe Asp Gly Cys Thr Pro Gln Phe	
125 130 135	

GGC GGT CTG CCC GGC CAG CGC TAC GGC GGC ATC TCG TCC CGC AAC GAG	701
Gly Gly Leu Pro Gly Gln Arg Tyr Gly Gly Ile Ser Ser Arg Asn Glu	
140 145 150	
TGC GAT CGG TTC CCC GAC GCC CTC AAG CCC GGC TGC TAC TGG CGC TTC	749
Cys Asp Arg Phe Pro Asp Ala Leu Lys Pro Gly Cys Tyr Trp Arg Phe	
160 165 170	
GAC TGG TTC AAG AAC GCC GAC AAC CCG AGC TTC AGC TTC CGT CAG GTC	797
Asp Trp Phe Lys Asn Ala Asp Asn Pro Ser Phe Ser Phe Arg Gln Val	
175 180 185	
CAA TGC CCA GCC GAG CTC GTC GCT CGC ACC GGA TGC CGC CGC AAC GAC	845
Gln Cys Pro Ala Glu Leu Val Ala Arg Thr Gly Cys Arg Arg Asn Asp	
190 195 200	
GAC GGC AAC TTC CCT GCC GTC CAG ATC CCC TCC AGC AGC ACC AGC TCT	893
Asp Gly Asn Phe Pro Ala Val Gln Ile Pro Ser Ser Ser Thr Ser Ser	
205 210 215	
CCG GTC GGC CAG CCT ACC AGT ACC AGC ACC ACC TCC ACC TCC ACC ACC	941
Pro Val Gly Gln Pro Thr Ser Thr Ser Thr Thr Ser Thr Ser Thr Thr	
220 225 230 235	
TCG AGC CCG CCC GTC CAG CCT ACG ACT CCC AGC GGC TGC ACT GCT GAG	989
Ser Ser Pro Pro Val Gln Pro Thr Thr Pro Ser Gly Cys Thr Ala Glu	
240 245 250	
AGG TGG GCT CAG TGC GGC GGC AAT GGC TGG AGC GGC TGC ACC ACC TGC	1037
Arg Trp Ala Gln Cys Gly Gly Asn Gly Trp Ser Gly Cys Thr Thr Cys	
255 260 265	

GTC GCT GGC AGC ACC TGC ACG AAG ATT AAT GAC TGG TAC CAT CAG TGC	1085
Val Ala Gly Ser Thr Cys Thr Lys Ile Asn Asp Trp Tyr His Gln Cys	
270 275 280	
CTG TAA ACGCAGGGCA GCCTGAGAAC CTTACTGGTT GCGCAACGAA ATGACACTCC	1141
Leu	
CAATCACTGT ATTAGTTCTT GTACATAATT TCGTCATCCC TCCAGGGATT GTCACATATA	1201
TGCAATGATG AATACTGAAC ACAAACCTGG CCGCTTGAAC TGGCCGAAGG AATGCC	1257