



(86) **Date de dépôt PCT/PCT Filing Date:** 2015/06/09
(87) **Date publication PCT/PCT Publication Date:** 2015/12/17
(45) **Date de délivrance/Issue Date:** 2020/10/27
(85) **Entrée phase nationale/National Entry:** 2016/12/06
(86) **N° demande PCT/PCT Application No.:** US 2015/034859
(87) **N° publication PCT/PCT Publication No.:** 2015/191551
(30) **Priorité/Priority:** 2014/06/09 (US62/009,467)

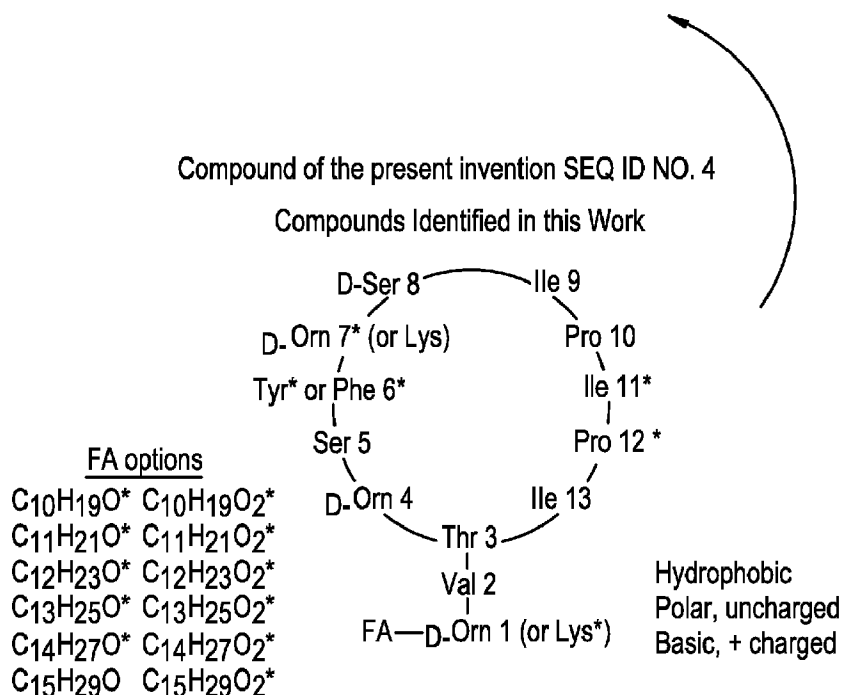
(51) **Cl.Int./Int.Cl. C07K 7/08** (2006.01),
A23K 20/147 (2016.01), **A23L 3/3526** (2006.01),
A23L 33/18 (2016.01), **A61K 38/16** (2006.01),
A61L 2/16 (2006.01), **A61P 31/00** (2006.01),
C07K 7/64 (2006.01), **C12P 21/02** (2006.01)

(72) **Inventeurs/Inventors:**
ZHENG, JIE, US;
KNOLHOFF, ANN MARIE, US;
BROWN, ERIC WAYNE, US;
CROLEY, TIMOTHY RAY, US

(73) **Propriétaire/Owner:**

(54) **Titre : PEPTIDES ANTIMICROBIENS, COMPOSITIONS PHARMACEUTIQUES ET PROCEDES D'UTILISATION CORRESPONDANTS**

(54) **Title: ANTIMICROBIAL PEPTIDES, PHARMACEUTICAL COMPOSITIONS, AND METHODS OF USE THEREOF**



(57) **Abrégé/Abstract:**

Disclosed herein are novel antimicrobial peptides, pharmaceutical compositions containing the peptides, and methods of use of the peptides to inhibit the growth or proliferation of microbes. The antimicrobial peptides are particularly useful to treat infections of

(73) **Propriétaires(suite)/Owners(continued):**

THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES, US

(74) **Agent:** NORTON ROSE FULBRIGHT CANADA LLP/S.E.N.C.R.L., S.R.L.

(57) **Abrégé(suite)/Abstract(continued):**

dangerous gram positive organisms such as MRSA and VRSA.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(10) International Publication Number
WO 2015/191551 A1

(43) International Publication Date
17 December 2015 (17.12.2015)

- (51) International Patent Classification:
C07K 7/08 (2006.01) *C07K 7/64* (2006.01)
- (21) International Application Number:
PCT/US2015/034859
- (22) International Filing Date:
9 June 2015 (09.06.2015)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/009,467 9 June 2014 (09.06.2014) US
- (71) Applicant: THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; National Institutes of Health, Office of Technology Transfer, 6011 Executive Boulevard, Suite 325, MSC 7660, Bethesda, Maryland 20892-7660 (US).

(72) Inventors: **ZHENG, Jie**; 5937 Cedar Fern Court, Columbia, Maryland 21044 (US). **KNOLHOFF, Ann Marie**; 8504 16th St., Apt. 413, Silver Spring, Maryland 20910 (US). **BROWN, Eric Wayne**; 28 Monocacy Circle, Taneytown, Maryland 21787 (US). **CROLEY, Timothy Ray**; 204 Gosvenor Lane, Severna Park, Maryland 21146 (US).

(74) Agent: **LECUYER, Karen A.**; Cantor Colburn LLP, 20 Church Street, 22nd Floor, Hartford, Connecticut 06103 (US).

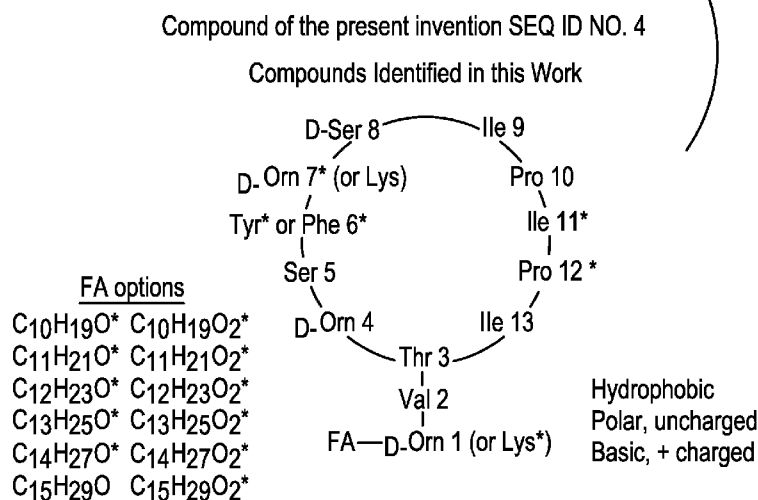
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,

[Continued on next page]

(54) Title: ANTIMICROBIAL PEPTIDES, PHARMACEUTICAL COMPOSITIONS, AND METHODS OF USE THEREOF

FIG. 8B



(57) Abstract: Disclosed herein are novel antimicrobial peptides, pharmaceutical compositions containing the peptides, and methods of use of the peptides to inhibit the growth or proliferation of microbes. The antimicrobial peptides are particularly useful to treat infections of dangerous gram positive organisms such as MRSA and VRSA.

WO 2015/191551 A1

TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

ANTIMICROBIAL PEPTIDES, PHARMACEUTICAL COMPOSITIONS, AND METHODS OF USE THEREOF

FIELD OF THE DISCLOSURE

[0001] The present disclosure is related to antimicrobial peptides, pharmaceutical compositions containing the antimicrobial peptides, and methods of treating microbial infections with the antimicrobial peptides.

BACKGROUND

[0002] Antibiotic-resistant pathogenic bacteria such as methicillin resistant *Staphylococcus aureus* (MRSA), fluoroquinolone resistant *Pseudomonas aeruginosa* and *Clostridium difficile*, and multi-drug resistant *Salmonella* spp. are an emerging problem in modern medicine. The loss of efficacy of current antibiotics makes the identification and development of new antibiotics more critical. The environment, for example, is an important source of microbial strains capable of producing potent antimicrobials which can be isolated and purified from their natural sources.

[0003] *Paenibacillus*, spore-forming species widely distributed in the environment, are a potential source of new antimicrobials. Strains of *Paenibacillus* can produce diverse antimicrobial agents including lantibiotics, lipopeptides, and macrolides. Lipopeptides, for example, are compounds that are generally not synthesized by ribosomes, and that are active against a wide range of bacteria, fungi, and oomycetes. Lipopeptides can act as antiviral and antitumor agents, immunomodulators or specific toxins and enzyme inhibitors.

[0004] There is thus a need to identify and develop antimicrobial agents that are effective against a broad spectrum of microbial pathogens such as Gram-positive and Gram-negative bacteria.

BRIEF SUMMARY

[0005] In an aspect, included herein is a peptide of the sequence
Xaa₁-Xaa₂-Xaa₃-Xaa₄-Xaa₅-Xaa₆-Xaa₇-Xaa₈-Xaa₉-Pro₁₀-Xaa₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 2),

wherein Xaa₆ is Tyr, Phe, or Trp; Xaa₁, Xaa₄ and Xaa₇ are each independently Lys or Orn; Xaa₂, Xaa₉ and Xaa₁₁ are each independently Leu, Ile, Val, or Ala; Xaa₃, Xaa₅, and Xaa₈ are each independently Cys, Tyr, Thr, or Ser; wherein the peptide optionally includes a saturated or unsaturated, substituted or unsubstituted, linear or branched C₄-C₂₀ fatty acid group, or a saturated or unsaturated, linear or branched C₄-C₂₀ ester covalently linked to Xaa₁.

[0006] In another aspect, included herein is a peptide of the sequence Xaa₁-Val₂-Thr₃-Xaa₄-Ser₅-Xaa₆-Xaa₇-Ser₈-Ile₉-Pro₁₀-Xaa₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 3),

wherein Xaa₆ is Tyr, Phe, or Trp; Xaa₁₁ is Leu, Ile, Val, or Ala; and Xaa₁, Xaa₄ and Xaa₇ are independently Lys or Orn, wherein the peptide optionally includes a saturated or unsaturated, substituted or unsubstituted, linear or branched C₄-C₂₀ fatty acid group, or a saturated or unsaturated, linear or branched C₄-C₂₀ ester covalently linked to Xaa₁.

[0007] In certain aspects, the peptides are cyclized through a bond between Thr₃ and Ile₁₃.

[0008] Further included herein is a composition comprising the peptides described above and a carrier, vehicle, excipient, or diluent.

[0009] In yet another aspect, a process for preparing the peptides described above comprises (a) cultivating a host cell under conditions that allow for production of the peptide; and (b) purifying and isolating the peptide.

[0010] In a further aspect, a process for preparing a composition comprising the peptides described above comprises (a) cultivating a host cell under conditions that allow for production of the peptide; (b) purifying and isolating the peptide, and (c) producing a composition comprising the isolated peptide and a carrier, vehicle, excipient, or diluent.

[0011] In a still further aspect, a method of inhibiting growth or proliferation of a microbe comprises contacting the microbe or a surface or product which may contain a microbe with the peptide described above.

[0012] In a still further aspect, a method of inhibiting growth or proliferation of a microbe in a subject comprises administering to the subject a composition comprising one or more of the peptides described above, wherein the peptide is administered in an amount effective to inhibit growth or proliferation of the microbe.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] Figures 1A and 1B show *in vitro* inhibition of foodborne pathogens and tomato bacterial phytopathogens by *Paenibacillus alvei* TS-15 (Figure 1A; ATCC PTA-121756) and A6-6i (Figure 1B; ATCC PTA-121885) on tryptic soy agar (TSA). The inhibition zones (mm) were measured against strains from *Salmonella* spp., *Escherichia coli* (*E. coli*), *Cronobacter sakazakii* (CS), *Listeria monocytogenes* (LM), *Shigella dysenteriae* (SD), Methicillin sensitive *Staphylococcus aureus* (MSSA), Methicillin resistant *Staphylococcus aureus* (MRSA), *Ralstonia solanacearum* race 5 (*R. solanacearum*), *Pseudomonas syringae* pv. tomato strain dc3000 (*P. syringae*), and *Erwinia carotovora* subsp. *carotovora* (*E. carotovora*). The plots represents the lowest, highest, and average measurements in each of the species listed above. The experiment was repeated twice.

[0014] Figures 2A and 2B show growth inhibition of major foodborne pathogens in *P. alvei* A6-6i cell free culture supernatant (CFCS). Brain Heart Infusion (BHI) broth was used as a control. Bacterial growth of A) *L. monocytogenes* (LM), *S. dysenteriae* (SD), *E. coli* O157, *C. sakazakii* (CS), and *S. Newport* strains; and B) Methicillin resistant *S. aureus* (MRSA) strains in *P. alvei* A6-6i CFCS and BHI was determined in five replicates by measuring O.D.₆₀₀ at 20-minute intervals for 24 hours. The experiment was repeated twice.

[0015] Figure 3 shows a polymyxin B standard on a *Salmonella* lawn. A 10 μ L volume of 2-fold serial polymyxin B dilutions starting from 1 mg/mL to 1 μ g/mL was spotted on a lawn of 10^6 cells of *Salmonella enterica* serovar Montevideo strain 29N. The zone of inhibition (ZI) was observed after 24 hour incubation at $35 \pm 2^\circ\text{C}$.

[0016] Figure 4 shows TS-15 1-minute active fractions against pathogens. A 10 μ L volume from the 1-minute fractions of *P. alvei* strains TS-15 was spotted on a lawn of 10^6 cells of A) *Escherichia coli* O157:H7 strain EDL933; B) Methicillin-resistant *Staphylococcus aureus* strain #12. After incubation at $35 \pm 2^\circ\text{C}$ for 24 hours, the antimicrobial activity exhibited by the 1-minute fractions was observed as a clear zone of inhibition (ZI). These experiments were also done for *P. alvei* strain A6-6i (data not shown).

[0017] Figure 5 shows MALDI-TOF MS results for fractions 20-26 from TS-15. The unlabeled arrows indicate a mass difference of 14 Da, which indicates a difference in CH_2 ; an asterisk indicates MW 1623, which is designated as the primary sequence; two examples of a mass difference of 2 and 16 Da are labeled, which correspond to the other molecular variants.

[0018] Figure 6 shows MS/MS similarity. MALDI-TOF MS/MS comparison of two compounds that differ by 14 Da in molecular weight. Mass-to-charge ratios with an asterisk indicate an observed mass difference of 14 Da in the comparison between the two spectra.

[0019] Figure 7 shows partial sequence information from MALDI-TOF MS/MS spectrum of MW 1623 elucidated by manual *de novo* sequencing.

[0020] Figures 8A and 8B show schematics of a previously published identified lipopeptide and the peptides discovered in current work. A. Molecule described in U.S. Publication No. US2013/0164317 and referred to as paenibacterin. B. Molecule(s) discovered in current work. The “R” group corresponds to an attached acyl chain in structure A and B; however, the “R” group for the structure in B. can also be an ester, for example, when Tyr is at position 6.

[0021] Figure 9 shows a representative example of different product ion series in an MSⁿ spectrum (MS³ 642²⁺). The corresponding sequences are coded to the amino acids in the inset schematic of the cyclic peptide structure.

[0022] Figures 10A and 10B show product ion assignments as a result of combined interpretation of multiple MSⁿ analyses. Representative examples of MS² spectra are shown for two of the most abundant compounds, containing a Phe (MW 1607) or Tyr (MW 1623) at position 6, respectively. These product ion assignments led to the chemical structure shown in Figure 8. .

[0023] Figure 11 shows MS² spectra. Three series of compounds within the class of antibiotics differ by an amino acid or a difference in their attached fatty acid. Compounds that contain a Tyr at position 6 have m/z 657²⁺ as a consistent product ion in their resulting MS/MS spectra, while Phe at position 6 results in m/z 649²⁺. The m/z values with an asterisk indicate a mass difference of 16 Da between the MS³ spectra, which is the mass difference between Phe and Tyr. The mass difference between MW 1623 and MW 1625 corresponds to one less CH₂ group and an additional oxygen in the attached fatty acid (-CH₂+O)..

[0024] Figure 12 shows three representative MS³ spectra that demonstrate sequence similarity between compounds of molecular weights that differ by 14 Da and their corresponding complementary ion pair MS³ spectra. MS² spectra are outlined and MS³ spectra are outlined, with product ions selected for MS³ highlighted .

[0025] Figure 13 shows MS/MS comparison of MW 1637 (top) and MW 1639 (bottom). The m/z values with an asterisk indicate a 1.979 Da mass difference between product ions in the two spectra, corresponding to one less CH₂ and an additional oxygen in the attached fatty acid compared to the dominant molecular species.

[0026] Figure 14 shows different MS/MS spectra for three compounds of the same molecular weight indicate Lys or Orn at Position 7 in Figure 8 and potential diversity in the

structure of the attached fatty acid due to distinct chromatographic peaks shown in the extracted ion chromatogram (EIC).

[0027] Figure 15 shows a schematic of a chemically synthesized, fatty acid modified peptide.

[0028] Figures 16A-16M are a *ptbA* alignment for strains A6-6i, TS-15 and OSY SE.

[0029] Figures 17A-17M are a *ptbB* alignment for strains A6-6i, TS-15 and OSY SE.

[0030] Figures 18A-18G are a *ptbC* alignment for strains A6-6i, TS-15 and OSY SE.

[0031] The above-described and other features will be appreciated and understood by those skilled in the art from the following detailed description, drawings, and appended claims.

DETAILED DESCRIPTION

[0032] Described herein are novel peptides, specifically cyclic peptides, that have antimicrobial and broad-spectrum antibacterial activity. Specifically, the peptides are active against dangerous Gram-positive organisms such as MRSA and VRSA.

[0033] Recently, tomatoes have been implicated as a primary vehicle in foodborne outbreaks of *Salmonella Newport* and other *Salmonella* serovars. Long-term intervention measures to reduce *Salmonella* prevalence on tomatoes remain elusive for growing and post-harvest environments. A naturally-occurring bacterium identified by 16S rDNA sequencing as *Paenibacillus alvei* was isolated epiphytically from plants native to the Virginia Eastern Shore tomato growing region. After initial antimicrobial activity screening against *Salmonella* and 10 other bacterial pathogens associated with the human food supply, strain TS-15 was further used to challenge an attenuated strain of *S. Newport* on inoculated tomatoes, leaves, and blossoms of tomato plants in an insect-screened high tunnel with a split-plot design. Survival of *Salmonella* after inoculation was measured for groups with and without the antagonist at days 0, 1, 2, 3, and 5 for blossoms and day 6 for tomatoes and leaves, respectively. TS-15 exhibited broad range antimicrobial activity against both major foodborne pathogens and major tomato plant-associated bacterial pathogens. After *P. alvei* strain TS-15 was applied onto the tomatoes, leaves, and blossoms of tomato plants, the concentration of *S. Newport* was significantly lower ($p \leq 0.05$) compared with controls. Surprisingly, more than 90% of the plants had no detectable levels of *Salmonella* by day 5 for blossoms. The naturally occurring antagonist strain TS-15 is highly effective in reducing carriage of *Salmonella Newport* on whole tomato plants. The application of *P. alvei* strain

TS-15 is a promising approach for reducing the risk of *Salmonella* contamination during tomato production. In addition, *Paenibacillus* strain A6-6i was found to retain comparable properties. Given to the fact that this activity can be attributed to the bactericidal compounds within, the present inventors have chemically isolated, purified, and identified certain compounds responsible for anti-*Salmonella* and other antibiotic effects including effects on dangerous gram positive organisms such as MRSA and VRSA.

[0034] More specifically, antimicrobial compounds were isolated from these *Paenibacillus* strains and a combination of low and high resolution mass spectrometry with multiple-stage tandem mass spectrometry was used for identification. A group of closely related cyclic lipopeptides was identified, differing primarily by fatty acid chain length and one of two possible amino acid substitutions. Variation in the fatty acid length resulted in mass differences of 14 Da and yields groups of related MSⁿ spectra. Despite the inherent complexity of MS/MS spectra of cyclic compounds, straightforward analysis of these spectra was accomplished by determining differences in complementary product ion series between compounds that differ in molecular weight by 14 Da.

[0035] An “antimicrobial compound” is a compound that exhibits antimicrobial activity or a compound that affects microbial activity, meaning a compound that slows or stops growth and/or proliferation, slows or stops the rate of growth and/or proliferation, or stuns, inactivates, or kills a microbe. Antimicrobial compounds include antibiotics, antibacterials (e.g., bactericidal or bacteriostatic agents), antivirals (e.g., virucidal agents), antifungals (e.g., fungicidal or fungistatic agents), mold-inhibiting agents, anthelmintics (e.g., vermifuge or vermucidal agents), antiparasitics, and the like. Antimicrobial activity can be determined using methods described herein as well as methods known in the art.

[0036] As used herein, amino acids include alpha-amino acids of the general formula H₂NCHR⁺COOH when free and HN⁺CHR⁺CO when in a polypeptide, wherein R is an amino acid side chain comprising an organic substituent, as well as uniquely structured amino acids such as, for example, proline. Amino acids include, for example, isoleucine, leucine, alanine, asparagine, glutamine, lysine, aspartic acid, glutamic acid, methionine, cysteine, phenylalanine, threonine, tryptophan, glycine, valine, proline, serine, tyrosine, arginine, histidine, norleucine, ornithine, taurine, selenocysteine, selenomethionine, lanthionine, 2-aminoisobutyric acid, dehydroalanine, hypusine, citrulline, 3-aminopropanoic acid, aminobutyric acid (alpha, beta, and gamma) diaminobutyric acid, and the like. The term

“amino acid side chain” refers to the various organic substituent groups (e.g., “R” in H₂NCHR₁COOH) that differentiate one amino acid from another.

[0037] In one aspect, an antimicrobial peptide has the sequence:

Xaa₁-Val₂-Thr₃-Xaa₄-Ser₅-Xaa₆-Xaa₇-Ser₈-Ile₉-Pro₁₀-Ile₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 1),

wherein Xaa₆ is Tyr, Phe, or Trp, specifically Tyr or Phe; Xaa₁, Xaa₄, and Xaa₇ are each independently Lys or Orn; or,

Xaa₁-Xaa₂-Xaa₃-Xaa₄-Xaa₅-Xaa₆-Xaa₇-Xaa₈-Xaa₉-Pro₁₀-Xaa₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 2)

wherein Xaa₆ is Tyr, Phe, or Trp, specifically Tyr or Phe; Xaa₁, Xaa₄, and Xaa₇ are each independently Lys or Orn; Xaa₂, Xaa₉ and Xaa₁₁ are each independently a hydrophobic amino acid selected from Leu, Ile, Val, and Ala; Xaa₃, Xaa₅, and Xaa₈ are each independently an amino acid selected from amino acids that can form a hydrogen bond, a disulfide bond, a thioether bond, or an ester bond, such as Cys, Tyr, Thr, and Ser.

[0038] In a more specific aspect, an antimicrobial peptide has the sequence:

Xaa₁-Val₂-Thr₃-Xaa₄-Ser₅-Xaa₆-Xaa₇-Ser₈-Ile₉-Pro₁₀-Xaa₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 3),

wherein Xaa₆ is Tyr, Phe, or Trp, specifically Tyr or Phe; Xaa₁, Xaa₄, and Xaa₇ are each independently Lys or Orn; Xaa₁₁ is a hydrophobic amino acid selected from Leu, Ile, Val, and Ala. In a specific embodiment, Xaa₁ and Xaa₄ are Orn and Xaa₁₁ is Ile; more specifically,

Orn₁-Val₂-Thr₃-Orn₄-Ser₅-Xaa₆-Xaa₇-Ser₈-Ile₉-Pro₁₀-Ile₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 4),

wherein Xaa₆ is Tyr or Phe; and Xaa₇ is Lys or Orn.

[0039] In a specific embodiment, the peptide is

Orn₁-Val₂-Thr₃-Orn₄-Ser₅-Tyr₆-Lys₇-Ser₈-Ile₉-Pro₁₀-Ile₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 77).

[0040] In certain embodiments, the peptides of SEQ ID NOs. 1-4 include a fatty acid group, particularly a saturated or unsaturated, substituted or unsubstituted, linear or branched C₄-C₂₀ fatty acid group, or a saturated or unsaturated, linear or branched C₄-C₂₀ ester covalently linked to the amino acid at the 1 position, e.g., Xaa₁. The fatty acid chain is

diverse in the number of $-\text{CH}_2$. For example, the molecular formula of the acyl chain can be $\text{C}_{10}\text{H}_{19}\text{O}$, $\text{C}_{11}\text{H}_{21}\text{O}$, $\text{C}_{12}\text{H}_{23}\text{O}$, $\text{C}_{13}\text{H}_{25}\text{O}$, $\text{C}_{14}\text{H}_{27}\text{O}$, or $\text{C}_{15}\text{H}_{29}\text{O}$. The ester group differs similarly ($\text{C}_{10}\text{H}_{19}\text{O}_2$, $\text{C}_{11}\text{H}_{21}\text{O}_2$, $\text{C}_{12}\text{H}_{23}\text{O}_2$, $\text{C}_{13}\text{H}_{25}\text{O}_2$, $\text{C}_{14}\text{H}_{27}\text{O}_2$, or $\text{C}_{15}\text{H}_{29}\text{O}_2$).

[0041] Depending on the functional groups, peptides can be cyclized C-terminus to N-terminus, C-terminus to side-chain, side-chain to N-terminus, or side-chain to side-chain. In one aspect, the peptide is cyclized through a bond between $\text{Thr}_{3/4}$ and $\text{Ile}_{13/14}$.

[0042] In a specific embodiment, when Xaa_6 is Tyr, the amino acid at the 1 position is covalently linked to a fatty acid group or an ester group, and when Xaa_6 is Phe, the amino acid at the 1 position is covalently linked to a fatty acid group.

[0043] It is noted that the presently disclosed compounds are distinct from the cyclic peptides disclosed in US2013/0164317, which requires that X_{12} is an amino acid with a charged side chain. In the present compounds, X_{12} is a Proline, which is expected to impart a unique structure and also antimicrobial properties to the disclosed peptides.

[0044] Also included herein are methods of producing an antimicrobial compound, specifically an antimicrobial peptide, by (a) cultivating a host cell under conditions that allow for production of the antimicrobial peptide; and optionally (b) purifying/isolating the antimicrobial peptide.

[0045] Host cells are cultivated in a nutrient medium suitable for production of the antimicrobial peptide using techniques known in the art. For example, the cell is cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. A suitable nutrient medium (e.g., a medium comprising carbon and nitrogen sources, inorganic salts, etc.) is used to cultivate the cells. In embodiments wherein the antimicrobial peptide is secreted from the cell into the nutrient medium, the antimicrobial peptide can be recovered directly from the medium. If the antimicrobial peptide is not secreted, it can be recovered from cell lysates or as inclusion bodies. The antimicrobial peptide may be recovered from the nutrient medium or cell lysates by procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, and precipitation.

[0046] The antimicrobial peptides disclosed herein may be purified by a variety of procedures including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g.,

preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, and extraction.

[0047] Bacterial cultures for the production of the antimicrobial peptides can be grown using suitable methods and media useful for bacterial cell growth, maintenance, and/or protein production.

[0048] In some embodiments of this process, the antimicrobial peptide is isolated and/or purified using a suitable technique known in the art, including liquid chromatography, phase separation, using organic solvents and/or aqueous solvent or buffer systems. In some embodiments the antimicrobial peptide is purified to about 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more. Analysis of purity can be made using an analytical method or technique such as, for example, mass spectrometry, gel electrophoresis, fluorescence, colorimetric assays, NMR, UV-Vis, total amino acid hydrolysis, chromatographic separation methods that utilize, for example, liquid chromatographic methods such as HPLC, FPLC, size exclusion, affinity binding, hydrophobic interaction, ionic charge, where purity can be assessed based on peak area.

[0049] In other embodiments, the antimicrobial peptides can be generated by standard chemical and/or protein and peptide synthetic techniques as are known in the art. Some embodiments relate to a synthetic strategy that incorporates a combination of chemical, peptide, and enzymatic (e.g., cyclase) synthetic steps. Chemical techniques for cyclizing peptides are well-known in the art.

[0050] Aspects of the disclosure relate to compositions and formulations, including pharmaceutical compositions and formulations that comprise an effective amount of at least one antimicrobial peptide. Such compositions and formulations comprise an effective amount of an antimicrobial peptide in combination with a carrier, vehicle, excipient, or diluent, including pharmaceutically and/or agriculturally acceptable carriers, vehicles, excipients, and diluents. An “effective amount” relates to a quantity of an antimicrobial peptide that is high enough to provide a significant positive result (e.g., slow or stop microbial activity) or positive modification of the subject's condition to be treated, and is suitably low enough to avoid serious side effects (at a reasonable benefit/risk ratio). Carriers, vehicles, excipients, and diluents are one or more compatible substances that are suitable for administration to a mammal such as, for example, solid or liquid fillers, diluents, hydrotopes, surface-active agents, and encapsulating substances. “Compatible” means that the components of the composition are capable of being mixed with the active agent, and with

each other, in a manner such that there is no interaction which would substantially reduce the efficacy of the composition under ordinary use situations. Carriers, vehicles, excipients, and diluents are suitably of sufficiently high purity and sufficiently low toxicity to render them suitable for administration to the subject being treated, such as a human subject. The carrier, vehicle, excipient, or diluent can be inert, or it can possess pharmaceutical benefits and/or aesthetic benefits, or both. Suitable carriers, vehicles, excipients, and diluents are known in the art.

[0051] The antimicrobial compositions are applicable in a variety of products and applications, ranging from, for example, products of low and high pH-values, highly concentrated and diluted products, products usable in the technical field (e.g., in detergents for industrial or house-hold use), in the pharmaceutical field (e.g., for cleaning/disinfection of equipment or in the preparation of pharmaceutical compositions or their packaging, and in surgical supplies and sterilization of tools/hospital operating rooms), in personal care (e.g., in manufacture of cosmetics, shampoos, creams and lotions), in the feed industry (e.g., for cleaning of equipment, in the manufacture, storage, handling and preparation of animal feed and drink products) and in the food and drink industry (post-harvest foods, food processing surfaces and packaging). The antimicrobial peptides are also useful in post-surgical bandage and wound dressing prep, external wound healing and cleansing. In addition, the antimicrobial peptides are useful in post-harvest and food preservation applications against spoilage organisms. In embodiments relating to use of the compositions in a product, the antimicrobial composition can be provided as an ingredient in the final product (e.g., cosmetic, detergent, pharmaceutical, food, or drink product). Accordingly, in some embodiments, the compositions are effective against certain yeasts, fungi, and bacteria commonly associated with food-spoilage. Standard methods can be used in the manufacture of such products that comprise one or more of the antimicrobial peptide described herein.

[0052] In some embodiments, the antimicrobial composition is present on the surface of the products or inside the products. In some embodiments, the disclosure includes a method for reducing or preventing the presence, growth or activity of a microbe (e.g., gram-positive or gram-negative bacteria) in a product, such as a food or drink product, wherein the method comprises contacting the food or drink product during one or more of the various stages in the food processing process including the stages of the manufacture, the handling, the storage and/or the preparation of the food or drink product with the antibacterial compositions that are disclosed herein. The antimicrobial composition may be applied or

introduced by a suitable route or method such as, for example, as a spray, a rinse or a wash solution or as solution wherein the various food products are dipped. Further, the antimicrobial composition may be used to treat containers or packaging film prior to, simultaneously with, or subsequently after packaging the products.

[0053] In one aspect, a method of inhibiting growth or proliferation of a microbe comprises contact of the microbe with an antimicrobial peptide as described herein. “Contacting,” as used herein as in “contacting a cell,” refers to contacting a cell directly or indirectly in vitro, ex vivo, or in vivo (i.e., within a subject, such as a mammal, including humans, mice, rats, rabbits, cats, and dogs). Contacting a cell, which also includes “reacting” a cell, can occur as a result of administration to a subject. Contacting encompasses administration to a cell, tissue, mammal, subject, patient, or human. Further, contacting a cell includes adding an agent to a cell culture. Other suitable methods may include introducing or administering an agent to a cell, tissue, mammal, subject, or patient using appropriate procedures and routes of administration as defined herein.

[0054] The antimicrobial compositions described herein may be provided in solid or liquid form. When in liquid form, the composition is typically an aqueous composition, which may be a solution, emulsion, or dispersion.

[0055] Accordingly, the methods described herein include administration of one or more pharmaceutical compositions, in which an antimicrobial peptide is admixed together with one or more pharmaceutically acceptable carriers, excipients, buffers, adjuvants, stabilizers, or other materials, as described herein.

[0056] “Pharmaceutically acceptable,” as used herein, pertains to compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of a subject (e.g., a human) without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio. Each carrier, excipient, etc. must also be “acceptable” in the sense of being compatible with the other ingredients of the formulation.

[0057] The formulations may conveniently be presented in unit dosage form and may be prepared by methods known in the art of pharmacy. Such methods include the step of bringing into association the active compound(s) with the carrier which constitutes one or more accessory ingredients. In general, the formulations are prepared by uniformly and intimately bringing into association the active compound with liquid carriers or finely divided solid carriers or both, and then if necessary shaping the product.

[0058] Formulations may be in the form of liquids, solutions, suspensions, emulsions, elixirs, syrups, tablets, lozenges, granules, powders, capsules, cachets, pills, ampoules, suppositories, pessaries, ointments, gels, pastes, creams, sprays, mists, foams, lotions, oils, boluses, electuaries, or aerosols.

[0059] Formulations suitable for oral administration (e.g., by ingestion) are typically presented as discrete units such as capsules, cachets or tablets, each containing a predetermined amount of the active compound; as a powder or granules; as a solution or suspension in an aqueous or nonaqueous liquid; or as an oil-in-water liquid emulsion or a water-in-oil liquid emulsion; as a bolus; as an electuary; or as a paste.

[0060] A tablet may be made by compression or molding, optionally with one or more accessory ingredients. Compressed tablets are prepared by compressing in a suitable machine the active compound in a free-flowing form such as a powder or granules, optionally mixed with one or more binders (e.g., povidone, gelatin, acacia, sorbitol, tragacanth, hydroxypropylmethyl cellulose); fillers or diluents (e.g., lactose, microcrystalline cellulose, calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc, silica); disintegrants (e.g., sodium starch glycolate, cross-linked povidone, cross-linked sodium carboxymethyl cellulose); surface-active or dispersing or wetting agents (e.g., sodium lauryl sulfate); and preservatives (e.g., methyl p-hydroxybenzoate, propyl p-hydroxybenzoate, sorbic acid). Molded tablets are made by molding in a suitable machine a mixture of the powdered compound moistened with an inert liquid diluent. The tablets are optionally coated or scored and may be formulated so as to provide slow or controlled release of the active compound therein using, for example, hydroxypropylmethyl cellulose in varying proportions to provide the desired release profile. Tablets are optionally provided with an enteric coating, to provide release in parts of the gut other than the stomach.

[0061] Formulations suitable for parenteral administration (e.g., by injection, including cutaneous, subcutaneous, intramuscular, intravenous and intradermal injection), include aqueous and nonaqueous isotonic, pyrogen-free, sterile injection solutions which may contain anti-oxidants, buffers, preservatives, stabilizers, bacteriostats, and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents, and liposomes or other microparticulate systems which are designed to target the compound to blood components or one or more organs. Examples of isotonic vehicles for use in such formulations include Sodium Chloride Injection, Ringer's Solution, or Lactated Ringer's

Injection. The formulations may be presented in unit-dose or multi-dose sealed containers, for example, ampoules and vials, and may be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules, and tablets. Formulations may be in the form of liposomes or other microparticulate systems which are designed to target the active compound to blood components or one or more organs.

[0062] Formulations suitable for topical administration (e.g., transdermal, intranasal, ocular, buccal, and sublingual) may be formulated as an ointment, cream, suspension, lotion, powder, solution, past, gel, spray, aerosol, or oil. Alternatively, a formulation may comprise a patch or a dressing such as a bandage or adhesive plaster impregnated with active compounds and optionally one or more excipients or diluents. Formulations suitable for topical administration to the eye also include eye drops wherein the active compound is dissolved or suspended in a suitable carrier, especially an aqueous solvent for the active compound.

[0063] Formulations suitable for topical administration via the skin include ointments, creams, and emulsions. When formulated in an ointment, the active compound may optionally be employed with either a paraffinic or a water-miscible ointment base. Alternatively, the active compounds may be formulated in a cream with an oil-in-water cream base. The topical formulations may desirably include a compound which enhances absorption or penetration of the active compound through the skin or other affected areas. Examples of such dermal penetration enhancers include dimethylsulfoxide.

[0064] It will be appreciated that appropriate dosages of the active compounds, and compositions comprising the active compounds, can vary from subject to subject. Determining the optimal dosage will generally involve the balancing of the level of therapeutic benefit against any risk or deleterious side effects of the treatments described herein. The selected dosage level will depend on a variety of factors including, but not limited to, the species of the particular subject, the activity of the particular compound, the route of administration, the time of administration, the rate of excretion of the compound, the duration of the treatment, whether other drugs, compounds, and/or materials are used in combination, and the age, sex, weight, condition, general health, and prior medical history of the subject. The amount of compound and route of administration will ultimately be at the discretion of the physician, although generally the dosage will be to achieve local

concentrations at the site of action which achieve the desired effect without causing substantial harmful or deleterious side-effects.

[0065] Administration in vivo can be effected in one dose, continuously or intermittently (e.g., in divided doses at appropriate intervals) throughout the course of treatment. Methods of determining the most effective means and dosage of administration are well known to those of skill in the art and will vary with the formulation used for therapy, the purpose of the therapy, the target cell being treated, and the subject being treated. Single or multiple administrations can be carried out with the dose level and pattern being selected by the treating physician.

[0066] In general, a suitable dose of the active compound is in the range of about 100 µg to about 250 mg per kilogram body weight of the subject per day, administered in a single or multiple doses per day.

[0067] A method of inhibiting growth or proliferation of a microbe in a subject comprises administering to the subject a composition comprising an antimicrobial peptide, wherein the antimicrobial peptide is administered in an amount effective to inhibit growth or proliferation of the microbe.

[0068] In an embodiment, a method of treating a condition or disease associated with the presence of a microbe comprises administering to a subject in need thereof a composition comprising an antimicrobial peptide, wherein the antimicrobial peptide is administered in an amount effective to treat the condition or disease.

[0069] In an embodiment, a method of treating a microbial infection comprises administering to a subject in need thereof a composition comprising an antimicrobial peptide, wherein the antimicrobial peptide is administered in an amount effective to treat the microbial infection. Exemplary infections include MRSA, VRSA, and CRE infections.

[0070] “Administration” or “administering,” as used herein, refers to providing, contacting, and/or delivery of a compound or compounds by an appropriate route to achieve the desired effect. Administration may include, but is not limited to, oral, sublingual, parenteral (e.g., intravenous, subcutaneous, intracutaneous, intramuscular, intraarticular, intraarterial, intrasynovial, intrasternal, intrathecal, intralesional, or intracranial injection), transdermal, topical, buccal, rectal, vaginal, nasal, ophthalmic, via inhalation, and implants.

[0071] As used herein, the terms “treatment,” “treating,” or “treat” refer to both therapeutic treatment and prophylactic or preventative measures. Those subjects in need of treatment include those already showing clinical signs of the particular disease, disorder, or

condition as well as those prone to having or developing the disease, disorder, or condition, or those in which the disease, disorder, or condition is to be prevented. Many diseases, disorders, and conditions relate to the presence of microbes and are known to those of skill in the art, including secondary conditions resulting from opportunistic infections arising from other primary diseases and disorders (e.g., immune-suppressing conditions). Thus, a variety of patient classes can benefit from the methods of treatment described herein.

[0072] As used herein, the term “subject” is intended to include human and non-human animals. Exemplary human subjects include a human patient having a disorder, e.g., a disorder described herein, or a normal subject. The term “non-human animals” includes all vertebrates, e.g., non-mammals (such as fowl (e.g., ducks, chickens, etc.), amphibians, reptiles) and mammals, such as non-human primates, domesticated and/or agriculturally useful animals (such as horses, goats, sheep, dogs, cats, cows, pigs, etc.), and rodents (such as mice, rats, hamsters, guinea pigs, etc.).

[0073] In some embodiments the “effective amount” is an amount sufficient to stop or slow the progression of the disease, disorder, or condition. In some embodiments the effective amount is an amount sufficient to reverse disease, disorder, or condition, or repair the clinical signs of a disease, disorder, or condition. In embodiments the amount is sufficient to stop or slow the progression of an infection that is directly or indirectly related to a microbe. In some embodiments the effective amount is sufficient to stop or slow the proliferation and/or growth of a microbe. In further embodiments, the effective amount is sufficient to kill a microbe.

[0074] “Co-administered,” as used herein, refers to simultaneous or sequential administration of multiple compounds or agents. A first compound or agent may be administered before, concurrently with, or after administration of a second compound or agent. The first compound or agent and the second compound or agent may be simultaneously or sequentially administered on the same day, or may be sequentially administered within 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 2 weeks, 3 weeks or one month of each other. Suitably, compounds or agents are co-administered during the period in which each of the compounds or agents are exerting at least some physiological effect and/or has remaining efficacy. In some embodiments, the methods described herein can comprise co-administering two or more active agents disclosed herein. In some embodiments, the methods comprising co-administering two or more active agents include at least one antimicrobial agent disclosed herein in combination with a known active agent

against a particular indication. In some further embodiments, the known active agent also exhibits antimicrobial activity.

[0075] Additional antimicrobial agents can be selected based on the particular method and indication, such that it can provide an additive or a synergistic antimicrobial effect when compared to administration of the antimicrobial agent alone. For example, other antibiotics such as polymyxin B can be concurrently applied to increase its effect against gram-negative bacteria. Furthermore, fungicides can be co-administered for broader protection.

[0076] The invention is further illustrated by the following non-limiting examples.

Examples

Methods

[0077] Bacterial Cell Culture: *Paenibacillus alvei* strains A6-6i and TS-15, naturally-occurring bacterium previously isolated from plant and soil native to the Virginia Eastern Shore tomato growing region, were propagated on tryptic soy agar (TSA) at 35°C. The indicator strains (Table 1) included were also propagated on TSA at 35°C. Stock cultures grown overnight at 35°C on TSA were resuspended in brain heart infusion broth (BHI) with 25% glycerol and stored at -80°C. Three tomato plant-associated bacterial pathogens including *Erwinia carotovora* subsp. *carotovora*, *Pseudomonas syringae* pv. tomato strain dc3000, and *Ralstonia solanacearum* race 5 were grown on TSA at 25°C (Table 1).

Table 1

Strain	Reference or source
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Newport #17	CFSAN laboratory collection
<i>Salmonella enterica</i> subsp. <i>enterica</i> Saintpaul	CFSAN laboratory collection
<i>Salmonella enterica</i> subsp. <i>enterica</i> Montevideo 42N	CFSAN laboratory collection
<i>Salmonella enterica</i> subsp. <i>enterica</i> Javiana	CFSAN laboratory collection
<i>Salmonella enterica</i> subsp. <i>enterica</i> Typhimurium 368477	CFSAN laboratory collection
<i>Salmonella enterica</i> subsp. <i>enterica</i> Typhimurium SAR C #1	SGSC ^a
<i>Salmonella enterica</i> subsp. <i>enterica</i> Typhi SAR C #3	SGSC
<i>Salmonella enterica</i> subsp. <i>arizonae</i> SAR C #5	SGSC
<i>Salmonella enterica</i> subsp. <i>arizonae</i> SAR C #7	SGSC
<i>Salmonella enterica</i> subsp. <i>arizonae</i> SAR C #9	SGSC
<i>Salmonella bongori</i> SAR C #11	SGSC
<i>Salmonella bongori</i> SAR C #13	SGSC
<i>Salmonella bongori</i> SAR C #15	SGSC
<i>Escherichia coli</i> O157:H7 IS O57	CFSAN laboratory collection
<i>Escherichia coli</i> O157:H7 EDL933	CFSAN laboratory collection
<i>Escherichia coli</i> ATCC 51434	ATCC ^b
<i>Escherichia coli</i> ATCC BAA-179	ATCC
<i>Shigella dysenteriae</i> 2457T	CFSAN laboratory collection
<i>Shigella dysenteriae</i> BS103	CFSAN laboratory collection
<i>Cronobacter sakazakii</i> E932	CFSAN laboratory collection
<i>Cronobacter sakazakii</i> E784	CFSAN laboratory collection
<i>Listeria monocytogenes</i> N1-225	CFSAN laboratory collection
<i>Listeria monocytogenes</i> R2-583	CFSAN laboratory collection
Methicillin-resistant <i>Staphylococcus aureus</i> #9	CFSAN laboratory collection
Methicillin-resistant <i>Staphylococcus aureus</i> #12	CFSAN laboratory collection
Methicillin-resistant <i>Staphylococcus aureus</i> #28	CFSAN laboratory collection
Methicillin-resistant <i>Staphylococcus aureus</i> #29	CFSAN laboratory collection
Methicillin-resistant <i>Staphylococcus aureus</i> #30	CFSAN laboratory collection
<i>Staphylococcus aureus</i> NRS70	NARSA ^c
<i>Staphylococcus aureus</i> NRS106	NARSA
<i>Staphylococcus aureus</i> NRS107	NARSA
<i>Staphylococcus aureus</i> NRS271	NARSA
<i>Salmonella enterica</i> Newport #17 $\Delta tolC::aph$	CFSAN laboratory collection
<i>Erwinia carotovora</i> subsp. <i>carotovora</i>	Dr. Dilip Lakshman, ARS ^d
<i>Pseudomonas syringae</i> pv. tomato strain dc3000	Dr. Dilip Lakshman, ARS
<i>Ralstonia solanacearum</i> race 5	Dr. Dilip Lakshman, ARS

^a SGSC, Salmonella Genetic Stock Centre, University of Calgary, Canada

^b ATCC, American Type Culture Collection, Manassas, VA, USA

^c NARSA, Network on Antimicrobial Resistance in *Staphylococcus aureus*, Chantilly, VA, USA

^d ARS, Agricultural Research Service, Department of Agriculture, Beltsville, MD, USA

[0078] Determination of Mode of Action and Spectrum of Antimicrobial Activities

To determine mode of action and antimicrobial spectrum of the bacterial antagonists, both agar plug assay (using bacterial culture) and bioscreen assay (using culture supernatant) were performed against a broad spectrum of major foodborne pathogens and bacterial phytopathogens (Table 1). In the agar plug assay, bactericidal effects against pathogenic bacterial strains in the zone of inhibition were confirmed when no viable cells were recovered on TSA plates. In the bioscreen assay, the antagonist supernatant from overnight culture was filter sterilized with a 0.22 μ m pore-size cellulose acetate (CA) membrane filter. Each 3 ml

TS-15 cell-free culture supernatant (CFCS) was inoculated with 3 μL of 10^8 cfu/mL bacterial culture (Table 1). Aliquots (200 μL) were then dispensed into sterile Bioscreen C microwell plates (Growth Curves USA, Piscataway, NJ) and incubated as described for the respective bacterial strains. Bacterial growth was determined in five replicates by measuring O.D.₆₀₀ at 20-min intervals for 24 hours.

[0079] Fraction Collection: The modified method was based on methods known in the art. Cells were removed from Petri dishes using cell scrapers and were deposited into Eppendorf tubes. A final volume of 100 μL of acetonitrile was added for every dish of scraped cells. The samples were shaken for 30 minutes and centrifuged at 7710 g for 15 minutes. The supernatant was removed and evaporated. The sample was reconstituted in water and was filtered with a 0.22 μm Nylon filter. Fraction collection by liquid chromatography (LC) was achieved using a Shimadzu Nexera with a Kinetex C18 column (1.7 μm , 100 $\text{\AA}$, 150 $\times$ 2.10 mm). The separation was performed with a column temperature of 60°C and a flow rate of 400 $\mu\text{L}/\text{min}$ using water with 0.1% formic acid (v/v) and acetonitrile with 0.1% formic acid (v/v) with the following gradient: 5 min hold at 95% water, 50 min linear gradient from 95% to 5% water, 5 min equilibration at 95% water. Fractions were concentrated and biological activity was tested against MRSA and *E. coli*. Fractions with activity were further examined by multiple mass spectrometry (MS) platforms.

[0080] Bioactivity Assay: Ten microliters of the 1-minute fractions from *Paenibacillus alvei* strain A6-6i or TS-15 were spotted directly on plates containing a lawn of 10^6 cells of *Escherichia coli* O157:H7 strain EDL933 and Methicillin-resistant *Staphylococcus aureus* strain #12, respectively. After incubation at $35 \pm 2^\circ\text{C}$ for 24 hours, the antimicrobial activity exhibited by 1-minute fractions was observed as a clear zone of inhibition (ZI). The fraction that exhibited the ZI was focused henceforth. As a control experiment, 10 μL of serial two-fold polymyxin B (Sigma-Aldrich, St. Louis, MO) dilutions starting from 1 mg/mL stock were spotted on a lawn of 10^6 cells of *Salmonella enterica* serovar Montevideo strain 29N.

[0081] Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF/MS) Analysis: Fractions that were found to be active were diluted 1:30 in water and 1 μL was placed onto the MALDI target with 1 μL of prepared CHCA matrix (20 mg/mL in 70% acetonitrile with 0.1% formic acid). The MALDI instrument used to analyze the samples was an Applied Biosystems/MDS Sciex 4800 MALDI TOF/TOF Analyzer. The laser power was optimized for each analysis to use the

minimum level required for sufficient ionization. Tandem mass spectrometry (MS/MS) was also performed on ions of interest using post-source decay (PSD).

[0082] LC/MS Analysis with High-Resolution Mass Spectrometry: The sample extract (prior to fraction collection) and resulting fractions were analyzed using the same LC conditions listed previously coupled to a high-resolution mass spectrometer (Q-Exactive™, Thermo Scientific). The Q-Exactive™ settings used were: 140,000 resolution, 1e6 AGC target, Maximum IT 60 ms, and a mass range of 300-4000 Da was monitored; the settings for the heated electrospray ionization probe (HESI-II) were: 4 kV spray voltage, 50 psi sheath gas, 15 (arbitrary units) auxiliary gas, 380 °C capillary temperature, and 300°C heater temperature. Active fractions were further analyzed with LC-MS/MS with the following modified conditions: 1 µL injection of the 1:30 diluted fraction, 35,000 resolution for full scan mode, and Maximum IT of 120 ms.

[0083] MSⁿ Analysis: Multiple-stage mass spectrometry experiments (MSⁿ) were performed using an Orbitrap Elite™ (Thermo Scientific). Fractions were diluted 1:30 in 70% methanol with 0.1% formic acid. Infusion for nanospray was accomplished using the Triversa Nanomate (Advion) with 1.5 kV voltage and 0.3 psi gas pressure. A mass range of 225 to 2000 was monitored in full MS mode with 120,000 resolution. Both collision-induced dissociation (CID) and electron-transfer dissociation (ETD) were used for MS/MS and MSⁿ experiments.

[0084] 16S rRNA Gene Amplification and Sequencing: Genomic DNA of potential bacterial antagonists was extracted using the Wizard® genomic DNA purification kit (Promega, Madison, WI). A pair of universal primers specific for bacterial 16S rRNA, Eubac27 and R1492, were used to amplify the corresponding gene. PCR amplification of the 16S rRNA was performed with a Hotstart Taq® plus DNA polymerase kit (QIAGEN, Valencia, CA) under the following conditions: after an initial 5-minute incubation at 95°C, the mixture was subjected to 30 cycles, each including 1 minute at 95°C, 1 minute at 58°C, and 1 minute at 72°C. A final extension was performed at 72°C for 10 minutes. Both strands of purified PCR products were directly Sanger sequenced using the following primers: 27F (5'-AGAGTTTATCTTGGCTCAG-3'; SEQ ID NO. 72), 1492R (5'-GGTTACCTTGTTACGACTT-3'; SEQ ID NO. 73), 357F (5'-CTCCTACGGGAGGAGCA-3'; SEQ ID NO. 74), 518R (5'-CGTATTACCGCGGCTGCTGG-3'; SEQ ID NO. 75), and 1100R (5'-AGGGTTGCGCTCGTTG-3'; SEQ ID NO. 76) Sequence fragments were edited and

assembled into contigs using Molecular Evolutionary Genetics Analysis software v.5.0(MEGA 5.0). The BLAST algorithm was used for a homology search against Genbank. Only results from the highest-score queries were considered for phylotype identification, with 99% minimum similarity.

[0085] Whole Genome Sequencing: Genomic DNA was isolated from an overnight culture of each strain using a QIAGEN DNeasy® blood and tissue kit (QIAGEN Inc., Valencia, CA). Genome sequencing was performed using 454 Titanium sequencing technology (Roche, Branford, CT), achieving > 25X average genome coverage. *De novo* assembly was created for each genome using the 454 Life Sciences Newbler software package, v.2.5.3 (Roche). The genomic DNA of *P. alvei* strains TS-15 and A6-6i was also sequenced using the Pacific Biosciences (PacBio) RS sequencing platform. A single 10-kb library was sequenced using C₂ chemistry on 8 single-molecule real-time (SMRT) cells with a 90-min collection protocol on the PacBio RS. The 10-kb continuous-long-read (CLR) data were *de novo* assembled using the PacBio hierarchical genome assembly process (HGAP)/Quiver software package, followed by Minimus 2, and they were polished with Quiver. The assembled contigs from both approaches were annotated with the NCBI Prokaryotic Genomes Automatic Annotation Pipeline.

[0086] Identification and Characterization of the pbt Gene Cluster: Genomic comparison of pbt gene cluster (NRPS genes involved in the biosynthesis of a non-ribosomal lipopeptide antibiotic) between these two strains and *P. thiaminolyticus* strain OSY-SE (accession #ALKF000000000) as described in U.S. Publication No. US2013/0164317 was performed. The nonribosomal peptide synthetase (NRPS) machinery is composed of modular multi-domain enzymes which act as an assembly line to incorporate each amino acid monomer by one module. A typical module (C-A-T) in an NRPS contains a carrier Thiolation (T) domain and two catalytic domains, an adenylation (A) domain for amino acid activation and selectivity and a condensation (C) domain catalyzing peptide bond formation. In the termination module (C-A-T-Te), the Te-domain is responsible for releasing the assembled peptide. Additionally, optional epimerase (E) domain may also be present for L- to D-epimerization of amino acids. The NRPS in *P. alvei* strains A6-6i and TS-15 genomes was analyzed by NRPSpredictor2, a webserver for predicting NRPS adenylation domain. The A domain possesses a conserved binding pocket for amino acid recognition and activation. The substrate specificity of A-domain for amino acid was identified using NRPSpredictor2, based

on the fingerprint residues at the substrate-binding site. In addition, epimerization (E) domains and the Te domain were identified (PKS/NRPS analysis webserver).

Example 1: Broad Antimicrobial Spectrum of *P. alvei* Strains A6-6i and TS-15

[0087] *In vitro* agar plug assays showed inhibition zones against all the indicator strains including six major foodborne pathogens and three major tomato bacterial phytopathogens when challenged with both *P. alvei* isolates (Figures 1A and 1B). Notably, the antagonist migrated outward from the plug after forming the inhibition zone with SD (*S. dysenteriae*) or LM (*L. monocytogenes*), and the antagonistic growth ring expanded with time, especially in the case of *Listeria*. Both A6-6i and TS-15 had a wide range of inhibition against MRSA strains with zone diameters from 15 to 35 mm, and 15 to 20 mm, respectively. It is also interesting to note that strain A6-6i showed strong inhibitory effects on various MRSA strains tested despite the fact that some strains were resistant to up to 14 different antimicrobial drugs.

[0088] When supernatants were tested against the panel of gram-negative and gram-positive bacteria using the Bioscreen assay, both A6-6i (Figure 2) and TS-15 (not shown) CFCs exhibited a broad spectrum of antimicrobial activity, in which the lag phase was significantly extended in all the pathogens tested and the cell density was largely reduced at the end of incubation. Furthermore, the lag phase in CS (*C. sakazakii*), SD (*S. dysenteriae*), LM (*L. monocytogenes*), and some MRSA strains were extended to almost 24 hours in both A6-6i and TS-15 CFCs. Compared to A6-6i, CFCs from TS-15 had a much stronger inhibitory effect when tested against SN (*S. Newport*) (not shown).

Example 2: Bioactivity Results

[0089] Polymyxin B showed clear antimicrobial dose response against *S. Montevideo* (control experiment). The minimum inhibitory concentration for polymyxin B to show clear ZI was 4 µg/mL on the lawn of *S. Montevideo* strain 29N (Figure 3). Seven of the 1-minute fractions showed antimicrobial activity against both *E. coli* O157:H7 strain EDL933 and Methicillin-resistant *S. aureus* strain #12 (Figure 4). Similarly, multiple 1-minute fractions from *P. alvei* strain A6-6i exhibited antimicrobial activities against both strains as well (results not shown).

Example 3: Identification of the Primary Peptide Sequence

[0090] MALDI-TOF MS analysis revealed a number of compounds that were present in each of the seven bioactive fractions, as shown in Figure 5. The MALDI spectra provided a view of the full complement of compounds from each bioactive fraction within a single spectrum where clusters of these compounds differed by 14 Da, indicated by unlabelled arrows. The compound with a molecular weight of 1623 will be referred to as the primary compound (ion indicated with an asterisk in Figure 5), although multiple variants of similar abundances are present (Figure 5).

[0091] The MALDI-TOF MS/MS analyses of these molecular species revealed similar fragmentation patterns, which confirmed that the compounds that differ by 14 Da were related. As illustrated in Figure 6, comparison of MS/MS spectra of the primary compound, MW 1623, and the compound of MW 1637, which differ in molecular weight by 14 Da, revealed a series of product ions that shared the same mass and a second product ion series that differed by 14 Da, suggesting that the mass discrepancy between compounds was localized to one region of the molecule. This enabled identification of complementary product ion pairs, with one direction corresponding to the product ion series retaining the region of the molecule that contained the 14 Da mass-shift and the other direction corresponding to product ions that were identical between the two peptides (Figure 7). Manual *de novo* sequencing resulted in a partial amino acid sequence, yielding a putative sequence assignment.

[0092] A compound of the present invention is shown in Figure 8 B.

[0093] Analysis of MALDI-TOF MS/MS spectra only revealed partial sequence tags. To improve sequence coverage, individual fractions were infused and analyzed with the Orbitrap Elite, allowing for the collection of MSⁿ data. Again, pairs of MSⁿ spectra were analyzed and complementary ion pairs were identified based on the presence of 14 Da mass differences. The combined MSⁿ data allowed a complete amino acid sequence to be determined (Figure 8B). This amino acid sequence was similar to a previously identified cyclic compound isolated from a different *Paenibacillus* strain which also showed broad-spectrum activity against MRSA and *E. coli*; the structure of that compound, designated as paenibacterin, is shown in Figure 8 A.

[0094] By determining which series of product ions did or did not contain the molecular component that results in the 14 Da difference, *de novo* sequencing by MSⁿ analysis was more straightforward. This was particularly critical because the compounds were cyclic and resulting MSⁿ spectra can be difficult to interpret due to multiple ring

opening events occurring at a distribution of sites. An example is shown in the MS³ spectrum in Figure 9 where the ring opens at different amino acid positions, yielding a number of different sequence series within the same spectrum; thus, *de novo* sequencing of the primary sequence of cyclic peptides can be challenging. However, by using the described approach, the assignment of product ions and the identification of the primary sequence and sequence variants were accomplished without linearizing the molecule. The cumulative ion assignments for the primary amino acid sequence can be found in in Figure 10A, B.

[0095] Subsequent analysis with UPLC coupled to high resolution MS provided accurate mass data which confirmed that there were actually three predominant compound series, with each series containing groups of compounds that differ by 14.02 Da. The most pronounced differences between the three observed series were either a decrease of 15.99 Da or an increase of 1.98 Da in mass compared to the primary compound series; examples are designated with arrows in Figure 5. A comprehensive list of these compounds and their accurate mass molecular weights can be found in Table 2. These are designated as F, Y, and Y, -CH₂+O in the table and throughout the figures; F and Y correspond to phenylalanine or tyrosine at position 6 in Figure 8 B and -CH₂+O corresponds to a molecular difference in the fatty acid chain.

Table 2

F		Y		Y -CH ₂ +O	
649.390 ²⁺		657.394 ²⁺		657.394 ²⁺	
Nominal Molecular Weight	Complementary Ion Pair	Nominal Molecular Weight	Complementary Ion Pair	Nominal Molecular Weight	Complementary Ion Pair
		1581	269.222 ¹⁺		
1579	283.238 ¹⁺	1595	283.238 ¹⁺		
1593	297.255 ¹⁺	1609	297.253 ¹⁺		
1607	311.269 ¹⁺	1623	311.269 ¹⁺	1625	313.248 ¹⁺
		1637	325.285 ¹⁺	1639	327.264 ¹⁺
		1651	339.301 ¹⁺	1653	341.279 ¹⁺

[0096] Two compound series in the bioactive fractions differed from each other by 15.99 Da and exhibited MS² spectra with all the compounds in one series yielding a product ion at m/z 657²⁺ while the other series generates a product at m/z 649²⁺ (Figure 11). Figure 12 illustrates the MSⁿ spectra of three precursor ions that differed by 14 Da, all of which generated an MS² product ion at m/z 657²⁺. Because these were conserved product ions

within a precursor series that included compounds that differed in molecular weight by 14 Da, we were able to conclude that the region of the compound that yields product ions 657^{2+} or 649^{2+} does not contain the fatty acid. The MS^3 spectra of 657^{2+} were consistent with one another confirming that the region of the peptide that generated this sequence was conserved between these compounds (Figure 12). The MS^3 spectra from the ion series that generated an MS^2 product ion at 649^{2+} were similar to the 657^{2+} MS^3 spectra, except for a series of product ions that differed by 16 Da (masses with asterisks in Figure 12). This enabled the distinction between tyrosine and phenylalanine at position 6 (Figure 8), designated as Y and F in the tables and figures, respectively; these amino acids differ in molecular weight by 16 Da. Assignments for the MS/MS spectrum for a peptide containing Phe are illustrated in Figure 10A-B, where a direct comparison can be observed between the peptides containing Phe and Tyr at position 6.

Example 4: Identification of the Attached Fatty Acid

[0097] The MS^2 spectra of the series of compounds that contain a Tyr at position 6 were dominated by product ion 657^{2+} and its complementary ion pair (Figure 11). While 657^{2+} was conserved within the Tyr compound series, its complementary ion contained the same 14 Da mass shift as its precursor (*i.e.*, m/z 311, 325, and 339 in Figure 12). The same trend was also present for the Phe ion series. A list of the multiple complementary ions for 657^{2+} and 649^{2+} are listed in Table 3. When these complementary ions were dissociated (examples shown in Figure 12), a loss of ornithine was observed. Subtracting the cyclized peptide sequence mass from the mass of the entire compound yields the mass attributed to an attached fatty acid; molecular formula generation of these masses yield the molecular formulae of the different fatty acid variants (Table 4 and Figure 8). These fatty acids are similar to what was observed in Guo, *et al.*, although the lengths of the carbon chains differ and both the TS-15 and A6-6i strains presented in the current work exhibit a greater variability in chain length and composition.

[0098] As mentioned previously, three major series of compounds were determined: two series containing a tyrosine at position 6 and one containing phenylalanine. The two compound series containing Tyr differed by a molecular weight of 1.979 Da. Similar to the MS spectral analysis methodology shown in Figure 2, these compound series also had similar MS^2 fragmentation patterns, where some product ion masses are conserved and others differ by 1.979 Da (Figure 13). This mass difference corresponded to one less CH_2 and an

additional oxygen (labeled as $-\text{CH}_2+\text{O}$) in the attached fatty acid compared to the tyrosine molecular series.

Table 3

F		Y		Y -CH ₂ +O	
649.390 ²⁺		657.394 ²⁺		657.394 ²⁺	
Nominal Molecular Weight	Complementary Ion Pair	Nominal Molecular Weight	Complementary Ion Pair	Nominal Molecular Weight	Complementary Ion Pair
		1581	269.222 ¹⁺		
1579	283.238 ¹⁺	1595	283.238 ¹⁺		
1593	297.255 ¹⁺	1609	297.253 ¹⁺		
1607	311.269 ¹⁺	1623	311.269 ¹⁺	1625	313.248 ¹⁺
		1637	325.285 ¹⁺	1639	327.264 ¹⁺
		1651	339.301 ¹⁺	1653	341.279 ¹⁺

Table 4

Y 657.394 ²⁺							
MS ² Product Ions ¹⁺	269.222	283.238	297.253*	311.269	325.285*	339.301*	
MS ³ Product Ions ¹⁺	251.212 115.086	265.227 115.086	279.243 115.086	293.258 115.086	307.274 115.086	321.290 115.086	Water loss Ornithine
Molecular Formula of the Fatty Acid	C ₁₀ H ₁₉ O	C ₁₁ H ₂₁ O	C ₁₂ H ₂₃ O	C ₁₃ H ₂₅ O	C ₁₄ H ₂₇ O	C ₁₅ H ₂₉ O	

Example 5: Multiple Compounds with the Same Molecular Weight

[0099] It was also observed that more than one of the complementary ion pairs were occasionally present within the same MS² spectrum in the infusion experiment data. This corresponded to two compounds of the same precursor mass being fragmented within the same isolation window. There were multiple examples in the UPLC and NanoLC/MS data that showed several eluting peaks for entities with the same mass (example shown in Figure 14). The MS² spectra of the ions in each of these chromatographic peaks showed small differences in the fragmentation pattern and thus, the primary sequences of these respective peptides. Nearly identical amino acid sequences were confirmed for compounds with an identical molecular weight (1579): a compound with Lys at position 7 and a compound with

ornithine at position 7 with an additional CH₂ in the attached fatty acid (Lys and ornithine differ by a CH₂ in their side chains). Specifically, the product ion at m/z 649²⁺ corresponds to Lys at position 7 and m/z 642²⁺ corresponds to ornithine at position 7. The two chromatographic peaks with Lys at position 7 may indicate a structural difference in the attached fatty acid resulting in the observed difference in retention time (Figure 14). A similar substitution was also found at position 1. Some of the subsequent MS³ analyses of the fatty acid containing fragment ions (m/z 325¹⁺ and 339¹⁺ in Figure 5B) indicated that Lys can also be present at position 1 rather than ornithine (m/z 129¹⁺ in MS³ spectra in Figure 12). A Lys at position 1 and a decrease of CH₂ in the attached fatty acid resulted in identical molecular weights for each compound.

Example 6: Identification and Characterization of the pbt Gene Cluster

[0100] The sequence shown in Figure 8B was confirmed by genome mining for non-ribosomal peptide synthesis. Many pharmacologically important peptides in bacteria are synthesized by nonribosomal peptide synthetases (NRPS). NRPS machinery is composed of modular multi-domain enzymes which act as an assembly line to incorporate each amino acid monomer by one module. A typical module in an NRPS contains an adenylation (A) domain which possesses a conserved binding pocket for the recruitment of amino acid monomers that are to be incorporated into the final peptide product. A single contig of 6536324 bp (G+C content, 46.63%) and a single contig of 6784766 bp (G+C content, 46.69%) representing the complete chromosome for *P. alvei* strains A6-6i and TS-15 was generated, respectively. The draft genome sequences of strain A6-6i and TS-15 are available in DDBJ/EMBL/GenBank under GenBank accession # ATMS000000000 and ATMT000000000, respectively. The DNA sequences of A6-6i ptbA, ptbB, and ptbC are SEQ ID NOs. 5, 6, and 7, respectively. The DNA sequences of TS-15 ptbA, ptbB, and ptbC are SEQ ID NOs. 8, 9, and 10, respectively. The protein sequences for A6-6i ptbA, ptbB, and ptbC are SEQ ID NOs. 11, 12, and 13, respectively. The protein sequences for TS-15 ptbA, ptbB, and ptbC are SEQ ID NOs. 14, 15, and 16, respectively. A local BLASTX analysis against pbt gene cluster in *P. thiaminolyticus* strain OSY-SE (accession #ALKF000000000; U.S. Publication No. US2013/0164317) identified a 49-kb DNA region responsible for the compounds biosynthesis (Table 5). The protein sequences for OSY-SE ptbA, ptbB, and ptbC are SEQ ID NOs. 17, 18, and 19, respectively. Comparative genomic analysis showed 64% to 70% similarities in DNA sequences of the pbt gene cluster between the two *P. alvei* strains and *P.*

thiaminolyticus strain OSY-SE; and only 60% to 67% similarities in amino acid sequences (Figures 16-18). This DNA region encodes three peptide synthetase units which consist of thirteen modules (Table 5) responsible for incorporating the thirteen amino acids in the compounds. The predicted peptide sequence agreed with the chemical structure of the compounds determined by MS/MS (Table 4). In addition, epimerization (E) domains were found in modules for Orn₁, Orn₄, Orn₇, and Ser₈, which indicated that those amino acids might be in the D-form.

[0101] The bacterially-produced cyclic peptides are synthesized by a class of enzymes known as the nonribosomal peptide synthetases (NRPSs). NRPSs are found in many organisms and synthesize a number of medically-important peptides such as antibiotics and immunosuppressants. By sequence analysis, 14 NRPS genes have been identified in the *P. alvei* A6-6i and TS-15 genomes. Three NRPS genes which covered 49 kb were found to control the production of the compounds in the current application. NRPSs are large enzymes that are organized into modules made up of functional domains. The entire length of the amino acid sequence showed homology to the Pbt encoded by the *pbt* gene cluster of *P. thiaminolyticus* strain OSY-SE in the prior art patent (66% similarity to *pbtA* gene, 66% similarity to *pbtB* gene, and 59% similarity to *pbtC* gene). Detailed analysis of the *pbtABC* gene cluster showed that each gene had domain and module organization as in Table 5B. To predict the substrate specificity-conferring amino acids in the adenylation (A) domain of each module, the structural regions, A3 and A6 motifs in the A domain, were blasted against the NCBI protein database and showed only 36% identity to the ptbB1 module, 38% identity to the ptbC1 module, and 35% identity to the ptbC2 module, indicating structurally unique antibiotics from paenibacterin in the prior art patent.

Table 5

A. Amino acid similarities between A6-6i, TS-15, and OSY-SE

	A6-6i	TS-15
pbtA	66.95%	66.73%
pbtB	66.66%	66.79%
pbtC	59.72%	59.31%

B. Modules and domains predicted in the gene cluster

pbtA	CA _{Orn1} TECA _{Val2} TCA _{Thr3} TCA _{Orn4} TECA _{Ser5} T
pbtB	CA _{Tyr6} TCA _{Orn7} TECA _{Ser8} TECA _{Ile9} TCA _{Pro10} T
pbtC	CA _{Ile11} TCA _{Pro12} TCA _{Ile13} TTe

[0102] Table 5 shows the predicted amino acids of the antimicrobial peptides generated by the PKS/NRPS web server. For example, the pbtA gene (encoding a non-ribosomal peptide synthetase) contains 5 modules, each comprised of a C (condensation) domain, an A (adenylation) domain, and a T (thiolation) domain. The A domain is responsible for amino acid activation and selectivity. Because the binding pocket of each A domain is a conserved sequence, the amino acid substrate can be predicted. Additionally, epimerization (E) domains were found in modules for Orn₁, Orn₄, Orn₇, and Ser₈, which indicate that those amino acids may be in the D-form. Likewise, pbtB gene contains 5 modules and pbtC gene contains 3 modules.

C. Predicted amino acid substrates in the adenylation (A) domain in each module.

A6-6i

Module	Active site residue with 8 A of the amino acid substrate	Binding pocket	Predicted substrate
PbtA1	LAWAFDVFTGDRESVVGSDLSYGVTEACVDACY SEQ ID NO. 20	DVGEVGSVDK SEQ ID NO. 21	D-Orn
PbtA2	LGASFDAATFEGWMLVGGDINGYGPTENTTFTCC SEQ ID NO. 22	DAFWLGGTFK SEQ ID NO. 23	Val
PbtA3	LNSHFDFSVWEGNQIFGGEINMYGITETTTHVHTY SEQ ID NO. 24	DFWNIGMVHK SEQ ID NO. 25	Thr
PbtA4	MAWAFDVFSGDRESIIGSDINSYGVTEACVDSSY SEQ ID NO. 26	DVGEIGSVDK SEQ ID NO. 27	D-Orn
PbtA5	RWMTFDVSVWEWHFFTSGEINLYGPTEATVDVTY SEQ ID NO. 28	DVWHFSLVDK SEQ ID NO. 29	Ser
PbtB1	AWRFFDGFVMSICTLAGEFNEYGPTENSVVATC SEQ ID NO. 30	DGMITAEVVK SEQ ID NO. 31	Tyr
PbtB2	MAWAFDVFSGDRDCAVGSDINSYGVTEATCIDASY SEQ ID NO. 32	DVGDAAGSIDK SEQ ID NO. 33	D-Orn
PbtB3	RWMTFDVSVWEWHFFTSGEINLYGPTEATVDVTY SEQ ID NO. 34	DVWHFSLVDK SEQ ID NO. 35	D-Ser
PbtB4	VETSFDGSTFDGFI LGGEKHVYGPTTESTVFATC SEQ ID NO. 36	DGFFLGVVFK SEQ ID NO. 37	Ile
PbtB5	LYQAFDVCYQESFIITAGEHNHYGPSETHVTTY	DVQFIAHVVK	Pro

	SEQ ID NO. 38	SEQ ID NO. 39	
PbtC1	INTSFDGSAFDGLILFGGEKHAYGPSESTVYATW SEQ ID NO. 40	DGFLLGAVYK SEQ ID NO. 41	Ile
PbtC2	LYQAFDVCYQESYIITAGEHNHYGPSETHVVTY SEQ ID NO. 42	DVQYIAHVVK SEQ ID NO. 43	Pro
PbtC3	VDASFDGSTFDGFIILFGGEKHVYGPTESTVFATS SEQ ID NO. 44	DGFFLGVVFK SEQ ID NO. 45	Ile

TS-15

Module	Active site residue with 8 Å of the amino acid substrate	Binding pocket	Predicted substrate
PbtA1	LAWAFDVFTGDRESVVGSDLSYGVTEACVDACY SEQ ID NO. 46	DVGEVGSVDK SEQ ID NO. 47	D-Orn
PbtA2	LAASFDAATFEGWMLVGGDINGYGPTENTTFTCC SEQ ID NO. 48	DAFWLGGTFK SEQ ID NO. 49	Val
PbtA3	LNSHFDVSVWEGNQIFGGEINMYGITETT VHVTY SEQ ID NO. 50	DFWNIGMVHK SEQ ID NO. 51	Thr
PbtA4	MAWAFDVFSGDRESIIGSDINSYGVTEACVDSSY SEQ ID NO. 52	DVGEIGSVDK SEQ ID NO. 53	D-Orn
PbtA5	RWMTFDVSVWEWHFFTSGEINLYGPTEATVDVTY SEQ ID NO. 54	DVWHFSLVDK SEQ ID NO. 55	Ser
PbtB1	AWRFFDGFVMSCTLAGEFNEYGPTENS SVVATC SEQ ID NO. 56	DGMITAEVVK SEQ ID NO. 57	Tyr
PbtB2	MAWAFDVFSGDRDCAVGSDINSYGVTEATCIDASY SEQ ID NO. 58	DVGDAGSIDK SEQ ID NO. 59	D-Orn
PbtB3	RWMTFDVSVWEWHFFTSGEINLYGPTEATVDVTY SEQ ID NO. 60	DVWHFSLVDK SEQ ID NO. 61	D-Ser
PbtB4	VETSFDGSTFDGFIILFGGEKHVYGPTESTVFATC SEQ ID NO. 62	DGFFLGVVFK SEQ ID NO. 63	Ile
PbtB5	LYQAFDVCYQESFIITAGEHNHYGPSETHVVTY SEQ ID NO. 64	DVQFIAHVVK SEQ ID NO. 65	Pro
PbtC1	INTSFDGSAFDGLILFGGEKHAYGPSESTVYATW SEQ ID NO. 66	DGFLLGAVYK SEQ ID NO. 67	Ile
PbtC2	LYQAFDVCYQESYIITAGEHNHYGPSETHVVTY SEQ ID NO. 68	DVQYIAHVVK SEQ ID NO. 69	Pro
PbtC3	VDASFDGSTFDGFIILFGGEKHVYGPTESTVFATS SEQ ID NO. 70	DGFFLGVVFK SEQ ID NO. 71	Ile

[0103] Table 5 shows the identification and characterization of the pbt gene cluster.

A. DNA sequence similarities and amino acid sequence similarities of the pbt gene cluster between *P. alvei* strains A6-6i and TS-15 and *P. thiaminolyticus* strain OSY-SE. B. Modules and domains identified in the NRPS subunits: C, A, T, E, and Te representing condensation domain, adenylation domain, thiolation domain, epimerization domain, and thioesterase domain, respectively. C. Substrate prediction for each of the 13 modules in the peptide.

[0104] It is worth noting that genome mining did not predict the presence of abundant sequence variants. As Lys and ornithine differ by CH₂, their binding affinities are likely similar which may be contributing to the observed molecular diversity. Likewise, Tyr and Phe differ by a hydroxyl group. Furthermore, NRPSpredictor 2 has lower single amino acid

substrate prediction scores for Phe and Lys, which may indicate why these were not additionally predicted in the primary sequence. NRPSpredictor2 also predicted that the epimerization (E) domain was in modules 1, 4, 7, and 8, which indicated that these resulting amino acid substrates may be in the D-form. As expected, NRPS analysis does not offer information about the presence or length of the alkyl chain.

[0105] The NRPS analysis confirmed both the presence and order of the amino acids of the peptide assignments made through the combination of MALDI-MS, high-resolution mass spectrometry, and MSⁿ analysis. NRPS can be used as a screening technique to identify potential nonribosomal peptides which may act as antibiotics; however, it does not yield information regarding any molecular variants that may be produced. Moving forward, NRPS analysis and mass spectrometry can be used to combine rapid prediction of candidate peptides with the molecular specificity of mass spectrometry to enable identification of cyclic antibiotics and their sequence and fatty acid variants despite the presence of molecular diversity and complicated spectra.

Example 6: Design and characterization of a synthetic peptide

[0106] Based on the identified sequence in Figure 8B, a peptide was chemically synthesized in accordance with Figure 15, referred to herein as synthetic depsipeptide A.

[0107] Minimum inhibitory concentrations (MICs) of synthetic depsipeptide A against selected bacteria, including antibiotic-resistant strains (Table 6), were determined by the broth microdilution method following the procedure of the Clinical and Laboratory Standards Institute (CLSI) as is known in the art. Briefly, synthesized depsipeptide was dissolved in Optima grade water to reach 25.6mg/ml as the stock concentration. After 100 times dilution in cation-adjusted Mueller-Hinton II broth (CAMHBII) (Becton, Dickinson & Co., Sparks, MD), the depsipeptide was then further two-fold serially diluted in CAMHBII in clear, sterile, non-treated round bottom 96-well plates (Nunc, Roskilde, Denmark). Bacterial cultures were suspended in demineralized water to achieve a turbidity equivalent to a 0.5 McFarland turbidity standard (Remel, Lenexa, KA). An equal volume of culture suspension was added to the diluted depsipeptide to give a final volume of 100µl/well in the assay plates. Polymyxin B (Sigma, St Louis, MO) and vancomycin (Sigma) were used as positive controls in the AST assays. Strains *Escherichia coli* ATCC 25853, *Pseudomonas aeruginosa* ATCC 27853, *Enterococcus faecalis* ATCC 29212, and *Staphylococcus aureus* ATCC 29213 were used as quality control strains in the assays. The MIC end point refers to the lowest

concentration of an antimicrobial agent that completely inhibits growth of bacterial cells after incubation at 35C for 20-24 h.

Table 6

Table 6. Minimum inhibitory concentrations (MICs) of synthesized depsipeptide and other antibiotics

Bacterial strain	MIC(μ g/ml)		
	Synthesized depsipeptide	PolymyxinB	Vancomycin
<i>E.coli</i> ATCC 25853	4	0.5	-
<i>Salmonella</i> Newport #17	16	1	-
<i>Enterobacter sakazakii</i> E784	8	0.5	-
<i>E. coli</i> O157:H7 EDL933	4	0.5	-
<i>Serratia marcescens</i> SBJ-9047 (CRE, clinical isolate)	32	>128	-
<i>Klebsiella pneumoniae</i> SBJ-9149 (clinical isolate)	4	0.5	-
<i>Serratia marcescens</i> SBJ-8283 (CRE, clinical isolate)	32	>128	-
<i>Enterobacter cloacae</i> SBJ-9395 (CRE, clinical isolate)	4	>128	-
<i>Enterobacter cloacae</i> SBJ-7612 (clinical isolate)	4	1	-
<i>Klebsiella pneumoniae</i> SBJ-9483 (clinical isolate)	4	0.5	-
<i>Klebsiella pneumoniae</i> SBJ-9388 (clinical isolate)	4	0.5	-
<i>Enterobacter cloacae</i> SBJ-9222 (clinical isolate)	8	1	-
<i>P. aeruginosa</i> ATCC-27853	8	1	-
<i>P. aeruginosa</i> SBJ-10884 (PMB-R, clinical isolate)	4		
<i>P. aeruginosa</i> -03 (PMB-R, clinical isolate)	8		
<i>P. aeruginosa</i> -02 (PMB-R, clinical isolate)	8		
<i>P. aeruginosa</i> -01 (PMB-R, clinical isolate)	8		
<i>P. aeruginosa</i> SBJ-10886 (PMB-R, clinical isolate)	4		
<i>E. faecalis</i> ATCC 29212	4	-	2
<i>Listeria monocytogenes</i> R2-583	2	-	1
<i>S. aureus</i> ATCC25923	2	-	1
<i>S. aureus</i> 19 (MRSA, MDR clinical isolate)	4	-	1
<i>S. aureus</i> 14 (MRSA, MDR clinical isolate)	4	-	1
<i>S. aureus</i> 12 (MRSA, MDR clinical isolate)	4	-	1
<i>S. aureus</i> ATCC 29213	4	-	1
<i>S. aureus</i> 10 (MRSA, MDR clinical isolate)	4	-	1
<i>S. aureus</i> 8 (MRSA, MDR clinical isolate)	4	-	1

<i>S. aureus</i> 6 (MRSA, MDR clinical isolate)	4	-	1
<i>S. aureus</i> 4 (MRSA, MDR clinical isolate)	4	-	1
<i>S. aureus</i> 2 (MRSA, MDR clinical isolate)	4	-	1

PMB-R, polymyxin B-resistant; MDR, multidrug-resistant; MRSA, methicillin-resistant *S. aureus*; CRE, carbapenem-resistant.

[0108] The synthetic depsipeptide showed a broad antimicrobial spectrum against both Gram-negative and Gram-positive bacteria, including against significant antibiotic-resistant clinical isolates (Table 1). Specifically, the synthesized depsipeptide showed very potent activity against those carbapenem-resistant (CRE) strains and MRSA strains. Additionally, bacteria showed much greater sensitivity to this peptide than to paenibacterin in U.S. Publication No. US2013/0164317.

[0109] It is striking that the peptides found in this study are amphiphilic with distinct hydrophilic and hydrophobic regions: hydrophobicity on one side, the other side being predominantly polar and charged amino acids, and a hydrophobic fatty acid chain (Figure 8B). Major differences between paenibacterin and the compounds discovered in this work are the length of the attached fatty acid, the different combinations of lysine and ornithine at positions 1 and 7, and the amino acids at position 6 and 12. This is particularly interesting because the amino acids at position 6 and 12 have different properties (e.g., hydrophobic, hydrophilic, or positively charged). Furthermore, the presence of D-amino acids influences the structure and properties of the peptide and will also make the compound more resistant to enzymatic degradation and thus inherently more stable.

[0110] Aspects of paenibacterin's mode of action have been previously studied. The results suggest that the compound has a high affinity to the negatively-charged outer membrane of gram-negative bacteria. This is likely due to the presence of positively charged amino acids in the molecule, which is similar to the mode of action of polymyxin. Three positively charged amino acids were found in the molecules discovered here compared to four in paenibacterin (Figure 8), which may result in varying degrees of effectiveness. However, the Lys to Pro substitution at position 12 also increases the hydrophobicity of that portion of the molecule which may result in a better affinity to the hydrophobic core of cellular membranes, a characteristic that may aid in its disruption. Similarly, the presence of Tyr at position 6 contributes to a more polar region of the molecule. The mode of action may also be due to the amphiphilic nature of the compound, acting as a surfactant to disrupt cell membranes. It is notable that polymyxin also has distinct hydrophilic and hydrophobic domains.

[0111] It was also determined that paenibacterin resulted in the permeabilization of both gram-positive and -negative cell membranes, which was probably disrupted by the attached fatty acid. The chain length will likely affect observed antimicrobial activity, although it is uncertain if it will be less or more effective with a longer/shorter chain. For polymyxin, it is hypothesized that the fatty acyl chain disrupts the cellular membrane. Studies on the fatty acid chain of polymyxin indicate that antimicrobial activity correlates with the length and bulkiness of this moiety. However, reports are varied and subsequent experiments to design fatty acid analogues for the compounds in this study will yield insight into how this affects antimicrobial activity. It is also interesting that a single strain of bacteria can produce such a large number of molecular variants. This may enable the strain to exert a more concerted antimicrobial effect and may have resulted from extensive selection pressure in the community from which it was isolated.

[0112] The major significance of the antimicrobial peptides described herein relates to overcoming the worldwide public health crisis of drug resistance by several major classes of bacterial human pathogens. Several of the most dangerous bacterial pathogens such as MRSA, VRSA, and CRE are resistant to nearly every antibiotic currently in the arsenal of human antimicrobial prophylaxis. CRE, for instance, has no known antibiotic weakness. What is most important is that these antimicrobial peptides may represent entirely new classes of antibiotics, each with the ability to control these deadly bacteria and cure associated pathology. Safety studies in rat with the host organism that produces the antimicrobial peptides was most encouraging, with rats showing no overt pathology from this strain.

[0113] The use of the terms “a” and “an” and “the” and similar referents (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms first, second etc. as used herein are not meant to denote any particular ordering, but simply for convenience to denote a plurality of, for example, layers. The terms “comprising”, “having”, “including”, and “containing” are to be construed as open-ended terms (i.e., meaning “including, but not limited to”) unless otherwise noted. Recitation of ranges of values are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. The endpoints of all ranges are included within the range and independently combinable. All methods

described herein can be performed in a suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”), is intended merely to better illustrate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention as used herein.

[0114] While the invention has been described with reference to an exemplary embodiment, it will be understood by those skilled in the art that various changes may be made and equivalents may be substituted for elements thereof without departing from the scope of the invention. In addition, many modifications may be made to adapt a particular situation or material to the teachings of the invention without departing from the essential scope thereof. Therefore, it is intended that the invention not be limited to the particular embodiment disclosed as the best mode contemplated for carrying out this invention, but that the invention will include all embodiments falling within the scope of the appended claims. Any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

Claims

1. A peptide of the sequence
Xaa₁-Xaa₂-Xaa₃-Xaa₄-Xaa₅-Xaa₆-Xaa₇-Xaa₈-Xaa₉-Pro₁₀-Xaa₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 2),
wherein Xaa₆ is Tyr, Phe, or Trp; Xaa₁, Xaa₄ and Xaa₇ are each independently Lys or Orn; Xaa₂, Xaa₉ and Xaa₁₁ are each independently Leu, Ile, Val, or Ala; Xaa₃, Xaa₅, and Xaa₈ are each independently Cys, Tyr, Thr, or Ser; wherein the peptide optionally includes a saturated or unsaturated, substituted or unsubstituted, linear or branched, C₄-C₂₀ fatty acid group, or a saturated or unsaturated, linear or branched C₄-C₂₀ ester, covalently linked to Xaa₁.
2. The peptide of claim 1 having the sequence
Xaa₁-Val₂-Thr₃-Xaa₄-Ser₅-Xaa₆-Xaa₇-Ser₈-Ile₉-Pro₁₀-Xaa₁₁-Pro₁₂-Ile₁₃, (SEQ ID NO. 3),
wherein Xaa₆ is Tyr, Phe, or Trp; Xaa₁₁ is Leu, Ile, Val, or Ala; and Xaa₁, Xaa₄, and Xaa₇ are independently Lys or Orn, wherein the peptide optionally includes a saturated or unsaturated, substituted or unsubstituted, linear or branched C₄-C₂₀ fatty acid group, or a saturated or unsaturated, linear or branched C₄-C₂₀ ester, covalently linked to Xaa₁.
3. The peptide of claim 2, wherein Xaa₆ is Tyr or Phe; Xaa₁₁ is Ile; Xaa₁ and Xaa₄ are Orn; and Xaa₇ is Lys or Orn.
4. The peptide of any one of claims 1-3, wherein the peptide comprises a saturated or unsaturated, substituted or unsubstituted, linear or branched C₄-C₂₀ fatty acid group covalently linked to Xaa₁.
5. The peptide of claim 4, wherein the fatty acid group has the formula C₁₀H₁₉O, C₁₁H₂₁O, C₁₂H₂₃O, C₁₃H₂₅O, C₁₄H₂₇O, or C₁₅H₂₉O.
6. The peptide of claim 5, wherein Xaa₆ is Phe.
7. The peptide of any one of claims 1-3, wherein Xaa₆ is Tyr and the peptide comprises a saturated or unsaturated, linear or branched C₄-C₂₀ ester covalently linked to Xaa₁.
8. The peptide of claim 7, wherein the ester has the formula C₁₀H₁₉O₂, C₁₁H₂₁O₂, C₁₂H₂₃O₂, C₁₃H₂₅O₂, C₁₄H₂₇O₂, or C₁₅H₂₉O₂.

9. The peptide of any one of claims 1-8, wherein the peptide is cyclized through a bond between Xaa₃ or Thr₃ and Ile₁₃.

10. A composition comprising the peptide of any one of claims 1-9 and a carrier, vehicle, excipient, or diluent.

11. The composition of claim 10, wherein the composition is a cleaning product, a disinfecting product, a personal care product, an animal feed product, or a food product.

12. The composition of claim 10, wherein the composition is a pharmaceutical composition comprising the pharmaceutically acceptable excipient.

13. A process for preparing the peptide of any one of claims 1-9, comprising
(a) cultivating a host cell under conditions that allow for production of the peptide;
and
(b) purifying and isolating the peptide.

14. A process for preparing a composition comprising the peptide of any one of claims 1-9, comprising
(a) cultivating a host cell under conditions that allow for production of the peptide;
(b) purifying and isolating the peptide, and
(c) producing the composition comprising the isolated peptide and a carrier, vehicle, excipient, or diluent.

15. The process of claim 14, wherein the composition is a cleaning product, a disinfecting product, a personal care product, an animal feed product, or a food product.

16. The process of claim 14, wherein the composition is a pharmaceutical composition comprising the pharmaceutically acceptable excipient.

17. A method of inhibiting growth or proliferation of a microbe, said method comprising contacting, outside of a human or animal body, the microbe or a surface or product which may contain the microbe with the peptide of any one of claims 1-9.

18. The method of claim 17, wherein the surface or product which may contain the microbe is a surface associated with post-harvest foods, a food process surface, a food package, a household surface, a food preparatory tool, a surface in a hospital, a surgical supply, a surgical tool, a post-surgical bandage or a wound dressing.

19. A composition comprising the peptide of any one of claims 1-8 for use in the prevention or treatment of a disease or disorder associated with the presence, in or on a subject, of a microbe.

20. The composition according to claim 19, wherein the subject has a MRSA, VRSA, or CRE infection.

21. The composition according to claim 19 or 20 wherein the subject is a mammal.

22. The composition according to claim 21 wherein the mammal is a human subject.

23. Use of a composition comprising the peptide of any one of claims 1-8 for preventing or treating of a disease or disorder associated with the presence, in or on a subject, of a microbe.

24. The use according to claim 23, wherein the subject has a MRSA, VRSA, or CRE infection.

25. The use according to claim 23 or 24 wherein the subject is a mammal.

26. The use according to claim 25 wherein the mammal is a human subject.

27. Use of a composition comprising the peptide of any one of claims 1-8 for the manufacture of a medicament for the prevention or treatment of a disease or disorder associated with the presence, in or on a subject, of a microbe.

28. The use according to claim 27, wherein the subject has a MRSA, VRSA, or CRE infection.

29. The use according to claim 27 or 28 wherein the subject is a mammal.
30. The use according to claim 29 wherein the mammal is a human subject.

1/48

FIG. 1A

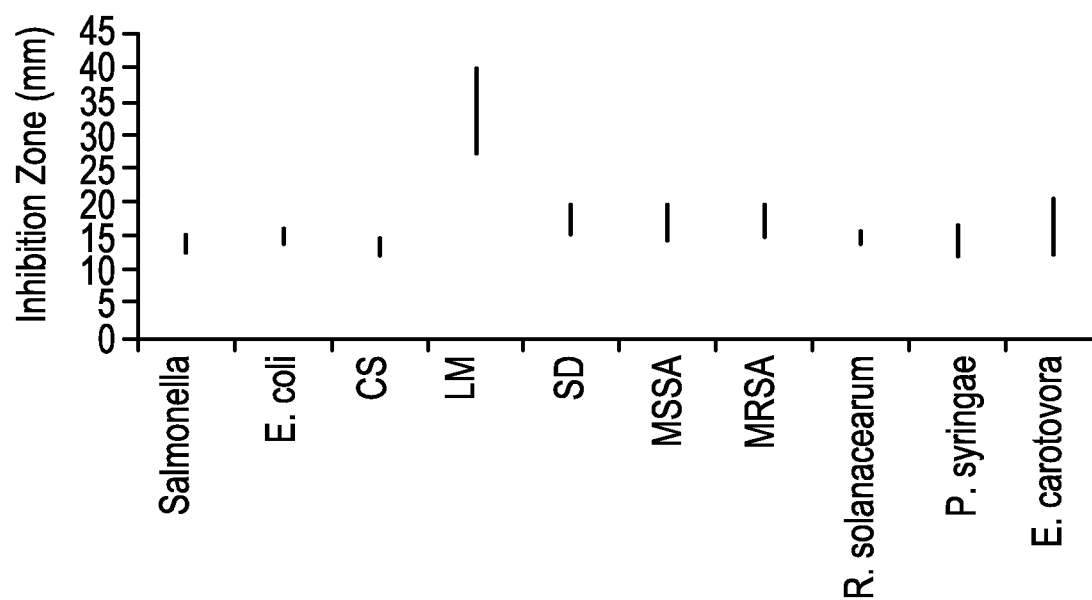
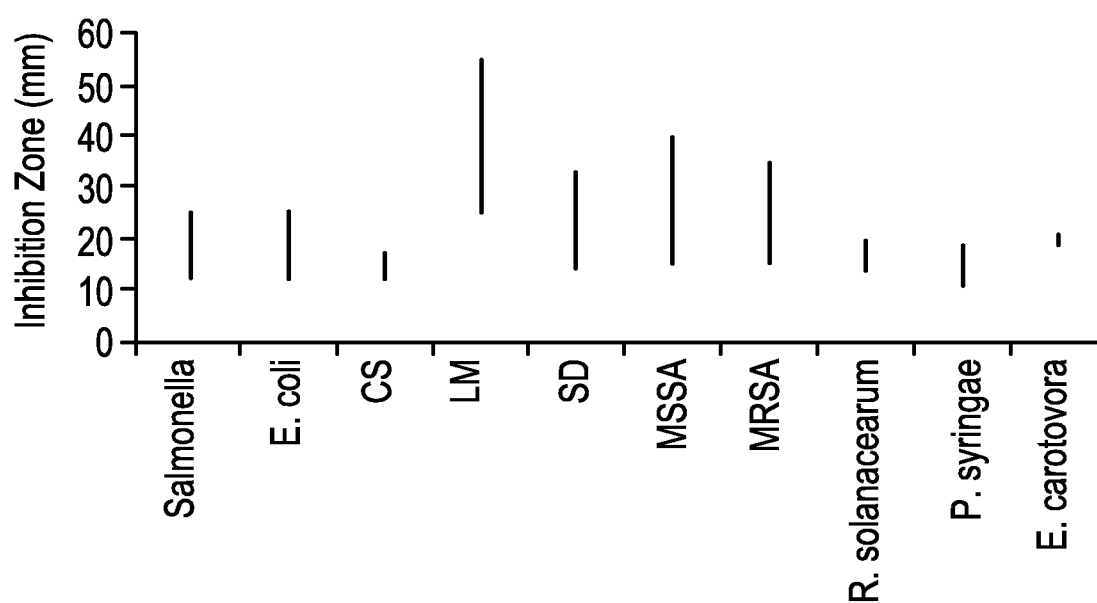


FIG. 1B



2/48

FIG. 2A

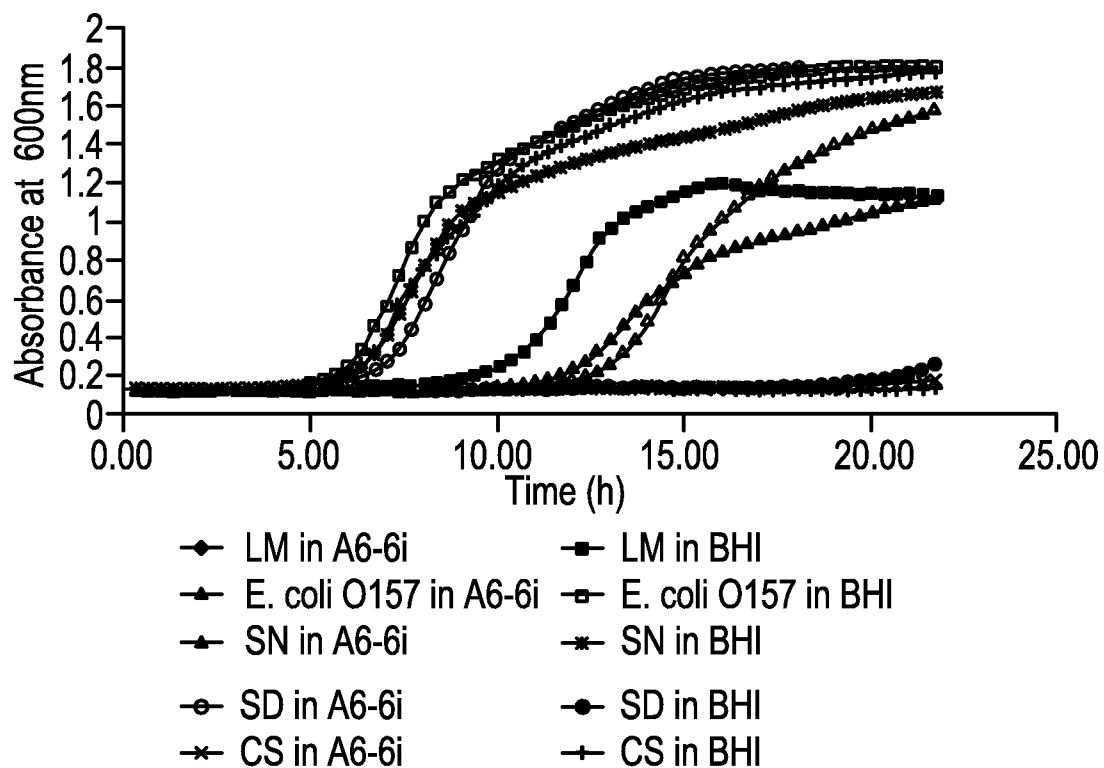
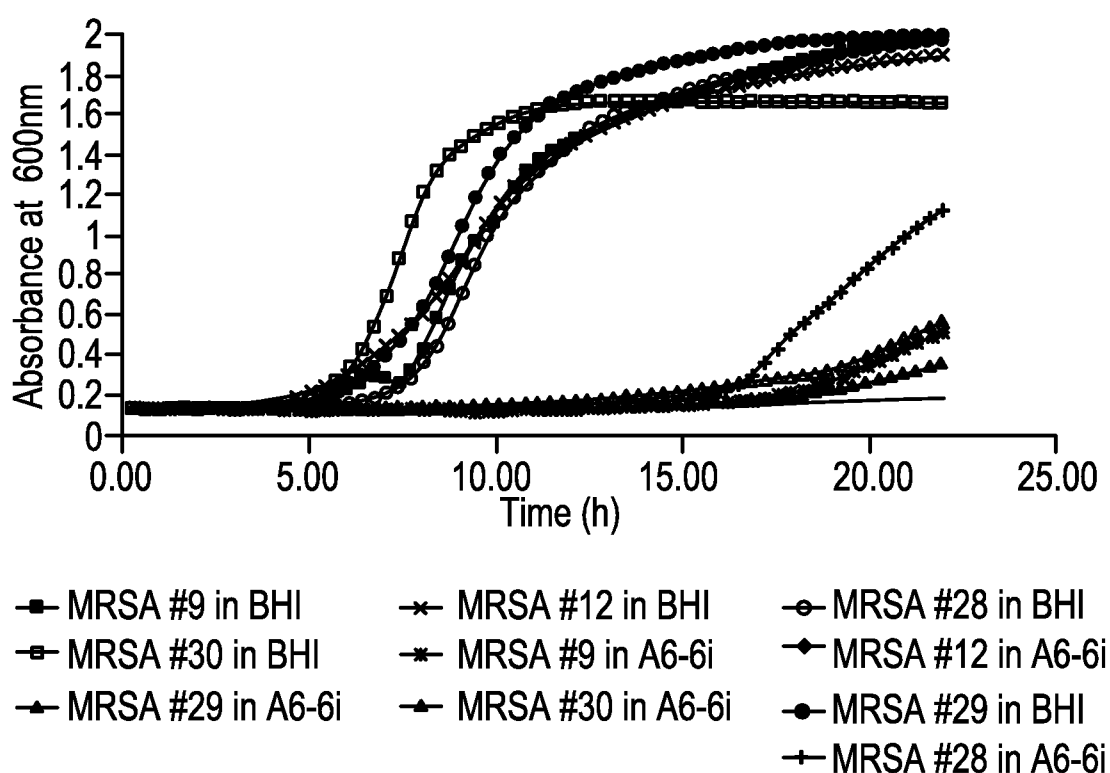
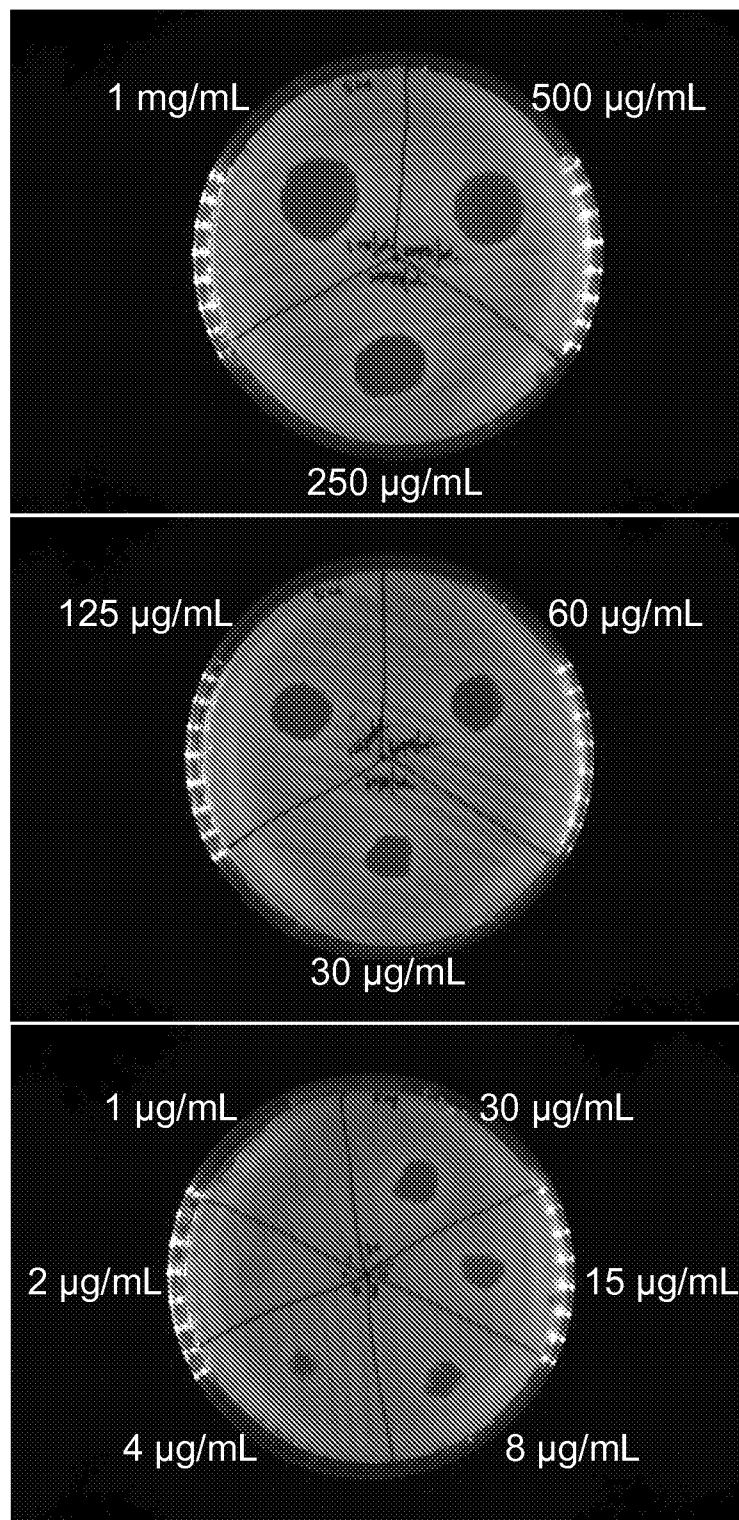


FIG. 2B



3/48

FIG. 3



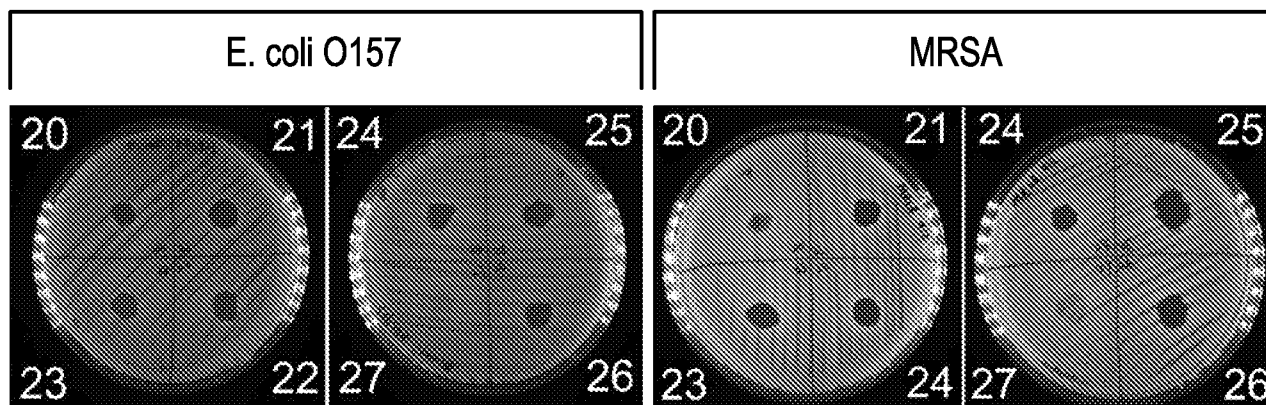
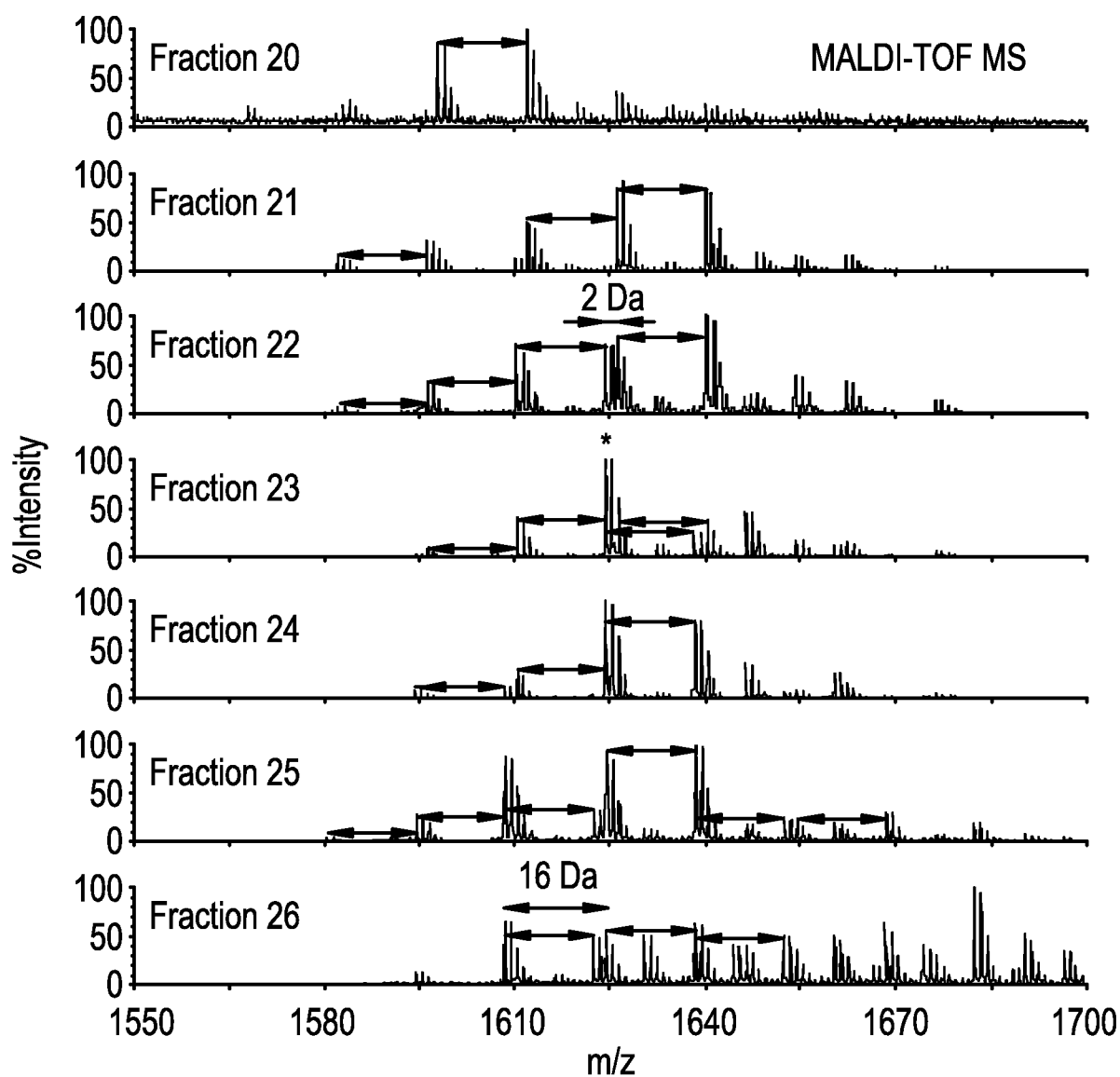
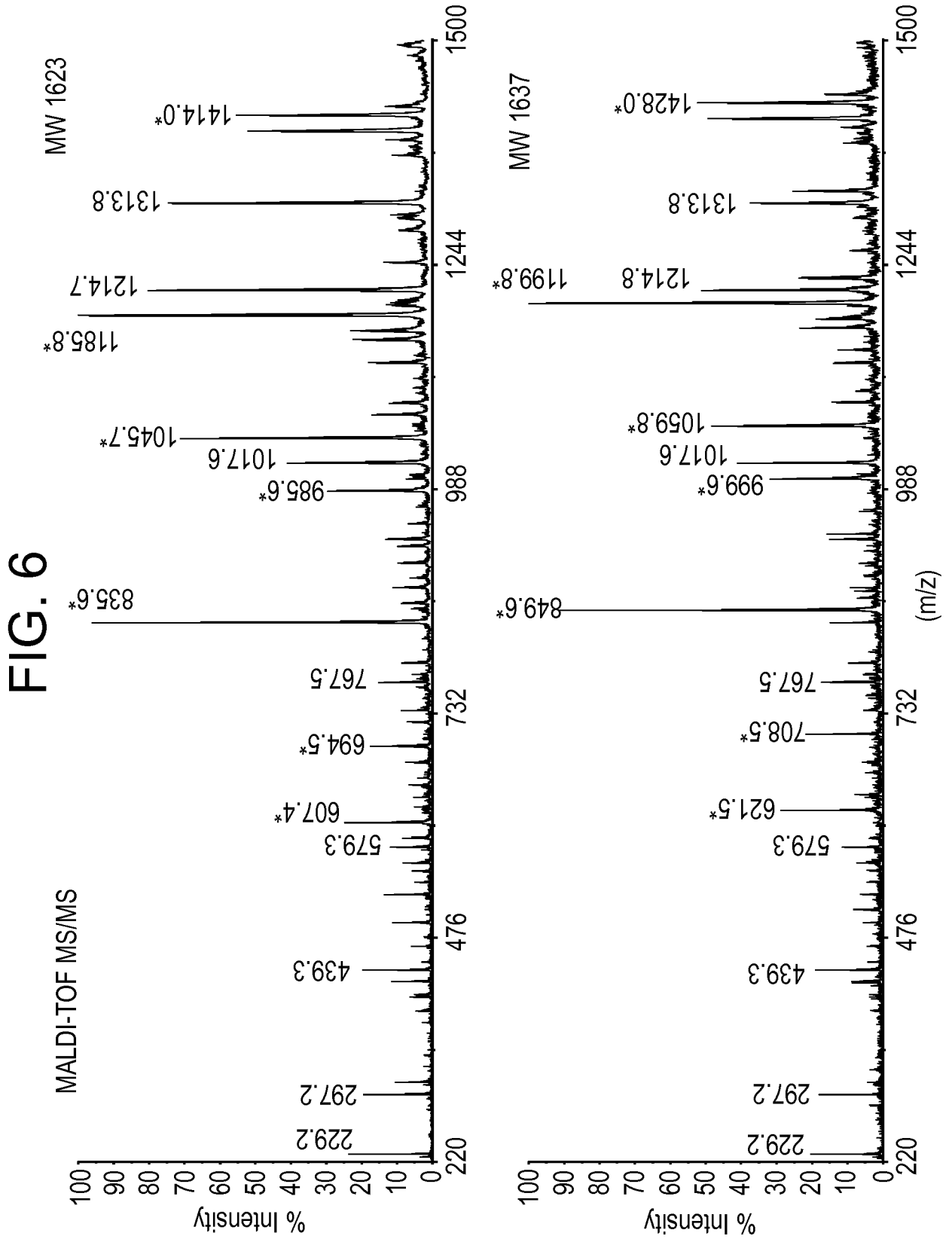


FIG. 5

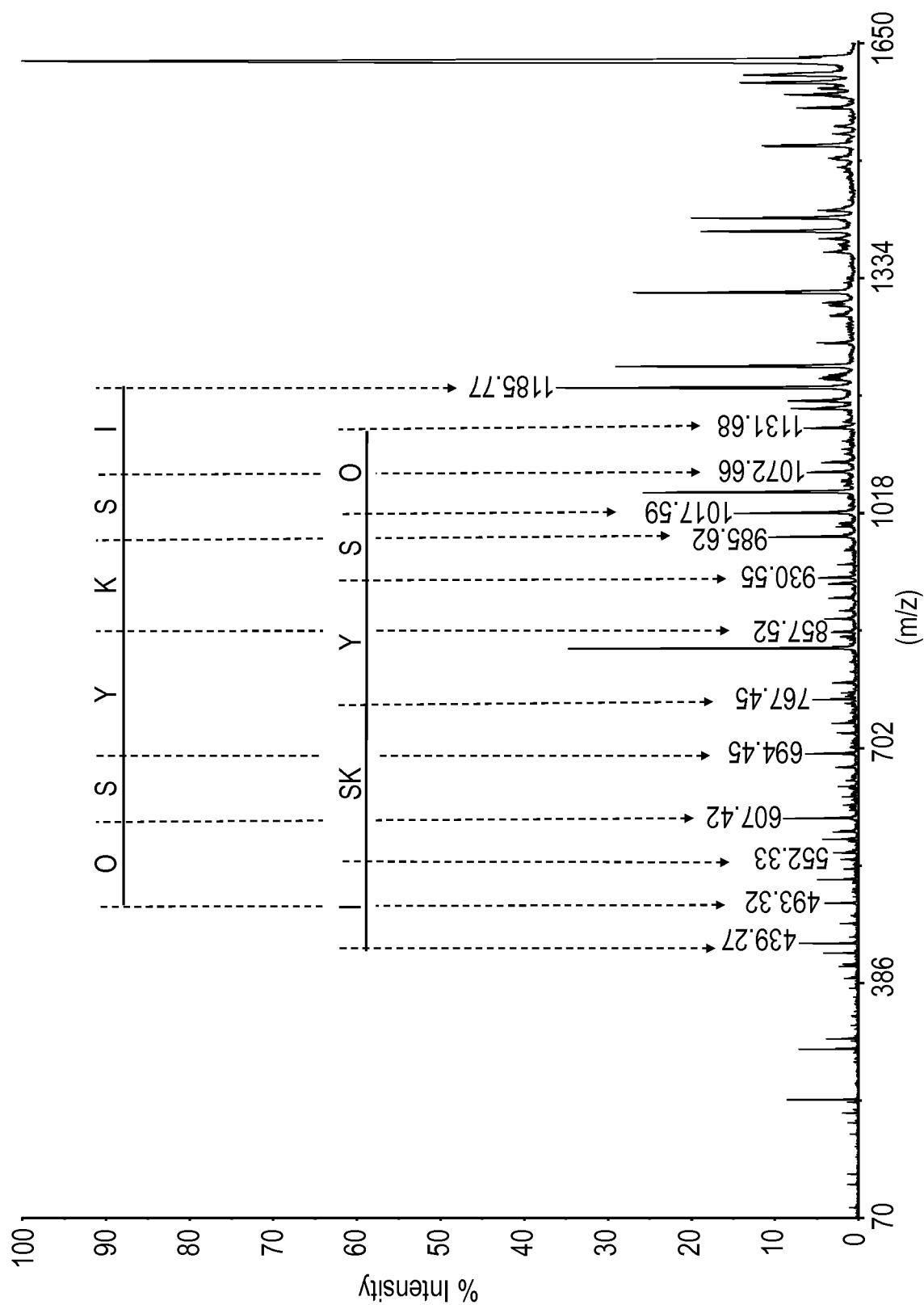


5/48



6/48

FIG. 7

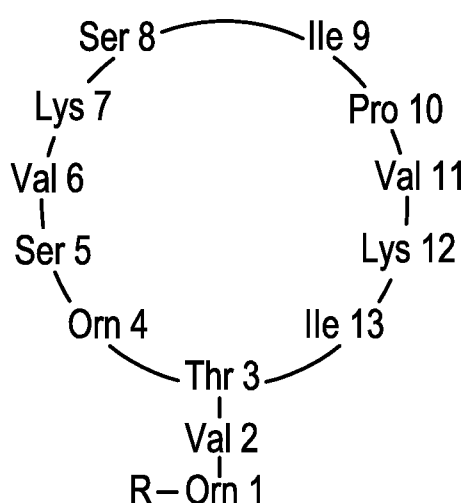


7/48

FIG. 8A

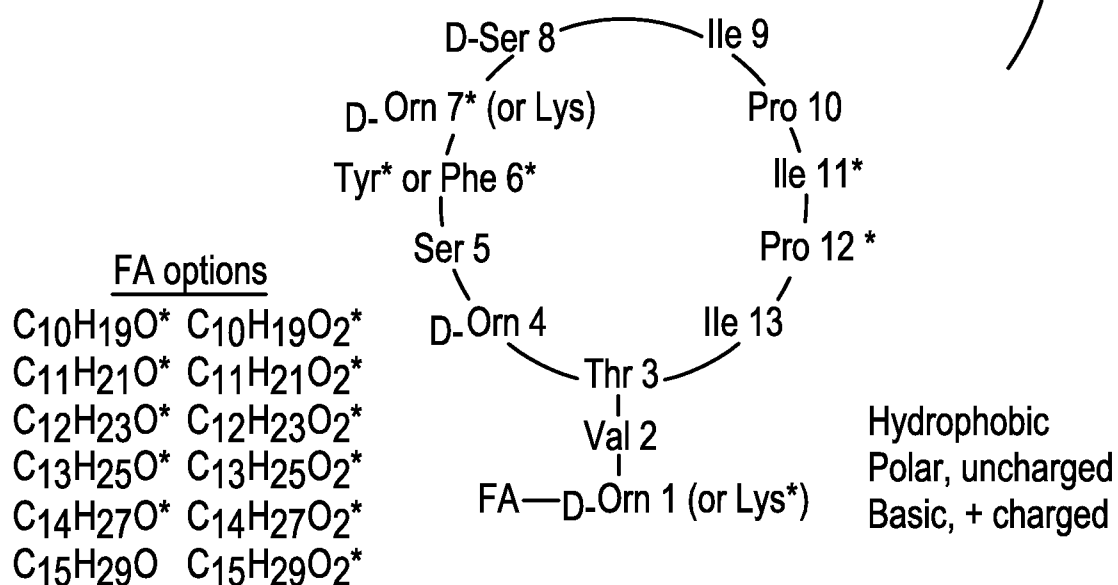
Prior Art

SEQ ID NO. 78

**FIG. 8B**

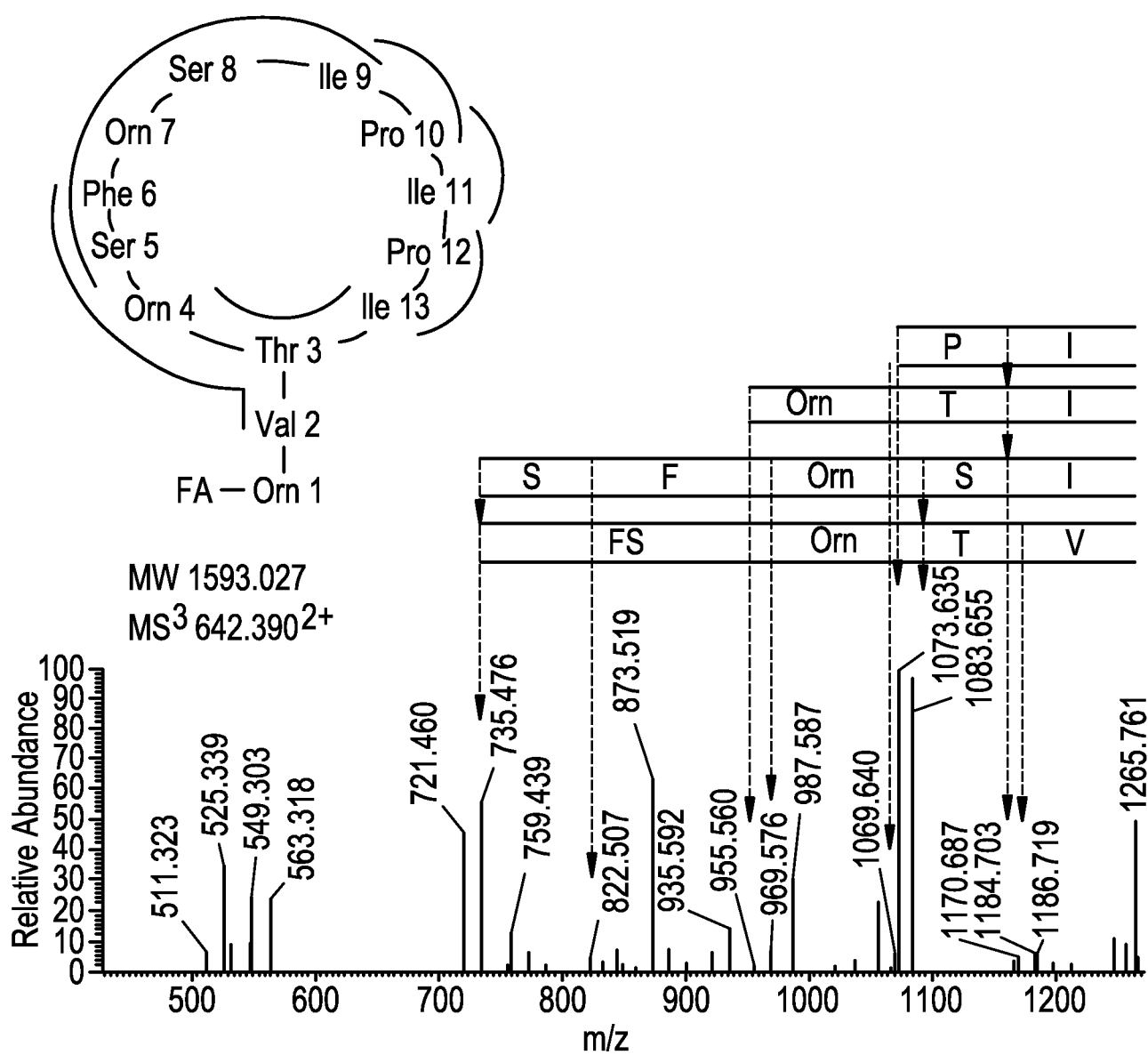
Compound of the present invention SEQ ID NO. 4

Compounds Identified in this Work



8/48

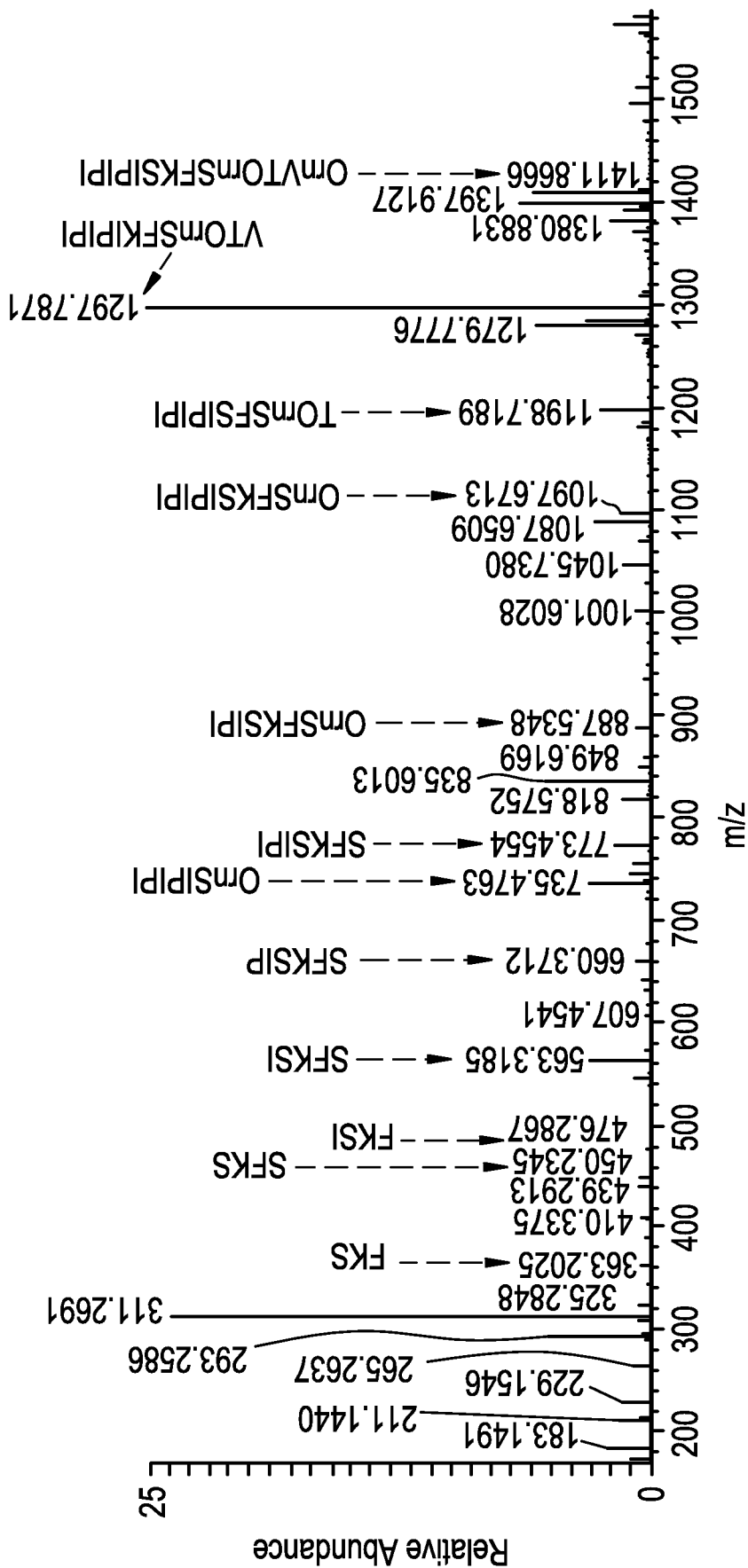
FIG. 9



8/46

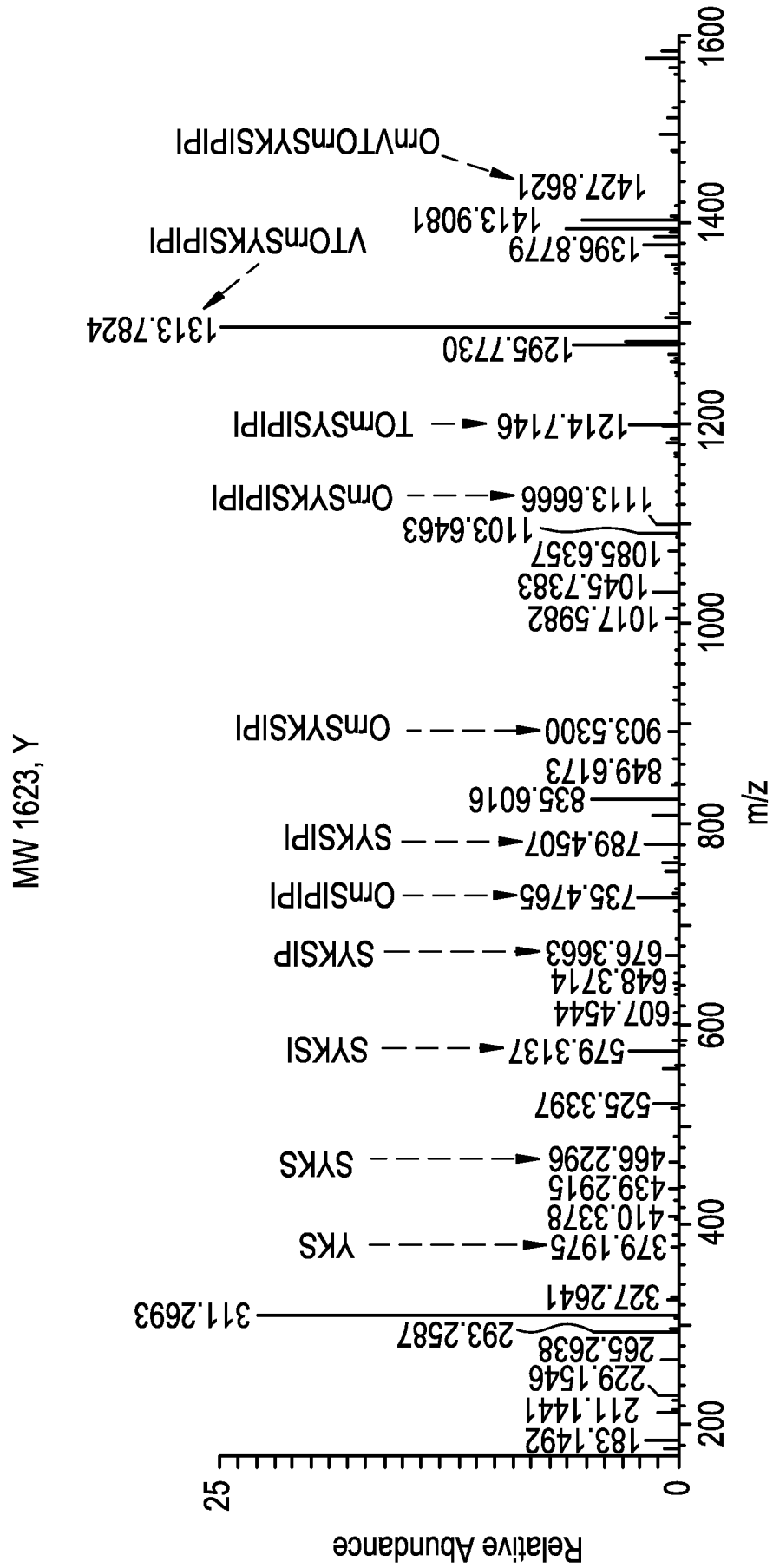
FIG. 10A

MV 1607, F



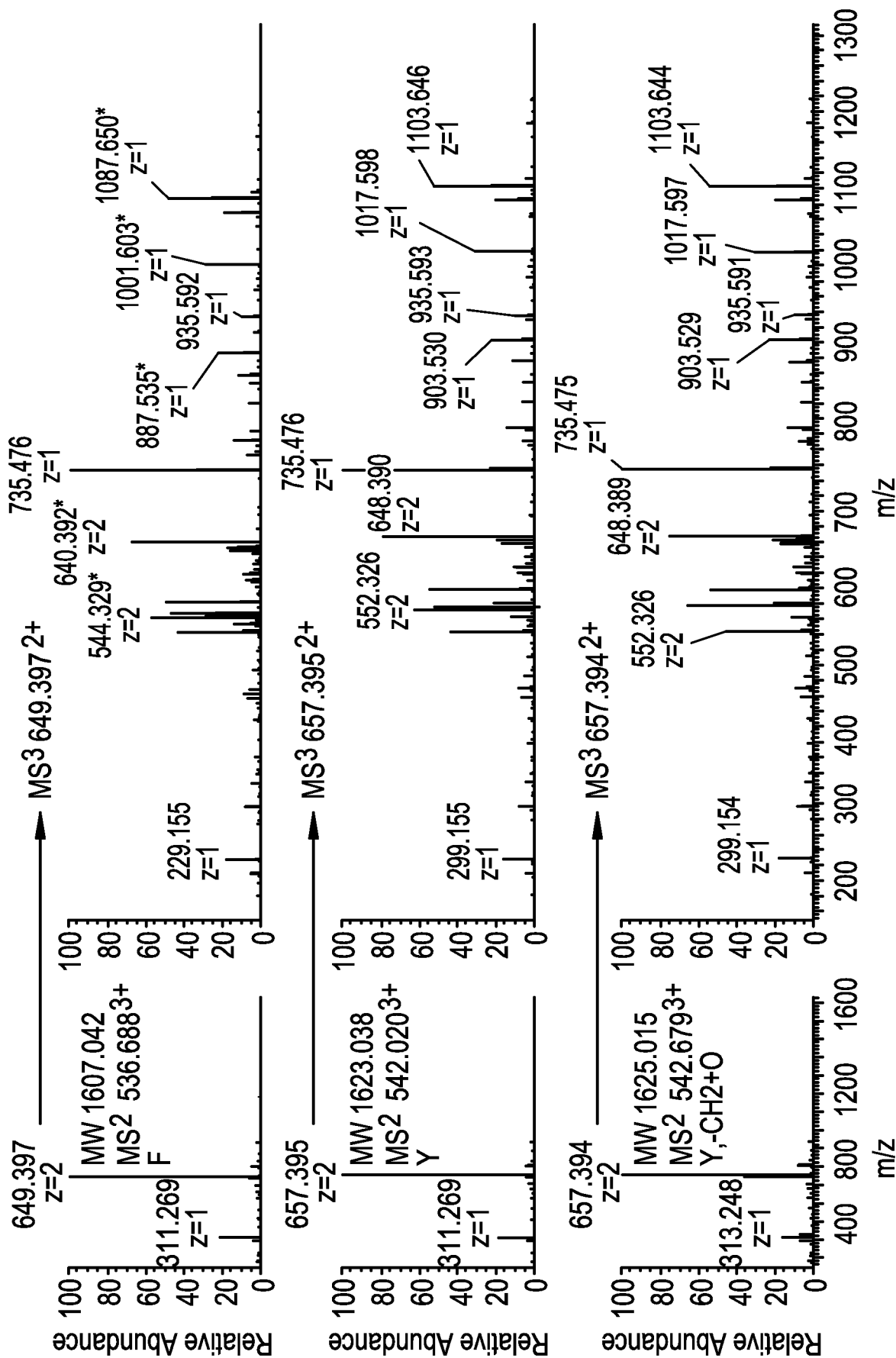
10/48

FIG. 10B



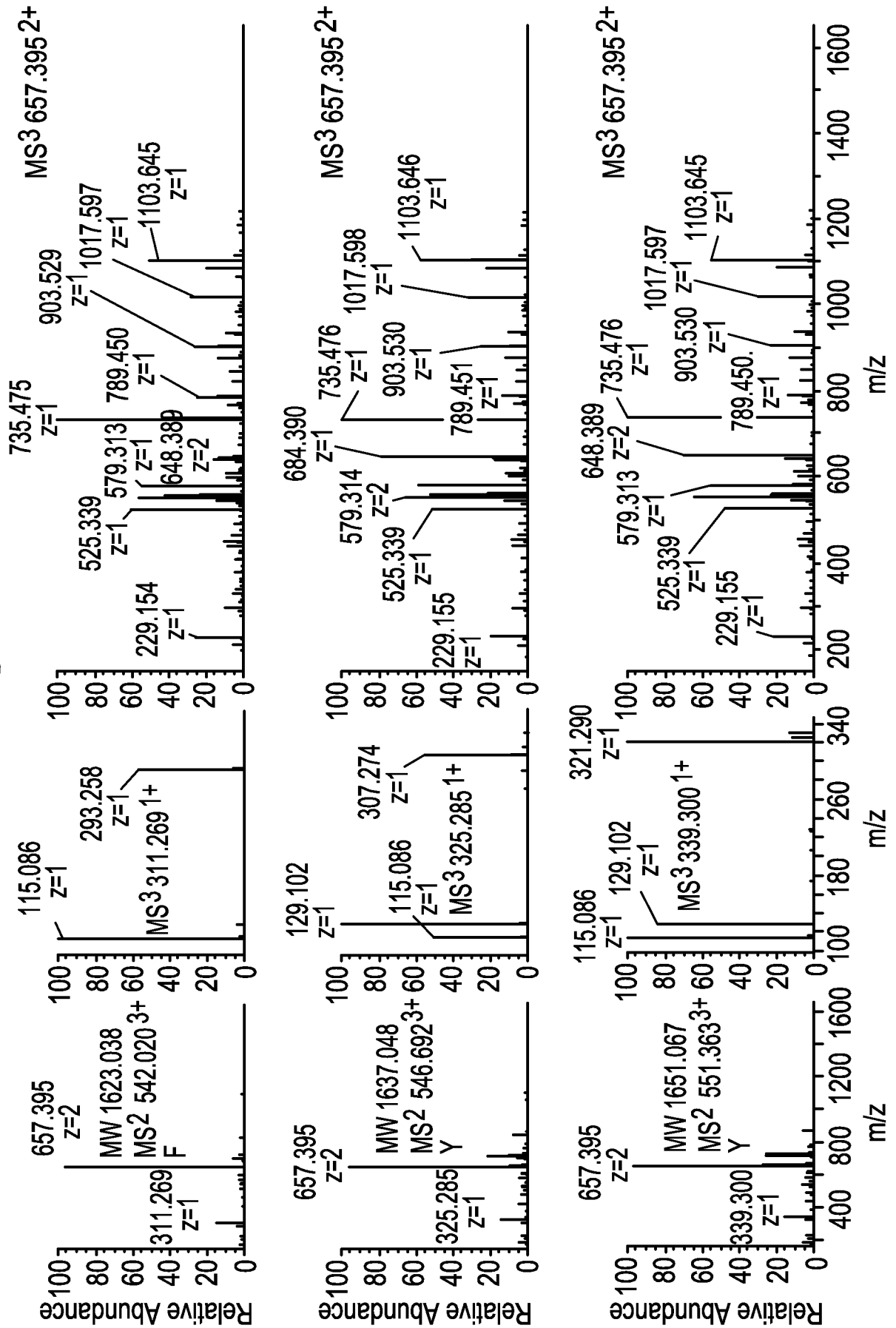
11/48

FIG. 11



12/48

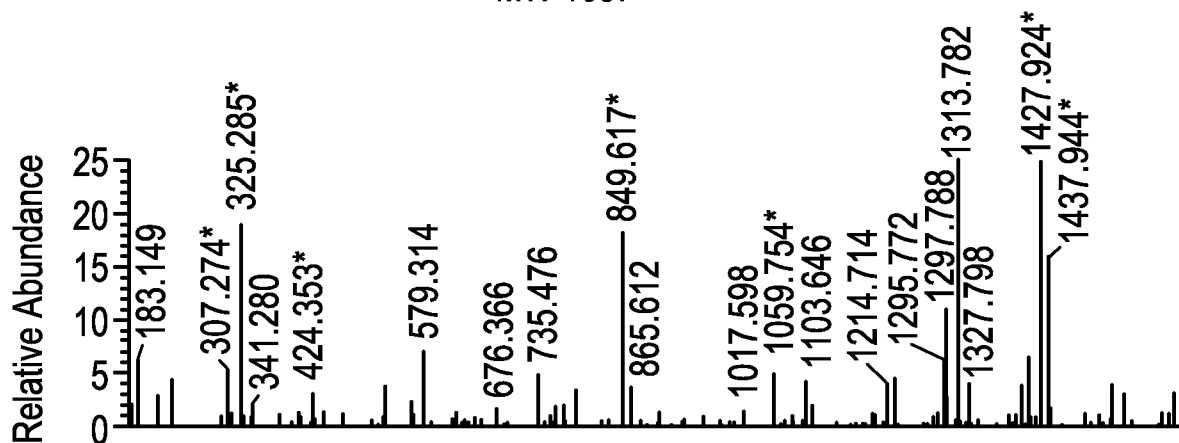
FIG. 12



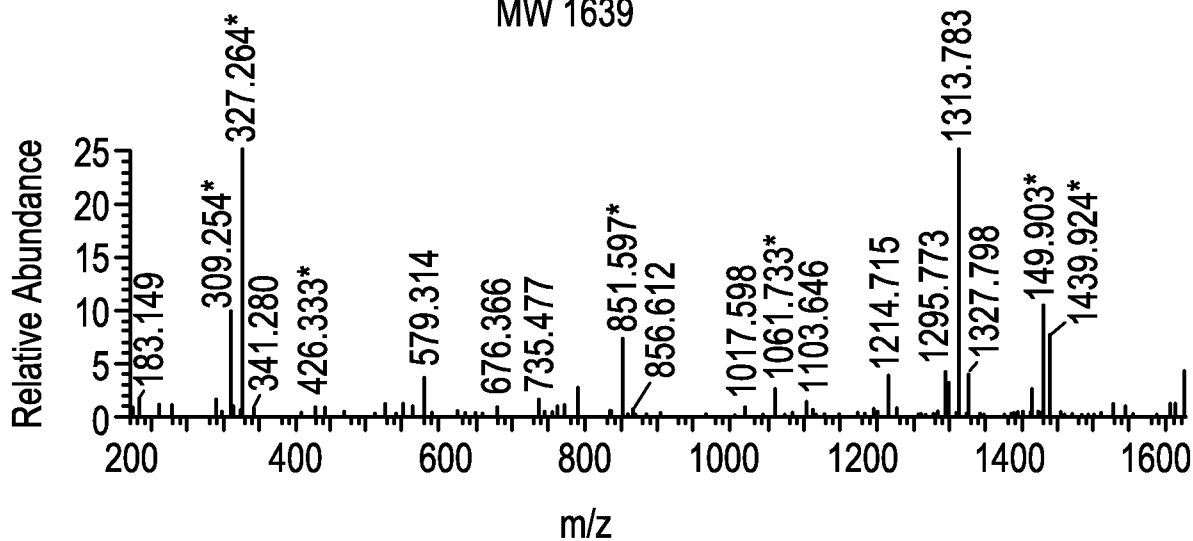
13/48

FIG. 13

MW 1637

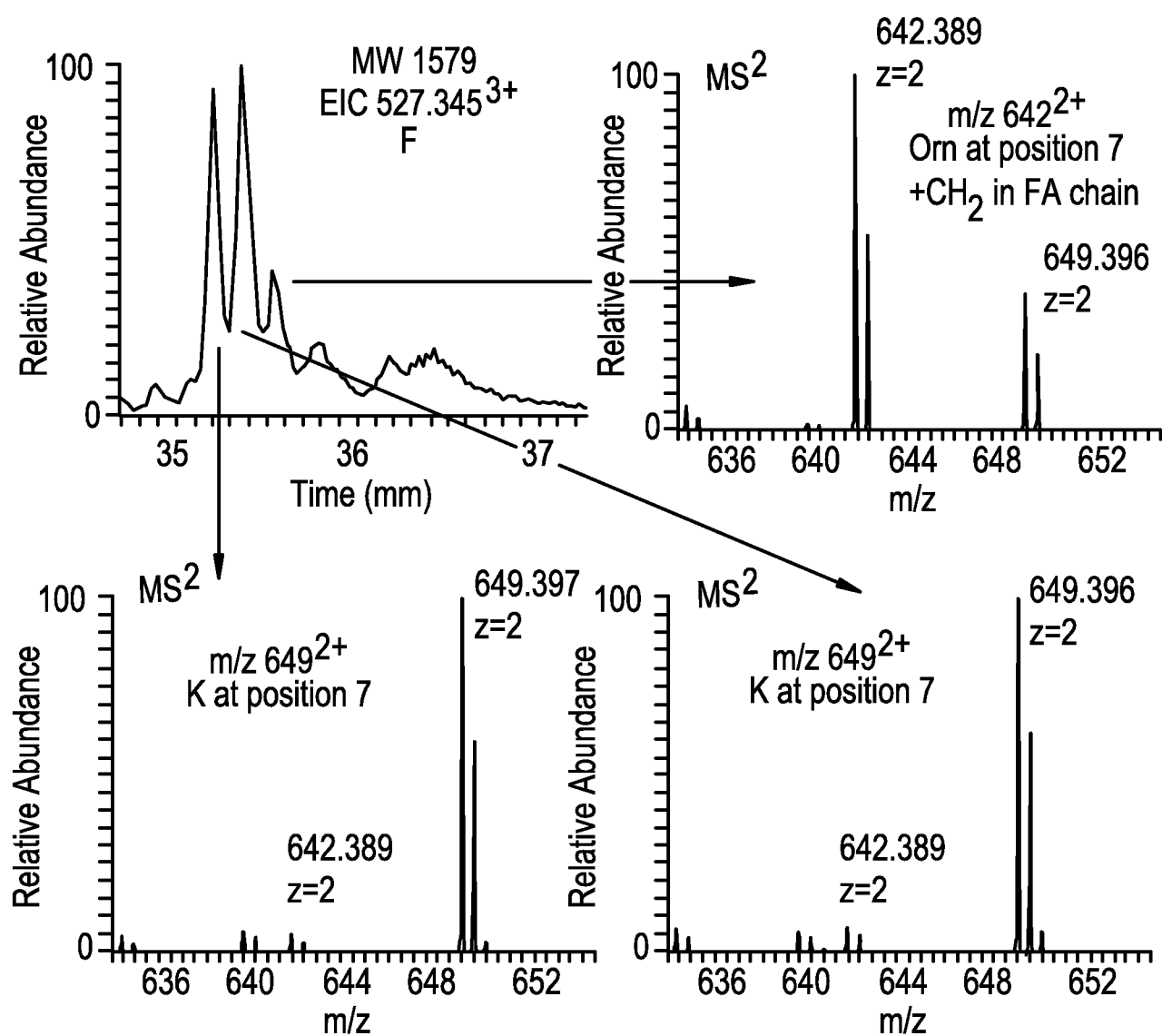


MW 1639



14/48

FIG. 14



15/48

FIG. 15
(SEQ ID NO. 77)

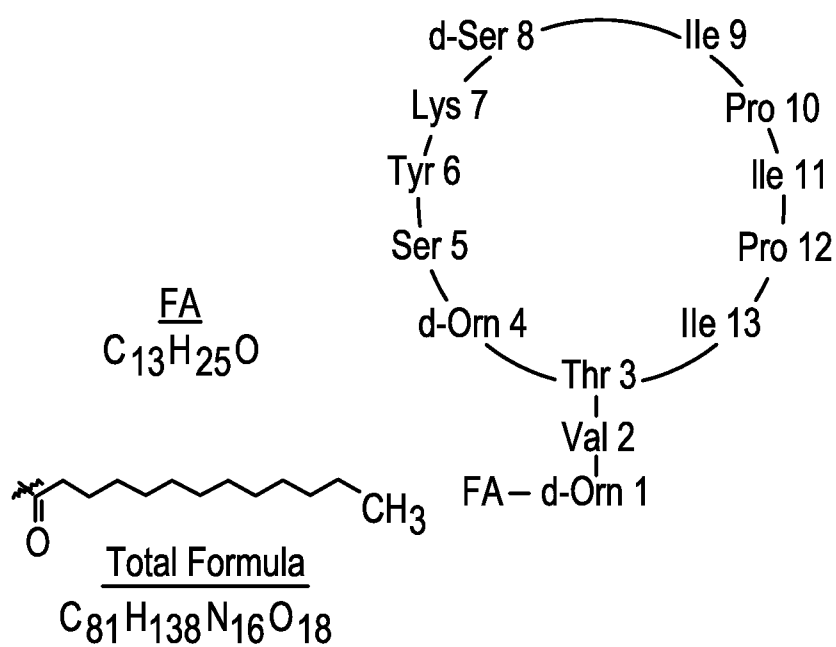


FIG. 16A

A6_61_pbtA	1	MNANPNENWYPLTQAQRRIWYTEMHPNNTSVTTVAGTMYIRGKVDVEILKMAIQVIMQHD
TS_15_pbtA	1	MNANPNENWYPLTQAQRRIWYTEMHPNNTSVTTVAGTMYIRGKVDVEILKMAIQVIMQHD
OSY_SE_pbtA	1	-----MEMNHPNTVSTVACTVTVRGEVDAALCKESICQVITQHD
A6_61_pbtA	61	AFRIHIAITNDNPKQAPVEQIVPHVDYLENDNQIEAESMLQRFNHIPIHMFDPALYHF
TS_15_pbtA	61	AFRIHIAITNDNPKQAPVEQIVPHVDYLENDNQIEAESMLQRFNHIPIHMFDPALYHF
OSY_SE_pbtA	41	AFRIHIAITNDNPKQAPVEQIVPHVDYLENDNQIEAESMLQRFNHIPIHMFDPALYHF
A6_61_pbtA	121	VVPHVNDDEA WFNLEKMHIAITDGVSSHIAIKIMNTYAMVSGNADTDEQESTYLDYIPA
TS_15_pbtA	121	VVPHVNDDEA WFNLEKMHIAITDGVSSHIAIKIMNTYAMVSGNADTDEQESTYLDYIPA
OSY_SE_pbtA	101	TVPHVNDDEA WFNLEKMHIAITDGVSSHIAIKIMNTYAMVSGNADTDEQESTYLDYIPA
A6_61_pbtA	181	EREYEQSDRYAKDKAYWLDKPFSTMPREVIGIKSYPPHSIGTEASRTSTVSGEMYEKLYRF
TS_15_pbtA	181	EREYEQSDRYAKDKAYWLDKPFSTMPREVIGIKSYPPHSIGTEASRTSTVSGEMYEKLYRF
OSY_SE_pbtA	161	EREYEQSDRYAKDKAYWLDKPFSTMPREVIGIKSYPPHSIGTEASRTSTVSGEMYEKLYRF
A6_61_pbtA	241	SOQHNISLFTPLGSLYAFLYKTTGNNDIAVGAAYANRTSKODKDALGMPVSTVAARLT
TS_15_pbtA	241	SOQHNISLFTPLGSLYAFLYKTTGNNDIAVGAAYANRTSKODKDALGMPVSTVAARLT
OSY_SE_pbtA	221	SOQHNISLFTPLGSLYAFLYKTTGNNDIAVGAAYANRTSKODKDALGMPVSTVAARLT
A6_61_pbtA	301	SPQDVLTPLHNVAKQKAILRHQKYPYNQLILDLREQNNNSVEIQDLYRISIDYMPIRWS
TS_15_pbtA	301	SPQDVLTPLHNVAKQKAILRHQKYPYNQLILDLREQNNNSVEIQDLYRISIDYMPIRWS
OSY_SE_pbtA	281	SPQDVLTPLHNVAKQKAILRHQKYPYNQLILDLREQNNNSVEIQDLYRISIDYMPIRWS
A6_61_pbtA	361	SYGELAAQRSSFCGHEVDDFAVHVEDMVDNDQIIFNIDYRKQLFSEHEVIRIIEQMNTI
TS_15_pbtA	361	SYGELAAQRSSFCGHEVDDFAVHVEDMVDNDQIIFNIDYRKQLFSEHEVIRIIEQMNTI
OSY_SE_pbtA	341	SYGELAAQRSSFCGHEVDDFAVHVEDMVDNDQIIFNIDYRKQLFSEHEVIRIIEQMNTI
A6_61_pbtA	421	VDQMISNPSC LQQLSMISDKEAQIILTRFSGNWNSTPQPVGRTIHQLFEEQVETPDQV
TS_15_pbtA	421	VDQMISNPSC LQQLSMISDKEAQIILTRFSGNWNSTPQPVGRTIHQLFEEQVETPDQV
OSY_SE_pbtA	401	VDQMISNPSC LQQLSMISDKEAQIILTRFSGNWNSTPQPVGRTIHQLFEEQVETPDQV
A6_61_pbtA	481	AVVFCGRHLYKELNEQANCPCARTLRAHGVAAEQFVCGIMADRSIEMVVGILAILKAGGAY
TS_15_pbtA	481	AVVFCGRHLYKELNEQANCPCARTLRAHGVAAEQFVCGIMADRSIEMVVGILAILKAGGAY
OSY_SE_pbtA	459	AVVFCGRHLYKELNEQANCPCARTLRAHGVAAEQFVCGIMADRSIEMVVGILAILKAGGAY

17/48

FIG. 16B

A6 61 PbtA	541	VPIDPEYPERILYMLSDSNARVLVSQSHLQTRIGYTGTVVLLDSNDYEANRDNLESVN
TS 15 PbtA	541	VPIDPEYPERILYMLSDSNARVLVSQSHLQTRIGYTGTVVLLDSNDYEANRDNLESVN
OSY SE PbtA	519	VPIDPEYPERILYMLSDSNARVLVSQSHLQTRIGYTGTVVLLDSNDYEANRDNLESVN
A6 61 PbtA	601	EAHHLAVVIYTSCTTKPKPGVMIEHKQITALGDAMKHAYQLDDESGIETLOWASFSFDVFT
TS 15 PbtA	601	EAHHLAVVIYTSCTTKPKPGVMIEHKQITALGDAMKHAYQLDDESGIETLOWASFSFDVFT
OSY SE PbtA	579	SPSNLSYVIYTSCTTKPKPGVMIEHKQITALGDAMKHAYQLDDESGIETLOWASFSFDVFT
A6 61 PbtA	661	GMVVRALLYGCGELIICPSEARANPEAICELIARHRIHIFESTPALVIPLMHYVHEQCKDV
TS 15 PbtA	661	GMVVRALLYGCGELIICPSEARANPEAICELIARHRIHIFESTPALVIPLMHYVHEQCKDV
OSY SE PbtA	639	GMVVRALLYGCGELIICPSEARANPEAICELIARHRIHIFESTPALVIPLMHYVHEQCKDV
A6 61 PbtA	721	SSLRLLVVGSDDHCPAAEYRKLKMERFGSQMRILNSYGVTEACVDACIYENCSVDLSITMLP
TS 15 PbtA	721	SSLRLLVVGSDDHCPAAEYRKLKMERFGSQMRILNSYGVTEACVDACIYENCSVDLSITMLP
OSY SE PbtA	699	SSLRLLVVGSDDHCPAAEYRKLKMERFGSQMRILNSYGVTEACVDACIYENCSVDLSITMLP
A6 61 PbtA	781	IGKPLPSVSMYILDENKALQPIGIVGELYIGGAGVGRGYLNRDDLTAEKFVDDPYSOCK
TS 15 PbtA	781	IGKPLPSVSMYILDENKALQPIGIVGELYIGGAGVGRGYLNRDDLTAEKFVDDPYSOCK
OSY SE PbtA	758	IGKPLPSVSMYILDENKALQPIGIVGELYIGGAGVGRGYLNRDDLTAEKFVDDPYSOCK
A6 61 PbtA	840	MYRTGDLARWLPGDNIEYLGRLDQVKIRGNRIEIGIEIETRMLOQTSLVREAVIVAREDEN
TS 15 PbtA	840	MYRTGDLARWLPGDNIEYLGRLDQVKIRGNRIEIGIEIETRMLOQTSLVREAVIVAREDEN
OSY SE PbtA	818	MYRTGDLARWLPGDNIEYLGRLDQVKIRGNRIEIGIEIETRMLOQTSLVREAVIVAREDEN
A6 61 PbtA	900	GLKALCAIIVADSEISVQOQLRSTLAEOVPDYMIPSPYFKLERPLTPNGKIDRNGLPAPS
TS 15 PbtA	900	GLKALCAIIVADSEISVQOQLRSTLAEOVPDYMIPSPYFKLERPLTPNGKIDRNGLPAPS
OSY SE PbtA	878	GLKALCAIIVADSEISVQOQLRSTLAEOVPDYMIPSPYFKLERPLTPNGKIDRNGLPAPS
A6 61 PbtA	960	GQDYSCKIYVEPRNQAEQTLASINWMLVGVKRVGILDHFFELGCGDSIKSIQVSSRMQQAG
TS 15 PbtA	960	GQDYSCKIYVEPRNQAEQTLASINWMLVGVKRVGILDHFFELGCGDSIKSIQVSSRMQQAG
OSY SE PbtA	938	GQDYSCKIYVEPRNQAEQTLASINWMLVGVKRVGILDHFFELGCGDSIKSIQVSSRMQQAG
A6 61 PbtA	1020	YKLDIRDLFKYPTIEQISPHLVEVQKAEQCEESGEVGLTPILRWYFRRDEVSLSHHYNQS
TS 15 PbtA	1020	YKLDIRDLFKYPTIEQISPHLVEVQKAEQCEESGEVGLTPILRWYFRRDEVSLSHHYNQS
OSY SE PbtA	998	YKLDIRDLFKYPTIEQISPHLVEVQKAEQCEESGEVGLTPILRWYFRRDEVSLSHHYNQS

FIG. 16C

A6 61	Pbta	1080	IMLHRKAGFDEAALRNALHKITTEHHDALRNVPFRTEQGEYAAWNRRIEKGELYRLDVLDDI
TS 15	Pbta	1080	IMLHRKAGFDEAALRNALHKITTEHHDALRNVPFRTEQGEYAAWNRRIEKGELYRLDVLDDI
OSY SE	Pbta	1088	IMLHRKAGFDEAALRNALHKITTEHHDALRNVPFRTEQGEYAAWNRRIEKGELYRLDVLDDI
A6 61	Pbta	1140	KERSACVSESESLHNHLAAZANVIGCFDIEACPLVGCAGLFRCPDGDHLLIVIHHAVIDA
TS 15	Pbta	1140	KERSACVSESESLHNHLAAZANVIGCFDIEACPLVGCAGLFRCPDGDHLLIVIHHAVIDA
OSY SE	Pbta	1117	KERSACVSESESLHNHLAAZANVIGCFDIEACPLVGCAGLFRCPDGDHLLIVIHHAVIDA
A6 61	Pbta	1200	VSMRILLLEDLATGYEQALCGSEIRLPDSTDSFRLWSELSATAQQSNNEELKYMEQVAAQ
TS 15	Pbta	1200	VSMRILLLEDLATGYEQALCGSEIRLPDSTDSFRLWSELSATAQQSNNEELKYMEQVAAQ
OSY SE	Pbta	1170	VSMRILLLEDLATGYEQALCGSEIRLPDSTDSFRLWSELSATAQQSNNEELKYMEQVAAQ
A6 61	Pbta	1260	TTITPLPTDY-AGIALQDSESVTVVWSASETELLKQAHRAINTOMDDLLLTALGIAFR
TS 15	Pbta	1260	TTITPLPTDY-AGIALQDSESVTVVWSASETELLKQAHRAINTOMDDLLLTALGIAFR
OSY SE	Pbta	1230	TTITPLPTDY-AGIALQDSESVTVVWSASETELLKQAHRAINTOMDDLLLTALGIAFR
A6 61	Pbta	1319	RWCGHERIRINLECHGRESILPDLDITRTVGMFTSEYFQQLLEVGSSEELPRIIKSVKEDL
TS 15	Pbta	1319	RWCGHERIRINLECHGRESILPDLDITRTVGMFTSEYFQQLLEVGSSEELPRIIKSVKEDL
OSY SE	Pbta	1290	RWCGHERIRINLECHGRESILPDLDITRTVGMFTSEYFQQLLEVGSSEELPRIIKSVKEDL
A6 61	Pbta	1379	RSIPNKGIGYGICRYLSNTGICENVNGTAPESVFNILGQDDQDFQNSGSPSPYSTGSGNIG
TS 15	Pbta	1379	RSIPNKGIGYGICRYLSNTGICENVNGTAPESVFNILGQDDQDFQNSGSPSPYSTGSGNIG
OSY SE	Pbta	1350	RSIPNKGIGYGICRYLSNTGICENVNGTAPESVFNILGQDDQDFQNSGSPSPYSTGSGNIG
A6 61	Pbta	1439	GDQGRPYLLDMNGMVSQCKLQDLISYGRTOYRAETIERLASLIRDSLELIIHCVAKERTI
TS 15	Pbta	1439	GDQGRPYLLDMNGMVSQCKLQDLISYGRTOYRAETIERLASLIRDSLELIIHCVAKERTI
OSY SE	Pbta	1410	GDQGRPYLLDMNGMVSQCKLQDLISYGRTOYRAETIERLASLIRDSLELIIHCVAKERTI
A6 61	Pbta	1499	ELTPSDVSLQRIISIQELEQIVERTSGICEVEDIYALTPMOKGMWFFHTAMDSQAGAYFELT
TS 15	Pbta	1499	ELTPSDVSLQRIISIQELEQIVERTSGICEVEDIYALTPMOKGMWFFHTAMDSQAGAYFELT
OSY SE	Pbta	1470	ELTPSDVSLQRIISIQELEQIVERTSGICEVEDIYALTPMOKGMWFFHTAMDSQAGAYFELT
A6 61	Pbta	1559	RITLLEGTNIEAPAAASNNELAAARHAFRTNPLVDSNGCEPLQVVFESKRI SVKHEDLRSIA
TS 15	Pbta	1559	RITLLEGTNIEAPAAASNNELAAARHAFRTNPLVDSNGCEPLQVVFESKRI SVKHEDLRSIA
OSY SE	Pbta	1530	RITLLEGTNIEAPAAASNNELAAARHAFRTNPLVDSNGCEPLQVVFESKRI SVKHEDLRSIA

FIG.16D

A6_61_pbtA	1619	AMEQAVAIENEAAKER	EQGPDLENG	VMRVSVLQ	TADENV	EVVWIS	HHIVMD	GWCLPLVA
TS_15_pbtA	1619	SVRQAVAIENEAAKER	EQGPDLENG	VMRVSVLQ	TADENV	EVVWIS	HHIVMD	GWCLPLVA
OSY_SE_pbtA	1590	PAKQRAH	IAENAVNEEK	LEGPDLEQ	DEMRVLT	ERTAC	ES	HNVLWSSHHIEMD
								GWCLPLVA
A6_61_pbtA	1679	AEVFN	TYTSALVEDK	KPILASVPS	YNH	YIQWLER	ODES	AAAAAYNN
TS_15_pbtA	1679	AEVFN	TYTSALVEDK	KPILASVPS	YNH	YIQWLER	ODES	AAAAAYNN
OSY_SE_pbtA	1650	KEVFD	TYSAVRH	EHLEKT	VP	YSQY	YIEMLE	CODE
								EAASAYMSEYIAC
								QOHTALPQCK
A6_61_pbtA	1739	GRHSG	QIIEAG	QVQID	LGTSLS	ALNQVAT	QH	VTINTLLQAS
TS_15_pbtA	1739	GRHSG	QIIEAG	QVQID	LGTSLS	ALNQVAT	QH	VTINTLLQAS
OSY_SE_pbtA	1710	DOGE	SEANAE	HEHED	CELCKD	LS	HE	NEAAKRN
								VTISTLLCTMGI
								LOKYNCTG
								DIIVFG
A6_61_pbtA	1799	SVVSG	RP	AE	LVG	IE	EM	IGLF
TS_15_pbtA	1799	SVVSG	RP	AE	LVG	IE	EM	IGLF
OSY_SE_pbtA	1770	GVVSG	RP	AE	LVG	IE	EM	IGLF
								INTIPRV
								VSSQA
								ERF
								EVMT
								NO
								ED
								AL
								SS
								AK
								HD
								I
								P
								LY
A6_61_pbtA	1859	IQAO	CTLK	CM	LI	TH	IM	VLEN
TS_15_pbtA	1859	IQAO	CTLK	CM	LI	TH	IM	VLEN
OSY_SE_pbtA	1830	IQAO	SA	QK	CE	LI	NH	IM
								VFN
								SE
								Q
								LES
								DC
								CK
								KL
								ED
								V
								MT
								S
								QT
								NY
								DN
								LI
								IP
A6_61_pbtA	1919	DN	IV	IR	PD	EN	KQ	AL
TS_15_pbtA	1919	DN	IV	IR	PD	EN	KQ	AL
OSY_SE_pbtA	1890	DE	IV	IR	PD	EN	KQ	AL
								ES
								AD
								DM
								N
								VL
								EQ
								AS
								NP
								RI
								SI
								GE
								LO
								Q
								AT
								DE
								ER
								AV
								MM
								SE
								PN
A6_61_pbtA	1979	DT	FV	AY	PRE	K	SI	HR
TS_15_pbtA	1979	DT	FV	AY	PRE	K	SI	HR
OSY_SE_pbtA	1950	NT	V	EY	PRE	K	SI	HR
								EE
								SI
								PD
								AV
								FG
								CM
								ONT
								IG
								AL
								NA
								AN
								RR
								MA
								RL
								ER
								AG
								VT
								MG
A6_61_pbtA	2039	EL	VG	IC	ADR	SL	EN	VV
TS_15_pbtA	2039	EL	VG	IC	ADR	SL	EN	VV
OSY_SE_pbtA	2010	TH	VG	IC	ADR	SL	EN	VV
								GL
								LI
								CI
								IF
								AG
								AY
								VP
								ID
								ES
								NI
								PI
								DG
								AG
								DD
								LA
								VI
								IT
								YS
								SG
								ST
								GT
								PK
								GC
								VC
								VT
								HR
A6_61_pbtA	2099	NP	EH	FN	VI	LD	-	VN
TS_15_pbtA	2099	NP	EH	FN	VI	LD	-	VN
OSY_SE_pbtA	2070	VP	NS	IS	DS	VS	LD	AA
								BE
								TI
								EG
								SW
								PN
								ET
								ID

								TA
								AT
								AD
								DD
								V
								AY
								II
								YT
								SG
								ST
								GT
								PK
								GC
								VC
								VT
								HR

20/48

FIG. 16E

A6 61	Pbta	2158	GVVRLVCAATYVEIDSSDVFLQGSTISPDAAATPEIWSLLNGAALAILPSCNVSLTDWSE
TS 15	Pbta	2157	GVVRLVCAATYVEIDSSDVFLQGSTISPDAAATPEIWSLLNGAALAILPSCNVSLTDWSE
OSY SE	Pbta	2126	GVVRLVADATYVEHSKDVFLQGSTISPDAAATPEIWSLLNGAALAILPSCNVSLTDWSE
A6 61	Pbta	2218	AIQRHRVTTLWMTAGLFQVAVVEQOIEGEYGVKQLLVGSDVVSPTHVKKVMEKHNGIRVIN
TS 15	Pbta	2217	AIQRHRVTTLWMTAGLFQVAVVEQOIEGEYGVKQLLVGSDVVSPTHVKKVMEKHNGIRVIN
OSY SE	Pbta	2186	AIQCHQVTHLWMTAGLFHVAVVEQOIALQGVQQLLVGSDVVSPTATKVSEYNGIRGIN
A6 61	Pbta	2278	GYGPTENTTFTCCHTITAADLDRGCSIPIGRPISNTRVYVLDDEAGNALPVGVCGLYAGG
TS 15	Pbta	2277	GYGPTENTTFTCCHTITAADLDRGCSIPIGRPISNTRVYVLDDEAGNALPVGVCGLYAGG
OSY SE	Pbta	2246	GYGPTENTTFTCCHEISAADETFPSIPIGRPISNTRVYVLDDEAGNALPVGVCGLYAGG
A6 61	Pbta	2338	DGLARGYLNRPELTAEKFFVNDPFIPGERLYRTGDLARNLPDGSIEFIGRCDSEQVKIRGYR
TS 15	Pbta	2336	DGLARGYLNRPELTAEKFFVNDPFIPGERLYRTGDLARNLPDGSIEFIGRCDSEQVKIRGYR
OSY SE	Pbta	2305	DGLAQGYLNRPELTAEKFFVDEPIVPAIRLYRTGDLARNLPDGSIEFIGRCDSEQVKIRGYR
A6 61	Pbta	2398	IEPGEVLAAYLLRIDEVGEAAVIAREDSSGQKELCAVPTTETELASASGLKETLARELPAYM
TS 15	Pbta	2396	IEPGEVLAAYLLRIDEVGEAAVIAREDSSGQKELCAVPTTETELASASGLKETLARELPAYM
OSY SE	Pbta	2365	IELGEVBAHLLKVEFVQSAAVIAKDESGQNMLCAVYANKKELTASELSALSQELPBYM
A6 61	Pbta	2458	IPSHFIQIEBELPLTPNGKVDRRALPQFCEGMHINICIQPRTTELEAKLALIMQDVLGLEN
TS 15	Pbta	2456	IPSHFIQIEBELPLTPNGKVDRRALPQFCEGMHINICIQPRTTELEAKLALIMQDVLGLEN
OSY SE	Pbta	2425	IPSHFIQIEBELPLTPNGKVDRRALPQFCEGMHINICIQPRTTELEAKLALIMQDVLGLEN
A6 61	Pbta	2517	VGVTDSFFELGCHSLRATTLVSKVHRELSEVVLPLQDVFRYPTIEQMSLAIQGMQKESFAS
TS 15	Pbta	2515	VGVTDSFFELGCHSLRATTLVSKVHRELSEVVLPLQDVFRYPTIEQMSLAIQGMQKESFAS
OSY SE	Pbta	2485	VSNKDSFFELGCHSLRATTLVSKVHRELSEVVLPLQDVFRYPTIEQMSLAIQGMQKESFAS
A6 61	Pbta	2577	IPRVEDREWPVSSAQKRLFVLHQMEGAELSYNNMPGVMAIEGKLRDRLEAAFEELIARR
TS 15	Pbta	2575	IPRVEDREWPVSSAQKRLFVLHQMEGAELSYNNMPGVMAIEGKLRDRLEAAFEELIARR
OSY SE	Pbta	2545	IPRVEDREWPVSSAQKRLFVLHQMEGAELSYNNMPGVMAIEGKLRDRLEAAFEELIARR
A6 61	Pbta	2637	EVLRTCFEMHNGCEPMQRIYSDEFTVEHIVGAASEAESVIRSFVRAFQLNKPPLLRVGL
TS 15	Pbta	2635	EVLRTCFEMHNGCEPMQRIYSDEFTVEHIVGAASEAESVIRSFVRAFQLNKPPLLRVGL
OSY SE	Pbta	2605	ETLRTCFEMHNGCEPMQRIYSDEFTVEHIVGAASEAESVIRSFVRAFQLNKPPLLRVGL

21/48

FIG. 16F

A6_61_Pbta	2697	IEVDAGRHLLEFDMHHIISDGCASMGILLDEPVALYSCEQLPELRLOQKDYASWQSESAVL
TS_15_Pbta	2695	IEVDAGRHLLEFDMHHIISDGCASMGILLDEPVALYSCEQLPELRLOQKDYASWQSESAVL
OSY_SE_Pbta	2664	IEVDAGRHLLEFDMHHIISDGCASMGILLDEPVALYSCEQLPELRLOQKDYASWQSESAVL
A6_61_Pbta	2757	SMEEQKAYWLETLRGELPVLQLPVDYTRPAPRSFAGSTLEPIVPADKTDOLKQLGAGSD
TS_15_Pbta	2755	SMEEQKAYWLETLRGELPVLQLPVDYTRPAPRSFAGSTLEPIVPADKTDOLKQLGAGSD
OSY_SE_Pbta	2724	DWSKQCEAYWLETLRGELPVLQLPVDYTRPAPRSFAGSTLEPIVPADKTDOLKQLGAGSD
A6_61_Pbta	2817	ATMYMVLALYTTALLHKYTGQEDVIVGMPIACRTHADIEPLIGMFEVNTLPLRHYPAGEKT
TS_15_Pbta	2815	ATMYMVLALYTTALLHKYTGQEDVIVGMPIACRTHADIEPLIGMFEVNTLPLRHYPAGEKT
OSY_SE_Pbta	2784	ATMYMVLALYTTALLHKYTGQEDVIVGMPIACRTHADIEPLIGMFEVNTLPLRHYPAGEKT
A6_61_Pbta	2877	FRSPLGEVRQSTLQAYEHQYEPPEELVDHIOPTRDVSRNPIDTVLVLQNTKGAWSIDG
TS_15_Pbta	2875	FRSPLGEVRQSTLQAYEHQYEPPEELVDHIOPTRDVSRNPIDTVLVLQNTKGAWSIDG
OSY_SE_Pbta	2844	FRSPLGEVRQSTLQAYEHQYEPPEELVDHIOPTRDVSRNPIDTVLVLQNTKGAWSIDG
A6_61_Pbta	2937	LAVTPNPIEHAFAKFDLTLHVEEEDGLACSIETATYALYHRETIERLACHFNQLEAVIS
TS_15_Pbta	2935	LAVTPNPIEHAFAKFDLTLHVEEEDGLACSIETATYALYHRETIERLACHFNQLEAVIS
OSY_SE_Pbta	2904	LAVTPNPIEHAFAKFDLTLHVEEEDGLACSIETATYALYHRETIERLACHFNQLEAVIS
A6_61_Pbta	2997	NPSARLQOLGIIITETEKQOLPECFNDTSADYPRDKTIHRLPEEQVETPDIAVTCGDGF
TS_15_Pbta	2995	NPSARLQOLGIIITETEKQOLPECFNDTSADYPRDKTIHRLPEEQVETPDIAVTCGDGF
OSY_SE_Pbta	2964	NPSARLQOLGIIITETEKQOLPECFNDTSADYPRDKTIHRLPEEQVETPDIAVTCGDGF
A6_61_Pbta	3057	LTYYQELNERANSLAWVLRAEGICADKLVGIMAEERTDMLVGLIAILKAGGAYVPIDPEYP
TS_15_Pbta	3055	LTYYQELNERANSLAWVLRAEGICADKLVGIMAEERTDMLVGLIAILKAGGAYVPIDPEYP
OSY_SE_Pbta	3024	LTYYQELNERANSLAWVLRAEGICADKLVGIMAEERTDMLVGLIAILKAGGAYVPIDPEYP
A6_61_Pbta	3117	DEERISYMLSDSGADILLPRHLEKQVAYEGTVLFLDDDEQTYSGDKSNPSSVNKPSPDLAYV
TS_15_Pbta	3115	DEERISYMLSDSGADILLPRHLEKQVAYEGTVLFLDDDEQTYSGDKSNPSSVNKPSPDLAYV
OSY_SE_Pbta	3084	DEERISYMLSDSGADILLPRHLEKQVAYEGTVLFLDDDEQTYSGDKSNPSSVNKPSPDLAYV
A6_61_Pbta	3177	IYTSGETTGKPKGTLIEHKNNVRLLFNSRNLFDPFRSTDTWTLPHSFCDFPSVWEMYGALLY
TS_15_Pbta	3175	IYTSGETTGKPKGTLIEHKNNVRLLFNSRNLFDPFRSTDTWTLPHSFCDFPSVWEMYGALLY
OSY_SE_Pbta	3144	IYTSGETTGKPKGTLIEHKNNVRLLFNSRNLFDPFRSTDTWTLPHSFCDFPSVWEMYGALLY

22/48

FIG. 16G

A6_61_Pbta	3237	GG	LVVVPQ	TAKNPAM	FLQLLA	EEERV	TI	LNQ	TP	Y	F	Q	L	I	R	E	A	L	L	D	G	S	P	E	L	N	R	N	V	I	F	G																											
TS_15_Pbta	3235	GG	LVVVPQ	TAKNPAM	FLQLLA	EEERV	TI	LNQ	TP	Y	F	Q	L	I	R	E	A	L	L	D	G	S	P	E	L	N	R	N	V	I	F	G																											
OSY_SE_Pbta	3204	GG	LVVVPQ	TAKNPAM	FLQLLA	EEERV	TI	LNQ	TP	Y	F	Q	L	I	R	E	A	L	L	D	G	S	P	E	L	N	R	N	V	I	F	G																											
A6_61_Pbta	3297	GE	ALSPQ	LKDWR	AKY	P	R	T	O	L	I	N	M	Y	G	I	T	E	T	T	V	H	V	T	K	E	I	T	E	A	E	I	E	O	A	R	S	N	I	G	F	P	I	P	T	L	R												
TS_15_Pbta	3295	GE	ALSPQ	LKDWR	AKY	P	R	T	O	L	I	N	M	Y	G	I	T	E	T	T	V	H	V	T	K	E	I	T	E	A	E	I	E	O	A	R	S	N	I	G	F	P	I	P	T	L	R												
OSY_SE_Pbta	3264	GE	ALSPQ	LKDWR	AKY	P	R	T	O	L	I	N	M	Y	G	I	T	E	T	T	V	H	V	T	K	E	I	T	E	A	E	I	E	O	A	R	S	N	I	G	F	P	I	P	T	L	R												
A6_61_Pbta	3357	Y	I	L	D	A	N	R	O	C	V	P	I	G	V	A	G	E	M	F	V	A	G	E	L	A	G	Y	L	N	R	P	E	L	T	A	D	R	F	V	D	N	P	F	E	P	G	S	K	M	I	K	T	G	D	L	A	K	
TS_15_Pbta	3355	Y	I	L	D	A	N	R	O	C	V	P	I	G	V	A	G	E	M	F	V	A	G	E	L	A	G	Y	L	N	R	P	E	L	T	A	D	R	F	V	D	N	P	F	E	P	G	S	K	M	I	K	T	G	D	L	A	K	
OSY_SE_Pbta	3324	Y	I	L	D	A	N	R	O	C	V	P	I	G	V	A	G	E	M	F	V	A	G	E	L	A	G	Y	L	N	R	P	E	L	T	A	D	R	F	V	D	N	P	F	E	P	G	S	K	M	I	K	T	G	D	L	A	K	
A6_61_Pbta	3417	L	P	D	G	N	I	E	V	L	G	R	I	D	H	O	V	K	I	R	G	Y	R	I	E	L	G	E	V	E	A	Q	V	T	K	V	S	V	R	E	A	I	V	I	A	R	E	E	-	N	G	E	K	L	L	C	A	Y	F
TS_15_Pbta	3415	L	P	D	G	N	I	E	V	L	G	R	I	D	H	O	V	K	I	R	G	Y	R	I	E	L	G	E	V	E	A	Q	V	T	K	V	S	V	R	E	A	I	V	I	A	R	E	E	-	N	G	E	K	L	L	C	A	Y	F
OSY_SE_Pbta	3384	L	P	D	G	N	I	E	V	L	G	R	I	D	H	O	V	K	I	R	G	Y	R	I	E	L	G	E	V	E	A	Q	V	T	K	V	S	V	R	E	A	I	V	I	A	R	E	E	-	N	G	E	K	L	L	C	A	Y	F
A6_61_Pbta	3476	A	D	R	O	L	T	V	G	E	M	R	T	E	L	A	O	E	L	P	A	M	I	P	S	F	V	Q	-	E	R	M	P	L	T	S	N	G	K	V	D	R	K	A	L	P	A	P	E	G	S	I	N	T	G	K	E	I	V
TS_15_Pbta	3474	A	D	R	O	L	T	V	G	E	M	R	T	E	L	A	O	E	L	P	A	M	I	P	S	F	V	Q	-	E	R	M	P	L	T	S	N	G	K	V	D	R	K	A	L	P	A	P	E	G	S	I	N	T	G	K	E	I	V
OSY_SE_Pbta	3444	A	A	R	E	L	T	V	G	E	M	R	T	E	L	A	O	E	L	P	A	M	I	P	S	F	V	Q	-	E	R	M	P	L	T	S	N	G	K	V	D	R	K	A	L	P	A	P	E	G	S	I	N	T	G	K	E	I	V
A6_61_Pbta	3536	A	P	R	T	P	M	E	A	S	L	A	R	M	W	E	L	L	G	I	E	O	V	G	V	T	D	N	F	F	E	L	G	H	S	L	R	A	T	A	L	V	N	R	V	H	Q	E	M	N	I	Q	L	P	L	R	D	V	
TS_15_Pbta	3534	A	P	R	T	P	M	E	A	S	L	A	R	M	W	E	L	L	G	I	E	O	V	G	V	T	D	N	F	F	E	L	G	H	S	L	R	A	T	A	L	V	N	R	V	H	Q	E	M	N	I	Q	L	P	L	R	D	V	
OSY_SE_Pbta	3504	A	P	R	T	P	M	E	A	S	L	A	R	M	W	E	L	L	G	I	E	O	V	G	V	T	D	N	F	F	E	L	G	H	S	L	R	A	T	A	L	V	N	R	V	H	Q	E	M	N	I	Q	L	P	L	R	D	V	
A6_61_Pbta	3596	R	F	S	T	I	E	E	L	A	A	A	M	S	E	M	A	E	S	S	I	P	V	A	E	V	Q	D	H	P	V	S	S	A	Q	K	R	L	I	L	H	Q	L	E	G	A	E	Q	G	N	M	P	G	I					
TS_15_Pbta	3594	R	F	S	T	I	E	E	L	A	A	A	M	S	E	M	A	E	S	S	I	P	V	A	E	V	Q	D	H	P	V	S	S	A	Q	K	R	L	I	L	H	Q	L	E	G	A	E	Q	G	N	M	P	G						
OSY_SE_Pbta	3564	R	F	S	T	I	E	E	L	A	A	A	M	S	E	M	A	E	S	S	I	P	V	A	E	V	Q	D	H	P	V	S	S	A	Q	K	R	L	I	L	H	Q	L	E	G	A	E	Q	G	N	M	P	G						
A6_61_Pbta	3656	L	L	E	G	E	L	D	R	S	R	F	E	A	A	P	R	K	L	I	A	H	D	I	L	R	T	G	F	E	L	V	G	E	A	V	Q	R	I	H	D	T	V	D	F	A	I	E	V	R	K	V	E	S	Q	E	V	Q	
TS_15_Pbta	3654	L	L	E	G	E	L	D	R	S	R	F	E	A	A	P	R	K	L	I	A	H	D	I	L	R	T	G	F	E	L	V	G	E	A	V	Q	R	I	H	D	T	V	D	F	A	I	E	V	R	K	V	E	S	Q	E	V	Q	
OSY_SE_Pbta	3624	L	L	E	G	E	L	D	R	S	R	F	E	A	A	P	R	K	L	I	A	H	D	I	L	R	T	G	F	E	L	V	G	E	A	V	Q	R	I	H	D	T	V	D	F	A	I	E	V	R	K	V	E	S	Q	E	V	Q	
A6_61_Pbta	3716	Q	V	Q	F	I	R	T	F	E	L	D	K	P	L	L	R	V	G	L	I	E	I	A	G	T	K	E	Q	S	V	L	L	D	M	H	I	I	S	D	G	V	S	I	G	I	V	L	Q	E	I	M	R	H	V				
TS_15_Pbta	3714	Q	V	Q	F	I	R	T	F	E	L	D	K	P	L	L	R	V	G	L	I	E	I	A	G	T	K	E	Q	S	V	L	L	D	M	H	I	I	S	D	G	V	S	I	G	I	V	L	Q	E	I	M	R	H	V				
OSY_SE_Pbta	3684	V	V	R	D	F	R	T	F	E	L	D	K	P	L	L	R	V	G	L	I	E	I	A	G	T	K	E	Q	S	V	L	L	D	M	H	I	I	S	D	G	V	S	I	G	I	V	L	Q	E	I	M	R	H	V				

FIG. 16H

A6_61_Pbta	3776	EQVPPLHIOYKDYAANQOSEAQKEQOLKHQKAYMLCQFQOGELPILLELPTDYARPAQQYCG
TS_15_Pbta	3774	EEVPPLHIOYKDYAANQOSEAQKEQOLKHQKAYMLDQFQOGELPILLELPTDYARPAQQYDG
OSY_SE_Pbta	3742	AAEPEHIOYKDYAANQOSEAQOQWKEQCEYRQVVEGELPLELPTDYARPAQQYAG
A6_61_Pbta	3836	LZLPPRIDKOVADGLNRIADTGTLLYAVLLAAAYTMLHXYTQGEDIVVGTPIAGRTHHE
TS_15_Pbta	3834	LTLPFRIDKOVADGLNRIADTGTLLYAVLLAAAYTMLHXYTQGEDIVVGTPIAGRTHHE
OSY_SE_Pbta	3802	STLSPRIDPQNSEGLRRIANETGTLLYAVLLAAAYTMLHXYTQGEDIVVGTPIAGRTHG
A6_61_Pbta	3896	LQPLIGMFVNTLAIRAYPEGKAKAFRSYLDDEIRSTMIGAYEHQOQYPPFEELVEGLQLTRDLS
TS_15_Pbta	3894	LQPLIGMFVNTLAIRAYPEGKAKAFRSYLDDEIRSTMIGAYEHQOQYPPFEELVEGLQLTRDLS
OSY_SE_Pbta	3862	LQPLIGMFVNTLAIRNYPAGEKTERSYLAEEZENTLGAZEHOYPPFEELVEGLQYARDLS
A6_61_Pbta	3956	RNPLPDTNPFALDNTDKADSLGSELQMKPYPLEYTIISKPDVSLDVKADEAGLDCSFEYATS
TS_15_Pbta	3954	RNPLPDTNPFALDNTDMVDSLGSELQMKPYPLEYTIISKPDVSLDVKADEAGLDCSFEYATS
OSY_SE_Pbta	3922	RNPLPDTNPFALNTEPETFPLEGLTLPYSEYTIISKPDVSLDVSEKNDLESCSEYATA
A6_61_Pbta	4016	LEKSETIHRMAEHFSEHLKDIVNHPPDAQGLGELMLTVQESDEILQVFNPTSHSLKAEKETI
TS_15_Pbta	4014	LEKSETIHRMAEHFSEHLKDIVNHPPDAQGLGELMLTVHESDEILQVFNPTSHSLKAEKETI
OSY_SE_Pbta	3982	LEKSETAERMAEHFQQLIDSIVQPEAKLVSLGMLTEESKAQIQHVFNRAEAGHSQSEKTV
A6_61_Pbta	4076	HELPEQAERTPEQPAVVFGNERMTYRELNERANKLARTLRAEGVEPDDLLIGVMADRSID
TS_15_Pbta	4074	HELPEQAERTPEQPAVVFGNERMTYRELNERANKLARTLRAEGVEPDDLLIGVMADRSID
OSY_SE_Pbta	4042	PELPEEQVETPPEKAVVHEKQETTYRELNERANKLARTLRAEGVEPEQLCGMADRSID
A6_61_Pbta	4136	MVVAVMAVLKSGGAYVPIDPEYPEDRIRYMLDQAKARILLTQCHLODKVPEGTWVLLED
TS_15_Pbta	4134	MVVAVMAVLKSGGAYVPIDPEYPEDRIRYMLDQAKARILLTQCHLODKVPEGTWVLLED
OSY_SE_Pbta	4102	MVVGNNALKSGGAYVPIDPEYPEDRIRYMLDQCHQVNRQKHQASVAFAGTAVVLDE
A6_61_Pbta	4196	EASVHEDDTNLEPNCEPCHLCYVIYTSGTTGNPKGVMIHRQLAAMAEAWKAEYELHEPG
TS_15_Pbta	4194	EASVHEDDTNLEPNCEPCHLCYVIYTSGTTGNPKGVMIHRQLAAMAEAWKAEYELHEPG
OSY_SE_Pbta	4162	EAPYHEDDTNLEPNNEPMHLSYVIYTSGTTGNPKGVMIHRQLVANAQWKEVYELHEEG
A6_61_Pbta	4256	IRMLQWASPSFDVPSGDLARTLLHGCELILCPSDTRANPGALAEILLRSSGQMFESTPAL
TS_15_Pbta	4254	IRMLQWASPSFDVPSGDLARTLLHGCELILCPSDTRANPGALAEILLRSSGQMFESTPAL
OSY_SE_Pbta	4222	IRMLQWASPSFDVPSGDLVRTLTYGGELILCPSEQARANPANISELIRKHQIQMFESTPAL

24/48

FIG. 16I

A6_61_Pbta	4316	VIPLMEHVVEECLDIDSLELLIIGSDLCPADDEFKLU	DRFSSHMRIINSYGVTACVDS8
TS_15_Pbta	4314	VIPLMEHVVEERVDIDSLELLIIGSDLCPADDEFKLU	DRFGSHRIINSYGVTACVDS8
OSY_SE_Pbta	4282	VIPFMQYVYNNLDISSLELLIIGSDHCPTAEFKLT	TERGSHMRIINSYGVTACVDSAC
A6_61_Pbta	4376	VVEPRLSDSVRSVPICKPLPYVSMYILGENLSLOPVGLAGELYI	IAGAGVGRGYWNRPEMT
TS_15_Pbta	4374	VVEPRLSDPVRSPVICKPLPYVSMYILGENLSLOPVGLAGELYI	IAGAGVGRGYWNRPEMT
OSY_SE_Pbta	4342	VVERTTPDAENRTPICKPLPYVSMYILDNLSLOP	CHTGELYIGACVGRGYWNRPT
A6_61_Pbta	4436	ADKFVRDPFADGORNRYRTGDLAKWLPDGNIELIGRTDHQVKIRGYRIEIGEVESKLCQETP	
TS_15_Pbta	4434	ADKFVRDPFADGQCMYRTGDLAKWLPDGNIELIGRTDHQVKIRGYRIEIGEVESKLCQETP	
OSY_SE_Pbta	4402	VVKFVDNPFMPGASMYRTGDLAWLPDGNIEYAGRT	DHOVKIRGYRIEIGEVESQLLAAA
A6_61_Pbta	4496	HIREAAVVAKEDGSGKVLCAVYTS	RELTAGENWRAALAKELPAYMIPSHFMRLERMPLT
TS_15_Pbta	4494	DIREAAVVAKEDGSGKVLCAVYTS	RELTAGENWRAALAKELPAYMIPSHFMRLERMPLT
OSY_SE_Pbta	4462	GVREAAVVAEDGSCQKVLCAVVAESALIVGELRA	ESACQLPPTINIPAHFVQLERMPLT
A6_61_Pbta	4556	PNGKLDKGLPAPEGAAVYTGTEYCAPRTDAEIALAAWQSVLHVERVGTNDHFFELGCD8	
TS_15_Pbta	4554	PNGKLDKGLPAPEGAAVYTGTEYCAPRTDAEIALAAWQSVLHVERVGTNDHFFELGCD8	
OSY_SE_Pbta	4522	PNGKVDKGLPAPEGNAVYTGTEYCAPRTDAEIALAAWQSVLHVERVGTNDHFFELGCD8	
A6_61_Pbta	4616	IKSIQVSSRLHQAGYKLEIRDLFKYPTIAQLSPLOQPIGRIADQGEV	GEVELTPIQCMY
TS_15_Pbta	4614	IKSIQVSSRLHQAGYKLEIRDLFKYPTIAQLSPLOQPIGRIADQGEV	GEVELTPIQCMY
OSY_SE_Pbta	4582	IKSIQVSSRLHQAGYKLEIRDLFKYPTIAELSPHETQPIGRIADQGEV	GEVSAALTPIQCMY
A6_61_Pbta	4676	FGDLDDMMHYNQSPFLYRQ	GFNEEALRKTLR
TS_15_Pbta	4674	FGDLDDMMHYNQSPFLYRQ	GFNEEALRKTLR
OSY_SE_Pbta	4642	FGQEFADPHHYNQSPFLYRQ	GFDEAERKTLR
A6_61_Pbta	4736	EEGELFDPLAFDIANSQDAEQVIEAKANDIQASIDLQGGPLVKAGLPRCEQGHLLIAIH	
TS_15_Pbta	4734	EEGELFDPLAFDIANSQDAEQVIEAKANDIQASIDLQGGPLVKAGLPRCEQGHLLIAIH	
OSY_SE_Pbta	4702	CEGELVSLNVADFTDDPACYRAIEAKANDIQSCINLQGGPLVKAGLPRCEQGHLLIAIH	
A6_61_Pbta	4796	HAVDGVSWRILLEDIAAGYEQACKGDDIRLPSTDSYAANSRSLVEYA	EHYDGLGHERSY
TS_15_Pbta	4794	HAVDGVSWRILLEDISAGYEQACKGDDIRLPSTDSYAANSRSLVEYA	SLTDLGHERSY
OSY_SE_Pbta	4762	HAVDGVSWRILLEDIAAGYEQACKGRIELPSTDSYLTNS	CLSCGYAQSPAEQHERSY

25/48

FIG.16J

A6 61	pbtA	4856	MRHVLNAGTNPLPKDFDTESLQODSNSVTVANNOQDTEHLLKRVHAYNTDMNEILLAA
TS 15	pbtA	4854	MRHVLNAGTNPLPKDFDTESLQODSNSVTVANNOQDTEHLLKRVHAYNTDMNEILLAA
OSY SE	pbtA	4822	MQRIACSNTPKPLKDRITVNVSLQEDSESVQVQSEREDTEQLKHHVHAYNTDMNEILLTA
A6 61	pbtA	4916	LATAIQKWSCHNQILVNLECHGCHREPIAGDLDISRTVGMPTSEYPVLLKAEERDGLAYHIK
TS 15	pbtA	4914	LATAIQKWSCHNQILVNLECHGCHREPIAGDLDISRTVGMPTSEYPVLLKAEERDGLAYHIK
OSY SE	pbtA	4882	LSMAIQKWSGDRMLVNLECHGCHRESIMADLDISRTVGMPTSEYPVLLKAEERDGLAYHIK
A6 61	pbtA	4976	RAKEELRQIPNKGIGYICRYLSEPDLSLENGAARPEISFNYLQDFDQDMSGGMMLSPYS
TS 15	pbtA	4974	RAKEELRQIPNKGIGYICRYLSEPDLSLENGAARPEISFNYLQDFDQDMSGGMMLSPYS
OSY SE	pbtA	4942	XVKEELRQIPNKGIGYICRYLSEPDLSLENGAARPEISFNYLQDFDQDMSGGMMLSPYS
A6 61	pbtA	5036	KGSDGSAHTRQYVLDINGAITDGLTLDMYSYSEKRYRKETMBELLACHFHESLLSILHHC
TS 15	pbtA	5034	KGSDGSAHTRQYVLDINGAITDGLTLDMYSYSEKRYRKETMBELLACHFHESLLSILHHC
OSY SE	pbtA	5002	SGSDGSAHTRQYVLDINGAITDGLTLDMYSYSEKRYRKETMBELLACHFHESLLSILHHC
A6 61	pbtA	5096	VSRQOTELTPSDLLQGLSIEQLAEETHELGIENNIMYLTIPROKGMWFHFNALDQCEG
TS 15	pbtA	5094	VSRQOTELTPSDLLQGLSIEQLAEETHELGIENNIMYLTIPROKGMWFHFNALDQCEG
OSY SE	pbtA	5062	AAERTELTPSDLLQGLSIEQLAEETHELGIENNIMYLTIPROKGMWFHFNALDQCEG
A6 61	pbtA	5156	AGAYFEQTRFTLRGELDPALFAQSLHSLAARHSVLRNTPCS-LDGEPVQVFPREGRIIP
TS 15	pbtA	5154	AGAYFEQTRFTLRGELDPALFAQSLHSLAARHSVLRNTPCS-LDGEPVQVFPREGRIIP
OSY SE	pbtA	5120	-AGAYFEQTRFTLRGELDPALFAQSLHSLAARHSVLRNTPCS-LDGEPVQVFPREGRIIP
A6 61	pbtA	5215	TYEDLSQLPAEQAVVWERVASDKLQGFDERDPLVRVTLNRTAESCHVLWSSHILM
TS 15	pbtA	5213	TYEDLSQLPAEQAVVWERVASDKLQGFDERDPLVRVTLNRTAESCHVLWSSHILM
OSY SE	pbtA	5179	VYEDLSQLPAEQAVVWERVASDKLQGFDERDPLVRVTLNRTAESCHVLWSSHILM
A6 61	pbtA	5275	DGNCPLPQTLDELPRIPYSAVTNHAGTTEATCTVCIIP---SASRNKEANLP-----DY
TS 15	pbtA	5273	DGNCPLPQTLDELPRIPYSAVTNHAGTTEATCTVCIIP---SASRNKEANLP-----DY
OSY SE	pbtA	5239	DGNCPLPQTLDELPRIPYSAVTNHAGTTEATCTVCIIP---SASRNKEANLP-----DY
A6 61	pbtA	5327	SRYIEWLAEQDMSAAAETWNGYLAGEIQOTRLPNKGIIVKDKPIVLEQASRKLGIQITSR
TS 15	pbtA	5328	SRYIEWLAEQDMSAAAETWNGYLAGEIQOTRLPNKGIIVKDKPIVLEQASRKLGIQITSR
OSY SE	pbtA	5276	SCYIEWLAEQDMSAAAETWNGYLAGEIQOTRLPNKGIIVKDKPIVLEQASRKLGIQITSR

26/48

FIG. 16K

A6_61_Pbta	5387	MIRIAKQHQVTLNLTLLQAANGIVLQKYNCTQDVVFGVSGRPADVPGVESMIGLPINTI
TS_15_Pbta	5388	MIRIAKQHQVTLNLTLLQAANGIVLQKYNCTQDVVFGVSGRPADVPGVESMIGLPINTI
OSY_SE_Pbta	5336	MIRIAKQHQVTLNLTLLQAANGIVLQKYNCTQDVVFGVSGRPADVPGVESMIGLPINTI
A6_61_Pbta	5447	PVRVSCAGASPSDVMSQLQNAALESGRYDYVPLYEIQSRTSQKSELISHIMVFENYPLD
TS_15_Pbta	5448	PVRVSCAGASPSDVMSQLQNAALESGRYDYVPLYEIQSRTSQKSELISHIMVFENYPLD
OSY_SE_Pbta	5396	PVRVSCAGASPSDVMSQLQNAALESGRYDYVPLYEIQSRTSQKSELISHIMVFENYPLD
A6_61_Pbta	5507	ERMEOGTGNDGALALTDVQAAREQINVDNLMVVPGEELIIRFDNSEVYDRGHMERLHH
TS_15_Pbta	5508	ERMEOGTGNDGALALTDVQAAREQINVDNLMVVPGEELIIRFDNSEVYDRGHMERLHH
OSY_SE_Pbta	5456	ERMEOGTGNDGALALTDVQAAREQINVDNLMVVPGEELIIRFDNSEVYDRGHMERLHH
A6_61_Pbta	5567	HLMBVLEQVTGNPAISIAEVQLSTEAERKAEVQSAFNDTVVDYPREQTIHMPFEEQVQRT
TS_15_Pbta	5568	HLMBVLEQVTGNPAISIAEVQLSTEAERKAEVQSAFNDTVVDYPREQTIHMPFEEQVQRT
OSY_SE_Pbta	5516	HLMBVLEQVTGNPAISIAEVQLSTEAERKAEVQSAFNDTVVDYPREQTIHMPFEEQVQRT
A6_61_Pbta	5627	DAAAVLYGDDVITYRELMNERANRLARTLRAAGVPPDQIVGIMAEERSLELMVGIMGILKAG
TS_15_Pbta	5628	DAAAVLYGDDVITYRELMNERANRLARTLRAAGVPPDQIVGIMAEERSLELMVGIMGILKAG
OSY_SE_Pbta	5576	DAAAVLYGDDVITYRELMNERANRLARTLRAAGVPPDQIVGIMAEERSLELMVGIMGILKAG
A6_61_Pbta	5687	GAYVPIAPDYPEERIRYMLDSDKAQVLLVQGSAGEAVDFVGTGRIINLDDAEAYDEDESNP
TS_15_Pbta	5688	GAYVPIAPDYPEERIRYMLDSDKAQVLLVQGSAGEAVDFVGTGRIINLDDAEAYDEDESNP
OSY_SE_Pbta	5636	GAYVPIAPDYPEERIRYMLDSDKAQVLLVQGSAGEAVDFVGTGRIINLDDAEAYDEDESNP
A6_61_Pbta	5747	PVNKPTDIAVITYTSGTTGTPKGVMEHTSVINRLLMQKRYPIGSEDTIMQKTAITFDV
TS_15_Pbta	5748	PVNKPTDIAVITYTSGTTGTPKGVMEHTSVINRLLMQKRYPIGSEDTIMQKTAITFDV
OSY_SE_Pbta	5696	PVNKPTDIAVITYTSGTTGTPKGVMEHTSVINRLLMQKRYPIGSEDTIMQKTAITFDV
A6_61_Pbta	5807	SVWELFWWAFVGSKVLMSVGGKKNPHAVDAIERHRIITTHHFVPSMLHAPLEHVEQITD
TS_15_Pbta	5808	SVWELFWWAFVGSKVLMSVGGKKNPHAVDAIERHRIITTHHFVPSMLHAPLEHVEQITD
OSY_SE_Pbta	5756	SVWELFWWAFVGSKVLMSVGGKKNPHAVDAIERHRIITTHHFVPSMLHAPLEHVEQITD
A6_61_Pbta	5867	AERERGLAPLRQVFTSGEALLASQVERPFRHYIAPASGAQLINLYGPTTEATVDVTFDCEP
TS_15_Pbta	5868	AERERGLAPLRQVFTSGEALLASQVERPFRHYIAPASGAQLINLYGPTTEATVDVTFDCEP
OSY_SE_Pbta	5816	AERERGLAPLRQVFTSGEALLASQVERPFRHYIAPASGAQLINLYGPTTEATVDVTFDCEP

27/48

FIG.16L

A6_61_pbtA	5927	GQTVSVPIGKPIDNTSIYIVNEHNQVQPIGVAGELCIAGVGLARGYNRPBELTAEKFPVT
TS_15_pbtA	5928	GQTVSVPIGKPIDNTSIHIVNEHNQVQPIGVAGELCIAGVGLARGYNRPBELTAEKFPVT
OSY_SE_pbtA	5976	GQPIYTSVPIGKPIDNTCIYIVNRQDQIPIGVAGELCIAGVGLARGYLRPBELTAEKFPVT
A6_61_pbtA	5987	IPSVYGERMYRTGDLARWLPDGNIEYLGRIHQVKIRGYRIELGELERALLNVOEIRKTV
TS_15_pbtA	5988	IPSVYGERMYRTGDLARWLPDGNIEYLGRIHQVKIRGYRIELGELERALLNVOEIRKTV
OSY_SE_pbtA	5936	IPFMPGARMYRTGDLARWLPDGSIEYLGRIHQVKIRGYRIELGELERACLLQVEFIREAV
A6_61_pbtA	6046	VVAREEEDGQKSLCAYIVADGDPVTCMLRAALAAELPSYNIIPSYFIQLEQMP LAPNGKLD
TS_15_pbtA	6047	VVAREEEDGQKSLCAYIVADGDPVTCMLRAALAAELPSYNIIPSYFIQLEQMP LAPNGKLD
OSY_SE_pbtA	5996	VVAREEESGQKALCAIFAADELPSVSELRSALAVELPSYNIIPSYFIQLEELPLSANGKLD
A6_61_pbtA	6106	RKALPAPKDVICTGTDHMAPRTALEVKLVIRINQEVVLGDIQIGVKDDPFELGCHSLRATAL
TS_15_pbtA	6107	RKALPAPKDVIRICTDEMAPRTALEVKLVIRINQEVVLGDIQIGVKDDPFELGCHSLRATAL
OSY_SE_pbtA	6056	RKALPAPGCSRECKEHAPRESLEVKLVIRINQEVVLGLAHVSVKDDPFELGCHSLRATL
A6_61_pbtA	6166	ASKVSKEMHVALPLRDIPIHYSTLEAMAQAIGELEKQEHRAIPIAPMAEHYPLASAKRRLY
TS_15_pbtA	6167	ASKVSKEMHVALPLRDIPIHYSTLEAMAQAIGELEKQEHRAIPIAPMAEHYPLASAKRRLY
OSY_SE_pbtA	6116	VSKHEKENINPLRDIPIHYSTLEDMNAIGRTQEQEFQTIQVEASESYPLSSAKRRLY
A6_61_pbtA	6226	ILHOAEGAQCSYNNPGAMSVSGCHIDRNRLAALLELIARHDTLRTSFEMVDGEPVORVHQ
TS_15_pbtA	6227	ILHOAEGAQCSYNNPGAMSVSGCHIDRNRLAALLCILARHDTLRTSFEMVDGEPVORVHQ
OSY_SE_pbtA	6176	ILQOVGEAEQSYNNPGAMSTRGQDRNQFEAAFRGLIARHDEVFRTSFEMVCGEPQORVHQ
A6_61_pbtA	6286	HVDPALEYSTAREKDEQVAKQFVRDEPDLQPPLEVLVGLVQLEQSEQHLLLEDMMHIIISD
TS_15_pbtA	6287	HVDPALEYSTAREKDEQVAKQFVRDEPDLQPPLEVLVGLVQLEQSEQHLLLEDMMHIIISD
OSY_SE_pbtA	6236	DTAFAYEYMQANEEAEAAHQFVRTFDLEQPSLLEVLVGLTELDREHHHLLFDMHIIISD
A6_61_pbtA	6346	GISMIDILVDELARLYDGEELPPLLEIQKDYVLNQQAASSEQMKKEHEEYWLRTLGNELPL
TS_15_pbtA	6347	GISMIDILVDELARLYDGEELPPLLEIQKDYVLNQQAASSEQMKKEHEEYWLRTLGNELPL
OSY_SE_pbtA	6295	GWSMCIIVSEFARLYSGEELPPLLEIQKDYANWQQSEACQSEKQCEAVWLDAIDGELPQ

FIG.16M

A6 6i Pbta	6406	LELPTEFARCEQORSYDGDKXLFPAIDGQLNKQLQRLASQSGATLYMVLLAAAYTTLHLHKYSC
TS 15 Pbta	6407	LELPTEFARCEQORSYDGDKXLFPAIDGQLNKQLQRLASQSGATLYMVLLAAAYTTLHLHKYSC
OSY SE Pbta	6355	LELPTEFARPAHQSHGDTLDFVIDSMSGGLQRLAHEHGTLLYVLLAAAYTTLHLHKYSD
A6 6i Pbta	6466	QNDLVVGTPIAGRTHVDVEPLIGMFVNSLAIRNYPNDDKTFRSYLEEVKESTLSAFEHQD
TS 15 Pbta	6467	QNDLVVGTPIAGRTHVDVEPLIGMFVNSLAIRNYPNDDKTFRSYLEEVKESTLSAFEHQD
OSY SE Pbta	6415	QNDLVVGTPIAGRTHADVEPLIGMFVNSLAIRNYPNDDKTFRSYLEEVKESTLSAFEHQD
A6 6i Pbta	6526	YPPDKLVEQLEDAAWVPGRNVPFDTMFVLQNAKARTINLGELEPEPLIPSHTVAKFDLTLE
TS 15 Pbta	6527	YPPDKLVEQLEDAAWVPGRNVPFDTMFVLQNAKARTINLGELEPEPLIPSHTVAKFDLTLE
OSY SE Pbta	6475	YPPDKLVEQLEDAAWVPGRNVPFDTMFVLQNAKARTINLGELEPEPLIPSHTVAKFDLTLE
A6 6i Pbta	6586	MAIEDGMLSGOFEYCTKLFSANNIANFAEDFELEILSQACEOPDIRLEDIQLSGSAYQEEE
TS 15 Pbta	6587	MAIEDGMLSGOFEYCTKLFSANNIANFAEDFELEILSQACEOPDIRLEDIQLSGSAYQEEE
OSY SE Pbta	6535	MAIEDGMLSGOFEYCTKLFSANNIANFAEDFELEILSQACEOPDIRLEDIQLSGSAYQEEE
A6 6i Pbta	6646	LEEEIDFAP
TS 15 Pbta	6647	LEEEIDFAP
OSY SE Pbta	6595	ASLEEEIDFAP

A6 6i Pbta is SEQ ID NO. 11
TS-15 Pbta is SEQ ID NO. 14
OSY SE Pbta is SEQ ID NO. 17

FIG. 17A

A6 61 pbtB	1	--SAPDKEIEFPWAKALDTEPTILPYTSTESSEAAHYSVSVPAPISERIIRMSKGS
TS 15 pbtB	1	--SAPDKEIEFPWAKALDTEPTILPYTSTESSEAAHYSVSVPAPISERIIRMSKGS
OSY SE pbtB	1	MNFAFKEETFPNAQFDEDPAILPYTASISVTARDHNSSTSSASQRIHMSGS
A6 61 pbtB	59	HCAAFMILLGGIOCLLHKYITSENRIVIGMPIVRKAGEKRLPINOVLKKNVNEELTFKS
TS 15 pbtB	59	HCAAFMILLGGIOCLLHKYITSENRIVIGMPIVRKAGEKRLPINOVLKKNVNEELTFKS
OSY SE pbtB	61	HLAE M I L L G I O C L L Y K Y T C H E S Y I G M P I V R K S K T R P I N V V L K N K G A N R T P K S
A6 61 pbtB	119	LLTSLKOSFTEAIRHQHDPFRLLITEOMNVOKNGLPVINTMAALKNIHTVNP IPTVADV
TS 15 pbtB	119	LLTSLKOSFTEAIRHQHDPFRLLITEOMNVOKNGLPVINTMAALKNIHTVNP IPTVADV
OSY SE pbtB	121	LLTELEKTLTAIHNHONIPPEKTEHHEHLEAVNGPNTNVSKEIHTTESSQVVSPD
A6 61 pbtB	179	LPOFEPEAENLLSVVYNERNVYDSVPISQIIEHLQRVLSIVLLEPNTNLGLRLSDEET
TS 15 pbtB	179	LPOFEPEAENLLSVVYNERNVYDSVPISQIIEHLQRVLSIVLLEPNTNLGLRLSDEET
OSY SE pbtB	181	LPOFEPEOEVLSLEVTNNERNYDKPPIQIINHHVSEFADVLYTPERVLODNLLESQET
A6 61 pbtB	239	SLLHGFNNTAAEYPRDRTIHELFTEQAKRTPDAAVAVGCOOCLTYAELNGRANRLARTL
TS 15 pbtB	239	SLLHGFNNTAAEYPRDRTIHELFTEQAKRTPDAAVAVGCOOCLTYAELNGRANRLARTL
OSY SE pbtB	241	AQLLYEPNDTAAVYPRDTTIHOLFEBQEEQTPDAVAVLEDKOLTYAELNANRANRLELL
A6 61 pbtB	299	ONAGVRSDOLVGIMAERSLEMIVGLLAIKAGGAYVPIDPEYPOERIRYMLEDSCAGTLL
TS 15 pbtB	299	ONAGVRSDOLVGIMAERSLEMIVGLLAIKAGGAYVPIDPEYPOERIRYMLEDSCAGTLL
OSY SE pbtB	301	RERCVARGTLVGCIVERSLEMTGLLIIKAGGAYVPIDPEYPERTNAMLEDALISLL
A6 61 pbtB	359	QODHLRERV--SYEGTIVDMNSEHNYHDDCTELASVSDSS-NLAIYVITSGTIGNPKGVN
TS 15 pbtB	359	QODHLRERV--SYEGTIVDMNSEHNYHDDCTELASVSDSS-NLAIYVITSGTIGNPKGVN
OSY SE pbtB	361	TQRIQTSMPNSIDSVELDAAAETALEGCEWPMETDAAGTADDYAYIYTSCGTCEKGYC
A6 61 pbtB	416	TEHRSVUNALLNRIRTYGLSSDDRILQLFSPSPDGFVMSAFCSLLSGAGLFLKKEEDAKD
TS 15 pbtB	416	TEHRSVUNALLNRIRTYGLSSDDRILQLFSPSPDGFVMSAFCSLLSGAGLFLKKEEDAKD
OSY SE pbtB	421	YTERGVNR-LTAAANYVDSSKDVFLQGSTISPDATFEINGSLNGRLAFLPPGNV-S
A6 61 pbtB	476	PLALHCAISQSGITHEFICVNPNLICALLNVMOAESVSTLRITVTLAGESVSSALVARSQOL
TS 15 pbtB	476	PLALHCAISQSGITHEFICVNPNLICALLNVMOAESVSTLRITVTLAGESVSSALVARSQOL
OSY SE pbtB	479	LTENTEAIQQHQWILNLTAGLEHVME-NQLQALQGVQQLLGGVVSXTHAKVLSRY

30/48

FIG. 17B

A6_61_pbtB	536	PMVKLFNSVGYPTENSVAATC	---AIGLEKQDQITIGTPISNASVLLINTSGELQPLHVPG
TS_15_pbtB	536	PMVKLFNSVGYPTENSVAATC	---AIGLEKQDQITIGTPISNASVLLINTSGELQPLHVPG
OSY_SE_pbtB	538	QDIHLVNSGYPTENTIFTCC	HEISADTERLSHPICRPIANTQVYLDGAGKLLPVCVVG
A6_61_pbtB	593	ELCIAGEGLARGYLNRPELT	TEKFAAHFVPFGERIYHGTDSARWLPNGTIEYLGRIIDHGV
TS_15_pbtB	593	ELCIAGEGLARGYLNRPELT	TEKFAAHFVPFGERIYHGTDSARWLPNGTIEYLGRIIDHGV
OSY_SE_pbtB	598	ELYTGGGLARGYNRP	ELTAEKFDVSPFVPGCTREYETGDIARWLPDGTIEYVGRIDHGV
A6_61_pbtB	653	KIRGPRIELGELIESSLLK	AGVREVIVDARPDGNGGQMLCAIMVADAEITVTELREALSS
TS_15_pbtB	653	KIRGPRIELGELIESSLLK	AGVREVIVDARPDGNGGQMLCAIMVADAEITVTELREALSS
OSY_SE_pbtB	658	KISGRIELCE	EAHLKVESVLDIAINFAQDESSGQTLCAVFTANAEIMACGIRAVLSQ
A6_61_pbtB	713	NLPDYNIPSHFPVOMEQL	PLTPSGKLDKRLSPDPOANEAIGTEYIAPRTPLEARLAQINQE
TS_15_pbtB	713	NLPDYNIPSHFPVOMEQL	PLTPSGKLDKRLSPDPOANEAIGTEYIAPRTPLEARLAQINQE
OSY_SE_pbtB	718	ELPATTNIP	HPFVQVQRPLTPNGZDRALPPEGLINTGCHVAPRTPLESMLAHENQE
A6_61_pbtB	773	SLGVEVVGIKDNFP	FALGGHSLRAATLASKLHKLNVNVPRLDLPNPTIEELALLMDGME
TS_15_pbtB	773	SLGVEVVGIKDNFP	FALGGHSLRAATLASKLHKLNVNVPRLDLPNPTIEELALLMDGME
OSY_SE_pbtB	778	VLCSEKVS	WKDSPFEEIGCHSLEATTLASKLHKLHVSPLRDIPRPHPTIEELACLDGME
A6_61_pbtB	833	QQEFSAI	ERVKESEHYVSSSAQKRLVVLQOLEGAEQSYNMPGAMLLEGCLDRERLEASFR
TS_15_pbtB	833	QQEFSAI	ERVKESEHYVSSSAQKRLVVLQOLEGAEQSYNMPGAMLLEGCLDRERLEASFR
OSY_SE_pbtB	838	RQSYRQIFL	DERGQVYPVSSAQKRLYVLEHQLEGAEQSYNMPGVMILEGCLDRERLEASFR
A6_61_pbtB	893	KLIARHET	TLRTGPELMDGEPVQKVYQSVSPAIEYMQSESEAAQKAREFIRAFDLMTPPL
TS_15_pbtB	893	KLIARHET	TLRTGPELMDGEPVQKVYQSVSPAIEYMQSESEAAQKAREFIRAFDLMTPPL
OSY_SE_pbtB	898	SLIGRHET	TLRTGPEVNGEPVQVTCRSEVNFSEMEQCASEEADAAIRSFIRPFDELEKPL
A6_61_pbtB	953	MRVGLIEMAPDRHVLL	YDMHHIISDGASMGVVEEFARLYGGEELPPLRIQYKDFAAWQQ
TS_15_pbtB	953	MRVGLIEMAPDRHVLL	YDMHHIISDGASMGVVEEFARLYGGEELPPLRIQYKDFAAWQQ
OSY_SE_pbtB	958	LRVGLI	ESQCDRHLLYDMHHIISDGVSMEIVVEEFVRLYGGEELPPLRIQYKDFAAWQQ
A6_61_pbtB	1013	SEACQKRS	KQQRHNLQTFGELPVLPLEPTDIARPAIOSYEGETIEFTVDSIDISTALQRL
TS_15_pbtB	1013	SEVQCKRS	QQEAYNLQAFGGELPVLPLEPTDIARPAIOSYEGETIEFTVDSIDISTALQRL
OSY_SE_pbtB	1018	SEPOCEL	MLKQQRHNLQAFGGELPVLPLEPTDIARPAIOSYEGETIEFTVDSIDISTALQRL

31/48

FIG. 17C

A6_61_pbtB	1073	AADSGTTLNVLAAAYTVLLHKYTGQEDIVVGTNAGRMHDDLQPLIGMFVNTLAIRNYP
TS_15_pbtB	1073	AADSGTTLNVLAAAYTVLLHKYTGQEDIVVGTNAGRMHDDLQPLIGMFVNTLAIRNYP
OSY_SE_pbtB	1078	AA[REDACTED]SGTTLNVLAAAYTVLLHKYTGQEDIVVGTNAGRMHDDLQPLIGMFVNTLAIRNYP
A6_61_pbtB	1133	AGESTFRAYLEQVKEQALAAFEHQEYPPFEELVEKL[REDACTED]RDMSRNPLPDTMPSLQNMENKD
TS_15_pbtB	1133	AGESTFRAYLEQVKEQALAAFEHQEYPPFEELVEKL[REDACTED]RDMSRNPLPDTMPSLQNMENKD
OSY_SE_pbtB	1138	AGSKTFLEYLEQVKE[REDACTED]SLCAFEN[REDACTED]YPPFEELVEKL[REDACTED]VARD[REDACTED]SRNPLPDTMPSL[REDACTED]ENESKD
A6_61_pbtB	1193	FELPGLOLKPYGFEHQISKFDLSLDVAEGADGLACSELYASSLYRQDTIVRMA[REDACTED]HYRQLL
TS_15_pbtB	1193	FELPGLOLKPYGFEHQISKFDLSLDVAEGADGLACSELYASSLYRQDTIVRMA[REDACTED]HYRQLL
OSY_SE_pbtB	1198	LE[REDACTED]A[REDACTED]RLKPYPA[REDACTED]YK[REDACTED]VAKPDL[REDACTED]SLDVAEGAG[REDACTED]AC[REDACTED]SELYAT[REDACTED]LYR[REDACTED]ESHERMA[REDACTED]H[REDACTED]FCOLL
A6_61_pbtB	1253	HSIAQSP[REDACTED]AQISVLGMLTPGEGEQIRFKFNHDPSEMEQKHTVHQLPFEQAA[REDACTED]LTPERTAVV
TS_15_pbtB	1253	HSIAQSP[REDACTED]AQISVLGMLTPGEGEQIRFKFNHDPSEMEQKHTVHQLPFEQAA[REDACTED]LTPERTAVV
OSY_SE_pbtB	1258	EA[REDACTED]TCEPEARE[REDACTED]SSLGMLTE[REDACTED]E[REDACTED]K[REDACTED]AQIQHVFNDAEAGRSQ[REDACTED]ETVPELFEQ[REDACTED]VERTPERTAVV
A6_61_pbtB	1313	HENEQLSYRELNERANRLARTLRQHGVOPEQLVGLADRS[REDACTED]LDMIVGIMAILKAGGAYVPI
TS_15_pbtB	1313	HENEQLSYRELNERANRLARTLRQHGVOPEQLVGLADRS[REDACTED]LDMIVGIMAILKAGGAYVPI
OSY_SE_pbtB	1318	HEDKQL[REDACTED]SYRELNERANRLARTLR[REDACTED]AEDV[REDACTED]PEQLVGI[REDACTED]ADRS[REDACTED]LDMIVGIMAILK[REDACTED]SGGAYVPI
A6_61_pbtB	1373	DPKYPEERIRYML[REDACTED]EDSKANVLVTQSHLQSL[REDACTED]SFDGTVLLDEESSYAE[REDACTED]AANLVSI[REDACTED]TEPQ
TS_15_pbtB	1373	DPKYPEERIRYML[REDACTED]EDSKANVLVTQSHLQSL[REDACTED]SFDGTVLLDEESSYAE[REDACTED]AANLVSI[REDACTED]TEPQ
OSY_SE_pbtB	1378	DPKYPEERIRYML[REDACTED]DSNA[REDACTED]QVLA[REDACTED]QEHLCARA[REDACTED]FSGRR[REDACTED]TLDDEEAF[REDACTED]YDE[REDACTED]G[REDACTED]SNLER[REDACTED]N[REDACTED]QPE
A6_61_pbtB	1433	HLAYVLYTSGTTGQPKGAMIEHRQLTVMAKAWEREYRLREESIRMMQWASPS[REDACTED]SDFVPSGDL
TS_15_pbtB	1433	HLAYVLYTSGTTGQPKGAMIEHRQLTVMAKAWEREYRLREESIRMMQWASPS[REDACTED]SDFVPSGDL
OSY_SE_pbtB	1438	HL[REDACTED]SYVLYTSGTTGQPKGAMIEHRQ[REDACTED]MAV[REDACTED]SAAN[REDACTED]SEYELREESIRMMQWASPS[REDACTED]SDFVPSGDL
A6_61_pbtB	1493	IRALLHGGELVLCPEHARANPAEIIYELIRK[REDACTED]RLHMFDC[REDACTED]TPSIVIPLM[REDACTED]EYVENKLDISSL
TS_15_pbtB	1493	IRALLHGGELVLCPEHARANPAEIIYELIRK[REDACTED]RLHMFDC[REDACTED]TPSIVIPLM[REDACTED]EYVENKLDISSL
OSY_SE_pbtB	1498	IRALLHGGELVLCPE[REDACTED]SRANPAEIIYELIRK[REDACTED]RLHMFDC[REDACTED]TPSIVIPLM[REDACTED]EYVENKLDISSL
A6_61_pbtB	1553	KLVAVGSDYCP[REDACTED]PDEPQKMLDRPGSQ[REDACTED]PRIINSYGVTE[REDACTED]TCIDASYE[REDACTED]ZPTTPTVP[REDACTED]RALPIGKP
TS_15_pbtB	1553	KLVAVGSDYCP[REDACTED]PDEPQKMLDRPGSQ[REDACTED]PRIINSYGVTE[REDACTED]TCIDASYE[REDACTED]ZPTTPTVP[REDACTED]RALPIGKP
OSY_SE_pbtB	1558	KLVAVGADHCP[REDACTED]ZEEFQKMLERPGSQ[REDACTED]PRIINSYGVTE[REDACTED]TIDSCY[REDACTED]EQASTEGLE[REDACTED]STPIGKP

32/48

FIG. 17D

A6_61_pbtB	1613	LPGVTVIMDQORSLLPVGVIGELYIGGPCVGRGYNRSEMTMEKFPVSDPPFLQDYRMVYT
TS_15_pbtB	1613	LPGVTVIMDQORSLLPVGVIGELYIGGPCVGRGYNRSEMTMEKFPVSDPPFLQDYRMVYT
OSY_SE_pbtB	1618	LPGVTVIMDQORSLLPVGVIGELYIGGPCVGRGYNRSEMTMEKFPVSDPPFLQDYRMVYT
A6_61_pbtB	1673	GDLARMPDGNIEYLGRIDHQC4KIRGYRIEIGEVESEKLLKVTRESVIVARQDPNGTKA
TS_15_pbtB	1673	GDLARMPDGNIEYLGRIDHQC4KIRGYRIEIGEVESEKLLKVTRESVIVARQDPNGTKA
OSY_SE_pbtB	1678	GDLARMPDGNIEYLGRIDHQC4KIRGYRIEIGEVESEKLLKVTRESVIVARQDPNGTKA
A6_61_pbtB	1733	LCAYFVADRNLTVSELSALADELPAYMIPSPVQDLRLPLTPNGKVDRKALPAPEAGAH
TS_15_pbtB	1733	LCAYFVADRNLTVSELSALADELPAYMIPSPVQDLRLPLTPNGKVDRKALPAPEAGAH
OSY_SE_pbtB	1738	LCAYFVADRNLTVSELSALADELPAYMIPSPVQDLRLPLTPNGKVDRKALPAPEAGAH
A6_61_pbtB	1793	TGIEYVAPRTESEELANVMQTVLGIERVGVQDHPFELGDSIKSIQVASRLQOAGYKLE
TS_15_pbtB	1793	TGIEYVAPRTESEELANVMQTVLGIERVGVQDHPFELGDSIKSIQVASRLQOAGYKLE
OSY_SE_pbtB	1798	TGIEYVAPRTESEELANVMQTVLGIERVGVQDHPFELGDSIKSIQVASRLQOAGYKLE
A6_61_pbtB	1853	IRDLCKYPTIAQLGSEHLQRAKSVADQGEVSGDVPLTPILGMPFEQQPADAHHYNQSIMLY
TS_15_pbtB	1853	IRDLCKYPTIAQLGSEHLQRAKSVADQGEVSGDVPLTPILGMPFEQQPADAHHYNQSIMLY
OSY_SE_pbtB	1858	IRDLCKYPTIAQLGSEHLQRAKSVADQGEVSGDVPLTPILGMPFEQQPADAHHYNQSIMLY
A6_61_pbtB	1913	RREGPDFAAIRNVLOAVTEHHDALRIIVPRNDQGDYTANNRAIEEGELPHLEVNLTCGA
TS_15_pbtB	1913	RREGPDFAAIRNVLOAVTEHHDALRIIVPRNDQGDYTANNRAIEEGELPHLEVNLTCGA
OSY_SE_pbtB	1918	RREGPDFAAIRNVLOAVTEHHDALRIIVPRNDQGDYTANNRAIEEGELPHLEVNLTCGA
A6_61_pbtB	1973	SGDSEQNVROIIIEAKA4EIQRSFDLHDGCLARAGLFRTEGDHLLLVMMHGVVDGVSWRI
TS_15_pbtB	1973	SGDSEQNVROIIIEAKA4EIQRSFDLHDGCLARAGLFRTEGDHLLLVMMHGVVDGVSWRI
OSY_SE_pbtB	1977	SGDSEQNVROIIIEAKA4EIQRSFDLHDGCLARAGLFRTEGDHLLLVMMHGVVDGVSWRI
A6_61_pbtB	2033	LLEDIAAGIEQALKGEFVRLPAKTDSPRTWANHLASVARSSEAMIEEQFWEQASANATSI
TS_15_pbtB	2033	LLEDIAAGIEQALKGEFVRLPAKTDSPRTWANHLASVARSSEAMIEEQFWEQASANATSI
OSY_SE_pbtB	2031	LLEDIAAGIEQALKGEFVRLPAKTDSPRTWANHLASVARSSEAMIEEQFWEQASANATSI
A6_61_pbtB	2093	GLPKDFEAEISLQDSESSEVSESEETDMLLKHVRAYNTDMDILLALGMAIQQNC
TS_15_pbtB	2093	GLPKDFEAEISLQDSESSEVSESEETDMLLKHVRAYNTDMDILLALGMAIQQNC
OSY_SE_pbtB	2090	GLPKDFEAEISLQDSESSEVSESEETDMLLKHVRAYNTDMDILLALGMAIQQNC

33/48

FIG. 17E

A6 61	pbtB	2153	CHEKALVTLECHGCHRENIMP	ELDI	SRTV	GNFT	SEY	PF	LL	ES	DP	NK	SL	SY	KIK	RM	KEN	L	R	I
TS 15	pbtB	2153	CHEKALVTLECHGCHRENIMP	ELDI	SRTV	GNFT	SEY	PF	LL	ES	DP	NK	SL	SY	KIK	RM	KEN	L	R	I
OSY SE	pbtB	2149	GRDEVLTLECHGCHRES	IM	TDI	ERTV	GNFT	SEY	PF	LL	EM	PD	SL	SR	KIK	RM	KED	LE	CI	
A6 61	pbtB	2213	PNKGICYGICHRVLS	SG	SC	EN	VS	RS	EA	SA	QP	EL	SP	YL	GF	QD	QD	LN	NE	ME
TS 15	pbtB	2213	PNKGICYGICHRVLS	SG	SC	EN	VS	RS	EA	SA	QP	EL	SP	YL	GF	QD	QD	LN	NE	ME
OSY SE	pbtB	2209	PNKGICYGICRV	SL	HD	EA	VN	-----	GR	EP	PI	SE	PN	YL	GF	QD	QD	KN	NE	ME
A6 61	pbtB	2273	AEISVRQARNTTLD	FN	GM	IS	AG	VL	AE	VS	SS	KO	YR	RD	TID	RLA	LL	KG	SL	Q
TS 15	pbtB	2273	AEISVRQARNTTLD	FN	GM	IS	AG	VL	AE	VS	SS	KO	YR	RD	TID	RLA	LL	KG	SL	Q
OSY SE	pbtB	2263	ME	SE	Q	Q	A	R	H	AL	IN	GM	T	A	D	GE	S	A	L	E
A6 61	pbtB	2333	KDKPELT	PS	D	V	L	N	G	L	G	I	E	D	L	E	R	I	A	E
TS 15	pbtB	2333	KDKPELT	PS	D	V	L	N	G	L	G	I	E	D	L	E	R	I	A	E
OSY SE	pbtB	2323	K	E	P	E	T	P	S	D	L	E	G	I	S	E	L	E	Q	I
A6 61	pbtB	2393	FEQTRFTIQ	G	E	L	D	V	Q	F	A	S	S	L	D	V	L	A	T	H
TS 15	pbtB	2393	FEQTRFTIQ	G	E	L	D	V	Q	F	A	S	S	L	D	V	L	A	T	H
OSY SE	pbtB	2383	P	B	Q	T	R	F	T	Q	A	L	D	V	E	P	A	K	S	
A6 61	pbtB	2453	TELPED	E	K	O	D	R	V	E	A	M	A	O	A	D	E	R	G	P
TS 15	pbtB	2453	TELPED	E	K	O	D	R	V	E	A	M	A	O	A	D	E	R	G	P
OSY SE	pbtB	2443	C	A	L	P	E	A	C	E	T	E	V	E	T	I	Q	E	D	
A6 61	pbtB	2513	P	O	L	T	Q	E	N	L	E	T	S	D	S	V	N	G	R	
TS 15	pbtB	2513	P	O	L	T	Q	E	N	L	E	T	S	D	S	V	N	G	R	
OSY SE	pbtB	2503	P	O	L	T	Q	E	N	L	E	T	S	D	S	V	N	G		
A6 61	pbtB	2573	P	O	K	S	S	G	R	S	D	V	I	A	D	N	L	V		
TS 15	pbtB	2573	P	O	K	S	S	G	R	S	D	V	I	A	D	N	L	V		
OSY SE	pbtB	2563	P	O	A	K	Q	G	R	S	D	E	M	A	E	R	I			
A6 61	pbtB	2633	AVFGGVV	S	G	R	P	A	E	I	P	G	I	E	S	M	I	G		
TS 15	pbtB	2633	AVFGGVV	S	G	R	P	A	E	I	P	G	I	E	S	M	I	G		
OSY SE	pbtB	2623	V	F	G	G	V	S	G	R	P	A	E	I	P	G	I			

34/48

FIG.17F

A6	61	Pb	bE	2693	PLYDIQALSTHKQDLINHLVFPENYPMEZQMEQAG - DERGQINITDVRVAEQTSYDFNLY
TS	15	Pb	bE	2693	PLYDIQALSTHKQDLINHLVFPENYPMEZQMEQAG - DERGQINITDVRVAEQTSYDFNLY
OSY	SE	Pb	bE	2683	PLYDIQALSTHKQDLINHLVFPENYPMEZQMEQAG - DERGQINITDVRVAEQTSYDFNLY
A6	61	Pb	bE	2752	VMPGEDMMIRLEYNNAVYDRADMER - RQHLIHVLQVTAADPAIAVKDVELATDDEKAELL
TS	15	Pb	bE	2752	VMPGEDMMIRLEYNNAVYDRADMER - RQHLIHVLQVTAADPAIAVKDVELATDDEKAELL
OSY	SE	Pb	bE	2743	VMPGEDMMIRLEYNNAVYDRADMER - RQHLIHVLQVTAADPAIAVKDVELATDDEKAELL
A6	61	Pb	bE	2812	TAFNDTEIRYPREOMIHRMPFEQVQRTPDATAVLMGAAANTYREMNERRANQLARTLRAAG
TS	15	Pb	bE	2812	TAFNDTEIRYPREOMIHRMPFEQVQRTPDATAVLMGAAANTYREMNERRANQLARTLRAAG
OSY	SE	Pb	bE	2803	TAFNDTEIRYPREOMIHRMPFEQVQRTPDATAVLMGAAANTYREMNERRANQLARTLRAAG
A6	61	Pb	bE	2872	VVPDQIVGIMAERSLELMVGIMGILKAGGAYVPIADYPPEERIRYMLDDSEAOVLIVQGS
TS	15	Pb	bE	2872	VVPDQIVGIMAERSLELMVGIMGILKAGGAYVPIADYPPEERIRYMLDDSEAOVLIVQGS
OSY	SE	Pb	bE	2863	VVPDQIVGIMAERSLELMVGIMGILKAGGAYVPIADYPPEERIRYMLDDSEAOVLIVQGS
A6	61	Pb	bE	2932	AGEAIDFAGREVINLDDADSDYDQDSSNLEMVNKPTDIAYIYTSGTTGRPKGVMEHTSVI
TS	15	Pb	bE	2932	AGEAIDFAGREVINLDDADSDYDQDSSNLEMVNKPTDIAYIYTSGTTGRPKGVMEHTSVI
OSY	SE	Pb	bE	2923	AGEAIDFAGREVINLDDADSDYDQDSSNLEMVNKPTDIAYIYTSGTTGRPKGVMEHTSVI
A6	61	Pb	bE	2992	NRLLWMQKRYPIGADDTIMOKTAITFDVSVWELFWAFVGSKVLMLPVGGEKNPAAIIVEA
TS	15	Pb	bE	2992	NRLLWMQKRYPIGADDTIMOKTAITFDVSVWELFWAFVGSKVLMLPVGGEKNPAAIIVEA
OSY	SE	Pb	bE	2983	NRLLWMQKRYPIGADDTIMOKTAITFDVSVWELFWAFVGSKVLMLPVGGEKNPAAIIVEA
A6	61	Pb	bE	3052	IEQYIETMHPVPSMLHAPLEHISQLPEARERER - LSPLQVFTSGEALLASQVERPHQVY
TS	15	Pb	bE	3052	IEQYIETMHPVPSMLHAPLEHISQLPEARERER - LSPLQVFTSGEALLASQVERPHQVY
OSY	SE	Pb	bE	3043	IEQYIETMHPVPSMLHAPLEHISQLPEARERER - LSPLQVFTSGEALLASQVERPHQVY
A6	61	Pb	bE	3111	VPASGARLINLYGPTTEATVDVTYFDCEPGQTVSVPIQKPIDNTRIYIVN EHFQVQPIGV
TS	15	Pb	bE	3111	VPASGARLINLYGPTTEATVDVTYFDCEPGQTVSVPIQKPIDNTRIYIVN EHFQVQPIGV
OSY	SE	Pb	bE	3103	VPASGARLINLYGPTTEATVDVTYFDCEPGQTVSVPIQKPIDNTRIYIVN EHFQVQPIGV
A6	61	Pb	bE	3171	AGELCIACVGLARGYWNRPTELTEKEFVLVP - SVGERMYRTGDLARWLDPDGNIEYLGRIIDH
TS	15	Pb	bE	3171	AGELCIACVGLARGYWNRPTELTEKEFVLVP - SVGERMYRTGDLARWLDPDGNIEYLGRIIDH
OSY	SE	Pb	bE	3163	AGELCIACVGLARGYWNRPTELTEKEFVLVP - SVGERMYRTGDLARWLDPDGNIEYLGRIIDH

FIG. 17G

A6_61_pbtB	3230	QVKIRGYRIELGELLETALLINIDAVRETVVAREDESQKSLCAYTVADGEATSDLRAL
TS_15_pbtB	3230	QVKIRGYRIELGELLETALLINIDAVRETVVAREDESQKSLCAYTVADGEATSDLRAL
OSY_SE_pbtB	3223	QVKIRGYRIELGESQAQLQAAARETVVAREDESQKSLCAYTAADSSETVSLEAL
A6_61_pbtB	3290	AAELPSYNIPIPSYFVRLQOMPLAPNGKIDRKALPAPERSSLOVESYVAPRTEAKOMLITVW
TS_15_pbtB	3290	AAELPSYNIPIPSYFVRLQOMPLAPNGKIDRKALPAPERSSLOVESYVAPRTEAKOMLITVW
OSY_SE_pbtB	3283	AAELPSYNIPIPSYFVRLQOMPLAPNGKIDRKALPAPERSSLOVESYVAPRTEAKOMLITVW
A6_61_pbtB	3350	QAVLGIERSVGIITDHPFELGCDISIKSIQVAARMQOAGFKLEIRDLPKYPTVTOLVPYMQPI
TS_15_pbtB	3350	QAVLGIERSVGIITDHPFELGCDISIKSIQVAARMQOAGFKLEIRDLPKYPTVTOLVPYMQPI
OSY_SE_pbtB	3343	QAVLGIERSVGIITDHPFELGCDISIKSIQVAARMQOAGFKLEIRDLPKYPTVTOLVPYMQPI
A6_61_pbtB	3410	NRTADQGEVVGVEVPMTPILHWPFEHQOFANPHHFNQSVMLYRKDGFVADAVCKALHKLVEH
TS_15_pbtB	3410	NRTADQGEVVGVEVPMTPILHWPFEHQOFANPHHFNQSVMLYRKDGFVADAVCKALHKLVEH
OSY_SE_pbtB	3403	NRTADQGEVVGVEVPMTPILHWPFEHQOFANPHHFNQSVMLYRKDGFVADAVCKALHKLVEH
A6_61_pbtB	3470	HDALRIVIQRTQCEYSLNRSLSVEGELEFSMGEIDLTDQSDFAAAIAEAAHIIQGSIDLQ
TS_15_pbtB	3470	HDALRIVIQRTQCEYSLNRSLSVEGELEFSMGEIDLTDQSDFAAAIAEAAHIIQGSIDLQ
OSY_SE_pbtB	3463	HDALRIVIQRTQCEYSLNRSLSVEGELEFSMGEIDLTDQSDFAAAIAEAAHIIQGSIDLQ
A6_61_pbtB	3530	AGPLVKAGLPHGSDGDHLLLVIIHHAVIDGVSWRILLEDLAGGYEQALNQQRVRLPMKTD
TS_15_pbtB	3530	AGPLVKAGLPHGSDGDHLLLVIIHHAVIDGVSWRILLEDLAGGYEQALNQQRVRLPMKTD
OSY_SE_pbtB	3523	AGPLVKAGLPHGSDGDHLLLVIIHHAVIDGVSWRILLEDLAGGYEQALNQQRVRLPMKTD
A6_61_pbtB	3590	PRTWAQQLVEYANSPAMDKESAYNLGVAQTEVAALPKDSECTVSLQORDSESVVLEWNRD
TS_15_pbtB	3590	PRTWAQQLVEYANSPAMDKESAYNLGVAQTEVAALPKDSECTVSLQORDSESVVLEWNRD
OSY_SE_pbtB	3583	PRTWAQQLVEYANSPAMDKESAYNLGVAQTEVAALPKDSECTVSLQORDSESVVLEWNRD
A6_61_pbtB	3650	TERLLKHVHRAVNTMDDILLTALGRALMKWRCIGIRLVTLLEGHGRESILQCMIDITRTV
TS_15_pbtB	3650	TERLLKHVHRAVNTMDDILLTALGRALMKWRCIGIRLVTLLEGHGRESILQCMIDITRTV
OSY_SE_pbtB	3643	TERLLKHVHRAVNTMDDILLTALGRALMKWRCIGIRLVTLLEGHGRESILQCMIDITRTV
A6_61_pbtB	3710	WFTTEYVPELGMENDSLGSGQIKKVKEDLRRIPNKGIGIGLFRYLSNSGKQAWNDAPTQ
TS_15_pbtB	3710	WFTTEYVPELGMENDSLGSGQIKKVKEDLRRIPNKGIGIGLFRYLSNSGKQAWNDAPTQ
OSY_SE_pbtB	3703	WFTTEYVPELGMENDSLGSGQIKKVKEDLRRIPNKGIGIGLFRYLSNSGKQAWNDAPTQ

36/48

FIG. 17H

A6_61	PbTb	3770	IRYNVLGQPDADLSNNE	SVSPYASCSSEISDEQERKYPDLINGVIAEQTLGLS	SVKE
TS_15	PbTb	3770	IRYNVLGQPDADLSNNE	SVSPYASCSSEISDEQERKYPDLINGVIAEQTLGLS	SVKE
OSY_SE	PbTb	3763	IRYNVLGQPDADLSNNE	SVSPYASCSSEISDEQERKYPDLINGVIAEQTLGLS	SVKE
A6_61	PbTb	3830	YKKTMEELGDL	LTESLKEIIAHCESQERTQLTSPDVLPKGLSLEWLDRISSQ	MOHIGEI
TS_15	PbTb	3830	YKKTMEELGDL	LTESLKEIIAHCESQERTQLTSPDVLPKGLSLEWLDRISSQ	MOHIGEI
OSY_SE	PbTb	3823	YKKTMEELGDL	LTESLKEIIAHCESQERTQLTSPDVLPKGLSLEWLDRISSQ	MOHIGEI
A6_61	PbTb	3890	ENVYALTTPMQKGMWPHSAMD	SLTGAYHEQTMFTLEG	LDVELPSSSLNELAKRHAVLRTN
TS_15	PbTb	3890	ENVYALTTPMQKGMWPHSAMD	SLTGAYHEQTMFTLEG	LDVELPSSSLNELAKRHAVLRTN
OSY_SE	PbTb	3883	ENVYALTTPMQKGMWPHSAMD	SLTGAYHEQTMFTLEG	LDVELPSSSLNELAKRHAVLRTN
A6_61	PbTb	3950	FISGPQGEPPQVVFERNKPIG	FSFQDVRALNEEQSPFIKEAVSSDQLLCPDLAOGALMRV	
TS_15	PbTb	3950	FISGPQGEPPQVVFERNKPIG	FSFQDVRALNEEQSPFIKEAVSSDQLLCPDLAOGALMRV	
OSY_SE	PbTb	3943	FISGPQGEPPQVVFERNKPIG	FSFQDVRALNEEQSPFIKEAVSSDQLLCPDLAOGALMRV	
A6_61	PbTb	4010	SAIRTGELSCRVLWSSHHILMD	GWCLPOLMQELFDITYAALLQKESPD	-TAVPAYSQYIE
TS_15	PbTb	4010	SAIRTGELSCRVLWSSHHILMD	GWCLPOLMQELFDITYAALLQKESPD	-TAVPAYSQYIE
OSY_SE	PbTb	4003	SAIRTGELSCRVLWSSHHILMD	GWCLPOLMQELFDITYAALLQKESPD	-TAVPAYSQYIE
A6_61	PbTb	4069	NLGOODEBAAGTIWSAYLD	DIDQVTEIPOESSAGIDSEPIKAEKWSRELDAGLSASISRT	
TS_15	PbTb	4069	NLGOODEBAAGTIWSAYLD	DIDQVTEIPOESSAGIDSEPIKAEKWSRELDAGLSASISRA	
OSY_SE	PbTb	4063	NLGOODEBAAGTIWSAYLD	DIDQVTEIPOESSAGIDSEPIKAEKWSRELDAGLSASISRA	
A6_61	PbTb	4129	AROHQVTLNTLLQAAANGVIL	OKYNGTNDVVFSGSVSGRPAEVP	PGIETMIGLFINITIPRV
TS_15	PbTb	4129	AROHQVTLNTLLQAAANGVIL	OKYNGTNDVVFSGSVSGRPAEVP	PGIETMIGLFINITIPRV
OSY_SE	PbTb	4122	AROHQVTLNTLLQAAANGVIL	OKYNGTNDVVFSGSVSGRPAEVP	PGIETMIGLFINITIPRV
A6_61	PbTb	4189	KCEGSTSPAELMGLLQEOALES	GKIDYIPLYEIQSRSAALKQNAIQIMVFN	YPMDEQLE
TS_15	PbTb	4189	KCEGSTSPAELMGLLQEOALES	GKIDYIPLYEIQSRSAALKQNAIQIMVFN	YPMDEQLE
OSY_SE	PbTb	4182	KCEGSTSPAELMGLLQEOALES	GKIDYIPLYEIQSRSAALKQNAIQIMVFN	YPMDEQLE
A6_61	PbTb	4249	QACGDEHGMPSLTDVAVEEQ	TNYDFNLIVPGEQISIRPFDYNANRFVQADMERL	MGLN
TS_15	PbTb	4249	QACGDEHGMPSLTDVAVEEQ	TNYDFNLIVPGEQISIRPFDYNANRFVQADMERL	MGLN
OSY_SE	PbTb	4242	QACGDEHGMPSLTDVAVEEQ	TNYDFNLIVPGEQISIRPFDYNANRFVQADMERL	MGLN

37/48

FIG. 171

A6_61_pbtB	4309	ILLEQIVDNPRJAVEDLELATEAKSEVQSFNDTFTINVPDRDMLHRLPFEQAEHRPDAVA
TS_15_pbtB	4309	ILLEQIVDNPREVAVEDLELATEAKSEVQSFNDTFTINVPDRDMLHRLPFEQAEHRPDAVA
OSY_SE_pbtB	4302	WLEQIAANFQVTVGELELATEAKSEVQSFNDTFTINVPDRDMLHRLPFEQAEHRPDAVA
A6_61_pbtB	4369	ISPRDFQMTYRDNDVNRRLARTLRAVGVTGDKLVGLMSERSPDMLIGILAILKAGGGYV
TS_15_pbtB	4369	ISPRDFQMTYRDNDVNRRLARTLRAVGVTGDKLVGLMSERSPDMLIGILAILKAGGGYV
OSY_SE_pbtB	4362	VMTENSQMTYRDLNRANRLARTLRAVGVTGDKLVGLMSERSPDMLIGILAILKAGGGYV
A6_61_pbtB	4429	PIDPEYPCERIRYMLEDSCARIMLAQQHLTGKIPVMDASPLDAIINLDTSTSYDSNGSNL
TS_15_pbtB	4429	PIDPEYPCERIRYMLEDSCARIMLAQQHLTGKIPVMDASPLDAIINLDTSTSYDSNGSNL
OSY_SE_pbtB	4422	PIDPEYPCERIRYMLEDSCARIMLAQQHLTGKIPVMDASPLDAIINLDTSTSYDSNGSNL
A6_61_pbtB	4489	EAANTDASSENLAQVIYTSCTGCKPKGNLTTHRNIVRVVRETEYIDITNHDNVLMQSSYAF
TS_15_pbtB	4489	EAANTDASSENLAQVIYTSCTGCKPKGNLTTHRNIVRVVRETEYIDITNHDNVLMQSSYAF
OSY_SE_pbtB	4477	DLNNGPA--DLAYVIYTSCTGCKPKGNLTTHRNIVRVVRETEYIDITNHDNVLMQSSYAF
A6_61_pbtB	4549	DGSTPDIYGALLNGAKLVLPVPHETLLEVRQLAELIVQEKISVMFITTAIFNVLVVDVQASC
TS_15_pbtB	4549	DGSTPDIYGALLNGAKLVLPVPHETLLEVRQLAELIVQEKISVMFITTAIFNVLVVDVQASC
OSY_SE_pbtB	4535	DGSTPDIYGALLNGAKLVLPVPHETLLEVRQLAELIVQEKISVMFITTAIFNVLVVDVQASC
A6_61_pbtB	4609	LSNIRAILPGGERVSVSHVRKALNHVAPGALKHVYGPTESTVFATCHDVHEVTENAVTVF
TS_15_pbtB	4609	LSNIRAILPGGERVSVSHVRKALNHVAPGALKHVYGPTESTVFATCHDVHEVTENAVTVF
OSY_SE_pbtB	4595	LRHIRAILPGGERVSVSHVRKALNHVAPGALKHVYGPTESTVFATCHDVHEVTENAVTVF
A6_61_pbtB	4669	HGRPISNTSIYIVDANNKLOPVGVAGELCVAGDGLARGYLNRPDLTAEEKFVDSPIVQGER
TS_15_pbtB	4669	HGRPISNTSIYIVDANNKLOPVGVAGELCVAGDGLARGYLNRPDLTAEEKFVDSPIVQGER
OSY_SE_pbtB	4655	HGRPISNTSIYIVDANNKLOPVGVAGELCVAGDGLARGYLNRPDLTAEEKFVDSPIVQGER
A6_61_pbtB	4729	MYRTGDLAKNLPDGSIEYVGRIDQOVKIRGYRIELGEIEAQLNVEDVQEAUVVARDTDT
TS_15_pbtB	4729	MYRTGDLAKNLPDGSIEYVGRIDQOVKIRGYRIELGEIEAQLNVEDVQEAUVVARDTDT
OSY_SE_pbtB	4715	MYRTGDLAKNLPDGSIEYVGRIDQOVKIRGYRIELGEIEAQLNVEDVQEAUVVARDTDT
A6_61_pbtB	4789	GEKQLCAYVYVAMRPLEANHLREVVMQAMPSYMLPAHFQLEQLPLTPNGKVDKRALPAPE
TS_15_pbtB	4789	GEKQLCAYVYVAMRPLEANHLREVVMQAMPSYMLPAHFQLEQLPLTPNGKVDKRALPAPE
OSY_SE_pbtB	4775	GEKQLCAYVYVAMRPLEANHLREVVMQAMPSYMLPAHFQLEQLPLTPNGKVDKRALPAPE

FIG. 17J

A6_61	Pb1E	4849	EC-RSGETFTPTPLEAQLVQINODVLGISESSVTDHFFELGCHSLKATLLVNRHQEL
TS_15	Pb1E	4849	EC-RSGETFTPTPLEAQLVQINODVLGISESSVTDHFFELGCHSLKATLLVNRHQEL
OSY_SE	Pb1E	4835	EHVQCTETAPRTTCEQELARINQVLCERKGVNPNFPELCHSLKATLLVNRHQEL
A6_61	Pb1E	4908	NIELPLKDVFOIPTLEVMAKLSNAEGSRHVSIPVAAPSOHYPVSSAQKRLYLILHQLEGA
TS_15	Pb1E	4908	NIELPLKDVFOIPTLEVMAKLSNAEGSRHVSIPVAAPSOHYPVSSAQKRLYLILHQLEGA
OSY_SE	Pb1E	4895	SHIELPLDVFKETLEAMAEELAGLCQOMYTEIIPVVEEQAHYPSAQKRLYLILHQLEGA
A6_61	Pb1E	4968	ELSYNMPNMELLECAVDLGRLEEAERLIERHETLRTGPEIVNGEPPVQRIIPEVDFAIER
TS_15	Pb1E	4968	ELSYNMPNMELLECAVDLGRLEEAERLIERHETLRTGPEIVNGEPPVQRIIPEVDFAIER
OSY_SE	Pb1E	4955	ELSYNMPNMILLECAVDLGRLEEAERLIERHETLRTGPEIVNGEPPVQRIIPEVDFAIER
A6_61	Pb1E	5028	VLAEEGASKLMQOPVRSFQLEKPPLLRIGIIELSSEERSILMFDMSHIIISDGTSMGILIN
TS_15	Pb1E	5028	VLAEEGASKLMQOPVRSFQLEKPPLLRIGIIELSSEERSILMFDMSHIIISDGTSMGILIN
OSY_SE	Pb1E	5015	VLAEEGASKLMQOPVRSFQLEKPPLLRIGIIELSSEERSILMFDMSHIIISDGTSMGILIN
A6_61	Pb1E	5098	EFVHLISCAELTPLRIQKDYAVWQOSDTPQKANKLQEGYWLKVLGGELPVLEMPDTSIR
TS_15	Pb1E	5098	EFVHLISCAELTPLRIQKDYAVWQOSDTPQKANKLQEGYWLKVLGGELPVLEMPDTSIR
OSY_SE	Pb1E	5075	EFVHLISCAELTPLRIQKDYAVWQOSDTPQKANKLQEGYWLKVLGGELPVLEMPDTSIR
A6_61	Pb1E	5148	PTTQSPRGDLLOPDLDPVRSAGLRRIAAENGATNIVLLALIKTTLHKYISQEDIIIVGTP
TS_15	Pb1E	5148	PTTQSPRGDLLOPDLDPVRSAGLRRIAAENGATNIVLLALIKTTLHKYISQEDIIIVGTP
OSY_SE	Pb1E	5135	PTTQSPRGDLLOPDLDPVRSAGLRRIAAENGATNIVLLALIKTTLHKYISQEDIIIVGTP
A6_61	Pb1E	5208	IAGRNHGDLOPLLCMFVNTLAIRSYPAAASKTFLSYLGEIKETTLGAFENQNYPFALVEQ
TS_15	Pb1E	5208	IAGRNHGDLOPLLCMFVNTLAIRSYPAAASKTFLSYLGEIKETTLGAFENQNYPFALVEQ
OSY_SE	Pb1E	5195	IAGRNHGDLOPLLCMFVNTLAIRSYPAAASKTFLSYLGEIKETTLGAFENQNYPFALVEQ
A6_61	Pb1E	5268	VQVMRDMSRNPVFDTNFILONADQGMKIDGRLQSVNPNHTVSKFDLTFOAEDEAEIV
TS_15	Pb1E	5268	VQVMRDMSRNPVFDTNFILONADQGMKIDGRLQSVNPNHTVSKFDLTFOAEDEAEIV
OSY_SE	Pb1E	5255	VQVMRDMSRNPVFDTNFILONADQGMKIDGRLQSVNPNHTVSKFDLTFOAEDEAEIV
A6_61	Pb1E	5328	CSIEVATDLFKRGTIERMARHFEQLVDVLDNMPQASLSNLSMVTNEEKALLQDKFNDDTM
TS_15	Pb1E	5328	CSIEVATDLFKRGTIERMARHFEQLVDVLDNMPQASLSNLSMVTNEEKALLQDKFNDDTM
OSY_SE	Pb1E	5315	CSIEVATDLFKRGTIERMARHFEQLVDVLDNMPQASLSNLSMVTNEEKALLQDKFNDDTM

39/48

FIG. 17K

A6_61_pbtB	5388	AHPSDKTVHFLPAEQVETPDAAVAVSC	CEQLSYGELNRKANQLANKLREYGVTAEQPVG
TS_15_pbtB	5388	AHPSDKTVHFLPAEQVETPDAAVAVSC	CEQLSYGELNRKANQLANKLREYGVTAEQPVG
OSY_SE_pbtB	5375	GYPREKTHQFEECAERTPDAAVAVSC	CEQLSYGELNRKANQLANKLREYGVTAEQPVG
A6_61_pbtB	5448	IIVERTLDTVAVAVLAVLKASGTEVPIDPEYPETRIKRYMLADSCAKLVLAQSEI	PGIIPDD
TS_15_pbtB	5448	IIVERTLDTVAVAVLAVLKASGTEVPIDPEYPETRIKRYMLADSCAKLVLAQSEI	PGIIPDD
OSY_SE_pbtB	5435	IIVERTLDTVAVAVLAVLKASGTEVPIDPEYPETRIKRYMLADSCAKLVLAQSEI	PGIIPDD
A6_61_pbtB	5508	VRLIDVRDESLYQGDGADVPNGSKPSNLLYIYTS	GTGTPKGVMLHGNMUNLLHYQOK
TS_15_pbtB	5508	VRLIDVRDESLYQGDGADVPNGSKPSNLLYIYTS	GTGTPKGVMLHGNMUNLLHYQOK
OSY_SE_pbtB	5495	VRLIDVRDESLYQGDGADVPNGSKPSNLLYIYTS	GTGTPKGVMLHGNMUNLLHYQOK
A6_61_pbtB	5568	GTNIPMPSRILOYASGSPDVCYQEMPSALLFGGSLYAVDNEMRKDPVRLPQEI	IEKHEIDV
TS_15_pbtB	5568	GTNIPMPSRILOYASGSPDVCYQEMPSALLFGGSLYAVDNEMRKDPVRLPQEI	IEKHEIDV
OSY_SE_pbtB	5555	GTNIPMPSRILOYASGSPDVCYQEMPSALLFGGSLYAVDNEMRKDPVRLPQEI	IEKHEIDV
A6_61_pbtB	5628	LYIPVAFLEKPIFAPEWAEAPPRCVRHIIITAGEQLVVTQVQACLRDLICLHNHYGPSE	
TS_15_pbtB	5628	LYIPVAFLEKPIFAPEWAEAPPRCVRHIIITAGEQLVVTQVQACLRDLICLHNHYGPSE	
OSY_SE_pbtB	5615	LYIPVAFLEKPIFAPEWAEAPPRCVRHIIITAGEQLVVTQVQACLRDLICLHNHYGPSE	
A6_61_pbtB	5688	THVVVTYMTPEIIEVGLPPIGKPIANTSIYVNDSPFELQPIGVKGEI	VSGACVGRGYM
TS_15_pbtB	5688	THVVVTYMTPEIIEVGLPPIGKPIANTSIYVNDSPFELQPIGVKGEI	VSGACVGRGYM
OSY_SE_pbtB	5675	THVVVTYMTPEIIEVGLPPIGKPIANTSIYVNDSPFELQPIGVKGEI	VSGACVGRGYM
A6_61_pbtB	5748	GRDILTSEKFLDNPFVPCERLYKTGDVARWLPDGSIEYVGRSDHQVKIRGFRIELGEVES	
TS_15_pbtB	5748	GRDILTSEKFLDNPFVPCERLYKTGDVARWLPDGSIEYVGRSDHQVKIRGFRIELGEVES	
OSY_SE_pbtB	5735	GRDILTSEKFLDNPFVPCERLYKTGDVARWLPDGSIEYVGRSDHQVKIRGFRIELGEVES	
A6_61_pbtB	5808	QLLHVPVQVQVAVLEADHAGQKQKQCAVFTAEISLTAGEIRAVLSQELPGYMI	PSYFVQL
TS_15_pbtB	5808	QLLHVPVQVQVAVLEADHAGQKQKQCAVFTAEISLTAGEIRAVLSQELPGYMI	PSYFVQL
OSY_SE_pbtB	5795	QLLHVPVQVQVAVLEADHAGQKQKQCAVFTAEISLTAGEIRAVLSQELPGYMI	PSYFVQL
A6_61_pbtB	5868	ERLPLTPNGKIDRRALPKPEGGIEGTGVVAPRTTEARLARIWQDVGLGPSVGVKDNFF	
TS_15_pbtB	5868	ERLPLTPNGKIDRRALPKPEGGIEGTGVVAPRTTEARLARIWQDVGLGPSVGVKDNFF	
OSY_SE_pbtB	5855	ERLPLTPNGKIDRRALPKPEGGIEGTGVVAPRTTEARLARIWQDVGLGPSVGVKDNFF	

40/48

FIG.17L

A6	61	pbtB	5928	ELGGHSLRATTLLVSRLLYKEMNVNFPPLRGVFRHPTIEMSCAISOQNETS	NTAIPAIABEQE
TS	15	pbtB	5928	ELGGHSLRATTLLVSRLLYKEMNVNFPPLRGVFRHPTIEMSCAISOQNETS	LYTAIPAIABEQE
OSY	SE	pbtB	5915	ELGGHSLRATTLLVSRLLYKEMNVNFPPLRGVFRHPTIEMSCAISOQNETS	LYTAIPAIABEQE
A6	61	pbtB	5988	YIPLSSAQLRLYIMSQLGSELSINMPGMLVLEGQLNRDQFOTAFKLIAKHETLRTGFE	
TS	15	pbtB	5988	YIPLSSAQLRLYIMSQLGSELSINMPGMLVLEGQLNRDQFOTAFKLIAKHETLRTGFE	
OSY	SE	pbtB	5975	VYPLSSAQLRLYIMSQLGSELSINMPGMLVLEGQLNRDQFOTAFKLIAKHETLRTGFE	
A6	61	pbtB	6048	NVDGEPNORIHNTSPAFIDYQVSE	-DEVPE
TS	15	pbtB	6048	NVDGEPNORIHNTSPAFIDYQVSE	-DEVPE
OSY	SE	pbtB	6035	NVDGEPNORIHNTSPAFIDYQVSE	-DEVPE
A6	61	pbtB	6107	ILVFDNHHIISDCGMSNLVDEFTRLYANESRPPPLRIQYKDYAVNQQA	RENLERKKROED
TS	15	pbtB	6107	ILVFDNHHIISDCGMSNLVDEFTRLYANESRPPPLRIQYKDYAVNQQA	SENLERKKROED
OSY	SE	pbtB	6095	ILVFDNHHIISDCGMSNLVDEFTRLYANESRPPPLRIQYKDYAVNQQA	SENLERKKROED
A6	61	pbtB	6167	YVNSVLQGDLPNAELP	DIDR
TS	15	pbtB	6167	YVNSVLQGDLPNAELP	DIDR
OSY	SE	pbtB	6155	YVNSVLQGDLPNAELP	DIDR
A6	61	pbtB	6227	SAYQILLSKYCGQDDIIVGTPVAGRNHADLEPLIGMPVNTLAIRNRPQGDXTFQSELAEV	
TS	15	pbtB	6227	SAYQILLSKYCGQDDIIVGTPVAGRNHADLEPLIGMPVNTLAIRNRPQGDXTFQSELAEV	
OSY	SE	pbtB	6215	SAYQILLSKYCGQDDIIVGTPVAGRNHADLEPLIGMPVNTLAIRNRPQGDXTFQSELAEV	

FIG.17M

A6 6i PbtB	6227	SAYQILLISKYCGODDIIVGTPVAGRNHADLEPLIGMFVNTLAIRNRPOGDKTFOSEFLAEV
TS 15 PbtB	6227	SAYQILLISKYCGODDIIVGTPVAGRNHADLEPLIGMFVNTLAIRNRPOGDKTFOSEFLAEV
OSY SE PbtB	6215	SAYTLLISKYSGQDIIVGTPVAGRNHADLEPLIGMFVNTLAIRNRPOGDKTFOSEFLAEV
A6 6i PbtB	6287	KESTLGAFEHQEYPPFEELIDLLKQWETSERNPLFDTVPVLQNTERRRAGIGGLTISPYVT
TS 15 PbtB	6287	KESTLGAFEHQEYPPFEELIDLLKQWETSERNPLFDTVPVLQNTERRRAGIGGLTISPYVT
OSY SE PbtB	6275	KETA LGAFEHQEYPPFEELERLQWQDANRNPEDTVPVQNTERRRAGIGGLTISPYVT
A6 6i PbtB	6347	DDSVSAKFDLTLSVSEEDDGMKGSFLYASKLPKAGIHRMRDYLISILSQVCENPRIRIQ
TS 15 PbtB	6347	DDSVSAKFDLTLSVSEEDDGMKGSFLYASKLPKAGIHRMRDYLISILSQVCENPRIRIQ
OSY SE PbtB	6335	DNPEIDAKFDLTLSVSEEDDGMKGSFLYASKLPKAGIHRMRDYLISILSQVCENPRIRIQ
A6 6i PbtB	6407	DISISGQOQTOEKSXIDTIEFAP
TS 15 PbtB	6407	DISISGQOQTOEKSXIDTIEFAP
OSY SE PbtB	6395	DIKCNQSVNNQRSIETIEFAP

A6 6i PbtB is SEQ ID NO. 12
TS-15 PbtB is SEQ ID NO. 15
OSY SE PbtB is SEQ ID NO. 18

42/48

FIG. 18A

A6_61_pbtC	1	MRSVYEKEEAYWNGMFDSDSDSMSILPYCSTHGREANADTNAAKVSMI	---	---	RALPPELS
TS_15_pbtC	1	MRSVYEKEEAYWNGMFDSDSDSMSILPYCSTHGREANADTNAAKVSMI	---	---	RALPPELS
OSY_SE_pbtC	1	MRSVYEKEEAYWNGMFDSDSDSMSILPYCSTHGREANADTNAAKVSMI	---	---	RALPPELS
A6_61_pbtC	57	RMNTLANGSDYALYMIAGLACVTCLLHHYTNRENVLVGMPTIDES	DDSSP	---	RDVLIAMKTN
TS_15_pbtC	57	RMNTLANGSDYALYMIAGLACVTCLLHHYTNRENVLVGMPTIDES	DDSSP	---	RDVLIAMKTN
OSY_SE_pbtC	55	RMNTLANGSDYALYMIAGLACVTCLLHHYTNRENVLVGMPTIDES	DDSSP	---	RDVLIAMKTN
A6_61_pbtC	116	LTRGSSPFRSVLGSIKITAVGCALEHRRHLPPRKAVHNLNLEMDTNGLPVMTVVS	SAIHTA	---	SAIHTA
TS_15_pbtC	116	LTRGSSPFRSVLGSIKITAVGCALEHRRHLPPRKAVHNLNLEMDTNGLPVMTVVS	SAIHTA	---	SAIHTA
OSY_SE_pbtC	115	VNKSSEPSSELCCHIAVSSEALEHCHLPFRKAFRLNLQVDPGCLPIVNTIIVSYHNIHTA	---	---	SAIHTA
A6_61_pbtC	176	SIDLSVSSDVFRPDAKDDGLHLEVLYDESRYDASYATLFEHFRRLHFLVLPQDPQIG	---	---	SAIHTA
TS_15_pbtC	176	SIDLSVSSDVFRPDAKDDGLHLEVLYDESRYDASYATLFEHFRRLHFLVLPQDPQIG	---	---	SAIHTA
OSY_SE_pbtC	175	SEQSAANEAACQEFVNDRIQLRMSEDENRYSDESMALHFRRLHFLVLPQDPQIG	---	---	SAIHTA
A6_61_pbtC	236	NTLLSAEKKHRLHFNVDVTDPPQATLYQFIEEHAERQPDALIAVSIEDTKLTYRELN	---	---	SAIHTA
TS_15_pbtC	236	NTLLSAEKKHRLHFNVDVTDPPQATLYQFIEEHAERQPDALIAVSIEDTKLTYRELN	---	---	SAIHTA
OSY_SE_pbtC	235	NAVLLSEKKHRLHFNVDVTDPPQATLYQFIEEHAERQPDALIAVSIEDTKLTYRELN	---	---	SAIHTA
A6_61_pbtC	296	APANRLARTLRSRGVCTEALVGLMAERSIDMIVGMLAVLKAGGCVVAIDPEYPERVRVM	---	---	SAIHTA
TS_15_pbtC	296	APANRLARTLRSRGVCTEALVGLMAERSIDMIVGMLAVLKAGGCVVAIDPEYPERVRVM	---	---	SAIHTA
OSY_SE_pbtC	295	APANRLARTLRSRGVCTEALVGLMAERSIDMIVGMLAVLKAGGCVVAIDPEYPERVRVM	---	---	SAIHTA
A6_61_pbtC	356	LEDSCARVILVQOHLQNRVPNTESAAALLTLDDQSYHEDASNL	---	---	SAIHTA
TS_15_pbtC	356	LEDSCARVILVQOHLQNRVPNTESAAALLTLDDQSYHEDASNL	---	---	SAIHTA
OSY_SE_pbtC	355	LEDSCARVILVQOHLQNRVPNTESAAALLTLDDQSYHEDASNL	---	---	SAIHTA
A6_61_pbtC	415	SGTTCNPKGNLTTHRNIVRIVKNTNVIETQODKVLQSSYPDGSAPDIFCALTNGAOL	---	---	SAIHTA
TS_15_pbtC	415	SGTTCNPKGNLTTHRNIVRIVKNTNVIETQODKVLQSSYPDGSAPDIFCALTNGAOL	---	---	SAIHTA
OSY_SE_pbtC	414	SGTTCNPKGNLTTHRNIVRIVKNTNVIETQODKVLQSSYPDGSAPDIFCALTNGAOL	---	---	SAIHTA
A6_61_pbtC	475	VLPVPHHTLLDAKLAELIETEQISVMLITTAFFNVLDVNM	SCURHIRAILFGGERSVA	---	SCURHIRAILFGGERSVA
TS_15_pbtC	475	VLPVPHHTLLDAKLAELIETEQISVMLITTAFFNVLDVNM	SCURHIRAILFGGERSVA	---	SCURHIRAILFGGERSVA
OSY_SE_pbtC	474	VLPVPHHTLLDAKLAELIETEQISVMLITTAFFNVLDVNM	SCURHIRAILFGGERSVA	---	SCURHIRAILFGGERSVA

43/48

FIG. 18B

A6 61 pbtC	535	HVR ~~AA~~ LE ~~OT~~ GPGR ~~LK~~ HA ~~Y~~ GPSE ~~ST~~ VIATWHDVTEI ~~SE~~ RAVSVPIGRPI ~~SN~~ TAIYIVNERN
TS 15 pbtC	535	HVR ~~QA~~ LE ~~OT~~ GPGR ~~LK~~ HA ~~Y~~ GPSE ~~ST~~ VIATWHDVTEI ~~SE~~ QAVSVPIGRPI ~~SN~~ TAIYIVNERN
OSY SE pbtC	533	HAA ~~AV~~ LE ~~RY~~ XD ~~IR~~ LVN ~~Y~~ GH ~~PE~~ NT ~~ET~~ TC ~~CE~~ EE ~~SA~~ ADMER ~~ES~~ PIGRPIANTO ~~Y~~ LEDEAG
A6 61 pbtC	595	DLQPIGVSGELCVAGEGLVRCYLNRP ~~ELTE~~ QK ~~FV~~ DNPPFVPGERM ~~YR~~ TGDLARWL ~~PD~~ GTIE
TS 15 pbtC	595	DLQPIGVSGELCVAGEGLVRCYLNRP ~~ELTA~~ CK ~~TV~~ DNPPFVPGERM ~~YR~~ TGDLARWL ~~PD~~ GTIE
OSY SE pbtC	593	KL ~~LP~~ GVV ~~GE~~ LYT ~~CG~~ GLA ~~Q~~ CIYNNRP ~~ELTA~~ EK ~~FV~~ DNPPFVPGER ~~IR~~ TGDLARWL ~~PD~~ GTIE
A6 61 pbtC	655	YVGRMD ~~DQ~~ VKIRGHR ~~IE~~ IG ~~EV~~ EAQ ~~LL~~ KVA ~~PI~~ QKATI ~~VV~~ RGREDG ~~EG~~ KQLCA ~~YV~~ ADRL ~~LS~~ PA
TS 15 pbtC	655	YVGRMD ~~DQ~~ VKIRGHR ~~IE~~ IG ~~EV~~ EAQ ~~LL~~ KVA ~~LI~~ QKATI ~~VV~~ RGREDG ~~EG~~ KQLCA ~~YV~~ ADRL ~~LS~~ PA
OSY SE pbtC	653	YVGRMD ~~DQ~~ VKIRGVR ~~IE~~ IG ~~EV~~ EA ~~HL~~ LVESV ~~LD~~ NI ~~VV~~ ARQ ~~DES~~ Q ~~KT~~ LCAY ~~ET~~ NA ~~HA~~ ELMA
A6 61 pbtC	715	GEIR ~~TL~~ LAKEL ~~PS~~ YMI ~~PA~~ YFIOLEQ ~~MP~~ LTINGKVD ~~RK~~ ALPEPEH ~~VQ~~ AETEIVAPR ~~SEQE~~
TS 15 pbtC	715	GEIR ~~TL~~ LAKEL ~~PS~~ YMI ~~PA~~ YFIOLEQ ~~MP~~ LTINGKVD ~~RK~~ ALPEPEH ~~VQ~~ AETEIVAPR ~~SEQE~~
OSY SE pbtC	713	GE ~~ER~~ AA ~~LS~~ QEL ~~FV~~ YMI ~~PT~~ HEL ~~Q~~ VD ~~Q~~ MP ~~LT~~ FNCKV ~~DR~~ ALPEPEGL ~~MT~~ GI ~~EH~~ VAPR ~~SF~~ LE
A6 61 pbtC	775	IRLARV ~~NO~~ EV ~~LG~~ LSRVGAK ~~DH~~ FFELG ~~CH~~ SLRAT ~~TL~~ VS ~~KL~~ HKKE ~~NI~~ SLSL ~~RD~~ VFRNP ~~TL~~ EA
TS 15 pbtC	775	IRLARV ~~NO~~ EV ~~LG~~ LSRVGAK ~~DH~~ FFELG ~~CH~~ SLRAT ~~TL~~ VS ~~KL~~ HKKE ~~NI~~ SLSL ~~RD~~ VFRNP ~~TL~~ EA
OSY SE pbtC	773	SLA ~~HT~~ Q ~~Q~~ EV ~~LG~~ LE ~~VS~~ VD ~~SF~~ FE ~~GG~~ HS ~~LR~~ AT ~~TL~~ AS ~~KL~~ HK ~~EL~~ HV ~~SL~~ SL ~~RD~~ VFRNP ~~TL~~ EE
A6 61 pbtC	835	NAALME ~~DA~~ QGR ~~KE~~ FAPIPTVEEK ~~DV~~ YVSSVQ ~~KR~~ LFI ~~LH~~ QLEGA ~~EQ~~ SYNMPG ~~AL~~ LEG ~~DVD~~
TS 15 pbtC	835	NAALME ~~DA~~ QGR ~~TE~~ FSPIPTVEEK ~~DV~~ YVSSVQ ~~KR~~ LFI ~~LH~~ QLEGA ~~EQ~~ SYNMPG ~~AL~~ LEG ~~DVD~~
OSY SE pbtC	833	PAR ~~LD~~ SGMERQAYRQI ~~PL~~ DEED ~~WT~~ YVSSA ~~Q~~ KRL ~~FI~~ LH ~~Q~~ LEGA ~~EQ~~ SYNMPG ~~AL~~ LEG ~~QDD~~
A6 61 pbtC	895	RNRLEYA ~~FR~~ OLITR ~~HE~~ TLRTGFEMVNGEPVQRI ~~YPTV~~ CFV ~~VE~~ MSA ~~LEG~~ SEVE ~~EQ~~ IRO ~~PI~~
TS 15 pbtC	895	RNRLEYA ~~FR~~ OLITR ~~HE~~ TLRTGFEMVNGEPVQRI ~~YPTV~~ CFV ~~VE~~ MSA ~~LEG~~ SEVE ~~EQ~~ IRO ~~PI~~
OSY SE pbtC	893	RNR ~~FE~~ EA ~~FG~~ SLICR ~~HE~~ TLRTGFEMVNGEPVQRCRE ~~VN~~ FSVEN ~~MQ~~ ASE-EEEA ~~AV~~ RS ~~FI~~
A6 61 pbtC	955	RAFDLST ~~AP~~ LFRAG ~~LI~~ ELAPORHILLFD ~~NH~~ HIISD ~~GT~~ SIGIMIEE ~~PT~~ SLYSGNE ~~LE~~ PLRI
TS 15 pbtC	955	RAFDLST ~~PP~~ LFRAG ~~LI~~ ELAPORHILLFD ~~NH~~ HIISD ~~GT~~ SIGIMIEE ~~PT~~ SLYSGNE ~~LE~~ PLRI
OSY SE pbtC	952	RFPDLE ~~EP~~ PLRVGL ~~IE~~ LQDRHILL ~~DN~~ HHIISD ~~GV~~ SGI ~~V~~ VEF ~~VR~~ LYCGE ~~EL~~ FP ~~RI~~
A6 61 pbtC	1015	QYKDFAA ~~NQ~~ RSSE ~~EQ~~ IEQL ~~NR~~ Q ~~EE~~ Y ~~N~~ LQ ~~ME~~ GVLPVLE ~~PTD~~ VYRPAVQ ~~SHD~~ GALF ~~PS~~ ED
TS 15 pbtC	1015	QYKDFAA ~~NQ~~ RSSE ~~EQ~~ IEQL ~~KS~~ Q ~~EE~~ Y ~~N~~ LQ ~~ME~~ GVLPVLE ~~PTD~~ VYRPAVQ ~~SHD~~ GALF ~~PS~~ ID
OSY SE pbtC	1012	QYKDFAA ~~NQ~~ CSSE ~~FP~~ Q ~~CE~~ L ~~KK~~ Q ~~EE~~ Y ~~N~~ LQ ~~SE~~ GVLPVLE ~~PLD~~ VYRPAVQ ~~SV~~ Y ~~CD~~ T ~~FE~~ FAID

FIG. 18C

A6_61_pbtC	1075	REQSODLRKLAADTRTLLYMLAAAYTILHKKYSQEDIVVGTPIAGRTHDDVQPLIGMP
TS_15_pbtC	1075	REQSODLRKLAADTRTLLYMLAAAYTILHKKYSQEDIVVGTPIAGRTHDDVQPLIGMP
OSY_SE_pbtC	1072	PRISEAHGVAESGTTLYMVLAAAYTILHKKYSQEDIVVGTPIAGRTHDDVQPLIGMP
A6_61_pbtC	1135	VNTLAIRNYPGSKSVLTYLEEIKETTLGAPFHKQDYPPEELVENVQISRDMSRHPVFTM
TS_15_pbtC	1135	VNTLAIRNYPGSKSVLTYLEEIKETTLGAPFHKQDYPPEELVENVQISRDMSRHPVFTM
OSY_SE_pbtC	1132	VNTLAIRNYPGSKSVLTYLEEIKETTLGAPFHKQDYPPEELVENVQISRDMSRHPVFTM
A6_61_pbtC	1195	PALENTZHRREDFDGLQVKKPYGAEYGMKAPDNLTVTDEGDDGLYCTMEYATALYNRSTIE
TS_15_pbtC	1195	PALENTZHRREDFDGLQVKKPYGAEYGMKAPDNLTVTDEGDDGLYCTMEYATALYNRSTIA
OSY_SE_pbtC	1192	PALQNMDDGLGLGLRKPYPAPYKAPDNLTVTDEGDDGLYCTMEYATALYNRSTIE
A6_61_pbtC	1255	RLCGHFLQVVGSMTHNPQAAISSLOMVTISEEKAQLQHEFNGTVMETSSDKTVHELPAEQV
TS_15_pbtC	1255	RLCGHFLQVVGSMTHNPQAAISSLOMVTISEEKAQLQHEFNGTVMETSSDKTVHELPAEQV
OSY_SE_pbtC	1252	RHAKHSGRELEAVAHPEEARLASLEMTTEEESEQIEHYVNDTEAGRSQQTVPPELFEBOV
A6_61_pbtC	1315	ERTPDAAVAVVSGSQLSYGLNKKANQLAMKLRKYCVTAQOPVGLIVERTLDTVVAVAV
TS_15_pbtC	1315	ERTPDAAVAVVSGSQLSYGLNKKANQLAMKLRKYCVTAQOPVGLIVERTLDTVVAVAV
OSY_SE_pbtC	1312	ERTPDAAVAVVSGSQLSYGLNKKANQLAMKLRKYCVTAQOPVGLIVERTLDTVVAVAV
A6_61_pbtC	1375	LKASGTFVPIDPEYPETRIRYMLADSGAKLVLAQSELPGLIISDDVRLIDVRDZSLYQGDG
TS_15_pbtC	1375	LKASGTFVPIDPEYPETRIRYMLADSGAKLVLAQSELPGLIISDDVRLIDVRDZSLYQGDG
OSY_SE_pbtC	1372	LKSSGAVPIDPQYPEEDRIHNMHDSNAQVLLAQRHFCQARAAPSGRRIMEDEEAFYGNBG
A6_61_pbtC	1435	ADVPNGSKPSNLLYIYTSGTTGNPKGVMLERHNMVLLHYQOQCTNIEMPS-RILQYAS
TS_15_pbtC	1435	ADVPNGSKPSNLLYIYTSGTTGNPKGVMLERHNMVLLHYQOQCTNIEMPS-RILQYAS
OSY_SE_pbtC	1432	SNREYVQPEEMSYIYTSGTTGNPKGVMLERHNMVLLHYQOQCTNIEMPS-RILQYAS
A6_61_pbtC	1494	GSPDVICYQEMFSALLPGSSLYMVDNEMKDPVRLPQSEIKKHEIDVNYIP----
TS_15_pbtC	1494	GSPDVICYQEMFSALLPGSSLYMVDNEMKDPVRLPQSEIKKHEIDVNYIP----
OSY_SE_pbtC	1492	FSPDVFSGLIRALLHGGELILCPEDANANPAEIVELIRKRIQMFDPVPSLVPLMEIV
A6_61_pbtC	1550	PAEPENWAEAPRCVRHIITAGEQLVVTQVQACLRKEL--DLCHNNHGPCSETHVVTYTM
TS_15_pbtC	1550	PAEPENWAEAPRCVRHIITAGEQLVVTQVQACLRKEL--DLCHNNHGPCSETHVVTYTM
OSY_SE_pbtC	1552	YENK-----LDISSMKLAVVGAADHCPKEEFQKLEERFGSQRRIVNSYGVETTTIDSCNFE

FIG. 18D

A6_61_pbtC	1608	TPEVIEVGLPPICKPIANTSIYIVNDSPFLOPIGVKGBLYVSGACVGRGYNGRTDLDLTERK
TS_15_pbtC	1608	MEVIEVGLPPICKPIANTSIIVDDSPFLOPIGVKGBLYVSGACVGRGYNGRTDLDLTERK
OSY_SE_pbtC	1607	QASTEGRTVPICKPFGVTYIYIDQHSLEPFGTCHLYGCFVGRGYNGRTDLDLTERK
A6_61_pbtC	1668	FLDNPFAPGERLYKTGDVARWLPDGSIEYVGRSDHQVKIRGFRIELGEVESQLLHVPAVQ
TS_15_pbtC	1668	FLDNPFAPGERLYKTGDVARWLPDGSIEYVGRSDHQVKIRGFRIELGEVESQLLHVPAVQ
OSY_SE_pbtC	1667	FNDNPFAPGERLYTGDVARWLPDGSIEYVGRSDHQVKIRGFRIELGEVESQLLRTTFFER
A6_61_pbtC	1728	EATVVALEDHAGOKOLCAIFPAERSLTAGELRCALSOELPGYIMIPSYFVQLERLPLTPNG
TS_15_pbtC	1728	EATVVALEDHAGOKOLCAIFPAERSLTAGELRCALSOELPGYIMIPSYFVQLERLPLTPNG
OSY_SE_pbtC	1727	EATVVALEDHAGOKOLCAIFPAERSLTAGELRCALSOELPGYIMIPSYFVQLERLPLTPNG
A6_61_pbtC	1788	KIDRRALPKPEGCIETGTEYVAPRTTTEARLARIQDVGLGPSVGKDNPFELGCHSLRA
TS_15_pbtC	1788	KIDRRALPKPEGCIETGTEYVAPRTTTEARLARIQDVGLGPSVGKDNPFELGCHSLRA
OSY_SE_pbtC	1787	KIDRRALPKPEGCIETGTEYVAPRTTTEARLARIQDVGLGPSVGKDNPFELGCHSLRA
A6_61_pbtC	1848	TTLVSRLYKEMNVNPLGCVPRHPTIEMAKAITEMHQELYTEIPIAEEKAYVPLSSAQK
TS_15_pbtC	1848	TTLVSRLYKEMNVNPLGCVPRHPTIEMAKAITEMHQELYTEIPIAEEKAYVPLSSAQK
OSY_SE_pbtC	1847	TTLVSRLYKEMNVNPLGCVPRHPTIEMAKAITEMHQELYTEIPIAEEKAYVPLSSAQK
A6_61_pbtC	1908	RLPIVSQLTGAESVYNMPCVLLILEGELDRAPERAFQKLLIDRHESLRTSFETVRGEPVQR
TS_15_pbtC	1908	RLPIVSQLTGAESVYNMPCVLLILEGELDRAPERAFQKLLIDRHESLRTSFETVRGEPVQR
OSY_SE_pbtC	1907	RLPIVSQLTGAESVYNMPCVLLILEGELDRAPERAFQKLLIDRHESLRTSFETVRGEPVQR
A6_61_pbtC	1968	IHSQVEPAIEYHLAAEQDAEALITHFVRPFQKQAPLLRVGLIETGHERHILMFDMHHII
TS_15_pbtC	1968	IHSQVEPAIEYHLAAEQDAEALITHFVRPFQKQAPLLRVGLIETGHERHILMFDMHHII
OSY_SE_pbtC	1967	IHSQVEPAIEYHLAAEQDAEALITHFVRPFQKQAPLLRVGLIETGHERHILMFDMHHII
A6_61_pbtC	2028	SDGVTAGHVVNFEFSRIYAGDQLPALRIQYKDYAVWQOSNEYAEKLAHOESYWLKQDQEL
TS_15_pbtC	2028	SDGVTAGHVVNFEFSRIYAGDQLPALRIQYKDYAVWQOSNEYAEKLAHOESYWLKQDQEL
OSY_SE_pbtC	2027	SDGVTAGHVVNFEFSRIYAGDQLPALRIQYKDYAVWQOSNEYAEKLAHOESYWLKQDQEL
A6_61_pbtC	2088	PTLELPPTDIYVRPAVQCFEGDVALLFTLTNSQAEQORLAANYGATLYMVLAAAYTVLLHKY
TS_15_pbtC	2088	PTLELPPTDIYVRPAVQCFEGDVALLFTLTNSQAEQORLAANYGATLYMVLAAAYTVLLHKY
OSY_SE_pbtC	2087	PTLELPPTDIYVRPAVQCFEGDVALLFTLTNSQAEQORLAANYGATLYMVLAAAYTVLLHKY

46/48

FIG.18E

A6	61	PbTc	2148	TGQDDII	VGTPI	ACRNHT	TEPLV	GMFVNT	TLAIR	NYPTG	EKSPA	ELLAE	VKDTA	LAA	APEN
TS	15	PbTc	2148	TGQDDII	VGTPI	ACRNHT	TEPLV	GMFVNT	TLAIR	NYPTG	EKSPA	ELLAE	VKDTA	LAA	APEN
OSY	SE	PbTc	2147	TGQDDII	VGTPI	ACRNHT	TEPLV	GMFVNT	TLAIR	NYPTG	EKSPA	ELLAE	VKDTA	LAA	APEN
A6	61	PbTc	2208	QDYPFET	LVEK	VHKSR	DMSRNP	VFDTIF	SVERHE	QOSSP	HIDG	LIRIS	PYPHSH	SVAK	FDLT
TS	15	PbTc	2208	QDYPFET	LVEK	VHKSR	DMSRNP	VFDTIF	SVERHE	QOSSP	HIDG	LIRIS	PYPHSH	SVAK	FDLT
OSY	SE	PbTc	2207	QDYPFET	LVEK	VHKSR	DMSRNP	VFDTIF	SVERHE	QOSSP	HIDG	LIRIS	PYPHSH	SVAK	FDLT
A6	61	PbTc	2268	PHAEQNE	EGILC	GLGYA	TALYAK	ETACRM	GEHFV	QLIDAI	IAEPN	AKIM	SLNMS	LOERE	
TS	15	PbTc	2268	PHAEQNE	EGILC	GLGYA	TALYAK	ETACRM	GEHFV	QLIDAI	IAEPN	AKIM	SLNMS	LOERE	
OSY	SE	PbTc	2267	PHAEQNE	EGILC	GLGYA	TALYAK	ETACRM	GEHFV	QLIDAI	IAEPN	AKIM	SLNMS	LOERE	
A6	61	PbTc	2328	OVKLV	PNDTIT	SYPR	EKTIO	HLFEE	QAKSP	DAVAV	QPC	EG	LT	YREL	SERNLARTLR
TS	15	PbTc	2328	OVKLV	PNDTIT	SYPR	EKTIO	HLFEE	QAKSP	DAVAV	QPC	EG	LT	YREL	SERNLARTLR
OSY	SE	PbTc	2327	OVKLV	PNDTIT	SYPR	EKTIO	HLFEE	QAKSP	DAVAV	QPC	EG	LT	YREL	SERNLARTLR
A6	61	PbTc	2388	GKGVK	AGRCV	GLMT	DRSL	DMIVAI	MA	TKAG	AYVP	IDP	DYPER	IR	YMNIDDSG
TS	15	PbTc	2388	GKGVK	AGRCV	GLMT	DRSL	DMIVAI	MA	TKAG	AYVP	IDP	DYPER	IR	YMNIDDSG
OSY	SE	PbTc	2387	GKGVK	AGRCV	GLMT	DRSL	DMIVAI	MA	TKAG	AYVP	IDP	DYPER	IR	YMNIDDSG
A6	61	PbTc	2448	QRHLQ	ANH	IPAD	CMV	VLVD	DECSY	HADG	TNLEQ	HNCAS	DLAY	VI	YTSCTT
TS	15	PbTc	2448	QRHLQ	ANH	IPAD	CMV	VLVD	DECSY	HADG	TNLEQ	HNCAS	DLAY	VI	YTSCTT
OSY	SE	PbTc	2447	QRHLQ	ANH	IPAD	CMV	VLVD	DECSY	HADG	TNLEQ	HNCAS	DLAY	VI	YTSCTT
A6	61	PbTc	2508	RNI	VRV	RD	AKYIE	IDO	BDT	VLQ	LS	YAF	DG	ST	FDIF
TS	15	PbTc	2508	RNI	VRV	RD	AKYIE	IDO	BDT	VLQ	LS	YAF	DG	ST	FDIF
OSY	SE	PbTc	2506	RNI	VRV	RD	AKYIE	IDO	BDT	VLQ	LS	YAF	DG	ST	FDIF
A6	61	PbTc	2568	ADTIE	SE	KIS	VMP	IT	TAY	PN	LL	V	D	LR	V
TS	15	PbTc	2568	ADTIE	SE	KIS	VMP	IT	TAY	PN	LL	V	D	LR	V
OSY	SE	PbTc	2566	ADTIE	SE	KIS	VMP	IT	TAY	PN	LL	V	D	LR	V

FIG.18F

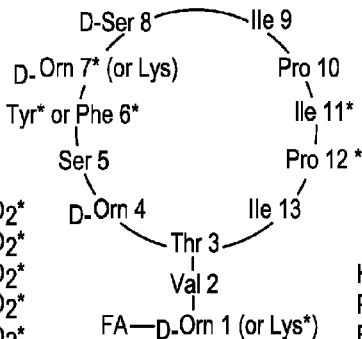
AS	61	pbtC	2628	KHVI	GPT	STVF	FAT	SHNV	DEV	AD	SA	VT	PI	GR	PI	GN	T	AV	YV	IG	GV	LQ	PI	GV	AG	EL	CV	A						
TS	15	pbtC	2628	KHVI	GPT	STVF	FAT	SHNV	DEV	AD	SA	VT	PI	GR	PI	GN	T	AV	YV	IG	GV	LQ	PI	GV	AG	EL	CV	A						
OSY	SE	pbtC	2626	KHVI	GPT	STVF	FAT	CHNV	EV	AD	SA	VT	PI	GR	PI	GN	T	AV	YV	NA	QD	HQ	PI	GV	AG	EL	CV	A						
AS	61	pbtC	2688	GDG	VA	AG	YL	NR	PE	LS	GA	KF	VN	NP	FP	GD	RM	YK	TG	DL	AR	WL	SD	ST	IE	IV	GR	DD	QV	KIR	GY			
TS	15	pbtC	2688	GDG	VA	AG	YL	NR	PE	LS	GA	KF	VN	NP	FP	GD	RM	YK	TG	DL	AR	WL	SD	ST	IE	IV	GR	DD	QV	KIR	GY			
OSY	SE	pbtC	2686	GDG	VA	AG	YL	NR	PE	LS	GA	KF	VN	NP	FP	GD	RM	YK	TG	DL	AR	WL	SD	ST	IE	IV	GR	DD	QV	KIR	GY			
AS	61	pbtC	2748	RI	EL	GE	VE	AH	LL	D	LE	AI	Q	E	AV	V	I	V	R	E	S	D	G	Q	R	L	C	A	Y	V	A			
TS	15	pbtC	2748	RI	EL	GE	VE	AH	LL	D	LE	AI	Q	E	AV	V	I	V	R	E	S	D	G	Q	R	L	C	A	Y	V	A			
OSY	SE	pbtC	2746	RI	EL	GE	VE	TH	L	R	V	E	I	Q	E	AV	V	I	R	E	S	D	S	G	E	R	L	C	A	Y	V	A		
AS	61	pbtC	2808	ML	PS	Y	V	Q	L	D	K	L	P	L	S	P	N	G	K	V	N	R	K	A	L	P	A	P	E	L	H	V	Q	A
TS	15	pbtC	2808	ML	PS	Y	V	Q	L	D	K	L	P	L	S	P	N	G	K	V	N	R	K	A	L	P	A	P	E	L	H	V	Q	A
OSY	SE	pbtC	2806	ML	PS	Y	V	Q	L	D	K	L	P	L	S	P	N	G	K	V	N	R	K	A	L	P	A	P	E	L	H	V	Q	A
AS	61	pbtC	2868	QV	G	V	K	D	N	F	E	L	G	H	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	
TS	15	pbtC	2868	QV	G	V	K	D	N	F	E	L	G	H	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	
OSY	SE	pbtC	2866	QV	G	V	K	D	N	F	E	L	G	H	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S	
AS	61	pbtC	2928	GN	SP	M	K	L	N	E	NC	H	I	P	V	F	C	F	P	P	G	L	G	S	L	E	L	A	K	O	L	D	H	C
TS	15	pbtC	2928	GN	SP	M	K	L	N	E	NC	H	I	P	V	F	C	F	P	P	G	L	G	S	L	E	L	A	K	O	L	D	H	C
OSY	SE	pbtC	2926	SS	Q	T	L	N	E	NC	L	N	V	F	C	F	P	P	G	L	G	S	L	E	L	A	K	O	L	D	H	C	L	N
AS	61	pbtC	2988	VN	AI	L	A	V	Q	P	P	P	L	L	G	S	L	G	G	N	L	T	F	E	V	A	K	A	L	E	C	R	G	Y
TS	15	pbtC	2988	VN	AI	L	A	V	Q	P	P	P	L	L	G	S	L	G	G	N	L	T	F	E	V	A	K	A	L	E	C	R	G	Y
OSY	SE	pbtC	2986	VD	AI	L	A	V	Q	P	P	P	L	L	G	S	L	G	G	N	L	T	F	E	V	A	K	A	L	E	C	R	G	Y

FIG.18G

A6_6i_PbC	3048	GDIDQMIDGVEELKEMLVHHPLELDQVKNEMRAYWSYATLVNSDIIDANIHALMAEPSE
TS_15_PbC	3048	GDIDQMIDGVEELKEMLVHHPLELDQVKNEMRAYWSYATLVNSDIIDANIHALMAEPSE
OSY_SE_PbC	3046	SDIDQMEAVGCEQEMQLSENPLETRERVKHKKRAYVTGCGQLVNTGAVEDANHYALHAEEDSE
A6_6i_PbC	3108	VNQNGEGAARDMAGCGYSRQVSGVQPTRCAKESASTAPFGSER-----
TS_15_PbC	3108	VNQADGGEQLATNQEATGSEVVEYMTREGVHEDVLQPPPLEANANNVQCNTIEHHDEQDMVTH
OSY_SE_PbC	3106	AVKFPENVTSALMDGATRDQAYCEHRLIGVHEDVLDPCEFTIEHNVKVIFHVVVHCHIEQHRGVH
A6_6i_PbC	---	---
TS_15_PbC	3168	-----
OSY_SE_PbC	3166	EVLSE
A6 6i PbC is SEQ ID NO. 13		
TS-15 PbC is SEQ ID NO. 16		
OSY SE PbC is SEQ ID NO. 19		

Compound of the present invention SEQ ID NO. 4

Compounds Identified in this Work



FA options

C₁₀H₁₉O* C₁₀H₁₉O₂*
C₁₁H₂₁O* C₁₁H₂₁O₂*
C₁₂H₂₃O* C₁₂H₂₃O₂*
C₁₃H₂₅O* C₁₃H₂₅O₂*
C₁₄H₂₇O* C₁₄H₂₇O₂*
C₁₅H₂₉O C₁₅H₂₉O₂*

Hydrophobic
Polar, uncharged
Basic, + charged