

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la  
Propriété Intellectuelle  
Bureau international



(43) Date de la publication internationale  
03 octobre 2024 (03.10.2024)

WIPO | PCT

(10) Numéro de publication internationale  
WO 2024/200976 A1

(51) Classification internationale des brevets :  
C12N 9/10 (2006.01) C12P 7/22 (2006.01)  
C12N 15/63 (2006.01)

(21) Numéro de la demande internationale :  
PCT/FR2024/050407

(22) Date de dépôt international :  
28 mars 2024 (28.03.2024)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :  
FR2302961 28 mars 2023 (28.03.2023) FR

(71) Déposant : COMPAGNIE GENERALE DES ETABLISSEMENTS MICHELIN [FR/FR] ; 23 place des Carmes-Déchaux, 63000 CLERMONT-FERRAND (FR).

(72) Inventeurs : CANU, Nicolas ; 71 rue Achille Viadieu – Bat. A, Appt 1002, 31400 TOULOUSE (FR). RAGHUNATHAN, Govindan ; 41 avenue Jean Moulin - Bat. A, Apt 42, 31400 TOULOUSE (FR). EL BACHIRI, Mohammed ; 32bis rue Saint Thomas d'Aquin, 31400 TOULOUSE (FR). PANEL, Nicolas ; 56 rue Raspail, 31400 TOULOUSE (FR). SHARMA, Rajendra ; 3 rue Colette – Apt 4, Bat. A, 31200 TOULOUSE (FR). NIGRO, Giuliano ; 65 Avenue Thomas Decaroli, 06700 SAINT LAURENT DU VAR (FR). BOZONNET, Sophie ; 10 allée des Demoiselles – Appt 10, 31400 TOULOUSE (FR). REMAUD-SIMEON, Magali ; 1 rue Benjamin Charrier, 31520 RAMONVILLE (FR). LAFAQUIERE, Vincent ; MANUFACTURE FRANCAISE DES PNEUMATIQUES MICHELIN, Service juridique- Propriété Intellectuelle, DCJ/PI- Site de Ladoux - F35, 23 place des Carmes Déchaux, 63040 CLERMONT FERRAND CEDEX 9 (FR). ANDRE, Isabelle ; 1, rue de l'aqueduc, 31500 TOULOUSE (FR).

(74) Mandataire : REGIMBEAU ; 20, rue de Chazelles, 75847 PARIS CEDEX 17 (FR).

(81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AO,

AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasiatique (AM, AZ, BY, KG, KZ, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publiée:

- avec rapport de recherche internationale (Art. 21(3))
- avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues (règle 48.2(h))
- avec la partie de la description réservée au listage des séquences (règle 5.2(a))

(54) Title: USE OF TYPE III POLYKETIDE SYNTHASES OF CYANOBACTERIA AS PHLOROGLUCINOL SYNTHASES

(54) Titre : UTILISATION DE POLYKETIDE SYNTHASES DE TYPE III DE CYANOBACTERIES COMME PHLOROGLUCINOL SYNTHASES

(57) Abstract: The present invention relates to the use of a polypeptide selected from type III polyketide synthases of cyanobacteria, a nucleic acid molecule coding for same, a vector comprising a nucleic acid molecule coding for same or a host cell expressing same, to produce phloroglucinol. The invention also provides particular nucleic acid molecules coding for a type III polyketide synthase of cyanobacteria, vectors and host cells, as well as methods of producing phloroglucinol.

(57) Abrégé : La présente invention concerne l'utilisation d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, d'une molécule d'acides nucléiques codant pour celui-ci, d'un vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou d'une cellule hôte exprimant celui-ci, pour produire du phloroglucinol. L'invention concerne également des molécules d'acides nucléiques particulières codant pour une polykétide synthase de type III de cyanobactéries, des vecteurs et des cellules hôtes, ainsi que des méthodes de production de phloroglucinol.



WO 2024/200976 A1

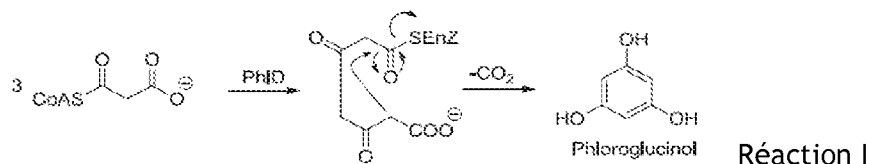
**DESCRIPTION****TITRE : Utilisation de polykétide synthases de type III de cyanobactéries comme phloroglucinol synthases****DOMAINE DE L'INVENTION**

5 La présente invention se situe dans les domaines de la biochimie microbienne et plus particulièrement dans le domaine de la synthèse par des enzymes microbiennes du phloroglucinol. Elle concerne l'utilisation de polykétide synthases de type III de cyanobactéries comme phloroglucinol synthases, des méthodes de production de phloroglucinol associées, ainsi que des acides nucléiques, vecteurs et cellules hôtes destinés  
10 à la production de phloroglucinol.

**ETAT DE LA TECHNIQUE**

Le phloroglucinol est un composé organique aromatique utilisé notamment dans la fabrication de produits pharmaceutiques et d'explosifs.

15 La synthèse de phloroglucinol est catalysée par certaines polykétide synthases de type III appelées phloroglucinol synthases. Les phloroglucinol synthases réalisent la condensation de trois molécules de malonyl-CoA pour former une molécule de phloroglucinol selon le schéma réactionnel suivant (Réaction I):



20

De nombreux oligomères peuvent ensuite être synthétisés à partir du phloroglucinol, tels que les phlorotannins. Les phlorotannins incluent notamment les fucols, les phloretols et les fucophloretols, qui sont des produits dérivés du phloroglucinol composant la paroi des algues brunes. En outre, diverses activités protectrices des algues brunes ont également été  
25 attribuées aux phlorotannins.

La synthèse naturelle de phloroglucinol a été décrite initialement chez les bactéries Gram-*Pseudomonas fluorescens* (Achkar et al., 2005 ; Zha et al., 2006) et chez l'algue brune *Ectocarpus siliculosus* (Meslet-Cladière et al., 2013). L'enzyme phloroglucinol synthase impliquée dans la synthèse du phloroglucinol a été identifiée chez ces deux espèces.

30 Chez *Pseudomonas fluorescens*, la phloroglucinol synthase est codée par le gène **PHLD** (Achkar et al., 2005 ; Zha et al., 2006). L'activité de phloroglucinol synthase de **PHLD** a pu être démontrée chez *Escherichia coli* exprimant un gène **PHLD** hétérologue (Achkar et al.,

2005). Cette activité a été confirmée *in vitro*, par des tests enzymatiques à petite échelle réalisés avec une PHLD recombinante exprimée et purifiée à partir de cultures d'*Escherichia coli* (Zha et al., 2006).

5 Chez *Ectocarpus siliculosus*, la phloroglucinol synthase est codée par le gène **PKS1** (Meslet-Cladière et al., 2013). L'activité de phloroglucinol synthase de **PKS1** a été démontrée *in vitro*, à partir de PKS1 recombinante exprimée et purifiée chez *Escherichia coli* et à partir d'extraits cellulaires d'*E. siliculosus* (Meslet-Cladière et al., 2013, WO 2013/045510).

Toutefois, les enzymes PHLD et PKS1 présentent des activités enzymatiques faibles. En outre, il a été montré que les enzymes PHLD de *Pseudomonas fluorescens* et PKS1 de  
10 *Ectocarpus siliculosus* ne produisent pas ou peu de phloroglucinol lorsque ces séquences sont exprimées chez la levure au lieu de chez *Escherichia coli* (WO2019/002799 ; WO2019/002798). Or, les systèmes eucaryotes peuvent être avantageux, notamment pour des productions à grande échelle. Ils permettent en effet l'obtention d'enzymes pouvant être modifiées au niveau post-traductionnel.

15 WO2019/002799 et WO2019/002798 décrivent des phloroglucinol synthases capables de synthétiser du phloroglucinol à un niveau bien plus élevé que la PKS1 de *Ectocarpus siliculosus* lorsque des séquences codant pour ces enzymes sont transfectées chez la levure. Ces phloroglucinol synthases sont issues de bactéries actinomycètes (WO2019/002799) ou d'algues eucaryotes (WO2019/002798).

20 Il existe cependant un besoin pour d'autres phloroglucinol synthases capables de synthétiser du phloroglucinol à haut niveau, notamment lorsque des séquences codant pour ces enzymes sont transfectées dans une cellule hôte, en particulier chez la levure.

Cependant, les polykétide synthases de type III sont une vaste classe d'enzymes rassemblant des protéines présentant des similarités de séquences tout en présentant des activités  
25 enzymatiques très dissemblables (Meslet-Cladière et al., 2013). Ainsi, l'activité phloroglucinol synthase n'est que l'une des multiples activités susceptibles d'être présentes chez une polykétide synthase de type III et seule une faible proportion des polykétide synthase de type III possède donc une activité phloroglucinol synthase. En outre, aucun motif de séquence associé à l'activité phloroglucinol synthase n'a été décrit, rendant l'identification  
30 des polykétide synthases de type III possédant une telle activité particulièrement difficile.

## EXPOSE DE L'INVENTION

De manière surprenante, car des gènes codant pour des phloroglucinol synthases n'avaient jusqu'ici été identifiés que chez *Pseudomonas fluorescens* et chez les bactéries  
35 Actinomycètes et les algues, les Inventeurs ont toutefois pu identifier de nouvelles phloroglucinol synthases issues de cyanobactéries (embranchement *Cyanobacteriota*,

différent de l'embranchement *Actinomycetota*, auquel appartiennent les bactéries Actinomycètes), capables de synthétiser du phloroglucinol à un niveau intéressant lorsque des séquences codant pour ces enzymes sont transfectées dans une cellule hôte, en particulier chez une levure ou chez une bactérie.

- 5 Ainsi, selon un premier aspect, l'invention concerne l'utilisation d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactérie, d'une molécule d'acides nucléiques codant pour celui-ci, d'un vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou d'une cellule hôte exprimant celui-ci, pour produire du phloroglucinol.
- 10 Selon un deuxième aspect, l'invention concerne une molécule d'acides nucléiques isolée comprenant une séquence d'acides nucléiques codant pour un polypeptide tel que défini dans l'utilisation selon l'invention, caractérisée en ce que :
- a) la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques ; ou
  - 15 b) la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques ;ou
  - c) la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, en particulier dans une levure ou une bactérie; ou
  - d) toute combinaison de a) à c).
- 20 Selon un troisième aspect, l'invention concerne un vecteur comprenant une molécule d'acides nucléiques selon l'invention, ledit vecteur étant de préférence un plasmide.

Selon un quatrième aspect, l'invention concerne une cellule hôte comprenant une molécule d'acides nucléiques selon l'invention, ou un vecteur selon l'invention.

- 25 Selon un cinquième aspect, l'invention concerne une méthode pour produire du phloroglucinol, comprenant les étapes consistant en :
- (i) la mise en contact d'une cellule hôte exprimant le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini dans l'utilisation selon l'invention, par exemple une cellule hôte selon l'invention, avec un substrat approprié ;
  - 30 (ii) la culture *in vitro* de la cellule hôte de l'étape (i) dans des conditions permettant la croissance de ladite cellule hôte et/ou l'expression de la molécule d'acides nucléiques contenue dans ladite cellule hôte, de manière à produire du phloroglucinol ;

(iii) optionnellement la récupération du milieu de culture comprenant le phloroglucinol, obtenu après l'étape (ii) ; et

(iv) optionnellement, la purification du phloroglucinol à partir du milieu de culture de l'étape (iii).

5 Selon un sixième aspect, l'invention concerne une méthode pour produire du phloroglucinol, comprenant les étapes consistant en :

(i) la mise en contact d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini dans l'utilisation selon l'invention avec du malonyl-CoA ;

10 (ii) l'incubation du mélange issu de l'étape (i) dans des conditions adaptées pour produire du phloroglucinol ;

(iii) optionnellement la récupération du milieu réactionnel comprenant le phloroglucinol, obtenu après l'étape (ii) ; et

15 (iv) optionnellement, la purification du phloroglucinol à partir du milieu réactionnel de l'étape (iii).

## DESCRIPTION DES FIGURES

[Fig. 1] La figure 1 représente l'alignement des séquences d'acides aminés SEQ ID NO :1 (PhlD de *Rivularia sp. PCC 7116* PlhD-Rs), SEQ ID NO :2 (PhlD de *Chamaesiphon minutus* 20 dénotée PlhD-Cm), SEQ ID NO :3 (PhlD de *Filamentous cyanobacterium CCP2* dénotée PlhD-Fc), SEQ ID NO :4 (PhlD de *Calothrix sp. HK-06* dénotée PlhD-Cs), et SEQ ID NO :5 (PhlD de *Gloeocapsa sp. PCC 7428* dénotée PlhD-Gs). Des fragments conservés (a) à (g) sont signalés par des encadrements en gras.

[Fig. 2] la figure 2 représente la carte du plasmide PBIM5 utilisé dans les exemples.

25

## DESCRIPTION DETAILLEE DE L'INVENTION

### Définitions

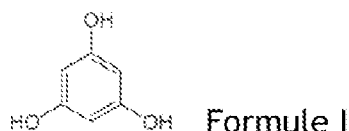
Par « un » ou « une », on entend un/une ou plusieurs. Autrement dit, lorsqu'on utilise « un » ou « une » vis-à-vis d'une caractéristique, on couvre à la fois les modes de réalisation avec 30 une seule fois la caractéristique d'intérêt et ceux avec plusieurs fois la caractéristique d'intérêt. En d'autres termes, sauf indication contraire (comme « un seul » ou « un unique » ou « seulement un »), « un(e) » est utilisé comme synonyme de « un(e) ou plusieurs » ou « au moins un(e) ».

Dans le présent document, lorsqu'ils sont utilisés pour définir des produits, des compositions et des méthodes, les termes « **comprenant** » (et toute forme de « comprenant », telle que « comprend »), « **ayant** » (et toute forme de « ayant », telle que « a »), « **incluant** » (et toute forme d'inclusion, telle que « inclut ») ou « **contenant** » (et toute forme de « contenant », telle que « contient ») sont ouverts et n'excluent pas des éléments ou des étapes de méthode supplémentaires, non mentionnés. Ainsi, un polypeptide « comprend » une séquence d'acides aminés lorsque la séquence d'acides aminés fait partie de la séquence finale d'acides aminés du polypeptide. Un tel polypeptide peut avoir jusqu'à plusieurs centaines de résidus d'acides aminés supplémentaires. Par « **consistant essentiellement en** » ou « **constitué essentiellement de** », on entend l'exclusion d'autres composants ou étapes d'une quelconque importance essentielle. Ainsi, un polypeptide "consiste essentiellement en" une séquence d'acides aminés lorsqu'une telle séquence d'acides aminés est présente avec éventuellement seulement quelques résidus d'acides aminés supplémentaires (par exemple, un peptide d'au plus 20 acides aminés, tel qu'une étiquette de 6 histidines Hisx6, peut en outre être présent). « **Consistant en** » ou « **constitué de** » signifie excluant plus que des traces d'autres composants ou étapes. Par exemple, un polypeptide "consiste en" une séquence d'acides aminés lorsque le polypeptide ne contient pas d'autres acides aminés que la séquence d'acides aminés indiquée.

Par « **polykétide synthase de type III** », on entend une enzyme multifonctionnelle ou un complexe enzymatique produisant des polykétides et qui n'utilise pas de domaine de protéine porteuse d'acyle (ou ACP, pour « acyl carrier protein »).

Par « **polykétide** », on entend une grande famille de métabolites secondaires chez les bactéries, les mycètes, les plantes et certaines lignées d'animaux qui proviennent de la condensation itérative de sous-unités acétyle ou malonyle par des enzymes polykétide-synthases. Les polykétides servent également de matières premières pour la fabrication d'un éventail large de produits naturels et semi-synthétiques.

Par « **phloroglucinol** », on entend un composé organique aromatique benzène-1,3,5-triol présentant la formule chimique suivante (Formule I) :



Par « **phloroglucinol synthase** », on entend une enzyme multifonctionnelle ou un complexe enzymatique appartenant à la famille des **polykétide synthases de type III** et catalysant la synthèse du phloroglucinol. Une phloroglucinol synthase catalyse la condensation de trois molécules de malonyl-CoA pour former une molécule de phloroglucinol.

5

Par « **activité enzymatique** » ou « **activité catalytique** » ou encore « **activité** » d'une enzyme, on entend l'efficacité d'une enzyme à convertir un substrat en produit dans un environnement donné. L'efficacité de l'enzyme prend en compte ici la vitesse de conversion du substrat en produit par l'enzyme et le taux de conversion du substrat en produit par l'enzyme. Par « **taux de conversion du substrat en produit par l'enzyme** » on entend ici le ratio entre la quantité de produit final obtenu par rapport à la quantité initiale de substrat pour une quantité définie d'enzyme. Par exemple, une activité enzymatique au sens de l'invention peut être exprimée en quantité de phloroglucinol produit dans un volume donné (en g/L).

15

Par « **bactérie** », on entend un organisme microscopique et procaryote (sans noyau intracellulaire) présent dans tous les milieux.

Par « **cyanobactérie** », on entend un organisme unicellulaire du règne des bactéries, et de l'embranchement des cyanobactéries ou « *Cyanobacteriota* », un embranchement comprenant des bactéries photosynthétiques, c'est à dire qui utilisent l'énergie lumineuse comme source d'énergie. Cet embranchement comprend principalement la classe des *Cyanophyceae*, qui elle-même comprend notamment les ordres des *Nostocales*, des *Synechococcales*, et des *Chroococcales*. L'ordre des *Nostocales* comprend notamment la famille des *Rivulariaceae*, qui elle-même comprend notamment les genres *Rivularia* (auquel appartiennent de nombreuses espèces, dont *Rivularia sp. PCC7116* dont est issue la PHID-Rs de séquence SEQ ID NO :1) et *Calothrix* (auquel appartiennent de nombreuses espèces, dont *Calothrix sp. HK-06* dont est issue la PHID-Cs de séquence SEQ ID NO :4). L'ordre des *Synechococcales* comprend notamment la famille des *Chamaesiphonaceae*, qui elle-même comprend le genre *Chamaesiphon* (auquel appartiennent de nombreuses espèces, dont *Chamaesiphon minutus* dont est issue la PHID-Rs de séquence SEQ ID NO :2). L'ordre des *Chroococcales* comprend notamment la famille des *Chroococcaceae*, qui elle-même comprend le genre *Gloeocapsa* (auquel appartiennent de nombreuses espèces, dont *Gloeocapsa sp. PCC 7428* dont est issue la PHID-Rs de séquence SEQ ID NO :5). L'espèce *Filamentous cyanobacterium CCP2* (dont est issue la PHID-Fc de séquence SEQ ID NO :3) n'a pas encore été attribuée à un genre, une famille ou un ordre particulier de cyanobactéries.

35

Par « **PHLD.Pf** », on entend indifféremment le gène codant pour la phloroglucinol synthase PHLD de *Pseudomonas fluorescens*, ou le polypeptide codé par ce gène.

Par « **PKS1.Es** » ou « **PHLD.Es** », on entend indifféremment le gène codant pour la phloroglucinol synthase PKS1 d'*Ectocarpus siliculosus*, ou le polypeptide codé par ce gène.

5 « **PhID** » ou « **PHLD** » désigne ici un gène candidat codant pour une enzyme phloroglucinol synthase candidate, ou le polypeptide codé par ce gène. Selon la nomenclature choisie par les Inventeurs, « **PhID.ii** » ou « **PHLD.ii** » désigne ici le gène candidat ou le polypeptide candidat issu d'un organisme donné. Les lettres « ii » représentent les initiales du genre et de l'espèce auxquels ledit organisme appartient. Par exemple, PhID.Hm correspond à une  
10 enzyme phloroglucinol synthase candidate d'une espèce *Hirsutella minnesotensis*, et PhID.At correspond à une enzyme phloroglucinol synthase candidate d'une espèce *Aspergillus tanneri*. Un nombre peut suivre les initiales lorsque plusieurs enzymes phloroglucinol synthases candidates ont été identifiées chez la même espèce.

Par « **acides aminés apolaires** » ou « **acides aminés non polaires** », on entend une famille  
15 d'acides aminés dont les positions moyennes des charges partielles positives et négatives sont confondues. Parmi les 20 acides aminés conventionnels, les acides aminés apolaires sont les suivants : la glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M) et cystéine (C). Ceux autres que la glycine (G) peuvent être subdivisés en trois sous-familles : 1) les « **acides  
20 aminés apolaires aliphatiques** », qui possèdent une chaîne latérale de type aliphatique et incluent l'alanine (A), la valine (V), la leucine (L), l'isoleucine (I), et la proline (P), 2) les « **acides aminés apolaires aromatiques** », qui possèdent une chaîne latérale de type aromatique et incluent le tryptophane (W), la phénylalanine (F), et la tyrosine (Y), et 3) les « **acides aminés apolaires soufrés** » dont la chaîne latérale comprend un atome de soufre  
25 et qui incluent la méthionine (M) et la cystéine (C).

Par « **acides aminés polaires** », on entend une famille d'acides aminés dont les positions moyennes des charges partielles positives et négatives ne sont pas confondues. Parmi les 20 acides aminés conventionnels, les acides aminés polaires sont les suivants : sérine (S), thréonine (T), asparagine (N), glutamine (Q), acide glutamique (E), acide aspartique (D),  
30 lysine (K), arginine (R), histidine (H). Ils peuvent être subdivisés en trois sous-familles : 1) les « **acides aminés polaires neutres (ou non chargés)** », qui possèdent une chaîne latérale non chargée et incluent la sérine (S), la thréonine (T), l'asparagine (N), et la glutamine (Q), 2) les « **acides aminés polaires chargés négativement** », qui possèdent une chaîne latérale chargée négativement et incluent l'acide glutamique (E) et acide aspartique (D), et 3) les

« **acides aminés polaires chargés positivement** » qui possèdent une chaîne latérale chargée positivement et qui incluent la lysine (K), l'arginine (R), et l'histidine (H).

Par « **molécule d'acides nucléiques** », on entend un polymère de n'importe quelle longueur d'acide désoxyribonucléique (ADN), ou polydésoxyribonucléotides, incluant notamment les  
5 ADN complémentaires ou ADNc, les ADN génomiques, les plasmides, les vecteurs, les génomes viraux, l'ADN isolé, les sondes, les amorces et tout mélange de ceux-ci ; ou un polymère de n'importe quelle longueur d'acide ribonucléique (ARN), ou polyribonucléotides, incluant notamment les ARN messagers ou ARNm, ARN antisens ; ou des polyribo-  
10 polydésoxyribonucléotides mélangés. Ils englobent des polynucléotides simples ou à double brins, linéaires ou circulaires, naturels ou synthétiques. En outre, un polynucléotide peut comprendre des nucléotides non naturels et peut être interrompu par des composants non nucléotidiques.

Dans le cadre de la présente invention, les termes "acide nucléique", "molécule d'acides nucléiques", "polynucléotide" et "séquence nucléotidique" sont utilisés de manière  
15 interchangeable.

Par « **molécule isolée** », on entend une molécule, notamment une protéine, un polypeptide, un peptide, une molécule d'acides nucléiques, un vecteur plasmidique, un vecteur viral ou une cellule hôte, qui est extrait de son environnement naturel (c'est-à-dire séparé d'au moins un autre composant auquel il est naturellement associé).

20 Par « **polypeptide** », "**protéine**" et "**peptide**", on entend des polymères de résidus d'acides aminés qui comprennent au moins neuf acides aminés liés par des liaisons peptidiques. Le polymère peut être linéaire, ramifié ou cyclique. Le polymère peut comprendre des acides aminés naturels et/ou analogues d'acides aminés et il peut être interrompu par des résidus non-acides aminés. Comme indication générale et sans y être cependant lié dans la présente  
25 demande, si le polymère d'acides aminés contient plus de 50 résidus d'acides aminés, il est de préférence appelé un polypeptide ou une protéine, alors que si le polymère est constitué de 50 acides aminés ou moins, il est de préférence appelé "peptide".

Par « **identité** », on entend une correspondance exacte de séquence entre deux polypeptides ou deux molécules d'acides aminés. Les « **pourcentages d'identité** » auxquels il est fait  
30 référence dans le cadre de l'exposé de la présente invention sont déterminés sur la base d'un alignement global des séquences (nucléiques ou protéiques) à comparer, c'est-à-dire sur un alignement des séquences prises dans leur intégralité sur toute leur longueur en utilisant tout algorithme bien connu de l'homme du métier tel que l'algorithme de Needleman et Wunsch-1970. Cette comparaison de séquences peut être effectuée à l'aide

de tout logiciel bien connu de l'homme du métier, par exemple le logiciel needle en utilisant le paramètre « Gap open » égal à 10.0, le paramètre « Gap extend » égal à 0.5 et une matrice « Blosum 62 ». Le logiciel needle est par exemple disponible sur le site internet ebi.ac.uk world wide sous la dénomination « Align ».

- 5 Par « **vecteur** », on entend un véhicule, de préférence une molécule d'acide nucléique ou une particule virale, qui contient les éléments nécessaires pour permettre l'administration, la propagation et/ou l'expression d'une ou plusieurs molécule(s) d'acides nucléiques dans une cellule hôte ou un organisme.

D'un point de vue fonctionnel, ce terme englobe des vecteurs pour la maintenance (vecteurs  
10 de clonage), des vecteurs pour l'expression dans diverses cellules ou organismes hôtes (vecteurs d'expression), des vecteurs extrachromosomiques (par exemple des plasmides multicopie) ou des vecteurs d'intégration (par exemple conçus pour s'intégrer dans le génome d'une cellule hôte et produire des copies supplémentaires de la molécule d'acides nucléiques qu'il contient lorsque la cellule hôte se réplique). Ce terme englobe également des vecteurs  
15 navettes (par exemple, fonctionnant à la fois dans des hôtes procaryotes et/ou eucaryotes) et des vecteurs de transfert (par exemple pour le transfert de molécule(s) d'acides nucléiques dans le génome d'une cellule hôte).

D'un point de vue structurel, les vecteurs selon l'invention peuvent être des sources  
20 génétiques naturelles, synthétiques ou artificielles, ou une combinaison d'éléments génétiques naturels et artificiels.

Ainsi, dans le contexte de l'invention, le terme "vecteur" doit être compris de manière large en incluant des vecteurs plasmidiques (ou plasmides) et viraux.

Un "**plasmide**" tel qu'utilisé ici désigne une construction d'ADN répliquable. Habituellement, les vecteurs plasmidiques contiennent des gènes marqueurs de sélection qui permettent aux  
25 cellules hôtes portant le plasmide d'être identifiées et/ou sélectionnées de manière positive ou négative en présence du composé correspondant au marqueur de sélection. Une variété de gènes marqueurs de sélection positifs et négatifs sont connus dans la technique. À titre d'illustration, un gène de résistance aux antibiotiques peut être utilisé comme un gène marqueur de sélection positif permettant de sélectionner une cellule hôte en présence de  
30 l'antibiotique correspondant.

Le terme "**vecteur viral**" tel qu'utilisé ici se réfère à un vecteur d'acide nucléique qui comprend au moins un élément d'un génome du virus et peut être conditionné dans une particule virale ou une particule virale. Les vecteurs viraux peuvent être compétents pour la répllication ou sélectifs (par exemple, conçus pour se répliquer mieux ou sélectivement

dans des cellules hôtes spécifiques), ou peuvent être génétiquement désactivés de manière à être défectueux ou déficients pour la réplication.

Par « **cellule hôte** », on entend une cellule contenant une molécule d'acides nucléiques hétérologue. Par « hétérologue » ou « exogène », on entend que la molécule d'acides nucléiques provient d'une espèce différente de l'espèce de la cellule hôte. Ainsi, une cellule hôte n'est pas une cellule qui existe à l'état naturel mais est un outil de biologie moléculaire obtenu par les techniques de manipulations génétiques.

La cellule hôte peut être constituée d'un type unique de cellules ou d'un groupe de différents types de cellules (on a alors un mélange de cellules hôtes de type distinct). La cellule hôte peut également être une cellule hybride, c'est-à-dire résultant de la fusion d'au moins deux cellules de types différent. La cellule hôte peut appartenir à des lignées cellulaires cultivées, à des cellules primaires, à des cellules souches ou à des cellules prolifératives. Dans le cadre de l'invention, le terme "cellule hôte" comprend des cellules procaryotes et des cellules eucaryotes. Par « **procaryote** », on entend un micro-organisme unicellulaire dont la structure cellulaire ne comporte pas de noyau. Les procaryotes incluent les règnes des bactéries et des archées. Par « **eucaryote** », on entend par opposition aux procaryotes tout organisme unicellulaire ou multicellulaire dont les cellules possèdent un noyau structuré. Les cellules eucaryotes incluent les cellules de levures, d'insectes, de plantes et d'animaux (notamment des cellules de mammifères non humaines). La cellule hôte peut par exemple être isolée ou organisée en tissu, en organe ou encore être au sein d'un organisme complet. Dans le cas où la cellule hôte est au sein d'un organisme complet, ledit organisme n'est pas humain.

Dans la description détaillée qui suit, les modes de réalisation pourront être pris seuls ou combinés de manière appropriée par l'homme du métier.

### **Utilisation de polypeptides isolés, molécules d'acides nucléiques, vecteurs, cellules hôtes pour la production de phloroglucinol**

On s'intéresse ici à l'utilisation de polypeptides isolés choisis parmi les polykétide synthases de type III de cyanobactéries.

En effet, de manière tout à fait surprenante, les Inventeurs ont identifié des gènes codant de nouvelles polykétide synthases de type III dans le génome de cyanobactéries. De plus, les polykétide synthases de type III sont une vaste classe d'enzymes rassemblant des protéines présentant des similarités de séquences tout en présentant des activités enzymatiques très

dissemblables (Meslet-Cladière et al., 2013). Ainsi, l'activité phloroglucinol synthase n'est que l'une des multiples activités susceptibles d'être présentes chez une polykétide synthase de type III. En outre, aucun motif de séquence associé à l'activité phloroglucinol synthase n'a été décrit, rendant l'identification des polykétide synthases de type III possédant une telle activité particulièrement difficile. Or, les Inventeurs ont en outre mis en évidence que les nouvelles polykétide synthases de type III identifiées chez des cyanobactéries ont une activité de phloroglucinol synthase.

Dans un premier aspect, l'invention concerne donc l'utilisation d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, d'une molécule d'acides nucléiques isolée codant pour celui-ci, d'un vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou d'une cellule hôte non humaine exprimant celui-ci, pour produire du phloroglucinol.

### ***Polypeptide utilisé***

Le polypeptide utilisé dans l'invention est choisi parmi les polykétide synthases de type III de cyanobactéries, et possède avantageusement une activité phloroglucinol synthase.

Il peut notamment être choisi parmi les polykétide synthases de type III de cyanobactéries des familles *Rivulariaceae*, *Chamaesiphonaceae* et *Chroococcaceae*, avantageusement des genres *Rivularia*, *Calothrix*, *Chamaesiphon* et *Gloeocapsa*, plus avantageusement des espèces *Rivularia sp. PCC 7116*, *Calothrix sp. HK-06*, *Chamaesiphon minutus*, *Gloeocapsa sp. PCC 7428*, et *Gloeocapsa sp. PCC 73106*. Alternativement, il peut être choisi parmi les polykétide synthases de type III de cyanobactéries de l'espèce *Filamentous cyanobacterium CCP2*.

Avantageusement, le polypeptide utilisé dans l'invention correspond à ceux identifiés ici comme ayant une activité phloroglucinol synthase ou à des dérivés de celui-ci. Notamment, le polypeptide utilisé dans l'invention comprend, est essentiellement constitué ou est constitué d'une séquence d'acides aminés ayant au moins 70% d'identité, au moins 75% d'identité, avantageusement au moins 80% d'identité, au moins 85% d'identité, plus avantageusement au moins 90% d'identité, au moins 91% d'identité, au moins 92% d'identité, au moins 93% d'identité, au moins 94% d'identité, encore plus avantageusement au moins 95% d'identité, au moins 96% d'identité, au moins 97% d'identité, au moins 98% d'identité, au moins 99% d'identité, au moins 99,5% d'identité, voire 100% d'identité avec une séquence choisie parmi SEQ ID NO :1 (PhlD de *Rivularia sp. PCC 7116* PhlD-Rs), SEQ ID NO :2 (PhlD de *Chamaesiphon minutus* dénotée PhlD-Cm), SEQ ID NO :3 (PhlD de *Filamentous*

*cyanobacterium CCP2* dénotée PhD-Fc), SEQ ID NO :4 (PhD de *Calothrix sp. HK-06* dénotée PhD-Cs), et SEQ ID NO :5 (PhD de *Gloeocapsa sp. PCC 7428* dénotée PhD-Gs).

Les séquences SEQ ID NO :1 à 5 sont présentées dans le **Tableau 1** ci-dessous :

[Table 1]

Identifiant	Nom	Espèce d'origine	Séquence d'acides aminés
SEQ ID NO :1	PHLD -Rs	<i>Rivularia sp. PCC 7116</i>	MTNTTIRVKQKGLIARLSNHNISNISSKQVLPTIEGIATGTPDNIISQK DAAEFVANLNLDNRNRHRIAKIYQNTRIEERLAVNLLSEESYEFSTRKQ GTIQQRMEMFEEYAFPLAEKVAKQALKSAEASTRDDGSLYSDNIEESI GLIVFVTSTGFVAPGIDTKLIKSLGLKRNIARIPVHFMGCAAAMNGLRI ACDYIKANPNQRALVVCELESSVNAVFEEDNINDTIIHSIFGDGCAAVL GGREENMLFGRNKIVIKDNFSYLVDDEDGIRLGINDSGITCKLSRDL SYIENGLNPIINNFLADCQLTKEQIDLWAVHPGGTRIEKAQSSLGLTD EQVADSWELRQYGNMLSCAILFVIERVMLRLESESFQDGKTIPEKLTG LSFSFAPGIGVEGILFEKY
SEQ ID NO :2	PHLD -Cm	<i>Chamaesiphon minutus</i>	MSNQFLPIVESIATGTPQHTVYQADAARFVSELYGVGQHQERIEKIYK NTRIENRHLALDLLTDEVVDFCRQPSNIKQRMDLYAEHAVPLATKVAS KALAIATARNRQERPNDTTKIEDEIGLIVFVTSTGFLAPGVDTKVIEHL GLRRNIARIPVNFMGCAAAITGLRVACDYLRAYPHSKALMVCELESSV NSSFGDNLNDIIHSIFGDGCAAVLGGARDIANLAAANRLVIRDNFSYL VEDTADGIVLDVQANGITCQLSPQLPGYIEAGVDPIITNLFSEHQLAK ADIDLWVHPGGTRIEKQVRSGLTDEQVDDSWELREYGNVLSCAV LFTLERMLIRLDSSSHFDRTEPLTGVAFSFSFPGVIGIEGILFEKY
SEQ ID NO :3	PHLD -Fc	<i>Filamentous cyanobacterium CCP2</i>	MLANAFQTTSQLSNSQHSNRPQSSYFPHPSIAHPSVMP TIERLATGT PEYIIPQQQAAQIIANLPHLEQSRDRIEKLYNNTRIHTRHLAVDLRSEA TIDRSHRSGTIQSRMDHFKA EAVPLAERVVQQALVGVSIETIGLVVVF TSTGFIAPGVDAELIERLGLRRNIGRLTVNFMGCAAAMNGLRAACDH VKAYPHQKALLVCELESSVNAVFDLNDVIIHSIFGDGCAAAIIGSQE VAEARARGQIIIRDHLSYLIEGTQDGITLGIRDNGITCRLSRQLPDYIEA GVGAVIGNFLAERQMTKADIDLWAVHPGGTRIIQKSQSALALTDEQV ACSWSVLRDYG NMLSILFVLERMLQQHQTVPSSQTATGLAFSFSF GVGVEGMLFEVA
SEQ ID NO :4	PHLD -Cs	<i>Calothrix sp. HK-06</i>	MLSNNILATIEIATGTPDNVILQSDAAKTVANLPHLQENASRIEKIYY NTRIYTRHLALNLLSDEVLAWCQNASIQERMRLFHSYGVPLAEKVAK KALANMTSTRNNLETVKDSIRQIVLVSSTGFVAPGIDTALIQLGLRRD ISRVTVNFMGCAAAMNGIRVASDHVLANPTHKSLVCELESSINAVFN DNINDVIIHSIFGDGCAAVVVGACEAEQARKQNKVVIRDYLSYLIENSE DGINLGVEDTGITCKLSRYLPNYIEAGVGSIIKDFLATNHLTKENIDLW

			AVHPGGTRILEKVQISLSLKDCQLDNSWKILWEYGNMLSASVLFVLEK MLFNsvgyqSNDASTGLAFsfSPGVGLEGILFEKGYVNI
SEQ ID NO :5	PHLD -Gs	<i>Gloeocapsa</i> <i>sp. PCC</i> 7428	MNTKVIEMNKAHsrPFVAPTQmKANQlDELcVIRskRLLPTIESIAT GTPNNVIRQSDAAKFVANfPTLEqNQfRIEKLYSNTRIDTRHLAVNLL SDEAIAlSGHNIPIQSRMQMFQeYAVPLAERVAKQALLSATASIDPDDP WQSELKIEDSIRLIVfVSSTGFVAPGVDTALIEKLGRRDIARVTvNFM GCAAAMNGLRVACDHVRANpTHKALVVCLELSSVNAVFADEVNDVV IHSIFSDGCAAVVIGACEEDRAIAQgKVVIRDHLSYLVEDTQDGIVLGI HDNGITCQLSRQLPEYIETGVGPIIERFLASHALTkESIDLWAVHPGG TRIIEKAQRSLGLSDRQVADsWEILRQYGNMLSPSVLFLERMLQRCE QDSTEMKPLTGIAfSFSPGVGVEGILFQKV

En outre, les Inventeurs ont identifié des motifs conservés entre les cinq séquences SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5 (voir l'alignement de la **Figure 1**), qui sont donc avantageusement présents dans la séquence du polypeptide utilisé dans l'invention.

- 5 Ainsi, dans un mode de réalisation avantageux général, le polypeptide utilisé dans l'invention comprend en outre (en sus d'un pourcentage d'identité minimal avec SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 ou SEQ ID NO :5 tel que décrit ci-dessus) au moins un fragment choisi parmi :
  - (a) ATGTP (SEQ ID NO:6) ;
  - 10 (b) STGFX<sub>1</sub>APG, dans laquelle X<sub>1</sub> est choisi parmi les acides aminés apolaires (i.e. glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M) et cystéine (C), SEQ ID NO :7);
  - (c) VX<sub>2</sub>FMGCAA, dans laquelle X<sub>2</sub> est choisi parmi les acides aminés polaires (i.e. sérine (S), thréonine (T), asparagine (N), glutamine (Q), acide glutamique (E), acide aspartique (D),
  - 15 lysine (K), arginine (R), et histidine (H), SEQ ID NO:8) ;
  - (d) VCLELSS (SEQ ID NO :9) ;
  - (e) IHSIFX<sub>3</sub>DGCAA, dans laquelle X<sub>3</sub> est n'importe quel acide aminé (i.e. X<sub>3</sub> est choisi parmi glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M), cystéine (C), sérine (S), thréonine (T),
  - 20 asparagine (N), glutamine (Q), acide glutamique (E), acide aspartique (D), lysine (K), arginine (R), et histidine (H), SEQ ID NO :10); et

(f) IDLWX<sub>4</sub>VHPGGTRI, dans laquelle X<sub>4</sub> est choisi parmi les acides aminés apolaires (i.e. glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M) et cystéine (C), SEQ ID NO :11).

5 Avantageusement, dans un mode de réalisation intermédiaire, le polypeptide utilisé dans l'invention comprend en outre (en sus d'un pourcentage d'identité minimal avec SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 ou SEQ ID NO :5 tel que décrit ci-dessus) au moins un fragment choisi parmi :

(a) ATGTP (SEQ ID NO:6) ;

10 (b) STGFX<sub>1</sub>APG, dans laquelle X<sub>1</sub> est choisi parmi les acides aminés apolaires aliphatiques (i.e. l'alanine (A), la valine (V), la leucine (L), l'isoleucine (I), et la proline (P), SEQ ID NO :12) ;

(c) VX<sub>2</sub>FMGCAAA, dans laquelle X<sub>2</sub> est choisi parmi les acides aminés polaires neutres ou polaires chargés positivement (i.e. sérine (S), thréonine (T), asparagine (N), glutamine (Q), lysine (K), arginine (R), et histidine (H), SEQ ID NO:13) ;

(d) VCLELSS (SEQ ID NO :9) ;

(e) IHSIFX<sub>3</sub>DGCAA, dans laquelle X<sub>3</sub> est choisi parmi les acides aminés apolaires ou polaires neutres (i.e. glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M), cystéine (C), sérine (S), thréonine (T), asparagine (N), et glutamine (Q), SEQ ID NO :14); et

(f) IDLWX<sub>4</sub>VHPGGTRI, dans laquelle X<sub>4</sub> est choisi parmi les acides aminés apolaires aliphatiques (i.e. l'alanine (A), la valine (V), la leucine (L), l'isoleucine (I), et la proline (P), SEQ ID NO:15).

25 Plus avantageusement encore, dans un mode de réalisation limité, le polypeptide utilisé dans l'invention comprend en outre (en sus d'un pourcentage d'identité minimal avec SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 ou SEQ ID NO :5 tel que décrit ci-dessus) au moins un fragment choisi parmi :

(a) ATGTP (SEQ ID NO:6) ;

30 (b) STGFX<sub>1</sub>APG, dans laquelle X<sub>1</sub> est choisi parmi la leucine (L), la valine (V), et l'isoleucine (I) (SEQ ID NO :16),

(c) VX<sub>2</sub>FMGCAAA, dans laquelle X<sub>2</sub> est choisi parmi l'asparagine (N) et l'histidine (H) (SEQ ID NO :17);

(d) VCLELSS (SEQ ID NO :9) ;

5 (e) IHSIFX<sub>3</sub>DGCAA, dans laquelle X<sub>3</sub> est choisi parmi la glycine (G) et la sérine (S) (SEQ ID NO :18) ; et

(f) IDLWX<sub>4</sub>VHPGGTRI, dans laquelle X<sub>4</sub> est choisi parmi l'alanine (A) et la valine (V) (SEQ ID NO :19).

10 Avantageusement, le polypeptide utilisé dans l'invention comprend en outre (en sus d'un pourcentage d'identité minimal avec SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 ou SEQ ID NO :5 tel que décrit ci-dessus) plusieurs des fragments (a) à (g) tels que décrits ci-dessus (mode de réalisation général, intermédiaire ou limité). Notamment, le polypeptide utilisé dans l'invention peut comprendre en outre (en sus d'un pourcentage d'identité minimal avec SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 ou SEQ ID NO :5 tel que

15 décrit ci-dessus (mode de réalisation général, intermédiaire ou limité)) :

- les fragments (b), (c), et (e),
- les fragments (d) et (f),
- les fragments (b) à (f), ou
- tous les fragments (a) à (f).

20

De manière particulièrement avantageuse, le polypeptide utilisé dans l'invention comprend, est essentiellement constitué, ou est constitué d'une séquence d'acides aminés choisie parmi SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5.

## 25 ***Molécule d'acides nucléiques isolée utilisée***

La molécule d'acides nucléiques isolée utilisée dans l'invention peut être toute molécule d'acides nucléiques comprenant une séquence d'acides nucléiques codant pour tout polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus.

Selon un mode de réalisation, la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques codant pour du polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus. De manière avantageuse, le promoteur est un promoteur exogène, en particulier un promoteur de levure, de préférence un promoteur choisi parmi ADH2 (pADH2), ce promoteur permet l'expression notamment lorsque le milieu de culture contient de l'éthanol comme source de carbone), CCW12 (pCCW12, ce promoteur permet l'expression notamment lorsque le milieu de culture contient du glucose comme source de carbone), et TEF1 (pTEF1, ce promoteur permet l'expression notamment lorsque le milieu de culture contient du glucose ou saccharose comme source de carbone), de préférence encore un promoteur choisi parmi ADH2 (pADH2) de *Saccharomyces cerevisiae* et CCW12 (pCCW12) de *S. cerevisiae*, et TEF1 (pTEF1) de *S. cerevisiae*, de préférence encore un promoteur choisi parmi ADH2 de séquence SEQ ID NO: 20, CCW12 séquence de SEQ ID NO: 21 et TEF1 de séquence SEQ ID NO: 22.

Selon un mode de réalisation alternatif ou combinable avec le précédent, la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription de la séquence d'acides nucléiques codant pour du polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus. De manière avantageuse, le terminateur est un terminateur exogène, en particulier un terminateur de levure, de préférence le terminateur RPL3 (tRPL3) ou le terminateur ADH1, de préférence encore le terminateur RPL3 de *S. cerevisiae* ou le terminateur de ADH1 de *S. cerevisiae*, de préférence encore le terminateur RPL3 de séquence SEQ ID NO: 23 ou le terminateur ADH1 de séquence SEQ ID NO:24.

Selon un mode de réalisation préféré, la molécule d'acides nucléiques isolée comprend en outre à la fois un promoteur et un terminateur qui sont tels que définis ci-dessus.

Selon un mode de réalisation alternatif ou combinable avec chacun des modes de réalisations visant la molécule d'acides nucléiques isolée utilisée dans l'invention, la molécule d'acides nucléiques comprend en outre une séquence d'export. Avantageusement, cette séquence d'export permet la sécrétion ou l'excrétion du ou des polypeptides codés par la molécule d'acides nucléiques dans le milieu cellulaire.

Selon un mode de réalisation préféré, la molécule d'acides nucléiques isolée comprend en outre à la fois un promoteur, un terminateur et une séquence d'export tels que définis ci-dessus.

La molécule d'acides nucléiques peut être isolée à partir de souches homologues en culture, de préférence choisies parmi les familles *Rivulariaceae*, *Chamaesiphonaceae* et *Chroococcaceae*, avantageusement des genres *Rivularia*, *Calothrix*, *Chamaesiphon* et *Gloeocapsa*, plus avantageusement des espèces *Rivularia sp. PCC 7116*, *Calothrix sp. HK-06*,  
5 *Chamaesiphon minutus*, *Gloeocapsa sp. PCC 7428*, et *Gloeocapsa sp. PCC 73106*, ou de l'espèce *Filamentous cyanobacterium CCP2*. Alternativement, la molécule d'acides nucléiques peut être isolée à partir d'un vecteur ou d'une cellule hôte hétérologue comprenant ladite molécule, ledit vecteur ou ladite cellule hôte étant tel(le) que défini(e)  
10 plus haut et que décrit(e) ci-après dans les sections « Cellules hôtes » ou « Vecteurs ». Alternativement, la molécule d'acides nucléiques isolée peut être synthétisée *in vitro* par des techniques de synthèse nucléique connues de la personne du métier.

Selon un mode de réalisation alternatif ou combinable avec chacun des modes de réalisations visant la molécule d'acides nucléiques isolée utilisée dans l'invention, la séquence d'acides nucléiques comprise dans la molécule d'acides nucléiques isolée utilisée dans l'invention est  
15 en outre optimisée pour l'expression dans une cellule hôte, en particulier dans une levure ou une bactérie.

Pour optimiser l'expression dans une cellule hôte, en particulier dans une levure ou dans une bactérie, la personne du métier sait utiliser le biais d'usage du code génétique, qui désigne l'utilisation préférentielle pour un organisme donné d'un des triplets de nucléotides  
20 ou codons possibles pour coder un même acide aminé. En effet, il existe en général plusieurs combinaisons de trois nucléotides (appelées « codon ») codant pour le même acide-aminé (sauf pour la méthionine et le tryptophane), appelés codons synonymes, mais certaines de ces combinaisons sont en général utilisées préférentiellement par un organisme donné. Pour la production d'une séquence d'acides aminés d'intérêt, une expression optimale peut ainsi  
25 être obtenue lorsque les codons choisis pour coder la séquence d'acides aminés sont ceux utilisés de manière préférentielle par l'organisme d'origine de la cellule hôte. Selon l'organisme de production choisi (notamment la levure ou la bactérie), des séquences nucléiques optimales différentes seront donc utilisées lorsque la séquence d'acides nucléiques comprise dans la molécule d'acides nucléiques isolée utilisée dans l'invention est  
30 en outre optimisée pour l'expression dans l'organisme de production choisi.

Différents logiciels sont accessibles à la personne du métier pour optimiser les codons pour une expression chez une cellule hôte, notamment chez une levure ou une bactérie. Des exemples de tels logiciels incluent le logiciel Twist Codon Optimization tool (fourni par Twist  
35 Biosciences), le logiciel GenSmart™ Codon Optimization (fourni par GenScript et dont les

fonctionnalités sont décrites dans la demande WO2020024917A1), le logiciel IDT Codon Optimization Tool (fourni par Integrated DNA technologies), et le logiciel Azenta’s codon optimization tool (fourni par Azenta).

Des exemples de séquences nucléiques optimisées codant respectivement pour les polypeptides de séquence d’acides aminés SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5 sont les séquences d’acides nucléiques SEQ ID NO :25 à SEQ ID NO :29, respectivement, présentées dans le Tableau 2 ci-dessous et qui sont optimisées (au moins partiellement) pour une expression chez une levure, notamment *Saccharomyces cerevisiae* :

[Table 2]

Nom	Identifiant séquence d’acides aminés	Séquence nucléique optimisée pour une expression chez une levure, notamment <i>Saccharomyces cerevisiae</i> (identifiant)
PHLD-Rs	SEQ ID NO :1	atgactaataacaacgataagggtaaaacaacaaaaggctcaattgcaaggcttcaaatcacaata tttctaattttcttcaaaacagggtttaccacaattgaaggaatcgcgactggtaaccaggacaat ataataagtcaaaaggacgccgccgagttcgtggtaacctaataaatcttgatcgtaacaggcata ggattgcaaaaatttaccaaaatacaaggatagaggagaggcgtctggccgtaaattgtgtctgaa gaatcatacgaatttagcagaaagcaaggaacaatccaacagagaatggaaatgttgaagagtac gcattcccacttggcgaagggttccaagcaggctttaaactctgctgaggcatctactagagatga tggttccttatactcagataacatcgaagagagattggcttaattgtgtttgtactctactggctt gtcgcaccgggtattgatacaaaagtttaataaagtcattgggattgaagagaatattgcgaggatacc tgtgcattttatgggttgcgcagctgcaatgaatggcttaaggattgcttgcgactatattaagctaa tcctaaccaacgtgcattagtagttgtctggagctgagtagcgtaaatgcggtattcaggagataat taacgacacaattattcactctatattggggatggctgtgcagcgggtgtttgggtggcgtagga aaatagctgttcggctgtaataaaatgtcataaaagacaatttctcataccttgggatgacacaga agatggtattagactaggtatcaacgatccggatcacctgtaaattaagttagagatctccctcata catagaaaatggctcaaatccaattatcaataatttcttagctgactgcaagcttcaaaaagagcaaa tcgaccttgggctgtacatccggcggtaccagaataattgaaaagcacagctcctccttgggattg accgatgagcaagtcgcagattcctgggaaatacttaggcaatagcgtaatatgttgcactgtctatt ctgtttgtcatagaaagattatgcttctgttggaaatcagagcttttcaagatggtaagactattccgg aaaactgacgggtctagtttagcttctcagcattggcgtcgagggaaatattatcgaaaaat actga (SEQ ID NO:25)
PHLD-Cm	SEQ ID NO :2	atgtcaaaccagttcctgccaatagttgaatcaattgccactggactccacagcatacggttacca agcggacgctgcgagattgtatcagaattgtatgggtgctggacaacatcaggaaagaatagaaaag atttataaaaacacaagaatagaaaatcgtcacttggcgtggatctattaaccgacgaagttgtga cttttcagacaacctccaataaaaacaagaatggattgtatgcggaacatgctgtacctcttgc tacaaaagtcgcttcaaaggcacttgaattgctacagcaagaacagacaagaacgtcctaacga cactacaaagattgaggatgaaattggattaatagtagttgttacgtctactgggttcttgcacctgg gtagacacaaaggtaatagaacatttgggtttaaagaagaacatagctagaattcccgtcaatttcat gggttgtgctgctatcacggctcgtgtggcctgtgattacctaagagcttatccacattctaa agctttaatggctgctgcaatattctcagtcataatagttcttccgggataactgaaatgacattata atacattctatttggcgacgggtgtgcagcagtggttttaggcgtagagatcgcctaatttagctg ctgcaaatagattggctataagagataatttagctatttagtagaagacactgctgacggatcgtt tagacgttcaagctaattgatttacctgtcaattgagccccagttgcccggttacattgaagcgggc gttgaccaattataactaatttttgcagagcatcagcttgcaggcagatatacgcattatgggtg gttcatccgggaggcagcagaattatagaaaaggttcaaagatcactgggttaccgatgaacaag tcgatgattcagggaaatttaagggaaatcaggaacgtattgagctgtgccctctttcacgttag aaagaatgttaggttagacagttcctcccactttagtaggacagaacccttaactggcgtagcc ttctatttaccagggtgtaggtattgagggtattctattcgaataattaa (SEQ ID NO:26)

PHLD-Fc	SEQ ID NO :3	<p>atggttagcgaatgcatttcaaacaccattcgcaattgtcaaatagtcaaacactctaatagaccgcaatc  tcttattttcccacccttcaatagcacatccacattccggtatgcctacaatcgaaagacttgcaac  agggactcctgaatacattatcctcaacaacaagcagcacaataatcgtaaccttcccacctg  gaacagtctaggatagaatcgagaagctgtacaacaatacaaggattcacacaagacacttagcc  gtcgacttaaggtctgaagctacaattgatagatcacatagatccggtactattcaatctaggatgga  tcacttaaggccgaagccgtaccattggctgagagagtcgtcaacaggcttagtggcgatctat  cgaactattgggttagtgttttgcacttccaccggcttatcgctcctgggtgcgacgccaattg  attgaaagattgggttgagaagaaataggtagactgactgtaattttatgggatgtgccggct  atgaatggtttaagagccgctgacaccacgtcaagcatacccgcatacagaaggcattgtggctg  cttggactgtcttctgtgaatgcagtttctgatgatgttaaatgatgttatcattcacagattttcg  gtgatggatgtctgcagcaatattgggtctcaggaggtgcagaagcaagagcgagagggcaatt  atcataagagatcatctatcatatctgattgagggcacacaagacgggataacgttaggtattagga  caacggaatcacatgcagactttccagacaattaccagactatattgaagcaggagttggagctgta  attgaaacttcttagctgaaagacagatgacaaaagctgatagattgtgggcagttcacccgg  cgggacgagaatcatcaaaaaatctcagagcgcattggcttaactgatgaacaagtgcagttctt  ggctcggtactaagggtattggaatatgttatccgttagtattttgtttgtttggaaagaatgctaca  acaacacaaaccgtccaagttctcaaccgccacaggattggcttcagctttctccagggtgtg  gtgtgaaaggaatggtgttgaagtagcttaa (SEQ ID NO:27)</p>
PHLD-Cs	SEQ ID NO :4	<p>atggtgagtaataacattttagcgacgatcgagactattgctacaggtacccccgataacgttatatta  cagtcggacgcagctaaaactgtggctaacctacccatttacaagaaaatgcactaggttagaga  aaatttattataaacacgagaatttatcacagtcacttagccctaaactattaagttagttagtcttg  cgtggtgcaaaacgcatacaggaagaatgagattatccatagttatgggtacctttagcc  gagaaggtagcaaaaaagcttggctaataatgactagcactagaataattggaaactgtcaaaag  actctattaggcaaatcgtccttcttcaaccgggttgggtgcccaggaatcgactgcccga  ttcagagcttggcctgagaaggacatctcaagagttaccgtcaatttcaggggtgtgctgccgta  tgaatggtataagagtgcatctgatcatgttttggcaaatcctactcacaagagtttagtggctggtt  agaattatcaagcataaatgctgtattcaacgataataaaacgacgtgataattcattcgatctcg  gagacggatgtgcagctgtcgtcgtgggtgcttgcgaagctgaacaagcccgtaaacaaaataaagt  agtgatcagggttacctttcgtatttaacgaaaatagtgaggacggtattaatttaggcgttagga  cactggcattacttgaattatcaagatatctaccaaatattgaggctgggtggttctatcatc  aaggactccttgaacaaatcatttaacaaaagagaatattgattatgggcagttcacccgggtgg  tactagaatttagaaaaagtaacaaatagcttatcttgaagactgccagctagataattcatgga  aaattctatgggaatacggtaataatgctttctgcatctgtttgtcgtactagagaaaatgtatttaa  tagcgttgggtaccaatccaatgatgcctcaaccggctagcatttagttctctccaggggtaggctt  agaaggaattcttttgagaaaggttacgtgaatatctaa (SEQ ID NO:28)</p>
PHLD-Gs	SEQ ID NO :5	<p>atgaataccaaagttatagaatgaacaaagccgctcattcaagaccattcgtggctccaacacaaa  tgaaggcgaatcaaatgttgatgagttgtgtgtattaggctcaagagactttgccaacgattgaat  ccattgctactggtactccgaacaatgtaatacaggcaaacgatgcggctaaatttggccaacttc  ccaacattggaacagaatcaattcaggattgaaaaactatattctaacaccagaatagatacaagac  acttggcagtgaaatctttaaagtgatgaagcaattgccttatctggtcacaacattccaattcaaagta  gaatgcaaatgttccaagaatacgtgttcccttggcagaaagagttgcaaaacaagcgttgctatca  gctacagctagtattgatcctgatgatccatggcaaacggaattgaaaatcgaggattctattaggctt  atcgttttgttcttaccgggttggctcctggcgttgatagcgtcctaaattgaaaaataggatt  gagaagagatattgctagagtaactgtaattcatgggtgtgctgcagcagatgaacgggttaagag  tggcgtgatcacgttagagcaaatccaaccacaaggccttggctgctgtttagaattgtctagcg  taaatgcgggttgcagatgaagtaaatgacgtcgtcatcattcaatttctcagatggttgtgctg  ctgtggtaatcgggtgatgtgaggaagatagagccatagcccaaggtaaagtgttattagggatcat  ctatcgtatttggcgaagatactcaagatggtattgtactgggcattcacgataatggcattacctgc  cagcttcaaggcagttgccagagtattgaaaccgggtaggtccaataattgaaaggttttggct  agccatgccttaactaaagaatctttagcttatggcctgacaccggggcgcacaaggatcatcg  aaaaggcacaaggtctttaggtctatctgatagacaagtggtgattcgtgggaaatcttgagacaa  tatgtaaatatgtgtcctcgagtgtgtattgttctggaaaggatgttacaagaatgtgaacaagatt  ccacagaaatgaagccataacaggatagctttctcttctccaggtgttggcgtagaggggaattt  tattcaaaaggtgtag (SEQ ID NO:29)</p>

Néanmoins, la personne du métier saurait générer d'autres séquences nucléiques optimisées pour une expression dans une levure, notamment *Saccharomyces cerevisiae*, ou dans un autre type de cellule hôte (notamment une autre espèce de levure ou une bactérie).

- Une fois la séquence nucléique optimisée définie, la personne du métier peut obtenir cette
- 5 séquence par synthèse *in vitro* directement avec les codons optimisés. Lorsque le nombre de nucléotides à modifier par rapport à la séquence d'origine n'est pas trop élevé, la séquence optimisée peut aussi être obtenue par mutagénèse dirigée *in vitro* à partir d'un échantillon de la molécule d'acides nucléiques dont les codons sont à adapter, à l'aide d'une amplification par réaction en chaîne par polymérase (PCR).
- 10 Selon un mode de réalisation préféré, la molécule d'acides nucléiques isolée comprend en outre à la fois un promoteur et un terminateur tels que définis ci-dessus, et la séquence d'acides nucléiques comprise dans la molécule d'acides nucléiques isolée utilisée dans l'invention est en outre optimisée pour l'expression dans une cellule hôte, avantageusement dans une levure ou une bactérie, de préférence dans une levure, tel que décrit ci-dessus.
- 15 Selon un autre mode de réalisation préféré, la molécule d'acides nucléiques isolée comprend en outre à la fois un promoteur, un terminateur et une séquence d'export tels que définis ci-dessus, et la séquence d'acides nucléiques comprise dans la molécule d'acides nucléiques isolée utilisée dans l'invention est en outre optimisée pour l'expression dans une cellule hôte, avantageusement dans une levure ou une bactérie, de préférence une levure, tel que
- 20 décrit ci-dessus.

### ***Vecteur utilisé***

Le vecteur utilisé dans l'invention peut être tout vecteur comprenant toute molécule d'acides nucléiques telle que définie ci-dessus.

- 25 Les vecteurs qui sont appropriés dans le cadre de la présente invention comprennent, sans limitation, des vecteurs bactériophages, plasmides ou cosmides pour l'expression dans des cellules hôtes procaryotes telles que des bactéries (par exemple *E. coli*, ou des bactéries du genre *Pseudomonas*); des vecteurs pour l'expression chez la levure (par exemple *Saccharomyces cerevisiae*, *Schyzosaccharomyces pombe*, *Pichia pastoris*); des vecteurs de
- 30 baculovirus pour l'expression dans des systèmes de cellules d'insectes (par exemple, des cellules de Sf 9); des vecteurs viraux et plasmidiques pour l'expression dans des systèmes de cellules végétales (par exemple le plasmide Ti, le virus de la mosaïque du chou-fleur, CaMV,

le virus de la mosaïque du tabac TMV); ainsi que des vecteurs viraux et plasmidiques pour l'expression dans des cellules ou des organismes de vertébrés, notamment de mammifères.

En fonction de l'hôte d'intérêt, la personne du métier connaît des vecteurs d'expression appropriés. Ces vecteurs sont généralement disponibles dans le commerce (par exemple, 5 auprès des fournisseurs tels Invitrogen, Stratagene, Amersham Biosciences, Promega, etc.), disponibles auprès d'institutions de dépôt telles que l'American Type Culture Collection (ATCC, Rockville, Md.), ou ont fait l'objet de nombreuses publications décrivant leur séquence, leurs structures et leurs méthodes de production, de sorte que la personne du métier peut les appliquer sans difficultés.

10 L'invention s'intéresse tout particulièrement à la production de phloroglucinol dans des levures, et le vecteur utilisé dans l'invention est donc avantageusement approprié et même optimisé pour la transfection de levures. Le vecteur utilisé dans l'invention peut notamment être avantageusement un vecteur plasmidique approprié pour la transfection de levures. Des 15 exemples représentatifs de vecteurs plasmidiques appropriés pour la transfection de levures comprennent, sans limitation, pREP4, pCEP4 (Invitrogen), pCI (Promega), pVAX (Invitrogen) et pgWiz (Gene Therapy System Inc) et YCplac22 (ATCC 87585), ainsi que tout dérivé de ces vecteurs (notamment un dérivé de YCplac22 dans lequel ont été intégrés le couple (promoteur PGK et terminateurs CYC) ainsi que le couple (promoteur TEF1 et terminateur ADH1)), mais tout autre vecteur approprié pour transfection de levures peut être utilisé.

20

### ***Cellule hôte utilisée***

La cellule hôte utilisée dans l'invention peut être toute cellule hôte exprimant tout polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini 25 ci-dessus. Elle peut notamment être choisie parmi les cellules hôtes comprenant une molécule d'acides nucléiques isolée telle que définie ci-dessus ou un vecteur tel que défini ci-dessus.

Tel que défini ci-dessus, une cellule hôte est une cellule contenant une molécule d'acides nucléiques hétérologue. Dans le cadre de l'invention, la molécule d'acides nucléiques hétérologue correspond à la molécule d'acides nucléiques isolée telle que définie ci-dessus 30 ou au vecteur tel que défini ci-dessus.

La cellule hôte utilisée peut être une cellule procaryote ou une cellule eucaryote. Parmi les cellules procaryotes, elle peut être choisie parmi les bactéries. Parmi les cellules

eucaryotes, elle peut être choisie parmi les cellules de levure, les cellules de champignons, les cellules d'algues, les cellules d'insectes, de plantes ou de mammifères non humaines.

La cellule hôte utilisée est de préférence une levure, ladite levure étant en particulier sélectionnée parmi les genres *Saccharomyces*, *Candida*, *Eremothecium*, *Dekkera*, *Pichia*  
5 (*Hansenula*), *Debaryomyces*, *Lodderomyces*, *Yarrowia*, *Zigosaccharomyces*,  
*Schizosaccharomyces*, *Torulaspota*, *Kluyveromyces*, *Brettanomyces*, *Cryptococcus* et  
*Malassezia*.

De manière encore plus particulière, la levure est sélectionnée parmi les espèces  
*Saccharomyces cerevisiae*, *Saccharomyces boulardii*, *Saccharomyces douglasii*,  
10 *Saccharomyces bayanus*, *Zigosaccharomyces bailii*, *Schizosaccharomyces pombe*, *Dekkera*  
*brucelensis*, *Dekkera intermedia*, *Brettanomyces custersii*, *Brettanomyces intermedius*,  
*Kluyveromyces themotolerens*, *Torulaspota globosa* et *Torulaspota glabrata*.

Encore plus particulièrement, la levure est du genre *Saccharomyces*, de préférence de  
l'espèce *Saccharomyces cerevisiae*.

15 La cellule hôte utilisée peut aussi être une bactérie, ladite bactérie étant en particulier  
sélectionnée parmi les protéobactéries, les bactéries actinomycètes et les bactéries  
Firmicutes. Parmi les protéobactéries, la cellule hôte peut avantageusement être choisie  
parmi les genres *Escherichia* (notamment parmi les souches de l'espèce *Escherichia coli*,  
espèce la plus utilisée pour la production de protéines recombinantes) et *Pseudomonas*.  
20 Parmi les bactéries actinomycètes, la cellule hôte peut avantageusement être choisie parmi  
les genres *Streptomyces* et *Corynebacterium*. Parmi les bactéries Firmicutes, cellule hôte  
peut avantageusement être choisie parmi les genres *Bacillus* et *Lactobacillus*.

La cellule hôte utilisée comprend au moins une copie de la molécule d'acides nucléiques  
isolée telle que définie ci-dessus intégrée dans son génome. Elle peut notamment  
25 comprendre une copie unique de la molécule d'acides nucléiques isolée telle que définie ci-  
dessus intégrée dans son génome.

Lorsque la cellule hôte utilisée est une cellule de levure (notamment du genre  
*Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*), la ou les copies de la  
molécule d'acides nucléiques peut être intégrée à différents loci, préférentiellement au  
30 locus URA3, au locus JLP1, au locus LEU2, ou au locus TRP1 du génome de ladite cellule de  
levure. Lorsque la cellule hôte est une cellule de levure et que plusieurs copies de la  
molécule d'acides nucléiques sont intégrées, les différentes copies peuvent être intégrées  
au même locus, ou encore à différents loci, préférentiellement à l'une quelconque des  
combinaisons des loci URA3, JLP1, LEU2, et/ou TRP1.

Avantageusement, les codons utilisés dans la molécule d'acides nucléiques codant pour le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries comprise dans la cellule hôte ou dans le vecteur compris dans la cellule hôte ont été adaptés pour une expression optimale dans la cellule hôte sélectionnée. Comme expliqué précédemment, la production d'une séquence d'acides aminés d'intérêt, une expression optimale peut être obtenue lorsque les codons choisis pour coder la séquence d'acides aminés sont ceux utilisés de manière préférentielle par l'organisme d'origine de la cellule hôte. Quel que soit le type de cellule hôte, la personne du métier saura trouver quels codons sont à privilégier dans la littérature ou en utilisant un logiciel d'optimisation des codons. Les exemples de logiciels d'optimisation des codons mentionnés ci-dessus pour la levure peuvent également être utilisés pour d'autres type de cellules hôte. Les levures, et notamment celles du genre *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*, étant tout particulièrement préférées comme cellules hôtes, les codons utilisés dans la molécule d'acides nucléiques codant pour le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries comprise dans la cellule hôte ou dans le vecteur compris dans la cellule hôte ont avantageusement été adaptés pour une expression optimale dans les levures, et notamment dans les levures du genre *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*.

Pour l'utilisation selon l'invention, les cellules hôtes peuvent être cultivées dans des bioréacteurs aérobie ou anaérobie, à petite et grande échelle, dans des flacons ou des boites de Petri. La culture peut être effectuée à une température, un pH, un milieu de culture et une teneur en oxygène appropriés pour une cellule hôte donnée. Pour les levures, et plus particulièrement pour *Saccharomyces cerevisiae*, la culture est avantageusement effectuée par n'importe quelle méthode appropriée. Les procédés de culture d'une souche *Saccharomyces cerevisiae* sont connus dans l'art, et l'homme du métier sait optimiser les conditions de culture pour chaque souche en fonction de sa nature. Des procédés classiques sont notamment décrits dans l'ouvrage de référence "Yeast Technology", 2ème Edition, 1991, Reed and Nagodawithana, publié par Van Nostrand Reinhold (ISBN 0-442-31892-8). Notamment, la culture d'une souche de levure, en particulier de l'espèce *Saccharomyces cerevisiae*, peut généralement être réalisée à une température comprise entre 20 et 37°C, dans un milieu liquide riche (par exemple le milieu YPD accessible auprès de VWR) ou synthétique (défini pour répondre précisément aux besoins de la souche), en culture aérobie ou anaérobie.

### **Utilisation**

Le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, la molécule d'acides nucléiques codant pour celui-ci, le vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou la cellule hôte exprimant celui-ci, chacun étant tel que défini ci-dessus, est utilisé pour produire du phloroglucinol, quelle que soit la méthode de production utilisée.

Selon la méthode, la personne du métier saura choisir le plus approprié entre le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, la molécule d'acides nucléiques codant pour celui-ci, le vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou la cellule hôte exprimant celui-ci, chacun étant tel que défini ci-dessus.

Notamment, le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus peut être utilisé directement pour produire du phloroglucinol in vitro, à condition de le mettre en contact avec du malonyl-CoA. En effet, les phloroglucinol synthases réalisent la condensation de trois molécules de malonyl-CoA pour former une molécule de phloroglucinol selon le schéma réactionnel Réaction I précédemment décrit.

La molécule d'acides nucléiques codant pour celui-ci et le vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci peuvent être utilisés pour transférer une cellule hôte, qui peut ensuite être utilisée soit pour produire le polypeptide (qui peut alors produire du phloroglucinol in vitro) soit pour produire directement du phloroglucinol lorsqu'elle est cultivée en présence d'un substrat approprié (tel qu'une source de carbone comme le glucose ou l'éthanol).

Des exemples de méthodes plus précises dans lesquelles le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, la molécule d'acides nucléiques codant pour celui-ci, le vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou la cellule hôte exprimant celui-ci tels que décrits ci-dessus sont décrits dans la section ci-dessous concernant des méthodes selon l'invention pour produire du phloroglucinol.

### **30 Molécule d'acides nucléiques isolée selon l'invention**

L'invention concerne également une molécule d'acides nucléiques isolée comprenant une séquence d'acides nucléiques codant pour un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus, caractérisée en ce que :

- a) la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques ; ou
- b) la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques ; ou
- 5 c) la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, notamment dans une levure ou une bactérie ; ou
- d) toute combinaison de a) à c).

De telles molécules d'acides nucléiques isolées sont utiles pour soit pour la production d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini

10 ci-dessus, soit pour produire directement du phloroglucinol lorsqu'elles sont transfectées (directement ou au sein d'un vecteur) dans une cellule hôte, cultivée en présence d'un substrat approprié.

Dans le mode de réalisation a), la molécule d'acides nucléiques isolée comprend en outre

15 un promoteur contrôlant l'expression de la séquence d'acides nucléiques.

Dans le mode de réalisation b), la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques.

Dans le mode de réalisation c), la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, notamment dans une levure ou dans une bactérie,

20 avantageusement dans une levure.

Dans d'autres modes de réalisation, la molécule d'acides nucléiques isolée peut combiner plusieurs des caractéristiques a) à c). En particulier, la molécule d'acides nucléiques isolée peut être caractérisée en ce que :

d) la molécule d'acides nucléiques isolée comprend en outre un promoteur et un terminateur

25 de transcription contrôlant l'expression de la séquence d'acides nucléiques (combinaison de a) et b)) ;

e) la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques et la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, notamment dans une levure ou dans

30 une bactérie, avantageusement dans une levure (combinaison de a) et c)) ;

f) la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques et la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, notamment dans

une levure ou dans une bactérie, avantageusement dans une levure (combinaison de b) et c)) ; ou

- g) la molécule d'acides nucléiques isolée comprend en outre un promoteur et un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques et la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, notamment dans une levure ou dans une bactérie, avantageusement dans une levure (combinaison de a), b) et c)).

Lorsque la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques (modes de réalisation a), d), e), et g) ci-dessus), le promoteur est avantageusement un promoteur exogène, en particulier un promoteur de levure, et de préférence un promoteur choisi parmi ADH2 (pADH2) et CCW12 (pCCW12) et TEF1 (pTEF1), de préférence encore un promoteur choisi parmi ADH2 (pADH2) de *Saccharomyces cerevisiae* et CCW12 de *S. cerevisiae*, CCW12 (pCCW12) de *S. cerevisiae* et TEF1 (pTEF1) de *S. cerevisiae* de préférence encore un promoteur choisi parmi ADH2 de séquence SEQ ID NO: 20, CCW12 de séquence SEQ ID NO: 21 et TEF1 de séquence SEQ ID NO : 22.

Lorsque la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques (modes de réalisation b), d), f), et g) ci-dessus), le terminateur de transcription est avantageusement un terminateur exogène, tel qu'un terminateur de levure, et de préférence le terminateur RPL3 (tRPL3) ou le terminateur ADH1 (tADH1), de préférence encore le terminateur RPL3 de *S. cerevisiae* ou le terminateur ADH1 de *S. cerevisiae*, de préférence encore le terminateur RPL3 de séquence SEQ ID NO: 23 ou le terminateur de ADH1 de séquence SEQ ID NO : 24.

Lorsque la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une levure (modes de réalisation c), e), f), et g) ci-dessus), elle est avantageusement optimisée pour l'expression dans une levure du genre *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*. En particulier, elle est avantageusement choisie parmi SEQ ID NO :25 à SEQ ID NO :29.

Lorsque la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une bactérie (modes de réalisation c), e), f), et g) ci-dessus), elle est avantageusement optimisée pour l'expression dans une bactérie des genres *Escherichia* (notamment des souches de l'espèce *Escherichia coli*, espèce la plus utilisée pour la production de protéines recombinantes) et *Pseudomonas*.

**Vecteur selon l'invention**

L'invention concerne également un vecteur comprenant une molécule d'acides nucléiques selon l'invention.

5 Le vecteur selon l'invention peut comprendre toute molécule d'acides nucléiques selon l'invention décrite ci-dessus. Il peut être choisi parmi tout type de vecteur décrit précédemment dans la section visant les vecteurs susceptibles d'être utilisés dans l'utilisation selon l'invention.

**Cellule hôte selon l'invention**

10 L'invention concerne également une cellule hôte comprenant une molécule d'acides nucléiques selon l'invention, ou un vecteur selon l'invention.

La cellule hôte selon l'invention peut comprendre toute molécule d'acides nucléiques selon l'invention décrite ci-dessus ou tout vecteur selon l'invention tel que décrit ci-dessus. Elle peut en outre comprendre toute caractéristique ou combinaison de caractéristique décrite  
15 précédemment dans la section visant les cellules hôtes susceptibles d'être utilisées dans l'utilisation selon l'invention.

**Méthodes pour produire un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries**

20 La présente invention concerne en outre une méthode pour produire un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus.

Selon un mode de réalisation, la méthode pour produire un tel polypeptide comprend, est essentiellement constituée, ou est constituée des étapes suivantes :

25 (i) l'introduction d'une molécule d'acides nucléiques ou d'un vecteur comprenant une molécule d'acides nucléiques tels que décrits ci-dessus (molécule d'acides nucléiques ou vecteur utilisés dans l'invention ou selon l'invention) dans une cellule hôte appropriée conforme à la description qui précède ; et

(ii) la culture *in vitro* de ladite cellule hôte obtenue à l'étape (i) dans des conditions  
30 permettant l'expression de ladite molécule d'acides nucléiques, de manière à produire ledit polypeptide.

Selon un autre mode de réalisation, la méthode pour produire un tel polypeptide comprend, est essentiellement constituée, ou est constituée au moins de l'étape consistant en :

- 5 (ii) la culture *in vitro* d'une cellule hôte exprimant ledit polypeptide, par exemple une cellule hôte telle que décrite ci-dessus (cellule hôte utilisée dans l'invention ou selon l'invention), dans des conditions permettant l'expression de la molécule d'acides nucléiques contenue dans ladite cellule hôte, de manière à produire ledit polypeptide.

10 L'étape (i) d'introduction de la molécule d'acides nucléiques ou du vecteur dans une cellule hôte appropriée peut être réalisée par toute méthode appropriée connue de la personne du métier, telle que la transfection par phosphate de calcium, la transfection par des liposomes comprenant la molécule d'acides nucléiques ou le vecteur à transférer, la transfection par des agents polycationiques, l'électroporation, et la transfection par choc thermique. Notamment, l'étape (i) peut être réalisée par transfection (par toute méthode décrite ci-dessus) dans la cellule hôte d'un plasmide de transfert comprenant la séquence codant pour  
15 le polypeptide choisi parmi les polykétide synthases de type III de champignons Ascomycètes tel que défini ci-dessus, encadrée en 5' et en 3' par des séquences homologues à des séquences génomiques de la cellule hôte, permettant ainsi une recombinaison homologue entre le plasmide de transfert et le génome de la cellule hôte. Cette méthode est notamment applicable lorsque la cellule hôte est une levure, en particulier du  
20 *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*, ou une bactérie.

L'étape (ii) de culture *in vitro* de la cellule hôte est réalisée dans des conditions permettant l'expression de la molécule d'acides nucléiques contenue dans ladite cellule hôte, de manière à produire ledit polypeptide. Ces conditions varient en fonction de la cellule hôte utilisée et la personne du métier saura les déterminer sur la base de ses connaissances  
25 générales. Notamment, lorsque la cellule hôte est une levure, en particulier du *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*, l'étape (ii) peut être réalisée par toute méthode appropriée, notamment tel que décrit dans la section ci-dessus concernant les cellules hôtes utilisées dans l'utilisation selon l'invention.

30 La méthode pour produire le polypeptide peut en outre comprendre au moins une étape additionnelle choisie parmi les étapes suivantes :

- (α) la récupération des cellules hôtes exprimant ledit polypeptide et/ou du surnageant comprenant le polypeptide, obtenus après l'étape de culture ; et

(B) la purification du polypeptide à partir des cellules hôtes et/ou du surnageant récupérées à l'étape (α).

L'étape optionnelle (α) de récupération des cellules hôtes exprimant ledit polypeptide et/ou du surnageant comprenant le polypeptide peut être réalisée par toute technique appropriée connue de la personne du métier. Les cellules et le surnageant peuvent notamment être séparés par décantation ou centrifugation.

L'étape optionnelle (β) de purification du polypeptide à partir des cellules hôtes et/ou du surnageant récupérées à l'étape (α) peut être réalisée par toute technique appropriée connue de la personne du métier. Lorsque le polypeptide est purifié à partir du surnageant, toute technique de purification de protéine en phase liquide peut être utilisée, comme par exemple les différents types de chromatographie (filtration sur gel, échange d'ions, par interaction hydrophobe, par affinité lorsque la protéine comprend une étiquette d'affinité, chromatographie liquide haute performance « HPLC »). Lorsque le polypeptide est purifié à partir des cellules ou des cellules et du surnageant, la purification comprend en outre une ou plusieurs étape(s) préalable(s) de lyse des cellules et éventuellement d'élimination des débris cellulaire, avant l'utilisation d'une technique de purification en phase liquide.

### **Méthodes pour produire du phloroglucinol**

Comme expliqué dans la sous-section « Utilisation » de la section visant l'utilisation selon l'invention, différentes méthodes de production de phloroglucinol peuvent être mises en œuvre selon qu'elles reposent sur le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, la molécule d'acides nucléiques codant pour celui-ci, le vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou la cellule hôte exprimant celui-ci, chacun étant tel que défini ci-dessus.

Notamment, la molécule d'acides nucléiques codant pour le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus et le vecteur comprenant une molécule d'acides nucléiques codant pour ce polypeptide peuvent être utilisés pour transférer une cellule hôte, qui peut ensuite être utilisée soit pour produire le polypeptide (qui peut alors produire du phloroglucinol in vitro) soit pour produire directement du phloroglucinol lorsqu'elle est cultivée en présence d'un substrat approprié.

Ainsi, selon un mode de réalisation M1, la méthode de production de phloroglucinol utilise une cellule hôte exprimant le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus (par exemple une cellule hôte selon l'invention),

qui est cultivée en présence d'un substrat appropriée pour produire du phloroglucinol. Ce mode de réalisation peut être subdivisé en deux sous-modes de réalisation, selon que la méthode comprend (mode de réalisation M1A) ou non (mode de réalisation M1B) une étape préalable d'obtention des cellules hôtes.

5 Dans le mode de réalisation M1A, l'invention concerne une méthode pour produire du phloroglucinol qui comprend, est essentiellement constituée, ou est constituée des étapes suivantes :

(i1) l'introduction d'une molécule d'acides nucléiques ou d'un vecteur comprenant  
une molécule d'acides nucléiques tels que décrits ci-dessus (molécule d'acides  
10 nucléiques ou vecteur utilisés dans l'invention ou selon l'invention) dans une cellule  
hôte appropriée conforme à la description qui précède ;

(ii1) la mise en contact des cellules hôtes obtenues à l'étape (i1) avec un substrat  
approprié ;

(iii1) la culture *in vitro* de la cellule hôte de l'étape (ii1) dans des conditions  
15 permettant l'expression de la molécule d'acides nucléiques contenue dans ladite  
cellule hôte, de manière à produire du phloroglucinol ;

(iv1) optionnellement la récupération du milieu de culture comprenant le  
phloroglucinol, obtenu après l'étape (iii1) ; et

(v1) optionnellement, la purification du phloroglucinol à partir du milieu de culture  
20 de l'étape (iv1).

Dans le mode de réalisation M1B, l'invention concerne une méthode pour produire du  
phloroglucinol, qui comprend, est essentiellement constituée, ou est constituée des étapes  
suivantes :

(ii1) la mise en contact d'une cellule hôte non humaine exprimant le polypeptide  
25 choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini  
ci-dessus (par exemple une cellule hôte selon l'invention), avec un substrat  
approprié ;

(iii1) la culture *in vitro* de la cellule hôte de l'étape (ii1) dans des conditions  
30 permettant l'expression de la molécule d'acides nucléiques contenue dans ladite  
cellule hôte, de manière à produire du phloroglucinol ;

(iv1) optionnellement la récupération du milieu de culture comprenant le  
phloroglucinol, obtenu après l'étape (iii1) ; et

(v1) optionnellement, la purification du phloroglucinol à partir du milieu de culture  
de l'étape (iv1).

Aux fins des méthodes M1A et M1B, et notamment lorsque la cellule hôte est une levure (en particulier du genre *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*) :

- 5 - l'étape (i1) d'introduction de la molécule d'acides nucléiques ou du vecteur dans une cellule hôte appropriée peut être réalisée par toute méthode décrite ici dans le contexte de l'étape (i) des méthodes de production du polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus.
- 10 - le substrat mise en contact avec la cellule hôte à l'étape (ii1) est avantageusement une source de carbone. Avantageusement, la source de carbone est une source de carbone pure ou un coproduit industriel (tel que la mélasse ou les égouts pauvres, par exemple issus de l'industrie sucrière). De préférence, le substrat dans la source de carbone pure ou le coproduit industriel est un sucre simple, tel que le glucose (ou dextrose), le fructose, le galactose, le mannose, le saccharose, le lactose, ou le maltose ; un sucre complexe, tel qu'un monosaccharide, un disaccharide ou trisaccharides, ou encore un polysaccharide comme l'amidon ; un alcool, tel que l'éthanol ; un acide ; un acide gras et un dérivé d'ester de celui-ci ; ou un mélange de sucres, d'alcools, d'acides et/ou d'acides gras ou leurs dérivés d'ester. De préférence, le substrat est le glucose ou l'éthanol.
- 15 - l'étape (iii1) de culture *in vitro* de la cellule hôte dans des conditions permettant l'expression de la molécule d'acides nucléiques contenue dans ladite cellule hôte, de manière à produire du phloroglucinol peut être réalisée par toute méthode décrite ici dans le contexte de l'étape (ii) des méthodes de production du polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus.
- 20 - lorsqu'elle est présente, l'étape (iv1) de récupération du milieu de culture comprenant le phloroglucinol peut être réalisée par toute technique décrite ici pour la séparation des cellules et du surnageant dans le contexte de l'étape optionnelle ( $\alpha$ ) des méthodes de production du polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus.
- 25 - lorsqu'elle est présente, l'étape (v1) de purification du phloroglucinol à partir du milieu de culture peut être réalisée par toute technique appropriée, telle que l'extraction liquide-liquide.

30

De manière alternative, le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus peut être utilisé directement pour produire du phloroglucinol *in vitro*, à condition de le mettre en contact avec du malonyl-CoA. En effet,

les phloroglucinol synthases réalisent la condensation de trois molécules de malonyl-CoA pour former une molécule de phloroglucinol selon le schéma réactionnel Réaction I précédemment décrit.

Ainsi, selon un mode de réalisation M2, la méthode de production de phloroglucinol utilise le polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que décrit ci-dessus pour produire du phloroglucinol in vitro en présence de malonyl-CoA. Ce mode de réalisation peut être subdivisé en deux sous-modes de réalisation, selon que la méthode comprend (mode de réalisation M2A) ou non (mode de réalisation M2B) une étape préalable de production du polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus.

Dans le mode de réalisation M2A, l'invention concerne méthode pour produire du phloroglucinol, qui comprend, est essentiellement constituée, ou est constituée des étapes suivantes :

(i2) la production d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus par l'une des méthodes de production de polypeptide décrite ci-dessus ;

(ii2) la mise en contact du polypeptide obtenu à l'étape (i2) de la méthode telle que décrite ci-dessus, avec du malonyl-CoA ;

(iii2) l'incubation du mélange issu de l'étape (ii2) dans des conditions adaptées pour produire du phloroglucinol ;

(iv2) optionnellement la récupération du milieu réactionnel comprenant le phloroglucinol, obtenu après l'étape (iii2) ; et

(v2) optionnellement, la purification du phloroglucinol à partir du milieu réactionnel de l'étape (iv2).

Dans le mode de réalisation M2B, l'invention concerne une méthode pour produire du phloroglucinol, qui comprend, est essentiellement constituée, ou est constituée des étapes suivantes :

(ii2) la mise en contact d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini ci-dessus avec du malonyl-CoA ;

(iii2) l'incubation du mélange issu de l'étape (ii2) dans des conditions adaptées pour produire du phloroglucinol ;

(iv2) optionnellement la récupération du milieu réactionnel comprenant le phloroglucinol, obtenu après l'étape (iii2) ; et

(v2) optionnellement, la purification du phloroglucinol à partir du milieu réactionnel de l'étape (iv2).

Aux fins de ces méthodes M2A et M2B, et notamment lorsque la cellule hôte est une levure (en particulier du genre *Saccharomyces*, de préférence de l'espèce *Saccharomyces cerevisiae*) :

- l'étape (i2) peut être réalisée par toute méthode décrite ici pour produire un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries.

- l'étape (ii2) consiste à mettre en contact le polypeptide avec son substrat le malonyl-CoA pour permettre la production de phloroglucinol. Cette mise en contact est faite dans un milieu et à une température qui n'altèrent pas l'activité enzymatique du polypeptide. Un milieu approprié peut notamment être choisi parmi un tampon phosphate de sodium à pH 7 (notamment à une concentration d'environ 25-75 mM) comprenant du NaCl (notamment à une concentration d'environ 5-15 mM) et de l'albumine sérique bovine (ASB, notamment à une concentration d'environ 150-250 mg/L). Une température appropriée se situe dans la gamme 20 à 37°C.

- l'étape (iii2) consiste à incuber le polypeptide et son substrat dans des conditions permettant la production de phloroglucinol. Ces conditions impliquent un milieu et une température d'incubation qui n'altèrent pas l'activité enzymatique du polypeptide, tel que décrit ci-dessus pour l'étape (ii2). La durée d'incubation est choisie en fonction de la quantité de substrat présente dans le milieu. La personne du métier saura adapter la concentration initiale de substrat à la durée d'incubation. Il est peut aussi être envisagé d'ajouter du substrat pendant la durée de l'incubation pour continuer la production de phloroglucinol.

- lorsqu'elle est présente, l'étape (iv2) de récupération du milieu réactionnel ne nécessite aucune technique particulière puisqu'il n'y a ici aucune cellule à exclure ou lyser avant purification.

- lorsqu'elle est présente, l'étape (v2) de purification du phloroglucinol à partir du milieu de culture peut être réalisée par toute technique appropriée, telle que l'extraction liquide-liquide.

Les exemples qui suivent visent à illustrer la présente invention.

## EXEMPLES

**Exemple 1. Identification de polykétide synthases de type III de cyanobactéries**

Les PhlD sont des polykétides synthases (PKS) de type III impliquées dans la synthèse biocatalytique du phloroglucinol à partir de trois molécules de molécules de malonyl-CoA (Zha et al., 2006). Pour identifier de nouvelles PhlD par une approche basée sur les séquences, nous avons utilisé la séquence de la PhlD de *Tsukamurella pulmonis* (PhlD\_Tpu). Cette dernière s'est en effet avérée être la meilleure enzyme produisant du phloroglucinol parmi les 12 enzymes précédemment testées dans *S. cerevisiae*, l'organisme hôte utilisé pour le clonage et la production de protéines (WO2019/002799).

Un blast (Johnson et al., 2008) [NCBI] de la séquence protéique de PhlD\_Tpu a été effectué contre les bases de données protéiques standards (séquences protéiques non-redondantes, *protein databank*, Swissprot, protéines issues de métagénomés et séquences protéiques brevetées, versions du 22.11.2019). Les séquences obtenues ont été filtrées en collectant toutes les séquences ayant un pourcentage d'identité supérieure à 20% par rapport à la séquence d'entrée. Les séquences redondantes ont ensuite été supprimées (CD-HIT, identité de séquence >95%), réduisant le jeu de données à environ 3 320 séquences. De plus, les séquences de 45 PKS caractérisées provenant de la littérature ont été incluses si elles n'étaient pas présentes dans l'ensemble de données générées précédemment. Les séquences ont enfin été regroupées en utilisant un réseau de similarité de séquence (« SSN » pour « *Sequence Similarity Networks* », Gerlt, *et al.*, 2015), chaque groupe correspondant à un embranchement ou à une classe différente. Les réseaux SSN sont utiles pour analyser les ensembles de données de séquences et étudier les relations séquence-fonction. Le SSN est un moyen pratique de visualiser les relations entre les séquences de protéines via la construction d'un graphe où chaque membre d'une famille de protéines est représenté par un nœud relié par une arête aux nœuds de tous les autres membres qui partagent une similarité de séquence supérieure à une valeur spécifiée par l'utilisateur. Ces réseaux de graphes sont faciles à visualiser et sont plus aisément interprétables que les approches traditionnelles des dendrogrammes et des arbres phylogénétiques. Cette représentation permet d'attribuer plus facilement une fonction putative à une protéine d'un cluster donné si un membre a déjà été caractérisé et de suivre les relations évolutives entre les clusters.

Ces graphes peuvent être visualisés à l'aide d'un logiciel de visualisation tel que Cytoscape (Shannon et al., 2003), où des sous-régions des graphes (arêtes et nœuds) peuvent être agrandies ou manipulées de manière itérative, permettant ainsi de mieux identifier la similarité des séquences à l'échelle microscopique du réseau.

Après analyse minutieuse des clusters, nous avons identifié des séquences PKS déjà caractérisées comme produisant du phloroglucinol ou d'autres composés apparentés connus

pour être produits par des enzymes de la famille PKS. La majorité de ces enzymes se retrouvent dans les groupes des plantes, des algues brunes et des actinobactéries. D'autres enzymes ont toutefois également été caractérisées dans d'autres clusters isolés. Les séquences *P. fluorescens* PhlD et PhlD\_Tpu de *P. fluorescens* connues pour produire du phloroglucinol ont également été identifiées, ce qui valide la méthode utilisée.

Pour approfondir la classification des séquences et leurs interconnexions, des clusters de séquences provenant de différentes classes ont été générés avec une valeur  $E < 10^{-130}$ . Cela a permis de retrouver deux PKS de type III provenant de macroalgues brunes (*Ectocarpus siliculosus*-PKS1.Es et *Sargassium binderi*-PHLD.Sbi) déjà signalées comme produisant du phloroglucinol dans *Saccharomyces cerevisiae* (WO2019/002799). L'utilisation d'un score seuil plus strict (valeur  $E < 10^{-170}$ ) a permis de séparer un cluster contenant des PKS de type III du genre *Tsukamurella*, également connues pour avoir une forte activité de production de phloroglucinol dans *S. cerevisiae* (WO2019/002799). Ces résultats valident encore une fois la méthode utilisée.

Enfin, nous avons regroupé les informations issues des deux analyses de clustering avec une valeur  $E < 10^{-130}$  et une valeur  $E < 10^{-170}$  et avons sélectionné au total 23 séquences issues de sept clusters et de différentes classes pour évaluer leur capacité à produire du phloroglucinol.

### **Exemple 2. Séquences de cinq polykétide synthases de type III de cyanobactéries identifiées à l'Exemple 1**

Les polypeptides de séquences d'acides aminés SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5 font partie des 23 séquences sélectionnées à l'Exemple 1 et sont toutes cinq issues de cyanobactéries.

Plus particulièrement :

- SEQ ID NO :1 correspond à la séquence d'acides aminés d'une PhlD de *Rivularia sp. PCC 7116* (dénotée PlhD-Rs) ;
- SEQ ID NO :2 correspond à la séquence d'acides aminés d'une PhlD de *Chamaesiphon minutus* (dénotée PlhD-Cm) ;
- SEQ ID NO :3 correspond à la séquence d'acides aminés d'une PhlD de *Filamentous cyanobacterium CCP2* (dénotée PlhD-Fc) ;
- SEQ ID NO :4 correspond à la séquence d'acides aminés d'une PhlD de *Calothrix sp. HK-06* (dénotée PlhD-Cs) ; et

- SEQ ID NO :5 correspond à la séquence d'acides aminés d'une PhLD de *Gloeocapsa sp. PCC 7428* (dénotée PhLD-Gs).

Un alignement multiple de leurs séquences d'acides aminés est présenté sur la **Figure 1**.

- 5 Leurs pourcentages d'identité deux à deux, ainsi que vis-à-vis d'autres PhLD de l'art antérieur sont en outre présentés dans le **Tableau 3** ci-dessous.

[Table 3]

	Référence dans les bases de données	Espèce	Rs	Cm	Fc	Cs	Gs
PHLD.Rs	WP_015119976.1	<i>Rivularia sp. PCC 7116</i>	100,0	58,2	53,0	54,7	61,2
PHLD.Cm	WP_015159767.1	<i>Chamaesiphon minutus</i>	58,2	100,0	50,2	55,8	57,1
PHLD.Fc	WP_106169877.1	<i>Filamentous cyanobacterium CCP2</i>	53,0	50,2	100,0	52,0	57,1
PHLD.Cs	WP_073620125.1	<i>Calothrix sp. HK-06</i>	54,7	55,8	52,0	100,0	60,0
PHLD.Gs	WP_015188686.1	<i>Gloeocapsa sp. PCC 7428</i>	61,2	57,1	57,1	60,0	100,0
PKS1.Es	CBN76919.1	<i>Ectocarpus siliculosus</i>	38,1	38,4	39,1	36,3	40,9
PHLD.Pf	AAY95147.1	<i>Pseudomonas fluorescens Pf-5</i>	22,8	25,9	24,1	24,1	22,6
PHLD.Tp	WP_013126955.1	<i>Tsukamurella paurometabola</i>	39,7	41,5	44,1	41,1	43,5
PHLD.Tt	WP_068525790.1	<i>Tsukamurella tyrosinosolvans</i>	38,7	41,6	43,2	39,4	41,7
PHLD.Tps	WP_068569959.1	<i>Tsukamurella pseudospumae</i>	38,7	42,5	44,1	41,4	43,7
PHLD.Tpu	WP_068567598.1	<i>Tsukamurella pulmonis</i>	39,7	42,1	44,1	41,0	42,9

PHLD.Tsp	WP_019202763.1	<i>Tsukamurella</i> <i>sp. 1534</i>	37,9	41,4	44,9	41,3	43,4
PHLD.Nf	WP_011211464.1	<i>Nocardia</i> <i>farcinica</i>	44,2	45,2	46,7	43,5	45,4
PHLD.Mma	WP_012393954.1	<i>Mycobacterium</i> <i>marinum</i>	45,0	45,3	45,3	44,7	46,0
PHLD.Mk	ZP_04749411.1	<i>Mycobacterium</i> <i>kansasii</i> ATCC 12478	42,4	44,1	46,0	44,4	46,9
PHLD.Mt	WP_031677738.1	<i>Mycobacterium</i> <i>tuberculosis</i>	35,5	34,9	36,6	34,4	36,9
PHLD.Gh	WP_066172590.1	<i>Gordonia</i> <i>hydrophobica</i>	39,9	45,4	44,2	44,1	42,6
PHLD.Sbi	ADK13089.1	<i>Sargassum</i> <i>binderi</i>	38,7	37,7	39,5	34,8	41,3
PHLD-1.Aa	SEQ ID NO :3 de WO2019/002798	<i>Aureococcus</i> <i>anophagefferens</i>	38,1	39,2	40,1	38,6	40,0

Le Tableau 3 ci-dessus montre que les 5 séquences nouvellement identifiées comme de possibles phloroglucinol synthases n'ont qu'une faible identité de séquence avec les séquences des phloroglucinol synthases précédemment identifiées, et appartiennent à un 5 règne différent (cyanobactéries versus bactéries actinomycètes ou algues eucaryotes).

**Exemple 3. Capacité de polykétide synthases de type III de cyanobactéries à produire du phloroglucinol chez la levure**

Les 5 phloroglucinol synthases putatives précédemment identifiées (SEQ ID NO :1, SEQ ID 10 NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5) ont été testées pour leur capacité à produire du phloroglucinol lorsqu'elles sont exprimées par une levure de l'espèce *Saccharomyces cerevisiae*. Leur activité a en outre été comparée à celle de PHLD de l'art antérieur.

15 Matériels et méthodes

*Cellules de levure*

Des cellules de la levure *S. cerevisiae* CEN.PK2-1D (voir Entian KD and Kötter P, 2007, commercialement accessible notamment chez Euroscarf) ont été utilisées.

*Préparation de clones cellulaires de S. cerevisiae CEN.PK2-1D exprimant les 5 phloroglucinol synthases putatives ou des phloroglucinol synthases de l'art antérieur*

- 5 Les séquences nucléiques SEQ ID NO :25 à 29 codant respectivement pour les 5 phloroglucinol synthases putatives précédemment identifiées (SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5) ont été modifiées en ajoutant à leurs extrémités 5' et 3' respectivement des séquences homologues au site souhaité d'intégration dans le génome des cellules de levure *S. cerevisiae* CEN.PK2-1D, conduisant aux séquences SEQ ID NO :30 à
- 10 34. Pour les phloroglucinol synthases PHL.D.Pf, PHL.D.Tsp et PHL.D.Tt, les séquences SEQ ID NO :35 à 37, qui comprennent des séquences homologues au site souhaité d'intégration dans le génome, ont été utilisées.

- Les séquences SEQ ID NO :30 à 37 ont ensuite été clonées dans un plasmide appelé « PBIM5 » ou « YCplac22\_PGK\_CYC\_TEF\_ADH » (dérivé du plasmide commercial YCplac22 par ajout du
- 15 couple (promoteur PGK et terminateur CYC, voir Gietz RD et al., 1988) ainsi que le couple (promoteur TEF1 et terminateur ADH1), voir Figure 2), sous contrôle du promoteur TEF1 et du terminateur ADH1. Les séquences SEQ ID NO :30 à 37 ont ensuite été intégrées dans le génome de cellules de la levure *S. cerevisiae* CEN.PK2-1D par recombinaison homologue in vivo, et les séquences clonées vérifiées par séquençage Sanger.

- 20 *Culture en plaque 96 puits pour la production de phloroglucinol*

Chaque puits a été rempli avec 500 µl de milieu synthétique complet Trp\_drop out commercialisé par Formedium.

Un clone cellulaire unique obtenu tel que décrit ci-dessus a été prélevé et inoculé dans un puits.

- 25 La plaque a été incubée à 30°C pendant environ 20 heures à 800 tours par minute dans un incubateur vibrant. Le jour suivant, une nouvelle plaque de 96 puits remplie de 500 µl de milieu synthétique complet a été inoculée avec 10 µl de culture. Trois plaques individuelles ont été inoculées pour des échantillonnages de 24, 48 et 72 heures.

*Quantification de la production de phloroglucinol par test colorimétrique*

- 30 Toutes les 24 heures, chaque plaque a été retirée de l'incubateur et les mesures ont été effectuées comme décrit ci-dessous :

-Tout d'abord, l'absorbance a été mesurée dans la plaque de microtitration à 600 nm dans le lecteur de plaque Tecan. Pour cela, 20 µl de culture ont été dilués avec 180 µl d'eau milliQ.

-Ensuite, la plaque de 96 puits a été centrifugée à 3700 rpm pendant 10 minutes et le surnageant a été utilisé pour le test colorimétrique.

- 5 -150 µl de 500mg/L de 4-hydroxy-3méthoxy-cinnamaldéhyde dans HCl/EtOH v:v (1:3) ont été ajoutés à l'échantillon et au standard Phloroglucinol. L'incubation a été réalisée à température ambiante pendant 30min.

### Résultats

- 10 Les quantités de phloroglucinol produites dans le test utilisé par les levures exprimant les différentes PhID sont présentées dans le **Tableau 4** ci-dessous.

[Table 4]

PhID	Production de phloroglucinol après 72h
PHLD.Rs	19,5 mg/L
PHLD.Cm	18,2 mg/L
PHLD.Fc	10,7 mg/L
PHLD.Cs	7,2 mg/L
PHLD.Gs	7,1 mg/L
PHLD.Pf	Pas de production de phloroglucinol
PhID.Tsp	47,5 mg/L
PhID.Tt	37,9 mg/L

- 15 Le **Tableau 4** montre que les 5 phloroglucinol synthases putatives précédemment identifiées (SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5), bien qu'issues de cyanobactéries, sont capables de produire du phloroglucinol lorsqu'elles sont exprimées dans une levure.

### REFERENCES BIBLIOGRAPHIQUES

Achkar J et al., (2005) « Biosynthesis of phloroglucinol » J Am Chem Soc. 127:5332-5333.

Entian K-D,e et al., (2007) « Yeast genetic strain and plasmid collections », *Methods in microbiology*, 36, 0580-9517.

Gietz R.D et al., (1988) « New yeast-escherichia coli shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites », *Gene*, 74, 527-534.

- 5 Gerlt, J. A. *et al.* Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta BBA - Proteins Proteomics* 1854, 1019-1037 (2015).

Johnson, M. et al. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, W5-W9 (2008).

- 10 Meslet-Cladière L, Delage L, Leroux CJ, Goultitquer S, Leblanc C, Creis E, Gall EA, Stiger-Pouvreau V, Czjzek M, and Potin P. (2013) “Structure/function analysis of a type III polyketide synthase in the brown alga *Ectocarpus siliculosus* reveals a biochemical pathway in phlorotannin monomer biosynthesis.” *Plant Cell.* 25:3089-3103.

Needleman et Wunsch. *J.Mol. Biol.* 48,443-453, 1970.

- 15 Shannon, P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498-2504 (2003).

WO2013/045510

WO2019/002798

WO2019/002799

WO2020/024917A1

- 20 Zha W, Rubin-Pitel SB and Zhao H. (2006) “Characterization of the substrate specificity of PHLD, a type III polyketide synthase from *Pseudomonas fluorescens*.” *J Biol Chem.* 281:32036-32047.

## REVENDICATIONS

1. Utilisation d'un polypeptide choisi parmi les polykétide synthases de type III de cyanobactéries, d'une molécule d'acides nucléiques codant pour celui-ci, d'un vecteur comprenant une molécule d'acides nucléiques codant pour celui-ci ou d'une cellule hôte non humaine exprimant celui-ci, pour produire du phloroglucinol ; l'utilisation étant caractérisée en ce que ledit polypeptide comprend une séquence d'acides aminés ayant au moins 70% d'identité avec une séquence choisie parmi SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5.
2. Utilisation selon la revendication 1, caractérisée en ce que ledit polypeptide est choisi parmi les polykétide synthases de type III :
- des familles *Rivulariaceae*, *Chamaesiphonaceae* et *Chroococcaceae*, avantageusement des genres *Rivularia*, *Calothrix*, *Chamaesiphon* et *Gloeocapsa*, plus avantageusement des espèces *Rivularia sp. PCC 7116*, *Calothrix sp. HK-06*, *Chamaesiphon minutus*, *Gloeocapsa sp. PCC 7428*, et *Gloeocapsa sp. PCC 73106* ; ou
  - de l'espèce *Filamentous cyanobacterium CCP2*.
3. Utilisation selon la revendication 1 ou la revendication 2, caractérisée en ce que ledit polypeptide comprend une séquence d'acides aminés ayant au moins 80% d'identité avec une séquence choisie parmi SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5.
4. Utilisation selon l'une quelconque des revendications 1 à 3, caractérisée en ce que ledit polypeptide comprend en outre au moins un fragment choisi parmi :
- (a) ATGTP (SEQ ID NO:6) ;
  - (b) STGFX<sub>1</sub>APG, dans laquelle X<sub>1</sub> est choisi parmi les acides aminés apolaires (i.e. glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M) et cystéine (C), SEQ ID NO :7);
  - (c) VX<sub>2</sub>FMGCAAA, dans laquelle X<sub>2</sub> est choisi parmi les acides aminés polaires (i.e. sérine (S), thréonine (T), asparagine (N), glutamine (Q), acide glutamique (E), acide aspartique (D), lysine (K), arginine (R), et histidine (H), SEQ ID NO:8) ;
  - (d) VCLELSS (SEQ ID NO :9) ;
  - (e) IHSIFX<sub>3</sub>DGCAA, dans laquelle X<sub>3</sub> est n'importe quel acide aminé (i.e. X<sub>3</sub> est choisi parmi glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M), cystéine (C), sérine (S), thréonine (T),

asparagine (N), glutamine (Q), acide glutamique (E), acide aspartique (D), lysine (K), arginine (R), et histidine (H), SEQ ID NO :10);

(f) IDLWX<sub>4</sub>VHPGGTRI, dans laquelle X<sub>4</sub> est choisi parmi les acides aminés apolaires (i.e. glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), 5 phénylalanine (F), tyrosine (Y), méthionine (M) et cystéine (C), SEQ ID NO :11).

5. Utilisation selon la revendication 4, dans laquelle caractérisée en ce que ledit polypeptide comprend en outre au moins un fragment choisi parmi :

(a) ATGTP (SEQ ID NO:6) ;

(b) STGFX<sub>1</sub>APG, dans laquelle X<sub>1</sub> est choisi parmi les acides aminés aliphatiques (i.e. l'alanine 10 (A), la valine (V), la leucine (L), l'isoleucine (I), et la proline (P), SEQ ID NO :12) ;

(c) VX<sub>2</sub>FMGCAAA, dans laquelle X<sub>2</sub> est choisi parmi les acides aminés polaires neutres ou polaires chargés positivement (i.e. sérine (S), thréonine (T), asparagine (N), glutamine (Q), lysine (K), arginine (R), et histidine (H), SEQ ID NO:13) ;

(d) VCLELSS (SEQ ID NO :9) ;

15 (e) IHSIFX<sub>3</sub>DGCAA, dans laquelle X<sub>3</sub> est choisi parmi les acides aminés apolaires ou polaires neutres (i.e. glycine (G), alanine (A), valine (V), leucine (L), isoleucine (I), proline (P), tryptophane (W), phénylalanine (F), tyrosine (Y), méthionine (M), cystéine (C), sérine (S), thréonine (T), asparagine (N), et glutamine (Q), SEQ ID NO :14);

(g) IDLWX<sub>4</sub>VHPGGTRI, dans laquelle X<sub>4</sub> est choisi parmi les acides aminés apolaires 20 aliphatiques (i.e. l'alanine (A), la valine (V), la leucine (L), l'isoleucine (I), et la proline (P), SEQ ID NO:15).

6. Utilisation selon la revendication 5, dans laquelle caractérisée en ce que ledit polypeptide comprend en outre au moins un fragment choisi parmi :

(a) ATGTP (SEQ ID NO:6) ;

25 (b) STGFX<sub>1</sub>APG, dans laquelle X<sub>1</sub> est choisi parmi la leucine (L), la valine (V), et l'isoleucine (I) (SEQ ID NO :16),

(c) VX<sub>2</sub>FMGCAAA, dans laquelle X<sub>2</sub> est choisi parmi l'asparagine (N) et l'histidine (H) (SEQ ID NO :17);

(d) VCLELSS (SEQ ID NO :9) ;

30 (e) IHSIFX<sub>3</sub>DGCAA, dans laquelle X<sub>3</sub> est choisi parmi la glycine (G) et la sérine (S) (SEQ ID NO :18) ;

(g) IDLWX<sub>4</sub>VHPGGTRI, dans laquelle X<sub>4</sub> est choisi parmi l'alanine (A) et la valine (V) (SEQ ID NO :19).

7. Utilisation selon l'une quelconque des revendications 1 à 6, caractérisée en ce que ledit polypeptide comprend une séquence d'acides aminés choisie parmi SEQ ID NO :1, SEQ ID NO :2, SEQ ID NO :3, SEQ ID NO :4 et SEQ ID NO :5.

8. Molécule d'acides nucléiques isolée comprenant une séquence d'acides nucléiques codant pour un polypeptide tel que défini dans l'une quelconque des revendications 1 à 7, caractérisée en ce que :

- a) la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques ; ou
- b) la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques ; ou
- c) la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une cellule hôte, notamment une levure ou une bactérie; ou
- d) toute combinaison de a) à c).

9. Molécule d'acides nucléiques isolée selon la revendication 8, caractérisée en ce que lorsque la molécule d'acides nucléiques isolée comprend en outre un promoteur contrôlant l'expression de la séquence d'acides nucléiques, le promoteur est un promoteur exogène, en particulier un promoteur de levure.

10. Molécule d'acides nucléiques isolée selon la revendication 8, caractérisée en ce que lorsque la molécule d'acides nucléiques isolée comprend en outre un terminateur de transcription contrôlant l'expression de la séquence d'acides nucléiques, le terminateur de transcription est un terminateur exogène, tel qu'un terminateur de levure.

11. Molécule d'acides nucléiques isolée selon la revendication 8, caractérisée en ce que la séquence d'acides nucléiques est en outre optimisée pour l'expression dans une levure, avantageusement elle est choisie parmi SEQ ID NO :25 à SEQ ID NO :29.

12. Vecteur comprenant une molécule d'acides nucléiques selon l'une quelconque des revendications 8 à 11, ledit vecteur étant de préférence un plasmide.

13. Cellule hôte non humaine comprenant une molécule d'acides nucléiques selon l'une quelconque des revendications 8 à 11, ou un vecteur selon la revendication 12.

14. Cellule hôte non humaine selon la revendication 13, caractérisée en ce que ladite cellule hôte est une levure, avantageusement sélectionnée parmi les genres *Saccharomyces*, *Candida*, *Ashbya*, *Dekkera*, *Pichia (Hansenula)*, *Debaryomyces*, *Clavispora*, *Lodderomyces*,

5 *Yarrowia*, *Zigosaccharomyces*, *Schizosaccharomyces*, *Torulaspota*, *Kluyveromyces*,  
*Brettanomyces*, *Cryptococcus* et *Malassezia* ; de manière plus particulière parmi les  
espèces *Saccharomyces cerevisiae*, *Saccharomyces boulardii*, *Saccharomyces douglasii*,  
*Saccharomyces bayanus*, *Zigosaccharomyces bailii*, *Schizosaccharomyces pombe*, *Dekkera*  
10 *brucelensis*, *Dekkera intermedia*, *Brettanomyces custersii*, *Brettanomyces intermedius*,  
*Kluyveromyces themotolerens*, *Torulaspota globosa* ou *Torulaspota glabrata* ; encore plus  
particulièrement, la levure est du genre *Saccharomyces*, de préférence de l'espèce  
*Saccharomyces cerevisiae*.

15. Méthode pour produire du phloroglucinol, qui comprend les étapes suivantes :

- 10 (ii1) la mise en contact d'une cellule hôte non humaine exprimant le polypeptide  
choisi parmi les polykétide synthases de type III de cyanobactéries tel que défini  
dans l'une quelconque des revendications 1 à 7, avec un substrat approprié ;  
(iii1) la culture *in vitro* de la cellule hôte de l'étape (ii1) dans des conditions  
permettant l'expression de la molécule d'acides nucléiques contenue dans ladite  
15 cellule hôte, de manière à produire du phloroglucinol ;  
(iv1) optionnellement la récupération du milieu de culture comprenant le  
phloroglucinol, obtenu après l'étape (iii1) ; et  
(v1) optionnellement, la purification du phloroglucinol à partir du milieu de culture  
de l'étape (iv1).

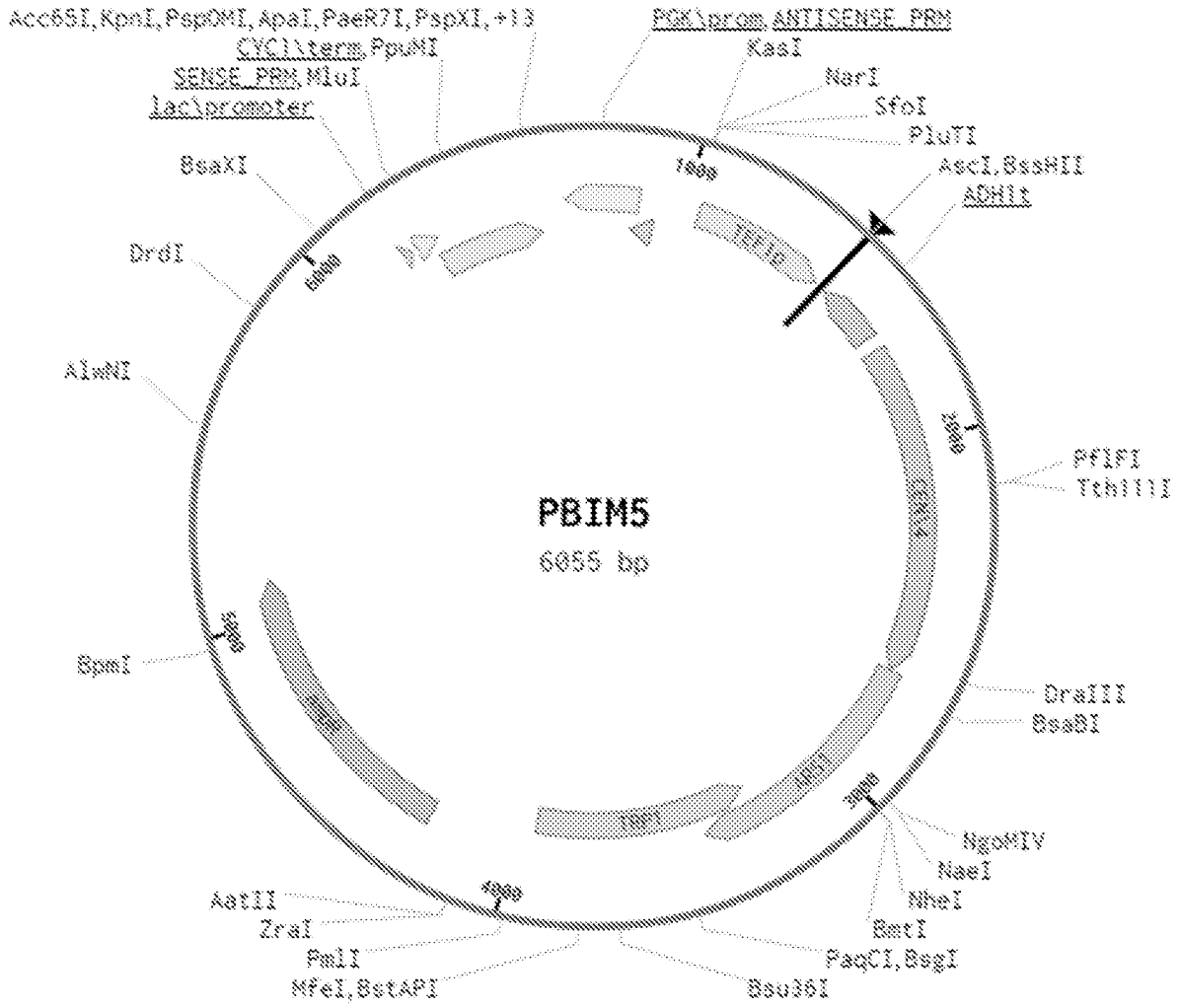
20 16. Méthode pour produire du phloroglucinol, qui comprend les étapes suivantes :

- (ii2) la mise en contact d'un polypeptide choisi parmi les polykétide synthases de  
type III de cyanobactéries tel que défini dans l'une quelconque des revendications  
1 à 7 avec du malonyl-CoA ;  
(iii2) l'incubation du mélange issu de l'étape (ii2) dans des conditions adaptées pour  
25 produire du phloroglucinol ;  
(iv2) optionnellement la récupération du milieu réactionnel comprenant le  
phloroglucinol, obtenu après l'étape (iii2) ; et  
(v2) optionnellement, la purification du phloroglucinol à partir du milieu réactionnel  
de l'étape (iv2).

30



[Fig. 2]



## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/FR2024/050407**

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> <i>C12N 9/10</i> (2006.01)i; <i>C12N 15/63</i> (2006.01)i; <i>C12P 7/22</i> (2006.01)i  According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>  Minimum documentation searched (classification system followed by classification symbols) C12N; C12P; C40B  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, Sequence Search		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2019002799 A1 (MICHELIN & CIE [FR]) 03 January 2019 (2019-01-03) cited in the application abstract page 24, line 7 - line 18 claims 1-18; figure 1; table 4; sequences 6,8	1-16
A	"SubName: Full=Putative naringenin-chalcone synthase [ECO:0000313 EMBL:AFY56411.1];" 06 March 2013 (2013-03-06), retrieved from EBI accession no. UNIPROT:K9RFY4, abstract No. Database accession no. K9RFY4, Retrieved from: UniProt [online] XP002810841 sequence	1-16
Y	"SubName: Full=Putative naringenin-chalcone synthase [ECO:0000313 EMBL:AFY93621.1];" 22 February 2023 (2023-02-22), retrieved from EBI accession no. UNIPROT:K9UGQ3, abstract No. Database accession no. K9UGQ3, Retrieved from: UniProt [online] XP002810842 the whole document	1-16
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
Date of the actual completion of the international search <b>28 June 2024</b>		Date of mailing of the international search report <b>18 July 2024</b>
Name and mailing address of the ISA/EP <b>European Patent Office p.b. 5818, Patentlaan 2, 2280 HV Rijswijk Netherlands (Kingdom of the)</b> Telephone No. (+31-70)340-2040 Facsimile No. (+31-70)340-3016		Authorized officer <b>Gurdjian, Didier</b>  Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR2024/050407

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	"SubName: Full=Naringenin-chalcone synthase {ECO:0000313 EMBL:PSB23125.1};" 22 February 2023 (2023-02-22), retrieved from EBI accession no. UNIPROT:A0A2T1DRP2, abstract No. Database accession no. A0A2T1DRP2, Retrieved from: UniProt [online] XP002810843 the whole document	1-16
A	"Calothrix sp. HK-06 naringenin-chalcone synthase" 27 January 2017 (2017-01-27), retrieved from EBI accession no. EMBLWGS:OKH51295, abstract No. Database accession no. OKH51295, Retrieved from: EMBL [online] XP002810844 the whole document	1-16
A	"SubName: Full=Naringenin-chalcone synthase {ECO:0000313 EMBL:AFZ30814.1}; EC=2.3.1.74 {ECO:0000313 EMBL:AFZ30814.1};" 22 February 2023 (2023-02-22), retrieved from EBI accession no. UNIPROT:K9XFD7, abstract No. Database accession no. K9XFD7, Retrieved from: UniProt [online] XP002810845 the whole document	1-16
A	AUSTIN MICHAEL B ET AL. "The chalcone synthase superfamily of type III polyketide synthases" <i>NATURAL PRODUCT REPORTS</i> , GB, Vol. 20, No. 1, 21 January 2003 (2003-01-21), pages 79-110 DOI: 10.1039/b100917f ISSN: 0265-0568, XP093172857 abstract page 105 - page 106; figure 18	1-16
A	NAAKE THOMAS ET AL. "Kingdom-wide analysis of the evolution of the plant type III polyketide synthase superfamily" <i>PLANT PHYSIOLOGY</i> , Rockville, Md, USA, Vol. 185, No. 3, 02 April 2021 (2021-04-02), pages 857-875, Retrieved from the Internet: <a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8133574/pdf/kiaa086.pdf">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8133574/pdf/kiaa086.pdf</a> DOI: 10.1093/plphys/kiaa086 ISSN: 0032-0889, XP093172870 abstract; figure 3	1-16

**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
  - a.  forming part of the international application as filed:
    - in the form of an Annex C/ST.25 text file.
    - on paper or in the form of an image file.
  - b.  furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
  - c.  furnished subsequent to the international filing date for the purposes of international search only:
    - in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
    - on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2.  In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.  
**PCT/FR2024/050407**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2019002799	A1	03 January 2019	BR	112019028038	A2	07 July 2020
				CA	3068730	A1	03 January 2019
				CN	111051515	A	21 April 2020
				EP	3645726	A1	06 May 2020
				FR	3068368	A1	04 January 2019
				US	2021147884	A1	20 May 2021
				WO	2019002799	A1	03 January 2019
-----							

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/FR2024/050407

<b>A. CLASSEMENT DE L'OBJET DE LA DEMANDE</b> INV. C12N9/10 C12N15/63 C12P7/22 ADD.				
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB				
<b>B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE</b> Documentation minimale consultée (système de classification suivi des symboles de classement) C12N C12P C40B				
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche				
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés) EPO-Internal, Sequence Search				
<b>C. DOCUMENTS CONSIDERES COMME PERTINENTS</b>				
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées		
Y	WO 2019/002799 A1 (MICHELIN & CIE [FR]) 3 janvier 2019 (2019-01-03) cité dans la demande abrégé page 24, ligne 7 - ligne 18 revendications 1-18; figure 1; tableau 4; séquences 6,8  <div style="text-align: center;">-----</div> <div style="text-align: center;">-/--</div>	1-16		
<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;"> <input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents                 </td> <td style="width: 50%; border: none;"> <input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe                 </td> </tr> </table>			<input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents	<input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe
<input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents	<input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe			
* Catégories spéciales de documents cités:				
"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent "E" document antérieur, mais publié à la date de dépôt international ou après cette date "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée	"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier "&" document qui fait partie de la même famille de brevets			
Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale			
28 juin 2024	18/07/2024			
Nom et adresse postale de l'administration chargée de la recherche internationale	Fonctionnaire autorisé			
Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Gurdjian, Didier			

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>DATABASE UniProt [Online]</p> <p>6 mars 2013 (2013-03-06), "SubName: Full=Putative naringenin-chalcone synthase {ECO:0000313 EMBL:AFY56411.1}";", XP002810841, extrait de EBI accession no. UNIPROT:K9RFY4 Database accession no. K9RFY4 séquence</p>	1-16
Y	<p>DATABASE UniProt [Online]</p> <p>22 février 2023 (2023-02-22), "SubName: Full=Putative naringenin-chalcone synthase {ECO:0000313 EMBL:AFY93621.1}";", XP002810842, extrait de EBI accession no. UNIPROT:K9UGQ3 Database accession no. K9UGQ3 le document en entier</p>	1-16
A	<p>DATABASE UniProt [Online]</p> <p>22 février 2023 (2023-02-22), "SubName: Full=Naringenin-chalcone synthase {ECO:0000313 EMBL:PSB23125.1}";", XP002810843, extrait de EBI accession no. UNIPROT:A0A2T1DRP2 Database accession no. A0A2T1DRP2 le document en entier</p>	1-16
A	<p>DATABASE EMBL [Online]</p> <p>27 janvier 2017 (2017-01-27), "Calothrix sp. HK-06 naringenin-chalcone synthase", XP002810844, extrait de EBI accession no. EMBLWGS:OKH51295 Database accession no. OKH51295 le document en entier</p>	1-16

-/--

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>DATABASE UniProt [Online]</p> <p>22 février 2023 (2023-02-22), "SubName: Full=Naringenin-chalcone synthase {ECO:0000313 EMBL:AFZ30814.1}; EC=2.3.1.74 {ECO:0000313 EMBL:AFZ30814.1}"; XP002810845, extrait de EBI accession no. UNIPROT:K9XFD7 Database accession no. K9XFD7 le document en entier</p> <p>-----</p>	1-16
A	<p>AUSTIN MICHAEL B ET AL: "The chalcone synthase superfamily of type III polyketide synthases", NATURAL PRODUCT REPORTS, vol. 20, no. 1, 21 janvier 2003 (2003-01-21), pages 79-110, XP093172857, GB ISSN: 0265-0568, DOI: 10.1039/b100917f abrégé page 105 - page 106; figure 18</p> <p>-----</p>	1-16
A	<p>NAAKE THOMAS ET AL: "Kingdom-wide analysis of the evolution of the plant type III polyketide synthase superfamily", PLANT PHYSIOLOGY, vol. 185, no. 3, 2 avril 2021 (2021-04-02) , pages 857-875, XP093172870, Rockville, Md, USA ISSN: 0032-0889, DOI: 10.1093/plphys/kiaa086 Extrait de l'Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/artic les/PMC8133574/pdf/kiaa086.pdf&gt; abrégé; figure 3</p> <p>-----</p>	1-16

## RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/FR2024/050407

### Cadre N° I Séquence(s) de nucléotides ou d'acides aminés (suite du point 1.c de la première feuille)

1. En ce qui concerne la ou les séquences de nucléotides ou d'acides aminés divulguées dans la demande internationale, la recherche internationale a été effectuée sur la base d'un listage des séquences :
  - a.  faisant partie de la demande internationale telle que déposée.
  - b.  remis postérieurement à la date de dépôt international aux fins de la recherche internationale (règle 13<sup>ter</sup>.1.a),  
 accompagné d'une déclaration selon laquelle le listage des séquences ne va pas au-delà de la divulgation faite dans la demande internationale telle que déposée.
2.  En ce qui concerne la ou les séquences de nucléotides ou d'acides aminés divulguées dans la demande internationale, le présent rapport a été établi dans la mesure où une recherche valable pouvait être effectuée en l'absence d'un listage des séquences conforme à la norme ST.26 de l'OMPI.
3. Commentaire complémentaires:

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/FR2024/050407

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 2019002799	A1	03-01-2019	
		BR 112019028038 A2	07-07-2020
		CA 3068730 A1	03-01-2019
		CN 111051515 A	21-04-2020
		EP 3645726 A1	06-05-2020
		FR 3068368 A1	04-01-2019
		US 2021147884 A1	20-05-2021
		WO 2019002799 A1	03-01-2019
-----			