

(51) International Patent Classification:
A61K 31/351 (2006.01)(21) International Application Number:
PCT/US2014/019526(22) International Filing Date:
28 February 2014 (28.02.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/771,322 1 March 2013 (01.03.2013) US(71) Applicant (for all designated States except US): **LOS ALAMOS NATIONAL SECURITY, LLC** [US/US]; Los Alamos National Laboratory, P.O. Box 1663, MS A187, Los Alamos, NM 87545 (US).(72) Inventors: **KOGLIN, Alexander**; 111 Paseo Nopal, Santa Fe, NM 87507 (US). **STRIEKER, Matthias**; 1608 Ben Hur Drive, Santa Fe, NM 87501 (US).(74) Agent: **CONNOLLY, Jodi, L.**; Klarquist Sparkman, LLP, One World Trade Center, Suite 1600, 121 SW Salmon Street, Portland, OR 97204 (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, LT, 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, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, 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:

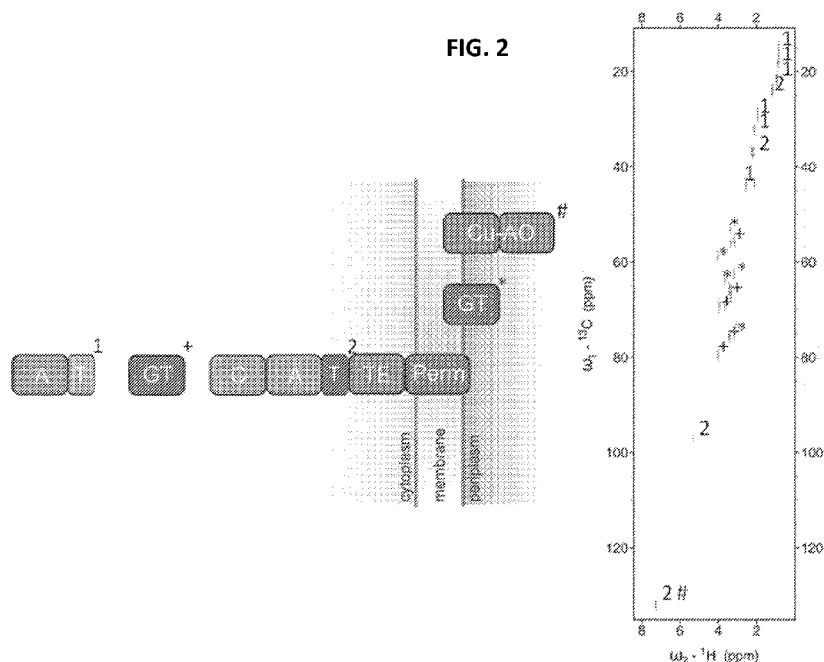
— of inventorship (Rule 4.17(iv))

Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: BROAD SPECTRUM ANTIBIOTIC COMPOUNDS AND USE THEREOF



(57) **Abstract:** Described herein is the discovery of a non-ribosomal peptide synthetase (NRPS) gene cluster in the genome of *Clostridium thermocellum* that produces a secondary metabolite that is assembled outside of the host membrane. Also described is the identification of homologous NRPS gene clusters several additional microorganisms. The secondary metabolites produced by the NRPS gene clusters exhibit broad spectrum antibiotic activity. Thus, provided herein are antibiotic compounds produced by the NRPS gene clusters, and analogs thereof, their use for inhibiting bacterial growth, and methods of making the antibiotic compounds.

BROAD SPECTRUM ANTIBIOTIC COMPOUNDS AND USE THEREOF

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 61/771,322, filed
5 March 1, 2013, which is herein incorporated by reference in its entirety.

FIELD

This disclosure concerns antibiotic compounds produced by non-ribosomal peptide
synthetase gene clusters found in microbial organisms, the use of such compounds for inhibiting
10 microbial growth and treating microbial infections, and methods of making the antibiotic
compounds.

ACKNOWLEDGMENT OF GOVERNMENT SUPPORT

This invention was made with government support under Contract No. DE-AC52-
15 06NA25396 awarded by the U.S. Department of Energy. The government has certain rights in the
invention.

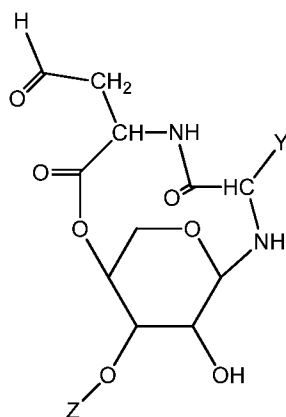
BACKGROUND

Antibiotics transformed medicine dramatically by providing efficient treatments against
20 bacterial infections causing severe illnesses and death. However, the emergence of antibiotic
resistance is significantly reducing the efficacy of antibiotic treatments and bacterial infections
have become a major public health threat again. Gram-negative bacteria like *Acinetobacter*
baumannii and *Pseudomonas aeruginosa*, and gram-positive bacteria such as *Staphylococcus*
aureus or *Clostridium difficile*, are rapidly evolving resistance against multiple antibiotics. The
25 endemic appearance of bacterial pathogens in hospitals, health care facilities and water treatment
plants is causing an additional threat.

Until 15-20 years ago, new antibiotics were developed in time to counter emerging resistant
bacteria. However, this pipeline of new antibiotics has dried up as the development of new
antibodies is no longer a focus for pharmaceutical companies. For example, *A. baumannii*, *S.*
30 *aureus* and *M. tuberculosis* are causing untreatable infections due to the emergence of resistance
even against antibiotics of last resort. While the emergence of drug resistance among natural
bacteria is of immense concern, man-made pathogens pose an additional threat. New antibiotics,
particularly those that can treat infections with multidrug resistant bacteria, are therefore in dire
need.

SUMMARY

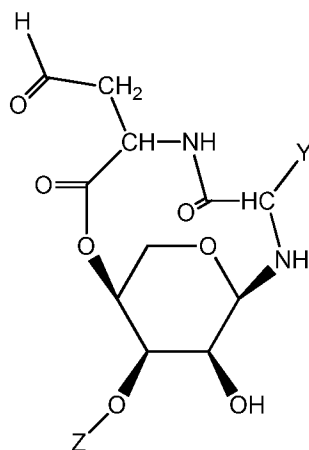
According to one embodiment disclosed herein, there is provided an antibiotic compound, or a pharmaceutically acceptable salt, ester, hydrate or solvate thereof, comprising Formula I:



5

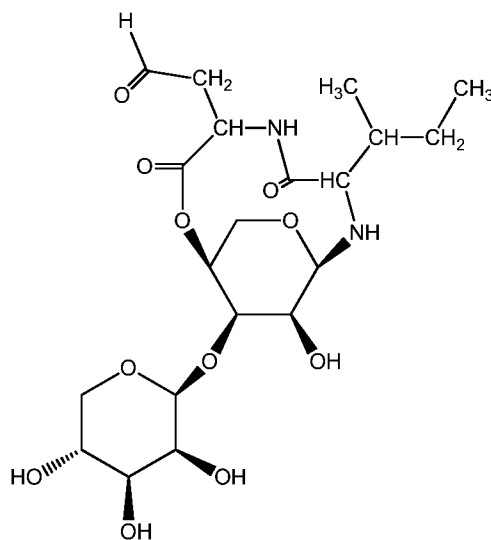
where Y is C₁-C₁₀ alkyl, and Z is hydrogen, -OH, or a monosaccharide.

In some embodiments, the antibiotic compound has the stereochemistry of Formula II:



10

In one non-limiting example, the antibiotic compound (referred to herein as Compound #1) comprises the structure:



Further provided is a method of inhibiting microbial growth, such as bacterial growth, by contacting the microorganisms with an antibiotic compound disclosed herein. In some embodiments, the method of inhibiting microbial growth is an *in vitro* method. In other embodiments, the method is an *in vivo* method that includes administering the antibiotic compound to a subject with a microbial, such as a bacterial, infection.

Also provided herein is a method of treating a microbial infection (such as a bacterial infection) in a subject by selecting a subject in need of treatment and administering to the subject a therapeutically effective amount of an antibiotic compound disclosed herein.

Methods of making the disclosed antibiotic compounds are also provided by the present disclosure.

The foregoing and other objects, features, and advantages of the invention will become more apparent from the following detailed description, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic of the non-ribosomal peptide synthetase (NRPS) gene cluster identified in *Clostridium thermocellum*. The genes identified, and the functions of the encoded proteins, are shown. C = condensation domain; A = adenylation domain; T = thiolation domain; TE = termination domain; Perm = maltose-permease homologue transporter.

FIG. 2 shows a schematic of the proteins identified in the NRPS cluster and their assignments in a ^{13}C - ^1H HSQC NMR spectra recorded of the purified compound. C = condensation domain; A = adenylation domain; T = thiolation domain; TE = termination domain; Perm = maltose-permease homologue transporter; GT = glycosyltransferase; Cu-AO = copper-containing amine oxidase.

FIG. 3 shows the results of radio-high pressure liquid chromatography (HPLC) of a *C. thermocellum* culture containing ^{14}C -labelled amino acids.

FIG. 4 is an image showing the results of thin layer chromatography (TLC) of a DMSO extract of the *C. thermocellum* secondary metabolite (Compound #1).

FIG. 5 shows the results of a disk diffusion assay to test the antibiotic properties of the DMSO extract of the *C. thermocellum* secondary metabolite. The DMSO extract was tested against *E. coli* and *B. subtilis* indicator strains.

FIG. 6 is a schematic of a NRPS gene cluster identified in *Herpetosiphon aurantiacus*.

SEQUENCE LISTING

The amino acid sequences listed in the accompanying Sequence Listing are shown using standard three letter code for amino acids, as defined in 37 C.F.R. 1.822. The Sequence Listing is submitted as an ASCII text file, created on February 20, 2014, 86.9 KB, which is incorporated by reference herein. In the accompanying Sequence Listing:

SEQ ID NO: 1 is the amino acid sequence of the amino acid adenylation domain from *Clostridium thermocellum* strain DSM4150.

SEQ ID NO: 2 is the amino acid sequence of the amino acid adenylation domain from *Herpetosiphon aurantiacus* strain DSM 785.

SEQ ID NO: 3 is the amino acid sequence of the amino acid adenylation enzyme/thioester reductase family protein from *Streptosporangium roseum* strain DSM 43021.

SEQ ID NO: 4 is the amino acid sequence of the amino acid adenylation enzyme/thioester reductase family protein from *Catenulispora acidiphila* strain DSM 44928.

SEQ ID NO: 5 is the amino acid sequence of the sugar transporter protein from *Pyrobaculum aerophilum*.

SEQ ID NO: 6 is the amino acid sequence of a putative NRPS related protein from *Emericella nidulans*.

DETAILED DESCRIPTION

I. Abbreviations

	CIP	ciprofloxacin
	DMSO	dimethylsulfoxide
5	HPLC	high pressure liquid chromatography
	MIC	minimum inhibitory concentration
	MTD	maximum tolerated dose
	NMR	nuclear magnetic resonance
	NRPS	non-ribosomal peptide synthetase
10	TLC	thin layer chromatography
	VAN	vancomycin

II. Terms and Methods

Unless otherwise noted, technical terms are used according to conventional usage.

15 Definitions of common terms in molecular biology may be found in Benjamin Lewin, *Genes V*, published by Oxford University Press, 1994 (ISBN 0-19-854287-9); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

20 In order to facilitate review of the various embodiments of the disclosure, the following explanations of specific terms are provided:

Administer: As used herein, administering a composition (*e.g.* an antibiotic compound) to a subject means to give, apply or bring the composition into contact with the subject.

Administration can be accomplished by any of a number of routes, such as, for example, topical,
25 oral, subcutaneous, intramuscular, intraperitoneal, intravenous, intrathecal and intramuscular.

Alkyl: A hydrocarbon group having a saturated carbon chain. The chain may be cyclic, branched or unbranched. The chain may unsubstituted, or substituted with one or more substituents. Exemplary substituents include, for example and without limitation, hydroxyl, amine, amide, sulfonamide, halo, cyano, carboxy, mercapto, alkyl, alkoxy, alkylthio, thioalkoxy, arylalkyl,
30 heteroaryl, alkylamino, oxy, and dialkylamino.

Amino acid: Amino acid refers to both natural and unnatural amino acids, including their D and L stereoisomers for chiral amino acids. Natural and unnatural amino acids are well known to those of ordinary skill in the art. Common natural amino acids include, without limitation, alanine (Ala), arginine (Arg), asparagine (Asn), aspartic acid (Asp), cysteine (Cys), glutamine (Gln),

glutamic acid (Glu), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tryptophan (Trp), tyrosine (Tyr), and valine (Val). Uncommon and unnatural amino acids include, without limitation, allyl glycine (AllylGly), biphenylalanine (Bip), citrulline (Cit), 4-

5 guanidinophenylalanine (Phe(Gu)), homoarginine (hArg), homolysine (hLys), 2-naphthylalanine (2-Nal), ornithine (Orn) and pentafluorophenylalanine (F5Phe).

Amino acids are typically classified in one or more categories, including polar, hydrophobic, acidic, basic and aromatic, according to their side chains. Examples of polar amino acids include those having side chain functional groups such as hydroxyl, sulfhydryl, and amide, as well as the acidic and basic amino acids. Polar amino acids include, without limitation, asparagine, 10 cysteine, glutamine, histidine, selenocysteine, serine, threonine, tryptophan and tyrosine. Examples of hydrophobic or non-polar amino acids include those residues having nonpolar aliphatic side chains, such as, without limitation, leucine, isoleucine, valine, glycine, alanine, proline, methionine and phenylalanine. Examples of basic amino acid residues include those having a basic side chain, 15 such as an amino or guanidino group. Basic amino acid residues include, without limitation, arginine, homolysine and lysine. Examples of acidic amino acid residues include those having an acidic side chain functional group, such as a carboxy group. Acidic amino acid residues include, without limitation aspartic acid and glutamic acid. Aromatic amino acids include those having an aromatic side chain group. Examples of aromatic amino acids include, without limitation, 20 biphenylalanine, histidine, 2-naphthylalanine, pentafluorophenylalanine, phenylalanine, tryptophan and tyrosine. It is noted that some amino acids are classified in more than one group, for example, histidine, tryptophan and tyrosine are classified as both polar and aromatic amino acids. Additional amino acids that are classified in each of the above groups are known to those of ordinary skill in the art.

25 **Analog, derivative or mimetic:** An analog is a molecule that differs in chemical structure from a parent compound, for example a homolog (differing by an increment in the chemical structure, such as a difference in the length of an alkyl chain), a molecular fragment, a structure that differs by one or more functional groups, and/or a change in ionization. Structural analogs are often found using quantitative structure activity relationships (QSAR), with techniques such as those disclosed in Remington (*The Science and Practice of Pharmacology*, 19th Edition (1995), 30 chapter 28). When the changes to the original compound are substantial, or many incremental changes are combined, the compound is no longer an analog. A derivative is a biologically active molecule derived from the base structure. A mimetic is a molecule that mimics the activity of

another molecule by mimicking the structure of such a molecule, such as a biologically active molecule. Thus, the term “mimetic” indicates a definite structure related to activity.

Antibiotic: A substance often produced by or derived from certain fungi, bacteria, and other organisms, that can destroy or inhibit the growth of other microorganisms. Antibiotics can also be synthetically produced.

***Aspergillus nidulans*:** A type of filamentous fungi in the phylum Ascomycota. It is one of the few species in its genus capable of forming sexual spores through meiosis. *A. nidulans* is also called *Emmericella nidulans* when referring to its sexual form.

***Catenulispora acidiphila*:** A Gram-positive, filamentous, acidophilic bacterium first isolated from forest soil.

***Clostridium thermocellum*:** An anaerobic, thermophilic bacterium capable of converting cellulosic substrates into ethanol. The complete genome sequence of *C. thermocellum* (ATCC 27405) is available in GENBANK™ under NCBI reference number NC_009012, which is herein incorporated by reference as it appears in the database on December 8, 2011.

Condensation: A reaction in which two molecules combine to form a single molecule, accompanied by the loss of a small molecule, *e.g.*, water, methanol, acetic acid, or a hydrogen halide.

Contacting: Placement in direct physical association; includes both in solid and liquid form.

Coupling: Joining a first unit to a second unit. This includes, but is not limited to, covalently bonding one molecule to another molecule, noncovalently bonding one molecule to another (*e.g.*, electrostatically bonding), non-covalently bonding one molecule to another molecule by hydrogen bonding, non-covalently bonding one molecule to another molecule by van der Waals forces, and any and all combinations of such couplings.

Gene cluster: A set of genetic elements grouped together on the chromosome, the protein products of which have a related function, such as forming a natural product biosynthetic pathway.

Gram-positive bacteria: Bacteria having a cell wall composed of a thick layer of peptidoglycan. Gram-positive bacteria retain the color of the crystal violet stain in the Gram method of staining. Exemplary Gram-positive bacteria include *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Bacillus subtilis*, and *Clostridium difficile*.

Gram-negative bacteria: Bacteria having a cell wall composed of a thin layer of peptidoglycan. Gram-negative bacteria lose the crystal violet stain and take the color of the red counterstain in Gram's Method of staining. Gram-negative bacteria include most of the bacteria normally found in the gastrointestinal tract. Exemplary Gram-negative bacteria include

Pseudomonas aeruginosa, *Salmonella enterica*, *Pseudomonas putida*, *Escherichia coli*, *Acinetobacter baumannii* and *Haemophilus influenzae*.

***Herpetosiphon aurantiacus*:** A filamentous, Gram-negative bacterium that grows in long, flexible filaments. This bacterium was first isolated from Lake Birch in Minnesota in 1968.

5 **Homologous amino acid sequence:** An amino acid sequence that differs from an amino acid sequence by one or more conservative amino acid substitutions. Homologous sequences also encompass allelic variants as well as sequences containing deletions or insertions which retain the functional characteristics of the polypeptide. In some instances, such a sequence is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 98% identical to any one of the amino acid sequences disclosed herein.

10 Homologous amino acid sequences include sequences that are identical or substantially identical to the amino acid sequences set forth in the sequence listing. By “substantially identical to the amino acid sequence” it is meant a sequence that is at least 90%, at least 95%, at least 97%, or at least 99% identical to an amino acid sequence of reference. In an example, the sequence is at least 90% identical and differs from the sequence of reference by conservative amino acid substitutions. Polypeptides having a sequence homologous to any one of the amino acid sequences disclosed herein include naturally-occurring allelic variants, as well as mutants or any other non-naturally occurring variants that retain the inherent characteristics (*e.g.*, biosynthetic activity) of any polypeptide providing herein. Homology can be measured using sequence analysis software such as Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705. Amino acid sequences can be aligned to maximize identity. Gaps can also be artificially introduced into the sequence to attain optimal alignment. Once the optimal alignment has been set up, the degree of homology is established by recording all of the positions in which the amino acids of both sequences are identical, relative to the total number of positions. Homologous polynucleotide sequences are defined in a similar way. In some embodiments, a homologous sequence is one that is at least 45%, 50%, 60%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or identical to any one of the sequences disclosed herein.

20 **Inhibiting bacterial growth:** Reducing or eliminating bacterial replication and/or spread, *in vitro* or *in vivo*. Inhibition need not be complete inhibition. In some embodiments, inhibition of bacterial growth is inhibition of at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70% or at least 80%. Similarly, **inhibiting microbial growth** refers to reducing or eliminating microbial (such as bacterial or fungal) replication and/or spread *in vitro* or *in vivo*.

Isolated: An isolated biological component (such as a nucleic acid molecule, organic compound or protein) is one that has been substantially separated or purified away from other biological components in the cell of the organism in which the component naturally occurs, such as other chromosomal and extra-chromosomal DNA and RNA, proteins and organelles. With respect to nucleic acids and/or polypeptides, the term can refer to nucleic acids or polypeptides that are no longer flanked by the sequences typically flanking them in nature. Nucleic acids and proteins that have been isolated include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids and proteins.

Mitsunobu reaction: An organic reaction that allows the conversion of primary and secondary alcohols to amines, esters, phenyl ethers, thioethers, and various other compounds. Conversion of an alcohol to an amine typically includes reaction of the alcohol with triphenylphosphine, diethyl azodicarboxylate (DEAD), and a nitrogen nucleophile (*e.g.*, phthalimide), followed by subsequent hydrolysis or selective reduction.

Monosaccharide: A simple, monomeric sugar. With few exceptions, a monosaccharide has the basic chemical formula $C_x(H_2O)_y$, where x and y are integers, and x is at least 3. Typically, $y = x$ or $y = x - 1$. Many monosaccharides are pentoses ($x = 5$) or hexoses ($x = 6$). Examples of monosaccharides include arabinose, fructose, galactose, glucose, ribose, and xylose, among others. A **disaccharide** is a dimer formed by two monosaccharides linked together by a glycosidic bond.

Disaccharides include both hetero-dimers and homo-dimers.

***Myxococcus xanthus*:** A rod-shaped, Gram-negative bacterium that grows in colonies that exhibit self-organizing behavior.

Nonribosomal peptide synthetase (NRPS): A large multi-functional protein that synthesizes polypeptides by a nonribosomal mechanism, often known as thiotemplate synthesis (Kleinkauf and von Doehren *Ann. Rev. Microbiol.* 41: 259-289, 1987). Such nonribosomal polypeptides can have a linear, cyclic, or branched cyclic structure and often contain amino acids not present in proteins or amino acids modified through methylation or epimerization. NRPSs are typically organized into modules. A “**module**” is a set of distinctive domains that encode all the enzyme activities necessary for one cycle of peptide chain elongation and associated modifications. The number and order of modules and the type of domains within a module on each NRPS protein determine the structural variations of the resulting peptide products by dictating the number, order, choice of the amino acid to be incorporated, and the modifications associated with a particular cycle of elongation. The modular architecture of NRPS (Cane *et al. Science* 282: 63-68, 1998, Stachelhaus *et al. Science* 269: 69-72, 1995; Stachelhaus *et al. Mol. Gen. Genet.* 257: 308-318,

1998; and Belshaw *et al. Science* 284:486-489, 1999) has been successfully used in combinatorial biosynthesis of diverse natural product analogs.

Typically, the enzyme organization of an NRPS gene cluster is such that the modules occur in the primary sequence in the same order that the amino acids are assembled into the peptide product (colinearity principle). There are three components to the minimum extending module. The adenylation (A) domain recognizes a specific amino acid and uses ATP to activate the carboxyl as the adenylate. This facilitates attachment of the carboxyl to the thiol of a 4'-phosphopantetheine (p-pant) cofactor attached to the peptidyl carrier (PCP, or thiolation (T)) domain. The third domain is the condensation (C) domain that catalyzes formation of peptide bonds between amino acids or peptides attached to PCP domains or adjacent modules, thereby promoting directional peptide chain elongation. The order of these domains in a typical module is C-A-T.

Oligomer: A general term for a polymeric molecule consisting of relatively few monomers, *e.g.*, ≤ 10 monomers.

Oligopeptide: An oligomer in which the monomers are amino acid residues that are joined together through amide bonds.

Pharmaceutically acceptable vehicles: The pharmaceutically acceptable carriers (vehicles) useful in this disclosure are conventional. *Remington's Pharmaceutical Sciences*, by E. W. Martin, Mack Publishing Co., Easton, PA, 15th Edition (1975), describes compositions and formulations suitable for pharmaceutical delivery of one or more therapeutic compositions, such as one or more antibiotic compounds.

In general, the nature of the carrier will depend on the particular mode of administration being employed. For instance, parenteral formulations usually comprise injectable fluids that include pharmaceutically and physiologically acceptable fluids such as water, physiological saline, balanced salt solutions, aqueous dextrose, glycerol or the like as a vehicle. For solid compositions (for example, powder, pill, tablet, or capsule forms), conventional non-toxic solid carriers can include, for example, pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. In addition to biologically-neutral carriers, pharmaceutical compositions to be administered can contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate.

Preventing, treating or ameliorating a disease: "Preventing" a disease (such as an infection) refers to inhibiting the full development of a disease. "Treating" refers to a therapeutic intervention that ameliorates a sign or symptom of a disease or pathological condition after it has

begun to develop. "Ameliorating" refers to the reduction in the number or severity of one or more signs or symptoms of a disease.

Protecting or protective group: When synthesizing organic compounds, often some specific functional groups cannot survive the required reagents or chemical environments. These groups must be protected. A protecting group, or protective group, is introduced into a molecule by chemical modification of a functional group in order to obtain chemoselectivity in a subsequent chemical reaction. Various exemplary protecting or protective groups are disclosed in Greene's Protective Groups in Organic Synthesis, by Peter G. M. Wuts and Theodora W. Greene (October 30, 2006), which is incorporated herein by reference.

Purified: The term **purified** does not require absolute purity; rather, it is intended as a relative term. Thus, for example, a purified protein preparation is one in which the protein referred to is more pure than the protein in its natural environment within a cell.

***Pyrobaculum aerophilum*:** A rod-shaped hyperthermophilic archaeobacterium with an optimal growth temperature of 100°C and pH of 7.0.

***Rickettsia conorii*:** A unicellular, gram-negative, obligate intracellular bacterium of the genus Rickettsia.

Secondary metabolite: An organic compound that is not directly involved in the normal growth, development or reproduction of an organism. Secondary metabolites often play a role in host defense (*e.g.*, antibiotics).

Sequence identity: The similarity between two nucleic acid sequences or between two amino acid sequences is expressed in terms of the level of sequence identity shared between the sequences. Sequence identity is typically expressed in terms of percentage identity; the higher the percentage, the more similar the two sequences.

Methods for aligning sequences for comparison are well known in the art. Various programs and alignment algorithms are described in: Smith and Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman and Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins and Sharp, *Gene* 73:237-244, 1988; Higgins and Sharp, *CABIOS* 5:151-153, 1989; Corpet *et al.*, *Nucleic Acids Research* 16:10881-10890, 1988; Huang, *et al.*, *Computer Applications in the Biosciences* 8:155-165, 1992; Pearson *et al.*, *Methods in Molecular Biology* 24:307-331, 1994; Tatiana *et al.*, (1999), *FEMS Microbiol. Lett.*, 174:247-250, 1999. Altschul *et al.* present a detailed consideration of sequence-alignment methods and homology calculations (*J. Mol. Biol.* 215:403-410, 1990).

The National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST™, Altschul *et al.*, *J. Mol. Biol.* 215:403-410, 1990) is available from several sources,

including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence-analysis programs blastp, blastn, blastx, tblastn and tblastx. A description of how to determine sequence identity using this program is available on the internet under the help section for BLAST™.

5 For comparisons of amino acid sequences of greater than about 30 amino acids, the “Blast 2 sequences” function of the BLAST™ (Blastp) program is employed using the default BLOSUM62 matrix set to default parameters (cost to open a gap [default = 5]; cost to extend a gap [default = 2]; penalty for a mismatch [default = -3]; reward for a match [default = 1]; expectation value (E) [default = 10.0]; word size [default = 3]; number of one-line descriptions (V) [default = 100];
10 number of alignments to show (B) [default = 100]). When aligning short peptides (fewer than around 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins (or nucleic acids) with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 50%, at least 60%, at
15 least 70%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity.

For comparisons of nucleic acid sequences, the “Blast 2 sequences” function of the BLAST™ (Blastn) program is employed using the default BLOSUM62 matrix set to default parameters (cost to open a gap [default = 11]; cost to extend a gap [default = 1]; expectation value (E) [default = 10.0]; word size [default = 11]; number of one-line descriptions (V) [default = 100];
20 number of alignments to show (B) [default = 100]). Nucleic acid sequences with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 98% sequence identity.

Nucleic acid sequences that do not show a high degree of identity can nevertheless encode
25 similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid molecules that all encode substantially the same protein.

***Streptosporangium roseum*:** An aerobic, Gram-positive bacterium first isolated from vegetable garden soil in 1955.

30 **Subject:** Living multi-cellular vertebrate organisms, a category that includes both human and non-human mammals.

Therapeutically effective amount: A quantity of a specified agent (such as an antibiotic compound) sufficient to achieve a desired effect in a subject being treated with that agent. For example, this may be the amount of an antibiotic compound useful for treating a bacterial infection

in a subject. In the context of the present disclosure, a therapeutically effective amount of an antibiotic compound, for example, is an amount sufficient to treat a microbial infection in a subject (e.g., reduce or eliminate infectious organisms from the subject) without causing a substantial cytotoxic effect in the subject. The effective amount of an antibiotic compound useful for treating a microbial infection in a subject will be dependent on, for example, the subject being treated, the type of infection being treated, the manner of administration of the therapeutic composition and other factors.

Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. The singular terms “a,” “an,” and “the” include plural referents unless context clearly indicates otherwise. Similarly, the word “or” is intended to include “and” unless the context clearly indicates otherwise. Hence “comprising A or B” means including A, or B, or A and B. It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including explanations of terms, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

III. Introduction

Disclosed herein is the discovery of a non-ribosomal peptide synthetase (NRPS) gene cluster that produces a bioactive compound that is assembled outside of its host membrane. The NRPS gene cluster was first identified in *Clostridium thermocellum*; however, several additional microorganisms were identified that encode an NRPS gene cluster. The metabolite was isolated from *Clostridium thermocellum*, and the structure was assigned by nuclear magnetic resonance spectroscopy and mass spectrometry. The compound (referred to herein as Compound #1) is a small, previously unknown glycopeptide that possesses broad-spectrum antibiotic activity against both Gram positive and Gram negative bacteria. Also described herein is the isolation of a second compound (Compound #2) produced by an NRPS gene cluster in *Herpetosiphon aurantiacus*. An antibiotic secondary metabolite has also been isolated from *Emericella nidulans* and is referred to herein as Compound #3.

IV. Overview of Several Embodiments

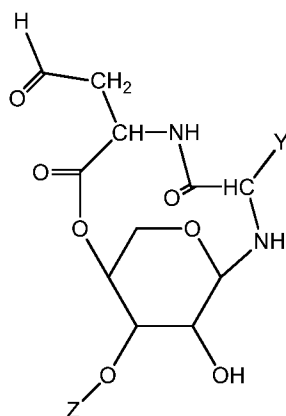
Disclosed herein is the identification of a NRPS gene cluster in *Clostridium thermocellum*, as well as homologous NRPS gene clusters in several additional microorganisms, including

5 *Herpetosiphon aurantiacus*, *Myxococcus xanthus*, *Streptosporangium roseum*, *Catenulispora acidiphila*, *Pyrobaculum aerophilum*, *Rickettsia conorii* and *Emericella nidulans*. The NRPS gene clusters produce a secondary metabolite with broad spectrum antibiotic activity.

Thus, provided herein are broad spectrum antibiotic compounds synthesized by NRPS gene clusters of microorganisms. Also provided are methods of making such compounds, either by

10 isolating the compounds from microbial cultures or by synthesizing the compounds. Methods of inhibiting bacterial growth *in vitro* and *in vivo* using the disclosed compounds are further provided by the present disclosure.

In some embodiments disclosed herein, provided is an antibiotic compound, or a pharmaceutically acceptable salt, ester, hydrate or solvate thereof, represented by the Formula I:

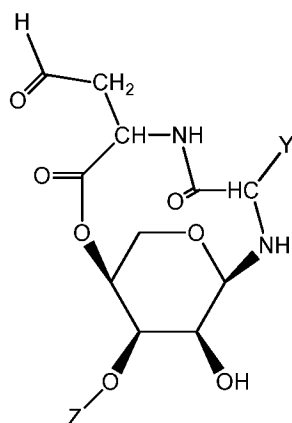


Formula I

where Y is C₁-C₁₀ alkyl, and Z is hydrogen, -OH, or a monosaccharide.

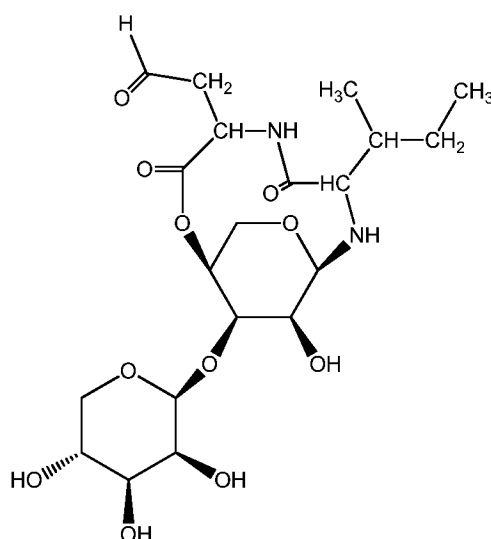
In some examples, Y is a branched C₄ alkyl, such as, but not limited to, -CH(CH₃)CH₂CH₃. In some examples, Z is a hexose, such as, but not limited to, xylose.

In some embodiments, the antibiotic compounds comprise the stereochemistry of Formula II:



Formula II

In one non-limiting example, the antibiotic compound has the structure:



5

Further provided is a method of inhibiting bacterial growth by contacting the bacteria with an antibiotic compound disclosed herein. In some embodiments, the bacteria are Gram-positive bacteria. For example, the Gram-positive bacteria may be selected from *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Bacillus subtilis* and *Clostridium difficile*. In other embodiments, the bacteria are Gram-negative bacteria. In some examples, the Gram-negative bacteria are selected from *Pseudomonas aeruginosa*, *Salmonella enterica*, *Pseudomonas putida*, *Escherichia coli*, *Acinetobacter baumannii* and *Haemophilus influenzae*. In some embodiments, the bacteria is a species of *Mycobacterium*, such as, but not limited to, *Mycobacterium bovis* or *Mycobacterium tuberculosis*.

In some embodiments of the present disclosure, the method of inhibiting bacterial growth is an *in vitro* method. In other embodiments, the method of inhibiting bacterial growth is an *in vivo*

method. In some examples of the *in vivo* method, contacting the bacteria with the compound comprises administering a therapeutically effective amount of the compound to a subject infected with the bacteria.

Also provided herein is a method of treating a bacterial infection in a subject by selecting a subject in need of treatment and administering to the subject a therapeutically effective amount of an antibiotic compound disclosed herein. In some embodiments, the subject in need of treatment is a subject with a bacterial infection, such as an infection with Gram-positive bacteria, Gram-negative bacteria or Mycobacteria.

In some examples of the disclosed methods, the subject is a non-human animal. In other examples, the subject is a human.

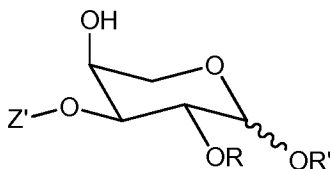
Also provided herein is a method of making a compound according to Formula I, where Y is C₁-C₁₀ alkyl, and Z is hydrogen, -OH, or a monosaccharide. In some embodiments, the method comprises:

providing a first amino acid according to the formula Y-CH(NHnOS)C(O)OH, where nOS is a protecting group;

providing a homoserine analog according to the formula R''OCH₂CH₂CH(NH₂)C(O)OH where R'' is a protecting group;

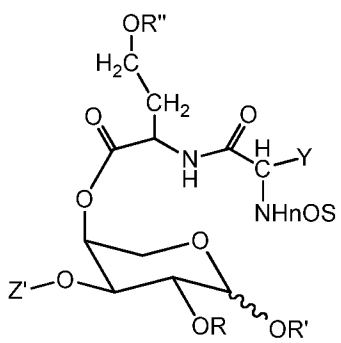
coupling the first amino acid to the homoserine analog to form an amino acid dimer;

providing a carbohydrate precursor having the structure

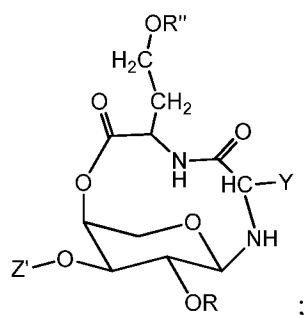


where R and R' are protecting groups, and Z' is hydrogen, R''' where R''' is a protecting group, or a monosaccharide precursor comprising one or more protecting groups in place of hydroxyl groups;

coupling the amino acid dimer to the carbohydrate precursor, thereby forming the structure



cyclizing the amino acid dimer and carbohydrate precursor, thereby forming the structure



removing protecting groups to form hydroxyl groups, wherein the protecting groups comprise R, R''' and, when Z' is R''', R'''; and

oxidizing the hydroxyl group formed by removal of R'', thereby forming the compound according to Formula I.

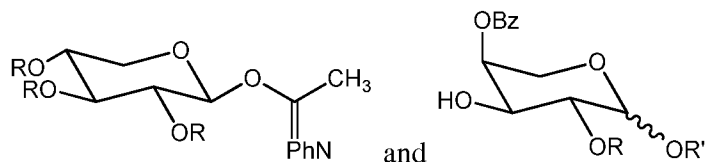
In some embodiments of the method for making the compound of Formula I, the first amino acid and the homoserine analog are coupled via an intermolecular condensation reaction using 1-mesitylene-2-fulsonyl-3-nitro-1,2,4-triazole.

In some embodiments, the amino acid dimer is coupled to the carbohydrate precursor via pentafluorophenyl ester coupling. In some examples, the amino acid dimer and carbohydrate precursor are cyclized via a Mitsunobu reaction.

In some embodiments of the method for making the compound of Formula I, R is p-methoxybenzyl ether, R' is n-pentenyl, and R'' is *tert*-butyldimethylsilyl ether.

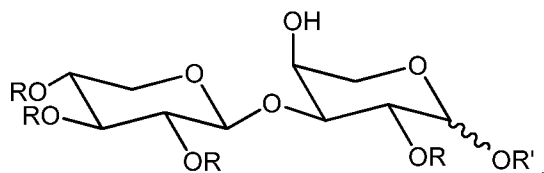
In some embodiments, Z is xylose and providing the carbohydrate precursor further comprises:

providing first and second monosaccharide precursors having the structures



where R, R', and Bz are protecting groups p-methoxybenzyl ether, n-pentenyl, and benzoyl, respectively, and Ph is phenyl; and

coupling the first and second monosaccharide precursors to form the carbohydrate precursor, wherein the carbohydrate precursor has the structure



In some examples, the first and second monosaccharide precursors are coupled via an intermolecular condensation reaction using trimethylsilyl trifluoromethanesulfonate.

The present disclosure provides antibiotic compounds substantially as hereinbefore described, wherein the antibiotic compounds are isolated from a microorganism or synthesized as
5 hereinbefore described.

Also provided is a method of isolating the disclosed antibiotic compounds as described in Example 2, 3 or 6.

V. Identification of Non-Ribosomal Peptide Synthetase (NRPS) Gene Clusters

10 Described herein is the discovery of a new non-ribosomal peptide synthetase (NRPS) cluster for the biosynthesis of an unusual, novel secondary metabolite in the sequenced genome of the hyper-thermophilic anaerobic bacteria *Clostridium thermocellum* (FIG. 1). The enzymatic functions of the encoded proteins of this novel NRPS assembly line were identified (FIG. 2). The assumed secondary metabolite appeared unusual as additional unique enzymatic functions were
15 associated with its biosynthesis, including glycosyl-transferases and a copper coordinating amine-oxidase. The chemical structure of the new small glyco-peptide was proposed and its biosynthetic pathway was reconstructed *in vitro* by analyzing the order of the enzymatic functions associated with this NRPS cluster. The chemical structure of the isolated natural product (referred to herein as Compound #1) was analyzed by mass spectrometry and NMR spectroscopy.

20 Also disclosed herein is the identification of homologous NRPS gene clusters in several additional microorganisms, including *Herpetosiphon aurantiacus*, *Myxococcus xanthus*, *Streptosporangium roseum*, *Catenulispora acidiphila*, *Pyrobaculum aerophilum*, *Rickettsia conorii* and *Emericella nidulans*. A schematic of the NRPS gene cluster identified in *Herpetosiphon aurantiacus* is shown in FIG. 6. In addition, the amino acid sequences of the adenylation domains
25 (A domains) or homologous proteins from *Clostridium thermocellum*, *Herpetosiphon aurantiacus*, *Streptosporangium roseum*, *Catenulispora acidiphila*, *Pyrobaculum aerophilum* and *Emericella nidulans* are set forth herein as SEQ ID NOs: 1-6, respectively. The antibiotic secondary metabolites produced by the NRPS clusters of *Herpetosiphon aurantiacus* and *Emericella nidulans* have also been isolated and are referred to in the Examples as Compound #2 and Compound #3,
30 respectively.

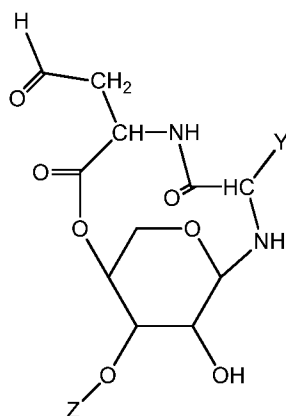
Further described herein are studies demonstrating the antibiotic activity of Compound #1, Compound #2 and Compound #3 against several Gram-positive and Gram-negative bacterial pathogens. As described in the Examples below, these antibiotic compounds are not toxic to mammalian cells and are well tolerated *in vivo*.

The antibiotic secondary metabolites produced by *C. thermocellum* and *H. aurantiacus* do not show any similarities to any currently known antibiotics. It is likely that the host organisms themselves are susceptible to their own metabolite based on the organization of the biosynthetic pathway that assembles cytosolic precursors in the periplasm. This suggests that evolution of antibiotic resistance against these compounds will not be easily accomplished.

VI. Antibiotic Compounds

A. Structure

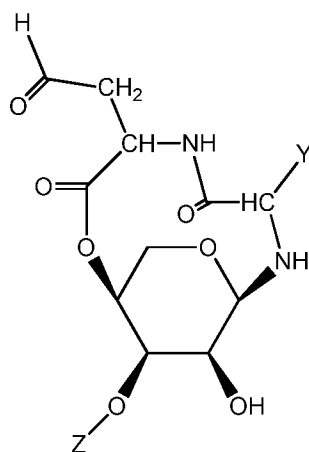
Embodiments of the disclosed antibiotic compounds comprise at least one monosaccharide moiety and an oligopeptide moiety, which forms a cyclic structure with the monosaccharide. In some embodiments, the oligopeptide moiety is a dipeptide. In certain embodiments, the antibiotic compound comprises a disaccharide moiety and an oligopeptide moiety, which forms a cyclic structure with one of the monosaccharides. In some embodiments, an antibiotic compound has a structure according to Formula I:



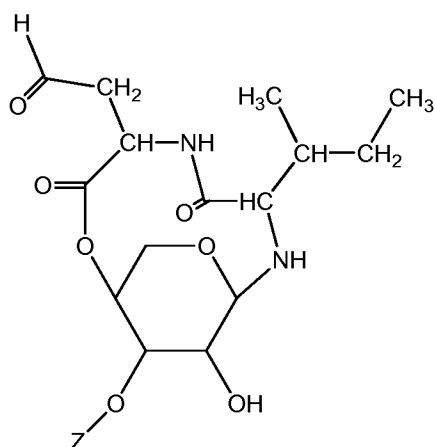
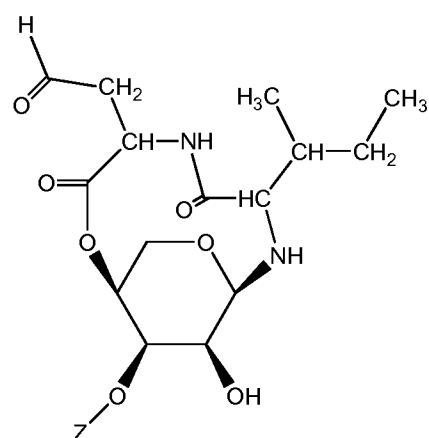
Formula I

With reference to Formula I, Y is C₁-C₁₀ alkyl, and Z is hydrogen, -OH, or a monosaccharide. In some embodiments, Y is a branched C₄ alkyl. In certain embodiments, Z is a hexose.

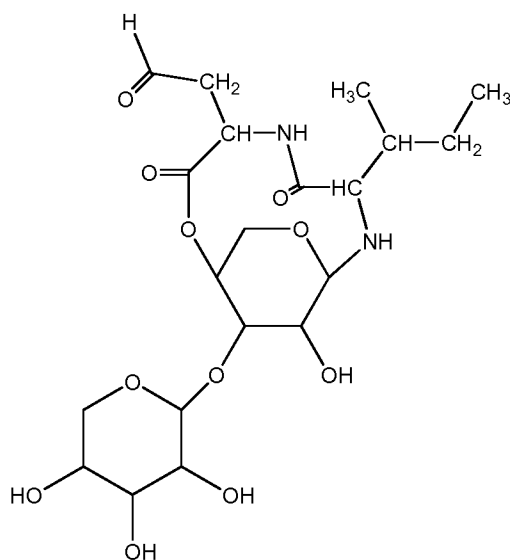
In certain embodiments, an antibiotic compound according to Formula I has a stereochemistry as shown in Formula II:

**Formula II**

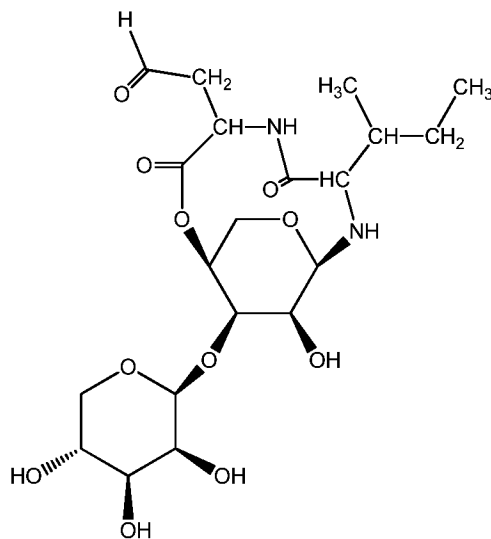
In some embodiments, as shown in Formulas IIIA and IIIB, the antibiotic compound has a structure according to Formula I where Y is $-\text{CH}(\text{CH}_3)\text{CH}_2\text{CH}_3$.

**Formula IIIA****Formula IIIB**

In another embodiment, as shown in Formula IV, the antibiotic compound has a structure according to Formula I where Y is $-\text{CH}(\text{CH}_3)\text{CH}_2\text{CH}_3$, and Z is xylose:

**Formula IV**

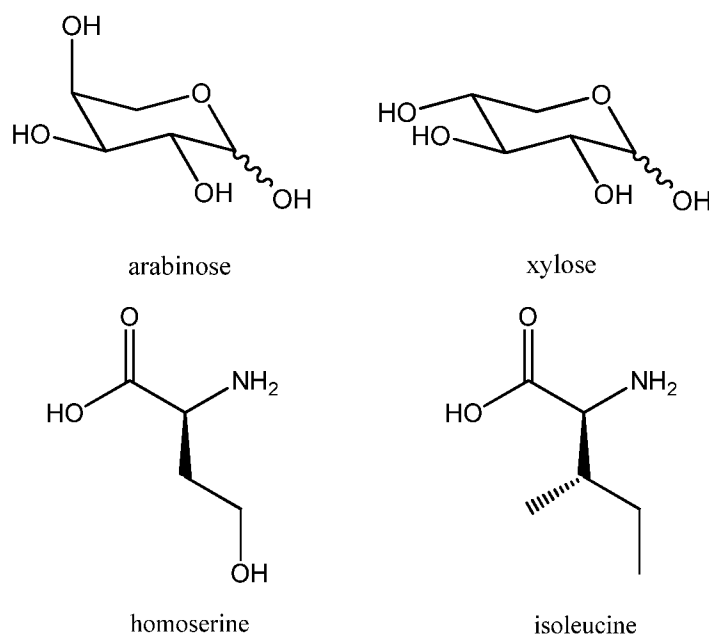
In a particular embodiment, the antibiotic compound has a structure according Formula III with a stereochemistry as shown below:

**Compound #1**

In some embodiments, an antibiotic compound according to Formula I is isolated from a microorganism. For example, an antibiotic compound according to Formula I may be isolated from *Clostridium thermocellum*, *Herpetosiphon aurantiacus*, *Myxococcus xanthus*, *Streptosporangium roseum*, *Catenulisporea acidiphila*, *Pyrobaculum aerophilum*, *Rickettsia conorii* or *Emericella nidulans*.

In certain embodiments, an antibiotic compound according Formula I is synthesized from its building blocks, which comprise one or two carbohydrate moieties and at least two amino acid

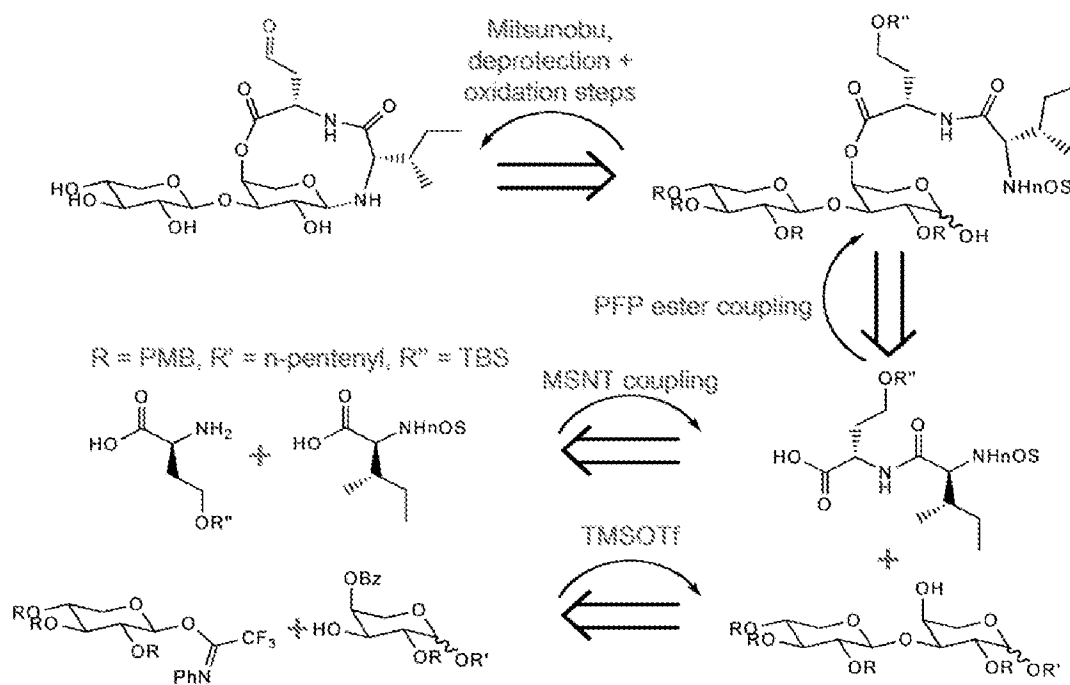
moieties. In one embodiment, Compound #1 is synthesized from two carbohydrate moieties, xylose and arabinose, and two amino acid moieties, homoserine and isoleucine.



5

B. Synthesis

An exemplary synthesis of Compound #1 is shown in Scheme 1, wherein Compound #1 is synthesized from its starting blocks of xylose, arabinose, homoserine, and isoleucine:



Scheme 1

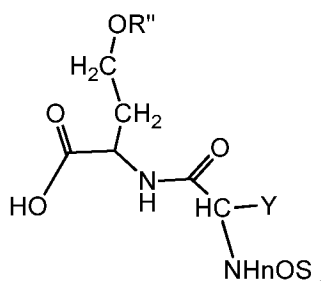
10

In Scheme 1, PMB is p-methoxybenzyl ether, TBS is *tert*-butyldimethylsilyl ether, Bz is benzoyl, TMSOTf is trimethylsilyl trifluoromethanesulfonate, MSNT is 1-(mesitylene-2-sulfonyl)-3-nitro-1,2,4-triazole, and PFP ester is pentafluorophenyl ester.

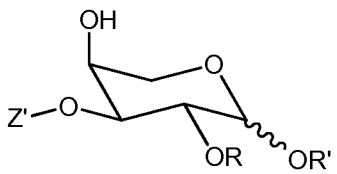
Non-reacting hydroxyl groups of arabinose and xylose may be protected with protecting groups according to conventional methods known to one of ordinary skill in the art of organic synthesis. The side-chain hydroxyl group of homoserine, and the amino group of isoleucine also may be protected.

The two carbohydrate precursors may be coupled together via an intermolecular condensation reaction using TMSOTf. The two amino acid precursors may be coupled together via an intermolecular condensation reaction using MSNT. The carbohydrate dimer and amino acid dimer then may be joined in another intermolecular condensation reaction via PFP ester coupling. Finally, the amino acid dimer and the arabinose moiety may be cyclized via a Mitsunobu reaction, the protected homoserine hydroxyl group may be oxidized, and the remaining protective groups may be removed to produce Compound #1.

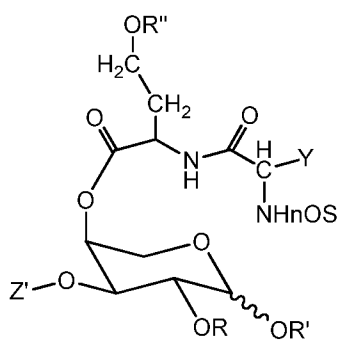
More generally, a compound according to Formula I is synthesized by providing a first amino acid having a formula $Y-CH(NHnOS)C(O)OH$, and a homoserine analog having the formula $R''OCH_2CH_2CH(NH_2)C(O)OH$. The first amino acid is coupled to the homoserine analog to form an amino acid dimer



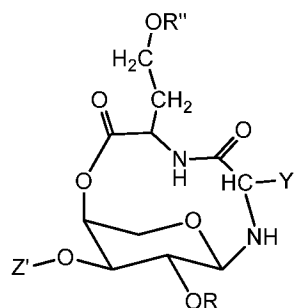
A carbohydrate precursor also is provided according to the structure:



where Z' is hydrogen, R''' where R''' is a protecting group, or a monosaccharide precursor comprising one or more protecting groups in place of hydroxyl groups. The amino acid dimer is coupled to the carbohydrate precursor, forming the structure:

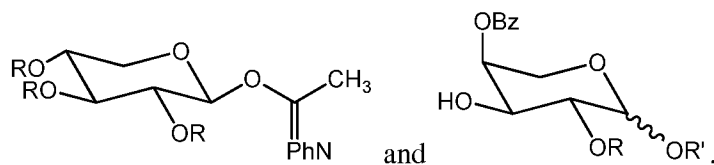


The amino acid dimer and carbohydrate precursor then are cyclized to form the structure:

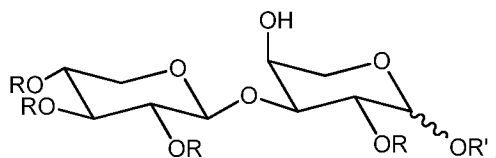


The protecting groups R and R'' are removed to form hydroxyl groups, and the hydroxyl group formed by removal of R'' is oxidized, thereby forming the compound according to Formula I.

In certain embodiments, Z is xylose, and providing the carbohydrate precursor includes providing first and second monosaccharide precursors having the structures:



where R, R', and Bz are protecting groups p-methoxybenzyl ether, n-pentenyl, and benzoyl, respectively, and Ph is phenyl. The first and second monosaccharide precursors are coupled to form the carbohydrate precursor



VII. Pharmaceutical Compositions and Administration Thereof

This disclosure includes pharmaceutical compositions comprising at least one antibiotic compound for use in human or veterinary medicine. Embodiments of pharmaceutical compositions include a pharmaceutically acceptable carrier and at least one active ingredient. Useful pharmaceutically acceptable carriers and excipients are known in the art. Active ingredients may

comprise, for example, at least one antibiotic compound as described herein. In addition, other medicinal or pharmaceutical agents, for example, with similar, related or complementary effects on the affliction being treated, may be included as active ingredients in pharmaceutical compositions.

The pharmaceutical compositions comprising the antibiotic compounds disclosed herein may be formulated in a variety of ways depending, for example, on the mode of administration and/or on the location and type of infection to be treated. For example, such pharmaceutical compositions may be formulated as pharmaceutically acceptable salts. As another example, parenteral formulations may comprise injectable fluids that are pharmaceutically and physiologically acceptable fluid vehicles such as water, physiological saline, other balanced salt solutions, aqueous dextrose, glycerol or the like. Excipients may include, for example, nonionic solubilizers, such as cremophor, or proteins, such as human serum albumin or plasma preparations. If desired, the pharmaceutical composition to be administered may also contain non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example, sodium acetate or sorbitan monolaurate.

The dosage form of the pharmaceutical composition can be determined, at least in part, by the mode of administration chosen. For example, in addition to injectable fluids, topical and oral formulations may be employed. Topical preparations may include eye drops, ointments, sprays and the like. Oral formulations may be liquid (*e.g.*, syrups, solutions or suspensions), or solid (*e.g.*, powders, pills, tablets, or capsules). For solid compositions, conventional non-toxic solid carriers may include pharmaceutical grades of mannitol, lactose, starch, or magnesium stearate. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in the art.

Certain embodiments of the pharmaceutical compositions comprising antibiotic compounds as described herein may be formulated in unit dosage form suitable for individual administration of precise dosages. The amount of a therapeutic compound administered will depend on the subject being treated, the type and severity of the infection, and the manner of administration, and is known to those skilled in the art. Within these bounds, the formulation to be administered will contain a quantity of the antibiotic compounds disclosed herein in an amount effective to achieve the desired effect in the subject being treated (*e.g.*, reducing or eliminating Gram-positive pathogens or Gram-negative pathogens).

The present disclosure contemplates treatments for infection of a subject by Gram-positive bacteria, Gram-negative bacteria and/or Mycobacteria. Such treatments include administering an antibiotic compound disclosed herein, or a combination of the antibiotic compound and one or more other pharmaceutical agents, to the subject in a pharmaceutically acceptable carrier and in an amount effective to treat a microbial infection. Subjects can be selected using more specific

criteria, such as a definitive diagnosis of a condition based on, for example, a biological specimen that has been provided to be tested for a bacterial infection.

The vehicle in which the antibiotic compound is delivered may include, for example, the pharmaceutical compositions described above. Routes of administration include but are not limited to oral and parenteral routes, such as intravenous, intraperitoneal, rectal, topical, ophthalmic, nasal, and transdermal.

Therapeutically effective doses of an antibiotic compound can be determined by one of skill in the art. An example of a dosage range is 0.1 to 200 mg/kg body weight orally in single or divided doses. Another example of a dosage range is 1.0 to 100 mg/kg body weight orally in single or divided doses. For oral administration, the compositions are, for example, provided in the form of a tablet containing 0.1 or 1.0 to 1000 mg of the active ingredient, particularly 1, 5, 10, 15, 20, 25, 50, 100, 200, 400, 500, 600, or 1000 mg of the active ingredient for the symptomatic adjustment of the dosage to the subject being treated.

The specific dose level and frequency of dosage for any particular subject may be varied and will depend upon a variety of factors, including the activity of the specific antibiotic compound, the metabolic stability and length of action of that compound, the age, body weight, general health, sex and diet of the subject, mode and time of administration, rate of excretion, drug combination, and type and severity of the infection in the subject being treated.

The following examples are provided to illustrate certain particular features and/or embodiments. These examples should not be construed to limit the disclosure to the particular features or embodiments described.

EXAMPLES

Example 1: Identification of a non-ribosomal peptide synthetase (NRPS) gene cluster in *Clostridium thermocellum* and other microorganisms

This example describes the discovery of a non-ribosomal peptide synthetase (NRPS) cluster for the biosynthesis of a novel secondary metabolite in the sequenced genome of *Clostridium thermocellum*.

C. thermocellum (ATCC 27405) is a hyper-thermophilic anaerobic bacterial strain with an optimal growth temperature of 60°C. *C. thermocellum* is Gram-positive, rod-shaped, and produces endospores. This bacterium produces an extracellular enzyme system capable of degrading crystalline cellulose to soluble sugars. A new NRPS gene cluster was identified in the sequenced genome of *C. thermocellum*. The enzymatic functions of the encoded proteins of this natural

product assembly line were identified (see FIG. 1 and FIG. 2). The assumed secondary metabolite appeared unusual since additional unique enzymatic functions are associated with its biosynthesis. These include two glycosyl-transferases with one active in the cytoplasm and the second one located in the periplasmic space. In addition, a copper amine-oxidase, responsible for a desamination reaction, was identified in the periplasm to be associated with the biosynthesis of this new natural product, referred to herein as Compound #1. The chemical structure of this small glyco-peptide was elucidated by NMR spectroscopy and mass spectrometry and its biosynthetic pathway was reconstructed *in vitro* by analyzing the order of the enzymatic functions associated with this NRPS cluster.

In addition to *C. thermocellum*, a number of additional strains were identified that carry genetic material in their genome coding for similar non-ribosomal peptide synthetase assembly systems. The identified strains are the following bacteria, archae bacterium and fungi:

Herpetosiphon aurantiacus (ATCC 23779)

Myxococcus xanthus DK 1622 (ATCC 19368/25232)

Streptosporangium roseum DSM 43021 (ATCC 12428)

Catenulispora acidiphila DSM 44928

Pyrobaculum aerophilum (NCBI ID 13773)

Rickettsia conorii (ATCC 613)

Emericella nidulans (also known as *Aspergillus nidulans*) (ATCC 11267)

A second compound (Compound #2), produced by *Herpetosiphon aurantiacus* (ATCC 23779), was purified and the elucidated structure confirms the chemical structure of the first compound. In addition, a third compound produced by *Emericella nidulans* has been isolated and is referred to herein as Compound #3.

Example 2: Isolation and characterization of an antibiotic compound produced by the *C. thermocellum* NRPS gene cluster

C. thermocellum was grown anaerobically in cellulose-containing medium. ¹⁴C-labelled amino acids (¹⁴C-L-Asp, ¹⁴C-L-Leu and ¹⁴C-L-Ile) were added to the culture media and radio-HPLC was performed. The radiolabelled amino acids were consumed by *C. thermocellum* and no new peak was detected, indicating the secondary metabolite is water insoluble (FIG. 3). To extract the water insoluble pellet, a radiolabelled 2 L culture was filtered through a Büchner funnel, the solids were resuspended in 50 mL DMSO and shaken for 2 hours. The orange suspension was centrifuged, the liquid layer separated from the cells and the DMSO was removed under reduced

pressure. The final yield from this process was 14.7 mg. Radio-thin layer chromatography (TLC) of 1 mg/mL solution in DMSO revealed an additional spot (FIG. 4).

The DMSO extract of the solids was tested for antimicrobial activity against *E. coli* and *B. subtilis* indicator strains. As shown in FIG. 5, 20 µl of 1 mg/ml of the DMSO extract was bioactive against *B. subtilis*.

Cloning and expression of *C. thermocellum* genes

Genomic DNA was purchased from ATCC. PCR was carried out using Phusion HF Master Mix (Finnzymes). DNA was cloned into the pET151-DTOPO expression vector, which encodes for a hexa-His-tag and a TEV protease cleavage site. Genes (including Cthe-1904 A domain, Cthe_1905, Cthe_1906 and Cthe_1907) were expressed in BL21 Star™ (DE3) competent *E. coli* cells. Competent cells were grown at 37°C until an OD of 0.6 was reached. Cells were induced with 0.1 mM IPTG final concentration, and further grown for 14 hours at 16°C. Protein was purified using Ni-NTA purification with increasing imidazole concentration.

Structure elucidation by NMR

C. thermocellum cultures were fed ¹³C-glucose, which was incorporated into the secondary metabolite. The metabolite was isolated by DMSO extraction, as described above. For structure elucidation, 2D spectra were recorded (¹³CHSQC, ¹³C refocused rotational NOESY, ¹H¹H-COSY). The structure of Compound #1 contains a dipeptide, which is bridged by a sugar.

Growth of *C. thermocellum* cultures

C. thermocellum can be grown in either small or large cultures, such as cultures of up to 5 gallons. A 5 gallons culture of *C. thermocellum* (ATCC 27405) usually grows for 18-24 days until sporulation has started. The production of the antibiotic secondary metabolite appears to be directly linked to sporulation and Compound #1 can be isolated by washing spores with di-methylsulfoxide (DMSO) and subsequently precipitated with ethanol and freeze dried to a DMSO solution. For each 5 gallon culture, approximately 85 mg of >95% pure compound is isolated and can be used for testing dissolved in 200 µl DMSO.

Example 3: NRPS gene cluster in *Herpetosiphon aurantiacus*

Herpetosiphon aurantiacus was originally isolated from Lake Birch in Minnesota in 1968. It is a Gram-negative bacterium which grows in long, flexible filaments with a transparent section of empty cell wall at the end of the filament. An NRPS gene cluster homologous to the one

identified in *C. thermocellum* was found in *H. aurantiacus* (FIG. 6). 1D-¹H-NMR of the extract of the water-insoluble materials after centrifugation of the *H. aurantiacus* culture indicated a structure of the metabolite similar to Compound #1.

Genomic DNA of *H. aurantiacus* (ATCC 23779) was prepared and subject to PCR using Phusion HF Master Mix (Finnzymes). Amplified DNA was cloned into the pET151-DTOPO expression vector, which encodes for a hexa-His tag and a TEV protease cleavage site. Several *H. aurantiacus* genes (Haur_3126, Haur_3127, Haur_3128 and Haur_3129) were expressed in BL21 StarTM (DE3) competent *E. coli* cells. Competent cells were grown at 37°C until an OD of 0.6 was reached. Cells were induced with 0.1 mM IPTG final concentration, and further grown for 14 hours at 16°C.

Example 4: Toxicity screening of antibiotic compounds

In vitro assays

In vitro cell toxicology assays were performed to evaluate the effect of Compound #1 on human cells. Compound #1, dissolved in DMSO, was added to vital cultures of HeLa cells and peripheral blood mononuclear cells (PMBC) one day after the cell cultures were started. For each cell line, pure DMSO (control #1), no DMSO (control #2) and Compound #1 in one times (T1 – 5 nM), five times (T5 – 25 nM) and ten times (T10 – 50 nM) of the bacteria-toxic concentration was added to the cell cultures for two sets each. All ten culture flasks per cell line (2x: T1, T5, T10, C#1, C#2) were incubated under identical conditions and simple cell counts were performed on days two, four and six while replacing the cell culture medium to maintain the concentration of Compound #1.

The observed death rates for treated and untreated cell cultures were essentially identical (less than 5% variation). A slightly reduced proliferation rate for T10 (50 nM)-treated HeLa cells was observed in both experiments. No pyrogenic reactions, and no changes in physical appearance of either cell line, were observed.

Maximum tolerated dose (MTD) studies

MTD studies were performed for Compound #1, Compound #2 and Compound #3. CD-1 mice, 6-8 weeks of age, were used for these studies. Groups of three mice were injected subcutaneously with 20 mg/kg, 40 mg/kg or 80 mg/kg of Compound #1, Compound #2 or Compound #3. Animals were observed 1 hour, 24 hours and 48 hours after administration of compound. No signs of illness were observed in the mice at any doses tested. These results indicate that all three compounds are well tolerated *in vivo*.

Example 5: Growth inhibition and minimum inhibitory concentration (MIC) assays

To evaluate the antibiotic properties of the compounds produced by the NRPS cluster, Compound #1 or Compound #2 (0.85 µg) was diluted in 2 µl DMSO and soaked onto 6 mm cellulose disks for a bacterial growth inhibition assay. The results demonstrated that Compound #1 and Compound #2 significantly inhibited bacterial growth of the following Gram(+) and Gram(-) bacterial strains:

E. coli K-12

E. coli (ATCC 25922)

Staphylococcus aureus (ATCC 2913)

Pseudomonas putida (ATCC 11172)

Listeria grayi (ATCC 19120)

Acinetobacter baumannii (ATCC 17904)

Bacillus subtilis PY79

As a control, disks were soaked in DMSO alone. The control disks did not cause any bacterial growth inhibition.

Additional studies were carried out to determine the minimum inhibitory concentration (MIC) of Compound #1, Compound #2 and Compound #3 against *Pseudomonas aeruginosa*, *Salmonella enterica*, *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Clostridium difficile*.

All compounds were completely dissolved in DMSO at a concentration of 25,600 µg/mL. Stocks of vancomycin hydrochloride (VAN) and ciprofloxacin (CIP) were prepared at a concentration of 3,200 µg/mL in water, and 400 µg/mL in slightly acidified water, respectively. Stocks of test compounds and standards were diluted 1:100 to provide working stocks in each growth medium.

For *P. aeruginosa*, *S. enterica* and *S. aureus*, eleven serial, one-half dilutions of the compounds and standards were prepared in 96-well polypropylene dilution blocks in cation-adjusted Mueller Hinton broth (CAMHB). The dilutions were transferred to 96-well assay plates. The assay concentration ranges of the test compounds were 128-0.12 µg/mL. Bacterial suspensions were prepared and added to each well at a concentration of approximately 5×10^5 colony-forming-units per milliliter. The inoculated plates were incubated for 16-20 hours at $35 \pm 1^\circ\text{C}$. At the completion of incubation the wells of each plate were evaluated visually for the presence of bacterial growth. The MIC was the concentration at which there was no growth.

MICs against *H. influenzae*, *S. pneumoniae* and *C. difficile* were determined similarly, but dilutions were prepared directly in the 96-well assay plates and MIC parameters were as indicated below (Table 1).

5 **Table 1. MIC parameters (all incubations were at 35°C)**

Species	Inoculum Medium	MIC Incubation Medium	Air	Incubation Time (h)
<i>P. aeruginosa</i> , <i>S. enterica</i> and <i>S. aureus</i>	Trypticase Soy Agar (TSA)	Cation Adjusted Mueller-Hinton Broth (CAMHB)	Ambient	16-20
<i>S. pneumoniae</i>	TSA + 5% Sheep's Blood (TSA+SB)	CAMHB + Lysed horse blood (LHB)	Ambient	20-24
<i>H. influenzae</i>	Chocolate Agar (enhanced CO ₂)	Haemophilus Test Medium Broth	Ambient	20-24
<i>C. difficile</i>	TSA+SB (anaerobic)	Supplemented Brucella Broth + LHB	Anaerobic	46-48

The standard for *H. influenzae*, *P. aeruginosa*, *S. enterica* and *S. aureus* was CIP (2-0.002 µg/mL); and the standard for *S. pneumoniae* and *C. difficile* was VAN (16-0.016 µg/mL). All testing was performed in duplicate.

10 Results

Table 2. Minimum inhibitory concentrations (MIC) of test compounds and standards against six species (µg/mL)

Species	ATCC ^a	Compound			Standard	
		#1	#2	#3	CIP ^b	VAN ^c
<i>P. aeruginosa</i>	27853	>128	>128	>128	0.5 ^d	--
<i>S. enterica</i>	14028	>128	>128	>128	0.25	--
<i>S. pneumoniae</i>	49619	64	16	64	--	0.25 ^e
<i>H. influenzae</i>	49247	2	1	2	0.016 ^f	--
<i>C. difficile</i>	700057	32	16	16	--	0.25

^aAmerican Type Culture Collection

^bCiprofloxacin

^cVancomycin hydrochloride

^dCLSI acceptable range: 0.25-1 µg/mL

^eCLSI acceptable range: 0.12-0.5 µg/mL

^fCLSI acceptable range: 0.004-0.03 µg/mL

15

The greatest activity observed was against the Gram-negative *H. influenzae*. Activity at high concentrations was also observed against the Gram-positive species *C. difficile* and *S. pneumoniae*.

MIC was further evaluated using EtOAc extracts of more highly purified Compound #1 and Compound #2. MIC of these compounds was compared to ACHN-490 (Plazomicin; an aminoglycoside effective primarily against Gram-negative bacteria), Colistin (a polymyxin antibiotic effective against most Gram-negative bacilli), and Meropenem (an ultra broad-spectrum, beta-lactam antibiotic). The results are shown in Table 3.

Table 3. MIC of Compounds #1 and #2 against seven bacterial species (µg/mL)

Species	ATCC	Compound #1	Compound #2	ACHN-490	Colistin	Meropenem
<i>S. aureus</i> (MRSA)	33591	16	8	0.25...4	32	>16
<i>S. enterica</i>	14028	16	8	0.25...4	>32	>16
<i>K. pneumoniae</i>	43816	8	4	0.25...64	0.12	0.5...>16
<i>H. influenzae</i>	49247	0.5	<0.125	N/A	N/A	N/A
<i>S. pneumoniae</i>	49619	16	8	N/A	N/A	N/A
<i>A. baumannii</i>	19606	8	4	8	>32	1
<i>C. difficile</i>	70057	16	8	N/A	N/A	N/A

Both Compound #1 and Compound #2 exhibited antimicrobial activity against all bacterial strains tested.

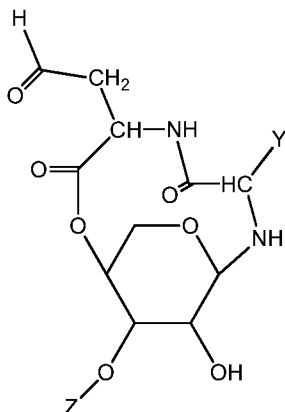
Example 6: General isolation procedure for purification of antibiotic secondary metabolites

Procedures for isolating antibiotic compounds produced by the NRPS gene clusters can vary depending upon, for example, the organism producing the secondary metabolite, the quantity of product desired and the solubility of the desired compound. The following procedure is a general procedure for isolation of the secondary metabolites. For solid cultures, such as spores of *C. thermocellum* or fungi cultures on plates, extractions of cells are performed using DMSO. Cultures are filtered and the clear DMSO extract is collected and washed with diethylether to remove fatty acids. The extract is precipitated and frozen (-20°C) with ethanol to remove nucleic acids and proteins, then lyophilized to remove DMSO. The product is then washed with ethanol. The ethanol is removed to get a dry powder.

In view of the many possible embodiments to which the principles of the disclosed invention may be applied, it should be recognized that the illustrated embodiments are only preferred examples of the invention and should not be taken as limiting the scope of the invention. Rather, the scope of the invention is defined by the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

CLAIMS

1. A compound according to Formula I:



Formula I

where Y is C₁-C₁₀ alkyl, and Z is hydrogen, -OH, or a monosaccharide.

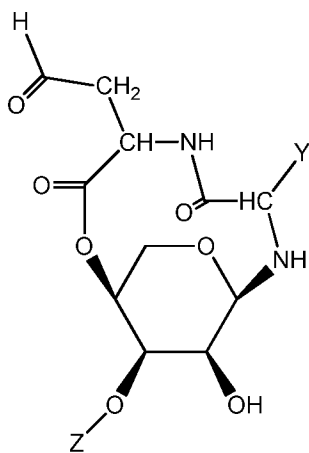
2. The compound of claim 1 where Y is a branched C₄ alkyl.

3. The compound of claim 2 where Y is -CH(CH₃)CH₂CH₃.

4. The compound of claim 1 where Z is a hexose.

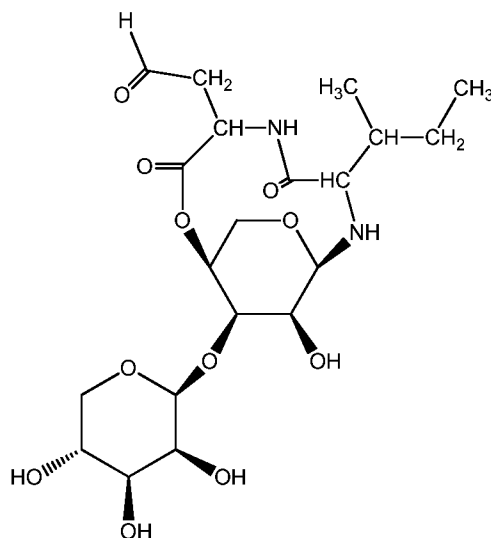
5. The compound of claim 4 where Z is xylose.

6. The compound of claim 1 having the stereochemistry of Formula II:



Formula II.

7. The compound of claim 1 having the structure:



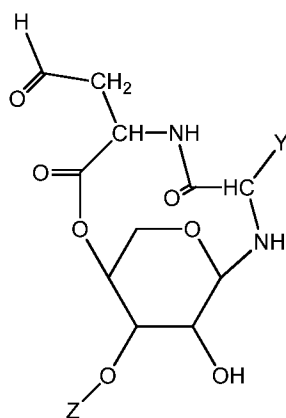
8. A method of inhibiting bacterial growth, comprising contacting bacteria with the
5 compound of any one of claims 1-7.
9. The method of claim 8, wherein the bacteria are Gram-positive bacteria.
10. The method of claim 9, wherein the Gram-positive bacteria are selected from the
10 group consisting of *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Bacillus subtilis* and *Clostridium difficile*.
11. The method of claim 8, wherein the bacteria are Gram-negative bacteria.
12. The method of claim 11, wherein the Gram-negative bacteria are selected from the
15 group consisting of *Pseudomonas aeruginosa*, *Salmonella enterica*, *Pseudomonas putida*, *Escherichia coli*, *Acinetobacter baumannii* and *Haemophilus influenzae*.
13. The method of claim 8, wherein the bacteria is a species of *Mycobacterium*.
14. The method of claim 13, wherein the *Mycobacterium* species is *Mycobacterium bovis* or *Mycobacterium tuberculosis*.
15. The method of any one of claims 8-14, wherein the method is an *in vitro* method.

16. The method of any one of claims 8-14, wherein the method is an *in vivo* method and contacting bacteria with the compound comprises administering a therapeutically effective amount of the compound to a subject infected with the bacteria.

17. The method of claim 16, wherein the subject is a non-human animal.

18. The method of claim 16, wherein the subject is a human.

19. A method of making a compound according to Formula I:



Formula I

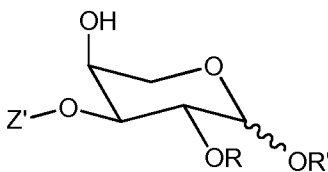
where Y is C₁-C₁₀ alkyl, and Z is hydrogen, -OH, or a monosaccharide, the method comprising:

providing a first amino acid according to the formula Y-CH(NHnOS)C(O)OH, where nOS is a protecting group;

providing a homoserine analog according to the formula R''OCH₂CH₂CH(NH₂)C(O)OH where R'' is a protecting group;

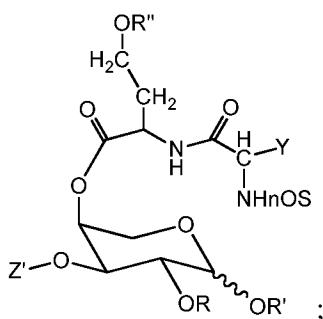
coupling the first amino acid to the homoserine analog to form an amino acid dimer;

providing a carbohydrate precursor having the structure

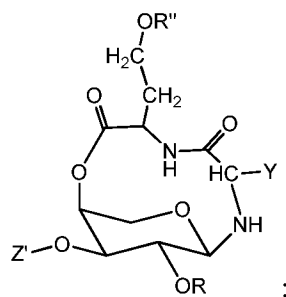


where R and R' are protecting groups, and Z' is hydrogen, R''' where R''' is a protecting group, or a monosaccharide precursor comprising one or more protecting groups in place of hydroxyl groups;

coupling the amino acid dimer to the carbohydrate precursor, thereby forming the structure



cyclizing the amino acid dimer and carbohydrate precursor, thereby forming the structure



removing protecting groups to form hydroxyl groups, wherein the protecting groups

5 comprise R, R'' and, when Z' is R''', R'''; and

oxidizing the hydroxyl group formed by removal of R'', thereby forming the compound according to Formula I.

20. The method of claim 19, wherein the first amino acid and the homoserine analog are coupled via an intermolecular condensation reaction using 1-mesitylene-2-fulsonyl-3-nitro-1,2,4-triazole.

21. The method of claim 19, wherein the amino acid dimer is coupled to the carbohydrate precursor via pentafluorophenyl ester coupling.

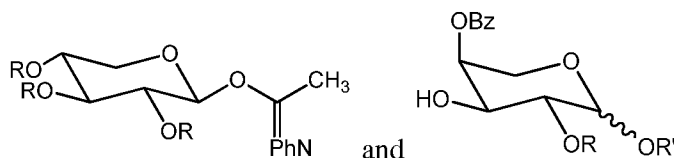
15

22. The method of claim 19, wherein the amino acid dimer and carbohydrate precursor are cyclized via a Mitsunobu reaction.

23. The method of claim 19, wherein R is p-methoxybenzyl ether, R' is n-pentenyl, and R'' is *tert*-butyldimethylsilyl ether.

24. The method of claim 19, wherein Z is xylose and providing the carbohydrate precursor further comprises:

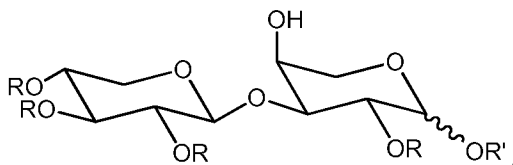
providing first and second monosaccharide precursors having the structures



where R, R', and Bz are protecting groups p-methoxybenzyl ether, n-pentenyl, and benzoyl, respectively, and Ph is phenyl; and

coupling the first and second monosaccharide precursors to form the carbohydrate

5 precursor, wherein the carbohydrate precursor has the structure

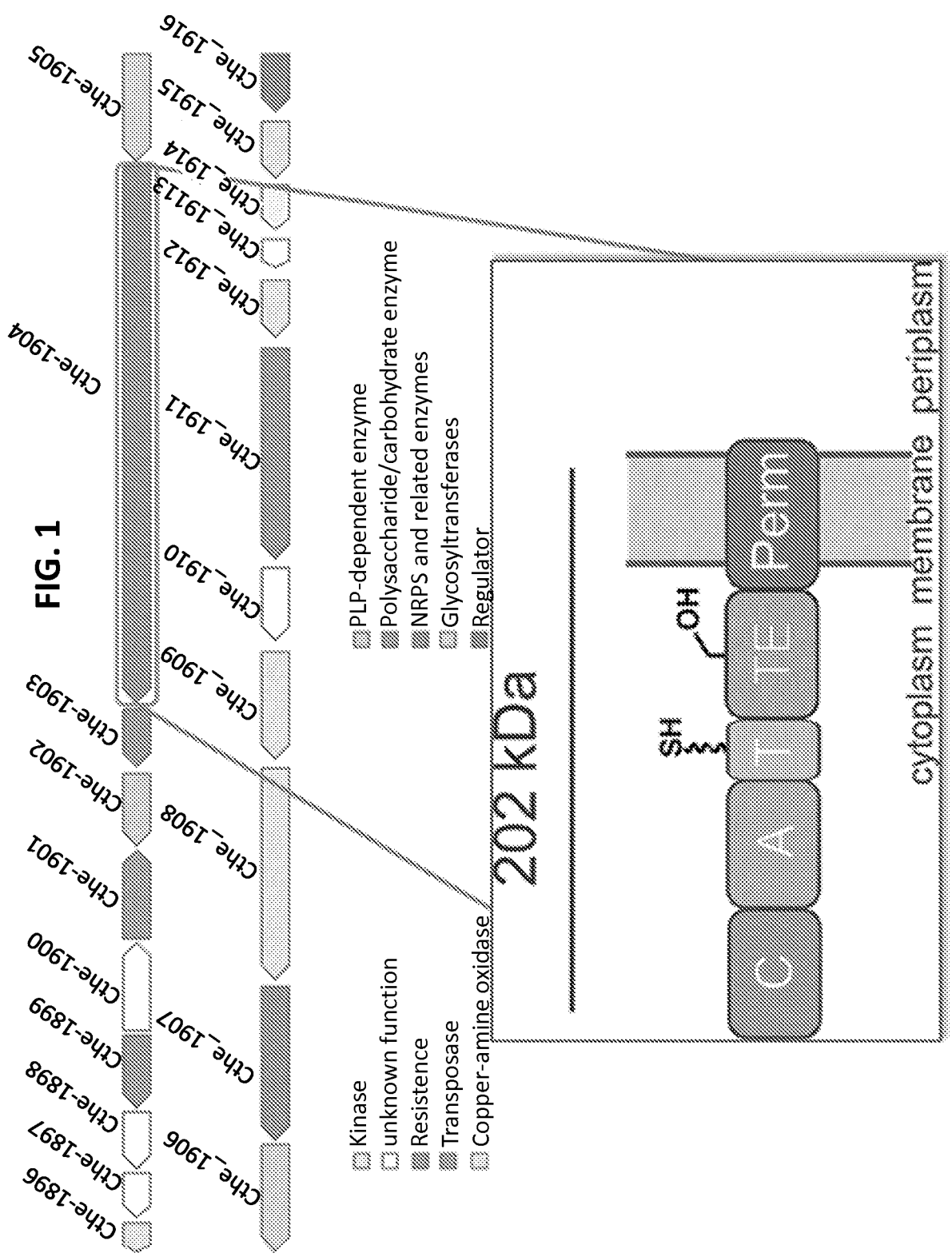


25. The method of 24, wherein the first and second monosaccharide precursors are coupled via an intermolecular condensation reaction using trimethylsilyl trifluoromethanesulfonate.

10

26. An antibiotic compound substantially as hereinbefore described, wherein the antibiotic compound is isolated from a microorganism or synthesized as hereinbefore described.

27. Use of an antibiotic compound for inhibiting bacterial growth substantially as
15 hereinbefore described, wherein the antibiotic compound is isolated from a microorganism or synthesized as hereinbefore described.



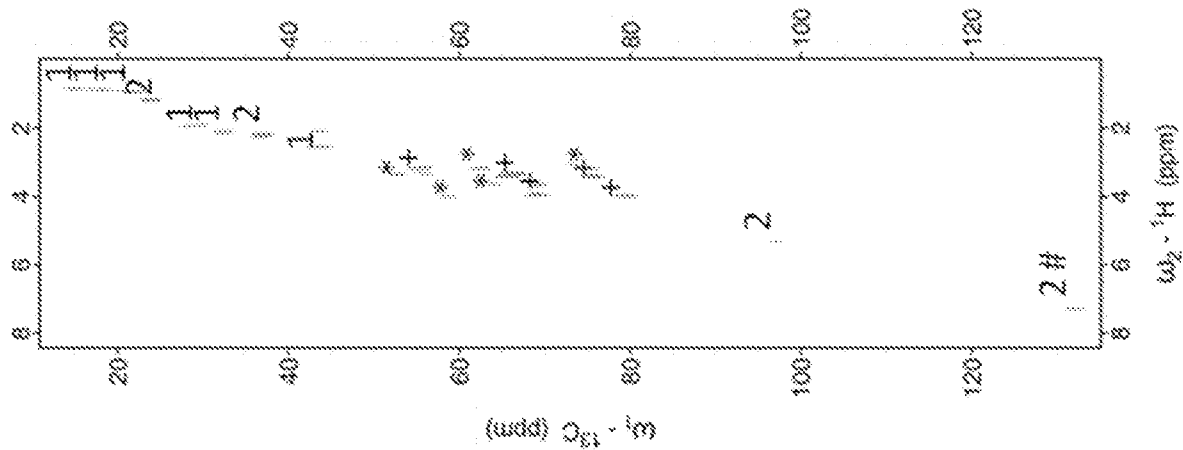


FIG. 2

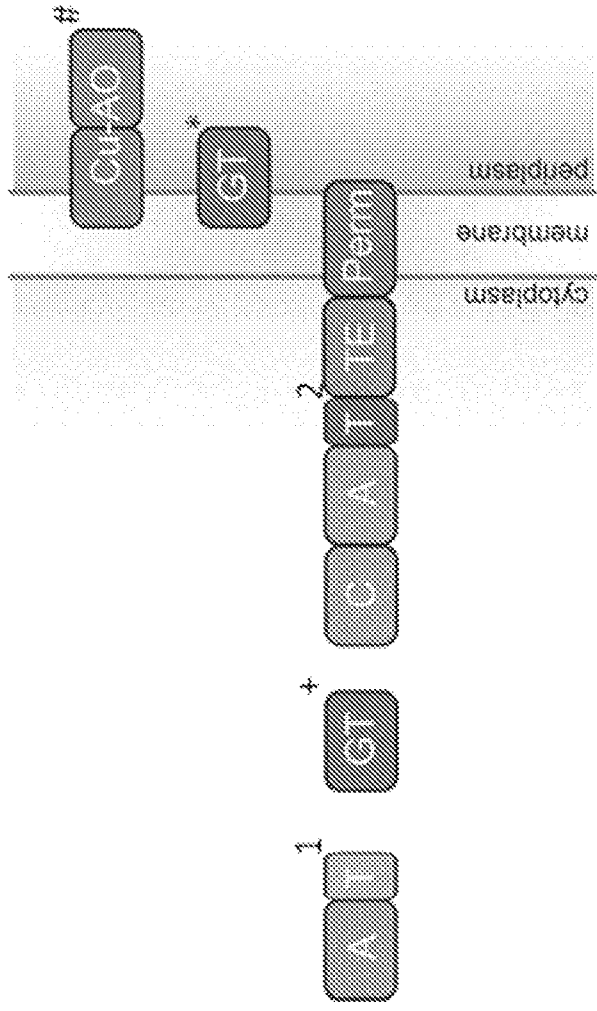


FIG. 3

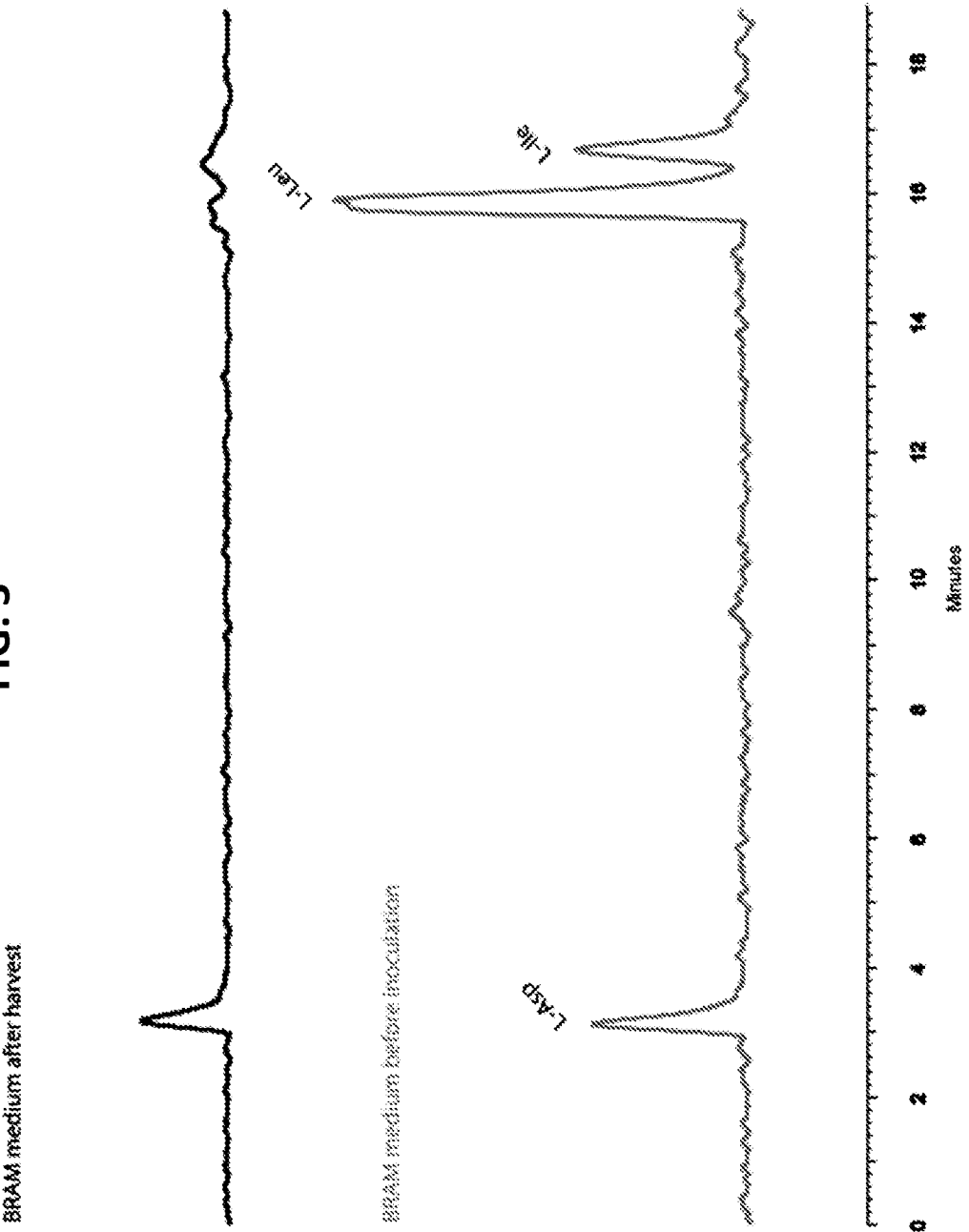


FIG. 4

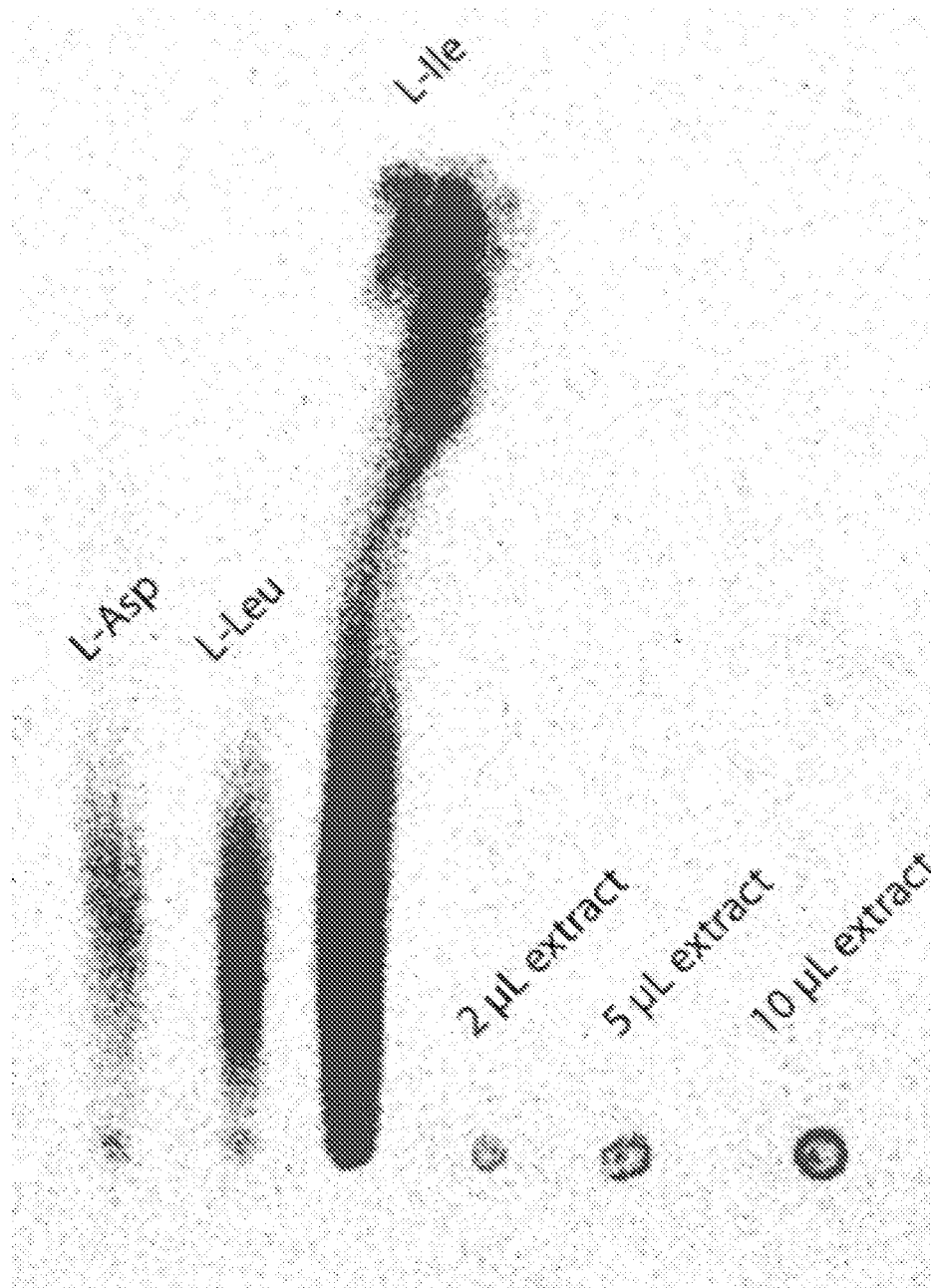
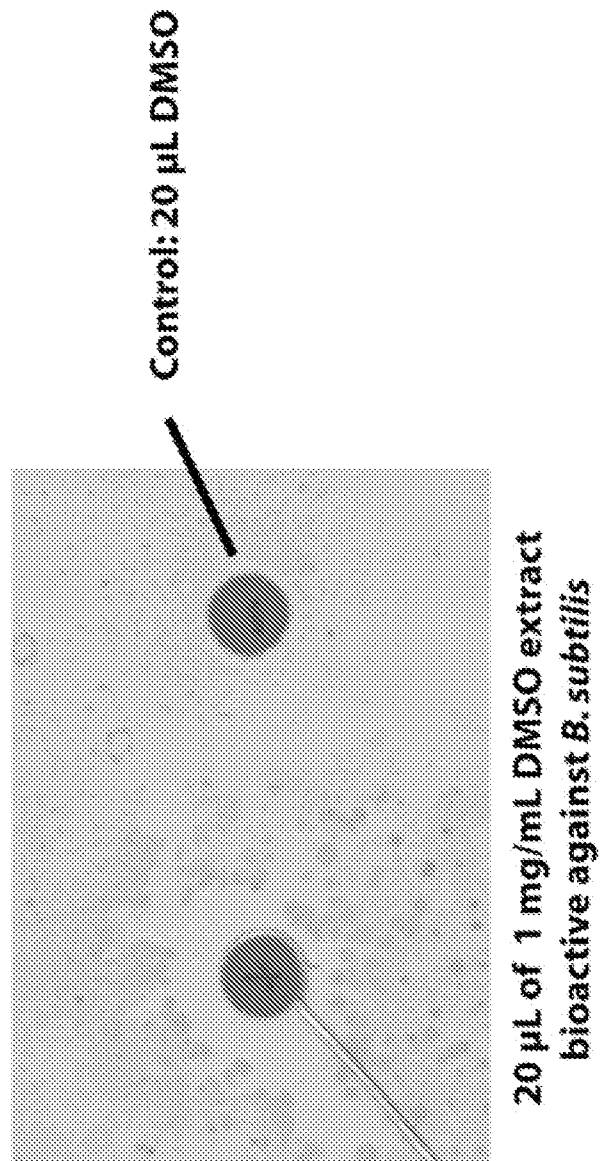
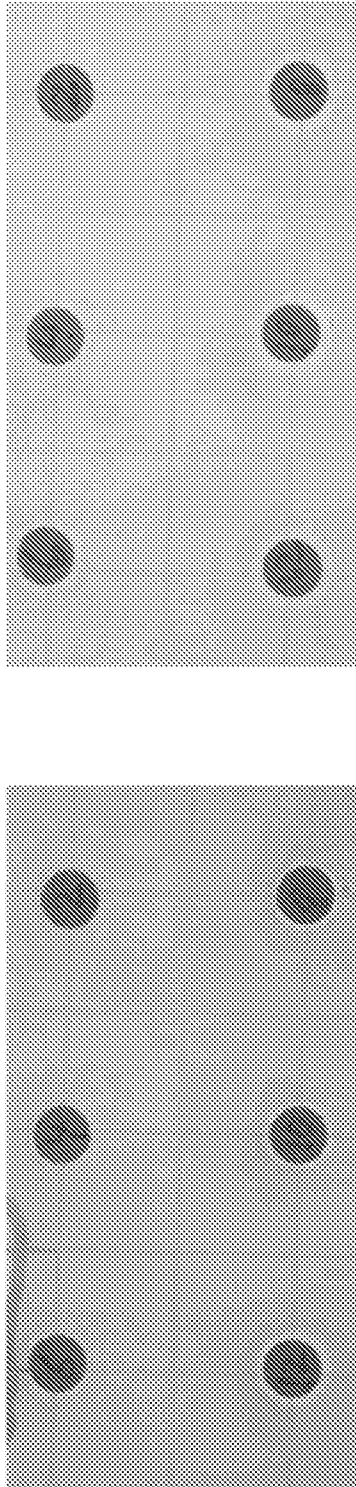
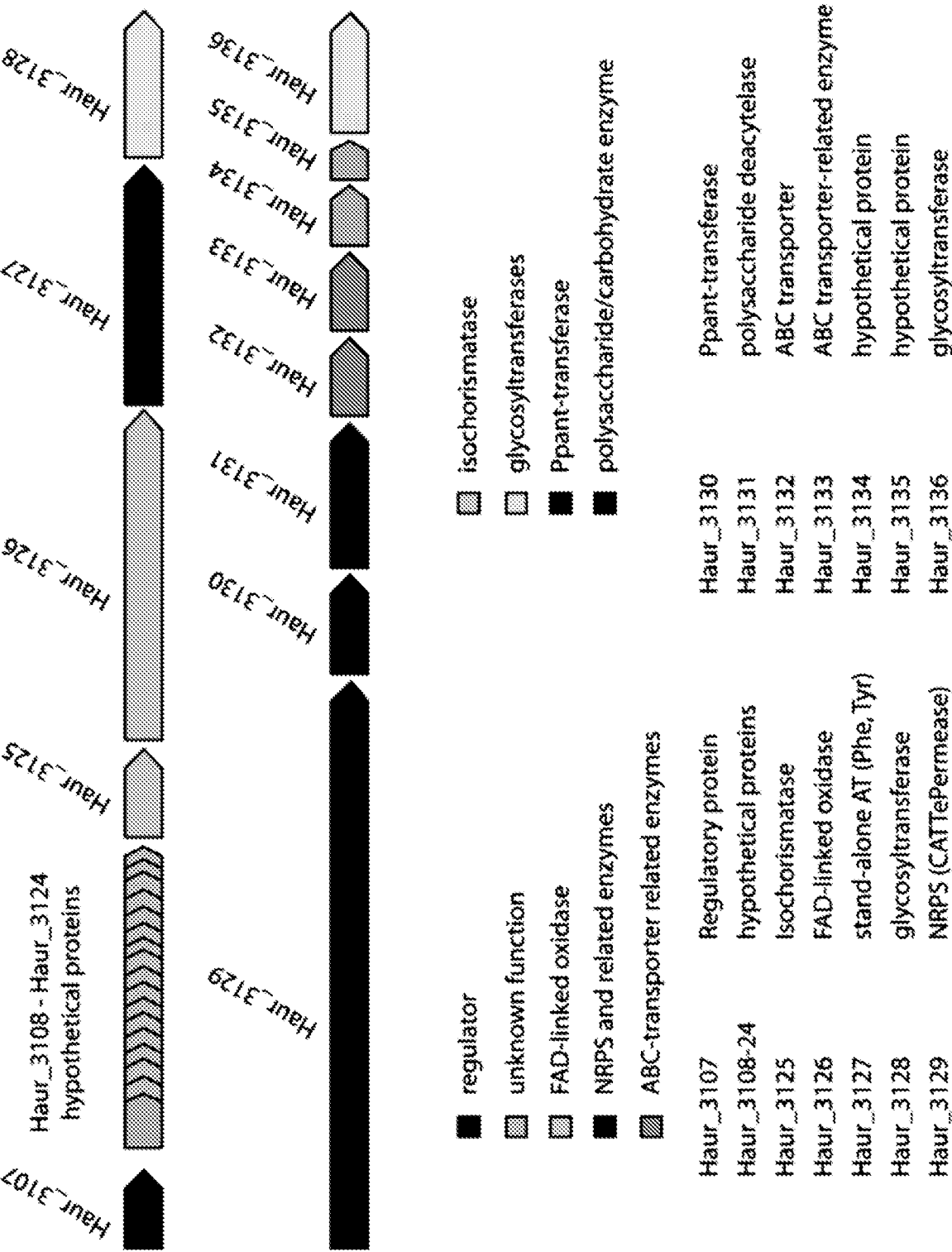


FIG. 5 Gram-Positive strain (*B. subtilis* PY79) Gram-negative strain (*E. coli* K12)





A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 31/351 (2014.01)

USPC - 540/451

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8) - A61K 31/351, 31/7052; A61P 31/04 (2014.01)

USPC - 514/9; 540/451; 549/356

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
CPC - A61K 31/351, 31/7052 (2014.02)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatBase, Google Patents, STN, Google Scholar, PubChem

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6,455,507 B1 (DRACH et al) 24 September 2002 (24.09.2002) entire document	1-14, 19-25
A	US 2012/0108530 A1 (GOFF et al) 03 May 2012 (03.05.2012) entire document	1-14, 19-25
A	US 2010/0222546 A1 (CRICH et al) 02 September 2010 (02.09.2010) entire document	1-14, 19-25
A	US 5,668,272 A (PRASAD et al) 16 September 1997 (16.09.1997) entire document	1-14, 19-25
A	STRIEKER et al. Nonribosomal Peptide Synthetases: Structures and Dynamics. Current opinion in structural biology 20(2): 234-240, 2010. [retrieved on 02 May 2014]. Retrieved from the Internet. <URL: http://www.ncbi.nlm.nih.gov/pubmed/20153164 >. Abstract	1-14, 19-25

☐ Further documents are listed in the continuation of Box C.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

02 May 2014

Date of mailing of the international search report

15 MAY 2014

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 15-18, 26, 27
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.