

**POLYPEPTIDES AND METHODS FOR SYNTHESISING MYCOSPORINE-LIKE
AMINO ACIDS AND MYCOSPORINE-LIKE AMINO ACIDS SYNTHESISED
THEREWITH**

5 Technical field

The present invention relates, in general terms, to polypeptides and methods for synthesising mycosporine-like amino acids (MAAs), and in particular to engineered D-Ala-D-Ala ligases and uses thereof for synthesising MAAs.

10

Background

The mutagenic nature of excessive UV exposure has the potential to cause sunburns, premature skin aging, and cancer. Increased awareness of these harmful effects, as well as the increase in incidence of skin cancer over the last decades, have resulted in more people using sunscreens for additional UV protection. However, widely-used sunscreen compounds, such as oxybenzone, zinc oxide and titanium dioxide have negative effects on human health and the environment, including but not limited to skin inflammation, genotoxicity and coral bleaching. As an evolved protective mechanism against the harmful effects of excessive UV exposure, microorganisms have developed the ability to produce UV-absorbing compounds, and the potential use of these compounds as a natural sunscreen ingredient make them an attractive target for various cosmeceutical companies.

Mycosporine-like amino acids, or MAAs, are one such class of natural sunscreen compounds produced in cyanobacteria, fungi, and marine algae. They contain a cyclohexenone or cyclohexenimine ring, conjugated with one or two amino acids. There are currently two main roadblocks to the use of MAAs as a primary sunscreen ingredient in commercial formulations: their low yield, and their general hydrophilicity. Although MAAs can be obtained from the extracts of their natural producers, they are produced in very low amounts insufficient for commercial use, and their isolation and purification processes are highly resource-intensive. Chemical synthesis of MAAs is also currently untenable in large scales, as the process involves about 15 steps with an overall yield of only 2%. Additionally, although the hydrophilic nature of MAAs means that they will not be readily absorbed by skin, thus limiting adverse effects, this property of MAAs means that they are easily removed

from the skin when in water.

It would be desirable to overcome or alleviate at least one of the above-described problems, or at least to provide a useful alternative.

5

Summary

Disclosed herein is an engineered polypeptide that is distinguished from a wild-type D-Ala-D-Ala ligase (Ddl) by at least one amino acid substitution, wherein the at least one amino acid substitution is at a position in the wild-type Ddl corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1.

Disclosed herein is an isolated polypeptide comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Disclosed herein is the use of a polypeptide as defined herein for synthesising a mycosporine-like amino acid (MAA).

Disclosed herein is a polynucleotide encoding a polypeptide as defined herein.

Disclosed herein is an expression vector comprising a polynucleotide as defined herein.

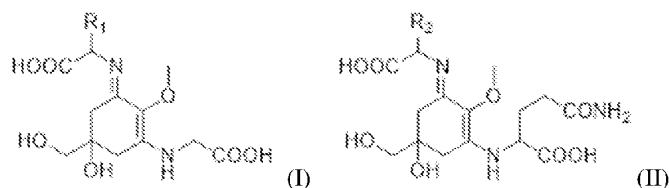
Disclosed herein is a host cell comprising a polynucleotide or an expression vector as defined herein.

Disclosed herein is a method for producing a polypeptide as defined herein, comprising culturing a host cell as defined herein under conditions suitable for expressing the polypeptide.

Disclosed herein is a method for synthesising a mycosporine-like amino acid (MAA), the method comprising culturing a host cell that expresses a polypeptide as defined herein under conditions for synthesis of the MAA by the host cell.

Disclosed herein is a mycosporine-like amino acid (MAA) obtained by a method as defined herein.

- 5 Disclosed herein is a mycosporine-like amino acid (MAA), wherein the MAA is a compound of formula (I) or (II):



- wherein R₁ is the side chain of leucine, isoleucine, methionine, asparagine or glutamine; and R₂ is the side chain of glycine, alanine, leucine, isoleucine, serine, threonine, cysteine,
 10 methionine, histidine, glutamine, glutamic acid or phenylalanine.

Disclosed herein is a UV-absorbing composition comprising one or more MAAs as defined herein.

- 15 Disclosed herein is a pharmaceutical composition comprising one or more MAAs as defined herein.

Disclosed herein is a mycosporine-like amino acid (MAA) as defined herein, for use as a medicament.

20

Disclosed herein is a method of treating a skin disease or condition in a subject in need thereof, the method comprising administering an MAA or a pharmaceutical composition as defined herein to the subject.

- 25 Disclosed herein is an MAA or a pharmaceutical composition as defined herein, for use in the treatment of a skin disease or condition in a subject in need thereof.

Disclosed herein is the use of an MAA or a pharmaceutical composition as defined herein in the manufacture of a medicament for the treatment of a skin disease or condition in a

- 30 subject in need thereof.

Brief description of the drawings

Embodiments of the present invention will now be described, by way of non-limiting
5 example, with reference to the drawings in which:

Figure 1 shows the biosynthetic pathway of MAA production, together with the structures of the intermediates and possible products. Fungal ATP-grasp ligases (ATPGLs) attach glutamine to 4-deoxygadusol to form mycosporine-glutamine, while non-fungal
10 counterparts attach glycine to form mycosporine-glycine. Various NRPS-like enzymes or Ddl ligases act upon the mycosporine glycine substrate and add a second amino acid to the structure. B) The workflow of screening putative MAA-producing enzymes. Putative ATPGL and Ddl enzymes are cloned into separate plasmids, transformed into the base DDGS-OMT strain of *S. cerevisiae*. Successful transformants are analysed for MAA
15 production via UV-visible spectrometry and mass spectrometry.

Figure 2 shows the mass spectrometry profiles of wild-type SpDdl1 and SpDdl2, as well as the swap mutants SpDdl12 and SpDdl21. The smaller blocks in the swap mutants indicate the swapped residues, where the asterisks indicate the residues predicted to be crucial based
20 on the AlphaFold structures, the yellow blocks indicate residues that are conserved in M2G producers, and the green blocks indicate residues conserved in P334 producers.

Figure 3 shows the proportion of the various MG-X products produced by the wild-type SpDdls and the SpDdl1 variants. The total fraction of mycosporine-glycine converted to an
25 MG-X product is reflected by the heights of the bars. Only the top 6 MG-X products observed are individually reflected in the chart, 'others' comprise: MG-isoleucine, MG-leucine, MG-methionine, MG-valine, MG-glutamate, and dehydroxy-porphyrin-334.

Figure 4 shows the proposed structures of novel mycosporine-glycine-X MAAs produced
30 by SpDdl1 Y253X mutants. Those marked with a red asterisk contain amino acid substituents which are more hydrophobic.

Figure 5 shows the *in vitro* catalytic efficiencies of SpDdl1, SpDdl2, SpDdl1 Y253L, Y253I and Y253A with the amino acids they were found to be most active with *in vivo*.

Figure 6 shows that wild-type SpDdl1 exhibits an increase in absorbance at 330 nm with time, in a reaction mixture including mycosporine-glycine, glycine, and ATP. Meanwhile, the mutant defective in ATP-binding, SpDdl1 D305A, does not exhibit any activity.

5

Figure 7 shows the proportion of mycosporine-glutamine-X (MGln-X) products produced by each original MGln mutant (OM). MGln-glycine and MGln-alanine are the most common products.

10 **Figure 8** shows the proportion of mycosporine-glutamine-X (MGln-X) products produced by additional MGln mutants. 'Others' include MGln-Met, MGln-Val, MGln-Cys, MGln-Glu, MGln-Ser, MGln-Gln, MGln-His, MGln-Asn and MGln-Lys.

Figure 9 shows the proposed structures of novel mycosporine-glutamine-X MAAs produced
15 by SpDdl1 mutants.

Figure 10 shows mass spectrometry profiles and UV absorbance spectra for: (a) mycosporine-glutamine-glycine and mycosporine-glycine-glutamine, (b) mycosporine-glutaminc-phenylalanine, (c) mycosporine-glutaminc-scrinc, (d) mycosporine-glutaminc-histidine, (e) mycosporine-glutamine-glutamine, and (f) mycosporine-glutamine-asparagine.
20

Figure 11 shows UV absorbance spectra for: (a) mycosporine-glutamine-valine, (b) mycosporine-glutamine-cysteine, (c) mycosporine-glutamine-alanine, (d) mycosporine-glutamine-leucine or mycosporine-glutamine-isoleucine, (e) mycosporine-glutamine-methionine, (f) mycosporine-glutaminc-threonine, and (g) mycosporine-glycine-methionine.
25

Detailed description

30

The inventors have discovered and characterised new D-Ala-D-Ala ligases from the stony coral *Stylophora pistillata* (SpDdls), which are able to synthesise mycosporine-like amino acids (MAAs) using mycosporine-glycine as substrate. The inventors also discovered that a single amino acid substitution in one of these SpDdls can surprisingly alter substrate

- specificity and increase substrate promiscuity, thus enabling the production of hitherto unknown mycosporine-glycine-X compounds (where X is a second amino acid). Further engineering of the enzyme using an unbiased mutagenesis approach uncovered additional positions where substitutions or deletions can produce gain-of-function variants that are able
- 5 to recognise mycosporine-glutamine as a substrate. Surprisingly, none of these positions are located near the predicted binding pocket of the mycosporine substrate. The engineered enzymes are able to synthesise a new class of mycosporine-glutamine-X compounds which are not previously accessible enzymatically.
- 10 The new MAAs that can be synthesised using the engineered Ddls include several that contain hydrophobic amino acids (e.g., phenylalanine or leucine attached to the central cyclohexenone structure). The increased hydrophobicity of these MAAs potentially makes them more water-resistant than many of the known MAAs, thus making them advantageous for use as natural UV-protective agents in sunscreens.
- 15 Accordingly, this disclosure provides wild-type and engineered D-Ala-D-Ala ligases (Ddls), methods of producing the Ddls, and use of the Ddls for synthesising mycosporine-like amino acids (MAAs), in particular novel MAAs based on mycosporine-glycine (MGly) and mycosporine-glutamine (MGln). Also provided are compositions containing novel MGly-
- 20 and MGln-based MAAs, and their use as UV absorbents or as medicaments.

General definitions

- The terms “polypeptide”, “proteinaceous molecule” and “protein” are used interchangeably
- 25 herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues is a synthetic non-naturally-occurring amino acid, such as a chemical analogue of a corresponding naturally-occurring amino acid, as well as to naturally-occurring amino acid polymers. These terms do not exclude modifications, for example, glycosylation,
- 30 acetylation, phosphorylation and the like. Soluble forms of the subject polypeptides are particularly useful. Included within the definition are, for example, polypeptides containing one or more analogues of an amino acid including, for example, unnatural amino acids or polypeptides with substituted linkages.

A polypeptide herein may comprise a peptide or polypeptide fusion partner at its N- or C-terminus. The fusion partner may be added for any purpose, including but not limited to facilitating purification, improving manufacturability, enhancing protein biophysical properties (e.g., solubility or stability), enabling protein labelling for identification, etc. For instance, a hexahistidine tag may be added to enable affinity chromatography during protein purification, or a solubility tag may be added to enhance protein solubility and/or stability during expression and purification.

As used herein “sequence identity” refers to the number (or fraction expressed as a percentage %) of identical or similar amino acids or nucleotide bases in a comparison between a test and a reference polypeptide or polynucleotide. Sequence identity can be determined by sequence alignment of nucleic acid or protein sequences to identify regions of similarity or identity. For purposes herein, sequence identity is generally determined by alignment to identify identical residues. Alignment can be local or global, but for purposes herein alignment is generally a global alignment where the full-length of each sequence is compared. Matches, mismatches and gaps can be identified between compared sequences. Gaps are null amino acids or nucleotides inserted between the residues of aligned sequences so that identical or similar characters are aligned. Generally, there can be internal and terminal gaps. Sequence identity can be determined by taking into account gaps as the number of identical residues/length of the shortest sequence $\times 100$. When using gap penalties, sequence identity can be determined with no penalty for end gaps (e.g., terminal gaps are not penalised). Alternatively, sequence identity can be determined without taking into account gaps as the number of identical positions/length of the total aligned sequence $\times 100$.

Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are known to those skilled in the art, for instance, using publicly available computer software available on internet web sites such as <http://blast.ncbi.nlm.nih.gov>, or <https://www.ebi.ac.uk/jdispatcher/msa/clustalo>, or <https://www.ebi.ac.uk/jdispatcher/msa/mafft>. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

As used herein, the phrase “consisting essentially of” in the context of a recited subunit sequence (e.g., amino acid sequence) indicates that the sequence may comprise at least one

- additional upstream subunit (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more upstream subunits; e.g., amino acids) and/or at least one additional downstream subunit (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more upstream subunits; e.g., amino acids), wherein the number of upstream subunits and the number of downstream subunits are independently selectable.
- 10 By “corresponding to” is intended the amino acid position relative to that position in SEQ ID NO: 1 when two (or more) sequences are aligned using standard alignment algorithms described elsewhere herein. By aligning the sequences, one skilled in the art can identify corresponding residues, for example, using conserved and identical amino acid residues as guides.
- 15 The terms “polynucleotide” and “nucleic acid,” used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, terms “polynucleotide” and “nucleic acid” encompass single-stranded DNA (ssDNA); double-stranded DNA (dsDNA); multi-stranded DNA; single-stranded RNA (ssRNA);
- 20 double-stranded RNA (dsRNA); multi-stranded RNA; genomic DNA; cDNA; DNA-RNA hybrids; and a polymer comprising purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases.
- By “expression vector” or “vector” is meant a nucleic acid molecule, preferably a DNA
- 25 molecule derived, for example, from a plasmid, virus, bacteriophage, or plant virus, into which a nucleic acid sequence may be inserted or cloned. A vector preferably contains one or more unique restriction sites and may be capable of autonomous replication in a defined host cell including a target cell or tissue or a progenitor cell or tissue thereof, or be integrable with the genome of the defined host such that the cloned sequence is reproducible.
- 30 Accordingly, the vector may be an autonomously replicating vector, i.e., a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a linear or closed circular plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into

- the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. A vector system may comprise a single vector or plasmid, two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The choice of the vector will typically depend
- 5 on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may also include a selection marker such as an antibiotic resistance gene that can be used for selection of suitable transformants. Examples of such resistance genes are well known to those of skill in the art.
- 10 An “expression cassette” comprises a DNA coding sequence operably linked to a promoter. “Operably linked” refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression.
- 15 “Recombinant” refers to polynucleotides synthesised or otherwise manipulated *in vitro* (“recombinant polynucleotides”) and to methods of using recombinant polynucleotides to produce gene products encoded by those polynucleotides in cells or other biological systems. For example, a cloned polynucleotide may be inserted into a suitable expression vector, such as a bacterial plasmid, and the plasmid can be used to transform a suitable host cell. A host
- 20 cell that comprises the recombinant polynucleotide is referred to as a “recombinant host cell”. The gene is then expressed in the recombinant host cell to produce, e.g., a “recombinant protein”. A recombinant polynucleotide may serve a non-coding function (e.g., promoter, origin of replication, ribosome-binding site, etc.) as well.
- 25 The terms “wild-type protein” or “parent protein” are used interchangeably herein and refer to the non-mutated version of a polypeptide as it appears naturally. The terms “mutant”, “variant”, “engineered polypeptide” and “engineered protein” are used interchangeably herein to refer to a polypeptide derived from a wild-type protein and comprising one or more amino acid modifications, e.g., an amino acid substitution, insertion and/or deletion. The
- 30 variants may be obtained by various techniques well known in the art, e.g., site-directed mutagenesis, random mutagenesis and synthetic oligonucleotide construction.

The term “modification” or “alteration” as used herein in relation to a position in a polypeptide sequence or an amino acid means that the amino acid in the particular position has been modified compared to the amino acid of the wild-type protein.

- 5 A “substitution” means that an amino acid residue is replaced by another amino acid residue. An amino acid residue may be replaced by another selected from the naturally-occurring standard 20 amino acid residues, rare naturally occurring amino acid residues (e.g. hydroxyproline, hydroxylysine, allohydroxylysine, 6-N-methyllysine, N-ethylglycine, N-methylglycine, N-ethylasparagine, allo-isoleucine, N-methylisoleucine, N-methylvaline,
- 10 pyroglutamine, aminobutyric acid, ornithine, norleucine, norvaline), and non-naturally occurring amino acid residue, often made synthetically, e.g. cyclohexyl-alanine. A substitution herein is preferably the replacement of an amino acid residue by another selected from the naturally-occurring standard 20 amino acid residues (G, P, A, V, L, I, M, C, F, Y, W, H, K, R, Q, N, E, D, S and T). The sign “+” indicates a combination of substitutions. The
- 15 following terminology is used herein to designate a substitution: Y253L denotes that the amino acid residue at position 253 (tyrosine, Y) of the parent sequence is changed to a leucine (L). Y253L/V denotes that the amino acid residue at position 253 (Y) of the parent sequence is substituted with either a leucine (L) or a valine (V).
- 20 A “conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, which can be generally sub-classified as follows:

Amino acid sub-classification

Sub-classes	Amino acids
Acidic	Aspartic acid, Glutamic acid
Basic	Noncyclic: Arginine, Lysine; Cyclic: Histidine
Charged	Aspartic acid, Glutamic acid, Arginine, Lysine, Histidine
Small	Glycine, Serine, Alanine, Threonine, Proline
Polar/neutral	Asparagine, Histidine, Glutamine, Cysteine, Serine, Threonine
Polar/large	Asparagine, Glutamine
Hydrophobic	Tyrosine, Valine, Isoleucine, Leucine, Methionine, Phenylalanine, Tryptophan
Aromatic	Tryptophan, Tyrosine, Phenylalanine
Residues that influence chain orientation	Glycine and Proline

Conservative amino acid substitution also includes groupings based on side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. For example, it is reasonable to expect that replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid will not have a major effect on the properties of the resulting variant polypeptide. Whether an amino acid change results in a functional polypeptide can readily be determined by assaying its activity. Conservative substitutions are shown in the table below under the heading of exemplary substitutions. Amino acid substitutions falling within the scope of the invention, are, in general, accomplished by selecting substitutions that do not differ significantly in their effect on maintaining (a) the structure of the peptide backbone in the area of the substitution, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. After the substitutions are introduced, the variants are screened for biological activity.

Exemplary amino acid substitutions

Original Residue	Exemplary Substitutions
Ala	Val, Leu, Ile
Arg	Lys, Gln, Asn
Asn	Gln, His, Lys, Arg
Asp	Glu
Cys	Ser
Gln	Asn, His, Lys,
Glu	Asp, Lys
Gly	Pro
His	Asn, Gln, Lys, Arg
Ile	Leu, Val, Met, Ala, Phe
Leu	Ile, Val, Met, Ala, Phe
Lys	Arg, Gln, Asn
Met	Leu, Ile, Phe
Phe	Leu, Val, Ile, Ala
Pro	Gly
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe, Thr, Ser
Val	Ile, Leu, Met, Phe, Ala

The term “mycosporine-like amino acid (MAA)” refers to a low molecular weight compound containing a 6-carbon ring structure (e.g., a cyclohexenone or cyclohexenimine ring), substituted with at least one amino acid moiety. An MAA containing one N-substituent (e.g., one amino acid moiety) on a cyclohexenone ring is also referred to as a mono-substituted MAA or oxo-MAA herein. Exemplary mono-substituted MAAs include mycosporine-glycine (MG) and mycosporine-glutamine (MGln). An MAA containing two N-substituents (e.g., two amino acid moieties) on a cyclohexenimine ring is also referred to as a di-substituted MAA or imino-MAA herein. Exemplary di-substituted MAAs include mycosporine-glycine-glycine (mycosporine-2-glycine), mycosporine-glycine-serine (shinorine), mycosporine-glycine-threonine (porphyra-334) and mycosporine-glutamine-alanine. MAAs are generally capable of absorbing UV light (e.g., light of wavelength 310 nm to 360 nm) and generally exhibit photoprotective functionality.

The term “UV” refers to light between the wavelengths of 280–400 nm. This is divided between “UV-A” and “UV-B”. “UV-A” refers to light between the wavelengths of 315–400 nm. “UV-B” refers to light between the wavelengths of 280–315 nm.

- 5 The terms “D-Alanyl-D-Alanine ligase”, “D-Alanine-D-Alanine ligase” or “D-Ala-D-Ala ligase” are used interchangeably herein to refer to an enzyme which catalyses the addition of a single amino acid to a mono-substituted MAA. The amino acid is preferably one of the 20 canonical amino acids but can also include α -amino acids (e.g., ornithine, homoserine or α -aminobutyric acid), β -amino acids, γ -amino acids and amino sulphonic acids (e.g.,
- 10 taurine). The D-Ala-D-Ala ligase may be capable of adding different amino acids to a mono-substituted MAA to form a family of di-substituted MAA compounds. For the avoidance of doubt, a D-Ala-D-Ala ligase herein is not restricted to catalysing the addition of alanine to a mono-substituted MAA or to mycosporine-alanine.
- 15 The terms “disruption” and “disrupted” are used interchangeably herein to refer to any genetic modification that decreases or eliminates expression and/or functional activity of a nucleic acid or an expression product thereof. For example, disruption of a gene includes within its scope any genetic modification that decreases or eliminates expression of the gene and/or the functional activity of a corresponding gene product (e.g., mRNA and/or protein).
- 20 Genetic modifications include complete or partial inactivation, suppression, deletion, interruption, blockage, or down-regulation of a nucleic acid (e.g., a gene). Illustrative genetic modifications include, but are not limited to, gene knockout, inactivation, mutation (e.g., insertion, deletion, point, or frameshift mutations that disrupt the expression or activity of the gene product), or use of inhibitory nucleic acids (e.g., inhibitory RNAs such as sense or
- 25 antisense RNAs, molecules that mediate RNA interference such as siRNA, shRNA, miRNA; etc.), inhibitory polypeptides (e.g., antibodies, polypeptide-binding partners, dominant negative polypeptides, enzymes etc.) or any other molecule that inhibits the activity of the gene or level or functional activity of an expression product of the gene.
- 30 By “pharmaceutically acceptable carrier” is meant a pharmaceutical vehicle comprised of a material that is not biologically or otherwise undesirable, i.e., the material may be administered to a subject along with the selected active agent without causing any or a substantial adverse reaction. Carriers may include excipients and other additives such as diluents, detergents, colouring agents, wetting or emulsifying agents, pH buffering agents,

preservatives, and the like.

The terms “treating” and “preventing” include: delaying or preventing the onset of symptoms of the disease, disorder or condition; reducing the severity of (i.e., alleviating) the symptoms of the disease, disorder or condition; reversing the symptoms of (i.e., ameliorating) the disease, disorder or condition; reducing morbidity of subjects having the disease, disorder or condition; reducing mortality of subjects having the disease, disorder or condition; delaying or preventing progression of the disease, disorder or condition (e.g., to a later stage); and/or otherwise inhibiting the symptoms or effects of the disease, disorder or condition for at least a period of time. It is to be understood that the terms “treating” and “preventing” do not imply that the disease, disorder or condition, or a symptom or effect thereof, is permanently delayed, reduced, alleviated, ameliorated or otherwise inhibited and therefore also encompasses the temporary delay, reduction, alleviation, amelioration or otherwise inhibition of the disease, disorder or condition, or a symptom or effect thereof.

The term “administering” refers to contacting, applying, injecting, transfusing or providing an inhibitor as referred to herein to a subject.

The terms “subject”, “patient”, “host” or “individual” used interchangeably herein, refer to any subject, particularly a vertebrate subject, and even more particularly a mammalian subject, for whom therapy or prophylaxis is desired. Suitable vertebrate animals that fall within the scope of the invention include, but are not restricted to, any member of the subphylum Chordata including primates (e.g., humans, monkeys and apes, and includes species of monkeys such as from the genus *Macaca* (e.g., cynomolgus monkeys such as *Macaca fascicularis*, and/or rhesus monkeys (*Macaca mulatta*)) and baboon (*Papio ursinus*), as well as marmosets (species from the genus *Callithrix*), squirrel monkeys (species from the genus *Saimiri*) and tamarins (species from the genus *Saguinus*), as well as species of apes such as chimpanzees (*Pan troglodytes*)), rodents (e.g., mice rats, guinea pigs), lagomorphs (e.g., rabbits, hares), bovinos (e.g., cattle), ovines (e.g., sheep), caprines (e.g., goats), porcines (e.g., pigs), equines (e.g., horses), canines (e.g., dogs), felines (e.g., cats), avians (e.g., chickens, turkeys, ducks, geese, companion birds such as canaries, budgerigars etc.), marine mammals (e.g., dolphins, whales), reptiles (snakes, frogs, lizards etc.), and fish. In one embodiment, the subject is a human subject.

By "effective amount", in the context of treating or preventing a condition is meant the administration of an amount of an agent or composition to an individual in need of such treatment or prophylaxis, either in a single dose or as part of a series, that is effective for the prevention of incurring a symptom, holding in check such symptoms, and/or treating existing
5 symptoms, of that condition. The effective amount will vary depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated, the formulation of the composition, the assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

10

As used herein, "and/or" refers to and encompasses any and all possible combinations of one or more of the associated listed items, as well as the lack of combinations when interpreted in the alternative (or).

15 As used in this application, the singular form "a," "an," and "the" include plural references unless the context clearly dictates otherwise. For example, the term "an agent" includes a plurality of agents, including mixtures thereof.

Throughout this specification and the claims which follow, unless the context requires
20 otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

Throughout this specification and the claims which follow, unless the context requires
25 otherwise, the phrase "consisting essentially of", and variations such as "consists essentially of" will be understood to indicate that the recited element(s) is/are essential i.e. necessary elements of the invention. The phrase allows for the presence of other non-recited elements which do not materially affect the characteristics of the invention but excludes additional unspecified elements which would affect the basic and novel characteristics of the method
30 defined.

D-Ala-D-Ala ligases

Polypeptides and engineered polypeptides herein are generally capable of catalysing the

addition of an amino acid to a mono-substituted MAA to form a di-substituted MAA.

Disclosed herein is an isolated polypeptide comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 2. The polypeptide may comprise an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In some embodiments, the polypeptide comprises or consists essentially of the amino acid sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In some embodiments, the polypeptide comprises an amino acid substitution at one or more positions corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1. In one embodiment, the polypeptide comprises an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1, wherein the amino acid at one or more positions corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1 has been substituted. In one embodiment, the polypeptide comprises an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 2, wherein the amino acid at one or more positions corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1 has been substituted.

Also disclosed herein is an engineered polypeptide that is distinguished from a wild-type D-Ala-D-Ala ligase (Ddl) by at least one amino acid substitution, wherein the at least one amino acid substitution is at a position in the wild-type Ddl corresponding to position 253, 28, 65, 66, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1.

30

In some embodiments, the wild-type D-Ala-D-Ala ligase (Ddl) comprises an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1 or 2. In one embodiment, the wild-type D-Ala-D-Ala ligase (Ddl) comprises or consists essentially of an amino acid sequence set forth in SEQ ID NO: 1 or 2.

The engineered polypeptide may contain 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 or 23 amino acid substitutions at one or more positions in the wild-type Ddl corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1. The engineered polypeptide may contain at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21 or at least 22 amino acid substitutions at one or more positions in the wild-type Ddl corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1.

In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 253 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 28 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 65 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 66 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 67 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 68 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 69 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 70 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 71 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 72 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 73 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 74 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 75 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding

to position 76 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 77 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 78 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 79 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 137 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 189 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 227 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 230 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 270 of SEQ ID NO: 1. In one embodiment, the engineered polypeptide contains an amino acid substitution at a position corresponding to position 357 of SEQ ID NO: 1.

In some embodiments, the amino acid substitution in the engineered polypeptide is at a position in the wild-type Ddl corresponding to one of the following positions or sets of positions of SEQ ID NO: 1:

- (a) a substitution at position 253;
- (b) a substitution at position 137;
- (c) a substitution at positions 28 and 189;
- (d) a substitution at positions 66 and 137;
- (e) a substitution at positions 69 and 357;
- (f) a substitution at positions 227 and 230;
- (g) a substitution at positions 65, 66 and 69;
- (h) a substitution at positions 65, 66, 69 and 79;
- (i) a substitution at each of positions 65 to 69;
- (j) a substitution at positions 65, 79, 227, 230 and 270;
- (k) a substitution at positions 65, 66, 69, 79 and 137;
- (l) a substitution at positions 65, 66, 69, 79 and 357;
- (m) a substitution at positions 65, 66, 69, 79, 137 and 189;
- (n) a substitution at positions 65, 66, 69, 79, 137 and 253;

- (o) a substitution at positions 65, 66, 69, 79, 137 and 357;
- (p) a substitution at positions 65, 66, 69, 79, 253 and 357;
- (q) a substitution at positions 65, 79, 227, 230, 253 and 270;
- (r) a substitution at positions 28, 65, 66, 69, 79, 137, 189, 227, 230, 270 and 357;
- 5 (s) a substitution at positions 28, 65, 66, 69, 79, 137, 189, 227, 230, 253, 270 and 357; or
- (t) a substitution at each of positions 65 to 79.

In some embodiments, the polypeptide comprises one or more amino acid substitutions selected from the following:

- 10 (a) a substitution to A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W or Y at a position corresponding to position 253 of SEQ ID NO: 1;
- (b) a substitution to A, I, L or V at a position corresponding to position 28 of SEQ ID NO: 1;
- (c) a substitution to G or P at a position corresponding to position 65 of SEQ ID NO: 1;
- 15 (d) a substitution to F, I, L or M at a position corresponding to position 66 of SEQ ID NO: 1;
- (e) a substitution to I, L or V at a position corresponding to position 67 of SEQ ID NO: 1;
- (f) a substitution to I, L or V at a position corresponding to position 68 of SEQ ID NO: 1;
- (g) a substitution to G or P at a position corresponding to position 69 of SEQ ID NO: 1;
- 20 (h) a substitution to A, I, L or V at a position corresponding to position 70 of SEQ ID NO: 1;
- (i) a substitution to A, I, L or V at a position corresponding to position 71 of SEQ ID NO: 1;
- (j) a substitution to A, I, L or V at a position corresponding to position 72 of SEQ ID NO: 1;
- 25 (k) a substitution to A, I, L or V at a position corresponding to position 73 of SEQ ID NO: 1;
- (l) a substitution to A, I, L or V at a position corresponding to position 74 of SEQ ID NO: 1;
- 30 (m) a substitution to A, I, L or V at a position corresponding to position 75 of SEQ ID NO: 1;
- (n) a substitution to A, I, L or V at a position corresponding to position 76 of SEQ ID NO: 1;
- (o) a substitution to A, I, L or V at a position corresponding to position 77 of SEQ ID NO: 1;

- 1;
- (p) a substitution to A, I, L or V at a position corresponding to position 78 of SEQ ID NO: 1;
- (q) a substitution to A, F, I, L, M or V at a position corresponding to position 79 of SEQ ID NO: 1;
- 5 (r) a substitution to F, I, L, M, S or T at a position corresponding to position 137 of SEQ ID NO: 1;
- (s) a substitution to G or P at a position corresponding to position 189 of SEQ ID NO: 1;
- (t) a substitution to A, I, L or V at a position corresponding to position 227 of SEQ ID NO: 1;
- 10 (u) a substitution to G or P at a position corresponding to position 230 of SEQ ID NO: 1;
- (v) a substitution to K, N, Q or R at a position corresponding to position 270 of SEQ ID NO: 1; and
- (w) a substitution to D, E or K at a position corresponding to position 357 of SEQ ID NO: 1.
- 15

In some embodiments, the polypeptide comprises one or more amino acid substitutions selected from the following:

- (a) a substitution to A, C, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1;
- 20 (b) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1;
- (c) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1;
- (d) a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1;
- (e) a substitution to A at a position corresponding to position 67 of SEQ ID NO: 1;
- 25 (f) a substitution to A at a position corresponding to position 68 of SEQ ID NO: 1;
- (g) a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1;
- (h) a substitution to A at a position corresponding to position 70 of SEQ ID NO: 1;
- (i) a substitution to A at a position corresponding to position 71 of SEQ ID NO: 1;
- (j) a substitution to A at a position corresponding to position 72 of SEQ ID NO: 1;
- 30 (k) a substitution to A at a position corresponding to position 73 of SEQ ID NO: 1;
- (l) a substitution to A at a position corresponding to position 74 of SEQ ID NO: 1;
- (m) a substitution to A at a position corresponding to position 75 of SEQ ID NO: 1;
- (n) a substitution to A at a position corresponding to position 76 of SEQ ID NO: 1;
- (o) a substitution to A at a position corresponding to position 77 of SEQ ID NO: 1;

- (p) a substitution to A at a position corresponding to position 78 of SEQ ID NO: 1;
 - (q) a substitution to A or L at a position corresponding to position 79 of SEQ ID NO: 1;
 - (r) a substitution to M or T at a position corresponding to position 137 of SEQ ID NO: 1;
 - (s) a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1;
 - 5 (t) a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1;
 - (u) a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1;
 - (v) a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1; and
 - (w) a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1.
- 10 In some embodiments, the polypeptide comprises one of the following:
- (a) a substitution to A, C, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1;
 - (b) a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1;
 - (c) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1, and a
 - 15 substitution to G at a position corresponding to position 189 of SEQ ID NO: 1;
 - (d) a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, and a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1;
 - (e) a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
 - 20 (f) a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1 and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
 - (g) a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, and a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1;
 - (h) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a
 - 25 substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, and a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1;
 - (i) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, and a
 - 30 substitution to L at a position corresponding to position 79 of SEQ ID NO: 1;
 - (j) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a

- substitution to L at a position corresponding to position 79 of SEQ ID NO: 1; and a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1;
- (k) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1; and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- (l) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1, and a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1;
- (m) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1, and a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1;
- (n) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1, and a substitution to I at a position corresponding to position 253 of SEQ ID NO: 1;
- (o) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- (p) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1,

substitution to C, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;

- 5 (q) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1, a substitution to A, C, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1, and a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1;
- 10 (r) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1, a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1, a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- 15 (s) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1, a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1, a substitution to A, C, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1, a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- 20
- 25
- 30

- (t) a substitution to A at a position corresponding to position 70 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 71 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 72 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 73 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 74 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 75 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 76 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 77 of SEQ ID NO: 1; a substitution to A at a position corresponding to position 78 of SEQ ID NO: 1, and a substitution to A at a position corresponding to position 79 of SEQ ID NO: 1.

In some embodiments, the polypeptide comprises one of the following:

- (a) a substitution to A, C or G at a position corresponding to position 253 of SEQ ID NO: 1;
- (b) a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1;
- (c) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1, and a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1;
- (d) a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, and a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1;
- (e) a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- (f) a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1 and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- (g) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, and a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1;
- (h) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1; and a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1;
- (i) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a

- substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1; and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- (j) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1, and a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1;
- (k) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1, and substitution to I at a position corresponding to position 253 of SEQ ID NO: 1; or
- (l) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1; or
- (m) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, substitution to G, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1.

In some embodiments, the polypeptide comprises one of the following amino acid substitutions:

- (a) Y253A;
- (b) Y253C;
- (c) Y253G;
- (d) I137M;

- (e) G28A and D189G;
- (f) L66M and I137T;
- (g) L69P and K357E;
- (h) I137T and K357E;
- 5 (i) T227A and E230G;
- (j) D65G, L66M and L69P;
- (k) D65G, L66M, L69P and M79L;
- (l) D65G, L66M, L69P, M79L and I137T;
- (m) D65G, L66M, L69P, M79L and K357E;
- 10 (n) D65G, M79L, T227A, E230G and T270R;
- (o) D65G, L66M, L69P, M79L, I137M and D189G;
- (p) D65G, L66M, L69P, M79L, I137M and Y253I;
- (q) D65G, L66M, L69P, M79L, I137T and K357E;
- (r) D65G, L66M, L69P, M79L, Y253C and K357E;
- 15 (s) D65G, L66M, L69P, M79L, Y253G and K357E;
- (t) D65G, L66M, L69P, M79L, Y253I and K357E;
- (u) D65G, L66M, L69P, M79L, Y253L and K357E;
- (v) D65G, M79L, T227A, E230G, Y253A and T270R;
- (w) D65G, M79L, T227A, E230G, Y253C and T270R;
- 20 (x) D65G, M79L, T227A, E230G, Y253G and T270R;
- (y) D65G, M79L, T227A, E230G, Y253I and T270R;
- (z) D65G, M79L, T227A, E230G, Y253L and T270R;
- (aa) G28A, D65G, L66M, L69P, M79L, I137M, D189G, T227A, E230G, T270R and K357E;
- 25 (bb) G28A, D65G, L66M, L69P, M79L, I137M, D189G, T227A, E230G, Y253A, T270R and K357E;
- (cc) G28A, D65G, L66M, L69P, M79L, I137M, D189G, T227A, E230G, Y253C, T270R and K357E;
- (dd) G28A, D65G, L66M, L69P, M79L, I137M, D189G, T227A, E230G, Y253G, T270R and K357E;
- 30 (ee) G28A, D65G, L66M, L69P, M79L, I137M, D189G, T227A, E230G, Y253I, T270R and K357E;
- (ff) G28A, D65G, L66M, L69P, M79L, I137M, D189G, T227A, E230G, Y253L, T270R and K357E; or

(gg) E70A, E71A, T72A, E73A, N74A, S75A, V76A, K77A, K78A, M79A.

In some embodiments, the polypeptide comprises one of the following amino acid substitutions:

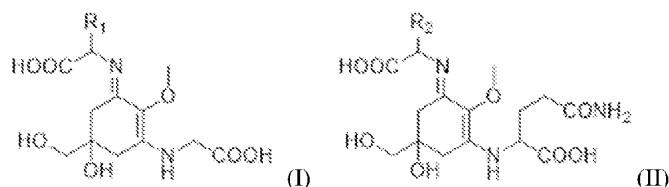
- 5 (a) Y253A;
- (b) Y253C;
- (c) Y253G;
- (d) I137M;
- (e) G28A and D189G;
- 10 (f) L66M and I137T;
- (g) L69P and K357E;
- (h) I137T and K357E;
- (i) D65G, L66M, L69P and M79L;
- (j) D65G, L66M, L69P, M79L and I137T;
- 15 (k) D65G, L66M, L69P, M79L and K357E;
- (l) D65G, M79L, T227A, E230G and T270R;
- (m) D65G, L66M, L69P, M79L, I137M and Y253I;
- (n) D65G, L66M, L69P, M79L, I137T and K357E;
- (o) D65G, L66M, L69P, M79L, Y253C and K357E;
- 20 (p) D65G, L66M, L69P, M79L, Y253G and K357E;
- (q) D65G, L66M, L69P, M79L, Y253I and K357E; or
- (r) D65G, L66M, L69P, M79L, Y253L and K357E.

- 25 Substitutions to A, C or G at the 253 position in SEQ ID NO: 1 (or at a position corresponding to the 253 position in SEQ ID NO: 1) can produce Ddls capable of producing more hydrophobic MAAs (such as MGly-Lcu, MGly-Ile and MGly-Met).

- 30 In some embodiments, the polypeptide comprises an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 3–36. The polypeptide may comprise an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to an amino acid sequence set forth in SEQ ID NO: 3–36.

In one embodiment, the polypeptide comprises or consists essentially of an amino acid sequence set forth in SEQ ID NO: 3–36.

- In some embodiments, polypeptides and engineered polypeptides herein are capable of synthesising an MAA of formula (I) or (II):



wherein R_1 and R_2 are independently the side chain of glycine, alanine, valine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, proline, asparagine, glutamine, aspartic acid, glutamic acid, lysine, arginine, phenylalanine, tyrosine or tryptophan.

- Compounds of formula (I) are also referred to herein as mycosporine-glycine-X, MGly-X or MG-X (where X may be any of the standard 20 amino acids). Compounds of formula (II) are also referred to herein as mycosporine-glutamine-X, MGln-X (where X may be any of the standard 20 amino acids).

- In particular, polypeptides and engineered polypeptides herein may be capable of synthesising the MAAs of Figures 4 and 9, i.e., MGly-Leu, MGly-Ile, MGly-Met, MGly-Arg, MGly-Gln, MGln-Gly, MGln-Ala, MGln-Leu, MGln-Ile, MGln-Ser, MGln-Thr, MGln-Cys, MGln-Met, MGln-His, MGln-Gln, MGln-Glu and MGln-Phe.

Expression constructs

Disclosed herein is a polynucleotide encoding a polypeptide as defined herein.

- Disclosed herein is an expression vector comprising a polynucleotide as defined herein. In addition to containing a sequence encoding a D-Ala-D-Ala ligase, the expression vector may also contain sequences encoding a 2-dimethyl 4-deoxygadusol synthase (DDGS), an O-methyltransferase (OMT), and/or an ATP-grasp ligase (ATPGL). Exemplary DDGS, OMT and ATPGL proteins are those endogenous to *Nostoc punctiforme* ATCC 29133 and *Anabaena variabilis* ATCC 29413.

Codons in the expression construct may be optimised for a particular expression host, as is well known in the art. Genes in the expression construct may be operably linked to one or more promoters, e.g., each gene may be controlled by an individual promoter. The promoters can be constitutive or regulatable. The regulatable promoters can be, for example, inducible.

- 5 The promoter sequences can be derived, for example, from the host cell, from another organism, or can be synthetically derived.

Any desired promoter can be used to regulate the expression of the inserted MAA biosynthesis genes. For example, the gene(s) may be under the transcriptional control of an
10 inducible promoter, such as a promoter inducible by nutrient source, nutrient starvation, growth phase, heat shock, cold shock, oxidative stress, salt stress, environmental stress, metal concentration, specific metabolites or organic compounds, light exposure, etc. In other embodiments, the gene(s) may be under the transcriptional control of a constitutive promoter. In this way, a sustained level of transcription and, therefore, enzymatic activity of
15 the corresponding protein can be maintained during the whole period of culture. The constitutive promoter may be endogenous to the host cell. This has the advantage that no recombinant transcription factor has to be present in the host cell. The endogenous promoter is usually well-recognised by the host cell without the need to introduce further genetic
20 modifications.

Expression hosts

Disclosed herein is a host cell comprising a polynucleotide or an expression vector as defined herein. Also disclosed herein is a host cell comprising a polypeptide as defined herein.

25

The host cell may be a prokaryotic cell (e.g., bacteria or cyanobacteria) or a eukaryotic cell (e.g., yeast, fungal, microalgal, insect or mammalian cell). The polynucleotide encoding the enzymes (i.e., D-Ala-D-Ala ligase and optionally DDGS, OMT and/or ATPGL) may be present as an extrachromosomal construct (e.g., a plasmid or another type of vector), or it
30 may be integrated into the host genome by random or targeted insertion. Gene or vector introduction can be performed by appropriately selecting a known transformation method, for example, electroporation, lipofection, heat shock, microinjection or particle bombardment.

In one embodiment, the host cell is a fungal cell. In one embodiment, the fungal cell is a yeast cell, such as *Saccharomyces spp.* (e.g., *S. cerevisiae*, *S. kluyveri*, *S. diastolicus*, *S. carlsbergensis* or *S. norbensis*), *Schizosaccharomyces pombe*, *Komagataella spp.* (e.g., *K. pastoris* (also known as *Pichia pastoris*), *K. pseudopastoris* and *K. phaffii*), *Yarrowia lipolytica*, *Kluyveromyces lactis*, *Ogataea polymorpha* (also known as *Hansenula polymorpha*), and *Candida spp.* (e.g., *Candida glabrata*). In one embodiment, the yeast is *Saccharomyces cerevisiae*.

The inventors have discovered that MAAs are produced primarily from sedoheptulose 7-phosphate (S7P), which is an intermediate of the pentose phosphate pathway. Redirection of the host metabolic flux towards the pentose phosphate pathway or towards the generation of S7P can thus increase MAA production. Reducing glycolytic flux may increase the carbon flux to the pentose phosphate pathway. For example, previous studies have shown that deleting the hexokinase gene (e.g., *HXK2*) can decrease glycolytic flux and glucose metabolism in favour of MAA production in yeast. The inventors have found that deletion of *TALI* (transaldolase), *PFK1* (phosphofructokinase 1) or *PFK2* (phosphofructokinase 2) can also increase accumulation of S7P and boost MAA production by more than 8-fold in yeast.

Thus, in some embodiments, the host cell comprises a disrupted *TALI* (transaldolase 1), *PFK1* (phosphofructokinase 1), *PFK2* (phosphofructokinase 2), and/or *HXK2* (hexokinase 2) gene.

In some embodiments, the host cell is a xylose-utilising strain. Xylose catabolism contributes to S7P production, thus MAA yields may be increased in xylose-utilising hosts. The xylose-utilising host cell may express, naturally or through genetic engineering, a xylose transporter for xylose uptake and/or one or more enzymes involved in xylose metabolism. Xylose transporters include, without limitation, Gal2, Hxt7, Xlt1, Gxf1, Cs4130, and any transporters (e.g., hexose transporters) capable of transporting xylose into the host cell. Enzymes for xylose metabolism may include, but are not limited to, xylose reductase or a functionally equivalent polypeptide for reduction of xylose to xylitol; xylitol dehydrogenase or a functionally equivalent polypeptide for conversion of xylitol to xylulose; xylose isomerase or a functionally equivalent polypeptide for direct conversion of xylose to xylulose; and xylulokinase or a functionally equivalent polypeptide for phosphorylation of

xylulose to xylulose-5-phosphate (which feeds into the pentose phosphate pathway). A functional equivalent of a protein includes polypeptide fragments or variants (such as mutants or natural homologues) that exhibit one or more bioactivities (such as one or more enzyme activities) of that protein. Non-limiting examples of xylose reductases include

5 XYL1 and GRE3. Non-limiting examples of xylitol dehydrogenases include XYL2 and XYLD. Xylose isomerases include, for example, XYLA. Non-limiting examples of xylulokinases include XYL3, XYLB and XKS1. Xylose-utilising strains may be cultured in media containing a mix of xylose and other fermentable sugars (such as sucrose, glucose and/or fructose).

10

In some embodiments, the host cell expresses one or more polypeptides selected from the group consisting of a xylose transporter, a xylose reductase, a xylitol dehydrogenase, a xylose isomerase, and a xylulokinase. In some embodiments, the host cell comprises a nucleic acid molecule encoding one or more proteins selected from: a xylose transporter, a

15 xylose reductase, a xylitol dehydrogenase, a xylose isomerase, and a xylulokinase. The nucleic acid may be a genomic, extrachromosomal or episomal nucleic acid.

Methods of ligase production

20 Disclosed herein is a method for producing a polypeptide as defined herein, comprising culturing a host cell as defined herein under conditions suitable for expressing the polypeptide. The host cell may be a prokaryotic host cell (e.g., bacteria or cyanobacteria) or a eukaryotic host cell (e.g., yeast, fungal, microalgal, insect or mammalian cell).

25 The host cells may be cultured using methods well known in art, e.g., by shake flask cultivation or 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 skilled person may select suitable media and culture conditions for a given host

30 cell.

The polypeptide may be secreted extracellularly into the nutrient medium, or expressed intracellularly. Accordingly, the polypeptide can be recovered directly from the culture supernatant or from cell lysate or permeabilised cells, using any method known in the art.

For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, collection, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation. Alternatively, the polypeptide may be partially or totally purified by a variety of procedures known in the art including but not limited to

5 chromatography (e.g., ion exchange, affinity, hydrophobic and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction to obtain substantially pure polypeptides.

10 Methods of MAA synthesis

Polypeptides herein, including both isolated wild-type and engineered polypeptides, have D-Ala-D-Ala ligase activity are thus capable of synthesising a mycosporine-like amino acid (MAA). Accordingly, this disclosure also provides the use of a polypeptide as defined herein

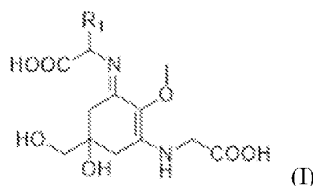
15 for synthesising a mycosporine-like amino acid (MAA).

Polypeptides herein enable enzymatic access to mycosporine-glycine-X and mycosporine-glutamine-X MAAs (where X can be any of the standard amino acids) which have not been previously described. Some of these newly accessible MAAs contain amino acids with

20 hydrophobic side groups (e.g., phenylalanine, leucine or isoleucine), which may confer greater hydrophobicity to the MAAs, making them more suitable in topical or cosmetic compositions that require a certain amount of water resistance. These MAAs may be used in a range of applications, including as UV absorbents, antioxidants, anti-aging compounds, and as therapeutic agents for inflammatory or proliferative conditions.

25

In some embodiments, polypeptides herein may be used to synthesise one or more MAA compounds of formula (I),



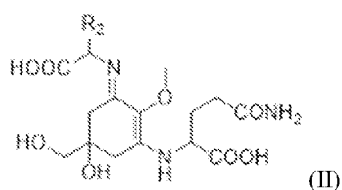
where R₁ is a side chain of leucine, isoleucine, methionine, asparagine or glutamine. The

30 enzymatic synthesis of these MAAs have not been described previously. Engineered variants

of SpDdl1 (SEQ ID NO: 1) and SpDdl2 (SEQ ID NO: 2) with an amino acid substitution at position 253 (such as polypeptides with amino acid sequence of SEQ ID NO: 3–21) are found to be capable of synthesising a variety of MGly-X compounds, including MGly-Leu, MGly-Ile, MGly-Met, MGly-Arg and MGly-Gln. Table 2 provides non-limiting examples of polypeptides that may be used for the synthesis of MGly-X compounds.

In some embodiments, a polypeptide comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1–21 is used to synthesise a mycosporine-glycine-X compound. In some embodiments, a polypeptide comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 3–21 is used to synthesise MGly-Leu, MGly-Ile, MGly-Met, MGly-Arg and/or MGly-Gln.

In some embodiments, polypeptides herein may be used to synthesise one or more MAA compound of formula (II),



where R₁ is a side chain of glycine, alanine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, glutamine, glutamic acid or phenylalanine. The enzymatic synthesis of these MAAs have not been described previously. Engineered variants of SpDdl1 (SEQ ID NO: 1) with amino acid substitutions or deletions at positions 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 253, 270 and/or 357 (such as polypeptides with amino acid sequence of SEQ ID NO: 22–36) are found to be capable of synthesising a variety of MGln-X compounds, including MGln-Gly, MGln-Ala, MGln-Leu, MGln-Ile, MGln-Ser, MGln-Thr, MGln-Cys, MGln-Met, MGln-His, MGln-Gln, MGln-Glu and MGln-Phe. FIGS. 7 and 8 provide non-limiting examples of polypeptides that may be used for the synthesis of MGln-X compounds.

In some embodiments, a polypeptide comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 22–36 is used to synthesise a mycosporine-glutamine-X compound. In some embodiments, a polypeptide

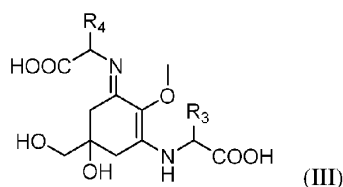
comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 22–36 is used to synthesise MGln-Gly, MGln-Ala, MGln-Leu, MGln-Ile, MGln-Ser, MGln-Thr, MGln-Cys, MGln-Met, MGln-His, MGln-Gln, MGln-Glu and/or MGln-Phe.

5

MAA synthesis may be determined by methods known in the art, such as using mass spectrometry, analytical chemistry, or spectrophotometry. For instance, an increase in UV absorbance in a reaction mix containing the D-Ala-D-Ala ligase and suitable substrates may indicate the production of an MAA.

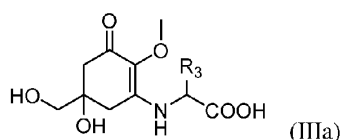
10

Disclosed herein is a method of synthesising an MAA of formula (III)



wherein R₃ is the side chain of glycine or glutamine; and R₄ is the side chain of glycine, alanine, valine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, proline, asparagine, glutamine, aspartic acid, glutamic acid, lysine, arginine, phenylalanine, tyrosine and tryptophan;

the method comprising contacting a polypeptide as defined herein with a compound of formula (IIIa) and an amino acid,



wherein the amino acid is selected to have a side chain corresponding to R₄. The amino acid is preferably an L-amino acid.

In some embodiments, the contacting is performed *in vivo* in a cell. The cell may be a host cell as defined herein that is capable of expressing the D-Ala-D-Ala ligase (e.g., from an expression vector introduced into the cell, or from a genomically-integrated nucleic acid construct).

Disclosed herein is a method for synthesising a mycosporine-like amino acid (MAA), the

method comprising culturing a host cell that expresses a polypeptide as defined herein under conditions for synthesis of the MAA by the host cell. In some embodiments, the polypeptide is an engineered polypeptide as defined herein. In some embodiments, the polypeptide comprises an amino acid sequence having at least 70% sequence identity (such as at least 5 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity) to an amino acid sequence as set forth in SEQ ID NO: 1–36.

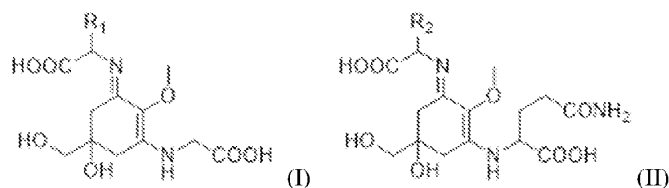
The mycosporine-like amino acids may be secreted into the medium or remain in the cells. 10 As such, some embodiments of the method comprise recovering the MAAs from the cell culture, such as from the culture medium or from the cell fraction. The MAAs may be recovered using any suitable method known in the art according to the culture method. For example, the recovery may be performed using centrifugation, filtration, ion exchange chromatography, crystallisation and HPLC, etc. The step of recovering the mycosporine-like 15 amino acid may additionally include a conventional separation process and/or purification step.

Disclosed herein is a mycosporine-like amino acid (MAA) obtained by a method as defined herein. 20

MAA compositions

This disclosure also relates to mycosporine-like amino acids (MAAs), and in particular to mycosporine-glycine-X (MGly-X) and mycosporine-glutamine-X (MGln-X) compounds 25 (where X can be any of the 20 standard amino acids) which are newly characterised and which are biosynthetically accessible using polypeptides as defined herein. MAAs herein may be useful in a range of applications, including as UV absorbents, antioxidants, anti-aging compounds, and anti-inflammatory agents.

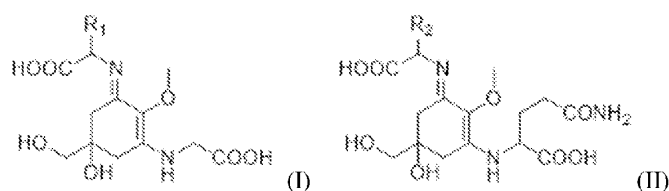
30 Accordingly, disclosed herein is a mycosporine-like amino acid (MAA), wherein the MAA is a compound of formula (I) or (II):



wherein R₁ is the side chain of leucine, isoleucine, methionine, asparagine or glutamine; and R₂ is the side chain of glycine, alanine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, glutamine, glutamic acid or phenylalanine.

5

Also disclosed herein are compositions comprising one or more MAAs of formula (I) or (II):



wherein R₁ is the side chain of leucine, isoleucine, methionine, asparagine or glutamine; and R₂ is the side chain of glycine, alanine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, glutamine, glutamic acid or phenylalanine.

10

Preferred MAAs are those which contain hydrophobic side groups, such as MGly-Leu, MGly-Ile, MGly-Met, MGln-Leu, MGln-Ile, MGln-Met and MGln-Phe. These MAAs are potentially advantageous in topical or cosmetic compositions where water resistance is beneficial.

15

Hydrophobic substances may be identified by their partition coefficient. The partition coefficient (P) is the ratio of concentrations of a substance in the two phases of a mixture of two immiscible liquids at equilibrium. Typically, one of the solvents chosen is water while the second is hydrophobic such as octanol. The logarithm of the ratio of the concentrations of the substance in the solvents (logP) may be used as a measure of hydrophobicity. LogP values < 0 are characteristic of hydrophilic compounds, while higher logP values are indicative of more hydrophobic compounds. In some embodiments, compositions herein comprise MAAs with positive logP values or logP values greater than -1.00.

20

Compositions herein may be UV-absorbing compositions. Such compositions may be formulated for application to a surface (such as a coating on a surface), or for topical

25

application to the skin. For example, the MAAs of this disclosure may be formulated into carriers such as water-based liquids, lotions, oils, oil-based solutions, powders, gels, emulsions, dispersions, or mixtures thereof for application to a surface or to the skin.

- 5 In one embodiment, the composition is a topical composition. The term “topical” as used herein includes any route of administration that enables the compounds to line the skin or mucosal tissues. Topical compositions may contain adjuvants and additives such as preservatives, organic solvents, browning agents, antioxidants, stabilisers, emollients, silicones, alpha-hydroxy acids, demulcents, anti-foaming agents, moisturising agents,
- 10 vitamins, fragrances, ionic or non-ionic thickeners, surfactants, fillers, sequestrants, polymers, propellants, alkalinising or acidifying agents, opacifiers, fatty compounds (e.g., oil, wax, alcohols, esters, fatty acids), colorants, or mixtures thereof or any other ingredient that may be used for the production of topical compositions.
- 15 Suitable carrier materials for topical administration include any carrier or vehicle commonly used as a base for creams, ointments, gels, emulsions, lotions, pastes, jellies, sprays, foams, powders, or paints for topical administration, including but not limited to emulsifying agents, inert carriers including hydrocarbon bases, emulsifying bases, water-soluble bases, or combinations thereof. In particular, suitable solvents, emollients and emulsifiers for
- 20 hydrophobic topical formulations include lanolin, paraffin, beeswax, emulsifying waxes, dimethicones, mineral oils, silicone oils, vegetable oils, fatty acids and alkyl esters of fatty acids or dicarboxylic acids, triglyceride esters, fatty alcohols and fatty alcohol ethers and sterols. Other suitable solvents, emollients and emulsifiers include polyhydric alcohols and polyether derivatives such as glycerol, sorbitol, polyethylene glycols, polypropylene glycols,
- 25 ionic and zwitterionic surfactants, amphoteric surfactants and non-ionic surfactants.

The compositions of the present disclosure may be in the form of, for example, an aqueous solution, an emulsion (oil-in-water or water-in-oil), a solid (e.g., a roll-on stick), an ointment, a lotion, a powder, a gel, a cream, or an aerosol (foams, sprays, or the like).

30

Ointments and creams may, for example, be formulated with an aqueous or oily base with the addition of suitable thickening and/or gelling agents. Lotions may be formulated with an aqueous or oily base and will in general also contain one or more emulsifying agents, stabilising agents, dispersing agents, or thickening agents. Liquid sprays may be delivered

from pressurised packs, for example, via a specially shaped closure.

In some embodiments, the composition is a cosmetic composition. The MAA compounds may be incorporated into cosmetic and/or personal care product compositions. For example,
5 the compounds may be included into formulations used in the preparation of cosmetic products such as make-up, for example in cream make-up, eye-care preparations, eye shadow preparations, mascara, eyeliner, eye creams or eye-fix creams; lip-care preparations, e.g., lipstick, lip gloss, lip contour pencils; nail-care preparations, such as nail varnish, nail varnish removers, nail hardeners or cuticle removers. These products are formulated
10 according to known methods in the art. The compounds may also be incorporated into formulations that may be used to protect hair against photochemical damage (e.g., discoloration or damage of a mechanical nature).

In some embodiments, the composition is a sunscreen composition. The specific amount of
15 MAAs needed to obtain a desired sun protection factor (SPF) for a sunscreen can be determined by techniques well known in the art. Sunscreen should provide a minimum protection against UV-A and/or UV-B rays. An increased sun protection factor (i.e., mainly UV-B protection) should include an increase in the UV-A protection as well. The composition may also include other UV-absorbing agents known in the art, such as,
20 hydrophilic or lipophilic UV-A and/or UV-B sunscreen agents.

The compounds of the present invention may also be included into pharmaceutical formulations and/or compositions along with a suitable pharmaceutical carrier. These formulations and/or compositions may be prepared according to known methods in the art.
25

Representative pharmaceutically acceptable carriers include any and all solvents, dispersion media, coatings, surfactants, antioxidants, preservatives (e.g., antibacterial agents, antifungal agents), isotonic agents, absorption delaying agents, salts, preservatives, drugs, drug stabilizers, gels, binders, excipients, disintegration agents, lubricants, sweetening
30 agents, flavoring agents, dyes, such like materials and combinations thereof, as would be known to one of ordinary skill in the art (see, for example, *Remington's Pharmaceutical Sciences*, 18th Ed. Mack Printing Company, 1990, pp. 1289-1329, incorporated herein by reference). Except insofar as any conventional carrier is incompatible with the active ingredient(s), its use in the pharmaceutical compositions is contemplated.

Medical uses

5 Mycosporine-like amino acids (MAAs) herein may be used as medicaments to treat a variety of diseases and conditions associated with inflammation, ageing and UV- or oxidation-induced cellular or tissue damage. MAAs have also been shown to have anti-cancer, anti-adipogenic and wound healing effects.

10 Disclosed herein is a method of treating a skin disease or condition in a subject in need thereof, the method comprising administering an MAA or a pharmaceutical composition as defined herein to the subject. Also disclosed herein is an MAA or a pharmaceutical composition as defined herein, for use in the treatment of a skin disease or condition in a subject in need thereof. Disclosed herein is the use of an MAA or a pharmaceutical composition as defined herein in the manufacture of a medicament for the treatment of a
15 skin disease or condition in a subject in need thereof.

The skin disease or condition may affect one or more layers of the skin, including but not limited to the epidermis (including sublayers such as the stratum corneum, stratum granulosum, stratum spinosum and stratum basale), dermis and subcutaneous (or
20 hypodermis) layer. The skin disease or condition may affect one or more cell or tissue types within the skin, including but not limited to fibroblasts, keratinocytes, melanocytes, nervous tissue, adipose tissue, vascular tissue, glandular tissue and follicular tissue.

In some embodiments, the skin disease or condition is a wound. As used herein, the term
25 “wound” refers to an injury to a tissue. Wounds include both open wounds (in which the underlying tissue is exposed to the outside environment, such as, for example, a laceration, puncture, burn or surgical incision) and closed wounds (in which the underlying tissue is not exposed to the outside environment, such as, for example, pressure sores, wounds induced by blunt trauma, and wounds caused by surgical implants). The wound may be an acute or a
30 chronic wound.

In some embodiments, the skin disease or condition is an inflammatory or proliferative skin disease or condition. Proliferative skin diseases and conditions include cancerous, pre-cancerous and non-cancerous growths, non-limiting examples of which include keratoses

(such as actinic or seborrheic keratosis), keratoacanthoma, basal cell carcinoma, squamous cell carcinoma and melanoma. Inflammatory skin diseases or conditions include, for example, psoriasis, vitiligo, dermatitis (e.g., contact dermatitis, seborrheic dermatitis, atopic dermatitis), xerosis, epidermolytic hyperkeratosis, ichthyosis, acne, folliculitis, pruritis, dermatoses and urticaria.

In some embodiments, the skin disease or condition is associated with extended or excessive UV exposure (such as extended or excessive solar UV exposure), non-limiting examples of which include sunburn, actinic keratosis, polymorphous light eruption, photoaging, solar lentigines, hyperpigmentation, and skin cancer. The skin cancer may be, for example, melanoma, squamous cell carcinoma or basal cell carcinoma.

Table 1. Amino acid sequences of exemplary wild-type and engineered D-Ala-D-Ala ligases

D-Ala-D-Ala ligase	Amino acid sequence
Wild-type SpDdl1	MKETVKILHVLGSPSDEFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGWSWFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 1)
Wild-type SpDdl2	MKSTMKILHVLGSPSDEYSFRLGILYGGSFPEDMQSKQYDYIFS ALVRPTGTWSFVKKLSDIVDLEKLEERENNVKKMDITTALQYIM DVLKPDALVHSHCLKGMTSHRALFEALNIPVVGSAESRYLTM DKLSTRGVLLSSGNVACADGLIYSKGDPLPSTEQVKYPCVVKASK GEDTKAVQLVKDPKFLNAAIEHALSFSHDVIERFIEGREIRCAVV ESVATGELKALRCIEYKVRENDIRTTEDKLLCDKKGLPVCKAPET KTWFLDPSKEEKLHRIQQQSCSAFRELGLQDFGLDFRVDLQGN PFFLECNLFCSFGPQSILNIIANSSGITGERLFDMMVKNALLRKNVI Q (SEQ ID NO: 2)

SpDdl1 Y253A	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGWSWFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKAICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 3)
SpDdl1 Y253C	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGWSWFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKCICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 4)
SpDdl1 Y253D	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGWSWFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKDICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 5)
SpDdl1 Y253E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGWSWFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKEICNEKGLPVSKPPEAKT

	WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 6)
SpDdl1 Y253F	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKFICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 7)
SpDdl1 Y253G	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKGICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 8)
SpDdl1 Y253H	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKHICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 9)
SpDdl1 Y253I	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM

	KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKIIICNEKGLPVSKPPEAKTW FLDPTKEEKLHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL ECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQEN NNVNL (SEQ ID NO: 10)
SpDdl1 Y253K	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKKICNEKGLPVSKPPEAKT WFLDPTKEEKLHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 11)
SpDdl1 Y253L	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKLICNEKGLPVSKPPEAKT WFLDPTKEEKLHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 12)
SpDdl1 Y253N	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKNICNEKGLPVSKPPEAKT WFLDPTKEEKLHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 13)

SpDdl1 Y253P	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDK P ICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 14)
SpDdl1 Y253Q	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDK Q ICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 15)
SpDdl1 Y253R	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDK R ICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 16)
SpDdl1 Y253S	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDK S ICNEKGLPVSKPPEAKT

	WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 17)
SpDdl1 Y253T	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKTICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 18)
SpDdl1 Y253V	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKVICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 19)
SpDdl1 Y253W	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKWICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 20)
SpDdl2 L254Y	MKSTMKILHVLGSPSDEYSFRLGILYGGSFPEDMQSKQYDYIFS ALVRPTGTWSFVKLSDIVDLEKLEERENNVKKMDITTALQYIM DVLKPDALVHSHCLKGMTSHRALFEALNIPVVGSAESRYLTM

	DKLSTRGVLLSSGNVACADGLIYSKGDPLPSTEQVKYPCVVKASK GEDTKAVQLVKDPKFLNAAIEHALSFSDHVIIEFIEGREIRCAVV ESVATGELKALRCIEYKVRENDIRTTEDKYLCDKKGLPVCKAPET KTWFLDPSKEEKLIHRIQQKSCSAFRELGLQDFGLDFRVDLQGN PFFLECNLFCSFGPQSILNIIANSSGITGERLFDMMVKNALLRKNVI Q (SEQ ID NO: 21)
SpDdl1 G28A, D189G	MKETVKILHVLGSPSDEFSFKINMLYGASFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKGSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 22)
SpDdl1 D65G, M79L, T227A, E230G, T270R	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGLKKLEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVES TATGGLKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKR WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 23)
SpDdl1 L66M, I137T	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDMKKLEETENSVKKMDITAALQHIK HVIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM DKLTTRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASK GEDTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVV ESTTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAK TWFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPF LLECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRKRKQ ENNVNL (SEQ ID NO: 24)

SpDdl1 L69P, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKPEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVE TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRERKQE NNNVNL (SEQ ID NO: 25)
SpDdl1 I137M	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLMTRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKG EDTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVE STTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 26)
SpDdl1 D65G, L66M, L69P, M79L, I137M, Y253I	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLMTRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKG EDTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVE STTTGELKALSCLEYKVRENDIRRTEDKIICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 27)
SpDdl1 D65G, L66M, L69P, M79L, Y253I, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSPPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVE TTTGELKALSCLEYKVRENDIRRTEDKIICNEKGLPVSKPPEAKTW

	FLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL ECNLFCFSGPQSVVNIIAKESGFTDESLFDMVENTLLRERKQEN NNVNL (SEQ ID NO: 28)
SpDdl1 D65G, L66M, L69P, M79L, Y253G, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKGICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMVENTLLRERKQE NNNVNL (SEQ ID NO: 29)
SpDdl1 D65G, L66M, L69P, M79L, Y253C, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKCICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMVENTLLRERKQE NNNVNL (SEQ ID NO: 30)
SpDdl1 D65G, L66M, L69P, M79L, Y253L, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKLICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCFSGPQSVVNIIAKESGFTDESLFDMVENTLLRERKQE NNNVNL (SEQ ID NO: 31)
SpDdl1 I137T, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVDLKKLEETENSVKKMDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM

	KLTTTRSVLVSYGNVSCPDPGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMMVENTLLRERKQE NNNVNL (SEQ ID NO: 32)
SpDdl1 D65G, L66M, L69P, M79L	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDPGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 33)
SpDdl1 D65G, L66M, L69P, M79L, I137T	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLTTTRSVLVSYGNVSCPDPGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMMVENTLLRKRKQE NNNVNL (SEQ ID NO: 34)
SpDdl1 D65G, L66M, L69P, M79L, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLITRSVLVSYGNVSCPDPGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMMVENTLLRERKQE NNNVNL (SEQ ID NO: 35)

SpDdl1 D65G, L66M, L69P, M79L, I137T, K357E	MKETVKILHVLGSPSDEFSFKINMLYGGSFPEEMQSPKHDYCFGY ALVRPTGSWSFAERLSDIVGMKKPEETENSVKKLDITAALQHIKH VIQPDMALVHFVCLKGITTYRALFDALNIPLMAGSVESRYLAM KLTTTRSVLVSYGNVSCPDGFIYNKGDILQSEQVKYPCVVKASKGE DTKAVRLVKDSEFLNAAIEHALSYSDHVIERYIEGREIRCAVVES TTTGELKALSCLEYKVRENDIRRTEDKYICNEKGLPVSKPPEAKT WFLDPTKEEKLIHRIQQQSCRAFRELGLQDFGLFDFRVDLEGNPFL LECNLFCSFGPQSVVNIIAKESGFTDESLFDMMVENTLLRERKQE NNNVNL (SEQ ID NO: 36)
--	--

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications, which fall within the spirit and scope. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

Unless otherwise defined, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

Certain embodiments of the invention will now be described with reference to the following examples which are intended for the purpose of illustration only and are not intended to limit the scope of the generality hereinbefore described.

EXAMPLES

Example 1: Identification of MAA-producing D-Ala-D-Ala ligases (Ddl) from *S. pistillata*

- The biosynthetic genes responsible for MAA production in the cyanobacterium *Nostoc punctiforme* have previously been characterised. Shinorine, the MAA studied, is synthesised from S7P over four enzymatic steps: S7P is firstly converted to 4-deoxygadusol (4-DG) by 2-demethyl 4-deoxygadusol synthase (DDGS) and O-methyltransferase (OMT), followed by the conjugation of glycine to 4-DG via an ATP-grasp ligase to form mycosporine-glycine (MG) (FIG. 1, top). A serine molecule is then conjugated to MG, either via a non-ribosomal peptide synthase (NRPS)-like enzyme (such as in *Anabaena variabilis*), or by a D-Ala-D-Ala ligase (Ddl) (as in *N. punctiforme*), to produce shinorine (FIG. 1, top). The specificity of the final two enzymes in the pathway for the amino acid to be incorporated into the cyclohexenone core determines the final identity and properties of the MAA product.
- A *S. cerevisiae* platform for the detection of MAA production was set up, through the integration of the first two conserved genes of the MAA biosynthetic pathway, DDGS and OMT, into the genome of a haploid W303 strain (FIG. 1, bottom). This strain is referred to as the “MAA base strain”. The open reading frames of the putative ATP-grasp ligases (ATPGLs) and Ddl homologs of an identified MAA biosynthetic cluster from *Nostoc punctiforme* were cloned into separate vectors, which are then co-transformed into the MAA base strain. The Ddl from the organism *A. millepora* resulted in mycosporine-2-glycine production, while those from *G. chorda* and *P. umbilicalis* resulted in porphyra-334 production. Interestingly, the first Ddl of *Stylophora pistillata* (SpDdl1, NCBI Reference Sequence: XP_022803927.1) attached a glycine residue to form mycosporine-2-glycine, while the second Ddl from the same coral (SpDdl2, NCBI Reference Sequence: XP_022806117.1) attached a threonine residue to form porphyra-334.

Example 2: Engineered SpDdl ligases expand classes of enzymatically accessible MAAs

- A single-residue substitution in SpDdl1 results in increased promiscuity towards the second amino acid of the MAA**
- “Swap mutants” SpDdl12 and SpDdl21 were made in an attempt to swap the amino acid substrate specificities of the two SpDdls, with the seven swapped residues made on the basis of their conservation among Ddl sequences that have the same amino acid specificity, and/or

their importance in amino acid substrate recognition in the Alphafold-predicted structure of SpDdl1. The swap mutants were successful in producing the other SpDdl's primary product, that is, SpDdl12 produced porphyrin-334 instead of mycosporine-2-glycine, and SpDdl21 produced mycosporine-2-glycine instead of porphyrin-334.

5

Through reversion mutations of the swap mutant SpDdl12, where each of the seven swap mutations are reverted back to the original residue, it was found that the single reversion of Y253L in SpDdl12 results in the Ddl producing almost solely mycosporine-2-glycine again (FIG. 2). A single Y253L mutation in SpDdl1 was also able to result in a product switch
10 from mycosporine-2-glycine to porphyrin-334. Subsequently, to see if other residues at the same position would exhibit the same amino acid specificity, Y253X mutations were made in SpDdl1, where X is all the other 19 amino acids available. These mutations resulted in the production of a wide variety of mycosporine-glycine-X compounds (FIG. 3), including several which have not been previously characterised, putatively identified based on
15 theoretical masses as: mycosporine-glycine-leucine/isoleucine, mycosporine-glycine-methionine, mycosporine-glycine-glutamine, and mycosporine-glycine-asparagine. A summary of the Y253X mutations and their MAA products are shown in Table 2.

The proposed structures of these new mycosporine-glycine-X compounds are provided in
20 FIG. 4, and the MS/MS profiles of mycosporine-2-glycine, mycosporine-glycine-alanine, mycosporine-glycine-threonine (porphyrin-334), mycosporine-glycine-serine (shinorine) and mycosporine-glycine-glutamine are shown in FIG. 10(a)-(e).

To further substantiate the ability of the SpDdl1 mutants to produce novel mycosporine-glycine-X variants, *in vitro* studies involving the purified enzymes were carried out. These
25 allow for the activities of the enzymes to be characterised in a defined environment. Defined amounts of purified SpDdl1 variants were incubated with ATP, mycosporine-glycine and varying concentrations of an amino acid substrate, and enzymatic activity was detected through monitoring the absorbance at 334 nm of the reaction mixture. An increase in the
30 absorbance at 334 nm indicates that the enzyme is active with the amino acid substrate X added into the mixture, and is able to produce a mycosporine-glycine-X compound.

Of the variants purified, SpDdl1, SpDdl2, SpDdl1 Y253L, SpDdl1 Y253I and SpDdl1 Y253A were found to be active *in vitro*. Their catalytic efficiencies with respect to the

various amino acid substrates tested are depicted in FIG. 5. Purified SpDdl1 Y253I was catalytically active with asparagine, supporting the formation of mycosporine-glycine-asparagine, while purified SpDdl1 Y253A was catalytically active with methionine, supporting the formation of mycosporine-glycine-methionine. To eliminate the possibility of non-specific activity being detected, a negative control in the form of the ATP-binding mutant SpDdl1 D305A was also subjected to the assay, which was found to be inactive at all concentrations of glycine tested (FIG. 6).

There was no clear trend in the type of MAAs produced with the substitution of Y253 with another amino acid, suggesting that this mutation results in an increase in amino acid promiscuity of SpDdl1. This could be attributed to steric interactions within the amino acid binding pocket of SpDdl1. In the AlphaFold model of the wild-type SpDdl1, the presence of glycine is relatively favourable, compared to the docking of threonine, which is sterically hindered. This could explain the promiscuity of some of the Y253X mutants, if the second amino acid introduced to the mycosporine molecule is limited primarily by its size.

Random mutagenesis of SpDdl1 allows for the synthesis of new mycosporine-glutamine-X class of MAAs

The surprising result of the importance of the residue Y253 in SpDdl1 gave rise to the question of whether the enzyme's specificity towards mycosporine-glycine can also be altered. Random mutagenesis of SpDdl1 was carried out with a GeneMorph random mutagenesis kit (Agilent), and 1440 of these random mutants were screened using a UV spectrometry-based method. Five mutants that are able to recognise mycosporine-glutamine as a substrate were identified in this screen (labelled OM1-5), which were able to produce a total of nine novel mycosporine-glutamine-X compounds, in varying proportions. These nine novel MAA compounds are: mycosporine-glutamine-glycine, mycosporine-glutamine-glutamine, mycosporine-glutamine-methionine, mycosporine-glutamine-alanine, mycosporine-glutamine-leucine/isoleucine, mycosporine-glutamine-serine, mycosporine-glutamine-phenylalanine, mycosporine-glutamine-valine and mycosporine-glutamine-threonine. The mutations present in the SpDdl1 OM mutants are described in Table 3, while the relative peak heights of the mycosporine-glutamine-X products are represented in FIG. 7.

Following these findings, several of the favourable mutations allowing SpDdl1 to recognise mycosporine-glutamine as a substrate were combined with selected Y253X mutations. The mutation combinations are described in Table 4. Mass spectrometry analysis of these new mutants revealed the production of additional new mycosporine-glutamine-X species such as mycosporine-glutamine-cysteine, mycosporine-glutamine-glutamate, and mycosporine-glutamine-histidine (FIG. 8).

The proposed structures of all new mycosporine-glutamine-X compounds are provided in FIG. 9, and the MS/MS profiles of mycosporine-glutamine-phenylalanine, mycosporine-glutamine-serine, mycosporine-glutamine-histidine and mycosporine-glutamine-glutamine are shown in FIG. 10(e)-(i).

Mycosporine-glutamine-X compounds have not been previously described. The ability of the SpDdl1 mutants to incorporate more hydrophobic amino acids (e.g., leucine and phenylalanine) onto the cyclohexenone ring of mycosporine-glutamine allows for the production of new MAAs which have additional hydrophobicity. This property of increased hydrophobicity could improve the water resistance of MAA-based sunscreens.

There does not seem to be an apparent rationale behind the residues identified as enabling the enzyme to recognise mycosporine-glutamine as a substrate. The mutations allowing for the recognition of mycosporine-glutamine as a substrate are primarily located near the surface of SpDdl1, suggesting the involvement of allosteric effects.

Discussion

The replacement of the single residue Y253 in the MAA-producing enzyme SpDdl1 to other naturally-occurring amino acid residues is able to result in the production of multiple new-to-nature mycosporine-like amino acids, specifically mycosporine-glycine-X where X is a second amino acid.

Random mutagenesis of SpDdl1 resulted in five mutants which are able to recognize mycosporine-glutamine instead of mycosporine-glycine as a substrate, allowing for the production of a new class of mycosporine-glutamine-X MAAs, where X is a second amino acid attached to the cyclohexenone ring.

Some of the new MAAs produced using both enzyme engineering approaches contain more hydrophobic groups (e.g. phenylalanine or leucine as the second amino acid attached to the central cyclohexenone structure). This makes the new MAA likely to be more hydrophobic and therefore more water-resistant than naturally-occurring MAAs, thus rendering it more
5 suitable for use as a natural sunscreen molecule.

The intricacies surrounding the factors determining an enzyme's substrate specificity means that often, the simultaneous mutation of multiple residues within the active site are required for a change in the enzyme's specificity. The fact that the mutation of a single residue Y253
10 in SpDdl1 was able to give rise to not only a change in substrate specificity, but also increase its promiscuity such that it is now able to recognise amino acids which are not typically incorporated into mycosporine-glycine, was highly unexpected. Previous studies have attempted to swap an entire omega-loop region spanning 43 amino acids in a Ddl to swap the substrate specificity, but this approach may not be generally applicable to other Ddls due
15 to greater structural disruptions associated with more changes in the enzyme. A technical advantage to the current approach lies in the relative ease of altering a single residue in the enzyme over mutating multiple residues, both in terms of molecular cloning work and in minimising the structural perturbations to the enzyme.

20 The random mutagenesis of SpDdl1 to alter its specificity towards mycosporine-glycine was also unexpected, considering that the wide range of untargeted mutations introduced would necessitate the screening of a large number of mutants for the identification of a desirable mutant. The serendipitous discovery of not just one, but five beneficial mutants that gained the ability to recognize mycosporine-glutamine as a substrate was also highly unexpected.
25 All the mutations identified in the beneficial mutants were also unique, and none were located near the predicted binding pocket of the mycosporine-glycine substrate; this highlights the advantage of taking an unbiased/random approach towards mutating the enzyme, since these residues would not have been predicted to be important based on a rational method. The technical advantage of this random mutagenesis method lies in the
30 relative ease of screening for beneficial SpDdl1 mutants that are able to recognise mycosporine-glutamine as a substrate, on the basis of the changes in the UV absorption spectra of the mutants. An "unsuccessful" mutant would have a clear peak at the 310 nm region (suggesting mycosporine-glutamine has not been converted into an MAA product),

while a “successful” mutant would have a slight peak at the 330 nm region, characteristic of doubly-substituted MAAs.

- 5 The use of a host cell as a platform for the identification of novel MAA-producing enzymes could allow for the screening of a wide variety of putative MAA genes from a broad range of organisms. This opens up the possibility of discovering currently unidentified MAAs, broadening the range of MAAs available for testing for use in sunscreen formulations.

10 **Table 2.** MAAs produced by wild-type and mutant SpDdls with a substitution at position 253 (or a corresponding position). Novel MAAs are in bold.

Sample	Most abundant MGly-X products according to relative peak height			MGly conversion	Notes
SpDdl1	MG-Gly	MG-Ala		0.594	Highest producer of MG-Gly
SpDdl2	MG-Thr	MG-Ser		0.510	
SpDdl2 L254Y	MG-Gly	MG-Ala	MG-Thr	0.630	Also produces MG-Gln Highest producer of MG-Val , MG-Ala
SpDdl1 Y253F	MG-Thr	MG-Ser	MG-Gly	0.658	Also produces MG-Ala
Y253S	MG-Gly	MG-Gln	MG-Ser	0.574	Also produces MG-Thr, MG-Met
Y253C	MG-Ser	MG-Gly	MG-Gln	0.698	Also produces MG-Met , MG-Thr, MG-Asn Highest producer of MG-Leu
Y253W	MG-Thr	MG-Ser	MG-Ala	0.655	Also produces MG-Gly, MG-Gln
Y253P	MG-Gly	MG-Gln	MG-Thr	0.312	Also produces MG-Ser

					Likely to be least active due to unreactive nature of proline
Y253H	MG-Thr	MG-Gly	MG-Ala	0.710	Also produces MG-Ser
Y253Q	MG-Thr	MG-Ser	MG-Gln	0.710	Also produces MG-Gly
Y253R	MG-Thr	MG-Gly	MG-Gln	0.428	Also produces MG-Ser
Y253I	MG-Thr	MG-Ser	MG-Asn	0.761	Also produces MG-Gly, MG-Gln Highest producer of MG-Glu, MG-Thr, MG-Ser, MG-Asn
Y253A	MG-Gly	MG-Thr	MG-Ser	0.707	Also produces MG-Gln, MG-Met Highest producer of MG-Met
Y253T	MG-Thr	MG-Ser	MG-Gln	0.730	Also produces MG-Gly, MG-Ala
Y253N	MG-Gly	MG-Ser	MG-Thr	0.509	Also produces MG-Gln
Y253K	MG-Thr	MG-Gln	MG-Ser	0.431	Also produces MG-Gly Likely to be less active due to positively-charged nature of lysine
Y253V	MG-Ser	MG-Gly	MG-Thr	0.657	Also produces MG-Asn, MG-Met
Y253D	MG-Thr	MG-Gln	MG-Ser	0.587	
Y253E	MG-Thr	MG-Gly	MG-Gln	0.618	Also produces MG-Ser
Y253G	MG-Gln	MG-Thr	MG-Ser	0.636	Highest producer of MG-Ile, MG-Gln

Table 3. SpDdl1 original mutants (OMs) that recognise MGln as substrate, and the mutations that are present.

MGln-recognising mutant	Mutations present
-------------------------	-------------------

OM1	G28A, D189G
OM2	D65G, M79L, T227A, E230G, T270R
OM3	L66M, I137T
OM4	L69P, K357E
OM5	I137M

Table 4. SpDdl1 mutants that recognise MGln as substrate, and the mutations that are present.

SpDdl1 species	Mutations present
MGlnY 1	D65G, L66M, L69P, M79L, I137M, Y253I
MGlnY 5	D65G, L66M, L69P, M79L, Y253I, K357E
MGlnY 6	D65G, L66M, L69P, M79L, Y253G, K357E
MGlnY 7	D65G, L66M, L69P, M79L, Y253C, K357E
MGlnY 8	D65G, L66M, L69P, M79L, Y253L, K357E
MGlnY 9	I137T, K357E
MGlnY 10	D65G, L66M, L69P, M79L
MGlnY 11	D65G, L66M, L69P, M79L, I137T
MGlnY 12	D65G, L66M, L69P, M79L, K357E
MGlnY 13	D65G, L66M, L69P, M79L, I137T, K357E

5 **Example 3: Metabolic engineering increases yield of MAAs in yeast**

In this work, *Saccharomyces cerevisiae* (more commonly known as baker’s yeast) is used as the heterologous host to express MAA-producing genes. Alternative organisms such as *Escherichia coli* and *Corynebacterium glutamicum* have previously been used for MAA
10 production, but *S. cerevisiae* has resulted in the highest known yields, and *S. cerevisiae* is the only host to have demonstrated gram/L scale of production in a bioreactor.

Certain genetic modifications made to the yeast genome were found to increase the yield of MAAs in yeast. These include:

- i) deletion of the transaldolase gene *TALI*, which leads to an accumulation of the MAA precursor sedoheptulose 7-phosphate (S7P);
- 5 ii) deletion of the phosphofructokinase gene *PFK2*, which also leads to S7P accumulation;
- iii) deletion of the hexokinase gene *HXK2*, which is involved in glycolysis and the regulation of gene transcription; and
- 10 iv) introduction of xylose-utilisation genes into the yeast genome, such that the yeast is able to metabolise xylose as an alternative carbon source.

Yeast cells are cultured in yeast extract peptone (YEP), synthetic complete (SC) or dropout media (6.7 g/L yeast nitrogen base without amino acids supplemented with the respective amino acid drop-out medium supplements). For small-scale fermentation, yeast cells were

15 grown in 10 ml of medium with either glucose only (2% w/v) or glucose (1% w/v) and xylose (1% w/v), and cultured in 50 ml falcon tubes or conical flasks incubated at 30°C with shaking (250 rpm). Xylose is a significant component of lignocellulosic waste, which can be used as a cheaper feedstock (compared to traditional carbon sources such as glucose) for cell culture to produce MAAs.

20

Table 5. MAA production in non-xylose-utilising strains

Strains	Yield (mg/L), 4-day culture in YPD media			Yield (mg/L), 5-day culture in YPD media		
	Pellet	Supernatant	Total	Pellet	Supernatant	Total
Wild-type	1.50	2.16	3.66	1.61	0.64	2.25
<i>tal1Δ</i>	11.57	27.12	38.69	12.37	28.43	40.80
<i>hvk2Δ</i>	34.63	73.96	108.59	35.62	75.97	111.60
<i>tal1Δhvk2Δ</i> (7)	27.25	59.10	86.35	27.10	57.73	84.83
<i>tal1Δhvk2Δ</i> (8)	33.40	56.03	89.43	33.34	72.19	105.53

Table 6. MAA production in xylose-utilising strains

Strains	Yield (mg/L), 4-day culture in CAA media		
	Pellet	Supernatant	Total
WT	2.680589	2.236056	4.916644
WT PsXYL (7)	-1.78081	1.334492	-0.44632
tal1Δ	21.66698	7.66552	29.3325
tal1Δ PsXYL	54.06345	26.72109	80.78454
hxxk2Δ	43.29914	11.25169	54.55083
tal1Δhxxk2Δ (7)	28.37782	9.421783	37.7996
tal1Δhxxk2Δ (7) PsXYL	18.91944	7.185727	26.10517
tal1Δhxxk2Δ (8)	38.54853	12.03275	50.58127
tal1Δhxxk2Δ (8) PsXYL	82.72022	17.49345	100.2137

Example 4: Sunscreens containing MAAs

- 5 A few cosmetic products containing MAAs are currently available: Helioguard® 365 from Mibelle AG Biochemistry, containing porphyra-334 and shinorine extracted from *Porphyra umbilicalis*, Helionori® by Gelyma, Ronacare® RenouMer containing a *Polysiphonia elongata* extract, the extract Alga-Gorria® from *Gelidium corneum*, and Aethic Sôvéc which contains palythine extracted from the seaweed *Chondrus yendoii*. Of these, Helioguard®,
- 10 Helionori® and Aethic Sôvéc are marketed as sunscreens, while Ronacare® RenouMer is purported to have hydrating and skin-smoothing effects, and Alga-Gorria® is marketed as an anti-ageing extract. MGln-Gly, MGln-Phe, MGln-Ser, MGln-His, MGln-Gln, MGln-Asn, MGln-Val, MGln-Cys, MGln-Ala, MGln-Leu, MGln-Ile, MGln-Met, MGly-Gln and MGly-Met are chemically similar to porphyra-334, shinorine and palythine, and
- 15 spectrophotometric analysis showed that these compounds are capable of absorbing UV light of both UVA (315–400 nm) and UVB (280–315 nm) wavelengths (Figures 11a–g), making them useful as UV absorbents in sunscreen and anti-ageing formulations.

The hydrophobicity of sunscreen compounds is a key property, particularly in the context of sunscreens used extensively during water sports or activities involving contact with water. Some of the new MAA compounds generated by the Ddl variants possess more hydrophobic groups compared to known MAAs (e.g. mycosporine-glutamine-phenylalanine, mycosporine-glycine-isoleucine). A comparison of logP values calculated by ChemDraw for the newly-generated MAAs showed that several of the MAAs (marked with an asterisk in Table 7) are more hydrophobic than (i.e., have a less negative logP value) known mycosporine-glycine-X or mycosporine-glutamine molecules. Positive logP values denote hydrophobic compounds which partition into the octanol phase, while negative logP values denote hydrophilic compounds which partition into the water phase. The more positive the logP value, the more hydrophobic the compound is, and the more negative the logP value, the more hydrophilic. These hydrophobic MAA compounds (e.g., mycosporine-glycine-methionine, mycosporine-glycine-leucine, mycosporine-glycine-isoleucine, mycosporine-glutamine-leucine, mycosporine-glutamine-isoleucine and mycosporine-glutamine-phenylalanine) may be used as alternative waterproof UV absorbents in sunscreens. To further increase the hydrophobicity of these compounds, the MAAs can be conjugated to more hydrophobic compounds or matrices. Such 'MAA-conjugates' have previously been made with chitosan matrices.

Table 7. Calculated logP values (ClogP) for various MAAs. Known MAAs are italicised.

MAA	ClogP value
<i>Mycosporine-glycine</i>	-1.2588
<i>Mycosporine-2-glycine</i>	-0.613724
<i>Porphyra-334 (mycosporine-glycine-threonine)</i>	-1.05907
<i>Shinorine</i>	-1.36807
Mycosporine-glycine-methionine*	-0.156725
Mycosporine-glycine-leucine*	1.15227
Mycosporine-glycine-isoleucine*	1.81153
Mycosporine-glycine-glutamine	-2.11272
Mycosporine-glycine-asparagine	-1.62707
Mycosporine-glutamine	-2.7578
Mycosporine-glutamine-glycine*	-2.11272
Mycosporine-glutamine-leucine*	-0.346725

Mycosporine-glutamine-isoleucine*	0.312525
Mycosporine-glutamine-phenylalanine*	-0.385725
Mycosporine-glutamine-methionine*	-1.65572
Mycosporine-glutamine-cysteine*	-1.82367
Mycosporine-glutamine-glutamate*	-2.68172
Mycosporine-glutamine-threonine*	-2.55807
Mycosporine-glutamine-glutamine*	-2.69172
Mycosporine-glutamine-histidine*	-2.36972
Mycosporine-glutamine-alanine*	-1.80372
Mycosporine-glutamine-serine	-2.86707

Example 5: Use of MAAs in biomedical applications

MAAs have been found to have antioxidant, anti-inflammatory, anti-photoageing, wound
5 healing, anti-cancer and anti-adipogenic properties. It was also recently discovered that
MAAs can act as an anti-viral agent against the SARS-CoV-2 infection. The potential of
MAAs to act as biomedical ingredients are demonstrated in some of their properties
described below:

- 10 i) MAAs can act as antioxidants by absorbing high-energy photons and releasing the
energy as heat. This property has been demonstrated in the MAA-producing
Stylophora pistillata, where the MAA provides rapid protection in oxidative stress
upon exposure of the organism to high temperatures, even before the induction of
antioxidant enzymes.
- 15 ii) The role of MAAs as anti-inflammatory agents have been demonstrated in immortal
human keratinocytes (HaCaT cells) and macrophage models. For example, MAAs can
inhibit genes involved in inflammation pathways, such as *COX-2* and inducible NO
synthase (iNOS).
- 20 iii) The anti-ageing properties of MAAs arise from their ability to suppress the
degradation of collagen and elastin. MAAs have also been found to inhibit protein
glycation, which is implicated in the progression of ageing and ageing-related diseases.
- iv) MAAs such as shinorine and porphyra-334 have been found to promote wound repair
in HaCaT cells, through the activation of kinases involved in wound healing and cell
proliferation.

- v) MAAs can act as an anti-cancer agent by inhibiting the proliferation of neoplastic cells. This effect of MAAs has been exhibited on a murine skin melanoma cell line and in HeLa cells from cervical cancer.
 - vi) MAAs have a potential effect in controlling obesity through its anti-adipogenic effects.
- 5 Shinorine and porphyrin-334 have been found to suppress lipid droplet accumulation and adipocyte differentiation in adipocytes, and reduce the expression of adipogenesis-related genes.

It will be appreciated that many further modifications and permutations of various aspects
10 of the described embodiments are possible. Accordingly, the described aspects are intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims.

CLAIMS

1. An engineered polypeptide that is distinguished from a wild-type D-Ala-D-Ala ligase (Ddl) by at least one amino acid substitution, wherein the at least one amino acid substitution is at a position in the wild-type Ddl corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1.
2. The engineered polypeptide of claim 1, wherein the polypeptide comprises one or more amino acid substitutions selected from the following:
 - (a) a substitution to A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W or Y at a position corresponding to position 253 of SEQ ID NO: 1;
 - (b) a substitution to A, I, L or V at a position corresponding to position 28 of SEQ ID NO: 1;
 - (c) a substitution to G or P at a position corresponding to position 65 of SEQ ID NO: 1;
 - (d) a substitution to F, I, L or M at a position corresponding to position 66 of SEQ ID NO: 1;
 - (e) a substitution to I, L or V at a position corresponding to position 67 of SEQ ID NO: 1;
 - (f) a substitution to I, L or V at a position corresponding to position 68 of SEQ ID NO: 1;
 - (g) a substitution to G or P at a position corresponding to position 69 of SEQ ID NO: 1;
 - (h) a substitution to A, I, L or V at a position corresponding to position 70 of SEQ ID NO: 1;
 - (i) a substitution to A, I, L or V at a position corresponding to position 71 of SEQ ID NO: 1;
 - (j) a substitution to A, I, L or V at a position corresponding to position 72 of SEQ ID NO: 1;
 - (k) a substitution to A, I, L or V at a position corresponding to position 73 of SEQ ID NO: 1;
 - (l) a substitution to A, I, L or V at a position corresponding to position 74 of SEQ ID NO: 1;

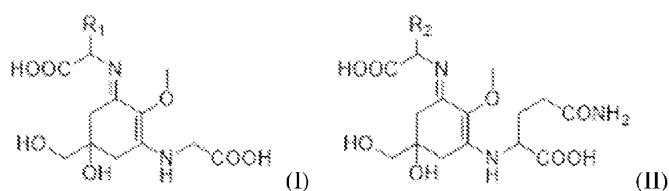
- (m) a substitution to A, I, L or V at a position corresponding to position 75 of SEQ ID NO: 1;
 - (n) a substitution to A, I, L or V at a position corresponding to position 76 of SEQ ID NO: 1;
 - (o) a substitution to A, I, L or V at a position corresponding to position 77 of SEQ ID NO: 1;
 - (p) a substitution to A, I, L or V at a position corresponding to position 78 of SEQ ID NO: 1;
 - (q) a substitution to A, F, I, L, M or V at a position corresponding to position 79 of SEQ ID NO: 1;
 - (r) a substitution to F, I, L, M, S or T at a position corresponding to position 137 of SEQ ID NO: 1;
 - (s) a substitution to G or P at a position corresponding to position 189 of SEQ ID NO: 1;
 - (t) a substitution to A, I, L or V at a position corresponding to position 227 of SEQ ID NO: 1;
 - (u) a substitution to G or P at a position corresponding to position 230 of SEQ ID NO: 1;
 - (v) a substitution to K, N, Q or R at a position corresponding to position 270 of SEQ ID NO: 1; and
 - (w) a substitution to D, E or K at a position corresponding to position 357 of SEQ ID NO: 1.
3. The engineered polypeptide of claim 2, wherein the polypeptide comprises one or more amino acid substitutions selected from the following:
- (a) a substitution to A, C, G, I, or L at a position corresponding to position 253 of SEQ ID NO: 1;
 - (b) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1;
 - (c) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1;
 - (d) a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1;
 - (e) a substitution to A at a position corresponding to position 67 of SEQ ID NO: 1;
 - (f) a substitution to A at a position corresponding to position 68 of SEQ ID NO: 1;
 - (g) a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1;
 - (h) a substitution to A at a position corresponding to position 70 of SEQ ID NO: 1;

- (i) a substitution to A at a position corresponding to position 71 of SEQ ID NO: 1;
 - (j) a substitution to A at a position corresponding to position 72 of SEQ ID NO: 1;
 - (k) a substitution to A at a position corresponding to position 73 of SEQ ID NO: 1;
 - (l) a substitution to A at a position corresponding to position 74 of SEQ ID NO: 1;
 - (m) a substitution to A at a position corresponding to position 75 of SEQ ID NO: 1;
 - (n) a substitution to A at a position corresponding to position 76 of SEQ ID NO: 1;
 - (o) a substitution to A at a position corresponding to position 77 of SEQ ID NO: 1;
 - (p) a substitution to A at a position corresponding to position 78 of SEQ ID NO: 1;
 - (q) a substitution to A or L at a position corresponding to position 79 of SEQ ID NO: 1;
 - (r) a substitution to M or T at a position corresponding to position 137 of SEQ ID NO: 1;
 - (s) a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1;
 - (t) a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1;
 - (u) a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1;
 - (v) a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1;
 - and
 - (w) a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1.
4. The engineered polypeptide of any one of claims 1 to 3, wherein the polypeptide comprises an amino acid substitution at positions corresponding to one of the following sets of positions of SEQ ID NO: 1:
- (a) a substitution at position 253;
 - (b) a substitution at position 137;
 - (c) a substitution at positions 28 and 189;
 - (d) a substitution at positions 66 and 137;
 - (e) a substitution at positions 69 and 357;
 - (f) a substitution at positions 227 and 230;
 - (g) a substitution at positions 65, 66 and 69;
 - (h) a substitution at positions 65, 66, 69 and 79;
 - (i) a substitution at each of positions 65 to 69;
 - (j) a substitution at positions 65, 79, 227, 230 and 270;
 - (k) a substitution at positions 65, 66, 69, 79 and 137;
 - (l) a substitution at positions 65, 66, 69, 79 and 357;

- (m) a substitution at positions 65, 66, 69, 79, 137 and 189;
 - (n) a substitution at positions 65, 66, 69, 79, 137 and 253;
 - (o) a substitution at positions 65, 66, 69, 79, 137 and 357;
 - (p) a substitution at positions 65, 66, 69, 79, 253 and 357;
 - (q) a substitution at positions 65, 79, 227, 230, 253 and 270;
 - (r) a substitution at positions 28, 65, 66, 69, 79, 137, 189, 227, 230, 270 and 357;
 - (s) a substitution at positions 28, 65, 66, 69, 79, 137, 189, 227, 230, 253, 270 and 357;
 - or
 - (t) a substitution at each of positions 65 to 79.
5. The engineered polypeptide of claim 4, wherein the polypeptide comprises one of the following:
- (a) a substitution to A, C, or G at a position corresponding to position 253 of SEQ ID NO: 1;
 - (b) a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1;
 - (c) a substitution to A at a position corresponding to position 28 of SEQ ID NO: 1, and a substitution to G at a position corresponding to position 189 of SEQ ID NO: 1;
 - (d) a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1 and a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1;
 - (e) a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1 and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
 - (f) a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1 and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
 - (g) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, and a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1;
 - (h) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1; and a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1;
 - (i) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a

- substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1; and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1;
- (j) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to A at a position corresponding to position 227 of SEQ ID NO: 1, a substitution to G at a position corresponding to position 230 of SEQ ID NO: 1, and a substitution to R at a position corresponding to position 270 of SEQ ID NO: 1;
- (k) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 137 of SEQ ID NO: 1, and a substitution to I at a position corresponding to position 253 of SEQ ID NO: 1;
- (l) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to T at a position corresponding to position 137 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1; or
- (m) a substitution to G at a position corresponding to position 65 of SEQ ID NO: 1, a substitution to M at a position corresponding to position 66 of SEQ ID NO: 1, a substitution to P at a position corresponding to position 69 of SEQ ID NO: 1, a substitution to L at a position corresponding to position 79 of SEQ ID NO: 1, a substitution to C, G, I or L at a position corresponding to position 253 of SEQ ID NO: 1, and a substitution to E at a position corresponding to position 357 of SEQ ID NO: 1.
6. The engineered polypeptide of any one of claims 1 to 5, wherein the wild-type D-Ala-D-Ala ligase (Ddl) comprises an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1 or 2.

7. The engineered polypeptide of claim 6, wherein the wild-type D-Ala-D-Ala ligase (Ddl) comprises or consists essentially of an amino acid sequence as set forth in SEQ ID NO: 1 or 2.
8. The engineered polypeptide of any one of claims 1 to 7, wherein the polypeptide comprises an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 3–36.
9. The engineered polypeptide of any one of claims 1 to 8, wherein the polypeptide is capable of synthesising a di-substituted mycosporine-like amino acid (MAA).
10. The engineered polypeptide of claim 9, wherein the MAA is a compound of formula (I) or (II):



wherein R_1 is the side chain of leucine, isoleucine, methionine, asparagine or glutamine;

and

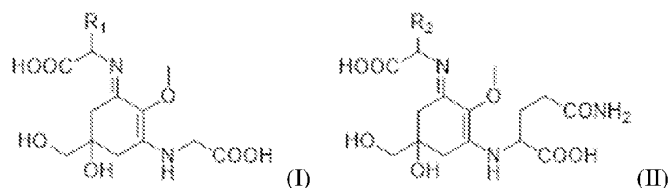
R_2 is the side chain of glycine, alanine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, glutamine, glutamic acid or phenylalanine.

11. An isolated polypeptide comprising an amino acid sequence having at least 70% sequence identity to an amino acid sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 2.
12. The isolated polypeptide of claim 11, wherein the polypeptide comprises or consists essentially of an amino acid sequence as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.
13. The isolated polypeptide of claim 11, wherein the polypeptide comprises an amino acid substitution at one or more positions corresponding to position 253, 28, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 137, 189, 227, 230, 270 or 357 of SEQ ID NO: 1.

14. Use of a polypeptide of any one of claims 1 to 13 for synthesising a mycosporine-like amino acid (MAA).
15. A polynucleotide encoding a polypeptide of any one of claims 1 to 13.
16. An expression vector comprising a polynucleotide of claim 15.
17. A host cell comprising a polynucleotide of claim 15 or an expression vector of claim 16.
18. A method for producing a polypeptide of any one of claims 1 to 13, comprising culturing a host cell of claim 17 under conditions suitable for expressing the polypeptide.
19. A method for synthesising a mycosporine-like amino acid (MAA), the method comprising culturing a host cell that expresses a polypeptide of any one of claims 1 to 13 under conditions for synthesis of the MAA by the host cell.
20. The method of claim 19, wherein the host cell is a fungal cell.
21. The method of claim 20, wherein the host cell is a yeast cell.
22. The method of claim 21, wherein the yeast is *Saccharomyces cerevisiae*.
23. The method of any one of claims 19 to 22, wherein the host cell comprises a disrupted *TAL1* (transaldolase 1), *PFK1* (phosphofructokinase 1), *PFK2* (phosphofructokinase 2), and/or *HXK2* (hexokinase 2) gene.
24. The method of any one of claims 19 to 23, wherein the host cell is a xylose-utilising strain.
25. The method of claim 24, wherein the host cell expresses one or more polypeptides selected from the group consisting of a xylose transporter, a xylose reductase, a xylitol dehydrogenase, a xylose isomerase, and a xylulokinase.

26. A mycosporine-like amino acid (MAA) obtained by a method of any one of claims 19 to 25.

27. A mycosporine-like amino acid (MAA), wherein the MAA is a compound of formula (I) or (II):



wherein R₁ is the side chain of leucine, isoleucine, methionine, asparagine or glutamine;
and

R₂ is the side chain of glycine, alanine, leucine, isoleucine, serine, threonine, cysteine, methionine, histidine, glutamine, glutamic acid or phenylalanine.

28. A UV-absorbing composition comprising one or more MAAs of claim 27.

29. The composition of claim 28, wherein the composition is a cosmetic composition.

30. The composition of claim 28 or 29, wherein the composition is a topical composition.

31. A pharmaceutical composition comprising one or more MAAs of claim 27.

32. A mycosporine-like amino acid (MAA) of claim 27, for use as a medicament.

33. A method of treating a skin disease or condition in a subject in need thereof, the method comprising administering an MAA of claim 27 or a pharmaceutical composition of claim 31 to the subject.

34. The method of claim 33, wherein the skin disease or condition is selected from the group consisting of skin cancer, actinic keratosis, photoaging, hyperpigmentation, and polymorphous light eruption.

35. The method of claim 34, wherein the skin cancer is melanoma, squamous cell carcinoma or basal cell carcinoma.
36. A mycosporine-like amino acid (MAA) of claim 27 or a pharmaceutical composition of claim 31, for use in the treatment of a skin disease or condition in a subject in need thereof.
37. Use of a mycosporine-like amino acid (MAA) of claim 27 or a pharmaceutical composition of claim 31 in the manufacture of a medicament for the treatment of a skin disease or condition in a subject in need thereof.