



(86) Date de dépôt PCT/PCT Filing Date: 2008/04/02
(87) Date publication PCT/PCT Publication Date: 2008/11/27
(85) Entrée phase nationale/National Entry: 2009/09/24
(86) N° demande PCT/PCT Application No.: FR 2008/050576
(87) N° publication PCT/PCT Publication No.: 2008/142318
(30) Priorité/Priority: 2007/04/03 (FR0754235)

(51) Cl.Int./Int.Cl. *C12N 9/10* (2006.01),
C12N 15/52 (2006.01)
(71) Demandeur/Applicant:
PHILIP MORRIS PRODUCTS S.A., CH
(72) Inventeurs/Inventors:
SALLAUD, CHRISTOPHE, FR;
RONTEIN, DENIS, FR;
TISSIER, ALAIN, FR
(74) Agent: FETHERSTONHAUGH & CO.

(54) Titre : GENES CODANT POUR LA Z,Z-FARNESYL DIPHOSPHATE SYNTHASE ET UNE SESQUITERPENE
SYNTHASE A PRODUITS MULTIPLES ET LEURS UTILISATIONS
(54) Title: GENES ENCODING Z,Z-FARNESYL DIPHOSPHATE SYNTHASE AND A SESQUITERPENE SYNTHASE
WITH MULTIPLE PRODUCTS AND USES THEREOF

(57) **Abrégé/Abstract:**

La présente invention concerne les gènes impliqués dans la voie de biosynthèse de sesquiterpènes de type SB (alpha-santalène, épi-beta-santalène, cis-alpha-bergamotène, trans-alpha-bergamotène et endo-beta-bergamotène,) et de son précurseur le ZZ-farnesyl diphosphate (ZZ-FPP) impliquant une Z,Z-FPP synthase et une sesquiterpene SB synthase et leurs utilisations pour la production de composés sesquiterpéniques de type SB.



(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international

PCT

(43) Date de la publication internationale
27 novembre 2008 (27.11.2008)(10) Numéro de publication internationale
WO 2008/142318 A3(51) Classification internationale des brevets :
C12N 9/10 (2006.01) C12N 15/52 (2006.01)(21) Numéro de la demande internationale :
PCT/FR2008/050576

(22) Date de dépôt international : 2 avril 2008 (02.04.2008)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0754235 3 avril 2007 (03.04.2007) FR(71) Déposant (pour tous les États désignés sauf US) : **LIBRO-PHYT** [FR/FR]; 24 rue Raspail, F-84120 Pertuis (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : **SAL-LAUD, Christophe** [FR/FR]; 233 avenue du Littoral, F-34070 Montpellier (FR). **RONTEIN, Denis** [FR/FR]; 49 Lou Cigaloun, Chemin de Sainte Annette, F-04800 Gréoux Les Bains (FR). **TISSIER, Alain** [FR/FR]; 24 rue Raspail, F-84120 Pertuis (FR).(74) Mandataires : **GALLOIS, Valérie** etc.; Becker & Associés, 25 rue Louis Le Grand, F-75002 Paris (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, BR, BW, BY, BZ, CA, CH,

CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Déclaration en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv))

Publiée :

— avec rapport de recherche internationale
— avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues
— avec la partie réservée au listage des séquences de la description publiée séparément sous forme électronique et disponible sur demande auprès du Bureau international

(88) Date de publication du rapport de recherche internationale: 15 janvier 2009

(54) Title: GENES ENCODING Z,Z-FARNESYL DIPHOSPHATE SYNTHASE AND A SESQUITERPENE SYNTHASE WITH MULTIPLE PRODUCTS AND USES THEREOF

(54) Titre : GENES CODANT POUR LA Z,Z-FARNESYL DIPHOSPHATE SYNTHASE ET UNE SESQUITERPENE SYNTHASE A PRODUITS MULTIPLES ET LEURS UTILISATIONS

(57) Abstract: The present invention relates to the genes involved in the biosynthetic pathway for sesquiterpenes of SB type (alpha-santalene, epi-beta-santalene, cis-alpha-bergamotene, trans-alpha-bergamotene and endo-beta-bergamotene), and for the precursor thereof, Z,Z-farnesyl diphosphate (ZZ-FPP), involving a Z,Z-FPP synthase and a sesquiterpene SB synthase, and to the uses thereof for producing sesquiterpene compounds of SB type.

(57) Abrégé : La présente invention concerne les gènes impliqués dans la voie de biosynthèse de sesquiterpènes de type SB (alpha-santalène, épi-beta-santalène, cis-alpha-bergamotène, trans-alpha-bergamotène et endo-beta-bergamotène,) et de son précurseur le ZZ-farnesyl diphosphate (ZZ-FPP) impliquant une Z,Z-FPP synthase et une sesquiterpene SB synthase et leurs utilisations pour la production de composés sesquiterpéniques de type SB.

WO 2008/142318 A3

**GENES CODANT POUR LA Z,Z-FARNESYL DIPHOSPHATE SYNTHASE ET
UNE SESQUITERPENE SYNTHASE A PRODUITS MULTIPLES ET LEURS
UTILISATIONS**

5 **Domaine de l'Invention**

La présente invention se rapporte à deux gènes responsables de la synthèse d'un mélange de sesquiterpènes et leur utilisation pour la préparation de ces composés dans des organismes vivants tels que les bactéries, les levures, les cellules animales et les plantes.

10

Introduction

Les terpènes sont des molécules présentes dans tous les organismes vivants tels que les bactéries, les champignons, les animaux, et les plantes. Ils forment la plus grande classe de molécules naturelles du monde vivant. Chez les plantes, ces molécules sont présentes dans le métabolisme primaire en tant qu'hormones (les gibbérellines, les cytokinines, les brassinostéroïdes et l'acide abscissique) ou en tant que composés impliqués dans la photosynthèse (caroténoïdes, chlorophylles, plastoquinones et phytol) ainsi que dans les processus de prénylation des protéines, et dans la structure des membranes (stérols). Cependant, les terpénoïdes issus du métabolisme secondaire représentent l'essentiel de la diversité structurale de cette classe de molécules. Les rôles des terpénoïdes dits secondaires se rangent essentiellement dans le domaine des interactions avec l'environnement comme par exemple l'attraction d'insectes bénéfiques pour la plante (pollinisation et prédateurs d'insectes phytophages), la défense contre les pathogènes et les insectes et la protection contre le stress photo-oxydatif (Tholl, 2006).

25 Les terpènes sont constitués de multiples d'unités isoprényles à 5 carbones. Le nombre d'unités isoprényles permet de les classer en monoterpènes (composés à dix atomes de carbone ou C₁₀), sesquiterpènes (C₁₅), diterpènes (C₂₀), sesterterpènes (C₂₅), triterpènes (C₃₀), tetraterpènes (C₄₀), et ainsi de suite. Les précurseurs universels de tous les terpènes sont l'isopentenyl diphosphate (IPP) et le diméthylallyl diphosphate (DMAPP). Deux voies distinctes conduisent à la formation de l'IPP et du DMAPP. L'une, la voie du mévalonate (MEV) est localisée dans le cytoplasme des cellules eucaryotes alors que l'autre, la voie du méthyl-erythritol (MEP) n'existe que chez certaines bactéries et les plantes où elle est localisée dans le chloroplaste (Rodriguez-Concepcion et Boronat, 2002).

30

Toutes les étapes et tous les gènes codant les enzymes des étapes de ces deux voies sont connus chez un certain nombre d'organismes.

La première étape de la biosynthèse des terpènes à partir des précurseurs communs IPP et DMAPP est réalisée par des prényl diphosphate synthases (ou prényltransférases) qui vont par condensation d'un précurseur homoallylique, l'IPP, et d'un précurseur allylique (le DMAPP, le géranyl diphosphate, le farnésyldiphosphate ou encore le géranylgéranyldiphosphate), générer des chaînes de prényldiphosphates de longueur variable. Il existe deux grandes classes principales de prényldiphosphate synthases : les *cis*-prényltransférases ou *Z*-prényltransférases et les *trans*-prényltransférases ou *E*-prényltransférases. Les *cis*-prényltransférases conduisent à la formation d'une double liaison en configuration *cis*, alors que les *trans*-prényltransférases conduisent à la formation d'une double liaison en configuration *trans* (Koyama 1999). Ainsi, l'addition séquentielle de 2 unités d'IPP sur un DMAPP par une *trans*-farnésyldiphosphate synthase, ou *E*-FPS produit du *E,E*-farnésyldiphosphate (*E,E*-FPP, C15). La réaction d'élongation commence par la formation d'un cation allylique après l'élimination de l'ion diphosphate pour former un prényl allylique. Ensuite, l'addition d'un IPP se produit avec une élimination stéréospécifique d'un proton en position 2 pour former une nouvelle liaison C-C et une nouvelle double liaison dans le produit. La répétition de cette condensation stéréospécifique de l'IPP avec un prényl diphosphate allylique, conduit à la synthèse d'un prényl diphosphate de longueur de chaîne et de stéréochimie données spécifiques de chaque enzyme. En particulier, la longueur de la chaîne du produit final peut varier considérablement allant d'un C10 (géranyl diphosphate) jusqu'aux polyprenyldiphosphates précurseurs du latex naturel composés de plusieurs milliers d'atomes de carbone.

Comme mentionné plus haut, les prényltransférases peuvent être divisées en deux familles génétiques différentes selon la stéréochimie (*E* ou *Z*) de la double liaison formée après chaque cycle d'élongation (Poulter, 2006 ; Liang *et al.*, 2002).

Les *E*-prényltransférases caractérisées à ce jour sont responsables de la synthèse de prényl phosphates à courte chaîne (C10 à C50) en configuration *E*. On peut citer par exemple la géranyl diphosphate synthase (GPS), la farnesyl diphosphate synthase (FPS), la géranyl géranyl diphosphate synthase (GGPS), l'octaprenyl diphosphate synthase (OPS), la solanesyl diphosphate synthase (SPS) et la decaprenyl diphosphate synthase (DPS) qui sont respectivement responsables de la synthèse du (*E*)-GPP (C10), (*E,E*)-FPP (C15),

(*E,E,E*)-GGPP (C20), (*E,E,E,E,E,E,E*)-OPP (C40), (*E,E,E,E,E,E,E,E*)-SPP (C45), et du (*E,E,E,E,E,E,E,E,E*)-DPP (C50). Il faut noter que le (*E,E*)-FPP est en général le substrat des prényltransférases responsables de la synthèse des chaînes à plus de 20 carbones. Le premier gène codant pour une (*E*)-prényl transférase caractérisé est celui de la FPS de rat (Clarke *et al.*, 1987). L'alignement des séquences codant pour des *E*-prényltransférases a permis de mettre en évidence deux motifs conservés de type DDXXD. La structure tridimensionnelle d'une *E*-prényltransférase, déterminée aux rayons X pour la première fois sur une FPS (Tarshis *et al.*, 1994), et les études de mutagenèse dirigée ont montré que les deux motifs DDXXD étaient impliqués dans la liaison avec les substrats et dans l'activité catalytique en résultant. Finalement, il a été démontré qu'un petit nombre d'acides aminés spécifiques, localisés en amont du premier motif DDXXD, déterminaient la longueur de l'élongation de la chaîne d'isoprényles synthétisée (Wang & Ohnuma, 1999).

Le premier gène codant pour une *Z*-prényltransférase a été cloné chez *Micrococcus luteus* (Shimizu *et al.*, 1998). Il code pour une undecaprényltransférase (UPS) qui utilise le *E*-FPP comme substrat pour former un prényl diphosphate à 55 carbones qui possède une stéréochimie de type *Z,E*. Cette molécule intervient dans la biosynthèse des peptidoglycanes de la paroi de certaines bactéries. Une séquence homologue a été caractérisée chez *Arabidopsis* comme étant responsable de la synthèse d'isoprényl diphosphate en conformation *Z,E* de 100 à 130 carbones (Oh *et al.*, 2000). L'analyse des séquences en acides aminés de ces enzymes ne montre aucune homologie avec les séquences des *E*-prényltransférases (Koyama, 1999). En particulier, les *Z*-prényltransférases ne possèdent pas de motifs riches en acide aspartique (DDXXD). Sept régions conservées qui interviennent toutes plus ou moins dans la fixation du substrat et la catalyse de celui-ci ont été identifiées. Toutes les *Z*-prényltransférases identifiées à ce jour synthétisent des prényls diphosphates ayant une longueur de chaîne supérieure ou égale à 55 carbones à l'exception d'une *Z,E*-FPS de *Mycobacterium tuberculosis* qui utilise le *E*-GPP comme substrat pour former du *Z,E*-FPP (Schulbach *et al.*, 2000). Une étude récente de mutagenèse dirigée sur l'UPS de *M. luteus* a permis d'identifier plusieurs acides aminés jouant un rôle clé dans le nombre de cycles d'élongation de la chaîne d'isoprényles synthétisée (Kharel *et al.*, 2006).

Les isoprényles diphosphates sont les substrats d'une classe d'enzyme, les terpènes synthases, qui permettent la formation du squelette de base des terpènes cycliques ou

acycliques en C10 (monoterpène), C15 (sesquiterpène), C20 (diterpène), et C30 (triterpène) (Cane 1999; McMillan and Beale, 1999; Wise and Croteau, 1999). La diversité de réaction de ces enzymes est à l'origine de la diversité des squelettes terpéniques cycliques que l'on retrouve dans la nature. Quatre stéréoisomères du farnésyl diphosphate

5 sont théoriquement possibles en fonction de la configuration stéréochimique des doubles liaisons en position 2 ou 6 de la molécule (*E,E*, *Z,E*, *E,Z* ou *Z,Z*). Cependant, à ce jour, toutes les sesquiterpènes synthases pour lesquelles les gènes ont été caractérisés utilisent comme substrat le *E,E*-FPP. C'est le cas par exemple de la 5-épi-aristolochène synthase de

10 tabac (Facchini & Chappell, 1992), de la (*E*)-alpha-bisabolène synthase d'*Abies grandis* (Bohlmann *et al.*, 1998), de la germacrène A synthase de chicorée (Bouwmeester *et al.*, 2002), de l'amorpha-4,11-diene synthase de l'armoise annuelle (Wallaart *et al.*, 2001), de la germacrène C synthase de tomate (Colby *et al.*, 1998), de la caryophyllène synthase de l'armoise annuelle (Cai *et al.*, 2002), et de la valencène synthase d'orange (Sharon-Asa *et al.*, 2003). Une étude récente a cependant montré qu'une sesquiterpène synthase de maïs

15 (tps4) responsable de la synthèse d'un mélange de 14 sesquiterpènes acceptait le *E,E*-FPP et le *Z,E*-FPP (Kollner *et al.*, 2006), montrant ainsi que les sesquiterpènes synthases pouvaient utiliser d'autres isomères que le *E,E*-FPP comme substrat. Néanmoins, aucune sesquiterpène synthase clonée à ce jour n'utilise de *Z,Z*-FPP.

Les trichomes glandulaires de type VI de *Solanum habrochaites* (auparavant

20 dénommée *Lycopersicon hirsutum*) excrètent de grandes quantités de sesquiterpènes oléfiniques appartenant à deux classes distinctes. La classe I contient notamment la germacrène B et la classe II contient entre autres l'alpha-santalène et l'alpha-bergamotène. L'analyse de la ségrégation entre un génotype de tomate cultivée (*Lycopersicon esculentum* = *Solanum lycopersicum*) et un génotype de tomate sauvage (*Lycopersicon*

25 *hirsutum* = *Solanum habrochaites*) a permis de démontrer que la biosynthèse de ces deux classes différentes de sesquiterpènes était contrôlée par deux loci différents (Sst1A et Sst2) localisés sur deux chromosomes différents (van der Hoeven *et al.* 2000). Le locus Sst1A est localisé sur le chromosome 6, et les gènes de ce locus sont responsables de l'accumulation des sesquiterpènes de classe I. L'allèle Sst2 de *S. habrochaites* localisé sur

30 le chromosome 8 contrôle l'accumulation de sesquiterpènes de classe II comme l'alpha-santalène mais aussi le cis-alpha-bergamotène, le trans-alpha-bergamotène, l'épi-beta-santalène et le beta-bergamotène et d'autres composés mineurs non identifiés. Tandis que les séquences codantes correspondant aux gènes localisés sur le locus Sst1-A ont été

identifiées (van der Hoeven *et al.* 2000), celles localisées sur le locus Sst2 ne l'ont pas été.

Résumé de l'Invention

5 La présente invention se rapporte à deux gènes responsables de la synthèse d'un mélange de sesquiterpènes composé principalement d'alpha-santalène, d'épi-beta-santalène, de cis-alpha-bergamotène, de trans-alpha-bergamotène, et de endo-beta-bergamotène et de leur précurseur le Z,Z-farnesyl diphosphate (Z,Z-FPP) et leur utilisation pour la préparation de ces composés dans des organismes vivants tels que les bactéries, les levures, les cellules
10 animales et les plantes.

Plus précisément, le présent document décrit l'identification et la caractérisation de deux gènes de *S. habrochaites* responsables de la synthèse des sesquiterpènes de la classe II (alpha-santalène, épi-beta-santalène, cis-alpha-bergamotène, trans-alpha-bergamotène et
15 endo-beta-bergamotène entre autres) de *S. habrochaites* que nous avons renommés sesquiterpènes de type SB (pour Santalène et Bergamotène) dans la suite du texte. Le premier gène code pour une Z,Z-farnesyl diphosphate synthase et le second pour une sesquiterpène synthase à produits multiples qui utilise le Z,Z-farnesyl diphosphate comme substrat. La production des protéines recombinantes correspondantes nous a permis de
20 démontrer que le profil sesquiterpénique obtenu *in vitro* est identique à celui du contrôlé par le locus Sst2 de tomate.

L'invention concerne les activités enzymatiques de la biosynthèse du Z,Z-FPP à partir d'IPP et de DMAP d'une part, et d'un mélange de sesquiterpènes composé principalement
25 d'alpha-santalène, d'épi-beta-santalène, de cis-alpha-bergamotène, de trans-alpha-bergamotène et de endo-beta-bergamotène à partir de Z,Z-FPP, les séquences peptidiques responsables de ces activités et les séquences d'acide nucléique codant ces séquences peptidiques. L'invention concerne également des méthodes utilisant ces activités enzymatiques pour la production de ces composés ou de dérivés de ceux-ci.

30

La présente invention concerne donc une méthode de production d'un mélange de sesquiterpènes, principalement composés d'alpha-santalène, d'épi-beta-santalène, de cis-alpha-bergamotène, de trans-alpha-bergamotène et de endo-beta-bergamotène, à partir de Z,Z-FPP dans une cellule ayant une source d'IPP et de DMAP, comprenant :

- 5 a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression comprenant un premier gène codant pour une Z,Z-FPS selon la présente invention et d'une construction portant une cassette d'expression comprenant un deuxième gène codant pour une sesquiterpène synthase dénommée SBS selon la présente invention ;
- b) la culture de la cellule transformée dans des conditions appropriées pour l'expression desdits premier et deuxième gènes ; et,
- 10 c) facultativement, la récolte du Z,Z-FPP ou des produits sesquiterpéniques de l'enzyme SBS ou de dérivés de ceux-ci contenus dans ladite cellule et/ou dans le milieu de culture.

La présente invention concerne une protéine isolée ou recombinante ayant une activité Z,Z-FPS. Cette protéine a de préférence une séquence présentant au moins 80 % d'identité avec la SEQ ID No 2. Elle concerne aussi un acide nucléique comprenant une séquence
15 nucléotidique codant pour une telle protéine, ou une séquence capable de s'hybrider à celui-ci dans des conditions stringentes.

La présente invention concerne également une protéine isolée ou recombinante ayant une activité de type SB synthase et ayant une séquence présentant au moins 80 % d'identité
20 avec la SEQ ID No 4. Elle concerne aussi un acide nucléique comprenant une séquence nucléotidique codant pour une telle protéine, ou une séquence capable de s'hybrider à celui-ci dans des conditions stringentes.

La présente invention concerne une cassette d'expression comprenant un acide nucléique selon la présente invention, un vecteur comprenant une cassette d'expression ou un acide
25 nucléique selon la présente invention, une cellule hôte comprenant un vecteur, une cassette d'expression ou un acide nucléique selon la présente invention, et un organisme transgénique non-humain comprenant une cellule, un vecteur, une cassette d'expression ou un acide nucléique selon la présente invention.

30

La présente invention concerne en outre une méthode de production du Z,Z-farnesyl diphosphate (Z,Z-FPP), ou de dérivés de ceux-ci à partir d'IPP et de DMAP dans une cellule ayant une source d'IPP et de DMAP, comprenant :

- a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression avec une séquence nucléique codant une Z,Z-FPS selon la présente invention ;
- b) la culture de la cellule transformée dans des conditions appropriées pour l'expression du gène ; et
- c) facultativement, la récolte du ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.

La présente invention concerne l'utilisation d'une protéine, d'un acide nucléique, d'une cellule hôte ou d'un organisme transgénique selon la présente invention pour la préparation de Z,Z-FPP ou des sesquiterpènes de type SB, tels que l'alpha-santalène, l'épi-beta-santalène, le cis-alpha-bergamotène, le trans-alpha-bergamotène et l'endo-beta-bergamotène, ou de dérivés de ceux-ci comme l'alpha-santalol, l'épi-beta-santalol, le cis-alpha-bergamotol, le trans-alpha-bergamotol et le beta-bergamotol.

La présente invention concerne une cellule ou un organisme transgénique non-humain, caractérisé en ce que la voie de synthèse des sesquiterpènes de classe II est bloquée par inactivation soit du gène codant pour une Z,Z-FPS selon la présente invention, soit du gène codant pour une SB synthase selon la présente invention, soit encore des deux gènes.

La présente invention concerne en outre l'utilisation d'un acide nucléique pour identifier des marqueurs moléculaires permettant d'introduire les séquences génomiques correspondantes dans d'autres espèces ou variétés de tomate cultivées (*Solanum lycopersicum*). Au titre de la présente invention, les séquences nucléotidiques SEQ ID N°1 et N°2 ou des variants de celles-ci peuvent également être utilisées comme marqueurs moléculaires pour l'introgession de ces séquences dans d'autres espèces ou variétés de tomates cultivées.

Ainsi, la présente invention concerne des marqueurs moléculaires comprenant tout ou partie d'un acide nucléique présentant au moins 80 % d'identité avec la SEQ ID No 1 ou SEQ ID No 3 pour identifier un polymorphisme génomique entre *S. habrochaites* et des espèces de type *Solanum* sexuellement compatible avec *S. habrochaites* afin d'introduire les gènes correspondant dans ces espèces. Elle concerne également une méthode consistant

à introduire les gènes *zFPS* et *SBS* dans une espèce non sexuellement compatible avec *Solanum habrochaites* par fusion de protoplastes.

La présente invention concerne aussi une méthode de production du Z,Z-farnesol par
5 déphosphorylation du Z,Z-FPP par une phosphatase.

Légende des Figures

Figure 1 : Les stéréoisomères possibles du FPP. Tous les stéréoisomères du FPP sont
10 biosynthétisés à partir d'IPP et de DMAPP. A ce jour seules étaient connues la *E,E*-
farnesyl diphosphate synthase (1) caractérisée dans de nombreux organismes, et la *Z,E*-
farnesyl diphosphate synthase (2) de *Mycobacterium tuberculosis* (Schulbach *et al.*, 2000).
(3) : *Z,Z*-farnesyl diphosphate synthase de tomate selon la présente invention ; (4) : *E,Z*-
farnesyl diphosphate synthase : aucune enzyme ayant cette activité n'a été décrite à ce jour.

15

Figure 2 : Représentation schématique des ADN-T portant les transgènes *Sh-zFPS* et *Sh-SBS*. *Km* : gène de résistance à la kanamycine ; 35S : promoteur du gène 35S du CaMV. CBTS-ter : terminateur du gène CBTS; eCBTS1.0 : promoteur de 1kb du gène CBTS fusionné avec l'amplificateur du promoteur 35S du CaMV. *Sh-zFPS* : phase codante du
20 gène codant pour la zFPS de tomate (*S. habrochaites LA1777*) ; *Sh-SBS* : phase codante du
gène codant pour la SB synthase de tomate (*S. habrochaites LA1777*); LB : Bordure
gauche du T-DNA; RB : Bordure droite du T-DNA; 35S-ter : Termineur du gène 35S du
CaMV; nos-ter : Termineur du gène de la nopaline synthase d'*Agrobacterium tumefaciens*. **Figure 2** : fragment ADN-T du vecteur binaire pLIBRO64 ; **Figure 2B** :
25 fragment ADN-T du vecteur binaire pLIBRO65.

Figure 3 : Analyse de l'expression du transgène *Sh-SBS* dans les feuilles de tabac transgéniques. L'expression a été mesurée par PCR quantitative en temps réel en utilisant des sondes fluorescente (®TAQ-MAN, ABI). L'une des sondes est spécifique du transgène
30 *Sh-SBS* tandis que l'autre est spécifique d'un gène d'actine de tabac (gène contrôle). Les
valeurs de l'axe des ordonnées représentent le rapport des puissances 2 des valeurs
obtenues avec la sonde *Sh-SBS* sur celles obtenues avec la sonde actine ($2^{\text{SBS}}/2^{\text{actine}}$). Ce
rapport exprime le rapport d'expression du transgène *Sh-SBS* sur celui de l'actine. Les
valeurs de l'axe des abscisses indiquent le numéro des lignées transgéniques. **Figure**

3A) Lignées de tabac transgéniques ayant intégré les transgènes *Sh-zFPS* et *Sh-SBS* (pLIBRO-064) **Figure 3B)** Lignées de tabac transgéniques ayant intégré uniquement le transgène *Sh-SBS* (pLIBRO-065).

- 5 **Figure 4** : Profil GC/MS du (Z,Z)-farnesol superposé aux standard de farnesols (mélange d'isomères). **Figure 4A** : Le chromatogramme gris (m/z: 69) correspond au produit obtenu *in vitro* avec la protéine recombinante *Sh-zFPS*-6His. *Sh-FPS*-6His a été incubée avec du DMAPP et de l'IPP. Le produit isoprényl-diphosphate a été déphosphorylé en un alcool et purifié par chromatographie en phase liquide. Le pic n° 1 correspond au Z,Z-farnesol. 1. Le
- 10 chromatogramme noir (m/z: 69) correspond au standard farnesol constitué d'un mélange d'isomères (Z,E)-, (E,Z)-, et (E,E)-farnesol (Fluka). Le pic n°2 correspond à un mélange de (Z,E)- et (E,Z)- farnesol et le pic n°3 au (E,E)-farnesol. **Figure 4B** : Spectre de masse du (Z,Z)-farnesol correspondant au pic n°1. **Figure 4C** : Spectre de masse du (Z,E)- et du (E,Z)-farnesol (mélange) correspondant au pic n°2. **Figure 4D** : Spectre de masse du
- 15 (E,E)-farnesol. correspondant au pic n°3.

Figure 5 : Profil GC/MS des tests d'activités *in vitro* réalisés avec les protéines recombinantes *Sh-zFPS*-6His et *Sh-SBS*-6His : Les deux enzymes recombinantes *Sh-zFPS*-6His et *Sh-SBS*-6His ont été incubées ensembles avec IPP et DMAPP. Le mélange

20 réactionnel a été extrait au pentane et analysé par GC-MS. **Figure 5A** : Chromatogramme du produit obtenu *in vitro* de la *Sh-zFPS* et *Sh-SBS* Le pic 2 est le produit majoritaire et correspond à l'alpha-santalène.. 1, *cis*-alpha-bergamotène ; 2, alpha-santalène; 3, *trans*-alpha-bergamotène; 4, épi-beta-santalène; 5, endo-beta-bergamotène. **Figure 5B** : Spectre de masse correspondant au pic n° 2 (à gauche) et structure de l'alpha-santalène (à droite).

25

Figure 6 : Superposition des profils GC entre les produits obtenus *in vitro* avec les enzymes recombinantes *Sh-zFPS*-6His et *Sh-SBS*-6His, et l'exsudat de la lignée recombinante isogénique de tomate TA517. Les tracés correspondent à l'ion extrait m/z 94. En gris: oléfines produites *in vitro* par *Sh-zFPS* et *Sh-KS*. En noir: exsudat de la lignée

30 isogénique TA517. 1, *cis*-alpha-bergamotène 2, alpha-santalène; 3, *trans*-alpha-bergamotène ; 4, épi-beta-santalène; 5, endo-beta-bergamotène.

Figure 7 : Profil GC/MS des molécules volatiles émises par une plante transgénique (#3877) portant les transgènes *Sh-zFPS* et *Sh-SBS*. La plante transgénique #3877 a été

cultivée pendant 24 heures en atmosphère contrôlée dans une enceinte de culture. Les molécules volatiles émises par les plantes ont été piégées sur une matrice Super® Q (Alltech) et analysées par GC/MS. Le chromatogramme (7A) montre plusieurs pics ayant des signatures m/z caractéristiques des sesquiterpènes de type SB. Les pics ont été
 5 identifiés par comparaison de leur temps de rétention et de leur spectre de masse avec ceux de la lignée TA517 (7B). 1, cis-alpha-bergamotène 2, alpha-santalène; 3, trans-alpha-bergamotène ; 4, épi-beta-santalène; 5, endo-beta-bergamotène.

Figure 8 : Polymorphisme du gène *zFPS* entre *S. lycopersicum* (*Sl*) et *S. habrochaites* (*Sh*):

10 Le gène complet *zFPS* a été amplifié par PCR à partir d'ADN génomique de *S. lycopersicum*, *S. habrochaites* et TA517. La séparation des produits PCR sur gel d'agarose à 0,8% permet d'identifier 3 bandes A, B, C d'une taille d'environ 3000, 2500 et 2000 nucléotides respectivement. Le produit PCR correspondant à la bande B est spécifique du génome *S.h* et correspond au gène *zFPS* présenté dans ce document. kb : kilobases

15

Abréviations et Définitions

ARNm : acide ribonucléique messenger

ADN : acide désoxyribonucléique

20 ADNc : ADN complémentaire produit par transcription inverse à partir d'ARNm.

CaMV : virus de la mosaïque du chou-fleur

DMAPP : diméthylallyl diphosphate

CBT-ol : cembratrien-ol

CBTS : cembratrien-ol synthase

25 GPP : Geranyl diphosphate

FPP : Farnesyl diphosphate

GC : chromatographie gazeuse

GGPP : géranyl géranyl diphosphate

ihpRNAi : intron hairpin ARN interférent

30 IPP : isopentenyl diphosphate

MS : spectrométrie de masse

PCR : "polymerase chain reaction"

RACE : "rapid amplification of cDNA ends"

A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V : Code standard des acides aminés selon la nomenclature universelle

www3.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?mode=t

A, C, G, T, B, D, H, K, M, N, R, S, V, W, Y : Code standard des bases selon la nomenclature

5 universelle www3.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?mode=t

RFLP : « Restriction fragment length polymorphism »

Par « Z,Z-farnesyl diphosphate synthase » ou zFPS est entendue une enzyme capable de produire du Z,Z-farnesyl diphosphate ((2Z,6Z)-farnesyldiphosphate ou *cis,cis-FPP*), à
10 partir d'isopentenyl diphosphate (IPP) et de diméthyl allyl diphosphate (DMAP). Par « activité Z,Z-FPS » est entendue la production de Z,Z- Farnesyl diphosphate à partir de IPP et DMAPP. L'activité Z,Z-FPS peut être mesurée en mesurant l'apparition du Z,Z-farnesyl diphosphate.

15 Par « SB synthase » ou SBS est entendue une enzyme capable de produire un mélange de sesquiterpènes principalement composé d'alpha-santalène, d'épi-beta-santalène, de cis-alpha-bergamotène, de trans-alpha-bergamotène et de endo-beta-bergamotène à partir de Z,Z-FPP. Par « activité de type SB » est entendue la production de sesquiterpènes de type SB à partir de Z,Z- Farnesyl diphosphate. L'activité de type SB peut être mesurée en
20 mesurant l'apparition d'un ou plusieurs sesquiterpènes de type SB. Un exemple de mesure d'activité de type SB est décrit dans les exemples.

Par « sesquiterpène de type SB » est entendue un mélange de sesquiterpènes principalement composé d'alpha-santalène, d'épi-beta-santalène, de cis-alpha-
25 bergamotène, de trans-alpha-bergamotène et de endo-beta-bergamotène.

« Conditions stringentes d'hybridation » : Généralement, des conditions stringentes, pour un polynucléotide de taille et de séquence données, sont obtenues en opérant à une température inférieure d'environ 5°C à 10°C à la température de fusion (T_m) de l'hybride
30 formé, dans le même mélange réactionnel, par ce polynucléotide et son complémentaire. Des conditions stringentes d'hybridation pour un polynucléotide donné, peuvent être identifiées par l'homme du métier en fonction de la taille et de la composition en bases du polynucléotide concerné, ainsi que de la composition du mélange d'hybridation

(notamment pH et force ionique). Des conditions de forte stringence incluent une étape de lavage avec du tampon 0,2 x SSC à 65°C.

Les conditions d'hybridations de stringences décrites ci-dessus peuvent être adaptées par l'homme du métier pour les polynucléotides de taille plus grande ou plus petite, selon les enseignements appropriés connus de l'homme du métier, notamment décrit dans Sambrook *et al.*, 1989, Molecular cloning, A laboratory Manual, Cold Spring Harbor ; Maniatis *et al.*, 1982, Molecular cloning, A laboratory Manual, Cold Spring Harbor Lab. CSH, N. Y. USA, ou l'une de ses récentes rééditions, et dans Ausubel *et al.*, Eds., 1995, Current Protocols in Molecular Biology, Chapter 2 (Greene Publishing and Wiley-Interscience, N. Y.).

Par « hétérologue » est entendu que le gène a été introduit par génie génétique dans la cellule. Il peut y être présent sous forme épisomale ou chromosomique. L'origine du gène peut être différente de la cellule dans laquelle il est introduit. Cependant, le gène peut également provenir de la même espèce que la cellule dans laquelle il est introduit mais il est considéré comme hétérologue en raison de son environnement qui ne sera pas naturel. Par exemple, le gène est dit hétérologue car il est sous le contrôle d'un promoteur autre que son promoteur naturel, il est introduit dans un endroit différent de celui-ci où il est situé naturellement. La cellule hôte peut contenir une copie du gène endogène au préalable de l'introduction du gène hétérologue ou elle peut ne pas contenir de copie endogène.

Par « pourcentage d'identité » entre deux séquences d'acide nucléique ou d'acides aminés au sens de la présente invention, on entend désigner un pourcentage de nucléotides ou de résidus d'acides aminés identiques entre les deux séquences à comparer, obtenu après le meilleur alignement, ce pourcentage étant purement statistique et les différences entre les deux séquences étant réparties au hasard et sur toute leur longueur. Le meilleur alignement ou alignement optimal est l'alignement pour lequel le pourcentage d'identité entre les deux séquences à comparer, comme calculé ci-après, est le plus élevé. Les comparaisons de séquences entre deux séquences d'acide nucléique ou d'acides aminés sont traditionnellement réalisées en comparant ces séquences après les avoir alignées de manière optimale, ladite comparaison étant réalisée par segment ou par fenêtre de comparaison pour identifier et comparer les régions locales de similarité de séquence. L'alignement optimal des séquences pour la comparaison peut être réalisé, outre manuellement, au moyen de l'algorithme d'homologie locale de Smith et Waterman (1981)

(Ad. App. Math. 2 : 482), au moyen de l'algorithme d'homologie locale de Neddleman et Wunsch (1970) (J. Mol. Biol. 48 : 443), au moyen de la méthode de recherche de similarité de Pearson et Lipman (1988) (Proc. Natl. Acad. Sci. USA 85 : 2444), au moyen de logiciels informatiques utilisant ces algorithmes (GAP, BESTFIT, FASTA et TFASTA dans le Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI). Le pourcentage d'identité entre deux séquences d'acide nucléique ou d'acides aminés est déterminé en comparant ces deux séquences alignées de manière optimale par fenêtre de comparaison dans laquelle la région de la séquence d'acide nucléique ou d'acides aminés à comparer peut comprendre des additions ou des délétions par rapport à la séquence de référence pour un alignement optimal entre ces deux séquences. Le pourcentage d'identité est calculé en déterminant le nombre de positions identiques pour lesquelles le nucléotide ou le résidu d'acide aminé est identique entre les deux séquences, en divisant ce nombre de positions identiques par le nombre total de positions dans la fenêtre de comparaison et en multipliant le résultat obtenu par 100 pour obtenir le pourcentage d'identité entre ces deux séquences.

Description Détaillée de l'Invention

La présente invention décrit donc pour la première fois les gènes codant pour les enzymes impliquées dans la voie de synthèse de plusieurs sesquiterpènes. En particulier, la présente invention concerne la caractérisation d'une sesquiterpène synthase de tomate à produits multiples (SB synthase) qui permet la production d'un mélange de sesquiterpènes dénommés composés de type SB et principalement composé d'alpha-santalène, d'épi-beta-santalène, de cis-alpha-bergamotène, de trans-alpha-bergamotène et de endo-beta-bergamotène à partir de Z,Z-FPP. En outre, elle concerne également la caractérisation de la Z,Z-farnesyl diphosphate synthase (zFPS) de tomate qui permet la production de Z,Z-FPP à partir d'IPP et de DMAP. Ces enzymes peuvent être utilisées pour la production de sesquiterpènes de type SB ou de Z,Z-FPP, *in vitro* ou *in vivo* dans des organismes transgéniques ou des cellules ou micro-organismes recombinants. Les organismes transgéniques, cellules ou micro-organismes tels que les bactéries, les levures, les champignons, les cellules animales, d'insectes ou de plantes et les plantes ou animaux transgéniques sont considérés. Elles permettent également la production de composés dérivés du Z,Z-FPP comme le Z,Z-farnesol. Les procédés de la présente invention sont plus particulièrement applicables à la production de Z,Z-FPP dans des microorganismes (bactérie, levure) et de composés de type SB dans les plantes possédant des trichomes

glanduleux de type sécréteurs. La production du mélange de sesquiterpènes de type SB ou de leurs dérivés peut également être diminuée ou supprimée dans des plantes, par exemple la tomate, par des techniques d'extinction de gènes ou par mutagenèse. En outre, les séquences identifiées dans la présente invention sont utiles pour l'identification et/ou le clonage de gènes codant pour des enzymes ayant les mêmes activités dans d'autres espèces d'organismes, en particulier de plantes. Enfin, des marqueurs moléculaires polymorphiques peuvent être identifiés à partir des acides nucléiques codant pour la zFPS et la SB synthase et permettront de suivre l'introduction des séquences génomiques fonctionnelles correspondantes dans d'autres espèces ou variétés de tomate cultivées (*Solanum lycopersicum*).

Le Z,Z-FPP est un substrat potentiel de sesquiterpènes synthases qui n'est pas disponible actuellement commercialement. Si la majorité des sesquiterpène synthases caractérisées à ce jour utilise le E,E-FPP, il est évident par ailleurs que l'absence de Z,Z-FPP disponible commercialement a très largement limité son utilisation à des fins expérimentales. La cadinène synthase de coton est le seul exemple qui décrit l'utilisation préférentielle du Z,Z-FPP par une sesquiterpène synthase (Heinstein *et al.*, 1970). Dans ce cas le Z,Z-FPP a été produit par synthèse chimique, et la Z,Z-FPP synthase décrite dans ce brevet fournit un moyen de produire cette molécule de manière aisée et peu coûteuse.

Par ailleurs, les sesquiterpènes de type SB contiennent l'alpha-santalène, lui-même précurseur direct de l'alpha-santalol. L'alpha-santalol est un des composants caractéristiques de l'huile essentielle de bois de santal. L'huile de bois de santal est extrêmement prisée en parfumerie et son coût a très fortement augmenté depuis 10 ans en raison d'une surexploitation du bois de santal, principalement en Inde. Cette surexploitation fait peser une menace sur la viabilité de la ressource naturelle de bois de Santal. L'alpha-santalol peut être obtenu à partir de l'alpha santalène par une simple hydroxylation.

La présente invention concerne donc une méthode de production de sesquiterpènes de type SB ou de dérivés de ceux-ci à partir de Z,Z-FPP dans une cellule ayant une source de DMAP et IPPP, comprenant :

- a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression comprenant un premier gène codant pour une zFPS selon la présente invention et d'une construction portant une cassette d'expression comprenant un deuxième gène codant pour une SB synthase selon la présente invention;

- b) la culture de la cellule transformée dans des conditions appropriées pour l'expression desdits premier et deuxième gènes ; et,
 - c) facultativement, la récolte de sesquiterpènes de type SB ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.
- 5 Dans un mode de réalisation particulier, les sesquiterpènes de type SB produits sont récoltés à l'étape c). Dans un mode de réalisation préféré, les sesquiterpènes de type SB produits sont sélectionnés parmi l'alpha-santalène, l'épi-beta-santalène, l'alpha-bergamotène, le beta-bergamotène, et l'endo-beta-bergamotène. Dans un autre mode de
- 10 réalisation particulier, les sesquiterpènes de type SB produits sont les produits de départ pour obtenir d'autres composés d'intérêt comme l'alpha-santalol. Dans un mode de réalisation préféré, les dérivés de sesquiterpènes de type SB sont sélectionnés parmi l'alpha-santalol, l'épi-beta-santalol, le cis-alpha-bergamotol, le trans-alpha-bergamotol et l'endo-beta-bergamotol.
- 15 Dans un mode de réalisation particulier, la présente invention concerne une méthode de production de sesquiterpènes de type SB ou de dérivés de ceux-ci à partir de Z,Z-FPP dans une cellule ayant une source de DMAP et IPPP, comprenant :
- a) la fourniture d'une cellule recombinante comprenant un premier gène hétérologue codant pour une zFPS selon la présente invention et un deuxième gène hétérologue
 - 20 codant pour une SB synthase selon la présente invention ;
 - b) la culture de la cellule dans des conditions appropriées pour l'expression desdits premier et deuxième gènes ; et,
 - c) facultativement, la récolte de sesquiterpènes de type SB ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.
- 25
- Dans un mode de réalisation préféré, ladite cellule produit du Z,Z-FPP. Dans un autre mode de réalisation, le Z,Z-FPP est fourni à la cellule.
- Dans un mode de réalisation particulier, les deux cassettes d'expression sont portées par
- 30 une même construction. Dans un autre mode de réalisation, les deux cassettes d'expression sont portées par deux constructions distinctes.
- La présente invention concerne une protéine isolée ou recombinante ayant une activité Z,Z-FPP synthase. En particulier, elle concerne un polypeptide isolé ou recombinant ayant une

activité de Z,Z-FPP synthase et dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 2. De préférence, le polypeptide présente au moins 95% d'identité avec la SEQ ID No 2. Dans un mode de réalisation particulier, le polypeptide comprend ou consiste en la séquence SEQ ID No 2. Ce polypeptide peut également comprendre une séquence additionnelle permettant de faciliter la purification de l'enzyme, par exemple une séquence étiquette comprenant plusieurs acides aminés histidine consécutifs.

La présente invention concerne également un acide nucléique isolé comprenant une séquence nucléotidique codant un polypeptide ayant une activité de Z,Z-FPP synthase et dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 2 ou une séquence capable de s'hybrider à celui-ci dans des conditions stringentes. Un exemple d'acide nucléique est l'acide nucléique comprenant ou consistant en une séquence présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 1 ou une séquence complémentaire de celle-ci. La présente invention concerne un acide nucléique isolé capable de s'hybrider dans des conditions de forte stringence à un acide nucléique qui code un polypeptide ayant une activité de Z,Z-FPP synthase et la séquence dudit polypeptide présentant au moins 95% ou 98% d'identité avec la SEQ ID No 2. Dans un mode de réalisation particulier, l'acide nucléique comprend ou consiste en la séquence SEQ ID No 1. La présente invention concerne également une cassette d'expression, un vecteur, une cellule hôte ou un organisme transgénique comprenant un acide nucléique selon la présente invention.

Les acides nucléiques peuvent être des ADN génomiques, des ADN complémentaires (ADNc) ou des ADN synthétiques. Ils peuvent être sous forme simple chaîne ou en duplexe ou un mélange des deux. Dans le cadre de la présente invention, les acides nucléiques transcrits sont de préférence des ADNc dénués d'introns. Les acides nucléiques transcrits peuvent être des molécules synthétiques ou semi-synthétiques, recombinantes, éventuellement amplifiées ou clonées dans des vecteurs, modifiées chimiquement ou comprenant des bases non naturelles. Il s'agit typiquement de molécules d'ADN isolées, synthétisées par des techniques recombinantes bien connues en soi de l'homme du métier.

La présente invention concerne également un polypeptide isolé ou recombinant ayant une activité de SB synthase et dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%,

98% ou 99% d'identité avec la SEQ ID No 4. Dans un mode de réalisation particulier, le polypeptide comprend ou consiste en la séquence SEQ ID No 4. Dans ce cas, la présente invention concerne un polypeptide isolé ou recombinant ayant une activité de SB synthase et dont la séquence présente au moins 95% d'identité avec la SEQ ID No 4. Ce

5 polypeptide peut également comprendre une séquence additionnelle permettant de faciliter la purification de l'enzyme, par exemple une séquence étiquette comprenant plusieurs acides aminés histidine consécutifs.

La présente invention concerne également un acide nucléique isolé comprenant une

10 séquence nucléotidique codant un polypeptide ayant une activité de SB synthase et dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4 ou une séquence capable de s'hybrider à celui-ci dans des conditions stringentes. Un exemple d'acide nucléique est l'acide nucléique comprenant ou consistant en une séquence présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité

15 avec la SEQ ID No 3 ou une séquence complémentaire de celle-ci. La présente invention concerne un acide nucléique isolé capable de s'hybrider dans des conditions de forte stringence à un acide nucléique qui code un polypeptide ayant une activité de SB synthase et la séquence dudit polypeptide présentant au moins 95% ou 98% d'identité avec la SEQ ID No 4. Dans un mode de réalisation particulier, l'acide nucléique comprend ou

20 consiste en la séquence SEQ ID No 3. La présente invention concerne également une cassette d'expression, un vecteur, une cellule hôte ou un organisme transgénique comprenant un acide nucléique selon la présente invention.

La présente invention concerne tout particulièrement un vecteur, une cellule hôte ou un

25 organisme transgénique comprenant un acide nucléique comprenant une séquence hétérologue codant un polypeptide ayant une activité de SB synthase et dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4 et un acide nucléique comprenant une séquence hétérologue codant un polypeptide ayant une activité de Z,Z-FPP synthase et dont la séquence présente au moins 80%, 85%, 90%,

30 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 2. Dans un mode de réalisation particulier de l'invention, l'organisme transgénique est une plante, et en particulier une plante à trichomes.

La présente invention concerne également un vecteur, une cellule hôte ou un organisme transgénique comprenant un acide nucléique comprenant une séquence hétérologue codant un polypeptide ayant une activité de SB synthase et dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4.

5

La présente invention concerne en outre un vecteur, une cellule hôte ou un organisme transgénique comprenant un acide nucléique comprenant une séquence hétérologue codant un polypeptide ayant une activité de Z,Z-FPP synthase, de préférence dont la séquence présente au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No

10 2.

La présente invention concerne une méthode de production d'un mélange de sesquiterpènes de type SB ou de dérivés de ceux-ci à partir d'IPP et de DMAP comprenant la mise en contact d'IPP et de DMAP avec une Z,Z-FPP synthase recombinante ou purifiée
15 présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 3 et une SB synthase recombinante ou purifiée présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4 dans des conditions appropriées et la récolte d'un mélange de sesquiterpène de type SB ou de dérivés de celui-ci obtenus. Elle concerne l'utilisation d'une Z,Z-FPP synthase recombinante ou purifiée présentant au
20 moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 2 et d'une SB synthase recombinante ou purifiée présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4 pour la préparation d'un mélange de sesquiterpène de type SB ou de dérivés de celui-ci à partir de d'IPP et de DMAP.

25 La présente invention concerne une méthode de production de Z,Z-FPP ou de dérivés de celui-ci à partir d'IPP et de DMAP comprenant la mise en contact d'IPP et de DMAP avec une Z,Z-FPP synthase recombinante ou purifiée présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 2 dans des conditions appropriées et la récolte du Z,Z-FPP ou de dérivés de celui-ci obtenus. Facultativement, la méthode
30 comprend en outre l'incubation du Z,Z-FPP obtenu avec une phosphatase alcaline d'intestin de veau et la récolte du Z,Z-farnesol. La présente invention concerne l'utilisation d'une Z,Z-FPP synthase recombinante ou purifiée présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 2 pour la préparation de Z,Z-FPP ou de dérivés de celui-ci à partir d'IPP et de DMAP.

La présente invention concerne une méthode de production d'un mélange de sesquiterpènes de classe II ou de dérivés de ceux-ci à partir de Z,Z-FPP comprenant la mise en contact du Z,Z-FPP avec une SB synthase recombinante ou purifiée présentant au moins
5 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4 dans des conditions appropriées et la récolte d'un mélange de sesquiterpène de type SB ou de dérivés de ceux-ci obtenus. Elle concerne l'utilisation d'une SB synthase recombinante ou purifiée présentant au moins 80%, 85%, 90%, 95%, 97%, 98% ou 99% d'identité avec la SEQ ID No 4 pour la préparation d'un mélange de sesquiterpène de classe II ou de dérivés
10 de ceux-ci à partir de Z,Z-FPP.

De manière générale, une cassette d'expression comprend tous les éléments nécessaires à la transcription et traduction du gène en une protéine. Notamment, elle comprend un promoteur, éventuellement un amplificateur, un terminateur de transcription et les éléments
15 permettant la traduction.

Le promoteur est adapté à la cellule hôte. Par exemple, si la cellule est procaryote, le promoteur peut être sélectionné parmi les promoteurs suivants : LacI, LacZ, pLacT, ptac, pARA, pBAD, les promoteurs d'ARN polymérase de bactériophage T3 or T7, le
20 promoteur de la polyhédrine, le promoteur PR ou PL du phage lambda. Si la cellule est eucaryote et animale, le promoteur peut être sélectionné parmi les promoteurs suivants : promoteur précoce du cytomegalovirus (CMV), promoteur de la thymidine kinase du virus de l'herpes simple (herpes simplex virus, HSV), promoteur précoce ou tardif du virus simian 40 (SV40), le promoteur de la métallothionéine-L de souris, et les régions LTR
25 (long terminal repeat) de certains rétrovirus. De manière générale, pour le choix d'un promoteur adapté, l'homme du métier pourra avantageusement se référer à l'ouvrage de Sambrook et Russell (2000) ou encore aux techniques décrites dans l'ouvrage d'Ausubel *et al.* (2006)

30 Le vecteur peut être un plasmide, un phage, un phagemide, un cosmide, un virus, un YAC, un BAC, un plasmide pTi d'*Agrobacterium*, etc. Le vecteur peut comprendre de préférence un ou plusieurs éléments sélectionnés parmi une origine de répllication, un site de clonage multiple et un marqueur. Le marqueur peut être un gène rapporteur qui engendre un signal détectable. Le signal détectable peut être un signal fluorescent, une coloration, une

émission de lumière. Le marqueur pourra donc être GFP, EGFP, DsRed, la beta-galactosidase, la beta-glucosidase, la luciférase, etc. Le marqueur est de préférence un marqueur de sélection procurant à la cellule une résistance à un antibiotique ou un herbicide. Par exemple, le gène peut être un gène de résistance à la kanamycine, néomycine, etc. Dans un mode de réalisation préféré, le vecteur est un plasmide. Des exemples de vecteurs procaryotes figurent dans la liste suivante qui n'est pas exhaustive : pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pBR322, et pRIT5 (Pharmacia), pET (Novagen) et pQE-30 (QIAGEN). Des exemples de vecteurs eucaryotes figurent dans la liste suivante qui n'est pas exhaustive : pWLNEO, pSV2CAT, pPICZ, pcDNA3.1 (+) Hyg (Invitrogen), pOG44, pXT1, pSG (Stratagene); pSVK3, pBPV, pCI-nco (Stratagene), pMSG, pSVL (Pharmacia);. Les vecteurs viraux peuvent être de manière non-exhaustive des adénovirus, des AAV, des HSV, des lentivirus, etc. De préférence, le vecteur d'expression est un plasmide ou un vecteur viral.

La cellule hôte peut être un procaryote, par exemple *Escherichia coli*, *Bacillus subtilis*, *Streptomyces* sp, et *Pseudomonas* sp ou un eucaryote. L'eucaryote peut être un eucaryote inférieur comme une levure (par exemple, *Saccharomyces cerevisiae*) ou un champignon filamenteux (par exemple du genre *Aspergillus*) ou un eucaryote supérieur comme une cellule d'insecte, de mammifère ou de plante. La cellule peut être une cellule mammifère, par exemple une cellule COS, CHO (US 4,889,803 ; US 5,047,335). Dans un mode de réalisation particulier, la cellule est non-humaine et non-embryonnaire. La cellule peut être isolée, par exemple dans un milieu de culture. Les cellules peuvent également être comprises dans des organismes, par exemple des animaux non-humains transgéniques ou des plantes transgéniques.

La présente invention concerne donc une méthode de production de Z,Z-FPP ou de dérivés de celui-ci à partir d'IPP et de DMAP dans une cellule ayant une source d'IPP et de DMAP, comprenant :

- a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression comprenant un gène codant pour une Z,Z-FPP synthase selon la présente invention ;

- b) la culture de la cellule transformée dans des conditions appropriées pour l'expression du gène ; et,
- c) facultativement, la récolte du Z,Z-FPP ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.
- 5 Dans un mode de réalisation particulier, le Z,Z-FPP produit est récolté à l'étape c). Dans un autre mode de réalisation particulier, le Z,Z-FPP produit est le produit de départ pour obtenir un autre composé d'intérêt comme le Z,Z-farnesol. Par exemple, le Z,Z-FPP peut être transformé en Z,Z-farnesol par une phosphatase. Ainsi, la méthode peut comprendre en outre à l'étape a) l'introduction d'une construction portant une cassette d'expression
- 10 comprenant un gène codant pour une phosphatase, permettant ainsi la récolte de Z,Z-farnesol à l'étape c). La présente invention concerne donc une méthode de production de Z,Z-FPP ou de dérivés de celui-ci à partir d'IPP et de DMAP dans une cellule recombinante ayant une source d'IPP et de DMAP et comprenant un gène hétérologue codant pour une Z,Z-FPP synthase selon la présente invention. La présente invention
- 15 concerne ainsi une méthode de production de Z,Z-FPP dans une cellule recombinante ayant une source de DMAP et IPPP, comprenant :
- a) la fourniture d'une cellule recombinante comprenant un gène hétérologue codant pour une zFPS selon la présente invention ;
- b) la culture de la cellule dans des conditions appropriées pour l'expression dudit
- 20 gène ; et,
- c) facultativement, la récolte du Z,Z-FPP ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.

La présente invention concerne donc une méthode de production d'un mélange de

25 sesquiterpènes de type SB ou de dérivés de ceux-ci à partir de Z,Z-FPP dans une cellule ayant une source de Z,Z-FPP, comprenant :

- a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression comprenant un gène codant pour une SB synthase selon la présente invention ;
- 30 b) la culture de la cellule transformée dans des conditions appropriées pour l'expression du gène ; et,
- c) facultativement, la récolte d'un mélange de sesquiterpène de type SB ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.

- Dans un mode de réalisation particulier, le mélange de sesquiterpène de type SB produit est récolté à l'étape c). Dans un autre mode de réalisation particulier, le mélange de sesquiterpènes de type SB produit est le produit de départ pour obtenir d'autres composés d'intérêt comme l'alpha-santalol. La présente invention concerne donc une méthode de
- 5 production d'un mélange de sesquiterpène de type SB ou de dérivés de ceux-ci à partir de Z,Z-FPP dans une cellule recombinante ayant une source de Z,Z-FPP et comprenant un gène hétérologue codant pour une SB synthase selon la présente invention. La présente invention concerne ainsi une méthode de production de sesquiterpènes de type SB ou de dérivés de ceux-ci à partir de Z,Z-FPP dans une cellule recombinante ayant une source de
- 10 Z,Z-FPP, comprenant :
- a) la fourniture d'une cellule recombinante comprenant un gène hétérologue codant pour une SB synthase selon la présente invention ;
 - b) la culture de la cellule dans des conditions appropriées pour l'expression dudit gène ; et,
 - 15 c) facultativement, la récolte de sesquiterpènes de type SB ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.

La cellule peut également être partie d'un organisme multicellulaire, par exemple une plante ou un animal non-humain. Dans ce cas, la présente invention concerne une méthode

20 de préparation d'un organisme multicellulaire comprenant :

- a) l'introduction dans une cellule de l'organisme d'une construction portant une cassette d'expression comprenant un premier gène codant pour une Z,Z-FPP synthase selon la présente invention et d'une construction portant une cassette d'expression comprenant un deuxième gène codant pour une SB synthase selon la
- 25 présente invention ; et,
- b) la reconstitution d'un organisme à partir de ladite cellule et la sélection des organismes transgéniques exprimant les deux gènes introduits.

Dans un mode de réalisation, l'organisme est un animal non-humain. Par exemple, l'organisme peut être une souris, un rat, un cobaye, un lapin, etc. Dans un autre mode de

30 réalisation préféré, l'organisme est une plante, de préférence une plante à trichomes glanduleux.

La présente invention concerne donc une méthode de production d'un mélange de sesquiterpènes de type SB comprenant la fourniture d'un organisme multicellulaire

transgénique exprimant une Z,Z-FPP synthase selon la présente invention et une SB synthase selon la présente invention et la récolte d'un mélange de sesquiterpènes de type SB produit ou de dérivés de ceux-ci dans lesdits organismes transgéniques.

- 5 La méthode permet de faire produire un mélange de sesquiterpène de type SB à des organismes ne produisant pas ces composés ou d'augmenter la quantité d'un mélange de sesquiterpène de type SB produit par des organismes le produisant déjà.

La présente invention concerne une méthode de préparation d'un organisme multicellulaire
10 comprenant :

- a) l'introduction dans une cellule de l'organisme d'une construction portant une cassette d'expression comprenant un gène codant pour une Z,Z-FPP synthase selon la présente invention; et
- b) la reconstitution d'un organisme à partir de ladite cellule et la sélection des
15 organismes transgéniques exprimant le gène introduit.

La présente invention concerne également une méthode de production de Z,Z-FPP ou de dérivés de celui-ci comprenant la fourniture d'un organisme multicellulaire transgénique exprimant une Z,Z-FPP synthase selon la présente invention et la récolte du
20 Z,Z-FPP produit ou de dérivés de celui-ci dans lesdits organismes transgéniques.

Dans un mode de réalisation, l'organisme est un animal non-humain. Par exemple, l'organisme peut être une souris, un rat, un cobaye, un lapin, etc. Dans un autre mode de réalisation préféré, l'organisme est une plante, de préférence une plante à trichomes glanduleux. La méthode permet de faire produire du Z,Z-FPP à des organismes ne
25 produisant pas ce composé ou d'augmenter la quantité de Z,Z-FPP produit par des organismes.

La présente invention concerne une méthode de préparation d'un organisme multicellulaire
30 comprenant :

- a) l'introduction dans une cellule de l'organisme d'une construction portant une cassette d'expression comprenant un gène codant pour une SB synthase selon la présente invention ; et

- b) la reconstitution d'un organisme à partir de ladite cellule et la sélection des organismes transgéniques exprimant le gène introduit.

La présente invention concerne une méthode de production d'un mélange de sesquiterpène de type SB ou de dérivés de ceux-ci comprenant la fourniture d'un organisme multicellulaire transgénique exprimant une SB synthase selon la présente invention et la

5 récolte d'un mélange de sesquiterpène de type SB produit ou de dérivés de ceux-ci dans lesdits organismes transgéniques.

Dans un mode de réalisation, l'organisme est un animal non-humain. Par exemple, l'organisme peut être une souris, un rat, un cobaye, un lapin, etc. Dans un autre mode de

10 réalisation préféré, l'organisme est une plante, de préférence une plante à trichomes sécréteurs. La méthode permet de faire produire un mélange de sesquiterpènes de type SB à des organismes ne produisant pas ce composé ou d'augmenter la quantité d'un mélange de sesquiterpènes de type SB produit par des organismes.

15 L'invention est notamment applicable à toutes les plantes de familles possédant des trichomes glanduleux, par exemple les Astéracées (tournesol, etc.), Solanacées (tomate, tabac, pomme de terre, poivron, aubergine, etc.), Cannabacées (ex *Cannabis sativa*) et Lamiacées (menthe, basilic, lavande, thym, etc.). Elle est particulièrement adaptée aux plantes de la famille des Solanacées, telles que par exemple des genres *Nicotiana*,

20 *Solanum*, *Capsicum*, *Petunia*, *Datura*, *Atropa*, etc., et en particulier des genres *Solanum* et *Nicotiana* comme par exemple la tomate cultivée (*Solanum lycopersicum*), la tomate sauvage *Solanum habrochaites*, le tabac cultivé (*Nicotiana tabacum*), et le tabac sylvestre (*Nicotiana glauca*). De manière non limitative, l'invention peut s'appliquer aux plantes des genres suivants : *Populus*, *Nicotiana*, *Cannabis*, *Pharbitis*, *Apteria*, *Psychotria*,

25 *Mercurialis*, *Chrysanthemum*, *Polypodium*, *Pelargonium*, *Mimulus*, *Matricaria*, *Monarda*, *Solanum*, *Achillea*, *Valeriana*, *Ocimum*, *Medicago*, *Aesculus*, *Plumbago*, *Pityrogramma*, *Phacelia*, *Avicennia*, *Tamarix*, *Frankenia*, *Limonium*, *Foeniculum*, *Thymus*, *Salvia*, *Kadsura*, *Beyeria*, *Humulus*, *Mentha*, *Artemisia*, *Nepta*, *Geraea*, *Pogostemon*, *Majorana*, *Cleome*, *Cnicus*, *Parthenium*, *Ricinocarpos*, *Hymenocallis*, *Larrea*, *Primula*, *Phacelia*,

30 *Dryopteris*, *Plectranthus*, *Cypripedium*, *Petunia*, *Datura*, *Mucuna*, *Ricinus*, *Hypericum*, *Myoporum*, *Acacia*, *Diplopeltis*, *Dodonaea*, *Halgania*, *Cyanostegia*, *Prostanthera*, *Anthocercis*, *Olearia*, *Viscaria*. De préférence, la plante est une plante de la famille des Asteracées, Cannabacées, Solanacées ou Lamiacées. Dans un mode de réalisation encore

plus préféré, la plante appartient aux genres *Solanum* ou *Nicotiana*, de préférence *Solanum esculentum*, *Solanum habrochaites*, *Nicotiana tabacum* ou *Nicotiana glauca*.

Dans ce mode de réalisation, les gènes sont placés sous le contrôle d'un promoteur permettant une expression, de préférence spécifique, dans les trichomes de la plante. De tels promoteurs existent et sont connus par l'homme du métier (Tissier *et al.*, 2004).

Au sens de l'invention, on entend par promoteur "spécifique" un promoteur principalement actif dans un tissu ou un groupe cellulaire donné. Il est entendu qu'une expression résiduelle, généralement plus faible, dans d'autres tissus ou cellules ne peut être entièrement exclue. Une caractéristique particulière de l'invention réside dans la capacité de construire des promoteurs spécifiques des cellules sécrétrices des trichomes glanduleux, permettant une modification de la composition des sécrétions foliaires de la plante, et notamment d'y exprimer les gènes de la présente invention permettant de préparer le mélange de sesquiterpène de type SB. Ainsi, le Z,Z-FPP, et le mélange de sesquiterpènes de type SB ou des dérivés de ceux-ci peuvent être récoltés au niveau des trichomes de ces plantes, en particulier dans l'exsudat des trichomes par extraction avec un solvant ou encore par distillation.

Par exemple, il a été montré qu'une séquence régulatrice de 1852 pb, située en amont de l'ATG du gène CYP71D16, permet de diriger l'expression du gène rapporteur *uidA* spécifiquement dans les cellules sécrétrices de trichomes de tabac (demande US2003/0100050 A1, Wagner *et al.*, 2003). D'autre part, plusieurs séquences promotrices extraites de différentes espèces ont été identifiées comme étant capables de diriger l'expression d'un gène hétérologue dans les trichomes de tabac.

Gène Abréviation	Nom du gène	Plantes	Références
LTP3	Lipid transfer protein	Cotton	Liu <i>et al.</i> , 2000, BBA, 1487:106111
LTP6		Cotton	Hsu <i>et al.</i> , 1999, Plant science, 143:6370
wax9D		B. oleracea	Pyee and Kolattukudy, 1995, Plant J. 7:4559
LTP1		Arabidopsis	Thoma <i>et al.</i> , 1994, Plant Physiol. 105 3545
CYC71D16	CBTol hydroxylase	Tobacco	Wang <i>et al.</i> , 2002, J.of Exp. Bot. 18911897

Dans un mode de réalisation préféré de la présente invention, le promoteur utilisé dans la cassette est dérivé des gènes NsTPS-02a, 02b, 03, et 04 de l'espèce *Nicotiana sylvestris* possédant une forte similarité de séquence avec CYC-2 (CBT-ol cyclase; NID : AF401234). Ces séquences promotrices sont plus amplement décrites dans la demande de
5 brevet WO2006040479 (Tissier *et al.*, 2004).

Parmi les séquences terminateurs préférées, on peut citer le terminateur NOS (Bevan *et al.*, 1983), et le terminateur du gène d'histone (EPO 633 317).

10 Dans un mode de réalisation particulier, la cassette d'expression peut comprendre une séquence permettant d'augmenter l'expression (« amplificateur »), par exemple certains éléments du promoteur CaMV35S et de gènes de l'octopine synthase (US 5 290 924). De préférence, les éléments activateurs du promoteur CaMV35S sont utilisés.

15 L'introduction des constructions portant un ou les deux gènes de l'invention dans une cellule ou un tissu végétal, y compris une graine ou plante, peut être réalisée par toute technique connue de l'homme du métier, y compris sous forme épisomale ou chromosomique. Les techniques de transgénèse végétale sont bien connues dans le domaine, et comprennent par exemple l'utilisation de la bactérie *Agrobacterium*
20 *tumefaciens*, l'électroporation, des techniques biolistiques, la transfection par un vecteur viral notamment, et toute autre technique connue de l'homme du métier.

Une technique communément utilisée repose sur l'utilisation de la bactérie *Agrobacterium tumefaciens*, qui consiste essentiellement à introduire la construction d'intérêt (acide
25 nucléique, cassette, vecteur, etc.) dans la bactérie *A. tumefaciens*, puis à mettre en contact cette bactérie transformée avec des disques de feuilles de la plante choisie. L'introduction de la cassette d'expression dans la bactérie est typiquement réalisée en utilisant comme vecteur le plasmide Ti (ou T-DNA), qui peut être transféré dans la bactérie par exemple par choc thermique ou par électroporation. L'incubation de la bactérie transformée avec les
30 disques foliaires permet d'obtenir le transfert du plasmide Ti dans le génome des cellules des disques. Ceux-ci peuvent éventuellement être cultivés dans des conditions appropriées pour reconstituer une plante transgénique dont les cellules comprennent la construction de l'invention. Pour plus de détails ou des variantes de mise en œuvre de la technique de

transformation par *A. tumefaciens*, on peut se référer par exemple à Horsch *et al.* (1985) ou Hooykaas and Schilperoort (1992).

5 Ainsi, dans un mode de réalisation particulier, la cassette d'expression ainsi constituée est insérée entre les bordures gauche et droite de l'ADN de transfert (T-DNA) d'un plasmide Ti désarmé pour le transfert dans des cellules végétales par *Agrobacterium tumefaciens*. Le T-DNA comprend également un gène dont l'expression confère une résistance à un antibiotique ou un herbicide et qui permet la sélection des transformants.

10 Une fois régénérées, les plantes transgéniques peuvent être testées pour l'expression hétérologue des gènes de la présente invention ou la production d'un mélange de sesquiterpènes de type SB, de Z,Z-FPP ou de Z,Z-farnesol dans les trichomes. Ceci peut être réalisé en récoltant l'exsudat des feuilles et en testant la présence du produit d'intérêt dans cet exsudat. L'analyse des volatiles émis par la plante peut également permettre
15 d'identifier les dits composés. Ceci peut également être fait en analysant la présence d'une ou des enzymes hétérologues de la présente invention dans les feuilles et, plus particulièrement, dans les cellules de trichome (par exemple en analysant les ARNm ou l'ADN génomique au moyen de sondes ou d'amorces spécifiques). Les plantes peuvent éventuellement être sélectionnées, croisées, traitées, etc. pour obtenir des plantes
20 présentant des niveaux d'expression améliorés.

Un autre objet de l'invention réside également dans une plante ou graine comprenant un acide nucléique, une cassette d'expression ou un vecteur tels que définis précédemment.

25 Un autre objet de l'invention concerne l'utilisation d'un acide nucléique de la présente invention comme marqueurs moléculaires pour permettre d'introduire les séquences génomiques correspondantes de *S. habrochaites* dans d'autres espèces ou variétés de tomate cultivée (*S. lycopersicum*) sexuellement compatibles. En effet, ces marqueurs permettent de contrôler l'introduction des séquences génomiques correspondantes de *S.*
30 *habrochaites* dans d'autres espèces ou variétés de tomate cultivée (*S. lycopersicum*) sexuellement compatibles et ainsi, par exemple, d'identifier et de sélectionner les individus issus d'un croisement entre l'espèce sauvage et une espèce cultivée dans lesquels l'introggression des acides nucléiques de la présente invention a eu lieu. L'utilisation des acides nucléiques de la présente invention comme marqueurs moléculaires peut être

réalisée par toute technique connue de l'homme du métier. Par exemple, une technique communément utilisée repose sur l'identification d'un polymorphisme de bande (RFLP) entre deux génomes digérés par différentes enzymes de restriction par hybridation moléculaire avec les acides nucléiques de la présente invention. Un autre exemple, consiste
5 à rechercher un polymorphisme en utilisant la technique de PCR pour amplifier sur les génomes à comparer tout ou partie d'un acide nucléique de la présente invention à partir d'amorces complémentaires de l'acide nucléique de la présente invention pour identifier après amplification PCR un polymorphisme de taille entre les deux génomes. Le polymorphisme peut également être révélé après digestion desdits produits PCR par une ou
10 des enzymes de restriction qui feront apparaître des fragments de taille différente entre les génomes comparés.

Une fois le polymorphisme identifié, cette information est utilisée, par exemple, pour identifier les individus d'une population F2 issus d'un croisement entre *S. habrochaites* et
15 une variété de tomate dans laquelle l'introggression des acides nucléiques de la présente invention est souhaitée.

Un autre objet de l'invention concerne l'introduction des séquences génomiques de *S. habrochaites* codant pour la zFPS et la SB synthase dans des espèces non compatibles
20 sexuellement par fusion de cellules, notamment de protoplastes. Les techniques de fusion de protoplastes sont connus de l'homme du métier (Zimmermann *et al.*, 1981, Bates *et al.*, 1983), et permettent de créer des cellules hybrides contenant des chromosomes appartenant aux deux espèces fusionnées. Lors de la fusion des chromosomes ou fragments de chromosome des deux espèces sont éliminés. Les cellules hybrides ayant conservé le ou les
25 fragments chromosomiques contenant les gènes de la présente invention peuvent alors être sélectionnées par PCR en utilisant des amorces spécifiques du gène codant pour la zFPS (SEQ ID N°1) ou de celui codant pour la SB synthase (SEQ ID N°2) ou des deux à la fois.

La présente invention concerne enfin une cellule ou un organisme transgénique non-
30 humain, caractérisé en ce que la voie de synthèse d'un mélange de sesquiterpène de type SB est bloquée par inactivation du gène codant pour une Z,Z-FPP synthase selon la présente invention ou du gène codant pour une SB synthase selon la présente invention ou des deux gènes. L'expression de ces gènes peut être bloquée par de nombreuses techniques disponibles et connues par l'homme du métier. Les gènes peuvent être délétés, mutés (par

exemple par mutagenèse chimique par Ethyl Methane Sulfonate ou par rayonnements ionisants) ou interrompus (mutagenèse insertionnelle). Par ailleurs, le blocage de l'expression de ces gènes peut également être réalisé par extinction de gènes en exprimant un transcrit inhibiteur. Le transcrit inhibiteur est un ARN qui peut se présenter sous la
5 forme d'un ARN double brin, d'un ARN antisens, d'un ribozyme, d'un ARN capable de former une triple hélice, et qui a une certaine complémentarité ou spécificité avec le transcrit du gène à bloquer.

La présente invention concerne également un acide nucléique capable de diminuer ou de
10 supprimer l'expression d'un gène codant une Z,Z-FPP synthase. En effet, un ARN inhibiteur qui peut se présenter sous la forme d'un ARN double brin, d'un ARN antisens, d'un ribozyme, d'un ARN capable de former une triple hélice, et qui a une certaine complémentarité ou spécificité avec le gène codant Z,Z-FPP synthase peut être préparé sur la base de l'enseignement de la présente invention. Ainsi, la présente invention concerne
15 un ARN inhibiteur du gène codant une Z,Z-FPP synthase comprenant au moins 21, 30, 40, 50, 60, 70, 80, 90 ou 100 nucléotides consécutifs de la SEQ ID No 1 ou d'une séquence complémentaire de celle-ci. De même, elle concerne l'utilisation d'un polynucléotide comprenant au moins 21, 30, 40, 50, 60, 70, 80, 90 ou 100 nucléotides consécutifs de la SEQ ID No 1 ou d'une séquence complémentaire de celle-ci pour inhiber l'expression du
20 gène codant une Z,Z-FPP synthase. L'inhibition de cette enzyme peut être utile pour stopper ou ralentir la voie de synthèse du mélange de sesquiterpène de type SB et rendre l'IPP et le DMAP disponible pour une autre voie de synthèse d'intérêt nécessitant un apport en Z,Z-FPP. Par exemple, l'inhibition de cette enzyme peut permettre d'augmenter le taux de production germacrene B par les cellules glandulaires des trichomes de feuilles
25 de tomate. La présente invention concerne donc des cellules ou des organismes transgéniques dans lesquels le gène codant une Z,Z-FPP synthase a été inactivé. L'inactivation peut être effectuée par des techniques d'ARN interférent, de délétion de gène, de mutation chimique ou d'inactivation par recombinaison homologue. Ainsi le taux de Z,Z-FPP ou du mélange de sesquiterpène de type SB peut être diminué dans ces cellules
30 ou organismes transgéniques.

En outre, la présente invention concerne également un acide nucléique capable de diminuer ou de supprimer l'expression d'un gène codant une SB synthase. En effet, un ARN inhibiteur qui peut se présenter sous la forme d'un ARN double brin, d'un ARN antisens,

d'un ribozyme, d'un ARN capable de former une triple hélice, et qui a une certaine complémentarité ou spécificité avec le gène codant une SB synthase peut être préparé sur la base de l'enseignement de la présente invention. Ainsi, la présente invention concerne un ARN inhibiteur du gène codant une SB synthase comprenant au moins 30, 40, 50, 60, 70, 80, 90 ou 100 nucléotides consécutifs de la SEQ ID No 3. De même, elle concerne l'utilisation d'un polynucléotide comprenant au moins 21, 30, 40, 50, 60, 70, 80, 90 ou 100 nucléotides consécutifs de la SEQ ID No 3 pour inhiber l'expression du gène codant une SB synthase. L'inhibition de cette enzyme peut être utile pour stopper la voie de synthèse d'un mélange de sesquiterpène de type SB et rendre le Z,Z-FPP disponible pour une autre voie de synthèse d'intérêt nécessitant un apport en ce composé, par exemple pour préparer des dérivés du Z,Z-FPP, par exemple, le Z,Z farnesol, ou d'autres sesquiterpènes qui sont le produit de synthases utilisant le Z,Z-FPP comme substrat.

La présente invention concerne donc des cellules ou des organismes transgéniques dans lesquels le gène codant une SB synthase est inactivé. L'inactivation peut être effectuée par des techniques d'ARN interférent, de délétion de gène, de mutation chimique ou d'inactivation par recombinaison homologue. Ainsi le taux du mélange de sesquiterpènes de type SB peut être diminué dans ces cellules ou organismes transgéniques.

La présente invention concerne également une méthode d'identification ou de clonage d'autres gènes codant une Z,Z-FPP synthase ou une SB synthase provenant d'une autre espèce dans laquelle une sonde présentant au moins 15, 30, 50, 75, 100, 200 ou 300 nucléotides consécutifs de la séquence SEQ ID No 1 ou 3, respectivement, est préparée et utilisée pour identifier ou sélectionner dans un échantillon une séquence nucléotidique capable de s'hybrider à ladite sonde. L'échantillon peut être par exemple une banque génomique ou de cDNA provenant d'un ou plusieurs organismes. La méthode peut comprendre une étape supplémentaire de caractérisation de la séquence identifiée ou sélectionnée. Cette étape supplémentaire peut comprendre le clonage, et/ou le séquençage, et/ou un alignement de séquence, et/ou un test d'activité enzymatique.

30

D'autres aspects et avantages de la présente invention apparaîtront à la lecture des exemples qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

Exemples**Synthèse d'ADNc de feuilles de tomate (*S. habrochaites*)**

Des ARNm ont été extraits à partir de feuilles de tomate (*Solanum habrochaites* = *L.*
 5 *Hirsutum LA1777*) en utilisant un kit d'extraction d'ARN totaux (RNeasy minikit, Qiagen). Les ADNc correspondants ont été synthétisés par transcription inverse à partir d'une amorce polyT en utilisant une transcriptase inverse (Roche, Cat. N° 03 531 317) dans les conditions réactionnelles préconisées par le fabricant.

10 **Clonage des séquences codantes pour la Z,Z-FPP synthase (*Sh-zFPS*) et la SB synthase (*Sh-SBS*) de tomate *Solanum habrochaites*.**

Les séquences nucléotidiques codantes et complètes pour les deux gènes *Sh-zFPS* et *Sh-SBS* ont été clonées à partir des informations publiques de séquences de la banque d'EST de trichomes de feuilles de tomate de *S. habrochaites* (Van der Hoeven, *et al.*, 2000, unpublished,) obtenu sur le site du SGN (« SOL Genomics Network »,
 15 <http://www.sgn.cornell.edu/>).

Les séquences de *Sh-zFPS* (SEQ ID N°1) et *Sh-SBS* (SEQ ID N°3) ont été obtenues par PCR à partir des ADNc de feuilles de tomate en utilisant les amorces 5'-ACCATGGGTTCTTTGGTTCTTCAATGTTGGA-3' (SEQ ID No 5) et 5'-ACTCGAGATATGTGTGTCCACCAAACGTCTATG-3' (SEQ ID No 6) pour la
 20 séquence ID N°1 et les amorces 5'-TCCATGGTAGTTGGCTATAGAAGCACAATCA-3' (SEQ ID No 7) et 5'-TCTCGAGCATAAATTCAAATTGAGGGATTAATGA-3' (SEQ ID No 8) pour la séquence ID N°3. Des sites d'enzymes de restriction (NcoI et XhoI) ont été ajoutés à l'extrémité 5' des amorces pour faciliter le clonage du cDNA en aval d'un promoteur. L'amplification par PCR a été réalisée avec 0,5 unité de Taq polymérase
 25 (Eurogentec) dans le tampon fourni par le fabricant. Le produit PCR a été purifié sur une colonne Qiaquick (Qiagen) et cloné par ligation avec une T4 DNA ligase (NEB) dans le vecteur de clonage pGEM-T (Promega). Les clones correspondant à la séquence attendue ont été sélectionnés après séquençage complet des inserts.

La séquence nucléotidique *Sh-zFPS* (SEQ ID N°1) a une longueur de 912 bases et code
 30 pour une protéine (SEQ ID N°2) de 303 acides aminés. Un domaine protéique conservé de type *cis* (Z)-isoprényles diphosphate synthases (*cis*-IPPS) ou également annoté undécaprényle synthase (UPPS) a été identifié « *in silico* » avec le système CCD (NCBI, Marchler-Bauer A, Bryant SH (2004). Les séquences similaires en acide aminé à *Sh-zFPS*

ont été recherchées dans la banque de protéine SwissProt et GenBank en utilisant le programme BlastP (Altschul *et al.*, 1997). Les meilleurs pourcentages d'identité ont été obtenus avec des séquences de plantes ayant une homologie avec des undécaprényles diphosphate synthases mais dont les fonctions n'ont pas été démontrées (Tab. 1). Par exemple, *Sh-zFPS* possède un pourcentage d'identité de 42% et 39% avec des UPPS putatives d'*Arabidopsis thaliana* BAA97345 et AAO63349 respectivement. On peut noter également que la séquence *Sh-zFPS* possède un très faible degré d'identité (24 %) avec la *Z,E-FPS* de *Micrococcus tuberculosis* (O53434).

10

Organisme	PID	Fonction	% identité
Arabidopsis thaliana	BAA97345	Putative UPPS	42
Arabidopsis thaliana	AAO63349	Putative UPPS	39
Arabidopsis thaliana	AAO42764	Putative UPPS	39
Anabaena sp. PCC 7120	DP58563	Putative UPPS	37
Oryza sativa	NP_915561	Putative UPPS	36
Synechococcus elongatus	Q8DI29	Putative UPPS	35
Micrococcus luteus	O82827	Putative UPPS	34
Gloeobacter violaceus	Q7NPE7	Putative UPPS	32
Micrococcus tuberculosis	P60479	UPPS	30
Micrococcus tuberculosis	O53434	<i>E,Z-FPS</i>	24

Tableau 1 : Séquences présentant le meilleur pourcentage d'identité (% identité) obtenu par recherche d'homologie avec le programme BlastP 2.2.13 (Altschul *et al.*, 1997) contre les banque de SwissProt et de GenBank avec la séquence *Sh-zFPS*.

La séquence nucléotidique *Sh-SBS* (SEQ ID N°3) a une longueur de 2334 bases et code pour une protéine (SEQ ID N°4) de 777 acides aminés. Un domaine protéique conservé de terpène cyclase (cd00684.2) a été identifié « *in silico* » avec le système CCD (NCBI, Marchler-Bauer A, Bryant SH (2004). Les séquences similaires en acide aminé à *Sh-SBS* ont été recherchées dans la banque de protéine SwissProt et GenBank en utilisant le programme BlastP (Altschul *et al.*, 1997). A l'exception d'une séquence de tabac de fonction inconnue (60% d'identité), les meilleurs homologues ont été obtenues avec des diterpènes synthases de plantes (Tab. 2) avec un pourcentage d'identité d'environ 40 à 45% avec des séquences codant pour des kaurène synthases de différentes plantes (*Arabidopsis thaliana*, *Cucurbita maxima*, *Stevia rebaudiana*, *Oryza sativa*). *Sh-SBS*

possède également des homologues plus faibles (26 à 33% d'identité) avec des terpènes synthases de fonction connue comme l'abietadiène synthase et la taxadiène synthase et l'alpha-bisabolène synthase.

5

Organisme	PID	Fonction	% identité
Nicotiana tabacum	AAS98912	Putative terpene synthase	60
Arabidopsis thaliana	AAC39443	Kaurene synthase	44
Lactuca sativa	BAB12441	Putative kaurene synthase	43
Cucurbita maxima	Q39548	Kaurene synthase	43
Medicago truncatula	ABE87878	Putative terpene synthase	43
Stevia rebaudiana	AAD34294	Kaurene synthase	42
Oryza sativa	AAQ72559	Kaurene synthase	41
Hordeum vulgare	AAT49066	Kaurene synthase	36
Abies grandis	AAC24192	Alpha-bisabolene synthase	33
Abies grandis	AAB05407	Abietadiene synthase	27
Taxus baccata	2211347A	Taxadiene synthase	26

Tableau 2 : Séquences présentant le meilleur pourcentage d'identité (% identité) obtenu par recherche d'homologie avec le programme BlastP 2.2.13 (Altschul *et al.*, 1997) contre les banques de SwissProt et de GenBank avec la séquence Sh-SBS.

Production des protéines recombinantes *Sh-zFPS* et *Sh-SBS*

- 10 Les séquences nucléiques codant pour *Sh-zFPS* et *Sh-SBS* (séquences ID N° 1 et 3 respectivement) ont été introduites séparément dans le vecteur d'expression pET-30b (Novagen). Pour permettre la purification des protéines recombinantes, les codons STOP de ces séquences ont été supprimés afin de créer une protéine de fusion avec une queue de six histidines à l'extrémité C-terminale. Les vecteurs *Sh-zFPS*-pET30b et *Sh-SBS*-pET30b
- 15 ainsi obtenus ont été introduits par électroporation dans des cellules d'*Escherichia coli* modifiées pour l'expression de protéines recombinantes (BL21-CodonPlus, Stratagene). Pour la production des protéines, les cultures (500 ml) de *E. coli* sont initiées à 37 °C jusqu'à une densité optique de 0,5 à 600 nm. A ce stade, la température de culture est abaissée à 16 °C et de l'éthanol (1%, v/v) est ajouté dans les milieux. Après une heure
- 20 d'incubation, l'IPTG est ajouté au milieu à une concentration de 1 mM pour induire l'expression de la protéine recombinante. La culture est alors incubée pour une durée de 18 heures. La culture est arrêtée par centrifugation des cellules à 4 °C. Le culot cellulaire est

repris dans un tampon phosphate 200mM. Les cellules sont lysées à la presse de French ($\Delta P = 19$ bar). Les lysats sont centrifugés à $15.000\times g$. Les surnageants contenant les protéines solubles sont dessalés sur colonne PD10 (Amersham), équilibrés avec un tampon phosphate (pH 7, 50 mM), puis appliqués sur une résine d'affinité nickel-nitrilotriacétique (Ni-NTA, Qiagen). Le protocole d'élution du fabricant est strictement suivi. Les fractions les plus enrichies en protéines Sh-zFPS-6His et Sh-SBS-6His sont identifiées par gel SDS-PAGE, puis dessalées sur PD10 équilibrées avec un tampon (Hepes 50 mM, pH 7.8, KCl 100 mM). Un facteur d'enrichissement d'environ 20 a été obtenu pour les deux protéines par rapport à la fraction protéique brute.

10 **Test d'activité enzymatique *in vitro* des protéines recombinantes Sh-zFPS et Sh-SBS**

Les tests d'activité sont réalisés dans un tampon ayant la composition suivante : Hepes 50 mM, pH 7.8, KCl 100 mM, $MgCl_2$ 7.5 mM, glycérol 5% (w/v), et DTT 5 mM. Pour les tests d'activité, 25 ou 50 μg de protéines provenant de la fraction enrichie ont été incubés avec les substrats IPP, DMAPP, GPP, FPP, GGPP (Sigma), seuls ou en combinaison, à une concentration de $65\mu M$ chacun à $32^\circ C$ pendant 2 heures.

Pour les tests réalisés avec Sh-zFPS-6His, les produits de la réaction sont ensuite déphosphorylés par ajout de 20 unités de phosphatase alcaline d'intestin de veau (New England Biolabs) pendant 1 heure à $37^\circ C$. Les terpènes alcools, produits de la déphosphorylation, sont ensuite extraits par un mélange chloroforme:méthanol:eau (0,5:1:0,4). Les phases aqueuses et organiques sont séparées par ajout d'eau (0,5 vol) et de chloroforme (0,5 vol), et centrifugées ($2.000\times g$). Les phases organiques sont prélevées et séchées sous courant d'azote gazeux, reprises dans du pentane et analysées par chromatographie gazeuse couplée à un spectromètre de masse (GC-MS 5973N, Agilent Technologies). Les terpènes alcools sont identifiés par comparaison de leur spectre de masse avec ceux de la base de données du National Institute of Standards and Technology (NIST), et par superposition des temps de rétention avec des standards de geraniol, farnesol et geranylgeraniol (Fluka).

Pour les tests réalisés avec Sh-SBS-6His, les terpènes produits dans la réaction sont directement extraits (à 3 reprises) des mélanges réactionnels avec du pentane (volume à volume). Les 3 extraits pentanes sont rassemblés, concentrés sous courant d'azote gazeux, et analysés par GC-MS. Les terpènes produits sont identifiés par interrogation de la base de

données NIST et comparaison avec un chromatogramme d'extrait d'exsudat de *Solanum habrochaites*.

La protéine recombinante Sh-zFPS-6His a été incubée avec les substrats IPP+DMAPP ou
 5 IPP+GPP (cf. tableau 1). Seule la première combinaison (IPP+DMAPP) conduit, après
 déphosphorylation, à la formation d'un terpène alcool unique de 15 atomes de carbones
 (C15), dont le spectre de masse permet de l'identifier comme un farnesol par comparaison
 avec le spectre standard présent dans la base de données NIST (figure 4). Le temps de
 10 rétention a été comparé à un standard de farnesol contenant un mélange des 4 isomères du
 farnesol (*E,E*-farnesol ; *E,Z*-farnesol ; *Z,E* farnesol et *Z,Z*-farnesol). Le farnesol produit par
 Sh-zFPS-6His a un temps de rétention identique à celui du *Z,Z*-farnesol

La protéine recombinante Sh-SBS-6His a été incubée avec les substrats GPP, FPP et
 GGPP. Quel que soit le substrat utilisé, aucun nouveau terpène n'a pu être détecté,
 signifiant que l'enzyme est inactive sur les substrats trans-allyliques (tableau 1).

Enz1	Enz2	Substrat	IPP	Produits
Sh-zFPS	absent	DMAP	IPP	A*
Sh-zFPS	absent	<i>E</i> -GPP	IPP	-
absent	Sh-SBS	<i>E</i> -GPP	No	-
absent	Sh-SBS	<i>E,E</i> -FPP	No	-
absent	Sh-SBS	<i>E,E,E</i> -GGPP	No	-
Sh-zFPS	Sh-SBS	DMAPP	IPP	B,C,D,E,F
Sh-zFPS	Sh-SBS	GPP	IPP	-

15 * : après déphosphorylation

Tableau 3 : Résumé des tests d'activité enzymatique obtenus *in vitro* avec les protéines recombinantes *Sh-zFPS-6His* et *Sh-SBS-6His* en fonction de différents substrats isopréniques utilisés. A, *Z,Z*-farnesol, B, alpha-santalène; C, epi-beta-santalène; D, endo-beta-bergamotène; E, cis-alpha-bergamotène; F, trans-alpha-bergamotène.

20 Les enzymes *Sh-zFPS-6His* et *Sh-SBS-6His* ont été incubées ensemble avec les substrats
 DMAPP et IPP ou GPP et IPP. Le produit de la réaction n'a pas été traité à la phosphatase,
 mais extrait avec du pentane dans lequel toute oléfine est soluble. Avec le DMAPP et l'IPP
 comme substrats, l'analyse chromatographique de l'extrait révèle la présence de plusieurs
 composés terpéniques (Fig. 5). Le profil GC a été comparé à celui de l'exsudat de la lignée
 25 de tomate TA517. Cette lignée est une lignée quasi-isogénique obtenue à partir des parents
S. lycopersicum et *S. habrochaites* (Monforte and Tanksley, 2000). La lignée TA517
 possède un fragment d'introgession de *S. habrochaites* responsable de la biosynthèse des

sesquiterpènes de classe II (van der Hoeven *et al.*, 2000). Les deux profils GC sont identiques ce qui confirme que les enzymes recombinantes ont une activité strictement identique à celles présentes dans la tomate (Fig. 6). Le composé majoritaire (Fig. 5, pic 2) est identifié comme l'alpha-santalène par comparaison de son temps de rétention et son spectre de masse avec celui de l'extrait d'exsudat de tomate TA517. Par ordre de quantités décroissantes, l'alpha-santalène (pic 4), l'endo-beta-bergamotène (pic 5), et l'alpha-bergamotène (pic 1) ont également été identifiés. L'enzyme Sh-SBS-6His est donc une enzyme dont l'activité conduit à la formation de plusieurs produits.

Il peut être conclu que Sh-zFPS est une enzyme qui agit sur l'IPP et le DMAPP pour produire un composé phosphorylé de type farnesyl diphosphate. Le temps de rétention est identique au standard (Z,Z)-farnesol indiquant que le farnesyldiphosphate produit par Sh-zFPS est caractérisé par une configuration de type Z,Z. L'inactivité de Sh-zFPS-6His sur le *E*-geranyl diphosphate en combinaison avec l'IPP indique que cette molécule n'est pas un produit intermédiaire de Sh-zFPS-6His et produit au moins une liaison de type Z en combinant un DMAPP et un IPP. Sh-SBS est actif uniquement sur le composé farnesyldiphosphate produit par l'enzyme Sh-zFPS, mais produit au moins 4 sesquiterpènes clairement identifiés dont le majoritaire est l'alpha-santalène. Enfin, la présence d'une extension en C-terminale par un *spacer* et 6 histidines, ne modifie en rien l'activité de l'enzyme par rapport aux enzymes natives de la plante d'origine (Fig. 3).

20 Expression des gènes Z,Z-FPP synthase (*Sh-zFPS*) et la SB synthase (*Sh-SBS*) dans le tabac

Un promoteur de tabac spécifique des trichomes (eCBTS02, Tissier *et al.*, 2004) a été cloné en amont des séquences codantes *Sh-zFPS* et *Sh-SBS* pour former les constructions eCBTS1.0-*Sh-zFPS* [#1] et eCBTS1.0-*Sh-SBS* [#2]. Ces deux constructions ont été introduites dans un même vecteur binaire pour la transformation par *Agrobacterium*, contenant un gène de résistance à la kanamycine pour donner le vecteur binaire pLIBRO-064 (Fig. 2A). La construction #2 a également été introduite de manière indépendante dans un vecteur binaire pLIBRO-065 (Fig. 2B). Les vecteurs ont été introduits dans la souche d'*Agrobacterium* LBA4404 par électroporation. Les souches contenant les différents vecteurs ont été utilisées pour réaliser la transformation génétique de *N. sylvestris* par la méthode des disques foliaires (Horsch *et al.*, 1985).

- Les T-DNA portant les transgènes ont été introduits par transformation génétique dans une lignée transgénique de *N. sylvestris* (lignée source ihpCBTS) dont l'expression du gène CBTS a été inhibée par ARN interférence (Tissier *et al.*, 2005). Dans cette lignée, la quantité de CBT-diol des feuilles est très réduite. *N. sylvestris*, un tabac sauvage ne produit pas de sesquiterpènes équivalents à ceux de la tomate. Environ 25 plantes résistantes à la kanamycine ont été obtenues pour chaque construction. La présence des transgènes dans chaque plante a été confirmée par PCR à partir d'ADN génomique de feuilles. Le niveau d'expression des transgènes dans les feuilles a été vérifié par PCR quantitative en temps réel à partir de mRNA de feuilles de tabac. Le niveau d'expression du transgène *Sh-SBS* a été comparé à celui d'un gène actine dont l'expression est constitutive dans les feuilles de tabac. Les résultats sont présentés sur la figure 3 en exemple. Ils indiquent que les lignées sélectionnées expriment toutes le transgène à des niveaux variables de 1 à 50 pour la construction pLIBRO-064 et de 1 à 10 pour la construction pLIBRO-065 par rapport au gène d'actine.
- Les sesquiterpènes oléfiniques étant des molécules volatiles, les molécules volatiles émises par les plantes ont été analysées en atmosphère contrôlée dans une enceinte de culture. Une partie de l'atmosphère a été captée dans un circuit de dérivation contenant un filtre SuperQ® (Altech) ayant la capacité d'adsorber les molécules terpéniques oléfiniques. Les molécules ont ensuite été éluées avec 1 ml de pentane et analysées par GC/MS. Un exemple de profil chromatographique des molécules volatiles émises par une plante transgénique est présenté sur la figure 7. Il montre que la lignée #3877 exprimant les gènes *Sh-zFPS* et *Sh-SBS* contient plusieurs nouveaux pics possédant une signature aux ions moléculaires 93, 94, 121 et 204 caractéristiques des sesquiterpènes de cette classe. Le profil chromatographique a été comparé à celui obtenu à partir de l'exsudat de la lignée de tomate TA517 qui produit naturellement ces composés (Fig. 7). Les deux profils sont identiques et les spectres de masse de chacun des pics sont identiques à ceux des pics de la lignée TA517. Les plantes qui expriment les deux transgènes *Sh-zFPS* et *Sh-SBS* produisent de l'alpha-santalène, de l'épi-beta-santalène, du cis-alpha-bergamotène, du trans-alpha-bergamotène et de l'endo-beta-bergamotène tandis que les lignées transgéniques ayant intégré le gène *Sh-SBS* seul ne produisent aucun de ces composés. Ces résultats indiquent que la présence des deux gènes est nécessaire à la synthèse des sesquiterpènes de type SB. Ils confirment les résultats obtenus *in vitro* avec les enzymes recombinantes. En conclusion, les gènes *Sh-zFPS* et *Sh-SBS* sont responsables de la

biosynthèse des sesquiterpènes de tomate localisés au niveau du locus Sst2 décrit par van der Hoeven *et al.*, 2000.

Polymorphisme génétique du gène *zFPS* entre *S. lycopersicum* et *S. habrochaites*.

Le gène complet *zFPS* a été amplifié par PCR à partir d'ADN génomique de *S.*
 5 *lycopersicum*, *S. habrochaites* et TA517 en utilisant les amorces
 ACCATGGGTTCTTTGGTTCTTCAATGTTGGA (SEQ ID No 9) et
 ACTCGAGATATGTGTGTCCACCAAAACGTCTATG (SEQ ID No 10). Deux produits
 ont été amplifiés dans les deux génomes indiquant qu'il existe au moins deux copies du
 gène dans les deux génomes (Fig. 8). L'un des deux produits qui possède une taille
 10 d'environ 3000 nucléotide (bande A) est commun aux deux génomes. Le second produit
 possède une taille d'environ 2000 nucléotides (bande C) chez *S. lycopersicum* et d'environ
 2500 nucléotides (bande B) chez *S. habrochaites*. La lignée TA517 montre un profil
 identique à celui de *S. habrochaites*. Ces résultats indiquent que le gène B spécifique de *S.*
habrochaites a été introduit dans la lignée TA517. Le séquençage du produit B a permis de
 15 confirmer qu'il correspond au gène *zFPS* présenté dans ce document. Ce résultat montre
 également que le gène *zFPS* de *S. habrochaites* peut être introduit dans *S. lycopersicum* en
 utilisant le polymorphisme qui existe entre les deux génomes pour ce gène.

Références

- 20 Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang,
 Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-
 BLAST: a new generation of protein database search programs", Nucleic Acids Res.
 25:3389-3402.
- Bates GW, Gaynor JJ, Shekhawat NS. (1983) Fusion of Plant Protoplasts by
 25 Electric Fields. Plant Physiol. 72:1110-1113..
- Bohlmann J, Crock J, Jetter R, Croteau R. (1998) Terpenoid-based defenses in
 conifers: cDNA cloning, characterization, and functional expression of wound-inducible
 (E)-alpha-bisabolene synthase from grand fir (*Abies grandis*). Proc Natl Acad Sci U S A.
 1998 95:6756-6761.
- 30 Bouwmeester HJ, Kodde J, Verstappen FW, Altug IG, de Kraker JW, Wallaart TE.
 (2002) Isolation and characterization of two germacrene A synthase cDNA clones from
 chicory. Plant Physiol. 29:134-144.
- Cai Y, Jia JW, Crock J, Lin ZX, Chen XY, Croteau R. (2002) A cDNA clone for
 beta-caryophyllene synthase from *Artemisia annua*. Phytochemistry. 61:523-529.

Cane DE. (1999) Sesquiterpene biosynthesis: Cyclization mechanisms. In comprehensive natural products chemistry, isoprenoids including carotenoids and steroids, Vol. 2, D.D. Cane, ed (Amsterdam, The Netherlands;Elsever), pp.155-200.

Carman RM, Duffield AR (1993) The Biosynthesis of Labdanoids - the Optical
5 Purity of Naturally-Occurring Manool and Abienol. Aust. J. Chem. 46 : 1105-1114.

Clarke CF, Tanaka RD, Svenson K, Wamsley M, Fogelman AM, Edwards PA (1987) Molecular cloning and sequence of a cholesterol-repressible enzyme related to prenyltransferase in the isoprene biosynthetic pathway. Mol. Cell. Biol 7: 3138-3146.

Colby SM, Crock J, Dowdle-Rizzo B, Lemaux PG, Croteau R. Germacrene C
10 synthase from *Lycopersicon esculentum* cv. VFNT cherry tomato: cDNA isolation, characterization, and bacterial expression of the multiple product sesquiterpene cyclase (1998) Proc Natl Acad Sci 95:2216-2221.

Facchini PJ, Chappell J. Gene family for an elicitor-induced sesquiterpene cyclase in tobacco (1992) Proc Natl Acad Sci 89(22):11088-11092.

Heinstein PF, Herman DL, Tove SB, Smith FH. Biosynthesis of gossypol.
15 Incorporation of mevalonate-2-¹⁴C and isoprenyl pyrophosphates (1970) J. Biol. Chem. 245: 4658-4665.

Horsch RB, Rogers SG, Fraley RT (1985) Transgenic plants. Cold Spring Harb Symp Quant Biol. 50: 433-437.

Kharel Y, Takahashi S, Yamashita S, Koyama T. (2006) Manipulation of prenyl
20 chain length determination mechanism of cis-prenyltransferases. FEBS J. 273:647-657.

Kollner TG, O'Maille PE, Gatto N, Boland W, Gershenzon J, Degenhardt J. (2006) Two pockets in the active site of maize sesquiterpene synthase TPS4 carry out sequential parts of the reaction scheme resulting in multiple products. Arch Biochem Biophys.
25 448:83-92.

Koyama T (1999) Molecular analysis of prenyl chain elongation enzymes. Biosci. Biotechnol. Biochem. 63:1671-1676.

Liang PH, Ko TP, Wang AH. (2002) Structure, mechanism and function of prenyltransferases. (2002) 269:3339-3354.

Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", Nucleic Acids Res.32(W)327-331.
30

McMillan J., and Beale M.H. (1999) Dipterene biosynthesis. In comprehensive natural products chemistry, isoprenoids including carotenoids and steroids Vol. 2, D.D. Cane, ed (Amsterdam, The Netherlands;Elsever), pp.217-243.

Monforte AJ, Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in an *L. esculentum* background: A tool for gene mapping and gene discovery. Genomic 43:803-813.

- 5 Oh SK, Han KH, Ryu SB, Kang H. (2000) Molecular cloning, expression, and functional analysis of a cis-prenyltransferase from *Arabidopsis thaliana*. Implications in rubber biosynthesis. J Biol Chem. 275:18482-18488.

Poulter DC (2006) Farnesyl Diphosphate Synthase. A Paradigm for Understanding Structure and Function Relationships in E-polyprenyl Diphosphate Synthases. Phytochem
10 review 5:17-26.

Rodriguez-Concepcion M, Boronat A. (2002) Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiol. 130(3):1079-89.

- Sharon-Asa L, Shalit M, Frydman A, Bar E, Holland D, Or E, Lavi U, Lewinsohn
15 E, Eyal Y. (2003) Citrus fruit flavor and aroma biosynthesis: isolation, functional characterization, and developmental regulation of *Cstps1*, a key gene in the production of the sesquiterpene aroma compound valencene. Plant J. 2003 36:664-74.

Shimizu N, Koyama T, Ogura K. (1998) Molecular cloning, expression, and purification of undecaprenyl diphosphate synthase. No sequence similarity between E- and
20 Z-prenyl diphosphate synthases. J Biol Chem. 273:19476-19481.

Schulbach MC, Brennan PJ, Crick DC. (2000) Identification of a short (C15) chain Z-isoprenyl diphosphate synthase and a homologous long (C50) chain isoprenyl diphosphate synthase in *Mycobacterium tuberculosis*. J Biol Chem. 275:22876-22881.

- Tarshis LC, Yan M, Poulter CD, Sacchettini JC. (1994) Crystal structure of
25 recombinant farnesyl diphosphate synthase at 2.6-A resolution. Biochemistry 33:10871-10877.

Tholl D. (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism Curr Opin Plant Biol. 9:297-304.

- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997)
30 The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 24:4876-4882.

Tissier A, Sallaud C, Rontcin D (2004) Promoteurs végétaux et utilisations PCT/FR 05/02530.

Tissier A, Sallaud C, Rontein D (2005) Système de production de terpènes dans les plantes Demande de brevet France FR 05 00855.

van Der Hoeven RS, Monforte AJ, Brecken D, Tanksley SD, Steffens JC. (2000) Free in PMC Genetic control and evolution of sesquiterpene biosynthesis in *Lycopersicon* 5 *esculentum* and *L. hirsutum*. Plant Cell. 12:2283-2294.

Wallaart TE, Bouwmeester HJ, Hille J, Poppinga L, Majers NC (2001) Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin. Planta. 212:460-465.

Wang K, Ohnuma SI (1999) Chain-length determination mechanism of isoprenyl 10 diphosphate synthases and implications for molecular evolution 24 :445-451.

Wise ML, Croteau RA (1999). Monoterpene biosynthesis. In comprehensive natural products chemistry, isoprenoids including carotenoids and steroids Vol. 2, D.D. Cane, ed (Amsterdam, The Netherlands;Elsevier), pp.97-153.

Zimmermann U, Scheurich P. (1981) High frequency fusion of plant protoplasts by 15 electric fields. Planta 151:26-32

Revendications

- 1- Méthode de production de sesquiterpènes de type SB ou de dérivés de ceux-ci à partir de Z,Z-FPP dans une cellule ayant une source de DMAP et IPP, comprenant :
- 5 a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression comprenant un premier gène codant pour une zFPS présentant au moins 80 % d'identité avec la SEQ ID No 2 et d'une construction portant une cassette d'expression comprenant un deuxième gène codant pour une SB synthase présentant au moins 80 % d'identité avec la SEQ ID No 4;
- 10 b) la culture de la cellule transformée dans des conditions appropriées pour l'expression desdits premier et deuxième gènes ; et,
- c) facultativement, la récolte de sesquiterpènes de type SB ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.
- 15 2- Protéine isolée ou recombinante ayant une activité Z,Z-FPP synthase et une séquence présentant au moins 80% d'identité avec la SEQ ID No 2.
- 3- Protéine selon la revendication 2, caractérisée en ce que sa séquence correspond à la SEQ ID No 2.
- 20 4- Protéine isolée ou recombinante ayant une activité de SB synthase et ayant une séquence présentant au moins 80 % d'identité avec la SEQ ID No 4.
- 5- Acide nucléique comprenant une séquence nucléotidique codant pour une protéine selon les revendications 2 à 4.
- 25 6- Cassette d'expression comprenant un acide nucléique selon la revendication 5.
- 7- Vecteur comprenant un acide nucléique selon la revendication 5 ou une cassette d'expression selon la revendication 6.
- 30 8- Cellule hôte comprenant un acide nucléique selon la revendication 5, une cassette d'expression selon la revendication 6, ou un vecteur selon la revendication 7.

9- Organisme transgénique non-humain comprenant un acide nucléique selon la revendication 5, une cassette d'expression selon la revendication 6, un vecteur selon la revendication 7, ou une cellule selon la revendication 8.

5 10- Organisme transgénique selon la revendication 9, caractérisé en ce que l'organisme est une plante à trichomes glanduleux, de préférence du genre *Nicotiana* ou *Solanum*.

10 11- Méthode de production du Z,Z-farnesyl diphosphate (Z,Z-FPP), ou de dérivés de ceux-ci à partir d'IPP et de DMAP dans une cellule ayant une source d'IPP et de DMAP, comprenant :

- a) l'introduction dans ladite cellule d'une construction portant une cassette d'expression avec une séquence nucléique codant une protéine selon la revendication 2 ou 3 ;
- b) la culture de la cellule transformée dans des conditions appropriées pour l'expression du gène ; et
- c) facultativement, la récolte du ou de dérivés de celui-ci contenus dans ladite cellule et/ou dans le milieu de culture.

20 12- Utilisation d'une protéine selon les revendications 2 à 4, d'un acide nucléique selon la revendication 5, d'une cellule hôte selon la revendication 8 ou d'un organisme transgénique selon la revendication 9 ou 10 pour la préparation de Z,Z-FPP, d'alpha-santalène, d'épi-beta-santalène, d'alpha-bergamotène, de beta-bergamotène, et d'endo-beta-bergamotène ou de dérivés de ceux-ci comme l'alpha-santalol, l'épi-beta-santalol, l'alpha-bergamotol, le beta-bergamotol et le Z,Z-farnesol.

25

13- Cellule ou organisme transgénique non-humain, caractérisé en ce que la voie de synthèse de sesquiterpène de type SB est bloquée par inactivation du gène codant pour une protéine selon la revendication 2 ou 3, ou du gène codant pour une protéine selon la revendication 4 ou des deux gènes.

30

14- Utilisation de marqueurs moléculaires comprenant tout ou partie d'un acide nucléique présentant au moins 80 % d'identité avec la SEQ ID No 1 ou SEQ ID No 3 pour identifier un polymorphisme génomique entre *S. habrochaites* et des espèces de type *Solanum*

WO 2008/142318

PCT/FR2008/050576

44

sexuellement compatible avec *S. habrochaites* afin de contrôler l'introggression des gènes correspondant dans ces espèces.

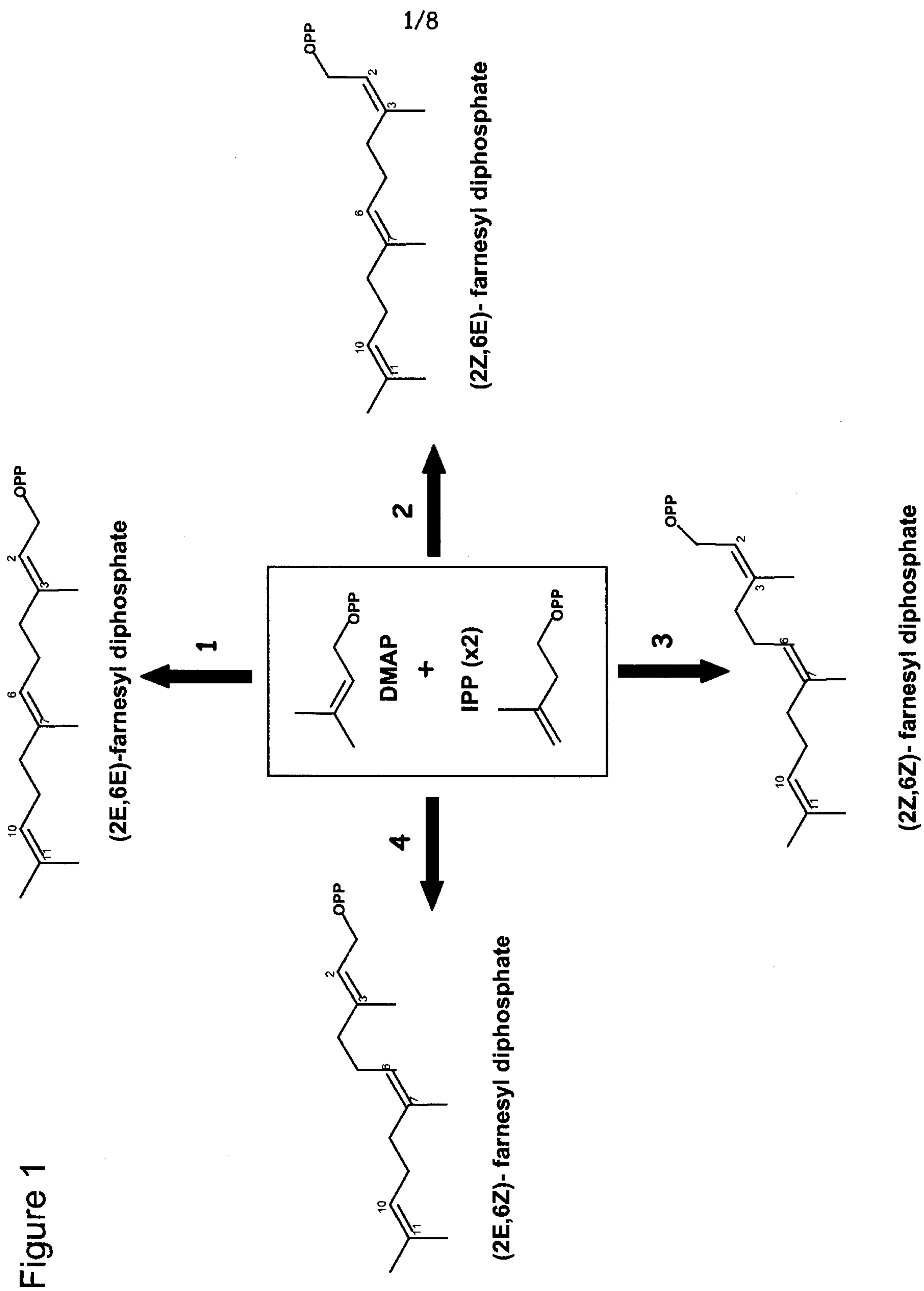
15. Utilisation selon la revendication 14, caractérisée en ce que les marqueurs moléculaires
5 comprennent des séquences nucléiques correspondant aux SEQ ID No. 9 et 10.

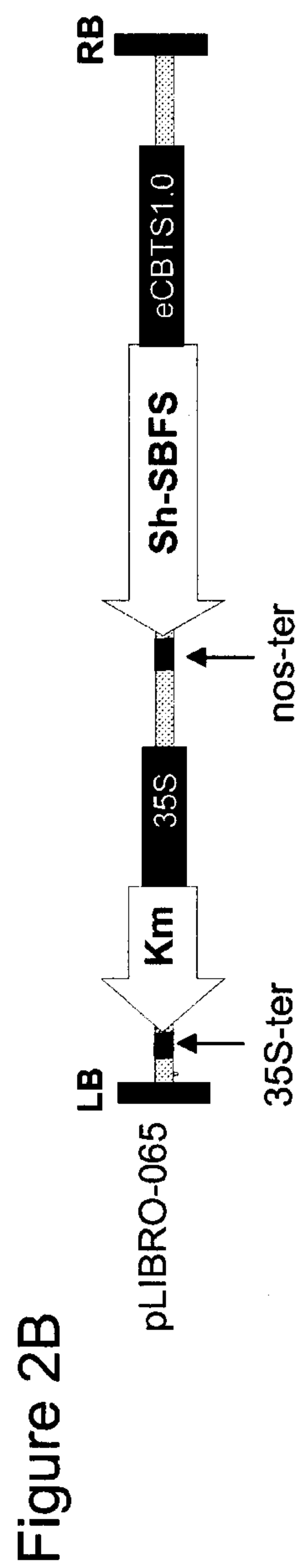
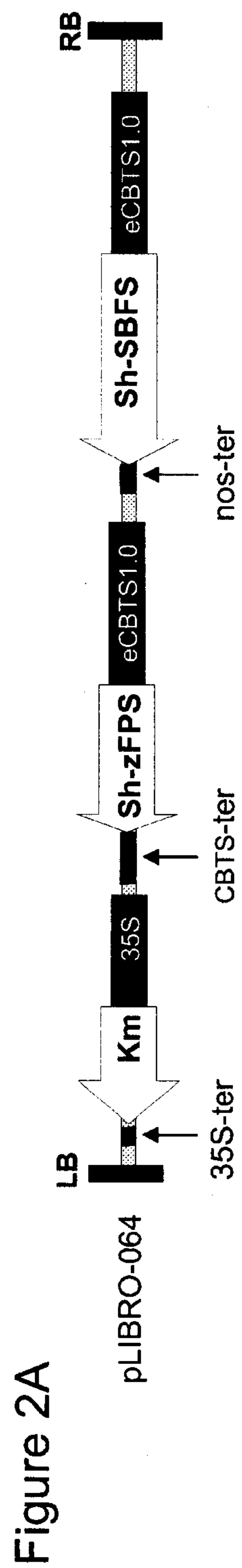
16-. Méthode consistant à introduire le gène *zFPS* codant pour une zFPS présentant au
moins 80 % d'identité avec la SEQ ID No 2 et le gène *SBS* codant pour une SB synthase
présentant au moins 80 % d'identité avec la SEQ ID No 4 dans une espèce non
10 sexuellement compatible avec *Solanum habrochaites* par fusion de protoplastes.

17- Méthode de production du Z,Z-farnesol par déphosphorylation du Z,Z-FPP obtenu
selon la méthode de la revendication 11 par une phosphatase.

WO 2008/142318

PCT/FR2008/050576





3/8

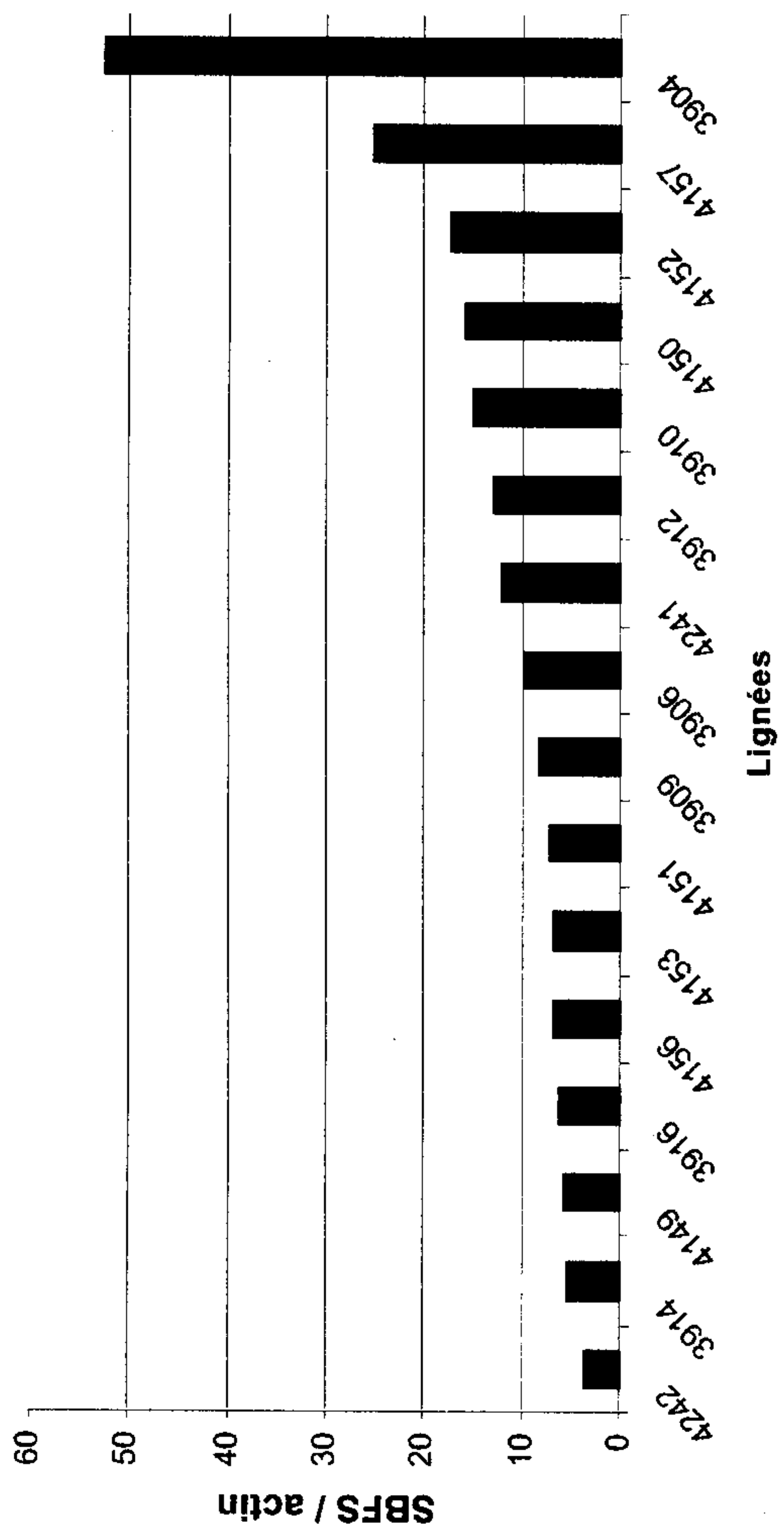


Figure 3A

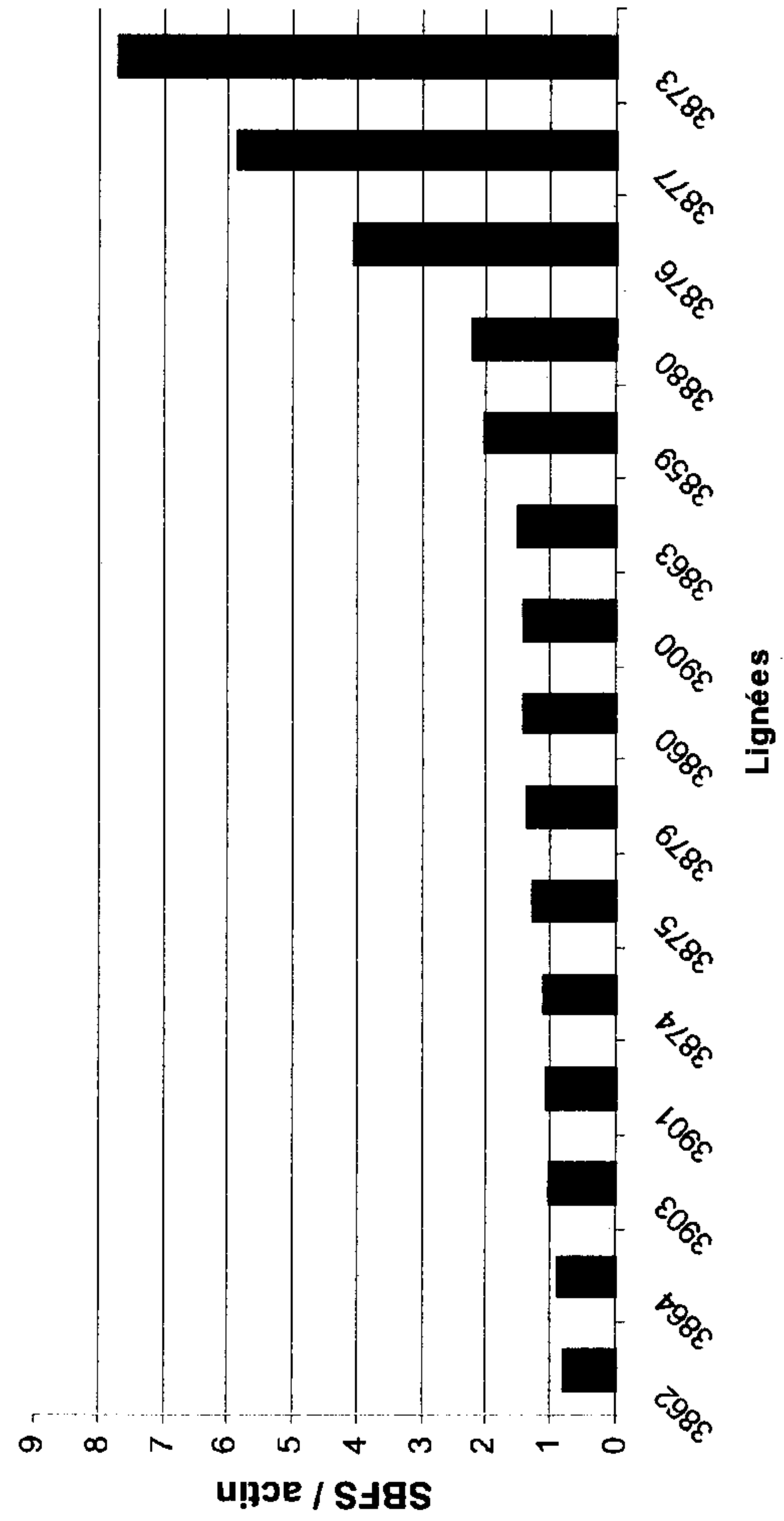


Figure 3B

Figure 4A

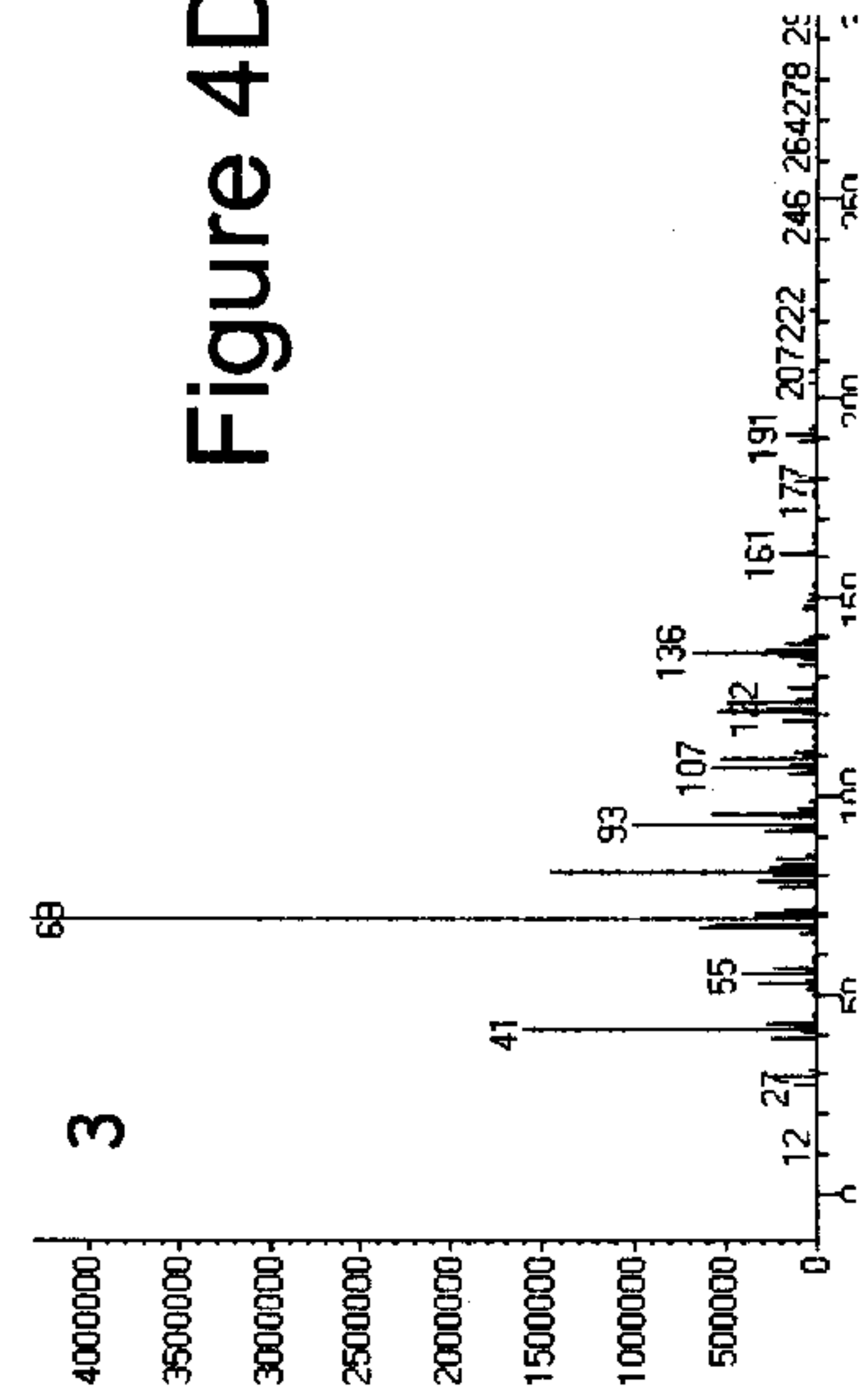
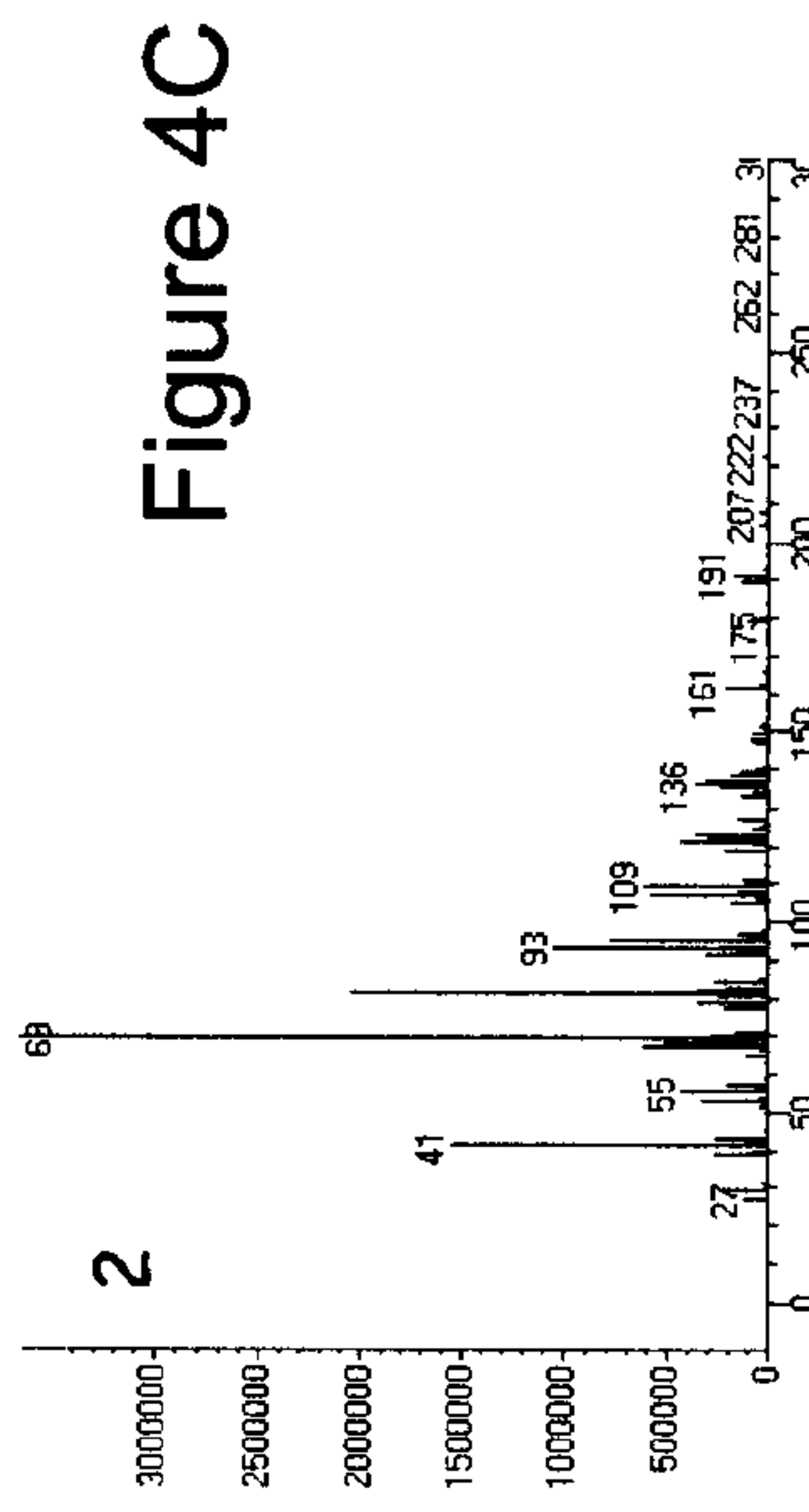
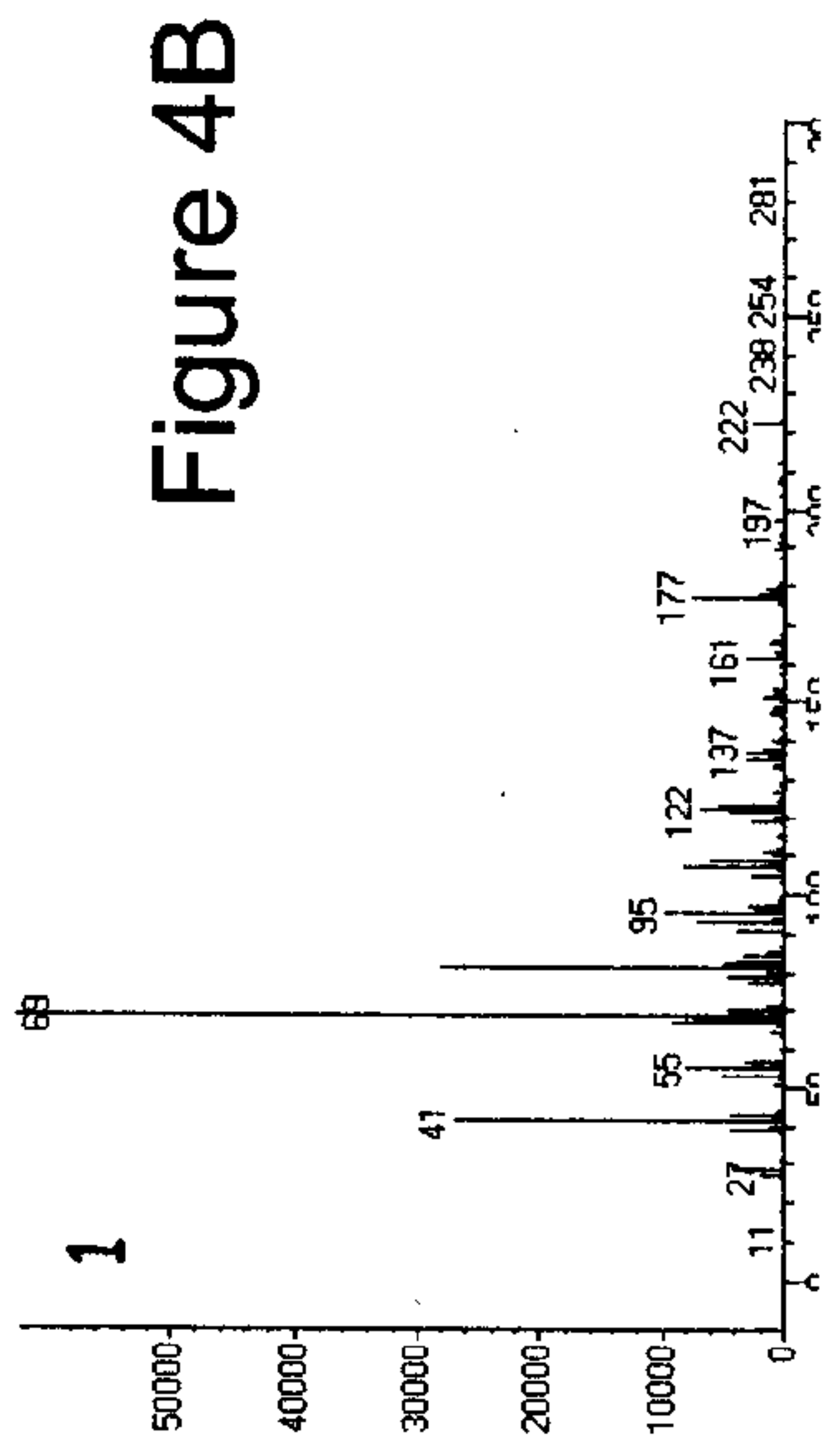
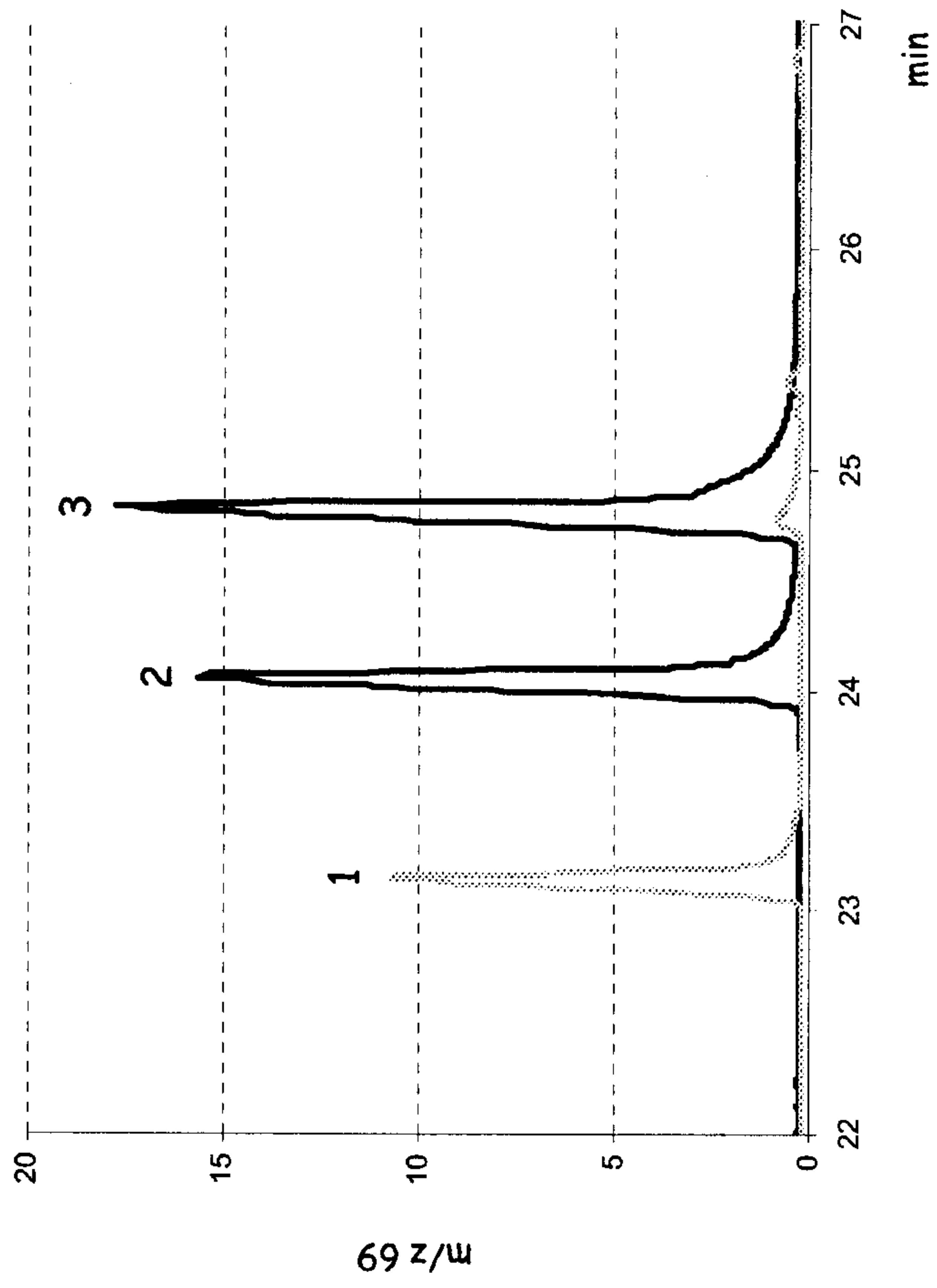
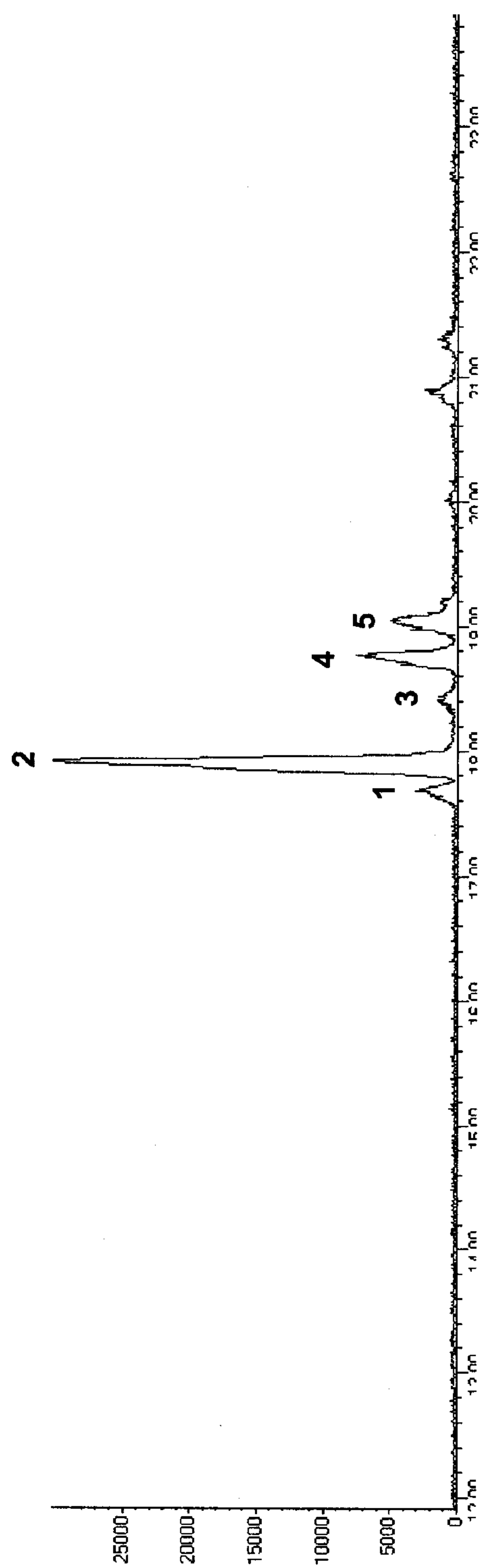
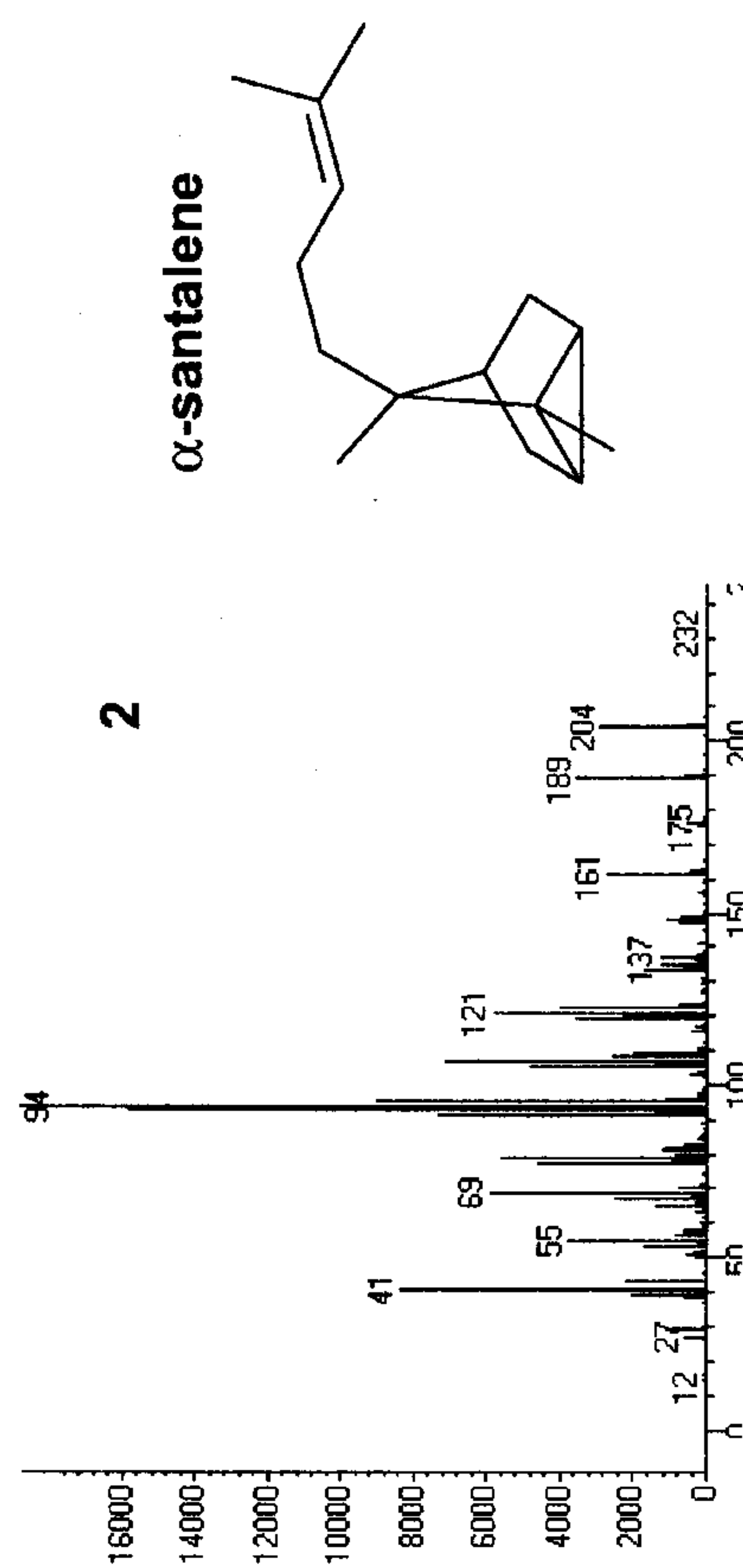
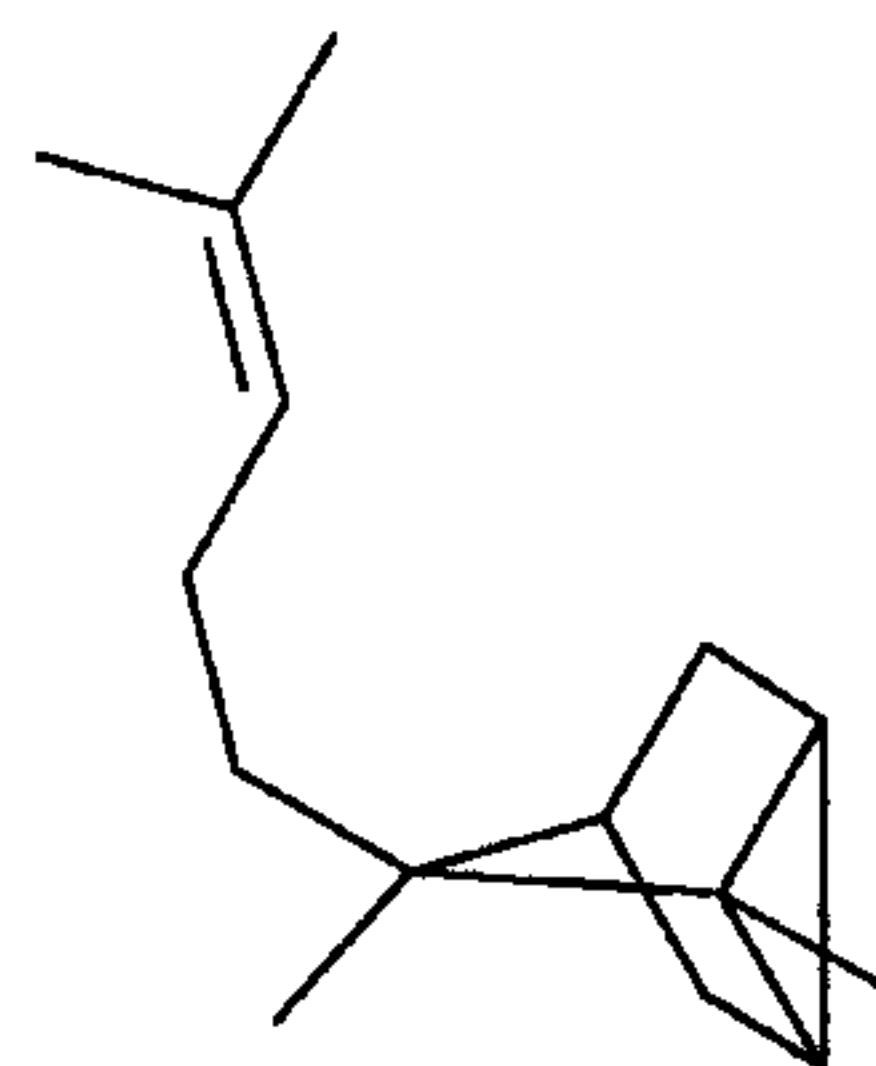


Figure 5A



5/8

Figure 5B

 α -santalene

6/8

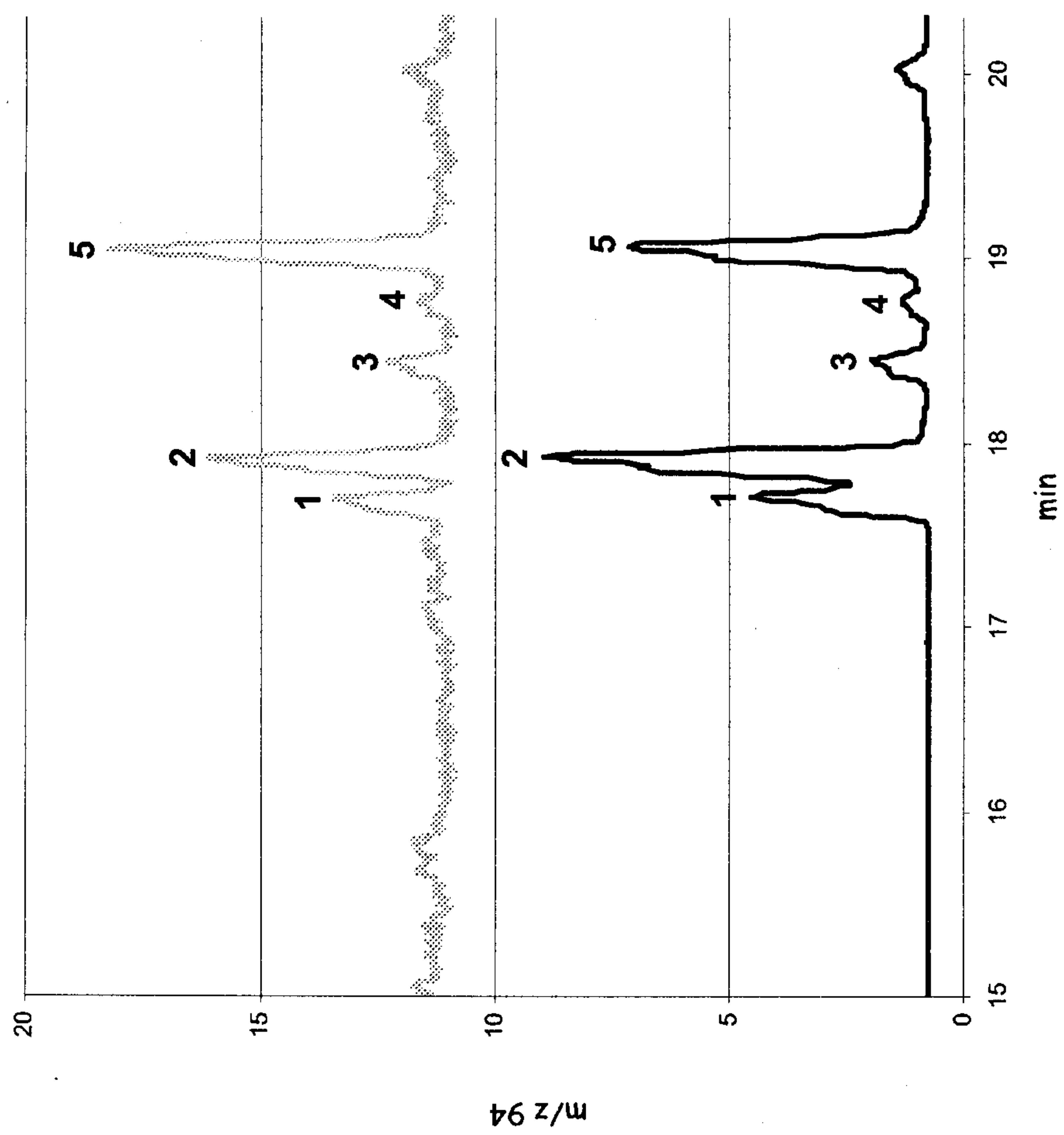


Figure 6

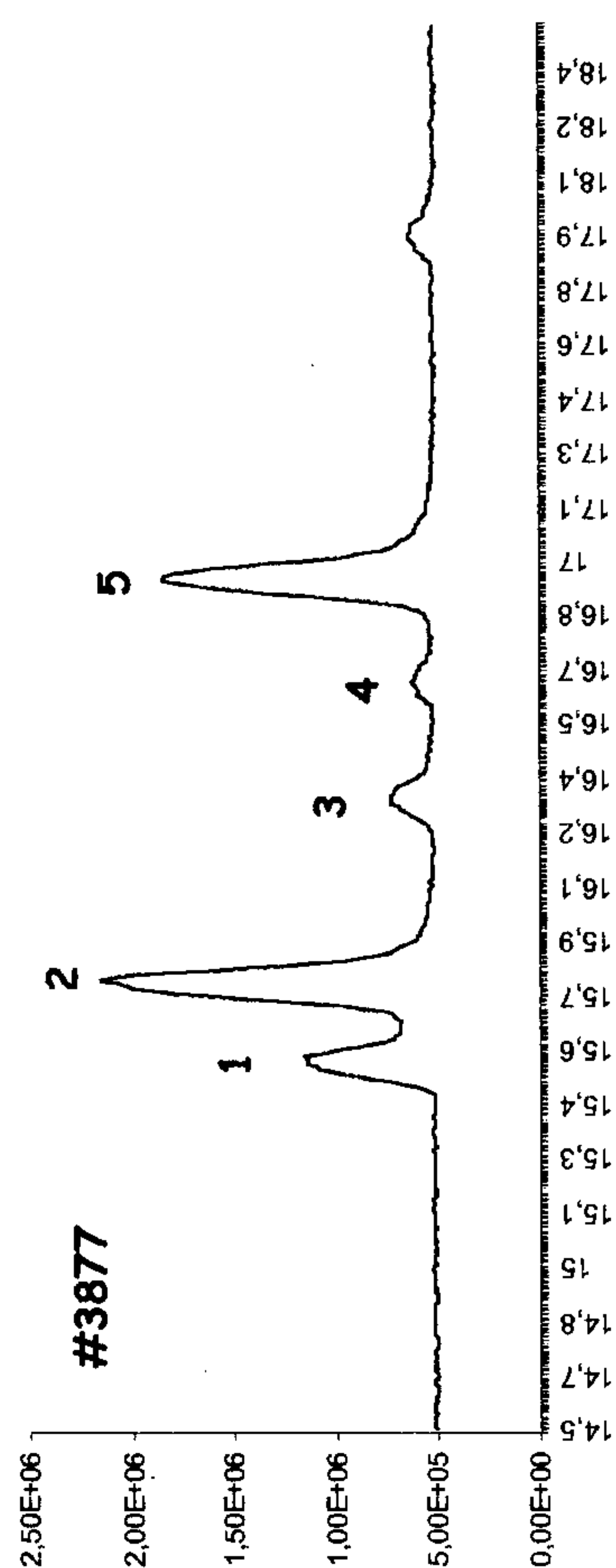


Figure 7A

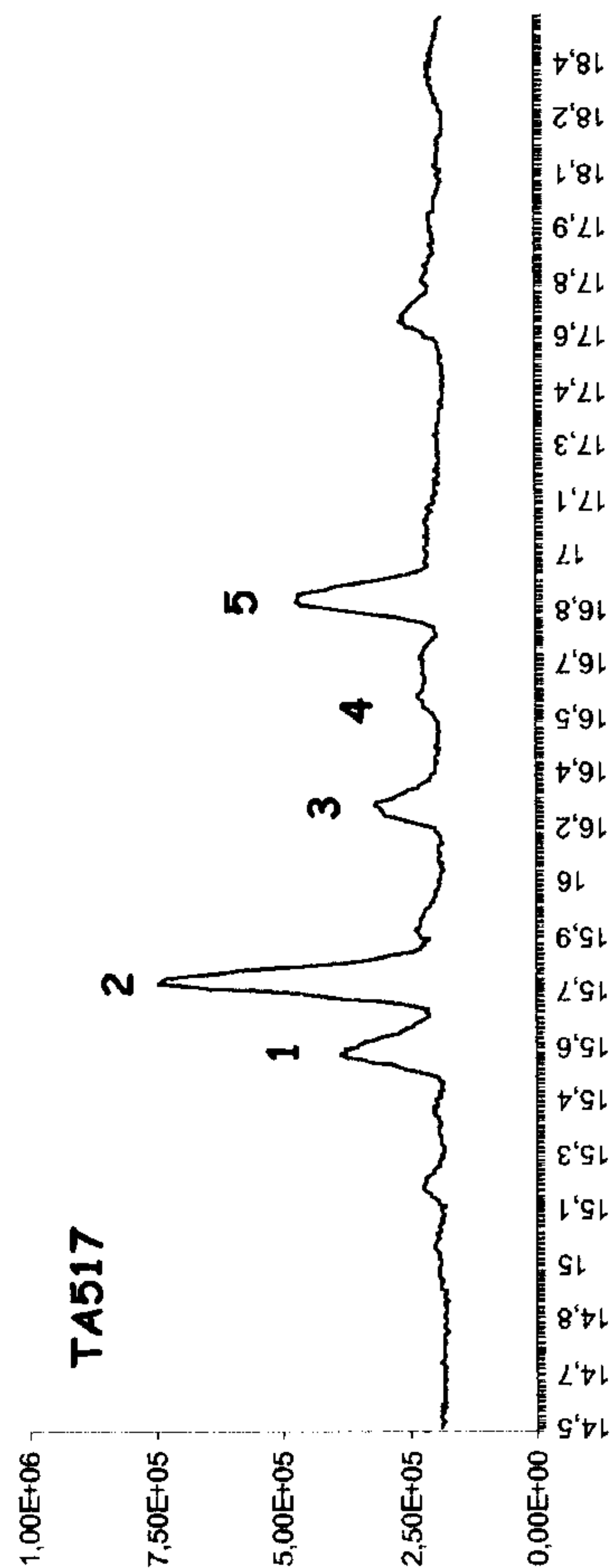


Figure 7B

8/8

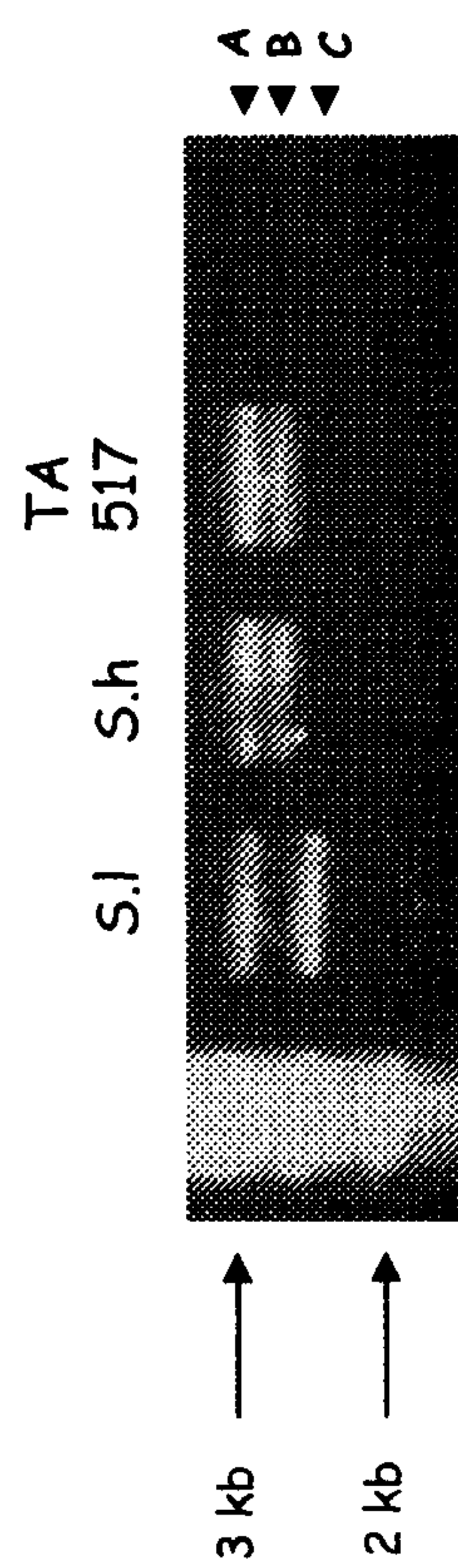


Figure 8