



(12)发明专利

(10)授权公告号 CN 104073507 B

(45)授权公告日 2016.08.17

(21)申请号 201410120626.6

C12P 17/18(2006.01)

(22)申请日 2014.03.27

C12R 1/465(2006.01)

(83)生物保藏信息

CCTCC NO.M2014081 2014.03.13

(56)对比文件

JP S60160888 A,1985.08.22,

Perez Marta 等.Tartrolon D, a

cytotoxic macrodiolide from the marine-derived actinomycete streptomyces sp MDG-04-17-069.《Journal of Natural Products》.2009,第72卷(第12期),第2192-2194页.

Jomon K.等.A new antibiotic, ikarugamycin.《the Journal of antibiotics》.1972,第25卷(第5期),271-280.

(73)专利权人 中国科学院南海海洋研究所

地址 510301 广东省广州市新港西路164号

(72)发明人 张长生 张光涛 张文军 张庆波 朱义广 张海波 李苏梅 马亮

(74)专利代理机构 广州科粤专利商标代理有限公司 44001

代理人 刘明星

审查员 李宁

(51)Int.Cl.

C12N 15/60(2006.01)

C12N 9/88(2006.01)

C12N 1/20(2006.01)

权利要求书1页 说明书10页

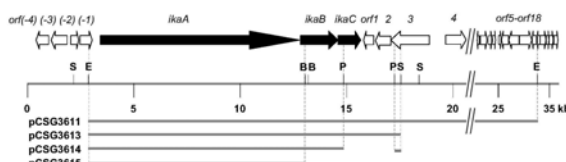
序列表11页 附图12页

(54)发明名称

一种斑鸠霉素的生物合成基因簇及其应用

(57)摘要

本发明公开了一种斑鸠霉素的生物合成基因簇及其应用。斑鸠霉素的生物合成基因簇,其核苷酸序列如SEQ ID NO.1的第3411-15678位所示,包含3个基因:杂合的聚酮/非核糖体聚肽合成酶基因ikaA、FAD依赖的氧化还原酶基因ikaB及类Michael环化反应酶基因ikaC。本发明所提供的斑鸠霉素的生物合成相关的所有基因和蛋白信息,有助于阐释多环tetramate大环内酰胺家族天然产物的生物合成机制,为进一步遗传改造提供理论基础和材料。本发明所提供的基因及蛋白可以用于自然界中挖掘或创制可用于医药卫生及工农业的化合物或基因和蛋白。



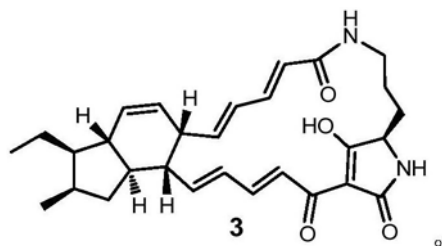
1. 一种斑鸠霉素的生物合成基因簇, 其特征在于, 其核苷酸序列如SEQ ID NO.1的第3411-15678位的碱基序列所示。

2. 权利要求1所述的斑鸠霉素的生物合成基因簇在制备斑鸠霉素中的应用。

3. 一种类Michael环化反应酶基因ikaC, 其特征在于, 其核苷酸序列如SEQ ID NO.1第14614-15678位碱基所示。

4. 一种权利要求3所述的类Michael环化反应酶基因ikaC编码的类Michael反应环化酶IkaC。

5. 权利要求4所述的类Michael反应环化酶IkaC在催化式3所述的化合物形成斑鸠霉素中的应用;



一种斑鸠霉素的生物合成基因簇及其应用

技术领域：

[0001] 本发明属于微生物基因工程领域，具体涉及一种斑鸠霉素的生物合成基因簇及其应用。

背景技术：

[0002] 斑鸠霉素(Ikarugamycin)属于多环tetramate大环内酰胺类化合物，其结构式如图1的式1所示，其具有独特5-6-5三环系统以及抗菌及抗肿瘤细胞等多种生物活性。目前斑鸠霉素已作为抗肿瘤先导化合物引起关注，相应的具体药理机制正在深入研究中。

[0003] 近年来蓬勃发展起来的组合生物合成技术，为改造斑鸠霉素的生产菌Streptomyces sp.ZJ306带来了新的机遇。在阐明了自然界的生物合成途径、理解自然界聚酮化合物天然组合生物合成机理的基础上，人们可以采用组合生物合成技术对抗生素的生物合成基因、调控基因进行体内敲除、突变、置换和重组等操作，不但能够生产“非天然”的天然产物结构类似物，而且还可以提高天然产物的产量，或定向积累所需要的天然产物，为天然产物的发现和药物开发提供化合物结构和生物活性多样化。

[0004] 迄今为止，有关珠江口源微生物中斑鸠霉素生物合成基因簇、类Michael环化反应酶等生物催化制备斑鸠霉素及基因敲除获得斑鸠霉素类似物等，国内外尚未见报道。

发明内容：

[0005] 本发明的目的是提供一种斑鸠霉素的生物合成基因簇。

[0006] 本发明是以珠江口沉积物源放线菌Streptomyces sp.ZJ306中斑鸠霉素为目标分子，在筛选到生物合成基因簇基础上，综合运用分子生物学、微生物学、合成生物学、生物化学及有机化学相结合的技术，阐明了斑鸠霉素的生物合成机制，为利用组合生物合成方法修饰改造斑鸠霉素类似物提供依据，为药物筛选提供化合物实体。

[0007] 本发明的斑鸠霉素生物合成基因簇来源于珠江口沉积物放线菌Streptomyces sp.ZJ306，其特征在于，该生物合成基因簇的核苷酸序列如SEQ ID NO.1的第3411-15678位的碱基序列所示，包括3个基因，具体为：

[0008] (1)负责斑鸠霉素的聚酮-鸟氨酸-聚酮骨架合成的基因，即ikaA，位于核苷酸序列(SEQ ID NO.1，下同)第3411-12785个碱基处，长度为9375个碱基对，编码杂合的聚酮/非核糖体聚肽合成酶，3124个氨基酸；

[0009] (2)负责斑鸠霉素骨架合成修饰的另一个基因，即ikaB，位于核苷酸序列第12782-14614个碱基处，长度为1833个碱基对，编码FAD依赖的氧化还原酶，610个氨基酸；

[0010] (3)负责斑鸠霉素的还原型环化反应的基因，即ikaC，位于核苷酸序列第14614-15678个碱基处，长度1065个碱基对，编码乙醇脱氢酶，354个氨基酸。

[0011] SEQ ID NO.1(序列表)所示序列的第3411-15678位碱基序列的互补序列可根据DNA碱基互补原则随时得到，且第3411-15678位碱基序列或部分核苷酸序列可以通过聚合酶链式反应(PCR)或用合适的限制性内切酶酶切相应DNA或DNA重组技术获得。

[0012] 本发明提供了构建至少包含部分SEQ ID NO.1所示序列的第3411-15678位中DNA片段的重组DNA载体途径。

[0013] 本发明还提供了创制斑鸠霉素生物合成基因被阻断或异源表达等其他基因改造的途径,至少其中之一基因包含SEQ ID NO.1所示序列的第3411-15678位中DNA片段。

[0014] 本发明所提供的核苷酸序列或部分核苷酸序列,可使用PCR探针法或Southern杂交等技术,从其他生物体中筛选获得斑鸠霉素的生物合成基因簇同源基因。

[0015] 本发明提供的核苷酸序列或部分核苷酸序列,可用于Streptomyces sp.ZJ306基因组文库中定位更多的文库质粒。这些文库质粒至少包括本发明中的部分序列,也包含Streptomyces sp.ZJ306基因组中邻近区域未克隆DNA。

[0016] 包含本发明所提供的核苷酸序列或至少部分核苷酸序列的DNA片段,可以被体内外突变等修饰,包括插入、置换、缺失、易错聚合酶链式反应、位点特异性突变、不同序列重组、定向进化等。

[0017] 包含本发明所提供的核苷酸序列或部分核苷酸序列的克隆,可以通过合适表达系统在外源宿主中表达生产相应酶或者提高生物活性化合物产量。这些外源宿主包括大肠杆菌、链霉菌、假单胞菌、芽孢杆菌、酵母及动植物等。

[0018] 本发明提供的氨基酸序列可以用于分离所需蛋白并可用于抗体制备。

[0019] 包含本发明所提供的氨基酸序列或者部分序列的多肽,可能在去除或者替代某些氨基酸之后仍有生物活性甚至新的生物活性,或提高了产量或优化了蛋白动力学特征或其他致力于得到的性质。

[0020] 包含本发明所提供的核苷酸序列或部分核苷酸序列的基因或基因簇可以在异源宿主中表达及揭示它们在宿主代谢中的功能。

[0021] 包含本发明所提供核苷酸序列或部分核苷酸序列的基因或基因簇可以通过DNA重组技术构建重组载体,以获得新型生物合成途径,也可以通过插入、置换、缺失或失活,进而获得其他基于生物合成途径的新结构化合物。

[0022] 本发明提供了斑鸠霉素的生物合成基因簇在制备斑鸠霉素及其类似物中的应用。

[0023] 本发明提供了一种类Michael环化反应酶基因ikaC,其特征在于,其核苷酸序列如SEQ ID NO.1第14614-15678位碱基所示。

[0024] 上述类Michael环化反应酶基因ikaC编码的类Michael反应环化酶IkaC。

[0025] 类Michael反应环化酶IkaC在制备斑鸠霉素中的应用,如催化化合物3形成斑鸠霉素化合物1。

[0026] 本发明提供了类Michael环化反应酶基因缺失突变株中斑鸠霉素类似物的创制应用。

[0027] 总之本发明所提供的包括斑鸠霉素生物合成相关的所有基因和蛋白信息,帮助人们理解多环tetramate大环内酰胺家族天然产物的还原型环化等关键机制,为进一步遗传改造提供材料和理论基础。本发明所提供的基因和蛋白质可以用来寻找和发现可用于医药、卫生或农业的多环tetramate大环内酰胺家族、基因或蛋白。

[0028] 本发明的放线菌Streptomyces sp.ZJ306于2014年3月13日保藏于中国典型培养物保藏中心(CCTCC),地址:中国武汉武汉大学,其保藏编号为:CCTCC NO.M2014081。

附图说明：

[0029] 图1是斑鸠霉素(化合物1)及生物合成中间体化合物3、以及化合物2的化学结构式。

[0030] 图2是Streptomyces sp.ZJ306中斑鸠霉素的生物合成基因簇结构示意图和限制性内切酶谱(S, SacI; E, EcoRI; B, BglIII; P, PvuI.)。

[0031] 图3A是遗传改造Streptomyces sp.ZJ306中斑鸠霉素生物合成基因簇获得的突变株,在发酵培养基中的丁酮粗提物的HPLC分析图:I:野生型Streptomyces sp.ZJ306;II:突变株 Δ ikaA;III:突变株 Δ ikaB;IV:突变株 Δ ikaC;图中阿拉伯数字表示斑鸠霉素及其中间体,其中1、2和3分别代表图1中的式1、式2和式3表示的化合物。

[0032] 图3B是ikaA、ikaAB和ikaABC在S.lividans TK64中表达的工程菌株发酵产物HPLC分析图。如图:V:TK64/pCSG3611(ika cluster);VI:TK64/pCSG3615(ikaA);VII:TK64/pCSG3614(ikaAB);VIII:TK64/pCSG3613(ikaABC);IX:TK64/pSET152。图中阿拉伯数字表示斑鸠霉素及其中间体,其中1、2和3分别代表图1中的式1、式2和式3表示的化合物。

[0033] 图4是orf1、orf2、orf3、orf4和orf5的双交换突变株,在发酵培养基中的丁酮粗提取的HPLC图。(i)野生型Streptomyces sp.ZJ306;(ii) Δ orf1突变株;(iii) Δ orf2突变株;(iv) Δ orf3突变株;(v) Δ orf4突变株;(vi) Δ orf5突变株,其中1、2和3分别代表图1中的式1、式2和式3表示的化合物。

[0034] 图5是化合物1的¹H核磁共振谱结果。

[0035] 图6是化合物1的¹³C和EDPT核磁共振谱结果。

[0036] 图7A和7B分别是化合物2的HRESIMS和MS/MS结果。

[0037] 图8-14分别是化合物3的HR-ESI-MS、¹H、¹³C、EDPT、HSQC、¹H-¹H COSY、HMBC和NOESY的核磁共振谱结果。

[0038] 图15A是纯化的重组IkaC的SDS-PAGE结果。

[0039] 图15B是图1中式3表示的化合物作为底物在IkaC催化作用下生成斑鸠霉素的HPLC分析图。(i)斑鸠霉素标准品1;(ii)标准反应体系:3, IkaC和NADPH;(iii)3, heat-denatured IkaC和NADPH;(iv)3和IkaC。

[0040] 图16提出的斑鸠霉素生物合成途径。

具体实施方式：

[0041] 以下是对本发明的进一步说明,而不是对本发明的限制。

[0042] 1. 克隆分析斑鸠霉素的生物合成基因簇

[0043] 斑鸠霉素(图1的式1所示)属于多环tetramate大环内酰胺家族(PTM)天然产物,由杂合的聚酮/非核糖体聚肽合成酶催化产生,因此针对聚酮/非核糖体聚肽合成酶基因的保守结构域,设计兼并引物C5F/C7R。通过PCR扩增获得一个约0.8kb的PCR片段,连接到T载体中,测序后,进行序列比对,结果表明与PTM的杂合聚酮/非核糖体聚肽合成酶基因具有高度同源性。根据测序结果,设计特异引物306-CF/306-CR,对Streptomyces sp.ZJ306基因组文库约2000个克隆进行筛选,获得10个不同的阳性克隆。通过限制性内切酶酶切分析,确定了10个不同的阳性克隆相对位置。经过末端测序分析,选取阳性克隆pCSG3500使用鸟枪法全

测序,其核苷酸序列如SEQ ID NO.1所示。

[0044] 2.斑鸠霉素的生物合成基因簇中各基因体内功能分析。

[0045] 本发明对斑鸠霉素的生物合成机制进行了研究。对pCSG3500测序结果,运用生物信息学分析,显示包含25个ORF,预测ikaA、ikaB和ikaC负责斑鸠霉素的生物合成(表1和图2)。

[0046] 构建了Streptomyces sp.ZJ306的遗传操作系统,在斑鸠霉素生产菌株进行体内基因敲除,构建双交换突变株(检测引物和基因敲除引物见表2)。

[0047] 对突变株进行发酵及代谢产物HPLC分析,进一步证实ikaA、ikaB和ikaC负责斑鸠霉素的生物合成(图3A),orf1-5不参与斑鸠霉素的生物合成(图4)。采用PCR-targeting技术,敲除了杂合的聚酮/非核糖体肽合成酶基因ikaA的突变株 Δ ikaA丧失了生产斑鸠霉素及相关化合物的能力;敲除了FAD依赖的氧化还原酶基因ikaB的突变株 Δ ikaB不再生产斑鸠霉素,而产生了新化合物2。敲除了类Michael环化反应酶基因ikaC的突变株,也不再生产斑鸠霉素,产生了新化合物3。

[0048] 表1斑鸠霉素的生物合成基因簇的基因及其功能

[0049]

Name	Size	Proposed function	Homolog (origin, accession no.)	identity/ similarity%
orf(-4)	263	hydroxylase	SCO0265 (<i>Streptomyces coelicolor</i> A3(2), CAB55527)	63/75
orf(-3)	251	deacetylase	SSPG_01452 (<i>Streptomyces lividans</i> TK24, EFD65812)	65/77
orf(-2)	123	hypothetical protein	SSFG_00280 (<i>Streptomyces ghanaensis</i> ATCC, ZP_06574565)	41/44
orf(-1)	193	hypothetical protein	DUF305 (<i>Streptomyces pristinaespiralis</i> , WP_005320715)	54/65
ikaA	3124	PKS/NRPS [KS-AT-DH-KR-ACP-C-A-PCP-TE]	HSAF PKS-NRPS (<i>Lysobacter enzymogenes</i> , ABL86391)	66/77
ikaB	594	FAD dependent oxidoreductase	Ox3 (<i>Lysobacter enzymogenes</i> , ACP19698)	68/81
ikaC	354	zinc-binding alcohol dehydrogenase	Ox4 (<i>Lysobacter enzymogenes</i> , ACP19697)	63/77
orf1	156	GCNS-related N-acetyltransferase	SPW_6416 (<i>Streptomyces</i> sp. W007, EHM25172)	71/82
orf2	242	predicted protein unkwon function	SSMG_08119 (<i>Streptomyces</i> sp. AA4, EFL12448)	46/55
orf3	602	asparagine synthase	SPW_6415 (<i>Streptomyces</i> sp. W007, EHM25171)	52/65
orf4	435	transcriptional regulator	KSE_58920 (<i>Kitatospora setae</i> KM-6054, BAJ31662)	45/59
orf5	131	ABC transporter	SSRG_03148 (<i>Streptomyces griseoflavus</i> , EFL40344)	87/93
orf6	454	conserved hypothetical protein	SCAB_76841 (<i>Streptomyces scabiei</i> 87.22, CBG74656)	93/96
orf7	247	hypothetical protein	Sfla_6619 (<i>Streptomyces flavogriseus</i> ATCC 33331, YP_004909524)	81/91
orf8	191	putative membrane transport protein	SCAB_76871 (<i>Streptomyces scabiei</i> 87.22, CBG74659)	93/96
orf9	160	putative acetyltransferase	SSOG_03986 (<i>Streptomyces hygrosopicus</i> ATCC 53653, EFL24272)	61/75
orf10	466	cytochrome P450	MCAG_03325 (<i>Micromonospora</i> sp. ATCC 39149, EEP72998)	48/65
orf11	672	putative alpha-glucosidase	Tfu_1613 (<i>Thermobifida fusca</i> YX, AAZ55648)	35/48
orf12	805	phosphoketolase	SSEG_09767 (<i>Streptomyces svuceus</i> ATCC 29083, EDY58979)	80/87
orf13	292	helix-turn-helix domain-containing protein	SSCG_04599 (<i>Streptomyces clavuligerus</i> ATCC 27064, EDY51621)	55/73
orf14	184	membrane protein	SSOG_00820 (<i>Streptomyces hygrosopicus</i> ATCC 53653, EFL21108)	70/78
orf15	382	NAD-dependent epimerase/dehydratase	SCLAV_p1163 (<i>Streptomyces clavuligerus</i> ATCC 27064, EFG04649)	54/60
orf16	184	RNA polymerase, sigma-24 subunit	Amir_6890 (<i>Actinosynnema mirum</i> DSM 43827, ACU40685)	62/75
orf17	290	unknown function	hypothetical protein (<i>Streptomyces</i> sp. AA0539, WP_019436027)	46/55
orf18	283	hypothetical protein	DUF3152 (<i>Streptomyces</i> sp. AA0539, WP_019436028)	81/86

[0050] a氨基酸数目;b括号内包含同源蛋白GenBank登录号;c一致性/相似性(identity/similarity)

[0051] 表2:本发明中使用的引物

[0052]

引物名称	序列
基因组文库筛选所用引物	
C5F	5'- GCSCTSTGGTTCCTYAAGCA -3'
C7R	5'- GGCARCGGGTTKACRAAGTA -3'
306-CF	5'- CGAACTTCCTGCTGGTGGAG -3'
306-CR	5'- GCTCAGCTCCGCGTCCAGCG -3'
构建基因敲除突变株所用引物	

[0053]

ikaAKF	5'- ATCGAGGAGTACCGCAAGCCGTACGACCTGGCGACCGATATTCGGGGATCCGTCGACC -3'
ikaAKR	5'- GCTCAGCAGCACCATGAAGACGGTGACGTTGTGCTCCCGTGTAGGCTGGAGCTGCTTC -3'
ikaBKF	5'- ATCGACGAGGTGGGCACGATCTCCCCGGCTACGTCTCGATTCCGGGGATCCGTCGACC -3'
ikaBKR	5'- CTGCGTGACCTGCCGGCCCGAGGCCACCGCCCGGATCAGTGTAGGCTGGAGCTGCTTC -3'
ikaCKF	5'- CCCGCCTCGCTGGCGGACGATCAGATGCTGCTGCGCACCATTCGGGGATCCGTCGACC -3'
ikaCKR	5'- CTTGCGGATCAGCCCGCCACCTCCTCGTGCAGCGCGTCTGTAGGCTGGAGCTGCTTC -3'
Orf1KF	5'- ATGATCGGATACGGCCGCCAGGCCCCCGAAAGTCTCGACATTCGGGGATCCGTCGACC -3'
Orf1KR	5'- GAAGTCCGATGCGAGGGCGTCCATGATGATCTCGTCGCATGTAGGCTGGAGCTGCTTC -3'
Orf2KF	5'- CCCGCCGCTGCGCTGCGGTGGTGTGCCGACCCGTACCATTCCGGGGATCCGTCGACC -3'
Orf2KR	5'- GACCGCCTGCACCGCGGCGGTGGCTTCGGCCTCGGTAGCTGTAGGCTGGAGCTGCTTC -3'
Orf3KF	5'- GTGCTGGGGCGTTGCGGTGCGTCCGAAGCCGAGCTGGGCATTCCGGGGATCCGTCGACC -3'
Orf3KR	5'- CTCTGGGTCCAGGACCGTCCGGTACTCGACCGCGCCTGTGTAGGCTGGAGCTGCTTC -3'
Orf4KF	5'- GACGCGCTGGACGTCTCGCCCGCACACGAGCCGCCCGATTCCGGGGATCCGTCGACC -3'
Orf4KR	5'- GAGGGACCACAGCACGCCACCTGTAGAGGGTGACAGTTGTAGGCTGGAGCTGCTTC -3'
Orf5KF	5'- ACCACCGCCGTCCGCCAGGCCCTGGCCGACCAGGGGTTTCATTCCGGGGATCCGTCGACC -3'
Orf5KR	5'- GACGGGTTTGAGGGAATCCAGGCCGGTGAGGGTGACCATTGTAGGCTGGAGCTGCTTC -3'
检测基因敲除突变株	
ikaAF	5'- CTGGTGGAGGGGCAGGCGGT -3'
ikaAR	5'- CCGAACTCCTCTGGGTGCG -3'
ikaBF	5'- GCGGCCGGTATCTCAACGAC -3'
ikaBR	5'- CGCTCCTCGACGATGTGACG -3'
ikaCF	5'- CCGTCACATCGTCGAGGAGC -3'
ikaCR	5'- GCGTCGACCATCTCCTCGGT -3'
Orf1F	5'- CGCAAGGGCGAGCTGGTCTA -3'
Orf1R	5'- ACCATCGGAGACCGCCATCA -3'
Orf2F	5'- GCAGCCACAGAATCGGGTCA -3'
Orf2R	5'- TCTGCTCCGCACGCAACTGG -3'
Orf3F	5'- TGATGGTGCTGGGGCGTTGC -3'
Orf3R	5'- GCAGCAGATCGGCGAGGTGA -3'

[0054]

Orf4F	5'-CTCGCCATCCCACAGAACCT-3'
Orf4R	5'-GAACTGTGCCGGGTTGAGAG-3'
Orf5F	5'-GGGGTGCTCTTCCCGGTTGT-3'
Orf5R	5'-ACGACTGCCGCTCCTCCACA-3'
<i>for construction of ikaC expression vectors</i>	
EikaCF	5'-GCGTCATATGCAGGTACACACATGGGT-3' (<i>NdeI</i>)
EikaCR	5'-GGCGGATCCCTACAGCGCGACCAGGACCT-3' (<i>BamHI</i>)

[0055] 3.斑鸠霉素的生物合成基因簇在模式菌株链霉菌*S.lividans* TK64的异源组合表达。

[0056] 本发明使用分子生物学技术,构建了4个基于质粒pSET152的载体pCSG3611, pCSG3613-3615(图2)。pCSG3611的插入片段包含ikaA-orf13,位于核苷酸序列SEQ ID NO.1第2866-32536个碱基处。pCSG3613的插入片段包含ikaA-ikaC,位于核苷酸序列SEQ ID NO.1第2866-17556个碱基处。pCSG3614的插入片段包括ikaAB,位于核苷酸序列SEQ ID NO.1第2866-14792个碱基处。pCSG3615的插入片段包含ikaA表达盒,位于核苷酸序列SEQ ID NO.1第2866-13065个碱基处。将上述片段分别插入质粒pSET152中,分别获得载体pCSG3611, pCSG3613-3615。然后再将载体pCSG3611, pCSG3613-3615分别转化进入*S.lividans* TK64菌中,分别得到TK64/pCSG3611、TK64/pCSG3613、TK64/pCSG3614、TK64/pCSG3615的链霉菌TK64转化子,以转入空载体pSET152作为对照(TK64/pSET152)。

[0057] pCSG3611的链霉菌TK64转化子(TK64/pCSG3611),成功异源表达了化合物1(斑鸠霉素)及中间体化合物3。该结果表明ikaA上游的orf(-1)-orf(-4)不是合成斑鸠霉素的必要基因。携带完整的杂合聚酮/非核糖体肽合成酶基因的ikaA的pCSG3615链霉菌TK64转化子(TK64/pCSG3615),表达了少量的化合物2。携带完整ikaAB的pCSG3614的链霉菌TK64转化子(TK64/pCSG3614)成功合成了化合物3和少量化合物2。携带完整ikaABC的pCSG3613的链霉菌TK64转化子(TK64/pCSG3613)成功合成了化合物1(斑鸠霉素)和3。

[0058] 由此充分证实ikaABC三个基因组合是合成斑鸠霉素的最小基因簇,并成功实现了ikaABC三个基因组合,在异源细胞工厂TK64中生产斑鸠霉素及其结构类似物,提供了创制斑鸠霉素及其中间体化合物的新策略(图3B)。

[0059] 4.斑鸠霉素及其中间体的分离鉴定:

[0060] 将培养皿上活化的*Streptomyces* sp.ZJ306等量分别每瓶含有50mL AM6培养基的250mL的锥形培养瓶中,28℃,200r·min⁻¹,培养48h,制得种子液。将种子液以10%的接种量转接至AM6发酵培养基(置于1000mL的锥形培养瓶,每瓶含有200ml培养基)中,28℃,200r·min⁻¹,振荡培养96h。发酵产物使用等体积丁酮萃取,静置分层。将丁酮萃取液与水相分离,用旋转蒸发器将丁酮蒸干,获得发酵粗提物。采用正相硅胶柱层析、反相中压液相色谱及凝胶层析等技术纯化获得1和3,并采用核磁共振谱法鉴定结构(图5-6和8-14)。将化合物1的氢谱碳谱与文献[Ito S&Hirata Y(1977)The structure of ikarugamycin,an acyltetramic acid antibiotic possessing a unique as-hydrindacene

skeleton. Bull Chem Soc Jpn 50(7):1813-1820.]比较,发现化合物1为已知化合物斑鸠霉素。分析化合物3的氢谱碳谱发现它与斑鸠霉素(化合物1)相似,不同之处在于化合物1中C-4, C-5, C-15, 和C-16的亚甲基或次甲基在化合物3中由这些烯键的碳所取代[$\delta_{\text{H}}6.35(\text{H}-4)$; $\delta_{\text{C}}132.3(\text{CH}, \text{C}-4)$], [$\delta_{\text{H}}5.76(\text{H}-5)$; $\delta_{\text{C}}141.3(\text{CH}, \text{C}-5)$], [$\delta_{\text{H}}5.88(\text{H}-15)$; $\delta_{\text{C}}141.1(\text{CH}, \text{C}-15)$], [$\delta_{\text{H}}6.34(\text{H}-16)$; $\delta_{\text{C}}133.7(\text{CH}, \text{C}-16)$], 这说明化合物3与化合物1的区别是化合物3失去了五元环。化合物1和化合物3的具体结构如图1中的式1和式2所示。

[0061] 5. 类Michael环化反应酶IkaC体外酶催化体系的构建。

[0062] 使用PCR技术、酶切连接等分子生物学方法,将ikaC基因克隆到大肠杆菌表达载体pET28a。将IkaC的表达载体转化至E. coli BL21(DE3)中诱导表达后,收集细胞,超声破碎获取粗酶液;镍柱亲和层析纯化、SDS-PAGE和Bradford法蛋白浓度测定等,最终获得高纯度IkaC重组酶(图15A)。每100u1 IkaC酶催化体系反应体系包含200um的化合物3、2mMNADPH、10ug纯化的IkaC重组酶。该反应于pH8.0的磷酸盐缓冲液中28度温育10小时完成。利用该体系成功将前体化合物3催化成斑鸠霉素(化合物1)(图15B)。

[0063] 以下进一步提供实施实例,这些实施实例有助于理解本发明,仅用作说明而不限制本发明的应用范围。

[0064] 实施例1

[0065] 斑鸠霉素生产菌Streptomyces sp. ZJ306的基因组DNA的提取:

[0066] 将新鲜Streptomyces sp. ZJ306的菌丝体按照5%的接种量接种于50mL的TSB培养基(胰蛋白胨17g,植物蛋白胨3g,氯化钠5g,磷酸氢二钾2.5g,葡萄糖2.5g,加水至1L, pH7.2-7.4)中,28-30°C,振荡培养约3-4天,4000rpm离心10分钟收集菌丝体。菌丝体用STE溶液(NaCl75mM, EDTA25mM, Tris-Cl20mM)洗涤两次,向洗涤后的菌丝体中加入30mL STE溶液和终浓度3mg/mL的溶菌酶,涡旋均匀,37°C温浴3小时,加入至终浓度0.1-0.2mg/mL的蛋白酶K,混匀,37°C温浴10分钟,加入至终浓度1-2%的SDS,混匀,放入55°C水浴约1小时,期间颠倒数次。加入等体积的酚-氯仿-异戊醇(V/V/V=25:24:1),混合均匀,置于冰上冷却30分钟。12000rpm,4°C离心10分钟,然后用剪过的大口径枪头小心吸取上清到新的离心管中,用同样的方法反复处理3次,然后用等体积的氯仿洗涤两次,12000rpm,4°C离心10分钟。用剪过的大口径枪头将水相吸出转移至新的离心管,加入1/10体积3mol/L NaAc(pH5.2),混匀后再加入等体积的异丙醇,混匀后冰上放置,沉淀DNA。小心地用玻璃棒将DNA纤维团转移至新的离心管中,用70%乙醇洗涤两次,将液体倾出,在37°C下略微烘干,加5mL TE溶解,并加入3-5U的RNA酶,由此得到Streptomyces sp. ZJ306的基因组DNA。

[0067] 实施例2

[0068] 斑鸠霉素生产菌Streptomyces sp. ZJ306全基因组文库的建立:

[0069] 首先通过一系列的稀释实验来确定限制性核酸内切酶Sau3A I的用量,在20 μ L体系中,含有17 μ L的基因组DNA,2 μ L的10 \times 反应缓冲液和1 μ L不同稀释度的Sau3A I,其终止反应为4 μ L0.5mol/L EDTA和合适的上样缓冲液。通过摸索确定了0.025-0.05U的酶活单位比较合适。在此基础上通过大量部分酶切得到略大于40kb的基因组DNA片段,用去磷酸化酶进行去磷酸化处理。

[0070] 用于构建文库的载体SuperCos 1质粒先用限制性核酸内切酶Xba I从两个cos序列中间切开,然后进行去磷酸化处理,再从多克隆位点处用限制性核酸内切酶Bam HI切开,

获得两个臂。处理后的载体与之前制备的部分酶切的约40kb的基因组DNA片段连接过夜，连接体系为10 μ L，含有1.25 μ g制备的基因组DNA和0.5 μ g处理后的SuperCos1质粒，1 μ L的10 \times Buffer，0.3U的连接酶。连接产物与65 $^{\circ}$ C处理15分钟，使连接酶失活。从-80 $^{\circ}$ C冰箱中取出一管包装混合物(50 μ L)置于冰上，将包装混合物在指间迅速融化，小心吸取一半包装混合物(25 μ L)至一个新的离心管中，加入10 μ L热处理后的连接产物，其余包装混合物于-80 $^{\circ}$ C保存。小心混匀，30 $^{\circ}$ C温浴90分钟，加入另外一半包装混合物(25 μ L)，30 $^{\circ}$ C温浴继续90分钟。加入500 μ L噬菌体稀释缓冲液(100mmol/L NaCl, 10mmol/L MgCl₂, 10mmol/L pH8.3Tris-HCl)，再加入25 μ L氯仿，轻轻混匀，于4 $^{\circ}$ C保存，由此得到包装液。

[0071] 将冻存于-80 $^{\circ}$ C的菌株*E. coli* LM392MP涂布在LB培养基上复苏。包装反应前一天，挑取单克隆接种于LB培养基中(添加0.2%麦芽糖和10mM MgSO₄)，37 $^{\circ}$ C振荡培养过夜，包装反应当天，取5mL过夜培养的菌液加入到50mL新鲜的LB培养基中(添加0.2%麦芽糖和10mM MgSO₄)，37 $^{\circ}$ C，200rpm振荡至培养物OD₆₀₀达到0.8-1时，4 $^{\circ}$ C保存备用。取100 μ L如上处理的宿主菌液和100 μ L适度稀释的包装液轻轻混匀，于37 $^{\circ}$ C温浴15分钟，然后涂布于含有氨苄青霉素和卡那霉素的LB平板上，37 $^{\circ}$ C培养过夜。将长出来的单个克隆子，用无菌牙签点种于25块含合适抗生素的LB的96孔板上，37 $^{\circ}$ C培养过夜，加入终浓度为20%的甘油，混合均匀，置于-80 $^{\circ}$ C保存，得到*Streptomyces* sp. ZJ306全基因组文库。根据测序结果，设计特异引物306-CF/306-CR，对*Streptomyces* sp. ZJ306基因组文库约2000个克隆进行筛选，获得10个不同的阳性克隆。通过限制性内切酶酶切分析，确定了10个不同的阳性克隆相对位置。经过末端测序分析，选取阳性克隆pCSG3500使用鸟枪法全测序。阳性克隆pCSG3500的核苷酸序列如SEQ ID NO.1所示。

[0072] 实施例3

[0073] 斑鸠霉素生产菌*Streptomyces* sp. ZJ306遗传转移系统的建立及基因中断突变菌株的获得，以敲除类Michael环化反应酶基因*ikaC*为例：

[0074] 利用PCR-targeting的方法获得体外敲除突变株。根据获得的斑鸠霉素生物合成基因簇序列，参照文献报道的PCR-targeting系统，设计一对*ikaC*基因的敲除引物，引物序列见于表2中的*ikaC*敲除引物*ikaCKF/ikaCKR*。然后参照PCR-targeting的方法构建体外敲除质粒然后转入到接合转移的供体菌中。具体步骤如下：(1)将cosmid质粒pCSG3500转入大肠杆菌*E. coli* BW25113/pIJ790中获得*E. coli* BW25113/pIJ790/pCSG3500，用10mmol/L的L-阿拉伯糖诱导 λ /red重组系统表达，并将其制备成为电转感受态细胞待用。(2)用内切酶EcoR I和Hind III酶切质粒pIJ773，回收其中约1.4kb含有转移原点和阿泊拉霉素抗性基因的DNA片段，以此作为PCR模板，用引物*ikaCKF/ikaCKR*通过PCR扩增出1.4kb的PCR产物，50 μ L的PCR反应体系：高保真DNA聚合酶3U，10 \times Buffer 5 μ L，dNTPs 0.5mmol/L，DMSO 2.5 μ L，引物各0.5 μ mol/L，DNA模板约1ng，加水补至50 μ L。PCR反应条件为：预变性94 $^{\circ}$ C 5min；扩增循环为94 $^{\circ}$ C变性45s，58 $^{\circ}$ C退火45s，72 $^{\circ}$ C延伸90s，30个循环；最后72 $^{\circ}$ C延伸10min。将1.4kb的PCR产物回收纯化待用。(3)将PCR产物电转入(1)步骤中制备的感受态细胞使其发生重组，涂布于LB筛选平板(含100 μ g/mL氨苄青霉素，50 μ g/mL卡那霉素，50 μ g/mL阿泊拉霉素)上，37 $^{\circ}$ C过夜培养。从平板上挑出阳性单克隆，抽提质粒，重组质粒命名为pCSG3553，该质粒中的*ikaC*基因的部分片段被转移原点和阿泊拉霉素抗性基因取代。(4)将构建好的重组突变质粒pCSG3553转化到*E. coli* ET12567/pUZ8002中，构建成*E. coli* ET12567/pUZ8002/pCSG3553，

作为接合转移的供体菌。

[0075] 放线菌ZJ306在38#培养基平板中划线培养3-5天,长出的孢子用无菌棉签收集于含3%海盐的TSB培养基中,涡旋振荡,使孢子分散。过滤分离菌丝体和孢子,孢子悬浮于5mL含3%海盐的TSB培养基中,50℃热激10min,然后于28℃萌发2小时,作为接合转移的受体菌。供体菌E.coli ET12567/pUZ8002/pCSG3553在50mL含50μg/mL卡那霉素,25μg/mL氯霉素和50μg/mL阿泊拉霉素的LB液体培养基中于37℃生长至OD₆₀₀值约为0.8时,离心收集菌体(4000rpm,10min),用LB清洗菌体3遍,悬浮于300μL LB培养基中,作为接合转移的供体菌。取上述受体菌400μL和供体菌100μL混合均匀,涂布于不含任何抗生素的ISP4固体培养基上,吹干后,于28℃培养18-20h。然后将平板取出,用含有抗生素的水覆盖平板,其终浓度为35μg/mL阿泊拉霉素和50μg/mL甲氧苄氨嘧啶,吹干后,置于28℃培养箱中,培养5-7天后观察。

[0076] 当接合转移平板上长出菌落后,用无菌牙签将其转接到含有35μg/mL阿泊拉霉素和50μg/mL甲氧苄氨嘧啶的ISP4平板上,28℃培养3天后,抽取各个突变株的基因组DNA,利用检测引物ikaCF/R(引物序列见于表2)通过PCR检测获得阳性克隆,即获得ikaC敲除双交换突变菌株ΔikaC。

[0077] 其他各个基因的敲除引物和突变株检测引物参见表2,参照上述方法,利用PCR-targeting技术获得各个基因敲除的突变菌株。

[0078] 具体如下:

[0079] 敲除了ikaA-杂合聚酮/非核糖体肽合成酶基因,获得突变株ΔikaA;敲除了ikaB-FAD依赖的氧化还原酶基因,获得突变株ΔikaB;此外分别敲除了orf1基因、orf2基因、orf3基因、orf4基因和orf5基因,分别获得突变株Δorf1、Δorf2、Δorf3、Δorf4、Δorf5。

[0080] 实施例4

[0081] 斑鸠霉素及其中间体的生物发酵与检测:

[0082] 将放线菌ZJ306野生菌或突变株(ΔikaA、ΔikaB、ΔikaC等)活化后,按1%的接种量分别加入到250ml摇瓶的50ml AM6液体培养基(可溶性淀粉20g/L,葡萄糖10g/L,蛋白胨5g/L,酵母粉5g/L,碳酸钙5g/L,3g/L海盐)中,于28℃,200r.p.m,培养5d后,加入等体积丁酮,超声破碎细胞10min,然后静置分层。将上层丁酮萃取液与水相分离,使用旋转蒸发仪蒸干。粗提物使用二甲基亚砜溶解,进行HPLC检测,分析条件为:菲罗门Luna C18(5μm,150×4.6mm)反相柱,流动相A相为10%乙腈,含0.1%三氟乙酸(TFA),流动相B相为90%乙腈;流速为1ml/min。检测波长为360nm和245nm。HPLC程序:0-15min,5%B到100%B(线性梯度),15-22min100%B,22-24min,100%B to5%B,24-30min5%B。

序列表

- <110> 中国科学院南海海洋研究所
- <120> 一种斑鸠霉素的生物合成基因簇及其应用
- <160> 1
- <210> 1
- <211> 36720
- <212> DNA
- <213> *Streptomyces* sp. ZJ306
- <400> 1

[0001]

gatcatgaac	ccgtcgacgg	acctgagcca	cctcgtgtgg	cgcttcagca	gttacagcgg	60
aggagagggc	gagtgcgtcg	aggtegccga	tgacataccc	ggggtcatac	ccgtgcggga	120
cagcaagcgg	cccgcggggc	ccgtgctggt	gttccggcac	gaggcgtggc	gggccttcgt	180
cggcaccttg	aggtgatacc	tcaggtctcc	ggggtcagca	ggacgaagtc	cgcgccttg	240
gggtcggcgc	agaccgacag	gcggccgacg	ccttccgctg	tttcggggcc	catctgcacc	300
gtgcccgcgt	tccgggtgac	cgtggccgtg	gccgcgtcgc	agtccgctgt	ggtgaagaca	360
gggtgccagt	acgccgcgcc	gttcgtcagg	gtcagggcgt	cggggggaag	ttgcatcagc	420
ccgccctgca	tgcggctctc	gtccgccgtg	gccggggtga	tcagggagta	cgtgccactg	480
tcgcccggca	ggggcatctc	ggagtactgc	cagccgaaga	ccgccgtgta	gaaggtcttc	540
gccgccgcgc	cgccccgggt	gtacagctcg	gtccaggaca	gggagccggg	gacgtccacc	600
gcctccacgc	cccggctgctc	gcggggttcg	aagaccgcga	accggccgcc	ctggggatcg	660
gtgaactggg	ccatgcggtg	ttcgtcgtcc	atgccggtgg	ggacgacccg	gacgggtgcc	720
cccgcgttcg	tcacggcgtc	ggcgggtggc	tccgcctcgg	tggtggcga	gtagagcatc	780
cacgcggagc	gggcgccctc	ctccgtgagg	cggcccaggc	cggccaccgt	gcggccctcg	840
ctctgcacga	ccccgtaccc	cgagccctcc	ggaccatgg	gccggaaggt	ccagcccagg	900
accgcgccgt	agaaggcggc	ggtegtctcc	gggtccgggg	cgccgaggtc	gagccagcag	960
ggggagccgg	ggacgaagtc	ggtggtcagc	acggagtcc	cccttgctac	gtcagcctca	1020
tctgccagcg	tgacaccggc	cgccccggc	cgccaccgct	cggggcggct	cacagccggc	1080
cggggggcgc	tgccaccggg	tcggtcagcg	gtagaccttg	cccggttcgg	gtacgccggg	1140
cgccatcagc	tgggagacgg	tgacgaaggt	gtatccgtcg	gccttcagtg	cctcgatgat	1200
gccggggacg	gcggggacgg	tgccctcgta	gatgtcgtgc	agcaggatga	tgccgtcgcg	1260
gtcgacctgg	tccaggacgc	gttcgggtgat	caggtccgag	tcgttggtcg	agtagtcttt	1320
ggcgtgacg	gaccacagga	cctgcgccag	gcccagtcc	tcgcagacct	gcggccctcg	1380
gtcgtcgggt	cggccctgcg	gcgggcgcat	gagcgtcggg	gtctgccccg	tgagttcctc	1440
gatggcctcg	ttggtgcgca	gcagctggtc	gcgcacctcg	tcccgggtcca	cctcgggtgag	1500
tcgggggtgg	gtccaggtgt	ggttgccag	ctcgtgccct	tctcggccg	agcgcaccac	1560
gagttcgggg	taggtgtcga	tgtgtctctt	gccgagcagg	aagagcgtcg	cggggacctc	1620
ttcctcggcc	aggatgtcca	gcagttcggc	cgtgtacggg	cccgggccgg	cgtcgaaggt	1680
gagggcgacg	cacttggcct	cgcggcagtc	caccgagccg	aacgcgccgc	cggacttctc	1740
gtcggccttg	tcgcgcaagc	tctcgggcga	tatccgctcc	gccgagcagc	cggtcagggc	1800
gagcgtcacg	gtcagggta	gggccgccgc	gacggcggtc	gcgggcttgt	tgagtcgcat	1860
ggagcgggaa	tccccccagg	tggaacggtc	gtttctctcc	ccacgccccg	gccggccgga	1920
cggtgaaaca	actatacgtt	ccgtgtatag	cggccccctg	ggtgggacct	ctcacggctc	1980
cgccggaccc	ccaccgttac	ccacgggggg	taccctctga	ggatggcggg	cgtgaaccgg	2040
gtccgtggcg	cgctgctgat	cctggcggta	ctcggccggg	tgttcgccat	gcacgcgctc	2100
gcgccggggg	tcgtcatctc	ccccgcgat	cacacggagg	ccgccgcgcc	cgcgcacgac	2160
gccggtggcg	cggcgcactc	cgaccacgcg	gaccgggact	gtgtggcggg	tggggtcggc	2220
ggggcggccg	ccccggtcgg	tgccggggcg	ggggtgacca	gtggggtcac	ggtgccggcc	2280
gtcgtcggcg	cgcgggagct	cgccgggccc	gccccgcct	gggggcagcc	ccggccggag	2340
gctggggcg	gcgcgccgcc	ttcgtgagt	tctctgcagt	cgctgcgtat	atagggcgt	2400
cgtgcaccga	cgcgtccgac	acctcagcac	tgacagaaa	gctgtttctc	caccatgacc	2460
tcgtaccgtt	tcgcggcgcg	ggcgtcagc	gccgccgccg	tcacctccgc	cttgcgtgtc	2520
accggcgct	gcgggagcga	ttccgatgag	ccgggcggac	attcccagca	gcggcgcgca	2580
ccggtttccg	agcccttcaa	cgcgccggat	ctgccttcg	tccagggcat	gatccccgac	2640
cacgagcagg	ccatcgagat	gtccgaactc	gcggccgggc	gcgccggatc	ggacgcgggtg	2700
cgggatctgg	ccgccgtcat	cgagcaggcc	caggggcccg	agatcgacac	actccggtca	2760
tggtctctcg	actgggacgc	cgagccggcg	gaccacggcg	ggcacgacat	gaccggcatg	2820
atgtcggacg	aggacatggc	ggagctggaa	ggctcctccg	gcgcggaatt	cgacgccgct	2880
ttcctgcgcc	tgatgatcgc	gcaccacgag	ggcgcggtcg	agatggcgcg	ggagctggag	2940
cggcaggggt	ccttcgcgcc	cgctctggag	atggcgcagg	acatcatctc	ctcccagttc	3000

	gctgaaatcg	atcaaatgac	cgaacttctg	aactgatatt	cgcacccgag	gggtgggcca	3060
	tccgaaatcc	ggaacgccc	ccccctttgt	atgaccggta	gcgcgcaaag	gcgctccgcg	3120
	cgccccccgc	gtaaccctat	tggcctattc	gatcttcata	aggcaggctt	caaccgcgca	3180
	gccgcaaac	acaatcccc	taaacaccta	cggggctgtc	atcggcctgc	gctgtccac	3240
	gagggggggt	tctcgttccg	cgttccgcca	tgccgttcgt	cagcggcatt	cgcatatctg	3300
	ttcttcagac	ttctctgtcc	ttaccgcgge	gctgtccccc	aacgggctga	ccaggattcc	3360
	cgctcgaatc	gectgatccg	cccacgcgcc	atactcaag	gattctcaag	atggattcca	3420
	tgcatacccc	tgcccccgtc	cccgtaccgc	aagtccccac	gcccgtcccg	tcccaggacg	3480
	acgcgttcgc	catcgtcggc	atcggctgcc	gectgcccgg	cggcgcagc	gactaccgga	3540
	ccttctggcg	caacctcctc	gacggcaagg	actgcatcac	cgacaccccc	gccgaccgct	3600
	acgacacccg	caccctgggc	agcggcgaca	aggccaagcc	cgcccggtg	gtcggcgagc	3660
	gcggtggata	catcgacggc	ttcgacgagt	tcgaccccgc	cttcttcggc	atcagcccgc	3720
	gcgaggccga	gcacatggac	ccccagcagc	ggaagctcct	ggaggtcgcc	tgggaggcgc	3780
	tggaggacgg	cgccctcaag	cccgcgagc	tgcccgccac	cgatgtcggg	gtgtacgtcg	3840
	ggcggttac	cctcgactac	aagatcctgc	agttcgccga	cctcggcttc	gagaccctgg	3900
	ccgcgcacac	cgccaccggc	accatgatga	cgatgggtgc	caaccggatc	tcgtactgct	3960
	tgacttccg	cgccccctcg	gtctccgctg	acaccgcgtg	cagcggctcc	ctggctcgccg	4020
	tccacttggc	ctgccagagc	ctgcgcgccc	gcgagacctc	cgctcccctg	gccggcgcca	4080
	ccctgtcga	catggcgccc	cagtacacca	tcgcccagag	caaggcgggg	ttcctctccc	4140
	ccgagggccg	ctccccgccc	ctggacgccc	ccgccaacgg	ctacgtgcgc	gccgaccgcg	4200
	tcggatgggt	cgccatcaag	cgctcgcggg	acgcgcagcg	cgacggcgac	cccattccacg	4260
	ccgtcatcat	cgccagcggc	gtcaaccagg	acggccgcac	caaccggatc	accgtgcccc	4320
	acccccagc	ccaggtcgcc	ctgategagc	gggtctgcgc	cgccgcggcg	gtcacccccg	4380
	gcagcctgca	gtacgtcgag	gcacacggca	cctccacccc	cgctcggcgac	ccgctggagg	4440
	ccaacgccct	cgcccgcgcg	ctgtccatcg	gccgcgagcc	ggcgcccggg	acgtacgtcg	4500
	gttcgggtgaa	gaccaacatc	gggcacaccg	agtcggccgc	cgccatcgcc	ggctgatca	4560
	agacgggtgct	cagcctcaag	cacaagatca	tcccgcgcca	catcaacctg	gagaagctca	4620
	acccgcagat	cgacgaggcg	tccctgcccgt	acgagatccc	ccgcgagccc	acccccctggc	4680
	ccgacacag	cgcccgggcc	cgcccgggcg	tcaactcctt	cgcttccggc	gggaccaacc	4740
	cccagctcct	gctccaggag	gcaccgcccc	ccgcgcagga	gccccgcca	cccgccaccg	4800
[0002]	acgggtactc	cgctgtgccc	ctcagcggcc	gcgaccccga	agcctttccc	cccatcgccc	4860
	cgccgctcgc	cgaacgactc	gccgaggggc	tgccgggtggg	cgacgcccgc	tacaccctcg	4920
	cccaccggcg	gcagcatctg	gagcagcggc	tgtccgtcgt	gtacgactcc	cccgaggccc	4980
	tcgacgaggt	cctcggcgcc	gtcggcccgc	gcgagagcca	tccgctgccc	gtcggcgcca	5040
	cccagcgcga	ggcctgggac	cgccggctgg	tgtgggtggt	caccggcatg	ggcccgcagt	5100
	ggtggggccat	ggcccgccag	ctgtacgcga	gcgagcccgt	ctaccgcgag	gtcatcgacc	5160
	gctgcgacca	ggagatcgcc	gcgctcaccg	gctggctccct	caccagggag	ctgaacgccc	5220
	acgaggccga	ctcccggatg	agcgagacct	ggctcgcccc	gcccgccaac	ttcgccgtgc	5280
	agategcctt	ggccgcccctg	tgccgcagca	aggcgctcca	gccggacgcc	gtgaccgggc	5340
	acagaccggg	tgaggtcgcc	gcgttctacg	aggccggggg	gtacacgctc	cccgaggccc	5400
	taagatcag	ggtgcaccgc	agccggctcc	agcagaagct	catcgccacc	ggctccatgc	5460
	tcgcccgtcag	cctcaccgag	cgccgagccc	cccgcggggt	gcgcccgcac	ggcaccgggg	5520
	tctccatcgc	cgccgtcaac	agccccacct	ccatcaccct	ggccggggac	accgaggcgc	5580
	tggaggtgat	cgccgcccag	ctgggcgccc	aggacatctt	cgcccgttcc	ctggaggctg	5640
	gcgtcccgtg	ccacagcccc	cgcatggagc	tgatcaagga	cgagctgctg	acctcgctcg	5700
	ccgatctcaa	gccgcagcag	gcgaagttgc	cgctgtatct	caccgcgctg	ccgggcaccg	5760
	tcgcgagggg	cacggagctg	gacggcgaact	actggtggcg	caatgtgcgc	gaggccgtgc	5820
	acttccgggc	cgccgtggac	cggtctgctg	acgacgggta	cgccgtcttc	ctggagatcg	5880
	gcccgcaccc	cgctgctcgc	cactccctgc	gtgagtgtctg	cgaggccggc	gacgcgcaca	5940
	ggctcaccct	ggcctccatc	cgccgcaagg	ccgacgagcg	cgaacgcctc	accctgtcgc	6000
	tcgcccgcct	gcacagcctc	ggcttcgccc	tggactggca	cgccctgcac	cccgcgggac	6060
	gcccggccga	actcccgcgc	tacccttccc	ggcgcgaccg	ctactgggtc	gagccggccc	6120
	cggtcgcgca	gatccggctc	ggccaccgcg	accaccgctt	gctggggcgc	cgaccgcgca	6180
	gcgcccagcc	ggtgtgggag	gtgaagctgg	acgcggaggc	cgccccgtac	ctggaggacc	6240
	accgatccca	gggcaccgtg	ctgttcccgg	ccgcccggta	tctggagatg	gccgcgcagg	6300
	ccatgcggtc	gctgaccggt	ggtgagcaca	gcaccgcccg	gctggccggc	gtcgagctgc	6360
	gcaaggcgct	gttcttccg	gacggcgagc	cgagacgggt	gcagctgtcc	ttctctccg	6420
	acgcccgcgc	gttctccatc	gccaccgtgg	gcgcgcccgg	cgccgagccc	accgtgcacg	6480
	ccaccgggtac	ggtacggggc	gcccagcgcc	gcccggctgac	cgccgctgctg	gacaccgtcg	6540
	ccgtccgggc	ccgcgcccgc	cgccatctga	cgcccccga	ctgctacgcc	gaactggccc	6600

[0003]

cgctcggeta	ccactacggc	cccgccttcc	agggcacgca	ggaggtgtgg	atcggcgagg	6660
gcgaggccct	ggcgcggate	cggccgccc	aggggctcgc	cccggacgcg	gcggcgacc	6720
acatgcatcc	ggtgctgctc	gactcctgct	tccagtcgct	gctgacccc	cagttgctca	6780
ccgcgcccgc	cgggcccggg	ggcaccggca	tccggctgcc	gctgtccatc	gccgaggtac	6840
ggctggacc	ggtcggcgac	cgcgaaactgt	gggtgcacgc	caccgtcacc	ggtgaggacg	6900
aggacgaact	caccggggac	atcgccgtgt	acgacggcgc	cgacggcacg	ccgctggggc	6960
gcatcgccgg	attccgcgcc	gccgacgtgg	agaaggccgc	caccaccgtg	gggctgtcca	7020
ccatcgacag	ctggctcacc	gaactgagct	gggtgccgcg	tcccctgcc	gaggcaccgt	7080
ccgcccctgc	cgcggccggg	cggcacgtgc	tgttcgccga	cgcgggcggg	gtcgcgcggc	7140
ggctggccgc	gctgatcggc	gaggccggcg	gcgaggtgct	tctggtgcgg	cccggtgcgg	7200
cgtagcgctt	ggacagcacg	gcgcgcaccg	ccaccgtcgt	ccccgatcc	gcggatgacc	7260
tgcggcggtt	gctcgccgat	ctcggcccgg	tggacggcgt	cgccacactg	tggaacctgg	7320
accggccggc	gctggccgac	gcgccgcgcg	ggcggttcgc	ggacatcgcc	tccaccggcg	7380
cgtagccct	gatcgccctc	accagggcgc	tgctcgccga	cccggaacgg	cacggcggca	7440
ccccctgca	catcgtcacc	agagccgccc	agtgcgtgct	ccccggtgag	ccggtggagc	7500
cgctggcgcc	gcccgcctgg	ggcatcggcc	gggtgctgtg	gcagcaggaa	ctggccgggg	7560
gcggcgcaaa	gctgatcgac	ctggcgccgc	acggcgacat	cgaggaggac	gcgtccgcgc	7620
tgctgcgcga	actggccgac	cccaccgggg	ccgccgagcg	cgaggacgag	atcgcgctgc	7680
gcgccgggga	gcggcacacc	agccggctgg	tggccgcgca	agggctgagc	aggccgctgc	7740
ccctcgggct	gcgcccggac	ggcagctacc	tggtgaccgg	cgcggttcggc	gcctgcccga	7800
ggctgctgtg	ccgcacgctg	gtcaggcgcg	ggcgcgccgc	gctgatcctg	gtggggcgca	7860
cccggctgcc	ggagcgcgag	cgctgggccc	accaggacc	gaactcgccc	gccgggccc	7920
acgtggcctt	cctcaaggag	ctggaggcgc	tggcgcgca	gccgattctc	gcgccgctgg	7980
acatcaccga	cgaggacgcg	ctggccggct	ggctggccgg	gtaccggcgc	gccaggggcc	8040
cgccgatccg	cggggtgttc	catctggccc	ggcaggtgcg	cgacaccctg	gtcccgagga	8100
tggaccggga	ggtgttcgac	gccgtccacg	acccgaaagt	ggtgggcgcg	gcgctgctgc	8160
accggcagct	gagcggcgaa	ccgctggagc	acttcgtgct	gttcgcctcg	gtcgcggcct	8220
ggctgacgac	ggccgggcag	accaactacg	cggcgggcaa	cgcttctctg	gacgcgctgg	8280
cgcaccaccg	ccgcgcgcag	ggcctgccc	cgctggcctg	ggactggggc	ccgtgggcca	8340
ccggcatgat	cgaggaactg	ggcctgatcg	accactaccg	caacagcccg	ggcatgtcct	8400
cgctggccgc	cgagggcggc	atggcggtac	tggagcgggt	catcgggcag	gaccggcgcg	8460
agctgctggt	ggccacggtc	gtggactggc	cggtgttcat	gtcctggtac	gctcgcggc	8520
cgcggtggt	cacggagctg	gcggccaccg	cccaggggccc	ggggtccgag	ggcgacggca	8580
gtttcttga	cgcgttccgg	gagacgaccg	cggacaagcg	gcggctgctg	ctgaccgagc	8640
ggttcacgac	gctggtggcg	ggtgtgctgc	gcgtgcgggc	cgagcaggtg	gatccggcgg	8700
tcagcctgaa	tctgctaggg	ctcgactcgc	tgctggcgat	ggagctgcgg	gcccggttgg	8760
tggccgaggt	gggcatcgcg	ctgccggtgg	tggcgctgct	gtccagcgcg	ccggccgggg	8820
acctgatcac	ccagctgcac	gagggcctgg	aggagtgtct	ggccgaggag	ggcagcggcg	8880
ccgccgtgac	ggcggtggag	cggttcgagg	acgaggccga	gttcccctg	acgcagaacc	8940
agaaggcctg	gtggttctctg	aagcagctga	acccgacgg	ttacgcgtac	aacatcgggc	9000
ggcgctcga	ggtgcgggtc	gagctggacc	cggacctgat	gttcgaggcg	ttccgcccgc	9060
tgctggccc	gcattcccgtg	ctgcgggca	acttctgct	ggtggagggg	caggccggtgc	9120
agcggatctc	cccggagatc	aaggaggaca	tcgcgctctt	cgacgtcggg	gaccgcgctg	9180
gggacgacat	ctaccgatg	atcatcgagg	agtaccgcaa	gccgtacgac	ctggcgaccg	9240
atccgctgat	ccggttccgc	ctcttccggc	gcggcccggg	ccgctgggtc	atcaccaggg	9300
ccgtccacca	catcatctcg	gacgccatct	ccaccttcac	cttcatcgag	gaactgctgt	9360
ccctgtacga	ggggctgcgg	cagggccacg	acgtcgaact	gccgccggtg	tcggcccctg	9420
acctggattt	cctcaactgg	cagaacgcgt	tcttgccgg	ccgcgaggcg	cagaagatgc	9480
tcgctactg	gcggggcgag	ctgccggagg	aggtgccggt	gctggcgctg	cccaccgaca	9540
agccgcgccc	ggcggtgctc	accacaacg	ggcgctccga	gttcttcgcg	ctggacgcgg	9600
agctgagcgc	ccgggtgcac	gcgctggcgc	gggagcaca	cgccaccgtc	ttcatggtgc	9660
tgctgagcgc	gtactacctg	ctgctgcacc	gctattcggg	gcaggacgac	atcatcgctg	9720
gctccccgt	caccggccgc	accagggag	agttcggctc	cgcttacggg	tacttcgtga	9780
accgctgcc	gctgcacgcc	tcgctggccc	gtgacccac	ggtggccgag	ctgctggacc	9840
aggtgcgcac	caccgtgctg	ggcgccctgg	atcaccagga	gtaccgcttc	acgctgctgg	9900
tggagcagct	ggggctggcc	cacgaccgca	gccggtcggc	ggtcttccag	gcgatgttca	9960
tctgtctgca	ccacaaggtg	gccaccgaga	agtacggcta	caagctggag	tacatcgagc	10020
tgccccagga	ggagggccag	ttcgacctga	cgctctccgc	gtacgaggag	gagggcgacc	10080
ggcggttcca	ctgcgtcttc	aagtacaaca	ccgacctctt	cgaggcggag	acgatccagc	10140
ggctcgccgg	gcactacacg	cggtgctggt	agtcgctgac	cgccgcgccc	gccgacgccc	10200

	ccaccggtgg	actgcggatg	ctgtcgggcg	gtgagcggga	gcggatcctc	accgggtgga	10260
	gcggtgccga	agtgcgcgcg	caggaggcgc	cggtgccggt	gcaccggctg	atcgcgagg	10320
	cggcgcaccg	tacccgcag	gccgtcgcgg	tggccgcgcc	cgccgagagc	ggggagacc	10380
	ggcggctgac	ctaccggcaa	ctggaggagc	gcgcgggcga	actggccggg	cggtcgcggg	10440
	cgcgcgcggt	gcgcgagggc	accgtcgtcg	cgctgtgcct	ggagaagtgc	cccgagctga	10500
	tcaccgcctt	gctggcggtc	ctcaaggcgg	gcggcgccct	tctgccgctg	gaccgggact	10560
	atccggccga	ccggctcgcg	tacatggtgc	gcaacgccgg	ggccgcgctg	gtgatcggcg	10620
	ggacgggcgg	cgcgccgaa	gggctgccgg	gcaccgtggt	cagcctggag	gaactgctcg	10680
	tcggcgaggc	cgcgaggtg	gttccggacg	ccgagccggg	gcccgaactc	cccgcctacg	10740
	tcactatac	ctccgctcc	accgggcgtc	ccaaggcggg	cgcggtcagc	caccgaacc	10800
	tggcctcgt	gtacgcgga	tggcgcgacg	cctaccgtct	ggaagaggga	gacatccggg	10860
	tccatctcca	gatggccagc	ccttccttcg	acgtcttcac	cggcgacctg	accgcgcgcc	10920
	tgtgctcggg	cggcacgtg	gtgctggtcg	gccgggagct	gctgttcaac	accgccggc	10980
	tgtacgagac	gatgcgcgcc	gaacgggtgg	actgcggcga	gttcgtgcc	gccgtggtgc	11040
	gcaccctggt	gcggcactgc	gaggacaccg	gggcccggct	ggacttctg	cggtgctca	11100
	tcgtgggctc	ggactcctgg	aaggccgagg	agtacgagcg	gctgcgcgcc	ctgggcgcgc	11160
	ggcggctggt	gaactcgtac	gggctcaccg	aggccacat	cgacagcgcc	tggttcgagg	11220
	gtcccgcgga	cgacctggag	ggcgccgga	tggtgcccat	cgggcggccg	ttcccgggca	11280
	gcgcgctgta	catcctggac	tcgcgcggcg	agccggtgcc	gcccgggtgc	cccggtgagc	11340
	tgtggatcgg	cggcgccggt	gtggcgctcg	gctacctggg	cgacgaggag	ctgaccgggg	11400
	aacggttccg	caccgcacc	ctggcggcg	acgctccggt	acggctgtac	cgaccgggtg	11460
	acctcgcacg	gtgggacgcg	gccggcaccg	tccatctgct	gggcccggcc	gactcgcaga	11520
	tcaaggtcgc	cgggcaccgc	atcgagatcg	gggagatcga	gtcgcagctg	gcggcctgcc	11580
	ccgagctggc	ccaggcgcag	gtcaccgtac	ggccggacgc	gggcccggag	aacgtgctgt	11640
	gcgcgtacgg	ggtggcggca	ccgggcgccg	tgctggactg	gcgcgaggtg	cgccggcgcc	11700
	tggcggacta	tctgccgacg	ttcatgatcc	ccaccactt	caccgagttg	cccgtcttgc	11760
	cgctacccc	gaacggcaag	gtggacgtgg	cggcgctgcc	cgccccgcgc	gccggcgacg	11820
	gcgcggacgg	gccggtgtac	gaggccccgg	tcacgctgta	cgagacgcgg	atggccgagc	11880
	actggcagcg	gctgctgggc	atcgaggccc	ccggccccgg	tctgggccac	gacttcttcg	11940
[0004]	agaccggcgg	cagctccatc	cggtgatcgc	agctgateta	ccacctgcag	gccgagttcg	12000
	ggatctccat	cccggtcagc	cggtgttcc	aggtgacgac	gctgcacggc	atggccaaga	12060
	cggtcgcagc	gatcgtcacc	ggggagatcg	aggggtcgct	gccgtatctg	cggttcaacg	12120
	agaacgccgc	ggcgggcaccg	gtgttctgct	tcccgcggcg	cggtggccac	ggcctggtct	12180
	accgggagtt	cgcggcgcgg	ctgcccggagt	tcgagttcct	cgcttcaac	tacctgatgg	12240
	gcgaggacaa	ggtgagcggg	tacgccgatc	tggtgccggg	gcaccggccc	gagggcgaga	12300
	tcgacctgct	cggtactcgc	ctgggcggca	atctcgcctt	cgaggtcgcc	aaggagctgg	12360
	agcggcgcgg	ccgcaccgtg	cggcacgtcg	tcctcatcga	ctcgtgcggg	gtgacggagt	12420
	cctacgagct	gggtccggag	cacctggccg	tcttcgagcg	cgagctggcc	gggcatctgc	12480
	gcaagcacac	cggtcggca	ctggtcgcgg	agaagacgcg	ggaacaggcc	aaggactacc	12540
	tggagtac	cgccgcacc	gccaaccccg	ggaccaccgg	cgcccggatc	gcggtgatca	12600
	gcgacgagga	gaacgccccc	gcgtacgaca	gcggcgcgca	gggcagctgg	cacggccct	12660
	cccgttagga	aaccgaagtg	ctgcgcgggg	tgggcggca	cgccgacatg	ctgatcccg	12720
	ggacggtcga	gcacaacgcg	cgcttggcgc	gcggcattct	caccggcggt	gatggcgagg	12780
	tatgagcct	ttcgttcagc	cggcggtcga	caccaaggag	cacagggccca	tgtcatcccc	12840
	caccacctcc	ggcaccgccg	gcaggcagtc	gatgatcctc	atcgccggcg	gcttgggggg	12900
	cctgtccacc	ggctgctacg	cgagatgaa	cggtacgcg	acgggggtct	tcgagatgca	12960
	cgagatccc	ggcggttctc	gcaccgcctg	ggagcgcggg	gacttcaact	tcgactggtg	13020
	cgtcagctgg	ctgctcggca	gcggtcccgg	caacgagatg	taccagatct	ggatggaact	13080
	ggggcggttg	cagggcaagg	agatgcgcca	gttcgacgtc	ttcaacatcg	tgcgggtgcg	13140
	cggcggccag	ccggtgtact	tctactccga	cccggaccgg	ctccaggcgc	acctgctgga	13200
	gatctcccc	gccgacgccc	gccgatcaa	gaacttctgc	gacgggttac	gcaccttcca	13260
	gaaggcgctg	gccgtctatc	cgttctctaa	gccggtgggg	ctgatggggc	ggtgggagcg	13320
	gtggaaaatg	ctggcctcgt	tctgcgta	cttcgaaccg	atccgcaagt	ccatcaccga	13380
	gctgatgacg	gactacgccg	agaagtcca	gcaccgggtg	ctgcgcgagg	cgttcaacta	13440
	cgctgcttac	gagaagcacg	ccgacttccc	cgctctgcg	ttctggttcc	agctggcctc	13500
	gcacgccaac	ggctcggcgg	gggtgcccga	gggcggctcg	ctggagctgg	cccgtcgggt	13560
	ggagcggcgc	tacacggggc	tcggcgggga	gatcacctac	aacccaagg	tggagaagat	13620
	cctcgtcgag	cacgacaagg	cggtgggggt	acggctcacc	gacggccgcg	agttccgcgc	13680
	ggacatcgtg	gtgtcggcgg	ccgatctgca	caccaccgcc	atggggatgc	tcggcggccg	13740
	gtatctcaac	gacacctggc	gcaagctgct	caccgagacc	atcgacgagg	tgggcacgat	13800

[0005]

ctccccggc	tacgtctcgc	tgttcctggg	gctgcgccgg	ccgttccccg	agggcgagcc	13860
gtgcaccacg	tacgtgctgg	aggacagcat	ggcggcgaag	ctcaccggca	tgcggcatcc	13920
cagcatgaac	gtgcagttcc	gcagctgcc	ctaccgggag	ctgtcggcgc	gcgagaccac	13980
ggtcatcttc	gccacgtact	tctcggaggc	cgagccgtgg	cgggcgctgc	gcgacgacgt	14040
gccggagcag	gcgggcccgg	tgcggcgcgg	ccaggtgctg	cacaccctgc	cggtgaagca	14100
cggcaaggcg	tacacgcagg	ccaagcggca	ggcgcggatc	accatcgaga	acttcctgga	14160
cgagcggttc	cccggctca	aggacgcggt	cgccgtgcgg	gacgtgtcca	cgccgctgac	14220
gcaggtgcgc	tacacgggca	cctacaacgg	cggttcccc	ggctggcagc	cattcgtgga	14280
caagttcgat	accgtggaag	tggagatcaa	caagaacggc	ccggtgctgc	cggggcttcc	14340
caacttctat	ctggccgggg	tgtgggtcac	cgtcggcggg	ctgatccggg	cggtggcctc	14400
gggcccgcag	gtcacgcagg	tgatctgccg	ggacgacggc	cgggagtcca	cggcgagcgt	14460
ggacgagagc	gcgcccgcgc	ccaccaggt	ctccatcgcg	gtgggcaagc	agcccggcgt	14520
gccggatctg	gcccccgggt	tccccccca	gaccaccggc	ggcgacaccg	ccgccacac	14580
cgacaacacc	gtcacatcgt	cgaggagcgt	gtagtgcagg	tacacacatg	ggtgatcgcc	14640
gagcacatcc	cggggggtccc	ggacaccgac	cggatctacc	ggaaggtgac	gcgggagttc	14700
gatcccgcct	cgctggcgga	cgatcagatg	ctgctgcgca	cccggtagct	gtcggtaggac	14760
ccctacctgg	tggggctctc	cctccagacg	ccgatccggg	acaccgtgcg	cggtgactcg	14820
atcatggagg	tggaggtggc	ggggcccgcgc	gcccgttccc	aggctgggga	cctggtgcag	14880
gggtacggcg	gctggtgcag	ccatctggtg	tccaccgggg	ggcccagcgg	gtgggacgac	14940
gaggggcccg	agttcgccgt	ccagttgccg	ccgttccgca	agctggacc	gcgcccgtac	15000
gacgaggtgc	tgccgctgtc	cacggcgcgt	ggcgtgatgg	gcaccccggg	catcacccgcg	15060
ttcggcacac	tcaagacgtt	cctgacgggt	ggctcagagg	acaccgtggt	gatcagtggtc	15120
gcgtccggga	cggtgggcac	cctggtgggt	cagctcgcca	agcgggcccg	ggcccgggtg	15180
gtgggcacca	cctcctcacc	ggggaaggcc	gcgtatctga	cgagctggg	cttcgacgcg	15240
gtggtgact	accggcaggg	cgacgacacg	gacacggtc	gcgaggccct	ggcggcggcg	15300
gcgcccacg	ggatcgacaa	gtacttcgac	aacctgggcg	gcaccgtgac	ggacgcggtg	15360
ttcacgatgc	tcaacgtgca	ctcccaggtc	gcggtgtgct	ggcagtgggc	caccacggtc	15420
aacggggact	ggacggggcc	gcggctgctg	ccgtacatca	tgttcccgcg	caccacgatc	15480
cgcgggatct	tcgccgacga	gtggtacacc	gaggagatgg	tcgacgcgct	gcacgaggag	15540
gtgggcgggc	tgatccgcaa	ggcgagactg	gtctaccacc	agaccatcca	ccagggcttc	15600
gacgcctgc	ccggcgcgta	ccgtccctg	tacaccggcc	aggagggcaa	cgggggcaag	15660
gtcctggtcg	cgctgtagcc	gccggcacc	gcacggcggc	ccgacgtcat	ggccccttcc	15720
gcgttcgctg	agcgcgggag	gggccagcgg	tgtgcgctgc	cgaccgcgca	gatgtgctgt	15780
cgtgcgcagg	atcgcggtga	ggccccgatc	tcaagaaggc	tgtgcggggt	cggagatgac	15840
ggagagacc	gtgaagtccg	atgcgagggc	gtccatgatg	atctcgtcgc	acggcttgcc	15900
tcgccagtac	ccggcctgac	gaagtccggc	ggccggttgg	aaaccgcgct	tctcgtaggc	15960
gcggagcgcg	gcggcgttcg	gggcgaggac	ggagagccag	accatccgca	ggttcgtgac	16020
gtgccaggcg	tagtccagaa	cgagccgggt	ggcggccgtc	cccaggccgt	gcccgcggaa	16080
ctcgttccg	agcataacga	catactccgc	acagcggacg	ctgtggtcgg	gcagcagagt	16140
ggccacccc	gccggcgtcg	gctcaccgcc	ggtggtgagg	tcgtagagg	tgaagcggat	16200
attgtcccg	ccctgctgcc	ggcgcgatgc	ctcgttctcg	cgctcgagac	tctcgggggc	16260
ctggcggcgg	tatccgatca	tgaggcctgg	gtccttctcc	caccgccagt	aggtctcgac	16320
gaggtcggtg	cggtacgggc	cgaggccgca	cttctcgttg	cgagccaca	gaatcgggtc	16380
agtcggggtg	ctgatggcgg	tctccgatgg	tggggacggg	ggtgtaggcg	ccggtggacc	16440
ccggctcggc	gacggtggcg	ccgtccacag	tctgcaccca	ggcatggagc	cgacccgggt	16500
ctggcgcgac	gccatgacac	cactgggccc	tgcggccctg	agcggccagc	agcagcaggg	16560
cgccggccga	ctcctccaga	cacgcggttc	tgccggcgca	gtaccagcct	gcccggcgga	16620
ccgcctgcac	cgcggcggtg	gcttcggcct	cggtagccgg	gcggttgacg	cgacattgg	16680
tgttggcgta	ggttagcagg	gagatgatcc	gctgcatgcc	ctgagtcccc	ctggtgacca	16740
cgaagaccg	ggcgagagcc	gcgcccgtc	ggacaacaga	cgggagagg	acatgggccc	16800
cccggacgat	accggcgcgg	tgctcgtcgc	cgctccacc	gactcgggca	gtggccggga	16860
tgagggcggg	ccacggctgg	gggtgctctg	tgggaacaat	caggcccagg	tcgagcaaac	16920
ggctggtgag	cggtgtctcg	agtggggtga	ggtcgcgggt	acgggctgcg	gcacaccacc	16980
gcagcgcagc	ggcgggcagg	aggcattgtg	ccctgccaga	acggtagttg	atgagaacca	17040
ggacatgccc	gaggtccacc	gcgcccagc	tcgggggaca	ggtggcatga	gcgggcggat	17100
cagacatggg	cgttttctc	ccgggtccag	acaagggccc	gggcacgggtg	gagggagtg	17160
agccatgcct	cggtcgtgag	gaactgttcg	atcgtcgcga	gtggcatggg	gacgccggcg	17220
gctcctcggg	ccagttgcgt	gcggagcaga	cccggctcga	tgatccccag	cccggccaga	17280
tgcccctcgg	tgaggtcgg	cagcgcggc	aggttcgcgc	gcaccccgg	gtagtgggct	17340
gtgttgaagc	tgcccttgg	ggtgcgggca	gcgagctggg	cgccagctac	acggcacatc	17400

[0006]

gcccggcgca	ggacgggctt	gtaggcgtac	aagggcgggc	gggaccgcag	ctcggtcgcg	17460
agcaccgcat	cgacaacgtg	cgcgtccaga	aacggattgt	gcagctcaac	gccctgctgg	17520
tcggccaact	gggctcggc	ggccgcagta	cgggcgagct	cccggatctc	atccaccagg	17580
acccggggcc	cgaggtccag	gccgggcagg	gggtccgctt	cggcggcggc	gcgccggagt	17640
ccgccggcaa	gcaggctgcg	agccttctcg	gtgatccacg	gtggcatcgg	gagctgcggc	17700
acccaggtga	catcaccccc	gtcctccccg	gcgggggtggc	cggcgcctc	cacggtgtgc	17760
gcgaggtcgg	cgaaggcatc	cgcccggctg	acaccggcca	gccgacgggc	attgttcage	17820
agtggcagga	cggctgtgtg	ccgcagccgg	gcccagcccc	tcgcttcgct	caccgcccgt	17880
gcccaccgga	gatgccgcag	cagatcggcg	aggtgagcgg	cgggtacgaa	cagcacatg	17940
tctccgccgt	gcgccgtcat	gtgcgtgcgt	gtgccgaact	cggacttcat	ccagtcgagt	18000
tggtattcca	gacgggcgcg	ggccagggtg	gaagggcggg	gctcgtcggg	ggccggcacc	18060
gcggtgatgt	ccgtgtaggg	cagttgctcg	gtgccgaggg	agagcaggtg	gtgcgcgatc	18120
cgggggtcgc	gggctggccc	cagccgcgcg	tagtgagggt	cggcaccgct	ggtgtcgtcg	18180
tccgggtgga	cggtgacgcc	gttcaggtgg	ccctcggcgg	gacgggcccg	agcggcgagc	18240
agtgccaatg	tgctggaatc	cagcccggcc	gacagatcgc	acgacacctc	tgggtccagg	18300
accgtccggt	actcgaccgc	ggcctggagc	acgaccggga	atcggcgggc	cggatcacc	18360
ggcaccgggt	ccggccgcca	cacggttgta	cagcgtggcg	tgccccgctc	cgaaggcagc	18420
cacaccgcg	agcccgggtg	cagcagccgt	acaccggtga	agaaggaccg	agcaccggcc	18480
gccggaacgt	acgaagcgaa	gaccgatacc	gccaggcggt	ccgcgctgag	agcggccccg	18540
gtcagctggg	cgagaacccc	ggccgagctc	gaccaggccc	agccgcccc	gaactccgcc	18600
gtgtagaggg	cgatggcccc	ggccggatcg	gtatgcagca	cgacgccttc	ggcccggctg	18660
tccacgacgg	tgtaggtacc	gggcccaccg	caggtcacat	cctccggcag	ctgcccggcg	18720
gccaccgtgc	ccagctcggc	ttcggacgca	ccgcaacgcc	ccagcaccat	caccgtgccg	18780
ccgtctgctc	cgctgacggc	ccggaccatc	atcgggcccg	tccacacgga	gcgggcccgc	18840
tcggcgtatt	cggtgccatg	agggcggagc	ccgtcactac	aaccgaccgt	gaatcctccg	18900
aagaacatgc	gtctccttcc	gggtctggga	cctgaacggc	ggagggtgtg	gcacccggtg	18960
tgggggcccg	ggtgccacac	ccgcacggga	tcagtgggtg	ttgtaggcgc	gacgcttgct	19020
ctcgtctgtg	cccttgccct	ggccgagggt	cagggcgacc	agttcgcgca	ggtcaacgac	19080
categgcgcc	tcgatctcgt	tcgtctcggg	gtggatgtcc	atcacagttc	tccttccgcg	19140
tgcgctcgg	gccaggttgc	ccggggctcg	tgctgcctc	ccatcacacc	gccgccccg	19200
ccggccaccg	agggatttcc	cgggccgtgc	acaccgcgtg	cgtgctgcgc	ataccgctc	19260
gaccggcccg	cgcccggtc	gctaccgtgg	atcaattgat	gtcccgtg	gtttctgggg	19320
gcggtcatgc	aggcacagcc	accctgacct	gactggta	cctcacgagc	ccgggcgctc	19380
atcaaggggg	gcgacatcgc	cgggctgate	cgttgggccc	gacagcaacg	cggtggagg	19440
caagcggacc	tcggcaagge	ggccgggtac	tcgcctcgg	ccgtcagccg	cctggagacg	19500
ggcgtacgcc	cgccgaccga	catcgacatg	ctcagccggc	tgactgacgc	cctcgccatc	19560
ccacagaacc	tgttcgccgc	agcactcgga	atggcaccgg	ccgctcctcc	tcgactaca	19620
gttaccccca	cccatcaggg	gggcccgtgg	ggttttgagg	aggatccgat	gaaacgcgg	19680
tcgtaactcg	gcgccgggtc	cgccgcccga	ccgctcagtg	tcctcaccgc	tcttgatgac	19740
gcgctggacg	tctcgcccgc	accacgagcc	gccccgacc	ctgaccgctg	tgctgccgag	19800
taccgcccgg	cccgcgcgct	gttcgacggg	ggcgataccc	cccggctcat	gcgggcctcg	19860
ccggaccctc	tcgccaccgc	cgaagcccgt	gcgacaccgg	gcgaggcgcc	cgactcggc	19920
cggtctgcgg	cctgttacga	cctggccacc	gaagctctct	ccaagatcgg	acagaggcgc	19980
tcgcccggga	tcaccgcccga	ccgctccacc	atctacgccc	gcatctcggg	ctccccactc	20040
gctgccgctg	cgctccgccc	agtaactcag	attgtctctg	ggcacgaagg	ccgcgcaacc	20100
atcgcccagc	acgtcactct	tcaggctgcc	gcccctctcg	acgctgcccg	cccgtcgacc	20160
cccgcccaag	cgccggccct	cgcacagatg	ctgtgcaccg	cctcatacgc	cgcagcagaa	20220
gcgggtgacc	gcgagcgggc	gctggagatg	atccacgacg	ccgaacgtgc	caccgcgcaa	20280
ctaccccgca	cgcgctcctg	ccacatctcc	accgcgactg	tcacctctta	caaggtgggc	20340
gtgctgtggt	ccctcggcga	cgccggggcc	gcccctcggc	caggccgcac	tctcaaccgg	20400
gcacagttct	ccactgccga	acgcccgggc	cgtaccaca	ccgacatggc	ccgggcatgg	20460
tgggatggg	gcaagcccga	ggagaccgcc	caagcactcc	ttggcgccca	ccacatgcc	20520
ccggccgaaa	tcgcgaccg	ggcctcgatc	cgcgccatcg	tactgaact	ctccaacgt	20580
catccccgtg	ccggcggcgt	ccgcgaactc	ctcactctga	ccggacttgc	gtgagggacg	20640
ttgcgcatgt	acgggcctac	gggcccggga	gacgaccggg	gtgctcttcc	cggttgtctc	20700
ttcggcccga	cgccgatggc	gtcgcgtgcc	ggggctcgcc	ggatgccgag	ccggggggag	20760
gtcaggtggc	ggccagggtg	ctgagcgcgg	cgctcagccc	ggcgggtggc	tcgtcggcga	20820
cgggtttgag	ggaatccagg	ccggtgaggg	tgaccatggt	gtccgggtcg	agtgcctgga	20880
cgagcacgtg	gtcgcctg	cgccggacga	cgacgttgca	gggcaggagc	agtccgatgg	20940
agcggtcggc	ctccagggcc	cggtgggcga	cgccggggtt	gcaggcggcc	aggatgaggt	21000

[0007]

agtcctccat	gtcgtggccg	agcttggcct	tcagcgtggc	ccggacatcg	atctcgggtga	21060
ggatgccgaa	cccctggtcg	gccagggcct	ggcggacggc	ggtggtggcg	gcgtcgaagt	21120
cggtgacggc	caggcggacg	gtgcggtcgt	agggcacggt	gcgctgctcc	ttcggtgagg	21180
cggagaacce	ggctccttga	taccccgccg	ggtaccgggc	ggggtggacg	caatgccgca	21240
cccggggatg	cggggacgcc	agcaccacag	gaataccccc	tggggtaccc	ctgttagggt	21300
gggtatccga	ataccccccg	gggtacacca	ttcctgtgga	ggagcggcag	tcgtgttctt	21360
cgtegacacc	cttgatttcg	agggtctggg	caaccgcage	tacctggcgg	gcggggccgc	21420
cggcgcggtg	gtcatcgatc	cgccgcgcga	catcgaccag	gtgatcgccc	ccgccgccag	21480
acgcggggtg	accatcgcgt	acgtggcgga	gacccacatc	cacaacgact	acgtcagcgg	21540
cgccctggag	ctggccccg	tcaccggcgc	ctcctatctg	gtgccccggc	cgcgccacgt	21600
gtccttcgcc	cgcacccccg	tcggcgacgg	ggacaccgtg	gacgtggaca	cgggactggt	21660
cctgcgcgcg	atcgccacgc	ccggacacac	cccgcaccac	acctcctacg	tcttgaggca	21720
ggacggggcg	ggcgtggcgg	cgttcaccgg	cggatcgctg	ctgatcggca	cggtgggccc	21780
ccccgacctg	gtggagcccc	ggctgaccga	gcggctggcc	cgggcccage	acgcctccgc	21840
ccaccgctg	gcccgcgagc	tgcacgacga	ggtgccgggt	ctgcccaccc	acgggttcgg	21900
cagcttctgc	tcttcgcccc	aggccgaggg	ggacgcgacc	acgatcggca	gggaacctgc	21960
cagcaacgac	gcgctcacc	tggacgtgga	gaccttcggt	cagcggacgc	tcgcccgggt	22020
ggaggacgtg	ccgcctact	acgcgcacat	gggcccggtc	aacgcgcggc	gccccgcct	22080
cctcgacctc	accccgcccg	agcgtgcgga	cgcgggcctg	atctcctccc	ggctggccgc	22140
ggcgtagtgg	gtggtggacc	tgcgcagcgg	cgtggccctc	gcccaggggc	acgtggccgg	22200
ctcgttcaac	ttcgaggcgc	agggcaaac	cgccacctac	ctggcctggc	tgatcccgtg	22260
gggcaagccc	gtcacctcgc	tcgcccacac	cccgcggcag	atcggcgcgg	cgccagcggga	22320
gctggcgcgg	gtgggcatcg	accgcccggc	cgcgcgcgcc	accggcgacc	cgcccggtcg	22380
gatccgggag	ggcgagcage	tcgcctcctt	cccgcgtgcc	cgcttcgcgg	acctgcggcg	22440
ggtgcgcgag	cgcggcgacg	aggtggccgt	cctggacgtg	cgccgggaat	cggagcgcgt	22500
gggcggttac	atcgacggct	cggtccacat	cccgatccac	gagctgcacg	gccgtatcgg	22560
cgaggtgccc	gacgggacgg	tgtgggtgca	ctgcgcgggc	gggatgcccg	cgccgatcgc	22620
cgctcctctg	ctcgatgccg	ccggccgcga	tgtggtgcc	gtggacgacg	gtttcgacgc	22680
cgcgacggat	gcccgggtgc	ccctgacctc	cgctgagtt	cgcccgctcc	cctcgcgagg	22740
aaggaagcgt	gtgatgttct	gctcccgcgg	aaacggggcc	cgtgtcacgg	tcgacgaggg	22800
acgcagcggt	accagcggca	tgaatgcggg	ggccggccgc	ggacaccccc	tcgtcgacc	22860
acgcggcaaa	accgggcaca	cagcgtgacg	gtcctcattc	tcgccctcat	cgccggggcc	22920
gtcateggtc	tggccctcgg	cgccctcggc	ggcggcggca	gtgtccttgc	cgtaaccggca	22980
ctgatctacc	tgtctggttt	cactccggct	gccgccacca	cgccagcct	gatcattgtc	23040
accgccacct	ccgccacggc	cctgtacgcc	cacacccgcg	acgggaacct	ggcctggagg	23100
acagggggcg	tgttcgccgc	cgcgggcac	gtccccgctt	tctcgcggcg	cgctgtcgcc	23160
ggacgcccgc	ccgcagcgg	actgacggcg	gcgttcgcgg	tcacgcggcg	gctggcggcc	23220
ctgcgatcgc	tgcgcccgtc	cgcgctccg	ccgcgggacc	ggatccgtcc	cgccaaggcg	23280
gcccgtgccc	gtgcccga	cggcgcccgt	acaggcctgc	tgggggtcgg	cgggggatcc	23340
ctcgcctg	ccgcccgtgt	gagcgtcctg	gggtcggcca	tgcgcccggc	ggtgggcacc	23400
agcctgctc	tcataccgt	gaactcactc	gcccgcctgc	ccgcccgcag	cgccaccacc	23460
ggcgaactcc	actgggagat	catcgcccc	ttcaccggag	ccgcgatcct	cgccgcccgtg	23520
gacggcaaac	gcctcgcgac	gaagatctcc	ggcaccaccc	tgcggaagat	tttcgcccggc	23580
ctcctgctgg	cggtggcggc	cttcatgctc	gtcgacgtga	tcgtctgaac	acgagggccc	23640
ccaccggctc	ggaccgggga	gcaggagaag	cttctccggc	cgcccgcgca	tctcgggggc	23700
cgtctccggt	ccggctccc	tcagatcaag	ctcagcacac	ccaccacgca	aggacacct	23760
ctgccatgac	cacaccacg	gccctcgcca	ccgacgagge	gcgcacccgc	ctgcacgaac	23820
tgaccatcat	cgacgtgcgc	accccggg	aatacgccgg	cgcccatctg	cccggcgccg	23880
tcaacatccc	cctcgaccag	atacggcg	cgctgcccga	catccgtcag	gccgcggccc	23940
gcccgcacat	cctggtcgtc	tgcgctcgg	gcgcccgtc	ccagaacgcc	tgcacgatcc	24000
tcgcccagaa	cggtcgtgacc	gcccaccacc	tgtcgggagg	caccggggcc	tggggcggcc	24060
acggcaacga	gctccaccgc	ccccaggcgg	tctcccgtgc	cgctggggga	atggagcggc	24120
aggtaaggct	caccgcggga	gtgatcgtc	tgtcggcct	gctgctcggg	ctcgtcgtgc	24180
accccgcctt	ccagttctc	tccgcccggca	tcgcccggcg	tctggtcttc	tccgcccctga	24240
ccaacacctg	cgcatggcc	gccatgctc	ccaagctccc	ccacaaccgc	ccccgcggcg	24300
ccgacctcga	caacacgctg	gccgcaactc	gcgaccgctg	aaccacacca	caaggggagg	24360
gaccgggagc	agggtcctc	cccttgtg	ggccagacca	ggcccgggtg	cgaaccgctc	24420
cggacagggc	gggtgatccg	ccctcggc	accccggctc	aggcgagaga	aaggaacagc	24480
ttctccagcc	gagcccgc	ctgctccc	tcaccggaag	cagccatgct	cgctcgggtc	24540
aggcagtget	gcagcccc	ggcgatgatc	gcgaaaccgg	ctttgtccaa	agccctggac	24600

[0008]

gccgcggcga	gctgggtgac	gacgtcctcg	cactcccgc	cctcctcgat	catcttgatc	24660
accccggcga	tctggccctg	gcgccggcgc	agccggttga	ccactgtctt	cagttcctcg	24720
gccgccatcg	ccagttccac	aacccacact	cctcagagat	acccccttgg	gtatcctatc	24780
gaaggggtaa	ccccaaccgg	gcgaagaaaa	atccccgccg	ctccccacca	ccgctcggcc	24840
ggcaccggg	cacaagccgg	ggaaccctct	ccggccggct	gaaccgagcc	ccggcggcgc	24900
gcgcttcteta	cccgactgcc	ggccccgtgcg	gcacactccc	ccgcatggtg	cgtctgcaca	24960
tcgtcacgtc	ggatgactgg	ccgctgtggc	aggaggctccg	cgctcggcgcg	ctgaccgatg	25020
cgccgtatgc	cttcaaggcc	cggctcggcc	actgggacag	ggcggccgg	gagcagtggc	25080
gtgctcgtct	ggacctgccc	ggctcccaca	acgtcgtcgc	gctgcgcgga	gacctgcccg	25140
tgggcatggt	cagaggcatc	cgcggtgacc	acgggacgag	cgagctcagg	tcgctgtggg	25200
tcagccccgc	ggcacggggt	gacggtgtcg	gagaccggct	gatcgaagcg	gtgacgact	25260
gggcgcggcg	gtcgggccc	accacgctca	ggctcgcgg	gatccccgga	aacgcgtccg	25320
ccctcgcct	ctaccggcgc	aacgggttcg	ccgccaccgg	agagttggga	gagctgttgc	25380
ccggcgggtg	gaccggggag	caggtgctgg	cccggagact	ttcctgatcc	gcgtcacgtc	25440
cggacgccg	cacaggcccc	ttctcgtggc	tcggccccgc	cgaccgtgcg	gtcggcgggg	25500
cggcctttcc	gctaccaggc	gacgggcagt	tcgtggaggc	cgttcatgaa	gtggccccgc	25560
ttccacttga	gagcgttctc	ggggatgccg	aagcgcaggt	cggggaagcg	ctcggtgagg	25620
gcttcgagga	ccacctgcag	ctcgaccggg	gccagatggg	cgccgatgca	gaagtgcggg	25680
ccctggccga	agcccagggtg	cgggagcccc	ttgcgttcga	ggtcggggcg	gccggcgtcc	25740
gggtagacgt	cggggtcgcg	gttggcggcg	gcgatccagc	ggatgacggc	ctcggcggcg	25800
cggacggta	ccccgctcat	ctccaggctc	tcggtggcgt	agcgggggaa	ggagaagccg	25860
ctgaggatcg	gcacgtggcg	catgagctcc	tcgacggcgg	ccgggatctg	ctccggcttg	25920
tccaccagg	gctggagcct	ttcggggcgg	gccagcagca	tggcctgaa	gltggcgatc	25980
tggtggcg	tgctgacgaa	gccggcgtg	agcaggctcg	aggcggcgc	gtgcatctcg	26040
atctcggta	tcggccctc	ctcgtcgcag	gcccgcacca	gcatgccgaa	catgtcgtcg	26100
gtgggtcggg	cgcggcgctc	ctccaccatg	gcgcccattg	aggcgggtaa	ctcctggtag	26160
cgccggatga	ggtcgtcggg	cggtgtgtcg	ttcgacagga	tcgtctcggc	ccagccccag	26220
aacaggtggt	ggtcgggctc	gggcacgccc	atcatctcgc	agaccacggc	ggtgggcac	26280
ggcacggcga	ggtggggcag	gagatcggcg	ggggaccggc	cgccctccat	ccggtccagg	26340
cactggtcgc	tcagctccc	caccctcgc	cgcagcttct	ccaccggcgc	cggggtgaag	26400
tcccggccca	ggacggagcg	cagccgggag	tgctcggggg	ggtcggggga	gaacaggccc	26460
ttgccgacgc	gggccccggc	ccgggtgcgt	ggctggtcct	ggccgaggga	ggcgacgatg	26520
gagaaccggg	agtccgacag	cacggtcttc	acatcggcgt	accggatggc	gagataggcg	26580
tcgtcggcgt	aggggagccg	gaccttgggt	agccgctgct	ggtacagctc	gtggtagacg	26640
gggggaaact	caagaccctg	cgaggcgaag	gggaaggtga	ggagttcctg	ctgctccggg	26700
tgctccgacg	gtgcctgctg	ttcctgctgt	ccgggcatgg	gggcgggtgt	gcctttcgtg	26760
gctgggacat	ggtgagggcg	tgtgccggcc	ggcacggtgg	cggaggggga	ggggcgccgg	26820
tgcccgtggg	acgacggggc	ggctgggtgc	cgggcggcac	acgccctgtc	cagcatgcgg	26880
ttccgggcac	cgccgacggg	caccgggacc	atcaagatca	ctccattgag	cgagcgtgcg	26940
cctacattgc	gttatgcgcg	ccggtgcttc	acagtcctat	ttcccggccc	gcgcgcacat	27000
agccggggaa	ctcctgcggc	ccgcaggca	cttcacgcca	gcccggcccc	gtcagggagc	27060
ccctggccag	cagttgccgc	cactcccccg	gcccacgta	cgatcgcgg	cgcccggccc	27120
gctccagcac	cggggccacc	agcacgtcag	gaccgagcag	gtactgcaac	tgcgcgtcgc	27180
gggcgcaccg	gtcctcgggg	aaggcgagcg	gcacggccg	catgaggggc	atcccgcct	27240
ccctggtggc	ctcggcggcg	cagccggcga	tgtacggggc	gagccggtgc	cgagccggc	27300
acgccgcgat	ggcggccgac	cgtaccggct	cggggtaggc	ggtcggctcc	ctgcgcccc	27360
tcccgtgaa	ccgcatgatc	ggggagaacg	ccgccactg	cgcccaccgc	ccgtacaact	27420
cgggctccac	gtccgcggtg	acggcgctccg	gtgtcatccg	ctgccgcacc	acgtaggcgg	27480
cttcgaagga	ctggggcgtg	tagtagccgc	ccgcgtcgtg	ggtgaccacc	gagcagccgc	27540
tgagcgccat	ggactccacg	gcgcgcaccg	tgagcccg	tcgctccag	ctggagggca	27600
tgtaaccgc	ccagtgcgcc	gggttgcgct	gcgagcccg	cgtgcccgg	cgggcgatcg	27660
ccgtgaagtc	cccgcgcgg	gcgcgcagtc	ccgctcgcg	caccacctcc	tggtagatca	27720
gcccgtagga	gttgcggcgc	tgcgtgcccc	ggctgccgtc	ggcgaagacc	gcgtcggcgg	27780
ggatctctc	gccgaagtgc	gcgagcacgg	cgctcgcgtg	ttcctcgcgc	agcgcgtcgg	27840
ccagccgctc	gcaccaccat	tcgcgggcct	cggggtgggt	gaagtgcagc	acgagggagt	27900
ccgggttgc	ggcgggtggc	gccggggcgc	cgtacacgtc	ccggacgagg	tagccgcgcc	27960
gggtcagctc	gtcggcccagc	ggagtgccgc	gtttgacgta	cgggttgatc	cacaggctgg	28020
tgcgcactcc	gcgttggggc	agttgccggg	accagtcggg	cggaaagcgg	cgccggctcg	28080
gcccgtacgt	ccaggcggcg	gcctccagga	cgctgccggc	cagccactcg	tcgacgtgca	28140
tcacgtccac	cgggcagccc	cgggcctgga	gttcgtcgac	ggtgcggacc	atctcgtccg	28200

[0009]

cggtgaagta	ggaggagcgg	ctcatccaca	ccccgaacgc	ccagtccggc	agcgggtgcg	28260
gacggccggt	gagggtaggg	tagcgggtcca	ggatgggtggg	ggcgtcgccg	ctgatgagga	28320
agaggtcgag	gtcatcgccc	tcgatctcga	tccgggcggc	ctcggcgtgg	gtggcgccga	28380
tgtcggcgtc	gacgggcccg	ccggtgtgcg	cgaacaggcc	ccagccggcg	tccgaccaca	28440
gcagcggcac	attgaggtag	gcggtgtcgc	gcccggcggc	gcggttctcc	tccggtgttc	28500
gcagcgtgcg	gcgcccggcg	cgcaggtcga	tgccctggta	gctctcgccg	ccgccgtaga	28560
cgcagtggcc	cgggttaggg	tggatggctc	ccaccacaga	ggcgggtcgg	tggtgccggc	28620
tcgcccggcg	gtaccgggag	cggaacggcg	cgataccggc	ctccggctcg	ctgaagcggg	28680
ggtaggtgcc	gaagtggaa	ccctgccccg	gcccgcgttc	ccacagcgcc	cgcaccccgg	28740
gcccggatgat	ctccaccccc	tcgggggtgg	tgtgcacccg	cgccgggccc	cgcggggcgt	28800
ccagcagcac	cggcgagtcg	tggtgggtga	cggcggcggg	ccggtcggcg	agccgcagcc	28860
gcaacacccc	gtccagcggg	acggtgcact	cgacgcggac	ggtgggtccc	gaggcgagcg	28920
tcaccgcag	gaacgggccc	tccgcctccg	gacctgcggc	cgcgtacggg	tccggtgtgc	28980
tcgacatccg	ccgcgggtac	ccggcaccgc	aggggcgaaa	cgggcgtttg	agggcggccc	29040
ggtccggcta	cccggcgccc	atgacactca	cggtggaacg	gcgcgcccgc	gtcctggggc	29100
cggacgagct	ggcggccctg	gacgcgcgct	ggcgggcccg	gaactatctg	accgccgggc	29160
agttgtacct	gctggacaac	ccgctgctgg	ccgagccact	gcggcccag	cacatcaagc	29220
cgcggctgct	ggggcactgg	ggtaccgcac	ccggcctcaa	cctgctgcac	gcgcatctga	29280
accgggtgat	cgacgcccgc	gccctggacg	ccatgtgctg	gtggggcccg	ggcgaccggc	29340
gcccgggtgat	ggtggcgggc	acctggctgg	acggcacctg	gagcgaggtc	taccgggatg	29400
tgccccggga	cggggcgggg	atggcgcggc	tcttcgccca	gttctccttc	cccggcggga	29460
tccccagcca	cgtggcgccc	gacgtccccg	gctcgatcaa	cgagggcggc	gagctgggct	29520
acagcctggc	gcacgcgtac	ggggcggcgc	tggacaatcc	ggggctgctg	gtggcgtgcg	29580
tcatcgggga	cggcgaggcg	gagaccgggc	cgctggcggg	ttcctggcac	gcgaacaagt	29640
tcctggacc	ggtgcacgac	ggggcgggtc	tgccgggtgt	gcacctcaac	ggctggaaga	29700
tcgccaatcc	cacggtcctg	gcccggctgc	ccgaggaaga	actgacacag	ttgctgcgcg	29760
ggtacggcca	cgaccgcgtg	ttcgtctccg	gcgaccggcc	cgaactcgtc	caccgggagc	29820
tgcccggggc	gctggacgag	gcggtggaac	ggatcggccc	gctccagcgc	gcggcacgcg	29880
gcggggcggc	cgggggcgta	cgcgagcgtc	cccgggtggc	gctgatcgtg	ctgcgcacgc	29940
cgaaggctg	gaccgggccc	gacgaggtgg	acgggctgcc	ggtggcggcg	acctggcgtc	30000
cccaccaggt	gcccgtggac	ggggtggggc	agaaccccgc	tcacctgcgg	cagctggaga	30060
cctggctgcg	ctcgtaccgc	cccgaggagc	tcttcgacga	gcggggcccg	ccgctgagg	30120
cgggtgctggc	gcaggtgccg	caaggcacgc	ggcggtggg	cagtagcggc	cacgccaacg	30180
gcgggctgct	gctgcgggac	ctgccggtgc	cgcgctgga	gggcgcggcg	gtgccggtgg	30240
agcggcccgg	cagcacgctg	cacgagccga	cccaggtgct	gggcaccctg	ctcaccgagc	30300
tgatgcgggc	caccgcggcg	cggcgcgact	tccgcatcgt	gggcccggac	gagacggcct	30360
ccaaccgct	ggacgcggtg	tacgaggcca	ccggcaagga	gtgggagagc	gcggtgctgg	30420
gcgtggaccg	gcatctggag	caaggcggcc	gggtgatgga	ggtgctctcg	gagcatctgt	30480
gccagggtcg	gctggagggc	tatctgctca	ccgggcggca	cgccctgttc	tctgctacg	30540
aggcgttcgt	gcacatcgtc	gactcgatgg	ccaaccagca	catcaagtgg	ctgcgtccc	30600
gtggcggca	ggaccacaac	ggcttctcgc	ctctccctgaa	ctacctgctg	acctcccacg	30660
tgaacaagag	cccggaggtg	gtgcgggtct	atctgccgcc	gggcttcgtc	gacatgtgc	30720
cgtgcgccga	tcactgctg	cacagccggg	actacgtgaa	cgtgggtggt	gccgggaagc	30780
agccgtgctt	cgactggctg	gacctggaga	gcgcgcgggc	gcaactgcgc	cgcggcggcc	30900
ggatctggga	gtgggcccgt	acggagtcgg	cgggttccgg	gcgcgagccg	gacgtggtgc	30960
tggcctgtgc	cggggacgtc	cccacgcagg	agacgctggc	ggcggcgtgg	ctgctgcgcc	31020
gggagctgcc	ggggctttcg	gtgcgggtgg	tcaactggtg	ggacatcgcg	cggtgctgc	31080
ccgccgagga	gcatccgcac	gggatggcgg	acggcgagta	cgacgcgctg	ttcaccgcgg	31140
accggccggt	gatcttcgcg	taccacggct	atccgtggct	gatccaccgg	ctgacgtacc	31200
ggcggccggg	gcacgaccag	ctgcatgtgc	ggcgtaccg	cgagcagggc	acgaccacga	31260
cgccgttcga	catggtggtg	cgcaacgatc	tggaccgcta	ccggctggtg	atggatgtga	31320
tcgaccgggt	gccggggctc	gcggtgcggg	cggcgggccc	acggcagcgg	atggcggatg	31380
tgcggttcgg	gcacgagcgg	tggatccggg	agcacggcga	cgatctgccg	gaggtggccg	31440
gctggcgggt	ggacgacggc	gagggcggcg	gatcgtgacc	ggcgggcccg	tcacaggacg	31500
ttgcccgcct	ccttgatcag	ccggagcgac	tgctgctggt	cgagggcgct	ggagagcagg	31560
tcgctgaagg	cgccccgta	cttctccacg	tcgtcggggc	ggcgggtgat	gccggagccc	31620
gccaggttct	ccaggtagac	gatgggcccg	atgcccggtc	ggcgggaagc	gatcaccacg	31680
aacgggcccg	tcatcccggc	gtgcgcgccc	gcgctgaacg	ggacgacgag	caccgggatg	31740
tgcggtgcc	cggcggccag	cgaggcgatg	tggtcgagct	gggcgcgcat	caactcacgc	31800

[0010]

ccgccgaccg	tacgccgcag	caccgcctcg	ttgaggatgg	cggtgagctt	gggggcgccc	31860
tcggggcaca	gcacctcget	gcgcgtcatg	cgcaaggcgg	tgaggcgctc	gacctggtcg	31920
gcggagaggt	cgggcgggcc	cagccggtgg	acggcgcggc	tgtactcctc	cgctctgcagc	31980
agccccggca	cgtactcggg	ctggtagctg	tgcagctggg	cgggcgctggc	ctccagcgcc	32040
aggtacgcct	ggaagcccgg	tgagatgacc	gagcggtaet	cgggcgactg	ccaccagtcc	32100
ttcttggctt	tggtgaccac	ggcgtagctc	tccaggcggg	cgggcggtctc	ggcgtcgggc	32160
tcgtacagac	ggcacagcag	catgacgtcg	gacggcttga	cgacgccgtc	gctgcccggtc	32220
tccaggcggg	tcatcttggg	ggcggcccag	tggaagcggc	ggggcgacctc	gcccggcggtg	32280
aggcgggact	cctcgcgcag	ccggcgcaet	tcggtgctca	ggtgcaetgc	gcacaacgcg	32340
ggtgacgcgg	attcgatggt	gcggaccatg	cgccctcctc	ggtatgtgce	ggggcgggcac	32400
cgcggtcgcg	gtacgcaatc	cgcgccgcag	tgcttccggg	agtatgccga	ggcgccgcac	32460
gattcttcca	acggatcgcg	ctggaacagc	agttcagggtg	tgatgccggg	gcccgcggcc	32520
gggacgccgg	gcacggaatt	ccgtgccgcg	cattcggcgg	cggtggatgc	gcggcgccag	32580
ccgccgggcg	agcacggcgg	gcagcagcag	ccacagcccc	gcgcacgcgg	cgaaggcacc	32640
ggccgcggcg	atcaccggg	ccgttcggga	gaccaccacg	tccaccacca	gcagcagcgc	32700
accggtcagg	gcgaggggca	gcgtcaccag	gccgagcgcg	gccaggcggg	aggagacctc	32760
gacgatgcgc	cgcttggcgc	cgcgccggaa	cagcgcccgg	tgcaagcggg	cgggggcggtt	32820
gaacaggatg	ctcgcggcca	tcgccagcag	cagcgtggtc	aggtaggtgc	cccgtggaa	32880
ggcgtcagc	accgggaagc	gggcggtgaa	ggcgagggtg	agcaggaagg	cgaagaggat	32940
ctgcaagcct	gtctgggtga	ccgcagctc	ctggagcagt	tcggcgaagt	tcggctctgc	33000
gcgtcgcct	ggggtttccc	tcctgctgtg	ggcgggtgct	ccctccatga	cagccattct	33060
tccgtccgcc	ggccggggccc	gcaggcgcag	cgcatggcgg	tctgcggcag	gatgcctcgc	33120
atgcgtatgc	ggatactggt	gctgggcggg	acatggttcc	tggggcgcgc	ggtcgtcgc	33180
gagggcctgc	ggcggggtgg	gaggtcacca	ccttcaaccg	gggctgtcc	ggggccgacg	33240
tccccggcgt	gcggccgggtg	cgcggggacc	gcggggcggt	cgccgatctg	cgggcgctgg	33300
cgagtgagg	gccctgggac	gcggtgatcg	acagctcggc	cgccgacctc	ggcccggcgg	33360
cggtcgcggc	gggcgcccag	gcgctggagc	cggtggccgc	acggtatgtg	ttcgtctcct	33420
ccgtcaacgc	ctaccggggc	tggccggccg	agccgctcac	cgagtccagc	gagctttctg	33480
caccgcggcg	gacggggtac	gggcgcgcga	aggccgccgc	cgagcggggc	gtgggcgacc	33540
ggttcggcac	cgggcgcacc	gcggtgctgc	ggcccggcct	gatcctgggc	cccggcgagt	33600
acgtggggcc	gctgccgtgg	tggtgcacc	gggcggcccg	gggcggggcg	atgctggcgc	33660
ccggcagccc	cgggcggacc	atccagccgg	tggacgtacg	ggatgtcgc	gccttcgccc	33720
tgcggtcgc	gggcgcccgcg	ggcccggctg	gcggggcgta	caactgggtg	gcgcccgatcg	33780
ggcgggagac	gatgggcggg	ctgctgcacg	actgtgtgac	ggtcacccggg	cgggcgggccg	33840
aaccgcgctg	ggtggcgagc	gaggtgctgc	tggcgcacgg	ggtgcccggc	tggaccgaac	33900
tgccgctgtg	gcgggcgatc	gacggcggct	gggaggtggc	atcggacgcg	gcgcaggcgg	33960
ccgggctgat	gtgcaggccg	ctggcggaga	ccgtgatgga	cacctgggcc	tggtgaacg	34020
gcggcggtct	tccggtggac	catccgcgcc	gggcccagca	cgggatcgac	ccggccaagg	34080
aggcggcgat	cctggcctcg	ccgggcgggc	cggtgacggg	agcgcgccgc	ccgcccgggc	34140
ccaaagccc	cgcgccgcac	cgccccctcg	cgtgtttctc	cccgcaccgg	gcccgcgat	34200
gcaaccttt	tggcggggcg	tccgtggtgg	ggggctgaa	ggggcctga	atgcccggcc	34260
acgaagaaca	gcacttccgg	gacttcgcag	cacaggaggc	cgccgcactg	cgccgcaccg	34320
cctacctgct	gtgcggggac	tggcatctgg	cggaggatct	gatgcagagc	gtactgatca	34380
agggtgaccg	gcactggtcg	cggtggagc	ggcgcgacgc	ggtgcagccc	ttcgtccggc	34440
gggtgctgct	gcgcacctgg	ctggacgaga	agcgcgaagc	ctggcggcgg	gcccagcaga	34500
gcttcgccac	cctgcccag	gtcatcgacc	cggcagcggg	ccccgatcag	gcccgcgacc	34560
ggctgcacac	ccgggcccgtg	gtgcggcacg	cgtgtctcca	gctgcccggc	gggcagcggg	34620
ccgccctggt	gctgcggttc	ttcgacgatc	tgctgggtctc	cgacacggcc	gcccgcgatgc	34680
gctgttcggg	gggcaacgtc	aagagccaga	cggcacgcgg	tctctcgacg	ctgcgccggc	34740
atctgggcct	gaaccgcaac	ccgcagccgg	ccggcgccgt	gtggcagcac	gggtgggcgt	34800
catgacggaa	cggggtcagg	aggaggctcc	ccacatctg	gcgaacgcca	cagcgcagca	34860
gcccccgtg	cggtgggaca	tggacgagct	ggtcggacgg	gcccgcgggg	agcgcagacg	34920
gcgggcacc	ctggtgggca	cggtggcggc	caccgcctc	atcgccgtct	ctggcacgct	34980
gctgcccgcc	ctgctgtcgc	cgccccgcc	gctgccccg	ccggggcccc	ccgccccga	35040
cgccccgacg	acggccccctc	cgggcgtccc	cgacgccgtg	gagccggatc	cggtggatcc	35100
ggattcggcc	ccggcgtccg	ggctgtggcc	cgacccggca	ccggtggact	acgaccggga	35160
cgcgttcgcg	gcggtgtcgg	cgtacgccac	gcaccggctg	gcgtcggtac	cgacgccat	35220
cggggacggg	ggcgagggtc	cggacgggga	gggctacacc	ctcggcgtgc	tcaccgaccg	35280
cagcgtgtat	ccgggacggg	gacagaaggt	cttctctctg	cacttcaccg	cggcgctgcg	35340
caccgcgcgg	ggggagttca	ccctggcct	ccgctggac	agccaatcca	ccgcttccc	35400

[0011]

ggagccgctg	tcgtgcegcc	ccgggacggc	cgaggacgga	ccgtgcccga	tcggggagcc	35460
gcaggcgcac	tccccccccg	acccggcggg	gggcggggac	agggacgggg	ccaagccctc	35520
gtcccggatc	gagatcgcgg	gggtggtgtc	gaccctggag	ttccggctca	gcgactttca	35580
gagcgtcacc	gtcaccgtca	ccggcgcaga	gcggaccccc	gtgccgctca	ccgagggcca	35640
aatggatgac	atcgccccgc	gtctcctcgc	atagcgcctac	caacgggcag	tcgacaaaga	35700
agagtatcgg	tgacggctcg	tcgaccgagg	gaccggtgaac	gccccccatg	aaaaccctc	35760
agcgcagact	gcggatcggc	tggttcctgg	gtgcctcggc	ggcctcggcg	gtgatattcc	35820
tcgccctcgc	cggaccgtcc	gcgggactgc	tgcccgagga	gacgcacggt	gcgaagaacg	35880
gcaccgggga	ccgcagcgtc	gaggcgtcgg	cgccgcggcc	cggccccag	ggcggcgcca	35940
cccccgcccc	cgggcggggag	gacgccggac	agcgggaccc	cgacacggcc	gggtaccggg	36000
aactgctgga	gacggtcatc	gagatcgacc	cggagctgac	cggttccggc	cggctggacc	36060
cggtgcccgg	cagcgaccgg	gcggccgagc	cggacgcgac	ccgggtctac	cgctaccggg	36120
tcgatgtgga	gcggggcctg	ccgggtggacc	cggagtctct	cgccgaggtc	gtgcaccgca	36180
ccctcaacga	cccgcgcagc	tggggcaacg	acgggcagcg	cggttcgccc	cggatctccg	36240
gcggctcccc	ggacttcgtc	atcaccttgg	ccagcccggg	gacgaccgcc	gagtgggtgcg	36300
ccaaggtcgg	cctggacacc	tgggaggaca	acgtctcctg	cgacgtggcc	tacaccgagc	36360
gcgtgatgat	caacgcgtgg	cggttcgcgc	agggtccaa	gaccttcggc	gaggacatcg	36420
ccggttaccg	ccacatggtc	atcaaccacg	aggtcggtea	ccgcctcggc	cacggccacc	36480
acatctgcac	gaaggaaggc	gcgctcgcgc	ccgtgatgat	gcagcagacc	aagaccctga	36540
ccacccccgg	caccggggtc	agctgcccgc	cgaaccctg	gccgaaccgg	gggaagtgc	36600
cagcagcate	gcgaaccggt	cctgcggatc	gtccaccag	tggtccaccg	cgagaccggc	36660
cgccgctcagt	tcttcggtga	gggtggcccc	gcggaacttg	gtggagatct	ccgtaccgct	36720

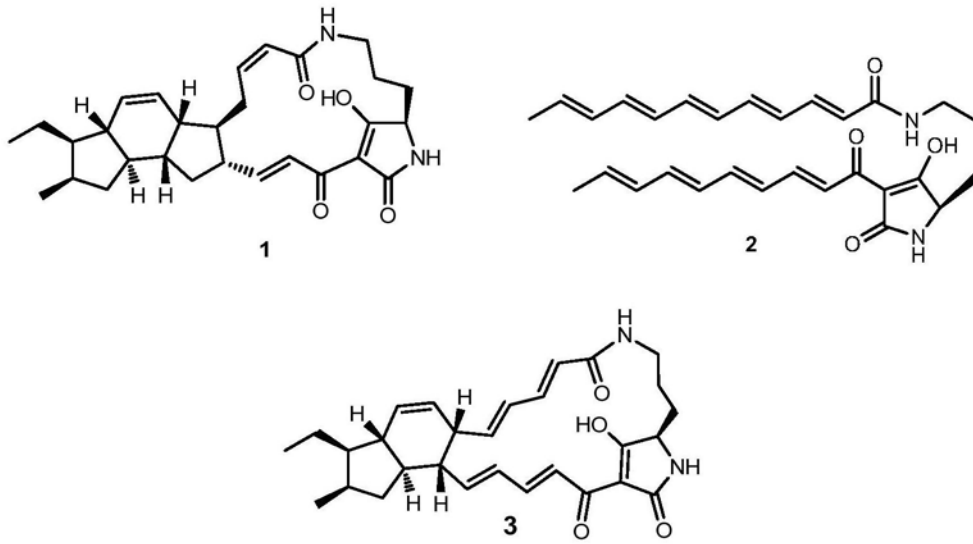


图1

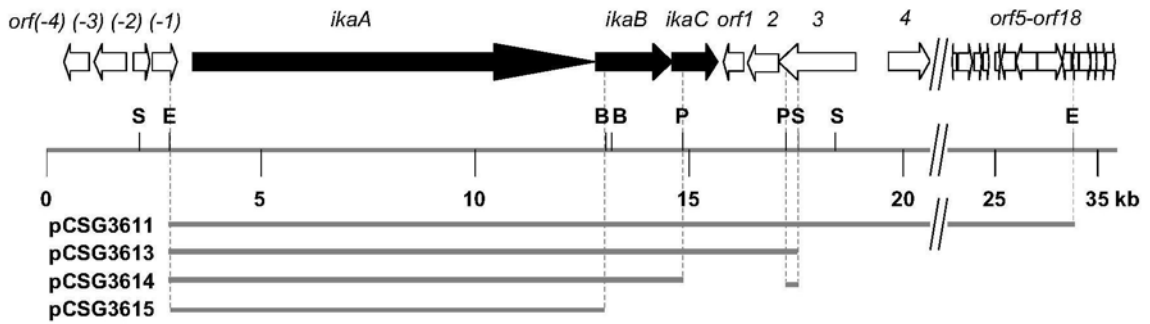


图2

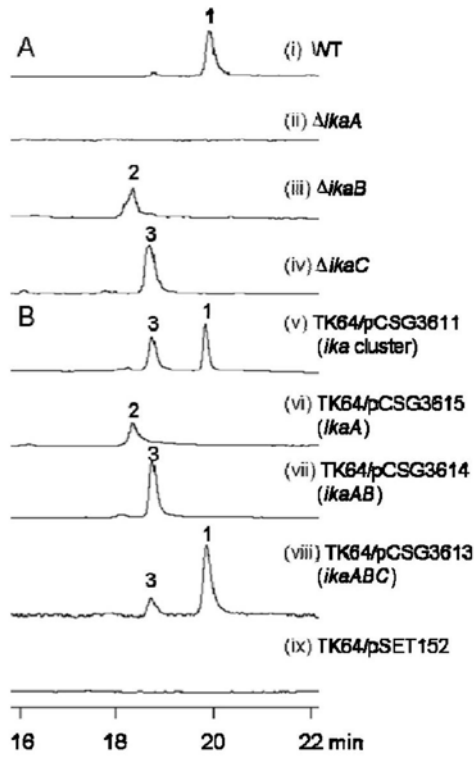


图3

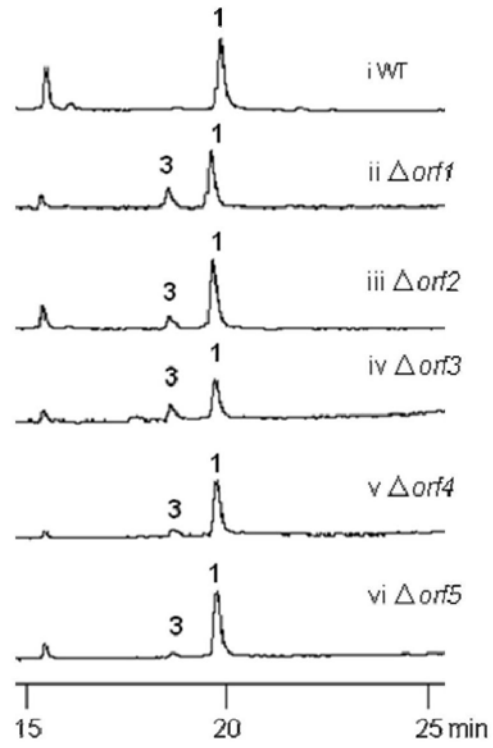


图4

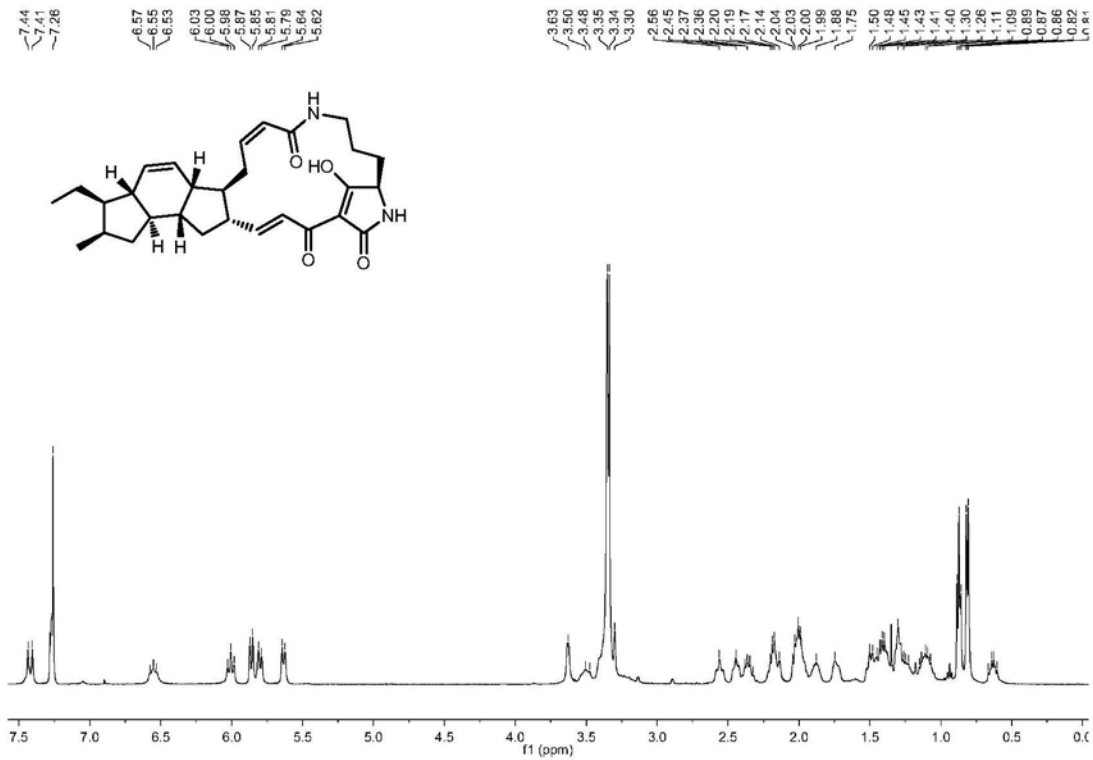


图5

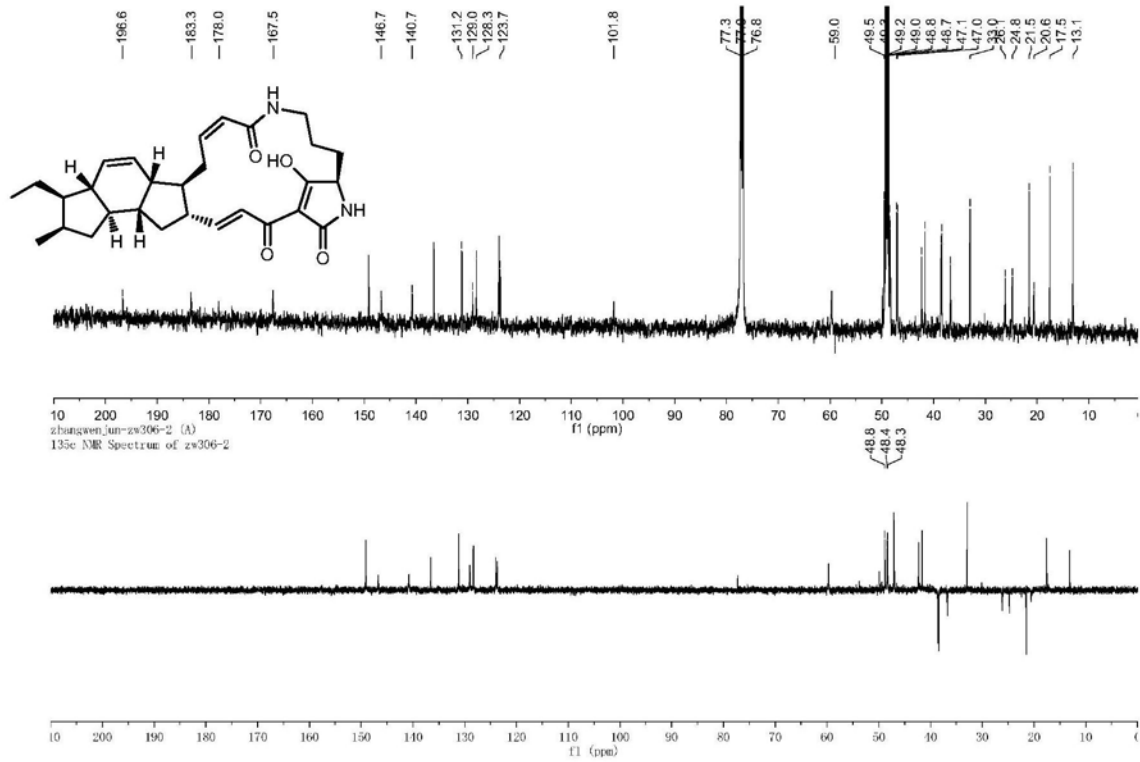


图6

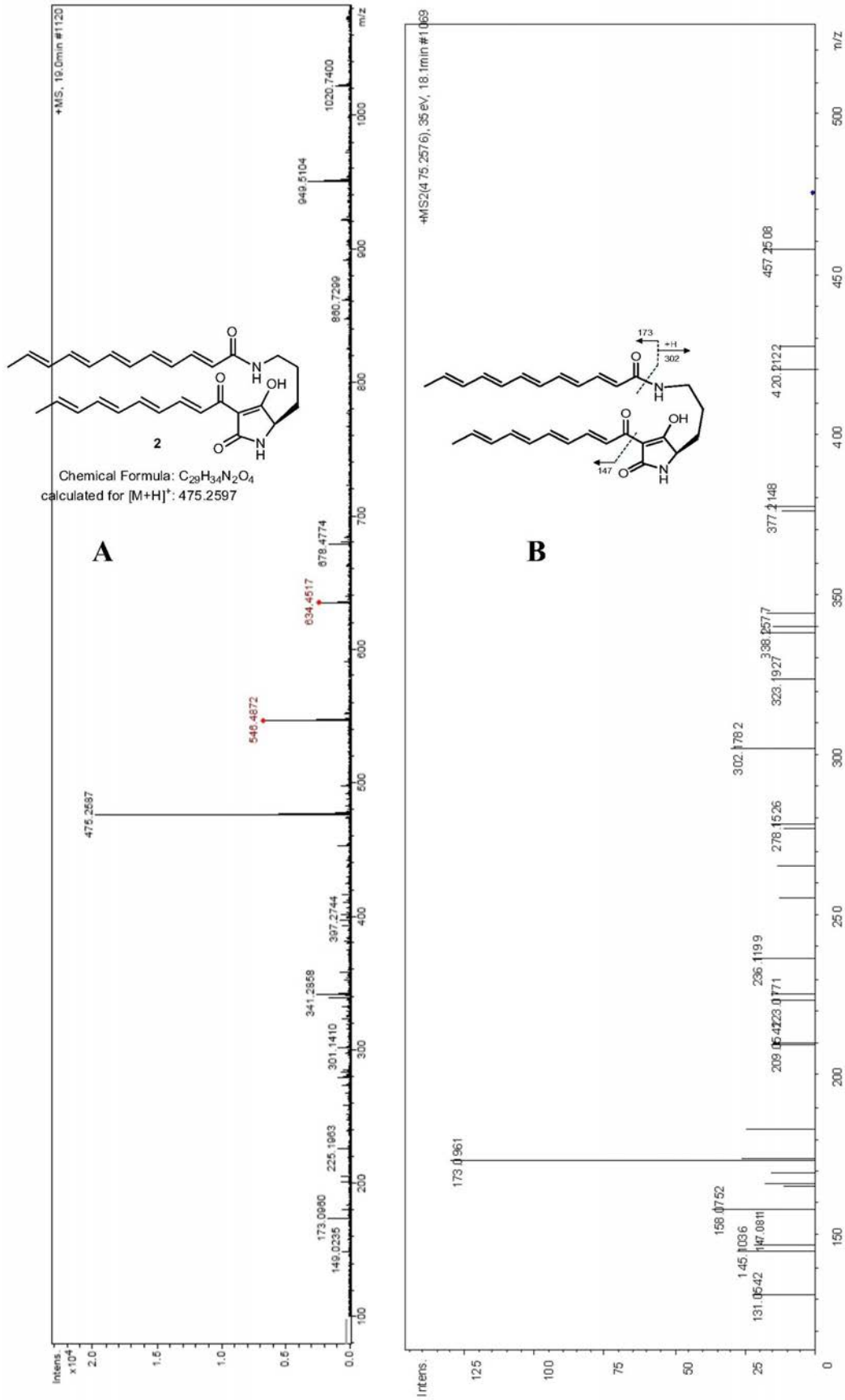


图7

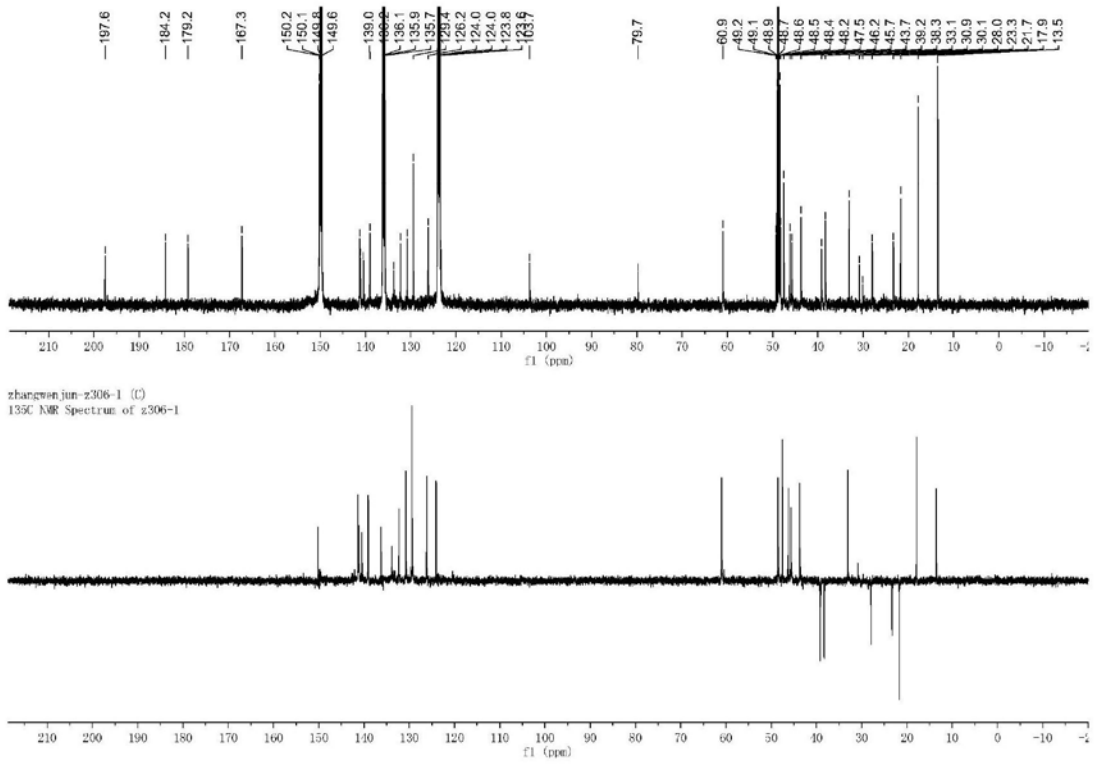


图10

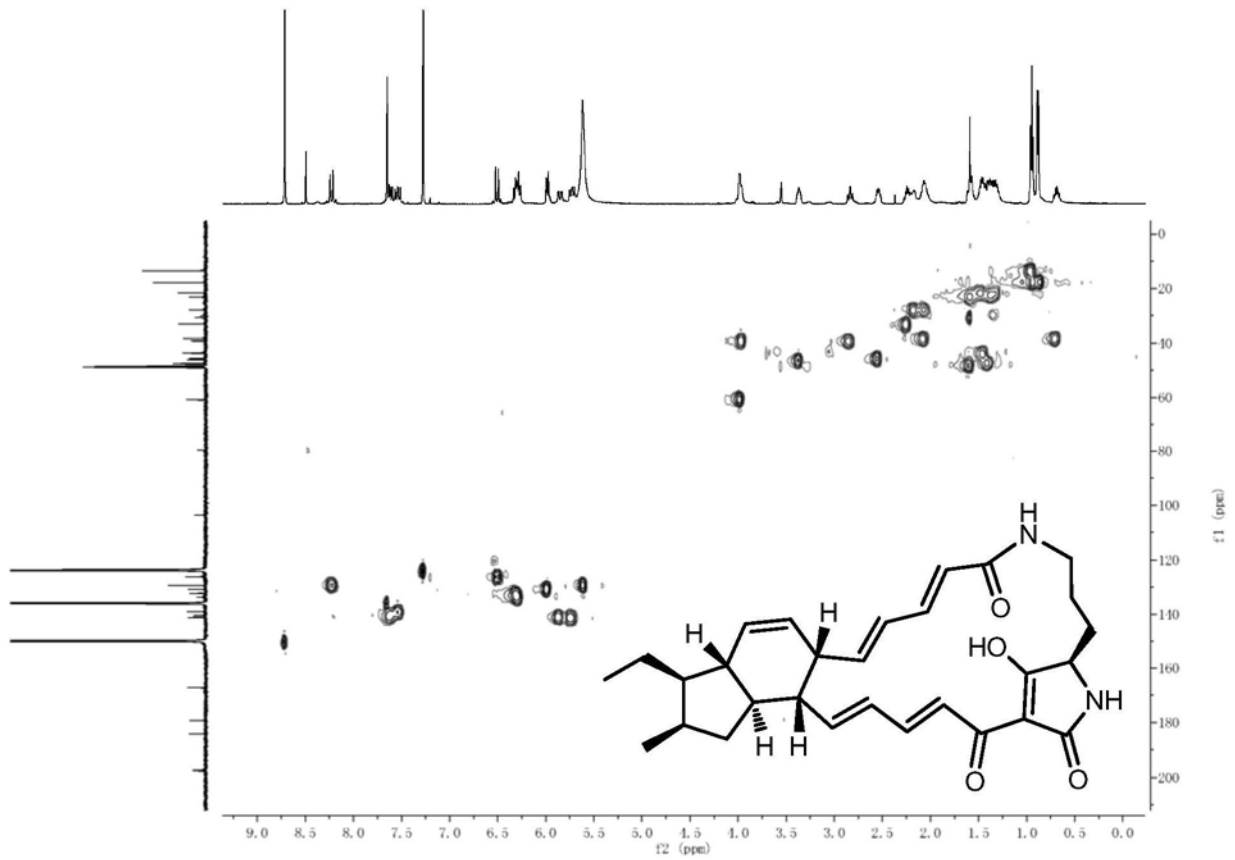


图11

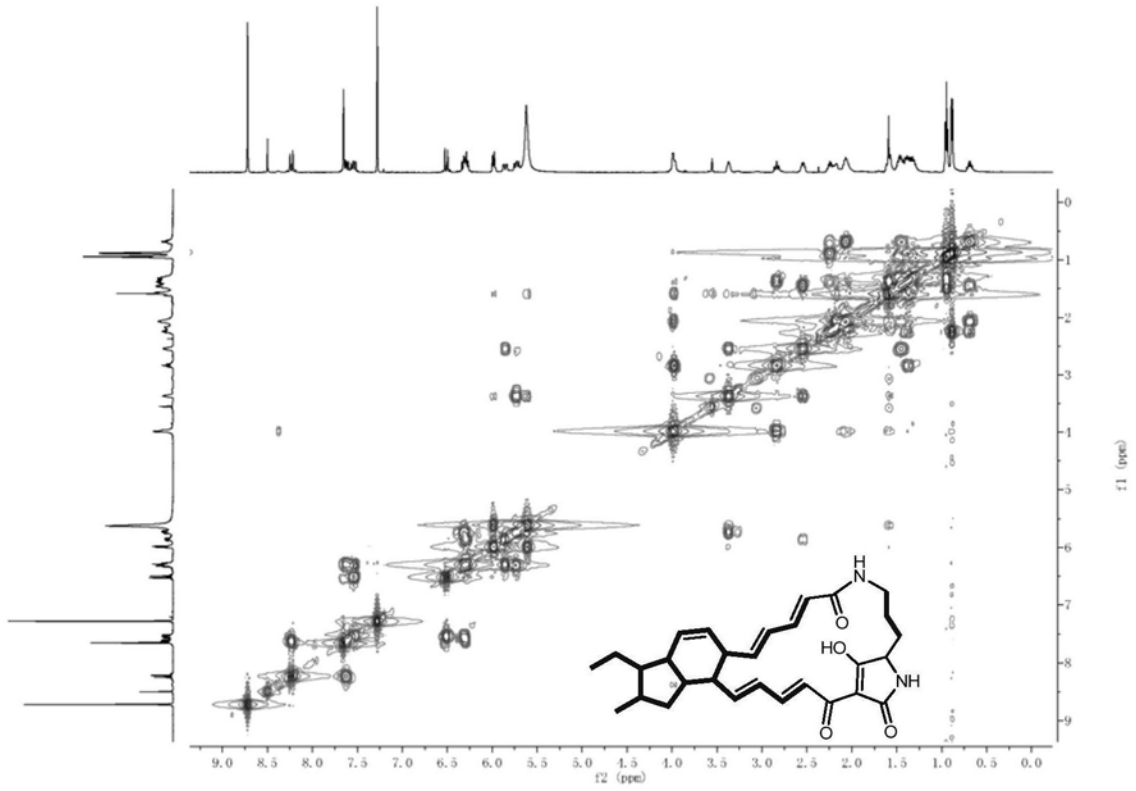


图12

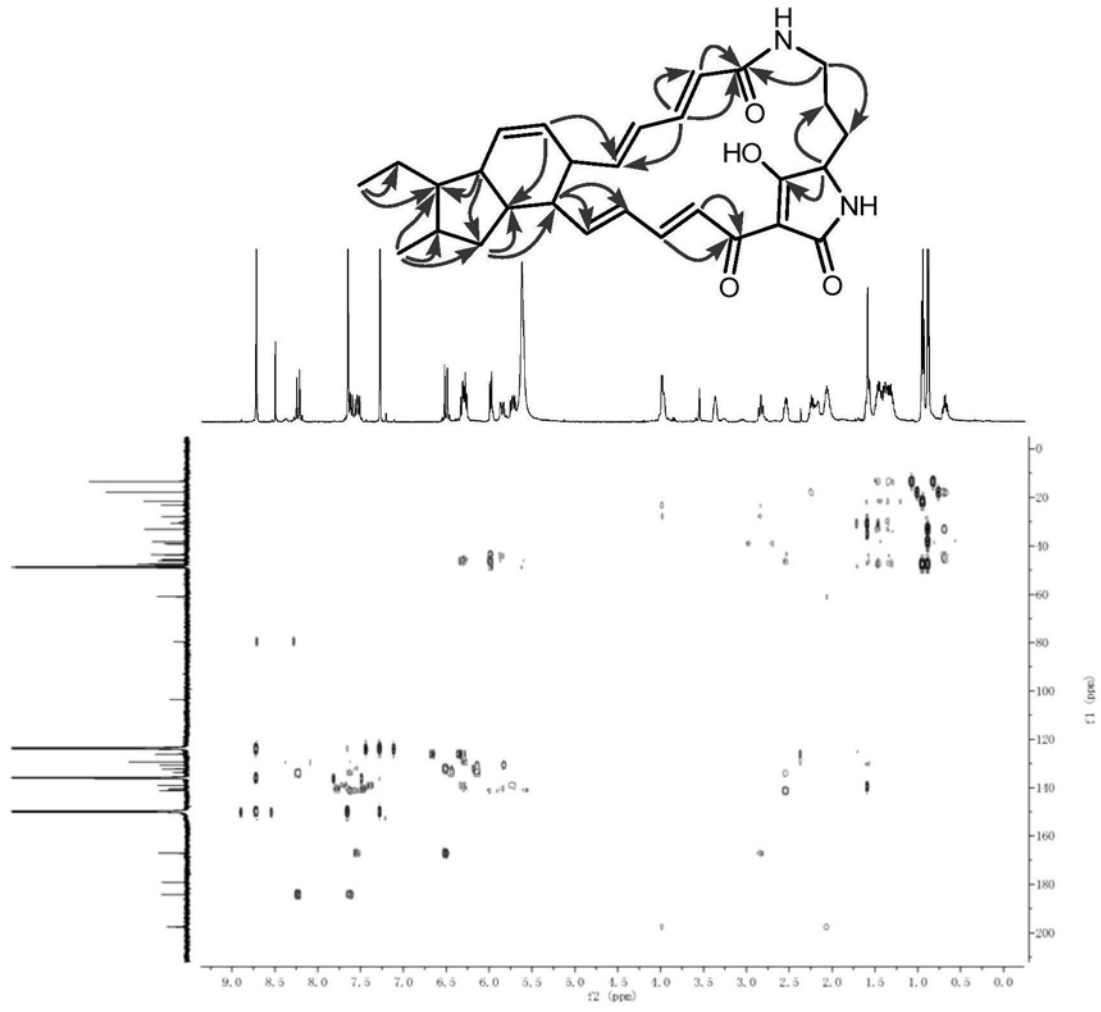


图13

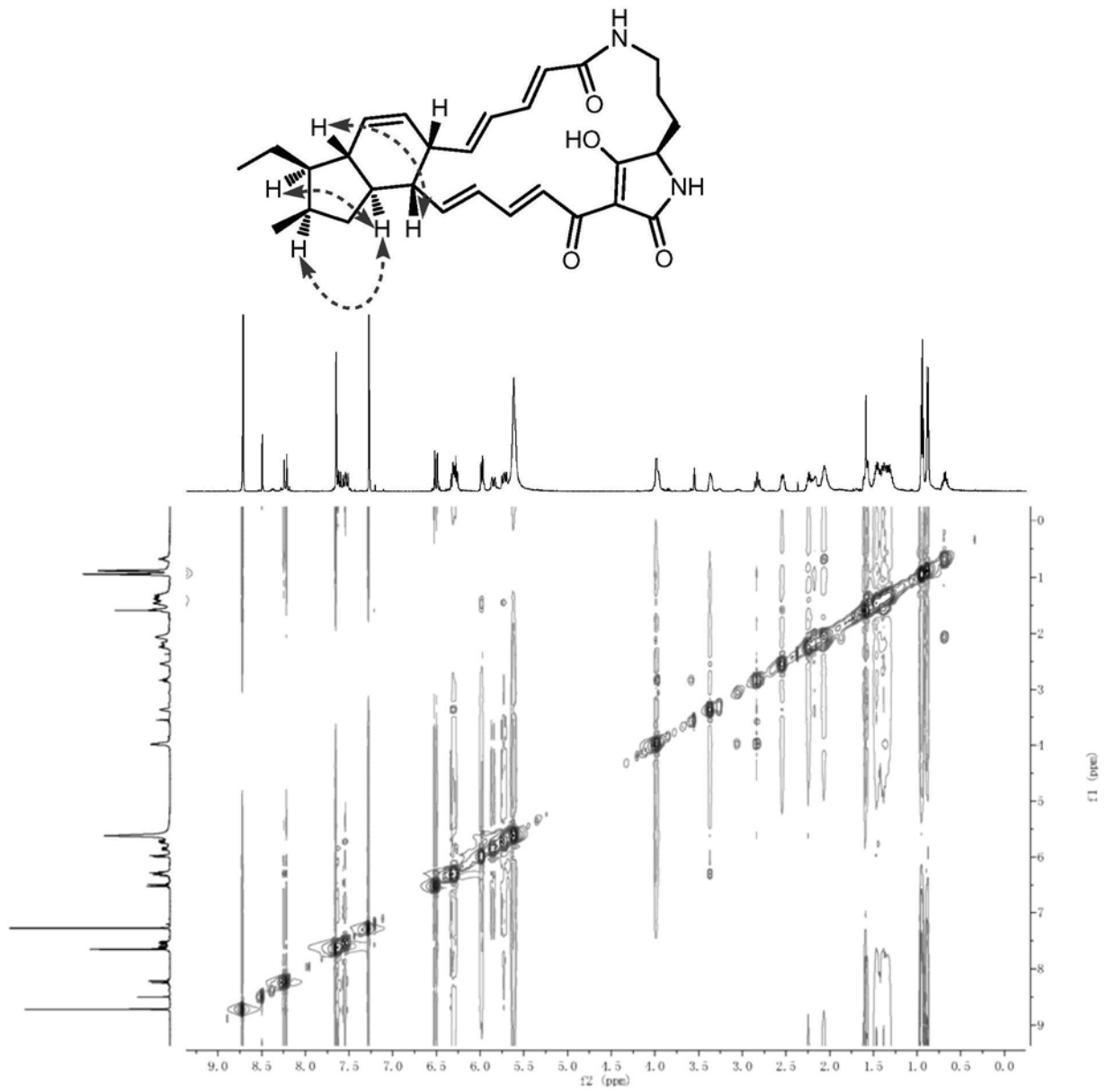


图14

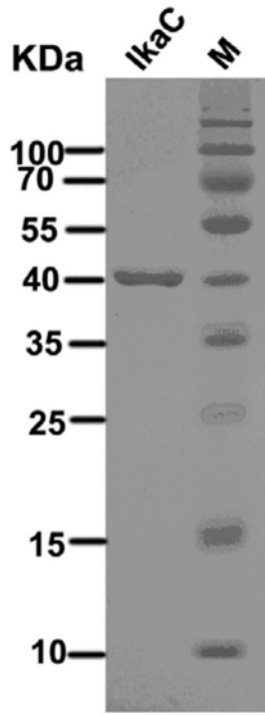


图15A

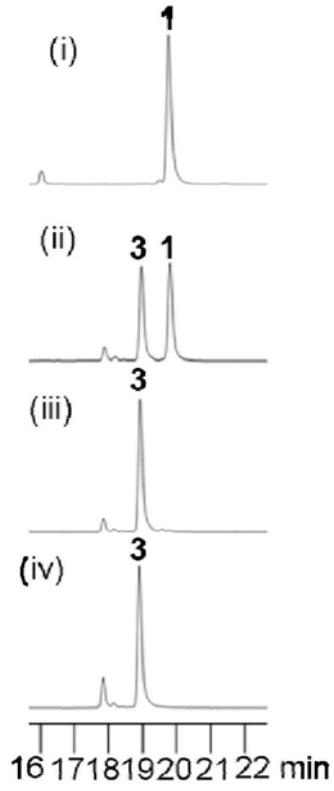


图15B

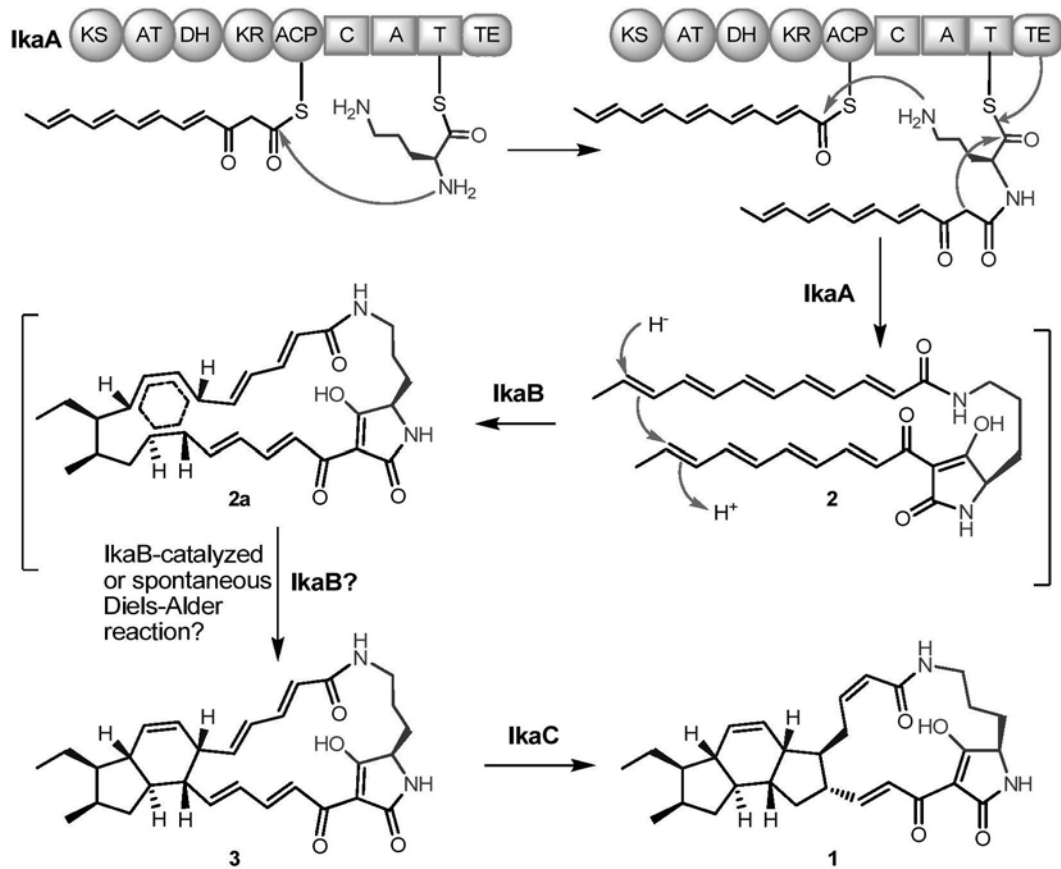


图16