(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 111918969 B
(45)授权公告日 2024.11.19

(21)申请号 201980022337.3

(22)申请日 2019.03.29

(65)同一申请的已公布的文献号
申请公布号 CN 111918969 A

(43)申请公布日 2020.11.10

(30)优先权数据
2018-070188 2018.03.30 JP

(85)PCT国际申请进入国家阶段日
2020.09.25

(86)PCT国际申请的申请数据
PCT/JP2019/013926 2019.03.29

(87)PCT国际申请的公布数据
WO2019/189724 JA 2019.10.03

(73)专利权人 株式会社API
地址 日本福冈县

(72)发明人 吉田丰和 石田浩一 三宅良磨 井浦崇敦 川端润

(74)专利代理机构 中原信达知识产权代理有限责任公司 11219
专利代理师 盛曼 金龙河

(51)Int.Cl.
C12N 15/55(2006.01)
C12N 1/15(2006.01)
C12N 1/19(2006.01)
C12N 1/21(2006.01)
C12N 5/10(2006.01)
C12N 9/18(2006.01)
C12P 41/00(2006.01)

(56)对比文件
NCBI.alpha/beta hydrolase [Pseudonocardia dioxanivorans].GenBank.2013,WP_013673631.
NCBI.alpha/beta hydrolase [Microbacterium chocolatum].GenBank.2015,WP_053548180.1.
Savory E.A.等.alpha/beta hydrolase [Rhodococcus sp. 06-621-2].GenBank.2017,OZC62876.1.

审查员 全弘扬

权利要求书1页 说明书19页
序列表17页 附图2页

(54)发明名称

新型水解酶和利用该酶的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的制造方法

(57)摘要

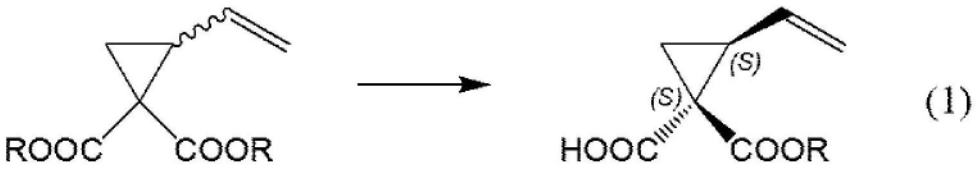
本发明提供能够高效率且低成本地在工业上制造光学纯度高的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的新型水解酶以及使用该酶的制造方法。

CN 111918969 B

1.一种水解酶,其是由序列号2所表示的氨基酸序列构成的多肽。

2.一种核酸,其编码权利要求1所述的水解酶。

3.(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的制造方法,其包括:使权利要求1所述的水解酶、具有生产所述酶的能力的微生物或细胞、所述微生物或细胞的包含所述酶的处理产物、和/或对所述微生物或细胞进行培养而得到的包含所述酶的培养液与式(I)所表示的2-乙烯基环丙烷-1,1-二羧酸二烷基酯接触,从而制造式(II)所表示的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸,



(I) (II) (1)

式(1)中,R为乙基。

4.如权利要求3所述的制造方法,其中,所述微生物或细胞为利用权利要求2所述的核酸进行了转化的微生物或细胞。

5.一种重组载体,其包含权利要求2所述的核酸。

6.一种转化体,其特征在于,包含权利要求5所述的重组载体。

新型水解酶和利用该酶的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的制造方法

技术领域

[0001] 本发明涉及可用于(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的制造的新型水解酶(酯酶)及其应用。

背景技术

[0002] (1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸是在作为丙型肝炎治疗药开发中的各种HCV NS3蛋白酶抑制剂等的制造中有用的中间体。

[0003] 非专利文献1、专利文献1和专利文献2中记载了利用酶将2-乙烯基环丙烷-1,1-二羧酸二甲酯水解而得到(1S,2S)-1-甲氧基羰基-2-乙烯基环丙烷羧酸的方法。

[0004] 但是,非专利文献1中使用的酶的光学选择性不充分,所得到的(1S,2S)-1-甲氧基羰基-2-乙烯基环丙烷羧酸的光学纯度为90%e.e.,并不充分,不适合作为药品中间体的工业制造法。另外,专利文献1和专利文献2中使用的酶的光学选择性不充分,除了作为目标的(1S,2S)-1-甲氧基羰基-2-乙烯基环丙烷羧酸以外,作为副产物的(1S,2R)-1-甲氧基羰基-2-乙烯基环丙烷羧酸的生成量也多,因此难以高效率地制造目标物质,另外,在工业制造的情况下需要大型的分离装置、纯化装置,成本增高。

[0005] 因此,期望一种高纯度、高效率且低成本地在工业上制造作为丙型肝炎治疗药等的制造用中间体有用的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的方法。

[0006] 现有技术文献

[0007] 专利文献

[0008] 专利文献1:日本专利第5613660号

[0009] 专利文献2:日本专利第5657560号

[0010] 非专利文献

[0011] 非专利文献1:C.Fliche等Synth.Commun.24(20),2873-2876(1994)

发明内容

[0012] 发明所要解决的问题

[0013] 本发明要解决的课题在于提供用于高光学纯度、高选择性地制造(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的新型水解酶(酯酶)。此外,本发明要解决的课题还在于提供高效率且低成本地在工业上制造光学纯度高的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的新方法。

[0014] 用于解决问题的方法

[0015] 本发明人为了解决上述课题进行了深入研究,结果发现,来源于红球菌属D32株(Rhodococcus sp.D32株)、嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)和巧克力微杆菌(Microbacterium chocolatum)的酶能够高选择性地水解2-乙烯基环丙烷-1,1-二羧酸二烷基酯,从而高效率地得到(1S,2S)-1-烷氧基

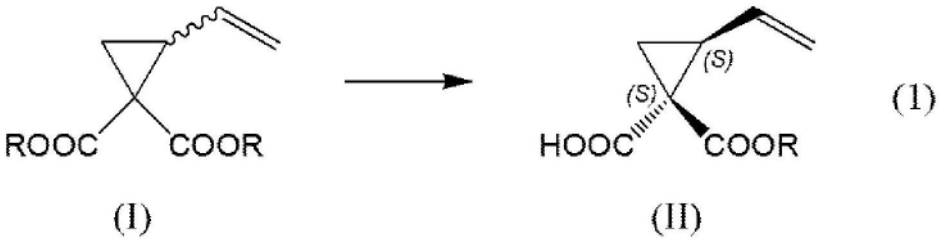
羰基-2-乙烯基环丙烷羧酸。另外发现,通过使该酶、具有生产该酶的能力的微生物或细胞、该微生物或细胞的处理产物、和/或对该微生物或细胞进行培养而得到的包含该酶的培养液(以下有时将这些统括地称为“酶等”)与2-乙烯基环丙烷-1,1-二羧酸二烷基酯接触,能够低成本地得到高光学纯度、高选择性且高浓度的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸。进一步还发现,通过使用这样得到的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸,还能够高效率且低成本地制造光学纯度高的(1R,2S)-1-氨基-1-烷氧基羰基-2-乙烯基环丙烷。本发明是基于这些见解而完成的。

[0016] 即,本发明如下所述。

[0017] [1]一种水解酶,其包含下述(a)~(e)中的任意一种多肽。

[0018] (a)具有序列号2所表示的氨基酸序列的多肽

[0019] (b)具有序列号4或6所表示的氨基酸序列、并且具有催化式(1)所示的反应的活性的多肽

[0020] 

(1)

[0021] (式中,R表示碳原子数1~6的烷基)

[0022] (c)具有在序列号2、4或6所示的氨基酸序列中缺失、插入、置换和/或添加一个或两个以上氨基酸而得到的氨基酸序列、并且具有催化式(1)所示的反应的活性的多肽

[0023] (d)具有与序列号2、4或6所示的氨基酸序列具有60%以上的序列一致性的氨基酸序列、并且具有催化上述式(1)所示的反应的活性的多肽

[0024] (e)具有与序列号2、4或6所示的氨基酸序列具有60%以上的序列一致性的氨基酸序列、并且在上述氨基酸序列中具有选自下述(i)~(iv)的基序序列中的一种以上基序序列的多肽

[0025] (i)由连续的5个残基(天冬氨酸(D)、丙氨酸(A)、苏氨酸(T)、精氨酸(R)、甘氨酸(G))构成的序列

[0026] (ii)由连续的4个残基(脯氨酸(P)、酪氨酸(Y)、甘氨酸(G)、苯丙氨酸(F))构成的序列

[0027] (iii)由连续的4个残基(天冬酰胺(N)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0028] (iv)由连续的5个残基(脯氨酸(P)、甘氨酸(G)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0029] [2]如[1]所述的水解酶,其中,式(1)中的R为乙基。

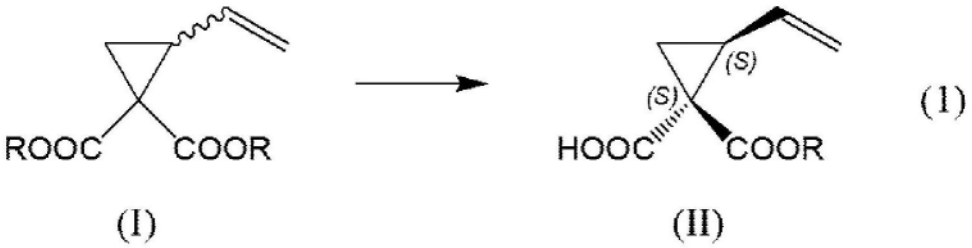
[0030] [3]一种核酸,其编码[1]或[2]所述的水解酶。

[0031] [4]如[3]所述的核酸,其中,上述核酸包含下述(f)、(g)、(h)或(i)所示的碱基序列。

[0032] (f)序列号1、3或5所表示的碱基序列

[0033] (g)序列号1、3或5所表示的碱基序列中具有一个或两个以上碱基的置换、缺失和/

或添加的碱基序列,并且是编码具有催化式(1)所示的反应的活性的多肽的核酸

[0034] 

(I) (II) (1)

[0035] (式中,R表示碳原子数1～6的烷基)

[0036] (h)具有与序列号1、3或5所表示的碱基序列具有60％以上的序列一致性的碱基序列,并且是编码具有催化上述式(1)所示的反应的活性的多肽的核酸

[0037] (i)在严格条件下与序列号1、3或5所表示的碱基序列的互补链杂交的碱基序列,并且是编码具有催化上述式(1)所示的反应的活性的多肽的核酸

[0038] [5](1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的制造方法,其包括:使[1]或[2]所述的水解酶、具有生产上述酶的能力的微生物或细胞、上述微生物或细胞的处理产物、和/或对上述微生物或细胞进行培养而得到的包含上述酶的培养液与式(I)所表示的2-乙烯基环丙烷-1,1-二羧酸二烷基酯接触,从而制造式(II)所表示的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸。

[0039] [6]如[5]所述的制造方法,其中,上述微生物或细胞为利用[3]或[4]所述的核酸进行了转化的微生物或细胞。

[0040] [7]一种重组载体,其包含[3]或[4]所述的核酸。

[0041] [8]一种转化体,其特征在于,包含[7]所述的重组载体。

[0042] 发明效果

[0043] 根据本发明,能够提供用于高光学纯度、高选择性地制造作为丙型肝炎治疗药等的制造用中间体有用的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的新型水解酶(酯酶)。另外,根据本发明,能够提供高效率且低成本地在工业上制造光学纯度高的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的新方法。此外,通过使用这样得到的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸,能够高效率且低成本地制造光学纯度高的(1R,2S)-1-氨基-1-烷氧基羰基-2-乙烯基环丙烷。

[0044] 利用本发明的方法制造的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸和使用其制造制造的(1R,2S)-1-氨基-1-烷氧基羰基-2-乙烯基环丙烷能够用作丙型肝炎的治疗药等的制造中的起始物质或制造用中间体。

附图说明

[0045] 图1是示出实施例1中使用具有氨基酸序列D32Est的菌体而得到的反应产物的HPLC分析结果的图。

[0046] 图2是示出使用具有氨基酸序列PnbA3027-m26的菌体而得到的反应产物的HPLC分析结果的图。

[0047] 图3是示出水解酶的比对图。

具体实施方式

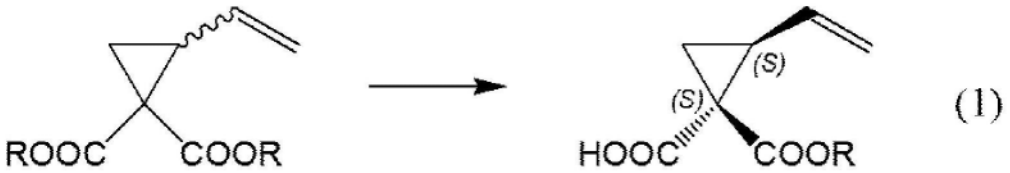
[0048] 以下,详细地对本发明进行说明。

[0049] 1.本发明的水解酶

[0050] 本发明的水解酶是包含具有序列号2、4或6所表示的氨基酸序列的多肽的水解酶、或者是包含具有与该氨基酸序列具有高一致性的氨基酸序列(以下有时称为“氨基酸序列的同源物”)且具有水解2-乙烯基环丙烷-1,1-二羧酸二乙酯的活性的多肽的水解酶(以下有时称为“水解酶的同源物”)。具体而言,本发明的新型水解酶包含下述(a)、(b)、(c)、(d)或(e)所示的多肽。

[0051] (a)具有序列号2所表示的氨基酸序列的多肽

[0052] (b)具有序列号4或6所表示的氨基酸序列、并且具有催化式(1)所示的反应的活性的多肽

[0053] 

(I) (II) (1)

[0054] (式中,R表示碳原子数1～6的烷基)

[0055] (c)具有在序列号2、4或6所表示的氨基酸序列中缺失、插入、置换和/或添加一个或两个以上氨基酸而得到的氨基酸序列、并且具有催化式(1)所示的反应的活性的多肽

[0056] (d)具有与序列号2、4或6所表示的氨基酸序列具有60％以上的序列一致性的氨基酸序列、并且具有催化上述式(1)所示的反应的活性的多肽

[0057] (e)具有与序列号2、4、或6所表示的氨基酸序列具有60％以上的序列一致性的氨基酸序列、并且在该氨基酸序列中具有选自下述(i)～(iv)的基序序列中的一种以上基序序列的多肽:

[0058] (i)由连续的5个残基(天冬氨酸(D)、丙氨酸(A)、苏氨酸(T)、精氨酸(R)、甘氨酸(G))构成的序列

[0059] (ii)由连续的4个残基(脯氨酸(P)、酪氨酸(Y)、甘氨酸(G)、苯丙氨酸(F))构成的序列

[0060] (iii)由连续的4个残基(天冬酰胺(N)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0061] (iv)由连续的5个残基(脯氨酸(P)、甘氨酸(G)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0062] 式(1)中,作为R所表示的碳原子数1～6的烷基,可以列举碳原子数1～6的直链状或支链状的烷基,更优选可以列举碳原子数1～3的直链状或支链状的烷基。作为具体例,可以列举甲基、乙基、异丙基、正丙基、正丁基、异丁基、仲丁基、叔丁基、正戊基、异戊基、叔戊基、新戊基、正己基等,其中优选甲基、乙基,特别优选乙基。需要说明的是,式(I)中的两个R可以相同也可以不同,优选相同。

[0063] 本发明中,具有序列号2、4或6所表示的氨基酸序列的水解酶的同源物包含上述

(c)、(d)或(e)所示的多肽。

[0064] 上述(c)所示的多肽是具有在序列号2、4或6所表示的氨基酸序列中缺失、插入、置换和/或添加一个或两个以上氨基酸而得到的氨基酸序列、并且具有催化式(1)所示的反应的活性的多肽。置换的情况下,优选一个或两个以上氨基酸进行保守性置换。本说明书中,“氨基酸进行保守性置换”是指化学性质等相似的氨基酸彼此的置换,可以列举例如用碱性氨基酸置换碱性氨基酸、用酸性氨基酸置换酸性氨基酸。

[0065] “一个或两个以上氨基酸”通常为1个～100个、优选为1个～50个、更优选为1个～20个、进一步优选为1个～10个、特别是优选为1个～5个氨基酸。

[0066] 上述(d)所示的多肽是具有与序列号2、4或6所表示的氨基酸序列具有60％以上的序列一致性的氨基酸序列、并且具有催化上述式(1)所示的反应的活性的多肽。优选为具有与序列号2、4或6所表示的氨基酸序列全长具有80％以上、更优选90％以上、进一步优选95％以上的序列一致性的氨基酸序列、并且具有催化上述式(1)所示的反应的活性的多肽。

[0067] 本说明书中的氨基酸序列的同源性(也称为一致性或相似性)可以使用同源性计算算法NCBI BLAST(National Center for Biotechnology Information Basic Local Alignment Search Tool,美国国家生物技术信息中心基本局部比对搜索工具),利用例如下述条件(期望值＝10;允许间隙;矩阵＝BLOSUM62;过滤＝OFF)来计算。作为用于确定氨基酸序列的同源性的其他算法,可以列举例如Karlin等,Proc.Natl.Acad.Sci.USA,90:5873-5877(1993)中记载的算法[该算法编入在NBLAST和XBLAST程序(2.0版)中(Altschul等,Nucleic Acids Res.,25:3389-3402(1997))]、Needleman等,J.Mol.Biol.,48:444-453(1970)中记载的算法[该算法编入在GCG软件包中的GAP程序中]、Myers和Miller,CABIOS,4:11-17(1988)中记载的算法[该算法编入在作为CGC序列比对软件包的一部分的ALIGN程序(2.0版)中]、Pearson等,Proc.Natl.Acad.Sci.USA,85:2444-2448(1988)中记载的算法[该算法编入在GCG软件包中的FASTA程序中]等,这些算法也同样可以优选使用。

[0068] 本发明中,催化式(1)所示的反应的活性是指催化将式(I)所表示的2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解而生成式(II)所表示的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的反应的活性。

[0069] 在此,2-乙烯基环丙烷-1,1-二羧酸二烷基酯存在(2R)体和(2S)体两种立体异构体,只要不是在发生2位上的立体翻转的特殊条件下,基本上只有(2S)体能够作为目标(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的原料。

[0070] 因此,催化式(1)所示的反应的活性是指催化具有下述选择性的反应的活性:在(2S)-2-乙烯基环丙烷-1,1-二羧酸二烷基酯的1位的前手性碳上键合的两个烷氧羰基中,仅优先水解pro-R的烷氧羰基,优先生成(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸。

[0071] 需要说明的是,pro-R是对于CX_2YZ上存在的2个X将它们区分开的表示法,对于C上的各取代基,按照CIP规则来决定优先顺序。此时,假设2个X中一者的优先顺序高于另一者,它们与Y、Z的优先顺序的关系不变。基于假设的优先顺序,通过RS表示法来确定中心碳的手性为R或为S。R体的情况下,将此时优先的X作为pro-R,S体的情况下作为pro-S。

[0072] 该选择性可以以通过水解而生成的1-烷氧基羰基-2-乙烯基环丙烷羧酸的(1S,2S)体与(1R,2S)体的比率作为指标来确认。与(1S,2S)体的生成量相比(1R,2S)体的生成量越低,(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的收率越提高,在工业化方面越有利。

[0073] 另外,(1S,2S)体与(1R,2S)体的生成比率可以以生成的Anti体与Syn体的比率作为指标来进行比较。Anti体、Syn体是指几何异构体,(1S,2S)体和(1R,2R)体为Anti体,(1R,2S)体和(1S,2R)体为Syn体。本发明中,水解得到的产物的主要成分为(1S,2S)体,因此,在对产物进行定量时,不仅可以使用单独利用(1S,2S)体来定量的评价系统,还可以使用能够将Anti体与Syn体分离的评价系统来对生成量进行定量。即,在对(1S,2S)体进行定量的情况下,可以通过对Anti体的生成量进行定量来代替。同样地,在对(1R,2S)体进行定量的情况下,可以通过对Syn体的生成量进行定量来代替。Syn体的生成比率越低,(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的收率越提高,在工业化方面越有利。

[0074] 这些选择性可以通过下述方法来确认:使作为对象的水解酶等与外消旋体的2-乙烯基环丙烷-1,1-二羧酸二烷基酯接触而生成1-烷氧基羰基-2-乙烯基环丙烷羧酸,测定所生成的1-烷氧基羰基-2-乙烯基环丙烷羧酸的(1S,2S)体和(1R,2S)体的生成量或者测定Anti体和Syn体的生成量。

[0075] 接触方法没有特别限定,可以列举例如在包含作为对象的上述水解酶的液体中添加外消旋体的2-乙烯基环丙烷-1,1-二羧酸二烷基酯,在适当的温度(例如约10℃～约45℃)、压力(例如约为大气压)下使其反应;等等。

[0076] 另外,将外消旋体的2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解时,(2R)体水解时可能生成(1S,2R)体和(1R,2R)体,优选将相当于作为目标的(1S,2S)体的对映异构体的(1R,2R)体的生成量抑制得较低。因此,本发明中,进一步优选具有下述选择性:在(2R)-2-乙烯基环丙烷-1,1-二羧酸二烷基酯的1位的前手性的碳上键合的两个烷氧羰基中,仅优先水解pro-R的烷氧羰基,优先生成(1S,2R)-1-烷氧基羰基-2-乙烯基环丙烷羧酸。同样地,进一步优选具有下述选择性:与(2R)-2-乙烯基环丙烷-1,1-二羧酸二烷基酯相比,优先水解(2S)-2-乙烯基环丙烷-1,1-二羧酸二烷基酯。任意一种选择性均可以通过由水解生成的1-烷氧基羰基-2-乙烯基环丙烷羧酸的(1S,2S)体相对于(1R,2R)体的对映异构体过剩率(%e.e.)来测定。(1S,2S)体的对映异构体过剩率越高,对之后的制造工序和所制造的药品的生理活性产生不良影响的可能性越低,在工业上越有利。

[0077] 上述(e)所示的多肽是具有与序列号2、4或6所表示的氨基酸序列具有60%以上的序列一致性的氨基酸序列、并且在该氨基酸序列中具有选自下述(i)～(iv)的基序序列中的一种以上基序序列的多肽。

[0078] (i)由连续的5个残基(天冬氨酸(D)、丙氨酸(A)、苏氨酸(T)、精氨酸(R)、甘氨酸(G))构成的序列

[0079] (ii)由连续的4个残基(脯氨酸(P)、酪氨酸(Y)、甘氨酸(G)、苯丙氨酸(F))构成的序列

[0080] (iii)由连续的4个残基(天冬酰胺(N)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列、和

[0081] (iv)由连续的5个残基(脯氨酸(P)、甘氨酸(G)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0082] 本发明中,基序序列是指在多肽的氨基酸序列中确认到的小的结构部分。已知在保持特定功能的一类多肽(蛋白质)中,共同地存在特征性的基序序列。例如已知,在特定种类的水解酶中,丝氨酸(S)、天冬氨酸(D)和组氨酸(H)是催化活性(水解功能)所需要的氨基

酸残基。具体而言,Catalysis Science and Technology,2015,vol.5,2622中所示的水解酶RhEst1具有丝氨酸(S)、天冬氨酸(D)和组氨酸(H)作为氨基酸残基。另外,如图3所示,美国专利第8298799号中记载的水解酶AFX20780也在特定位置具有丝氨酸(S)、天冬氨酸(D)和组氨酸(H)作为氨基酸残基。

[0083] 如图3所示可知,本发明的序列号2、4或6所表示的氨基酸序列也与上述水解酶RhEst1和AFX20780同样地在特定位置具有丝氨酸(S)、天冬氨酸(D)和组氨酸(H),具有催化活性(水解功能)。

[0084] 另外,如图3所示,本发明的序列号2、4或6所表示的氨基酸序列具有选自下述(i)～(iv)的基序序列中的一种以上基序序列作为特征性的基序序列。

[0085] (i)由连续的5个残基(天冬氨酸(D)、丙氨酸(A)、苏氨酸(T)、精氨酸(R)、甘氨酸(G))构成的序列

[0086] (ii)由连续的4个残基(脯氨酸(P)、酪氨酸(Y)、甘氨酸(G)、苯丙氨酸(F))构成的序列

[0087] (iii)由连续的4个残基(天冬酰胺(N)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0088] (iv)由连续的5个残基(脯氨酸(P)、甘氨酸(G)、色氨酸(W)、脯氨酸(P)、甘氨酸(G))构成的序列

[0089] 这些基序序列是除了上述的丝氨酸(S)、天冬氨酸(D)和组氨酸(H)以外在本发明的序列号2、4或6所表示的氨基酸序列中共同具有的,因此推测,这些基序序列是具有该氨基酸序列的本发明的酶所具有的特征性功能、即立体选择性地对2-乙烯基环丙烷-1,1-二羧酸二烷基酯进行水解所需要的基序序列。这些基序序列可能单独发挥功能,也可能通过多个基序序列同时起作用而发挥所期望的功能。

[0090] 如上所述,本发明的水解酶为包含具有序列号2、4或6所表示的氨基酸序列的多肽的水解酶、或者为包含作为该氨基酸序列的同源物的具有2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解活性的多肽的水解酶。包含具有该氨基酸序列的同源物的多肽的水解酶的2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解活性的程度在定量上可以与包含具有序列号2、4或6所表示的氨基酸序列的多肽的水解酶是同等的,也可以在可允许的范围(例如约0.1倍～约5倍、优选约0.3倍～约3倍)内不同。

[0091] 序列号2、4和6所表示的氨基酸序列分别来源于红球菌属D32株(Rhodococcus spD32株)、嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)和巧克力微杆菌(Microbacterium chocolatum)。经本发明人鉴定,这些氨基酸序列为2-乙烯基环丙烷-1,1-二羧酸二烷基酯的水解酶。

[0092] 具有序列号2所表示的氨基酸序列的多肽为来源于Rhodococcus sp D32株的多肽,该菌株是从土壤样品中回收的菌。通过该菌生产的2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解酶的N末端序列的分析和Rhodococcus spD32株的染色体DNA的分析,得到了序列号2所表示的氨基酸序列。

[0093] 具有序列号4和6所表示的氨基酸序列的多肽是由各自所来源的微生物的基因组信息得到的序列。序列号4和6所表示的氨基酸序列分别与预测为编码蛋白质的DNA序列所编码的氨基酸序列GenBank登录号WP-013673631和WP-053548180相同。均没有作为蛋白质

被分离等确认到实际存在的报告例,关于作为蛋白质的功能当然也完全不清楚。

[0094] 本发明中,在具有序列号2、4和6所表示的氨基酸序列的多肽中,出于(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的收率高、选择性也高的考虑,优选具有序列号2所表示的氨基酸序列的多肽。

[0095] 本发明的包含具有序列号2、4或6所表示的氨基酸序列的多肽的水解酶也可以利用公知的方法分别从红球菌属D32株(Rhodococcus sp.D32株)、嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)和巧克力微杆菌(Microbacterium chocolatum)纯化分离而得到。在此,作为纯化方法,可以列举例如溶剂提取、蒸馏、柱色谱法、液相色谱法、重结晶、它们的组合等。

[0096] 另外,本发明的水解酶也可以通过对含有编码该水解酶的核酸的转化体进行培养,从所得到的培养物中分离纯化该水解酶而制造。编码本发明的水解酶的核酸可以为DNA,也可以为RNA,或者也可以为DNA/RNA嵌合体。优选可以列举DNA。另外,该核酸可以为双链,也可以为单链。双链的情况下,可以为双链DNA、双链RNA或DNA:RNA的杂交体。单链的情况下,可以为有义链(即编码链),也可以为反义链(即非编码链)。

[0097] 作为编码本发明的水解酶的DNA,可以列举合成DNA等。例如,可以通过下述方法来取得:使用由来源于红球菌属D32株(Rhodococcus sp.D32株)、嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)或巧克力微杆菌(Microbacterium chocolatum)的细胞或组织制备的总RNA或mRNA级分作为模板,通过逆转录酶PCR直接进行扩增,对于所得到的全长水解酶cDNA,使用公知的试剂盒、例如MutanTM-super Express Km(TAKARA BIO INC.)、MutanTM-K(TAKARA BIO INC.)等,根据ODA-LA PCR法、缺口双链法、Kunkel法等自身公知的方法或者基于这些方法的方法进行转换。或者,也可以通过将上述的总RNA或mRNA的片段插入适当的载体中,从所制备的cDNA文库中通过集落杂交法或噬菌斑杂交法或PCR法等克隆出cDNA,根据上述方法对克隆出的cDNA进行转换来取得。文库中使用的载体可以为噬菌体、质粒、粘粒、噬粒等中的任意一种。

[0098] 作为编码具有序列号2、4或6所表示的氨基酸序列的多肽的核酸,可以列举分别包含序列号1、3或5所表示的碱基序列的核酸。只要编码具有催化式(1)所示的反应的活性的多肽,也可以是包含与序列号1、3或5的碱基序列具有高一致性的碱基序列的核酸(以下有时称为“核酸的同源物”)。即,作为编码该多肽的核酸,可以列举具有下述(f)、(g)、(h)或(i)所示的碱基序列的核酸。

[0099] (f)具有序列号1、3或5所表示的碱基序列的核酸

[0100] (g)具有在序列号1、3或5所表示的碱基序列中置换、缺失和/或添加一个或两个以上碱基而得到的碱基序列,并且编码具有催化上述式(1)所示的反应的活性的多肽的核酸

[0101] (h)具有与序列号1、3或5所表示的碱基序列具有60%以上的序列一致性的碱基序列,并且编码具有催化上述式(1)所示的反应的活性的多肽的核酸

[0102] (i)具有在严格条件下与序列号1、3或5所表示的碱基序列的互补链杂交的碱基序列,并且编码具有催化上述式(1)所示的反应的活性的多肽的核酸

[0103] 作为上述(g)所示的核酸的同源物,可以列举包含在序列号1、3或5所表示的碱基序列中缺失、置换、插入和/或添加一个或两个以上碱基而得到的碱基序列、并且编码具有催化上述式(1)所示的反应的活性的多肽的核酸。置换、插入或添加的情况下,优选置换、插

入或添加一个或两个以上碱基。在此,“一个或两个以上碱基”例如为1个～300个、优选为1个～150个、更优选为1个～60个、进一步优选为1个～30个、特别优选为1个～15个碱基。

[0104] 需要说明的是,序列号1、3或5所表示的碱基序列是分别将来源于红球菌属(Rhodococcus sp.)D32株的水解酶RsD32Est(序列号2)、来源于嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)的水解酶PdEst(序列号4)和来源于巧克力微杆菌(Microbacterium chocolatum)的水解酶McEst(序列号6)的基因按大肠杆菌表达用途对密码子进行了优化的碱基序列。这样根据作为转化对象的宿主对密码子进行了优化的DNA当然也包含在本发明中可使用的编码具有催化上述式(1)所示的反应的活性的多肽的核酸中。

[0105] 作为上述(h)所示的核酸的同源物,可以列举具有与序列号1、3或5所表示的碱基序列具有60％以上的序列一致性的碱基序列、并且编码具有催化上述式(1)所示的反应的活性的多肽的核酸。优选为具有与序列号1、3或5所表示的碱基序列具有80％以上、更优选90％以上、进一步优选95％以上、更进一步优选98％以上的同源性(也称为一致性)的碱基序列、并且编码具有催化上述式(1)所示的反应的活性的多肽的核酸。

[0106] 本说明书中的碱基序列的同源性(也称为一致性)可以使用同源性计算算法NCBI BLAST(National Center for Biotechnology Information Basic Local Alignment Search Tool,美国国家生物技术信息中心基本局部比对搜索工具),利用例如下述条件(期望值＝10;允许间隙;过滤＝ON;匹配得分＝1;错配得分＝-3)来计算。作为用于确定碱基序列的同源性的其他算法,可以同样地优选例示上述的氨基酸序列的同源性计算算法。

[0107] 作为上述(h)所示的核酸的同源物,只要编码具有催化上述式(1)所示的反应的活性的多肽,也可以是在严格条件下与序列号1、3或5的碱基序列的互补链杂交的核酸。在此,作为“严格条件”,可以参考已经报道的条件(例如Current Protocols in Molecular Biology,John Wiley&Sons,6.3.16.3.6,1999)来适当设定,具体而言,可以列举例如在相当于通常的Southern杂交的清洗条件、即60℃、1×SSC、0.1％SDS、优选0.1×SSC、0.1％SDS、进一步优选65℃、0.1×SSC、0.1％SDS或68℃、0.1×SSC、0.1％SDS等(高严格条件)的盐浓度和温度下清洗1次、更优选2～3次的条件等。

[0108] 本领域技术人员可以通过使用定点诱变法(Nucleic Acids Res.10,pp.6487(1982)、Methods in Enzymol.100,pp.448(1983)、Molecular Cloning、PCR A Practical Approach IRL Press pp.200(1991))等对序列号1、3或5所表示的核酸适当进行置换、缺失、插入和/或添加,导入所期望的突变而得到如上所述的核酸的同源物。

[0109] 本发明的核酸能够编码具有催化上述式(1)所示的反应的活性的多肽。本发明的核酸具有序列号1、3或5所示的碱基序列、或者与序列号1、3或5所示的碱基序列具有高一致性的碱基序列的情况下,包含由该核酸编码的多肽的水解酶的2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解活性的程度在定量上可以与包含具有序列号2、4或6所表示的氨基酸序列的多肽的水解酶、或者包含具有该氨基酸序列的同源物的多肽的水解酶是同等的,也可以在可允许的范围(例如约0.1倍～约5倍、优选约0.3倍～约3倍)内不同。

[0110] 另外,也可以基于序列号2、4或6的氨基酸序列或其一部分、或序列号1、3或5所表示的碱基配或其一部分,对于例如日本DNA数据库(DNA Databank of JAPAN;DDBJ)等数据库进行同源性检索,获得具有催化式(1)所示的反应的活性的多肽的氨基酸序列信息或编

码该多肽的DNA的碱基序列信息。

[0111] 本发明的水解酶具有以高于现有公知的包含序列号24所示的多肽的水解酶的选择性来催化上述式(1)所示的反应的活性。另外,本发明的水解酶在将2-乙烯基环丙烷-1,1-二羧酸二烷基酯水解时,1-烷氧基羰基-2-乙烯基环丙烷羧酸的(1S,2R)体、(1R,2S)体和/或(1R,2R)体的生成比率低,特别是与包含序列号24所示的多肽的水解酶相比,该(1S,2R)体、(1R,2S)体和/或(1R,2R)体的生成比率低。

[0112] 在后述的本发明的制造方法中,可以直接将上述水解酶用于与2-乙烯基环丙烷-1,1-二羧酸二烷基酯反应,但优选使用具有生产该酶的能力的微生物或细胞、该微生物或细胞的处理产物、和/或对该微生物或细胞进行培养而得到的包含该酶的培养液。

[0113] 作为具有生产本发明的水解酶的微生物或细胞,可以使用原本就具有生产该水解酶的能力的微生物或细胞,也可以是通过育种而赋予了上述生产能力的微生物或细胞。作为微生物或细胞,无论其死活均可,例如可以适当使用休眠菌体等。作为具有生产本发明的水解酶的微生物或细胞的种类,可以列举下文中作为“宿主微生物”或“宿主细胞”所记述的微生物或细胞。

[0114] 作为通过育种而赋予上述生产能力的方法,可以采用基因重组处理(转化)、突变处理等公知的方法。作为转化的方法,可以列举导入目标DNA的方法、在染色体上对启动子等表达调控序列进行改造而强化目标DNA的表达的方法等。

[0115] 这些之中,优选使用利用上述的编码本发明的多肽的DNA进行了转化的微生物或细胞。

[0116] 编码本发明的多肽(水解酶)的核酸(DNA)可以如上所述通过使用来源于红球菌属D32株(Rhodococcus sp.D32株)、嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)或巧克力微杆菌(Microbacterium chocolatum)的染色体DNA作为模板,使用适当的引物进行PCR而进行克隆。

[0117] 另外,编码本发明的多肽(水解酶)的核酸(DNA)可以如上所述通过使用来源于红球菌属D32株(Rhodococcus sp.D32株)、嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)或巧克力微杆菌(Microbacterium chocolatum)的总RNA或mRNA作为模板,利用RT-PCR法直接扩增而制备出全长水解酶cDNA后,使用适当的引物进行PCR而进行克隆。

[0118] 例如,通过将如上得到的编码本发明的多肽的DNA以能够进行表达的配置插入到公知的表达载体中,可提供本发明的多肽基因表达载体。并且,通过利用该表达载体对宿主细胞进行转化,能够得到导入有编码本发明的多肽的DNA的转化体。转化体也可以通过利用同源重组等方法将编码本发明的多肽的DNA以能够进行表达的方式整合到宿主的染色体DNA中而得到。

[0119] 本说明书中,“表达载体”是用于通过插入编码具有所期望的功能的蛋白质的多核苷酸并导入到宿主生物中而使具有所期望的功能的蛋白质在上述宿主生物中复制和表达的遗传因子。可以列举例如质粒、病毒、噬菌体、粘粒等,但不限定于这些。优选表达载体为质粒。

[0120] 本说明书中,“转化体”是指使用上述表达载体等导入了目的基因、能够表现出与具有所期望的功能的蛋白质相关的所期望的性状的微生物或细胞。

[0121] 作为转化体的制作方法,具体而言,可以例示向在宿主细胞中稳定存在的质粒载体、噬菌体载体或病毒载体中导入编码本发明的多肽的DNA,将所构建的表达载体导入该宿主细胞中的方法;或直接将该DNA导入宿主基因组中,对其遗传信息进行转录、翻译的方法。这种情况下,优选在宿主中在DNA的5'-侧上游连接适当的启动子,进一步更优选在3'-侧下游连接终止子。作为这样的启动子和终止子,只要是已知在用作宿主的细胞中发挥功能的启动子和终止子则没有特别限定,可以使用例如"微生物学基础讲座8基因工程·共立出版"中详述的载体、启动子和终止子。

[0122] 作为用于表达本发明的水解酶的作为转化对象的宿主微生物,只要宿主自身不会对2-乙烯基环丙烷-1,1-二羧酸二烷基酯、(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸产生不良影响则没有特别限定,可以列举例如以下所示的微生物。

[0123] 属于埃希氏菌(Escherichia)属、芽孢杆菌(Bacillus)属、假单胞菌(Pseudomonas)属、沙雷氏菌(Serratia)属、短杆菌(Brevibacterium)属、棒杆菌(Corynebacterium)属、链球菌(Streptococcus)属、乳杆菌(Lactobacillus)属等的建立了宿主载体系统的细菌。

[0124] 属于红球菌(Rhodococcus)属、链霉菌(Streptomyces)属等的建立了宿主载体系统的放线菌。

[0125] 属于酵母(Saccharomyces)属、克鲁维酵母(Kluyveromyces)属、裂殖酵母(Schizosaccharomyces)属、接合酵母(Zygosaccharomyces)属、耶氏酵母(Yarrowia)属、丝孢酵母(Trichosporon)属、红冬孢酵母(Rhodosporidium)属、汉逊酵母(Hansenula)属、毕赤酵母(Pichia)属、假丝酵母(Candida)属等的建立了宿主载体系统的酵母。

[0126] 属于脉孢霉(Neurospora)属、曲霉(Aspergillus)属、头孢霉(Cephalosporium)属、木霉(Trichoderma)属等的建立了宿主载体系统的霉菌。

[0127] 用于转化体制作的程序、适合于宿主的重组载体的构建和宿主的培养方法可以依据分子生物学、生物工程、基因工程的领域中惯用的技术来进行(例如Molecular Cloning中记载的方法)。

[0128] 以下,具体地列举优选的宿主微生物、各微生物中的优选的转化方法、载体、启动子、终止子等的示例,但本发明不限定于这些示例。

[0129] 对于埃希氏菌属、特别是大肠杆菌(Escherichia coli),作为质粒载体,可以列举pBR、pUC系质粒等,可以列举来源于lac(β-半乳糖苷酶)、trp(色氨酸操纵子)、tac、trc(lac、trp的融合)、λ噬菌体PL、PR等的启动子等。另外,作为终止子,可以列举来源于trpA、来源于噬菌体、来源于rrnB核糖体RNA的终止子等。

[0130] 对于芽孢杆菌属,作为载体,可以列举pUB110系质粒、pC194系质粒等,另外,也可以整合到染色体中。作为启动子和终止子,可以利用碱性蛋白酶、中性蛋白酶、α-淀粉酶等的酶基因的启动子、终止子等。

[0131] 对于假单胞菌属,作为载体,可以列举在恶臭假单胞菌(Pseudomonas putida)、洋葱假单胞菌(Pseudomonas cepacia)等中建立的一般的宿主载体系统、或以参与甲苯化合物的分解的质粒、TOL质粒为基础的广宿主载体(包含来源于RSF1010等的自主复制所需要的基因)pKT240(Gene,26,273-82(1983))等。

[0132] 对于短杆菌属、特别是乳糖发酵短杆菌(Brevibacterium lactofermentum),作为

载体,可以列举pAJ43(Gene 39,281(1985))等质粒载体。作为启动子和终止子,可以利用在大肠杆菌中使用的各种启动子和终止子。

[0133] 对于棒杆菌属、特别是谷氨酸棒杆菌(Corynebacterium glutamicum),作为载体,可以列举pCS11(日本特开昭57-183799号公报)、pCB101(Mol.Gen.Genet.196,175(1984))等质粒载体。

[0134] 对于酵母(Saccharomyces)属、特别是酿酒酵母(Saccharomyces cerevisiae),作为载体,可以列举YRp系、YEp系、YCp系、YIp系质粒等。另外,可以利用醇脱氢酶、3-磷酸甘油醛脱氢酶、酸性磷酸酶、β-半乳糖苷酶、磷酸甘油酸激酶、烯醇化酶等各种酶基因的启动子、终止子。

[0135] 对于裂殖酵母(Schizosaccharomyces)属,作为载体,可以列举Mol.Cell.Biol.6,80(1986)中记载的来源于粟酒裂殖酵母的质粒载体等。特别是pAUR224由宝生物株式会社销售,可以容易地利用。

[0136] 对于曲霉(Aspergillus)属,黑曲霉(Aspergillus niger)、米曲霉(Aspergillus oryzae)等在霉菌中研究得最多,可以利用质粒或向染色体中的整合,可以利用来源于菌体外蛋白酶或淀粉酶的启动子(Trendsin Biotechnology 7,283-287(1989))。

[0137] 另外,除了上述以外,还建立了与各种微生物相应的宿主载体系统,可以适当使用这些宿主载体系统。

[0138] 另外,除了微生物以外,还在植物、动物中建立了各种各样的宿主/载体系统,特别是建立了在昆虫(例如蚕)等动物中(Nature315,592-594(1985))、菜籽、玉米、马铃薯等植物中大量表达不同种类的蛋白质的系统、以及使用大肠杆菌无细胞提取液或小麦胚芽等的无细胞蛋白质合成系统的系统,可以适当地利用。

[0139] 作为具有生产本发明的水解酶的能力的微生物或细胞的处理产物,可以列举例如将该微生物或细胞用丙酮、二甲基亚砜(DMSO)、甲苯等有机溶剂或表面活性剂进行处理而得到的处理产物、进行冷冻干燥处理而得到的处理产物、以物理方式或利用酶破碎得到的处理产物等细胞制备物、将微生物或细胞中的酶级分以粗制物或纯化物的方式取出而得到的处理产物、进一步将它们固定于以聚丙烯酰胺凝胶、卡拉胶等为代表的载体上而得到的处理产物等。

[0140] 作为对具有生产本发明的水解酶的能力的微生物或细胞进行培养而得到的包含该酶的培养液,可以列举例如该细胞和液体培养基的悬浮液、在该细胞为分泌表达型细胞的情况下通过离心分离等除去该细胞而得到的上清或其浓缩物。

[0141] 2.本发明的组合物

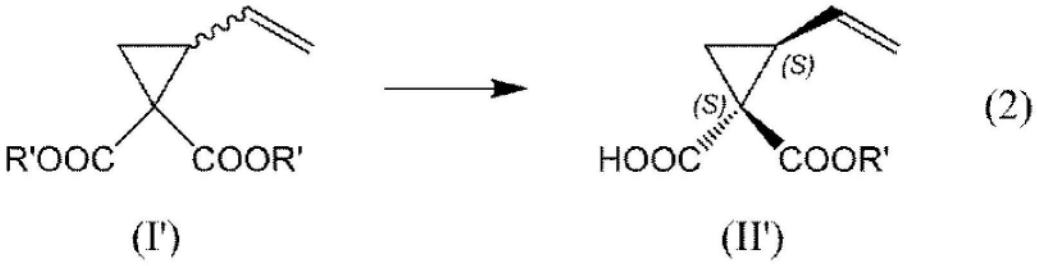
[0142] 本发明的组合物(酶剂)包含本发明的水解酶、具有生产该酶的能力的微生物或细胞、该微生物或细胞的处理产物、和/或对该微生物或细胞进行培养而得到的包含上述酶的培养液,以2-乙烯基环丙烷-1,1-二羧酸二烷基酯作为底物,催化生成(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的反应。通过使用本发明的组合物作为催化剂,能够高效率且低成本地在工业上制造光学纯度高的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸,因此有用。

[0143] 本发明的组合物中,除了有效成分(酶等)以外,还可以含有赋形剂、缓冲剂、助悬剂、稳定剂、保存剂、防腐剂、生理盐水等。作为赋形剂,可以使用乳糖、山梨醇、D-甘露醇、白糖等。作为缓冲剂,可以使用磷酸盐、柠檬酸盐、乙酸盐等。作为稳定剂,可以使用丙二醇、抗

坏血酸等。作为保存剂,可以使用苯酚、苯扎氯铵、苯甲醇、氯丁醇、对羟基苯甲酸甲酯等。作为防腐剂,可以使用苯扎氯铵、对羟基苯甲酸、氯丁醇等。

[0144] 3.使用本发明的水解酶的下式(II')所表示的羧酸单酯的制造

[0145] 根据本发明,可以提供下式(II')所表示的羧酸单酯的制造方法,该制造方法包括式(2)所表示的制造,即,使本发明的水解酶作用于下式(I')所表示的二羧酸酯,制造下式(II')所表示的羧酸单酯。

[0146] 

(I') (II') (2)

[0147] (式中,R'表示可以被取代的碳原子数1~10的烷基、碳原子数7~20的芳烷基、或碳原子数6~12的芳基)

[0148] 式(2)中,作为R'所表示的"可以被取代的碳原子数1~10的烷基"中的"1~10的烷基",可以列举碳原子数1~10的直链状或支链状或环状的烷基,更优选可以列举碳原子数1~6的直链状或支链状或环状的烷基。这些之中,优选可以列举例如甲基、乙基、异丙基、正丙基、正丁基、异丁基、仲丁基、叔丁基、正戊基、异戊基、叔戊基、新戊基、正己基、环丙基、环丁基、环戊基、环己基等,更优选碳原子数1~4的烷基,进一步优选直链状或支链状的碳原子数1~4的烷基,特别优选乙基。作为"可以被取代的碳原子数1~10的烷基"中的取代基,可以列举卤原子(例如氯原子、溴原子)、碳原子数1~6的直链状或支链的烷氧基(例如甲氧基、乙氧基)、硝基等。

[0149] 作为R'所表示的"可以被取代的碳原子数7~20的芳烷基"中的"碳原子数7~20的芳烷基",优选为碳原子数7~12的芳烷基。这些之中,可以列举苄基、2-苯基乙基、1-苯基乙基、3-苯基丙基等。作为"可以被取代的碳原子数7~20的芳烷基"中的取代基,可以列举卤原子(例如氯原子、溴原子)、碳原子数1~6的直链状或支链的烷基(例如甲基、乙基)、碳原子数1~6的直链状或支链的烷氧基(例如甲氧基、乙氧基)等。作为"可以被取代的碳原子数7~20的芳烷基",优选可以列举苄基、4-甲基苄基、4-甲氧基苄基、4-氯苄基、4-溴苄基、2-苯基乙基、1-苯基乙基、3-苯基丙基等。

[0150] 作为R'所表示的"可以被取代的碳原子数6~12的芳基"中的"碳原子数6~12的芳基",可以列举例如苯基、1-萘基、2-萘基等。作为"可以被取代的碳原子数6~12的芳基"中的取代基,可以列举卤原子(例如氯原子、溴原子)、碳原子数1~6的直链状或支链的烷基(例如甲基、乙基)、碳原子数1~6的直链状或支链的烷氧基(例如甲氧基、乙氧基)、硝基等。作为"可以被取代的碳原子数6~12的芳基",优选可以列举例如苯基、1-萘基、2-萘基、邻甲苯基、间甲苯基、对甲苯基、邻甲氧基苯基、间甲氧基苯基、对甲氧基苯基、2,3-二甲基苯基、2,4-二甲基苯基、2,5-二甲基苯基、2,6-二甲基苯基、3,4-二甲基苯基、3,5-二甲基苯基、2,3,5-三甲基苯基、2,3,6-三甲基苯基、2,4,6-三甲基苯基、邻硝基苯基、间硝基苯基、对硝基苯基等。

[0151] 作为R',特别优选为甲基、乙基、叔丁基或苄基,更优选为乙基。

[0152] 需要说明的是,式(I')中的两个R'可以相同也可以不同,优选相同。

[0153] 使本发明的水解酶与式(I')所表示的二羧酸二酯接触时,可以通过使纯化或粗纯化的本发明的水解酶、具有生产本发明的水解酶的能力的微生物或细胞(例如具有编码本发明的多肽的DNA的转化体等)、该微生物或细胞的处理产物、和/或对该微生物或细胞进行培养而得到的包含该酶的培养液与式(I')所表示的二羧酸二酯接触,从而制造式(II')所表示的羧酸单酯。

[0154] 本发明的水解酶可以直接用于反应,但优选使用具有生产该酶的能力的微生物或细胞、该微生物或细胞的处理产物、和/或对该微生物或细胞进行培养而得到的包含该酶的培养液,这些之中,优选使用具有编码本发明的多肽的DNA的转化体。

[0155] 关于添加到反应液中的微生物或细胞、该微生物或细胞的处理产物、和/或对该微生物或细胞进行培养而得到的包含该酶的培养液的量,在添加微生物或细胞的情况下,按照该微生物或细胞的浓度通常以湿菌体重计为约0.1w/v%~约50w/v%、优选为1w/v%~20w/v%的方式添加到反应液中,在使用处理产物或培养液的情况下,求出酶的比活性,添加时添加达到上述细胞浓度的量。在此,w/v%表示重量(weight)/体积(volume)%。

[0156] 作为本发明的水解酶的反应底物的式(I')所表示的二羧酸二酯通常使用(2S)体或外消旋体。

[0157] 反应的方法没有特别限定,可以向包含本发明的水解酶的液体中添加作为底物的式(I')所表示的二羧酸二酯,在适当的温度和压力(例如约大气压)下使其反应。由此,能够制造式(II')所表示的羧酸单酯。

[0158] 作为反应底物的式(I')所表示的二羧酸二酯通常在底物浓度为0.01w/v%~90w/v%、优选为0.1w/v%~30w/v%的范围内使用。反应底物可以在反应开始时全部一起添加,但从减少存在酶的底物抑制的情况下的影响的观点和提高产物的蓄积浓度的观点出发,优选连续地或间歇地添加。

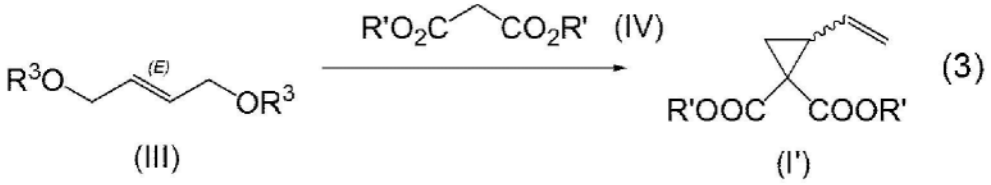
[0159] 反应通常在水性介质中、或者水性介质与有机溶剂的混合物中进行。作为水性介质,可以列举例如水或缓冲液。作为有机溶剂,可以使用甲醇、乙醇、1-丙醇、2-丙醇、1-丁醇、叔丁醇、丙酮、二甲基亚砜等对于作为反应底物的式(I')所表示的二羧酸二酯的溶解度高的溶剂。另外,作为有机溶剂,也可以使用对于反应副产物的除去等有效的乙酸乙酯、乙酸丁酯、甲苯、氯仿、正己烷等。

[0160] 反应在通常为4℃~60℃、优选为10℃~45℃的反应温度下、并且在通常为pH 3~11、优选为pH 5~8的条件下进行。反应时间通常为约1小时~约72小时。

[0161] 对于利用本发明的制造方法生成的式(II')所表示的羧酸单酯,可以在反应结束后通过离心分离、膜处理等本领域技术人员公知的分离或纯化方法分离出反应液中的菌体、蛋白质等,然后通过将利用1-丁醇、叔丁醇等有机溶剂的提取、蒸馏、利用离子交换树脂或硅胶等的柱色谱法、等电点结晶或利用一盐酸盐、二盐酸盐、钙盐等的结晶等适当组合而进行纯化。

[0162] 本发明的水解酶可以特别适合用于将2-乙烯基环丙烷-1,1-二羧酸二乙酯水解而制造(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸的方法。

[0163] 另外,作为底物的式(I')所表示的二羧酸二酯可以通过下述式(3)的反应来制造。

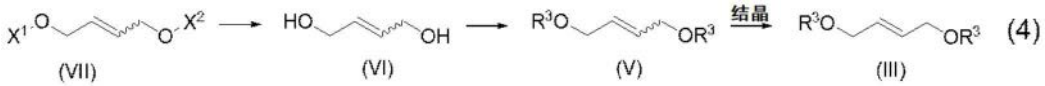
[0164] 

[0165] (式中,R^3表示可以被取代的碳原子数6～12的芳基磺酰基、可以被取代的碳原子数1～10的烷基磺酰基、或者可以被取代的碳原子数7～20的芳烷基磺酰基,R'与上述的含义相同。)

[0166] 即,可以通过在碱金属醇盐或碱金属氢化物的存在下使式(III)所表示的化合物与式(IV)所表示的丙二酸酯类反应而制造。

[0167] 式(III)中,作为R^3所表示的"可以被取代的碳原子数6～12的芳基磺酰基"中的"碳原子数6～12的芳基磺酰基",可以列举例如苯磺酰基、1-萘磺酰基、2-萘磺酰基等。作为"可以被取代的碳原子数1～10的烷基磺酰基"中的"碳原子数1～10的烷基磺酰基",可以列举例如甲磺酰基、乙磺酰基、丙磺酰基等。作为"可以被取代的碳原子数7～20的芳烷基磺酰基"中的"碳原子数7～20的芳烷基磺酰基",可以列举例如苄基磺酰基等。作为"可以被取代的碳原子数6～12的芳基磺酰基"、"可以被取代的碳原子数1～10的烷基磺酰基"和"可以被取代的碳原子数7～20的芳烷基磺酰基"中的取代基,可以列举卤原子(例如氯原子、溴原子)、碳原子数1～6的直链状或支链的烷基(例如甲基、乙基)、碳原子数1～6的直链状或支链的烷氧基(例如甲氧基、乙氧基)、硝基等。作为R^3的适当的具体例,可以列举例如苯磺酰基、对甲苯磺酰基、1-萘磺酰基、2-萘磺酰基、甲磺酰基、乙磺酰基、丙磺酰基、三氟甲磺酰基和苄基磺酰基等。R^3优选为甲磺酰基、苯磺酰基和对甲苯磺酰基,特别优选为对甲苯磺酰基。

[0168] 在此,作为式(3)的原料化合物的式(III)所表示的化合物可以依照公知的方法、例如Frederic Dolle等、Bioorg.Med.Chem.2006,14,1115-1125等中记载的方法而制造。另外,可以通过下述式(4)的反应来制造。

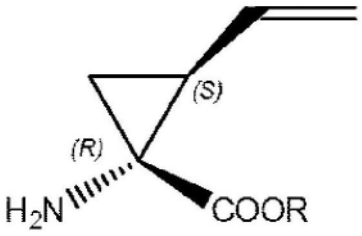
[0169] 

[0170] 即,可以通过使1,4-丁烯二醇(VI)与R^3X(X表示氯原子、溴原子、碘原子等卤原子)反应,使所得到的式(V)的化合物结晶而制造;或者通过将市售的原料化合物(VII)利用酸或碱进行水解而得到1,4-丁烯二醇,进一步与R^3X反应,使所得到的式(V)的化合物结晶而制造。

[0171] 式(4)中,X^1表示氢原子或R^1,X^2表示氢原子或R^2,R^1和R^2各自独立地表示碳原子数2～11的烷基羰基、碳原子数8～21的芳烷基羰基或碳原子数7～13的芳基羰基。但是,除外X^1和X^2同时为氢原子的情况。优选X^1为R^1且X^2为R^2,从能够以市售品而获得的观点出发,更优选R^1和R^2均为乙酰基、乙基羰基、叔丁基羰基或苯甲酰基,进一步优选均为乙酰基。

[0172] 通过使用本发明中得到的式(II')所表示的羧酸单酯,能够制造在作为丙型肝炎治疗药开发中的各种HCV NS3蛋白酶抑制剂等的制造中有用的中间体(1R,2S)-1-氨基-1-乙氧基羰基-2-乙烯基环丙烷(VIII)和其盐。进一步,通过使用这样得到的式(II')所表示的羧酸单酯来进行制造,还能够高效率、低成本地制造光学纯度高的(1R,2S)-1-氨基-1-烷

氧基羰基-2-乙烯基环丙烷。

[0173] 

(VIII)

[0174] 以下,利用实施例进一步详细地说明本发明,但本发明并不限定于此。

[0175] 实施例

[0176] 实施例1:水解酶基因的克隆

[0177] 以水解活性为指标从土壤样品中进行探索,得到红球菌属(Rhodococcus sp.)D32株。根据本微生物产生的水解酶的分析和基因信息的分析得到新的水解酶RsD32Est(序列号2)。基于本氨基酸序列信息,由DNA2.0公司人工合成编码RsD32Est的按大肠杆菌表达用途进行了密码子优化的基因序列(序列号1),得到pJ411RsD32Est。

[0178] 由DNA2.0公司人工合成编码来源于嗜二氧杂环己烷假诺卡氏菌CB1190株(Pseudonocardia dioxanivorans CB1190株)、巧克力微杆菌(Microbacterium chocolatum)的推定上的水解酶PdEst(GenBank登录号WP_01367363、序列号4)、McEst(GenBank登录号WP_053548180、序列号6)和水解酶AFX20780(GenBank登录号AFX20780、序列号8)的按大肠杆菌表达用途进行了密码子优化的基因序列(pedst_Ecodon、序列号3、mcest_Ecodon、序列号5和afx20780_Ecodon),插入到pJExpress411(DNA2.0公司制造)中,制作质粒(分别为pJ411PdEst、pJ411McEst和pJ411AFX20780)。将各基因序列所编码的氨基酸序列分别命名为RsD32Est、PdEst、McEst、AFX20780。各氨基酸序列分别如序列号2(RsD32Est)、序列号4(PdEst)、序列号6(McEst)、序列号8(AFX20780)所示。

[0179] 使用所得到的各质粒,按照常规方法对大肠杆菌(Eschelichia coli)BL21(DE3)(Invitrogen公司制造)进行转化,得到重组大肠杆菌BL21(DE3)/pJ411RsD32Est、BL21(DE3)/pJ411PdEst、BL21(DE3)/pJ411McEst和BL21(DE3)/pJ411AFX20780。为了得到表达所导入的基因的菌体,使用包含卡那霉素和lac启动子诱导物质的液体LB培养基在30℃下对各重组大肠杆菌进行培养,在培养后约20小时收集细菌。

[0180] 对于通过后述参考例1制作的大肠杆菌克隆No.26,也使用包含卡那霉素和lac启动子诱导物质的液体LB培养基在30℃下进行培养,在培养后约20小时收集细菌。

[0181] 实施例2

[0182] 使用实施例1中得到的各菌体,在包含30g/L的外消旋体的2-乙烯基环丙烷-1,1-二羧酸二乙酯(以下称为VCPDE)和5%的二甲基亚砜的100mmol/L磷酸钾缓冲液中在30℃、pH 7的条件下使其反应21小时。作为各菌体,使用通过离心分离从实施例1中制作的培养液0.4mL中得到的菌体。需要说明的是,VCPDE是依照日本专利第5657560号的合成例中记载的方法合成的。

[0183] 用乙腈将20小时反应后的反应液稀释为适当浓度后,利用下述条件的高效液相色谱(HPLC)进行分析。分析条件(1)为对光学纯度进行评价的分析条件。分析条件(2)为用于

对生成的以1-乙氧基羰基-2-乙烯基环丙烷羧酸的(1S,2S)体作为主要成分的Anti体的定量和Syn体相对于Anti体的生成比率进行评价的分析条件。

[0184] 分析条件(1)

[0185] 柱:CHIRALPAK AD-3(4.6×250mm、Daicel化学公司制造)

[0186] 洗脱液:己烷:乙醇:三氟乙酸＝95:5:0.1

[0187] 流速:0.8ml/分钟

[0188] 温度:30℃

[0189] 检测:UV 210nm

[0190] 分析条件(2)

[0191] 柱:COSMOSIL 5C18-MS-II(4.6×250mm、Nacalai Tesque公司制造)

[0192] 洗脱液A:含有0.1%三氟苯乙酮的水溶液

[0193] 洗脱液B:含有0.1%三氟苯乙酮的乙腈

[0194] 流速:0.8ml/分钟

[0195] 温度:40℃

[0196] 检测:UV 210nm

[0197] 表1.梯度洗脱程序

[0198]

分钟	洗脱液A(%)	洗脱液B(%)
0	75	25
20	30	70
23	30	70
30	75	25

[0199] 表1中示出在使用各菌体的情况下生成的1-乙氧基羰基-2-乙烯基环丙烷羧酸的Anti体的量、Syn体相对于Anti体的生成比率和光学纯度。

[0200] 表2.

[0201]

氨基酸序列	Anti体 (g/L)	Syn体相对于Anti体的生成比率 (%)	光学纯度 (%e.e.)
RsD32Est	13.6	未检出	99.1
PdEst	4.28	未检出	80.9
McEst	4.71	0.2	64.4
AFX20780	2.71	0.5	-45.9
PnbA3027-m26	2.69	14.2	99.0

[0202] 由表2表明,与具有序列号8(AFX20780)或序列号24(PnbA3027-m26)的氨基酸序列的现有的水解酶相比,具有序列号2(RsD32Est)、序列号4(PdEst)和序列号6(McEst)的氨基酸序列的本发明的水解酶所生成的1-乙氧基羰基-2-乙烯基环丙烷羧酸的Anti体的量多,Syn体相对于Anti体的生成比率低。

[0203] 特别是具有序列号2(RsD32Est)的本发明的水解酶所生成的1-乙氧基羰基-2-乙烯基环丙烷羧酸的Anti体的量极多,未生成Syn体,光学纯度也极高。

[0204] 另外,图1是使用具有氨基酸序列RsD32Est的菌体得到的反应产物的HPLC分析结

果,图2是使用具有氨基酸序列PnbA3027-m26的菌体得到的反应产物的HPLC分析结果。由图1表明,具有氨基酸序列RsD32Est的本发明的水解酶几乎未生成1-乙氧基羰基-2-乙烯基环丙烷羧酸的(1S,2R)体、(1R,2S)体和(1R,2R)体,能够高选择性地生成(1S,2S)体。由图2表明,具有氨基酸序列PnbA3027-m26的现有的水解酶的1-乙氧基羰基-2-乙烯基环丙烷羧酸的(1S,2R)体的生成量多,与本发明的水解酶相比选择性低。

[0205] 实施例3:(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸的制造

[0206] 在5L发酵罐中将1mol/L磷酸钾缓冲液(pH 7.0)160mL、脱盐水1332mL、2-乙烯基环丙烷-1,1-二羧酸二乙酯84.2g和实施例1中得到的大肠杆菌克隆BL21(DE3)/pJ411RsD32Est的湿菌体24g混合,设定反应温度为30℃、反应时的pH为7.0,在充分的搅拌速度下使其反应24小时。反应终点时的(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸的浓度为22.3g/L。

[0207] 从反应结束后的液体1100mL(纯的(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸为22.2g)中去除菌体和菌体残渣后,添加甲苯500mL,在室温下搅拌20分钟后,分离甲苯层,去除残留的2-乙烯基环丙烷-1,1-二羧酸二乙酯。对所得到的水层添加6N(=3mol/L)的硫酸,将pH调节为2.0。然后,添加甲苯250mL,在室温下搅拌20分钟后,获得包含(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸的甲苯层。再次对所得到的水层添加甲苯250mL,在室温下搅拌20分钟后,获得包含(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸的甲苯层。通过HPLC分析对所得到的甲苯层进行纯度分析和光学纯度分析,结果纯的(1S,2S)-1-乙氧基羰基-2-乙烯基环丙烷羧酸为19.3g,光学纯度为99.7%e.e.。

[0208] 参考例1:对硝基苄基酯酶基因(pnbA3027;序列号9)的克隆

[0209] 基因的克隆

[0210] 将枯草杆菌(Bacillus subtilis)NBRC3027在菌株保藏机构指定的液体培养基中分别培养一晚,使用DNeasy血液和组织试剂盒(Qiagen公司制造)从所得到的菌体中分别制备染色体DNA。

[0211] 基于编码来源于基因组序列公知株枯草杆菌(Bacillus subtilis)ATCC23857的对硝基苄基酯酶(PNBE23857、GenBank登录号AQR87688、序列号12)的基因序列(pnbA23857、序列号11),设计、合成用于扩增全长对硝基苄基酯酶基因的引物。将各自的碱基序列示于序列表的序列号13(pnbA F)和14(pnbA R)。

[0212] 以所制备的染色体DNA作为模板,以pnbA F、pnbA R作为引物,通过PCR扩增出约1.5kbp的DNA片段。

[0213] 表达载体的制备

[0214] 以依照日本特开2005-34025号公报记载的方法制备的质粒pKV32作为模板,使用序列号15所示的引物(pKVXmaIFW)和序列号16所示的引物(pKVXmaIRV),通过PCR扩增出约4kbp的片段。将扩增出的片段用XmaI进行消化,利用Ligation-Convenience试剂盒(日本基因公司制造)使其自身闭环,将所得到的质粒命名为pKW32。

[0215] 表达用质粒的制备

[0216] 按照常规方法将上述基因的克隆中得到的DNA片段导入到上述的“表达载体的制备”中制备的克隆载体pKW32中。以下,将所得到的质粒作为ppnbA3027。

[0217] 对插入到质粒中的DNA序列进行分析,确认其包含由1467bp的ORF构成的基因。将

所得到的基因命名为pnbA3027,序列如序列号9所示。将该DNA序列所编码的氨基酸序列命名为PNBE3027,其氨基酸序列如序列号10所示。PNBE3027的氨基酸序列相对于已知的PNBE23857的序列显示出90％的序列一致性。

[0218] 参考例2:通过突变导入进行的对硝基苄基酯酶基因的改造

[0219] 以参考例1中得到的质粒ppnbA3027作为模板,使用序列表的序列号17所示的引物(L70FW)和序列号18所示的引物(L70RV),利用QuikChange多点突变试剂盒(Stratagene公司制造)在编码氨基酸序号第70号的亮氨酸的碱基处导入随机突变。同样地,使用序列表的序列号19所示的引物(L313FW)和序列号20所示的引物(L313RV),对编码氨基酸序号第313号的亮氨酸残基的碱基导入随机突变,并且,使用序列表的序列号21所示的引物(L270L273FW)和序列号22所示的引物(L270L273RV),对编码氨基酸序号第270和273号的亮氨酸残基的碱基导入随机突变,使用所得到的突变导入质粒按照常规方法对大肠杆菌(Escherichia coli)JM109(宝生物株式会社制造)进行转化。

[0220] 从所得到的重组大肠杆菌中分离出克隆No.26,对该克隆所具有的质粒中插入的DNA序列进行分析,结果其序列(命名为pnbA3027-26)如序列号23所示。将该DNA序列所编码的氨基酸序列命名为PNBE3027-m26,其氨基酸序列如序列号24所示。

[0221] 产业上的可利用性

[0222] 通过使用本发明的水解酶等,能够高效率且低成本地在工业上制造光学纯度高的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸。

[0223] 本发明以在日本提交的日本特愿2018-070188为基础,其内容全部包含在本说明书中。

[0001] 序列表

[0002] <110> 株式会社API(API Corporation)

[0003] <120> 新型水解酶和利用该酶的(1S,2S)-1-烷氧基羰基-2-乙烯基环丙烷羧酸的制造方法

[0004] (Novel hydrolase and method for producing (1S,2S)-1-alkoxycarbonyl-2-vinylcyclopropanecarboxylic acid by utilizng thereof)

[0005] <130> 092851

[0006] <150> JP2018-070188

[0007] <151> 2018-03-30

[0008] <160> 24

[0009] <170> PatentIn version 3.5

[0010] <210> 1

[0011] <211> 1062

[0012] <212> DNA

[0013] <213> 人工序列(Artificial Sequence)

[0014] <220>

[0015] <223> Ecoli opt

[0016] <400> 1

[0017] atgaatctgc caccgggtgt ccgctctgtt acgactcaaa cttcacgtct gcgcttgcat 60

[0018] cacttggagg ccggtccggt cgatggcgtg ccactggttc tggtgcacgg taatctgagc 120

[0019] tccggtcgtt tctatgaaga tgtgatgccg gccctcgcga aaacctatcg cgtcatcgca 180

[0020] ccggatatgc gcggtttcgg tgatagcgag cgcgttaccc tggatgcgac gcgcggtctg 240

[0021] gcggactggg cagatgacat tgcggcgctg ctggaagcgt tagacattga ccaggctccg 300

[0022] catctgctgg gctggagcac gggtgcaggc gcgattaccc gttacgtcct ggatggtcgt 360

[0023] accgccgctt cgctgacgtt gatggacccg gtcccgccgt acggtttcgt cggtatgcac 420

[0024] gcagacggca cgccgtggtt tagcgactat gcgggctgtg gtgctggtgt aatgaacacg 480

[0025] gaatttaccg agcgtattgc tgctggtgac cgttccaccg attccctggc atctccgcgt 540

[0026] aacgttgcgg caggcttttg gggtgaggcg ccgccgatca gccaagagcg tgttgacgtg 600

[0027] ctgattgacg agctgttgaa aacctgggtt agcgaagata attttccggg tagcgttgtg 660

[0028] ccgagcaaaa actggccggg tatcgcccct ggcaccaccg gcatcctgaa cgcactgagc 720

[0029] ccgaagtact gcgactggag ccgcatcacc gagctgggta gcaagccgcc gattatgtgg 780

[0030] atccagggtg gccaagatga tgtgatcagc aatgcgagcc acaacgaccc ggcgaccctg 840

[0031] ggcgcagcgg gcctgatccc gggttggccg ggcgaagatg tctgcccggc gcagccgatg 900

[0032] atcacccaga ttcgtgatgt gttgcaagca tacgaagatg ccggcggccg tacgcgtacc 960

[0033] gagtggttcg aggcaagcca ccatctgcct atgattgaag aaccggaccg ttggctgcaa 1020

[0034] gccgtgtcta gctttgtggc ggaagccgac gcgattgcgt aa 1062

[0035] <210> 2

[0036] <211> 353

[0037] <212> PRT

[0038] <213> Rhodococcus sp. D32

[0039] <400> 2

[0040] Met Asn Leu Pro Pro Gly Val Arg Ser Val Thr Thr Gln Thr Ser Arg
[0041] 1 5 10 15
[0042] Leu Arg Leu His His Leu Glu Ala Gly Pro Val Asp Gly Val Pro Leu
[0043] 20 25 30
[0044] Val Leu Val His Gly Asn Leu Ser Ser Gly Arg Phe Tyr Glu Asp Val
[0045] 35 40 45
[0046] Met Pro Ala Leu Ala Lys Thr Tyr Arg Val Ile Ala Pro Asp Met Arg
[0047] 50 55 60
[0048] Gly Phe Gly Asp Ser Glu Arg Val Thr Leu Asp Ala Thr Arg Gly Leu
[0049] 65 70 75 80
[0050] Ala Asp Trp Ala Asp Asp Ile Ala Ala Leu Leu Glu Ala Leu Asp Ile
[0051] 85 90 95
[0052] Asp Gln Ala Pro His Leu Leu Gly Trp Ser Thr Gly Ala Gly Ala Ile
[0053] 100 105 110
[0054] Thr Arg Tyr Val Leu Asp Gly Arg Thr Ala Ala Ser Leu Thr Leu Met
[0055] 115 120 125
[0056] Asp Pro Val Pro Pro Tyr Gly Phe Val Gly Met His Ala Asp Gly Thr
[0057] 130 135 140
[0058] Pro Trp Phe Ser Asp Tyr Ala Gly Cys Gly Ala Gly Val Met Asn Thr
[0059] 145 150 155 160
[0060] Glu Phe Thr Glu Arg Ile Ala Ala Gly Asp Arg Ser Thr Asp Ser Leu
[0061] 165 170 175
[0062] Ala Ser Pro Arg Asn Val Ala Ala Gly Phe Trp Gly Glu Ala Pro Pro
[0063] 180 185 190
[0064] Ile Ser Gln Glu Arg Val Asp Val Leu Ile Asp Glu Leu Leu Lys Thr
[0065] 195 200 205
[0066] Trp Val Ser Glu Asp Asn Phe Pro Gly Ser Val Val Pro Ser Lys Asn
[0067] 210 215 220
[0068] Trp Pro Gly Ile Ala Pro Gly Thr Thr Gly Ile Leu Asn Ala Leu Ser
[0069] 225 230 235 240
[0070] Pro Lys Tyr Cys Asp Trp Ser Arg Ile Thr Glu Leu Gly Ser Lys Pro
[0071] 245 250 255
[0072] Pro Ile Met Trp Ile Gln Gly Gly Gln Asp Asp Val Ile Ser Asn Ala
[0073] 260 265 270
[0074] Ser His Asn Asp Pro Ala Thr Leu Gly Ala Ala Gly Leu Ile Pro Gly
[0075] 275 280 285
[0076] Trp Pro Gly Glu Asp Val Cys Pro Ala Gln Pro Met Ile Thr Gln Ile
[0077] 290 295 300
[0078] Arg Asp Val Leu Gln Ala Tyr Glu Asp Ala Gly Gly Arg Thr Arg Thr
[0079] 305 310 315 320
[0080] Glu Trp Phe Glu Ala Ser His His Leu Pro Met Ile Glu Glu Pro Asp
[0081] 325 330 335

[0082] Arg Trp Leu Gln Ala Val Ser Ser Phe Val Ala Glu Ala Asp Ala Ile

[0083] 340 345 350

[0084] Ala

[0085] <210> 3

[0086] <211> 1053

[0087] <212> DNA

[0088] <213> 人工序列

[0089] <220>

[0090] <223> Ecoli opt

[0091] <400> 3

[0092] atgcagctgt tgcctggtgt tcgtagcgtg gtggtcccaa ctgatcgtct ggaagttcat 60

[0093] ttggtggagt atggtcctga agagggcgtt ccggttgtga tgctgcacgg taacctttct 120

[0094] accggtcgtt tctttgagca cctgttcccg ggtgccccgc agggttaccg tatcatcgca 180

[0095] ccggacatgc gttgctttgg cgatacggaa cgtctgccgt tagacgcaac ccgtggcctg 240

[0096] gctgactggg ccgatgacgt tgcggctctg ctgcgtgcac tgcgcgttga gcgtccggtt 300

[0097] cacctcctgg gctggagcac tgcgggcgcg gccattgtgg attttgcaag cgcgcatccg 360

[0098] gtcacctctc tgaccttcct ggatccggtc agcccgtacg gtttcggtgg cgtgctggcg 420

[0099] gatggtacgc cgtgcttccc ggactttgcc ggttccggcg gcggcattgt gaatccagaa 480

[0100] gtcgtccgtc gcctggccga gggtgacgat acgaccgaga gcccgttttc tattcgcagc 540

[0101] gtcatgcgtg gctcgtattg gttggaaagc catagcgaac cgcgcgaaga tctgctggta 600

[0102] gcagaagtgc tgaaaaccgt gaccggcgac gacaattatc ctggtgactc cgtcgcgagc 660

[0103] ccgaattggc cgggtagcgc accgggtacg accggtatca tcaacgctct gagcccgaag 720

[0104] tactgtaact ggacgcgcgt tgttgacctg gatccgaaac cgccggtttt gtggacgcat 780

[0105] ggtgccgagg acaccgtcgt cgcggatgcg agcatgcaag acctgggtac cctgggcgag 840

[0106] ctgggttacg tgccgggctg gccgggtgca gatgtgttcc cgagccagcc gatggtgagc 900

[0107] caaattcgcg aagtcctggg ccgttacgcg gcggcgggtg gccacgttcg caccgagatc 960

[0108] ctgccgggtg cgggccactc cccgcacatt gagctgccag aactgtggag cggtgtgttt 1020

[0109] tgggacttcg ttggtgcggc ggagcgtggt taa 1053

[0110] <210> 4

[0111] <211> 350

[0112] <212> PRT

[0113] <213> Pseusdonocardia dioxanivorans CB1190

[0114] <400> 4

[0115] Met Gln Leu Leu Pro Gly Val Arg Ser Val Val Val Pro Thr Asp Arg

[0116] 1 5 10 15

[0117] Leu Glu Val His Leu Val Glu Tyr Gly Pro Glu Glu Gly Val Pro Val

[0118] 20 25 30

[0119] Val Met Leu His Gly Asn Leu Ser Thr Gly Arg Phe Phe Glu His Leu

[0120] 35 40 45

[0121] Phe Pro Gly Ala Pro Gln Gly Tyr Arg Ile Ile Ala Pro Asp Met Arg

[0122] 50 55 60

[0123] Cys Phe Gly Asp Thr Glu Arg Leu Pro Leu Asp Ala Thr Arg Gly Leu

[0124] 65 70 75 80

[0125] Ala Asp Trp Ala Asp Asp Val Ala Ala Leu Leu Arg Ala Leu Arg Val

[0126] 85 90 95

[0127] Glu Arg Pro Val His Leu Leu Gly Trp Ser Thr Ala Gly Ala Ala Ile

[0128] 100 105 110

[0129] Val Asp Phe Ala Ser Ala His Pro Val Thr Ser Leu Thr Phe Leu Asp

[0130] 115 120 125

[0131] Pro Val Ser Pro Tyr Gly Phe Gly Gly Val Leu Ala Asp Gly Thr Pro

[0132] 130 135 140

[0133] Cys Phe Pro Asp Phe Ala Gly Ser Gly Gly Gly Ile Val Asn Pro Glu

[0134] 145 150 155 160

[0135] Val Val Arg Arg Leu Ala Glu Gly Asp Asp Thr Thr Glu Ser Pro Phe

[0136] 165 170 175

[0137] Ser Ile Arg Ser Val Met Arg Gly Ser Tyr Trp Leu Glu Ser His Ser

[0138] 180 185 190

[0139] Glu Pro Arg Glu Asp Leu Leu Val Ala Glu Val Leu Lys Thr Val Thr

[0140] 195 200 205

[0141] Gly Asp Asp Asn Tyr Pro Gly Asp Ser Val Ala Ser Pro Asn Trp Pro

[0142] 210 215 220

[0143] Gly Ser Ala Pro Gly Thr Thr Gly Ile Ile Asn Ala Leu Ser Pro Lys

[0144] 225 230 235 240

[0145] Tyr Cys Asn Trp Thr Arg Val Val Asp Leu Asp Pro Lys Pro Pro Val

[0146] 245 250 255

[0147] Leu Trp Thr His Gly Ala Glu Asp Thr Val Val Ala Asp Ala Ser Met

[0148] 260 265 270

[0149] Gln Asp Leu Gly Thr Leu Gly Glu Leu Gly Tyr Val Pro Gly Trp Pro

[0150] 275 280 285

[0151] Gly Ala Asp Val Phe Pro Ser Gln Pro Met Val Ser Gln Ile Arg Glu

[0152] 290 295 300

[0153] Val Leu Gly Arg Tyr Ala Ala Ala Gly Gly His Val Arg Thr Glu Ile

[0154] 305 310 315 320

[0155] Leu Pro Gly Ala Gly His Ser Pro His Ile Glu Leu Pro Glu Leu Trp

[0156] 325 330 335

[0157] Ser Gly Val Phe Trp Asp Phe Val Gly Ala Ala Glu Arg Gly

[0158] 340 345 350

[0159] <210> 5

[0160] <211> 1113

[0161] <212> DNA

[0162] <213> 人工序列

[0163] <220>

[0164] <223> Ecoli opt

[0165] <400> 5

[0166] atgaccctgt ttgatggcat cacttcgcgt attgttgaca ccgatcgtct gaccgtgaat 60
[0167] attctggagc gtgcggcaga cgatccgcaa acgccgcctg atcgcacggt tgtttttgtg 120
[0168] catggtaacg tcagcagcgc gctgttttgg caagaaatta tgcaagacct cccgagcgat 180
[0169] ctgcgtgcta tcgcggtgga tctgcgtggc ttcggtggca gcgagcacgc accggtcgac 240
[0170] gcgacccgtg gtgttcgcga tttcagcgat gacttgcacg caacgctgga agcactggat 300
[0171] atcccggtcg cacacctggt cggttggagc atgggtggcg gcgtggttat gcagtatgca 360
[0172] ttagaccatc cggttttgag cctgacgctg cagtccccgg tcagcccgta cggtttcggt 420
[0173] ggcacccgtc gtgacggcag ccgtctgact gatgacgatg ctggctgcgg tggcggcggc 480
[0174] gccaatccgg atttcattca gcgtctgatc gaccacgaca cgtccgacga cgcgcagacc 540
[0175] agcccgcgct ccgttttccg tgcgggttac gtagcgagcg attacaccac ggaccacgaa 600
[0176] gatgtgtggg tggagagcat gctgaccacc tctaccgcgg acggcaatta cccaggtgac 660
[0177] gccgtcccga gcgacaactg gccaggtttt gccgcgggtc gccacggtgt tctgaacacg 720
[0178] atggcgcctc aatacttcga tgtcagcggt atcgttgatc tggcggagaa accgccgatt 780
[0179] ttgtggatcc atggtaccgc ggatgcgatt gtgtccgacg cgagctttta tgacctgaac 840
[0180] tatctgggtc agctgggtat cgtgccgggt tggccgggtg aagatgtcgc tccggcccaa 900
[0181] gaaatggtta gccagacgcg cgacgtcctg ggccgttacg ctgcgggtgg cggcaccgtt 960
[0182] accgaagttg ccgttgaggg tgcgggtcac agcgcacatt tggagcgtcc ggccgtgttc 1020
[0183] cgccacgcgc tgctggaaat tattggttat gtgggtgcgg ctgctgaccc ggcaccgccg 1080
[0184] actgaagcaa tcatcatccg ctctgccgac taa 1113
[0185] <210> 6
[0186] <211> 370
[0187] <212> PRT
[0188] <213> Microbacterium chocolatum
[0189] <400> 6
[0190] Met Thr Leu Phe Asp Gly Ile Thr Ser Arg Ile Val Asp Thr Asp Arg
[0191] 1 5 10 15
[0192] Leu Thr Val Asn Ile Leu Glu Arg Ala Ala Asp Asp Pro Gln Thr Pro
[0193] 20 25 30
[0194] Pro Asp Arg Thr Val Val Phe Val His Gly Asn Val Ser Ser Ala Leu
[0195] 35 40 45
[0196] Phe Trp Gln Glu Ile Met Gln Asp Leu Pro Ser Asp Leu Arg Ala Ile
[0197] 50 55 60
[0198] Ala Val Asp Leu Arg Gly Phe Gly Gly Ser Glu His Ala Pro Val Asp
[0199] 65 70 75 80
[0200] Ala Thr Arg Gly Val Arg Asp Phe Ser Asp Asp Leu His Ala Thr Leu
[0201] 85 90 95
[0202] Glu Ala Leu Asp Ile Pro Val Ala His Leu Val Gly Trp Ser Met Gly
[0203] 100 105 110
[0204] Gly Gly Val Val Met Gln Tyr Ala Leu Asp His Pro Val Leu Ser Leu
[0205] 115 120 125
[0206] Thr Leu Gln Ser Pro Val Ser Pro Tyr Gly Phe Gly Gly Thr Arg Arg
[0207] 130 135 140

[0208] Asp Gly Ser Arg Leu Thr Asp Asp Asp Ala Gly Cys Gly Gly Gly Gly
[0209] 145 150 155 160
[0210] Ala Asn Pro Asp Phe Ile Gln Arg Leu Ile Asp His Asp Thr Ser Asp
[0211] 165 170 175
[0212] Asp Ala Gln Thr Ser Pro Arg Ser Val Phe Arg Ala Gly Tyr Val Ala
[0213] 180 185 190
[0214] Ser Asp Tyr Thr Thr Asp His Glu Asp Val Trp Val Glu Ser Met Leu
[0215] 195 200 205
[0216] Thr Thr Ser Thr Ala Asp Gly Asn Tyr Pro Gly Asp Ala Val Pro Ser
[0217] 210 215 220
[0218] Asp Asn Trp Pro Gly Phe Ala Ala Gly Arg His Gly Val Leu Asn Thr
[0219] 225 230 235 240
[0220] Met Ala Pro Gln Tyr Phe Asp Val Ser Gly Ile Val Asp Leu Ala Glu
[0221] 245 250 255
[0222] Lys Pro Pro Ile Leu Trp Ile His Gly Thr Ala Asp Ala Ile Val Ser
[0223] 260 265 270
[0224] Asp Ala Ser Phe Tyr Asp Leu Asn Tyr Leu Gly Gln Leu Gly Ile Val
[0225] 275 280 285
[0226] Pro Gly Trp Pro Gly Glu Asp Val Ala Pro Ala Gln Glu Met Val Ser
[0227] 290 295 300
[0228] Gln Thr Arg Asp Val Leu Gly Arg Tyr Ala Ala Gly Gly Gly Thr Val
[0229] 305 310 315 320
[0230] Thr Glu Val Ala Val Glu Gly Ala Gly His Ser Ala His Leu Glu Arg
[0231] 325 330 335
[0232] Pro Ala Val Phe Arg His Ala Leu Leu Glu Ile Ile Gly Tyr Val Gly
[0233] 340 345 350
[0234] Ala Ala Ala Asp Pro Ala Pro Pro Thr Glu Ala Ile Ile Ile Arg Ser
[0235] 355 360 365
[0236] Ala Asp
[0237] 370
[0238] <210> 7
[0239] <211> 894
[0240] <212> DNA
[0241] <213> 人工序列
[0242] <220>
[0243] <223> Ecoli opt
[0244] <400> 7
[0245] atggaaaagc gcagcattac tctgaaaaac ggtgaagttt acaaatacgt ggaacagggt 60
[0246] cagggcgacc aagttctgct gctgatccat ggcaacttta gctccagcct gcacttcacc 120
[0247] ccgttgctgg agcgcctgcc gaagaacatt aaggttattg cgccggactt gcgtggttat 180
[0248] ggtgattcat cttactatcg tcgtatcagc tctctgaacg acttcgcgga agatgtccac 240
[0249] atgtttatgg aagccaaaga aattaaaagc taccacgtag ttggctggag cctgggtggc 300

[0250] ggcgtggcgc tggaactggc ggcacatcat ccagaggctg tggagtcctt ggtgctcatc 360
[0251] aacagcacca cgcacaaggg ttatccggtg ttcaagaaag gtgcggatgg taagccactg 420
[0252] gtcggtgaag tctatcaaag cgccgatgag atggcaaatg atccggtcca agtcctgccg 480
[0253] ctgctgaagg ctcaaaaaga taagaatttt gactttgtta gctatatctt cgatgttacc 540
[0254] atttacaccg tgaataagcc ttcggtggat gacaacaagc tgtggattaa cgaaagcctg 600
[0255] aaacagcgta atctgccgga tgcagactgg gcgctggcga atttgaacat gagcgaccag 660
[0256] cacaatttct acaatgccgg tatgaacaat atctccaagg ttaaggcacc ggtgctgcac 720
[0257] acgtggggcg acaaagatat taccgtgccg gagtatatga tcaaagacaa tgtgaaagcg 780
[0258] ctggaagaac aaagcaaact ggtcgtctac gagaactgcg gccatagccc gctggttgat 840
[0259] gttccggacc agctgacgaa agacatcctg gactttatcg gttacaaggg ttaa 894
[0260] <210> 8
[0261] <211> 297
[0262] <212> PRT
[0263] <213> Unknown
[0264] <220>
[0265] <223> patented sequence
[0266] <400> 8
[0267] Met Glu Lys Arg Ser Ile Thr Leu Lys Asn Gly Glu Val Tyr Lys Tyr
[0268] 1 5 10 15
[0269] Val Glu Gln Gly Gln Gly Asp Gln Val Leu Leu Leu Ile His Gly Asn
[0270] 20 25 30
[0271] Phe Ser Ser Ser Leu His Phe Thr Pro Leu Leu Glu Arg Leu Pro Lys
[0272] 35 40 45
[0273] Asn Ile Lys Val Ile Ala Pro Asp Leu Arg Gly Tyr Gly Asp Ser Ser
[0274] 50 55 60
[0275] Tyr Tyr Arg Arg Ile Ser Ser Leu Asn Asp Phe Ala Glu Asp Val His
[0276] 65 70 75 80
[0277] Met Phe Met Glu Ala Lys Glu Ile Lys Ser Tyr His Val Val Gly Trp
[0278] 85 90 95
[0279] Ser Leu Gly Gly Gly Val Ala Leu Glu Leu Ala Ala His His Pro Glu
[0280] 100 105 110
[0281] Ala Val Glu Ser Leu Val Leu Ile Asn Ser Thr Thr His Lys Gly Tyr
[0282] 115 120 125
[0283] Pro Val Phe Lys Lys Gly Ala Asp Gly Lys Pro Leu Val Gly Glu Val
[0284] 130 135 140
[0285] Tyr Gln Ser Ala Asp Glu Met Ala Asn Asp Pro Val Gln Val Leu Pro
[0286] 145 150 155 160
[0287] Leu Leu Lys Ala Gln Lys Asp Lys Asn Phe Asp Phe Val Ser Tyr Ile
[0288] 165 170 175
[0289] Phe Asp Val Thr Ile Tyr Thr Val Asn Lys Pro Ser Val Asp Asp Asn
[0290] 180 185 190
[0291] Lys Leu Trp Ile Asn Glu Ser Leu Lys Gln Arg Asn Leu Pro Asp Ala

[0292] 195 200 205

[0293] Asp Trp Ala Leu Ala Asn Leu Asn Met Ser Asp Gln His Asn Phe Tyr

[0294] 210 215 220

[0295] Asn Ala Gly Met Asn Asn Ile Ser Lys Val Lys Ala Pro Val Leu His

[0296] 225 230 235 240

[0297] Thr Trp Gly Asp Lys Asp Ile Thr Val Pro Glu Tyr Met Ile Lys Asp

[0298] 245 250 255

[0299] Asn Val Lys Ala Leu Glu Glu Gln Ser Lys Leu Val Val Tyr Glu Asn

[0300] 260 265 270

[0301] Cys Gly His Ser Pro Leu Val Asp Val Pro Asp Gln Leu Thr Lys Asp

[0302] 275 280 285

[0303] Ile Leu Asp Phe Ile Gly Tyr Lys Gly

[0304] 290 295

[0305] <210> 9

[0306] <211> 1470

[0307] <212> DNA

[0308] <213> Bacillus subtilis NBRC3027

[0309] <400> 9

[0310] atgactcatc aaatagtaac gactcaatac ggcaaagtaa aaggcacaac agaaaatggc 60

[0311] gtacataaat ggaaaggcat cccctatgcc agaccgcctg tcgggccatt gcgttttaaa 120

[0312] gcaccggaac ctccggaagc gtgggagaac gaactggacg caacagcgta cggccctatt 180

[0313] tgcccgcagc cgtctgattt gctgtcactt tcgtataatg agctgccccg ccagtctgag 240

[0314] aattgcttgt atgtcaatgt atttgcgcct gatactccaa gtcaaaacct gcctgtcatg 300

[0315] gtgtggattc acggcggcgc tttttatctt ggagcgggca gtgagccatt attcgatggg 360

[0316] tcaagacttg cggcgcaggg agaagtcatt gtcgttacac tgaattatcg actggggccg 420

[0317] tttggatttt tacatttgtc ttcgtttgat gaggcgtatt ccgataacct tggtcttttg 480

[0318] gaccaagccg ccgcactgaa atgggtgcga gacaatatct cagcatttgg cggtgatccg 540

[0319] gataacgtaa cggtatttgg agaatccgct ggcggcatga gcattgccgc gctgctcgca 600

[0320] atgcctgcgg caaaaggcct gttccagaaa gcgatcatgg aaagcggcgc ttctagaaca 660

[0321] atgacaacag aaaaagcggc tagcactgca gcagcctttt tacaggtcct tgggattaac 720

[0322] gagagccaat tggacagact gcacactgta tctgcggaag atttgcttaa agcggccgat 780

[0323] aagcttcgga aagcagaaaa tgaaaatctc tttcagctgt tattccagcc cgcccttgat 840

[0324] ccgaaaacgc tgcctgctga accagaaaaa gcaatcgcag gaggtgctgc tgccgacatt 900

[0325] ccgctgttaa tcggaacaaa ccgcgatgaa ggatatttat ttttcacccc ggactcagac 960

[0326] gttcattctc aagaaacgtt taatgccgcg cttgagtatt tattggaaca gccgctggca 1020

[0327] aagaaagccg ccgatctgta tccgcgttca ctagaaagcc aaattcatat gatgactgat 1080

[0328] ttgttatttt ggcgcccggc cgtcgcctat gcctccgctc agtctcaata cgcgcctgta 1140

[0329] tggatgtacc gatttgattg gcactctgat aagccgccat acaataaggc gtttcacgca 1200

[0330] ttagagcttc cttttgtttt cggaaatctg gacgggttag aacgaatggc aaaagcagag 1260

[0331] gttacggatg aggtgaaacg gctttctcat accatacaat cagcatggat cacgtttgcc 1320

[0332] aaaacaggaa acccaagcac cgaagatgta aaatggccgg cgtatcatga agaaacaaga 1380

[0333] cagacgctga ttttagattc agagattacg atcgaaaacg atcctgaatc tgaaaaaagg 1440

[0334] cagaagctat tcccttcaaa aggagaataa 1470

[0335] <210> 10

[0336] <211> 489

[0337] <212> PRT

[0338] <213> Bacillus subtilis NBRC3027

[0339] <400> 10

[0340] Met Thr His Gln Ile Val Thr Thr Gln Tyr Gly Lys Val Lys Gly Thr

[0341] 1 5 10 15

[0342] Thr Glu Asn Gly Val His Lys Trp Lys Gly Ile Pro Tyr Ala Arg Pro

[0343] 20 25 30

[0344] Pro Val Gly Pro Leu Arg Phe Lys Ala Pro Glu Pro Pro Glu Ala Trp

[0345] 35 40 45

[0346] Glu Asn Glu Leu Asp Ala Thr Ala Tyr Gly Pro Ile Cys Pro Gln Pro

[0347] 50 55 60

[0348] Ser Asp Leu Leu Ser Leu Ser Tyr Asn Glu Leu Pro Arg Gln Ser Glu

[0349] 65 70 75 80

[0350] Asn Cys Leu Tyr Val Asn Val Phe Ala Pro Asp Thr Pro Ser Gln Asn

[0351] 85 90 95

[0352] Leu Pro Val Met Val Trp Ile His Gly Gly Ala Phe Tyr Leu Gly Ala

[0353] 100 105 110

[0354] Gly Ser Glu Pro Leu Phe Asp Gly Ser Arg Leu Ala Ala Gln Gly Glu

[0355] 115 120 125

[0356] Val Ile Val Val Thr Leu Asn Tyr Arg Leu Gly Pro Phe Gly Phe Leu

[0357] 130 135 140

[0358] His Leu Ser Ser Phe Asp Glu Ala Tyr Ser Asp Asn Leu Gly Leu Leu

[0359] 145 150 155 160

[0360] Asp Gln Ala Ala Ala Leu Lys Trp Val Arg Asp Asn Ile Ser Ala Phe

[0361] 165 170 175

[0362] Gly Gly Asp Pro Asp Asn Val Thr Val Phe Gly Glu Ser Ala Gly Gly

[0363] 180 185 190

[0364] Met Ser Ile Ala Ala Leu Leu Ala Met Pro Ala Ala Lys Gly Leu Phe

[0365] 195 200 205

[0366] Gln Lys Ala Ile Met Glu Ser Gly Ala Ser Arg Thr Met Thr Thr Glu

[0367] 210 215 220

[0368] Lys Ala Ala Ser Thr Ala Ala Ala Phe Leu Gln Val Leu Gly Ile Asn

[0369] 225 230 235 240

[0370] Glu Ser Gln Leu Asp Arg Leu His Thr Val Ser Ala Glu Asp Leu Leu

[0371] 245 250 255

[0372] Lys Ala Ala Asp Lys Leu Arg Lys Ala Glu Asn Glu Asn Leu Phe Gln

[0373] 260 265 270

[0374] Leu Leu Phe Gln Pro Ala Leu Asp Pro Lys Thr Leu Pro Ala Glu Pro

[0375] 275 280 285

[0376] Glu Lys Ala Ile Ala Gly Gly Ala Ala Ala Asp Ile Pro Leu Leu Ile
[0377] 290 295 300
[0378] Gly Thr Asn Arg Asp Glu Gly Tyr Leu Phe Phe Thr Pro Asp Ser Asp
[0379] 305 310 315 320
[0380] Val His Ser Gln Glu Thr Phe Asn Ala Ala Leu Glu Tyr Leu Leu Glu
[0381] 325 330 335
[0382] Gln Pro Leu Ala Lys Lys Ala Ala Asp Leu Tyr Pro Arg Ser Leu Glu
[0383] 340 345 350
[0384] Ser Gln Ile His Met Met Thr Asp Leu Leu Phe Trp Arg Pro Ala Val
[0385] 355 360 365
[0386] Ala Tyr Ala Ser Ala Gln Ser Gln Tyr Ala Pro Val Trp Met Tyr Arg
[0387] 370 375 380
[0388] Phe Asp Trp His Ser Asp Lys Pro Pro Tyr Asn Lys Ala Phe His Ala
[0389] 385 390 395 400
[0390] Leu Glu Leu Pro Phe Val Phe Gly Asn Leu Asp Gly Leu Glu Arg Met
[0391] 405 410 415
[0392] Ala Lys Ala Glu Val Thr Asp Glu Val Lys Arg Leu Ser His Thr Ile
[0393] 420 425 430
[0394] Gln Ser Ala Trp Ile Thr Phe Ala Lys Thr Gly Asn Pro Ser Thr Glu
[0395] 435 440 445
[0396] Asp Val Lys Trp Pro Ala Tyr His Glu Glu Thr Arg Gln Thr Leu Ile
[0397] 450 455 460
[0398] Leu Asp Ser Glu Ile Thr Ile Glu Asn Asp Pro Glu Ser Glu Lys Arg
[0399] 465 470 475 480
[0400] Gln Lys Leu Phe Pro Ser Lys Gly Glu
[0401] 485
[0402] <210> 11
[0403] <211> 1470
[0404] <212> DNA
[0405] <213> Bacillus subtilis ATCC23857
[0406] <400> 11
[0407] atgactcatc aaatagtaac gactcaatac ggcaaagtaa aaggcacaac ggaaaacggc 60
[0408] gtacataagt ggaaaggcat cccctatgcc aagccgcctg tcggacaatg gcgttttaaa 120
[0409] gcacctgagc cgcctgaagt gtgggaagat gtgcttgatg ccacagcgta cggctctatt 180
[0410] tgcccgcagc cgtctgattt gctgtcactt tcgtatactg agctgccccg ccagtccgag 240
[0411] gattgcttgt atgtcaatgt atttgcgcct gacaccccaa gtaaaaatct tcctgtcatg 300
[0412] gtgtggattc acggaggcgc tttttatcta ggagcgggca gtgagccatt gtatgacgga 360
[0413] tcaaaacttg cggcacaggg agaagtcatt gtcgttacat tgaactatcg gctggggccg 420
[0414] tttggctttt tgcacttgtc ttcatttaat gaggcgtatt ctgataacct tgggctttta 480
[0415] gaccaagccg ccgcgctgaa atgggtgcga gagaatattt cagcgtttgg cggtgatccc 540
[0416] gataacgtaa cagtatttgg agaatccgcc ggcgggatga gcattgccgc gctgcttgct 600
[0417] atgcctgcgg caaaaggcct gttccagaaa gcaatcatgg aaagcggcgc ttctcgaacg 660

[0418] atgacgaaag aacaagcggc gagcacctcg gcagcctttt tacaggtcct tgggattaac 720
[0419] gagggccaac tggataaatt gcatacggtt tctgcggaag atttgctaaa agcggctgat 780
[0420] cagcttcgga ttgcagaaaa agaaaatatc tttcagctgt tcttccagcc cgcccttgat 840
[0421] ccgaaaacgc tgcctgaaga accagaaaaa gcgatcgcag aaggggctgc ttccggtatt 900
[0422] ccgctattaa ttggaacaac ccgtgatgaa ggatatttat ttttcacccc ggattcagac 960
[0423] gttcattctc aggaaacgct tgatgcagcg ctcgagtatt tactagggaa gccgctggca 1020
[0424] gagaaagttg ccgatttgta tccgcgttct ctggaaagcc aaattcatat gatgactgat 1080
[0425] ttattatttt ggcgccctgc cgtcgcctat gcatccgcac agtctcatta cgcccctgtc 1140
[0426] tggatgtaca ggttcgattg gcacccgaag aagccgccgt acaataaagc gtttcacgca 1200
[0427] ttagagcttc cttttgtctt tggaaatctg gacggattgg aacgaatggc aaaagcggag 1260
[0428] attacggatg aggtgaaaca gctttctcac acgatacaat cagcgtggat cacgttcgcc 1320
[0429] aaaacaggaa acccaagcac cgaagctgtg aattggcctg cgtatcatga agaaacgaga 1380
[0430] gagacgctga ttttagactc agagattacg atcgaaaacg atcccgaatc tgaaaaaagg 1440
[0431] cagaagctat tcccttcaaa aggagaataa 1470
[0432] <210> 12
[0433] <211> 489
[0434] <212> PRT
[0435] <213> Bacillus subtilis ATCC23857
[0436] <400> 12
[0437] Met Thr His Gln Ile Val Thr Thr Gln Tyr Gly Lys Val Lys Gly Thr
[0438] 1 5 10 15
[0439] Thr Glu Asn Gly Val His Lys Trp Lys Gly Ile Pro Tyr Ala Lys Pro
[0440] 20 25 30
[0441] Pro Val Gly Gln Trp Arg Phe Lys Ala Pro Glu Pro Pro Glu Val Trp
[0442] 35 40 45
[0443] Glu Asp Val Leu Asp Ala Thr Ala Tyr Gly Ser Ile Cys Pro Gln Pro
[0444] 50 55 60
[0445] Ser Asp Leu Leu Ser Leu Ser Tyr Thr Glu Leu Pro Arg Gln Ser Glu
[0446] 65 70 75 80
[0447] Asp Cys Leu Tyr Val Asn Val Phe Ala Pro Asp Thr Pro Ser Lys Asn
[0448] 85 90 95
[0449] Leu Pro Val Met Val Trp Ile His Gly Gly Ala Phe Tyr Leu Gly Ala
[0450] 100 105 110
[0451] Gly Ser Glu Pro Leu Tyr Asp Gly Ser Lys Leu Ala Ala Gln Gly Glu
[0452] 115 120 125
[0453] Val Ile Val Val Thr Leu Asn Tyr Arg Leu Gly Pro Phe Gly Phe Leu
[0454] 130 135 140
[0455] His Leu Ser Ser Phe Asn Glu Ala Tyr Ser Asp Asn Leu Gly Leu Leu
[0456] 145 150 155 160
[0457] Asp Gln Ala Ala Ala Leu Lys Trp Val Arg Glu Asn Ile Ser Ala Phe
[0458] 165 170 175
[0459] Gly Gly Asp Pro Asp Asn Val Thr Val Phe Gly Glu Ser Ala Gly Gly

[0460] 180 185 190

[0461] Met Ser Ile Ala Ala Leu Leu Ala Met Pro Ala Ala Lys Gly Leu Phe

[0462] 195 200 205

[0463] Gln Lys Ala Ile Met Glu Ser Gly Ala Ser Arg Thr Met Thr Lys Glu

[0464] 210 215 220

[0465] Gln Ala Ala Ser Thr Ser Ala Ala Phe Leu Gln Val Leu Gly Ile Asn

[0466] 225 230 235 240

[0467] Glu Gly Gln Leu Asp Lys Leu His Thr Val Ser Ala Glu Asp Leu Leu

[0468] 245 250 255

[0469] Lys Ala Ala Asp Gln Leu Arg Ile Ala Glu Lys Glu Asn Ile Phe Gln

[0470] 260 265 270

[0471] Leu Phe Phe Gln Pro Ala Leu Asp Pro Lys Thr Leu Pro Glu Glu Pro

[0472] 275 280 285

[0473] Glu Lys Ala Ile Ala Glu Gly Ala Ala Ser Gly Ile Pro Leu Leu Ile

[0474] 290 295 300

[0475] Gly Thr Thr Arg Asp Glu Gly Tyr Leu Phe Phe Thr Pro Asp Ser Asp

[0476] 305 310 315 320

[0477] Val His Ser Gln Glu Thr Leu Asp Ala Ala Leu Glu Tyr Leu Leu Gly

[0478] 325 330 335

[0479] Lys Pro Leu Ala Glu Lys Val Ala Asp Leu Tyr Pro Arg Ser Leu Glu

[0480] 340 345 350

[0481] Ser Gln Ile His Met Met Thr Asp Leu Leu Phe Trp Arg Pro Ala Val

[0482] 355 360 365

[0483] Ala Tyr Ala Ser Ala Gln Ser His Tyr Ala Pro Val Trp Met Tyr Arg

[0484] 370 375 380

[0485] Phe Asp Trp His Pro Lys Lys Pro Pro Tyr Asn Lys Ala Phe His Ala

[0486] 385 390 395 400

[0487] Leu Glu Leu Pro Phe Val Phe Gly Asn Leu Asp Gly Leu Glu Arg Met

[0488] 405 410 415

[0489] Ala Lys Ala Glu Ile Thr Asp Glu Val Lys Gln Leu Ser His Thr Ile

[0490] 420 425 430

[0491] Gln Ser Ala Trp Ile Thr Phe Ala Lys Thr Gly Asn Pro Ser Thr Glu

[0492] 435 440 445

[0493] Ala Val Asn Trp Pro Ala Tyr His Glu Glu Thr Arg Glu Thr Leu Ile

[0494] 450 455 460

[0495] Leu Asp Ser Glu Ile Thr Ile Glu Asn Asp Pro Glu Ser Glu Lys Arg

[0496] 465 470 475 480

[0497] Gln Lys Leu Phe Pro Ser Lys Gly Glu

[0498] 485

[0499] <210> 13

[0500] <211> 33

[0501] <212> DNA

[0502] <213> 人工序列
[0503] <220>
[0504] <223> 引物(primer)
[0505] <400> 13
[0506] gcgaattcat gactcatcaa atagtaacga ctc 33
[0507] <210> 14
[0508] <211> 33
[0509] <212> DNA
[0510] <213> 人工序列
[0511] <220>
[0512] <223> 引物
[0513] <400> 14
[0514] gctctagatt attctccttt tgaagggaat agc 33
[0515] <210> 15
[0516] <211> 32
[0517] <212> DNA
[0518] <213> 人工序列
[0519] <220>
[0520] <223> 引物
[0521] <400> 15
[0522] tccccccggg tcaaggcgca ctcccgttct gg 32
[0523] <210> 16
[0524] <211> 35
[0525] <212> DNA
[0526] <213> 人工序列
[0527] <220>
[0528] <223> 引物
[0529] <400> 16
[0530] tccccccggg tggggtgcct aatgagtgag ctaac 35
[0531] <210> 17
[0532] <211> 40
[0533] <212> DNA
[0534] <213> 人工序列
[0535] <220>
[0536] <223> 引物
[0537] <220>
[0538] <221> misc_feature
[0539] <222> (19)..(20)
[0540] <223> n is a, c, g, or t
[0541] <400> 17
[0542] ccgtctgatt tgctgtcann ktcgtataat gagctgcccc 40
[0543] <210> 18

[0544] <211> 40
[0545] <212> DNA
[0546] <213> 人工序列
[0547] <220>
[0548] <223> 引物
[0549] <220>
[0550] <221> misc_feature
[0551] <222> (21)..(22)
[0552] <223> n is a, c, g, or t
[0553] <400> 18
[0554] ggggcagctc attatacgam nntgacagca aatcagacgg 40
[0555] <210> 19
[0556] <211> 32
[0557] <212> DNA
[0558] <213> 人工序列
[0559] <220>
[0560] <223> 引物
[0561] <220>
[0562] <221> misc_feature
[0563] <222> (17)..(18)
[0564] <223> n is a, c, g, or t
[0565] <400> 19
[0566] ccgcgatgaa ggatatnnkt ttttcacccc gg 32
[0567] <210> 20
[0568] <211> 32
[0569] <212> DNA
[0570] <213> 人工序列
[0571] <220>
[0572] <223> 引物
[0573] <220>
[0574] <221> misc_feature
[0575] <222> (15)..(16)
[0576] <223> n is a, c, g, or t
[0577] <400> 20
[0578] ccggggtgaa aaamnnatat ccttcatcgc gg 32
[0579] <210> 21
[0580] <211> 53
[0581] <212> DNA
[0582] <213> 人工序列
[0583] <220>
[0584] <223> 引物
[0585] <220>

[0586] <221> misc_feature
[0587] <222> (26)..(27)
[0588] <223> n is a, c, g, or t
[0589] <220>
[0590] <221> misc_feature
[0591] <222> (35)..(36)
[0592] <223> n is a, c, g, or t
[0593] <400> 21
[0594] gcttcggaaa gcagaaaatg aaaatnnktt tcagnnktta ttccagcccg ccc 53
[0595] <210> 22
[0596] <211> 53
[0597] <212> DNA
[0598] <213> 人工序列
[0599] <220>
[0600] <223> 引物
[0601] <220>
[0602] <221> misc_feature
[0603] <222> (18)..(19)
[0604] <223> n is a, c, g, or t
[0605] <220>
[0606] <221> misc_feature
[0607] <222> (27)..(28)
[0608] <223> n is a, c, g, or t
[0609] <400> 22
[0610] gggcgggctg gaataamnnc tgaaamnnat tttcattttc tgctttccga agc 53
[0611] <210> 23
[0612] <211> 1470
[0613] <212> DNA
[0614] <213> Bacillus subtilis NBRC3027
[0615] <400> 23
[0616] atgactcatc aaatagtaac gactcaatac ggcaaagtaa aaggcacaac agaaaatggc 60
[0617] gtacataaat ggaaaggcat cccctatgcc agaccgcctg tcgggccatt gcgttttaaa 120
[0618] gcaccggagc ctccggaagc gtgggagaac gaactggacg caacagcgta cggccctatt 180
[0619] tgcccgcagc cgtctgattt gctgtcagat tcgtataatg agctgccccg ccagtctgag 240
[0620] aattgcttgt atgtcaatgt atttgcgcct gatactccaa gtcaaaacct gcctgtcatg 300
[0621] gtgtggattc acggcggcgc tttttatctt ggagcgggca gtgagccatt attcgatggg 360
[0622] tcaagacttg cggcgcaggg agaagtcatt gtcgttacac tgaattatcg actggggccg 420
[0623] tttggatttt tacatttgtc ttcgtttgat gaggcgtatt ccgataacct tggtcttttg 480
[0624] gaccaagccg ccgcactgaa atgggtgcga gacaatatct cagcatttgg cggtgatccg 540
[0625] gataacgtaa cggtatttgg agaatccgct ggcggcatga gcattgccgc gctgctcgca 600
[0626] atgcctgcgg caaaaggcct gttccagaaa gcgatcatgg aaagcggcgc ttctagaaca 660
[0627] atgacaacag aaaaagcggc tagcactgca gcagcctttt tacaggtcct tgggattaac 720

[0628] gagagccaat tggacagact gcacactgta tctgcggaag atttgcttaa agcggccgat 780
[0629] aagcttcgga aagcagaaaa tgaaaatcag tttcagcggt tattccagcc cgcccttgat 840
[0630] ccgaaaacgc tgcctgctga accagaaaaa gcaatcgcag gaggtgctgc tgccgacatt 900
[0631] ccgctgttaa tcggaacaaa ccgcgatgaa ggatatatgt ttttcacccc ggactcagac 960
[0632] gttcattctc aagaaacgtt taatgccgcg cttgagtatt tattggaaca gccgctggca 1020
[0633] aagaaagccg ccgatctgta tccgcgttca ctagaaagcc aaattcatat gatgactgat 1080
[0634] ttgttatttt ggcgcccggc cgtcgcctat gcctccgctc agtctcaata cgcgcctgta 1140
[0635] tggatgtacc gatttgattg gcactctgat aagccgccat acaataaggc gtttcacgca 1200
[0636] ttagagcttc cttttgtttt cggaaatctg gacgggttag aacgaatggc aaaagcagag 1260
[0637] gttacggatg aggtgaaacg gctttctcat accatacaat cagcatggat cacgtttgcc 1320
[0638] aaaacaggaa acccaagcac cgaagatgta aaatggccgg cgtatcatga agaaacaaga 1380
[0639] cagacgctga ttttagattc agagattacg atcgaaaacg atcctgaatc tgaaaaaagg 1440
[0640] cagaagctat tcccttcaaa aggagaataa 1470
[0641] <210> 24
[0642] <211> 489
[0643] <212> PRT
[0644] <213> Bacillus subtilis NBRC3027
[0645] <400> 24
[0646] Met Thr His Gln Ile Val Thr Thr Gln Tyr Gly Lys Val Lys Gly Thr
[0647] 1 5 10 15
[0648] Thr Glu Asn Gly Val His Lys Trp Lys Gly Ile Pro Tyr Ala Arg Pro
[0649] 20 25 30
[0650] Pro Val Gly Pro Leu Arg Phe Lys Ala Pro Glu Pro Pro Glu Ala Trp
[0651] 35 40 45
[0652] Glu Asn Glu Leu Asp Ala Thr Ala Tyr Gly Pro Ile Cys Pro Gln Pro
[0653] 50 55 60
[0654] Ser Asp Leu Leu Ser Asp Ser Tyr Asn Glu Leu Pro Arg Gln Ser Glu
[0655] 65 70 75 80
[0656] Asn Cys Leu Tyr Val Asn Val Phe Ala Pro Asp Thr Pro Ser Gln Asn
[0657] 85 90 95
[0658] Leu Pro Val Met Val Trp Ile His Gly Gly Ala Phe Tyr Leu Gly Ala
[0659] 100 105 110
[0660] Gly Ser Glu Pro Leu Phe Asp Gly Ser Arg Leu Ala Ala Gln Gly Glu
[0661] 115 120 125
[0662] Val Ile Val Val Thr Leu Asn Tyr Arg Leu Gly Pro Phe Gly Phe Leu
[0663] 130 135 140
[0664] His Leu Ser Ser Phe Asp Glu Ala Tyr Ser Asp Asn Leu Gly Leu Leu
[0665] 145 150 155 160
[0666] Asp Gln Ala Ala Ala Leu Lys Trp Val Arg Asp Asn Ile Ser Ala Phe
[0667] 165 170 175
[0668] Gly Gly Asp Pro Asp Asn Val Thr Val Phe Gly Glu Ser Ala Gly Gly
[0669] 180 185 190

[0670] Met Ser Ile Ala Ala Leu Leu Ala Met Pro Ala Ala Lys Gly Leu Phe
[0671] 195 200 205
[0672] Gln Lys Ala Ile Met Glu Ser Gly Ala Ser Arg Thr Met Thr Thr Glu
[0673] 210 215 220
[0674] Lys Ala Ala Ser Thr Ala Ala Ala Phe Leu Gln Val Leu Gly Ile Asn
[0675] 225 230 235 240
[0676] Glu Ser Gln Leu Asp Arg Leu His Thr Val Ser Ala Glu Asp Leu Leu
[0677] 245 250 255
[0678] Lys Ala Ala Asp Lys Leu Arg Lys Ala Glu Asn Glu Asn Gln Phe Gln
[0679] 260 265 270
[0680] Arg Leu Phe Gln Pro Ala Leu Asp Pro Lys Thr Leu Pro Ala Glu Pro
[0681] 275 280 285
[0682] Glu Lys Ala Ile Ala Gly Gly Ala Ala Ala Asp Ile Pro Leu Leu Ile
[0683] 290 295 300
[0684] Gly Thr Asn Arg Asp Glu Gly Tyr Met Phe Phe Thr Pro Asp Ser Asp
[0685] 305 310 315 320
[0686] Val His Ser Gln Glu Thr Phe Asn Ala Ala Leu Glu Tyr Leu Leu Glu
[0687] 325 330 335
[0688] Gln Pro Leu Ala Lys Lys Ala Ala Asp Leu Tyr Pro Arg Ser Leu Glu
[0689] 340 345 350
[0690] Ser Gln Ile His Met Met Thr Asp Leu Leu Phe Trp Arg Pro Ala Val
[0691] 355 360 365
[0692] Ala Tyr Ala Ser Ala Gln Ser Gln Tyr Ala Pro Val Trp Met Tyr Arg
[0693] 370 375 380
[0694] Phe Asp Trp His Ser Asp Lys Pro Pro Tyr Asn Lys Ala Phe His Ala
[0695] 385 390 395 400
[0696] Leu Glu Leu Pro Phe Val Phe Gly Asn Leu Asp Gly Leu Glu Arg Met
[0697] 405 410 415
[0698] Ala Lys Ala Glu Val Thr Asp Glu Val Lys Arg Leu Ser His Thr Ile
[0699] 420 425 430
[0700] Gln Ser Ala Trp Ile Thr Phe Ala Lys Thr Gly Asn Pro Ser Thr Glu
[0701] 435 440 445
[0702] Asp Val Lys Trp Pro Ala Tyr His Glu Glu Thr Arg Gln Thr Leu Ile
[0703] 450 455 460
[0704] Leu Asp Ser Glu Ile Thr Ile Glu Asn Asp Pro Glu Ser Glu Lys Arg
[0705] 465 470 475 480
[0706] Gln Lys Leu Phe Pro Ser Lys Gly Glu
[0707] 485

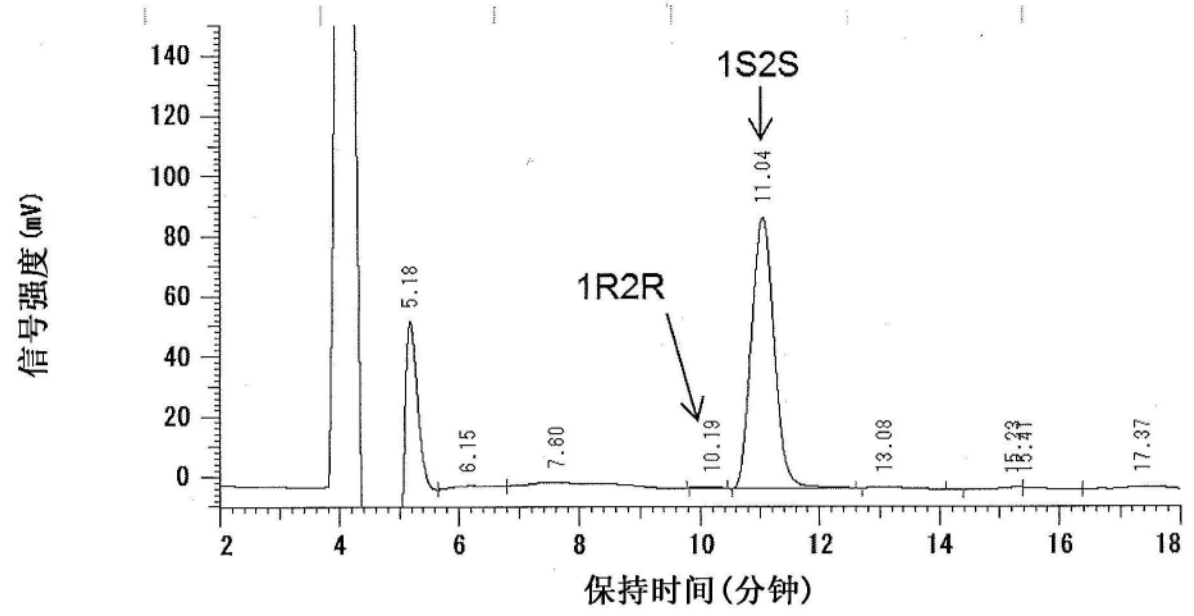

信号强度(mV)
140
120
100
80
60
40
20
0
1S2S
11.04
1R2R
5.18
6.15
7.60
10.19
13.08
15.23
15.41
17.37
2
4
6
8
10
12
14
16
18
保持时间(分钟)

图1

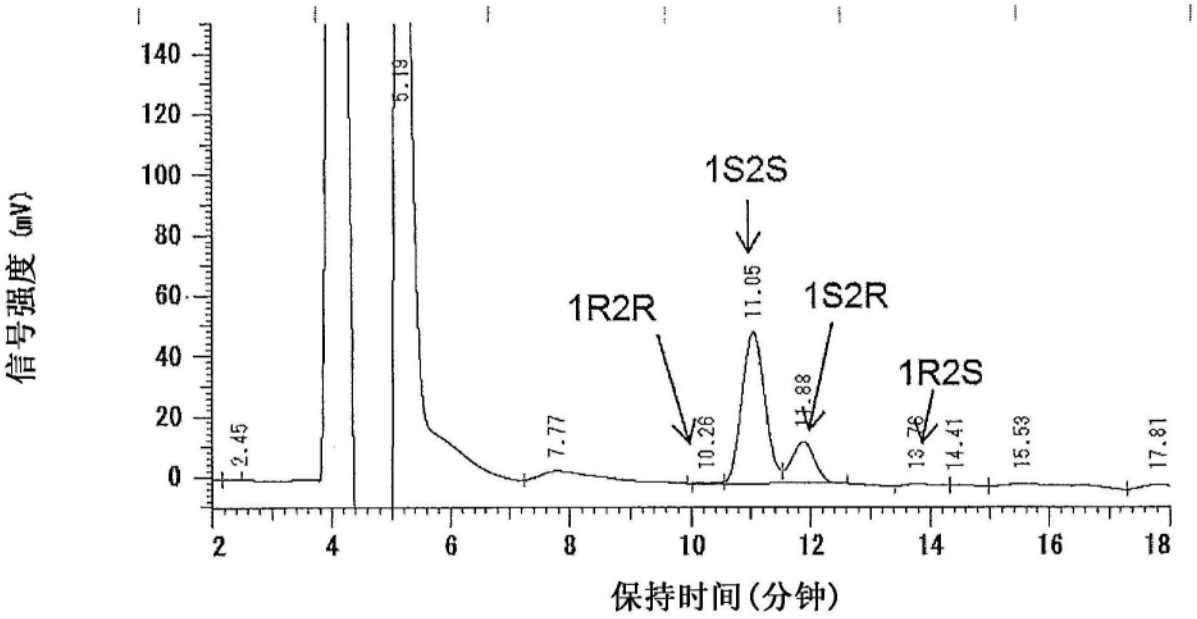

信号强度(mV)
140
120
100
80
60
40
20
0
5.19
1S2S
11.05
1R2R
1S2R
1R2S
2.45
7.77
10.26
11.88
13.26
14.41
15.53
17.81
2
4
6
8
10
12
14
16
18
保持时间(分钟)

图2

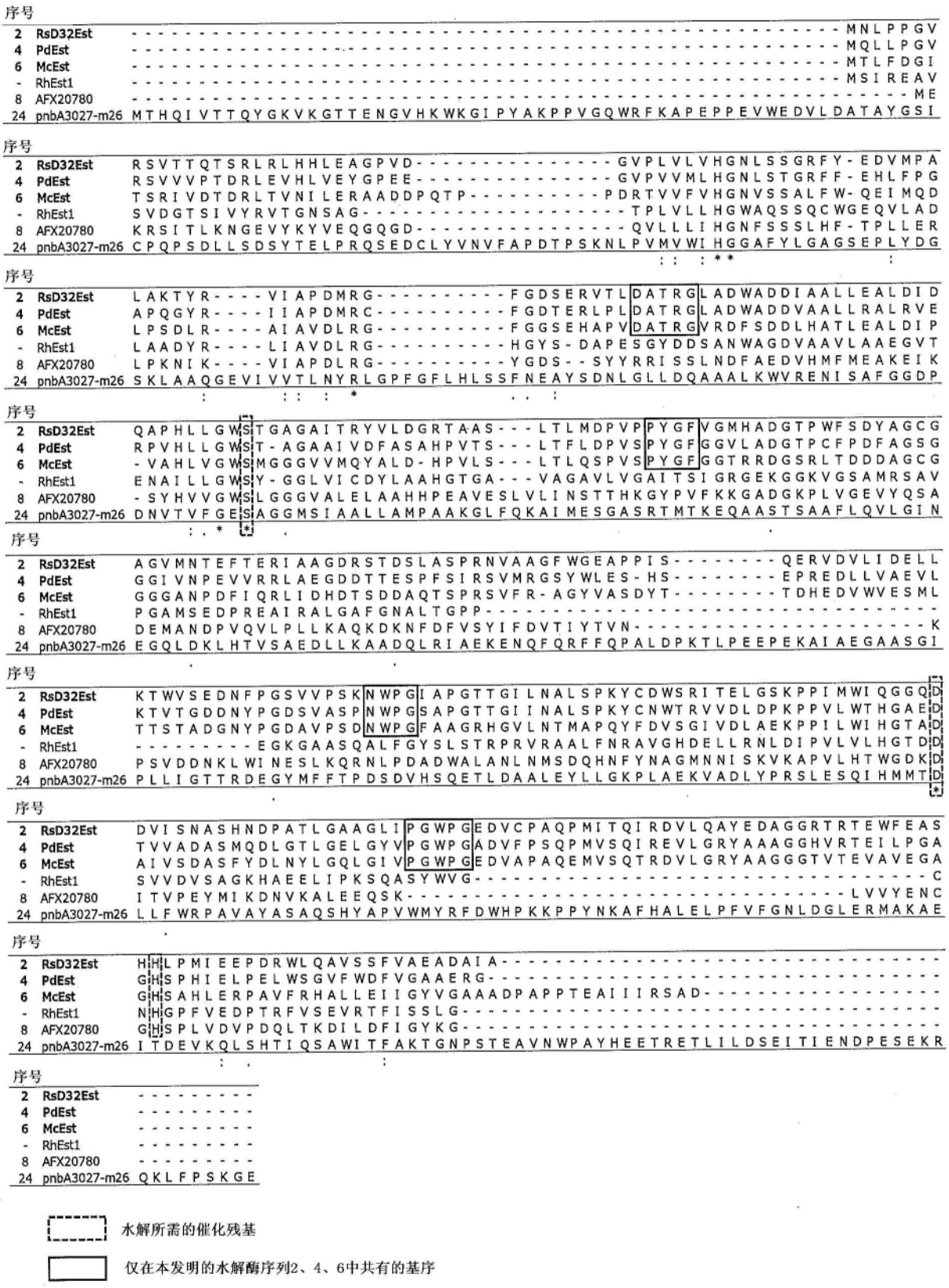

序号
2 RsD32Est
4 PdEst
6 McEst
- RhEst1
8 AFX20780
24 pnbA3027-m26
水解所需的催化残基
仅在本发明的水解酶序列2、4、6中共有的基序

图3