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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 October 2012 (26.10.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/142659 A1

(51) International Patent Classification:

C07K 1/107 (2006.01)	*A61K 38/04* (2006.01)
C07K 1/113 (2006.01)	*A61K 38/08* (2006.01)
C07K 1/13 (2006.01)	*A61K 38/10* (2006.01)
C07K 5/00 (2006.01)	*A61K 38/14* (2006.01)
C07K 5/062 (2006.01)	*A61K 51/08* (2006.01)
C07K 5/083 (2006.01)	*C12P 21/00* (2006.01)
C07K 5/103 (2006.01)	*C12P 21/02* (2006.01)
C07K 7/06 (2006.01)	*C07K 14/00* (2006.01)
C07K 7/08 (2006.01)	*C07K 14/54* (2006.01)
C07K 9/00 (2006.01)	*C07K 16/00* (2006.01)
C07K 17/06 (2006.01)	*A61K 38/05* (2006.01)
C07K 17/08 (2006.01)	*A61K 38/06* (2006.01)
C07K 17/10 (2006.01)	*A61K 38/07* (2006.01)
C07K 17/14 (2006.01)	*A61K 38/16* (2006.01)
C07K 19/00 (2006.01)	*A61K 38/20* (2006.01)
C12N 9/52 (2006.01)	*A61K 51/10* (2006.01)

(21) International Application Number:
PCT/AU2012/000400

(22) International Filing Date:
19 April 2012 (19.04.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/476,902 19 April 2011 (19.04.2011) US

(71) Applicants *(for all designated States except US)*: **BAKER IDI HEART AND DIABETES INSTITUTE HOLDINGS LIMITED** [AU/AU]; 75 Commercial Road, Melbourne, Victoria 3004 (AU). **THE UNIVERSITY OF MELBOURNE** [AU/AU]; Grattan Street, Parkville, Victoria 3052 (AU). **STARPHARMA PTY LIMITED** [AU/AU]; Baker Building, 75 Commercial Road, Melbourne, Victoria 3004 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HAGEMEYER, Christoph** [AU/AU]; 58 Tulip Grove, Cheltenham, Victoria 3102 (AU). **PETER, Karlheinz** [AU/AU]; 24 Clive Road, Hawthorn, Victoria 3122 (AU). **JOHNSTON, Angus Philip Rayner** [AU/AU]; 2/61 Lygon Street, Brunswick East, victoria 3057 (AU). **OWEN, David** [AU/AU]; 4 Leonard Court, Vermont South, victoria 3133 (AU).

(74) Agents: **MORRIS, Kathryn B.** et al.; c/o Davies Collison Cave, Level 10, 301 Coronation Drive, Milton, Queensland 4064 (AU).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

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

(54) Title: SITE-SELECTIVE MODIFICATION OF PROTEINS

(57) Abstract: The present invention relates to methods and reagents for use in site-selective modification of proteins and other entities with functionalized peptides using a chemoenzymatic sortase mediated reaction. The functionalized entities may be further reacted to conjugate two large entities together. Functionalized peptides including a glycine motif recognised by Sortase enzymes are also described.

WO 2012/142659 A1

SITE-SELECTIVE MODIFICATION OF PROTEINS

Field of the Invention

The present invention relates to reagents for use in site-selective modification of proteins. Methods of conjugating proteins to a second entity using a chemoenzymatic method with a sortase enzyme are also described.

Background of the Invention

The development of recombinant protein technology has provided large scale, commercial production of a plethora of recombinant proteins products for a variety of applications. However, many applications require the recombinant proteins to be further modified by conjugation to a modifying agent in order to optimize function. For example, protein-plasma half-life and/or immunogenicity may be improved by conjugation of the protein to biologically benign polymers such as polyethylene glycol (PEG), polyglutamate or polyvinylpyrolidone.

For further applications, the recombinant protein may represent a single function of a multi-functional product. There are many examples of such products in the art, including protein-drug conjugates for targeted drug delivery; protein-label conjugates for targeted imaging; protein-protein conjugates wherein each of the constituent recombinant proteins provides a specific function, for example, antibody-toxin conjugates have been developed for targeted anticancer products, or antibody-directed enzyme therapy. In yet other examples, there may be a requirement for a recombinant protein to be conjugated to living cells. Such products may have use in targeted stem cell therapies.

Some methods of modifying proteins are not site-selective and therefore modification occurs at one or more sites of the protein where there is suitable functionality for modification to occur. This suitable functionality may occur in or close to the site required for protein function and therefore modification results in reduced protein functional performance or inactivation of the protein. Furthermore, often modification using non-

site-selective methods results in a mixture of modified proteins where modification occurs at different sites in the protein or a combination of different sites in the protein.

The advantages of site-selective modification methods include modification at a site distal to the site responsible for protein function and therefore the modified protein retains functional performance and modification at a single controlled site results in the production of a single product which has advantages in quality, safety and registrability of the modified protein product.

Current coupling approaches for site-selective conjugation are based on the introduction of unique functional groups such as ketones and azides into proteins not present in natural amino acids. They can be incorporated into proteins by chemical modification of a protein's N-terminus or C-terminus, unnatural amino acid mutagenesis or by the use of enzymes that transfer prosthetic groups to proteins.

An example of site-selective protein modification has been described in US 5,824,784 in which the N terminus of a recombinant protein can be selectively modified with a modifying agent that contains an aldehyde moiety. In order for the conjugation to selectively modify the N-terminus of the protein, the reductive amination which achieves the conjugation must be carried out at a sufficiently acidic pH to ensure that only the lysine side chain epsilon amino groups are protonated.

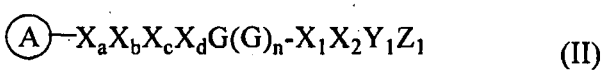
Another method involves the use of the *Staphylococcus aureus* enzyme Sortase A (WO 2010/087994). By inclusion of the LPXTG motif in a recombinant protein and by providing a label or second protein with the required complimentary polyglycine nucleophile, the Sortase enzyme will under some circumstances provide the conjugation of the two component moieties.

However, these described methods face practical limitations in terms of feasibility, scalability and efficacy.

Summary of the Invention

The present invention is predicated in part on the discovery that protein modification can be achieved in good yield by sortase conjugation of a LPXTG modified protein with a functionalized GGG peptide moiety, for example, a GGG peptide moiety comprising a spacer group or steric hinderance near the GGG moiety and a functional group for further attachment of the peptide to another moiety, comprising a complimentary functional group.

In a first aspect, there is provided a method of preparing a compound of formula (II):



$$(A)\text{—}X_aX_bX_cX_dG(G)_n\text{-}X_1X_2Y_1Z_1 \qquad (II)$$

wherein

A is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a photosensitizer, DNA, RNA, a carbohydrate, a contrast agent, a catalyst, a lipid or a virus particle,

X_a is selected from a leucine, serine, asparagine, valine, isoleucine, tyrosine and glutamine residue;

X_b is selected from a proline, alanine, serine and glycine residue;

X_c is any amino acid residue;

X_d is selected from a threonine, arginine, serine, alanine and glycine residue;

G is a glycine residue;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

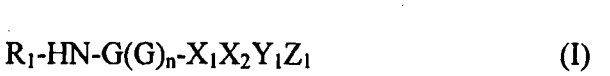
X_2 is absent or is an amino acid residue with a reactive functional group in its side chain;

Y_1 is absent or is a spacer group;

Z_1 is a reactive functional group; and

n is 0 or an integer of 1 to 4;

said method comprising contacting a compound of formula (I):



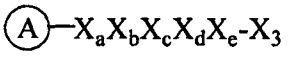
$$R_1\text{-}HN\text{-}G(G)_n\text{-}X_1X_2Y_1Z_1 \qquad (I)$$

wherein

R_1 is hydrogen or an amino protecting group; and

G, X_1 X_2, Y_1, Z_1 and n are defined above;

with a compound of formula (III):

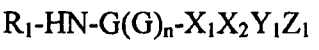


(III)

wherein A, X_a, X_b, X_c and X_d are defined above and X_e is selected from a glycine, alanine, asparagine, serine or valine residue and X_3 is absent, is a purification tag or when A is a protein or peptide, X_3 is optionally a bond to A;

in the presence of a sortase enzyme.

In another aspect of the invention, there is provided a compound having formula (IV):



wherein G is a glycine residue;

R_1 is hydrogen or an amino protecting group;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

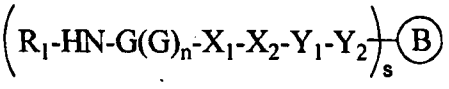
X_2 is an amino acid residue with a reactive functional group in its side chain;

Y_1 is a spacer group which is attached to X_2 through the reactive functional group in the side chain of X_2;

Z_1 is a reactive functional group; and

n is 0 or an integer of 1 to 4.

In yet another aspect of the invention there is provided a method of preparing a compound of formula (V):



(V)

wherein G is a glycine residue;

each R_1 is independently selected from hydrogen or an amino protecting group;

each X_1 is independently absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

each X_2 is independently absent or is an amino acid residue with a reactive functional group in its side chain;

each Y_1 is independently absent or is a spacer group;

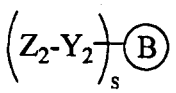
each Y_2 is independently absent or is a linker group;

B is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a polymer, a photosensitizer, DNA, RNA, a carbohydrate, a contrast agent, a catalyst, a lipid or a virus particle;

s is an integer from 1 to 100; and

n is 0 or an integer of 1 to 4.

said method comprising reacting a compound of formula (I) described above with a compound of formula (VI):



(VI)

wherein B, Y_2 and s are defined above and Z_2 is a reactive functional group complementary to Z_1 of formula (I).

In a further aspect of the invention, there is provided a kit comprising a compound of formula (I) and a sortase enzyme.

Brief Description of the Figures

Figure 1 shows labeling efficiency of site-selective modified proteins in which an scFv antibody with an LPETG tag prepared in Example 1 is conjugated using a sortase enzyme to a variety of GGG peptides comprising a reactive function Z_1, as described in Example 3. Efficiency of the protein-GGG peptide conjugation is shown as a percentage conversion.

The GGG peptides used in the sortase conjugation were selected from GGGWWSSK-PEG_4-N_3, GGGWWK-PEG_4-N_3, GGGK-PEG_4-N_3 and GGGWWGA-pG, where pG is a propargyl glycine. The Sortase A mediated conjugation of scFv-LPETG with GGG-eGFP is provided for comparison.

Figure 2 is a graphical representation of optimization of reaction conditions for Sortase A mediated conjugation of Lissamine Rhodamine B-LPETGGHHHHHH with GGGWWSSK-PEG_4-functionalized particles. Figure 2A shows optimization of concentration of Sortase A enzyme at 0.1 g/L. Figure 2B shows optimization of concentration of Lissamine Rhodamine B-LPETGGHHHHHH at 0.1 g/L. Figure 2C shows optimization of incubation time at 1 hour. Figure 2D shows optimized sortase mediated reaction with 0.1 g/L sortase, 0.1 g/L Lissamine Rhodamine B-LPETGGHHHHHH, 1 hour incubation time, where white bar is particles alone, single hatch bar is reaction with unfunctionalized particles and Lissamine Rhodamine B-LPETGGHHHHHH, double hatch bar is unfunctionalized particles, Lissamine Rhodamine B-LPETGGHHHHHH and Sortase A, grey bar is GGGWWSSK-PEG_4-coated particles and Lissamine Rhodamine B-LPETGGHHHHHH and the black bar is GGGWWSSK-PEG_4-coated particles, Lissamine Rhodamine B-LPETGGHHHHHH and Sortase A.

Figure 3 is a graphical representation of human *in vitro* thrombi incubated with sortase coupling between scFv-LPETG-GGG modified $(PVPON_{Alk})_5$ core shell particles with DL800 label as imaged by near infrared imaging. The thrombi that was untreated (black bar) had fluorescence set to one. Thrombi treated with uncoated $(PVPON_{Alk})_5$ core shell particles (single hatch bar), scFv(-) coated capsules (double hatch bar), scFv(+) coated capsules (grey bar) and blocked thrombi treated with scFv(+) coated capsules are shown. Strong specific binding was observed with the unblocked thrombi treated with scFv(+) coated particles (scFv(+) = anti-GPIIb/IIIa scFv, scFv(-) = mutated scFv).

Figure 4 is a graphical presentation of binding of non-targeted CHO cells (black bars) and scFv-coupled CHO cells (white bars) to activated platelets immobilized on coverslips. There is significant specific binding of scFv-coupled cells to a layer of activated platelets.

The numbers of bound scFv-cells were significantly higher than those of control non-targeted cells.

Description of the Invention

Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For the purposes of the present invention, the following terms are defined below.

Throughout this 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 or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The articles "*a*" and "*an*" are used herein to refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

As used herein, the term "*about*" refers to a quantity, level, value, dimension, size, or amount that varies by as much as 30%, 20%, or 10% to a reference quantity, level, value, dimension, size, or amount.

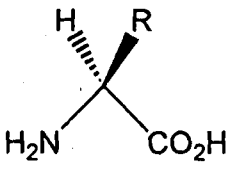
As used herein, the term "*peptide*" refers to two or more naturally occurring or non-naturally occurring amino acids joined by peptide bonds. Generally, peptides will range from about 2 to about 80 amino acid residues in length, usually from about 5 to about 60 amino acid residues in length and more usually from about 2 to about 40 amino acid residues in length. The peptide may also be a retro-inverso peptide. The peptide may contain α-amino acid residues, β-amino acid residues, D-amino acid residues, L-amino acid

residues, naturally occurring amino acid residues or non-naturally occurring amino acid residues.

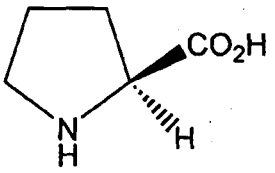
The amino acid may also be further substituted in the α-position or the β-position with a group selected from -C_1-C_6alkyl, -$(CH_2)_nCOR_1$, -$(CH_2)_nR_2$, -PO_3H, -$(CH_2)_n$heterocyclyl or -$(CH_2)_n$aryl where R_1 is -OH, -NH_2, -NHC_1-C_3alkyl, -OC_1-C_3alkyl or -C_1-C_3alkyl and R_2 is -OH, -SH, -SC_1-C_3alkyl, -OC_1-C_3alkyl, -C_3-C_{12}cycloalkyl, -NH_2, -NHC_1-C_3alkyl or -NHC(C=NH)NH_2 and where each alkyl, cycloalkyl, aryl or heterocyclyl group may be substituted with one or more groups selected from -OH, -NH_2, -NHC_1-C_3alkyl, -OC_1-C_3alkyl, -SH, -SC_1-C_3alkyl, -CO_2H, -CO_2C_1-C_3alkyl, -$CONH_2$ or -$CONHC_1$-C_3alkyl.

Amino acid structure and single and three letter abbreviations used throughout the specification are defined in Table 1, which lists the twenty naturally occurring amino acids which occur in proteins as L-isomers.

Table 1



(1)



(2)

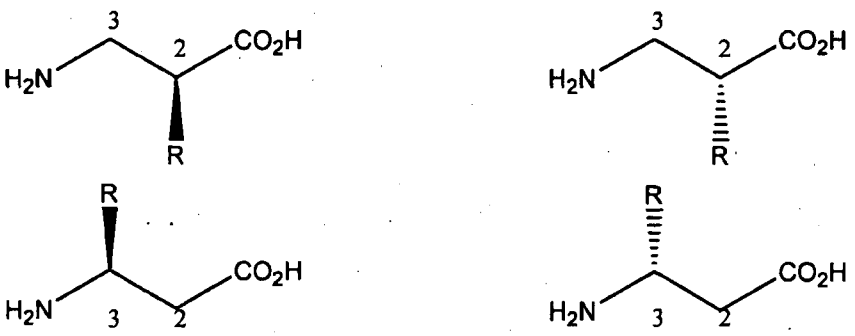
Amino Acid	Three-letter Abbreviation	One-letter symbol	Structure of side chain (R)
Alanine	Ala	A	-CH_3
Arginine	Arg	R	-$(CH_2)_3NHC(=N)NH_2$
Asparagine	Asn	N	-CH_2CONH_2
Aspartic acid	Asp	D	-CH_2CO_2H
Cysteine	Cys	C	-CH_2SH
Glutamine	Gln	Q	-$(CH_2)_2CONH_2$
Glutamic acid	Glu	E	-$(CH_2)_2CO_2H$

Glycine	Gly	G	-H
Histidine	His	H	-CH_2(4-imidazolyl)
Isoleucine	Ile	I	-$CH(CH_3)CH_2CH_3$
Leucine	Leu	L	-$CH_2CH(CH_3)_2$
Lysine	Lys	K	-$(CH_2)_4NH_2$
Methionine	Met	M	-$(CH_2)_2SCH_3$
Phenylalanine	Phe	F	-CH_2Ph
Proline	Pro	P	see formula (2) above for structure of amino acid
Serine	Ser	S	-CH_2OH
Threonine	Thr	T	-$CH(CH_3)OH$
Tryptophan	Trp	W	-CH_2(3-indolyl)
Tyrosine	Tyr	Y	-CH_2(4-hydroxyphenyl)
Valine	Val	V	-$CH(CH_3)_2$

The term "*α-amino acid*" as used herein, refers to a compound having an amino group and a carboxyl group in which the amino group and the carboxyl group are separated by a single carbon atom, the α-carbon atom. An α-amino acid includes naturally occurring and non-naturally occurring L-amino acids and their D-isomers and derivatives thereof such as salts or derivatives where functional groups are protected by suitable protecting groups. The α-amino acid may also be further substituted in the α-position with a group selected from -C_1-C_6alkyl, -$(CH_2)_nCOR_1$, -$(CH_2)_nR_2$, -PO_3H, -$(CH_2)_n$heterocyclyl or -$(CH_2)_n$aryl where R_1 is -OH, -NH_2, -NHC_1-C_3alkyl, -OC_1-C_3alkyl or -C_1-C_3alkyl and R_2 is -OH, -SH, -SC_1-C_3alkyl, -OC_1-C_3alkyl, -C_3-C_{12}cycloalkyl, -NH_2, -NHC_1-C_3alkyl or -NHC(C=NH)NH_2 and where each alkyl, cycloalkyl, aryl or heterocyclyl group may be substituted with one or more groups selected from -OH, -NH_2, -NHC_1-C_3alkyl, -OC_1-C_3alkyl, -SH, -SC_1-C_3alkyl, -CO_2H, -CO_2C_1-C_3alkyl, -$CONH_2$ or -$CONHC_1$-C_3alkyl.

As used herein, the term "*β-amino acid*" refers to an amino acid that differs from an α-amino acid in that there are two (2) carbon atoms separating the carboxyl terminus and the amino terminus. As such, β-amino acids with a specific side chain can exist as the *R* or

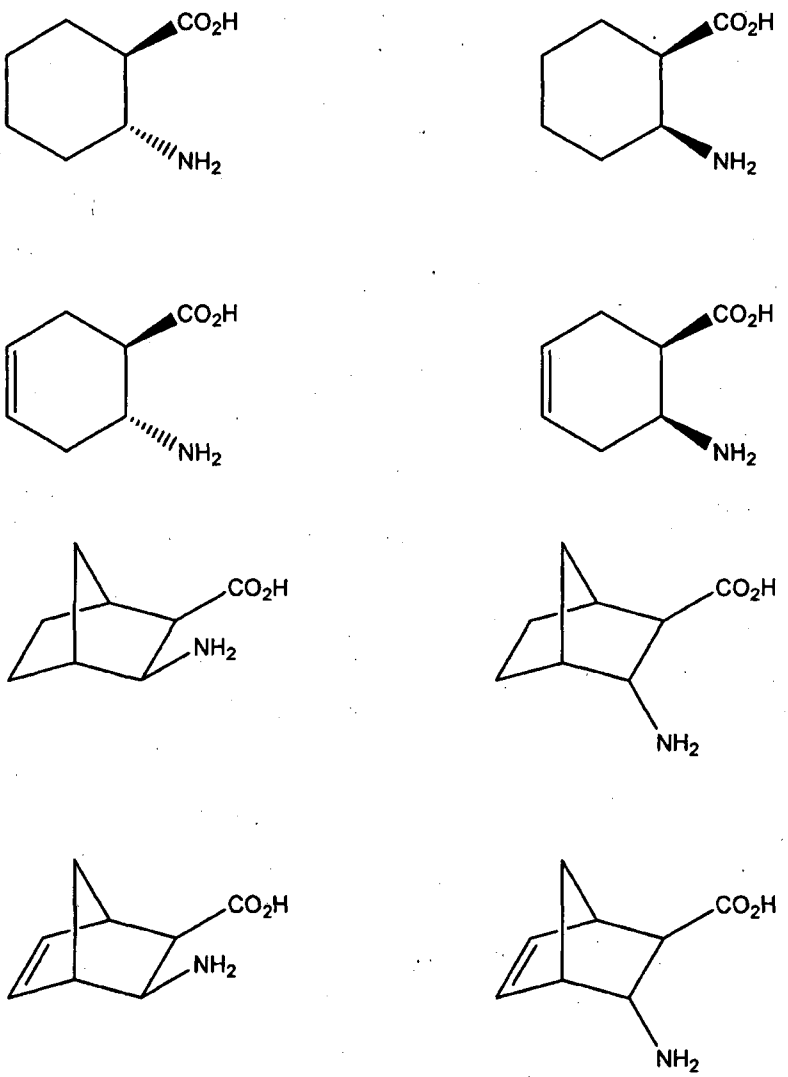
S enantiomers at either of the α (C2) carbon or the β (C3) carbon, resulting in a total of 4 possible isomers for any given side chain. The side chains may be the same as those of naturally occurring α-amino acids (see Table 1 above) or may be the side chains of non-naturally occurring amino acids (see Table 2 below).



Furthermore, the β-amino acids may have mono-, di-, tri- or tetra-substitution at the C2 and C3 carbon atoms. Mono-substitution may be at the C2 or C3 carbon atom. Di-substitution includes two substituents at the C2 carbon atom, two substituents at the C3 carbon atom or one substituent at each of the C2 and C3 carbon atoms. Tri-substitution includes two substituents at the C2 carbon atom and one substituent at the C3 carbon atom or two substituents at the C3 carbon atom and one substituent at the C2 carbon atom. Tetra-substitution provides for two substituents at the C2 carbon atom and two substituents at the C3 carbon atom. Suitable substituents include -C_1-C_6alkyl, -$(CH_2)_nCOR_1$, -$(CH_2)_nR_2$, -PO_3H, -$(CH_2)_n$heterocyclyl or -$(CH_2)_n$aryl where R_1 is -OH, -NH_2, -NHC_1-C_3alkyl, -OC_1-C_3alkyl or -C_1-C_3alkyl and R_2 is -OH, -SH, -SC_1-C_3alkyl, -OC_1-C_3alkyl, -C_3-C_{12}cycloalkyl, -NH_2, -NHC_1-C_3alkyl or -NHC(C=NH)NH_2 and where each alkyl, cycloalkyl, aryl or heterocyclyl group may be substituted with one or more groups selected from -OH, -NH_2, -NHC_1-C_3alkyl, -OC_1-C_3alkyl, -SH, -SC_1-C_3alkyl, -CO_2H, -CO_2C_1-C_3alkyl, -$CONH_2$ or -$CONHC_1$-C_3alkyl.

Other suitable β-amino acids include conformationally constrained β-amino acids. Cyclic β-amino acids are conformationally constrained and are generally not accessible to enzymatic degradation. Suitable cyclic β-amino acids include, but are not limited to, *cis*-

and *trans*-2-aminocyclopropyl carboxylic acids, 2-aminocyclobutyl and cyclobutenyl carboxylic acids, 2-aminocyclopentyl and cyclopentenyl carboxylic acids, 2-aminocyclohexyl and cyclohexenyl carboxylic acids and 2-amino-norbornane carboxylic acids and their derivatives, some of which are shown below:



Suitable derivatives of β-amino acids include salts and may have functional groups protected by suitable protecting groups.

The term "*non-naturally occurring amino acid*" as used herein, refers to amino acids having a side chain that does not occur in the naturally occurring L-α-amino acids listed in

Table 1. Examples of non-natural amino acids and derivatives include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, citrulline, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids. A list of unnatural amino acids that may be useful herein is shown in Table 2.

Table 2

Non-conventional amino acid	Code	Non-conventional amino acid	Code
α-aminobutyric acid	Abu	L-N-methylalanine	Nmala
α-amino-α-methylbutyrate	Mgabu	L-N-methylarginine	Nmarg
aminocyclopropane-carboxylate	Cpro	L-N-methylasparagine	Nmasn
		L-N-methylaspartic acid	Nmasp
aminoisobutyric acid	Aib	L-N-methylcysteine	Nmcys
aminonorbornyl-carboxylate	Norb	L-N-methylglutamine	Nmgln
		L-N-methylglutamic acid	Nmglu
cyclohexylalanine	Chexa	L-N-methylhistidine	Nmhis
cyclopentylalanine	Cpen	L-N-methylisoleucine	Nmile
D-alanine	Dal	L-N-methylleucine	Nmleu
D-arginine	Darg	L-N-methyllysine	Nmlys
D-aspartic acid	Dasp	L-N-methylmethionine	Nmmet
D-cysteine	Dcys	L-N-methylnorleucine	Nmnle
D-glutamine	Dgln	L-N-methylnorvaline	Nmnva
D-glutamic acid	Dglu	L-N-methylornithine	Nmorn
D-histidine	Dhis	L-N-methylphenylalanine	Nmphe
D-isoleucine	Dile	L-N-methylproline	Nmpro
D-leucine	Dleu	L-N-methylserine	Nmser
D-lysine	Dlys	L-N-methylthreonine	Nmthr
D-methionine	Dmet	L-N-methyltryptophan	Nmtrp
D-ornithine	Dorn	L-N-methyltyrosine	Nmtyr

	D-phenylalanine	Dphe	L-N-methylvaline	Nmval
	D-proline	Dpro	L-N-methylethylglycine	Nmetg
	D-serine	Dser	L-N-methyl-t-butylglycine	Nmtbug
	D-threonine	Dthr	L-norleucine	Nle
5	D-tryptophan	Dtrp	L-norvaline	Nva
	D-tyrosine	Dtyr	α-methyl-aminoisobutyrate	Maib
	D-valine	Dval	α-methyl-aminobutyrate	Mgabu
	D-α-methylalanine	Dmala	α-methylcyclohexylalanine	Mchexa
	D-α-methylarginine	Dmarg	α-methylcylcopentylalanine	Mcpen
10	D-α-methylasparagine	Dmasn	α-methyl-α-napthylalanine	Manap
	D-α-methylaspartate	Dmasp	α-methylpenicillamine	Mpen
	D-α-methylcysteine	Dmcys	N-(4-aminobutyl)glycine	Nglu
	D-α-methylglutamine	Dmgln	N-(2-aminoethyl)glycine	Naeg
	D-α-methylhistidine	Dmhis	N-(3-aminopropyl)glycine	Norn
15	D-α-methylisoleucine	Dmile	N-amino-α-methylbutyrate	Nmaabu
	D-α-methylleucine	Dmleu	α-napthylalanine	Anap
	D-α-methyllysine	Dmlys	N-benzylglycine	Nphe
	D-α-methylmethionine	Dmmet	N-(2-carbamylethyl)glycine	Ngln
	D-α-methylornithine	Dmorn	N-(carbamylmethyl)glycine	Nasn
20	D-α-methylphenylalanine	Dmphe	N-(2-carboxyethyl)glycine	Nglu
	D-α-methylproline	Dmpro	N-(carboxymethyl)glycine	Nasp
	D-α-methylserine	Dmser	N-cyclobutylglycine	Ncbut
	D-α-methylthreonine	Dmthr	N-cycloheptylglycine	Nchep
	D-α-methyltryptophan	Dmtrp	N-cyclohexylglycine	Nchex
25	D-α-methyltyrosine	Dmty	N-cyclodecylglycine	Ncdec
	D-α-methylvaline	Dmval	N-cylcododecylglycine	Ncdod
	D-N-methylalanine	Dnmala	N-cyclooctylglycine	Ncoct
	D-N-methylarginine	Dnmarg	N-cyclopropylglycine	Ncpro
	D-N-methylasparagine	Dnmasn	N-cycloundecylglycine	Ncund
30	D-N-methylaspartate	Dnmasp	N-(2,2-diphenylethyl)glycine	Nbhm
	D-N-methylcysteine	Dnmcys	N-(3,3-diphenylpropyl)glycine	Nbhe
	D-N-methylglutamine	Dnmgln	N-(3-guanidinopropyl)glycine	Narg
	D-N-methylglutamate	Dnmglu	N-(1-hydroxyethyl)glycine	Nthr

	D-N-methylhistidine	Dnmhis	N-(hydroxyethyl))glycine	Nser
	D-N-methylisoleucine	Dnmile	N-(imidazolylethyl))glycine	Nhis
	D-N-methylleucine	Dnmleu	N-(3-indolylyethyl)glycine	Nhtrp
	D-N-methyllysine	Dnmlys	N-methyl-γ-aminobutyrate	Nmgabu
5	N-methylcyclohexylalanine	Nmchexa	D-N-methylmethionine	Dnmmet
	D-N-methylornithine	Dnmorn	N-methylcyclopentylalanine	Nmcpen
	N-methylglycine	Nala	D-N-methylphenylalanine	Dnmphe
	N-methylaminoisobutyrate	Nmaib	D-N-methylproline	Dnmpro
	N-(1-methylpropyl)glycine	Nile	D-N-methylserine	Dnmser
10	N-(2-methylpropyl)glycine	Nleu	D-N-methylthreonine	Dnmthr
	D-N-methyltryptophan	Dnmtrp	N-(1-methylethyl)glycine	Nval
	D-N-methyltyrosine	Dnmtyr	N-methylnapthylalanine	Nmanap
	D-N-methylvaline	Dnmval	N-methylpenicillamine	Nmpen
	γ-aminobutyric acid	Gabu	N-(*p*-hydroxyphenyl)glycine	Nhtyr
15	L-*t*-butylglycine	Tbug	N-(thiomethyl)glycine	Ncys
	L-ethylglycine	Etg	penicillamine	Pen
	L-homophenylalanine	Hphe	L-α-methylalanine	Mala
	L-α-methylarginine	Marg	L-α-methylasparagine	Masn
	L-α-methylaspartate	Masp	L-α-methyl-*t*-butylglycine	Mtbug
20	L-α-methylcysteine	Mcys	L-methylethylglycine	Metg
	L-α-methylglutamine	Mgln	L-α-methylglutamate	Mglu
	L-α-methylhistidine	Mhis	L-α-methylhomophenylalanine	Mhphe
	L-α-methylisoleucine	Mile	N-(2-methylthioethyl)glycine	Nmet
	L-α-methylleucine	Mleu	L-α-methyllysine	Mlys
25	L-α-methylmethionine	Mmet	L-α-methylnorleucine	Mnle
	L-α-methylnorvaline	Mnva	L-α-methylornithine	Morn
	L-α-methylphenylalanine	Mphe	L-α-methylproline	Mpro
	L-α-methylserine	Mser	L-α-methylthreonine	Mthr
	L-α-methyltryptophan	Mtrp	L-α-methyltyrosine	Mtyr
30	L-α-methylvaline	Mval	L-N-methylhomophenylalanine	Nmhphe
	N-(N-(2,2-diphenylethyl) carbamylmethyl)glycine	Nnbhm	N-(N-(3,3-diphenylpropyl) carbamylmethyl)glycine	Nnbhe

1-carboxy-1-(2,2-diphenyl-ethylamino)cyclopropane	Nmbc	propargyl glycine	pG

As used herein, the term "*amino acid residue with a bulky side chain*" refers to amino acid residues that have side chains with branched or cyclic substituents. Examples of amino acid residues with a bulky side chain include tryptophan, tyrosine, phenylalanine, homophenylalanine, leucine, isoleucine, histidine, α-methyltryptophan, α-methyltyrosine, α-methylphenylalanine, α-methylleucine, α-methylisoleucine, α-methylhistidine, cyclopentylalanine, cyclohexylalanine and naphthylalanine.

As used herein, the term "*amino acid residue with a reactive functional group in its side chain*" refers to an amino acid residue that has a side chain that bears a functional group that is able to react with another functional group. Suitable functional groups include amino groups, carboxylic acid groups, thiol groups and hydroxyl groups. Examples of amino acid residues with a reactive functional group in their side chains include lysine, ornithine, glutamic acid, aspartic acid, serine, threonine, cysteine, α-methyllysine, α-methylornithine, α-methylglutamic acid, α-methylaspartic acid, α-methylserine, α-methylcysteine and α-methylthreonine.

The term "*pharmacokinetic modifying polymer*" as used herein refers to a polymer that alters pharmacokinetic properties of the molecule in which it is incorporated. For example, a polymer that increases solubility of the compound in which it is incorporated or a polymer that increases the plasma half life of the molecule in which it is incorporated or a polymer which assists in targeting the molecule in which it is incorporated to a particular site in the body. Suitable pharmacokinetic modifying polymers include, but are not limited to, polyethylene glycol (PEG), polypropylene glycol (PPG), polyethyleneoxide (PEO), poly(alkyloxazolines) such as poly(ethyloxazoline) (PEOX), polyvinylpyrrolidone (PVPON), polylysine, polyglutamic acid, polymethacrylic acid and polypropacrylic acid.

As used herein, the term "*sortase recognition tag*" refers to the amino acid sequence recognized by the sortase enzyme that enables conjugation with the $G(G)_n$ motif. For example, the sortase recognition tag for Sortase A from *S. aureus* is LPXTG (SEQ ID NO:1).

The term "*alkyl*" as used herein refers to straight chain or branched saturated hydrocarbon groups. Suitable alkyl groups include, but are not limited to methyl, ethyl, propyl, isopropyl, butyl, isobutyl, pentyl, hexyl, heptyl, octyl, nonyl, decyl, dodecyl, tetradecyl, hexadecyl and octadecyl. The term alkyl may be prefixed by a specified number of carbon atoms to indicate the number of carbon atoms or a range of numbers of carbon atoms that may be present in the alkyl group. For example, C_1-C_3alkyl refers to methyl, ethyl, propyl and isopropyl.

The term "*alkenyl*" as used herein refers to straight chain, branched or cyclic hydrocarbon groups having at least one carbon-carbon double bond. The term alkenyl may be prefixed by a specified number of carbon atoms to indicate the number of carbon atoms or a range of numbers of carbon atoms that may be present in the alkenyl group. For example, C_2-C_3alkenyl refers to ethenyl, propenyl and isopropenyl. Suitable alkenyl groups include, but are not limited to ethenyl, propenyl, isopropenyl, butenyl, isobutenyl, pentenyl, cyclopentenyl, hexenyl, cyclohexenyl, heptenyl, octenyl, nonenyl, decenyl, dodecenyl, tetradecenyl, hexadecenyl and octadecenyl.

The term "*alkynyl*" as used herein refers to straight chain, branched or cyclic hydrocarbon groups having at least one carbon-carbon triple bond. The term alkynyl may be prefixed by a specified number of carbon atoms to indicate the number of carbon atoms or a range of numbers of carbon atoms that may be present in the alkynyl group. For example, C_2-C_3alkynyl refers to ethynyl and propynyl. Suitable alkynyl groups include, but are not limited to ethynyl, propynyl, butynyl, pentynyl, hexynyl, heptynyl, cycloheptynyl, octynyl, cyclooctynyl, nonynyl, decynyl, dodecynyl, tetradecynyl, hexadecynyl and octadecynyl.

As used herein, the term "*alkylene*" refers to a linear chain of divalent -(CH_2)- groups that link two groups together. Typically an alkylene group has 1 to 20 -(CH_2)- groups, especially 1 to 15, 1 to 10, 2 to 9, 2 to 8, 2 to 7, 2 to 6, 2 to 5, 2 to 4 or 2 to 3 -(CH_2)- groups.

As used herein, the term "*alkenylene*" refers to an alkylene in which two -(CH_2)- groups are replaced by at least one -CH=CH- group and which links two groups together. Typically an alkenylene group has 2-20 carbon atoms and may comprise one or more -CH=CH- groups and optionally further -CH_2- groups. The alkenylene group may have 2-15, 2-10, 2-9, 2-8, 2-7, 2-6, 2-5, 2-4 or 2-3 carbon atoms including at least one -CH=CH- group.

As used herein, the term "*alkynylene*" refers to an alkylene in which two -(CH_2)- groups are replaced by at least one -C≡C- group and which links two groups together. Typically an alkynylene group has 2-20 carbon atoms and may comprise one or more -C≡C- groups and optionally further -CH_2- groups. The alkynylene group may have 2-15, 2-10, 2-9, 2-8, 2-7, 2-6, 2-5, 2-4 or 2-3 carbon atoms including at least one -C≡C- group.

The term "*aryl*" as used herein, refers to a C_6-C_{12} aromatic cyclic hydrocarbon groups in which at least one ring is aromatic such as phenyl, naphthyl, biphenyl and tetrahydronaphthyl.

The term "*cycloalkyl*" as used herein, refers to cyclic hydrocarbon groups. Suitable cycloalkyl groups include, but are not limited to, cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, cyclooctyl, cyclononyl, cyclodecyl, cycloundecyl and cyclododecyl.

The term "*heterocyclyl*" as used herein refers to 5 or 6 membered saturated, partially unsaturated or aromatic cyclic hydrocarbon groups in which at least one carbon atom has been replaced by N, O or S. Optionally, the heterocyclyl group may be fused to a phenyl ring. Suitable heterocyclyl groups include, but are not limited to pyrrolidinyl, piperidinyl,

pyrrolyl, thiophenyl, furanyl, oxazolyl, imidazolyl, thiazolyl, isoxazolyl, pyrazolyl, isothiazolyl, pyridinyl, quinolinyl, isoquinolinyl, indolyl, benzofuranyl, benzothiophenyl, oxadiazolyl, tetrazolyl, triazolyl and pyrimidinyl.

The term "heteroaryl" as used herein, represents a stable monocyclic, bicyclic or tricyclic ring of up to 7 atoms in each ring, wherein at least one ring is aromatic and at least one ring contains from 1 to 4 heteroatoms selected from the group consisting of O, N and S. When more than one ring is present the rings may be fused. The heteroaryl group may also include a carbonyl group attached to an unsaturated carbon in the ring system. Examples of suitable heteroaryl groups include pyrrolyl, furanyl, thienyl, pyrazolyl, imidazolyl, 1,2,4-triazolyl, 1,2,3-triazolyl, isoxazolyl, oxazolyl, thiazolyl, isothiazolyl, oxadiazolyl, oxatriazolyl, pyridinyl, pyridazinyl, pyrimidinyl, pyrazinyl, triazinyl, azepinyl, oxepinyl, thiepinyl, diazepinyl, coumaranyl, benzofuranyl, isobenzofuranyl, benzothienyl, indolyl, indolinyl, isoindolyl, benzimidazolyl, benzisoxazolyl, benzoxazolyl, benzisoxazolyl, benzothiazolyl, 1-benzopyranyl, 2-benzopyranyl, benzopyran-2-on-yl, benzopyran-1-on-yl, quinolinyl, tetrahydroquinolinyl, isoquinolinyl, quinazolinyl, cinnolinyl, quinoxalinyl, tetrahydroquinoxalinyl, naphthyridinyl, acridinyl, carbazolyl, xanthenyl, phenazinyl, phenothiazinyl, phenoxazinyl, 1,4-benzodiazepin-2-on-yl, 1,5-benzodiazepin-2-on-yl, 1,4-benzodiazepin-2,5-dion-yl, pyrrolo[2,1-*c*][1,4]benzodiazepine-5,11-dion-yl, 1,4-benzothiazepin-5-on-yl, 5,11-dihydro-benzo[*e*]pyrido[3,2-*b*][1,4]-diazepin-6-on-yl, chromonyl, pyranocoumarinyl, 3,4-dihydroquinoxalin-2-on-yl, quinazolinonyl, quinazolindionyl, imidazoquinoxalinyl, 2,3-dihydrospiro[indene-1,4'-piperidine] and spiro[indoline-3,4'-piperidine].

As used herein, a "*hydroxylamine*" is a moiety that includes a –O-NH- group. Either or both of the oxygen and amine may be substituted.

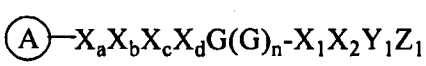
The term "*hydrazine*" is a moiety that includes a –N-N- group, where either or both nitrogen atoms may be substituted.

The term "*oxo*" as used herein refers to a substituent that is =O. An oxo substituent may be substituted on a saturated carbon atom in a ring or alkyl chain to form a carbonyl group.

As used herein, the term "*thioester*" refers to a moiety that includes a -C(=O)S-R group, where the R group is a substituent such as an alkyl group, an alkenyl group, cycloalkyl group or aryl group. The carbon atom may also be substituted.

The term "*complementary reactive functional group*" as used herein refers to a reactive functional group that is included in a compound or molecule and which may be paired up with another reactive functional group to undergo a particular reaction following methodologies that have been well described in the art. For example, when the reactive functional group is an alkyne group, the complementary functional group is an azide and when the reactive functional group is an azide, the complementary functional group is an alkyne, together these groups undergo a 1,3-dipolar cycloaddition reaction. Other examples include, but are not limited to, a reactive functional group being a hydroxylamine or hydrazine and the complementary group being an aldehyde or ketone, or *vice versa*, where the groups react together to form oximes or hydrazones; the reactive functional group may be an amino group and the complementary group a carboxylic acid, or *vice versa*, where the two groups undergo amide formation; the reactive functional group may be a thiol group and the complementary group may be an alkene, or *vice versa*, and the two groups undergo a thiol-ene reaction or the reactive functional group may be a thioester and the complementary group a cysteine or equivalent thereof, or *vice versa*, and the two groups undergo native chemical ligation.

In one aspect, the present invention provides a method of preparing a compound of formula (II):



(II)

wherein

A is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a photosensitizer, DNA, RNA, a carbohydrate, a contrast agent, a catalyst, a lipid or a virus particle,

X_a is selected from a leucine, serine, asparagine, valine, isoleucine, tyrosine and glutamine residue;

X_b is selected from a proline, alanine, serine and glycine residue;

X_c is any amino acid residue;

X_d is selected from a threonine, arginine, serine, alanine and glycine residue;

G is a glycine residue;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

X_2 is absent or is an amino acid residue with a reactive functional group in its side chain;

Y_1 is absent or is a spacer group;

Z_1 is a reactive functional group; and

n is 0 or an integer of 1 to 4;

said method comprising contacting a compound of formula (I):

$$R_1\text{-HN-G(G)}_n\text{-}X_1X_2Y_1Z_1 \qquad \text{(I)}$$

wherein

R_1 is hydrogen or an amino protecting group; and

G, X_1 X_2, Y_1, Z_1 and n are defined above;

with a compound of formula (III):

$$\text{(A)}\text{—}X_aX_bX_cX_dX_e\text{-}X_3 \qquad \text{(III)}$$

wherein A, X_a, X_b, X_c and X_d are defined above and X_e is selected from a glycine, alanine, asparagine, serine or valine residue and X_3 is absent, is a purification tag or when A is a protein or peptide, X_3 is optionally a bond to A;

in the presence of a sortase enzyme.

In this aspect of the invention, the compound of formula (I) described above is contacted with a group A having a sortase recognition tag in the presence of a sortase enzyme. The sortase recognition tag is recognized by the sortase enzyme which then ligates the tag with the $G(G)_n$ portion of formula (I).

The sortase enzyme useful in the method of the invention is any sortase that recognizes the recognition tag present in formula (III). Sortase enzymes have been classified into four classes, designated A, B, C and D based on sequence alignment and phylogenetic analysis (Dramsi *et al.*, Res. Microbiol., 2005, 156(3):289-97). Sortase enzymes may be derived from any bacterial species or strain, such as Gram positive bacterial species. Different sortase enzymes may recognize different sortase recognition tags and the sortase recognition tag used in formula (III) will correspond to the sortase recognition tag of the sortase enzyme used in the method.

In some embodiments, the sortase enzyme is a Sortase A enzyme, which recognizes the recognition tag:

$X_aX_bX_cX_dX_e$

in which

X_a is leucine or isoleucine, especially leucine;

X_b is proline or glycine, especially proline;

X_c is any amino acid residue, especially glutamic acid, aspartic acid, alanine, asparagine, lysine or arginine, most especially glutamic acid;

X_d is threonine or alanine; and

X_e is glycine or alanine.

In a particular embodiment, the recognition tag is LPX_cTG (SEQ ID NO:1) where X_c is any amino acid residue, especially glutamic acid, aspartic acid, alanine, asparagine, lysine or arginine, most especially glutamic acid.

In some embodiments, the sortase enzyme is a Sortase B enzyme, which recognizes the recognition tag:

$X_aX_bX_cX_dX_e$

in which

X_a is asparagine;

X_b is proline, serine or alanine, especially proline;

X_c is any amino acid residue, especially glutamic acid, aspartic acid or lysine;

X_d is threonine or serine; and

X_e is asparagine, glycine or serine.

In some embodiments, the sortase enzyme is a Sortase C enzyme which recognizes the recognition tag:

LPX_cTG

in which X_c is any amino acid residue.

In some embodiments, the sortase enzyme is a Sortase D enzyme which recognizes the recognition tag:

$X_aX_bX_cX_dX_e$

in which

X_a is asparagine or leucine;

X_b is alanine or proline;

X_c is any amino acid residue, especially glutamic acid, serine, histidine and asparagine;

X_d is threonine; and

X_e is glycine or alanine.

Sortase enzymes may be derived from a number of Gram positive bacteria including *Staphylococcus aureus*, *Bacillus anthracis*, *Enterococcus faecalis*; *Listeria monocytogenes*, *Streptococcus gordonii*, *Streptococcus pneumoniae*, *Streptococcus sanquinis*, *Streptococcus suis*, *Streptococcus agalactiae*, *Streptococcus pyrogenes*, *Streptococcus mutans* and *Actinomyces naeslundii*.

In a particular embodiment, the sortase enzyme is Sortase A from *Staphylococcus aureus* and the recognition of $X_aX_bX_cX_dX_e$ is LPX_cTG (SEQ ID NO:1), especially LPETG (SEQ ID NO:380).

X_3 is absent or is a purification tag such as His_6, cMyc protein, Hemagglutinan (HA), FLAG-tag, HSV-tag or V5-tag, especially His_6. This can be particularly helpful in purification of the reaction product as unreacted formula (III) and the $G\text{-}X_3$ released in the sortase mediated reaction may be readily removed by chromatography such as affinity chromatography. In some embodiments, when A is a protein or peptide, the sortase recognition tag is attached to the C-terminus of the protein or peptide. In other embodiments, the sortase recognition tag is located within the peptide or protein.

In some embodiments, it has been observed that the sortase mediated reaction is more efficacious if the sortase recognition tag is located at or near the C-terminus of the protein or peptide. For example, the sortase recognition tag or sortase recognition tag and purification tag are located at the C-terminus.

In some embodiments, when A is a protein or peptide which requires a free C-terminal carboxy group for activity, X_3 is a covalent bond to A. In these embodiments, the sortase recognition tag is located in an accessible loop within the protein or peptide. By "accessible loop" is meant that the recognition tag appears in the sequence of the protein or peptide which is on the surface of the tertiary structure of the protein or peptide and is accessible to both the peptide of formula (I) having the $G(G)_n$ sequence and the sortase enzyme. Although the sortase enzyme will cleave the recognition tag, the protein or peptide may have sufficient tertiary structure, such as disulfide bonds, to maintain the protein intact.

A is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a photosensitizer, a virus particle, a carbohydrate, a contrast agent, a catalyst, a lipid, a DNA molecule or an RNA molecule;

especially a protein or peptide, more especially a protein, more especially an antibody or a fragment thereof.

The protein or peptide may be any protein or peptide that is useful when conjugated to another moiety, such as another peptide or protein, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a polymer, a photosensitizer, a virus particle, a carbohydrate, a contrast agent, a catalyst, a lipid, a DNA molecule or an RNA molecule, or may be useful conjugated to a solid surface.

The protein or peptide may be biologically active, for example, an antibody or antibody chain (heavy or light chain), or a portion comprising an immunoglobulin domain (constant or variable domain), a cell proliferation factor, an apoptosis factor, an extracellular signaling molecule, a receptor, an angiogenesis factor or a cell interaction factor.

In some embodiments, the antibody is IgG, IgM, IgA or IgE or a fragment thereof. In some embodiments, the antibody is a polyclonal or monoclonal antibody and may be chimeric, humanized or bispecific. The antibody may be selected for its interaction with a specific antigen. Examples of suitable antibodies include, but are not limited to, muromonab, abciximab, rituximab, daclizumab, basiliximab, palivizumab, infliximab, trastuzumab, gemtuzumab, alemtuzumab, ibritumomab, adelimumab, omalizumab, tositumomab, efalizumab, cetuximab, bevacizumab, natalizumab, ranibizumab, panitumumab, eculizumab, certolizumab, pexelizumab, ipilimumab, zanolimumab, denosumab, adalimumab, belimumab and motauizumab.

In some embodiments, the protein or peptide is an antibody fragment, such as Fab, Fab', F(ab)'2, F(ab)'3, Dab, Fv, single chain Fv (scFv) fragment scFv-Fc (scFv)2, intrabodies (Kontermann, Methods, 2004, 34:163-170), diabodies, triabodies or tetrabodies, especially a scFv fragment. Examples of suitable antibody fragments include, but are not limited to, anti-ED-B-scFv (extra domain B fibronectin), anti-LIBS GPIIb/IIIa scFv (platelet integrin), anti-E-selectin scFv, anti-P-selectin scFv, anti-Mac-1 scFv (leukocyte integrin), 59D8 (fibrin), MA-15CF, A33-scFv, anti-γ-SM scFv, ch-TNT-3 Fab and L49 scFv.

Examples of other proteins and peptides include proteins and peptide involved in recognition of other proteins and peptides, including, but are not limited to protein kinases such as mitogen activated protein (MAP) kinase, and kinases that directly or indirectly phosphorylate MAP kinase, Januse kinase (JAKI) and cyclin dependent kinases, epidermal growth factor (EGF) receptor, platelet-derived growth factor (PDGF) receptor, fibroblast-derived growth factor (FGF) receptor, insulin receptor and insulin-like growth factor (IGF) receptor, protein phosphatases such as PTPIB, PP2A and PP2C, GDP/GTP binding proteins such as Ras, Raf, ARF, Ran and Rho, GTPase activating proteins (GAF), guanidine nucleotide exchange factors (GEFs), proteases such as caspase 3, 8 and 9, peroxidases, ubiquitin ligases such as MDM2 and E3 ubiquitin ligases, acetylation and methylation proteins such as p300/CBP and histone acetyl transferase, tumor suppressor proteins such as p53, cytokines such as interleukins (IL-1, IL-1ra, IL-2, IL-10, IL-11, IL-6), interferons (IFN-α, IFN-β), erythropoietin (EPO), thrombopoietin (TPO), TGF-β, and chemokines (CCL-1 to CCL-28, α-chemokines, β-chemokines, lymphotactin-α, lymphotactin-β and fractalkine), and their receptors such as G-protein coupled receptors, haemopoeitic growth factor, tumor necrosis factors (TNF).

Other suitable proteins include DARPins (designed ankyrin repeat proteins), which are small, single domain proteins that can be selected to bind to a given target protein with high affinity and specificity (Stumpp *et al.*, Drug Discovery Today, 2008, 13:695-701).

In some embodiments, the protein may be an albumin, such as human serum albumin, and this modifies the pharmacokinetic properties of the conjugate.

Suitable labels include but are not limited to fluorescent labels, phosphorescent labels, chemiluminescent labels, streptavadin, biotin, poly(histidine) or a radioactive label.

Suitable small molecule drugs may be any drugs that may be delivered in a more targeted manner or may be more soluble when part of a conjugate. For example, the small molecule drug may be toxic to normal cells as well as diseased cells or microbial cells and

conjugation allows targeting to the diseased or microbial cell reducing toxicity to normal host cells. The drug may be a poorly soluble drug that has no or low bioavailability, especially oral bioavailability and which for solubility reasons is difficult to formulate for parenteral administration, and the conjugate improves solubility and bioavailability. Suitable small molecule drugs include diabetes therapies such as biguanides, sulfonylureas, glitazones and insulin; anti-obesity drugs such as orlistat, rimonabant and sibutramine; anti-microbial drugs such as antibiotics, anti-fungal agents and anti-protozoal agents, anti-viral agents such as zovirax and Azt; cholesterol lowering agents such as HMG-CoA inhibitors (statins), cardiovascular drugs such as calcium channel blockers, β-blockers, ACE inhibitors and diuretics, analgesics such as opioids and non-steroidal anti-inflammatory drugs, anti-convulsant drugs such as benzodiazepines, barbiturates and GABA analogues, anti-histamines such as diphenhydramine and cimetidine; asthma drugs such as salbutamol and fluticasone, cancer therapies such as chemotherapeutic drugs including cisplatin, carboplatin, taxol and related taxane compounds; immune suppressing drugs such as cyclosporine, hormones and hormone analogues such as insulin and thyroxine; vitamins such as vitamin B group compounds, vitamin E, vitamin C and vitamin A; contraceptives, muscle relaxants such as tubocurarine, succinylcholine and pancuronium; sedatives including barbiturates and benzodiazapines; arthritis drugs such as COX2-selective inhibitors; heparin; fXa blockers such as rivaroxaban; and thrombin inhibitors such as argatroban, lepirudin, desirudin and bivalirudin.

Suitable radioisotopes include but are not limited to, ^{3}H, ^{47}Ca, ^{11}C, ^{14}C, ^{57}Co, ^{58}Co, ^{64}Cu, ^{18}F, ^{67}Ga, ^{68}Ga, ^{18}F, ^{111}In, ^{123}I, ^{124}I, ^{125}I, ^{131}I, ^{32}P, ^{75}Se, ^{153}Sm, ^{13}N, ^{22}Na, ^{24}Na, ^{15}O, ^{89}Sr, ^{99}Tc, ^{201}Tl, ^{133}Xe and ^{90}Y.

Suitable targeting agents are compounds that bind to cell receptors in specific parts of the body such as tissues or organs, or to cell receptors on unwanted cells such as microbial cells or cancer cells. Antibodies or parts thereof may be used to target particular cells or tissues in the body. Another example is folate as folate receptors overexpressed on the surface of some cancer cells. A further example is cytokines and chemokines and their cell surface receptors.

Suitable nanoparticles include, but are not limited to, iron oxide particles used in MRI imaging (MPIO) (McAteer *et al.*, Methods Mol. Biol., 2011, 680:103-15; Guenther *et al.*, Invest. Radiol., 2010, 45(10):586-91; von Zur Muhlen *et al.*, Circulation, 2008, 118:258-67; von Zur Muhlen *et al.*, J. Vasc. Res., 2009, 46:6-14; von Zur Muhlen *et al.*, J. Clin. Invest., 2008, 118:1198-207; McAteer *et al.*, Arterioscler. Thromb. Vasc. Biol., 2008, 28:77-83; McAteer *et al.*, Nat. Med., 2007, 13(10):1253-8), LbL click capsules (Leung *et al.*, Small, 2011, 21 Mar. doi:10.1002/smll. 201002258; Chong *et al.*, Langmuir, 2011, 1 Mar, 27(5):1724-30; Such *et al.*, Chem. Soc. Rev., 2011, 40(1):19-29; Becker *et al.*, Small, 2010, 6(17):1836-52), lipid nanoparticles, liposomes, micelles, GdO particles (Faure *et al.*, Small, 2009, 5(22):2565-75; Ou *et al.*, J. Colloid Interface Sci., 2009, 333(2):684-9; Alric *et al.*, J. Am. Chem. Soc., 2008, 130(18):5908-15) and F click capsules.

Suitable cells include, but are not limited to, stem cells, cells from specific tissues or organs in the body, T-regulatory cells or Natural Killer T cells.

Suitable dendrimers include, but are not limited to, polyamidoamine (PAMAM), poly(ethyleneimine) (PEI), polylysine, polyglutamine, poly(etherhydroxylamine) (PEHAM) and poly(propyleneimine) (PPI) dendrimers. Other suitable dendrimers incorporate diamino building units such as analogues of lysine. Architecture of lysine and lysine analogue dendrimers has been described by Denkewalter in US Patent No. 4,289,872. More preferably the dendrimer includes the at least two generations of building units and all contain one or more branches originating from a core molecule. In particular embodiments, the dendrimer is a polylysine dendrimer of two to eight generations.

Suitable photosensitizers, include but are not limited to, porphyrines, chlorophylls and dyes such as aminolevulinic acid, levulinic acid, methyl aminolevulinic acid, photofrin, Visudyne, Foscan, Metvix, Hexvix®, Cysview™, Laserphryin, Antrin, Photochlor, Photosens, Photrex, Civira, Visonac, Amphinex and azadipyrromethenes.

Suitable virus particles, include but are not limited to, virus particles able to carry DNA for use in gene therapy. Examples of suitable virus particles include, but are not limited to adenovirus, adeno-associated virus, lentivirus, herpes simplex virus, replication-competent vectors, envelope protein pseudotyping of viral vectors and shells from disease causing viruses.

Suitable carbohydrates include carbohydrates that provide cell signalling functions or those that are unique to the cell surfaces of microbes or viruses. Examples include carbohydrates comprising one or more monomers selected from sialic acid, glucose, galactose, mannose and fucose, particularly Sialyl Le^X and L-rhamnopyranose.

Suitable contrast agents include but are not limited to paramagnetic particles or metal ions including Cr^{3+}, Mn^{2+}, Fe^{3+}, Co^{2+}, Cu^{2+}, Pr^{3+}, Eu^{3+}, Gd^{3+}, Tb^{3+}, Tb^{4+}, Dy^{3+}, Ho^{3+} and Er^{3+}.

Suitable catalysts include enzymes, such as thrombolytic enzymes and proteases.

Suitable lipids include long chain fatty acids, such as myristic acid, that provide an increase in plasma half-life.

Suitable DNA and RNA is any suitable DNA or RNA or fragment thereof that has therapeutic value. Examples include SiRNA (gene silencing), antagomirs (microRNA inhibition) and microRNA (effector RNA).

The GGG peptide moiety is a compound having formula (I):

$$R_1\text{-HN-G(G)}_n\text{-}X_1X_2Y_1Z_1 \qquad (I)$$

wherein G is a glycine residue;

R_1 is hydrogen or an amino protecting group;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

X_2 is absent or is an amino acid residue with a reactive functional group in its side chain;

Y_1 is absent or is a spacer group;

Z_1 is a reactive functional group; and

n is 0 or an integer of 1 to 4.

In some embodiments, n is 0 and only one glycine residue is present at the N-terminus of formula (I). In some embodiments n is 1 and two glycine residues are present at the N-terminus of formula (I). In other embodiments, n is 2 and three glycine residues are present at the N-terminus of formula (I). In some embodiments, n is 3 and four glycine residues are present at the N-terminus of formula (I). In other embodiments, n is 4 and five glycine residues are present at the N-terminus of formula (I). In particular embodiments, n is 2.

In some embodiments, X_1 is absent. In other embodiments, X_1 is an amino acid residue or a peptide of 2-10 amino acid residues in length, especially 2-9, 2-8, 2-7, 2-5, 2-4 or 2-3 amino acid residues in length, most especially 2-5 amino acid residues in length. While X_1 may be any amino acid residue or peptide of 2-10 amino acid residues in length, in some embodiments, X_1 comprises one or more amino acid residues with a bulky side chain, especially where there is an amino acid residue with a bulky side chain attached to the C-terminal glycine residue of the $G(G)_n$ sequence.

In one embodiment, X_1 is a peptide of the sequence:

$Xaa_1Xaa_2Xaa_3Xaa_4$

wherein Xaa_1 is an amino acid residue having a bulky side chain;

Xaa_2 is absent or is an amino acid residue having a bulky side chain;

Xaa_3 is absent or any amino acid residue; and

Xaa_4 is absent or any amino acid residue.

In some embodiments Xaa_1 is selected from tryptophan, tyrosine, phenylalanine, leucine, isoleucine and histidine, especially tryptophan, tyrosine and phenylalanine, most especially tryptophan.

In some embodiments, where Xaa_2 is present, Xaa_2 is selected from tryptophan, tyrosine, phenylalanine, leucine, isoleucine and histidine, especially tryptophan, tyrosine and phenylalanine, most especially tryptophan.

In some embodiments, where Xaa_3 is present, Xaa_3 is selected from serine, alanine, aspartic acid, methionine, threonine and valine, especially serine or threonine, most especially serine.

In some embodiments, where Xaa_4 is present, Xaa_4 is selected from serine, alanine, aspartic acid, methionine, threonine and valine, especially serine or threonine, most especially serine.

In some embodiments, one or more of Xaa_1, Xaa_2, Xaa_3 or Xaa_4 is absent. For example, in some embodiments, Xaa_3 and Xaa_4 are absent and Xaa_1 and Xaa_2 make a peptide of two amino acid residues or Xaa_4 is absent and Xaa_1, Xaa_2 and Xaa_3 make a peptide of three amino acid residues.

In some embodiments, X_1 is selected from -W-, -WS-, -WG-, -WA-, -WW-, -WWS-, -WWG-, -WWA-, -WWSS- (SEQ ID NO:2), -WWSG- (SEQ ID NO:3), -WWSA- (SEQ ID NO:4), -WWGS- (SEQ ID NO:5), -WWGG- (SEQ ID NO:6), -WWGA- (SEQ ID NO:7), -WWAS- (SEQ ID NO:8), -WWAG- (SEQ ID NO:9) and -WWAA- (SEQ ID NO:10).

In some embodiments, X_2 is absent and Y_1 or Z_1 is attached directly to the C-terminus of X_1 or the $G(G)_n$ sequence. In other embodiments, X_2 is an amino acid with a reactive functional group in its side chain and Y_1 or Z_1 is attached to the C-terminal carboxy group of X_2 or to the reactive functional side chain group of X_2. In some embodiments, X_2 is

selected from lysine, ornithine, aspartic acid and glutamic acid, especially lysine or ornithine, most especially lysine.

Y_1 may be absent or may be a spacer group that connects Z_1 to the C-terminus of the G(G)n sequence, the C-terminus of X_1 or the C-terminus or reactive functional group of X_2. The spacer group may be any sort of spacer group and may provide relief from steric hindrance around the Z_1 functional group or may be a pharmacokinetic modifying agent.

In some embodiments, the spacer group is a group containing 1-100 carbon atoms covalently connected by single or multiple bonds, wherein one or more carbon atoms are optionally replaced with heteroatoms sch as O, N, S, or P. Examples include linear or branched alkylene, alkenylene or alkynylene groups or divalent cycloalkyl, aryl, heterocyclyl or heteroaryl groups or combinations thereof, wherein one or more carbon atoms in an alkylene, alkenylene or alkynylene group is optionally replaced by a heteroatom selected from O, N, S or P, and wherein each alkylene, alkenylene, alkynylene, cycloalkyl, aryl, heterocyclyl or heteroaryl group is optionally substituted, for example, with one or more alkyl, alkenyl, alkynyl, cycloalkyl, aryl, heterocyclyl, heteroaryl, $-N(R_2)_2$, oxo group, $-OR_2$, $-SR_2$, $-COR_2$, $-CO_2R_2$, $-CONH_2$, $-NHC(O)-NH_2$, $-NH-C(=NH)-NH_2$, $-PO_4H_3$ or $-S(O)_2N(R_2)_2$ wherein each R_2 is independently selected from hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, aryl, heteroalkyl or heteroaryl.

In some embodiments, the spacer has a molecular weight in the range of 85 Da – 1200 Da. In some embodiments, the spacer is polyethylene glycol (PEG) having a molecular weight between 85 and 1200 Da, for example PEG_2, PEG_4, PEG_6, PEG_8, PEG_{12} or PEG_{24}, especially PEG_4, PEG_6, PEG_8 or PEG_{12}.

In some embodiments, the spacer is a pharmacokinetic modifying agent. In particular embodiments, the pharmacokinetic modifying agent is an oligomer or polymer comprising monomers selected from ethyleneoxy, propyleneoxy, alkyloxazoline such as ethyloxazoline, vinylpyrrolidone and amino acids such as lysine, glutamate and aspartate. In some embodiments the pharmacokinetic modifying agent is selected from polyethylene

glycol (PEG), polypropylene glycol (PPG), polyethylene oxide (PEO), a poly(alkyloxazoline) such as poly(ethyloxazoline (PEOX), polyvinylpyrrolidone (PVPON), polylysine, polyglutamate and polyaspartate, especially PEG, PPG, PEOX, polylysine, polyglutamate and polyaspartate, or mixtures thereof, more especially PEG.

In some embodiments, the polymer has a molecular weight in the range of 5 kDa – 100 kDa. In some embodiments, the pharmacokinetic modifying polymer is PEG having a molecular weight between 5 kDa and 60 kDa, for example, 10 kDa, 15 kDa, 20 kDa, 30 kDa or 40 kDa, especially 20 kDa.

In a particular embodiment, X_2 is a lysine, ornithine, glutamic acid or aspartic acid residue and the spacer, especially PEG, is attached to the ε-amino group of lysine, the δ-amino group of ornithine, the γ-carboxy group of aspartic acid or the δ-carboxy group of glutamic acid, especially where X_2 is lysine and the spacer is attached to the ε-amino group.

In a particular embodiment, X_2 is a lysine, ornithine, glutamic acid or aspartic acid residue and the pharmacokinetic modifying agent, especially PEG or a polylysine, polyaspartic acid or polyglutamic acid group, is attached to the ε-amino group of lysine, the δ-amino group of ornithine, the γ-carboxy group of aspartic acid or the δ-carboxy group of glutamic acid, especially where X_2 is lysine and the modifying agent is attached to the ε-amino group.

In embodiments where the spacer or pharmacokinetic modifying agent is attached to the side chain functional group of X_2, the C-terminal carboxy group of X_2 may optionally have another group attached, for example, a label such as a fluorescent label, a radioactive label or an imaging label.

Z_1 is a reactive functional group that is capable of reacting with another moiety. In some embodiments, the reactive functional group is a group complementary to a reactive functional group on the moiety to which it is proposed to be attached. For example, Z_1 may react with another functional group via a 1,3-cycloaddition reaction, hydrazone

formation, oxime formation, amide formation such as an amide formed by native chemical ligation, thio-ene reaction or Diels-Alder reaction.

In some embodiments, Z_1 is selected from an azide, an alkyne, a hydroxylamine, a hydrazine, an aldehyde, a ketone, an amino group, a carboxylic acid, a thiol group, an alkene, a thioester or a cysteine residue. In particular embodiments, Z_1 is an azide, an alkyne, a hydroxylamine, a hydrazine, an aldehyde, an amino group, a carboxylic acid or a thioester. In some embodiments, the carboxylic acid is not the C-terminal carboxylic acid of the glycine moiety, X_1 or X_2.

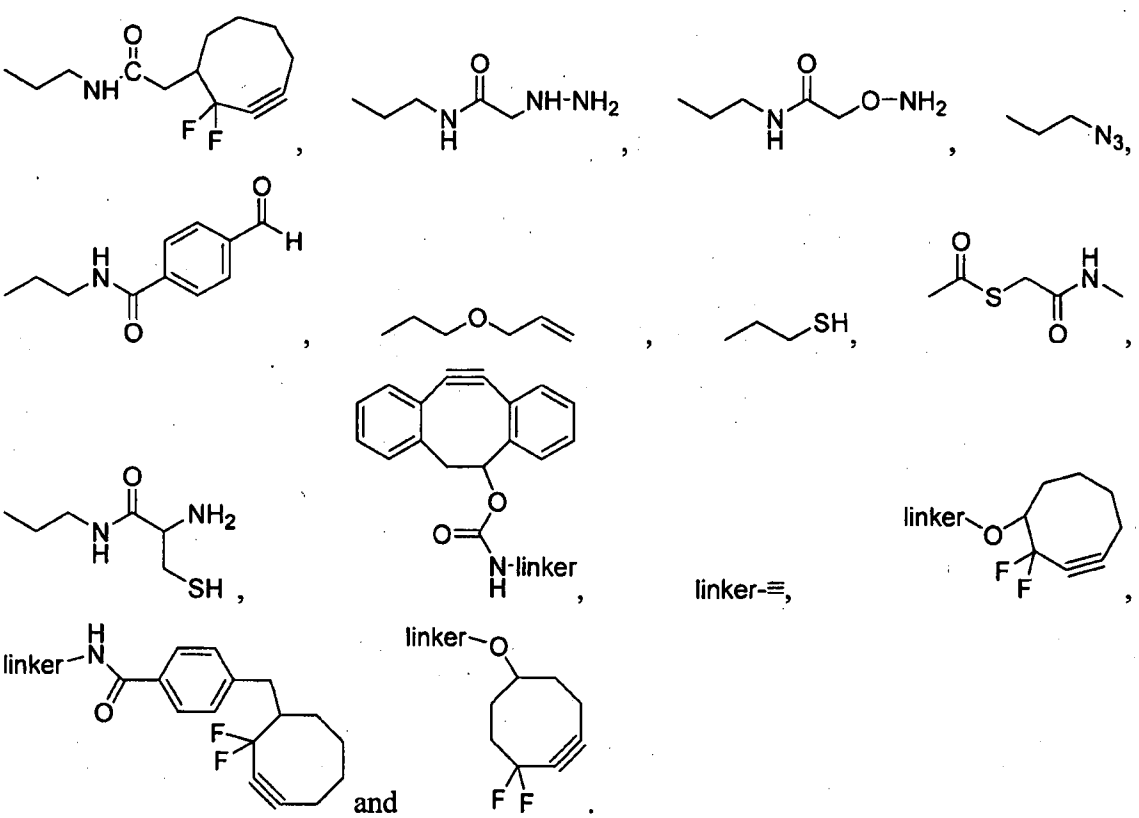
The Z_1 group may be attached directly to the Y_1 or the C-terminal carboxy group of X_1 or X_2 or the C-terminal carboxy group of the $G(G)_n$ sequence or may be derived from the terminal group of Y_1 or the C-terminal carboxy group of the $G(G)_n$ sequence, X_1 or X_2. Alternatively, Z_1 may be attached to the $G(G)_n$ sequence, X_1, X_2 or Y_1 through a linker group.

For example, where the group Y_1 is a polymer that terminates with a carboxy group, it may be reduced to an aldehyde group by methods known in the art. Where Y_1 is a polymer that terminates in an amino group, it may be used directly or modified to an azide group, a thiol group or an alkene by methods known in the art.

Alternatively, the terminal group of Y_1 or the C-terminal carboxy group of X_1 or the $G(G)_n$ sequence may be reacted with a linker group that bears the functional group Z_1. Suitable linker groups include alkylene groups ($(CH_2)_x$) in which x is an integer from 1 to 10 and in which one or more CH_2 groups may be replaced by -O-, -NH-, -NH-C(O)-, -C(O)-NH, -S- or $-C_6H_4-$.

Suitable Z_1 groups, optionally including a linker, include:
$-(CH_2)_xN_3$, $-(CH_2)_xONH_2$, $-(CH_2)_xN{=}NH_2$, $-(CH_2)_x$-SH, $-(CH_2)_xCHO$, $-(CH_2)_xC(O)SR$, $-(CH_2)_xC(O)R$, $-(CH_2)_x$alkyne, $-(CH_2)_xCH{=}CH_2$, $-(CH_2)_xNH$-C(O)CH(NH_2)(CH_2SH), $-(CH_2)_mNHC(O)(CH_2)_p$alkyne, $-(CH_2)_m$-NHC(O)-$(CH_2)_pNHNH_2$, $-(CH_2)_mNHC(O)(CH_2)_p$-

ONH_2, $-(CH_2)_mNHC(O)(CH_2)_p$-SH, $-(CH_2)_m$-NHC(O)$(CH_2)_p$CHO, $-(CH_2)_mNHC(O)(CH_2)_pC(O)R$, $-(CH_2)_mNHC(O)(CH_2)_pC(O)SR$, $-(CH_2)_mNHC(O)(CH_2)_p$-$CH=CH_2$, $-(CH_2)_mC(O)NH(CH_2)_p$alkyne, $-(CH_2)_mC(O)NH-(CH_2)_pNHNH_2$, $-(CH_2)_mC(O)NH-(CH_2)_p-ONH_2$, $-(CH_2)_m-C(O)NH-(CH_2)_pSH$, $-(CH_2)_m$-$C(O)NH(CH_2)_pCHO$, $-(CH_2)_mC(O)NH-(CH_2)_pC(O)R$, $-(CH_2)_mC(O)NH-(CH_2)_pSR$, $-(CH_2)_mC(O)(CH_2)_p-CH=CH_2$, $-(CH_2)_qNHCO-C_6H_4-CHO$ and $-(CH_2)_qNHCOC_6H_4-C(O)R$ wherein m is an integer from 1 to 10, m + p = x-1 and q = x-2 and R is an alkyl group or alkyl group substituted with an amide group. Some suitable linker-Z_1 groups include, but are not limited to:



R_1 is hydrogen or an amino protecting group. In some cases where the compound of formula (I) may need to be reacted with another moiety before sortase conjugation of the $G(G)_n$ moiety with an LPXTG (SEQ ID NO:1) motif, the N-terminal amino group of the $G(G)_n$ moiety may need to be protected. Suitable protecting groups are known in the art, such as in Greene and Wutz, Protective Groups in organic synthesis, 3rd Ed., Wiley

Interscience, 1999. Commonly used amino protecting groups include, but are not limited to, benzyloxycarbonyl (Z), t-butoxycarbonyl (Boc), 2-(4-biphenylyl)-isopropoxycarbonyl (Bpoc), 9-fluorenylmethoxycarbonyl (Fmoc), triphenylmethyl (trityl, trt) and 2-nitrophenylsulphenyl (Nps), especially those groups which can be cleaved using conditions which do not lead to undue loss of protein function, such as Boc and Fmoc.

In some embodiments, the compound of formula (I) is selected from:

GGGW-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:11),

GGGW-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:12),

GGGW-PEG$_4$-CH$_2$C≡CH (SEQ ID NO:13),

GGGW-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:14),

GGGW-PEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:15),

GGGW-PEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:16),

GGGW-PEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:17),

GGGW-PEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:18),

GGGW-PEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:19),

GGGW-PEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:20),

GGGW-PEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:21),

GGGK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:22),

GGGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:23),

GGGK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:24),

GGGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:25),

GGGK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:26),

GGGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:27),

GGGK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:28),

GGGK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:29),

GGGK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:30),

GGGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:31),

GGGK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:32),

GGG-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:33),

GGG-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:34),

GGG-PEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:35),

GGG-PEG_4-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:36),

GGG-PEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:37),

GGG-PEG_4-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:38),

GGG-PEG_4-$(CH_2)_2SH$ (SEQ ID NO:39),

GGG-PEG_4-$CH_2CH{=}CH_2$ (SEQ ID NO:40),

GGG-PEG_4-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:41),

GGG-PEG_4-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:42),

GGG-PEG_4-$(CH_2)_2$-Maleimide (SEQ ID NO:43),

GGGWW-PEG_4-$(CH_2)_2N_3$ (SEQ ID NO:44),

GGGWW-PEG_4-$(CH_2)_2NHC(O)CH_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:45),

GGGWW-PEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:46),

GGGWW-PEG_4-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:47),

GGGWW-PEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:48),

GGGWW-PEG_4-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:49),

GGGWW-PEG_4-$(CH_2)_2SH$ (SEQ ID NO:50),

GGGWW-PEG_4-$CH_2CH{=}CH_2$ (SEQ ID NO:51),

GGGWW-PEG_4-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:52),

GGGWW-PEG_4-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:53),

GGGWW-PEG_4-$(CH_2)_2$-Maleimide (SEQ ID NO:54),

GGGWWK-εPEG_4-$(CH_2)_2N_3$ (SEQ ID NO:55),

GGGWWK-εPEG_4-$(CH_2)_2NHC(O)CH_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:56),

GGGWWK-εPEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:57),

GGGWWK-εPEG_4-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:58),

GGGWWK-εPEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:59),

GGGWWK-εPEG_4-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:60),

GGGWWK-εPEG_4-$(CH_2)_2SH$ (SEQ ID NO:61),

GGGWWK-εPEG_4-$CH_2CH{=}CH_2$ (SEQ ID NO:62),

GGGWWK-εPEG_4-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:63),

GGGWWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:64),
GGGWWK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:65),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:66),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:67),
GGGWWSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:68),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:69),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:70),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:71),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:72),
GGGWWSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:73),
GGGWWSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:74),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:75),
GGGWWSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:76),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:77),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:78),
GGGWWSSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:79),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:80),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:81),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:82),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:83),
GGGWWSSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:84),
GGGWWSSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:85),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:86),
GGGWWSSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:87),
GGGWK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:88),
GGGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:89),
GGGWK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:90),
GGGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:91),

GGGWK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:92),
GGGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:93),
GGGWK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:94),
GGGWK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:95),
GGGWK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:96),
GGGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:97),
GGGWK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:98),
GGW-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:99),
GGW-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:100),
GGW-PEG$_4$-CH$_2$C≡CH (SEQ ID NO:101),
GGW-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:102),
GGW-PEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:103),
GGW-PEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:104),
GGW-PEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:105),
GGW-PEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:106),
GGW-PEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:107),
GGW-PEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:108),
GGW-PEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:109),
GGK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:110),
GGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:111),
GGK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:112),
GGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:113),
GGK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:114),
GGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:115),
GGK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:116),
GGK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:117),
GGK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:118),
GGK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:119),
GGK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:120),
GG-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:121),
GG-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:122),

GG-PEG$_4$-$CH_2C{\equiv}CH$ (SEQ ID NO:123),

GG-PEG$_4$-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:124),

GG-PEG$_4$-$(CH_2)_2NHNH_2$ (SEQ ID NO:125),

GG-PEG$_4$-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:126),

GG-PEG$_4$-$(CH_2)_2SH$ (SEQ ID NO:127),

GG-PEG$_4$-$CH_2CH{=}CH_2$ (SEQ ID NO:128),

GG-PEG$_4$-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:129),

GG-PEG$_4$-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:130),

GG-PEG$_4$-$(CH_2)_2$-Maleimide (SEQ ID NO:131),

GGWW-PEG$_4$-$(CH_2)_2N_3$ (SEQ ID NO:132),

GGWW-PEG$_4$-$(CH_2)_2NHC(O)CH_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:133),

GGWW-PEG$_4$-$CH_2C{\equiv}CH$ (SEQ ID NO:134),

GGWW-PEG$_4$-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:135),

GGWW-PEG$_4$-$(CH_2)_2NHNH_2$ (SEQ ID NO:136),

GGWW-PEG$_4$-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:137),

GGWW-PEG$_4$-$(CH_2)_2SH$ (SEQ ID NO:138),

GGWW-PEG$_4$-$CH_2CH{=}CH_2$ (SEQ ID NO:139),

GGWW-PEG$_4$-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:140),

GGWW-PEG$_4$-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:141),

GGWW-PEG$_4$-$(CH_2)_2$-Maleimide (SEQ ID NO:142),

GGWWK-εPEG$_4$-$(CH_2)_2N_3$ (SEQ ID NO:143),

GGWWK-εPEG$_4$-$(CH_2)_2NHC(O)CH_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:144),

GGWWK-εPEG$_4$-$CH_2C{\equiv}CH$ (SEQ ID NO:145),

GGWWK-εPEG$_4$-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:146),

GGWWK-εPEG$_4$-$(CH_2)_2NHNH_2$ (SEQ ID NO:147),

GGWWK-εPEG$_4$-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:148),

GGWWK-εPEG$_4$-$(CH_2)_2SH$ (SEQ ID NO:149),

GGWWK-εPEG$_4$-$CH_2CH{=}CH_2$ (SEQ ID NO:150),

GGWWK-εPEG$_4$-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:151),

GGWWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:152),
GGWWK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:153),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:154),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:155),
GGWWSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:156),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:157),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:158),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:159),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:160),
GGWWSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:161),
GGWWSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:162),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:163),
GGWWSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:164),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:165),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:166),
GGWWSSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:167),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:168),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:169),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:170),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:171),
GGWWSSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:172),
GGWWSSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:173),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:174),
GGWWSSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:175),
GGWK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:176),
GGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:177),
GGWK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:178),
GGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:179),

GGWK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:180),

GGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:181),

GGWK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:182),

GGWK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:183),

GGWK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:184),

GGWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:185),

GGWK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:186),

GW-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:187),

GW-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:188),

GW-PEG$_4$-CH$_2$C≡CH (SEQ ID NO:189),

GW-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:190),

GW-PEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:191),

GW-PEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:192),

GW-PEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:193),

GW-PEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:194),

GW-PEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:195),

GW-PEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:196),

GW-PEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:197),

GK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:198),

GK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:199),

GK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:200),

GK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:201),

GK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:202),

GK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:203),

GK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:204),

GK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:205),

GK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:206),

GK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:207),

GK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:208),

G-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:209),

G-PEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:210),

G-PEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:211),

G-PEG_4-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:212),

G-PEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:213),

G-PEG_4-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:214),

G-PEG_4-$(CH_2)_2SH$ (SEQ ID NO:215),

G-PEG_4-$CH_2CH{=}CH_2$ (SEQ ID NO:216

G-PEG_4-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:217)

G-PEG_4-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:218),

G-PEG_4-$(CH_2)_2$-Maleimide (SEQ ID NO:219),

GWW-PEG_4-$(CH_2)_2N_3$ (SEQ ID NO:220),

GWW-PEG_4-$(CH_2)_2NHC(O)CH_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:221),

GWW-PEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:222),

GWW-PEG_4-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:223),

GWW-PEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:224),

GWW-PEG_4-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:225),

GWW-PEG_4-$(CH_2)_2SH$ (SEQ ID NO:226),

GWW-PEG_4-$CH_2CH{=}CH_2$ (SEQ ID NO:227),

GWW-PEG_4-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:228),

GWW-PEG_4-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:229),

GWW-PEG_4-$(CH_2)_2$-Maleimide (SEQ ID NO:230),

GWWK-εPEG_4-$(CH_2)_2N_3$ (SEQ ID NO:231),

GWWK-εPEG_4-$(CH_2)_2NHC(O)CH_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:232),

GWWK-εPEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:233),

GWWK-εPEG_4-$(CH_2)_2NHC(O)CH_2ONH_2$ (SEQ ID NO:234),

GWWK-εPEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:235),

GWWK-εPEG_4-$(CH_2)_2NHC(O)$-C_6H_4-CHO (SEQ ID NO:236),

GWWK-εPEG_4-$(CH_2)_2SH$ (SEQ ID NO:237),

GWWK-εPEG_4-$CH_2CH{=}CH_2$ (SEQ ID NO:238),

GWWK-εPEG_4-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:239),

GWWK-εPEG_4-$(CH_2)_2NHC(O)$-$CH(NH_2)CH_2SH$ (SEQ ID NO:240),

GWWK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:241),
GWWSK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:242),
GWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:243),
GWWSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:244),
GWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:245),
GWWSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:246),
GWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:247),
GWWSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:248),
GWWSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:249),
GWWSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:250),
GWWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:251),
GWWSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:252),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:253),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:254),
GWWSSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:255),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:256),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:257),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:258),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:259),
GWWSSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:260),
GWWSSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:261),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:262),
GWWSSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:263),
GWK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:264),
GWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:265),
GWK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:266),
GWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:267),
GWK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:268),
GWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:269),

GWK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:270),
GWK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:271),
GWK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:272),
GWK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:273),
GWK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:274),
GGGWWGA-pG (SEQ ID NO:275),
GGGAGAGAC (SEQ ID NO:276),
GGGWWSSK-εPEG$_4$-SANH (SEQ ID NO:277),
GGGWWSSK-εPEG$_4$-DBCO (SEQ ID NO:278),
GGGWWSSK-εPEG$_4$-penenoic acid (SEQ ID NO:279),
GGGWSK-εPEG$_4$-(CH$_2$)N$_3$ (SEQ ID NO:280),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:281),
GGGWSK-εPEG$_4$-CH$_2$C≡CH (SEQ ID NO:282),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:283),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:284),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:285),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:286),
GGGWSK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:287),
GGGWSK-εPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:288),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:290),
GGGWSK-εPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:290),
GGG-EEE-SH (SEQ ID NO:291),
GGWE-PEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:292),
GGGWSK-(ε-NH)-CH$_2$C≡CH (SEQ ID NO:293),
GGGG-EEE-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:294),
GGYK-εPEG$_{12}$-(CH$_2$)$_2$NHCOC$_6$H$_4$CHO (SEQ ID NO:295),
GGYK-εPEG$_6$-(CH$_2$)$_2$NHCOC$_6$H$_4$CHO (SEQ ID NO:296),
GE-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:297),
GGSEC (SEQ ID NO:298),
GGGFDK-εKKK-(CH$_2$)$_2$N$_3$ (SEQ ID NO:299),

GGGFDK-εKKK-$(CH_2)_2$NHC(O)CH_2-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:300),

GGGFDK-εKKK-$CH_2C{\equiv}CH$ (SEQ ID NO:301),

GGGFDK-εKKK-$(CH_2)_2$NHC(O)CH_2ONH_2 (SEQ ID NO:302),

GGGFDK-εKKK-$(CH_2)_2NHNH_2$ (SEQ ID NO:303),

GGGFDK-εKKK-$(CH_2)_2$NHC(O)-C_6H_4-CHO (SEQ ID NO:304),

GGGFDK-εKKK-$(CH_2)_2SH$ (SEQ ID NO:305),

GGGFDK-εKKK-$CH_2CH{=}CH_2$ (SEQ ID NO:306),

GGGFDK-εKKK-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:307),

GGGFDK-εKKK-$(CH_2)_2$NHC(O)-$CH(NH_2)CH_2SH$ (SEQ ID NO:308),

GGGFDK-εKKK-$(CH_2)_2$-Maleimide (SEQ ID NO:309),

GGGWSOrn (δPEG_4)-$(CH_2)_2N_3$ (SEQ ID NO:310),

GGGWSOrn (δPEG_4)-$(CH_2)_2$NHC(O)CH_2-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:311),

GGGWSOrn (δPEG_4)-$CH_2C{\equiv}CH$ (SEQ ID NO:312),

GGGWSOrn (δPEG_4)-$(CH_2)_2$NHC(O)CH_2ONH_2 (SEQ ID NO:313),

GGGWSOrn (δPEG_4)-$(CH_2)_2NHNH_2$ (SEQ ID NO:314),

GGGWSOrn (δPEG_4)-$(CH_2)_2$NHC(O)-C_6H_4-CHO (SEQ ID NO:315),

GGGWSOrn (δPEG_4)-$(CH_2)_2SH$ (SEQ ID NO:316),

GGGWSOrn (δPEG_4)-$CH_2CH{=}CH_2$ (SEQ ID NO:317),

GGGWSOrn (δPEG_4)-$CH_2C(O)SCH_2C(O)NHCH_3$ (SEQ ID NO:318),

GGGWSOrn (δPEG_4)-$(CH_2)_2$NHC(O)-$CH(NH_2)CH_2SH$ (SEQ ID NO:319),

GGGWSOrn (δPEG_4)-$(CH_2)_2$-Maleimide (SEQ ID NO:320),

GGGWSOrn (δPEG_4)-$(CH_2)_2SCOCH_3$ (SEQ ID NO:321),

GGGE-δPEG_4-$(CH_2)_2N_3$ (SEQ ID NO:322),

GGGE-δPEG_4-$(CH_2)_2$NHC(O)CH_2-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:323),

GGGE-δPEG_4-$CH_2C{\equiv}CH$ (SEQ ID NO:324),

GGGE-δPEG_4-$(CH_2)_2$NHC(O)CH_2ONH_2 (SEQ ID NO:325),

GGGE-δPEG_4-$(CH_2)_2NHNH_2$ (SEQ ID NO:326),

GGGE-δPEG_4-$(CH_2)_2$NHC(O)-C_6H_4-CHO (SEQ ID NO:327),

GGGE-δPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:328),
GGGE-δPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:329),
GGGE-δPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:330),
GGGE-δPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:331),
GGGE-δPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:332),
GGGE-δPEG$_4$-(CH$_2$)$_2$SCOCH$_3$ (SEQ ID NO:333),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:334),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:335),
GGWYSOrn-δPEG$_6$-CH$_2$C≡CH (SEQ ID NO:336),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:337),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:338),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:339),
GGWYSOrn-δPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:340),
GGWYSOrn-δPEG$_6$-CH$_2$CH=CH$_2$ (SEQ ID NO:341),
GGWYSOrn-δPEG$_6$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:342),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:343),
GGWYSOrn-δPEG$_6$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:344),
GGGWE-δPEG$_4$-(CH$_2$)N$_3$ (SEQ ID NO:345),
GGGWE-δPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:346),
GGGWE-δPEG$_4$-CH$_2$C≡CH (SEQ ID NO:347),
GGGWE-δPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:348),
GGGWE-δPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:349),
GGGWE-δPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:350),
GGGWE-δPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:351),
GGGWE-δPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:352),
GGGWE-δPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:353),
GGGWE-δPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:354),
GGGWE-δPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:355),
GGGGK-εPEG$_8$-(CH$_2$)N$_3$ (SEQ ID NO:356),

GGGGK-εPEG$_8$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:357),

GGGGK-εPEG$_8$-CH$_2$C≡CH (SEQ ID NO:358),

GGGGK-εPEG$_8$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:359),

GGGGK-εPEG$_8$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:360),

GGGGK-εPEG$_8$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:361),

GGGGK-εPEG$_8$-(CH$_2$)$_2$SH (SEQ ID NO:362),

GGGGK-εPEG$_8$-CH$_2$CH=CH$_2$ (SEQ ID NO:363),

GGGGK-εPEG$_8$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:364),

GGGGK-εPEG$_8$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:365),

GGGGK-εPEG$_8$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:366),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:367),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$-4-(3,3-difluoro)-cyclooctyne (SEQ ID NO:368),

GGGGGWK-ε-EEEPEG$_4$-CH$_2$C≡CH (SEQ ID NO:369),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:370),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$NHNH$_2$ (SEQ ID NO:371),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$NHC(O)-C$_6$H$_4$-CHO (SEQ ID NO:372),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$SH (SEQ ID NO:373),

GGGGGWK-ε-EEEPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:374),

GGGGGWK-ε-EEEPEG$_4$-CH$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO:375),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$NHC(O)-CH(NH$_2$)CH$_2$SH (SEQ ID NO:376),

GGGGGWK-ε-EEEPEG$_4$-(CH$_2$)$_2$-Maleimide (SEQ ID NO:377),

GGGGGWDDK-ε-lipoic acid (SEQ ID NO:378),

GGDK-ε-CH(NH$_2$)CH$_2$SH (SEQ ID NO:379),

In one aspect of the present invention are compounds of formula (IV) which are a subset of compounds of formula (I):

R_1-HN-G(G)$_n$-$X_1X_2Y_1Z_1$

wherein G is a glycine residue;

R_1 is hydrogen or an amino protecting group;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

X_2 is an amino acid residue with a reactive functional group in its side chain;

Y_1 is a spacer group which is attached to X_2 through the reactive functional group in the side chain of X_2;

Z_1 is a reactive functional group; and

n is 0 or an integer of 1 to 4.

Preferred embodiments of R_1, X_1, X_2, Y_1, Z_1 and n are as described for formula (I) above.

In the compounds of formula (IV), X_2 is an amino acid group which has a reactive functional group in the side chain and Y_1 is a spacer group attached to the reactive functional group of the side chain of X_2. In some embodiments, X_2 is selected from lysine, ornithine, aspartic acid and glutamic acid.

Particular sequences include those of SEQ ID NOs:22-32, 55-98, 110-120, 143-186, 198-208, 231-274, 277-290, 295, 296, 299-377 and 379.

The compounds of formula (I) and formula (IV) may be prepared by methods known in the art. The peptide sequence $G(G)_n$-X_1X_2 may be prepared by solid phase synthetic methods using Fmoc chemistry as described in Fmoc Solid Phase Synthesis, A practical approach, edited by W. C. Chan, P.D. White, Oxford Press, 2000 or Boc chemistry as described by Schnoltzer *et al.*, Int. J. Peptide Protein Res., 40, 180 (1992). Further reactions to incorporate Y_1 and Z_1 may be performed while $G(G)_n$-X_1X_2 is still resin bound. Following deprotection and cleavage from the solid support, the peptides may be purified using high performance liquid chromatography (HPLC).

Alternatively, the peptide sequence $G(G)_n$-X_1X_2 may be prepared by recombinant DNA technology where a DNA sequence encoding the desired peptide sequence, or its

precursor, may be inserted into a suitable vector and protein expressed in a suitable expression system.

Where present, Y_1 may be incorporated into the peptide by reaction of a terminal functional group on the spacer with the C-terminal carboxylic acid of the $G(G)_n$-X_1X_2 peptide or with a functional group on the side chain of X_2 if it is present.

For example, Y_1 may include a terminal amino group which is able to react with the C-terminal carboxylic acid of G, X_1 or X_2 or a carboxy side chain group of X_2, which may be activated by formation of an activated ester by methods known in the art, to form an amide.

Another option is where Y_1 includes a terminal carboxy group, which upon activation is able to react with an amino group in the side chain of X_2 to form an amide bond.

Y_1 may be reacted with the peptide $G(G)_n$-X_1X_2 before reaction with another group to introduce Z_1 or Y_1 may be reacted with another group to introduce Z_1 before reaction with the peptide $G(G)_n$-X_1X_2.

The Z_1 group may also be introduced by methods known in the art such as amide formation, nucleophilic substitution, elimination, oxidation and reduction. For example, Z_1 may be bifunctional and include an amine functional group capable of amide formation with the C-terminal carboxy group of $G(G)_n$ or $G(G)_n$-X_1 or G(G)n-X_1X_2 or with a terminal carboxy group of Y_1. Alternatively where Z_1 includes a carboxy functional group, this may react with a side chain amine in X_2 or with a terminal amino group of Y_1.

When Z_1 is an azide, the azide may be introduced onto Y_1, X_1, or X_2 or the C-terminal carboxy group of the $G(G)_n$ sequence by reaction of an aminoalkylN_3 group with an activated carboxy group to form an amide or reaction of an activated HO_2CalkylN_3 group with a side chain amino group of X_2 or a terminal amino group of Y_1. Alternatively, the azide may be formed *in situ* if Y_1 terminates with an alkylhydroxy group by using the

method described in Tetrahedron, 2009, 65: 7329. Briefly the hydroxy group may be reacted with mesyl chloride to form a mesylate which can be reacted with sodium azide (NaN_3) to form an azide.

When Y_1 includes a terminal hydroxy group, Z_1 as a thiol group may be introduced by conversion of the terminal hydroxyl group to a thioacetate, which is a thiol precursor, by reaction of the hydroxy group with tosyl chloride to form a tosylate which is then reacted with tetrabutylammonium iodide (TBAI) and potassium thioacetate (KSAc) (WO 05/010481, p16).

When Y_1 includes a terminal amino group, Z_1 as a cysteine residue may be introduced by reacting the carboxyl moiety of a protected cysteine with the amino group of Y_1 under amide forming conditions. A suitable protected cysteine is N-(t-butoxycarbonyl)-thiazolidone-2-carboxylic acid.

When Y_1 includes a terminal carboxy group, Z_1 as a thioester may be introduced by reacting an thiol group, such as N-(methyl)mercaptoacetamide, with the activated carboxy group of Y_1.

When Y_1 includes a terminal hydroxy group, Z_1 as an alkene may be introduced by the method outlined in Angew Chem., 2008, 47, 3192, Supp. p7. Briefly, the hydroxy group is reacted with allylchloride and tetrabutylammonium sulphate in dichloromethane in the presence of aqueous NaOH.

For any of the reactions above, functional groups may need to be protected and deprotected in some cases selectively. Suitable protecting groups are known in the art and may be found in, for example, Protective Groups in Organic Synthesis, Greene and Wutz, third edition, Wiley Interscience, 1999.

The sortase mediated reaction may be performed on any scale and will depend on the amount of conjugated product required.

The reaction mixture is maintained at a temperature suitable for conjugation to occur, for example, between 15°C and 50°C, especially from 23°C to 40°C, more especially 30°C to 40°C, most especially about 37°C.

Suitable ratios of sortase enzyme to compound of formula (III) are between 1:1 to 1:1000. In some embodiments, the ratio of enzyme to compound of formula (III) is 1:1 or greater, 1:2 or greater, 1:3 or greater, 1:4 or greater, 1:5 or greater, 1:6 or greater, 1:7 or greater, 1:8 or greater, 1:9 or greater or 1:10 or greater. In some embodiments the ratio of enzyme to compound of formula (III) is about 1:1 to 1:2.

In some embodiments, the compound of formula (III) is present in an amount between 1 μM and 10 mM, especially 1 μM to 5 mM, more especially 1 μM to 1 mM, especially about 10 μM.

In some embodiments, the compound of formula (I) or formula (IV) is present in a ratio with the compound of formula (III) of between 1:10 to 10:1. In some embodiments, especially where X_3 is a purification tag, the ratio of formula (I) or formula (IV) to formula (III) is between 1:1 and 1:10, especially 1:1 to 1:5, more especially about 1:3. In some embodiments the amount of compound of formula (I) or formula (IV) is in the range of 1 μM and 10 mM, especially 10 μM to 5 mM, more especially 10 μM to 1 mM.

In some embodiments, the sortase enzyme is present in an amount of between 1 μM and 500 μM, especially about 5 μM to 200 μM, or 5 μM to 100 μM, more especially about 10 μM.

The sortase mediated reaction may be conveniently carried out in aqueous media such as water or buffer. In some embodiments, the aqueous media is a buffer that maintains a pH between 6 to 8.5, especially 6.5 to 8.5, 7 to 8.5, 7.5 to 8.5, more especially about 8. One suitable buffer is 50 mM Tris with 150 mM sodium chloride at pH 8.

The reaction mixture may also comprise calcium ions preferably in the form of calcium chloride, in an amount between 0.1 mM and 20 mM, especially 0.1 to 15 mM. In particular instances where the group A is an antibody or fragment thereof, the calcium chloride concentration is between 0.1 mM and 5 mM, especially between 0.1 and 1 mM, more especially about 0.5 mM.

The reaction time required may be monitored by chromatographic methods or by SDS-PAGE and the reaction stopped at a desired amount of conjugation product has been produced. Suitable reaction times range between 0.5 hour and 24 hours, especially 0.5 to 15 hours, 0.5 to 10 hours or 1 to 5 hours, especially 1 to 5 hours.

When fluorescent labels, photosensitizers or light sensitive molecules are present in the reaction mixture, the reaction may be carried out in the dark.

After completion of the reaction, the mixture may be purified by methods known in the art such as centrifugation followed by chromatography methods such as dialysis ultrafiltration, ultracentrifugation, size exclusion chromatography, affinity chromatography, HPLC or fast protein chromatography (FPLC).

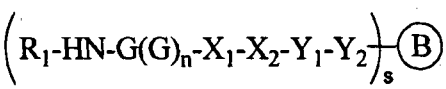
In one embodiment, the method of the invention further comprises the step of conjugating two compounds of formula (II) with one another wherein Z_1 of one compound of formula (II) is a complementary reactive functional group to Z_1 of the second compound of formula (II).

In some embodiments, one Z_1 is an azide and the other Z_1 is an alkyne and the conjugation occurs by 1,3-cycloaddition reaction further mediated by copper or by strain inherent to the alkyne. In other embodiments, one Z_1 is a hydroxylamine and the other Z_1 is an aldehyde or ketone and an oxime is formed under acid catalysis conditions. In further embodiments, one Z_1 is a hydrazine and the other Z_1 is an aldehyde or ketone and a hydrazone is formed under acid catalysis conditions. In another embodiment, one Z_1 is a carboxylic acid and the other Z_1 is an amine and an amide is formed under amide forming conditions. In yet

further embodiments, one Z_1 is a thiol group and the other Z_1 is an alkene or alkyne and the conjugation occurs by thiol-ene or hydrothiolation reaction where SH is added to the double bond either by radical mediated reaction or where the alkene is polarized by conjugation with an electron withdrawing group as in α,β-unsaturated ester, by Michael addition. In yet other embodiments, one Z_1 is a thioester and the other Z_1 is a cysteine residue or equivalent thereof and conjugation occurs by native chemical ligation.

In some embodiments, each A is the same. In other embodiments, each A is different and each A may be selected for a complementary function. For example, one A may be a small molecule drug and the other a targeting compound selected to deliver the drug to its site of action. One A may be a contrast agent and the other a targeting agent selected to deliver the contrast agent to an organ being studied. One A may be a nanoparticle or dendrimer and the other a protein, peptide, small molecule drug, contrast agent or radioisotope. In particular embodiments, one A is a protein or peptide, especially an antibody or a fragment thereof.

In yet another aspect of the invention, the compound of formula (I) may be reacted with another entity before sortase conjugation with a compound of formula (III), thereby preparing a compound of formula (V). Thus there is provided a method of preparing a compound of formula (V):



(V)

wherein G is a glycine residue;

each R_1 is independently selected from hydrogen or an amino protecting group;

each X_1 is independently absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

each X_2 is independently absent or is an amino acid residue with a reactive functional group in its side chain;

each Y_1 is independently absent or is a spacer group;

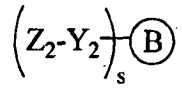
each Y_2 is independently absent or is a linker group;

B is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a polymer, a photosensitizer, DNA, RNA, a carbohydrate, a contrast agent, a catalyst, a lipid or a virus particle;

s is an integer from 1 to 100; and

n is 0 or an integer of 1 to 4;

said method comprising reacting a compound of formula (I) described above with a compound of formula (VI):



(VI)

wherein B, Y_2 and s are defined above and Z_2 is a reactive functional group complementary to Z_1.

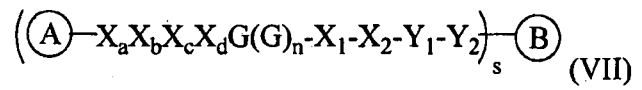
In some embodiments, Z_1 is an azide and Z_2 is an alkyne or *vice versa* and the conjugation occurs by 1,3-cycloaddition reaction further mediated by copper or by strain inherent to the alkyne. In yet other embodiments, Z_1 is a hydroxylamine and Z_2 is an aldehyde or ketone or *vice versa* and an oxime is formed under acid catalysis conditions. In further embodiments, Z_1 is a hydrazine and Z_2 is an aldehyde or ketone or *vice versa* and a hydrazone is formed under acid catalysis conditions. In another embodiment, Z_1 is a carboxylic acid and Z_2 is an amine or *vice versa* and an amide is formed under amide forming conditions. In yet further embodiments, one Z_1 is a thiol group and the other Z_2 is an alkene or alkyne or *vice versa* and the conjugation occurs by thiol-ene or hydrothiolation reaction where SH is added to the double bond either by radical mediated reaction or where the alkene is polarized by conjugation with an electron withdrawing group as in α,β-unsaturated ester, by Michael addition. In yet other embodiments, one Z_1 is a thioester and the other Z_2 is a cysteine residue or equivalent thereof or *vice versa* and conjugation occurs by native chemical ligation.

Y_2 may be absent or may be a linker that connects Z_2 to B. The spacer may be any sort of spacer and may provide relief from steric hindrance allowing better reaction between Z_1 and Z_2. In some embodiments the spacer group is 1 to 10 atoms to length and may be form example, alkylene groups ($(CH_2)_x$) in which x is an integer from 1 to 10 and in which one or more CH_2 groups may be replaced by -O-, -NH-, -NH-C(O)-, -C(O)-NH, -S- or -C_6H_4-.

In some embodiments, B may bear more than one -Y_2-Z_2 group allowing multiple compounds of formula (I) to be attached to the surface of B. In this case, s is greater than 1, for example 2 to 100 or 2 to 80 or 2 to 65 or 2 to 45 or 2 to 35 or 2 to 20, 2 to 18, 2 to 16, 2 to 14, 2 to 12, 2 to 10, 2 to 8, 2 to 6 or 2 to 4. This embodiment is particularly suitable when B is a nanoparticle, a cell or dendrimer in which the resulting dendrimer, cell or nanoparticle is decorated with multiple copies of the $G(G)_n$-motif.

In some embodiments where s is greater than 1, each R_1 may be the same or different. When more than one type of R_1 is present, a first type of R_1 may be selectively removed to allow a first subset of $G(G)_n$-moieties on B to be modified with one type of A group by sortase conjugation, removal of a second type of R_1 allows a second subset of $G(G)_n$-moieties on B to be modified with a second type of A group by sortase conjugation. This may be further repeated until all types of R_1 have been removed and all $G(G)_n$-moieties modified.

This method may then further comprise the step of sortase mediated conjugation with a compound of formula (III) to form a compound of formula (VII):



(VII)

where B, X_a, X_b, X_c, X_d, G, X_1, X_2, Y_1, Y_2 and s are defined above.

The conditions used in the sortase mediated conjugation are the same as those described above. However, the ratio of the compound of formula (III) and the compound of formula

(V) will depend on the number of $G(G)_n$-motifs that are attached to B in the compound of formula (V) (s) and which are not protected by R_1. When s is 1, the ratio of compound of formula (V) to formula (III) will be in the range of 1:10 to 10:1, especially 1:1 to 1:10, 1:1 to 1:5, more especially 1:3, particularly where X_3 is a purification tag. When s is greater than 1, the ratio of formula (V) to formula (III) will be in the order of 1:5 x s to 1:10 x s, for example 1:1 x s to 1:5 x s, especially about 1:3 x s.

In some embodiments, A and B are the same. In other embodiments, A and B are different. In some embodiments, each A is independently selected and may be the same as another or a group of other A groups or may be different from another or group of other A groups. In particular embodiments, at least one A is a protein or peptide, especially an antibody or a fragment thereof. In some embodiments, B is a protein or peptide, especially an antibody or a fragment thereof, and may be the same or different from A.

In yet another embodiment there is a kit comprising a compound of formula (I) and a sortase enzyme.

In some embodiments, the kit further comprises a compound of formula (III), especially where A in the compound of formula (III) is a protein or peptide. In some embodiments, the sortase enzyme is a Sortase A enzyme, especially a Sortase A enzyme from *S. aureus*. Optionally the kit further comprises a buffer composition suitable for use in a sortase mediated conjugation reaction.

Applications

The site-selective modifications described herein may have many applications, particularly in the pharmaceutical sciences, for example, in drug delivery, imaging, cell therapy, diagnostics, vaccination, radiotherapy, phototherapy and gene therapy.

The site-selective modifications described herein may be used to provide high purity fusion proteins, where the fusion of proteins is mediated by way of the site specific conjugation described above. For example, a fusion of two antibodies or scFvs with

different specificities. Examples include antibody-effector fusions wherein the effector may be a toxin, an interleukin (eg: IL2), a cytokine, an RNAse or perforin. Known methods of producing fusion proteins often have low yield and/or low purity. The methods of the present invention allow each protein to be prepared and modified separately and purified if required, then conjugated with high yield and purity.

The site-selective modifications described herein may be used to attach a protein, such as an antibody, with a nanoparticle such as a dendrimer, a polymer, liposomes, micelles, a microparticle of iron oxide (MPIO) or an LbL click capsule. Fusions with MPIO or a dendrimer-gadolinium complex may provide targeted delivery of the MPIO to a specific site in the body for magnetic resonance imaging (MRI). Fusions with click capsules or polymers may also be used to deliver drugs, genetic material, radiolabels or contrast agents in a targeted manner.

When a protein, such as an antibody or a fragment thereof is attached to a cell, such as a stem cell or a NK cell, the conjugate may be used to deliver the cells to a specific site in the body. For example, stem cells could be delivered to a tissue in need of regeneration or NK cells delivered to a tumour.

Another application is diagnostic or *in vitro* testing where an antibody or protein is modified in a site selective manner to facilitate attachment of the protein to a surface or a moiety that directly or indirectly provides a signal for detection. The site-specific attachment of a protein to a surface, for example, by site specific attachment of biotin and subsequent binding to an avidin coated surface, provides a means by which the function of that protein, for example, specific binding of an analyte, may be optionally preserved. The site specific attachment of a moiety that directly or indirectly provides a signal, for example, attachment of a fluorescent moiety, provides a means by which the function of that protein, for example, specific binding of an analyte, may be optionally preserved. Such components are important elements of *in vitro* diagnostic tools which are used to detect particular analytes.

The protein or antibody may be conjugated with a dendrimer or polymer for imaging or drug delivery. For example, the antibody may provide targeting to a particular site in the body and the dendrimer or polymer may carry at least one small molecule drug, photosensitizer, radiolabel or contrast agent. The conjugate may then be used to deliver the drug, photosensitizer, radiolabel or contrast agent to a specific site in the body.

Fusion proteins may be formed between recombinant protein or antibody or fragment thereof and human serum albumin (HAS) (monomeric) or human Fc (dimeric). Such fusion proteins may have stable and improved *in vivo* pharmacokinetic properties.

Some of the methods of the present invention may also allow reactions to proceed under conditions which are more amenable to drug synthesis, for example, the absence of metal ions that may be traditionally used as catalysts (eg: Cu(I) ions).

The methods of the present invention may be used to produce conjugates of virus particles carrying specific genetic material with targeting proteins so that the virus particles are delivered to the cells to which the genetic material is to be delivered. These conjugates may be particularly useful in gene therapy.

The methods of the present invention may also be used to produce conjugates of targeting proteins with enzymes. The targeting protein allows the enzyme to be attached to a cell surface and after delivery of a prodrug to the vicinity of the cells, the enzyme can act on the prodrug to produce the active drug in the site it is required to act. Following this theory, conjugates may be produced by the present invention that are suitable for antibody-directed enzyme prodrug therapy (ADEPT), gene-directed enzyme prodrug therapy (GDEPT), clostridial-directed enzyme prodrug therapy (CDEPT) and polymer-directed enzyme prodrug therapy (PDEPT) (Schellmann *et al.*, Mini-Reviews in Medicinal Chemistry, 2010, 10:887-904).

The invention will now be described with reference to the accompanying Figures and Examples. However, it is understood that the particularity of the following description is not to supersede the generality of the preceding description of the invention.

EXAMPLES

Example 1: Generation and expression of single chain anti-LIBS with C-terminal LPETG-$(His)_6$ motif (scFv(+)) and non-binding single chain mutMA2 with C-terminal LPETG-$(His)_6$ motif (scFv(-))

Step 1: Single chain anti-LIBS with C-terminal LPETG-$(His)_6$ motif

The generation of the anti-LIBS scFv from a hybridoma cell line expressing a monoclonal antibody against LIBS epitopes on GPIIb/IIIa has been described previously (Schwarz *et al., J. Pharmacol. Exp. Ther.*, **308**, 1002-1011, 2004; Stoll *et al., Arterioscler. Thromb. Vasc. Biol.*, **27**, 1206-1212, 2007). To express anti-LIBS in insect cells, the scFv was cleaved at NcoI and NotI restriction sites and sub-cloned into a pMT vector system (Invitrogen, USA). To clone LPETG motif into pMT vector for generating the anti-LIBS-LPETG (SEQ ID NO:380), we designed two complementary primers containing the LPETG-$(His)_6$ (SEQ ID NO:381) sequence flanked by 5' NotI and 3' ApaI as following: sense strand: 5'-GCGGCCGCTCTGCCGGAAACCGGCGGCGGGCCCA-3' (SEQ ID NO:381), antisense strand: 5'-GGGCCCGCCGCCGGTTTCCGGCAGAGCGGCCGCA-3' (SEQ ID NO:382) (LPETG sequence underlined). Primers were also designed with A overhangs for sub-cloning into pGEM-T Easy vector (Invitrogen, USA). The primers were annealed to a double strand product, ligated into the pGEM-T Easy vector and transformed into DH5α *E.Coli* cells (Invitrogen, USA) for amplification of the vectors. pGEM-T Easy-LPETG was then digested with EcoRI and then with NotI and ApI. Subsequently, the amplified LPETG (SEQ ID NO:380) strands were cloned into pMT-anti-LIBS at NotI and ApaI restriction sites. The resulting plasmid constructs were then transformed into TG1 *E.Coli* cells (Invitrogen, USA). The transformed cells were grown in LB media containing 100 μg/mL ampicillin and 100 mM glucose at 37°C and the plasmids were purified using Plasmid Maxi Kit (Qiagen, Australia).

Drosophila S2 cells (Invitrogen, USA) were transfected with pMT-anti-LIBS-LPETG using a method described by Han *et al.*, Nucleic Acids Res., 24, 4362-4363, 1996. Briefly, cells were diluted to 1×10^6 cells/mL and mixed with 80 ng/mL anti-LIBS-LPETG-$(His)_6$ (SEQ ID NO:280) DNA preincubated with 250 ng/mL dimethyldioctadecylammonium bromide for 30 min. The cells were then cultured in Express Five SFM medium containing 18 mM L-glutamine and 1% penicillin/streptomycin at 28°C in ventilated polycarbonate Erlenmeyer flasks (Corning, Acton, MA, USA) under constant rotation (100 rpm, Bench top Orbital Shaker Incubator, Ratek Instruments, Australia). After two days, protein production was induced by 500 µM $CuSO_4$. Six days later, the cell supernatant was collected by centrifugation at 15,000 g for 15 min. The cell supernatant was applied to a chelating Sepharose fast flow column (20 mL bed volume, 5 mL/min flow rate, GE Healthcare, Uppsala, Sweden). The column was washed to baseline with PBS, 0.5 M NaCl and 10 mM imidazole in 50 mM Tris, pH 8.0 to remove non-specifically bound proteins. Elution was carried out with 250 mM imidazole in 50 mM Tris, pH 8.0. Fractions of 50 mL were collected. Fractions containing significant amounts of product were pooled and dialyzed against PBS. A second purification was done with a nickel-based metal affinity chromatography column, Ni-NTA column (Invitrogen), according to the manufacturer's instruction manual. Fractions of 2 mL were collected and dialysed against PBS.

Step 2: Non-binding single chain mutMA2 with C-terminal LPETG-$(His)_6$ (SEQ ID NO:384)motif (scFv(-))

The non-binding scFv mutMA2 is based on the blocking antibody MA2 as described in Schwarz, FASEB, 2004, 18:1704-1706 and Schwarz, Circ. Res., 2006, 99:25-33.

The sequence of the mutMA2 with LPETG tag with a CDR3 mutation of the RND motif responsible for the binding to GPIIbIIIa was de novo synthesized by Geneart (Regensburg).

DNA (SEQ ID NO:383)

ATG GAG ACA GAC ACA CTC CTG CTA TGG GTA CTG CTG CTC TGG GTT CCA GGT TCC ACT GGT GAC GCG GCC CAG CCG GCC AGG CGC GCC GTA CGA AGC TTG GTA CCG AGC TCG GAT CCA CTC CAG TGT GGT GGA ATT CTC GCG

GAA GTG CAG CTG GTG CAG TCT GGA GCT GAG GTG AAT AAG CCT GGG
GCC TCA GTG AAG GTC TCC TGC AAG GCT TCT GGA TAC ACC TTC ACC GGC
TAC TAT ATG CAC TGG GTG CGA CAG GCC CCT GGA CAA GGG CTT GAG
TGG ATG GGA TGG ATC AAC CCT AAC AGT GGT GGC ACA AAC TAT GCA
CAG AAG TTT CAG GGC TGG GTC ACC ATG ACC AGG GAC ACG TCC ATC
AGC ACC GCC TAC ATG GAG CTG AGC AGG CTG AGA TCT GAC GAC ACG
GCC GTG TAT TAC TGT GCG AGA GGC CGT GCT TTG TAT AAC GCG AAC GAC
CGG TCC CCC AAC TGG TTC GAC CCC TGG GGC CAG GGA ACC CTG GTC ACC
GTC TCC TCA GGG AGT GCA TCC GCC CCA ACC CTT AAG CTT GAA GAA GGT
GAA TTT TCA GAA GCA CGC GTA CAG GCT GTG CTG ACT CAG CCG CCC TCG
GTG TCA GTG GCC CCA GGA CAG ACG GCC AGG ATT ACC TGT GGG GGA
AAC AAC ATT GGA AGT AAA AGT GTG CAG TGG TAC CAG CAG AAG CCA
GGC CAG GCC CCT GTG CTG GTC GTC TAT GAT GAT AGC GAC CGG CCC TCA
GGG ATC CCT GAG CGA TTC TCT GGC TCC AAC TCT GGG AAC ATG GCC ACC
CTG ACC ATC AGC AGG GTC GAA GCC GGG GAT GAG GCC GAC TAT TAC
TGT CAG GTG TGG GAT AGT AGT AGT GAT CAT GTG GTA TTC GGC GGA
GGG ACC AAG CTG ACC GTC CTA GGT GGG AAA CCC ATT CCT AAC CCA
CTG CTG GGA CTG GAT AGC ACA CTG CCT GAG ACT GGC GGG CTG GAA
GAG GCG GCC GCT CGA GGA GGG CCC GAA CAA AAA CTC ATC TCA GAA
GAG GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA

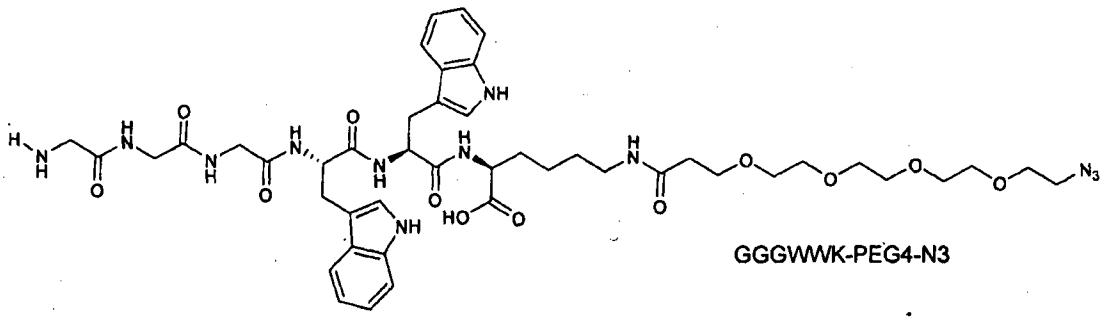
Amino Acid Sequence (SEQ ID NO:384)

G D P S W L A T M E T D T L L L W V L L L W V P G S T G D A A Q P A R R A V R
S L V P S S D P L Q C G G I L A E V Q L V Q S G A E V N K P G A S V K V S C K A
S G Y T F T G Y Y M H W V R Q A P G Q G L E W M G W I N P N S G G T N Y A Q
K F Q G W V T M T R D T S I S T A Y M E L S R L R S D D T A V Y Y C A R G R A
L Y N A N D R S P N W F D P W G Q G T L V T V S S G S A S A P T L K L E E G E F
S E A R V Q A V L T Q P P S V S V A P G Q T A R I T C G G N N I G S K S V Q W Y
Q Q K P G Q A P V L V V Y D D S D R P S G I P E R F S G S N S G N M A T L T I S R
V E A G D E A D Y Y C Q V W D S S S D H V V F G G G T K L T V L G G K P I P N P

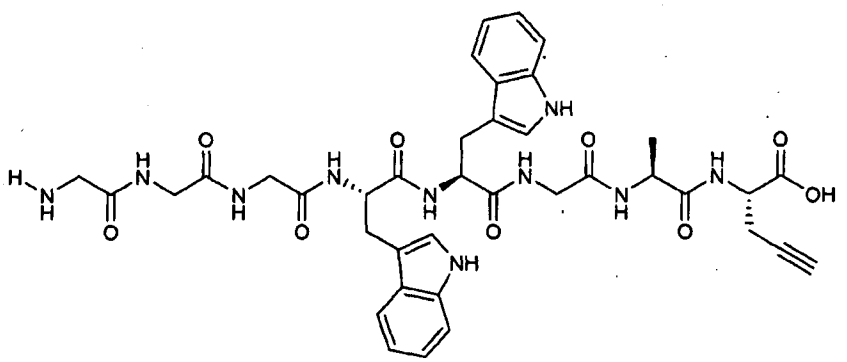
LLGLDSTLPETGGLEEAAARGGPEQKLISEEDLNSAVDHH HHHH

The scFv was cloned in the expression vector pET20b (catalogue 69739, Merck) and transformed into *Escherichia coli* BL21(DE3) Competent Cells (catalogue 69450, Merck) Bacteria were induced with 0.25 mmol/L isopropyl-D-galactoside (Sigma) and incubated for 16 hours at 200 rpm and 23°C. The bacteria were lysed with an ice-cold hyperosmotic shock solution (20% sucrose, EDTA, Tris), and scFvs were purified by FPLC-procedure, using metal-affinity chromatography (Ni-NTA column, Superflow Columns, Qiagen catalogue 30622). Elution was done with Elution Buffer (50mM NaH_2PO_4, 300mM NaCl, 250mM Imidazol) Fractions with highest protein content were pooled and dialysed against PBS.

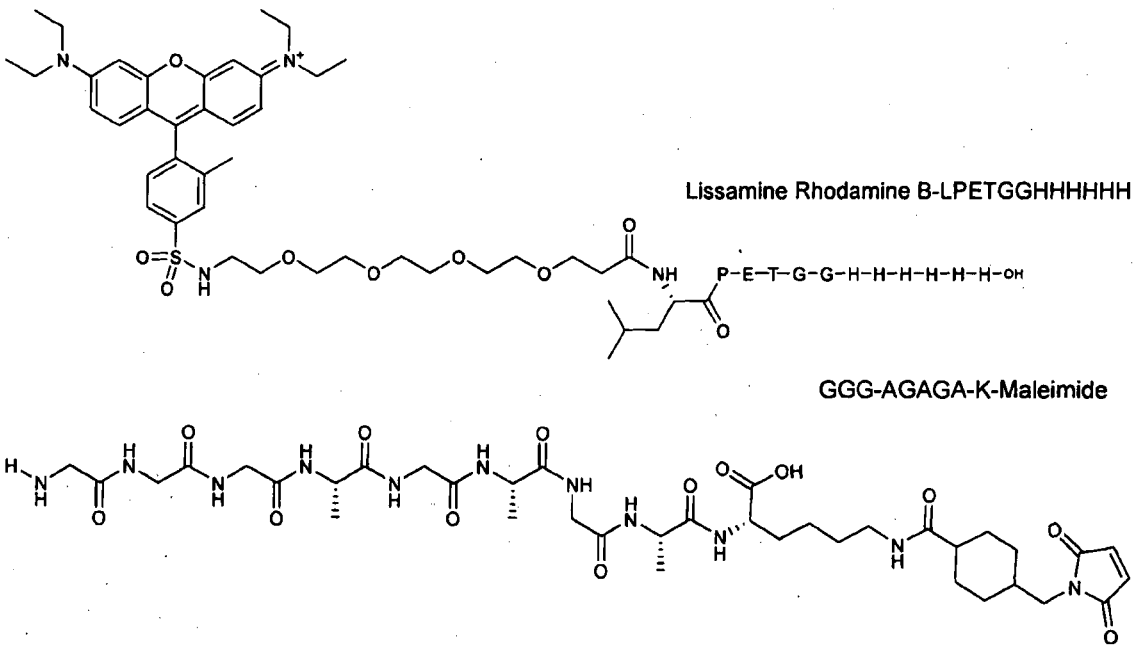
Example 2: Preparation of GGG-R compounds



GGGWWK-PEG4-N3



GGG-WW-GA-pG



Peptides were supplied by various commercial peptide manufacturers and were prepared using peptide synthesis grade reagents including Fmoc-protected amino acids for example Fmoc-propargyl glycine (pG). In a representative example, the sequences were assembled on a Symphony synthesizer (Protein Technologies Inc., USA) using O-Benzotriazol-1-yl-N,N,N',N'-tetramethyluroiumhexafluorophosphate (HBTU) diisopropylethylamine (DIEA) activation in N,N-Dimethylformamide (DMF) with piperidine deprotection of the Fmoc groups. CO-$[(CH_2)_2O]_4(CH_2)_2N_3$ was added manually to sequences as the last coupling using Azido-dPEG®4-NHS ester (Item #:10501 Quanta Biodesign Ohio USA) activated with DIEA in DMF. In a similar manner, Lissamine Rhodamine B-NH-PEG_4-CO was added manually to the sequence as the last coupling following piperidine deprotection of the terminal Fmoc NH protecting group, using Lissamine Rhodamine B sulphonamide-$dPEG_4$-acid (item #10229) activated by O-Benzotriazol-1-yl-N,N,N',N'-tetramethyluroiumhexafluorophosphate (HBTU) and diisopropylethylamine (DIEA) in NN-Dimethylformamide (DMF).

The crude peptides were cleaved from the resin support using a mixture of Trifluoroacetic acid (TFA), water, dithiothreitol (DTT) and Triisopropylsilane (TIPS) (ratio 92.5:2.5:2.5:2.5 respectively). Crude cyclic peptides were then purified by RP-HPLC

(Waters, USA) using C-18 columns with TFA and aqueous acetonitrile buffers and the sequences confirmed by mass spectrometry (ESI-MS) and amino acid analysis. The peptides prepared were:

Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385)

GGGWWSSK-εPEG$_4$-N$_3$ (SEQ ID NO:77)

GGGWWK-εPEG$_4$-N$_3$ (SEQ ID NO:55)

GGGK-εPEG$_4$-N$_3$ (SEQ ID NO:22)

GGG-WW-GA-pG (SEQ ID NO:386)

GGG-AGAGAK-Maleimide (SEQ ID NO:387)

GGG-eGFP was used as a comparative example and produced by the following method:

Generation, expression and production of GGG-eGFP

The following primers were designed to introduce the NdeI and XhoI restriction sites and GGG Sortase ligand into pEGFP-C1 vector (Clontech, USA): sense strand 5'-CATATG**GGAGGCGGCGGTTCA**ATGGTGAGCAAGGGCGAG-3' (SEQ ID NO:388), antisense strand 5'-CTCGAGCTTGTACAGCTCGTCCATG-3' (SEQ ID NO:389) (GGGWW in bold; NdeI and XhoI underlined, respectively). The amplification of GGG-eGFP sequence was performed by PCR using these primers. The PCR products were then cloned into a pET-20b(+) vector system at the NdeI and XhoI restriction sites. Amplification of the plasmids was done using XL1-B *E.Coli* cells (Invitrogen, USA) and the plasmid purification was performed using a Plasmid Mini-Prep Kit (Qiagen, Australia). GGG-eGFP was expressed in BL21-DE3 *E.Coli* (Invitrogen, USA). The cells were cultured in LB media containing 100 μg/mL ampicillin until the OD600 of 0.8 was reached. GGG-eGFP production was induced with 1 mmol/L of isopropyl β-D-1-thiogalactopyranoside (IPTG) for 4 hours at 37°C. Bacteria were then isolated by centrifugation at 4000 rpm for 10 min. Proteins were purified using Ni-NTA column (Invitrogen), according to the manufacturer's protocol. Fractions of 2 mL were collected and dialysed against PBS.

Example 3: Sortase A mediated reactions

ScFv mut-MA2 with LPETG tag (SEQ ID NO:384) (30 uM) prepared in Example 1 was incubated with the different GGG substrates prepared in Example 2 (GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77), GGGWWK-PEG_4-N_3 (SEQ ID NO:55), GGGK-PEG_4-N_3, (SEQ ID NO:22) GGGWWGA-pG (SEQ ID NO:386) and GGG-eGFP) (20 uM) in the presence of Sortase A (10 uM) and $CaCl_2$ (0.5 μM) at 37 °C for 3 h with gentle shaking. The reaction was performed in Sortase coupling buffer (50 mmol/L Tris, 150 mmol/L NaCl, pH 8.0).

The sortase conjugation product was treated in one of two different ways:

A. When reaction was complete, the reaction mixture was passed down a Ni-NTA column (Superflow Columns, Qiagen catalogue 30622) and flushed with PBS (Phosphate Buffered Saline, 8 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, 0.24 g KH_2PO_4 in 1L, pH 7.6). Unreacted scFv-mutMA2-LPETGH_6 and the sortase enzyme was retained and the yield of sortase-reacted scFv-mutMA2 in the reaction was determined as total protein recovered in the flow through using the BCA assay protocol (BCA Protein Assay Kit, Pierce catalogue 23225). This was expressed as a percentage of scFv-mutMA2-LPETGH_6 (SEQ ID NO:384) in Figure 1.

B. When reaction was complete, 200 μl Ni-NTA resin (Superflow resin, Qiagen catalogue 30410) per 1 mL reaction mixture was added and reaction was agitated at 4 °C for 1 hour. The reaction was centrifuged at 100g and the supernatant isolated. Unreacted scFv-mutMA2-LPETGH_6 and Sortase was retained by the resin and the yield of sortase-reacted scFv-mutMA2 in the reaction was determined as total protein recovered in the supernatant using the BCA assay protocol as for A. This was expressed as a percentage of scFv-mutMA2-LPETGH_6 (SEQ ID NO:384) in Figure 1.

Results

The Sortase A mediated conjugation of the scFv antibody-LPETG (SEQ ID NO:384) of Example 1 with the GGGR compounds of Example 2 provided scFv-antibody-LPETGGG-R conjugates with varying efficacy as shown in Figure 1.

As can be seen in Figure 1, the reaction of a GGG-tagged protein (GGGeGFP) proceeded with an efficacy of about 50% as did the reaction with GGGKPEG_4-N_3 (SEQ ID NO:22).

The sortase mediated reactions of the present invention are at least as efficacious as those of the prior art.

However, use of GGGWW (SEQ ID NO:390) motif improved efficacy to greater than 60% and addition of a $KPEG_4$ motif improved efficacy of the sortase reaction to at least 80%.

Example 4: Sortase coupling between scFv-LPETG and GGG-modified $(PVPON_{Alk})_5/PEG_{Alk}$ capsules with DL800 Label

The assembly of alkyne functionalized layer-by-layer (LbL) click capsules is described in Leung et al. (Assembly and Degradation of Low-Fouling Click-Functionalized Poly(ethylene glycol)-Based Multilayer Films and Capsules. Leung MKM, Such GK, Johnston APR, Biswas DP, Zhu Z, Yan Y, Lutz JF, Caruso F. Small. 2011 Mar 21. doi: 10.1002/smll.201002258. [Epub ahead of print]) and Kinnane et al, (Low-Fouling Poly(N-vinyl pyrrolidone) Capsules with Engineered Degradable Properties. Kinnane CR, Such GK, Antequera-Garcia G, Yan Y, Dodds SJ, Liz-Marzan LM, Caruso F, Biomacromolecules 2009, 10, 2839).

General

High-purity (Milli-Q) water with a resistivity greater than 18 MΩ cm was obtained from an in-line Millipore RiOs/Origin water purification system. Poly(methacrylic acid) (PMA, 30 wt %, Mw ~ 15000) was purchased from Polysciences (U.S.A.) and used as received.

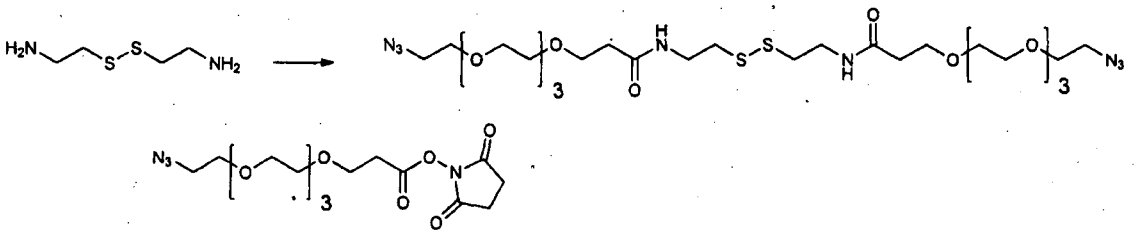
SiO_2 particles (3.25 μm diameter) were purchased from MicroParticles GmbH as a 5 wt % suspension.

Oligo-(ethylene glycol) methyl ether methacrylate (OEGMA, Mw = 475 Da, mean degree of polymerization, DP, is 8 to 9), 2-(2-methoxyethoxy)ethyl methacrylate (MEO_2MA), O-(2-Aminoethyl)-O'-(2-azidoethyl)pentaethylene glycol (azido-PEG-amine, Mw = 350.41 Da) and methyl 2-bromo propionate (MBP) were obtained from Aldrich. OEGMA and MEO_2MA monomers were purified by being passed through a neutral alumina column to remove the inhibitor. The reversible addition-fragmentation chain transfer (RAFT) agent 2-

cyanoprop-2-yl dithiobenzoate (CPDB) was prepared using literature procedures. (Macromol. Rapid Commun., 2006, 27:821) 2,2'-Azobisisobutyronitrile (AIBN) was recrystallized from 95% ethanol.

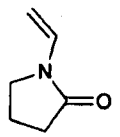
Solutions were adjusted to the required pH using a pH meter (Mettler Toledo) with 0.1 M HCl or 0.1 M NaOH. Unless noted otherwise, all chemicals were purchased from Sigma-Aldrich and used without further purification.

Step 1. Synthesis of Linker

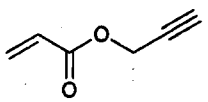


A solution of cystamine dihydrochloride (58 mg, 0.25 mmol) and triethylamine (103 mg, 1.02 mmol) in methanol (2 mL) was added to a solution of azido-dPEG$_4$-NHS ester (200 mg, 0.51 mmol, product #10501 Quanta Biodesign Ohio USA) in dioxane (2 mL). The resulting reaction mixture was stirred at room temperature for 12 h. The white precipitate formed was filtered off and the filtrate was concentrated to dryness. The residue was purified by column chromatography over a silica gel (49:50:1 hexane/diethyl ether/ethanol) to give the desired product [(*N*,*N*'-(dithiodiethane-2,1-diyl)bis(1-azido tertaethyleneglycol acetamide), 144 mg, 10.37 mmol, 73% yield] as a pale yellow oil. This product is herein denoted as the bisazide linker. NMR spectra were recorded on a Varian Gemini 2000 spectrometer operating at 400 MHz for ^{1}H NMR and 100 MHz for ^{13}C NMR. Chemical shifts are reported in ppm: ^{13}C NMR (100 MHz, D_2O): 176.7 (CO), 69.7, 69.4, 66.9, 66.1, 50.3, 46.8, 38.2 ($CONHCH_2$-), 36.8, 36,2, 25.3 (-S-CH_2-), 14.2, 8.4.

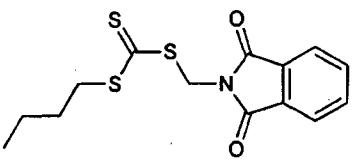
Step 2. Synthesis of Poly(N-vinyl pyrrolidone-ran-propargyl acrylate) or $PVPON_{Alk}$



vinylpyrrolidone



propargyl acrylate



RAFT initiator

Poly(*N*-vinyl pyrrolidone-*ran*-propargyl acrylate) was synthesized using the following procedure: propargyl acrylate (7.0 mg, 6.4 × 10-2 mmol), *N*-vinyl pyrrolidone (0.98 g, 8.9 mmol), the phthalimidomethyl xanthate reversible addition-fragmentation chain transfer (RAFT) agent (12.8 mg, 4.6 × 10-2 mmol), and azobisisobutyronitrile AIBN (0.4 mg, 2.4 × 10-3 mmol) were added to a vial. The synthesis of the RAFT initiator used has been described previously (Postma, A.; Davis, T. P.; Evans, R. A.; Li, G. X.; Moad, G.; O'Shea, M. S. *Macromolecules* 2006, *39*, 5293–5306).

These components were then dissolved in 3 mL of dioxane and transferred to a Schlenk tube. The reaction was then purged of oxygen using four freeze-thaw cycles on a Schlenk line. The Schlenk tube was then sealed and the polymerization was conducted in a constant temperature oil bath at 60 °C for 5 h. After polymerization, the reaction was dialyzed in water for several days to remove excess monomer and then the final product was freeze-dried. Gel permeation chromatography showed that the *M*n of the polymer was approximately 19000 with a polydispersity of 1.38.

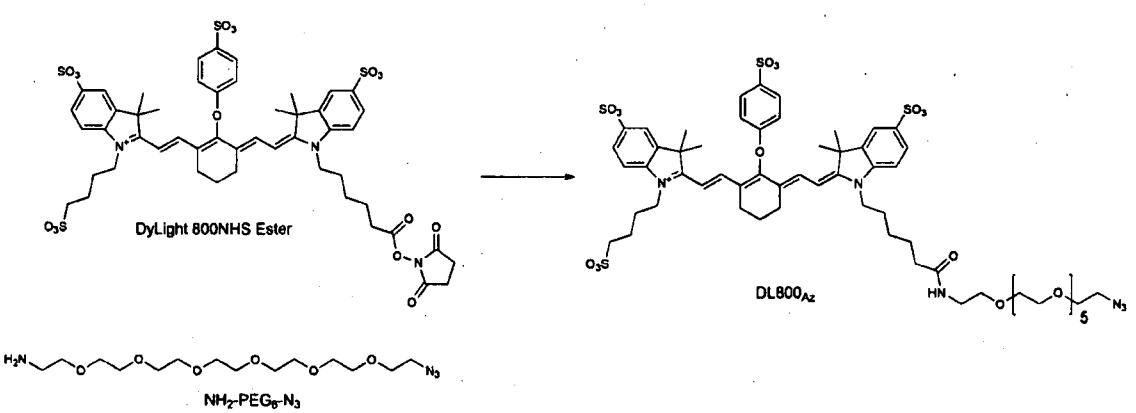
Step 3. Synthesis of PVPON/(PMA/$PVPON_{Alk}$)$_5$/PMA core-shell particles

Briefly, silica particles (~3 μm diameter, 5 wt% suspension, 100 μL) were first washed with sodium acetate (NaOAc) solution (50 mM, pH 4) by centrifugation (1000 *g*, 1 min) and redispersed in 200 μL of NaOAc solution (50 mM, pH 4). A total of three centrifugation/redispersion cycles were conducted. This procedure was used as the standard washing procedure. To form the multilayers, an equal volume of poly(*N*-vinyl pyrrolidone) (PVPON, *M*w ~ 55 000, 1 g L^{-1} in 50 mM NaOAc, pH 4) was added to the particle suspension for adsorption (15 min) with constant shaking. The particles were then washed via three centrifugation (1000 *g*, 1 min)/redispersion (200 μL) cycles. After washing, 200 μL of poly(methacrylic acid) (PMA, 1 g L^{-1} in 50 mM NaOAc, pH 4) was

adsorbed with constant shaking for 15 min, and the particles were then washed again. Alkyne-functionalized poly(*N*-vinyl pyrrolidone) ($PVPON_{Alk}$, 1 g L^{-1} in 50 mM NaOAc, pH 4) was adsorbed onto the PMA layer (using the same process as described for the previous two layers) to create a $PMA/PVPON_{Alk}$ bilayer.

This $PMA/PVPON_{Alk}$ adsorption procedure was repeated to form $PVPON/(PMA/PVPON_{Alk})_5/PMA$ core-shell particles.

Step 4. Synthesis of $DL800_{Az}$



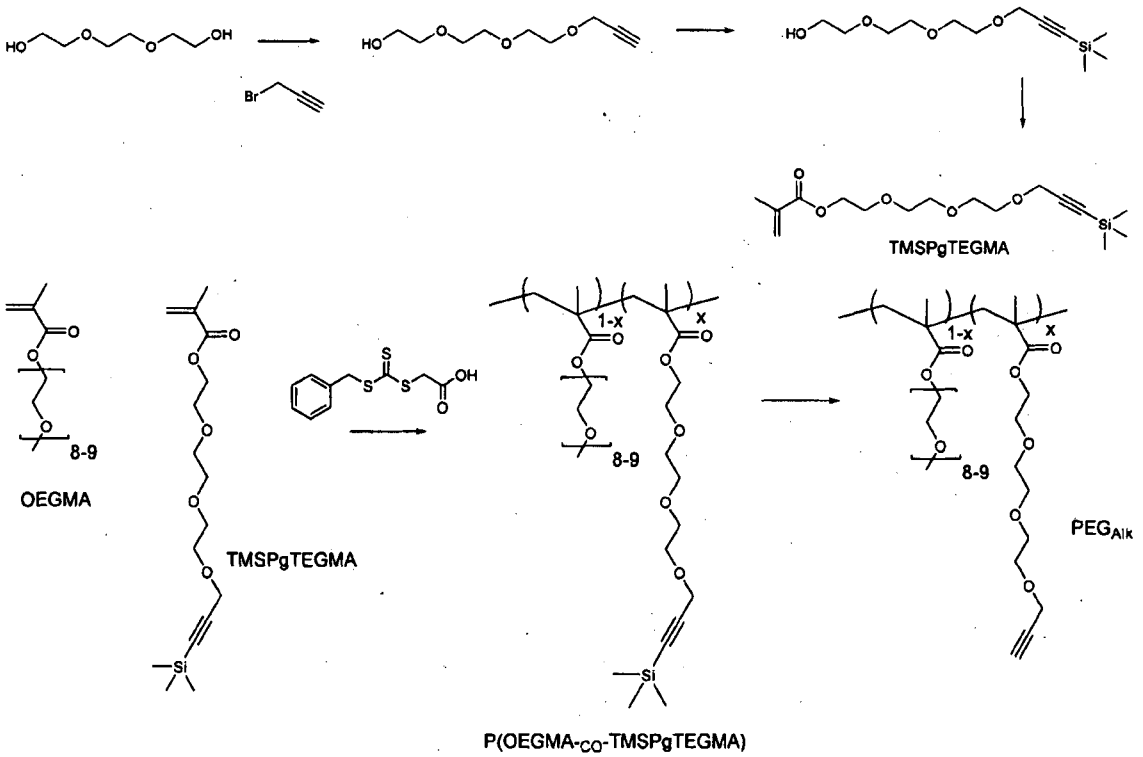
DyLight 800 NHS-Ester (Thermo Scientific) (0.2 mg) was dissolved in DMSO (20 µL) and mixed with *O*-(2-Aminoethyl)-*O'*-(2-azidoethyl)pentaethylene glycol (Aldrich) (2 µL) for 4 h to give DL800-PEG-azido ($DL800_{Az}$). This solution was used without further purification.

Step 5. Synthesis of $PVPON/(PMA/PVPON_{Alk})_5/PMA$ core-shell particles with $DL800_{Az}$ Label

For targeting studies, the particles were labeled with $DL800_{Az}$ for near infrared analysis after the completion of the assembly of partially assembled, $PVPON/(PMA/PVPON_{Alk})_2/PMA$, being an intermediate stage in the synthesis of $PVPON/(PMA/PVPON_{Alk})_5/PMA$ particles. Without purification, 0.5 µL of the $DL800_{Az}$ suspension together with sodium ascorbate (4.4 g L^{-1}, 50 µL) and copper sulfate (1.75 g L^{-1}, 50 µL) (all in 50 mM NaOAc, pH 4) were added to the particles (in 150 µL of 50 mM

NaOAc, pH 4) and labeling was allowed to proceed for 1 h. Following extensive washing, three bilayers of $PVPON_{Alk}$/PMA were assembled on the $DL800_{Az}$-labeled particles as described in Step 3.

Step 6. Reaction of $DL800_{Az}$-labeled PVPON/(PMA/$PVPON_{Alk}$)$_5$/PMA core-shell particles with PEG_{Alk}.



Step 6a. Synthesis of 2-(2-(2-(3-(Trimethylsilyl)prop-2-ynyloxy)ethoxy)ethoxy)ethyl Methacrylate (TMSPgTEGMA) **:**

Sodium hydride (2.0 g, 50 mmol, 60% dispersion in mineral oil) was slowly added to a THF (50 mL) solution of triethylene glycol (11.6 g, 76 mmol) at 0 °C. The mixture was stirred for 20 min before propargyl bromide (4.2 mL, 38 mmol) was added dropwise. The mixture was then stirred at room temperature for 20 h before THF was removed in vacuum. The residual was dissolved in DCM and then washed successively with saturated sodium hydrogen carbonate ($NaHCO_3$) (2 × 50 mL) and water (50 mL). The extracts were dried over magnesium sulfate ($MgSO_4$), filtered and concentrated in vacuum. The crude product was purified by column chromatography, eluting with a 2:3 mixture of *n*-hexane

and ethyl acetate (EtOAc) to give triethylene glycol alkyne 2-(2-(2-(prop-2-ynyloxy)ethoxy)ethoxy)ethanol as a light yellow oil (4.5 g, 63%). ^{1}H NMR (400 MHz, $CDCl_3$, tetramethylsilane (TMS)): δ H 4.13 (*s*, 2H, $OCH_2C{\equiv}CH$), 3.61–3.58 (*m*, 10H, CH_2O), 3.50 (*t*, 2H, $HOCH_2$), 2.75 (*br*, 1H, **HO**), 2.38 (*s*, 1H, $CH{\equiv}C$) ppm; ^{13}C NMR (100 MHz, $CDCl_3$, TMS) δ C 77.2 ($C{\equiv}CH$), 75.6 ($C{\equiv}CH$), 70.4 (CH_2O), 70.3 (CH_2O), 70.1 (CH_2O), 69.6 (CH_2O), 69.2 (CH_2O), 61.5 ($HOCH_2$), 60.1 ($OCH_2C{\equiv}CH$) ppm.

Protection of the triethylene glycol alkyne (3.39 g, 18 mmol) was achieved by mixing with silver chloride (0.24 g, 1.8 mmol) suspended in anhydrous DCM (25 mL), followed by addition of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (3.5 g, 23 mmol). The reaction mixture was heated to 40°C and chlorotrimethylsilane (2.8 g, 26 mmol) was added dropwise. After stirring for 24 h at 40°C the mixture was cooled to room temperature and diluted with 200 mL of *n* -hexane. The organic phase was washed successively with $NaHCO_3$ (2 × 50 mL), 0.1 M HCl (2 × 50 mL) and water (50 mL). The organic phase was dried over $MgSO_4$, filtered, and concentrated in vacuum. The crude product was purified by column chromatography, eluting with a 8:1 mixture of *n*-hexane and EtOAc to give TMS-alkyne triethylene glycol 2-(2-(2-(3-(trimethylsilyl)-prop-2-ynyloxy)ethoxy)-ethoxy)ethanol as a light yellow colorless liquid (2.11 g, 45%). ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ H 4.13 (*s*, 2H, $OCH_2C{\equiv}CH$), 3.61–3.58 (*m*, 10H, CH_2O), 3.50 (*t*, 2H, $HOCH_2$), 2.15 (*br*, 1H, **HO**), 0.21 (*s*, 9H, $Si(CH_3)_3$) ppm; ^{13}C NMR (100 MHz, $CDCl_3$, TMS) δ C 105.5 ($CH_2C{\equiv}C$), 91.2 ($C{\equiv}CSi$), 70.8 (CH_2O), 70.7 (CH_2O), 70.3 (CH_2O), 69.1 (CH_2O), 69.0 (CH_2O), 61.1 ($HOCH_2$), 60.0 ($OCH_2C{\equiv}CH$), 5.2 ($SiCH_3$) ppm.

Methacryloyl chloride (0.78 mL, 8 mmol) dissolved in DCM (10 mL) was added dropwise over 1 h to a mixture of the TMS-alkyne triethylene glycol (2.0 g, 7.7 mmol) and triethylamine (0.81 g, 8 mmol) in DCM (50 mL) at 0°C. The mixture was then stirred at room temperature for 18 h and filtered to remove triethylamine hydrochloride. The filtrate was washed with saturated $NaHCO_3$ (2 × 50 mL) and water (50 mL). The organic phase was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by column chromatography, eluting DCM to give 2-(2-(2-(3-(trimethylsilyl)prop-2-ynyloxy)ethoxy)-ethoxy)ethyl methacrylate (TMSPgTEGMA) as a

light yellow colorless liquid (2.1 g, 88%). ^{1}H NMR (400 MHz, $CDCl_3$, TMS): δ H 6.15 (*s*, 1H, CH_2=C(CH_3)), 5.51 (*s*, 2H, CH_2=C(CH_3)), 4.13 (*s*, 2H, $OCH_2C{\equiv}CH$), 3.61–3.58 (*m*, 10H, CH_2O), 3.50 (*t*, 2H, $HOCH_2$) , 2.0 (*s*, 3H, CH_2=C(CH_3)), 0.21 (*s*, 9H, $Si(CH_3)_3$) ppm; ^{13}C NMR (100 MHz, $CDCl_3$, TMS) δ C 168.3 (CO), 133.8 (=CCH_3), 131.5 (=CH_2), 106.1 ($CH_2C{\equiv}C$), 91.6 ($C{\equiv}CSi$), 70.3 (CH_2O), 70.2 (CH_2O), 70.0 (CH_2O), 69.4 (CH_2O), 69.2 (CH_2O), 61.8 ($HOCH_2$), 60.2 ($OCH_2C{\equiv}CH$), 18.8 (=CCH_3), 4.1 ($SiCH_3$) ppm.

Step 6b. Synthesis of Alkyne-Functionalized Copolymers Poly(OEGMA-co-PgTEGMA) as PEG_{Alk}

Poly(OEGMA-*co*-TMSPgTEGMA) was prepared by RAFT polymerization. OEGMA (1.0 g, 2.1 mmol), TMSPgTEGMA (70 mg, 0.21 mmol), 2-cyanoprop-2-yl dithiobenzoate (CPDB) (prepared as per *Macromol. Rapid Commun.* **2006**, *27*, 821) (7.7 mg, 0.035 mmol), and AIBN (0.7 mg, 0.004 mmol) were dissolved in dioxane (2 mL). The mixture was degassed through freeze-pump-thaw (3×) and then reacted at 70°C for 20 h. The mixture was quenched in liquid nitrogen and diluted with THF before it was precipitated into diethyl ether. This precipitation procedure was repeated three times. The resulting product was dried in vacuum to yield poly(OEGMA-*co*-TMSPgTEGMA) as a red viscous solid, 0.8 g (74%). GPC-MALLS (THF): M_n = 34 kDa, M_w/M_n = 1.13. The degree of polymerization (DP) of the copolymer was determined to be 85 and the alkyne ratio in the copolymer was 5% from ^{1}H NMR (ratio of peak at 4.1 ppm (2H) to 0.2 ppm (9H).).

Poly(OEGMA-*co*-TMSPgTEGMA) (0.5 g, 0.43 mmol based on the alkyne-trimethylsiyl groups) and acetic acid (37 μL, 0.65 mmol) were dissolved in THF (10 mL). Argon was bubbled through the solution for 10 min at 0°C. A 0.2 M solution of tetra-*n*-butyl ammonium fluoride in THF (3.2 mL, 0.65 mmol) was added slowly with a syringe with vigorous stirring. The resulting solution was stirred at room temperature for 12 h and then passed through a neutral alumina column. The solution was then concentrated in vacuum and precipitated into *n*-hexane. The isolated product was dried in vacuum to yield poly(OEGMA-*co*-PgTEGMA), herein denoted as PEG_{Alk}, as a light red viscous solid (0.40 g, 80%) with ^{1}H NMR analysis confirming 100% removal of the TMS-protecting groups.

Step 6c. Reaction of $DL800_{Az}$-labeled PVPON/(PMA/$PVPON_{Alk}$)$_5$/PMA core-shell particles with PEG_{Alk}

To achieve PEGylation, PEG_{Alk} (1 g L^{-1} in 150 mM NaOAc, pH 5) was deposited onto the *$DL800_{Az}$-labeled* PVPON/(PMA/$PVPON_{Alk}$)$_5$/PMA particles according to the procedure described in Step 3 giving PVPON/(PMA/$PVPON_{Alk}$)$_5$/PMA/PEG_{Alk} core-shell particles. The particles (in 50 µL of 150 mM NaOAc, pH 5) were then incubated with the bisazide linker from Step 1 (1 g L^{-1}, 150 µL) in the presence of sodium ascorbate (4.4 g L^{-1}, 50 µL) and copper sulfate (1.75 g L^{-1}, 50 µL) (all in 150 mM NaOAc, pH 5) overnight for crosslinking.

Step 7. Reaction of GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77) with $DL800_{Az}$-labeled PVPON/(PMA/$PVPON_{Alk}$)$_5$/PMA/PEG_{Alk} core-shell particles.

To a solution of the *$DL800_{Az}$-labeled* crosslinked particles from Step 6 (in 50 µL of 150 mM NaOAc, pH 5) was added a mixture of GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77) from Example 2 (0.2 g L^{-1}, 150 µL; i.e., final concentration = 0.1 g L^{-1}), sodium ascorbate (4.4 g L^{-1}, 50 µL) and copper sulfate (1.75 g L^{-1}, 50 µL) (all in 150 mM NaOAc, pH 5) for 30 min incubation. The GGGWWSSK-PEG_4-N_3-functionalized particles were then redispersed in 200 µL of NaOAc (50 mM, pH 4) and five alternating layers of PMA/PVPON, i.e., (PMA/PVPON)$_2$PMA, (all at 1 g L^{-1} in 50 mM NaOAc, pH 4) were deposited as protective capping layers using the procedure described in Step 3.

Step 8. Formation of $DL800_{Az}$-labeled GGG modified ($PVPON_{Alk}$)$_5$/PEG_{Alk} capsules I.

Following washing, the protected particles ($DL800_{Az}$-labeled) prepared in Step 7 were redispersed in 200 µL of NaOAc (50 mM, pH 4). To this suspension were added 200 µL of ammonium fluoride (NH_4F, 8 M) and 100 µL of hydrofluoric acid (HF, 2 M) for core dissolution, followed by three centrifugation (1500 *g*, 4 min)/redispersion (200 µL) cycles. *Caution! Hydrofluoric acid and ammonium fluoride are very toxic. Extreme care should be taken when handling HF solution, and only small quantities should be prepared.* The capsules were redispersed in 100 µL of PBS for 12 h to remove PVPON and PMA, and GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77)-functionalized ($PVPON_{Alk}$)$_5$/PEG_{Alk} capsules were obtained.

Step 9. Sortase coupling between $DL800_{Az}$-labeled GGG modified $(PVPON_{Alk})_5/PEG_{Alk}$ capsules and scFv-LPTEG (SEQ ID NO:282)

The GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77)-functionalized capsules were incubated in 300 μL of sortase coupling buffer (50 mM Tris, 150 mM NaCl, pH 8) with Sortase A, scFv-LPETG from Example 1 and $CaCl_2$. Sortase A, scFv-LPETG and $CaCl_2$ were added to a final concentration of 0.1 g L^{-1}, 0.1 g L^{-1} and 0.5 mM, respectively. The mixture was incubated at 37°C for 1 h with gentle shaking. After conjugation, the scFv-LPETG-functionalized capsules were washed with PBS to remove any unattached antibody. This ligation procedure was applied to immobilize both anti-GPIIb/IIIa scFv (denoted as scFv(+)) and the mutated scFv (denoted as scFv(-)).

Example 5: Sortase-mediated conjugation of Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:386) with GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77) modified $(PVPON_{Alk})_5$ core-shell particles

Step 1. Preparation of $(PVPON_{Alk})_5$ core-shell particles

Briefly, silica particles (~3 μm diameter, 5 wt% suspension, 100 μL) were first washed with NaOAc (50 mM, pH 4) by centrifugation (1000 *g*, 1 min) and redispersed in 200 μL of NaOAc (50 mM, pH 4). To form the multilayers, an equal volume of PVPON was added to the particle suspension for adsorption (15 min) with constant shaking. The particles were then washed via three centrifugation (1000 *g*, 1 min)/redispersion (200 μL) cycles. After the first layer, the adsorption process was repeated with five bilayers of $PMA/PVPON_{Alk}$ (both at 1 g L^{-1} in 50 mM NaOAc, pH 4) to form $PVPON/(PMA/PVPON_{Alk})_5$ core-shell particles. The particles (in 50 μL of 50 mM NaOAc, pH 4) were then incubated with the bisazide linker from Example 4 (1 g L^{-1}, 150 μL) in the presence of sodium ascorbate (4.4 g L^{-1}, 50 μL) and copper sulfate (1.75 g L^{-1}, 50 μL) (all in 50 mM NaOAc, pH 4) overnight for crosslinking.

Step 2. Sortase-mediated conjugation of Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:284) with GGGWWSSK-PEG4-N_3 (SEQ ID NO:77) modified (PVPONAlk)5 core-shell particles

For the immobilization of Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) peptide from Example 2, the crosslinked $(PVPON_{Alk})_5$ core-shell particles from Step 1 (in 25 μL of 150 mM NaOAc, pH 5) were incubated with a mixture of GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77) (0.2 g L^{-1}, 75 μL; i.e., final concentration = 0.1 g L^{-1}), sodium ascorbate (4.4 g L^{-1}, 25 μL) and copper sulfate (1.75 g L^{-1}, 25 μL) (all in 150 mM NaOAc, pH 5) for 30 min. The GGGWWSSK-PEG_4-N_3-functionalized particles were redispersed in 100 μL of PBS for 12 h to remove PVPON and PMA.

To evaluate the optimal concentration of Sortase A for the functionalization, the GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77)-coated particles were incubated in 300 μL of sortase coupling buffer (50 mM Tris, 150 mM NaCl, pH 8) with Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385), $CaCl_2$ and Sortase A. Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) and $CaCl_2$ were added to a final concentration of 0.1 g L^{-1} and 0.5 mM, respectively. Sortase A was examined at different final concentrations (0, 0.05, 0.1 and 0.5 g L^{-1}). The mixture was incubated at 37°C for 1 h with gentle shaking. The same procedure was employed for the investigation of the optimal concentration of Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) and ligation time. Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) was tested at 0.05, 0.1 and 0.5 g L^{-1} in the presence of Sortase A (0.1 g L^{-1}). For the optimization of ligation time, Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) (0.1 g L^{-1}) was immobilized on the GGGWWSSK-PEG_4-N_3-coated particles for 0.5, 1 and 2 h in the presence of Sortase A (0.1 g L^{-1}). After conjugation, the Lissamine Rhodamine B-LPETGGHHHHHH(SEQ ID NO:385)-functionalized particles were washed with PBS extensively.

Fluorescence intensity of the particles not treated with Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) peptide was set at 1. The results in Figure 2A show that the optimal Sortase A concentration is 0.1 g/L.

The optimal concentration of Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) peptide was 0.1 g/L as shown in Figure 2B.

The results in Figure 2C show that the optimal incubation time is 1 hour.

A further test of the Sortase mediated conjugation was undertaken using the following conditions. The GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77)-functionalized and unfunctionalized $(PVPON_{Alk})_5$ core-shell particles were incubated with Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) peptide (0.1 g/L) at 37°C for 1 hour at pH 8 (50 mM Tris, 150 mM NaCl) in the presence or absence of Sortase A (0.1 g/L). Fluorescence intensity of the untreated unfunctionalized capsules was set at 1. The results are shown in Figure 2D showing untreated unfunctionalized capsules (white bar), unfunctionalized particles with Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) peptide (hatched bar), unfunctionalized particles, Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) peptide and Sortase A (double hatched bar), GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77)-functionalized particles and Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) (gray bar) and GGGWWSSK-PEG_4-N_3 (SEQ ID NO:77)-functionalized particles, Lissamine Rhodamine B-LPETGGHHHHHH (SEQ ID NO:385) and Sortase A (black bar).

Example 6: targeting of scFv-LPETG (SEQ ID NO:282) conjugated $DL800_{Az}$-labeled $(PVPON_{Alk})_5$/PEG_{Alk} capsules to in vitro thrombi

The capsules prepared in Example 4 were tested for binding to *in vitro* human thrombi.

Blood from a healthy volunteer taking no medication was anti-coagulated with citric acid and centrifuged at 1000 rpm for ten minutes. To every 1 mL of the resulting platelet rich plasma (PRP) was added a mixture of adenosinediphosphate (ADP, möLab GmbH, Langenfeld, Germany) (100 µL, 200 mM in water), actin (Dade Behring, Marburg, Germany) (88 µL) and calcium chloride (25 µL, 1 M). To form thrombi, aliquots (100 µL) of the PRP mixture were incubated at 37°C for 12 min. After washing in PBS (with

Ca/Mg), each thrombus was exposed to ~2 × 10^5 $DL800_{Az}$-labeled scFv(+)- and scFv(-)-functionalized, and unfunctionalized capsules in 500 µL of PBS (with Ca/Mg) at 37°C for 30 min with gentle mixing. The thrombi were then washed in PBS (with Ca/Mg) three times to remove unbound capsules. The fluorescence intensities of the capsule-bound thrombi were analyzed using a Licor Odyssey near infrared imaging system. (Both scFv(+) and scFv(-) were LPETG (SEQ ID NO:380) encoded. scFv(+) = anti-GPIIb/IIIa scFv; scFv(-) = mutated scFv).

To perform completive binding, the activated GP IIb/IIIa receptors on thrombi were pre-blocked with scFv(+). Thrombi were exposed to scFv(+) (300 µL, 2 g L^{-1}) in 500 µL of PBS (with Ca/Mg) for 30 min at 37°C with gently mixing. Following incubation, the blocked thrombi were washed three times in PBS (with Ca/Mg) to remove attached antibody before addition of the scFv(+)-functionalized capsules. Incubation, washing and analysis were performed as described above.

As shown Figure 3, strong specific binding of anti-GPIIb/IIIa scFv conjugated particles was observed.

Example 7: Sortase conjugation of ScFv-LIBS-LPETG with GGG modified Cells

Cell-anti-LIBS conjugation and functional assessment CHO cells with surface amines were functionalized with sulfhydryl groups according to the protocol described below and reacted with H_2N-GGG-AGAGA-K-Malemide (SEQ ID NO:387) as prepared in Example 2. The GGG-functionalized cells were then conjugated with anti-LIBS-LPETG-scFv using sortase methodology. The protocol is set out below.

Step 1. Introduction of sulfhydryls to cell surface via reaction with primary amines using 2-Iminothiolane or Traut's reagent

CHO cells were trypsinized and washed once with PBS. 1 x 10^6 cells were resuspended in 200 µL of modified PBS buffer with EDTA (PBS without Ca and Mg, 4500 mg/L glucose, 15 mM HEPES, 2 mM EDTA, pH 7.3). Traut's reagent was added at the final concentration of 0.6 mM (16 µg). Cells were incubated for 30 min at room temperature

with shaking. After the incubation, cells were washed once with modified PBS with EDTA.

Step 2. Labelling cells with NH_2-GGG-tags via specific reaction of sulfhydryls on cell surface and maleimide groups on NH_2-GGG-maleimide peptides:

1 x 10^6 cells were resuspended in 200 µL of modified PBS buffer with EDTA. H_2N-GGG-AGAGA-K-Maleimide (SEQ ID NO:387) peptide was added to the final concentration of 12 µM (2 µg). The reaction mixture was incubated for 30 min at room temperature with shaking. Cells were then washed once with modified PBS buffer without EDTA (PBS with Ca, Mg, 4.5 g/L glucose, 15 mM HEPES, pH 7.3).

Step 3. Sortase-mediated coupling between anti-LIBS-LPETG and NH_2-GGG-groups on cell membrane

1 x 10^6 cells were resuspended in 100 µL of modified PBS buffer without EDTA. Anti-LIBS-LPETG and Sortase A enzyme were added to the final concentration of 10 µM and 10 µM, respectively. Cells were incubated at 37°C for 1 hour with shaking, then pelleted and washed twice with PBS. The supernatant was collected for evaluation of the coupling efficiency. Non-coupled anti-LIBS was quantified by SDS-PAGE.

Step 4. Page Analysis

30 µL of each sample (remaining reaction mixture) and 6 µL of 5X reducing SDS loading buffer were added to 1.5-mL tube and denatured at 96°C for 5 min. 36 µL of each sample was run on SDSPAGE gel in SDS running buffer at 30 mA for 2 hours. The gel was then stained with Coommassie Brilliant Blue for 1 hour and subsequently destained for at least 12 hours with Coommassie destaining solution. The gel was visualised and analysed using a BioRad Gel-Doc system with Quantity One software (Australia).

SDS-PAGE analysis of the reaction mixture revealed that the average coupling efficiency was 10.08 ± 3.15 pg (n=3) of anti-LIBS per CHO cell.

Step 5. Anti-LIBS-conjugated cells were assessed in in vitro static adhesion assays for their binding ability to activated platelets.

100 µL of diluted purified platelets was added into each well of a 96-well plate and stimulated with 18 µM ADP. After incubating at 37°C for 20 min, wells were washed twice with PBS (Ca, Mg) and then blocked with 100 µL of 1% BSA in JNL buffer (pH 7.4) for 30 min at 37°C. After washing twice, platelet-coated wells were incubated with predetermined numbers of either targeting cells (anti-LIBS coupled cells) or non-targeting cells (GGG-cells) in PBS with 0.5% BSA at 37°C for 10 min. Wells were then washed twice with PBS and fixed with Cellfix. DIC/fluorescent images were taken by IX81 Olympus Fluorescence Microscope and Cell^P 1692 (AnalySIS Image Processing) software. Binding particles/cells were quantified using Image J software.

It was found that the binding of the cells was significantly stronger than the control (GGG-) cells as shown in Figure 4. The degree of binding increased with increasing cell numbers in a dose dependant manner.

Step 6. Flow Experiments.

After demonstrating the targeting capability of sortase-coupled cells in a static adhesion assay, anti-LIBS-coupled cells were investigated further in flow experiments Glass capillaries (0.20 x 2.0 mm I.D., 10 cm in length, Vitrotubes™, USA) were coated with 42 µL of 100 µg/mL collagen-1 overnight at 4°C. Capillaries were then blocked with 1% BSA in PBS for 1 hour at 37°C. Blood was taken from healthy donors, citrated and perfused through the capillary at a shear rate of 100 s^{-1}. After 5 min, thrombi with desirable sizes formed. The capillary was then washed by perfusing with PBS (Ca, Mg) for 5 min at a shear rate of 500 s^{-1} until no blood cells were observed. Particles in PBS containing 0.5% BSA (3 x 10^6 particles or cells/mL) was perfused through capillary for 5 min at different shear rates. Movies and images were taken using Olympus Fluorescence Microscope and Cell^P 1692 (AnalySIS Image Processing) software. Capillaries were washed with PBS and thrombi were specifically stained with FITC-PAC-1 or CD62P-PE. Binding particles were quantified using Image J software. Thrombus area was estimated by Image-Pro Plus 6.0.

After perfusing targeted cells through a platelet-coated capillary at 100 s^{-1} for 5 min, approximately 400 cells/mm^2 platelet aggregates were immobilized while less than 10 cells bound to a unit area (mm^2) of platelet aggregates. When the shear rate increased to 250 s^{-1}, specific targeting was still observed, however the number of bound cells decreased to 150 cells/mm^2 platelet aggregate, indicating shear dependent targeting efficacy.

Step 7. Cell Staining with CellTracker™ Green CMFDA

To facilitate visualization, cells were stained with CellTracker™ Green CMFDA by the following method.

CHO cells within 20 passages were cultured in 175-cm^2 flask until they reached approximately 95% of confluency. The medium was removed and cells were washed once with PBS. CellTracker™ Green CMFDA (Invitrogen, Australia) dye working solution (4 μM in plain DMEM buffer without serum) was added to the flask containing the adherent cells. After incubating for 40 min at 37°C/5% CO_2, the dye working solution was replaced with fresh and prewarmed plain DMEM medium (without serum). The cells were incubated for another 30 min at 37°C/5% CO_2. They were then trypsinized, harvested and washed once with PBS before undergoing the coupling process.

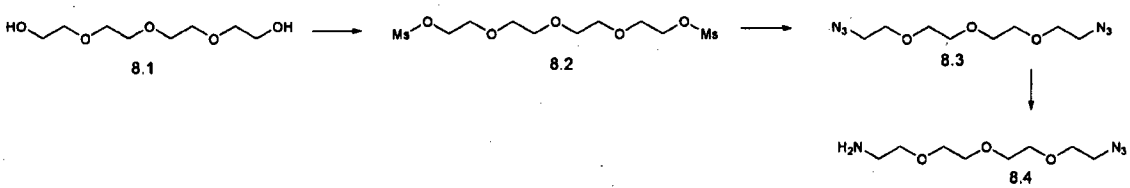
Step 8. Intravital Microscopy

Intravital microscopy was used to evaluate the targeting of anti-LIBS-coupled cells to activated platelets in mesenteric veins of mice. 4-5 week old C57BL/6 wild-type mice weighing 15-17 g were anaesthetized by intraperitoneal injection of a ketamine:xylazine mixture (100:20 mg/kg body weight). Mesenteric arteries/veins (diameters 100-150 μm) were observed with inverted microscope (IX81, Olympus) and the images were recorded with digital B/W camera (XM10, Olympus). The injury of arteries/veins was induced by the micro-drop of 10% $FeCl_3$. When the thrombus started building up, 200 μL of PBS containing the predetermined amounts of anti-Libs particles was injected into the blood flow system via a jugular vein cannula. DIC and fluorescence images of the thrombus were taken before, during and after the injection. Particles were seen by their TRITC autofluorescence. The number of anti-LIBS-conjugated cells binding on the venous

thrombi was significantly higher compared to the control (GGG-tagged) cells, indicating successful *in vivo* cell targeting with this approach.

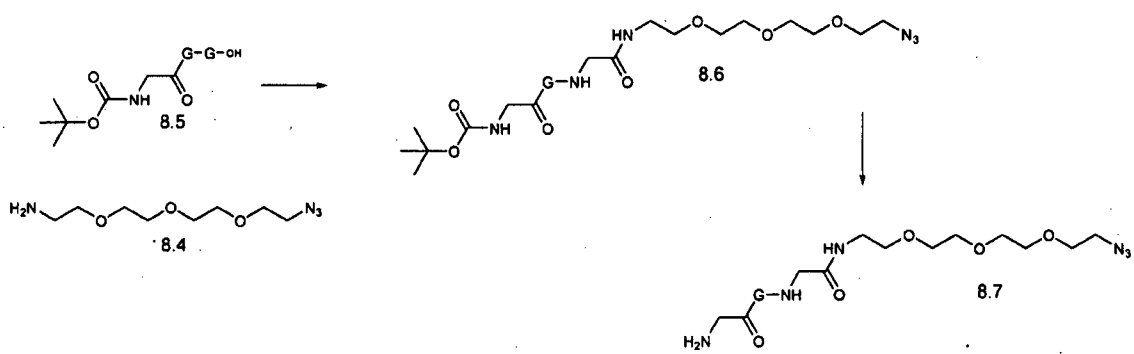
Example 8: Synthesis of GGG-PEG$_3$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:391)

a. Synthesis of $H_2N[(CH_2)_2O]_3(CH_2)_2N_3$



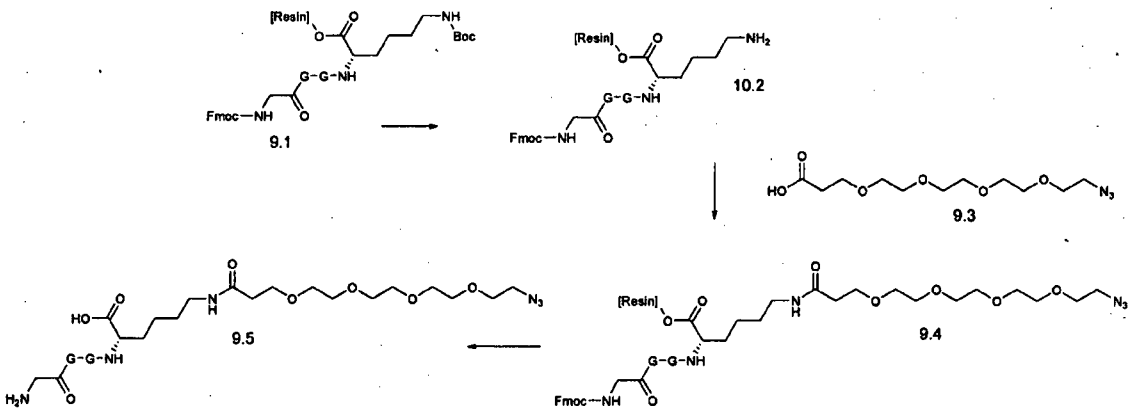
A polyethylene glycol compound having a terminal amino group and a terminal azide was prepared following known methods (Tetrahedron, 2009, 65, 7329). Briefly, tetraethyleneglycol **8.1** was treated with mesyl chloride (MsCl) in dichloromethane and triethylamine at 0°C to produce the tetraethyleneglycol dimesylate **8.2**. The dimesylated compound was treated with sodium azide (NaN_3) in ethanol/dimethyl acetamide (DMAc) at reflux to give the diazo compound **8.3**. One azo group of the diazo compound was then selectively reduced with triphenylphosphine (PPh_3) in a solution of tetrahydrofuran/ether in the presence of 1M HCl to give the monoazo, monoamino PEG compound **8.4** after chromatography.

b. Synthesis of GGG-PEG$_3$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:391)



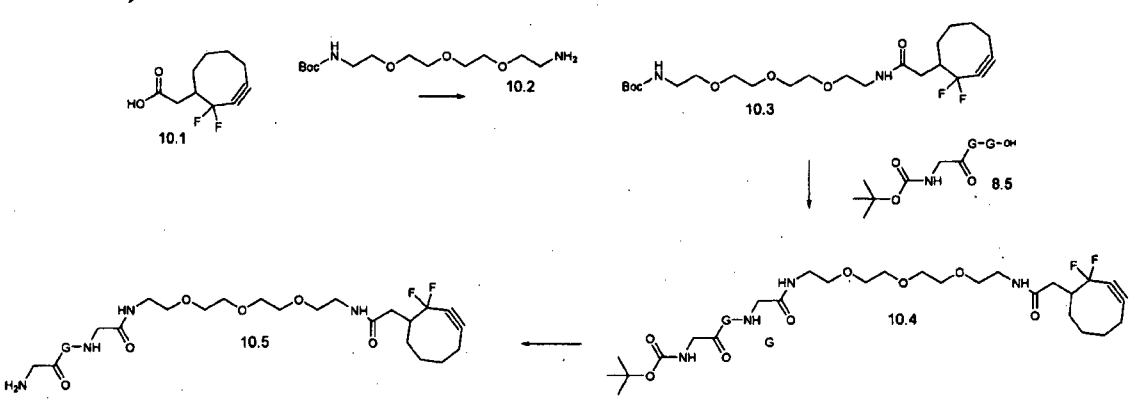
Boc-GGG **8.5** was activated with dicarbodiimide (DCC) in the presence of 1.5 equivalents of N-hydroxysucciniamide (NHS) in DMF then treated with **8.4** from a. to form Boc-GGG-PEG$_3$-(CH$_2$)$_2$N$_3$ **8.6**. This compound was then treated with trifluoroacetic acid (TFA) to deprotect the N-terminus and provide GGG-PEG$_3$-(CH$_2$)$_2$N$_3$ **8.7**.

Example 9: Synthesis of GGGK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:22)



Using standard solid phase synthesis techniques, FmocGGG-εBoc-K **9.1** was synthesised on resin. This compound, while still attached to the resin, was treated with TFA in dichloromethane to remove the Boc protecting group from the ε-amino group of the lysine to give FmocGGG-εNH$_2$-K **9.2**. HO$_2$C-PEG$_4$-N$_3$ **9.3** (product #16502 purchased from Quanta Biodesign) was activated by treatment with DCC and NHS in dimethylformamide (DMF) and reacted with the deprotected ε-amino group of the lysine to give **9.4**. The compound was then deprotected and cleaved from the resin to provide GGGK-εPEG$_4$-N$_3$ **9.5**.

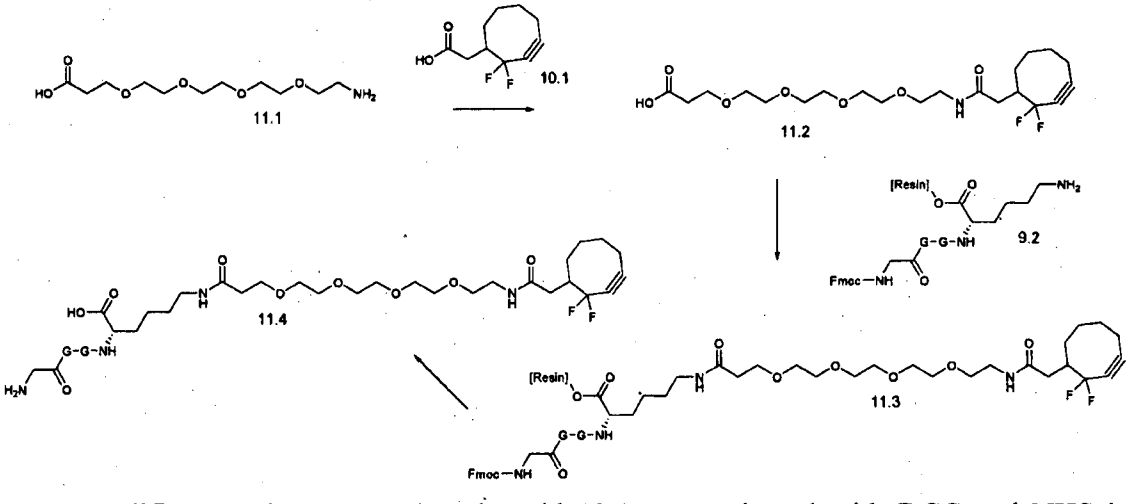
Example 10: Synthesis of GGG-PEG$_4$-NHCOCH$_2$-4-(3,3-difluorooctacyclooctyne) (SEQ ID NO:34)



2-[4-(3,3-difluorocyclooct-1-yne)]acetic acid **10.1** was activated with DCC and NHS in DMF and treated with Boc-NH-[(CH$_2$)$_2$O]$_3$(CH$_2$)$_2$NH$_2$ **10.2** to give BocNH-PEG$_3$-(CH$_2$)$_2$NHCOCH$_2$-4-(3,3-difluorocyclooct-1-yne) **10.3** which was then treated with TFA to remove the Boc protecting group. The resulting H$_2$N-PEG$_3$-(CH$_2$)$_2$NHCOCH$_2$-4-(3,3-

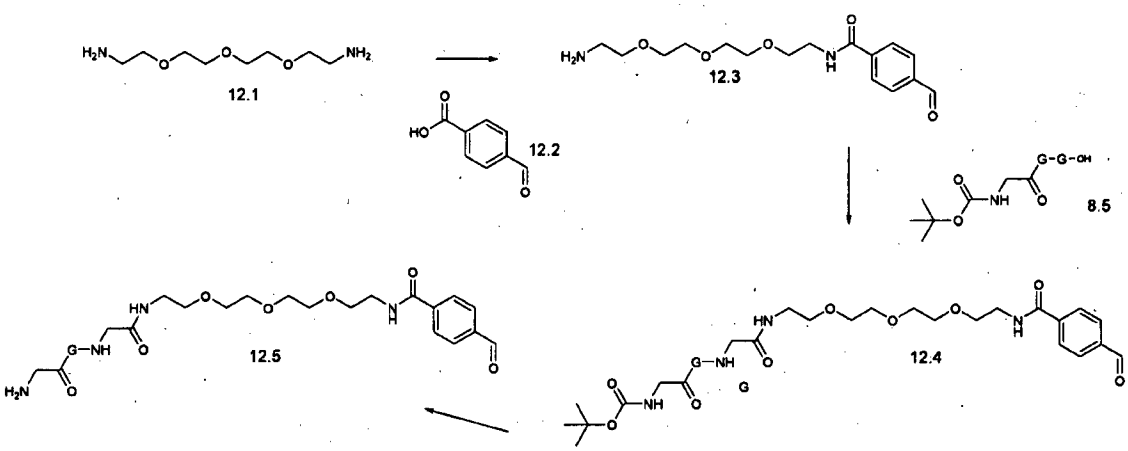
difluorocyclooct-1-yne) **10.3** was reacted with BocGGG **8.5**, which had been activated with DCC in DMF, to give BocGGG-PEG_3-$(CH_2)_2NHCOCH_2$-4-(3,3-difluorocyclooct-1-yne) **10.4**. Finally the N-terminal Boc group was deprotected with TFA to give alkyne **10.5**.

Example 11: Synthesis of GGGK-εPEG$_3$-(CH$_2$)$_2$NHCOCH$_2$-4-(3,3-difluorooctacyclooct-1-yne) (SEQ ID NO:23)



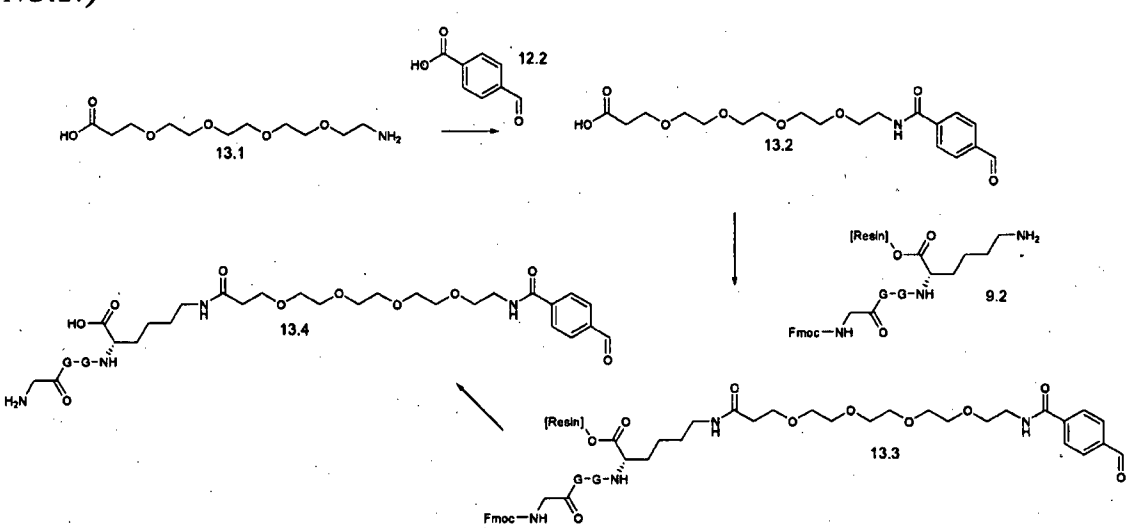
2-[4-(3,3-difluorocyclooct-1-yne)]acetic acid **10.1** was activated with DCC and NHS in DMF and treated with HO_2C-$[(CH_2)_2O]_3(CH_2)_2NH_2$ **11.1**. After chromatography, the carboxylic acid group of HO_2C-PEG_3-$(CH_2)_2NHCOCH_2$-4-(3,3-difluorooctacyclooct-1-yne) **11.2** was activated with DCC and NHS in DMF and reacted with resin bound FmocGGG-εNH_2-K-resin **10.2** to produce resin bound GGGK-εPEG_3-$(CH_2)_2NHCOCH_2$-4-(3,3-difluorooctacyclooct-1-yne) **11.3** which was then removed from the resin as described in Example 9 to provide alkyne **11.4**.

Example 12: Synthesis of GGG-PEG$_3$-(CH$_2$)$_2$NHCO-4-formylbenzene (SEQ ID NO:38)



4-Formylbenzoic acid **12.2** was activated with DCC and NHS in DMF and treated with H_2N-[$(CH_2)_2O$]$_3$$(CH_2)_2NH_2$ **12.1** to give H_2N-PEG$_3$-$(CH_2)_2$NHCO-4-formylbenzene **12.3**. The resulting H_2N-PEG$_3$-$(CH_2)_2$NHCO-4-formylbenzene **12.3** was reacted with BocGGG **8.5**, which had been activated with DCC and NHS in DMF, to give BocGGG-PEG$_3$-$(CH_2)_2$NHCO-4-formylbenzene **12.4**. Finally the N-terminal Boc group was deprotected with TFA to give aldehyde **12.5**.

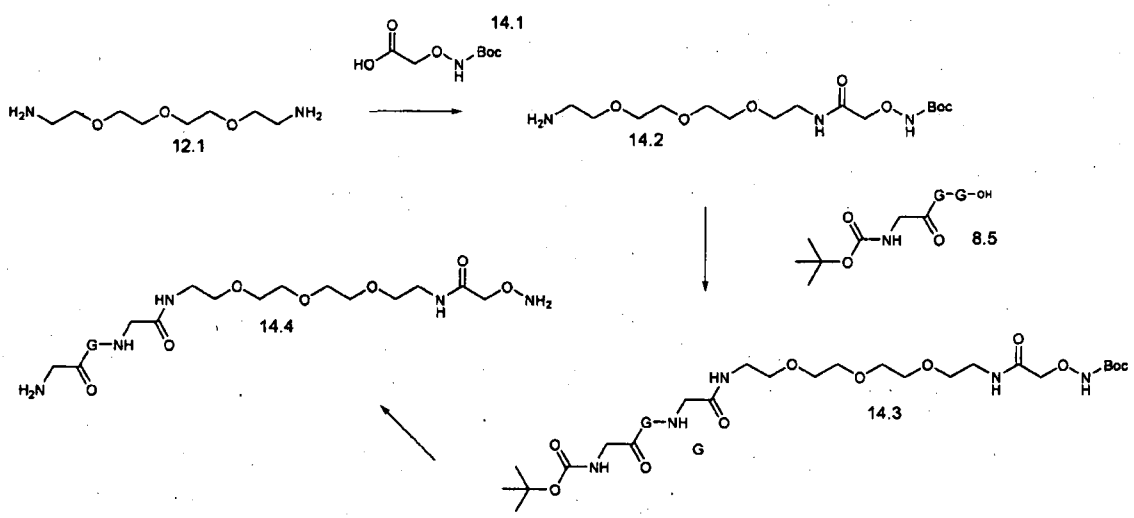
Example 13: Synthesis of GGGK-εPEG$_4$-(CH$_2$)$_2$NHCO-4-formylbenzene (SEQ ID NO:27)



4-Formylbenzoic acid **12.2** was activated with DCC and NHS in DMF and reacted with HO_2C-[$(CH_2)_2O$]$_3$$(CH_2)_2NH_2$ **13.1**. After chromatography, the carboxylic acid group of HO_2C-PEG$_3$-$(CH_2)_2$NHCO-4-formylbenzene **13.2** was activated with DCC and NHS in

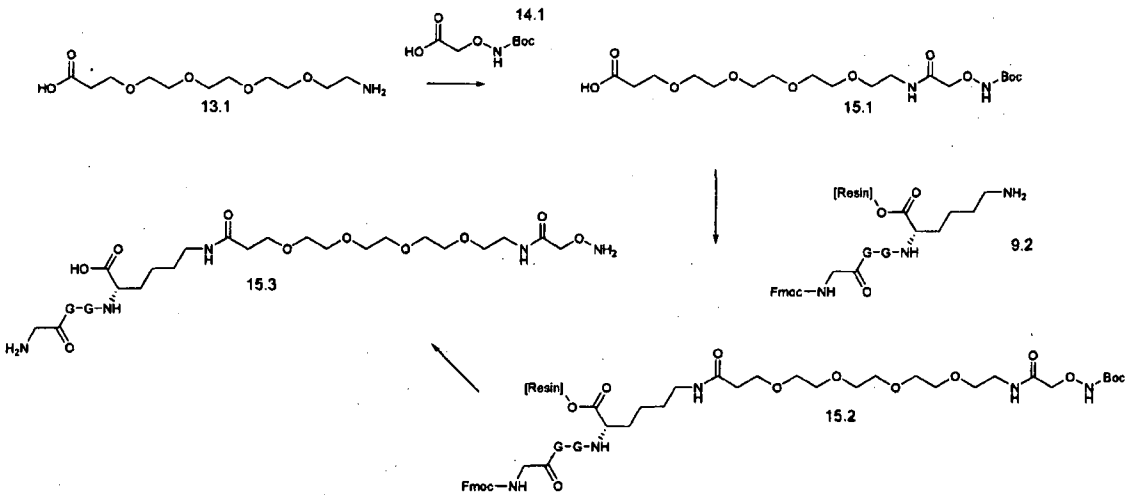
DMF and reacted with resin bound FmocGGG-εNH$_2$-K-resin **9.2** to produce resin bound FmocGGGK-εPEG$_3$-(CH$_2$)$_2$NHCO-4-formylbenzene **13.3** which was then deprotected and removed from the resin to give aldehyde **13.4**.

Example 14: Synthesis of GGG-PEG$_3$-(CH$_2$)$_2$NHCOCH$_2$ONH$_2$ (SEQ ID NO:36)



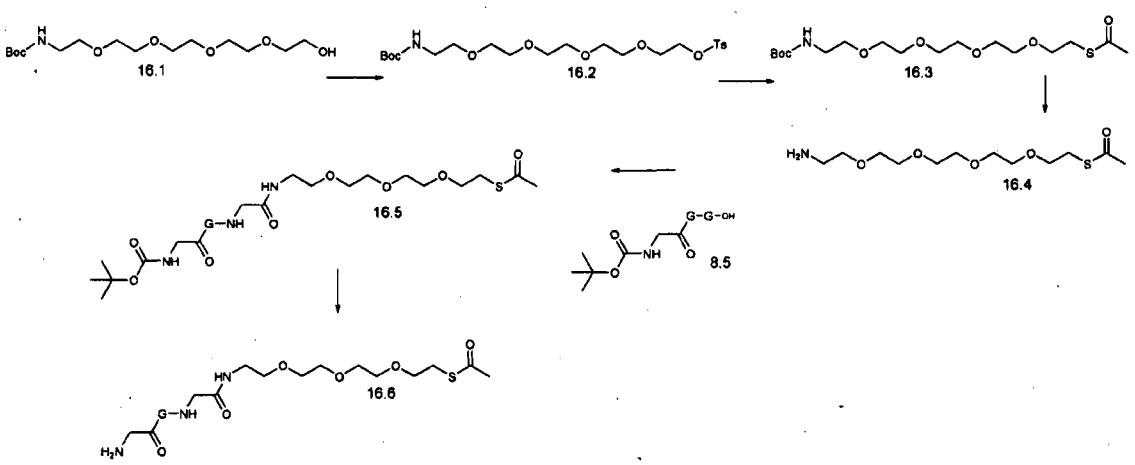
HO$_2$CCH$_2$ONH-Boc **14.1** was activated with DCC and NHS in DMF and treated with an excess of H$_2$N-[(CH$_2$)$_2$O]$_3$(CH$_2$)$_2$NH$_2$ **12.1** to give H$_2$N-PEG$_3$-(CH$_2$)$_2$NHCOCH$_2$ONHBoc **14.2**. The resulting H$_2$N-PEG$_3$-(CH$_2$)$_2$NHCOCH$_2$ONHBoc was reacted with BocGGG **8.5**, which had been activated with DCC and NHS in DMF, to give BocGGG-PEG$_3$-(CH$_2$)$_2$NHCOCH$_2$ONHBoc **14.3**. Finally the N-terminal Boc groups were deprotected with TFA to give alkoxyamine **14.4**.

Example 15: Synthesis of GGGK-εPEG$_4$-(CH$_2$)$_2$NHCOCH$_2$ONH$_2$ (SEQ ID NO:25)



$HO_2CCH_2ONHBoc$ **14.1** was activated with DCC and NHS in DMF and reacted with HO_2C-$[(CH_2)_2O]_3(CH_2)_2NH_2$ **13.1**. After chromatography, the carboxylic acid group of HO_2C-PEG_3-$(CH_2)_2NHCOCH_2ONHBoc$ **15.1** was activated with DCC and NHS in DMF and reacted with resin bound FmocGGG-εNH_2-K-resin to produce resin bound FmocGGGK-εPEG_3-$(CH_2)_2NHCOCH_2ONHBoc$ **15.2** which was then deprotected and removed from the resin to provide alkoxyamine **15.3**.

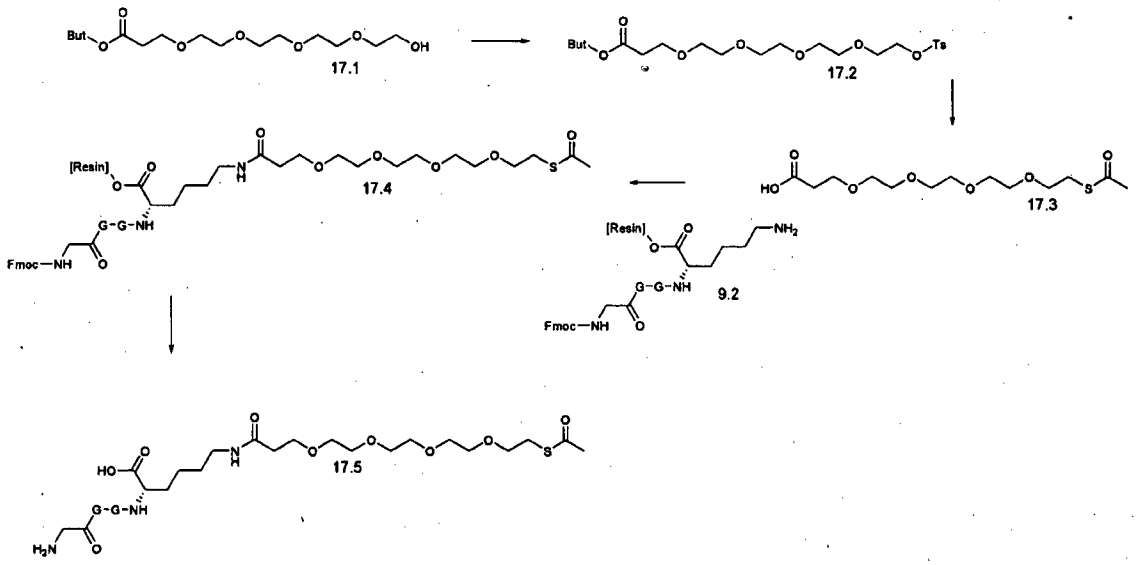
Example 16: Synthesis of GGG-PEG$_4$(CH$_2$)$_2$SCOCH$_3$ (SEQ ID NO:392)



BocNH-$[(CH_2)_2O]_4(CH_2)_2OH$ **16.1** was treated with tosyl chloride and pyridine in dichloromethane to give BocNH-PEG_4-$(CH_2)_2OTs$ **16.2**. The resulting Boc-NH-PEG_4-$(CH_2)_2OTs$ **16.2** was reacted with tetrabutylammonium iodide (TBAI) and potassium thioacetate (KSAc) in DMF to give Boc-NH-PEG_4-$(CH_2)_2SCOCH_3$ **16.3** and then treated

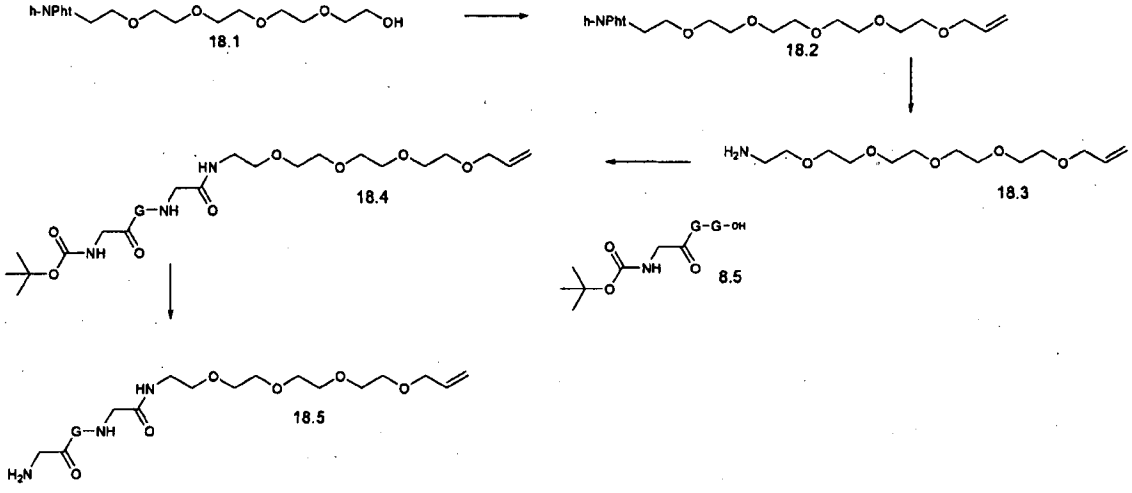
with TFA to provide H_2N-PEG_4-$(CH_2)_2SCOCH_3$ **16.4**. BocGGG **8.5**, was activated with DCC and NHS in DMF and reacted with H_2N-PEG_4-$(CH_2)_2SCOCH_3$ **16.4** to give BocGGG-PEG_4-$(CH_2)_2SCOCH_3$ **16.5**. Finally the N-terminal Boc group was deprotected with TFA to give thioacetate **16.6**.

Example 17: Synthesis of GGGK-εPEG$_4$-(CH$_2$)$_2$SCOCH$_3$ (SEQ ID NO:393)



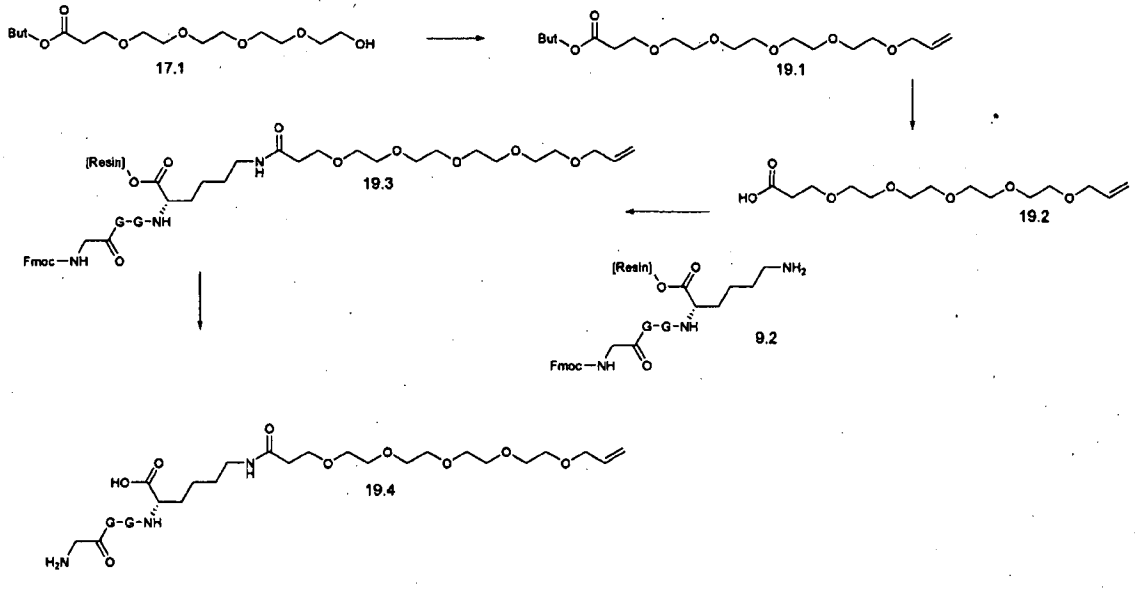
tBuOCO-$[(CH_2)_2O]_4(CH_2)_2OH$ **17.1** was treated with tosyl chloride and pyridine in dichloromethane to give tBuOCO-PEG_4-$(CH_2)_2OTs$ **17.2**. The resulting tBuOCO-PEG_4-$(CH_2)_2OTs$ **17.2** was reacted with tetrabutylammonium iodide (TBAI) and potassium thioacetate (KSAc) in DMF then the t-Butylester was cleaved with TFA to give HO_2C-PEG_4-$(CH_2)_2SCOCH_3$ **17.3**. This compound was activated with DCC and NHS in DMF and reacted with resin bound FmocGGG-εNH_2-K(SEQ ID NO:293)-resin **9.2** to produce resin bound FmocGGGK-εPEG$_4$-$(CH_2)_2SCOCH_3$ (SEQ ID NO:294) **17.4** which was then deprotected and removed from the resin to give thioacetate **17.5**.

Example 18: Synthesis of GGG-PEG$_4$CH$_2$CH=CH$_2$ (SEQ ID NO:40)



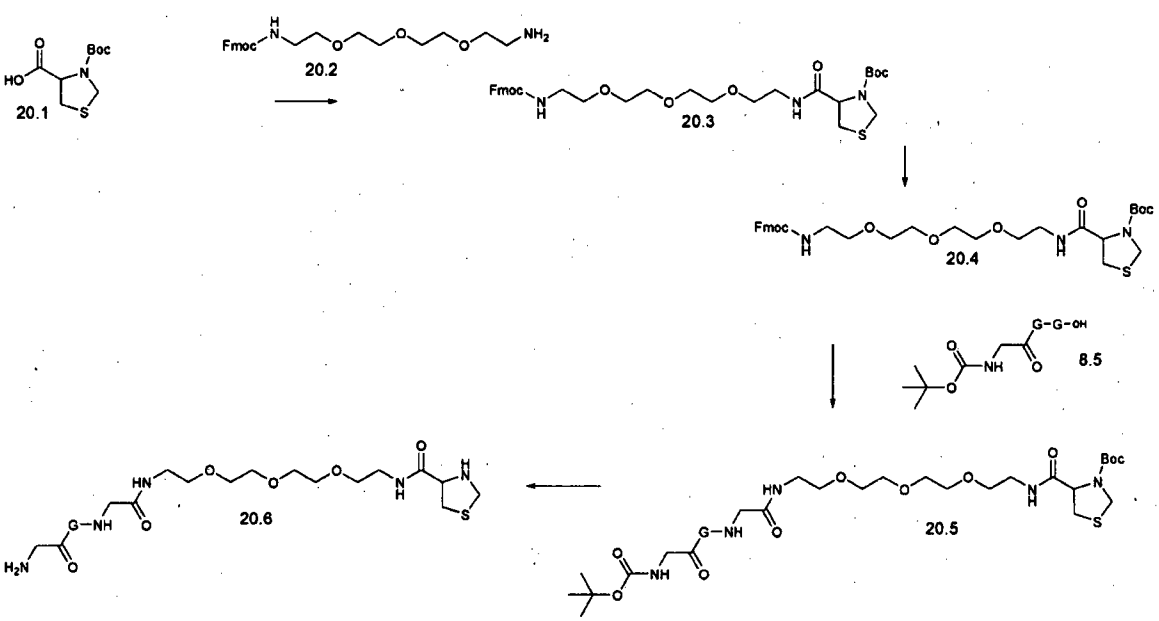
Phth-NH-[(CH$_2$)$_2$O]$_4$(CH$_2$)$_2$OH **18.1** was treated with allyl chloride and tetrabutylammonium sulphate in the presence of NaOH in dichloromethane to give Phth-NH-PEG$_4$-(CH$_2$)$_2$OCH$_2$CH=CH$_2$ **18.2**. The resulting Phth-NH-PEG$_4$-(CH$_2$)$_2$OCH$_2$CH=CH$_2$ **18.2** was reacted with hydrazine to give H_2N-PEG$_4$-(CH$_2$)$_2$OCH$_2$CH=CH$_2$ **18.3**. BocGGG **8.5**, was activated with DCC and NHS in DMF and reacted with H_2N-PEG$_4$-(CH$_2$)$_2$OCH$_2$CH=CH$_2$ **18.3** to give BocGGG-PEG$_4$-(CH$_2$)$_2$OCH$_2$CH=CH$_2$ **18.4**. Finally the N-terminal Boc group was deprotected with TFA to give the allylic ether **18.5**.

Example 19: Synthesis of GGGK-εPEG$_4$-CH$_2$CH=CH$_2$ (SEQ ID NO:29)



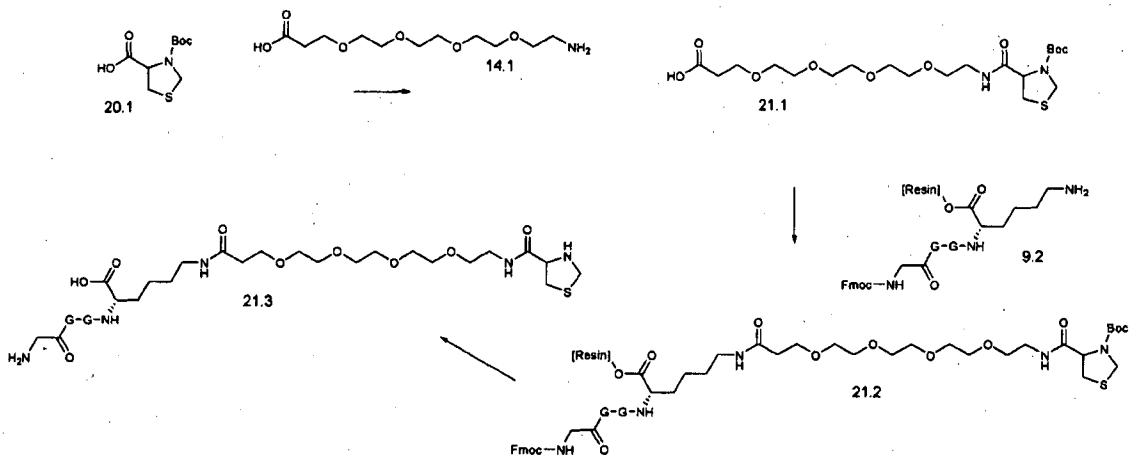
tBuOCO-$[(CH_2)_2O]_4(CH_2)_2OH$ **17.1** was treated with allyl chloride and tetrabutylammonium sulphate in the presence of aqueous NaOH in dichloromethane to give tBuOCO-PEG_4-$(CH_2)_2OCH_2CH=CH_2$ **19.1**. The resulting tBuOCO-PEG_4-$(CH_2)_2OCH_2CH=CH_2$ **19.1** was treated with TFA to provide HO_2C-PEG_4-$(CH_2)_2OCH_2CH=CH_2$ **19.2**. The carboxylic acid of **19.2** was activated with DCC and NHS in DMF and reacted with resin bound FmocGGG-εNH_2-K-resin **9.2** to produce resin bound FmocGGGK-εPEG_4-$(CH_2)_2OCH_2CH=CH_2$ **19.3** which was then deprotected and removed from the resin to provide allylic ether **19.4**.

Example 20: Synthesis of GGG-PEG_3-$(CH_2)_2$NHCO(4-thiazoline) (SEQ ID NO:394)



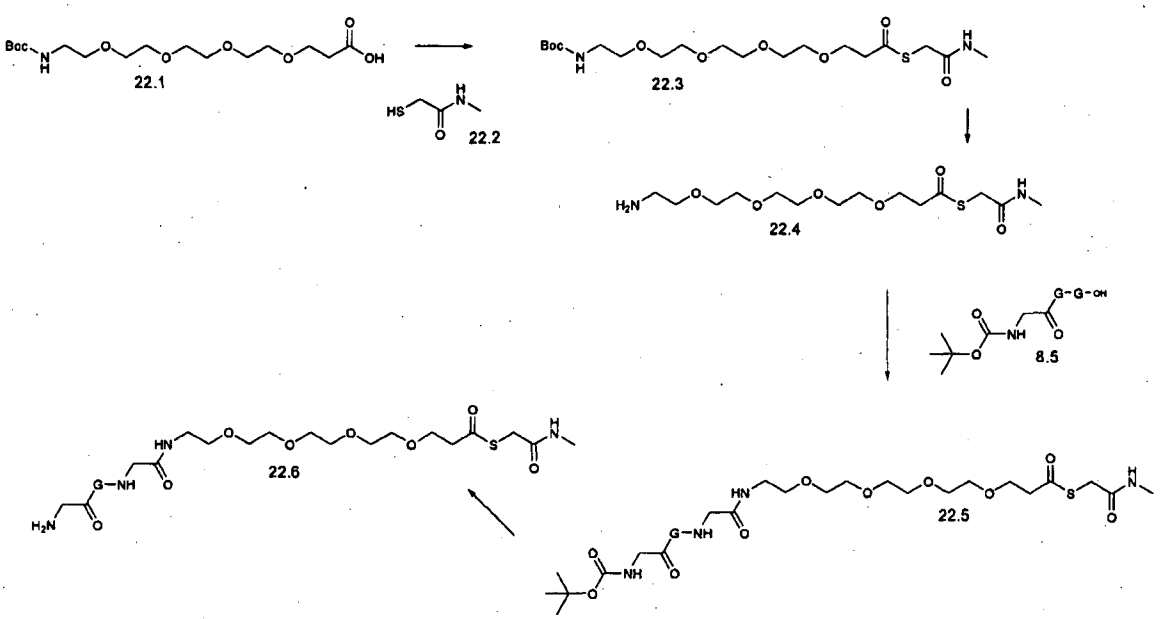
Protected cysteine **20.1** was activated with DCC and NHS in DMF then reacted with Fmoc-NH-$[(CH_2)_2O]_3(CH_2)_2NH_2$ **20.2** to give Fmoc-NH-PEG_3-$(CH_2)_2$NHCO(4-N-Boc-thiazoline) **20.3**. The resulting Fmoc-HN-PEG_3-$(CH_2)_2$NHCO(4-N-Boc-thiazoline) **20.3** was Fmoc deprotected with pyridine in DMF to give H_2N-PEG_3-$(CH_2)_2$NHCO-(4-N-Boc-thiazoline) **20.4**. BocGGG **8.5**, was activated with DCC and NHS in DMF in the presence of triethylamine and reacted with H_2N-PEG_3-$(CH_2)_2$NHCO-(4-N-Boc-thiazoline) **20.4** to give BocGGG-PEG_3-$(CH_2)_2$NHCO-(4-N-Boc-thiazoline) **20.5**. Finally the Boc groups were deprotected with TFA to provide the protected cysteine **21.6**.

Example 21: Synthesis of GGGK-εPEG$_4$-(CH$_2$)$_2$NHCO(4-thiazoline) (SEQ ID NO:395)



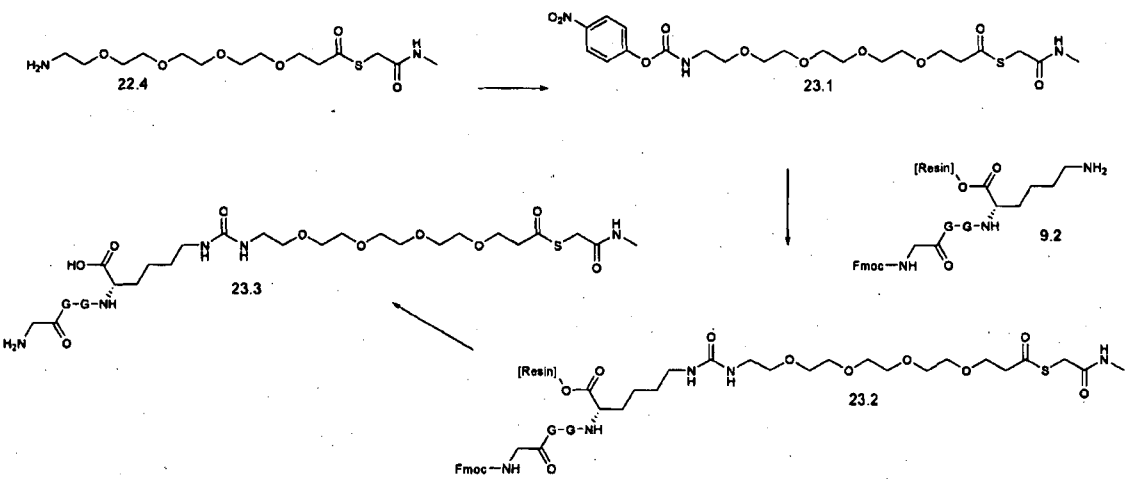
Protected cysteine **20.1** was activated with DCC and NHS in DMF then reacted with HO_2C-$[(CH_2)_2O]_4(CH_2)_2NH_2$ **14.1** to give HO_2C-PEG_4-$(CH_2)_2NHCO$(4-N-Boc-thiazoline) **21.1**. This compound **21.1** was activated with DCC and NHS in DMF and reacted with resin bound FmocGGG-εNH_2-K-resin **9.2** to produce resin bound FmocGGGK-εPEG_4-$(CH_2)_2NHCO$-(4-N-Boc-thiazoline) **21.2** which was then deprotected and removed from the resin as described in Example 9 to provide protected cysteine **21.3**.

Example 22: Synthesis of GGG-PEG$_4$(CH$_2$)$_2$COSCH$_2$CONHCH$_3$ (SEQ ID NO:41)



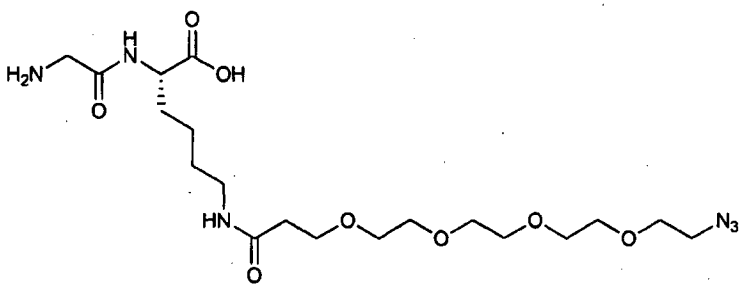
Boc-NH-$[(CH_2)_2O]_4(CH_2)_2CO_2H$ **22.1** was activated with DCC and NHS in DMF then treated with N-(methyl)metcaptoacetamide **22.2** to give thioester **22.3**. The Boc group of thioester **23.3** was removed by treatment with TFA to give amine-thioester **22.4**. The carboxylic acid of Boc-GGG **8.5** was activated by treatment of DCC and NHS in DMF and then reacted with amine-thioester **22.4** in the presence of triethylamine to give Boc-GGG-PEG_4-$(CH_2)_2COSCH_2CONHCH_3$ **22.5**. The Boc group of the Boc-amino-thioester **23.5** was then removed by treatment with TFA to give thioester **22.6**.

Example 23: Synthesis of GGGK-εPEG$_4$-(CH$_2$)$_2$COSCH$_2$CONHCH$_3$ (SEQ ID NO:30)



Amine-thioester **22.4** was reacted with 1 eq of *p*-nitrophenylchloroformate and triethylamine in DMF to give *p*-nitrophenylcarbamate **23.1**. A solution of p-nitrophenylcarbamate **23.1** was added to a suspension of resin bound Fmoc-GGG-εNH_2-K(SEQ ID NO:293)-resin **9.2** in DMF to produce resin bound Fmoc-GGG-K-εPEG_4-$(CH_2)_2CO$-$SCH_2CONHCH_3$ (SEQ ID NO:298) **23.4** which was then deprotected and removed from the resin as described in Example 9 to provide thioester **23.3**.

Example 24: Synthesis of GK-εPEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:198)



The peptide was synthesized on resin. Fmoc-Lys(Dde)-Wang Resin (909 mg, loading 0.33 mmol/g) was allowed to swell in DMF (20 mL) for 30 min. The DMF was drained and the Fmoc protecting group was removed with 20% piperidine/DMF (20 mL) under N_2 for 30 min. The resin was then washed with DMF (6x10 mL). The glycine residue was conjugated to the lysine residue by adding Boc-Gly-OH (158 mg, 0.9 mmol), HBTU (324 mg, 0.855 mmol), NMM (204 µL, 1.8mmol) in DMF (10 mL) to the resin. The reaction mixture under an N_2 atmosphere was placed on a shaker for 30 min and the resin was then washed with DMF (3x10 mL). Kaiser test confirmed the reaction was complete.

The ε-NH_2 group of lysine was deprotected by treatment with 5% Hydrazine hydrate/DMF (20 mL) under N_2 for 30 min. The resin was then washed with DMF (6x10 mL). The deprotected lysine amino group was pegylated by treating the resin with HOOC-PEG$_4$-(CH$_2$)$_2$N$_3$ (263 mg, 0.9 mmol), HATU (324 mg, 0.855 mmol) and NMM (204 µL, 1.8 mmol) in DMF (10 mL). The reaction was allowed to proceed under N_2 for 120 min. The resin was then washed with DMF (3x10 mL). Kaiser test confirmed the reaction was complete. The resin was washed with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x1 mL) and then air-dried (1060 mg).

The above peptide was removed from the resin and deprotected by treating with 10.6 mL of a mixture solution (2% H_2O, 3% Tris, 95% TFA) on a shaker at 25°C for 2 hours. The above mixture was filtered to give a filtrate in a 100 mL centrifuge tube. To the filtrate was added 50 mL diethyl ether to triturate the peptide. The solid was washed with diethyl ether (5x50 mL) and then dried under vacuum for 12 hours to give the crude peptide (108 mg).

The crude peptide (108 mg) was dissolved in water (10 mL) and purified by HPLC using a Global Chromatography 30*250mm SP-120-10-C18-AP and a solvent gradient of 18-28 Solvent A/30min. Purity by HPLC: 94.6% MS: $C_{19}H_{37}N_6O_8^+$: Calc 477.27, found: 477.5.

Column:	30*250mm, SP-120-10-C18-AP		
Solvent A:	0.1% Trifluoroacetic in 100% Acetonitrile		
Solvent B:	0.1% Trifluoroacetic in 100% Water		
Gradient:		A	B
	0.01 min	18%	82%
	3 min	28%	72%
	30.01 min	100%	0%
	35 min	Stop	
Flow rate:	25.0 mL/min		
Wavelength:	220 nm		
Volume:	10 mL		

HPLC Analysis

Type of machine

BeiJing chuangxin LC3000 HPLC

Buffer

Solvent A : 0.1% Trifluoroacetic in 100% Acetonitrile

Solvent B : 0.1% Trifluoroacetic in 100% Water

MS conditions:

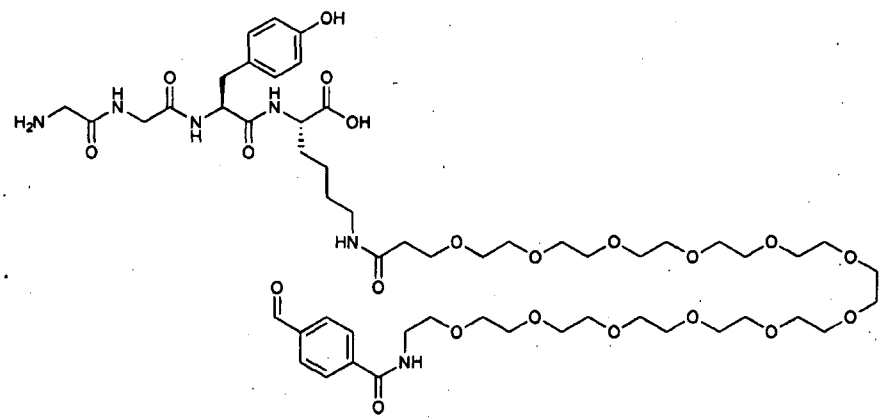
Type of machine

DaoJing LCMS-2020

Ionization

Probe: ESI, Nebulizer Gas Flow: 1.5L/min, Drying Gas Flow: 12L/min, DL Temperature: 250°C, Block Temperature: 200°C, Detector voltage: 1.0kv.

Example 25: Synthesis of GGYK(ε-NH)-PEG$_{12}$-(CH$_2$)$_2$NHC(O)C$_6$H$_4$CHO (SEQ ID NO:295)



Boc-GGYK on resin was prepared in a similar manner to Example 24 using Fmoc-Lys(Dde)-Wang resin, Fmoc-Tyr(tBu)-OH Fmoc-Gly-OH and Boc-Gly-OH followed by deprotection of the ε-NHDde protecting group.

The lysine ε-NH_2 group was reacted with Fmoc-PEG_{12}-OH (252 mg, 0.3 mmol) in the presence of HATU (108 mg, 0.285 mmol) and NMM (68 µL, 0.6 mmol) in DMF (10 mL) under an N_2 atmosphere for 120 min. After washing with DMF (3x10 mL), the Fmoc group was deprotected with 20% piperidine/DMF (20 mL) under an N_2 atmosphere for 30 min. After washing with DMF (6x10 mL), the resin was suspended in DMF (10 mL) and 4-Carboxybenzaldehyde (45 mg, 0.3 mmol), HATU (108 mg, 0.285 mmol) and NMM (68 µL, 0.6 mmol) was added. The reaction proceeded under N_2 for 120 min followed by washing with DMF (3x10 mL), MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL). The resin was then air-dried (420 mg).

The peptide was then removed from the resin and Boc deprotected as described in Example 24 to give the crude peptide (87 mg) which was purified by HPLC as described in Example 24 and analysed by Mass spectrometry. Purity by HPLC: 91.4% MS: $C_{54}H_{87}B_6O_{21}^+$: Calc: 1155.59, Found: 1156.3.

MS conditions:

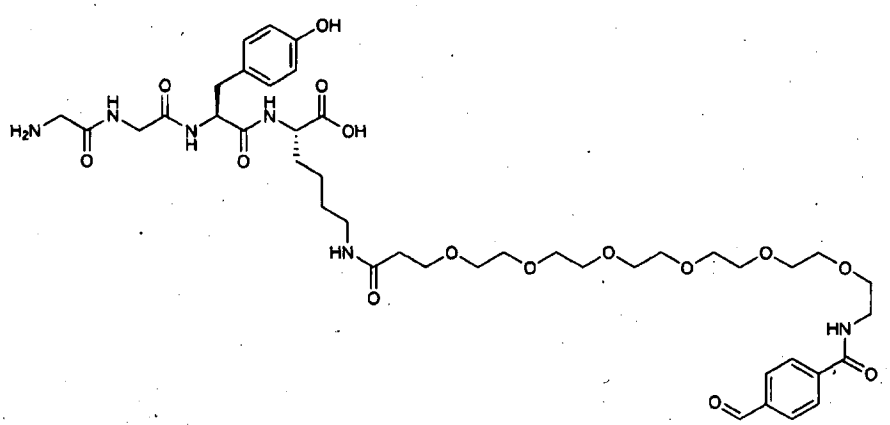
Type of machine

Waters ZQ2000

Ionization

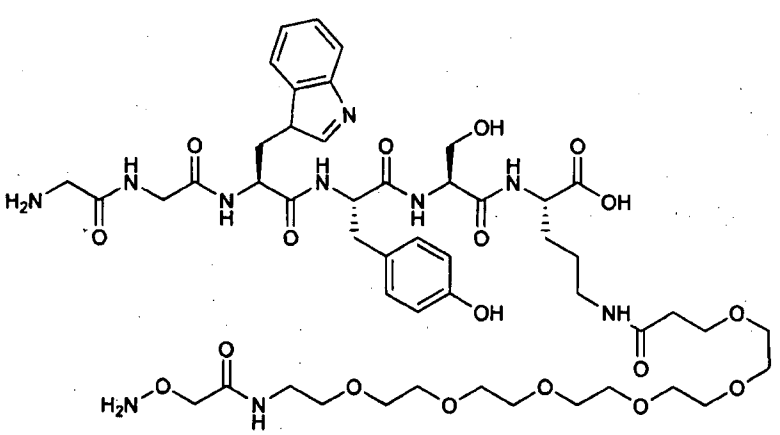
Probe: ESI, Cone: 50v, Capillary: 3.00KV, Extractor: 5v, Desovation Temp: 350, Gas Flow: 350.

Example 26: Synthesis of GGYK(ε-NH)-PEG$_6$-(CH$_2$)$_2$NHC(O)C$_6$H$_4$CHO (SEQ ID NO:296)



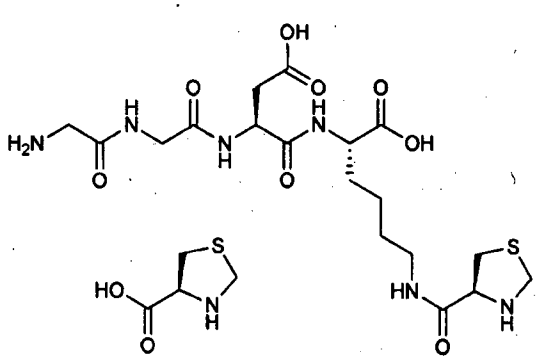
This peptide was prepared in an analogous manner to SEQ ID NO: 295 in Example 25 but using Fmoc-NH-(CH$_2$)$_2$-PEG$_6$CO$_2$H. Purity by HPLC: 94.9% MS: $C_{42}H_{63}N_6O_{15}^+$: Calc: 891.43, Found 891.76.

Example 27: Synthesis of GGWYSOrn(δ-NH)-PEG$_6$-(CH$_2$)$_2$NHC(O)CH$_2$ONH$_2$ (SEQ ID NO:337)



Boc-GGWYSOrn-δ-NHC(O)$PEG_6(CH_2)_2NH_2$ was synthesized in an analogous manner to that described in Example 25 using Fmoc-Orn(Dde)-Wang resin, Fmoc-Ser(tBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH, Boc-Gly-OH and $HOOCPEG_6(CH_2)_2NHFmoc$. The peptide in DMF (10 mL) was treated with Bis-Boc-amino-oxyacetic acid BOC_2-Aoa (163 mg, 0.56 mmol), HATU (213 mg, 0.56 mmol) and NMM (171 µL, 1.5 mmol) under an N_2 atmosphere for 120 min. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (560 mg). The peptide was then removed from the resin and Boc deprotected as described in Example 24 to give the crude peptide (213 mg) which was purified by HPLC as described in Example 24 using a Boston Crest ODS 5um 100A column and a solvent gradient of 15-40 Solvent A/30min followed by analysis by Mass spectrometry as described in Example 24. Purity by HPLC: 93% MS: $C_{49}H_{76}N_{10}O_{18}^{2+}$: Calc$(M+2H)^{2+}$: 546.26: Found: 545.75.

Example 28: Synthesis of GGDK(ε-NH)Cys (SEQ ID NO:379)

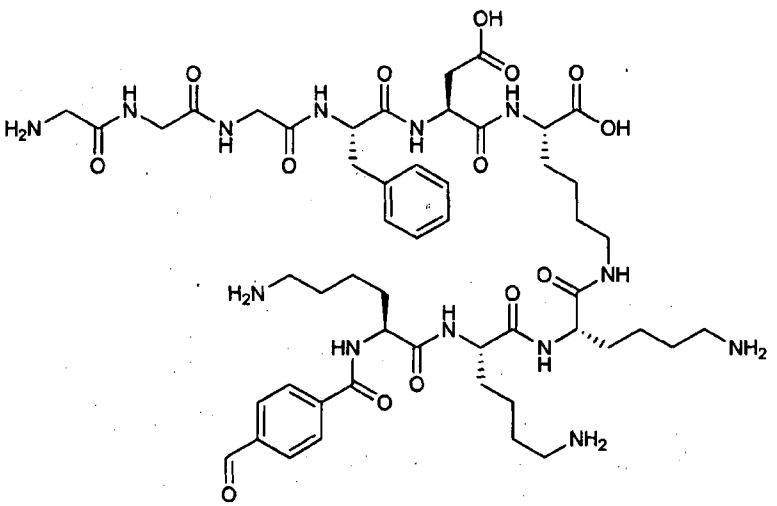


(*S*)-thiazolidine-4-carboxylic acid

Boc-GGDK(ε-NH_2) was synthesized in an analogous manner to that described in Example 25 using Fmoc-Lys(Dde)-Wang resin, Fmoc-Asp(OtBu)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGDK(ε-NH_2) in DMF (10 mL) was treated with Fmoc-thiazolidine-COOH (Fmoc-Thz-OH) (110 mg, 0.31 mmol), HATU (118 mg, 0.31 mmol) and NMM (100 µL, 0.88 mmol) under an N_2 atmosphere for 120 min then washed with DMF (3x10 mL). The resin was further washed with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (580 mg). The peptide was then removed from the resin and Boc deprotected as described in Example 24 to give the crude peptide (103 mg) which was

purified by HPLC as described in Example 24 using 4.6×250 mm, Sinochrom ODS-BP-5 column and a solvent gradient of 10-35 Solvent A/30min followed by analysis by Mass spectrometry as described in Example 24. Purity by HPLC: 97% MS: $C_{18}H_{31}N_6O_8S^+$: Calc: 491.19, Found: 491.55.

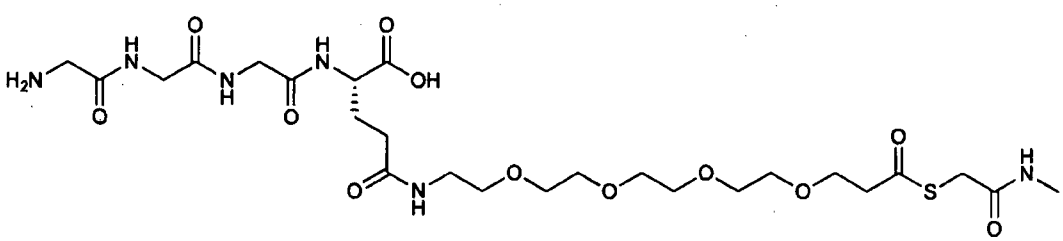
Example 29: Synthesis of GGGFDK(ε-NH)-KKK-C(O)C$_6$H$_4$CHO (SEQ ID NO: 304)



Boc-GGGFDKε-NH_2 was synthesized in a similar manner as described in Example 24 using Fmoc-Lys(Dde)-Wang resin, Fmoc-Asp(OtBu)-OH, Fmoc-Phe-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGFDKε-NH_2 in DMF (10 mL) was treated with Fmoc-Lys(Boc)-OH (141 mg, 0.3 mmol), HBTU (108 mg, 0.285 mmol) and NMM (68 µL, 0.6 mmol) under an N_2 atmosphere for 30 min and was then washed with DMF (3x10 mL). The Fmoc group was removed using 20% piperidine/DMF (20 mL) under an N_2 atmosphere for 30 min followed by washing with DMF (6x10 mL). This reaction and deprotection was repeated twice more with Fmoc-Lys(Boc)-OH to give GGGFDK(ε-NH)-KKK. GGGFDK(ε-NH)-KKK in DMF (10 mL) was then treated with 4-Carboxybenzaldehyde (45 mg, 0.3 mmol), HATU (108 mg,0.285 mmol) and NMM (68 µL, 0.6 mmol) under an N_2 atmosphere for 120 min. The resin was then washed with DMF (3x10 mL), followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (438 mg). The peptide was then removed from the resin and Boc deprotected as described in Example 24 to give the crude peptide (92 mg) which

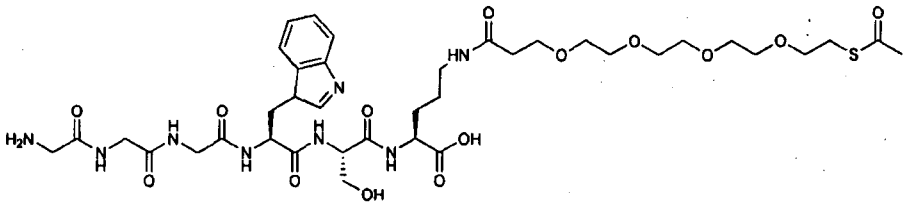
was purified by HPLC as described in Example 24 using a SP-120-10-C18-AP column and the purification gradient is 19-31 Solvent A/30min followed by analysis by Mass spectrometry as described in Example 24. Purity by HPLC: 91.9% MS: $C_{51}H_{78}N_{13}O_{14}^+$: Calc: 1096.58, Found: 1097.4.

Example 30: Synthesis of GGGE(δ-CONH)-PEG$_4$-(CH$_2$)$_2$C(O)SCH$_2$C(O)NHCH$_3$ (SEQ ID NO: 330)



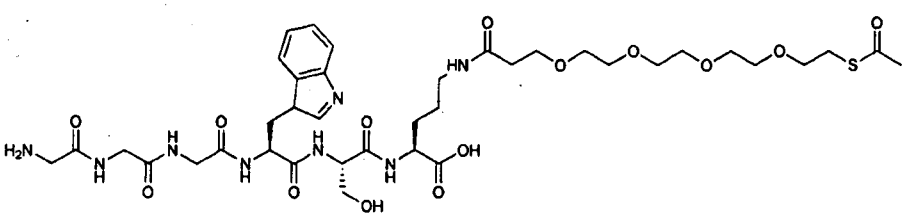
Boc-GGGE(α-OtBu)(δ-CONH) PEG_4CO_2H was synthesized from Fmoc-PEG_4-2-Cl-Frt-Cl resin using Fmoc-Glu-OtBu, Fmoc-Gly-OH and Boc-Gly-OH and standard peptide synthesis techniques. The peptide was removed from the resin using 2% TFA/DCM (20 mL) under N_2 for 5 min. The peptide was triturated with diethylether (50 mL) and the solid washed with diethylether (5 x 50 mL) and dried under vacuum to give crude Boc protected peptide (378 mg). The crude peptide in DMF (10 mL) was treated with Z-mercapto-N-methylacetamide (64 mg, 0.56 mmol), HATU (152 mg, 0.4 mmol) and NMM (122 µL, 1.07 mmol) under an N_2 atmosphere for 12 hours. The peptide was triturated with diethyl ether (150 mL) and the solid washed with diethylether (5 x 50 mL) then dried under vacuum for 12 hours to give crude peptide (252 mg). The peptide was purified by HPLC as described in Example 24 using a VYDAC-C18 5µm 100A column and a solvent gradient of 10-40 Solvent A/30 min. The peptide was analysed by mass spectrometry as described in Example 24. Purity by HPLC: 95.1% MS: $C_{25}H_{45}N_6O_{12}S^+$: Calc: 653.28, Found 653.6.

Example 31: Synthesis of GGGWSOrn(δ-NH)PEG$_4$-(CH$_2$)$_2$SAc (SEQ ID NO:321)



Boc-GGGWSOrn(δ-NH_2) on resin was synthesized in an analogous manner to Example 25 using Fmoc-Orn(Dde)-Wang resin, Fmoc-Ser(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGWSOrn(δ-NH_2) in DMF (10 mL) was added to S-acetyl-dPEG$_4$-NHS ester (63.22 mg, 0.15 mmol) and NMM (107 µL, 0.6 mmol) under an N_2 atmosphere for 30 min, followed by washing with DMF (3x10 mL), DCM (3x15 mL) and MeOH (3x 15 mL) and then air-dried (342 mg). GGGWSOrn(δ-NH)PEG$_4$ -(CH$_2$)$_2$SAc was removed from the resin and deprotected as described in Example 24 to give the crude peptide (98 mg). The crude peptide was purified by HPLC using a Venusil XBP C18(L) column and a solvent gradient of 24-49 Solvent A/30min as described in Example 24 and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 93.1% MS: $C_{38}H_{45}N_6O_{12}S^+$: Calc: 883.39, Found: 883.82.

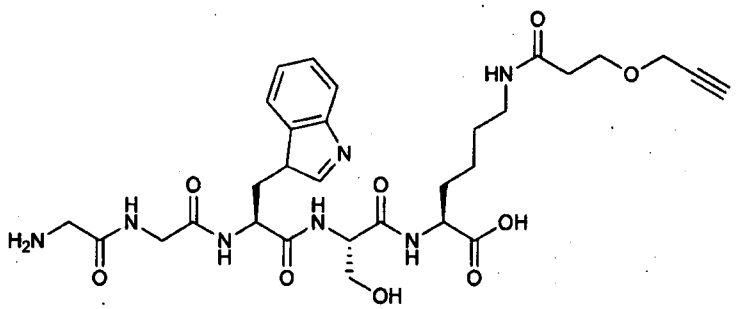
Example 32: Synthesis of GGGWSOrn(δ-NH)PEG$_4$-(CH$_2$)$_2$SAc amide



GGGWSOrn(δ-NH_2) was synthesized as described in Example 25 using Rink Amide-AM Resin, Fmoc-Orn(Dde)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. GGGWSOrn(δ-NH_2) in DMF (10 mL) was treated with S-acetyl-dPEG$_4$-NHS ester (63.22 mg, 0.15 mmol) and NMM (107 µL, 0.6 mmol) under an N_2 atmosphere for 30 min followed by washing with DMF (3x15 mL), DCM (3x15 mL) and MeOH (3x 15 mL) and then air-dried (323 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (89 mg) which was purified by HPLC as described in Example 24 using a Venusil XBP C18(L) column and a solvent gradient of 24-49 Solvent A/30min. The peptide was analysed by mass

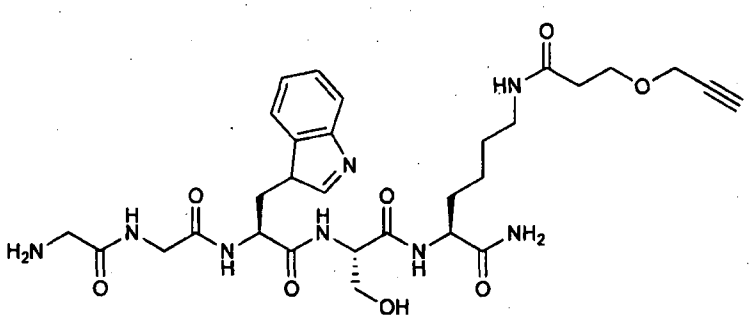
spectrometry as described in Example 25. Purity by HPLC: 90.8% MS: $C_{38}H_{60}N_9O_{13}S^+$: Calc: 882.40, Found 882.06.

Example 33: Synthesis of GGGWSK(ε-NH)C(O)$(CH_2)_2OCH_2C{\equiv}CH$ (SEQ ID NO:396)



Boc-GGGWSK(ε-NH_2) was synthesized as described in Example 25 using Fmoc-Lys(Dde)-Wang resin, Fmoc-Ser(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGWSK(ε-NH_2) in DMF (10 mL) was treated with Propargyl-dPEG®1-NHS ester (136 mg, 0.6 mmol), HATU (216 mg, 0.57 mmol) and NMM (136 µL, 0.2 mmol) under an N_2 atmosphere for 120 min. The resin was then washed with DMF (3x10 mL), MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (775 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (125 mg) which was purified by HPLC using a SP-120-10-C18-AP column and a solvent gradient of 15-25 Solvent A/30min as described in Example 24. The peptide was analysed by mass spectrometry as described in Example 25. Purity by HPLC: 96.7% MS: $C_{32}H_{45}N_8O_{10}^+$: Calc: 701.33, Found 701.22.

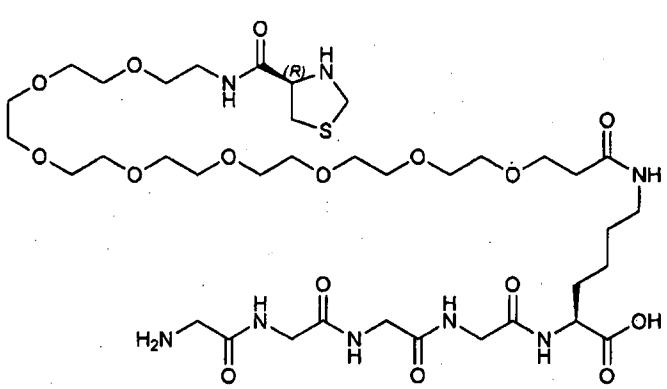
Example 34: Synthesis of GGGWSK(ε-NH)C(O)$(CH_2)_2OCH_2C{\equiv}CH$ amide



Boc-GGGWSK(ε-NH_2) was synthesized as described in Example 25 using Rink Amide MBHA resin, Fmoc-Lys(Dde)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-

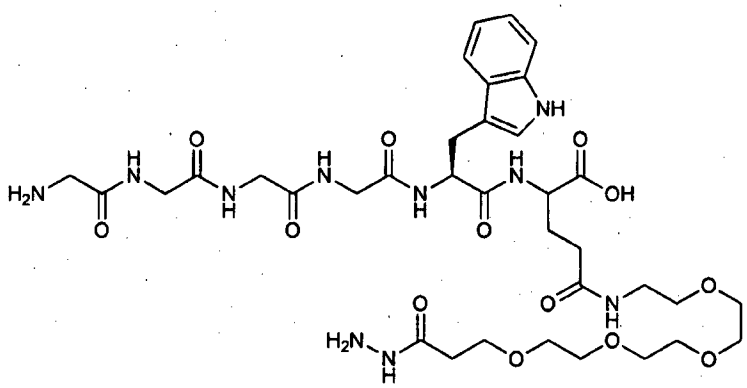
OH and Boc-Gly-OH. Boc-GGGWSK(ε-NH_2) in DMF (10 mL) was treated with Propargyl-dPEG®1-NHS ester (136 mg, 0.6 mmol), HATU (216 mg, 0.57 mmol) and NMM (136 µL, 1.2 mmol) under N_2 for 120 min. The resin was then washed with DMF (3x10 mL). Kaiser test confirmed the reaction was complete. The resin was washed with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (760 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (110 mg) which was purified by HPLC using a SP-120-10-C18-AP column and a solvent gradient of 25-35 Solvent A/30min and analysed by mass spectrometry as described in Example 24. Purity by HPLC: 93.7% MS: $C_{32}H_{46}N_9O_9^+$: Calc: 700.34, Found 700.80.

Example 35: Synthesis of GGGGK(ε-NH)-PEG$_8$-(CH$_2$)$_2$NHCO-thiazolidinone (SEQ ID NO:397)



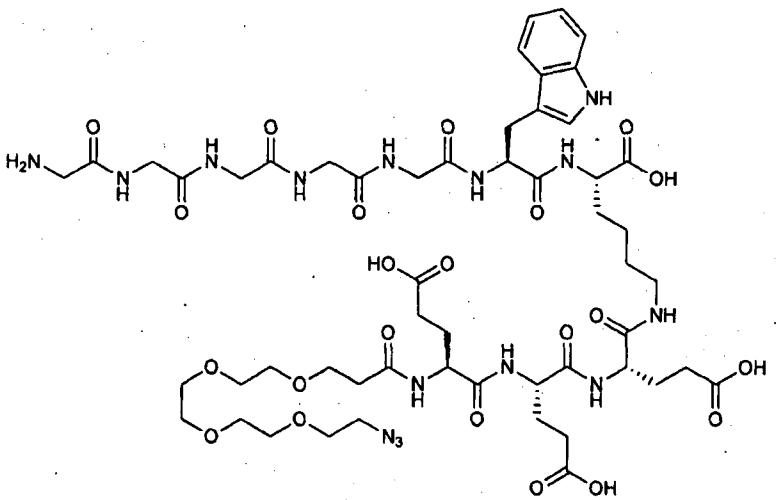
Boc-GGGGK(ε-NH-PEG_8-NH_2) was synthesized as described in Example 25 using Fmoc-Lys(Dde)-Wang Resin, Fmoc-Gly-OH, Boc-Gly-OH and Fmoc-PEG_8-OH. Boc-GGGGK(ε-NH –PEG_8-NH_2) on resin in DMF (10 mL) was added to Fmoc-Thz-OH (110 mg, 0.31 mmol), HATU (118 mg, 0.31 mmol) and NMM (100 µL, 0.88 mmol) under an N_2 atmosphere for 120 min. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (635mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (286 mg) which was purified by HPLC using a VYDAC-C18 column and a solvent gradient of 20-45 Solvent A/30min and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 93% MS: $C_{37}H_{69}N_8O_{16}S^+$: Calc: 913.45, Found 913.74.

Example 36: Synthesis of GGGGWE(δ-CONH)PEG$_4$-(CH$_2$)$_2$C(O)NHNH$_2$ (SEQ ID NO:398)



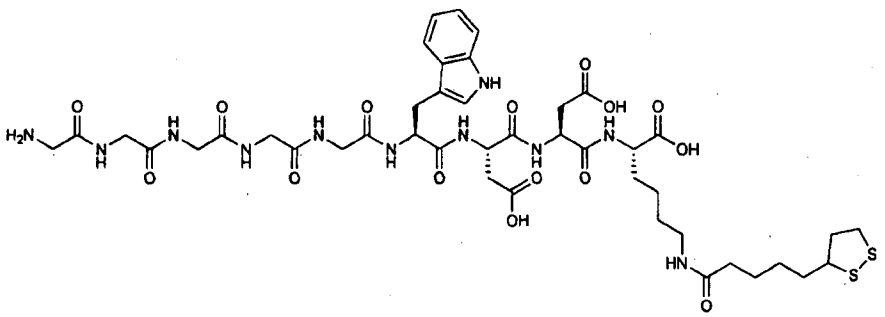
Boc-GGGGWE(δ-COOH) was synthesized as described in Example 25 using Fmoc-Glu(ODmab)-2cl (Trt) Resin, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGGWE(δ-COOH) on resin in DMF (10 mL) was added to Amino-dPEG$_4$-t-Boc-hydrazide (75.89 mg , 0.2 mmol), TBTU (96.3 mg, 0.3 mmol), DIEA (107 µL, 0.6 mmol) under an N_2 atmosphere for 30 min. The resin was then washed with DMF (3x10 mL). Kaiser test confirmed the reaction was complete. The resin was washed with DMF (3x 15 mL), DCM (3x15 mL) and MeOH (3x 15 mL) and then air-dried (442 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (105 mg) which was purified by HPLC using a kromasil C18-5 column and a solvent gradient of 15-40 Solvent A/30min and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 92.3% MS: $C_{35}H_{55}N_{10}O_{13}^+$: Calc: 823.39, Found 821.88.

Example 37: Synthesis of GGGGGWK(ε-NH)-EEE-PEG$_4$-(CH$_2$)$_2$-N$_3$ (SEQ ID NO:367)



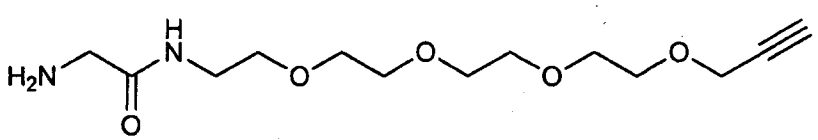
Boc-GGGGGWK(ε-NH_2)-on resin was synthesized as described in Example 25 using Fmoc-Lys(Dde)-Wang Resin, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. The Boc-GGGGGWK(ε-NH_2)-on resin in DMF (10 mL) was added to Fmoc-Glu(Otbu)-OH (128 mg, 0.3 mmol), HBTU (108 mg, 0.285 mmol), NMM (68 µL, 0.6 mmol) under an N_2 atmosphere for 30 min. The resin was then washed with DMF (3x10 mL) and deprotected with 20% piperidine/DMF (20 mL) under an N_2 for 30 min. The resin was then washed with DMF (6x10 mL). This reaction was repeated twice more to provide Boc-GGGGGWK-ε-NH-EEE-CO_2H. Boc-GGGGGWK-ε-NH-EEE-CO_2H on resin in DMF (10 mL) was added to N_3-$(CH_2)_2$-PEG_4–COOH (88 mg, 0.3 mmol), HATU (108 mg, 0.285 mmol) and NMM (68 µL, 0.6 mmol) under an N_2 atmosphere for 120 min. The resin was then washed with DMF (3x10 mL) followed by washes with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (476 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (112 mg) which was purified by HPLC using a SP-120-10-C18-AP column and a solvent gradient of 25-35 Solvent A/30min and analysed by mass spectrometry as described in Example 24. Purity by HPLC: 91.7% MS: $C_{53}H_{78}N_{17}O_{22}^+$: Calc: 1278.9, Found 1279.4.

Example 38: Synthesis of GGGGGWDDK(ε-NH)-Lipoic acid (SEQ ID NO:378)



Boc-GGGGGWDDK(ε-NH_2)-on resin was synthesized as described in Example 25 using Fmoc-Lys(Dde)-Wang Resin, Fmoc-Asp(otBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGGGWDDK(ε-NH_2)-on resin in DMF (10 mL) was added to Lipoic acid (64 mg, 0.31 mmol), HATU (118 mg, 0.31 mmol) and NMM (100 µL, 0.88 mmol) under an N_2 atmosphere for 120 min. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (745mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (340 mg) which was purified by HPLC using a Boston Crest ODS column and a solvent gradient of 21-46/30min and analysed by mass spectrometry as described in Example 24. Purity by HPLC: 90.3% MS: $C_{43}H_{62}N_{11}O_{15}S^+$: Calc: 1036.39, Found 1034.85.

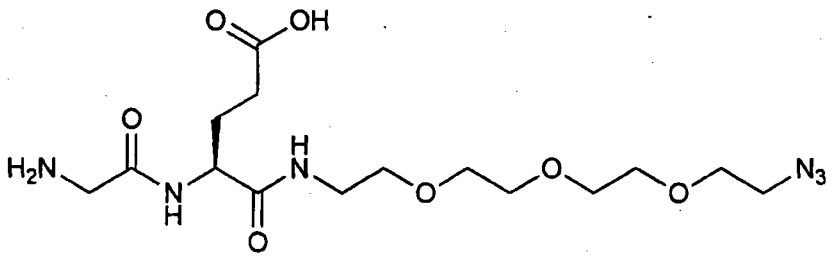
Example 39: Synthesis of G-PEG$_4$-CH$_2$C≡CH (SEQ ID NO:211)



Boc-Gly-OH (100 mg, 0.57 mmol) in DMF (10 mL) was added to NH_2-PEG_4-CH_2-C≡CH (100 mg, 0.432 mmol), HBTU (332 mg, 0.875 mmol) and DIEA (110.94 µL 0.86 mmol) and allowed to react for 24 h. Water was added to triturate the peptide which was then treated with 11.7 mL of a mixture solution (2% H_2O, 3% Tis, 95% TFA) on a shaker at 25°C for 2 hours. The mixture was filtered and 50 mL diethyl ether was added to the filtrate to triturate the peptide. The solid was washed with diethyl ether (5x50 mL) and dried under vacuum for 12 hours to give the crude peptide (105 mg). The crude peptide was dissolved in water (10 mL) and purified by HPLC using a Venusil XBP-C18 5um

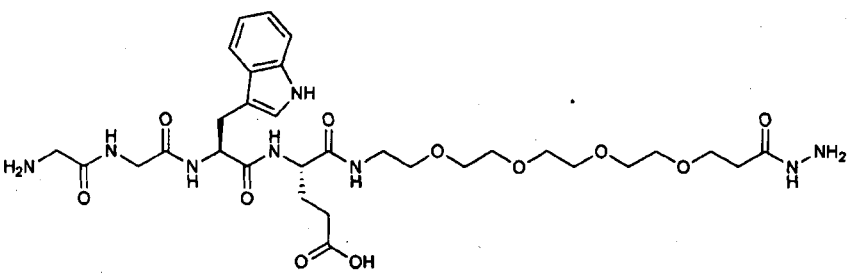
100A column and a solvent gradient of 10-35 Solvent A/30min as described in Example 24. Mass spectral analysis was also performed as in Example 24. Purity by HPLC: 93% MS: $C_{13}H_{25}N_2O_5^+$: Calc: 289.18, Found 289.25.

Example 40: Synthesis of GE-PEG$_4$-(CH$_2$)$_2$-N$_3$ (SEQ ID NO:399)



Boc-GE-(α-COOH) was synthesized on resin as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin and Boc-Gly-OH. Boc-GE-(α-COOH) on resin in DMF (10 mL) was added to NH_2-PEG_4-$(CH_2)_2$-N_3 (197 mg, 0.9 mmol), HATU (324 mg, 0.855 mmol) and NMM (204 µL, 1.8 mmol) under an N_2 atmosphere for 12 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (1086 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (102 mg) which was purified by HPLC using a Galaksil UP-C18 10um 120A column and a solvent gradient of 17-27/30min and analysed by mass spectrometry as described in Example 24. Purity by HPLC: 92.2% MS: $C_{15}H_{29}N_6O_7^+$: Calc: 405.21, Found 405.0.

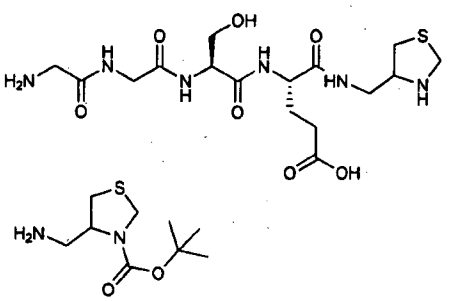
Example 41: Synthesis of GGWE-PEG$_4$-(CH$_2$)$_2$CONHNH$_2$ (SEQ ID NO:400)



Boc-GGWE-(α-COOH) was synthesized on resin as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGWE-(α-COOH) on resin in DMF (10 mL) was treated with NH_2-PEG_4-$(CH_2)_2$-

CONHNH-(Boc) (152 mg, 0.4 mmol), HATU (152 mg, 0.4 mmol) and NMM (122 µL, 1.07 mmol) under an N_2 atmosphere for 2 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (750mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (180 mg) which was purified by HPLC using a VYDAC-C18 5um 100A column and a solvent gradient of 10-35 Solvent A/30min and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 93.6% MS: $C_{31}H_{49}N_8O_{11}^+$: Calc: 709.35, Found 709.3.

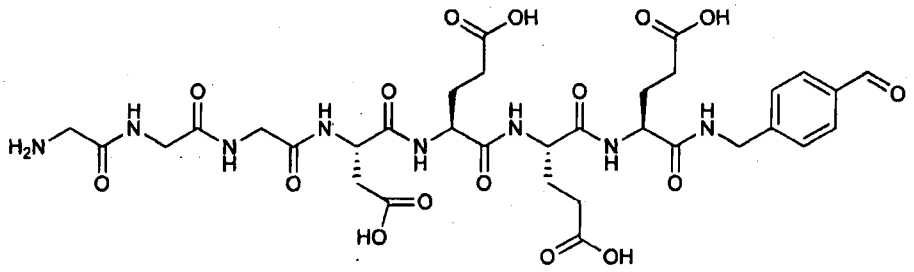
Example 42: Synthesis of GGSE-thiazolidine (SEQ ID NO:401)

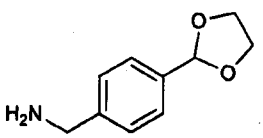


Aminomethylthiazolidine-3-carboxylic acid tert-butyl ester

Boc-GGSE-α-COOH was synthesized as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin, Fmoc-Ser(tBu)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGSE-α-COOH on resin in DMF (10 mL) was treated with (R)-4-Aminomethylthiazolidine-3-carboxylic acid tert-butyl ester (131 mg, 0.6 mmol), HATU (152 mg, 0.4 mmol) and NMM (122 µL, 1.07 mmol) under an N_2 atmosphere for 2 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and air-drying (780mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (210 mg) which was purified by HPLC using a VYDAC-C18 5um 100A column and a solvent gradient of 2-25 Solvent A/30min and analysed by mass spectrometry as described in Example 24. Purity by HPLC: 93.1% MS: $C_{16}H_{29}N_6O_7S^+$: Calc: 449.18, Found 449.4.

Example 43: Synthesis of GGGD-EEE-CONHCH$_2$C$_6$H$_4$CHO (SEQ ID NO:402)

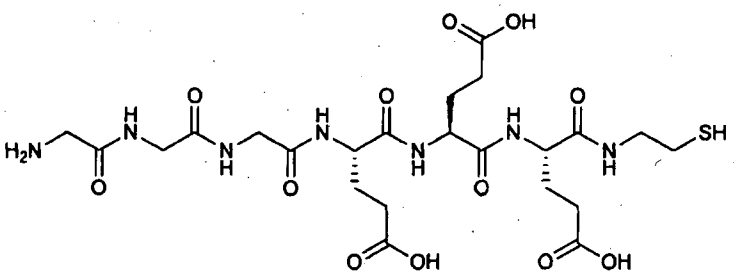




(4-(1.3-Dioxolan-2-yl)phenyl)methanamine

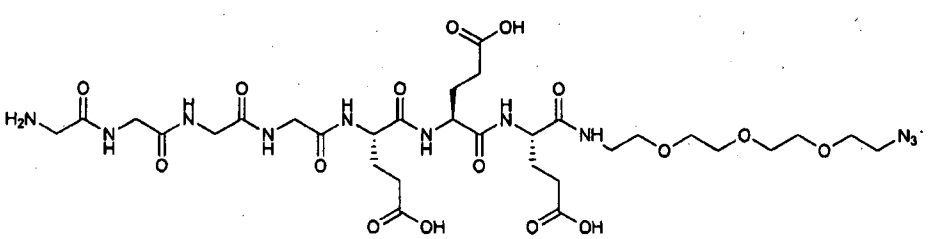
Boc-GGGDEEE-α-COOH was synthesized as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin, Fmoc-Asp(OtBu)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGDEEE-α-COOH on resin in DMF (10 mL) was treated with (4-(1.3-Dioxolan-2-yl)phenyl)methanamine (162 mg, 0.9 mmol), HATU (324 mg, 0.855 mmol) and NMM (204 µL, 1.8mmol) under an N_2 atmosphere for 12 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (1170 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (205 mg) which was purified by HPLC using a Galaksil UP-C18 10um 120A column and a solvent gradient of 11-21 Solvent A/30min and analysed by mass spectrometry as described in Example 24. Purity by HPLC: 92.8% MS: $C_{33}H_{43}N_8O_{16}^-$: Calc: 807.28, Found 807.6.

Example 44: Synthesis of GGG-EEE-NH(CH$_2$)$_2$SH (SEQ ID NO:403)



Boc-GGG-EEE-α-COOH was synthesized as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin, Fmoc-Glu(OtBu)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGG-EEE-α-COOH on resin DMF (10 mL) was treated with (Trt) S-CH_2CH_2-NH_2 (259 mg, 0.9 mmol), HATU (324 mg, 0.855 mmol) and NMM (204 µL, 1.8 mmol) under an N_2 atmosphere for 12 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (1170 mg). The peptide was removed from the resin and deprotected as described by treatment with 11.7 mL of a mixture solution (2.5% H_2O, 2.5% phenol, 2.5% EDT, 5% methylphenylsulfide, 87.5% TFA) on a shaker at 25°C for 2 hours. The above mixture was filtered and the filtrate was treated with 50 mL diethyl ether to triturate the peptide. The solid was washed with diethyl ether (5x50 mL) and then dried under vacuum for 12 hours to give the crude peptide (165 mg) which was purified by HPLC as described in Example 24 using a SP-120-10-C18-AP column and a solvent gradient of 20-30/30min. Mass spectral analysis was also performed as described in Example 25. Purity by HPLC: 94.1% MS: $C_{23}H_{38}N_7O_{12}S^+$: Calc: 636.23, Found 636.25.

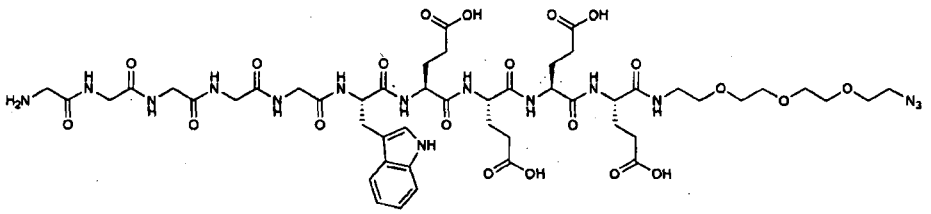
Example 45: Synthesis of GGGG-EEE-PEG$_4$-(CH$_2$)$_2$-N$_3$ (SEQ ID NO:404)



Boc-GGGG-EEE-α-COOH was synthesized as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin, Fmoc-Glu(OtBu)-OH, Fmoc-Gly-OH and Boc-Gly-OH. Boc-GGGG-EEE-α-COOH on resin in DMF (10 mL) was treated with NH_2-PEG_4-$(CH_2)_2$-N_3 (197 mg, 0.9 mmol), HATU (324 mg, 0.855 mmol) and NMM (204 µL, 1.8mmol) under an N_2 atmosphere for 12 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (1150mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (189 mg) which was purified by HPLC using a Galaksil UP-C18 10um 120A column and a solvent gradient of 14-24 Solvent

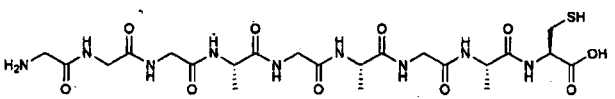
A/30min and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 90.5% MS: $C_{31}H_{52}N_{11}O_{16}^+$: Calc: 834.36, Found 834.41.

Example 46: Synthesis of GGGGGWE-EEE-PEG$_4$-(CH$_2$)$_2$-N$_3$ (SEQ ID NO:405)



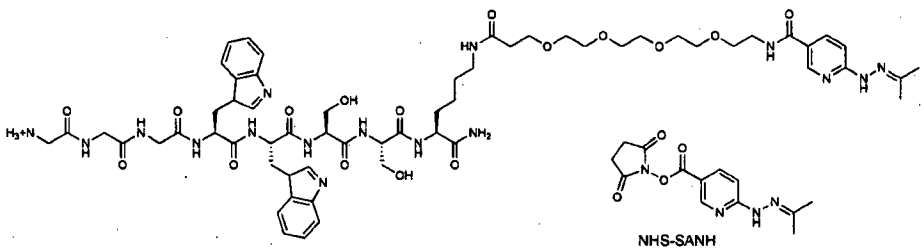
GGGGGWE-EEE-α-COOH was synthesized as described in Example 24 using Fmoc-Glu-α-ODmab Wang Resin, Fmoc-Glu(OtBu)-OH, Fmoc-Trp(Boc)-OH, Fmoc-Gly-OH and Boc-Gly-OH. GGGGGWE-EEE-α-COOH on resin in DMF (10 mL) was treated with NH_2-PEG_4-$(CH_2)_2$-N_3 (197 mg, 0.9 mmol), HATU (324 mg, 0.855 mmol) and NMM (204 µL, 1.8 mmol) under an N_2 atmosphere for 12 hours. The resin was then washed with DMF (3x10 mL) followed by washing with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-drying (1250mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide (257 mg) which was purified by HPLC using a Galaksil UP-C18 10um 120A column and a solvent gradient of 18-28 Solvent A/30min and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 95.6% MS: $C_{49}H_{70}N_{15}O_{21}^-$: Calc: 1204.49, Found 1204.86.

Example 47: Synthesis of GGGAGAGAC (SEQ ID NO:276)



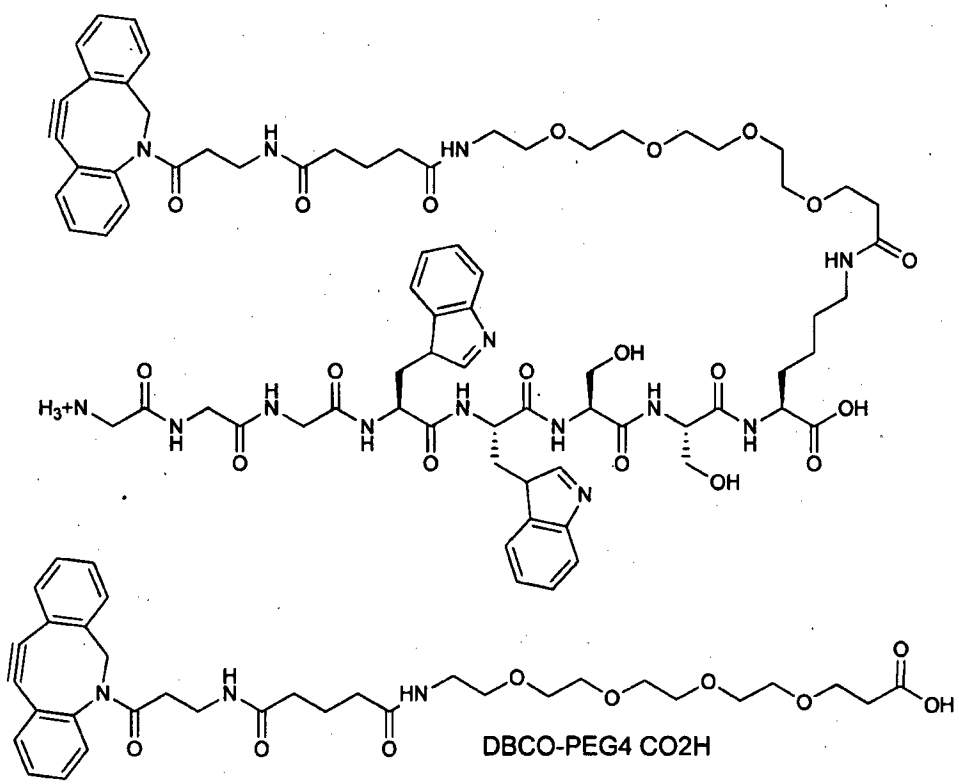
This peptide was prepared and cleaved from the resin according to the standard procedures. The crude peptide (257 mg) was purified by HPLC as described in Example 24 using a Galaksil UP-C18 10um 120A column and the solvent gradient was 18-28/30min. Mass spectral analysis was performed as for Example 25. Purity by HPLC: 100%. MS: $C_{22}H_{38}N_9O_{10}S^+$: Calc: 620.25, Found: 619.15.

Example 48: Synthesis of GGGWWSSK-(ε-NH)-PEG$_4$-(CH$_2$)$_2$NHCOC$_5$H$_4$N-NHN=C(CH$_3$)$_2$ (SANH) (SEQ ID NO:277)



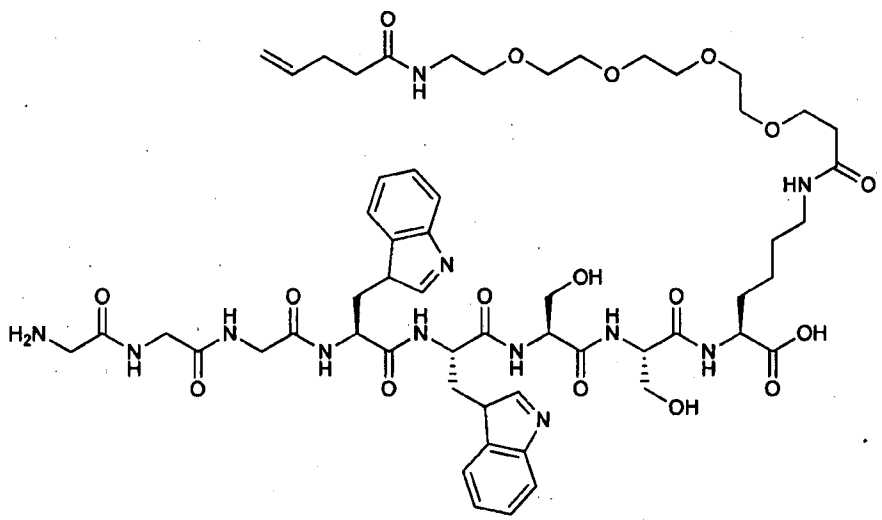
Boc-GGGWWSSK-(ε-NH$_2$) on resin was prepared and the Dde protecting group of the lysine ε-NH$_2$ moiety deprotected according to standard procedures. Boc-GGGWWSSK-(ε-NH$_2$) on resin in DMF (10 mL) was treated with Fmoc-PEG$_4$-CO$_2$H (0.3 mmol), HATU (108 mg, 0.285 mmol), NMM (68 μL, 0.6 mmol) under an N$_2$ atmosphere for 120 min followed by washing with DMF (3x10 mL). The Fmoc group was removed with 20% piperidine/DMF (20 mL) and the resin was washed with DMF (6x10 mL). Boc-GGGWWSSK-(ε-NH$_2$)PEG$_4$-NH$_2$ in DMF (10 mL) and treated with NHS-SANH (0.3 mmol) and NMM (68 μL, 0.6 mmol) under an N$_2$ atmosphere for 120 min followed by washing with DMF (3x10 mL), then MeOH (1x10 mL), DCM (3x10mL) and MeOH (2x10 mL) and then air-dried (400 mg). The peptide was removed from the resin and deprotected as described in Example 24 to give the crude peptide. The crude peptide (257 mg) was purified by HPLC as described in Example 24 using a Galaksil UP-C18 10um 120A column and a solvent gradient of 15-50 Solvent A/30min. Mass spectral analysis was performed as described in Example 25. Purity by HPLC: 95.4%. MS: $C_{60}H_{85}N_{16}O_{16}^+$: Calc: 1285.63, Found: 1285.4.

Example 49: Synthesis of GGGWWSSK-(ε-NH)-PEG$_4$-DBCO (SEQ ID NO:278)



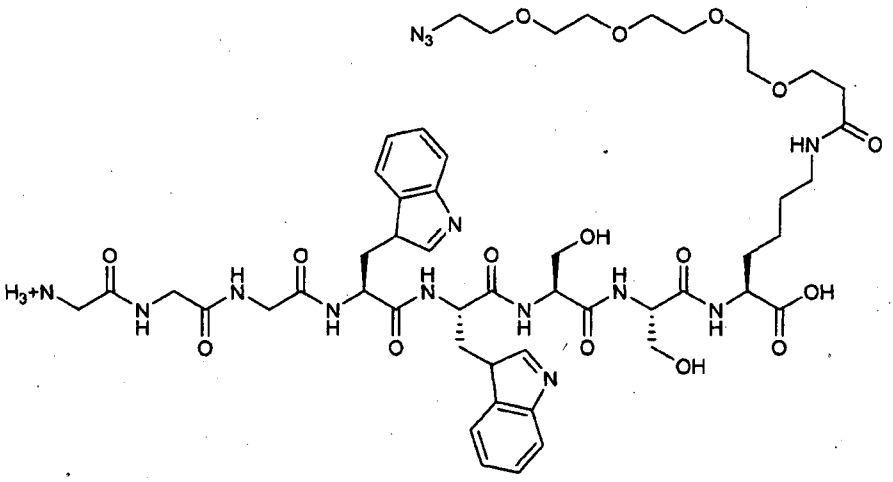
Boc-GGGWWSSK-(ε-NH_2) was prepared and the Dde protecting group of the lysine ε-NH_2 moiety deprotected according to standard procedures. Boc-GGGWWSSK-(ε-NH_2) in DMF (10 mL) was treated with DBCO-PEG_4-CO_2H (0.3mmol), HATU (108 mg, 0.285 mmol), NMM (68 µL, 0.6 mmol) under an N_2 atmosphere for 120 min followed by washing with DMF (3x10 mL) then MeOH (1x10 mL), DCM(3x10 mL) and MeOH (2x10 mL) and then air-dried (400 mg). The peptide was removed from the resin and deprotected with 5%H_2O, 95%TFA) at 25°C for 2 hours, followed by filtration, to which 50 mL diethyl ether was added to triturate the peptide. The solid was washed with diethyl ether (5x50 mL) and dried under vacuum for 12 hours to give the crude peptide. The crude peptide (257mg) was purified as described in Example 24 using a Galaksil UP-C18 10um 120A column and a solvent gradient of 15-50 Solvent A/30min. Mass spectral analysis was carried out as described in Example 25. Purity by HPLC: 95.3%. MS: $C_{74}H_{95}N_{14}O_{19}^+$: Calc: 1483.69, Found 1485.3.

Example 50: Synthesis of GGGWWSSK-(ε-NH)-PEG$_4$-Pentenoic acid (SEQ ID NO:279)



Boc-GGGWWSSK-(ε-NH_2) was prepared and the Dde protecting group of the lysine ε-NH_2 moiety deprotected according to standard procedures. Boc-GGGWWSSK-(ε-NH_2) on resin in DMF (10 mL) was treated with Fmoc-PEG_4-CO_2H (0.3 mmol), HATU (108 mg, 0.285 mmol), NMM (68 µL, 0.6mmol) under N_2 for 120 min. The resin was then washed with DMF (3x10 mL) and the Fmoc group removed. The resin in DMF (10 mL) was treated with 4-Pentenoic acid (0.3mmol), HATU (108 mg, 0.285 mmol) and NMM (68 µL, 0.6 mmol) under N_2 for 120 min followed by with DMF (3x10 mL) then with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (400 mg). The peptide was removed from the resin and deprotected as described in Example 48 to give the crude peptide. The crude peptide (257mg) was purified and analysed by mass spectroscopy as described in Example 48. Purity by HPLC: 97.2%. MS: $C_{56}H_{79}N_{14}O_{17}^-$, Calc: 1191.57. Found: 1192.15.

Example 51: Synthesis of GGGWWSSK-(ε-NH)-PEG$_4$-(CH$_2$)$_2$N$_3$ (SEQ ID NO:77)



GGGWWSSK-(ε-NH_2) was prepared and the Dde protecting group of the lysine ε-NH_2 moiety deprotected according to standard procedures. GGGWWSSK-(ε-NH_2) on resin in DMF (10 mL) was treated with N_3-PEG_4-CO_2H (0.3mmol), HATU (108 mg, 0.285 mmol), NMM (68 µL, 0.6mmol) under N_2 for 120 min. The resin was then washed with DMF (3x10 mL) then with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (400 mg). The peptide was removed from the resin, deprotected, purified and analysed by mass spectrometry as described in Example 48. Purity by HPLC: 96.7%. MS: $C_{51}H_{73}N_{14}O_{16}^+$: Calc: 1137.53, Found: 1138.58.

Example 52: Recombinant preparation of LPETG tagged proteins and sortase proteins

Cloning and DNA preparation

DNA coding for the protein to be expressed was ordered from Geneart/Invitrogen optimised for the expression host (see table). All constructs were transformed into NEB Turbo Competent *E. coli* (High Efficiency) (C2984) and grown at 37°C under shaking in the presence of ampicillin (100 µg/mL) overnight. DNA was extracted using QIAprep Spin Miniprep Kit (Qiagen) and digested at 37°C for 3h with the corresponding Enzymes (NEB, see table).

Samples were run on 0.8% agarose gel and digested fragments extracted from the gel using QIAquick Gel Extraction Kit (Qiagen). Inserts were ligated into the corresponding vectors (see table) over night at 4°C using T4 DNA Ligase (NEB). Ligated vectors were

transformed into NEB Turbo Competent *E. coli* (High Efficiency) (Qiagen) and positive colonies selected by PCR. DNA was prepared using QIAprep Spin Miniprep Kit (Qiagen) for construct to be expressed in *E. coli* and EndoFree Plasmid Maxi Kit (Qiagen) for constructs to be expressed in S2 insect cells, HEL293 or CHO cells. DNA for *E. coli* expression was transformed into BL21 Star™ Chemically Competent Cells for expression.

Protein production

in *E. coli* (BL21DE3*)

Confirmed clones are plated on ampicillin (100 µg/mL) containing agar plates and incubated at 37°C overnight. One clone is inoculated with 5 mL of LB broth with ampicillin (100 µg/mL) and grown overnight at 37°C under shaking. The starter culture is transferred to a 1L production culture and expanded until the OD reaches 0.6-0.8 units. Protein production is started by adding 1 mM IPTG for 4h. Bacteria are lysed (BugBuster, Novagen), spun down (15,000 g for 15 min at 4°C) and the dialysed (against PBS) supernatant purified via FPLC (IMAC).

in S2 insect cells

Drosophila S2 cells (Invitrogen, USA) are transfected with a mixture of DNA and DDAB. Cells are diluted to 2mill cells/mL and mixed with 80 ng/mL DNA preincubated with 250 ng/mL dimethyldioctadecylammonium bromide for 20 min. The cells are then cultured in Express Five SFM medium containing 18 mmol/L L-glutamine and 1% penicillin/streptomycin at 28°C in ventilated polycarbonate Erlenmeyer flasks (Corning, Acton, MA, USA) under constant rotation (100 rpm, Bench top Orbital Shaker Incubator, Ratek Instruments, Australia). Five days later, the cell supernatant is collected by centrifugation at 15,000 g for 15 min and the dialysed (against PBS) supernatant purified via FPLC (IMAC).

in HEK293 and CHO mammalian cells

HEK 293 and CHO cells are transfected using Lipofectamine™ 2000 (invitrogen). Briefly, 30 µg of each DNA were incubated with 75 µg Lipofectamine in 3.6 mL of Opti-MEM® Reduced Serum Medium, GlutaMAX™ (invitrogen). DNA is mixed and added to adherent

HEK and CHO cells in a 75 cm^2 flask in 18 mL of Opti-MEM® Reduced Serum Medium, GlutaMAX™ (invitrogen). After 5 hours the medium was replaced with Gibco® DMEM (invitrogen) containing 10% FCS. Cells were grown for 48h, supernatant was harvested and FPLC purified (Protein G affinity and FLAG tag purification).

Protein purification

Dialysed supernatant (S2 insect cells or HEK293 cells) and bacterial lysate was purified on an automated FPLC system (Biorad Duoflow) using a 5 mL Co-NTA column (clonetech) at a flow rate of 5 mL per min. Supernatant was applied and the column then washed with 10 column volumes of wash buffer (0.5 mol/L NaCl, 20 mmol/L imidazole in 50 mM Tris, pH 8.0) and protein eluted with elution buffer (0.5 mol/L NaCl, 250 mmol/L imidazole in 50 mM Tris, pH 8.0). Fractions with high protein concentration as measured at 280 nm were pooled and dialysed against PBS (without Ca/Mg). A summary of the protein preparation is given in Table 3:

Table 3

Protein (all LPETG) enzymes	Size (kD)	Vector	Production system	Restriction
SE scFv	35	pSecTag2A	HEK293	NotI/EcoRI
SCE5 scFv	35	pSecTag2A	HEK293	NotI/EcoRI
mutMA2 scFv	36	pSecTag2A	HEK293	NotI/EcoRI
IL-11	25	pET20b+	BL21DE3*	NcoI/XhoI
IL1-ra	21	pET20b+	BL21DE3*	NcoI/XhoI
Streptavidin	17	pET20b+	BL21DE3*	NcoI/XhoI
eGFP	29	pET20b+	BL21DE3*	NcoI/XhoI
Sortase B	22	pET20b+	BL21DE3*	NcoI/XhoI
Sortase A	20	pQE30	BL21DE3*	BamHI
IgG 528 Vh	53	pEE6.4	CHO/HEK293	HindIII/ApaI
IgG 528 Vl	28	pEE12.4	CHO/HEK293	HindIII/RsrII

Peroxidase	38	pAC-M3	S2 cells	NotI/NcoI
mouse i-Domain	32	pHOG	NEB turbo	NcoI/XhoI

The sequences of the protein constructs are:

Sortase A (SEQ ID NO: 406)

MASSHHHHHHDYDIPTTENLYFQGSQAKPQIPKDKSKVAG
YIEIPDADIKEPVYPGPATPEQLNRGVSFAEENESLDDQNIS
IAGHTFIDRPNYQFTNLKAAKKGSMVYFKVGNETRKYKMT
SIRDVKPTDVGVLDEQKGKDKQLTLITCDDYNEKTGVWEK
RKIFVATEVK

Sortase B (SEQ ID NO:407)

MVVRDELRDLQKLNKDMVGWLTIIDTEIDYPILQSKDNDY
YLHHNYKNEKARAGSIFKDYRNTNEFLDKNTIIYGHNMKD
GSMFADLRKYLDKDFLVAHPTFSYESGLTNYEVEIFAVYE
TTTDFYYIETEFPETTDFEDYLQKVKQQSVYTSNVKVSGK
DRIITLSTCDTEKDYEKGRMVIQGKLVTK

Peroxidase (SEQ ID NO:408)

AMAMQLTPTFYDNSCPNVSNIVRDTIVNELRSDPRIAASIL
RLHFHDCFVNGCDASILLDNTTSFRTEKDAFGNANSARGF
PVIDRMKAAVESACPRTVSCADLLTIAAQQSVTLAGGPSW
RVPLGRRDSLQAFLDLANANLPAPFFTLPQLKDSFRNVGL
NRSSDLVALSGGHTFGKNQCRFIMDRLYNFSNTGLPDPTL
NTTYLQTLRGLCPLNGNLSALVDFDLRTPTIFDNKYYVNL
EEQKGLIQSDQELFSSPNATDTIPLVRSFANSTQTFFNAFVE
AMDRMGNITPLTGTQGQIRLNCRVVNSNSGKPIPNPLLGL
DSTLPETGGLEHHHHHHHH

IL-11 (SEQ ID NO:409)

MVNCVCRLVLVVLSLWPDRVVAPGPPAGSPRVSSDPRADL
DSAVLLTRSLLADTRQLAAQMRDKFPADGDHSLDSLPTLA
MSAGTLGSLQLPGVLTRLRVDLMSYLRHVQWLRRAGGPS
LKTLEPELGALQARLERLLRRLQLLMSRLALPQAAPDQPVI
PLGPPASAWGSIRAAHAILGGLHLTLDWAVRGLLLLKTRL
GKPIPNPLLGLDSTLPETGGLEHHHHHHHH

IL1-ra (SEQ ID NO:410)

MVASEAACRPSGKRPCKMQAFRIWDTNQKTFYLRNNQLIA
GYLQGPNIKLEEKLDMVPIDLHSVFLGIHGGKLCLSCAKSG
DDIKLQLEEVNITDLSKNKEEDKRFTFIRSEKGPTTSFESAA
CPGWFLCTTLEADRPVSLTNTPEEPLIVTKFYFQEDQGKPI
PNPLLGLDSTLPETGGLEHHHHHHHH

Streptavidin (SEQ ID NO:411)

MVAEAGITGTWYNQLGSTFIVTAGADGALTGTYESAVGN
AESRYVLTGRYDSAPATDGSGTALGWTVAWKNNYRNAHS
ATTWSGQYVGGAEARINTQWLLTSGTTEANAWKSTLVGH
DTFTKVKPSAASGKPIPNPLLGLDSTLPETGGLEHHHHHHH
H

mouse i-Domain (SEQ ID NO:412)

MKYLLPTAAAGLLLLAAQPAMAMVLRPPQQFPEALRECPQ
QESDIVFLIDGSGSINNIDFQKMKEFVSTVMEQFKKSKTLF
SLMQYSDEFRIHFTFNDFKRNPSPRSHVSPIKQLNGRTKTA
SGIRKVVRELFHKTNGARENAAKILVVITDGEKFGDPLDY
KDVIPEADRAGVIRYVIGVGNAFNKPQSRRELDTIASKPAG
EHVFQVDNFEALNTIQNQLQEKIFAIEGTQTGSTSSFEHEM

SQEGFSASRGGPEQKLISEEDLNSAVDLPETGGEAAALEHH
HHHH

IgG 528 light chain (SEQ ID NO:413)

MKLPVRLLVLMFWIPASSSDVLMTQTPLSLPVSLGDQASIS
CRSSQNIVHNNGITYLEWYLQRPGQSPKLLIYKVSDRFSGV
PDRFSGSGSGTDFTLKISRVEAEDLGIYYCFQGSHIPPTFGG
GTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYP
REAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL
SKADYEKHKVYACEVTHQGLSSPVTKSFNRGECLPETGGD
YKDDDDK

IgG 528 heavy chain (SEQ ID NO:414)

MGWSCIILFLVATATGAHSQVQLQQSGSEMARPGASVKLP
CKASGDTFTSYWMHWVKQRHGHGPEWIGNIYPGSGGTNY
AEKFKNKVTLTVDRSSRTVYMHLSRLTSEDSAVYYCTRSG
GPYFFDYWGQGTSLTVSSASTKGPSVFPLAPSSKSTSGGTA
ALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLY
SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCD
KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCV
VVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYR
VVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG
QPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEW
ESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGN
VFSCSVMHEALHNHYTQKSLSLSPGKLPETGGDYKDDDDK

SE scFv (SEQ ID NO:415)

DAAQPARRAVRSLVPSSDPLQCGGILEVQLVESGGGLVQP
GGSLRLSCAASGFMFSRYAMSWVRQAPGKGPEWVSGISGS
GGSTYYADSVKGRFTVSRDNSKNTLYLQMNSLRAEDTAV
YYCARIFTHRSRGDVPDQTSFDYWGQGTLVTVSSGSASAP

KLEEGEFSEARVSSELTQDPAVSVALGQTVRITCQGDSLRN
FYASWYQQKPGQAPTLVIYGLSKRPSGIPDRFSASSSGNTA
SLTITGAQAEDEADYYCLLYYGGGQQGVFGGGTKLTVLRG
KPIPNPLLGLDSTLPETGGLEEAAARGGPEQKLISEEDLNS
AVDHHHHHH

SCE5 scFv (SEQ ID NO:416)

DAAQPARRAVRSLVPSSDPLQCGGILQVQLQESGGGLVQP
GGSLRLSCAASGFMFSRYAMSWVRQAPGKGPEWVSGISGS
GGSTYYADSVKGRFTVSRDNSKNTLYLQMNSLRAEDTAV
YYCARGATYTSRSDVPDQTSFDYWGQGTLVTVSSGSASAP
KLEEGEFSEARVSELTQDPAVSVALGQTVRITCQGDSLRNF
YASWYQQKPGQAPTLVIYGLSKRPSGIPDRFSASSSGNTAS
LTITGAQAEDEADYYCLLYY.GGGQQGVFGGGTKLTVLRG
KPIPNPLLGLDSTLPETGGLEEAAARGGPEQKLISEEDLNS
AVDHHHHHH

Egfp (SEQ ID NO:417)

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY
GKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMK
QHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL
VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKN
GIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELY.KL
PETGGLEHHHHHH

Example 53: Conjugation of peptides with SE scFv tagged protein and GFP-tagged protein

Sortase reaction

LPETG proteins (60 μM) were modified with a peptides (180 μM) under Sortase catalysis (180 μM) for 5 hours at 37°C under shaking (900 rpm) in the presence of 0.5 mmol/L

$CaCl_2$ in Tris reaction buffer (50 mmol/L Tris, 150 mmol/L sodium chloride, pH 8). Non-reacted LPETG protein and Sortase were removed by addition of 200 µL Co-NTA beads (Talon, Clonetech) for 3 h under shaking at 4°C. The beads were spun down and the supernatant containing the modified protein dialysis against PBS (10kD cut-off) overnight at 4°C to remove non-reacted peptide. Coupling efficacy was determined by protein concentration, SDS-PAGE and MALDI. Functional integrity was determined by flow cytometry.

Determination of protein concentration

A bicinchoninic acid assay (BCA) was used to estimate concentration of the purified, post-dialysed samples. A Pierce BCA protein assay kit from Thermo Scientific (Prod # 23227, Lot # MG159654) was used. Standards were pre-prepared at concentrations of 1, 0.5, 0.25, 0.125, 0.0625, 0.03125, 0.0156 and 0.0075 mg/mL in phosphate buffered saline (without Ca^{2+}/Mg^{2+}) from the 2 mg/mL vial of bovine serum albumin (BSA, Pierce #23209) provided in the kit. 4 mL of working reagent (WR) was prepared according to the provided instructions from reagents A and B provided in the kit. 200 uL of WR was added to each well of a Costar 96 well cell culture cluster flat bottom plate (Corning Incorporated). 25 µL of each standard was added to each well. 25 µL of each unknown sample was added to each well. The plate was placed in a 37°C incubator for 30 minutes and read at 560 nm in a Perkin Elmer Victor microplate reader (Perkin Elmer, 1420 Multilabel Counter, Victor 3V) and analysed using a Wallac 1420 Workstation software package.

SDS-PAGE

30 µL of each sample and 6 µL of 5X reducing SDS loading buffer were added to 1.5-mL tube and denatured at 96°C for 5 min. 36 µL of each sample was run on 12 % SDS-PAGE gel in SDS running buffer at 30 mA for 2 hours. The gel was then stained with Coommassie Brilliant Blue for 1 hour and subsequently destained for at least 12 hours with Coommassie destaining solution. The gel was visualised and analysed using a BioRad Gel-Doc system with Quantity One software to confirm protein conjugation had occurred. Protein concentration determination before and after reaction was used to determine % efficacy. The results are given in Tables 4 and 5.

Table 4: SE scFv (SEQ ID NO:415) conjugation with peptides

Sortase studies SE

SEQ ID NO	Peptide	Starting amount	Recovered	%
22	GGGK-εPEG$_4$-N$_3$	1 mg	0.367	37
55	GGGWWK-εPEG$_4$-N$_3$	1 mg	0.396	40
77	GGGWWSSK-εPEG$_4$-N$_3$	1 mg	0.393	39
386	GGGWWGA-pG	1 mg	0.329	33
276	GGGAGAGAC	1 mg	0.427	43
277	GGGWWSSK-PEG$_4$-SANH	1 mg	0.289	29
278	GGGWWSSK-PEG$_4$-DBCO	1 mg	0.499	50
279	GGGWWSSK-PEG4-pentenoic acid	1 mg	0.421	42
198	GK(ε-NH)-PEG$_4$-Azide	1 mg	0.389	39
367	GGGGGWK(ε-NH)-EEE–PEG$_4$-N$_3$	1 mg	0.426	43
282	GGGWSK(ε-NH)-Alkyne	1 mg	0.405	41
211	G-PEG$_4$-Alkyne	1 mg	0.396	40
296	GGYK(ε-NH)-PEG$_6$-Aldehyde	1 mg	0.313	31
295	GGYK(ε-NH)-PEG$_{12}$-Aldehyde	1 mg	0.291	29
304	GGGFDK(ε-NH)-KKK-Aldehyde	1 mg	0.320	32
321	GGGWSOrn(δ-NH)-PEG$_4$-SAc	1 mg	0.339	34
378	GGGGGWDDK(ε-NH)-Lipoic acid	1 mg	0.370	37
379	GGDK(ε-NH)-Thiazolidine	1 mg	0.393	39
362	GGGGK(ε-NH)-PEG$_8$-Thiazolidine	1 mg	0.327	33
291	GGG-EEE-SH	1 mg	0.483	48
292	GGWE-PEG$_4$-hydrazine	1 mg	0.497	50
293	GGGWSK(ε-NH)-Alkyne	1 mg	0.425	42
333	GGGE(δ-CONH)-PEG$_4$-Thioester	1 mg	0.341	34
337	GGWYSOrn (δ-NH)-PEG$_6$-ONH$_2$	1 mg	0.343	34
404	GGGG EEE-PEG$_4$-Azide	1 mg	0.403	40
297	GE-PEG$_4$-Azide	1 mg	0.440	44
298	GGSE-Thiazolidine	1 mg	0.488	49
349	GGGGWE(δ-CONH)PEG$_4$-hydrazine	1 mg	0.447	45

Table 5: GFP (SEQ ID NO:418) conjugation with peptides

Sortase studies GFP

SEQ ID NO	Peptide	Starting amount	Recovered	%
22	GGGK-εPEG$_4$-N$_3$	1 mg	0.87	87
55	GGGWWK-εPEG$_4$-N$_3$	1 mg	0.65	65
77	GGGWWSSK-εPEG$_4$-N$_3$	1 mg	0.82	82

	Sortase studies GFP			
386	GGGWWGA-pG	1 mg	0.79	79
276	GGGAGAGAC	1 mg	0.79	79
277	GGGWWSSK-PEG_4-SANH	1 mg	0.79	79
278	GGGWWSSK-PEG_4-DBCO	1 mg	0.83	83
279	GGGWWSSK-PEG_4-pentenoic acid	1 mg	0.83	82
198	GK(ε-NH)-PEG_4-Azide	1 mg	0.81	80
367	GGGGGWK(ε-NH)-EEE-PEG_4-N_3	1 mg	0.86	86
282	GGGWSK(ε-NH)-Alkyne	1 mg	0.84	84
211	G-PEG_4-Alkyne	1 mg	0.77	77
296	GGYK(ε-NH)-PEG_6-Aldehyde	1 mg	0.51	51
295	GGYK(ε-NH)-PEG_{12}-Aldehyde	1 mg	0.63	62
304	GGGFDK(ε-NH)-KKK-Aldehyde	1 mg	0.76	76
321	GGGWSOrn(δ-NH)-PEG_4-SAc	1 mg	0.73	73
378	GGGGGWDDK(ε-NH)-Lipoic acid	1 mg	0.63	63
379	GGDK(ε-NH)-Thiazolidine	1 mg	0.78	78
362	GGGGK(ε-NH)-PEG_8-Thiazolidine	1 mg	0.88	87
291	GGG-EEE-SH	1 mg	0.81	81
292	GGWE-PEG_4-hydrazine	1 mg	0.66	66
293	GGGWSK(ε-NH)-Alkyne	1 mg	0.73	73
333	GGGE(δ-CONH)-PEG_4-Thioester	1 mg	0.64	64
337	GGWYSOrn (δ-NH)-PEG_6-ONH_2	1 mg	0.59	59
404	GGGG EEE-PEG_4-Azide	1 mg	0.69	69
297	GE-PEG_4-Azide	1 mg	0.64	64
298	GGSE-Thiazolidine	1 mg	0.70	70
349	GGGGWE(δ-CONH)PEG_4-hydrazine	1 mg	0.71	71

Example 54: Binding of SE scFv protein conjugates with activated platelets

To confirm that the SE scFv protein was still active, its binding to activated platelets was assessed using FACS studies with anti-His-tag and anti-VS-tag antibodies.

Flow cytometry

Blood (3 mL) from healthy volunteers taking no medication was collected by venepuncture, anticoagulated with citric acid and centrifuged at 150 g for 10 min at room temperature in an Eppendorf 5810 centrifuge. The upper yellow platelet-rich plasma (PRP) layer was removed and diluted 1:20 times in phosphate buffered saline (PBS, with Ca^{2+}/Mg^{2+}) and divided into two aliquots (activated and non-activated platelets). 100 µL of

20 mM adenosine diphosphate (ADP) was added to a 1 mL aliquot to activate the platelets. 50 µL of diluted PRP was added to each FACS tube.

Negative controls for the experiment were as follows-

Activated platelets (0.8 µL Qiagen, Penta-His, Alexa-Fluor 488 Conjugate)

Activated platelets (10 µL of 1:10 dilution, Invitrogen, anti-V5 FITC, 1.11 mg/mL, P/N 46-0308)

Positive controls were as follows-

Activated platelets (0.8 µL Qiagen, Penta-His, Alexa-Fluor 488 Conjugate plus 0.5 ug/mL unconjugated antibody)

Activated platelets (10 µL of 1:10 dilution, Invitrogen, anti-V5 FITC, 1.11 mg/mL, P/N 46-0308 plus 0.5 µg/mL unconjugated antibody)

Activation controls were as follows-

Non-activated platelets (1 µL PAC-1 FITC, BD Biosciences, Cat. #- 340507, 0.025 mg/mL)

Activated platelets (1 µL PAC-1, BD Biosciences, Cat. #- 340507, 0.025 mg/mL)

Staining was carried out as follows: Unconjugated antibody was added to the positive control tubes to achieve a concentration of 0.5 µg/mL in the tube. Conjugated antibody test samples were added at the same concentration. Samples were then incubated at room temperature for 10 minutes. After primary incubation, PAC-1 FITC was added to the non-activated platelet and activated platelet controls.

Penta-His and anti V5 FITC antibodies were added in the volumes mentioned above to the control tubes as well as to each sample tube so that each test sample was incubated separately with Penta-His and with anti-V5 FITC. Tubes were left to incubate for a further 10 minutes in the dark at room temperature. BD CellFIX was diluted 1:10 times and 500 µL was added to each FACS tube. Tubes were run on a BD FACSCalibur and analysed using FlowJo (TreeStar) flow cytometry software.

The LPETG sortase tag was placed before the His tag and after the V5 tag, so during the conjugation reaction the His tag is cleaved off and the V5 is still attached. As depicted in Table 6 providing mean fluorescence values, all products are negative for the His tag but positive for the V5 tag, confirming successful conjugation to the modifying peptides.

Table 6

SEQ ID NO	Conjugate peptide and SE scFv	Mean fluorescence	
	SE alone + anti His or anti V5	216.14	212.59
22	GGGK-εPEG$_4$-N$_3$	7.01	127.66
55	GGGWWK-εPEG$_4$-N$_3$	6.99	122.19
77	GGGWWSSK-εPEG$_4$-N$_3$	7.70	43.94
386	GGGWWGA-pG	8.34	76.32
276	GGGAGAGAC	8.68	210.87
277	GGGWWSSK-PEG$_4$-SANH	6.97	132.15
278	GGGWWSSK-PEG$_4$-DBCO	32.57	157.22
279	GGGWWSSK-PEG$_4$-pentenoic acid	12.54	192.50
198	GK(ε-NH)-PEG$_4$-Azide	11.06	202.30
367	GGGGGWK(ε-NH)-EEE-PEG$_4$-N$_3$	9.86	104.72
282	GGGWSK(ε-NH)-Alkyne	9.30	208.51
211	G-PEG$_4$-Alkyne	8.35	190.85
296	GGYK(ε-NH)-PEG$_6$-Aldehyde	8.55	147.72
295	GGYK(ε-NH)-PEG$_{12}$-Aldehyde	8.01	200.51
304	GGGFDK(ε-NH)-KKK-Aldehyde	7.51	60.64
321	GGGWSOrn(δ-NH)-PEG$_4$-SAc	7.81	216.66
378	GGGGGWDDK(ε-NH)-Lipoic acid	7.93	174.38
379	GGDK(ε-NH)-Thiazolidine	8.62	103.20
362	GGGGK(ε-NH)-PEG$_8$-Thiazolidine	6.98	64.64
291	GGG-EEE-SH	7.25	221.26
292	GGWE-PEG$_4$-hydrazine	7.4	121.43
293	GGGWSK(ε-NH)-Alkyne	6.23	169.90
333	GGGE(δ-CONH)-PEG$_4$-Thioester	5.98	225.07
337	GGWYSOrn (δ-NH)-PEG$_6$-ONH$_2$	6.68	199.57
404	GGGG EEE-PEG$_4$-Azide	6.44	108.42
297	GE-PEG$_4$-Azide	6.28	184.47
298	GGSE-Thiazolidine	6.82	107.82
349	GGGGWE(δ-CONH)PEG$_4$-hydrazine	6.60	60.42

Example 55: MALDI studies of protein peptide conjugates

Method protein analysis (CHCA) Linear mode peptide: Samples are mixed 1:1 with Matrix, 10 mg/mL α-cyano-4-hydroxycinnamic acid (Laser BioLabs, Sophia-Antipolis, France) in 50% Acetonitrile 0.1% TFA) and spotted onto the MALDI target plate. Proteins are analysed in Linear mode with a mass range of 10kDa to 50kDa and a focus mass of 35kDa and 29kDa. Laser shots are fired randomly across the sample well and the summed spectrum consists of spectra collected at the rate of 2500 shots/spectrum. The spectrum was internally calibrated against co-spotted Enolase using 2 point calibrate masses, 46672 and 23336 Da Mass Spectrometry Instrument: MALDI TOF/TOF, model 4700 Proteomics Analyser from Applied Biosystems (Foster City, CA, USA) running Software 4000 Series Explorer version 3.0. The results are shown in Tables 7A and 7B.

Table 7A

SEQ ID NO	structure	MW	Calc: GFP+Peptide	Calc: GFP Hydrolysis	measured GFP	Measured vs GFP-Pep	Measured vs GFP-Hyd
22	GGGK-εPEG$_4$-N$_3$	589.65	*27971.63*	*27381.98*	27887.99	-0.00299914	0.01814437
55	GGGWWK-εPEG$_4$-N$_3$	962.08	*28344.06*	*27381.98*	28274.45	-0.00246194	0.03156454
77	GGGWWSSK-εPEG$_4$-N$_3$	1136.31	*28518.29*	*27381.98*	28444.46	-0.00259558	0.037352792
386	GGGWWGA-pG	784.84	*28166.82*	*27381.98*	28057.89	-0.00388233	0.024089837
276	GGGAGAGAC	619.68	*28001.66*	*27381.98*			
277	GGGWWSSK-PEG$_4$-SANH	1288	*28669.98*	*27381.98*			
278	GGGWWSSK-PEG$_4$-DBCO	1483.61	*28865.59*	*27381.98*			
279	GGGWWSSK-PEG$_4$-pentenoic acid	1193.25	*28575.23*	*27381.98*			
198	GK(ε-NH)-PEG$_4$-Azide	476.54	*27858.52*	*27381.98*	27830.75	-0.00099782	0.01612497
367	GGGGGWK(ε-NH)-EEE–PEG$_4$-N$_3$	1278.32	*28660.3*	*27381.98*	28654.13	-0.00021533	0.044396741
282	GGGWSK(ε-NH)-Alkyne (amide)	699.64	*28081.62*	*27381.98*	28023.56	-0.00207183	0.022894308
211	G-PEG$_4$-Alkyne	288	*27669.98*	*27381.98*	27750.83	0.002913426	0.013291494
296	GGYK(ε-NH)-PEG$_6$-Aldehyde	891.05	*28273.03*	*27381.98*	27370.7	-0.032967	-0.00041212
295	GGYK(ε-NH)-PEG$_{12}$-Aldehyde	1155.36	*28537.34*	*27381.98*	27363.09	-0.04291365	-0.00069035

295	GGYK(ε-NH)-PEG_{12}-Aldehyde	1155.36	*28537.34*	*27381.98*	28421.42	-0.00407861	0.036572416
304	GGGFDK(ε-NH)-KKK-Aldehyde	1096.27	*28478.25*	*27381.98*	28276.99	-0.00711745	0.03165153
321	GGGWSOrn(δ-NH)-PEG_4-SAc	882.61	*28264.59*	*27381.98*	28136.49	-0.00455281	0.026816067
378	GGGGGWDDK(ε-NH)-Lipoic acid	1036.18	*28418.16*	*27381.98*	28278.6	-0.00493518	0.031706662
379	GGDK(ε-NH)-Thiazolidine	490.69	*27872.67*	*27381.98*			
362	GGGGK(ε-NH)-PEG_8-Thiazolidine	913.15	*28295.13*	*27381.98*	28203.73	-0.00324071	0.029136217
291	GGG-EEE-SH	636	*28017.98*	*27381.98*	28566.23	0.019192242	0.041456293
292	GGWE-PEG_4-hydrazine	708.09	*28090.07*	*27381.98*			
293	GGGWSK(ε-NH)-Alkyne	700.64	*28082.62*	*27381.98*			
333	GGGE(δ-CONH)-PEG_4-Thioester	652.99	*28034.97*	*27381.98*			
337	GGWYSOrn (δ-NH)-PEG_6-ONH_2	1090.17	*28472.15*	*27381.98*			
404	GGGG EEE-PEG_4-Azide	833.13	*28215.11*	*27381.98*	28268.33	0.001882672	0.031354877
297	GE-PEG_4-Azide	404.44	*27786.42*	*27381.98*	27836.08	0.001784016	0.01631336
297	GE-PEG_4-Azide	404.44	*27786.42*	*27381.98*	28171.05	0.013653378	0.028009961
298	GGSE-Thiazolidine	448	*27829.98*	*27381.98*	28416.13	0.020627369	0.036393063
349	GGGGWE(δ-CONH)PEG_4-hydrazide	822.56	*28204.54*	*27381.98*	28262	0.002033119	0.031137924

GFP alone	28561.2	28561.2	*28561.2*	28558.26	-0.00010295	-0.00010295

Table 7B

SEQ ID NO	structure	MW	Calc: SE+Peptide	Calc: SE Hydrolysis	measured SE	Measured vs SE-Pep	Measured vs SE-Hyd
22	GGGK-εPEG$_4$-N$_3$	589.65	*32251.93*	*31662.28*			
55	GGGWWK-εPEG$_4$-N$_3$	962.08	*32624.36*	*31662.28*			
77	GGGWWSSK-εPEG$_4$-N$_3$	1136.31	*32798.59*	*31662.28*			
386	GGGWWGA-pG	784.84	*32447.12*	*31662.28*			
276	GGGAGAGAC	619.68	*32281.96*	*31662.28*			
277	GGGWWSSK-PEG$_4$-SANH	1288	*32950.28*	*31662.28*			
278	GGGWWSSK-PEG$_4$-DBCO	1483.61	*33145.89*	*31662.28*			
279	GGGWWSSK-PEG$_4$-pentenoic acid	1193.25	*32855.53*	*31662.28*			
198	GK(ε-NH)-PEG$_4$-Azide	476.54	*32138.82*	*31662.28*			
367	GGGGGWK(ε-NH)-EEE–PEG$_4$-N$_3$	1278.32	*32940.6*	*31662.28*			
282	GGGWSK(ε-NH)-Alkyne (amide)	699.64	*32361.92*	*31662.28*			
211	G-PEG$_4$-Alkyne	288	*31950.28*	*31662.28*			
296	GGYK(ε-NH)-PEG$_6$-Aldehyde	891.05	*32553.33*	*31662.28*			
295	GGYK(ε-NH)-PEG$_{12}$-Aldehyde	1155.36	*32817.64*	*31662.28*			
295	GGYK(ε-NH)-PEG$_{12}$-Aldehyde	1155.36	*32817.64*	*31662.28*			
304	GGGFDK(ε-NH)-KKK-Aldehyde	1096.27	*32758.55*	*31662.28*			
321	GGGWSOrn(δ-NH)-PEG$_4$-SAc	882.61	*32544.89*	*31662.28*			
378	GGGGGWDDK(ε-NH)-Lipoic acid	1036.18	*32698.46*	*31662.28*			

379	GGDK(ε-NH)-Thiazolidine	490.69	*32152.97*	*31662.28*				
362	GGGGK(ε-NH)-PEG_8-Thiazolidine	913.15	*32575.43*	*31662.28*				
291	GGG-EEE-SH	636	*32298.28*	*31662.28*				
292	GGWE-PEG_4-hydrazine	708.09	*32370.37*	*31662.28*	32090.52	-0.00872064	0.013344751	
293	GGGWSK(ε-NH)-Alkyne	700.64	*32362.92*	*31662.28*	32115.17	-0.00771442	0.014102058	
333	GGGE(δ-CONH)-PEG_4-Thioester	652.99	*32315.27*	*31662.28*	32164.84	-0.00467685	0.015624514	
337	GGWYSOm (δ-NH)-PEG_6-ONH_2	1090.17	*32752.45*	*31662.28*	31798.54	-0.02999855	0.004285102	
404	GGGG EEE-PEG_4-Azide	833.13	*32495.41*	*31662.28*				
297	GE-PEG_4-Azide	404.44	*32066.72*	*31662.28*				
297	GE-PEG_4-Azide	404.44	*32066.72*	*31662.28*				
298	GGSE-Thiazolidine	448	*32110.28*	*31662.28*				
349	GGGGWE(δ-CONH)PEG_4-hydrazide	822.56	*32484.84*	*31662.28*				
SE alone	35223.04		35223.04	*35223.04*	34202.68	-0.02983275	-0.02983275	
SE alone	35223.04		35223.04	*35223.04*	35579	0.010004778	0.010004778	

The Sortase mediated modification of a protein with an LEPTG motif has at least two competing mechanisms, being cleavage at LEPT_G followed by amide bond formation to the N-terminal of the Glycine peptide, or cleavage at LEPT_G followed by hydrolysis The MALDI data presented in Tables 7A and 7B shows the observed molecular weight for proteins modified by Sortase as described in Example 53. The columns A and B represent a calculated ratio: (Observed MW – Calculated MW) ÷ Observed MW for each of the two possible Sortase pathways. In most cases, the Observed MW is more consistent with integration of the modifier peptide into the GFP or SE protein.

Example 56: Conjugation of two proteins with complementary functional groups

Conjugation reaction

Hydrazine/Aldehyde

Proteins modified with complementing reactive groups (Hydrazine/Aldehyde) were mixed in a molar ratio of 1:1, 1:3 or 3:1 and reacted in reaction buffer (100 mM citrate, 150mM NaCl, pH6.0) in a final volume of 300 µL for 12h at 4°C. Successful conjugation was confirmed by SDS gel and FACS (high molecular weight product and binding to activated platelets)

Cu free click (DBCO to N_3)

Proteins modified with complementing reactive groups (DBCO/Azide) were mixed in a molar ratio of 1:1 and reacted in reaction buffer (PBS without Ca/Mg) for 45 min at RT. Successful conjugation was confirmed by SDS gel and FACS (high molecular weight product and binding to activated platelets)

Cu click (Alkyne to N_3)

300 µL of 1 mg/ml antibody solution (modified with Alkyne containing peptide) was mixed with a 1:2 molar excess of antibody solution (modified with N_3 containing peptide) in the presence of CAC (copper/ascorbate/chelate mixture) for 40 min at 4°C. CAC = 10 µL each of 1.8 mg/mL Cu in MilliQ, 4.4 mg/mL ascorbate in PBS and 4.89 mg/mL chelate in DMSO. Conjugation was confirmed by SDS PAGE.

NCL (Native Chemical Ligation)

For NCL (Native Chemical Ligation) between Thiazolidine to Thioester, the Thiazolidine is first deprotected with 250 mM Methoxyamine HCl in 50mM Acetate buffer (pH 4) for

2h at 37°C and dialysed against PBS. The product was then reacted in 100 mM phosphate buffer pH 7.0-7.5 in the presence of 2% (v/v) thiophenol and 2% (v/v) benzyl mercaptan overnight at 4°C. Successful conjugation was confirmed by SDS gel and FACS (high molecular weight product and binding to activated platelets)

Conjugation to NIR dye

300 µL of 1 mg/mL modified antibody solution (pG peptide) was mixed with a 1:3 molar excess of NIR-dye (lumiprobe) in the presence of CAC (copper/ascorbate/chelate mixture) for 40 min at 4°C. Excess dye was removed by 3 wash steps using a 30K spincolumn. CAC = 10 µL each of 1.8 mg/mL Cu in MilliQ, 4.4 mg/ml ascorbate in PBS and 4.89 mg/mL chelate in DMSO. Labeling was confirmed by SDS PAGE and imaging on an Odyssey NIR system (Licor Biosystems).

Example 57: Conjugation of SE Scv with GFP

Reaction of two protein conjugates of the invention was undertaken. The protein conjugates were prepared in Example 53.

SE scFv modified with GGGWWSSK-PEG_4-SANH (SEQ ID NO:277) was reacted with GFP modified with GGYK(ε-NH)PEG_{12}-Aldehyde (SEQ ID NO:295)

and

SE scFv modified with GGYK(ε-NH)PEG_{12}-Aldehyde (SEQ ID NO:295) was reacted with GFP modified with GGGWWSSK-PEG_4-SANH (SEQ ID NO:277)

in molar ratios of 1:1 and 1:3. The products were assessed by SDS PAGE and showed successful conjugation with a product of 61 kD.

Binding of the SE-GFP conjugates to activated platelets was confirmed using FACS. Results are shown in Table 8.

Table 8: Mean fluorescence in binding to activated platelets

Reaction	Ratio	GFP mean fluorescence (FL1-H)
SE277 + GFP295	1:1	6.92
SE277 + GFP295	1:3	13.73
SE295 + GFP277	1:1	8.41

SE295 + GFP277	1:3	18.4
SE + GFP295	1:3	7.51

SE-GFP conjugates reacted in a ratio of 1:3 have a higher mean fluorescence compared to the control.

Example 58: Conjugation of SE scFv with GFP using click reaction

Reaction of two protein conjugates of the invention was undertaken using click chemistry. The protein conjugates were prepared in Example 53.

SE scFv modified with GK(ε-NH)PEG_4-$(CH_2)_2N_3$ (SEQ ID NO:198) was reacted with GFP modified with GGGWSK(ε-NH)CHC≡CH (SEQ ID NO:282)

and

SE scFv modified with GGGWSK(ε-NH)CHC≡CH (SEQ ID NO:282) was reacted with GFP modified with GK(ε-NH)$PEG_4(CH_2)_2N_3$ (SEQ ID NO:198).

The reaction was carried out as described in Example 56 with the components presented in a 1:2 ratio. SDS PAGE showed successful conjugation with a band at 61 kD. Binding of the SE-GFP conjugates to activated platelets was confirmed by FACS. The results are shown in Table 9.

Table 9: Mean fluorescence in binding to activated platelets

Reaction	**Ratio**	**GFP mean fluorescence (FL1-H)**
SE198 + GFP282	1:2	19.58
GFP198 + SE282	1:2	20.15
SE + GFP282	1:2	9.64

Example 59: Conjugation of SE scFv with GFP using native chemical ligation

The conjugation of two proteins using native chemical ligation was undertaken. The peptides used were prepared in Example 53.

SE scFv modified with GGGE(δ-CONH)-PEG_4-thioester (SEQ ID NO:333) and GFP modified with GGDK(ε-NH)-thiazolidine (SEQ ID NO:379)

and

SE scFv modified with GGDK (ε-NH)-thiazolidine (SEQ ID NO:379) and GFP modified with GGGE(δ-CONH)PEG_4-thioester (SEQ ID NO:333)

and

SE scFv modified with GGGGK(ε-NH)-PEG_8-thiazolidine (SEQ ID NO:362) and GFP modified with GGGE(δ-CONH)-PEG_4-thioester (SEQ ID NO:333).

The following reactions were carried out in a 1:3 or 3:1 ratio:

SE379 : SFP333	3:1
SE379 : SFP333	1:3
GFP379 : SE333	3:1
GFP379 : SE333	1:3
SE362 : GFP333	3:1
SE362 : GFP333	1:3

In all cases, the conjugation product (61 kD) was observed by SDS PAGE indicating successful conjugation.

Example 60: Conjugation of protein and dendrimer

Two dendrimers were prepared as follows:

Dendrimer A: Azidobenzamide-NEOEOEN[SU(NPN)$_2$][Lys]$_{10}$[NH_2.TFA]$_{32}$

Azidobenzamide-NEOEOEN[Su(NPN)$_2$] [Lys]$_{16}$ [BOC]$_{32}$

[NH_2]EOEOEN[Su(NPN)$_2$][Lys]$_{16}$[BOC]$_{32}$(prepared as described in WO 2008/017125 Example 1(xiii)) was reacted with azidobenzoic acid using standard PyBOP coupling conditions.

Azidobenzamide-NEOEOEN[Su(NPN)$_2$] [Lys]$_{16}$ [NH_2.TFA]$_{32}$

A solution of azidobenzamide-NEOEOEN[Su(NPN)$_2$] [Lys]$_{16}$ [BOC]$_{32}$ (190 mg, 25.1 μmol) in DCM/TFA (2.5 mL / 2.5 mL) was stirred for 3 h then added to an ice-cooled, stirred solution of diethyl ether (15 mL).The off-white precipitate was collected by filtration and washed with diethyl ether, then dissolved in water and freeze dried.

Dendrimer B: Azidobenzamid-PEG_{12}-NEOEOEN[SU$(NPN)_2$][Lys]$_{16}$[PEG_{570}]$_{32}$ [NH_2]EOEOEN[Su$(NPN)_2$][Lys]$_{16}$[PEG_{570}]$_{32}$ (as described in WO 2008/017125 Example 1(xvii)) was reacted with azidobenzamide-PEG_{12}-CO_2H employing standard PyBOP coupling conditions.

SE scFv modified with GGGWSK(ε-NH)-alkyne amide (SEQ ID NO:282) or G-PEG_4-Alkyne (SEQ ID NO:211) as described in Example 53. The SE-alkynes were conjugated to Dendrimers A and B using a copper catalyzed click reaction with a molar ratio of SE protein to dendrimer of 1:2. Successful conjugation was confirmed by SDS PAGE – SE-Dendrimer A 41 kD, SE-Dendrimer B 56 kD.

Confirmation that activity of the SE scFv protein was preserved was obtained by assessing binding to activated platelets as described above. Activity was maintained for the conjugates as shown in Table 10.

Table 10: SE scFv dendrimer conjugates mean fluorescence

Product	**Mean fluorescence (anti V5 antibody)**
SE282/Dendrimer A	8.78
SE282/Dendrimer B	5.3
SE211/Dendrimer A	7.07
SE211/Dendrimer B	6.31
SE + Dendrimer control	7.78

Example 61: Effect of position of LPETG tag on efficacy of reaction with GGG peptides

Sortase A mediated conjugation of GGG peptides and LPETG proteins in which the sortase recognition tag is near the C-terminus: or within the protein was compared.

Three proteins were used, GFP in which the LPETG tag is located 13-9 amino acids from the C-terminus

EKRDHMVLLEFVTAAGITLGMDELYK**LPETG**GLEHHHHHH (SEQ ID NO:418).

Mut MA2 protein in which the LPETG tag is located 37-33 amino acid residues from the C-terminus:

L P E T G G L E E A A A R G G P E Q K L I S E E D L N S A V D H H H H H H
(SEQ ID NO:419)

and

SCE5/SE protein in which the LPETG tag is located 37-33 amino acid residues from the C-terminus:

L P E T G G L E E A A A R G G P E Q K L I S E E D L N S A V D H H H H H H
(SEQ ID NO:420)

The conjugation efficacy was assessed as described in Example 53 and is shown in Table 11.

Table 11: Conjugation efficacy

SEQ ID NO:	LPETG tag	Starting amount	Recovered	%
277	GFP	1 mg	0.79	79
279	GFP	1 mg	0.83	82
295	GFP	1 mg	0.63	62
362	GFP	1 mg	0.88	87
277	SE	1 mg	0.289	29
279	SE	1 mg	0.421	42
295	SE	1 mg	0.291	29
365	SE	1 mg	0.327	33
277	mut MA2	1 mg	0.390	39
279	mut MA2	1 mg	0.455	46
295	mut MA2	1 mg	0.273	27
365	mut MA2	1 mg	0.470	47

Example 62: Modification of other proteins

Several proteins with LPETG tags were modified with GGYK(ε-NH)-PEG_{12}-Aldehyde (SEQ ID NO:295) or GGGWWSSK-PEG_4-SANH (SEQ ID NO:277). The proteins modified were IgG (SEQ ID NO:413), peroxidase (SEQ ID NO:408), streptavidine (SEQ ID NO:411), IL-11(SEQ ID NO:409, IL-1ra (SEQ ID NO:410), miD (SEQ ID NO: 412), GFP (SEQ ID NO:417) and SE scFv (SEQ ID NO:415). The efficacy of sortase reaction was determined by assessing protein by BCA. The results are shown in Table 12:

Table 12

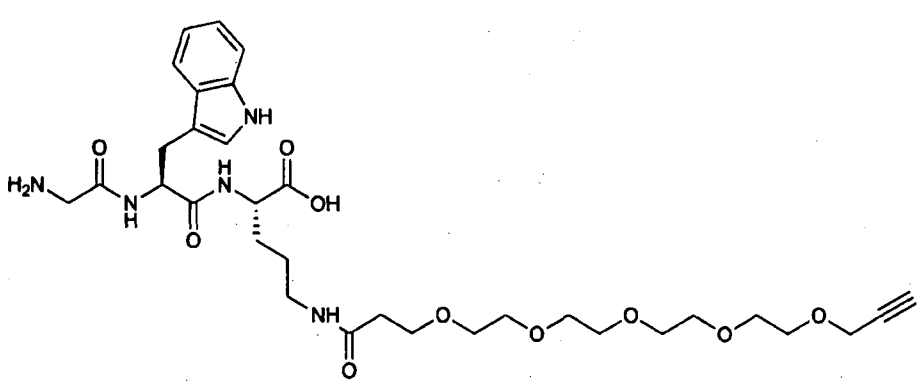
Protein	SEQ ID NO:	Amount in reaction	% coupling
IL-11	SEQ ID NO:295	170	17
IL-1ra	SEQ ID NO:295	210	11
Strept	SEQ ID NO:295	150	15
miD	SEQ ID NO:295	300	30
peroxidase	SEQ ID NO:295	700	70
IgG	SEQ ID NO:295	561	28
GFP	SEQ ID NO:277	1 mg	79
SE scFv	SEQ ID NO:277	1 mg	29

Example 63: Conjugation with Sortase B

Sortase B studies

Sortase B model substrate MAPAALWVALVFELQLWATGHTGA**NPQTN** was modified with GGYK(ε-NH)-PEG_{12}-Aldehyde (SEQ ID NO:295) under Sortase B catalysis conditions and conjugated to GFP modified with GGGWWSSK-PEG_4-SANH (SEQ ID NO:277). Successful conjugation (addition of 4kD to GFP) was demonstrated by SDS PAGE.

Example 64: Synthesis of GWOrn(δNH)-$PEG_4CH_2C{\equiv}CH$



GWOrn(δNH_2) was synthesised as described in Example 24 using Fmoc-Orn(Dde)-Wang Resin, Fmoc-Trp(Boc)-OH and Boc-Gly-OH followed by deprotection of the δ-amino group. GWOrn(δNH_2) in DMF (10 mL) was treated with Alkyne-PEG_4–CO_2H (183 mg, 0.6 mmol), HATU (216 mg, 0.570 mmol) and NMM (136 µL, 0.6 mmol) under N_2 for 120 min. The resin was then washed with DMF (3x10 mL). Kaiser test confirmed the reaction was complete. The resin was washed with MeOH (1x10 mL), DCM (3x10 mL) and MeOH (2x10 mL) and then air-dried (800 mg). The peptide was then removed from the resin and

deprotected to give the crude peptide (89 mg) as described in Example 24. The peptide was purified by HPLC using a Galaksil UP-C18 10um 120A column and a solvent gradient of 30-40 Solvent A/30min as described in Example 24 and analysed by mass spectrometry as described in Example 25. Purity by HPLC: 97.6%. MS: $C_{32}H_{48}N_5O_{10}^+$ Calc: 662.34, Found: 661.88.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method of preparing a compound of formula (II):

$$(A)\text{—}X_aX_bX_cX_dG(G)_n\text{-}X_1X_2Y_1Z_1 \qquad (II)$$

wherein

A is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a photosensitizer, DNA, RNA, a carbohydrate, a contrast agent, a catalyst, a lipid or a virus particle,

X_a is selected from a leucine, serine, asparagine, valine, isoleucine, tyrosine and glutamine residue;

X_b is selected from a proline, alanine, serine and glycine residue;

X_c is any amino acid residue;

X_d is selected from a threonine, arginine, serine, alanine and glycine residue;

G is a glycine residue;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

X_2 is absent or is an amino acid residue with a reactive functional group in its side chain;

Y_1 is absent or is a spacer group;

Z_1 is a reactive functional group; and

n is 0 or an integer of 1 to 4;

said method comprising contacting a compound of formula (I):

$$R_1\text{-NH-}G(G)_n\text{-}X_1X_2Y_1Z_1 \qquad (I)$$

wherein

R_1 is hydrogen or an amino protecting group; and

G, X_1 X_2, Y_1, Z_1 and n are defined above;

with a compound of formula (III):

$$(A)—X_aX_bX_cX_dX_e\text{-}X_3 \qquad \text{(III)}$$

wherein A, X_a, X_b, X_c and X_d are defined above and X_e is selected from a glycine, alanine, asparagine, serine or valine residue and X_3 is absent, is a purification tag or when A is a protein or peptide, X_3 is optionally a bond to A;
in the presence of a sortase enzyme.

2. The method according to claim 1 wherein A is a protein or peptide.

3. The method according to claim 2 wherein the protein or peptide is an antibody or fragment thereof.

4. The method according to any one of claims 1 to 3 wherein the sortase enzyme is Sortase A and the recognition tag is $X_aX_bX_cX_dX_e$:
wherein:
X_a is leucine or isoleucine,
X_b is proline or glycine,
X_c is any amino acid residue,
X_d is threonine or alanine, and
X_e is glycine or alanine.

5. The method according to claim 4 wherein the recognition tag is LPX_cTG.

6. The method according to any one of claims 1 to 5 wherein Xaa_3 is a purification tag.

7. The method according to claim 6 wherein the purification tag is His_6.

8. The method according to any one of claims 1 to 7 wherein X_1 comprises one or more amino acid residues having a bulky side chain.

9. The method according to claim 8 wherein X_1 comprises the sequence

-$Xaa_1Xaa_2Xaa_3Xaa_4$-

wherein Xaa_1 is an amino acid residue having a bulky side chain;

Xaa_2 is absent or is an amino acid residue having a bulky side chain;

Xaa_3 is absent or is any amino acid; and

Xaa_4 is absent or is any amino acid.

10. The method according to claim 9 wherein Xaa_1 is selected from a tryptophan, tyrosine, phenylalanine, leucine, isoleucine and histidine residue.

11. The method according to claim 10 wherein Xaa_1 is a tryptophan residue.

12. The method according to any one of claims 1 to 11 wherein Xaa_2 is selected from tryptophan, tyrosine, phenylalanine, leucine, isoleucine and histidine residue.

13. The method according to claim 12 wherein Xaa_2 is a tryptophan residue.

14. The method according to any one of claims 1 to 13 wherein Xaa_3 is selected from a serine, alanine, asparagine, aspartic acid, methionine, threonine and valine residue.

15. The method according to claim 14 wherein Xaa_3 is a serine or threonine residue.

16. The method according to any one of claims 1 to 15 wherein Xaa_4 is selected from a serine, alanine, asparagine, aspartic acid, methionine, threonine and valine residue.

17. The method according to claim 16 wherein Xaa_4 is a serine or threonine residue.

18. The method according to any one of claims 1 to 17 wherein X_2 is selected from a lysine, ornithine, aspartic acid or glutamic acid residue.

19. The method according to claim 18 wherein X_2 is a lysine residue.

20. The method according to any one of claims 1 to 19 wherein Y_1 is attached to X_2 through the reactive functional group in the side chain of X_2.

21. The method according to any one of claims 1 to 20 wherein Y_1 is a pharmacokinetic modifying agent.

22. The method according to claim 21 wherein the pharmacokinetic modifying agent is selected from polyethylene glycol, polypropylene glycol, polyethyleneoxide, poly(alkyloxazolines), polyvinylpyrrolidinone, polylysine and polyglutamate or mixtures thereof.

23. The method according to claim 22 wherein the pharmacokinetic modifying agent is polyethylene glycol.

24. The method according to any one of claims 1 to 23 wherein Z_1 is a reactive functional group selected from an alkyne, an azide, a hydroxylamine group, a hydrazine group, an aldehyde, a ketone, an amino group, a carboxylic acid, a thiol group, an alkene, an thioester and a cysteine residue.

25. The method according to claim 24 wherein Z_1 is selected from an alkyne, an azide, a hydrazine group, a hydroxylamine group, an aldehyde, a carboxylic acid and an amino group.

26. The method according to any one of claims 1 to 25 wherein A is a protein or peptide.

27. The method according to claim 26 wherein the protein or peptide is an antibody or fragment thereof.

28. The method according to any one of claims 1 to 25 wherein A is a dendrimer.

29. The method according to claim 28 wherein the dendrimer is a polylysine dendrimer.

30. The method of any one of claims 1 to 29 further comprising the step of conjugating two compounds of formula (II) with one another wherein Z_1 of one compound of formula (II) is a complementary reactive functional group to Z_1 of the second compound of formula (II).

31. A method according to claim 30 wherein one Z_1 is an alkyne and the other is an azide, one Z_1 is an aldehyde and the other is an hydrazine or hydroxylamine, one Z_1 is an amine and the other is a carboxylic acid, one Z_1 is an alkene and the other is a thiol group or one Z_1 is a thioester and the other is a cysteine residue.

32. The method according to claim 30 or claim 31 wherein each A is the same.

33. The method according to claim 30 or claim 31 wherein each A is different.

34. The method according to any one of claims 30 to 33 wherein at least one A is a protein or peptide.

35. The method according to claim 34 wherein the protein or peptide is an antibody or fragment thereof.

36. A compound having formula (IV):

$$R_1\text{-HN-G(G)}_n\text{-}X_1X_2Y_1Z_1$$

wherein G is a glycine residue;

R_1 is hydrogen or an amino protecting group;

X_1 is absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

X_2 is an amino acid residue with a reactive functional group in its side chain;
Y_1 is a spacer group which is attached to X_2 through the reactive functional group in the side chain of X_2;
Z_1 is a reactive functional group; and
n is 0 or an integer of 1 to 4.

37. A compound according to claim 36 wherein X_2 is selected from a lysine, ornithine, glutamic acid or aspartic acid residue.

38. A compound according to claim 36 or 37 wherein X_2 is a lysine residue and Y_1 is attached to the ε-amino group of the lysine side chain.

39. A compound according to claim 36 or 37 wherein X_2 is an ornithine residue and Y_1 is attached to the δ-amino group of the ornithine side chain.

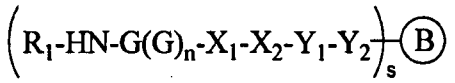
40. A compound according to claim 36 or 37 wherein X_2 is a glutamic acid residue and Y_1 is attached to the δ-carboxylic acid of the glutamic acid side chain.

41. A compound according to claim 36 or 37 wherein X_2 is an aspartic acid residue and Y_1 is attached to the δ-carboxylic acid of the aspartic acid group.

42. A compound according to any one of claims 36 to 41 wherein Y_1 is a pharmacokinetic modifying agent.

43. A compound according to claim 42 wherein the pharmacokinetic modifying agent is selected from polyethylene glycol, polylysine, polyglutamic acid, polyethylene glycol and polyethylene glycol and mixtures thereof.

44. A method of preparing a compound of formula (V):



(V)

wherein G is a glycine residue;

each R_1 is independently selected from hydrogen or an amino protecting group;

each X_1 is independently absent or is an amino acid residue or a peptide of 2-10 amino acid residues;

each X_2 is independently absent or is an amino acid residue with a reactive functional group in its side chain;

each Y_1 is independently absent or is a spacer group;

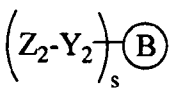
each Y_2 is independently absent or is a linker group;

B is a protein, a peptide, a label, a radioisotope, a small molecule drug, a targeting molecule, a nanoparticle, a cell, a dendrimer, a polymer, a photosensitizer, DNA, RNA, a carbohydrate, a contrast agent, a catalyst, a lipid or a virus particle;

s is an integer from 1 to 100; and

n is 0 or an integer of 1 to 4;

said method comprising reacting a compound of formula (I) defined in claim 1 with a compound of formula (VI):



(VI)

wherein B, Y_2 and s are defined above and Z_2 is a reactive functional group complementary to Z_1 of formula (I).

45. A method according to claim 44 wherein one of Z_1 and Z_2 is an alkyne group and the other is an azide group; one of Z_1 and Z_2 is an aldehyde group and the other is a hydrazine or hydroxylamine group; one of Z_1 and Z_2 is an amino group and the other is a carboxylic acid group; one of Z_1 and Z_2 is an alkene group and the other is a thiol group or one of Z_1 and Z_2 is a thioester and the other is a cysteine residue.

46. The method according to claim 44 or claim 45 further comprising the step of reacting a compound of formula (IV) with a compound of formula (III) as defined in claim 1, in the presence of a sortase enzyme.

47. The method according to claim 46 wherein at least one of A and B is a protein or peptide.

48. The method according to claim 47 wherein the protein or peptide is an antibody or fragment thereof.

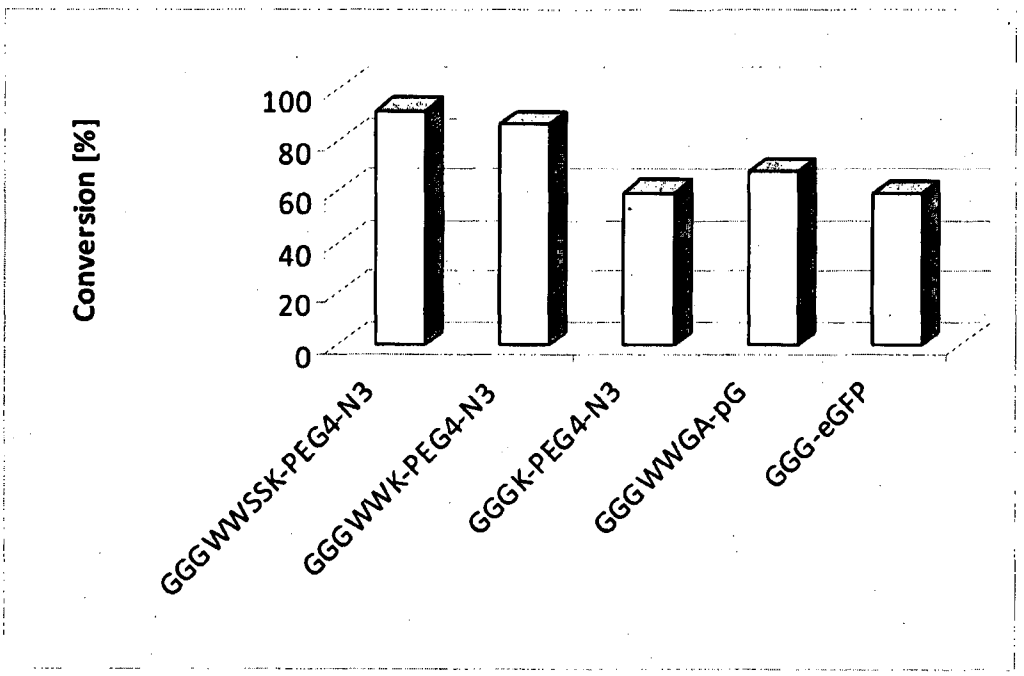


Figure 1

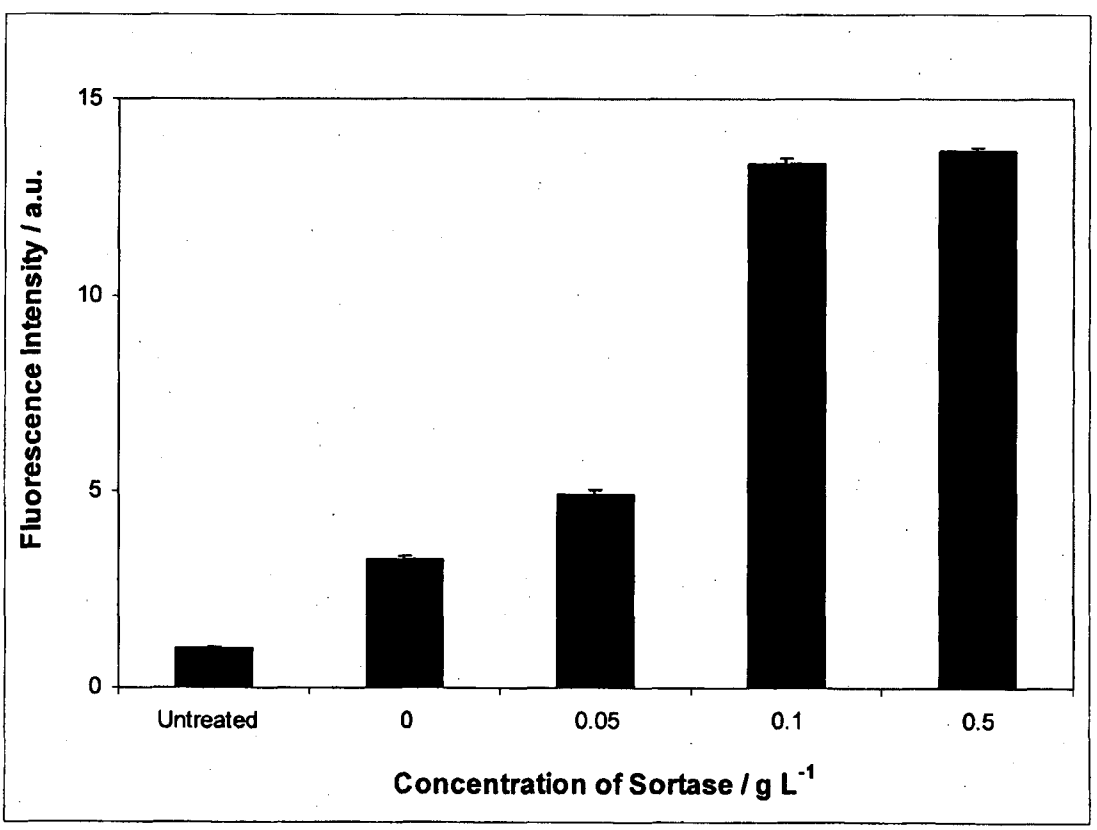


Figure 2A

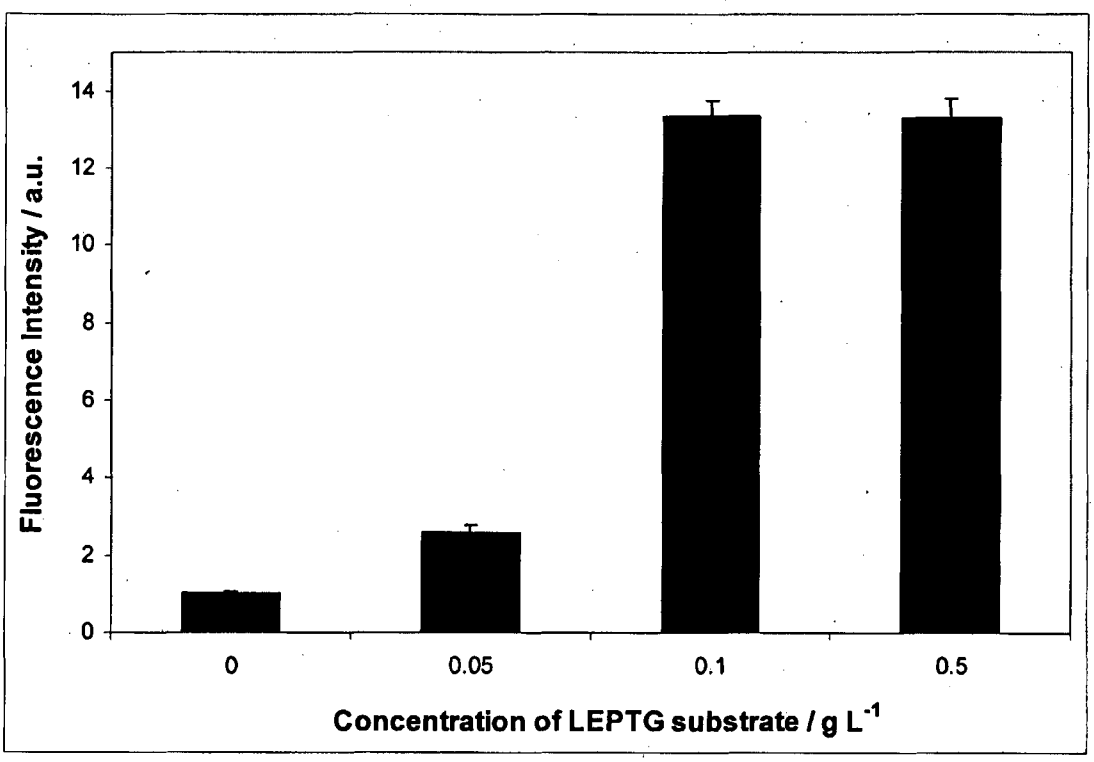


Figure 2B

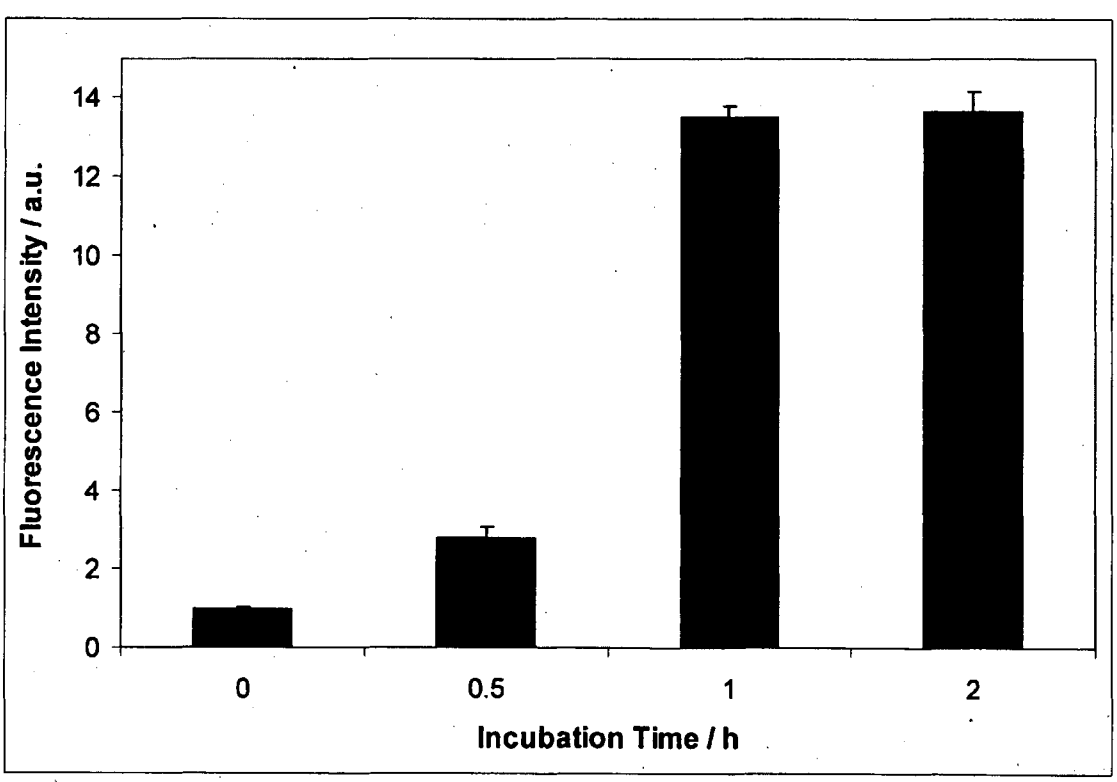


Figure 2C

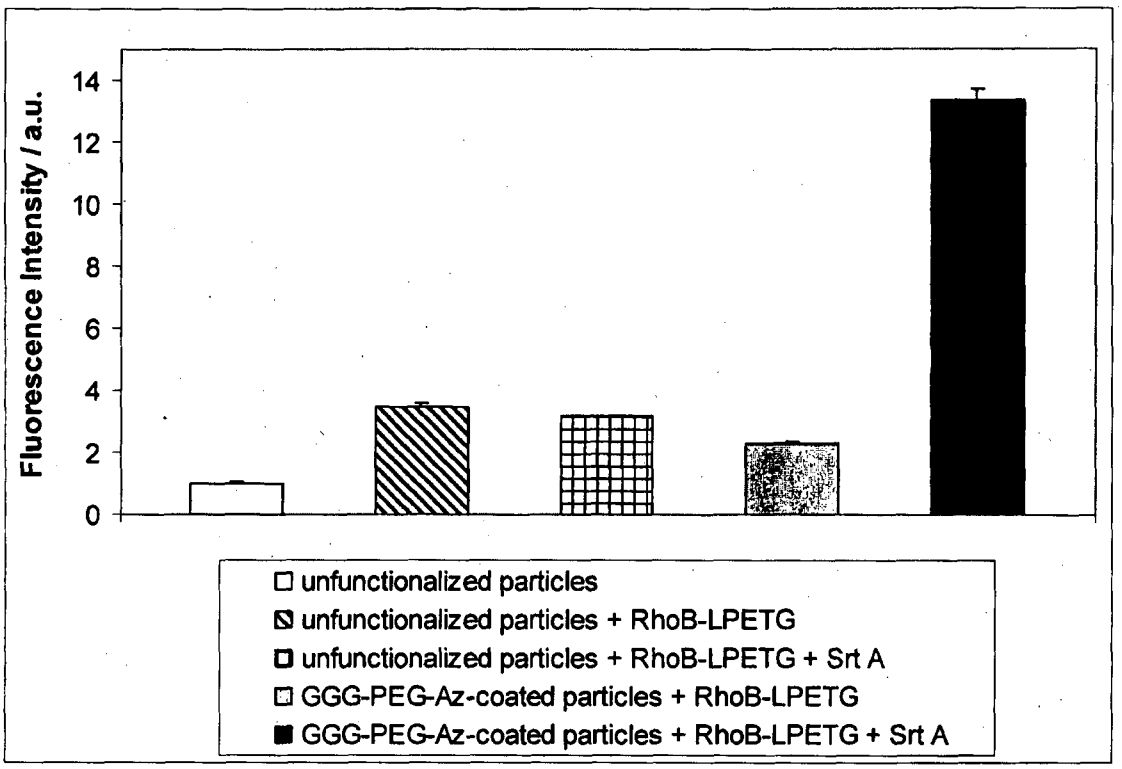


Figure 2D

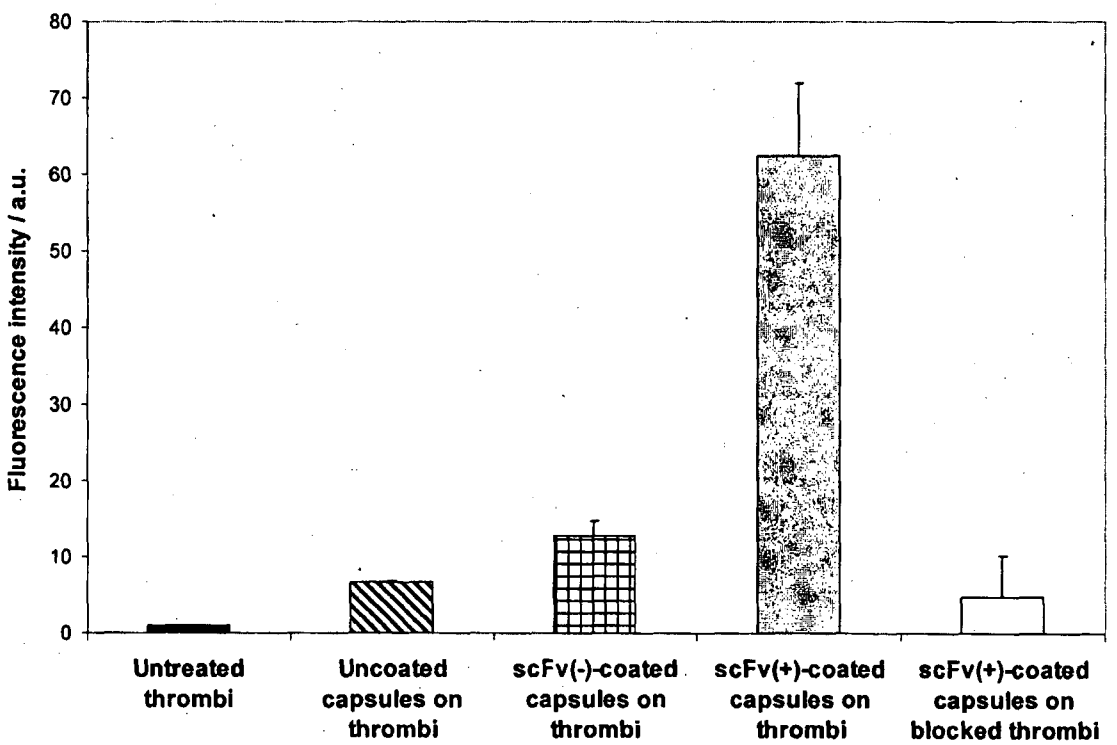


Figure 3

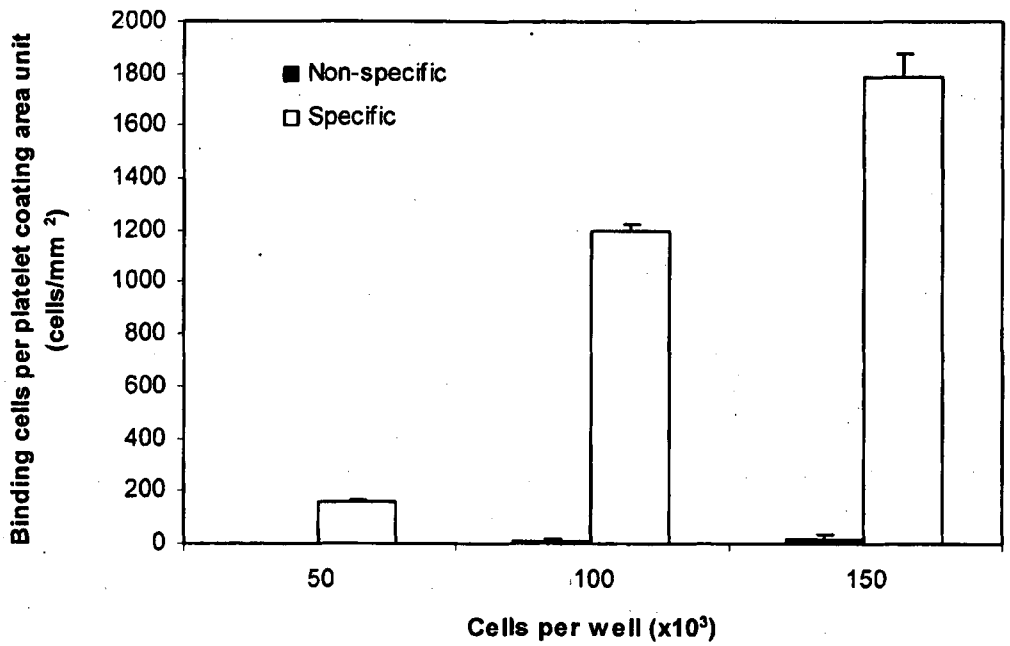

Binding cells per platelet coating area unit (cells/mm 2)
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Non-specific
Specific
50
100
150
Cells per well (x10³)

Figure 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2012/000400

A. CLASSIFICATION OF SUBJECT MATTER

[See Supplemental Sheet]

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Medline, Caplus, Biotechabs, Biosis, WPIDS: Keyword Search based on (Sortase, conjugate, link, ligate, reaction, stapler, catalyse, complex, glycine, tag, motif, affinity, diglycine, site specific, antibody, nanoparticle, label); EPODOC, WPI: Keyword Search based on Sortase and A61K; STN Registry, Caplus: Structure Search based on compounds of formula (IV); EPODOC, WPI: Keyword Search based on (conjugate, link, ligate, click, antibody, nanoparticle, dye, bead, peptide, protein, glycine, hydrazine, azide, Diels-Alder, ncl).

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 September 2012	18 September 2012
Name and mailing address of the ISA/AU	**Authorised officer**
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustralia.gov.au Facsimile No.: +61 2 6283 7999	Tina Bryant AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61 2 62837992

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2012/000400

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

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

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

See Supplemental Box for Details

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

Remark on Protest

☒ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		International application No. PCT/AU2012/000400
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Mao, H. et al.,"Sortase-Mediated Protein Ligation: A New Method for Protein Engineering", J. Am. Chem. Soc., 2004, Vol. 126, pages 2670-2671. Page 2671 Table 2, Page 2670 right column second paragraph – page 2671 left column second paragraph.	1, 2, 4-18, 19, 21, 24-26, 36-38
X	WO 2005/051976 A2 (ANSATA THERAPEUTICS, INC.) 09 June 2005 Paragraph [0006], [0022]-[0024], [0044]. [0079], Examples 1, 3, 4, 6, Claims 1, 2, 8, 13 and 20.	1-18, 19, 21-27
X	Antos J. M. et al., "Lipid Modification of Proteins through Sortase-Catalyzed Transpeptidation", Journal of the American Chemical Society, 2008, Vol. 130, No. 48, pages 16338-16343. Page 16340 left column lines 5-25, Page 16340 Scheme 2, page 16340 Figure 1a, page 16341 Figure 2, page 16341 Table 1.	1, 2, 4-7, 24-26
X	Popp M. W. et al., "Sortagging: a versatile method for protein labeling", Nature Chemical Biology, 2007, Vol. 3, No. 11, pages 707-708. Page 707 Figure 1a, Page 708 left column lines 29-31, Supplementary Figure 6e, Supplementary Material pages 19 and 20 CXCL14 Substrate Cloning, Expression, and Purification and CXCL14 Labeling.	1, 2, 4-7, 18-20, 24-26, 36-38
X	Parthasarathy R. et al., "Sortase A as a Novel Molecular "Stapler" for Sequence-Specific Protein Conjugation", Bioconjugate Chemistry, 2007, Vol. 18, No. 2, pages 469-476. Page 470 right column lines 17-35, page 471 left column lines 6-9, page 473 right column line 49 – page 474 right column line 7, page 474 Figure 6.	1, 2, 4-7, 21-23, 26
X	Popp, M. W. et al., "Sortase-catalyzed transformations that improve the properties of cytokines", PNAS, 2011, Vol. 108, No. 8, pages 3169-3174 (February 22 2011). Page 3169 Abstract, page 3170 left column – right column Modification of IFNa2, page 3170 Figure 2, page 3170 right column – page 3171 column Modification of a Single Protein Preparation by Multiple Nucleophiles, page 3171 left column Superior Properties Endowed upon Sortase Substrates Are Applicable to Multiple Proteins, page 3171, Figure 3, page 3172 Figure 4.	1-7, 18-27
X	Orth, R. et al., "Chemical Probes for Labeling of the Bacterial Glucosaminidase NagZ via the Huisgen Cycloaddition", Synthesis, 2010, Vol. 13, pages 2201-2206 and CAS Registry Number RN 1245731-06-9. Page 2206 left column lines 1-50 and CAS RN 1245731-06-9	36-41
X	Minta, E. et al., "Synthesis of cyclooctapeptides: constraints analogs of the peptidic neurotoxin, ω-agatoxine IVB – an experimental point of view", Journal of Peptide Science, 2008, Vol. 14, No. 3, pages 267-277. Page 276 Synthesis of EM 10: Fmoc-c(Glu-Cys-Thr-AzaTrp-Gly-Gly-Thr-Lys)-NH2	36-41
X	US 2004/0208828 A1 (LEHMANN et al.) 21 October 2004 Example 16 Paragraphs [0193]-[0196]	36-41
X	Nishimura, H. et al., "Mass spectra of sulphur-containing amino acids and peptides', Tetrahedron, 1972, Vol. 28, No. 17, pages 4503-4513. Page 4510 Compound 11, page 4512 line 16 – page 4513 line 15.	36-41
X	Kumar, I. et al., "Transpeptidation Reactions of a Specific Substrate Catalyzed by the Streptomyces R61 DD-Peptidase: Characterization of a Chromogenic Substrate and Acyl Acceptor Design", Biochemistry, 2005, Vol. 44, No. 30, pages 9971-9979. Page 9972 Compound 3, page 9972 right column line 13 – page 9973 right column line 25, page 9973 Scheme 2.	36-41

INTERNATIONAL SEARCH REPORT C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		International application No. **PCT/AU2012/000400**
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Lamango, N. S. et al., "Porcine Liver Carboxylesterase Requires Polyisoprenylation for High Affinity Binding to Cysteinyl Substrates", The Open Enzyme Inhibition Journal, 2009, Vol. 2, pages 12-27. Page 17 left column lines 5-30, page 19 Table Compound RD-18.	36-41
X	GB 2089816 A (RHONE-POULENC SANTE) 30 June 1982 Example 15	36-41
X	Chong, P. C. S. et al., "A New Heterobifunctional Cross-linking Reagent for the Study of Biological Interactions between Proteins", The Journal of Biological Chemistry, 1981, Vol. 256, No. 10, pages 5064-5070. Page 5066 Figure 1, page 5066 left column lines 6-26, page 5066 right column line 45 – page 5067 left column line 5, Page 5067 Figure 2.	36-41
X	Avaeva, S. M. et al., "Serine phosphates and pyrophosphates. Some properties of bisserine pyrophosphates", The Journal of General Chemistry of the USSR, 1968, Vol. 38, No. 10, pages 2158-2161. Page 2158 Figure Compound (II).	36-41
X	WO 1993/022677 A1 (PB DIAGNOSTICS SYSTEMS, INC.) 11 November 1993 Example II.	36-41
X	CAS Registry Number RN 1017528-75-4, STN Entry Date 27 April 2008. RN 1017528-75-4	36-38
X	CAS Registry Number RN 1126082-58-3, STN Entry Date 24 March 2009. RN 1126082-58-3	36-38
X	CAS Registry Number RN 218787-41-8, STN Entry Date 31 January 1999. RN 218787-41-8	36-38
X	CAS Registry Number RN 1042667-46-8, STN Entry Date 22 August 2008. RN 1042667-46-8	36-38
X	CAS Registry Number RN 81666-59-3, STN Entry Date 16 November 1984. RN 81666-59-3	36, 37, 39
X	CAS Registry Number RN 62105-03-7, STN Entry Date 16 November 1984. RN 62105-03-7	36, 37, 39
X	CAS Registry Number RN 439151-57-2, STN Entry Date 17 July 2002. RN 439151-57-2	36
X	CAS Registry Number RN 475273-01-9, STN Entry Date 06 December 2002. RN 475273-01-9	36
X	CAS Registry Number RN 38570-35-3, STN Entry Date 16 November 1984. RN 38570-35-3	36
X	CAS Registry Number RN 862706-87-4, STN Entry Date 08 September 2005. RN 862706-87-4	36
	WO 2008/134761 A2 (INTEZYNE TECHNOLOGIES, INC.) 06 November 2008	

INTERNATIONAL SEARCH REPORT C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		International application No. **PCT/AU2012/000400**
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Page 31 second compound, Paragraphs [0075], [0092], [0095], [00124], Tables 3, 4, 5, 8, 9, 14, 19, 20, 21, 32 Examples 1 and 5, Claims 1-20.	44-48
X	WO 2006/106348 A2 (ISIS INNOVATION LIMITED) 12 October 2006 Page 2 lines 27-29, page 6 lines 20-27, page 25 lines 11-16, Claims 1-4.	44-48
X	US 4864019 A (VALE et al.) 05 September 1989 Examples 1-4.	44-48
X	WO 1986/001210 A1 (SCRIPPS CLINIC AND RESEARCH FOUNDATION) 27 February 1986 Page 10 lines 16 - page 11 line 18, page 62 line 1 – page 63 line 28, Claims 1 and 4.	44-48
X	EP 0037110 A2 (TAKEDA YAKUHIN KOGYO KABUSHIKI KAISHA) 07 October 1981 Reference Example 3, Example 4.	44-48
X	Tornøe, C. W. et al., "Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides", J. Org. Chem., 2002, Vol. 67, No. 9, pages 3057-3064. Page 3058 Scheme 2, page 3059 Table 1, page 3062 left column lines 12-30.	44-48

INTERNATIONAL SEARCH REPORT	International application No. PCT/AU2012/000400

Supplemental Box

Continuation of: **Box III**
This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

This Authority has found that there are different inventions based on the following features that separate the claims into distinct groups:

- Claims 1-35 are directed towards a method of synthesising compounds of formula (II) via a Sortase mediated enzymatic reaction. The feature of a method of synthesising compounds of formula (II) from compounds of formula (I) in the presence of a Sortase enzyme is specific to this group of claims.

- Claims 36-43 are directed towards compounds of formula (IV). The feature of compounds of formula (IV) (that are a subset of compounds of formula (I)) is specific to this group of claims.

- Claims 44-48 are directed towards a method of synthesising compounds of formula (V). The feature of a method of synthesising compounds of formula (V) from compounds of formula (I) is specific to this group of claims.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

When there is no special technical feature common to all the claimed inventions there is no unity of invention.

In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. The only feature common to all of the claimed inventions and which provides a technical relationship among them is the compounds of formula (IV) (that are a subset of compounds of formula (I)). However this feature does not make a contribution over the prior art because it is disclosed in many documents, including D1, D4 and D7-D16:

D1: Mao, H. et al.,"Sortase-Mediated Protein Ligation: A New Method for Protein Engineering", J. Am. Chem. Soc., 2004, Vol. 126, pages 2670-2671.
(Page 2671 Table 2, Page 2670 right column second paragraph – page 2671 left column second paragraph).

D4: Popp M. W. et al., "Sortagging: a versatile method for protein labeling", Nature Chemical Biology, 2007, Vol. 3, No. 11, pages 707-708.
(Page 707 Figure 1a, Page 708 left column lines 29-31, Supplementary Figure 6e, Supplementary Material pages 19 and 20 CXCL14 Substrate Cloning, Expression, and Purification and CXCL14 Labeling).

D7: Orth, R. et al., "Chemical Probes for Labeling of the Bacterial Glucosaminidase NagZ via the Huisgen Cycloaddition", Synthesis, 2010, Vol. 13, pages 2201-2206 and CAS Registry Number RN 1245731-06-9.
(Page 2206 left column lines 1-50 and RN 1245731-06-9).

D8: Minta, E. et al., "Synthesis of cyclooctapeptides: constraints analogs of the peptidic neurotoxin, ω-agatoxine IVB – an experimental point of view", Journal of Peptide Science, 2008, Vol. 14, No. 3, pages 267-277.
(Page 276 Synthesis of EM 10: Fmoc-c(Glu-Cys-Thr-AzaTrp-Gly-Gly-Thr-Lys)-NH_2).

D9: US2004/0208828 A1 (LEHMANN et al.), 21 October 2004.
(Example 16 Paragraphs [0193]-[0196]).

D10: Nishimura, H. et al., "Mass spectra of sulphur-containing amino acids and peptides', Tetrahedron, 1972, Vol. 28, No. 17, pages 4503-13.

Supplemental Box

(Page 4510 Compound 11, page 4512 line 16 – page 4513 line 15).

D11: Kumar, I. et al., "Transpeptidation Reactions of a Specific Substrate Catalyzed by the *Streptomyces* R61 DD-Peptidase: Characterization of a Chromogenic Substrate and Acyl Acceptor Design", Biochemistry, 2005, Vol. 44, No. 30, pages 9971-9979.
(Page 9972 Compound 3, page 9972 right column line 13 – page 9973 right column line 25, page 9973 Scheme 2).

D12: Lamango, N. S. et al., "Porcine Liver Carboxylesterase Requires Polyisoprenylation for High Affinity Binding to Cysteinyl Substrates", The Open Enzyme Inhibition Journal, 2009, Vol. 2, pages 12-27.
(Page 17 left column lines 5-30, page 19 Table Compound RD-18).

D13: GB2089816 A (RHONE-POULENC SANTE), 30 June 1982.
(Example 15).

D14: Chong, P. C. S. et al., "A New Heterobifunctional Cross-linking Reagent for the Study of Biological Interactions between Proteins", The Journal of Biological Chemistry, 1981, Vol. 256, No. 10, pages 5064-5070.
(Page 5066 Figure 1, page 5066 left column lines 6-26, page 5066 right column line 45 – page 5067 left column line 5, Page 5067 Figure 2).

D15: Avaeva, S. M. et al., "Serine phosphates and pyrophosphates. Some properties of bisserine pyrophosphates", The Journal of General Chemistry of the USSR, 1968, Vol. 38, No. 10, pages 2158-2161. (Page 2158 Figure Compound (II).

D16: WO1993/022677 A1 (PB DIAGNOSTICS SYSTEMS, INC.), 11 November 1993. (Example II).

Therefore in the light of these documents this common feature (compounds of formula (IV)) cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a posteriori*.

INTERNATIONAL SEARCH REPORT	International application No. PCT/AU2012/000400

Supplemental Box – IPC Marks

C07K 1/107 (2006.01)

C07K 1/113 (2006.01)

C07K 1/13 (2006.01)

C07K 5/00 (2006.01)

C07K 5/062 (2006.01)

C07K 5/083 (2006.01)

C07K 5/103 (2006.01)

C07K 7/06 (2006.01)

C07K 7/08 (2006.01)

C07K 9/00 (2006.01)

C07K 17/06 (2006.01)

C07K 17/08 (2006.01)

C07K 17/10 (2006.01)

C07K 17/14 (2006.01)

C07K 19/00 (2006.01)

C12N 9/52 (2006.01)

A61K 38/04 (2006.01)

A61K 38/08 (2006.01)

A61K 38/10 (2006.01)

A61K 38/14 (2006.01)

A61K 51/08 (2006.01)

C12P 21/00 (2006.01)

C12P 21/02 (2006.01)

C07K 14/00 (2006.01)

C07K 14/54 (2006.01)

C07K 16/00 (2006.01)

A61K 38/05 (2006.01)

A61K 38/06 (2006.01)

A61K 38/07 (2006.01)

A61K 38/16 (2006.01)

A61K 38/20 (2006.01)

A61K 51/10 (2006.01)

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. **PCT/AU2012/000400**

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		**Patent Family Member/s**	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
WO 2005/051976 A2	09 Jun 2005	WO 2005051976 A2	09 Jun 2005
US 2004/0208828 A1	21 Oct 2004	AR 043021 A1	13 Jul 2005
		AU 2003296740 A1	30 Aug 2004
		DE 10305463 A1	12 Aug 2004
		EP 1590317 A1	02 Nov 2005
		EP 1590317 B1	24 Jan 2007
		JP 2006514081 A	27 Apr 2006
		JP 4878119 B2	15 Feb 2012
		US 2004208828 A1	21 Oct 2004
		US 2012064002 A1	15 Mar 2012
		WO 2004069787 A1	19 Aug 2004
GB 2089816 A	30 Jun 1982	AU 555294 B2	18 Sep 1986
		AU 7859081 A	24 Jun 1982
		BE 891539 A1	18 Jun 1982
		CA 1250398 A1	21 Feb 1989
		DE 3150295 A1	29 Jul 1982
		DK 563181 A	20 Jun 1982

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. **PCT/AU2012/000400**

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		**Patent Family Member/s**	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
		DK 154436 B	14 Nov 1988
		FR 2496654 A1	25 Jun 1982
		FR 2496654 B1	16 Dec 1983
		GB 2089816 A	30 Jun 1982
		GB 2089816 B	18 Jul 1984
		IE 52113 B1	24 Jun 1987
		IT 1142148 B	08 Oct 1986
		JP 57154151 A	22 Sep 1982
		JP 3037560 B	05 Jun 1991
		LU 83843 A1	07 Jul 1982
		NL 8105587 A	16 Jul 1982
		NZ 199310 A	11 Oct 1985
		SE 454699 B	24 May 1988
		SE 8107634 A	20 Jun 1982
		US 4401658 A	30 Aug 1983
		ZA 8108737 A	24 Nov 1982
WO 1993/022677 A1	11 Nov 1993	AU 662627 B2	07 Sep 1995
		AU 3976393 A	29 Nov 1993
		CA 2101734 A1	24 Oct 1993
		EP 0606411 A1	20 Jul 1994
		EP 0606411 B1	03 Mar 1999
		JP H06508639 A	29 Sep 1994
		US 5294536 A	15 Mar 1994
		WO 9322677 A1	11 Nov 1993

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. PCT/AU2012/000400

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
WO 2008/134761 A2	06 Nov 2008	EP 2155177 A2	24 Feb 2010
		JP 2010526091 A	29 Jul 2010
		US 2009110662 A1	30 Apr 2009
		WO 2008134761 A2	06 Nov 2008
WO 2006/106348 A2	12 Oct 2006	AU 2006232642 A1	12 Oct 2006
		BR PI0609088 A2	16 Nov 2010
		CA 2603936 A1	12 Oct 2006
		CN 101198619 A	11 Jun 2008
		EA 200702193 A1	28 Apr 2008
		EA 013265 B1	30 Apr 2010
		EP 1871784 A2	02 Jan 2008
		JP 2008534665 A	28 Aug 2008
		KR 20080007575 A	22 Jan 2008
		MX 2007012544 A	25 Feb 2008
		NZ 562996 A	28 Aug 2009
		US 2011059501 A1	10 Mar 2011
		WO 2006106348 A2	12 Oct 2006
		ZA 200709105 A	27 Aug 2008
US 4864019 A	05 Sep 1989	US 4864019 A	05 Sep 1989
WO 1986/001210 A1	27 Feb 1986	AU 585529 B2	22 Jun 1989
		AU 643188 B2	11 Nov 1993
		AU 4400089 A	05 Mar 1990
		AU 4721085 A	07 Mar 1986
		DK 21191 A	04 Apr 1991

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. PCT/AU2012/000400

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
		DK 175828 B1	14 Mar 2005
		DK 156786 A	07 Apr 1986
		DK 172035 B1	22 Sep 1997
		EP 0191813 A1	27 Aug 1986
		EP 0191813 B1	30 Nov 1994
		EP 0428629 A1	29 May 1991
		EP 0428629 B1	28 Feb 1996
		ES 8700442 A1	01 Jan 1987
		FI 102381 B	30 Nov 1998
		FI 102381 B1	30 Nov 1998
		GR 89100501 A	30 Dec 1991
		IE 65391 B1	18 Oct 1995
		IL 76037 A	31 Jul 1989
		JP H04503664 A	02 Jul 1992
		JP 3032221 B2	10 Apr 2000
		JP S62500027 A	08 Jan 1987
		JP 6078359 B	05 Oct 1994
		NO 910479 A	05 Apr 1991
		NO 309720 B1	19 Mar 2001
		NZ 213035 A	29 Nov 1988
		PT 91397 A	08 Mar 1990
		PT 91397 B	04 May 1995
		US 4654419 A	31 Mar 1987
		US 5116725 A	26 May 1992
		US 5122448 A	16 Jun 1992
		US 5717065 A	10 Feb 1998
		WO 8601210 A1	27 Feb 1986
		WO 9001495 A1	22 Feb 1990
		ZA 8505971 A	26 Mar 1986
EP 0037110 A2	07 Oct 1981	CA 1162848 A1	28 Feb 1984
		CA 1181342 A1	22 Jan 1985
		DK 137181 A	01 Oct 1981
		DK 159276 B	24 Sep 1990
		EP 0037110 A2	07 Oct 1981
		EP 0037110 B1	29 Jan 1986

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. **PCT/AU2012/000400**

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		**Patent Family Member/s**	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
		EP 0049898 A1	21 Apr 1982
		EP 0049898 B1	12 Mar 1986
		EP 0049898 B2	14 Aug 1991
		EP 0111216 A2	20 Jun 1984
		JP 57067858 A	24 Apr 1982
		JP 1036065 B	28 Jul 1989
		JP 56138248 A	28 Oct 1981
		JP 63062698 B	05 Dec 1988
		JP 57118790 A	23 Jul 1982
		JP 2018836 B	26 Apr 1990
		JP 57006363 A	13 Jan 1982
		JP 63041424 B	17 Aug 1988
		US 4496658 A	29 Jan 1985
		US 4517290 A	14 May 1985

End of Annex

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.