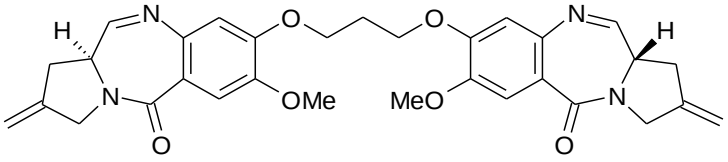
COMPOUNDS AND CONJUGATES

The present invention relates to azetidobenzodiazepine (ABD) dimers, conjugates comprising said dimers and the precursor drug linkers used to make such conjugates.

Background to the invention

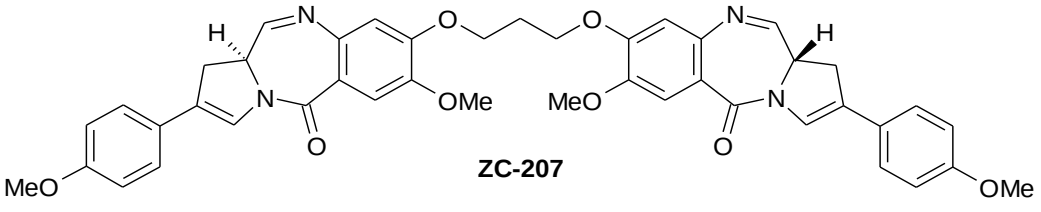
Pyrrolobenzodiazepine (PBD) dimers have been shown to be cytotoxic compounds.

For example, SG2000 (SJG-136):



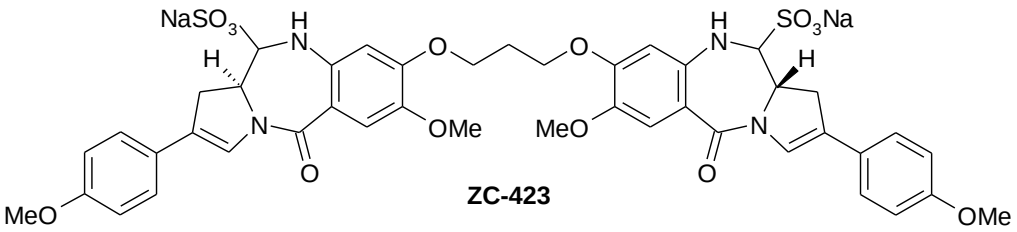
(Gregson, S.J., et al., Chem. Commun., 1999, 797-798. doi: 10.1039/A809791G; Gregson, S., *et al.*, *J. Med. Chem.*, **44**, 737-748 (2001); Alley, M.C., *et al.*, *Cancer Research*, **64**, 6700-6706 (2004); and Hartley, J.A., *et al.*, *Cancer Research*, **64**, 6693-6699 (2004)) has been involved in clinical trials as a standalone agent, for example, NCT02034227 investigating its use in treating Acute Myeloid Leukemia and Chronic Lymphocytic Leukemia (see: https://www.clinicaltrials.gov/ct2/show/NCT02034227).

Dimeric PBD compounds bearing C2 aryl substituents alongside endo-unsaturation, such as SG2202 (ZC-207), are disclosed in WO 2005/085251:



ZC-207

and in WO2006/111759, bisulphites of such PBD compounds, for example SG2285 (ZC-423):

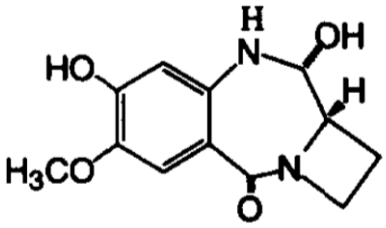


ZC-423

These compounds have been shown to be highly useful cytotoxic agents (Howard, P.W., *et al.*, *Bioorg. Med. Chem.* (2009), doi: 10.1016/j.bmcl.2009.09.012).

Dimer PBD compounds having linker groups for connection to a cell binding agent, such as an antibody, are described in WO 2011/130598. The linker in these compounds is attached to one of the available N10 positions, and are generally cleaved by action of an enzyme on the linker group. WO 2014/057074 and WO 2015/052322 describe specific PBD dimer conjugates bound via the N10 position on one monomer.

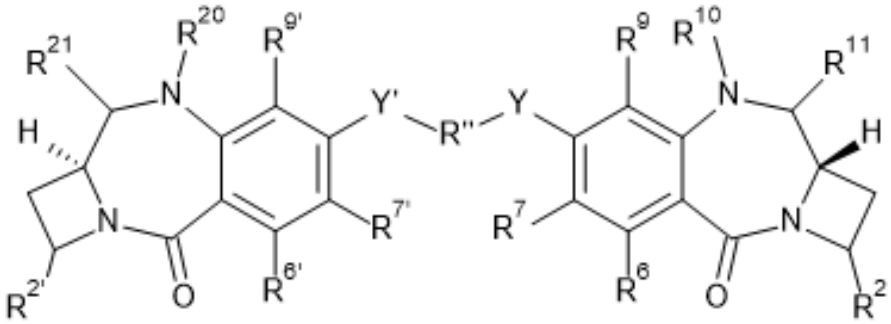
At a relatively early stage in the development of PBDs as molecules of interest, it was reported in 1997 (Bose, D.S., et al., Tetrahedron Letters, 38(33), 5839-5842, 1997; doi: 10.1016/S0040-4039(97)01297-5) that the following compound:



had been synthesised and would be evaluated as a potential DNA-binding ligand and cytotoxic agents. No further publication about this compound was made, so it appears that they were either not stable for testing or not active.

Disclosure of the invention

A first aspect of the present invention provides a compound of formula **IV**:



IV

and salts and solvates thereof, wherein:

R^2 and $R^{2'}$ are H;

R^6 and R^9 are independently selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo;

where R and R' are independently selected from optionally substituted C_{1-12} alkyl, C_{3-20} heterocyclyl and aromatic ring systems comprising 5-20 ring atoms~~C_{5-20} aryl groups~~;

either

(a) R^7 is selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo; $R^{7'}$ is selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo; or

(b) R^7 and $R^{7'}$ together form a group which is: (i) -O-$(CH_2)_n$-O-, where n is from 7 to 16; or (ii) -O-$(CH_2CH_2O)_m$-, where m is 2 to 5;

R″ is a C_{3-12} alkylene group, which chain may be interrupted by one or more heteroatoms~~,~~ ~~e.g.~~selected from O, S, NR^{N2} ~~(~~where R^{N2} is H or C_{1-4} alkyl~~)~~, and/or aromatic rings~~,~~ ~~e.g.~~selected from benzene or pyridine;

Y and Y' are selected from O, S, or NH;

$R^{6'}$ and $R^{9'}$ are selected from the same groups as R^6 and R^9 respectively;

either

(i-a) R^{10} and R^{11} together form a double bond between the N and C atoms to which they are bound; or

(i-b) R^{10} is H and R^{11} is selected from OH and OR^A, where R^A is C_{1-4} alkyl; or

(i-c) R^{10} and R^{11} are both H;

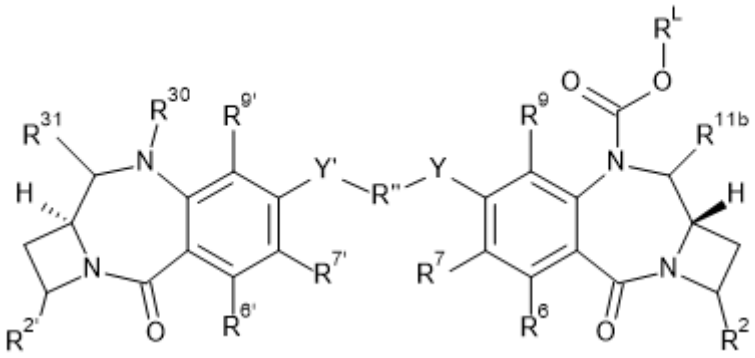
either

(ii-a) R^{20} and R^{21} together form a double bond between the N and C atoms to which they are bound; or

(ii-b) R^{20} is H and R^{21} is selected from OH and OR^B, where R^B is C_{1-4} alkyl; or

(ii-c) R^{20} and R^{21} are both H.

A second aspect of the present invention comprises a compound with the formula **I**:



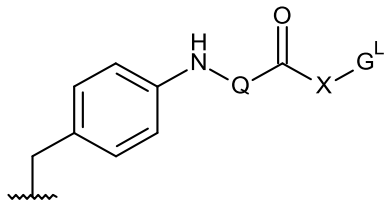
and salts and solvates thereof, wherein:

Y, Y', R", R^2, $R^{2'}$, R^6, $R^{6'}$, R^7, $R^{7'}$, R^9 and $R^{9'}$ are as defined in the first aspect of the invention;

R^{11b} is selected from OH, OR^A, where R^A is C_{1-4} alkyl; and

R^L is a linker for connection to a cell binding agent, which is selected from:

(iiia):

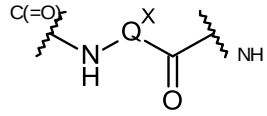


IIIa

,

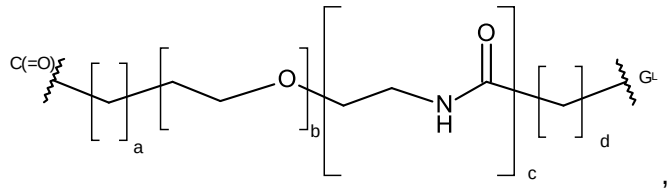
wherein

Q is:



, where Q^X is such that Q is an amino-acid residue, a dipeptide residue or a tripeptide residue;

X is:

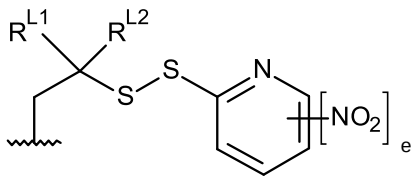


,

where a = 0 to 5, b = 0 to 16, c = 0 or 1, d = 0 to 5;

G^L is a linker for connecting to a Ligand Unit; and

(iiib):



IIIb

,

where R^{L1} and R^{L2} are independently selected from H and methyl, or together with the carbon atom to which they are bound form a cyclopropylene or cyclobutylene group;

and e is 0 or 1;

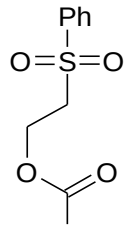
either:

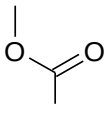
(a) R^{30} and R^{31} together form a double bond between the N and C atoms to which they are bound; or

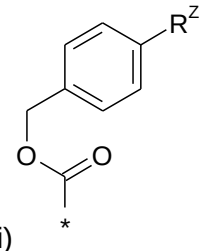
(b) R^{30} is H and R^{31} is selected from OH and OR^B, where R^B is C_{1-4} alkyl;

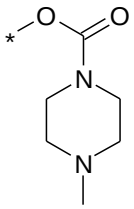
(c) R^{30} and R^{31} are both H; or

(d) R^{31} is OH or OR^B, where R^B is C_{1-4} alkyl and R^{30} is selected from:

(i)  ;

(ii)  ;

(iii)  , where R^Z is selected from:

(z-i)  ;

(z-ii) $OC(=O)CH_3$;

(z-iii) NO_2;

(z-iv) OMe;

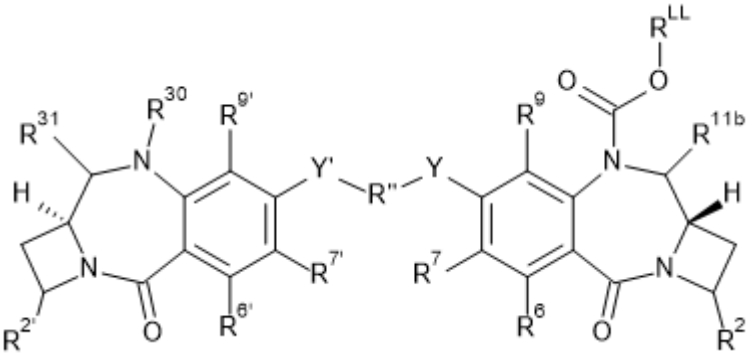
(z-v) glucoronide;

(z-vi) NH-C(=O)-X_1-NHC(=O)X_2-NH-C(=O)-R^{ZC}, where -C(=O)-X_1-NH- and -C(=O)-X_2-NH- represent natural amino acid residues and R^{ZC} is selected from Me, OMe, CH_2CH_2OMe, and $(CH_2CH_2O)_2Me$.

A third aspect of the present invention provides Conjugates of formula II:

L - $(D^L)_p$ (II)

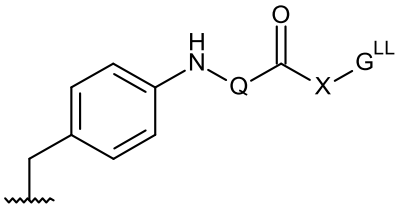
wherein L is a Ligand unit (i.e., a targeting agent), D^L is a Drug Linker unit of formula I'*:



wherein R^2, $R^{2'}$, R^6, R^7, R^9, R^{11b}, Y, R", Y', $R^{6'}$, $R^{7'}$, $R^{9'}$, R^{30} and R^{31} are as defined in the second aspect of the invention;

R^{LL} is a linker for connection to a cell binding agent, which is selected from:

(iiia):

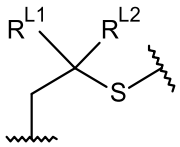


IIIa'

,

where Q and X are as defined in the first aspect and G^{LL} is a linker connected to a Ligand Unit; and

(iiib):



IIIb'

,

where R^{L1} and R^{L2} are as defined in the first aspect;

wherein p is an integer of from 1 to 20.

The Ligand unit, described more fully below, is a targeting agent that binds to a target moiety. The Ligand unit can, for example, specifically bind to a cell component (a Cell Binding Agent) or to other target molecules of interest. The Ligand unit can be, for example, a protein, polypeptide or peptide, such as an antibody, an antigen-binding fragment of an antibody, or other binding agent, such as an Fc fusion protein.

A fourth aspect of the present invention provides the use of a conjugate of the third aspect of the invention in the manufacture of a medicament for treating a proliferative disease. The fourth aspect also provides a conjugate of the third aspect of the invention for use in the treatment of a proliferative disease. The fourth aspect also provides a method of

treating a proliferative disease comprising administering a therapeutically effective amount of a conjugate of the second aspect of the invention to a patient in need thereof.

One of ordinary skill in the art is readily able to determine whether or not a candidate conjugate treats a proliferative condition for any particular cell type. For example, assays which may conveniently be used to assess the activity offered by a particular compound are described in the examples below.

A fifth aspect of the present invention provides the synthesis of a conjugate of the third aspect of the invention comprising conjugating a compound (drug linker) of the second aspect of the invention with a Ligand Unit.

Compounds of formula IV are the warheads released by conjugates of the third aspect.

Definitions

Substituents

The phrase “optionally substituted” as used herein, pertains to a parent group which may be unsubstituted or which may be substituted.

Unless otherwise specified, the term “substituted” as used herein, pertains to a parent group which bears one or more substituents. The term “substituent” is used herein in the conventional sense and refers to a chemical moiety which is covalently attached to, or if appropriate, fused to, a parent group. A wide variety of substituents are well known, and methods for their formation and introduction into a variety of parent groups are also well known.

Examples of substituents are described in more detail below.

C_{1-12} alkyl: The term “C_{1-12} alkyl” as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from a carbon atom of a hydrocarbon compound having from 1 to 12 carbon atoms, which may be aliphatic ~~or alicyclic,~~ and which may be saturated ~~or unsaturated (e.g. partially unsaturated, fully unsaturated)~~. The term “C_{1-4} alkyl” as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from a carbon atom of a hydrocarbon compound having from 1 to 4 carbon atoms, which ~~may be~~is aliphatic ~~or alicyclic~~, and which ~~may be~~is saturated ~~or unsaturated (e.g. partially~~

~~unsaturated, fully unsaturated)~~. The~~us, the~~ terms ~~"alkyl" includes the sub-classes~~ alkenyl, alkynyl, cycloalkyl, etc., are discussed below.

Examples of saturated alkyl groups include, but are not limited to, methyl (C_1), ethyl (C_2), propyl (C_3), butyl (C_4), pentyl (C_5), hexyl (C_6) and heptyl (C_7).

Examples of saturated linear alkyl groups include, but are not limited to, methyl (C_1), ethyl (C_2), n-propyl (C_3), n-butyl (C_4), n-pentyl (amyl) (C_5), n-hexyl (C_6) and n-heptyl (C_7).

Examples of saturated branched alkyl groups include iso-propyl (C_3), iso-butyl (C_4), sec-butyl (C_4), tert-butyl (C_4), iso-pentyl (C_5), and neo-pentyl (C_5).

C_{2-12} Alkenyl: The term "C_{2-12} alkenyl" as used herein, pertains to an alkyl group having one or more carbon-carbon double bonds.

Examples of unsaturated alkenyl groups include, but are not limited to, ethenyl (vinyl, -$CH=CH_2$), 1-propenyl (-$CH=CH-CH_3$), 2-propenyl (allyl, -$CH-CH=CH_2$), isopropenyl (1-methylvinyl, -$C(CH_3)=CH_2$), butenyl (C_4), pentenyl (C_5), and hexenyl (C_6).

C_{2-12} alkynyl: The term "C_{2-12} alkynyl" as used herein, pertains to an alkyl group having one or more carbon-carbon triple bonds.

Examples of unsaturated alkynyl groups include, but are not limited to, ethynyl (-C≡CH) and 2-propynyl (propargyl, -CH_2-C≡CH).

C_{3-12} cycloalkyl: The term "C_{3-12} cycloalkyl" as used herein, pertains to an alkyl group which is also a cyclyl group; that is, a monovalent moiety obtained by removing a hydrogen atom from an alicyclic ring atom of a cyclic hydrocarbon (carbocyclic) compound, which moiety has from 3 to 7 carbon atoms, including from 3 to 7 ring atoms.

Examples of cycloalkyl groups include, but are not limited to, those derived from:

saturated monocyclic hydrocarbon compounds:

cyclopropane (C_3), cyclobutane (C_4), cyclopentane (C_5), cyclohexane (C_6), cycloheptane (C_7), methylcyclopropane (C_4), dimethylcyclopropane (C_5), methylcyclobutane (C_5), dimethylcyclobutane (C_6), methylcyclopentane (C_6), dimethylcyclopentane (C_7) and methylcyclohexane (C_7);

unsaturated monocyclic hydrocarbon compounds:
cyclopropene (C_3), cyclobutene (C_4), cyclopentene (C_5), cyclohexene (C_6), methylcyclopropene (C_4), dimethylcyclopropene (C_5), methylcyclobutene (C_5), dimethylcyclobutene (C_6), methylcyclopentene (C_6), dimethylcyclopentene (C_7) and methylcyclohexene (C_7); and

saturated polycyclic hydrocarbon compounds:
norcarane (C_7), norpinane (C_7), norbornane (C_7).

C_{3-20} heterocyclyl: The term "C_{3-20} heterocyclyl" as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from a ring atom of a heterocyclic compound, which moiety has from 3 to 20 ring atoms, of which from 1 to 10 are ring heteroatoms. Preferably, each ring has from 3 to 7 ring atoms, of which from 1 to 4 are ring heteroatoms.

In this context, the prefixes (e.g. C_{3-20}, C_{3-7}, C_{5-6}, etc.) denote the number of ring atoms, or range of number of ring atoms, whether carbon atoms or heteroatoms. For example, the term "C_{5-6}heterocyclyl", as used herein, pertains to a heterocyclyl group having 5 or 6 ring atoms.

Examples of monocyclic heterocyclyl groups include, but are not limited to, those derived from:
N_1: aziridine (C_3), azetidine (C_4), pyrrolidine (tetrahydropyrrole) (C_5), pyrroline (e.g., 3-pyrroline, 2,5-dihydropyrrole) (C_5), 2H-pyrrole or 3H-pyrrole (isopyrrole, isoazole) (C_5), piperidine (C_6), dihydropyridine (C_6), tetrahydropyridine (C_6), azepine (C_7);
O_1: oxirane (C_3), oxetane (C_4), oxolane (tetrahydrofuran) (C_5), oxole (dihydrofuran) (C_5), oxane (tetrahydropyran) (C_6), dihydropyran (C_6), pyran (C_6), oxepin (C_7);
S_1: thiirane (C_3), thietane (C_4), thiolane (tetrahydrothiophene) (C_5), thiane (tetrahydrothiopyran) (C_6), thiepane (C_7);
O_2: dioxolane (C_5), dioxane (C_6), and dioxepane (C_7);
O_3: trioxane (C_6);
N_2: imidazolidine (C_5), pyrazolidine (diazolidine) (C_5), imidazoline (C_5), pyrazoline (dihydropyrazole) (C_5), piperazine (C_6);
N_1O_1: tetrahydrooxazole (C_5), dihydrooxazole (C_5), tetrahydroisoxazole (C_5), dihydroisoxazole (C_5), morpholine (C_6), tetrahydrooxazine (C_6), dihydrooxazine (C_6), oxazine (C_6);
N_1S_1: thiazoline (C_5), thiazolidine (C_5), thiomorpholine (C_6);

N_2O_1: oxadiazine (C_6);
O_1S_1: oxathiole (C_5) and oxathiane (thioxane) (C_6); and,
$N_1O_1S_1$: oxathiazine (C_6).

Examples of substituted monocyclic heterocyclyl groups include those derived from saccharides, in cyclic form, for example, furanoses (C_5), such as arabinofuranose, lyxofuranose, ribofuranose, and xylofuranse, and pyranoses (C_6), such as allopyranose, altropyranose, glucopyranose, mannopyranose, gulopyranose, idopyranose, galactopyranose, and talopyranose.

C_{5-20} aryl: The term "C_{5-20} aryl", as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from an aromatic ring atom of an aromatic compound, which moiety has from 3 to 20 ring atoms. The term "C_{5-7} aryl", as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from an aromatic ring atom of an aromatic compound, which moiety has from 5 to 7 ring atoms and the term "C_{5-10} aryl", as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from an aromatic ring atom of an aromatic compound, which moiety has from 5 to 10 ring atoms. Preferably, each ring has from 5 to 7 ring atoms.

In this context, the prefixes (e.g. C_{3-20}, C_{5-7}, C_{5-6}, C_{5-10}, etc.) denote the number of ring atoms, or range of number of ring atoms, whether carbon atoms or heteroatoms. For example, the term "C_{5-6} aryl" as used herein, pertains to an aryl group having 5 or 6 ring atoms.

The ring atoms may be all carbon atoms, as in "carboaryl groups".
Examples of carboaryl groups include, but are not limited to, those derived from benzene (i.e. phenyl) (C_6), naphthalene (C_{10}), azulene (C_{10}), anthracene (C_{14}), phenanthrene (C_{14}), naphthacene (C_{18}), and pyrene (C_{16}).

Examples of aryl groups which comprise fused rings, at least one of which is an aromatic ring, include, but are not limited to, groups derived from indane (e.g. 2,3-dihydro-1H-indene) (C_9), indene (C_9), isoindene (C_9), tetraline (1,2,3,4-tetrahydronaphthalene (C_{10}), acenaphthene (C_{12}), fluorene (C_{13}), phenalene (C_{13}), acephenanthrene (C_{15}), and aceanthrene (C_{16}).

Alternatively, the ring atoms may include one or more heteroatoms, as in “heteroaryl groups”. Examples of monocyclic heteroaryl groups include, but are not limited to, those derived from:

N_1: pyrrole (azole) (C_5), pyridine (azine) (C_6);

O_1: furan (oxole) (C_5);

S_1: thiophene (thiole) (C_5);

N_1O_1: oxazole (C_5), isoxazole (C_5), isoxazine (C_6);

N_2O_1: oxadiazole (furazan) (C_5);

N_3O_1: oxatriazole (C_5);

N_1S_1: thiazole (C_5), isothiazole (C_5);

N_2: imidazole (1,3-diazole) (C_5), pyrazole (1,2-diazole) (C_5), pyridazine (1,2-diazine) (C_6), pyrimidine (1,3-diazine) (C_6) (e.g., cytosine, thymine, uracil), pyrazine (1,4-diazine) (C_6);

N_3: triazole (C_5), triazine (C_6); and,

N_4: tetrazole (C_5).

Examples of heteroaryl which comprise fused rings, include, but are not limited to:

C_9 (with 2 fused rings) derived from benzofuran (O_1), isobenzofuran (O_1), indole (N_1), isoindole (N_1), indolizine (N_1), indoline (N_1), isoindoline (N_1), purine (N_4) (e.g., adenine, guanine), benzimidazole (N_2), indazole (N_2), benzoxazole (N_1O_1), benzisoxazole (N_1O_1), benzodioxole (O_2), benzofurazan (N_2O_1), benzotriazole (N_3), benzothiofuran (S_1), benzothiazole (N_1S_1), benzothiadiazole (N_2S);

C_{10} (with 2 fused rings) derived from chromene (O_1), isochromene (O_1), chroman (O_1), isochroman (O_1), benzodioxan (O_2), quinoline (N_1), isoquinoline (N_1), quinolizine (N_1), benzoxazine (N_1O_1), benzodiazine (N_2), pyridopyridine (N_2), quinoxaline (N_2), quinazoline (N_2), cinnoline (N_2), phthalazine (N_2), naphthyridine (N_2), pteridine (N_4);

C_{11} (with 2 fused rings) derived from benzodiazepine (N_2);

C_{13} (with 3 fused rings) derived from carbazole (N_1), dibenzofuran (O_1), dibenzothiophene (S_1), carboline (N_2), perimidine (N_2), pyridoindole (N_2); and,

C_{14} (with 3 fused rings) derived from acridine (N_1), xanthene (O_1), thioxanthene (S_1), oxanthrene (O_2), phenoxathiin (O_1S_1), phenazine (N_2), phenoxazine (N_1O_1), phenothiazine (N_1S_1), thianthrene (S_2), phenanthridine (N_1), phenanthroline (N_2), phenazine (N_2).

The above groups, whether alone or part of another substituent, may themselves optionally be substituted with one or more groups selected from themselves and the additional substituents listed below.

Halo: -F, -Cl, -Br, and -I.

Hydroxy: -OH.

Ether: -OR, wherein R is an ether substituent, for example, a C_{1-7} alkyl group (also referred to as a C_{1-7} alkoxy group, discussed below), a C_{3-20} heterocyclyl group (also referred to as a C_{3-20} heterocyclyloxy group), or a C_{5-20} aryl group (also referred to as a C_{5-20} aryloxy group), preferably a C_{1-7}alkyl group.

Alkoxy: -OR, wherein R is an alkyl group, for example, a C_{1-7} alkyl group. Examples of C_{1-7} alkoxy groups include, but are not limited to, -OMe (methoxy), -OEt (ethoxy), -O(nPr) (n-propoxy), -O(iPr) (isopropoxy), -O(nBu) (n-butoxy), -O(sBu) (sec-butoxy), -O(iBu) (isobutoxy), and -O(tBu) (tert-butoxy).

Acetal: $-CH(OR^1)(OR^2)$, wherein R^1 and R^2 are independently acetal substituents, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group, or, in the case of a “cyclic” acetal group, R^1 and R^2, taken together with the two oxygen atoms to which they are attached, and the carbon atoms to which they are attached, form a heterocyclic ring having from 4 to 8 ring atoms. Examples of acetal groups include, but are not limited to, $-CH(OMe)_2$, $-CH(OEt)_2$, and -CH(OMe)(OEt).

Hemiacetal: $-CH(OH)(OR^1)$, wherein R^1 is a hemiacetal substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of hemiacetal groups include, but are not limited to, -CH(OH)(OMe) and -CH(OH)(OEt).

Ketal: $-CR(OR^1)(OR^2)$, where R^1 and R^2 are as defined for acetals, and R is a ketal substituent other than hydrogen, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples ketal groups include, but are not limited to, $-C(Me)(OMe)_2$, $-C(Me)(OEt)_2$, -C(Me)(OMe)(OEt), $-C(Et)(OMe)_2$, $-C(Et)(OEt)_2$, and -C(Et)(OMe)(OEt).

Hemiketal: $-CR(OH)(OR^1)$, where R^1 is as defined for hemiacetals, and R is a hemiketal substituent other than hydrogen, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of hemiacetal groups include,

but are not limited to, -C(Me)(OH)(OMe), -C(Et)(OH)(OMe), -C(Me)(OH)(OEt), and -C(Et)(OH)(OEt).

Oxo (keto, -one): =O.

Thione (thioketone): =S.

Imino (imine): =NR, wherein R is an imino substituent, for example, hydrogen, C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably hydrogen or a C_{1-7} alkyl group. Examples of ester groups include, but are not limited to, =NH, =NMe, =NEt, and =NPh.

Formyl (carbaldehyde, carboxaldehyde): -C(=O)H.

Acyl (keto): -C(=O)R, wherein R is an acyl substituent, for example, a C_{1-7} alkyl group (also referred to as C_{1-7} alkylacyl or C_{1-7} alkanoyl), a C_{3-20} heterocyclyl group (also referred to as C_{3-20} heterocyclylacyl), or a C_{5-20} aryl group (also referred to as C_{5-20} arylacyl), preferably a C_{1-7} alkyl group. Examples of acyl groups include, but are not limited to, -C(=O)CH_3 (acetyl), -C(=O)CH_2CH_3 (propionyl), -C(=O)$C(CH_3)_3$ (t-butyryl), and -C(=O)Ph (benzoyl, phenone).

Carboxy (carboxylic acid): -C(=O)OH.

Thiocarboxy (thiocarboxylic acid): -C(=S)SH.

Thiolocarboxy (thiolocarboxylic acid): -C(=O)SH.

Thionocarboxy (thionocarboxylic acid): -C(=S)OH.

Imidic acid: -C(=NH)OH.

Hydroxamic acid: -C(=NOH)OH.

Ester (carboxylate, carboxylic acid ester, oxycarbonyl): -C(=O)OR, wherein R is an ester substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group,

preferably a C_{1-7} alkyl group. Examples of ester groups include, but are not limited to, -C(=O)OCH_3, -C(=O)OCH_2CH_3, -C(=O)$OC(CH_3)_3$, and -C(=O)OPh.

Acyloxy (reverse ester): -OC(=O)R, wherein R is an acyloxy substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of acyloxy groups include, but are not limited to, -OC(=O)CH_3 (acetoxy), -OC(=O)CH_2CH_3, -OC(=O)$C(CH_3)_3$, -OC(=O)Ph, and -OC(=O)CH_2Ph.

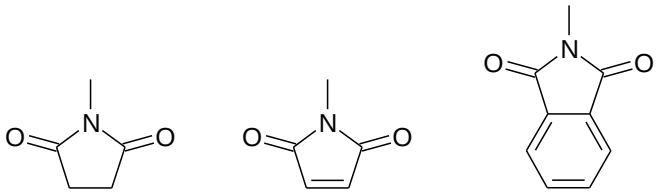
Oxycarboyloxy: -OC(=O)OR, wherein R is an ester substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of ester groups include, but are not limited to, -OC(=O)OCH_3, -OC(=O)OCH_2CH_3, -OC(=O)$OC(CH_3)_3$, and -OC(=O)OPh.

Amino: -NR^1R^2, wherein R^1 and R^2 are independently amino substituents, for example, hydrogen, a C_{1-7} alkyl group (also referred to as C_{1-7} alkylamino or di-C_{1-7} alkylamino), a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably H or a C_{1-7} alkyl group, or, in the case of a “cyclic” amino group, R^1 and R^2, taken together with the nitrogen atom to which they are attached, form a heterocyclic ring having from 4 to 8 ring atoms. Amino groups may be primary (-NH_2), secondary (-NHR^1), or tertiary (-NHR^1R^2), and in cationic form, may be quaternary (-$^+NR^1R^2R^3$). Examples of amino groups include, but are not limited to, -NH_2, -$NHCH_3$, -$NHC(CH_3)_2$, -$N(CH_3)_2$, -$N(CH_2CH_3)_2$, and -NHPh. Examples of cyclic amino groups include, but are not limited to, aziridino, azetidino, pyrrolidino, piperidino, piperazino, morpholino, and thiomorpholino.

Amido (carbamoyl, carbamyl, aminocarbonyl, carboxamide): -C(=O)NR^1R^2, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of amido groups include, but are not limited to, -C(=O)NH_2, -C(=O)$NHCH_3$, -C(=O)$N(CH_3)_2$, -C(=O)$NHCH_2CH_3$, and -C(=O)$N(CH_2CH_3)_2$, as well as amido groups in which R^1 and R^2, together with the nitrogen atom to which they are attached, form a heterocyclic structure as in, for example, piperidinocarbonyl, morpholinocarbonyl, thiomorpholinocarbonyl, and piperazinocarbonyl.

Thioamido (thiocarbamyl): -C(=S)NR^1R^2, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of amido groups include, but are not limited to, -C(=S)NH_2, -C(=S)$NHCH_3$, -C(=S)$N(CH_3)_2$, and -C(=S)$NHCH_2CH_3$.

Acylamido (acylamino): -$NR^1C(=O)R^2$, wherein R^1 is an amide substituent, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably hydrogen or a C_{1-7} alkyl group, and R^2 is an acyl substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20}aryl group, preferably hydrogen or a C_{1-7} alkyl group. Examples of acylamide groups include, but are not limited to, -$NHC(=O)CH_3$, -$NHC(=O)CH_2CH_3$, and -NHC(=O)Ph. R^1 and R^2 may together form a cyclic structure, as in, for example, succinimidyl, maleimidyl, and phthalimidyl:



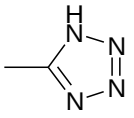
succinimidyl maleimidyl phthalimidyl

Aminocarbonyloxy: -$OC(=O)NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of aminocarbonyloxy groups include, but are not limited to, -$OC(=O)NH_2$, -OC(=O)NHMe, -$OC(=O)NMe_2$, and -$OC(=O)NEt_2$.

Ureido: -$N(R^1)CONR^2R^3$ wherein R^2 and R^3 are independently amino substituents, as defined for amino groups, and R^1 is a ureido substituent, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably hydrogen or a C_{1-7} alkyl group. Examples of ureido groups include, but are not limited to, -$NHCONH_2$, -NHCONHMe, -NHCONHEt, -$NHCONMe_2$, -$NHCONEt_2$, -$NMeCONH_2$, -NMeCONHMe, -NMeCONHEt, -$NMeCONMe_2$, and -$NMeCONEt_2$.

Guanidino: -NH-$C(=NH)NH_2$.

Tetrazolyl: a five membered aromatic ring having four nitrogen atoms and one carbon atom,



Imino: =NR, wherein R is an imino substituent, for example, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably H or a C_{1-7}alkyl group. Examples of imino groups include, but are not limited to, =NH, =NMe, and =NEt.

Amidine (amidino): -C(=NR)NR_2, wherein each R is an amidine substituent, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably H or a C_{1-7} alkyl group. Examples of amidine groups include, but are not limited to, -C(=NH)NH_2, -C(=NH)NMe_2, and -C(=NMe)NMe_2.

Nitro: -NO_2.

Nitroso: -NO.

Azido: -N_3.

Cyano (nitrile, carbonitrile): -CN.

Isocyano: -NC.

Cyanato: -OCN.

Isocyanato: -NCO.

Thiocyano (thiocyanato): -SCN.

Isothiocyano (isothiocyanato): -NCS.

Sulfhydryl (thiol, mercapto): -SH.

Thioether (sulfide): -SR, wherein R is a thioether substituent, for example, a C_{1-7} alkyl group (also referred to as a C_{1-7}alkylthio group), a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of C_{1-7} alkylthio groups include, but are not limited to, -SCH_3 and -SCH_2CH_3.

Disulfide: -SS-R, wherein R is a disulfide substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group (also referred to herein as C_{1-7} alkyl disulfide). Examples of C_{1-7} alkyl disulfide groups include, but are not limited to, -$SSCH_3$ and -$SSCH_2CH_3$.

Sulfine (sulfinyl, sulfoxide): -S(=O)R, wherein R is a sulfine substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfine groups include, but are not limited to, $-S(=O)CH_3$ and $-S(=O)CH_2CH_3$.

Sulfone (sulfonyl): $-S(=O)_2R$, wherein R is a sulfone substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group, including, for example, a fluorinated or perfluorinated C_{1-7} alkyl group. Examples of sulfone groups include, but are not limited to, $-S(=O)_2CH_3$ (methanesulfonyl, mesyl), $-S(=O)_2CF_3$ (triflyl), $-S(=O)_2CH_2CH_3$ (esyl), $-S(=O)_2C_4F_9$ (nonaflyl), $-S(=O)_2CH_2CF_3$ (tresyl), $-S(=O)_2CH_2CH_2NH_2$ (tauryl), $-S(=O)_2Ph$ (phenylsulfonyl, besyl), 4-methylphenylsulfonyl (tosyl), 4-chlorophenylsulfonyl (closyl), 4-bromophenylsulfonyl (brosyl), 4-nitrophenyl (nosyl), 2-naphthalenesulfonate (napsyl), and 5-dimethylamino-naphthalen-1-ylsulfonate (dansyl).

Sulfinic acid (sulfino): $-S(=O)OH$, $-SO_2H$.

Sulfonic acid (sulfo): $-S(=O)_2OH$, $-SO_3H$.

Sulfinate (sulfinic acid ester): -S(=O)OR; wherein R is a sulfinate substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfinate groups include, but are not limited to, $-S(=O)OCH_3$ (methoxysulfinyl; methyl sulfinate) and $-S(=O)OCH_2CH_3$ (ethoxysulfinyl; ethyl sulfinate).

Sulfonate (sulfonic acid ester): $-S(=O)_2OR$, wherein R is a sulfonate substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfonate groups include, but are not limited to, $-S(=O)_2OCH_3$ (methoxysulfonyl; methyl sulfonate) and $-S(=O)_2OCH_2CH_3$ (ethoxysulfonyl; ethyl sulfonate).

Sulfinyloxy: -OS(=O)R, wherein R is a sulfinyloxy substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfinyloxy groups include, but are not limited to, $-OS(=O)CH_3$ and $-OS(=O)CH_2CH_3$.

Sulfonyloxy: $-OS(=O)_2R$, wherein R is a sulfonyloxy substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group.

Examples of sulfonyloxy groups include, but are not limited to, -$OS(=O)_2CH_3$ (mesylate) and -$OS(=O)_2CH_2CH_3$ (esylate).

Sulfate: -$OS(=O)_2OR$; wherein R is a sulfate substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfate groups include, but are not limited to, -$OS(=O)_2OCH_3$ and -$SO(=O)_2OCH_2CH_3$.

Sulfamyl (sulfamoyl; sulfinic acid amide; sulfinamide): -$S(=O)NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of sulfamyl groups include, but are not limited to, -$S(=O)NH_2$, -$S(=O)NH(CH_3)$, -$S(=O)N(CH_3)_2$, -$S(=O)NH(CH_2CH_3)$, -$S(=O)N(CH_2CH_3)_2$, and -S(=O)NHPh.

Sulfonamido (sulfinamoyl; sulfonic acid amide; sulfonamide): -$S(=O)_2NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of sulfonamido groups include, but are not limited to, -$S(=O)_2NH_2$, -$S(=O)_2NH(CH_3)$, -$S(=O)_2N(CH_3)_2$, -$S(=O)_2NH(CH_2CH_3)$, -$S(=O)_2N(CH_2CH_3)_2$, and -$S(=O)_2$NHPh.

Sulfamino: -$NR^1S(=O)_2OH$, wherein R^1 is an amino substituent, as defined for amino groups. Examples of sulfamino groups include, but are not limited to, -$NHS(=O)_2OH$ and -$N(CH_3)S(=O)_2OH$.

Sulfonamino: -$NR^1S(=O)_2R$, wherein R^1 is an amino substituent, as defined for amino groups, and R is a sulfonamino substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfonamino groups include, but are not limited to, -$NHS(=O)_2CH_3$ and -$N(CH_3)S(=O)_2C_6H_5$.

Sulfinamino: -$NR^1S(=O)R$, wherein R^1 is an amino substituent, as defined for amino groups, and R is a sulfinamino substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfinamino groups include, but are not limited to, -$NHS(=O)CH_3$ and -$N(CH_3)S(=O)C_6H_5$.

Phosphino (phosphine): -PR_2, wherein R is a phosphino substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphino groups include, but are not limited to, -PH_2, -$P(CH_3)_2$, -$P(CH_2CH_3)_2$, -P(t-Bu)$_2$, and -$P(Ph)_2$.

Phospho: $-P(=O)_2$.

Phosphinyl (phosphine oxide): $-P(=O)R_2$, wherein R is a phosphinyl substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group or a C_{5-20} aryl group. Examples of phosphinyl groups include, but are not limited to, $-P(=O)(CH_3)_2$, $-P(=O)(CH_2CH_3)_2$, $-P(=O)(t\text{-}Bu)_2$, and $-P(=O)(Ph)_2$.

Phosphonic acid (phosphono): $-P(=O)(OH)_2$.

Phosphonate (phosphono ester): $-P(=O)(OR)_2$, where R is a phosphonate substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphonate groups include, but are not limited to, $-P(=O)(OCH_3)_2$, $-P(=O)(OCH_2CH_3)_2$, $-P(=O)(O\text{-}t\text{-}Bu)_2$, and $-P(=O)(OPh)_2$.

Phosphoric acid (phosphonooxy): $-OP(=O)(OH)_2$.

Phosphate (phosphonooxy ester): $-OP(=O)(OR)_2$, where R is a phosphate substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphate groups include, but are not limited to, $-OP(=O)(OCH_3)_2$, $-OP(=O)(OCH_2CH_3)_2$, $-OP(=O)(O\text{-}t\text{-}Bu)_2$, and $-OP(=O)(OPh)_2$.

Phosphorous acid: $-OP(OH)_2$.

Phosphite: $-OP(OR)_2$, where R is a phosphite substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphite groups include, but are not limited to, $-OP(OCH_3)_2$, $-OP(OCH_2CH_3)_2$, $-OP(O\text{-}t\text{-}Bu)_2$, and $-OP(OPh)_2$.

Phosphoramidite: $-OP(OR^1)\text{-}NR^2{}_2$, where R^1 and R^2 are phosphoramidite substituents, for example, -H, a (optionally substituted) C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphoramidite groups include, but are not limited to, $-OP(OCH_2CH_3)\text{-}N(CH_3)_2$, $-OP(OCH_2CH_3)\text{-}N(i\text{-}Pr)_2$, and $-OP(OCH_2CH_2CN)\text{-}N(i\text{-}Pr)_2$.

Phosphoramidate: -OP(=O)(OR^1)-NR$^2_2$, where R^1 and R^2 are phosphoramidate substituents, for example, -H, a (optionally substituted) C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphoramidate groups include, but are not limited to, -OP(=O)(OCH$_2$CH$_3$)-N(CH$_3$)$_2$, -OP(=O)(OCH$_2$CH$_3$)-N(i-Pr)$_2$, and -OP(=O)(OCH$_2$CH$_2$CN)-N(i-Pr)$_2$.

Alkylene

C_{3-12} alkylene: The term "C_{3-12} alkylene", as used herein, pertains to a bidentate moiety obtained by removing two hydrogen atoms, either both from the same carbon atom, or one from each of two different carbon atoms, of a hydrocarbon compound having from 3 to 12 carbon atoms (unless otherwise specified), which ~~may be~~are aliphatic~~ or alicyclic~~, and which ~~may be~~are saturated~~, partially unsaturated, or fully unsaturated~~. Th~~eus, the~~ terms ~~"alkylene" includes the sub-classes~~ alkenylene, alkynylene, cycloalkylene, etc., are discussed below.

Examples of linear saturated C_{3-12} alkylene groups include, but are not limited to, -(CH$_2$)$_n$- where n is an integer from 3 to 12, for example, -CH$_2$CH$_2$CH$_2$- (propylene), -CH$_2$CH$_2$CH$_2$CH$_2$- (butylene), -CH$_2$CH$_2$CH$_2$CH$_2$CH$_2$- (pentylene) and -CH$_2$CH$_2$CH$_2$CH-$_2$CH$_2$CH$_2$CH$_2$- (heptylene).

Examples of branched saturated C_{3-12} alkylene groups include, but are not limited to, -CH(CH$_3$)CH$_2$-, -CH(CH$_3$)CH$_2$CH$_2$-, -CH(CH$_3$)CH$_2$CH$_2$CH$_2$-, -CH$_2$CH(CH$_3$)CH$_2$-, -CH$_2$CH(CH$_3$)CH$_2$CH$_2$-, -CH(CH$_2$CH$_3$)-, -CH(CH$_2$CH$_3$)CH$_2$-, and -CH$_2$CH(CH$_2$CH$_3$)CH$_2$-.

Examples of ~~linear partially unsaturated C_{3-12} alkylene groups (~~C_{3-12} alkenylene, and alkynylene groups~~)~~ include, but are not limited to, -CH=CH-CH$_2$-, -CH$_2$-CH=CH$_2$-, -CH=CH-CH$_2$-CH$_2$-, -CH=CH-CH$_2$-CH$_2$-CH$_2$-, -CH=CH-CH=CH-, -CH=CH-CH=CH-CH$_2$-, -CH=CH-CH=CH-CH$_2$-CH$_2$-, -CH=CH-CH$_2$-CH=CH-, -CH=CH-CH$_2$-CH$_2$-CH=CH-, and -CH$_2$-C≡C-CH$_2$-.

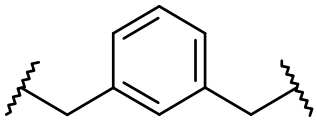
Examples of branched ~~partially unsaturated C_{3-12} alkylene groups (~~C_{3-12} alkenylene and alkynylene groups~~)~~ include, but are not limited to, -C(CH$_3$)=CH-, -C(CH$_3$)=CH-CH$_2$-, -CH=CH-CH(CH$_3$)- and -C≡C-CH(CH$_3$)-.

Examples of ~~alicyclic saturated C_{3-12} alkylene groups (~~C_{3-12} cycloalkylenes~~)~~ include, but are not limited to, cyclopentylene (e.g. cyclopent-1,3-ylene), and cyclohexylene (e.g. cyclohex-1,4-ylene).

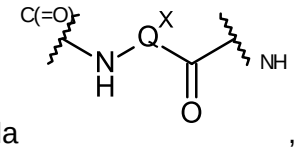
Examples of ~~alicyclic partially unsaturated C_{3-12} alkylene groups (~~C_{3-12} cycloalkylenes~~)~~ include, but are not limited to, cyclopentenylene (e.g. 4-cyclopenten-1,3-ylene), cyclohexenylene (e.g. 2-cyclohexen-1,4-ylene; 3-cyclohexen-1,2-ylene; 2,5-cyclohexadien-1,4-ylene).

Where the C_{3-12} alkylene group is interrupted by a heteroatom, the subscript refers to the number of atoms in the chain including the heteroatoms. For example, the chain -C_2H_4-O-C_2H_4- would be a C_5 group.

Where the C_{3-12} alkylene group is interrupted by an aromatic ring, the subscript refers to the number of atoms directly in the chain including the aromatic ring. For example, the chain



would be a C_5 group.



Connection labels: In the formula , the superscripted labels $^{C(=O)}$ and ^{NH} indicate the group to which the atoms are bound. For example, the NH group is shown as being bound to a carbonyl (which is not part of the moiety illustrated), and the carbonyl is shown as being bound to a NH group (which is not part of the moiety illustrated).

Ligand Unit

The Ligand Unit may be of any kind, and include a protein, polypeptide, peptide and a non-peptidic agent that specifically binds to a target molecule. In some embodiments, the Ligand unit may be a protein, polypeptide or peptide. In some embodiments, the Ligand unit may be a cyclic polypeptide. These Ligand units can include antibodies or a fragment of an antibody that contains at least one target molecule-binding site, lymphokines, hormones, growth factors, or any other cell binding molecule or substance that can specifically bind to a target.

The terms “specifically binds” and “specific binding” refer to the binding of an antibody or other protein, polypeptide or peptide to a predetermined molecule (e.g., an antigen). Typically, the antibody or other molecule binds with an affinity of at least about $1x10^7$ M^{-1}, and binds to the predetermined molecule with an affinity that is at least two-fold greater than its affinity for binding to a non-specific molecule (e.g., BSA, casein) other than the predetermined molecule or a closely-related molecule.

Examples of Ligand units include those agents described for use in WO 2007/085930, which is incorporated herein.

In some embodiments, the Ligand unit is a Cell Binding Agent that binds to an extracellular target on a cell. Such a Cell Binding Agent can be a protein, polypeptide, peptide or a non-peptidic agent. In some embodiments, the Cell Binding Agent may be a protein, polypeptide or peptide. In some embodiments, the Cell Binding Agent may be a cyclic polypeptide. The Cell Binding Agent also may be antibody or an antigen-binding fragment of an antibody. Thus, in one embodiment, the present invention provides an antibody-drug conjugate (ADC).

Cell Binding Agent

A cell binding agent may be of any kind, and include peptides and non-peptides. These can include antibodies or a fragment of an antibody that contains at least one binding site, lymphokines, hormones, hormone mimetics, vitamins, growth factors, nutrient-transport molecules, or any other cell binding molecule or substance.

Peptides

In one embodiment, the cell binding agent is a linear or cyclic peptide comprising 4-30, preferably 6-20, contiguous amino acid residues. In this embodiment, it is preferred that one cell binding agent is linked to one monomer or dimer azetidobenzodiazepine compound.

In one embodiment the cell binding agent comprises a peptide that binds integrin $\alpha_v\beta_6$. The peptide may be selective for $\alpha_v\beta_6$ over XYS.

In one embodiment the cell binding agent comprises the A20FMDV-Cys polypeptide. The A20FMDV-Cys has the sequence: NAVPNLRGDLQVLAQKVARTC. Alternatively, a variant of the A20FMDV-Cys sequence may be used wherein one, two, three, four, five, six,

seven, eight, nine or ten amino acid residues are substituted with another amino acid residue. Furthermore, the polypeptide may have the sequence NAVXXXXXXXXXXXXXXRTC.

Antibodies

The term “antibody” herein is used in the broadest sense and specifically covers monoclonal antibodies, polyclonal antibodies, dimers, multimers, multispecific antibodies (*e.g.,* bispecific antibodies), multivalent antibodies and antibody fragments, so long as they exhibit the desired biological activity (Miller *et al* (2003) *Jour. of Immunology* 170:4854-4861). Antibodies may be murine, human, humanized, chimeric, or derived from other species. An antibody is a protein generated by the immune system that is capable of recognizing and binding to a specific antigen. (Janeway, C., Travers, P., Walport, M., Shlomchik (2001) *Immuno Biology, 5th Ed*., Garland Publishing, New York). A target antigen generally has numerous binding sites, also called epitopes, recognized by CDRs on multiple antibodies. Each antibody that specifically binds to a different epitope has a different structure. Thus, one antigen may have more than one corresponding antibody. An antibody includes a full-length immunoglobulin molecule or an immunologically active portion of a full-length immunoglobulin molecule, *i.e.*, a molecule that contains an antigen binding site that immunospecifically binds an antigen of a target of interest or part thereof, such targets including but not limited to, cancer cell or cells that produce autoimmune antibodies associated with an autoimmune disease. The immunoglobulin can be of any type (e.g. IgG, IgE, IgM, IgD, and IgA), class (e.g. IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule. The immunoglobulins can be derived from any species, including human, murine, or rabbit origin.

"Antibody fragments" comprise a portion of a full length antibody, generally the antigen binding or variable region thereof. Examples of antibody fragments include Fab, Fab', $F(ab')_2$, and scFv fragments; diabodies; linear antibodies; fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, CDR (complementary determining region), and epitope-binding fragments of any of the above which immunospecifically bind to cancer cell antigens, viral antigens or microbial antigens, single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

The term “monoclonal antibody” as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e. the individual antibodies comprising the population are identical except for possible naturally occurring mutations

that may be present in minor amounts. Monoclonal antibodies are highly specific, being directed against a single antigenic site. Furthermore, in contrast to polyclonal antibody preparations which include different antibodies directed against different determinants (epitopes), each monoclonal antibody is directed against a single determinant on the antigen. In addition to their specificity, the monoclonal antibodies are advantageous in that they may be synthesized uncontaminated by other antibodies. The modifier "monoclonal" indicates the character of the antibody as being obtained from a substantially homogeneous population of antibodies, and is not to be construed as requiring production of the antibody by any particular method. For example, the monoclonal antibodies to be used in accordance with the present invention may be made by the hybridoma method first described by Kohler *et al* (1975) *Nature* 256:495, or may be made by recombinant DNA methods (see, US 4816567). The monoclonal antibodies may also be isolated from phage antibody libraries using the techniques described in Clackson et al (1991) Nature, 352:624-628; Marks et al (1991) J. Mol. Biol., 222:581-597 or from transgenic mice carrying a fully human immunoglobulin system (Lonberg (2008) Curr. Opinion 20(4):450-459).

The monoclonal antibodies herein specifically include chimeric antibodies, humanized antibodies and human antibodies.

Examples of cell binding agents include those agents described for use in WO 2007/085930, which is incorporated herein.

Tumour-associate antigens and cognate antibodies for use in embodiments of the present invention are listed below, and are described in more detail on pages 14 to 86 of WO 2017/186894, which is incorporated herein.

(1) BMPR1B (bone morphogenetic protein receptor-type IB)
(2) E16 (LAT1, SLC7A5)
(3) STEAP1 (six transmembrane epithelial antigen of prostate)
(4) 0772P (CA125, MUC16)
(5) MPF (MPF, MSLN, SMR, megakaryocyte potentiating factor, mesothelin)
(6) Napi3b (NAPI-3B, NPTIIb, SLC34A2, solute carrier family 34 (sodium phosphate), member 2, type II sodium-dependent phosphate transporter 3b)
(7) Sema 5b (FLJ10372, KIAA1445, Mm.42015, SEMA5B, SEMAG, Semaphorin 5b Hlog, 25 sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5B)

(8) PSCA hlg (2700050C12Rik, C530008O16Rik, RIKEN cDNA 2700050C12, RIKEN cDNA
2700050C12 gene)
(9) ETBR (Endothelin type B receptor)
(10) MSG783 (RNF124, hypothetical protein FLJ20315)
(11) STEAP2 (HGNC_8639, IPCA-1, PCANAP1, STAMP1, STEAP2, STMP, prostate cancer
associated gene 1, prostate cancer associated protein 1, six transmembrane epithelial antigen of prostate 2, six transmembrane prostate protein)
(12) TrpM4 (BR22450, FLJ20041, TRPM4, TRPM4B, transient receptor potential cation 5 channel, subfamily M, member 4)
(13) CRIPTO (CR, CR1, CRGF, CRIPTO, TDGF1, teratocarcinoma-derived growth factor)
(14) CD21 (CR2 (Complement receptor 2) or C3DR (C3d/Epstein Barr virus receptor) or Hs.73792)
(15) CD79b (CD79B, CD79β, IGb (immunoglobulin-associated beta), B29)
(16) FcRH2 (IFGP4, IRTA4, SPAP1A (SH2 domain containing phosphatase anchor protein 1a), SPAP1B, SPAP1C)
(17) HER2 (ErbB2)
(18) NCA (CEACAM6)
(19) MDP (DPEP1)
(20) IL20R-alpha (IL20Ra, ZCYTOR7)
(21) Brevican (BCAN, BEHAB)
(22) EphB2R (DRT, ERK, Hek5, EPHT3, Tyro5)
(23) ASLG659 (B7h)
(24) PSCA (Prostate stem cell antigen precursor)
(25) GEDA
(26) BAFF-R (B cell -activating factor receptor, BLyS receptor 3, BR3)
(27) CD22 (B-cell receptor CD22-B isoform, BL-CAM, Lyb-8, Lyb8, SIGLEC-2, FLJ22814)
(27a) CD22 (CD22 molecule)
(28) CD79a (CD79A, CD79alpha), immunoglobulin-associated alpha, a B cell-specific protein that covalently interacts with Ig beta (CD79B) and forms a complex on the surface with Ig M molecules, transduces a signal involved in B-cell differentiation), pI: 4.84, MW: 25028 TM: 2 [P] Gene Chromosome: 19q13.2).
(29) CXCR5 (Burkitt's lymphoma receptor 1, a G protein-coupled receptor that is activated by the CXCL13 chemokine, functions in lymphocyte migration and humoral defense, plays a

10 role in HIV-2 infection and perhaps development of AIDS, lymphoma, myeloma, and leukemia); 372 aa, pI: 8.54 MW: 41959 TM: 7 [P] Gene Chromosome: 11q23.3,

(30) HLA-DOB (Beta subunit of MHC class II molecule (Ia antigen) that binds peptides and 20 presents them to CD4+ T lymphocytes); 273 aa, pI: 6.56, MW: 30820.TM: 1 [P] Gene Chromosome: 6p21.3)

(31) P2X5 (Purinergic receptor P2X ligand-gated ion channel 5, an ion channel gated by extracellular ATP, may be involved in synaptic transmission and neurogenesis, deficiency may contribute to the pathophysiology of idiopathic detrusor instability); 422 aa), pI: 7.63, MW: 47206 TM: 1 [P] Gene Chromosome: 17p13.3).

(32) CD72 (B-cell differentiation antigen CD72, Lyb-2); 359 aa, pI: 8.66, MW: 40225, TM: 1 5 [P] Gene Chromosome: 9p13.3).

(33) LY64 (Lymphocyte antigen 64 (RP105), type I membrane protein of the leucine rich repeat (LRR) family, regulates B-cell activation and apoptosis, loss of function is associated

with increased disease activity in patients with systemic lupus erythematosis); 661 aa, pI: 6.20, MW: 74147 TM: 1 [P] Gene Chromosome: 5q12).

(34) FcRH1 (Fc receptor-like protein 1, a putative receptor for the immunoglobulin Fc domain

that contains C2 type Ig-like and ITAM domains, may have a role in B-lymphocyte 20 differentiation); 429 aa, pI: 5.28, MW: 46925 TM: 1 [P] Gene Chromosome: 1q21-1q22)

(35) IRTA2 (Immunoglobulin superfamily receptor translocation associated 2, a putative immunoreceptor with possible roles in B cell development and lymphomagenesis; deregulation of the gene by translocation occurs in some B cell malignancies); 977 aa, pI: 6.88, MW: 106468, TM: 1 [P] Gene Chromosome: 1q21)

(36) TENB2 (TMEFF2, tomoregulin, TPEF, HPP1, TR, putative transmembrane 35 proteoglycan, related to the EGF/heregulin family of growth factors and follistatin); 374 aa)

(37) PSMA – FOLH1 (Folate hydrolase (prostate-specific membrane antigen) 1)

(38) SST (Somatostatin Receptor; note that there are5 subtypes)

(38.1) SSTR2 (Somatostatin receptor 2)

(38.2) SSTR5 (Somatostatin receptor 5)

(38.3) SSTR1

(38.4) SSTR3

(38.5) SSTR4

AvB6 – Both subunits (39+40)

(39) ITGAV (Integrin, alpha V)

(40) ITGB6 (Integrin, beta 6)

(41) CEACAM5 (Carcinoembryonic antigen-related cell adhesion molecule 5)

(42) MET (met proto-oncogene; hepatocyte growth factor receptor)

(43) MUC1 (Mucin 1, cell surface associated)

(44) CA9 (Carbonic anhydrase IX)

(45) EGFRvIII (Epidermal growth factor receptor (EGFR), transcript variant 3,

(46) CD33 (CD33 molecule)

(47) CD19 (CD19 molecule)

(48) IL2RA (Interleukin 2 receptor, alpha); NCBI Reference Sequence: NM_000417.2);

(49) AXL (AXL receptor tyrosine kinase)

(50) CD30 - TNFRSF8 (Tumor necrosis factor receptor superfamily, member 8)

(51) BCMA (B-cell maturation antigen) - TNFRSF17 (Tumor necrosis factor receptor superfamily, member 17)

(52) CT Ags – CTA (Cancer Testis Antigens)

(53) CD174 (Lewis Y) - FUT3 (fucosyltransferase 3 (galactoside 3(4)-L-fucosyltransferase, Lewis blood group)

(54) CLEC14A (C-type lectin domain family 14, member A; Genbank accession no. NM175060)

(55) GRP78 – HSPA5 (heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)

(56) CD70 (CD70 molecule) L08096

(57) Stem Cell specific antigens. For example:

- 5T4 (see entry (63) below)
- CD25 (see entry (48) above)
- CD32
- LGR5/GPR49
- Prominin/CD133

(58) ASG-5

(59) ENPP3 (Ectonucleotide pyrophosphatase/phosphodiesterase 3)

(60) PRR4 (Proline rich 4 (lacrimal))

(61) GCC – GUCY2C (guanylate cyclase 2C (heat stable enterotoxin receptor)

(62) Liv-1 – SLC39A6 (Solute carrier family 39 (zinc transporter), member 6)

(63) 5T4, Trophoblast glycoprotein, TPBG – TPBG (trophoblast glycoprotein)

(64) CD56 – NCMA1 (Neural cell adhesion molecule 1)

(65) CanAg (Tumor associated antigen CA242)

(66) FOLR1 (Folate Receptor 1)

(67) GPNMB (Glycoprotein (transmembrane) nmb)

(68) TIM-1 – HAVCR1 (Hepatitis A virus cellular receptor 1)

(69) RG-1/Prostate tumor target Mindin – Mindin/RG-1

(70) B7-H4 – VTCN1 (V-set domain containing T cell activation inhibitor 1

(71) PTK7 (PTK7 protein tyrosine kinase 7)

(72) CD37 (CD37 molecule)

(73) CD138 – SDC1 (syndecan 1)

(74) CD74 (CD74 molecule, major histocompatibility complex, class II invariant chain)

(75) Claudins – CLs (Claudins)

(76) EGFR (Epidermal growth factor receptor)

(77) Her3 (ErbB3) – ERBB3 (v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian))

(78) RON - MST1R (macrophage stimulating 1 receptor (c-met-related tyrosine kinase))

(79) EPHA2 (EPH receptor A2)

(80) CD20 – MS4A1 (membrane-spanning 4-domains, subfamily A, member 1)

(81) Tenascin C – TNC (Tenascin C)

(82) FAP (Fibroblast activation protein, alpha)

(83) DKK-1 (Dickkopf 1 homolog (Xenopus laevis)

(84) CD52 (CD52 molecule)

(85) CS1 - SLAMF7 (SLAM family member 7)

(86) Endoglin – ENG (Endoglin)

(87) Annexin A1 – ANXA1 (Annexin A1)

(88) V-CAM (CD106) - VCAM1 (Vascular cell adhesion molecule 1)

An additional tumour-associate antigen and cognate antibodies of interest are:

(**89**) ASCT2 (ASC transporter 2, also known as SLC1A5).

ASCT2 antibodies are described in WO 2018/089393, which is incorporated herein by reference

The cell binding agent may be labelled, for example to aid detection or purification of the agent either prior to incorporation as a conjugate, or as part of the conjugate. The label may be a biotin label. In another embodiment, the cell binding agent may be labelled with a radioisotope.

Methods of Treatment

The compounds of the present invention may be used in a method of therapy. Also provided is a method of treatment, comprising administering to a subject in need of

treatment a therapeutically-effective amount of a conjugate of formula **II**. The term "therapeutically effective amount" is an amount sufficient to show benefit to a patient. Such benefit may be at least amelioration of at least one symptom. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, e.g. decisions on dosage, is within the responsibility of general practitioners and other medical doctors.

A conjugate may be administered alone or in combination with other treatments, either simultaneously or sequentially dependent upon the condition to be treated. Examples of treatments and therapies include, but are not limited to, chemotherapy (the administration of active agents, including, e.g. drugs; surgery; and radiation therapy).

Pharmaceutical compositions according to the present invention, and for use in accordance with the present invention, may comprise, in addition to the active ingredient, i.e. a conjugate of formula **II**, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material will depend on the route of administration, which may be oral, or by injection, e.g. cutaneous, subcutaneous, or intravenous.

Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may comprise a solid carrier or an adjuvant. Liquid pharmaceutical compositions generally comprise a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or polyethylene glycol may be included. A capsule may comprise a solid carrier such as gelatin.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

The Conjugates can be used to treat proliferative disease and autoimmune disease. The term "proliferative disease" pertains to an unwanted or uncontrolled cellular proliferation of

excessive or abnormal cells which is undesired, such as, neoplastic or hyperplastic growth, whether *in vitro* or *in vivo*.

Examples of proliferative conditions include, but are not limited to, benign, pre-malignant, and malignant cellular proliferation, including but not limited to, neoplasms and tumours (e.g., histocytoma, glioma, astrocyoma, osteoma), cancers (e.g. lung cancer, small cell lung cancer, gastrointestinal cancer, bowel cancer, colon cancer, breast carinoma, ovarian carcinoma, prostate cancer, testicular cancer, liver cancer, kidney cancer, bladder cancer, pancreatic cancer, brain cancer, sarcoma, osteosarcoma, Kaposi's sarcoma, melanoma), leukemias, psoriasis, bone diseases, fibroproliferative disorders (e.g. of connective tissues), and atherosclerosis. Other cancers of interest include, but are not limited to, haematological; malignancies such as leukemias and lymphomas, such as non-Hodgkin lymphoma, and subtypes such as DLBCL, marginal zone, mantle zone, and follicular, Hodgkin lymphoma, AML, and other cancers of B or T cell origin.

Examples of autoimmune disease include the following: rheumatoid arthritis, autoimmune demyelinative diseases (e.g., multiple sclerosis, allergic encephalomyelitis), psoriatic arthritis, endocrine ophthalmopathy, uveoretinitis, systemic lupus erythematosus, myasthenia gravis, Graves' disease, glomerulonephritis, autoimmune hepatological disorder, inflammatory bowel disease (e.g., Crohn's disease), anaphylaxis, allergic reaction, Sjögren's syndrome, type I diabetes mellitus, primary biliary cirrhosis, Wegener's granulomatosis, fibromyalgia, polymyositis, dermatomyositis, multiple endocrine failure, Schmidt's syndrome, autoimmune uveitis, Addison's disease, adrenalitis, thyroiditis, Hashimoto's thyroiditis, autoimmune thyroid disease, pernicious anemia, gastric atrophy, chronic hepatitis, lupoid hepatitis, atherosclerosis, subacute cutaneous lupus erythematosus, hypoparathyroidism, Dressler's syndrome, autoimmune thrombocytopenia, idiopathic thrombocytopenic purpura, hemolytic anemia, pemphigus vulgaris, pemphigus, dermatitis herpetiformis, alopecia arcata, pemphigoid, scleroderma, progressive systemic sclerosis, CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, and telangiectasia), male and female autoimmune infertility, ankylosing spondolytis, ulcerative colitis, mixed connective tissue disease, polyarteritis nedosa, systemic necrotizing vasculitis, atopic dermatitis, atopic rhinitis, Goodpasture's syndrome, Chagas' disease, sarcoidosis, rheumatic fever, asthma, recurrent abortion, anti-phospholipid syndrome, farmer's lung, erythema multiforme, post cardiotomy syndrome, Cushing's syndrome, autoimmune chronic active hepatitis, bird-fancier's lung, toxic epidermal necrolysis, Alport's syndrome, alveolitis, allergic alveolitis, fibrosing alveolitis,

interstitial lung disease, erythema nodosum, pyoderma gangrenosum, transfusion reaction, Takayasu's arteritis, polymyalgia rheumatica, temporal arteritis, schistosomiasis, giant cell arteritis, ascariasis, aspergillosis, Sampter's syndrome, eczema, lymphomatoid granulomatosis, Behcet's disease, Caplan's syndrome, Kawasaki's disease, dengue, encephalomyelitis, endocarditis, endomyocardial fibrosis, endophthalmitis, erythema elevatum et diutinum, psoriasis, erythroblastosis fetalis, eosinophilic faciitis, Shulman's syndrome, Felty's syndrome, filariasis, cyclitis, chronic cyclitis, heterochronic cyclitis, Fuch's cyclitis, IgA nephropathy, Henoch-Schonlein purpura, graft versus host disease, transplantation rejection, cardiomyopathy, Eaton-Lambert syndrome, relapsing polychondritis, cryoglobulinemia, Waldenstrom's macroglobulemia, Evan's syndrome, and autoimmune gonadal failure.

In some embodiments, the autoimmune disease is a disorder of B lymphocytes (e.g., systemic lupus erythematosus, Goodpasture's syndrome, rheumatoid arthritis, and type I diabetes), Th1-lymphocytes (e.g., rheumatoid arthritis, multiple sclerosis, psoriasis, Sjögren's syndrome, Hashimoto's thyroiditis, Graves' disease, primary biliary cirrhosis, Wegener's granulomatosis, tuberculosis, or graft versus host disease), or Th2-lymphocytes (e.g., atopic dermatitis, systemic lupus erythematosus, atopic asthma, rhinoconjunctivitis, allergic rhinitis, Omenn's syndrome, systemic sclerosis, or chronic graft versus host disease). Generally, disorders involving dendritic cells involve disorders of Th1-lymphocytes or Th2-lymphocytes. In some embodiments, the autoimmune disorder is a T cell-mediated immunological disorder.

In some embodiments, the amount of the Conjugate administered ranges from about 0.01 to about 10 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.01 to about 5 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.05 to about 5 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.1 to about 5 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.1 to about 4 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.05 to about 3 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.1 to about 3 mg/kg per dose. In some embodiments, the amount of the Conjugate administered ranges from about 0.1 to about 2 mg/kg per dose.

Drug loading

The drug loading (p) is the average number of ABD drugs per cell binding agent, e.g. antibody. Where the compounds of the invention are bound to cysteines, drug loading may range from 1 to 8 drugs (D) per cell binding agent, i.e. where 1, 2, 3, 4, 5, 6, 7, and 8 drug moieties are covalently attached to the cell binding agent. Compositions of conjugates include collections of cell binding agents, e.g. antibodies, conjugated with a range of drugs, from 1 to 8. Where the compounds of the invention are bound to lysines, drug loading may range from 1 to 80 drugs (D) per cell binding agent, although an upper limit of 40, 20, 10 or 8 may be preferred. Compositions of conjugates include collections of cell binding agents, e.g. antibodies, conjugated with a range of drugs, from 1 to 80, 1 to 40, 1 to 20, 1 to 10 or 1 to 8.

The average number of drugs per antibody in preparations of ADC from conjugation reactions may be characterized by conventional means such as UV, reverse phase HPLC, HIC, mass spectroscopy, ELISA assay, and electrophoresis. The quantitative distribution of ADC in terms of p may also be determined. By ELISA, the averaged value of p in a particular preparation of ADC may be determined (Hamblett et al (2004) Clin. Cancer Res. 10:7063-7070; Sanderson et al (2005) Clin. Cancer Res. 11:843-852). However, the distribution of p (drug) values is not discernible by the antibody-antigen binding and detection limitation of ELISA. Also, ELISA assay for detection of antibody-drug conjugates does not determine where the drug moieties are attached to the antibody, such as the heavy chain or light chain fragments, or the particular amino acid residues. In some instances, separation, purification, and characterization of homogeneous ADC where p is a certain value from ADC with other drug loadings may be achieved by means such as reverse phase HPLC or electrophoresis. Such techniques are also applicable to other types of conjugates.

For some antibody-drug conjugates, p may be limited by the number of attachment sites on the antibody. For example, an antibody may have only one or several cysteine thiol groups, or may have only one or several sufficiently reactive thiol groups through which a linker may be attached. Higher drug loading, e.g. p >5, may cause aggregation, insolubility, toxicity, or loss of cellular permeability of certain antibody-drug conjugates.

Typically, fewer than the theoretical maximum of drug moieties are conjugated to an antibody during a conjugation reaction. An antibody may contain, for example, many lysine residues that do not react with the Drug Linker. Only the most reactive lysine groups may react with an amine-reactive linker reagent. Also, only the most reactive cysteine thiol

groups may react with a thiol-reactive linker reagent. Generally, antibodies do not contain many, if any, free and reactive cysteine thiol groups which may be linked to a drug moiety. Most cysteine thiol residues in the antibodies of the compounds exist as disulfide bridges and must be reduced with a reducing agent such as dithiothreitol (DTT) or TCEP, under partial or total reducing conditions. The loading (drug/antibody ratio) of an ADC may be controlled in several different manners, including: (i) limiting the molar excess of Drug Linker relative to antibody, (ii) limiting the conjugation reaction time or temperature, and (iii) partial or limiting reductive conditions for cysteine thiol modification.

Certain antibodies have reducible interchain disulfides, i.e. cysteine bridges. Antibodies may be made reactive for conjugation with linker reagents by treatment with a reducing agent such as DTT (dithiothreitol). Each cysteine bridge will thus form, theoretically, two reactive thiol nucleophiles. Additional nucleophilic groups can be introduced into antibodies through the reaction of lysines with 2-iminothiolane (Traut's reagent) resulting in conversion of an amine into a thiol. Reactive thiol groups may be introduced into the antibody (or fragment thereof) by engineering one, two, three, four, or more cysteine residues (e.g., preparing mutant antibodies comprising one or more non-native cysteine amino acid residues). US 7521541 teaches engineering antibodies by introduction of reactive cysteine amino acids.

Cysteine amino acids may be engineered at reactive sites in an antibody and which do not form intrachain or intermolecular disulfide linkages (Junutula, et al., 2008b Nature Biotech., 26(8):925-932; Dornan et al (2009) Blood 114(13):2721-2729; US 7521541; US 7723485; WO2009/052249). The engineered cysteine thiols may react with linker reagents or the drug-linker reagents of the present invention which have thiol-reactive, electrophilic groups such as maleimide or alpha-halo amides to form ADC with cysteine engineered antibodies and the ABD drug moieties. The location of the drug moiety can thus be designed, controlled, and known. The drug loading can be controlled since the engineered cysteine thiol groups typically react with thiol-reactive linker reagents or drug-linker reagents in high yield. Engineering an IgG antibody to introduce a cysteine amino acid by substitution at a single site on the heavy or light chain gives two new cysteines on the symmetrical antibody. A drug loading near 2 can be achieved with near homogeneity of the conjugation product ADC.

Where more than one nucleophilic or electrophilic group of the antibody reacts with a drug-linker intermediate, or linker reagent followed by drug moiety reagent, then the resulting

product is a mixture of ADC compounds with a distribution of drug moieties attached to an antibody, e.g. 1, 2, 3, etc. Liquid chromatography methods such as polymeric reverse phase (PLRP) and hydrophobic interaction (HIC) may separate compounds in the mixture by drug loading value. Preparations of ADC with a single drug loading value (p) may be isolated, however, these single loading value ADCs may still be heterogeneous mixtures because the drug moieties may be attached, via the linker, at different sites on the antibody.

Thus the antibody-drug conjugate compositions of the invention include mixtures of antibody-drug conjugate compounds where the antibody has one or more ABD drug moieties and where the drug moieties may be attached to the antibody at various amino acid residues.

In one embodiment, the average number of dimer azetidobenzodiazepine groups per cell binding agent is in the range 1 to 20. In some embodiments the range is selected from 1 to 8, 2 to 8, 2 to 6, 2 to 4, and 4 to 8.

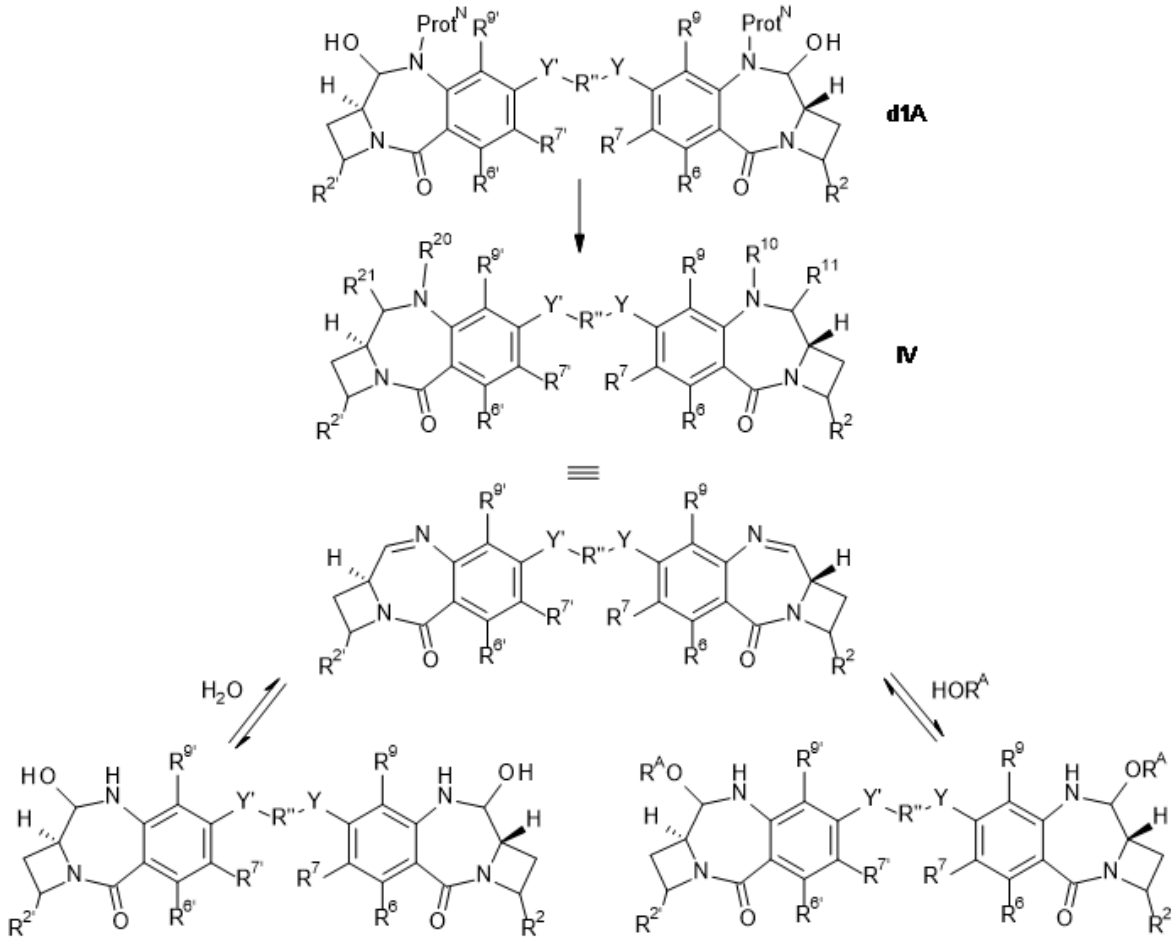
In some embodiments, there is one dimer azetidobenzodiazepine group per cell binding agent.

General synthetic routes

A large number of suitable N-Prot^{N}, O-Prot^{O} and Y-Prot^{Y} protecting groups are described in Greene, T.W. and Wuts, G.M., Protective Groups in Organic Synthesis, 3rd Edition, John Wiley & Sons, Inc., 1999, which is incorporated herein by reference.

Synthesis of Compounds of Formula IV

A possible step in the synthesis of the compounds of the first aspect of the invention, particularly the compound of formula IV, is illustrated in Scheme 1. This starts from a N10 protected ABD dimer (**d1A**).

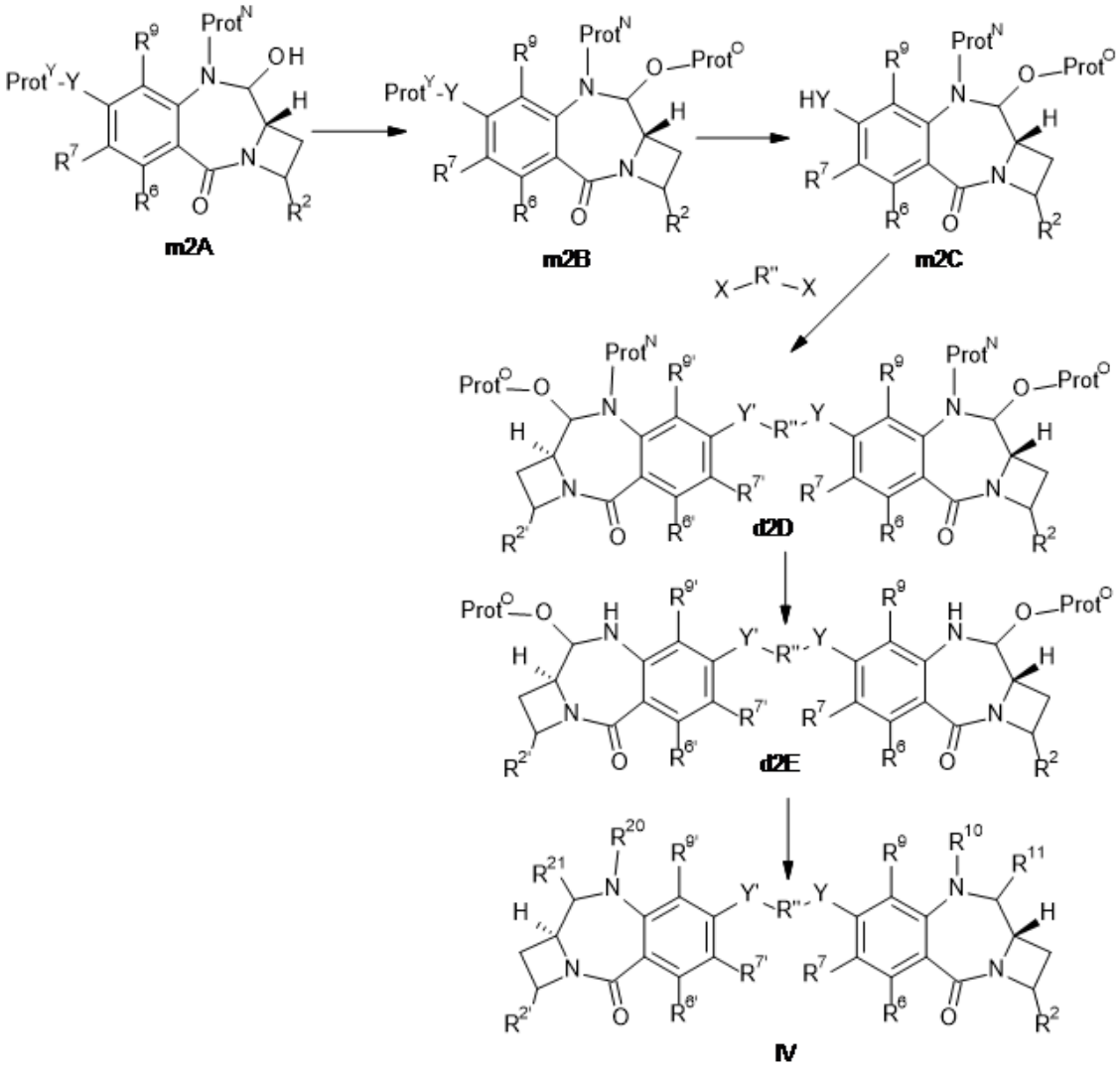


Scheme 1

The dimer **d1A** is deprotected at the N10 position by standard methods to afford the compound of formula IV. In cases where $Prot^N$ is Alloc, deprotection is carried out using palladium. The compound produced may be in its carbinolamine or carbinolamine ether form depending on the solvents used, in equilibrium with an imine.

In the case of ABDs, the ring strain for the four membered azetidine ring means the carbinolamine form is dominant in the equilibrium.

An alternative step in the synthesis of the compounds of formula IV is illustrated in Scheme 2. This starts from the N10 nitrogen protected ABD monomer (**m2A**).



Scheme 2

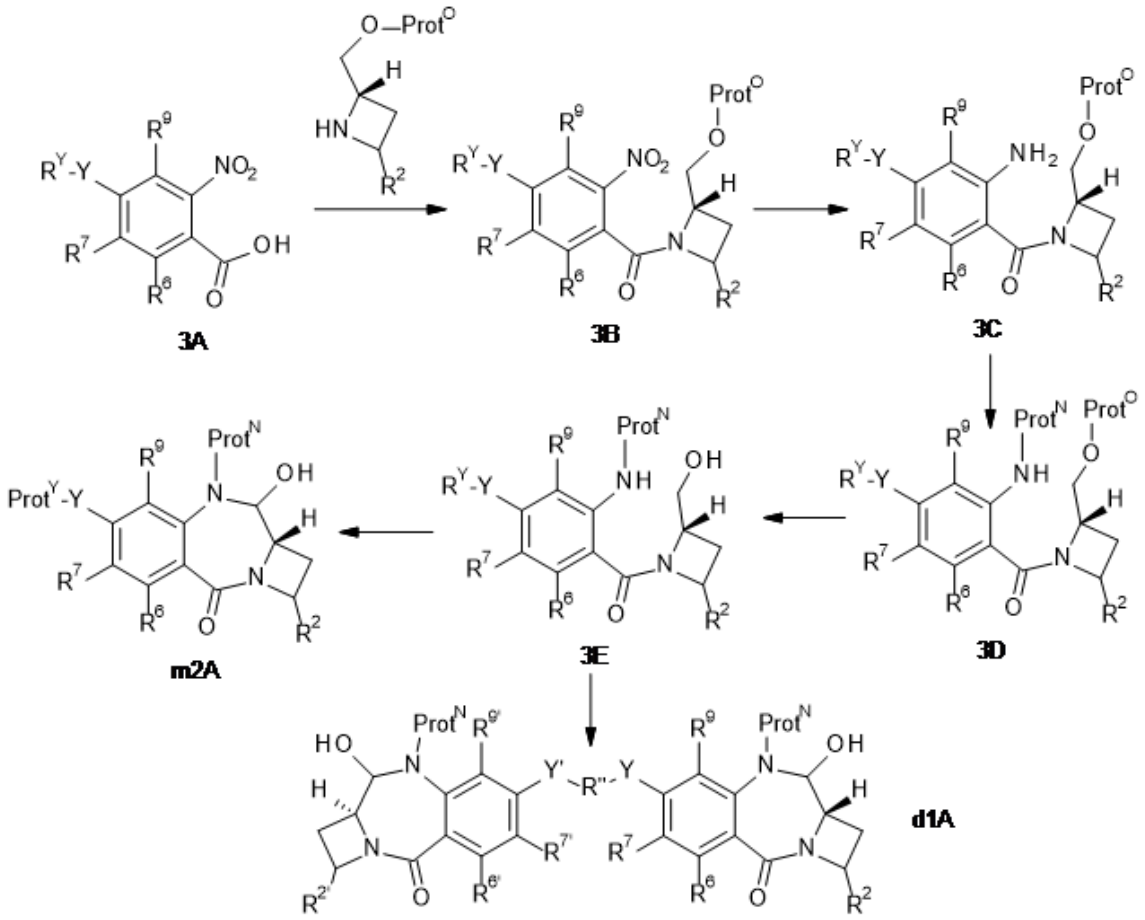
The N10 and Y8 position protected ABD monomer **m2A** is protected at the C11 position alcohol to give **m2B**. Preferably, $Prot^O$ is TBS and the protection is achieved by addition of excess TBS-Cl. Subsequent deprotection of the $Prot^Y$-Y protecting group provides a dimerisable species (**m2C**). When $Prot^Y$ is TIPS, deprotection may be achieved with LiOAc in DMF and water.

m2C is reacted with a dimer linker R''$(X)_2$ to afford the dimer **d2D**. Typically, Y is O and X is a halogen (preferably Br). In this case, a double Williamson ether synthesis forms the dimer, using a TBAI additive.

The N10 protecting group is removed from the dimer product to give **d2E**. For example if $Prot^N$ is Alloc and $Prot^O$ is an oxygen protecting group for synthesis, then the deprotection

is carried using palladium to remove the N10 protecting group, followed by the elimination of the oxygen protecting group for synthesis. If $Prot^N$ is Troc and $Prot^O$ is an oxygen protecting group for synthesis, then the deprotection is carried out using a Cd/Pb couple. If $Prot^N$ is SEM, or an analogous group, and $Prot^O$ is an oxo group, then the oxo group can be removed by reduction, which leads to a protected carbinolamine intermediate, which can then be treated to remove the SEM protecting group followed by the elimination of water. Removal of the C11 position alcohol protecting group provides the compound of formula IV. If $Prot^O$ is TBS, the alcohol deprotection may occur concomitantly with the aforementioned Alloc N-deprotection using palladium and pyrrolidine in DCM.

The dimer **d1A** and monomer **m2A** required for Schemes 1 and 2 respectively, may be synthesised by several routes. One possible route, via oxidative ring closure, is illustrated in Scheme 3.



Scheme 3

The compounds **3A**, **3B**, **3C**, **3D** and **3E** may be dimeric (where the group R^Y represents R'' connected to a similar ABD precursor) or monomeric (where the group R^Y represents a suitable protecting group).

Monomeric **3A** is a nitrobenzoic acid derivative. Many such derivatives are commercially available and others can be synthesised by conventional methods (e.g. Althuis, T. H. and Hess, H. J., J Medicinal Chem., **20(1)**, 146-266 (1977)). Often the nitrobenzoic acid is derived from the ester, by ester hydrolysis under mild conditions (such as with LiOH). Dimeric **3A** may be made by various strategies disclosed in the prior art (e.g. Scheme 3 of WO 00/12508). For example, appropriate benzoic acid esters may be dimerised about a suitable diol by Mitsunobo etherification, followed by nitration and hydrolysis. Alternatively, benzoic acid esters may be dimerised about a suitable dihalide by Williamson ether synthesis. Further transformations required to afford monomeric and dimeric **3A** are available in the literature.

The azetidine starting material may be synthesised via modification of comparable proline syntheses disclosed in the prior art (e.g. Scheme 4 of WO 00/12508). Strategies pertaining specifically to azetidine are also known in the literature (e.g. Bose, D.S., et al., Tetrahedron Letters, 38(33), 5839-5842, 1997; doi: 10.1016/S0040-4039(97)01297-5). For example, commercially available azetidine-2-carboxylic acid, may be protected at the azetidine nitrogen by a suitable protecting group, such as Cbz, before acidic esterification to achieve the methyl ester. The ester may be reduced with $LiBH_4$ in THF to yield Cbz protected 2-(hydroxymethyl)azetidine. In some approaches, an appropriate protecting group ($Prot^O$), such as TBS, may be added to the alcohol by reaction with TBS-Cl. In other approaches, the alcohol is left unprotected. In scheme 3, $Prot^O$ may represent either a suitable protecting group or H - a suitable $Prot^O$ group should be able to withstand NO_2 reduction conditions. Next, the nitrogen protecting group is removed, typically by reduction under H_2 gas, to yield the azetidine starting material required in Scheme 3.

The compound **3A** is condensed with the azetidine starting material to afford **3B**. Often the condensation is achieved via a DCC coupling or via an acid chloride (formed from the carboxylic acid with oxalyl chloride or $SOCl_2$), or with HOBt in DCM at low temperature.

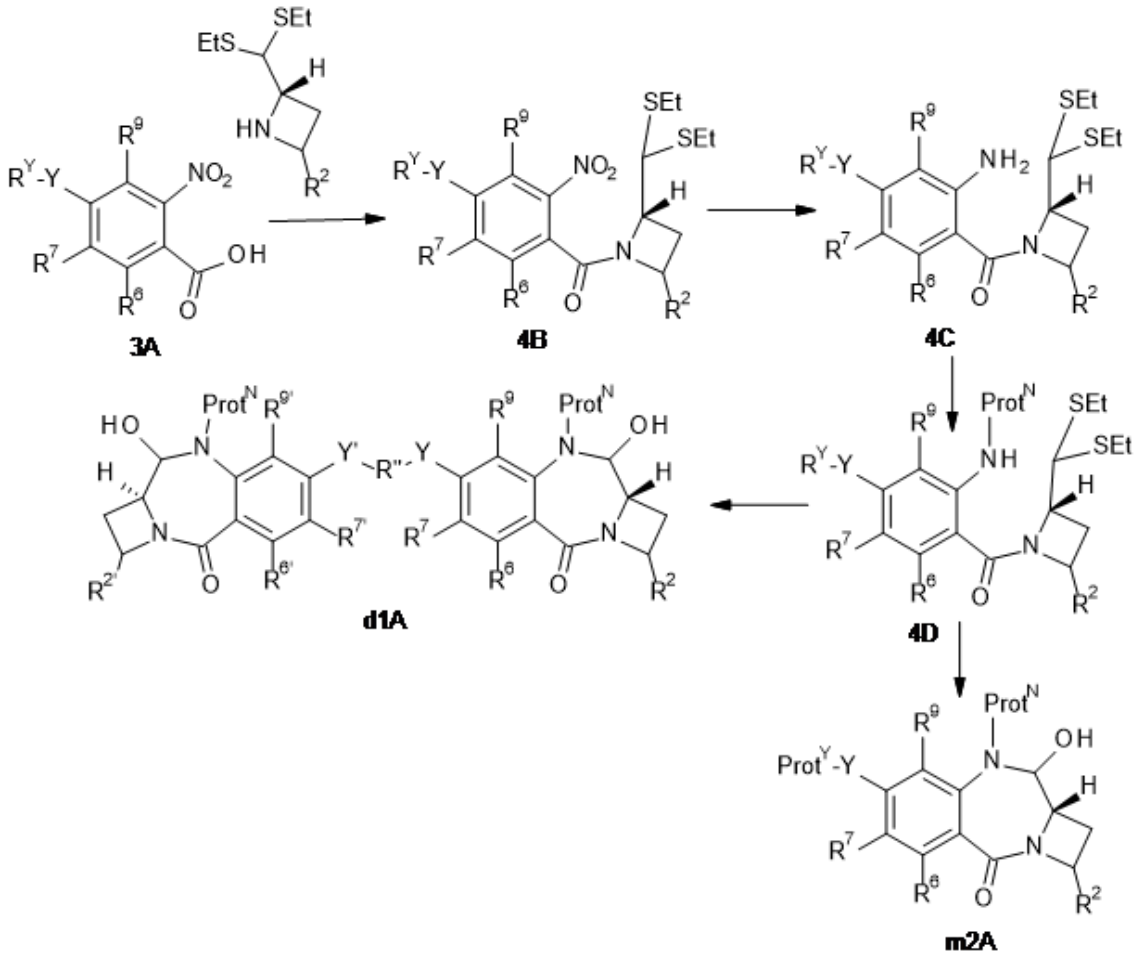
The nitro group of **3B** is reduced to the amine (**3C**), using standard procedures such as $SnCl_2$ in MeOH, or Zinc in MeOH/H_2O/formic acid (90:5:5), or sodium dithionite, or Raney

Nickel and hydrazine, or catalytic hydrogenation over palladium on charcoal. The method selected depends on the requirements of the hydroxyl protecting group.

The resulting amine is singly protected by a suitable protecting group to afford **3D**. The N-$Prot^N$ group is preferably a carbamate, such as N-Alloc. The nucleophilicity of the amine is reduced upon protection with Alloc, so singular protection is favoured. Typically, this is achieved by reaction with pyridine and one equivalent of allyl chloroformate. When $Prot^O$ is H, then **3D** is equivalent to **3E**. When $Prot^O$ is a protecting group it is removed to give alcohol **3E** under standard conditions. If $Prot^O$ is an acetate protecting group it may be removed under mild basic conditions (e.g. K_2CO_3), or if $Prot^O$ is a silyl ether protecting group, such as TBS, it may be removed by using TBAF or mild acid.

Oxidative ring closure of via the aldehyde, or a functional equivalent, from dimeric **3E** affords **d1A** (for further reaction as per Scheme 1) and from monomeric **3E** afford **m2A** (for further reaction as per Scheme 2). The selective alcohol-aldehyde oxidation may be achieved by exposure to tetrapropylammoniumperruthenate (TPAP) in N-methylmorpholine N-oxide (NMO) over molecular sieves, or by Swern oxidation (with DMSO and oxalyl chloride), or by Dess-Martin oxidation (with DMP) or preferably by Cu(I)/TEMPO radical oxidation (with Tetrakisacetonitrile copper(I) triflate, 1-hydroxy-2,2,6,6-tetramethyl-piperidine (TEMPO), 1-methylimidazole and 2-(2-pyridyl)pyridine). The latter is favoured as it does not require rigorous anhydrous conditions and there is no evidence of over oxidation to the ABD dilactam species. The aldehyde species undergoes spontaneous B-ring closure involving attack thereon by the singly protected N10 position.

An alternative route to the dimer **d1A** and monomer **m2A** is illustrated in Scheme 4. This route uses aldehyde unmasking to mediate ring closure.



Scheme 4

The compounds **3A**, **4B**, **4C** and **4D** may be dimeric (where the group R^Y represents R'' connected to a similar ABD precursor) or monomeric (where the group R^Y represents a suitable protecting group).

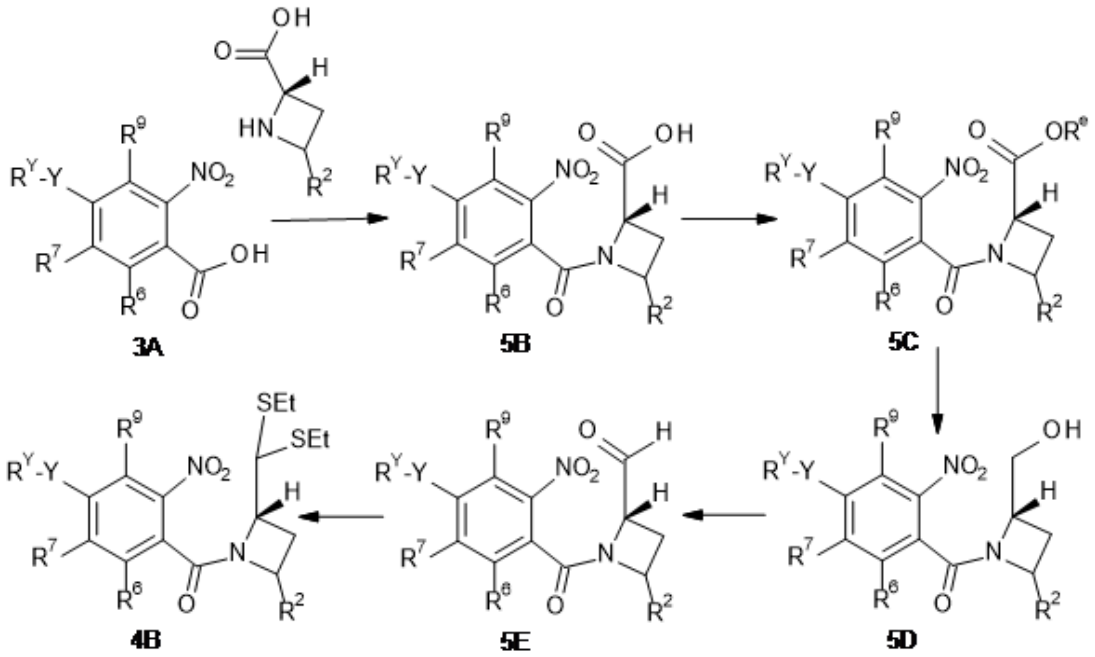
3A monomer and dimer variants may be generated by the strategies discussed above, in relation to Scheme 3. The azetidine starting material features a thioacetal in the 2-position (although other masked aldehyde equivalents may be used). The diethyl thioacetal azetidine may be prepared by the modification of similar proline synthetic strategies (e.g. Langley, D.R. & Thurston, D. E., J Organic Chemistry, **52**, 91-97 (1987)). Routes pertaining specifically to azetidine are also known in the literature (e.g. Bose, D.S., et al., Tetrahedron Letters, 38(33), 5839-5842, 1997; doi: 10.1016/S0040-4039(97)01297-5). For example, Cbz protected 2-(hydroxymethyl)azetidine may be prepared as described above (for Scheme 3). The alcohol is then typically reoxidized to the aldehyde by Dess-Martin oxidation (with DMP) or IBX in DMSO. The resulting aldehyde is preferably condensed with a thiol such as EtSH, with a mild acid catalyst, such as TMSCl in protic solvent, to

achieve the thioacetal. The thioacetal is incompatible with H_2 gas reduction, so the N-protecting group (e.g. Cbz) is often removed with TMS-I in DCM. This results in the diethyl thioacetal azetidine starting material.

Direct condensation of **3A** with the thioacetal azetidine starting material affords **4B**. The nitro group of **4B** may be reduced to the amine (**4C**) via the methods discussed above in relation to Scheme 3, preferably by the Tin(II)chloride method ($SnCl_2$ in MeOH) or Zinc in MeOH/H_2O/formic acid (90:5:5). Reduction is preferably not via direct hydrogenation due to incompatibility of the thioacetal group. The amine is singly protected by a suitable amine protecting group, such as Alloc, by reaction with the corresponding chloroformate or acid chloride. The N-Prot^N group of **4D** is preferably a carbamate, such as N-Alloc, as these species favour single protection.

Selective unmasking of the thioacetal to the aldehyde results in spontaneous cyclisation of the B-ring, by attack thereon by the singly protected N10 position. Typically, un-masking is mediated by Mercury(II), for example $HgCl_2$ with $CaCO_3$ in acetonitrile:water. For dimeric **4D** this provides **d1A** (for further reaction as per Scheme 1) and for monomeric **4D** this provides **m2A** (for further reaction as per Scheme 2).

Dimeric or monomeric thioacetal **4B** (as per Scheme 4) may be synthesised via an alternative route, illustrated in Scheme 5. This route generates the thioacetal in situ.



Scheme 5

Synthetic strategies to achieve monomeric and dimeric **3A** are discussed above with relation to Scheme 3. **3A** is condensed with commercially available azetidine-2-carboxylic acid to afford **5B**. The route from **5B** to **4B** follows a similar approach as the synthesis of the thioacetal azetidine starting material from azetidine-2-carboxylic (as discussed with relation to Scheme 4).

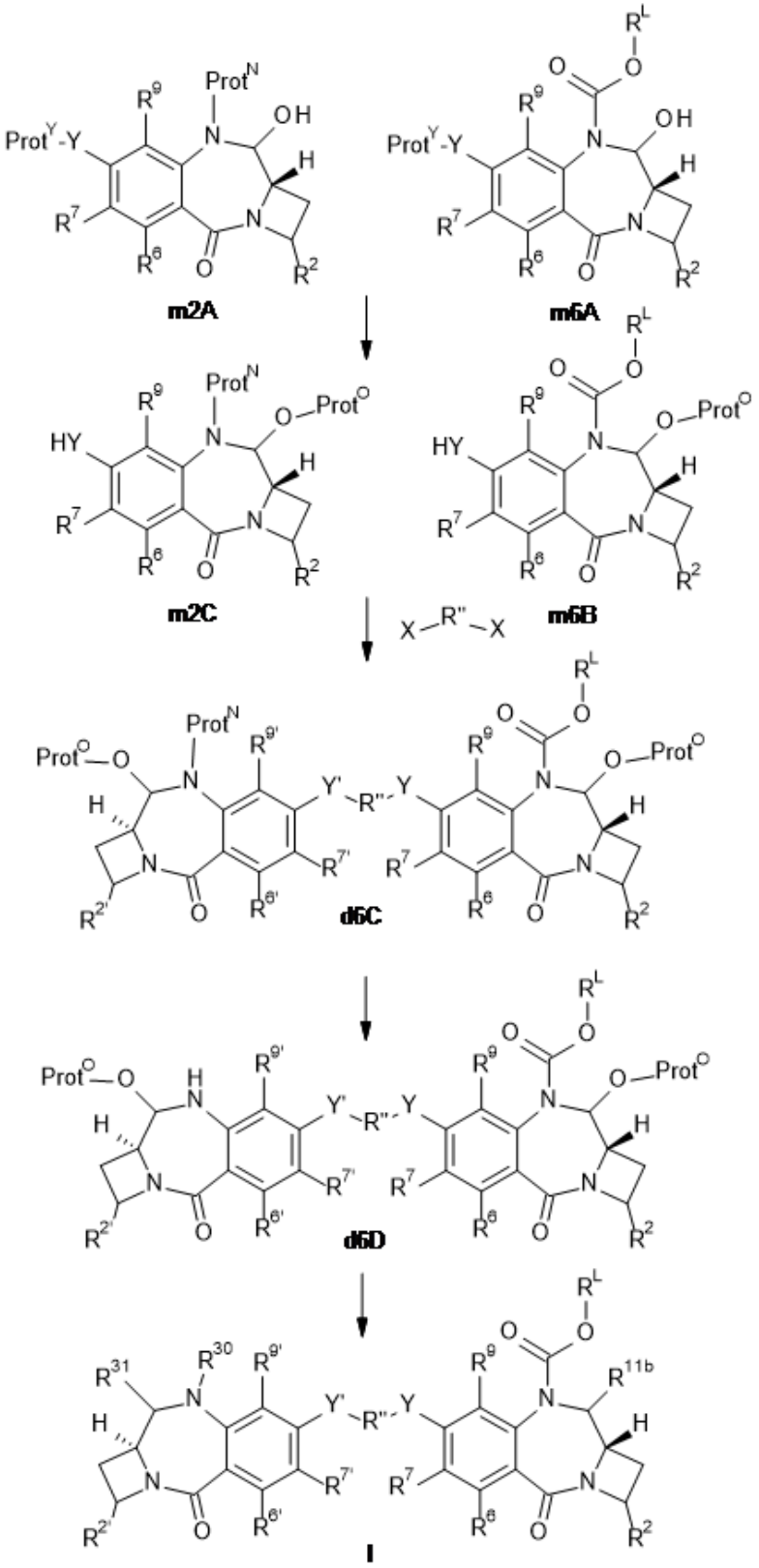
5C is reduced by a hydride reducing agent, typically by $LiBH_4$, to the secondary alcohol **5D**. **5D** is then reoxidised to the aldehyde (**5E**), often by a hypervalent iodine species (e.g. IBX or DMP). The thioacetal is generated in situ, preferably using EtSH under acidic conditions, to provide the compound **4B**. This can be further reacted as per Scheme 4 to reach the desired ABD species.

Synthesis of Compounds of Formula I

A possible step in the synthesis of the compounds of the second aspect of the invention, particularly the compound of formula I, is illustrated in Scheme 6. This starts with two cyclised ABD monomers: **m2A** with a $Prot^N$ protected N10 position and **m6A** with a R^L appended N10-nitrogen.

The compound of formula I may exist in equilibrium between the imine and the carbinolamine or carbinolamine ether form, depending on the solvent used (analogous to the equilibrium illustrated in Scheme 1 for Formula IV).

In some embodiments, $Prot^N$ may be equivalent to the R^{30} substituent of formula I (described by option (d) i, ii and iii of the second aspect of the invention).



Scheme 6

The **m2A** starting material may be prepared as per Schemes 3, 4, and 5. For compounds where $Prot^N$ is equivalent to the R^{30} carbamate linker groups, then **m2A** is prepared via the isocyanate (i.e. the same route as **m6A** - discussed below).

m2A and **m6A** may be dimerised about the Y8 position with a dimer linker R'', using a similar strategy as described in relation to Scheme 2. The C11 position alcohol is protected by $Prot^O$, where $Prot^O$ is preferably TBS and is introduced by reaction with TBS-Cl. The subsequent removal of the $Prot^Y$ group, where $Prot^Y$ is TIPS, may occur with LiOAc in DMF and water, to afford **m2C** and **m6B** respectively.

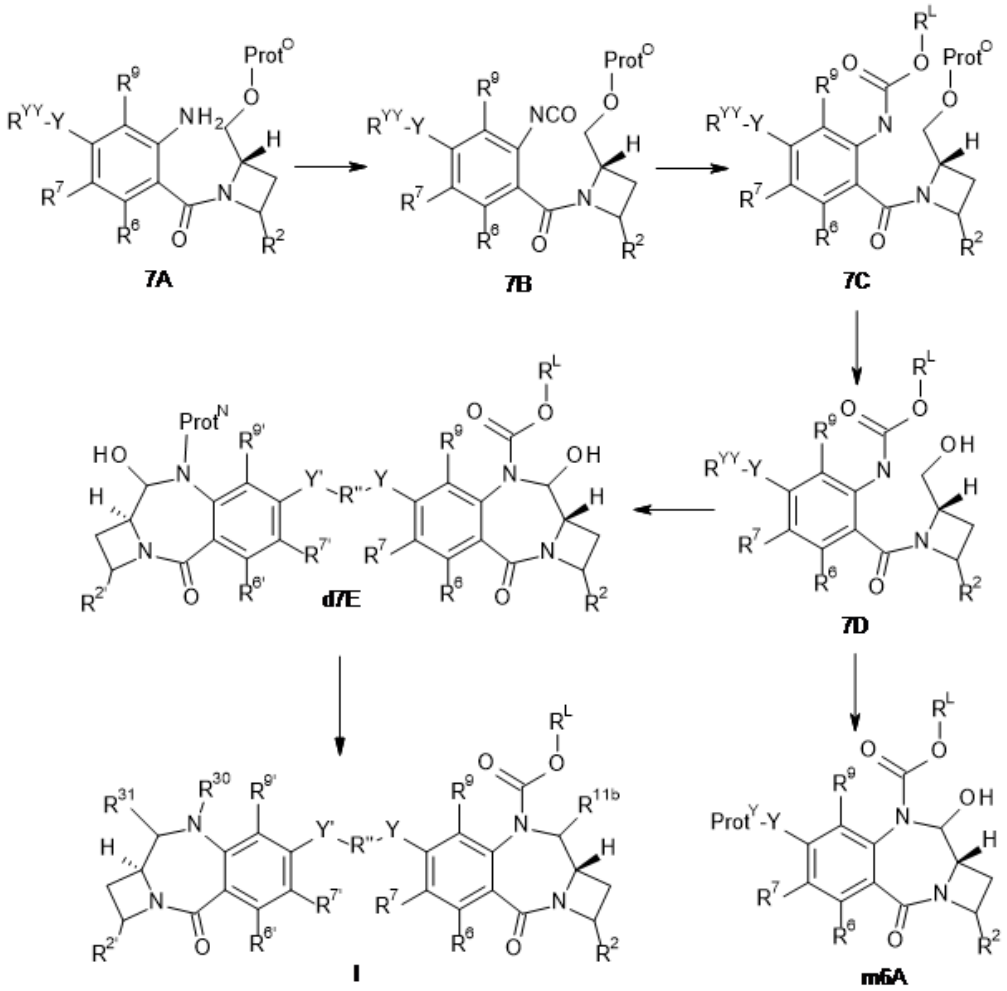
m2C and **m6B** are further reacted with $R''(X)_2$ to afford the dimer **d6C**. Typically, Y is O and X is a halogen (preferably Br). In this case, TBAI additive may drive a double Williamson ether synthesis to form the dimer. Alternative strategies for dimerisation are also known in the art, for example, via Mitsonobu etherification.

In some embodiments, the N10 protecting group is removed from the non-linker ABD to afford the asymmetric dimer **d6D**. Various deprotection strategies are discussed in relation to Scheme 1 and 2. In cases where $Prot^N$ is Alloc, then the deprotection may be carried out with palladium.

In other embodiments, the N10 protecting group is not removed. **d6C** is transformed directly to a compound of formula I by removal of $Prot^O$.

Removal of the C11 position alcohol protecting groups provides the asymmetrical compound of formula I. If $Prot^O$ is TBS, the alcohol deprotection may occur concomitantly with the Alloc N10 position deprotection with palladium and pyrrolidine in DCM.

A possible synthesis of monomer **m6A** (required for Scheme 6) and an alternative route to the compounds of formula I are illustrated in Scheme 7.



Scheme 7

The compounds **7A**, **7B**, **7C** and **7D** may be dimeric (where the group R^{YY} represents R'' connected to an ABD precursor singly protected at the N10 position) or monomeric (where the group R^{YY} represents a suitable protecting group). Monomeric **7A** is equivalent to monomeric **3C** and can be synthesised via a similar route (as illustrated in Scheme 3).

The amine of **7A** is transformed to isocyanate **7B**. Strategies for isocyanate formation are detailed in the prior art (for example, WO 2005/023814). Typically, phosgene is used, preferably triphosgene under basic conditions; the solid triphosgene crystals are safer and easier to handle than toxic phosgene gas. The reaction should be carried out in an anhydrous and non-hydroxylic organic solvent, which is preferably non-polar. Suitable solvents include anhydrous DCM and anhydrous toluene. The reaction may be carried out at room temperature, and is conveniently monitored by infrared spectroscopy at about 2265 cm^{-1}.

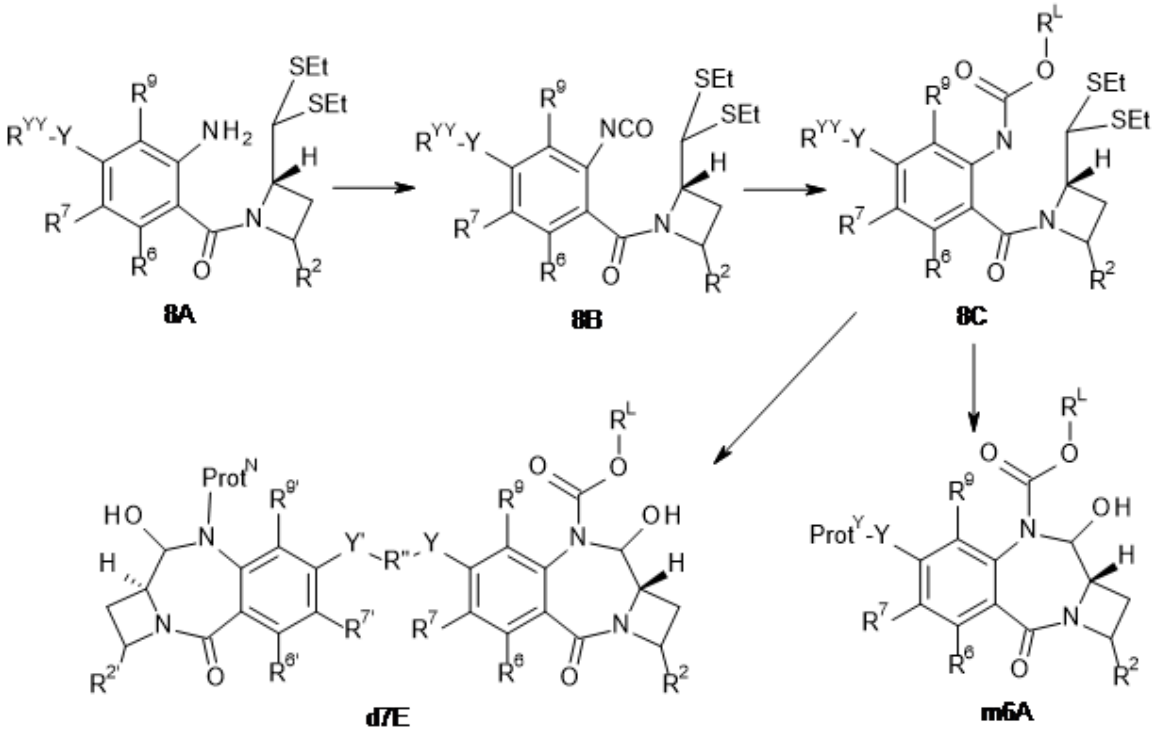
The carbamate **7C** is formed from the isocyanate via attack thereon by R^L-OH. The carbamate formation is often achieved via a one-pot-method, where the isocyanate is formed by triphosgene with TEA in DCM, and R^L-OH is added directly to the reaction mixture. This approach reduces the residence time of the isocyanate prior to carbamate formation, which lessens the chance of side reactions.

The $Prot^O$ protecting group is removed by a suitable method to provide the secondary alcohol **7D**, typically under acidic conditions (e.g. Acetic acid in THF:Water solvent).

Oxidative ring closure of **7D** via the aldehyde, or a functional equivalent, may be achieved by exposure to tetrapropylammoniumperruthenate (TPAP) in N-methylmorpholine N-oxide (NMO) over molecular sieves, or by Swern oxidation (DMSO and oxalyl chloride), or preferably by Cu(I)/TEMPO radical oxidation (Tetrakisacetonitrile copper(I) triflate, 1-hydroxy-2,2,6,6-tetramethyl-piperidine (TEMPO), 1-methylimidazole and 2-(2-pyridyl)pyridine). This affords **m6A** (for further reaction in Scheme 6) for the monomeric **7D** variant (i.e. where $R^{YY} = Prot^Y$), or affords **d7E** for the dimeric **7D** variant (i.e. where R^{YY} = R'' connected to a N10 protected ABD).

Removal of the nitrogen protecting group ($Prot^N$), typically Alloc removed with palladium, provides the asymmetrical compound of formula I.

An alternative route to the monomer **m6A** (Scheme 6) and the dimer **d7E** (Scheme 7) is illustrated in Scheme 8.



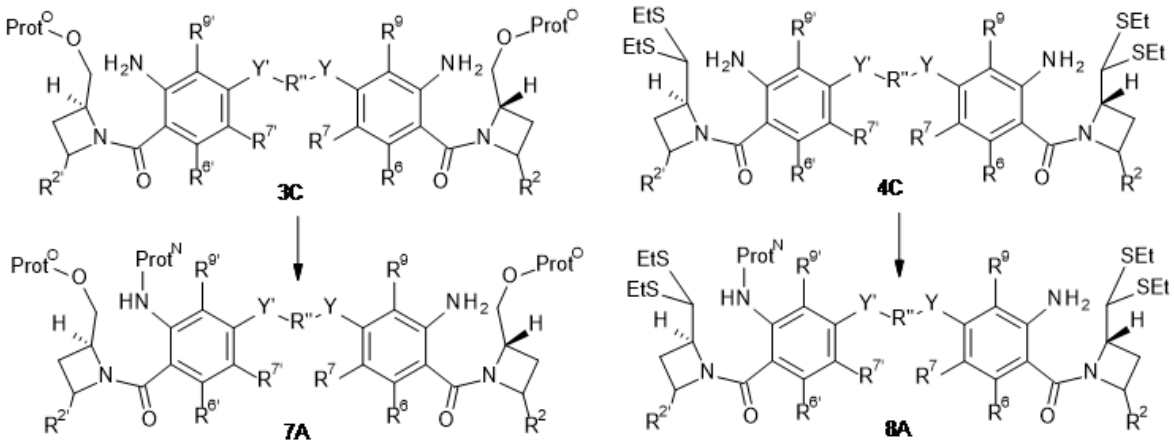
Scheme 8

The compounds **8A**, **8B** and **8C** may be dimeric (where the group R^{YY} represents R'' connected to an ABD precursor singly protected at the N10 position) or monomeric (where the group R^{YY} represents a suitable protecting group, $Prot^{Y}$). **8A** is transformed to isocyanate **8B**, which in turn is reacted with R^{L}-OH to append R^{L} by a carbamate. The preferred strategy is similar to that discussed in relation to Scheme 7.

Selective unmasking of the thioacetal **8C** to the aldehyde results in spontaneous cyclisation of the B-ring, by attack thereon by the singly protected N10 position. Typically, un-masking is mediated by Mercury(II), for example $HgCl_2$ with $CaCO_3$ in acetonitrile:water. The cyclisation provides monomeric **m6A** (from monomeric **8C**) and dimeric **d7E** (from dimeric **8C**). **m6A** may be reacted on as per Scheme 6, and **d7E** as per Scheme 7 to produce compounds of formula I.

The starting materials of Scheme 7 may be realised via a similar route to Scheme 3 (monomeric **3A** to **3C**) for monomeric **7A** (i.e. where R^{YY} is $Prot^{Y}$). Similarly, the starting materials of Scheme 8 may be realised via a similar route to Scheme 4 (monomeric **3A** to **4C**) for monomeric **8A** (i.e. where R^{YY} is $Prot^{Y}$).

A route to dimeric **7A** and **8A** (i.e. where R^{YY} is R'' connected to an ABD precursor singly protected at the N10 position) is illustrated in Scheme 9.



Scheme 9

The dimeric variants of **3C** and **4C** are generated via the strategies discussed in Schemes 3, 4 and 5. **3C** and **4C** are protected at only one N10 position to yield the asymmetric ABD dimers **7A** and **8A** respectively. This is achieved by the addition of one equivalent of protecting reagent, typically allyl chloroformate when $Prot^N$ is Alloc, and subsequent purification to remove un-protected or doubly protected products.

Synthesis of Conjugates of Formula II

A possible step in the synthesis of the conjugates of the third aspect of the invention, particularly the constituent Drug Linker unit (D^L) of formula II, involves connection of the linker to a Ligand Unit, thereby converting the group R^L (as per compounds of formula I) to the group R^{LL} (as per compounds of formula I').

Conjugates can be prepared as previously described. Antibodies can be conjugated to the Drug Linker compound as described in Doronina et al., Nature Biotechnology, 2003, 21, 778-784). Briefly, antibodies (4-5 mg/mL) in PBS containing 50 mM sodium borate at pH 7.4 are reduced with tris(carboxyethyl)phosphine hydrochloride (TCEP) at 37 ºC. The progress of the reaction, which reduces interchain disulfides, is monitored by reaction with 5,5'-dithiobis(2-nitrobenzoic acid) and allowed to proceed until the desired level of thiols/mAb is achieved. The reduced antibody is then cooled to 0°C and alkylated with 1.5 equivalents of maleimide drug-linker per antibody thiol. After 1 hour, the reaction is quenched by the addition of 5 equivalents of N-acetyl cysteine. Quenched drug-linker is removed by gel filtration over a PD-10 column. The ADC is then sterile-filtered through a 0.22 µm syringe filter. Protein concentration can be determined by spectral analysis at 280 nm and 329 nm, respectively, with correction for the contribution of drug absorbance at 280

nm. Size exclusion chromatography can be used to determine the extent of antibody aggregation, and RP-HPLC can be used to determine the levels of remaining NAC-quenched drug-linker.

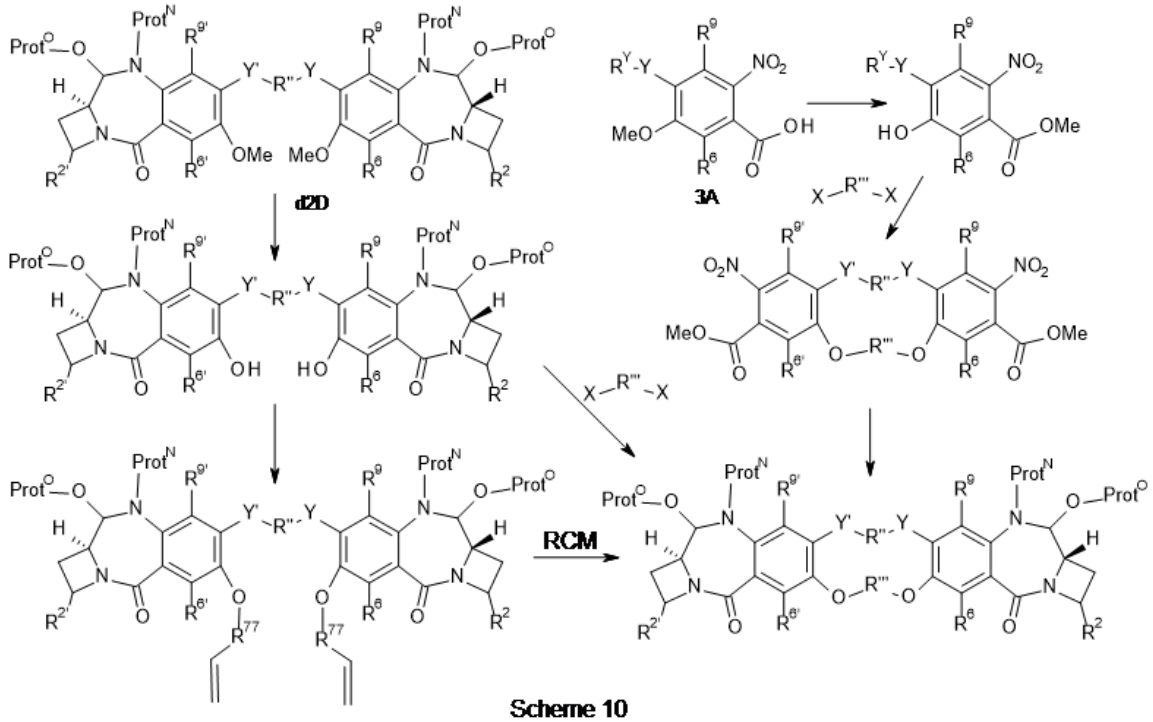
Synthesis of macrocyclic embodiments

In some embodiments of the first, second and third aspects of the invention, the R^7 and $R^{7'}$ substituents may together form a group which is: (i) -O-$(CH_2)_n$-O- where n is from 7 to 16, or (ii) -O-$(CH_2CH_2O)_m$- where m is from 2 to 5, to give a macrocyclic ABD dimer.

Various strategies may be employed to introduce a R^7–$R^{7'}$ linker, as illustrated in Scheme 10 below. Starting at **d2D,** where R^7 and $R^{7'}$ both represent -OR, the R group may be removed by addition of BBr_3 in DCM to reveal the alcohol. A substitution reaction of a dibromoalkane in base, such as 1,7-dribromoheptane with K_2CO_3, affords the macrocyclic product by attack thereon by both C7 position alcohols.

An alternative route starting at **d2D** involves substituting the C7 position alcohol with an n-bromoalk-1-ene. This provides two terminally unsaturated alkyenyl chains, which can readily undergo ring-closing metathesis (RCM). For example, substitution may be achieved with 5-bromopent-1-ene and RCM with Grubs-II catalyst. Macrocyclisation via RCM is generally high yielding.

A preferred route to the macrocycle starts at dimeric **3A** or an ester precursor, where R^7 and $R^{7'}$ both represent –OR. Removal of the R group and substitution by a dibromoalkane (using similar conditions as above) provides a macrocyclic compound. The ABD then may then be reached as per Schemes 3 or 4.



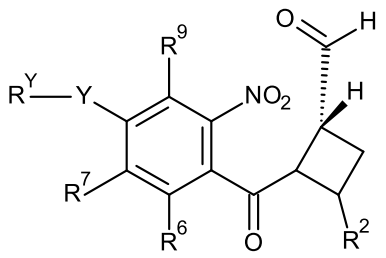
Scheme 10

The resulting product may be reacted on via Scheme 2 or Scheme 6 to achieve the compounds of formula IV and I respectively.

Further transformation details required to afford such macrocyclic products are available in the literature (Donnell, A.F., Zhang, Y., Stang, E.M., Wei, D.D., Tebben, A.J., Perez, H.L., Schroeder, G.M., Pan, C., Rao, C., Borzilleri, R.M., Vite, G.D., Gangwar, S., Macrocyclic pyrrolobenzodiazepine dimers as antibody-drug conjugate payloads, Bioorganic & Medicinal Chemistry Letters (2017), doi: https://doi.org/10.1016/j.bmcl.2017.10.028 and WO 2016/209951).

Synthesis of secondary amine embodiments

Compounds where the N10-C11 group is -NH-CH_2- (i.e. secondary amines) may be synthesised by a modification of the above procedures. In particular, reductive amination of the compound 3B* can yield a modified version of **m2A** or **d1A** for use in the further steps:



3B*

Compounds 3B* can be synthesised from a precursor alcohol by oxidation, the precursor alcohol being reachable by analogous steps to those used to synthesise **3B**.

Synthesis of Drug Conjugates

Conjugates can be prepared as previously described. Antibodies can be conjugated to the Drug Linker compound as described in Doronina et al., Nature Biotechnology, 2003, 21, 778-784). Briefly, antibodies (4-5 mg/mL) in PBS containing 50 mM sodium borate at pH 7.4 are reduced with tris(carboxyethyl)phosphine hydrochloride (TCEP) at 37 ºC. The progress of the reaction, which reduces interchain disulfides, is monitored by reaction with 5,5'-dithiobis(2-nitrobenzoic acid) and allowed to proceed until the desired level of thiols/mAb is achieved. The reduced antibody is then cooled to 0°C and alkylated with 1.5 equivalents of maleimide drug-linker per antibody thiol. After 1 hour, the reaction is quenched by the addition of 5 equivalents of N-acetyl cysteine. Quenched drug-linker is removed by gel filtration over a PD-10 column. The ADC is then sterile-filtered through a 0.22 µm syringe filter. Protein concentration can be determined by spectral analysis at 280 nm and 329 nm, respectively, with correction for the contribution of drug absorbance at 280 nm. Size exclusion chromatography can be used to determine the extent of antibody aggregation, and RP-HPLC can be used to determine the levels of remaining NAC-quenched drug-linker.

Brief Description of the Figures

Figure 1 shows the effect on the growth of a tumour cell line when treated by a control or by a conjugate of the present invention.

Further Preferences

The following preferences may apply to all aspects of the invention as described above, or may relate to a single aspect. The preferences may be combined together in any combination.

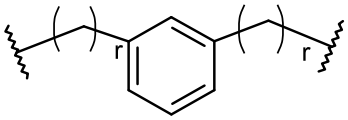
$R^{6'}$ and $R^{9'}$ are selected from the same groups as R^6 and R^9 respectively. In some embodiments, $R^{6'}$, $R^{7'}$, $R^{9'}$, and Y' are the same as R^6, R^7, R^9, and Y respectively.

Dimer link

In some embodiments, Y and Y' are both O.

In some embodiments, R'' is a C_{3-7} alkylene group with no substituents. In some of these embodiments, R'' is a C_3, C_5 or C_7 alkylene. In particular, R'' may be a C_3 or C_5 alkylene.

In other embodiments, R" is a group of formula:



where r is 1 or 2.

The phenylene group may be replaced by a pyridylene group.

R^6 to R^9

In some embodiments, R^9 is H.

In some embodiments, R^6 is selected from H, OH, OR, SH, NH_2, nitro and halo, and may be selected from H or halo. In some of these embodiments R^6 is H.

In some embodiments, R^7 is selected from H, OH, OR, SH, SR, NH_2, NHR, NRR', and halo. In some of these embodiments R^7 is selected from H, OH and OR, where R is selected from optionally substituted C_{1-7} alkyl, C_{3-10} heterocyclyl and C_{5-10} aryl groups. R may be more preferably a C_{1-4} alkyl group, which may or may not be substituted. A substituent of interest is a C_{5-6} aryl group (e.g. phenyl). Particularly preferred substituents at the 7- positions are OMe and OCH_2Ph. Other substituents of particular interest are dimethylamino (i.e. $–NMe_2$); $-(OC_2H_4)_qOMe$, where q is from 0 to 2; nitrogen-containing C_6 heterocyclyls, including morpholino, piperidinyl and N-methyl-piperazinyl.

These embodiments and preferences apply to $R^{9'}$, $R^{6'}$ and $R^{7'}$ respectively.

In other embodiments, R^7 and $R^{7'}$ together form a group which is $-O-(CH_2)_n-O-$, where n is from 7 to 16. n may be at least 7, 8, 9, 10 or 11. N may be at most 16, 15, 14 or 13.

In other embodiments, R^7 and $R^{7'}$ together form a group which is -O-$(CH_2CH_2O)_m$-, where m is 2 to 5. m may be at least 2, 3 or 4. m may be at most 5, 4 or 3.

R^{10}, R^{11}, R^{20}, R^{21} (Formula IV)

In some embodiments, R^{10} and R^{11} together form a double bond between the N and C atoms to which they are bound. In some of these embodiments, R^{20} and R^{21} together form a double bond between the N and C atoms to which they are bound. In other of these embodiments, R^{20} and R^{21} are both H.

In some embodiments, R^{10} is H and R^{11} is selected from OH and OR^A, where R^A is C_{1-4} alkyl. In some of these embodiments, R^{20} is H and R^{21} is selected from OH and OR^B, where R^B is C_{1-4} alkyl. In other of these embodiments, R^{20} and R^{21} are both H.

In some embodiments, R^{10} and R^{11} are both H. In some of these embodiments, R^{20} and R^{21} together form a double bond between the N and C atoms to which they are bound. In other of these embodiments, R^{20} is H and R^{21} is selected from OH and OR^B, where R^B is C_{1-4} alkyl.

In some embodiments, R^A is methyl. In some embodiments, R^B is methyl.
In some embodiments, only one of the pairs of R^{10} and R^{11} and R^{20} and R^{21} are both H. In other embodiments, neither of the pairs of R^{10} and R^{11} and R^{20} and R^{21} are both H.

In some embodiments, R^{10}, R^{11}, R^{20} and R^{21} are all H.

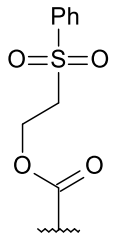
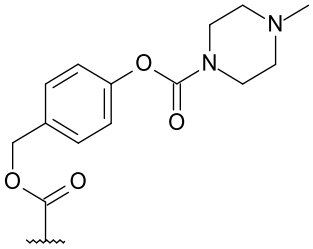
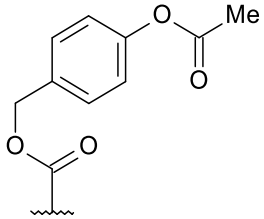
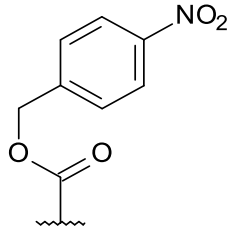
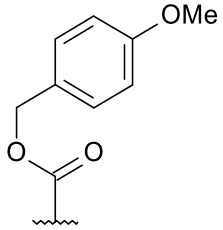
N10'-C11' (Formulae I and I)*

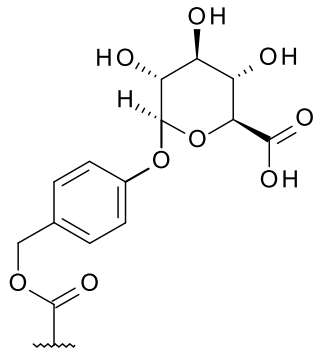
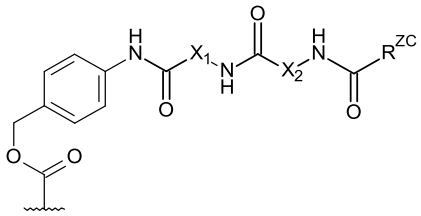
In some embodiments, R^{30} and R^{31} together form a double bond between the N and C atoms to which they are bound.

In some embodiments, R^{30} is H and R^{31} is selected from OH and OR^B, where R^B is C_{1-4} alkyl. In some of these embodiments, R^B is methyl.

In some embodiments, R^{30} is H and R^{31} is H.

In some embodiments, R^{31} is OH or OR^B, where R^B is C_{1-4} alkyl and R^{30} is selected from:

R^{30a}	
R^{30b}	
R^{30c}	
R^{30d}	
R^{30e}	
R^{30f}	

R^{30g}	
R^{30h}	

-C(=O)-X_1-NHC(=O)X_2-NH- represent a dipeptide. The amino acids in the dipeptide may be any combination of natural amino acids. The dipeptide may be the site of action for cathepsin-mediated cleavage.

In one embodiment, the dipeptide, -C(=O)-X_1-NHC(=O)X_2-NH-, is selected from:

-Phe-Lys-,
-Val-Ala-,
-Val-Lys-,
-Ala-Lys-,
-Val-Cit-,
-Phe-Cit-,
-Leu-Cit-,
-Ile-Cit-,
-Phe-Arg-,
-Trp-Cit-

where Cit is citrulline.

Preferably, the dipeptide, -C(=O)-X_1-NHC(=O)X_2-NH-, is selected from:

-Phe-Lys-,
-Val-Ala-,
-Val-Lys-,
-Ala-Lys-,
-Val-Cit-.

Most preferably, the dipeptide, -C(=O)-X_1-NHC(=O)X_2-NH-, is -Phe-Lys- or -Val-Ala-.

Other dipeptide combinations may be used, including those described by Dubowchik *et al.*, *Bioconjugate Chemistry*, 2002, 13,855-869, which is incorporated herein by reference.

In one embodiment, the amino acid side chain is derivatised, where appropriate. For example, an amino group or carboxy group of an amino acid side chain may be derivatised.

In one embodiment, an amino group NH_2 of a side chain amino acid, such as lysine, is a derivatised form selected from the group consisting of NHR and NRR'.
In one embodiment, a carboxy group COOH of a side chain amino acid, such as aspartic acid, is a derivatised form selected from the group consisting of COOR, $CONH_2$, CONHR and CONRR'.

In one embodiment, the amino acid side chain is chemically protected, where appropriate. The side chain protecting group may be a group as discussed above. The present inventors have established that protected amino acid sequences are cleavable by enzymes. For example, it has been established that a dipeptide sequence comprising a Boc side chain-protected Lys residue is cleavable by cathepsin.

Protecting groups for the side chains of amino acids are well known in the art and are described in the Novabiochem Catalog. Additional protecting group strategies are set out in Protective Groups in Organic Synthesis, Greene and Wuts.

Possible side chain protecting groups are shown below for those amino acids having reactive side chain functionality:

Arg: Z, Mtr, Tos;
Asn: Trt, Xan;
Asp: Bzl, t-Bu;
Cys: Acm, Bzl, Bzl-OMe, Bzl-Me, Trt;
Glu: Bzl, t-Bu;
Gln: Trt, Xan;
His: Boc, Dnp, Tos, Trt;
Lys: Boc, Z-Cl, Fmoc, Z, Alloc;

Ser: Bzl, TBDMS, TBDPS;

Thr: Bz;

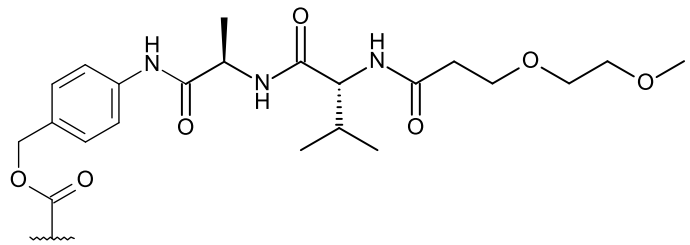
Trp: Boc;

Tyr: Bzl, Z, Z-Br.

In one embodiment, the side chain protection is selected to be orthogonal to a group provided as, or as part of, a capping group, where present. Thus, the removal of the side chain protecting group does not remove the capping group, or any protecting group functionality that is part of the capping group.

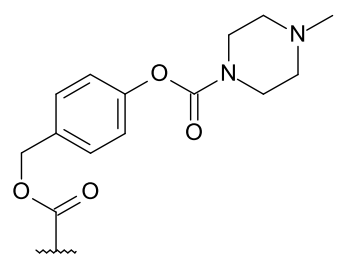
In other embodiments of the invention, the amino acids selected are those having no reactive side chain functionality. For example, the amino acids may be selected from: Ala, Gly, Ile, Leu, Met, Phe, Pro, and Val.

It is particularly preferred in the present invention, that if Q comprises a dipeptide, then -C(=O)-X_1-NHC(=O)X_2-NH- is the same dipeptide. An example of a preferred group is:

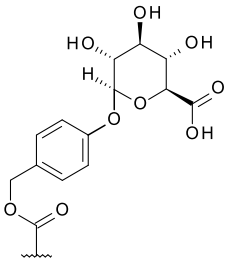


.

Other preferred R^{30} groups include:



and



.

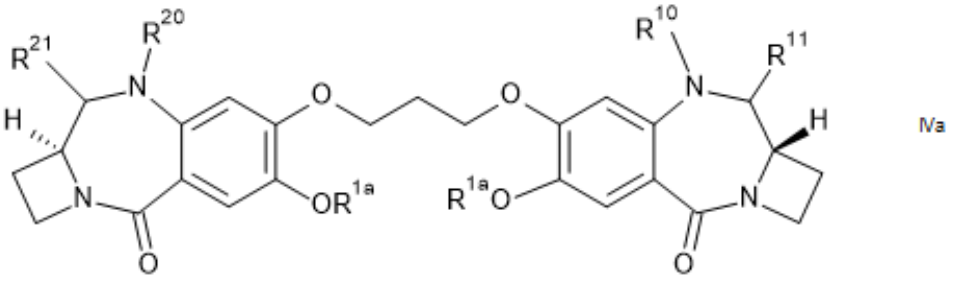
R^{11b} (Formulae I and I)*

In some embodiments, R^{11b} is OH.

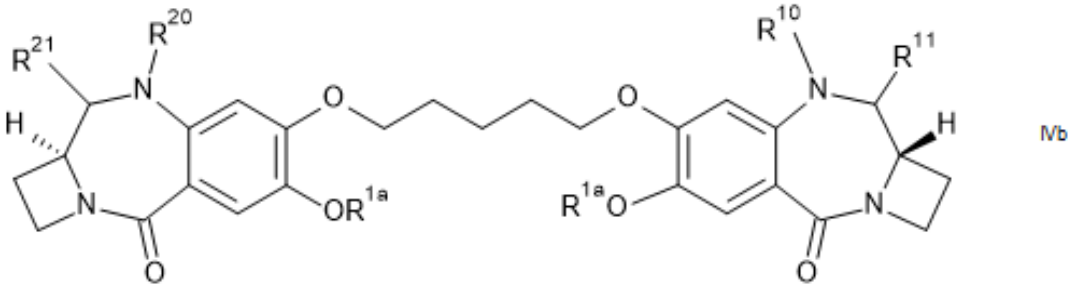
In some embodiments, R^{11b} is OR^A, where R^A is C_{1-4} alkyl. In some of these embodiments, R^A is methyl.

Further formulae

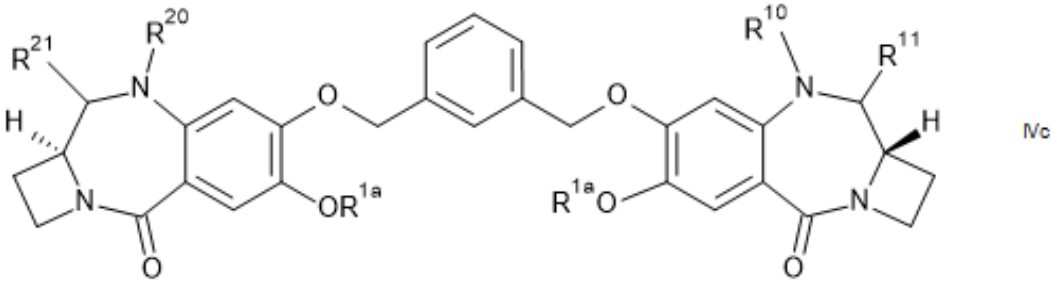
In some embodiments of the first aspect of the present invention are of formulae IVa, IVb or IVc:



IVa



IVb

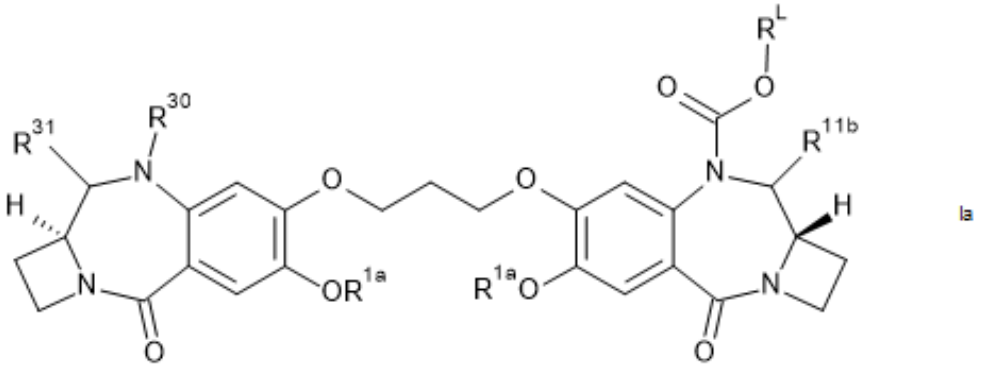


IVc

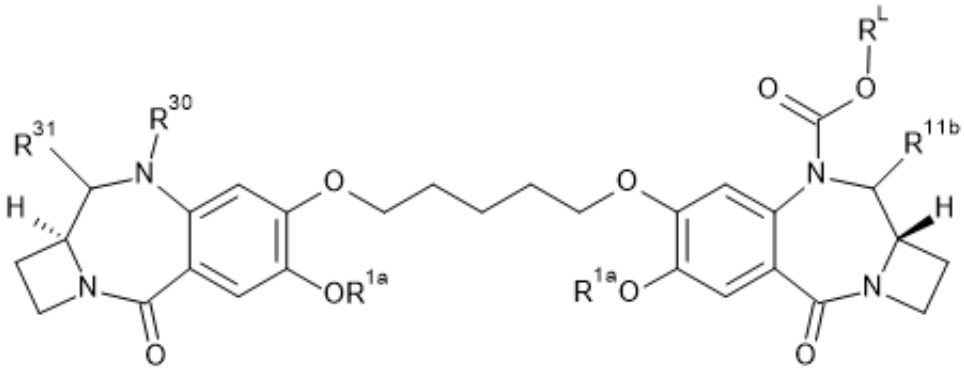
where R^{1a} is selected from methyl and benzyl;

R^{10}, R^{11}, R^{20} and R^{21} are as defined above.

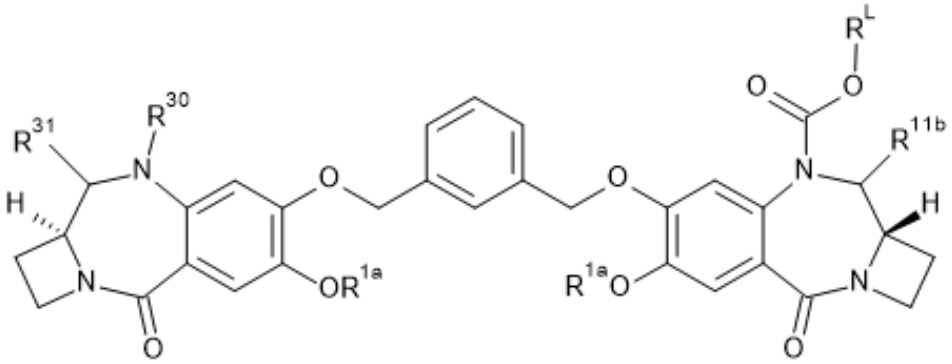
In some embodiments of the second aspect of the present invention are of formulae Ia, Ib or Ic:



Ia



Ib



Ic

where R^{1a} is selected from methyl and benzyl;

R^{30}, R^{31}, R^{L} and R^{11b} are as defined above.

These embodiments and preferences also apply to the third aspect of the invention.

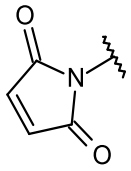
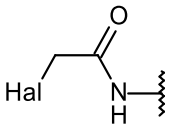
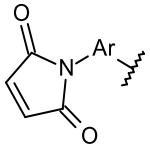
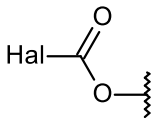
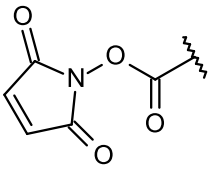
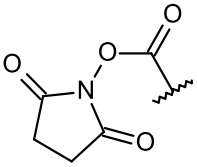
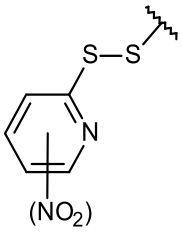
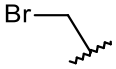
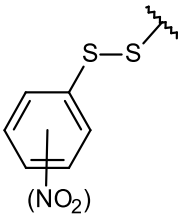
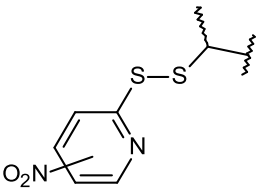
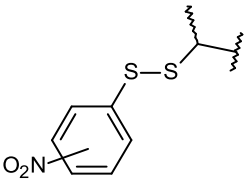
Linker (R^L)

In some embodiments, R^L is of formula IIIa.

In some embodiments, R^{LL} is of formula IIIa'.

G^L

G^L may be selected from

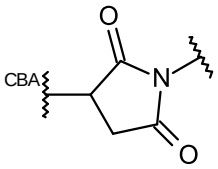
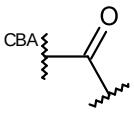
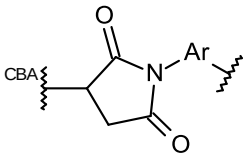
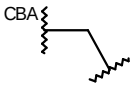
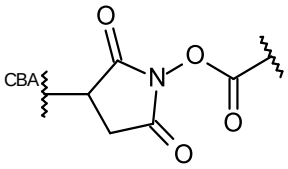
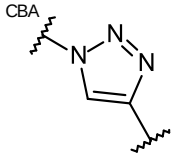
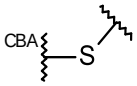
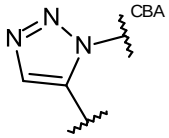
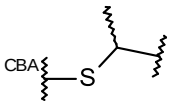
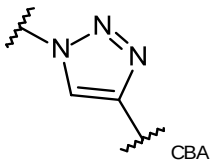
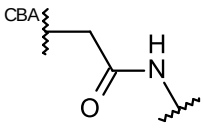
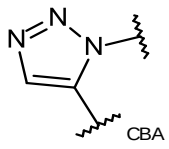
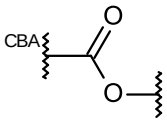
(G^{L1-1})		(G^{L4})	 Where Hal = I, Br, Cl
(G^{L1-2})		(G^{L5})	
(G^{L2})		(G^{L6})	
(G^{L3-1})	 where the NO_2 group is optional	(G^{L7})	
(G^{L3-2})	 where the NO_2 group is optional	(G^{L8})	
(G^{L3-3})	 where the NO_2 group is optional	(G^{L9})	
(G^{L3-4})	 where the NO_2 group is optional		

where Ar represents ~~a C_{5-6} arylene group, e.g.~~ phenylene.

In some embodiments, G^{L} is selected from G^{L1-1} and G^{L1-2}. In some of these embodiments, G^{L} is G^{L1-1}.

G^{LL}

G^{LL} may be selected from:

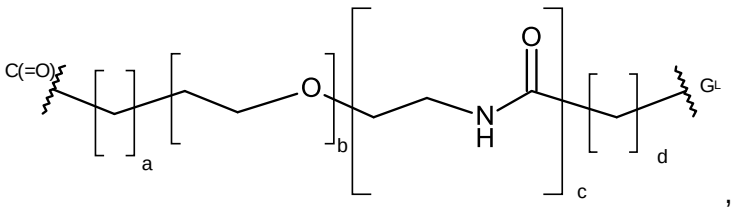
(G^{LL1-1})		(G^{LL6})	
(G^{LL1-2})		(G^{LL7})	
(G^{LL2})		(G^{LL8-1})	
(G^{LL3-1})		(G^{LL8-2})	
(G^{LL3-2})		(G^{LL9-1})	
(G^{LL4})		(G^{LL9-2})	
(G^{LL5})			

where Ar represents ~~a C_{5-6} arylene group, e.g.~~ phenylene.

In some embodiments, G^{LL} is selected from G^{LL1-1} and G^{LL1-2}. In some of these embodiments, G^{LL} is G^{LL1-1}.

X

X is:



,

where a = 0 to 5, b = 0 to 16, c = 0 or 1, d = 0 to 5.

a may be 0, 1, 2, 3, 4 or 5. In some embodiments, a is 0 to 3. In some of these embodiments, a is 0 or 1. In further embodiments, a is 0.

b may be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or 16. In some embodiments, b is 0 to 12. In some of these embodiments, b is 0 to 8, and may be 0, 2, 4 or 8.

c may be 0 or 1.

d may be 0, 1, 2, 3, 4 or 5. In some embodiments, d is 0 to 3. In some of these embodiments, d is 1 or 2. In further embodiments, d is 2.

In some embodiments of X, a is 0, c is 1 and d is 2, and b may be from 0 to 8. In some of these embodiments, b is 0, 4 or 8.

Q

In one embodiment, Q is an amino acid residue. The amino acid may a natural amino acids or a non-natural amino acid.

In one embodiment, Q is selected from: Phe, Lys, Val, Ala, Cit, Leu, Ile, Arg, and Trp, where Cit is citrulline.

In one embodiment, Q comprises a dipeptide residue. The amino acids in the dipeptide may be any combination of natural amino acids and non-natural amino acids. In some embodiments, the dipeptide comprises natural amino acids. Where the linker is a cathepsin labile linker, the dipeptide is the site of action for cathepsin-mediated cleavage. The dipeptide then is a recognition site for cathepsin.

In one embodiment, Q is selected from:

^{CO}-Phe-Lys-^{NH},

^{CO}-Val-Ala-^{NH},
^{CO}-Val-Lys-^{NH},
^{CO}-Ala-Lys-^{NH},
^{CO}-Val-Cit-^{NH},
^{CO}-Phe-Cit-^{NH},
^{CO}-Leu-Cit-^{NH},
^{CO}-Ile-Cit-^{NH},
^{CO}-Phe-Arg-^{NH}, and
^{CO}-Trp-Cit-^{NH};

where Cit is citrulline.

Preferably, Q is selected from:

^{CO}-Phe-Lys-^{NH},
^{CO}-Val-Ala-^{NH},
^{CO}-Val-Lys-^{NH},
^{CO}-Ala-Lys-^{NH},
^{CO}-Val-Cit-^{NH}.

Most preferably, Q is selected from ^{CO}-Phe-Lys-^{NH}, ^{CO}-Val-Cit-^{NH} and ^{CO}-Val-Ala-^{NH}.

Other dipeptide combinations of interest include:

^{CO}-Gly-Gly-^{NH},
^{CO}-Pro-Pro-^{NH}, and
^{CO}-Val-Glu-^{NH}.

Other dipeptide combinations may be used, including those described by Dubowchik et al., *Bioconjugate Chemistry*, 2002, 13,855-869, which is incorporated herein by reference.

In some embodiments, Q^X is a tripeptide residue. The amino acids in the tripeptide may be any combination of natural amino acids and non-natural amino acids. In some embodiments, the tripeptide comprises natural amino acids. Where the linker is a cathepsin labile linker, the tripeptide is the site of action for cathepsin-mediated cleavage. The tripeptide then is a recognition site for cathepsin. Tripeptide linkers of particular interest are:

^{CO}-Glu-Val-Ala-^{NH}
^{CO}-Glu-Val-Cit-^{NH}

^{CO}-αGlu-Val-Ala-^{NH}

^{CO}-αGlu-Val-Cit-^{NH}

In one embodiment, the amino acid side chain is chemically protected, where appropriate. The side chain protecting group may be a group as discussed below. Protected amino acid sequences are cleavable by enzymes. For example, a dipeptide sequence comprising a Boc side chain-protected Lys residue is cleavable by cathepsin.

Protecting groups for the side chains of amino acids are well known in the art and are described in the Novabiochem Catalog, and as described above.

In some embodiments, R^L is of formula IIIb.
In some embodiments, R^{LL} is of formula IIIb'.

R^{L1} and R^{L2} are independently selected from H and methyl, or together with the carbon atom to which they are bound form a cyclopropylene or cyclobutylene group.

In some embodiments, both R^{L1} and R^{L2} are H.

In some embodiments, R^{L1} is H and R^{L2} is methyl.

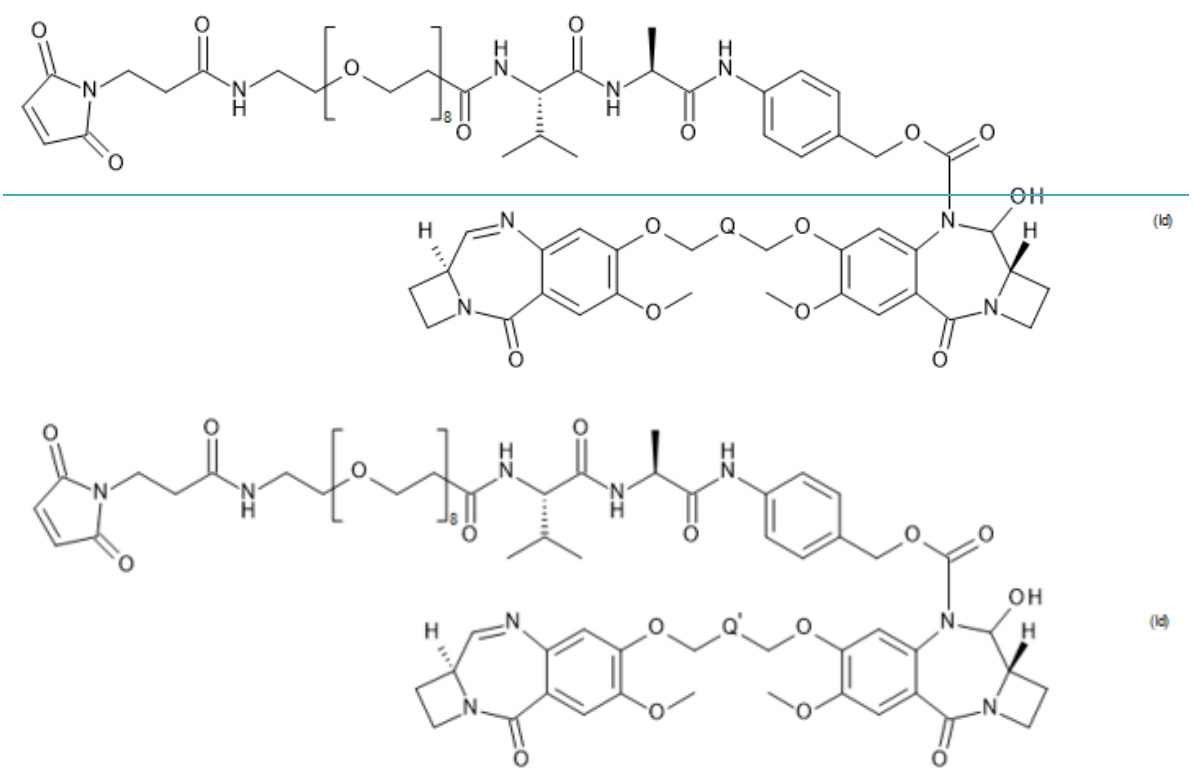
In some embodiments, both R^{L1} and R^{L2} are methyl.

In some embodiments, R^{L1} and R^{L2} together with the carbon atom to which they are bound form a cyclopropylene group.

In some embodiments, R^{L1} and R^{L2} together with the carbon atom to which they are bound form a cyclobutylene group.

In the group IIIb, in some embodiments, e is 0. In other embodiments, e is 1 and the nitro group may be in any available position of the ring. In some of these embodiments, it is in the ortho position. In others of these embodiments, it is in the para position.

In one particular embodiment, the second aspect of the invention comprises a compound of formula Id:



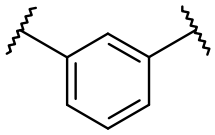
(Id)

(Id)

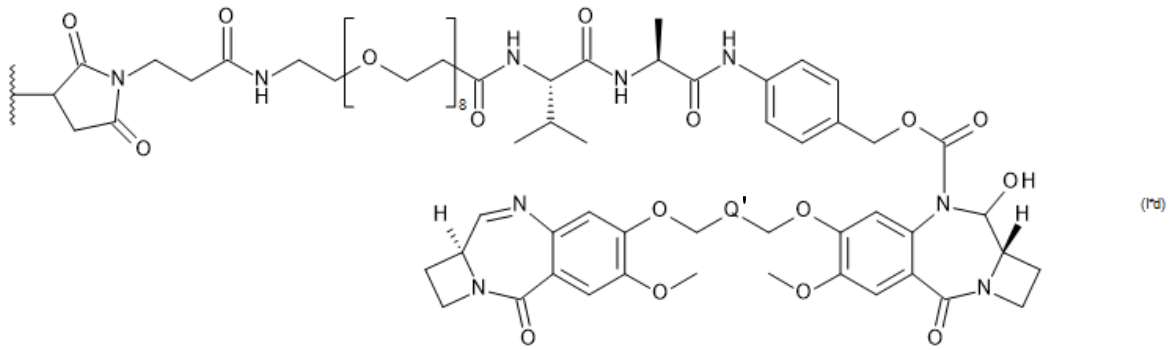
where Q' is selected from:

(a) $-CH_2-$;

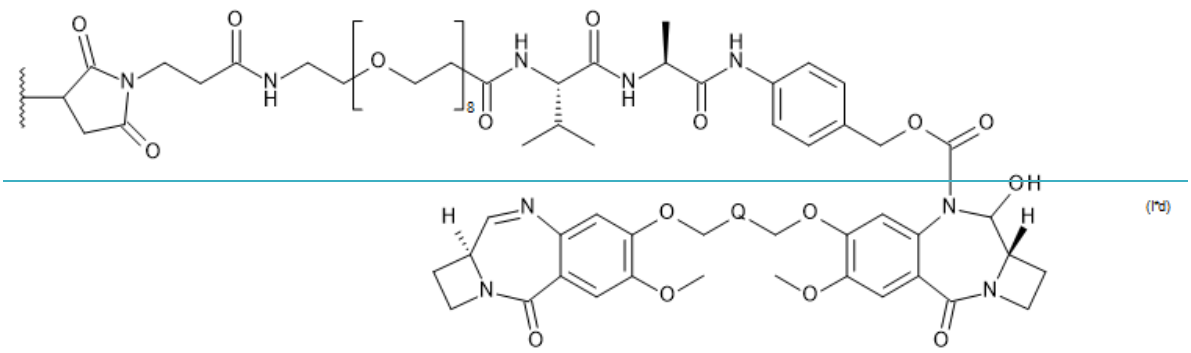
(b) $-C_3H_6-$; and

(c)  .

In one particular embodiment, the third aspect of the invention, the Drug linker (D^L) is of formula (I*d'):



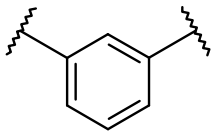
(I*d)



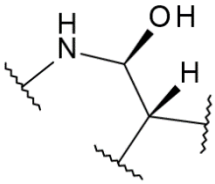
where Q' is selected from:

(a) $-CH_2-$;

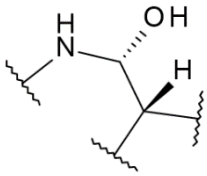
(b) $-C_3H_6-$; and

(c) .

In some embodiments of the present invention, the C11 substituent may be in the following stereochemical arrangement relative to neighbouring groups:



In other embodiments, the C11 substituent may be in the following stereochemical arrangement relative to neighbouring groups:



Examples

General Information

Manual flash chromatography was performed using Merck Kieselgel 60 F254 silica gel. Extraction and chromatography solvents were bought and used without further purification from Fisher Scientific, U.K. All chemicals were purchased from Aldrich, Lancaster or BDH. Automated flash chromatography was performed using a Biotage Isolera 1™ using gradient elution starting from either 88% hexane/EtOAc or 99.9% DCM/MeOH until all UV active components (detection at 214 and 254 nm) eluted from the column. The gradient

was manually held whenever substantial elution of UV active material was observed. Fractions were checked for purity using thin-layer chromatography (TLC) using Merck Kieselgel 60 F254 silica gel, with fluorescent indicator on aluminium plates. Visualisation of TLC was achieved with UV light or iodine vapour unless otherwise stated. Extraction and chromatography solvents were bought and used without further purification from VWR U.K. All fine chemicals were purchased from Sigma-Aldrich or TCI Europe unless otherwise stated. Pegylated reagents were obtained from Quanta biodesign US via Stratech UK.

The LC/MS conditions were as follow:

Positive mode electrospray mass spectrometry was performed using a Waters Aquity H-class. Mobile phases used were solvent A (water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid).

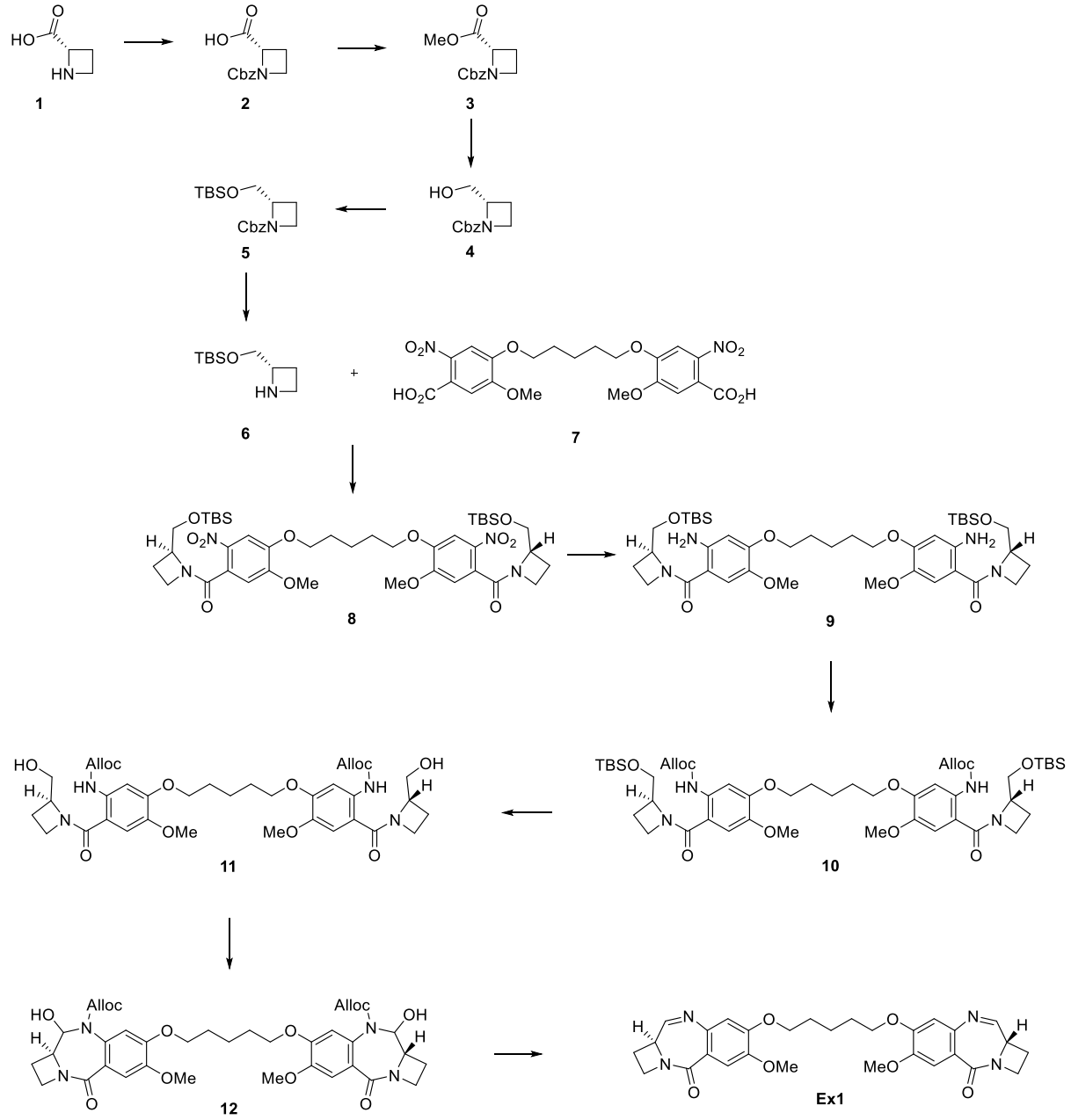
LCMS 3 min: initial composition was 5% B held over 0.25 min, then increase from 5% B to 100% B over a 2 min period. The composition was held for 0.50 min at 100% B, then returned to 5% B in 0.05 minutes and hold there for 0.05 min. Total gradient run time equals 3 min. Flow rate 0.8 mL/min. Detection was at 254 nm. Columns: Waters Acquity UPLC® BEH Shield RP18 1.7µm 2.1 x 50 mm at 50 °C fitted with Waters Acquity UPLC® BEH Shield RP18 VanGuard Pre-column, 130A, 1.7µm, 2.1 mm x 5 mm.

LCMS 15 min: initial composition 5% B held over 1 min, then increase from 5% B to 100% B over a 9 min period. The composition was held for 2 min at 100% B, then returned to 5% B in 0.10 minutes and hold there for 3 min. Total gradient run time equals 15 min. Flow rate 0.6 mL/min. Wavelength detection range: 190 to 800 nm. Oven temperature: 50°C. Column: ACE Excel 2 C18-AR, 2 µ, 3.0 x 100mm.

Preparative HPLC:

Reverse-phase ultra-fast high-performance liquid chromatography (UFLC) was carried out on a Shimazdzu Prominence® machine using a Phenomenex® Gemini NX 5µ C18 column (at 50 °C) dimensions: 150 x 21.2 mm. Eluents used were solvent A (H_2O with 0.1% formic acid) and solvent B (CH_3CN with 0.1% formic acid). All UFLC experiments were performed with gradient conditions: Initial composition 13% B increased to 60% B over a 15 minute period then increased to 100% B over 2 minutes. The composition was held for 1 minute at 100% B, then returned to 13% B in 0.1 minute and held there for 1.9 minutes. The total duration of the gradient run was 20.0 minutes. Flow rate was 20.0 mL/minute and detection was at 254 and 280 nm.

Example 1



a) 1-((benzyloxy)carbonyl)azetidine-2-carboxylic acid (**2**)

(2*S*)-Azetidine-2-carboxylic acid **1** (3 g, 29.674 mmol) and sodium bicarbonate (6.3 g, 75 mmol) were solubilised in H_2O (25 mL, 1387.75 mmol) and *N*-(benzyloxycarbonyl)succinimide (8.5 g, 34 mmol) in THF (25 mL, 307 mmol, 100 mass%) were added dropwise. After stirring at room temperature for 12 h, the two phases were allowed to separate. The aqueous phase was washed with diethyl ether (50 mL), cooled in an ice bath, and then acidified to pH=2 with conc. HCl. The aqueous layer was extracted with ethyl acetate (2 x 50 mL) and the combined organic extracts were dried ($MgSO_4$) and the excess solvent evaporated *in vacuo* to give crude product as a clear oil. The crude material was used without purification in the next step. LCMS 3min: ES^+ = 1.34 min, *m/z* 258.2 $[M + Na]^{\cdot+}$.

b) (1-benzyl 2-methyl (*S*)-azetidine-1,2-dicarboxylate **(3)**

In a dry round bottom flask, (2*S*)-1-benzyloxycarbonylazetidine-2-carboxylic acid **2** (6.98 g, 29.7 mmol) was solubilised in MeOH (65 mL) and sulfuric acid (3 mL) was added. The mixture was heated to reflux and left to stir overnight. The mixture was left to cool down to r.t. and quenched with Net_3 (to pH=7) before being stirred for 1h. Methanol was removed in vacuo. The residue was taken up in EtOAc, washed with H_2O and brine before being dried with $MgSO_4$ and filtered. The organics were removed in vacuo to give crude product **3** (8.004 g, 32.11 mmol) as a clear oil. LCMS 3min: ES^+ = 1.53 min, *m/z* no ionisation

c) Benzyl (*S*)-2-(hydroxymethyl)azetidine-1-carboxylate **(4)**

*O*1-benzyl *O*2-methyl (2*S*)-azetidine-1,2-dicarboxylate **3** (7.6 g, 30 mMol) was solubilised in THF (75 mL, 922 mmol), cooled to 0°C and $LiBH_4$ (1 g, 45 mMol,) was added. The mixture was allowed to warm to r.t. and stirred for a further hour at which point the reaction is complete. The reaction mixture was cooled to 0°C before being quench with H_2O and 1M HCl. The volatiles were removed in vacuo. The residue was taken up in EtOAc and washed with brine (2 x 50 mL), dried with $MgSO_4$, filtered and the solvent removed by rotary evaporation under reduced pressure. Purification by silica gel column chromatography (Hex/EtOAc, 100% to 1:2) afforded product a **4** as a clear oil (4.076 g, 60% yield over 3 steps). LCMS 3min: ES^+ =1.36 min, *m/z* 222.3 $[M + H]^{\cdot+}$.

d) Benzyl (*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carboxylate (**5**)

Benzyl (2*S*)-2-(hydroxymethyl)azetidine-1-carboxylate **4** (4.0766 g, 18.425 mmol) was solubilised in dry CH_2Cl_2 (20 mL, 312.0 mmol) and the mixture was cooled to 0°C before adding imidazole (2.508 g, 36.84 mmol) and TBS-Cl (4.16 g, 27.6 mmol). The mixture was

allowed to warm to room temperature and left to stir. LCMS shows reaction was complete within 5 min. The organics were washed with sat. NH_4Cl, water, brine, dried with $MgSO_4$, filtered and the volatiles removed in vacuo. Purification by silica gel column chromatography (Hex/EtOAc, 100% to 9:1) afforded product a **5** (6.90 g, not completely dry, quantitative). LCMS 3min: ES^+ =2.15 min, *m/z* 336.9 $[M + H]^{\cdot+}$.

e) (*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine **(6)**

Palladium on carbon (10%) (100 mg, 0.93 mMol) was treated with EtOAc (5 mL) dropwise and the resulting slurry added to a suspension of **5** (6.9027 g, 20.57 mmol) in EtOH (100 mL) at room temperature in a Parr hydrogenation bottle. The reaction mixture was subjected to H_2 gas at 137895 Pa (20 psi) then the bottle was evacuated under vacuum (repeated 3 times). The bottle was then topped up to 262001 Pa (38 psi) H_2 and shaken for 1 hour. The pressure dropped to ~206843 Pa (~30 psi) during this time and the bottle was topped up again to 275790 Pa (40 psi) and shaken for a further hour. No further decreases in pressure were observed and the reaction was deemed complete. This was confirmed by LC-MS. The mixture was filtered through celite and the filtrate evaporated in vacuo to provide the crude product **6** as a brown oil (3.761 g, 90% yield). LCMS 3min: ES^+ =1.70 min, *m/z* no ionisation.

f) ((*S*)-2-(((*Tert*-butyldimethylsilyl)oxy)methyl)azetidin-1-yl)(4-(6-(4-((2*R*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)cyclobutane-1-carbonyl)-2-methoxy-5-nitrophenoxy)hexyl)-5-methoxy-2-nitrophenyl)methanone **(8)**

DCC (3.8 g, 18mmol) was added to a solution of **7** (3.9 g, 7.9 mMol) and HOBt (2.3g, 17 mMol) in CH_2Cl_2 (200 mL) at 0°C. The cold bath was removed and the reaction was allowed to proceed for 30 min at room temperature, at which time a solution of **6** (3.65 g, 18 mMol) and triethylamine (3.2 mL, 23 mmol) in CH_2Cl_2 (200 mL) was added rapidly at -10 °C under argon The reaction mixture was allowed to stir at room temperature and monitored by LC/MS. After 2 min, the reaction was complete. The solids were removed by filtration over celite and the organic phase was washed with cold aqueous 0.1 M HCl until the pH was measured at 2. The organic phase was then washed with water, followed by saturated aqueous sodium bicarbonate, brine, dried with $MgSO_4$, filtered and vacced down under reduced pressure. Purification by silica gel column chromatography (Hex/EtOAc/CH_2Cl_2, 100% to 1:2:1) afforded product **8** (5.9 g, 87% yield). The product is contaminated with some mono-coupled product (impurity does not separate upon chromatography). LCMS 3min: ES^+ =2.35 min, *m/z* 862.2 $[M+H]^{\cdot+}$.

g) ((Pentane-1,5-diylbis(oxy))bis(2-amino-5-methoxy-4,1-phenylene))bis(((*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidin-1-yl)methanone) **(9)**

Zinc (4.65 g, 71.1 mmol) was slowly added to a solution of **8** (2.45 g, 2.85 mmol) in a mixture of MeOH/H_2O/formic acid 90:5:5 (66 mL). The resulting exotherm was controlled using an ice bath to maintain the temperature of the reaction mixture below 40°C. Upon completion, the solids were removed by filtration over celite and the organic phase was washed with water and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Crude material **9** (2.28 g, quantitative) was used as such in the next step. LCMS 3min: ES^+ =2.32 min, *m/z* 802.3 $[M+H]^{\cdot+}$.

h) Diallyl ((pentane-1,5-diylbis(oxy))bis(6-((*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carbonyl)-4-methoxy-3,1-phenylene))dicarbamate **(10)**

Compound **9** (2.23 g, 2.78 mmol) was solubilised in CH_2Cl_2 (50 mL) under an argon atmosphere. The mixture was cooled to -78°C before pyridine (0.99 mL, 12.3 mMol) and allyl chloroformate (0.738 mL, 2.49 mmol) were added. The reaction was left to stir at -78°C for 10 min before being allowed to warm up to room temp. After 15 min the reaction was complete. The organics were washed with sat. $CuSO_4$, H_2O, brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Crude product **10** (1.47 g, 1.52 mMol, quantitative) was used as such in the next step. LCMS 3min: ES^+ =2.53 min, *m/z* 970.3 $[M+H]^{\cdot+}$.

i) Diallyl ((pentane-1,5-diylbis(oxy))bis(6-((*S*)-2-(hydroxymethyl)azetidine-1-carbonyl)-4-methoxy-3,1-phenylene))dicarbamate **(11)**

Compound **10** (1.47 g, 1.52 mMol) was solubilised in a 3:1:1 mixture of H_2O/THF/acetic acid (16 mL) and the reaction was left to stir over the weekend. The mixture was extracted with CH_2Cl_2 and washed with sat. $NaHCO_3$, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography (Hex/EtOAc, 100% to 1:1) afforded product **11** (859 mg, 76.5% Yield) as a clear oil. LCMS 3min: ES^+ =1.75 min, *m/z* 742.0 $[M+H]^{\cdot+}$.

j) Diallyl 7,7'-(pentane-1,5-diylbis(oxy))(10a*S*,10a'*S*)-*bis*(10-hydroxy-6-methoxy-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate) **(12)**

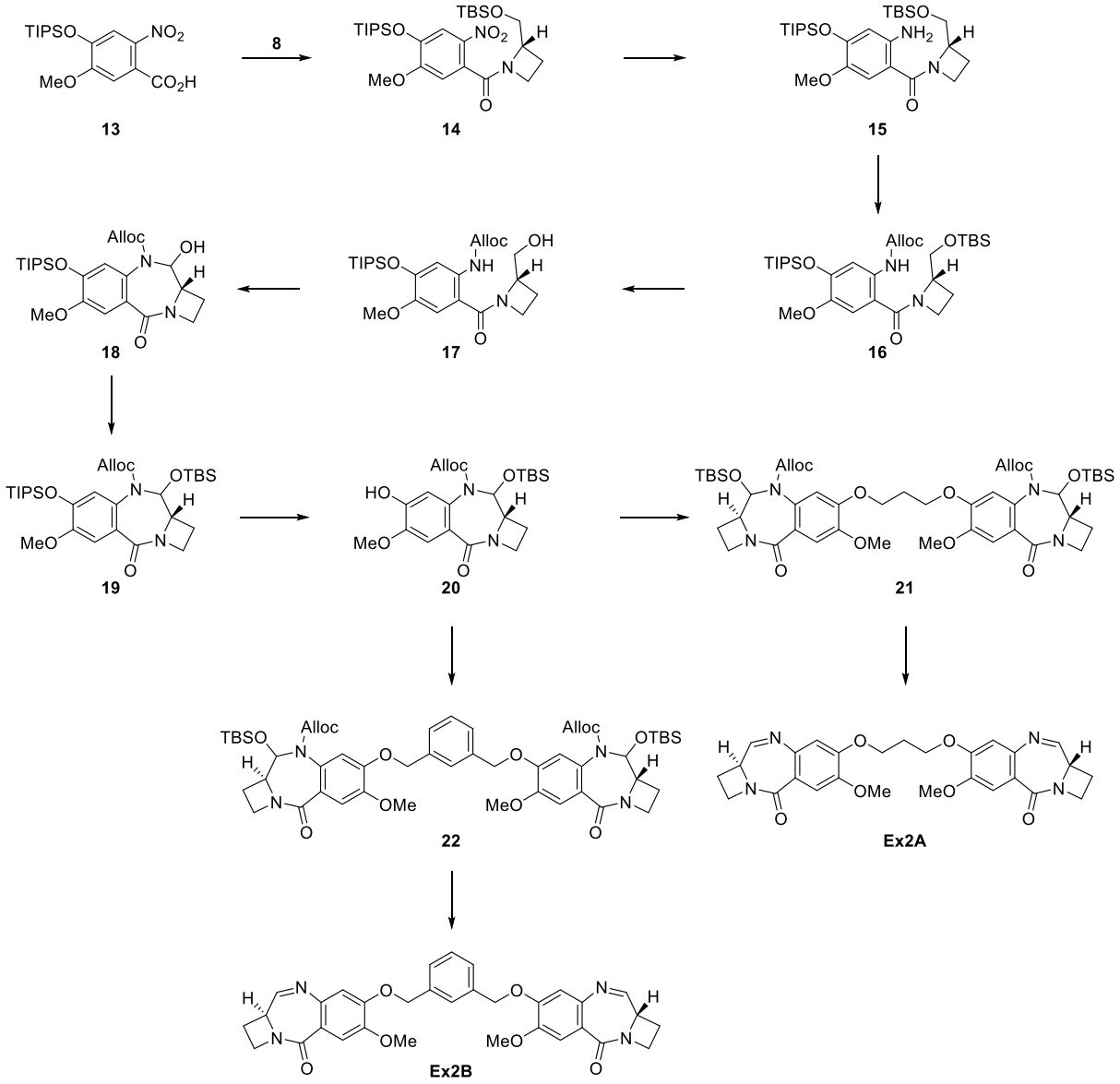
Compound **11** (850 mg, 1.14 mMol) was solubilised in CH_2Cl_2 (60 mL). 1-hydroxy-2,2,6,6-tetramethyl-piperidine; 1-methylimidazole; 2-(2-pyridyl)pyridine (0.7 mL, 1140 mmol, 0.2 mMl/L) and tetrakisacetonitrile copper(I) triflate (55 mg, 0.145 mMol) were subsequently

added ant the mixture stirred at 35°C with 2 balloon of air pressing in. The reaction was left to stir overnight before being vacuumed to dryness in a rotary evaporator. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 95:5) afforded product **12** (346 g, 0.47 mMol, 41% yield). LCMS 3min: ES^+ = 1.48 min, *m/z* 737.9 $[M+H]^{\cdot+}$.

k) (10a*S*,10a'*S*)-7,7'-(pentane-1,5-diylbis(oxy))bis(6-methoxy-1,10a-dihydroazeto[1,2-*a*]benzo[*e*][1,4]diazepin-4(2*H*)-one) **(Ex1)**

Compound **12** (335 mg, 0.45 mmol) was solubilised in CH_2Cl_2 (20 mL) in a flask under Argon. Pyrrolidine (650 µL, 7.8 mMol) and $Pd(PPh_3)_4$ (50 mg, 0.004 mMol) were subsequently added and the mixture left to stir at r.t until complete. The organics were washed with sat. NH_4Cl, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by isolera chromatography (CH_2Cl_2/(CH_2Cl_2+10%MeOH) 92:7 to 10:90. Two fractions containing the product were isolate but with insufficient purity. The fractions were combined and repurified by manual chromatography and pure product **Ex1** was isolated (146 mg, 0.27 mMol, 24% yield). LCMS 3 min: ES^+ = 1.32 min, *m/z* 533.8 $[M+H]^{\cdot+}$. LCMS 15 min: ES^+ = 4.83 min, *m/z* 533.9 $[M+H]^{\cdot+}$.

Example 2



a) ((*S*)-(2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidin-1-yl)(5-methoxy-2-nitro-4-((tri*iso*propylsilyl)oxy)phenyl)methanone **(13)**

DCC (4.021 g, 19.49 mmol) was added to a solution of 5-methoxy-2-nitro-4-triisopropylsilyloxy-benzoic acid **13** (6 g, 16.24 mmol), and HOPO (1.984 g, 17.86 mMol) in CH_2Cl_2 (100 mL) at 0°C. The cold bath was removed and the reaction was allowed to proceed for 30 min at room temperature, at which time a solution of [(2*S*)-azetidin-2-yl]methoxy-*tert*-butyl-dimethyl-silane **6** (3.761 g, 18.68 mmol) and triethylamine (3.39 mL, 33.5 mmol) in CH_2Cl_2 (100 mL) was added rapidly at -10°C under argon The reaction mixture was allowed to stir at room temperature and monitored by LC/MS. After 2 min, the reaction was complete. The solids were removed by filtration over celite and the organic

phase was washed with cold aqueous 0.1 M HCl until the pH was measured at 2. The organic phase was then washed with water, followed by saturated aqueous sodium bicarbonate, brine, dried with $MgSO_4$, filtered and vacced down under reduced pressure. Purification by silica gel column chromatography (Hex/EtOAc, 100% to 1:1) afforded product **14** (8.6737 g, 96.63% yield). LCMS 3min: ES^+ =2.44 min, *m/z* 554.2 $[M+H]^{\cdot+}$.

b) (*S*)-(2-amino-5-methoxy-4-((triisopropylsilyl)oxy)phenyl)(2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidin-1-yl)methanone **(15)**

Zinc (10 g, 152.9 mMol) was slowly added to a solution of **14** (8.6737 g, 15.69 mMol) in a mixture of MeOH/H_2O/formic acid 90:5:5 (200 mL). The resulting exotherm was controlled using an ice bath to maintain the temperature of the reaction mixture below 40°C. Upon completion, the solids were removed by filtration over celite and the organic phase was washed with water and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Crude material **15** (7.6343 g, 14.6 mMol, 93.05% yield) was used as such in the next step. LCMS 3min: ES^+ =2.42 min, *m/z* 524.4 $[M+H]^{\cdot+}$.

c) allyl (*S*)-(2-(2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carbonyl)-4-methoxy-5-((tri*iso*propylsilyl)oxy)phenyl)carbamate **(16)**

Compound **15** (7.6343 g, 14.60 mMol) was solubilised in CH_2Cl_2 (100 mL) under an argon atmosphere. The mixture was cooled to -78°C before pyridine (2.6 mL, 32 mMol) and allyl chloroformate (1.7 mL, 16 mMol) were added. The reaction was left to stir at -78°C for 10 min before being allowed to warm up to room temp. After 15 min the reaction was complete. The organics were washed with sat. $CuSO_4$, H_2O, brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Crude product **16** (8.9129 g, 14.69 mMol, quantitative) was used as such in the next step. LCMS 3min: ES^+ =2.53 min, *m/z* 608.2 $[M+H]^+$.

d) Allyl (*S*)-(2-(2-(hydroxymethyl)azetidine-1-carbonyl)-4-methoxy-5-((triisopropylsilyl)oxy)phenyl)carbamate **(17)**

Compound **16** (8.9129 g, 14.69 mMol) was solubilised in a 3:1:1 mixture of H_2O/THF/acetic acid (80 mL) and the reaction was left to stir over the weekend. The mixture was extracted with CH_2Cl_2 and washed with sat. $NaHCO_3$, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography (Hex/EtOAc, 100% to 1:1) afforded product **17** (5.5572 g, 76.80% yield) as a clear oil. LCMS 3min: ES^+ =1.97 min, *m/z* 494.0 $[M+H]^+$.

e) Allyl (10a*S*)-10-hydroxy-6-methoxy-4-oxo-7-((triisopropylsilyl)oxy)-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate **(18)**.
Compound **17** (5.5572 g, 11.28 mMol) was solubilised in CH_2Cl_2 (40 mL). 1-hydroxy-2,2,6,6-tetramethyl-piperidine; 1-methylimidazole; 2-(2-pyridyl)pyridine (6 mL, 1 mMol) and tetrakisacetonitrile copper(I) triflate (425 mg, 1.1279 mMol) were subsequently added ant the mixture stirred at 35°C with 2 balloon of air pressing in. The reaction was left to stir overnight before being vacuumed to dryness in a rotary evaporator. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 97:3) afforded product **18** (5.3835 g, 10.97 mMol, 97.27% yield) as a light orange foam. LCMS 3min: ES^+ =2.00 min, *m/z* 491.8 $[M+H]^{\cdot+}$.

f) Allyl (10aS)-10-((*tert*-butyldimethylsilyl)oxy)-6-methoxy-4-oxo-7-((tri*iso*propylsilyl)oxy)-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate **(19)**
Compound **18** (5.3835 g, 10.97 mMol) was solubilised in CH_2Cl_2 (50 mL) and the mixture was cooled to -78°C. 2,6-Lutidine (2.55 mL, 21.9 mMol) and TBS-OTf (3.78 mL, 16.4 mMol) were subsequently added. The mixture was left for 10 min before removing the cooling bath and allowing to warm to r.t. The organics were washed with, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 95:5) afforded product **19** (6.8532 g, quantitative). LCMS 3min: ES^+ =2.47 min, *m/z* 606.0 $[M+H]^{\cdot+}$.

g) Allyl (10a*S*)-10-((*tert*-butyldimethylsilyl)oxy)-7-hydroxy-6-methoxy-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate **(20)**
Compound **19** (6.8 g, 14 mMol) was solubilised in DMF (10 mL). LiOAc.$2H_2O$ (1.4 g, 14 mmMl) and H_2O (3 mL or as much as possible) were added. When the solution becomes clear again, add a few drops of water. Keep repeating the process until the reaction is complete. The organics were diluted with $CHCl_3$ and washed with a citric acid solution (pH=3), H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 95:5) afforded product **20** (5.2885 g, 11.79 mMol, 85% yield) as a yellow oil. LCMS 3min: ES^+ =1.86 min, *m/z* 449.8 $[M+H]^{\cdot+}$.

h) Diallyl 7,7'-(propane-1,3-diylbis(oxy))(10a*S*,10a'*S*)-bis(10-((*tert*-butyldimethylsilyl)oxy)-6-methoxy-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate) **(21)**
1,3-Dibromopropane (204.9 mg, 1.015 mMol) and compound **20** (1 g, 2.030 mMol) were

solubilised in CH_2Cl_2 (50 mL) under an argon atmosphere. K_2CO_3 (280 mg, 2.026 mMol) and TBAI (149 mg, 0.2 mMol) were subsequently added and the mixture was allowed to stir at 40°C until complete. The mixture was left to stir overnight but the reaction does not go to completion instead an impurity formed. The organics were washed with H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 97:3) afforded product **21** (482 mg, 0.471 mMol, 46.50% yield), contaminated with an inseparable impurity (r.t=9.95 min on LCMS 15 min). LCMS 15 min: ES^+ =9.86 min, *m/z* 938.3 $[M+H]^{\cdot+}$.

i) (10a*S*,10a'*S*)-7,7'-(propane-1,3-diylbis(oxy))bis(6-methoxy-1,10a-dihydroazeto[1,2-*a*]benzo[*e*][1,4]diazepin-4(2*H*)-one) **(Ex2A)**

Compound **21** (482 mg, 0.5143 mMol) was solubilised in CH_2Cl_2 (20 mL) in a flask under argon. Pyrrolidine (786 µL, 9.44 mMol) and $Pd(PPh_3)_4$ (54 mg, 0.046 mMol) were subsequently added and the mixture left to stir at r.t until complete. The organics were washed with sat. NH_4Cl, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by isolera chromatography (CH_2Cl_2/(CH_2Cl_2+10%MeOH) 98:2 to 30:70. Two fractions containing the product were isolated but with insufficient purity. The fractions were combined and repurified by isolera chromatography (same solvent system) and pure product **Ex2A** was isolated (35.1 mg, 0.135 mMol, 13.5% yield). LCMS 3 min: ES^+ = 1.23 min, *m/z* 505.8 $[M+H]^{\cdot+}$.

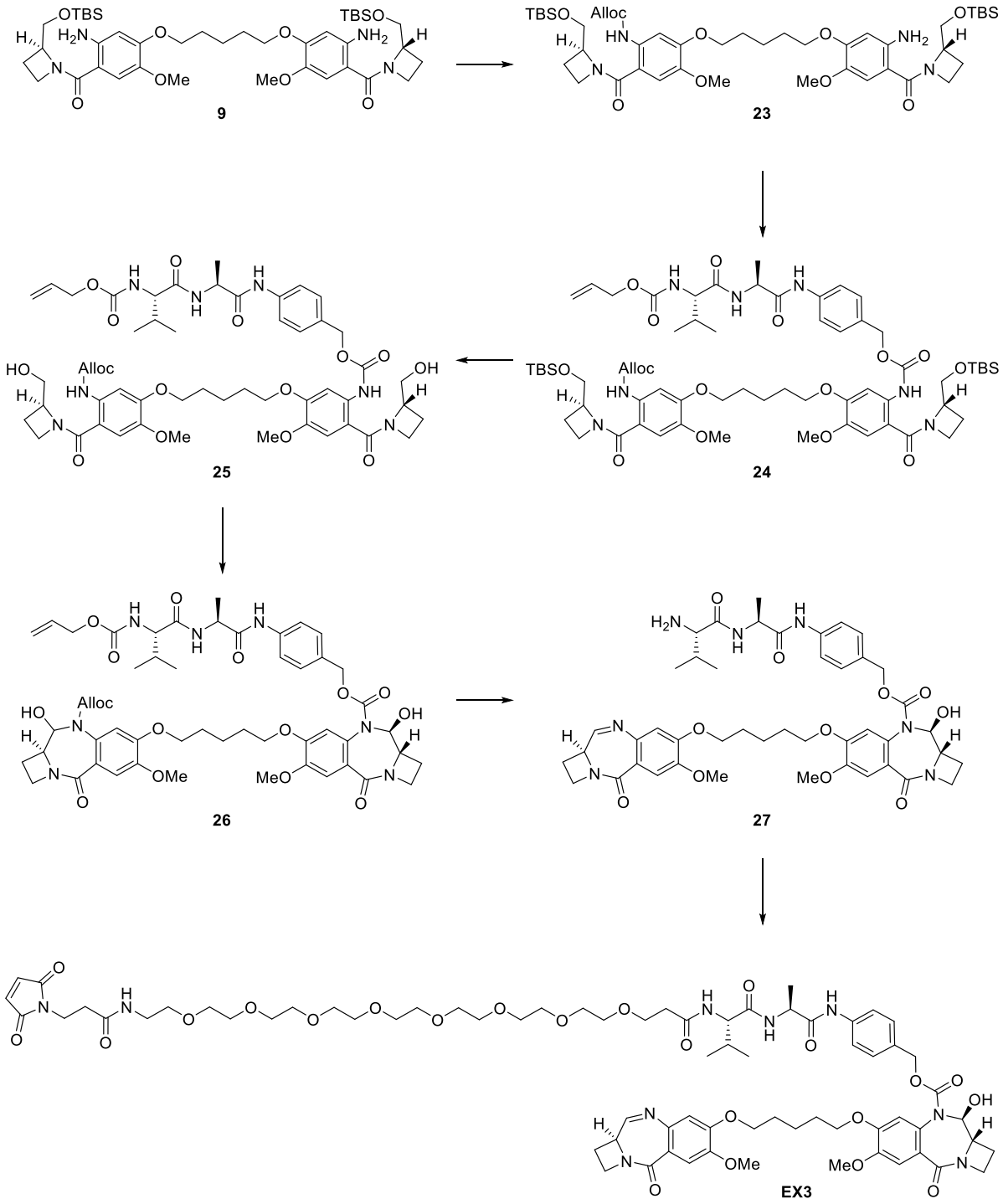
j) Diallyl 7,7'-((1,3-phenylenebis(methylene))bis(oxy))(10aS,10a'*S*)-bis(10-((*tert*-butyldimethylsilyl)oxy)-6-methoxy-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate) **(22)**

1,3-Bis(bromomethyl)benzene (267.9 mg, 1.011 mMol) and compound **20** (1 g, 2.030 mMol) were solubilised in DMF (5 mL) under an argon atmosphere. K_2CO_3 (280 mg, 2.026 mMol) and TBAI (749 mg, 2.027 mMol) were subsequently added and the mixture was allowed to stir at 40°C until complete. The mixture was left to stir overnight but the reaction did not go to completion and an impurity formed. The mixture was diluted with CH_2Cl_2 and washed with H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 97:3) afforded product **22** (467 mg, 0.43 mMol, 42.47% yield) + 398 mg of mixed fractions. LCMS 3 min: ES^+ =2.30 min, *m/z* 1000.5 $[M+H]^{\cdot+}$.

k) (10a*S*,10a'*S*)-7,7'-((1,3-phenylenebis(methylene))bis(oxy))bis(6-methoxy-1,10a-dihydroazeto[1,2-*a*]benzo[*e*][1,4]diazepin-4(2*H*)-one) **(Ex2B)**

Compound **22** (455 mg, 0.419 mMol) was solubilised in CH_2Cl_2 (20 mL) in a flask under Argon. Pyrrolidine (600 µL, 7.2 mMol) and $Pd(PPh_3)_4$ (48 mg, 0.041 mMol) were subsequently added and the mixture left to stir at r.t until complete. The organics were washed with sat. NH_4Cl, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by isolera chromatography (CH_2Cl_2/(CH_2Cl_2+10%MeOH) 98:2 to 30:70. Two fractions containing the product were isolate but with insufficient purity. The fractions were combined and repurified by isolera chromatography (same solvent system) and pure product **Ex2B** was isolated (214.5 mg, 0.378 mMol, 90.5% yield) as a white solid. LCMS 3 min: ES^+ = 1.38 min, *m/z* 567.8 $[M+H]^{\cdot+}$.

Example 3



a) Allyl (5-((5-(5-amino-4-((*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carbonyl)-2-methoxyphenoxy)pentyl)oxy)-2-((*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carbonyl)-4-methoxyphenyl)carbamate **(23)**

Compound **9** (1.192 g, 1.488 mMol) was solubilised in CH_2Cl_2 (250 mL) under an argon atmosphere. The mixture was cooled to -78°C before pyridine (0.241 mL, 2.98 mMol) and allyl chloroformate (0.158 mL, 1.484 mMol) were added. The reaction was left to stir at -

78°C for 10 min before being allowed to warm up to room temp. After 15 min the reaction was complete. The organics were washed with sat. $CuSO_4$, H_2O, brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH) afforded a mixture of mono and bis-alloc which was purified further with a second column (Hex/EtOAc) to give pure product **23** (499.2 g, 37.9% yield out of 50% possible). LCMS 3min: ES^+ =2.41 min, *m/z* 886.6 $[M+H]^{\cdot+}$.

b) Allyl (5-((5-(5-((((4-((*S*)-2-((*S*)-2-(((allyloxy)carbonyl)amino)-3-methylbutanamido)propanamido)benzyl)oxy)carbonyl)amino)-4-((*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carbonyl)-2-methoxyphenoxy)pentyl)oxy)-2-((*S*)-2-(((*tert*-butyldimethylsilyl)oxy)methyl)azetidine-1-carbonyl)-4-methoxyphenyl)carbamate **(24)**

Triphosgene (68.8 mg, 0.232 mMol) was added in one portion to a mixture of **23** (620 mg, 0.7 mMol) and TEA (203 µL, 1.46 mMol) in CH_2Cl_2 (50 mL) at 0°C. The ice bath was removed, and after 15 min, Alloc-Val-Ala-PAB-OH (275 mg, 0.728 mMol) was added in one portion as a fine powder, followed by more TEA (73 µL, 0.524 mMol,) and Dibutyltin dilaurate (39.6 µL, 0.07 mMol). The reaction mixture was allowed to stir at 37°C for 4h, followed by stirring at room temperature overnight. The organics were washed with H_2O, sat. NH_4Cl and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH) to give pure product **24** (414 g, 45.9% yield). LCMS 3min: ES^+ =2.43 min, *m/z* 1289.5 $[M+H]^{\cdot+}$.

c) Allyl (5-((5-(5-((((4-((*S*)-2-((*S*)-2-(((allyloxy)carbonyl)amino)-3-methylbutanamido)propanamido)benzyl)oxy)carbonyl)amino)-4-((*S*)-2-(hydroxymethyl)azetidine-1-carbonyl)-2-methoxyphenoxy)pentyl)oxy)-2-((*S*)-2-(hydroxymethyl)azetidine-1-carbonyl)-4-methoxyphenyl)carbamate **(25)**

Compound **24** (414 mg, 0.32 mMol) was solubilised in a 3:1:1 mixture of H_2O/THF/acetic acid (10 mL) and the reaction was left to stir over the weekend. The mixture was extracted with CH_2Cl_2 and washed with sat. $NaHCO_3$, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 94:6) afforded product **25** (326 mg, 95.7% Yield). LCMS 3min: ES^+ =1.80 min, *m/z* 1060.1 $[M+H]^{\cdot+}$.

d) Allyl (10a*S*)-7-((5-(((10*S*,10a*S*)-9-(((4-((*S*)-2-((*S*)-2-(((allyloxy)carbonyl)amino)-3-methylbutanamido)propanamido)benzyl)oxy)carbonyl)-10-hydroxy-6-methoxy-4-oxo-1,2,4,9,10,10a-hexahydroazeto[1,2-a]benzo[e][1,4]diazepin-7-yl)oxy)pentyl)oxy)-10-hydroxy-6-methoxy-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate **(26)**

Compound **25** (202.4 mg, 0.3 mMol) was solubilised in CH_2Cl_2 (20 mL). 1-hydroxy-2,2,6,6-tetramethyl-piperidine; 1-methylimidazole; 2-(2-pyridyl)pyridine (0.4 mL, 0.03 mMol) and tetrakisacetonitrile copper(I) triflate (11 mg, 0.03 mMol) were subsequently added ant the mixture stirred at 35°C with 2 balloon of air pressing in. The reaction was left to stir overnight before being vacuumed to dryness in a rotary evaporator. Purification by silica gel column chromatography ($CHCl_3$/MeOH, 100% to 97:3) afforded product **26** (313 mg, 0.19 mMol, 64.5% yield). LCMS 3min: ES^+ =1.59 min, *m/z* 1057.1 $[M+H]^{\cdot+}$.

e) 4-((*S*)-2-((*S*)-2-amino-3-methylbutanamido)propanamido)benzyl (10*S*,10a*S*)-10-hydroxy-6-methoxy-7-((5-(((*S*)-6-methoxy-4-oxo-1,2,4,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepin-7-yl)oxy)pentyl)oxy)-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepine-9(4H)-carboxylate) **(27)**

Compound **26** (195 mg, 0.184 mMol) was solubilised in CH_2Cl_2 (10 mL) in a flask under Argon. Pyrrolidine (262 µL, 3.15 mMol) and $Pd(PPh_3)_4$ (21 mg, 0.018 mMol) were subsequently added and the mixture left to stir at r.t until complete. The organics were washed with sat. NH_4Cl, H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by isolera chromatography (CH_2Cl_2/(CH_2Cl_2+10%MeOH) 98:2 to 30:70 gave product **27** (141 mg, 0.16 mMol, 87.7% yield). LCMS 3 min: ES^+ = 1.23 min, *m/z* 870.9 $[M+H]^{\cdot+}$.

f) 4-((2*S*,5*S*)-37-(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)-5-isopropyl-2-methyl-4,7,35-trioxo-10,13,16,19,22,25,28,31-octaoxa-3,6,34-triazaheptatriacontanamido)benzyl (10*S*,10a*S*)-10-hydroxy-6-methoxy-7-((5-(((*S*)-6-methoxy-4-oxo-1,2,4,10a-tetrahydroazeto[1,2-*a*]benzo[*e*][1,4]diazepin-7-yl)oxy)pentyl)oxy)-4-oxo-1,2,10,10a-tetrahydroazeto[1,2-*a*]benzo[*e*][1,4]diazepine-9(4*H*)-carboxylate (**Ex3)**

The reaction was conducted in a glovebox. Compound **27** (70 mg, 0.080 mMol) was solubilised in CH_2Cl_2 (10 mL) in a flask under argon at room temperature. Mal-dPEG$_8$-OH (50 mg, 0.084 mMol) and EDCI.HCl (15.4 mg, 0.080 mMol) were added and the mixture was stirred until completion. The organics were washed with H_2O and brine before being dried with $MgSO_4$, filtered and the volatiles removed under reduced pressure. Purification by isolera chromatography (CH_2Cl_2/(CH_2Cl_2+10%MeOH) 98:2 to 30:70 gave an impure

product. Further purification by reverse phase isolera gave pure **Ex3** (4 mg, 0.027 mMol, 3.4% yield) plus some unclean fractions (22 mg). LCMS 3 min: ES^+ = 1.51 min, *m/z* 1445.6 $[M+H]^{\cdot+}$.

Example 4

ConjA (Her2-Ex3)

A 10 mM solution of Tris(2-carboxyethyl)phosphine (TCEP) in phosphate-buffered saline pH 7.4 (PBS) was added (50 molar equivalent/antibody, 7.6 micromoles, 762.7 μL) to a 20.8 mL solution of tratuzumab (22.9 mg, 153 nanomoles) in reduction buffer containing 30 mM histidine / histidine HCl, 30 mM arginine, pH 6.8 and 1 mM ethylenediaminetetraacetic acid (EDTA) and a final antibody concentration of 1.1 mg/mL. The reduction mixture was allowed to react at 37°C for 2 hours (or until full reduction is observed by UHPLC) in an orbital shaker with gentle (60 rpm) shaking. The reduced antibody solution was buffer exchanged (to remove all the excess reducing agent), *via* spin filter centrifugation, into a conjugation buffer containing 30 mM histidine / histidine HCl, 30 mM arginine and 1 mM EDTA for a final antibody concentration of 1.1 mg/mL. **Ex3** was added as a DMSO solution (12.5 molar equivalent/antibody, 1.9 micromoles, in 2.1 mL DMSO) to 18.6 mL of this reduced antibody solution (20.5 mg, 136 nanomoles) for a 10% (*v/v*) final DMSO concentration. The solution was mixed for 17 hours at room temperature, then the conjugation was quenched by addition of *N*-acetyl cysteine (8.5 micromoles, 68 μL at 100 mM), then purified *via* spin filter centrifugation using a 15mL Amicon Ultracell 30KDa MWCO spin filter, sterile-filtered and analysed.

UHPLC analysis on a Shimadzu Prominence system using a Thermo Scientific MAbPac 50 mm x 2.1 mm column eluting with a gradient of water and acetonitrile on a reduced sample of ConjA at 214 nm and 330 nm (SG3931 specific) shows a mixture of unconjugated light chains, light chains attached to a single molecule of SG3931, unconjugated heavy chains and heavy chains attached to up to three molecules of SG3931, consistent with a drug-per-antibody ratio (DAR) of 7.32 molecules of SG3931 per antibody.

UHPLC analysis on a Shimadzu Prominence system using a Tosoh Bioscience TSKgel SuperSW mAb HTP 4 μm 4.6 x 150 mm column (with a 4 μm 3.0 x 20 mm guard column) eluting with 0.3 mL/minute sterile-filtered SEC buffer containing 200 mM potassium phosphate pH 6.95, 250 mM potassium chloride and 10% isopropanol (*v/v*) on a sample of ConjA at 280 nm shows a monomer purity of 94.2%. UHPLC SEC analysis gives a

concentration of final ConjA at 1.29 mg/mL in 5.8 mL, obtained mass of ConjA is 7.5 mg (37% yield).

Example 5 – Cytoxicity assay

The potency of the molecules were measured via *in vitro* cytotox assays in the carcinoma cell line NCI-N87.

Solid material was dissolved in DMSO to a 2 mM stock solution, from which eight serial dilutions were made at a 1:10 ratio in DMSO and stored at -20°C until use.

Adherent NCI-N87 cells were washed with D-PBS and detached with Trypsin-EDTA, cell density and viability were then determined in duplicate by Trypan blue exclusion assay using an automated cell counter (LUNA-II™). Cell suspension was diluted to 1 x 10^5 cells/ml in growth media (RPMI 1640 with Glutamax + 10% (v/v) HyClone™ Fetal Bovine Serum) and vortexed before dispensing 2mL per well into sterile 3 mL polypropylene plates. Warhead dilutions were then dispensed into the appropriate wells at 10 µl/well and mixed by repeat pipetting. For control wells 10 µl of DMSO was dispensed onto 2 mL cell suspension, and thoroughly mixed. 100µl of each sample was then aliquoted into 2 replicate wells of a sterile flat 96-well microplate and incubated in a 37°C CO_2-gassed (5%) incubator. At the end of the incubation period time (7 days), cell viability was measured by CellTiter 96 Aqueous One (MTS) assay, which was dispensed at 20µl/well and incubated for 4 hours at 37°C, 5% CO_2. Plates were then read on an EnVision Multi-label Plate Reader (Perkin Elmer) using absorbance at 490 nm.

Cell survival percentage was calculated from the mean absorbance of the 2 replicate wells for each sample, compared to the mean absorbance in the two control wells treated with DMSO only (100%). The IC_{50} was determined by fitting each data set to sigmoidal dose-response curves with a variable slope using the non-linear curve fit algorithm on the GraphPad Prism software (San Diego, CA).

All the experiments in this report were carried out and tested in three independent experiments. Data are reported as the mean of the three independent replicates.

	IC_{50} (nM)
Ex2A	70.11
Ex1	2.202
Ex2B	4.035

Example 6 – ADC Cytoxicity assay

The concentration and viability of cells from a sub-confluent (80-90% confluency) T75 flask are measured by trypan blue staining and counted using the LUNA-II™ Automated Cell Counter. Cells were diluted to $2x10^5$/ml, dispensed (50 µl per well) into 96-well flat-bottom plates.

A stock solution (1 ml) of the test antibody drug conjugate (ADC) (20 µg/ml) was made by dilution of filter-sterilised ADC into cell culture medium. A set of 8x 10-fold dilutions of stock ADC were made in a 24-well plate by serial transfer of 100 µl into 900 µl of cell culture medium. ADC dilution was dispensed (50 µl per well) into 4 replicate wells of the 96-well plate, containing 50 µl cell suspension seeded the previously. Control wells received 50 µl cell culture medium. The 96-well plate containing cells and ADCs was incubated at 37C in a CO_2-gassed incubator for the exposure time.

At the end of the incubation period, cell viability was measured by MTS assay. MTS (Promega) was dispensed (20 µl per well) into each well and incubated for 4 hours at 37°C in the CO_2-gassed incubator. Well absorbance was measured at 490 nm. Percentage cell survival was calculated from the mean absorbance in the 4 ADC-treated wells compared to the mean absorbance in the 4 control untreated wells (100%). IC_{50} was determined from the dose-response data using GraphPad Prism using the non-linear curve fit algorithm: sigmoidal dose-response curve with variable slope.

ADC incubation times were 4 days with MDA-MB-468 and 7 days for NCI-N87. MDA-MB-468 and NCI-N87 were cultured in RPMI 1640 with Glutamax + 10% (v/v) HyClone™ Fetal Bovine Serum.

The EC_{50} values were determined by fitting data to a sigmoidal dose-response curve with variable slope using GraphPad Prism software v6.05 (GraphPad, San Diego, CA).

	EC_{50} (µg/ml)	
	NCI-N87	MDA-MB-468
ConjA	0.002285	15.71

Example 7 – Xenograft testing

NCI-N87 Xenografted Mice

Female severe combined immune-deficient mice (Fox Chase SCID®, C.B-17/Icr-*Prkdcscid*, Charles River) were eight weeks old with a body weight (BW) range of 16.5 to 21.6 grams

on Day 1 of the study. The animals were fed *ad libitum* water (reverse osmosis, 1 ppm Cl), and NIH 31 Modified and Irradiated Lab Diet® consisting of 18.0% crude protein, 5.0% crude fat, and 5.0% crude fibre. The mice were housed on irradiated Enricho'cobs ™ Laboratory Animal Bedding in static micro-isolators on a 12-hour light cycle at 20–22°C (68–72°F) and 40–60% humidity. CR Discovery Services specifically complies with the recommendations of the *Guide for Care and Use of Laboratory Animals* with respect to restraint, husbandry, surgical procedures, feed and fluid regulation, and veterinary care. The animal care and use program at CR Discovery Services is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC), which assures compliance with accepted standards for the care and use of laboratory animals.

Tumour Cell Culture

Human NCI-N87 gastric carcinoma lymphoma cells were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 100 units/mL penicillin G sodium, 100 µg/mL streptomycin sulfate and 25 µg/mL gentamicin. The cells were grown in tissue culture flasks in a humidified incubator at 37 °C, in an atmosphere of 5% CO_2 and 95% air.

In Vivo Implantation and Tumour Growth

The NCI-N87 cells used for implantation were harvested during log phase growth and Re-suspended in phosphate buffered saline (PBS) containing 50% Matrigel™ (BD Biosciences). On the day of tumour implant, each test mouse was injected subcutaneously in the right flank with 1 x 10^7 cells (0.1 mL cell suspension), and tumour growth was monitored as the average size approached the target range of 100 to 150 mm^3. Fourteen days later, designated as Day 1 of the study, mice were sorted according to calculated tumour size into groups each consisting of ten animals with individual tumour volumes ranging from 108 to 144 mm^3 and group mean tumour volumes of 115 mm^3.

Tumours were measured in two dimensions using calipers, and volume was calculated using the formula:

Tumour Volume (mm^3) = $\frac{w^2 x\, l}{2}$

where *w* = width and *l* = length, in mm, of the tumour. Tumour weight may be estimated with the assumption that 1 mg is equivalent to 1 mm^3 of tumour volume.

Treatment

Treatment began on Day 1 in groups of 10 mice (n=10) with established subcutaneous NCI-N87 tumours (108–144 mm^3). ConjA (4 mg/kg) was administered intravenously once on Day 1 (qd x 1). A vehicle-treated group served as the control group for efficacy analysis. Tumours were measured twice per week until the study was ended on Day 79. Each mouse was euthanized when its tumour reached the endpoint volume of 800 mm^3 or on the final day, whichever came first. The time to endpoint (TTE) was calculated for each mouse.

The results are illustrated in Figure 1 which shows the change in normalised tumour growth (■ – control; ♦ - ConjA).

Endpoint and Tumor Growth Delay (TGD) Analysis

Tumors were measured using calipers twice per week, and each animal was euthanized when its tumor reached the endpoint volume of 800 mm^3 or at the end of the study (Day 79), whichever came first. Animals that exited the study for tumor volume endpoint were documented as euthanized for tumor progression (TP), with the date of euthanasia. The time to endpoint (TTE) for analysis was calculated for each mouse by the following equation:

$$\mathrm{TTE} = \frac{\log_{10}(\text{endpoint volume}) - b}{m}$$

where TTE is expressed in days, endpoint volume is expressed in mm^3, b is the intercept, and m is the slope of the line obtained by linear regression of a log-transformed tumor growth data set. The data set consisted of the first observation that exceeded the endpoint volume used in analysis and the three consecutive observations that immediately preceded the attainment of this endpoint volume. The calculated TTE is usually less than the TP date, the day on which the animal was euthanized for tumor size. Animals with tumors that did not reach the endpoint volume were assigned a TTE value equal to the last day of the study (Day 79). In instances in which the log-transformed calculated TTE preceded the day prior to reaching endpoint or exceeded the day of reaching tumor volume endpoint, a linear interpolation was performed to approximate the TTE. Any animal classified as having died from NTR (non-treatment-related) causes due to accident (NTRa) or due to unknown etiology (NTRu) were excluded from TTE calculations (and all further analyses). Animals classified as TR (treatment-related) deaths or NTRm (non-treatment-related death due to metastasis) were assigned a TTE value equal to the day of death. Treatment outcome was evaluated from tumor growth delay (TGD), which is defined as the increase in the median time to endpoint (TTE) in a treatment group compared to the control group:

$$TGD = T - C,$$

expressed in days, or as a percentage of the median TTE of the control group:

$$\%TGD = \frac{T - C}{C} \times 100$$

where:

T = median TTE for a treatment group, and

C = median TTE for the designated control group.

Tumour growth inhibition

Tumor growth inhibition (TGI) analysis evaluates the difference in median tumor volumes (MTVs) of treated and control mice. For this study, the endpoint for determining TGI was Day 19, which was the last day that all evaluable control mice remained in the study. The MTV (n), the median tumor volume for the number of animals, n, on the day of TGI analysis, was determined for each group. Percent tumor growth inhibition (%TGI) was defined as the difference between the MTV of the designated control group and the MTV of the drug-treated group, expressed as a percentage of the MTV of the control group:

$$\%\,TGI = \left(\frac{MTV_{control} - MTV_{drug\text{-}treated}}{MTV_{control}}\right) x\,100 = [1-(MTV_{drug\text{-}treated}/MTV_{control})]\ x\ 100$$

The data set for TGI analysis included all animals in a group, except those that died due to treatment-related (TR) or non-treatment-related (NTR) causes prior to the day of TGI analysis.

MTV and Criteria for Regression Responses

Treatment efficacy may be determined from the tumor volumes of animals remaining in the study on the last day. The MTV (n) was defined as the median tumor volume on the last day of the study in the number of animals remaining (n) whose tumors had not attained the endpoint volume. Treatment efficacy may also be determined from the incidence and magnitude of regression responses observed during the study. Treatment may cause partial regression (PR) or complete regression (CR) of the tumor in an animal. In a PR response, the tumor volume was 50% or less of its Day 1 volume for three consecutive measurements during the course of the study, and equal to or greater than 13.5 mm^3 for one or more of these three measurements. In a CR response, the tumor volume was less than 13.5 mm^3 for three consecutive measurements during the course of the study. Animals were scored only once during the study for a PR or CR event and only as CR if

both PR and CR criteria were satisfied. An animal with a CR response at the termination of a study was additionally classified as a tumor-free survivor (TFS). Animals were monitored for regression responses.

Toxicity

Animals were weighed daily on Days 1–5, then twice per week until the completion of the study. The mice were observed frequently for overt signs of any adverse, treatment-related (TR) side effects, and clinical signs were recorded when observed. Individual body weight was monitored as per protocol, and any animal with weight loss exceeding 30% for one measurement or exceeding 25% for three consecutive measurements was euthanized as a TR death. Group mean body weight loss was also monitored according to CR Discovery Services protocol. Acceptable toxicity was defined as a group mean body weight (BW) loss of less than 20% during the study and no more than 10% TR deaths. Dosing was suspended in any group where mean weight loss exceeded acceptable limits. If group mean body weight recovered to acceptable levels, then dosing was modified to lower levels and/or reduced frequency then resumed. Deaths were classified as TR if it was attributable to treatment side effects as evidenced by clinical signs and/or necropsy. A TR classification was also assigned to deaths by unknown causes during the dosing period or within 14 days of the last dose. A death was classified as non-treatment-related (NTR) if there was no evidence that death was related to treatment side effects. NTR deaths are further categorized as follows: NTRa describes deaths due to accidents or human error; NTRm is assigned to deaths thought to result from tumor dissemination by invasion and/or metastasis based on necropsy results; NTRu describes deaths of unknown causes that lack available evidence of death related to metastasis, tumor progression, accident or human error. It should be noted that treatment side effects cannot be excluded from deaths classified as NTRu.

Statistical and Graphical Analyses

GraphPad Prism 8.0 for Windows was used for all statistical analysis and graphical presentations. Study groups experiencing toxicity beyond acceptable limits (>20% group mean body weight loss or greater than 10% treatment-related deaths) or having fewer than five evaluable observations, were not included in the statistical analysis. The logrank test was employed to assess the significance of the difference between the overall survival experiences of two groups. The logrank test analyzes the individual TTEs for all animals in a group, except those lost to the study due to NTR death. Statistical analyses of the differences between Day 19 median tumor volumes (MTVs) of control and treated groups

were accomplished using the Mann-Whitney *U*-test. For statistical analyses, two-tailed tests were conducted at significance level $P = 0.05$. Prism summarizes test results as not significant (ns) at $P > 0.05$, significant (symbolized by "*") at $0.01 < P \leq 0.05$, very significant ("**") at $0.001 < P \leq 0.01$, and extremely significant ("***") at $P \leq 0.001$. Because tests of statistical significance do not provide an estimate of the magnitude of the difference between groups, all levels of significance were described as either significant or not significant within the text of this report.

		n	Median TTE	T-C	%TGD	MTV (n), Day 79
Vehicle	-	10	24.8	--	-	466(10)
ConjA	4 mg/kg	10	79.0	54.2	219	32 (9)

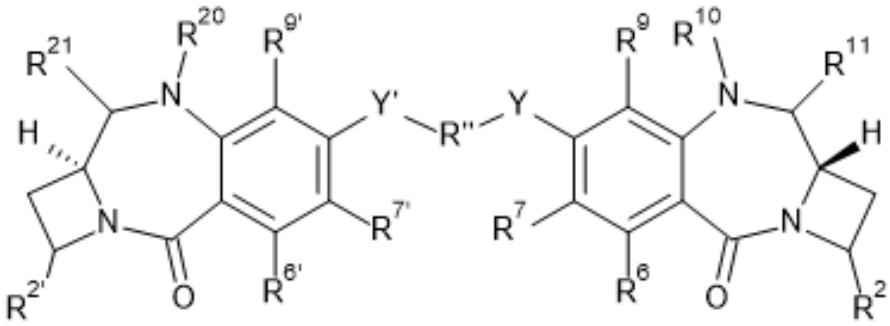
	PR	CR	TFS	BW Nadir	TR	NTRm	NTR
Vehicle	0	0	0	-2.0 (2)	0	0	0
ConjA	6	4	0	-1.9 (2)	0	0	0

The Day 19 MTV(10) for animals treated with ConjA was 32 mm^3, or a significant 93% TGI ($P < 0.001$, Mann-Whitney). Nine animals survived the study and the assigned median TTE was 79.0 days; this represents the maximally possible, significant 219% TGD ($P < 0.001$, logrank). The MTV(9) on Day 79 was 320 mm^3 and there were six PRs and four CRs.

All documents and other references mentioned above are herein incorporated by reference.

STATEMENTS OF INVENTION

1. A compound of formula **IV**:



IV

and salts and solvates thereof, wherein:

R^2 and $R^{2'}$ are H;

R^6 and R^9 are independently selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo;

where R and R' are independently selected from optionally substituted C_{1-12} alkyl, C_{3-20} heterocyclyl and aromatic ring systems comprising 5-20 ring atoms~~C_{5-20} aryl groups~~;

either

(a) R^7 is selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo;

$R^{7'}$ is selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo; or

(b) R^7 and $R^{7'}$ together form a group which is: (i) $-O-(CH_2)_n-O-$, where n is from 7 to 16; or (ii) $-O-(CH_2CH_2O)_m-$, where m is 2 to 5;

R″ is a C_{3-12} alkylene group, which chain may be interrupted by one or more heteroatoms~~, e.g.~~ selected from O, S, NR^{N2} ~~(~~where R^{N2} is H or C_{1-4} alkyl~~)~~, and/or by aromatic rings~~, e.g.~~ selected from benzene or pyridine;

Y and Y' are selected from O, S, or NH;

$R^{6'}$ and $R^{9'}$ are selected from the same groups as R^6, and R^9 respectively;

either

(i-a) R^{10} and R^{11} together form a double bond between the N and C atoms to which they are bound; or

(i-b) R^{10} is H and R^{11} is selected from OH and OR^A, where R^A is C_{1-4} alkyl; or

(i-c) R^{10} and R^{11} are both H;

either

(ii-a) R^{20} and R^{21} together form a double bond between the N and C atoms to which they are bound; or

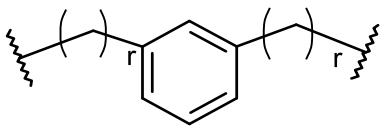
(ii-b) R^{20} is H and R^{21} is selected from OH and OR^B, where R^B is C_{1-4} alkyl; or

(ii-c) R^{20} and R^{21} are both H.

2. A compound according to statement 1, wherein both Y and Y' are O.

3. A compound according to either statement 1 or statement 2, wherein R'' is C_{3-7} alkylene.

4. A compound according to either statement 1 or statement 2, wherein R" is a group of formula:



where r is 1 or 2.

5. A compound according to any one of statements 1 to 4, wherein R^9 is H.

6. A compound according to any one of statements 1 to 5, wherein R^6 is H.

7. A compound according to any one of statements 1 to 6, wherein R^7 is selected from H, OH and OR and $R^{7'}$ is selected from H, OH and OR

8. A compound according to statement 7, wherein R^7 is a C_{1-4} alkyloxy group and $R^{7'}$ is a C_{1-4} alkyloxy group.

9. A compound according to any one of statements 1 to 8, wherein $R^{2'}$ is the same as R^2, $R^{6'}$ is the same group as R^6, $R^{7'}$ is the same group as R^7, $R^{9'}$ is the same group as R^9 and Y' is the same group as Y.

10. A compound according to any one of statements 1 to 9, wherein R^{10} and R^{11} together form a double bond between the N and C atoms to which they are bound.

11. A compound according to any one of statements 1 to 9, wherein R^{10} is H and R^{11} is selected from OH and OR^A.

12. A compound according to statement 11, wherein R^A is methyl.

13. A compound according to any one of statements 1 to 9, wherein R^{10} and R^{11} are both H.

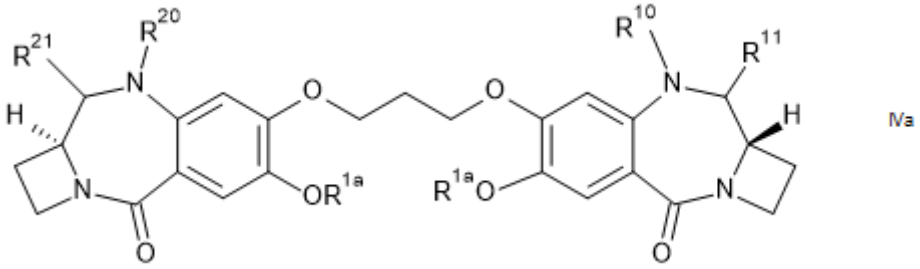
14. A compound according to any one of statements 1 to 13, wherein R^{20} and R^{21} together form a double bond between the N and C atoms to which they are bound.

15. A compound according to any one of statements 1 to 13, wherein R^{20} is H and R^{21} is selected from OH and OR^B.

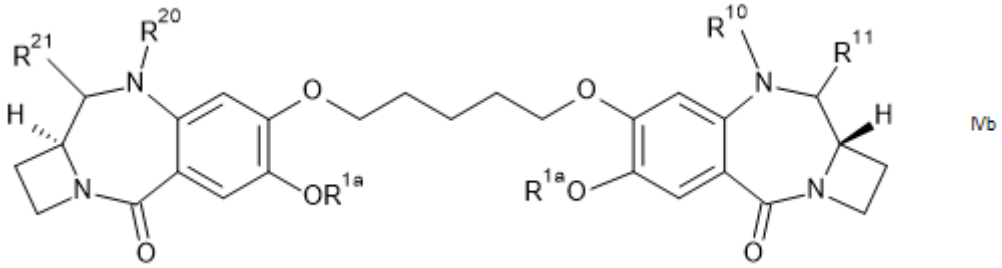
16. A compound according to statement 14, wherein R^B is methyl.

17. A compound according to any one of statements 1 to 13, wherein R^{20} and R^{21} are both H.

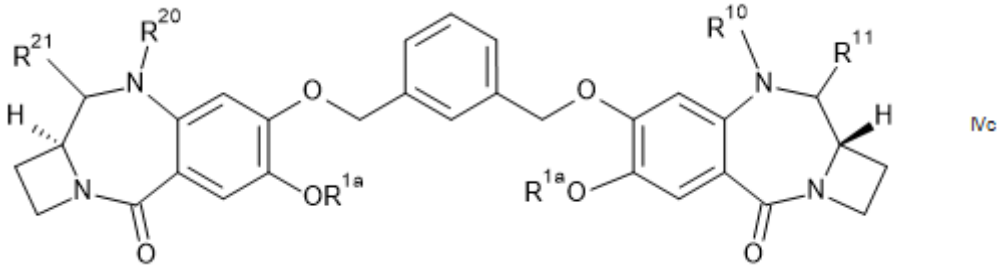
18. A compound according to statement 1 which is of formulae IVa, IVb or IVc:



IVa



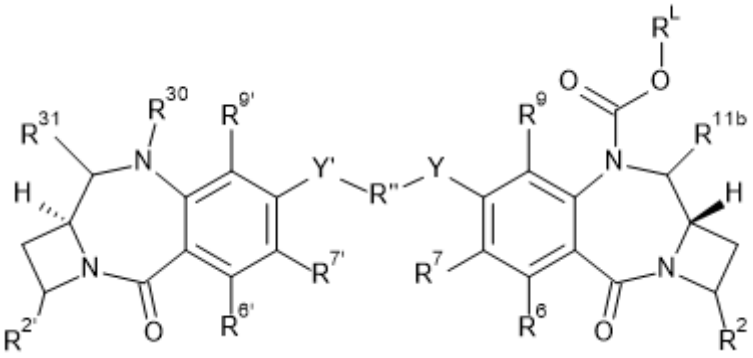
IVb



IVc

where R^{1a} is selected from methyl and benzyl.

19. A compound of formula **I**:



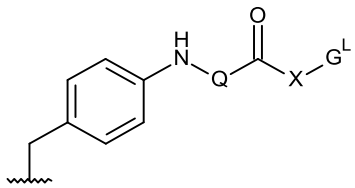
and salts and solvates thereof, wherein:

Y, Y', R", R^2, $R^{2'}$, R^6, $R^{6'}$, R^7, $R^{7'}$, R^9 and $R^{9'}$ are as defined in any one of statements 1 to 18;

R^{11b} is selected from OH, OR^A, where R^A is C_{1-4} alkyl; and

R^L is a linker for connection to a cell binding agent, which is selected from:

(iiia):

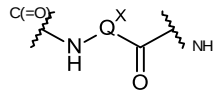


IIIa

,

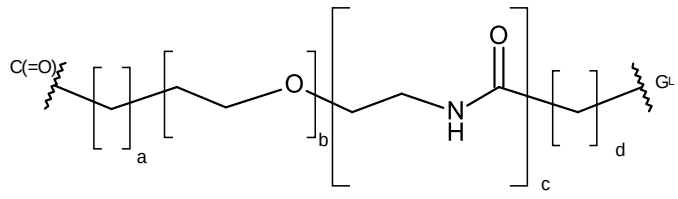
wherein

Q is:



, where Q^X is such that Q is an amino-acid residue, a dipeptide residue or a tripeptide residue;

X is:

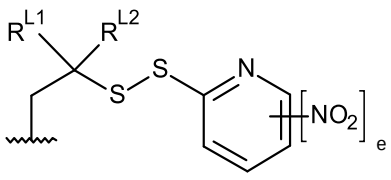


,

where a = 0 to 5, b = 0 to 16, c = 0 or 1, d = 0 to 5;

G^L is a linker for connecting to a cell binding agent ~~Ligand Unit~~; and

(iiib):



IIIb

,

where R^{L1} and R^{L2} are independently selected from H and methyl, or together with the carbon atom to which they are bound form a cyclopropylene or cyclobutylene group;

and e is 0 or 1;

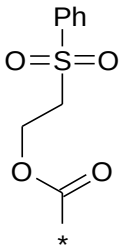
either:

(a) R^{30} and R^{31} together form a double bond between the N and C atoms to which they are bound; or

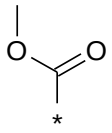
(b) R^{30} is H and R^{31} is selected from OH and OR^B, where R^B is C_{1-4} alkyl; or

(c) R^{30} and R^{31} are both H; or

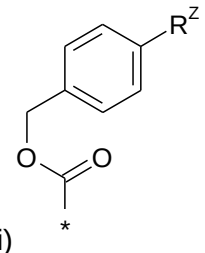
(d) R^{31} is OH or OR^B, where R^B is C_{1-4} alkyl and R^{30} is selected from:

(i) 

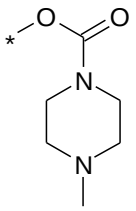
;

(ii) 

;

(iii) 

, where R^Z is selected from:



(z-i) ;

(z-ii) $OC(=O)CH_3$;

(z-iii) NO_2;

(z-iv) OMe;

(z-v) glucoronide;

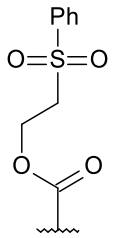
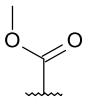
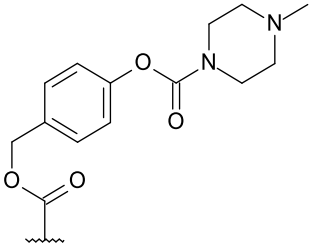
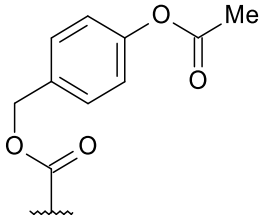
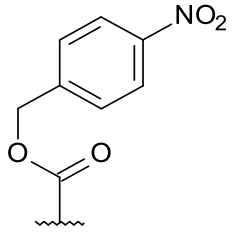
(z-vi) $NH\text{-}C(=O)\text{-}X_1\text{-}NHC(=O)X_2\text{-}NH\text{-}C(=O)\text{-}R^{ZC}$, where $-C(=O)-X_1-$ NH- and $-C(=O)-X_2-$NH- represent natural amino acid residues and R^{ZC} is selected from Me, OMe, CH_2CH_2OMe, and $(CH_2CH_2O)_2Me$.

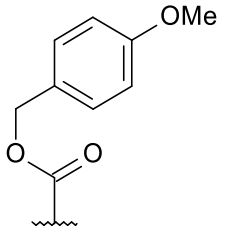
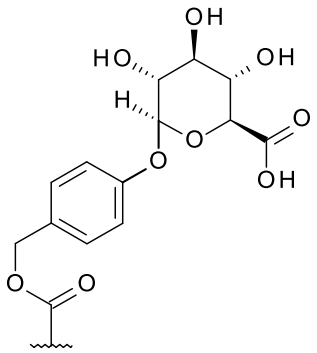
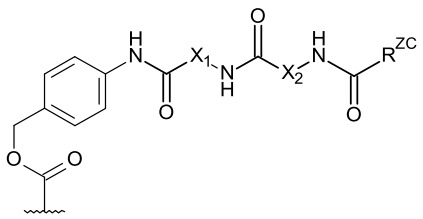
20. A compound according to statement 19, wherein R^{30} and R^{31} together form a double bond between the N and C atoms to which they are bound.

21. A compound according to statement 19, wherein R^{30} is H and R^{31} is selected from OH and OR^B, where R^B is C_{1-4} alkyl.

22. A compound according to statement 19, wherein R^{30} and R^{31} are both H.

23. The compound according to statement 19, wherein R^{31} is OH or OR^A and R^{30} is selected from:

R^{30a}	
R^{30b}	
R^{30c}	
R^{30d}	
R^{30e}	

R^{30f}	
R^{30g}	
R^{30h}	

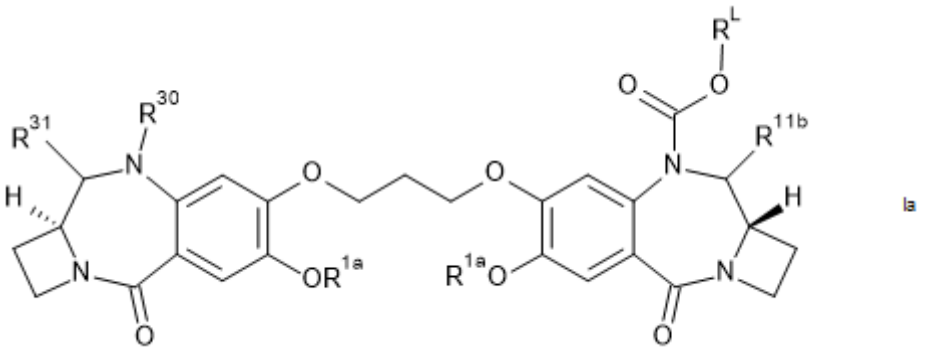
24. The compound according to statement 23, wherein -C(=O)-X_1-NHC(=O)X_2-NH-, is selected from: -Phe-Lys-, -Val-Ala-, -Val-Lys-, -Ala-Lys-, and -Val-Cit-.

25. The compound according to statement 23, wherein -C(=O)-X_1-NHC(=O)X_2-NH-, is selected from: -Phe-Lys-, and -Val-Ala-.

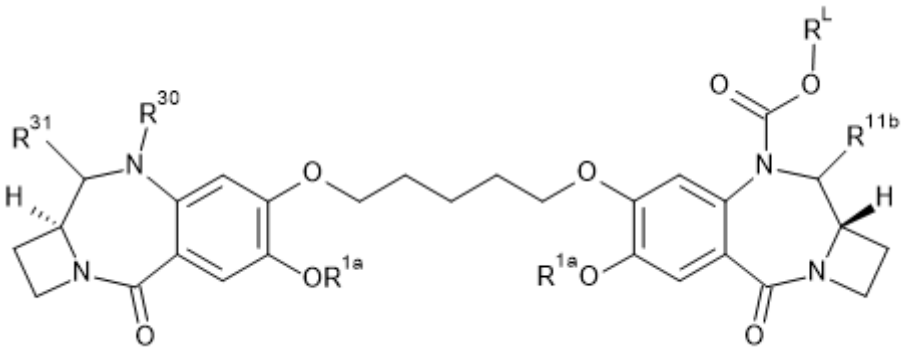
26. The compound according to any one of statements 23 to 25 wherein R^{ZC} is selected from CH_2CH_2OMe, and $(CH_2CH_2O)_2Me$.

27. The compound according to statement 26 wherein R^{ZC} is $(CH_2CH_2O)_2Me$.

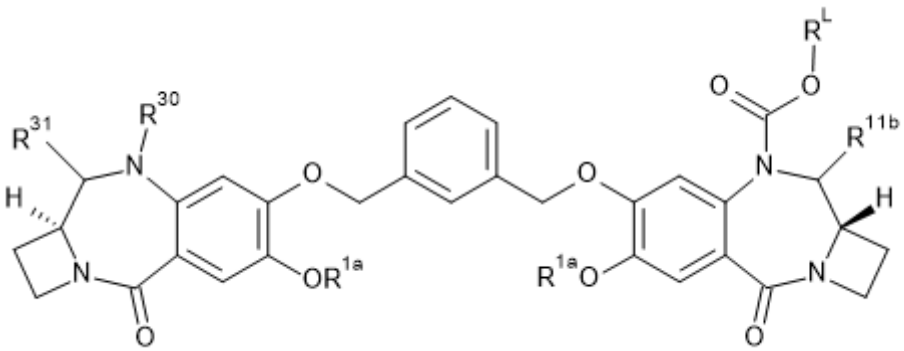
28. A compound according to statement 19, which is of formulae Ia, Ib or Ic:



Ia



Ib



Ic

where R^{1a} is selected from methyl and benzyl.

29. A compound according to any one of statements 19 to 28, wherein R^{11b} is OH.

30. A compound according to any one of statements 19 to 29, wherein R^{11b} is OR^{A}, where R^{A} is C_{1-4} alkyl.

31. A compound according to statement 30, wherein R^{A} is methyl.

32. A compound according to any one of statements 19 to 31, wherein R^{L} is of formula IIIa, and Q is an amino acid residue selected from Phe, Lys, Val, Ala, Cit, Leu, Ile, Arg, and Trp.

33. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIa, and Q is a dipeptide residue selected from:

^{CO}-Phe-Lys-^{NH},

^{CO}-Val-Ala-^{NH},

^{CO}-Val-Lys-^{NH},

^{CO}-Ala-Lys-^{NH},

^{CO}-Val-Cit-^{NH},

^{CO}-Phe-Cit-^{NH},

^{CO}-Leu-Cit-^{NH},

^{CO}-Ile-Cit-^{NH},

^{CO}-Phe-Arg-^{NH}, and

^{CO}-Trp-Cit-^{NH}.

34. A compound according to statement 33, wherein Q is selected from ^{CO}-Phe-Lys-^{NH}, ^{CO}-Val-Cit-^{NH} and ^{CO}-Val-Ala-^{NH}.

35. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIa, and Q is a tripeptide residue selected from:

^{CO}-Glu-Val-Ala-^{NH},

^{CO}-Glu-Val-Cit-^{NH},

^{CO}-αGlu-Val-Ala-^{NH}, and

^{CO}-αGlu-Val-Cit-^{NH}.

36. A compound according to any one of statements 19 to 35, wherein R^L is of formula IIIa and a is 0 to 3.

37. A compound according to statement 36, wherein a is 0.

38. A compound according to any one of statements 19 to 37, wherein R^L is of formula IIIa and b is 0 to 12.

39. A compound according to statement 38, wherein b is 0 to 8.

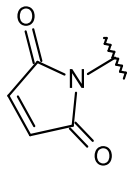
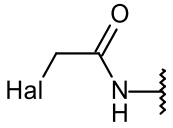
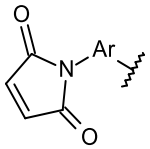
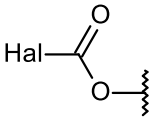
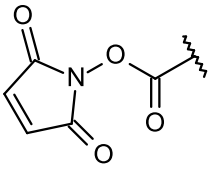
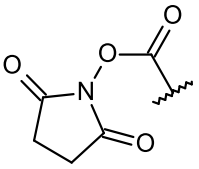
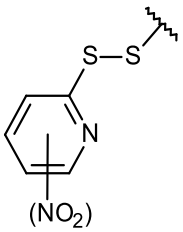
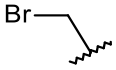
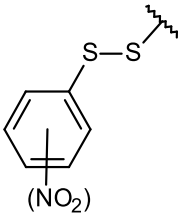
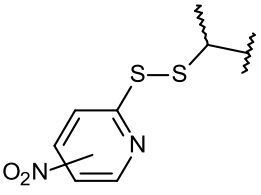
40. A compound according to any one of statements 19 to 39, wherein R^L is of formula IIIa and d is 0 to 3.

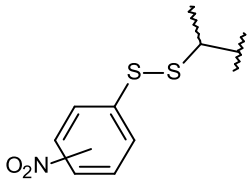
41. A compound according to statement 38, wherein d is 2.

42. A compound according to any one of statements 19 to 35, wherein R^L is of formula IIIa and, a is 0, c is 1 and d is 2, and b is from 0 to 8.

43. A compound according to statement 42, wherein b is 0, 4 or 8.

44. A compound according to any one of statements 19 to 43 wherein R^L is of formula IIIa and G^L is selected from:

(G^{L1-1})		(G^{L4})	 Where Hal = I, Br, Cl
(G^{L1-2})		(G^{L5})	
(G^{L2})		(G^{L6})	
(G^{L3-1})	 where the NO_2 group is optional	(G^{L7})	
(G^{L3-2})	 where the NO_2 group is optional	(G^{L8})	
(G^{L3-3})		(G^{L9})	

	where the NO_2 group is optional		
(G^{L3-4})	 where the NO_2 group is optional		

where Ar represents a C_{5-6} arylene group.

45. A compound according to statement 44, wherein Ar is a phenylene group.

46. A compound according to either statement 44 or statement 45, wherein G^L is selected from G^{L1-1} and G^{L1-2}.

47. A compound according to statement 46, wherein G^L is G^{L1-1}.

48. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIb, and both R^{L1} and R^{L2} are H.

49. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIb, R^{L1} is H and R^{L2} is methyl.

50. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIb, and both R^{L1} and R^{L2} are methyl.

51. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIb, and, R^{L1} and R^{L2} together with the carbon atom to which they are bound form a cyclopropylene group.

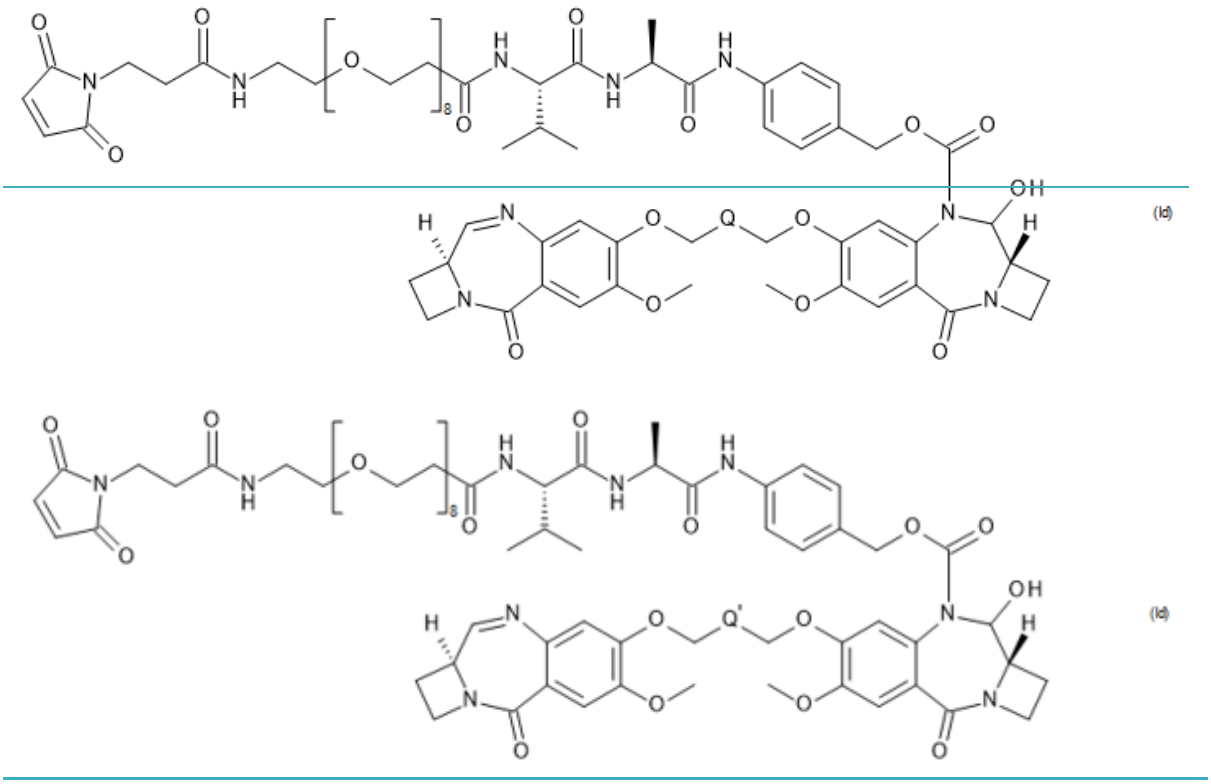
52. A compound according to any one of statements 19 to 31, wherein R^L is of formula IIIb, and, R^{L1} and R^{L2} together with the carbon atom to which they are bound form a cyclobutylene group.

53. A compound according to any one of statements 19 to 31 and 48 to 52, wherein R^L is of formula IIIb, and e is 0.

54. A compound according to any one of statements 19 to 31 and 48 to 52, wherein R^L is of formula IIIb, and e is 1.

55. A compound according to statement 54, wherein the nitro group is in the para position.

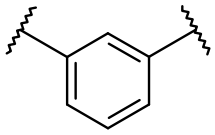
56. A compound according to statement 19, wherein the compound is of formula Id:



where Q' is selected from:

(a) $-CH_2-$;

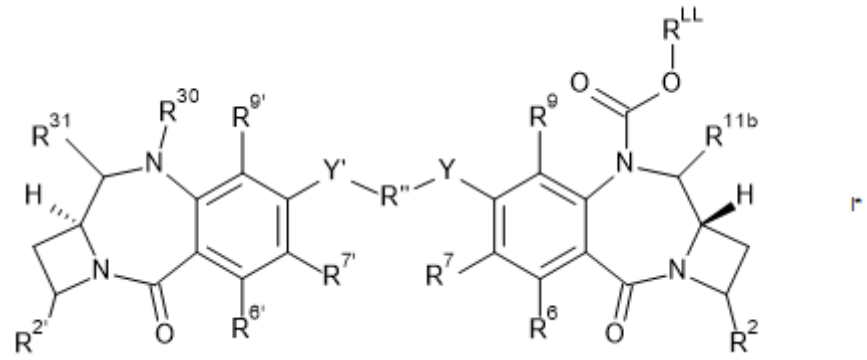
(b) $-C_3H_6-$; and

(c)  .

57. A conjugate of formula II:

$$L - (D^L)_p \qquad (II)$$

wherein L is a Ligand unit (i.e., a targeting agent), D^L is a Drug Linker unit of formula I'*:



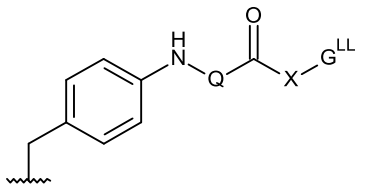
wherein:

Y, Y', R", R^2, $R^{2'}$, R^6, $R^{6'}$, R^7, $R^{7'}$, R^9 and $R^{9'}$ are as defined in any one of statements 1 to 18;

R^{11b}, R^{30} and R^{31} are as defined in any one of statements 19 to 27 and 29 to 31;

R^{LL} is a linker for connection to a cell binding agent, which is selected from:

(iiia):

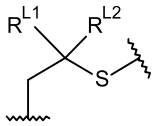


IIIa'

,

where Q and X are as defined in any one of statements 19 and 32 to 43 and G^{LL} is a linker connected to a Ligand Unit; and

(iiib):



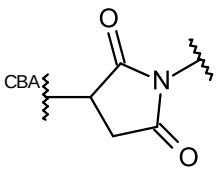
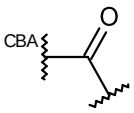
IIIb'

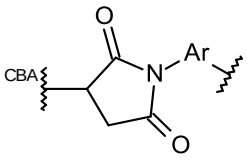
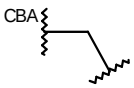
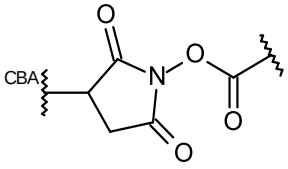
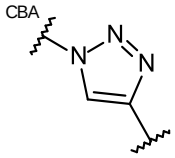
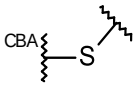
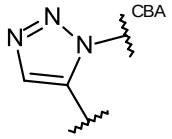
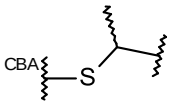
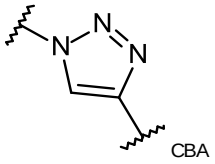
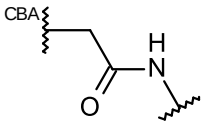
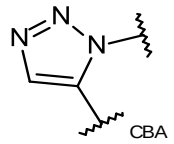
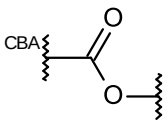
,

where R^{L1} and R^{L2} are as defined in any one of statements 19 and 48 to 52;

wherein p is an integer of from 1 to 20.

58. A conjugate according to statement 57, wherein G^{LL} is selected from:

(G^{LL1-1})		(G^{LL6})	

(G^{LL1-2})		(G^{LL7})	
(G^{LL2})		(G^{LL8-1})	
(G^{LL3-1})		(G^{LL8-2})	
(G^{LL3-2})		(G^{LL9-1})	
(G^{LL4})		(G^{LL9-2})	
(G^{LL5})			

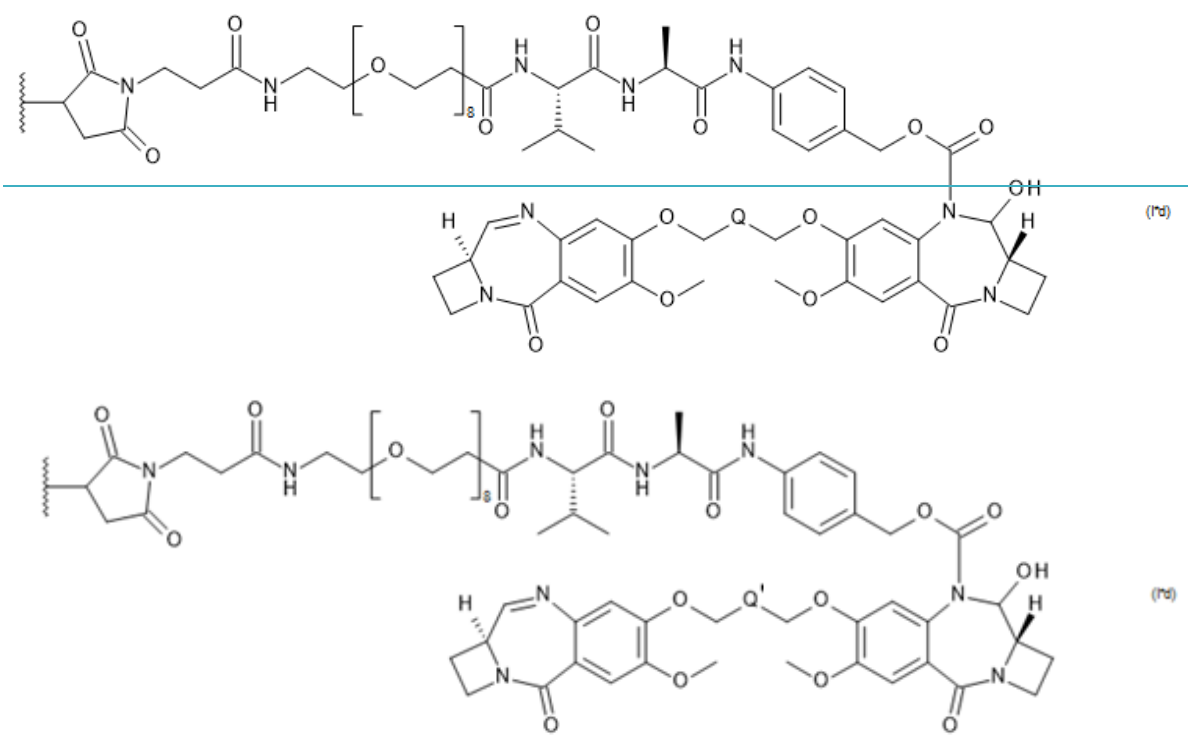
where Ar represents a C_{5-6} arylene group.

59. A conjugate according to statement 58, wherein Ar is a phenylene group.

60 A conjugate according to either statement 58 or statement 59, wherein G^{LL} is selected from G^{LL1-1} and G^{LL1-2}.

61. A conjugate according to statement 60, wherein G^{LL} is G^{LL1-1}.

62. A conjugate according to statement 57, wherein D^{L} is of formula (I*d1):



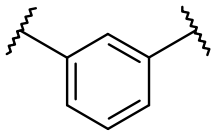
(IId)

(IId)

where Q' is selected from:

(a) $-CH_2-$;

(b) $-C_3H_6-$; and

(c) .

63. A conjugate according to any one of statements 57 to 62, wherein the Ligand Unit is an antibody or an active fragment thereof.

64. The conjugate according to statement 63, wherein the antibody or antibody fragment is an antibody or antibody fragment for a tumour-associated antigen.

65. The conjugate according to statement 64, wherein the antibody or antibody fragment is an antibody which binds to one or more tumor-associated antigens or cell-surface receptors selected from (1)-(88):

(1) BMPR1B;

(2) E16;

(3) STEAP1;

(4) 0772P;

(5) MPF;

(6) Napi3b;

(7) Sema 5b;

(8) PSCA hlg;

(9) ETBR;

(10) MSG783;

(11) STEAP2;

(12) TrpM4;

(13) CRIPTO;

(14) CD21;

(15) CD79b;

(16) FcRH2;

(17) HER2;

(18) NCA;

(19) MDP;

(20) IL20R-alpha;

(21) Brevican;

(22) EphB2R;

(23) ASLG659;

(24) PSCA;

(25) GEDA;

(26) BAFF-R;

(27) CD22;

(28) CD79a;

(29) CXCR5;

(30) HLA-DOB;

(31) P2X5;

(32) CD72;

(33) LY64;

(34) FcRH1;

(35) IRTA2;

(36) TENB2;

(37) PSMA – FOLH1;

(38) SST;

(38.1) SSTR2;

(38.2) SSTR5;

(38.3) SSTR1;

(38.4)SSTR3;

(38.5) SSTR4;

(39) ITGAV;

(40) ITGB6;

(41) CEACAM5;

(42) MET;

(43) MUC1;

(44) CA9;

(45) EGFRvIII;

(46) CD33;

(47) CD19;

(48) IL2RA;

(49) AXL;

(50) CD30 - TNFRSF8;

(51) BCMA - TNFRSF17;

(52) CT Ags – CTA;

(53) CD174 (Lewis Y) - FUT3;

(54) CLEC14A;

(55) GRP78 – HSPA5;

(56) CD70;

(57) Stem Cell specific antigens;

(58) ASG-5;

(59) ENPP3;

(60) PRR4;

(61) GCC – GUCY2C;

(62) Liv-1 – SLC39A6;

(63) 5T4;

(64) CD56 – NCMA1;

(65) CanAg;

(66) FOLR1;

(67) GPNMB;

(68) TIM-1 – HAVCR1;

(69) RG-1/Prostate tumor target Mindin – Mindin/RG-1;

(70) B7-H4 – VTCN1;

(71) PTK7;

(72) CD37;

(73) CD138 – SDC1;
(74) CD74;
(75) Claudins – CLs;
(76) EGFR;
(77) Her3;
(78) RON - MST1R;
(79) EPHA2;
(80) CD20 – MS4A1;
(81) Tenascin C – TNC;
(82) FAP;
(83) DKK-1;
(84) CD52;
(85) CS1 - SLAMF7;
(86) Endoglin – ENG;
(87) Annexin A1 – ANXA1;
(88) V-CAM (CD106) - VCAM1;
(89) ASCT2 (SLC1A5).

66. The conjugate of any one of statements 63 to 65 wherein the antibody or antibody fragment is a cysteine-engineered antibody.

67. The conjugate according to any one of statements 57 to 66 wherein p is an integer from 1 to 8.

68. The conjugate according to statement 67, wherein p is 1, 2, 3, or 4.

69. A composition comprising a mixture of conjugates according to any one of statements 57 to 68, wherein the average p in the mixture of conjugate compounds is ~~about~~ 1 to ~~about~~ 8.

70. The conjugate according to any one of statements 57 to 68, for use in therapy.

71. A pharmaceutical composition comprising the conjugate of any one of statements 57 to 68, and a pharmaceutically acceptable diluent, carrier or excipient.

72. The conjugate according to any one of statements 57 to 68 or the pharmaceutical composition according to statement 71, for use in the treatment of a proliferative disease in a subject.

73. The conjugate for use according to statement 72, wherein the disease treated is cancer.

74. Use of a conjugate according to any one of statements 57 to 68 or the pharmaceutical composition according to statement 71 in a method of medical treatment.

75. A method of medical treatment comprising administering to a patient the pharmaceutical composition of statement 71.

76. The method of statement 75 wherein the method of medical treatment is for treating cancer.

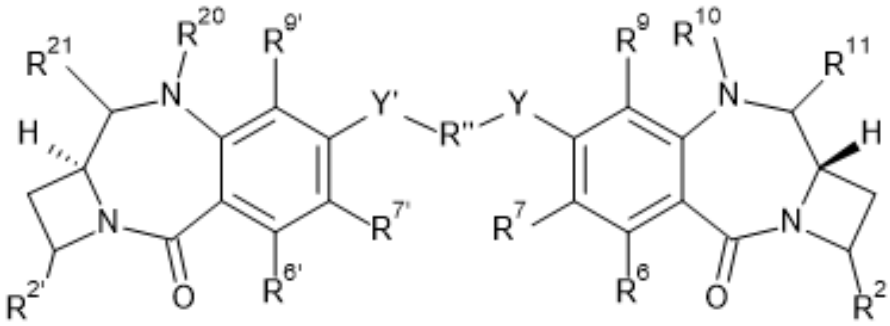
77. The method of statement 76, wherein the patient is administered a chemotherapeutic agent, in combination with the conjugate.

78. Use of a conjugate according to any one of statements 57 to 68 in a method of manufacture of a medicament for the treatment of a proliferative disease.

79. A method of treating a mammal having a proliferative disease, comprising administering an effective amount of a conjugate according to any one of statements 57 to 68 or the pharmaceutical composition according to statement 71.

CLAIMS

1. A compound of formula **IV**:



IV

and salts and solvates thereof, wherein:

R^2 and $R^{2'}$ are H;

R^6 and R^9 are independently selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo;

where R and R' are independently selected from optionally substituted C_{1-12} alkyl, C_{3-20} heterocyclyl and ~~C_{5-20} aryl groups~~ aromatic ring systems comprising 5-20 ring atoms;

either

(a) R^7 is selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo; $R^{7'}$ is selected from H, R, OH, OR, SH, SR, NH_2, NHR, NRR', nitro, Me_3Sn and halo; or

(b) R^7 and $R^{7'}$ together form a group which is: (i) $-O-(CH_2)_n-O-$, where n is from 7 to 16; or (ii) $-O-(CH_2CH_2O)_m-$, where m is 2 to 5;

R″ is a C_{3-12} alkylene group, which chain may be interrupted by one or more heteroatoms~~,~~ ~~e.g.~~ selected from O, S, NR^{N2} ~~(~~where R^{N2} is H or C_{1-4} alkyl~~)~~, and/or by aromatic rings~~,~~ ~~e.g.~~ selected from benzene or pyridine;

Y and Y' are selected from O, S, or NH;

$R^{6'}$ and $R^{9'}$ are selected from the same groups as R^6, and R^9 respectively;

either

(i-a) R^{10} and R^{11} together form a double bond between the N and C atoms to which they are bound; or

(i-b) R^{10} is H and R^{11} is selected from OH and OR^A, where R^A is C_{1-4} alkyl; or

(i-c) R^{10} and R^{11} are both H;

either

(ii-a) R^{20} and R^{21} together form a double bond between the N and C atoms to which they are bound; or

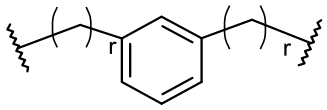
(ii-b) R^{20} is H and R^{21} is selected from OH and OR^B, where R^B is C_{1-4} alkyl; or

(ii-c) R^{20} and R^{21} are both H.

2. A compound according to claim 1, wherein both Y and Y' are O.

3. A compound according to either claim 1 or claim 2, wherein R'' is C_{3-7} alkylene.

4. A compound according to either claim 1 or claim 2, wherein R" is a group of formula:

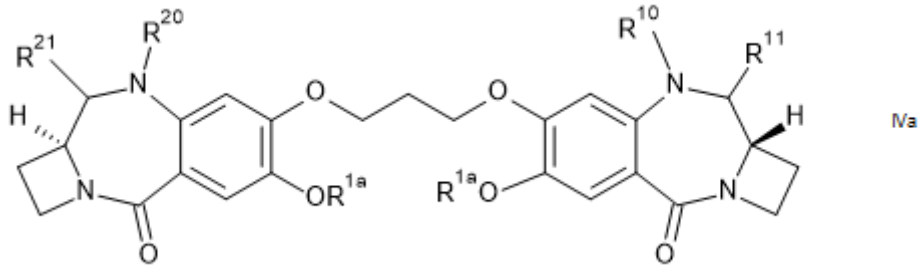


where r is 1 or 2.

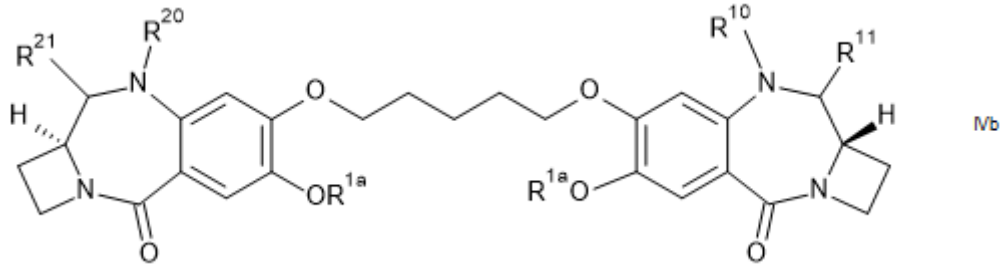
5. A compound according to any one of claims 1 to 4, wherein R^9 is H, R^6 is H, R^7 and $R^{7'}$ are independently a C_{1-4} alkyloxy group.

6. A compound according to any one of claims 1 to 4, wherein $R^{6'}$ is the same group as R^6, $R^{7'}$ is the same group as R^7, $R^{9'}$ is the same group as R^9 and Y' is the same group as Y.

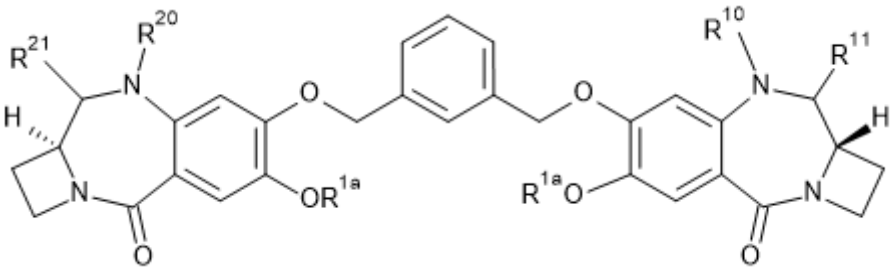
7. A compound according to claim 1 which is of formulae IVa, IVb or IVc:



IVa



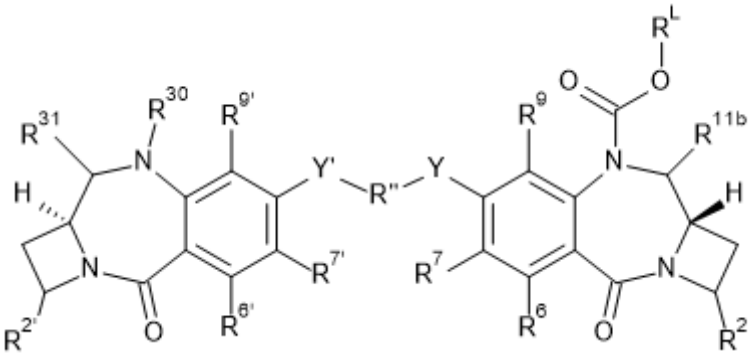
IVb



IVc

where R^{1a} is selected from methyl and benzyl.

8. A compound of formula **I**:



I

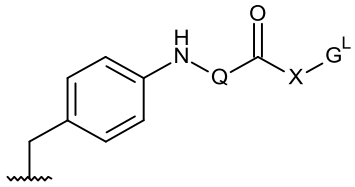
and salts and solvates thereof, wherein:

Y, Y', R", R^2, $R^{2'}$, R^6, $R^{6'}$, R^7, $R^{7'}$, R^9 and $R^{9'}$ are as defined in any one of claims 1 to 9;

R^{11b} is selected from OH, OR^A, where R^A is C_{1-4} alkyl; and

R^L is a linker for connection to a cell binding agent, which is selected from:

(iiia):

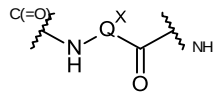


IIIa

,

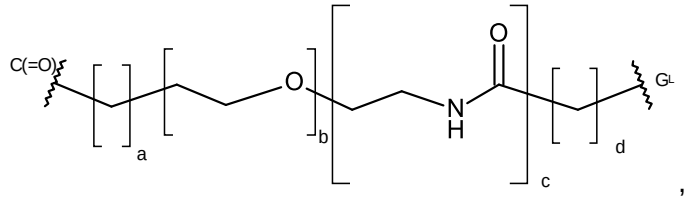
wherein

Q is:



, where Q^X is such that Q is an amino-acid residue, a dipeptide residue or a tripeptide residue;

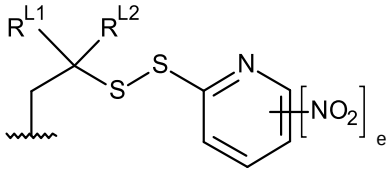
X is:



where a = 0 to 5, b = 0 to 16, c = 0 or 1, d = 0 to 5;

G^L is a linker for connecting to a ~~Ligand Unit~~cell binding agent; and

(iiib):



IIIb

where R^{L1} and R^{L2} are independently selected from H and methyl, or together with the carbon atom to which they are bound form a cyclopropylene or cyclobutylene group;

and e is 0 or 1;

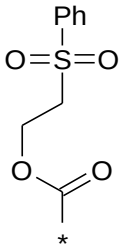
either:

(a) R^{30} and R^{31} together form a double bond between the N and C atoms to which they are bound; or

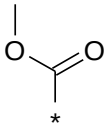
(b) R^{30} is H and R^{31} is selected from OH and OR^B, where R^B is C_{1-4} alkyl;~~,~~

(c) R^{30} and R^{31} are both H; or

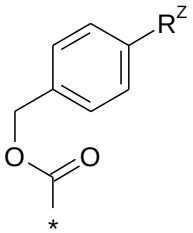
(d) R^{31} is OH or OR^B, where R^B is C_{1-4} alkyl and R^{30} is selected from:



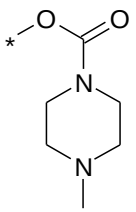
(i) ;



(ii) ;



(iii) , where R^Z is selected from:



(z-i) ;

(z-ii) $OC(=O)CH_3$;

(z-iii) NO_2;

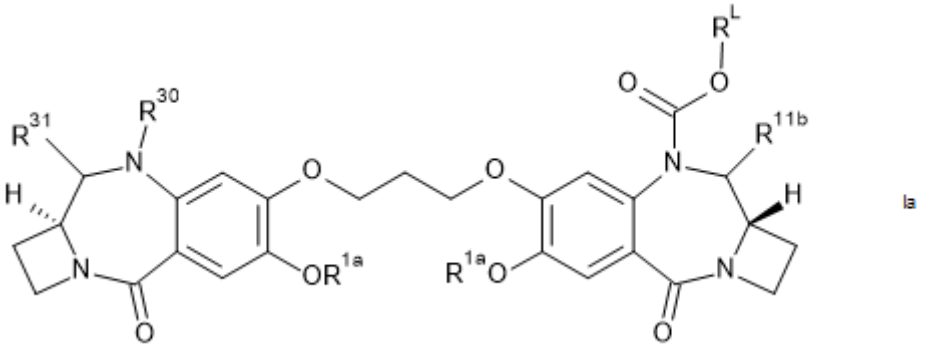
(z-iv) OMe;

(z-v) glucoronide;

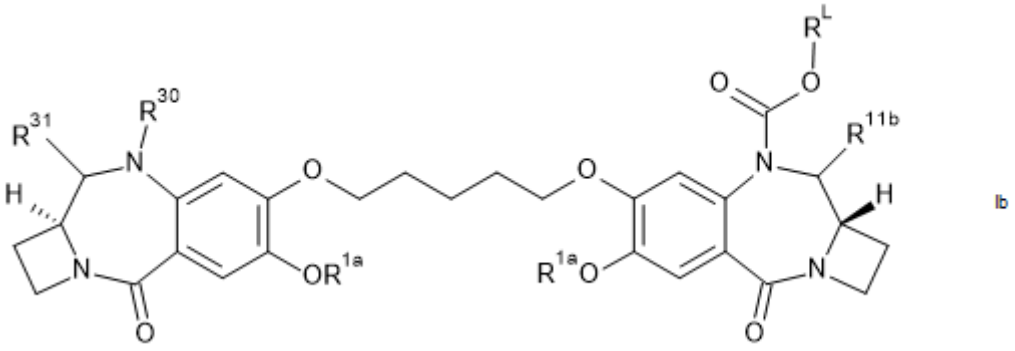
(z-vi) $NH\text{-}C(=O)\text{-}X_1\text{-}NHC(=O)X_2\text{-}NH\text{-}C(=O)\text{-}R^{ZC}$, where $-C(=O)-X_1-NH-$ and $-C(=O)-X_2-NH-$ represent natural amino acid residues and R^{ZC} is selected from Me, OMe, CH_2CH_2OMe, and $(CH_2CH_2O)_2Me$.

9. A compound according to claim 8, wherein R^{30} and R^{31} together form a double bond between the N and C atoms to which they are bound.

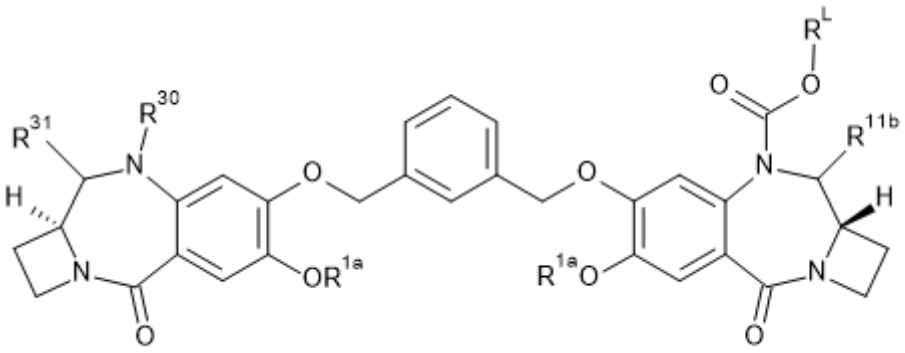
10. A compound according to claim 8, which is of formula Ia, Ib or Ic:



Ia



Ib



Ic

where R^{1a} is selected from methyl and benzyl.

11. A compound according to any one of claims 8 to 10, wherein R^L is of formula IIIa, and Q is a dipeptide residue selected from:

^{CO}-Phe-Lys-^{NH},

^{CO}-Val-Ala-^{NH},

^{CO}-Val-Lys-^{NH},

^{CO}-Ala-Lys-^{NH},

^{CO}-Val-Cit-^{NH},

^{CO}-Phe-Cit-^{NH},

^{CO}-Leu-Cit-^{NH},

^{CO}-Ile-Cit-^{NH},

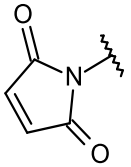
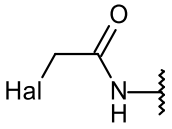
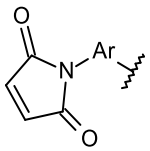
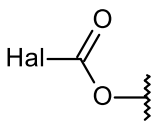
^{CO}-Phe-Arg-^{NH}, and

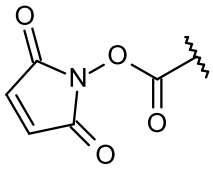
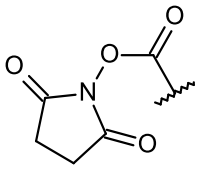
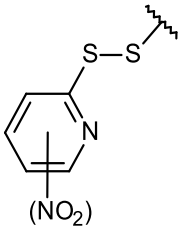
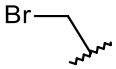
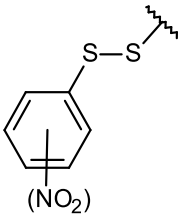
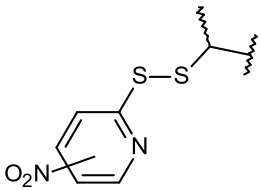
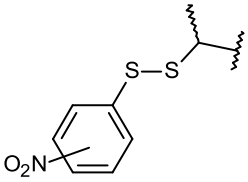
^{CO}-Trp-Cit-^{NH}.

12. A compound according to any one of claims 8 to 10 wherein R^L is of formula IIIa and, a is 0, c is 1 and d is 2, and b is from 0 to 8.

13. A compound according to claim 12, wherein b is 0, 4 or 8.

14. A compound according to any one of claims 8 to 13, wherein R^L is of formula IIIa and G^L is selected from:

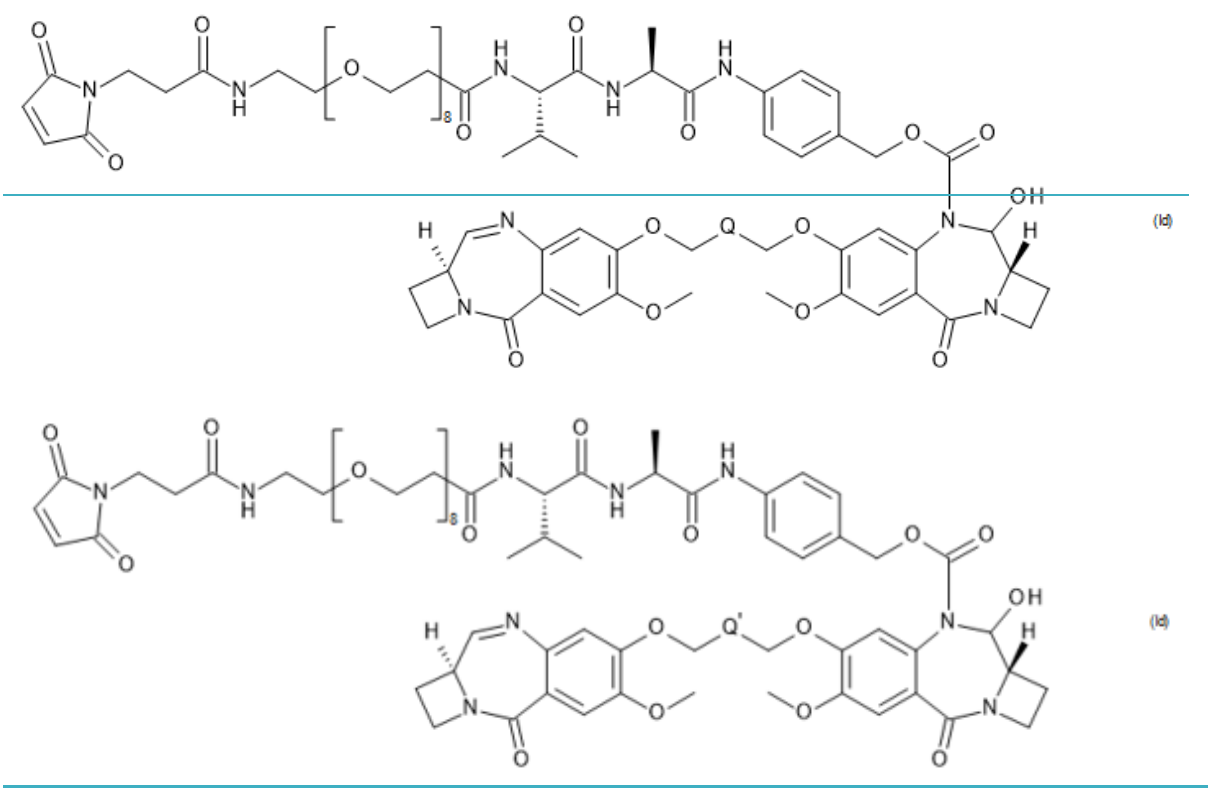
(G^{L1-1})		(G^{L4})	 Where Hal = I, Br, Cl
(G^{L1-2})		(G^{L5})	

(G^{L2})		(G^{L6})	
(G^{L3-1})	 where the NO_2 group is optional	(G^{L7})	
(G^{L3-2})	 where the NO_2 group is optional	(G^{L8})	
(G^{L3-3})	 where the NO_2 group is optional	(G^{L9})	
(G^{L3-4})	 where the NO_2 group is optional		

where Ar represents phenylene~~a C_{5-6} arylene group~~.

15. A compound according to claim 14, wherein G^L is G^{L1-1}.

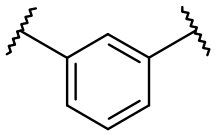
16. A compound according to claim 8, wherein the compound is of formula Id:



where Q' is selected from:

(a) $-CH_2-$;

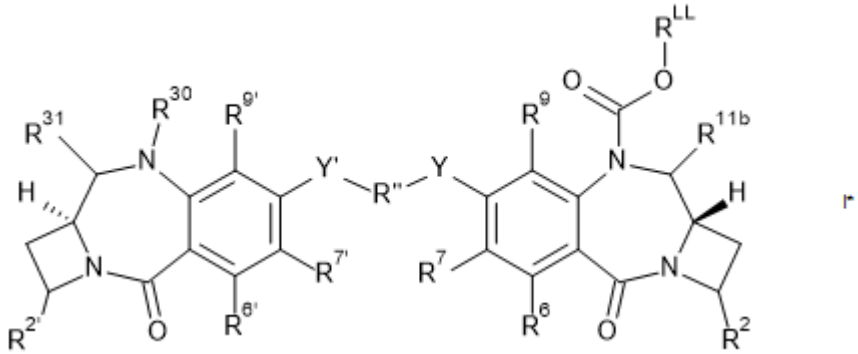
(b) $-C_3H_6-$; and

(c)  .

17. A conjugate of formula II:

$$L - (D^L)_p \qquad (II)$$

wherein L is a Ligand unit (i.e., a targeting agent), D^L is a Drug Linker unit of formula I'*:



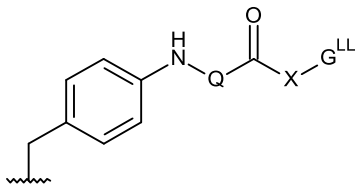
wherein:

Y, Y', R", R^2, $R^{2'}$, R^6, $R^{6'}$, R^7, $R^{7'}$, R^9 and $R^{9'}$ are as defined in any one of claims 1 to 6;

R^{11b}, R^{30} and R^{31} are as defined in either of claims 8 and 9;

R^{LL} is a linker for connection to a cell binding agent, which is selected from:

(iiia):

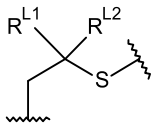


IIIa'

,

where Q and X are as defined in either claims 8 and 11 to 13 and G^{LL} is a linker connected to a Ligand Unit; and

(iiib):



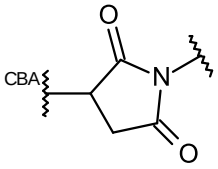
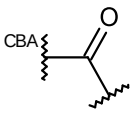
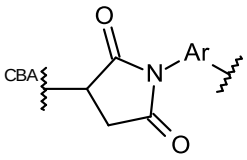
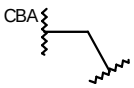
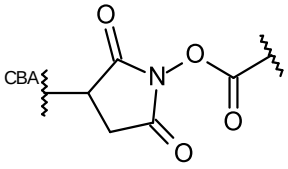
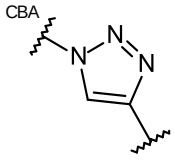
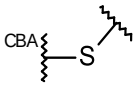
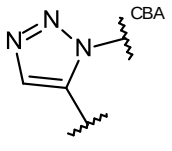
IIIb'

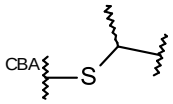
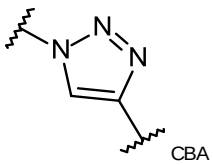
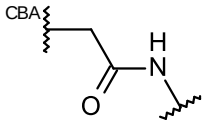
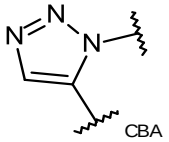
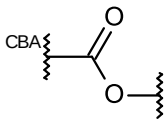
,

where R^{L1} and R^{L2} are as defined in claim 8;

wherein p is an integer of from 1 to 20.

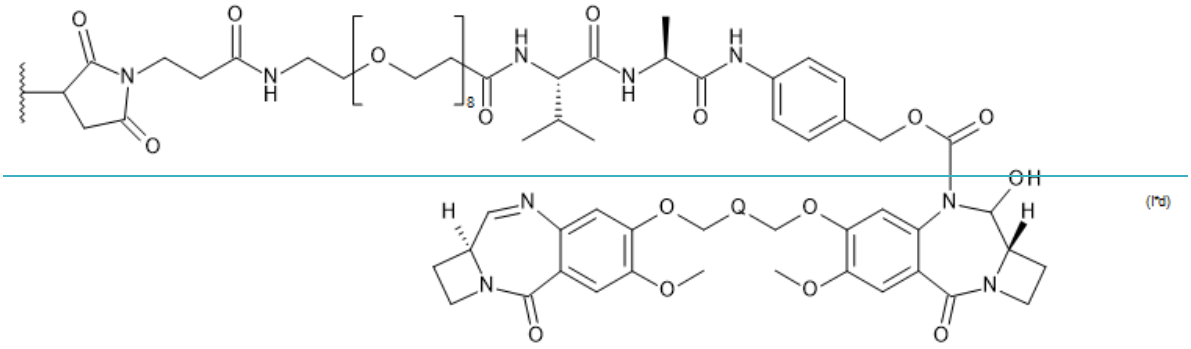
18. A conjugate according to claim 17, wherein G^{LL} is selected from:

(G^{LL1-1})		(G^{LL6})	
(G^{LL1-2})		(G^{LL7})	
(G^{LL2})		(G^{LL8-1})	
(G^{LL3-1})		(G^{LL8-2})	

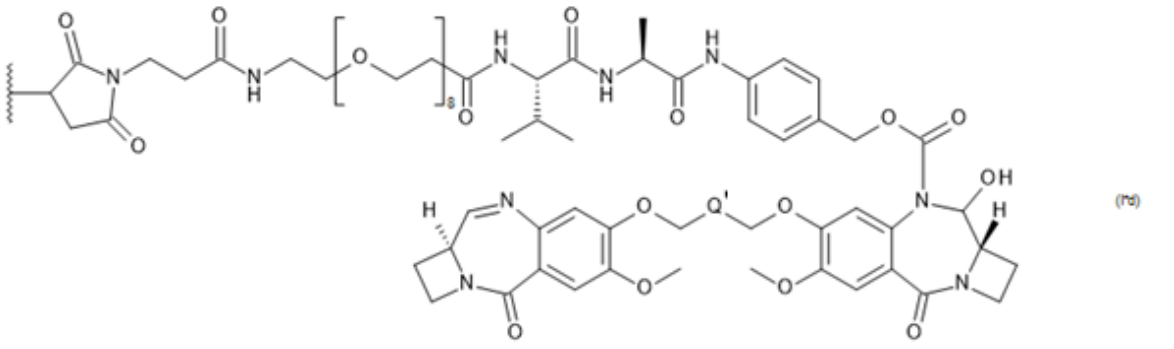
(G^{LL3-2})		(G^{LL9-1})	
(G^{LL4})		(G^{LL9-2})	
(G^{LL5})			

where Ar represents ~~a C_{5-6} arylene group~~phenylene.

19. A conjugate according to claim 17, wherein D^L is of formula (I*d'):



(I*d)

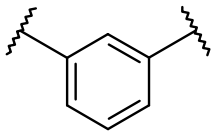


(I*d)

where Q' is selected from:

(a) $-CH_2-$;

(b) $-C_3H_6-$; and

(c)  .

20. A composition comprising a mixture of conjugates according to any one of claims 17 to 19, wherein the average p in the mixture of conjugate compounds is ~~about~~ 1 to ~~about~~ 8.

21. The conjugate according to any one of claims 17 to 19, for use in therapy.

22. A pharmaceutical composition comprising the conjugate of any one of claims 17 to 19, and a pharmaceutically acceptable diluent, carrier or excipient.

23. The conjugate according to any one of claims 17 to 19 or the pharmaceutical composition according to claim 22, for use in the treatment of a proliferative disease in a subject.

24. Use of the conjugate according to any one of claims 17 to 19 in the manufacture of a medicament for treating a proliferative disease.