

Amino acid derivatives

The invention relates to amino acid derivatives for use in bioconjugation processes.

INTRODUCTION

Pyrrolysine is a natural amino acid, the only one that is authentically specified by an amber codon. It uses a 21st aminoacyl-tRNA synthetase (PylRS), naturally evolved to be orthogonal to all other amino acids and tRNAs. Blight et al., 2004 showed that PylRS and its counterpart tRNA (tRNA^{Pyl}) can incorporate pyrrolysine at amber codons in *E. coli*. They also showed that the wt PylRS is naturally promiscuous and can incorporate analogs of lysine.

Yokoyama et al. (EP1911840) demonstrated that the PylRS/tRNA^{Pyl} system is orthogonal in eukaryotic cells and showed the incorporation of several non natural amino acids (nnAAs) into a target proteins encoded by amber codons in bacterial cells. These authors also identified key amino acid residues in pylRS that form the amino acid binding pocket and function in selecting pyrrolysine over other canonical amino acids. Mutations at this site generated mutants able to recognize and aminoacylate the tRNA^{Pyl} with AzZ-lys (Yanagisawa, 2008).

This orthogonality extends to bacteria and eukaryotic cells.

PylRS is a naturally promiscuous synthetase that has naturally evolved to exclude lysine, but will incorporate lysine analogs without mutation, including azides, alkynes and alkenes (Yanagisawa et al., 2008; Neumann et al., 2008; Mukai et al., 2008; Nguyen et al., 2009). The basis of this specificity is dependent on hydrophobic interactions between amino acid residues of the pylRS binding pocket with the pyrrole ring of pyrrolysine that stabilizes and correctly positions the amino acid in the active site of the synthetase (Kavran et al., 2007). This RS/tRNA pair has been introduced via transient transfection into bacterial, yeast and mammalian cells and shown to be effective for incorporation of a number of non-natural amino acids into target proteins.

For instance, EP 1911840 demonstrates incorporation of N-ε-boc-Lysine into a target protein in *E. coli* cells.

Pyrrolysine analogs, defined as amino acid derivatives recognized by either native or genetically evolved PylRS and incorporated into proteins at amber codon sites, have been disclosed in the past few years and reviewed, for instance, by Fekner et al. (Fekner, Li & Chan, 2010) and Liu et

al.. Analogs bearing functional groups or post translational modifications have been site-specifically incorporated into proteins using pylRS-tRNA^{pyl} systems. Several studies, see e.g., Yanagisawa et al., focused on mutations within the PylRS enzyme in order to accommodate analogs in which the N6 substituent was an aromatic ring within the binding pocket pyrrolysine. Others, for instance Nguyen et al. (also in WO2010/139948), and Li et al. (also in WO2011/044255) focused on identification of pyrrolysine analogs which do not carry a bulky N6 substituent, with the result of obtaining simpler analogs which would be simple to synthesize and interact with native pylRS/tRNA^{pyl} pairs. Furthermore, Chin et al. developed two analogs with terminal alkyne and azide groups, amenable to use for protein labeling via Copper catalyzed click chemistry (CUAAC).

There remains a need to develop further pyrrolysine analogs. Whilst pyrrolysine analogs made thus far have been restricted to those evolved from a lysine backbone, the present inventors have generated pyrrolysine analogs successfully incorporated into proteins with native pylRS/tRNA^{pyl} pairs starting from a variety of amino acid structures

SUMMARY OF THE INVENTION:

According to the invention there are provided pyrrolysine analogues of formulae I to VIII as described herein.

There is also provided a mutant protein containing as non-natural amino acid one or more (e.g., one) pyrrolysine analogues of formulae I to VIII as described herein.

There is also provided a mutant protein as aforesaid which is conjugated via the one or more (e.g., one) non-natural amino acids to one or more (e.g., one) moieties selected from proteins, cytotoxic agents, drugs and polymers.

There is also provided use of a pyrrolysine analogue as aforesaid in the manufacture of a mutant protein containing one or more non-natural amino acids.

Amino acid analogs described in the present invention are new and useful and have the merit of being straightforward to prepare, in being readily incorporated into proteins (typically

without loss of bioactivity when used appropriately) and in providing useful means for bioconjugation.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1: PEGylation of azide containing monoclonal antibodies. Lane 1: Antibody with Formula IA.1 analog incorporated into heavy chain and subjected to PEGylation conditions; Lane 2: Antibody with Formula III.1 analog incorporated into heavy chain and subjected to PEGylation conditions.

BRIEF DESCRIPTION OF THE SEQUENCES OF THE SEQUENCE LISTING

SEQ ID No 1: PylRS *Methanosarcina mazei* WT nucleotide sequence

SEQ ID No 2: PylRS *Methanosarcina mazei* WT amino acid sequence

SEQ ID No 3: PylRS *Methanosarcina mazei*, Y384F mutant nucleotide sequence

SEQ ID No 4: PylRS *Methanosarcina mazei*, Y384F mutant amino acid sequence

SEQ ID No 5: tRNA^{pyl} *Methanosarcina mazei* Go1

SEQ ID No 6: U6 snRNA Promoter

SEQ ID No 7: U6-tRNA^{pyl} construct

SEQ ID No 8: GFP nucleotide sequence

SEQ ID No 9: GFP amino acid sequence

SEQ ID No 10: GFPY40 nucleotide sequence

SEQ ID No 11: GFPY40 amino acid sequence

SEQ ID No 12: anti-Her2 (4D5) gamma nucleotide sequence

SEQ ID No 13: anti-Her2 (4D5) gamma amino acid sequence

SEQ ID No 14: anti-Her2 (4D5) gamma_K274amber nucleotide sequence

SEQ ID No 15: anti-Her2 (4D5) gamma_K274amber amino acid sequence

SEQ ID No 16: anti-Her2 (4D5) Kappa nucleotide sequence

SEQ ID No 17: anti-Her2 (4D5) Kappa amino acid sequence

DETAILED DESCRIPTION OF THE INVENTION

Definitions

The term “amide” refers to a —C(=O)—NH— linkage.

The term “carbamate” refers to a —O—C(=O)—NH— linkage.

The term “ester” refers to a —C—C(=O)—O—C linkage.

The term “alkyl” refers to an aliphatic linkage or substituent, typically containing 1-6 e.g., 1-4 carbon atoms and can be straight chain or branched. Examples include methyl, ethyl, n-propyl, i-propyl, n-butyl and t-butyl.

The term “alkoxy” refers to the group —O—alkyl .

The term “alkenyl”, “alkene” or “olefin” refers to an aliphatic linkage or substituent, typically containing 2-6 e.g., 2-4 carbon atoms and can be straight chain or branched and which is unsaturated in respect of containing at least one C=C moiety. Examples include ethenyl, propen-1-yl, propen-2-yl, and 2-methyl-propen-2-yl. An alkenyl group may be optionally substituted e.g., by one or more (e.g., 1) substituents such as halogen (e.g., Cl) or an ether group (e.g., $\text{—O—C}_{1-6}\text{alkyl}$) although suitably it is not substituted.

The term “alkynyl” or “alkyne” refers to an aliphatic linkage or substituent, typically containing 2-6 e.g., 2-4 carbon atoms and can be straight chain or branched and which is unsaturated in respect of containing at least one $\text{C}\equiv\text{C}$ moiety. Examples include $\text{—C}\equiv\text{CH}$ and $\text{—C}\equiv\text{C—CH}_3$. An alkynyl group may be optionally substituted e.g., by one or more (e.g., 1) substituents such as halogen (e.g., Cl) or an ether group (e.g., $\text{—O—C}_{1-6}\text{alkyl}$) although suitably it is not substituted.

The term “cycloalkyl” refers to an alicyclic and unsaturated compound typically containing 3 to 8 cyclic carbon atoms. Cycloalkyl groups may contain branching. The total number of carbon atoms will typically be 3 to 10. Exemplary groups include cyclopropyl, cyclobutyl, cyclopentyl, 3-methyl-cyclopropyl and cyclohexyl.

The term “cycloalkenyl” refers to an alicyclic compound typically containing 5 to 8 cyclic carbon atoms and containing at least one C=C moiety. Cycloalkenyl groups may contain branching.

The total number of carbon atoms will typically be 5 to 10. Exemplary groups include cyclopentenyl, 3-methyl-cyclopropenyl and cyclohexenyl.

The term “heterocyclyl” refers to a cycloalkyl or cycloalkenyl moiety in which the ring contains one or more (e.g., one, two or three, such as one or two, especially one) heteroatoms selected from O, N and S. Examples include azetidine, pyrrolidine, piperidine, piperazine, N-methylpiperazine, morpholine and thiomorpholine.

The term “aryl” refers to an aromatic ring structure that can be part of a linkage or part of a substituent. Aryl moieties may contain one ring (e.g., phenyl) or two rings (e.g., naphthyl). Aryl groups may be substituted e.g., by one or more (e.g., one or two, such as one) substituents selected from alkyl, alkenyl, alkynyl, fluoroalkyl, halogen, alkoxy, nitro and cyano. An exemplary aryl is phenyl.

The term “heteroaryl” refers to a heteroaromatic ring structure that can be part of a linkage or part of a substituent. The heteroaromatic ring may contain 1-4 (more usually 1-3 e.g., one or two) heteroatoms selected from O, N and S. Heteroaryl moieties may contain one ring or two rings. Example groups containing one 6 membered ring include pyridine and pyrimidine.

Example groups containing one 5 membered ring include pyrrole, furan, thiophene, oxazole, thiazole, diazole, thiadiazole and tetrazole. Heteroaryl moieties that contain two rings may contain heteroatoms in one or both rings. Examples include quinoline and isoquinoline.

Heteroaryl groups may be substituted e.g., by one or more (e.g., one or two, such as one) substituents selected from alkyl, alkenyl, alkynyl, fluoroalkyl, halogen, alkoxy, nitro and cyano.

The term “aromatic halide” refers to an aromatic ring (typically phenyl) which is substituted by at least one (e.g., one) halo group such as fluorine, chloride, bromide or iodine. Said aromatic ring may contain further substituents e.g., those mentioned for aryl.

The term “azide” and “azido” refers to a $\text{N}=\text{N}(+)=\text{N}(-)$ functional group.

The term “cycloalkyne” refers to a cyclic arrangement of carbon atoms (typically 6-9 membered, especially 8-9 membered) which includes a carbon-carbon triple bond captured in the ring structure. Examples include cyclooctyne and cyclononyne. A further example is

benzyne. Cycloalkyne groups may contain branching. The total number of carbon atoms will typically be 6 to 12 e.g., 6 to 10.

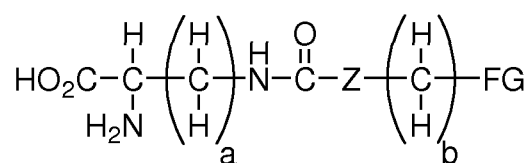
The term "ketone" refers to a C-C(=O)-C linkage.

The term "pyrrolysine analog" means an amino acid derivative recognized by either native or genetically evolved PylRS and incorporated into proteins at an amber codon site.

The expression "the side chain of one of the 20 natural amino acids" refers to the group R in the formula HOOC-CHR-NH₂ relating to the 20 natural amino acids known by their single letter codes A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W and Y. Either L or D stereochemistry (or a mixture thereof) is intended, although L stereochemistry is preferred.

The present invention discloses pyrrolysine analogs.

Some pyrrolysine analogs of the present invention have the structure of Formula I:



wherein

Z = bond, CH₂, CH-NH₂, CH-OH, NH, O, S or CH-NH₂;

a is an integer 3-7;

b is 0 or an integer 1-7; and

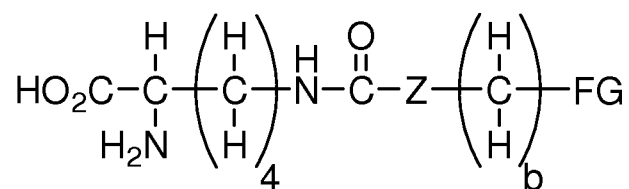
FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne.

Certain compounds of Formula (I) are known.

Thus, compounds of Formula (I) wherein a represents 4, Z represents O and -(CH₂)_b-FG represents -CH₂-C≡CH, -CH₂CH₂-N₃, -CH₂-CH=CH₂ and -CH₂-Ph are disclosed in WO2012/032181. Further compounds of Formula (I) wherein a represents 4, Z represents O and -(CH₂)_b-FG represents -CH₂CH₂CH₂-C(=O)CH₃ and -CH₂CH₂CH₂-CH=CH₂ are disclosed in WO2010/139948.

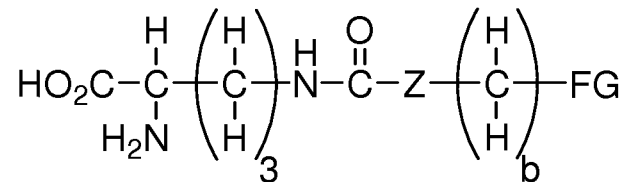
Compounds of Formula (I) wherein a represents 4, Z represents a bond, b represents 0 and $-(CH_2)_b-FG$ represents $-C(=O)Bn$ and $-C(=O)Me$ are disclosed in WO2012/032181. Further a compound of Formula (I) wherein a represents 4, Z represents a bond and $-(CH_2)_b-FG$ represents $-CH_2CH_2-C\equiv CH$ is disclosed in WO2010/139948.

In an embodiment of the present invention, a compound of Formula I has $a=4$, resulting in compounds of general Formula IA :



Formula IA

In an alternative embodiment of the present invention, $a=3$, resulting in compounds of general Formula IB:



Formula IB

In an embodiment of Formulae I, IA and IB, Z is bond, for example, Z is bond, b is 0 and FG is $C(=O)$ -aryl or $C(=O)$ -alkyl.

In an embodiment of Formulae I, IA and IB, Z is CH_2 , $CH-NH_2$, $CH-OH$, NH , O , S or $CH-NH_2$. For example, Z may represent O . Alternatively Z may represent NH . Alternatively Z may represent CH_2 , $CH-NH_2$, $CH-OH$, S or $CH-NH_2$.

In an embodiment, b is 1-4.

In an embodiment, FG is azide.

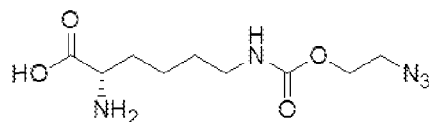
In an embodiment, FG is alkyne (e.g., ethynyl) or cycloalkyne.

In an embodiment, FG is alkene e.g., ethenyl.

In formulae I, IA and IB, when FG represents aryl, an example is aromatic halide e.g., 4-halo phenyl such as 4-iodo phenyl.

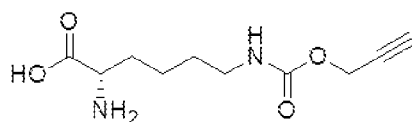
Exemplary compounds of Formula IA are the following:

(2S)-2-amino-6-[[2-azidoethoxy]carbonyl]amino}hexanoic acid



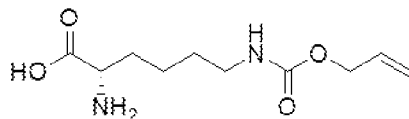
Formula IA.1;

(2S)-2-amino-6-[[prop-2-yn-1-yloxy]carbonyl]amino}hexanoic acid



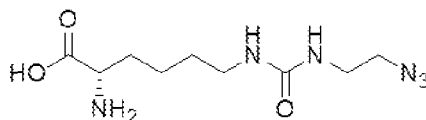
Formula IA.2;

(2S)-2-amino-6-[[prop-2-en-1-yloxy]carbonyl]amino}hexanoic acid



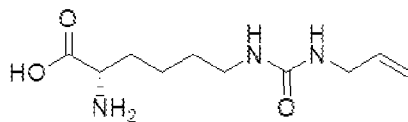
Formula IA.3;

(2S)-2-amino-6-[[2-azidoethyl]carbonyl]amino}hexanoic acid



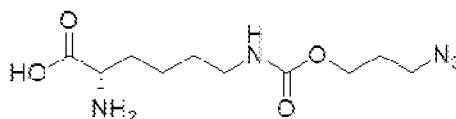
Formula IA.4;

(2S)-2-amino-6-[[prop-2-en-1-yl]carbonyl]amino}hexanoic acid



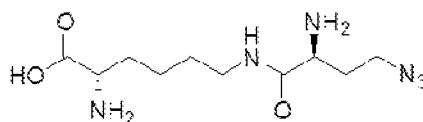
Formula IA.5;

(2S)-2-amino-6-[[[3-azidopropoxy]carbonyl]amino]hexanoic acid



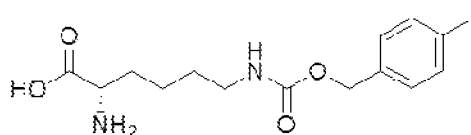
Formula IA.6;

(2S)-2-amino-6-[(2S)-2-amino-4-azidobutanamido]hexanoic acid



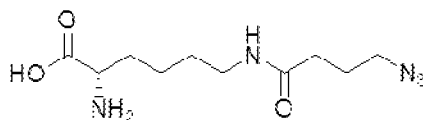
Formula IA. 7;

(2S)-2-amino-6-([(4-iodophenyl)methoxy]carbonyl)amino)hexanoic acid



Formula IA.8; and

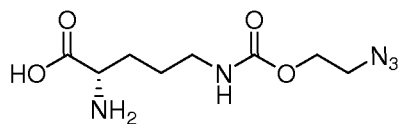
(2S)-2-amino-6-(4-azidobutanamido)hexanoic acid



Formula IA. 9;

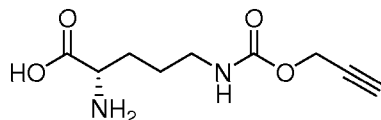
The following compounds are examples of compounds of Formula IB:

(2S)-2-amino-5-[(2-azidoethoxy)carbonyl]amino)pentanoic acid



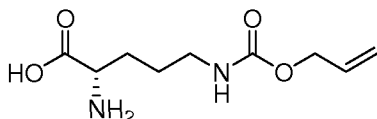
Formula IB.1;

(2S)-2-amino-5-[[[(prop-2-yn-1-yloxy)carbonyl]amino]pentanoic acid



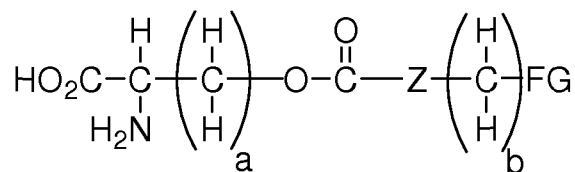
Formula IB.2; and

(2S)-2-amino-5-[[[(prop-2-en-1-yloxy)carbonyl]amino]pentanoic acid



Formula IB.3.

Alternative pyrrollysine analogs of the present invention have the structure of Formula II:



Formula II

wherein

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a = an integer 3 or 5-7; and

b = an integer 1-4.

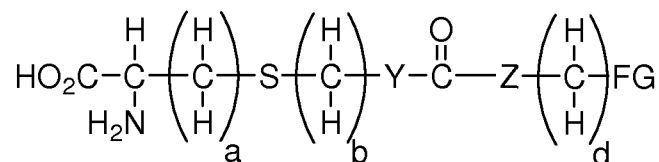
In an embodiment, FG is azide. In an embodiment, FG is alkyne e.g., ethynyl or cycloalkyne. In an embodiment, FG is alkene e.g., ethenyl. In an embodiment, Z represents NH. In an

embodiment b represents 1 or 2. In an embodiment, $Z(\text{CH}_2)_b\text{FG}$ represents $\text{NH}(\text{CH}_2)_2\text{N}_3$ or $\text{NHCH}_2\text{C}\equiv\text{CH}$.

In formula II, when FG represents aryl, an example is aromatic halide e.g., 4-halo phenyl such as 4-iodo phenyl.

Moiety $Z(\text{CH}_2)_b\text{FG}$ may, for example, represent CO-aryl e.g., CO-phenyl or -COalkyl e.g., -COMe.

Further pyrrollysine analogs of the present invention have the structure of Formula III:



Formula III

wherein

$\text{Y} = \text{CH}_2, \text{NH}, \text{O}$ or S ;

$\text{Z} = \text{CH}_2, \text{CH-NH}_2, \text{CH-OH}, \text{NH}, \text{O}$ or S ;

$\text{FG} = \text{azide, alkene, alkyne, ketone, ester, aryl}$ or cycloalkyne;

$a = \text{an integer } 1-7$;

$b = \text{an integer } 1-7$ save that when Z is NH, O or S then b is an integer $2-7$;

provided that $a+b$ is in the range $2-8$;

and

$d = \text{an integer } 1-4$.

In an embodiment, $a+b$ is in the range $3-6$ e.g., $3-4$.

In an embodiment a is 1 .

In an embodiment b is 2 . In another embodiment b is 3 .

In an embodiment, FG is azide.

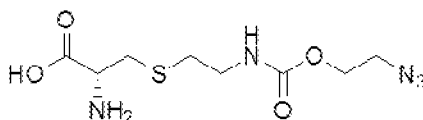
In an embodiment, FG is alkyne (e.g., ethynyl) or cycloalkyne.

In an embodiment, FG is alkene e.g., ethenyl.

In formula III, when FG represents aryl, an example is aromatic halide e.g., 4-halo phenyl such as 4-iodo phenyl.

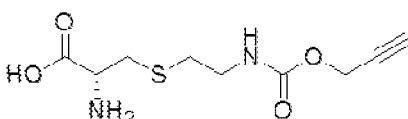
Exemplary compounds of Formula III are:

(2S)-2-amino-3-[(2-[[[(2-azidoethoxy)carbonyl]amino]ethyl)sulfanyl]propanoic acid



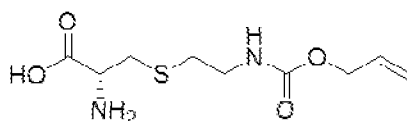
Formula III.1;

(2S)-2-amino-3-[(2-[[[(prop-2-yn-1-yloxy)carbonyl]amino]ethyl)sulfanyl]propanoic acid



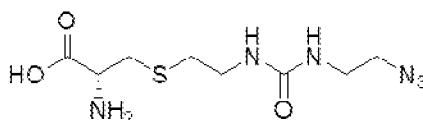
Formula III.2;

(2S)-2-amino-3-[(2-[[[(prop-2-en-1-yloxy)carbonyl]amino]ethyl)sulfanyl]propanoic acid



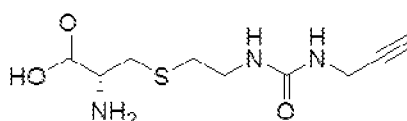
Formula III.3;

(2S)-2-amino-3-[(2-[[[(2-azidoethyl)carbamoyl]amino]ethyl)sulfanyl]propanoic acid



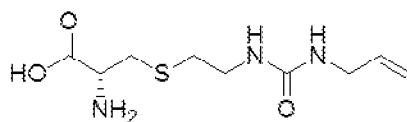
Formula III.4;

(2S)-2-amino-3-[(2-[[[(prop-2-yn-1-yl)carbamoyl]amino]ethyl)sulfanyl]propanoic acid



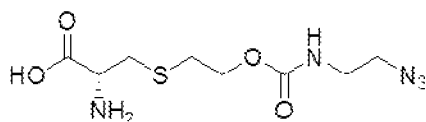
Formula III. 5;

(2S)-2-amino-3-[(2-[[[(prop-2-en-1-yl)carbamoyl]amino]ethyl)sulfanyl]propanoic acid



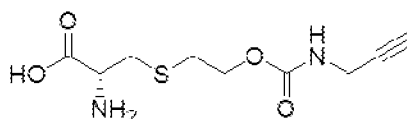
Formula III. 6;

(2S)-2-amino-3-[(2-[(2-azidoethyl)carbamoyl]oxy)ethyl]sulfanyl]propanoic acid



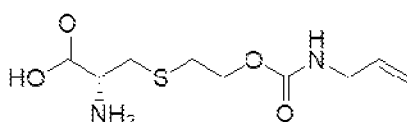
Formula III.7;

(2S)-2-amino-3-[(2-[(prop-2-yn-1-yl)carbamoyl]oxy)ethyl]sulfanyl]propanoic acid



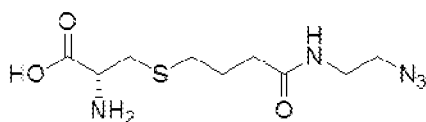
Formula III. 8;

(2S)-2-amino-3-[(2-[(prop-2-en-1-yl)carbamoyl]oxy)ethyl]sulfanyl]propanoic acid



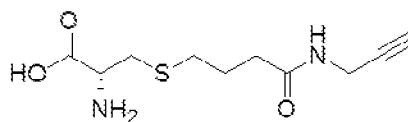
Formula III.9;

(2S)-2-amino-3-[(3-[(2-azidoethyl)carbamoyl]propyl)sulfanyl]propanoic acid



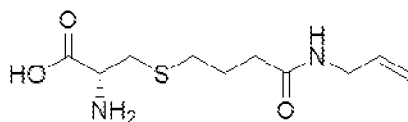
Formula III.10;

(2S)-2-amino-3-({3-[(prop-2-yn-1-yl)carbamoyl]propyl}sulfanyl)propanoic acid



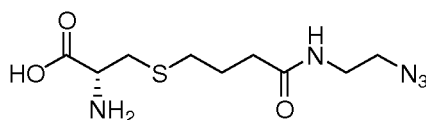
Formula III.11;

(2S)-2-amino-3-({3-[(prop-2-en-1-yl)carbamoyl]propyl}sulfanyl)propanoic acid



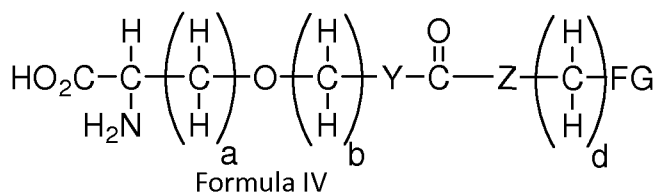
Formula III.12; and

(2S)-2-amino-3-({3-[(2-azidoethyl)carbamoyl]propyl}sulfanyl)propanoic acid



Formula III.13.

Still further pyrrollysine analogs have the structure of Formula IV:



wherein

Y = CH₂, NH, O or S;

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a = an integer 1-7;

b = an integer 1-7 save that when Z is NH, O or S then b is an integer 2-7;

provided that a+b is in the range 2-8;

and d = an integer 1-4.

In an embodiment, a+b is in the range 3-6 e.g., 3-4.

In an embodiment, a is 1.

In an embodiment, b is 2. In another embodiment b is 3.

In an embodiment, Y is NH and Z is O. In an embodiment, Y is NH and Z is NH. In an embodiment, Y is O and Z is O. In an embodiment, Y is CH₂ and Z is NH.

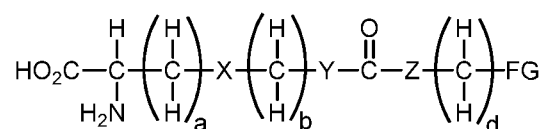
In an embodiment, FG is azide.

In an embodiment, FG is alkyne (e.g., ethynyl) or cycloalkyne.

In an embodiment, FG is alkene e.g., ethenyl.

In formula IV, when FG represents aryl, an example is aromatic halide e.g., 4-halo phenyl such as 4-iodo phenyl.

More generally, compounds of Formula III and IV are referred to as compounds of formula X:

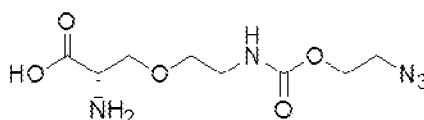


Formula X

wherein X represents O or S and all other variables are as defined for compounds of formula III or IV.

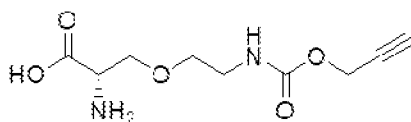
Exemplary compounds of Formula IV are as follows:

(2S)-2-amino-3-(2-([(2-azidoethoxy)carbonyl]amino)ethoxy)propanoic acid



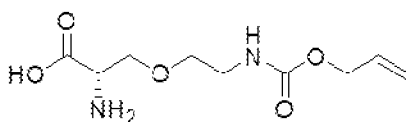
Formula IV.1;

(2S)-2-amino-3-(2-([(prop-2-yn-1-yloxy)carbonyl]amino)ethoxy)propanoic acid



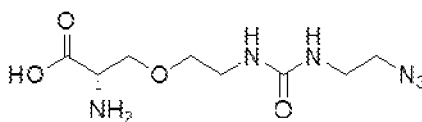
Formula IV.2;

(2S)-2-amino-3-(2-(((prop-2-en-1-yloxy)carbonyl)amino)ethoxy)propanoic acid



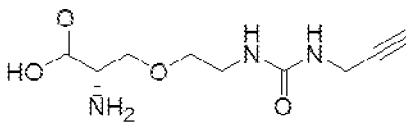
Formula IV. 3;

(2S)-2-amino-3-(2-(((2-azidoethyl)carbonyl)amino)ethoxy)propanoic acid



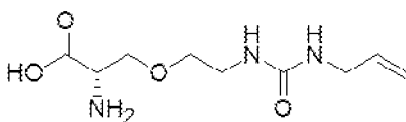
Formula IV.4;

(2S)-2-amino-3-(2-(((prop-2-yn-1-yl)carbonyl)amino)ethoxy)propanoic acid



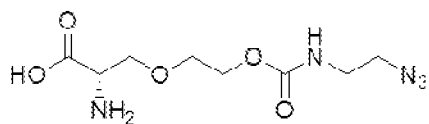
Formula IV.5;

(2S)-2-amino-3-(2-(((prop-2-en-1-yl)carbonyl)amino)ethoxy)propanoic acid



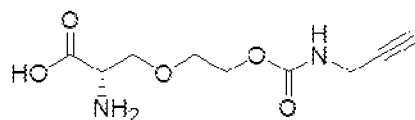
Formula IV.6;

(2S)-2-amino-3-(2-(((2-azidoethyl)carbonyl)oxy)ethoxy)propanoic acid



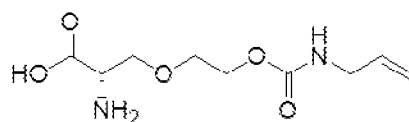
Formula IV.7;

(2S)-2-amino-3-(2-(((prop-2-yn-1-yl)carbamoyl)oxy)ethoxy)propanoic acid



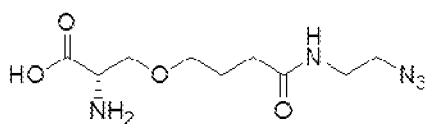
Formula IV. 8;

(2S)-2-amino-3-(2-(((prop-2-en-1-yl)carbamoyl)oxy)ethoxy)propanoic acid



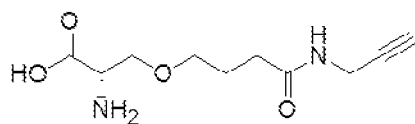
Formula IV. 9;

(2S)-2-amino-3-(3-(((2-azidoethyl)carbamoyl)propoxy)propanoic acid



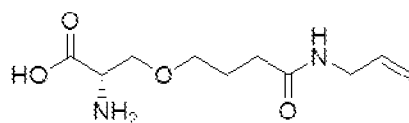
Formula IV. 10;

(2S)-2-amino-3-(3-(((prop-2-yn-1-yl)carbamoyl)propoxy)propanoic acid



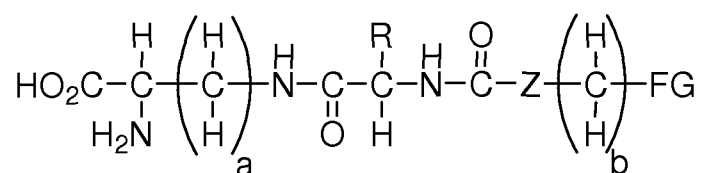
Formula IV. 11; and

(2S)-2-amino-3-(3-(((prop-2-en-1-yl)carbamoyl)propoxy)propanoic acid



Formula IV.12.

Additional pyrrolysine analogs have the structure of Formula V:



Formula V

wherein

R = the side chain of one of the 20 natural amino acids;

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a=1; and

b= an integer 1 to 4.

In an embodiment, Z is O.

In an embodiment, FG is azide.

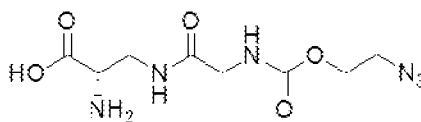
In an embodiment, FG is alkyne (e.g., ethynyl) or cycloalkyne.

In an embodiment, FG is alkene e.g., ethenyl.

In formula V, when FG represents aryl, an example is aromatic halide e.g., 4-halo phenyl such as 4-iodo phenyl.

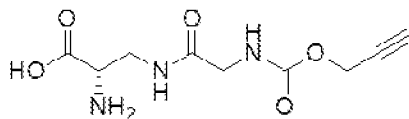
Exemplary compounds of Formula V are as follows:

(2S)-2-amino-3-(2-[[[(2-azidoethoxy)carbonyl]amino]acetamido]propanoic acid



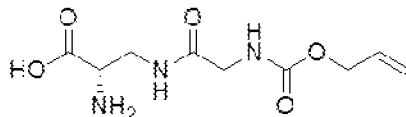
Formula V.1;

(2S)-2-amino-3-(2-[[[(prop-2-yn-1-yloxy)carbonyl]amino]acetamido]propanoic acid



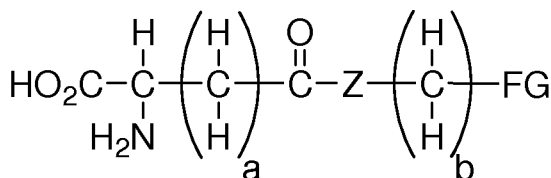
Formula V.2; and

(2S)-2-amino-3-(2-[[[(prop-2-en-1-yloxy)carbonyl]amino]acetamido]propanoic acid



Formula V.3.

Additional pyrrolysine analogs have structures of Formula VI:



Formula VI

wherein

Z = CH₂, CH-NH₂, CH-OH, NH, O and S;

FG = azide, alkene, alkyne, ketone, ester, aryl and cycloalkyne;

a = 4 or 5; and

b = an integer 1 to 4;

In an embodiment, Z is NH.

In an embodiment, a is 5.

In an embodiment, FG is azide.

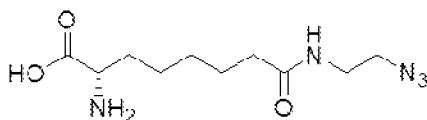
In an embodiment, FG is alkyne (e.g., ethynyl) or cycloalkyne.

In an embodiment, FG is alkene e.g., ethenyl.

In Formula VI, when FG represents aryl, an example is aromatic halide e.g., 4-halo phenyl such as 4-iodo phenyl.

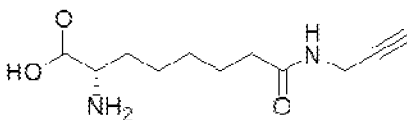
The following are exemplary compounds of Formula VI:

(2S)-2-amino-7-[(2-azidoethyl)carbamoyl]heptanoic acid



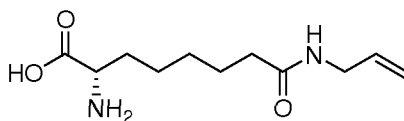
Formula VI.1;

(2S)-2-amino-7-[(prop-2-yn-1-yl)carbamoyl]heptanoic acid



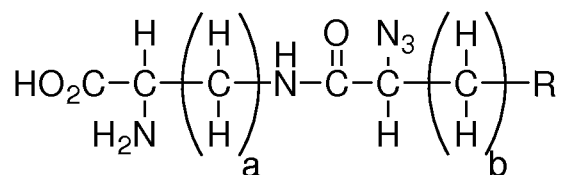
Formula VI.2;

(2S)-2-amino-7-[(prop-2-en-1-yl)carbamoyl]heptanoic acid



Formula VI.3;

Alternative pyrrolysine analogs have structure of Formula VII:



Formula VII

wherein

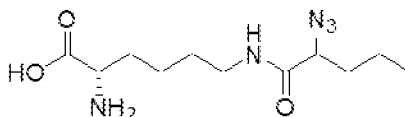
R = alkyl, alkenyl, cycloalkyl, aryl, heteroaryl or heterocyclyl;

a = an integer 1 to 7; and

b = an integer 1 to 3.

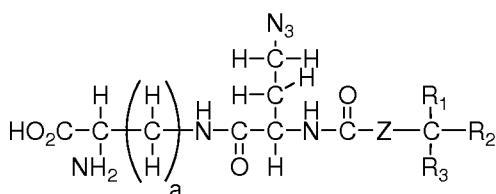
In some preferred embodiments, a=4.

A preferred compound of Formula VII is (2S)-2-amino-6-(2-azidopentanamido)hexanoic acid



Formula VII.1.

Alternative pyrrollysine analogs have structure of Formula (VIII):



Formula VIII

wherein

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

R₁ = H, alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl;

R₂ = alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl;

R₃ = H, alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl; and

a=1.

Suitably R₁ represents H or alkyl, such as H or methyl, especially H.

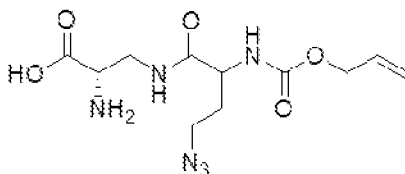
Suitably R₃ represents H or alkyl, such as H or methyl, especially H.

Suitably Z represents O.

R₂ may, for example, represent C₃₋₄alkenyl e.g., CH₂CH=CH₂ or C₁₋₄alkyl e.g., n-propyl or t-butyl.

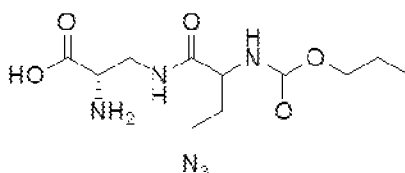
Exemplary compounds of Formula VIII are as follows:

(2S)-2-amino-3-(4-azido-2-[[[(prop-2-en-1-yloxy)carbonyl]amino]butanamido]propanoic acid



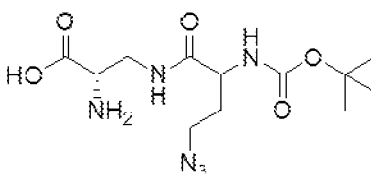
Formula VIII.1 ;

(2S)-2-amino-3-{4-azido-2-[(propoxycarbonyl)amino]butanamido}propanoic acid



Formula VIII.2; and

(2S)-2-amino-3-{4-azido-2-[(t-butoxycarbonyl)amino]butanamido}propanoic acid



Formula VIII.3

In structures of Formulae I to VI and X, when FG represents alkene, it suitably represents $-\text{CH}=\text{CH}_2$ or $-\text{CH}=\text{CH}-\text{CH}_3$, preferably $-\text{CH}=\text{CH}_2$. Examples of $-(\text{CH}_2)_b\text{-FG}$ and $-(\text{CH}_2)_d\text{-FG}$ are $-\text{CH}_2-\text{CH}=\text{CH}_2$ and $-\text{CH}_2-\text{CH}_2-\text{CH}=\text{CH}_2$.

In structures of Formulae I to VI and X, when FG represents alkyne, it suitably represents $-\text{C}\equiv\text{CH}$ or $-\text{C}\equiv\text{C}-\text{CH}_3$, preferably $-\text{C}\equiv\text{CH}$. An example of $-(\text{CH}_2)_b\text{-FG}$ and $-(\text{CH}_2)_d\text{-FG}$ is $-\text{CH}_2-\text{C}\equiv\text{CH}$.

In structures of Formulae I to VI and X, when FG represents ketone, it suitably represents $-\text{C}(=\text{O})-\text{CH}_3$ or $-\text{C}(=\text{O})-\text{CH}_2-\text{CH}_3$, preferably $-\text{C}(=\text{O})-\text{CH}_3$.

In structures of Formulae I to VI and X, when FG represents ester, it suitably represents $-\text{C}(=\text{O})\text{-Oalkyl}$ e.g., $-\text{C}(=\text{O})\text{-Oethyl}$.

In structures of Formulae I to VI and X, when FG represents aryl, it suitably represents phenyl substituted by halogen, especially iodine (e.g., 4-iodo-phenyl).

In structures of Formulae I to VI and X, when FG represents cycloalkyne, it suitably represents cyclooctyne, e.g., cyclooct-4,5-yne.

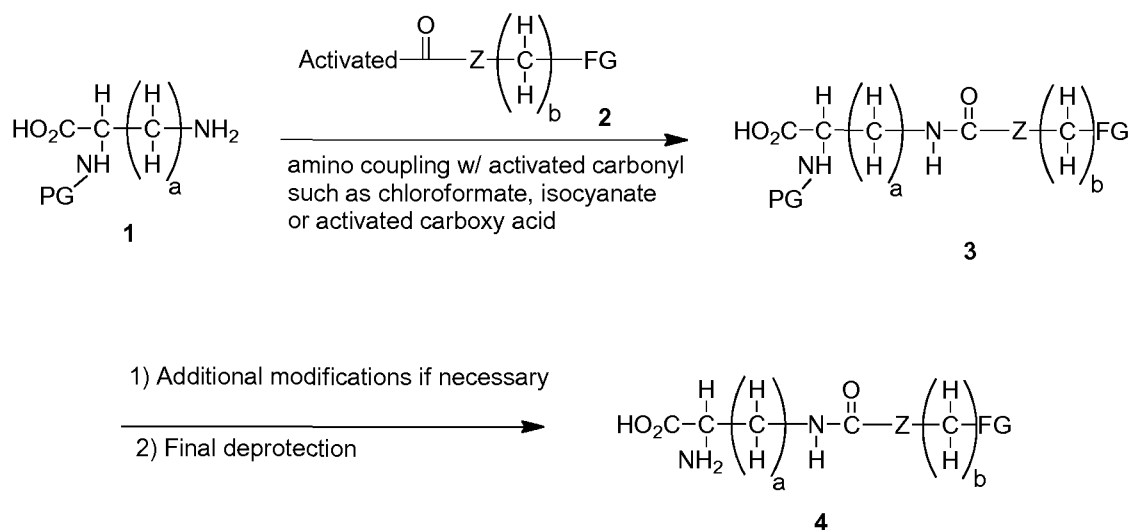
In structures of Formulae I to VI and X, when FG represents azide, an example of $-(CH_2)_b-FG$ or $-(CH_2)_d-FG$ is $-(CH_2)_2-N_3$.

In structures of Formulae VII and VIII when R , R^1 , R^2 and R^3 represent aryl, it suitably represents phenyl.

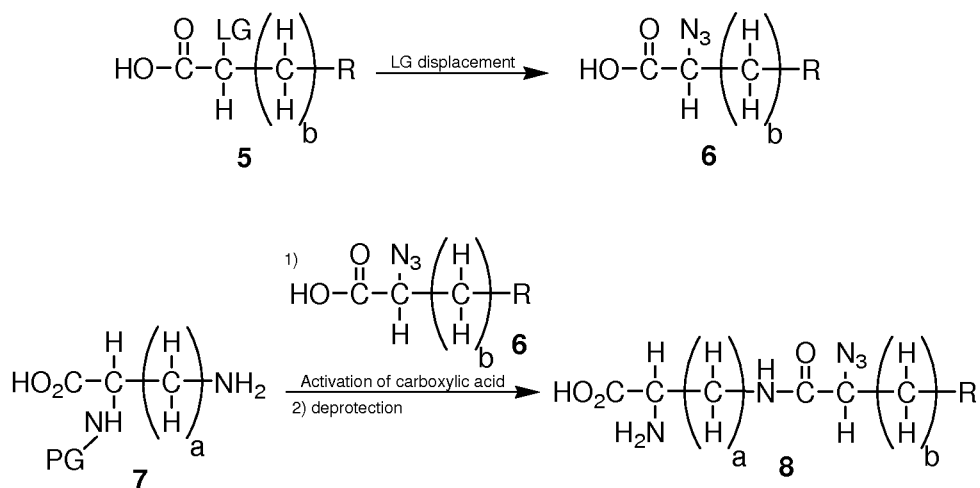
The pyrrollysine analogs disclosed herein can be made using various methods. The reaction conditions can generally be determined by one of the ordinary skill in the art.

Formula I analogs are readily prepared by the addition of an activated carbonyl group, such as a chloroformate, activated carboxylic acid ester, isocyanate, activated carbonate or sulfonyl halide to a mono-protected diamino substrate of type **1**, in which the α -amino group is protected by a protecting group ("PG") such as a Boc, Cbz, TFA, Acetyl or Fmoc group (see Scheme 1). The coupled product **3** can undergo further modifications, such as the displacement of halides with an azido nucleophile to install the desired functionality. Otherwise, the intermediate **3** is deprotected to remove the α -amino acid masking group to afford the desired Formula 1 analog.

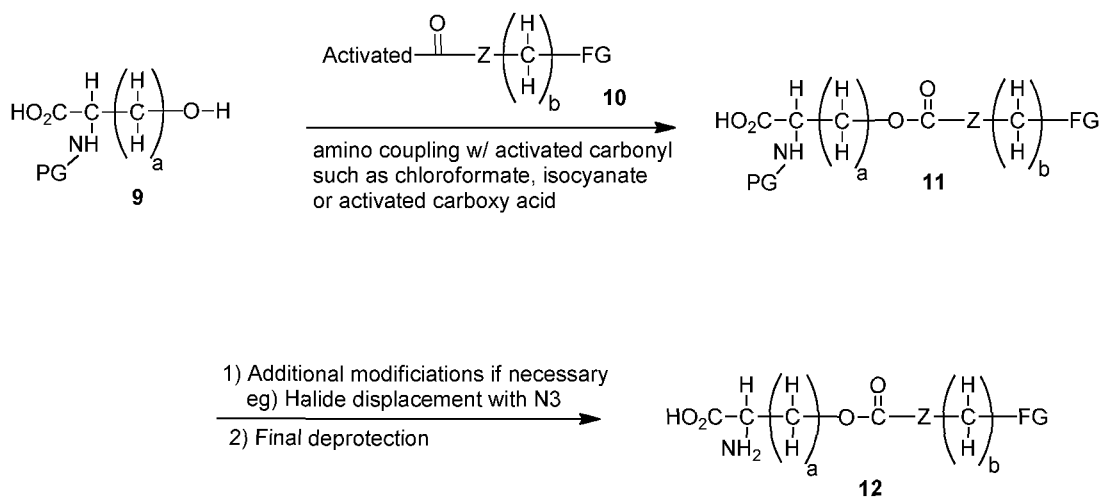
Scheme 1:



Formula VII analogs are readily prepared by first preparing the desired α -azido precursor. One strategy for accomplishing this is to displace a leaving group (LG) such as a halide (Cl, Br, or I) or alkyl sulfonate (tosylate, mesylate, triflate) with sodium azide to afford the α -azido acid **6**. This can then be coupled to a mono-protected diamino substrate of type **7** by activating the carboxylic acid with a group such as HBTU, DCC and NHS or CDI. The product **8** is obtained by removing the protecting group masking the α -amino group. Protecting groups may be used as per Scheme 1. See Scheme 2:

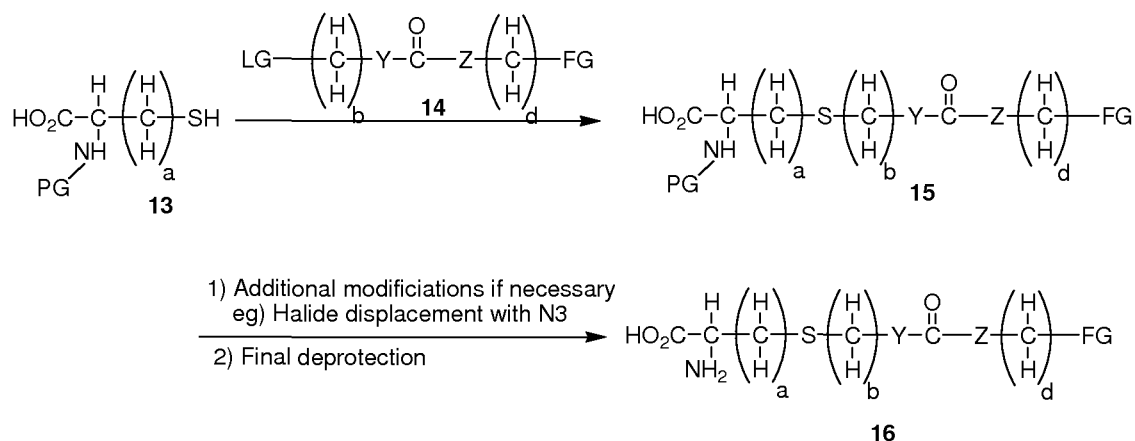


Formula II analogs were prepared by conjugation of hydroxyl amino acids **9** to substrates with activated carbonyls such as carboxylic acid ester, isocyanate, acid chlorides, activated carbonates or sulfonyl halides **10**. The coupled product **11** can undergo further modifications, such as the installation of the azide functional group by displacement of leaving groups such as halides or activated alcohols. The desired amino acid analog **12** is obtained by final deprotection to remove the α -amino acid masking group. Protecting groups may be used as per Scheme 1. See Scheme 3:

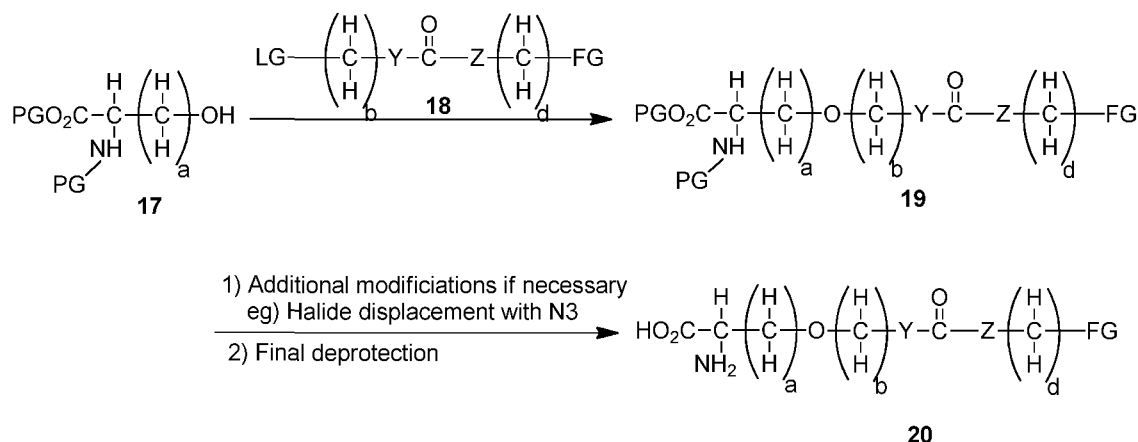


Formula III analogs which contain a thioether linkage are readily prepared by a key S-alkylation step. A thioamino acid such as cysteine **13** is treated with the desired functionalized alkylating agent **14** to affect the S-alkylation. Additional functionalization is then done to fix the amino acid in its final form (e.g., installation of azide group). Final deprotection as necessary reveals the desired amino acid **16**. Protecting groups may be used as per Scheme 1.

See Scheme 4:

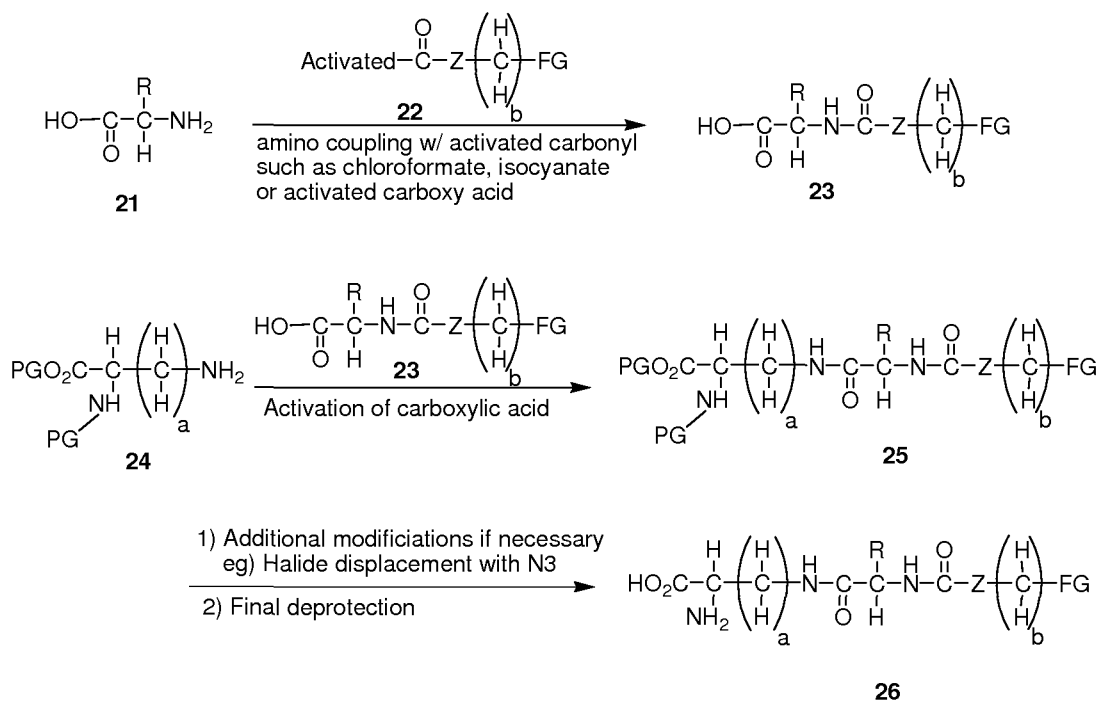


Formula IV analogs can be prepared in an analogous fashion to cysteine analogs in which the ether linkage is prepared by hydroxyl alkylation. A protected hydroxyl amino acid such as serine **17** is treated with alkylating agent **18** to affect the O-alkylation. Additional functionalization is then done to fix the amino acid in its final form (e.g., installation of azide group). Final deprotection as necessary reveals the desired amino acid **20**. Protecting groups may be used as per Scheme 1. See Scheme 5:

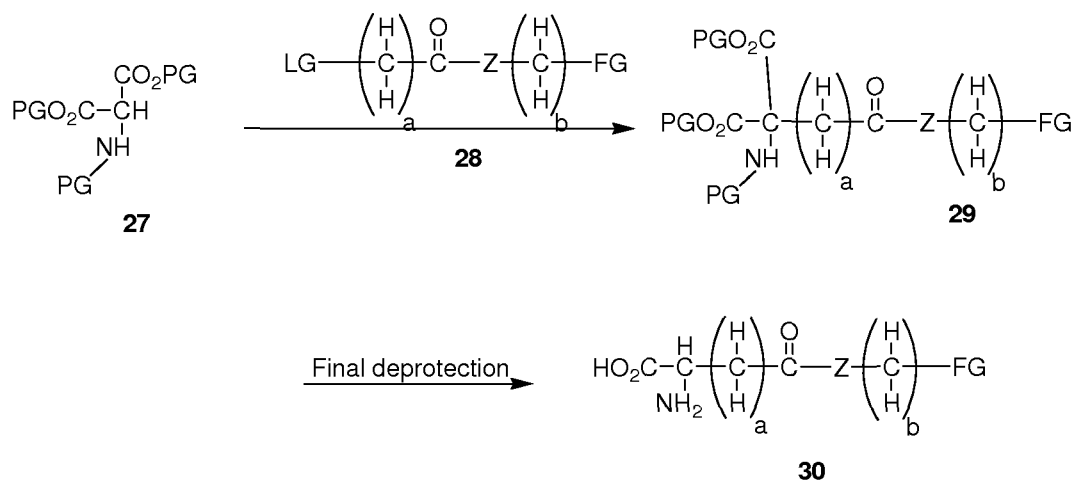


Formula V analogs were prepared in a straightforward manner by a key peptide coupling with a monoprotected diamino derivative **24**. The process begins with the coupling of bifunctional amine such as an α or β amino acid **21** to a activated carbonyl such as carboxylic acid ester, isocyanate, acid chlorides, activated carbonate or sulfonyl halides to afford intermediate **23**. The carboxylic acid of **23** is then activated with a reagent or combination of

reagents such as HBTU, DCC and NHS or CDI and used to acylate the monoprotected diamine such as diaminopropionic acid **24** to afford the peptide **25**. The intermediate **25** is either deprotected or further functionalized (e.g., azidation) just prior to deprotection to the desired amino acid **26**. Protecting groups may be used as per Scheme 1. See Scheme 6:

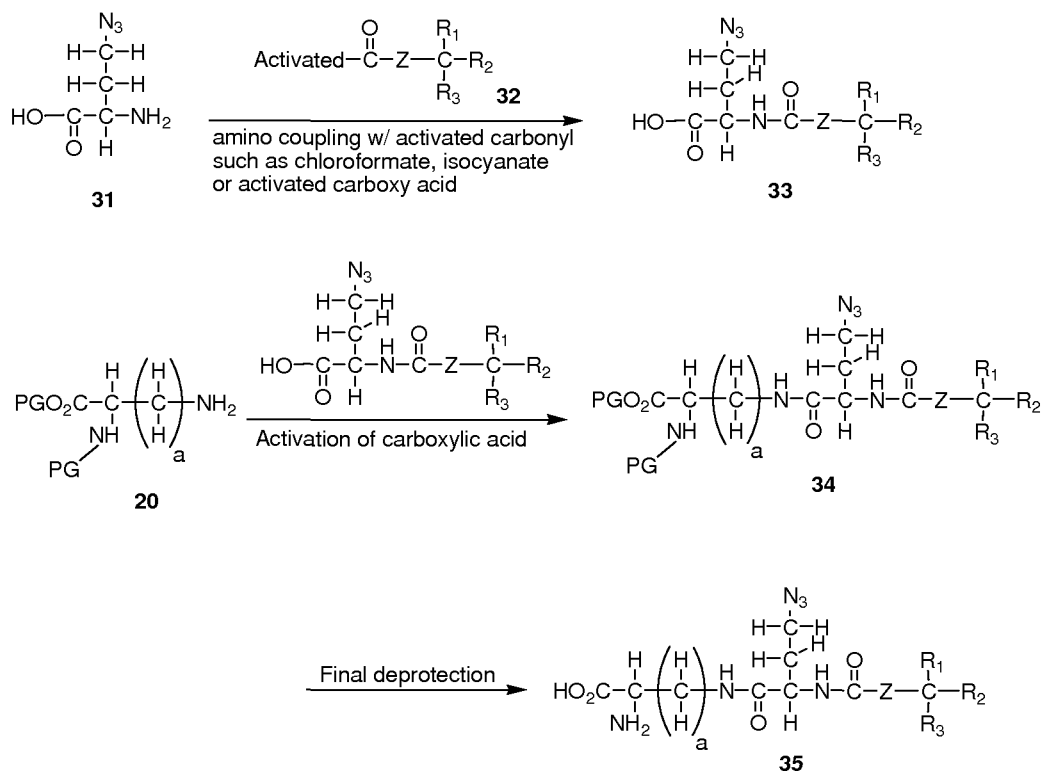


Formula VI analogs can be prepared by alkylation of the α -center of a glycine equivalent. In this sequence a protected glycine equivalent such as diethylacetamidomalonate **27** is treated with an alkylating agent such as **28** under basic conditions to effect alkylation at the α -position of the glycine derivative. Subsequent deprotection under acidic or basic conditions affords the desired amino acid **30**.



Formula VIII analogs were prepared by coupling an azidohomoalanine derivatives to the sidechain amine of a bifunctional amine such as **24**. The process begins with the coupling of azidohomoalanine **31** with an activated carbonyl such as a carboxylic acid ester, isocyanate, acid chlorides, activated carbonate or sulfonyl halides to afford intermediate **33**. The carboxylic acid of **33** is then activated with a reagent or combination of reagents such as HBTU, DCC and NHS or CDI and used to acylate the monoprotected diamine such as diaminopropionic acid **24** to afford the peptide **34**. The intermediate **34** is deprotected to afford **35**. Protecting groups may be used as per Scheme 1.

See Scheme 8:



Incorporation of non-natural amino acid into proteins

The pyrrolysine analogs disclosed herein can be incorporated into recombinant proteins. In particular, site specific incorporation of the analog into a recombinant protein can be achieved through amber suppression, wherein a nonsense (amber) codon is inserted within the nucleotide sequence encoding the recombinant protein, at a site where the pyrrolysine analog is to be inserted. The mutated nucleotide sequence, along with one or more plasmids encoding the PyIRS and tRNA_{pyl} are inserted into a cell of a cell free expression system.

The host cell may be a eukaryotic cell line which is transformed with a vector comprising a DNA construct as aforesaid.

Alternatively, a cell-free expression system is provided, wherein a synthesis reaction lysate obtained from a host cell comprises at least one component required for the synthesis of polypeptides. The synthesis reaction lysate is obtained from bacterial or eukaryotic cells.

Preferably, the synthesis reaction lysate is obtained from eukaryotic cells, more preferably, from rabbit reticulocytes or wheat germ.

Preferably, the cell-free expression system is capable of expressing WT PylRS and tRNA^{Pyl} of the present invention, wherein tRNA^{Pyl} is introduced into the cells used to obtain the synthesis reaction lysate with DNA constructs of the invention.

Cell-free expression systems suitable for use in the present invention are described for instance in WO2010081110, WO2010081111, WO2010083148, incorporated in their entirety herein by reference.

When the pyrrolysine analog is added to the cell or expression system, said analog is incorporated in the recombinant protein at the specified position. The nnAA and the tRNA^{Pyl} are bound by the pylRS and the tRNA^{Pyl} is subsequently aminoacylated with the nnAA. This tRNA^{Pyl} containing an amber anticodon is released into the cytosol where in response to an amber stop codon can interact with the ribosome and the nnAA released to form a peptide bond with the growing polypeptide chain.

Recombinant proteins modified to incorporate a pyrrolysine analog of the invention include all recombinant proteins amenable to site specific post translational modifications, e.g., therapeutic proteins, for instance cytokines, antibodies and antibody derivatives (such as Fab fragments, or single chain antibodies, e.g., single chain variable fragments (scfv)), peptides, enzymes, fusion proteins, decoy receptors, protein vaccines, protein hormones, e.g., insulin, growth factors, (e.g., human growth hormone, hGH, hGCSF, hFSH, hHCG). Further proteins modifiable with pyrrolysine analogs of the invention include diagnostic labels, imaging reagents.

Suitably, proteins may be modified site specifically to incorporate one or more than one nnAA (pyrrolysine analog) of the invention. For instance, an antibody may incorporate a nnAA of the invention at the heavy chain, or at the light chain, or at both light and heavy chain.

Site specific conjugation of proteins with incorporated non-natural amino acids

Proteins having incorporated pyrrolysine analogs of the present invention may be used for the preparation of functionalized protein conjugates. Molecules that may be conjugated to

proteins having incorporated non-natural amino acids include (i) other proteins, e.g., antibodies especially monoclonal antibodies; (ii) polymers e.g., PEG groups or other groups that may cause half life extension in the system; (iv) cytotoxic agents e.g., Auristatin F; and (v) drug moieties e.g., doxorubicin and moieties containing radioactive isotopes. Moreover these modified proteins can be conjugated to drugs or nucleotides for targeted delivery of these potent compounds.

More details of certain embodiments are given below in the discussion of antibody drug conjugates.

Pyrrolysine analogs may conveniently contain a unique chemical group permitting conjugation in a targeted fashion without risk of side reaction with other amino acids. For example non-natural amino acids may contain azide or alkyne groups permitting reaction with a molecule to be conjugated which contains a corresponding alkyne or azide group using the Huisgen 1,3-dipolar cycloaddition reaction.

Preferred conjugation chemistries of the invention include reactions which are orthogonal to the natural twenty amino acids. Such reactions do not interact or cause side reactions with the native 20 amino acids, they are specific to the functional groups associated with the reaction. Suitably the necessary functional groups are incorporated into the target protein via the pyrrolysine analogs of the present invention.

Further, said reactions proceed under conditions which are not destructive to the protein, for instance aqueous solvents, with a pH range which is acceptable to the protein and maintains its solubility, at a temperature which does not lead to deleterious effects upon the protein.

Increasing the stability of the attachment moiety between the protein and the linker can be advantageous. Conventional methods conjugate to the thiol groups of cysteine by reaction with a maleimide forming a thiol ether. The thiol ether can undergo the reverse reaction releasing the linker drug derivative from the antibody. In an embodiment of the invention, the conjugation chemistry employed between an azide and an alkyne results in an aromatic triazole which is significantly more stable, and not as prone to reversibility.

In addition, the product of the reaction, the linkage between protein and payload, ought to be stable, equal to or greater than the stability associated with conventional linkages (amide, thiol ether). Though not an impediment to conjugation, it is often advantageous if the conjugation reactions can be done under native conditions, as this will eliminate an extra refolding processing step.

Preferred chemical conjugations for production of conjugates of the invention include: a 3+2 alkyne-azide cycloaddition; 3+2 dipolar cycloaddition; Huisgen 3+2 cycloaddition; Copper promoted azide-alkyne cycloaddition (CuAAC); Ruthenium promoted azide alkyne cycloaddition (RAAC); metal promoted azide alkyne cycloaddition (MAAC); strain promoted azide alkyne cycloaddition (SPAAC); palladium based couplings including the Heck reaction, Sonogashira reaction, Suzuki reaction, Stille coupling, Hiyama/Denmark reaction olefin metathesis, Diels-alder reaction carbonyl condensation with hydrazine, hydrazide, alkoxy amine or hydroxyl amine; strain promoted cycloadditions with nitriles and nitrile oxides; electron promoted cycloaddition; fragment extrusion cycloaddition; alkene cycloaddition followed by a β -elimination reaction.

According to one preferred embodiment, the incorporated amino acid contains an azide or an alkyne group and the process of chemical modification comprises reacting said azide or alkyne group with a reagent comprising an alkyne or azide group. The envisaged reaction is a Huisgen 1,3-dipolar cycloaddition reaction which leads to production of a triazole linkage. The reagent comprising an alkyne or azide group may be a protein (e.g., an antibody) or a cytotoxic agent or a drug or a substance suitable for half life extension (e.g., a PEG group) which carries an alkyne or azide group optionally via a linker.

The site specific conjugations between the incorporated nnAA and the target payload can be done with fully folded proteins such as antibodies, antibody fragments, and cytokines.

Alternatively, the conjugation can be done on denatured proteins in the presence of denaturants such as sodium dodecylsulfate and urea. The copper catalyzed azide alkyne addition can be done in the presence of denaturants and reducing agents such as dithiothreitol and 2-mercaptoethanol.

When more than one nnAA is incorporated into a target protein (e.g., an antibody), the chemical modification may be the same or different. For example if two nnAAs are incorporated, one may be modified to be conjugated to a drug moiety and one may be modified to be conjugated to a PEG moiety.

Conveniently, upon incorporation of more than one nnAA of the invention bearing different but complementary reactive groups, said nnAAs can react with each other to generate an intramolecular link.

In an embodiment, conjugation chemistry of the invention is used for preparing an antibody drug conjugate. The conjugation chemistry may also be used to assemble antibody-protein conjugates, protein protein conjugates such as bispecifics composed of antibody fragments. The conjugation chemistry may also be used to conjugate polymer bond drug conjugates to targeting agents such antibodies and antibody fragments. The conjugation chemistry can also be used to attach polymers such as PEG to proteins to manipulate pharmacokinetic properties.

PEG moieties

Target proteins may be conjugated to PEG moieties. PEG moieties may be incorporated into antibody drug conjugates. The PEG moiety may typically have a molecular weight ranging between 5kDa and 40 kDa. More preferably, the PEG moiety may have a molecular weight of around 20 kDa. PEG moieties may be straight chain or branched.

Antibody Drug Conjugates (ADCs)

Pyrrolysine analogs according to the invention are particularly useful for production of Antibody Drug Conjugates (recombinant antibody covalently bound by a synthetic linker to a given drug, typically a cytotoxic drug, or else a protein or a PEG group) which are homogeneous nature, in which the number of drugs (or other conjugated molecule) per antibody and position of those drugs upon the antibody are explicitly controlled, whereby monoclonal antibodies containing incorporated non-natural amino acids are obtained and site specifically conjugated to a linker carrying a drug moiety (or other conjugated molecule) through orthogonal chemistry.

ADCs obtained with pyrrolysine analogs of the present invention may be manufactured following methods including the following steps:

1. Introducing into a stable cell line of the invention one or more plasmids carrying the DNA sequence coding for a full length antibody, whereby a stop codon is introduced at specific positions within the sequence
2. Purify the modified antibody with the pyrrolysine analog (nnAA) installed at desired position(s).
3. React a cytotoxin-linker derivative modified to include a functional group complimentary to the nnAA installed in the antibody with the modified antibody containing a complementary reactive group through an orthogonal chemistry
4. Purify the resulting ADC

Thus, the present invention also provides ADCs whereby the antibody component has been modified to incorporate non natural amino acids bearing a unique reactive functional group at desired positions, whereby such functional group allows conjugation to a drug moiety (or protein or PEG group).

In an embodiment the present invention provides an antibody conjugate comprising an anti-Her-2 antibody which is conjugated to one or more moieties (e.g., one, two, three or four, preferably one or two, especially one) selected from protein, drug and PEG moieties via linkers comprising a triazole moiety.

In particular, the triazole moiety may be formed by reaction of an azide or alkyne moiety in the side chain of a non-natural amino acid incorporated into the sequence of the anti-Her-2 antibody and an alkyne or azide moiety attached to the protein, drug or PEG moiety.

In one embodiment, the triazole moiety is formed by reaction of an azide or alkyne moiety in the side chain of a non-natural amino acid incorporated into the sequence of the anti-Her-2 antibody and an alkyne or azide moiety attached to the protein, drug or PEG moiety under conditions of Cu(I) catalysis.

In one embodiment a copper azide alkyne cycloaddition is used for the conjugation. Suitably, the reaction utilizes a cytotoxic agent such as auristatin, amanitin, taxol or doxorubicin bearing

a terminal alkyne. Further, the reaction employs a copper source such as copper sulfate, copper acetate, copper iodide or copper bromide; a reducing agents such as sodium ascorbate, hydrazine, hydroxylamine, sodium bisulfite, dithiothreitol, cysteine, b-mercaptoethanol; a copper chelating ligand such as Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) or Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA). Suitably, the reaction may be performed at 4 – 50°C. Suitably, the reaction time ranges from 0.5 to 48 hrs. In an alternative embodiment, a strain promoted azide alkyne cycloaddition is used for conjugation. Suitably, the reaction utilizes a dye, a PEG polymer, or cytotoxic agent such as auristatin bearing a cyclooctyne group. Suitably, the reaction is allowed to incubate at room temperature for 0.5 – 48 hrs.

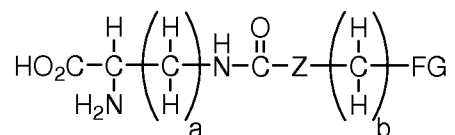
Salts

Pyrrolysine analogues described herein may optionally be employed in the form of a salt. Any such salts form an aspect of the invention. Salts of carboxylic acids may include salts formed with Group 1 and Group 2 metals, especially soluble salts such as sodium and potassium salts. Salts of amines may include salts formed with weak and strong acids, such as HCl, HBr or acetic acid.

Embodiments of the invention

Embodiments of the invention are represented as follows:

1. A pyrrolysine analogue of formula I:



wherein

Z = bond, CH₂, CH-NH₂, CH-OH, NH, O, S or CH-NH₂;

a is an integer 3-7;

b is 0 or an integer 1-7; and

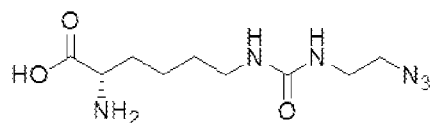
FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne.

2. A pyrrolysine analogue according to embodiment 1 wherein a is 3.

3. A pyrrolysine analogue according to embodiment 1 wherein a is 4.

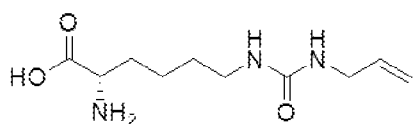
4. A pyrrolysine analogue according to embodiment 3 selected from

Formula IA.4,



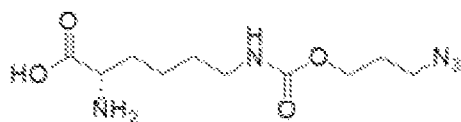
;

Formula IA.5,



;

Formula IA.6,



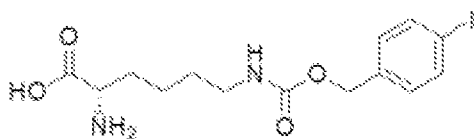
;

Formula IA.7,



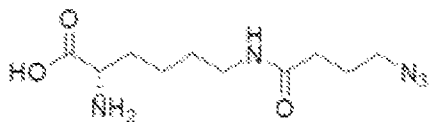
;

Formula IA.8,



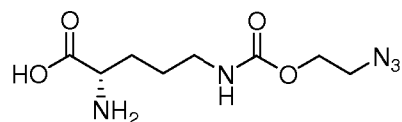
; and

Formula IA. 9,



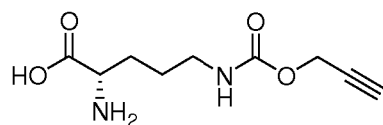
5. A pyrrolysine analogue according to embodiment 2 selected from

Formula IB.1,



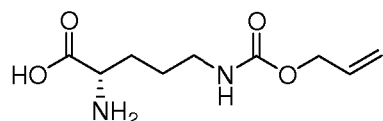
;

Formula IB.2,

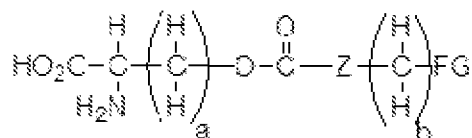


; and

Formula IB.3,



6. A pyrrolysine analogue of formula II:



wherein

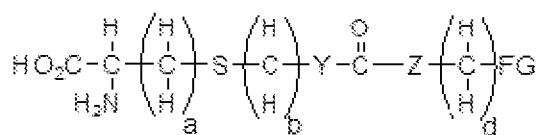
Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a = an integer 3 or 5-7; and

b = an integer 1-4.

7. A pyrrolysine analogue of formula III:



wherein

Y = CH₂, NH, O or S;

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a = an integer 1-7;

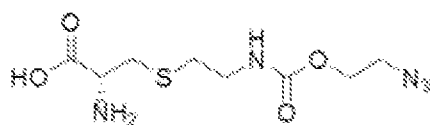
b = an integer 1-7 save that when Z is NH, O or S then b is an integer 2-7;

provided that a+b is in the range 2-8;

and d = an integer 1-4.

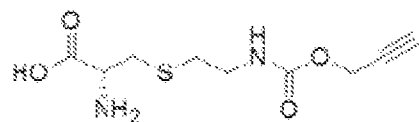
8. A pyrrollysine analogue according to embodiment 7 selected from

Formula III.1



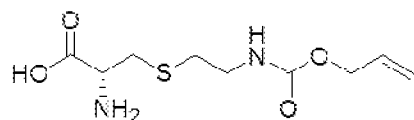
;

Formula III.2



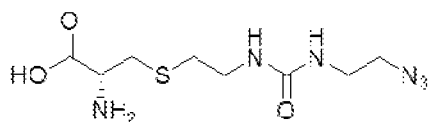
;

Formula III.3



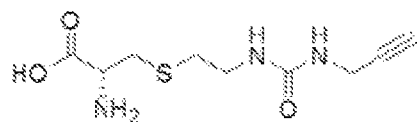
;

Formula III.4



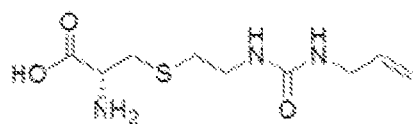
;

Formula III. 5



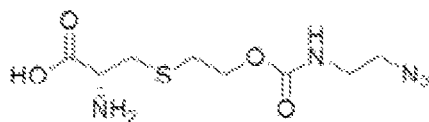
;

Formula III. 6



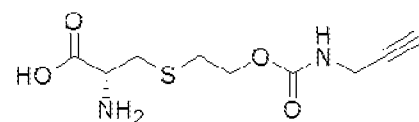
;

Formula III.7



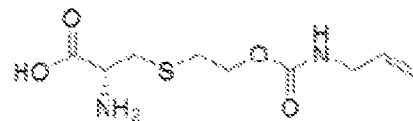
;

Formula III. 8;



;

Formula III.9



;

Formula III.10



;

Formula III.11



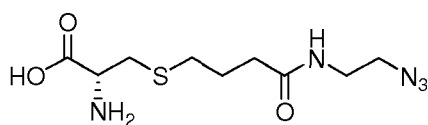
;

Formula III.12

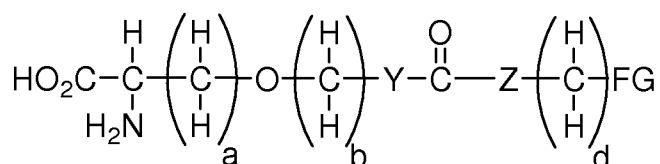


; and

Formula III.13



9. A pyrrolysine analogue of formula IV:



wherein

Y = CH₂, NH, O or S;

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a = an integer 1-7;

b = an integer 1-7 save that when Z is NH, O or S then b is an integer 2-7;

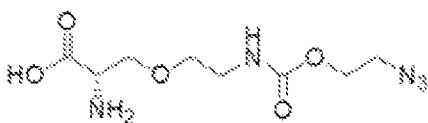
provided that a+b is in the range 2-8;

and

d = an integer 1-4.

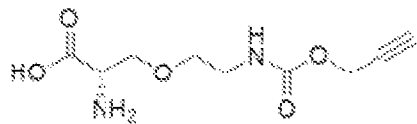
10. A pyrrolysine analogue according to embodiment 9 selected from

Formula IV.1



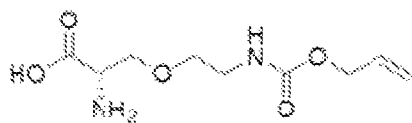
;

Formula IV.2



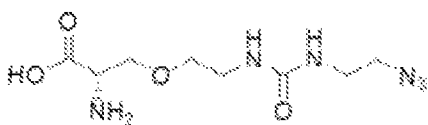
;

Formula IV. 3



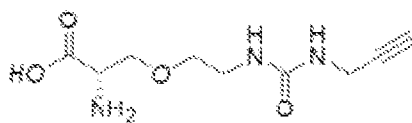
;

Formula IV.4



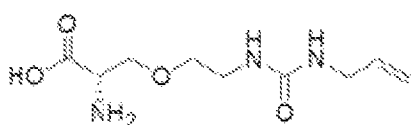
;

Formula IV.5



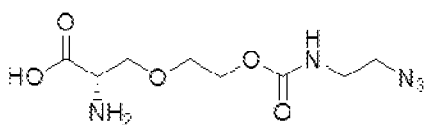
;

Formula IV.6



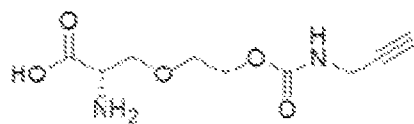
;

Formula IV.7



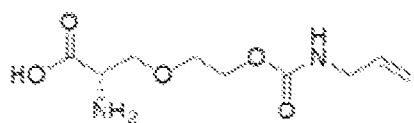
;

Formula IV. 8



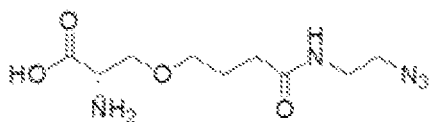
;

Formula IV. 9



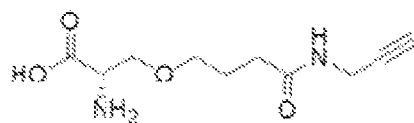
;

Formula IV.10



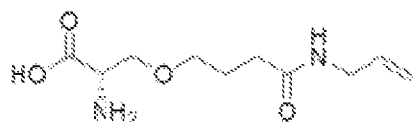
;

Formula IV. 11

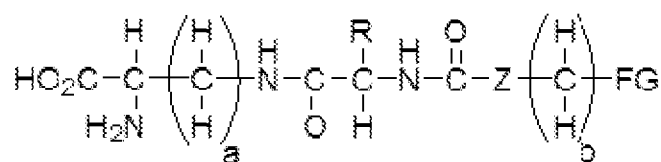


; and

Formula IV.12



11. A pyrrolysine analogue of formula V:



wherein

R = the side chain of one of the 20 natural amino acids;

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

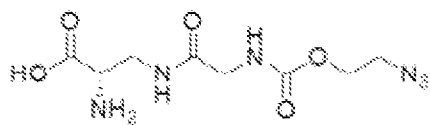
FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a=1; and

b= an integer 1 to 4.

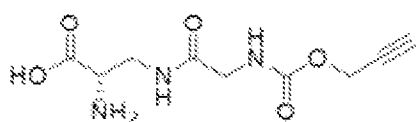
12. A pyrrolysine analogue according to embodiment 11 selected from

Formula V.1



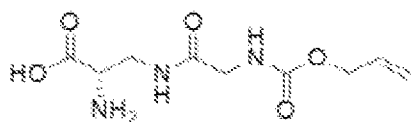
;

Formula V.2

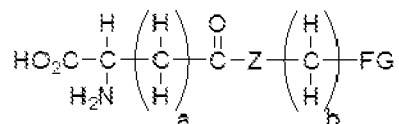


; and

Formula V.3



13. A pyrrolysine analogue of formula VI:



wherein

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

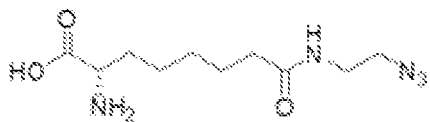
FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a=4 or 5; and

b= an integer 1 to 4;

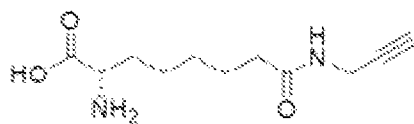
14. A pyrrolysine analogue according to embodiment 13 selected from

Formula VI.1,



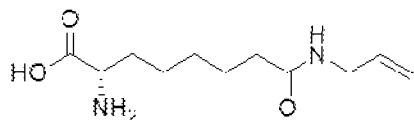
;

Formula VI.2,

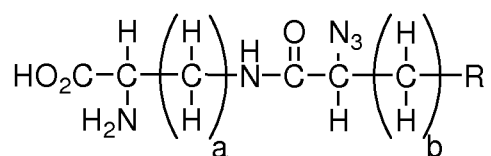


; and

Formula VI.3.



15. A pyrrolysine analogue of formula VII:



wherein

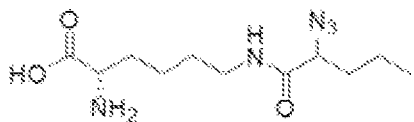
R = alkyl, alkenyl, cycloalkyl, aryl, heteroaryl or heterocyclyl;

a = an integer 1 to 7; and

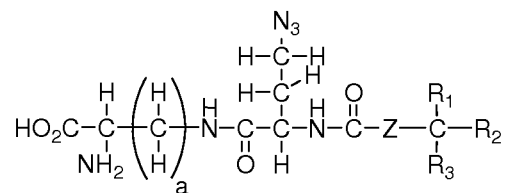
b = an integer 1 to 3.

16. A pyrrolysine analogue according to embodiment 15 which is:

Formula VII.1



17. A pyrrolysine analogue of formula VIII:



wherein

Z = CH₂, CH-NH₂, CH-OH, NH, O or S;

R₁ = H, alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl;

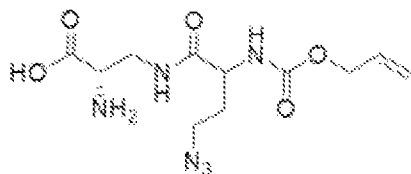
R_2 = alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl;

R_3 = H, alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl; and

$a=1$.

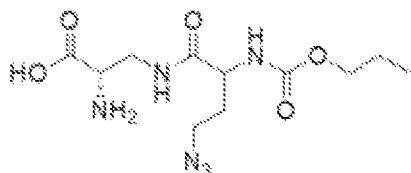
18. A pyrrolysine analogue according to embodiment 17 selected from

Formula VIII.1,



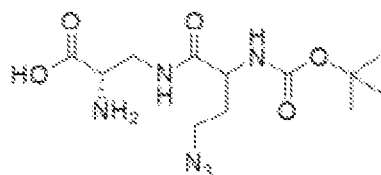
;

Formula VIII.2,



; and

Formula VIII.3



19. A mutant protein containing as non-natural amino acid one or more pyrrolysine analogues according to any one of embodiments 1 to 18.

20. A mutant protein according to claim 19 containing as non-natural amino acid one pyrrolysine analogue according to any one of embodiments 1 to 18.

21. A mutant protein according to embodiment 19 or embodiment 20 which is conjugated via the one or more non-natural amino acids to one or more moieties selected from proteins, cytotoxic agents, drugs and polymers.

22. A mutant protein according to embodiment 21 which is conjugated to a PEG moiety.

23. A mutant protein according to embodiment 21 which is conjugated to an antibody moiety.

24. A mutant protein according to embodiment 21 which is conjugated to a cytotoxic agent moiety.
25. A mutant protein according to embodiment 21 which is conjugated to a drug moiety.
26. Use of a pyrrolysine analogue according to any one of embodiments 1 to 18 in the manufacture of a mutant protein containing one or more non-natural amino acids.

EXAMPLES

Example 1. Preparation of Formula I and IB Analogs

Preparation of (2S)-2-amino-6-[(2S)-2-amino-4-azidobutanamido]hexanoic acid (Formula IA.7). In a 20mL vial with magnetic stirrer was placed N-Boc-azidohomoalanine (200mg, 1eq) and HBTU (311mg, 1eq) in 4mL of DMF. The mixture was stirred for 15min and then a solution N-Boc-Lysine (235mg, 1eq) in DMF (3mL) was added, followed by triethylamine (228uL). The vial was capped and stirred for 4h. The mixture was partitioned between 250mM citric acid and ethyl acetate. The organic layer was retained and the aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired intermediate. Additional purification was conducted via silica gel chromatography.

The Boc protected intermediate was placed in a 20mL vial and suspended in acetonitrile (4mL). To this was added a solution of hydrochloric acid in dioxane (4N, 4mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized. Additional purification by ion exchange resin (Dowex-50) afforded the desired amino acid. The product was confirmed by mass spectrometry.

Preparation of ((2S)-2-amino-6-[(2-azidoethyl)carbamoyl]amino}hexanoic acid (Formula IA.4). In a 4mL vial with magnetic stirrer was placed Boc-N-6-Lysine (50mg, 1eq) and DMF (1mL). To this was added 2-chloroethyl isocyanate (17.3mg, 1.0eq) and pyridine (32.3uL, 2 eq). The vial was capped and allowed to stir for 4h. The solution was transferred to an extraction funnel, diluted with ethylacetate and 100mM citric acid. The mixture shaken and the layers separated. The aqueous layer was extracted with ethyl acetate two additional times. The organic layers combined, washed with 5% lithium chloride, dried with sodium sulfate, filtered

and concentrated. The product was identified by mass spectrometry and taken forward into the next step directly.

In a 4mL vial with magnetic stirrer was placed the crude chloro derivative and DMSO (1mL). Sodium azide (130mg, 5eq) and pyridine (32.3uL, 2 eq) were added to the mixture and the vial was capped. The mixture was stirred overnight at 60°C. The mixture was transferred to an extraction funnel and diluted with 100mM citric acid and ethyl acetate. The mixture was shaken and the layers separated. The aqueous layer was extracted with ethyl acetate two additional times. The organic layers combined, washed with 5% lithium chloride, dried with sodium sulfate, filtered and concentrated. The intermediate was identified by mass spectrometry and carried on to the next step.

In a 20mL vial was placed the crude and acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized to a semi solid and used in translational testing. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-6-(((prop-2-en-1-yl)carbamoyl)amino)hexanoic acid (Formula IA.5). In a 4mL vial with magnetic stirrer was placed Boc-N-6-hydroxynorleucine (50mg, 1eq) and DMF (1.5mL). To this was added allyl isocyanate (18.0uL, 1.0eq) and pyridine (32.3uL, 2 eq). The vial was capped and allowed to stir for 4h. The solution was transferred to an extraction funnel, diluted with ethylacetate and 100mM citric acid. The mixture shaken and the layers separated. The aqueous layer was extracted with ethyl acetate two additional times. The organic layers combined, washed with 5% lithium chloride, dried with sodium sulfate, filtered and concentrated. The product was identified by mass spectrometry and taken forward into the next step directly.

In a 20mL vial was placed the crude hydroxyl leucine – allyl carbamate derivative in acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2.5mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized to a semi solid and used in translational testing. The product was confirmed by mass

spectrometry. Additional purification could be done with ion exchange chromatography (DOWEX-50).

Preparation of (2S)-2-amino-5-[[[(2-azidoethoxy)carbonyl]amino]pentanoic acid (Formula IB.1). In a 20mL vial with magnetic stirrer was placed N-Boc-Ornithine (500mg, 1eq) in a solution of potassium carbonate (1M, 2.5mL) and dioxane (2.5mL). 2-chloroethyl chloroformate (223uL, 1eq) was added and the vial was capped and stirred for 4h. The mixture was acidified to near pH 2 with 1M Citric acid. The solution was transferred to an extraction funnel and the aqueous layer extracted with ethyl acetate (3X). The organic layers combined, dried with sodium sulfate and concentrated. LC/MS confirmed the major product was the desired mass. The crude intermediate was carried forward to the next step without additional purification.

The crude chloride intermediate was placed in a 20mL vial and dissolved in DMSO (15mL). Sodium azide (625mg, 4eq) and pyridine (155uL) were added and the vial was capped and stirred overnight at 60°C. The mixture was poured into an extraction funnel with 250mM citric acid and extracted with ethyl acetate (3X). The organic layers were combined, dried with sodium sulfate and concentrated. LC/MS confirmed the formation of the product. Additional purification was conducted via silica gel chromatography.

The Boc Intermediate was placed in a 20mL vial and suspended in acetonitrile (4mL). To this was added a solution of hydrochloric acid in dioxane (4N, 4mL). The solution was stirred for 2h and then concentrated under reduced pressure. The concentrate was taken up in water and applied to an ion exchange resin (Dowex-50). The capture material was washed with water and eluted with ammonium bicarbonate. The enriched fractions with the amino acid were identified by TLC, pooled and lyophilized to afford the desired product. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-5-[[[(prop-2-en-1-yloxy)carbonyl]amino]pentanoic acid (Formula IB.3). In a 20mL vial with magnetic stirrer was placed N-Boc-Ornithine (300mg, 1eq) in a solution of potassium carbonate (1M, 2mL) and dioxane (2mL). Allyl chloroformate (137uL, 1eq) was added and the vial was capped and stirred for 4h. The mixture was acidified to near

pH 2 with 1M Citric acid. The solution was transferred to an extraction funnel and the aqueous layer extracted with ethyl acetate (3X). The organic layers combined, dried with sodium sulfate and concentrated. LC/MS confirmed the major product was the desired mass. Additional purification was done by silica gel chromatography.

The Boc Intermediate was placed in a 20mL vial and suspended in acetonitrile (3mL). To this was added a solution of hydrochloric acid in dioxane (4N, 3mL). The solution was stirred for 2h and then concentrated under reduced pressure. The concentrate was taken up in water and applied to an ion exchange resin (Dowex-50). The capture material was washed with water and eluted with ammonium bicarbonate. The enriched fractions with the amino acid were identified by TLC, pooled and lyophilized to afford the desired product. The product was confirmed by mass spectrometry.

Example 2. Preparation of compounds of Formula III

Preparation of (2S)-2-amino-3-[(2-[(2-azidoethoxy)carbonyl]amino)ethyl)sulfanyl]propanoic acid (Formula III.1). In a 4mL vial with magnetic stirrer was placed cysteine (75mg, 1eq) and a solution of sodium carbonate (2.8M, 620uL). To this was added 2-chloroethyl-N-(2-bromoethyl)carbamate (99mg, 1.0eq) in DMSO (620uL). The vial was capped and allowed to stir for 8h. Mass spectrometry indicated the formation of the desired s-alkylated intermediate. Following confirmation of the intermediate mass, sodium azide (120mg) and DMSO (1mL) were added, the vial was capped once again and the mixture heated to 60°C overnight. Mass spectrometry indicated the successful formation of the desired mass. The pH of the crude mixture was adjusted to 2, lyophilized to a paste. The material was suspended in water and captured on an ion exchange resin (DOWEX-50), washed with water and eluted with ammonium hydroxide. The eluted fractions were pooled and lyophilized to afford the desired amino acid. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-3-[(2-[(prop-2-yn-1-yloxy)carbonyl]amino)ethyl)sulfanyl]propanoic acid (Formula III.2). In a 4mL vial with magnetic stirrer was placed cysteine (75mg, 1eq) and a solution of sodium carbonate (2.8M, 620uL). To this was added prop-2-yn-1-yl N-(2-bromoethyl)carbamate (86mg, 1.0eq) in dioxane

(620ul). The vial was capped and allowed to stir overnight. Mass spectrometry indicated the formation of the desired S-alkylated product. The pH of the crude mixture was adjusted to 2, lyophilized to a paste. The material was suspended in water and captured on an ion exchange resin (DOWEX-50), washed with water and eluted with ammonium hydroxide. Enriched fractions were pooled and lyophilized to afford the desired amino acid. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-3-[(2-[(prop-2-en-1-

yl)oxy)carbonyl]aminoethyl)sulfanyl]propanoic acid (Formula III.3). In a 4mL vial with magnetic stirrer was placed cysteine (75mg, 1eq) and a solution of sodium carbonate (2.8M, 620uL). To this was added prop-2-en-1-yl N-(2-bromoethyl)carbamate (88mg, 1.0eq) in dioxane (620uL). The vial was capped and allowed to stir overnight. Mass spectrometry indicated the formation of the desired S-alkylated product. The pH of the crude mixture was adjusted to 2, lyophilized to a paste. The material was suspended in water and captured on an ion exchange resin (DOWEX-50), washed with water and eluted with ammonium hydroxide. Enriched fractions were pooled and lyophilized to afford the desired amino acid. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-3-[(3-[(2-azidoethyl)carbonyl]propyl)sulfanyl]propanoic acid (Formula III.13). In a 20mL vial with magnetic stirrer was placed cysteine (75mg, 1eq) and a solution of sodium bicarbonate (1M, 1.7mL). To this was added N-(2-azidoethyl)-4-bromobutanamide (86mg, 1.0eq) in dioxane (1.5mL). The vial was capped and allowed to stir for 8h. Mass spectrometry indicated the formation of the desired S-alkylated product. The pH of the crude mixture was adjusted to 2 and applied to an ion exchange resin (DOWEX-50), washed with water and eluted with ammonium hydroxide. Fractions with the amino acid as identified by TLC, were pooled and lyophilized to afford the desired amino acid. The product was confirmed by Mass spectrometry.

Preparation of (2S)-2-amino-3-[(3-[(prop-2-en-1-yl)carbonyl]propyl)sulfanyl]propanoic acid (Formula III.12). In a 20mL vial with magnetic stirrer was placed cysteine (285mg, 1eq) and a solution of sodium carbonate (2.8M, 1.2mL). To this was added 4-bromo-N-(prop-2-en-1-yl)butanamide (333mg, 1.0eq). The vial was capped and allowed to stir overnight. Mass

spectrometry indicated the formation of the desired s-alkylated product. The pH of the crude mixture was adjusted to 2, lyophilized to a paste. The material was suspended in water and captured on an ion exchange resin (DOWEX-50), washed with water and eluted with ammonium hydroxide. Enriched fractions were pooled and lyophilized to afford the desired amino acid. The product was confirmed by mass spectrometry.

Example 3. Preparation of compounds of Formula V

Preparation of (2S)-2-amino-3-(2-{{(2-azidoethoxy)carbonyl}amino}acetamido)propanoic acid (Formula V.1).

In a 4mL vial with magnetic stirrer was placed 2-{{(2-chloroethoxy)carbonyl}amino}acetic acid (50mg, 1eq) and HBTU (105mg, 1eq) in 1mL of DMF. The mixture was stirred for 15min and then a solution of (2S)-3-amino-2-{{(tert-butoxy)carbonyl}amino}propanoic acid (66mg, 1eq) in DMF (1mL) was added, followed by triethylamine (77uL). The vial was capped and stirred for 4h. The mixture was poured onto 250mM citric acid and ethyl acetate and shaken. The layers were allowed to separate and the organic layer was retained. The aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired product. The crude material was taken forward into the next step without additional purification.

The chloride intermediate was placed in a 20mL vial and dissolved in DMSO (5mL). Sodium azide (72mg, 4eq) was added and the vial was capped and stirred at 60°C overnight. The mixture was poured onto 250mM citric acid and ethyl acetate and shaken. The layers were allowed to separate and the organic layer was retained. The aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired product. Additional purification could be accomplished by silica gel chromatography.

The boc group was removed by simple acidification. The Boc-azide intermediate was placed in a 20mL vial and suspended in acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized to a paste. Additional purification with ion

exchange resin (Dowex-50) afforded the desired amino acid. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-3-(2-(((prop-2-en-1-yloxy)carbonyl)amino)acetamido)propanoic acid (Formula V.3). In a 4mL vial with magnetic stirrer was placed 2-(((prop-2-en-1-yloxy)carbonyl)amino)acetic acid (50mg, 1eq) and HBTU (120mg, 1eq) in 1mL of DMF. The mixture was stirred for 15min and then a solution of (2S)-3-amino-2-(((tert-butoxy)carbonyl)amino)propanoic acid (75mg, 1eq) in DMF (1mL) was added, followed by triethylamine (88uL). The vial was capped and stirred for 8h. The mixture was partitioned between 250mM citric acid and ethyl acetate. The organic layer was retained and the aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired product. Additional purification could be accomplished by silica gel chromatography.

The Boc-protected intermediate was placed in a 20mL vial and suspended in acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized to a paste. Additional purification by ion exchange resin (Dowex-50) afforded the desired amino acid. The product was confirmed by mass spectrometry.

Preparation of (2S)-2-amino-3-(2-(((prop-2-yn-1-yloxy)carbonyl)amino)acetamido)propanoic acid (Formula V.2). In a 4mL vial with magnetic stirrer was placed 2-(((prop-2-yn-1-yloxy)carbonyl)amino)acetic acid (50mg, 1eq) and HBTU (121mg, 1eq) in 1mL of DMF. The mixture was stirred for 15min and then a solution of (2S)-3-amino-2-(((tert-butoxy)carbonyl)amino)propanoic acid (76mg, 1eq) in DMF (1mL) was added, followed by triethylamine (89uL). The vial was capped and stirred for 8h. The mixture was partitioned between 250mM citric acid and ethyl acetate. The organic layer was retained and the aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired product. Additional purification could be accomplished by silica gel chromatography.

The Boc-protected intermediate, (2S)-2-[[[(tert-butoxy)carbonyl]amino]-3-(2-[[[(prop-2-yn-1-yloxy)carbonyl]amino]acetamido)propanoic acid was placed in a 20mL vial and suspended in acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized to a paste. Additional purification by ion exchange resin (Dowex-50) afforded the desired amino acid. The product was confirmed by mass spectrometry.

Example 4. Preparation of Formula VII analogs

Preparation of (2S)-2-amino-6-(2-azidopentanamido)hexanoic acid (Formula VII.1). In a 20mL vial with magnetic stirrer was placed 2-azidopentanoic acid (100mg, 1eq) and HBTU (266mg, 1eq) in 3mL of DMF. The mixture was stirred for 15min and then a solution N-Boc-Lysine (200mg, 1eq) in DMF (3mL) was added, followed by triethylamine (195uL). The vial was capped and stirred for 8h. The mixture was partitioned between 250mM citric acid and ethyl acetate. The organic layer was retained and the aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired intermediate. Additional purification was conducted via silica gel chromatography.

Intermediate was placed in a 20mL vial and suspended in acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized. Additional purification by ion exchange resin (Dowex-50) afforded the desired amino acid. The product was confirmed by mass spectrometry.

Example 5. Preparation of Formula VIII analogs

Preparation of (2S)-2-amino-3-(4-azido-2-[[[(prop-2-en-1-yloxy)carbonyl]amino]butanamido)propanoic acid (Formula VIII.1). In a 4mL vial with magnetic stirrer was placed (2S)-3-(4-azido-2-[[[(prop-2-en-1-yloxy)carbonyl]amino]butanamido)-2-[[[(tert-butoxy)carbonyl]amino]propanoic acid (65mg, 1eq) and HBTU (108mg, 1eq) in 1mL of DMF. The mixture was stirred for 15min and then a solution of (2S)-3-amino-2-[[[(tert-butoxy)carbonyl]amino]propanoic acid (68mg, 1eq) in DMF (1mL) was

added, followed by triethylamine (79uL). The vial was capped and stirred for 4h. The mixture was poured onto 250mM citric acid and ethyl acetate, shaken and the organic layer retained. The aqueous layer was extracted two additional times with ethyl acetate. The organic layers were combined, dried with sodium sulfate, filtered and concentrated. Mass spectrometry indicated the formation of the desired product. Additional purification could be accomplished by silica gel chromatography.

The Boc-protected intermediate was placed in a 20mL vial and suspended in acetonitrile (2mL). To this was added a solution of hydrochloric acid in dioxane (4N, 2mL). The solution was stirred for 2h and then concentrated under reduced pressure. The mixture was lyophilized. Additional purification by ion exchange resin (Dowex-50) afforded the desired amino acid. The product was confirmed by mass spectrometry.

Example 6. Translational testing with a GFP assay.

An in vitro cell based assay was developed to assess the compatibility of the pylRS/tRNA pair and the pyrrolysine analogs of the present invention (nnAAs) by and the efficiency of nnAAs integration into a target protein. For this, HEK293 cells stably expressing pylRS (SEQ ID NO 4) were transiently transfected with plasmids for the expression of tRNA^{pyl} (SEQ ID NO 7) and a reporter construct encoding GFPY40 (containing amber codon in place of tyrosine at amino acid residue number 40 (where 1 is the initiator methionine)) (SEQ ID NOs 10 and 11) using standard transfection protocols. Transfected cells were incubated with nnAAs at 2mM for 2-3 days GFP production was analyzed qualitatively by visual inspection under the microscope. The GFP fluorescence was quantified by flow cytometry using an Accuri flow cytometer and the geometric mean of the fluorescent cells determined.

This cell based assay was used to determine whether the different nnAAs were suitable substrates for the pylRS and allowed its translation into a target protein. Cells expressing the PylRS/tRNA^{pyl} pair and containing a vector encoding the GFPY40 reporter gene were incubated in the presence of the nnAAs. nnAAs that are readily utilized by the PylRS/tRNA^{pyl} pair support the translation of the nnAA into the amber site of GFP and allow read-through of the gene producing full length GFP (fluorescent protein). The fluorescence intensity of the cells depends

on the efficiency of nnAA incorporation. Thus, nnAAs that are poorly utilized produce weakly fluorescent or non-fluorescing cells. Microscopic observation identified a number of nnAAs usable by the pylRS (Table 1, Positive GFP). Furthermore, the relative expression levels in each sample was compared to those generated by substrates known to be efficiently utilized by pylRS. Formula 1A.1 (MFI = 931,289), Formula 1A.2 (MFI=1,676,250) and Formula 1A.3 (MFI=2,250,000) (see Table 1) supported high levels of GFP expression with a geometric mean.

Analog Formulas IA.4, IA.5, III.1, III.2, III.3, III.12, III.13, V.1, V.2, V.3, VII.1 of the present invention were found by the inventors to be incorporated in the GFP reporter gene and yield green cells under the experimental conditions used. Among these, analogs Formula III.1 (MFI=1,989,750), Formula III.2 (MFI=1,847,250) and Formula III.3, (MFI=2,187,000) (see Table 2) supported high levels of GFP expression and represent analogues that are efficiently utilized by the pylRS/tRNA pair under the experimental conditions tested.

Table 1 : Formula I analog GFP results

Formula	IUPAC Name	Positive GFP	MFI
1A.1	(2S)-2-amino-6-{[(2-azidoethoxy)carbonyl]amino}hexanoic acid	Yes	931289
1A.2	(2S)-2-amino-6-{[(prop-2-yn-1-yloxy)carbonyl]amino}hexanoic acid	Yes	1676250
1A.3	(2S)-2-amino-6-{[(prop-2-en-1-yloxy)carbonyl]amino}hexanoic acid	Yes	2250000
1A.4	(2S)-2-amino-6-{[(2-azidoethyl)carbamoyl]amino}hexanoic acid	Yes	254250

Table 2 : Formula III analog GFP results

Formula	IUPAC Name	Positive GFP Assay	MFI
III.1	(2S)-2-amino-3-[(2-[(2-azidoethoxy)carbonyl]amino)ethyl)sulfanyl]propanoic acid	Yes	1989750
III.2	(2S)-2-amino-3-[(2-[(prop-2-yn-1-yloxy)carbonyl]amino)ethyl)sulfanyl] propanoic acid	Yes	1847250
III.3	(2S)-2-amino-3-[(2-[(prop-2-en-1-yloxy)carbonyl]amino)ethyl)sulfanyl] propanoic acid	Yes	2187000
III.13	(2S)-2-amino-3-[(3-[(2-azidoethyl)carbamoyl]propyl)sulfanyl]propanoic acid	Yes	
III.12	(2S)-2-amino-3-[(3-[(prop-2-en-1-yl)carbamoyl]propyl)sulfanyl]propanoic acid	Yes	

Table 3: Formula V analog GFP results

Formula	IUPAC Name	Positive GFP Assay	MFI
V.1	(2S)-2-amino-3-(2-[(2-azidoethoxy)carbonyl]amino)acetamido)propanoic acid	Yes	
V.3	(2S)-2-amino-3-(2-[(prop-2-en-1-yloxy)carbonyl]amino)acetamido)propanoic acid	Yes	

Table 4: Formula VII analogs GFP results

Formula	IUPAC Name	Positive GFP Assay	MFI
VII.1	(2S)-2-amino-6-(2-azidopentanamido)hexanoic acid	Yes	

Table 5: Formula VIII analogs GFP results

Formula	IUPAC Name	Positive GFP Assay	MFI
VIII.1	(2S)-2-amino-3-(4-azido-2-[[[(prop-2-en-1-yloxy)carbonyl]amino]butanamido]propanoic acid	Yes	

Example 7. Site specific insertion of pyrrolysine analogs in antibodies and conjugation**Construction and expression of anti-Her2 antibody containing pyrrolysine analog residues**

A full length anti-Her2 antibody containing two non natural amino acids (one in each heavy chain) (4D5-2AZ ab) (SEQ ID NO 15) was expressed in mammalian cells. CHO cells stably expressing Her2 antibody encoding amber sites at position K274 (SEQ ID NO 14) and expressing the pylRS/tRNA pair were exposed to nnAA of Formula III.1 and Formula 1A.1. The nnAAs of Formula III.1 or Formula 1A.1, containing an azide moiety, were incorporated at the heavy chain position K274 and purified by affinity chromatography using either protein A resin (GE Healthcare) or by IgSelect (GE Healthcare, 17096901). The purified material was then concentrated and subjected to a conjugation reaction.

An antibody directed to the extracellular domain of Her2/neu was generated by cloning the variable regions of both the heavy and light chains of the mouse antibody 4D5 into vectors containing genes encoding human IgG. The variable regions of 4D5 were generated by gene

synthesis using overlapping oligomers and cloned into the human IgG1 frameworks encoded by pFUSE-CHlg-hG1 (IgG1 heavy chain;gamma) (SEQ ID NOs 12 and 13) and pFUSE-CHLlg-hK (light chain; kappa; Invivogen)(SEQ ID NOs 16 and 17) to generate a mouse-human hybrid. An amber codon was introduced into the heavy chain (gamma) at position K274 by site directed mutagenesis (SEQ ID NOs 14 and 15). Clones containing the amber codon were identified by DNA sequencing. To generate an integrating construct the promoters and ORF for the heavy chain was amplified by PCR and cloned by restriction enzyme digestion and ligation into pOptivec (Life Technologies). The light chain and a single copy of the tRNA were joined by two step PCR method using overlapping oligomers and cloned into available sites into the pOptivec plasmid containing the heavy chain. The construct was then transfected into a CHO-DG44 cell line containing the pylRS/tRNA pair and transfectants selected by virtue of their growth in medium lacking hypoxanthine and thymidine. Selected cells were then cloned and high IgG expressors capable of efficient introduction of nnAAs into amber codons (expression of full length IgG) isolated. The selected cell line was utilized to generate IgG containing the nnAAs described above. The cells were grown to a density of $1-2 \times 10^6$ cells/mL in Excel DHFR-medium (Sigma-Aldrich) and nnAA of Formula III.1 or Formula IA.1 added to culture to a final concentration of 1 mM. Cells were incubated for 5 days and IgG purified from the growth medium. Supernatants were harvested and subjected to centrifugation to collect suspended cells and other debris. The supernatant was then filtered through a 0.22um filter to remove any particulate material prior to application to a chromatography column. The filtered supernatant was applied to a 1mL -5mL prepacked HiTrap protein A Sepharose at 1-5 mL/min flow rate using an AKTA chromatography system. The bound material and resin were washed with PBS to remove loosely bound proteins and the bound material eluted with 100mM glycine (pH 3.0) at a flow rate of 1 mL/min. Peak fractions containing the target protein were neutralized with 0.1 fraction volumes of 1M Tris-HCl (pH8.0). All constructs were dialyzed to PBS at 4°C for 16 hours into the final phosphate buffer. The antibody with Formula III.1 as nnAA incorporated into both of its heavy chains at position 274 was called "4D5-2AzAb-HC-274-(2S)-2-amino-3-[(2-[(2-azidoethoxy)carbonyl]amino)ethyl)sulfanyl] propanoic acid".

PEGylation of 4D5-2AzAb-HC-274-(2S)-2-amino-3-[(2-[(2-azidoethoxy)carbonyl]amino)ethyl)sulfanyl]propanoic acid

In a 200uL PCR tube was placed a solution of 4D5-2AzAb-HC274-(2S)-2-amino-3-[(2-[(2-azidoethoxy)carbonyl]amino)ethyl)sulfanyl]propanoic acid (Formula III.1) (100uL, 0.05 mg/mL) followed by a solution of 20KPEG cyclooctyne (33.3, 60mg/mL). The solution was mixed vigorously on a vortexer and allowed to incubate overnight. The mixture was diluted to 200uL and applied to Protein-A magnetic beads. The mixture was vortexed and allowed to rotate to mix the beads for 90min. The beads were immobilized and the run through material disposed. The beads were washed with PBS (2X) and then suspended in reducing gel buffer. Vortexed and then heated to 95C for 3min. The suspension was loaded directly onto an SDS-PAGE gel. Commassie staining of the SDS-PAGE gel indicated the selective PEGylation of the Heavy chain (Figure 1, lane 2).

REFERENCES

- Fekner, T., Li, X., & Chan, M. K. (2010). Pyrrolysine Analogs for Translational Incorporation into Proteins. *European Journal of Organic Chemistry*, 4171-4179.
- Kavran, J. M., Gundllapalli, S., O' Donoghue, P., Englert, M., Soll, D., & Steltz, T. A. (2007). Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proceedings National Academy of Sciences*, 104 (27), 11268-11273.
- Kobayashi, T., Yanagisawa, T., Sakamoto, K., & Yokoyama, S. (2009). Recognition of Non-α-amino Substrates by Pyrrolysyl-tRNA Synthetase. *J. Mol. Biol.* (1352-1360), 385.
- Liu, Chang C, and Peter G Schultz. "Adding New Chemistries to the Genetic Code." *Annual Review of Biochemistry*, 2010: 413-444.
- Chan, Michael K, Tomasz Fekner, Xin Li, Marianne Lee, and Jennifer J Ottesen. International Patent WO2011/044255A1. 2011.
- Nguyen, D. P., Lusic, H., Neumann, H. K., Deiters, A., & Chin, J. W. (2009). Genetic Encoding and Labeling of Aliphatic Azides and Alkynes in Recombinant Proteins via a Pyrrolysyl-tRNA

Synthetase/tRNA^{Acua} Pair and Click Chemistry. *Journal of the American Chemical Society* , 8720-8721.

Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., & Yokoyama, S. (2008). Crystallographic Studies on Multiple Conformational States of Active-site loops in Pyrrolysyl-tRNA synthetase. *J. Mol. Biol.* , 378, 634-652.

Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., & Yokoyama, S. (2008). Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode Ne-(o-Azidobenzyloxycarbonyl)lysine for Site Specific Protein Modification. *Chemistry and Biology* , 15, 1187-1197.

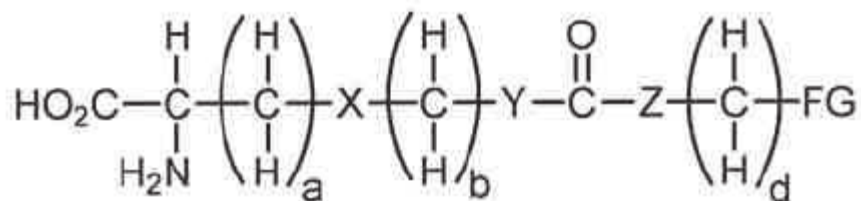
Yanagisawa, T., Sumida, T., Ishii, R., & Yokoyama, S. (2013). A novel crystal form of pyrrolysyl-tRNA synthetase reveals the pre- and post-aminoacyl-tRNA synthesis conformational states of the adenylate and aminoacyl moieties and an asparagine residue in the catalytic site. *Acta Crystallographica Section D* , D69, 5-15.

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

All patents and patent applications referred to herein are incorporated by reference in their entirety.

Claims

1. A pyrrolysine analogue of formula X:



wherein

X=O or S;

Y=CH₂, NH, O or S;

Z=CH₂, CH-NH₂, CH-OH, NH, O or S;

FG=azide, alkene, alkyne, ketone, ester, or cycloalkyne;

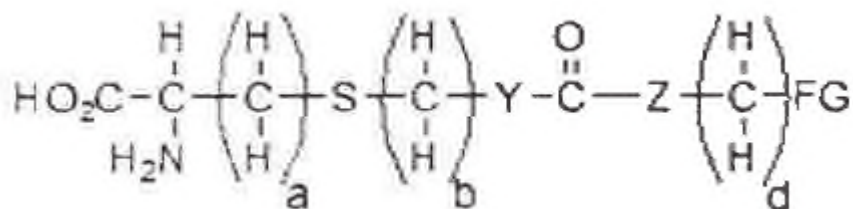
a=an integer 1-7;

b=an integer 1-7 save that when Z is NH, O or S then b is an integer 2-7;

provided that a+b is in the range 2-8;and

d=an integer 1-4.

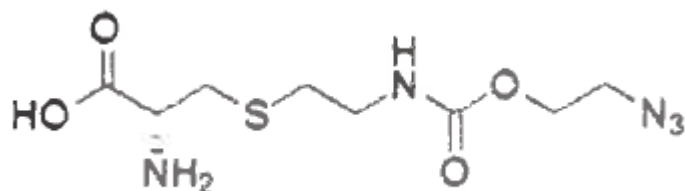
2. A pyrrolysine analogue according to claim 1 which is a pyrrolysine analogue of formula III:



wherein a, b, Y, Z, d and FG are as defined in claim 1.

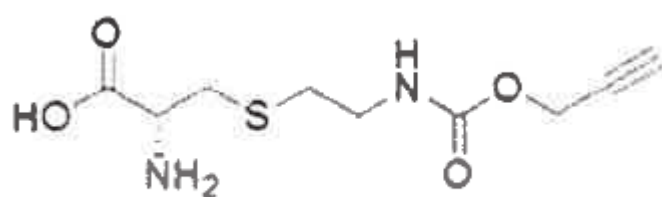
3. A pyrrolysine analogue according to claim 2 selected from

Formula III.1



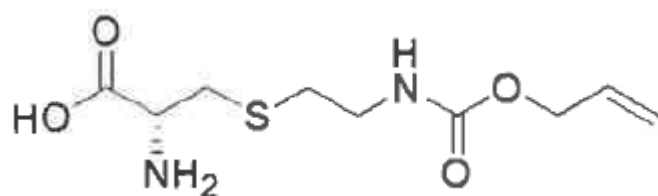
;

Formula III.2



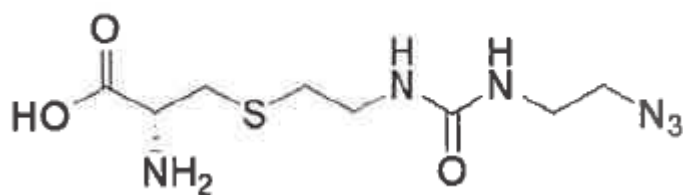
;

Formula III.3



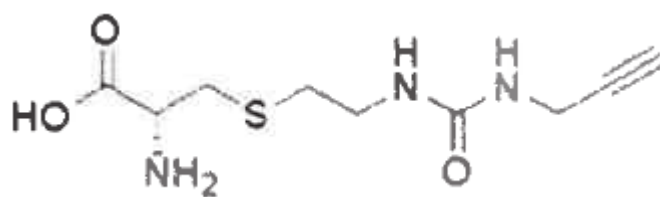
;

Formula III.4



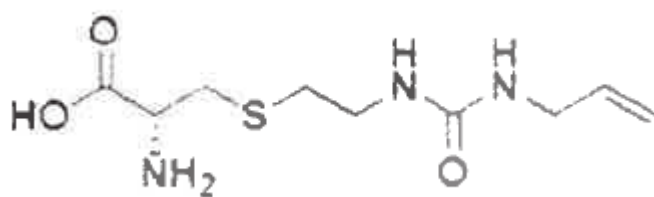
;

Formula III.5



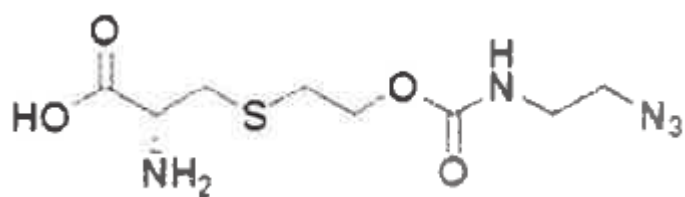
;

Formula III.6



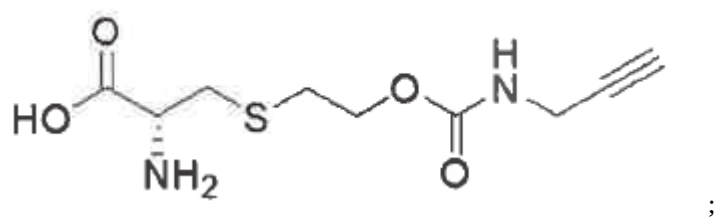
;

Formula III.7

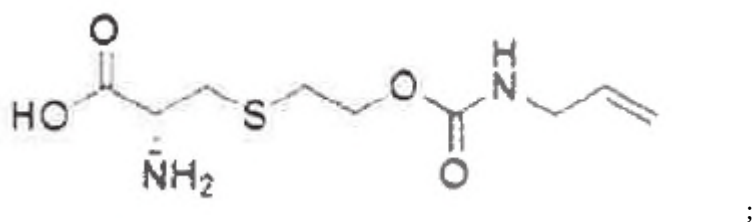


;

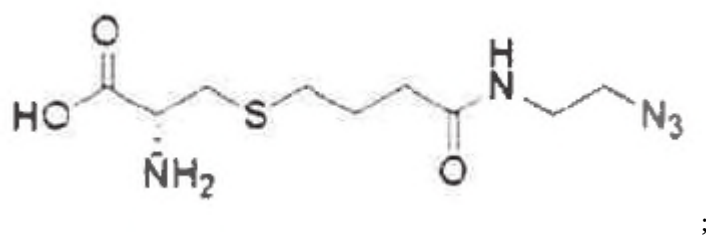
Formula III.8



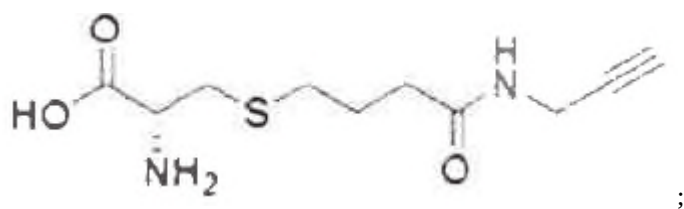
Formula III.9



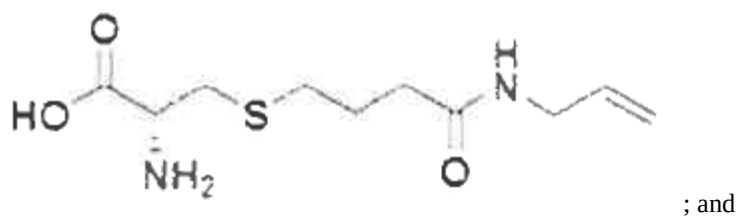
Formula III.10



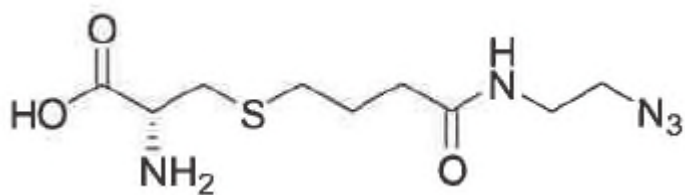
Formula III.11



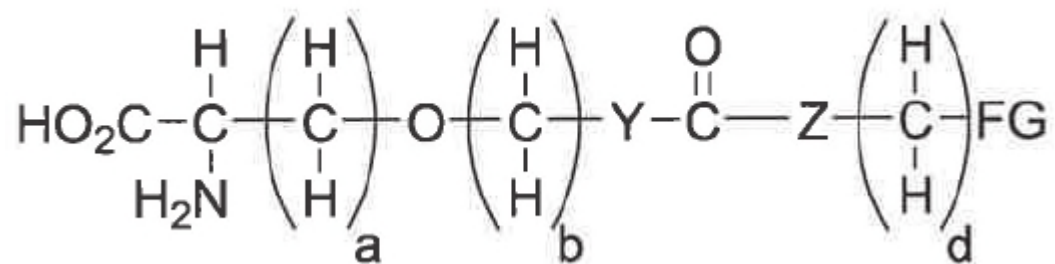
Formula III.12



Formula III.13



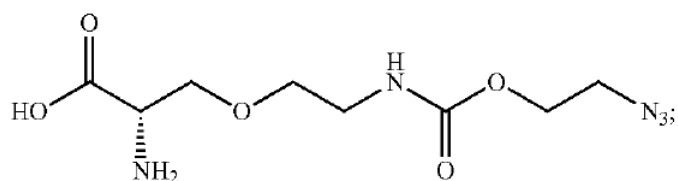
4. A pyrrolysine analogue according to claim 1 which is a pyrrolysine analogue of formula IV:



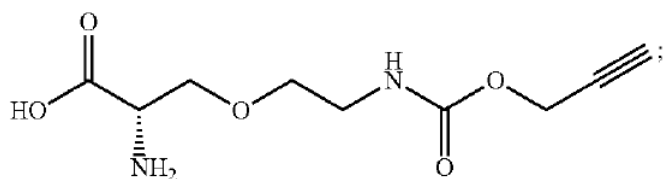
wherein a, b, Y, Z, d and FG are as defined in claim 1.

5. A pyrrolysine analogue according to claim 4 selected from

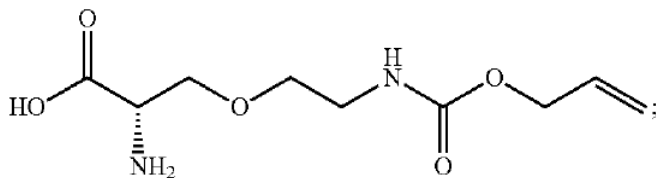
Formula IV.1



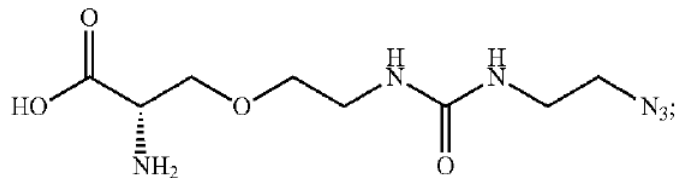
Formula IV.2



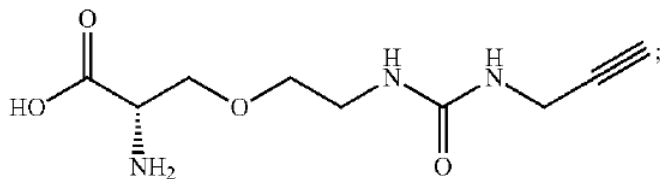
Formula IV.3



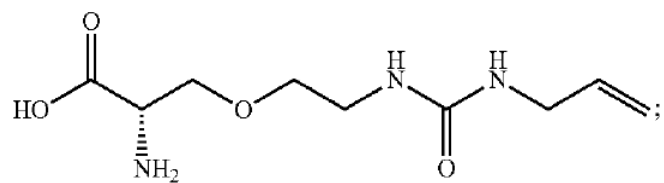
Formula IV.4



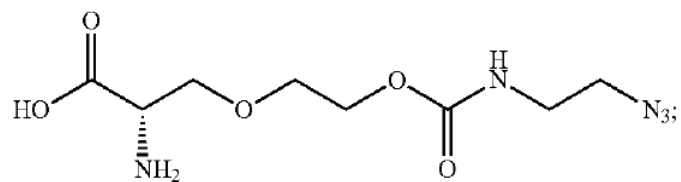
Formula IV.5



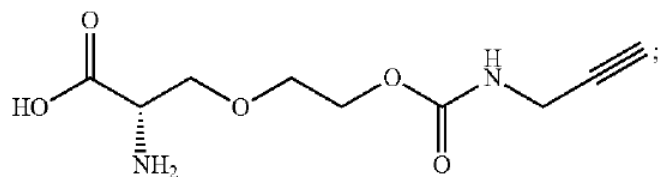
Formula IV.6



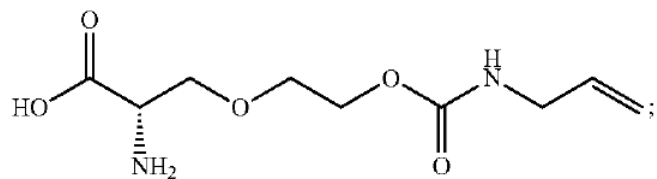
Formula IV.7



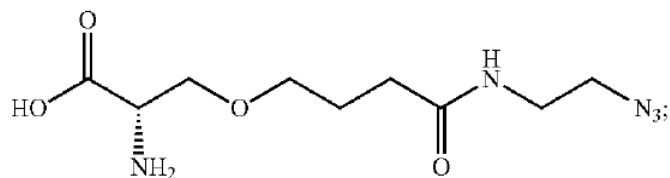
Formula IV.8



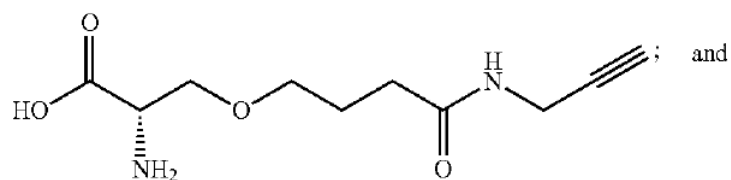
Formula IV.9



Formula IV.10

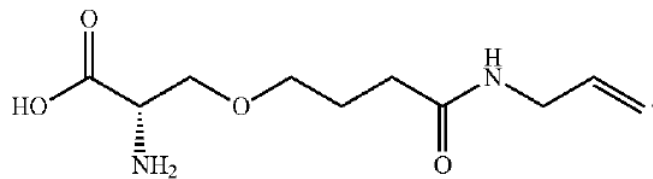


Formula IV.11

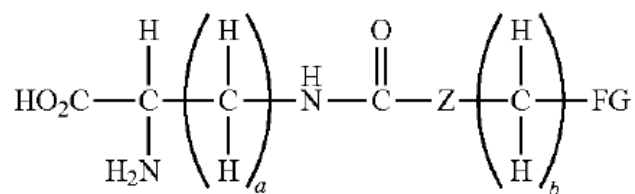


and

Formula IV.12



6. A pyrrolysine analogue of formula I:



Wherein

Z=bond, CH₂, CH—NH₂, CH—OH, NH, O, S or CH—NH₂;

a is an integer 3-7;

b is 0 or an integer 1-7; and

FG=azide, alkene, alkyne, ketone, ester, or cycloalkyne;

with the provisos that:

when a represents 4 and Z represents O, —(CH₂)_b-FG does not represent —CH₂-C≡CH, —CH₂CH₂-N₃, —CH₂-CH=CH₂ or —CH₂-Ph;

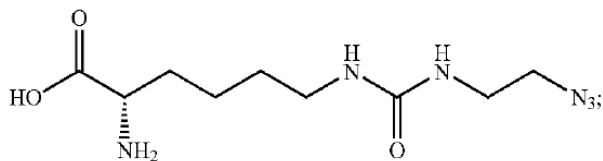
when a represents 4 and Z represents O, —(CH₂)_b-FG does not represent —CH₂CH₂CH₂-C(=O)CH₃ or —CH₂CH₂CH₂-CH=CH₂;

when a represents 4, Z represents a bond and b represents 0, —(CH₂)_b-FG does not represent —C(=O)Bn or —C(=O)Me; and

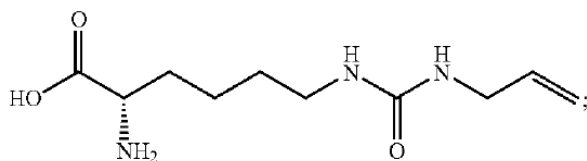
when a represents 4 and Z represents a bond, —(CH₂)_b-FG does not represent —CH₂CH₂-C≡CH.

7. A pyrrolysine analogue according to claim 6 selected from

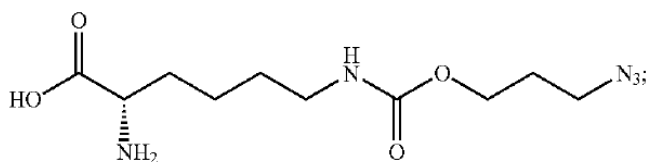
Formula IA.4,



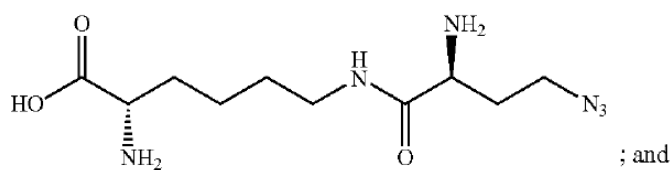
Formula IA.5,



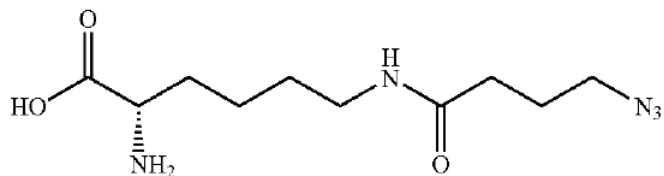
Formula IA.6



Formula IA.7,

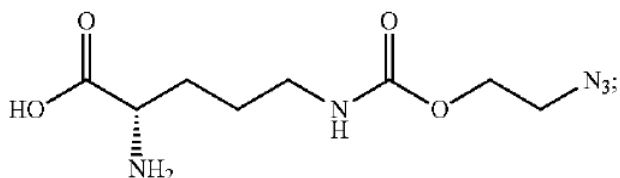


Formula IA.9,

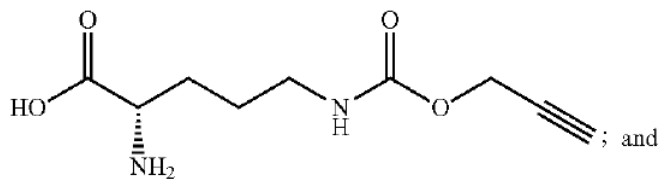


8. A pyrrolysine analogue according to claim 6 selected from

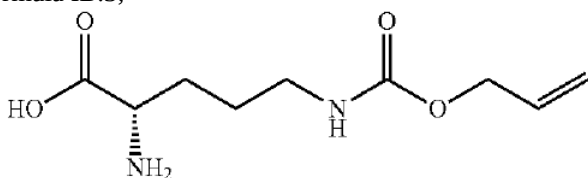
Formula IB.1,



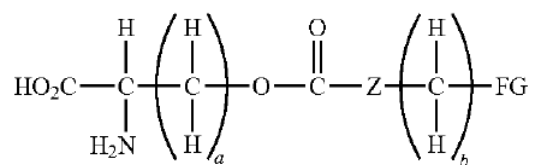
Formula IB.2,



Formula IB.3,



9. A pyrrolysine analogue of formula II:



wherein

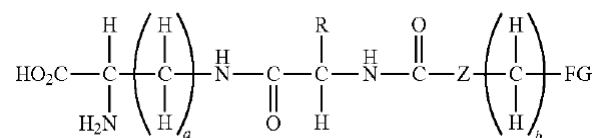
Z=CH₂, CH—NH₂, CH—OH, NH, O or S;

FG=azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a=an integer 3 or 5-7; and

b=an integer 1-4.

10. A pyrrolysine analogue of formula V:



wherein

R = the side chain of one of the 20 natural amino acids;

Z=CH₂, CH—NH₂, CH—OH, NH, O or S;

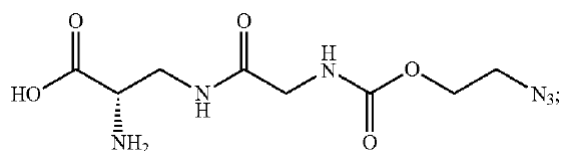
FG = azide, alkene, alkyne, ketone, ester, aryl or cycloalkyne;

a=1; and

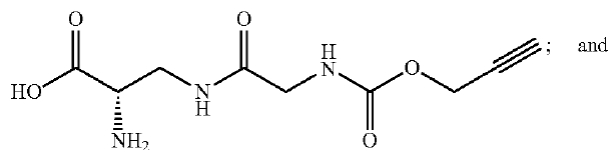
b=an integer 1 to 4.

11. A pyrrolysine analogue according to claim 10 selected from

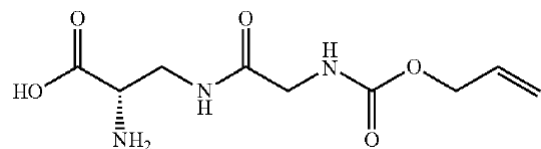
Formula V.1,



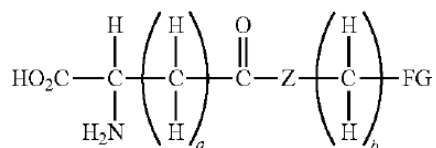
Formula V.2,



Formula V.3,



12. A pyrrolysine analogue of formula VI:



wherein

Z=CH₂, CH—NH₂, CH—OH, NH, O or S;

FG=azide, alkene, alkyne, ketone, ester, or cycloalkyne;

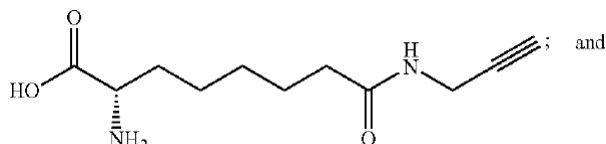
a=4 or 5; and

b=an integer 1 to 4.

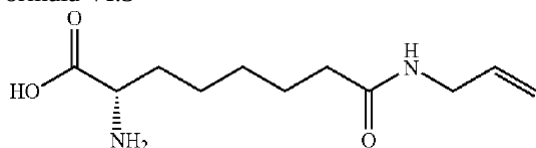
13. A pyrrolysine analogue according to claim 12 selected from Formula VI.1,



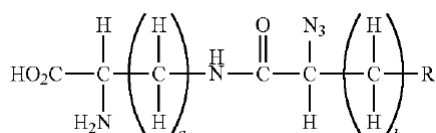
Formula VI.2,



Formula VI.3



14. A pyrrolysine analogue of formula VII:



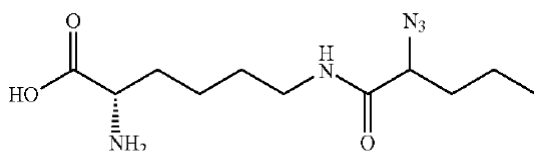
wherein R=alkyl, alkenyl, cycloalkyl, aryl, heteroaryl or heterocyclyl;

a=an integer 1 to 7; and

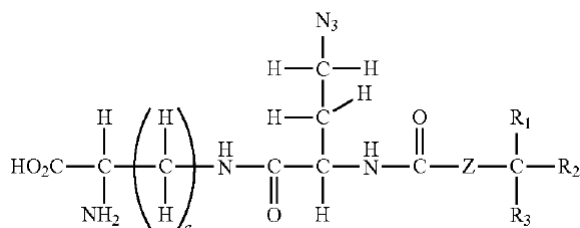
b=an integer 1 to 3.

15. A pyrrolysine analogue according to claim 14 which is:

Formula VII.1



16. A pyrrolysine analogue of formula VIII:



wherein

Z=CH₂, CH—NH₂, CH—OH, NH, O or S;

R₁ = H, alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl;

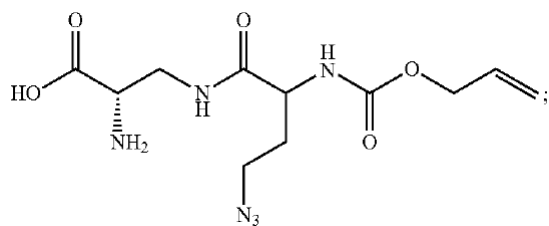
R₂ = alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl;

R₃ = H, alkyl, alkenyl, cycloalkyl, cycloalkenyl, aryl, heteroaryl or heterocyclyl; and

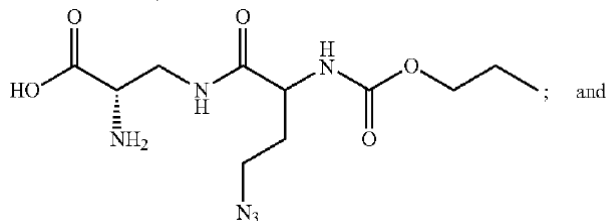
a=1.

17. A pyrrolysine analogue according to claim 16 selected from

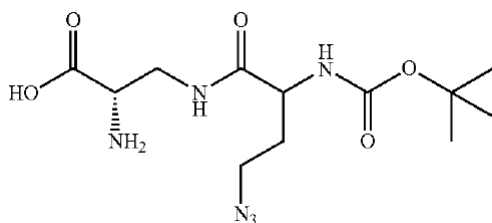
Formula VIII.1,



Formula VIII.2,

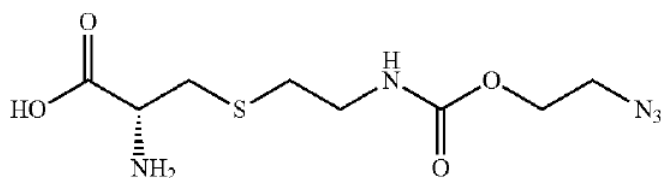


Formula VIII.3,



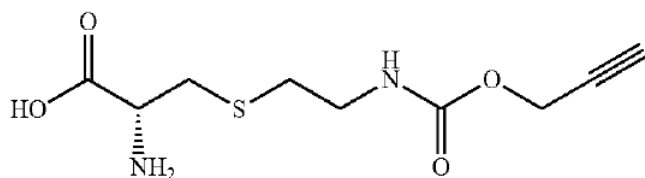
18. A pyrrolysine analogue according to claim 3 wherein the pyrrolysine analogue is

Formula III.1

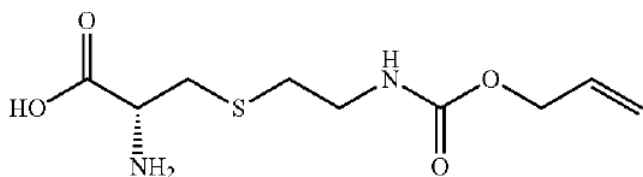


19. A pyrrolysine analogue according to claim 3 wherein the pyrrolysine analogue is

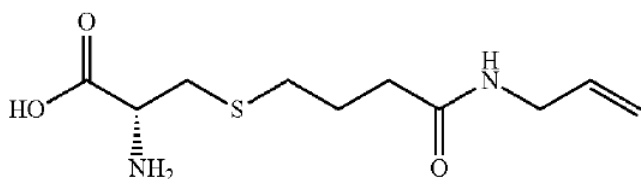
Formula III.2



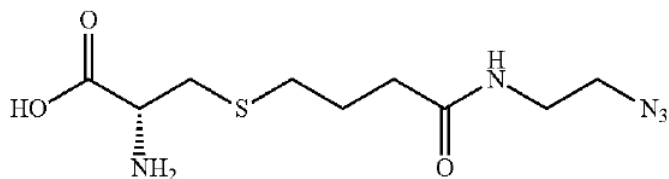
20. A pyrrolysine analogue according to claim 3 wherein the pyrrolysine analogue is
Formula III.3



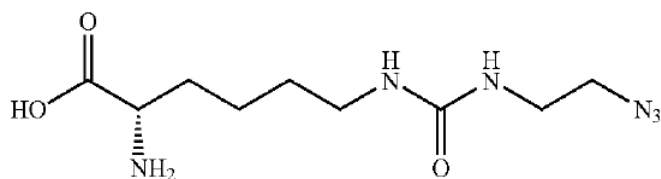
21. A pyrrolysine analogue according to claim 3 wherein the pyrrolysine analogue is
Formula III.12



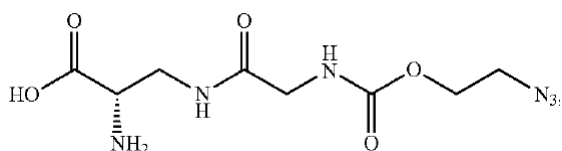
22. A pyrrolysine analogue according to claim 3 wherein the pyrrolysine analogue is
Formula III.13



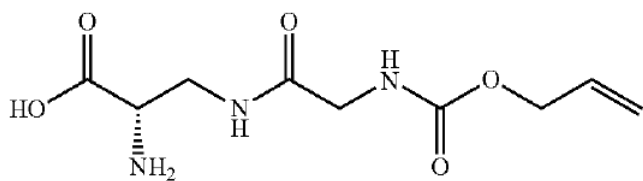
23. A pyrrolysine analogue according to claim 7 wherein the pyrrolysine analogue is
Formula IA.4



24. A pyrrolysine analogue according to claim 11 wherein the pyrrolysine analogue is
Formula V.1



25. A pyrrolysine analogue according to claim 11 wherein the pyrrolysine analogue is
Formula V.3



26. A pyrrolysine analogue according to claim 17 wherein the pyrrolysine analogue is
Formula VIII.1

