

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6163557号
(P6163557)

(45) 発行日 平成29年7月12日(2017.7.12)

(24) 登録日 平成29年6月23日(2017.6.23)

(51) Int.Cl.

F I

C O 7 D 307/68

(2006.01)

C O 7 D 307/68

C 1 2 P 7/44

(2006.01)

C 1 2 P 7/44

Z N A

C 1 2 N 15/09

(2006.01)

C 1 2 N 15/00

A

請求項の数 21 (全 74 頁)

(21) 出願番号	特願2015-533235 (P2015-533235)	(73) 特許権者	509262736
(86) (22) 出願日	平成25年9月20日 (2013.9.20)		シンセティック ジェノミクス インコーポレーテッド
(65) 公表番号	特表2015-536641 (P2015-536641A)		アメリカ合衆国 カリフォルニア州 ラホヤ ノース トレー パインズ ロード
(43) 公表日	平成27年12月24日 (2015.12.24)		1 1 1 4 9
(86) 国際出願番号	PCT/US2013/061036	(74) 代理人	100102978
(87) 国際公開番号	W02014/047510		弁理士 清水 初志
(87) 国際公開日	平成26年3月27日 (2014.3.27)	(74) 代理人	100102118
審査請求日	平成28年9月8日 (2016.9.8)		弁理士 春名 雅夫
(31) 優先権主張番号	61/704, 408	(74) 代理人	100160923
(32) 優先日	平成24年9月21日 (2012.9.21)		弁理士 山口 裕孝
(33) 優先権主張国	米国 (US)	(74) 代理人	100119507
早期審査対象出願			弁理士 刑部 俊

最終頁に続く

(54) 【発明の名称】 化学物質およびその誘導体を生成するための組成物および方法

(57) 【特許請求の範囲】

【請求項 1】

60 を超える温度で 5 - デヒドロ - 4 - デオキシグルカレート (DDG) をアルコールおよび無機酸と接触させて、2, 5 - フランジカルボン酸 (FDCA) の誘導体を形成するステップ

を含む、FDCA の誘導体を合成するための方法であって、
該 FDCA の誘導体が FDCA のモノエステルまたはジエステルである、方法。

【請求項 2】

前記アルコールがブタノールまたはエタノールであり、前記 FDCA の誘導体が、それぞれ FDCA のブチルもしくはジブチルエステル、またはエチルもしくはジエチルエステルである、請求項 1 に記載の方法。

【請求項 3】

少なくとも 25% モルの収率を有する、請求項 1 または 2 に記載の方法。

【請求項 4】

前記アルコールが、ブタノール、エタノール、メタノール、およびプロパノールから選択され、

前記酸が硫酸であり、

前記接触が、70 より高い温度で生じ、かつ

それによって、それぞれ FDCA のブチルもしくはジブチルエステル、エチルもしくはジエチルエステル、メチルもしくはジメチルエステル、またはプロピルもしくはジプロピ

ルエステルが合成される、
請求項 1 に記載の方法。

【請求項 5】

前記接触が、150 を超える温度で気相中で生じる、請求項 4 に記載の方法。

【請求項 6】

5 - デヒドロ - 4 - デオキシグルカレート (DDG) の誘導体は無機酸と接触させて、
2, 5 - フランジカルボン酸 (FDCA) の誘導体を生成するステップ
を含む、FDCA の誘導体を合成するための方法であって、
該 DDG の誘導体が DDG のモノエステルまたはジエステルであり、かつ
該 FDCA の誘導体が FDCA のモノエステルまたはジエステルである、方法。

10

【請求項 7】

25 % モルを超える収率を有する、請求項 6 に記載の方法。

【請求項 8】

前記 DDG の誘導体が、メチル - DDG (DDG のモノメチルエステル)、エチル - DDG (DDG のモノエチルエステル)、ブチル - DDG (DDG のモノブチルエステル)、ジメチル DDG (DDG のジメチルエステル)、ジエチル - DDG (DDG のジエチルエステル)、およびジブチル DDG (DDG のジブチルエステル) からなる群から選択され、かつ前記 FDCA の誘導体が、それぞれ FDCA のメチルエステル、エチルエステル、ブチルエステル、ジメチルエステル、ジエチルエステル、またはジブチルエステルである、請求項 6 または 7 に記載の方法。

20

【請求項 9】

前記接触が、少なくとも 90 の温度で気相中で生じる、請求項 8 に記載の方法。

【請求項 10】

5 - デヒドロ - 4 - デオキシグルカレート (DDG) の誘導体は無機酸と接触させて、
2, 5 - フランジカルボン酸 (FDCA) の誘導体を生成するステップ；および
得られた FDCA の誘導体を脱エステル化して FDCA を得るステップ
を含む、FDCA を合成するための方法であって、
該 DDG の誘導体が DDG のモノエステルまたはジエステルであり、かつ
該 FDCA の誘導体が FDCA のモノエステルまたはジエステルである、方法。

【請求項 11】

前記 FDCA の誘導体を生成するステップが 25 % モルを超える収率を有する、請求項 10 に記載の方法。

30

【請求項 12】

前記 DDG の誘導体が、メチル - DDG (DDG のモノメチルエステル)、エチル - DDG (DDG のモノエチルエステル)、ブチル - DDG (DDG のモノブチルエステル)、ジメチル DDG (DDG のジメチルエステル)、ジエチル - DDG (DDG のジエチルエステル)、およびジブチル DDG (DDG のジブチルエステル) からなる群から選択され、かつ前記 FDCA の誘導体が、それぞれ FDCA のメチルエステル、エチルエステル、ブチルエステル、ジメチルエステル、ジエチルエステル、またはジブチルエステルである、請求項 10 または 11 に記載の方法。

40

【請求項 13】

前記接触が、少なくとも 90 の温度で気相中で生じる、請求項 12 に記載の方法。

【請求項 14】

70 より高い温度で 5 - デヒドロ - 4 - デオキシグルカレート (DDG) を無機酸と接触させて 2, 5 - フランジカルボン酸 (FDCA) を合成するステップ
を含む、FDCA を合成するための方法。

【請求項 15】

前記無機酸が硫酸である、請求項 14 に記載の方法。

【請求項 16】

120 を超える温度で気相中で 5 - デヒドロ - 4 - デオキシグルカレート (DDG)

50

を無機酸と接触させて 2, 5 - フランジカルボン酸 (F D C A) を合成するステップを含む、F D C A を合成するための方法。

【請求項 1 7】

前記無機酸が硫酸である、請求項 1 6 に記載の方法。

【請求項 1 8】

無水反応条件下で 5 - デヒドロ - 4 - デオキシグルカレート (D D G) を無機酸と接触させて 2, 5 - フランジカルボン酸 (F D C A) を合成するステップを含む、F D C A を合成するための方法。

【請求項 1 9】

前記接触が、80 を超える温度で生じる、請求項 1 8 に記載の方法。

【請求項 2 0】

D D G が、10 % 未満の水 (w / w) を含有する溶媒中に含まれる、請求項 1 8 または 1 9 に記載の方法。

【請求項 2 1】

D D G が、5 % 未満の水 (w / w) を含有する溶媒中に含まれる、請求項 2 0 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2012年8月21日に出願された米国特許仮出願第61/704,408号の恩恵を主張し、この米国特許仮出願は、表、図面、および特許請求の範囲を含めて全体を参照により援用する。

【0002】

配列表の援用

添付の配列表のデータを参照により本出願中に援用する。添付の S G I 1 6 6 0 __ 1 W O __ P C T __ S e q u e n c e L i s t i n g __ S T 2 5 という名称の配列表テキストファイルは、2013年8月__日に作成されたものであり、__K B である。このファイルは、Windows OS を使用するコンピュータにおいて Microsoft Word を用いて評価することができる。

【背景技術】

【0003】

発明の背景

近年、有機化学物質の生成のために再生可能な原料を用いる新規かつ有効な方法を特定するためのより多くの努力が注がれている。多量の下流化学物質の処理技術のうち、バイオマス由来の糖から付加価値のある化学物質への変換は、極めて重要であるとみなされている。詳細には、6炭素糖質、すなわち、フルクトースおよびグルコースなどのヘキソースは、自然に存在している単糖のうち最も豊富に存在するものとして広く認識されているため、化学原料として、適宜かつ経済的に利用することが可能である。

【0004】

糖からのフランおよびフラン誘導体の生成は、化学においておよび触媒反応研究において注目が集まっており、持続可能なエネルギー供給および化学物質生成を達成するための主要なルートの1つを提供する可能性を有するものと考えられている。実際、一体型バイオマス変換プロセスを備えた生物精製所内において利用可能な糖の脱水および/または酸化を行えば、広範囲のフランおよびフラン誘導体を含む多様な製品を大量に得ることが可能になる。

【0005】

最も商業的価値の高いフランのうち、フラン - 2, 5 - ジカルボン酸 (2, 5 - フランジカルボン酸としても知られる、本明細書中以下 F D C A と略称する) は、薬剤、殺虫剤、抗菌剤、香料、農業用化学物質を含むいくつかの産業において、およびポリマー材料、

10

20

30

40

50

例えばバイオプラスチック樹脂の広範囲の製造用途において多様な用途を有する貴重な中間体である。そのため、FDCAは、毎年世界中で生成されている最大量の石油化学物質の1つである石油ベースのモノマーであるテレフタル酸(TPA)に対する環境に優しい代替物であるとみなされている。実際、米国エネルギー省は、FDCAを、将来の「環境に優しい」化学を確立するための付加価値のある化学物質となる糖から生成された、優先度が12位である化合物の1つとして認定しており、そのため再生可能な中間体化学物質の「眠れる巨人」の1つとして指名されている(Werpy and Petersen, Top Value Added Chemicals from Biomass, US Department of Energy, Biomass, Vol 1, 2004 (非特許文献1))。

10

【0006】

FDCAの商業規模生成については多様な方法が提案されている(総説については、例えばTong et al., Appl. Catalysis A: General, 385, 1-13, 2010 (非特許文献2)を参照)ものの、現在のところ、FDCAの主要な産業合成は、中間体5-ヒドロキシメチルフルフルール(5-HMF)へのグルコースまたはフルクトースなどのヘキソースの化学的脱水の後にFDCAへの化学的酸化を行うことに依存している。しかし、現在の脱水を介したFDCA生成プロセスは非選択的であることが多いため、それらの形成直後で無い限り、不安定な中間体生成物がより安定した物質へ転換し得ることが報告されている。そのため、FDCAの生成および利用における主な技術的課題は、バイオマス由来の糖からの有効かつ選択的な脱水プロセスの開発である。

20

【0007】

したがって、石油ベース原料に代わる再生可能エネルギーとして機能するだけでなく、エネルギーおよび資本集約的技術が少量ですむ代替的手段による、この極めて重要な化合物および他の多数の化学物質および代謝物の生成方法の開発が望まれている。詳細には、糖脱水の選択的制御は極めて強力な技術になり得るため、広範囲のさらなる低コストの構成要素がもたらされ得る。

【先行技術文献】

【非特許文献】

【0008】

30

【非特許文献1】Werpy and Petersen, Top Value Added Chemicals from Biomass, US Department of Energy, Biomass, Vol 1, 2004

【非特許文献2】Tong et al., Appl. Catalysis A: General, 385, 1-13, 2010

【発明の概要】

【0009】

本発明は、1つ以上の酵素的経路の生成物を生成するための方法を提供する。本発明の方法において用いられる経路は、1つ以上の変換ステップ、例えば、グルロン酸からD-グルカレートへの酵素的変換(ステップ7)、5-ケトグルコネート(5-KGA)からL-イズロン酸への酵素的変換(ステップ15)、L-イズロン酸からイダル酸への酵素的変換(ステップ7b)、および5-ケトグルコネートから4,6-ジヒドロキシ2,5-ジケトヘキサノエート(2,5-DDH)への酵素的変換(ステップ16)を含む。いくつかの実施形態において、本発明の方法は、2,5-フランジカルボン酸(FDCA)を生成物として生成する。これらの方法は、酵素的変換および化学的変換の双方をステップとして含む。グルコースから5-デヒドロ-4-デオキシ-グルカレート(DDG)への変換のため、およびグルコースからFDCAへの変換のための多様な経路も提供される。これらの方法は、特異性および効率が高い反応を行う遺伝子操作された酵素の使用も用い得る。

40

【0010】

50

第1の態様において、本発明は、出発基質から酵素的経路または化学的経路の生成物を生成するための方法を提供する。経路は、以下の変換ステップのうちいずれか1つ以上を含み得る：グルロン酸からD-グルカレートへの酵素的変換（ステップ7）、5-ケトグルコネート（5-KGA）からL-イズロン酸への酵素的変換（ステップ15）、L-イズロン酸からイダル酸への酵素的変換（ステップ7b）、および、5-ケトグルコネートから4,6-ジヒドロキシ2,5-ジケトヘキサノエート（2,5-DDH）への酵素的変換（ステップ16）、1,5-グルコノラクトンからグルロノラクトン（guluronolactone）への酵素的変換（ステップ19）。

【0011】

一実施形態において、酵素的経路の生成物は、5-デヒドロ-4-デオキシ-グルカレート（DDG）である。様々な実施形態において、方法の基質はグルコースであり得、かつ生成物は5-デヒドロ-4-デオキシ-グルカレート（DDG）であり得る。方法は、D-グルコースから1,5-グルコノラクトンへの酵素的変換（ステップ1）、1,5-グルコノラクトンからグルロノラクトンへの酵素的変換（ステップ19）、グルロノラクトンからグルロン酸への酵素的変換（ステップ1B）、グルロン酸からD-グルカレートへの酵素的変換（ステップ7）、および、D-グルカレートから5-デヒドロ-4-デオキシ-グルカレート（DDG）への酵素的変換（ステップ8）のステップを含み得る。

【0012】

本発明の別の方法において、基質はグルコースであり、かつ生成物はDDGであり、かつ方法は、D-グルコースから1,5-グルコノラクトンへの変換（ステップ1）、1,5-グルコノラクトンからグルコン酸への変換（ステップ1a）、グルコン酸から5-ケトグルコネート（5-KGA）への変換（ステップ14）、5-ケトグルコネート（5-KGA）からL-イズロン酸への変換（ステップ15）、L-イズロン酸からイダル酸への変換（ステップ7b）、および、イダル酸からDDGへの変換（ステップ8a）のステップを含む。

【0013】

本発明の別の方法において、基質はグルコースであり、かつ生成物はDDGであり、かつ方法は、D-グルコースから1,5-グルコノラクトンへの変換（ステップ1）、1,5-グルコノラクトンからグルコン酸への変換（ステップ1a）、グルコン酸から5-ケトグルコネート（5-KGA）への変換（ステップ14）、5-ケトグルコネート（5-KGA）から4,6-ジヒドロキシ2,5-ジケトヘキサノエート（2,5-DDH）への変換（ステップ16）、4,6-ジヒドロキシ2,5-ジケトヘキサノエート（2,5-DDH）から4-デオキシ-5-トレオ-ヘキソスロースウロン酸（DTHU）への変換（ステップ4）、および、4-デオキシ-5-トレオ-ヘキソスロースウロン酸（DTHU）からDDGへの変換（ステップ5）のステップを含む。

【0014】

本発明の別の方法において、基質はグルコースであり、かつ生成物はDDGであり、かつ方法は、D-グルコースから1,5-グルコノラクトンへの変換（ステップ1）、1,5-グルコノラクトンからグルコン酸への変換（ステップ1a）、グルコン酸から5-ケトグルコネート（5-KGA）への変換（ステップ14）、5-ケトグルコネート（5-KGA）からL-イズロン酸への変換（ステップ15）、L-イズロン酸から4-デオキシ-5-トレオ-ヘキソスロースウロン酸（DTHU）への変換（ステップ7B）、および、4-デオキシ-5-トレオ-ヘキソスロースウロン酸（DTHU）からDDGへの変換（ステップ5）のステップを含む。

【0015】

本明細書中に開示される方法のいずれも、DDGを2,5-フラン-ジカルボン酸（FDCA）へ変換するステップをさらに含み得る。これらの方法のいずれかにおけるDDGからFDCAへの変換は、DDGを無機酸と接触させてDDGをFDCAへ変換することを含み得る。

【0016】

別の態様において、本発明は、誘導体化（エステル化）F D C Aを合成するための方法を提供する。方法は、60 を超える温度でD D Gをアルコール、無機酸と接触させて、誘導体化F D C Aを形成するステップを含む。異なる実施形態において、アルコールは、メタノール、ブタノール、またはエタノールである。

【0017】

別の態様において、本発明は、F D C Aの誘導体を合成するための方法を含む提供。方法は、D D Gをアルコール、無機酸、および共溶媒と接触させてD D Gの誘導体を生成するステップ、任意選択的にD D Gの誘導体を精製するステップ、およびD D Gの誘導体を無機酸と接触させてF D C Aの誘導体を生成するステップを含む。無機酸は硫酸であり得、アルコールはエタノールまたはブタノールであり得る。様々な実施形態において、共溶媒は、T H F、アセトン、アセトニトリル、エーテル、酢酸ブチル、ジオキサン、クロロホルム、塩化メチレン、1, 2 - ジクロロエタン、ヘキサン、トルエン、およびキシレンのいずれかであり得る。

10

【0018】

一実施形態において、D D Gの誘導体はジエチルD D Gであり、F D C Aの誘導体はジエチルF D C Aであり、別の実施形態において、D D Gの誘導体はジブチルD D Gであり、F D C Aの誘導体はジブチルF D C Aである。

【0019】

別の態様において、本発明は、F D C Aを合成するための方法を提供する。方法は、気相中でD D Gを無機酸と接触させるステップを含む。

20

【0020】

別の態様において、本発明は、F D C Aを合成するための方法を提供する。方法は、120 を超える温度でD D Gを無機酸と接触させるステップを含む。

【0021】

別の態様において、本発明は、F D C Aを合成するための方法を提供する。方法は、無水反応条件下でD D Gを無機酸と接触させるステップを含む。

[本発明1001]

出発基質から酵素的経路または化学的経路の生成物を生成するための方法であって、該経路が、

グルロン酸からD - グルカレートへの酵素的変換（ステップ7）、
5 - ケトグルコネート（5 - K G A）からL - イズロン酸への酵素的変換（ステップ15）

30

、
L - イズロン酸からイダル酸への酵素的変換（ステップ7b）、および
5 - ケトグルコネートから4, 6 - ジヒドロキシ2, 5 - ジケトヘキサノエート（2, 5 - D D H）への酵素的変換（ステップ16）、

1, 5 - グルコノラクトンからグルコノラクトン（g u l u r o n o - l a c t o n e）への酵素的変換（ステップ19）

からなる群から選択される1つ以上の変換ステップを含む、方法。

[本発明1002]

前記1つ以上の変換ステップが、グルロン酸からD - グルカレートへの酵素的変換（ステップ7）である、本発明1001の方法。

40

[本発明1003]

前記1つ以上の変換ステップが、5 - ケトグルコネート（5 - K G A）からL - イズロン酸への酵素的変換（ステップ15）である、本発明1001の方法。

[本発明1004]

前記1つ以上の変換ステップが、L - イズロン酸からイダル酸への酵素的変換（ステップ7b）である、本発明1001の方法。

[本発明1005]

前記1つ以上の変換ステップが、5 - ケトグルコネートから4, 6 - ジヒドロキシ2, 5 - ジケトヘキサノエート（2, 5 - D D H）への酵素的変換（ステップ16）である、本発明1001

50

の方法。

[本発明1006]

前記1つ以上の変換ステップが、1,5-グルコノラクトンからグルロノラクトンへの酵素的変換(ステップ19)である、本発明1001の方法。

[本発明1007]

前記酵素的経路の前記生成物が5-デヒドロ-4-デオキシ-グルカレート(DDG)である、本発明1001の方法。

[本発明1008]

前記基質がグルコースでありかつ前記生成物が5-デヒドロ-4-デオキシ-グルカレート(DDG)であり、前記方法が、

10

D-グルコースから1,5-グルコノラクトンへの酵素的変換(ステップ1)、
1,5-グルコノラクトンからグルロノラクトンへの酵素的変換(ステップ19)、
グルロノラクトンからグルロン酸への酵素的変換(ステップ1B)、
グルロン酸からD-グルカレートへの酵素的変換(ステップ7)、
D-グルカレートから5-デヒドロ-4-デオキシ-グルカレート(DDG)への酵素的変換(ステップ8)

のステップを含む、本発明1001の方法。

[本発明1009]

前記基質がグルコースでありかつ前記生成物がDDGであり、前記方法が、

20

D-グルコースから1,5-グルコノラクトンへの変換(ステップ1)、
1,5-グルコノラクトンからグルコン酸への変換(ステップ1a)、
グルコン酸から5-ケトグルコネート(5-KGA)への変換(ステップ14)、
5-ケトグルコネート(5-KGA)からL-イズロン酸への変換(ステップ15)、
L-イズロン酸からイダル酸への変換(ステップ7b)、および
イダル酸からDDGへの変換(ステップ8a)

のステップを含む、本発明1001の方法。

[本発明1010]

前記基質がグルコースでありかつ前記生成物がDDGであり、前記方法が、

30

D-グルコースから1,5-グルコノラクトンへの変換(ステップ1)、
1,5-グルコノラクトンからグルコン酸への変換(ステップ1a)、
グルコン酸から5-ケトグルコネート(5-KGA)への変換(ステップ14)、
5-ケトグルコネート(5-KGA)から4,6-ジヒドロキシ2,5-ジケトヘキサノエート(2,5-DDH)への変換(ステップ16)、
4,6-ジヒドロキシ2,5-ジケトヘキサノエート(2,5-DDH)から4-デオキシ-5-トレオ-ヘキソスロースウロン酸(DTHU)への変換(ステップ4)、および
4-デオキシ-5-トレオ-ヘキソスロースウロン酸(DTHU)からDDGへの変換(ステップ5)

のステップを含む、本発明1001の方法。

[本発明1011]

前記基質がグルコースでありかつ前記生成物がDDGであり、前記方法が、

40

D-グルコースから1,5-グルコノラクトンへの変換(ステップ1)、
1,5-グルコノラクトンからグルコン酸への変換(ステップ1a)、
グルコン酸から5-ケトグルコネート(5-KGA)への変換(ステップ14)、
5-ケトグルコネート(5-KGA)からL-イズロン酸への変換(ステップ15)、
L-イズロン酸から4-デオキシ-5-トレオ-ヘキソスロースウロン酸(DTHU)への変換(ステップ7c)、および

4-デオキシ-5-トレオ-ヘキソスロースウロン酸(DTHU)からDDGへの変換(ステップ5)

のステップを含む、本発明1001の方法。

[本発明1012]

50

前記 D D G を 2, 5 - フラン - ジカルボン酸 (F D C A) へ変換するステップをさらに含む、本発明1008の方法。

[本発明1013]

前記 D D G を 2, 5 - フラン - ジカルボン酸 (F D C A) へ変換するステップをさらに含む、本発明1009の方法。

[本発明1014]

前記 D D G を 2, 5 - フラン - ジカルボン酸 (F D C A) へ変換するステップをさらに含む、本発明1010の方法。

[本発明1015]

前記 D D G を 2, 5 - フラン - ジカルボン酸 (F D C A) へ変換するステップをさらに含む、本発明1011の方法。

10

[本発明1016]

前記 D D G を F D C A へ変換するステップが、D D G を酸と接触させて該 D D G を F D C A へ変換することを含む、本発明1012の方法。

[本発明1017]

前記 D D G を F D C A へ変換するステップが、D D G を酸と接触させて該 D D G を F D C A へ変換することを含む、本発明1013の方法。

[本発明1018]

前記 D D G を F D C A へ変換するステップが、D D G を酸と接触させて該 D D G を F D C A へ変換することを含む、本発明1014の方法。

20

[本発明1019]

前記 D D G を F D C A へ変換するステップが、D D G を酸と接触させて該 D D G を F D C A へ変換することを含む、本発明1015の方法。

[本発明1020]

60 を超える温度で D D G をアルコール、無機酸と接触させて、F D C A を形成するステップを含む、F D C A の誘導体を合成するための方法。

[本発明1021]

前記アルコールがブタノールまたはエタノールである、本発明1020の方法。

[本発明1022]

30

少なくとも25%モルの収率を有する、本発明1020の方法。

[本発明1023]

D D G をアルコール、無機酸、および任意選択的に共溶媒と接触させて、D D G の誘導体を生成するステップを含む、D D G の誘導体を合成する方法。

[本発明1024]

(a) 前記アルコールがエタノールまたはブタノールであり、

(b) 前記無機酸が硫酸であり、かつ

(c) 前記共溶媒が、T H F、アセトン、アセトニトリル、エーテル、酢酸エチル、酢酸ブチル、ジオキサン、クロロホルム、塩化メチレン、1, 2 - ジクロロエタン、ヘキサン、ヘプタン、トルエン、四塩化炭素、石油エーテル、およびキシレンからなる群から選択される、

40

本発明1023の方法。

[本発明1025]

D D G の誘導体を無機酸と接触させて、F D C A の誘導体を生成するステップを含む、F D C A の誘導体を合成するための方法。

[本発明1026]

25%モルを超える収率を有する、本発明1025の方法。

[本発明1027]

前記誘導体が、メチル - D D G、エチル - D D G、ブチル - D D G、ジメチル D D G、

50

ジエチル - D D G、およびジブチル D D G からなる群から選択される D D G である、本発明1026の方法。

[本発明1028]

前記 F D C A の誘導体を脱エステル化して F D C A を得るステップをさらに含む、本発明1025の方法。

[本発明1029]

前記 F D C A の誘導体を重合するステップをさらに含む、本発明1025の方法。

[本発明1030]

気相中で D D G を無機酸と接触させるステップを含む、F D C A を合成するための方法。

10

[本発明1031]

120 を超える温度で D D G を無機酸と接触させるステップを含む、F D C A を合成するための方法。

[本発明1032]

無水反応条件下で D D G を無機酸と接触させるステップを含む、F D C A を合成するための方法。

[本発明1033]

出発基質から酵素的経路または化学的経路の生成物を生成するための方法であって、該経路が、

D T H U から D D G への変換 (ステップ5)、

20

グルコン酸からグルロン酸への変換 (ステップ6)、

D E H U から D D H への変換 (ステップ7A)、

グルロン酸から D E H U への変換 (ステップ17A)

からなる群から選択される1つ以上の変換ステップを含む、方法。

[本発明1034]

前記基質がグルコースでありかつ前記生成物が D D G であり、前記方法が、

D - グルコースから1,5 - グルコノラクトンへの変換 (ステップ1)、

1,5 - グルコノラクトンからグルコン酸への変換 (ステップ1a)、

グルコン酸から3 - デヒドロ - グルコン酸 (D H G) への変換 (ステップ2)、

3 - デヒドロ - グルコン酸 (D H G) から4,6 - ジヒドロキシ2,5 - ジケトヘキサノエート (2,5 - D D H) への変換 (ステップ3)、

30

2,5 D D H から4 - デオキシ - 5 - トレオ - ヘキソスロースウロン酸 (D T H U) への変換 (ステップ4)、

D T H U から D D G への変換 (ステップ5)

のステップを含む、本発明1033の方法。

[本発明1035]

前記基質がグルコースでありかつ前記生成物が D D G であり、前記方法が、

D - グルコースから1,5 - グルコノラクトンへの変換 (ステップ1)、

1,5 - グルコノラクトンからグルコン酸への変換 (ステップ1a)、

グルコン酸からグルロン酸への変換 (ステップ6)、

40

グルロン酸からグルカレートへの変換 (ステップ7)、

グルカレートから D D G への変換 (ステップ8)

のステップを含む、本発明1033の方法。

[本発明1036]

前記基質がグルコースでありかつ前記生成物が D D G であり、前記方法が、

D - グルコースから1,5 - グルコノラクトンへの変換 (ステップ1)、

1,5 - グルコノラクトンからグルコン酸への変換 (ステップ1a)、

グルコン酸から5 - ケトグルコネート (5 - K G A) への変換 (ステップ14)、

5 - ケトグルコネート (5 - K G A) から4,6 - ジヒドロキシ2,5 - ジケトヘキサノエート (2,5 - D D H) への変換 (ステップ16)、

50

4, 6 - ジヒドロキシ2, 5 - ジケトヘキサノエート (2, 5 - D D H) から4 - デオキシ - 5 -
 - トレオ - ヘキソスロースウロン酸 (D T H U) への変換 (ステップ4)、および
 4 - デオキシ - 5 - トレオ - ヘキソスロースウロン酸 (D T H U) からD D G への変換 (ステップ5)
 のステップを含む、本発明1033の方法。

[本発明1037]

前記基質がグルコースでありかつ前記生成物がD D Gであり、前記方法が、
 D - グルコースから1, 5 - グルコノラクトンへの変換 (ステップ1)、
 1, 5 - グルコノラクトンからグルコン酸への変換 (ステップ1a)、
 グルコン酸から5 - ケトグルコネート (5 - K G A) への変換 (ステップ14)、
 5 - ケトグルコネート (5 - K G A) からL - イズロン酸への変換 (ステップ15)、
 L - イズロン酸から4 - デオキシ - 5 - トレオ - ヘキソスロースウロン酸 (D T H U) へ
 の変換 (ステップ7B)、
 4 - デオキシ - 5 - トレオ - ヘキソスロースウロン酸 (D T H U) からD D G への変換 (ステップ5)
 のステップを含む、本発明1033の方法。

[本発明1038]

前記基質がグルコースでありかつ前記生成物がD D Hであり、前記方法が、
 D - グルコースから1, 5 - グルコノラクトンへの変換 (ステップ1)、
 1, 5 - グルコノラクトンからグルロン酸ラクトンへの変換 (ステップ19)、
 グルロン酸ラクトンからグルロン酸への変換 (ステップ1B)、
 グルロン酸からD E H U への変換 (ステップ17A)、
 D E H U からD D H への変換 (ステップ7A)
 のステップを含む、本発明1033の方法。

[本発明1039]

前記基質がグルコースでありかつ前記生成物がD D Hであり、前記方法が、
 D - グルコースから1, 5 - グルコノラクトンへの変換 (ステップ1)、
 1, 5 - グルコノラクトンからグルコン酸への変換 (ステップ1a)、
 グルコン酸からグルロン酸への変換 (ステップ6)、
 グルロン酸からD E H U への変換 (ステップ17A)、
 D E H U からD D H への変換 (ステップ7A)
 のステップを含む、本発明1033の方法。

[本発明1040]

前記1つ以上の変換ステップが、D T H U からD D G への変換 (ステップ5) である、本発明1033の方法。

[本発明1041]

前記1つ以上の変換ステップが、グルコン酸からグルロン酸への変換 (ステップ6) である、本発明1033の方法。

[本発明1042]

前記1つ以上の変換ステップが、D E H U からD D H への変換 (ステップ7A) である、本発明1033の方法。

[本発明1043]

前記1つ以上の変換ステップが、グルロン酸からD E H U への変換 (ステップ17A) である、本発明1033の方法。

【図面の簡単な説明】

【0022】

【図1】タンパク質474、475、および476の粗溶解物および精製酵素の電気泳動ゲルである。

【図2】図2A ~ 2Hはそれぞれルート1、2、2A、2C、2D、2E、および2Fの経路の模式図である。

10

20

30

40

50

【図3】図3A～3Cはそれぞれルート3、4、および5の経路の模式図である。

【図4】pSGI-359によるDHGの生成のためのグルコン酸デヒドラターゼを用いたグルコネートの脱水に関するHPLC-MS分析である。

【図5】グルコン酸デヒドラターゼの活性の測定のためのセミカルバジドアッセイプロットの図表である。

【図6】図6A～図6Bは、本発明の3つの酵素によるグルクロネートおよびイズロネートの酸化のためのラインウィーバー-バークプロットを提供する。

【図7】図7Aは、EC5.3.1.17ファミリー中の酵素DTHUISイソメラーゼを用いた5KGAおよびイズロネートの異性化のためのHPLC分析の時点の結果を示す。対照：対照として、死滅酵素は、加熱不活性化した酵素である。Med B1は、イソメラーゼ添加無しの反応を指す。時点、x軸1=0.5時間；2=1；3=2時間；4=16時間である。図7Bは、EC5.3.1.17ファミリー中の酵素を用いた5KGAおよびイズロネートの異性化の時点のHPLC分析を示す。対照：対照として、死滅酵素は、加熱不活性化した酵素である。Med B1は、イソメラーゼ添加無しの反応を指す。時点、X軸：1=0時間；2=1時間；3=2時間；4=17時間である。

【図8】EC5.3.1.n1ファミリー中の酵素を用いた5KGAおよびイズロネートの異性化に関する生成物形成を示す。本データは、酵素的アッセイ法から得られた。

【図9】2,5-DDHの形成のHPLC分析と、5KGA濃度の経時的低減とを示す。2,5-DDHの合計イオンカウントが示されている。

【図10】1,5-グルコノラクトンからのグルロン酸ラクトンの生成を示すHPLC-MSクロマトグラムである。真正グルロン酸のトレースのオーバーレイが示されている。

【図11】スキーム6の反応経路の模式図である。

【図12】図12Aおよび図12Bは、それぞれ、5-KGAおよびDDG反応生成物を示すLC-MSクロマトグラムである。

【図13】FDCAおよびFDCAジブチルエステル誘導体反応生成物を示すLC-MSクロマトグラムである。

【図14】図14Aは、DDGとエタノールとの反応からのジエチル-FDCA合成の粗反応試料のGC-MS分析である。単一ピークは、ジエチル-FDCAに対応していた。図14bは、DDGとエタノールとの反応からの主要生成物のMS断片化である。

【図15】図15Aは、DDGとエタノールとの反応からのジエチル-FDCA合成の粗反応試料のGC-MS分析である。単一ピークは、ジエチル-FDCAに対応していた。図15bは、DDGとエタノールとの反応からの主要生成物のMS断片化である。

【図16】DTHUISからのFDCAおよびその誘導体の合成の模式図である。

【図17】スキーム1の模式図である。グルコースからのDDGの無細胞酵素合成。酵素を以下に示す：ST-1：グルコースオキシダーゼ；ST-1A：加水分解-化学物質；ST-14：グルコン酸デヒドロゲナーゼ(pSGI-504)；ST-15：5-デヒドロ-4-デオキシ-D-グルクロン酸イソメラーゼ(DTHUIS、pSGI-434)；ST-7B：ウロン酸デヒドロゲナーゼ(UroDH、pSGI-476)；ST-8A：グルカル酸デヒドラターゼ(GlucDH、pSGI-353)；ST-A：NAD(P)Hオキシダーゼ(NADH_OX、pSGI-431)；ST-B：カタラーゼ。図17Bは、最初の3時間における、HPLCによって分析された反応中間体の濃度を示す。双方の反応において、DDG形成が示される。

【発明を実施するための形態】

【0023】

本発明は、酵素的経路の生成物を生成するための方法を提供する。これらの方法は、基質から生成物への酵素的変換を含み得る。本発明の酵素的および化学的経路を用いることにより、広範な多様な生成物を効率的かつ経済的に合成することが可能になる。本発明の方法および経路によって生成することが可能な1つの生成物として、本発明に従って商業規模で生成することが可能な2,5-フラニルジカルボン酸(FDCA)がある。方法は、本明細書中に開示される、1つ以上の酵素的基質および/または化学的基質から生成物

10

20

30

40

50

への変換ステップを含み得る。

【0024】

本発明の経路は、1つ以上のステップを含む。本発明の経路のステップは、正反応または逆反応を含み得て、すなわち、基質Aは中間体Bおよび生成物Cへ変換される一方、逆反応においては、基質Cは中間体Bおよび生成物Aへ変換されることが理解される。これらの方法において、他に特記無き限り、正反応および逆反応をステップとして説明する。

【0025】

これらの方法は、酵素的経路であり得る経路の生成物を生成するステップを含む。いくつかの実施形態において、これらの経路は、1つ以上の化学的ステップを含み得る。これらの方法は、基質から生成物への変換の、1つ以上の酵素的および/または化学的変換ステップを含む。方法に含まれ得るステップは、例えば、グルロン酸からD-グルカレートへの酵素的変換(ステップ7)、L-イズロン酸からイダル酸への酵素的変換(7B)、L-イズロン酸から4-デオキシ-5-トレオ-ヘキサスロースウロン酸(DTHU)への酵素的変換(7B)、5-ケトグルコネート(5-KGA)からL-イズロン酸への酵素的変換(ステップ15)、L-イズロン酸からイダル酸への酵素的変換(ステップ7B)、5-ケトグルコネートから4,6-ジヒドロキシ2,5-ジケトヘキサノエート(2,5-DDH)への酵素的変換(ステップ16)、1,5-グルコノラクトンからグルコノラクトンへの酵素的変換(ステップ19)の任意の1つ以上を含む。上記のステップの任意の1つ以上が、本発明の方法または経路に含まれ得る。酵素的ステップまたは経路は、ステップの進行のために酵素を反応における触媒として必要とするステップまたは経路である。化学的ステップは、酵素を反応における触媒として用いる必要無く行うことが可能である。方法に記載されるステップの任意の1つ以上は、酵素的ステップであり得る。いくつかの実施形態において、経路の各ステップは全て酵素的ステップであり、他の実施形態において、経路における1つ以上のステップは化学的ステップである。

【0026】

いくつかの実施形態において、方法のいずれも、変換を行う酵素を含む反応混合物への反応の基質の添加を含むステップを含み得る。よって、グルロン酸をD-グルカレートへ変換する方法(ステップ7)は、反応混合物への、出発基質としてのグルロン酸の添加を含み得、L-イズロン酸からイダル酸への酵素的変換(7B)は、反応混合物への、出発基質としてのL-イズロン酸の添加を含み得、L-イズロン酸から4-デオキシ-5-トレオ-ヘキサスロースウロン酸(DTHU)への酵素的変換(7B)は、反応混合物への、出発基質としてのDTHUの添加を含み得る。方法のいずれにも含まれ得る別のステップは、反応混合物から反応生成物を精製するステップである。よって、D-グルカレートまたはL-イズロン酸、またはイダル酸、または4,6-ジヒドロキシ2,5-ジケトヘキサノエートを精製するステップは、本明細書中に記載される方法のいずれにも含まれ得る。本明細書中に開示される方法のいずれも、反応混合物からDDGまたはFDCAを分離または精製するステップを含み得る。

【0027】

方法において用いられる反応混合物は、酵素的変換を行う1つ以上の酵素を含む細胞の細胞溶解物であり得るが、反応混合物の形成のためにユーザが添加した成分を含む反応混合物であってもよいが、あるいは、細胞溶解物から精製された成分を含んでもよいが、あるいは細胞生体触媒全体中に含めてもよい。

【0028】

様々な実施形態において、本発明の方法は、グルコースからDDGへ変換する方法、またはグルコースからFDCAへ変換する方法、またはグルコースからDTHUもしくはDEHUへ変換する方法、あるいはDDGからFDCAへ変換するための方法である。これらの方法は、方法において出発基質を生成物へ変換するステップを含み得る。出発基質は、方法を開始すると考えられる化学物質であり、生成物は、方法の最終生成物とみなされる化学物質である。中間体とは、(一時的にであれまたは恒久的にであれ)方法において生成される、出発基質と生成物との間に存在する化学物質である。様々な実施形態におい

て、本発明の方法および経路は、約4個または約5個の中間体または4～5個の中間体、または約3個の中間体または3～5個の中間体、または6個未満または7個未満または8個未満または9個未満または10個未満または15個未満または20個未満の中間体を有し、すなわち、これらの値は出発基質または最終生成物の数に含まれない。

【0029】

本発明は、高い収率を有する、グルコースからF D C Aおよび/またはD D Gを生成する方法を提供する。理論的収量とは、反応が理想的な条件下で完了した場合に形成されるであろう生成物の量である。異なる実施形態において、本発明の方法は、D D Gをグルコース、フルクトース、またはガラクトースから生成し、理論的収量が少なくとも50%モル、または少なくとも60%モルまたは少なくとも70%モル、または少なくとも80%モル、少なくとも90%モルまたは少なくとも95%モルまたは少なくとも97%モルまたは少なくとも98%モルまたは少なくとも99%モル、あるいは理論的収量が100%モルである。本発明の方法はまた、少なくとも80%または少なくとも90%または少なくとも99%、または100%の炭素保持で生成物を提供し、すなわち、初期基質中に存在する特定の炭素原子は、方法の最終生成物中に記載のパーセンテージで存在する。いくつかの実施形態において、方法は、グルコースから脱水反応を介してD D Gおよび/またはF D C Aを生成する。

【0030】

合成ルート

本発明はまた、所望の生成物を合成および生成するための特定の経路を提供する。以下の記載のルートまたは経路のいずれも、グルコースで開始し得、所望の生成物へ流れ得る。いくつかの実施形態において、D - グルコースは出発基質であり、経路の任意の中間体または最終生成物への経路の方向は下流方向とみなされ、グルコースへの反対方向は上流方向と見なされる。ルートまたは経路は、下流方向または上流方向に流れ得ることが認識される。また、グルコース、フルクトース、ガラクトース、またはいずれかの経路中の任意の中間体も、本発明の方法における出発基質であり得、D D G、F D C A、または本発明のルートまたは経路のいずれか中の任意の中間体も、本発明の方法の最終生成物であり得ることが理解される。よって、開示の方法は、任意の出発基質または中間体を開示のルートまたは経路のいずれか中のステップの1つ以上を用いて、開示のルートまたは経路中の任意の最終生成物または中間体へ変換するための本発明のルートまたは経路のいずれか中の開示の任意の1つ以上のステップを含む。よって、例えば、方法は、グルコースからD D Gへ、もしくはグルコースからグルロン酸へ、もしくはグルコースからガラクトレートへ、もしくはグルコースからD T H Uへ、もしくはグルコースからD E H Uへ変換するための方法、またはグルコースからグルロン酸への変換するための方法、またはグルコースからイズロン酸への変換するための方法、またはグルコースからイダル酸への変換するための方法、またはグルコースからグルカル酸への変換するための方法、またはガラクトレートからD D Gへの変換するための方法、またはグルロン酸からD - グルカレートへの変換するための方法、または5 - K G AからL - イズロン酸への変換するための方法、またはL - イズロン酸からイダル酸への変換するための方法、または5 - K G Aから2, 5 - D D H、もしくはD T H Uへの変換するための方法、またはD H GからD E H Uへの変換するための方法であり得る。これらの実施形態において、方法は、方法において開示されるステップと、出発基質としてのグルコースから関連最終生成物への本発明の経路とを用いる。

【0031】

ルート1を図2Aに示す。ルート1は、D - グルコース（または経路中の任意の中間体）を一連の示されるステップを介して酵素的経路を介して5 - デヒドロ - 4 - デオキシ - グルカレート（D D G）へ変換する。ルート1は、1, 5 - グルコノラクトン、グルコン酸、3 - デヒドロ - グルコン酸（D H G）、4, 6 - ジヒドロキシ2, 5 - ジケトヘキサノエート（2, 5 - D D H）、および4 - デオキシ - L - トレオ - ヘキサスロースウロン酸（D T H U）を中間体として有しかつD D Gを最終生成物として有する回路を介して、

D - グルコースを D D G へ変換する。経路のいずれについても、図示しないさらなる中間体が存在し得る。これらのステップは、D - グルコースから 1, 5 - グルコノラクトンへの酵素的変換（ステップ 1）と、1, 5 - グルコノラクトンからグルコン酸への酵素的変換（ステップ 1 A）と、グルコン酸から 3 - デヒドロ - グルコン酸（D H G）への酵素的変換（ステップ 2）と、3 - デヒドロ - グルコン酸（D H G）から 4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート（2, 5 - D D H）への酵素的変換（ステップ 3）と、4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート（2, 5 - D D H）から 4 - デオキシ - L - トレオ - ヘキソスロースウロン酸（D T H U）への酵素的変換（ステップ 4）と、4 - デオキシ - L - トレオ - ヘキソスロースウロン酸（D T H U）から 5 - デヒドロ - 4 - デオキシグルカレート（D D G）への酵素的変換（ステップ 5）とである。ルート 1 はまた、経路中のグルコースまたは任意の中間体が最終生成物としての任意の他の下流中間体へ変換されるサブルートを含み、生成物サブルートへの各基質は、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

10

【0032】

ルート 2 を図 2 B に示し、ルート 2 は、D - グルコースを D D G へ変換する。ルート 2 経路中のステップは、D - グルコースから 1, 5 - グルコノラクトンへの酵素的変換（ステップ 1）と、1, 5 - グルコノラクトンからグルコン酸への酵素的変換（ステップ 1 A）と、グルコン酸からグルロン酸への酵素的変換（ステップ 6）と、グルロン酸から D - グルカレートへの酵素的変換（ステップ 7）と、D - グルカレートから D D G への酵素的変換（ステップ 8）とである。ルート 2 はまた、経路中のグルコースまたは任意の中間体が最終生成物としての任意の他の下流中間体へ変換されるサブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

20

【0033】

ルート 2 A を図 2 C に示す。ルート 2 A におけるステップは、D - グルコースから 1, 5 - グルコノラクトンへの酵素的変換（ステップ 1）と、1, 5 - グルコノラクトンからグルロン酸ラクトンへの酵素的変換（ステップ 1 9）と、グルロン酸ラクトンからグルロン酸への酵素的変換（ステップ 1 B）と、グルロン酸から D - グルカレートへの酵素的変換（ステップ 7）と、D - グルカレートから 5 - デヒドロ - 4 - デオキシ - グルカレート（D D G）への酵素的変換（ステップ 8）とである。ルート 2 A はまた、出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

30

【0034】

ルート 2 B を図 2 D に示す。ルート 2 B におけるステップは、D - グルコースからグルコン酸への酵素的変換（ステップ 1 および 1 A）と、グルコン酸から 5 - ケトグルコネート（5 - K G A）への酵素的変換（ステップ 1 4）と、5 - K G A から L - イズロン酸への酵素的変換（ステップ 1 5）と、L - イズロン酸からイダル酸への酵素的変換（ステップ 7 B）と、イダル酸から D D G への酵素的変換（ステップ 8 A）とである。ルート 2 B はまた、出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含む。各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

40

【0035】

ルート 2 C を図 2 E に示す。ルート 2 C におけるステップは、D - グルコースからグルコン酸への酵素的変換（ステップ 1 および 1 A）と、グルコン酸から 5 - ケトグルコネート（5 - K G A）への酵素的変換（ステップ 1 4）と、5 - K G A から 4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート（2, 5 - D D H）への酵素的変換（ステップ 1 6）と、4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート（2, 5 - D D H）から 4 - デオキシ - 5 - トレオ - ヘキソスロースウロン酸（D T H U）への酵素的変換（ステップ 4）と、D T H U から D D G への酵素的変換（ステップ 5）とである。ルート 2 C はまた、

50

出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

【 0 0 3 6 】

ルート 2 D を図 2 F に示す。ルート 2 D におけるステップは、D - グルコースからグルコン酸への酵素的変換（ステップ 1 および 1 A ）と、グルコン酸から 5 - ケトグルコネート（ 5 - K G A ）への酵素的変換（ステップ 1 4 ）と、5 - K G A からイズロン酸への酵素的変換（ステップ 1 5 ）と、L - イズロン酸から D T H U への酵素的変換（ステップ 1 7 ）と、D T H U から D D G への酵素的変換（ステップ 5 ）とである。ルート 2 D はまた、出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

10

【 0 0 3 7 】

ルート 2 E を図 2 G に示す。ルート 2 D におけるステップは、D - グルコースから 1 , 5 - グルコノラクトンへの酵素的変換（ステップ 1 ）と、1 , 5 - グルコノラクトンからグルロン酸ラクトンへの酵素的変換（ステップ 1 9 ）と、グルロン酸ラクトンからグルロン酸への酵素的変換（ステップ 1 B ）と、グルロン酸から 4 - デオキシ - エリトロ - ヘキソスロースウロン酸（ D E H U ）への酵素的変換（ステップ 1 7 A ）と、D E H U から 3 - デオキシ - D - エリトロ - 2 - ヘキスロサル酸（ D D H ）への酵素的変換（ステップ 7 A ）とである。ルート 2 E はまた、出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

20

【 0 0 3 8 】

ルート 2 F を図 2 H に示す。ルート 2 F におけるステップは、D - グルコースからグルコン酸への酵素的変換（ステップ 1 および 1 A ）と、グルコン酸からグルロン酸への酵素的変換（ステップ 6 ）と、グルロン酸から 4 - デオキシ - エリトロ - ヘキソスロースウロン酸（ D E H U ）への酵素的変換（ステップ 1 7 ）と、D E H U から 3 - デオキシ - D - エリトロ - 2 - ヘキスロサル酸（ D D H ）への酵素的変換（ステップ 7 A ）とである。ルート 2 F はまた、出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

30

【 0 0 3 9 】

ルート 3 を図 3 A に示す。ルート 3 におけるステップは、D - グルコースからグルコン酸への酵素的変換（ステップ 1 および 1 A ）と、グルコン酸から 3 - デヒドロ - グルコン酸（ D H G ）への酵素的変換（ステップ 2 ）と、D H G から 4 - デオキシ - エリトロ - ヘキソスロースウロン酸（ D E H U ）への酵素的変換（ステップ 6 A ）と、D E H U から D D G への酵素的変換（ステップ 7 A ）とである。ルート 3 はまた、出発基質としての経路中のグルコースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

40

【 0 0 4 0 】

ルート 4 を図 3 B に示す。ルート 4 におけるステップは、D - グルコースから a - D - グルコ - ヘキシジアルド - 1 , 5 - ピラノースへの酵素的変換（ステップ 9 ）と、a - D - グルコ - ヘキシジアルド - 1 , 5 - ピラノースから a - D - グルコピラヌロン酸への酵素的変換（ステップ 1 0 ）と、a - D - グルコピラヌロン酸から D - グルカル酸 1 , 5 - ラクトンへの酵素的変換（ステップ 1 1 ）と、D - グルカル酸 1 , 5 - ラクトンから D - グルカレートへの酵素的変換（ステップ 1 C ）と、D - グルカレートから D D G への酵素的変換（ステップ 8 ）とである。ルート 4 はまた、出発基質としての経路中のグルコース

50

または任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

【0041】

ルート5を図3Cに示す。ルート5におけるステップは、D - ガラクトースからD - ガラクト - ヘキソジアルドースへの酵素的変換（ステップ9A）と、D - ガラクト - ヘキソジアルドースからガラクトツロネートへの酵素的変換（ステップ10A）と、ガラクトツロネートからガラクタレートへの酵素的変換（ステップ11A）と、ガラクタレートからDDGへの酵素的変換（ステップ13）とである。ルート5はまた、出発基質としての経路中のガラクトースまたは任意の中間体が、最終生成物としての任意の他の下流中間体へ変換される、サブルートを含み、各サブルートは、本明細書中にあたかも完全に記載されているかのように開示されているものとしてみなされる。

10

【0042】

酵素的ステップ

本明細書中に概要を示す方法のステップを行うことが可能な広範な様々な酵素（およびこれらの酵素をコードする核酸）が開示される。本明細書中に開示される本発明のステップを行うための酵素のファミリーおよびクラスに加えて、任意の酵素または本明細書中に開示される酵素のクラスのメンバーに対して配列同一性を有するさらなる酵素（または酵素をコードする核酸）も、任意の酵素または本明細書中に開示される酵素のクラスのメンバーに対して少なくとも40%または少なくとも50%または少なくとも60%または少なくとも70%または少なくとも80%または少なくとも90%または少なくとも95%または少なくとも97%または少なくとも98%または少なくとも99%の配列同一性を有するものであれば、本発明において有用である。アミノ酸またはヌクレオチド配列に対するパーセント単位の配列同一性または相同性は、本明細書中、最大パーセントの同一性または相同性を達成するように配列を最大パーセントの同一性のためにアライメントさせ必要ならばギャップを導入した後の、公知のポリペプチドと同一である候補配列中のアミノ酸またはヌクレオチド残基のパーセンテージとして定義される。ヌクレオチドまたはアミノ酸配列レベルにおける相同性または同一性は、配列類似検索に合わせて個別調整された当該分野において公知の方法によって決定することができ、この公知の方法は非限定的に、プログラムであるblastp、blastn、blastx、tblastn、およびtblastxによって用いられるアルゴリズムを用いたBLAST（Basic Local Alignment Search Tool）分析（Altschul（1997）、Nucleic Acids Res. 25, 3389 - 3402、およびKarlin（1990）、Proc. Natl. Acad. Sci. USA 87, 2264 - 2268）を含む。あるいは、本明細書中に開示される酵素（または当該酵素をコードする核酸）のいずれかの機能性断片を用いてもよい。「機能性断片」という用語は、アミノ末端欠失および/またはカルボキシ末端欠失を有するポリペプチドを指し、残りのアミノ酸配列は、参照配列中の対応する位置に対して少なくとも約40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%の配列同一性を有し、完全長ポリペプチドの活性の約75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%を保持する。機能性断片は、例えば、完全長ポリペプチドの90%以下、80%以下、70%以下、60%以下、50%以下、40%以下、30%以下、または20%以下を含み得、例えば、完全長ポリペプチドの約50%、55%、60%、65%、70%、75%、80%、85%、90%、92%、93%、94%、95%、96%、97%、98%、または99%までを含み得る。記載のEC番号は、国際生化学・分子生物学連合の命名法委員会の酵素命名法を用いている。

20

30

40

【0043】

ステップ1：グルコースから1,5 - グルコノラクトンへの変換（酸化または脱水素）

50

。このステップは、様々な酵素、例えば、ファミリー酸素依存グルコースオキシダーゼ (EC 1.1.3.4) または NAD(P) 依存グルコースデヒドロゲナーゼ (EC 1.1.1.118、EC 1.1.1.119) のものを用いて行うことができる。グルコノバクター・オキシダント (*Glucanobacter oxydans*) は、発酵槽中で増殖させた場合にグルコースを効率的にグルコン酸および 5 - ケトグルコネート (5 - KGA) へ酸化させることが分かっている。グルコノバクター (*Glucanobacter*) および他の酸化細菌中に見受けられる可溶性の膜結合性の PQQ 依存酵素のファミリーの酵素 (EC 1.1.99.35 および EC 1.1.5.2) を使用することができる。キノタンパク質グルコースは、このステップの実行において有用な別の酵素である。選択される特定の酵素は、表 1 に示す所望の反応条件と反応中に存在し得る必要な補因子とに依存する。

10

【0044】

ステップ 1A : 1, 5 - グルコノラクトンからグルコネートへの変換 (例えば、加水分解)。このステップは、化学的に水媒体中で行うことができ、加水分解速度は、pH に依存する (Shimahara, K, Takahashi, T., *Biochim. Biophys. Acta* (1970), 201, 410)。加水分解は、塩基性 pH (例えば、pH 7.5) においてはより高速であり、酸性 pH においてはよりゆっくりである。多数の微生物も、特定の 1, 5 - グルコノラクトンヒドロラーゼを含み、そのうちいくつかはクローン化および特性決定されている (EC 3.1.1.17; Shinagawa, E. *Biosci. Biotechnol. Biochem.* 2009, 73, 241-244)。

20

【0045】

ステップ 1B : グルロン酸ラクトンからグルロン酸への変換。グルロン酸ラクトンの化学加水分解は、水溶液中の自然反応によって行うことができる。この加水分解を触媒することが可能な酵素が、多数のラクトナーゼ (EC 3.1.1.XX およびより詳細には 3.1.1.17、3.1.1.25) の中で特定される。

【0046】

ステップ 2 : グルコン酸から 3 - デヒドログルコン酸 (DHG) への変換。いくつかの酵素、例えば、グルコン酸デヒドラターゼを、グルコン酸のデヒドログルコン酸 (DHG) への脱水において用いることができる。例を挙げると、グルコン酸デヒドラターゼファミリー (EC 4.2.13.9) に属するものがある。このようなデヒドラターゼの特定の例が、グルコネートを脱水させることが分かっている (Kim, S. Lee, S. B. *Biotechnol. Bioprocess Eng.* (2008), 13, 436)。このファミリー由来の酵素についての特定の実施例およびそのクローニングを実施例 1 に示す。

30

【0047】

ステップ 3 : 3 - デヒドロ - グルコン酸 (DHG) から 4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート (2, 5 - DDH) への変換。この変換を行うための酵素である 2 - デヒドロ - 3 - デオキシ - D - グルコン酸 5 - デヒドロゲナーゼ (または DHG デヒドロゲナーゼ) (EC 1.1.1.127) について記載されている。

40

【0048】

ステップ 4 : 4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート (2, 5 - DDH) から 4 - デオキシ - L - トレオ - ヘキサスロースウロン酸 (DTHU) への変換。ファミリー EC 5.3.1.12 の酵素をこのステップにおいて用いることができ、ステップ 15 は、5 個のこのような酵素がクローン化され、5 - KGA の脱水についての活性を有することが分かったことを示す。これらの酵素は、2, 5 - DDH および DTHU に対する活性も示すであろう。

【0049】

ステップ 5 : DTHU から 5 - デヒドロ - 4 - デオキシ - グルカレート (DDG) への変換。DDG は、例えばアルコールの存在下でアルデヒドを酸化させることが可能な穏や

50

かな化学触媒を用いて、D T H Uの化学的酸化または酵素的酸化から生成することができる。アルデヒドオキシダーゼを用いて、この酸化を触媒することができる。酸化細菌、例えば、アセトバクター (Acetobacter) およびグルコノバクター (Hollmann et al Green Chem. 2011, 13, 226) は、スクリーニングにおいて有用であろう。以下のファミリーの酵素が、この反応を行うことができる：アルデヒドオキシダーゼ E C 1 . 2 . 3 . 1、アルデヒドフェレドキシンオキシドレダクターゼ (E C 1 . 2 . 7 . 5)、およびファミリーの E C 1 . 2 . 1 . - X Xの全てのファミリー中のもの。ウロン酸デヒドロゲナーゼの酵素 (E C 1 . 1 . 1 . 203) (例えば、ステップ7を参照) も、この活性を有する。アルコール酸化活性およびアルデヒド酸化活性の双方を有する他の酵素を用いることができる (例えば、アルジトールオキシダーゼファミリー中の酵素 (ステップ19および6を参照)。他の広範な基質オキシダーゼは、可溶性および膜結合性の P Q Q 依存アルコール / アルデヒドオキシダーゼを含む。より詳細には、可溶性ペリプラスム P Q Q オキシダーゼ酵素および I 型に属するその相同体 (E C 1 . 1 . 9 . 1) および I I (E C 1 . 1 . 2 . 8) ファミリーと、E C 1 . 1 . 5 . X ファミリーに属する膜結合性の P Q Q オキシダーゼとが有用である。他の実施形態において、D T H U 上に作用するアルデヒドデヒドロゲナーゼ / オキシダーゼが使用可能である。

【0050】

ステップ6および6A：グルコン酸からグルロン酸(6)への変換および3-デヒドロ-グルコン酸(DHG)から4-デオキシ-5-エリトロ-ヘキサスロースウロン酸(DEHU)(6A)への変換。ステップ5中に記載される酵素は、これらの変換において有用である。他の有用な酵素は、E C 1 . 1 . 1 . X X ファミリー中の N A D (P) 依存デヒドロゲナーゼ、およびより詳細にはグルクロン酸デヒドロゲナーゼ (E C 1 . 1 . 1 . 19)、グルクロノラクトンレダクターゼ (E C 1 . 1 . 1 . 20) を含む。加えて、糖を含む広範な基質範囲を有する多数の O₂ 依存アルコールオキシダーゼは、有用であると考えられ (E C 1 . 1 . 3 . X X)、ソルビトール / マンニトールオキシダーゼ (E C 1 . 1 . 3 . 40)、ヘキサースオキシダーゼ (E C 1 . 1 . 3 . 5)、アルコールオキシダーゼ (E C 1 . 1 . 3 . 13) およびパニリンオキシダーゼ (E C 1 . 1 . 3 . 38) を含む。酸化細菌中に存在する P Q Q 依存酵素および酵素を、これらの変換のために用いることができる。

【0051】

ステップ7および7B：グルロン酸からD-グルカル酸(7)への変換およびL-イズロン酸からイダル酸(7B)への変換。これらのステップは、本明細書中に記載のようなウロン酸デヒドロゲナーゼ (E C 1 . 1 . 1 . 203) またはオキシダーゼのファミリーの酵素により達成することができる。

【0052】

ステップ7A：4-デオキシ-5-エリトロ-ヘキサスロースウロン酸(DEHU)から3-デオキシ-D-エリトロ-2-ヘキサロサル酸(DDH)への変換。ステップ5中に記載される同じ酵素が、この変換の実行において有用となるであろう。

【0053】

ステップ8および8A：D-グルカル酸から5-デヒドロ-4-デオキシ-グルカレート(DDG)への変換(ステップ8)およびイダル酸からDDGへの変換(ステップ8A)。グルカル酸デヒドラターゼのファミリーの酵素 (E C 4 . 2 . 1 . 40) を用いて、これらのステップを行うことができる。このファミリーの酵素がクローン化されており、グルカレートからDDGへの変換を効率的に行うことが判明している。2つのD-グルカル酸デヒドラターゼ (E C 4 . 2 . 1 . 40) を、以下のクローン化されたグルカル酸デヒドラターゼについての表に示した通りにクローン化した。双方の酵素は、ステップ2に述べるように、セミカルバジドアッセイ法を用いたグルカレートのDDGへの脱水についての極めて高い活性を示した。

【0054】

クローン化されたグルカル酸デヒドラターゼ

生物	pSGI (ベクター)	遺伝子ID	WT/SYN
大腸菌(E. coli)	353 (pET28)	P0AES2	WT
シュードモナス(Pseudomonas)(SGI)	244	#8114	WT

【 0 0 5 5 】

ステップ 9 および 9 A : D - グルコースから - D - グルコ - ヘキソジアルド - 1 , 5 - ピラノース (9) への変換および D - ガラクトースから D - ガラクト - ヘキソジアルドース (9 A) への変換。オキシダーゼ、例えば、ガラクトースオキシダーゼファミリー (E C 1 . 1 . 3 . 9) のものを、このステップにおいて用いることができる。また、突然変異ガラクトースオキシダーゼをグルコースに関する活性を有するように遺伝子操作することが記載されている (A r n o l d , F . H . e t a l C h e m B i o C h e m , 2 0 0 2 , 3 (2) , 7 8 1) 。

10

【 0 0 5 6 】

ステップ 1 0 : - D - グルコ - ヘキソジアルド - 1 , 5 - ピラノースから - D - グルコピラヌロン酸への変換 (ステップ 1 0) および D - ガラクト - ヘキソジアルドースからガラクトツロネートへの変換 (1 0 A) 。このステップは、アルデヒドデヒドロゲナーゼのファミリーの酵素を用いて行うことができる。

【 0 0 5 7 】

ステップ 1 1 および 1 1 A : - D - グルコピラヌロン酸からグルクロン酸 1 , 5 - ラクトンへの変換。ステップ 5 に記載のようなアルデヒドデヒドロゲナーゼおよびオキシダーゼが、このステップの実行において有用であろう。ステップ 7 および 7 B 中に記載のウロン酸デヒドロゲナーゼも、このステップの実行において有用であろう。ステップ 1 1 A は、ガラクトツロネートからガラクタレートへの変換である。例えばステップ 7 および 7 B に記載のようなウロン酸デヒドロゲナーゼ (E C 1 . 1 . 1 . 2 0 3) が、このステップの実行において有用であろう。

20

【 0 0 5 8 】

ステップ 1 2 : フルクトースからグルコースへの変換。グルコースおよびフルクトースイソメラーゼ (E C 5 . 3 . 1 . 5) が、このステップの実行において有用であろう。

30

【 0 0 5 9 】

ステップ 1 3 : ガラクタレートから 5 - デヒドロ - 4 - デオキシ - D - グルカレート (D D G) への変換。ガラクトアル酸デヒドロゲナーゼのファミリーの酵素 (E C 4 . 2 . 1 . 4 2) を用いてこのステップを行うことができ、さらなる酵素をこのステップのために遺伝子操作することができる。

【 0 0 6 0 】

ステップ 1 4 : グルコネートから 5 - ケトグルコネート (5 - K G A) への変換。N A D (P) 依存デヒドロゲナーゼのファミリーの複数の酵素 (E C 1 . 1 . 1 . 6 9) がクローン化され、グルコネートの酸化または 5 K G A の還元についての活性を有することが分かった。例えば、グルコノバクター (E x p a s y P 5 0 1 9 9) 由来の N A D P H 依存グルコン酸 5 - デヒドロゲナーゼを、本明細書中に示すように大腸菌における最適な発現のために合成して、p E T 2 4 (p S G I - 3 8 3) 中にクローン化した。酵素を発現させ、必要な活性を有することが分かった。このステップの実行において有用なさらなる酵素は、グルコノバクター中に存在する P Q Q 依存酵素のファミリーのもの (P e t e r s , B . e t a l . A p p l . M i c r o b i o l B i o t e c h n o l . , (2 0 1 3) , 9 7 , 6 3 9 7) と、ステップ 6 中に記載される酵素とを含む。これらのファミリー由来の酵素は、グルコネートから 5 K G A を合成する際にも用いることができる。

40

【 0 0 6 1 】

ステップ 1 5 : 5 - K G A から L - イズロン酸への変換。このステップは、実施例 4 中にさらに述べるように、異なるイソメラーゼファミリー由来の様々な酵素を用いて行うこ

50

とができる。

【 0 0 6 2 】

ステップ 1 6 : 5 - K G A から (4 S) - 4 , 6 - ジヒドロキシ 2 , 5 - ジケトヘキサノエート (2 , 5 - D D H) への変換。この脱水は、グルコン酸デヒドラターゼファミリー (E C 4 . 2 . 3 . 3 9) 中の酵素、例えば、実施例 5 またはステップ 1 7 に記載のものをを用いて行うことができる。

【 0 0 6 3 】

ステップ 1 7 および 1 7 A : L - イズロネートから 4 - デオキシ - 5 - トレオ - ヘキススロースウロン酸 (D T H U) へおよびグルロネートから 4 - デオキシ - 5 - ヘキススロースウロン酸 (D H U) へ。このステップのパフォーマンスにおいて使用することが可能である、デヒドラターゼのファミリーの酵素が特定された。グルコン酸デヒドラターゼまたはグルカル酸デヒドラターゼのファミリー由来の酵素は、これらのステップを行うための所望の活性を持つであろう。さらに、ファミリーの多数のデヒドラターゼ (E C 4 . 2 . 1 . X) は、これらのステップのパフォーマンスにおいて有用となるであろう。詳細には、ファミリーの酵素 : E C 4 . 2 . 1 . 6 (ガラクトン酸デヒドラターゼ)、E C 4 . 2 . 1 . 8 (マンノネートデヒドラターゼ)、E C 4 . 2 . 1 . 2 5 (アラボネートデヒドラターゼ)、E C 4 . 2 . 1 . 3 9 (グルコン酸デヒドラターゼ)、E C 4 . 2 . 1 . 4 0 (グルカル酸デヒドラターゼ)、E C 4 . 2 . 1 . 6 7 (フコネートデヒドラターゼ)、E C 4 . 2 . 1 . 8 2 (キシロネートデヒドラターゼ)、E C 4 . 2 . 1 . 9 0 (ラムノン酸デヒドラターゼ)、およびジヒドロキシ酸デヒドラターゼ (4 . 2 . 1 . 9) などの、2 - ケト酸を選択的に生成するように 1、2 - ジヒドロキシ酸を脱水させる酵素が有用であると考えられる。既知の酵素選択性はアルファ - ケト酸の生成であるため、特定された酵素から、D E H U および D T H U が反応生成物としてそれぞれ生成される。

【 0 0 6 4 】

ステップ 1 9 : 1 , 5 - グルコノラクトンからグルロン酸ラクトンへの変換。このステップは、アルジトールオキシダーゼのファミリーの酵素 (E C 1 . 1 . 3 . 4 1) またはステップ 6 中に記載される酵素により、行うことができる。

【 0 0 6 5 】

D D G から F D C A へ変換する方法ならびにエステル化 D D G およびエステル化 F D C A を作製する方法

本発明は、D D G を F D C A および F D C A エステルへ変換する新規な方法も提供する。F D C A のエステルは、ジエチルエステル、ジブチルエステル、および他のエステルを含む。方法は、D D G をアルコール、無機酸、および任意選択的に共溶媒と接触させて D D G の誘導体を生成することにより、D D G を D D G エステルへ変換するステップを含む。アルコールは、メタノール、エタノール、プロパノール、ブタノール、または任意の C 1 - C 2 0 アルコールであり得る。無機酸は、硫酸であり得る。共溶媒は、T H F、アセトン、アセトニトリル、エーテル、酢酸ブチル、ジオキサン、クロロホルム、塩化メチレン、1 , 2 - ジクロロエタン、ヘキサン、トルエン、およびキシレンのいずれかまたは任意の混合物であり得る。その後、エステル化 D D G をエステル化 F D C A へ変換することができる。この方法を行う前に、D D G を 1 つのステップとして任意選択的に精製することができる。D D G を精製することは、D D G を含む溶媒から水を除去すること、例えば、8 7 % を超える水または 9 0 % を超える水または 9 5 % を超える水または 9 7 % を超える水または 9 8 % を超える水または 9 9 % を超える水を、D D G を含む溶媒から除去することを含み得る。2 5 % または 3 0 % または 3 5 % または 4 0 % または 4 5 % を超えるモルの収率を得ることができる。

【 0 0 6 6 】

D D G 精製

D D G を酸性化させることにより、例えば、p H 約 2 . 5 までの濃 H C l の添加による反応物 p H の低下により、脱水またはエステル化のための D D G 精製を行った。この p H において、タンパク質および任意の残留グルカレート沈殿物を濾過によって除去し、混合

10

20

30

40

50

物を凍結乾燥させて、D D Gおよび反応塩からなる白色粉末を得た。このD D Gを脱水させて、2, 5 - F D C Aを得てもよい、あるいはエステル化してジブチルD D G（またはジエチルD D G）を得た後に脱水してもよい。このD D Gを精製またはエステル化する方法は、本明細書中に開示される方法および経路のいずれにおいても1つのステップとして付加され得る。

【0067】

F D C AおよびF D C A誘導体を合成するための方法

本発明は、F D C Aを合成するための多様な方法も提供する。F D C Aを合成するための1つの方法は、高温でD D Gをアルコール、無機酸と接触させて、F D C Aを形成するステップを含む。アルコールは任意のアルコールであり得、例として（非限定的に）、メタノール、エタノール、プロパノール、およびブタノールが挙げられる。ジオールを用いてもよい。高温は、F D C Aを形成するための70 を超える温度または80 を超える温度または90 を超える温度または100 を超える温度または110 を超える温度または120 を超える温度または130 を超える温度または140 を超える温度または150 を超える温度であり得る。20 %を超えるかまたは30 %を超えるかまたは35 %を超えるかまたは40 %を超える反応収率を達成することができる。

【0068】

本発明は、F D C Aの誘導体を合成するための方法も提供する。方法は、D D Gの誘導体を無機酸と接触させてF D C Aの誘導体を生成するステップを含む。無機酸は、例えば硫酸であり得る。任意選択的に、D D Gの誘導体を第2の無機酸と接触させる前に精製してもよい。使用可能なD D Gの誘導体の非限定的な例は、メチルD D G、エチルD D G、プロピルD D G、ブチルD D G、イソブチルD D G、ジメチルD D G、ジエチルD D G、ジプロピルD D G、ジブチルD D Gを含む。生成されたF D C Aの誘導体は、メチルF D C A、エチルF D C A、プロピルF D C A、ブチルF D C A、ジメチルF D C A、ジエチルF D C A、ジプロピルF D C A、ジブチルF D C A、およびイソブチルF D C Aであり得る。生成されたF D C Aの誘導体は、本方法において用いられるD D Gの誘導体に対応する。その後、F D C Aの誘導体を脱エステル化してF D C Aを生成することができる。方法は、例えば以下に述べるパラメータを用いて気相中で行ってもよい。

【0069】

F D C AまたはF D C A誘導体を合成するための別の方法は、気相中でD D GまたはD D G誘導体（本明細書中に記載される任意のもの）を無機酸と接触させるステップを含み、この接触は、短い滞留時間、例えば、10秒未満または8秒未満、または6秒未満または5秒未満または4秒未満または3秒未満または2秒未満または1秒未満の滞留時間で行われ得る。滞留時間とは、反応装置を通じた高温流れの反応ゾーン中に試料が存在する時間を指す。方法は、高温で、例えば、150 よりも高温、200 よりも高温、250 よりも高温、300 よりも高温、または350 よりも高温で行ってもよい。25 %を超えるかまたは30 %を超えるかまたは40 %を超えるかまたは45 %を超えるかまたは50 %モルを超える収率が入手可能である。F D C Aを合成するための別の方法は、80 または90 または100 または110 または120 を超える温度でD D Gを無機酸と接触させるステップを含む。F D C Aを合成するための別の方法は、無水反応条件下でD D Gを無機酸と接触させるステップを含む。様々な実施形態において、無水条件は、D D Gが10重量%未満または9重量%未満または8重量%未満または7重量%未満または6重量%未満または5重量%未満または4重量%未満または3重量%未満または2重量%未満の水を含むように、本明細書中に開示されるF D C Aを合成する任意の方法におけるD D Gを凍結乾燥することにより、確立することができる。

【0070】

本明細書中に記載されるF D C Aを合成するための本発明の方法は、今まで利用可能であったものよりもずっと高い収率を提供する。異なる実施形態において、10 %よりも高いかまたは15 %よりも高いかまたは20 %よりも高いかまたは25 %よりも高いかまたは30 %よりも高いかまたは35 %よりも高いかまたは40 %よりも高いかまたは45 %

10

20

30

40

50

よりも高いかまたは50%よりも高い、(DDGに対する)FDCAのモル収率を得ることができる。

【実施例】

【0071】

実施例1 - ステップ2: グルコン酸から3-デヒドロ-グルコン酸(DHG)へ

グルコネートの脱水に対して自然活性を有する酵素が発見されている(EC4.2.1.39)。このファミリー由来の3つの酵素を表1に示すようにクローン化した。酵素pSGI-365がクローン化され、グルコネート脱水についての高い活性を有する広範な基質範囲を持つデヒドラターゼであることが判明した(Kim, S. Lee, S. B. Biototechnol. Bioprocess Eng. 2008, 13, 436)。

10

【0072】

(表1) 本実験において用いられる酵素、および同一性相同性。全てP. ルオレッセンス(P. fluorescens)において発現

生物	pSGI (ベクター)	遺伝子ID	WT/SYN	発現宿主
アクロモバクター (<i>Achromobacter</i>)	365 (pRANGER)	E3HJU7	Syn	P. フルオレッセンス (<i>P. fluorescens</i>)
アクロモバクター	359 (pRANGER)	#0385	wt	P. フルオレッセンス
アシネトバクター (<i>Acinetobacter</i>)	360 (pRANGER)	#0336	wt	P. フルオレッセンス

20

	359_アクロモバクター	365_E3HJU7
pSGI-360_アシネトバクター(SGI)	78	79
pSGI-359_アクロモバクター(SGI)		95
pSGI-365_アクロモバクター		

30

【0073】

タンパク質359、360、および365は、グルコネート脱水の合成(ゲルは図示せず)において、粗酵素溶解物活性1mgあたり2~5μmol/分を示した。硫酸アンモニウムによる沈殿によりpSGI-359を分離させ、緩衝液中に再溶解させ、セミカルバジドアッセイ法によってアッセイした。グルコネート脱水についての、46.2U/mLまたは5.3U/mg(1単位=μmol/分)の活性をセミカルバジドアッセイプロットから計算した。Kpi10mM pH8.0を2mM MgCl2および3.5gr(0.016mol)のグルコン酸ナトリウムと共に含む反応緩衝液(93mL)を、7mLの前のグルコン酸デヒドラターゼ溶液と混合した。反応物を45で16時間培養した後、1つのアリコートでHPLC-MS(図4)によって分析した。図4に示すように、DHGの分子量を有する1つの新規主要生成物を生成した。この生成物は、DHGデヒドラターゼによる活性を有することも判明した。

40

【0074】

全タンパク質をpRANGER(商標)(Lucigen, Middleton, WI)発現ベクターにおいてクローン化し、蛍光菌株において発現させた。pRANGER(

50

商標)は、遺伝子の誘導発現のための p B B R 1 レプリコン、カナマイシン耐性および p B A D プロモーターを含む広範な宿主の市販のプラスミドベクターである。酵素アッセイ法のため、アルファケト酸定量化のためのセミカルバジドアッセイ法の改変版を用いて、各酵素の活性を計算した (Kim, S.; Lee, S. B. Biochem J. 2005, 387, 271)。SEQ ID NO: 30 ~ 32 および 33 ~ 35 は、グルコン酸デヒドラターゼ # 0385、# 0336、および E 3 H J U 7 のアミノ酸およびヌクレオチド配列をそれぞれ示す。

【 0 0 7 5 】

実施例 2 - ステップ 3 : 3 - デヒドロ - グルコン酸 (DHG) から (4S) - 4, 6 - ジヒドロキシ 2, 5 - ジケトヘキサノエート (2, 5 - DDH) へ

ファミリー (EC 1.1.1.127) の酵素を用いて、このステップを行うことができる。2つの例は、2 - デヒドロ - 3 - デオキシ - D - グルコン酸 5 - デヒドロゲナーゼおよび DHG デヒドロゲナーゼである。このファミリーから 5 個の酵素を以下の表 2 に示すようにクローン化した。p R A N G E R (商標) ベクターを各場合において用いた。

【 0 0 7 6 】

(表 2) DHG オキシドレダクターゼ (または 2 - デヒドロ - 3 - デオキシ - D - グルコン酸 5 - デヒドロゲナーゼ) のクローン化

生物	pSGI (ベクター)	遺伝子 ID	WT/SYN	発現宿主
アグロバクテリウム (Agrobacterium) 種 (SGI)	374	#9041	WT	P. フルオレッセンス
アグロバクテリウム・ツメファシエンズ (Agrobacterium tumefaciens) (SGI)	375	#8939	WT	P. フルオレッセンス
大腸菌	376	P37769	WT	P. フルオレッセンス
スフィンゴモナス (Sphingomonas) (SGI)	395	#5112	WT	P. フルオレッセンス
ホエフリー・フォトリフィカ (Hoeftlea phototrophica) (SGI)	396	#7103	WT	P. フルオレッセンス

【 0 0 7 7 】

ステップ 2 においてグルコネート脱水から調製された生成物を、表 2 の溶解物をアッセイするための基質として用いた。以下の表 3 に示すように、酵素は、NADH 形成を測定するアッセイ法において DHG の酸化についての活性を示すことが分かった (340 nm における吸光度増加)。

【 0 0 7 8 】

(表 3) DHG オキシドレダクターゼを用いた 2, 5 - DDH への DHG の酸化に関する活性計算

単位 = NADH の μ モル / 分

酵素	U/mg (100 mM DHG)		
	pH=7.5	pH=8.5 (10 mM DHG)	pH=9.5
pSGI_395	0.012	0.070 (0.02)	0.120
pSGI_396	0.033	0.139 (0.018)	0.418
pSGI_374	0.007	0.043 (0.012)	0.091
pSGI_376	0.007	0.121 (0.01)	1.610

10

【 0 0 7 9 】

これらの酵素による 2, 5 - D D H の形成のさらなる検証は、ステップ 1 6 において、p S G I - 3 9 5 による (5 K G A の脱水で生成された) 2, 5 - D D H の還元を示すステップ 1 6 に示した。

【 0 0 8 0 】

実施例 3 - ステップ 7 および 7 B : グルロン酸から D - グルカル酸 (7) への変換ならびに L - イズロン酸からイダル酸 (7 B) への変換

20

ステップ 7 および 7 B を示すため、以下の試験を行った。ウロン酸デヒドロゲナーゼ (E C 1 . 1 . 1 . 2 0 3) は、グルクロン酸およびガラクトン酸を酸化させる酵素である。公知のウロン酸デヒドロゲナーゼに配列が類似する 3 つの酵素 (E x p a s y : Q 7 C R Q 0 ; P r a t h e r , K . J , e t a l . , J . B a c t e r i o l . 2 0 0 9 , 1 9 1 , 1 5 6 5) を、表 4 および 5 に示すような菌種からクローン化した。

【 0 0 8 1 】

(表 4) クローン化されたウロン酸デヒドロゲナーゼ

生物	pSGI (pET28)	遺伝子 ID	発現
アグロバクテリウム	474	#8807	BL21DE3
根粒菌(Rhizobium)	475	#8958	BL21DE3
シュードモナス	476	#1770	BL21DE3

30

【 0 0 8 2 】

(表 5) 配列同一性

	475	476	Q7CRQ0
474_アグロバクテリウム	73	49	90
475_根粒菌		51	74
476_シュードモナス			50

40

【 0 0 8 3 】

p E T 2 8 由来の H i s タグを有する各タンパク質を発現させ、精製した後、スクリーニングを施した。粗溶解物のタンパク質ゲルおよび精製酵素を、図 1 のゲルにおいて示す。精製後、全酵素について、グルクロネートに対する活性とグルロネートおよびイズロネ

50

ートに対する活性とを試験した。異なる基質濃度における反応速度測定を行い、各酵素の計算された活性およびK m値を表6に示す。全酵素が、グルクロネートならびにL - イズロネートおよびグルロネートに対して良好な活性を示した。

【0084】

(表6) 精製ウロン酸デヒドロゲナーゼの活性およびK m値

酵素	Vmax (μM/ 分 /mg);および Km (mM)		
	グルクロネート	イズロネート	グルロネート (V _m のみ)
474	128.2 ; 0.37	0.96 ; 29.8	0.017
475	47.4 ; 0.22	0.59 ; 42.1	0.016
476	90.9 ; 0.34	1.36 ; 29.6	0.014

10

【0085】

大腸菌細胞において、表4に示す各プラスミドをBL21DE3形質転換した。清澄化した溶解物を均等量(25 mL)の平衡緩衝液と混合し、Ni NTAカラムで精製した。0.050 mLの多様な希釈度の各精製酵素を0.95 mLの反応緩衝液(100 mM Tris HCl、pH 8.0、50 mM NaCl、0.75 mM NAD+)と混合することにより、各精製酵素の活性を測定した。340 nmにおけるNADH形成の監視により、反応進行を測定した。図6Aおよび図6Bは、グルクロネートおよびイズロネートの酸化のラインウィーバー - バークプロットを提供する、3つの酵素全てを図6に示す。明確な正の傾きが得られ、全酵素から上記表に示す活性が得られた。ウロン酸デヒドロゲナーゼのタンパク質配列をSEQ ID NO: 1~3として示し、遺伝子をSEQ ID NO: 4~6として示す。

20

【0086】

実施例4 - ステップ15: 5 - ケトグルコネート(5 - KGA)からL - イズロン酸(15)またはグルロン酸(15A)への変換

この実施例は、5 - KGAからイズロン酸への異性化(ステップ15)またはグルロン酸への異性化(ステップ15A)が可能である酵素の特定を示す。3つの異なるイソメラーゼファミリー由来の13個の酵素を表7に示すようにクローン化し、その配列同一性%を表8に示す。

30

【0087】

(表7) クローン化されたイソメラーゼ

EC	生物	pSGI (pET28)	遺伝子ID Archetype(登録商標) またはExpasy	WT/SYN
5.3.1.17	根粒菌	433	#8938	WT
5.3.1.17	大腸菌	434	Q46938 (Expasy)	WT
5.3.1.17	根粒菌	435	#3891	WT
5.3.1.17	パノニバクター (Pannonibacter)	436	#7102	WT
5.3.1.n1	ラクトバシラス (Lactobacillus)	458	A5YBJ4 (Expasy)	SYN
5.3.1.n1	アキドピルム (Acidophilum)	440	F0J748 (Expasy)	SYN
5.3.1.n1	バシラス (Bacillus)	437	#9209	WT
5.3.1.n1	オクロバクトラム (Ochrobactrum)	438	#9732	WT
5.3.1.n1	ハロモナス (Halomonas)	439	#7403	WT
5.3.1.12	スフィンゴバクテリア (Sphingobacteria)	478	#1874	WT
5.3.1.12	サーモトガ (Thermotoga)	479	Q9WXR9	SYN
5.3.1.12	バシラス	480	Q9KFI6	SYN
5.3.1.12	バシラス	481	O34808	SYN

10

20

【 0 0 8 8 】

(表 8) イソメラーゼの同一性 %

30

	EC	436	434	435	458	440	437	438	439	481	480	479	478
433	5.3.1.17	<u>65</u>	<u>44</u>	<u>43</u>	16	13	18	11	14	6	11	11	7
436	5.3.1.17		<u>45</u>	<u>46</u>	18	14	15	12	13	5	10	11	7
434	5.3.1.17			<u>46</u>	17	10	15	10	13	6	10	12	7
435	5.3.1.17				18	16	18	14	16	9	11	13	7
458	5.3.1.n1					<u>37</u>	<u>57</u>	<u>41</u>	<u>44</u>	6	7	9	5
440	5.3.1.n1						<u>40</u>	<u>67</u>	<u>50</u>	6	6	6	5
437	5.3.1.n1							<u>46</u>	<u>51</u>	8	7	10	6
438	5.3.1.n1								<u>52</u>	5	5	6	4
439	5.3.1.n1									6	7	8	5
481	5.3.1.12										<u>7</u>	<u>36</u>	<u>54</u>
480	5.3.1.12											<u>7</u>	<u>7</u>
479	5.3.1.12												<u>37</u>
478	5.3.1.12												

10

20

【 0 0 8 9 】

表 8 に示すように、各ファミリー内における中程度の相同性（下線）を有する酵素をクロニング対象として選択した。L - イズロネートを主要生成物として生じさせる 5 - K G A の異性化についての活性を全ファミリー由来の酵素が示したことがデータから分かった。5 . 3 . 1 . 1 7 ファミリー（4 3 3 および 4 3 4）由来の 2 つの酵素も本実施例において用いたところ、5 - ケトグルコネート（5 K G A）からの D D G 形成がみられた。

30

【 0 0 9 0 】

2 つの異なる酵素に対する活性により生成物形成を検出した酵素的方法を用いて、表 7 からの酵素を用いた 5 K G A およびイズロネートの異性化についての活性を測定した。例えば、生成物イズロネートの活性を、ウロン酸デヒドロゲナーゼ（p S G I - 4 7 6）を用いて測定することにより、5 K G A の異性化を検出した。生成物 5 K G A の活性 5 K G A レダクターゼ（p S G I - 3 8 3、E C 1 . 1 . 1 . 6 9）を測定することにより、イズロネートの異性化を検出した。生成物の存在は、G C - M S によっても検出された。

【 0 0 9 1 】

全ファミリー由来の酵素は、5 K G A およびイズロネートの異性化についての活性の変動を示した。E C 5 . 3 . 1 . 1 2 由来の 2 つの酵素を無細胞の反応において用いて 5 K G A を異性化させ、最終的に実施例に記載のような D D G を生成した。また、これらの酵素をゲル電気泳動によって精製し、単一バンドを示した。溶解物および 5 K G A またはイズロネートを含む緩衝液を用いた反応において精製イソメラーゼを用いた。H P L C 法および上記の酵素的方法の双方を用いて、生成物形成を決定した。H P L C アッセイ法および酵素アッセイ法の双方を用いた 1 7 時間の培養の結果を図 7 A に示す。全酵素は、5 K G A およびイズロネートの双方の異性化において良好な活性を示した。S G I 4 3 3、p S G I 4 3 4、p S G I 4 3 5、および p S G I 4 3 6 によるイズロネート異性化に関する収率は、酵素的に測定され場合はそれぞれ 5 6 %、4 8 %、4 2 %、（4 3 6 は測定せず）であり、また、H P L C アッセイ法によって測定された場合はそれぞれ 7 8 . 8 %、

40

50

78.5%、73.3%、および76.6%であった。同一酵素による5KGA異性化を16時間行った後の収率を酵素的アッセイ法によって測定したところ、それぞれ18%、17%、および19%（436は測定せず）であり、HPLCアッセイ法によって測定したところ、それぞれ16.6%、17.8%、16.3%、および16.9%であった。

【0092】

EC 5.3.1.12 酵素

また、EC 5.3.1.12 ファミリー（グルクロネートイソメラーゼ）由来の酵素を、ゲル電気泳動によって精製し、分離し、使用して、5 mMの5KGAまたはイズロン酸を含む緩衝液（50 mM HEPES、1 mM ZnCl₂、pH 8.0）との混合により反応物を調製した。これらの反応物を30 において培養し、HPLC法および酵素的方法の双方を用いて生成物形成について分析した。結果を図7Bに示す。

10

【0093】

5.3.1.17 酵素

酵素 pSGI-478 および pSGI-479（5-デヒドロ-4-デオキシ-D-グルクロン酸イソメラーゼ）は、5KGAおよびイズロネートの双方について異性化活性を示した。この活性は、上記のように酵素的アッセイ法によって確認された。pSGI-478 および -479 によるイズロネートの異性化に関する収率は、酵素的に測定した場合にそれぞれ50%および37%であり、HPLCによって測定した場合に20%および18%であった。5KGA異性化に関する収率は、酵素的に測定した場合にそれぞれ23%および26%であり、HPLCによって測定した場合にそれぞれ24%および16%であった。結果を図7Aに示す。

20

【0094】

5.3.1.n1 酵素

このファミリー中の酵素を、ゲル電気泳動によって精製した。生成物形成を上記のような酵素的アッセイ法を用いて測定し、結果を図8に示す。このファミリーにおいてクローン化された全酵素は、5KGAおよびイズロネートの異性化についての活性を有することが分かった。

【0095】

各場合において、BL21DE3においてプラスミドを形質転換し、タンパク質をNiNTAカラムで精製した。

30

【0096】

実施例5 - ステップ16：5-ケト-グルコネート（5KGA）から（4S）-4,6-ジヒドロキシ2,5-ジケトヘキサノエート（2,5-DDH）へ

ステップ2（実施例1）に述べた3つのグルコン酸デヒドラターゼを、ステップ8からの精製グルカル酸デヒドラターゼと共に、実施例1に示すように発現させた。活性に対する酵素反応を行い、HPLC-MS分析から2,5-DDHの形成が示され（図9）、これは、グルコン酸デヒドラターゼを含む試料中のみの5-KGAの還元を伴う新規生成物の形成と、DHGデヒドラターゼ（pSGI-395）を用いた酵素的アッセイ法とによっても確認された。340 nmにおける良好な傾きから、NADH、pSGI-395 溶解物、および前回反応のアリコートを混合した場合（データは図示せず）に高い酵素活性が得られたことが分かる。この結果と、HPLC分析とから、グルコン酸デヒドラターゼが5KGAから2,5-DDHへの脱水を試験されたことが分かる。

40

【0097】

実施例6 - ステップ19：1,5-グルコノラクトンからグルロン酸 - ラクトンへの変換

1,5-グルコノラクトン酸化は、アルジトールオキシダーゼ（EC 1.1.3.41）ファミリー由来の酵素の副活性（side activity）である。これらの酵素は、ソルビトール、キシリトール、グリセロール、および他のものなどの多様なアルジトールを酸化させる。酵素は、以下の表6に示すように、1,5-グルコノラクトンの酸化についての活性を有することが特定された。

50

【 0 0 9 8 】

(表 6) 1, 5 - グルコノラクトンに関する活性を有するアルジトールオキシダーゼ

酵素	酵素源	ソルビトール U/mg	1,5- グルコノラクトン			
			U/mg	反応設定		
				酵素 mg	基質 mg/mM	収率
AO#13	テリグロバズ・ロゼアス (<i>Terriglobuds roseus</i>)	0.23	0.02	5.3	15 / 85	7%
AO#22	グラニユリセラ・マレンシス (<i>Granulicella mallensis</i>)	0.27	0.015	7.6	15 / 85	9%
AO#28	ストレプトミセス・アシディスカ (<i>Streptomyces acidiscabies</i>)	1.30	0.010	15	15 / 85	8%
AO#36	放線菌目 (<i>Actinomycetales</i>) (SGI)	1.83	0.102	25	90 / 35	46%
AO#51	フランキア (<i>Frankia</i>) 種	0.59	0.019	未試験	未試験	未試験
AO#57	プロピオニバクテリア (<i>Propionibacteriaceae</i>) (SGI)	1.47	0.051	40	70 / 57	6%
AO#76	ストレプトミセス (<i>Streptomyces</i>) 種	1.45	0.045	8.2	15 / 85	23%
AO#251*	パエニバチルス (<i>Paenibacillus</i>) 種	0.47	0.003	24	15 8.5	約2%

* 粗溶解物

【 0 0 9 9 】

表 6 に示す全精製酵素の溶解物を用いて、反応物を調製した。反応物を 50 mM の K - リン酸緩衝液 pH 7.0 および 0.5 mg / mL のカタラーゼ中に調製し、30 で培養した。新規の生成物が HPLC - MS 分析によって観察され、真正基準との比較後にグルコネートと同じ保持時間が示された (図 10)。これは GC - MS によって確認され、この生成物は、グルコネートと同じ MS フィンガープリントも持っていた。したがって、表中に記載の全アルジトールオキシダーゼは、1, 5 - グルコノラクトンの 6 - OH を酸化させてグルロン酸ラクトンを生成することが明らかである。全アルジトールオキシダーゼを、His タグを有する pET28a 中にクローン化し、BL21DE3 中に発現させ、NiNTA カラムで精製した。

【 0 1 0 0 】

実施例 7 - FDCA および他の中間体の合成

2, 5 - FDCA への脱水のために、精製 DDG モノカリウム塩を用いた。硫酸を DDG へ添加し、反応物を 60 において攪拌した。原位置収率を (HPLC - MS によって) 計算したところ、約 24% および約 27% であると分かった。

【 0 1 0 1 】

これらの反応溶液を一緒にして、次いで氷に移すことにより (熱をとるため) 希釈した。ほぼ均等量の THF を添加し、溶液を分液漏斗へ移した。分離が達成されるまで、塩化ナトリウム塩を添加した。最高の可能な溶解のため、添加の間に溶液を攪拌した。水性層を除去し、THF 層を飽和 NaCl 溶液により 3 回超洗浄した。硫酸ナトリウムを添加し、溶液を 1 晩放置した。2 つの層が 1 晩で再度形成された。水性層を廃棄した後、シリカゲルを溶液へ添加した。その後、それを固体になるまで *rotovap* を介して濃縮した。固体をシリカフラッシュカラム中へロードした後、クロマトグラフィーにより分離した。画分を濃縮して乾燥させた。分離収率は 173.9 mg であった。補正収率: 24.9%。¹H および ¹³C NMR および HPLC - MS 分析により生成物が確認された。

【 0 1 0 2 】

BuOH / H₂SO₄ 中での DDG からジブチル - 2, 5 - FDCA への脱水

10

20

30

40

50

脱水塩を含む非誘導体化凍結乾燥DDGのBuOH中での脱水を、ディーンスターク装置を用いて行った。これらの条件下で、DDGをBuOHへ添加した後、H₂SO₄を添加し、反応物を140℃で加熱した。4時間攪拌後、HPLC-MS分析により、DDG消失およびジブチル-2,5-FDCAの形成が示された。原位置収率を(HPLC-MSによって)計算したところ、36.5%であった。

【0103】

混合物を水、1%NaOHで抽出し、かつ再度水で抽出した。その後、有機層を最終質量37.21gまで濃縮した。この質量の一部(3.4423g)を除去し、0.34gのジブチル-2,5-FDCAを、HPLCを用いて精製した。反応物から分離された合計化合物量(37.21g)に対する分離生成物の収率を推定し、元々のDDG中に存在する塩量(約60%の重量純度)を考慮すると、反応収率は42%と計算された。¹Hおよび¹³C NMRおよびHPLC-MS分析により生成物が確認された。

10

【0104】

ジブチルDDGの合成

別の態様において、本発明は、DDGの誘導体を合成するための方法を提供する。方法は、DDGをアルコール、無機酸、および任意選択的に共溶媒と接触させて、DDGの誘導体を生成するステップを含む。任意選択的に、DDGの誘導体が精製され得る。この反応は、少なくとも10%のモル収率または少なくとも15%のモル収率または少なくとも20%のモル収率または少なくとも25%のまたは少なくとも30%のまたは少なくとも35%のモル収率または少なくとも40%のモル収率のDDGの誘導体の収率を持ち得る。無機酸は硫酸であり得、アルコールは、メタノール、エタノール、プロパノール、ブタノール、イソブタノール、または任意のC₁-C₂₀アルコールであり得る。様々な実施形態において、共溶媒は、THF、アセトン、アセトニトリル、エーテル、酢酸ブチル、ジオキサン、クロロホルム、塩化メチレン、1,2-ジクロロエタン、ヘキサン、トルエン、およびキシレンのいずれかであり得る。アルコールがエタノールである場合、DDG誘導体は、DDGモノエチルエステルおよび/またはDDGジエチルエステルである。アルコールがブタノールである場合、DDG誘導体は、DDGモノ-ブチルエステルおよび/またはDDGジブチルエステルである。

20

【0105】

DDGモノカリウム塩を以下の手順に従って誘導体化した。機械攪拌器および加熱マントルを備えた1Lモートン型の目的とする反応槽に、60:40DDG:KC1(31.2ミリモル)、BuOH、およびヘプタンを注入した。別個の小びんにおいて、硫酸を水へ添加し、溶解後に冷却させた。その後、溶液をフラスコへ添加した。溶液を30℃で保持した。

30

【0106】

沈殿物を濾過除去し濃縮した。残りのゲルをEtOAc中に溶解させた後、TLCプレートを溶液でスポットし、プレートをリンモリブデン酸混合物で噴霧した後、少なくとも150℃までホットプレート上で加熱して、DDG-DBE画分を特定した。分離収率: 4.62g(15.2ミリモル、47%収率)、98%超の純度。¹Hおよび¹³C NMRおよびHPLC-MS分析により生成物が確認された。

40

【0107】

THF、アセトン、アセトニトリル、エーテル(ジブチル、ジチエルなど)、酢酸ブチルなどのエステル、1,6-ジオキサン、クロロホルム、塩化メチレン、1,2-ジクロロエタン、ヘキサン、トルエン、およびキシレンなどの共溶媒とのBuOH(5%~95%v/v)の混合物などの様々な溶媒を、DDGエステルの合成において用いることができ、共溶媒として用いることができる。酸(硫酸、塩酸、ポリリン酸、もしくはDOWEXなどの固定酸、または塩基(ピリジン、エチル-アミン、ジエチル-アミン、三フッ化ホウ素)、またはカルボン酸のエステル化に一般的に用いられる他の触媒などの、反応触媒。

【0108】

50

n -BuOH/H₂SO₄ 中でのジブチル-DDG からジブチルFDCA への脱水

DDG-DBE (ジブチルエステル) の保存溶液を、ブタノール中に作製し、攪拌子を備えた清潔な乾燥した100 mL 丸底フラスコへ移した。フラスコへ、25 mL の濃硫酸を添加した。フラスコを密閉した後、60 で2時間攪拌した。原位置収率は、約56% であると計算された。反応溶液を濃縮し、残留物を、MTBE 中に溶解させ、分液漏斗へ移した後、水で洗浄した。回収した有機層を濃縮した後、HPLC を介して分離し、分離収率に関しては250.7 mg (約90% 純度)、および35% 分離収率 (純度に対して補正) であった。¹C および ¹³C NMR および HPLC-MS 分析により生成物が確認された。

【0109】

10

実施例8-DDG およびFDCA および誘導体の、5-KGA (ルート2A) からの無細胞合成

この実施例は、スキーム6 (2B のサブスキーム) に従った精製酵素を用いた5KGA からDDG への酵素的変換を示し、生成されたDDG を化学的ステップにより脱水してFDCA にすることも示す。このスキームは、5KGA の異性化 (ステップ15) およびその後のイダル酸への酸化 (ステップ7B) を含む。また、DDG を異なる化学条件下で脱水させてFDCA を得た。最終ステップ (ステップ8A) を、大腸菌由来のグルカ酸デヒドラターゼを用いて行った。

【0110】

20

スキーム6 を図11 に示す。このスキームは、5-KGA からのDDG の無細胞酵素合成を用いて行った。このスキームは、ステップ15、7B、および8A の実行を含む。2つのさらなるタンパク質を用いて反応経路を完了し、そのうち第1の反応経路は、NADH-オキシダーゼ (ステップA) であり、酸素の存在下でNAD+ 補因子と、NADH オキシダーゼの作用から生成された過酸化物を分解するカタラーゼ (ステップB) とを再循環させる。これらの酵素を以下の表7 に示す。全酵素はHis タグを含んでおり、Ni-NTA カラムを用いて精製した。このDDG 合成の収率は、少なくとも88~97% であると計算された。

【0111】

(表7)

ステップ	酵素	EC	生物
15	pSGI-433 (DTHU_IS)	5.3.1.17	根粒菌(SGI)
15	pSGI-434 (DTHU_IS)	5.3.1.17	大腸菌
7B	pSGI-476 (UroDH)	1.1.1.203	シュードモナス(SGI)
8A	pSGI-353 (GlucDH)	4.2.1.40	大腸菌
A	pSGI-431 (NADH_OX)	1.6.3.1	サーマス・サーモフィルス (<i>Thermus thermophilus</i>)
B	カタラーゼ	1.11.1.6	コリネバクテリウム (<i>Corynebacterium</i>)

30

40

【0112】

500 mL の液体培養を、反応のための各イソメラーゼについて精製した。表7 に示す

50

酵素に加えて、各反応物は、50 mM TrisHCl (pH 8.0)、50 mM NaCl、1 mM $ZnCl_2$ および 2 mM $MgCl_2$ 、1 mM $MnCl_2$ および 1 mM NAD^+ を含んでいた。16時間の培養後に反応物をHPLCによって分析し、図12はクロマトグラムを示す。

【0113】

FDC Aへの脱水のため、双方の試料の反応混合物を一緒にして、凍結乾燥して白色粉末状にし、この粉末を2つの試料に分割し、それぞれを、0.25 M H_2SO_4 を含むAcOH中または0.25 M H_2SO_4 を含む4.5 mL BuOH中に溶解させた。双方の反応物を密閉小びん中で2～4時間、120℃において加熱した。反応生成物を図13に示す。

10

【0114】

試料1および2はそれぞれ、真正基準、および、AcOH/ H_2SO_4 中の反応から3時間の時点を示す。試料1による試料2のスパイクにより単一ピークがみられ、FDC A生成物がさらに確認された。試料1および3(図13)はそれぞれ、真正基準、および、BuOH/ H_2SO_4 中の反応から4時間の時点を示す。酵素反応によるFDC A形成により、これらの試料中のDDGの存在がさらに確認される。

【0115】

実施例9 - グルコースおよびグルコネートからのDDG合成

この実施例は、グルコースおよびグルコネートからDDGへの酵素的変換を示す。反応は、精製酵素と、触媒としての粗溶解物とによって行った。酵素および基質を、以下の表に示すように生体反応装置において組み合わせた。

20

	基質	ST-1	ST-14 pSGI-504	ST-15 pSGI-434	ST-7B pSGI-476	ST-8A pSGI-353	ST-A pSGI-431	ST-B
反応-1	グルコース 600 mg	2 mg	7 mL ¹	50 mL ²	7.5 mL ¹	1 mL ³	4 mL ⁴	2 mg
反応-2	グルコネート 700 mg	-	7 mL	50 mL	7.5 mL	1 mL	4 mL	2 mg

30

1. プラスミドを有する組み換え大腸菌の500 mL液体培養物由来の溶解物

2. BL21 DE3 / pSGI - 434の2 L液体培養物由来の溶解物

3. 精製酵素、活性 約30単位(または精製GlucD 3 mg)

4. 250 mLの培養物由来の溶解物

【0116】

反応物を35℃で培養し、溶解した酸素およびpHをそれぞれ20%および8で保持した。複数の時点を経たHPLC-MSによって分析し、結果を図17Bに示す。抽出されたクロマトグラムにおいて、DDG質量(図示せず)および対応するMS断片化が確認された。これらの結果は、グルコースまたはグルコネートを伴う酵素の培養中のDDG生成を明確に示した。

40

【0117】

実施例10 - 組み換えグルカル酸デヒドラターゼのための発現カセットの構築

以下の実施例は、大腸菌細胞中の異種発現のためのD-グルカル酸デヒドラターゼ活性のコード配列(GDH、EC 4.2.1.40)を含む組み換え核酸構築物の生成を記載する。

【0118】

大腸菌(Expressy: P0AES2;)由来のD-グルカル酸デヒドラターゼ、アシ

50

ネトバクター (*Acinetobacter*) ADP1 (Expasy: P0AES2) をコードする遺伝子、ならびに所有権のあるシュードモナス菌種 (BP1MICT2128114) をゲノムDNAからPCR増幅した。

【0119】

その後、PCR増幅遺伝子のそれぞれを細菌形質転換ベクターpET24a(+)へクローン化し、GDH遺伝子それぞれの発現をT7プロモーターの制御下に置いた。PCR増幅インサートそれぞれのヌクレオチド配列も、配列決定確認により確かめた。

【0120】

実施例11 - 組み換えグルカル酸デヒドラターゼを発現する大腸菌株

実施例9に述べるように構築した発現ベクターのそれぞれを、熱ショック媒介形質転換によりNovablue(DE3)大腸菌中に導入した。推定形質転換体を、カナマイシン(50 µg/ml)を追加したLB寒天上で選択した。コロニーPCRアッセイ法において適切なPCRプライマーを用いて、発現ベクターそれぞれを含む陽性クローンを確認した。

【0121】

各発現ベクターについて、細菌コロニーを、形質転換プレートから取り出し、カナマイシン(50 µg/ml)を追加した液体LB培地において30℃で2日間増殖させた。その後、培養液を、15%グリセロールを含む小びん中へ移し、凍結された高純度培養液として-80℃で保存した。

【0122】

実施例11 - GDH酵素を発現する組み換え大腸菌細胞の細胞溶解物を用いることによるDDGのインビトロ合成の実証

この実施例は、大腸菌細胞中に生成された組み換えGDH酵素を用いていかにしてDDG中間体のインビトロ合成を達成したかについて述べる。

【0123】

細胞溶解物の調製：実施例2において上で記載したように構築された組み換え菌種を、カナマイシン(50 µg/ml)を追加した3mLの液体LB培地中30℃で、回転速度を250rpmに事前設定された回転振盪機において1日、個別に増殖させた。この事前培養液を用いて、カナマイシン(50 µg/ml)を含む100mLのTB培地を植え付けた後、30℃で250rpmに事前設定された回転振盪機において早期対数期(OD₆₀₀ 約0.5~0.6)まで2~3時間培養した後、イソプロピルD-1チオガラクトピラノシド(IPTG; 0.25mM最終濃度)を添加して、タンパク質発現を誘導した。細胞をさらに18時間30℃で増殖させた後、遠心分離によって回収し、15mLの溶解緩衝液(10mMリン酸緩衝液、pH7.8、2mM MgCl₂)中で再度懸濁させ、超音波処理によって溶解させた。大腸菌細胞中での組み換え酵素の生成を基準プレキャストSDS-PAGEゲルシステム(BioRad)を用いて定量し、特定の活性をGulick et al. (Biochemistry 39, 4590-4602, 2000)によって記載される手順に従って測定した。その後、細胞溶解物を、以下に詳述するグラム量のグルカレートをDDGへ変化する能力について試験した。

【0124】

グルカレートの酵素的脱水：5グラムのグルカル酸一カリウム(約0.02モル)を、10mM MgCl₂を含む85mLの5mM リン酸カリウム緩衝液へ添加した。5MのNaOH約2mLの添加後、基質グルカレートが徐々に溶解したことが分かった。反応物のpHを約7.8に調節した。その後、15mLの細胞溶解物は、実施例3に示すような10mMリン酸緩衝液(pH7.8)中の3つの組み換えデヒドラターゼそれぞれを含む。緩やかに攪拌しながら培養を30℃で1~2時間行った後、HPLC-MS技術を用いて反応物を分析した。HPLC-MS結果は、予測される生成物DDGに対応する分子量を有する唯一の主要生成物としての新規ピークと、微量のグルカル酸一カリウム基質とを示した。HPLC-MS分析において副生成物は何も検出されず、組み換え酵素それぞれによって触媒された変換反応は極めて高効率であり、特異性も高いことが分かった。

【0125】

酵素反応由来のDDG生成物の精製

酵素的脱水を介して精製されたDDGを、以下の2つの技術のうちいずれかを用いることにより精製した。

【0126】

酵素的脱水反応物を6MのHClによりpH約2.0まで酸性化させ、濾過して沈殿タンパク質を除去した後、凍結乾燥した。メタノール(MeOH)を凍結乾燥粉末へ添加した後、10～15分間緩やかに攪拌して、DDG生成物を溶解させたが、脱水反応混合物中の他の塩(例えば、KClおよびホスフェート)は溶解させなかった。懸濁液の濾過およびMeOHの蒸発の後、実質的に高純度のDDG酸が得られた。

10

【0127】

いくつかの場合、DDG塩の精製のために代替の手順が用いられ、ここで、第1のMeOH濾液を約15～25mLの量になるまで濃縮した後、0.5MKOHを含む同体積のMeOHと混合した。KOH添加後のDDGのカリウム塩沈殿物を、次いで濾過により分離した。

【0128】

HPLC-MS分析の結果は、これら2つの精製技術から得られた試料中の合計生成物のうち少なくとも95%をDDG生成物が構成していたことを示した。

【0129】

実施例12 - ワンステップ化学反応におけるDDGからのFDCAのインビトロ合成の実証

20

本出願人は、 H_2SO_4 の存在下におけるDDGからFDCAへの化学的変換により、FDCA合成(すなわち、遊離酸形態)を達成することができることを発見した。反応は、以下のように行った。約20mgのDDG酸(実施例3に述べるように事前精製された塩による粗凍結乾燥粉末)と、0.25Mの H_2SO_4 とを、1mLの水および1mLのDMSOを含む気密シーリングされた管中に添加した。DDGがこの溶液中に完全溶解したことが分かった。反応物を105℃で18時間攪拌した。粗反応試料に対して行われたHPLC-MS分析の結果は、FDCA遊離酸(FDCA: 2,5-フランジカルボン酸)が主要生成物として形成されたことと、微量のいくつかの他の未同定の副生成物が形成されたことを示した。HPLC-MS分析における対照として、市販のFDCAを同じ条件で分析した。

30

【0130】

実施例13 - FDCAエステル(ジメチルエステル、ジエチルエステル、ジブチルエステル、およびイソプロピルエステル)のインビトロ合成の実証

精製DDGからのジエチル-2,5FDCAの合成

気密シーリングされた管において、18mLのEtOH、実施例11に述べるように事前精製された0.2グラム(1ミリモル)のDDG酸、および0.25Mの H_2SO_4 を添加した。DDG酸は、この溶液中に完全には溶解しなかった。反応物を105℃で18時間緩やかに攪拌した。粗反応試料のGC-MS分析の結果は、ジエチル-FDCA主要生成物の形成を示した。対照として、真正FDCAを化学的に合成し、エステル化してジエチル-FDCAにし、同じ条件で分析した。

40

【0131】

実施例14 - 精製DDGからのジブチル-2,5FDCAの合成

気密シーリングされた管において、18mLのn-BuOH、実施例11に述べるように事前精製された0.2グラム(1ミリモル)のDDG酸、および0.25Mの H_2SO_4 を添加した。DDG酸は、この溶液中に完全には溶解しなかった。反応物を105℃で18時間緩やかに攪拌した。図15に示すように、反応試料のGC-MS分析の結果は、ジブチル-FDCA(FDCA: 2,5-フランジカルボン酸)が主要生成物として形成されたことを示した。対照として、真正FDCAを化学的に合成し、エステル化してジブチル-FDCAにし、同じ条件で分析した。

50

【 0 1 3 2 】

実施例 1 5 - 粗 D D G (未精製) からのジブチル - 2 , 5 F D C A の合成

実施例 1 1 に述べるようなグルカレートの酵素的脱水から直接得られた未精製の凍結乾燥粉末である 0 . 2 グラム (1 ミリモル) の粗 D D G 酸を、 1 8 m L の n - B u O H を含む気密シーリングされた管中に添加した後、 0 . 2 5 M の H_2SO_4 を添加した。粗 D D G 酸は、この溶液中に完全には溶解しなかった。反応物を 1 0 5 で 1 8 時間緩やかに攪拌した。粗反応試料の G C - M S 分析の結果は、ジブチル - F D C A (F D C A : 2 , 5 - フランジカルボン酸) が主要生成物として形成されたことを示した。G C - M S 結果は、粗凍結乾燥粉末 / 未精製凍結乾燥粉末中の汚染物質の塩の存在が反応結果に大きな影響を与えなかったことを示した。対照として、真正 F D C A を化学的に合成し、エステル化してジブチル - F D C Aにし、同じ条件で分析した。

10

【 0 1 3 3 】

実施例 1 6 - 固定酸を用いた F D C A および / またはエステルのインビトロ生成

工業的な実施において、固定酸は、いくつかの種類の溶媒 (水性、有機、または混合など) において機能することが多いため、脱水の実行に対して多数の利点を提供する。加えて、固定酸は、再循環および再利用が容易である。固定化 A M B E R L Y S T (登録商標) 1 5 (R o h m a n d H a a s , P h i l a d e l p h i a , P A) および D O W E X (登録商標) 5 0 W X 8 (D o w C h e m i c a l C o , M i d l a n d , M I) を用いた F D C A のエステルの合成のいくつかの例に従う。

20

【 0 1 3 4 】

D O W E X (登録商標) 5 0 W X 8 の使用による粗 D D G からのジブチル - F D C A の合成

気密シーリングされた管において、 2 m L の n - ブタノール、 2 0 m g の粗 D D G 酸 (塩を含む未精製の凍結乾燥粉末)、および 2 0 0 m g の D O W E X (登録商標) 5 0 W X 8 を一緒にした。D D G は、この溶液中に完全には溶解しなかった。反応物を 1 0 5 で 1 8 時間緩やかに攪拌した。粗反応試料の G C - M S 分析の結果は、ジブチル - F D C A (F D C A : 2 , 5 - フランジカルボン酸) が主要生成物として形成されたことを示した。この G C - M S 結果は、粗凍結乾燥粉末 / 未精製凍結乾燥粉末中の汚染物質の塩 (ホスフェートおよび N a C l) の存在が反応結果に大きく影響しなかったことを示した。対照として、真正 F D C A を化学合成しエステル化してジブチル - F D C Aにし、同じ条件で分析した。

30

【 0 1 3 5 】

A M B E R L Y S T (登録商標) 1 5 の使用による粗 D D G からのジブチル F D C A の合成

気密シーリングされた管において、 2 m L の n - ブタノール、 2 0 m g の粗 D D G 酸 (塩を含む粗凍結乾燥粉末)、および 2 0 0 m g の A M B E R L Y S T (登録商標) 1 5 (R o h m a n d H a a s , P h i l a d e l p h i a , P A) を一緒にした。D D G は、この溶液中に完全には溶解しなかった。反応物を 1 0 5 で 1 8 時間緩やかに攪拌した。粗反応試料の G C - M S 分析の結果は、ジブチル - F D C A (F D C A : 2 , 5 - フランジカルボン酸) が主要生成物として形成されたことを示した。この G C - M S 結果は、粗凍結乾燥粉末 / 未精製凍結乾燥粉末中の汚染物質の塩 (ホスフェートおよび N a C l) の存在が反応結果に大きく影響しなかったことを示した。対照として、真正 F D C A を化学合成しエステル化してジブチル - F D C Aにし、同じ条件で分析した。

40

【 0 1 3 6 】

A M B E R L Y S T (登録商標) 1 5 の使用による粗 D D G からのエチル - F D C A の合成

気密シーリングされた管において、 2 m L のエタノール、 2 0 m g の粗 D D G 酸 (塩を含む未精製の凍結乾燥粉末)、および 2 0 0 m g の A M B E R L Y S T (登録商標) 1 5 (R o h m a n d H a a s , P h i l a d e l p h i a , P A) を一緒にした。D D G は、この溶液中に完全には溶解しなかった。反応物を 1 0 5 で 1 8 時間緩やかに攪拌

50

した。粗反応試料のGC-MS分析の結果は、ジエチル-FDCA (FDCA: 2, 5-フランジカルボン酸) が主要生成物として形成されたことを示した。このGC-MS結果は、粗凍結乾燥粉末/未精製凍結乾燥粉末中の汚染物質の塩(ホスフェートおよびNaCl)の存在が反応結果に大きく影響しなかったことを示した。対照として、市販のFDCAを化学的にエステル化してジエチル-FDCAにし、同じ条件で分析した。

【0137】

DOWEX (登録商標) 50 WX 8の使用による粗DDGからのジエチル-FDCAの合成

気密シーリングされた管において、2 mLのエタノール、20 mgの粗DDG酸(塩を含む未精製の凍結乾燥粉末)、および200 mgのDOWEX (登録商標) 50 WX 8を一緒にした。DDGは、この溶液中に完全には溶解しなかった。反応物を105℃で18時間緩やかに攪拌した。粗反応試料のGC-MS分析の結果は、ジエチル-FDCA (FDCA: 2, 5-フランジカルボン酸) が主要生成物として形成されたことを示した。このGC-MS結果は、粗凍結乾燥粉末/未精製凍結乾燥粉末中の汚染物質の塩(ホスフェートおよびNaCl)の存在が反応結果に大きく影響しなかったことを示した。対照として、市販のFDCAを化学的にエステル化してジエチル-FDCAにし、同じ条件で分析した。

【0138】

実施例17 - FDCA誘導体の生成

有用性の高い複数のFDCA誘導体の合成を図16に記載し、ここで、DTHUの脱水により、フルフラール-5-カルボン酸、すなわちFCAが生成され、その後これを、FDCAに化学的または酵素的に酸化させるか、FCHに還元するか、または、(化学還元的アミノ化もしくはアミノ基転移酵素を用いて) アミノ酸-AFCへアミノ基転移させる。

【0139】

実施例18 - 気相反応におけるジブチルFDCAの生成

この実施例において、GCの入口を、ジブチルDDGからジブチルFDCAへの脱水を触媒する高温反応装置として用いた。その結果得られた生成物をクロマトグラフィーにより分離し質量分析によって検出した。ブタノール中のジブチルDDG (10 mM) および硫酸 (100 mM) の溶液をGC小びん中に入れた。小びんがGCに投入され、FDCAジブチルエステルが観察された。反応は、300℃の入口(滞留時間=4秒)において生じた。6回の注入の平均収率は54%であった。

GC設定: 直接液体注入/MS検出器。

入口: 300℃、全流量29.51 mL/分、分割比10:1、分流24.1 mL/分、隔膜パージ流3 mL/分。

GCライナー: 4 mm、ガラスウール (P/N 5183-4647)。

カラム流: 2.41 mL/分 He一定圧制御。

オーブンプログラム: 40℃で2分保持後、25℃/分で275℃まで上昇、その後40℃/分で325℃まで上昇、2分保持。

カラム: HP-5 MS、Agilent Technologies、30 m x 0.25 mm x 0.25 μm。

合計実行時間: 14.65分。

MSDトランスファーライン: 290℃。

MS源: 250℃。

MS四重極: 150℃。

保持時間:

2, 3-FDCAジブチルエステル: 9.3分、

2, 5-FDCAジブチルエステル: 9.7分。

【0140】

本明細書中に言及される全ての公開文献および特許出願は、個々の公開文献および特許

10

20

30

40

50

出願が具体的かつ個別に参考のため援用されているかのような程度まで、本明細書中に参考のため援用する。

【 0 1 4 1 】

いかなる参照も従来技術を構成していることを認めるものではない。参考文献中の考察は、それらの筆者の主張を述べるものであり、本出願人は、引用文献の精度および適切性に反論する権利を保持する。複数の従来技術の公開文献について本明細書中に参照しているが、この参照は、これらの文書のいずれも、当該分野における一般的知識の一部を形成するものとして認めるものではないことが明確に理解されるであろう。

【 0 1 4 2 】

また、上記の実施例は、本発明の例示のためのものであり、本発明を限定するものではないことが理解されるべきである。

<SEQ ID NO: 1>タンパク質 #474

MAMKRLLVLTGAAGQLGRVMRKRLASMAEIVRLADLAPLDPAGPNEECMQCDLAD
ADAVDAMVAGCDGIVHLGGISVEKPFEQILQGNIIGLYNLYEAARAHGQPRIIFASSN
HTIGYYPQTERLGPDPVFRPDGLYGVSKCFGESLARMYFEKFGQETALVRIGSCTPEP
LNYRMLSTWFSHDDFVSLIEAAFRAPVLGCPIVWGASANDASWWDNSHLGFIGWKP
KDNAEAFRRKIAETTPQPDARDPIVRFQGGVFVDNPIFKET*

<SEQ ID NO: 2> タンパク質 #475

MKRLITGAAGALGRVMRERLAPMATILRLSDIPIGAARQNEEIVQCDLADAKAVH
ALVEDCDGIVHLGGVSVERKFSQIVAGNIVGLYNLYEAARAHMRMPRIVFASSNHTIGF
YPQTERLSVDHPYRPDGLYGVSKCFGESLAHMYHEKFGQETALVRIGSCVTEPVNH
RMLSTWLSYDDFVSLIEAVFRAPKLGCPIVWGASNNDAGWWDNSAAGFLGWKPKD
NAEIFRSKIEAACERPGSDDPAARWQGGLFTQDPIFPEDE*

10

<SEQ ID NO: 3> タンパク質 #476

MTTAYTPFNRLLLTGAAGGLGKVLRESLRPYANVLRVSDIAAMSPATGAHEEVQVC
DLADKAAVHQLVEGVDAIILHFGGVSVERPFEEILGANICGVFHIYEAARRHGVKRV
FASSNHVIGFYKQDETIDANCPRRPDSYYGLSKSYGEDMASFYFDRIYGIETVSIRIGSS
FPEPHNRRMMSTWLSFADLTQLLERALYTPNVGHTVVYGMSANKNVWWDNHLAA
HLGFQPKDSSEVFRAQIDAQPMMPAADDPAMVVFQGGAFVAAGPFGDD*

20

SEQ ID NO: 4 pSGI-474-#8807-DNA

ATGGCAATGAAACGGCTTCTTGTTACCGGTGCTGCGGGCCAGCTTGGCCGCGTTA
TGCGCAAACGCCTTGCATCGATGGCCGAGATCGTTCGCCTTGCCGATCTCGCCCC
GCTCGATCCGGCAGGCCCGAACGAGGAATGCATGCAATGCGACCTTGC GGATGC
AGACGCCGTTGACGCCATGGTTGCCGGTTGCGACGGCATCGTTCACCTCGGCGGC
ATATCGGTGGAGAAGCCTTTCGAACAAATCCTTCAGGGCAACATCATCGGGCTGT
ATAATCTCTATGAGGCCGCCCCGCGCCACGGCCAGCCGCGCATCATCTTCGCCAG
TTCGAACCATACGATCGGTTATTACCCGCAGACGGAGAGGCTTGGACCGGATGTT
CCCTTCCGCCCCGGATGGGCTTTACGGCGTCTCCAAATGTTTCGGCGAGAGCCTTG
CCCGCATGTATTTGAGAAATTCGGCCAGGAGACCGCACTTGTCCGCATCGGCTC
CTGCACGCCCGGAACCCCTTAATTACCGCATGCTGTCCACCTGGTTTTTCGCATGAC
GATTTCTGCTCTGCTGATCGAGGCGGCGTTCCGCGCCCCCGTGCTCGGCTGCCCCA
TCGTCTGGGGGGCGTCGGCCAACGATGCGAGCTGGTGGGACAATTCGCATCTCG
GCTTTATTGGATGGAAACCGAAGGACAATGCCGAGGCCTTCCGCCGGAAGATTG
CCGAAACGACGCCGCAGCCGGACGCGCGCGACCCGATTGTCCGCTTTCAGGGTG
GCGTGTTTGTGACAACCCGATCTTCAAGGAGACGTGA

30

40

SEQ ID NO: 5 pSGI-475-#7895- DNA

ATGAAGAGACTTCTGATTACCGGCGCAGCGGGTGCCTGGGCGCGTGATGCGG
GAAAGGCTCGCACCCATGGCAACGATTCTGCGCCTTTCCGATATCGCCCCGATTG
GAGCGGCCCCGCCAGAACGAGGAAATCGTCCAGTGCGATCTTGCCGATGCCAAAG
CAGTGCATGCTCTGGTCTGAAGATTGCGACGGGATCGTCCATCTCGGTGGCGTCTC
AGTAGAGCGCAAGTTCTCGCAGATCGTCGCCGGCAACATCGTCGGCCTTTACAAT
CTCTACGAAGCCGCACGCGCGCATCGGATGCCGCGCATCGTCTTTGCAAGTTCCA
ATCACACCATCGGCTTTTATCCGCAAACCGAACGGTTGTTCGGTGGACCATCCCTA
TCGTCCGGACGGGCTCTACGGCGTATCGAAATGTTTCGGCGAGTCTCTGGCGCAT
ATGTACCATGAGAAGTTCGGGCAGGAGACGGCACTCGTGCGCATCGGGTCCTGC
GTGACCGAACCGGTCAACCATCGCATGCTTTCCACCTGGCTTTCCTACGATGATT
TCGTCTCGCTTATCGAGGCCGTATTCCTGTGCGCCGAAACTCGGCTGCCCCGTAT
CTGGGGCGCGTCGAACAACGATGCAGGATGGTGGGACAATTCCGCCGCCGGCTT
TCTCGGCTGGAAGCCGAAAGACAATGCCGAAATCTTCCGTTTCAAGATCGAAGC
CGCTTGCGAACGCCCCGGTTCTGATGATCCGGCCGCCCGCTGGCAAGGCGGGCTC
TTCACGCAGGACCCGATCTTCCCAGAGGACGAGTAA

10

20

SEQ ID NO: 6 pSGI-476-#1770-DNA

ATGACCACAGCCTACACCCCCTTCAATCGCCTGCTACTCACCGGAGCGGCAGGCG
GCCTCGGCAAGGTCCTGCGCGAAAGCCTGCGACCTTATGCCAACGTCCTGCGCGT
CTCCGACATCGCGGCCATGAGCCCTGCCACAGGCGCCCATGAAGAAGTCCAGGT
CTGCGACCTCGCCGATAAAGCGGCGGTCCATCAACTGGTCGAAGGCGTCGACGC
AATCCTGCACTTCGGTGGCGTATCGGTGGAGCGGCCCTTCGAGGAAATCCTCGGG
GCCAATATCTGCGGCGTGTTTCATATCTATGAAGCGGCGCGCCGGCATGGCGTAA
AGCGGGTGATCTTCGCCAGCTCCAACCACGTCATCGGTTTTTATAAGCAGGACGA
AACCATCGACGCCAACTGCCCCGCGCCGCCCGACAGCTACTACGGTCTGTCCAA
GTCCTACGGCGAAGACATGGCCAGCTTCTACTTCGACCGCTACGGCATCGAGACC
GTGAGCATCCGCATCGGCTCCTCGTTCCCCGAGCCGCACAATCGCCGCATGATGA
GCACCTGGCTGAGCTTTGCCGACCTGACGCAGCTGCTCGAACGCGCGCTGTACAC
CCCCAACGTCGGCCACACCGTGGTCTACGGCATGTCCGCTAACAAGAACGTCTG
GTGGGACAACCACCTGGCCGCGCACCTGGGCTTCCAACCGAAGGACAGCTCCGA
GGTGTTCGTCGCGAGATCGATGCCAGCCGATGCCCGCCGCCGATGACCCGGC
GATGGTCTTTCAAGGCGGCGCCTTTGTGCGAGCCGGGCGGTTTCGGCGACGACTGA

30

40

SEQ ID NO: 7 pSGI-433 #8938-タンパク質

MLNVETRHA VHADHARSLDTEGLRRHFLAQGLFAEGEIRLIYTHYDRFVMGGAVPD
GAPLVLDHVEETKTPGFDRREMGI V NIGAEGSVHAGNESWSLNRGDVLYLGMGAG
PVT FEGAGRFYLV SAPAHRSLPNRLVTPADSKEVKLGAL ETSNKRTINQFIHPLVMES
CQLVLGYTTLEDGSVWNTMPAHVHDDRMEAYLYFGMD ETSRVLHLMGEPQQTRH
LFVANEEGAISPWSIHAGAGIGSYTFIWAMAGDNVDYTDMEFIQPGDLR*

SEQ ID NO: 8 pSGI-434_Q46938-タンパク質

MDVRQSIHSAHAKTLDTQGLRNEFLVEKVFVADEYTMVYSHIDRIIVGGIMPITKTVS
VGGEV GKQLGVSYFLERRELGVINIGGAGTITVDGQCYEIGHRDALYVGKGAK EVV
FASIDTGT PAKFYYNCA PAHTTYPTKKVTPDEVSPVTLGDNLT SNRRTINKYFVPDVL
ETCQLSMGLTELAPGNLWNTMPCHTHERRMEVYFYFNMDDDACVFHMMGQPQET
RHIVMHNEQAVISPSWSIHSGVGTKAYTFIWGMVGENQVFDDMDHVA VKDLR

10

SEQ ID NO: 9 pSGI-435; 遺伝子 #3891-タンパク質

MTMKILY GAGPEDVKGYDTQRLRDAFLDDLFADDRVSFTYTHVDRLILGGAVPVT
TSLTFGSGTEIGTPYLLSAREMGIANLGGTGTIEVDGQRFTLENRDVLYVGRGARQM
TASSLSAERPARFYMN SVPAGADFP HRLITRGEAKPLDLGDARRSNRRRLAMYIHPE
VSPSCLLLMGITDLAEGSAWNTMPPHLHERRMEAYCYFDLSPEDRVIHMMGRPDET
RHLVVADGEAVLSPAWSIHMGAGTGPYAFVWGMTGENQEYNDVAPVAVADLK*

20

SEQ ID NO: 10 pSGI-436; 遺伝子 #7102-タンパク質

MLTVETRHAIDPQTAKRMDTEELRKHFHMGSLFAAGEIRLVYTHYDRMIVGA AVPS
GAPLVLDQVKECGTASILDRREMAVVNVGASGKVSAAGETYAMERGDVLYLPLGS
GKVTFEGEGRFYILSAPAHAAYPARLIRIGEA EKVKLGSAETSNDRTIYQFVHPAVMT
SCQLVVG YTQLHNGSVWNTMPAHVHDDRMEAYLYFDMKPEQRFVHFMGEPQETR
HLVMKNEDAVVSPWSIHCGAGTGSYTFIWAMAGDNVDYKDVEMVAMEDLR*

30

SEQ ID NO: 11 pSGI-437; 遺伝子 #9209 -タンパク質

MSYLLRKPQSNEVSNGVKLVHEVTKSNSDLTYVEFKVLDLASGSSYAEELKKQEICI
VAVTGNITVTDHESTFENIGTRESVFERKPTDSVYISNDRSFEITAVSDARVALCYSPS
EKQLPTKLIKAEDNGIEHRGKFSNKRTVHNILPDS DPSANSLLVVEVYTD SGNWSSYP
PHKHDQDNLPEESFLEETYYHELDPGQG FVFQRVYTDDRSIDETMTVENENVVIVPA
GYHPVGVPDGYTSYYLNV MAGPTRKWKFHNDPAHEWILER*

40

SEQ ID NO: 12 pSGI-438; 遺伝子 #9732 -タンパク質

MANLLRKPNGTHGKVHDITPENAKWGYVGFGLFRLKSGESVSEKTGSTEVILVLVE
GKAKISASGEDFGEMGERLNVFEKLPPHCLYVPAESDWHATATTD CVLAVCTAPGK
PGRKAQKLGPESLTLEQRGKGANTRFIHNIAMESRDVADSLLVTEVFTPQGNWSSYP

PHRHDEDNFPDMTYLEETYYHRLNPAQGFGFQRVFTEDGSLDETMAVSDGDVVLVP
KGHHPCGAPYGYEMYLLNVMAGPLRKWRFKNHPDHDWIFKRDNP*

SEQ ID NO: 13 pSGI-439; 遺伝子 #7403-タンパク質

MASLLVRPTAPDAQGTVIDVTPESAGWTHVGFVRVHKLAKGQRLEASSDDQEVCLVL
LTGRATVTCGEHRFEDIGQRMDFEQIPPYAVYLPDHVSYAVEATTDLELAVCTAPG
HGNHAPRLIAPDNIKQSTRGQGTNTRHVHDILPETEPADSLVVEVFTPAGNWSSYPP
HKHDVDNLPHEHLEETYYHRINPEQGFAFQRVYTDDRSLDETMAVENGCCVLVPK
GYHPVGASHGYSLLYLLNVMAGPKRAWKFHNDPDHEWLMNAG*

10

SEQ ID NO: 14 pSGI-440; 遺伝子 F0J748-タンパク質

MPDLLRKPFGTHGKVHDITPAAAGWRHVGFGLYRLRAGEFAAEATGGNEVILVMV
EGKASIRAAGRDWGVLGERSVFEKSPPHSLYVPNGAEWALVAETDCIVAVCSAPG
RGGHAARRIGPEGIVLTARGE GTNTRHINNIA MEAEDYCDALLVTEVFTPAGHWSSY
PSHRHDEDDDPRTYLEETYYHRLNPASGFGVQRVYTDDRALDQTMVSDGDVVLV
PRGHHPCAAPYGIEMYLLNVMAGPLRKWRFLPDPELGIK

SEQ ID NO: 15 pSGI-458; 遺伝子 A5YBJ4-タンパク質

20

MSLLYHKQNQELSSGVRLIQDVNASNSPMKYTAVKVLEFSADSSYEETLEAFEAGIV
VLEGKVTITADDQTFEDVGQRTSIFDKIPTDSVYVSTGLAFGIRAKQAAKILIA YAPT
N QTFP

VRLIRGNIHQVEHRGKYNNKRLVQNILPDNLFPADKLLLVEVYTDSANWSSYPPHRH
DHDDLPAESLLEEIYYHEMRPKQGFVFQRVYTDDLSDLDETMAVQNQDVVVVVPKGY
HPVGVPDGYDSYLLNVMAGPTRVWHFHNAPEHAWIIDRQ

SEQ ID NO: 16 pSGI-478; 遺伝子 #1874-タンパク質

30

MKKFMDENFLLQTETAQKLYHNHAANMPIFDYHCHINPKDIAEDRMFKTITEIWLY
GDHYKWRAMRTNGVDERFCTGDASDWEKFEKWAETVPHTLRNPLYHWHLELKK
FFGINEILSPKNAREIYDACNEKLQTPAYSCRNIIRMANVHTICTTDDPVD TLEYHQQI
KEDGFEVAVLPWRPDKAMMVEDPKFFNDYMDQLAEAAGIHIESFEDLMEALDTR
HQYFHDNGCRLSDHGLDTVFAEDYTEEEIKAIFKKIRGGSRLSETEILKFKSCMLY
EY GVM DHSRGWTQQLHIGAQRNNNTRLFKKLGPDTGFDSIGDKPIAEPLAKLLDR
LDQ ENKLCKTVLYNLNPRDNELYATMLGNFQDGSVPGKIQYGSWWFLDQKDGM
IKQ MNALS NLGLLSRFVGM LTDSRSFLSYTRHEYFRRTL CNLLGNDVENGEIPAD
MELLG SMVENICFNNAKNYFNF*

40

SEQ ID NO: 17 pSGI-479; 遺伝子 Q9WXR9-タンパク質

MFLGEDYLLTNRAAVRLFNEVKDLPIVDPHNHLDAKDIVENKPWNDIWEVEGATDH
YVWELMRRCGVSEEYITGSRSNKEKWLALAKVFPRFVGNPTYEWIHLDLWRRFNIK

KVISEETAEEIWEETKKKLPEMTPQKLLRDMKVEILCTTDDPVSTLEHHRKAKEAVE
GVTILPTWRPDRAMNVDKEGWREYVEKMGERYGEDTSTLDGFLNALWKSHEHFKE
HGCVASDHALLEPSVYYYVDENRARAVHEKAFSGEKLTQDEINDYKAFMMVQFGKM
NQETNWWVTQLHIGALRDYRDSLFKTLGPDSGGDISTNFLRIAEGRLYFLNEFDGKLKI
VLYVLDPTHLPTISTIARAFPNVYVGAPWWFNDSPFGMEMHLKYLASVDLLYNLAG
MVTDSRKLLSFGSRTEMFRRVLSNVVGEMVEKGQIPIKEARELVKHVSYDGPKALFF
G

SEQ ID NO: 18 pSGI-480; 遺伝子 Q9KFI6-タンパク質

MSINSREVLAEKVKNVNNQPVTDMHHLFSPNFGEILLWDIDELLTYHYLVAEVM
RWTDVSIKAFWAMSKREQADLIWEELFIKRSPVSEACRGVLTCLQGLGLDPATRDLO
VYREYFAKKTSEEQVDTVLQLANVSDVVMTNDPFDDNERISWLEGKQPDSRFHAAL
RLDPLLNEYEQTKHRLRDWGYKVNDWNEGSIQEVKRFLTDWIERMDPVYMAVSL
PPTFSFPEESNRGRIIRDCLLPVAEKHNIPFAMMIGVKKRVHPALGDAGDFVGKASM
DGVEHLLREYPNNKFLVTMLSRENQHELVLARKFSNLMIFGCWWFMNNPEIINEM
TRMRMEMLGTSFIPQHSDARVLEQLIYKWHHSKSIIAEVLIDKYDDILQAGWEVTEE
EIKRDVADLFSRNFWRVGRNDHVTSVKVEQQT

SEQ ID NO: 19 pSGI-481; 遺伝子 O34808-タンパク質

MEPFMGKNFLLKNETAVSLYHNYAKDMPIIDYHCHLSPKEIYENKTFQNITEAWLYG
DHYKWRIMRANGIEETYITGDAPDEEKFMAWAKTVPMAGNPLYNWNTHLELQRRFFG
IYEILNEKSGSAIWKQTNKLLKGEGFGARDLIVKSNVKVVCTTDDPVDSLEYHLLLK
EDKDFPVSVLPGRFPDKGLEINREGFPEWVQALEDAAAISITTYDEFLKALEKRVRF
HSAGGRVSDHAIDTMVFAETTKEEAGRIFSDRLQGTEVSCDEKKFKTYTLQFLCGL
YAELDWAMQFHINALRNTNTKMMKRLGPDGTGYDSMNDEEIAKPLYKLLNSVEMKN
QLPKTILYSLNPNDNYVIASMINSFQDGITPGKIQFGTAWWFNDTKDGMLDQMKALS
NVGLFSRFIGMLTDSRSFLSYTRHEYFRRIVCNLIGEWVENGEVPRDMELLGSIVQGI
CYDNAKHVFQFQEEKANV

SEQ ID NO: 20 pSGI-433; 遺伝子 #8938-DNA

ATGCTCAACGTGGAAACGAGGCACGCCGTTACGCGGATCACGCGAGATCACTC
GACACAGAGGGCCTGCGCCGGCACTTCCTGGCCCAGGGCCTGTTTGCGGAGGGC
GAGATACGGCTGATCTATACGCATTATGATCGATTCTGTCATGGGAGGGCGCCGTGC
CGGACGGCGCGCCACTTGTGCTCGATCATGTGCGAGGAGACGAAAACGCCGGGCT
TTCTCGACCGACGGGAGATGGGAATCGTCAATATCGGTGCTGAGGGCAGCGTGC
ATGCCGGCAACGAAAGCTGGTCGCTGAACCGTGGTGACGTACTTTATCTCGGCAT
GGGGGCGGGACCGGTCACCTTCGAAGGGGCTGGGCGCTTCTACCTCGTCTCGGC

10

20

30

40

ACCGGCGCATCGCAGCCTGCCCAACCGGCTCGTCACGCCGGCCGACAGCAAGGA
GGTCAAGCTTGGCGCTCTCGAGACTTCCAACAAACGCACCATCAATCAGTTCATT
CATCCCCTGGTCATGGAAAGCTGCCAGCTCGTGCTGGGATATACCACGCTGGAGG
ACGGCTCGGTCTGGAATACCATGCCCCGCGCATGTGCACGACCGACGCATGGAGG
CCTATCTCTATTTTCGGCATGGATGAGACATCGCGGGTTCTGCATCTGATGGGCGA
GCCGCAGCAAACGAGGCATCTCTTCGTCGCCAATGAGGAAGGGGGCGATCTCTCC
GCCGTGGTCCATCCATGCGGGAGCAGGCATTGGCAGCTATACCTTCATCTGGGCC
ATGGCGGGCGACAATGTCGATTATACCGACATGGAGTTCATCCAGCCGGGAGAT
CTTCGATGA

10

SEQ ID NO: 21 pSGI-434; 遺伝子 Q46938-タンパク質

ATGGACGTAAGACAGAGCATCCACAGTGCGCACGCAAAAACGCTGGATACCCAA
GGGCTGCGCAATGAATTTTTGGTTGAAAAGGTATTTGTGCGCCGATGAGTACACCA
TGGTTTACAGCCACATTGACCGAATTATTGTTGGCGGCATTATGCCGATAACTAA
AACGGTTTCCGTTGGCGGGGAAGTTGGTAAACAACCTCGGCGTAAGCTATTTCTT
GAACGTCGCGAGTTAGGTGTTATCAATATTGGCGGTGCCGGTACGATTACTGTGCG
ATGGCCAATGCTATGAAATCGGTACCGCGACGCCCTGTATGTTGGTAAAGGTGC
AAAAGAAGTTGTCTTTGCCAGTATTGATACCGGCACTCCGGCGAAGTTTTATTAC
AATTGCGCACCCGCGCATACGACGTATCCCACCAAAAAAGTCACACCGGACGAA
GTATCTCCAGTCACGTTAGGCGATAACCTCACCAGTAACCGTCGCACGATTAACA
AATATTTTGTCCCGGATGTACTGGAAACCTGCCAATTGAGTATGGGGCTGACGGA
GCTGGCTCCGGGTAACTTGTGGAACACCATGCCGTGTACACCCACGAGCGCCG
GATGGAAGTTTATTTCTATTTCAATATGGATGATGACGCCTGCGTTTTCCACATGA
TGGGGCAGCCGCAAGAAACGCGTCATATTGTGATGCATAACGAGCAGGCGGTGA
TCTCCCCGAGCTGGTCGATCCATTCCGGTGTGCGAACC AAAAGCTTATACCTTTAT
CTGGGGCATGGTCGGTGAAAACAGGTCTTTGATGATATGGACCATGTGGCCGTT
AAAGATTTGCGCTAG

20

30

SEQ ID NO: 22 pSGI-435; 遺伝子 #3891-タンパク質

ATGACGATGAAGATACTCTACGGCGCCGGACCGGAGGATGTGAAAGGGTATGAC
ACGCAGCGCCTGCGCGACGCCTTCCTGCTGGACGACCTCTTCGCCGACGACCGGG
TCAGTTTCACATATACCATGTGCGATCGCCTCATCCTCGGCGGGGGCCGTCCCGGT
GACGACGAGCCTCACCTTCGGCTCCGGCACGGAGATCGGAACGCCCTACCTGCTT
TCCGCCCCGCGAGATGGGGATCGCCAATCTCGGCGGCACGGGCACGATCGAGGTG
GATGGCCAGCGCTTCACGCTCGAAAACCGCGACGTGCTCTATGTCGGTCGCGGC
GCCCCGCAGATGACCGCCTCCAGCCTGTGCGCGGAGAGGCCAGCCCGCTTCTAC

40

ATGAATTCCGTGCCCCGCCGGCGCCGATTTCCCGCACCGTCTGATCACCCGCGGAG
AGGCCAAGCCCCCTCGATCTCGGCGATGCGCGCCGCTCGAACAGGCGCCGGCTCG
CAATGTACATCCATCCGGAGGTCTCGCCGTCCTGCCTGCTGCTCATGGGCATCAC
CGATCTTGCCGAGGGGACGCGCCTGGAACACCATGCCGCCGCATCTGCACGAGCG
GCGGATGGAGGCCTATTGCTACTTCGATCTCTCGCCCCGAGGACCGGGTCATCCAC
ATGATGGGTGCGCCCGGACGAAACCCGCCACCTTGTCGTGGCCGACGGCGAGGCG
GTCCTCTCTCCCGCCTGGTCGATCCATATGGGTGCCGGGACGGGGCCCTACGCCT
TCGTCTGGGGCATGACCGGCGAAAACCAGGAATACAACGACGTCGCTCCCGTAG
CCGTGGCTGATCTCAAATGA

10

SEQ ID NO: 23 pSGI-436; 遺伝子 #7102-タンパク質

ATGCTGACCGTCGAAACCCGCCACGCCATTGATCCGCAGACCGCAAAGCGGATG
GACACGGAAGAGCTGCGCAAGCATTTCACATGGGCAGCCTGTTTGCTGCCGGT
GAAATCCGCCTCGTCTACACCCACTATGACCGCATGATCGTCGGCGCTGCCGTGC
CCTCGGGCGCGCCGCTGGTGCTGGATCAGGTCAAGGAATGCGGCACCGCCAGCA
TCCTCGACCGCCGCGAGATGGCTGTCGTCAACGTGCGCGCCAGCGGCAAGGTCT
CTGCAGCAGGCGAAACCTACGCCATGGAACGCGGGCGACGTGCTCTATCTGCCGC
TGGGCTCCGGCAAGGTGACCTTCGAAGGCGAAGGCCGCTTCTACATTCTCTCCGC
TCCGGCCCACGCTGCTTACCCGGCCCCGCTGATCCGCATCGGCGAGGCCGAGAA
GGTCAAGCTCGGCTCGGCCGAGACCTCCAACGACCGCACCATCTACCAGTTCGTG
CATCCGGCGGTGATGACTTCCTGCCAACTCGTCGTCGGCTACACCCAGCTGCACA
ACGGCTCTGTCTGGAACACCATGCCCCGCCACGTGCATGACCGGCGCATGGAGG
CCTATCTCTATTTGACATGAAGCCGGAGCAGCGCGTGTTCCACTTCATGGGCGA
GCCGCAGGAAACCCGCCATCTGGTCATGAAGAACGAGGATGCGGTGGTCTCCCC
GCCCTGGTCCATCCACTGCGGCGCAGGCACCGGCAGCTACACCTTCATCTGGGCC
ATGGCCGGCGACAACGTCGACTACAAGGACGTGGAAATGGTCGCCATGGAGGAT
CTGCGGTGA

20

30

SEQ ID NO: 24 pSGI-437; 遺伝子 #9209 –DNA

ATGAGTTATTTGTTGCGTAAGCCGCAGTCGAATGAAGTGTCTAATGGGGTCAAAC
TGGTGCACGAAGTAACGAAATCCAACCTCTGATCTCACCTATGTAGAGTTTAAAGT
GTTAGATCTCGCTTCCGGTTCCAGCTATGCAGAAGAATTGAAAAACAGGAAAT
CTGTATTGTCGCGGTAACGGGAAACATTACAGTGACCGATCACGAGTCGACTTTT
GAGAATATCGGCACGCGTGAAAGCGTATTTCGAACGAAAACCGACAGACAGCGTC
TATATTTCAAATGACCGTTCCTTTGAGATCACAGCGGTCAGCGACGCAAGAGTGG
CGCTTTGCTATTCTCCATCGGAAAAACAGCTTCCGACAAAGCTGATCAAAGCGGA

40

AGACAATGGCATTGAGCATCGCGGGAAGTTTTCAAACAAACGTACTGTTCAACA
CATTCTTCCGGATTTCAGACCCTTCAGCTAACAGCCTATTAGTAGTTGAAGTCTAT
ACAGACAGCGGCAACTGGTCCAGCTATCCGCCTCATAAACATGATCAAGACAAT
TTGCCGGAGGAATCTTTTTTAGAAGAAACGTACTACCATGAGTTAGACCCGGGAC
AGGGCTTTGTGTTTCAGCGTGTATACACAGATGACCGCTCGATTGACGAGACAAT
GACTGTAGAAAATGAAAACGTTGTCATCGTTCCTGCAGGATACCACCCGGTAGG
CGTGCCGGACGGATACACATCCTACTATTTAAATGTCATGGCAGGGCCGACGCG
GAAATGGAAGTTTCATAATGACCCGGCGCATGAGTGGATTTTAGAACGTAA

10

SEQ ID NO: 25 pSGI-438; 遺伝子 #9732 -DNA

ATGGCCAATTTGTTGCGCAAGCCCAACGGCACGCATGGCAAGGTCCACGACATC
ACTCCGGAAAACGCCAAATGGGGTTATGTCCGGGTTCCGGGCTCTTTCGTCTCAAAT
CCGGCGAGAGTGTCTCCGAAAAGACCGGATCGACGGAGGTGATCCTTGTCTTGT
GGAAGGCAAGGCAAAGATTTCCGCTTCTGGCGAGGATTTCCGGCGAGATGGGTGA
ACGCTTAAACGTGTTTCGAGAACTGCCGCCACACTGCCTCTATGTGCCTGCTGAA
AGCGACTGGCATGCAACCGCCACGACAGATTGTGTTCTGGCTGTTTGCACCGCAC
CGGGCAAGCCAGGCCGCAAGGCACAGAAGCTTGGGCCGAAAGCTTGACACTTG
AACAACGCGGAAAAGGTGCCAATACCCGCTTTATCCATAATATCGCAATGGAAA
GCCGCGATGTTGCCGATAGCCTTCTTGTTACCGAGGTATTCACACCGCAGGGAAA
CTGGTCGTCCTATCCACCCACAGACACGACGAAGACAATTTTCCGGATATGACC
TATCTGGAAGAGACCTATTATCACCGTCTCAACCCGGCGCAGGGCTTCGGCTTCC
AGCGTGTTTTACCCGAAGACGGAAGCCTTGATGAAACCATGGCGGTCTCTGACG
GAGACGTCGTGCTTGTACCAAAGGCCACCATCCATGTGGCGCGCCCTATGGCTA
CGAGATGTATTATCTCAATGTGATGGCCGGTCCCTTGCGCAAATGGCGCTTCAAG
AACCATCCCGACCATGACTGGATTTTCAAACGCGACAATCCGTAA

20

30

SEQ ID NO: 26 pSGI-439; 遺伝子 #7403-DNA

ATGGCTTCCCTACTGGTACGCCCCACCGCCCCAGATGCCAGGGCACCGTGATTG
ACGTTACCCCTGAATCTGCTGGCTGGACGCACGTTGGCTTTCGGGTGCATAAACT
CGCCAAGGGCCAGCGCCTGGAGGCCAGCAGCGATGATCAGGAAGTCTGCCTGGT
GCTGCTCACCGGTCGCGCCACGGTAACTTGCGGCGAGCACCGCTTTGAAGATATT
GGCCAGCGTATGGATATTTTTGAGCAGATCCCTCCCTATGCGGTTTACCTACCTG
ACCATGTTAGCTACGCGGTGGAAGCGACCACAGACTTAGAGCTAGCGGTGTGCA
CCGCCCCCTGGGCATGGCAACCATGCCCCACGGCTCATCGCGCCTGACAACATCA
AGCAAAGCACCCGTGGCCAGGGCACCAACACCCGCCATGTTACGATATTCTGC
CGGAAACCGAGCCCGCCGATAGCCTATTAGTAGTCGAAGTATTCACACCTGCGG

40

GTA ACTGGTCGAGCTACCCGCCCCACAAACACGATGTGGATAACTTACCCACG
AATCACATCTGGAAGAGACCTACTACCACCGCATTAACCCTGAACAAGGGTTTCG
CCTTCCAGCGCGTTTACACCGATGACCGCAGCCTTGATGAAACCATGGCGGTGGA
AAACGGCTGCTGTGTGTTGGTTCCCAAGGGTTACCATCCGGTGGGCGCCTCCCAT
GGCTACTCGCTCTACTACTTAAATGTGATGGCGGGGCCCAAGCGGGCATGGAAA
TTTCACAACGACCCCGACCACGAATGGCTGATGAACGCTGGATAG

SEQ ID NO: 27 pSGI-440; 遺伝子 F0J748-DNA

ATGCCGGACTTACTGAGAAAACCGTTTGGCACCCATGGCAAAGTGCACGATATT
ACCCAGCAGCAGCAGGTTGGAGACATGTTGGTTTTGGCTTATATCGCTTAAGAG
CGGGCGAATTTGCAGCAGAAGCGACAGGCGGCAATGAAGTTATTCTGGTGATGG
TTGAGGGCAAAGCGTCTATTAGAGCAGCAGGCAGAGATTGGGGCGTTTTAGGCG
AACGTATGAGCGTCTTCGAAAAAAGTCCACCACATTCCCTGTATGTCCCGAATGG
TGCAGAATGGGCCTTAGTAGCCGAAACAGATTGCATTGTAGCAGTGTGTAGCGCT
CCGGGTAGAGGAGGTCATGCTGCAAGAAGAATTGGTCCTGAAGGTATTGTGTTA
ACCGCCAGAGGTGAAGGCACCAATACACGCCACATCAACAACATCGCCATGGAA
GCCGAAGATTATTGTGATGCCCTGTTAGTCACCGAAGTGTTACCCAGCCGGCC
ATTGGAGCTCTTATCCATCTCATCGTCATGATGAAGACGACGATCCGCGCATCAC
CTATTTAGAAGAGACCTACTATCATCGCTTAAATCCTGCCTCGGGCTTTGGCGTTC
AACGCGTCTATACCGATGATCGCGCCTTAGATCAAACCATGGCGGTTTCTGATGG
CGATGTTGTTTTAGTTCCTCGCGGCCATCATCCGTGTGCAGCCCCGTATGGTATTG
AAATGTATTACCTGAACGTCATGGCCGGCCCGTTACGTAAATGGCGCTTTTTACC
TGATCCTGAACTTGGCATTGCGAAATAA

SEQ ID NO: 28 pSGI-458; 遺伝子 A5YBJ4-DNA

ATGTCTCTGCTGTACCACAAGCAGAACCAGGAACTGAGTAGTGGTGTGCGCCTG
ATCCAAGATGTTAATGCCAGCAATAGCCCGATGAAATATACCGCCGTGAAAGTG
CTGGAGTTTAGCGCCGATAGCAGCTATGAGGAAACCTTAGAGGCCTTTGAAGCC
GGCATTGTTGTGTTAGAGGGCAAAGTGACCATCACCGCCGACGATCAAACCTTCG
AAGATGTGGGTCAAAGAACCTCGATCTTCGACAAAATCCCGACCGATAGCGTTT
ATGTGTCTACCGGTTTAGCCTTCGGTATTCGCGCCAAACAAGCCGCCAAAATCTT
AATCGCGTATGCTCCGACCAATCAGACCTTCCAGTTCGCTTAATTCGCGGCAAT
ATCCACCAGGTGGAACATCGCGGCAAGTACAACAACAAACGCTTAGTGCAGAAC
ATTCTCCCGGATAATCTCCCGTTCGCCGATAAATTACTGCTGGTTGAGGTGTACA
CCGATAGCGCCAATTGGAGCTCCTATCCGCCGCATAGACATGATCACGATGATTT
ACCGGCCGAAAGTCTGTTAGAGGAGATCTACTATCACGAAATGCGCCCGAAGCA

GGGCTTCGTCTTTCAACGCGTGTATACCGATGATCTGAGTCTGGATGAGACCATG
GCCGTTCAAAATCAAGATGTTGTCGTTGTCCCGAAAGGCTATCATCCGGTTGGTG
TCCCGACGGCTATGATTCGTATTACCTGAACGTGATGGCCGGCCCGACAAGAGT
GTGGCATTTTCATAATGCTCCGGAACATGCCTGGATTATTGATCGCCAGTAA

SEQ ID NO: 29 pSGI-478; 遺伝子 #1874-DNA

ATGAAAAAATTTATGGATGAAAATTTTCTGTTGCAAACCGAAACAGCGCAGAAA
TTGTATCATAATCACGCGGCAAACATGCCGATTTTCGATTACCACTGCCACATTA
ACCCCAAAGACATCGCGGAAGACCGGATGTTTAAAACCATCACCGAAATCTGGT
TGTACGGCGATCATTATAAATGGCGCGCCATGCGTACAAACGGCGTTGACGAGC
GCTTTTGCACCGGCGATGCAAGCGATTGGGAAAAGTTTGAAAAGTGGGCCGAAA
CGGTTTCCTCATACCCTGCGTAATCCGCTTTATCACTGGACACACCTGGAGCTAAA
GAAATTTTTCGGGATTAACGAGATCCTGAGTCCGAAAAATGCCCGGGAAATTTAT
GATGCCTGTAACGAAAAACTGCAAACGCCCGCGTATAGTTGCCGCAACATCATC
CGGATGGCCAATGTGCATACAATCTGTACCACCGACGACCCGGTTGACACACTG
GAATATCATCAGCAAATTAAAGAAGACGGCTTTGAAGTGGCGGTTTTACCTGCCT
GGCGTCCGGATAAAGCGATGATGGTGGAAAGACCCGAAGTTCTTTAACGACTATA
TGGACCAGTTGGCCGAAGCTGCCGGTATCCATATCGAATCGTTTGAGGATTTGAT
GGAAGCCTTGGATACGCGTCACCAGTATTTTCATGATAATGGTTGCCGTTTGTCC
GACCACGGGCTGGATACCGTTTTTGTCTGAAGATTATACGGAGGAAGAAATTAAA
GCGATCTTCAAAAAAATCCGTGGCGGCAGCAGGCTTAGCGAAACGGAAATCCTG
AAATTCAAGTCCTGCATGTTGTACGAATATGGGGTGATGGACCATTCGCGCGGCT
GGACACAACAATTGCACATTGGCGCACAACGCAACAACAACACCCGTTTGTTC
AAAAATTAGGTCCCGACACTGGTTTTCGATTTCGATTGGCGATAAGCCGATCGCTGA
ACCATTGGCCAAATTGCTCGACCGCTGGATCAGGAAAACAAATTGTGCAAAAC
GGTTTTGTATAATCTGAATCCGCGTGATAACGAGTTGTACGCTACCATGTTGGGC
AACTTTCAGGACGGATCGGTTCCCGGGAAAATTCAATACGGCTCGGGTTGGTGGT
TTCTCGATCAGAAAGACGGCATGATTAAACAGATGAATGCCCTTCCAATCTGGG
TTTGCTGAGCCGTTTCGTAGGCATGCTGACCGACTCAAGGAGCTTCCTTTCGTAC
ACCCGTCACGAATATTTCCGTCGTACCCTTTGCAACCTGCTTGGGAATGATGTTG
AAAACGGGGGAGATTCCGGCAGATATGGAGCTTTTGGGCAGTATGGTTGAGAATA
TTTGTTTTAATAACGCGAAGAAGTATTTTAATTTTAG

SEQ ID NO: 30 pSGI-479; 遺伝子 Q9WXR9-DNA

ATGTTTCTGGGCGAAGACTATCTGCTGACCAATCGTGCGGCAGTTCGTCTGTTCA
ACGAAGTGAAAGATCTGCCGATCGTTGATCCGCATAACCACCTGGATGCGAAAG

10

20

30

40

ATATCGTGGAACAAACCGTGGAACGACATCTGGGAAGTGGAAGGTGCGACCG
ATCACTATGTGTGGGAAGTATGCGTCGTTGTGGTGTAGCGAAGAATATATTAC
CGGCTCTCGTAGCAACAAAGAAAAATGGCTGGCGCTGGCGAAAGTGTTTCCGCG
TTTTGTGGGTAATCCGACGTACGAATGGATCCACCTGGATCTGTGGCGTCGTTTC
AACATCAAAAAAGTCATCAGCGAAGAAACCGCGGAAGAAATCTGGGAAGAAAC
CAAAAAAAAGTGGCGGAGATGACCCCGCAGAACTGCTGCGCGACATGAAAGT
GGAAATCCTGTGCACCACCGATGATCCGGTGTCTACCCTGGAACATCACCGTAAA
GCGAAAGAAGCCGTGGAAGGCGTGACCATTTTACCGACCTGGCGTCCGGATCGT
GCAATGAATGTTGATAAAGAAGGTTGGCGTGAATATGTTGAAAAAATGGGTGAA
CGCTATGGCGAAGATAACCAGCACCTGGATGGTTTTCTGAATGCCCTGTGGAAAA
GCCACGAACACTTCAAAGAACACGGCTGTGTGGCGAGCGATCATGCGCTGCTGG
AACCGAGCGTGTACTACGTGGATGAAAACCGCGCGCGTGCAGTTCATGAAAAAG
CATTTTCTGGTGAAAACTGACTCAAGATGAAATCAACGACTATAAAGCGTTCAT
GATGGTGCAGTTCGGCAAAATGAACCAGGAAACCAACTGGGTGACCCAGCTGCA
CATTGGTGCCCTGCGCGATTACCGCGATAGCCTGTTCAAACCCCTGGGCCCGGAT
TCTGGTGGCGATATCAGCACCAACTTTCTGCGTATTGCTGAAGGTCTGCGTTATTT
TCTGAACGAATTTGATGGTAACTGAAAATTGTGCTGTACGTGCTGGATCCGACC
CATTTACCGACCATTTTCGACCATTGCACGTGCGTTCGCCAACGTGTATGTGGGTG
CACCGTGGTGGTTCAACGATAGCCCGTTCGGCATGGAAATGCACCTGAAATACCT
GGCGAGCGTTGATCTGCTGTACAATCTGGCTGGTATGGTTACCGATTCACGTAAA
TTACTGAGTTTTGGTTCTCGTACCGAAATGTTTCGTCGCGTTCTGTCTAATGTGGT
TGCGGAAATGGTGGAAAAAGGCCAGATCCCGATCAAAGAAGCGCGCGAACTGGT
GAAACACGTGAGCTACGACGGCCCGAAAGCCCTGTTCTTTGGCTGA

10

20

30

SEQ ID NO: 31 pSGI-480; 遺伝子 Q9KFI6-DNA

ATGAGCATCAACAGCCGTGAAGTTCTGGCGGAAAAAGTGAAAAACGCGGTGAAC
AACCAGCCGGTTACCGATATGCATACCCACCTGTTTAGCCCGAACTTTGGCGAAA
TTCTGCTGTGGGACATCGATGAACTGCTGACCTATCACTACCTGGTTGCGGAAGT
TATGCGTTGGACCGATGTGAGCATTGAAGCGTTTTGGGCAATGAGCAAACGTGA
ACAGGCCGATCTGATTTGGGAAGAACTGTTTCATCAAACGCAGCCCGGTGAGCGA
AGCATGTCGTGGCGTTCTGACCTGTTTACAAGGTTTAGGTCTGGATCCGGCAACT
CGTGATTTACAGGTGTATCGTGAATACTTCGCCAAAAAAACCAGCGAGGAACAG
GTGGATACCGTTCTGCAGCTGGCAAATGTGAGCGATGTGGTGATGACCAATGATC
CGTTCGATGATAATGAACGCATCAGCTGGCTGGAAGGCAAACAGCCGGATAGCC
GCTTTCATGCAGCGTTACGTCTGGATCCGCTGCTGAATGAATATGAACAGACCAA

40

ACATCGTCTGCGTGATTGGGGTTATAAAGTGAACGACGAATGGAACGAAGGCAG
CATCCAGGAAGTGAAACGCTTTCTGACCGACTGGATTGAACGTATGGATCCGGTG
TATATGGCGGTGAGCTTACCGCCGACCTTCAGCTTTCCGGAAGAATCGAACCGTG
GCCGCATTATCCGTGATTGTCTGTTACCGGTTGCAGAAAAACATAACATCCCGTT
TGCAATGATGATTGGCGTGAAAAACGCGTGCATCCGGCGTTAGGTGATGCAGG
CGATTTTGTGGGTAAAGCAAGTATGGATGGCGTTGAACACCTGCTGCGCGAATAC
CCGAACAACAAATTTCCTGGTGACCATGCTGAGCCGCGAAAAACCAGCACGAACTG
GTGGTTCTGGCGCGTAAATTTAGTAACCTGATGATTTTTGGTTGTTGGTGGTTTAT
GAACAACCCGGAGATCATCAACGAAATGACCCGCATGCGCATGGAAATGCTGGG
TACCAGCTTTATCCCGCAGCACAGCGATGCCCGTGTCTGGAACAGCTGATCTAT
AAATGGCACCACAGCAAAAAGCATCATCGCGGAAGTCCTGATCGACAAATACGAC
GACATCCTGCAAGCAGGTTGGGAAGTTACCGAAGAAGAAATCAAACGTGATGTG
GCAGATCTGTTTAGCCGCAACTTTTGGCGCTTTGTGGGCCGTAACGATCACGTGA
CCAGCGTGAAAGTGGAACAGCAGACCTGA

10

SEQ ID NO: 32 pSGI-481; 遺伝子 O34808-DNA

20

ATGGAACCGTTTATGGGCAAAAACCTTCCTGCTGAAAAACGAGACCGCGGTGAGC
CTGTACCACAACCTACGCGAAAGATATGCCGATCATCGACTACCATTGCCATCTGA
GCCCGAAAGAAATCTACGAGAACAAAACCTTCCAGAACATCACCGAAGCGTGGC
TGTACGGCGATCACTACAAATGGCGCATCATGCGTGCGAATGGCATCGAAGAAA
CCTATATTACCGGTGATGCACCGGACGAAGAAAAATTCATGGCGTGGGCGAAAA
CCGTGCCGATGGCCATTGGTAATCCGCTGTATAACTGGACCCATCTGGAACCTGCA
ACGTTTTTTTTGGCATCTACGAAATCCTGAACGAAAAAAGCGGCAGCGCGATCTGG
AAACAGACCAACAACTGCTGAAAGGCGAAGGCTTTGGTGCGCGTGATCTGATC
GTGAAAAGCAACGTATAAGTGGTGTGCACCACCGACGATCCGGTGGATTCTCTG
GAATACCATCTGCTGCTGAAAGAAGACAAAGACTTCCCGGTTAGCGTTTTACCGG
GTTTTTCGTCCGGATAAAGGTCTGGAAATCAACCGTGAAGGCTTTCCGGAATGGGT
TCAAGCCCTGGAAGATGCGGCCGCAATTAGCATTACGACCTATGATGAATTTCTG
AAAGCGCTGGAAAAACGCGTGCGCTTCTTCCATAGTGCGGGTGGTCGTGTTAGCG
ATCATGCAATCGATACCATGGTTTTCCGCCGAAACCACCAAAGAAGAAGCGGGTC
GCATTTTTAGTGATCGTCTGCAAGGCACCGAAGTTAGCTGCGAAGACGAGAAAA
AATTCAAAACCTACACCCTGCAGTTTCTGTGTGGCCTGTATGCCGAACTGGACTG
GGCAATGCAGTTTCACATCAACGCGCTGCGCAACACCAACACCAAAATGATGAA
ACGCCTGGGTCCGGATACCGGTTATGATAGCATGAACGATGAAGAAATCGCGAA
ACCGCTGTACAACTGCTGAACAGCGTGGAATGAAAAACCAACTGCCGAAAAC

30

40

CATCCTGTACAGCCTGAACCCGAACGACAACACTACGTGATCGCGAGCATGATCAA
CAGCTTCCAGGATGGCATCACCCCGGGCAAATTCAGTTTGGCACCGCATGGTGG
TTCAACGATACCAAAGATGGTATGCTGGATCAGATGAAAGCACTGAGCAATGTG
GGCCTGTTTAGCCGTTTTATTGGCATGCTGACCGATAGCCGTAGCTTTCTGAGCTA
TACCCGTCACGAATACTTTCGCCGCATTGTGTGTAACTGATCGGCCGAATGGGTG
GAAAACGGCGAAGTTCCGCGCGATATGGAAGTCTGGGTAGTATTGTGCAAGGT
ATTTGCTACGATAACGCGAAACATTACTTCCAGTTCCAGGAGGAAAAAGCGAAC
GTGTGA

10

SEQ ID NO: 33 pSGI-359-0385-タンパク質

MSQTPRKLRSQKWFDPAHADMTAIYVERYLNYGLTRQELQSGRPIIGIAQTGSDLAPCNRH
HLALAERVKAGIRDAGGIPMEFPVHPLAEQGRPTAALDRNLAYLGLVEILHGYPLDGVVLT
TGCDKTTTACLMMAATVDLPAIVLSGGPMLDGWHDGQRVGSGTVIWHARNLMAAGKLDY
EGFMTLATASSPSVGHCMGTALSMNSLAEALGMSLPTCASIPAPYRERAQMAYATGMRI
CDMVREDLRPSHILTRQAFENAIIVASALGASTNCPPHLIAMARHAGIDLSLDDWQRLGEDV
PLLVNCVPAGEHLGEGFHRAGGVPVAVMHLEFAAGRLHPDCPTVSGKTIGDIAAGAKTRDAD
VIRSCAAPLKHAGFIVLSGNFFDSAIKMSVVGAEFRRAYLSEPGSENAFEARAIVFEGPEDY
HARIEDPALNIDEHCILVIRGAGTVGYPGSAEVVNMAPPSHLIKRGVDSLPLCLGDGRQSGTSG
SPSILNMSPEAAVGGGLALLRTGDKIRVDLNQRSVTALVDDAEMARRKQEPYQAPASQTP
WQELYRQLVGQLSTGGCLEPATLYLKVIETRGDPRHSH

20

SEQ ID NO: 34 pSGI-360-0336-タンパク質

MSERIKKMNDQNKRIFLRSQEWFDDEHADMTALYVERYMNYGLTRAELQSGRPIIGIAQTG
SDLTPCNRHHKELAERVKAGIRDAGGIPMEFPVHPHIAEQTRRPTAALDRNLAYLGLVEILHGY
PLDGVVLTGCDKTTTACLMMAATTDIPAIVLSGGPMLDGHFKGELIGSGTVLWHARNLLAT
GEIDYEGFMEMTTSASPSVGHCMGTALSMNALAEALGMSLPTCASIPAPYRERGMAYM
TGKRICEMVLEDLRPSKIMNKQSFENAIIVASALGASSNCPPHLIAIARHMGIELSLEDWQRV
GENIPLIVNCMPAGKYLGEFHRAGGVPVAVLHELQKASVLHEGCASVSGKTMGEIAKNAKT
SNVDVIFPYEQPLKHGAGFIVLSGNFFDSAIMKMSVVGAEFKKTYLSDPNGENSFEARAIVFE
GPEDYHARINDPALDIDEHCILVIRGAGTVGYPGSAEVVNMAPPAELIKKGIDSLPLCLGDGRQ
SGTSASPSILNMSPEAAVGGGIALKTNDRRLRIDLNKRSVNVLISDEELEQRRREWKPTVSSSQ
TPWQEMYRNMVGQLSTGGCLEPATLYMRVINQDNLPRHSH

30

SEQ ID NO: 35 pSGI-365 E3HJU7-タンパク質

MSQTPRKLRSQKWFDPAHADMTAIYVERYLNYGLTRQELQSGRPIIGIAQTGSDLAPCNRH
HLALAERIKAGIRDAGGIPMEFPVHPLAEQGRPTAALDRNLAYLGLVEILHGYPLDGVVLT
GCDKTTTACLMMAATVDIPAIVLSGGPMLDGWHDGQRVGSGTVIWHARNLMAAGKLDYEG
FMTLATASSPSIGHCMGTALSMNSLAEALGMSLPTCASIPAPYRERGMAYATGLRICDM
VREDLRPSHVLTRQAFENAIIVASALGASSNCPPHLIAMARHAGIDLSLDDWQRLGEDVPLL

40

VNCVPAGEHLGEGFHRAGGVPAVLHELAAAGRLHMDCATVSGKTIGEIAAAAKTNNADVIR
SCDAPLKHRAGFIVLSGNFFDSAIHKMSVVGEAFRRAYLSEPGSENAFEARAIVFEGPEDYHAR
IEDPTLNIDEHCILVIRGAGTVGYPGSAEVVNMAPPSHLLKRGIDSLPCLGDGRQSGTSASPSIL
NMSPEAAVGGGLALLRTGDRIRVDLNQRSVIALVDQTEMERRKLEPPYQAPESQTPWQELY
RQLVGQLSTGGCLEPATLYLKVVETRGRDPRHS

SEQ ID NO: 36 pSGI-359-0385-DNA

ATGTCTCAGACACCCGCAAGTTGCGCAGCCAGAAATGGTTCGACGACCCTGCGCATGC
CGATATGACGGCGATTTACGTCGAGCGTTATCTGAATTACGGCCTGACGCGGCAAGAGTT 10
GCAGTCCGGGCGGCCGATCATCGGCATCGCCAGACCGGCAGCGATCTGGCGCCCTGCA
ACCGCCATCACCTGGCGCTGGCCGAGCGCGTCAAAGCGGGCATCCGGGACGCGGGCGGC
ATCCCGATGGAGTTCCCGTGCACCCGCTGGCCGAACAAGGCCGGCGGCCACGGCCGC
GCTGGACCGCAACCTGGCCTATCTGGGCCTGGTCGAAATCCTGCACGGCTACCCCTTGGA
CGGGGTGGTGCTGACGACTGGCTGCGACAAGACCACGCCTGCCTGCCTGATGGCCGCCG
CCACGGTCGACCTGCCCCGCCATCGTGCTGTCCGGCGGCCCATGCTGGACGGCTGGCACG
ACGGCCAGCGCGTCGGTTCGGGCACCGTCATCTGGCACGCGCGCAACCTGATGGCGGCC
GGCAAGCTTGATTACGAAGGCTTCATGACGCTGGCCACCGCGTCTTCGCCGTCGGTCGGC 20
CACTGCAACACCATGGGCACGGCGTTGTCGATGAATTCGCTGGCCGAAGCGCTGGGCAT
GTCGCTGCCACCTGCGCCAGCATTCCCGCCCCCTACCGCGAACGCGCCCAGATGGCCTA
CGCCACCGGCATGCGCATCTGCGACATGGTGCGCGAAGACCTGCGACCCTCCCACATCCT
GACACGGCAGGCATTTCGAGAACGCCATCGTTCGTGGCATCGGCGCTGGGCGCGTCCACCA
ATTGCCCCGCCGACCTGATCGCGATGGCCCCGCCACGCCGGCATCGACCTTAGCCTGGACG
ACTGGCAGCGCCTGGGTGAAGACGTGCCGCTGCTGGTCAACTGCGTGCCGGCGGGCGAG
CATCTGGGCGAGGGCTTCCACCGCGCGGGCGGCGTCCCCGCGGTCATGCATGAACTGTTC
GCCGCCGGGCGCCTTCACCCGACTGCCCCACCGTATCCGGCAAGACCATCGGGGACAT 30
CGCCGCGGGCGCCAAGACCCGCGACGCCGACGTCATCCGCAGCTGCGCCGCCCCGCTGA
AACACCGGGCAGGCTTCATCGTGCTGTGCGGCAATTTCTTCGACAGCGCCATCATCAAGA
TGTCGGTCGTAGGCGAAGCGTTCCGCCGCGCCTACCTGTCCGAACCCGGCTCAGAGAAC
GCCTTCGAGGCCCGCGCCATCGTGTTTCGAAGGCCCGAGGACTACCACGCGCGCATCGA
AGACCCGGCGCTGAACATCGACGAACACTGCATCCTTGTATCCGCGGCGCCGGCACCG
TGGGCTACCCGGGCAGCGCCGAAGTGGTCAACATGGCGCCGCGTCCCACCTGATCAAG
CGCGGCGTGGATTCCCTGCCGTGCCTGGGGGATGGCAGGCAAAGCGGCACTTCCGGCAG
CCCGTCCATTTTGAACATGTCCCTGAAGCAGCAGTCGGGGGAGGATTGGCGCTGCTGCG 40
CACCGGCGACAAGATCCGTGTCGATCTGAACCAGCGCAGCGTCACCGCCTTGGTCGACG
ACGCGGAAATGGCAAGACGGAAGCAAGAACCGCCCTACCAGGCACCGGCCTCGAAAC
GCCCTGGCAAGAGCTGTACCGGCAACTGGTTCGGCCAGTTGTGACGGGCGGCTGCCTGG
AGCCCGCGACGCTATATCTGAAAGTCATCGAAACGCGCGGCGATCCCCGGCACTCTCACT
GA

SEQ ID NO: 37 pSGI-360-0336-DNA

ATGAGTGAAAGGATCAAAAAAATGAATGATCAAAATAAACGGATTTTTTTACGTAGCCA
AGAATGGTTTTGATGATCCTGAACATGCTGACATGACAGCACTCTATGTTGAGCGTTATAT
GAATTATGGCCTGACCCGTGCCGAGCTACAATCAGGCCGCCCGATTATTGGTATTGCACA
AACTGGCAGTGATTTAACTCCATGTAACCGTCACCACAAAGAACTTGCTGAACGGGTAA
AGCAGGTATTCGAGATGCGGGAGGTATTCCCATGGAATTCCCCGTTACCCGATTGCAGA
ACAAACCCGTCGCCCTACTGCTGCACTTGATAGAAATTTAGCTTACTTAGGCTTAGTTGA
AATATTGCATGGTTATCCGCTTGATGGTGTGGTGCTAACCACAGGTTGTGACAAAACACTAC
ACCTGCTTGTTTAATGGCTGCCGCAACGACAGATATACCAGCCATTGTGTTGTCTGGTGG
ACCAATGCTAGATGGTCATTTTAAAGGTGAGTTAATTGGTTCTGGGACTGTGCTTTGGCA
TGCAAGAAATTTACTTGCCACGGGTGAAATTGATTATGAAGGGTTCATGGAAATGACCA
CTTCAGCATCGCCTTCGGTCGGACATTGCAACACCATGGGCACTGCACTTTCTATGAATG
CCTTGGCAGAAGCTTTGGGCATGTCTTTACCGACATGTGCAAGTATTCCAGCGCCGTATC
GCGAACGAGGGCAAATGGCCTATATGACAGGCCAAAAGAATTTGTGAAATGGTTTTAGAA
GATTTACGCCCTTCTAAAATCATGAACAAACAATCATTTGAAAATGCCATCGCGGTAGCT
TCAGCATTAGGGGCATCAAGTAATTGCCCTCCTCACCTCATTGCAATTGCCCGTCATATG
GGCATTGAGCTCAGTTTAGAAGACTGGCAACGCGTTGGGGAGAACATTCCTCTCATTGTG
AACTGTATGCCTGCGGGTAAATATTTAGGTGAAGGTTTTACCGTGCTGGCGGTGTTCTT
GCTGTTTTGCATGAATTACAAAAGGCCAGCGTTTTACATGAAGGCTGTGCATCAGTCAGC
GGTAAAACGATGGGAGAAATTGCTAAAAATGCTAAAACCTCCAATGTAGATGTTATTTTT
CCATATGAACAACCATTAAAACATGGTGCAGGTTTTATTGTGCTTAGTGGCAATTTCTTC
GACAGCGCCATTATGAAAATGTCTGTTGTGGGTGAAGCATTTAAGAAAACCTATTTATCT
GACCCAAATGGGGAAAATAGCTTTGAAGCACGGGCAATCGTTTTTTGAAGGGCCAGAGGA
CTACCATGCACGAATTAATGATCCAGCCTTAGACATTGATGAACATTGTATTTTGGTCAT
TCGTGGCGCTGGAACAGTGGGCTATCCAGGTAGTGCAGAAGTTGTAAATATGGCTCCAC
CCGCAGAGTTAATTA AAAAAGGCATCGATTCACTGCCTTGCTTAGGAGATGGCCGCCAA
AGTGGTACGTCTGCCAGCCCTTCTATTTTAAATATGTCACCCGAAGCGGCGGTAGGCGGT
GGAATTGCATTATTAAGACCAATGACCGTTTACGCATTGATCTCAATAAACGCTCCGTC
AACGTACTCATTTCTGACGAAGAGTTAGAACAACGCCGCCGTGAGTGGAAACCGACGGT
CTCTTCATCTCAAACACCTTGGCAAGAAATGTATCGCAACATGGTGGGTCAATTATCCAC
TGGCGGTTGTTTGGAAACCTGCAACTTTATATATGCGAGTCATAAATCAAGACAACCTTCC
AAGACACTCTCATTA

10

20

30

40

SEQ ID NO: 38 pSGI-365 E3HJU7-DNA

ATGAGCCAAACACCGCGTAAATTACGCAGCCAGAAGTGGTTTGACGATCCTGCACATGC
CGATATGACCGCCATCTATGTTGAACGCTACCTGAACTATGGCTTAACCCGCCAAGAACT
GCAAAGTGGTCGCCCCGATTATTGGTATTGCCCAAACCGGCAGCGATTTAGCCCCGTGTAA
TCGCCATCATTTAGCCTTAGCCGAACGCATTAAAGCAGGCATTAGAGATGCAGGCGGCA

TTCCTATGGAATTTCCCGTTCATCCGCTGGCCGAACAAGGTAGACGTCCTACAGCAGCAT
TAGATCGCAATTTAGCCTATTTAGGCCTGGTGGAAATTTTACACGGCTATCCCCTGGACG
GTGTGGTGCTGACAACCGGTTGCGATAAAACAACACCGGCGTGTTTAATGGCAGCTGCA
ACAGTTGATATTCCGGCGATCGTGTTATCAGGTGGTCCGATGTTAGATGGCTGGCATGAT
GGCCAAAGAGTTGGCAGTGGTACCGTGATTTGGCATGCACGCAATTTAATGGCAGCAGG
CAAACCTGGATTATGAAGGCTTCATGACCCTGGCGACAGCCTCTTCTCCGAGTATTGGACA
CTGTAATACCATGGGCACAGCCTTAAGCATGAATAGTCTGGCAGAAGCCCTGGGTATGTC
TTTACCGACCTGTGCGTCTATTCCAGCCCCGTATAGAGAACGCGGTCAAATGGCGTATGC
TACTGGTTTACGCATTTGCGATATGGTGCGCGAAGATTTACGCCCCGTCACATGTTTTAAC
CCGCCAAGCCTTCGAAAATGCCATTGTTGTTGCCTCAGCCTTAGGTGCAAGCTCTAATTG
TCCCCCTCATTTAATTGCCATGGCCCCGTCATGCCGGTATCGACTTAAGCCTGGATGACTG
GCAACGCTTAGGCCGAAGATGTTCCGTTACTGGTCAATTGTGTGCCTGCCGGTGAACATTT
AGGTGAAGGATTTTCATCGCGCGGGTGGTGTTTCTGCTGTTTTACATGAATTAGCTGCCGC
AGGTCGTTTACATATGGATTGTGCTACCGTTTCTGGCAAGACCATCGGCGAAATTGCAGC
TGCCGCAAAAACCAACAACGCAGACGTGATTCGCTCGTGTGATGCCCCGTTAAACATA
GAGCCGGCTTTATTGTGTTAAGCGGCAATTTCTTCGACTCCGCCATCATCAAGATGTCCG
TTGTGGGTGAAGCCTTTCGCAGAGCCTATTTAAGTGAACCTGGCAGCGAAAATGCCTTTG
AAGCCCGTGCCATCGTGTTTGAAGGCCCGGAAGACTATCATGCCCCGATTGAAGATCCG
ACCCTGAATATTGATGAACACTGCATTCTGGTGATTTCGCGGCGCAGGTACCGTTGGTTAT
CCTGGTAGTGCTGAAGTTGTGAATATGGCCCCGCCGAGCCATTTATTAACCGCGGTATT
GATTCATTACCTTGCCTGGGAGATGGCCGCCAAAGTGGTACCTCAGCTAGTCCGTCTATC
CTGAATATGAGCCCTGAAGCCGCCGTTGGAGGAGGTTTAGCATTATTAAGAACCGGTGA
TCGCATTTCGCGTCGATCTGAATCAACGCTCAGTCATTGCATTAGTCGACCAGACCGAAAT
GGAACGCCGCAAATTAGAACCAACCGTATCAAGCACCTGAAAGCCAAACCCCGTGGCAAG
AACTGTATCGCCAATTAGTCGGTCAACTGTCAACAGGCGGCTGCCTGGAACCAGCCACCT
TATATTTAAAAGTCGTGGAAACCCGTGGAGATCCTCGTCATAGCCATTAA

10

20

30

SEQ ID NO: 39 - AO#13-0573

MDRRELLKTSALLMAAAPLARAANVPEDHANVPRTNWSKNFHYSTSRVYAPTTPEEVPAIV
LENGHLKGLGSRHCFNNIADSQY AQISMREVKGIQIDEAAQTVTVGAGIAYGELAPVLDKAG
FALANLASLPHISVGGTIATATHGSGVGNKNLSSATRAIEIVKADGSILRLSRDTDGERFRMA
VVHLGALGVLTKVTLDIVPRFDMSQVVYRNLSFDQLEHNLDLSSGYSVSLFTDWQRNRVN
QVWIKDKATADAPQKPLPPMFYGATLQTAKLHPIDDH PADACTEQMGSVGPWYLRLPHFK
MEFTPSSGEELQTEYFVAR KDGYRAIRAVEKL RDKITPHLFITEIRTIAADDLPMSMAYQRDS
MAIHFTWKPEEPTVRKLLPEIEAALAPFGVRPHWGKIFEIPPSYLHKQYPALPRFRAMAQALD
PGGKFRNAYLDRNIFGA

40

SEQ ID NO: 40 - AO#22-8001

MDKRDFLKGSATTAVALMMGLNESKAFADDSVPRTNWSGNYHYSTNKNVLQPASVAETQD
AVRSVAGVRALGTRHSFNIGIADSQIAQISTLKLKDVSLDAKSSTVTVGAGIRYGD LAVQLDA
KGFALHNLASLPHISVGGACATATHGSGMGNGNLATAVKAVEFVAADGSVHTLSRDRDGD
RFAGSVVGLGALGVVTHLTLQVQPRFEMTQVVYRDLPFSELEHHLPEIMGAGYSVSLFTDW
QNGRAGEVWIKRRVDQGGASAPPARFFNATLATTKLHPILDHPAEACTDQLNTVGPWYERL
PHFKLNFTPSSGQELQTEFFVPFDRGYDAIRAVETLRDVITPHLYITELRAVAADDLWMSMAY
QRPSLAIHFTWKPETDAVLKLLPQIEAKLAPFGARPHWAKVFTMKSSHVAPLYPRLKDFLVL
AKSFDPKGKFQNAFLQDHVDIA

10

SEQ ID NO: 41 - AO#28-9635.1

MTASVTNWAGNISFVAKDVVRPGGVEALRKVVAGNDRVRVLGSGHSFNRIAEPEGADGVLV
SLDALPQVIDVDTERRTVRVGGGVKYAELARHVNESGLALPNMASLPHISVAGSVATGTHGS
GVNNGPLATPVREVELLTADGSLVTIGKDDARFPGAVTSLGALGVVVALTLDLEPAYGVEQ
YTFTELPLEGLDFEAVASAAYSVSLFTDWREAGFRQVWVKRRIDEPYAGFPWAAPATEKLHP
VPGMPAENCTDQFGAAGPWHERELPHFKAFTPSGDELQSEYLLPREHALAALDAVGNVRE
TVSTVLQICEVRTIAADTQWLSPAYGRDSVALHFTWTDDMDAVLPAVRAVESALDGFGARP
HWGKVFTTAPAALRERYPRLDDFRTLRLDELDPAGKFTNAFVRDVLEG

20

SEQ ID NO: 42 - AO#36-7049

MTLERNWAGTHTFAAPRIVNATSIDEVRALVAEAARTGTRVRALGTRHSFTDLADSDGTLIT
VLDIPADPVFDEAAGSVTIGAGTRYGIAAAWLAEHGLAFHNMGSLPHISVGGAIATGTHGSG
NDNGILSSAVSGLEYVDATGELVHVRRGDPGFDGLVVGLGAYGIVVRVTVDVQPAYRVRQD
VYRDVPWDAVLADFEGVTGGAYSVSIFTNWLGD TVEQIWWKTRLVAGDDELPPVPESWLG
VQRDSL TAGNLVETDPDNLTLQGGVPGDWWERLPHFRLESTPSNGDEIQTEYFIDRADGPAA
ITALRALGDRIAPLLLVTTELRTAAPDKLWLSGAYHREMLAVHFTWRNLPEEVRAVLPAIEEA
LAPFDARPHWGKLNLLTAERIAEVVPRADARDLFEELDPAGTFSNAHLERIGVRLPR

30

SEQ ID NO: 43 - AO#51-9823

MRDAAAANWAGNVRFGAARVVAPESVGELQEIVAGSRKARALGTGHSFSRIADTDGTLIAT
ARLPRRIQIDDGSVTVSGGIRYGD LARELAPNGWALRNLGSLPHISVAGACATGTHGSGDRN
GSLATSVAALELVTASGELVSVRRGDEDFDGHVIALGALGVTVAVTLDLVPGFQVRQLVYE
GLTRDTLLESVQEIFAASYSVSVFTGWDPESSQLWLKQRVDGPGDDGEPPAERFGARLATRP
LHPVPGIDPHTTTQQLGVPGPWHERELPHFRLDFTPSAGDELQTEYFVAREHAAAAIEALFAIG
AVVRPALQISEIRTVAADALWLSPAYRRDVMALHFTWISAEGTVMPAVA AVERALAPFDPV
PHWGKVFALPPAAVRAGYPRAAEFLALAAARRDPEAVFRNQYLDAYLPAA

40

SEQ ID NO: 44 - AO#57-0794

MTQRNWAGNVSYSSSRVAEPASVDDLTALVESEPRVRPLGSRHCFNDIADTPGVHVSLARLR
GEEPRLTAPGTLRTPAWLRYGDLVPVLREAGAALANLASLPHISVAGAVQTGTHGSGDRIGT
LATQVSALELVTGTGEVLRRLERGEPDFDGAVVGLGALGVLTHVELDVSPARDVAQHVEYEGV

RLDDVLADLGAVTGAGDSVSMFTHWQDPAVVSQVWVKSGGDVDDAAIRDAGGRPADGPR
HPIAGIDPTPCTPQLGEPGPWYDRLPHFRLEFTPSVGEELQSEYLVD RDDAVDAIRAVQDLAP
RIAPLLFVCEIRTMASDGLWLSPAQGRD TVGLHFTWRPDES AVRQLLPEIERALPASARPHW
GKVFTLPGHDVAARYPRWADFVALRRRLDPERRFANAYLERLGL

SEQ ID NO: 45 - AO#76-BAA19135

MTPAEKNWAGNITFGAKRLCVPRSVRELRETVAASGAVRPLGTRHSFNTVADTSGDHVSLA
GLPRVVDIDVPGRAVSLSAGLRFGEFAAELHARGLALANLGSLPHISVAGAVATGTHGSGVG
NRSLAGAVRALSLVTADGETRTLRR TDEDFA GAVVSLGALGVVTSLELDLVPAFEVRQWVY
EDLPEATLAARFDEVMSAAYS SVSFTDWRPGPVGQVWLKQRVGDEGARSVMPAEWLGAR
LADGPRHPVPGMPAGNCTAQQGVPGP WHERLPHFRMEFTPSNGDELQSEYFVARADAVAA
YEALARLRDRIAPVLQVSELRTVAADDLWLSPAHGRDSVAFHFTWVPDAAAVAPVAGAIEE
ALAPFGARPHWGKVFSTAPEVLRTLYPRYADFEELVGRHDPEGTFRNAFLDRYFRR

10

SEQ ID NO: 46 - AO#251-F3MC79

MGDKLNWAGNYRYSMELLEPKSLEE VKDLVVSRTSIRVLGSCHSFNGIADTGGSHLSLRK
MNRVIDLDRVQRTVTVEGGIRYGDLCRYLNDHGYALHNLASLPHISVAGAVATATHGSGDL
NASLASSVRAIELMKSDGEVTVL TRGTDPEFDGAVVGLGGLGVVTKLKL DLVPSFQVSQTVY
DRLPFSALDHGIDEILSSAYS VSLFTDWA EPIFNQVWVKRKVGINGEDETSPDFFGALPAPEKR
HMLVGQSVVNCSEQMGDPGPWYERLPHFRMEFTPSAGNELQSEYFVPRRHAVEAMRALGK
LRDRIAPLLFISEIRTIASDTFWMSPCYRQDSVGLHFTWKPDWERVRQLLPLIERELEPFAARP
HWAKLFTMESEMIQARYERLADFRQLLLRYDPIGKFRNTFLDHYIMH

20

SEQ ID NO: 47 - AO#13-0573-DNA

ATGGATCGTCGTGAACTGCTGAAAACCTCTGCACTGCTGATGGCAGCAGCACCGTTAGCA
CGTGCAGCAAATGTTCCGGAAGATCATGCAAATGTTCCGCGTACCAATTGGAGCAAAAA
CTTCCACTATAGCACCAGCCGCGTTTATGCACCGACTACCCCGGAAGAAGTTCCGGCAAT
TGTTCTGGAAAATGGTCATCTGAAAGGTCTGGGTTCTCGTCACTGCTTCAACAACATCGC
CGATAGCCAGTATGCGCAGATCAGCATGCGCGAAGTTAAAGGCATT CAGATCGATGAAG
CCGCACAAACCGTTACCGTGGGTGCAGGTATTGCGTATGGTGAATTAGCACCGGTGCTGG
ATAAAGCGGGTTTTGCACTGGCAAATTTAGCAAGTTTACCGCATATCAGCGTGGGTGGCA
CCATTGCAACCGCAACACATGGCTCTGGCGTTGGTAACAAAAACCTGTCTTCTGCAACCC
GTGCAATTGAAATCGTGAAAGCGGATGGCAGCATTCTGCGTCTGTGCGTGATACTGATG
GTGAACGTTTTTCGTATGGCGGTGGTTCATCTGGGTGCATTAGGTGTTTTAACC AAAGTTA
CCCTGGATATCGTGCCGCGCTTCGATATGTCTCAGGTGGTGTATCGCAACCTGTCTTTGA
TCAGCTGGAACACAACCTGGATACCATTCTGAGCTCTGGCTATAGCGTTAGCCTGTTAC
CGACTGGCAGCGTAATCGTGTTAATCAGGTGTGGATCAAAGATAAAGCGACCGCGGATG
CACCGCAAAAACCGTTACCTCCGATGTTTTATGGTGCGACCCTGCAAACCGCAAAACTGC
ATCCGATCGATGATCATCCGGCAGATGCATGTACCGAACA AATGGGTAGTGTTGGTCCGT
GGTATTTACGTCTGCCGCATTTCAAAATGGAGTTTACCCCGAGCAGCGGTGAAGAATTAC

30

40

AGACCGAATACTTCGTGGCGCGCAAAGATGGCTATCGCGCAATTCGTGCCGTGGAAAAA
CTGCGCGATAAAATTACCCCGCACCTGTTTATCACCGAAATCCGCACCATTGCAGCAGAT
GATCTGCCGATGAGCATGGCATATCAACGTGACAGTATGGCGATTCATTTTACCTGGAAA
CCGGAAGAACCGACCGTGCGTAAATTACTGCCGGAAATCGAAGCAGCACTGGCGCCGTT
TGGTGTTTCGTCCGCATTGGGGCAAAATTTTTGAAATTCCGCCGAGCTATCTGCATAAACA
GTATCCGGCACTGCCGCGTTTTTCGCGCAATGGCACAGGCATTAGATCCTGGTGGCAAATT
TCGTAATGCATATCTGGATCGTAACATCTTTGGCGCGTAG

SEQ ID NO: 48 - AO#22-8001-DNA

10

ATGGACAAACGCGATTTCTGAAAGGTAGCGCAACCACCGCAGTTGCACTGATGATGGG
TCTGAATGAAAGCAAAGCGTTTGCGGATGATAGCGTTCCGCGTACCAATTGGAGCGGCA
ACTACCATTATAGCACCAACAAAGTGCTGCAGCCGGCAAGTGTTGCAGAAACCAAGAT
GCAGTTCGTAGTGTTGCAGGTGTTTCGTGCATTAGGTACTCGTCATAGCTTTAACGGCATC
GCGGATAGCCAGATTGCCAGATTAGTACCCTGAAACTGAAAGATGTGAGCCTGGATGC
GAAAAGCTCGACCGTGACCGTTGGTGCAGGTATTCGTTATGGTGATCTGGCGGTTTCAGCT
GGATGCGAAAGGTTTTGCTCTGCATAATCTGGCAAGTCTGCCGCATATTTCTGTTGGTGG
TGCATGTGCAACTGCGACCCATGGTTCAGGTATGGGTAATGGTAATTTAGCAACCGCAGT
TAAAGCGGTGGAATTTGTTGCGGCGGATGGTAGCGTGCATACCCTGTCTCGTGATCGTGA
TGGTGATCGTTTTGCGGGCTCTGTTGTTGGTCTGGGTGCATTAGGTGTTGTTACCCATTTA
ACCCTGCAAGTTCAGCCACGTTTTCGAAATGACCCAGGTGGTGTACCGTGATCTGCCATTT
AGTGAACTGGAACATCATCTGCCGGAATTATGGGTGCCGGTTATAGCGTGTCCCTGTTT
ACCGATTGGCAGAATGGTCGTGCAGGTGAAGTGTGGATCAAACGTCGCGTGGATCAAGG
TGGTGCAAGTGCTCCTCCAGCTCGTTTTTTTTAATGCAACCTTAGCAACCACCAAACCTGCA
CCCGATCCTGGATCATCCTGCTGAAGCATGTACCGATCAGTTAAATACCGTAGGTCCGTG
GTATGAACGTTTACCGCACTTCAAACCTGAACTTCACCCCGAGCAGTGGCCAAGAATTACA
GACCGAGTTTTTCGTGCCGTTTCGATCGCGGCTATGACGCCATTTCGTGCCGTTGAACTTT
ACGTGATGTGATTACCCCGCACCTGTATATCACCGAACTGCGTGCAGTTGCAGCTGATGA
TTTATGGATGAGCATGGCATATCAACGTCCGAGTCTGGCAATCCATTTTACCTGGAAACC
GGAAACCGATGCAGTGCTGAAATTACTGCCGCAGATTGAAGCGAAACTGGCCCCGTTTG
GTGCTCGTCCGCATTGGGGCAAAAGTTTTTACCATGAAAAGCAGCCATGTGGCACCGCTGT
ATCCGCGCCTGAAAGATTTTCTGGTTCTGGCAAAATCCTTTGATCCGAAAGGCAAATTCC
AAAACGCGTTTCTGCAGGACCATGTGGACATCGCATAG

20

30

SEQ ID NO: 49 - AO#28-9635-DNA

40

ATGACCGCATCTGTGACCAATTGGGCGGGTAACATCAGCTTTGTGGCGAAAGATGTTGTT
CGTCCGGGTGGTGTGTAAGCACTGCGTAAAGTTGTTGCGGGTAATGATCGTGTTTCGTGTT
CTGGGTTCTGGTCATAGCTTTAACCGTATCGCTGAACCGGGTGCTGATGGTGTTCTGGTT
AGCCTGGATGCATTACCGCAAGTGATTGATGTTGATAACCGAACGTCGTACCGTGCGTGTT
GGTGGTGGTGTAAATACGCGGAACTGGCTCGTCATGTGAATGAATCTGGTCTGGCACTG

CCGAATATGGCATCTCTGCCGCATATTTCTGTTGCAGGTTCTGTTGCAACTGGTACCCATG
GTTCTGGTGTGAATAATGGCCCGTTAGCAACCCCGGTTCTGTAAGTTGAATTATTAACCG
CGGATGGCTCTCTGGTGACCATCGGTAAAGATGATGCGCGTTTTCCGGGTGCAGTTACTT
CTCTGGGTGCGCTGGGTGTTGTTGTTGCACTGACCTTAGATTTAGAACCGGCGTATGGTG
TTGAACAGTATACCTTTACCGAATTACCGCTGGAAGGTCTGGACTTCGAAGCAGTTGCGA
GTGCAGCATATTCTGTTAGCCTGTTACCGGATTGGCGTGAAGCTGGTTTTCCGCCAAGTTTG
GGTGAAACGCCGCATTGATGAACCGTACGCGGGCTTTCCGTGGGCAGCACCGGCAACTG
AAAAATTACATCCGGTTCCGGGTATGCCAGCAGAAAATTGTACTGATCAATTTGGTGCAG
CAGGTCCATGGCATGAACGTTTACCGCATTTTAAAGCGGAATTTACCCCGTCTAGCGGTG
ATGAATTACAGAGCGAATATCTGCTGCCGCGTGAACATGCACTGGCGGCACTGGATGCA
GTGGGCAACGTGCGTGAAACCGTTTCTACCGTGCTGCAGATTTGCGAAGTTCGTACCATT
GCAGCAGATACCCAGTGGTTAAGTCCGGCTTATGGTCTGTGATAGTGTTCATTACATTTT
ACTTGGACCGATGATATGGATGCAGTTTTACCTGCAGTTCGTGCCGTTGAAAGCGCGCTG
GATGGCTTTGGTGCTCGCCCGCATTTGGGGTAAAGTGTTCACCACCGCACCGGCAGCATT
CGTGAACGTTATCCGCGTCTGGATGATTTTCGTACCCTGCGTGATGAATTAGATCCGGCA
GGCAAATTTACTAATGCATTTGTTTCGTGATGTTCTGGAAGGTTAG

10

20

SEQ ID NO: 50 - AO#36-7049-DNA

ATGACCCTGGAACGTAATTGGGCAGGTACCCATACCTTTGCAGCACCGCGTATTGTTAAT
GCAACCAGCATCGATGAAGTTCGTGCGTTAGTGGCAGAAGCAGCACGTACCGGTACCCG
TGTTTCGTGCATTAGGTACTCGTCATTCTTTTACCGATCTGGCAGATAGCGATGGTACCCTG
ATTACCGTGCTGGATATTCCGGCAGATCCAGTTTTTCGATGAAGCAGCAGGTAGCGTTACC
ATTGGTGCAGGTACCCGTTATGGTATTGCAGCAGCATGGTTAGCAGAACATGGTCTGGCG
TTTCACAACATGGGTAGCCTGCCGCATATTAGCGTTGGTGGTGCAATTGCAACCGGTACC
CATGGTAGTGGTAATGATAACGGCATTCTGAGTAGCGCAGTTAGTGGTCTGGAATATGTT
GATGCGACCGGTGAACTGGTTCATGTGCGTCGTGGTGATCCTGGTTTTGATGGTCTGGTT
GTTGGTTTAGGCGCGTATGGTATTGTGGTTCGTGTGACGGTGGATGTTCAACCGGCATAT
CGTGTTCCGCCAGGATGTGTATCGTGATGTTCCGTGGGATGCAGTTCTGGCAGATTTTGAA
GGTGTTACAGGTGGTGCGTATAGCGTTAGCATCTTTACCAACTGGCTGGGTGATACGGTG
GAACAGATTTGGTGGAAAACCCGTCTGGTTGCAGGTGATGATGAACTGCCGGTGGTTCC
GGAAAGCTGGCTGGGTGTTCAACGTGATTCTTTAACCGCAGGTAATCTGGTTGAAACCGA
TCCGGATAATTTAACCCTGCAAGGTGGTGTTCGGGTGATTGGTGGGAACGTTTACCGCA
TTTTCGTCTGGAAAGTACCCCGTCTAATGGTGATGAAATCCAGACCGAATACTTCATCGA
TCGCGCGGATGGTCCGGCGGCAATTACCGCACTGCGTGCAATTAGGTGATCGTATTGCTCC
GTTACTGTTAGTTACCGAATTACGTACCGCAGCTCCAGATAAACTGTGGCTGAGTGGCGC
ATATCATCGCGAAATGTTAGCGGTCCATTTTACCTGGCGTAATTTACCGGAAGAAGTGCG
TGCAGTTTTACCAGCGATCGAAGAAGCCCTGGCGCCGTTTGATGCTCGTCCGCATTGGGG
TAAACTGAATCTGTTAACCGCAGAACGTATTGCAGAAGTTGTTCCGCGTCTGGCTGATGC

30

40

ACGTGATCTGTTTGAAGAACTGGACCCGGCTGGTACCTTTTCTAATGCTCATCTGGAACG
TATTGGTGTTCGTTTACCGCGTTAG

SEQ ID NO: 51 - AO#51-9823-DNA

ATGCGTGATGCAGCAGCAGCAAATTGGGCAGGTAATGTGCGTTTTGGTGCAGCACGTGTT
GTTGCACCGGAAAGTGTTGGTGAAGTGCAGGAAATTGTTGCAGGTAGCCGTAAAGCACG
TGCATTAGGTACCGGTCATAGCTTTAGCCGTATTGCAGATACCGATGGTACCCTGATTGC
TACCGCACGTTTACCACGTCGTATTCAGATCGATGATGGCAGCGTTACCGTTTCTGGTGG
TATCCGTTATGGCGATCTGGCCCGTGAATTAGCACCGAATGGTTGGGCATTACGTAATCT
GGGTTCTTTACCGCACATTTTCAGTTGCAGGTGCATGTGCAACCGGTACCCATGGTTCAGG
TGATCGTAATGGTAGTCTGGCAACCTCTGTTGCAGCGTTAGAATTAGTTACCGCGTCTGG
TGAATTAGTGAGCGTTTCGTCTGGCGATGAAGATTTTCGATGGCCATGTGATTGCGCTGGG
TGCACTGGGTGTTACTGTTGCAGTTACCTGGATTTAGTTCCGGGTTTTTCAGGTTTCGTCA
CTGGTGTATGAAGGTCTGACCCGTGATACCTTACTGGAAAGTGTGCAGGAAATCTTTGCT
GCGAGCTATAGTGTTAGCGTGTTTACCGGTTGGGACCCGGAAAGTTCTCAACTGTGGCTG
AAACAGCGCGTTGATGGTCCGGGCGATGATGGTGAACCACCGGCAGAACGTTTTGGTGC
ACGTTTAGCAACTCGTCCGTTACATCCAGTTCCGGGTATTGATCCGACTCATACTACTCA
ACAATTAGGTGTTCCAGGTCCGTGGCATGAACGTTTACCGCATTTTTCGTCTGGATTTTACC
CCTTCTGCAGGTGATGAACTGCAAACCGAATACTTCGTGGCCCGCGAACATGCAGCGGC
GGCGATTGAAGCACTGTTTGCATTGGTGCAGTTGTTTCGTCCGGCATTACAAATTAGCGA
AATTTCGTACCGTTGCAGCTGATGCATTATGGCTGTCTCCGGCATATCGTTCGTGATGTTATG
GCGTTACATTTTACCTGGATTAGCGCAGAAGGTACCGTTATGCCAGCAGTTGCAGCAGTG
GAACGTGCACTGGCGCCGTTTATCCGGTTCCTCATTGGGGTAAAGTTTTTTCGCTGCCG
CCAGCAGCAGTTTCGTGCTGGTTATCCTCGTGCAGCAGAATTTTATGATTAGCAGCTCGT
CGTGATCCGGAAGCAGTTTTTTCGTAATCAGTATTTAGATGCATATTTACCGGCAGCATAG

10

20

30

SEQ ID NO: 52 - AO#57-0794-DNA

ATGACCCAGCGTAATTGGGCGGGTAATGTGAGCTATAGTAGCAGCCGTGTTGCAGAACC
AGCAAGTGTGGATGATTTAACCGCACTGGTTGAAAGTGAACCGCGTGTTTCGTCCGTTAGG
TAGTCGTCATTGCTTCAACGATATCGCCGATACCCAGGTGTTTCATGTTTCTCTGGCACGT
CTGCGTGGTGAAGAACCGCGTTTAAACAGCACCGGGTACCTTACGTACTCCAGCTTGGTTA
CGTTATGGTGATTTAGTTCCGGTCTGCGTGAAGCAGGTGCAGCATTAGCAAATTTAGCA
TCTCTGCCGCATATTAGCGTTGCAGGTGCAGTTCAAACCGGTACCCATGGTTCAGGTGAT
CGTATTGGCACTCTGGCAACCCAAAGTTAGCGCCCTGGAATTAGTGACCGGCACCGGTGA
AGTTTTACGCTTAGAACGTGGTGAACCTGATTTTGATGGTGCGGTTGTTGGTTTAGGTGC
GTTAGGTGTTCTGACTCATGTGGAATTAGATGTTAGTCCGGCGCGTGATGTTGCACAGCA
CGTGTATGAAGGTGTTTCGTCTGGATGATGTTCTGGCGGATTTAGGCGCGGTTACTGGCGC
AGGTGATTCGGTGAGCATGTTTACCCATTGGCAAGATCCGGCAGTTGTTAGTCAGGTTTG
GGTAAAAGTGGCGGTGATGTGGATGATGCAGCAATTCGTGATGCAGGTGGTCGTCCGG

40

CAGATGGTCCGCGTCATCCAATTGCAGGTATTGATCCGACTCCATGTACTCCACAATTAG
GTGAACCAGGTCCGTGGTATGATCGTCTGCCGCATTTTCGTCTGGAATTTACCCCGAGTG
TTGGTGAAGAACTGCAAAGTGAATATCTGGTTGATCGCGATGATGCCGTTGATGCAATTC
GTGCGGTGCAGGATTTAGCCCCGCGTATTGCGCCGCTGCTGTTTGTTCGCAAATTCGTA
CCATGGCAAGTGATGGTTTATGGCTGAGCCCGGCACAAGGTCGTGATACCGTTGGTCTGC
ATTTTACCTGGCGTCCTGATGAATCTGCAGTTCGTCAATTATTACCGGAAATTGAACGTG
CTTTACCGGCAAGTGCTCGTCCGCATTGGGGTAAAGTGTTTACCCTGCCGGGGCCATGATG
TTGCAGCACGTTATCCGCGTTGGGCAGATTTTGTTCATTACGTCGTCGTTTAGATCCGGA
ACGTCGTTTCGCGAATGCATACCTGGAACGTTTAGGTCTGTAG

10

SEQ ID NO: 53 - AO#76-BAA19135-DNA

ATGACTCCGCGCGGAAAAAATTGGGCGGGCAACATCACCTTTGGTGCAAAACGTCTGTG
TGTTCCGCGTTCTGTTTCGTGAACTGCGTGAAACCGTTGCAGCATCTGGTGCAGTTCGTCC
GTTAGGTACTCGTCATAGCTTTAATACCGTTGCAGATACCAGTGGTGATCATGTTAGTCT
GGCAGGTTTACCGCGTGTGTGGACATCGATGTTCCGGGTCGTGCAGTTTCTCTGTCTGCT
GGTCTGCGTTTTTGGTGAATTTGCGGCTGAATTACATGCACGTGGTCTGGCGCTGGCAAAT
TTAGGTTCTCTGCCGCATATTAGCGTTGCAGGTGCAGTTGCAACCGGTACTCATGGTTCT
GGTGTGGTAATCGTTCTTTAGCAGGTGCAGTTCGTGCTTTATCTCTGGTAACCGCCGATG
GTGAAACCCGTACCTTACGTCGTACCGATGAAGATTTTGCAGGTGCAGTGGTTTCTCTGG
GTGCACTGGGTGTTGTTACTTCTCTGGAAGTGGATTTAGTTCCGGCGTTTCAAGTGCGTC
AGTGGGTGTACGAAGATCTGCCGGAAGCAACTTTAGCAGCTCGTTTTGATGAAGTTATGT
CAGCAGCGTATAGCGTGTCCGTGTTACCGATTGGCGTCCGGGTCCTGTTGGTCAAGTTT
GGCTGAAACAACGTGTTGGTGATGAAGGTGCTCGTAGTGTTATGCCAGCAGAATGGTTA
GGTGCACGTTTAGCAGATGGTCCGCGTCATCCAGTTCAGGTATGCCTGCAGGTAATTGT
ACAGCACAACAAGGTGTTCCAGGTCCGTGGCATGAACGTTTACCGCATTTTCGCATGGAA
TTTACCCCGTCTAACGGCGATGAACTGCAAAGCGAATATTTTGTGGCGCGTGCAGATGCA
GTTGCAGCGTATGAAGCATTAGCACGTCTGCGTGATCGTATTGCGCCGTTCTGCAAGTT
AGCGAATTACGTACCGTTGCAGCAGATGATCTGTGGCTGAGTCCGGCACATGGTTCGTGAT
AGTGTGCGTTTTCATTTTACCTGGGTTCCGGATGCAGCAGCAGTTGCACCGGTTGCAGGT
GCTATTGAAGAAGCATTAGCACCGTTTGGTGCACGTCCACATTGGGGTAAAGTTTTTAGC
ACCGCACCGGAAGTTTTACGTACCTTATATCCGCGTTATGCCGATTTTCAAGAACTGGTG
GGCCGCCATGATCCGGAAGGCACCTTTTCGTAATGCATTTTATAGATCGCTACTTTCGTGCT
AG

20

30

40

SEQ ID NO: 54 - AO#251-F3MC79-DNA

ATGGGCGATAAACTGAATTGGGCGGGCAACTATCGTTATCGCAGCATGGAAGTCTGGA
ACCGAAAAGCCTGGAAGAAGTGAAAGATCTGGTGGTTAGCCGTACCAGCATTCGTGTTT
TGGGTAGCTGTCATAGCTTTAACGGCATTGCGGATACCGGTGGTAGTCATCTGAGTCTGC
GCAAAATGAACCGCGTGATTGATCTGGATCGTGTTACGCGTACCGTTACCGTTGAAGGTG

GTATTCGTTACGGTGATCTGTGCCGCTATCTGAACGATCATGGTTATGCCCTGCATAATCT
GGCAAGCTTACCGCACATCAGCGTTGCAGGTGCAGTTGCAACCGCAACCCATGGTTCTGG
TGATCTGAATGCAAGTCTGGCAAGCTCTGTTTCGTGCAATTGAACTGATGAAAAGCGATGG
CGAAGTTACGGTTCTGACCCGTGGTACCGATCCGGAATTTGATGGTGCAGTTGTTGGTCT
GGGTGGTTTAGGTGTTGTGACCAAACCTGAAACTGGATCTGGTTCCGAGCTTTACAGGTGTC
GCAGACCGTGTATGATCGTCTGCCGTTTAGCGCACTGGATCATGGCATCGATGAAATTCT
GAGTAGTGCATATAGCGTTAGCCTGTTACCGATTGGGCGGAACCGATCTTTAATCAGGT
GTGGGTGAAACGCAAAGTGGGCATTAACGGCGAAGATGAAACCAGTCCGGATTTTTTTTG 10
GCGCATTACCGGCACCGGAAAAACGCCACATGGTTCTGGGTGAGAGCGTGGTGAATTGC
AGCGAACAAATGGGTGATCCTGGTCCGTGGTATGAACGTTTACCGCATTTTTCGCATGGAA
TTTACCCCGAGTGCAGGCAATGAATTACAGAGCGAATATTTTGTGCCGCGTCGTCATGCG
GTTGAAGCAATGCGTGCGTTAGGTAAACTGCGTGATCGTATTGCACCACTGCTGTTTCATC
AGCGAAATCCGCACCATTGCGAGCGATACCTTCTGGATGAGCCCGTGTTATCGTCAGGAT
TCTGTTGGTCTGCATTTTACCTGGAAACCGGATTGGGAACGTGTTTCGTCAGTTATTACCGC
TGATTGAACGTGAACTGGAACCGTTTGCGGCACGTCCGCATTGGGCGAAACTGTTTACCA
TGGAAGCGAAATGATTCAGGCGCGCTATGAACGTCTGGCGGATTTTCGTCAGCTGCTGC 20
TGCGTTATGATCCGATTGGCAAATTCCGTAACACCTTTCTGGATCACTACATCATGCACT
AA

SEQ ID NO: 55 pSGI-431 Q72LK2-タンパク質

MEATLPVLDAKTAALKRRSIRRYRKDPVPEGLLREILEAALRAPSAWNLQPWRIVVVRDPAT
KRALREAAFGQAHVEEAPVVLVLYADLEDALAHLDEVIHPGVQGERREAQKQAIQRAFAA
MGQEARKAWASGQSYILLGYLLLLLEAYGLGSPMLGFDPERVKAILGLPSHAAIPALVALG
YPAEEGYPSHRLPLERVVLWR

SEQ ID NO: 56 pSGI-431 Q72LK2-DNA

ATGGAAGCAACCTTACCGGTGTTAGACGCGAAAAACCGCAGCACTGAAACGTCGTAGCAT
TCGCCGTTATCGCAAAGATCCAGTTCCGGAAGGTTTACTGCGCGAAATTCTGGAAGCAGC
ATTACGTGCACCGTCTGCATGGAATTTACAACCGTGGCGTATTGTGGTGGTTCGTGATCC
GGCAACTAAACGTGCATTACGTGAAGCAGCATTTGGTCAAGCCCATGTGGAAGAAGCAC
CGGTTGTTCTGGTTCTGTACGCAGATCTGGAAGATGCACTGGCACATCTGGATGAAGTGA
TTCATCCGGGCGTTCAAGGTGAACGTCGTGAAGCGCAGAAACAAGCAATTCAGCGTGCA
TTTGCAGCAATGGGTGAGGAAGCTCGTAAAGCTTGGGCAAGCGGTCAAAGTTATATTCTG 40
CTGGGTTATCTGCTGCTGCTGCTGGAAGCATATGGTCTGGGTTCTGTTCCGATGCTGGGTT
TTGATCCTGAACGTGTTAAAGCGATTCTGGGCCTGCCGTCACATGCAGCGATTCCGGCAT
TAGTTGCACTGGGTTATCCGGCTGAAGAAGGTTATCCGAGTCATCGTTTACCGCTGGAAC
GTGTTGTTTTATGGCGTTGA

SEQ ID NO: 57: pSGI-374 #9041 タンパク質

MKNPFSLQGRKALVTGANTGLGQAIIVGLAAAGAEVVCAARRAPDETLEMIASDGGKASA
LSIDFADPLAAKDSFAGAGFDILVNNAGIIRRADSVFSELDWDEVMDVNLKALFFTTQAFK
ELLAKGRSGKVVNIASSLFSQGGIRVPSYTAACHGVAGLTKLLANEWAAGKINVNAIAPGYI
ETNNTALRADAARNKAILERIPAGRWGRSEDIAGAAVFLSSAAADYVHGAILNVDGGWLA
R

SEQ ID NO: 58 pSGI-375 #8939 タンパク質

MIAGVGGEARELALDLSDPMAAKDVFAEGAYDLLINNAGIIRRADAVDFSEDDWDAVMDV
NLKAVFFTSQAFARALMSRNASGKIVNIASLLSFQGGIRVASYTAACHGVAGITRLLANEWA
SRGINVNAIAPGYIATNNTALRADEERNAAILARIPAGRWGRAEDIAGTAVYLCSPAADYV
HGAILNVDGGWLAR

10

SEQ ID NO: 59 pSGI-376 P37769-タンパク質

MILSAFSLEGKVAVVTGCDTGLGQGMALGLAQAGCDIVGINIVEPTETIEQVTALGRRFLSLT
ADLRKIDGIPALLDRAVAEFGHIDILVNNAGLIRREDALEFSEKDWDVNMNLNIKSVFFMSQA
AAKHFAAQNGGKIINIASMLSFQGGIRVPSYTASKSGVMGVTRLMANEWAKHNINVNAIAP
GYMATNNTQQLRADEQRS AEILDRIPAGRWGLPSDLMGPVFLASSASDYVNGYTIADVGG
WLAR

20

SEQ ID NO: 60 pSGI-395 #5112 タンパク質

MPGMTTPFDLHGKTAIVTGANTGIGQAIALSQAAGADIAAVGRTPAQDTVDQVRALGRR
DIISADLSTIEPVQRVLDETLEKLGALDILVNNAGIIRRADSVDFTEEDWDAVIDTNLKTTF
FLCQAAGRHLMAQAGAKIINIASLLSFQGGIRVPSYTASKSGVAGLTKLLANEWAAGKINVNAIA
PGYIATNNTAALQADETRNRQIERIPAGRWGDPADIGGA VFLASSAADYIHGHTLAVDGG
WLAR

SEQ ID NO: 61 pSGI-396 #7103-タンパク質

MNPFSLEGKTALVTGANTGIGQAIAMALGRAGADVICAGRSSCAETVALIAGSKGKARELVL
DFADPMAARDVFAAEPVDILVNNAGIIRRADAVDFTEADWDEVMDVNLKAVFFTCQAFGK
AVLGRGGNGKIVNIASLLSFQGGIRVPSYTASKHGVAGITKLLANEWAAGKINVNAIAPGYIE
TNNTALRADPVRNKAILERIPAGRWGQASDIGEAAVFLASPAANYIHGAVLNVDGGWLAR

30

SEQ ID NO: 62 pSGI-374 #9041 DNA

ATGAAGAATCCCTTTTCGCTTCAGGGGCGTAAGGCGCTCGTCACCGGCGCGAATACGGGGCTTGGC
CAGGCGATTGCGGTTGGGCTCGCCGCGGCCGGTGCGGAGGTGGTCTGCGCCGCCCGCGCGCC
GGATGAAACGCTGGAGATGATCGCCAGCGACGGCGGCAAGGCCAGCGCATTGTCCATCGATTTG
CCGATCCGCTGGCGGCGAAGGACAGTTTTGCCGCGCCGGTTTCGATATTCTCGTCAACAATGCCG
GTATCATCCCGCTGCCGATTCCGTCGAGTTCTCCGAACTCGACTGGGACGAGGTGATGGACGTCA
ATCTCAAGGCGCTGTTTTTACCACCCAGGCTTTTGC GAAAGAGCTGCTGGCGAAAGGCCGGTCCG
GCAAGGTGGTCAATATCGCTTCGCTCCTTTCTTTT CAGGGCGGTATTCGCGTGCCGTCCTATACGGC
GGCGAAACATGGTGTGCGCGGCCCTAACCAAACCTCTGGCGAATGAATGGGCCGCCAAGGGCATCA

40

ATGTGAATGCCATTGCGCCCGGTTATATCGAAACCAACAATACCGAGGCGCTACGCGCCGATGCG
GCTCGTAACAAGGCCATTCTCGAGCGCATCCCGGCCGGCCGCTGGGGGCGCTCGGAAGACATCGC
CGGGGCGGCGGTTTTCTGTCTATCTGCGGCGGCGGACTATGTGCATGGCGCCATTCTCAACGTCGA
TGGCGGCTGGCTGGCGCGCTGA

SEQ ID NO: 63 pSGI-375 #8939 DNA

ATGATCGCCGCGTGGGGGAGAAGCAAGGGAGCTGGCGCTCGATCTGTCCGATCCCATGGCGGC
AAAAGATGTTTTTGTGAAGGCGCTTACGACCTCCTCATCAACAATGCCGGCATCATCCGCCGTGC
CGATGCAGTCGATTTCTCCGAGGATGACTGGGACGCGGTGATGGACGTGAACCTGAAAGCCGTCT
TCTTCACCTCGCAAGCCTTTGCGCGGGCTCTCATGTCCAGAAACGCAAGCGGAAAGATCGTTAACA
TTGCATCCCTTCTGTCTGTTTCAAGGCGGCATTTCGCGTTGCCTCCTACACGGCCGCCAAGCACGGTGT
GGCAGGCATCACCAGACTGTTGGCAAACGAATGGGCGTCCCGCGGCATCAACGTCAATGCGATAG
CGCCCGGTTACATTGCCACGAACAACACGGAAGCGCTTCGAGCCGACGAGGAGCGCAACGCGGCG
ATCCTCGCACGCATTCCGGCTGGCCGCTGGGGGCGGGCGGAGGATATTGCGGGTACTGCTGTCTAT
CTTTGTTGCGCCGCGAGCCGATTATGTTTCATGGCGCCATTCTAAACGTCGATGGCGGTTGGCTCGCG
CGCTGA

10

SEQ ID NO: 64 pSSI-376 P37769-DNA

ATGATTTTAAGTGCATTTTCTCTCGAAGGTAAAGTTGCGGTCGTCCTGGTTGTGATACTG
GACTGGGTCAGGGGATGGCGTTGGGGCTGGCGCAAGCGGGCTGTGACATTGTTGGCATT
AACATCGTTGAACCGACTGAAACCATCGAGCAGGTCACAGCGCTGGGGCGTCGTTTTTTA
AGCCTGACCGCCGATCTGCGAAAGATTGATGGTATTCCAGCACTGCTGGATCGCGCGGTA
GCGGAGTTTGGTCATATTGATATCCTGGTGAATAACGCCGGATTGATTCGCCGCGAAGAT
GCTCTCGAGTTCAGCGAAAAGGACTGGGACGATGTCATGAACCTGAATATCAAGAGCGT
ATTCTTCATGTCTCAGGCAGCGGCGAAACACTTTATCGCGCAAGGCAATGGCGGCAAGA
TTATCAATATCGCGTCAATGCTCTCCTTCAGGGCGGGATCCGTGTGCCTTCTTATACCGC
ATCAAAAAGCGGCGTGATGGGTGTGACGCGATTGATGGCGAACGAATGGGCTAAACACA
ACATTAATGTTAATGCGATAGCCCCGGGTACATGGCGACCAACAATACTCAACAACACTAC
GGGCAGATGAACAACGTAGCGCGGAAATTCTCGACCGCATTCCAGCTGGTCGTTGGGGA
CTGCCGAGTGACCTGATGGGGCCGATAGTGTTCCTTGCCCTCCAGCGCTTCAGATTATGTG
AATGGTTATAACCATTGCCGTGGATGGCGGTTGGCTGGCGCGTTAA

20

30

SEQ ID NO: 65 pSGI-395 #5112 DNA

ATGCCCCGGCATGACCACTCCTTTTCGATCTTCATGGCAAGACCGCGATCGTCACCGGCGCCAATACC
GGCATCGGCCAGGCCATTGCCCTGTCTGCTCGCGCAGGCCGGCGCGGATATCGCCGCCGTGCGCCG
CACGCCCCGACAGGACACGGTCGATCAGGTCCGCGCGCTCGGCCGCCGGGCGGACATTATCTCGG
CCGACCTTTTCGACCATCGAACCAGTCCAGCGCGTCCTCGACGAAACGCTGGAAAAGCTTGGTGCCT
TGGACATACTGGTCAACAATGCCGGCATCATCCGCCGCGCCGACAGCGTCGATTTACCGAGGAG
GATTGGGACGCGGTGATCGACACCAATCTCAAGACCACCTTCTTCCTCTGTACGGCCGCCGGTTCG
CACATGCTTGCCCAAGGCGCTGGCAAAGATCATCAACATCGCCTCGCTTCTTTCCTTCCAGGGCGGC
ATTGCGGTGCCGAGCTACACCGCGTCCAAAAGCGGCGTCGCGGGCCTGACCAAGCTGCTCGCCAA
CGAATGGGCGGCCAAGGGCGTCAATGTGAACGCCATCGCGCCGGGCTATATCGCCACCAACAACA

40

CCGCCGCGCTCCAGGCCGACGAAACCCGCAACCGCCAGATCCAGGAGCGCATCCCGGCTGGCCGC
TGGGGCGACCCCGCCGACATTGGCGGCGCGGCCGTGTTCTTGCGCTCCAGCGCCGCCGATTATATC
CATGGCCACACGCTCGCCGTCGACGGCGGCTGGCTCGCGCGCTGA

SEQ ID NO: 66 pSGI-396 #7103-DNA

ATGAACCCCTTCTCGCTTGAGGGCAAGACCGCCCTTGTGACCGGTGCCAATACGGGCATCGGTCAG
GCCATCGCCATGGCGCTTGCGCGCGCCGGGGCGGACGTCATCTGCGCGGGACGCTCGTCCTGTGCG
GAGACCGTTGCCCTCATCGCTGGCAGCAAGGGCAAGGCGCGCGAACTGGTGCTCGACTTCGCCGA
CCCATGGCCGCCCGTGACGTGTTCCCGCCGAACCGGTGGACATCCTCGTCAACAACGCGGGCA
TCATCCGGCGCGCCGATGCAGTGGATTTCACCGAGGCCGACTGGGATGAGGTGATGGACGTGAAC
CTGAAGGCCGTGTTCTTACCTGCCAGGCCTTCGGCAAGGCCGTTCTTGGCCGTGGAGGAAACGGC
AAGATCGTCAACATTGCCTCGCTCCTGTCATTCCAGGGTGGTATCCGGGTGCCGTCCTACACGGCC
TCGAAGCATGGTGTTCAGGCATCACCAAGCTTCTGGCCAACGAATGGGCGGCGAAGGGCATCAA
TGTGAATGCCATCGCCCCCGGTTACATCGAAACGAACAATACCGAAGCACTGCGGGCGGACCCGG
TGCGCAACAAGGCCATCCTTGAGCGTATCCCTGCCGGCCGCTGGGGCCAGGCCTCGGACATCGGC
GAAGCCGCCGTGTTCTTGCCTCTCCGGCTGCCAATTACATCCATGGTGCAGTGTGAATGTTGAC
GGAGGCTGGCTTGCCCGCTGA

10

SEQ ID NO: 67 pSGI-353 P0AES2

MSSQFTTPVVTQMVIPVAGHDSMLMNLGAHAPFFTRNIVIIKDNSGHTGVGEIPGGEKIRK
TLEDAIPLVVGKTLGEYKNVLTIVRNTFADRDAGGRGLQTFDLRTTIHVVTGIEAAMLDLLG
QHLGVNVASLLGDGQQRSEVEMLGYLFFVGNRKATPLPYQSQPDDSCDWYRLRHEEAMTP
DAVVRLAEAAAYEKYGFNDFKLKGGVLAGEEEEAESIVALAQRFPQARITLDPNGAWSLNEAIK
IGKYLKGSLAYAEDPCGAEEQGFSGREVMAEFRRATGLPTATNMIATDWRQMGHTLSLQSV
IPLADPHFWTMQGSVRVAQMCHEFGLTWGSHSNHFDISLAMFTHVAAAAPGKITAIDTHW
IWQEGNQRLTKEPFEIKGGLVQVPEKPGLGVEIDMDQVMKAHELQYKHGLGARD DAMGMQ
YLIPGWTFDNKRPCMVR

20

30

SEQ ID NO: 68 pSGI-244 #8114

MTTAMSGTPRITELTVVPVAGQDSMLMNLGAHGPFTRNIIKDSAGHVGVGEVPGGEAI
RQTLDDARALLVGEPIGQYNALLGKVRRAFADRDAGGRGLQTFDLRIAHAVTALESALLDL
LGQHLEVPVAALLGEGQQRDEVEMLGYLFFIGDRNRDLDGYRDESNDDAWFRVRNEEAM
TPERIVRQAEAAAYERYGFKDFKLKGGVLRGEEVEAIRALAQRFPDARVTLDPNGAWSLDEA
SGLCRDLHGVLAYAEDPCGAENGYSGREVMAEFRRATGLPTATNMIATDWRQM SHAVCLH
SVDIPLADPHFWTMAGSVRVAQMCADFGLTWGS SHSNHFDISLAMFTHVAAAAPGRVTAID
THWIWQDGQHLTREPLKIVSGKVAVPQKPGLGVELDWDALEQAHAHYQEKGLGARD DAIA
MQYLIPNWTFFNNKKPCMVR

40

SEQ ID NO: 69 pSGI-353 P0AES2-DNA

ATGAGTTCTCAATTTACGACGCCTGTTGTTACTGAAATGCAGGTTATCCCGGTGGCGGGTCATGAC
AGTATGCTGATGAATCTGAGTGGTGACACGCACCGTTCTTTACGCGTAATATTGTGATTATCAA
GATAATTCTGGTCACACTGGCGTAGGGGAAATTCCCGGCGGCGAGAAAATCCGTAAAACGCTGGA

AGATGCGATTCCGCTGGTGGTAGGTAAAACGCTGGGTGAATACAAAACGTTCTGACGCTGGTGC
GTAATACTTTTGCCGATCGTGATGCTGGTGGGCGCGGTTTGCAGACATTTGACCTACGTACCACTA
TTCATGTAGTTACCGGGATAGAAGCGGCAATGCTGGATCTGCTGGGGCAGCATCTGGGGGTAAAC
GTGGCATCGCTGCTGGGCGATGGTCAACAGCGTAGCGAAGTCGAAATGCTCGGTTATCTGTTCTTC
GTCGGTAATCGCAAAGCCACGCCGCTGCCGTATCAAAGCCAGCCGGATGACTCATGCGACTGGTA
TCGCCTGCGTCATGAAGAAGCGATGACGCCGGATGCGGTGGTGCCTGGCGGAAGCGGCATATG
AAAAATATGGCTTCAACGATTTCAAACCTGAAGGGCGGTGTACTGGCCGGGGAAGAAGAGGCCGAG
TCTATTGTGGCACTGGCGCAACGCTTCCCGCAGGCGCGTATTACGCTCGATCCTAACGGTGCCTGG
TCGCTGAACGAAGCGATTAAAATCGGTAAATACCTGAAAGGTTGCTGGCTTATGCAGAAGATCC
GTGTGGTGGGAGCAAGGTTTCTCCGGGCGTGAAGTGATGGCAGAGTTCCGTCGCGCGACAGGTC
TACCGACTGCAACCAATATGATCGCCACCGACTGGCGGCAAATGGGCCATACGCTCTCCCTGCAAT
CCGTTGATATCCCGCTGGCGGATCCGCATTTCTGGACAATGCAAGGTTTCGGTACGTGTGGCGCAA
TGTGCCATGAATTTGGCCTGACCTGGGGTTCACACTCTAACAACCACTTCGATATTTCCCTGGCGAT
GTTTACCCATGTTGCCGCCGCTGCACCGGGTAAAATTACTGCTATTGATACGCACTGGATTTGGCA
GGAAGGCAATCAGCGCCTGACCAAAGAACCGTTTGAGATCAAAGGCGGGCTGGTACAGGTGCCAG
AAAAACCGGGGCTGGGTGTAGAAATCGATATGGATCAAGTGATGAAAGCCCATGAGCTGTATCAG
AAACACGGGCTTGGCGCGCGTGACGATGCGATGGGAATGCAGTATCTGATTCCTGGCTGGACGTT
CGATAACAAGCGCCCGTGTCATGGTGCCTTAA

10

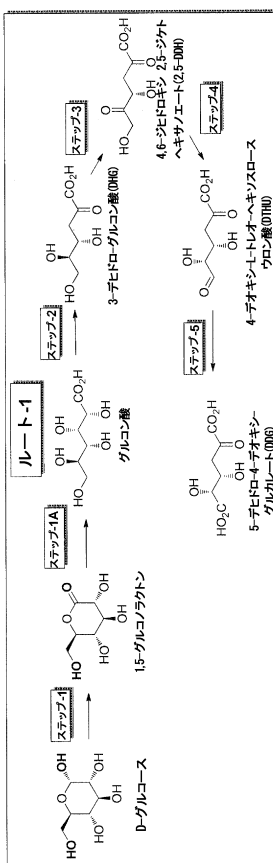
20

SEQ ID NO: 70 pSGI-244 #8114

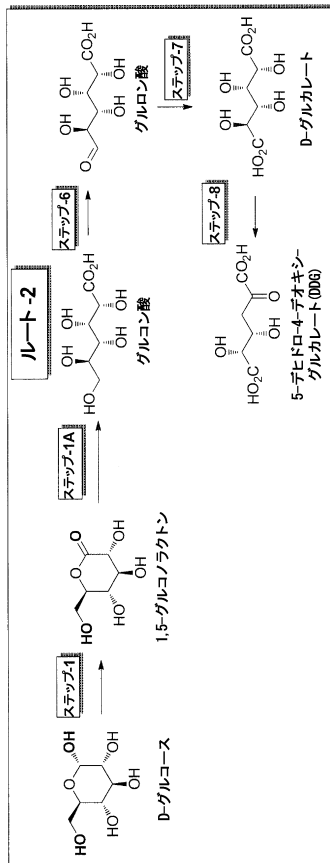
ATGACCACCGCCATGTCGGGCACGCCCCGCATCACCGAACTCACCGTCGTGCCCCGTCGCCGGGCA
GGACAGCATGCTGATGAACCTCAGCGGCGCCCATGGGCCCTGGTTCACCCGCAACATCCTCATCCT
CAAGGACAGCGCCGGCCACGTGCGCGTCGGCGAAGTGCCGGGCGGCGAAGCCATCCGCCAGACCC
TCGACGATGCCCCGTGCCCTGCTGGTTCGGCGAACCAGTCGGCCAGTACAACGCGCTGCTCGGCAAG
GTGCGCCGCGCCTTCGCCGACCGTGACGCCGGCGGCCGCGGCTGCAGACCTTCGACCTGCGCATC
GCCATTACGCCGTCACCGCGCTGGAGTCGGCGCTGCTCGACCTGCTCGGCCAGCACCTCGAGGTG
CCGTCGCCGCCTTGCTCGGCGAAGGCCAGCAGCGTGACGAAGTGGAATGCTCGGCTACCTGTT
CTTCATCGGCGATCGCAACAGGACCGACCTCGGCTACCGCGACGAATCCAACCTCCGACGACGCCT
GGTTTCGCGTGCGCAACGAGGAGGCCATGACGCCGAGCGCATCGTCCGCCAGGCCGAGGCGGCC
TACGAGCGCTACGGCTTCAAGGACTTCAAGCTCAAGGGCGGCGTACTGCGCGGCGAAGAGGAAGT
CGAGGCGATCCGCGCCCTGGCCCAGCGCTTCCCCGACGCCCGCGTGACTCTGGACCCCAACGGCG
CCTGGTCGCTGGACGAAGCCAGCGGCCTGTGTGCGGACCTGCACGGCGTGCTGGCCTATGCCGAA
GACCCCTGCGGTGCCGAGAACGGCTATTCCGGCCGCGAGGTGATGGCCGAGTTCCGCCGCGCCAC
CGGTCTGCCCACCGCGACCAACATGATCGCCACCGACTGGCGACAGATGAGTCACGCGGTGTGCC
TGCACTCGGTGGACATCCCGCTGGCCGACCCGCACTTCTGGACCATGGCCGGCTCTGTGCGCGTGG
CGCAGATGTGCGCCGACTTCGGCCTGACCTGGGGTTCGCACTCGAACAACCACTTCGACATCTCCC
TGGCGATGTTACCCACGTGGCGGCCGCCGCGCCGGGTGCGGTCACCGCCATCGACACCCACTGG
ATCTGGCAGGACGGCCAGCACCTGACCCGCGAGCCGCTGAAGATCGTCAGCGGCAAGGTTGCGGT
GCCGCGAAGCCGGGGCTGGGCGTCGAGCTGGACTGGGATGCCCTGGAGCAGGCGCATGCCACT
ACCAAGAGAAAAGGCTGGGTGCCCGCGATGACGCCATCGCCATGCAGTACCTGATCCCCAACTGG
ACCTTCAACAACAAGAAGCCGTGCATGGTGCCTGA

30

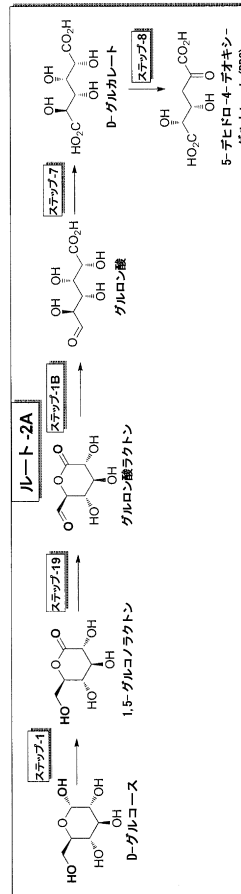
40



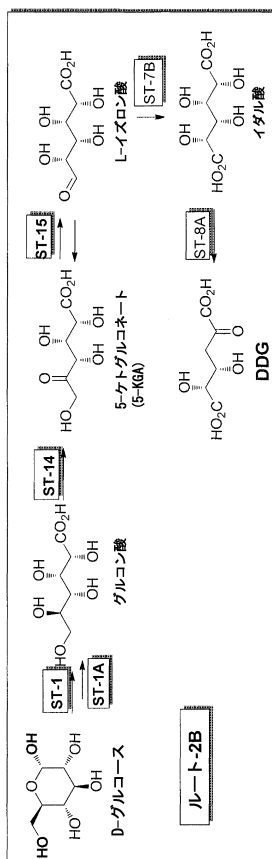
【図 2 B】



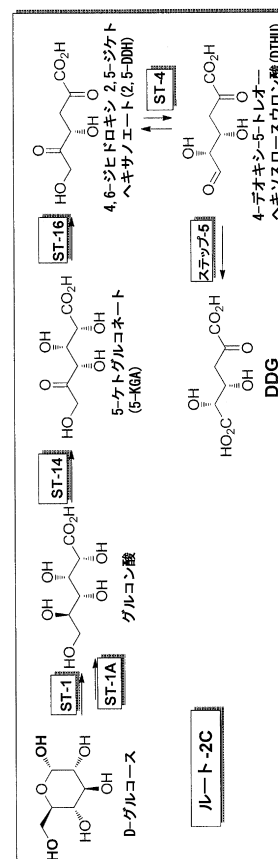
【図 2 C】



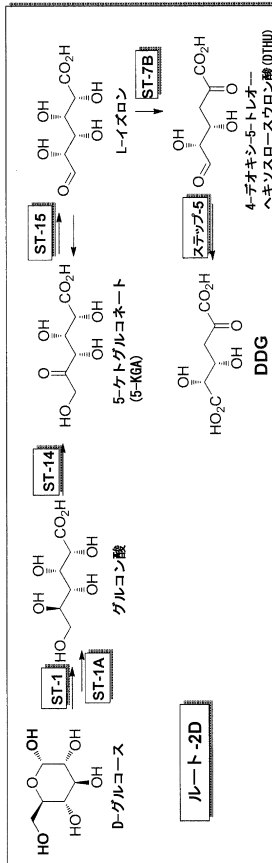
【図 2 D】



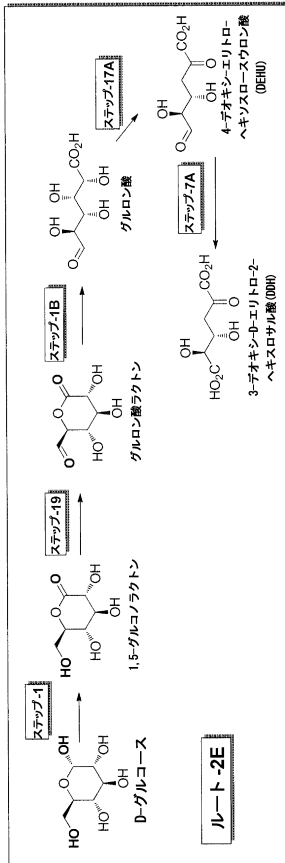
【図 2 E】



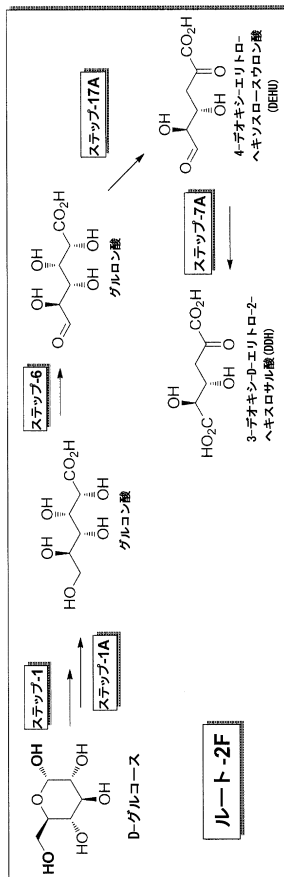
【 図 2 F 】



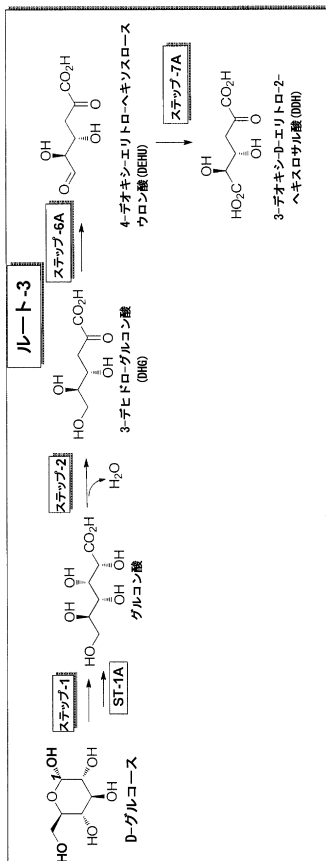
【 図 2 G 】



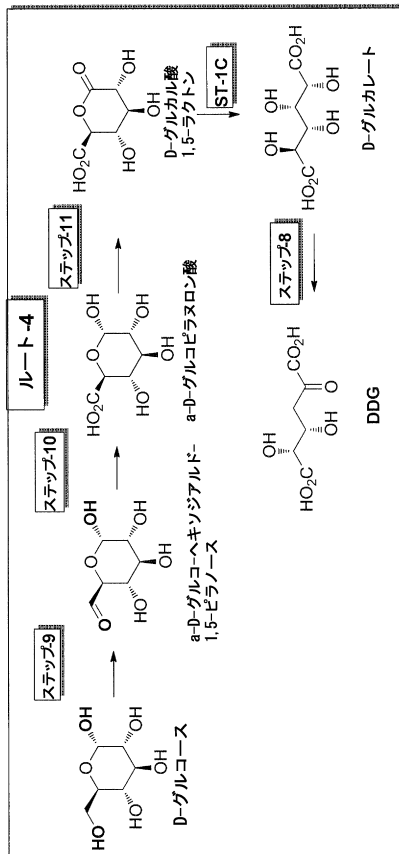
【 図 2 H 】



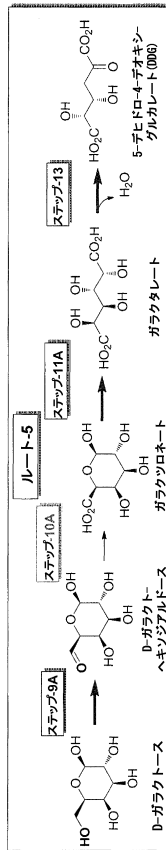
【 図 3 A 】



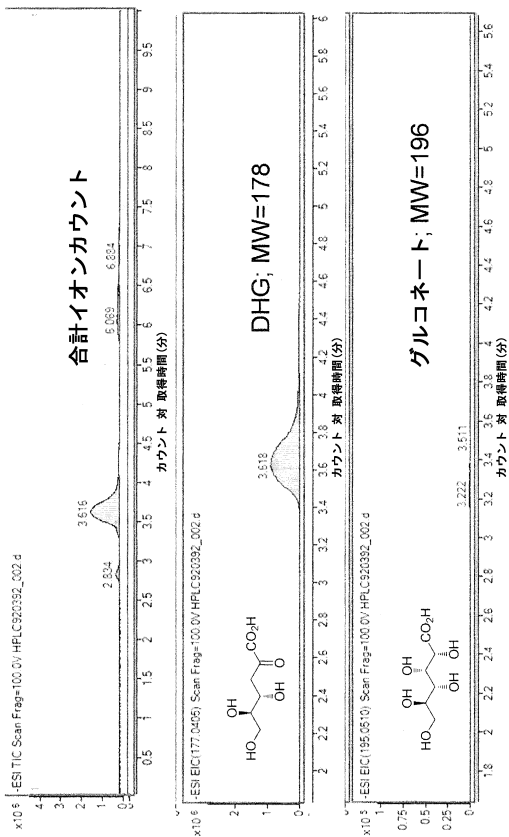
【図 3 B】



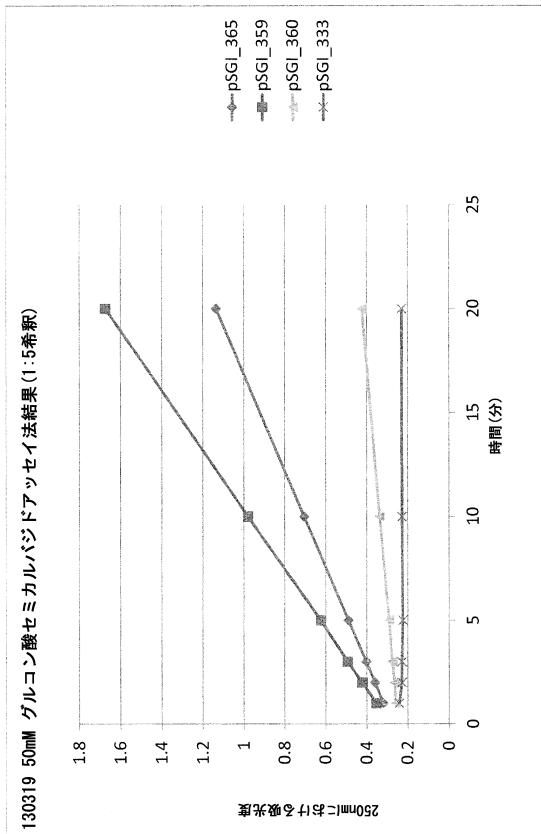
【図 3 C】



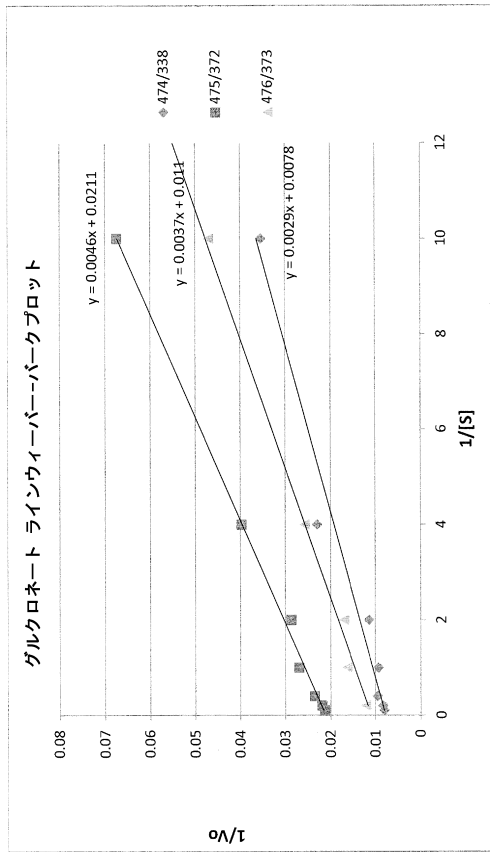
【図 4】



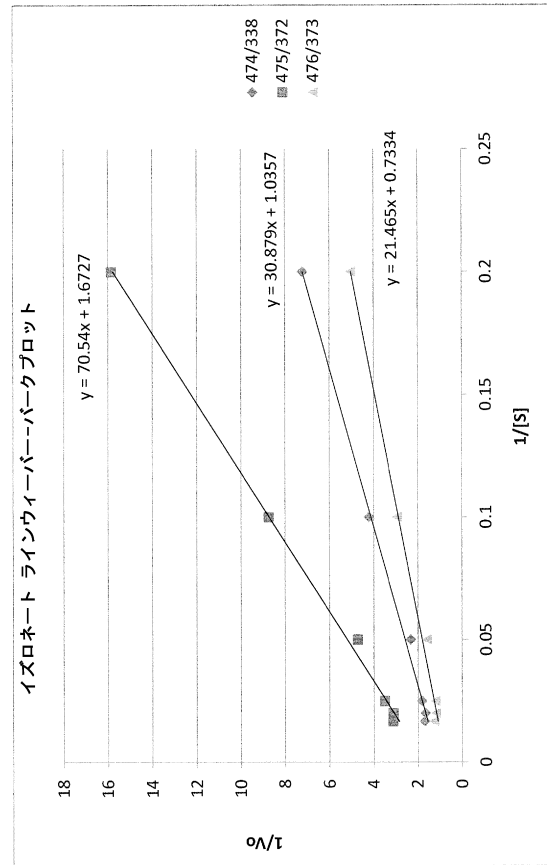
【図 5】



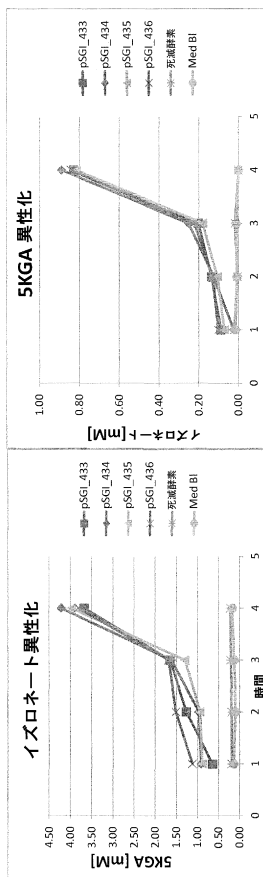
【図 6 A】



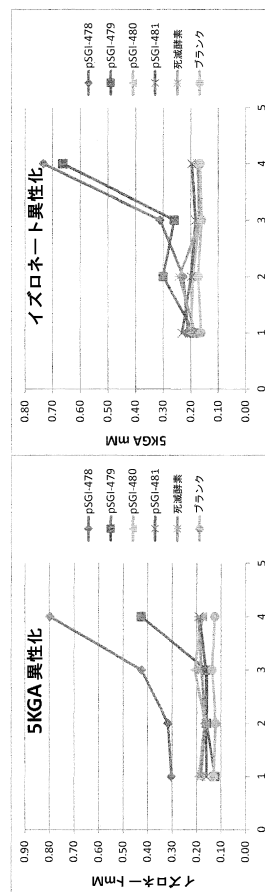
【図 6 B】



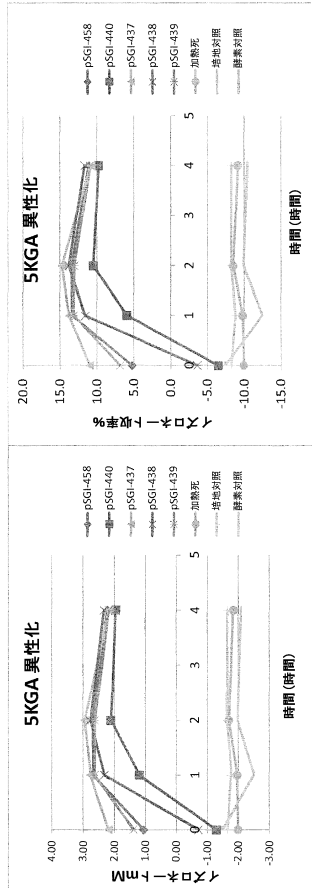
【図 7 A】



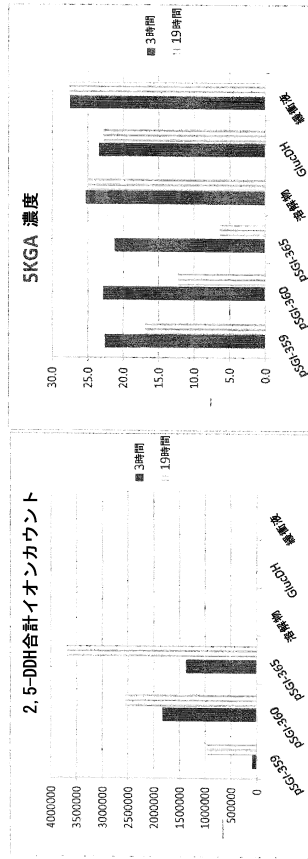
【図 7 B】



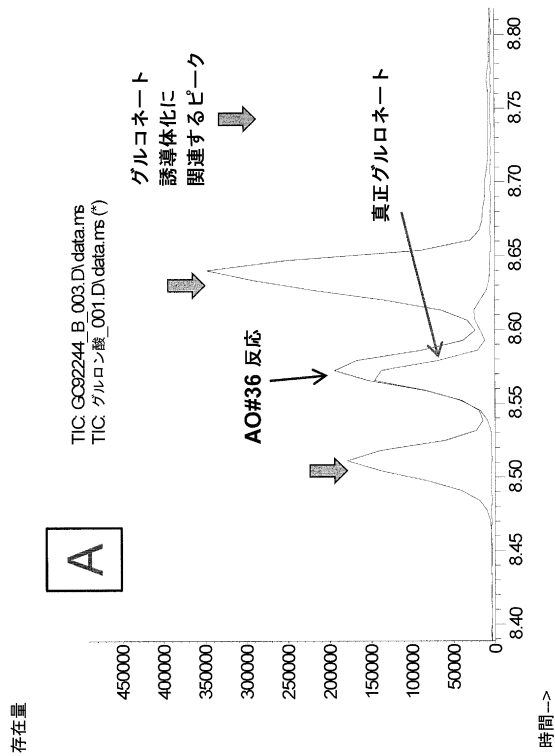
【図 8】



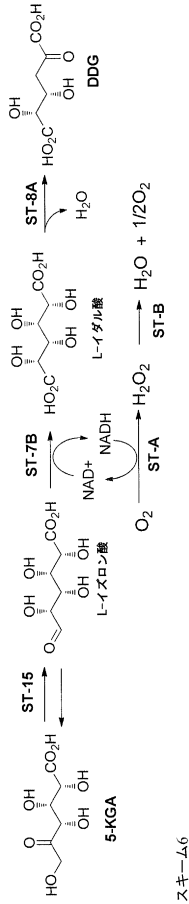
【図 9】



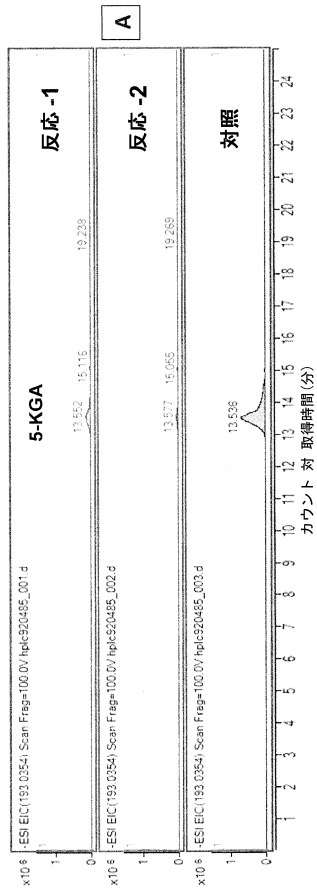
【図 10】



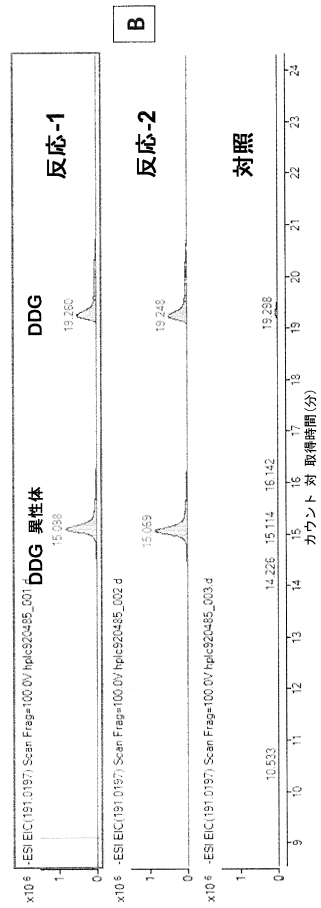
【図 11】



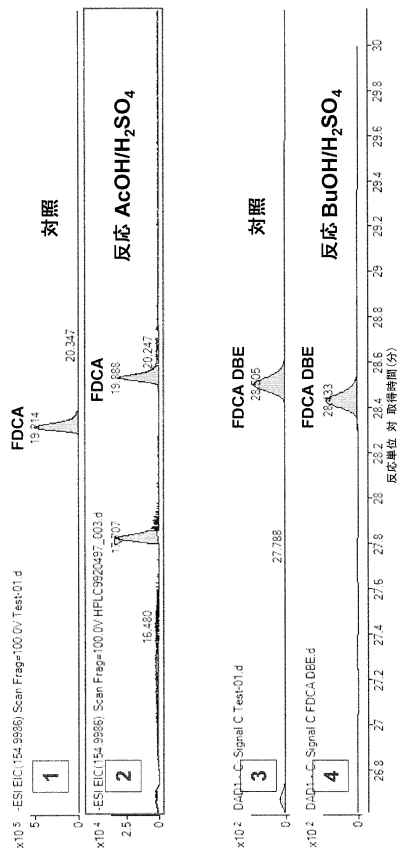
【図 1 2 A】



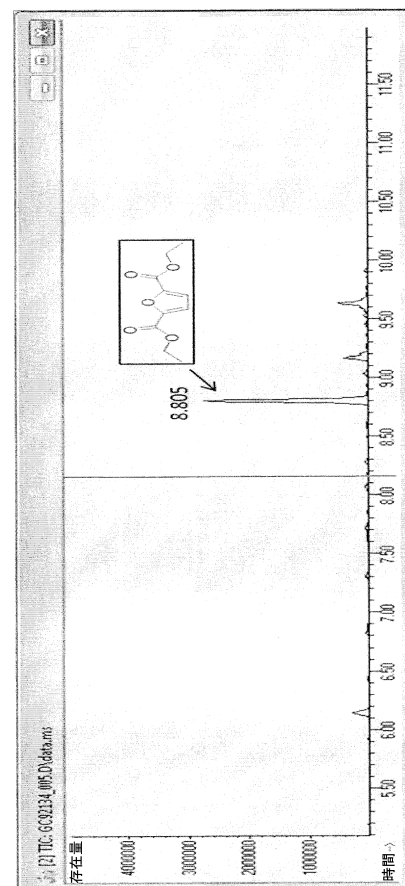
【図 1 2 B】



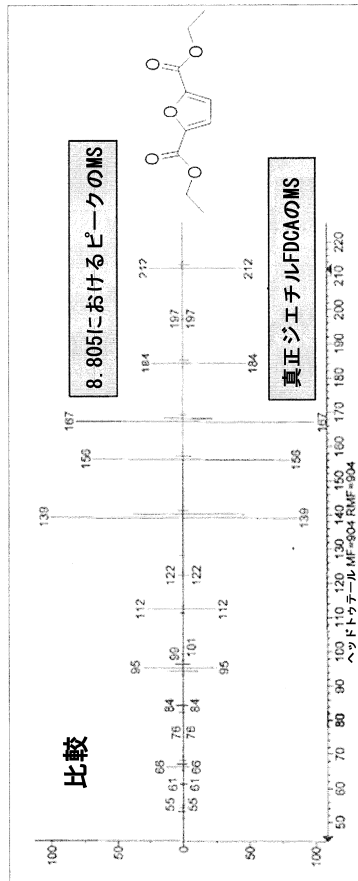
【図 1 3】



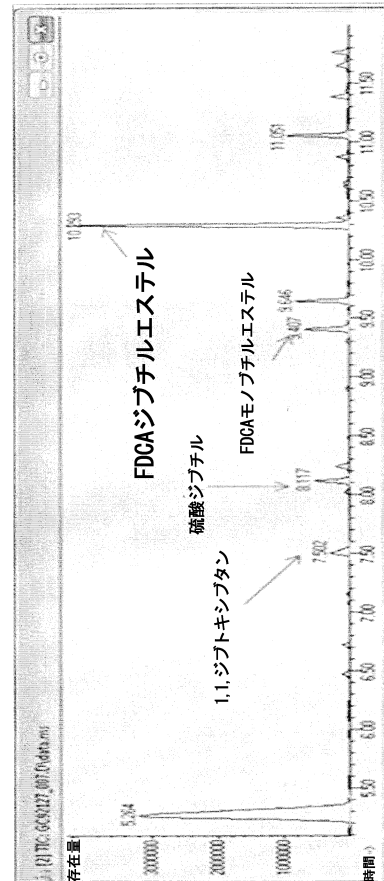
【図 1 4 A】



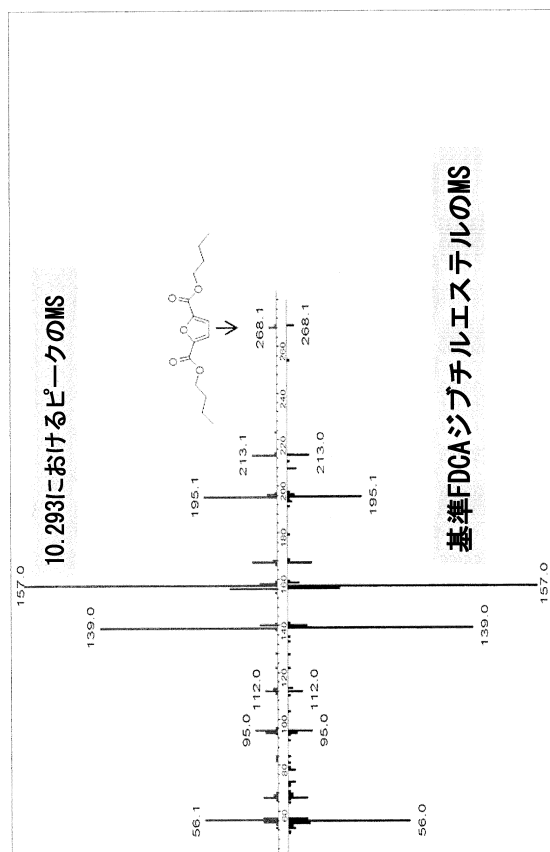
【図 14 B】



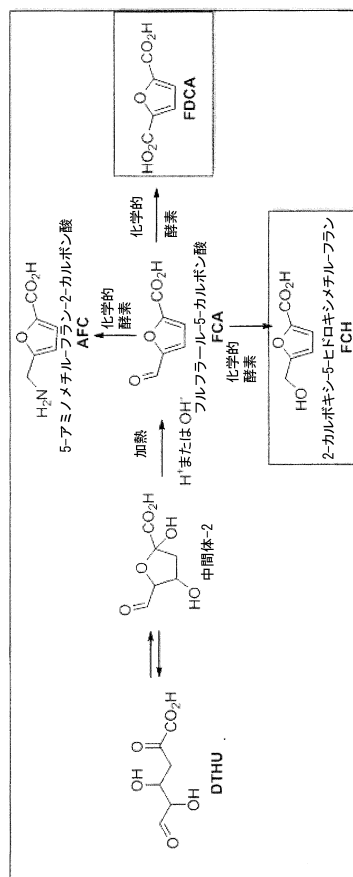
【図 15 A】



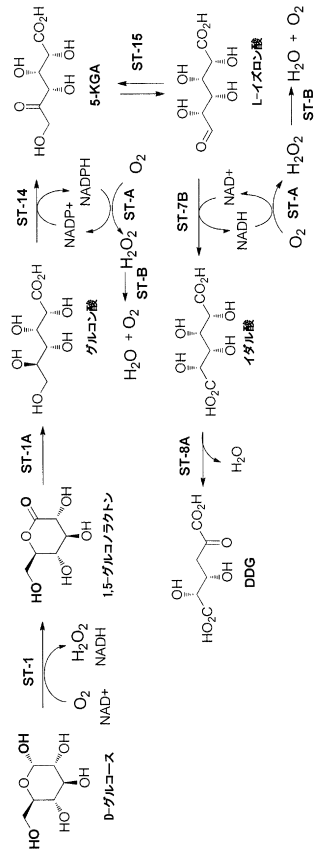
【図 15 B】



【図 16】



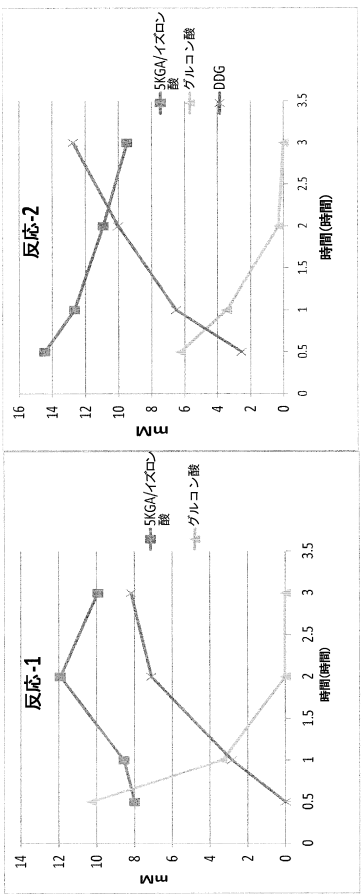
【図 17 A】



【配列表】

0006163557000001.app

【図 17 B】



フロントページの続き

- (74)代理人 100142929
弁理士 井上 隆一
- (74)代理人 100148699
弁理士 佐藤 利光
- (74)代理人 100128048
弁理士 新見 浩一
- (74)代理人 100129506
弁理士 小林 智彦
- (74)代理人 100114340
弁理士 大関 雅人
- (74)代理人 100114889
弁理士 五十嵐 義弘
- (74)代理人 100121072
弁理士 川本 和弥
- (72)発明者 カンブーレイキス スピロス
アメリカ合衆国 カリフォルニア州 サンディエゴ ケンジントン ドライブ 4 8 8 5
- (72)発明者 グリフィン ベンジャミン エム .
アメリカ合衆国 カリフォルニア州 サンディエゴ マエストロ コート 1 2 5 6 2
- (72)発明者 マーティン ケビン ブイ .
アメリカ合衆国 カリフォルニア州 ソラナ ビーチ バーバラ アベニュー 2 4 6

審査官 早川 裕之

- (56)参考文献 特開 2 0 0 8 - 1 2 7 2 8 2 (J P , A)
国際公開第 2 0 1 3 / 0 4 9 7 1 1 (W O , A 1)
仏国特許出願公開第 0 2 7 2 3 9 4 6 (F R , A 1)

- (58)調査した分野(Int.Cl. , D B 名)
- C 0 7 D 3 0 7 / 6 8
C 1 2 N 1 5 / 0 9
C 1 2 P 7 / 4 4
C A p l u s / R E G I S T R Y (S T N)
C A S R E A C T (S T N)