

5 **AROMATIC BORON-CONTAINING COMPOUNDS
AND INSULIN ANALOGS**

CROSS REFERENCES TO RELATED APPLICATIONS

 This application claims priority from U.S. Provisional Patent Application Nos.
10 63/116,050, filed November 19, 2020, 63/122,338, filed December 7, 2020, 63/210,968, filed
June 15, 2021, and 63/249,868, filed September 29, 2021, all of which are incorporated herein
by reference.

FIELD OF THE DISCLOSURE

 The present disclosure relates to novel compounds that include one or more aromatic
15 boron-containing groups. The present disclosure further relates to kits and the use of the
compounds and/or pharmaceutical compositions comprising the disclosed compounds for the
treatment of disorders, such as hyperglycemia, type 2 diabetes, impaired glucose tolerance, type
1 diabetes, obesity, metabolic syndrome X, or dyslipidemia, diabetes during pregnancy, pre-
diabetes, Alzheimer's disease, MODY 1, MODY 2 or MODY 3 diabetes, mood disorders, and
20 psychiatric disorders.

SEQUENCE LISTING

 The present application is being filed along with a Sequence Listing in electronic
format. The Sequence Listing is provided as a file titled "X23056 2Nov2021_ST25," created
November 2, 2021, and is 4,084,000 bytes in size. The information in the electronic format of
25 the Sequence Listing is incorporated herein by reference in its entirety.

BACKGROUND OF THE DISCLOSURE

 Boronic acids are generally considered Lewis acids that have a tendency to bind to
hydroxyls, because, as Lewis acids, boronic acids can form complexes with Lewis bases such
30 as, for example, hydroxide anions. Thus, molecules containing boronates including boronic
acids have a general tendency to bind hydroxyl groups. This binding tendency can be used for
detection of hydroxyl-containing groups by boronated labeling reagents wherein the boronate
groups bind to the hydroxyls and, depending on the solvent and buffer conditions, the boronates
can form hydrolysable boronate-ester bonds to the hydroxyl groups of hydroxyl containing
35 molecules, such as the hydroxyl groups present in diols (e.g., glucose). Although boron-
containing compounds can bind to diol containing molecules, achieving selectivity using
boron-containing compounds has been challenging because of their ability to bind various

5 diols, including cis diols, to varying degrees. While improved binding affinity of boron-containing compounds towards a specific vicinal diol of interest may be achieved, this may result in a loss of selectivity.

Glucose is the main fuel for the human body, and blood glucose values are tightly regulated in healthy individuals. For example, between meals, blood glucose is near 5 mmol/L
10 (mM), and when blood glucose concentrations rise after a meal, the value is quickly adjusted back toward 5 mM by the action of insulin. The hormone insulin is secreted from pancreatic beta cells, and when insulin binds to insulin receptors on cells all over the body (for example muscle and fat), the cells are stimulated to absorb glucose by translocation of glucose transporters from storage vesicles to the cell surface (GLUT4). (Huang, S.H. et al. (2007) Cell
15 Metabolism 5:237-252.)

People with diabetes may lose their ability to produce insulin due to autoimmunity against the beta cells (type 1) or have low sensitivity to insulin in combination with impaired insulin secretion (type 2). For example, those with type 1 diabetes may rely on multiple daily insulin injections, both for basal coverage, typically once a day, and with meals (bolus) to
20 control their glucose levels. (Polonsky, K.S. et al. (1988) The Journal of Clinical Investigation 81: 442-448.) Because glucose values can fluctuate unpredictably, perfect insulin dosing day after day is extremely difficult. Indeed, despite many technological advances in diabetes treatment, researchers are currently seeing, partly due to lifestyle problems, a worsening of long-term glucose control and/or overall metabolic health.

25 Thus, there is an unmet medical need for novel compounds, such as glucose-responsive insulin analogues/conjugates, that can control blood glucose levels.

SUMMARY OF THE DISCLOSURE

The present disclosure provides, according to some embodiments, novel Z1c
30 compounds (e.g., FF1-FF224) comprising aromatic boron-containing functionalities (e.g., F1-F12). In some embodiments, the present disclosure provides compounds that comprise a drug substance (e.g., X1) and at least one Z1c with aromatic boron-containing functionalities (e.g., aromatic boron-containing groups, Z1c scaffolds comprising FF1-FF224 and F1-F12). In some
embodiments, the compounds disclosed comprise one or more molecular scaffolds (e.g., FF
35 scaffolds). In some embodiments, the compounds comprise a drug substance. In at least some embodiments, the drug substance is a polypeptide or a small-molecule.

5 In at least some embodiments, the disclosed compounds have at least two aromatic boron-containing functionalities. In some embodiments, the compounds are selective towards specific sugars, such as glucose, while showing reduced affinity for other sugars and wherein there is at least two aromatic boron-containing functionalities in the aromatic boron-containing portion of the compounds. In certain embodiments, the aromatic boron-containing portion may
10 be covalently conjugated, directly or indirectly, to a drug substance comprising an amine, a drug substance that is covalently conjugated to an amine containing linker, an amine configured to be covalently conjugated to a drug substance, NH₂, or OH (e.g., X1). In some embodiments, the drug substance is covalently conjugated to an amine containing linker and the amine group is conjugated to the aromatic boron-containing portion (Z1c). In certain embodiments the
15 aromatic boron-containing portion (Z1c) has an architecture comprising of tethering groups (FF formulae) and aromatic boron containing groups that collectively make the aromatic boron-containing portion of the disclosed compounds.

In some embodiments, one or more Z1c may be linked (e.g., conjugated, connected) to a drug substance (e.g., X1) via one or more small-molecule linkers (e.g., Z1b) and/or one or
20 more amino acids connected together using amide or peptide bonds (e.g., Z1a). In at least some embodiments, rotational constraining of the boron functionalities via the tethering group (e.g., FF formulae) enhances the binding affinity of the compounds (e.g., conjugates) towards specific diols (such as glucose) and away from other diols in the body and thereby provides selectivity to specific diols. In some embodiments, the compounds may exhibit therapeutic
25 pharmacokinetics and/or pharmacodynamics in response to endogenous and/or exogenous small molecules in the body, such as glucose. Changes in the physiological concentration of glucose, result in the activation and/or release of the drug molecule or conversely, deactivation and/or sequestering of the drug molecule (e.g., a peptide hormone), and/or modulation of the activity of the drug molecule, through interaction of glucose with the boron-containing
30 architectures conjugated to the drug molecule.

In some embodiments, the present disclosure provides a compound comprising X1 and one or more Z1c, or a tautomer, stereoisomer or a mixture of stereoisomers, or a pharmaceutically acceptable salt, or hydrate, or isotopic derivative thereof, wherein:

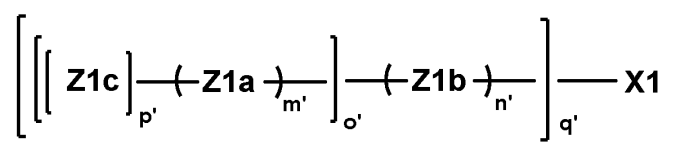
X1 comprises:

- 35 i. NH₂ or OH (e.g. X1 is NH₂ or OH);
ii. a drug substance comprising an amine;

- iii. a drug substance that is covalently conjugated to an amine containing linker; or
- iv. an amine configured to be covalently conjugated to a drug substance;

wherein each Z1c is independently selected from Formulae FF1-FF224; and wherein
10 each Z1c is covalently conjugated, directly or indirectly, to an amine in X1 or to OH when X1
is OH.

In some embodiments, the compound is a molecular conjugate represented by Formula I, or a stereoisomer or a mixture of stereoisomers, or pharmaceutically acceptable salts:



(Formula I)

15 wherein X1 comprises:

- 20
- i. NH_2 or OH (e.g. X1 is NH_2 or OH);
 - ii. a polypeptide drug substance comprising an amine;
 - iii. a polypeptide drug substance that is covalently conjugated to an amine containing linker; or
 - iv. an amine configured to be covalently conjugated to a polypeptide drug substance;

each Z1c is independently selected from Formulae FF1-FF224 and covalently conjugated either directly, or via Z1a and/or Z1b, to X1;

25 each Z1a comprises 1 to 50 amino acids connected together using amide or peptide bonds;

each Z1b is a small-molecule linker;

each m' is independently 0 or 1;

each n' is independently 0 or a positive integer;

30 each α_i is independently an integer of 1 or greater;

each p' is a positive integer; and

q' is a positive integer of at least 1 and not more than two times the total number of amine groups in X1, with the proviso that when any of n' , o' , p' , or q' is 2 or more, the

5 corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different;

wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1; and wherein optionally the molecular conjugate may comprise one or more isotopes at any position of the molecular conjugate of Formula I.

10 In some embodiments, X1 comprises human insulin or a human insulin analogue comprising an A-chain and a B-chain, wherein the A-chain comprises a sequence selected from SEQ ID NOs 1 and 3 to 33, and the B-chain comprises a sequence selected from SEQ ID NOs 2 and 34 to 74, 24047, and 24048;

each Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, FF163, FF193, FF194, FF203, and FF221-FF224 and covalently conjugated either 15 directly, or indirectly via a linker (an indirect linker), to Z1a and/or Z1b, or to X1;

each Z1a is independently absent or independently comprises a sequence selected from K, GK, KGS (SEQ ID NO:24049), KGSHK (SEQ ID NO:4238), KNSTK (SEQ ID NO:5085), GKASHK (SEQ ID NO:12414), GKKEEK (SEQ ID NO:12677), GKKEHK (SEQ ID NO:12680), KGSHK (SEQ ID NO:13120), KGSH (SEQ ID NO:24050), KGSHK (SEQ ID NO:13198), KGSTK (SEQ ID NO:13205), GKHEK (SEQ ID NO:13271), 20 GKNSHK (SEQ ID NO:13982), GKNSTK (SEQ ID NO:13989), GKQSSK (SEQ ID NO:14380), GKYQFK (SEQ ID NO:15128), KGSKK (SEQ ID NO:24045), GKPGKK (SEQ ID NO:24046), KGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and 25 GSHKGSHK (SEQ ID NO:24042);

each said linker is selected from FL1, FL3, FL4, and FL5;

each m' is independently 0 or 1;

each n' is independently 0, 1, 2, or 3;

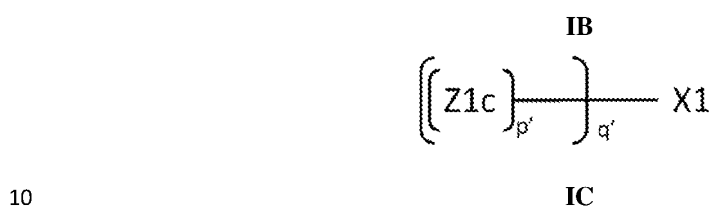
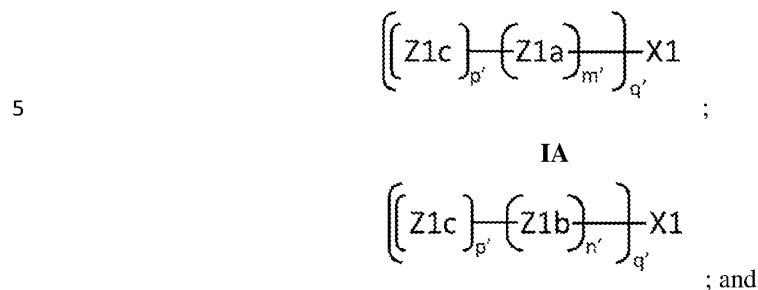
each o' is independently 1, 2, 3, 4, or 5;

30 each p' is 1, 2, 3, 4, or 5; and

q' is 1, 2, 3, or 4, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different; and

wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1.

35 In some embodiments, the compound of Formula I is selected from:



Also disclosed herein is a method of treating or preventing an endocrine and/or metabolic disease, in a subject in need thereof, comprising administering to said subject a compound of Formula (I), or a pharmaceutically acceptable salt thereof, or a pharmaceutical composition comprising a compound of Formula (I) or a pharmaceutically acceptable salt thereof. In some embodiments, the polypeptide drug substance is a polypeptide hormone such as insulin, an insulin analog, an incretin or an incretin analog. In some embodiments, the disease (e.g., disorder) is chosen from hyperglycemia, type 2 diabetes, impaired glucose tolerance, type 1 diabetes, obesity, metabolic syndrome X, or dyslipidemia, diabetes during pregnancy, pre-diabetes, Alzheimer's disease, MODY 1, MODY 2 or MODY 3 diabetes, mood disorders, and psychiatric disorders.

In at least one embodiment, the pharmaceutical composition of the present disclosure may be for use in (or in the manufacture of medicaments for) the treatment of diabetes in the subject. In at least one embodiment, a therapeutically-effective amount of a pharmaceutical composition of the present disclosure may be administered to a subject diagnosed with diabetes or metabolic disease. In at least one embodiment, the pharmaceutical composition of the present disclosure comprises at least one compound disclosed herein (e.g., Formula I, Z1c) and at least one additional component selected from pharmaceutically acceptable carriers, pharmaceutically acceptable vehicles, and pharmaceutically acceptable excipients.

In at least one embodiment, the present disclosure is directed to a human insulin analog comprising an A-chain and a B-chain, wherein the sequence of the A-chain comprises:

- 5 $X_{aa'}X_{bb'}X_{cc'}X_{dd'}X_{ee'}X_{ff'}X_{gg'}VEQCCX_{hh'}X_{ii'}ICSLYQLENYCNX_{jj'}X_{kk'}X_{ll'}X_{mm'}X_{nn'}X_{oo'}X_{pp'}$ (SEQ ID NO:24015); and wherein the sequence of the B-chain comprises:

(i)

$X_{aa'}X_{bb'}X_{cc'}X_{dd'}KX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}KX_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24016),

- 10 wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y and W,

(ii)

- 15 $X_{aa'}X_{bb'}X_{cc'}X_{dd'}KPX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}X_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24017),

- wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y and W, and wherein X_{ee} is selected from amino acid residues A, E, F, H, I, K, L, N, P, Q, R, S, T, V, Y and W,

(iii)

$X_{aa'}X_{bb'}X_{cc'}X_{dd'}KX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}KX_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24018),

- 25 wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is present and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is G,

(iv)

30 $X_{aa'}X_{bb'}X_{cc'}X_{dd'}KX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}KX_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24019),

- wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A,

- 5 D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least one of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ is present and at least one of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ is S, or

(v)

$X_{aa}X_{bb}X_{cc}X_{dd}KX_{ee}X_{ff}X_{gg}X_{hh}X_{ii}X_{jj}KX_{kk}X_{ll}X_{mm}X_{nn}QHLCGSHLVEALYLVCX_{oo}X_{pp}X_{qq}GFFYT$
 $X_{rr}X_{ss}X_{tt}X_{uu}X_{vv}X_{ww}$ (SEQ ID NO:24020),

- 10 wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least two of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ are present and at least one of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ is S, and another is G.
- 15 In some embodiments, one or more lysine residues and/or the N-terminus of the insulin A- or B-chain are covalently conjugated as described by Formula I. In some embodiments the insulins described herein are used as intermediate compounds for the manufacture of conjugates described by Formula I. In some embodiments, the insulins described herein are used in methods for preventing and/or treating an endocrine and/or metabolic disease, for
- 20 example comprising administering any compound (e.g., modified insulin) of the embodiments described herein to a subject in need thereof, thereby treating the endocrine and/or metabolic disease.

DETAILED DESCRIPTION OF THE DISCLOSURE

- While aromatic boron-containing compounds (e.g., groups) can bind to diol containing
- 25 molecules, achieving selectivity using aromatic boron-containing compounds (which can act as molecular sensors) is challenging because of their ability to bind various diols, including cis diols, to varying degrees. Improved binding affinity of aromatic boron-containing compounds (which can act as sensors) towards a specific vicinal diol of interest may result in a loss of selectivity.
- 30 Scaffolds that position the boron functionality (e.g., sensors) of the aromatic boron-containing compounds in a specific or particular ensemble of geometries can increase selectivity towards a specific vicinal diol while simultaneously maintaining affinity for the diol of interest. According to some embodiments, aromatic boron-containing compounds disclosed

5 herein have different pendant groups on the aromatic boron-based scaffolds along with which specific scaffold geometries that impact binding to hydroxyl containing molecules.

According to some embodiments, the compounds of the present disclosure comprise aromatic boron-containing compounds that orient the boron functionalities in three dimensional space, so that the boron-containing compounds are spatially oriented to engage hexoses
10 containing vicinal diols, such that the boron groups can appropriately engage the hydroxyls in the vicinal diol molecule and provide enhancement of selectivity. In some embodiments, the aromatic boron-containing compounds are modified with specific functional groups on the aromatic ring that, together with an appropriate or suitable scaffold, may provide higher selectivity and/or affinity for binding towards a vicinal diol of interest and away from other
15 diols in the body.

In some embodiments, the aromatic boron-containing compounds are conjugated to a drug substance (e.g., small-molecule, polypeptide) wherein the aromatic boron-containing compounds provide intramolecular and intermolecular interactions with the drug substance and/or with proteins in the body, such as circulating proteins in the blood and/or plasma
20 including albumin and/or globulins. In some embodiments, the selective binding of the sensors to specific vicinal diols changes the extent of those intramolecular and intermolecular bindings and thereby modulates the pharmacokinetics and overall activity of the drug substance in the body; this effect can be controlled by the level of the vicinal diols present.

In some embodiments the drug substance is a peptide hormone. In some embodiments,
25 the peptide hormone is a human peptide hormone such as insulin, glucagon, or another incretin hormone. In some embodiments the sensors are selective towards the vicinal diols in glucose, and this selectivity is enhanced while maintaining affinity to glucose and simultaneously reducing affinity to other sugars in the blood. In some embodiments, the scaffolds as well as (e.g., in combination with) the pendant groups on the aromatic core of the boron-containing
30 compounds enable controlling the overall activity and/or pharmacokinetics of the conjugated drug substances based on levels of glucose and/or other vicinal diols in the blood.

5 In some embodiments, the aromatic boron-containing compounds comprise specific scaffold molecules (i.e., FF structures) with conjugated boron functionalities (i.e., F1-F12 groups), wherein the scaffolds have been used to orient the boron functionalities in three dimensional geometries so that the boron functionalities are oriented near each other and within a distance that helps engage specific hydroxyl orientations of select hexoses such as glucose.

10 Without wishing to be bound by theory, it is believed that the aromatic boron-containing compounds (e.g., molecules) disclosed herein enhance selectivity through at least one or more of the following three mechanisms: (1) the FF scaffold facilitates matching the orientation of the hydroxyl and/or alkoxy groups on boron groups in the aromatic boron-containing compounds and the hydroxyls in the vicinal diol molecule which enhances selectivity; (2)

15 further selectivity gain is obtained by identifying specific functional groups attached to, or near, for example, the aromatic core of the boron-containing compound which impact the electronic structure of the aromatic boron-containing compound and thereby favor reversible binding to the vicinal diols at physiological pH; and (3) functional groups attached to the aromatic boron-containing compound (e.g., the sensor scaffold) help to provide steric hindrance to reduce

20 binding to unwanted hexoses while maintaining binding to the sugar of interest such as glucose. These effects as combined together in the present disclosure provide desired or suitable selectivity of binding towards a vicinal diol-containing molecule of interest and away from other diols in the body.

In some embodiments, the aromatic boron-containing compounds are conjugated to a

25 drug substance wherein the aromatic boron-containing compounds provide intramolecular and/or intermolecular interactions with proteins in the body. Such proteins may include circulating proteins in the blood and/or human plasma such as albumin, glycosylated proteins and/or immunoglobulins. In some embodiments the selective binding of the sensors to specific vicinal diols in a molecule of interest changes the extent of intramolecular and intermolecular

30 bindings and thereby modulates the pharmacokinetics and overall activity of the drug substance in the body. In some embodiments the drug substance is a peptide hormone and in certain

5 embodiments thereof the peptide hormone is an incretin hormone such as insulin and the
vicinal diol containing molecule is glucose, but the present disclosure is not limited thereto.

Definitions

10 Unless otherwise defined, all terms (including technical and scientific terms) used
herein have the same meaning as commonly understood by one of ordinary skill in the art to
which the present disclosure belongs. It will be further understood that terms, such as those
defined in commonly used dictionaries, should be interpreted as having a meaning that is
consistent with their meaning in the context of the relevant art and/or the present specification,
and should not be interpreted in an idealized or overly formal sense, unless expressly so defined
15 herein.

20 Unless specifically described herein, functional groups, functional moieties, and
reactions referred to herein are understood to have meanings consistent with standard
descriptions in and/or general principles of organic chemistry, for example, as described in
Organic Chemistry, Thomas Sorrell, University Science Books, Sausalito, 1999; Larock,
Comprehensive Organic Transformations, VCH Publishers, Inc., New York, 1989; Carruthers,
Some Modern Methods of Organic Synthesis, 3rd Edition, Cambridge University Press,
Cambridge, 1987; Smith and March, March's Advanced Organic Chemistry, 5th Edition, John
Wiley & Sons, Inc., New York, 2001. Generic functional groups (such as alkyl, aryl, acetyl,
etc.) encompass specific examples or species falling within those functional group categories as
25 generally defined in the field of organic chemistry, and those having ordinary skill in the art are
capable of identifying specific example embodiments of functional groups.

30 Unless specifically described herein, chemical terms, functional groups, and general
terms used throughout the specification are identified in accordance with the Periodic Table of
the Elements, CAS version, Handbook of Chemistry and Physics, 75th Ed., inside cover. In
certain embodiments the terms "a," "an," and "the" and similar referents used herein are to be
construed to cover both the singular and the plural unless their usage in context indicates
otherwise. The term "CAS #" as used herein is also referred to as CASRN or CAS Number, is a
unique numerical identifier assigned by Chemical Abstracts Service (CAS) to every chemical
substance described in the open scientific literature.

5 As used herein, nomenclature for compounds including organic compounds, can be given using common names, IUPAC, IUBMB, or CAS recommendations for nomenclature. One of skill in the art can readily ascertain the structure of a compound if given a name, either by systemic reduction of compound structure using naming conventions, or by commercially available software, such as CHEMDRAW™ (Cambridgesoft Corporation, U.S.A.).

10 The terminology used herein is for the purpose of describing embodiments and is not intended to be limiting of the present disclosure. It will be further understood that the terms “comprises,” “comprising,” “includes,” and “including,” when used in this specification, specify the presence of the stated features, integers, acts, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features,
15 integers, acts, operations, elements, components, and/or groups thereof. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items. Expressions such as “at least one of,” when preceding a list of elements, modify the entire list of elements and do not modify the individual elements of the list.

 As used herein, the terms “substantially,” “about,” and similar terms are used as terms
20 of approximation and not as terms of degree, and are intended to account for the inherent deviations in measured or calculated values that would be recognized by those of ordinary skill in the art. The term “about” used throughout is used to describe and account for small variations. For instance, “about” may mean the numeric value may be modified by $\pm 5\%$, $\pm 4\%$, $\pm 3\%$, $\pm 2\%$, $\pm 1\%$, $\pm 0.5\%$, $\pm 0.4\%$, $\pm 0.3\%$, $\pm 0.2\%$, $\pm 0.1\%$ or $\pm 0.05\%$. Numeric values modified
25 by the term “about” include the specific identified value. For example, “about 5.0” includes 5.0.

 Further, the use of “may” when describing embodiments of the present disclosure refers to “one or more embodiments of the present disclosure.” As used herein, the terms “use,” “using,” and “used” may be considered synonymous with the terms “utilize,” “utilizing,” and
30 “utilized,” respectively. Also, the term “exemplary” is intended to refer to an example or illustration.

 Also, any numerical range recited herein is intended to include all sub-ranges of the same numerical precision subsumed within the recited range. For example, a range of “1 to 10” is intended to include all subranges between (and including) the recited minimum value of 1
35 and the recited maximum value of 10, that is, having a minimum value equal to or greater than 1 and a maximum value equal to or less than 10, such as, for example, 2 to 7. Any maximum

5 numerical limitation recited herein is intended to include all lower numerical limitations subsumed therein, and any minimum numerical limitation recited in this specification is intended to include all higher numerical limitations subsumed therein. Accordingly, Applicant reserves the right to amend this specification, including the claims, to expressly recite any sub-range subsumed within the ranges expressly recited herein.

10 As used herein, “aromatic boron-containing group” refers to a compound having at least one boron atom covalently bonded to an aromatic group and/or a compound having at least one boron atom covalently incorporated within an aromatic group. The term “aromatic” as used herein may include “heterocycle,” “heterocyclyl,” or “heterocyclic.” As used herein the terms “heterocycle,” “heterocyclyl,” or “heterocyclic” each refer to an unsaturated 3- to 18-
15 membered ring containing one, two, three, or four heteroatoms independently selected from nitrogen, oxygen, phosphorus, and sulfur. In some embodiments, the term “aromatic” may include an “aryl.” The term “aryl” as used herein refers to a mono-, bi-, or other multi carbocyclic, aromatic ring system with 5 to 14 ring atoms. The aryl group can optionally be fused to one or more rings selected from aryls, cycloalkyls, heteroaryl, and heterocyclyls.
20 Exemplary aryl groups also include but are not limited to a monocyclic aromatic ring system, wherein the ring comprises 6 carbon atoms.

The term “heteroaryl” as used herein refers to a mono-, bi-, or multi-cyclic, aromatic ring system containing one or more heteroatoms, for example 1-3 heteroatoms, such as nitrogen, oxygen, and sulfur. Heteroaryls can be substituted with one or more substituents.
25 Heteroaryls can also be fused to non-aromatic rings. Exemplary heteroaryl groups include, but are not limited to, a monocyclic aromatic ring, wherein the ring comprises 2-5 carbon atoms and 1-3 heteroatoms. In some embodiments, the aromatic boron-containing group may include but is not limited to aryl- and heteroaryl boronic acids, aryl and heteroaryl boronate esters, and/or boroxoles. Exemplary aromatic boron-containing groups useful according to certain
30 embodiments, include, e.g., those described herein as FF1-FF224, F1-F10 and further include, e.g., those as disclosed in patent application PCT/US2021/025261 (filed Mar. 31, 2021) as compounds F1-F9, F12-F43, F500-F520; the disclosure of which is herein expressly incorporated by reference in its entirety.

The term “small-molecule linker” as used herein refers to a chemical group (e.g.,
35 scaffold, moiety) comprising a first attachment point toward X1 and a second attachment point toward Z1b, Z1a, or Z1c. In some embodiments, the first attachment point is toward X1 and the

5 second attachment point is toward Z1c. In some embodiments, the first attachment point is toward X1 and the second attachment point is toward Z1a. In some embodiments, the small molecule linker is a moiety/chemical group selected from Formulae IIa-IIai and Formulae IIIa-IIIai. In some embodiments, the small molecule linker is a moiety/chemical group selected from Formulae FL1-FL19 and an L- or D-amino acid comprising at least one amine group
10 directly conjugated to Z1c, wherein an acid functional group of the amino acid is conjugated toward X1 in Formula I.

The term “indirect linker” as used herein refers to a chemical group (e.g., scaffold, moiety) comprising a first attachment point toward X1 and a second attachment point toward Z1b, Z1a, or Z1c. In some embodiments, the first attachment point is toward X1 and the second
15 attachment point is toward Z1c. In some embodiments, the first attachment point is toward Z1a and the second attachment point is toward Z1c. In some embodiments, the indirect linker is a moiety/chemical group selected from Formulae FL1-FL19 and an L- or D-amino acid comprising at least one amine group directly conjugated to Z1c, wherein an acid functional group of the amino acid is conjugated toward Z1a or X1, independently, in Formula I.

20 The term “alkyl” as used herein refers to a saturated straight or branched hydrocarbon, such as a straight or branched group of 1-30 carbon atoms, referred to herein as C₁₋₃₀ alkyl. In some embodiments, the alkyl group is a C₁-C₂₂ alkyl group. In some embodiments, the alkyl group is a C₁-C₂₀ alkyl group. In some embodiments, the alkyl group is a C₁-C₁₈ alkyl group. In some embodiments, the alkyl group is a C₁-C₁₆ alkyl group. In some embodiments, the alkyl
25 group is a C₁-C₁₄ alkyl group. In some embodiments, the alkyl group is a C₁-C₁₂ alkyl group. In some embodiments, the alkyl group is a C₁-C₁₀ alkyl group. In some embodiments, the alkyl group is a C₁-C₈ alkyl group. In some embodiments, the alkyl group is a C₁-C₆ alkyl group. In some embodiments, the alkyl group is a C₁-C₄ alkyl group. Exemplary alkyl groups include, but are not limited to, methyl, ethyl, propyl, isopropyl, 2-methyl-1-propyl, 2-methyl-2-propyl,
30 2-methyl-1-butyl, 3-methyl-1-butyl, 2-methyl-3-butyl, 2,2-dimethyl-1-propyl, 2-methyl-1-pentyl, 3-methyl-1-pentyl, 4-methyl-1-pentyl, 2-methyl-2-pentyl, 3-methyl-2-pentyl, 4-methyl-2-pentyl, 2,2-dimethyl-1-butyl, 3,3-dimethyl-1-butyl, 2-ethyl-1-butyl, butyl, isobutyl, t-butyl, pentyl, isopentyl, neopentyl, hexyl, heptyl, and octyl. In some embodiments, “alkyl” is a straight-chain hydrocarbon. In some embodiments, “alkyl” is a branched hydrocarbon.

35 The term “cycloalkyl” as used herein refers to a saturated or unsaturated cyclic, bicyclic, or bridged bicyclic hydrocarbon group of 3-16 carbons, or 3-8 carbons, referred to

5 herein as “(C₃-C₈)cycloalkyl,” derived from a cycloalkane. Exemplary cycloalkyl groups include, but are not limited to, cyclohexanes, cyclohexenes, cyclopentanes, and cyclopentenes. Cycloalkyl groups may be substituted with alkoxy, aryloxy, alkyl, alkenyl, alkynyl, amide, amino, aryl, arylalkyl, carbamate, carboxy, cyano, cycloalkyl, ester, ether, formyl, halogen, haloalkyl, heteroaryl, heterocyclyl, hydroxyl, ketone, nitro, phosphate, sulfide, sulfinyl, 10 sulfonyl, sulfonic acid, sulfonamide and thioketone. Cycloalkyl groups can be fused to other cycloalkyl (saturated or partially unsaturated), aryl, or heterocyclyl groups, to form a bicycle, tetracycle, etc. The term “cycloalkyl” also includes bridged and spiro-fused cyclic structures which may or may not contain heteroatoms.

The term “acyl” as used herein refers to R-C(O)- groups such as, but not limited to, 15 (alkyl)-C(O)-, (alkenyl)-C(O)-, (alkynyl)-C(O)-, (aryl)-C(O)-, (cycloalkyl)-C(O)-, (heteroaryl)-C(O)-, and (heterocyclyl)-C(O)-, wherein the group is attached to the parent molecular structure through the carbonyl functionality. In some embodiments, it is a C₁₋₁₀ acyl radical which refers to the total number of chain or ring atoms of the, for example, alkyl, alkenyl, alkynyl, aryl, cycloalkyl, or heteroaryl, portion plus the carbonyl carbon of acyl. For example, 20 a C₄-acyl has three other ring or chain atoms plus carbonyl. In some embodiments, it is a C₁-C₂₂acyl group. In some embodiments, it is a C₁-C₂₀acyl group. In some embodiments, it is a C₁-C₁₈acyl group. In some embodiments, it is a C₁-C₁₆acyl group. In some embodiments, it is a C₁-C₁₄acyl group. In some embodiments, it is a C₁-C₁₂acyl group. In some embodiments, it is a C₁-C₁₀acyl group. In some embodiments, it is a C₁-C₈acyl group.

25 The term “haloalkyl” as used herein refers to an alkyl group substituted with one or more halogens. Examples of haloalkyl groups include, but are not limited to, trifluoromethyl, difluoromethyl, pentafluoroethyl, trichloromethyl, etc. In some embodiments, it is a C₁-C₂₂ haloalkyl group. In some embodiments, it is a C₁-C₂₀ haloalkyl group. In some embodiments, it is a C₁-C₁₈ haloalkyl group. In some embodiments, it is a C₁-C₁₆ haloalkyl group. In some 30 embodiments, it is a C₁-C₁₄ haloalkyl group. In some embodiments, it is a C₁-C₁₂ haloalkyl group. In some embodiments, it is a C₁-C₁₀ haloalkyl group. In some embodiments, it is a C₁-C₈ haloalkyl group.

The term “aryl” as used herein refers to a mono-, bi-, or other multi-carbocyclic, aromatic ring system with 5 to 14 ring atoms. The aryl group can optionally be fused to one or 35 more rings selected from aryls, cycloalkyls, heteroaryls, and heterocyclyls. The aryl groups of this present disclosure can be substituted with groups selected from alkoxy, aryloxy, alkyl,

- 5 alkenyl, alkynyl, amide, amino, aryl, arylalkyl, carbamate, carboxy, cyano, cycloalkyl, ester, ether, formyl, halogen, haloalkyl, heteroaryl, heterocyclyl, hydroxyl, ketone, nitro, phosphate, sulfide, sulfinyl, sulfonyl, sulfonic acid, sulfonamide, and thioketone. Exemplary aryl groups include, but are not limited to, phenyl, tolyl, anthracenyl, fluorenyl, indenyl, azulenyl, and naphthyl, as well as benzo-fused carbocyclic moieties such as 5,6,7,8-tetrahydronaphthyl.
- 10 Exemplary aryl groups also include but are not limited to a monocyclic aromatic ring system, wherein the ring comprises 6 carbon atoms.

“Isomers” means compounds having the same number and kind of atoms, and hence the same molecular weight, but differing with respect to the arrangement or configuration of the atoms in space.

- 15 “Stereoisomer” or “optical isomer” means a stable isomer that has at least one chiral atom or restricted rotation giving rise to perpendicular dissymmetric planes (e.g., certain biphenyls, allenes, and spiro compounds) and can rotate plane-polarized light. Because asymmetric centers and other chemical structure exist in the compounds of the disclosure which may give rise to stereoisomerism, the disclosure contemplates stereoisomers and
- 20 mixtures thereof. The compounds of the disclosure and their salts include asymmetric carbon atoms and may therefore exist as single stereoisomers, racemates, and as mixtures of enantiomers and diastereomers. In some embodiments, such compounds will be prepared as a racemic mixture. In some embodiments, such compounds can be prepared or isolated as pure stereoisomers, e.g., as individual enantiomers or diastereomers, or as stereoisomer-enriched
- 25 mixtures. As discussed in more detail below, individual stereoisomers of compounds may be prepared by synthesis from optically active starting materials containing the desired chiral centers or by preparation of mixtures of enantiomeric products followed by separation or resolution, such as conversion to a mixture of diastereomers followed by separation or recrystallization, chromatographic techniques, use of chiral resolving agents, or direct
- 30 separation of the enantiomers on chiral chromatographic columns. Starting compounds of particular stereochemistry are either commercially available or are made by the methods described below and resolved by techniques well-known in the art.

The term “pharmaceutically acceptable salt(s)” refers to salts of acidic or basic groups that may be present in compounds used in the present compositions.

- 35 As used herein, “drug substance” refers to small-molecule compounds and/or polypeptide containing compounds. According to some embodiments, a drug substance suitable

5 for use in the compounds and methods described herein is a therapeutically, prophylactically and/or diagnostically active drug substance.

It will be understood that, although terms such as “first,” “second,” “third,” etc., may be used herein to describe various elements (such as molecules, components, groups, and/or moieties, etc.), those elements should not be limited by these terms. These terms are merely
10 used to distinguish one element from another element. Thus, a first element described below could be termed a second element without departing from the spirit and scope of the present disclosure. It will be understood that when an element or group is referred to as being “connected to,” “conjugated with,” “linked,” or “coupled to” another element or group, the two elements may be directly connected, or one or more intervening elements may be present. It
15 will be understood that conjugations and linkages described herein have the option of being direct conjugations or direct linkages, unless expressly excluded or precluded by the context.

As used herein, the terms “directly” or “directly covalently conjugated” or “covalently conjugated directly” may be interchangeably used to indicate that a first group is “directly” or “directly covalently conjugated” or “covalently conjugated directly” to a second group, which
20 means the first and second groups are covalently bonded together without additional intervening groups.

As used herein, the terms “indirectly” or “indirectly covalently conjugated” or “covalently conjugated indirectly” may be interchangeably used to indicate that a first group is “indirectly” or “indirectly covalently conjugated” or “covalently conjugated indirectly” to a
25 second group, which means the first and second groups are covalently bonded together with at least one additional intervening group (e.g., a small-molecule, a linker, a spacer, a linear sequence of amino acids and/or nonlinear sequence of amino acids).

In at least some embodiments, one or more groups (e.g., X1a, Z1a, Z1b, Z1c) are covalently conjugated directly or indirectly to each other. For example, according to certain
30 embodiments Z1c is covalently conjugated, directly or indirectly, to an amine in X1 or to OH when X1 is OH. As another example, according to certain embodiments one or more drug substances (X1) are covalently conjugated to one or more amine containing linkers. In some embodiments, X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH

35 In certain embodiment, each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1.

5 As used herein, terms such as “attachment point toward [group],” “attachment to,” and “covalent linkage toward [group]” express that the indicated atom, attachment, or linkage is closer to the indicated group than the other attachment point or covalent linkage variables within the structure formula. In some embodiments an attachment point or covalent linkage may be directly adjacent to the indicated group, and in some embodiments other atoms or
10 groups may be present therebetween.

As used herein, the term “percentage homology” refers to the percentage of sequence identity between two sequences after optimal alignment. Identical sequences have a percentage homology of 100%. Optimal alignment may be performed by homology alignment algorithms described by the search for similarity method of Pearson and Lipman, *Proc. Natl. Acad. Sci.*
15 *USA* 85:2444 (1988), by general methods described for search for similarities by Needleman and Wunsch, *J. Mol. Biol.* 48:443 (1970), including implementation of these algorithms or visual comparison. As used herein, “insulin A-chain” is the chain of insulin that has the highest percentage homology to the A-chain of wild-type human insulin. As used herein, “insulin B-chain” is the chain of insulin that has the highest percentage homology to the B-chain of wild-
20 type human insulin.

In some embodiments, the terms “covalently connected,” “covalently conjugated,” or “through a covalent conjugation” may be interchangeably used to indicate that two or more atoms, groups, or chemical moieties are bonded or connected via a chemical linkage. In some embodiments, the chemical linkage (which in some embodiments may be referred to as a
25 covalent linkage) may be (e.g., consist of) one or more shared electron pairs (e.g., in a single bond, a double bond, or a triple bond) between two atoms, groups, or chemical moieties.” In some embodiments, the chemical (covalent) linkage may further include one or more atoms or functional groups, and may be referred to using the corresponding name of that functional group in the art. For example, a covalent linkage including a –S–S– group may be referred to as
30 a disulfide linkage; a covalent linkage including a –(C=O)– group may be referred to as a carbonyl linkage; a covalent linkage including a –(CF₂)– group may be referred to as a difluoromethylene linkage, etc. The type of linkage or functional group within the covalent bond is not limited unless expressly stated, for example when it is described as including or being selected from certain groups. The types or kinds of suitable covalent linkages will be
35 understood from the description and/or context.

5 In some embodiments, side chains of amino acids may be covalently connected (e.g., linked or cross-linked) through any number of chemical bonds (e.g., bonding moieties) as generally described in *Bioconjugate Techniques* (Third edition), edited by Greg T. Hermanson, Academic Press, Boston, 2013. For example, the side chains may be covalently connected through an amide, ester, ether, thioether, isourea, imine, triazole, or any suitable covalent
10 conjugation chemistry available in the art for covalently connecting one peptide, protein, or synthetic polymer to a second peptide, protein, or synthetic polymer. The term polymer includes polypeptide. The term “covalent conjugation chemistry” may refer to one or more functional groups included in the bonding moiety, and/or the chemical reactions used to form the bonding moiety.

15 The term “vicinal diol” refers to a group of molecules in which two hydroxyl groups occupy vicinal positions, that is, they are attached to adjacent atoms. Such molecules may include, but are not limited to, sugars such as hexoses, glucose, mannose and fructose.

 In some embodiments, the term “albumin” means human serum albumin or a protein with at least 60% percentage homology to human serum albumin protein. It is to be understood
20 that in some embodiments the albumin may be further chemically modified for the purposes of conjugation. In some embodiments, modifications may include one or more covalently connected linkers.

 The term “treatment” is meant to include both the prevention and minimization of the referenced disease, disorder, or condition (i.e., “treatment” refers to both prophylactic and
25 therapeutic administration of a compound of the present invention or a composition comprising a compound of the present invention unless otherwise indicated or clearly contradicted by context). The route of administration may be any route which effectively transports a compound of this disclosure to the desired or appropriate place in the body, such as parenterally, for example, subcutaneously, intramuscularly, orally, or intravenously. For
30 parenterally administration, a compound of the disclosure is formulated analogously with the formulation of known insulins. Furthermore, for parenterally administration, a compound of this disclosure is administered analogously with the administration of known insulins and the physicians are familiar with this procedure. The amount of a compound of this disclosure to be administered, the determination of how frequently to administer a compound of this disclosure,
35 and the election of which compound or compounds of this disclosure to administer, optionally

5 together with another antidiabetic compound, is decided in consultation with a practitioner who is familiar with the treatment of the condition (e.g. diabetes) to be treated.

In some embodiments, “therapeutic composition” and “pharmaceutical composition” as used herein means a composition that is intended to have a therapeutic effect such as pharmaceutical compositions, genetic materials, biologics, and other substances.

10 Pharmaceutical compositions may be configured to function in the body with therapeutic qualities; concentration may be altered to reduce the frequency of replenishment, and the like. In some embodiments “therapeutically effective amount” and “prophylactically effective amount” refer to an amount that provides a therapeutic benefit in the treatment, prevention, or management of a disease or an overt symptom of the disease. The therapeutically effective
15 amount may treat a disease or condition, a symptom of disease, or a predisposition toward a disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve, or affect the disease, the symptoms of disease, or the predisposition toward disease. The set or specific amount that is therapeutically effective can be readily determined by an ordinary medical practitioner, and may vary depending on factors known in the art, such as, *e.g.* the type
20 of disease, the patient's history and age, the stage of disease, and the administration of other therapeutic agents. In some embodiments, modified insulins described herein are delivered to the body by injection or inhalation, or by other routes, and can reversibly bind to soluble glucose in a non-depot form. In some embodiments, modified insulins described herein are released over an extended period of time from a local depot in the body or from bound forms to
25 proteins in the serum such as albumin. In some embodiments, the release of modified insulin is accelerated at elevated glucose levels, and in some embodiments such release rate may be dependent on blood sugar levels or levels of other small molecules in the blood including diol containing molecules. In some embodiments the release, bioavailability, and/or solubility of modified insulins described herein is controlled as a function of blood or serum glucose
30 concentrations or concentrations of other small molecules in the body.

Additionally, unless otherwise stated, structures described herein are also meant to include compounds that differ only in the presence of one or more isotopically enriched atoms. For example, compounds having the present structures except for the replacement of hydrogen by deuterium (^2H) or tritium (^3H), or the replacement of a carbon by a ^{13}C - or ^{14}C -carbon atom
35 are within the scope of this disclosure. Such compounds may be useful as, for example, analytical tools, probes in biological assays, or therapeutic agents.

5 In some embodiments, functional groups can be covalently conjugated or linked via any suitable covalent conjugation chemistry (linker) that can be used to covalently conjugate one functional group or amino acid side chain to another functional group, non-limiting examples include an amide, an ester, an ether, a thioether, an isourea, an imine, and a triazole linker. In some embodiments functional groups are covalently conjugated through click chemistry
10 reactions as defined in the art. These include, for example, cycloaddition reactions including but not limited to 3+2 cycloadditions, strain-promoted alkyne-nitrone cycloaddition, reactions of strained alkenes, alkene and tetrazine inverse-demand Diels-Alder, Copper(I)-Catalyzed Azide-Alkyne Cycloaddition (CuAAC), Strain-promoted azide-alkyne cycloaddition, Staudinger ligation, nucleophilic ring-opening reactions, and additions to carbon-carbon
15 multiple bonds. Some of these reactions are described for example by H. C. Kolb, M. G. Finn and K. B. Sharpless (2001); Click Chemistry: Diverse Chemical Function from a Few Good Reactions, *Angewandte Chemie International Edition* 40 (11): 2004–2021; Kolb and Sharpless, *Drug Discovery Today* 8:1128-1137, 2003; Huisgen, R. *Angew. Chem. Int. Ed. Engl.* 1963, 2, 565; Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* 2006,
20 1, 644. One skilled in the art will be capable of selecting suitable buffers, pH and reaction conditions for such click reactions. In some embodiments, covalent conjugation is the result of a “bioorthogonal reaction” as defined in the art. Such reactions are, for example, described by Sletten, Ellen M.; Bertozzi, Carolyn R. (2009). Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality, *Angewandte Chemie International Edition* 48 (38): 6974–
25 98.; Prescher, Jennifer A; Bertozzi, Carolyn R (2005). Chemistry in living systems, *Nature Chemical Biology* 1 (1): 13–21.

In some embodiments, functional groups may be linked using native chemical ligation as described for example by Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. (1994) Synthesis of proteins by native chemical ligation, *Science* 266 (5186): 776–778. As used
30 herein, terms such as “linkage,” “covalent conjugation,” etc. may refer to any of the chemistries described above in some embodiments. The terms “amine,” “amino group,” and/or “amine group,” when used to describe part of a covalent bond or connectivity, may be interchangeably used to indicate an amino group or an amine group to which the described element is covalently linked. In some embodiments, the amino group or amine group may be a primary amine, a
35 secondary amine, or a fragment such as $\text{NH---}\oplus$ to which a conjugation is made and described.

In some embodiments, the amino group or amine group may be the NH_2 group at the N-

5 terminus of a peptide or peptide chain, or the NH₂ group of a lysine side chain, but
embodiments of the present disclosure are not limited thereto. In some embodiments, the
connectivity of a first group to a second group is described by reference to an amine or amino
group, originating from the second group, that is part of a covalent linkage between the first
group and the second group. For example, an amine of a lysine side chain on X1 may be
10 referred to as an amine, and furthermore may be described as being conjugated through an
amide bond in order to specify the structure and connectivity of the functional groups that
constitute the covalent bond. If a covalent linkage is via an amine bond or amine linkage, then
it is referred to as an amine linkage. It is to be understood that a carbonyl connected to an
amine (e.g., a (C=O)-NH moiety) constitutes an amide bond, and thus by definition, an amine
15 linkage is not directly connected to a carbonyl group. Stated another way, the terms “amide
bond” and/or “amide linkage,” when used to describe a covalent bond or connectivity, may be
interchangeably used to indicate a carbonyl connected to an amine (e.g., a (C=O)-NH moiety).

In some embodiments further modifications include attachment of a chemical entity
(e.g., moiety or functional group) such as a carbohydrate group, one or more cis-diol containing
20 groups, one or more phosphate groups, one or more catechol groups, a farnesyl group, an
isofarnesyl group, a fatty acid group, or a linker for conjugation, functionalization, or other
modifications meant to impact the pharmacokinetics, pharmacodynamics, and/or biophysical
solution characteristics of insulin.

In some embodiments, a compound, such as a molecular conjugate, includes a human
25 peptide hormone (e.g., as X1). In some embodiments, the peptide hormone is a polypeptide
hormone of the human pancreas. In some embodiments, X1 in Formula I is NH₂. In some
embodiments, a compound of Formula I is conjugated to a drug substance via an optional
covalent-spacer. In some embodiments, a compound, such as a molecular conjugate, includes a
human insulin or a human insulin analogue. In some embodiments, two different amine groups
30 in insulin are covalently conjugated to as described by Formula I.

It will be understood that “human peptide hormone,” “polypeptide hormone of the
human pancreas,” “Insulin,” “human insulin,” “modified insulin,” and “human insulin
analogue” may be used interchangeably in some of the described embodiments; that is, for
example, in certain embodiments “human insulin analogue” may instead be used in
35 embodiments described as using human insulin. In some embodiments a compound, such as a
molecular conjugate, includes a human insulin or a human insulin analogue. In some

5 embodiments, a molecular conjugate includes a human insulin or a human insulin analogue as described by Formula I for $p'=1$ wherein a single amino group in insulin is conjugated to as described by Formula I. In some embodiments, the amino group is the N-terminus of the B-chain of insulin or an amino group of the side chain of a lysine. In some embodiments, two or more different amine groups in insulin are each independently covalently conjugated to as described by Formula I. In some embodiments, at least one amine groups is the N-terminus of the B-chain of insulin. In some embodiments, amino groups comprise amino groups of side chains of lysine residues in insulin.

 Various suitable modifications of the peptide hormone (e.g., human polypeptide hormone, for example, Insulin), known to those having ordinary skill in the art, are included in the scope of the disclosure. In some embodiments, the polypeptide of Z1a or the optionally extended polypeptide at the N-terminus of B-chain or C-terminus of A-chain of insulin contain sequences with up to 70% sequence homology to a human polypeptide sequence. In some embodiments, the polypeptide of Z1a or the optionally extended polypeptide at the N-terminus of B-chain or C-terminus of A-chain of insulin contain one or more lysine residues that are optionally next to a proline residue, such that the proline is C-terminal to lysine. In some embodiments, the amino group of lysine residues is each independently conjugated as described by Formula I.

 In some embodiments, insulin is further modified through conjugation to a sugar- or diol-containing molecule. In some embodiments, the human polypeptide hormone is a dual or triple hybrid peptide comprising sequences of two or more human peptide hormones and which can act through multiple receptors; for example, a glucose-dependent insulinotropic polypeptide (GIP) and GLP-1 receptor agonist or GLP-1/GIP/glucagon triple agonist. In some embodiments, the human polypeptide hormone is a gut hormone. In some embodiments, the human polypeptide hormone is chosen from c-peptide, adrenocorticotrophic hormone (ACTH), amylin, angiotensin, atrial natriuretic peptide (ANP), calcitonin, cholecystokinin (CCK), gastrin, ghrelin, glucagon, growth hormone, follicle-stimulating hormone (FSH), insulin, leptin, melanocyte-stimulating hormone (MSH), oxytocin, parathyroid hormone (PTH), prolactin, renin, somatostatin, thyroid-stimulating hormone (TSH), thyrotropin-releasing hormone (TRH), vasopressin, vasoactive intestinal peptide, a neuropeptide, a peptide hormone that impacts cardiovascular health or appetite, a hybrid of one or more of these peptides, and an analogue of one of these peptides. In some embodiments, compounds comprise a human polypeptide

5 hormone further modified, for example, through the covalent conjugation to polymers, XTEN
protein sequences or aliphatic chains. In some embodiments, polymer modified compounds
have a longer circulation time in the blood. In some embodiments, polymer modified
compounds, such long-acting variants require once a day injection, or one a week injection or
once a month injection. In some embodiments, a human polypeptide hormone or the analogue
10 thereof includes one or more L- or D- amino acids that are each independently one of the
twenty canonical amino acids or a non-canonical amino acid.

In some embodiments, a human polypeptide hormone or the analogue thereof includes
one or more residues that are 2-aminoisobutyric acid. In some embodiments, the C-terminus of
the B-chain of insulin is covalently conjugated to the N-terminus of the A-chain. In some
15 embodiments, the C-terminus of the B-chain of insulin is covalently conjugated to the N-
terminus of the A-chain and the connecting peptide is a C-peptide. In some embodiments, the
C-terminus of the B-chain of insulin is covalently conjugated to the N-terminus of the A-chain
and the connecting peptide is a C-peptide and further includes any intermediate compounds that
comprise a conjugate of Formula I. In some embodiments, insulin includes insulin lispro, or a
20 glargine-type of modification, or any suitable modification to human insulin analogue that
impacts the pharmacokinetics or half-life of insulin in the blood.

In some embodiments, the polypeptide hormone is glucagon. In some embodiments,
glucagon has additional mutations and modifications that are known to impact solubility and
solution stability of glucagon. In some embodiments, a compound, such as a molecular
25 conjugate, comprises a conjugation of a Z1a to the N-terminus of the B-chain of insulin through
a peptide bond, and at least one additional conjugation described by Formula I to insulin. In
some embodiments, the additional conjugation is to a lysine residue in insulin. In some
embodiments, at least one such lysine is a residue between position 15 and the C-terminus of
the B-chain of insulin. In some embodiments, the lysine residue is optionally next to a proline,
30 glycine, arginine, threonine or serine. In some embodiments, one or more amino acids in
Formula I is a D-amino acid. In some embodiments, any secondary or primary amine in a
compound, such as a molecular conjugate described by Formula I, is each independently
optionally acetylated. In some embodiments, a compound, such as a molecular conjugate has a
polypeptide hormone X1 further conjugated to a drug molecule, an imaging agent, a chelator, a
35 contrast agent, a radioactive isotope or a molecule that engages immune cells. In some
embodiments, X1 is a polypeptide hormone comprising a peptide ligand that binds to an

5 extracellular protein receptor. In some embodiments, X1 comprises a polypeptide analogue of a human polypeptide hormone that has at least 50% homology to a natural human polypeptide hormone. In some embodiments, X1 is an analogue of human insulin with up to 10 additional residues added to the A-chain or the B-chain of insulin.

In some embodiments, the term "glucose responsiveness" refers to the change in
10 activity in the presence and absence of glucose or in a difference of lower levels and higher levels of glucose (e.g., 3 mM glucose vs 20 mM glucose). In some embodiments the activity of a conjugated insulin is assessed by the concentration of insulin (in nanomolar units (nM) of insulin) required to induce the half maximum response (EC50) in a cell-based assay. Lower EC50 concentrations of conjugated insulins have a higher activity than insulins with higher
15 EC50 concentrations (e.g., an insulin with an EC50 of 3 nM is more active than an insulin with an EC50 of 50 nM). A "glucose response" is observed when the insulin changes from a less active EC50 (higher nM) to a more active EC50 (lower nM) in the absence and presence of glucose or in lower and higher levels of glucose, respectively.

In some embodiments, the compound, such as a molecular conjugate, comprises one or
20 more L- or D-artificial amino acids which are not one of the twenty naturally occurring amino acids. In some embodiments, the side chains of such artificial amino acids can be covalently conjugated through a number of reactions, including bio-orthogonal reactions such as, for example, described by: Rostovtsev, V.V., Green, L.G., Fokin, V.V. & Sharpless, K.B. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed.* 41, 2596-2599 (2002), or by: Liang, Y., Mackey, J.L., Lopez, S.A., Liu, F. & Houk, K.N. Control and design of mutual orthogonality in
25 bioorthogonal cycloadditions. *J. Am. Chem. Soc.* 134, 17904-17907 (2012). In some embodiments, Z1a contains one or more L- or D-artificial amino acids that are not one of the twenty naturally occurring amino acids. In some embodiments, the side chains of two amino
30 acids in Z1a are covalently conjugated together through a triazole bond.

Insulin hormone is an important regulator of blood glucose (sugar) levels. In a normal individual, insulin is present and, when released by the pancreas, it acts to reduce blood sugar levels, for example, by binding to and activating the insulin receptor, triggering glucose absorption by liver, fat, and skeletal muscle cells. Diabetes mellitus (DM), commonly referred
35 to as diabetes, is a group of metabolic diseases characterized by the persistence of high blood sugar levels over a prolonged period.

5 As used herein, “insulin” encompasses both wild-type and altered forms of insulin capable of binding to and activating the insulin receptor, or capable of causing a measurable reduction in blood glucose when administered in vivo and encompasses both wild-type and altered forms of human insulin capable of binding to and activating the human insulin receptor, or capable of causing a measurable reduction in blood glucose when administered in vivo to a human.

10 In some embodiments, insulin includes insulin from any species whether in purified, synthetic, or recombinant form and includes human insulin, porcine insulin, bovine insulin, sheep insulin and rabbit insulin. In some embodiments, insulin has two chains: a B- and an A-chain. In some embodiments, the chains are connected together through peptides such as, for example, c-peptide as is known in the art, or a shortened version of the c-peptide, and in other embodiments the insulin may be provided as a proinsulin (insulin precursor) that can be further processed into mature insulin. A variety of altered forms of insulin are known in the art and may be chemically altered such as by addition of a chemical moiety such as a PEG group or a fatty acyl chain. Altered insulins may be mutated including additions, deletions or substitutions of amino acids. In some embodiments the term “desB30” refers to an insulin lacking the B30 amino acid residue.

In some embodiments, insulin analogues include insulin that is chemically altered as compared to wild type human insulin, such as, but not limited to, by addition of a chemical moiety such as a PEG group or a fatty acyl chain. In some embodiments, altered insulins or insulin analogues may be mutated including additions, deletions or substitutions of amino acids. Different protomers of insulin may result from these changes and be incorporated into some embodiments. In some embodiments, active forms of insulins have fewer than 11 such modifications (e.g., 1-4, 1-3, 1-9, 1-8, 1-7, 1-6, 2-6, 2-5, 2-4, 1-5, 1-2, 2-9, 2-8, 2-7, 2-3, 3-9, 3-8, 3-7, 3-6, 3-5, 3-4, 4-9, 4-8, 4-7, 4-6, 4-5, 5-9, 5-8, 5-7, 5-6, 6-9, 6-8, 6-7, 7-9, 7-8, 8-9, 9, 8, 7, 6, 5, 4, 3, 2 or 1). As used herein, the wild-type sequence of human insulin (A-chain and B-chain), has an A-chain with the amino acid sequence GIVEQCCTSICSLYQLENYCN (SEQ ID NO:1), and a B-chain having the amino acid sequence FVNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:2).

Human insulin differs from rabbit, porcine, bovine, and sheep insulin in amino acids A8, A9, A10, and B30, which are in order the following: Thr, Ser, Ile, Thr for human; Thr, Ser, Ile, Ser for rabbit; Thr, Ser, Ile, Ala for porcine; Ala, Gly, Val, Ala for sheep; and Ala, Ser, Val,

5 Ala for bovine. In some embodiments, a modified insulin may be mutated at position B1, B2, B28 or B29, or at positions B28 and B29 of the B-chain. In some embodiments, a modified insulin may be mutated at A1, A2, A21 or other positions of the A-chain. For example, insulin lispro is a fast-acting modified insulin in which the lysine and proline residues on the C-terminal end of the B-chain have been reversed. Insulin aspart is a fast-acting modified insulin
10 in which proline has been substituted with aspartic acid at position B28. It is contemplated in some embodiments of the present disclosure that insulins mutated at B28 and B29 may further include additional mutations. For example, insulin glulisine is a fast-acting modified insulin in which aspartic acid has been replaced by a lysine residue at position B3, and lysine has been replaced by a glutamic acid residue at position B29. In some embodiments, longer acting and
15 higher stability insulin analogs are covalently modified as described by Formula I, and may contain mutations such as tyrosine at A14 replaced with glutamic acid, the tyrosine at B16 replaced with histidine, and the phenylalanine at B25 replaced with a histidine.

In some embodiments, the isoelectric point of insulins herein may be shifted relative to wild-type human insulin using any suitable method, for example by addition or substitution of
20 suitable amino acids. In some embodiments, the isoelectric point of the modified insulins may be modulated by glucose (e.g., by interaction with glucose). For example, insulin glargine is a basal insulin in which two arginine residues have been added to the C-terminus of the B-peptide, and A21 has been replaced by glycine. In some embodiments, the insulin may not have one or more of the residues B1, B2, B3, B26, B27, B28, B29, and B30 (e.g., the insulin may be
25 a deletion mutant at one or more of the listed residues). In some embodiments, the insulin molecule contains up to five additional amino acid residues on the N- or C-terminus of the A-chain or B-chain. In some embodiments, one or more amino acid residues are located at positions A1, A21, B1, B29, B30 and/or B31 or are missing. In some embodiments, an insulin molecule of the present disclosure is mutated such that one or more amino acids are replaced
30 (substituted) with their acidic forms. In some embodiments, an asparagine is replaced with aspartic acid or glutamic acid. In some embodiments, glutamine is replaced with aspartic acid or glutamic acid. In some embodiments, A21 may be an aspartic acid, B3 may be an aspartic acid, or both positions may contain an aspartic acid. One skilled in the art will recognize that it is possible to make any previously reported, or widely accepted mutations or modifications to
35 insulin that retains biological activity, and that such an insulin analogue can be used in embodiments of the present disclosure. In some embodiments, an insulin may be linked at any

5 position to a fatty acid, or acylated with a fatty acid at any amino group, including those on lysine side chains and the alpha-amino group on the N-terminus of insulin, and the fatty acid may include a C8, C9, C10, C11, C12, C14, C15, C16, C17, or C18 chain. In some embodiments, the fatty acid chain is 8-20 carbons long. In some embodiments, is an insulin detemir, in which a myristic acid is covalently conjugated to lysine at B29, and B30 is deleted
 10 or absent. In some embodiments, position B28 of the insulin molecule is lysine and the epsilon(ε)-amino group of this lysine is conjugated to a fatty acid.

In some embodiments, the N- or C-terminal end of the A- or B-chain of the modified insulin is ligated using a peptide ligase. In some embodiments, a polypeptide is added to the C-terminus of the insulin A- and/or B-chain or to the N-terminus of insulin A- and/or B-chain using a
 15 protein ligase, and in some embodiments thereof the ligase is chosen from sortases, butelases, Trypsiligases, Subtilisins, Peptiligases or enzymes having at least 75% homology to these ligases. In some embodiments, ligation is achieved through expressed protein ligation as described in: Muir TW, Sondhi D, Cole PA. "Expressed protein ligation: a general method for protein engineering." *Proc Natl Acad Sci U S A*. 1998; 95(12):6705–6710. In some
 20 embodiments, the polypeptide is linked to the modified insulin using Staudinger ligation, utilizing the Staudinger reaction and as described for example in Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. (2000). "Staudinger ligation: A peptide from a thioester and azide". *Org. Lett.* 2 (13): 1939–1941. In some embodiments, a polypeptide is conjugated to the modified insulin using Ser/Thr ligation as, for example, described in: Zhang Y, Xu C, Kam HY, Lee CL, Li X.
 25 2013, "Protein chemical synthesis by serine/threonine ligation." *Proc. Natl. Acad. Sci. USA*. 17:6657-6662. In some embodiments, the B-chain itself has less than 32 amino acids or 34 amino acids, and in some embodiments the insulin has 4 disulfide bonds instead of 3. There are disulfide bonds present in the A and B chains of insulin. For example, a disulfide bond exists between the cysteine at position 6 of SEQ ID NO:1 and the cysteine at position 11 of SEQ ID
 30 NO:1, a disulfide bond exists between the cysteine at position 7 of SEQ ID NO:1 and the cysteine at position 7 of SEQ ID NO:2, and a disulfide bond exists between the cysteine at position 20 of SEQ ID NO:1 and the cysteine at position 19 of SEQ ID NO:2.

In some embodiments, a modified insulin of the present disclosure comprises one or more mutations and/or chemical modifications including, but not limited to one of the
 35 following insulin molecules: N^{εB29}-octanoyl-Arg^{B0}Gly^{A21}Asp^{B3}Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-Arg^{A0}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-

- 5 Gly^{A21}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-Gly^{A21}Gln^{B3}Lys^{B28}Pro^{B30}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-Arg^{A0}Gly^{A21}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-Arg^{A0}Gly^{A21}Gln^{B3}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-Arg^{A0}Gly^{A21}Asp^{B3}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-myristoyl-Arg^{A0}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-
- 10 Gly^{A21}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-Gly^{A21}Gln^{B3}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-Arg^{A0}Gly^{A21}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB29}-palmitoyl-HI, N^{εB29}-myristoyl-HI, N^{εB28}-palmitoyl-Lys^{B28}Pro^{B29}-HI, N^{εB28}-myristoyl-Lys^{B28}Pro^{B29}-HI, N^{εB29}-palmitoyl-des(B30)-HI, N^{εB30}-myristoyl-Thr^{B29}Lys^{B30}-HI, N^{εB30}-palmitoyl-Thr^{B29}Lys^{B30}-HI, N^{εB29}-(N-palmitoyl-γ-glutamyl)-des(B30)-HI, N^{εB29}-(N-lithocolyl-γ-glutamyl)-des(B30)-HI, N^{εB29}-(ω-carboxyheptadecanoyl)-des(B30)-HI, N^{εB29}-(ω-carboxyheptadecanoyl)-HI, N^{εB29}-octanoyl-HI, N^{εB29}-myristoyl-Gly^{A21}Arg^{B31}Arg^{B31}-HI, N^{εB29}-myristoyl-Gly^{A21}Gln^{B3}Arg^{B31}Arg^{B32}-HI, N^{εB29}-myristoyl-Arg^{A0}Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-Arg^{A0}Gly^{A21}Gln^{B3}Arg^{B31}Arg^{B32}-HI, N^{εB29}-myristoyl-Arg^{A0}Gly^{A21}Asp^{B3}Arg^{B31}Arg^{B32}-HI, N^{εB29}-myristoyl-Arg^{B31}Arg^{B32}-HI, N^{εB29}-myristoyl-Arg^{A0}Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-
- 20 Gly^{A21}Gln^{B3}Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-Arg^{A0}Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-Arg^{A0}Gly^{A21}Gln^{B3}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-Arg^{A0}Gly^{A21}Gln^{B3}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-Arg^{A0}Gly^{A21}Asp^{B3}Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB28}-octanoyl-Lys^{B28}Pro^{B29}Arg^{B31}Arg^{B32}-HI, N^{εB29}-pentanoyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-hexanoyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{αA1}-heptanoyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-octanoyl-N^{αB1}-octanoyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-propionyl-N^{αA1}-propionyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{αA1}-acetyl-N^{αB1}-acetyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-formyl-N^{αA1}-formyl-N^{αB1}-formyl-Gly^{A21}Arg^{B31}Arg^{B32}-HI, N^{εB29}-formyl-des(B26)-HI, N^{αB1}-acetyl-Asp^{B28}-HI, N^{εB29}-propionyl-N^{αA1}-propionyl-N^{αB1}-propionyl-Asp^{B1}Asp^{B3}Asp^{B21}-HI, N^{εB29}-pentanoyl-
- 30 Gly^{A21}-HI, N^{αB1}-hexanoyl-Gly^{A21}-HI, N^{αA1}-heptanoyl-Gly^{A21}-HI, N^{εB29}-octanoyl-N^{αB1}-octanoyl-Gly^{A21}-HI, N^{εB29}-propionyl-N^{αA1}-propionyl-Gly^{A21}-HI, N^{αA1}-acetyl-N^{αB1}-acetyl-Gly^{A21}-HI, N^{εB29}-formyl-N^{αA1}-formyl-N^{αB1}-formyl-Gly^{A21}-HI, N^{εB29}-butyryl-des(B30)-HI, N^{αB31}-butyryl-des(B30)-HI, N^{αA1}-butyryl-des(B30)-HI, N^{εB29}-butyryl-N^{αB31}-butyryl-des(B30)-HI, N^{εB29}-butyryl-N^{αA1}-butyryl-des(B30)-HI, N^{αA1}-butyryl-N^{αB31}-butyryl-des(B30)-HI, N^{εB29}-butyryl-N^{αA1}-butyryl-N^{αB31}-butyryl-des(B30)-HI, Lys^{B28}Pro^{B29}-HI (insulin lispro), Asp^{B28}-HI (insulin aspart), Lys^{B3}Glu^{B29}-HI (insulin glulisine), Arg^{B31}Arg^{B32}-HI (insulin glargine), N^{εB29}-
- 35

- 5 myristoyl-des(B30)-HI (insulin detemir), Ala^{B26}-HI, Asp^{B1}-HI, Arg^{A0}-HI, Asp^{B1}Glu^{B13}-HI, Gly^{A21}-HI, Gly^{A21}Arg^{B31}Arg^{B32}-HI, Arg^{A0}Arg^{B31}Arg^{B32}-HI, Arg^{A0}Gly^{A21}Arg^{B31}Arg^{B32}-HI, des(B30)-HI, des(B27)-HI, des(B28-B30)-HI, des(B1)-HI, des(B1-B3)-HIN^{εB29}-tridecanoyl-des(B30)-HI, N^{εB29}-tetradecanoyl-des(B30)-HI, N^{εB29}-decanoyl-des(B30)-HI, N^{εB29}-dodecanoyl-des(B30)-HI, N^{εB29}-tridecanoyl-Gly^{A21}-des(B30)-HI, N^{εB29}-tetradecanoyl-Gly^{A21}-des(B30)-HI, N^{εB29}-decanoyl-Gly^{A21}-des(B30)-HI, N^{εB29}-dodecanoyl-Gly^{A21}-des(B30)-HI, N^{εB29}-tridecanoyl-Gly^{A21}Gln^{B3}-des(B30)-HI, N^{εB29}-tetradecanoyl-Gly^{A21}Gln^{B3}-des(B30)-HI, N^{εB29}-decanoyl-Gly^{A21}-Gln^{B3}-des(B30)-HI, N^{εB29}-dodecanoyl-Gly^{A21}-Gln^{B3}-des(B30)-HI, N^{εB29}-tridecanoyl-Ala^{A21}-des(B30)-HI, N^{εB29}-tetradecanoyl-Ala^{A21}-des(B30)-HI, N^{εB29}-decanoyl-Ala^{A21}-des(B30)-HI, N^{εB29}-dodecanoyl-Ala^{A21}-des(B30)-HI, N^{εB29}-tridecanoyl-Ala^{A21}-Gln^{B3}-des(B30)-HI, N^{εB29}-tetradecanoyl-Ala^{A21}Gln^{B3}-des(B30)-HI, N^{εB29}-decanoyl-Ala^{A21}Gln^{B3}-des(B30)-HI, N^{εB29}-dodecanoyl-Ala^{A21}Gln^{B3}-des(B30)-HI, N^{εB29}-tridecanoyl-Gln^{B3}-des(B30)-HI, N^{εB29}-tetradecanoyl-Gln^{B3}-des(B30)-HI, N^{εB29}-decanoyl-Gln^{B3}-des(B30)-HI, N^{εB29}-dodecanoyl-Gln^{B3}-des(B30)-HI, N^{εB29}-Z1-Gly^{A21}-HI, N^{εB29}-Z2-Gly^{A21}-HI, N^{εB29}-Z4-Gly^{A21}-HI, N^{εB29}-Z3-Gly^{A21}-HI, N^{εB29}-Z1-Ala^{A21}-HI, N^{εB29}-Z2-Ala^{A21}-HI, N^{εB29}-Z4-Ala^{A21}-HI, N^{εB29}-Z3-Ala^{A21}-HI, N^{εB29}-Z1-Gly^{A21}Gln^{B3}-HI, N^{εB29}-Z2-Gly^{A21}Gln^{B3}-HI, N^{εB29}-Z4-Gly^{A21}Gln^{B3}-HI, N^{εB29}-Z3-Gly^{A21}Gln^{B3}-HI, N^{εB29}-Z1-Ala^{A21}Gln^{B3}-HI, N^{εB29}-Z2-Ala^{A21}Gln^{B3}-HI, N^{εB29}-Z4-Ala^{A21}Gln^{B3}-HI, N^{εB29}-Z3-Ala^{A21}Gln^{B3}-HI, N^{εB29}-Z1-Gln^{B3}-HI, N^{εB29}-Z2-Gln^{B3}-HI, N^{εB29}-Z4-Gln^{B3}-HI, N^{εB29}-Z3-Gln^{B3}-HI, N^{εB29}-Z1-Glu^{B30}-HI, N^{εB29}-Z2-Glu^{B30}-HI, N^{εB29}-Z4-Glu^{B30}-HI, N^{εB29}-Z3-Glu^{B30}-HI, N^{εB29}-Z1-Gly^{A21}Glu^{B30}-HI, N^{εB29}-Z2-Gly^{A21}Glu^{B30}-HI, N^{εB29}-Z4-Gly^{A21}Glu^{B30}-HI, N^{εB29}-Z3-Gly^{A21}Glu^{B30}-HI, N^{εB29}-Z1-Gly^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z2-Gly^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z3-Gly^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z1-Ala^{A21}Glu^{B30}-HI, N^{εB29}-Z2-Ala^{A21}Glu^{B30}-HI, N^{εB29}-Z4-Ala^{A21}Gln^{B30}-HI, N^{εB29}-Z3-Ala^{A21}Glu^{B30}-HI, N^{εB29}-Z1-Ala^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z2-Ala^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z4-Ala^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z3-Ala^{A21}Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z1-Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z2-Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z4-Gln^{B3}Glu^{B30}-HI, N^{εB29}-Z3-Gln^{B3}Glu^{B30}-HI and where Z1 is tridecanoyl, Z2 is tetradecanoyl, Z3 is dodecanoyl, Z4 is decanoyl, and HI is human insulin.

In some embodiments, insulin includes one or more of the following mutations and/or chemical modifications: N^{εB28}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{αB1}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{αA1}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{εB28}-XXXXX-N^{αB1}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{εB28}-XXXXX-N^{αA1}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{αA1}-XXXXX-N^{αB1}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{εB28}-XXXXX-N^{αA1}-XXXXX-N^{αB1}-XXXXX-Lys^{B28}Pro^{B29}-HI, N^{εB29}-XXXXX-HI, N^{αB1}-

5 XXXXX-HI, N^{αA1}-XXXXX-HI, N^{εB29}-XXXXX-N^{αB1}-XXXXX-HI, N^{εB29}-XXXXX-N^{αA1}-XXXXX-HI, N^{αA1}-XXXXX-N^{αB1}-XXXXX-HI, N^{εB29}-XXXXX-N^{αA1}-XXXXX-N^{αB1}-XXXXX-HI, N^{εB29}-YYYYY-HI, N^{αB1}-YYYYY-HI, N^{αA1}-YYYYY-HI, N^{εB29}-YYYYY-N^{αB1}-YYYYY-HI, N^{εB29}-YYYYY-N^{αA1}-YYYYY-HI, N^{αA1}-YYYYY-N^{αB1}-YYYYY-HI, N^{εB29}-YYYYY-N^{αA1}-YYYYY-N^{αB1}-YYYYY-HI, N^{εB28}-YYYYY-Lys^{B28}Pro^{B29}-HI, N^{εB21}-YYYYY-Lys^{B28}Pro^{B29}-HI, N^{αA1}-YYYYY-Lys^{B28}Pro^{B29}-HI, N^{εB28}-YYYYY-N^{αB1}-YYYYY-Lys^{B28}Pro^{B29}-HI, N^{εB28}-YYYYY-N^{αA1}-YYYYY-Lys^{B28}Pro^{B29}-HI, N^{αA1}-YYYYY-N^{αB1}-YYYYY-Lys^{B28}Pro^{B29}-HI, N^{εB28}-YYYYY-N^{αA1}-YYYYY-N^{αB1}-YYYYY-Lys^{B28}Pro^{B29}-HI, and where YYYYY is one of acetyl or formyl and where XXXXX is one of: propionyl, butyryl, pentanoyl, hexanoyl, heptanoyl, octanoyl, nonanoyl or decanoyl and HI is human insulin.

15 In some embodiments, insulin may be conjugated through a reactive moiety that is naturally present within the insulin structure or is added prior to conjugation, including, for example, carboxyl or reactive ester, amine, hydroxyl, aldehyde, sulfhydryl, maleimidyl, alkynyl, azido, etc. moieties. Insulin naturally includes reactive alpha-terminal amine and epsilon-amine lysine groups to which NHS-ester, isocyanates or isothiocyanates can be covalently conjugated. In some embodiments, a modified insulin may be employed in which a suitable amino acid (e.g., a lysine or a non-natural amino acid) has been added or substituted into the amino acid sequence in order to provide an alternative point of conjugation in addition to the modified amino acids of the embodiments described herein. In some embodiments, the conjugation process may be controlled by selectively blocking certain reactive moieties prior to conjugation. In some embodiments, insulin may include any combination of modifications and the present disclosure also encompasses modified forms of non-human insulins (e.g., porcine insulin, bovine insulin, rabbit insulin, sheep insulin, etc.) that comprise any one of the aforementioned modifications. It is understood that some embodiments include these and other previously described modified insulins such as those described in United States patent numbers 30 5,474,978; 5,461,031; 4,421,685; 7,387,996; 6,869,930; 6,174,856; 6,011,007; 5,866,538; 5,750,497; 906,028; 6,551,992; 6,465,426; 6,444,641; 6,335,316; 6,268,335; 6,051,551; 6,034,054; 5,952,297; 5,922,675; 5,747,642; 5,693,609; 5,650,486; 5,547,929; 5,504,188; US 2015/0353619, including non-natural amino acids described or referenced herein and including such modifications to the non-human insulins described herein. It is also to be understood that 35 in some embodiments the insulin may be covalently conjugated to polyethylene glycol

5 polymers, such as polyethylene glycol polymers of no more than Mn 60,000, or covalently conjugated either through permanent or reversible bonds to albumin.

In some embodiments, a compound, such as a molecular conjugate, is conjugated to a chelator, and in some embodiments the chelator can be used to capture a radioactive payload, such as gallium 68, copper 64, lutetium 177, or actinium 225. In some embodiments, the
10 chelator is based on DOTA, NOTA, TETA, or 4-arm DOTA, and in some embodiments, the chelator can be linked to the peptide using a PEG linker through amide bonds to the chelator and to the peptide.

In some embodiments, the activity, bioavailability, solubility, isoelectric point, charge and/or hydrophobicity of the modified insulins can be controlled through chemical
15 modifications and/or as result of interaction of a small molecule such as a sugar with the compounds, such as a molecular conjugates, described herein which are either covalently linked or mixed with insulin.

In some embodiments one or more elements, functional groups, or atoms may be specifically omitted or excluded from a depicted structure (e.g., a terminal functional group
20 may be replaced by a hydrogen atom, or a linking group may be replaced by a bond), for example in Formulae FF1-FF224, and it will be understood that such omitted or excluded elements make these groups (structures) distinct and non-equivalent. For example, if an alternative version (variation) of a formula structure does not have a nitro group in R1 for B1 or B2, that variation is not equivalent (e.g., is structurally and chemically inequivalent) to a
25 structure that includes the nitro group, at least because the nitro group changes the pKa of B1 and B2 in physiological conditions and hence the overall affinity of Z1c for glucose.

Rotational Constrained Tethered Boron Conjugates.

In some embodiments, aromatic boron-containing compounds and/or aromatic boron-
30 containing groups are rotationally constrained tether boron conjugates. In some embodiments, rotationally constrained tether boron conjugates presented in this disclosure contain scaffolds that are rotationally hindered by disfavored steric interactions (e.g. gauche vs anti interactions of substituents), hindered rotation due to bond hybridization (e.g., cis- vs trans- amide rotations), or through rigid covalent bonds (e.g., (*E*) vs (*Z*) configurations for alkene moieties).
35 For example, formulae FF50 – FF62, FF116, and FF121-134 contain alkyl functionalities geminal (e.g., attached to the same atom) to the amine groups that are covalently conjugated to

5 the boronic acid functionalized moieties. Alkyl functionalities may limit the accessible dihedral angles and the rotation freedom around the C-C or C-X bond (commonly referred to as χ (chi) dihedral angles in amino acids). For example the hydroxyl sidechain on a serine residue can access dihedral angles of 60°, 180°, or 240° (-60°) with near equal distribution while the hydroxyl sidechain of threonine may only adopt dihedral angles of 180° or 240° (-60°). The
10 presence of a methyl group geminal to the hydroxy on threonine may provide steric bulk, creating unfavorable interactions when other bulky substituents are in a gauche conformation relative to the methyl. Formulae FF50 – FF62, FF116, and FF121-134 contain geminal alkyl substituents which may limit the accessible dihedral angles that the boron conjugated amines adopt, influencing adopted dihedral angles and placing the boronic functionalized groups closer
15 together and allowing for increased binding of the conjugates to target molecules such as proteins or sugars.

In some embodiments, the stereochemistry of isomeric structures (e.g., the stereochemistry of a compound, such as a molecular conjugate, for example within the Z1c moiety) may selectively increase the affinity of the conjugate (e.g., the Z1c moiety) for a
20 specific target diol, such as glucose. For example, in some embodiments one or more stereoisomers (e.g., *cis*- or *trans*-, (*R*) or (*S*), and (*E*) or (*Z*)) of Z1c may be selected to increase or decreases the affinity of Z1c (and the molecular architecture or conjugate as a whole) for glucose. In some embodiments, the *cis* form of Formulae FF1-FF224 is used when applicable (e.g., Z1c includes a structure having *cis* stereochemistry). In some embodiments, the *trans*
25 form of Formulae FF1-FF224 is used when applicable, e.g., when Formulae FF1-FF224 includes two stereocenters linked by a bond, (e.g., Z1c includes a structure having *trans* stereochemistry). In some embodiments, the *R* form of Formulae FF1-FF224 is used when applicable, e.g., when Formulae FF1-FF224 includes at least one stereocenter, (e.g., Z1c includes a structure having *R* stereochemistry). In some embodiments, the *S* form of Formulae
30 FF1-FF224 is used when applicable, e.g., when Formulae FF1-FF224 includes at least one stereocenter, (e.g., Z1c includes a structure having *S* stereochemistry). In some embodiments, the *S,S* form of Formulae FF1-FF224 is used when applicable, e.g., when Formulae FF1-FF224 includes two stereocenters linked by a bond, (e.g., Z1c includes a structure having *S,S* stereochemistry). In some embodiments, the *S,R* form of Formulae FF1-FF224 is used when
35 applicable, e.g., when Formulae FF1-FF224 includes two stereocenters linked by a bond, (e.g., Z1c includes a structure having *S,R* stereochemistry). In some embodiments, the *R,R* form of

5 Formulae FF1-FF224 is used when applicable, e.g., when Formulae FF1-FF224 includes two stereocenters linked by a bond, (e.g., Z1c includes a structure having *R,R* stereochemistry). In some embodiments, the *R,S* form of Formulae FF1-FF224 is used when applicable, e.g., when Formulae FF1-FF224 includes two stereocenters linked by a bond, (e.g., Z1c includes a structure having *R,S* stereochemistry). In some embodiments, a compound, such as a molecular
10 conjugate, includes one or more tautomers of a compound, such as a molecular conjugate, disclosed herein. In some embodiments, a compound, such as a molecular conjugate, includes one or more stereoisomer or a mixture of stereoisomers of a compound, such as a molecular conjugate, disclosed herein.

In some embodiments, a compound, such as a molecular conjugate, is covalently
15 conjugated to glucagon, GLP-1, GLP-2 or a variation of any of these (e.g., any variation with deletions, insertions and/or replacements of one or more amino acids). In some embodiments any suitable chemical modifications made to insulin discussed herein can be made to glucagon. In some embodiments the conjugate is mixed with a second or drug substance or one or more compounds chosen from: aminoethylglucose, aminoethylbimannose, aminoethyltrimannose, D-
20 glucose, D-galactose, D-Allose, D-Mannose, D-Gulose, D-Idose, D-Talose, N-Azidomannosamine (ManNAz) or N-Azidogalactoseamine (GalNAz) or N-azidoglucoseamine (GlcNAz), 2'-fluororibose, 2'-deoxyribose, glucose, sucrose, maltose, mannose, derivatives of these (e.g., glucosamine, mannosamine, methylglucose, methylmannose, ethylglucose, ethylmannose, etc.), sorbitol, inositol, galactitol, dulcitol, xylitol, arabitol and/or higher order
25 combinations of these (such as linear and/or branched bimannose, linear and/or branched trimannose), molecules containing cis-diols, catechols, tris, and DOPA molecules such as L-DOPA or L-3,4-dihydroxyphenylalanine.

Moreover, one skilled in the art will recognize that in some embodiments, one or more of suitable proteinogenic artificial amino acids can be used (included) in Z1a. For example, in
30 some embodiments one or more of the following artificial amino acids can be used based on the methods described in and referenced through, and the list of amino acids provided in: Liu, C. C.; Schultz, P. G. (2010). "Adding new chemistries to the genetic code." *Annual Review of Biochemistry* 79: 413–44. One skilled in the art will recognize that, in some embodiments, artificial amino acids can be incorporated by peptide synthesis in Z1a which is then covalently
35 conjugated to the drug or insulin, and these include the amino acids referenced herein as well as previously reported non-proteinogenic amino acids. In some embodiments, artificial amino

5 acids exist (e.g., may be included) in the insulin, and in some embodiments thereof, proteinogenic artificial amino acids can be incorporated through recombinant protein expression using suitable methods and approaches, including those described in United States patent and patent applications including: US 2008/0044854, US 8518666, US 8980581, US 2008/0044854, US 20140045261, US 2004/0053390, US 7229634, US 8236344, US 10 2005/0196427, US 2010/0247433, US 7198915, US 7723070, US 2002/0042097, US 2004/0058415, US 2008/0026422, US 2008/0160609, US 2010/0184193, US 2012/0077228, US 2014/025599, US 7198915, US 7632492, US 7723070, and other proteinogenic artificial amino acids may be introduced recombinantly using methods and approaches described in: US 7736872, US 7816320, US 7829310, US 7829659, US 7883866, US 8097702, US 8946148.

15 In some embodiments cyclic amino acids such as 3-hydroxyproline, 4-hydroxyproline, aziridine-2-carboxylic acid, azetidine-2-carboxylic acid, piperidine-2-carboxylic acid, 3-carboxy-morpholine, 3-carboxy-thiomorpholine, 4-oxaproline, pyroglutamic acid, 1,3-oxazolidine-4-carboxylic acid, 1,3-thiazolidine-4-carboxylic acid, 3-thiaproline, 4-thiaproline, 3-selenoproline, 4-selenoproline, 4-ketoproline, 3,4-dehydropyrolidine, 4-aminoproline, 4-fluoroproline, 4,4-difluoroproline, 4-chloroproline, 4,4-dichloroproline, 4-bromoproline, 4,4-dibromoproline, 4-methylproline, 4-ethylproline, 4-cyclohexyl-proline, 3-phenylproline, 4-phenylproline, 3,4-phenylproline, 4-azidoproline, 4-carboxyproline, α -methylproline, α -ethylproline, α -propylproline, α -allylproline, α -benzylproline, α -(4-fluorobenzyl)-proline, α -(2-chlorobenzyl)-proline, α -(3-chlorobenzyl)-proline, α -(2-bromobenzyl)-proline, α -(4-bromobenzyl)-proline, α -(4-methylbenzyl)-proline, α -(diphenylmethyl)-proline, α -(naphthylmethyl)-proline, D-proline, or S-homoproline, (2S, 4S)-4-fluoro-L-proline, (2S, 4R)-4-fluoro-L-proline, (2S)-3,4-dehydro-L-proline, (2S, 4S)-4-hydroxy-L-proline, (2S, 4R)-4-hydroxy-L-proline, (2S,4S)-4-azido-L-proline, (2S)-4,4-difluoro-L-proline, (2S)-azetidine-2-carboxylic acid, (2S)-piperidine-2-carboxylic acid, or (4R)-1,3-thiazolidine-4-carboxylic acid
25 can be used in the molecular architecture that is conjugated to insulin.

It is to be understood that in some embodiments, a specific orientation of amino acids is achieved using, for example, methods of Albericio, F. (2000). *Solid-Phase Synthesis: A Practical Guide* (1 ed.). Boca Raton: CRC Press. p. 848. In some embodiments a compound, such as a molecular conjugate, can bind to a diol, a catechol, a hexose sugar, glucose, xylose,
35 fucose, galactosamine, glucosamine, mannosamine, galactose, mannose, fructose, galacturonic acid, glucuronic acid, iduronic acid, mannuronic acid, acetyl galactosamine, acetyl

5 glucosamine, acetyl mannosamine, acetyl muramic acid, 2-keto-3-deoxy-glycero-galacto-
nononic acid, acetyl neuraminic acid, glycolyl neuraminic acid, a neurotransmitter, dopamine,
or a disaccharide or polymer of saccharides or diols.

In some embodiments modifications or intermediates may include the use of an N-methyliminodiacetic acid (MIDA) group to make a MIDA conjugated boronate or a MIDA
10 boronate; such modifications can be used during preparation of boronates towards the final
structures of use (e.g., in embodiments of methods for preparing the conjugates described
herein). In some embodiments boronic acid pinacol esters are used towards the final structures
wherein the pinacol group can be readily removed using standard techniques by one skilled in
the art. The MIDA-protected boronate esters are easily handled, stable under air, compatible
15 with chromatography, and unreactive under standard anhydrous cross-coupling conditions and
easily deprotected at room temperature under mild aqueous basic conditions such as 1M NaOH,
or even NaHCO₃, or as described by Lee, S. J. et al. (2008). *J. Am. Chem. Soc.* 130:466.

The biological mechanism by which wild type insulin binds to the insulin receptor is
previously reported in Menting, J.G. et al. (2013). *Nature* 493, 241-245; and Menting, J.G. et
20 al. (2014). "Protective hinge in insulin opens to enable its receptor engagement." *Proc. Natl.*
Acad. Sci. U.S.A. 111, E3395-3404. The activity of such insulin can be measured using any
suitable technique, for example, by using *in vitro* insulin receptor binding with TyrA14-¹²⁵I
human insulin as tracer and utilizing antibody binding beads with an insulin receptor
monoclonal antibody. In some embodiments, animal models can be used for *in vivo* assessment
25 of insulin activity during glucose challenge using methods that are known to one skilled in the
art. In some embodiments, a compound, such as a molecular conjugate, is partially or fully
expressed along with a recombinant protein of interest such as insulin. The processes for
expression of insulin in *E. coli* are known and can be easily performed by one skilled in the art
e.g., by using the procedures outlined in Jonasson (1996). *Eur. J. Biochem.* 236:656-661;
30 Cowley (1997). *FEBS Lett.* 402:124- 130; Cho (2001). *Biotechnol. Bioprocess Eng.* 6: 144-
149; Tikhonov (2001). *Protein Exp. Pur.* 21: 176-182; Malik (2007). *Protein Exp. Pur.* 55:
100-111; and Min (2011). *J. Biotech.* 151:350-356. In the most common process, the protein is
expressed as a single-chain proinsulin construct with a fission protein or affinity tag. The
compound, such as a molecular conjugate, that includes Z1a when Z1a is present can be
35 expressed as part of proinsulin, then modified chemically to conjugate, through amide linkages,
to structures of interest. This approach provides good yield and reduces experimental

5 complexity by decreasing the number of processing steps and allows refolding in a native-like insulin, see for example, Jonasson, *Eur. J. Biochem.* 236:656-661 (1996); Cho, *Biotechnol Bioprocess Eng.* 6: 144- 149 (2001); Tikhonov, *Protein Exp. Pur.* 21 : 176-182 (2001); Min, *J. Biotech.* 151 :350-356 (201 1)). When expressed in *E. coli*, proinsulin is usually found in inclusion bodies and can be easily purified by one skilled in the art.

10 In some embodiments, a compound, such as a molecular conjugate, may be formulated for injection. For example, it may be formulated for injection into a subject, such as a human. In some embodiments, the composition may be a pharmaceutical composition, such as a sterile, injectable pharmaceutical composition. In some embodiments, the composition may be formulated for subcutaneous injection. In some embodiments, the composition is formulated

15 for transdermal, intradermal, transmucosal, nasal, inhalable or intramuscular administration. In some embodiments, the composition may be formulated in an oral dosage form or a pulmonary dosage form. Pharmaceutical compositions suitable for injection may include sterile aqueous solutions containing, for example, sugars, polyalcohols such as mannitol and sorbitol, phenol, *meta*-cresol, and sodium chloride and dispersions may be prepared in glycerol, liquid

20 polyethylene glycols, and mixtures thereof and in oils and the carrier can, for example, be a solvent or dispersion medium containing, for example, water, saccharides, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. One skilled in the art recognizes that specific formulations can be developed to best suit the application and method of use of the molecular architectures of the invention.

25 General considerations in the formulation and manufacture of pharmaceutical compositions, routes of administering and including suitable pharmaceutically acceptable carriers may be found, for example, in *Remington's Pharmaceutical Sciences*, 19th ed., Mack Publishing Co., Easton, Pa., 1995. In some embodiments, the pharmaceutical composition may include zinc, *e.g.*, Zn^{2+} along with insulin if the compound, such as a molecular conjugate, comprises insulin.

30 Such zinc formulations are, for example, described in Unites States Patent No. 9,034,818. For example, the pharmaceutical composition may comprise zinc at a molar ratio to the modified insulin of about M:N where M is 1-11 and N is 6-1. In some embodiments, such modified insulins may be stored in a pump, and that pump being either external or internal to the body releases the modified insulins. In some embodiments, a pump may be used to release a constant

35 amount of modified insulin wherein the insulin is glucose responsive and can automatically adjust activity based on the levels of glucose in the blood and/or the release rate from the

5 injection site. In some embodiments, the compositions may be formulated in dosage unit form for ease of administration and uniformity of dosage. In some embodiments, the pharmaceutical composition may further include a second insulin type to provide fast-acting or basal-insulin in addition to the effect afforded by the molecular architecture. In some embodiments, a compound, such as a molecular conjugate, is injected separately from insulin but modulates the activity of insulin by binding to insulin, and in some embodiments this activity change is dependent on glucose.

In some embodiments, the pharmaceutical composition comprises one or more of the compounds disclosed herein and at least one additional component selected from pharmaceutically acceptable carriers, pharmaceutically acceptable vehicles, and pharmaceutically acceptable excipients. In some embodiments, the pharmaceutical composition comprises a compound of Formula I and at least one additional component selected from pharmaceutically acceptable carriers, pharmaceutically acceptable vehicles, and pharmaceutically acceptable excipients.

In some embodiments the present disclosure includes compounds that can be part of a kit, wherein the kit includes a compound, such as a molecular conjugate, such as a compound comprising modified insulin, as well as a pharmaceutically acceptable carrier, and for injections may include a syringe or pen. In some embodiments, a kit may include a syringe or pen that is pre-filled with a pharmaceutical composition that includes the compound, such as a molecular conjugate, with a liquid carrier. Alternatively, a kit may include a separate container such as a vial with a pharmaceutical composition that includes the compound, such as a molecular conjugate, with a dry carrier and an empty syringe or pen. In some embodiments, such a kit may include a separate container that has a liquid carrier, which can be used to reconstitute a given composition that can then be taken up into the syringe or pen. In some embodiments, a kit may include instructions. In some embodiments, the kit may include blood glucose measuring devices that either locally or remotely calculate an appropriate dose of the modified insulin that is to be injected at a given point in time, or at regular intervals. Such a dosing regimen is unique to the patient and may, for example, be provided as instruction to program a pump either by a person or by a computer. The kit may include an electronic device which transfers blood glucose measurements to a second computer, either locally or elsewhere (for example, in the cloud) which then calculate the correct amount of compound, such as a

5 molecular conjugate, comprising, e.g., a modified insulin that needs to be used by the patient at a certain time.

In some embodiments, the invention relates to a method for treating a disease or condition in a subject, comprising administering to the subject a composition including a compound, such as a molecular conjugate, described herein. In some embodiments, the disease
10 or condition may be hyperglycemia, type 2 diabetes, impaired glucose tolerance, type 1 diabetes, obesity, metabolic syndrome X, or dyslipidemia, diabetes during pregnancy, pre-diabetes, Alzheimer's disease, MODY 1, MODY 2 or MODY 3 diabetes, mood disorders and psychiatric disorders. It will be appreciated that this combination approach may also be used with insulin resistant patients who are receiving an insulin sensitizer or a secondary drug for
15 diabetes (such as, for example, a biguanide such as metformin, a glitazone) or/and an insulin secretagogue (such as, for example, a sulfonylurea, GLP-1, exendin-4 etc.) or amylin.

In some embodiments, a compound, such as a molecular conjugate, of the present disclosure may be administered to a patient who is receiving at least one additional therapy or taking at least one additional drug or therapeutic protein. In some embodiments, the at least one
20 additional therapy is intended to treat the same disease or disorder as the administered compound, such as a molecular conjugate. In some embodiments, the at least one additional therapy is intended to treat a side-effect of the compound, such as a molecular conjugate, or as an adjuvant. The timeframe of the two therapies may differ or be the same; they may be administered on the same or different schedules as long as there is a period when the patient is
25 receiving a benefit from both therapies. The two or more therapies may be administered within the same or different formulations as long as there is a period when the patient is receiving a benefit from both therapies. Any of these approaches may be used to administer more than one anti-diabetic drug to a subject.

In some embodiments a therapeutically effective amount of the compound, such as a
30 molecular conjugate, which is sufficient amount to treat (meaning for example to ameliorate the symptoms of, delay progression of, prevent recurrence of, or delay onset of) the disease or condition at a reasonable benefit to risk ratio will be used. In some embodiments, this may involve balancing of the efficacy and additional safety to toxicity. By additional safety, it is meant that, for example, the compound, such as a molecular conjugate, can be responsive to
35 changes in blood glucose levels or level of other molecules, even when the patient is not actively monitoring the levels of that molecule, such as blood glucose levels at a given

5 timeframe, for example during sleep. In some embodiments, therapeutic efficacy and toxicity may be determined by standard pharmacological procedures in cell cultures or in vivo with experimental animals, and for example measuring ED₅₀ and LD₅₀ for therapeutic index of the drug. In some embodiments, the average daily dose of insulin with the molecular architecture is in the range of 5 to 400 U, (for example 30-150 U where 1 Unit of insulin is about 0.04 mg). In
10 some embodiments, an amount of compound, such as a molecular conjugate, with these insulin doses is administered on a daily basis or bi-daily basis or by every three days or by every 4 days. In some embodiments the basis is determined by an algorithm, which can be computed by a computer. In some embodiments, an amount of compound, such as a molecular conjugate, with 5 to 10 times these doses is administered on a weekly basis or at regular intervals. In some
15 embodiments, an amount of conjugate with 10 to 20 times these doses is administered on a bi-weekly basis or at regular intervals. In some embodiments, an amount of compound, such as a molecular conjugate, with 20 to 40 times these doses is administered on a monthly basis or at regular intervals. In some embodiments, the C-terminus of the A-chain of insulin may be further extended with a peptide (amino acid sequence) including 1-20 amino acid residues. In
20 some embodiments the insulin analogue is a desB30 insulin.

In some embodiments, Z1a is an amino acid or a peptide. In at least some embodiments, the Z1a includes (is composed of) 1-50 amino acid residues, for example, 1 residue, 50 residues, or any intermediate number of residues (such as e.g., 10, 15, 25, 30, 42 residues, etc.). In some embodiments, the Z1a includes 1-15 amino acids. In at least some embodiments, the
25 peptide Z1a includes 1-8 amino acids. In some embodiments, Z1a includes 5 to 6 amino acids. In some embodiments, Z1a comprises at least one amino acid independently selected from the : Alanine (A), Asparagine (N), Glutamine (Q), Threonine (T), Methionine (M), Histidine (H), Cysteine (C), Valine (V), Isoleucine (I), Lysine (K), and Leucine (L), and the rest of the amino acids each independent selected from any of the twenty naturally occurring amino acids. In
30 some embodiments, Z1a may include diaminopropionic acid, diaminobutyric acid, or ornithine. In some embodiments, Z1a includes 1 to 5 lysines (K). In some embodiments, Z1a includes 1 to 3 K amino acids. In some embodiments Z1a includes 5 to 6 amino acids and at least one or more amino acids are K. In some embodiments, Z1a includes 5 to 6 amino acids and 1 to 3 amino acids are K. In some embodiments, Z1a is selected from any of KA, KD, KE, KF, KG,
35 KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI,

5 KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, 10 KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, 15 KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75-24014, 24037-24046. In some embodiments, Z1a is selected from KSNAPQK (SEQ ID NO:24037), KNASPQK (SEQ ID NO:24038), KLWAVK (SEQ ID NO:24039), KGARLK (SEQ ID NO:24040), ADKKTLLN (SEQ ID NO:24041), KGSHK (SEQ ID NO:4238), KNSTK (SEQ ID NO:5085), GKNSTK (SEQ ID NO:13989), GKGSHK (SEQ ID NO:13198), GSHKGSHK (SEQ ID NO:24042), GKPSHKP (SEQ ID NO:24043), GKGPSK (SEQ ID NO:24044), GKGSKK (SEQ ID NO:24045), and GKKPGKK (SEQ ID NO:24046).

In some embodiments Z1a is appended to the N-terminus and/or C-terminus, and/or 25 inserted into the sequence of the A-chain or B-chain of insulin.

Compounds

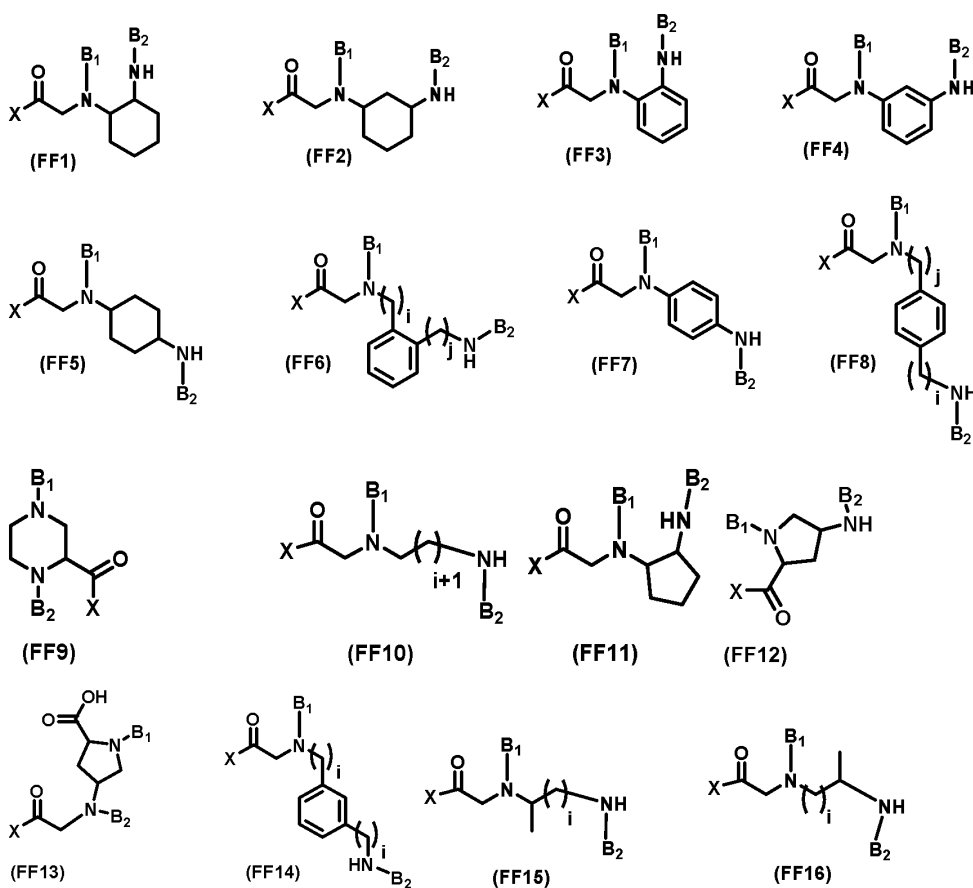
In some embodiments, the present disclosure provides a compound comprising X1 and one or more Z1c, or a tautomer, stereoisomer or a mixture of stereoisomers, or a 30 pharmaceutically acceptable salt, or hydrate, or isotopic derivative thereof, wherein: X1 comprises:

- (i) NH₂ or OH (for example, X1 is NH₂ or OH);
- (ii) a drug substance comprising an amine;
- (iii) a drug substance that is covalently conjugated to an amine containing linker; or 35 (iv) an amine configured to be covalently conjugated to a drug substance;

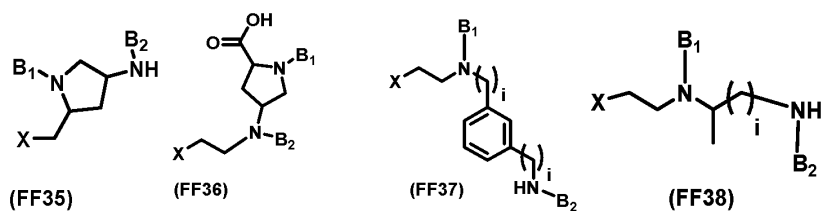
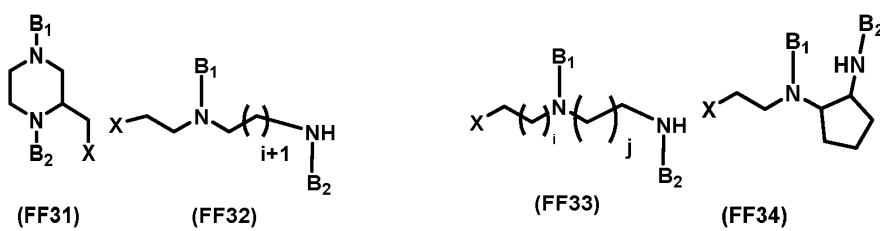
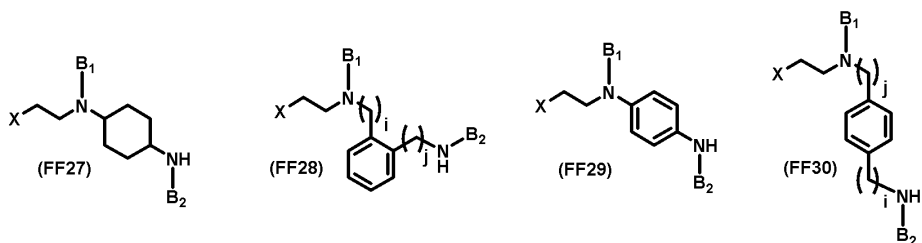
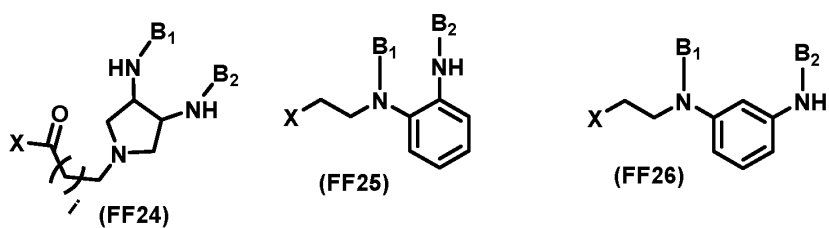
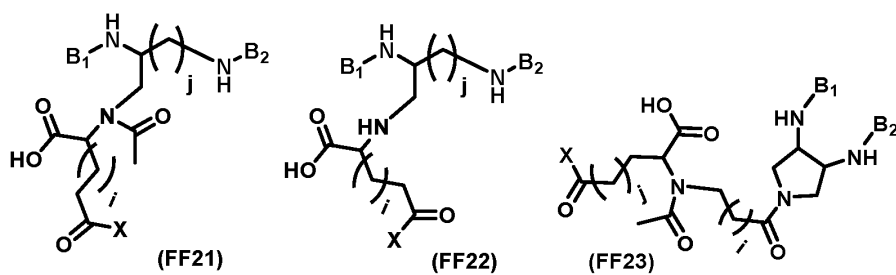
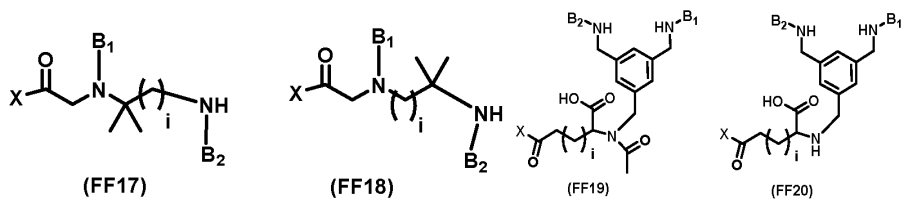
- 5 wherein each Z1c is independently selected from Formulae FF1-FF224; and wherein each Z1c is covalently conjugated, directly or indirectly, to an amine in X1 or to OH when X1 is OH.

In some embodiments, Z1c is independently selected from Formulae FF1-FF48,
Formulae FF49-FF88, FF89-FF112, FF113-FF136, FF137-FF160, FF161-FF164, FF165-
10 FF166, FF167-FF192, FF193-FF209, and FF210-FF224.

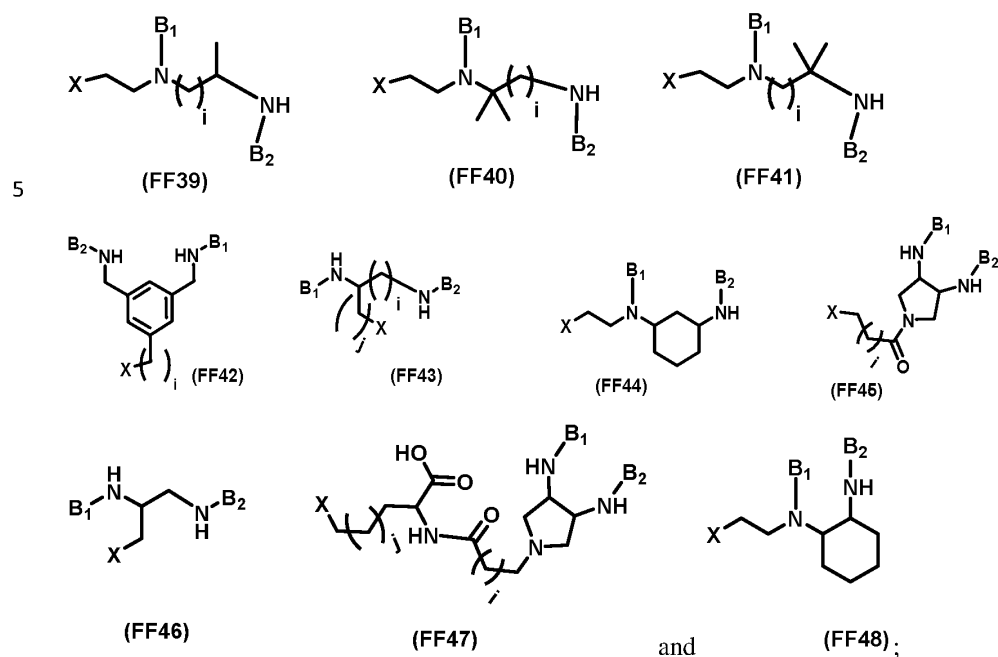
Formulae FF1-FF48 are:



5



10

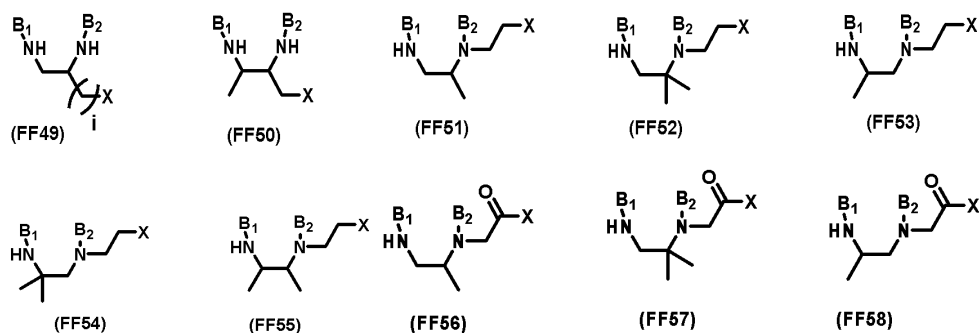


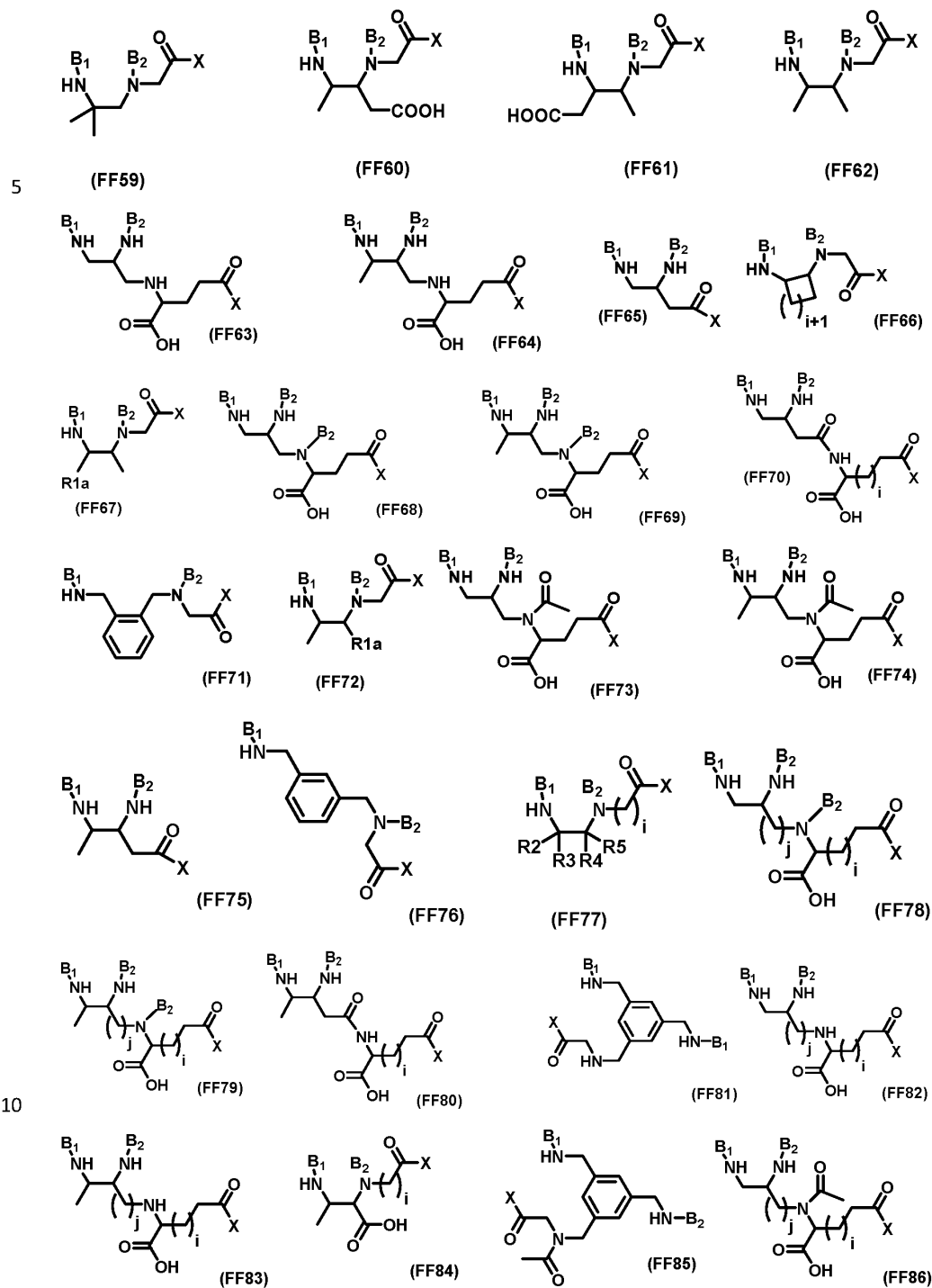
wherein X represents a point of covalent attachment either directly to an amine in X1 or
 10 to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

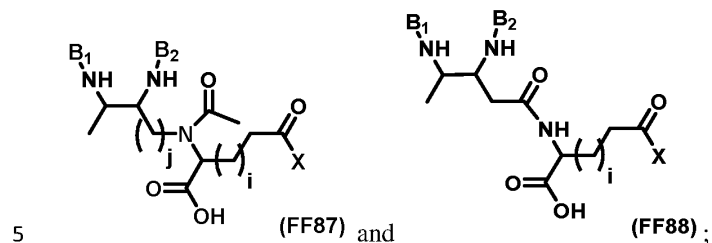
i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7; and B₁ and B₂, which may be identical or different, each
 independently represents an aromatic boron-containing group.

15 Formulae FF49-FF88 are:





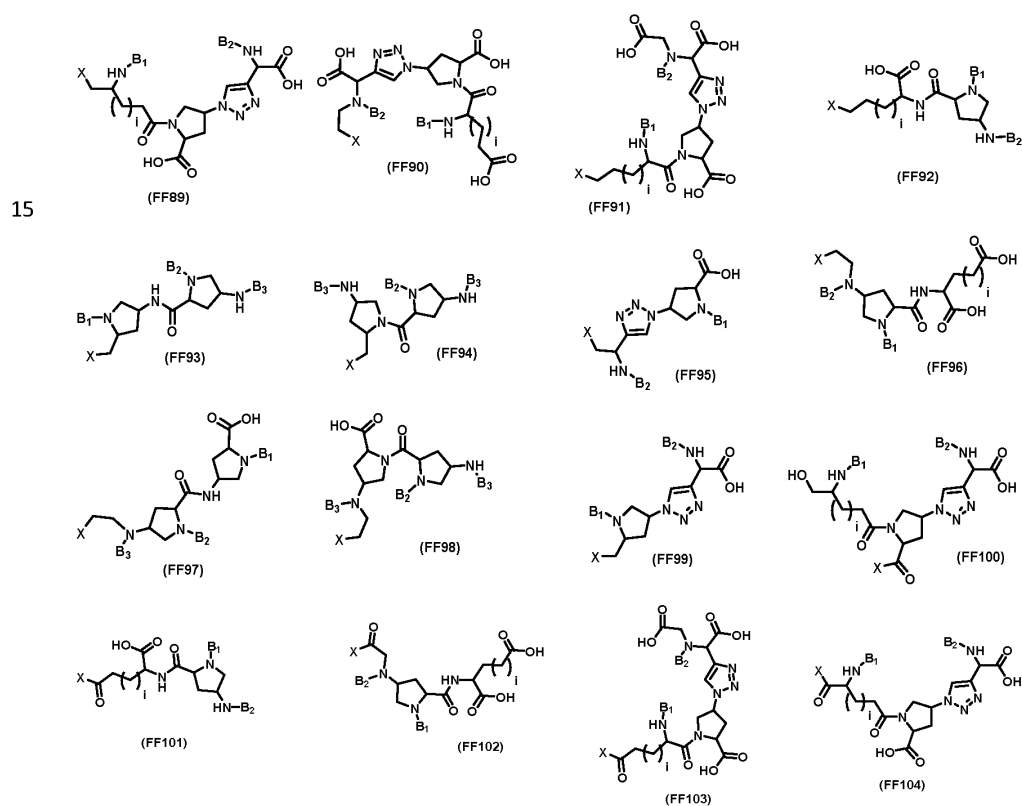


wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

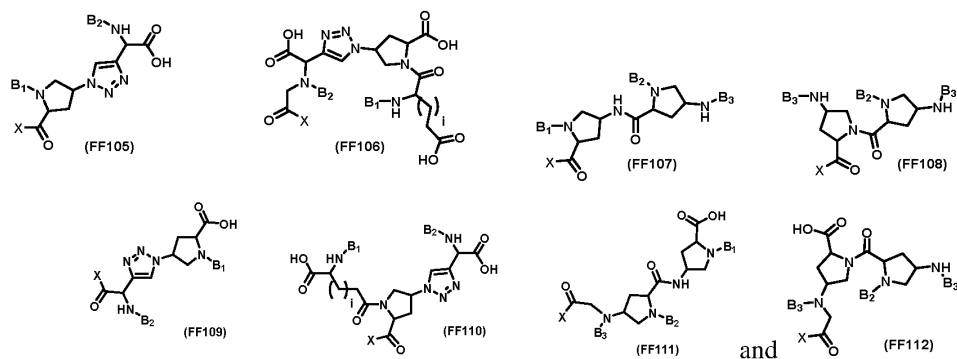
i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7; R1a is selected from COOH, CH₃, H, and OH; R2, R3, R4 and R5 are each independently selected from CH₃, H, OH, and COOH, and at least one of R2, R3, R4 and R5 is CH₃ or OH; and B₁ and B₂, which may be identical or different, are each independently an aromatic boron-containing group.

Formulae FF89-FF112 are:



5

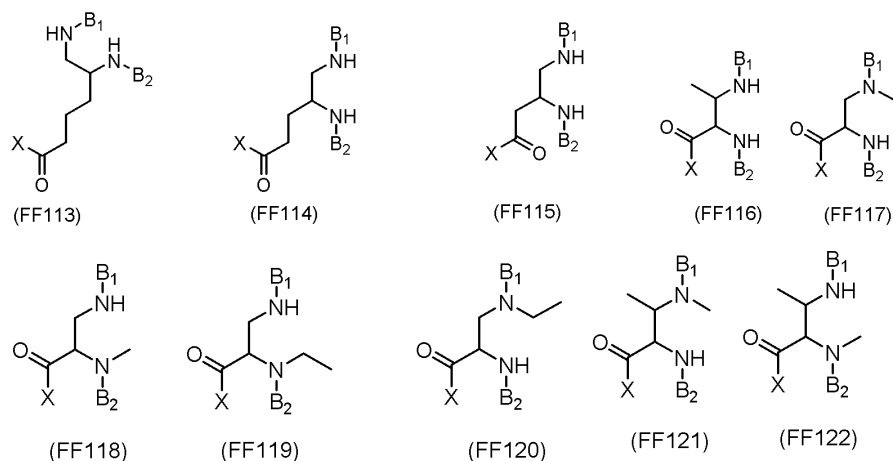


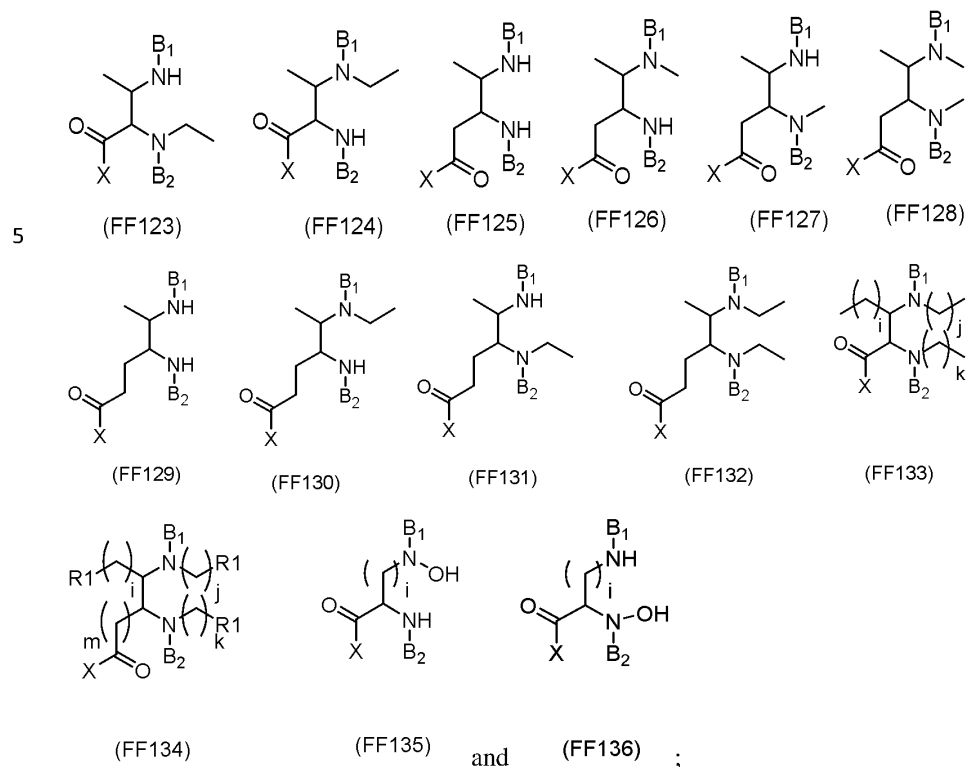
wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

i is 1, 2, 3, 4, 5, 6, or 7; and

- 10 B₁, B₂ and B₃, which may be identical or different, are each independently an aromatic boron-containing group, a carboxylic acid derivative, or a H, wherein in each FF89-FF112 structure containing B₁, B₂ and B₃ groups, at least two of the B₁, B₂ and B₃ groups are independently an aromatic boron-containing group.

- 15 Formulae FF113-FF136 are:





wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

10 i is 0, 1, 2, 3, 4, 5, 6, or 7;

j is 0, 1, 2, 3, 4, 5, 6, or 7;

k is 0, 1, 2, 3, 4, 5, 6, or 7;

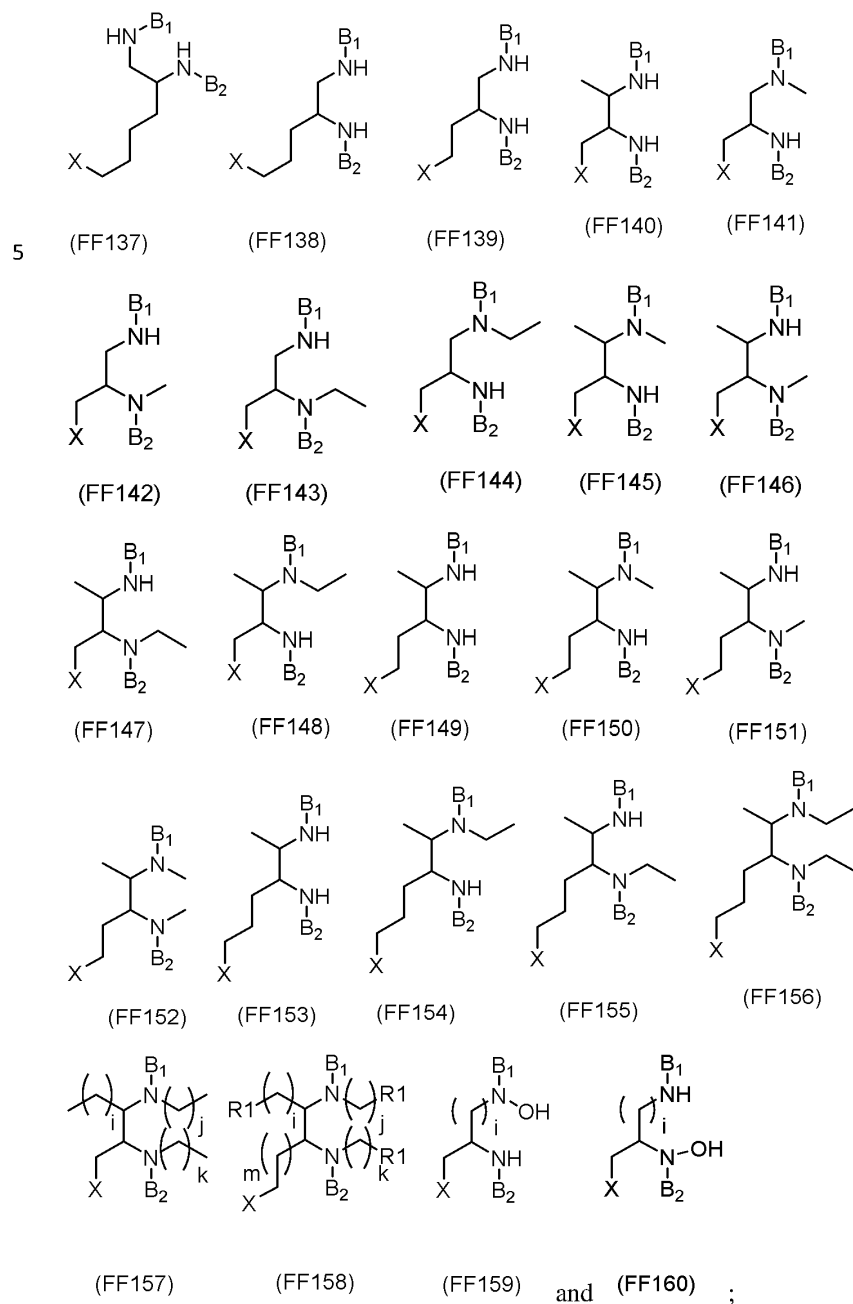
m is 0, 1, 2, 3, 4, 5, 6, or 7;

wherein $i+j+k+m$ is greater than 0;

15 each R1 is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; and

20 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group.

Formulae FF137-FF160 are:



10

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

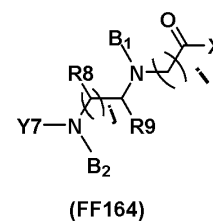
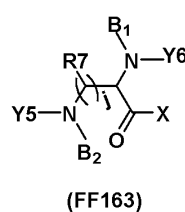
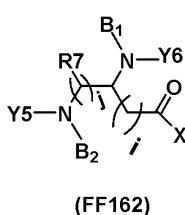
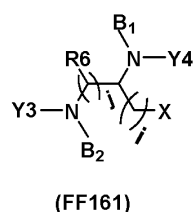
- 5 i is 0, 1, 2, 3, 4, 5, 6, or 7;
 j is 0, 1, 2, 3, 4, 5, 6, or 7;
 k is 0, 1, 2, 3, 4, 5, 6, or 7;
 m is 0, 1, 2, 3, 4, 5, 6, or 7;
 wherein $i+j+k+m$ is greater than 0;

10

each R1 is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; and

- 15 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group. In one embodiment, B₁ and B₂ may be identical or different.

Formulae FF161-FF164 are:



20

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

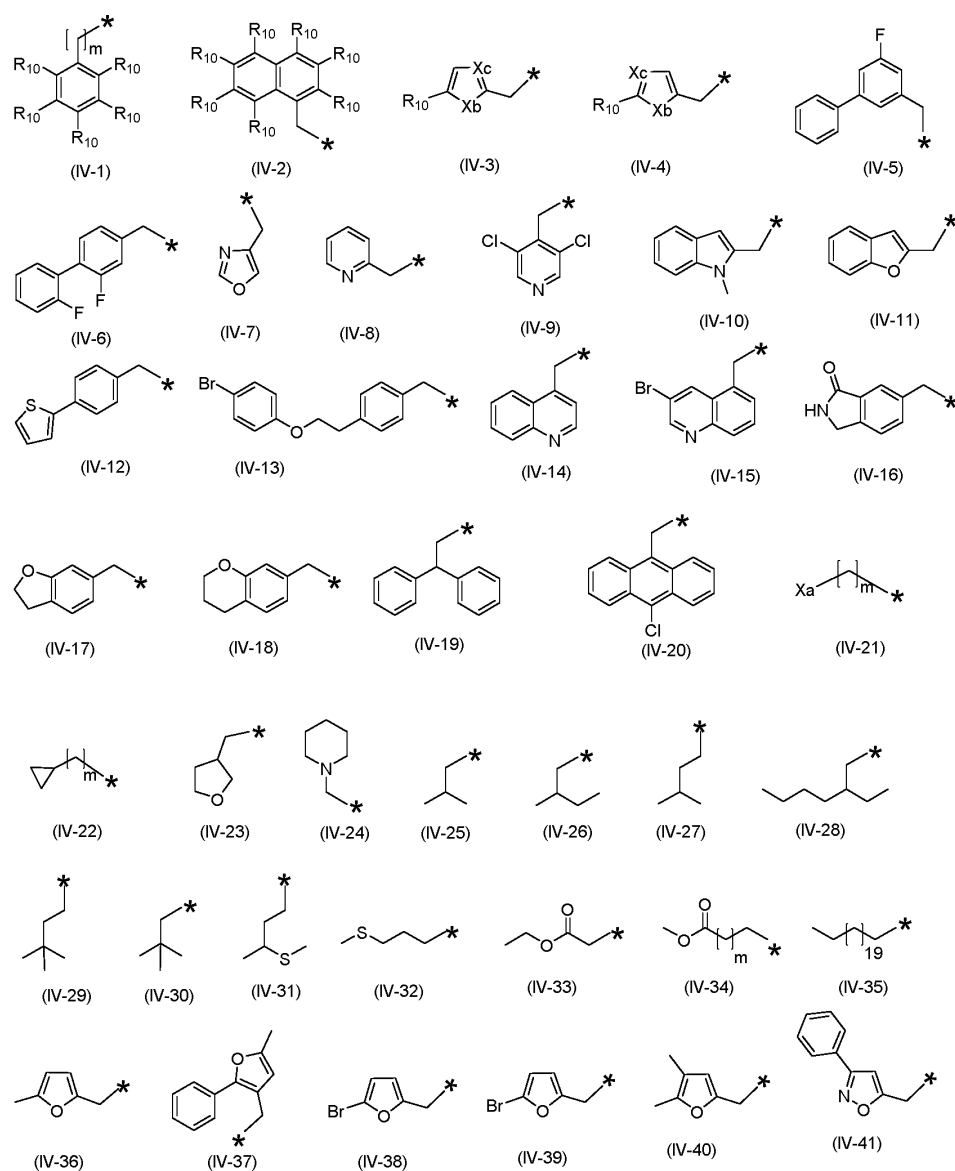
- i is 1, 2, 3, 4, or 5 (e.g., 1, 2, 3, or 5);
 j is 1, 2, 3, 4, or 5 (e.g., 1, 2, 3, or 5);
 25 each R6, R7, R8, and R9 for different values of j is independently selected from H, CF₃, CH₃, CHF₂, and (CH₂)_mCH₃, wherein m is 1, 2, 3, 4, or 5;

Y3, Y4, Y5, Y6 and Y7 are each independently selected from H, CH₂—X₄, and Formulae IV-1 to IV-135;

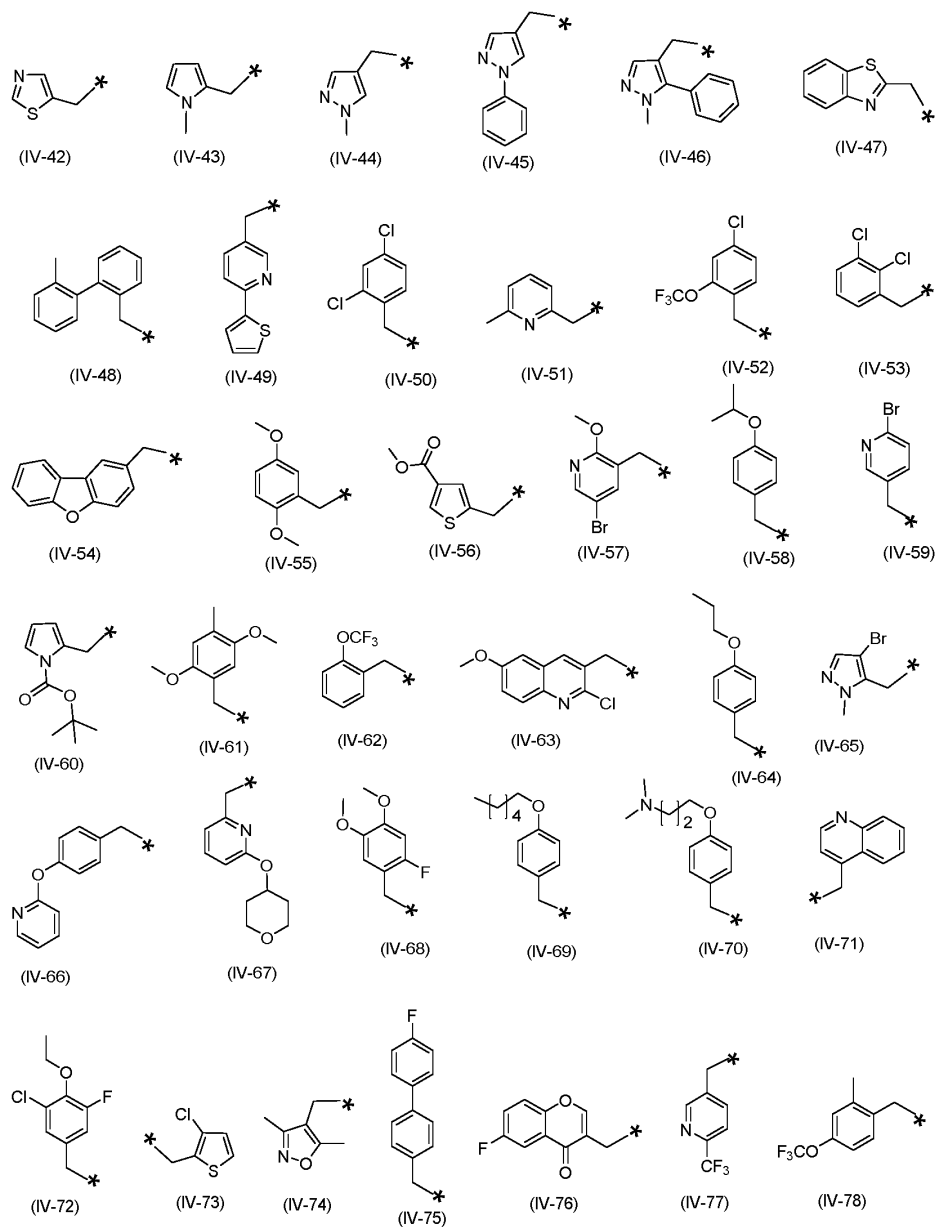
- wherein X₄ is selected from -COOH, -(CH₂)_mCOOH, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each X₄ optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; wherein m is 1, 2, 3, 4, or 5;
- 30

- 5 wherein at least one of Y5, Y6 and Y7 in Formulae FF162 and FF163 is not H and at least one of Y7, R8 and R9 in FF164 is not H; and

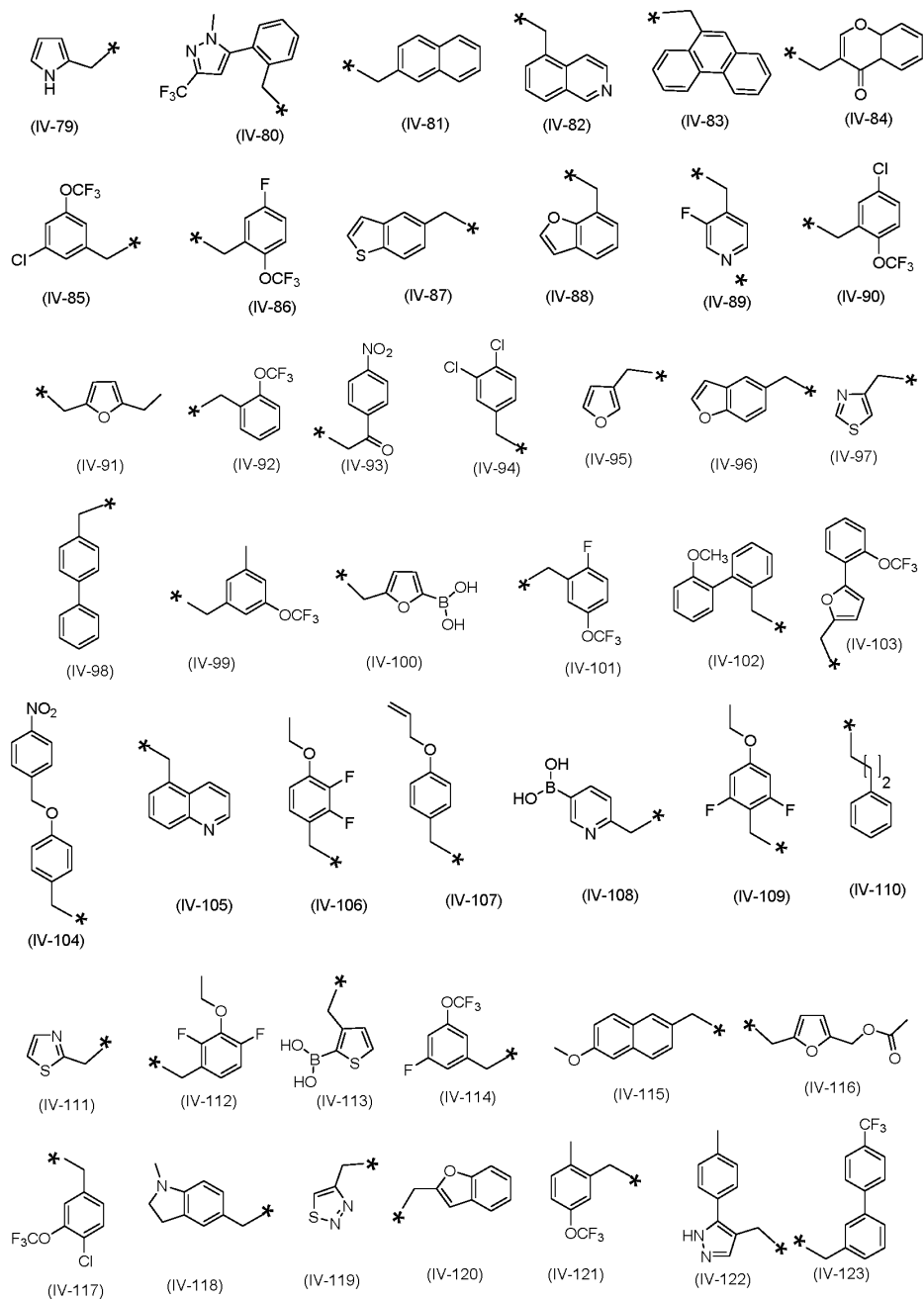
wherein Formulae IV-1 to IV-135 are:

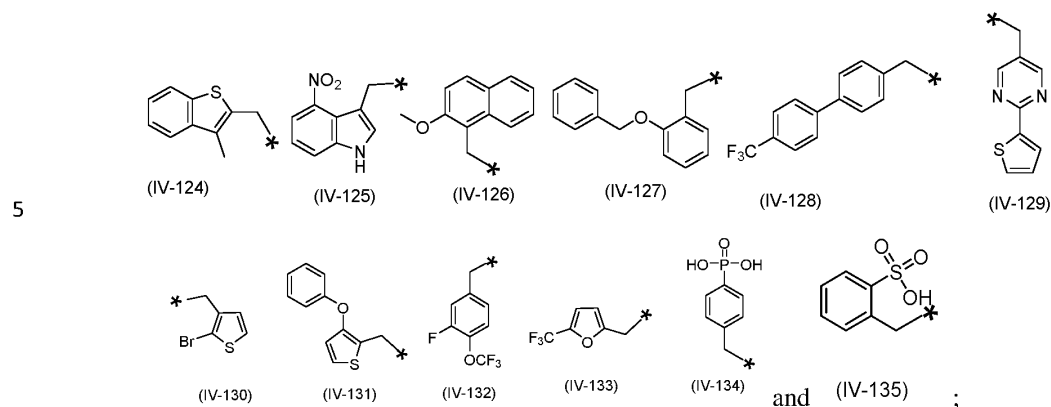


5



5





wherein Xa represents CH=O, CHF₂, CF₃, CH₂SH, COOH, CH₂OH, CH₂NO₂, CH₂NH₂, CH₃, C(CH₃)₃, CH(CH₃)₂, CH((CH₂)₃ CH₃)₂, or CH(CH₂ CH₃)₂;

Xb represents O, NH, CH₂, or S;

10 Xc represents CH or N;

each R₁₀ is independently selected from H, F, Cl, Br, CH₃, CF₃, CH=O, OH, COOH, and (CH₂)_nCH₃,

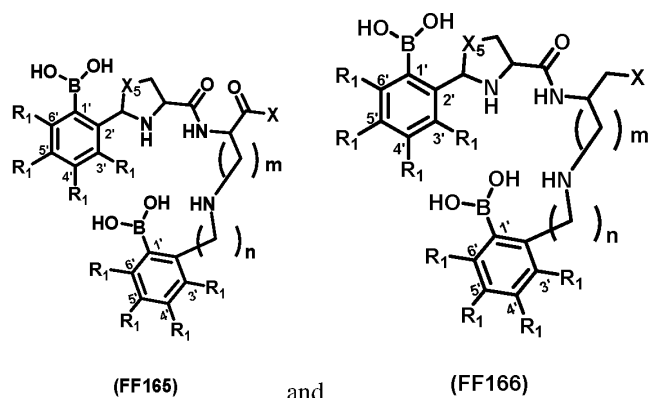
m is 1, 2, 3, 4, or 5; and n is 1, 2, 3, 4, or 5;

15 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group; and

* in Formulae IV-1 to IV-135 represents the point of attachment to corresponding Formulae FF161-164.

In some embodiments, when j is 4, X is not NH₂ for FF163.

20 Formulae FF165-FF166 are:



5

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

m is 1, 2, 3, 4, 5, 6, or 7;

n is 1, 2, 3, 4, 5, 6, or 7;

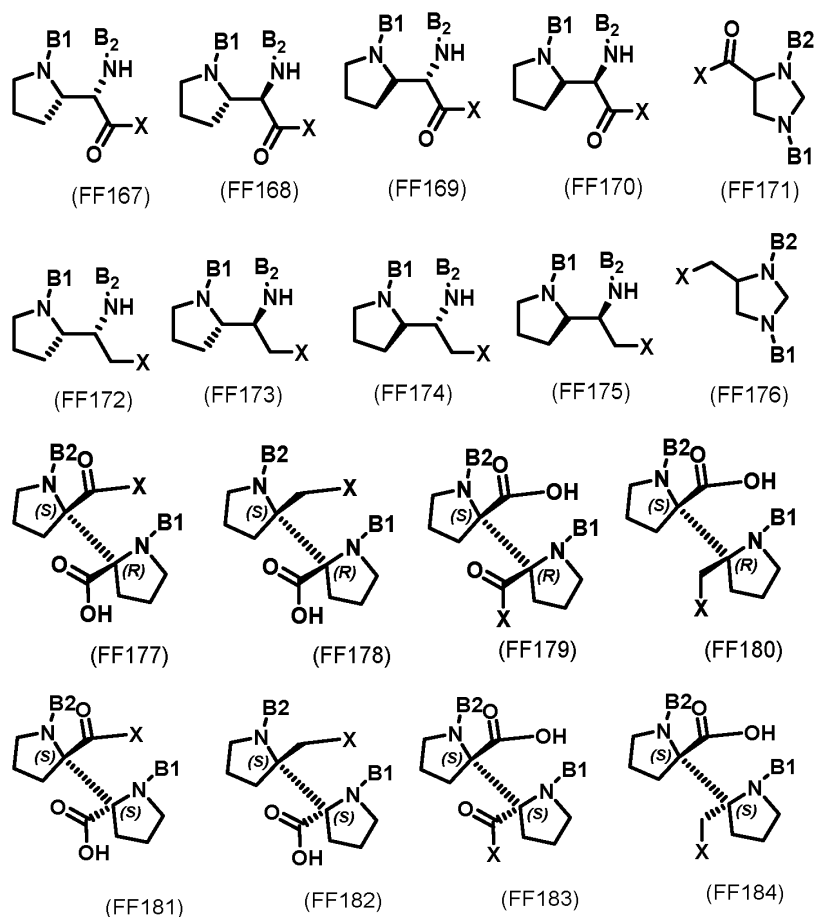
10

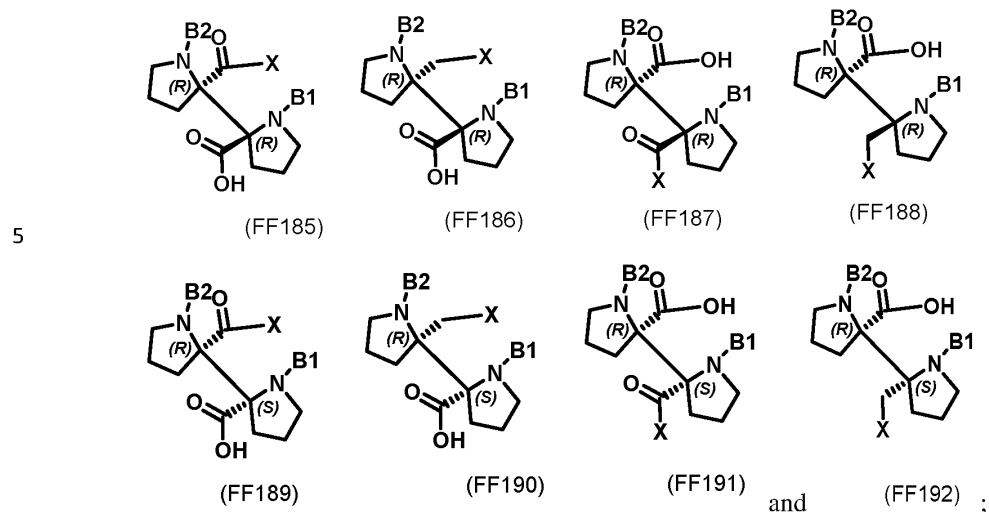
X5 is S, O, or NH; and

each R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7.

15

Formulae FF167-FF192 are:

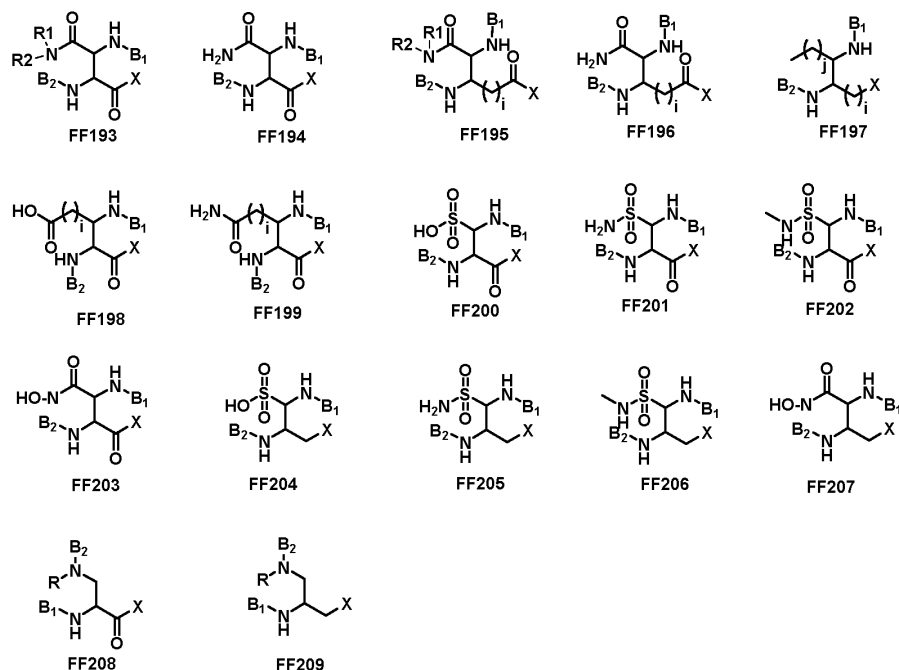




wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH; and

10 B_1 and B_2 , which may be identical or different, each independently represents an aromatic boron-containing group.

Formulae FF193-FF209 are:



5

wherein R in FF208 and FF209 is an alkyl, aryl or halide that is covalently conjugated through at least one CH₂ group to the amino group in the side chain of FF208 or FF209,

R1 and R2 are independently selected from H, CH₃, alkyl, and formulae IV-1 to IV-135;

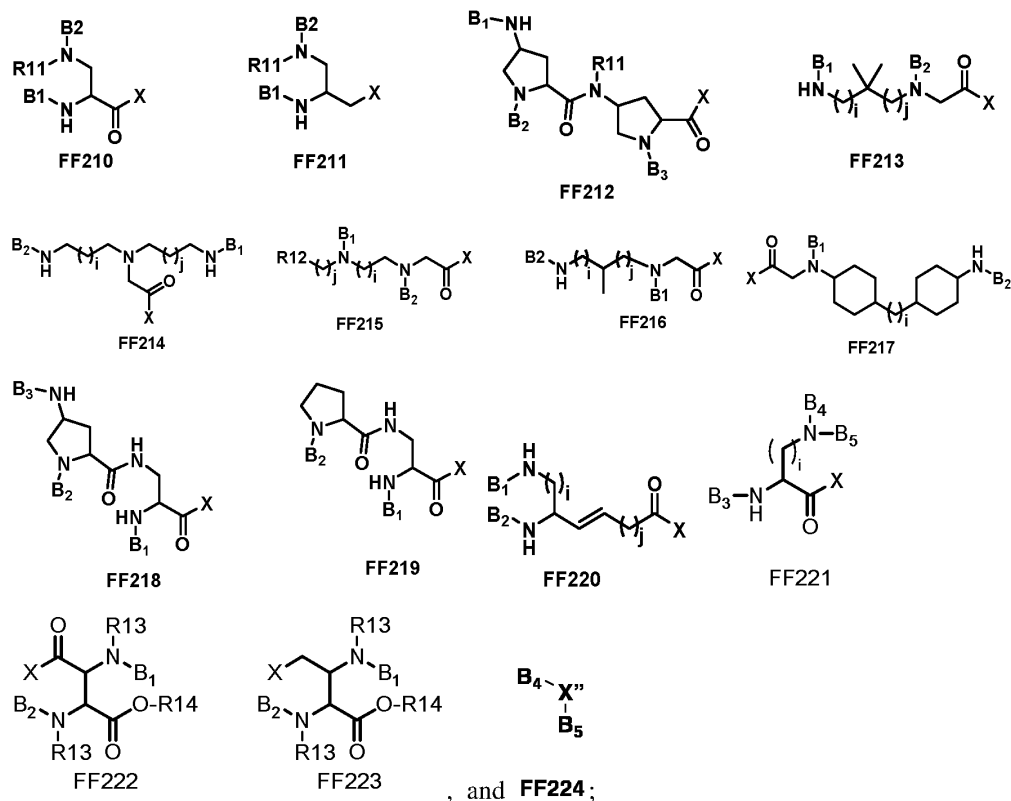
10 i is 1, 2, 3, 4, or 5;

j is 1, 2, 3, 4, or 5; and

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH; and

15 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group.

5 Formulae FF210-FF224 are:



10

wherein R11 in FF210 to FF212 is selected from Formulae IV-1 to IV-135 and R12 is selected from an amine, a hydroxyl, an alkyl, and a halide group;

wherein each R13 is independently selected from H, CH₃, alkyl, aryl and Formulae IV-1 to IV-135; R14 is selected from H, CH₃, alkyl, aryl and heteroaryl;

15 wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

X'' represents a point of covalent attachment to an amine --N in the compound, wherein -- represents a single covalent bond to a CH₂ or CH group in the compound;

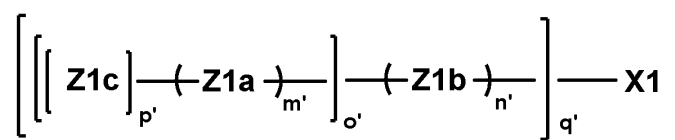
i is 1, 2, 3, 4, or 5;

20 j is 1, 2, 3, 4, or 5; and

5 B₁, B₂, B₃, B₄, B₅, and B₆ each independently represents an aromatic boron-containing group, wherein in each FF structure containing B₁, B₂ and B₃ groups, at least two of the B₁, B₂ and B₃ groups are independently an aromatic boron-containing group; and

wherein at least one primary or secondary amine in FF1-FF223 is optionally covalently conjugated to B₆.

10 In some embodiments, the present disclosure provides a compound of Formula (I) or a molecular conjugate represented by Formula I, or a stereoisomer or a mixture of stereoisomers, or pharmaceutically acceptable salt thereof:



(Formula I)

15 wherein X1 comprises:

- (i) NH₂ or OH (for example, X1 is NH₂ or OH),
- (ii) a polypeptide drug substance comprising an amine,
- (iii) a polypeptide drug substance that is covalently conjugated to an amine containing linker, or
- 20 (iv) an amine configured to be covalently conjugated to a polypeptide drug substance;

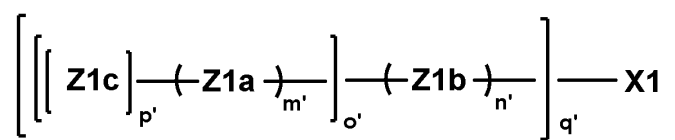
each Z1c is independently selected from Formulae FF1-FF224 and covalently conjugated either directly, or via Z1a and/or Z1b, to X1; each Z1a independently comprises 1 to 50 amino acids connected together using amide or peptide bonds; each Z1b is independently
 25 a small-molecule linker; each m' is independently 0 or 1; each n' is independently 0 or a positive integer; each o' is independently an integer of 1 or greater; each p' is a positive integer; and q' is a positive integer of at least 1 and not more than two times the total number of amine groups in X1, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different;
 30 wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1; and wherein optionally the molecular conjugate may comprise one or more isotopes at any position of the molecular conjugate of Formula I.

In at least some embodiments, X1 comprises one of:

- (i) NH₂ or OH (for example, X1 is NH₂ or OH),

- 5 (ii) a polypeptide drug substance comprising an amine,
 (iii) a polypeptide drug substance that is covalently conjugated to an amine containing linker, or
 (iv) an amine configured to be covalently conjugated to a polypeptide drug substance.

10 In at least some embodiments, the compound is a molecular conjugate represented by Formula I, or a stereoisomer or a mixture of stereoisomers, or pharmaceutically acceptable salts:



(Formula I)

15 wherein:

X1 is NH₂ or OH; or

X1 comprises:

- i. a polypeptide drug substance comprising an amine;
 ii. a polypeptide drug substance that is covalently conjugated to an amine
 20 containing linker; or
 iii. an amine configured to be covalently conjugated to a polypeptide drug substance;

each Z1c is independently selected from Formulae FF1-FF224 and covalently conjugated either directly, or via Z1a and/or Z1b, to X1;

25 each Z1a independently comprises 1 to 50 amino acids connected together using amide or peptide bonds; each Z1b is independently a small-molecule linker;

each m' is independently 0 or 1;

each n' is independently 0 or a positive integer;

each o' is independently an integer greater than or equal to 1;

30 each p' is a positive integer; and

q' is a positive integer of at least 1 and not more than two times the total number of amine groups in X1, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different; wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of

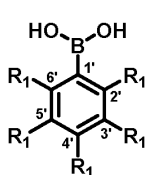
5 Z1a, to an amine of Z1b, or to X1; and wherein optionally the molecular conjugate may comprise one or more isotopes at any position of the molecular conjugate of Formula I. In some embodiments, the compound is additionally covalently conjugated as described by Formula I, and/or wherein one or more amines are each independently acetylated and/or independently alkylated.

10 In some embodiments, the compound of Formula I is covalently conjugated to B₁ using a covalent linkage X-B₁, wherein X is an amino group in Formula I.

In some embodiments, X1 comprises a polypeptide drug substance and the covalent conjugation to X1 is to amino group(s) in one or more lysine residues and/or to the N-terminal amino groups in X1.

15 In some embodiments, the compound comprises at least one of B₁, B₂ and B₃ independently selected from Formulae F1-F12 or wherein the compound comprises at least one of B₄, B₅ and B₆ independently selected from Formulae F1-F10,

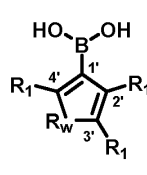
wherein Formulae F1-F10 are:



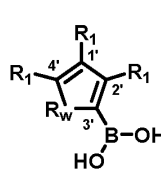
(F1)



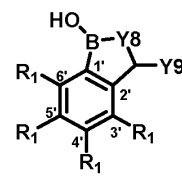
(F2)



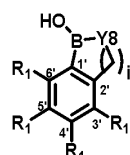
(F3)



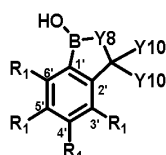
(F4)



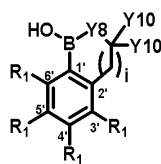
(F5)



(F6)



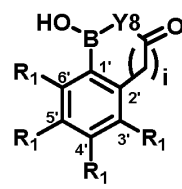
(F7)



(F8)



(F9)



(F10)

20 (F6) (F7) (F8) (F9) , and (F10) :

wherein for B₁, B₂, and B₃:

one R₁ represents (C=O)---*, S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein ---* represents an attachment point to the rest of Z1c, and m is 1, 2, 3, 4, 5, 6, or 7; and

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂,

25 (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃,—
(SO₂)NHCH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

wherein for B₄ and B₅:

one R₁ for B₄ represents (CH₂)_m---ø, wherein ---ø represents an attachment point to the rest of Z1_C and one R₁ for B₅ represents (C=O)---*, S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein ---* represents an attachment point to the rest of Z1_C, and m is 1, 2, 3, 4, 5, 6, or 7; and

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂,
 10 (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —
 (SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

wherein for B₆:

one R₁ for B₆ represents (CH₂)_m---ø, wherein ---ø represents an attachment point to the rest of the compound, and m is 1, 2, 3, 4, 5, 6, or 7; and

15 each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

wherein,

for Formulae F3-F4:

20 R_w is O or S;

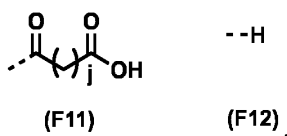
for Formulae F5-F10:

Y8 is selected from O, N, and NR, wherein R is an alkyl group or H;

Y9 is H, CH₃, or an alkyl group, provided that when Y8 is O, Y9 is a CH₃ or an alkyl group;

25 each Y10 is independently selected from H, CH₃, F, CF₃, and OCH₃; and
i is 1, 2, or 3; and

wherein Formulae F11-F12 are:



j is 1, 2, 3, 4, 5, 6, or 7; and

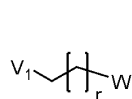
--- represents an attachment point to the rest of Z1c.

In at least some embodiments, B₁, B₂ and B₃ may be identical or different. If B₁, B₂ and B₃ are all present in a compound of the present disclosure, then each is independently an

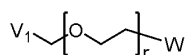
- 5 aromatic boron-containing group, a carboxylic acid derivative, or a H, wherein in each FF structure (i.e., FF1 to FF224) containing B₁, B₂ and B₃ groups, at least two of the B₁, B₂ and B₃ are independently an aromatic boron-containing group.

In some embodiments, the compound comprises at least one Z1b selected from

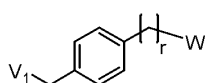
- 10 Formulae IIa-IIai or Formulae IIIa-IIIai,
wherein Formulae IIa-IIai are:



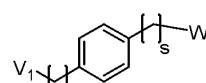
Formula IIa



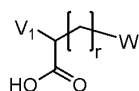
Formula IIb



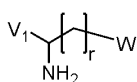
Formula IIc



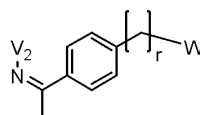
Formula IId



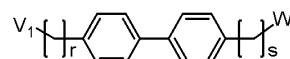
Formula IIe



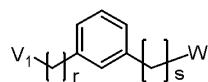
Formula IIff



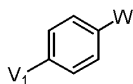
Formula IIg



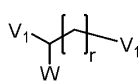
Formula IIh



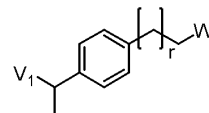
Formula IIi



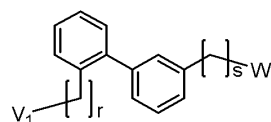
Formula IIj



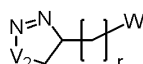
Formula IIk



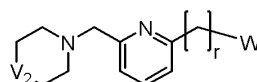
Formula IIl



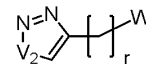
Formula IIm



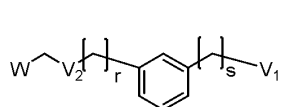
Formula IIn



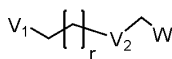
Formula IIo



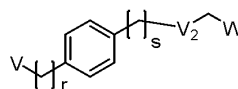
Formula IIp



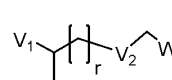
Formula IIq



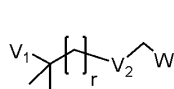
Formula IIr



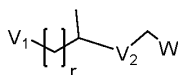
Formula IIs



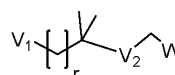
Formula IIu



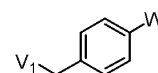
Formula IIv



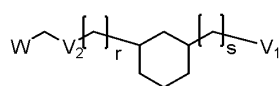
Formula IIw



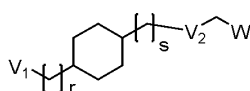
Formula IIx



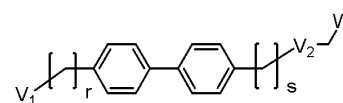
Formula IIy



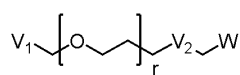
Formula IIz



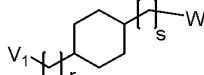
Formula IIaa



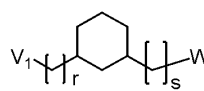
Formula IIab



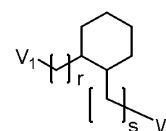
Formula IIac



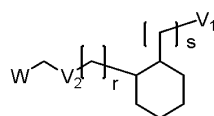
Formula IIad



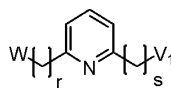
Formula IIae



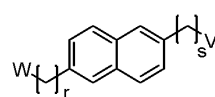
Formula IIaf



Formula IIag

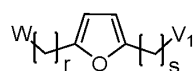


Formula IIah



Formula IIai

and



Formula IIaj ;

wherein:

10 r is 0, 1, 2, 3, 4, or 5;

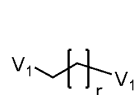
s is 0, 1, 2, 3, 4, or 5;

W represents CH₂--- or (C=O)---, wherein --- is a covalent linkage to X1; and

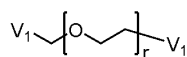
each V₁ is independently selected from NH---, CH₂---, and (C=O)--- and each V₂ is N---, wherein --- is a covalent linkage towards successive Z1b, Z1a or Z1c, provided that V₁ is NH--- when connected to Z1c; and the covalent linkages between Z1a and Z1b units each independently comprise an amine linkage or an amide linkage; and when n'=0 and m'=1, Z1a is directly conjugated to X1 by an amine linkage or amide linkage, and

5

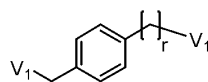
wherein Formulae IIIa-IIIai are:



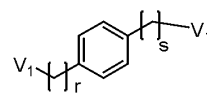
Formula IIIa



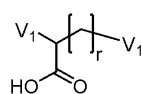
Formula IIIb



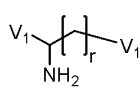
Formula IIIc



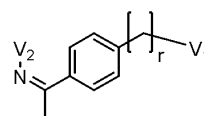
Formula IIId



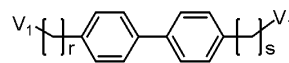
Formula IIIe



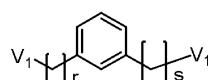
Formula IIIf



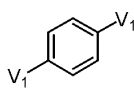
Formula IIIg



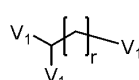
Formula IIIh



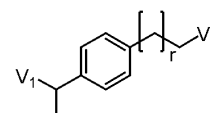
Formula IIIi



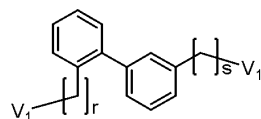
Formula IIIj



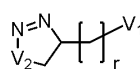
Formula IIIk



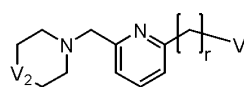
Formula IIIl



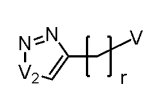
Formula IIIm



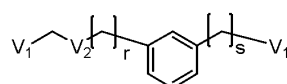
Formula IIIn



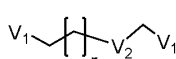
Formula IIIo



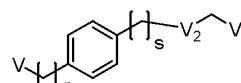
Formula IIIp



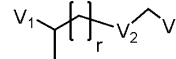
Formula IIIq



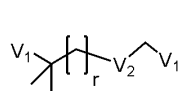
Formula IIIr



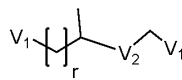
Formula IIIs



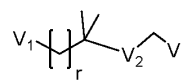
Formula IIIt



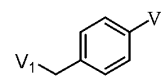
Formula IIIu



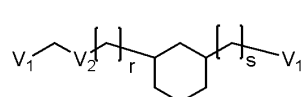
Formula IIIv



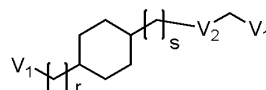
Formula IIIw



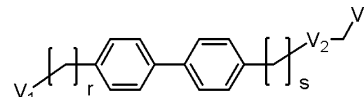
Formula IIIx



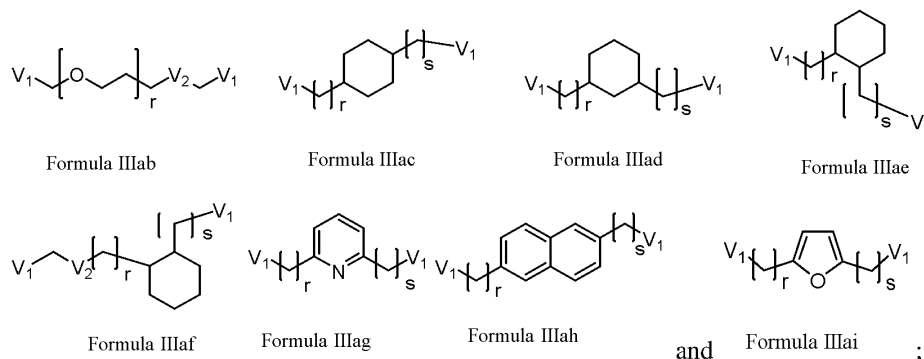
Formula IIIy



Formula IIIz



Formula IIIaa



wherein:

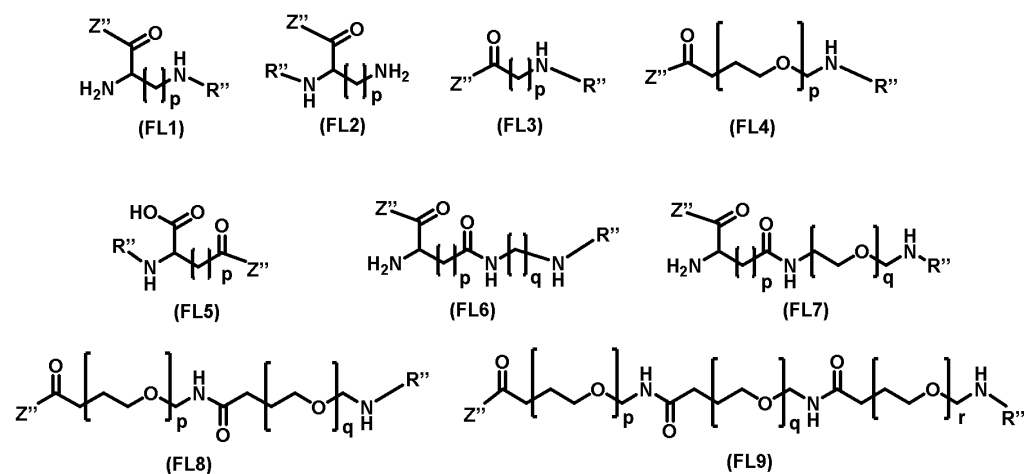
r is 1, 2, 3, 4, or 5;

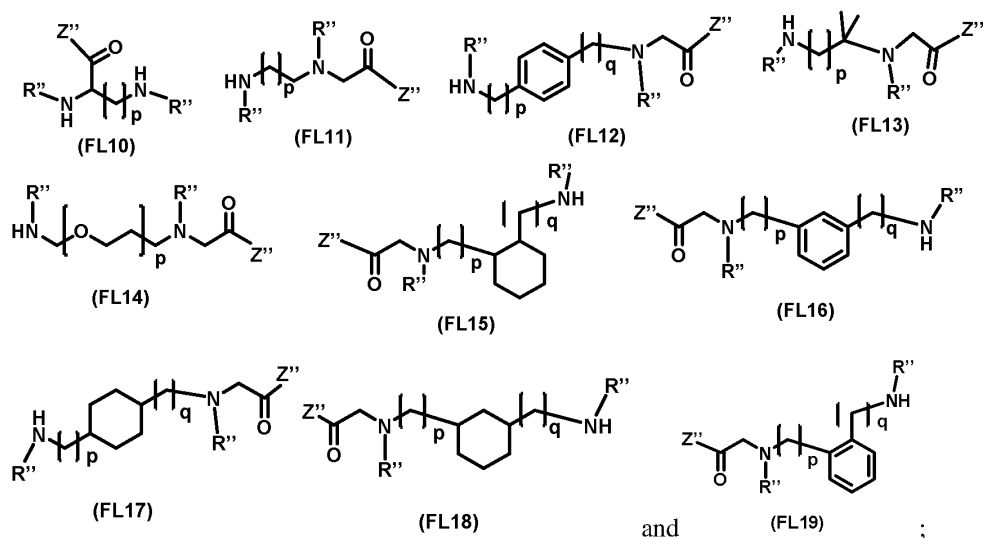
s is 1, 2, 3, 4, or 5; and

- 10 each V_1 is independently selected from $\text{NH---}\dagger$, $\text{CH}_2\text{---}\dagger$, and $(\text{C}=\text{O})\text{---}\dagger$ and each V_2 is $\text{N---}\dagger$, wherein $\text{---}\dagger$ is a covalent linkage towards successive Z1b, Z1a or Z1c, provided that V_1 is $\text{NH---}\dagger$ when connected to Z1c; and the covalent linkages between Z1a and Z1b units each independently comprise an amine linkage or an amide linkage; and when $n'=0$ and $m'=1$, Z1a is directly conjugated to X1 by an amine linkage or amide linkage.

15

In some embodiments, at least one Z1c is covalently conjugated indirectly via a linker (indirect linker) to the compound (e.g., the compound of Formula I). In some embodiments, the linker is selected from (i) Formulae FL1-FL19:





wherein, in Formulae FL1 to FL19:

Z'' represents an attachment point toward X1;

R'' represents an attachment point toward Z1c;

10 p is 1, 2, 3, 4, or 5,

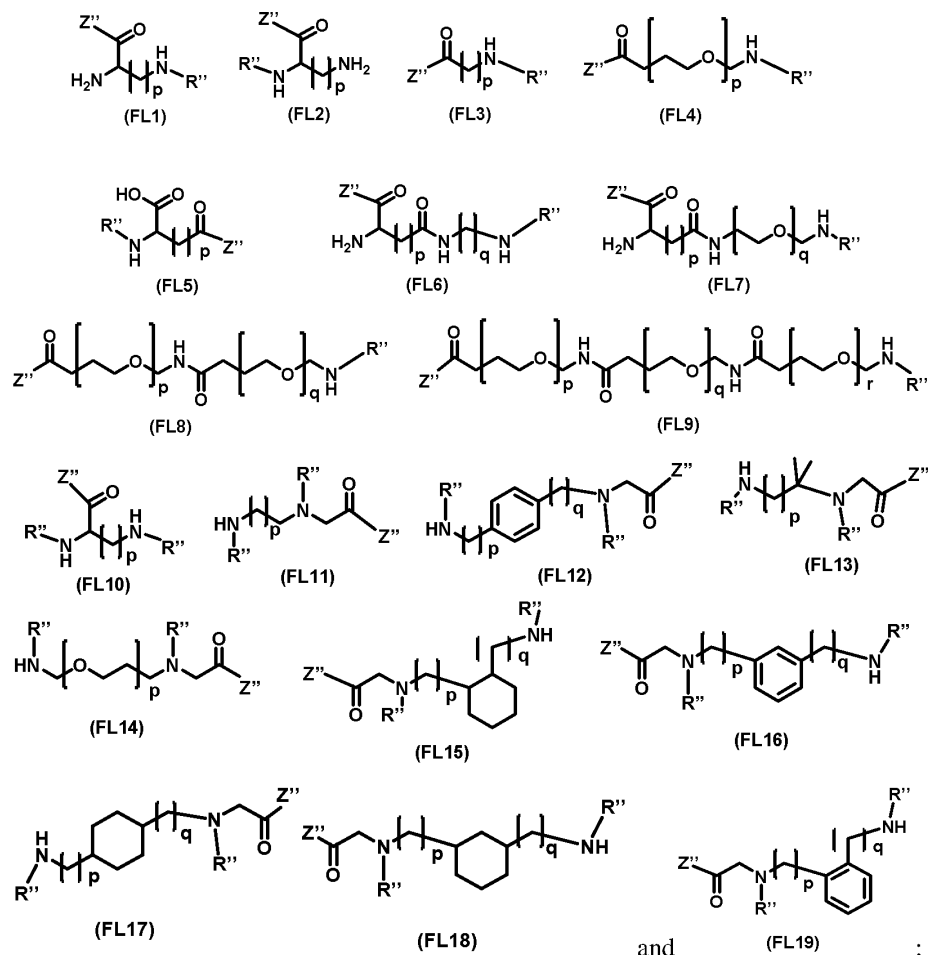
q is 1, 2, 3, 4, or 5,

r is 1, 2, 3, 4, or 5; and

any primary amine is optionally acetylated or alkylated; and

(ii) an L- or D-amino acid comprising at least one amine group directly conjugated to
 15 Z1c, wherein an acid functional group of the amino acid is conjugated toward X1 in Formula I.

In some embodiments, n' is 1 and each of the Z1b is independently selected from (i)
 Formulae FL1-FL19:



wherein, in Formulae FL1 to FL19:

Z'' represents an attachment point toward X1;

10 R'' represents an attachment point toward Z1c;

p is 1, 2, 3, 4, or 5,

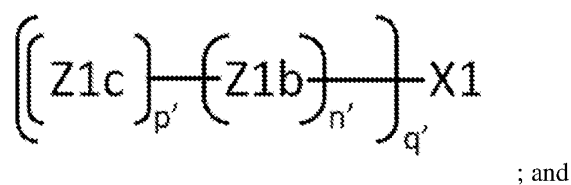
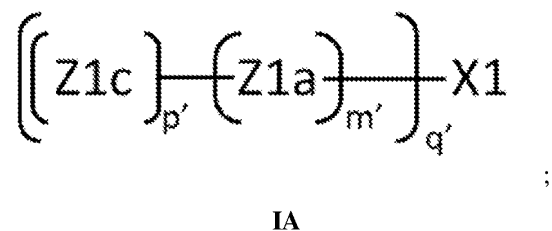
q is 1, 2, 3, 4, or 5,

r is 1, 2, 3, 4, or 5; and

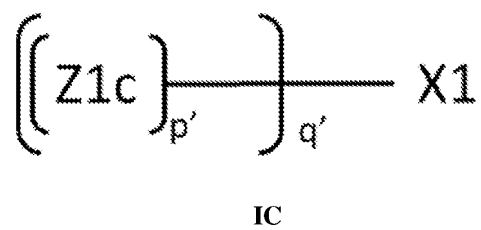
any primary amine is optionally acetylated or alkylated; and

15 (ii) an L- or D-amino acid comprising at least one amine group directly conjugated to Z1c, wherein an acid functional group of the amino acid is conjugated toward X1 in Formula I.

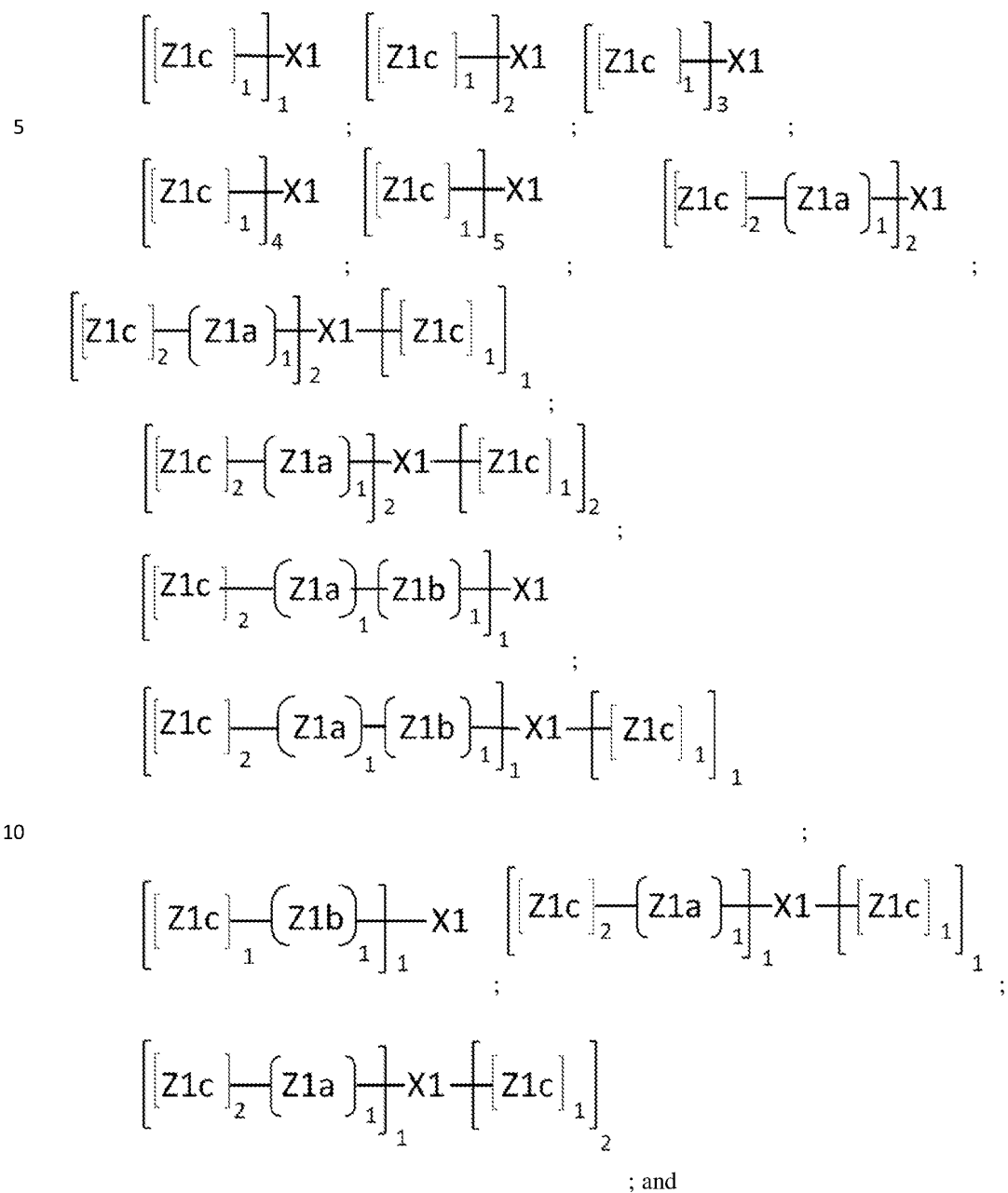
5 In some embodiments, the compound of Formula (I) is selected from:

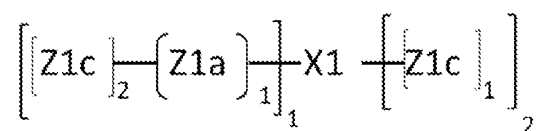


10



In some embodiments, the compound of Formula I is selected from:

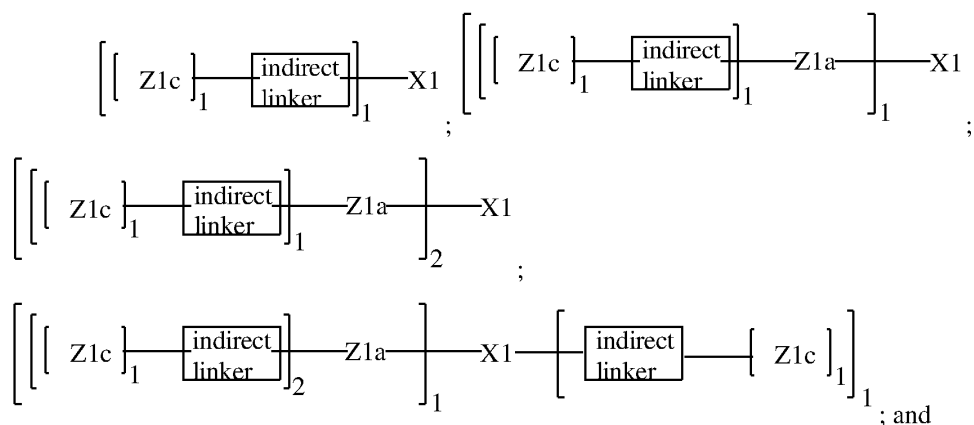




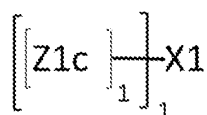
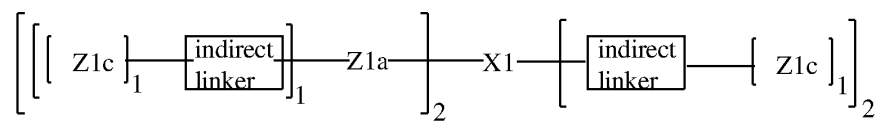
5

. In some embodiments, the

compound of Formula I is selected from:



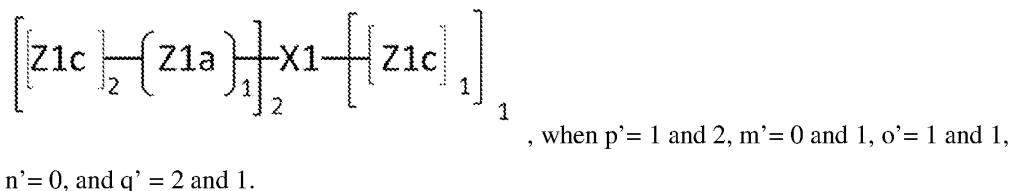
10



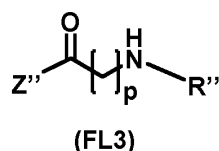
In some embodiments, the compound is , when $p'=1$, $m'=0$, $o'=1$, $n'=0$, and $q'=1$.

15

In some embodiments, the compound is

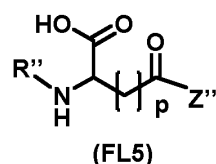


- 5 In some embodiments, in Formula I, Z1c is directly connected to X1 through an optional covalent-spacer, and the optional covalent-spacer is independently selected from gamma-glutamic acid, beta-alanine, and



Formula FL3, wherein p is 1 or 2; and

10



Formula FL5, wherein p is 2, 3, or 4;

In some embodiments, X1 is OH or NH₂, and X1 further comprises a drug substance covalently conjugated directly or indirectly to the compound.

- 15 In at least some embodiments, the compound of the present disclosure comprises a drug substance comprising a polypeptide hormone, a human polypeptide hormone and/or insulin, or an analogue thereof, or a hybrid polypeptide comprising one or more combinations thereof.

- 20 In at least some embodiments, the compound of the present disclosure comprises an amine in the compound that is conjugated via an amide linkage to an aromatic boron-containing compound (e.g., group). In some embodiments, the aromatic boron-containing group is selected from a phenylboronic acid, boroxole, and phenylboronate.

In at least some embodiments, the compound of the present disclosure is dehydrated (loses) by 1, 2, 3, 4, 5, 6, 7, 8, or more water molecules.

- 25 In at least some embodiments, the compound of the present disclosure is formulated in a solution comprising one or more of a buffer, stabilizer, vasodilator, preservative, surfactant, salt, sugar, or compounds containing one or more hydroxyls, alcohols, diols, or phenols. For example, the solution could comprise one or more of citrate, zinc, and/or cresol.

- 30 In at least some embodiments, X1 comprises a human polypeptide hormone of the human pancreas, insulin, glucagon, GLP-1, a somatostatin, a gastric inhibitory polypeptide, a

- 5 glucose-dependent insulintropic polypeptide, a hybrid peptide comprising sequences from two or more human polypeptide hormones, or an analogue thereof.

In some embodiments, X1 comprises human insulin or a human insulin analogue comprising an A-chain and a B-chain, wherein the A-chain comprises a sequence selected from SEQ ID NOs 1 and 3 to 33, and the B-chain comprises a sequence
10 selected from SEQ ID NOs 2 and 34 to 74, 24047, and 24048;

each Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, FF163, FF193, FF194, FF203, and FF221-FF224 and covalently conjugated either directly, or indirectly via the linker, to Z1a and/or Z1b, or to X1;

each Z1a is independently absent or independently comprises a sequence selected from
15 K, GK, KGSH (SEQ ID NO:24049), KGSHK (SEQ ID NO:4238), KNSTK (SEQ ID NO:5085), GKASHK (SEQ ID NO:12414), GKEEHK (SEQ ID NO:12677), GKEEHK (SEQ ID NO:12680), GKGHSHK (SEQ ID NO:13120), GKGSK (SEQ ID NO:24050), GKGSK (SEQ ID NO:13198), KGSTK (SEQ ID NO:13205), KHENK (SEQ ID NO:13271), GKNSHK (SEQ ID NO:13982), GKNSTK (SEQ ID NO:13989), GKQSSK (SEQ ID
20 NO:14380), GKYQFK (SEQ ID NO:15128), KGSKK (SEQ ID NO:24045), GKPGKK (SEQ ID NO:24046), KGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGSHK (SEQ ID NO:24042);

each linker is selected from FL1, FL3, FL4, and FL5;

each m' is independently 0 or 1;

25 each n' is independently 0, 1, 2, or 3;

each o' is independently 1, 2, 3, 4, or 5;

each p' is 1, 2, 3, 4, or 5; and

q' is 1, 2, 3, or 4, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different; and

30 wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1.

In some embodiments, X1 comprises the human insulin or human insulin analogue comprising an A-chain and a B-chain, wherein the A-chain comprises SEQ ID NO:1; and the B-chain is selected from SEQ ID NOs 2, 36, 24047, and 24048;

5 each Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, FF193, FF194, FF203, and FF221-FF224 and covalently conjugated either directly, or indirectly via the linker, to Z1a and/or Z1b, or to X1;

each Z1a independently comprises a sequence selected from K, GK, KGSH (SEQ ID NO:24049), KGSHK (SEQ ID NO:4238), KNSTK (SEQ ID NO:5085), GKASHK (SEQ ID NO:12414), GKEEEK (SEQ ID NO:12677), GKEEHK (SEQ ID NO:12680), GKGHSK (SEQ ID NO:13120), GKGSH (SEQ ID NO:24050), GKGSHK (SEQ ID NO:13198), GKGSTK (SEQ ID NO:13205), GKHENK (SEQ ID NO:13271), GKNSHK (SEQ ID NO:13982), GKNSTK (SEQ ID NO:13989), GKQSSK (SEQ ID NO:14380), GKYQFK (SEQ ID NO:15128), GKGSKK (SEQ ID NO:24045), GKKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGSBK (SEQ ID NO:24042);

each linker is independently absent or independently selected from FL3 and FL5;
each m' is independently 0 or 1;
each n' is independently 0 or 2;
20 each o' is independently 1, 2, or 3;
each p' is 1, 2, or 3; and
q' is 1, 2, or 3, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different;
wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1.

25 In some embodiments, each of the Z1a is independently absent or independently comprises a sequence selected from K, GK, KGSH (SEQ ID NO:24049), GKGSH (SEQ ID NO:24050), KGSHK (SEQ ID NO:4238), and GKGSHK (SEQ ID NO:13198).

In some embodiments, each of the Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, and FF221-FF224. In some embodiments, B₁ and B₂ are independently selected from Formulae F1 and F2. In some embodiments, the B₁ and the B₂ are F2. In some embodiments, at least one R1 in B₁ or B₂ is F or CF₃. In some embodiments, Z1b is independently absent, FL3, or FL5. In some embodiments, each of the Z1c is independently selected from FF10, FF12, FF116, FF221, FF222, and FF224. In some embodiments, each B₁ and B₂ is F2 and is covalently conjugated to Z1c using an amide linkage,

5 each Z1b is independently absent; FL3 wherein p is 1, 2, or 3; or FL5 wherein p is 2, 3, or 4;

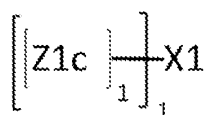
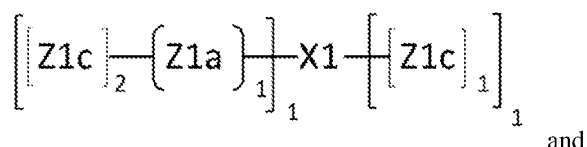
each FF is independently selected from FF10, FF12, FF116, FF134, FF163, FF193, FF203, FF221, FF222 and FF224; wherein FF12 and FF222 has either (*S,R*) or (*S,S*) stereochemistry;

10 each Z1c is conjugated either directly or indirectly through FL3 or FL5 to the amine group in one or more lysine side chain in X1 or the N-terminus in X1; and

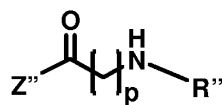
X1 is a polypeptide drug substance and/or an insulin optionally having from 0 to 4 residues replaced, inserted, or mutated to lysines, and wherein the lysines are each conjugated directly or indirectly to a Z1c.

15 In some embodiments, Z1c is FF224, n' is 0, and Z1a is an amine containing amino acid.

In some embodiments, the compound is selected from

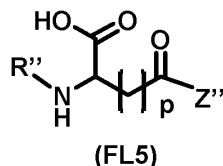


20 In some embodiments, Z1c is covalently conjugated directly to X1 via a linker, and wherein the linker is independently selected from gamma-glutamic acid, beta-alanine, and



Formula FL3,

wherein p is 1, 2, or 3; and



5

Formula FL5,
wherein p is 2, 3, or 4.

In some embodiments, the compound further comprises a drug substance covalently conjugated directly or indirectly to the compound.

10 In some embodiments, the compound of Formula I is selected from Examples 315, 318, 320, 605-608, 610-612, 589-595, 562-574, and 803-876.

In some embodiments, X1 is a polypeptide drug substance and/or an insulin optionally having from 0 to 4 residues replaced, inserted, or mutated to lysines, and wherein the lysines are each conjugated to a Z1c.

15 In some embodiments, one or more amines are each independently acetylated and/or independently alkylated.

In some embodiments, X1 comprises a polypeptide drug substance and the covalent conjugation to X1 is to amino group(s) in one or more lysine residues and/or to the N-terminal amino groups in X1.

20 In some embodiments, each R1 in FF1-FF224 is independently selected from a C₁-C₂₂ alkyl group, a C₁-C₂₂ acyl group, a (C₃-C₈)cycloalkyl group, a C₁-C₂₂ haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises one or more C₁-C₂₂ alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, C₁-C₂₂ alkyl, or aryl groups.

In some embodiments, X4 is selected from -COOH, -(CH₂)_mCOOH, a C₁-C₂₂ alkyl group, a C₁-C₂₂ acyl group, a (C₃-C₈)cycloalkyl group, a C₁-C₂₂ haloalkyl group, an aryl group, and a heteroaryl group, each X4 optionally comprises one or more C₁-C₂₂ alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, C₁-C₂₂ alkyl, or aryl groups; wherein m is 1, 2, 3, 4, or 5.

In some embodiments, the alkyl group of Y9 is a C₁-C₂₂ alkyl. In some embodiments, Y9 is CH₃.

30

In some embodiments, at least one primary or secondary amine in FF1-FF223 is covalently conjugated to B6.

5 In some embodiments, an amine in the compound is conjugated via an amide linkage to an aromatic boron-containing group.

In some embodiments, the aromatic boron-containing group is selected from a phenylboronic acid, boroxole, and phenylboronate.

10 In some embodiments, the compound is formulated in a solution comprising one or more of a buffer, stabilizer, vasodilator, preservative, surfactant, salt, sugar, or compounds containing one or more hydroxyls, alcohols, diols, or phenols. In some embodiments, the solution comprises one or more of citrate, zinc, and/or cresol.

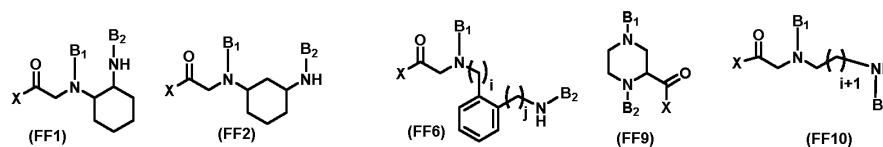
In some embodiments, Z1c is conjugated to a cysteine.

15 In some embodiments, the compound (e.g., the compound of Formula I) is covalently conjugated either directly or through a linker to a diol, sugar, carbohydrate or a diol containing molecule.

In some embodiments, the compound (e.g., the compound of Formula I) is covalently conjugated to an antibody, albumin or a fragment thereof, or covalently conjugated either directly or through a linker to a molecule that can bind to at least one protein present in human plasma. In at least one embodiment, the present disclosure provides a method to administer the compounds disclosed herein to a human subject as a therapeutic or prophylactic agent.

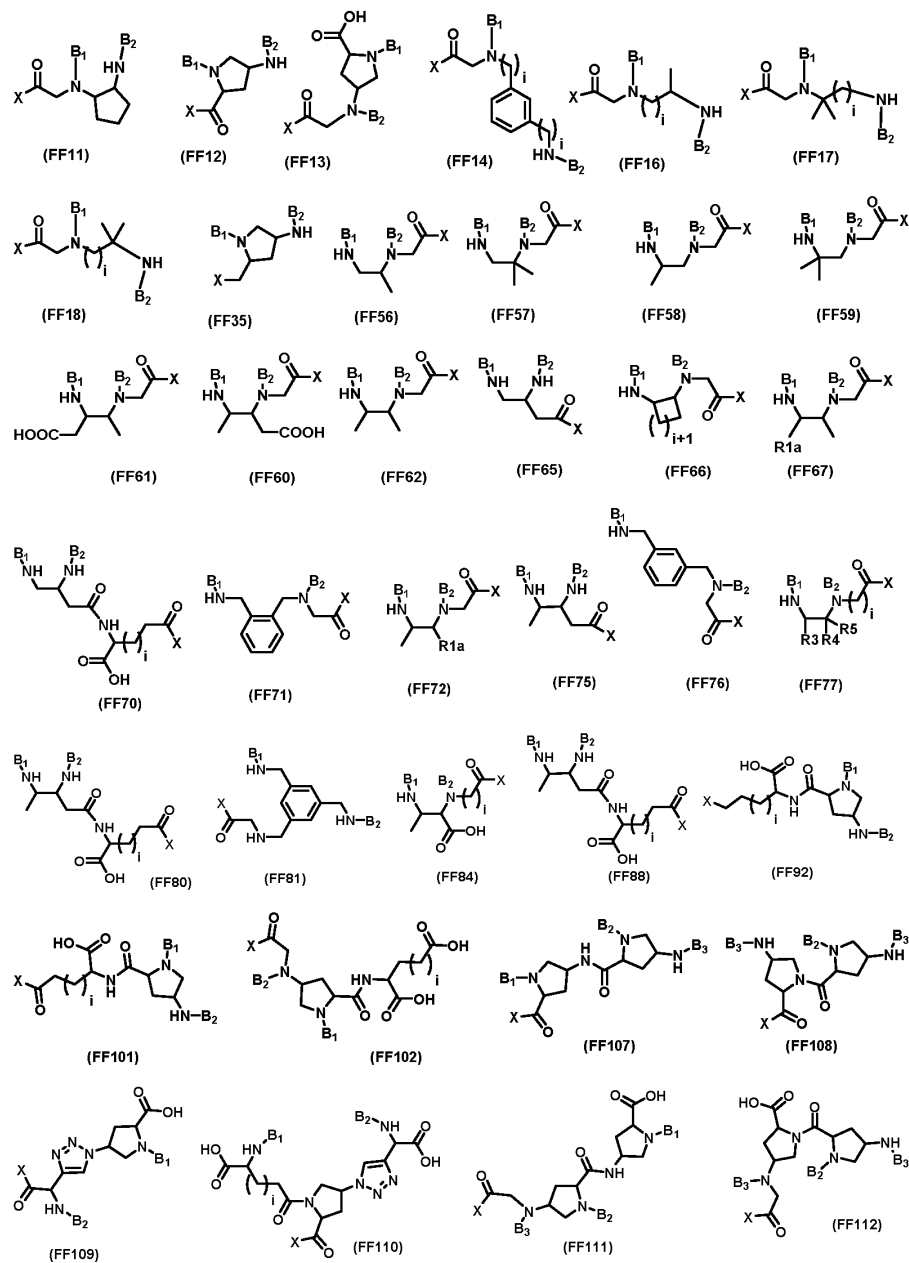
In some embodiments, the compounds disclosed herein are used as intermediate compounds for the manufacture of any compounds disclosed herein.

25 In some embodiments, the compounds disclosed herein comprise at least one Z1c. In at least some embodiments, the Z1c is a boron containing compound. In some embodiments, a subset of boron containing compounds is selected from a non-aromatic and/or an aromatic boron-containing group. In some embodiments, Z1c is an aromatic boron-containing group. In at least one embodiment, the compound of the present disclosure comprises at least one Z1c selected from:



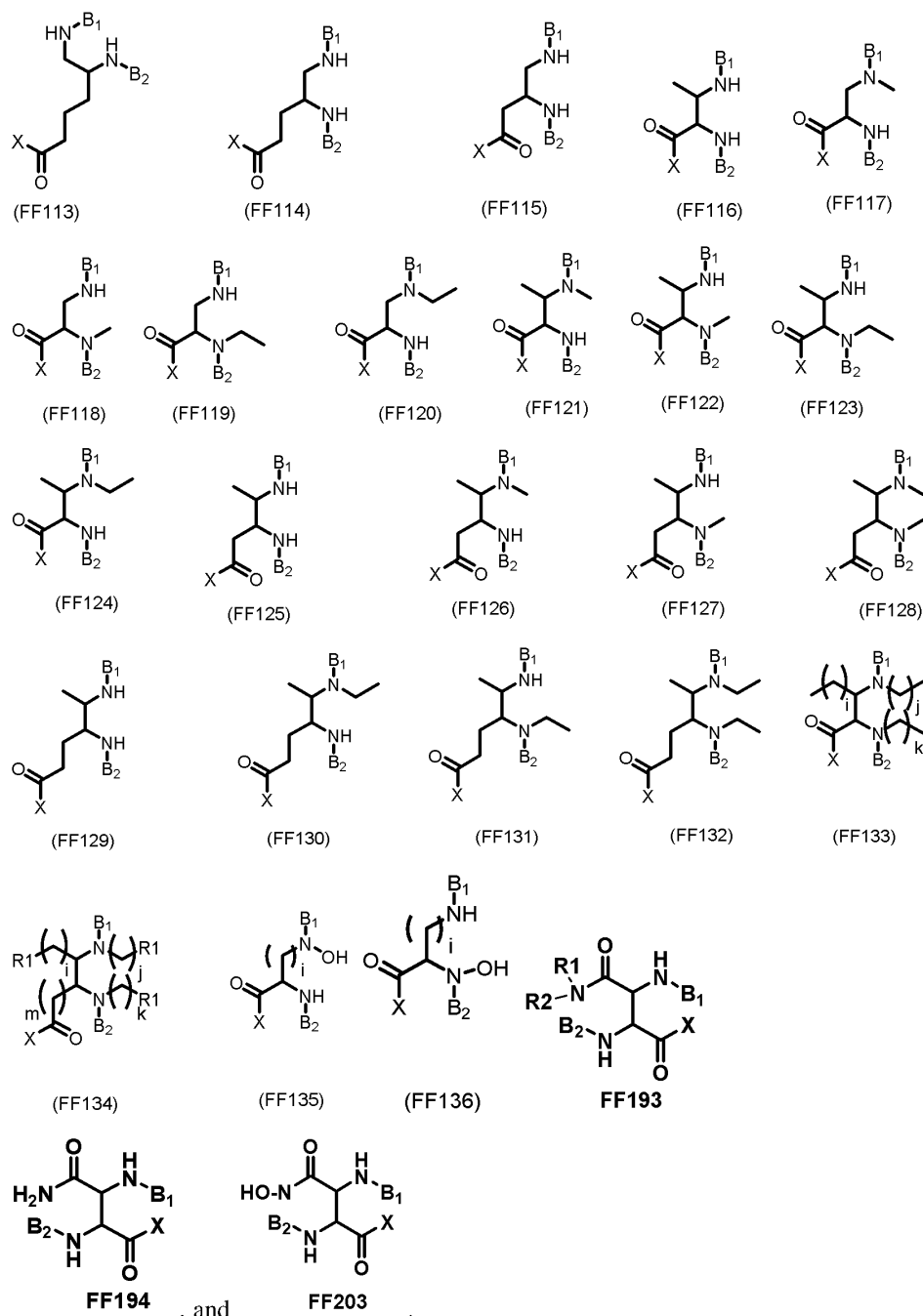
30

5



10

5



10

5 In at least some embodiments, the Z1c is selected from FF1-FF224. In some embodiments, the compound comprises at least one Z1c having at least one chiral center and selected from FF1, FF2, FF5, FF9, FF11-FF13, FF15-FF24, FF27, FF31, FF34-FF36, FF38, FF39, FF43-FF58, FF60-FF70, FF72-FF75, FF77-FF80, FF82-FF84, FF86-FF212, FF216-FF220, FF222, FF223, and combinations thereof.

10 In some embodiments, the compound comprises at least one FF12 and/or FF116. In some embodiments, the stereochemistry of FF12 and FF116 is independently selected from (*S,S*); (*S,R*); (*R,R*); and (*R,S*).

 In some embodiments, X1 comprises human insulin or a human insulin analogue comprising an A-chain and a B-chain, wherein the C-terminus of the A-chain of the human
15 insulin analogue is optionally extended with a polypeptide of up to 20 residues, and/or the N-terminus of the B-chain of the human insulin analogue is optionally extended with a polypeptide of up to 10 residues. In some embodiments, one to six residues of the insulin A-chain and/or the insulin B-chain are deleted or mutated.

 In some embodiments, X1 comprises at least one lysine having an amine side chain, and
20 Z1c is covalently conjugated directly to the amine side chain. In some embodiments, the compound of the present disclosure comprises at least one Z1a comprising one or more amino acids having an amine side chain, and wherein the one or more amino acids are selected from lysine, diaminopropionic acid, diaminobutyric acid, and ornithine; and wherein Z1c is covalently conjugated, directly or indirectly, to the amine side chain.

25 In some embodiments, the compound of the present disclosure may include one or more isotopes selected from deuterium, tritium, carbon-13, carbon-14, and iodine-124. In at least one embodiment, the compound comprises deuterium.

 In some embodiments, X1 comprises a drug substance covalently conjugated to at least one Z1c through an acid containing linker. In some embodiments, a composition of the present
30 disclosure comprises at least one compound as disclosed herein (e.g., a compound comprising X1 and one or more Z1c, Formula I), or a tautomer, stereoisomer or a mixture of stereoisomers, or a pharmaceutically acceptable salt, or hydrate, or isotopic derivative thereof formulated together with one or more pharmaceutically acceptable carriers.

 In some embodiments, the present disclosure also provides a composition or a mixture
35 comprising at least one compound as disclosed herein, for use as a medicament for the

5 treatment of diabetes, for control of blood sugar levels, or to control the release of a drug based on physiological levels of diol containing small molecules or sugars.

In some disclosed embodiments are a method of administering a compound as disclosed herein to a human subject as a therapeutic or prophylactic agent.

In some embodiments, the disclosure provides a method of making a compound as
10 disclosed herein comprising at least one alkylation and/or amidation step.

In some embodiments, the disclosure provides a method of treating a subject by administering a device or formulation comprising a compound as disclosed herein, such as Examples 1-880. For example, the device can be a fixed dose injector, microdosing injector, an internal or external patch.

15 In some embodiments, the disclosure provides a method of treating or preventing diabetes, impaired glucose tolerance, hyperglycemia, or metabolic syndrome (metabolic syndrome X, insulin resistance syndrome) comprising administering to a subject in need thereof a therapeutically effective amount of a compound disclosed herein.

In at least some embodiments, the present disclosure is directed to a compound of
20 Formulae FF1-FF224, or a tautomer, stereoisomer or a mixture of stereoisomers, or a pharmaceutically acceptable salt, or hydrate, or isotopic derivative thereof. In at least some embodiments, the present disclosure is directed to a compound selected from Formulae FF1-FF48, Formulae FF49-FF88, FF89-FF112, FF113-FF136, FF137-FF160, FF161-FF164, FF165-FF166, FF167-FF192, FF193-FF209, and FF210-FF224.

25 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF1-FF48,

wherein X is selected from an amine, OH, and halogen; and

i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7; and

30 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group.

In some embodiments, the present disclosure is directed to a compound selected from Formulae FF49-FF88,

wherein X is selected from an amine, OH, and halogen;

35 i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7;

- 5 R1a is selected from COOH, CH₃, H, and OH;
R2, R3, R4 and R5 is each independently selected from CH₃, H, OH, and COOH, and at least one of R2, R3, R4 and R5 is CH₃ or OH; and
B₁ and B₂, which may be identical or different, are each independently an aromatic boron-containing group.
- 10 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF89-FF112,
wherein X is selected from an amine, OH, and halogen;
i is 1, 2, 3, 4, 5, 6, or 7; and
B₁, B₂ and B₃, which may be identical or different, each independently represents an
15 aromatic boron-containing group, a carboxylic acid derivative, or a H, wherein at least two of B₁, B₂ and B₃ in each FF structure are independently an aromatic boron-containing group.
In some embodiments, the present disclosure is directed to a compound selected from Formulae FF113-FF136,
wherein X is selected from an amine, OH, and halogen;
20 i is 0, 1, 2, 3, 4, 5, 6, or 7;
j is 0, 1, 2, 3, 4, 5, 6, or 7;
k is 0, 1, 2, 3, 4, 5, 6, or 7;
m is 0, 1, 2, 3, 4, 5, 6, or 7;
wherein i+j+k+m is greater than 0
25 each R1 is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; and
B₁ and B₂, which may be identical or different, each independently represents an
30 aromatic boron-containing group.
In some embodiments, the present disclosure is directed to a compound selected from Formulae FF137-FF160,
wherein X is selected from an amine, OH, and halogen;
i is 0, 1, 2, 3, 4, 5, 6, or 7;
35 j is 0, 1, 2, 3, 4, 5, 6, or 7;
k is 0, 1, 2, 3, 4, 5, 6, or 7;

- 5 m is 0, 1, 2, 3, 4, 5, 6, or 7;
 wherein $i+j+k+m$ is greater than 0;
 each R₁ is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R₁ optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl
10 groups; and
 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group.
 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF161-FF164,
15 wherein X is selected from an amine, OH, and halogen;
 i is 1, 2, 3, 4, or 5 (e.g., 1, 2, 3, or 5);
 j is 1, 2, 3, 4, or 5 (e.g., 1, 2, 3, or 5);
 each R₆, R₇, R₈, and R₉ for different values of j is independently selected from H, CF₃, CH₃, CHF₂, and (CH₂)_mCH₃, wherein m is 1, 2, 3, 4, or 5;
20 Y₃, Y₄, Y₅, Y₆ and Y₇ are each independently selected from H, CH₂—X₄, and Formulae IV-1 to IV-135 (as previously defined);
 wherein X₄ is selected from -COOH, -(CH₂)_mCOOH, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each optionally comprising one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl
25 or aryl groups; wherein m is 1, 2, 3, 4, or 5;
 wherein at least one of Y₅, Y₆, and Y₇ in Formulae FF162 and FF163 is not H; and
 at least one of Y₇, R₈ and R₉ in FF164 is not H; and
 X_a represents CH=O, CHF₂, CF₃, CH₂SH, COOH, CH₂OH, CH₂NO₂, CH₂NH₂, CH₃, C(CH₃)₃, CH(CH₃)₂, CH((CH₂)₃ CH₃)₂, or CH(CH₂ CH₃)₂;
30 X_b represents O, NH, CH₂, or S;
 X_c represents CH or N;
 each R₁₀ is independently selected from H, F, Cl, Br, CH₃, CF₃, CH=O, OH, COOH, and (CH₂)_nCH₃,
 m is 1, 2, 3, 4, or 5; and n is 1, 2, 3, 4, or 5; and
35 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group. In some embodiments, when j is 4, X is not NH₂ for FF163.

5 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF165-FF166,

wherein X is selected from an amine, OH, and halogen;

m is 1, 2, 3, 4, 5, 6, or 7;

n is 1, 2, 3, 4, 5, 6, or 7;

10 X5 is S, O, or NH; and

each R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7.

15 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF167-FF192,

wherein X is selected from an amine, OH, and halogen; and

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group.

20 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF193-FF209,

wherein R in FF208 and FF209 is an alkyl, aryl or halide that is covalently conjugated through at least one CH₂ group to the amino group in the side chain of FF208 or FF209;

R₁ and R₂ are independently selected from H, CH₃, alkyl, and formulae IV-1 to IV-135;

25 i is 1, 2, 3, 4, or 5;

j is 1, 2, 3, 4, or 5; and

wherein X is selected from an amine, OH, and halogen; and

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group.

30 In some embodiments, the present disclosure is directed to a compound selected from Formulae FF210-FF224,

wherein R₁₁ in FF210 to FF212 is independently selected from Formulae IV-1 to IV-135 and R₁₂ is selected from an amine, a hydroxyl, an alkyl, and a halide group;

35 wherein each R₁₃ is independently selected from H, CH₃, alkyl, aryl, and formulae IV-1 to IV-135; R₁₄ is selected from H, CH₃, alkyl, aryl, and heteroaryl; wherein X is independently selected from an amine, OH, and halogen;

5 X'' is an amine;

i is 1, 2, 3, 4, or 5;

j is 1, 2, 3, 4, or 5; and

B₁, B₂, B₃, B₄, B₅, and B₆, each independently represents an aromatic boron-containing group, wherein in any compound containing B₁, B₂ and B₃ groups, at least two groups are
10 independently an aromatic boron-containing group.

In at least some embodiments, when X is an amine in any one of Formulae FF1 to FF223, X is optionally acetylated or alkylated.

In some embodiments, the compound comprises at least one of B₁, B₂ and B₃ independently selected from Formulae F1-F12 or wherein the compound comprises at least one
15 of B₄, B₅ and B₆ independently selected from Formulae F1-F10. In at least some embodiments, B₁, B₂ and B₃ may be identical or different. If B₁, B₂ and B₃ are all present in a compound of the present disclosure, then each is independently an aromatic boron-containing group, a carboxylic acid derivative, or a H, with the proviso that in each FF structure (i.e., FF1 to FF224) containing B₁, B₂ and B₃ groups, at least two groups are independently an aromatic
20 boron-containing group.

In some embodiments, for B₁, B₂, B₃:

one R₁ represents (C=O)---*, S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein
---* represents an attachment point to the rest of Z1c, and m is 1, 2, 3, 4, 5, 6, or 7;

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂,
25 (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

In some embodiments, for B₄, B₅:

one R₁ for B₄ represents (CH₂)_m---ø, wherein ---ø represents the attachment point (representing a covalent bond) to an amine in X1 and one R₁ for B₅ represents (C=O)---*,
30 S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein ---* represents the attachment point to the same amine in X1, and m is 1, 2, 3, 4, 5, 6, or 7;

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7.

35 In some embodiments, for B₆:

- 5 one R₁ for B₆ represents (CH₂)_m---ø, wherein ---ø represents the attachment point (representing a covalent bond) to the rest of the compound, and m is 1, 2, 3, 4, 5, 6, or 7; each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7.
- 10 In some embodiments, for Formulae F3-F4:
R_w is O or S;
for Formulae F5-F10:
Y₈ is selected from O, N, and NR, wherein R is an alkyl group or H;
15 Y₉ is H, CH₃, or an alkyl group, provided that when Y₈ is O, Y₉ is a CH₃ or higher alkyl group;
each Y₁₀ is independently selected from H, CH₃, F, CF₃, and OCH₃; and
i is 1, 2, or 3.
In some embodiments, the compound is selected from:
- 20 N-(3-(3-borono-5-nitrobenzamido)propyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS01);
N-(4-((4-(3-borono-5-nitrobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS02);
N-(4-((3-borono-5-nitrobenzamido)methyl)benzyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS03);
25 N-(3-((3-borono-5-nitrobenzamido)methyl)benzyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS04);
N-(4-(3-borono-5-nitrobenzamido)butyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS05);
N-(3-(3-borono-5-fluorobenzamido)propyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS06);
30 N-(3-(3-borono-5-fluorobenzamido)-2,2-dimethylpropyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS07);
bis(3-(3-borono-5-fluorobenzamido)propyl)glycine (DS08);
N-(4-((3-borono-5-fluorobenzamido)methyl)benzyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS09);
35 N-(3-((3-borono-5-fluorobenzamido)methyl)benzyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS10);

- 5 N-(2-(3-borono-5-fluorobenzamido)cyclohexyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS11);
- N-(3-(3-borono-4-fluorobenzamido)propyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS12);
- N-(4-((4-(3-borono-4-fluorobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-4-
10 fluorobenzoyl)glycine (DS13);
- N-(3-(3-borono-4-fluorobenzamido)-2,2-dimethylpropyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS14);
- N-(4-((3-borono-4-fluorobenzamido)methyl)benzyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS15);
- 15 N-(3-((3-borono-4-fluorobenzamido)methyl)benzyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS16);
- N-((1S,2R)-2-(3-borono-4-fluorobenzamido)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS17);
- N-((1S,2S)-2-(3-borono-4-fluorobenzamido)cyclohexyl)-N-(3-borono-4-
20 fluorobenzoyl)glycine (DS18);
- N-(3-(3-borono-5-bromobenzamido)propyl)-N-(3-borono-5-bromobenzoyl)glycine (DS19);
- N-(4-((4-(3-borono-5-bromobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-5-bromobenzoyl)glycine (DS20);
- 25 bis(3-(3-borono-5-bromobenzamido)propyl)glycine (DS21);
- N-(4-((3-borono-5-bromobenzamido)methyl)benzyl)-N-(3-borono-5-bromobenzoyl)glycine (DS22);
- N-(3-((3-borono-5-bromobenzamido)methyl)benzyl)-N-(3-borono-5-bromobenzoyl)glycine (DS23);
- 30 N-(2-(3-borono-5-bromobenzamido)cyclohexyl)-N-(3-borono-5-bromobenzoyl)glycine (DS24);
- N-(3-(4-borono-3-fluorobenzamido)propyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS25);
- N-(4-((4-(4-borono-3-fluorobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(4-borono-3-
35 fluorobenzoyl)glycine (DS26);

- 5 N-(3-(4-borono-3-fluorobenzamido)-2,2-dimethylpropyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS27);
bis(3-(4-borono-3-fluorobenzamido)propyl)glycine (DS28);
N-(4-((4-borono-3-fluorobenzamido)methyl)benzyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS29);
- 10 N-(3-((4-borono-3-fluorobenzamido)methyl)benzyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS30);
N-((1S,2R)-2-(4-borono-3-fluorobenzamido)cyclohexyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS31);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-(1-hydroxy-1,3-
- 15 dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS32);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pentyl)glycine (DS33);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-2,2-dimethylpropyl)glycine (DS34);
- 20 bis(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS35);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)benzyl)glycine (DS36);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-((1S,2R)-2-(1-
- 25 hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine (DS37);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butyl)glycine (DS38);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-((1S,2S)-2-(1-
- 30 hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine (DS39);
(R)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS40);
(S)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS41);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-
- 35 dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine (DS42);

- 5 N-(3-(4-borono-3,5-difluorobenzamido)propyl)-N-(4-borono-3,5-difluorobenzoyl)glycine (DS43);
N-(3-(4-borono-2-fluorobenzamido)propyl)-N-(4-borono-2-fluorobenzoyl)glycine (DS44);
N-(2-(N-ethyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)ethyl)-N-
10 (1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)glycine (DS45);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-N-(2-hydroxyethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)ethyl)glycine (DS46);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)hexyl)glycine (DS47);
15 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(4-((4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)methyl)cyclohexyl)glycine (DS48);
(2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS49);
20 ((2S,4S)-4-(3-borono-4-fluorobenzamido)-1-(3-borono-4-fluorobenzoyl)pyrrolidine-2-carbonyl)glycine (DS50);
((2S,4S)-4-(3-borono-5-nitrobenzamido)-1-(3-borono-5-nitrobenzoyl)pyrrolidine-2-carbonyl)glycine (DS51);
((2S,4S)-4-(5-borono-2-fluorobenzamido)-1-(5-borono-2-fluorobenzoyl)pyrrolidine-2-
25 carbonyl)glycine (DS52);
(S)-(1,4-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)piperazine-2-carbonyl)glycine (DS53);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-benzyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS54);
30 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(4-(trifluoromethyl)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS55);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-ethyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS56);
35 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-propyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS57);

- 5 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-isobutyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS58);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-((5-(thiophen-2-yl)pyridin-2-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS59);
- 10 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-isopentyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS60);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(quinolin-5-ylmethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS61);
- 15 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(2-(trifluoromethoxy)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS62);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(4-(methylsulfonyl)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS63);
- 20 (3-((2S,4S)-4-(5-borono-2-(methylsulfonyl)benzamido)-2-carbamoylpyrrolidine-1-carbonyl)-4-(methylsulfonyl)phenyl)boronic acid (DS64);
(4-(((3S,5S)-1-(4-borono-2,6-difluorobenzoyl)-5-carbamoylpyrrolidin-3-yl)carbamoyl)-3,5-difluorophenyl)boronic acid (DS65);
- 25 (R,E)-4,5-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pent-2-enoic acid (DS66);
(2S,4S)-1-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxamide (DS67);
- 30 N,N'-((2S,3S)-1-amino-1-oxobutane-2,3-diyl)bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide) (DS68);
(R)-3,4-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid (DS69);
3-((2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxamido)propanoic acid (DS70);
- 35

- 5 (S)-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid (DS71);
(R)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pentanoic acid (DS72);
(2S,4R)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS73);
10 (2S,4R)-1-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS74);
(2S,3S)-3-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-2-(1-hydroxy-7-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)butanoic acid (DS75);
15 (R)-5-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-(1-hydroxy-7-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)pentanoic acid (DS76);
20 ((2S,4S)-1-(5-borono-2-nitrobenzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS77);
((2S,4S)-1-(5-borono-2-(methylsulfonyl)benzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS78);
((2S,4S)-1-(3-borono-2,6-difluorobenzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS79);
25 (S)-(3-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS80);
(S)-(3-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS81);
30 (S)-(3-((3-boronobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS82);
(S)-(3-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS83);
(S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS84);
35

- 5 (S)-(5-((3-borono-N-(5,6-diamino-6-oxohexyl)-4-fluorobenzamido)methyl)-2-fluorophenyl)boronic acid (DS85);
(S)-(5-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS86);
(S)-(3-((3-borono-N-(5,6-diamino-6-oxohexyl)-4-fluorobenzamido)methyl)
10 phenyl)boronic acid (DS87);
(S)-(5-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS88);
(S)-(5-((4-borono-3-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS89);
15 (S)-(4-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS90);
(S)-(4-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS91);
(S)-(4-((3-boronobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic
20 acid (DS92);
(S)-(4-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS93);
(S)-(4-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS94);
25 (S)-(5-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)-2-fluorophenyl)boronic acid (DS95);
(S)-(3-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-bromophenyl)boronic acid (DS96);
(S)-(3-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)
30 phenyl)boronic acid (DS97);
(S)-(3-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)-5-methoxyphenyl)boronic acid (DS98);
(S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-bromophenyl)boronic acid (DS99);
35 (S)-(3-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic acid (DS100);

- 5 (S)-(3-((4-borono-3-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic acid (DS101);
- (S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic acid (DS102);
- (S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-2-fluorophenyl)boronic acid (DS103);
- 10 (S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-2,6-difluorophenyl)boronic acid (DS104);
- (S)-(3-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)phenyl)boronic acid (DS105);
- 15 (S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-3-methoxyphenyl)boronic acid (DS106);
- (S)-N-(5,6-diamino-6-oxohexyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS107);
- 20 (S)-N-(4-amino-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-oxobutyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS108);
- (S)-N-(6-amino-5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-6-oxohexyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-
- 25 dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS109);
- (2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS110);
- (2S,3S)-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid (DS111); and
- 30 (2S,4R)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS112).

In at least one embodiment, the compound of the present disclosure may be used, as an intermediate in the manufacture of a drug substance or a therapeutic of a prophylactic compound.

- 35 In another aspect, the disclosure provides a human insulin analog, comprising an A-chain and a B-chain, wherein the sequence of the A-chain comprises:

- 5 $X_{aa'}X_{bb'}X_{cc'}X_{dd'}X_{ee'}X_{ff'}X_{gg'}VEQCCX_{hh'}X_{ii'}ICSLYQLENYCNX_{jj'}X_{kk'}X_{ll'}X_{mm'}X_{nn'}X_{oo'}X_{pp'}$ (SEQ ID NO:24015); and wherein the sequence of the B-chain comprises:

(i)

$X_{aa'}X_{bb'}X_{cc'}X_{dd'}KX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}KX_{jj'}X_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24016),

- 10 wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y and W,

(ii)

- 15 $X_{aa'}X_{bb'}X_{cc'}X_{dd'}KPX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}X_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24017),

- wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y and W, and wherein X_{ee} is selected from amino acid residues A, E, F, H, I, K, L, N, P, Q, R, S, T, V, Y and W,

(iii)

$X_{aa'}X_{bb'}X_{cc'}X_{dd'}KX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}KX_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24018),

- 25 wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is present and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is G,

(iv)

30 $X_{aa'}X_{bb'}X_{cc'}X_{dd'}KX_{ee'}X_{ff'}X_{gg'}X_{hh'}X_{ii'}X_{jj'}KX_{kk'}X_{ll'}X_{mm'}X_{nn'}QHLGSHLVEALYLVCX_{oo'}X_{pp'}X_{qq'}GFFYT$
 $X_{rr'}X_{ss'}X_{tt'}X_{uu'}X_{vv'}X_{ww'}$ (SEQ ID NO:24019),

- wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A,

- 5 D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least one of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ is present and at least one of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ is S, or

(v)

$X_{aa}X_{bb}X_{cc}X_{dd}KX_{ee}X_{ff}X_{gg}X_{hh}X_{ii}X_{jj}KX_{kk}X_{ll}X_{mm}X_{nn}QHLCGSHLVEALYLVCX_{oo}X_{pp}X_{qq}GFFYT$
 $X_{rr}X_{ss}X_{tt}X_{uu}X_{vv}X_{ww}$ (SEQ ID NO:24020),

- 10 wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$, $X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least two of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ are present and at least one of $X_{ce}, X_{ff}, X_{gg}, X_{hh}, X_{ii}, X_{jj}$ is S, and another is G.

- 15 In some embodiments, the A-chain comprises a sequence selected from SEQ ID NOs 1 and 3 to 33, and is optionally appended at the N-terminus and/or at the C-terminus by at least one selected from KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY and SEQ ID NOs 75 to 24014, 24037-24046, and wherein the B-chain comprises at least one of SEQ ID NOs 2 and 34 to 74, 24047, and 24048, and is optionally appended at the N-terminus and/or at the C-terminus by at least one selected from KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN,
- 35

5 KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, 10 KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, 15 KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY and SEQ ID NOs 75 to 24014, 24037-24046.

In some embodiments, the A-chain comprises a sequence selected from SEQ ID NOs 1 and 3 to 33, the B-chain comprises at least one of SEQ ID NOs 2 and 34 to 74, 24047, and 20 24048, and

(a)

the A-chain and the B-chain are each independently and optionally appended at the N-terminus and/or at the C-terminus by at least one selected from:

KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, 25 KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, 30 KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, 35 KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY,

- 5 KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75 to 24014, 24037-24046,
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD,
- 10 KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII,
- 15 KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, QKD, KQE, QKF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT,
- 20 KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY and , KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 75 to 3224,
- 25 KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 3225 to 6374,
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 6375 to 15194,
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs
- 30 15195 to 18134,
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 18135 to 21074, and
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 21075 to 24014, 24037-24046, or
- 35 (b)

- 5 both the N-terminus and the C-terminus of the B-chain are independently covalently conjugated, via a peptide bond, to one selected from:
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE,
- 10 KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN,
- 15 KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, QKD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY,
- 20 KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75 to 24014, 24037-24036,
- KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD,
- 25 KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN,
- 30 KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY,
- 35

5 KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 75 to 3224,

KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 10 3225 to 6374,

KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 6375 to 15194,

KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 15195 to 18134,

15 KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 18135 to 21074, and

KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY and SEQ ID NOs 21075 to 24014, 24037-24036.

In some embodiments, the A-chain comprises a sequence selected from SEQ ID NOs 1
20 and 3 to 33, and is optionally appended at the N-terminus and/or at the C-terminus by at least one selected from KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD,
25 KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI,
30 KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS,
35 KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75 to 24014, KGSH (SEQ ID NO:24049), GKGS (SEQ ID NO:24050),

5 GKGSKK (SEQ ID NO:24045), GKKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGSHK (SEQ ID NO:24042); and wherein the B-chain comprises a sequence selected from SEQ ID NOs 2 and 34 to 74, 24047, and 24048, and is optionally appended at the N-terminus and/or at the C-terminus by at least one selected from KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, 10 KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, 15 KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, QKD, 20 KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75 to 25 24014, KGSB (SEQ ID NO:24049), GKGSB (SEQ ID NO:24050), GKGSKK (SEQ ID NO:24045), GKKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGSHK (SEQ ID NO:24042).

In some embodiments, no more than 4 residues are added or deleted from the A-chain and/or the B-chain of the insulin.

30 In some embodiments, a K residue is present at the N-terminus of the A-chain and/or the B-chain, and/or wherein no more than three K residues are present at the N-terminus of the A-chain and/or the B-chain, and/or wherein in (i) the tyrosine at A14 is replaced with glutamic acid, and/or (ii) the tyrosine at B16 is replaced with histidine, and/or (iii) the phenylalanine at B25 is replaced with a histidine, and/or wherein one to three residues selected from residues 35 B20, B21, and B22-B29 of the B-chain, residues A4 or A8 of the A-chain, and residues of an

5 optionally extended polypeptide, are lysine residues, and/or wherein only one K residue is present within 10 residues of the N-terminus of B-chain.

In some embodiments, X1 comprises an insulin and/or insulin analog as disclosed herein.

In some embodiments, X1 comprises an insulin and/or insulin analog as disclosed
10 herein, and the insulin and/or insulin analog further comprises an optional covalent-spacer.

In some embodiments, an amino group of one or more side chain(s) of one to four lysine residues of insulin is each independently covalently conjugated as described by Formula I.

In some embodiments, the insulin comprises at least two amines that are covalently
15 conjugated as described by Formula I, wherein one amine is the N-terminus amino group of the B-chain of insulin, and the other amine(s) are the side chain amine of a lysine that is 0 to 5 residues away from residue B22 of the B-chain of insulin, and/or the side chain amine of a lysine that is 1 to 5 residues away from residue A21 of the A-chain of insulin.

In some embodiments, an amino group at the N-terminus of the B-chain of insulin is
20 covalently conjugated as described by Formula I, and q' is optionally 2 or more, and the insulin includes at least one additional covalent conjugation to X1 as independently described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, such as in Examples 1-880, and wherein 1 to 4 residues are optionally added or deleted from
25 the A-chain and/or B-chain of the insulins shown in Examples 1-880.

In another aspect, the disclosure provides an insulin (e.g., a modified insulin) that may be used as an intermediate for the manufacture of a conjugate described by Formula I.

In some embodiments, each secondary amine in Formulae FF1-FF224 that is not conjugated to any of B₁, B₂ or B₃, is optionally independently acetylated or alkylated.

30 In some embodiments, the N-terminus of the A-chain and/or the N-terminus of the B-chain of insulin are additionally each independently covalently conjugated to at least two aromatic boron-containing groups, and/or wherein the C-terminus of the B-chain is further extended with a polypeptide of up to 20 residues, or the C-terminus of the A-chain is further extended with a polypeptide of up to 40 residues, each polypeptide independently comprising at
35 least one lysine residue in which the amino group of the lysine side chain is covalently conjugated as described by Formula I.

5 In some embodiments, X1 is insulin having a sequence comprising one selected from: a lysine at residue B21 of the B-chain and an arginine at residue B29 of the B-chain; a lysine at residue B21 of the B-chain; and a lysine at residue B29 of the B-chain; wherein an amino group of at least one lysine of the sequence is covalently conjugated as described by Formula I.

 In some embodiments, X1 is insulin having a sequence comprising: a lysine at residue
10 B21 of the B-chain; and at least one lysine at the N-terminus of the B-chain; wherein an amino group of at least one lysine of the sequence is covalently conjugated as described by Formula I.

 In some embodiments, the C-terminus of Z1a is conjugated through an amide linkage to a Z1b, and the Z1b is conjugated to the N-terminus of the B-chain of insulin through an amine linkage, and wherein the insulin is optionally further conjugated as described by Formula I.

15 In some embodiments, the compound comprises at least one Z1a comprising at least three amino acid residues having a side chain; the side chain of two residues of Z1a are conjugated together through a covalent bond included in at least one selected from a triazole linkage, an amide linkage, a disulfide linkage, a thioether linkage, a thiolene linkage, and an amine linkage; and the two conjugated residues are at least one residue apart.

20 In some embodiments, the N-terminal amine of a Z1a is covalently conjugated to a Z1c.

 In some embodiments, the compound comprises at least one Z1a comprising one or more amino acids selected from lysine, diaminopropionic acid, glycine, diaminobutyric acid, serine, histidine, and ornithine, and at least one or more of the side chains of the one or more amino acids, is covalently conjugated as described by Formula I.

25 In some embodiments, the compound comprises at least one Z1a comprising one or more glutamic or aspartic acid residues, and optionally other naturally occurring amino acids, and at least one lysine residue that is covalently conjugated as described by Formula I.

 In some embodiments, the compound comprises at least one Z1a comprising at least one lysine residue that is covalently conjugated as described by Formula I, wherein the majority of
30 the residues are negatively charged residues.

 In some embodiments, the insulin is covalently conjugated as described by Formula I, and Z1b is absent and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of insulin through a peptide bond.

 In some embodiments, the insulin is covalently conjugated as described by Formula I,
35 $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of insulin through a peptide bond; and Z1a comprises at least one amino acid from each of

5 groups M1 and M2 such that no two adjacent residues are from the same group and Z1a contains at least one lysine that is covalently conjugated as described by Formula I, wherein: (i) group M1 comprises lysine and alanine and group M2 comprises glycine, glutamic acid, serine, threonine, alanine and proline; or (ii) group M1 consists of lysine and alanine; and group M2 consists of glycine, glutamic acid, serine, threonine, alanine and proline.

10 19. In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of insulin through a peptide bond;

Z1a comprises at least one amino acid selected from K, P, E, G, S, T, A, and R, such that the sequence comprises at least one lysine, at least one proline, and at least one amino acid
15 selected from H, R, A and T; and

the amino group of least one lysine side chain in Z1a is covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of
20 Insulin through a peptide bond; and Z1a comprises at least one amino acid selected from the alanine, glycine, aspartic acid, threonine, histidine, methionine, cysteine, isoleucine, leucine, valine and glutamine; and at least one lysine having a side chain amino group that is covalently conjugated as described by Formula I; and the rest of the amino acids in Z1a are each independently selected from the twenty naturally occurring amino acids.

25 In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of Insulin through a peptide bond; Z1a has a sequence selected from: KPA, KPH, GKPA, GKPS, KP, GKPSG, and GKPGS; and Z1a comprises at least one lysine having a side chain amino group that is covalently conjugated as described by Formula I.

30 In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of Insulin through a peptide bond; Z1a comprises two or more copies of a sequence selected from: EGE, SGS, GSG, KP, GEG, E, GG, S, T, A, and R, such that no two adjacent copies are the same; Z1a optionally contains one or more of H, A, N and R; and the amino group of least one
35 lysine side chain in Z1a is covalently conjugated as described by Formula I.

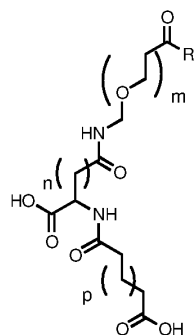
5 In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of Insulin through a peptide bond; Z1a comprises one or more amino acids selected from K, P, E, G, S, T, A, and R, such the sequence comprises at least one lysine, at least one proline, and at least one amino acid selected from H, R, A and T; and the amino group of least one lysine side
10 chain in Z1a is covalently conjugated as described by Formula I.

 In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of Insulin through a peptide bond; and at least one copy of KP is comprised in the polypeptide sequence of Z1a or insulin, wherein the amino group of the lysine side chain in KP is
15 covalently conjugated as described by Formula I.

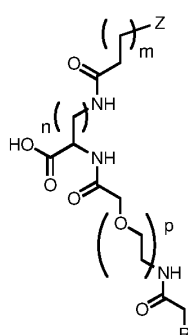
 In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the side chains of two residues of Z1a are conjugated via a covalent bond selected from a triazole, an amide bond, a disulfide bond, a thioether, a thiolene, and an amine; the two conjugated residues are separated by at least one amino acid; and the amino group of least one
20 lysine side chain in Z1a is covalently conjugated as described by Formula I.

 In some embodiments, the insulin is covalently conjugated as described by Formula I, $n'=0$ and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain of insulin through a peptide bond; Z1a is a polypeptide selected from the a polypeptide from the sequence of human insulin, a polypeptide from the sequence of human glucagon, a polypeptide
25 from the sequence of human C-peptide, a polypeptide from the sequence of human GLP-1, a polypeptide from the sequence of human GIP, a polypeptide from the sequence of human Extendin, and a human polypeptide hormone, and wherein the polypeptide comprises at least one lysine or Z1a contains at least one copy of dipeptide KP, and wherein the amino group of at least one lysine side chain in Z1a is covalently conjugated as described by Formula I.

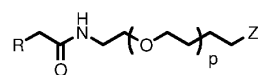
30 In some embodiments, at least one lysine residue, an inserted cysteine residue, or a residue that has been mutated to cysteine is covalently conjugated to a structure independently selected from Formulae F411-F416:



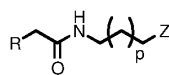
(F411)



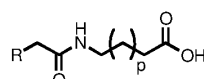
(F412)



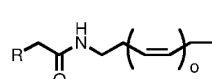
(F413)



(F414)



(F415)



(F416)

, and

5

wherein in Formulae F411-F416, R represents an attachment point to the amine group of the lysine side chain, or the thiol group of the cysteine side chain; n is an integer in the range of 1 to 14, m is an integer between 1 and 12, o is an integer between 1 and 6, p is an integer between 1 and 12; and Z represents one of $-(C=O)-OH$, $-NH_2$, $-CH_3$, a cholesterol, 7-OH cholesterol, 7,25-dihydroxycholesterol, cholic acid, chenodeoxycholic acid, lithocholic acid, deoxycholic acid, glycocholic acid, glycodeoxycholic acid, glycolithocholic acid, glycochenodeoxycholic acid, α -tocopherol, β -tocopherol, γ -tocopherol, δ -tocopherol, atocotrienol, β -tocotrienol, γ -tocotrienol, or δ -tocotrienol.

In some embodiments, the residue at position B29 of the B-chain of Insulin is a lysine covalently conjugated through an amide bond to the side chain of an L- or D- glutamic acid, and wherein the L- or D- glutamic acid is covalently conjugated through an amide bond to one of acid selected from hexanoic acid, myristic acid, octanoic acid, nonanoic acid, decanoic acid, lauric acid, stearic acid, and palmitic acid.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with the sequence $(XA_1X)_m$, wherein: A_1 is an L- or D-amino acid; m is an integer in the range of 1 to 4; and each X is K or KP; and the epsilon amine group

20

5 of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with the sequence $(XA_1A_2X)_m$ (SEQ ID NO:24021), wherein: A_1 and A_2 are each independently an L- or D-amino acid; m is an integer in the range of 1 to 4; each X is K or KP; and the epsilon amine group of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with the sequence $(XA_1A_2A_3X)_m$ (SEQ ID NO:24022), wherein: A_1 , A_2 , and A_3 are each independently an L- or D-amino acid; m is an integer in the range of 1 to 4; each X is K or KP; and the epsilon amine group of at least one lysine side chain in Z1a is covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with a sequence selected from $(XA_1X)_m(GGGGS)_n$ (SEQ ID NO:24023), $(XA_1A_2X)_m(GGGGS)_n$ (SEQ ID NO:24024), $(XA_1A_2A_3X)_m(GGGGS)_n$ (SEQ ID NO:24025), $(XA_1X)_m(GGGGS)_n(XA_2X)_o$ (SEQ ID NO:24026), and $(XA_1A_2X)_m(GGGGS)_n(XA_3A_4X)_o$ (SEQ ID NO:24027), wherein: A_1 , A_2 , A_3 , and A_4 are each independently an L- or D-amino acid; m is an integer in the range of 1 to 4; n is an integer in the range of 1 to 4; o is an integer in the range of 1 to 4; each X is K or KP; and the epsilon amine group of each lysine side chain of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with the sequence $(GX)_m$, wherein: X is KV; m is an integer in the range of 1 to 4, and the epsilon amine group of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with a sequence selected from: $GXA_1KG_2A_2XT)_m(GGSGSSS)_n$ (SEQ ID NO:24028), $(GXGXA_3GSSSGSSSXT)_o$ (SEQ ID NO:24029), $(GXGXA_3GSSSGSSSXT)_o$ (SEQ ID NO:24028), $(GXA_1ESA_2LYL)_m$ (SEQ ID NO:24029), $(TXEX)_m(GPGS)_n$ (SEQ ID NO:24030), $(GXESA_1VA)_m(KA_2K)_n$ (SEQ ID NO:24031), $(GXEA_1A_2)_m(GGS)_n(TYA_3XXT)_o$ (SEQ ID NO:24032), and $(TXAXYT)_m(TSSS)_n$ (SEQ ID NO:24033), wherein: each X is KV or KP; A_1 , A_2 , A_3 are each independently an L- or D-amino acid; m is an integer in the range of 1 to 4; n is an integer in the range of 1 to 4; and o is

5 an integer in the range of 1 to 4; and the epsilon amine group of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.

In some embodiments, the insulin is covalently conjugated as described by Formula I, Z1a comprises a polypeptide with a sequence selected from (TKPYA1KEVETA2GSGS)_m (GGGGS)_n (SEQ ID NO:24034),
10 (YTPLEA1KPYSTSYKPYSEA1L)_m(GKPTSLEA2FLVEA2LYTKP)_n (SEQ ID NO:24035), and
(GKEALYLTPLESALYKP)_m(TKPLEALYLKPEILSLKPESLA)_n(GKPGSSSKPDTSSSGTPKTAAGS)_o (SEQ ID NO:24036), wherein: A₁ and A₂ are each independently an L- or D-amino acid; m is an integer in the range of 1 to 4; n is an integer in the range of 1 to 4; and the
15 epsilon amine group of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.

In some embodiments, the compound is conjugated either directly, or via an optional covalent-spacer, to a drug molecule, an imaging agent, a contrast agent, a radioactive isotope, a radiotherapy agent, or a molecule that engages immune cells in the body.

20 In some embodiments, X1 is human glucagon or an analogue of human glucagon, and optionally covalently conjugated to one or more diol- or sugar- containing molecules, or X1 is an analogue of a human peptide hormone that is modified so that it binds to its cognate receptor but has diminished or null ability to activate the receptor in the body, or X1 is an analogue of a human peptide hormone that is modified so that it selectively binds or activates a subset of its
25 cognate receptors or subsets of receptors of human polypeptide hormones.

In some embodiments, the aromatic boron-containing groups are modified to be MIDA protected, pinacol protected, or in an ester form. In some embodiments, the aromatic boron-containing group is MIDA protected or pinacol protected.

In some embodiments, the modified aromatic boron-containing groups are used as
30 intermediates for the synthesis of a conjugate of Formula I.

In some embodiments, X1 comprises: (i) a human polypeptide hormone or an analogue of a human polypeptide hormone, wherein the covalent linkage to X1 is to an amine or via an optional covalent-spacer to an amine in X1; (ii) an amine configured to be covalently conjugated via an optional covalent-spacer to a human polypeptide hormone or an analogue of
35 a human polypeptide hormone, or (iii) NH₂, and wherein the amine in X1 is covalently conjugated twice as described by Formula I, wherein the first covalent conjugation is through

5 an amine bond and the second covalent conjugation is through an amide bond, and wherein each covalent conjugation is the same or different.

In some embodiments, residue B21 of the B-chain of Insulin is K, residue B22 of the B-chain of Insulin is P, and residue B29 of the B-chain of Insulin is R; and the N-terminus of the Insulin B-chain is covalently conjugated as described by Formula I to the C-terminus of Z1a,
10 wherein: $n'=0$; Z1a has the sequence GKPGHKP; and one Z1c is attached to each lysine side chain in Z1a, wherein each Z1c is independently represented by Formula FF12, and each of B1 and B2 are represented by Formula F2, wherein one R1 at position 5' is the covalent amide bond to Formula FF12, and the amino group of the lysine at B21 (residue 21 of the B-chain of Insulin) is covalently conjugated as described by Formula I, wherein $n'=0$; $m'=0$; and Z1c is
15 described by Formula FF12, wherein each of B1 and B2 are represented by formula F2, and wherein one R1 at position 5' is the covalent amide bond to Formula FF12.

In at least one embodiment, the present disclosure is directed to an insulin analog. In some embodiments, the insulin analog is desB30 human insulin; wherein the N-terminus of the Insulin B-chain is covalently conjugated as described by Formula I to the C-terminus of Z1a,
20 wherein: $n'=0$; Z1a has the sequence KPGSEHESA, and one Z1c is attached to each lysine side chain in Z1a, wherein each Z1c is described by Formula FF1, and each of B1 and B2 are described by Formula F1, wherein one R1 at position 3' is the covalent amide bond to Formula FF1 and wherein one R1 at position 5' is F, and wherein the amino group of the lysine at B29 (residue 29 of the B-chain of Insulin) is covalently conjugated as described by Formula I,
25 wherein $n'=0$; $m'=0$; and Z1c is described by Formula FF1, and each of B1 and B2 are described by Formula F1, wherein one R1 at position 3' is the covalent amide bond to Formula FF1 and one R1 at position 5' is F.

In some embodiments, the A- and/or B-chain sequence of the insulin is appended at the N-terminus or C-terminus by $KX'K$, KX' , or $X'K$ wherein X' represents a continuous sequence
30 of 2, 3, 4, or 5 residues selected from within wild-type A-chain (SEQ ID NO:1) and wild-type B-chain (SEQ ID NO:2). In some embodiments, each K residue is optionally and independently covalently conjugated as described by Formula I. In some embodiments, X' is a polypeptide of up to 30 residues with amino acids independently selected from: K, G, S, E, H, E, N, Q, D, A, P, R and C and each K residue is optionally and independently covalently conjugated as
35 described by Formula I.

5 In some embodiments, the N-terminus of the A-chain and/or B-chain are optionally and independently covalently conjugated as described by Formula I.

 In some embodiments, each K residue when present in insulin (insulin analog) is optionally and independently covalently conjugated as described by Formula I, wherein Z1c is any one of formulae FF1-FF224 and the B1 and B2 are each independently selected from F1
10 and F2.

 In some embodiments, each K residue when present in insulin is optionally and independently covalently conjugated as described by Formula I, wherein Z1c is any one of formulae:

 Formulae FF1-F22 and the B1 and B2 are each independently selected from F1 and F2;
15 Formulae FF23-FF48 and the B1 and B2 are each independently selected from F1 and F2;

 Formulae FF49-FF88 and the B1 and B2 are each independently selected from F1 and F2;

 Formulae FF89-FF112 and the B1 and B2 are each independently selected from F1 and
20 F2;

 Formulae FF113-FF136 and the B1 and B2 are each independently selected from F1 and F2;

 Formulae FF137-FF160 and the B1 and B2 are each independently selected from F1 and F2;

25 Formulae FF160-FF166 and the B1 and B2 are each independently selected from F1 and F2; or

 Formulae FF167-FF224 and the B1 and B2 are each independently selected from F1 and F2.

 In some embodiments, Z1b is optionally selected from Formula IIa-Formula IIIi; and/or
30 optionally selected from Formula IIIa-Formula IIIi.

 In at least some embodiments, the insulin does not comprise Z1a and/or Z1b. In some embodiments, Z1a is not present. In some embodiments, Z1b is not present. In at least some embodiments, Z1a and/or Z1b are not present.

 In some embodiments, each K residue when present in insulin is optionally and
35 independently covalently conjugated as described by Formula I, and wherein Z1c is any one of formulae FF1-FF224 and the B1 and B2 are each independently selected from F3 and F4.

- 5 In some embodiments, each K residue when present in insulin is optionally and independently covalently conjugated as described by Formula I, wherein Z1c is any one of formulae:
- Formulae FF1-F22 and the B1 and B2 are each independently selected from F3 and F4;
Formulae FF23-FF48 and the B1 and B2 are each independently selected from F3 and
10 F4;
Formulae FF49-FF88 and the B1 and B2 are each independently selected from F3 and F4;
Formulae FF89-FF112 and the B1 and B2 are each independently selected from F3 and F4;
15 Formulae FF113-FF136 and the B1 and B2 are each independently selected from F3 and F4;
Formulae FF137-FF160 and the B1 and B2 are each independently selected from F3 and F4;
Formulae FF160-FF166 and the B1 and B2 are each independently selected from F3
20 and F4; or
Formulae FF167-FF224 and the B1 and B2 are each independently selected from F3 and F4.
- In some embodiments, Z1b is optionally selected from Formula IIa-Formula IIIi and Formula IIIa-Formula IIIi.
- 25 In some embodiments, Z1a and/or Z1b are not present.
- In some embodiments, each K residue is optionally and independently covalently conjugated as described by Formula I, wherein Z1c is any one of Formulae FF1-F224 and the B1 and B2 are each independently selected from F5, F6, F7, and F8.
- In some embodiments, each K residue when present in insulin is optionally and
30 independently covalently conjugated as described by Formula I, wherein Z1c is any one of formulae:
- Formulae FF1-F22 and the B1 and B2 are each independently selected from F5, F6, F7, and F8;
Formulae FF23-FF48 and the B1 and B2 are each independently selected from F5, F6,
35 F7, and F8;

5 Formulae FF49-FF88 and the B1 and B2 are each independently selected from F5, F6, F7, and F8;

 Formulae FF89-FF112 and the B1 and B2 are each independently selected from F5, F6, F7, and F8;

 Formulae FF113-FF136 and the B1 and B2 are each independently selected from F5, F6, F7, and F8;

 Formulae FF137-FF160 and the B1 and B2 are each independently selected from F5, F6, F7, and F8;

 Formulae FF160-FF166 and the B1 and B2 are each independently selected from F5, F6, F7, and F8; or

15 Formulae FF167-FF224 and the B1 and B2 are each independently selected from F5, F6, F7, and F8.

 In some embodiments, Z1b is optionally selected from Formula IIa-Formula IIIi and Formula IIIa-Formula IIIi.

 In some embodiments, Z1a and/or Z1b are not present.

20 In some embodiments, each K residue when present in insulin is optionally and independently covalently conjugated as described by Formula I, wherein Z1c is any one of formulae FF1-F224 and the B1 and B2 are each independently selected from F9 and F10.

 In some embodiments, each K residue when present in insulin is optionally and independently covalently conjugated as described by Formula I, wherein Z1c is any one of
25 formulae:

 Formulae FF1-F22 and the B1 and B2 are each independently selected from F9 and F10;

 Formulae FF23-FF48 and the B1 and B2 are each independently selected from F9 and F10;

30 Formulae FF49-FF88 and the B1 and B2 are each independently selected from F9 and F10;

 Formulae FF89-FF112 and the B1 and B2 are each independently selected from F9 and F10;

 Formulae FF113-FF136 and the B1 and B2 are each independently selected from F9
35 and F10;

5 Formulae FF137-FF166 and the B1 and B2 are each independently selected from F9 and F10; or

 Formulae FF167-FF224 and the B1 and B2 are each independently selected from F9 and F10.

 In some embodiments, Z1b is not present in insulin (i.e., insulin analog) and wherein
10 Z1a is a polypeptide that is covalently linked by a peptide bond to the N-terminus of the B-chain of insulin and/or Z1a is a polypeptide that is covalently linked by a peptide bond to the N-terminus of the B-chain of insulin C-terminus of A-chain of insulin.

 In some embodiments, Z1b may be present in insulin. If present in insulin, Z1b is optionally selected from Formula IIa-Formula IIIi and Formula IIIa-Formula IIIi.

15 In some embodiments, Z1b is not present in insulin. In some embodiments, Z1a is not present. In at least some embodiments, Z1b and/or Z1a are not present.

Methods of Preparation

 In at least one embodiment, the present disclosure provides a method to prepare a
20 compound comprising an aromatic boron-containing compound and/or an aromatic boron-containing group (e.g., Z1c, Formula I) or a pharmaceutical preparation comprising one or more compounds of the present disclosure.

 In at least one embodiment, the disclosure provides a method for preparing rotationally constrained tether boron conjugates that contain scaffolds (Z1c) that are rotationally hindered
25 by disfavored steric interactions (e.g. gauche vs anti interactions of substituents), hindered rotation due to bond hybridization (e.g., cis- vs trans- amide rotations), or through rigid covalent bonds (e.g., (*E*) vs (*Z*) configurations for alkene moieties). For example, formulae FF50 – FF62, FF116, and FF121-134 contain alkyl functionalities geminal (e.g., attached to the same atom) to the amine groups that are covalently conjugated to the boronic acid
30 functionalized moieties. As another example, Formulae FF50 – FF62, FF116, and FF121-134 contain geminal alkyl substituents which may limit the accessible dihedral angles that the boron conjugated amines adopt, influencing adopted dihedral angles and placing the boronic functionalized groups closer together and allowing for increased binding of the conjugates to target molecules such as proteins or sugars.

35

Methods of Treatment

5 In at least one embodiment, the disclosure provides a method of treating a subject suffering from, or susceptible to, a disease that is beneficially treated by a compound disclosed herein or a pharmaceutical preparation comprising one or more of the compounds disclosed herein. In some embodiments, the method comprises the step of administering to a subject in need thereof an effective amount of a pharmaceutical preparation/composition of
 10 the present disclosure. In at least one embodiment, the compound(s) and/or pharmaceutical preparations of the present disclosure may be for use in (or in the manufacture of medicaments for) the treatment or prevention of disorders, including hyperglycemia, type 2 diabetes, impaired glucose tolerance, type 1 diabetes, obesity, metabolic syndrome X, or dyslipidemia, diabetes during pregnancy, pre-diabetes, Alzheimer's disease, MODY 1, MODY 2 or MODY 3
 15 diabetes, mood disorders, and psychiatric disorders. In at least one embodiment, a therapeutically-effective amount of a compound and/or pharmaceutical preparation of the present disclosure is administered to a subject suffering from diabetes.

The following examples and experimental data are provided for illustrative purposes only, and do not limit the scope of the embodiments of the present disclosure.

20 The following abbreviations have the definitions set forth below:

Abbreviation	Full Name
Acm	s-Acetamidomethyl group
ACN	Acetonitrile
ARS	Alizarin Red S
Boc	tert-Butyloxycarbonyl
DCM	Dichloromethane
Dde	1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)ethyl
DIPEA, DIEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DTDP	2,2,-Dithiopyridine
EDC	3-(Ethyliminomethyleneamino)-N,N-dimethylpropan-1-amine
Fmoc	Fluorenylmethyloxycarbonyl Chloride
GLP-1	Glucagon-Like Peptide 1
HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
HCl	Hydrochloric Acid
IGF1	Insulin-like Growth Factor 1
LC-MS	Liquid Chromatography-Mass Spectrometry

Abbreviation	Full Name
MeOH	Methanol
MIDA	N-methyliminodiacetic acid
MODY	Maturity Onset Diabetes of the Young
NHS	N-Hydroxysuccinimide
Oxyma	Ethyl cyanohydroxyiminoacetate
RAM	Rink Amide Matrix
PEG	Polyethylene Glycol
SDS	Sodium Dodecyl Sulfate
tBu	tert-Butyl
TFA	Trifluoroacetic Acid

5

Examples

A. Preparation of aromatic boron-containing compounds.

The disclosed compounds can be prepared according to the following schemes. The following schemes represent the general methods used in preparing these compounds.

- 10 However, the synthesis of these compounds is not limited to these representative methods, as they can also be prepared through various other methods by those skilled in the art of synthetic chemistry.

Method 1: Synthesis of Diboronates DS01-DS48:

- Chlorotriptyl resin (300 mg, 0.45 mmol) was swelled in dry DCM (5 mL) for 30 mins.
- 15 Solvent was removed with nitrogen and a solution of bromo acetic acid (0.139 g, 1M, 1 mL) in DCM with DIPEA (1M, 0.13 g 1 mL) was added immediately and gently mixed for 1 hr. The mixture was washed with DCM and unreacted sites were capped with a solution of 20% MeOH in a solution of DCM and DIEA (1M) and mixed for 1hr. The resin was washed with DCM (3x5 mL) then DMF (3x5 mL). A solution of FF-diamine linker propane-1,3-diamine (0.37 g, 1M) in DMF (5mL) was added to the resin and heated at 50 °C for 10 minutes. The resin was
- 20 washed with DMF (3x5 mL) and a solution of 3-borono-5-nitrobenzoic acid (0.2M, 0.21 mg, 5 mL) in DMF with N, N'-diisopropylcarbodiimide (DIC) (0.126 g, 1M, 1mL), Oxyma (0.5 M, 0.142 g, 2mL) in DMF and heated at 50 °C for 30 min. The resin was washed with DMF (3x 5mL) then DCM (3x 5 mL). A solution of trifluoroacetic acid with triisopropyl silane and water
- 25 (95:2.5:2.5, 5 mL) was added to the resin and mixed for 90 minutes. The solution was collected and dried under vacuum, dissolved in DMSO (100 uL) and fractionated by reverse-phase (RP) flash chromatography on a C18 column with a gradient of 20% ACN in water with 0.1% TFA

5 to 60% ACN in water with 0.1% TFA over 10 minutes. Pure fractions were isolated, combined, frozen, and lyophilized to yield N-(3-(3-borono-5-nitrobenzamido)propyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS01) as a white powder (7 mg).

Similar procedures are followed for the synthesis of aromatic boron-containing groups DS02-DS48.

10 **Method 2: Synthesis of Symmetric Diboronates With C-Terminus Linker DS49-DS53**

Tentagel-S-NH₂ resin (250 mg, 0.05 mmol) was swelled in DMF (5 mL) for 2hr. The solution was removed under a stream of nitrogen and a solution of Boc-Gly-HMBA (0.2 mmol) was coupled using 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU, 76 mg, 0.2 mmol), and DIPEA (70 ul) in DMF (2 mL) was added to the resin and mixed at room temperature for 45 minutes. The resin was washed with DMF (3x5 mL) and DCM (3x5 mL). A solution of 50% trifluoroacetic acid in DCM (5mL) was added to the resin and mixed for 20 minutes to remove the Boc protecting group. This step was repeated twice. The resin was washed with DCM (3x5 mL) and DMF (3x5 mL) and was treated with a solution of 10% DIEA in DMF (5 mL) for 10 minutes, the cycle was repeated twice, and resin was washed with DMF(3x5 mL). A solution of 1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)pyrrolidine-2-carboxylic acid (0.115 g, 0.2 mmol) with 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU, 0.2 mmol), and DIPEA (70 ul) in DMF (2 mL) was added to the resin and mixed for 45 minutes. The resin was washed with DMF (3x5 mL) and a solution of 20% piperidine in DMF (5 mLx3) was added to resin and mixed for 5 minutes. The resin was washed with DMF (3x5 mL) and a solution of 1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxylic acid (0.078 g, 0.4 mmol) with HATU (0.4 mmol) and DIPEA (140 uL) in DMF (2 mL) was added to the resin and mixed for 45 minutes. The resin was washed with DMF (3x5 mL) then DCM (3x5 mL). A solution of 0.1 M NaOH in 1:5 water:THF was added to the resin and mixed for 90 minutes. The solution was filtered and adjusted the pH~2 using 1.0 M HCl and fractionated by reverse-phase (RP) flash chromatography on a C18 column with a gradient of 20% ACN in water with 0.1% TFA to 60% ACN in water with 0.1% TFA over 10 minutes. Pure fractions were isolated, combined, frozen, and lyophilized to yield ((2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS49) as a white powder (10 mg).

5 Similar procedures are followed for the synthesis of aromatic boron-containing groups DS50-DS53.

Method 3: Synthesis of Diboronates with Reductive Alkylation on Side Chain Amine DS54-DS63

Rink-amide resin (0.05 mmol, 263 mg) was swelled in DMF (5 mL) for 20 minutes.

10 The solution was removed under a stream of nitrogen and a solution of 20% piperidine in DMF (5mL) was added to resin and mixed for 5 minutes. The resin was washed with DMF (3x5 mL). A solution of ((S)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-((diphenyl(*p*-tolyl)methyl)amino) propanoic acid (116 mg, 0.2 mmol) with 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide

15 hexafluorophosphate (HATU, 76 mg, 0.2 mmol), and DIPEA (70 ul) in DMF (2 mL) was added to the resin and mixed at room temperature for 45 minutes. The resin was washed with DMF (3x5 mL) and DCM (3x5 mL). A solution of 0.2% trifluoroacetic acid in DCM (5 mL) was added to the resin and mixed for 10 minutes to remove Mtt protecting group. This step was repeated twice. The resin was washed with DCM (3x5 mL) and DMF (3x5 mL) and was treated

20 with a solution of 10% DIEA in DMF (5 mL) for 10 minutes, the cycle was repeated twice, and resin was washed with DMF(3x5 mL). A solution of benzaldehyde (0.053 g, 0.5 mol) in trimethyl orthoformate (TMOF) (2mL) was added to the resin and mixed for 1hr. The solution was filtered, and resin was washed with DMF (3x5 mL) and 10% acetic acid in methanol (2x5 mL). The solution of NaCNBH₃ (10 equivalents) in 10% acetic acid in methanol was added to

25 the resin and mixed for 1hr, washed with DMF (3x5 mL). The resin was treated with 20% piperidine in DMF (3x5mL) to deprotect the Fmoc, washed with DMF (3x5 mL) and coupled with 1-hydroxy-1,3-dihydrobenzo[*c*][1,2]oxaborole-6-carboxylic acid (0.078 g, 0.4 mmol) using 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide

30 hexafluorophosphate (HATU, 152 mg, 0.4 mmol), and DIPEA (140 ul) in DMF (2 mL) for 45 minutes. After coupling reaction, the resin was washed with DMF (3x5mL) then with DCM (2x5 mL). A solution of trifluoroacetic acid with triisopropyl silane and water (95:2.5:2.5, 5 mL) was added to the resin and mixed for 90 minutes. The solution was collected and dried under vacuum, dissolved in DMSO (100 uL) and fractionated by reverse-phase (RP) flash chromatography on a C18 column with a gradient of 20% ACN in water with 0.1% TFA to

35 60% ACN in water with 0.1% TFA over 10 minutes. Pure fractions were isolated, combined, frozen, and lyophilized to yield (S)-N-(3-amino-2-(1-hydroxy-1,3-

- 5 dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-benzyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS54) as a white powder (7 mg).

Similar procedures are followed for the synthesis of aromatic boron-containing groups DS55-DS63.

Method 4: Synthesis of Symmetric Diboronates Based on Amino Acids DS64-DS76,

10 **DS109-DS111**

- Rink-amide resin (0.05 mmol, 263 mg) was swelled in DMF (5 mL) for 20 minutes. The solution was removed under a stream of nitrogen and a solution of 20% piperidine in DMF (5 mL) was added to resin and mixed for 5 minutes. The resin was washed with DMF (3x5 mL). A solution of 1-(((9H-fluoren-9-yl)methoxy)carbonyl)-4-(((9H-fluoren-9-yl)methoxy)carbonyl) amino) pyrrolidine-2-carboxylic acid (0.115 g, 0.2 mmol) with 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU, 0.2 mmol), and DIPEA (70 μ L) in DMF (2 mL) was added to the resin and mixed at 50 °C for 20 minutes. The resin was washed with DMF (3x5 mL) and a solution of 20% piperidine in DMF (5 mL) was added to the resin and mixed for 5 minutes.
- 20 The resin was washed with DMF (3x5 mL) and a solution of 1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxylic acid (0.078 g, 0.4 mmol) with HATU (0.4 mmol) and DIPEA (140 μ L) in DMF (2mL) was added to the resin and mixed at 50 °C for 30 minutes. The resin was washed with DMF (3x5mL) then with DCM (3x5 mL). A solution of trifluoroacetic acid with triisopropyl silane and water (95:2.5:2.5, 5 mL) was added to the resin
- 25 and mixed for 90 minutes. The solution was collected and dried under vacuum, dissolved in DMSO (100 μ L) and fractionated by reverse-phase (RP) flash chromatography on a C18 column with a gradient of 20% ACN in water with 0.1% TFA to 60% ACN in water with 0.1% TFA over 10 minutes. Pure fractions were isolated, combined, frozen, and lyophilized to yield
- 30 (3-((2S,4S)-4-(5-borono-2-(methylsulfonyl)benzamido)-2-carbamoylpyrrolidine-1-carbonyl)-4-(methylsulfonyl)phenyl)boronic acid (DS64) as a white powder (10 mg).

Similar procedures are followed for the synthesis of aromatic boron-containing groups DS65-DS76. Similar procedure can be followed for the synthesis of diboronate sensors DS109-DS111 noting that the resin used is 2-chlorotrityl resin and not the Amine Rink-amide resin.

Method 5: Asymmetric Diboronate Synthesis DS77-DS79

5 Tentagel-S-NH₂ resin (250 mg, 0.05 mmol) was swelled in DMF (5 mL) for 2hr. The solution was removed under a stream of nitrogen and a solution of Boc-Gly-HMBA (0.2 mmol) was coupled using 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU, 76mg, 0.2mmol), and DIPEA (70 ul) in DMF (2 mL) was added to the resin and mixed at room temperature for 45 minutes. The resin was washed with
10 DMF (3x5 mL) and DCM (3x5 mL). A solution of 50% trifluoroacetic acid in DCM (5mL) was added to the resin and mixed for 20 minutes. To remove Boc protecting group. This step was repeated twice. The resin was washed with DCM (3x5 mL) and DMF (3x5 mL) and was treated with a solution of 10% DIEA in DMF (5 mL) for 10 minutes, the cycle was repeated twice, and resin was washed with DMF(3x5 mL). A solution of 1-(((9H-fluoren-9-yl)methoxy)carbonyl)-
15 4-((*tert*-butoxycarbonyl)amino)pyrrolidine-2-carboxylic acid (0.09 g, 0.2 mmol) with 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU, 0.2 mmol), and DIPEA (70 ul) in DMF (2 mL) was added to the resin and mixed for 45 minutes. The resin was washed with DMF (3x5mL) and a solution of 20% piperidine in DMF (5 mLx3) was added to resin and mixed for 5 minutes. The resin was
20 washed with DMF (3x5 mL) and a solution of 1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxylic acid (0.039 g, 0.2 mmol) with HATU (0.076 g, 0.2 mmol) and DIPEA (140 uL) in DMF (2 mL) was added to the resin and mixed for 45 minutes. and DCM (3x5 mL). A solution of 50% trifluoroacetic acid in DCM (5 mL) was added to the resin and mixed for 20 minutes, to remove Boc protecting group. This step was repeated twice. The resin was washed with DCM
25 (3x5 mL) and DMF (3x5 mL) and was treated with a solution of 10% DIEA in DMF (5 mL) for 10 minutes, the cycle was repeated twice, and resin was washed with DMF(3x5 mL). A solution of 5-borono-2-nitrobenzoic acid (0.042 g, 0.2 mmol) with HATU (0.076 g, 0.2 mmol) and DIPEA (140 uL) in DMF (2mL) was added to the resin and mixed for 45 minutes. The resin was washed with DMF (3x5 mL) then DCM (3x5 mL). A solution of 0.1M NaOH in 1:5
30 water:THF was added to the resin and mixed for 90 minutes. The solution was filtered and adjusted the pH~2 using 1.0 M HCl and fractionated by reverse-phase (RP) flash chromatography on a C18 column with a gradient of 20% ACN in water with 0.1% TFA to 60% ACN in water with 0.1% TFA over 10 minutes. Pure fractions were isolated, combined, frozen, and lyophilized to yield ((2S,4S)-1-(5-borono-2-nitrobenzoyl)-4-(1-hydroxy-1,3-
35 dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS77) as a white powder (8 mg).

5 Similar procedures are followed for the synthesis of aromatic boron-containing groups DS78-DS79.

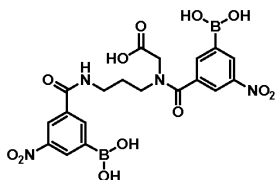
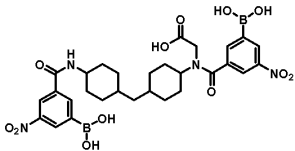
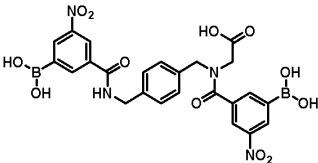
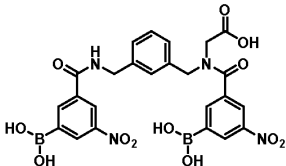
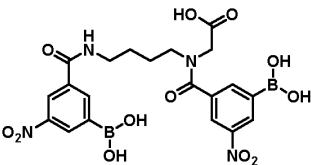
Method 6: Synthesis of Diboronates with Reductive Alkylation on Side Chain DS80-DS109

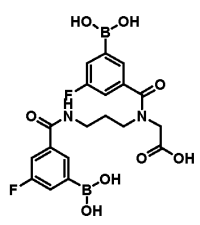
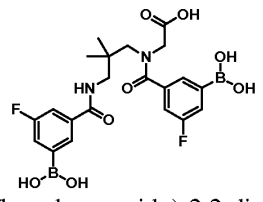
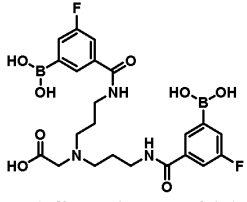
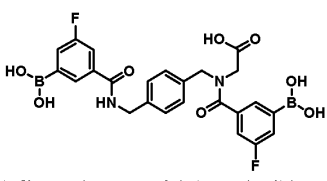
Amine Rink-amide resin (0.05 mmol, 263 mg) was swelled in DMF (5 mL) for 20 minutes. The solution was removed under a stream of nitrogen and a solution of 20% piperidine
10 in DMF (5 mL) was added to resin and mixed for 5 minutes. The resin was washed with DMF (3 x 5 mL). A solution of *N*⁶-(((9*H*-fluoren-9-yl)methoxy)carbonyl)-*N*2-(*tert*-butoxycarbonyl)lysine (93.6 mg, 0.2 mmol) with 1-[Bis(dimethylamino)methylene]- 1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU, 76 mg, 0.2 mmol), and DIPEA (70 μ L) in DMF (2 mL) was added to the resin and mixed at room temperature for 45 minutes.
15 The resin was washed with DMF (3x5 mL). The Fmoc was removed with 20% piperidine in DMF (2x5 mL) and then washed with additional DMF (3 x10 mL) A solution of 4-methyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzaldehyde (123 mg, 0.5 mol) in trimethyl orthoformate (TMOF) (2 mL) was added to the resin and mixed for 1hr. The solution was filtered, and resin was washed with DMF (3x5 mL) and a solution of NaBH₄ (10 equivalents)
20 in 20% methanol in DMF was added to the resin and mixed for 1hr, washed with DMF (3x5 mL). The reduced amine was then coupled with 3-nitro-5-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzoic acid (117 mg, 0.4 mmol) using 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate (HATU, 152 mg, 0.4 mmol), and DIPEA (140 μ L) in DMF (2 mL) for 45 minutes. After coupling reaction, the resin was
25 washed with DMF (3x5 mL) then with DCM (2x5 mL). A solution of trifluoroacetic acid with triisopropyl silane and water (95:2.5:2.5, 5 mL) was added to the resin and mixed for 90 minutes. The solution was collected and dried under vacuum, dissolved in DMSO (100 μ L) and fractionated by reverse-phase (RP) flash chromatography on a C18 column with a gradient of 20% ACN in water with 0.1% TFA to 60% ACN in water with 0.1% TFA over 10 minutes.
30 Pure fractions were isolated, combined, frozen, and lyophilized to yield diboronate sensor DS80 as a white powder (10 mg). Similar procedure was followed for the synthesis of diboronate sensors DS81-DS109.

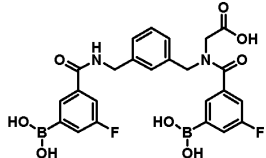
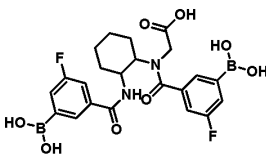
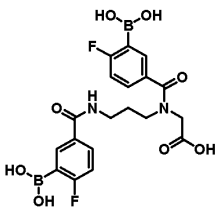
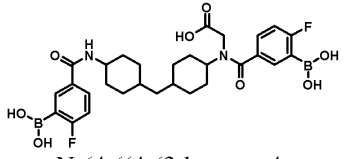
The chemical structure and IUPAC name of DS1 to DS109 are summarized in Table 1 below.

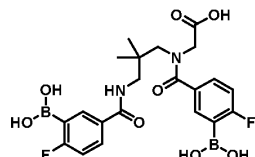
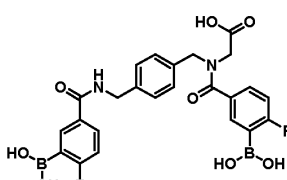
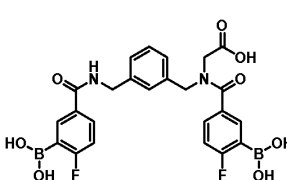
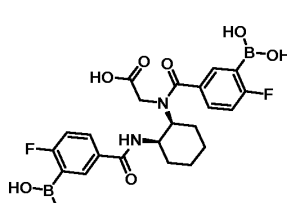
35

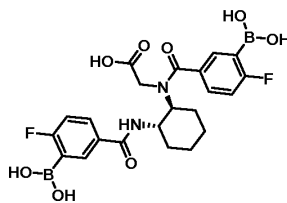
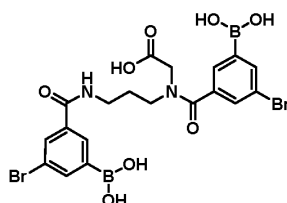
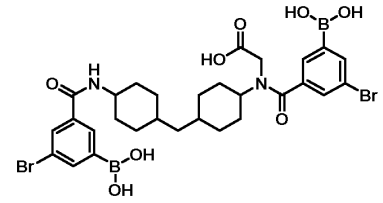
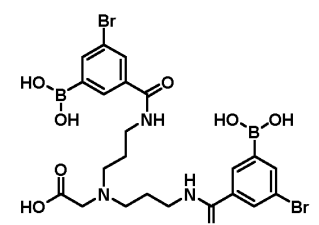
Table 1

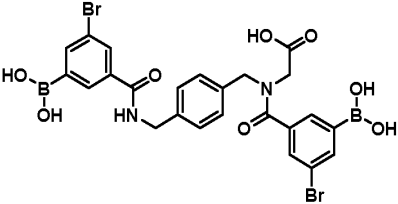
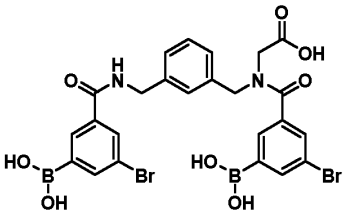
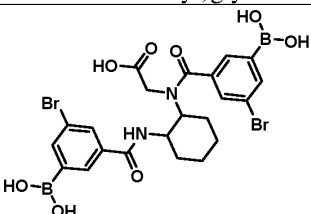
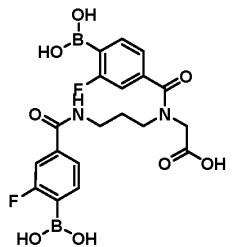
Compd #	FF	F	Structure/IUPAC Name
DS01	FF10	F1	 N-(3-(3-borono-5-nitrobenzamido)propyl)-N-(3-borono-5-nitrobenzoyl)glycine
DS02	FF217	F1	 N-(4-((4-(3-borono-5-nitrobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-5-nitrobenzoyl)glycine
DS03	FF8	F1	 N-(4-((3-borono-5-nitrobenzamido)methyl)benzyl)-N-(3-borono-5-nitrobenzoyl)glycine
DS04	FF14	F1	 N-(3-((3-borono-5-nitrobenzamido)methyl)benzyl)-N-(3-borono-5-nitrobenzoyl)glycine
DS05	FF10	F1	 N-(4-(3-borono-5-nitrobenzamido)butyl)-N-(3-borono-5-nitrobenzoyl)glycine

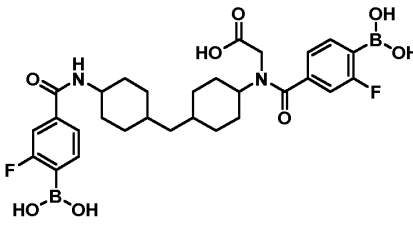
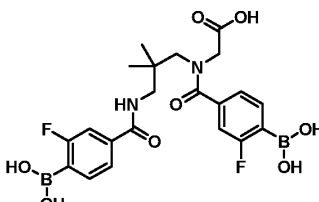
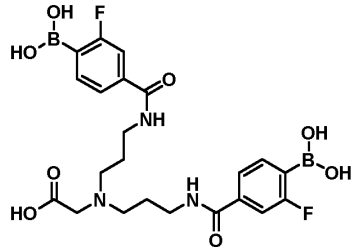
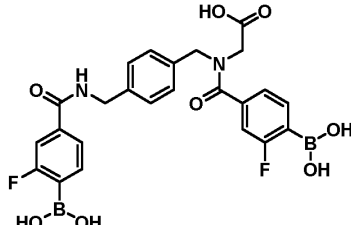
Compd #	FF	F	Structure/IUPAC Name
DS06	FF10	F1	 N-(3-(3-borono-5-fluorobenzamido)propyl)-N-(3-borono-5-fluorobenzoyl)glycine
DS07	FF213	F1	 N-(3-(3-borono-5-fluorobenzamido)-2,2-dimethylpropyl)-N-(3-borono-5-fluorobenzoyl)glycine
DS08	FF214	F1	 bis(3-(3-borono-5-fluorobenzamido)propyl)glycine
DS09	FF8	F1	 N-(4-((3-borono-5-fluorobenzamido)methyl)benzyl)-N-(3-borono-5-fluorobenzoyl)glycine

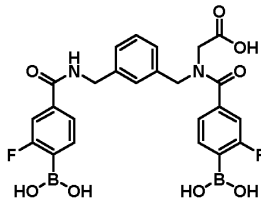
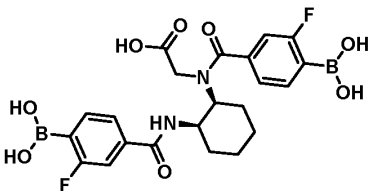
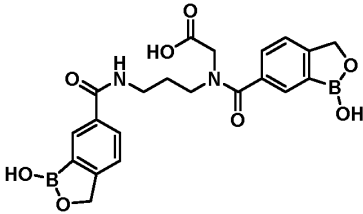
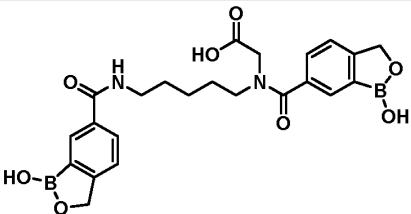
Compd #	FF	F	Structure/IUPAC Name
DS10	FF14	F	 N-(3-((3-borono-5-fluorobenzamido)methyl)benzyl)-N-(3-borono-5-fluorobenzoyl)glycine
DS11	FF1	F1	 N-(2-(3-borono-5-fluorobenzamido)cyclohexyl)-N-(3-borono-5-fluorobenzoyl)glycine
DS12	FF10	F1	 N-(3-(3-borono-4-fluorobenzamido)propyl)-N-(3-borono-4-fluorobenzoyl)glycine
DS13	FF217	F1	 N-(4-((4-(3-borono-4-fluorobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine

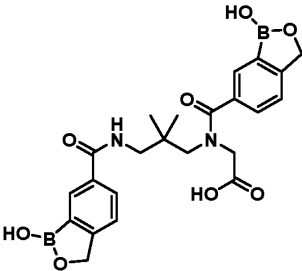
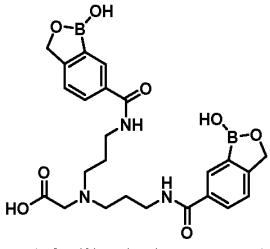
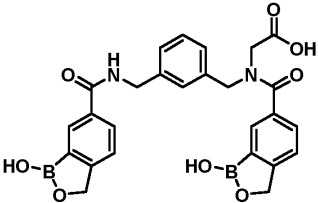
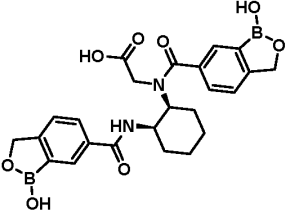
Compd #	FF	F	Structure/IUPAC Name
DS14	FF213	F1	 N-(3-(3-borono-4-fluorobenzamido)-2,2-dimethylpropyl)-N-(3-borono-4-fluorobenzoyl)glycine
DS15	FF8	F1	 N-(4-((3-borono-4-fluorobenzamido)methyl)benzyl)-N-(3-borono-4-fluorobenzoyl)glycine
DS16	FF14	F1	 N-(3-((3-borono-4-fluorobenzamido)methyl)benzyl)-N-(3-borono-4-fluorobenzoyl)glycine
DS17	FF1	F1	 N-((1S,2R)-2-(3-borono-4-fluorobenzamido)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine

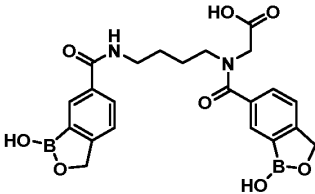
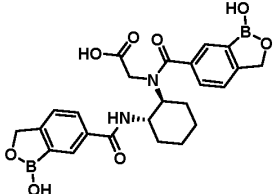
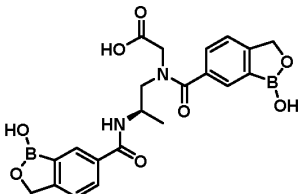
Compd #	FF	F	Structure/IUPAC Name
DS18	FF1	F1	 N-((1S,2S)-2-(3-borono-4-fluorobenzamido)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine
DS19	FF10	F1	 N-(3-(3-borono-5-bromobenzamido)propyl)-N-(3-borono-5-bromobenzoyl)glycine
DS20	FF217	F1	 N-(4-((4-(3-borono-5-bromobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-5-bromobenzoyl)glycine
DS21	FF214	F1	 bis(3-(3-borono-5-bromobenzamido)propyl)glycine

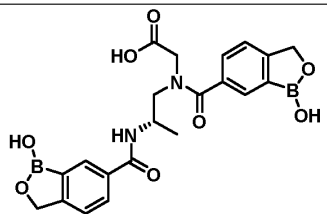
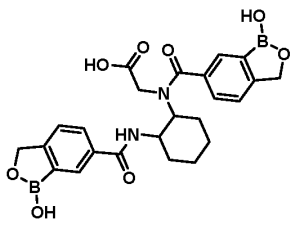
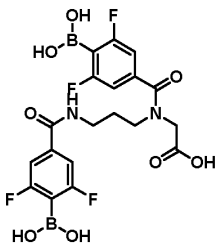
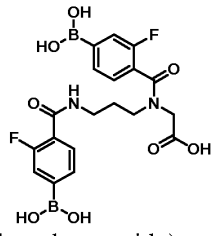
Compd #	FF	F	Structure/IUPAC Name
DS22	FF8	F1	 N-(4-((3-borono-5-bromobenzamido)methyl)benzyl)-N-(3-borono-5-bromobenzoyl)glycine
DS23	FF14	F1	 N-(3-((3-borono-5-bromobenzamido)methyl)benzyl)-N-(3-borono-5-bromobenzoyl)glycine
DS24	FF1	F1	 N-(2-(3-borono-5-bromobenzamido)cyclohexyl)-N-(3-borono-5-bromobenzoyl)glycine
DS25	FF10	F1	 N-(3-(4-borono-3-fluorobenzamido)propyl)-N-(4-borono-3-fluorobenzoyl)glycine

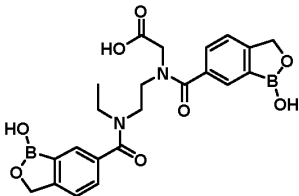
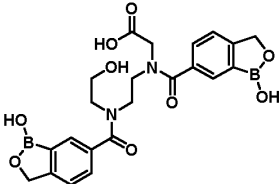
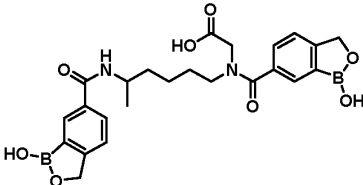
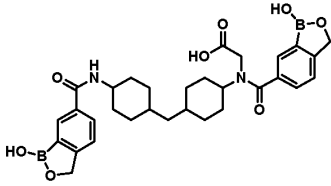
Compd #	FF	F	Structure/IUPAC Name
DS26	FF217	F1	 N-(4-((4-(4-borono-3-fluorobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(4-borono-3-fluorobenzoyl)glycine
DS27	FF213	F1	 N-(3-(4-borono-3-fluorobenzamido)-2,2-dimethylpropyl)-N-(4-borono-3-fluorobenzoyl)glycine
DS28	FF214	F1	 bis(3-(4-borono-3-fluorobenzamido)propyl)glycine
DS29	FF8	F1	 N-(4-((4-borono-3-fluorobenzamido)methyl)benzyl)-N-(4-borono-3-fluorobenzoyl)glycine

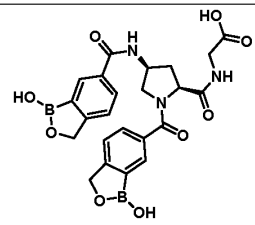
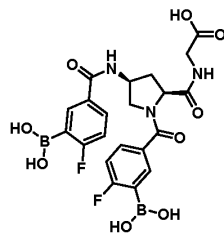
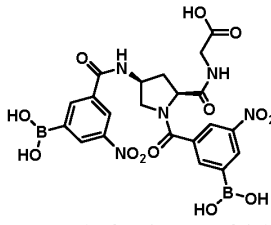
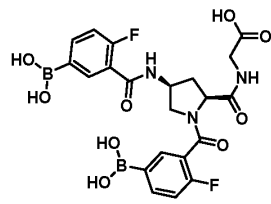
Compd #	FF	F	Structure/IUPAC Name
DS30	FF14	F1	 N-(3-((4-borono-3-fluorobenzamido)methyl)benzyl)-N-(4-borono-3-fluorobenzoyl)glycine
DS31	FF1	F1	 N-((1S,2R)-2-(4-borono-3-fluorobenzamido)cyclohexyl)-N-(4-borono-3-fluorobenzoyl)glycine
DS32	FF10	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine
DS33	FF10	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pentyl)glycine

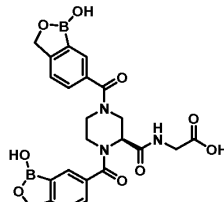
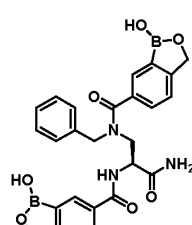
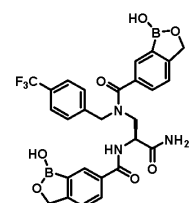
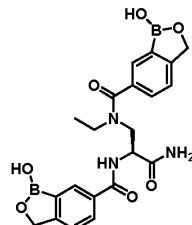
Compd #	FF	F	Structure/IUPAC Name
DS34	FF213	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-2,2-dimethylpropyl)glycine
DS35	FF214	F2	 bis(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine
DS36	FF14	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)benzyl)glycine
DS37	FF1	F2	

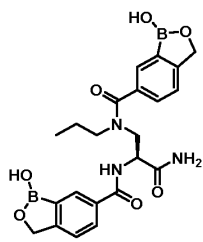
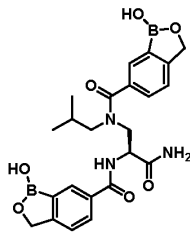
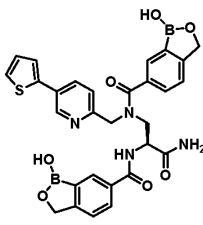
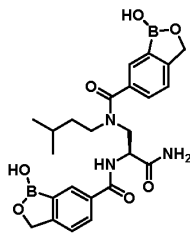
Compd #	FF	F	Structure/IUPAC Name
			N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-((1S,2R)-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine
DS38	FF10	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butyl)glycine
DS39	FF1	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-((1S,2S)-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine
DS40	FF15	F2	 (R)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine
DS41	FF15	F2	

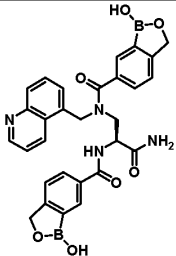
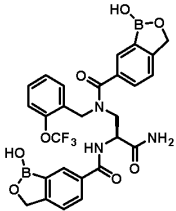
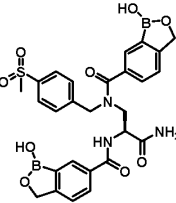
Compd #	FF	F	Structure/IUPAC Name
			 (S)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine
DS42	FF1	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine
DS43	FF10	F1	 N-(3-(4-borono-3,5-difluorobenzamido)propyl)-N-(4-borono-3,5-difluorobenzoyl)glycine
DS44	FF10	F1	 N-(3-(4-borono-2-fluorobenzamido)propyl)-N-(4-borono-2-fluorobenzoyl)glycine

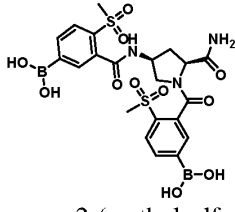
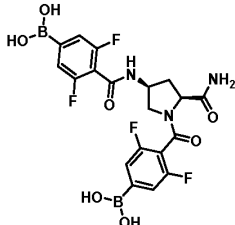
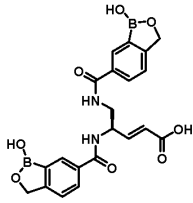
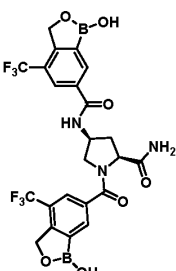
Compd #	FF	F	Structure/IUPAC Name
DS45	FF215	F2	 N-(2-(N-ethyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)ethyl)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)glycine
DS46	FF215	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-N-(2-hydroxyethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)ethyl)glycine
DS47	FF216	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)hexyl)glycine
DS48	FF217	F2	 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(4-((4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)methyl)cyclohexyl)glycine

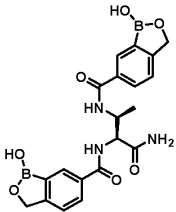
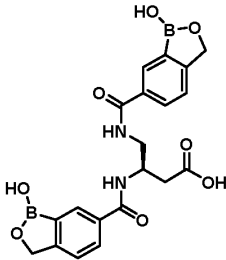
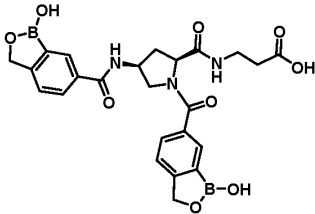
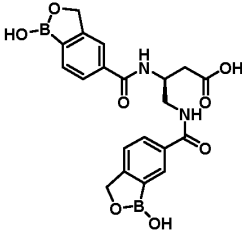
Compd #	FF	F	Structure/IUPAC Name
DS49	FF12	F2	 ((2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine
DS50	FF12	F1	 ((2S,4S)-4-(3-borono-4-fluorobenzamido)-1-(3-borono-4-fluorobenzoyl)pyrrolidine-2-carbonyl)glycine
DS51	FF12	F1	 ((2S,4S)-4-(3-borono-5-nitrobenzamido)-1-(3-borono-5-nitrobenzoyl)pyrrolidine-2-carbonyl)glycine
DS52	FF12	F1	 ((2S,4S)-4-(5-borono-2-fluorobenzamido)-1-(5-borono-2-fluorobenzoyl)pyrrolidine-2-carbonyl)glycine

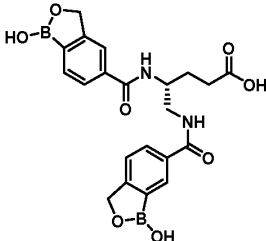
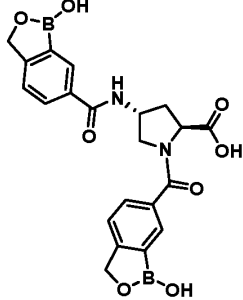
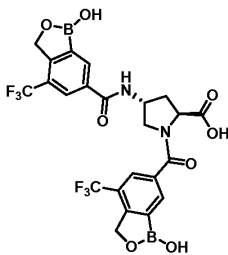
Compd #	FF	F	Structure/IUPAC Name
DS53	FF9	F2	 (S)-(1,4-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)piperazine-2-carbonyl)glycine
DS54	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-benzyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS55	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(4-(trifluoromethyl)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS56	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-ethyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide

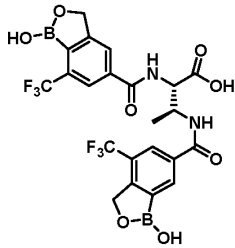
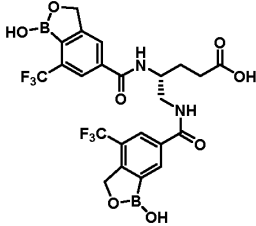
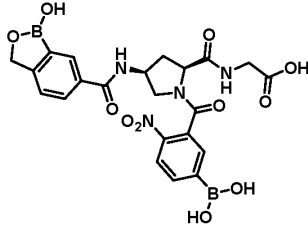
Compd #	FF	F	Structure/IUPAC Name
DS57	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-propyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS58	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-isobutyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS59	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-((5-(thiophen-2-yl)pyridin-2-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS60	FF208	F2	

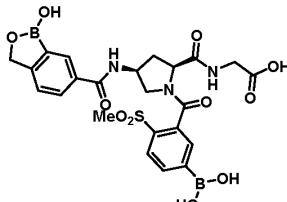
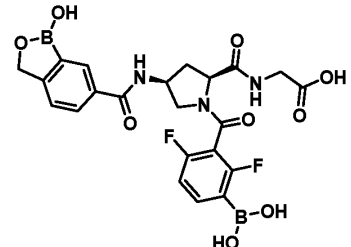
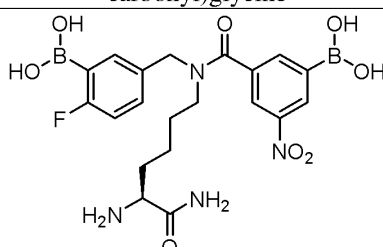
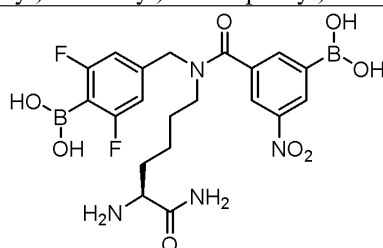
Compd #	FF	F	Structure/IUPAC Name
			(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-isopentyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS61	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(quinolin-5-ylmethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS62	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(2-(trifluoromethoxy)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS63	FF208	F2	 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(4-(methylsulfonyl)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide

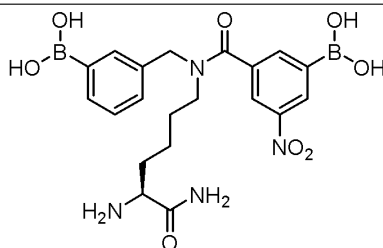
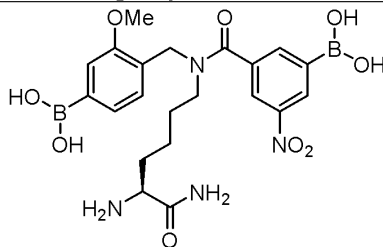
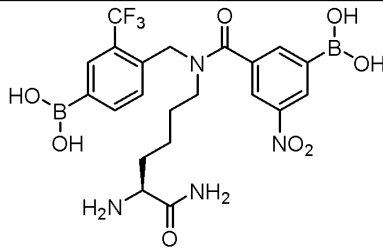
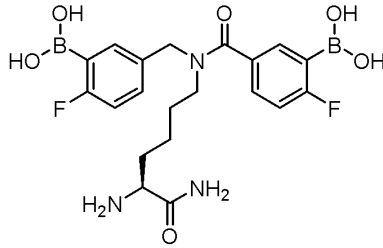
Compd #	FF	F	Structure/IUPAC Name
DS64	FF12	F1	 (3-((2S,4S)-4-(5-borono-2-(methylsulfonyl)benzamido)-2-carbamoylpyrrolidine-1-carbonyl)-4-(methylsulfonyl)phenyl)boronic acid
DS65	FF12	F1	 (4-(((3S,5S)-1-(4-borono-2,6-difluorobenzoyl)-5-carbamoylpyrrolidin-3-yl)carbamoyl)-3,5-difluorophenyl)boronic acid
DS66	FF220	F2	 (R,E)-4,5-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pent-2-enoic acid
DS67	FF12	F2	 (2S,4S)-1-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-2-carbamoylpyrrolidine-1-carbonyl

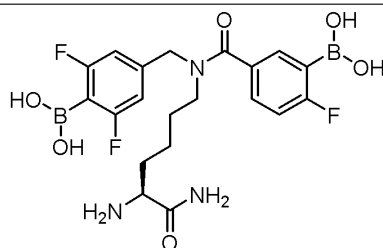
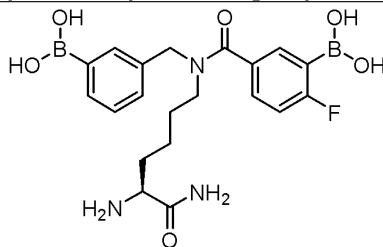
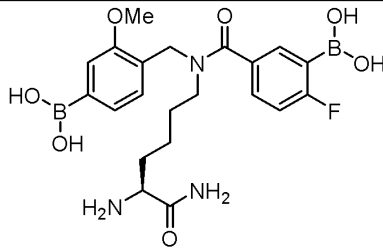
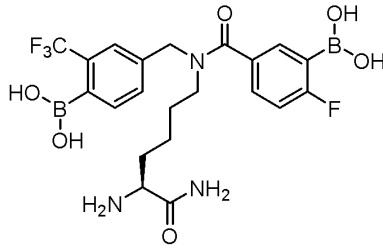
Compd #	FF	F	Structure/IUPAC Name
			(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxamide
DS68	FF116	F2	 N,N'-((2S,3S)-1-amino-1-oxobutane-2,3-diyl)bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide)
DS69	FF115	F2	 (R)-3,4-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid
DS70	FF12	F2	 3-((2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxamido)propanoic acid
DS71	FF115	F2	

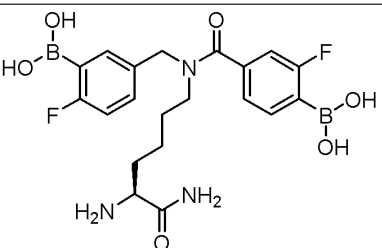
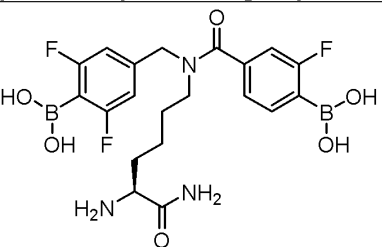
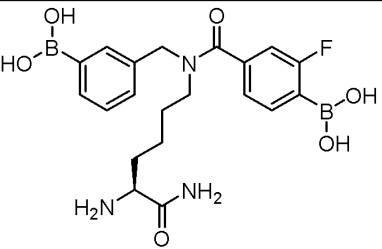
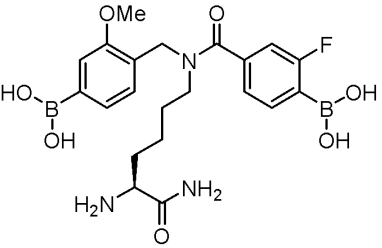
Compd #	FF	F	Structure/IUPAC Name
			(S)-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid
DS72	FF114	F2	 (R)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pentanoic acid
DS73	FF12	F2	 (2S,4R)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid
DS74	FF12	F2	 (2S,4R)-1-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid

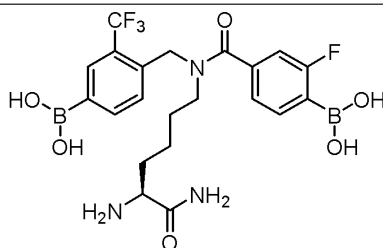
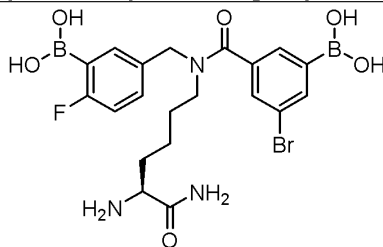
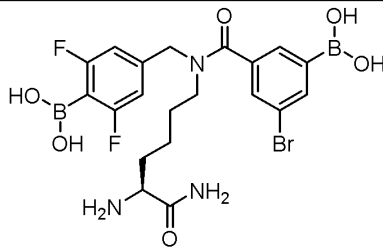
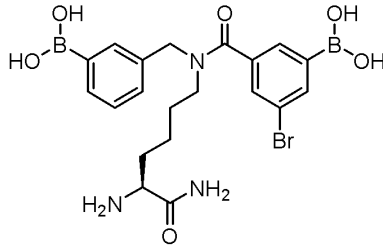
Compd #	FF	F	Structure/IUPAC Name
DS75	FF116	F2	 (2S,3S)-3-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-2-(1-hydroxy-7-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)butanoic acid
DS76	FF114	F2	 (R)-5-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-(1-hydroxy-7-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)pentanoic acid
DS77	FF12	F1 and F2	 ((2S,4S)-1-(5-borono-2-nitrobenzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine

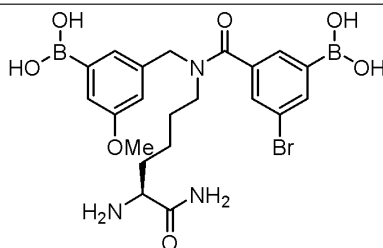
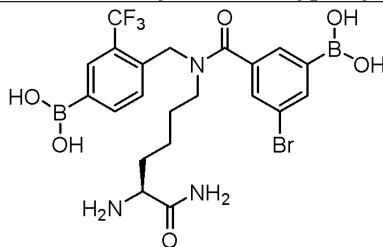
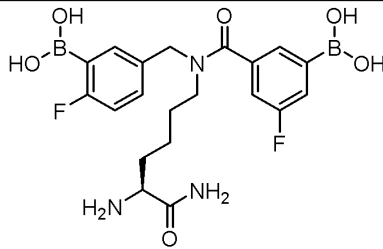
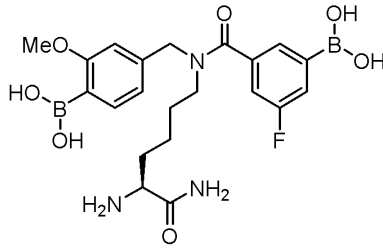
Compd #	FF	F	Structure/IUPAC Name
DS78	FF12	F1 and F2	 ((2S,4S)-1-(5-borono-2-(methylsulfonyl)benzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine
DS79	FF12	F1 and F2	 ((2S,4S)-1-(3-borono-2,6-difluorobenzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine
DS80	FF221	F1	 (S)-(3-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid
DS81	FF221	F1	 (S)-(3-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid

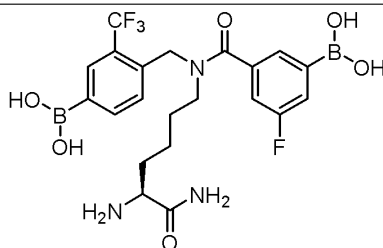
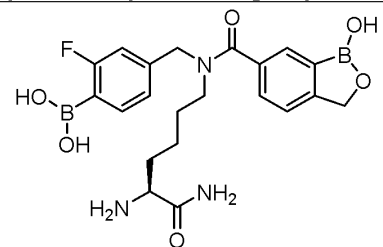
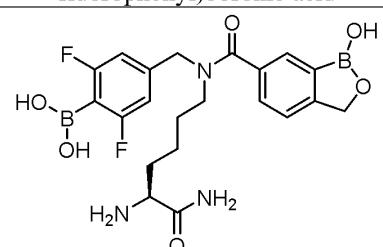
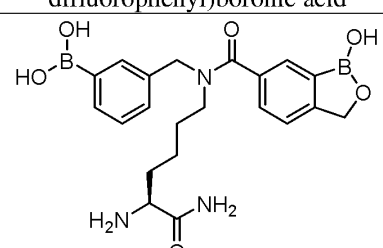
Compd #	FF	F	Structure/IUPAC Name
DS82	FF221	F1	 (S)-3-((3-boronobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenylboronic acid
DS83	FF221	F1	 (S)-3-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenylboronic acid
DS84	FF221	F1	 (S)-3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenylboronic acid
DS85	FF221	F1	 (S)-5-((3-borono-N-(5,6-diamino-6-oxohexyl)-4-fluorobenzamido)methyl)-2-fluorophenylboronic acid

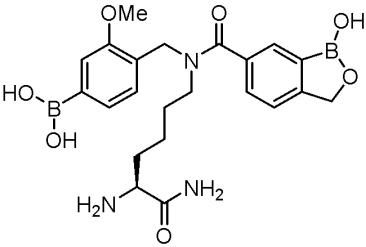
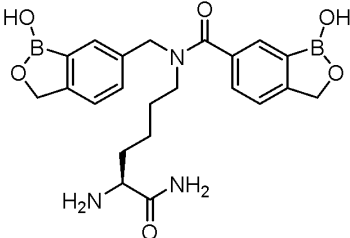
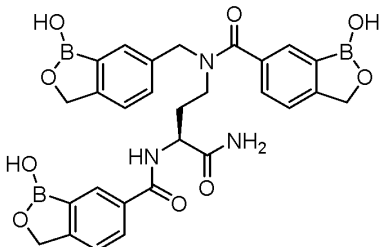
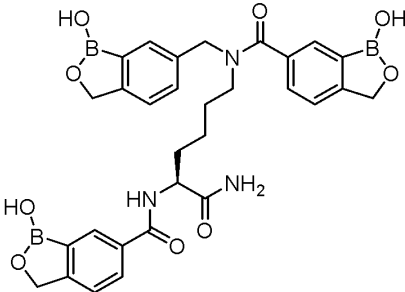
Compd #	FF	F	Structure/IUPAC Name
DS86	FF221	F1	 (S)-((5-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid
DS87	FF221	F1	 (S)-((3-((3-borono-N-(5,6-diamino-6-oxohexyl)-4-fluorobenzamido)methyl)phenyl)boronic acid
DS88	FF221	F1	 (S)-((5-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid
DS89	FF221	F1	 (S)-((5-((4-borono-3-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid

Compd #	FF	F	Structure/IUPAC Name
DS90	FF221	F1	 (S)-4-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenylboronic acid
DS91	FF221	F1	 (S)-4-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenylboronic acid
DS92	FF221	F1	 (S)-4-((3-boronobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenylboronic acid
DS93	FF221	F1	 (S)-4-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenylboronic acid

Compd #	FF	F	Structure/IUPAC Name
DS94	FF221	F1	 (S)-4-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenylboronic acid
DS95	FF221	F1	 (S)-5-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)-2-fluorophenylboronic acid
DS96	FF221	F1	 (S)-3-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-bromophenylboronic acid
DS97	FF221	F1	 (S)-3-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)phenylboronic acid

Compd #	FF	F	Structure/IUPAC Name
DS98	FF221	F1	 (S)-3-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)-5-methoxyphenylboronic acid
DS99	FF221	F1	 (S)-3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-bromophenylboronic acid
DS100	FF221	F1	 (S)-3-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenylboronic acid
DS101	FF221	F1	 (S)-3-((4-borono-3-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenylboronic acid

Compd #	FF	F	Structure/IUPAC Name
DS102	FF221	F1	 (S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic acid
DS103	FF221	F1 and F2	 (S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-2-fluorophenyl)boronic acid
DS104	FF221	F1 and F2	 (S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-2,6-difluorophenyl)boronic acid
DS105	FF221	F1 and F2	 (S)-(3-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)phenyl)boronic acid

Compd #	FF	F	Structure/IUPAC Name
DS106	FF221	F1 and F2	 (S)-4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-3-methoxyphenylboronic acid
DS107	FF221	F2	 (S)-N-(5,6-diamino-6-oxohexyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS108	FF221	F2	 (S)-N-(4-amino-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-oxobutyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide
DS109	FF221	F2	 (S)-N-(4-amino-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-oxobutyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide

Compd #	FF	F	Structure/IUPAC Name
			(S)-N-(6-amino-5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-6-oxohexyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide

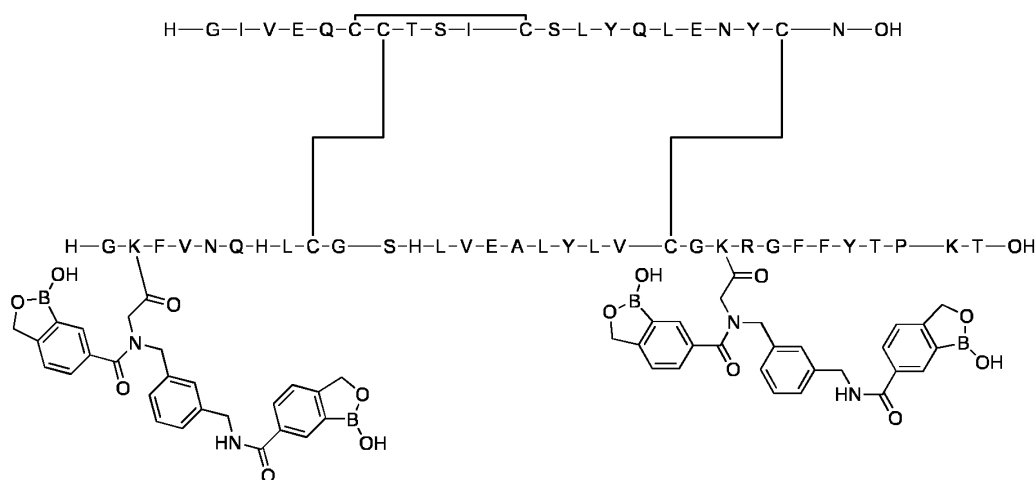
5

Synthesis of Compounds of Formula I

An illustrative synthesis protocol is provided that can be used to synthesize any of the examples described.

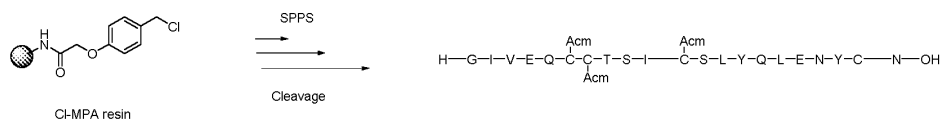
The lines connecting cysteine residues are disulfide bonds. For the sake of clarity, the H- at the N-terminus of the A- and B-chain of insulin is not histidine, it is the hydrogen of the N-terminus. The -OH shown at the C-terminal end of the A- and B-chain is the C-terminus of the respective chain.

Synthesis of Modified Insulins



Example 870

15 **A- chain synthesis:**



Sequence: GIVKQC(Acm)C(Acm)TSIC(Acm)SLYQLENYCN

5

Synthesis of A-chain and modified A-chain of Example 870 was accomplished using conventional solid-phase peptide synthesis (SPPS)

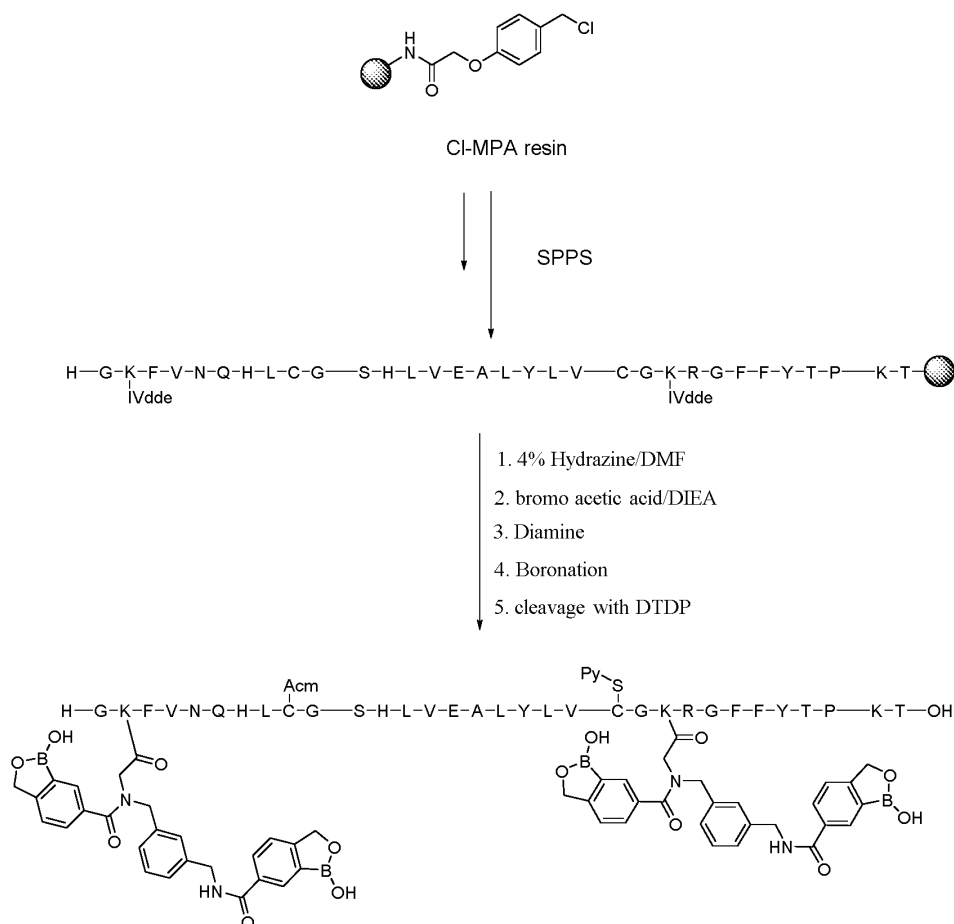
MPA resin (0.22 mmol/eq) was swelled in a mixture of DMF:DCM (50:50, v:v). A solution of potassium iodide with DIPEA (1 M) in DMF was added to the reaction vessel along with Fmoc-Asn(Trt)-OH (0.2 M). The reaction vessel was heated to 75°C. Each amino acid coupling step involved i) deprotection with 20% piperidine in DMF at 90°C; ii) washing with DMF; iii) activation and coupling of Fmoc protected amino acids with 0.5 M N,N'-diisopropylcarbodiimide (DIC, 1mL), 0.5 M Oxyma, and 0.2 M Fmoc-amino acid in DMF at 90°C; iv) washing with DMF.

Global deprotection and isolation.

Crude peptide was globally deprotected in TFA:TIPS:H₂O (95:2.5:2.5) and gently agitated for 2h. Crude solution was filtered and peptide was precipitated in cold ether, centrifuged and washed with additional cold ether. Supernatant was decanted and the crude peptide was dried under a gentle stream of nitrogen gas. Crude peptide was dissolved in 20% ACN in water and fractionated by RP-HPLC on a C18 column.

B-chain synthesis:

Sequence: GK FVNQHLC(Acm)GSHLVEALYLVC(DTDP)GERGFFYTPK



5

Synthesis of modified B-chain insulins using solid-phase peptide synthesis (SPPS).

MPA resin (0.22 mmol/eq) was swelled in a mixture of DMF:DCM (50:50, v:v). A solution of potassium iodide (125 mM) with DIPEA (1 M) in DMF was added to the reaction vessel along with Fmoc-Lys(ivDde)-OH (0.2 M). The reaction vessel was heated to 75°C. Each amino acid coupling step involved i) deprotection with 20% piperidine in DMF at 90°C; ii) washing with DMF; iii) activation and coupling of Fmoc protected amino acids with 0.5 M N,N'-diisopropylcarbodiimide (DIC), 0.5 M Oxyma, and 0.2 M Fmoc-amino acid in DMF at 90°C; iv) washing with DMF. Fmoc-Arg(Pbf)-OH was coupled twice using the methods described above.

10

15

5 **Deprotection IVdde and add diboronates to B chain.**

The ivDde protecting group on the lysine residue was removed with 4% hydrazine in DMF, then washed with DMF. A solution of bromo acetic acid (0.2 M, 2 mL) in DMF with DIC (0.5 M, 2 mL) was added immediately and gently mixed for 4 hr. The resin was washed with DMF (3x5 mL). A solution of 1,3-phenylenedimethanamine (1M) in DMF (5mL) was
10 added to the resin and heated at 50 C for 10 minutes. The resin was washed with DMF (3x 5mL) and a solution of 1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxylic acid (0.2 M, 5 mL) in DMF with 1 M N,N'-diisopropylcarbodiimide (DIC, 1M, 1mL), Oxyma (0.5 M, 2 mL) in DMF and heated at 50 C for 30 min.

15 **Cleavage and addition of DTDP to B-chain.**

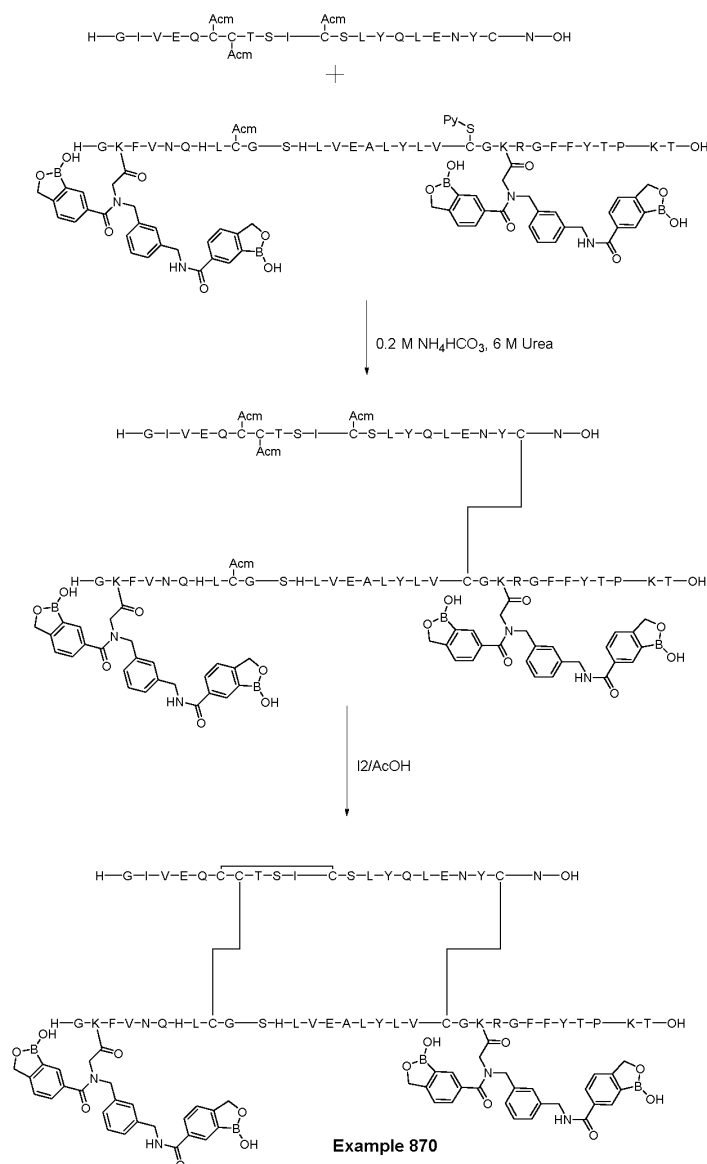
Crude peptide was globally deprotected with 2,2'-Dithiopyridine (DTDP) in TFA:TIPS:H₂O (95:2.5:2.5) and gently agitated. Crude solution was filtered and peptide was precipitated in cold ether, centrifuged, decanted, washed with additional cold ether, and centrifuged again. Supernatant was decanted and the crude peptide was dried under a gentle
20 stream of nitrogen gas. Crude peptide was dissolved in 20% ACN in water and fractionated by RP-HPLC on a C18 column.

Combination of A and B chains of insulin and modified insulins.

A chain of insulin was combined with B chain in 0.2 M NH₄HCO₃ with 6M urea and at pH
25 8. Mixture was gently agitated, diluted with water and fractionated by RP-HPLC on a C18 column.

Deprotection of Acm protecting groups and final oxidation.

The combined intermediate was dissolved in glacial acetic acid and water and vortexed
30 vigorously. A solution of iodine in glacial acetic acid (20 equiv) was added to the reaction mixture and gently agitated. A solution of ascorbic acid (5mM) was added directly to the reaction mixture. The mixture was diluted in 20% ACN in water and fractionated by RP-HPLC on a Higgins C18 column to give Example 870.



5

Similar procedures can be followed for the synthesis of Examples 1 to 869 and 871 to 876.

B. Testing of Compounds for Activity in Biological Assays

Exemplary compounds DS01-DS79 of the present disclosure were tested using an alizarin red S (ARS) displacement assay.

5 **Procedure for determination of the glucose, fructose, and lactate binding (Kd) using ARS displacement assay**

The association constant for the binding event between Alizarin Red S (ARS) and the exemplary compounds tested was determined using standard methods in the art. Triplicate titrations of 10^{-5} M ARS in 0.1 M phosphate buffer, pH 7.4, were performed in a 96-well plate against serial dilutions of example compounds, ranging in concentration from 0 – 0.1M at 25°C. The example compound-ARS solution was incubated for 5-45 minutes at 25 °C, and absorbance intensity was measured using excitation wavelength 468 nm and emission wavelength 585 nm. Changes in intensity were plotted against the concentration of the example compound, and the intensity data was fitted to yield an association constant for ARS binding.

15 The association constant for the binding between a target sugar compound (e.g., glucose) and the tested aromatic boron-containing groups was determined via the displacement of ARS bound to the example compounds. Triplicate wells of 10^{-5} M ARS and 0.1 M example compounds in 0.1 M phosphate buffer, pH 7.4, were titrated in a 96-well plate against serial dilutions of the target sugar compound, ranging in concentration from 0 – 2.0 M at 25 °C. The boron-ARS-carbohydrate solution was incubated for 30-60 minutes at 25 °C and the intensity of each well was measured in a plate reader at excitation wavelength 468 nm and emission wavelength 585 nm.

Changes in intensity were plotted against concentration of the target sugar compound, and the data was fitted to a one -site competition equation:

25
$$y = \min(y) + (\max(y) - \min(y)) / (1 + 10^{x - \log EC50})$$

to yield an association constant for the boron compound-target sugar compound binding event.

The binding constants of DS01-DS109 to glucose, fructose, and/or lactate were tested and were calculated, except DS76 was not tested for glucose binding, DS57 and DS75 were not tested for fructose or lactate binding. The tested compounds had Kd values ranging from ~0.8 mM to ~486 mM for glucose, ~0.9 mM to ~52 mM for fructose, and ~24 to ~425 mM for lactate.

***In Vitro* Demonstration of Activity for Compounds of Formula I**

CHO cells constitutively expressing Human Insulin Receptor Isoform β were plated in a 96-well tissue culture microplate at 35,000 cells/well and grown overnight in RPMI media

5 supplemented with Glutamine and 10% Fetal Bovine Serum (growth media). The next morning, growth media was replaced with fresh growth media.

A separate microplate was prepared with a stepwise serial dilution of glucose-responsive insulin in DMEM media, without glucose, without phenol red, with 4% w/v serum albumin; wells of serially diluted compounds of Formula I were prepared in triplicate with an appropriate
10 “high” and “low” concentration of glucose to determine change in potency of compounds of Formula I at various potential blood glucose levels.

Growth media on cells was then replaced with DMEM media, no glucose, no phenol red for 5 minutes. The media was aspirated and replaced with the contents of the prepared plate (spiking media) for 10 minutes. The spiking media was aspirated and the cells were fixed with 10%
15 neutral buffered formalin for 10 minutes. The neutral buffered formalin was aspirated, and the microplate was stringently washed with PBS, pH 7.4. The microplate was then blocked with PBS, pH 7.4 supplemented with 10% v/v Fetal Bovine Serum and 0.1% Triton X-100 for 30 minutes. The plate was then stained at 4°C overnight with 5% FBS in PBS + 1:680 v:v of Rb α -phospho-Y1150/Y1151 IR antibody (*Cell Signaling Technologies #3024*). After stringent
20 washes with PBS, pH 7.4, the microplate was incubated at 37°C in 5% FBS in PBS + 1:1000 of 2° Ab, HRP α -Rabbit (*Cell Signaling Technologies, #7074*) for 100 minutes. The plate was stringently washed with PBS, pH 7.4, and colorimetric readout was developed for 15 minutes at 37°C using TMB substrate. Color development was stopped with the addition of 0.1 M hydrochloric acid and absorbance measured at 450 nm. Triplicate absorbance values were
25 plotted in GraphPad Prism and analyzed using a four-parameter logistic regression to generate dose-response curves, and the EC50 of the dose-response curves were compared to assess fold activation of the exemplary compounds of Formula I from low to high glucose concentration.

Examples 315, 318, 320, 565, 590, 606, 611, and 803 - 880 had an IR Phosphorylation (fold change) ranging from ≥ 1.2 to 45.

30 ***In Vivo* Demonstration of Activity for Compounds of Formula I**

Diabetes was induced in 12-week-old B6 mice using streptozocin (STZ) treatment. Three weeks after the final STZ treatment, blood glucose of the mice was sampled from lateral tail vein to confirm diabetes. Mice with blood glucose of $>200\text{mg/dl}$ were deemed to be diabetic and fasted for 1-6 hours prior to injection. Human insulin and a compound of Formula I in

5 sterile phosphate buffered saline, pH 7.4 were injected either subcutaneously via neck scruff or intraperitoneally, depending on the experiment. Blood glucose was sampled with a glucometer via lateral tail vein at 15-minute intervals. After one hour of initial stabilization of blood glucose levels post-insulin injection, mice were challenged with an intraperitoneal injection of glucose (e.g., 2 g/kg, 4 g/kg, or 6 g/kg; the actual dose depends on the example and experiment)
10 in sterile phosphate buffered saline. The exemplary compound activated to lower blood glucose upon the introduction of a glucose bolus, while human insulin did not activate in a glucose-dependent manner.

The above experiment showed *in vivo* preferential activity response of an exemplary compound of the disclosure to glucose.

15 Streptozotocin-treated mice (55 mg/kg, 5 days) undergo surgical catheterization of a carotid artery and jugular vein for blood sampling and infusions. After a recovery period of 3-4 days, mice are placed in an experimental chamber, connected to sampling/infusion lines, and briefly fasted. Somatostatin (5 mg/kg/min) is continuously infused throughout the study. At time 0 min, biosynthetic human insulin (BHI) or a compound of Formula I are infused at 4 mU/kg/min
20 and glucose is infused at variable rates to achieve steady state ("clamped") at pre-determined glycemic levels. Blood glucose (BG) is clamped in windows of stepwise increasing blood glucose concentration. Steady-state Glucose Infusion Rate (GIR) is measured for each step increase in clamped blood glucose and plotted to assess the increase in GIR as a function of increasing blood glucose. As BG increases, compounds of Formula I require greater GIR to
25 maintain clamped BG than does BHI, demonstrating a glucose-responsive increase in compound glucose-lowering action with increased BG concentrations.

It is also observed in cell-based experiments on compounds containing formulae FF50-FF62, FF116, and FF121-FF134 that sensors with geminal alkyl substituent on the same carbon as the nitrogen conjugated to the boroxole or boronates provided between 5-56% higher glucose
30 responsiveness in the range of 3-20 mM glucose in comparison to variants that do not have the geminal alkyl substituents. For example, when a Z1C represented by one of formulae FF50 – FF62, FF116, and FF121-134 is conjugated to lysine residues in insulin wherein the boronates (B1,B2) in formulae FF50 – FF62, FF116, and FF121-134 are represented by F2, the resulting insulin is observed to be between 11-56% more responsive to changes in glucose levels between
35 3-20 mM glucose than if instead of one of formulae FF50 – FF62, FF116, and FF121-134, one

5 uses 2,3-diaminopropionic acid. This data shows that the presence of the geminal alkyl substituent on the same carbon as the nitrogen conjugated to the boroxole or boronates improves glucose responsiveness of the resulting insulin conjugate, and in tested variations in the 3-20mM glucose range. Without wishing to be bound by theory, it is believed that this general principle extends to other formulae FF50-FF62, FF116, and FF121-FF134 providing a
10 framework for enhancement of glucose responsiveness by at least 5%, at least 10%, at least 20%, or at least 40% in the 3-20mM or 2-50mM glucose ranges. Without wishing to be bound by theory, it is believed based on observations of glucose responsiveness trends, that the presence of the carbonyl group adjacent to- or within less than two carbon centers away from the amine groups in FF formulae (to which aromatic boron-containing groups are conjugated)
15 enhances glucose responsiveness through impacting ability to turn off activity of drug substance through plasma protein interactions such as with albumin and that this is independent of glucose affinity such that glucose affinity is not impacted by the position of this carbonyl group. Without wishing to be bound by theory, it is believed that the pharmacokinetics of the molecules and potential albumin or blood proteins binding is impacted by the position of this
20 carbonyl group, and thereby enhances overall glucose responsiveness whilst the absolute glucose affinity is maintained or nearly identical. Therefore, in certain embodiments of the present invention, the carbonyl group (as part of an acid, amid or linkage to X in FF formulae) is placed within less than three-, or within less than two-carbon center away from one of the two amines to which the boron-containing compounds are conjugated. In certain embodiments,
25 the placement of amines within two carbon centers from each other enables the spatial and geometric constraining of the aromatic boron containing groups to enhance glucose binding and selectively, and furthermore the presence of a carbonyl group (for example, as part of an amide linkage) which is within less than two carbon centers, from one of the two amines (to which aromatic boron containing groups are attached) ensures differential albumin binding in a
30 manner that results in the compound exhibiting glucose responsiveness in the blood and in the body. In some embodiments, the combination of geometrical constraining of the two amines to which the aromatic boron containing groups are conjugated, as well as the presence of the carbonyl within one to two carbon centers from one of the amines provides the necessary requirements for glucose responsiveness in physiological blood and plasma glucose levels.
35 Experiments on cell-based assays of insulins with lysines conjugated with one of formulae FF50 – FF62, FF116, and FF121-134 demonstrated that the enhanced glucose responsiveness

5 of the insulins is increased when one or more lysines are modified as described by Formula I using one of formulae FF50 – FF62, FF116, and FF121-134, and wherein the lysines are present in insulin (as insertions or mutations) or in a polypeptide that is appended to the N- or C- terminus of the B-chain of insulin or the C-terminus of the A-chain of insulin, and wherein there are additional lysine residues within the insulin sequence that are similarly modified. The results are further corroborated by testing of the compounds of Formula I in STZ diabetic mouse models wherein the activity of the insulin is measured through bolus injections of the compounds of Formula I followed by glucose challenges and measurements of blood glucose, or through glucose clamp assays in which activity of the insulins is measured as a function of blood glucose levels. The results further showed that exemplary compounds of Formula I disclosed herein function in the body and are responsive to physiological changes in blood glucose and provide dynamic insulin action in the body in response to changes in blood glucose levels.

In certain embodiments a sequence is appended to the N-terminus and/or C-terminus, and/or inserted into the sequence of the A-chain of insulin, wherein the A-chain of insulin comprises one of the following sequences, optionally with up to four additional deletions and/or mutations:

GIVEQCCTSICSLYQLENYCN (SEQ ID NO:1),
GIVKQCCTSICSLYQLENYCN (SEQ ID NO:3),
GIVEQCCHSICSLYQLENYCN (SEQ ID NO:4),
25 GIVEQCCASICSLYQLENYCN (SEQ ID NO:5),
GIVEQCCTRICSLYQLENYCN (SEQ ID NO:6),
GIVEQCCTKICSLYQLENYCN (SEQ ID NO:7),
GIVEQCCTSICSEYQENYCN (SEQ ID NO:8),
GIVKQCCTSICSLYQLENYCG (SEQ ID NO:9),
30 GIVEQCCHSICSLYQLENYCG (SEQ ID NO:10),
GIVEQCCASICSLYQLENYCG (SEQ ID NO:11),
GIVEQCCTRICSLYQLENYCG (SEQ ID NO:12),
GIVEQCCTKICSLYQLENYCG (SEQ ID NO:13),
GIVEQCCTSICSEYQENYCG (SEQ ID NO:14),
35 GIVEQCCTSICSEYQENYC (SEQ ID NO:15),
GIVEQCCTSICSLYQLENYCNK (SEQ ID NO:16),

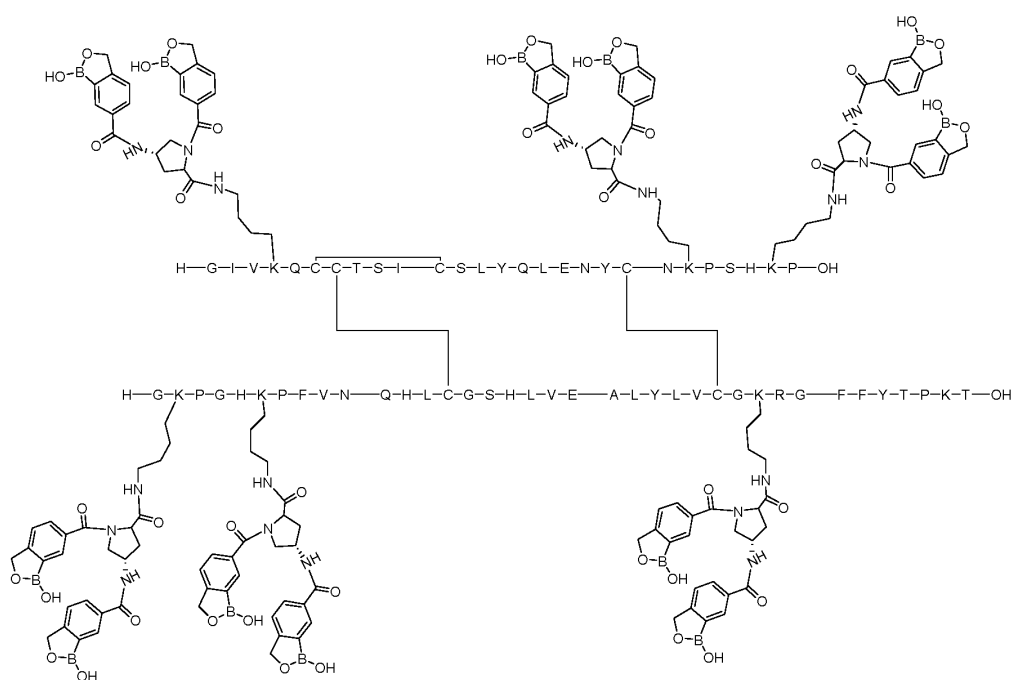
- 5 KPGIVEQCCTSICSLYQLENYCN (SEQ ID NO:17),
KPIVEQCCTSICSLYQLENYCN (SEQ ID NO:18),
KPVEQCCTSICSLYQLENYCN (SEQ ID NO:19),
KPGVEQCCTSICSLYQLENYCN (SEQ ID NO:20),
GEKPVEQCCTSICSLYQLENYCN (SEQ ID NO:21),
10 KPGEKPVEQCCTSICSLYQLENYCN (SEQ ID NO:22),
KPVEQCCTSICSLYQLENYCNK (SEQ ID NO:23),
KPVEQCCTSICSLYQLENYCNEKP (SEQ ID NO:24),
GIVEQCCTSICSLYQLENYCGK (SEQ ID NO:25),
KPGIVEQCCTSICSLYQLENYCG (SEQ ID NO:26),
15 KPIVEQCCTSICSLYQLENYCG (SEQ ID NO:27),
KPVEQCCTSICSLYQLENYCG (SEQ ID NO:28),
KPGVEQCCTSICSLYQLENYCG (SEQ ID NO:29),
GEKPVEQCCTSICSLYQLENYCG (SEQ ID NO:30),
KPGEKPVEQCCTSICSLYQLENYCG (SEQ ID NO:31),
20 KPVEQCCTSICSLYQLENYCGK (SEQ ID NO:32),
KPVEQCCTSICSLYQLENYCGEKP (SEQ ID NO:33),

and/or a sequence is appended to the N-terminus and/or C-terminus, and/or inserted into the sequence of the B-chain of insulin, wherein the B-chain of insulin comprises one of following sequences, and optionally with up to four additional deletions and/or mutations:

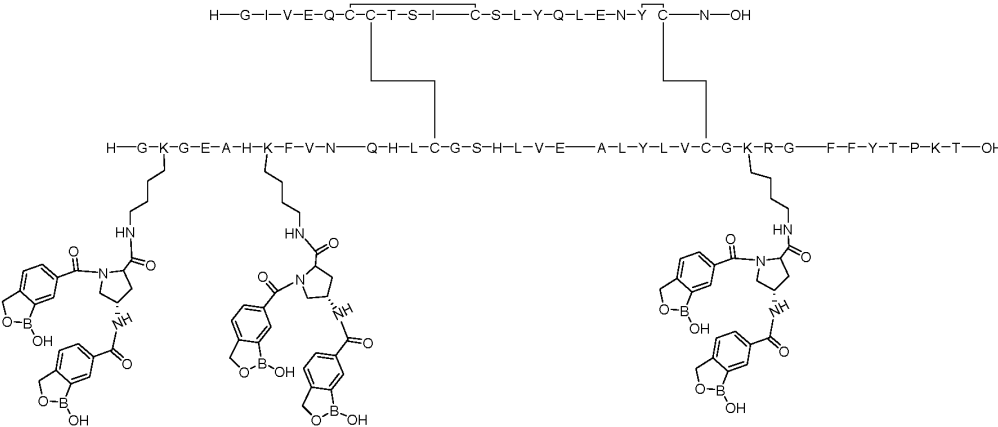
- 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:2),
FVNQHLCGSHLVEALYLVCGERGFFYTP (SEQ ID NO:34),
FVNQHLCGSHLVEALYLVCGKRGFFYTP (SEQ ID NO:35),
FVNQHLCGSHLVEALYLVCGKRGFFYTPRT (SEQ ID NO:36),
FVNQHLCGSHLVEALYLVCGKRGFFYT (SEQ ID NO:37),
30 VNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:38),
NQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:39),
QHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:40),
PFVNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:41),
PFVNQHLCGSHLVEALYLVCGKRGFFYTPRT (SEQ ID NO:42),
35 PFVNQHLCGSHLVEALYLVCGKEGFFYTPRT (SEQ ID NO:43),
PFVNQHLCGSHLVEALYLVCGKRGFFYTPR (SEQ ID NO:44),

5 PFVNQHLCGSHLVEALYLVCGRGFFYTRPT (SEQ ID NO:45),
PFVNQHLCGSHLVEALYLVCGRGFFYTRP (SEQ ID NO:46),
PFVNQHLCGSHLVEALYLVCGKNGFFYTPRT (SEQ ID NO:47),
PFVNQHLCGSHLVEALYLVCGKNGFFYTPRT (SEQ ID NO:48),
PFVNQHLCGSHLVEALYLVCGKNGFFYTPR (SEQ ID NO:49),
10 PFVNQHLCGSHLVEALYLVCGKNGFFYTRPT (SEQ ID NO:50),
PFVNQHLCGSHLVEALYLVCGKNGFFYTRP (SEQ ID NO:51),
PFVNQHLCGSHLVEALYLVCGKEGFFYTPRT (SEQ ID NO:52),
PFVNQHLCGSHLVEALYLVCGKEGFFYTPRT (SEQ ID NO:53),
PFVNQHLCGSHLVEALYLVCGKEGFFYTPR (SEQ ID NO:54),
15 PFVNQHLCGSHLVEALYLVCGKEGFFYTRPT (SEQ ID NO:55),
PFVNQHLCGSHLVEALYLVCGKEGFFYTRP (SEQ ID NO:56),
PFVNQHLCGSHLVEALYLVCGRGFFYTPR (SEQ ID NO:57),
PVNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:58),
PVNQHLCGSHLVEALYLVCGRGFFYTPRT (SEQ ID NO:59),
20 PVNQHLCGSHLVEALYLVCGRGFFYTPR (SEQ ID NO:60),
PNQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:61),
PNQHLCGSHLVEALYLVCGRGFFYTPRT (SEQ ID NO:62),
PNQHLCGSHLVEALYLVCGRGFFYTPR (SEQ ID NO:63),
PQHLCGSHLVEALYLVCGERGFFYTPKT (SEQ ID NO:64),
25 PQHLCGSHLVEALYLVCGRGFFYTPRT (SEQ ID NO:65),
PQHLCGSHLVEALYLVCGRGFFYTPR (SEQ ID NO:66),
PFVNQHLCGSHLVEALYLVCGERGFFYTKPT (SEQ ID NO:67),
PFVNQHLCGSHLVEALYLVCGERGFFYTKP (SEQ ID NO:68),
PVNQHLCGSHLVEALYLVCGERGFFYTKPT (SEQ ID NO:69),
30 PVNQHLCGSHLVEALYLVCGERGFFYTKP (SEQ ID NO:70),
PNQHLCGSHLVEALYLVCGERGFFYTKPT (SEQ ID NO:71),
PNQHLCGSHLVEALYLVCGERGFFYTKP (SEQ ID NO:72),
PQHLCGSHLVEALYLVCGERGFFYTKPT (SEQ ID NO:73),
FVNQHLCGSHLVEALYLVCGERGFFYTPK (SEQ ID NO:74),
35 FVNQHLCGSHLVEALYLVCGRGFFYTPKT (SEQ ID NO:24047), and
FVNQHLCGSHLVEALYLVCGRGFFYTPR (SEQ ID NO:24048).

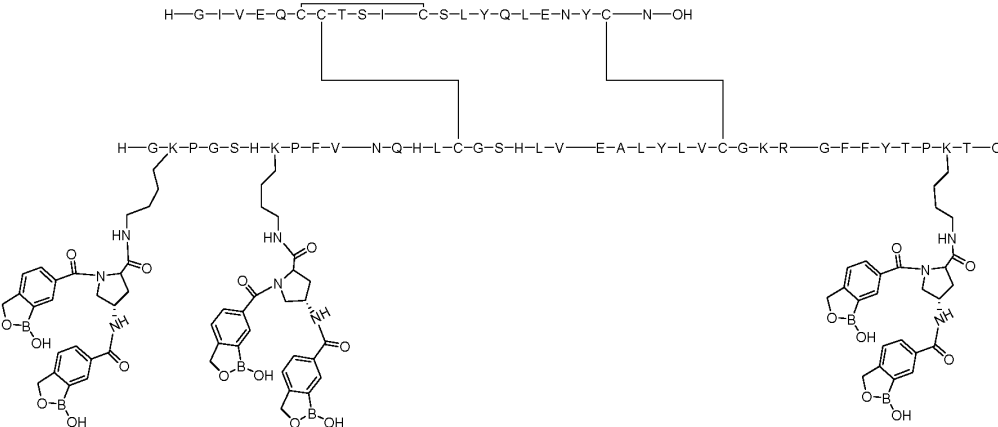
5 Example 1:



Example 2:

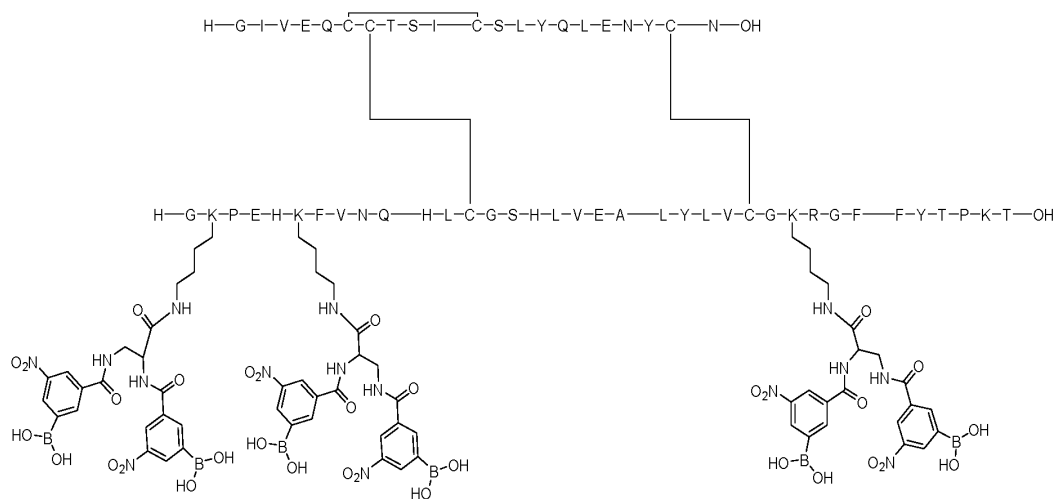


Example 3:

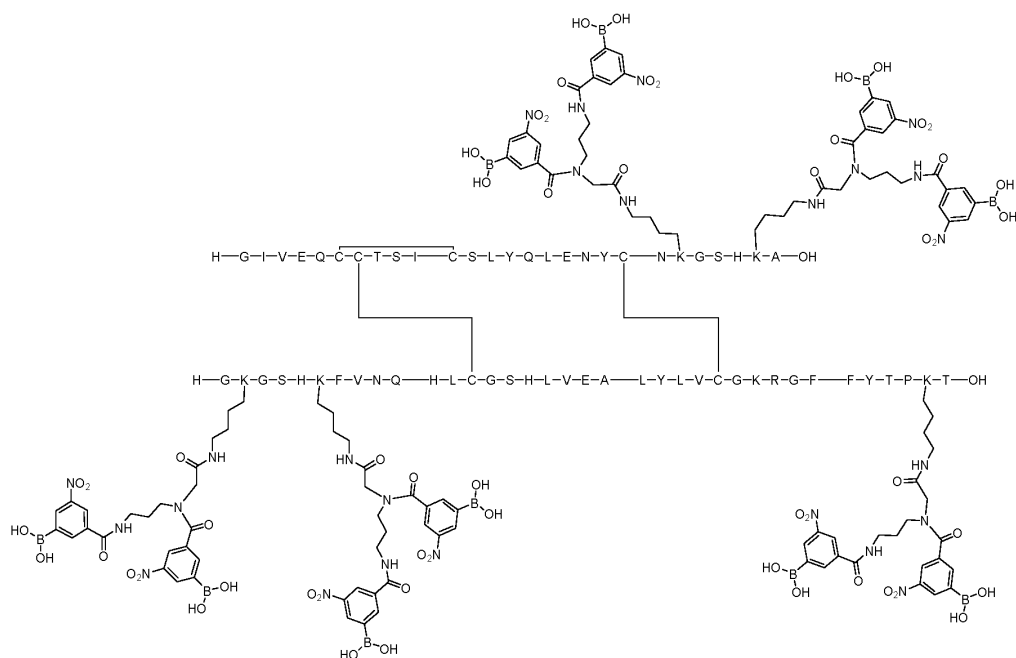


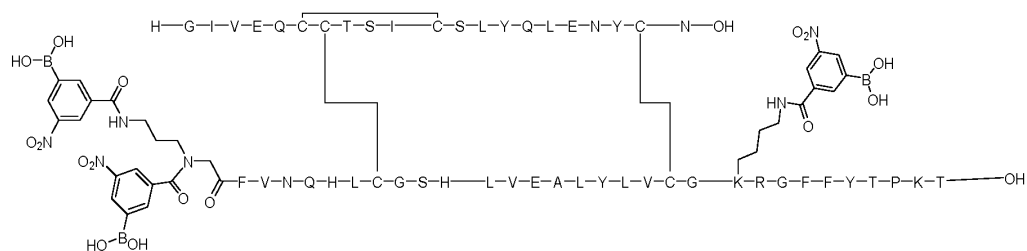
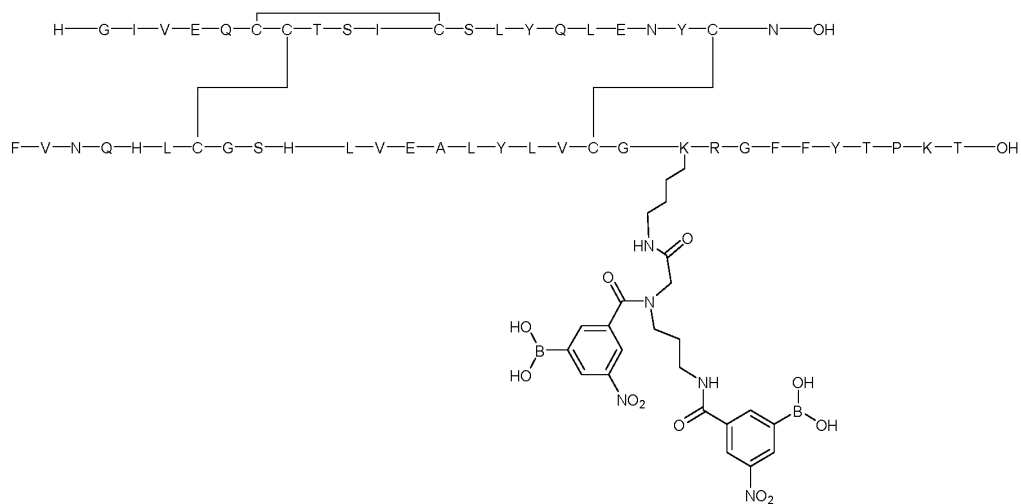
5

10

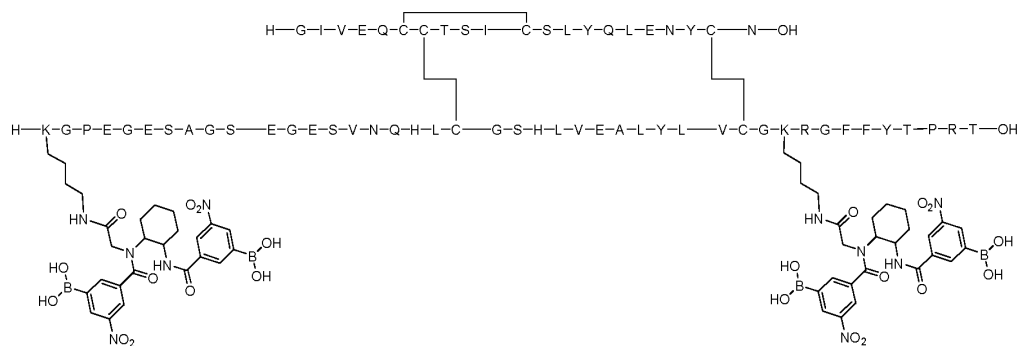
Example 4:

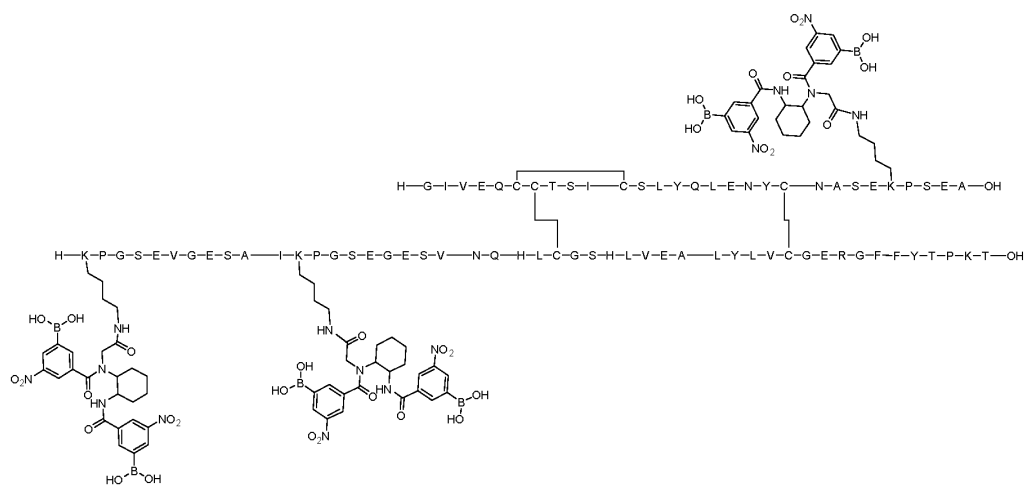
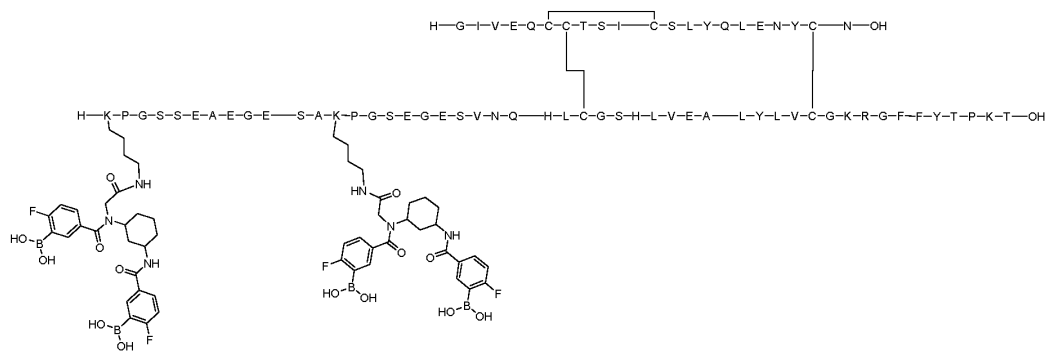
5

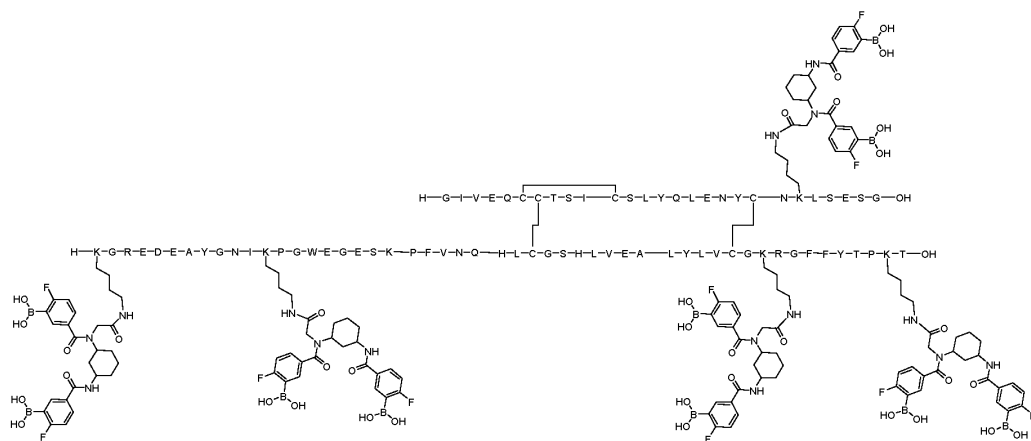
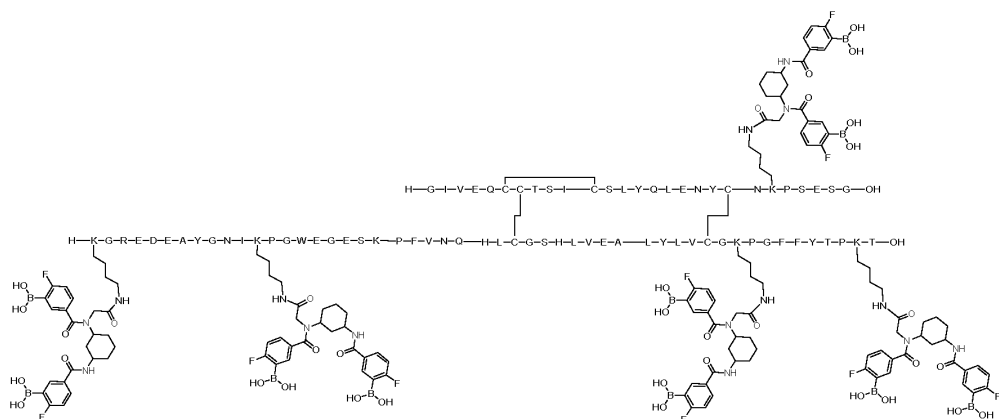
Example 5:

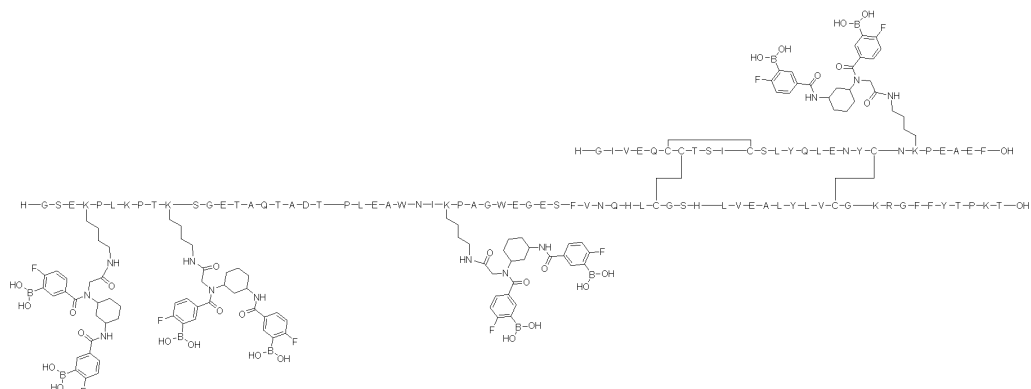
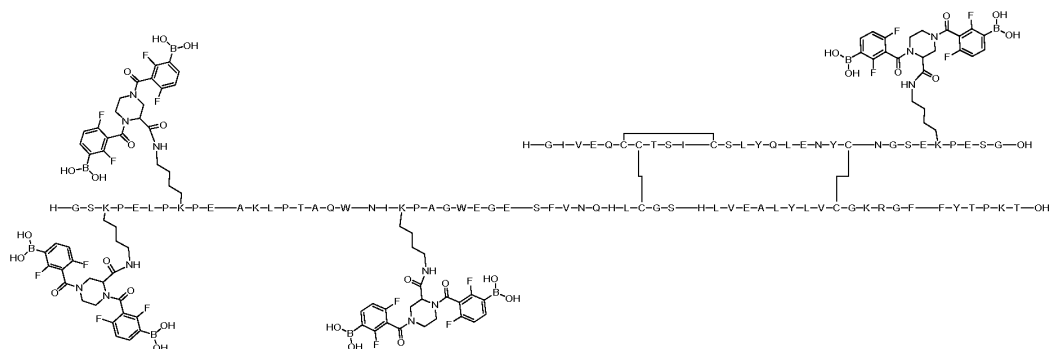
Example 6:**Example 7:**

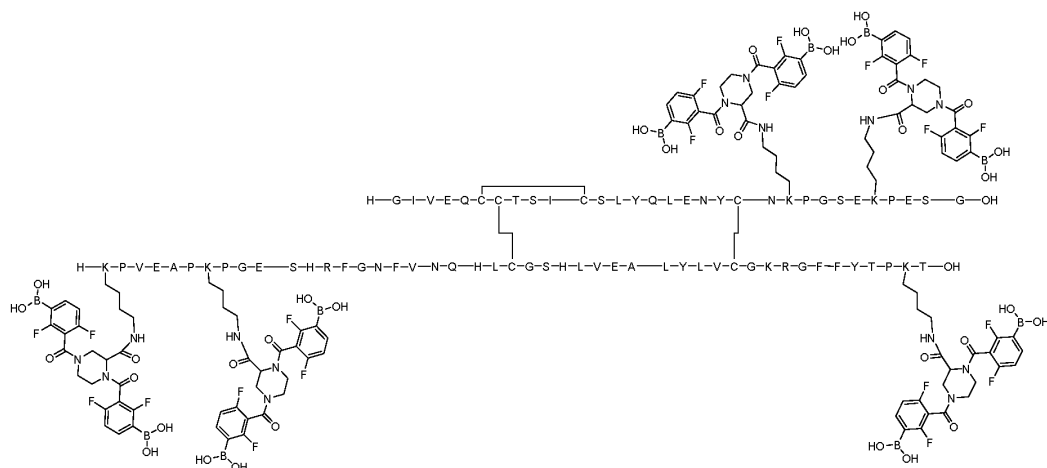
5

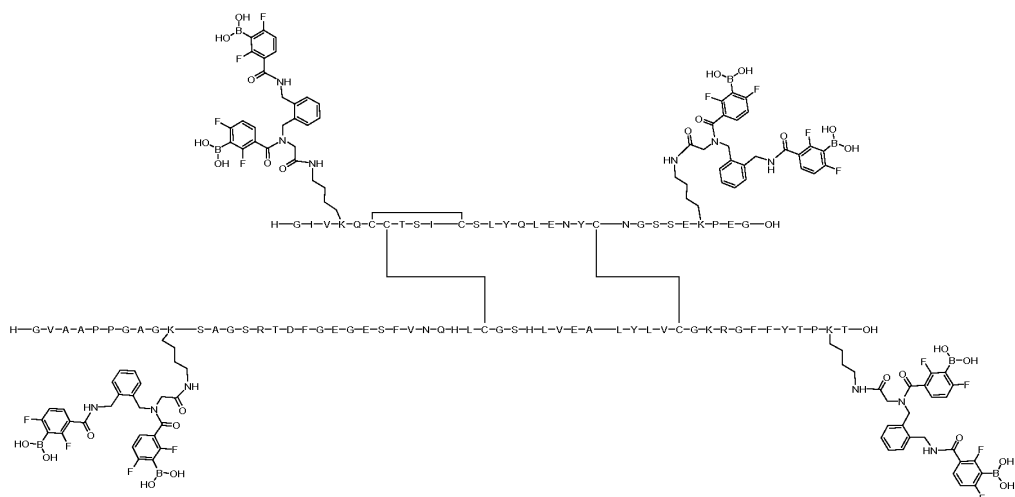
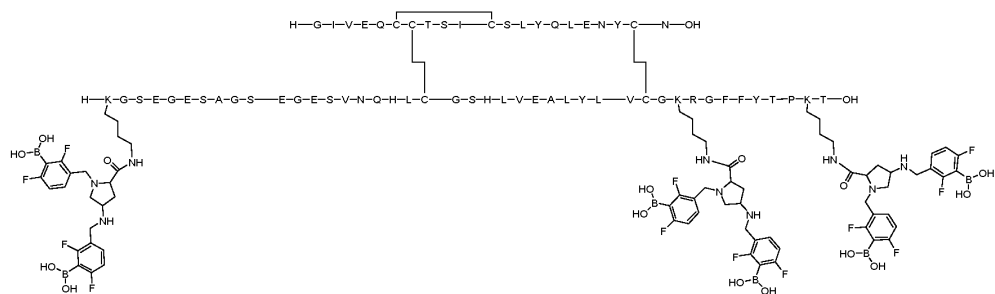
Example 8:

Example 9:**5 Example 10:**

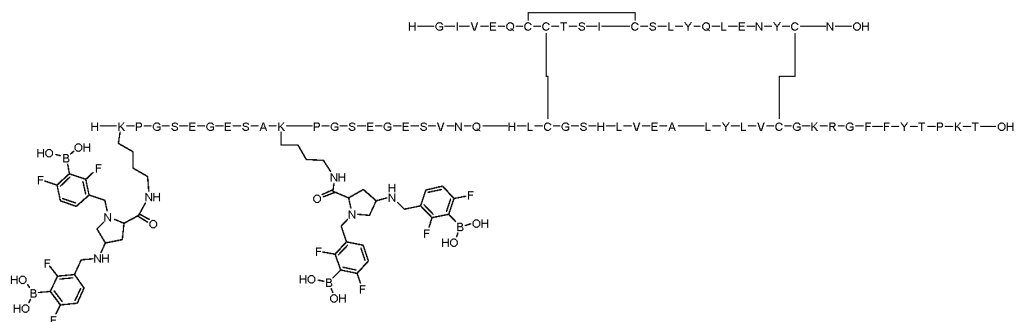
Example 11:**Example 12:**

Example 13:**Example 14:**

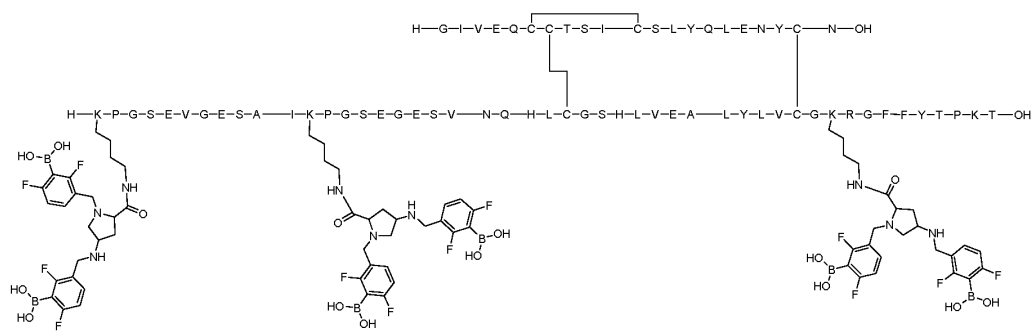
Example 15:**5 Example 16:**

Example 17:**Example 18:**

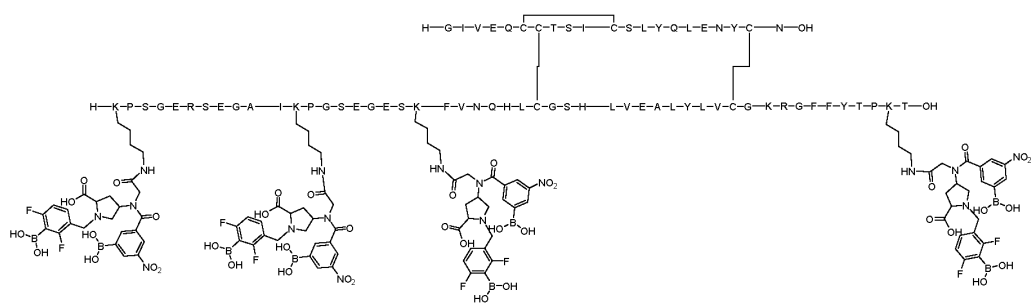
Example 19:

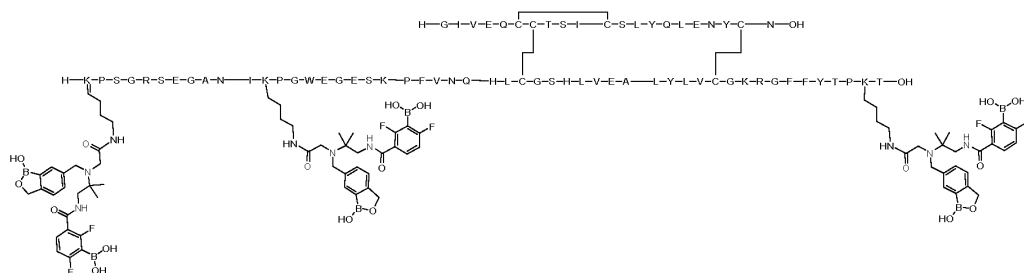
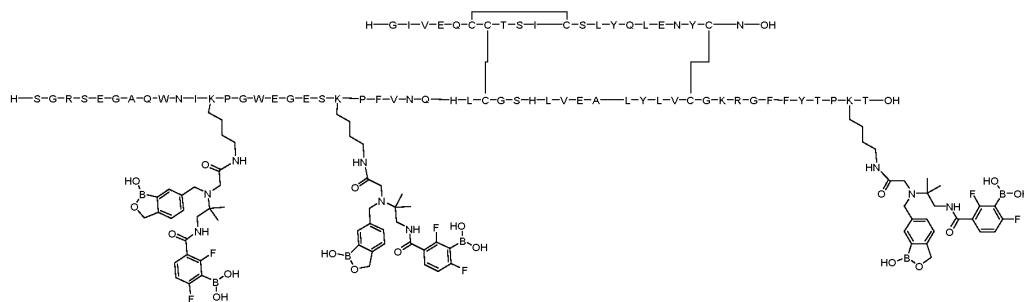


Example 20:

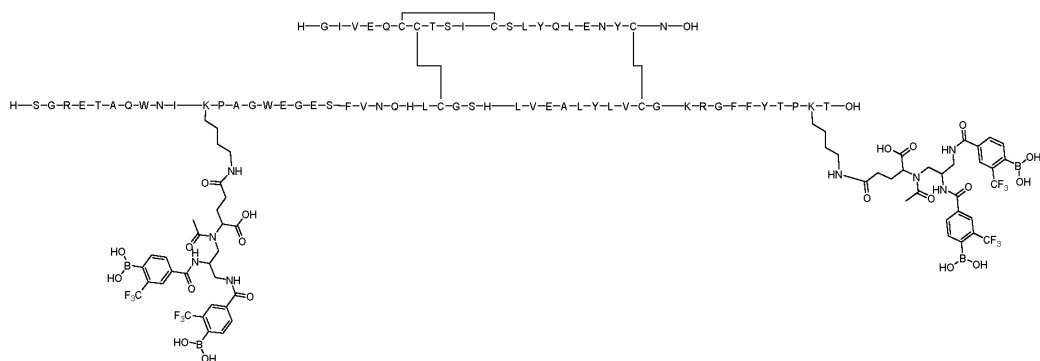


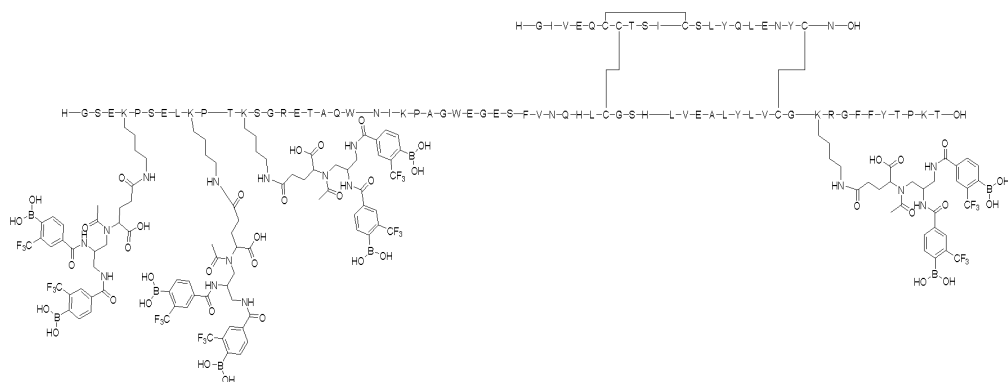
Example 21:



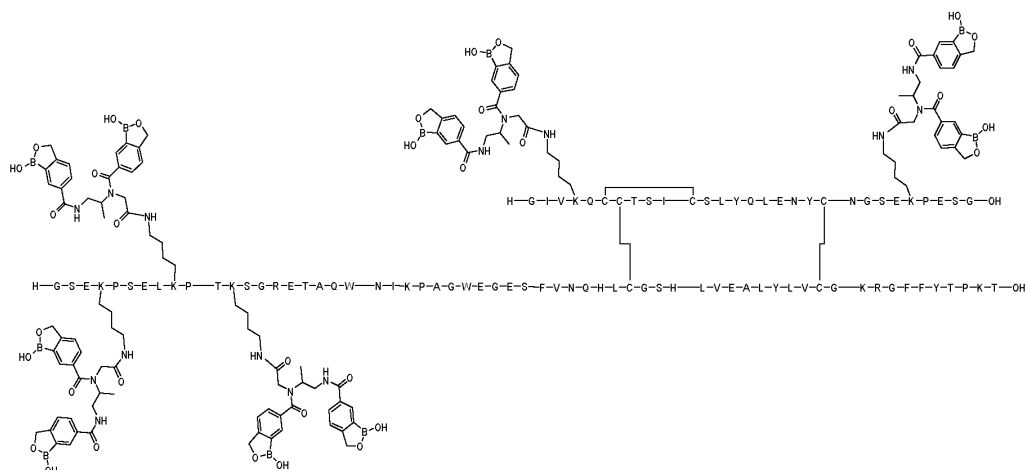
Example 22:**Example 23:**

5

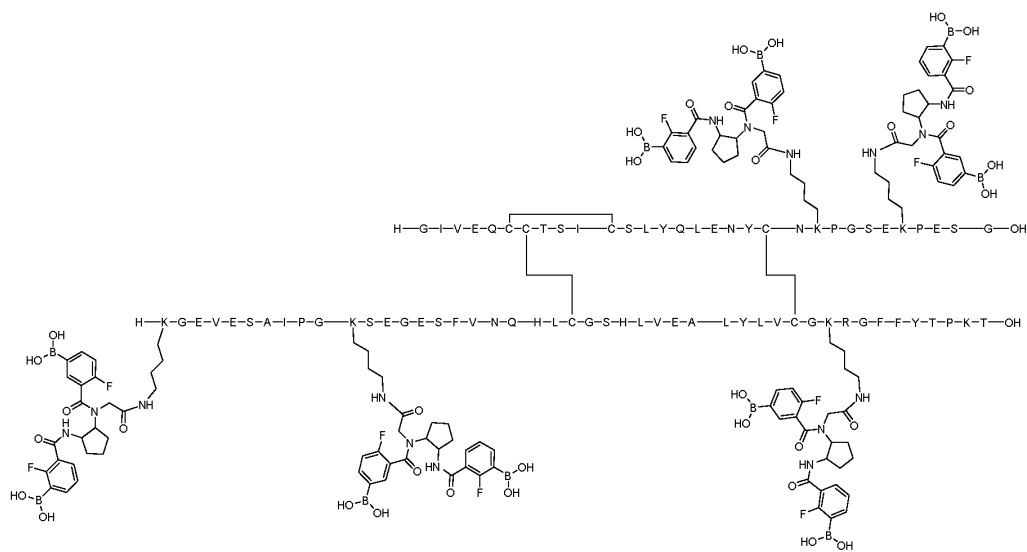
Example 24:

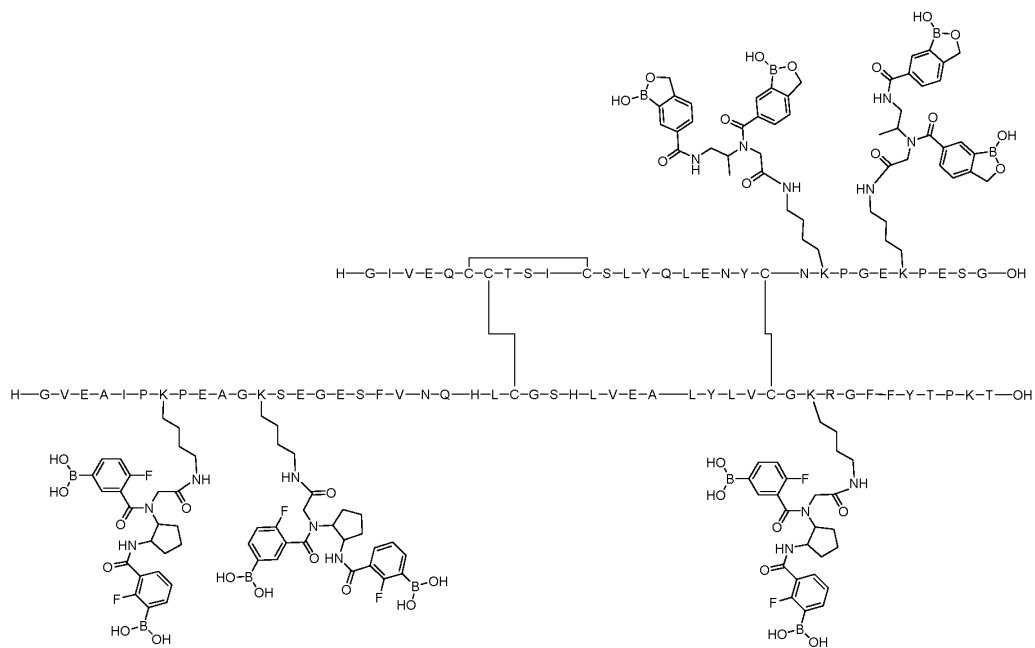


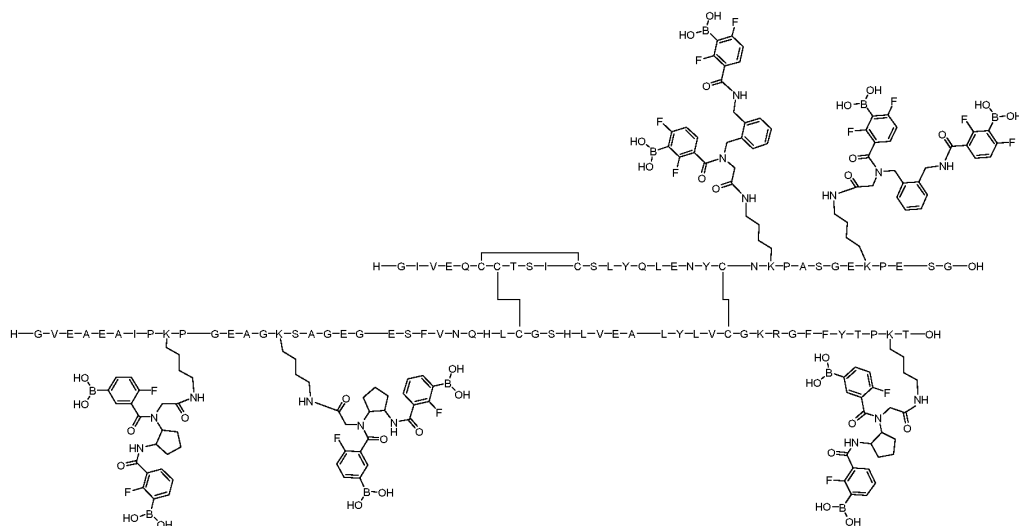
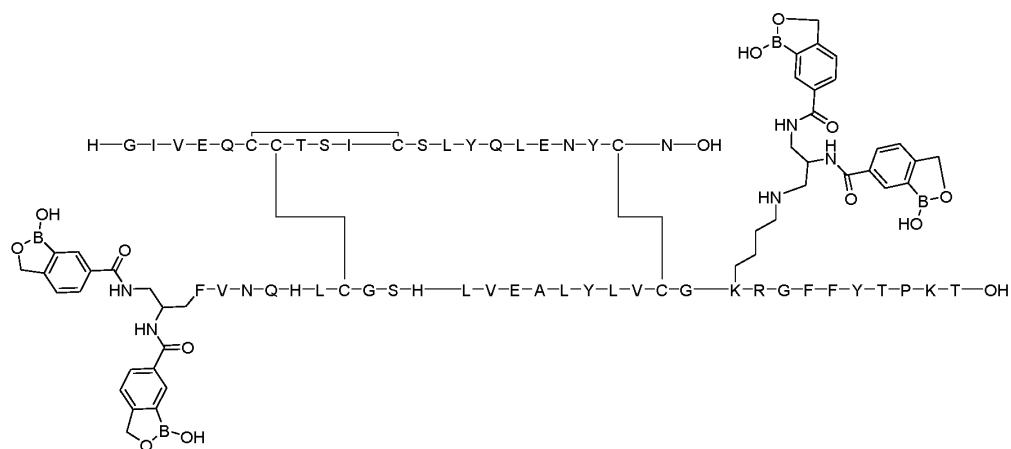
The diagram illustrates the chemical structure of the protein backbone of the 19S proteasome. The sequence is shown as a linear chain of amino acids: H-S-G-R-S-E-G-A-Q-W-N-I-K-P-G-W-E-G-E-S-K-P-F-V-N-Q-H-L-C-G-S-H-L-V-E-A-L-Y-L-V-C-G-K-R-G-F-F-Y-T-P-K-T-OH. The structure is decorated with various chemical groups, including boronic acid and boronate ester groups, and is shown in a 3D representation.

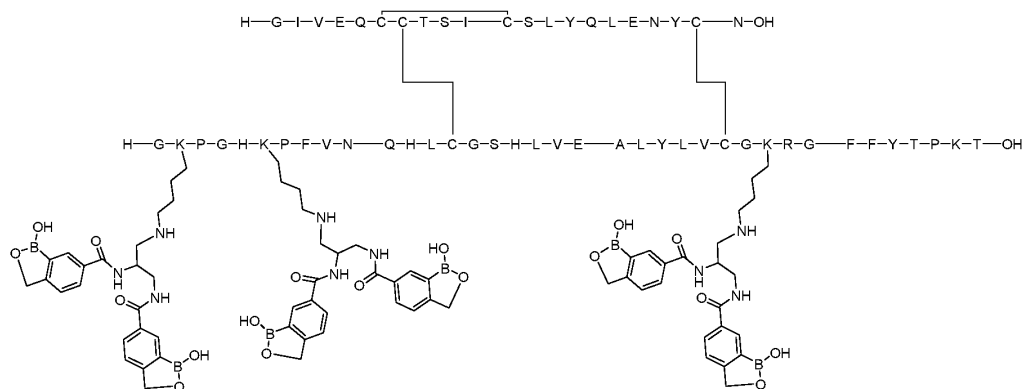
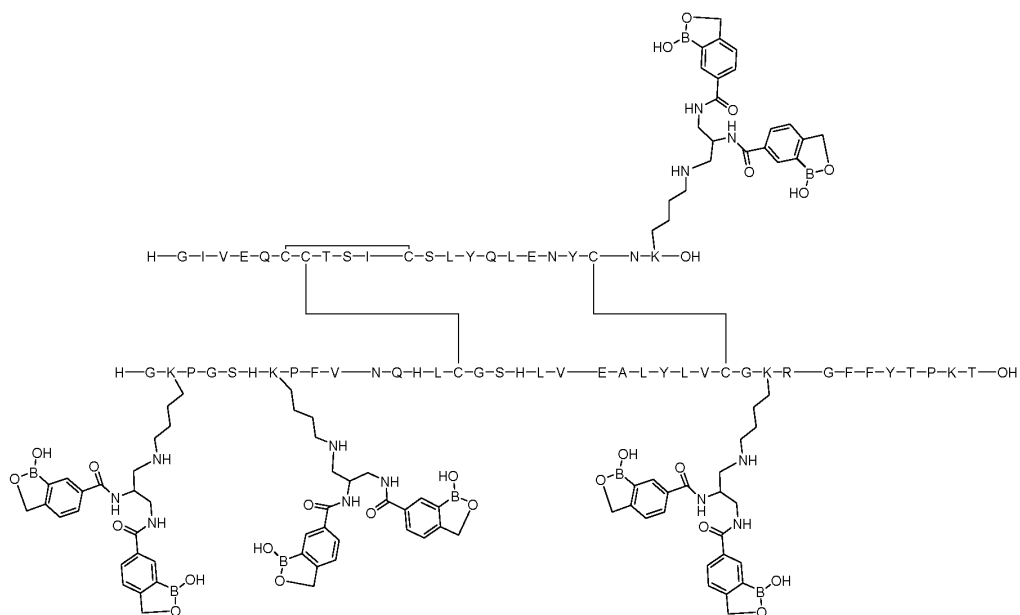


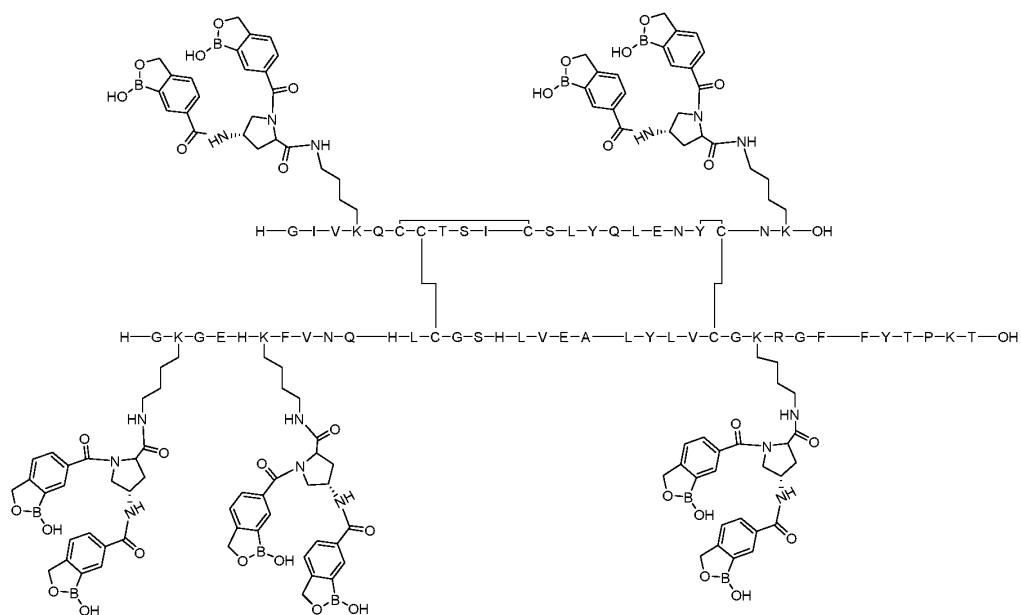
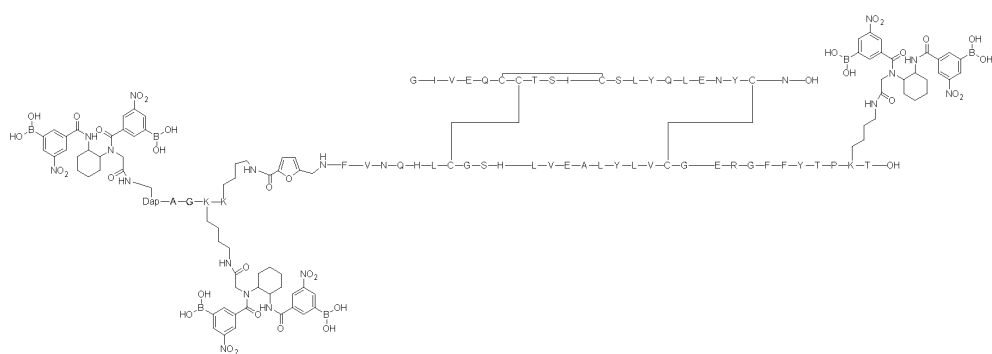
Example 28:

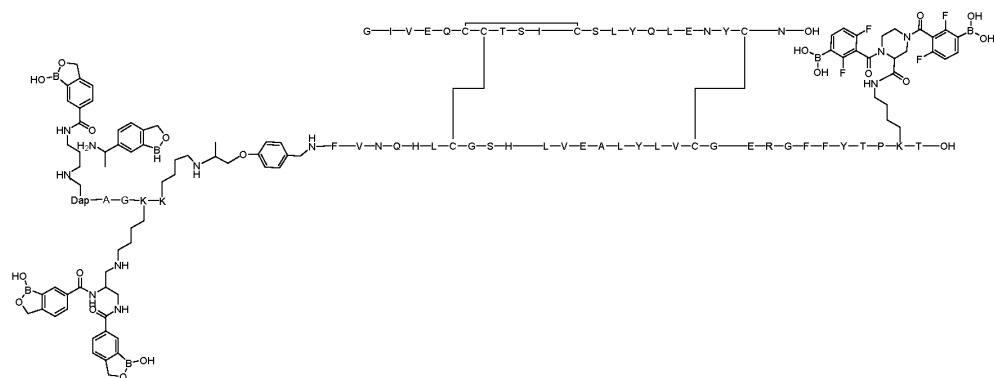
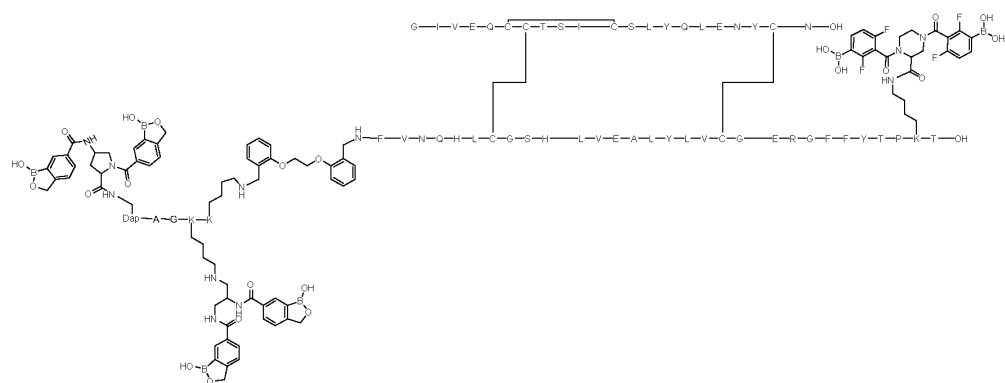
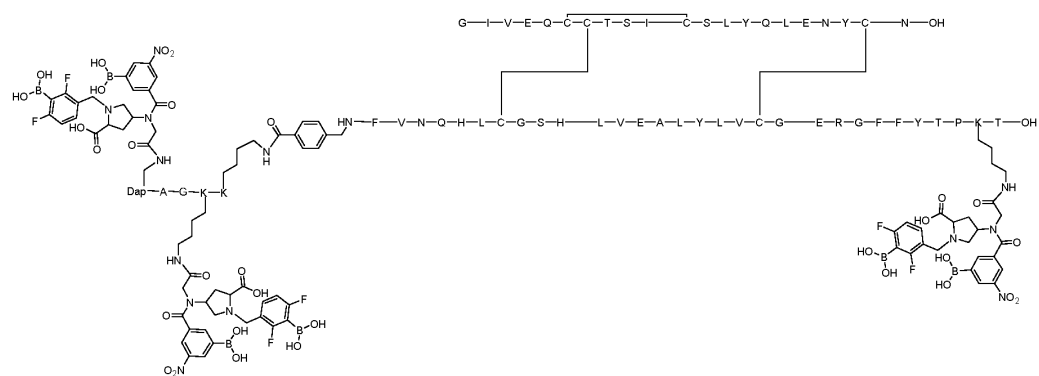


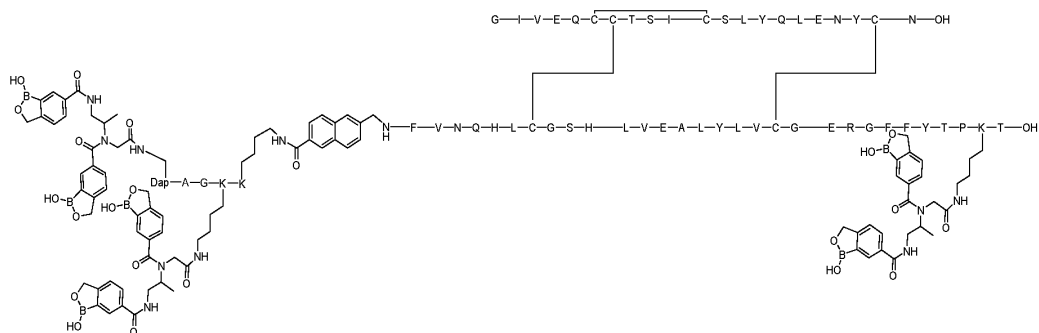
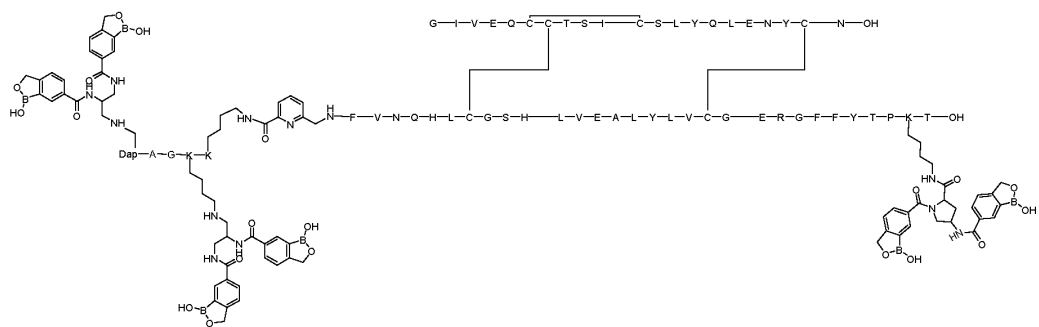
Example 29:

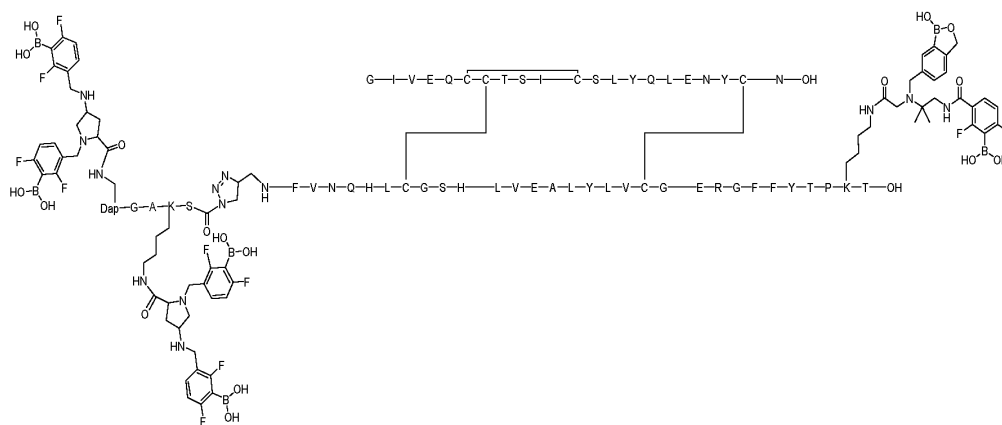
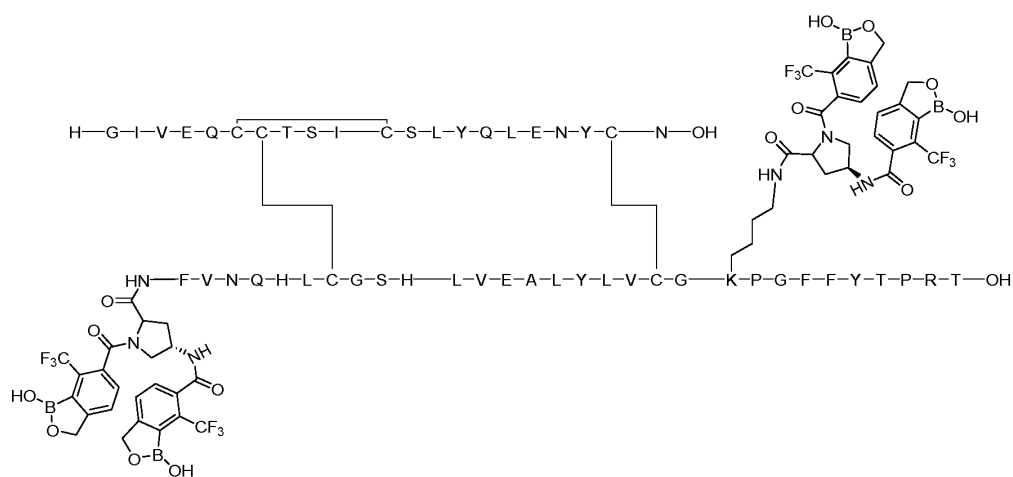
Example 30:**Example 31:**

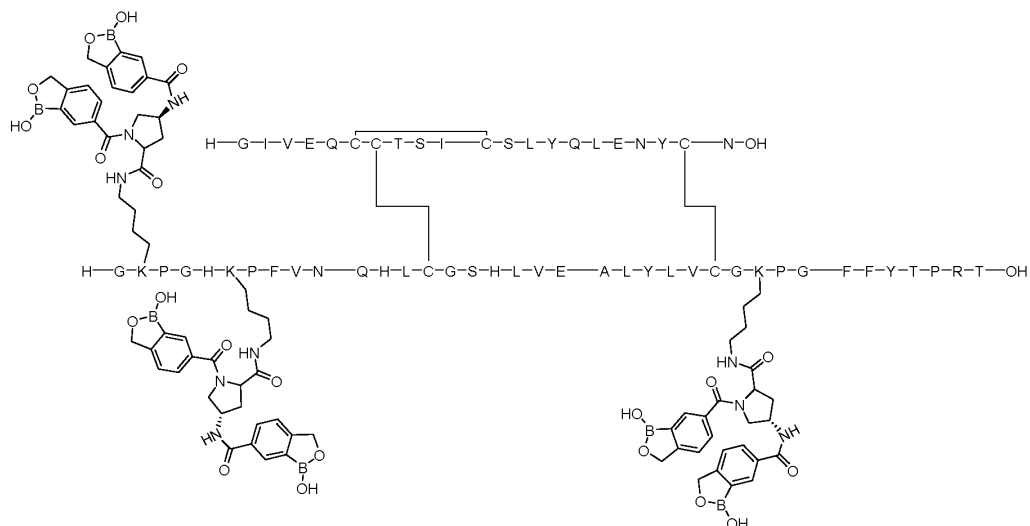
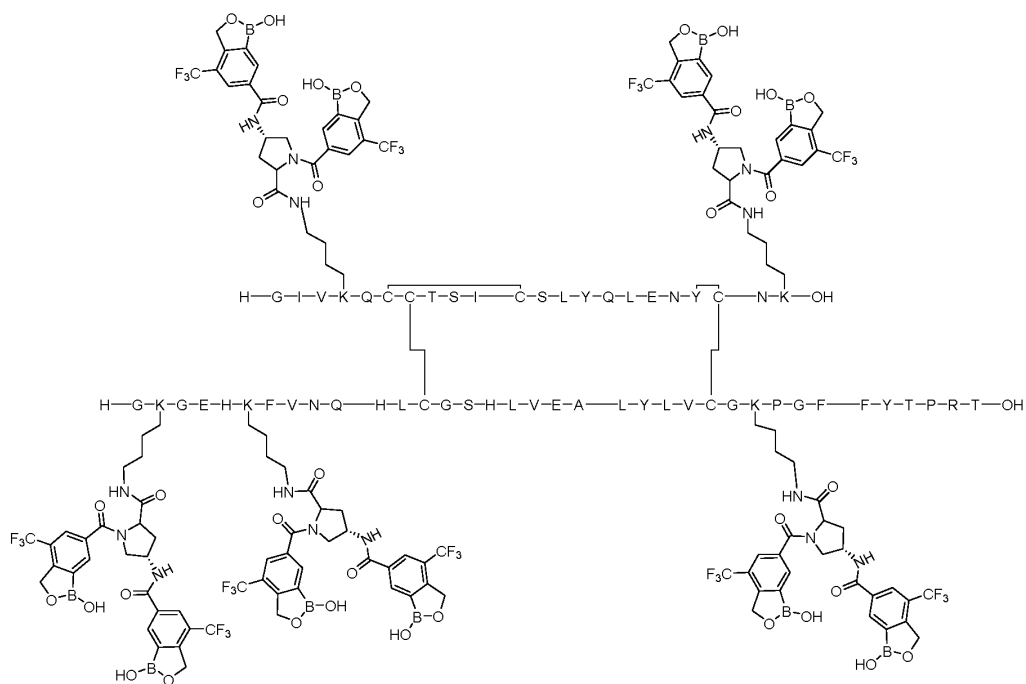
Example 32:**Example 33:**

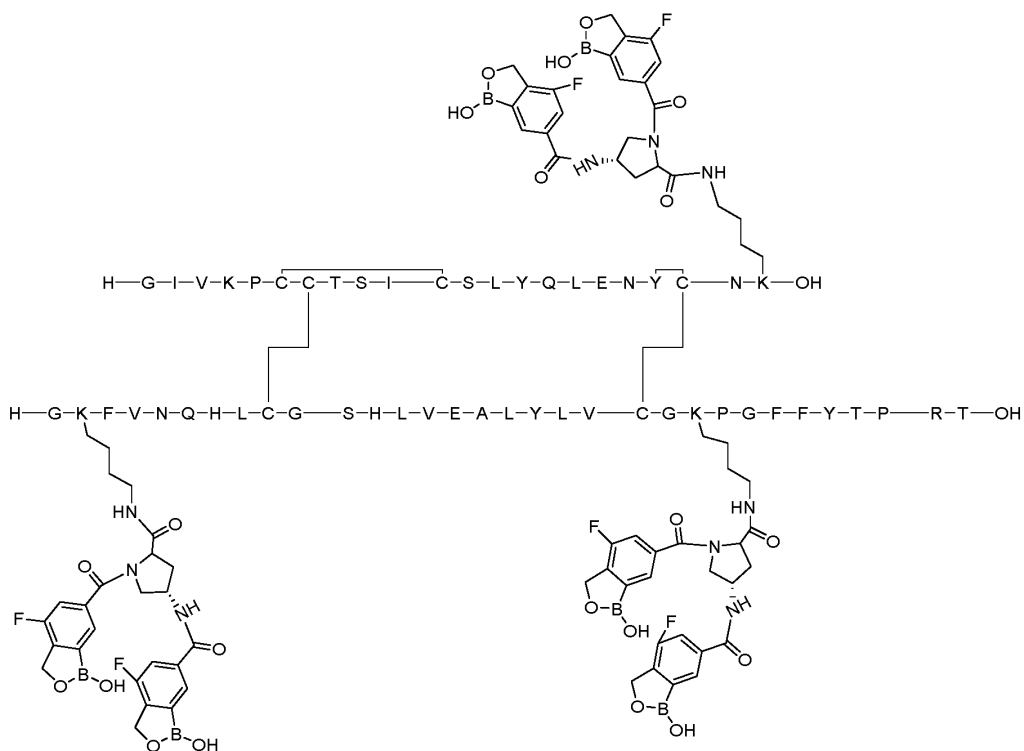
Example 34:**Example 35**

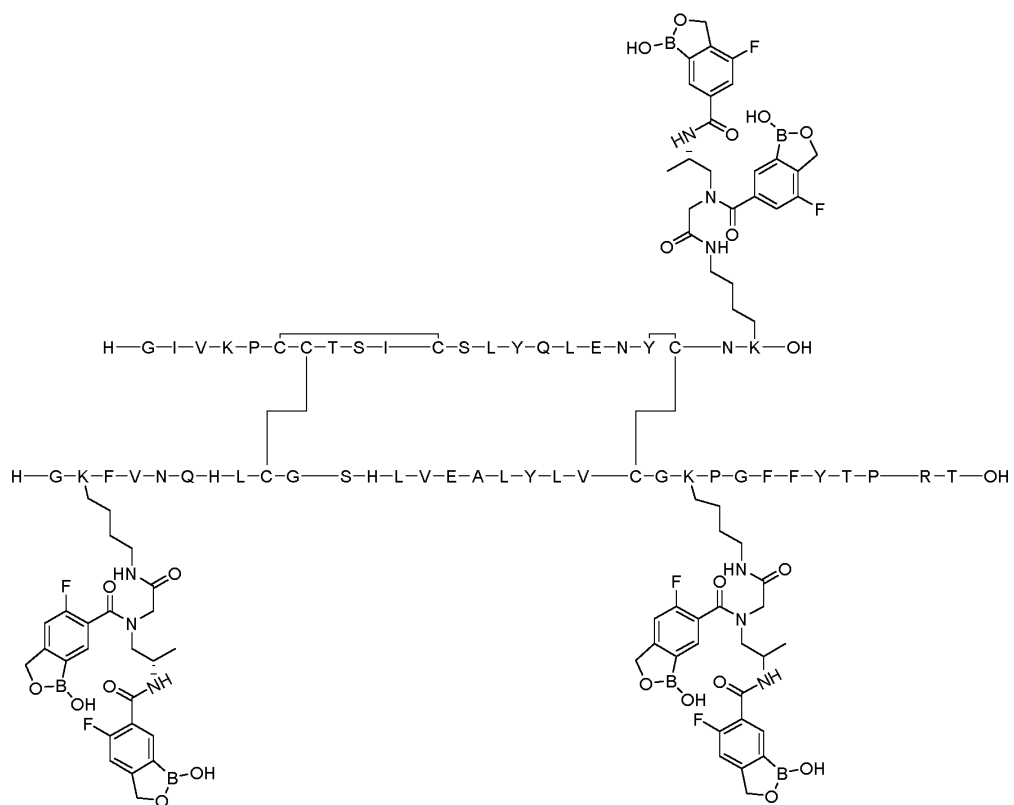
Example 36:**Example 37:****5 Example 38:**

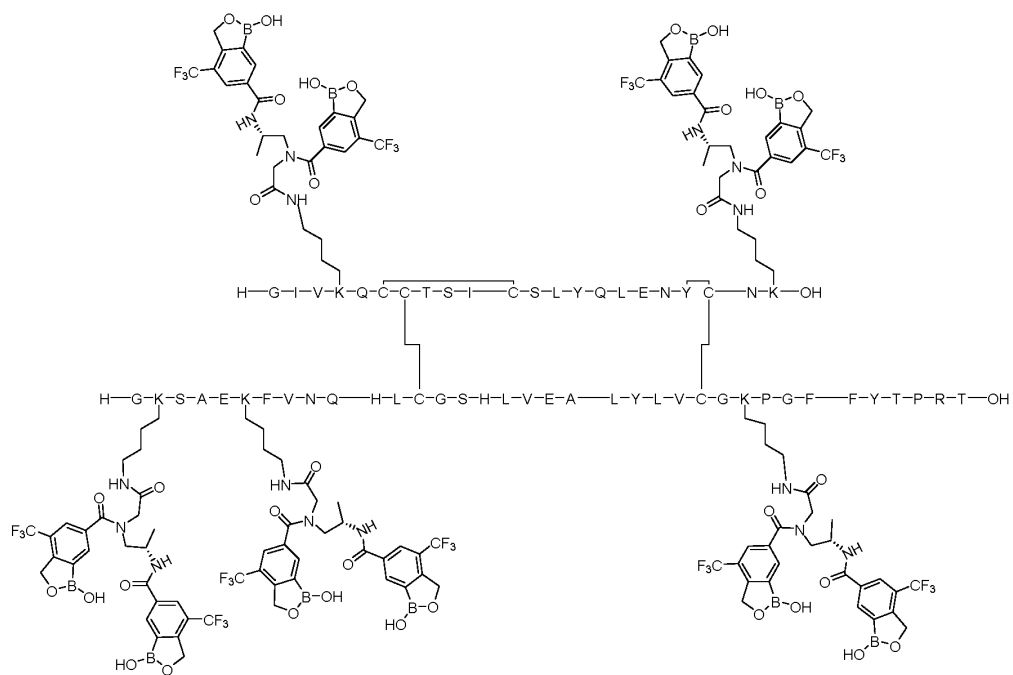
Example 39:**Example 40:**

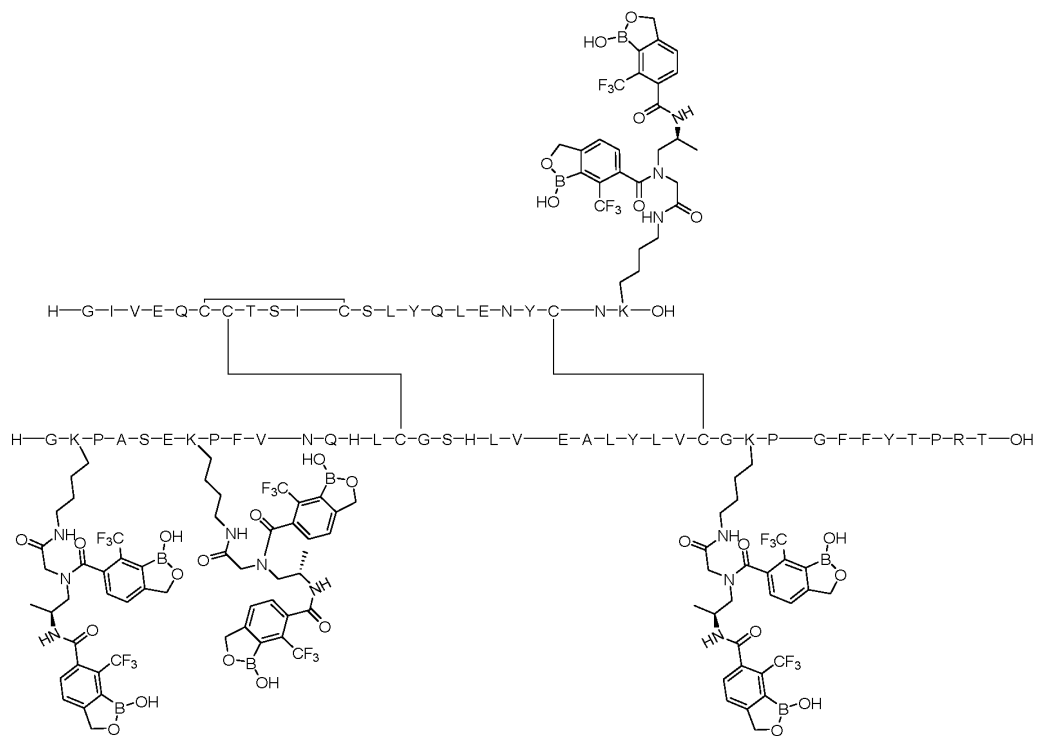
Example 41:**Example 42:**

Example 43:**Example 44:**

Example 45:

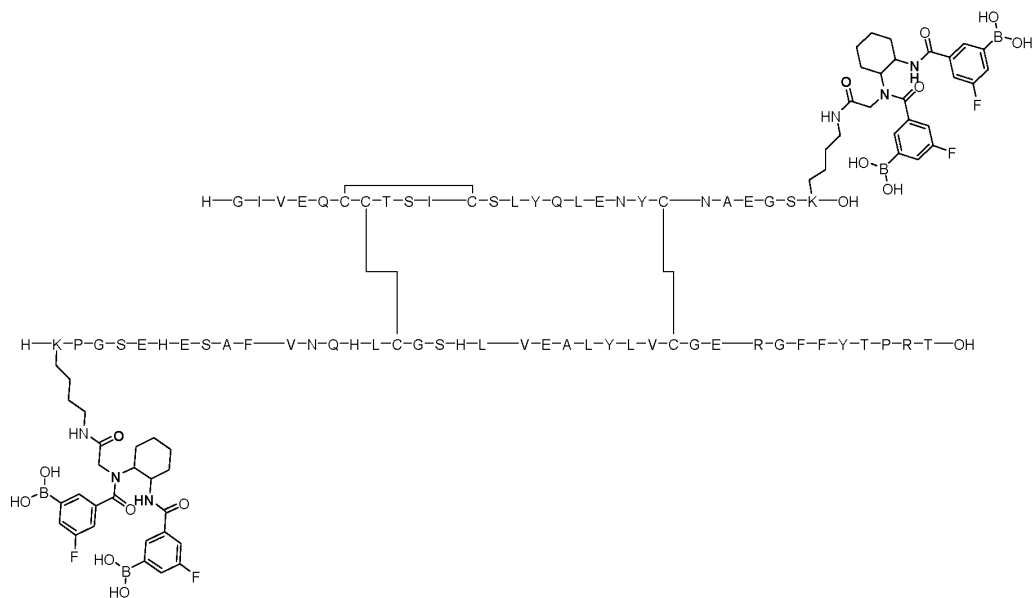
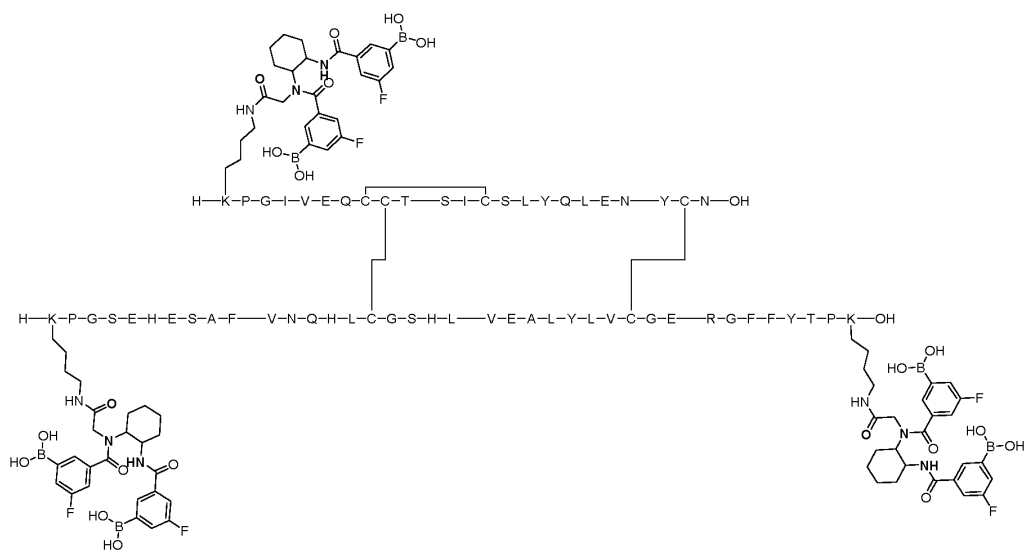
Example 46:

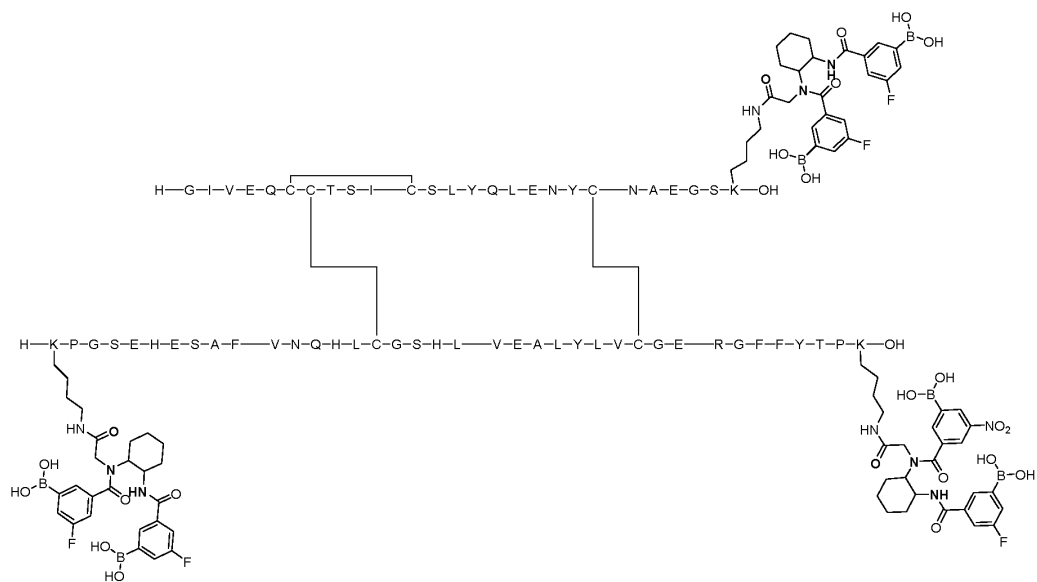
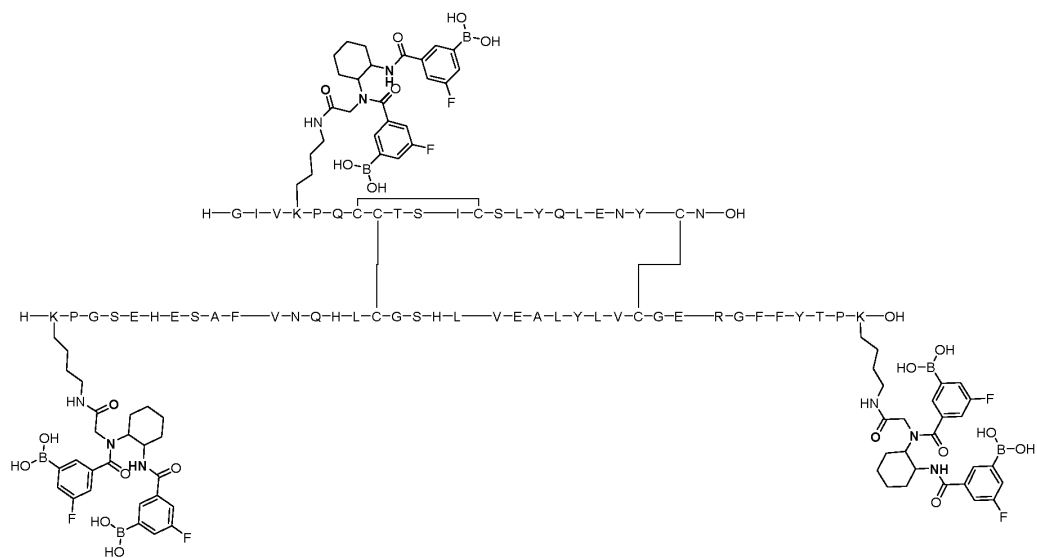
Example 47:

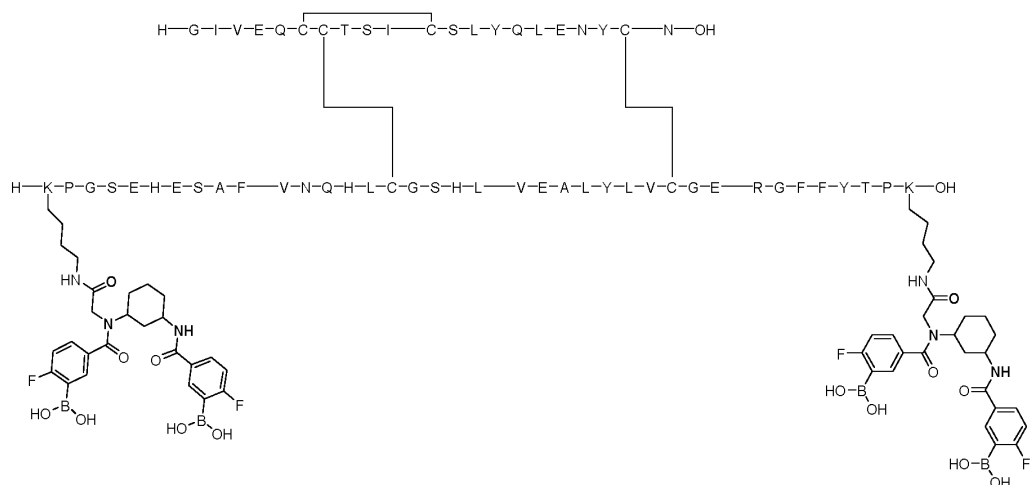
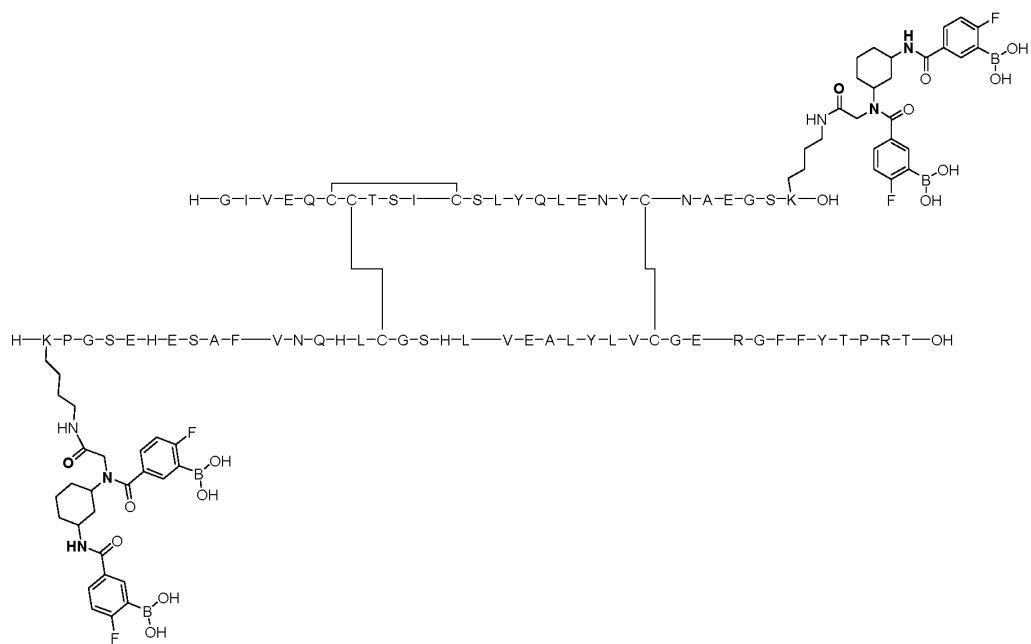


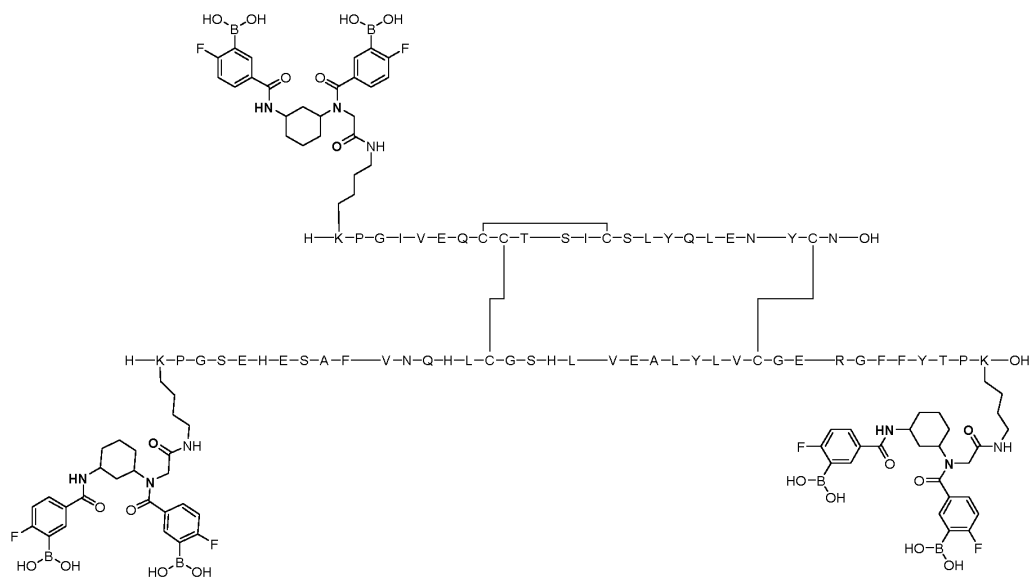
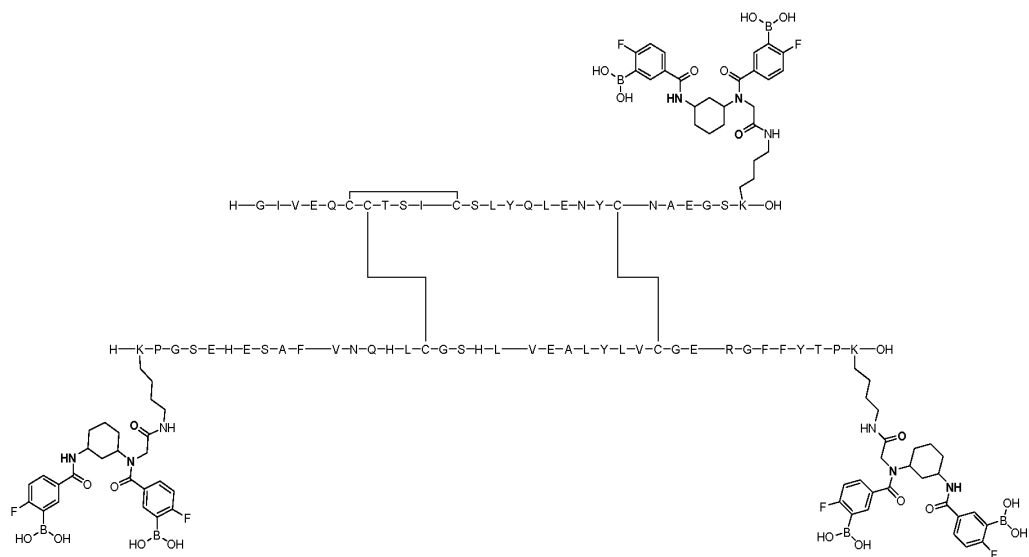
$$\text{H}-\text{G}-\text{I}-\text{V}-\text{E}-\text{Q}-\text{C}-\text{C}-\text{T}-\text{S}-\text{I}-\text{C}-\text{S}-\text{L}-\text{Y}-\text{Q}-\text{L}-\text{E}-\text{N}-\text{Y}-\text{C}-\text{N}-\text{OH}$$

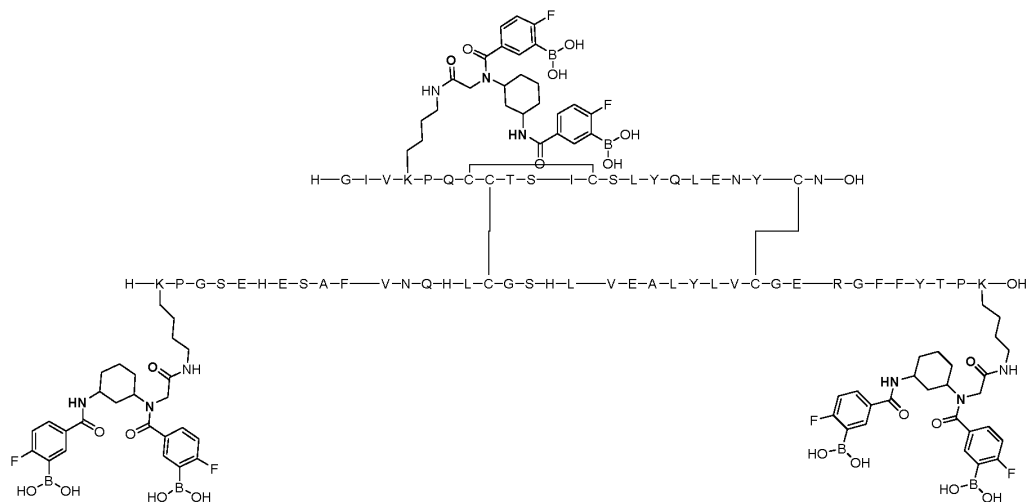
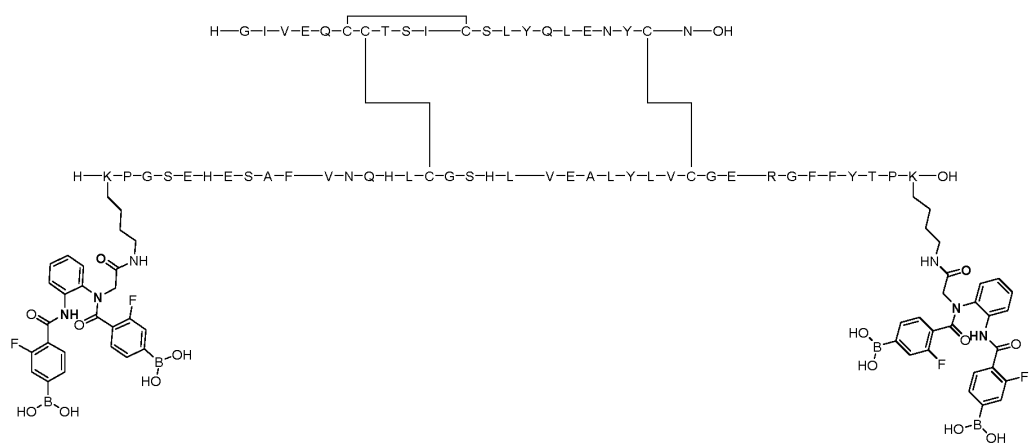
$$\text{H}-\text{K}-\text{P}-\text{G}-\text{S}-\text{E}-\text{H}-\text{E}-\text{S}-\text{A}-\text{F}-\text{V}-\text{N}-\text{Q}-\text{H}-\text{L}-\text{C}-\text{G}-\text{S}-\text{H}-\text{L}-\text{V}-\text{E}-\text{A}-\text{L}-\text{Y}-\text{L}-\text{V}-\text{C}-\text{G}-\text{E}-\text{R}-\text{G}-\text{F}-\text{F}-\text{Y}-\text{T}-\text{P}-\text{K}-\text{OH}$$

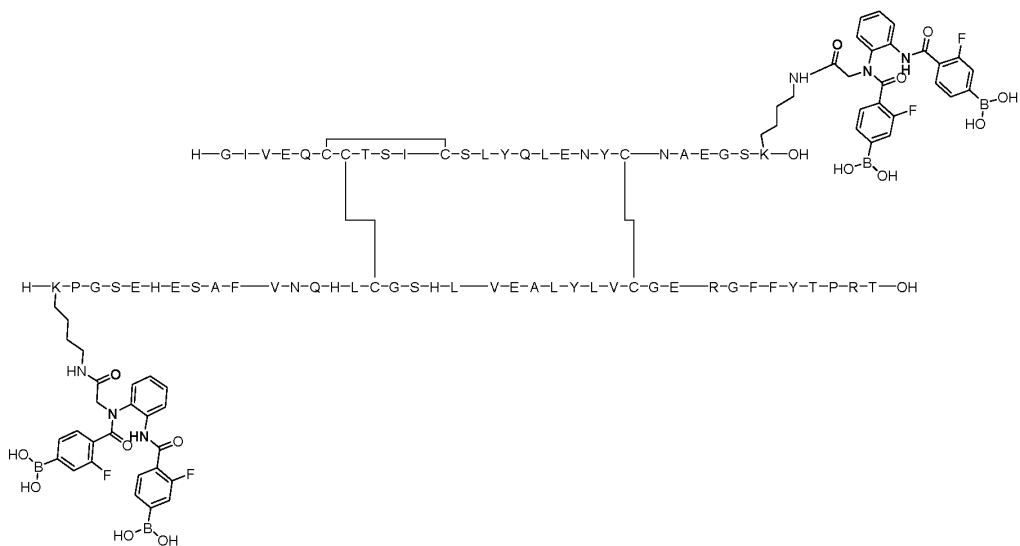
Example 50:**Example 51:**

Example 52:**Example 53:**

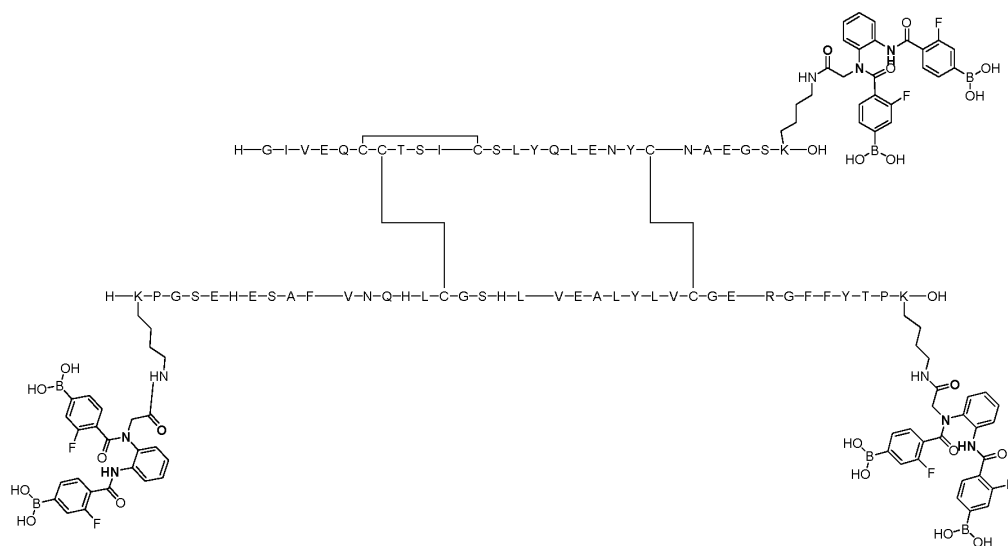
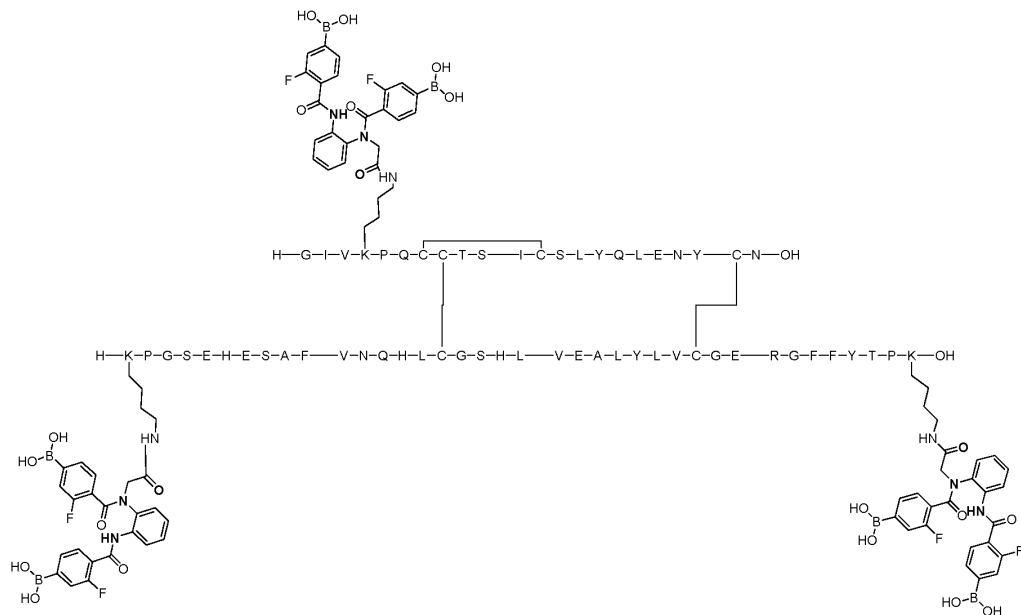
Example 54:**Example 55:**

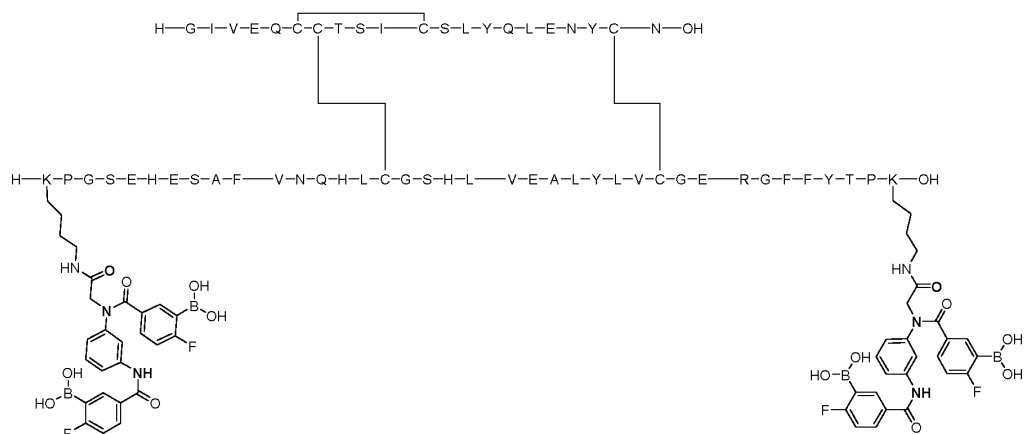
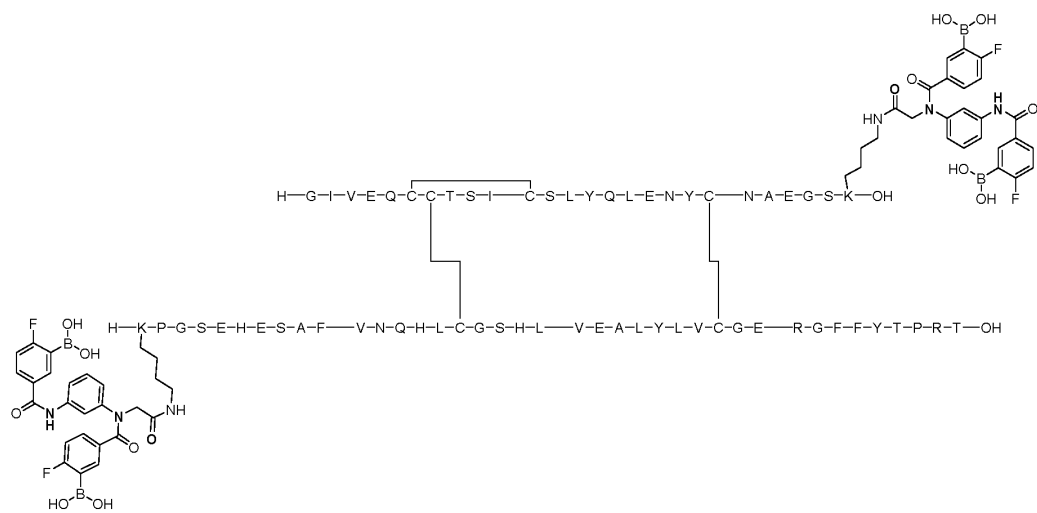
Example 56:**Example 57:**

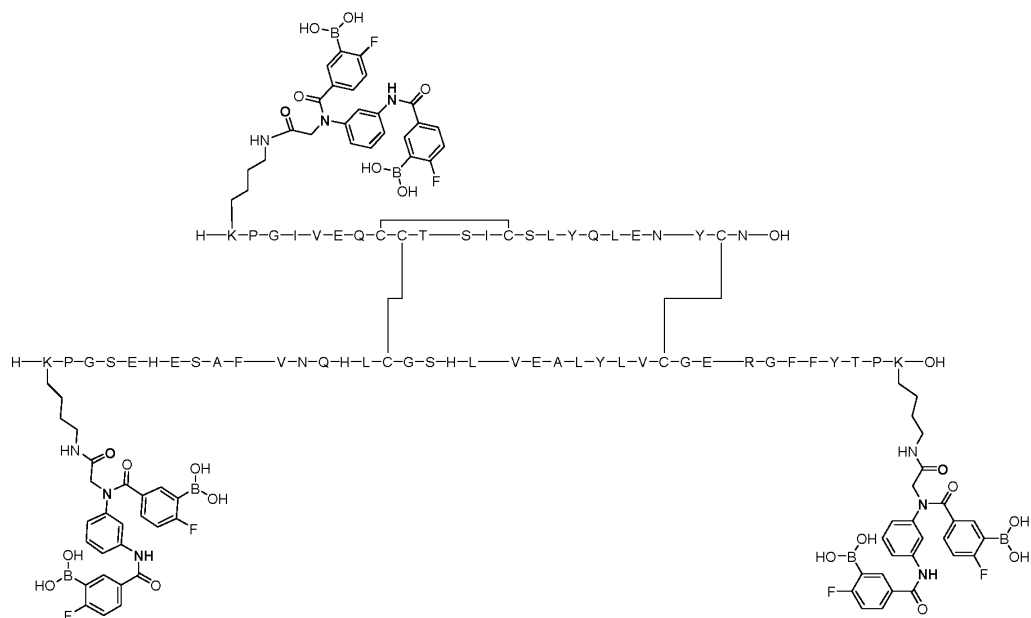
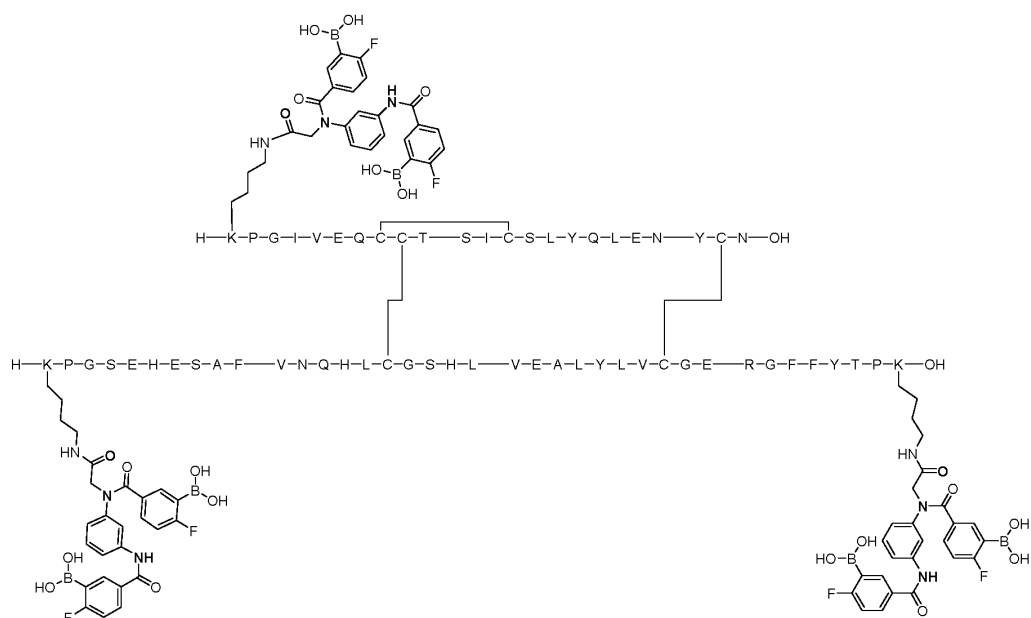
Example 58:**Example 59:**

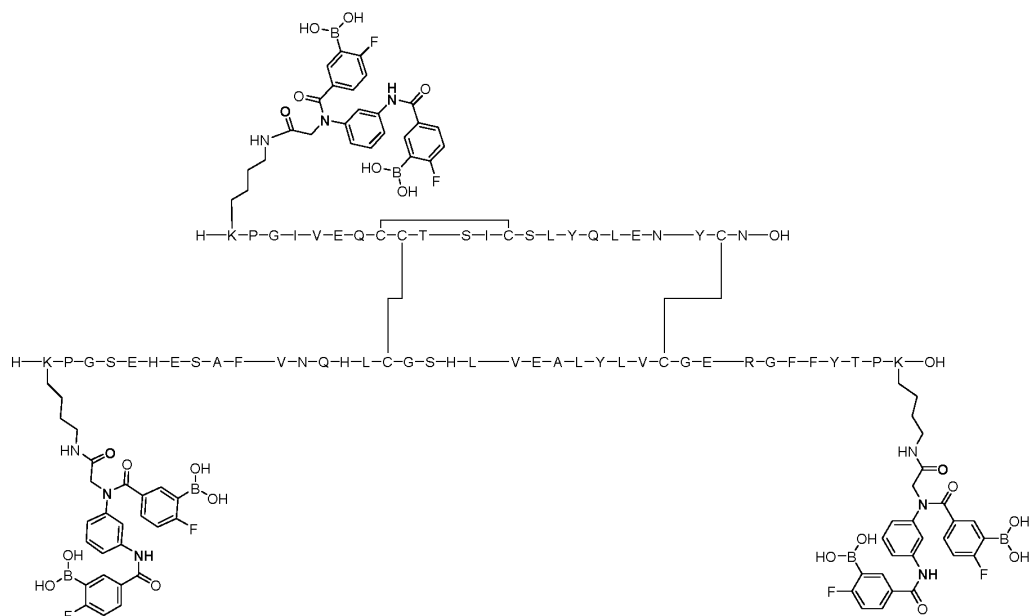
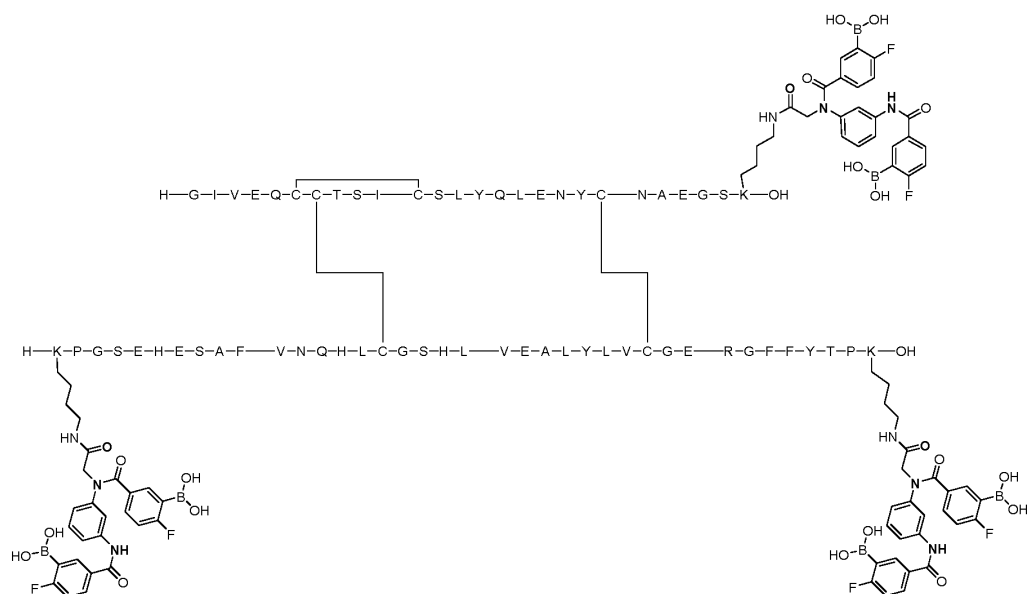


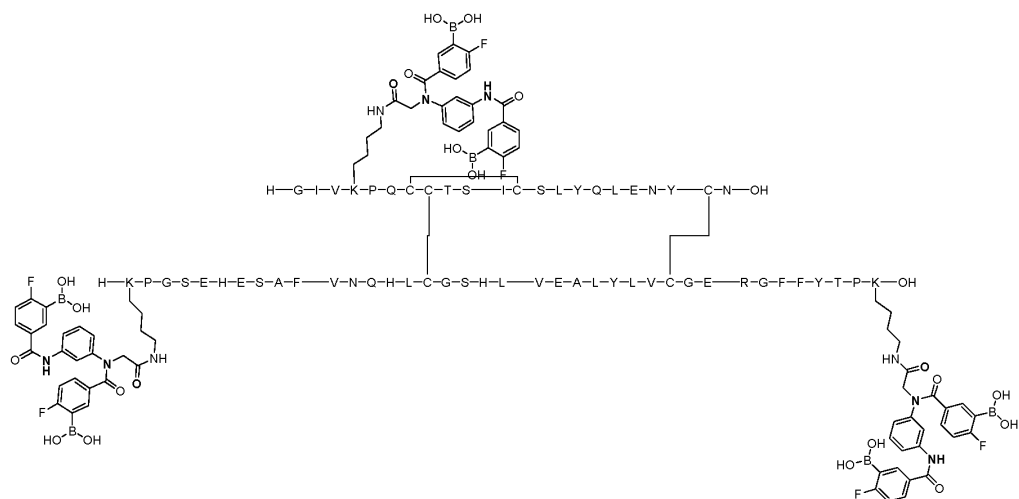
The diagram illustrates the chemical structure of the 12th generation Z-synthetic peptide library. It features a central peptide backbone with two side chains. The top side chain is a 12th generation Z-synthetic peptide, and the bottom side chain is a 12th generation Y-synthetic peptide. Both side chains are attached to the central backbone via a common peptide linkage. The top side chain is a 12th generation Z-synthetic peptide, and the bottom side chain is a 12th generation Y-synthetic peptide. Both side chains are attached to the central backbone via a common peptide linkage.

Example 62:**Example 63:**

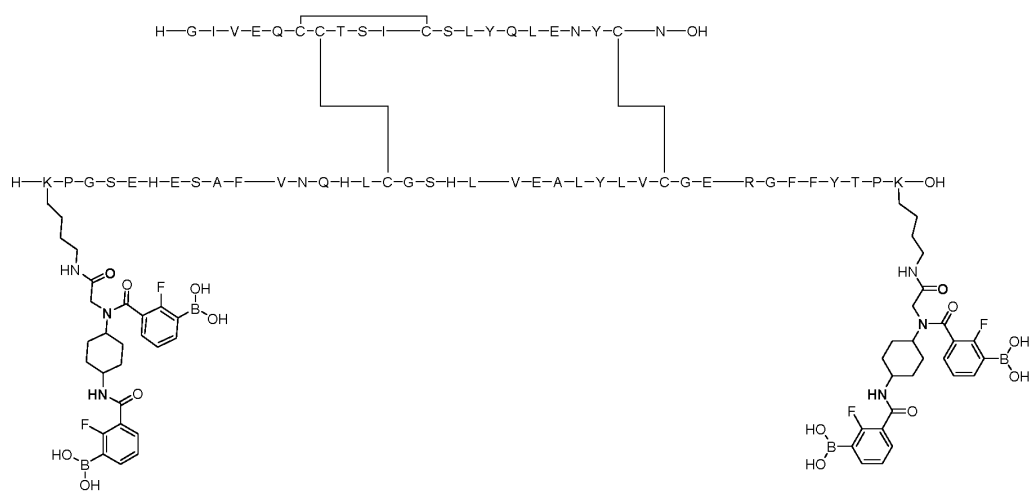
Example 64:**Example 65:**

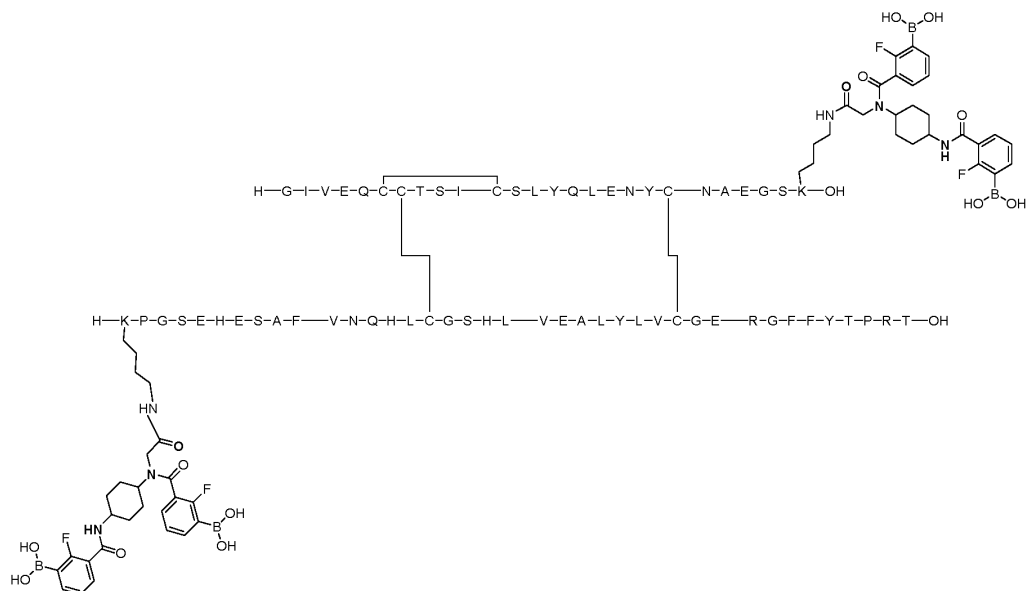
Example 66:**Example 67:**

Example 68:**Example 69:**

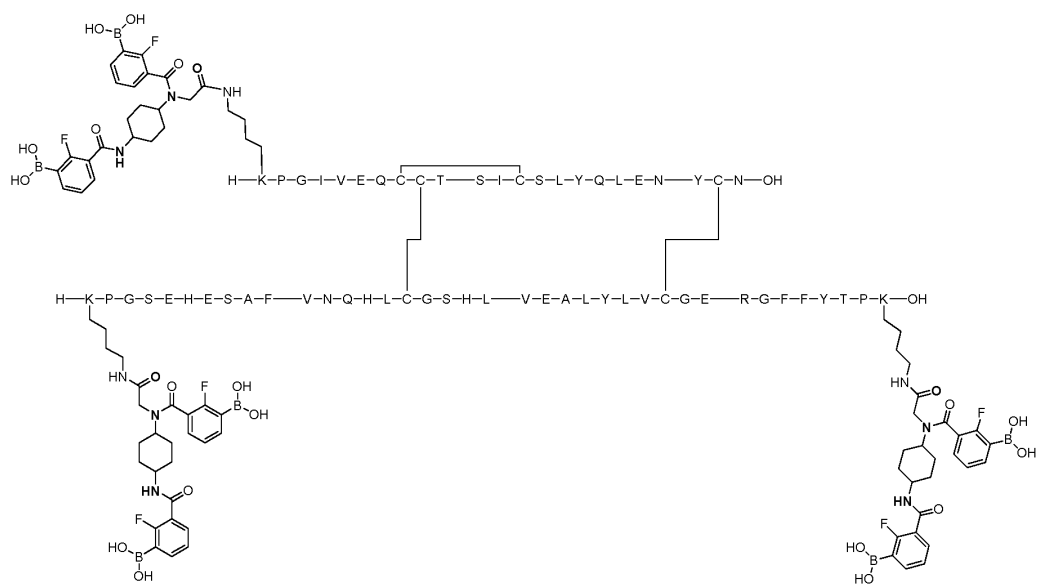


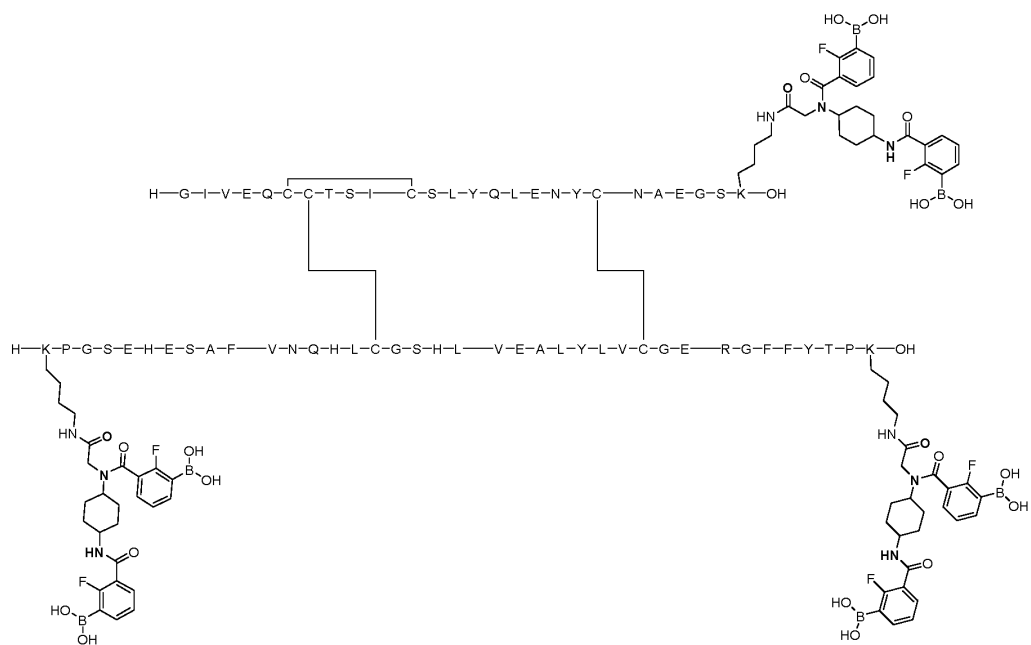
Example 71:

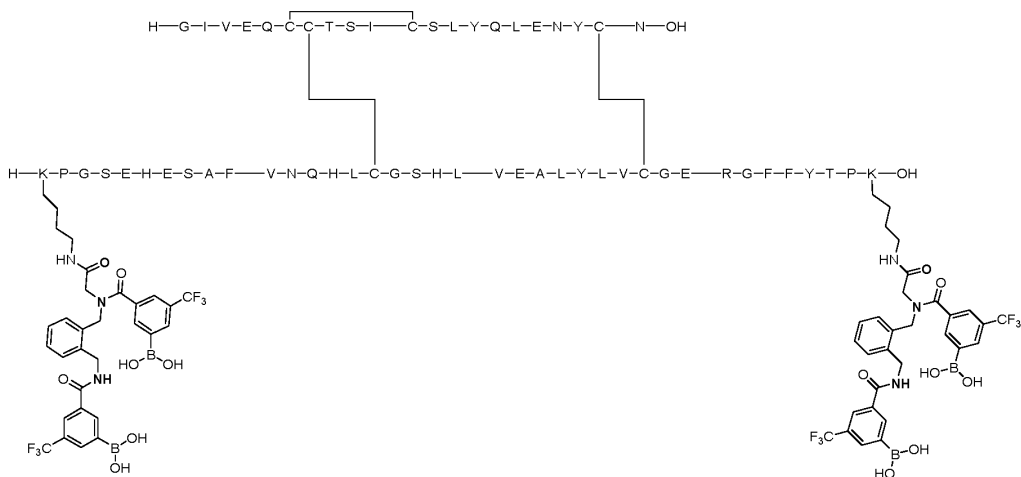


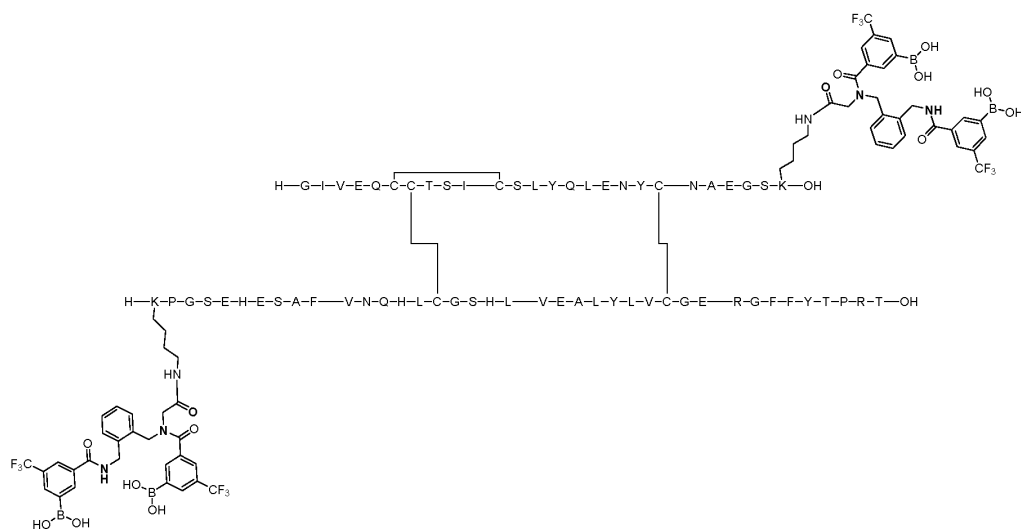
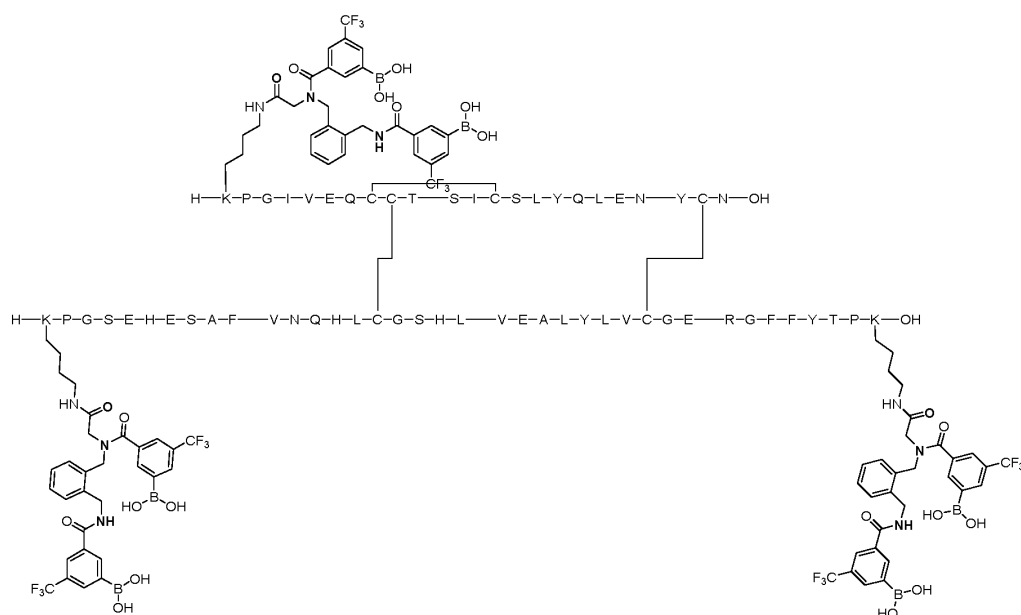


Example 73:

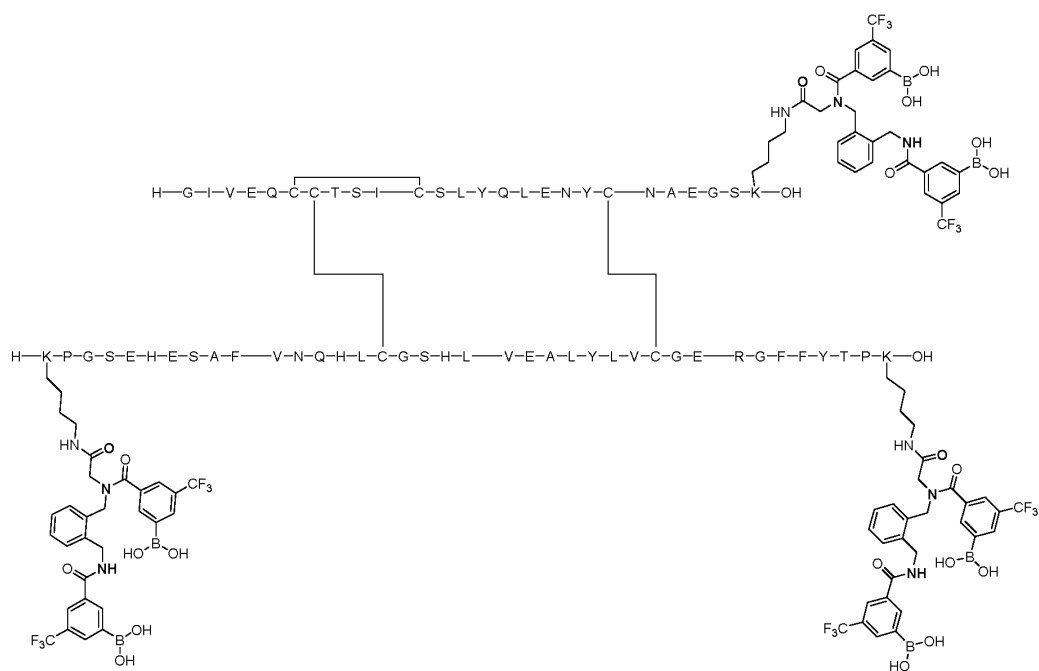




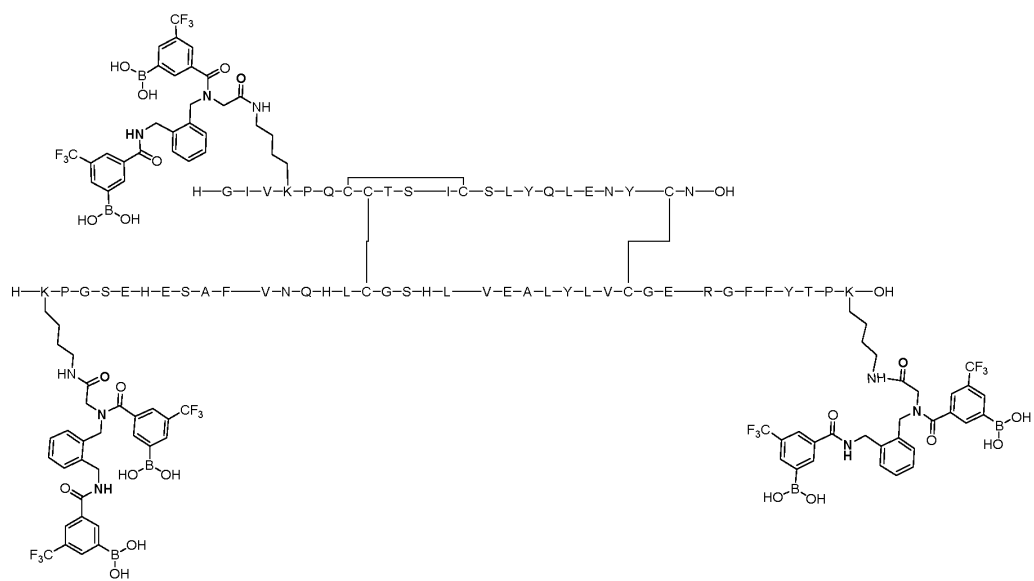


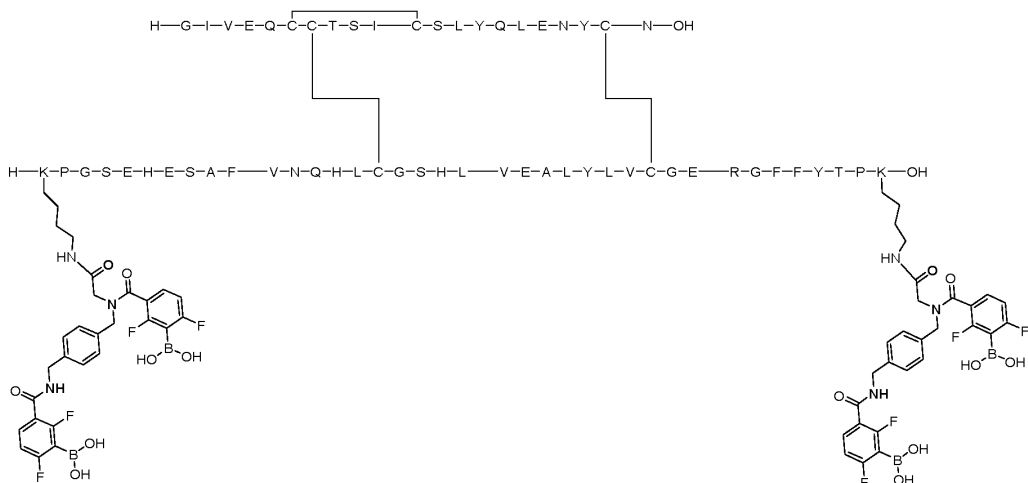
Example 77:**Example 78:**

Example 79:

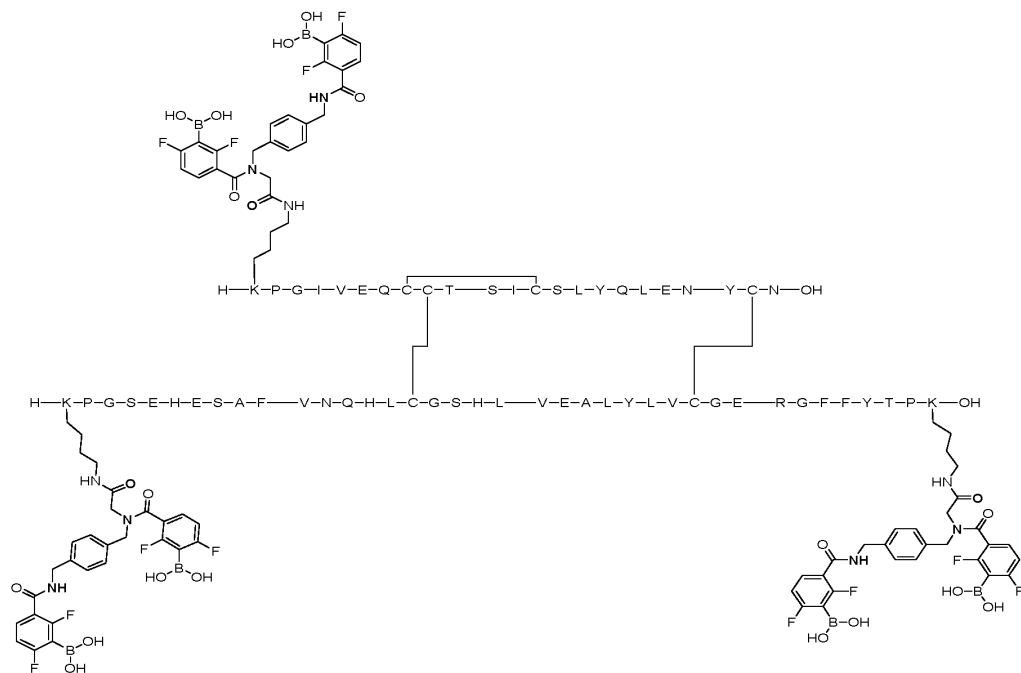


Example 80:

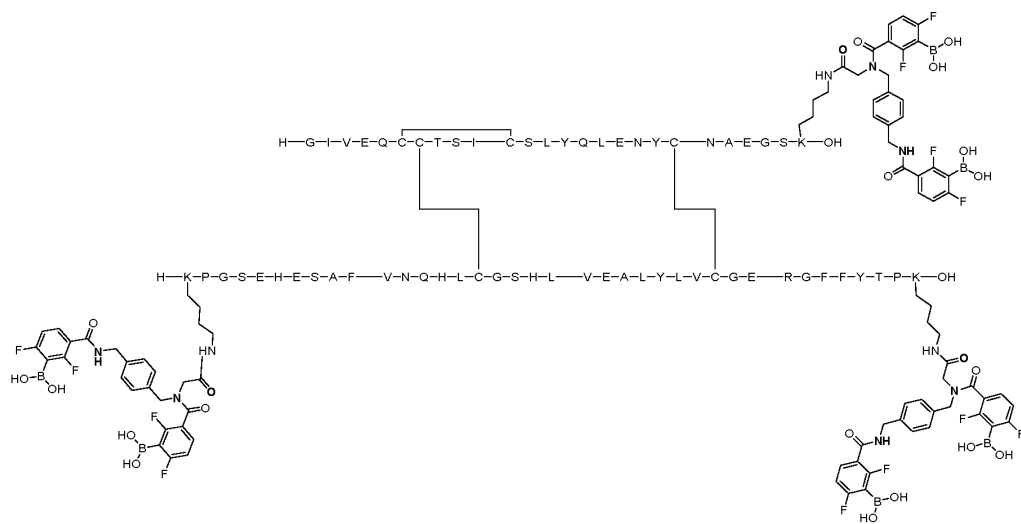


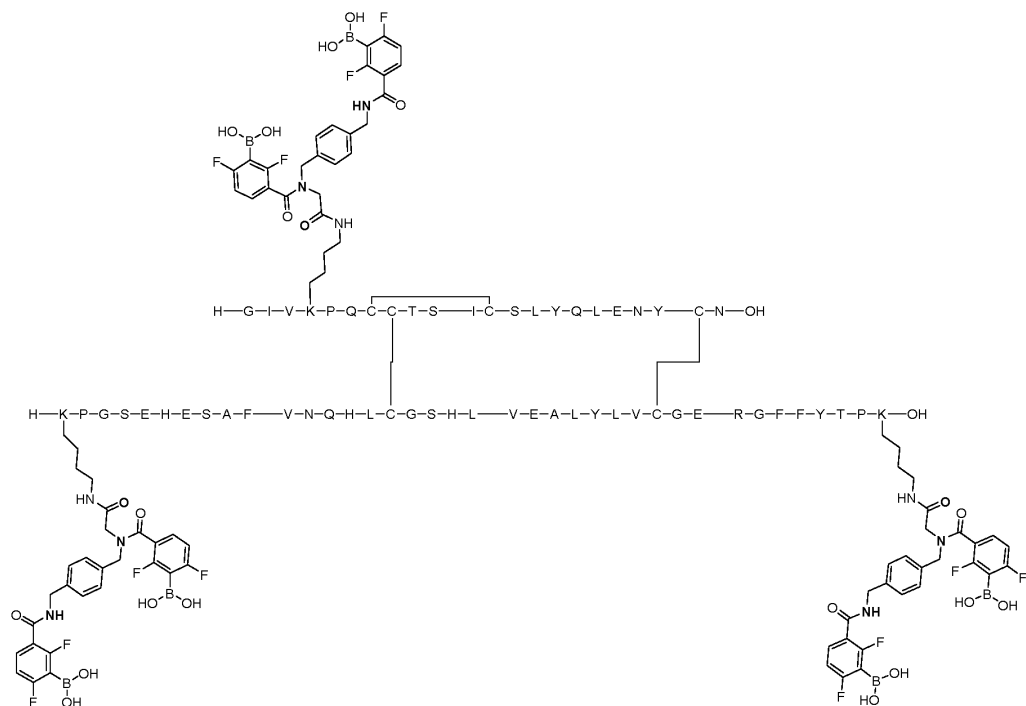


Chemical structure of the protein sequence: H-G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-A-E-G-S-K-OH. The structure shows the protein backbone and side chains, including fluorinated aromatic rings and boronic acid groups.

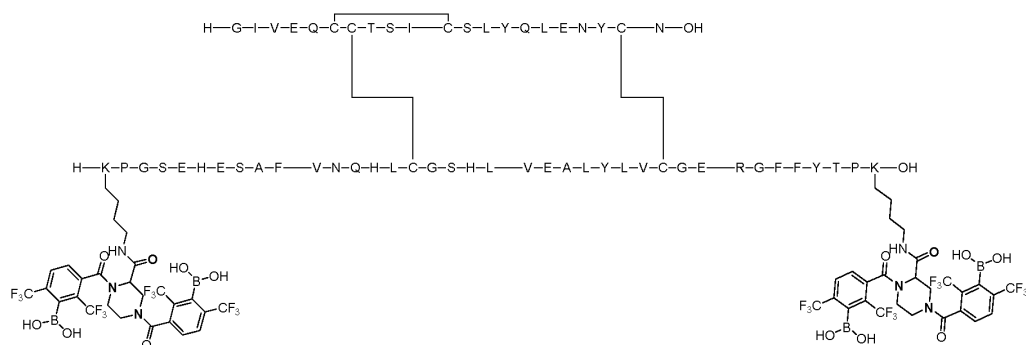


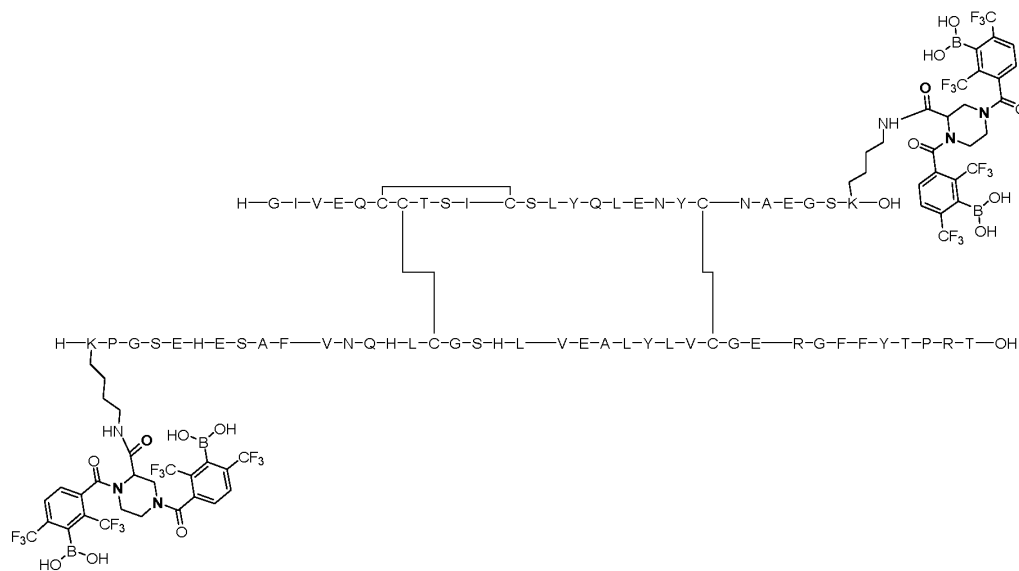
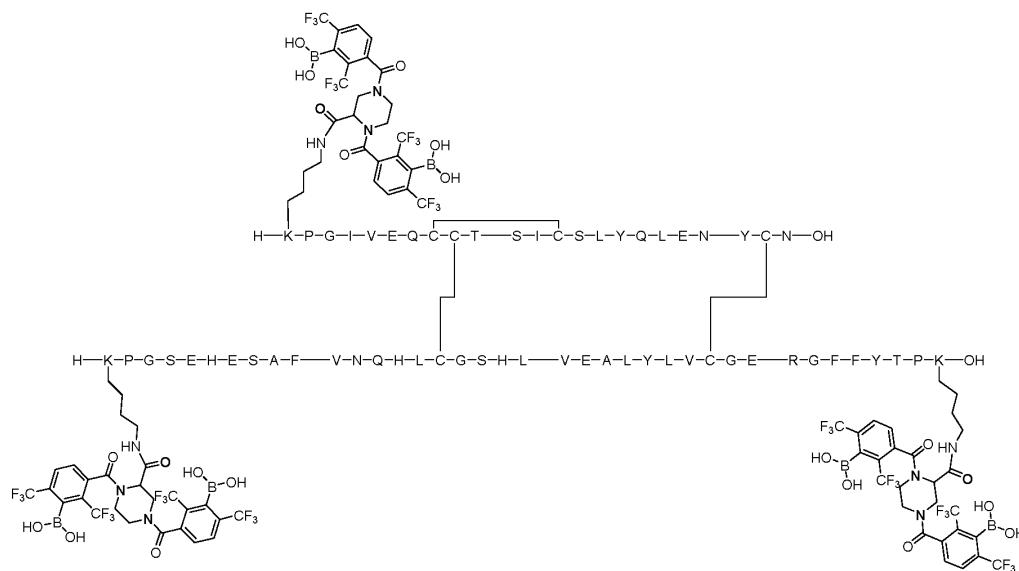
Example 84:

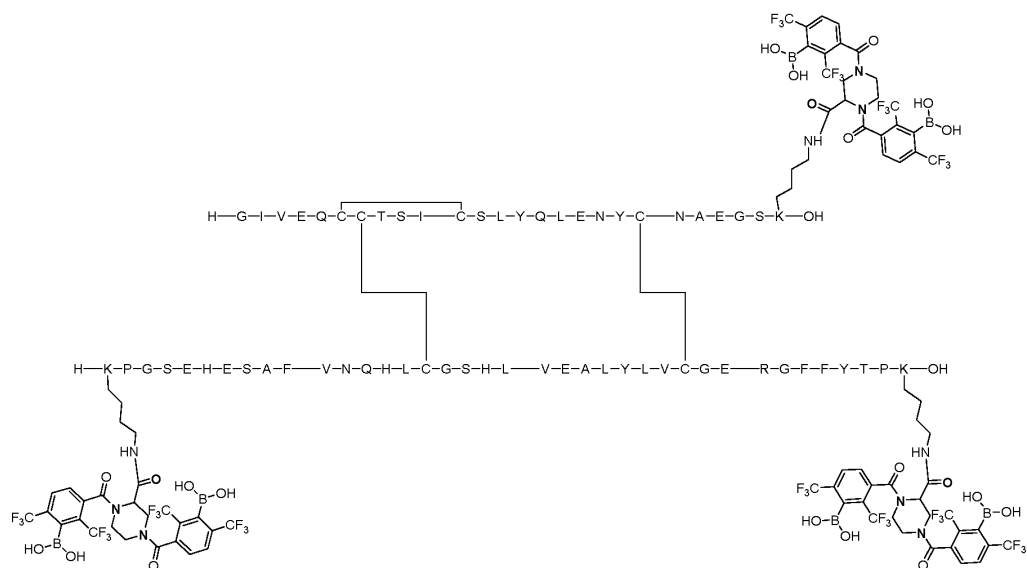
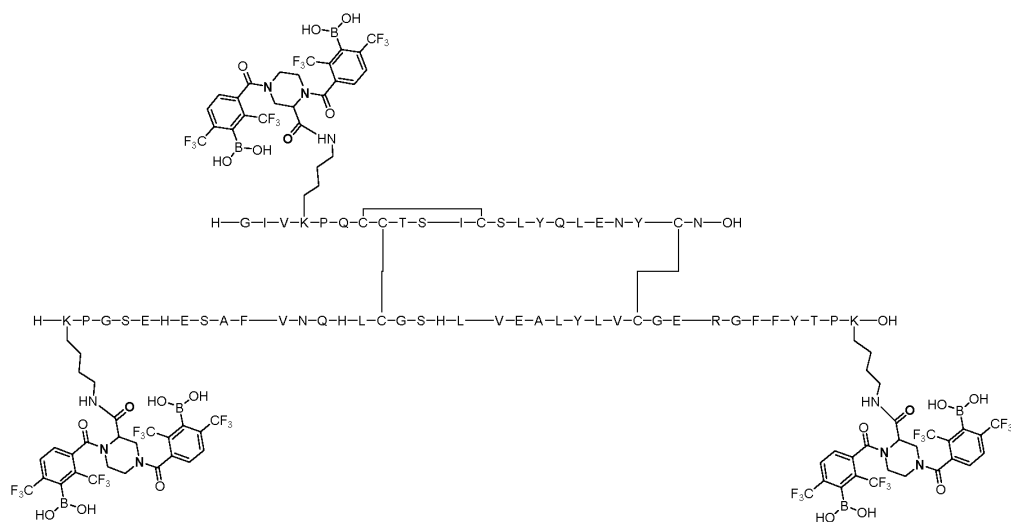


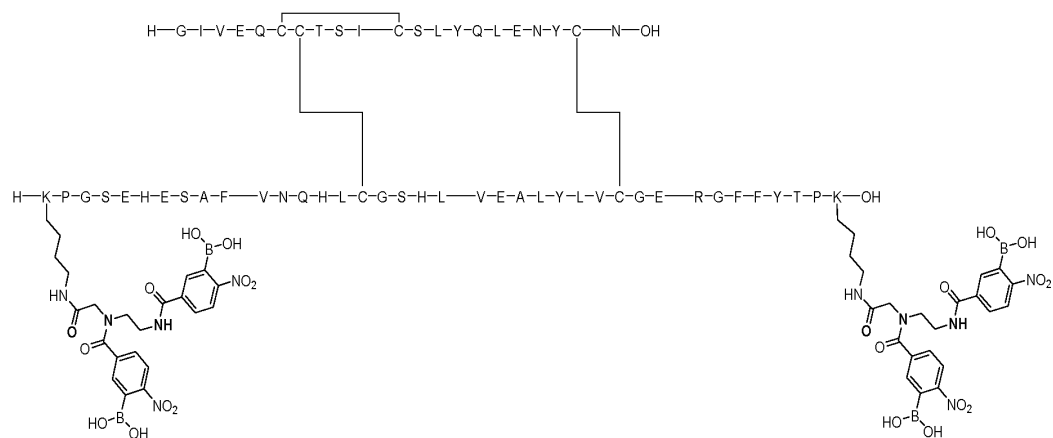
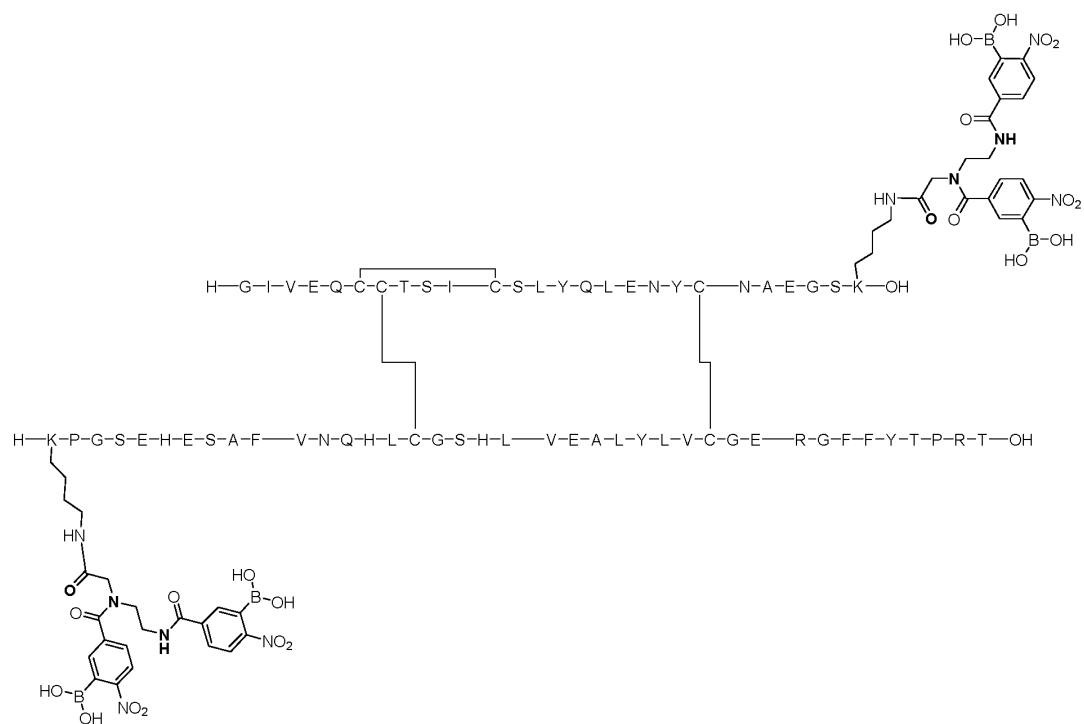


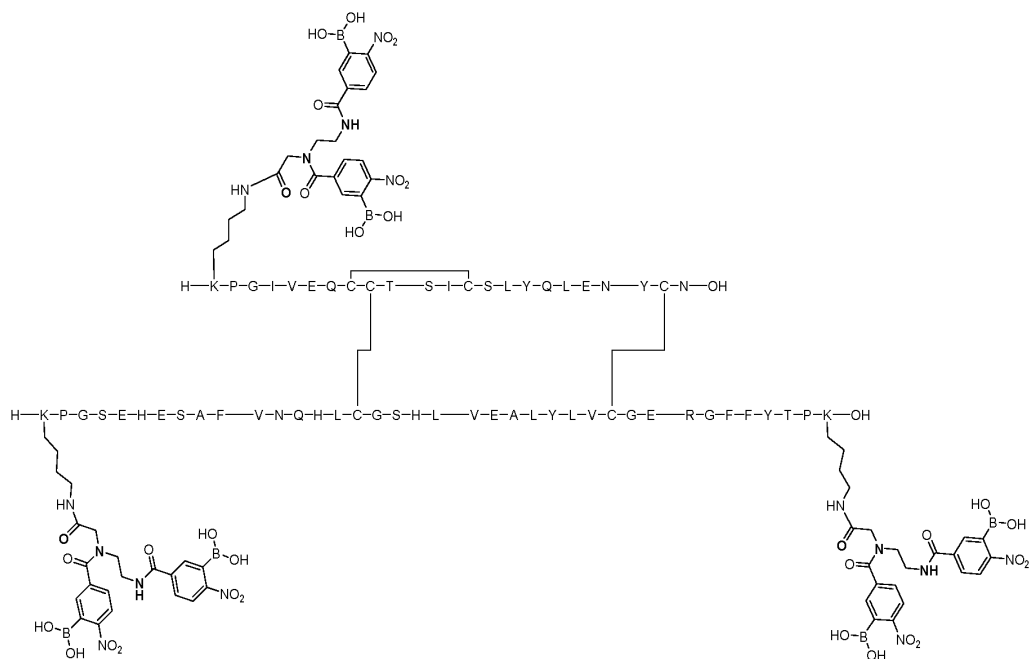
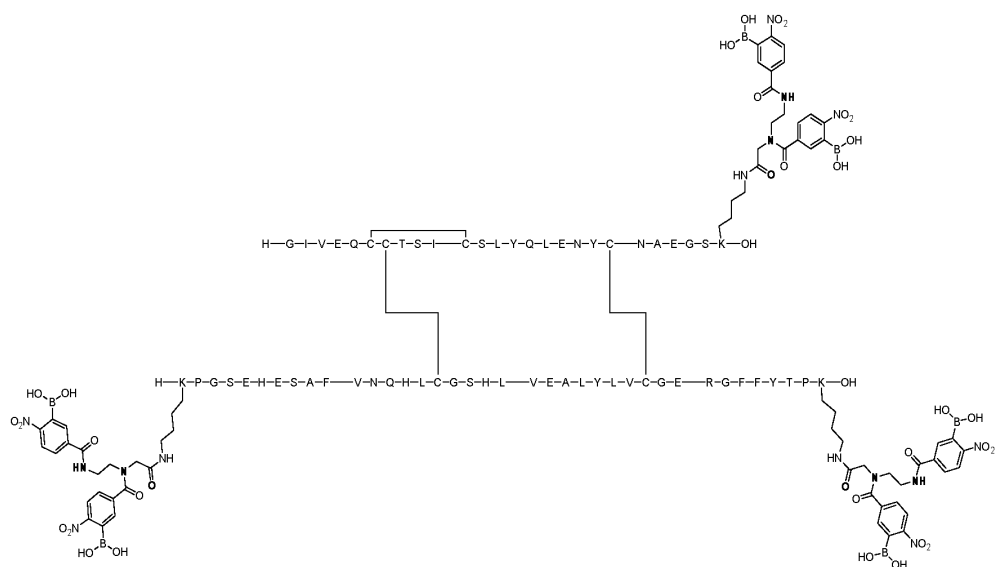
5

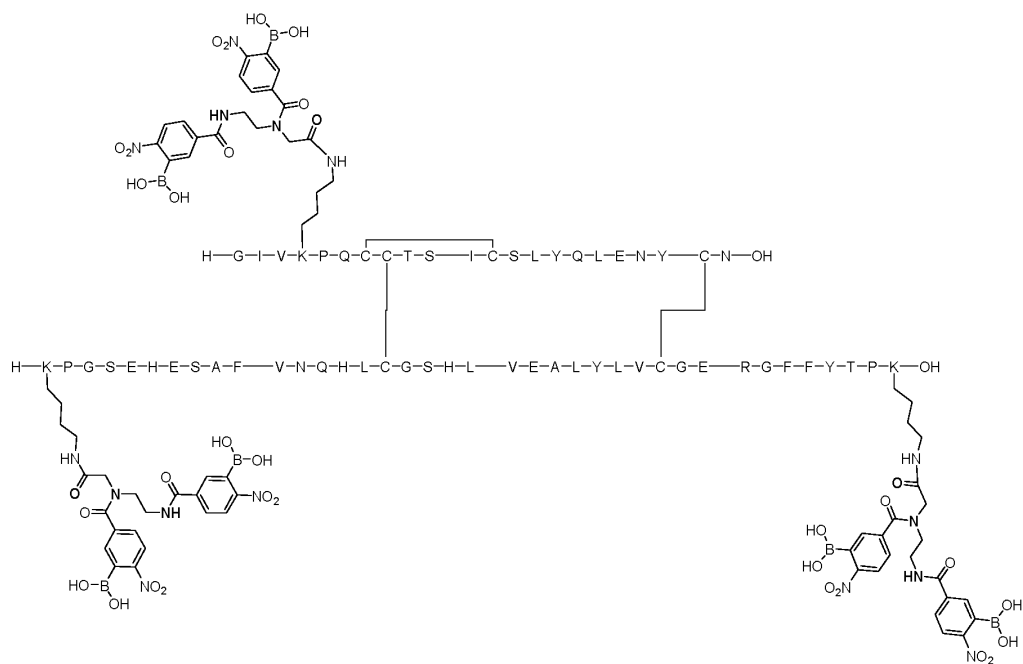
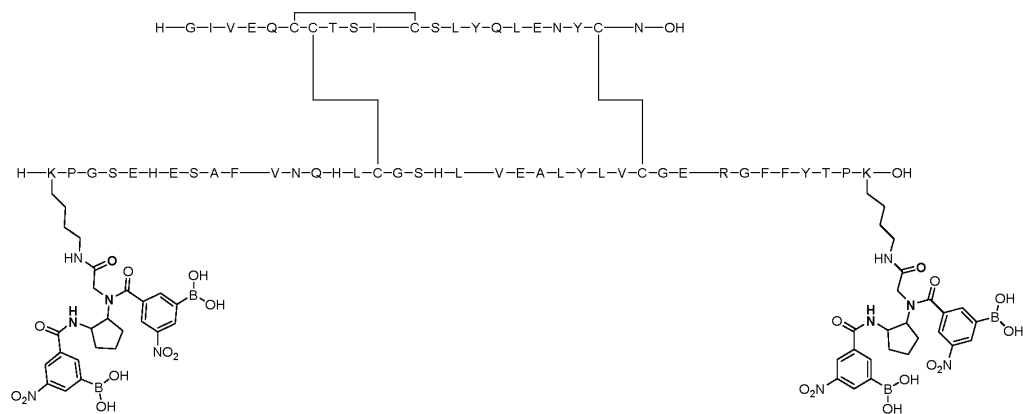


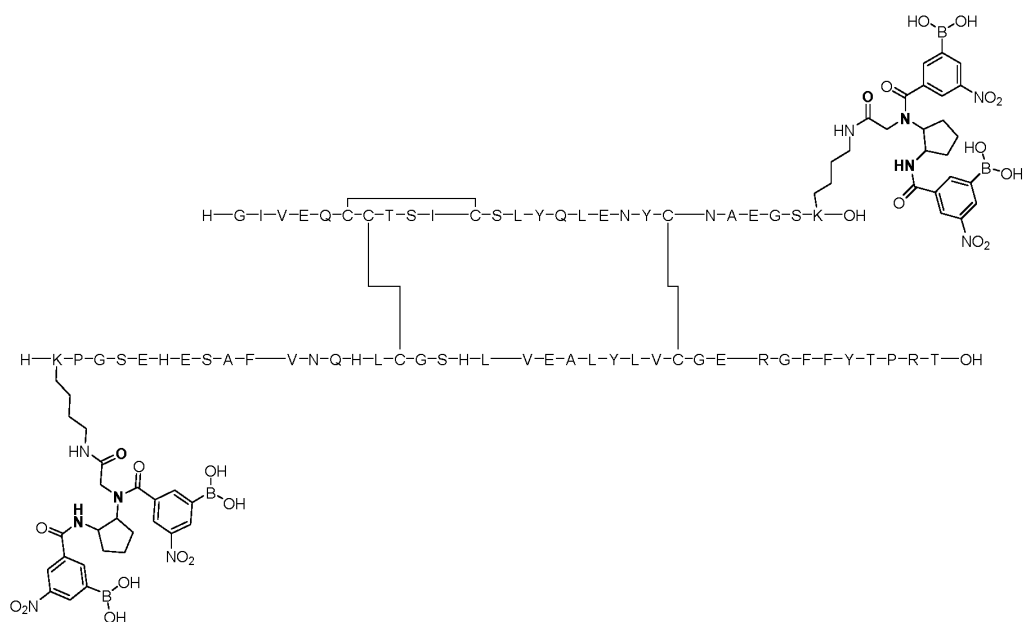
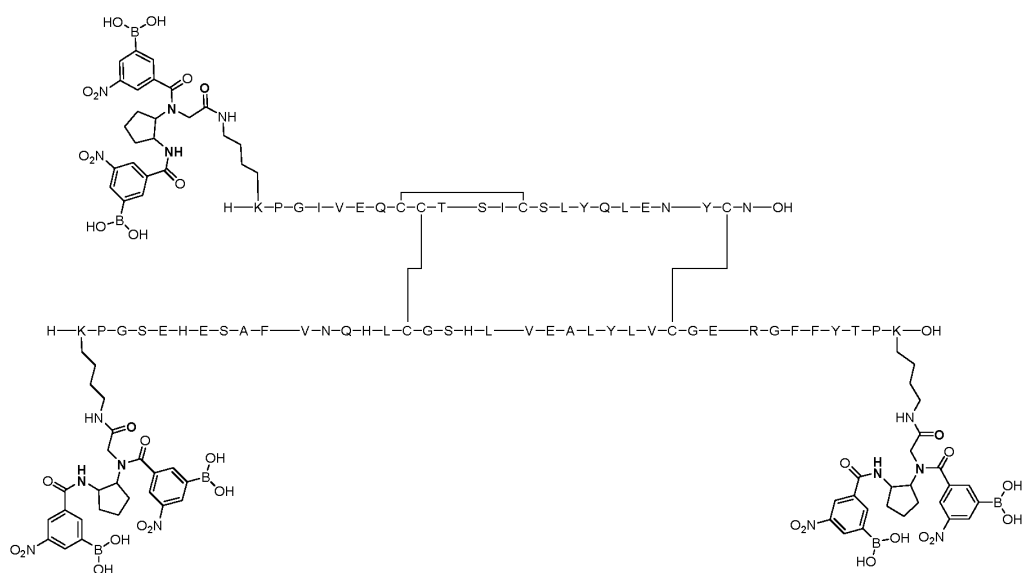
Example 87:**Example 88:**

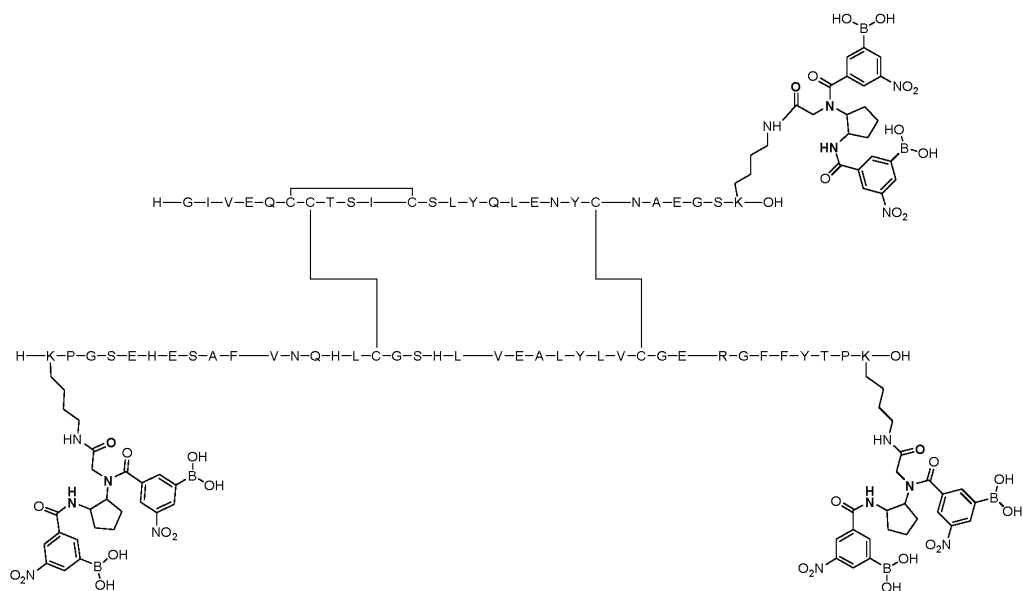
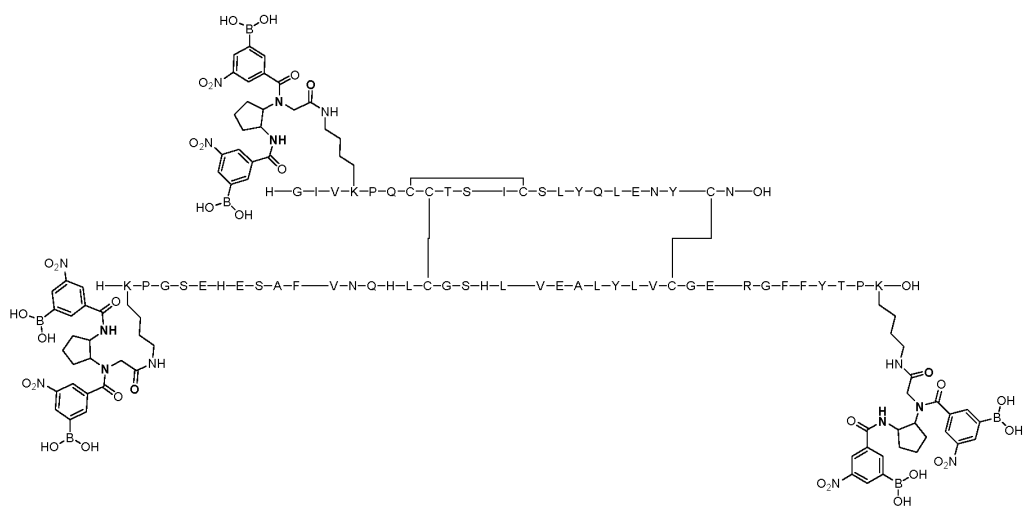
Example 89:**Example 90:**

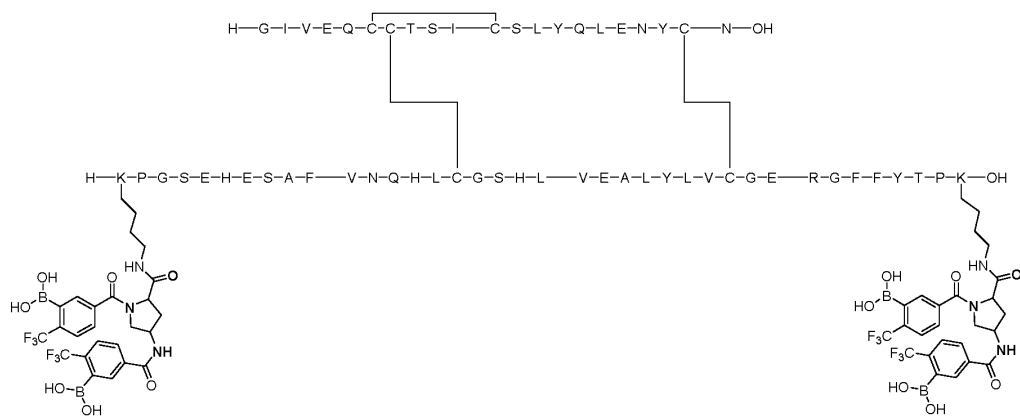
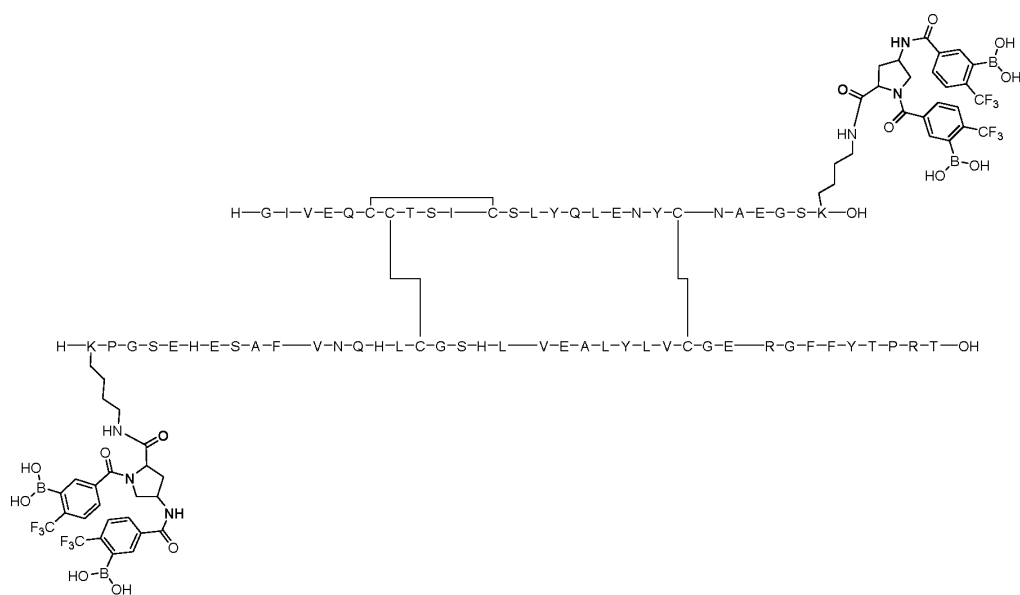
Example 91:**Example 92:**

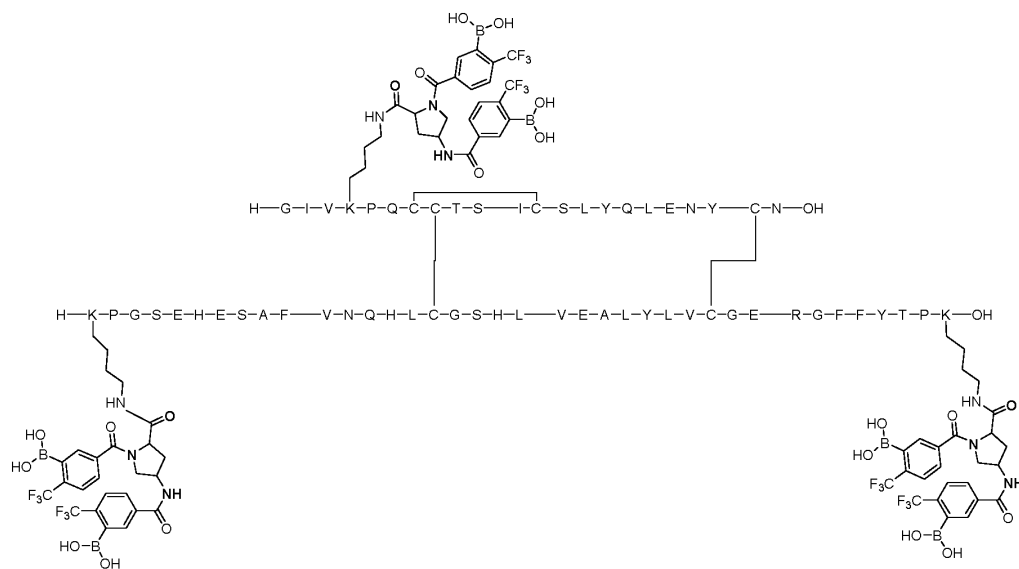
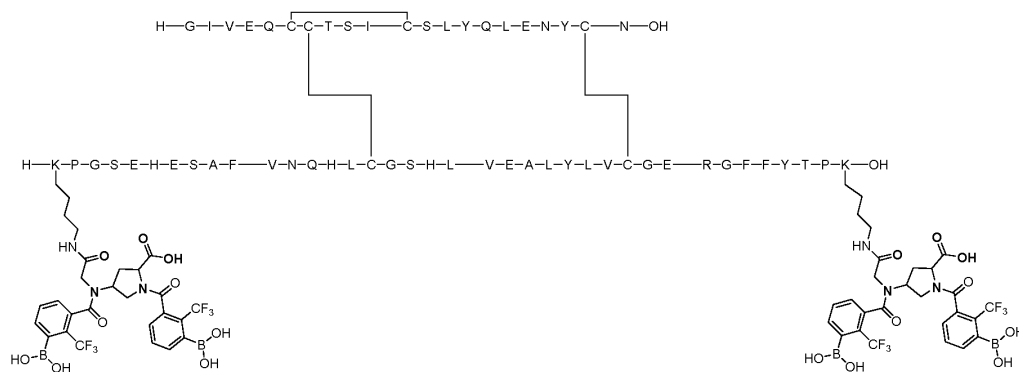
Example 93:**Example 94:**

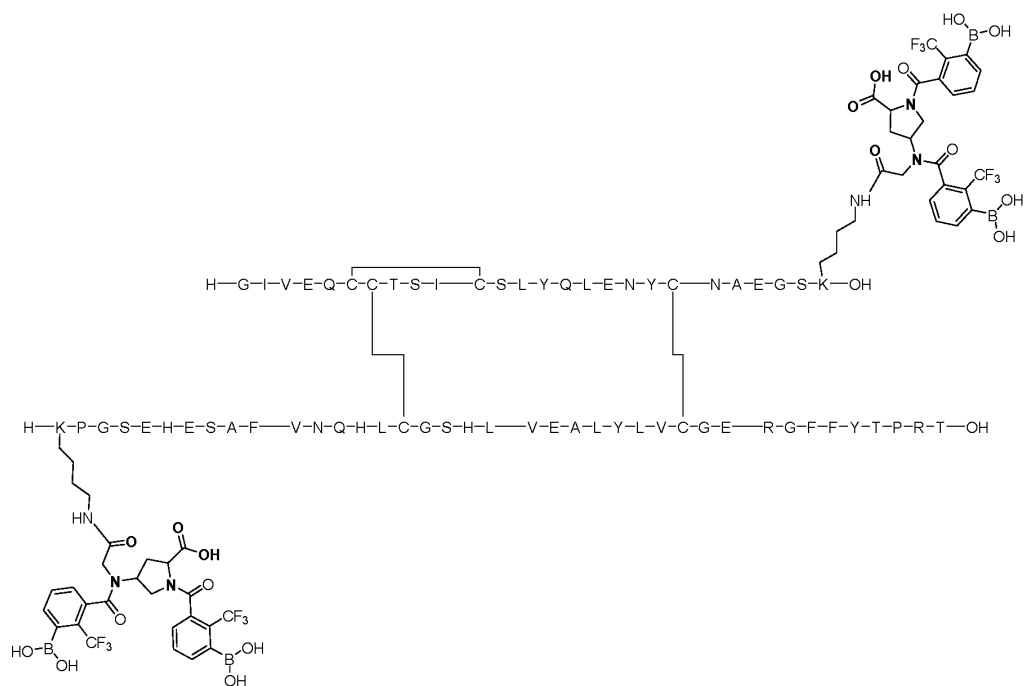
Example 95:**Example 96:**

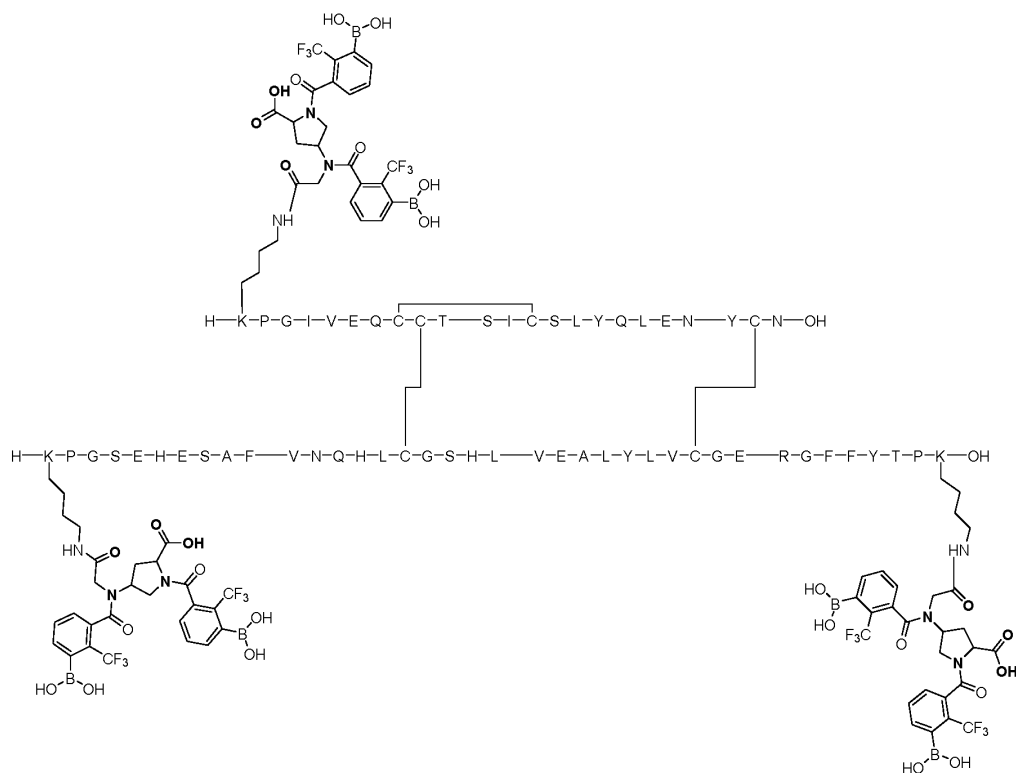
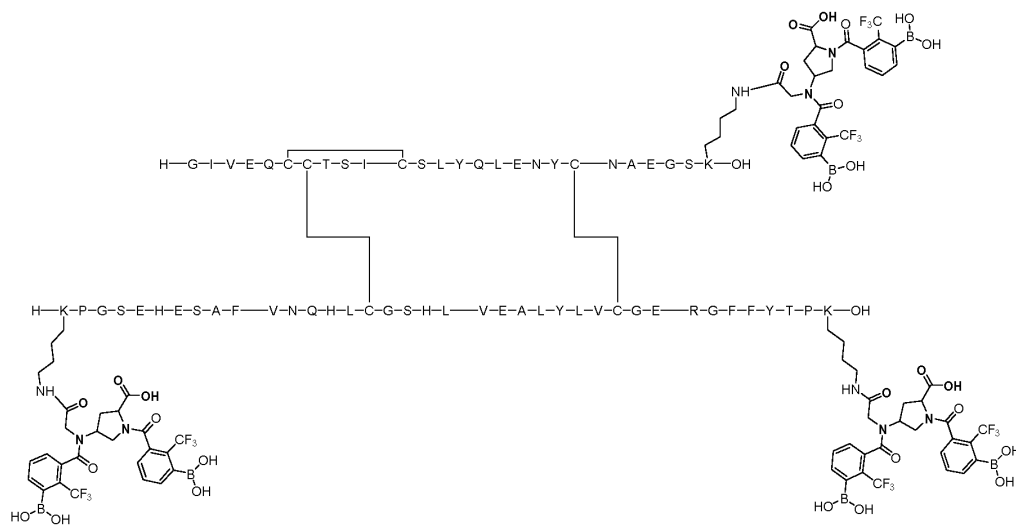
Example 97:**Example 98:**

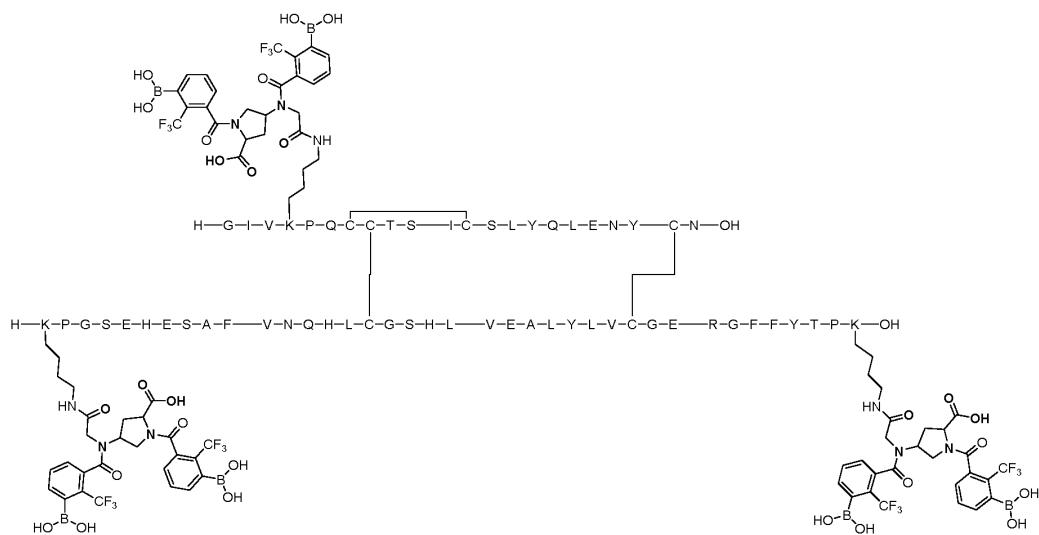
Example 99:**Example 100:**

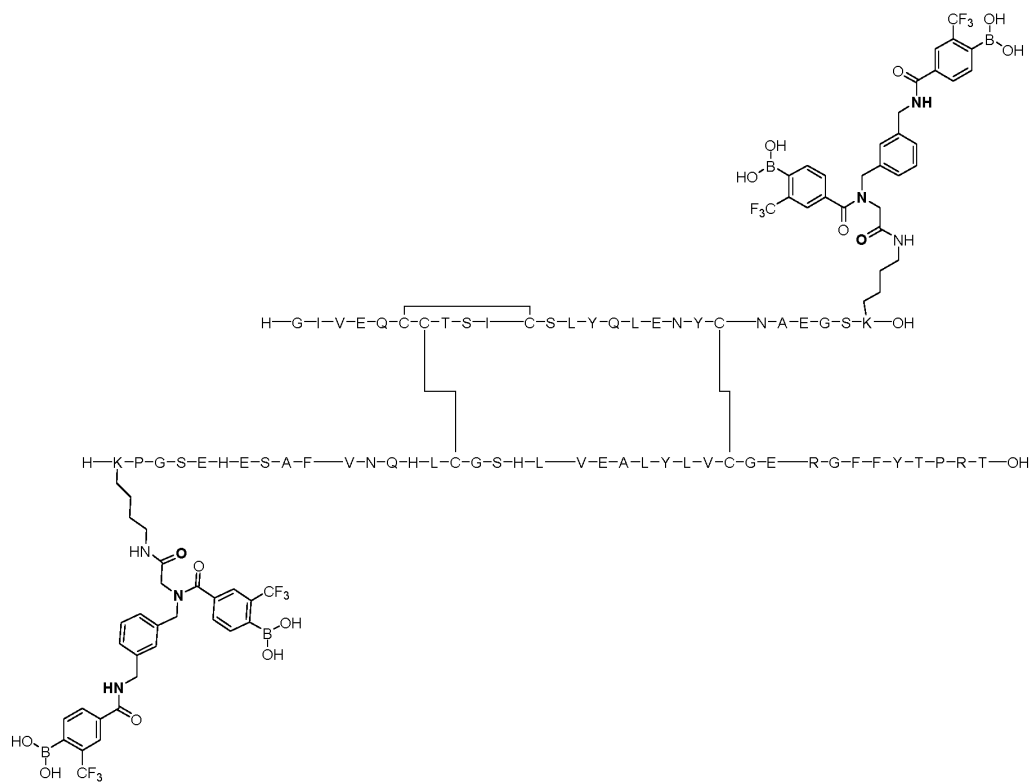
Example 101:**Example 102:**

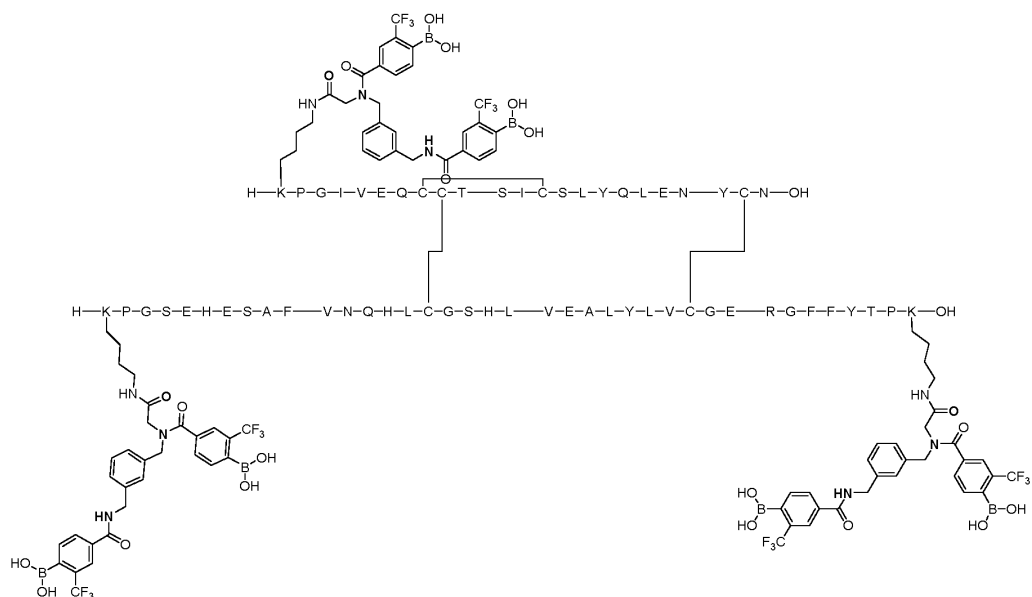
Example 103:**Example 104:**

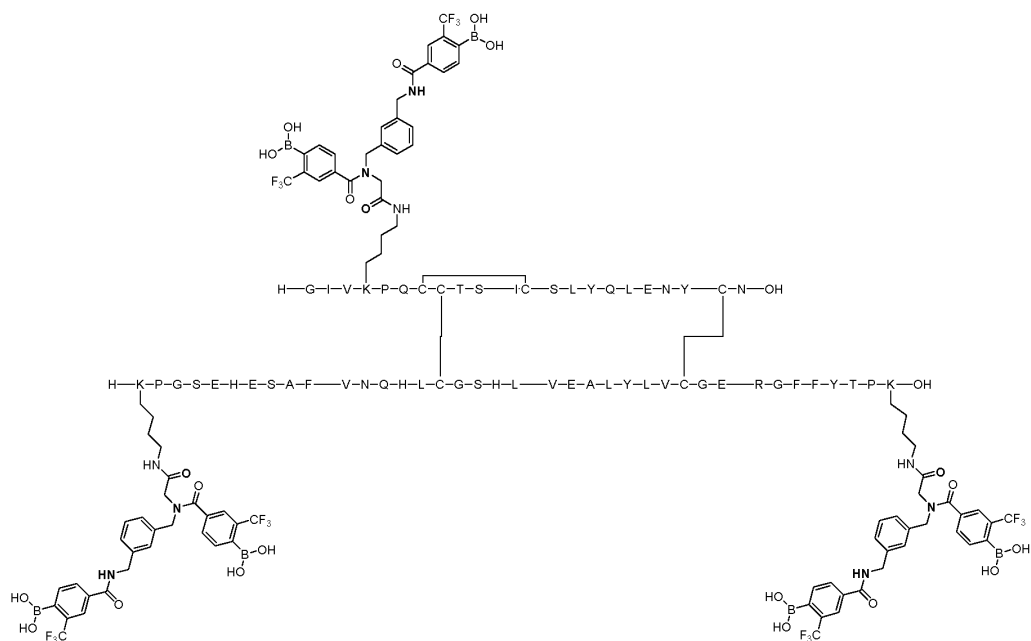
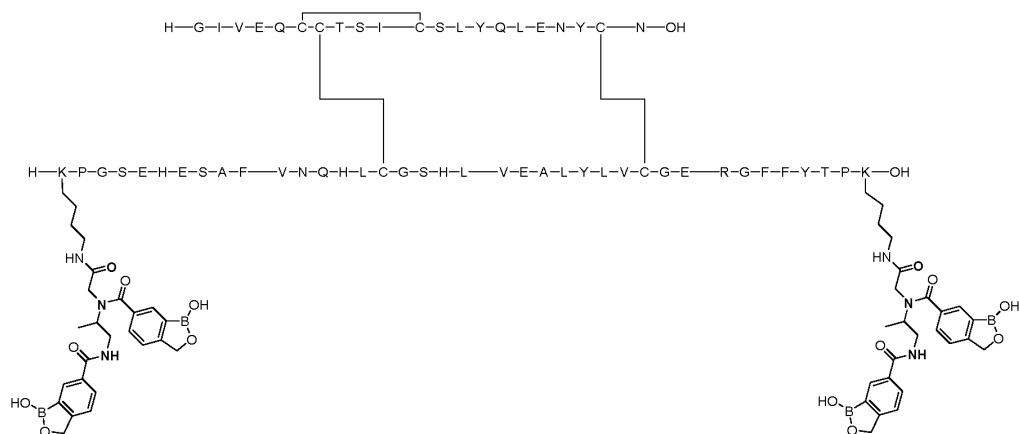
Example 105:

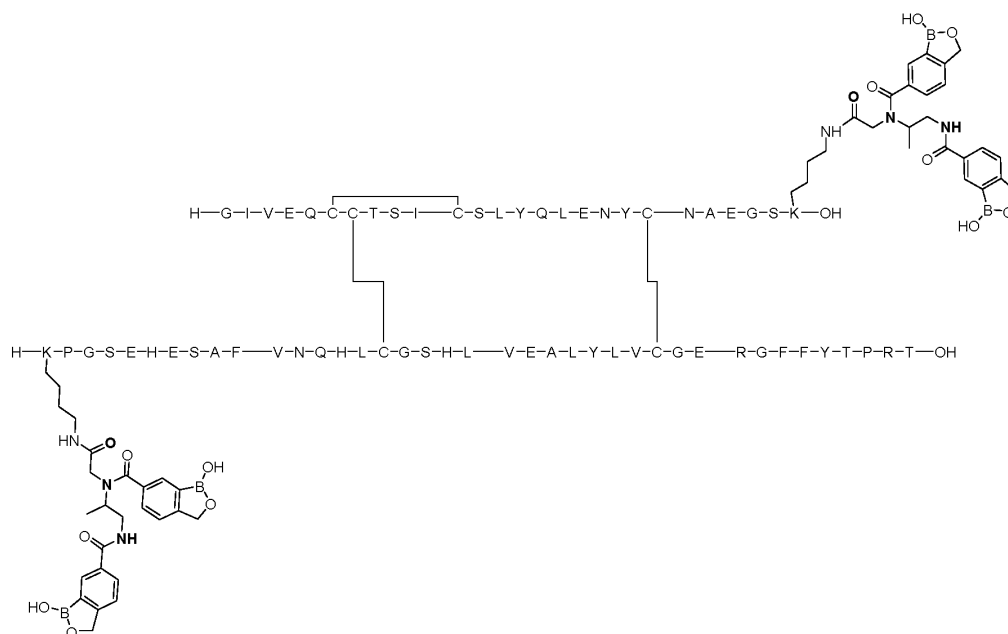
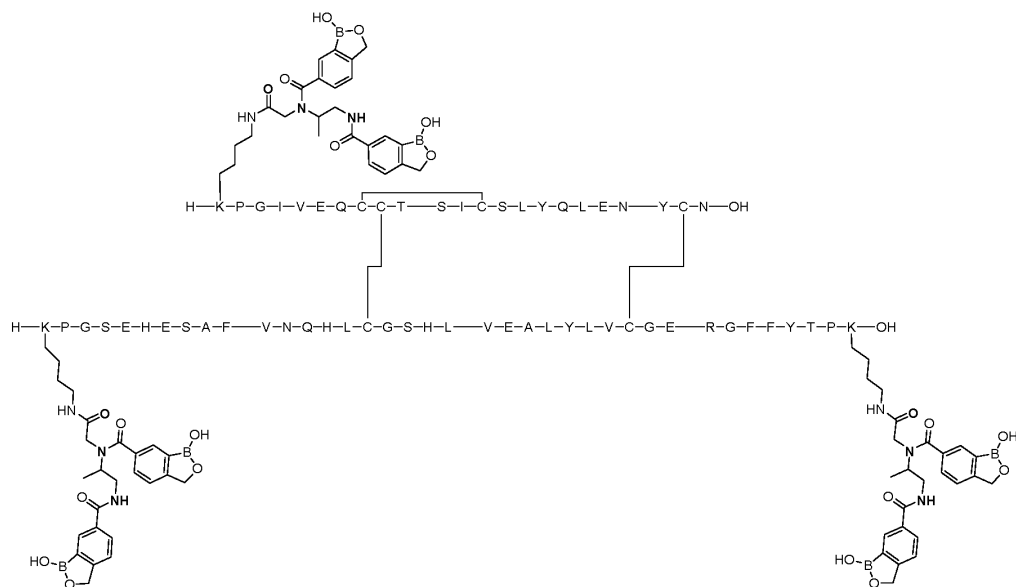
Example 106:**Example 107:**

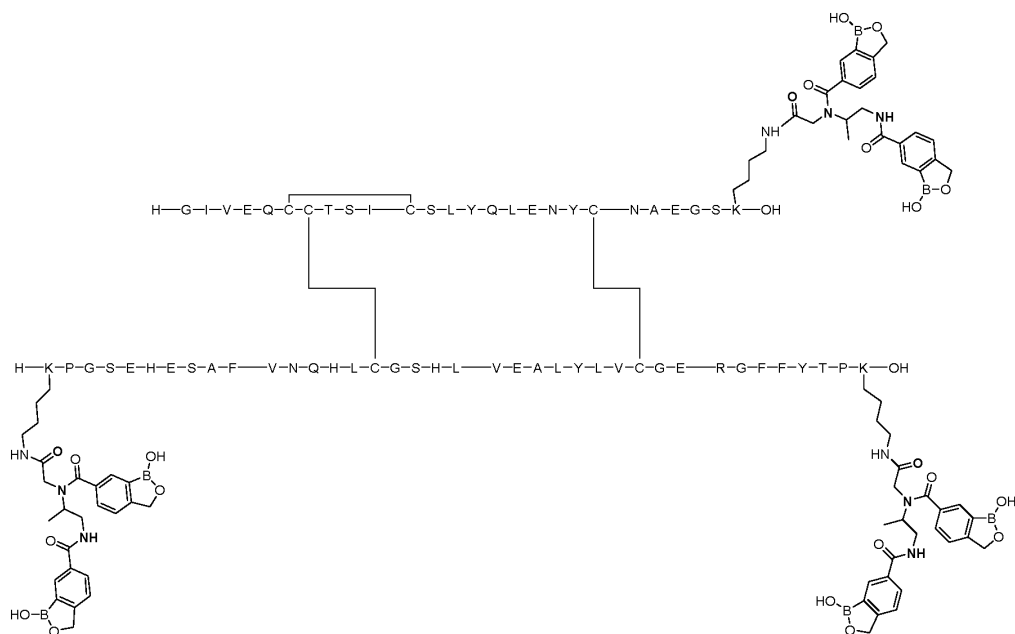
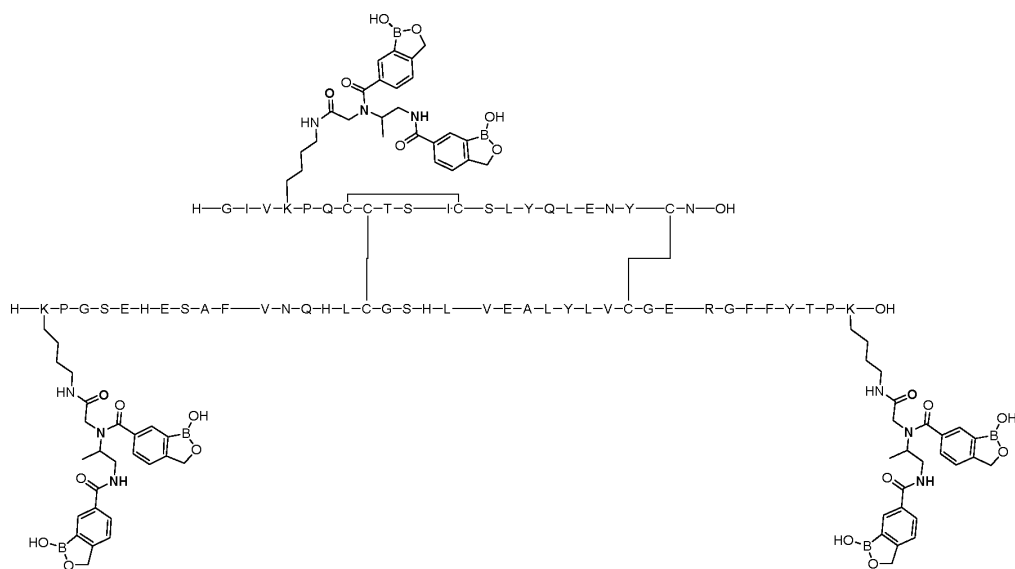
[illegible]

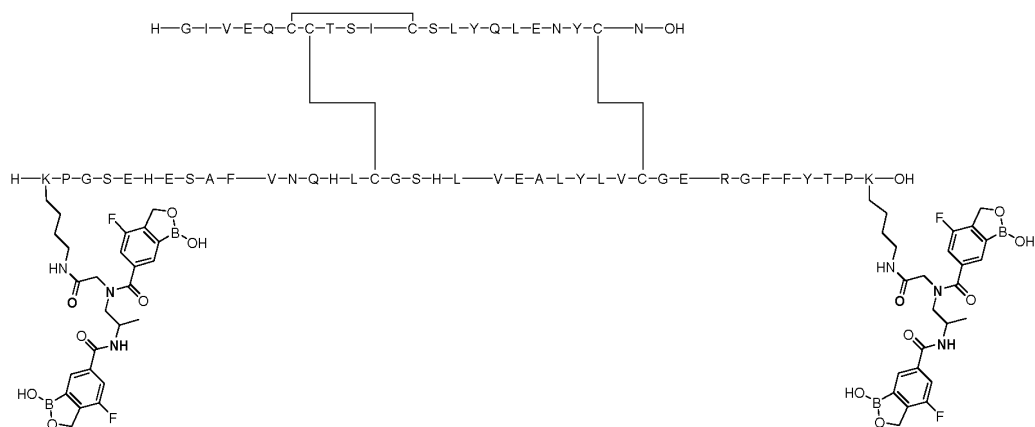
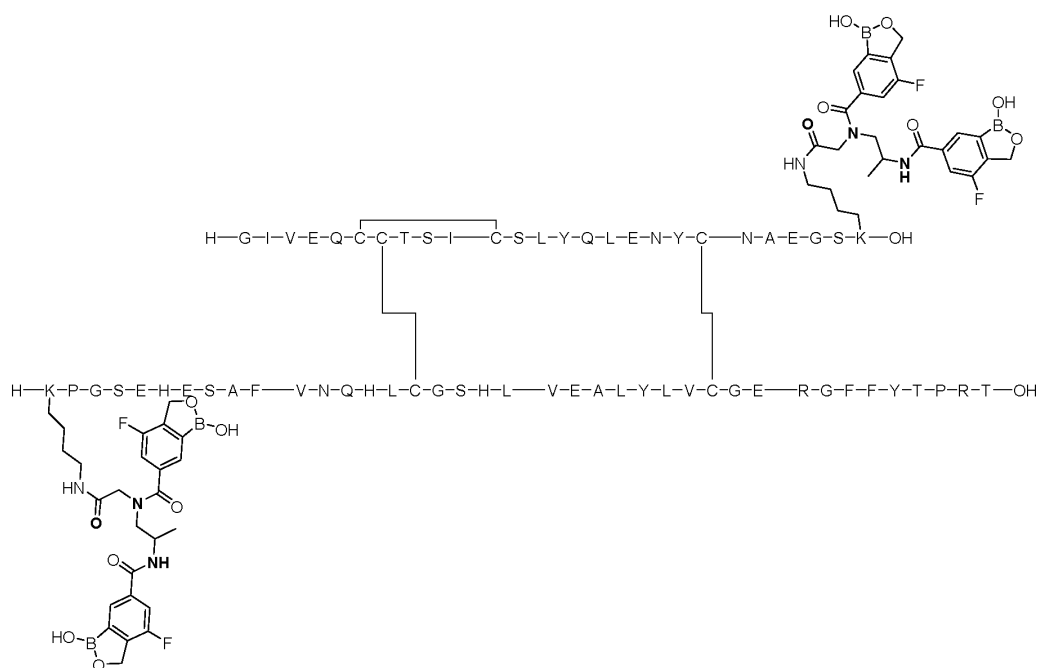


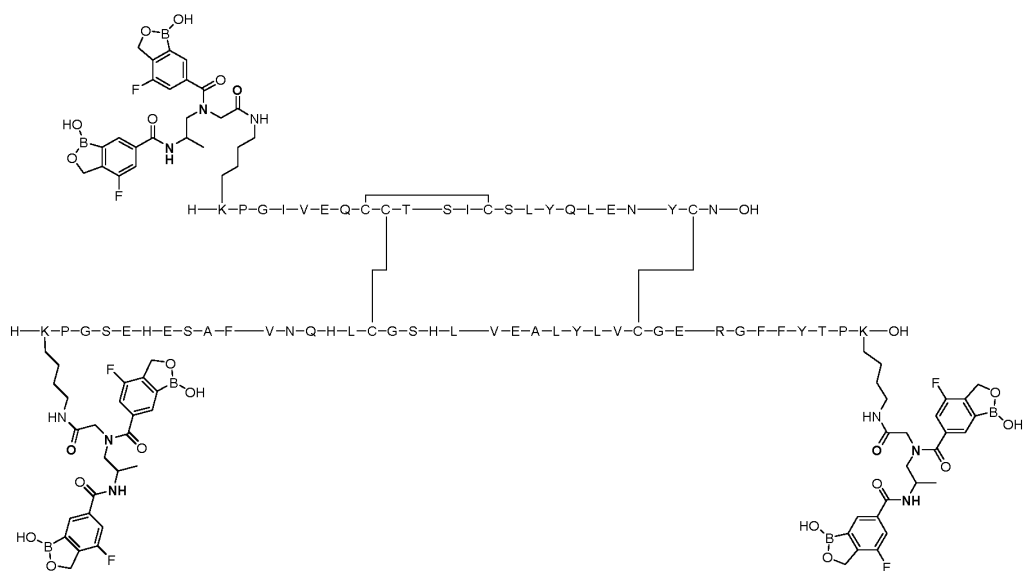
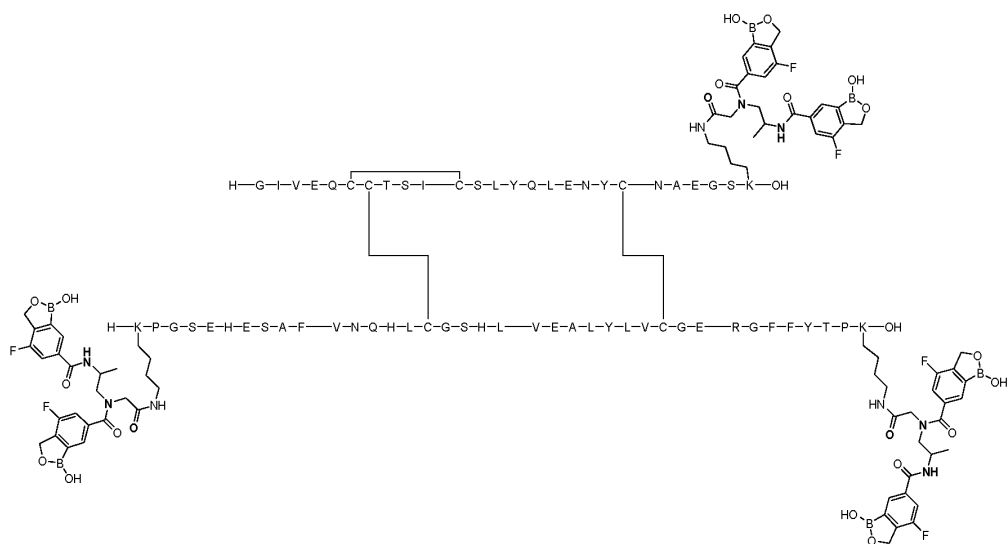


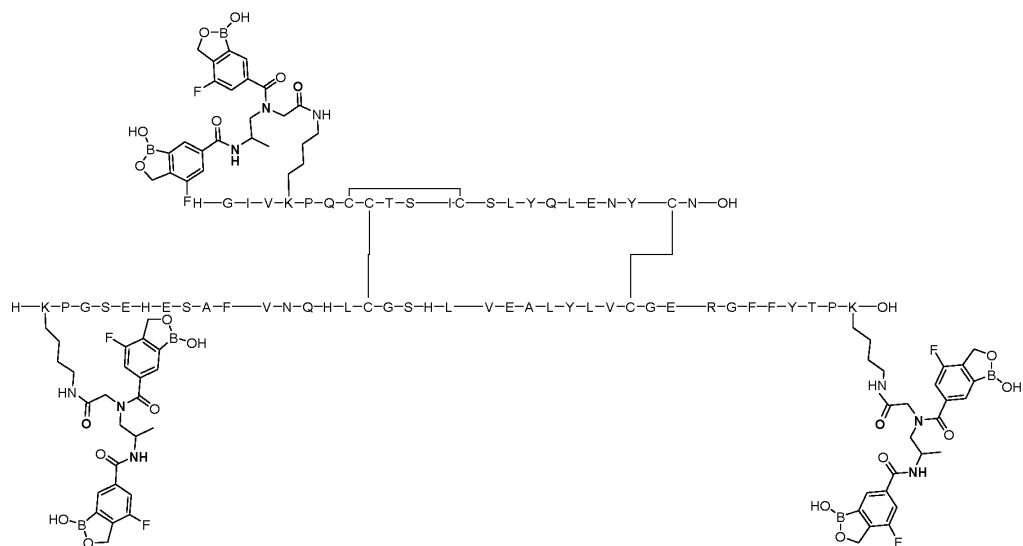
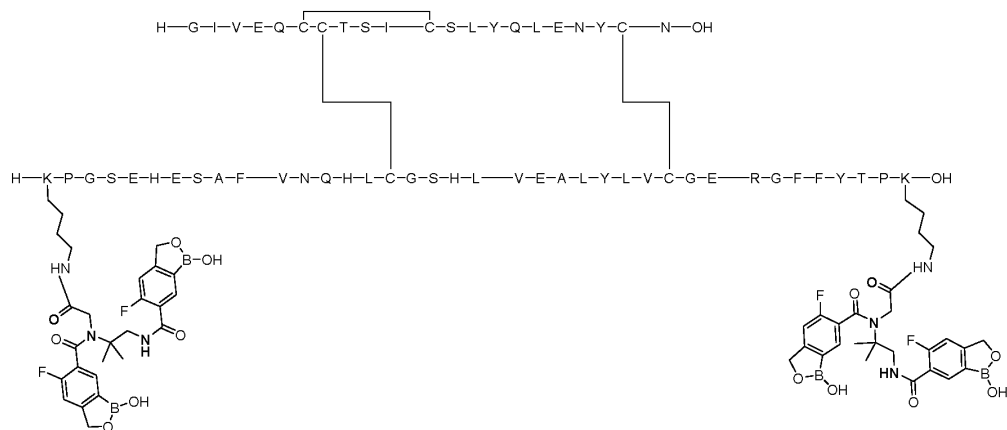
Example 113:**Example 114:**

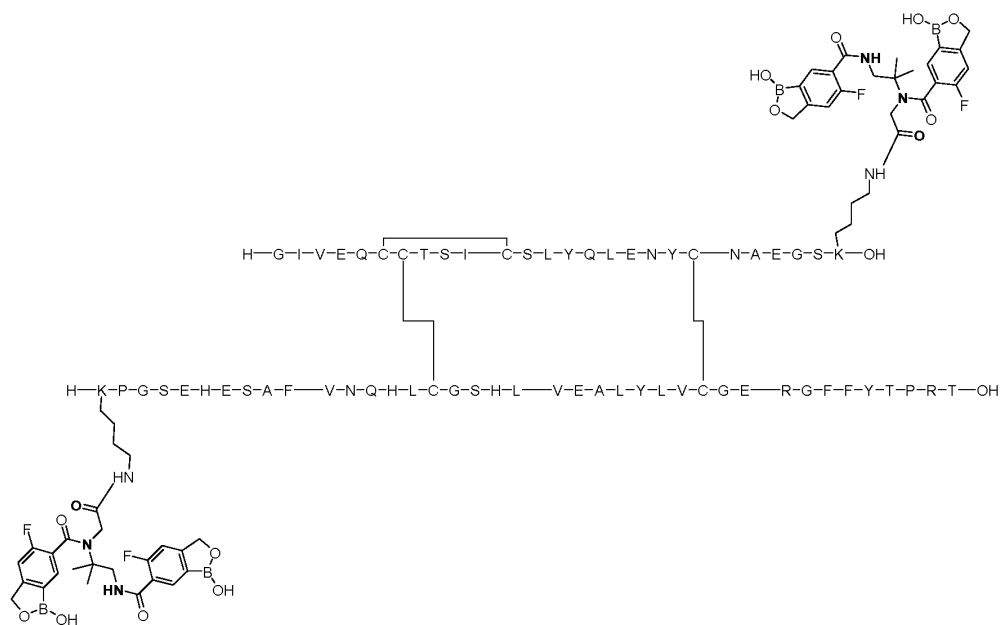
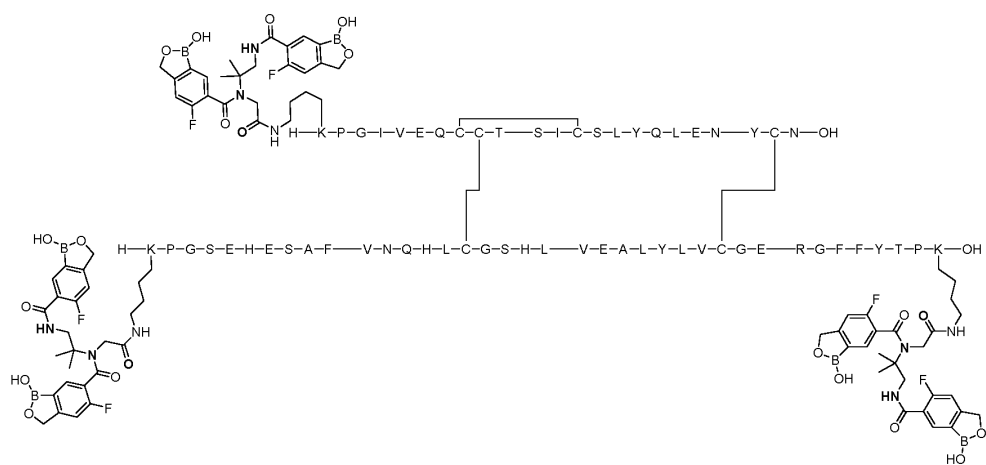
Example 115:**Example 116:**

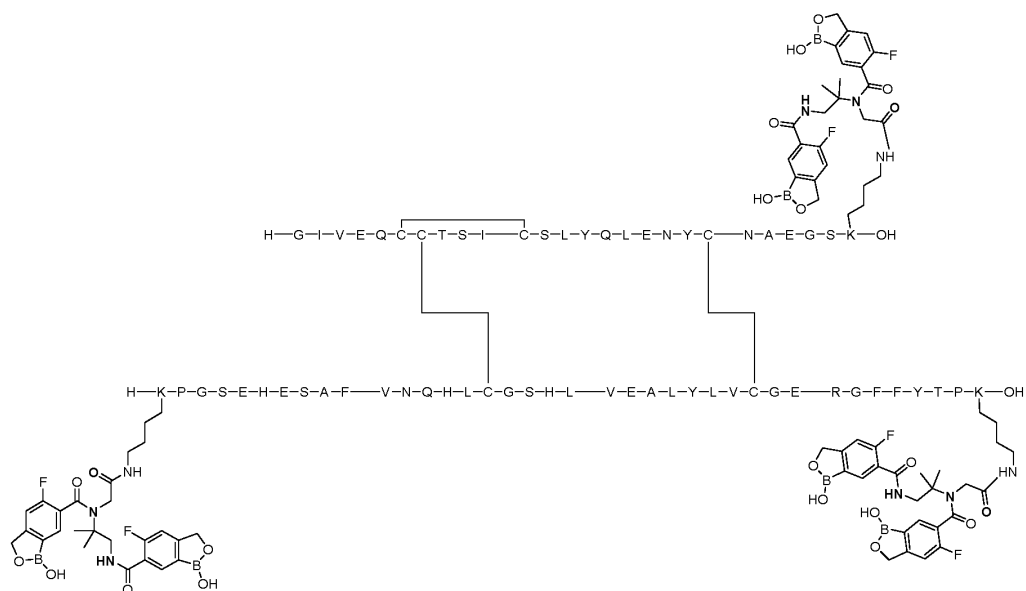
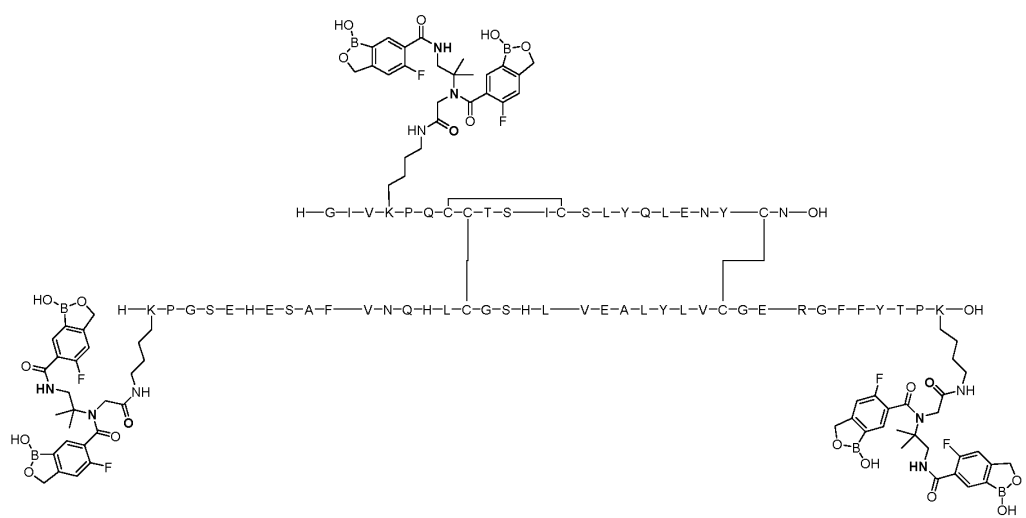
Example 117:**Example 118:**

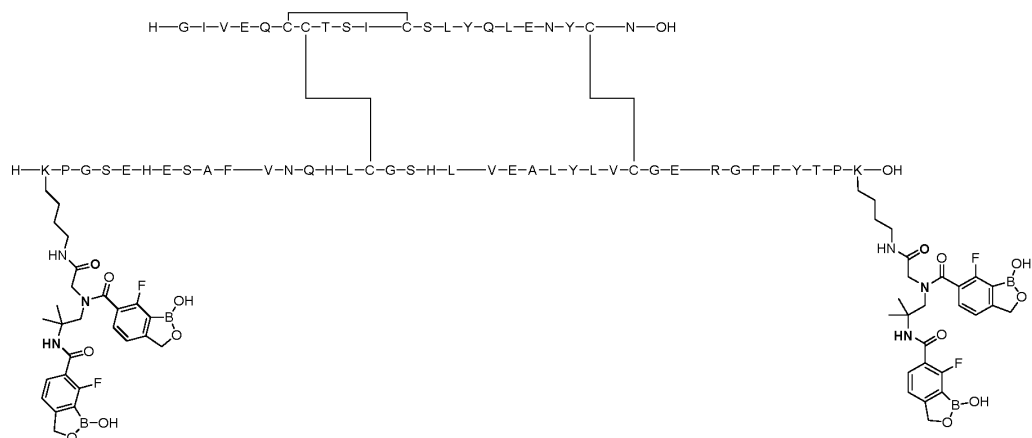
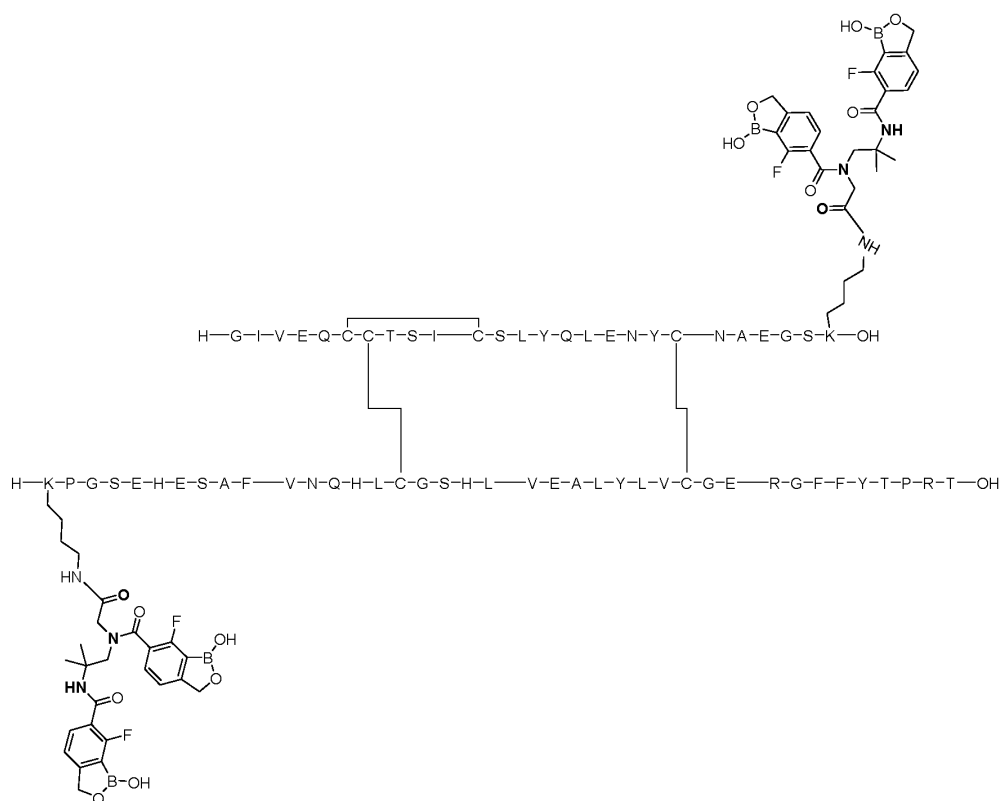
Example 119:**Example 120:**

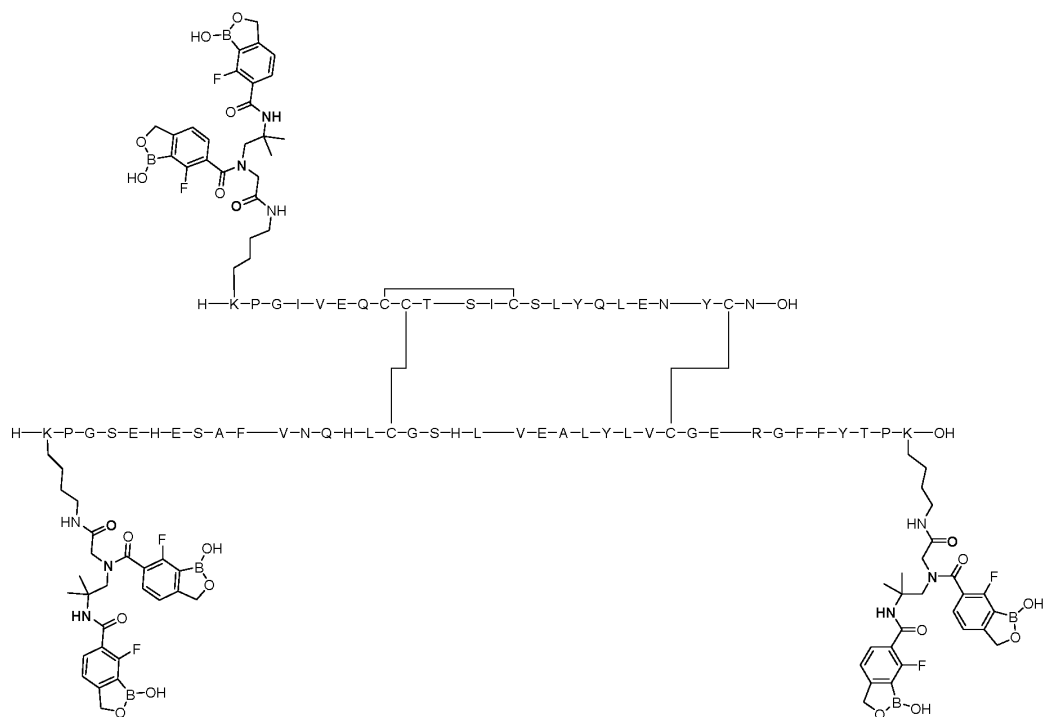
Example 121:**Example 122:**

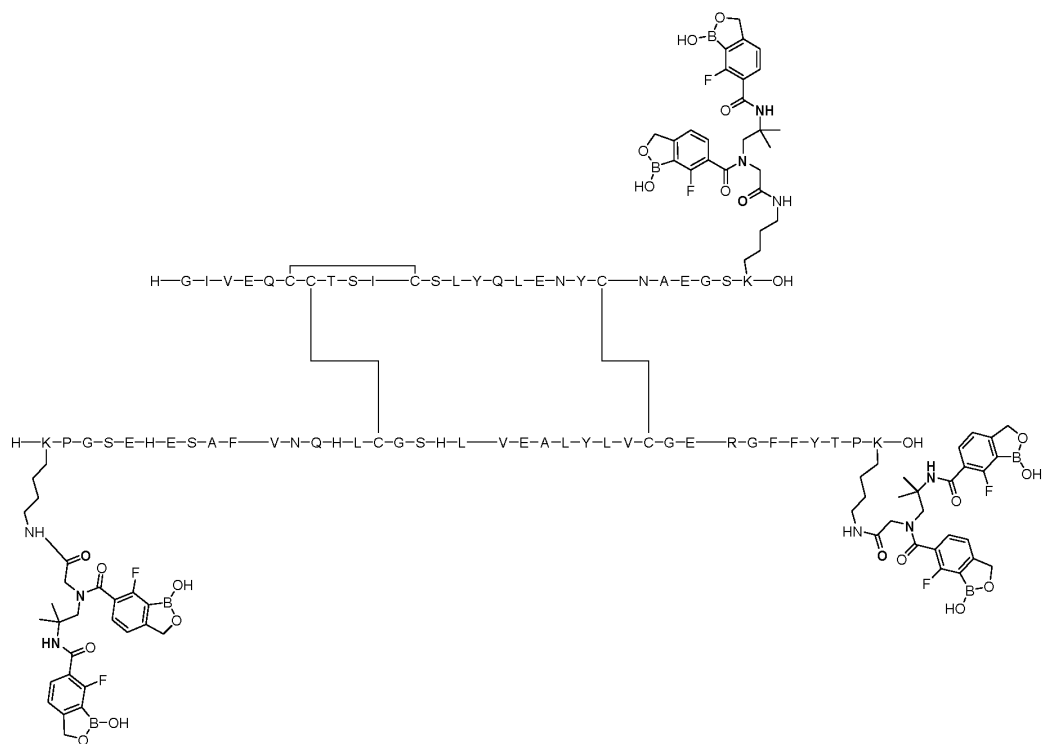
Example 123:**Example 124:**

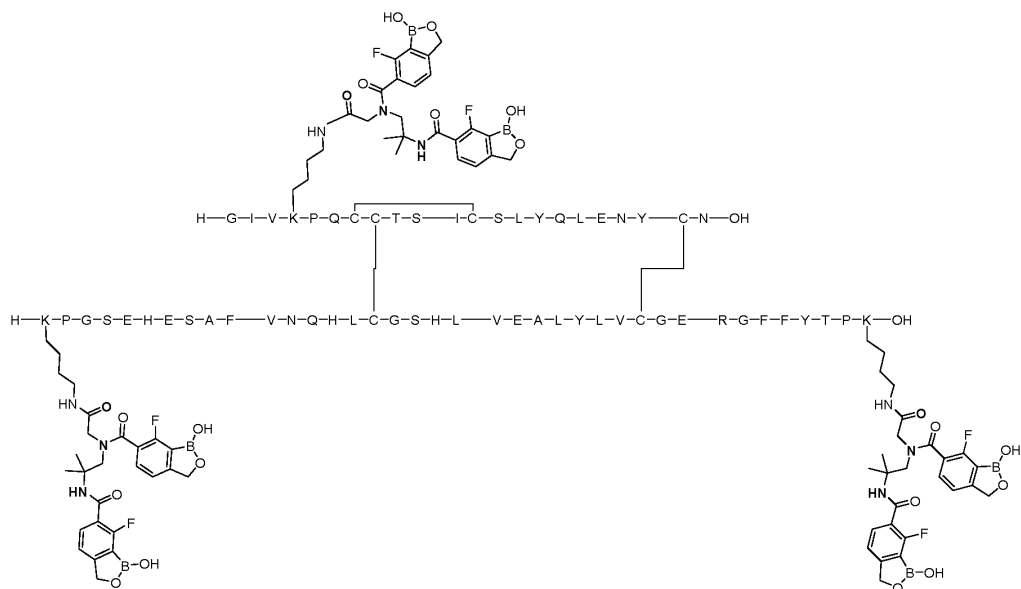
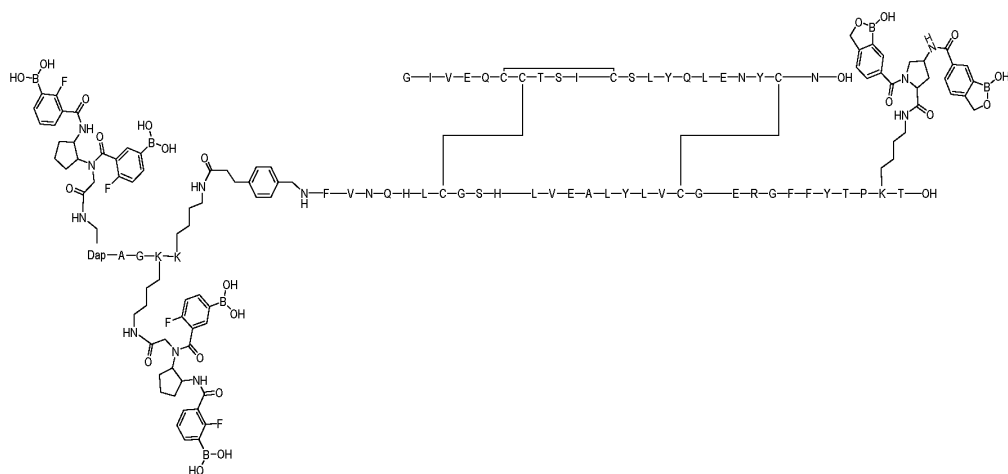
Example 125:**Example 126:**

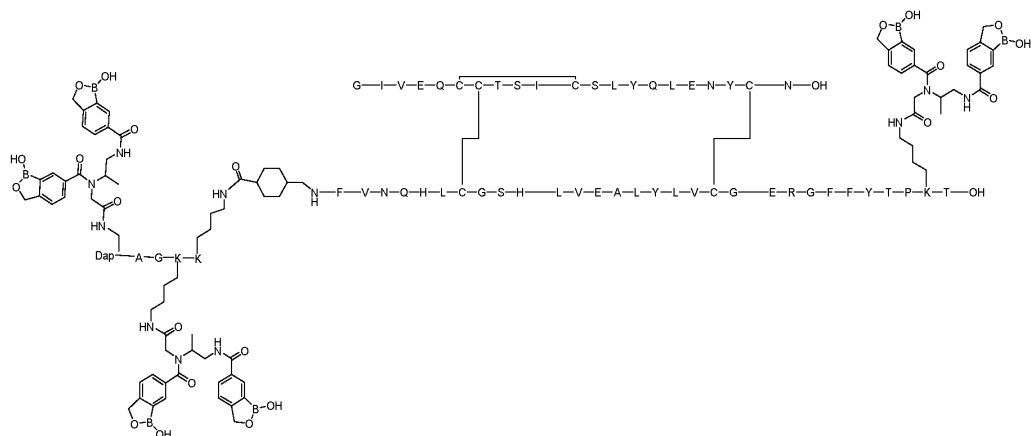
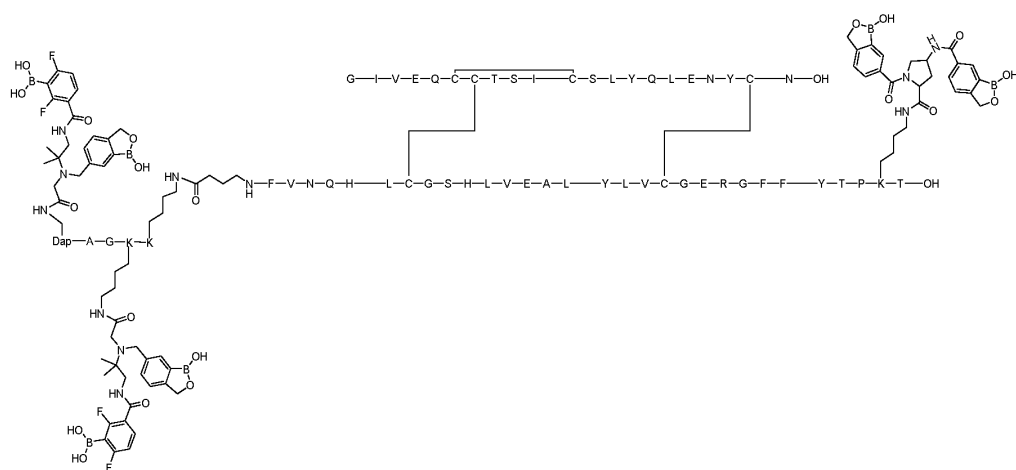
Example 127:**Example 128:**

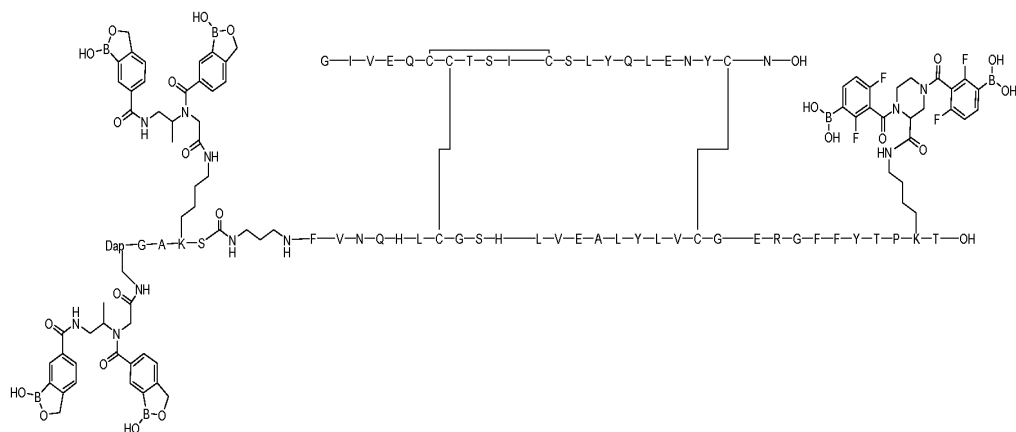
Example 129:**Example 130:**

Example 131:

Example 132:

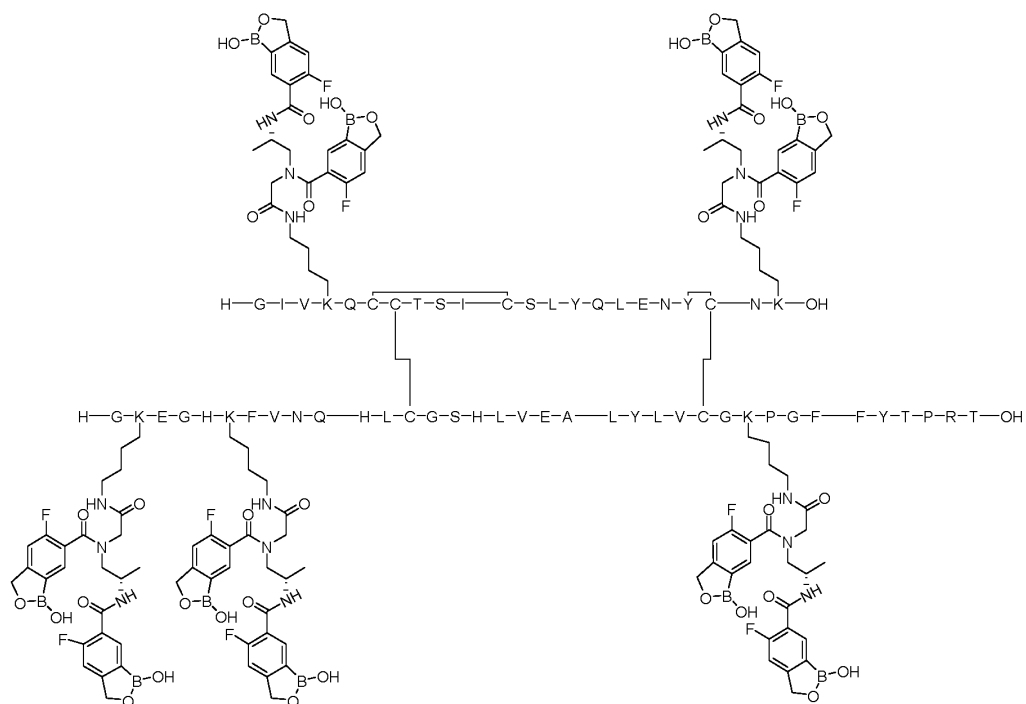
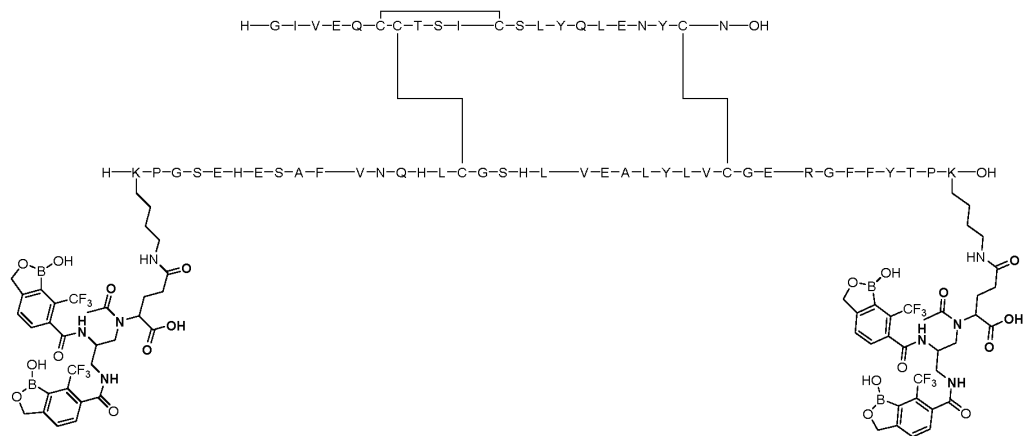
Example 133:**Example 134:**

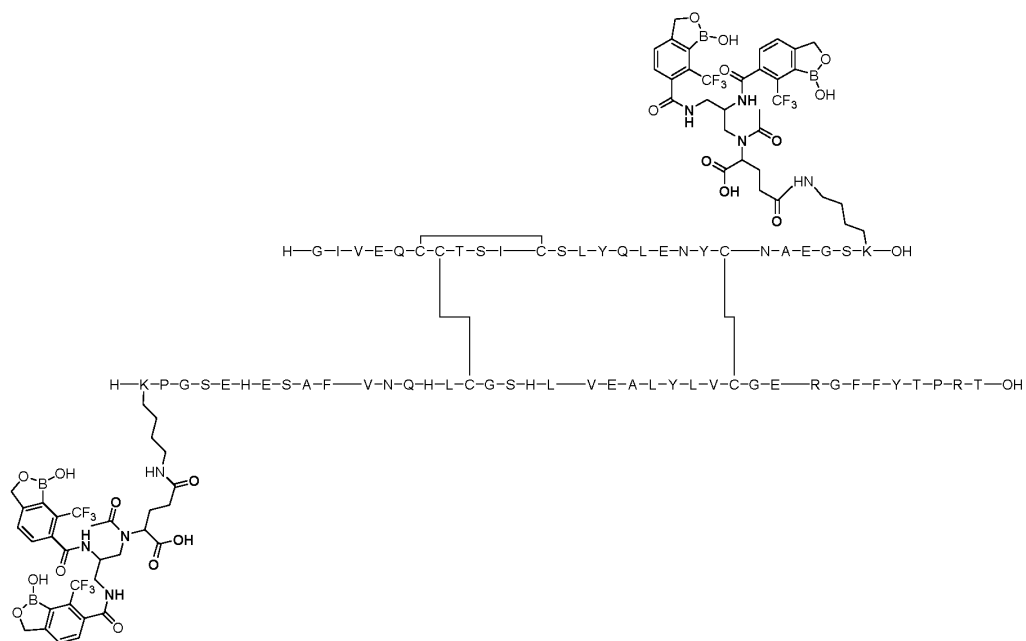
**Example 135:**



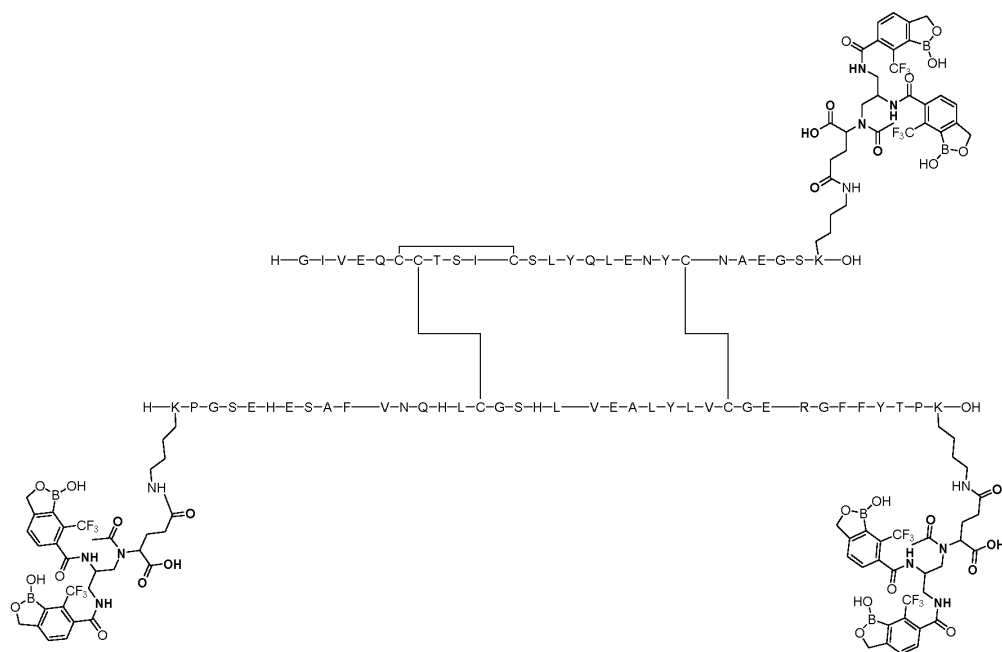
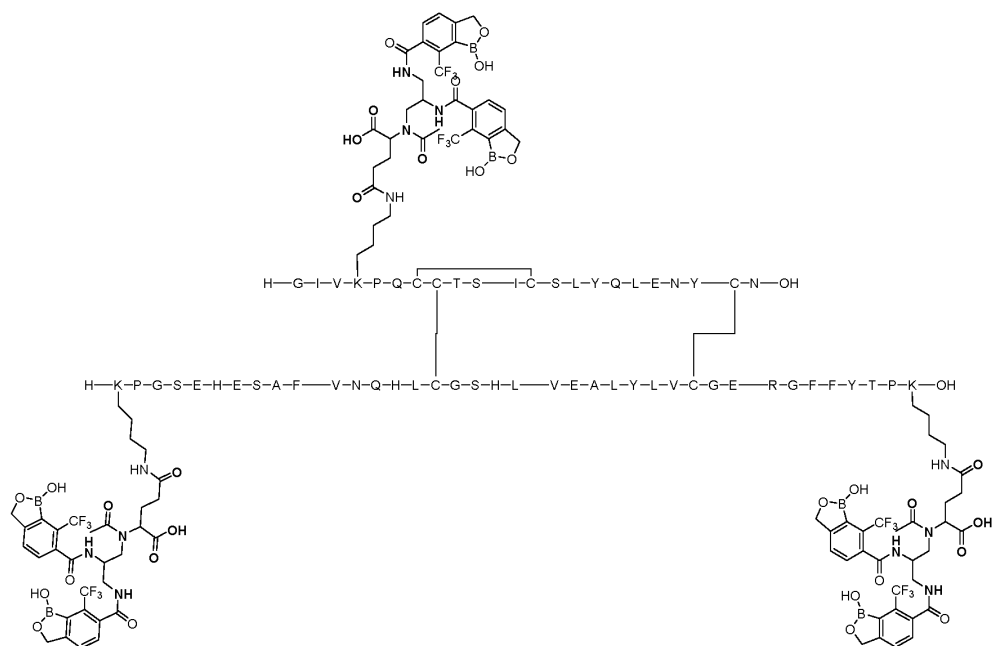
The chemical structure shows the protein complex with the following components:

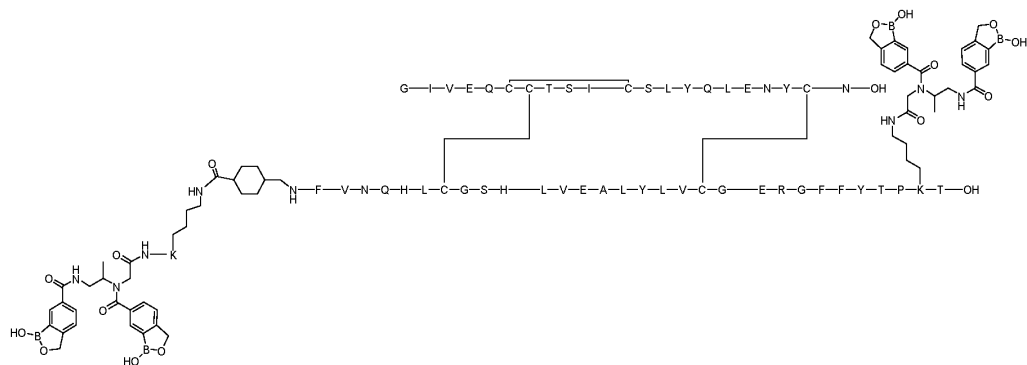
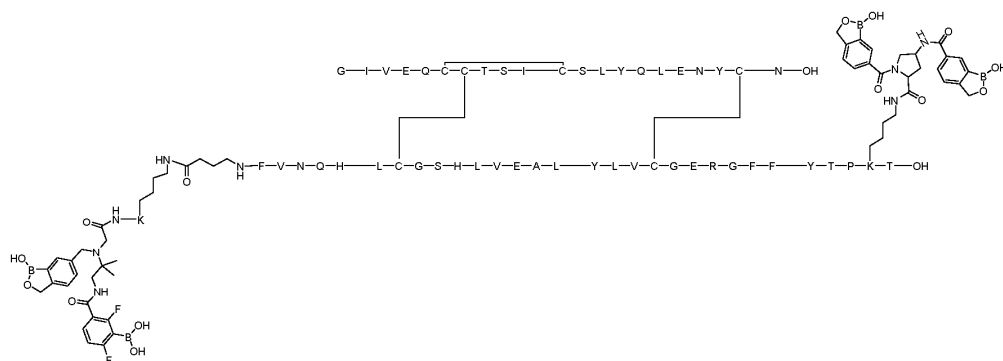
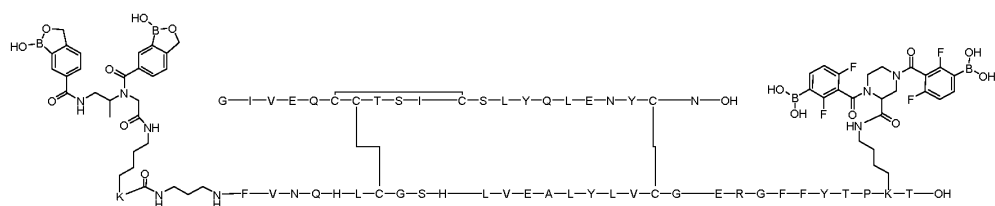
- Protein Backbone:** H-G-I-V-K-Q-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-K-OH
- Ligands:**
 - Top Ligand:** A complex molecule with a central boron atom (B) bonded to a phenyl ring (C₆H₄) and a pyridine ring (C₅H₄N). The boron atom is also bonded to a hydroxyl group (OH) and a methoxy group (OCH₃). The pyridine ring is substituted with a fluorine atom (F) and a methoxy group (OCH₃). The phenyl ring is substituted with a fluorine atom (F) and a methoxy group (OCH₃). The pyridine ring is also bonded to a hydroxyl group (OH) and a methoxy group (OCH₃).
 - Bottom Ligand:** A complex molecule with a central boron atom (B) bonded to a phenyl ring (C₆H₄) and a pyridine ring (C₅H₄N). The boron atom is also bonded to a hydroxyl group (OH) and a methoxy group (OCH₃). The pyridine ring is substituted with a fluorine atom (F) and a methoxy group (OCH₃). The phenyl ring is substituted with a fluorine atom (F) and a methoxy group (OCH₃). The pyridine ring is also bonded to a hydroxyl group (OH) and a methoxy group (OCH₃).

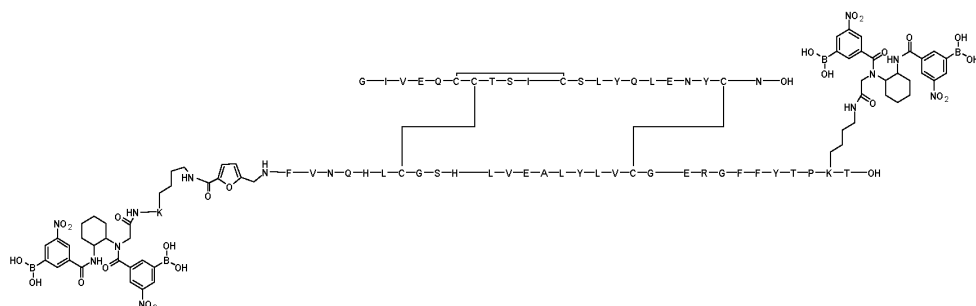
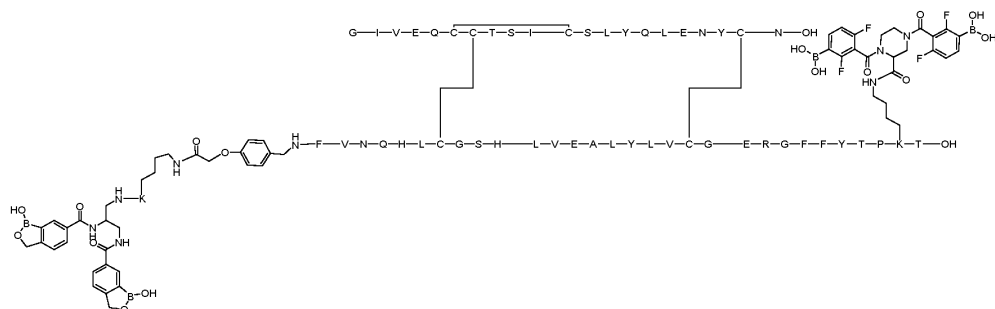
Example 138:**Example 139:**



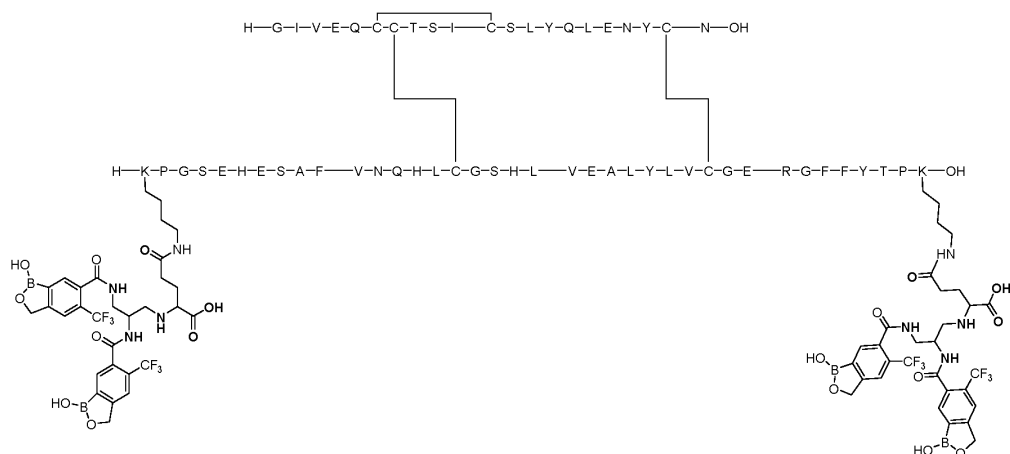
Chemical structure of the *in vitro* synthesized peptide **1** (H-K-P-G-I-V-E-Q-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH) is shown, which is a truncated version of the full-length peptide. The structure includes the peptide backbone and the side chains of the amino acids, with the C-terminal amino group (NH₂) and the N-terminal carboxyl group (COOH) highlighted in red. The peptide is shown in its zwitterionic form, with the N-terminal amino group protonated and the C-terminal carboxyl group deprotonated.

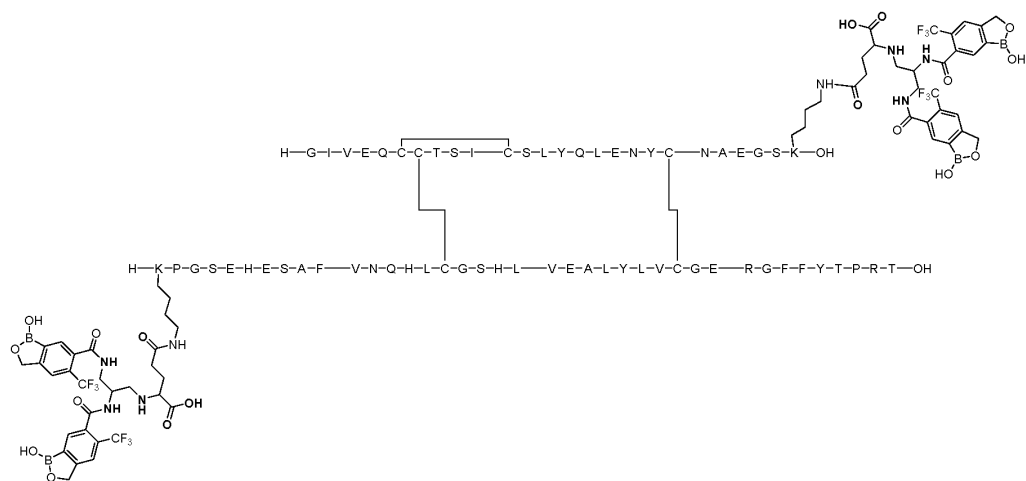
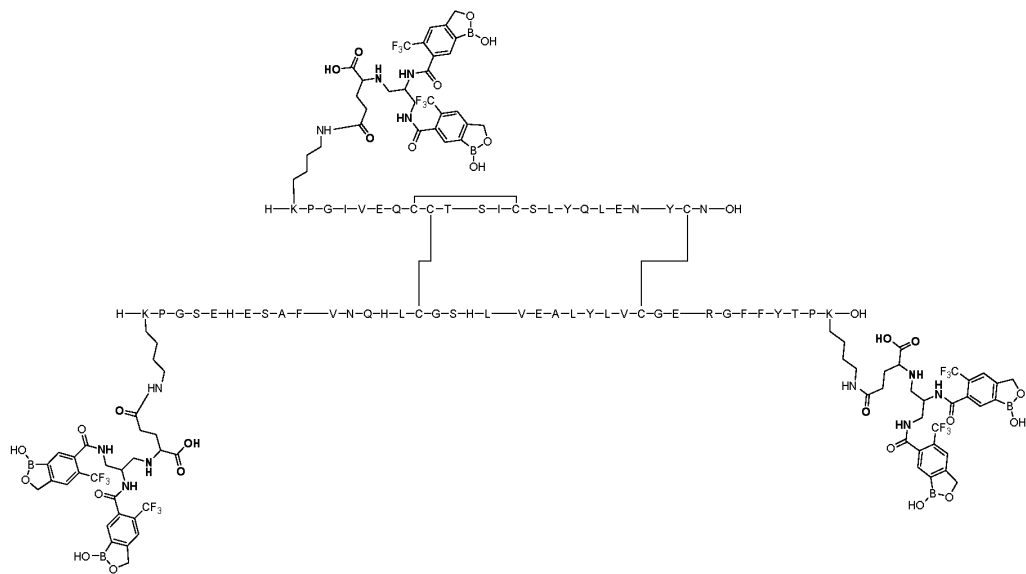
Example 142:**Example 143:**

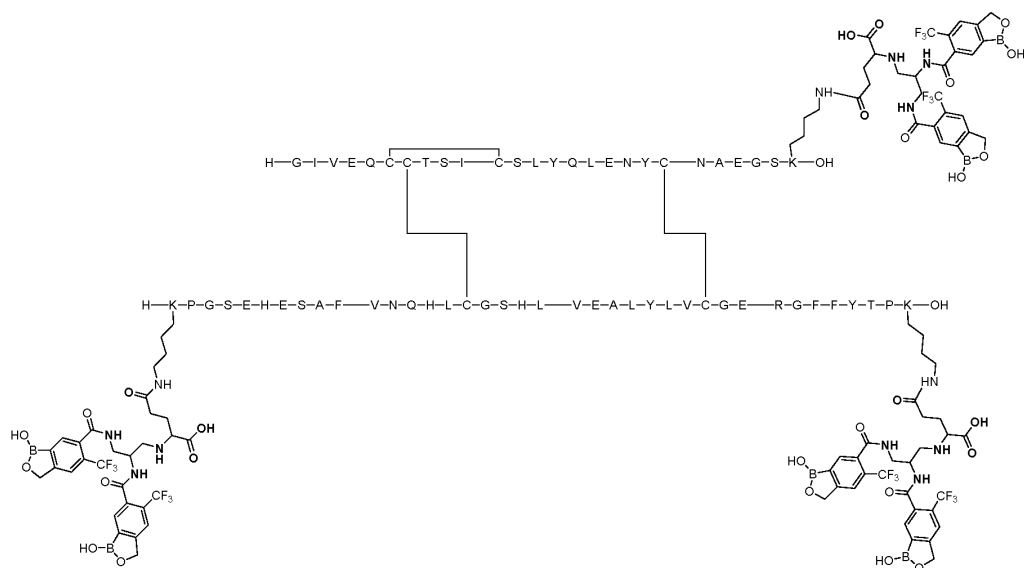
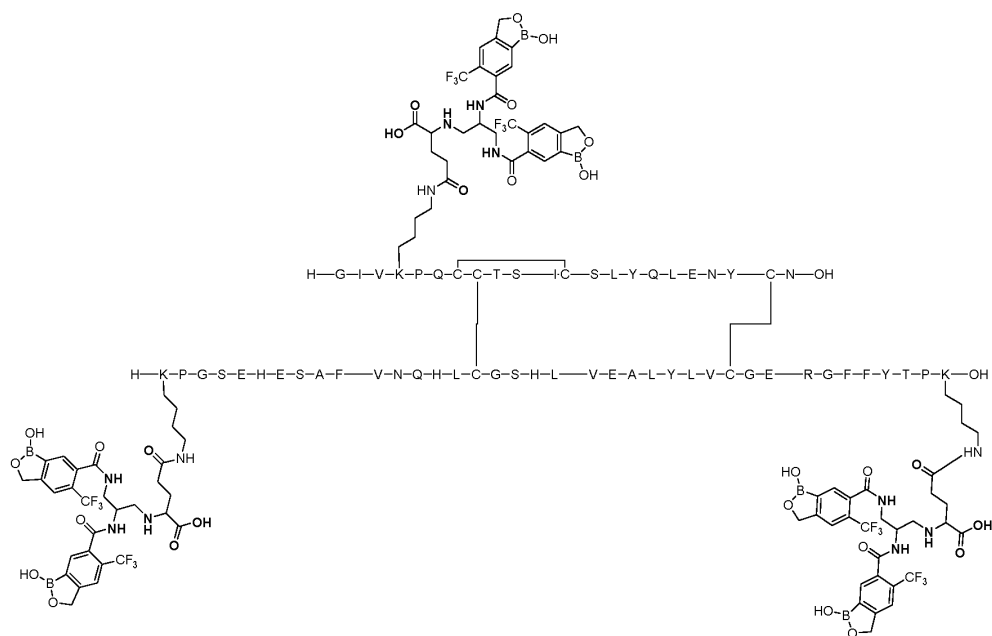
Example 144:**Example 145:****5 Example 146:**

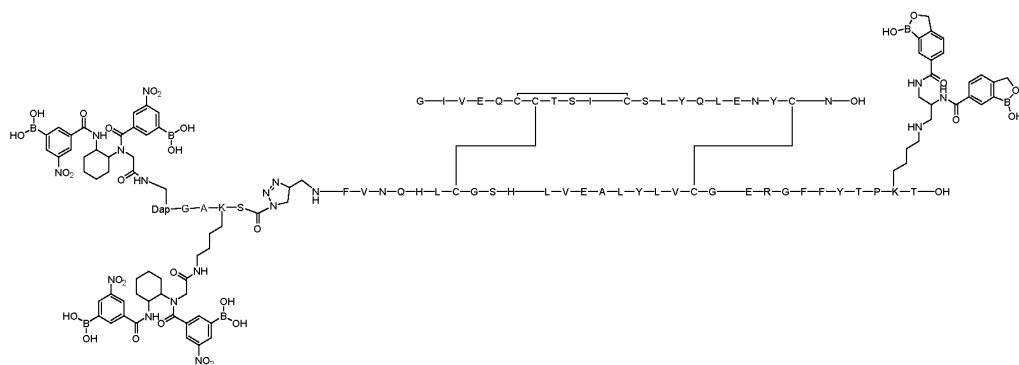
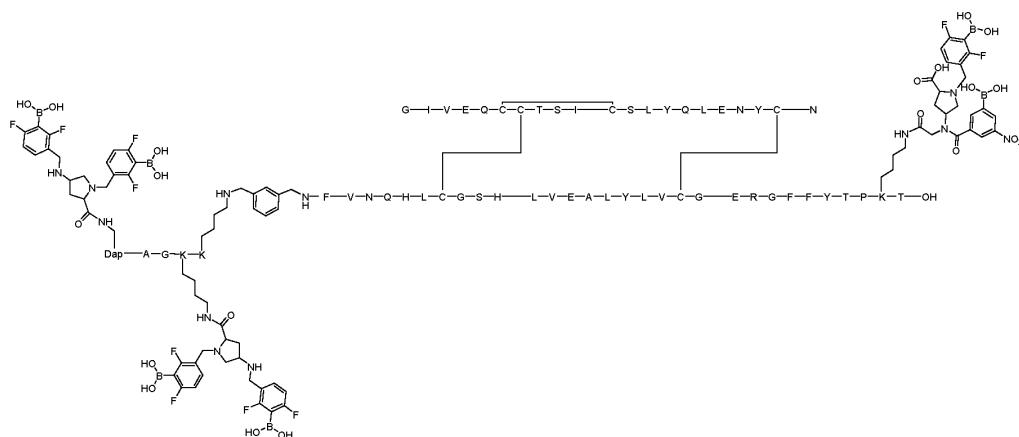
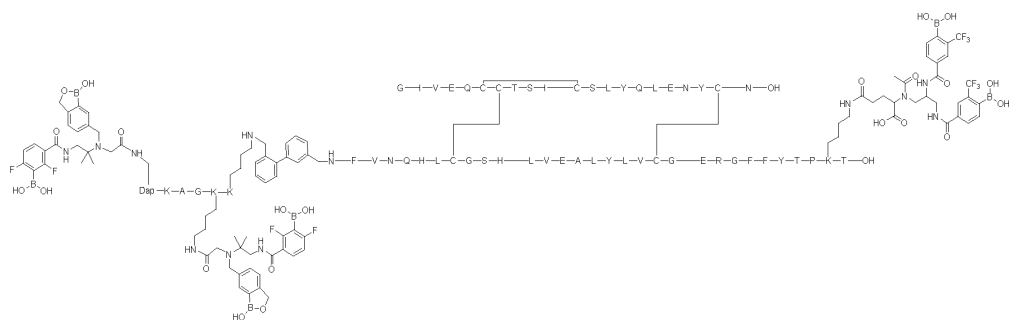
Example 147:**Example 148:**

5

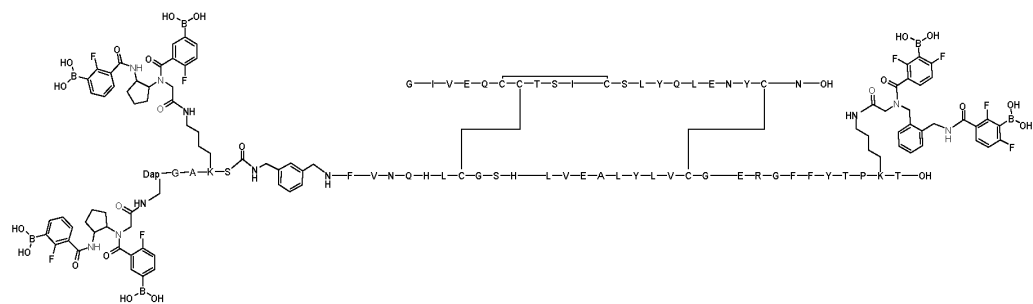
Example 149:

Example 150:**Example 151:**

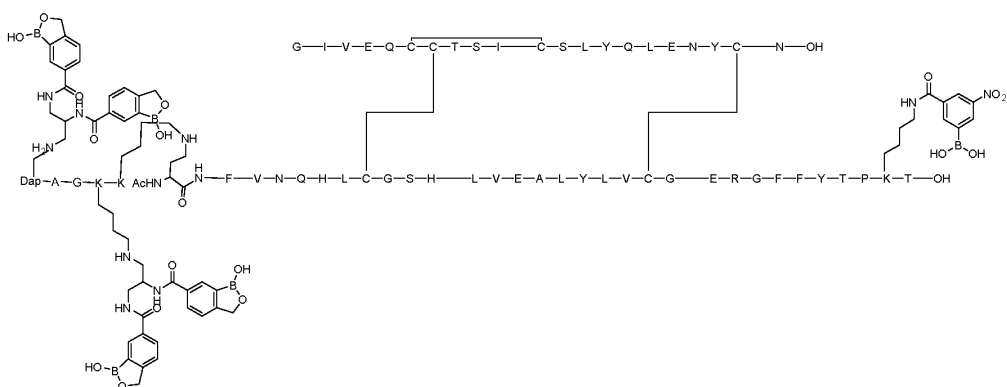
Example 152:**Example 153:**

Example 154:**Example 155:****5 Example 156:**

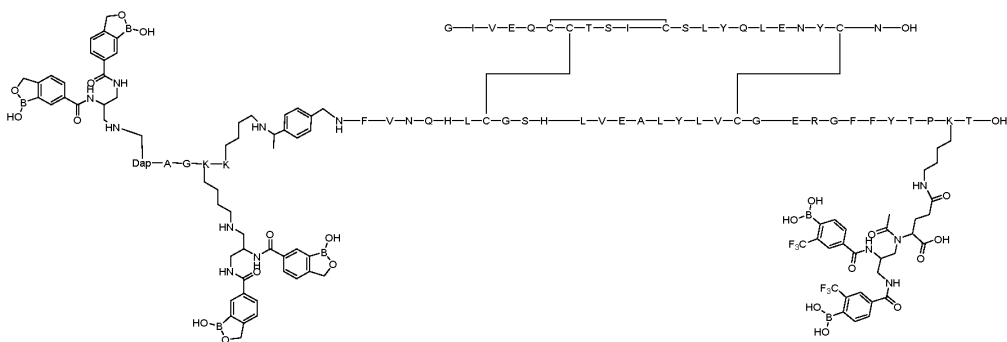
Example 157:

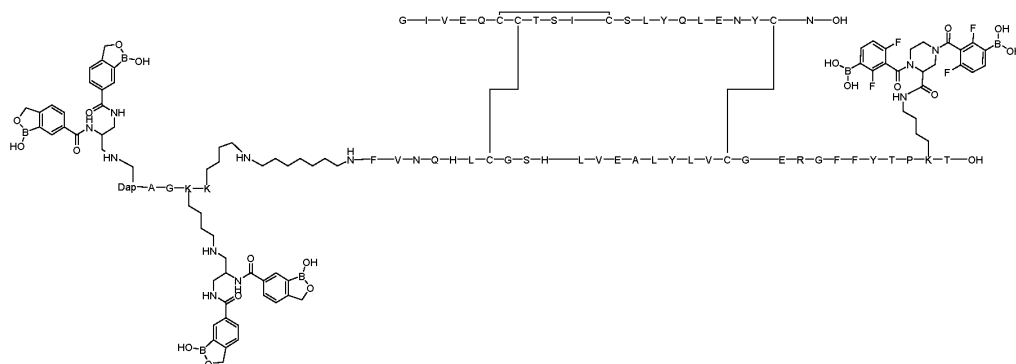
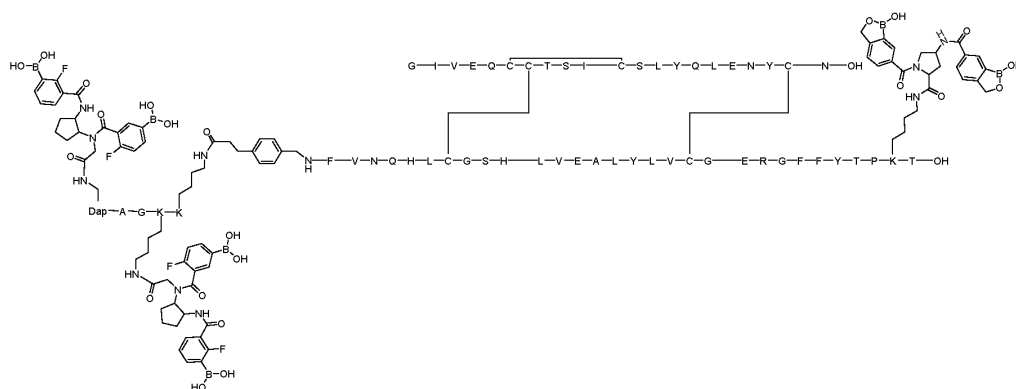
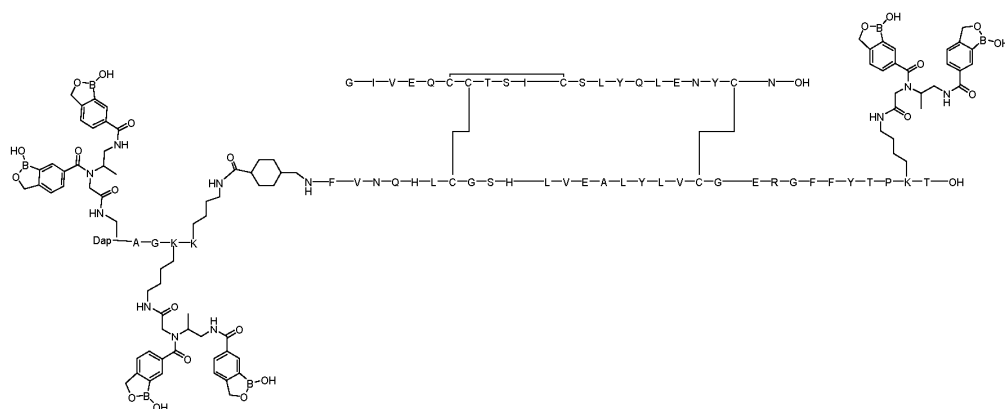


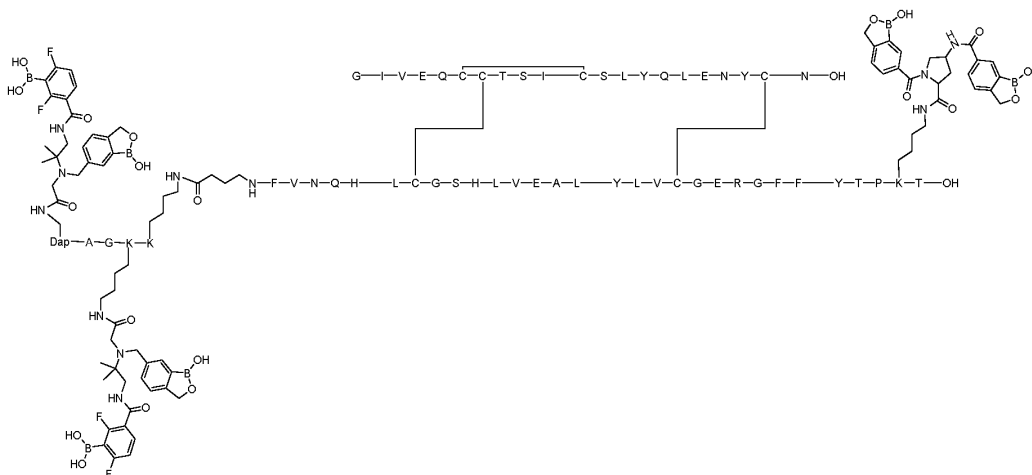
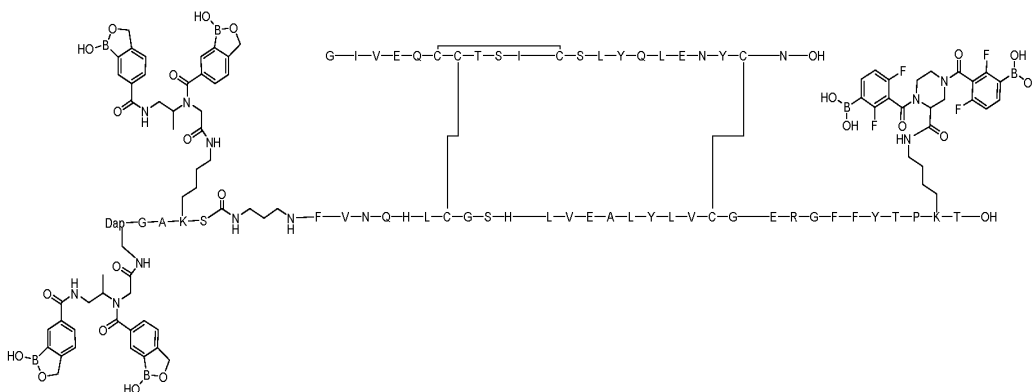
Example 158:



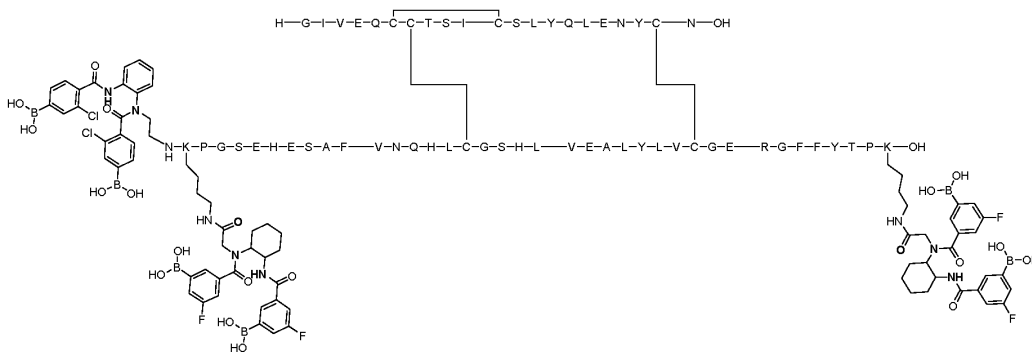
5 Example 159:

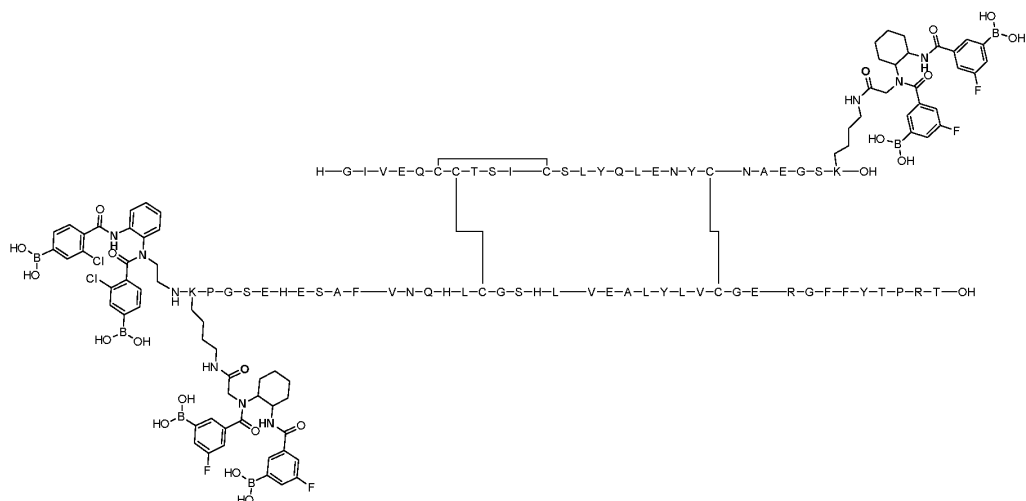
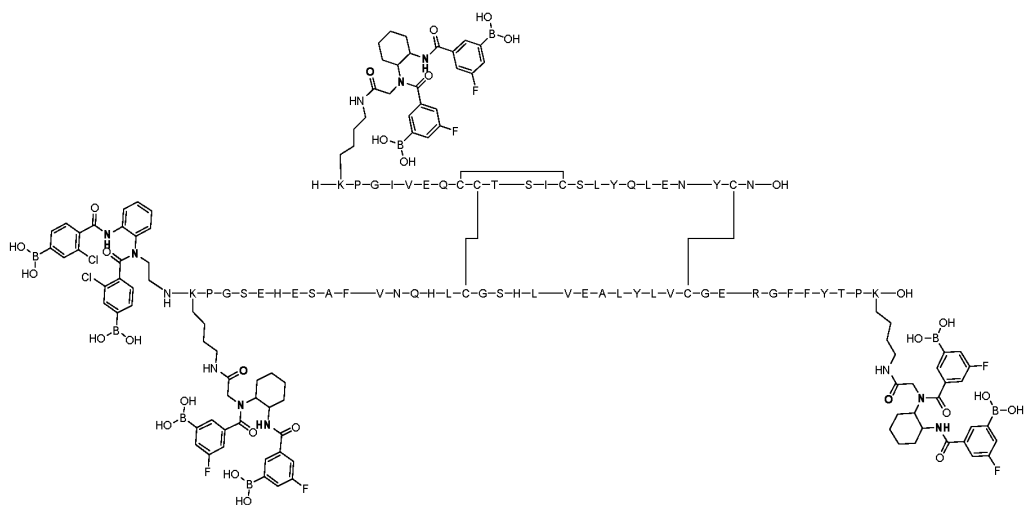


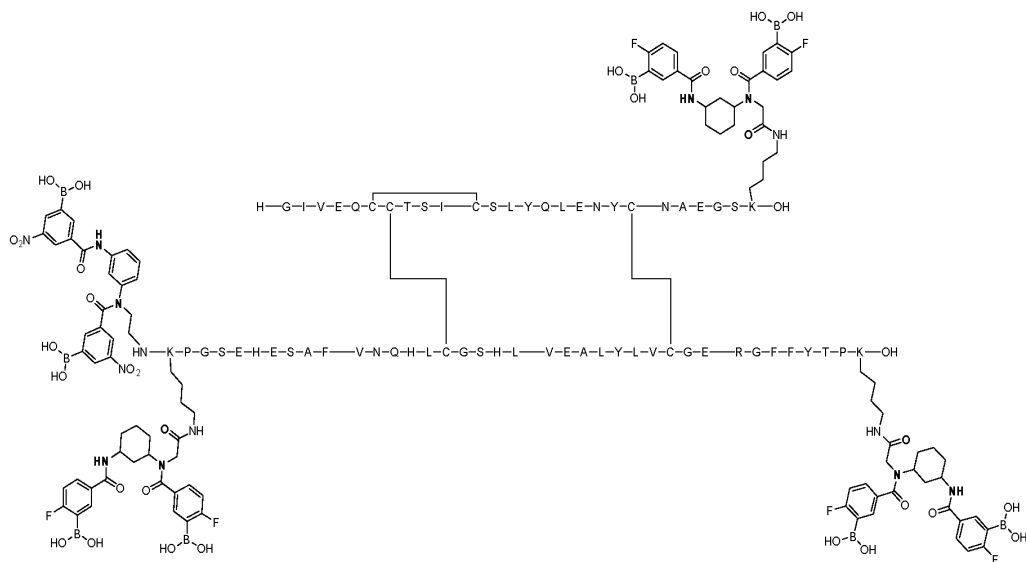
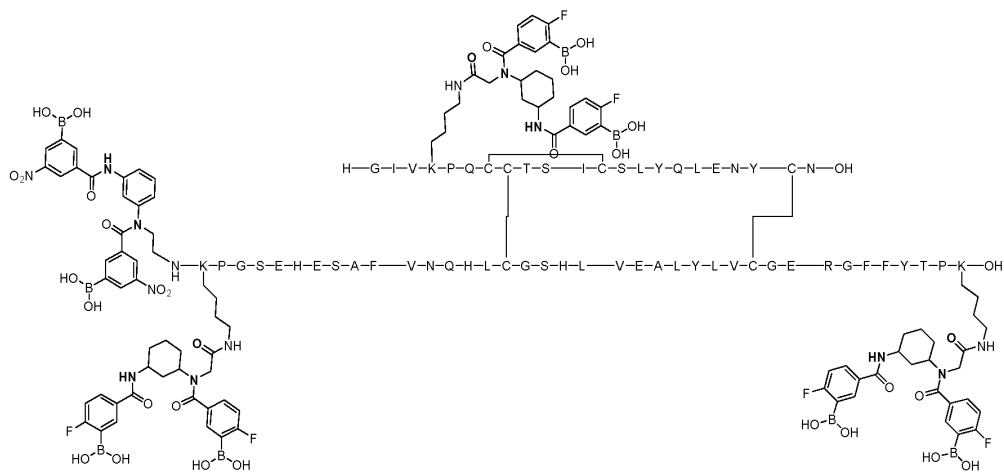
Example 160:**Example 161:****5 Example 162:**

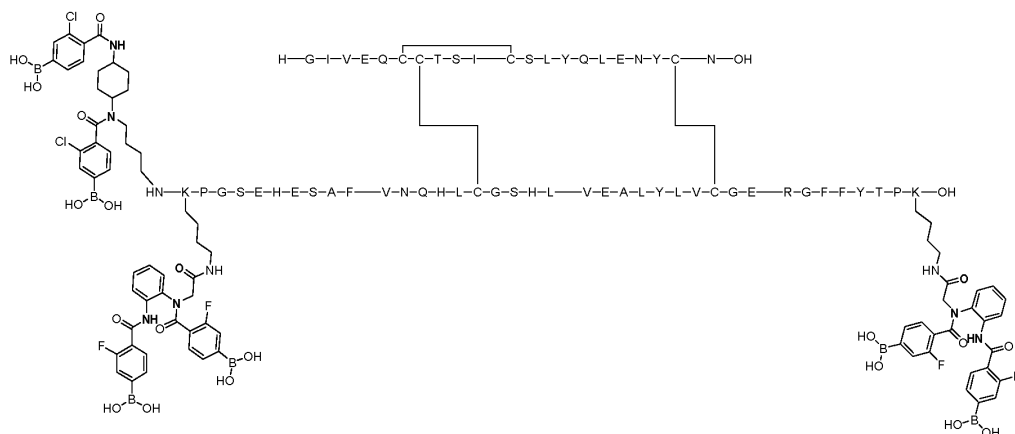
Example 162:**Example 163:**

5

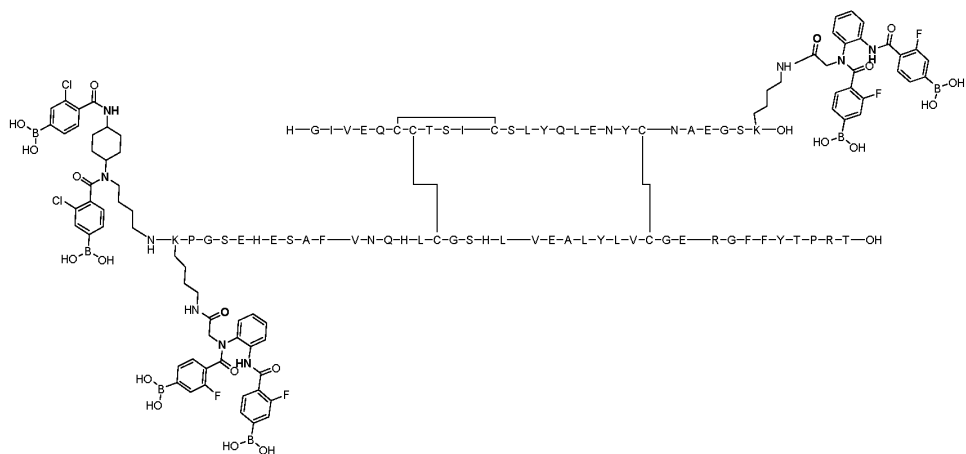
Example 164:

Example 165:**Example 166:**

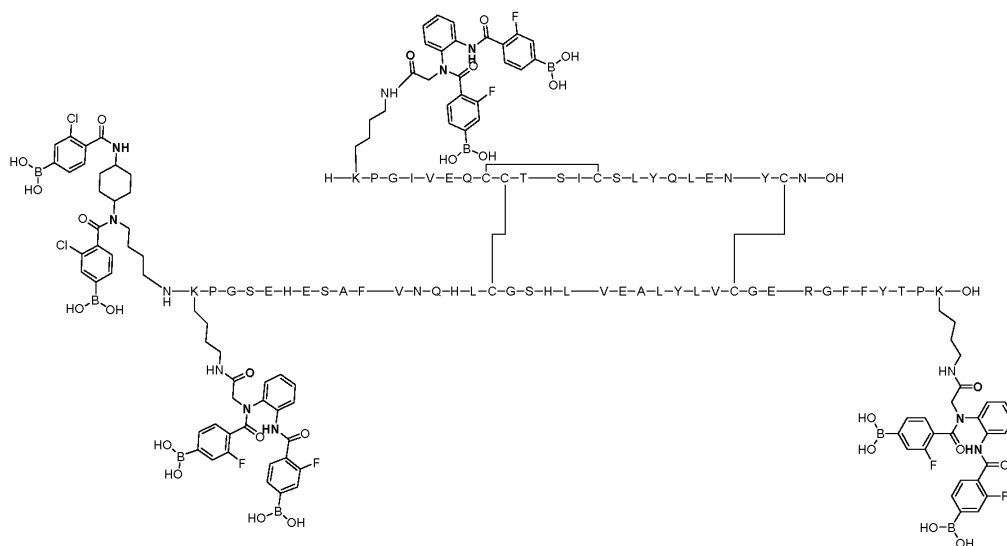
Example 167:**Example 168:**

Example 169:

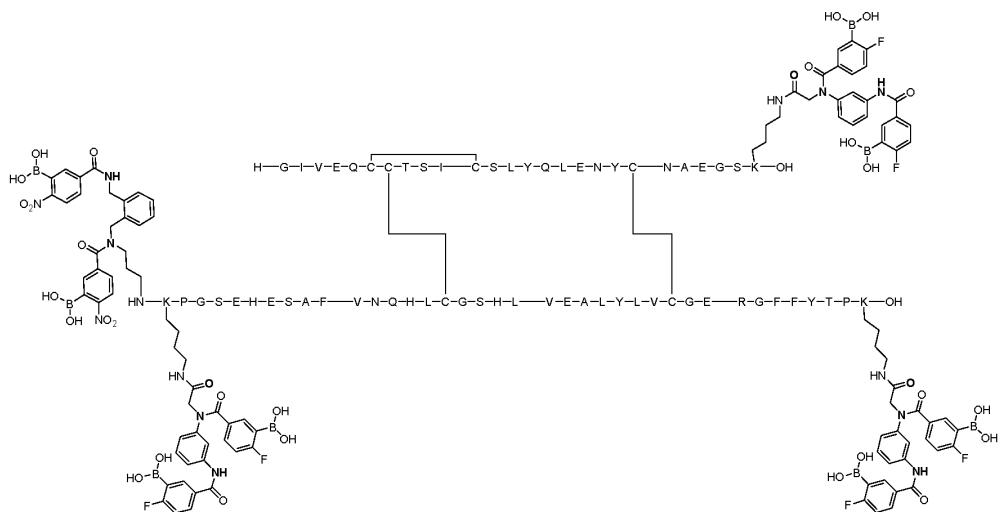
Example 170:

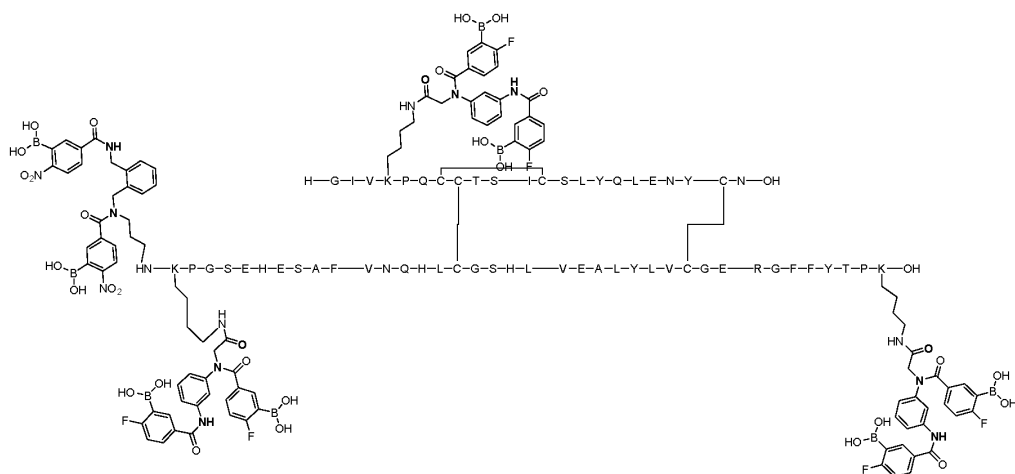


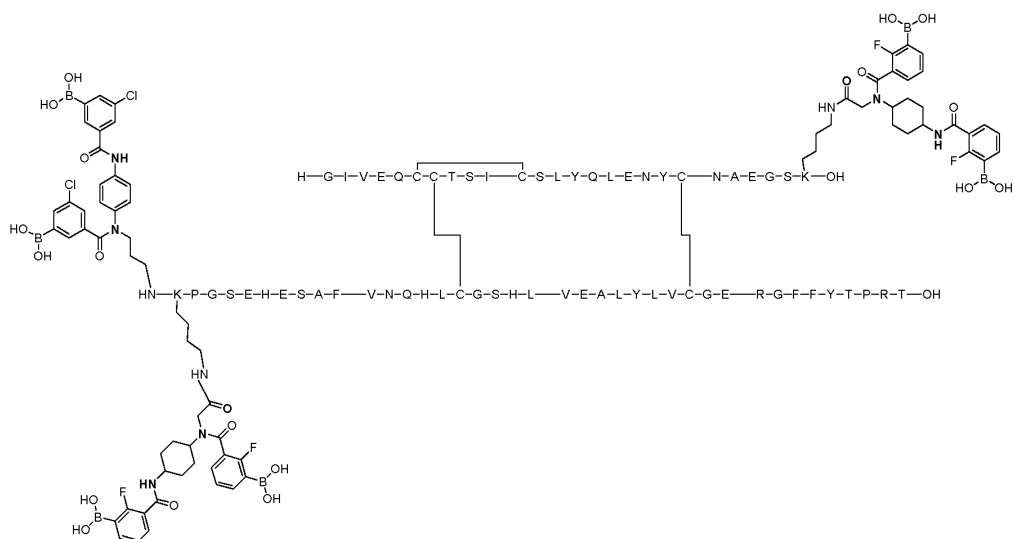
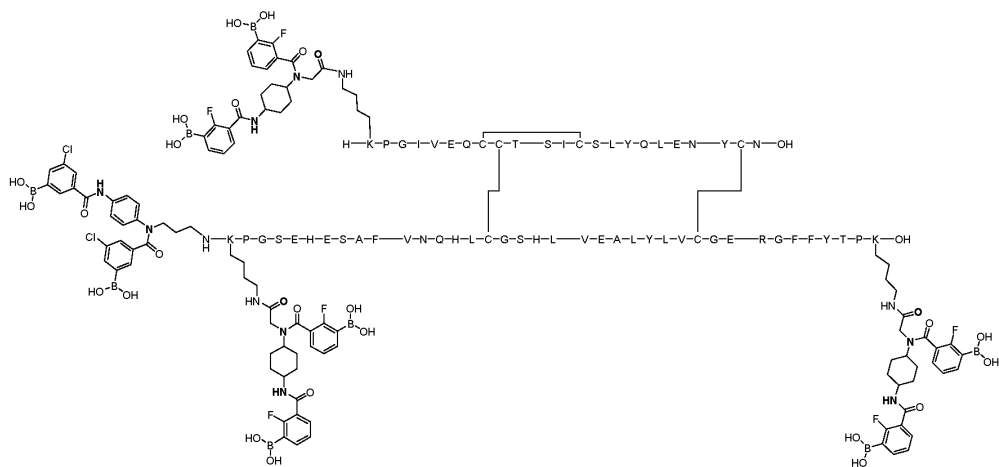
Example 171:



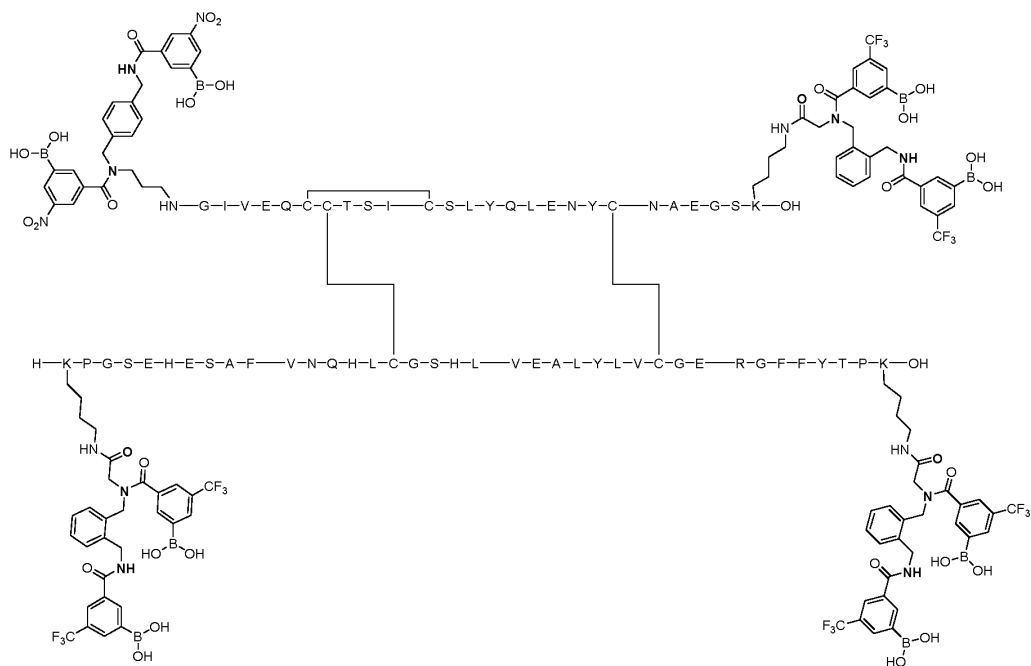
Example 172:



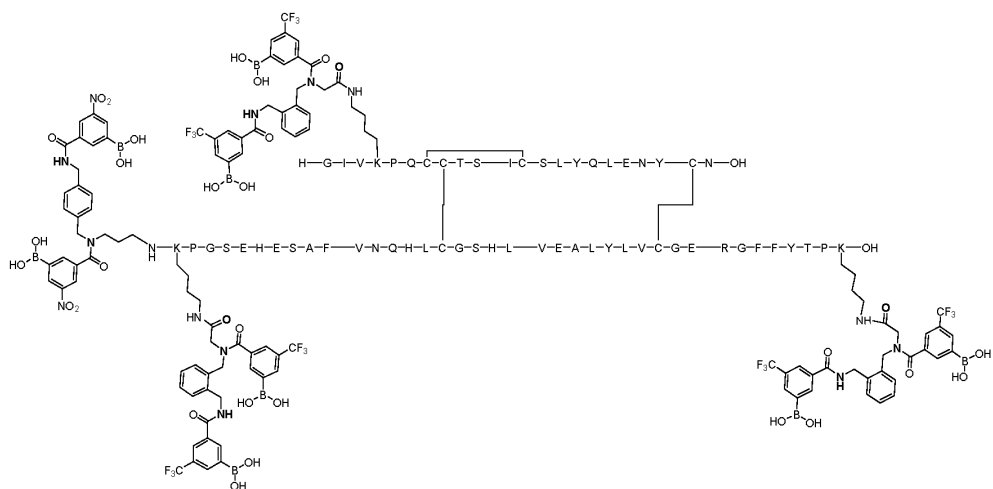


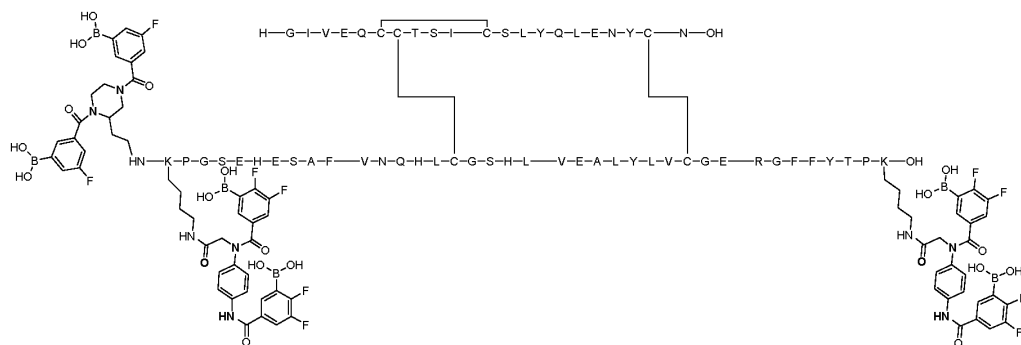
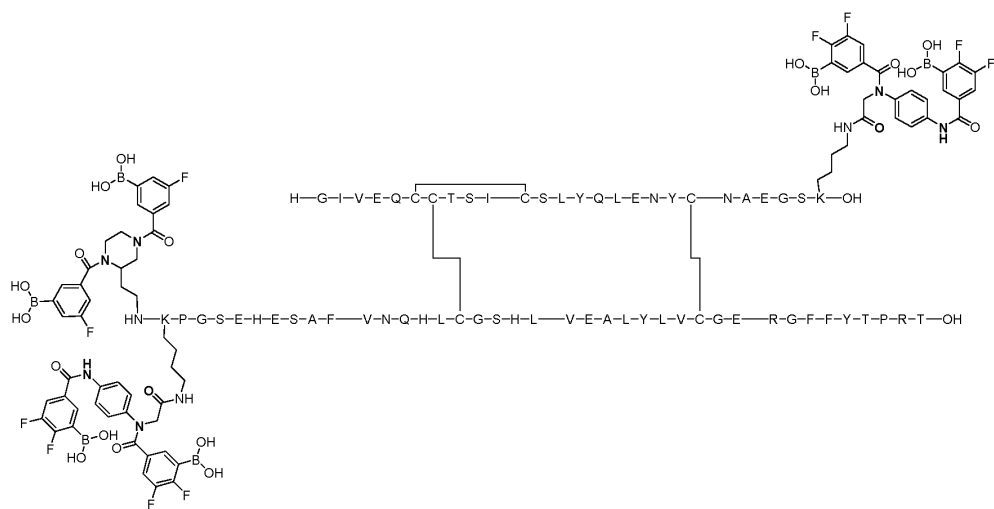
Example 175:**Example 176:**

Example 177:

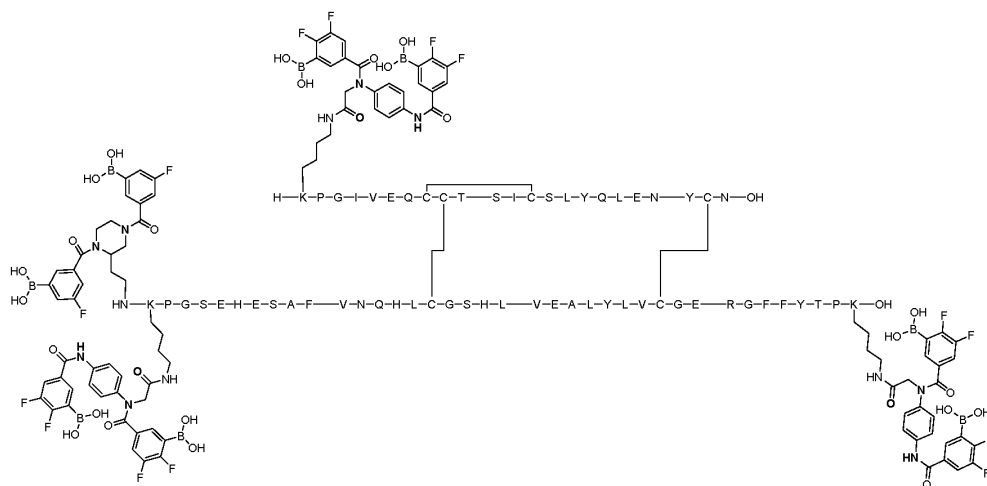


Example 178:

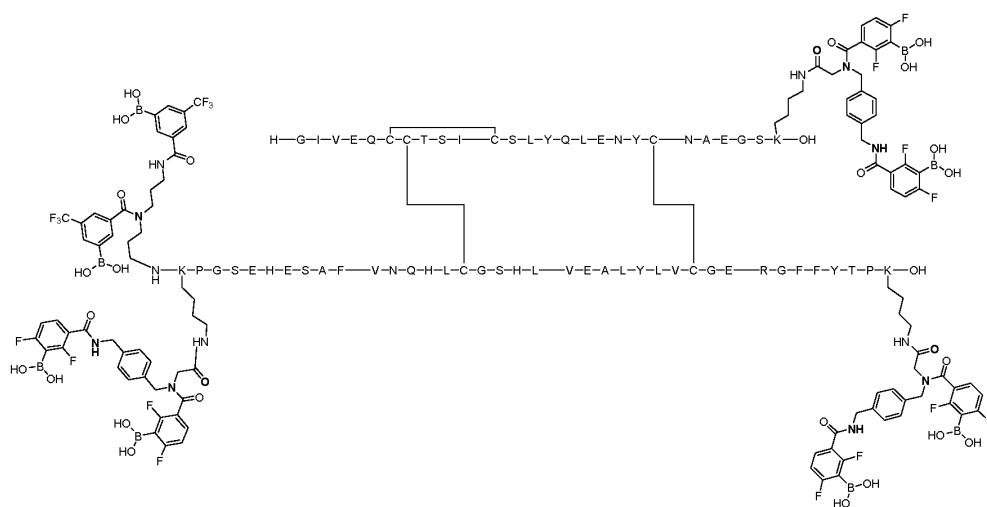


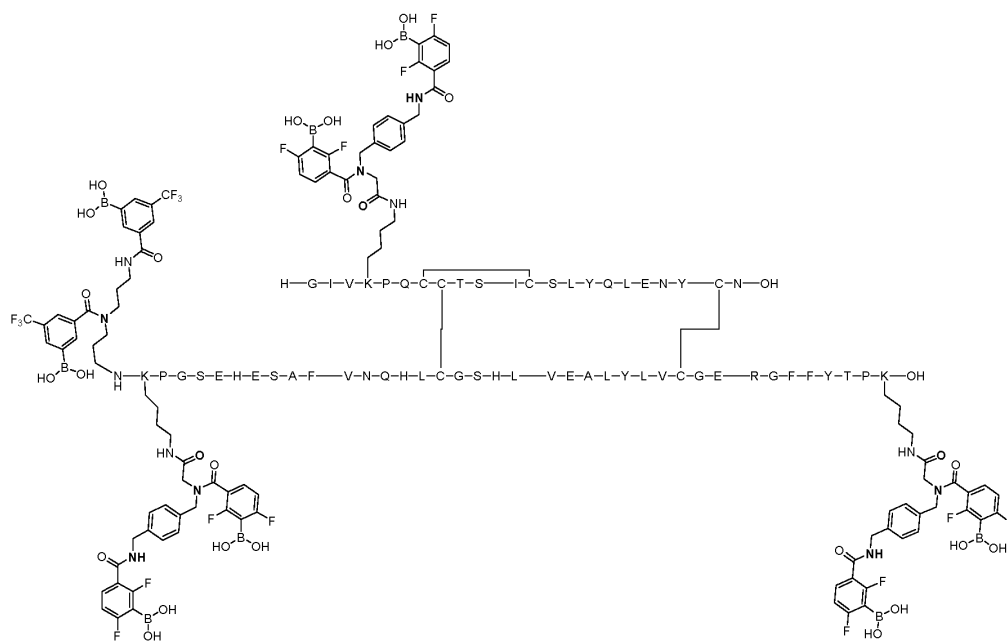
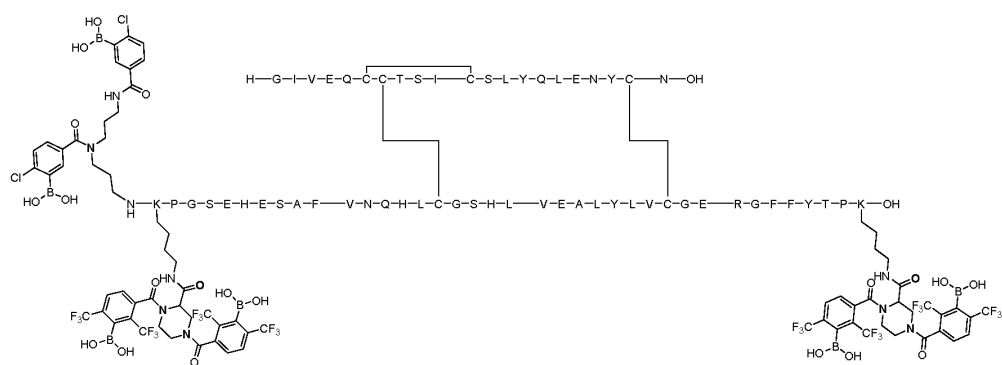
Example 179:**Example 180:**

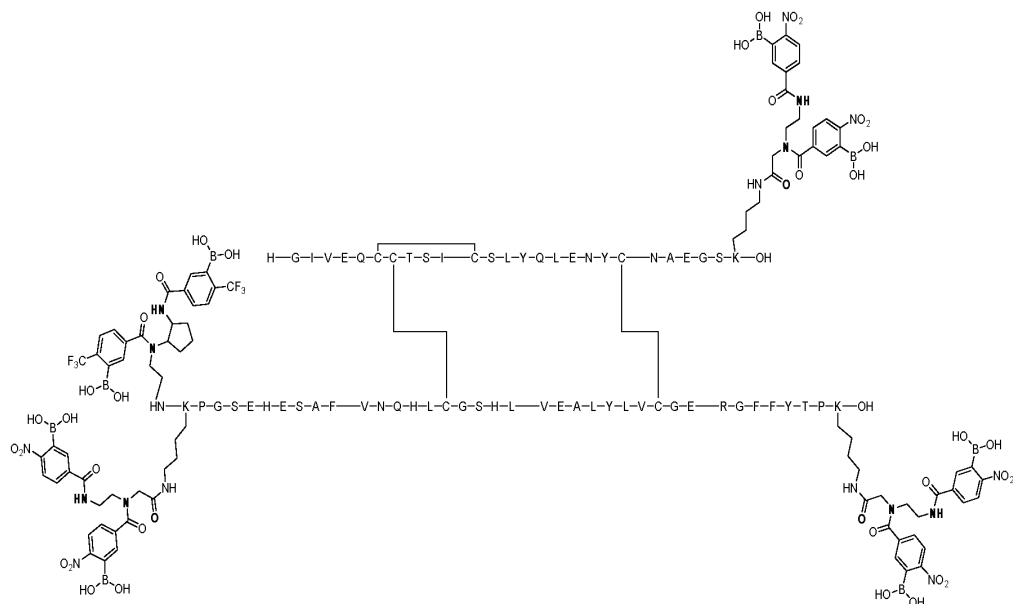
Example 181:

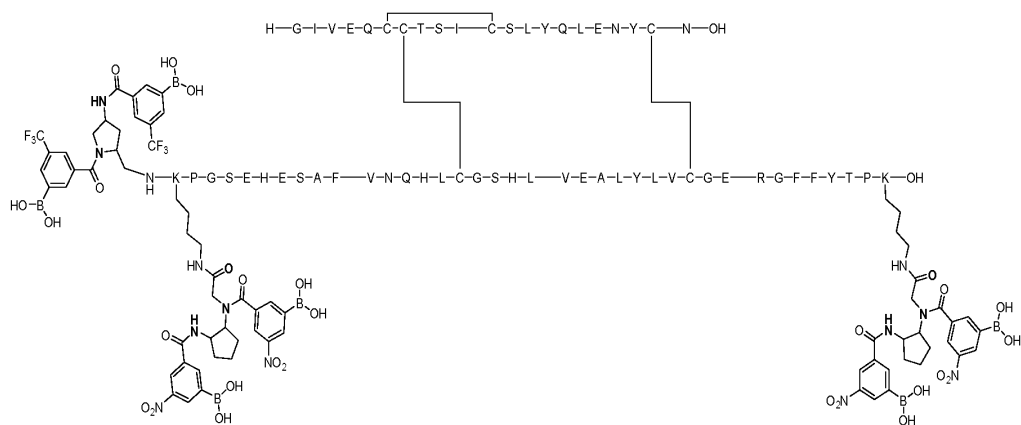
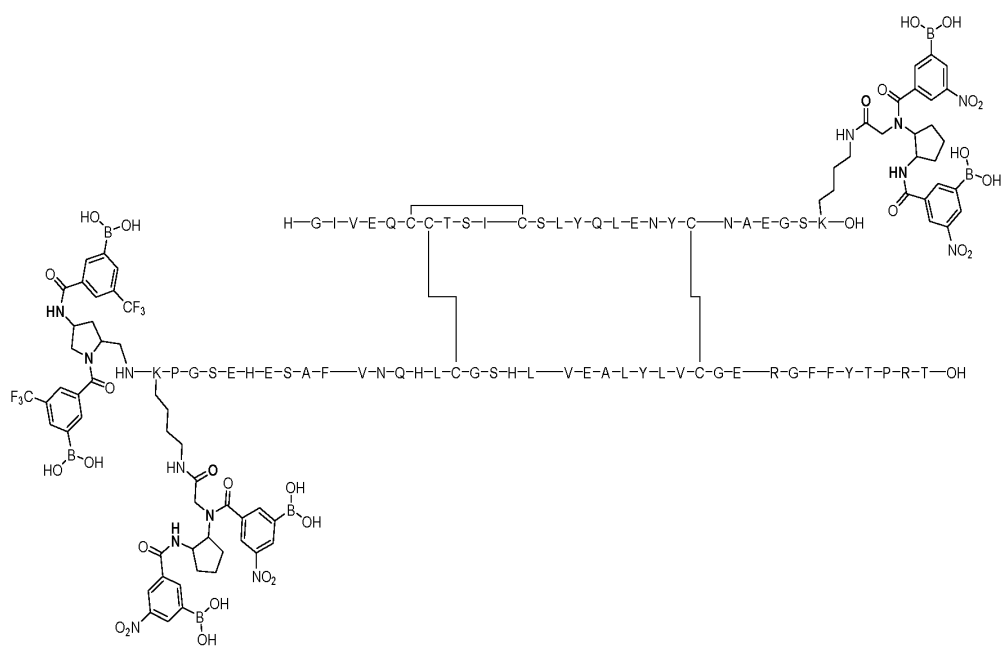


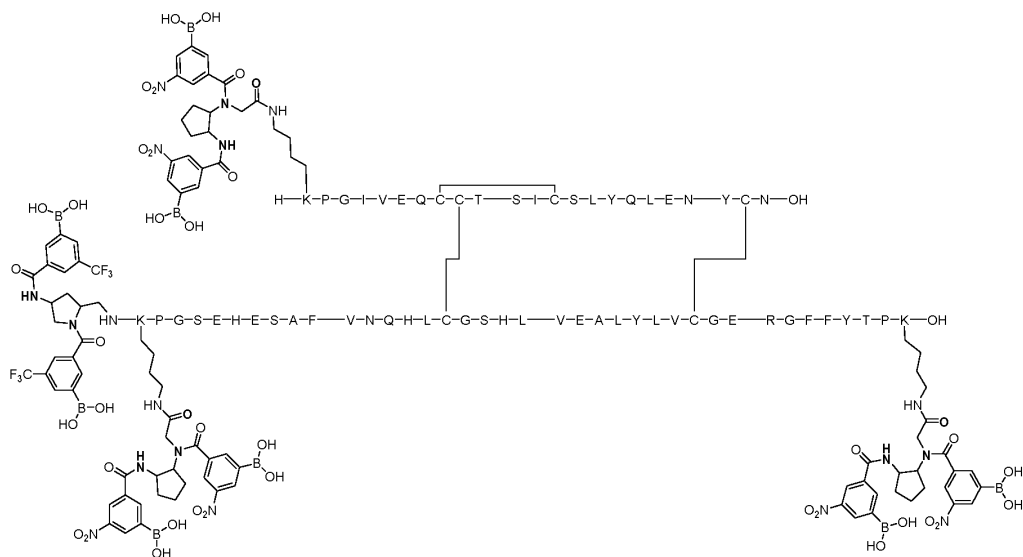
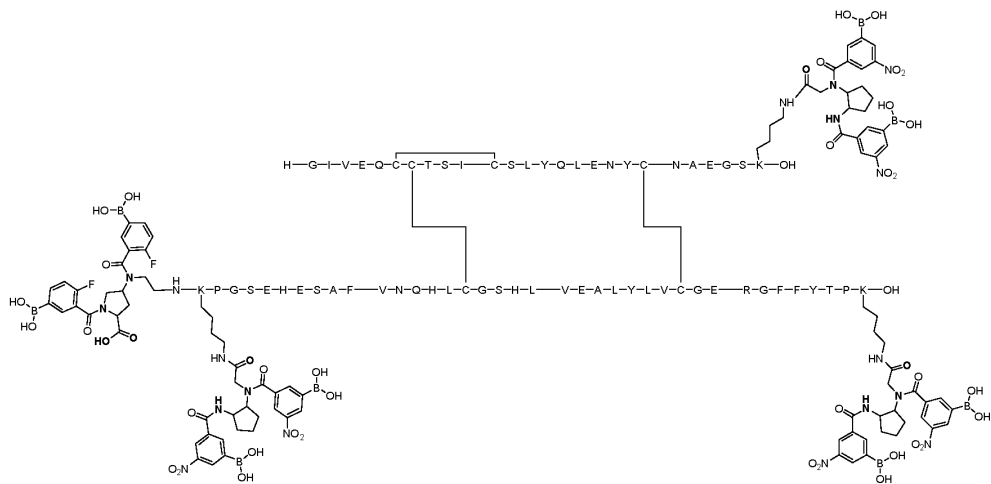
Example 182:



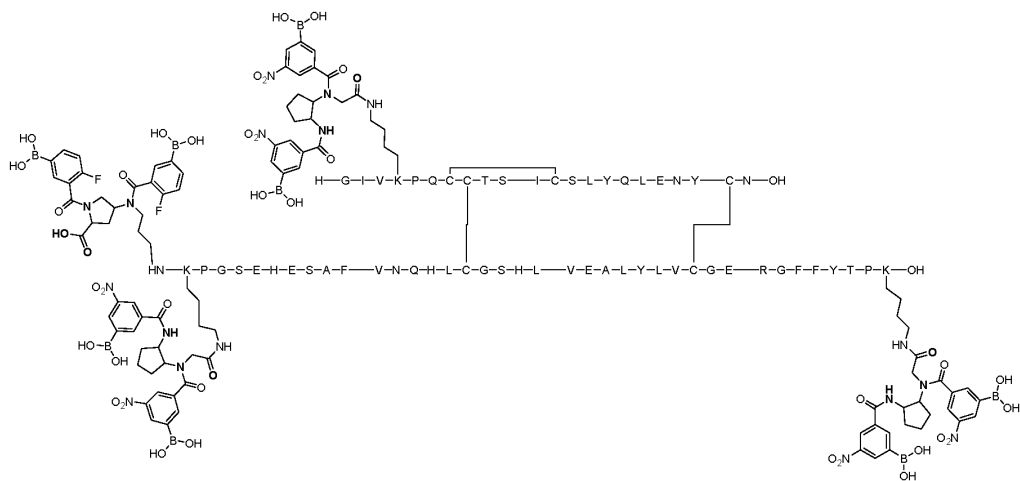
Example 183:**Example 184:**

[illegible]

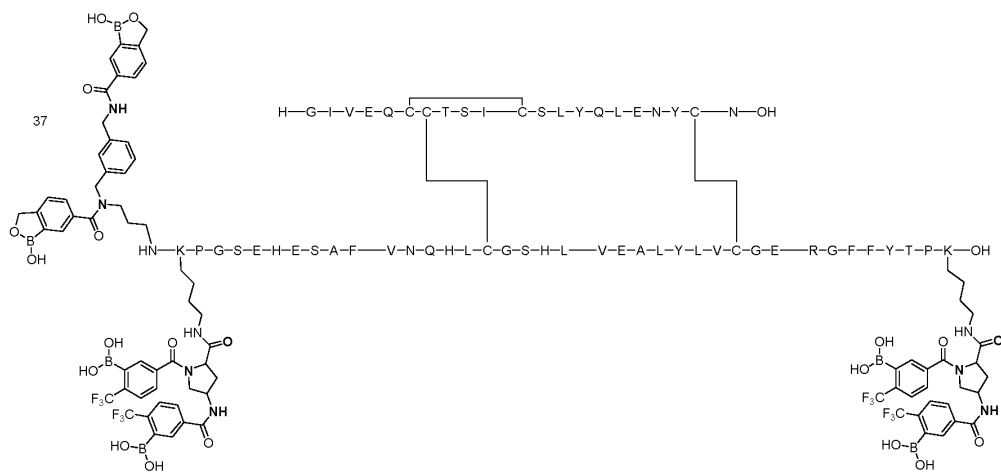
Example 189:**Example 190:**

Example 191:**Example 192:**

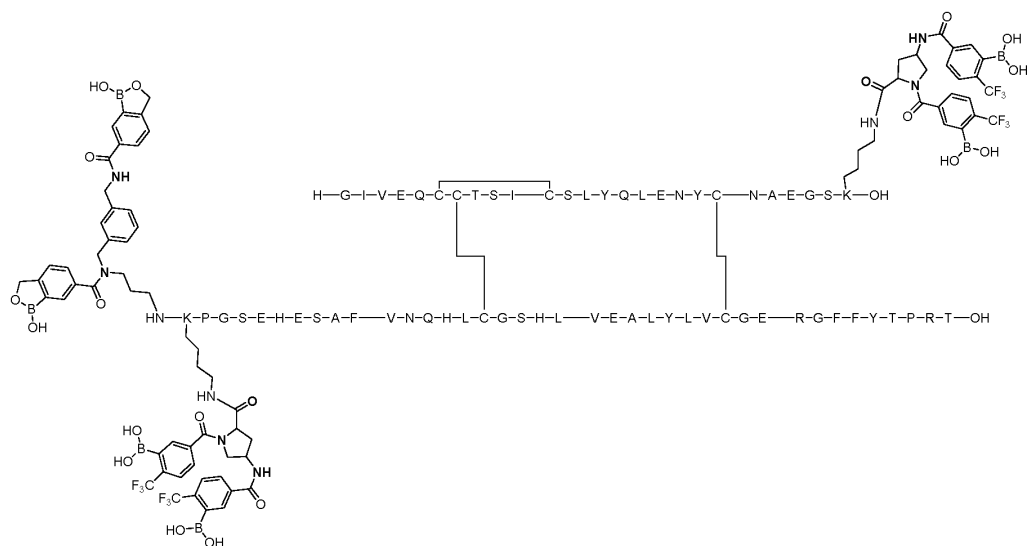
Example 193:



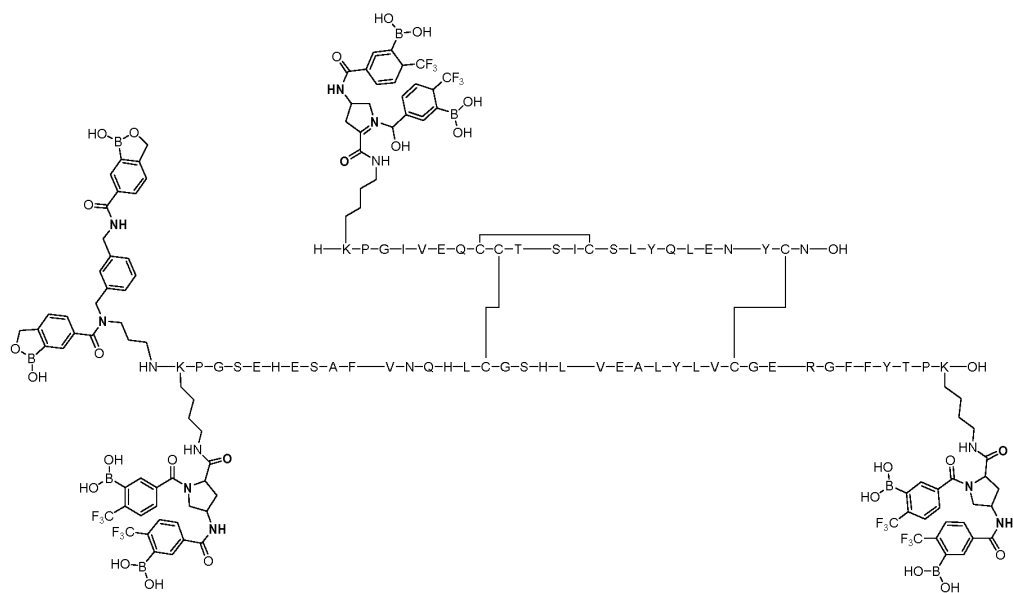
Example 194:

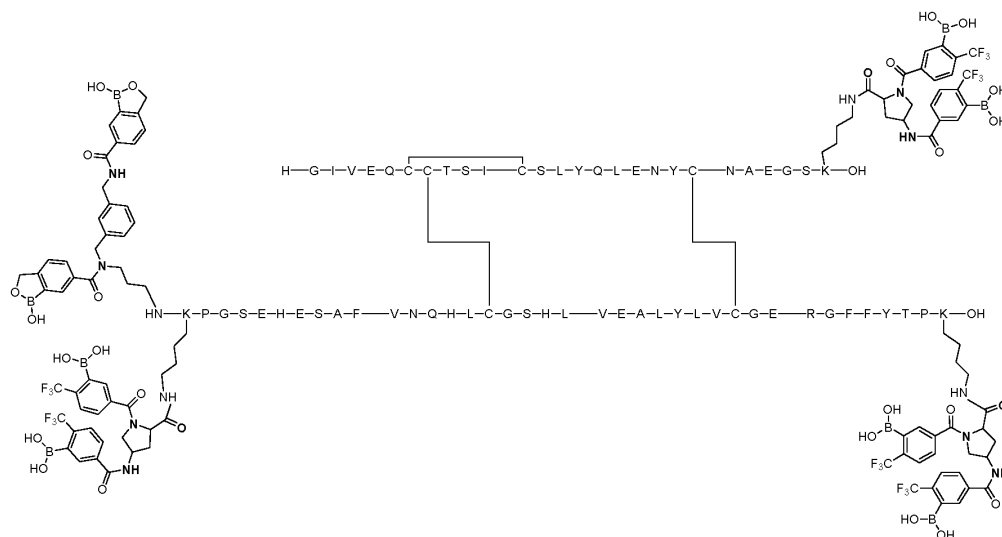
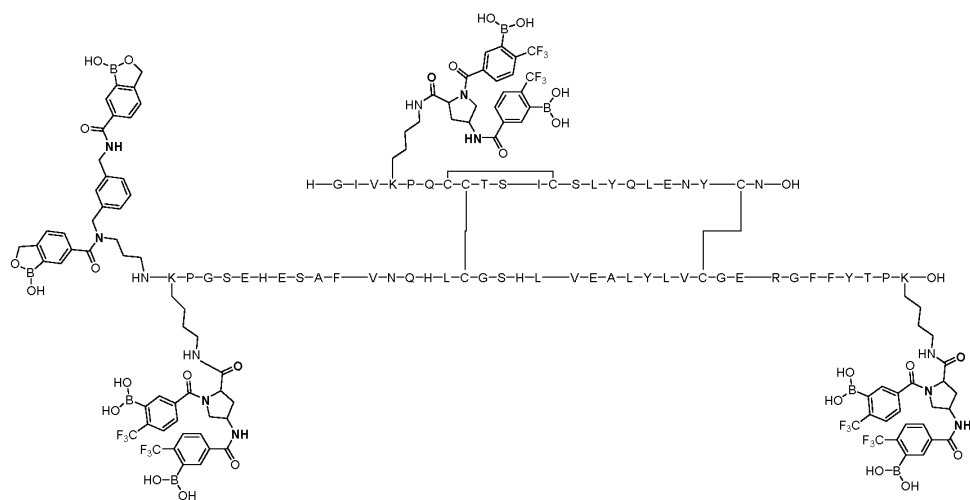


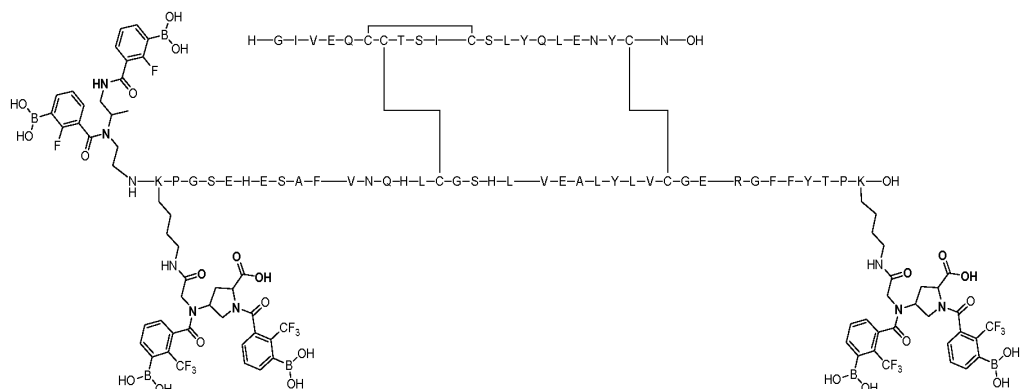
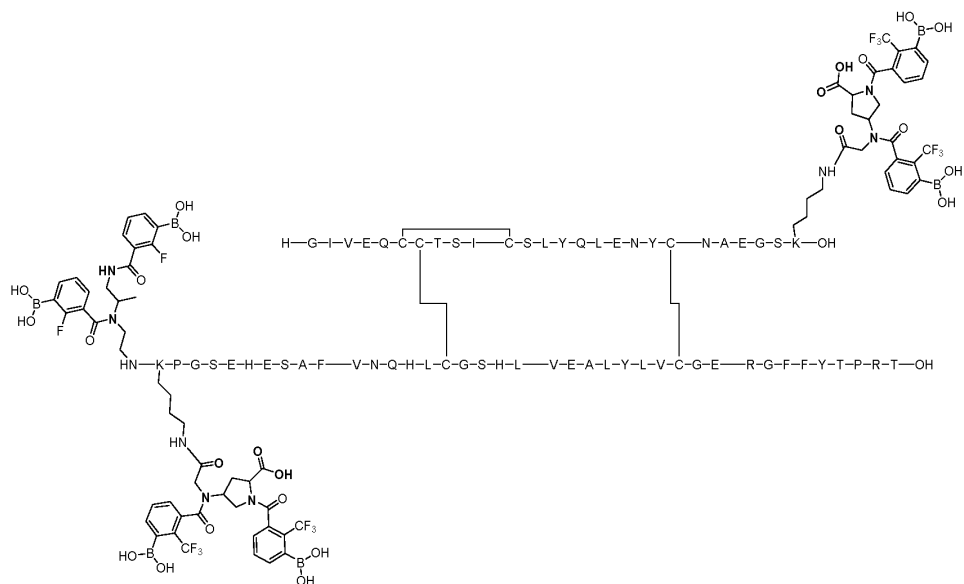
Example 195:

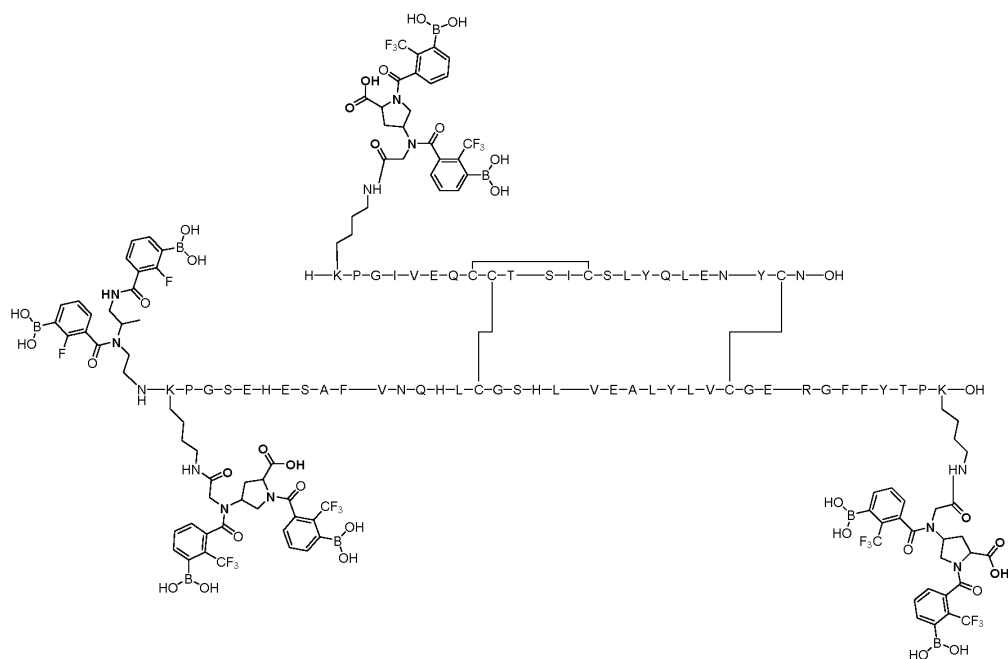
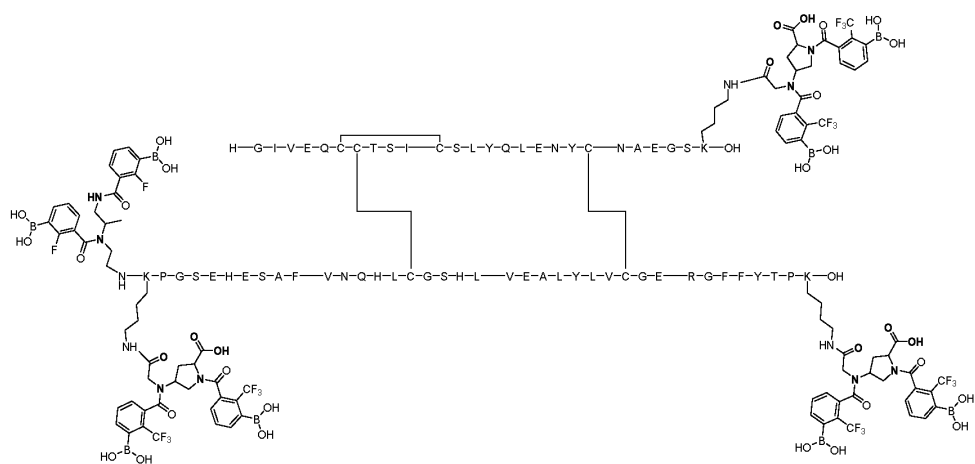


Example 196:



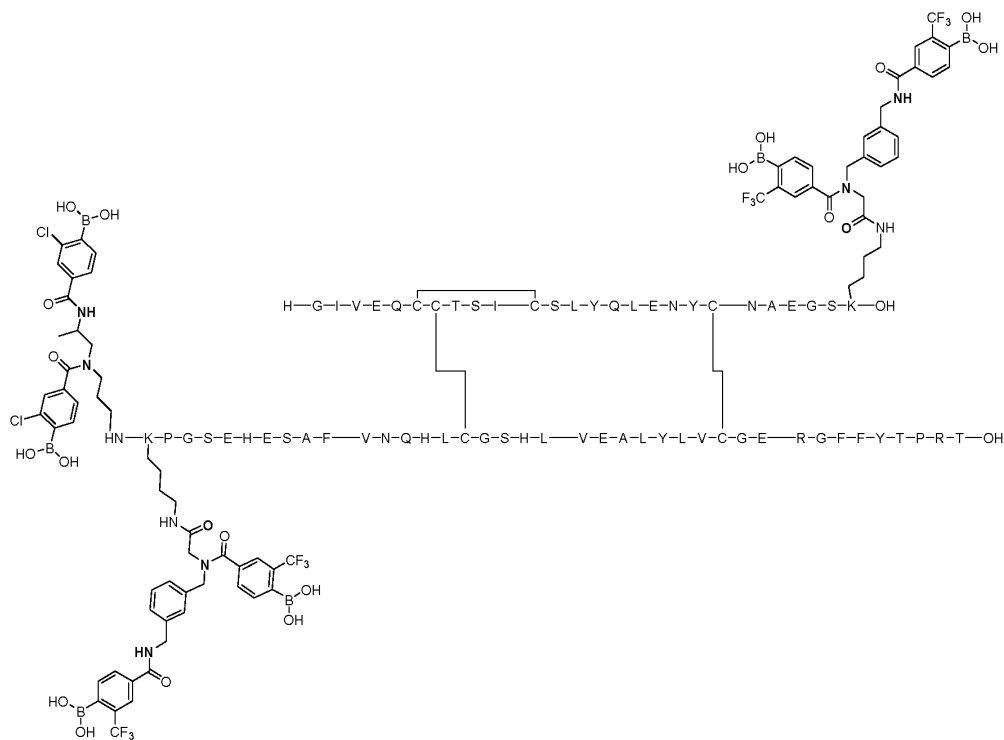
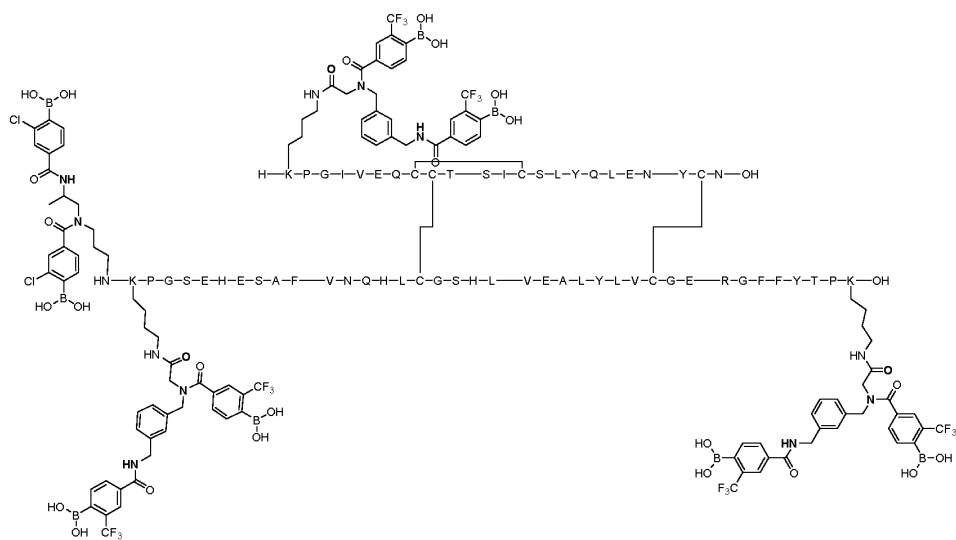
Example 197:**Example 198:**

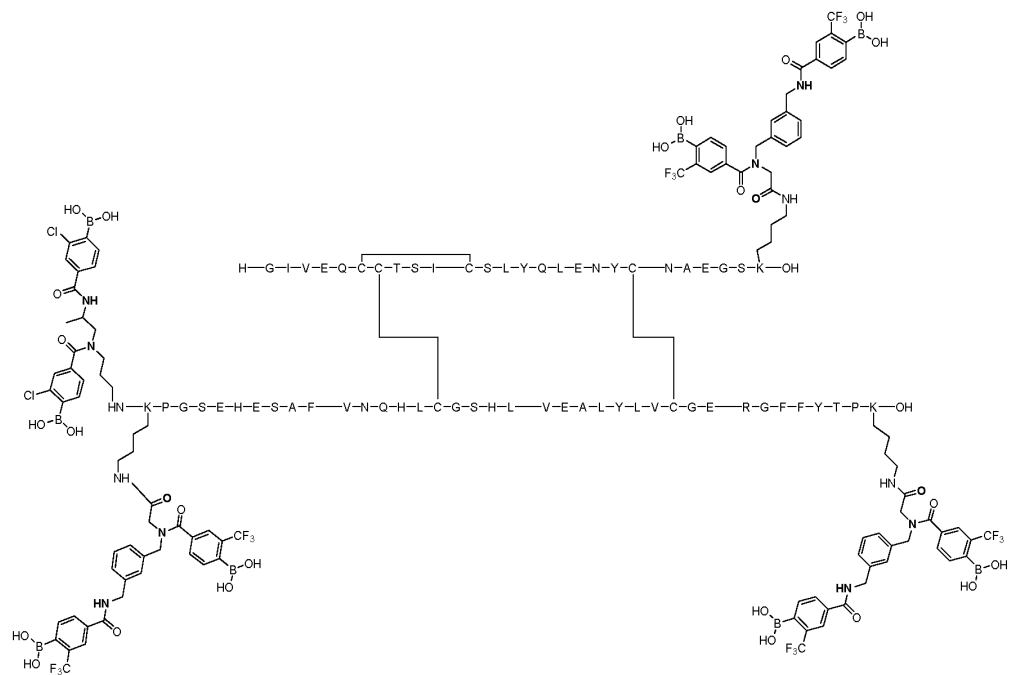
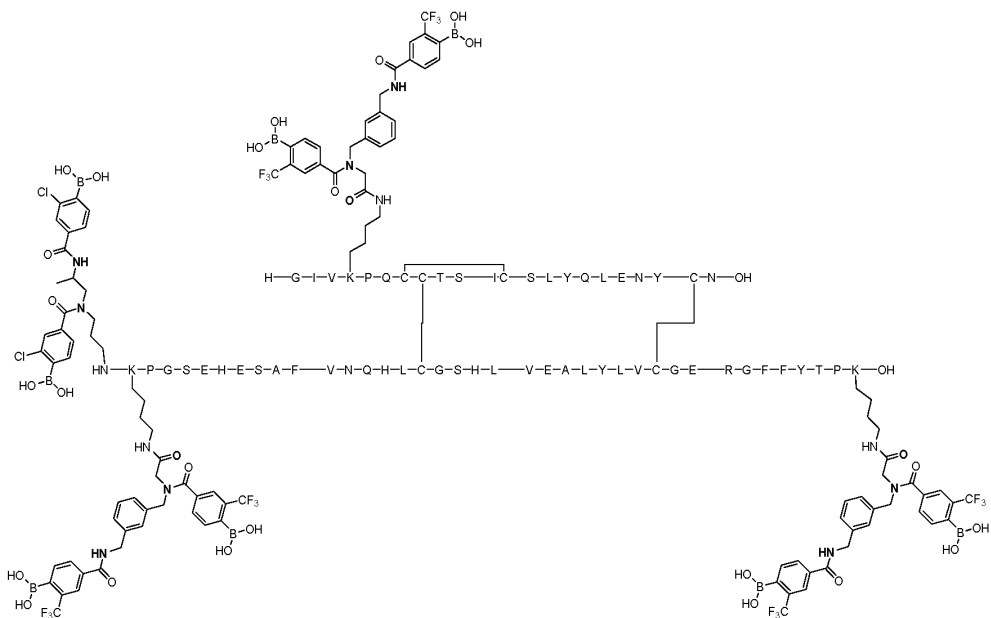
Example 199:**Example 200:**

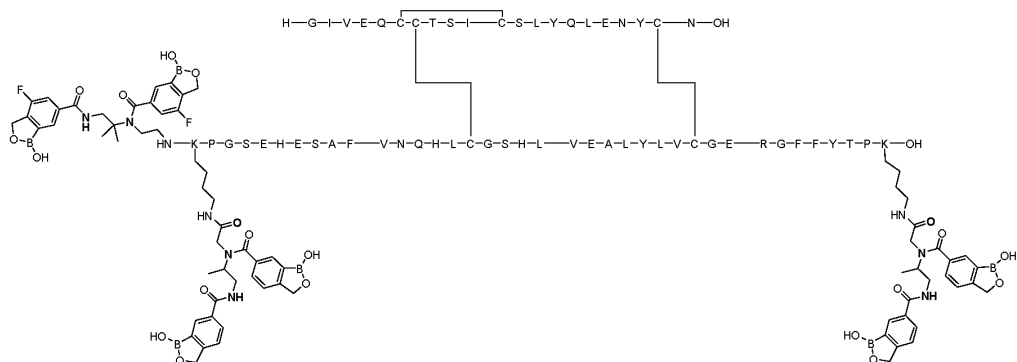
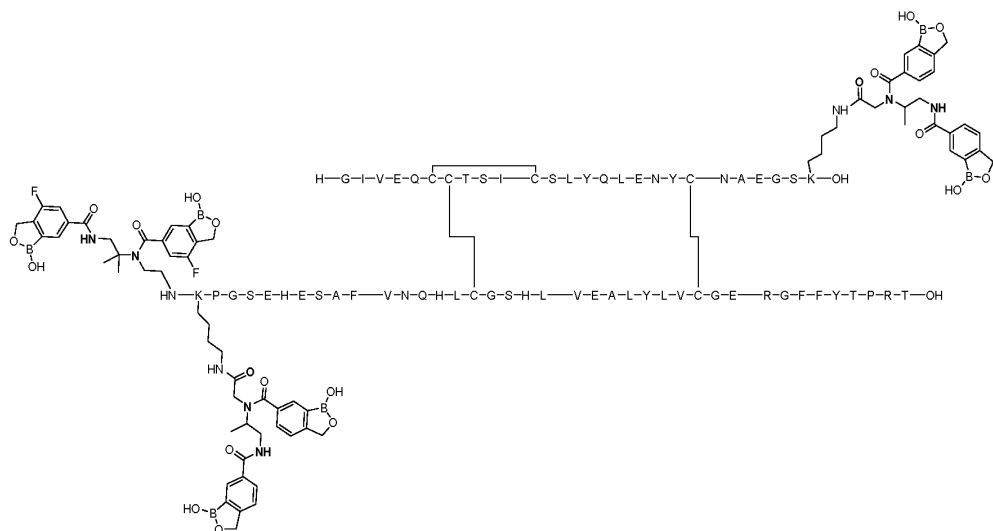
Example 201:**Example 202:**

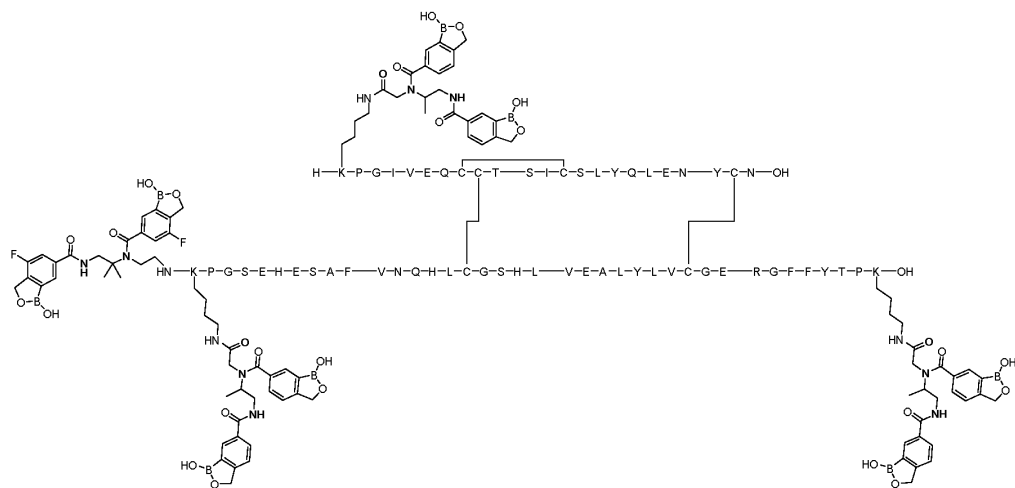
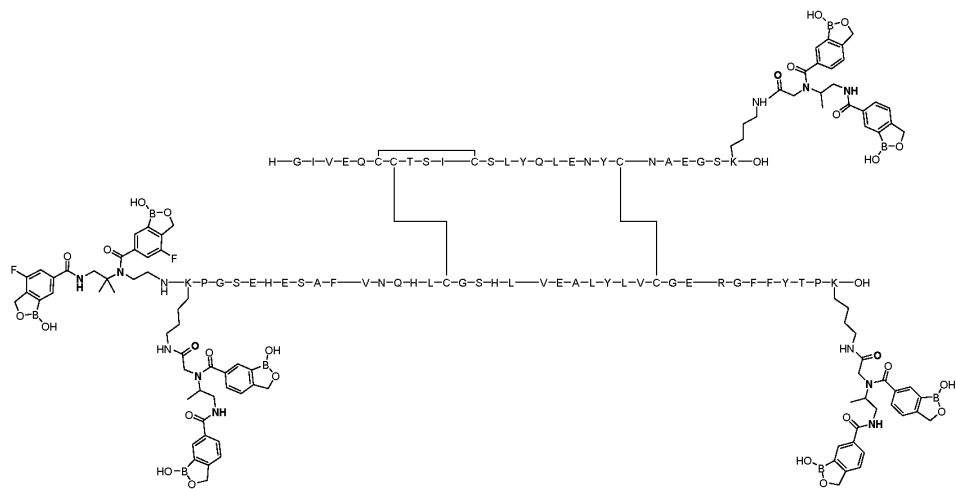
The chemical structure of the 12th generation BSLA peptide is shown. It features a central peptide backbone with the sequence: H-G-I-V-K-P-Q-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH. The side chains are highly complex and include fluorinated boronic acid groups, a cyclic peptide structure, and a complex cyclic peptide structure. The structure is highly branched and includes various functional groups such as carboxylic acids, amides, and boronic acids.

Chemical structure of the 100-residue protein 100A, showing its primary structure with various modifications. The protein sequence is H-G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH. The structure includes a disulfide bond between Cys2 and Cys10, and a phosphorylation site at Ser5. The N-terminus is modified with a 4-chlorophenylboronic acid group. The C-terminus is modified with a 4-(4-chlorophenyl)boronic acid group. The protein is also modified with a 4-(4-chlorophenyl)boronic acid group at the C-terminus. The structure is shown in a ribbon representation with various modifications highlighted in different colors.

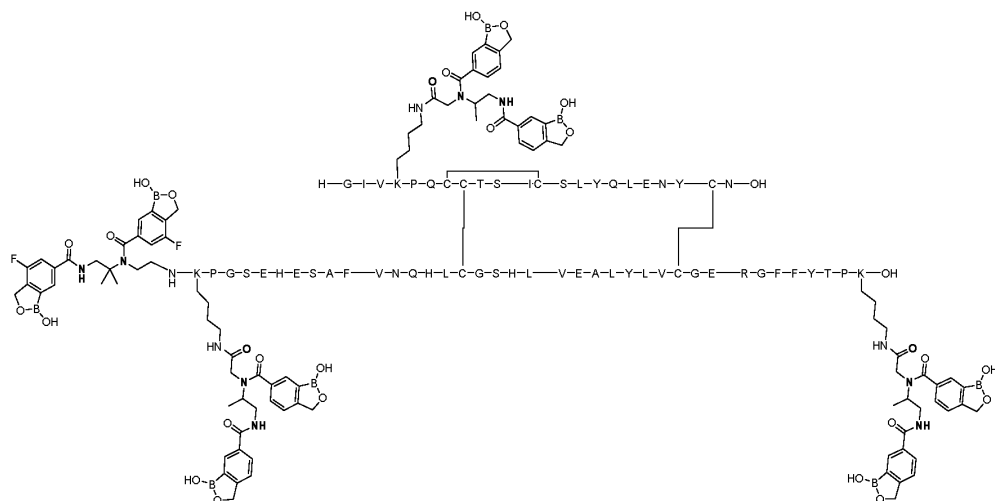
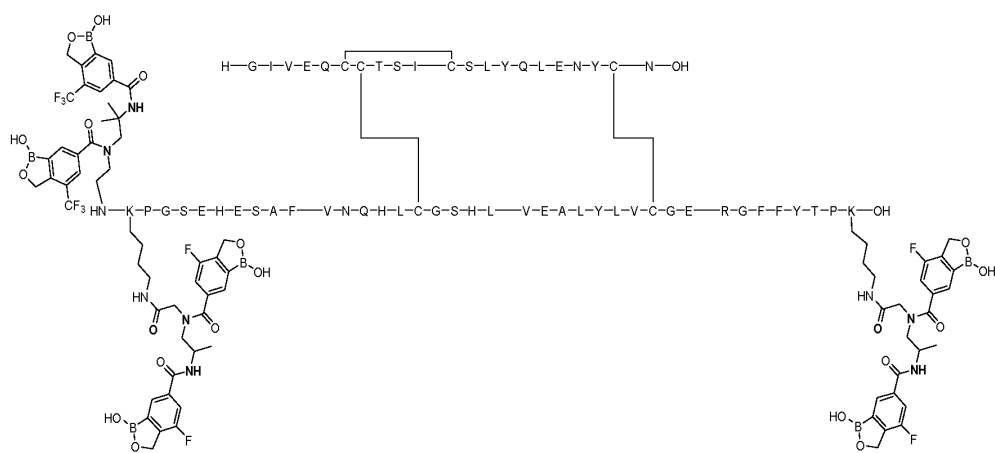
Example 205:**Example 206:**

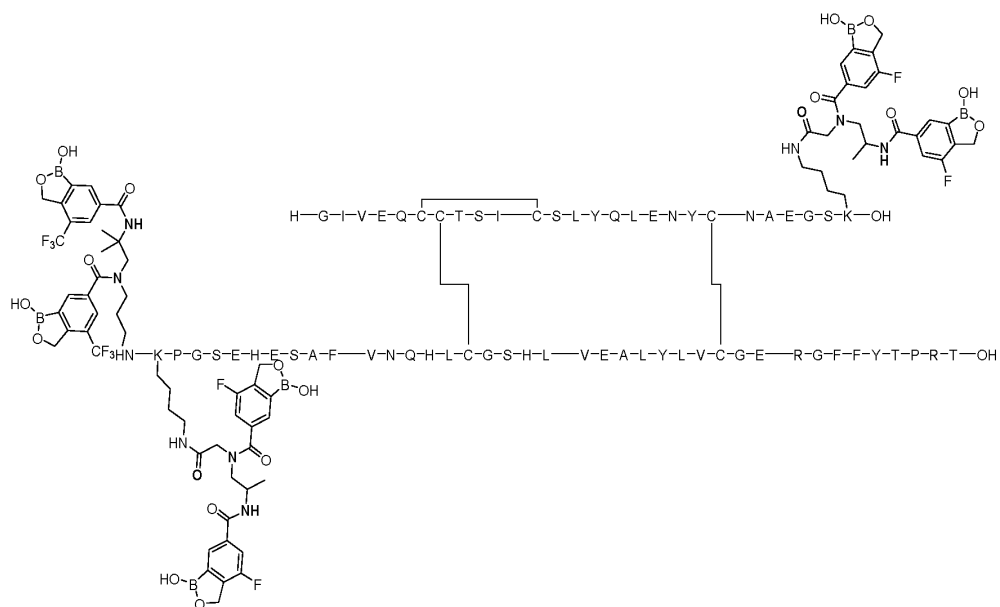
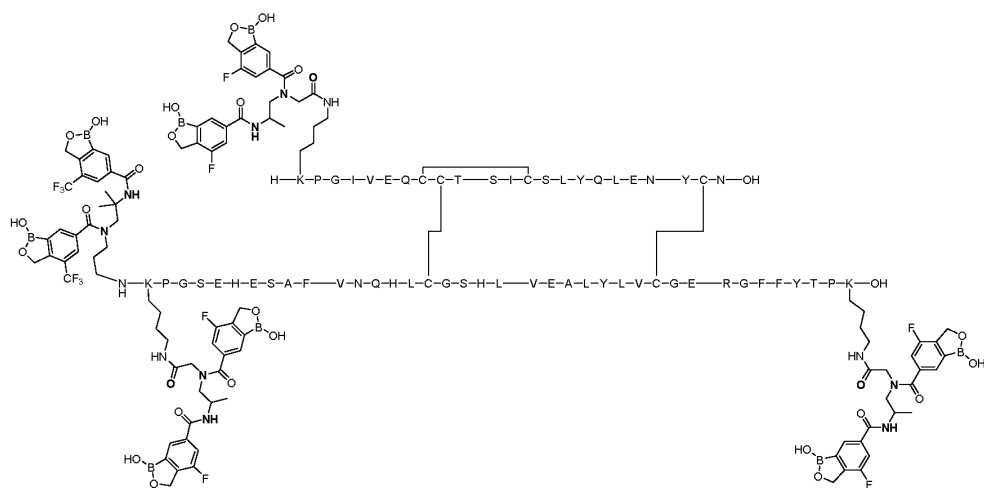
Example 207:**Example 208:**

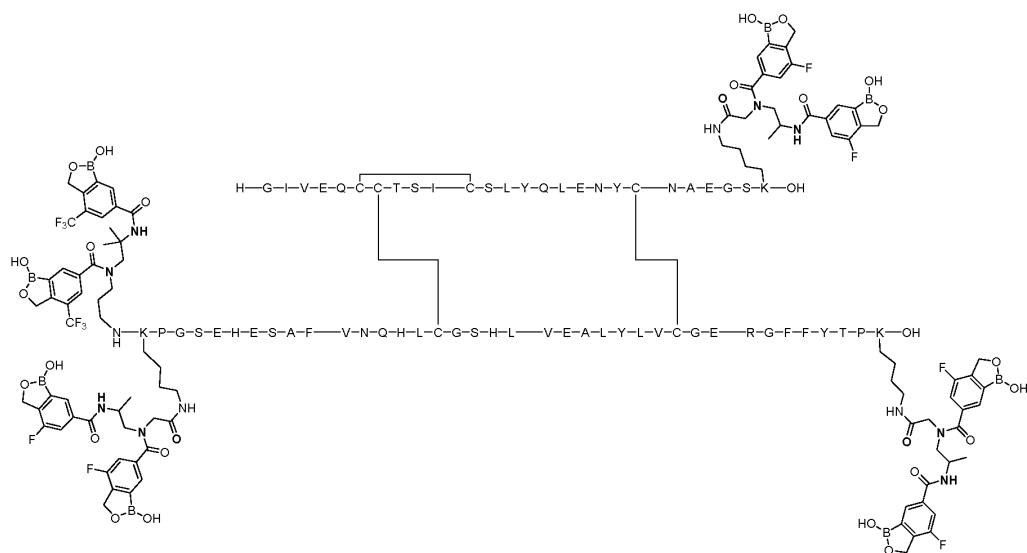
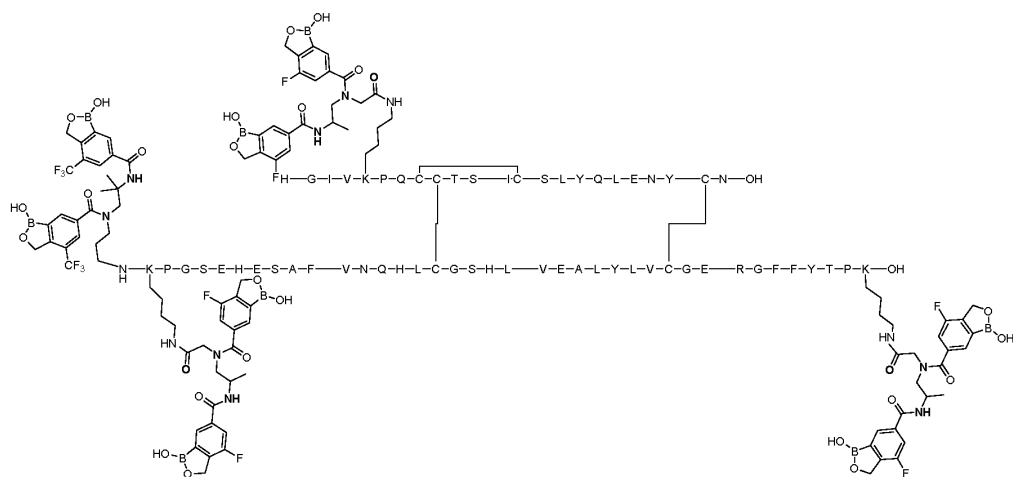
Example 209:**Example 210:**

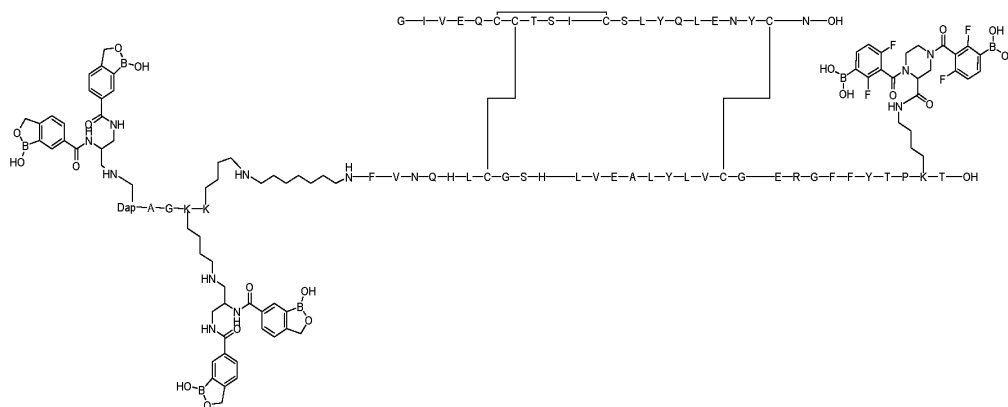
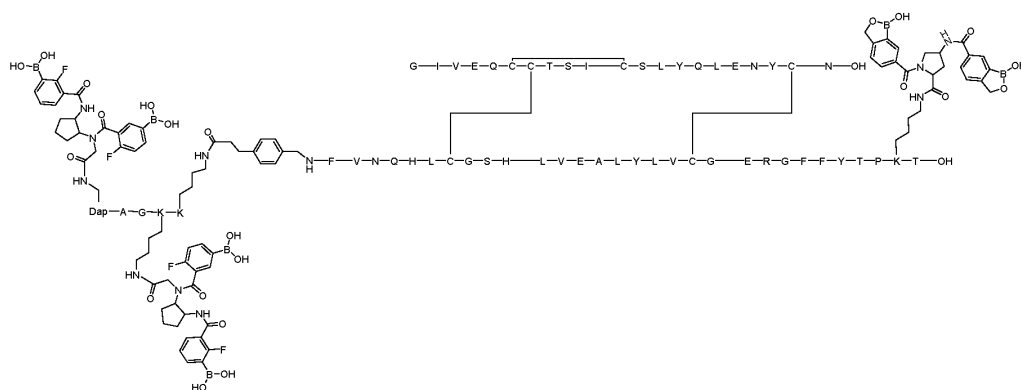
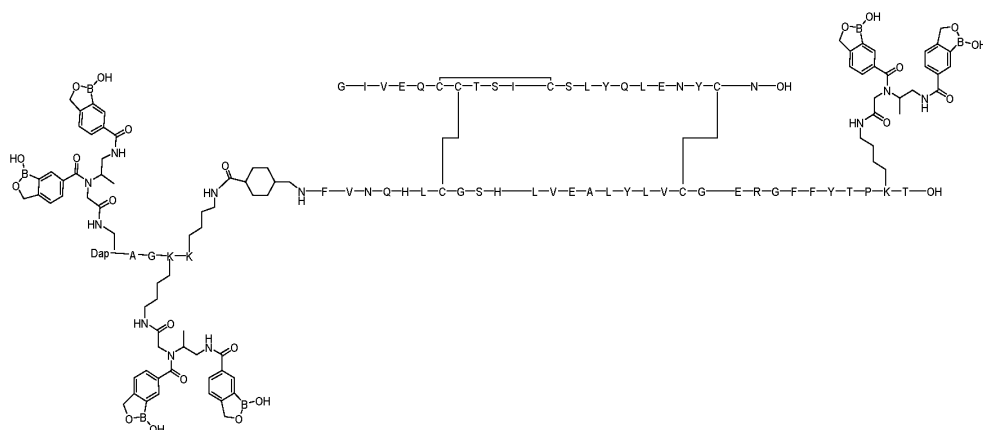
Example 211:**Example 212:**

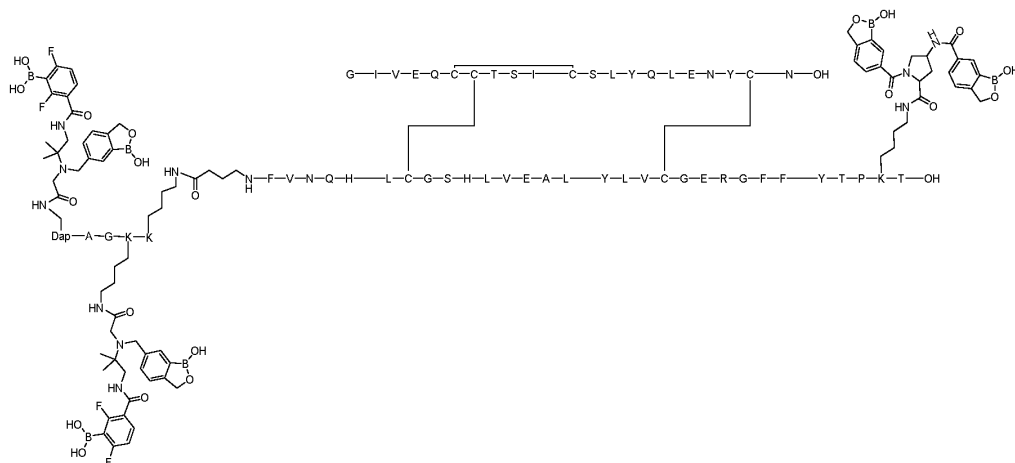
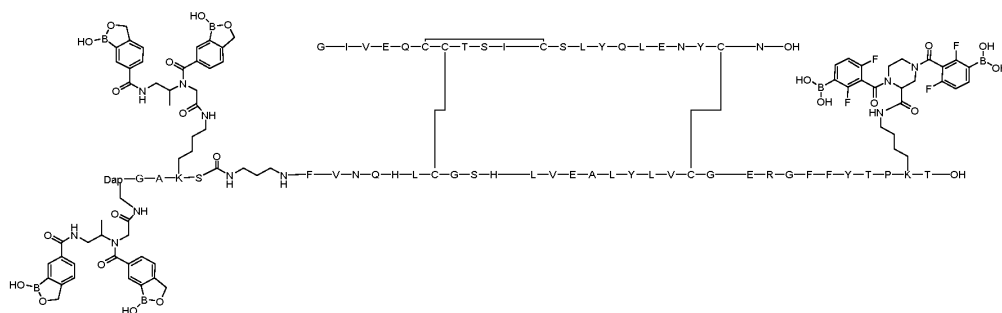
Example 213:

**Example 214:**

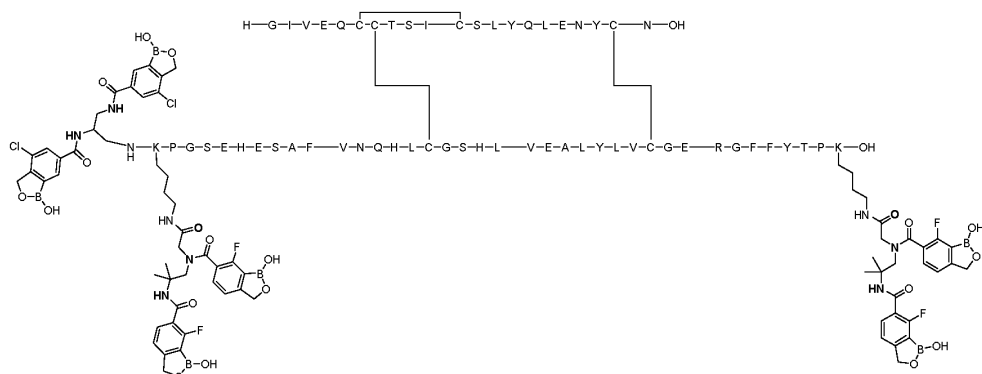
Example 215:**Example 216:**

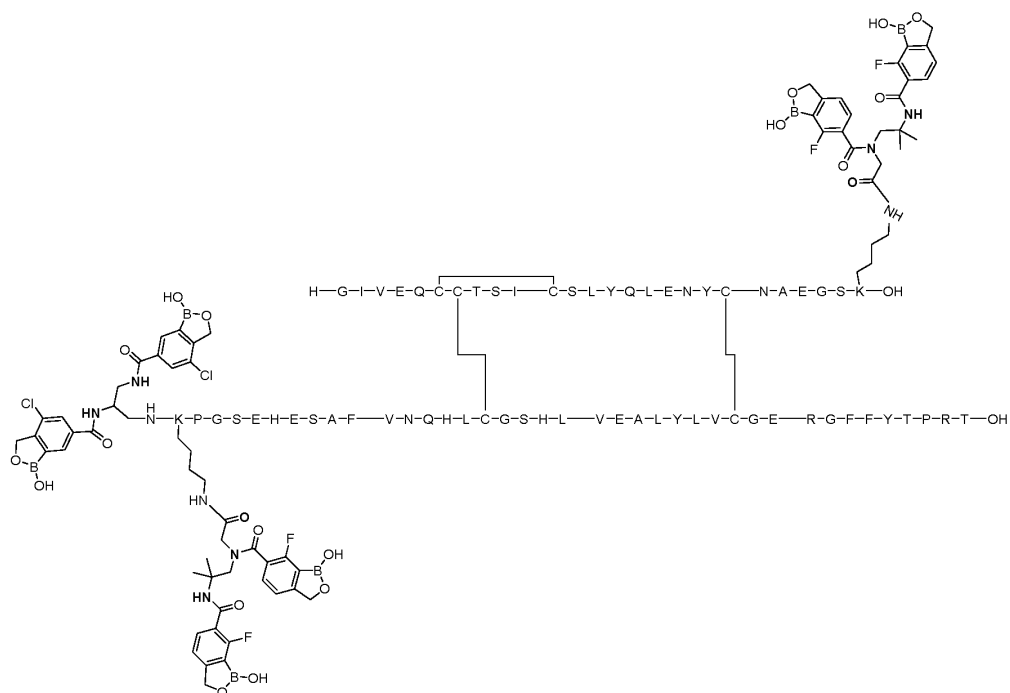
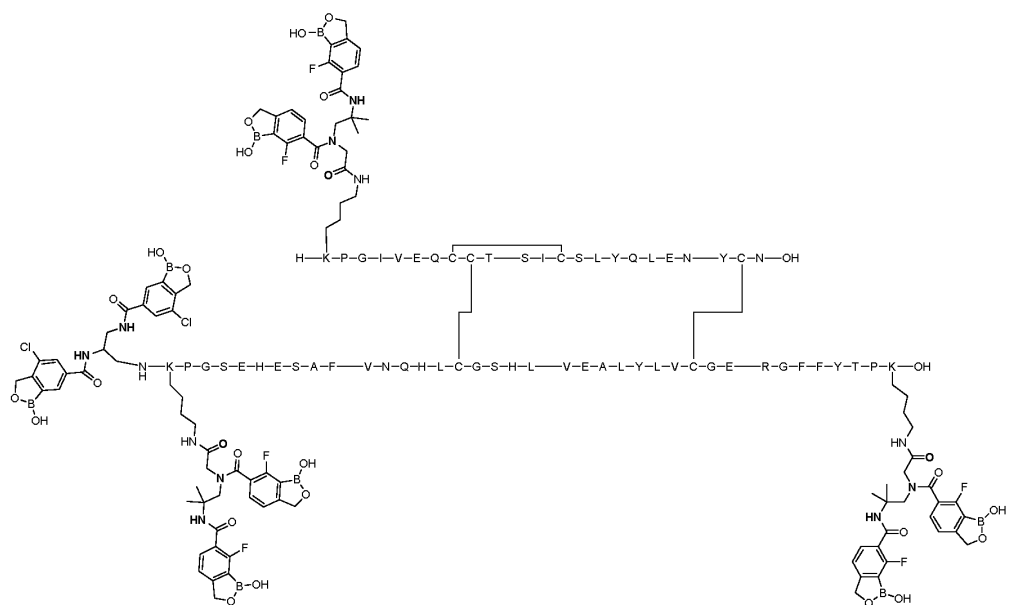
Example 217:**Example 218:**

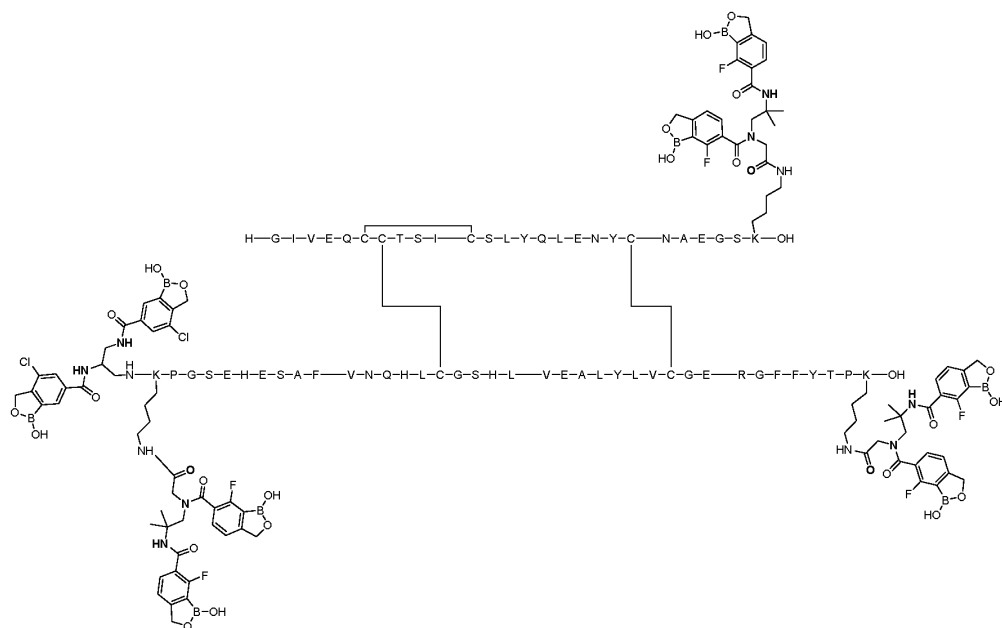
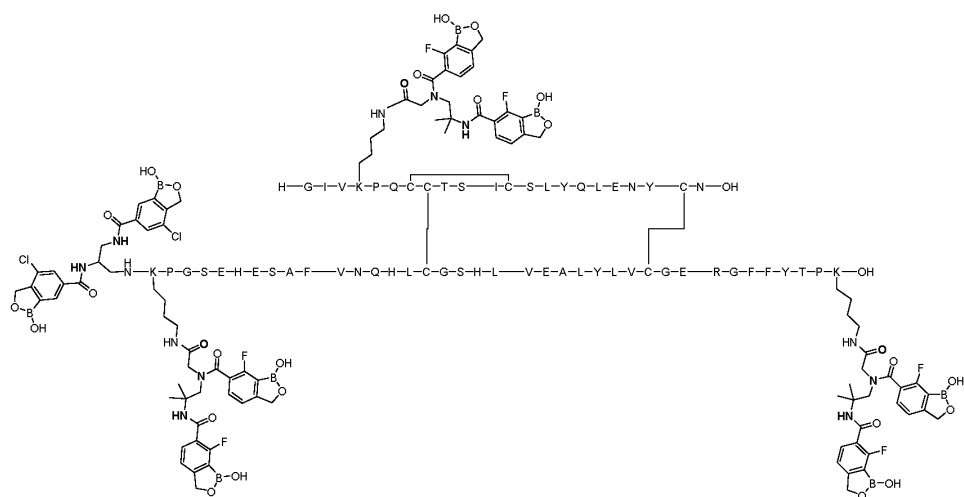
Example 219:**Example 220:****5 Example 221:**

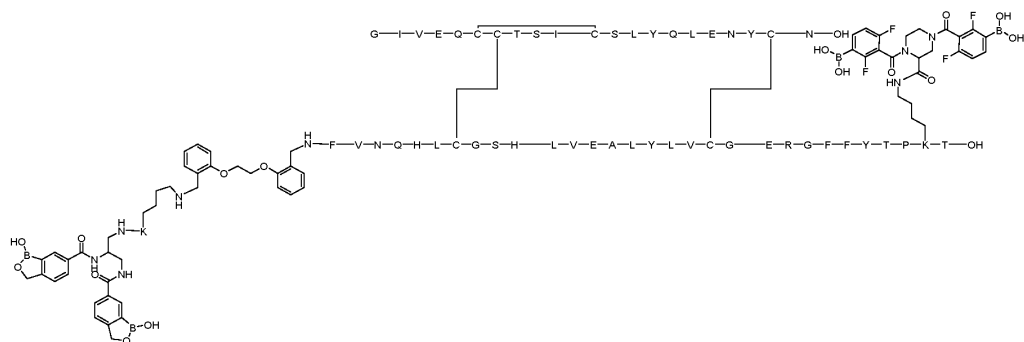
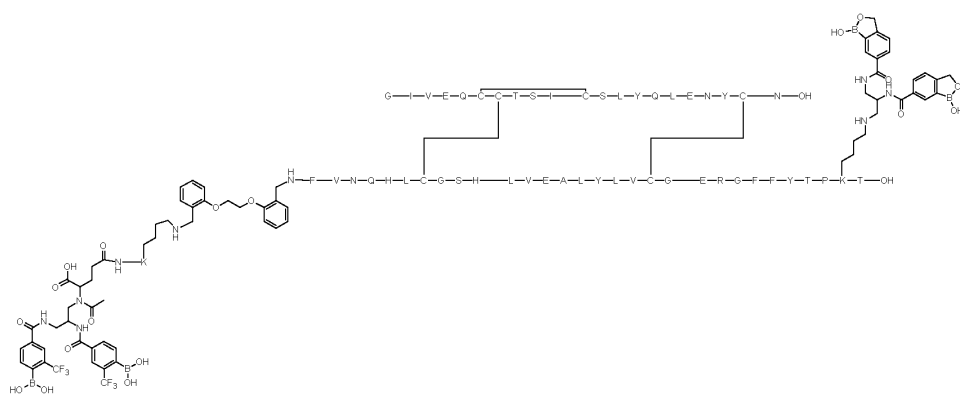
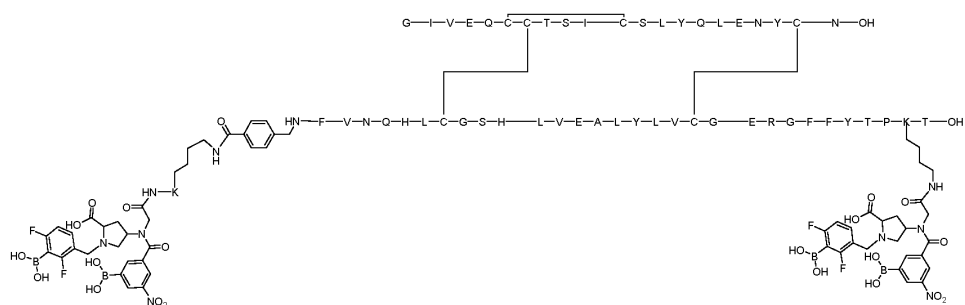
Example 222:**Example 223:**

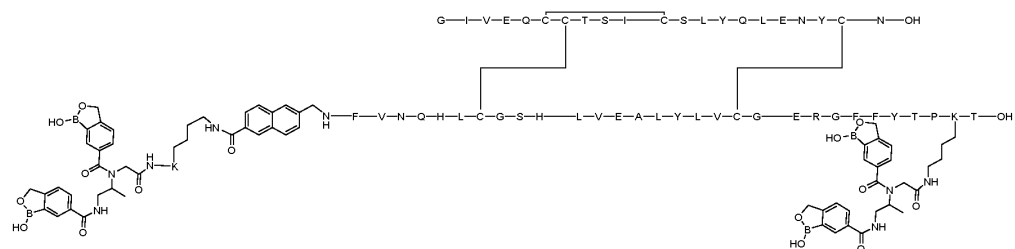
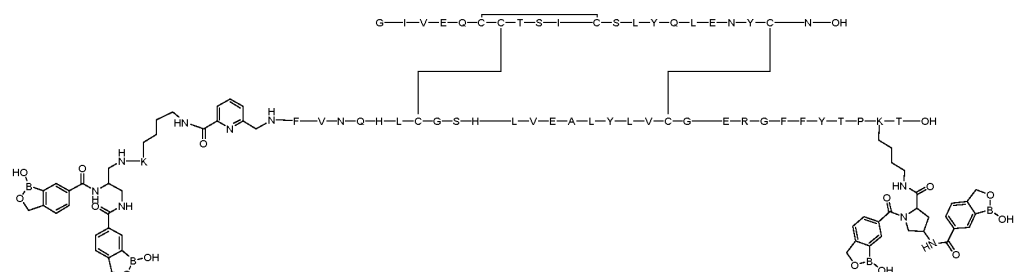
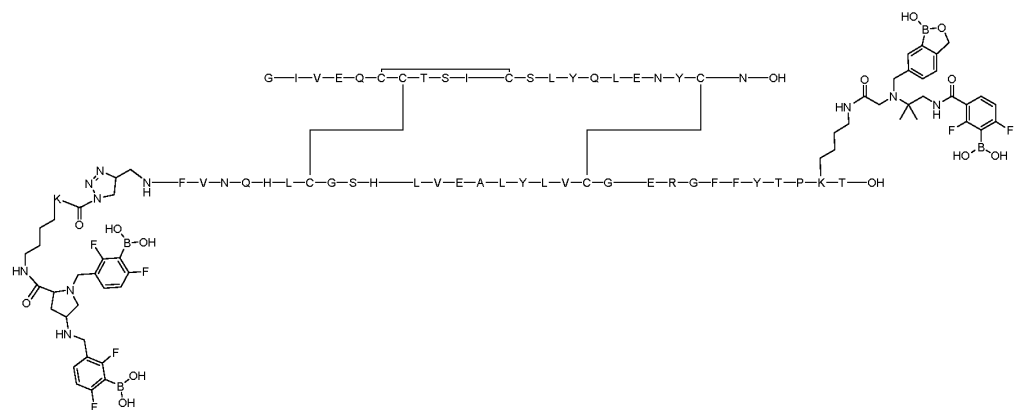
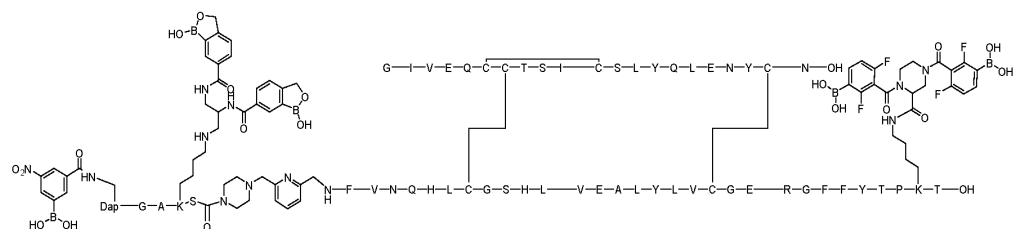
5

Example 224:

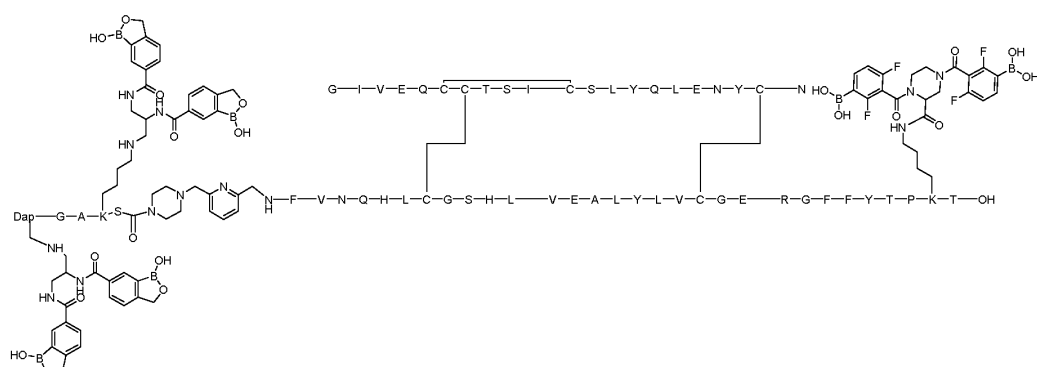
Example 225:**Example 226:**

Example 227:**Example 228:**

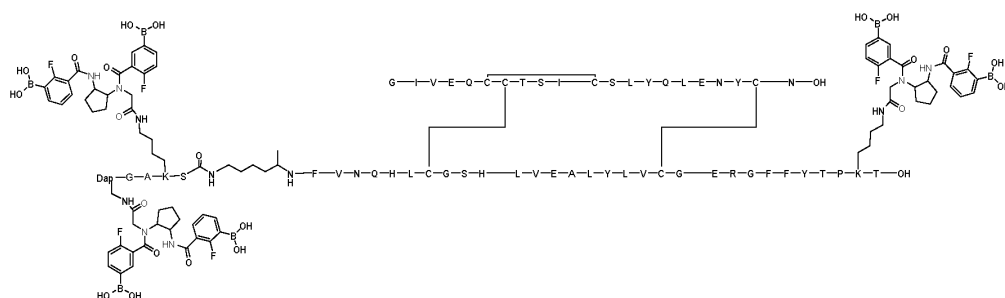
Example 229:**Example 230:****5 Example 231:**

Example 232:**Example 233:****5 Example 234:****Example 235:**

Example 236:

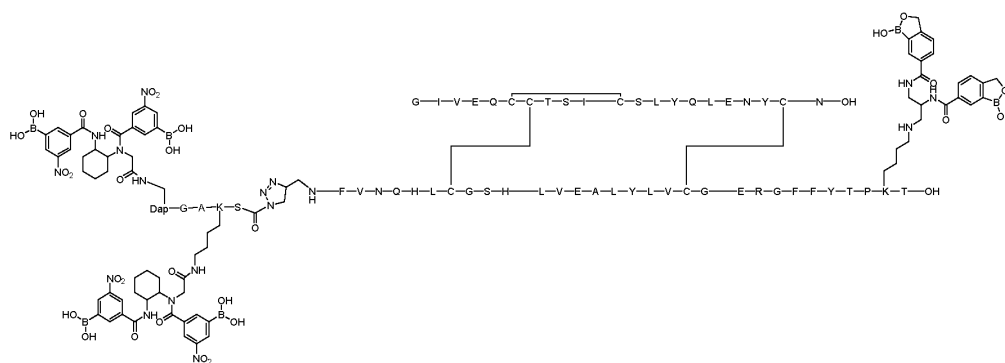


Example 237:

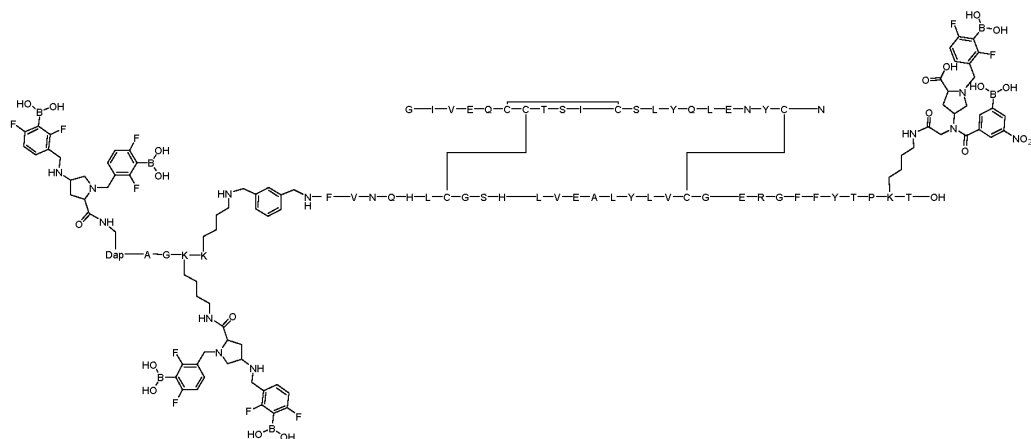


5

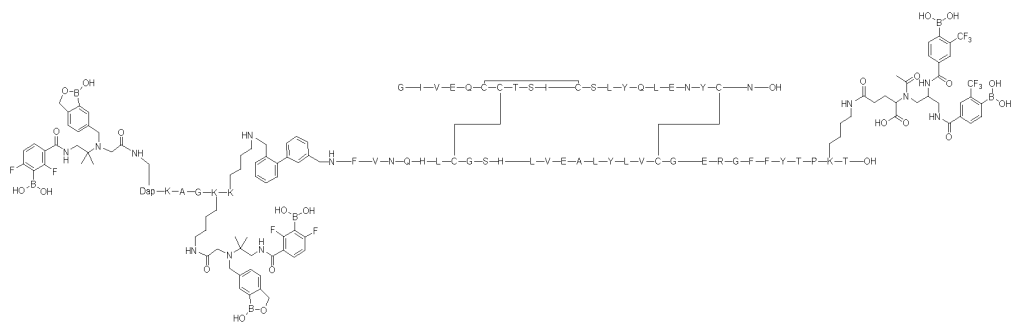
Example 238:



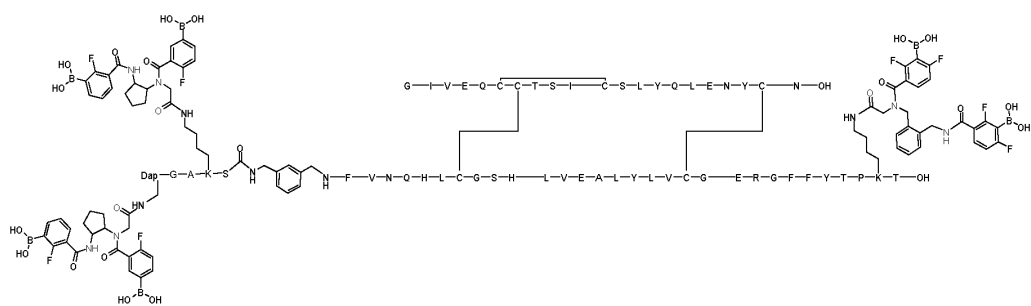
Example 239:



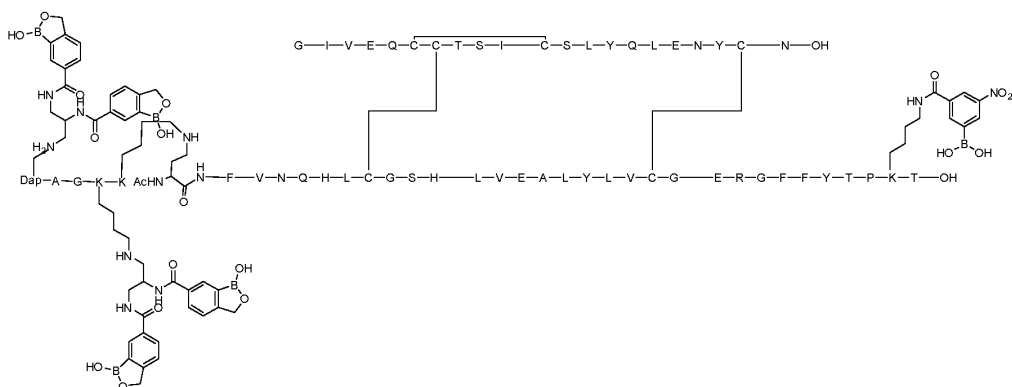
Example 240:



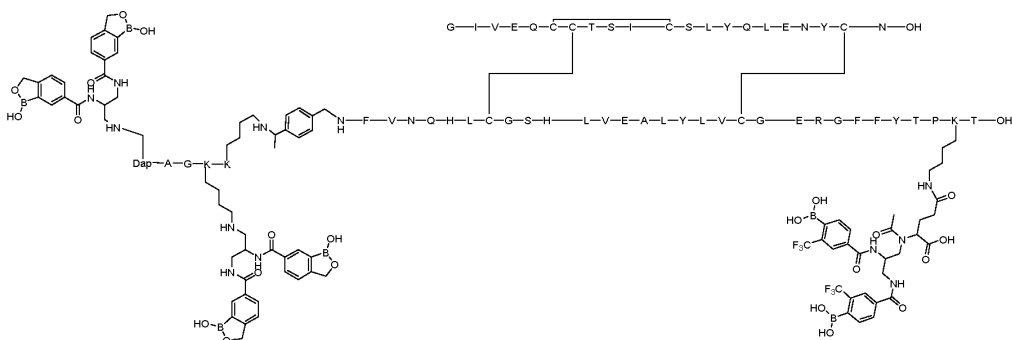
5 Example 241:



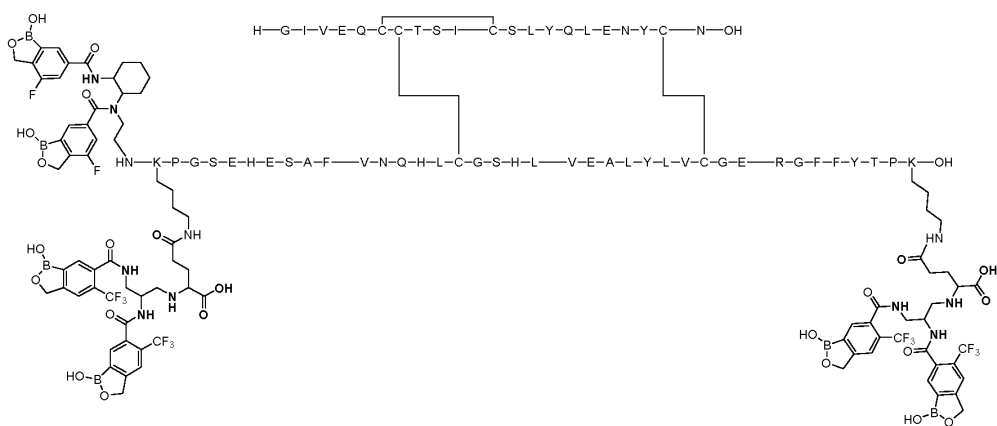
Example 242:

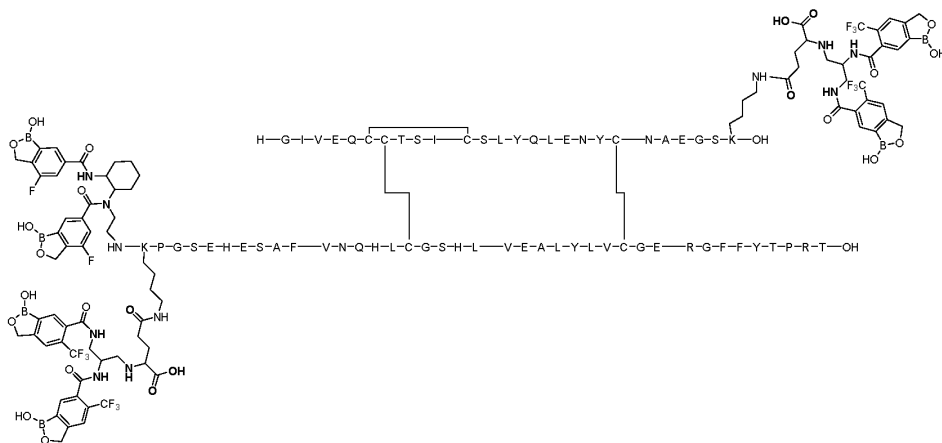
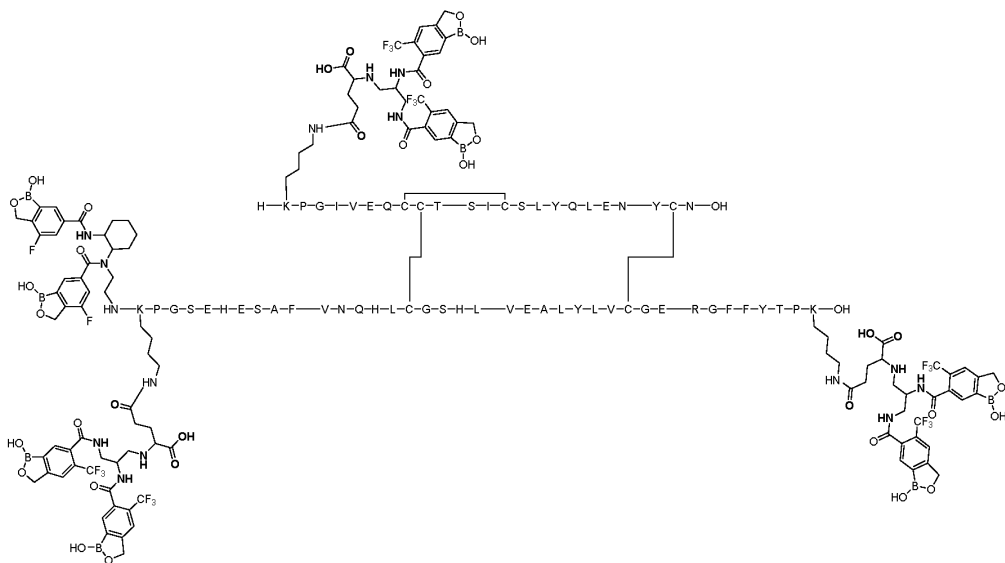


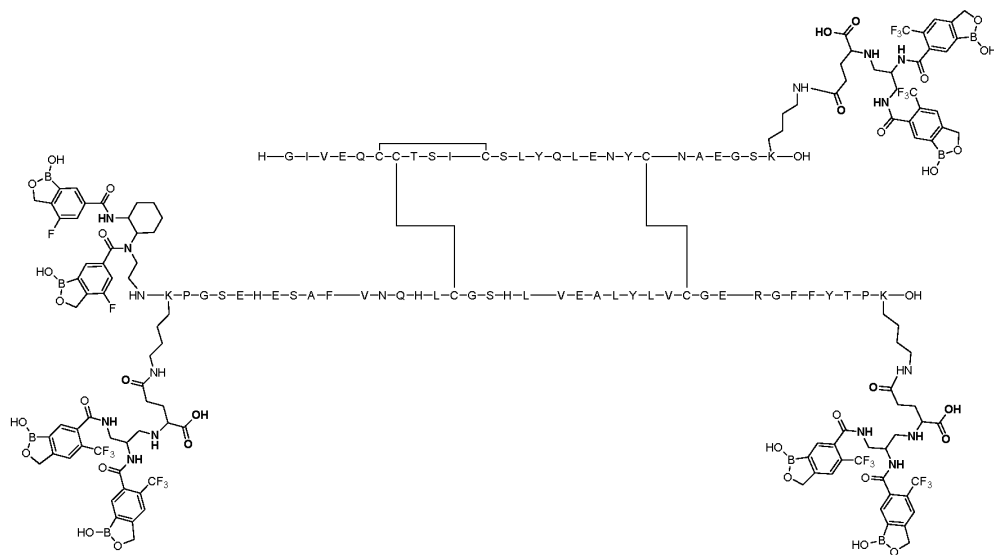
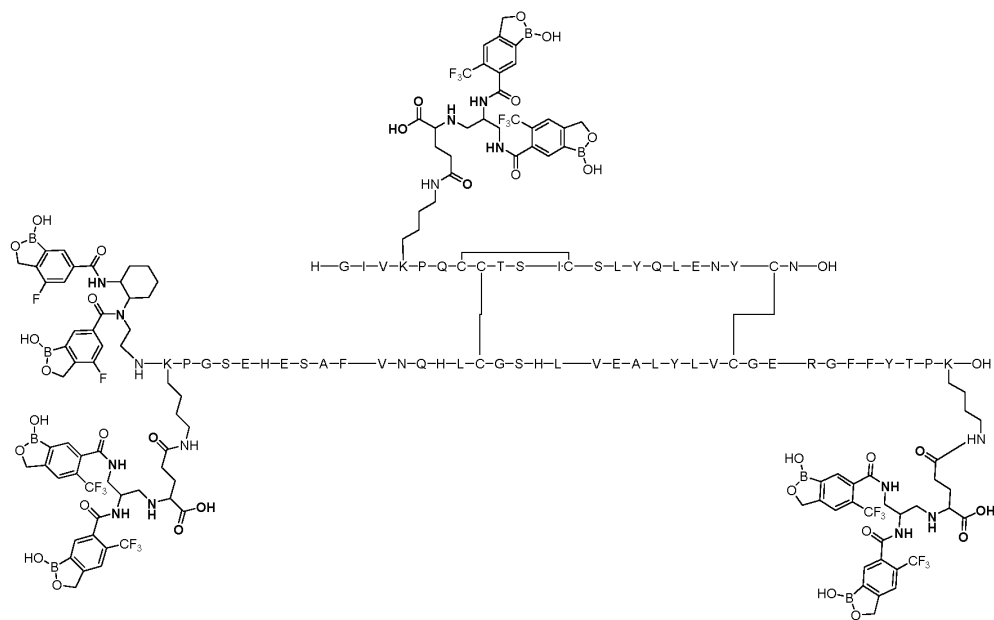
Example 243:

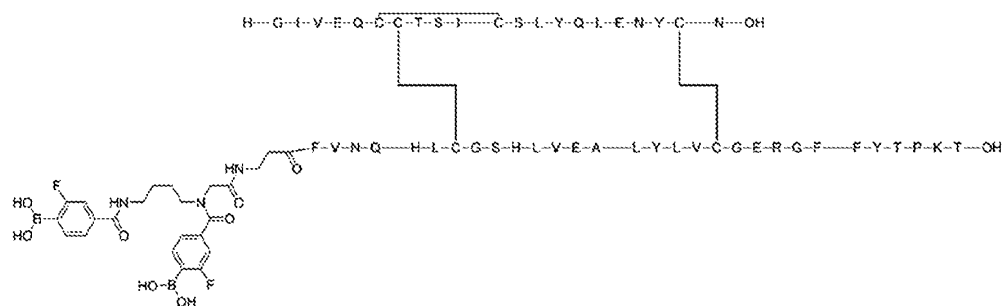
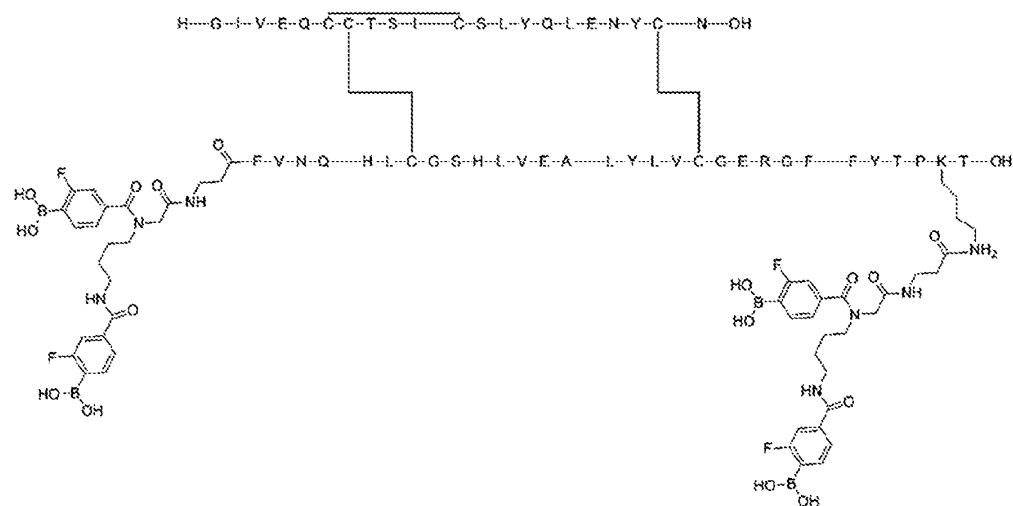


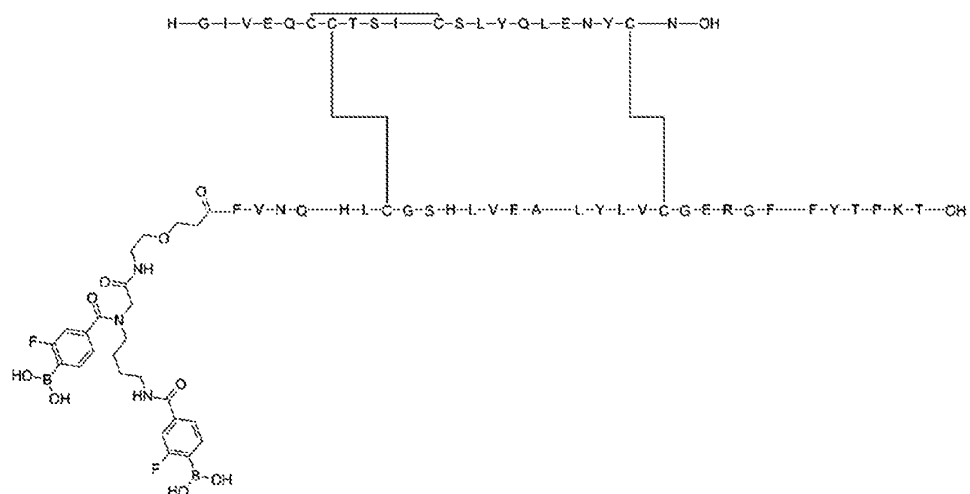
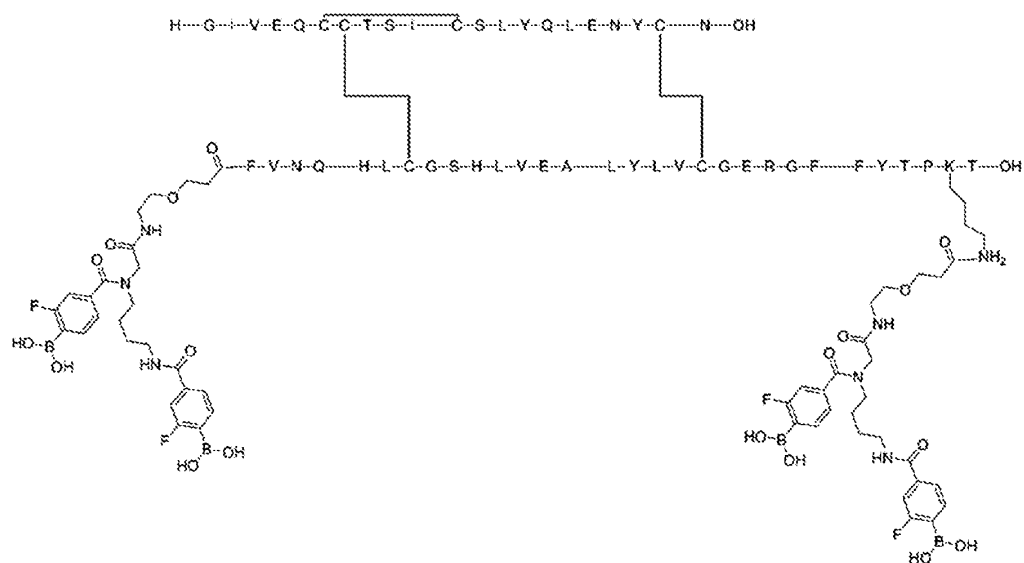
5 Example 244:

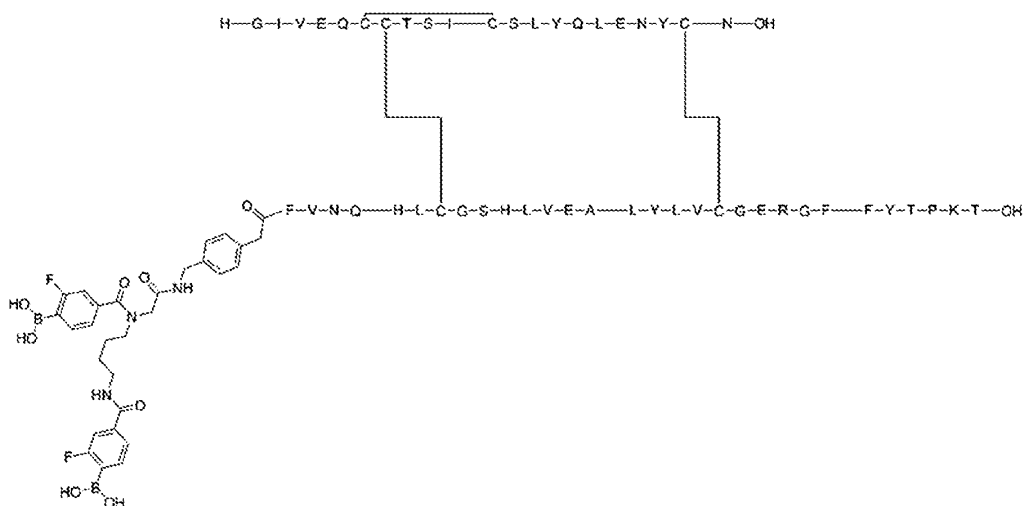
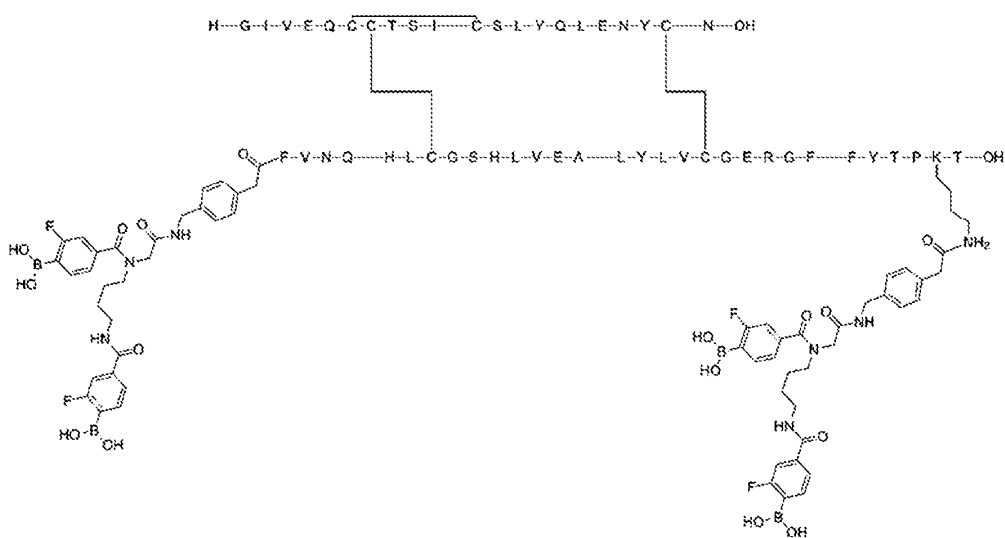


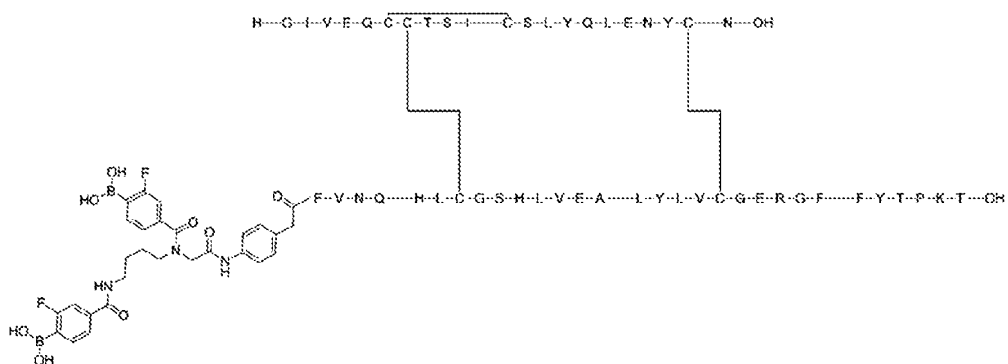
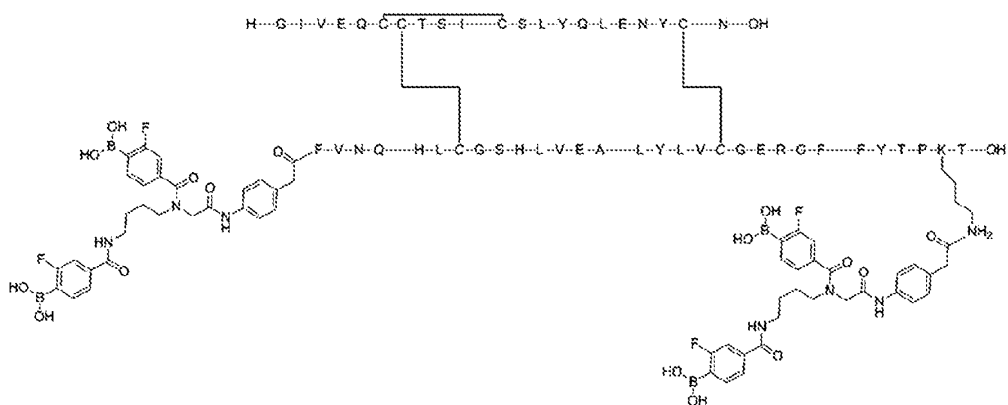
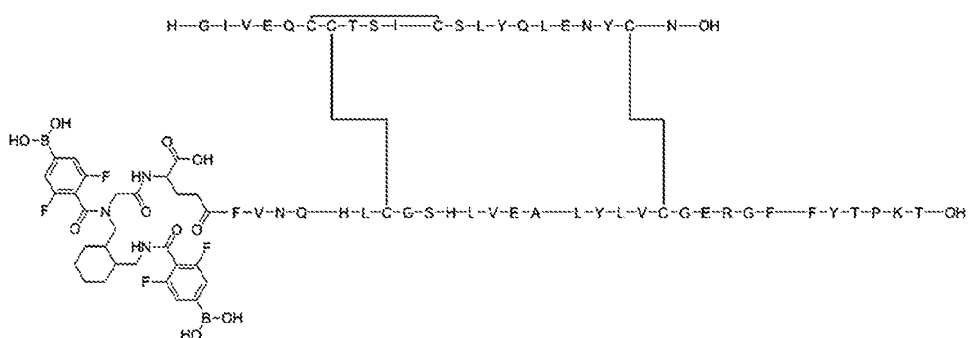
Example 245:**Example 246:**

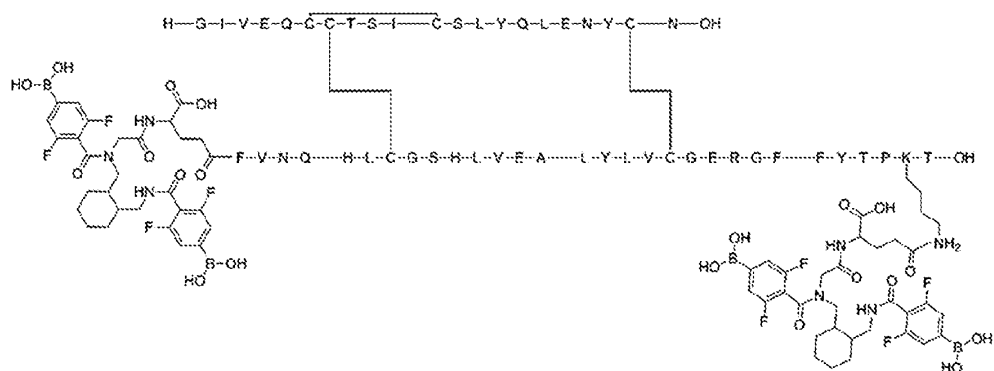
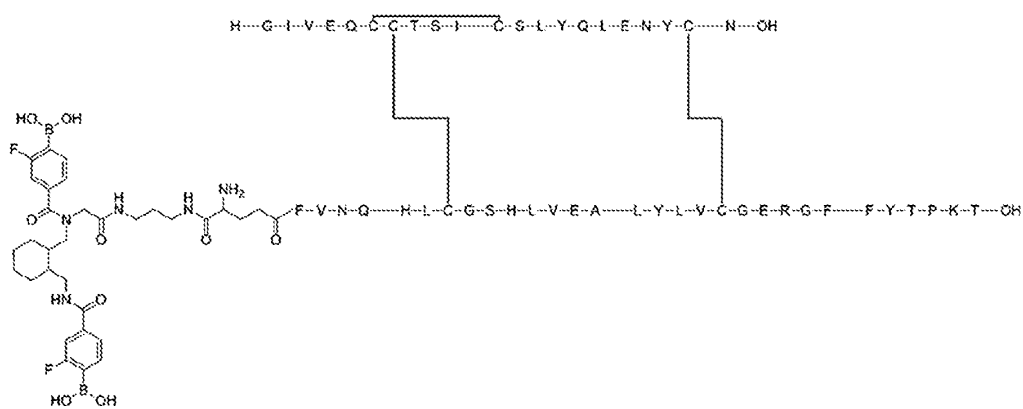
Example 247:**Example 248:**

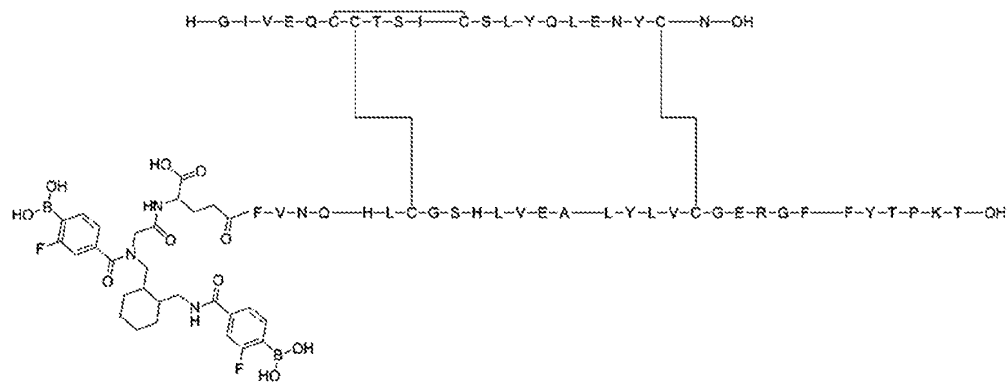
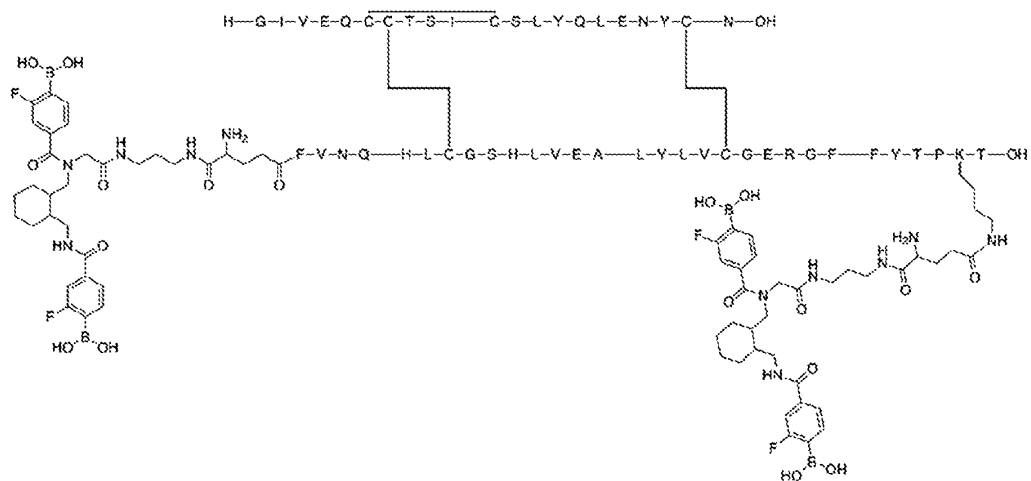
Example 249:**Example 250:**

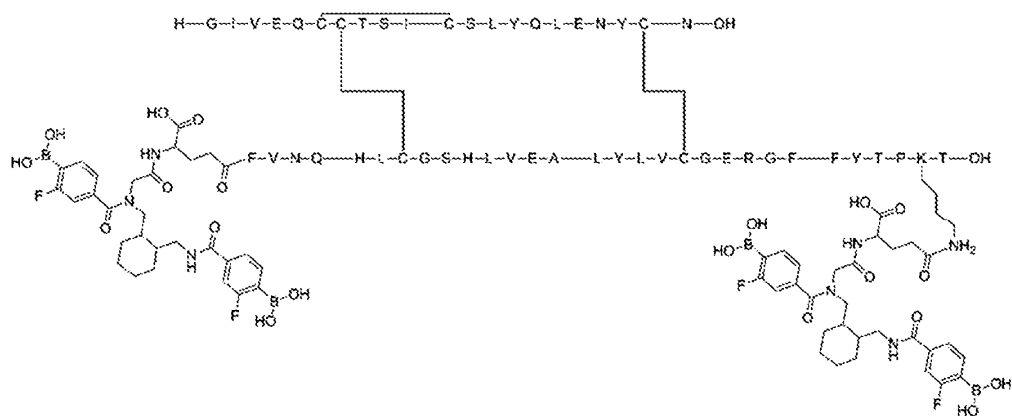
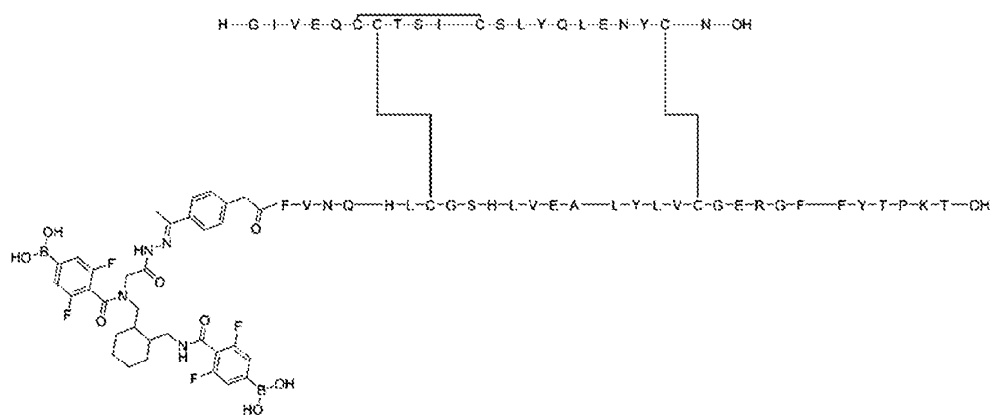
Example 251:**Example 252:**

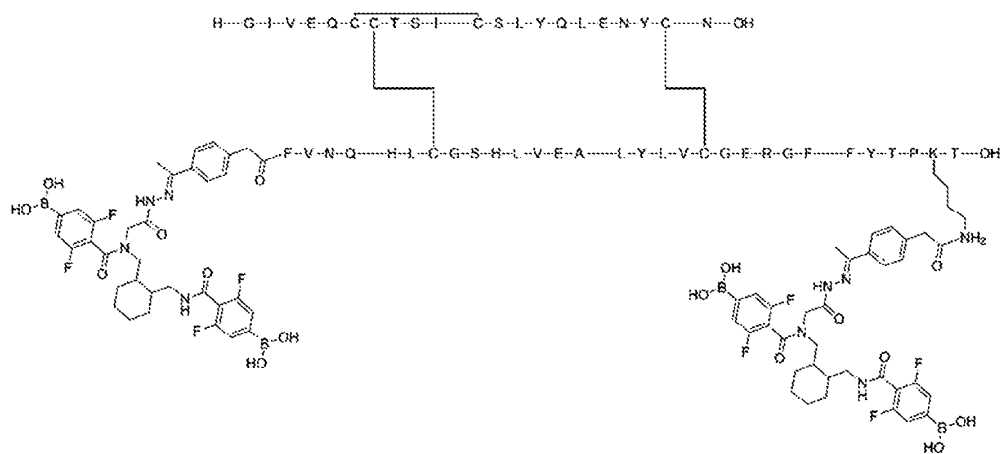
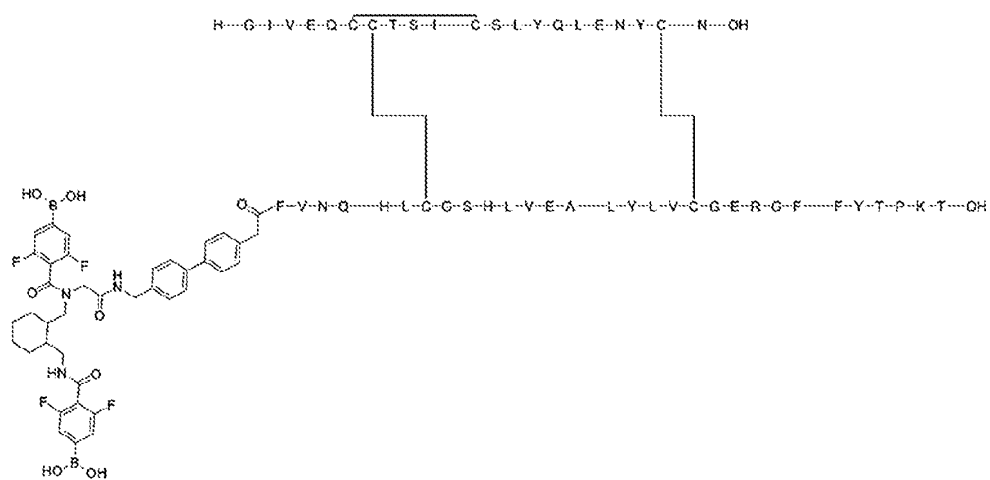
Example 253:**Example 254:**

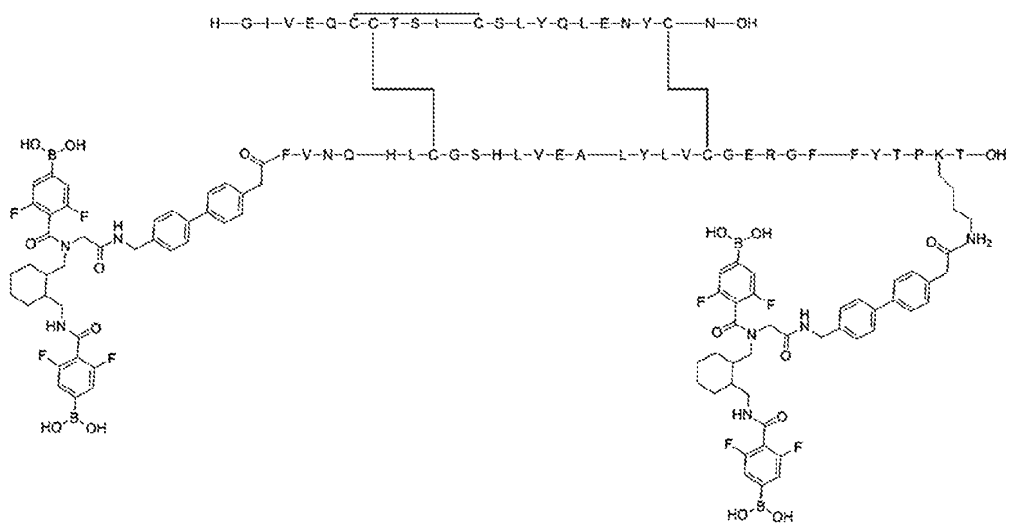
Example 255:**Example 256:****5 Example 257:**

Example 258:**Example 259:**

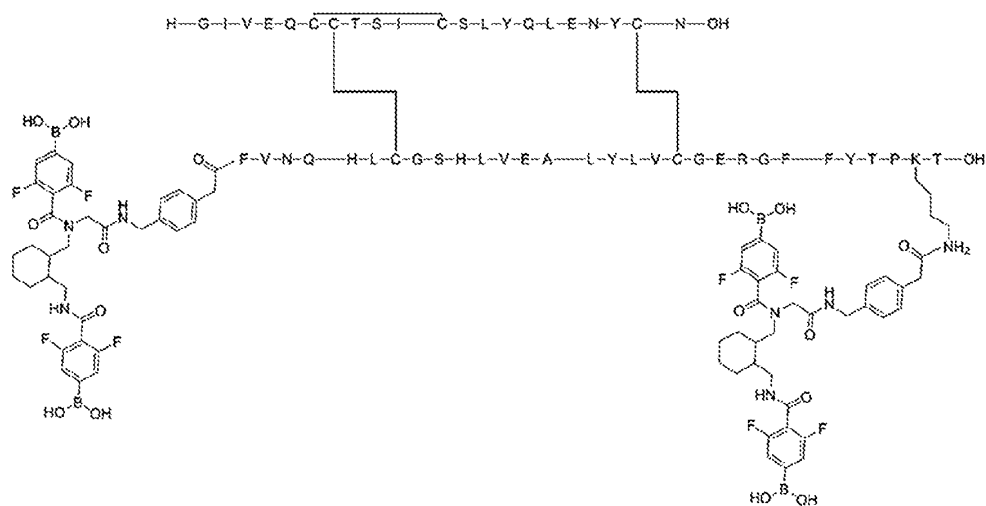
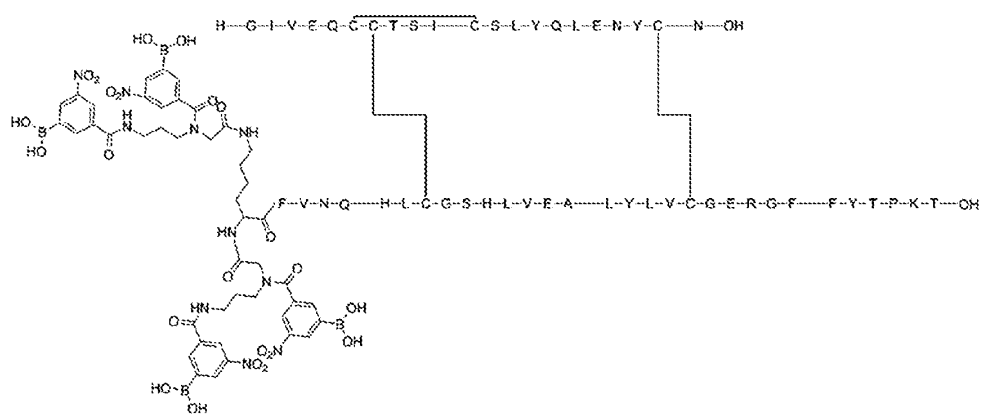


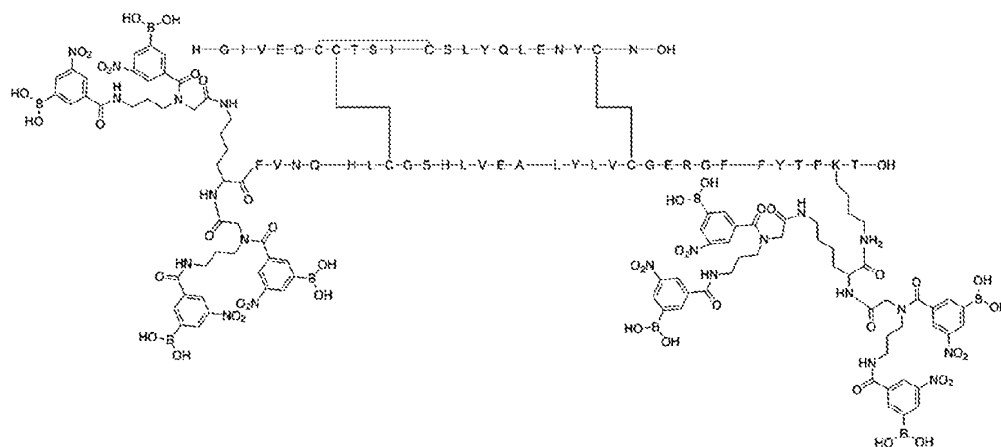
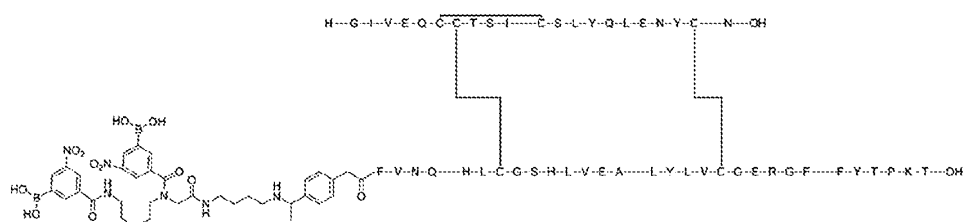
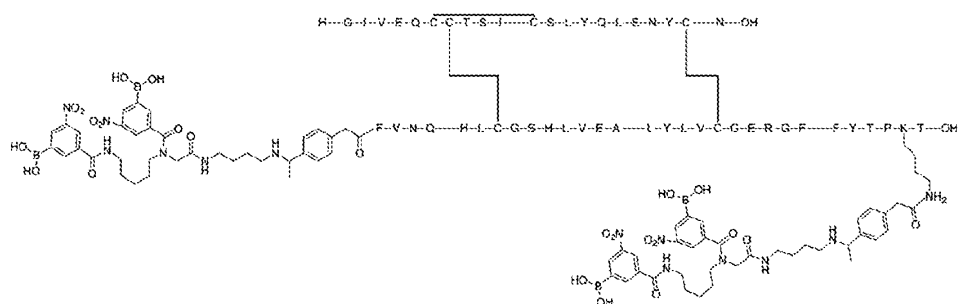
Example 262:**Example 263:**

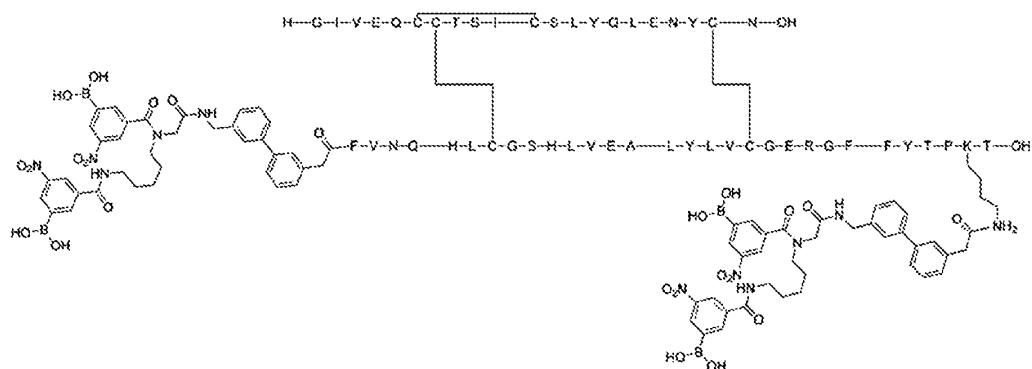
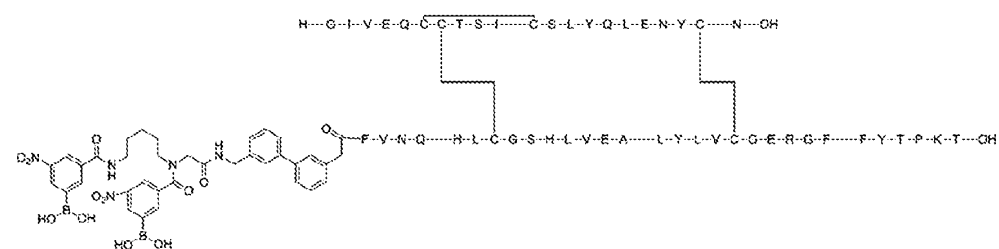
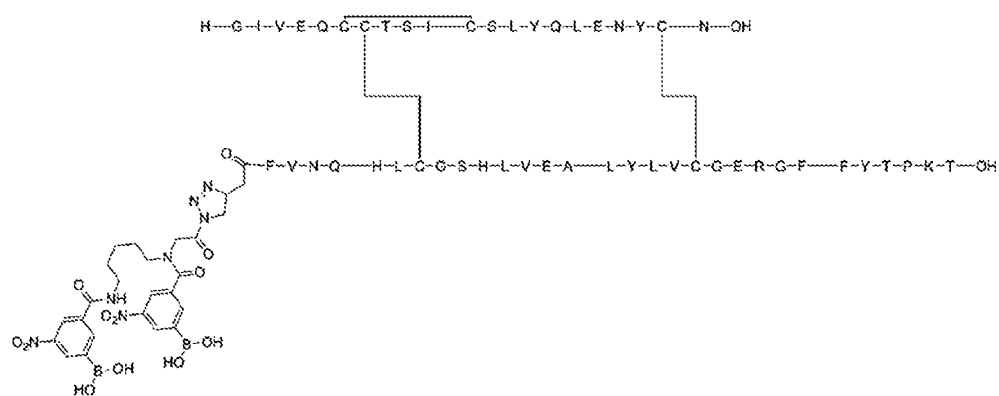
Example 264:**Example 265:**

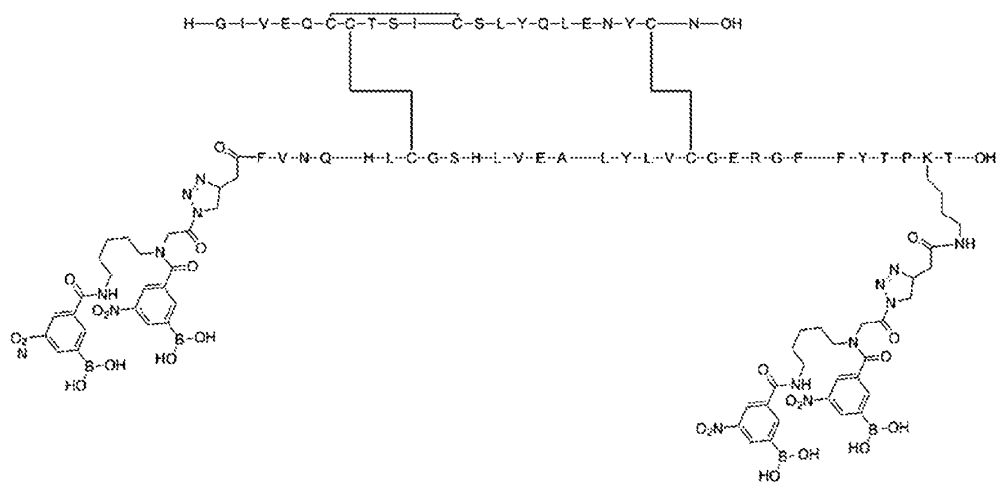
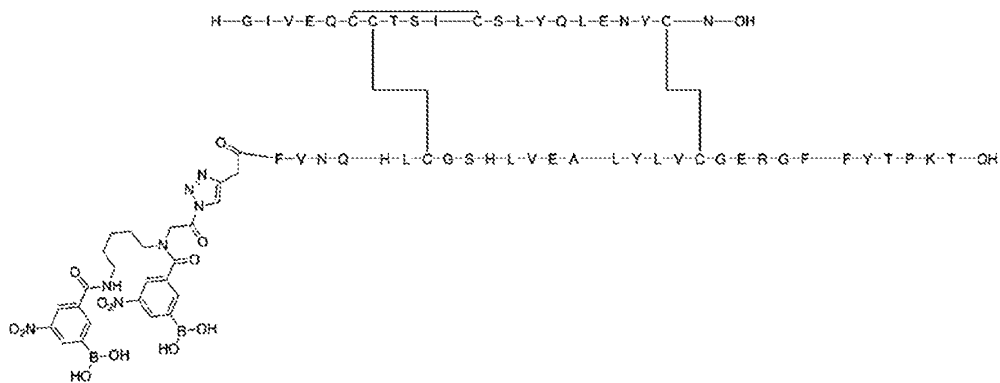


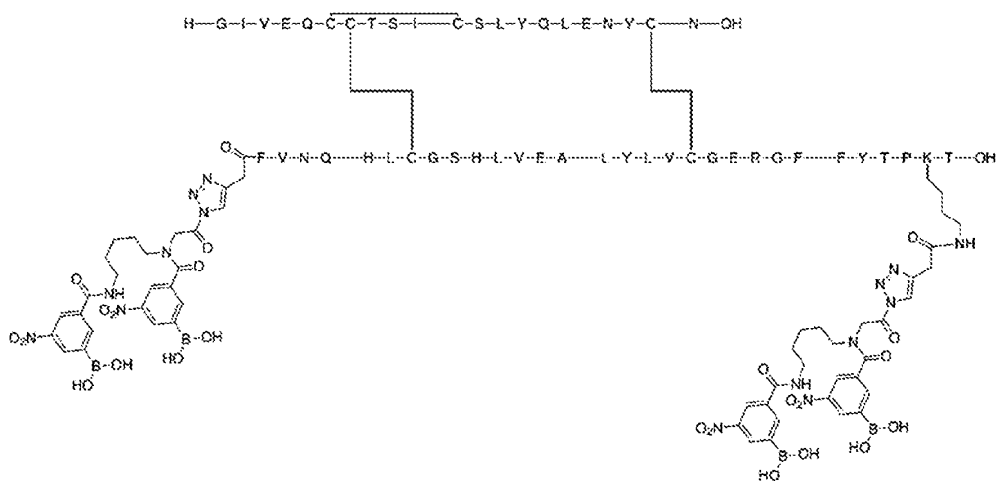
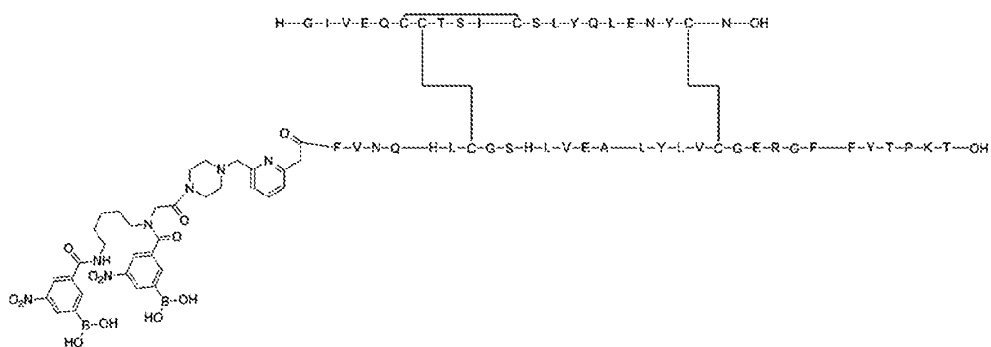
Chemical structure of a peptide-drug conjugate. The peptide sequence is H-G-I-V-E-Q-C-C-T-S-I-C-S-L-V-G-L-F-N-Y-C-N-OH. The Cysteine at position 10 is linked via a disulfide bridge to a Cysteine at position 19 of a second chain: F-V-N-Q-H-L-C-G-S-H-L-V-E-A-L-Y-L-V-C-G-E-R-G-F-F-Y-T-P-K-T-OH. The Cysteine at position 10 is also linked via a thioether bond to a complex molecule containing a 2,4,6-trifluorophenyl group, a cyclohexylmethyl group, and a 2,4,6-trifluorophenylboronic acid group.

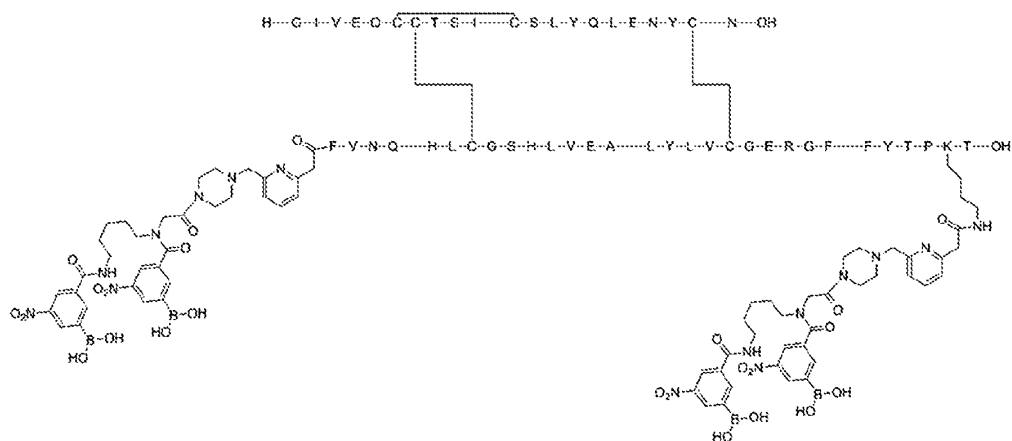
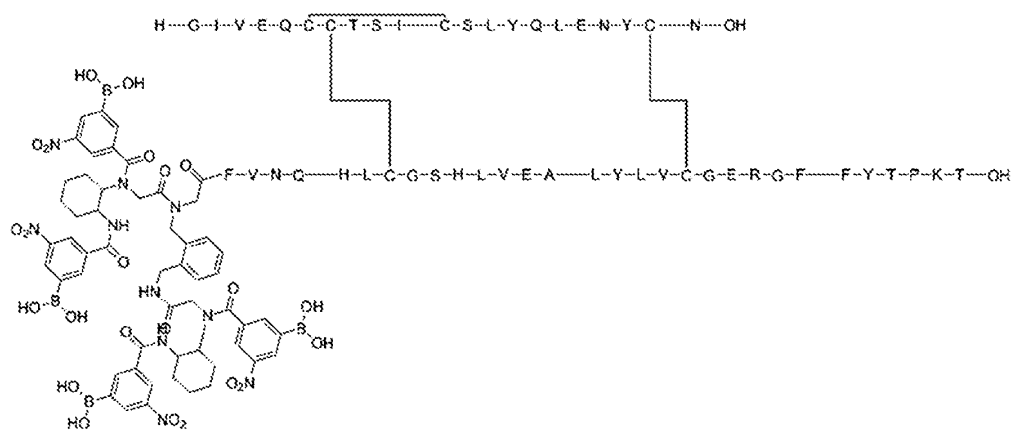
Example 268:**Example 269:**

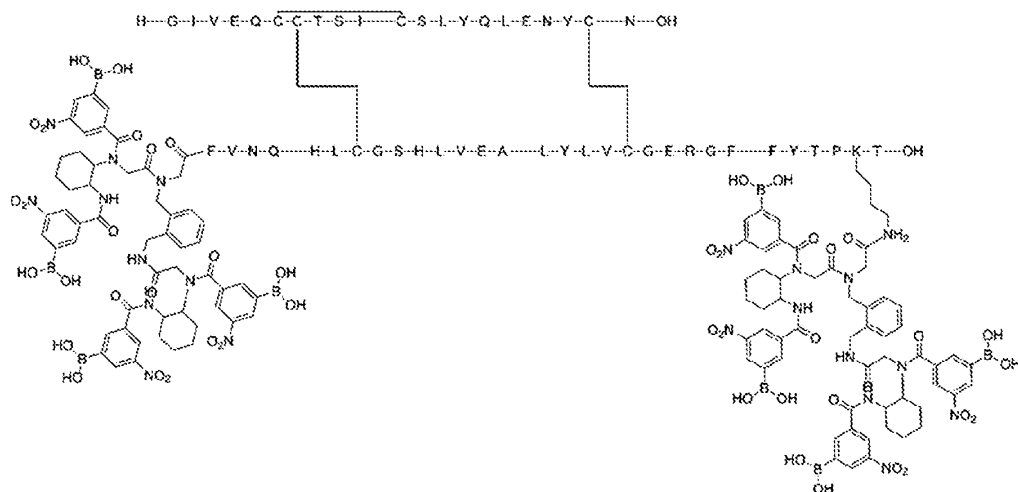
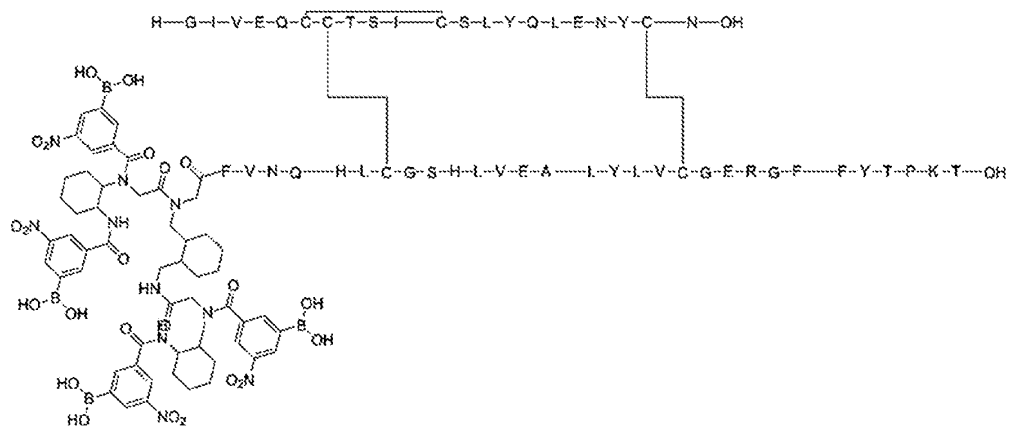
Example 270:**Example 271:****5 Example 272:**

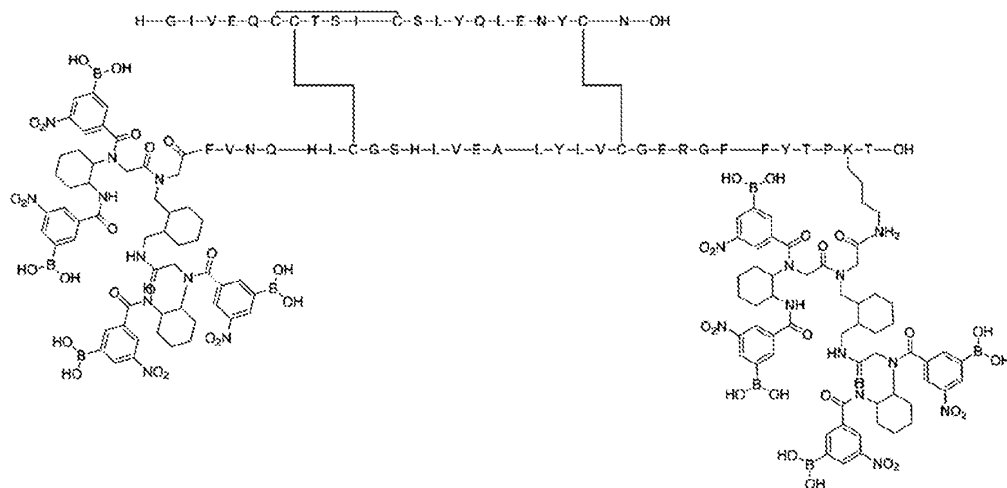
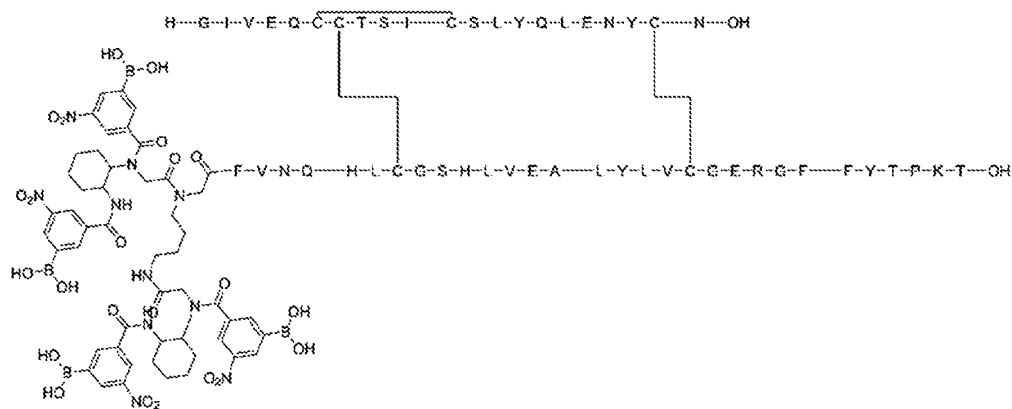
Example 273:**Example 274:****5 Example 275:**

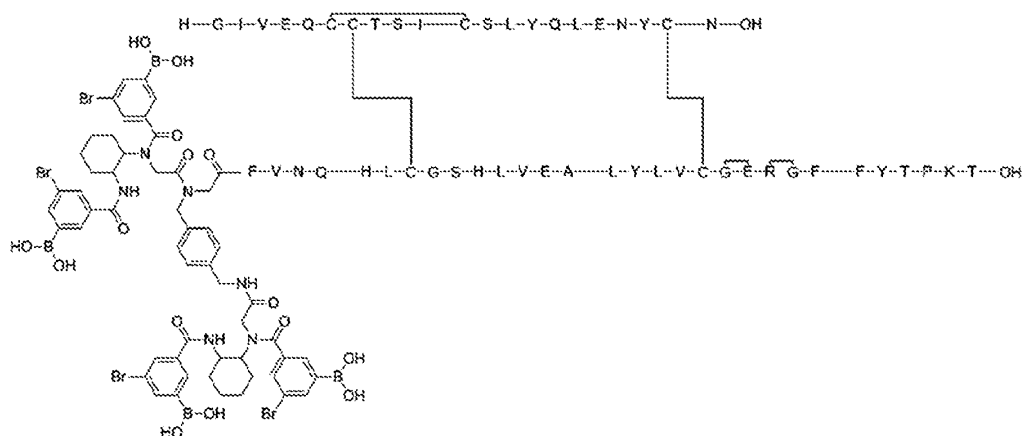
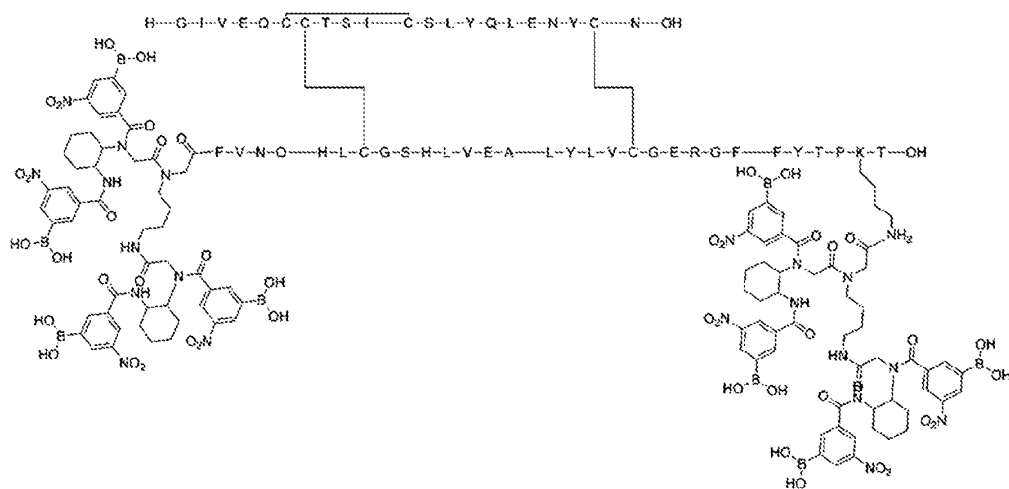
Example 276:**Example 277:**

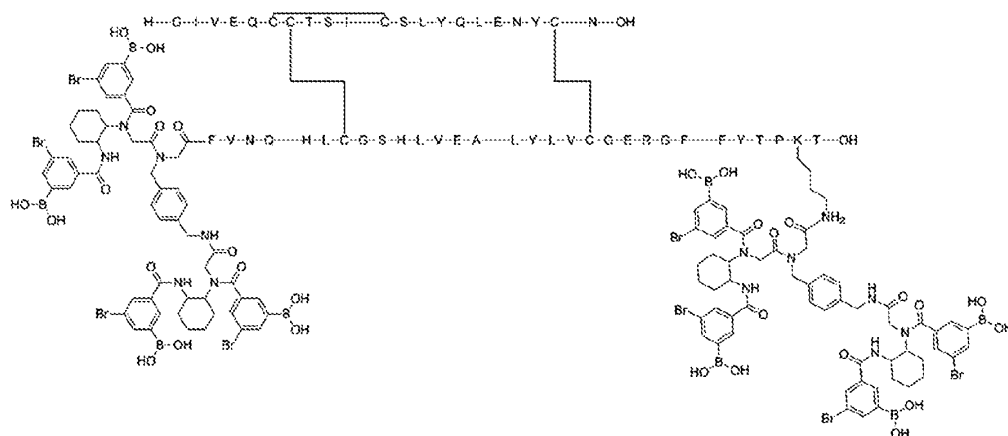
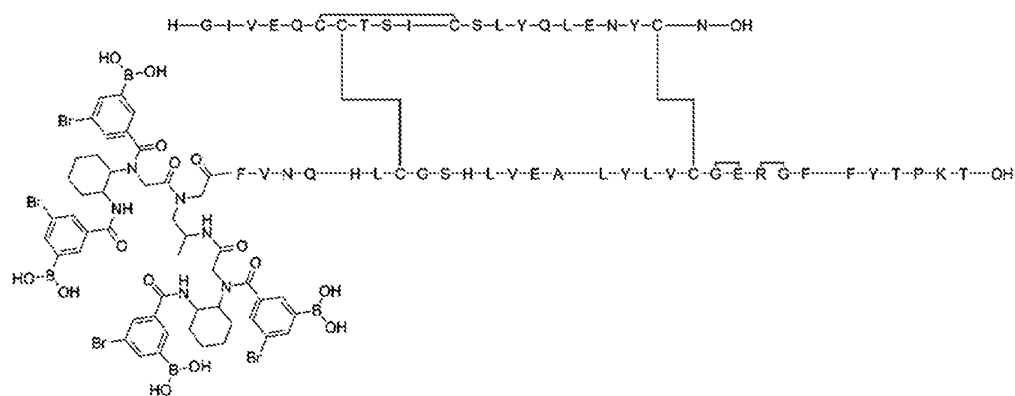
Example 278:**Example 279:**

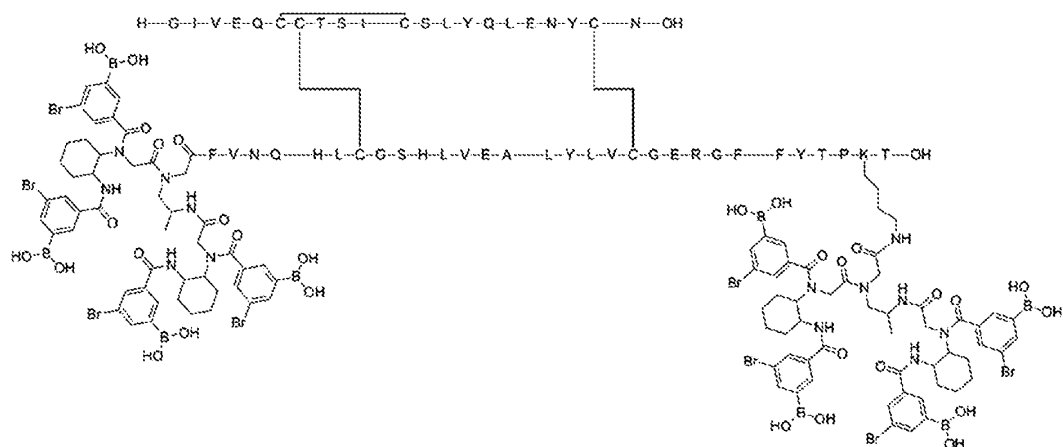
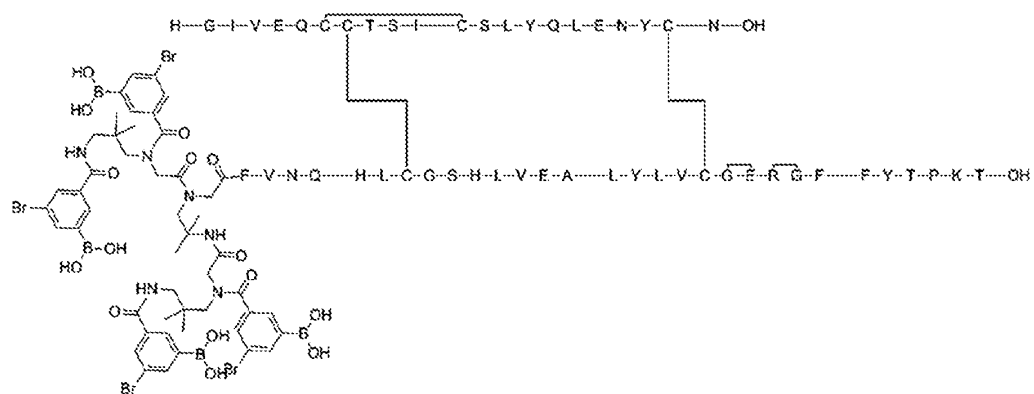
Example 280:**Example 281:**

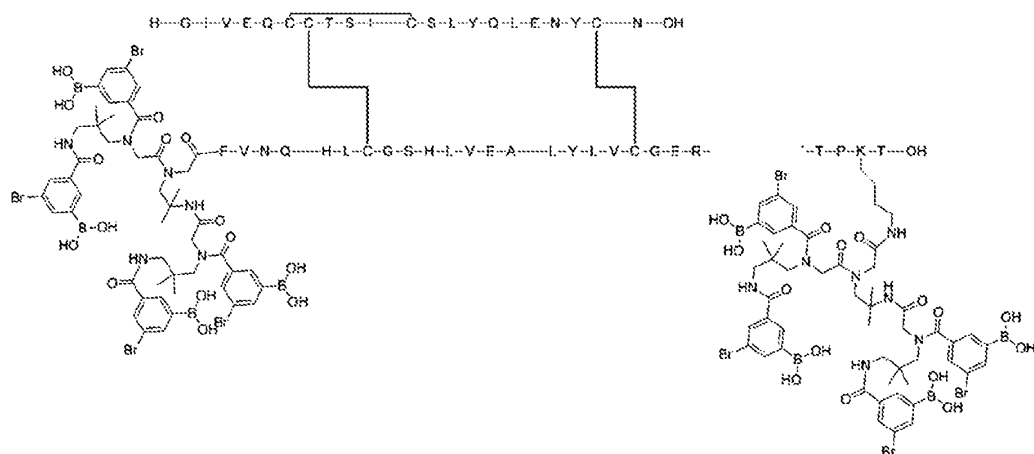
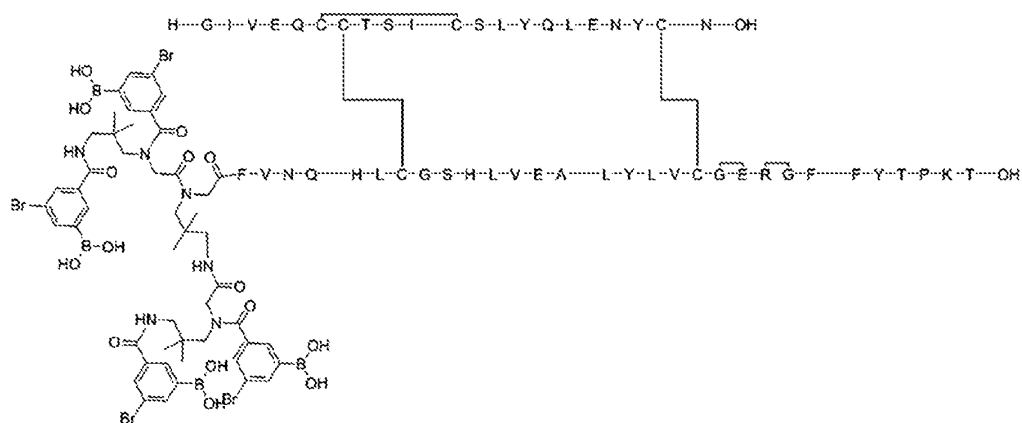
Example 282:**Example 283:**

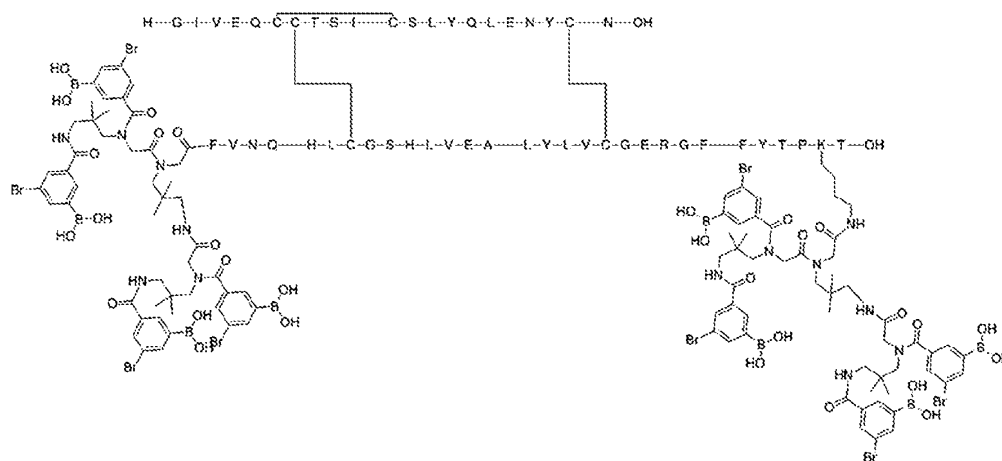
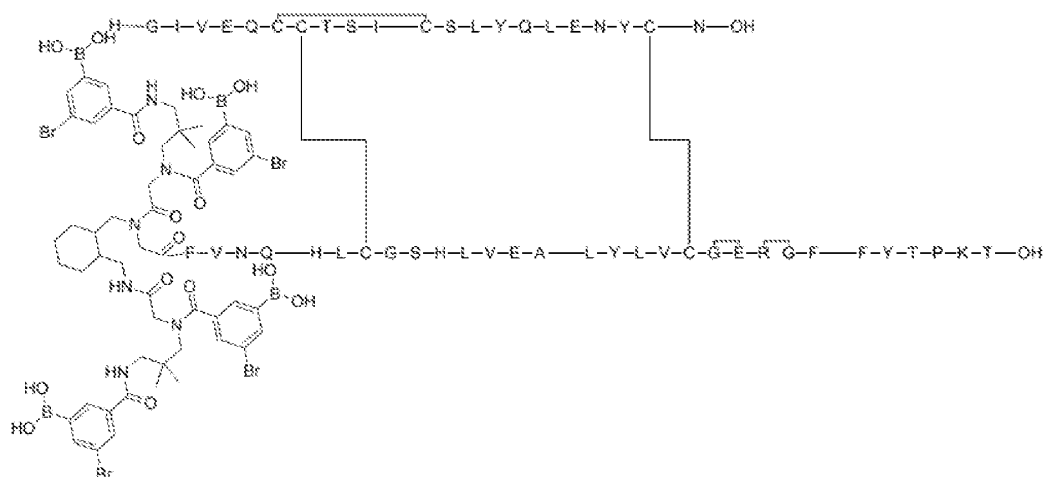
Example 284:**Example 285:**

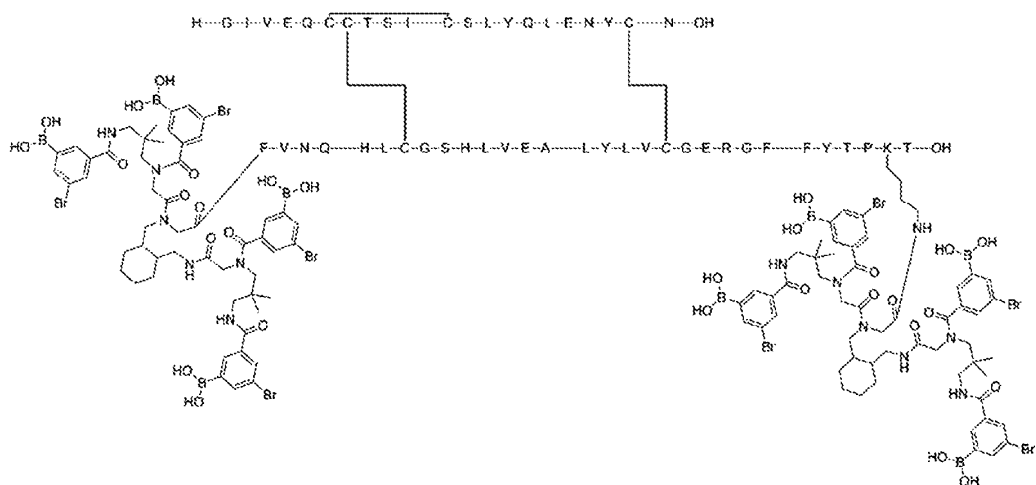
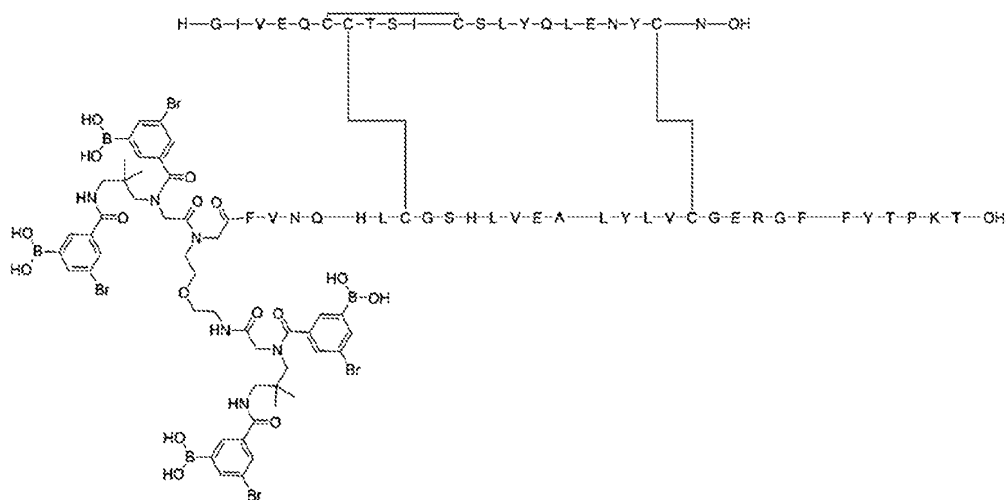
Example 286:**Example 287:**

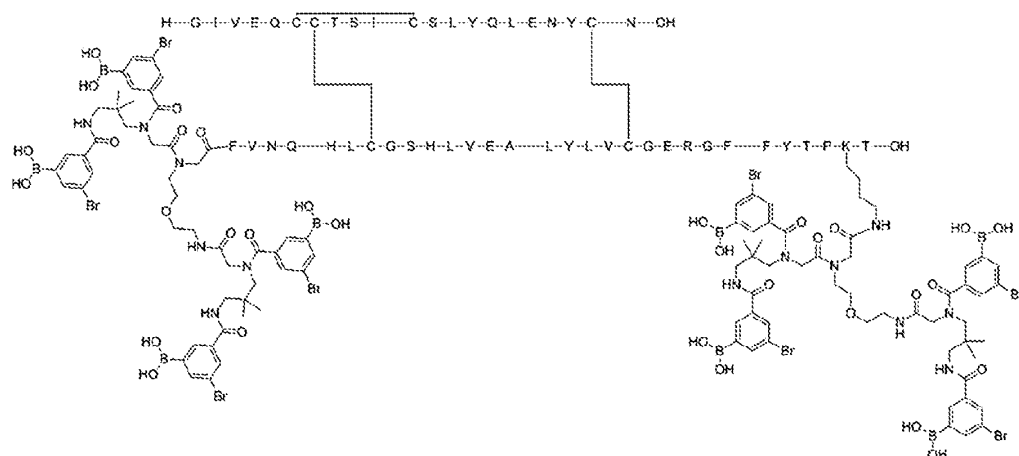
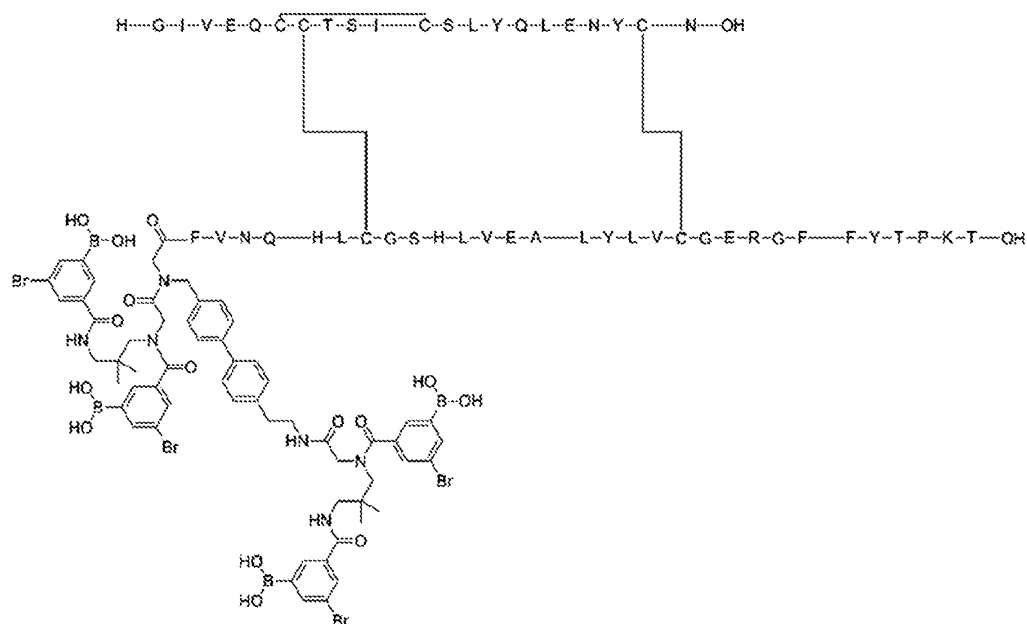
Example 288:**Example 289:**

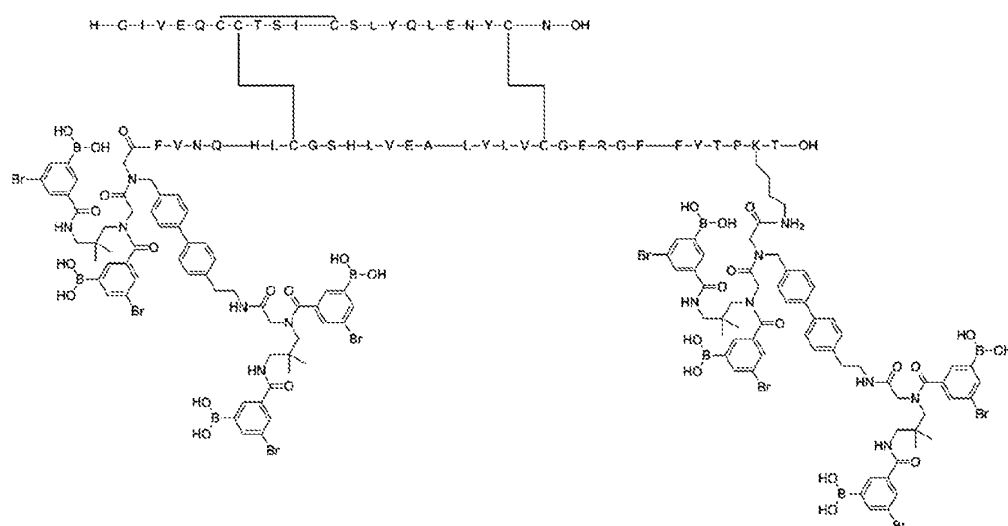
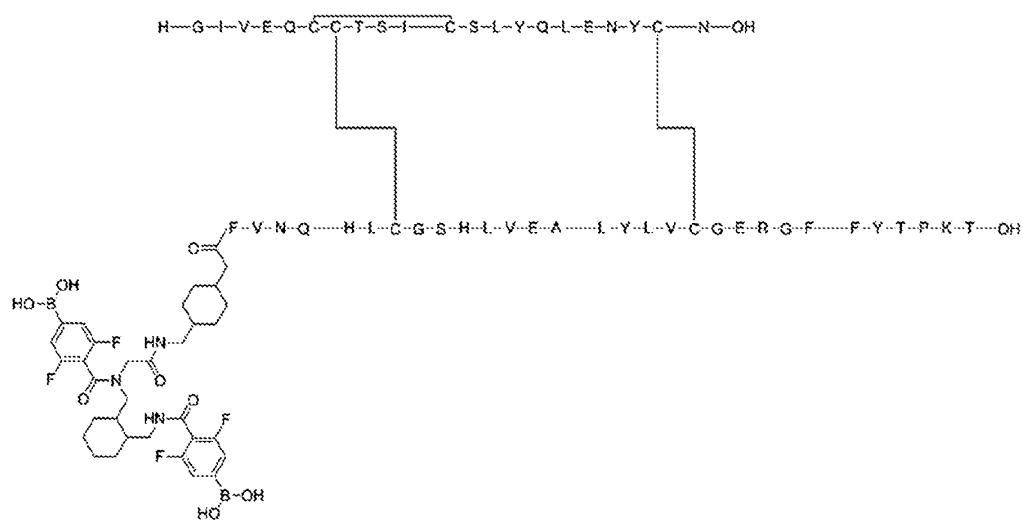
Example 290:**Example 291:**

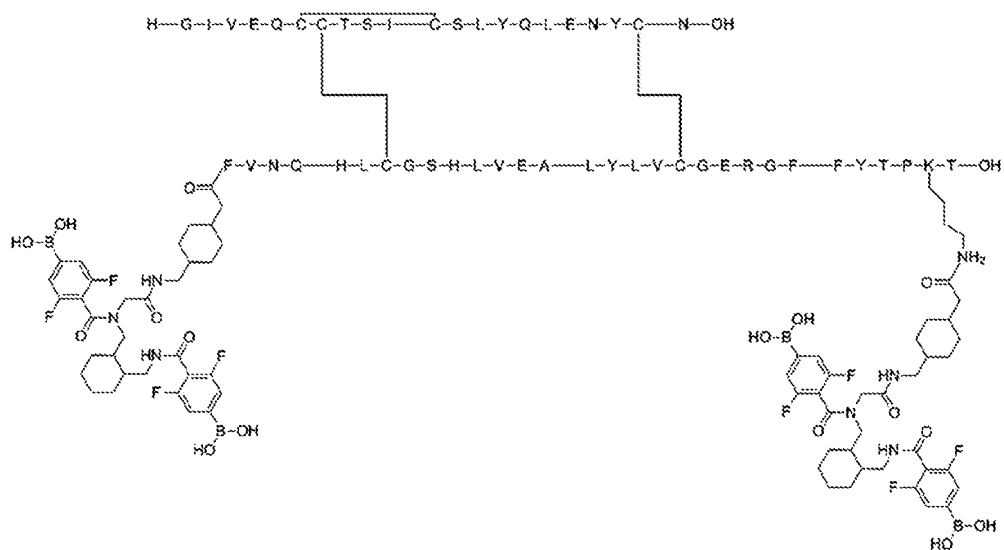
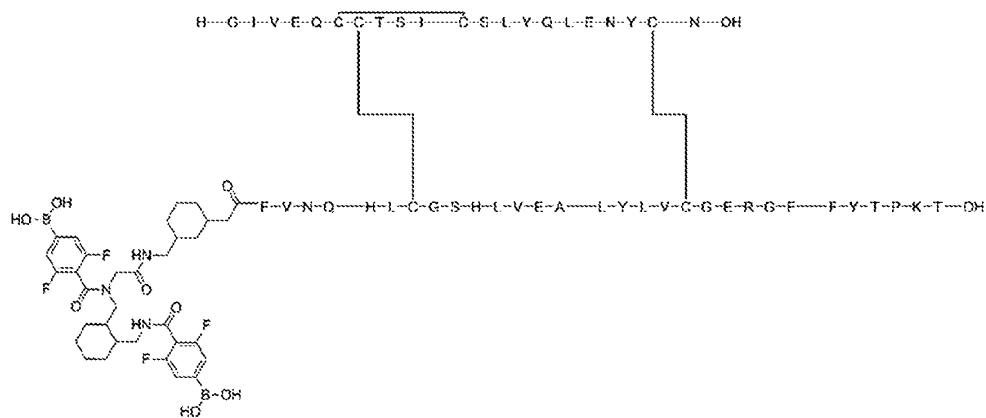
Example 292:**Example 293:**

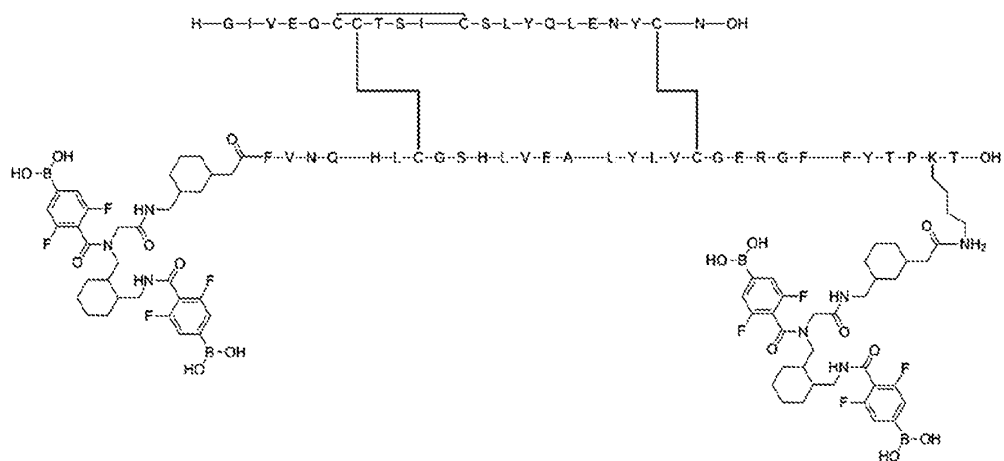
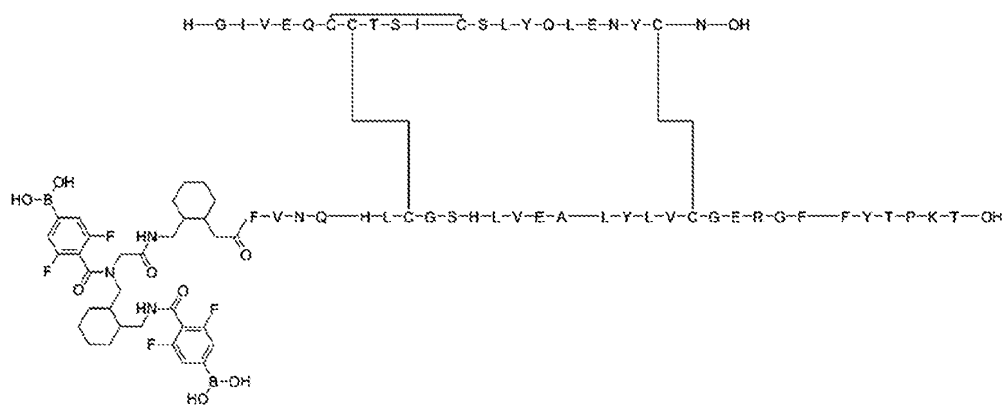
Example 294:**Example 295:**

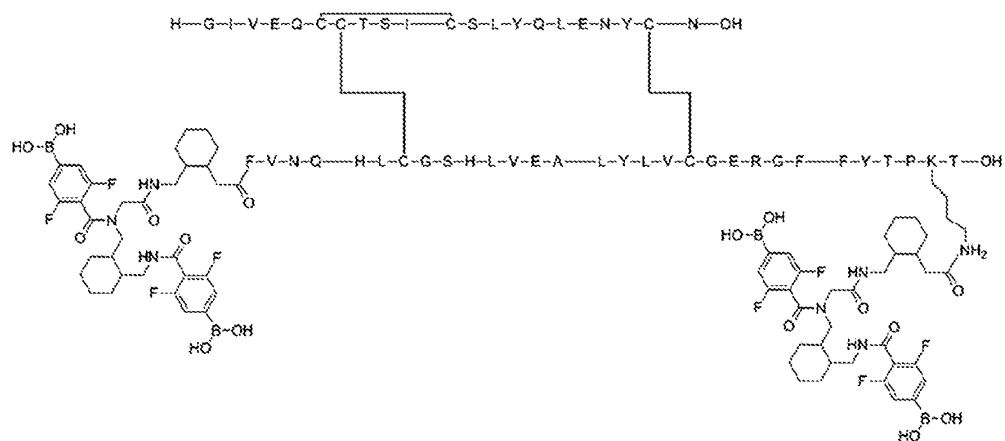
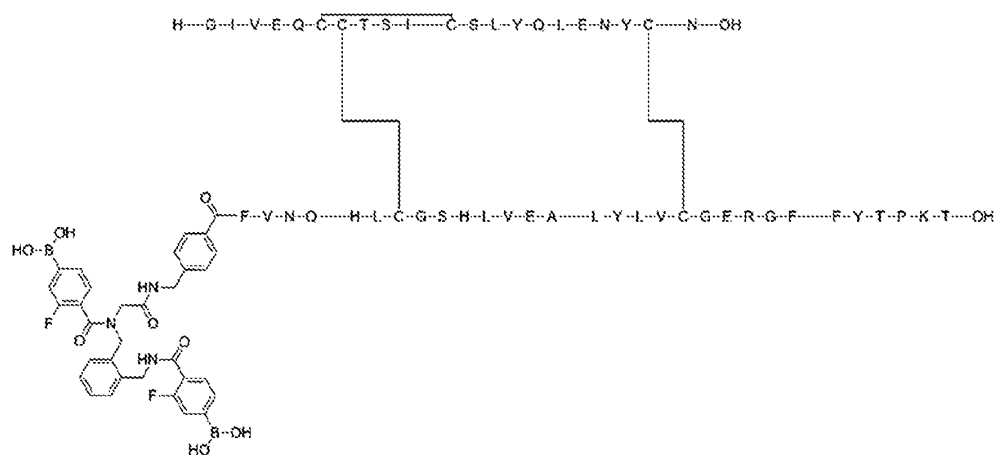
Example 296:**Example 297:**

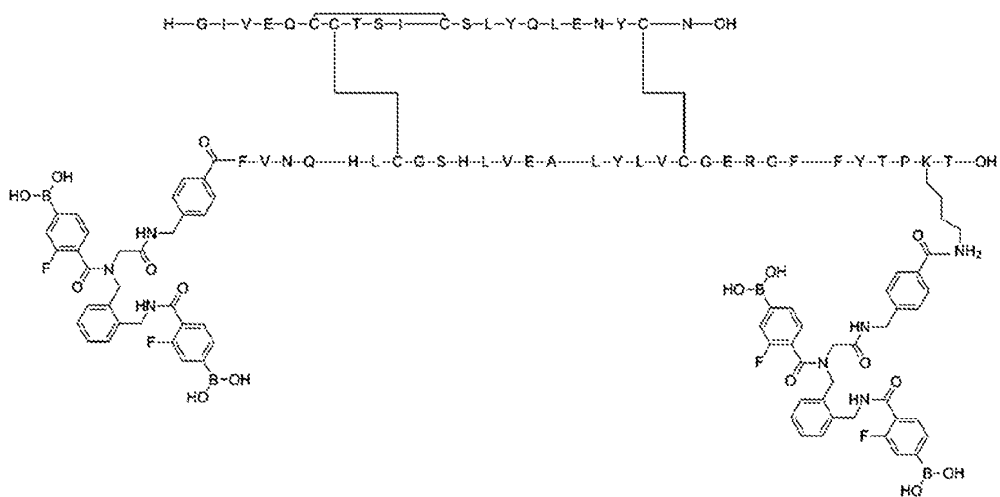
Example 298:**Example 299:**

Example 300:**Example 301:**

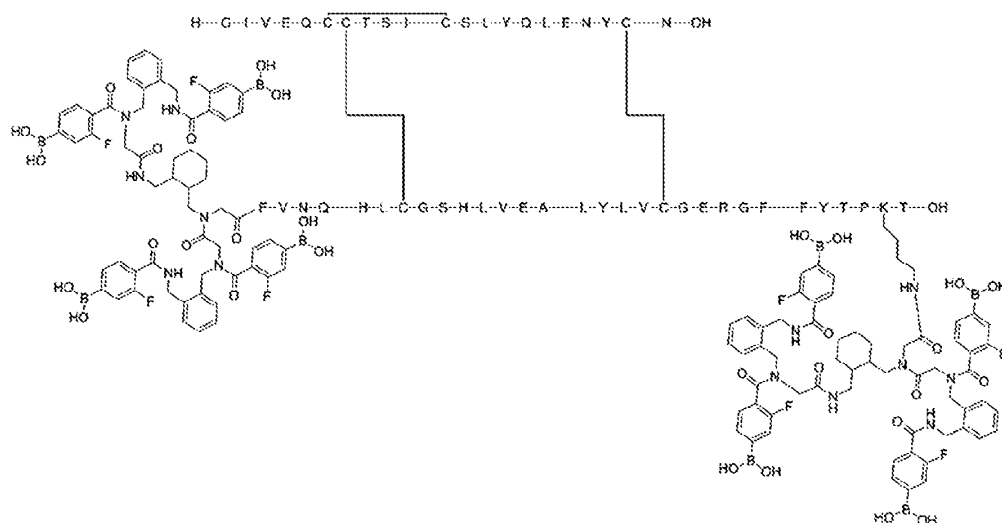
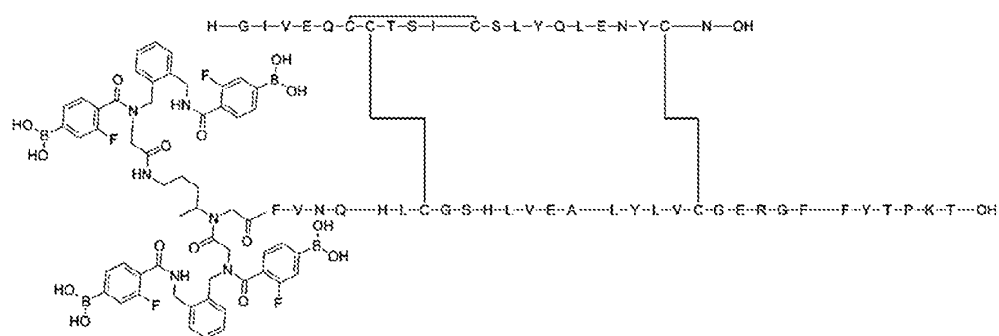
Example 302:**Example 303:**

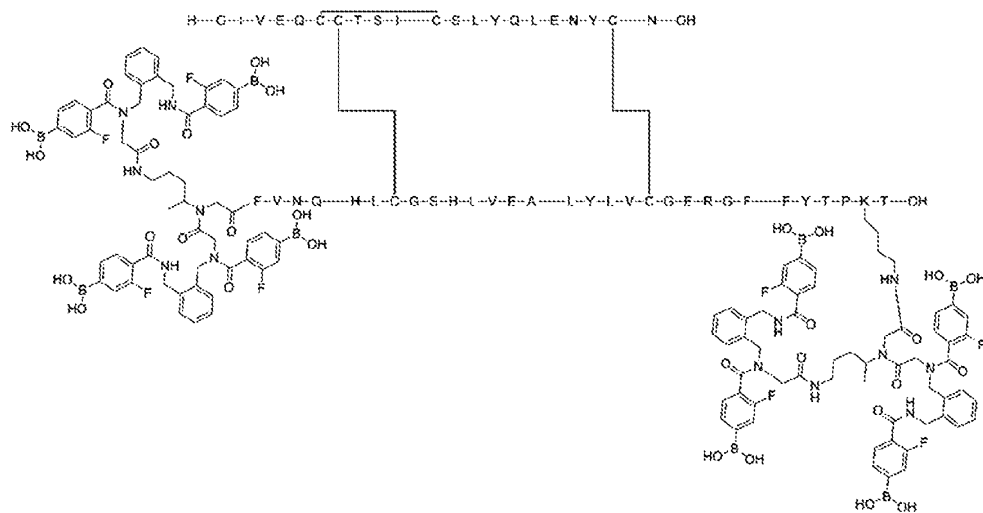
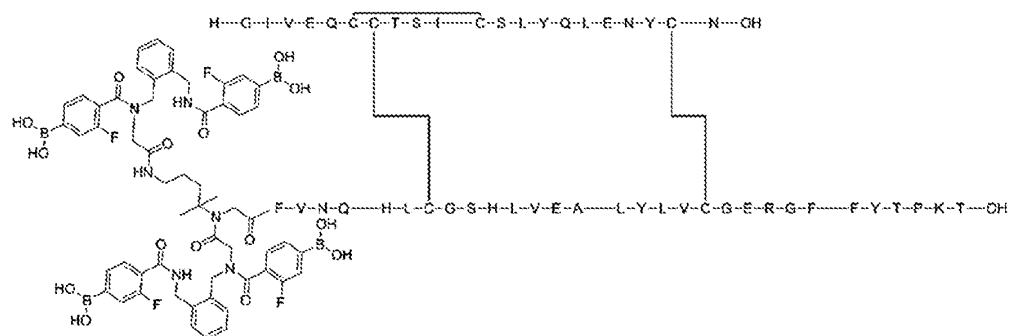
Example 304:**Example 305:**

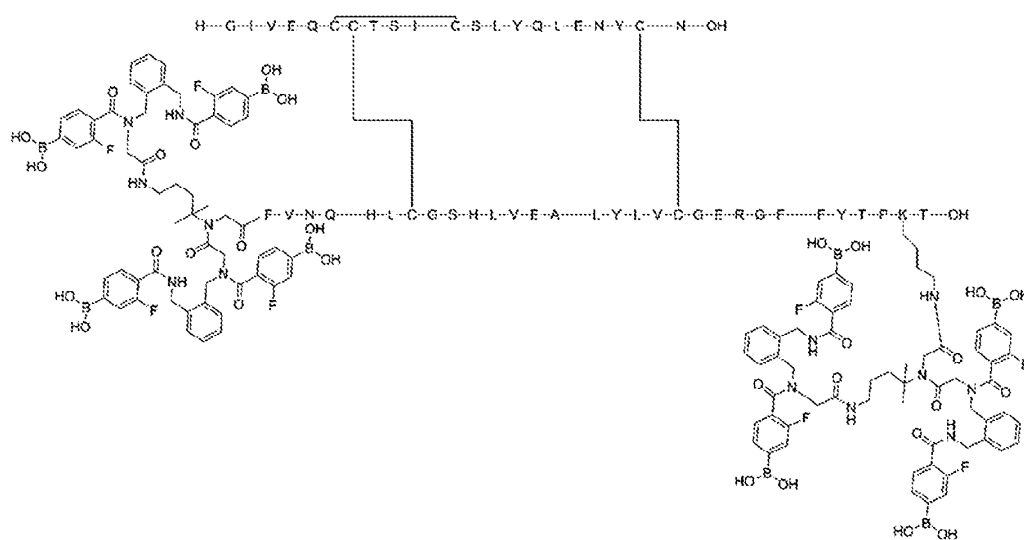
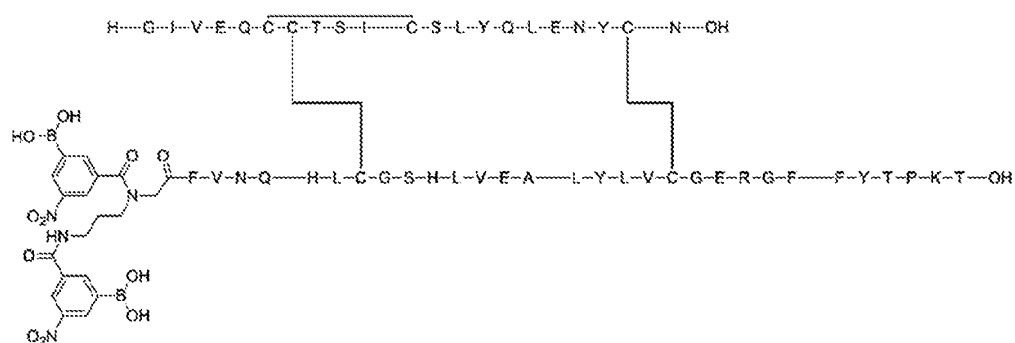
Example 306:**Example 307:**

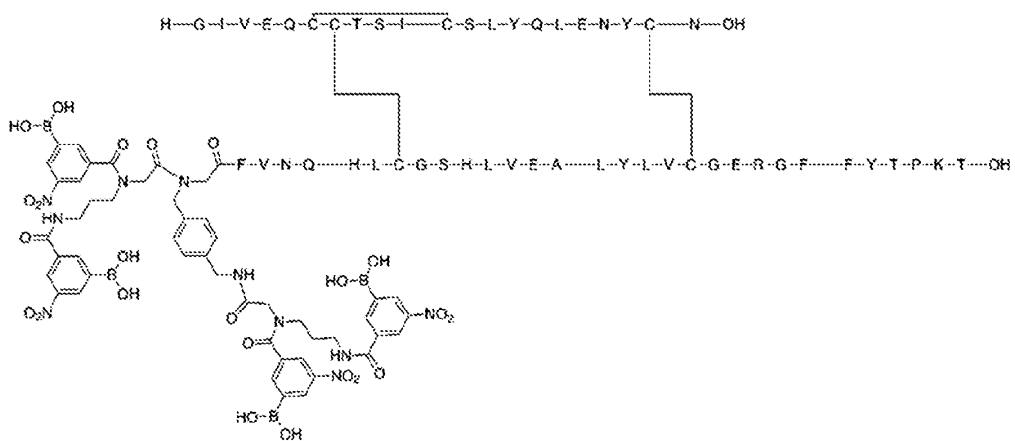
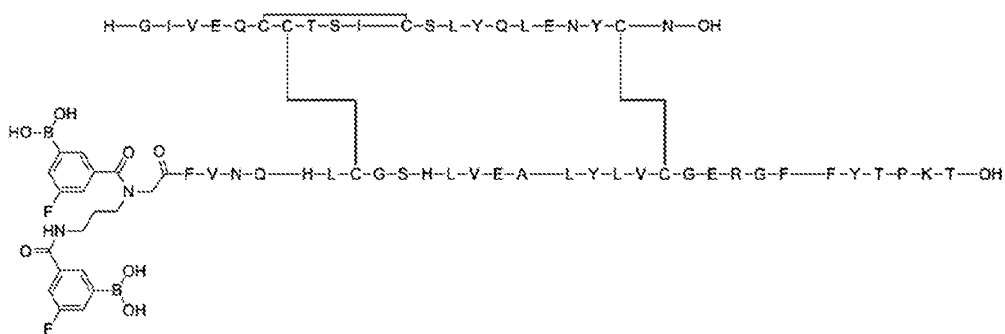


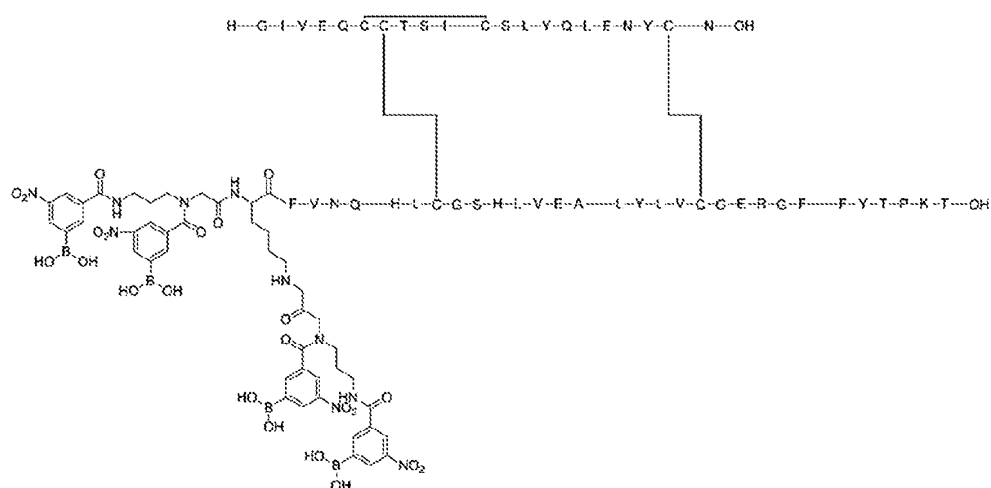
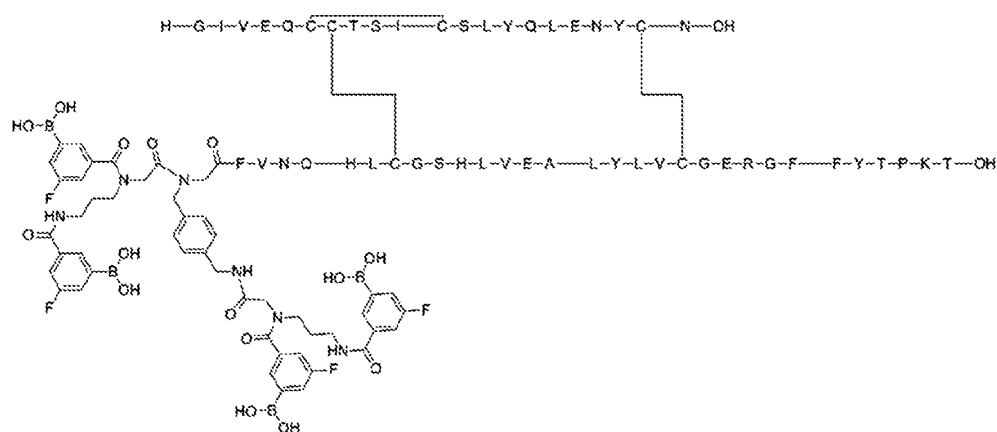
Chemical structure of a protein-ligand complex. The protein backbone is shown as a line drawing with side chains. The ligand is a complex molecule with multiple aromatic rings, a boronic acid group, and a hydroxyl group. The ligand is bound to the protein, with its boronic acid group interacting with a specific site on the protein.

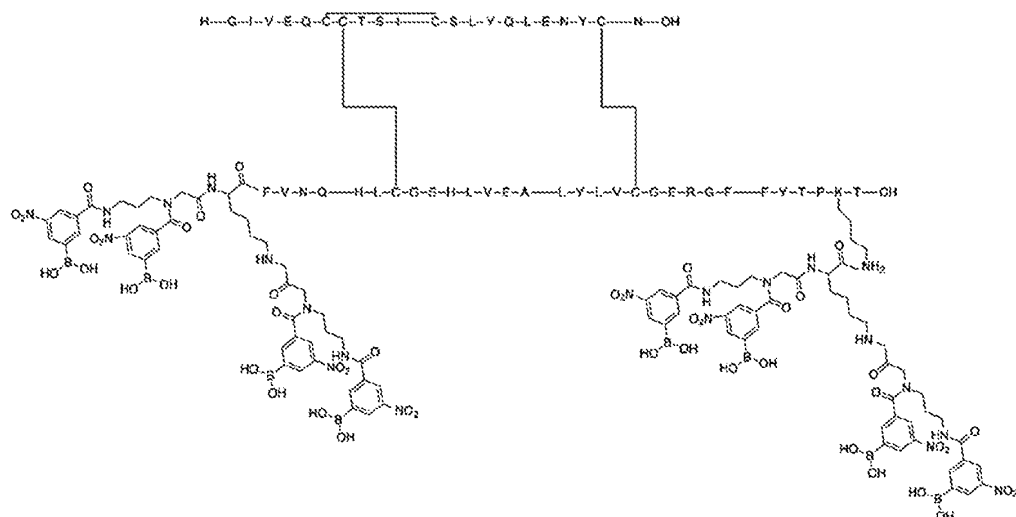
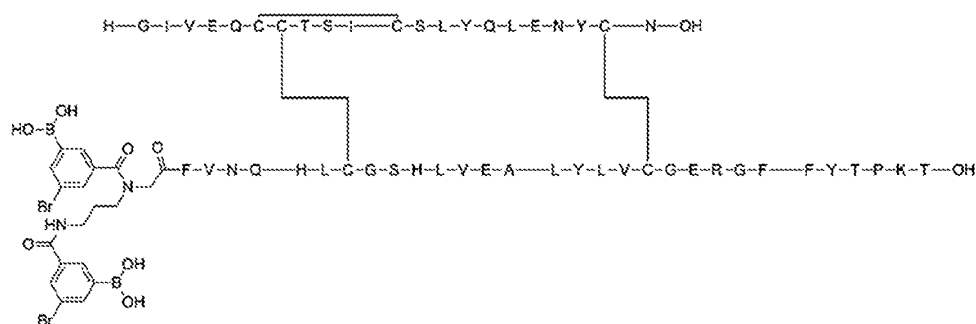
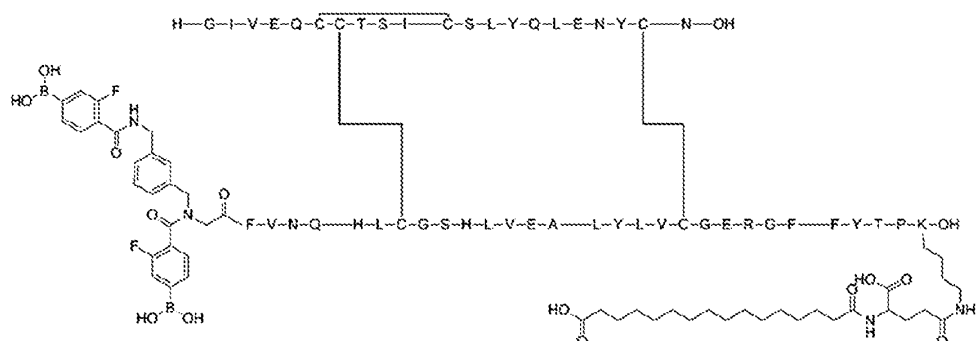
Example 310:**Example 311:**

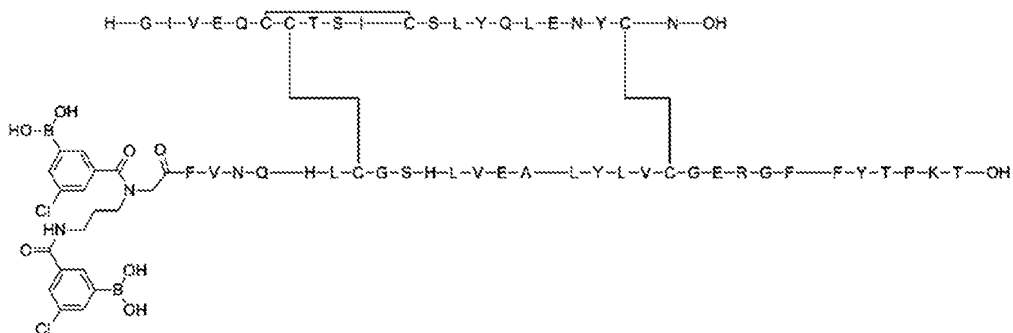
Example 312:**Example 313:**

Example 314:**Example 315:**

Example 316:**Example 317:**

Example 318:**Example 319:**

Example 320:**Example 321:****5 Example 322:**

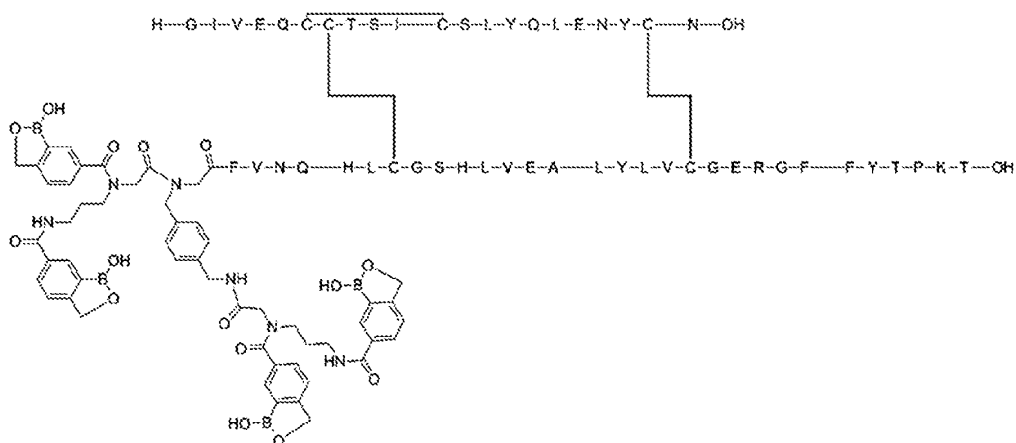
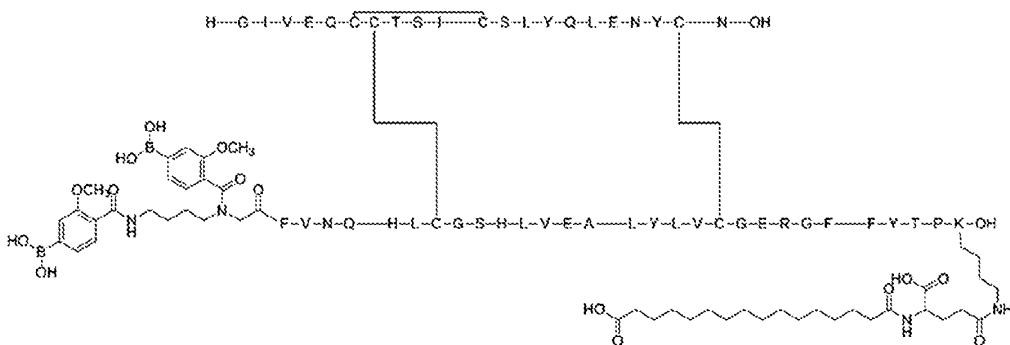
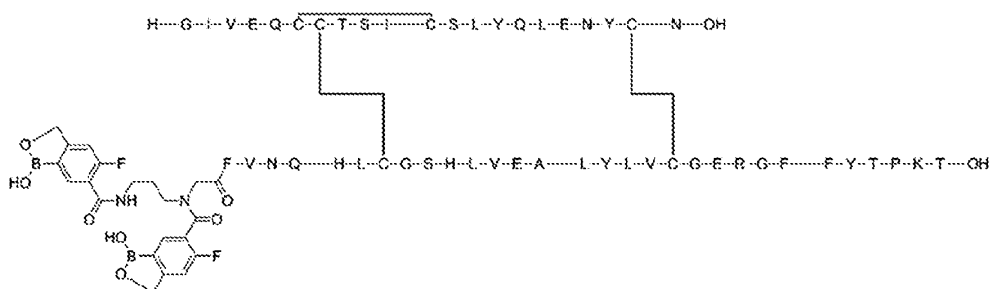


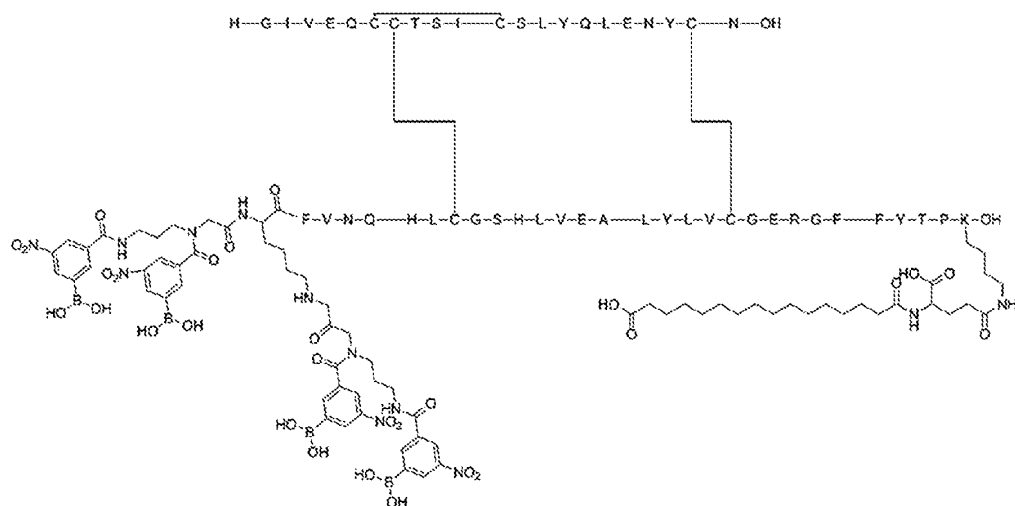
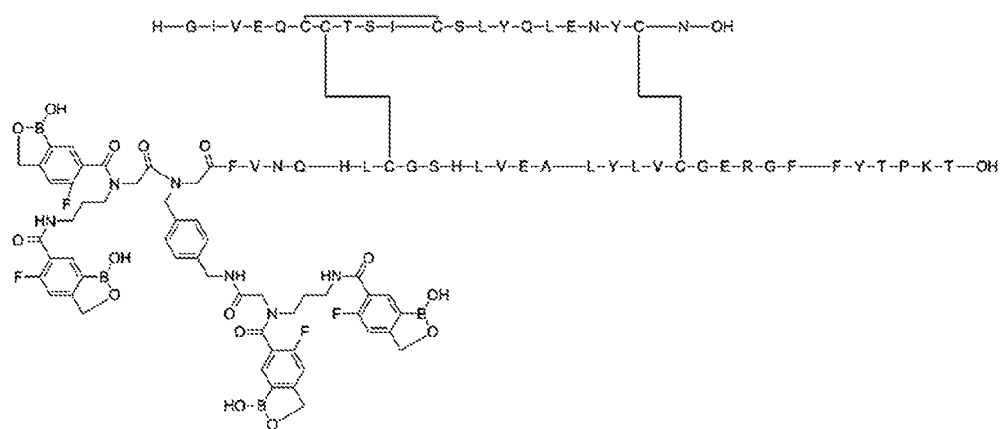
The chemical structure of the conjugate of the 12th antibody is shown. It features a protein backbone with various side chains. Key components include aromatic rings, hydroxyl groups, and a long fatty acid chain. The structure is complex, with multiple functional groups and a long hydrophobic tail.

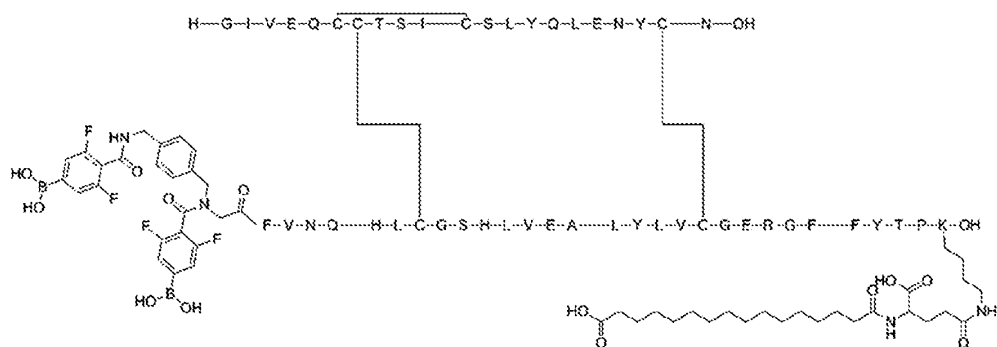
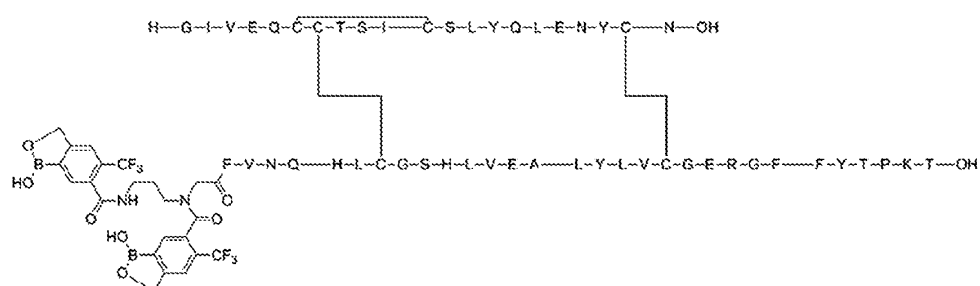
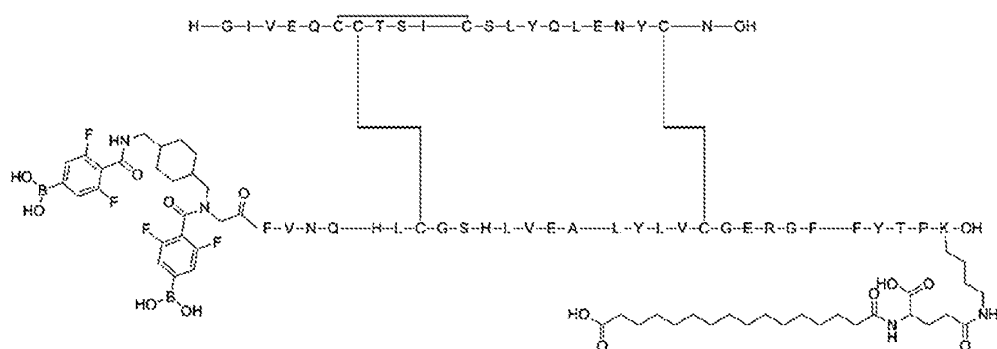
The diagram shows a protein backbone with the sequence: H-G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH. Two fluorinated boronic acid groups are attached to the protein. The first group is attached to the side chain of the first Cysteine (C1) via a disulfide bond. The second group is attached to the side chain of the second Cysteine (C2) via a disulfide bond. Both groups consist of a benzene ring with a fluorine atom and a boronic acid group (-B(OH)2). A long fatty acid chain is attached to the side chain of the last Lysine (K) via an amide bond. The fatty acid chain is a long, straight hydrocarbon chain with a terminal carboxylic acid group (-COOH).

O=C(NCCN(C(=O)c1ccc(O)cc1)C(=O)NCCN(C(=O)c2ccc(O)cc2)C(=O)NCCN(C(=O)c3ccc(O)cc3)C(=O)NCCN(C(=O)c4ccc(O)cc4)C(=O)NCCN(C(=O)c5ccc(O)cc5)C(=O)NCCN(C(=O)c6ccc(O)cc6)C(=O)NCCN(C(=O)c7ccc(O)cc7)C(=O)NCCN(C(=O)c8ccc(O)cc8)C(=O)NCCN(C(=O)c9ccc(O)cc9)C(=O)NCCN(C(=O)c10ccc(O)cc10)C(=O)NCCN(C(=O)c11ccc(O)cc11)C(=O)NCCN(C(=O)c12ccc(O)cc12)C(=O)NCCN(C(=O)c13ccc(O)cc13)C(=O)NCCN(C(=O)c14ccc(O)cc14)C(=O)NCCN(C(=O)c15ccc(O)cc15)C(=O)NCCN(C(=O)c16ccc(O)cc16)C(=O)NCCN(C(=O)c17ccc(O)cc17)C(=O)NCCN(C(=O)c18ccc(O)cc18)C(=O)NCCN(C(=O)c19ccc(O)cc19)C(=O)NCCN(C(=O)c20ccc(O)cc20)C(=O)NCCN(C(=O)c21ccc(O)cc21)C(=O)NCCN(C(=O)c22ccc(O)cc22)C(=O)NCCN(C(=O)c23ccc(O)cc23)C(=O)NCCN(C(=O)c24ccc(O)cc24)C(=O)NCCN(C(=O)c25ccc(O)cc25)C(=O)NCCN(C(=O)c26ccc(O)cc26)C(=O)NCCN(C(=O)c27ccc(O)cc27)C(=O)NCCN(C(=O)c28ccc(O)cc28)C(=O)NCCN(C(=O)c29ccc(O)cc29)C(=O)NCCN(C(=O)c30ccc(O)cc30)C(=O)NCCN(C(=O)c31ccc(O)cc31)C(=O)NCCN(C(=O)c32ccc(O)cc32)C(=O)NCCN(C(=O)c33ccc(O)cc33)C(=O)NCCN(C(=O)c34ccc(O)cc34)C(=O)NCCN(C(=O)c35ccc(O)cc35)C(=O)NCCN(C(=O)c36ccc(O)cc36)C(=O)NCCN(C(=O)c37ccc(O)cc37)C(=O)NCCN(C(=O)c38ccc(O)cc38)C(=O)NCCN(C(=O)c39ccc(O)cc39)C(=O)NCCN(C(=O)c40ccc(O)cc40)C(=O)NCCN(C(=O)c41ccc(O)cc41)C(=O)NCCN(C(=O)c42ccc(O)cc42)C(=O)NCCN(C(=O)c43ccc(O)cc43)C(=O)NCCN(C(=O)c44ccc(O)cc44)C(=O)NCCN(C(=O)c45ccc(O)cc45)C(=O)NCCN(C(=O)c46ccc(O)cc46)C(=O)NCCN(C(=O)c47ccc(O)cc47)C(=O)NCCN(C(=O)c48ccc(O)cc48)C(=O)NCCN(C(=O)c49ccc(O)cc49)C(=O)NCCN(C(=O)c50ccc(O)cc50)C(=O)NCCN(C(=O)c51ccc(O)cc51)C(=O)NCCN(C(=O)c52ccc(O)cc52)C(=O)NCCN(C(=O)c53ccc(O)cc53)C(=O)NCCN(C(=O)c54ccc(O)cc54)C(=O)NCCN(C(=O)c55ccc(O)cc55)C(=O)NCCN(C(=O)c56ccc(O)cc56)C(=O)NCCN(C(=O)c57ccc(O)cc57)C(=O)NCCN(C(=O)c58ccc(O)cc58)C(=O)NCCN(C(=O)c59ccc(O)cc59)C(=O)NCCN(C(=O)c60ccc(O)cc60)C(=O)NCCN(C(=O)c61ccc(O)cc61)C(=O)NCCN(C(=O)c62ccc(O)cc62)C(=O)NCCN(C(=O)c63ccc(O)cc63)C(=O)NCCN(C(=O)c64ccc(O)cc64)C(=O)NCCN(C(=O)c65ccc(O)cc65)C(=O)NCCN(C(=O)c66ccc(O)cc66)C(=O)NCCN(C(=O)c67ccc(O)cc67)C(=O)NCCN(C(=O)c68ccc(O)cc68)C(=O)NCCN(C(=O)c69ccc(O)cc69)C(=O)NCCN(C(=O)c70ccc(O)cc70)C(=O)NCCN(C(=O)c71ccc(O)cc71)C(=O)NCCN(C(=O)c72ccc(O)cc72)C(=O)NCCN(C(=O)c73ccc(O)cc73)C(=O)NCCN(C(=O)c74ccc(O)cc74)C(=O)NCCN(C(=O)c75ccc(O)cc75)C(=O)NCCN(C(=O)c76ccc(O)cc76)C(=O)NCCN(C(=O)c77ccc(O)cc77)C(=O)NCCN(C(=O)c78ccc(O)cc78)C(=O)NCCN(C(=O)c79ccc(O)cc79)C(=O)NCCN(C(=O)c80ccc(O)cc80)C(=O)NCCN(C(=O)c81ccc(O)cc81)C(=O)NCCN(C(=O)c82ccc(O)cc82)C(=O)NCCN(C(=O)c83ccc(O)cc83)C(=O)NCCN(C(=O)c84ccc(O)cc84)C(=O)NCCN(C(=O)c85ccc(O)cc85)C(=O)NCCN(C(=O)c86ccc(O)cc86)C(=O)NCCN(C(=O)c87ccc(O)cc87)C(=O)NCCN(C(=O)c88ccc(O)cc88)C(=O)NCCN(C(=O)c89ccc(O)cc89)C(=O)NCCN(C(=O)c90ccc(O)cc90)C(=O)NCCN(C(=O)c91ccc(O)cc91)C(=O)NCCN(C(=O)c92ccc(O)cc92)C(=O)NCCN(C(=O)c93ccc(O)cc93)C(=O)NCCN(C(=O)c94ccc(O)cc94)C(=O)NCCN(C(=O)c95ccc(O)cc95)C(=O)NCCN(C(=O)c96ccc(O)cc96)C(=O)NCCN(C(=O)c97ccc(O)cc97)C(=O)NCCN(C(=O)c98ccc(O)cc98)C(=O)NCCN(C(=O)c99ccc(O)cc99)C(=O)NCCN(C(=O)c100ccc(O)cc100)C(=O)NCCN(C(=O)c101ccc(O)cc101)C(=O)NCCN(C(=O)c102ccc(O)cc102)C(=O)NCCN(C(=O)c103ccc(O)cc103)C(=O)NCCN(C(=O)c104ccc(O)cc104)C(=O)NCCN(C(=O)c105ccc(O)cc105)C(=O)NCCN(C(=O)c106ccc(O)cc106)C(=O)NCCN(C(=O)c107ccc(O)cc107)C(=O)NCCN(C(=O)c108ccc(O)cc108)C(=O)NCCN(C(=O)c109ccc(O)cc109)C(=O)NCCN(C(=O)c110ccc(O)cc110)C(=O)NCCN(C(=O)c111ccc(O)cc111)C(=O)NCCN(C(=O)c112ccc(O)cc112)C(=O)NCCN(C(=O)c113ccc(O)cc113)C(=O)NCCN(C(=O)c114ccc(O)cc114)C(=O)NCCN(C(=O)c115ccc(O)cc115)C(=O)NCCN(C(=O)c116ccc(O)cc116)C(=O)NCCN(C(=O)c117ccc(O)cc117)C(=O)NCCN(C(=O)c118ccc(O)cc118)C(=O)NCCN(C(=O)c119ccc(O)cc119)C(=O)NCCN(C(=O)c120ccc(O)cc120)C(=O)NCCN(C(=O)c121ccc(O)cc121)C(=O)NCCN(C(=O)c122ccc(O)cc122)C(=O)NCCN(C(=O)c123ccc(O)cc123)C(=O)NCCN(C(=O)c124ccc(O)cc124)C(=O)NCCN(C(=O)c125ccc(O)cc125)C(=O)NCCN(C(=O)c126ccc(O)cc126)C(=O)NCCN(C(=O)c127ccc(O)cc127)C(=O)NCCN(C(=O)c128ccc(O)cc128)C(=O)NCCN(C(=O)c129ccc(O)cc129)C(=O)NCCN(C(=O)c130ccc(O)cc130)C(=O)NCCN(C(=O)c131ccc(O)cc131)C(=O)NCCN(C(=O)c132ccc(O)cc132)C(=O)NCCN(C(=O)c133ccc(O)cc133)C(=O)NCCN(C(=O)c134ccc(O)cc134)C(=O)NCCN(C(=O)c135ccc(O)cc135)C(=O)NCCN(C(=O)c136ccc(O)cc136)C(=O)NCCN(C(=O)c137ccc(O)cc137)C(=O)NCCN(C(=O)c138ccc(O)cc138)C(=O)NCCN(C(=O)c139ccc(O)cc139)C(=O)NCCN(C(=O)c140ccc(O)cc140)C(=O)NCCN(C(=O)c141ccc(O)cc141)C(=O)NCCN(C(=O)c142ccc(O)cc142)C(=O)NCCN(C(=O)c143ccc(O)cc143)C(=O)NCCN(C(=O)c144ccc(O)cc144)C(=O)NCCN(C(=O)c145ccc(O)cc145)C(=O)NCCN(C(=O)c146ccc(O)cc146)C(=O)NCCN(C(=O)c147ccc(O)cc147)C(=O)NCCN(C(=O)c148ccc(O)cc148)C(=O)NCCN(C(=O)c149ccc(O)cc149)C(=O)NCCN(C(=O)c150ccc(O)cc150)C(=O)NCCN(C(=O)c151ccc(O)cc151)C(=O)NCCN(C(=O)c152ccc(O)cc152)C(=O)NCCN(C(=O)c153ccc(O)cc153)C(=O)NCCN(C(=O)c154ccc(O)cc154)C(=O)NCCN(C(=O)c155ccc(O)cc155)C(=O)NCCN(C(=O)c156ccc(O)cc156)C(=O)NCCN(C(=O)c157ccc(O)cc157)C(=O)NCCN(C(=O)c158ccc(O)cc158)C(=O)NCCN(C(=O)c159ccc(O)cc159)C(=O)NCCN(C(=O)c160ccc(O)cc160)C(=O)NCCN(C(=O)c161ccc(O)cc161)C(=O)NCCN(C(=O)c162ccc(O)cc162)C(=O)NCCN(C(=O)c163ccc(O)cc163)C(=O)NCCN(C(=O)c164ccc(O)cc164)C(=O)NCCN(C(=O)c165ccc(O)cc165)C(=O)NCCN(C(=O)c166ccc(O)cc166)C(=O)NCCN(C(=O)c167ccc(O)cc167)C(=O)NCCN(C(=O)c168ccc(O)cc168)C(=O)NCCN(C(=O)c169ccc(O)cc169)C(=O)NCCN(C(=O)c170ccc(O)cc170)C(=O)NCCN(C(=O)c171ccc(O)cc171)C(=O)NCCN(C(=O)c172ccc(O)cc172)C(=O)NCCN(C(=O)c173ccc(O)cc173)C(=O)NCCN(C(=O)c174ccc(O)cc174)C(=O)NCCN(C(=O)c175ccc(O)cc175)C(=O)NCCN(C(=O)c176ccc(O)cc176)C(=O)NCCN(C(=O)c177ccc(O)cc177)C(=O)NCCN(C(=O)c178ccc(O)cc178)C(=O)NCCN(C(=O)c179ccc(O)cc179)C(=O)NCCN(C(=O)c1

Chemical structure of a protein with two modified residues. The protein sequence is H-G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH. The first Cysteine (C) is modified with a 4-methoxyphenylboronic acid group. The second Cysteine (C) is modified with a 4-methoxyphenylboronic acid group. The structure shows the protein backbone with side chains, and the two modified residues are highlighted with dashed boxes.

Example 329:**Example 330:****5 Example 331:**

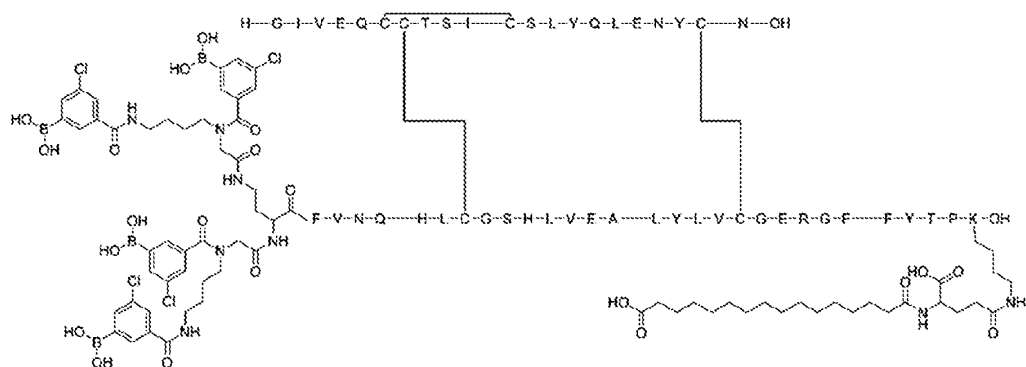
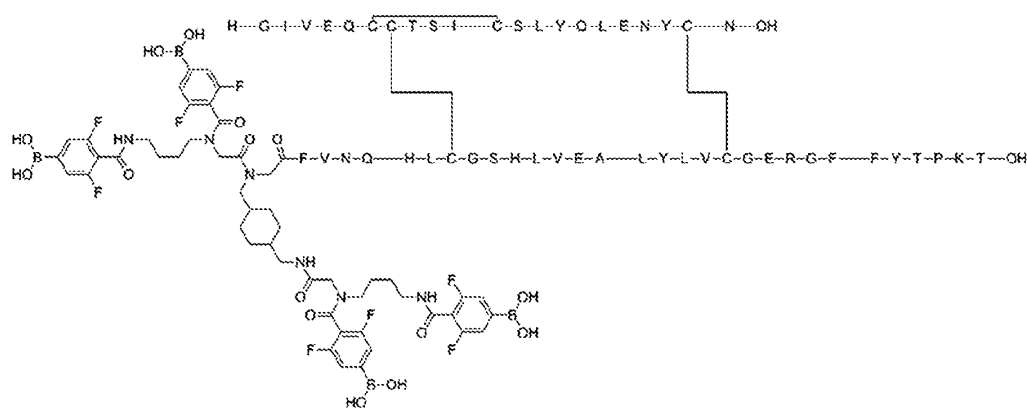
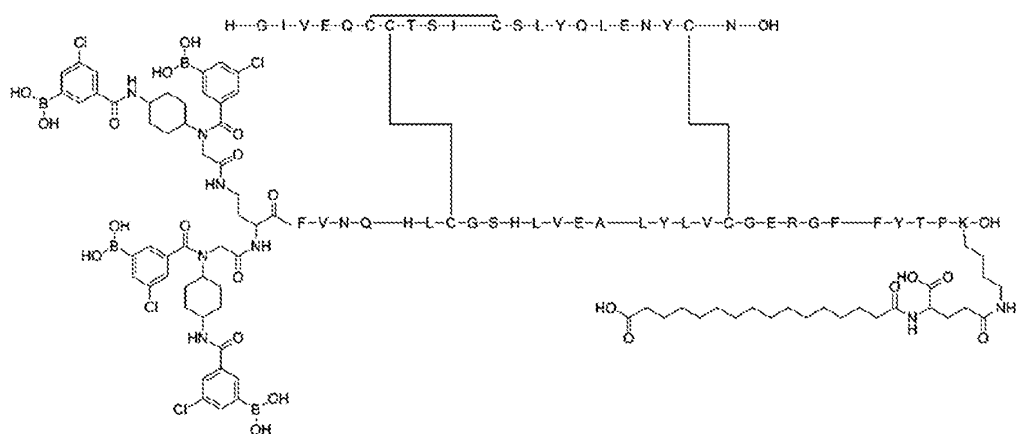
Example 332:**Example 333:**

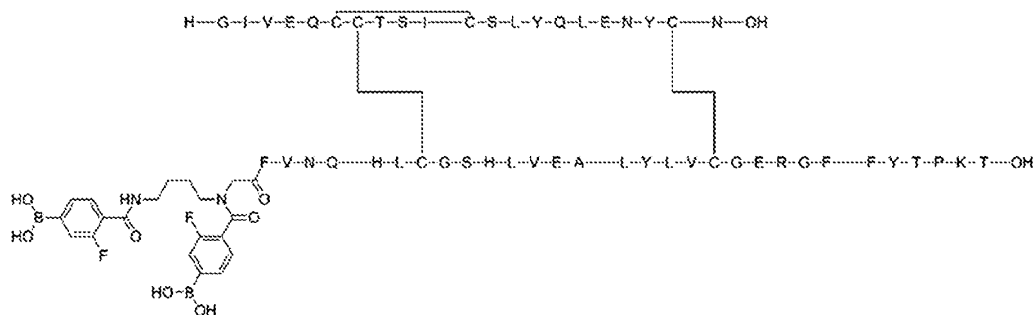
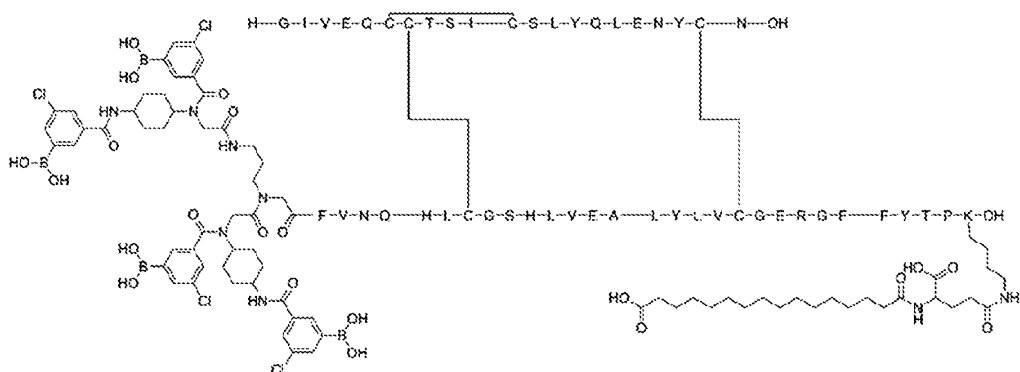
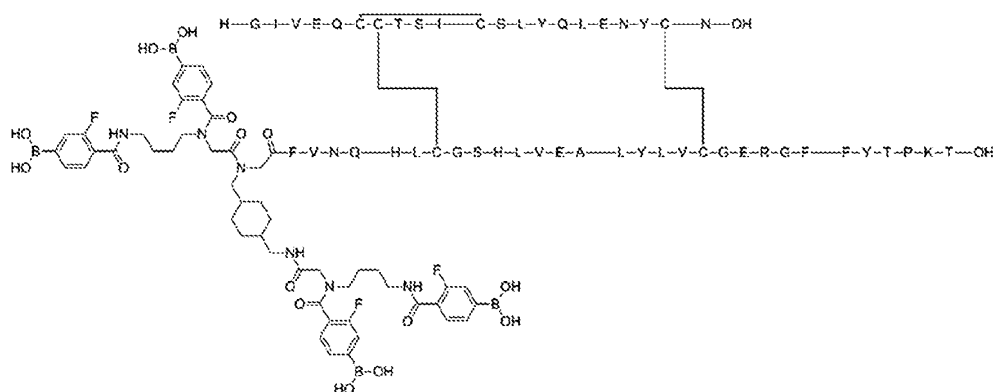
Example 334:**Example 335:****5 Example 336:**

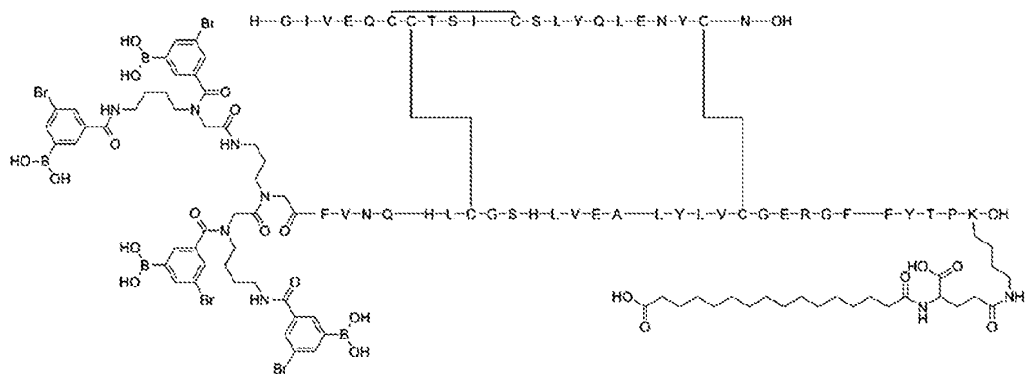
The chemical structure shows a peptide chain: H-G-I-V-E-Q-C-C-T-S-I-C-S-I-Y-Q-L-E-N-Y-C-N-OH. The peptide is conjugated to a fluorinated boronate ester (F₃C-C₆H₃(OH)-B(OH)₂) via a linker: -F-V-N-Q-H-L-C-G-S-H-L-V-E-A-I-Y-I-V-C-G-E-R-G-F-F-Y-T-P-K-T-OH. The linker is further conjugated to a fluorinated boronate ester (F₃C-C₆H₃(OH)-B(OH)₂) via a linker: -F₃C-C₆H₃(OH)-B(OH)₂.

The diagram illustrates the chemical structure of a protein-ligand complex. The protein backbone is represented by a line drawing, showing the sequence of amino acids. The side chains are explicitly drawn, including a long alkyl chain and a carboxylic acid group. The ligand is a complex molecule with a central core and several functional groups, including a long alkyl chain and a carboxylic acid group. The structure is shown in a perspective view, with the protein backbone and side chains in black and the ligand in red.

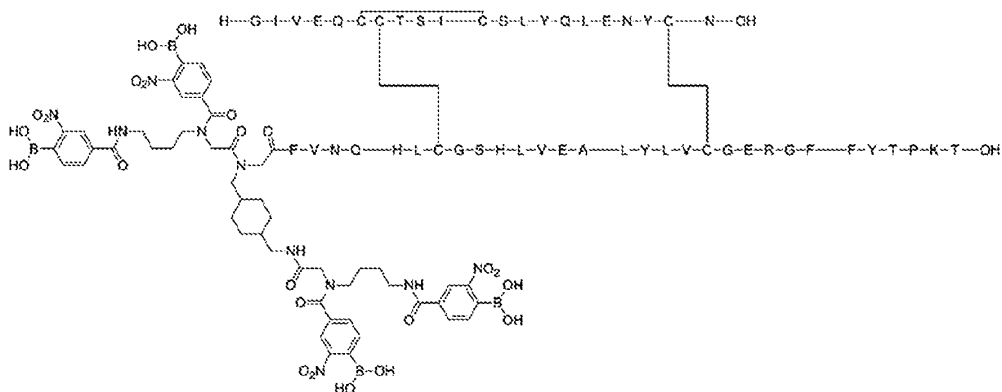
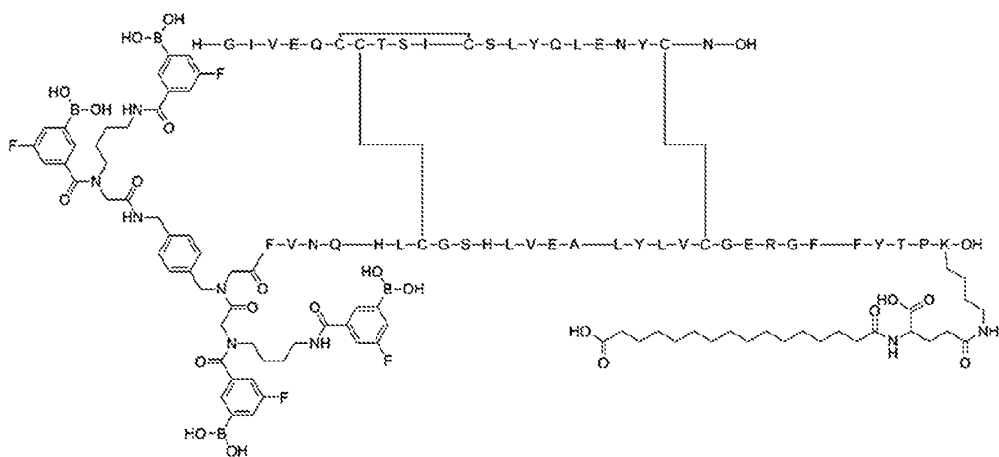
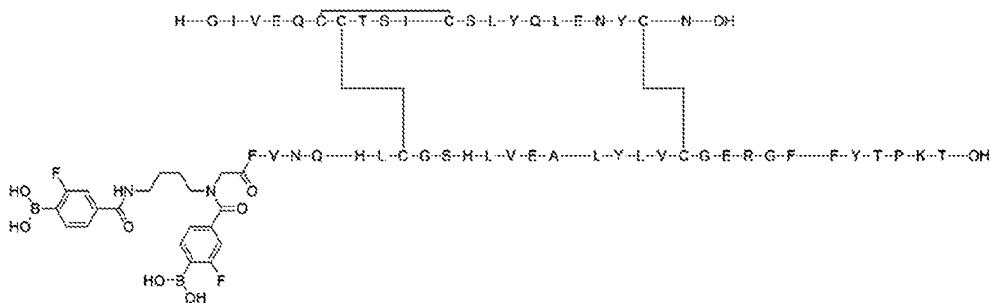
Oc1cc(F)c(NC(=O)NCC2C(=O)N(C3C(F)=CC(=C(C3)F)B(O)(O)O)C2=O)c(F)c1

Example 340:**Example 341:****5 Example 342:**

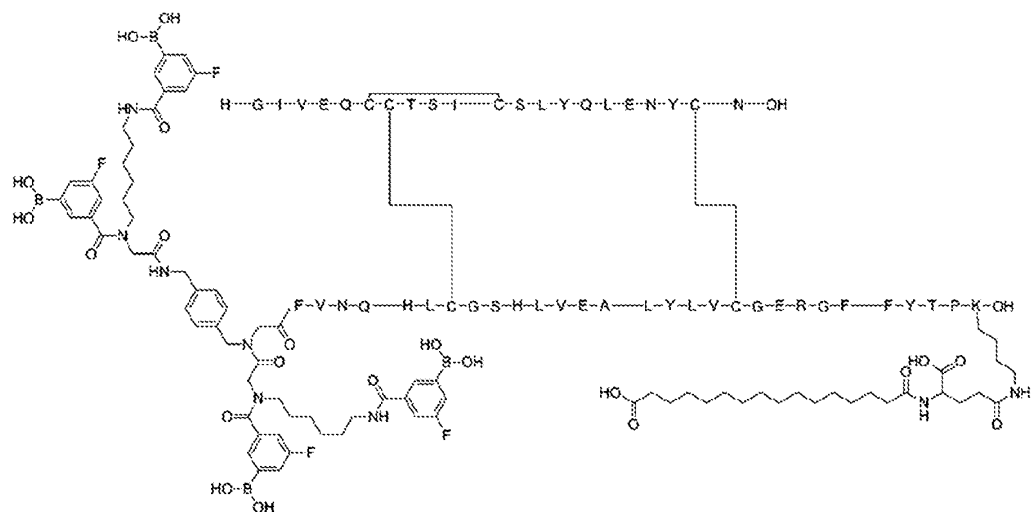
Example 343:**Example 344:****5 Example 345:**



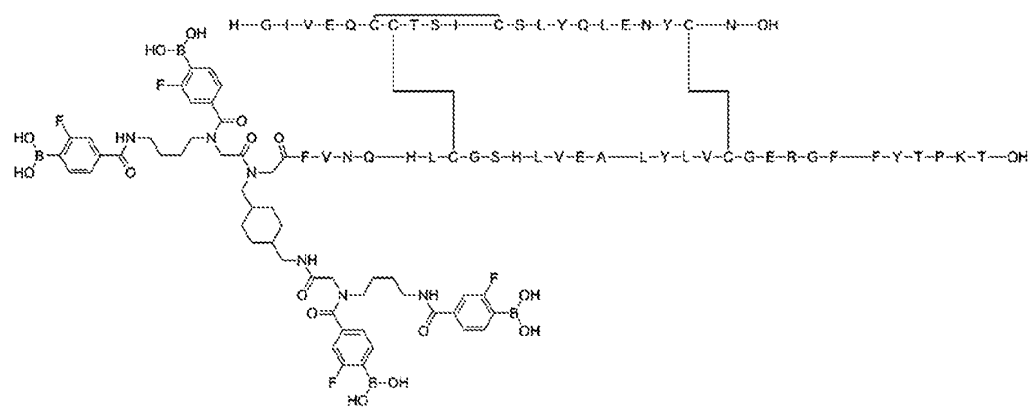
NC(CCC(CCN(C(C(=O)O)C1=CC=C(C=C1)C(=O)N2CCN(C2)C(=O)O)C3=CC=C(C=C3)[N+](=O)[O-])C(=O)O)C(=O)O)C(=O)O

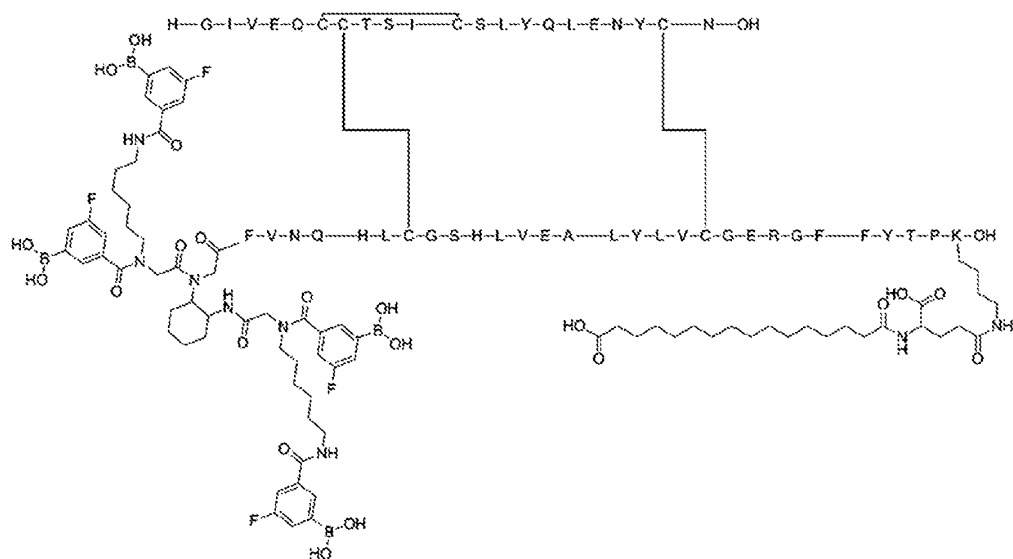
Example 349:**Example 350:****5 Example 351:**

Example 352:

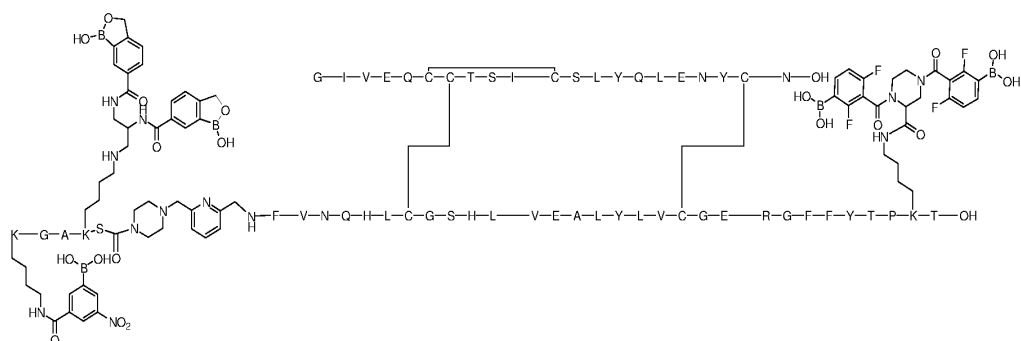


Example 353:

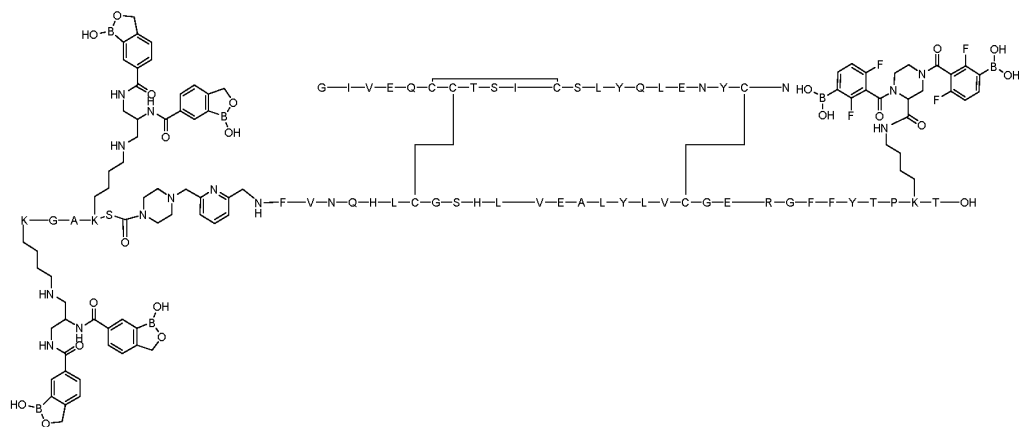




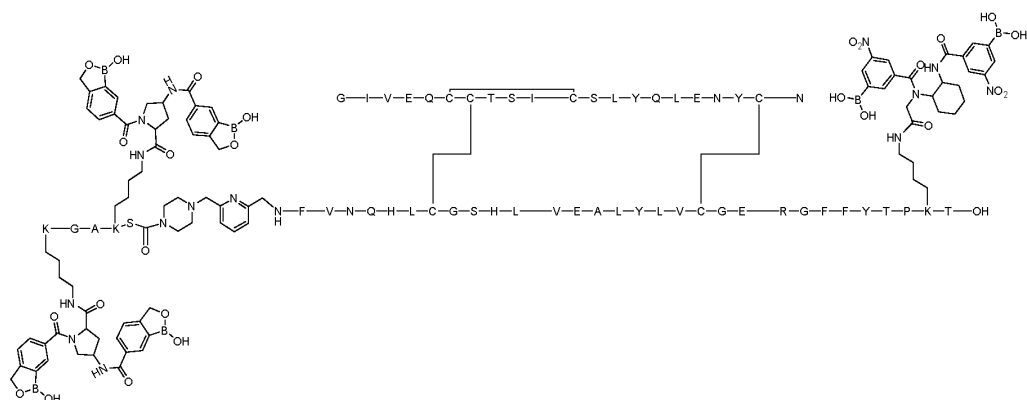
Example 355:



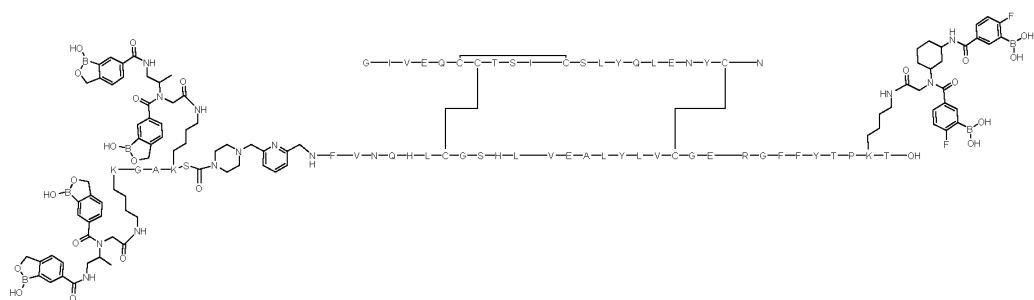
Example 356:

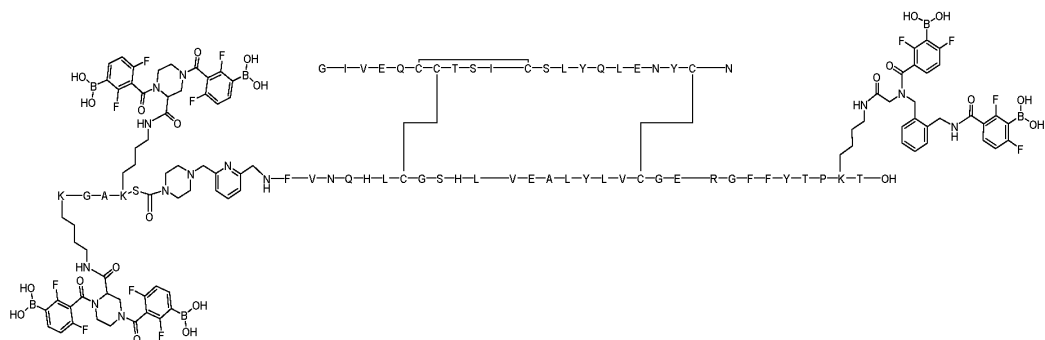


Example 357:



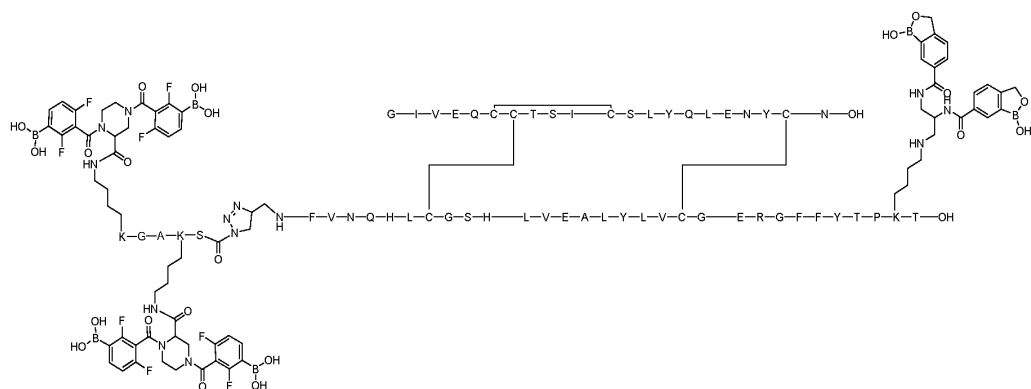
5 Example 358:



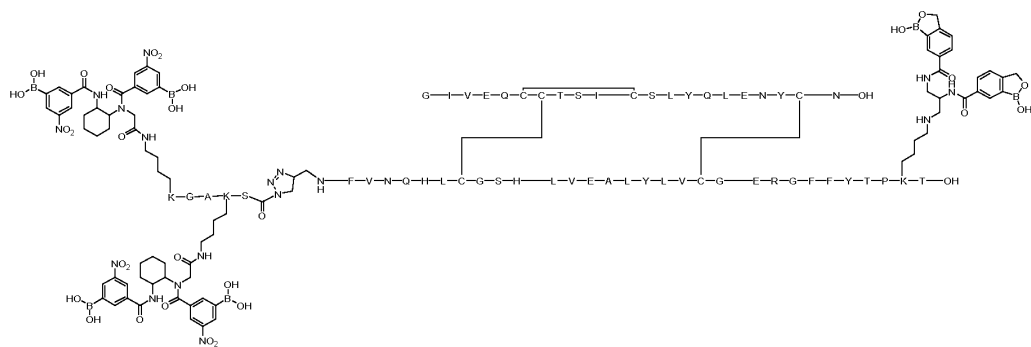


The diagram illustrates the chemical structure of the 120 kDa subunit of the 20S proteasome core particle. The structure is a complex polypeptide chain with various modifications and a central catalytic core. The chain is shown in a ribbon representation, with the main chain and side chains clearly defined. The structure is highly branched, with multiple side chains extending from the main chain. The central part of the structure features a catalytic core, which is a highly conserved region of the protein. The structure is shown in a ribbon representation, with the main chain and side chains clearly defined. The structure is highly branched, with multiple side chains extending from the main chain. The central part of the structure features a catalytic core, which is a highly conserved region of the protein. The structure is shown in a ribbon representation, with the main chain and side chains clearly defined. The structure is highly branched, with multiple side chains extending from the main chain. The central part of the structure features a catalytic core, which is a highly conserved region of the protein.

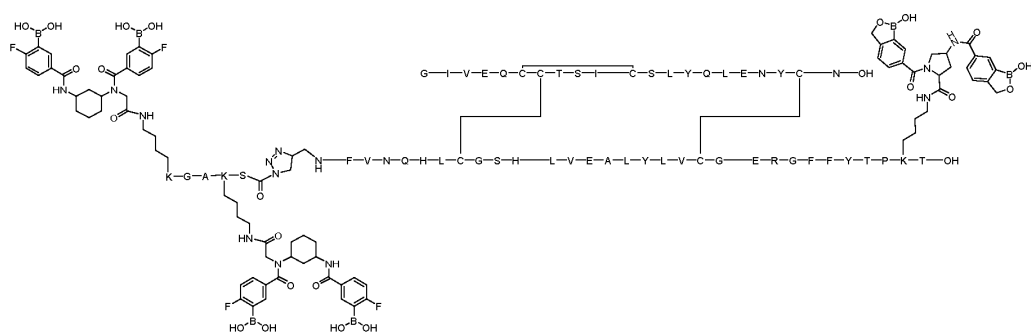
Example 362:



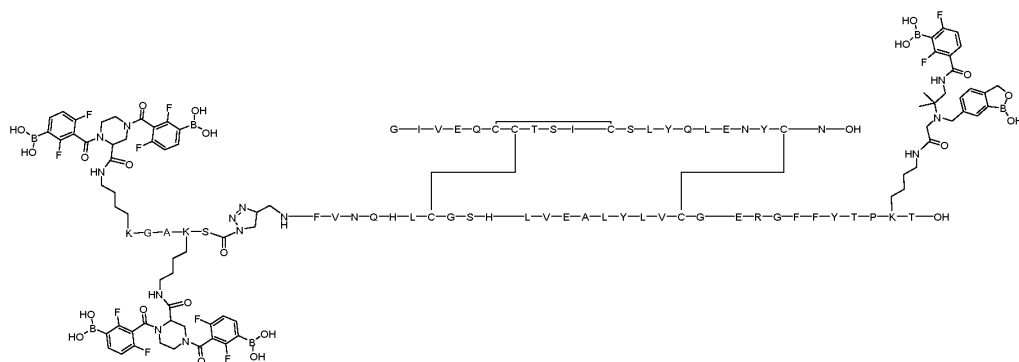
Example 363:



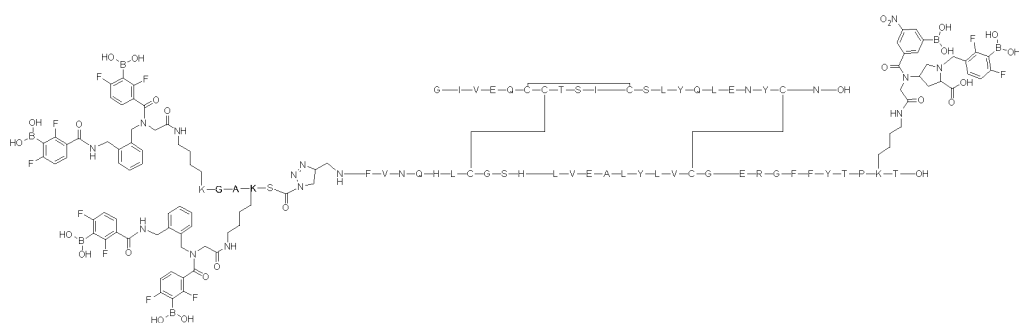
5 Example 364:



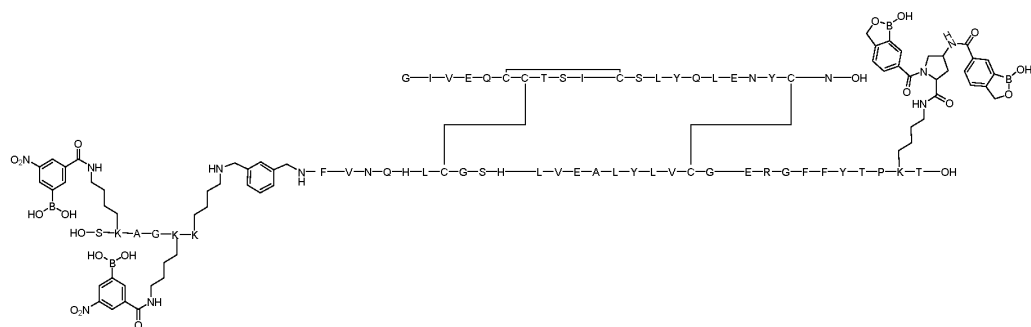
Example 365:



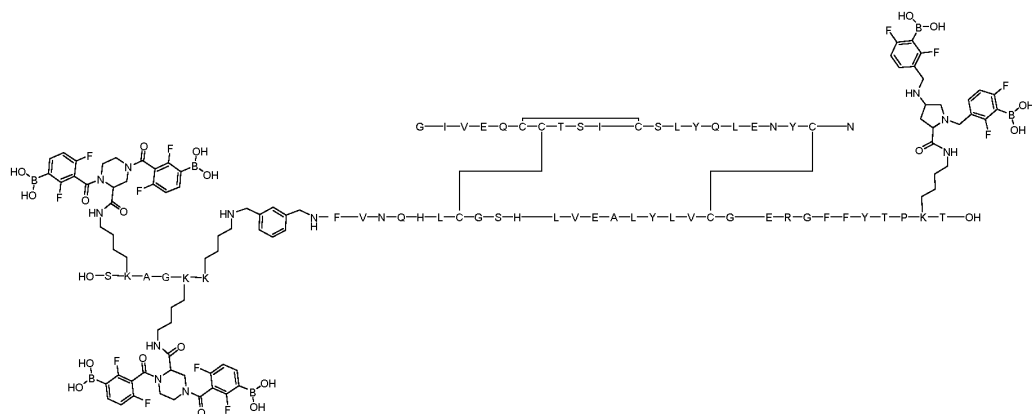
Example 366:



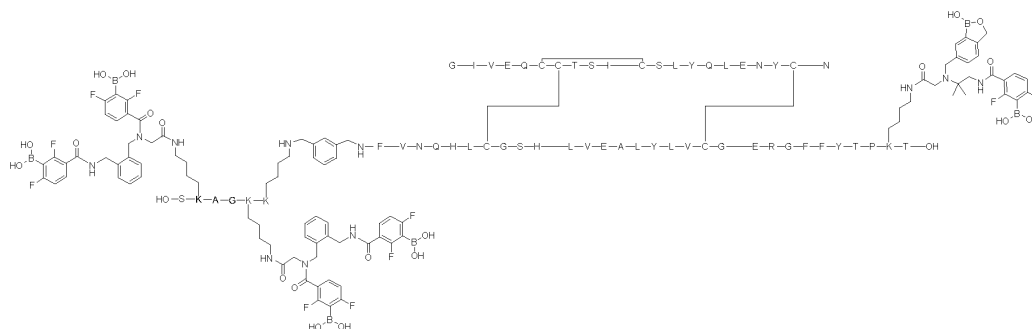
5 Example 367:



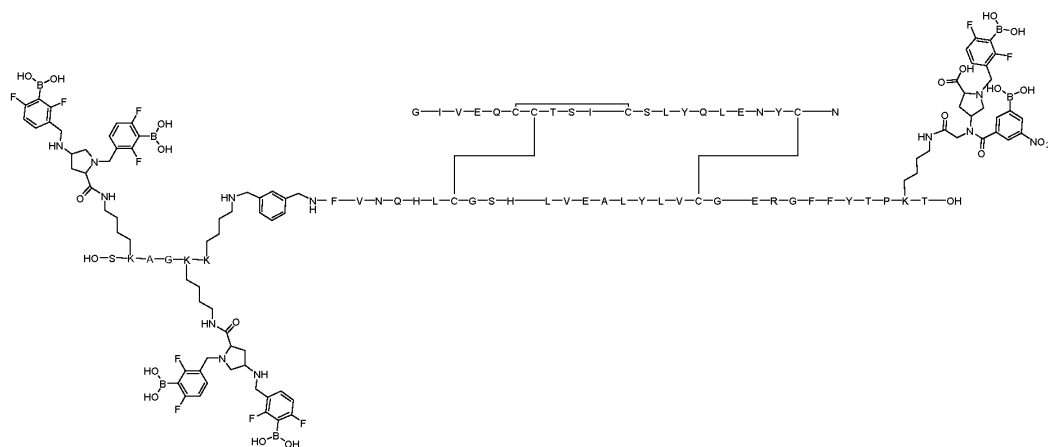
Example 368:



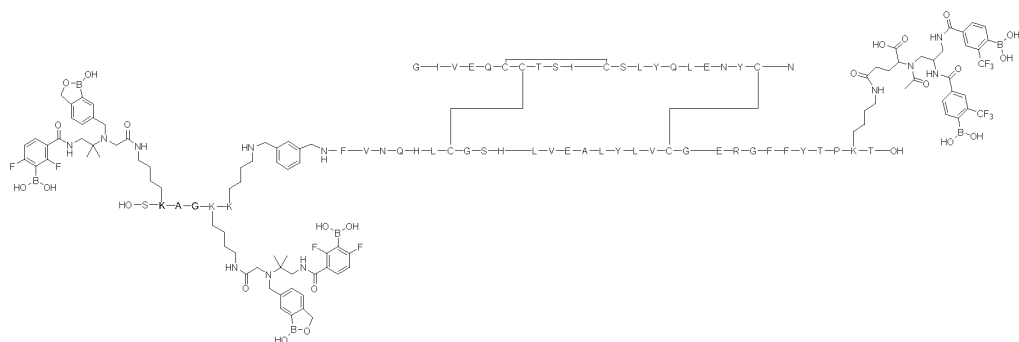
Example 369:



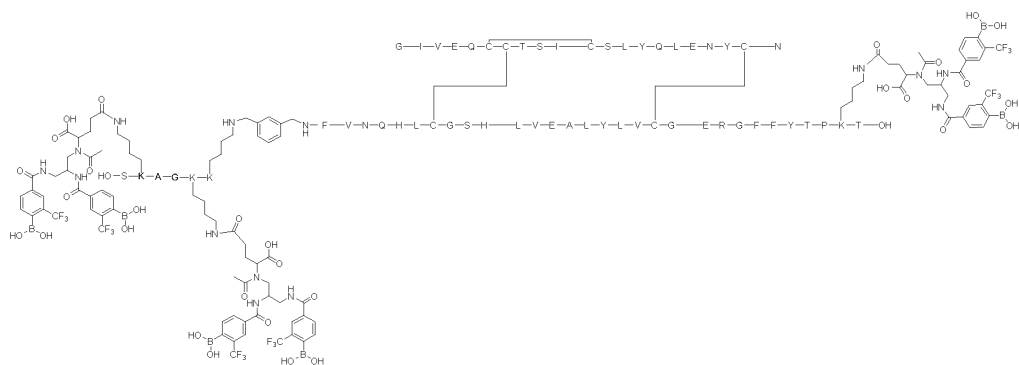
5 Example 370:



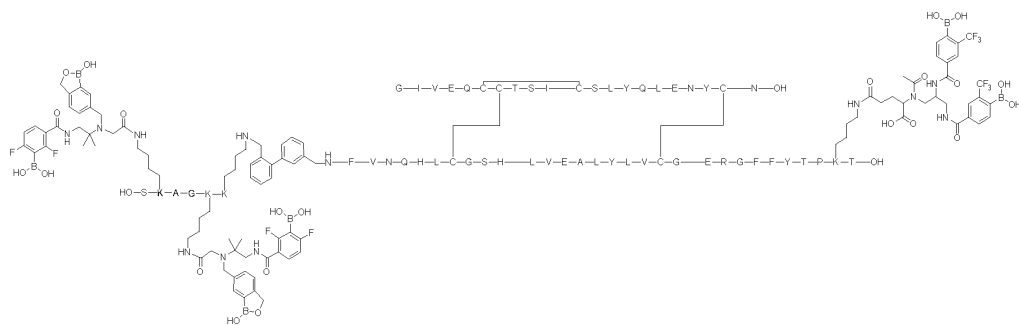
Example 371:

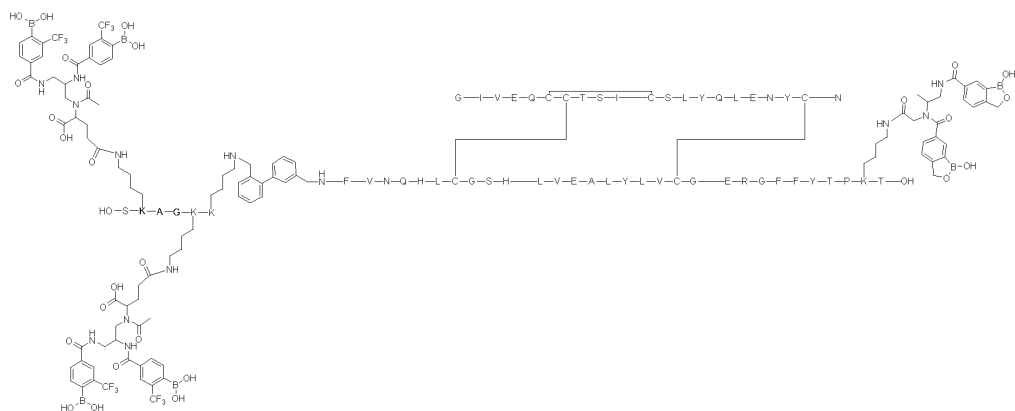
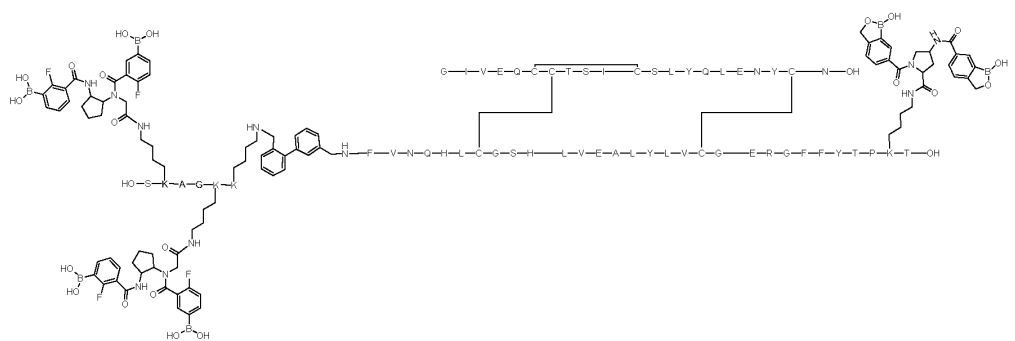
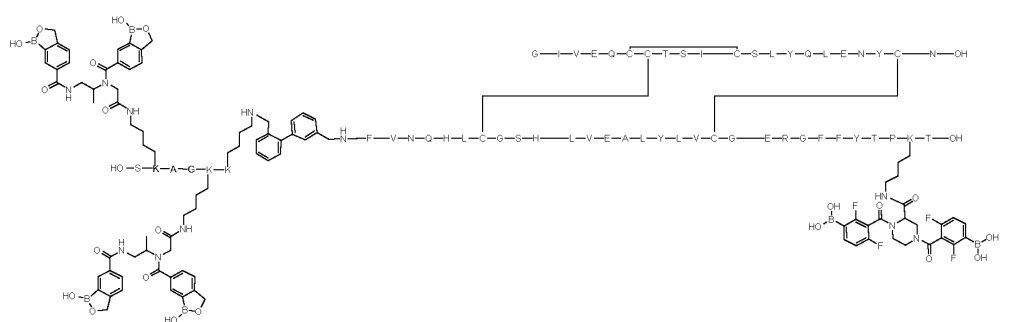


Example 372:

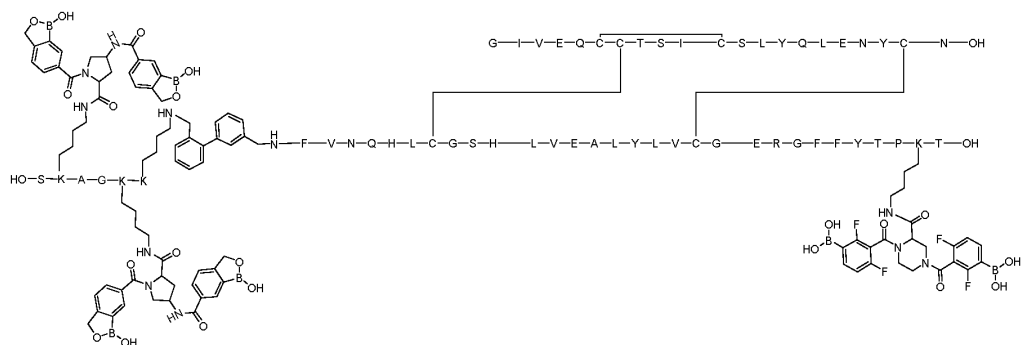


5 Example 373:

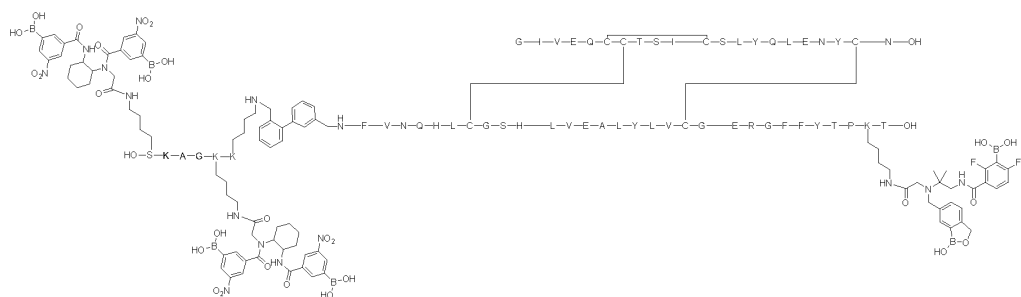


Example 374:**Example 375:****5 Example 376:**

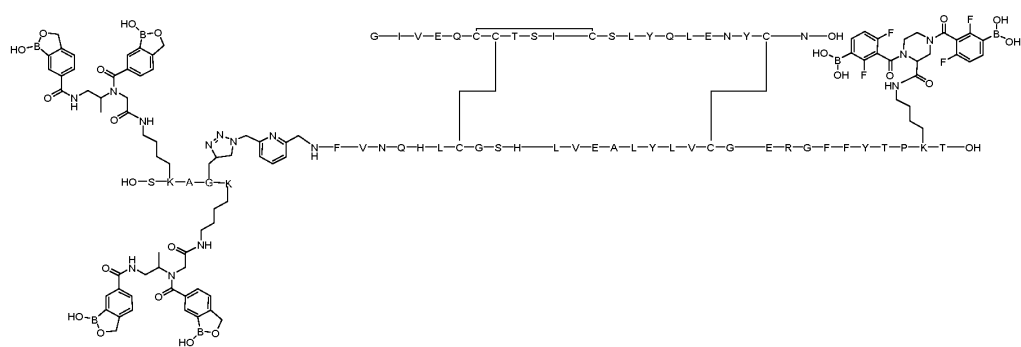
Example 377:

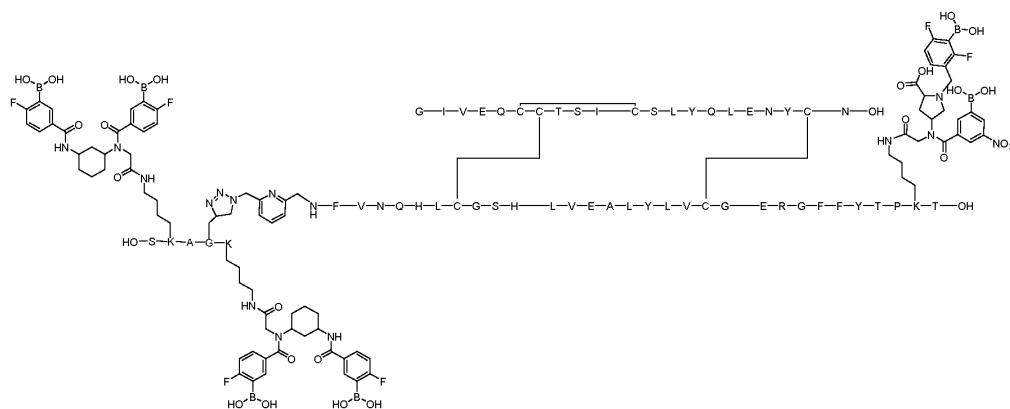
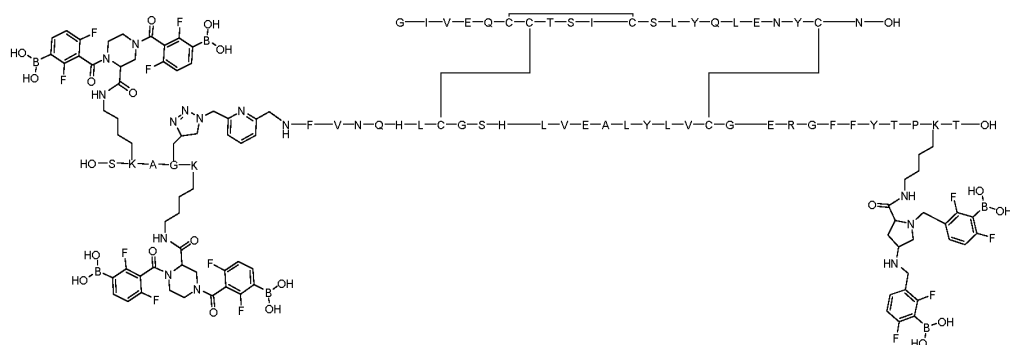
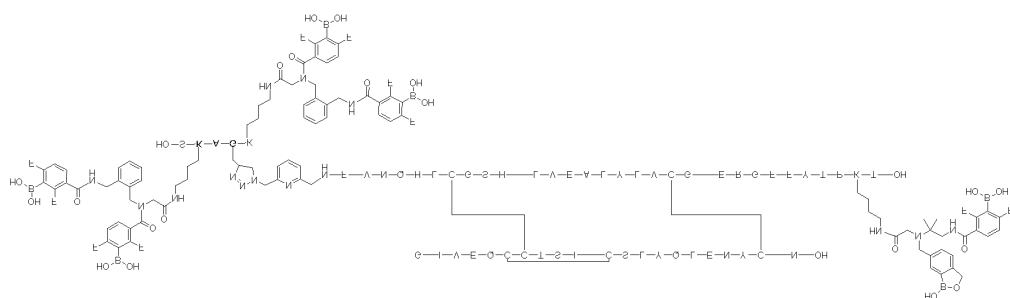


Example 378:

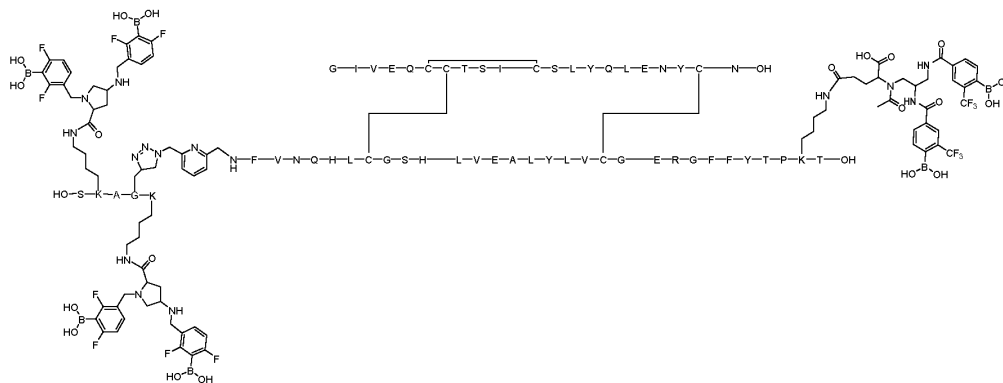


5 Example 379:

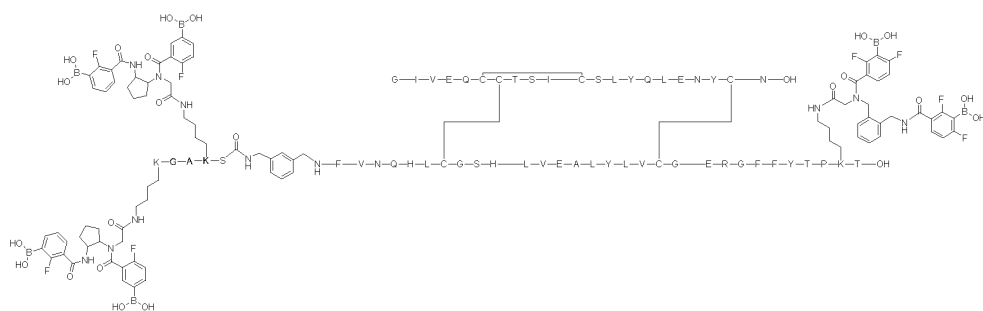


Example 380:**Example 381:****5 Example 382:**

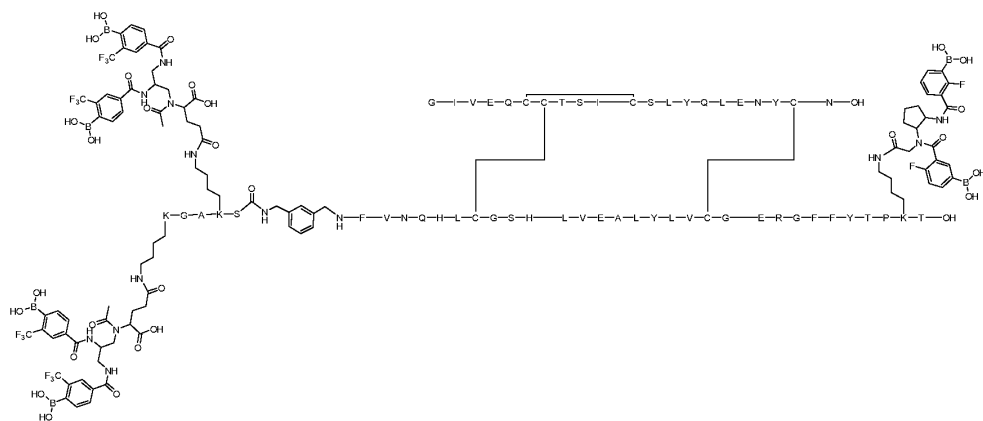
Example 383:



Example 384:



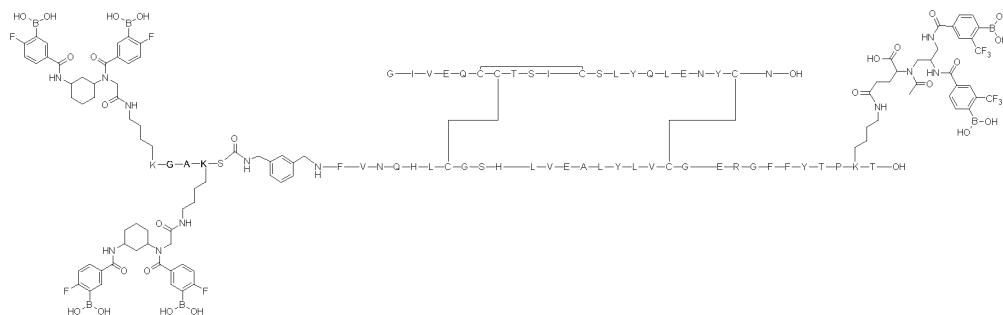
5 Example 385:



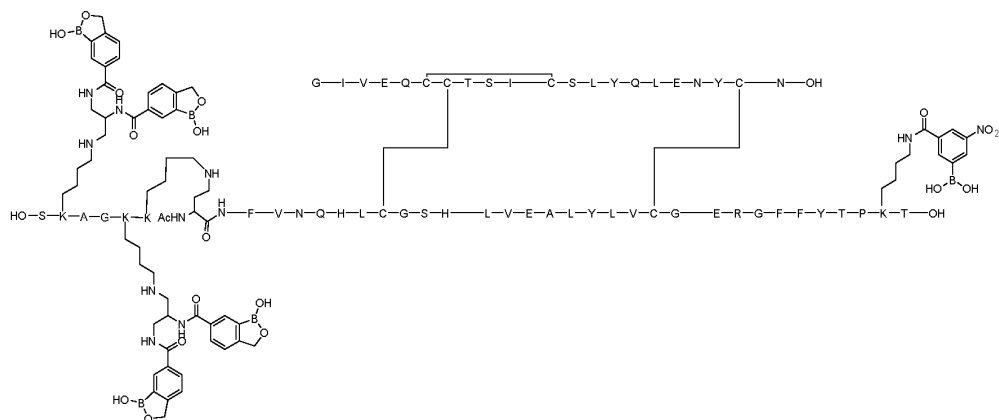
[illegible][illegible]

The chemical structure of the 100% de novo designed protein is shown. The protein backbone is represented by a series of alpha-carbons (C) and beta-carbons (C) connected by peptide bonds. The side chains are shown as R groups. The structure includes a catalytic site with residues E1, E2, E3, E4, E5, E6, E7, E8, E9, E10, E11, E12, E13, E14, E15, E16, E17, E18, E19, E20, E21, E22, E23, E24, E25, E26, E27, E28, E29, E30, E31, E32, E33, E34, E35, E36, E37, E38, E39, E40, E41, E42, E43, E44, E45, E46, E47, E48, E49, E50, E51, E52, E53, E54, E55, E56, E57, E58, E59, E60, E61, E62, E63, E64, E65, E66, E67, E68, E69, E70, E71, E72, E73, E74, E75, E76, E77, E78, E79, E80, E81, E82, E83, E84, E85, E86, E87, E88, E89, E90, E91, E92, E93, E94, E95, E96, E97, E98, E99, E100. The structure also shows the substrate binding site, which is a pocket formed by residues E1, E2, E3, E4, E5, E6, E7, E8, E9, E10, E11, E12, E13, E14, E15, E16, E17, E18, E19, E20, E21, E22, E23, E24, E25, E26, E27, E28, E29, E30, E31, E32, E33, E34, E35, E36, E37, E38, E39, E40, E41, E42, E43, E44, E45, E46, E47, E48, E49, E50, E51, E52, E53, E54, E55, E56, E57, E58, E59, E60, E61, E62, E63, E64, E65, E66, E67, E68, E69, E70, E71, E72, E73, E74, E75, E76, E77, E78, E79, E80, E81, E82, E83, E84, E85, E86, E87, E88, E89, E90, E91, E92, E93, E94, E95, E96, E97, E98, E99, E100. The structure is shown in a ribbon representation, with the backbone in blue and the side chains in red. The catalytic site is highlighted in yellow, and the substrate binding site is highlighted in green.

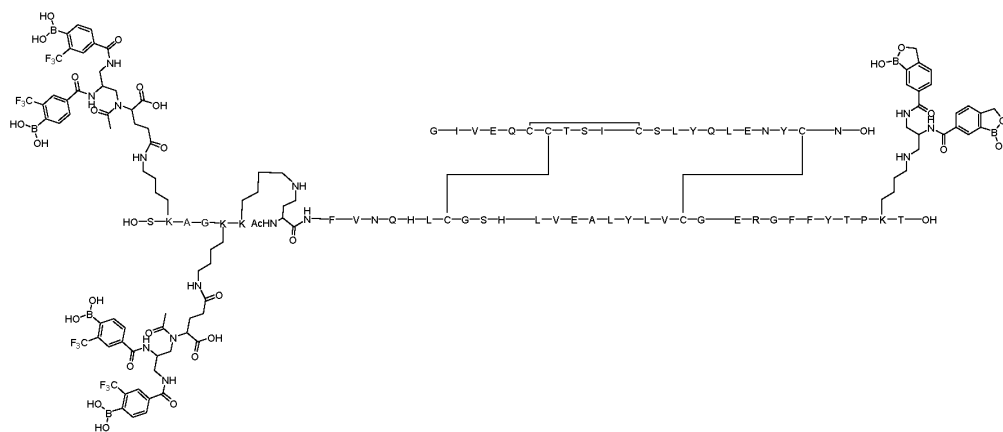
Example 389:

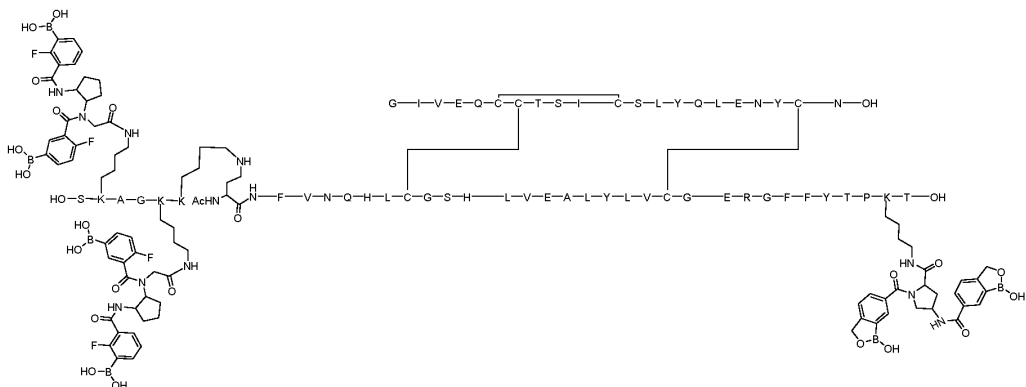


Example 390:

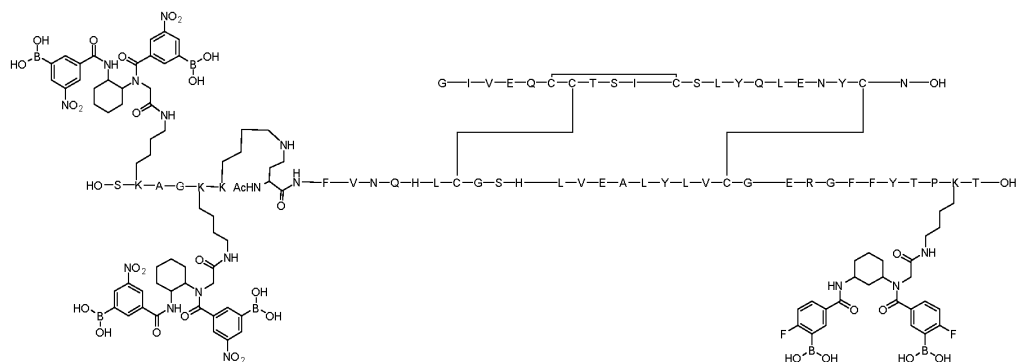


5 Example 391:

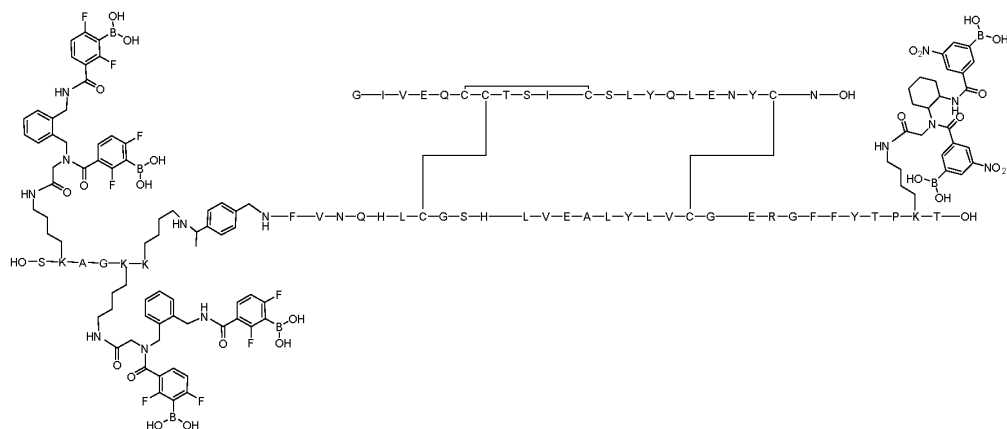




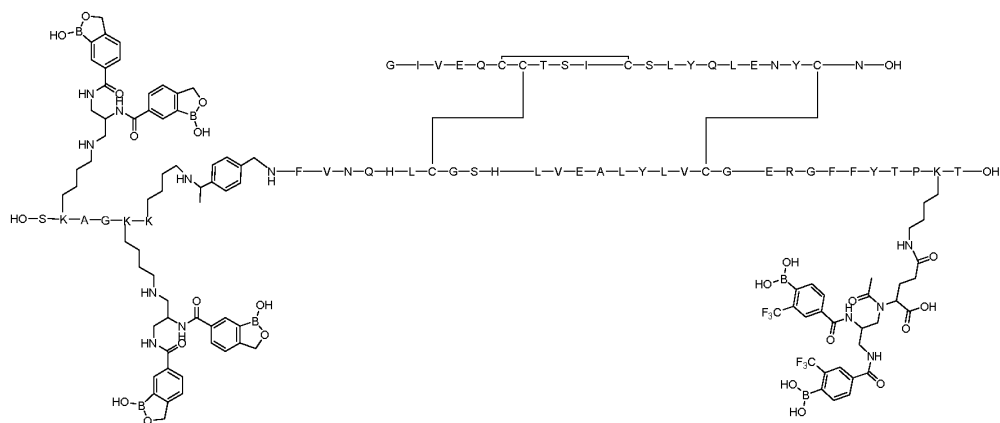
Example 395:



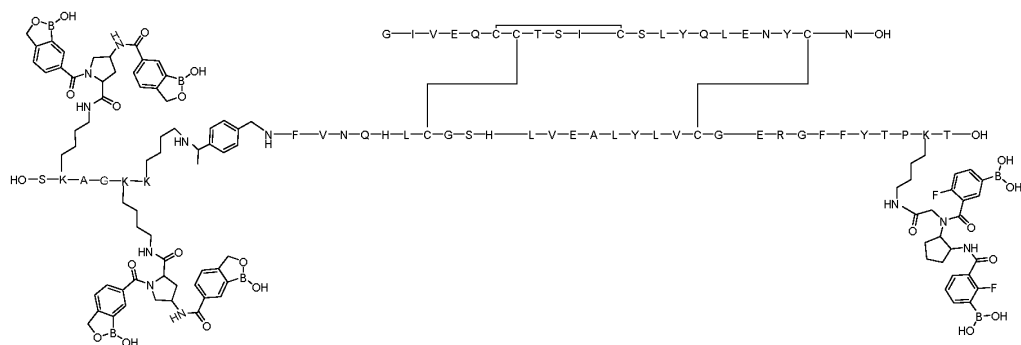
Example 396:



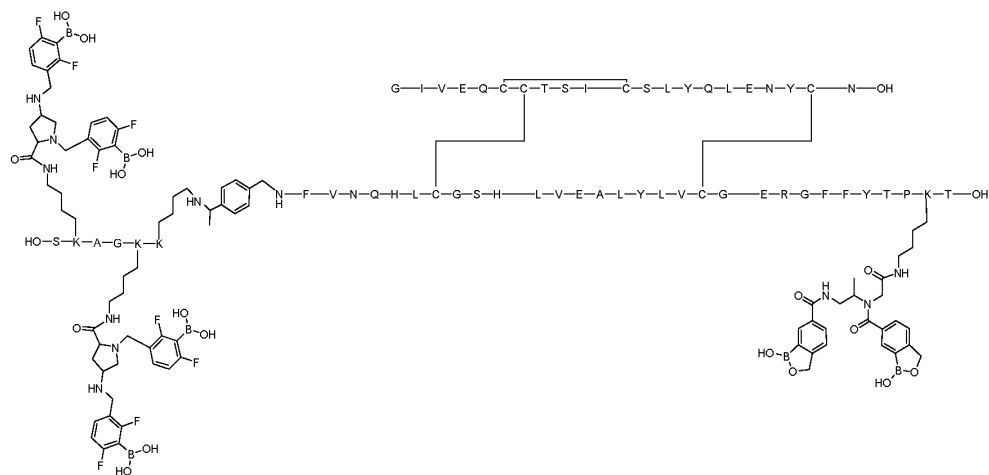
5 Example 397:



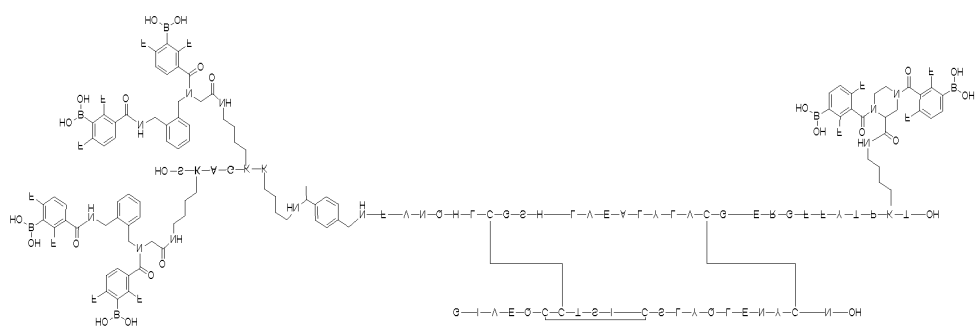
Example 398:

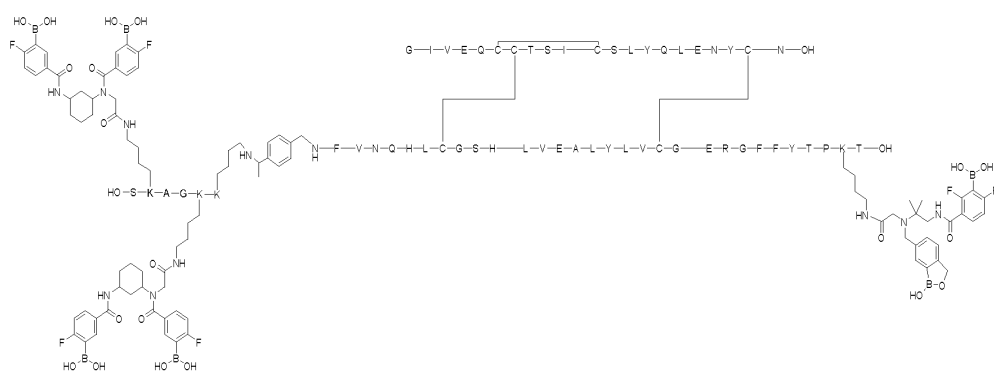


Example 399:

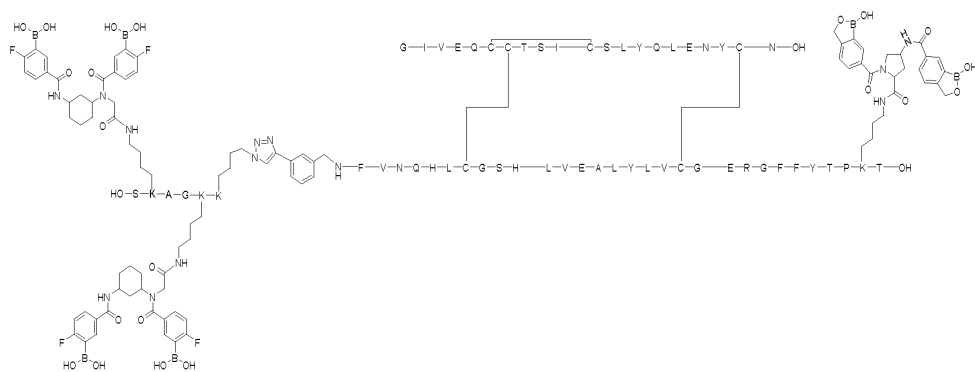
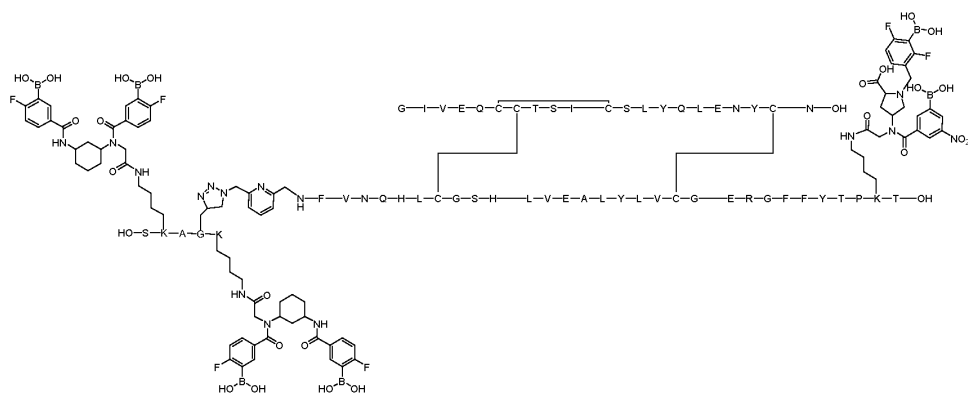


5 Example 400:

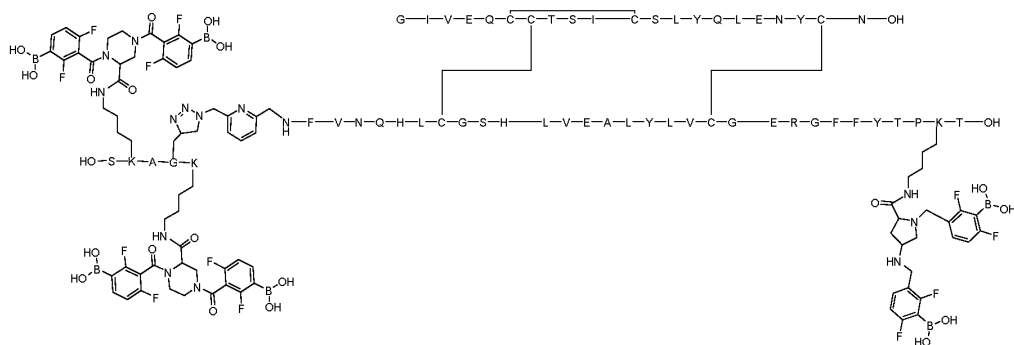


Example 401:**Example 402:**

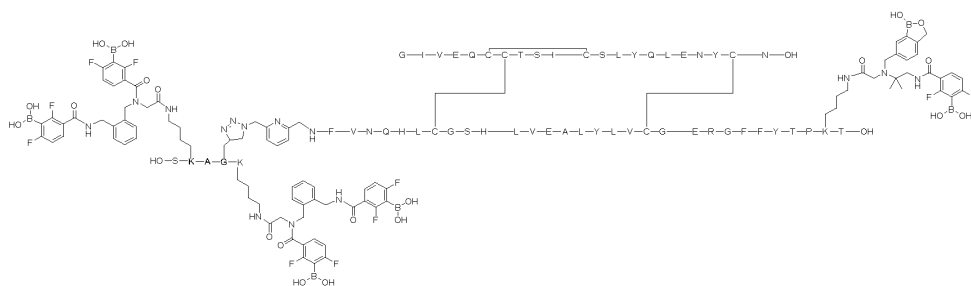
5

**Example 403:**

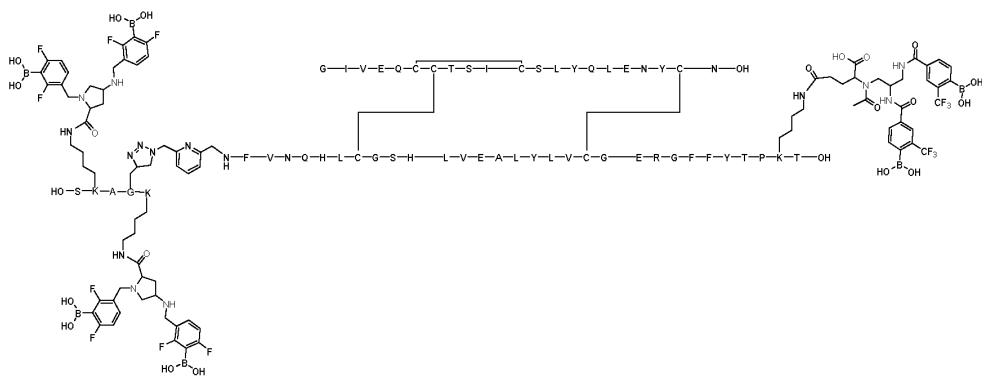
Example 404:



Example 405:



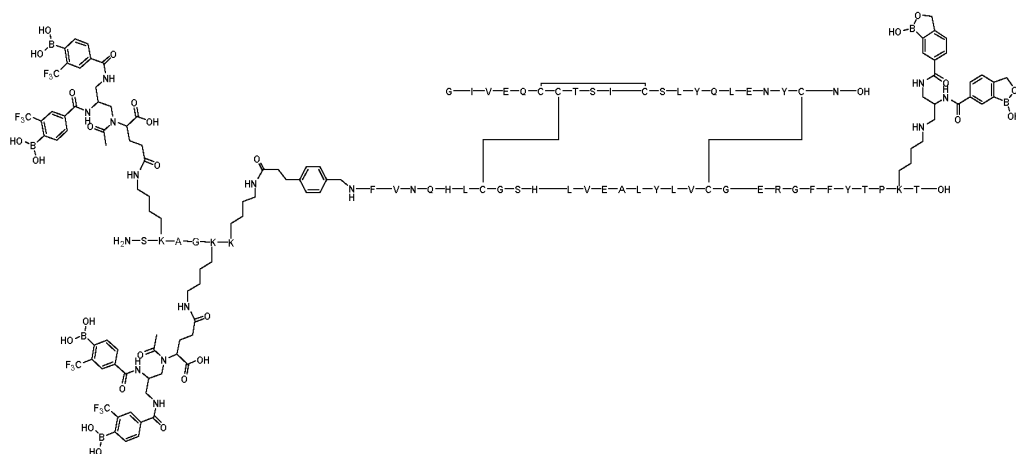
5 Example 406:



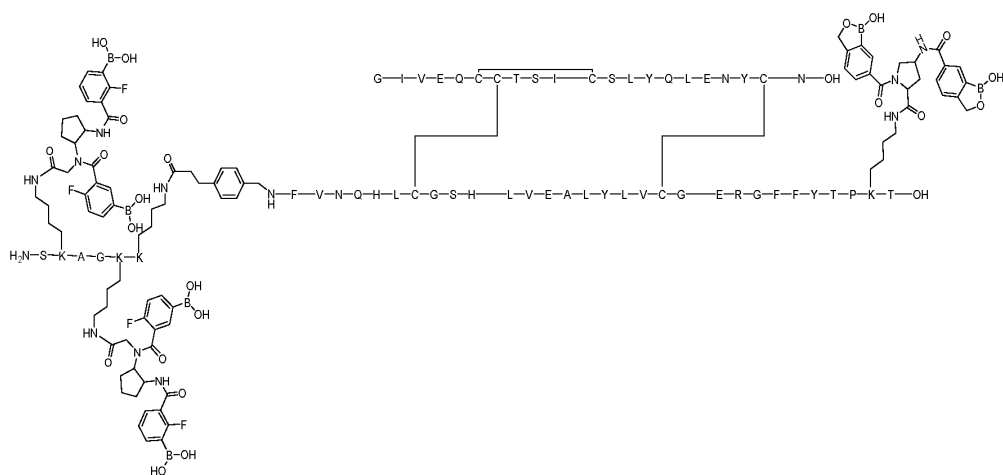
Chemical structure of the 100-residue protein 100R, showing its linear sequence and various modifications. The sequence is G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH. The structure includes a C-terminal signal peptide (residues 1-24) with a hydrophobic core, a proline-rich region (residues 25-40), a flexible loop (residues 41-50), a catalytic domain (residues 51-100) with a zinc-binding site, and a C-terminal domain (residues 101-120) with a zinc-binding site. The structure is shown in a ribbon representation with various modifications highlighted in different colors.

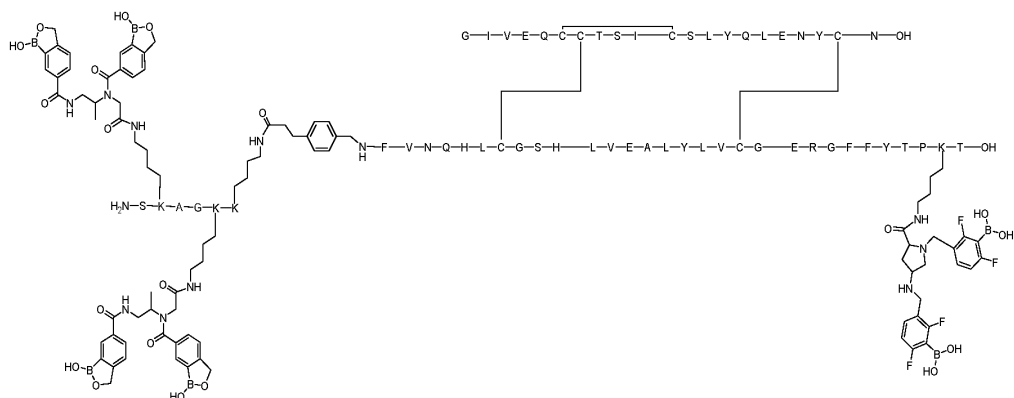
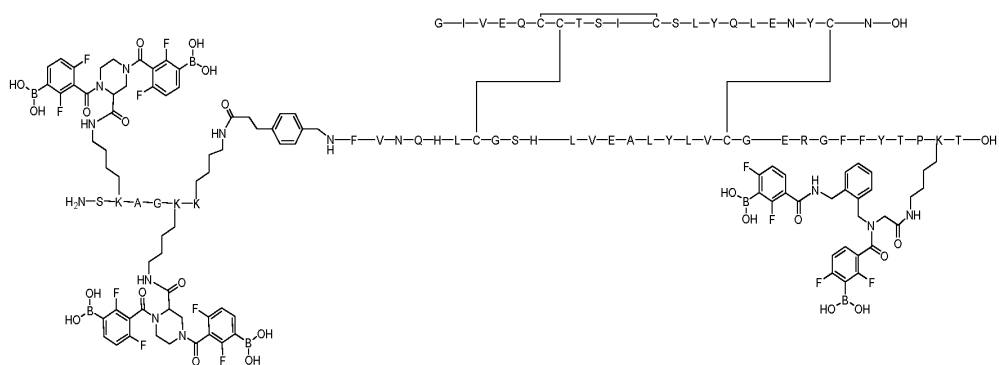
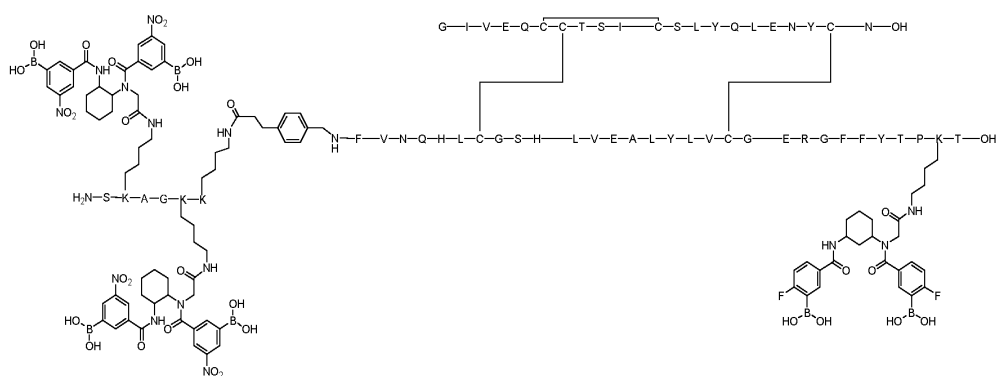
The image displays a complex chemical structure of a peptide molecule, likely a viral capsid protein. The structure features a central peptide backbone with various side chains, including aromatic rings, hydroxyl groups, and a boron-containing moiety. The molecule is highly branched and contains several functional groups, such as carboxylic acids and alcohols. The structure is labeled with various amino acid abbreviations (e.g., H⁺N, C, N, O, H, G, E, V, F, L, A, C, E, B, G, L, L, A, L, L, K, I, OH) and includes a detailed representation of the boron-containing moiety at the bottom right.

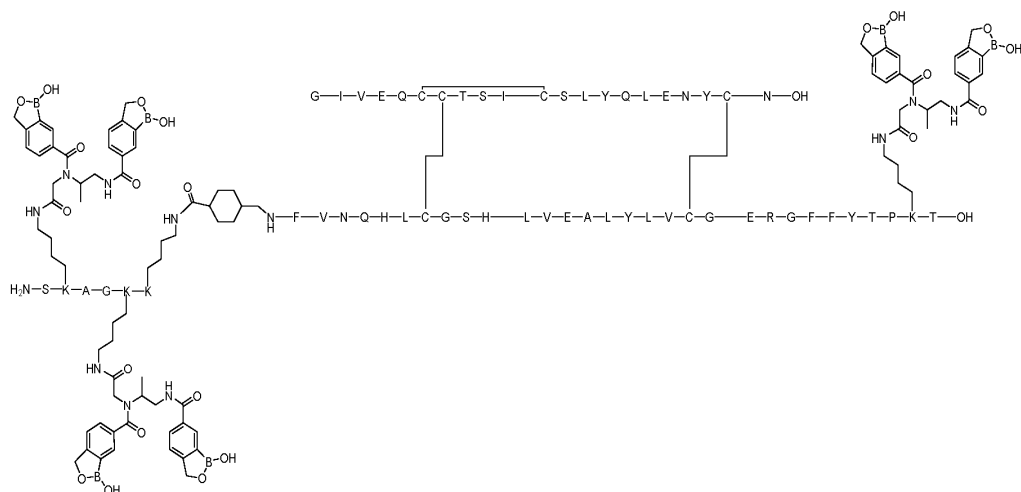
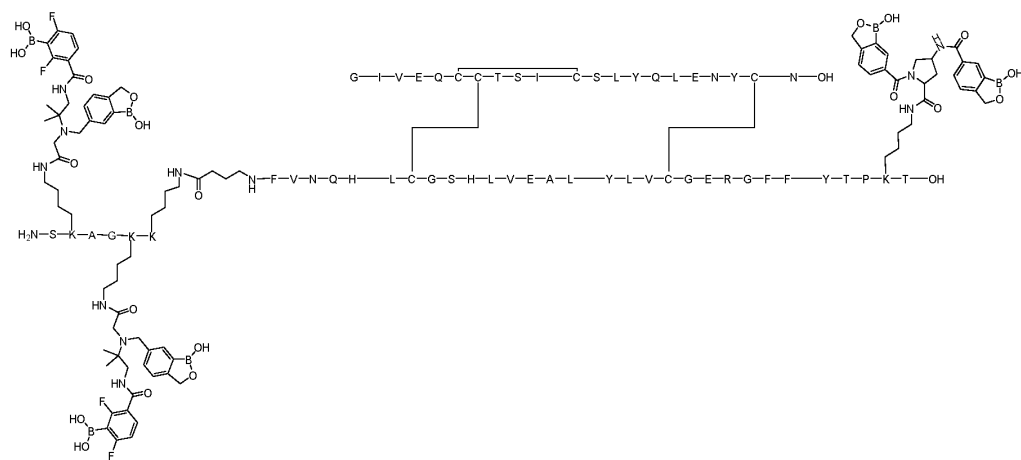
Example 410:

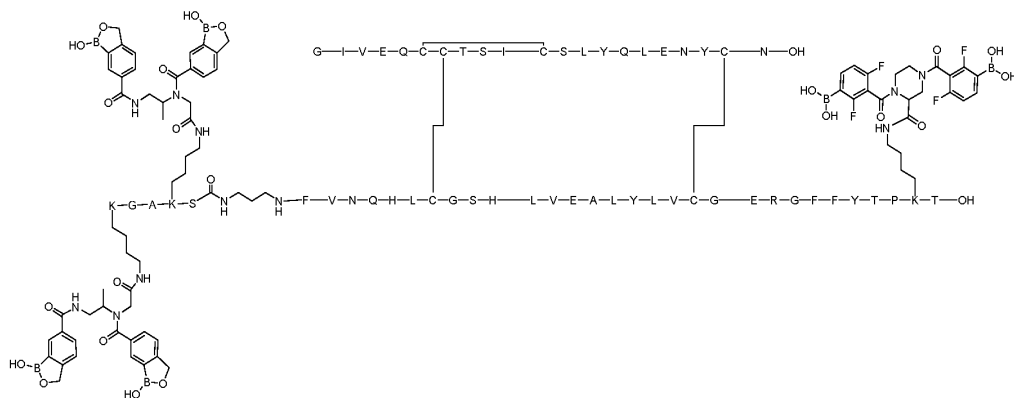
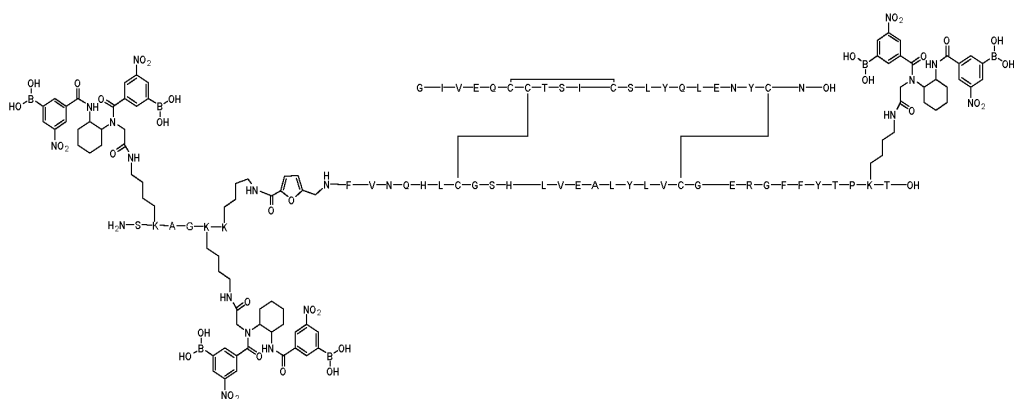


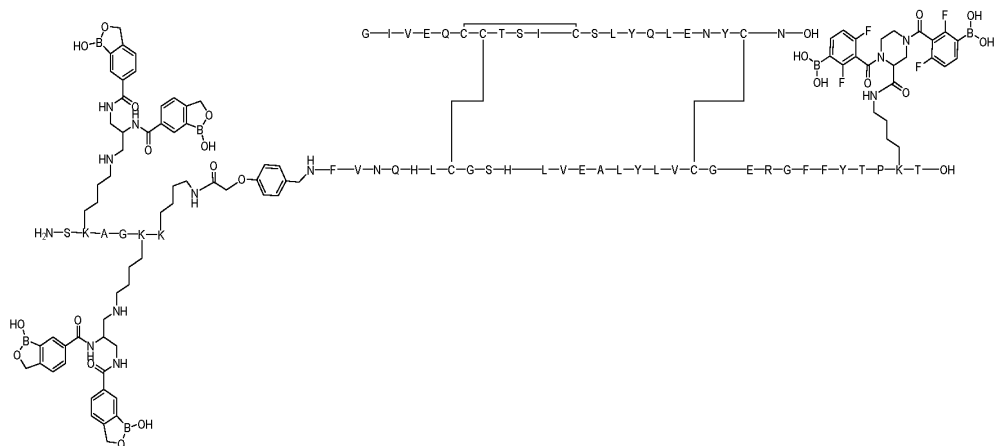
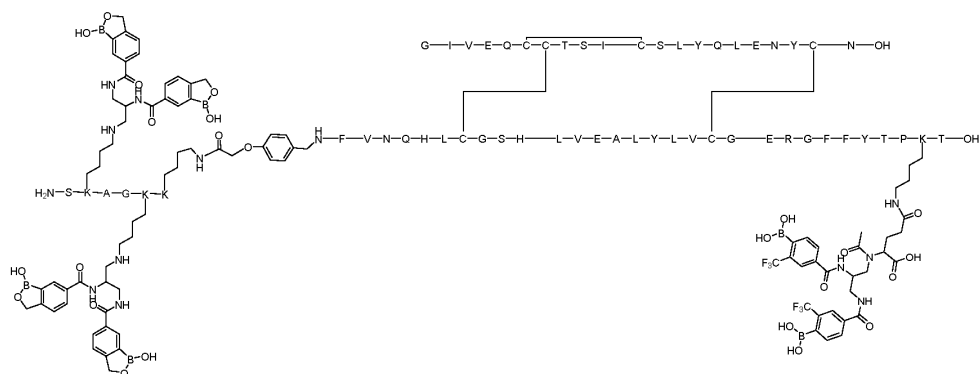
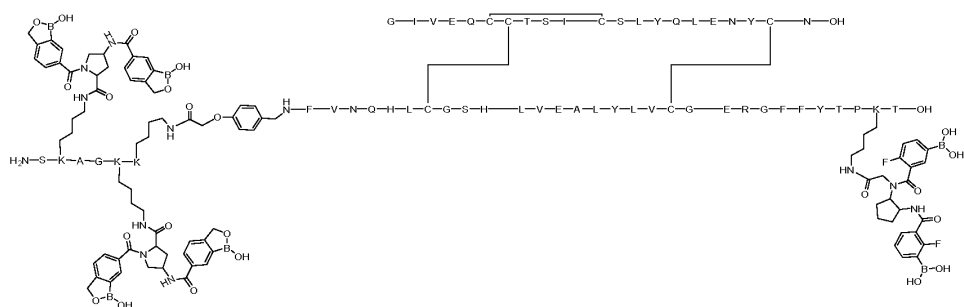
Example 411:



Example 412**Example 413:****5 Example 414:**

Example 415:**Example 416:**

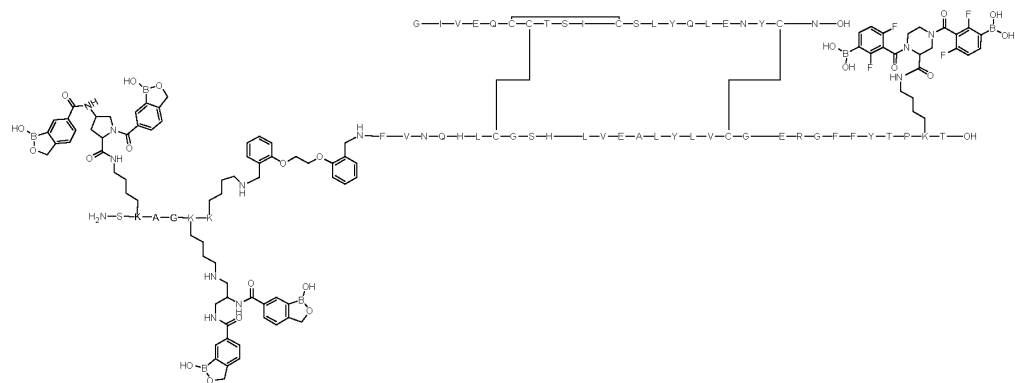
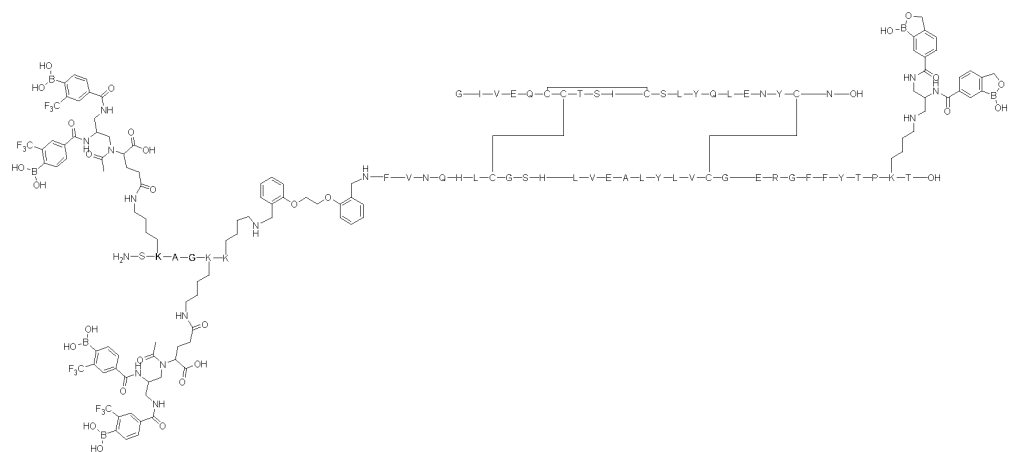
Example 417:**Example 418:**

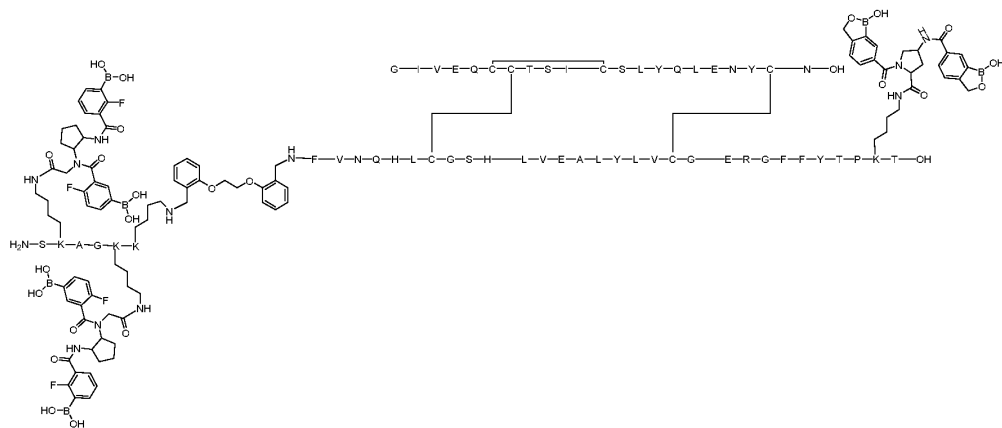
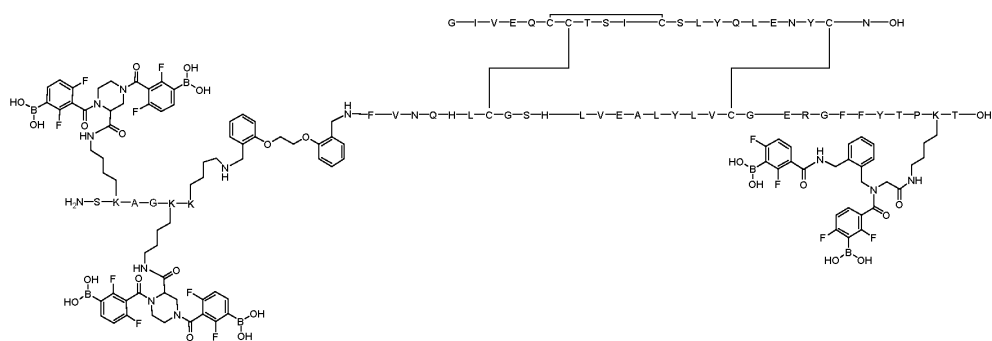
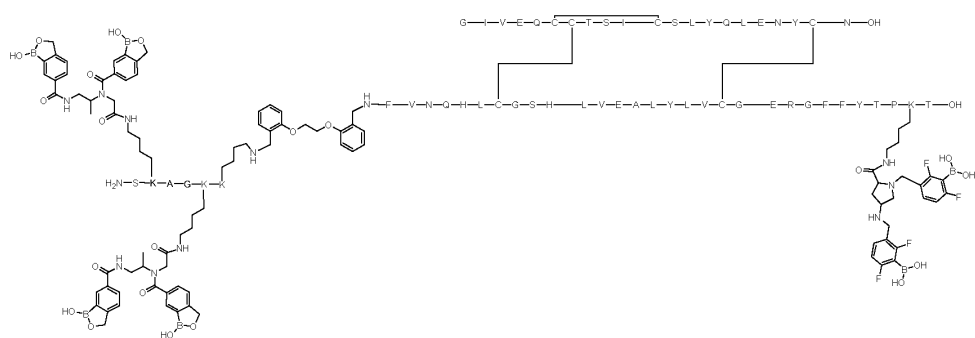
Example 419:**Example 420:****5 Example 421:**

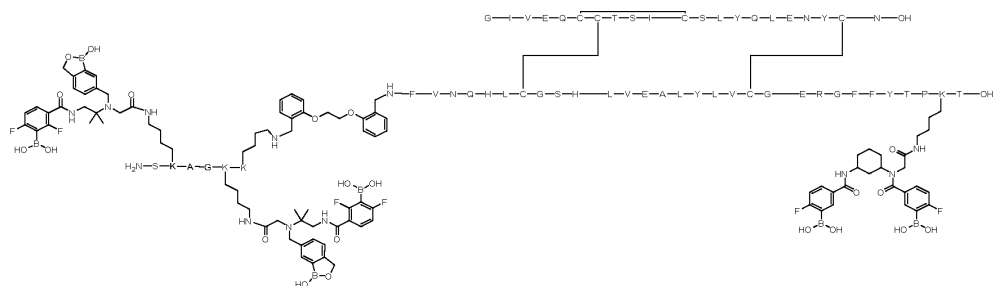
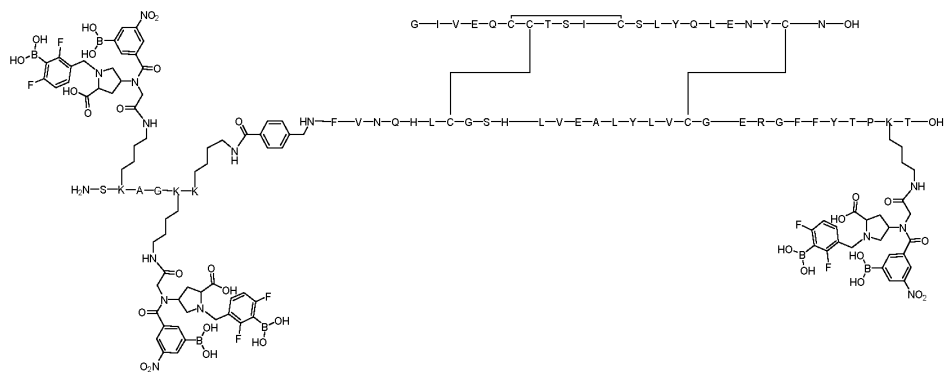
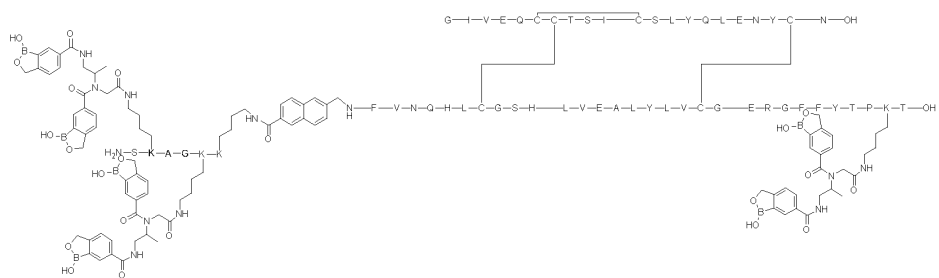
[illegible]

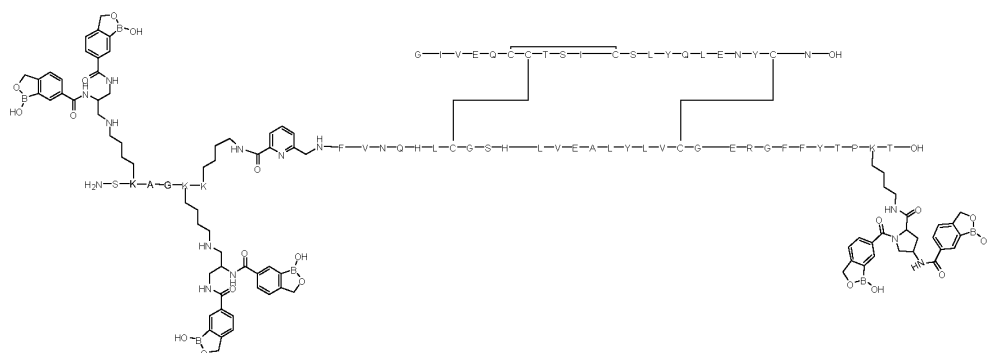
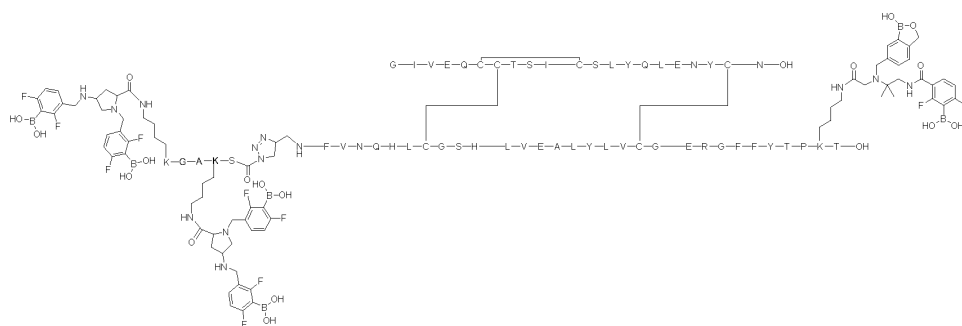
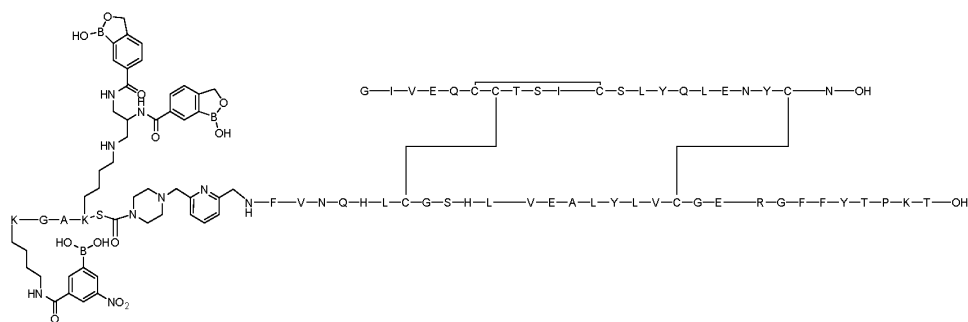
The chemical structure shows a central poly(amide amine) backbone, represented as $\text{G}-(\text{V-E-Q})_n-(\text{C-T-S})_m-(\text{C-S-L-Y-Q-L-E-N-H-Y})_p-(\text{C-N-OH})_q$. This backbone is functionalized with several side chains:

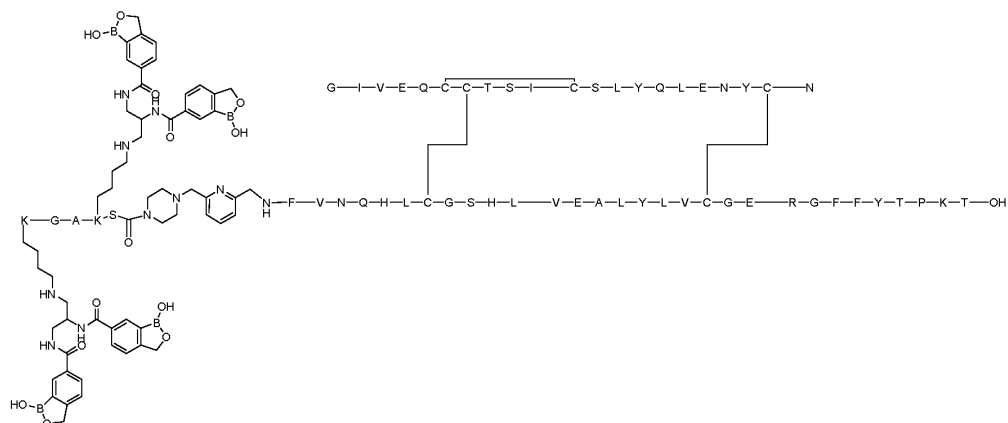
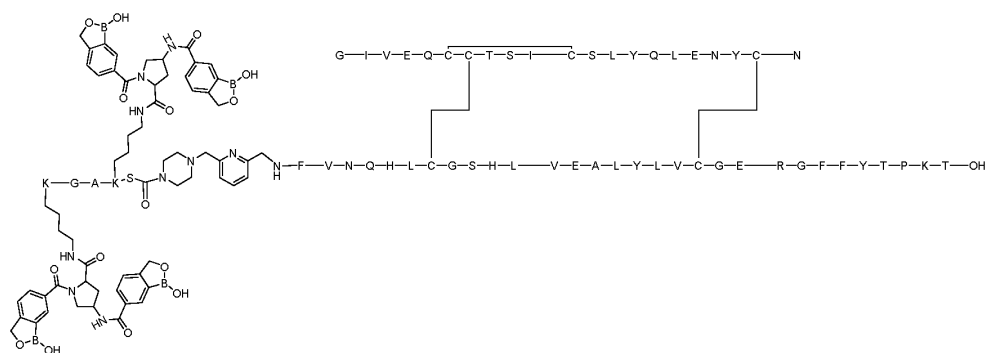
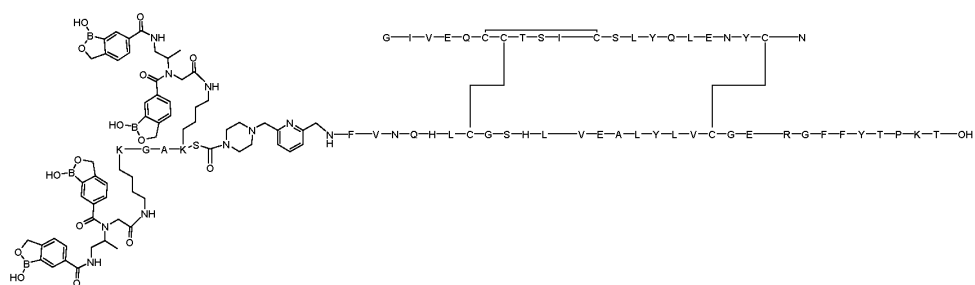
- Left side:** A carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$ is linked via an amide bond to a cyclohexane ring. This ring is further linked via an amide bond to a side chain containing a carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$ and a carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$.
- Right side:** A carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$ is linked via an amide bond to a cyclohexane ring. This ring is further linked via an amide bond to a side chain containing a carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$ and a carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$.
- Bottom:** A carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$ is linked via an amide bond to a cyclohexane ring. This ring is further linked via an amide bond to a side chain containing a carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$ and a carboxylic acid group $\text{HO}_2\text{C}-\text{C}_6\text{H}_3(\text{F})-\text{C}(=\text{O})-$.

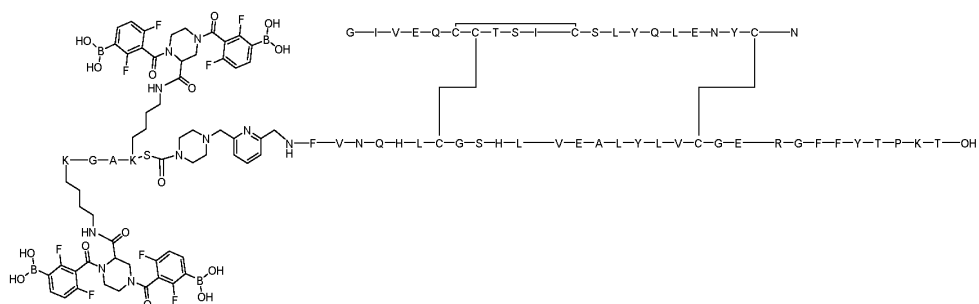
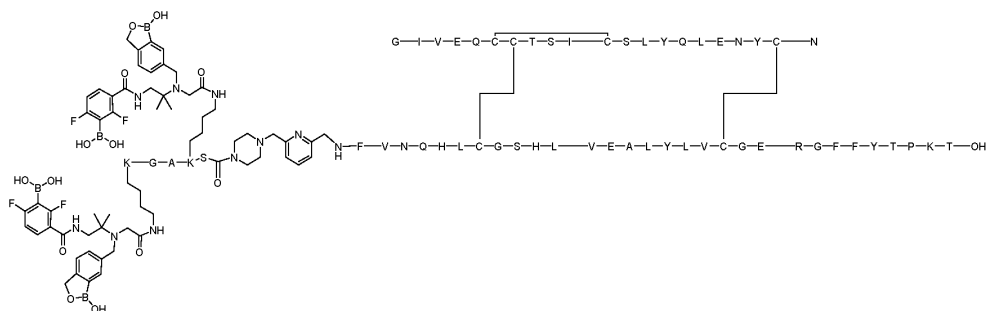
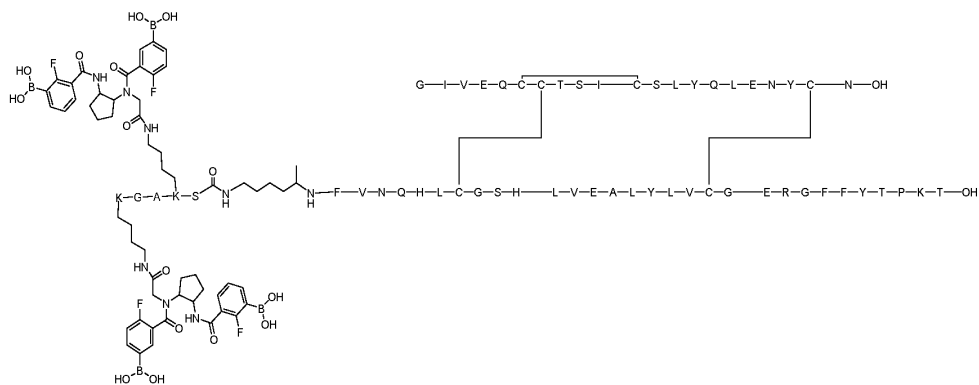
Example 425:**Example 426:**

Example 427:**Example 428:****5 Example 429:**

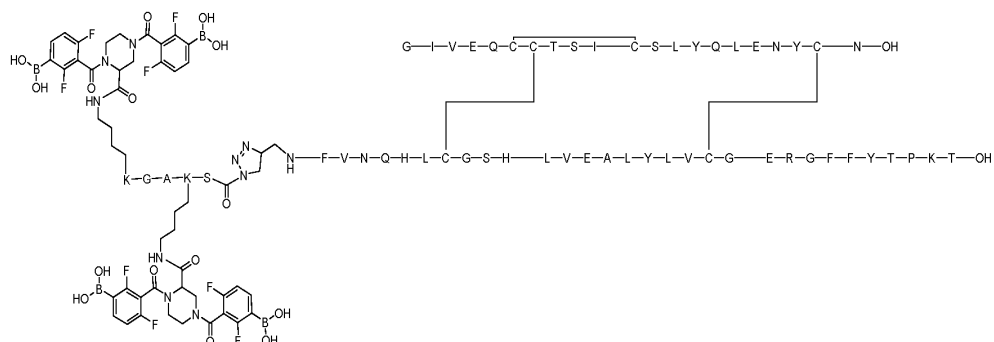
Example 430:**Example 431****5 Example 432:**

Example 433:**Example 434:****5 Example 435:**

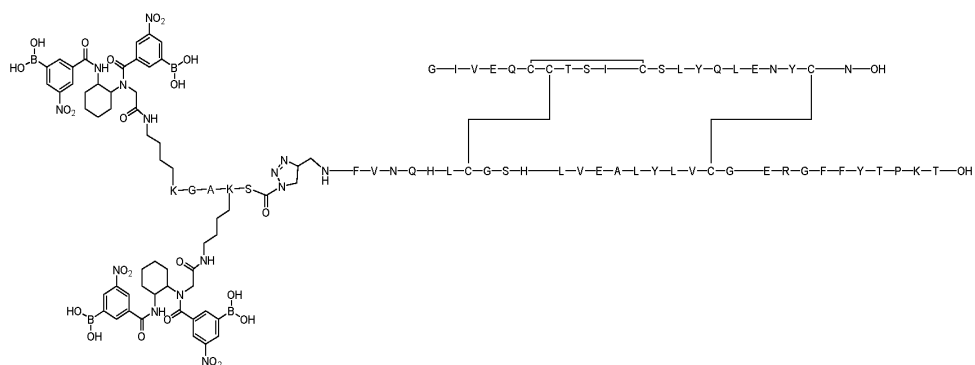
Example 436:**Example 437:****5 Example 438:**

Example 439:**Example 440:****5 Example 441:**

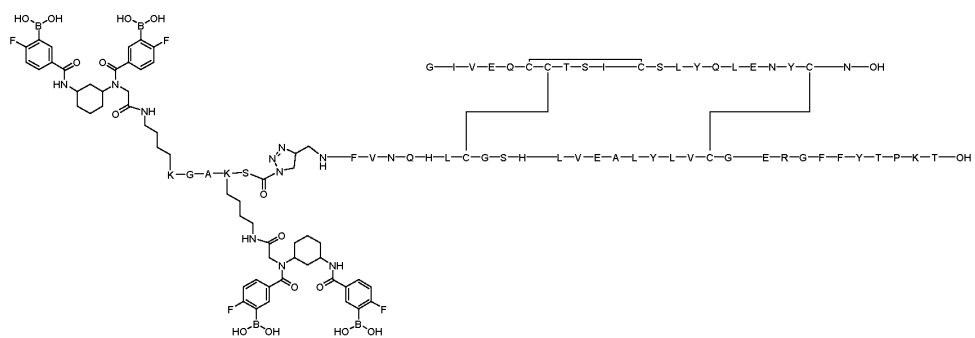
Example 442:

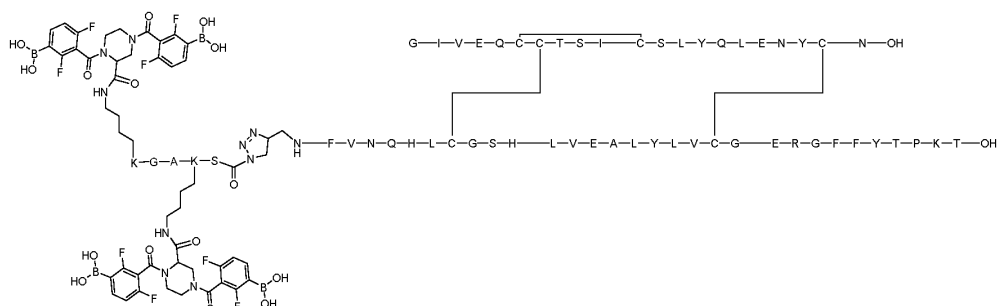


Example 443:

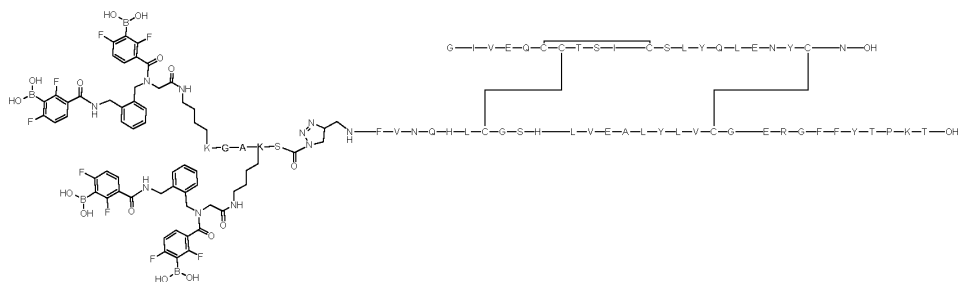


5 Example 444:

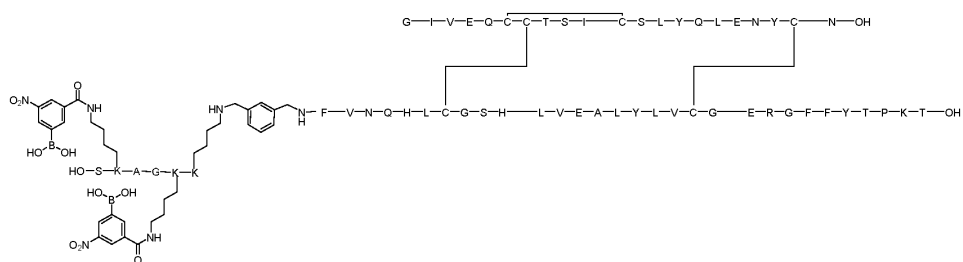


Example 445:

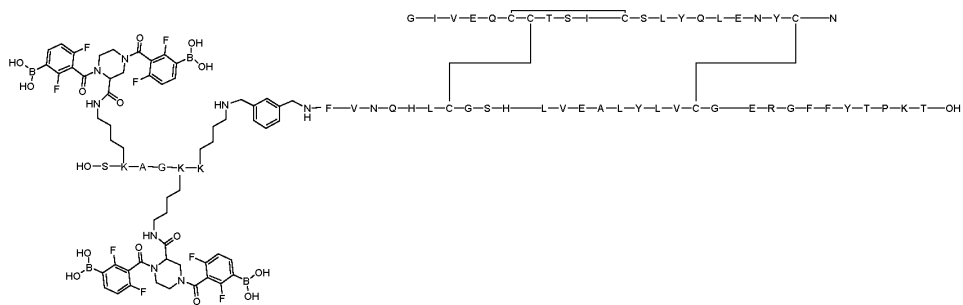
Example 446:

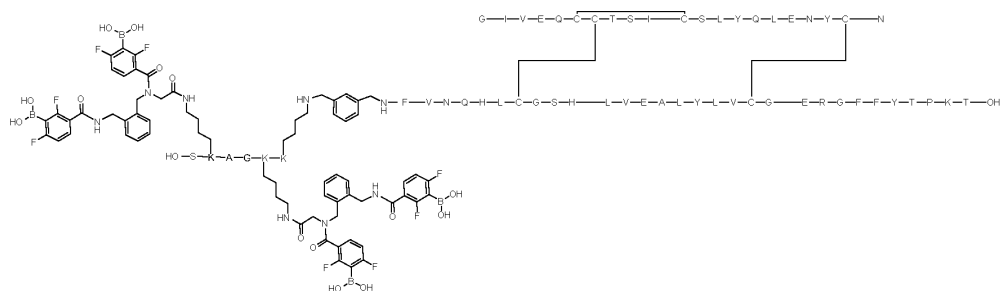
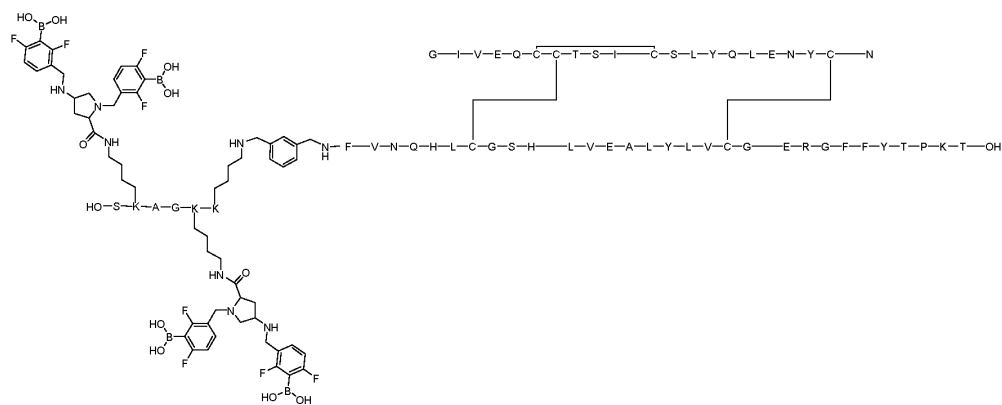
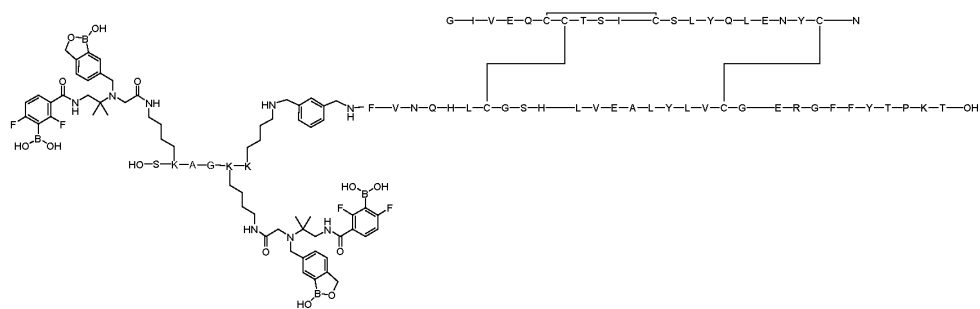


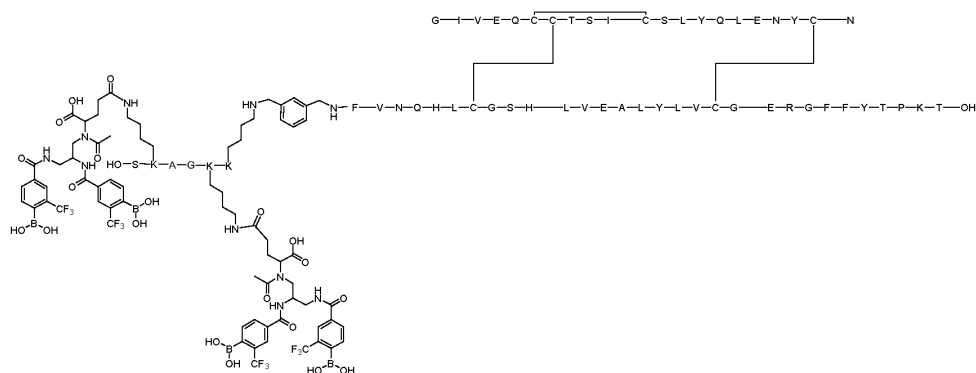
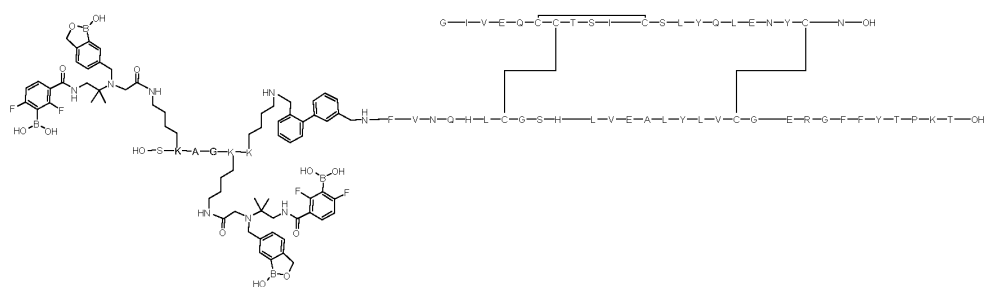
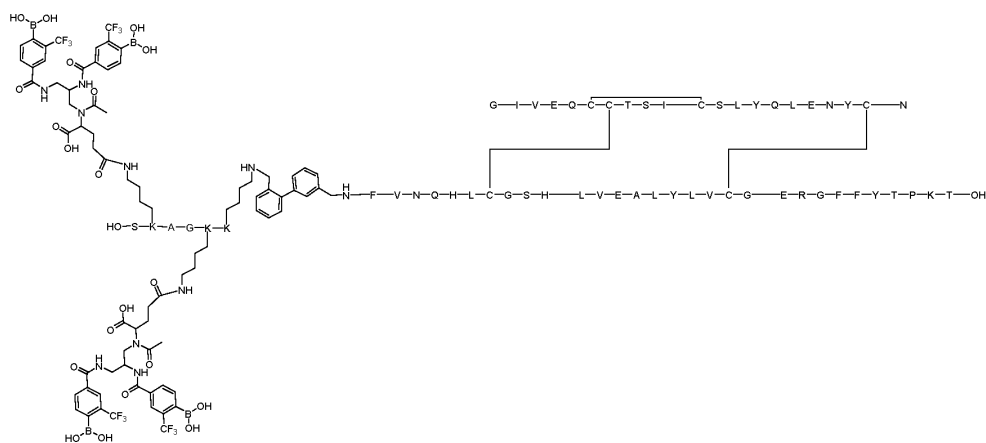
5 Example 447:

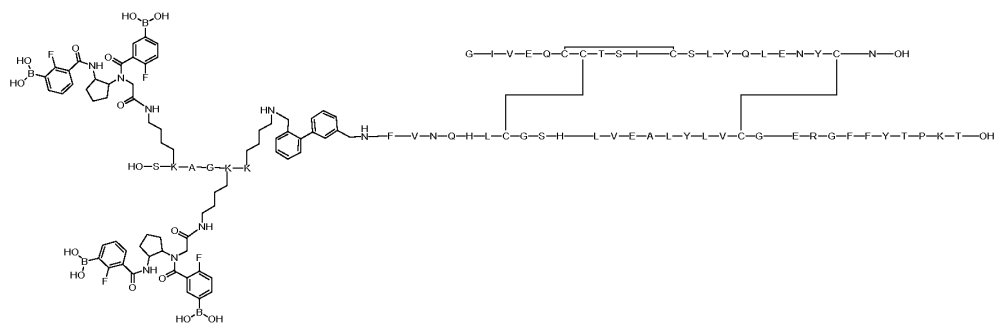
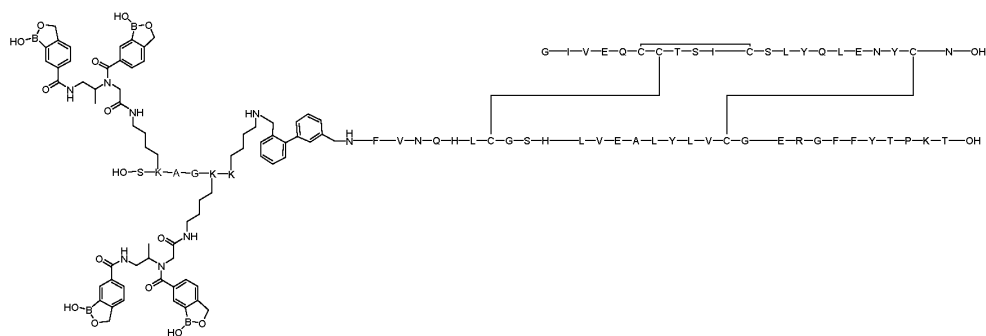
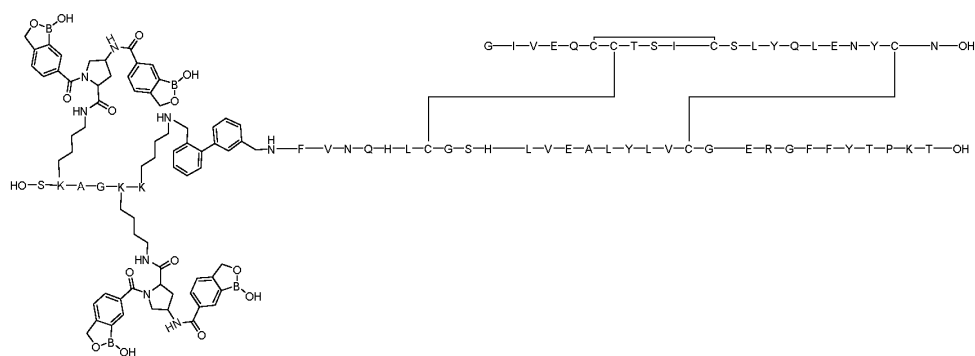


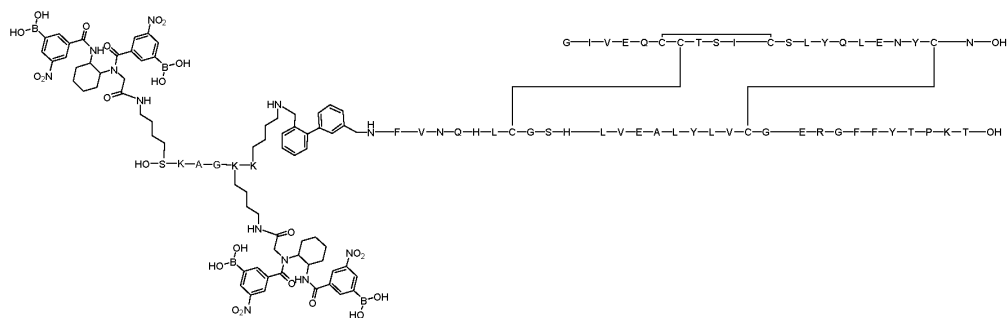
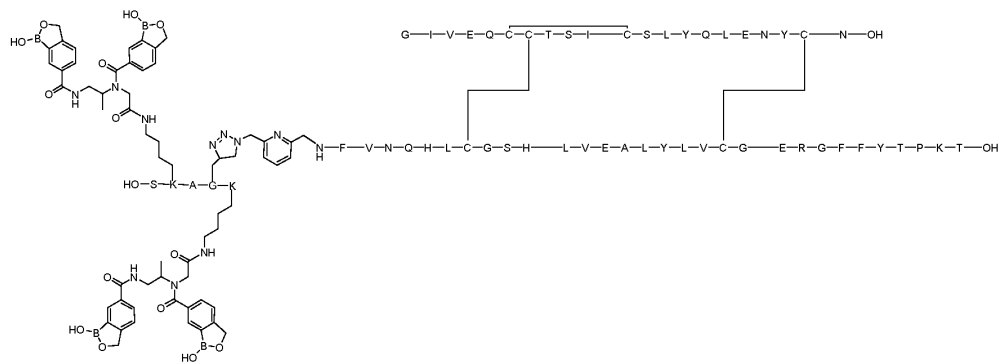
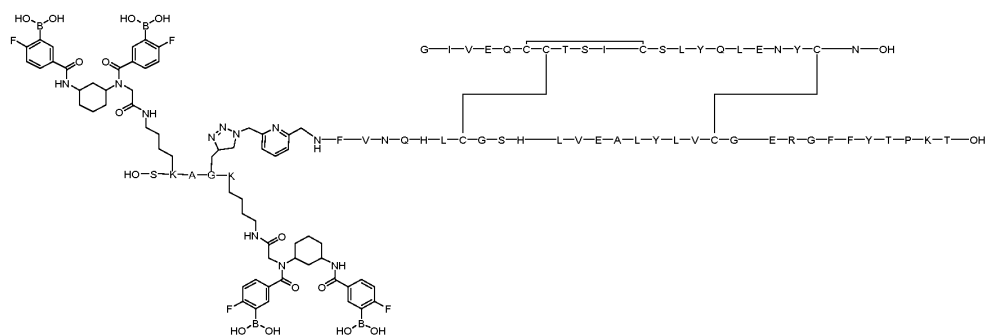
Example 448:



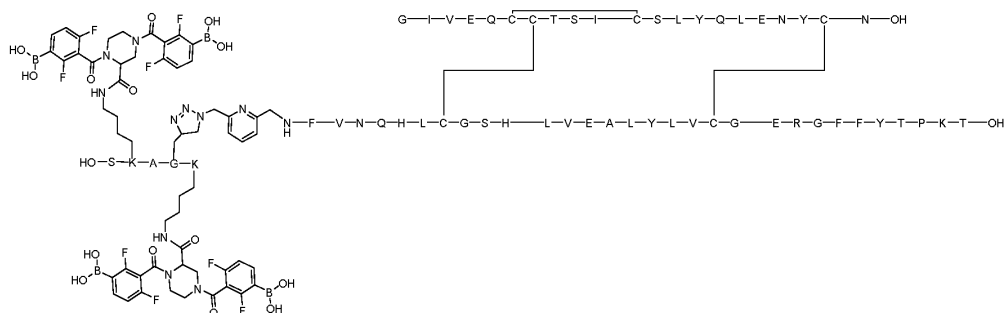
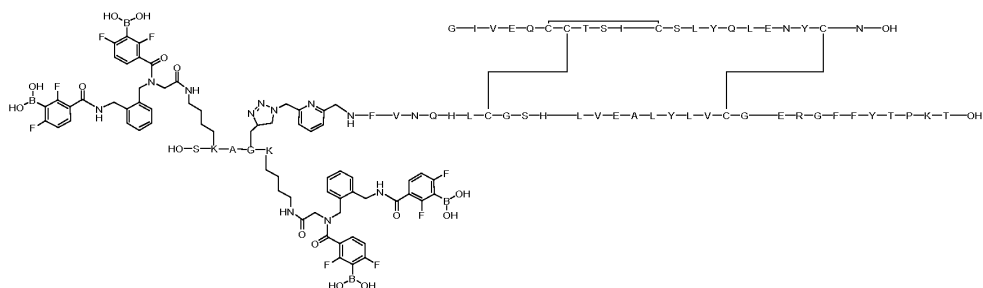
Example 449:**Example 450:****5 Example 451:**

Example 452:**Example 453:****5 Example 454:**

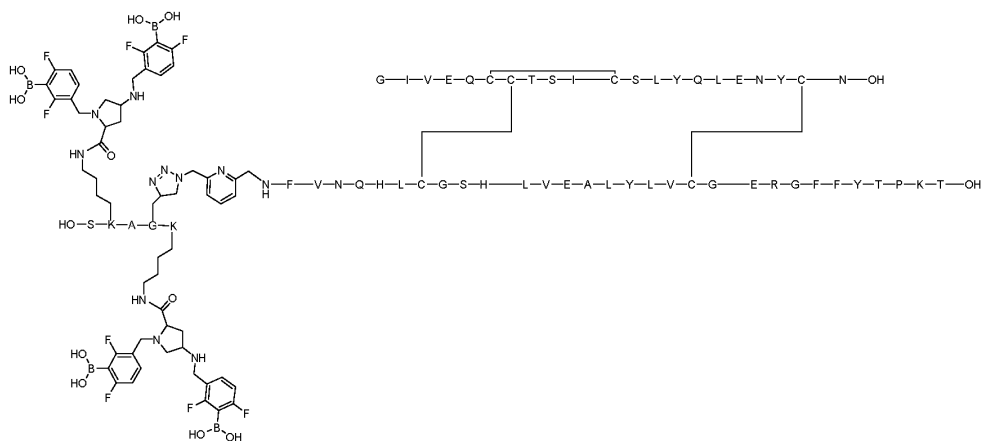
Example 455:**Example 456:****5 Example 457:**

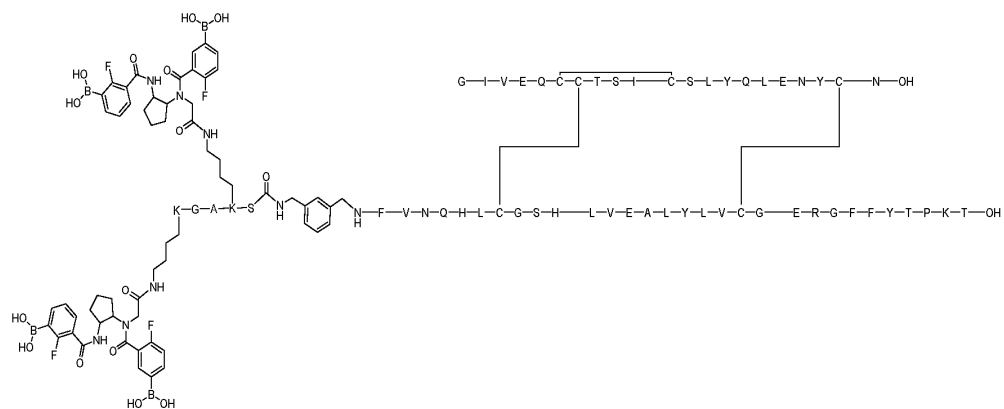
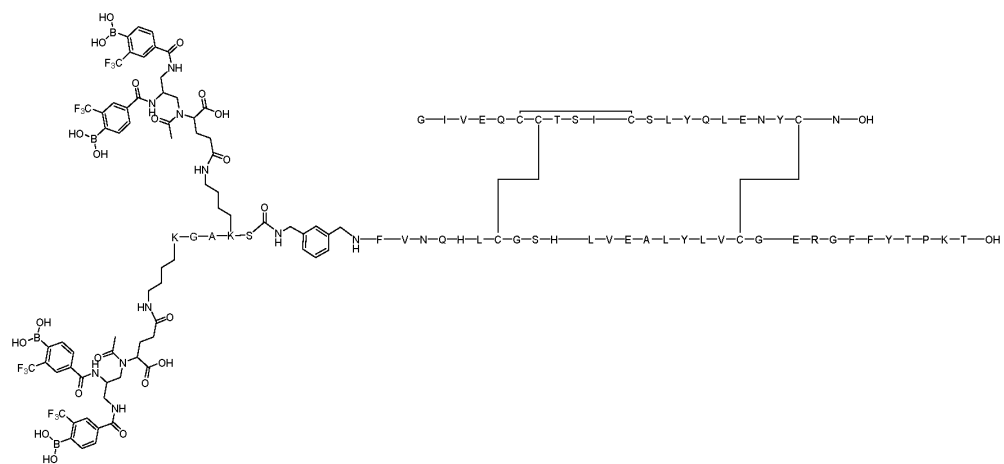
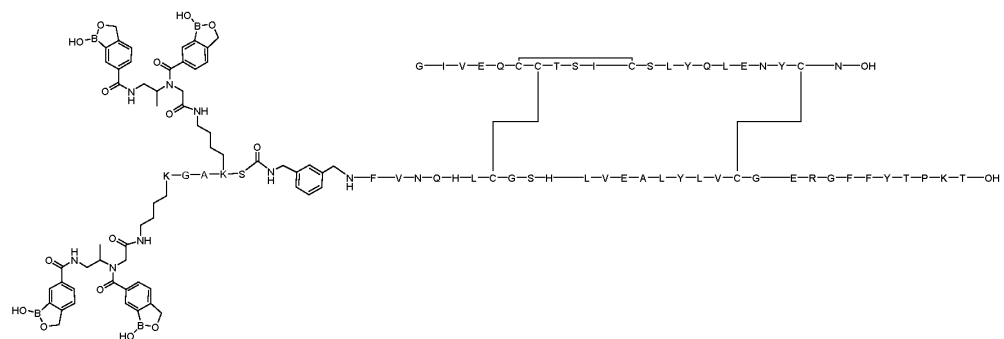
Example 458:**Example 459:****5 Example 460:**

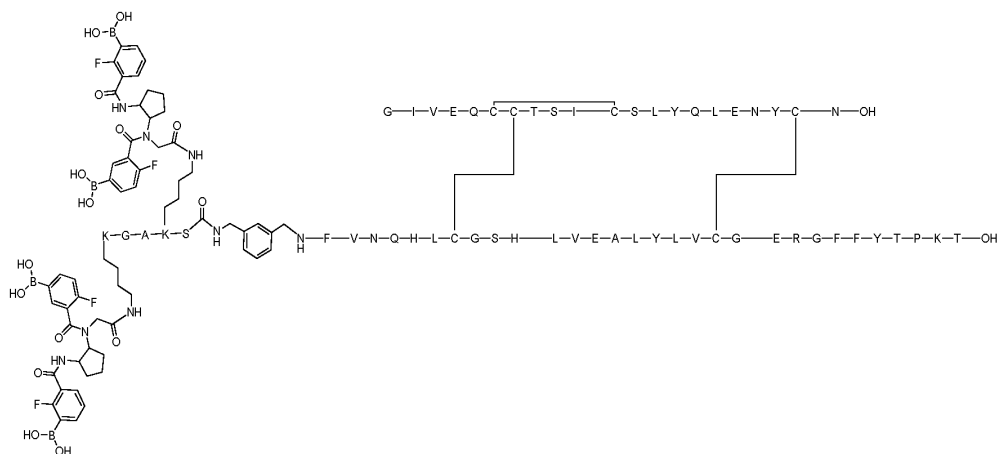
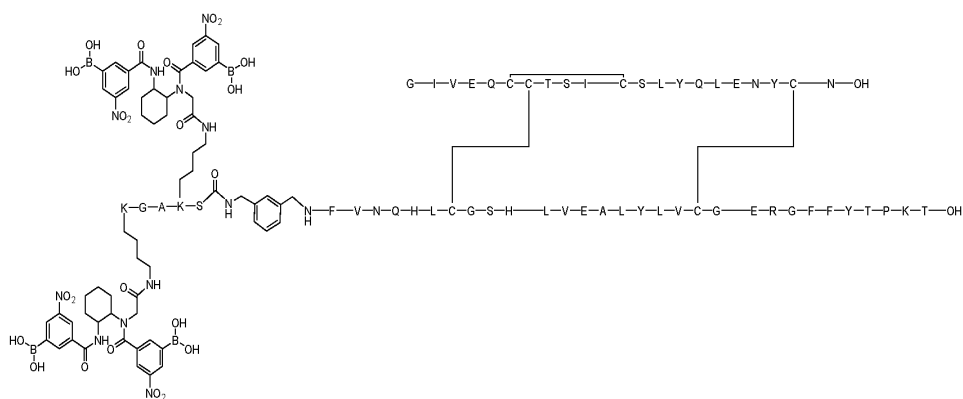
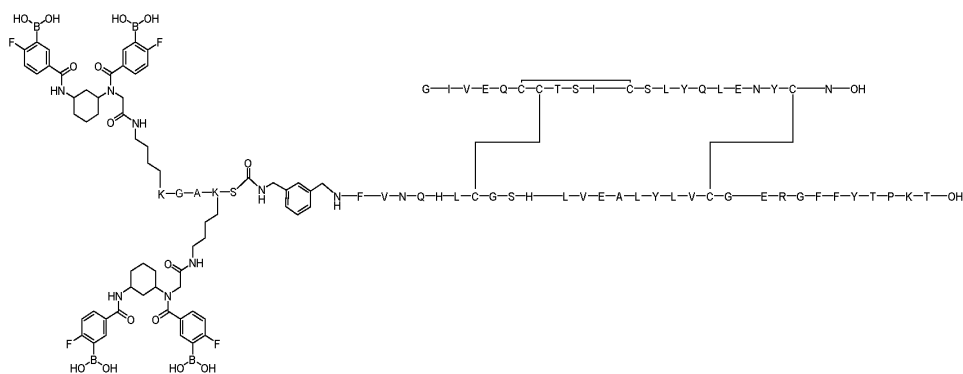
Example 461:

**Example 462:**

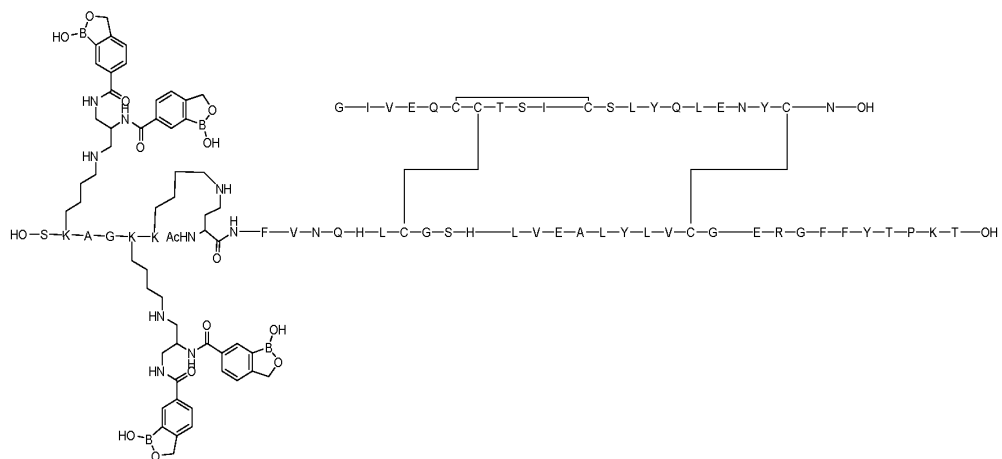
5 Example 463:



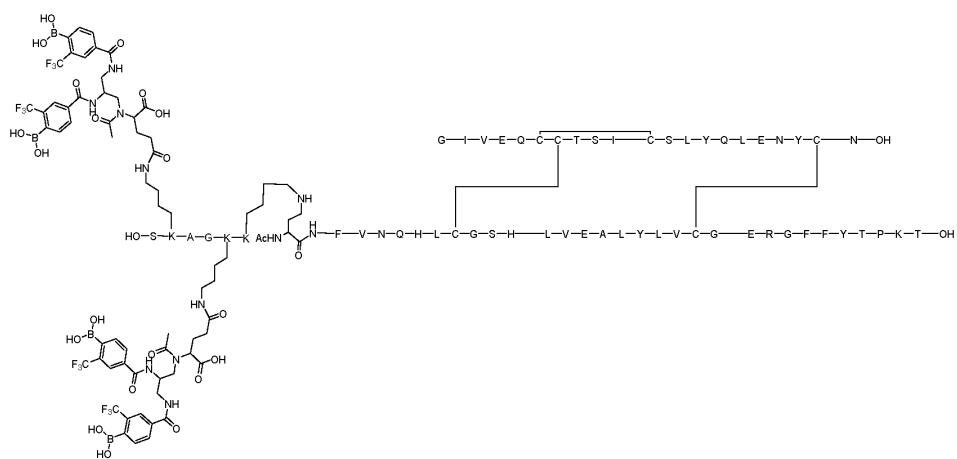
Example 464:**Example 465:****5 Example 466:**

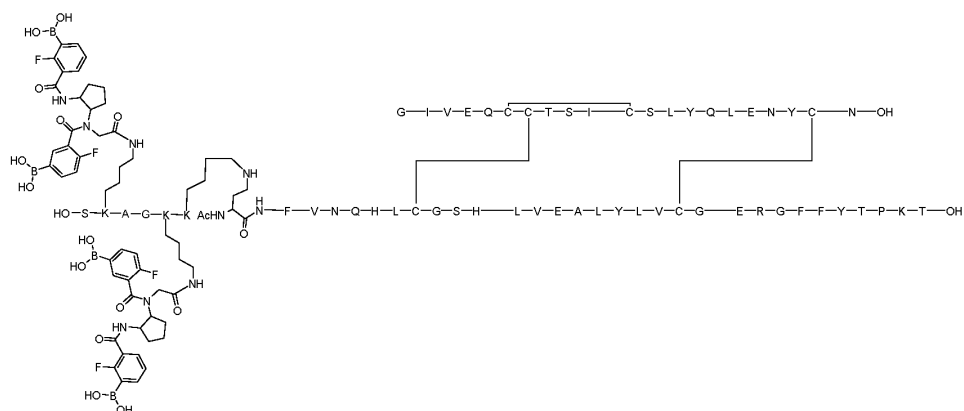
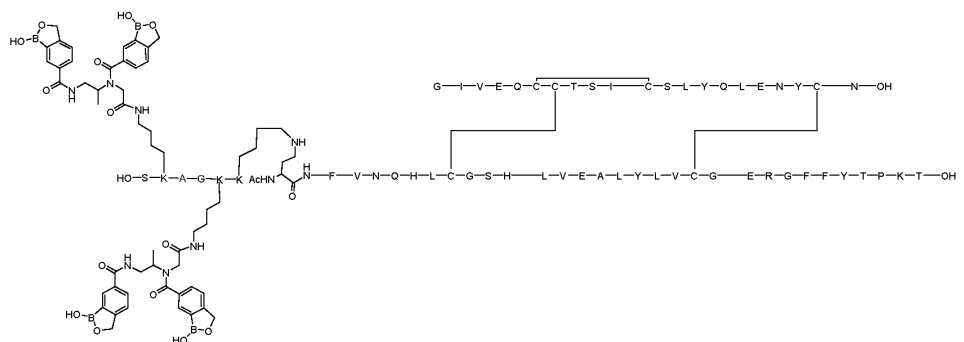
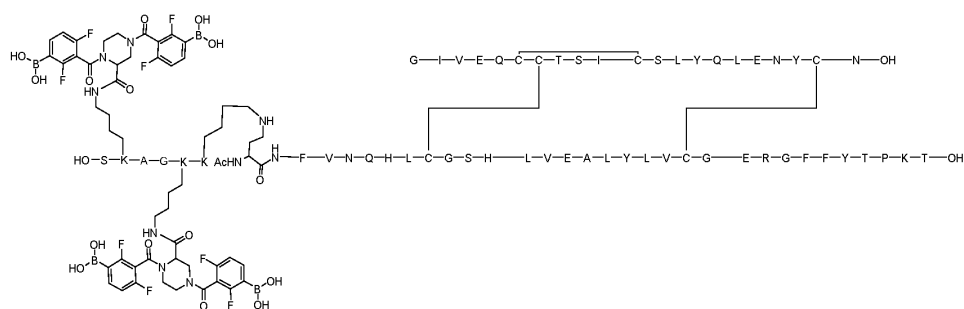
Example 467:**Example 468:****5 Example 469:**

Example 470:

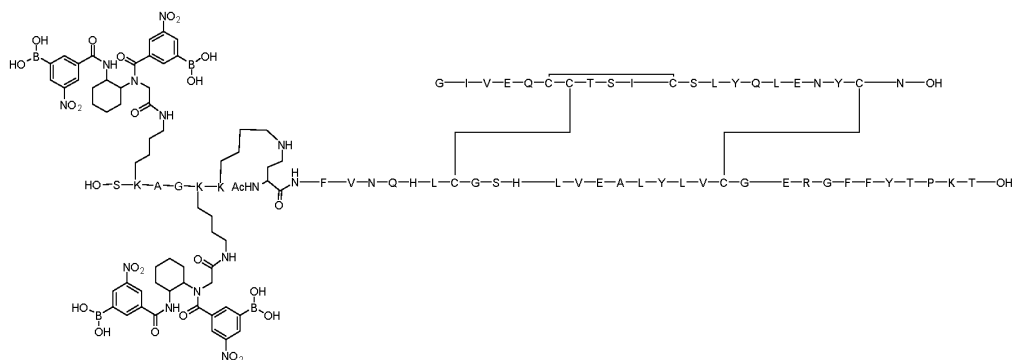


Example 471:

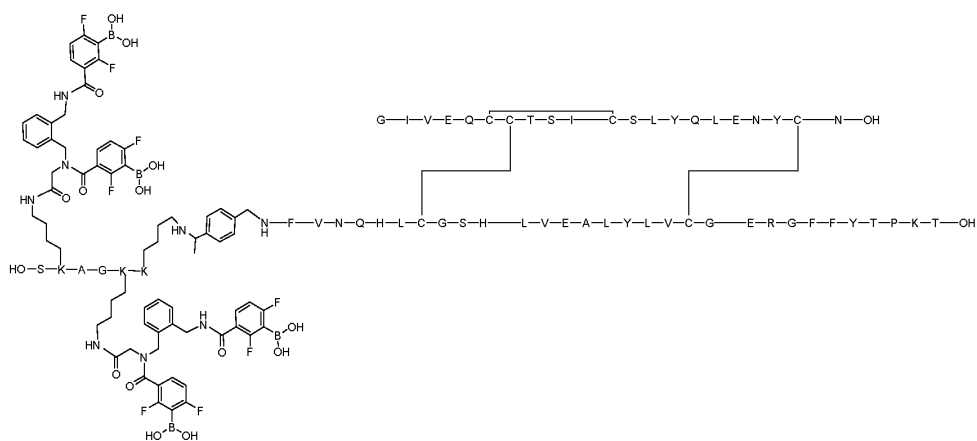


Example 472:**Example 473:****5 Example 474:**

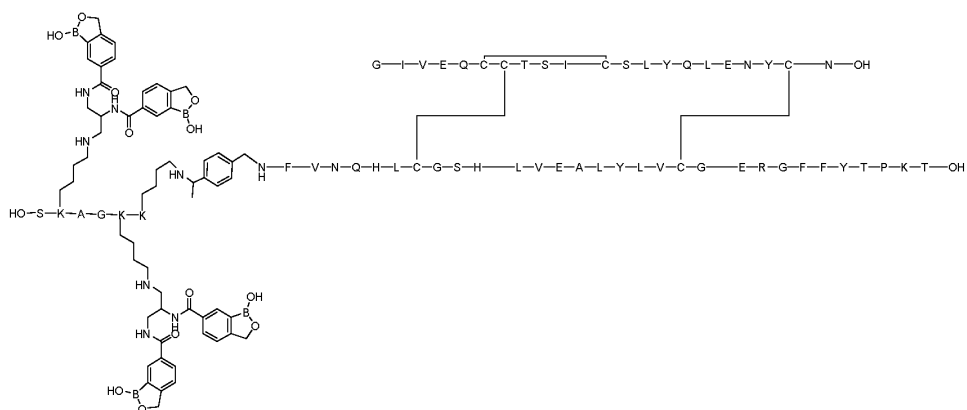
Example 475:



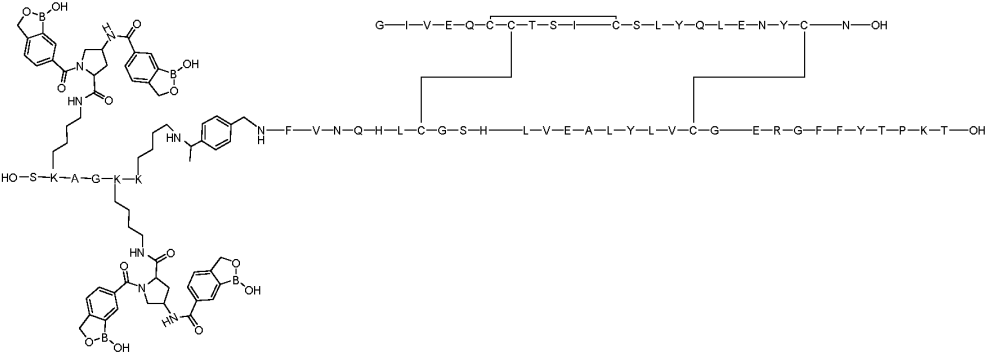
Example 476:



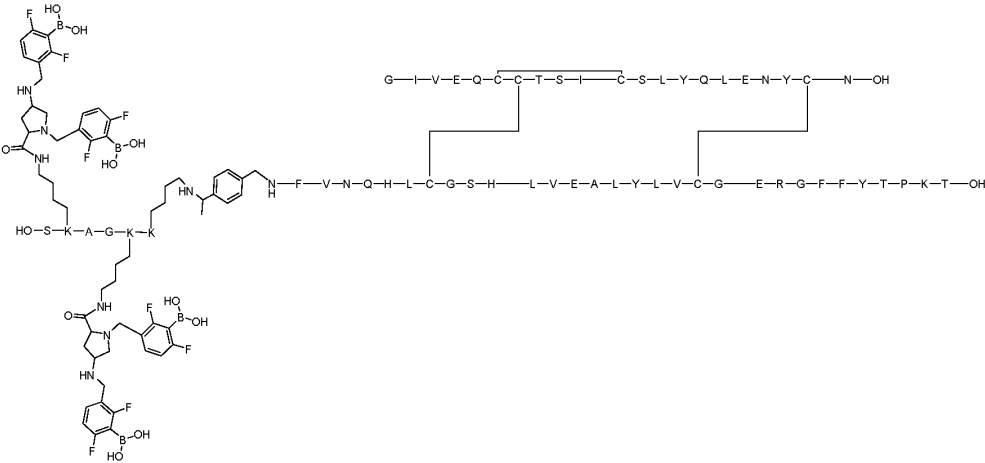
5 Example 477:



Example 478:

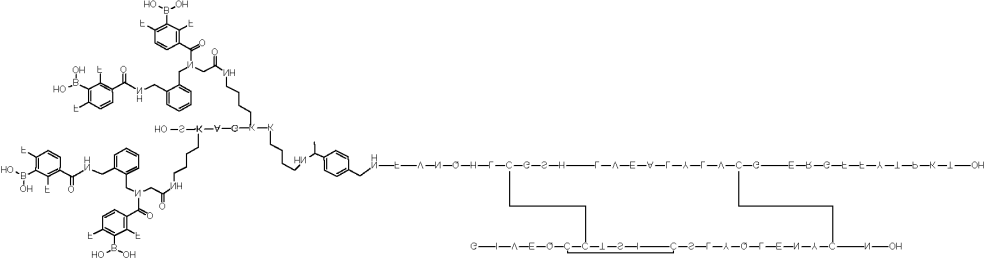


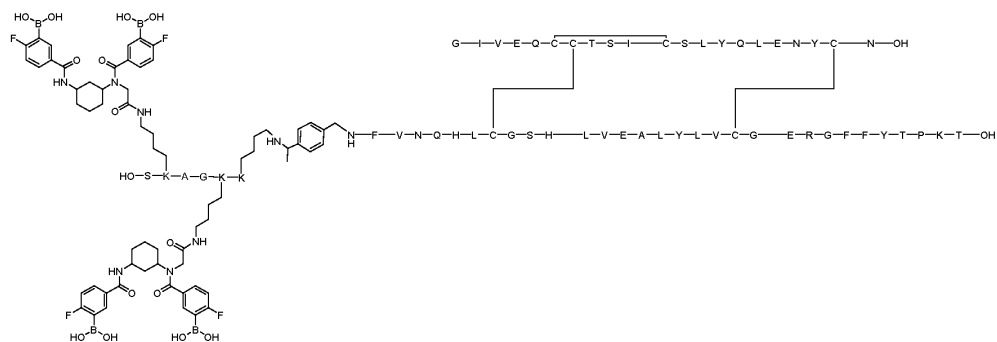
Example 479:

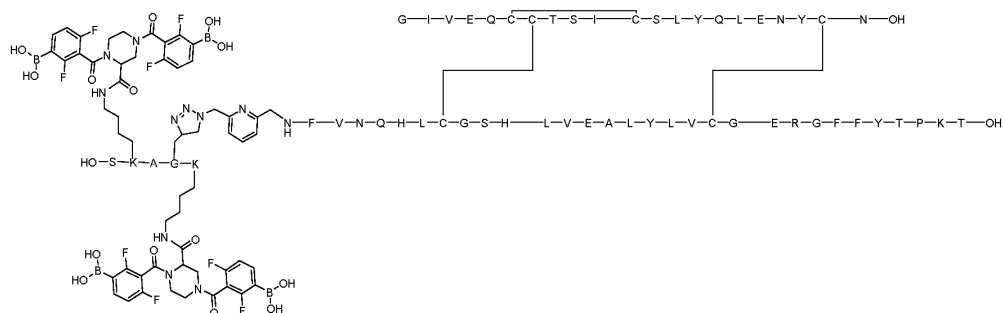


5

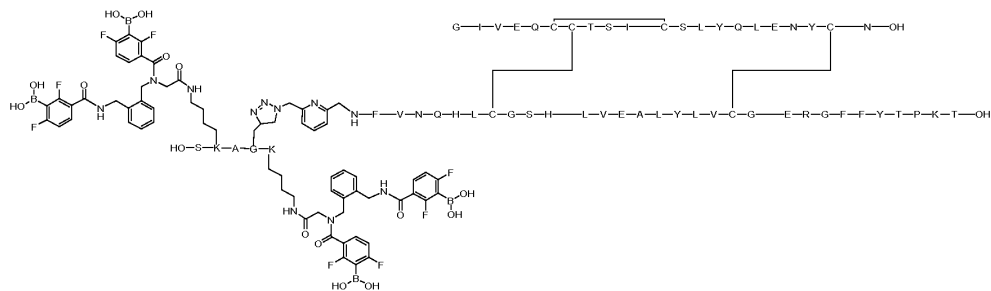
Example 480:



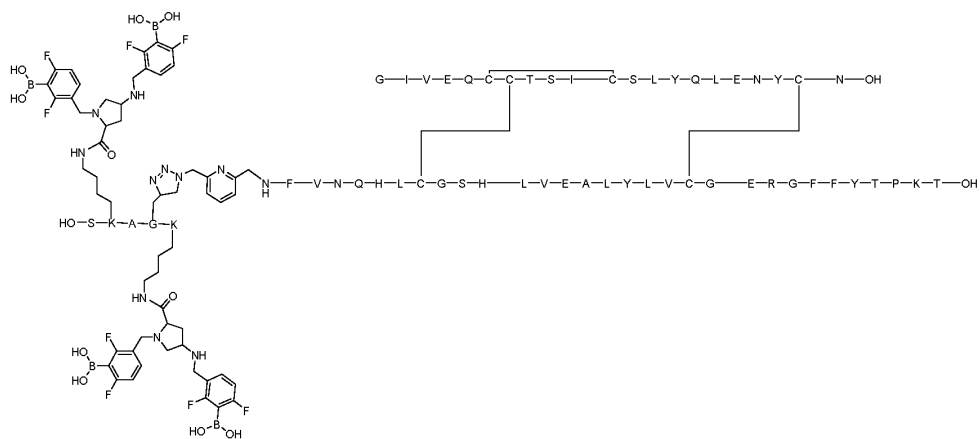
[illegible]

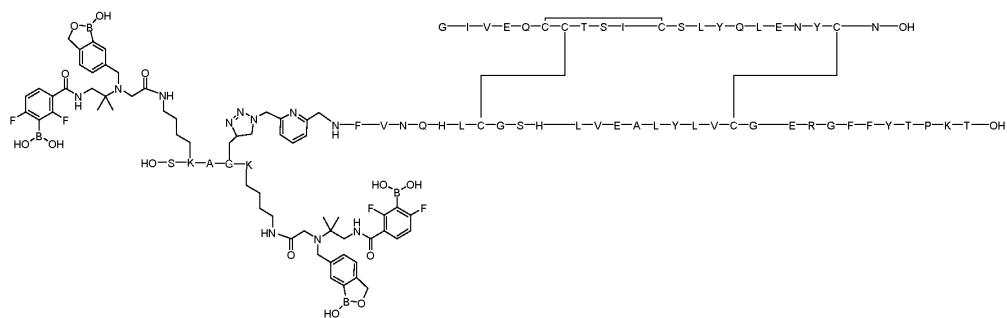
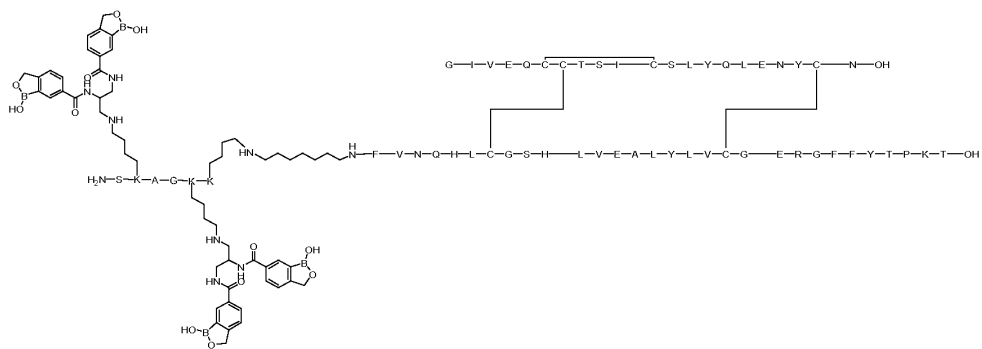
Example 484:

Example 485:



5 Example 486:



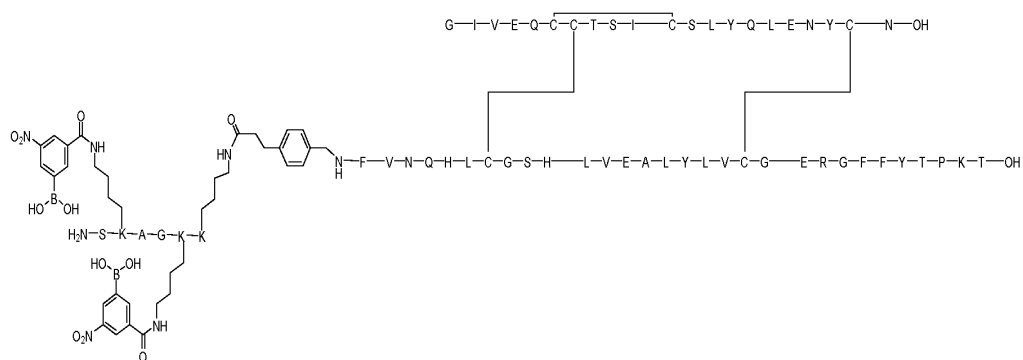
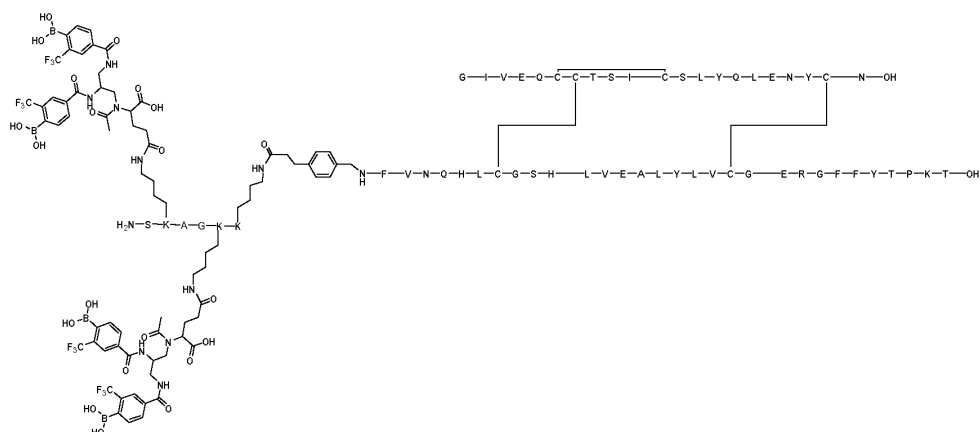
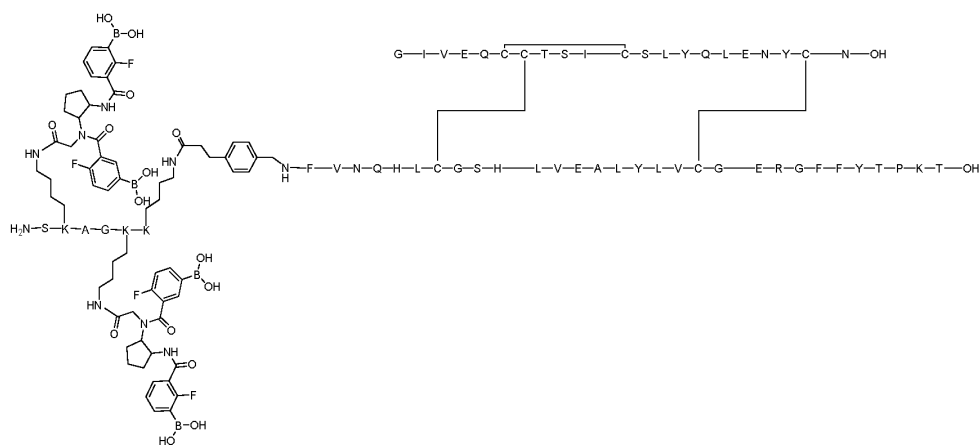
Example 487:**Example 488:**

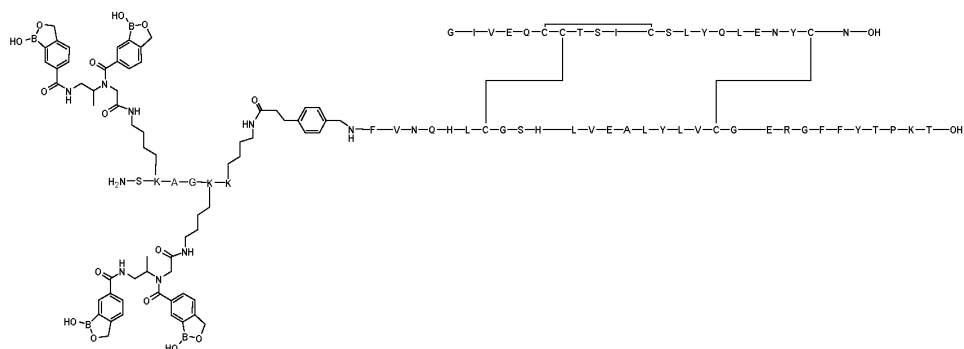
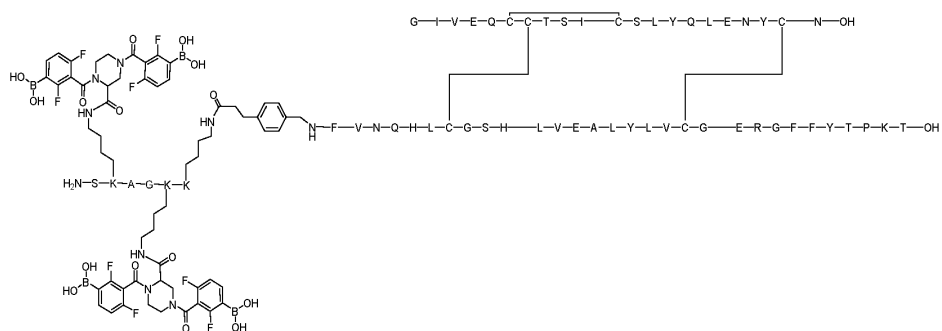
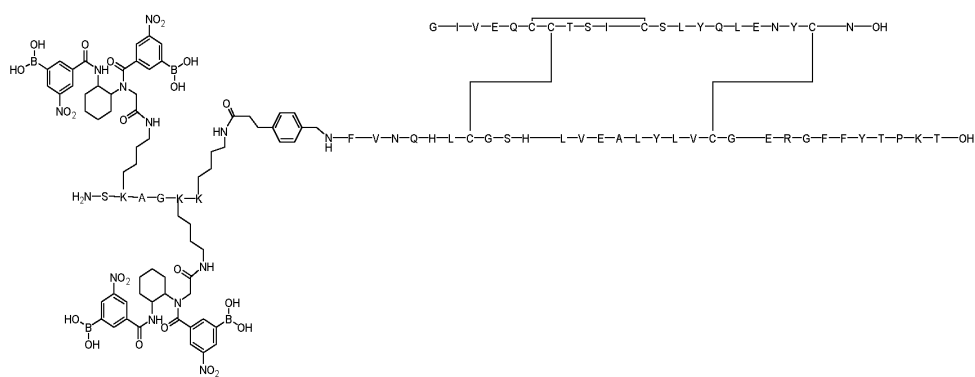
5

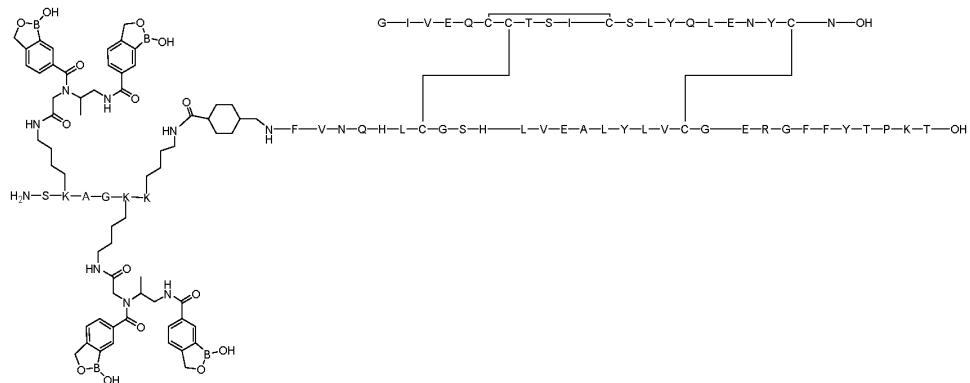
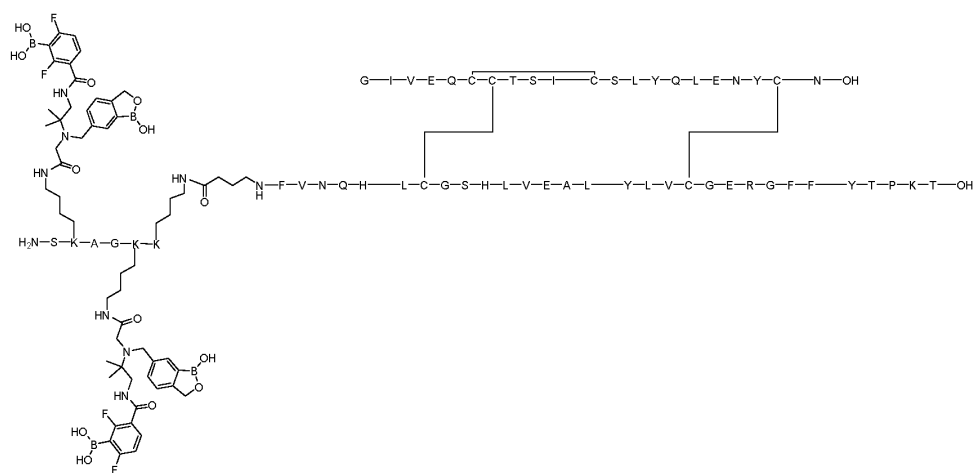
10

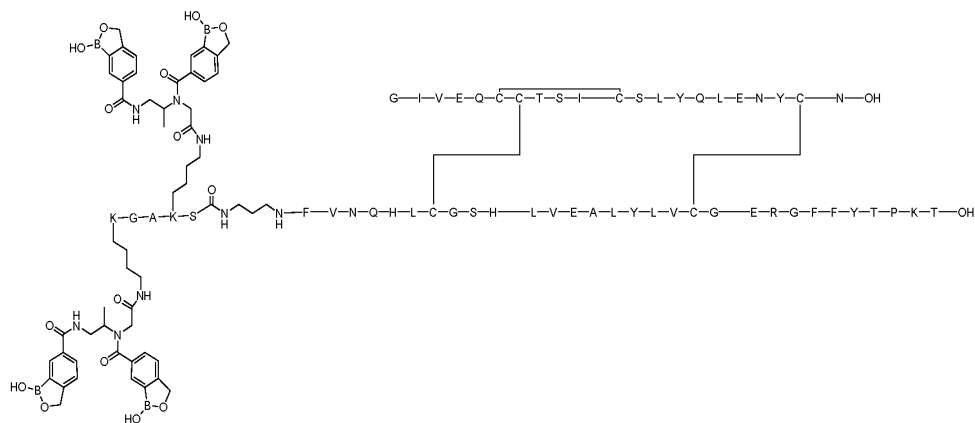
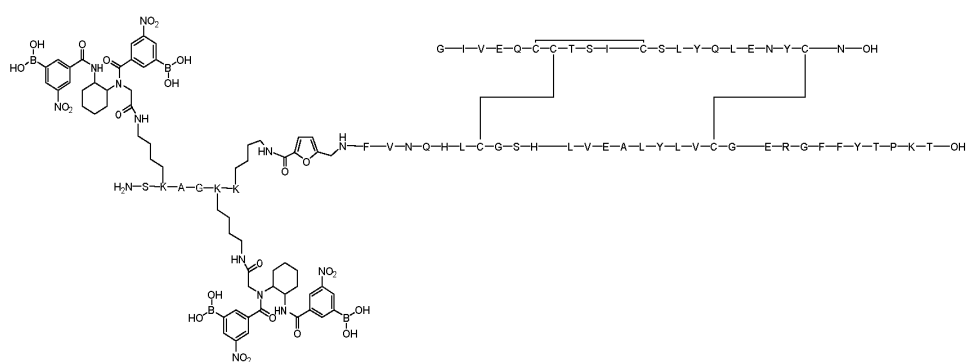
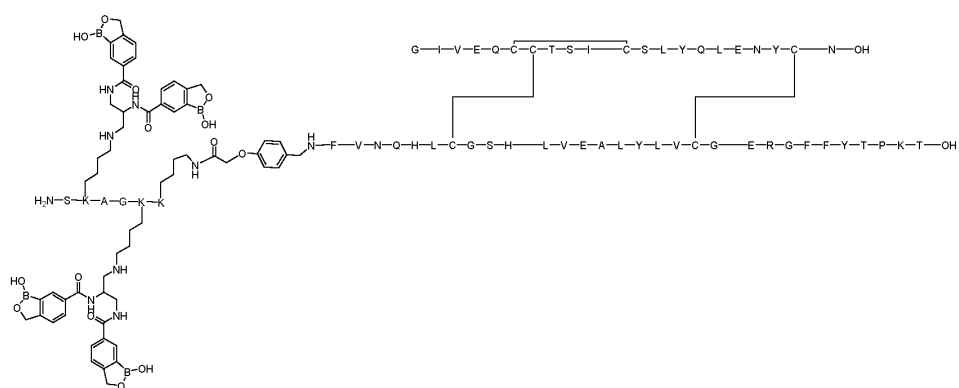
15

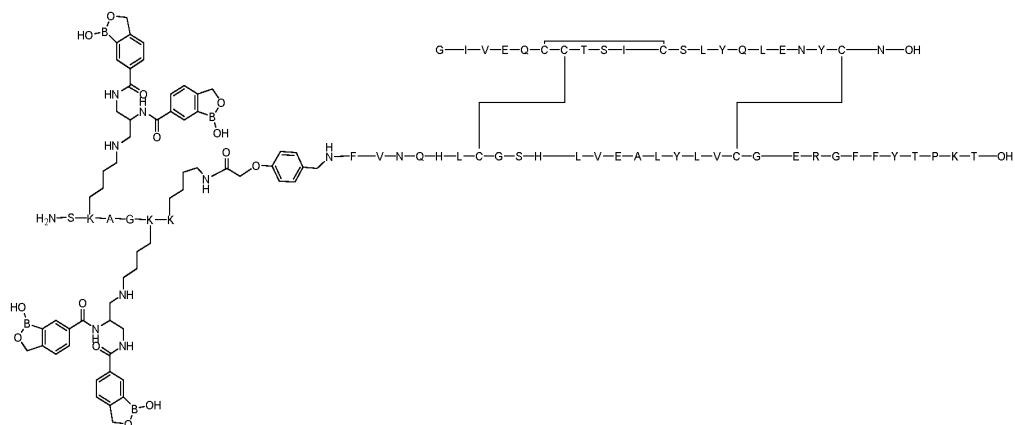
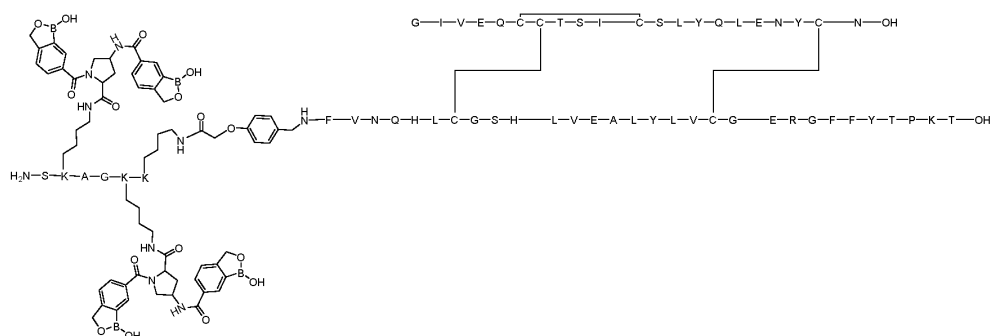
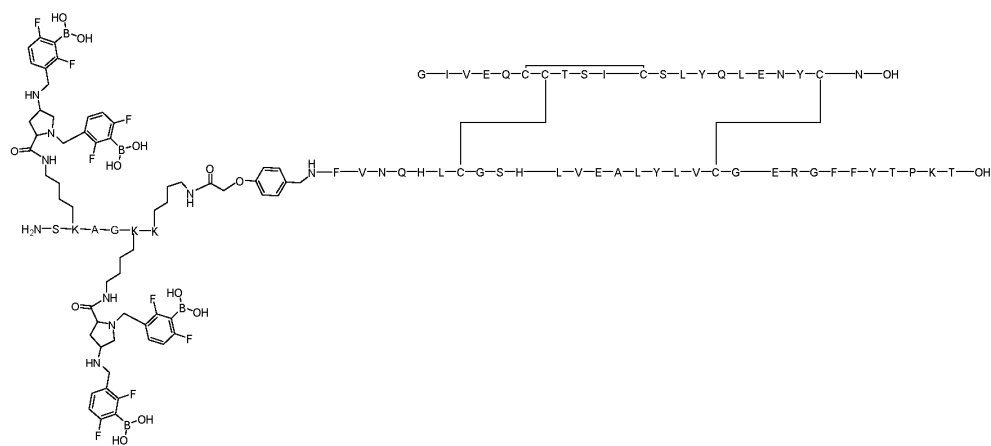
20

Example 489:**Example 490:****5 Example 491:**

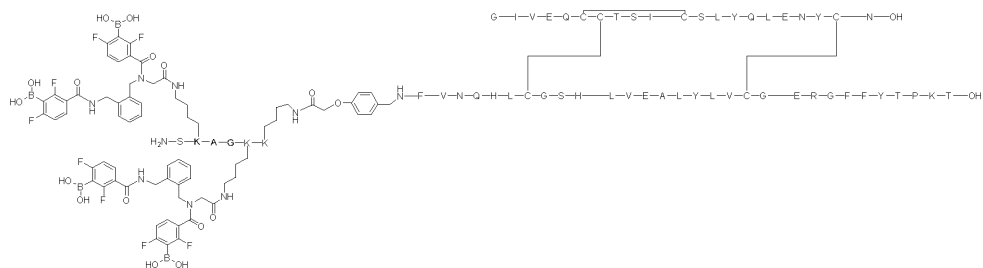
Example 492:**Example 493:****5 Example 494:**

Example 495:**Example 496:**

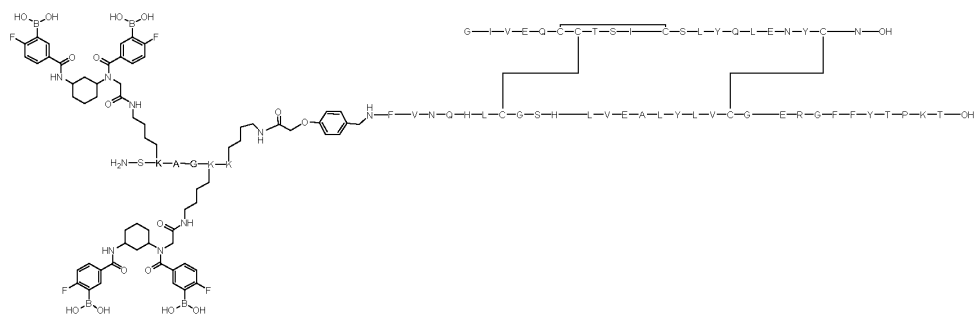
Example 497:**Example 498:****5 Example 499:**

Example 500:**Example 501:****5 Example 502:**

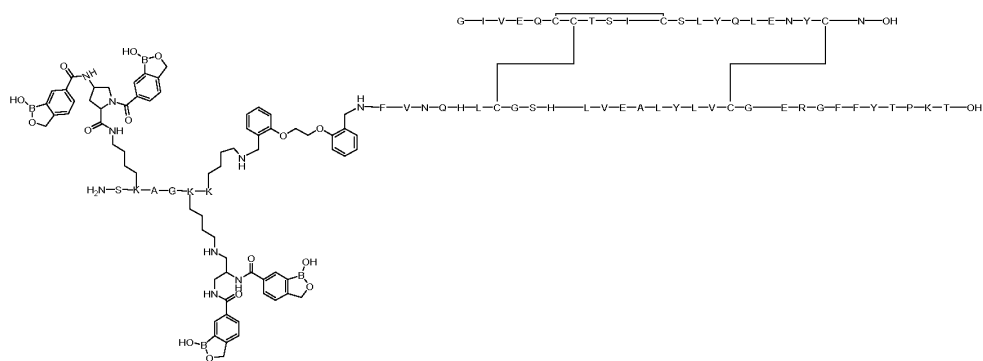
Example 503:

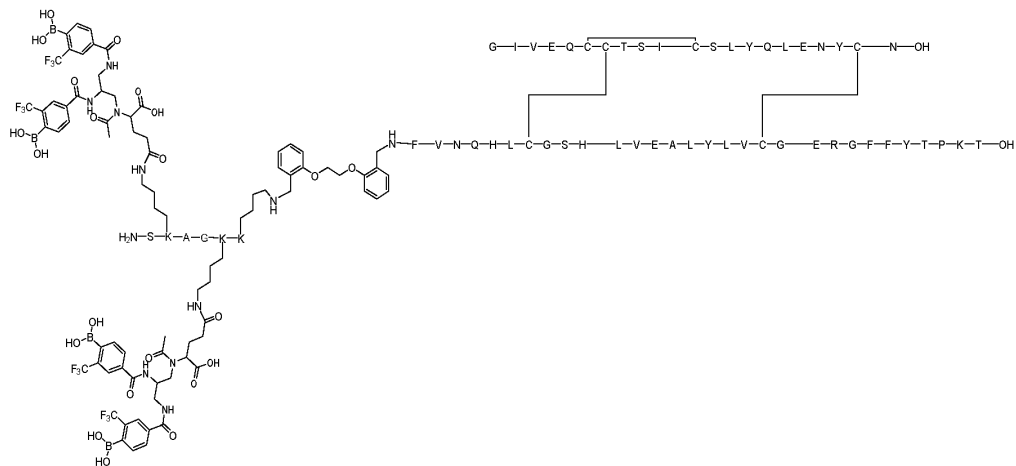
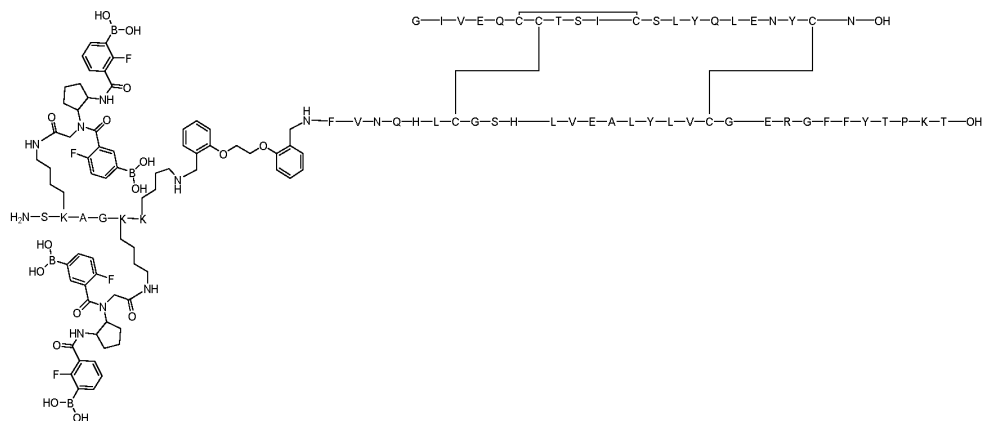


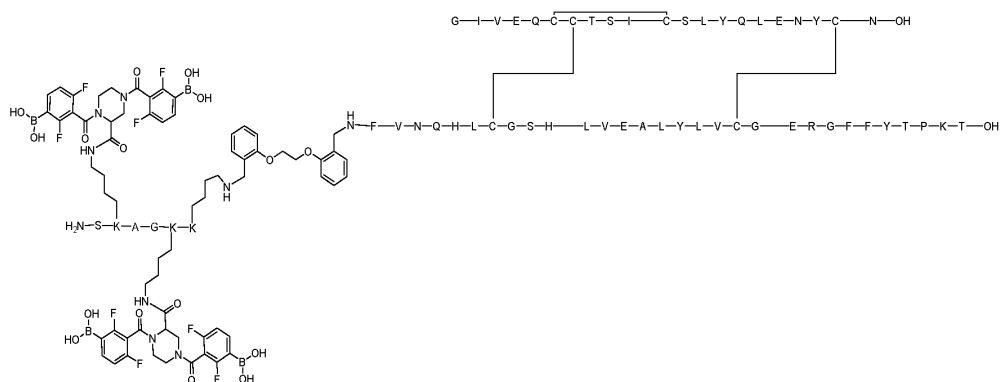
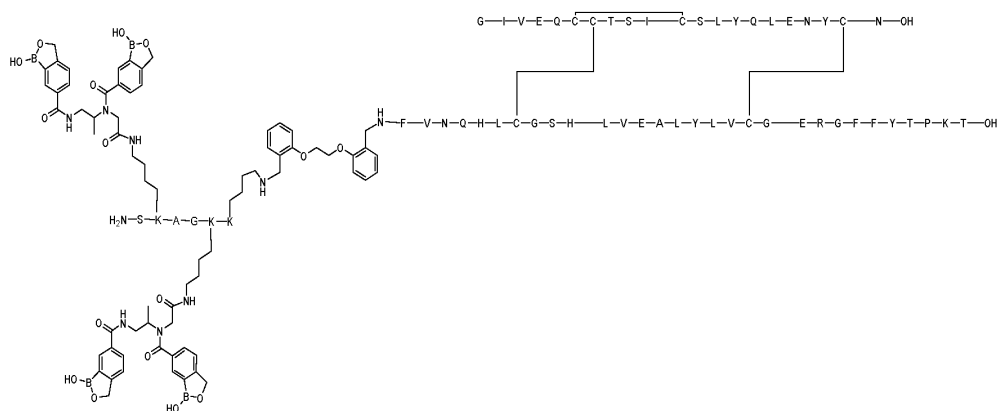
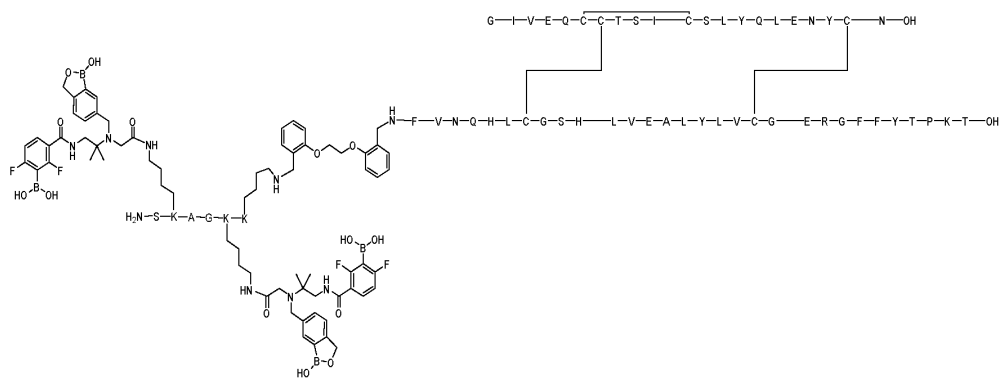
Example 504:



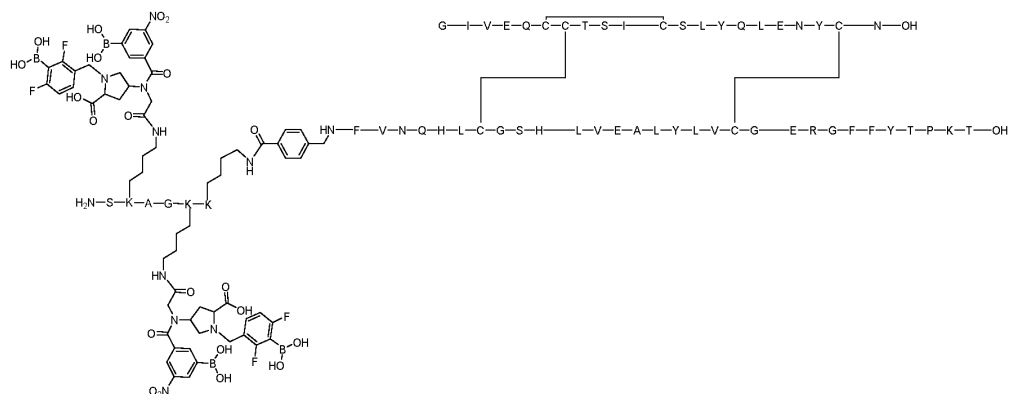
5 Example 505:



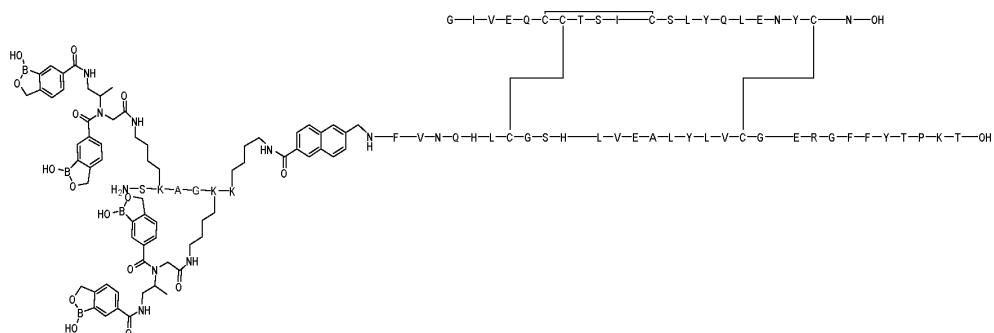
Example 506:**Example 507:**

Example 508:**Example 509:****5 Example 510:**

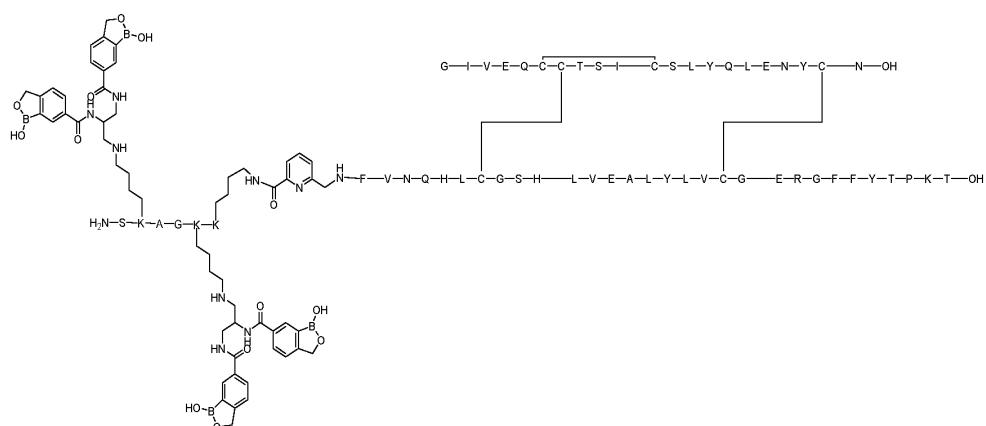
Example 511:

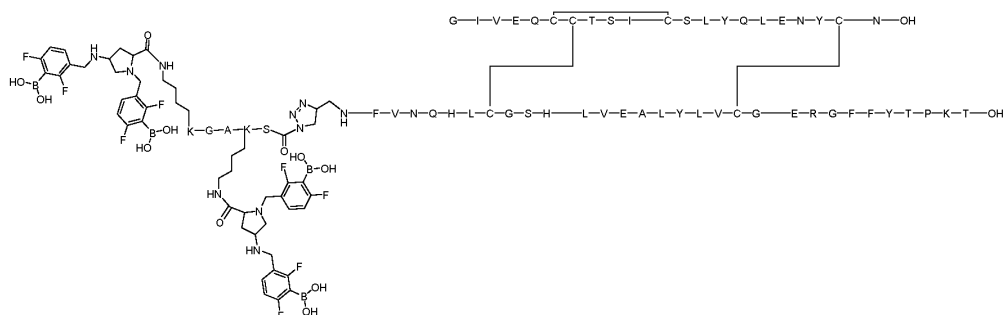
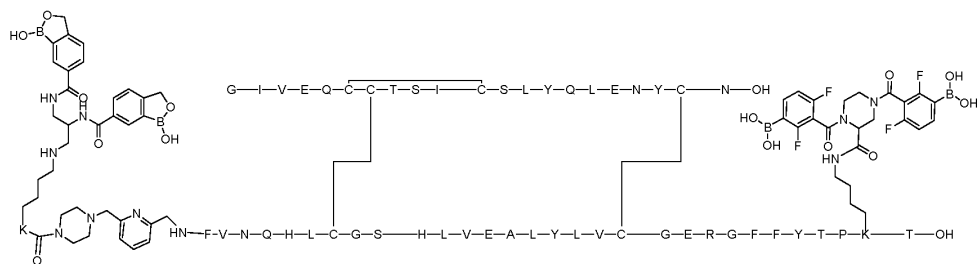
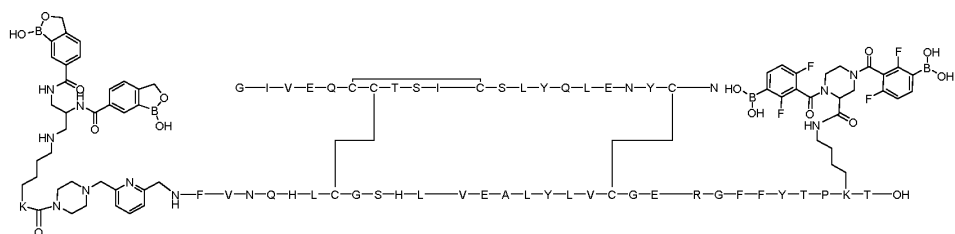
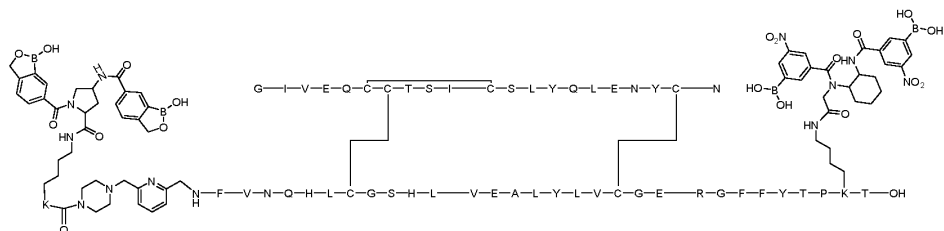


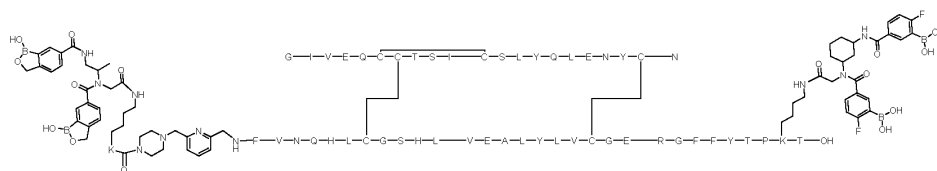
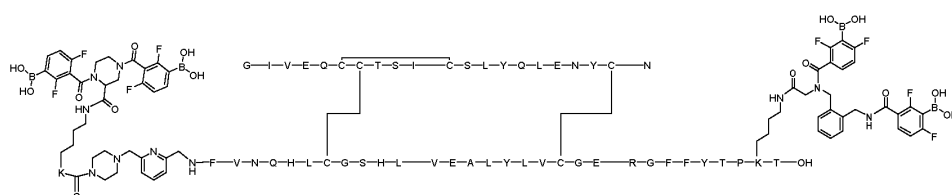
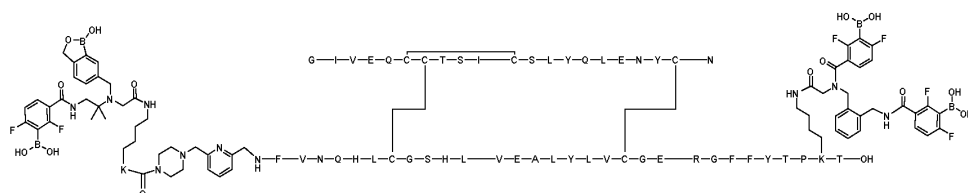
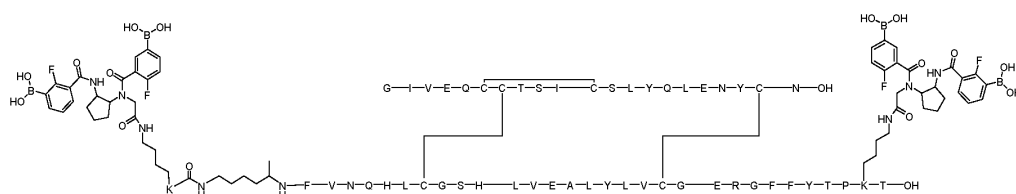
Example 512:



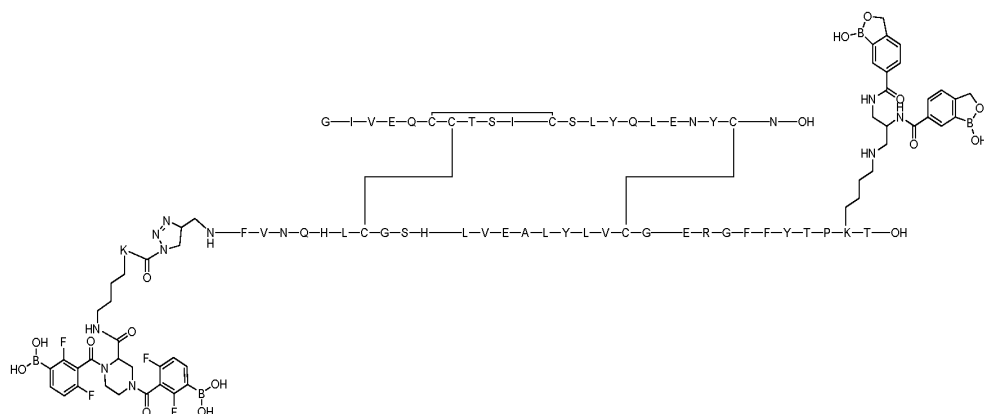
5 Example 513:



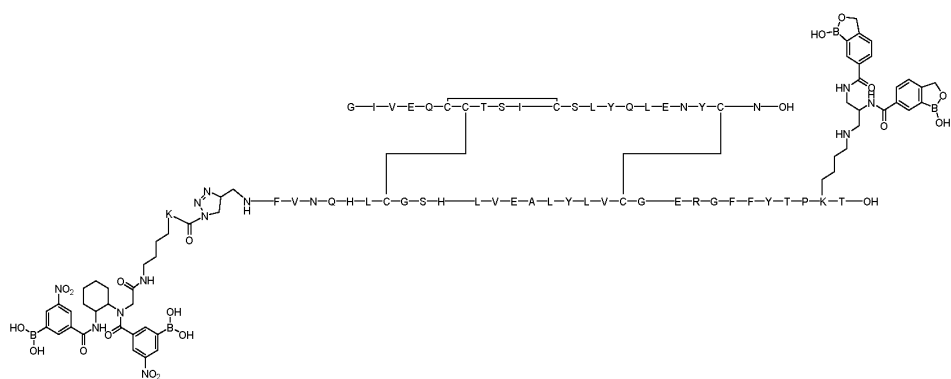
Example 514:**Example 515:****5 Example 516:****Example 517:**

Example 518:**Example 519:****5 Example 520:****Example 521:**

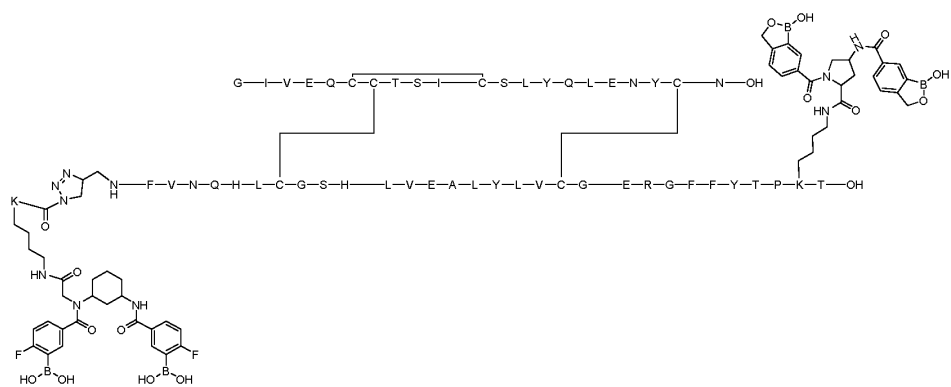
Example 522:

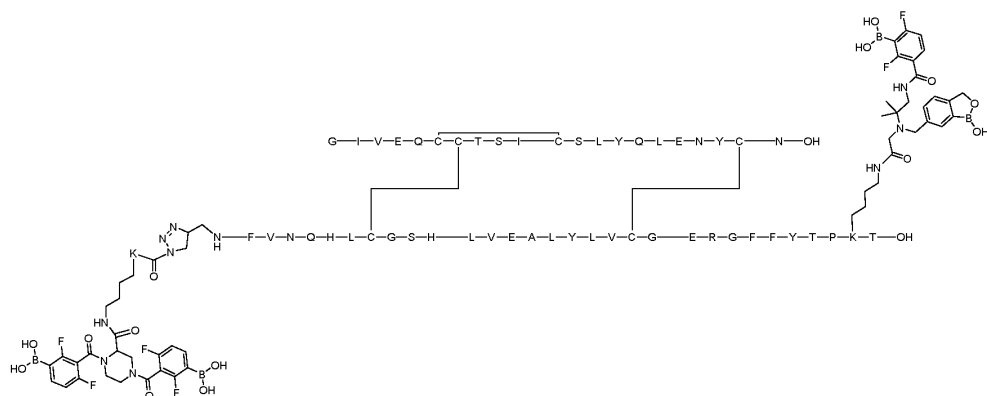
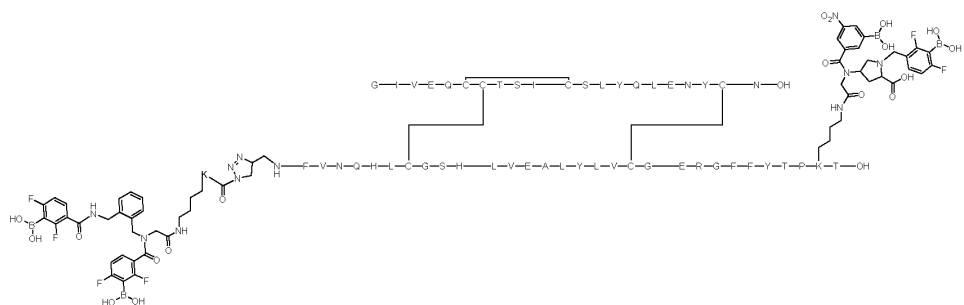
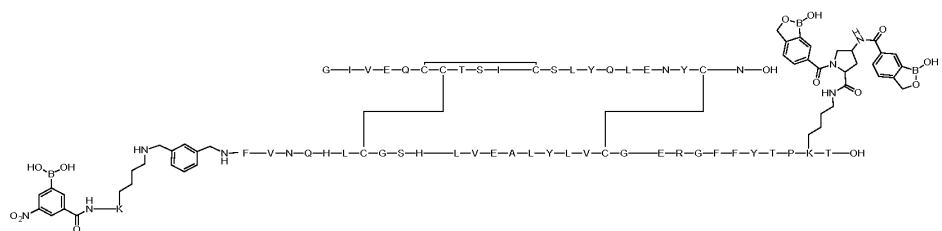


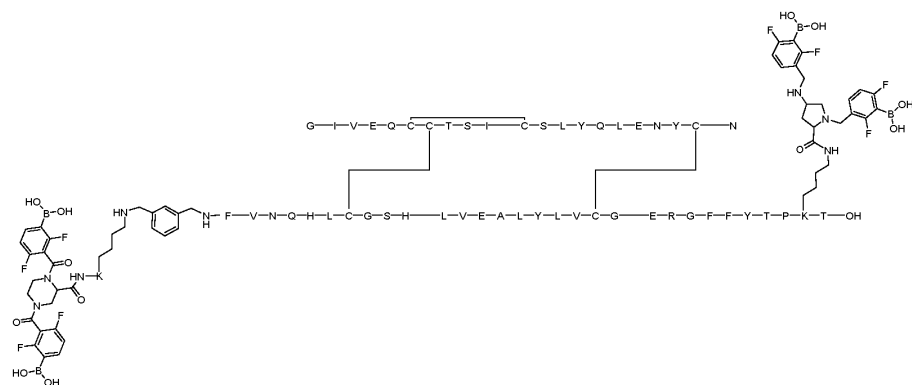
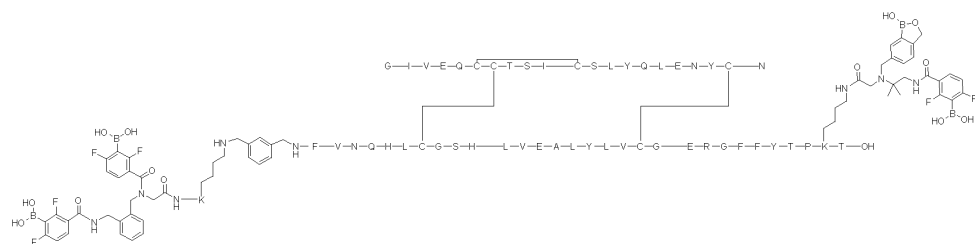
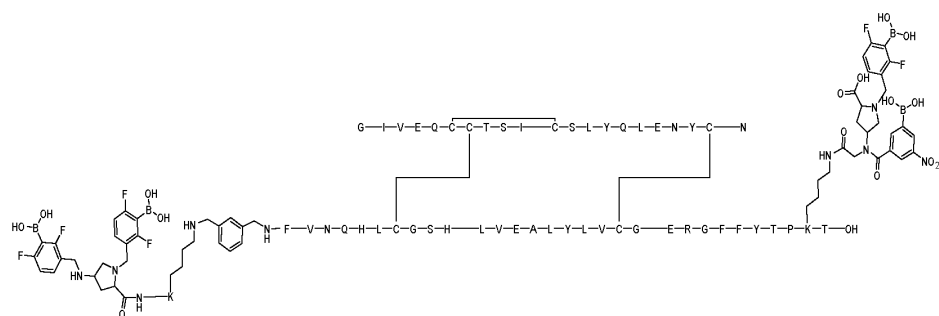
Example 523:

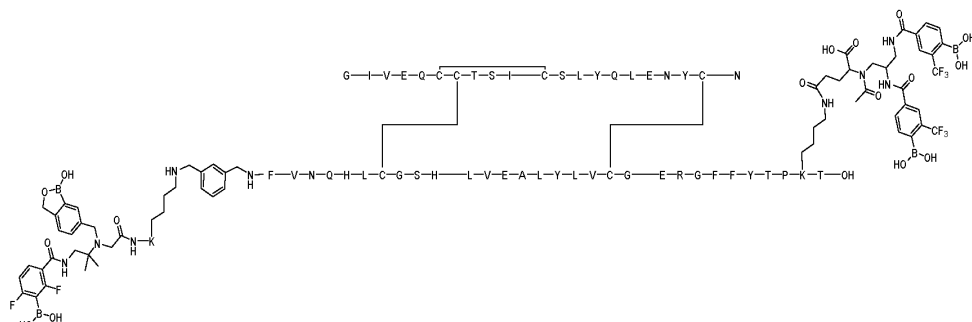
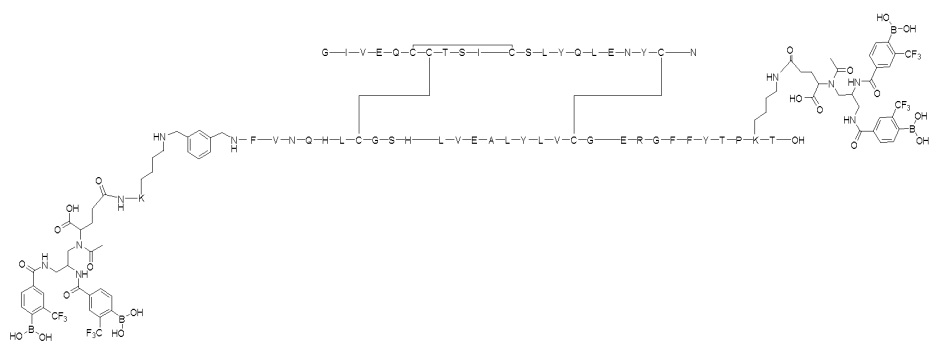
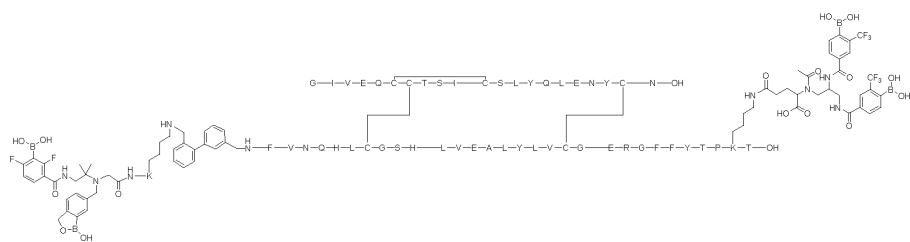


5 Example 524:

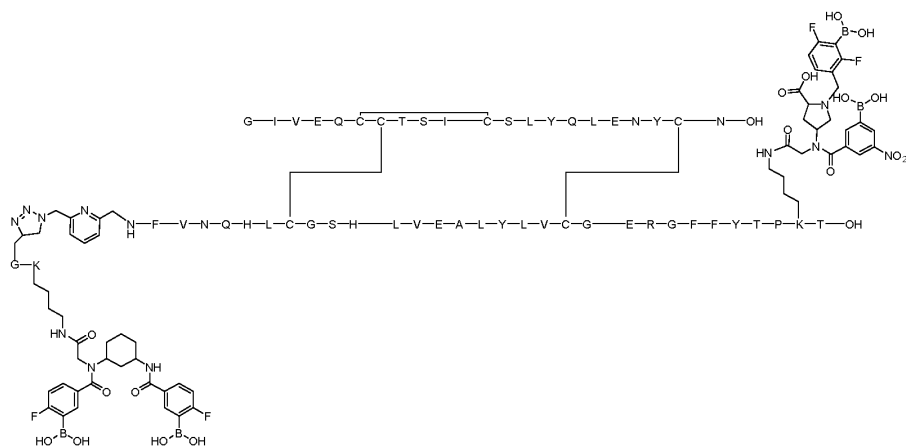
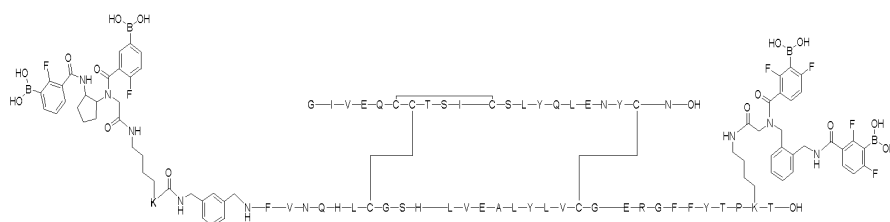
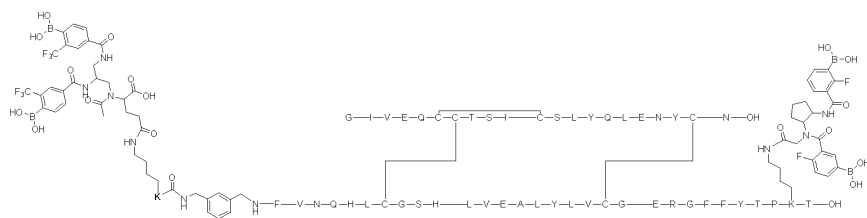
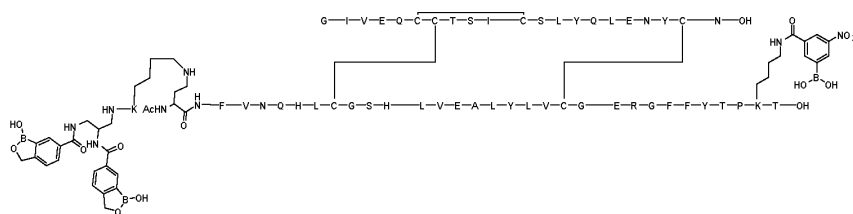


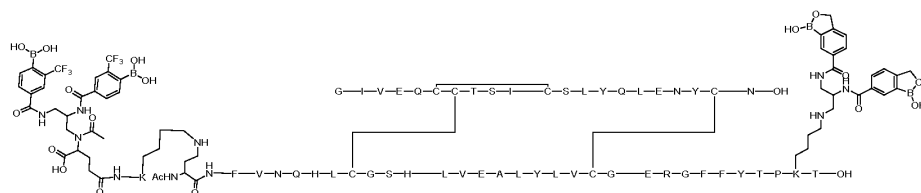
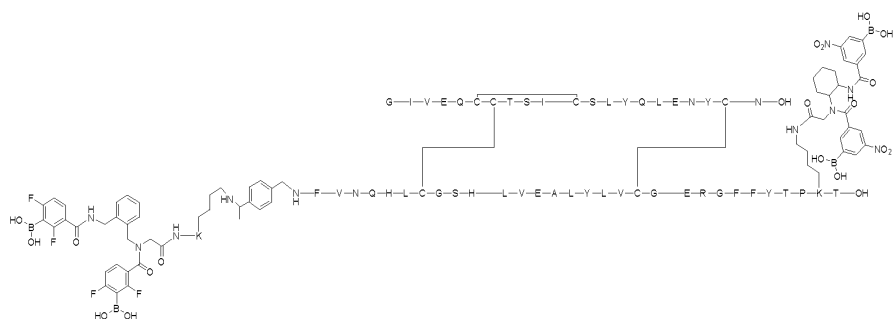
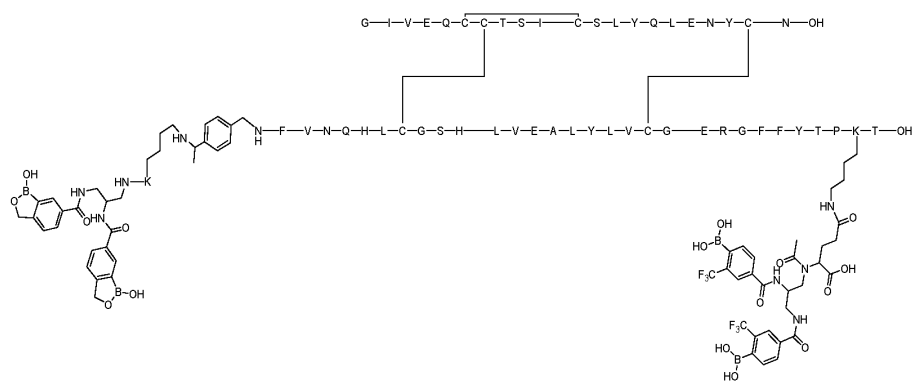
Example 525:**Example 526:****5 Example 527:**

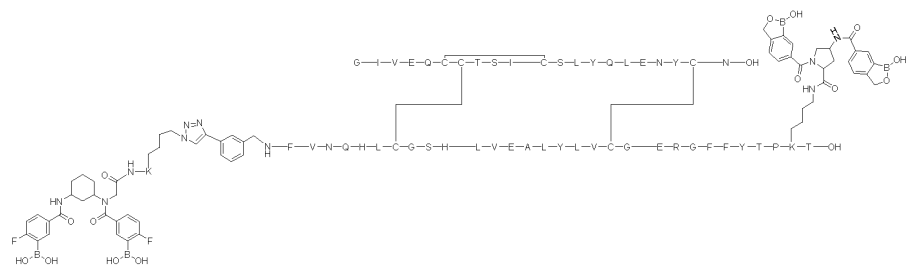
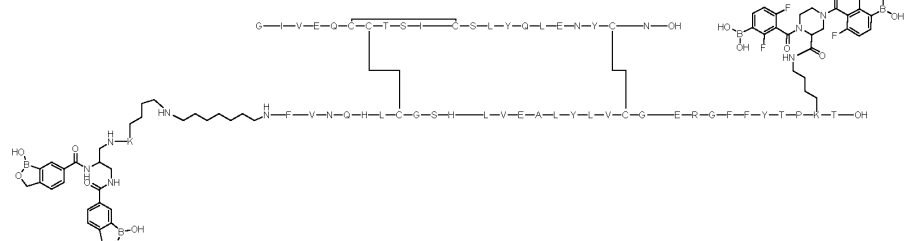
Example 528:**Example 529:****5 Example 530:**

Example 531:**Example 532:****5 Example 533:**

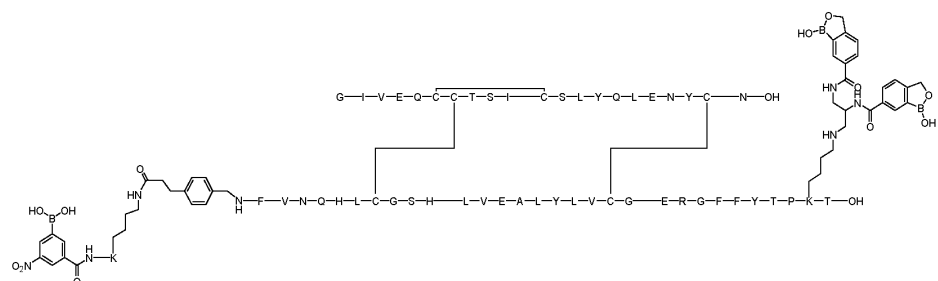
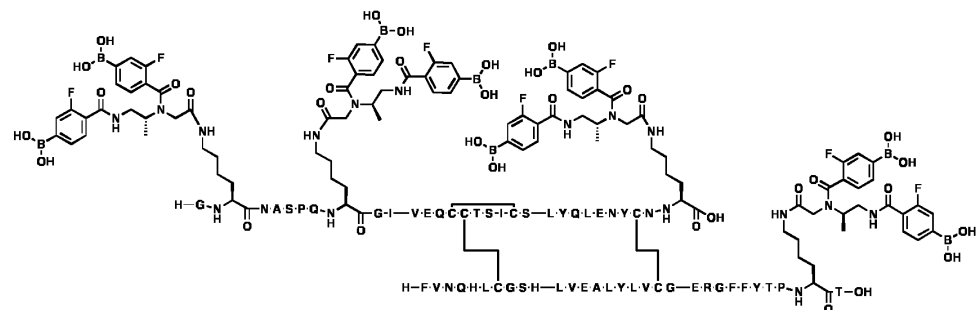


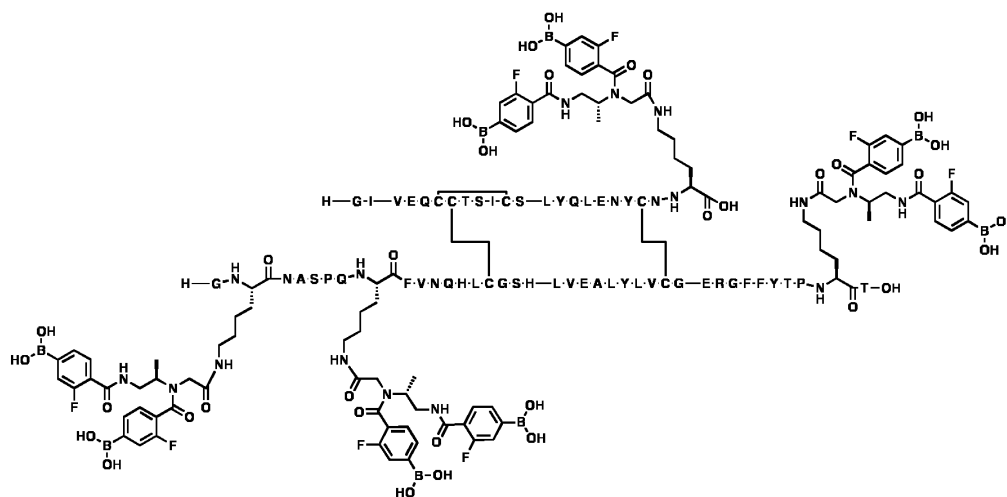
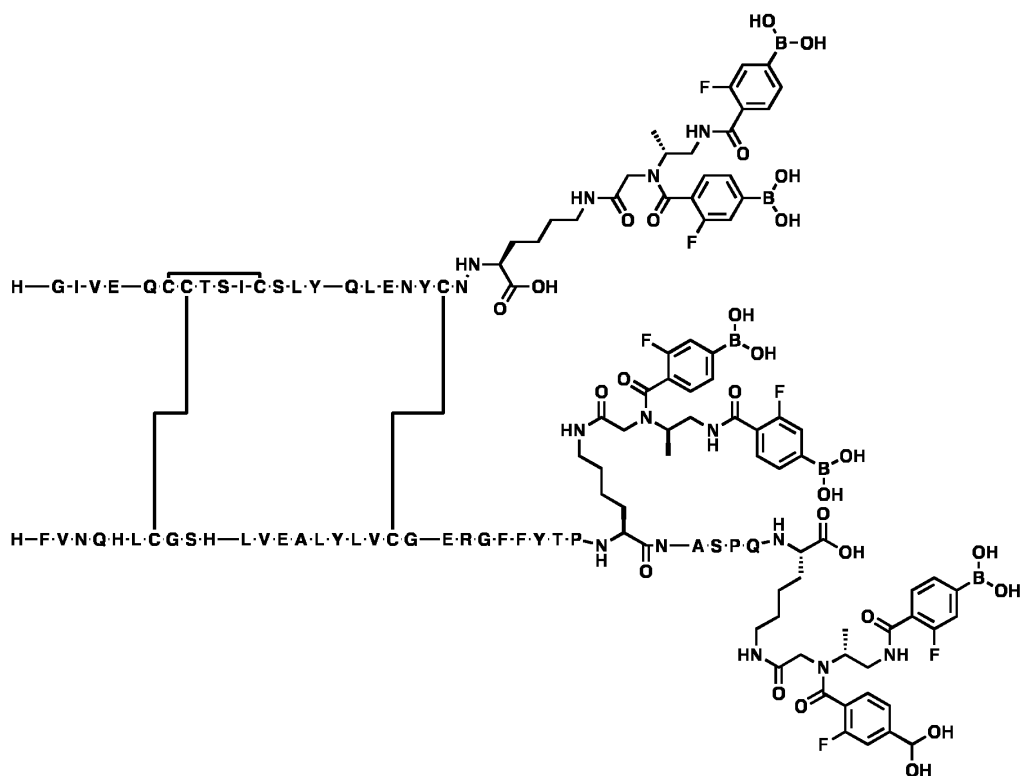
Example 537:**Example 538:****5 Example 539:****Example 540:**

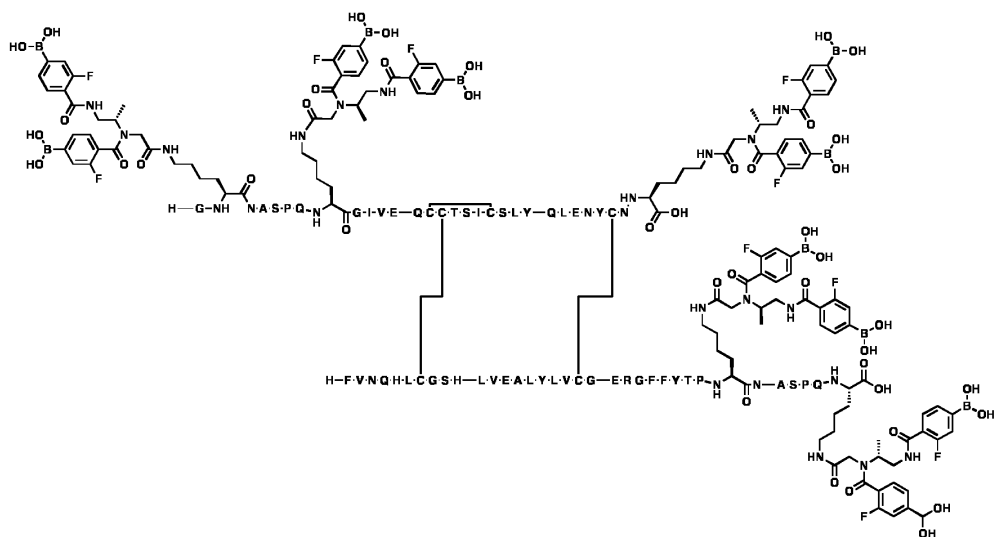
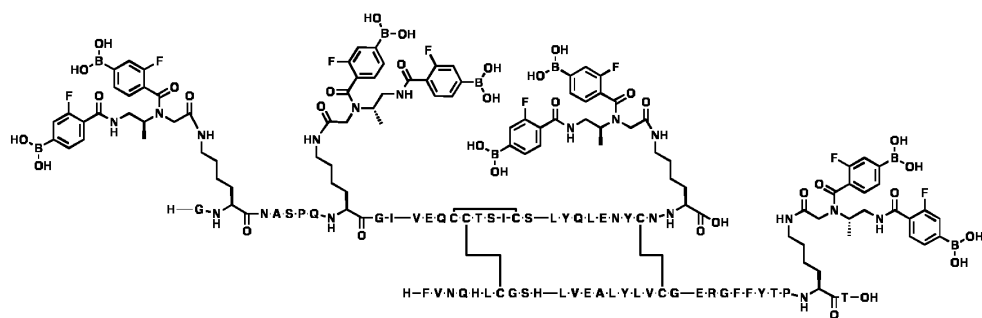
Example 541:**Example 542:****5 Example 543:**

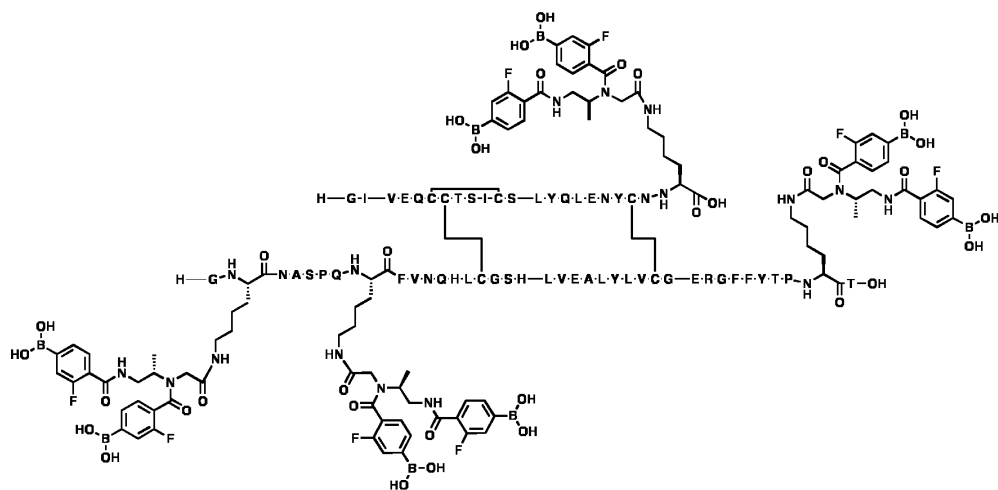
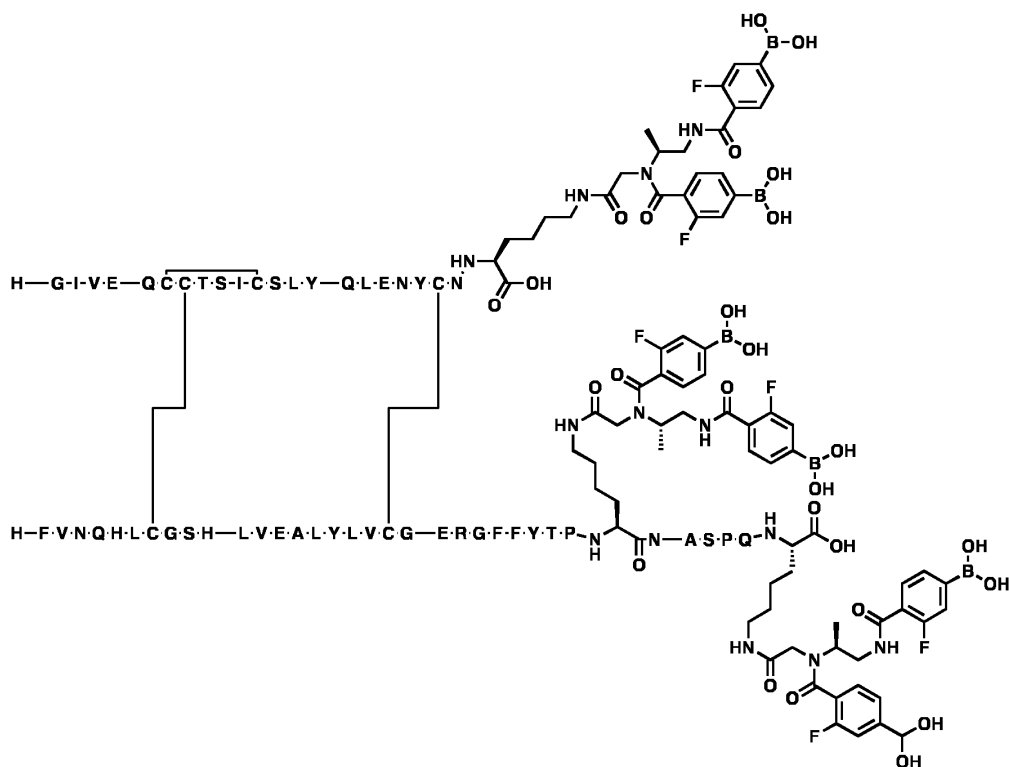
Example 544:**Example 545:**

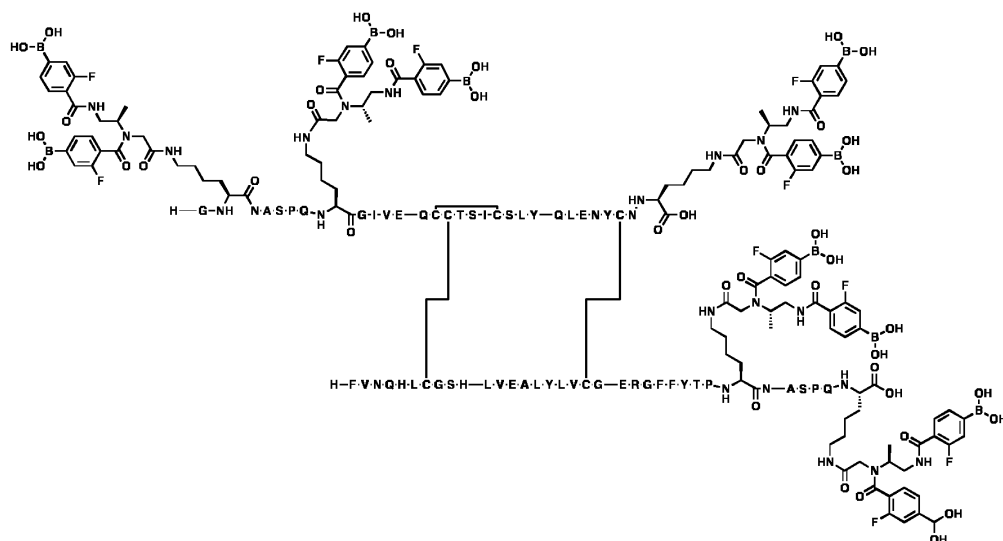
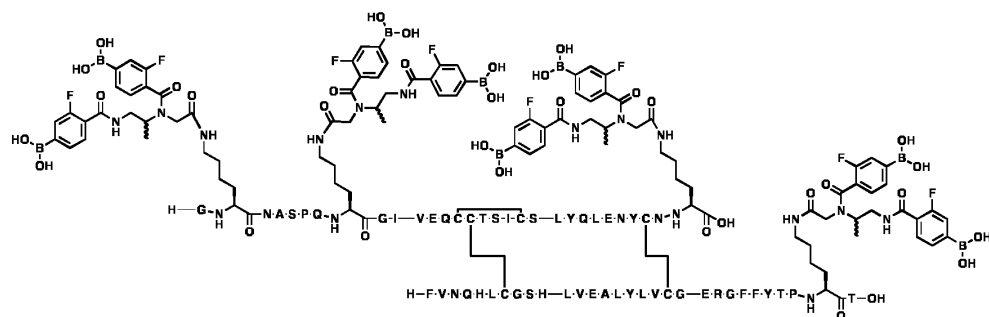
5

Example 546:**Example 547:**

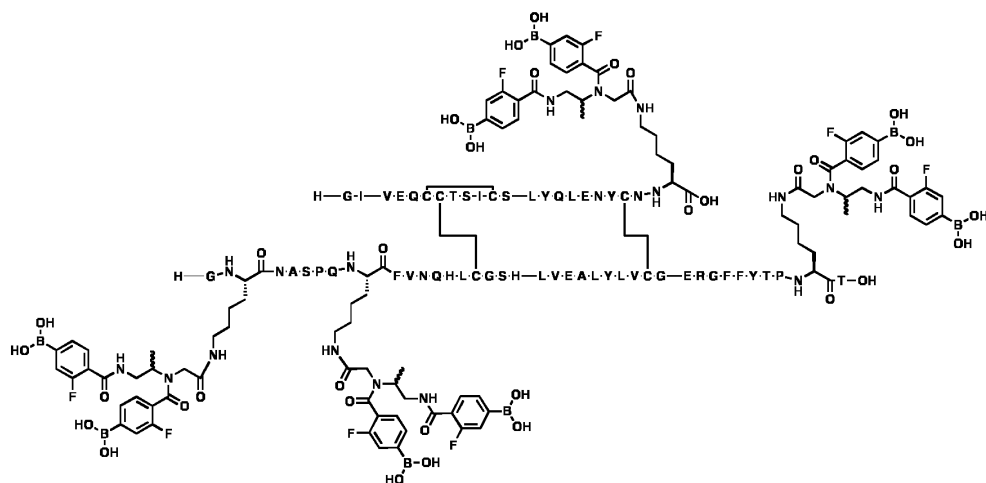
Example 548:**Example 549:**

Example 550:**Example 551:**

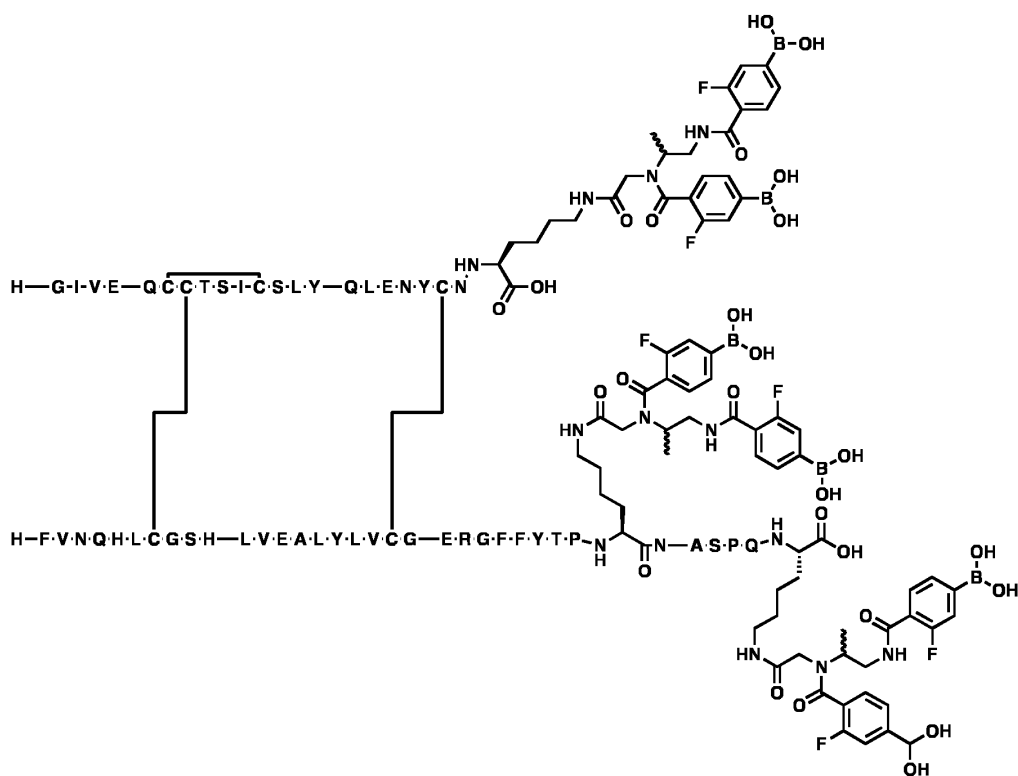
Example 552:**Example 553:**

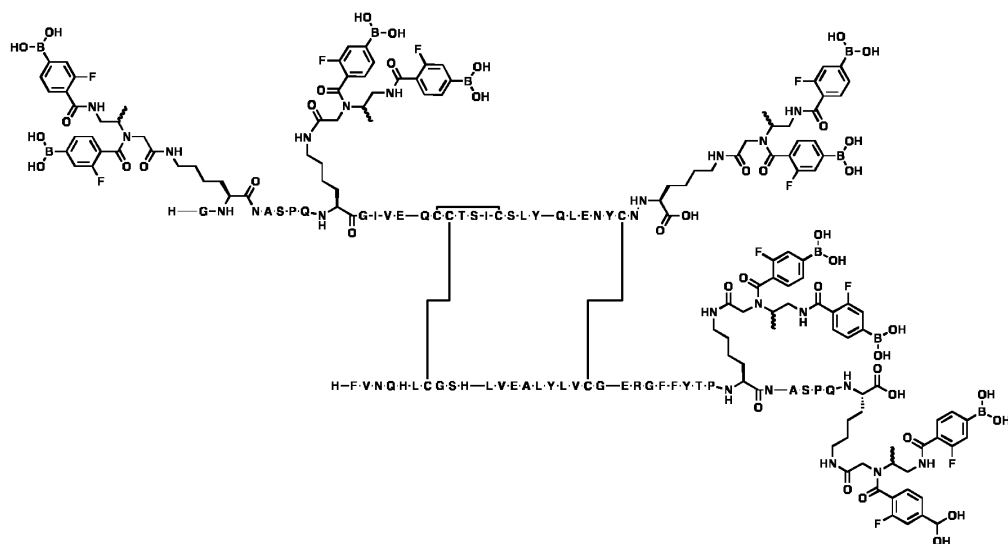
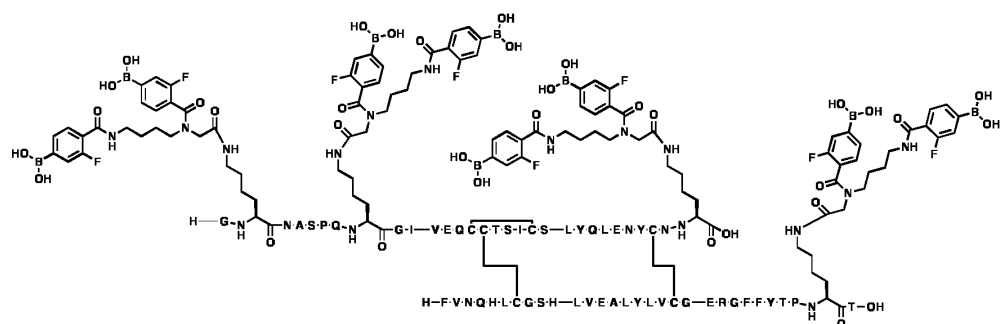
Example 554:**Example 555:**

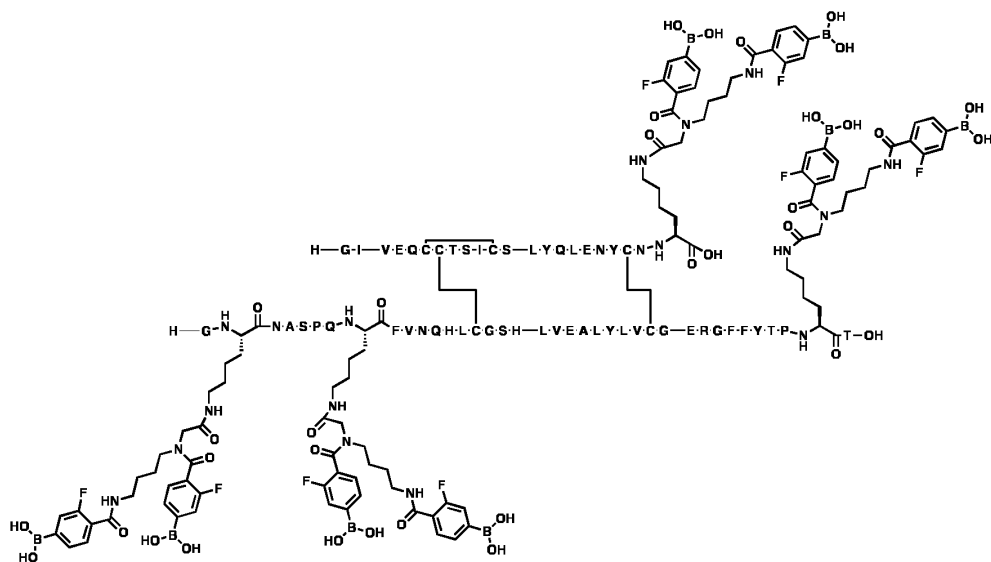
Example 556:



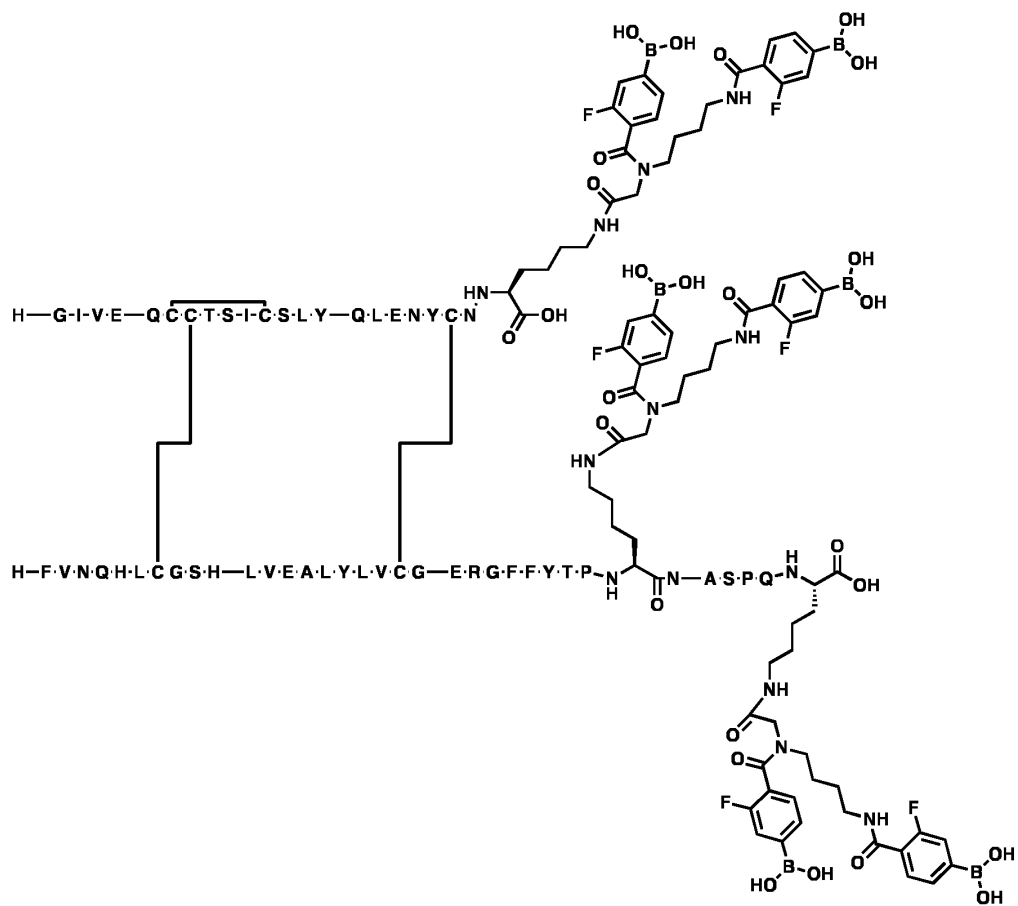
Example 557:

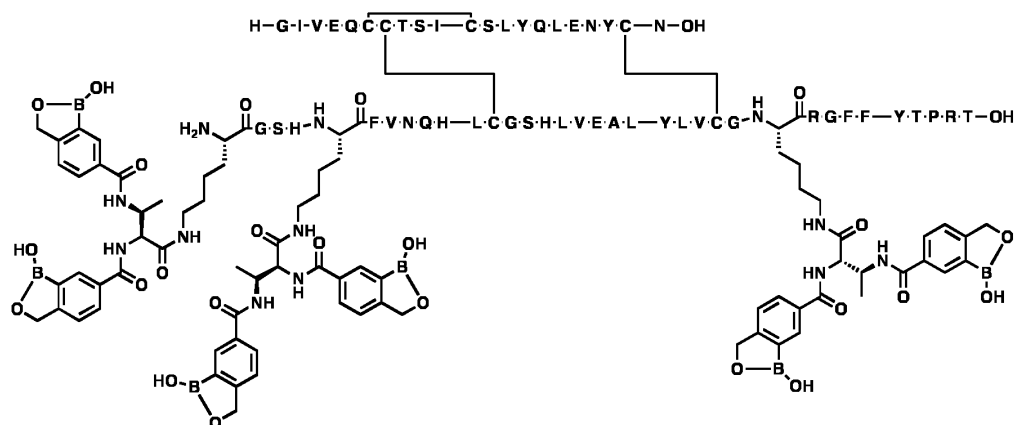
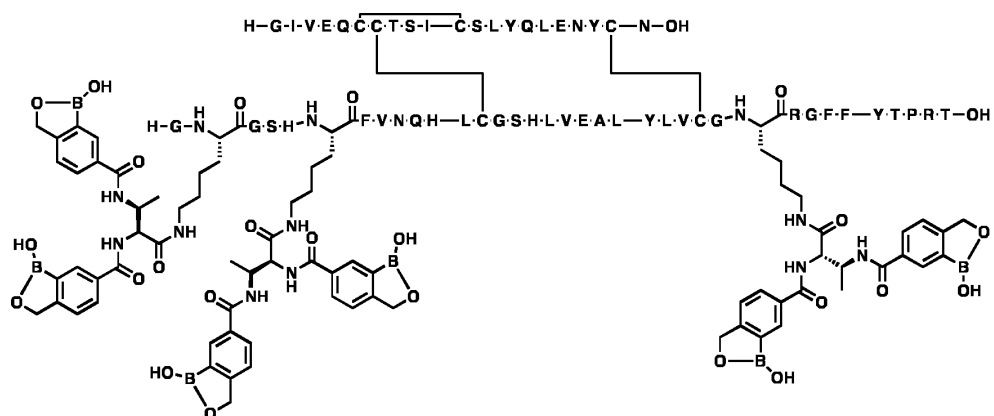


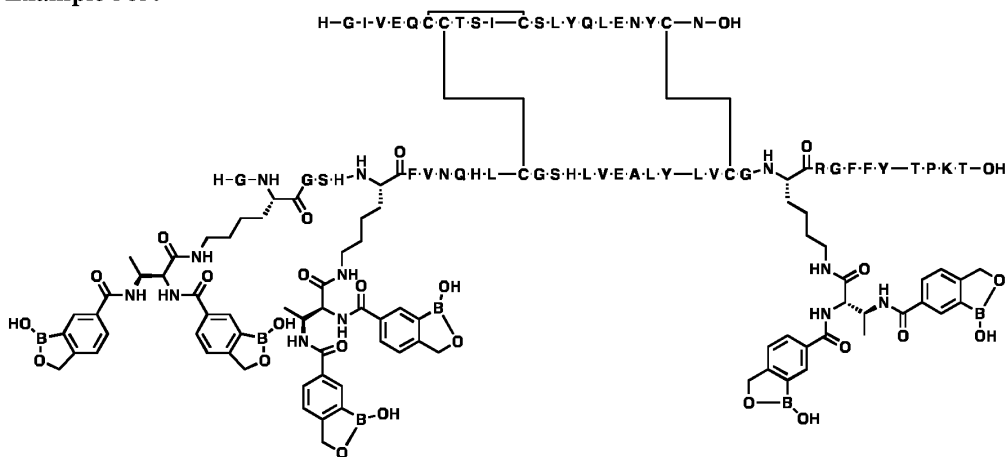
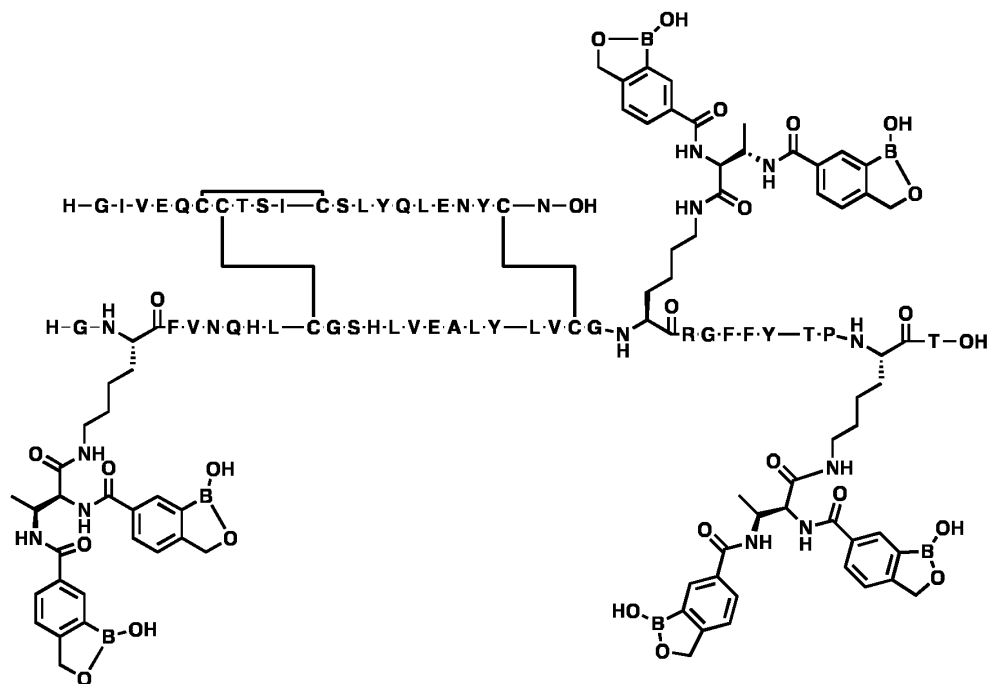
Example 558:**Example 559:**

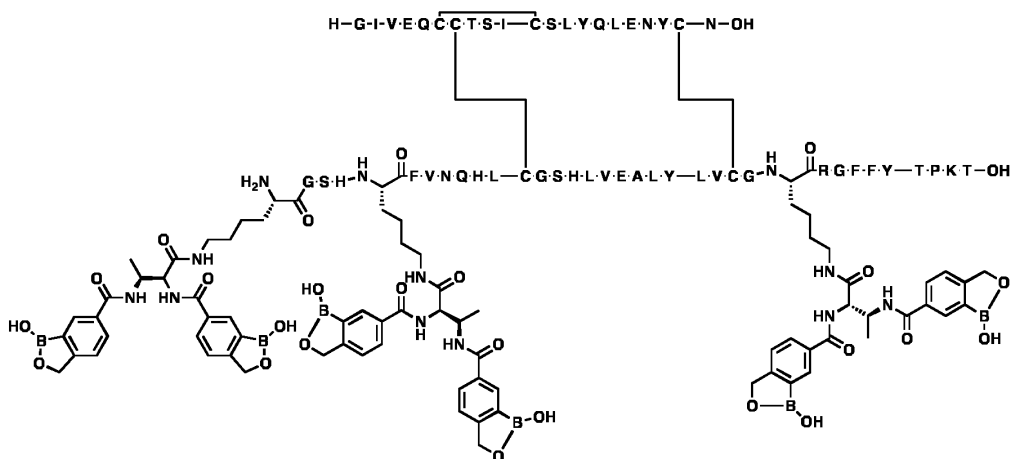
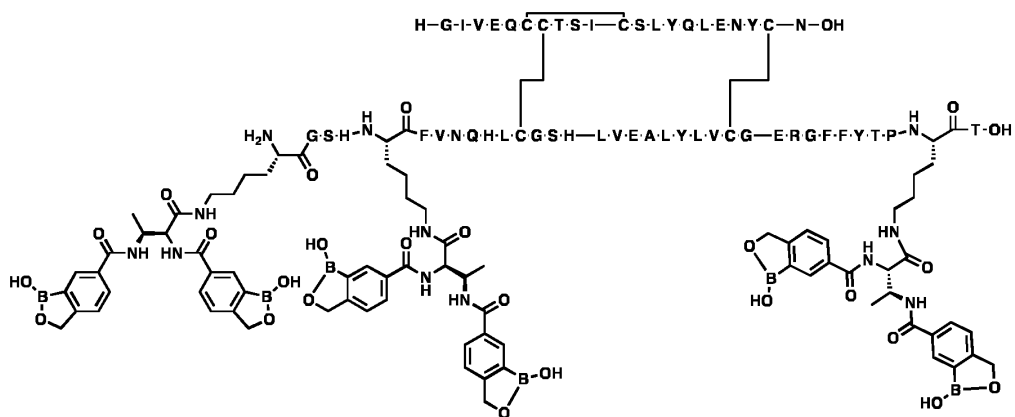
Example 560:

Example 561:

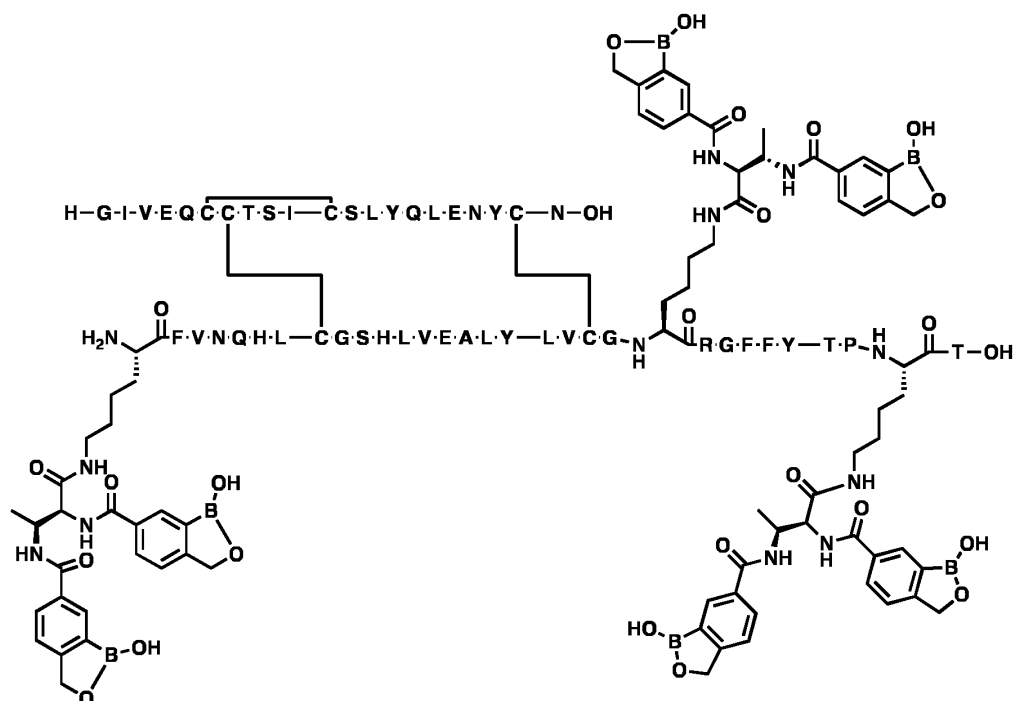


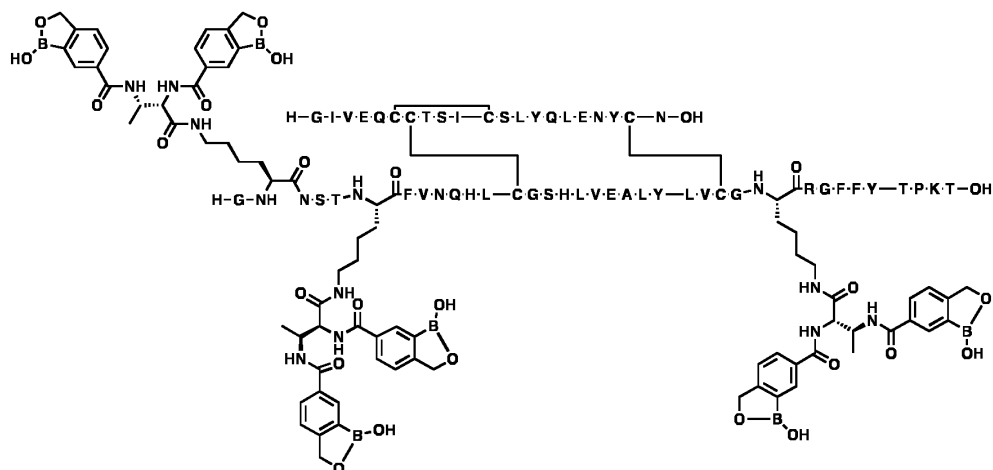
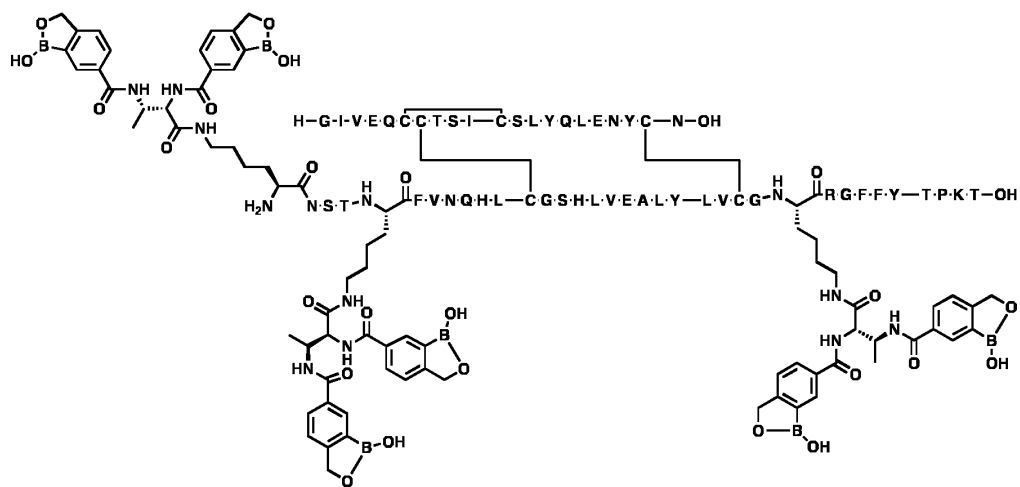
Example 562:**Example 563:**

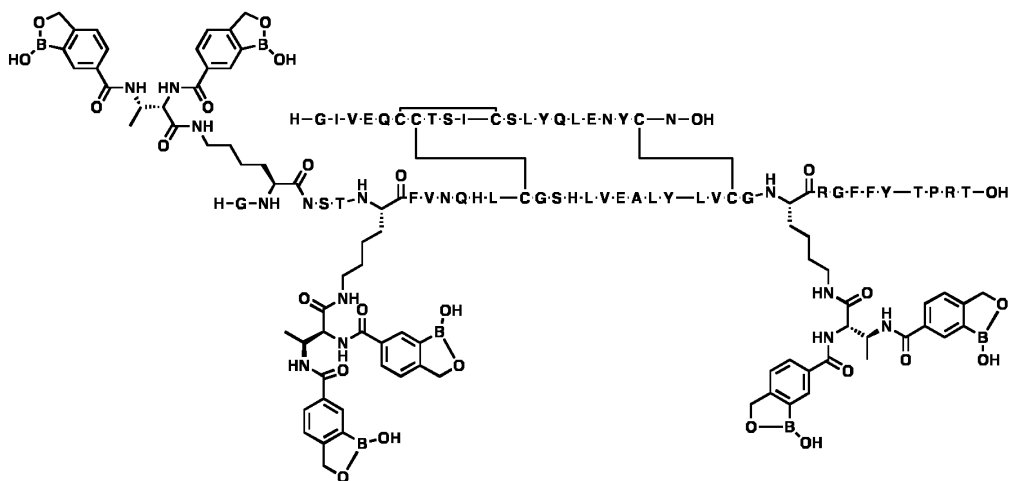
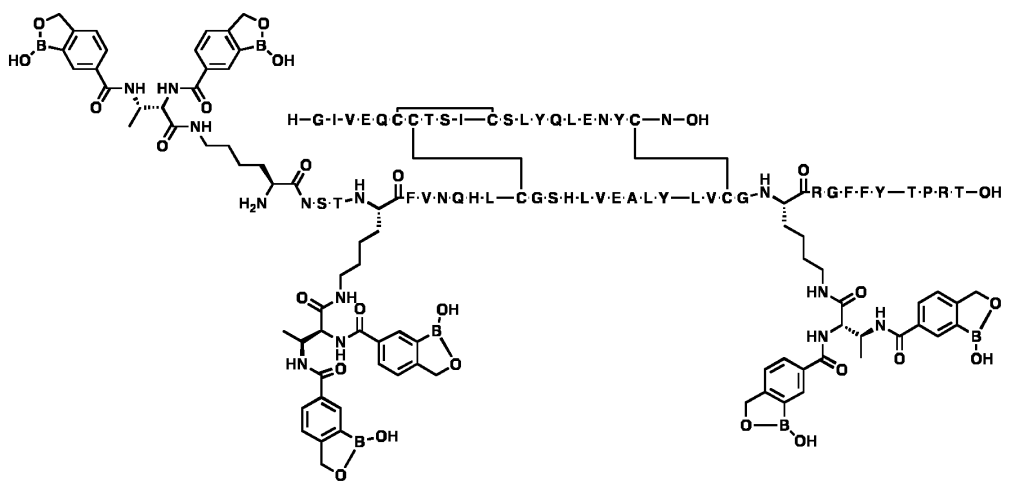


Example 566:**Example 567:**

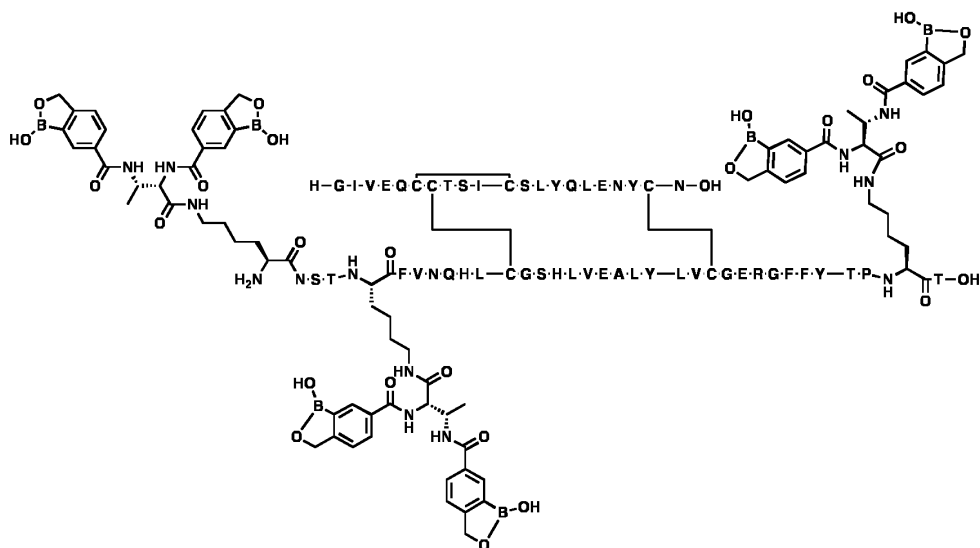
Example 568:



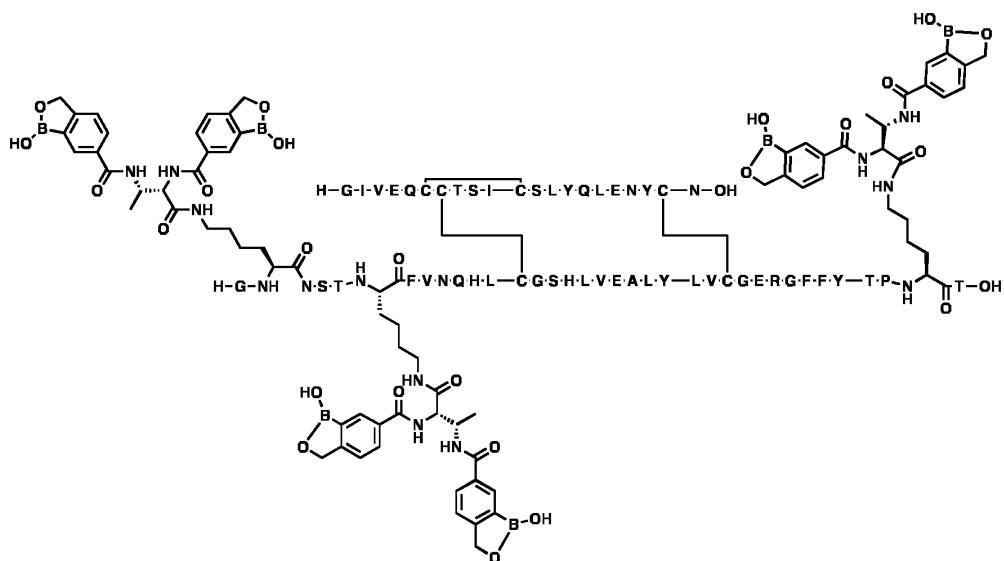
Example 569:**Example 570:**

Example 571:**Example 572:**

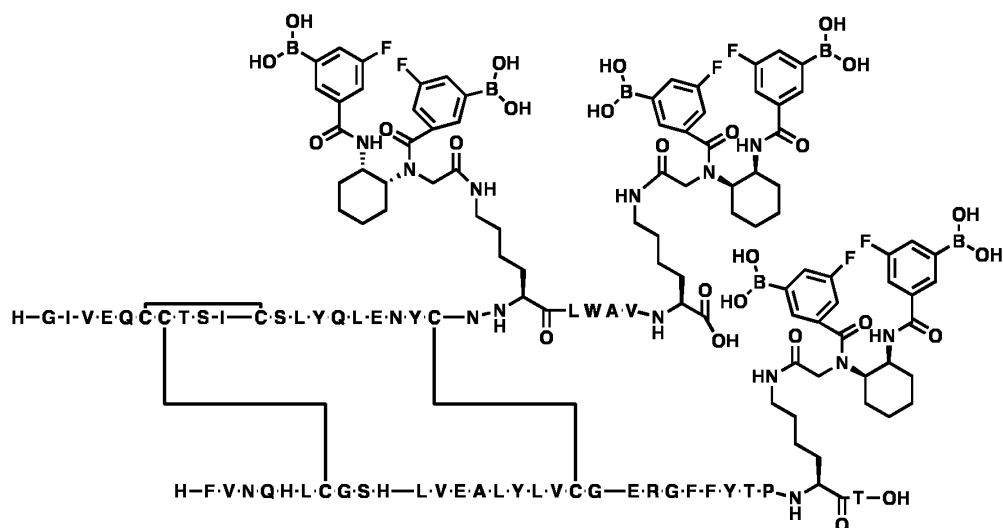
Example 573:



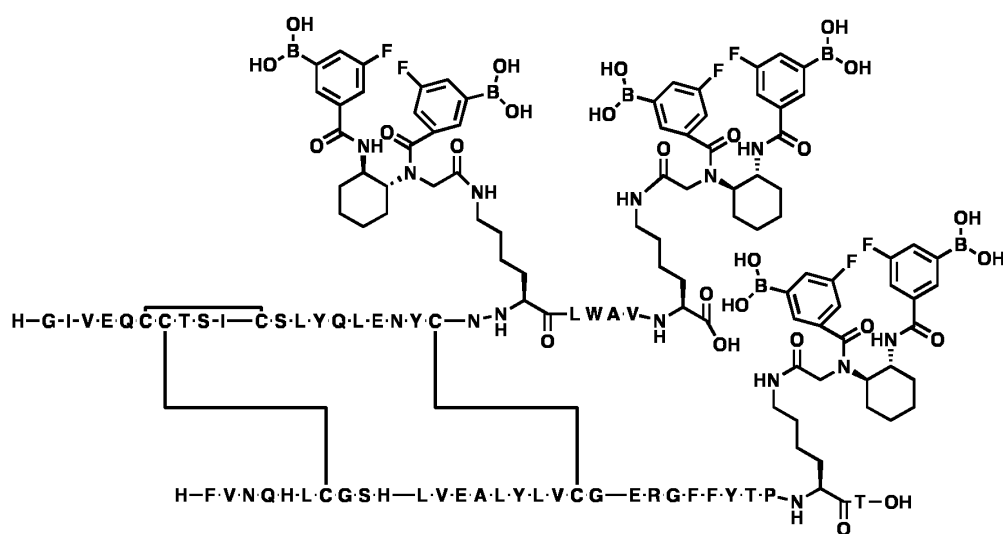
Example 574:

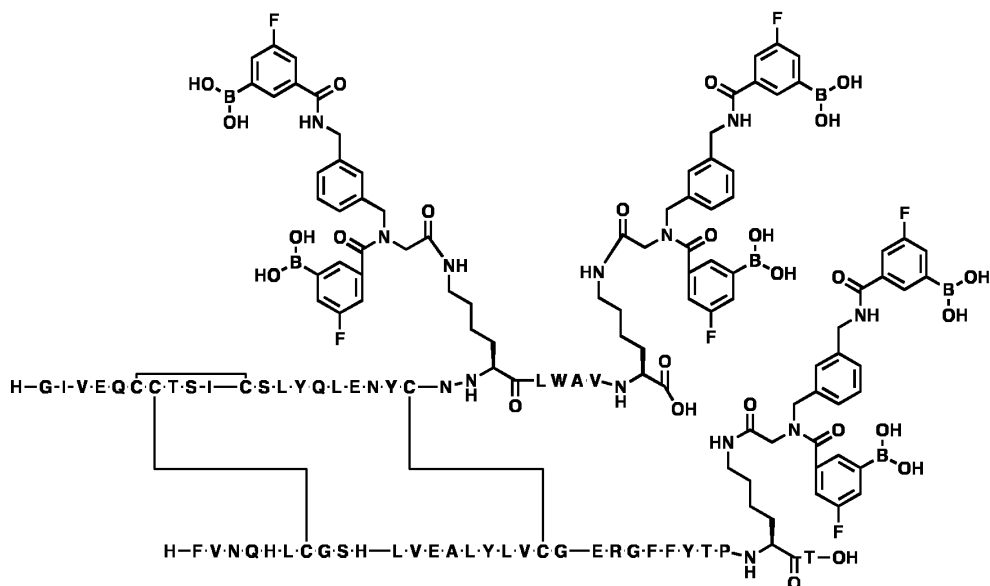
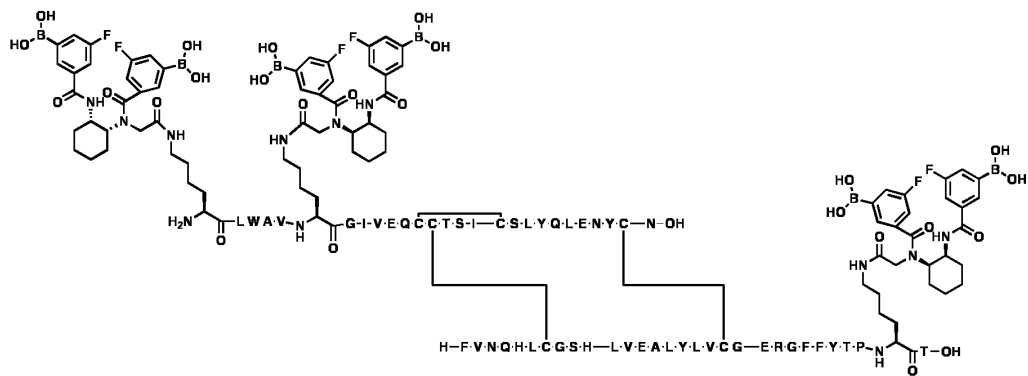


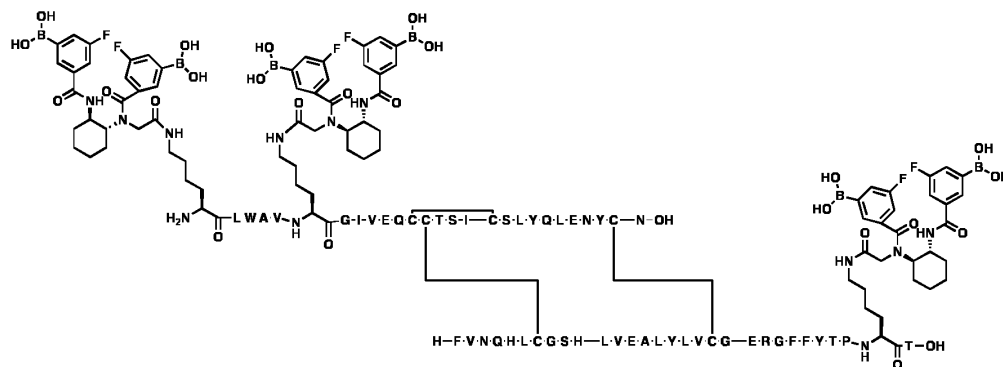
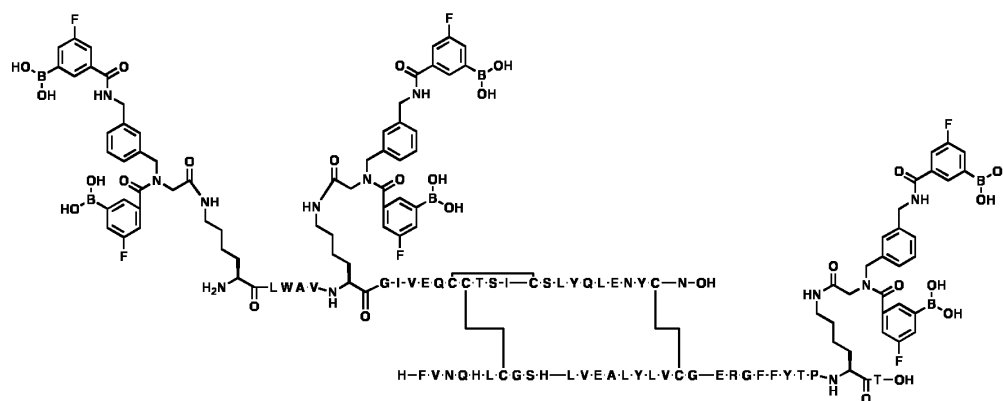
Example 575:

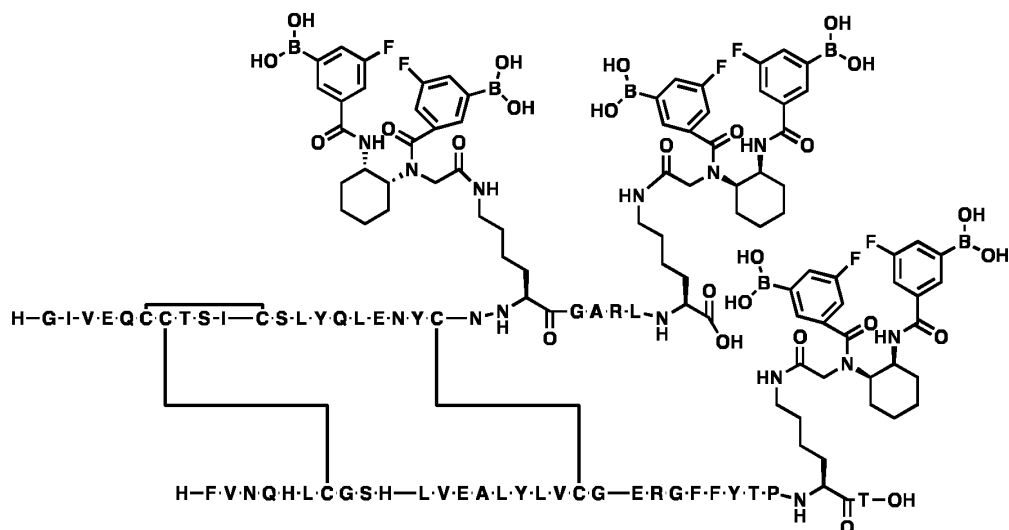


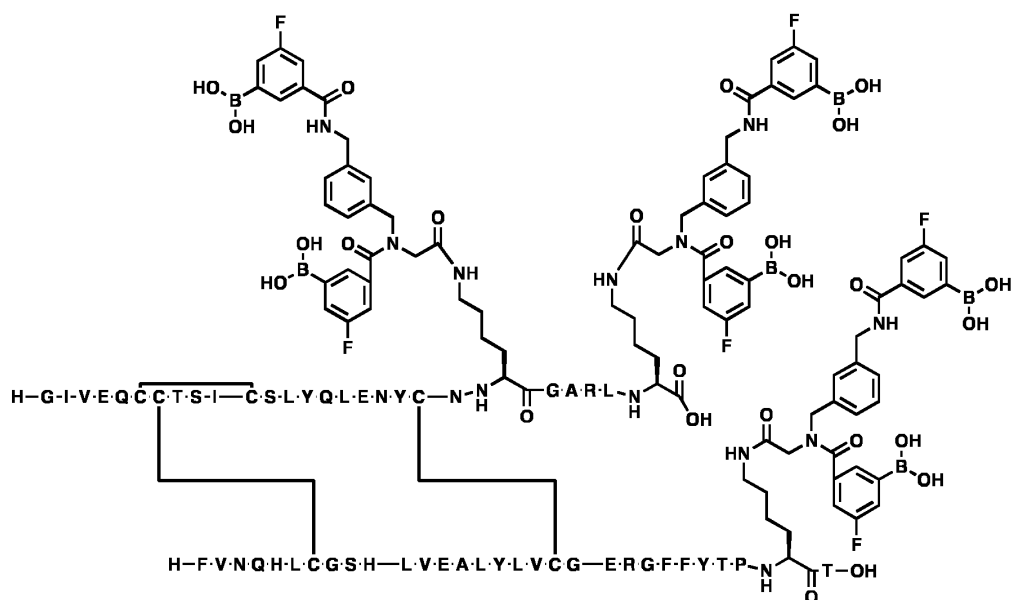
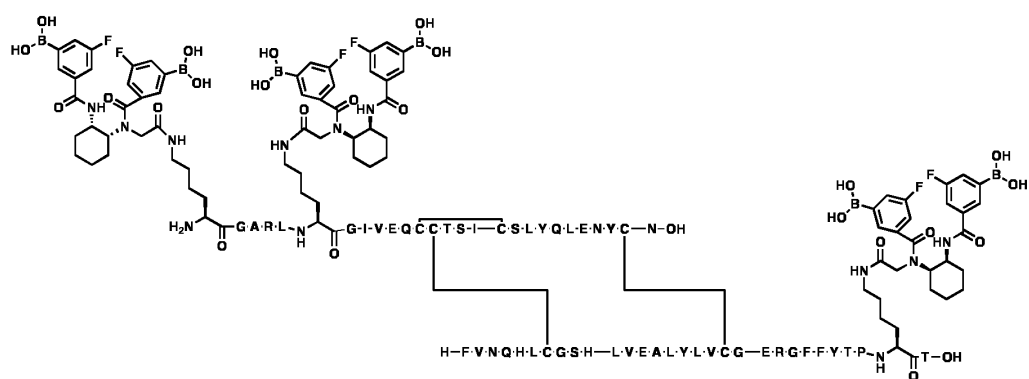
Example 576:

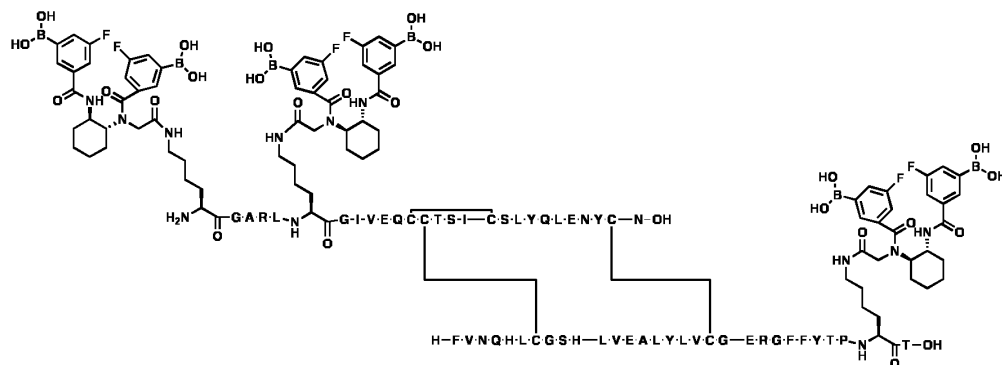
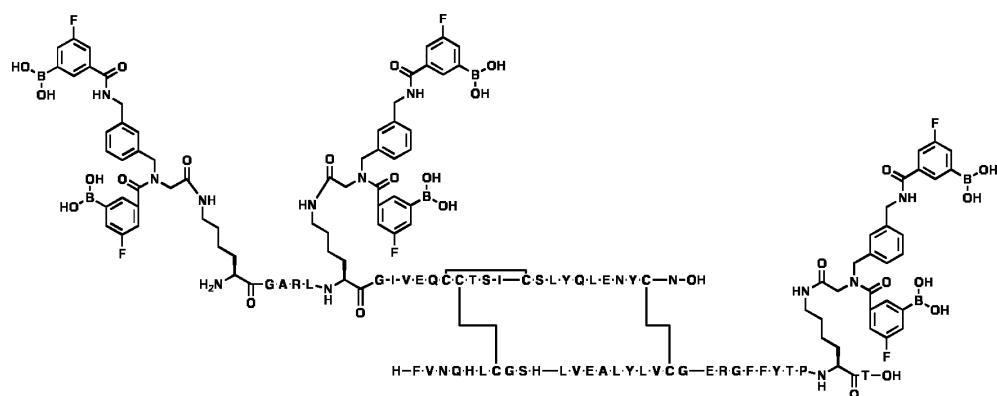
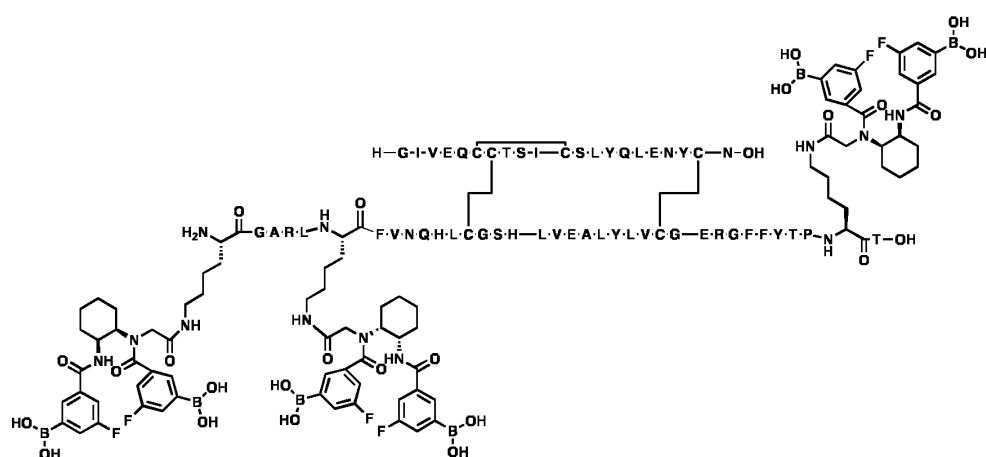


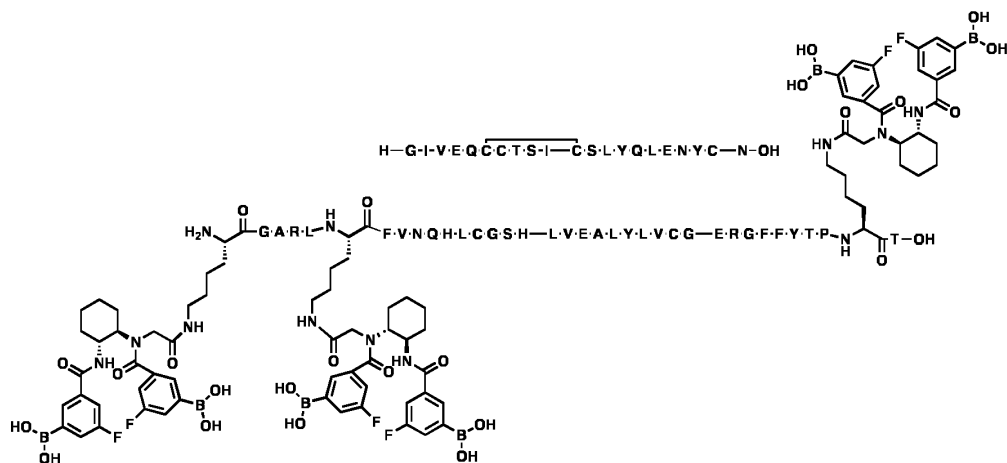
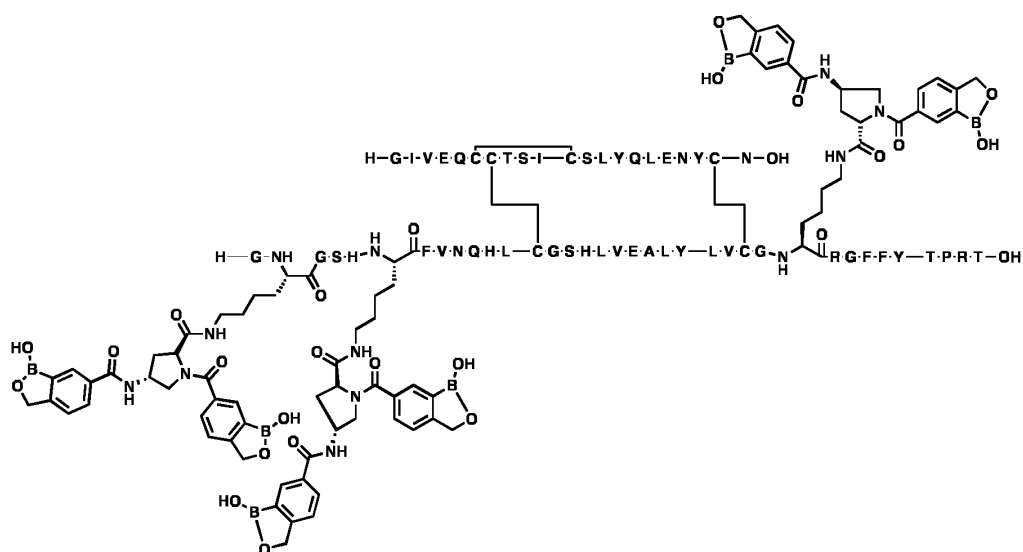
Example 577:**Example 578:**

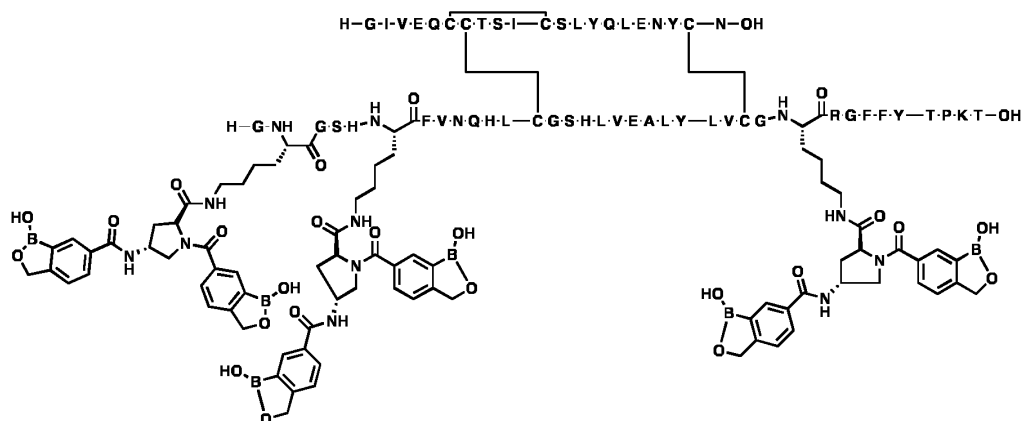
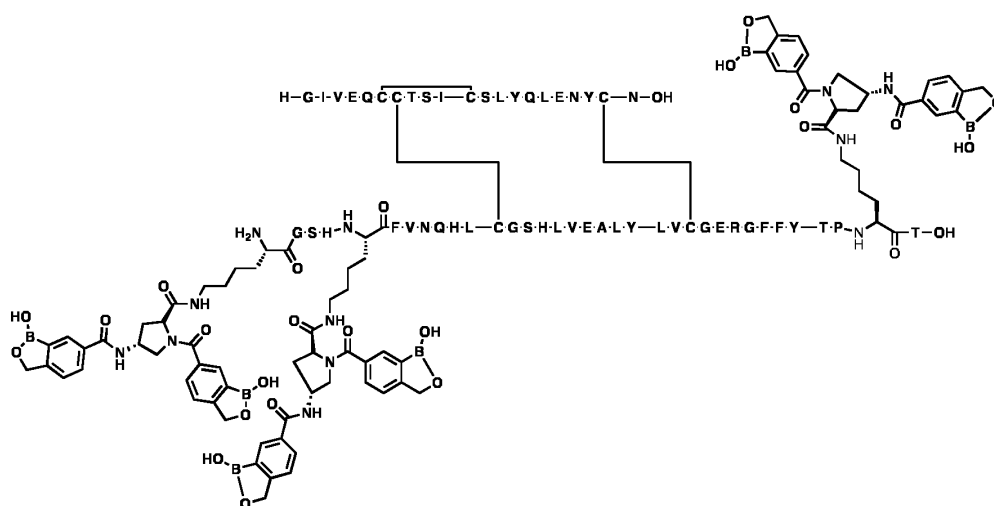
Example 579:**Example 580:**

[illegible]

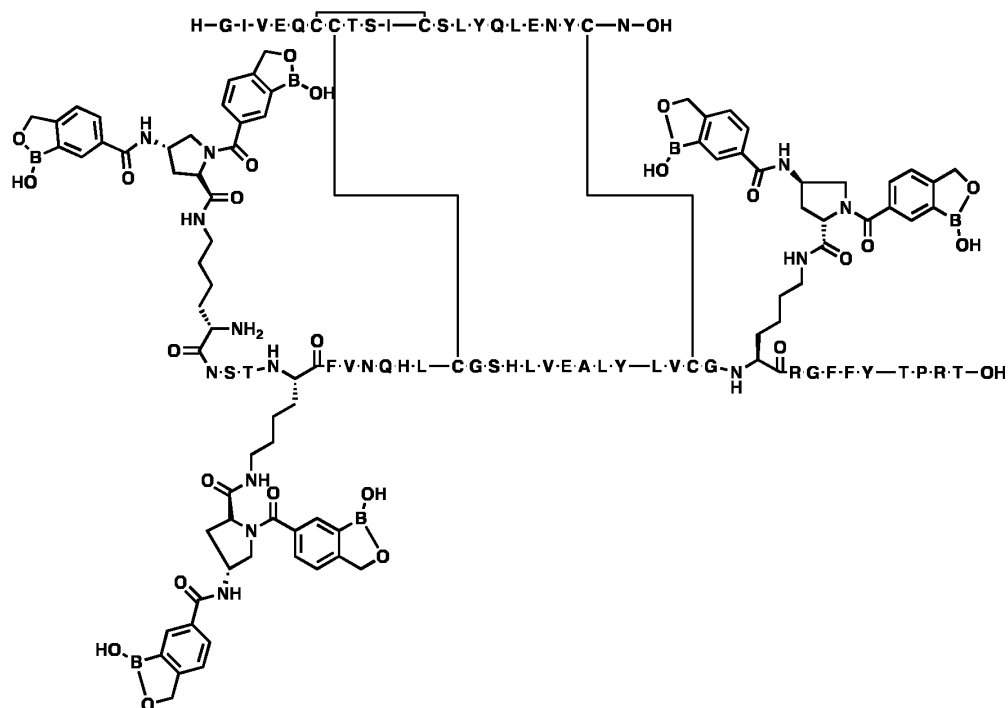
Example 583:**Example 584:**

Example 585:**Example 586:****5 Example 587:**

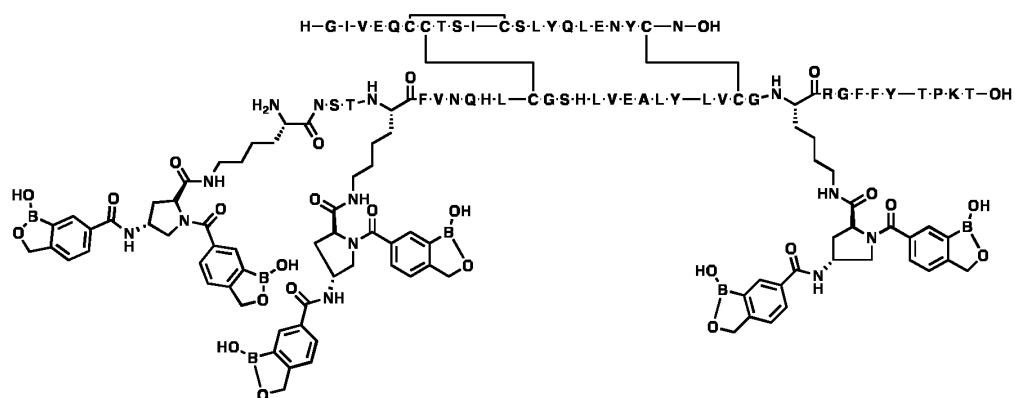
Example 588:**Example 589:**

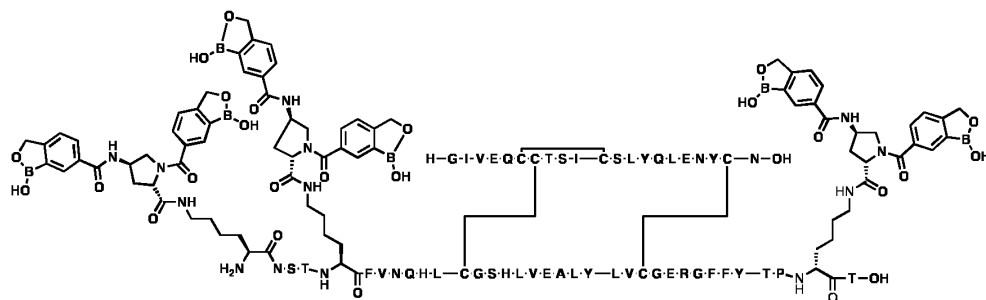
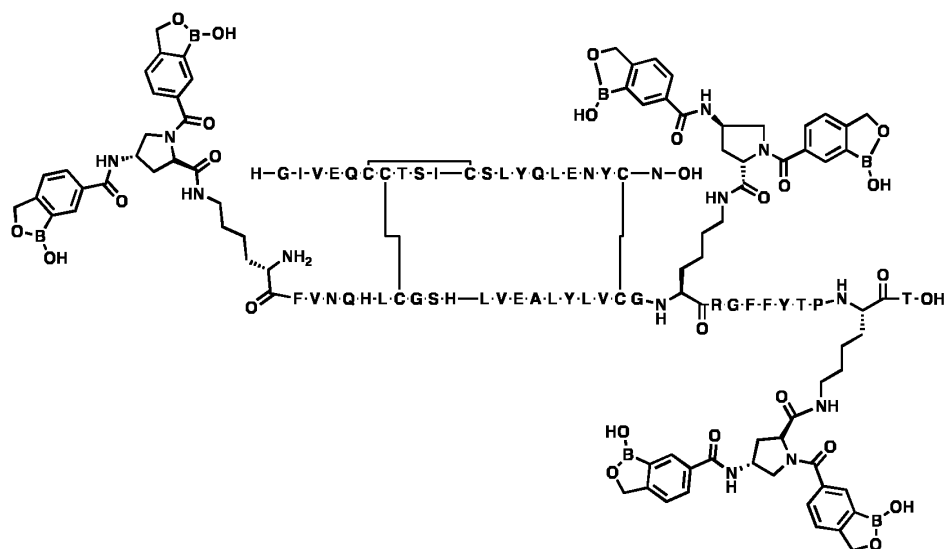
Example 590:**Example 591:**

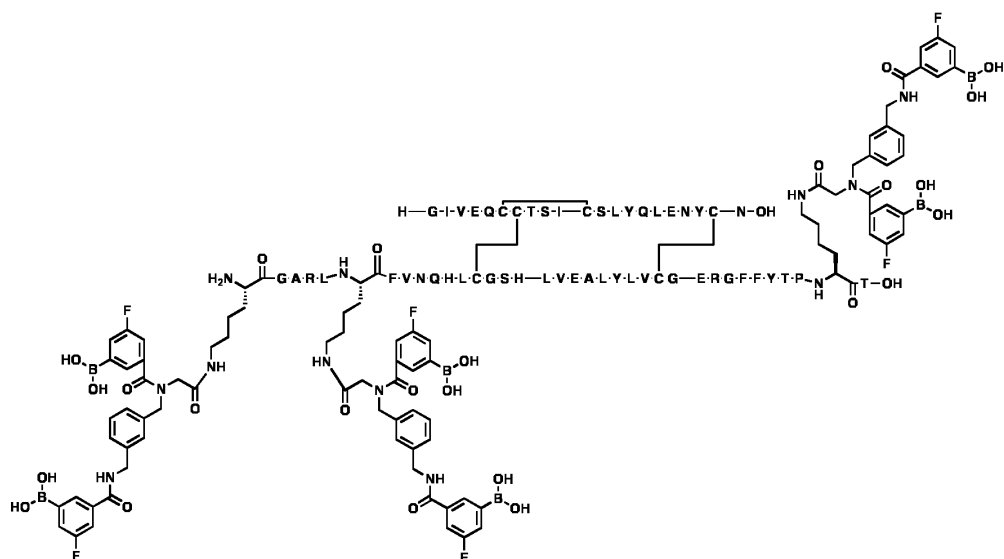
Example 592:

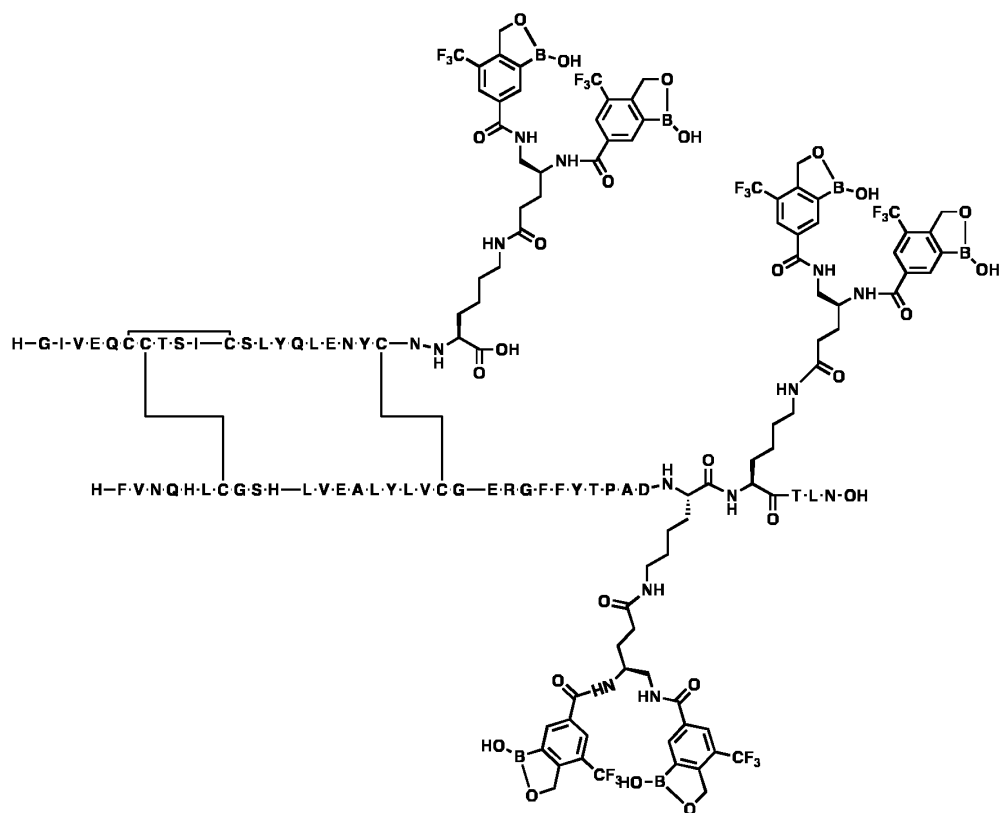


Example 593:

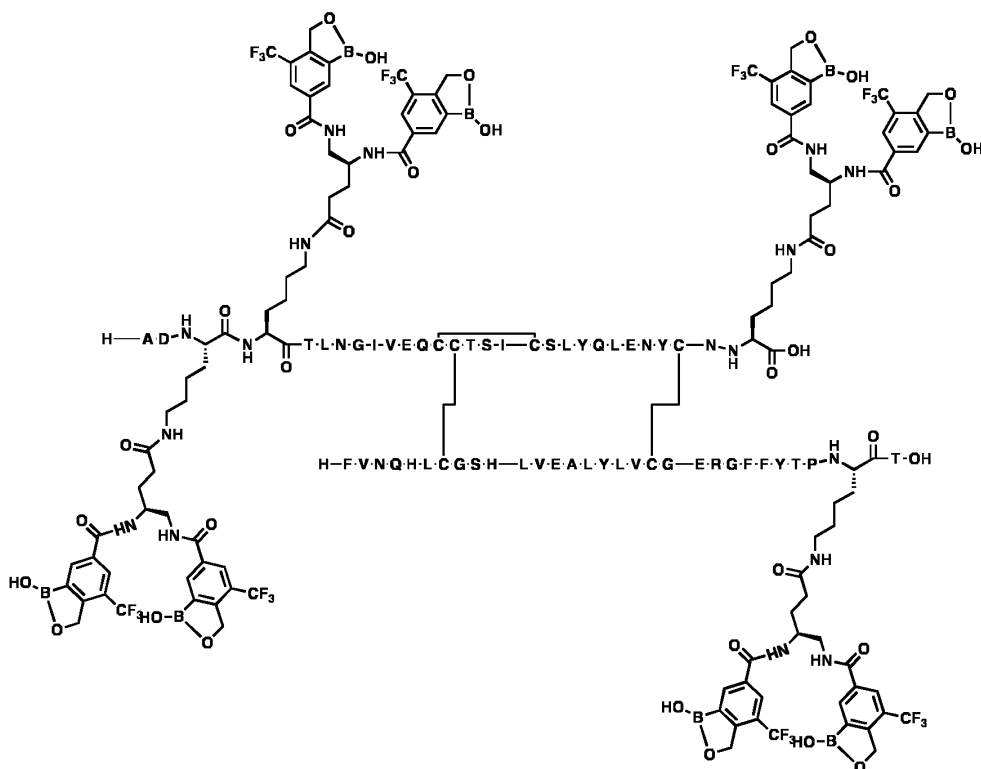


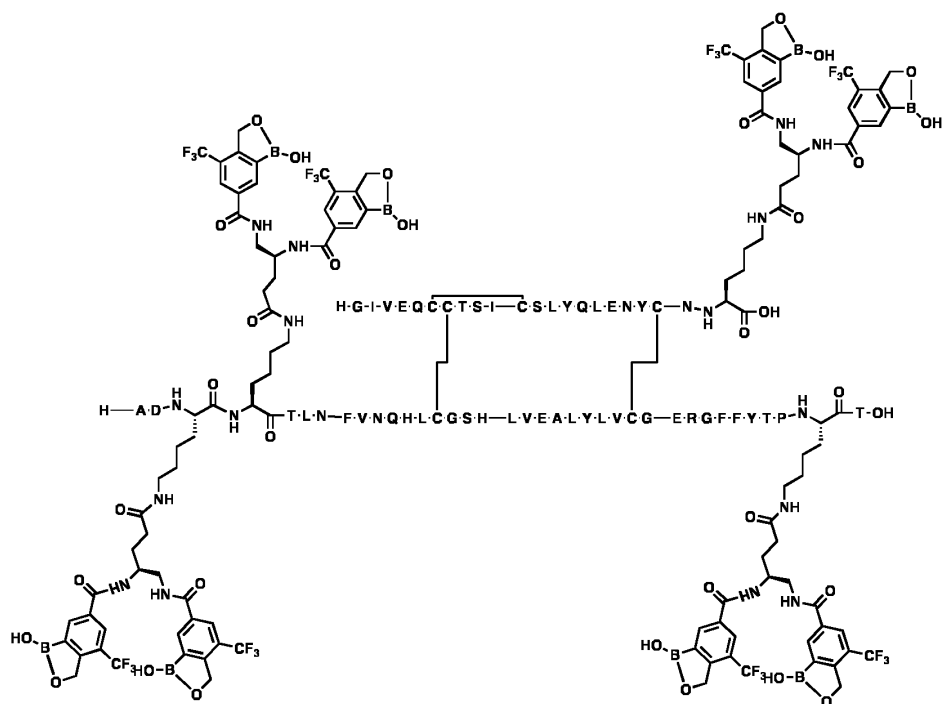
Example 594:**Example 595:**

Example 596:

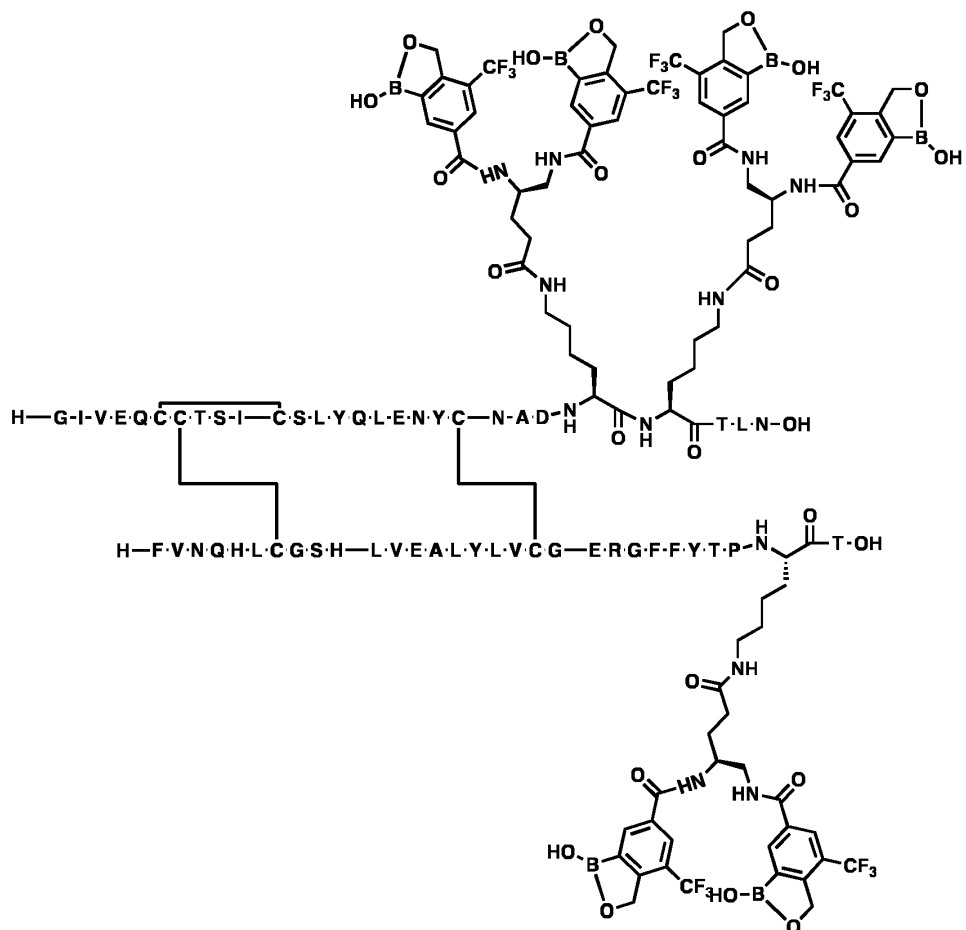
Example 597:

Example 598:

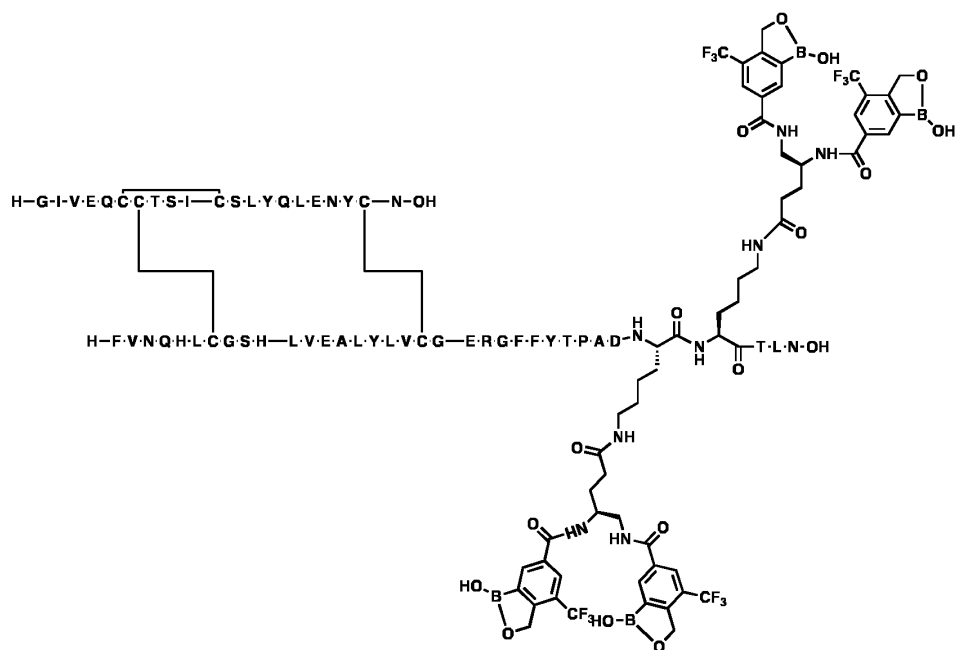


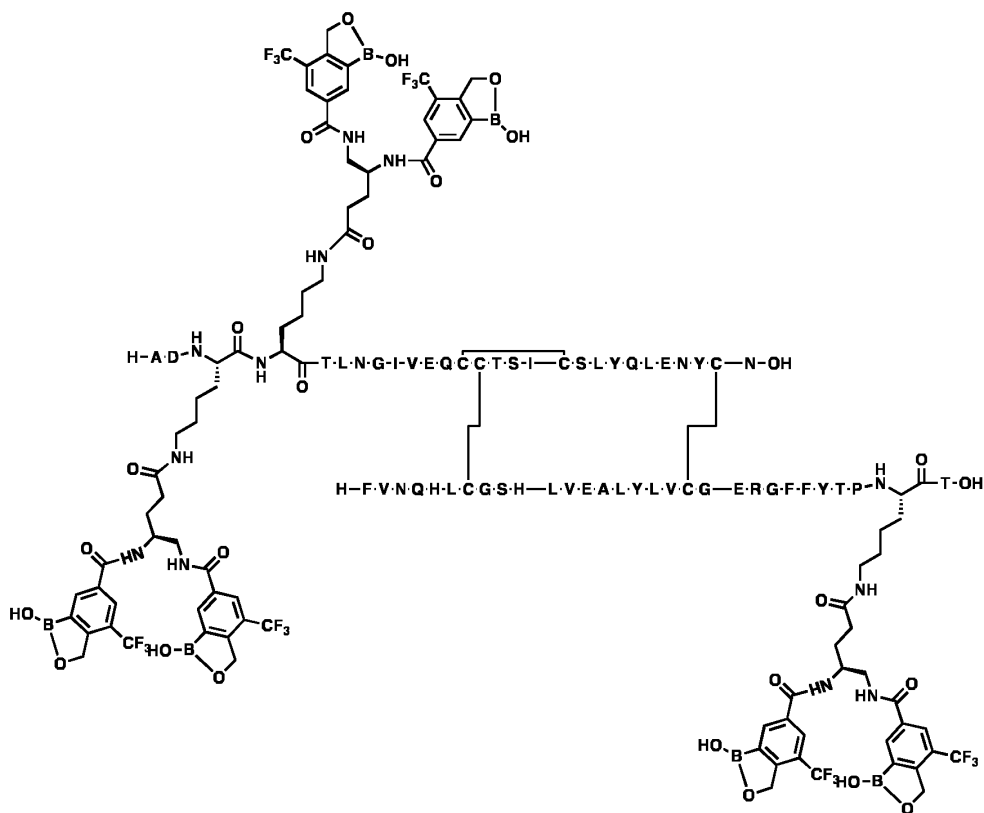
Example 599:

Example 600:

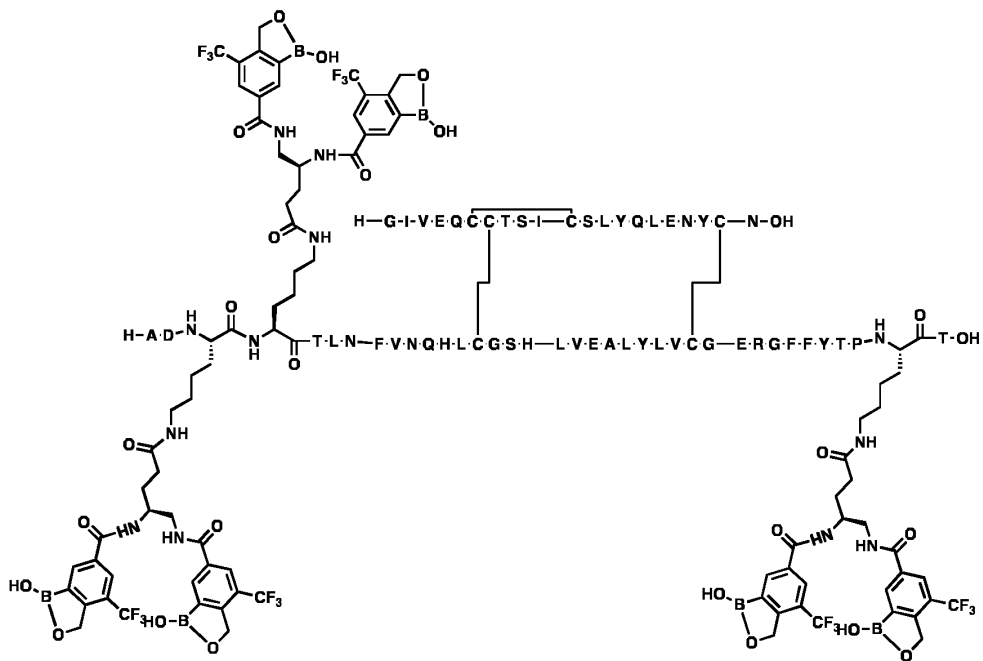


Example 601:

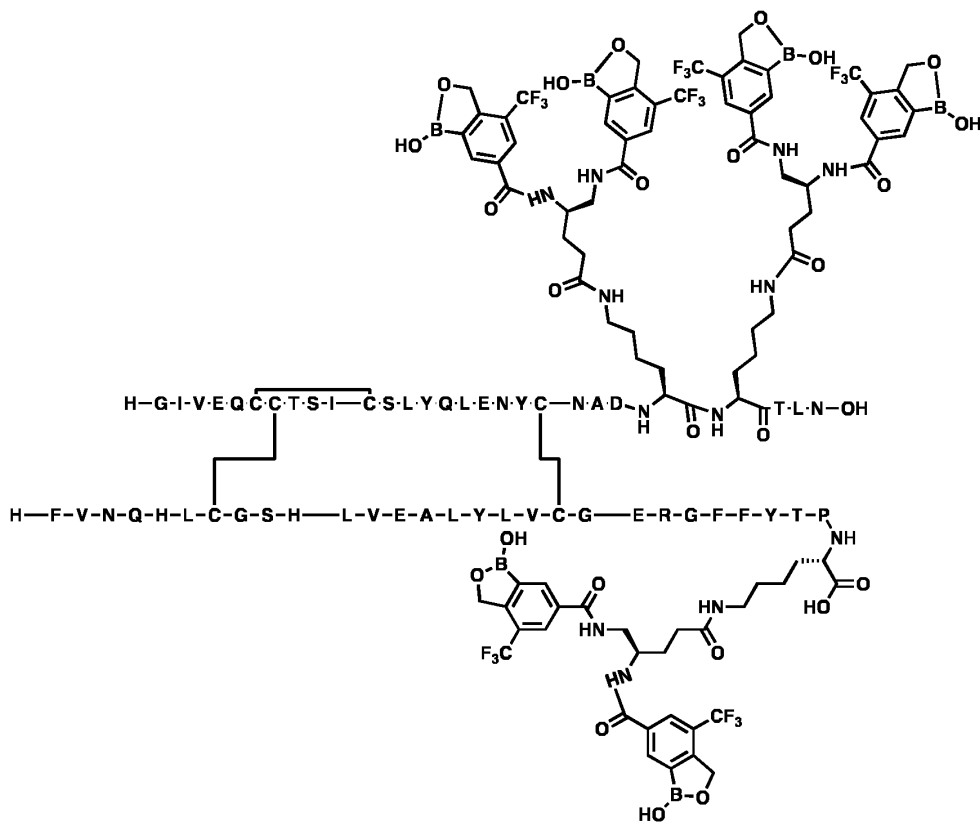




Example 603:

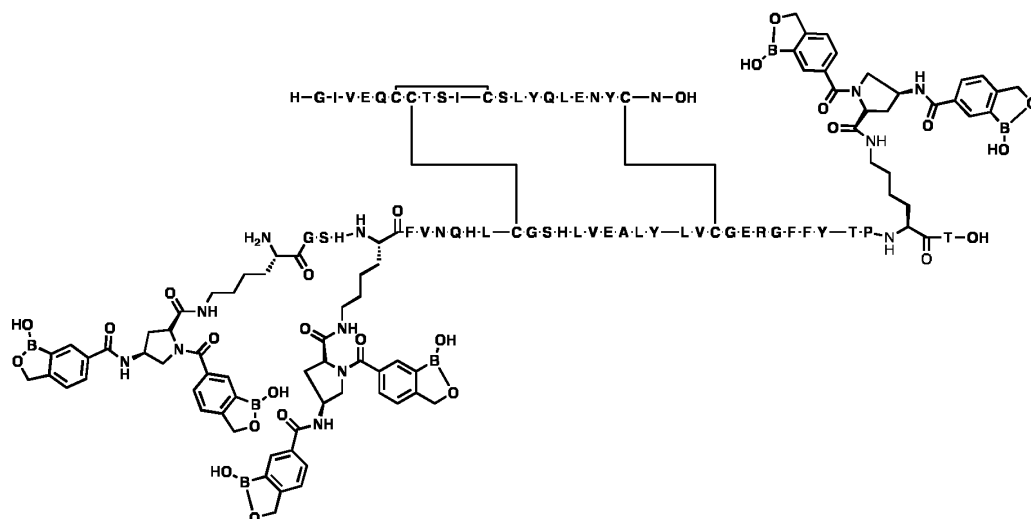


Example 604:

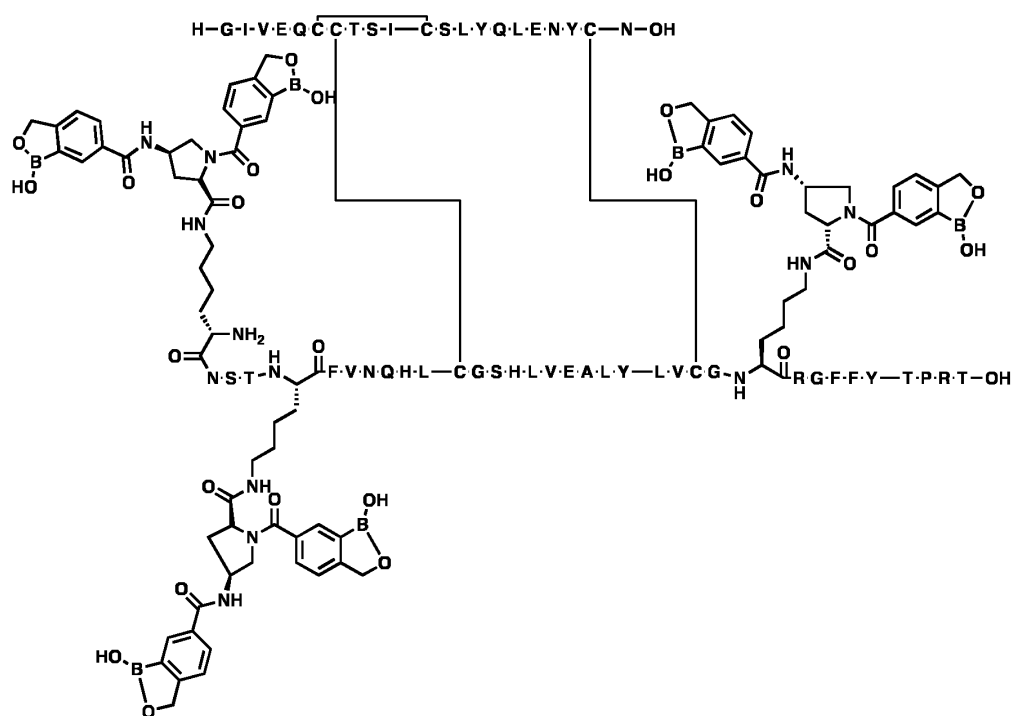


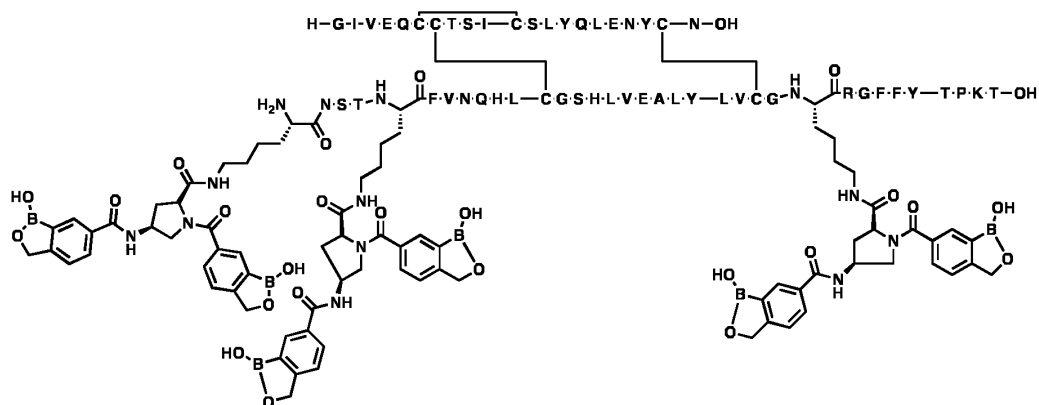
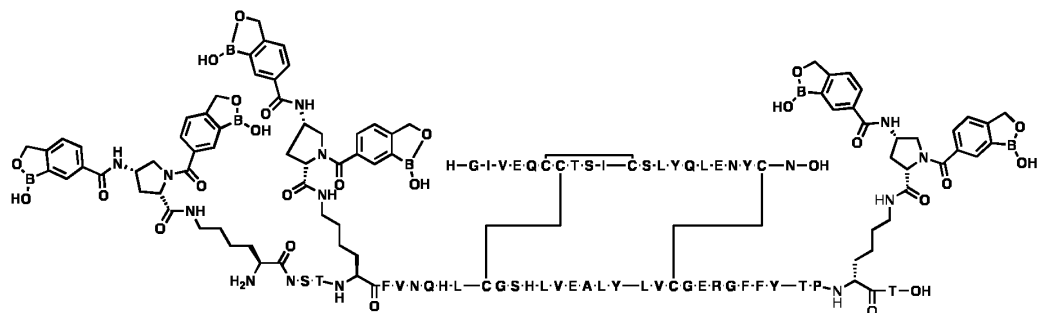
H-G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH
H-G-NH-C(=O)-S-H-NH-C(=O)-F-V-N-Q-H-L-C-G-S-H-L-V-E-A-L-Y-L-V-C-G-NH-C(=O)-R-G-F-F-Y-T-P-K-T-OH

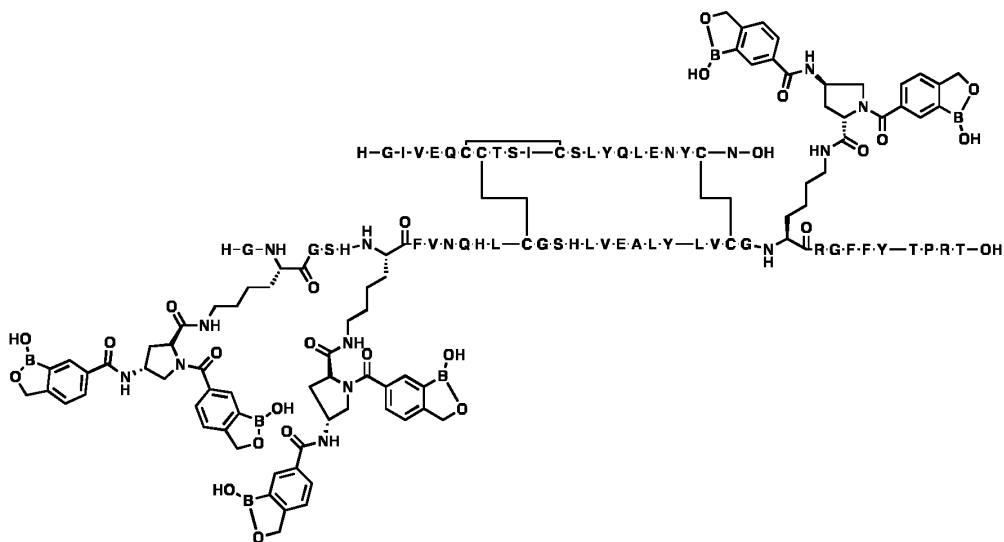
Example 607:



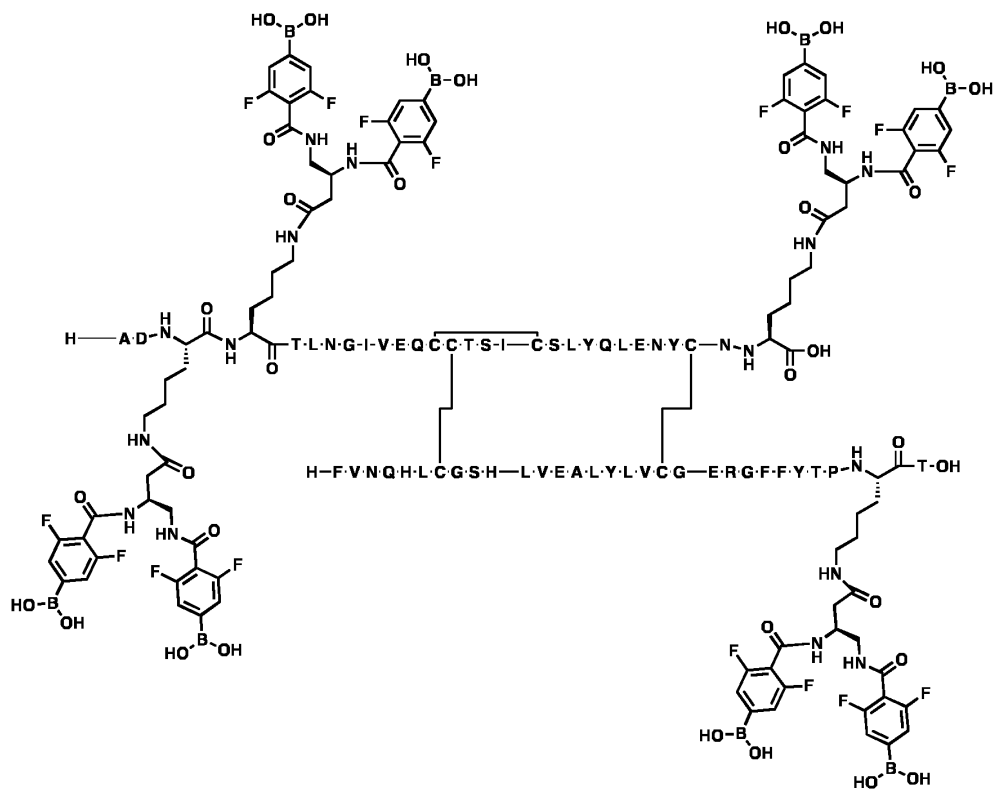
Example 608:



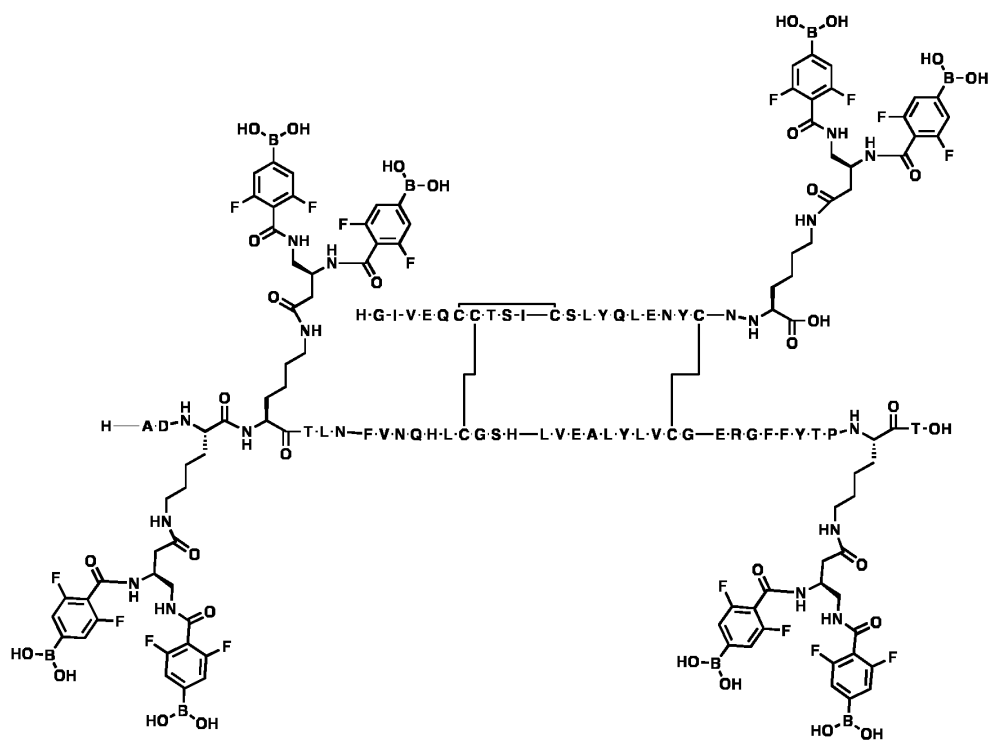
Example 609:**Example 610:**

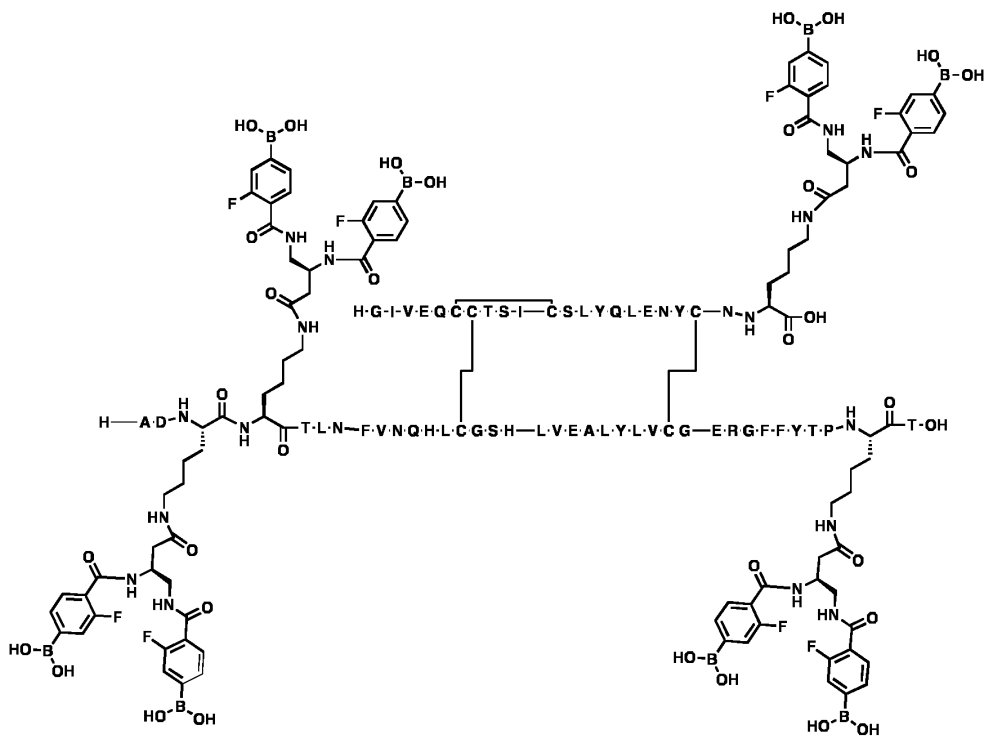


Chemical structure of the 125-residue protein structure, showing the full sequence and the three fluorinated boronate ester modifications. The sequence is H-G-I-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-NH₂. The modifications are attached to the Cys2, Cys10, and Cys11 residues. The Cys2 modification is a 4-fluorophenylboronate ester. The Cys10 modification is a 4-fluorophenylboronate ester. The Cys11 modification is a 4-fluorophenylboronate ester. The structure is shown in a ribbon representation with the modifications highlighted in red.

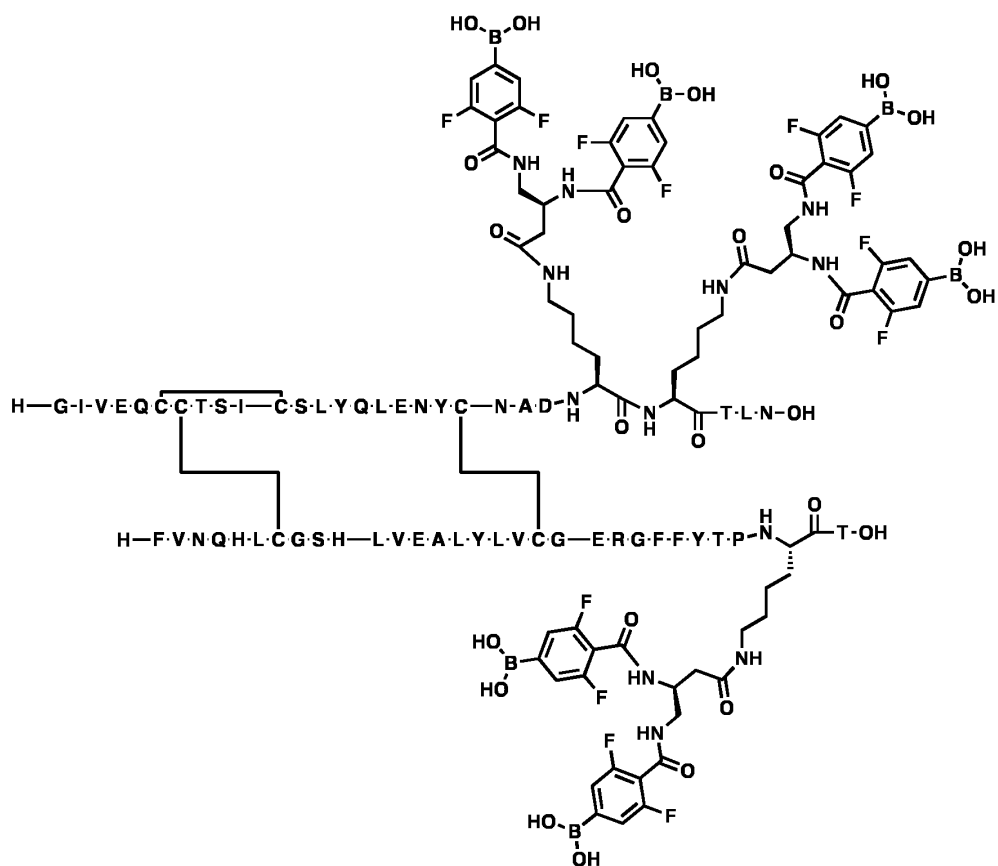


Example 615:

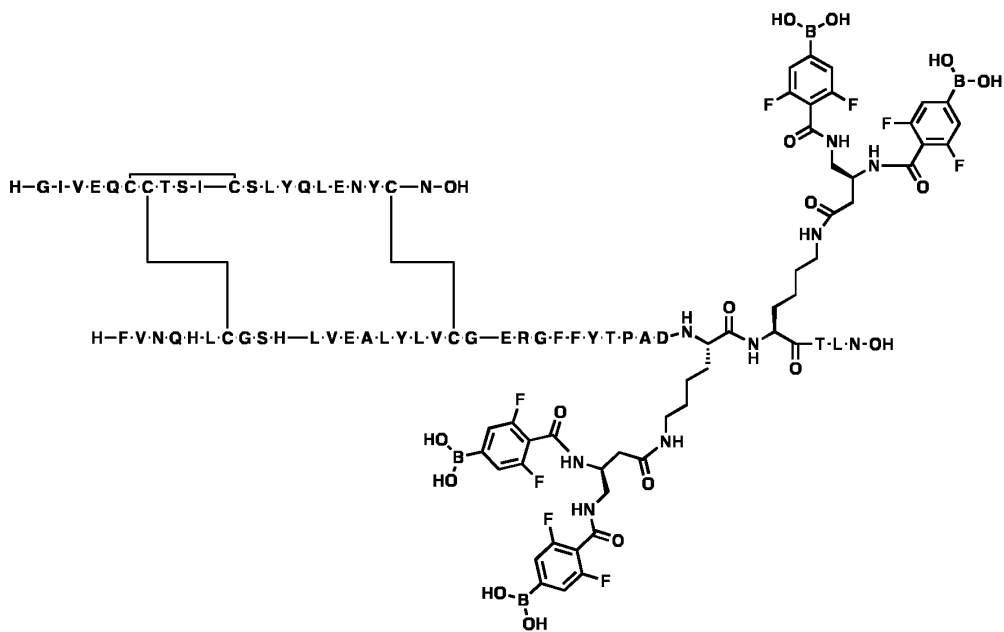




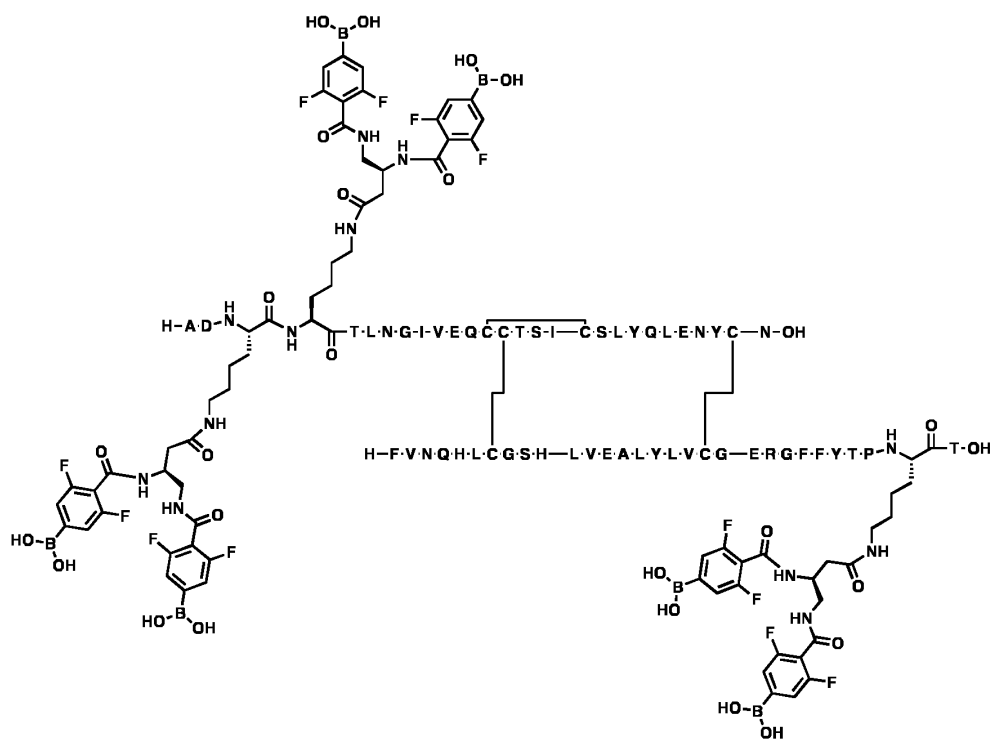
Example 617:



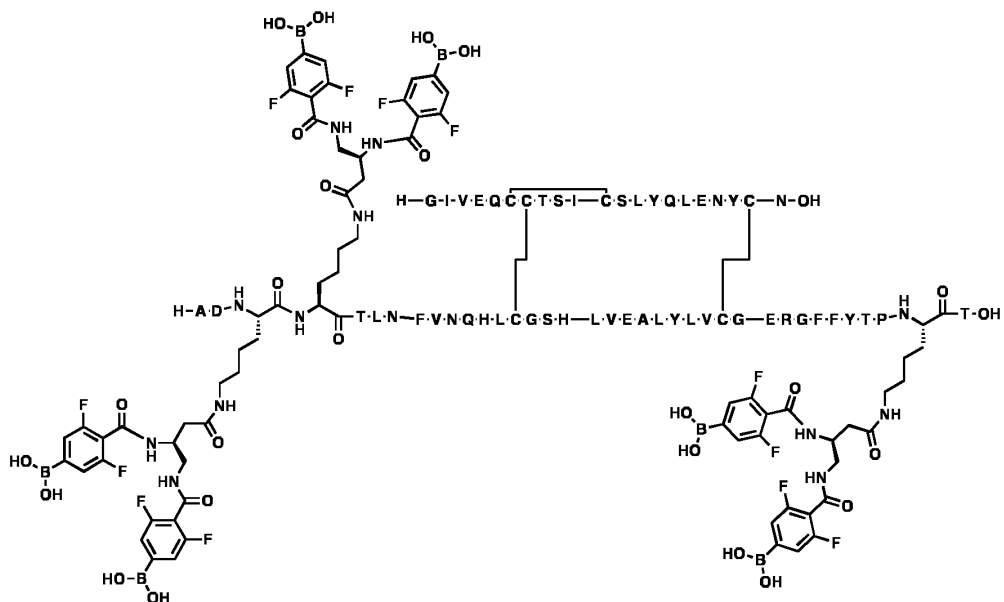
Example 618:



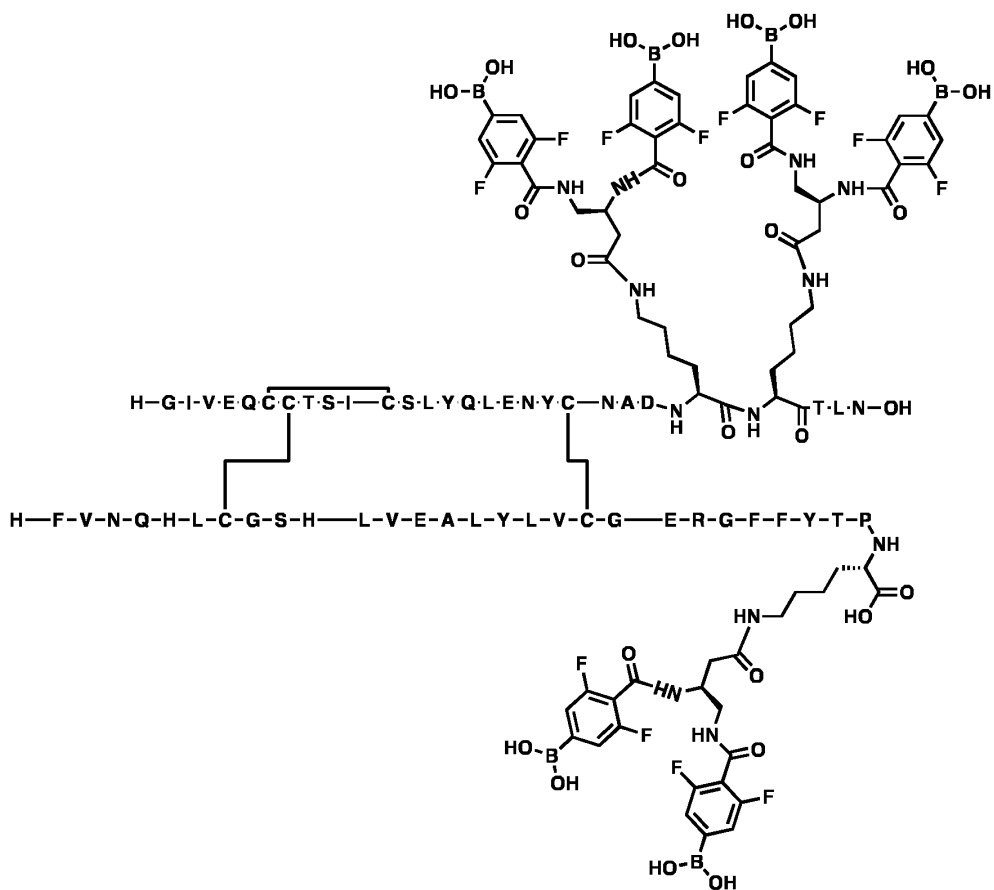
Example 619:



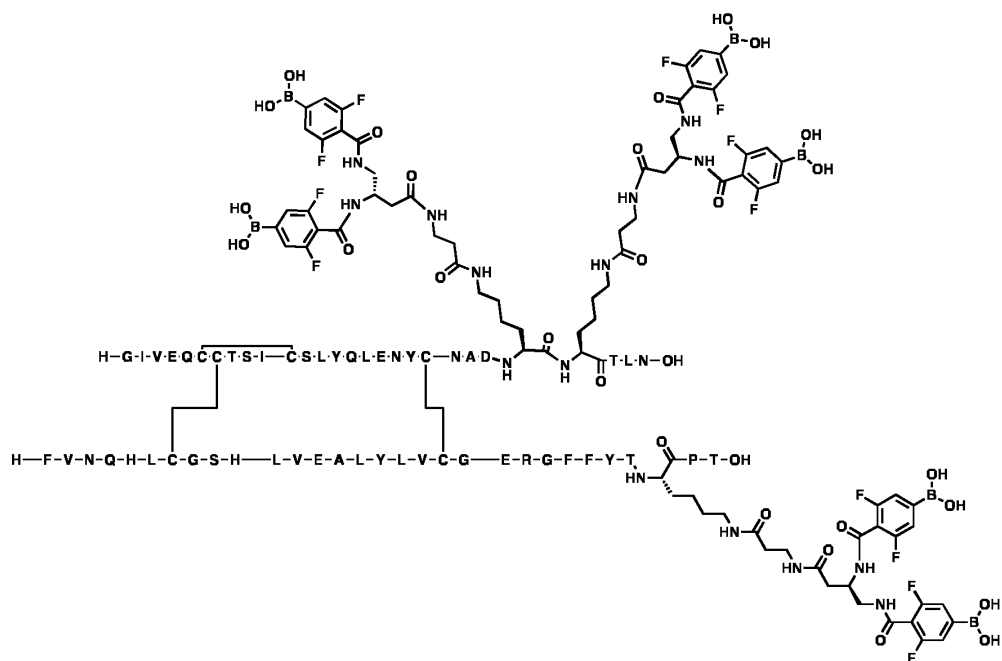
Example 620:



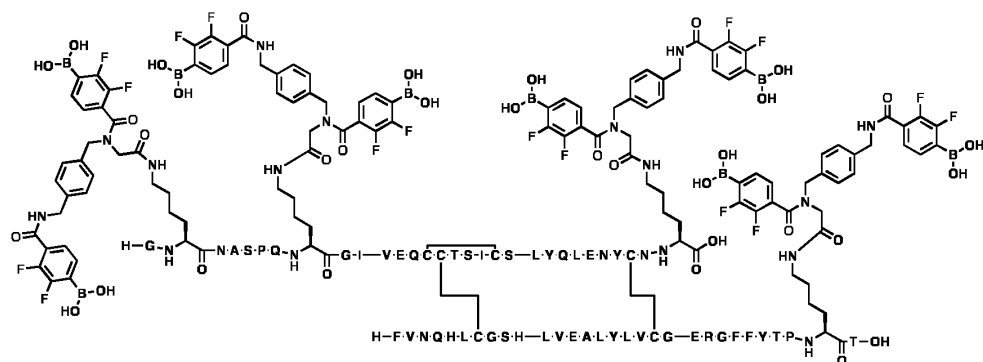
Example 621:



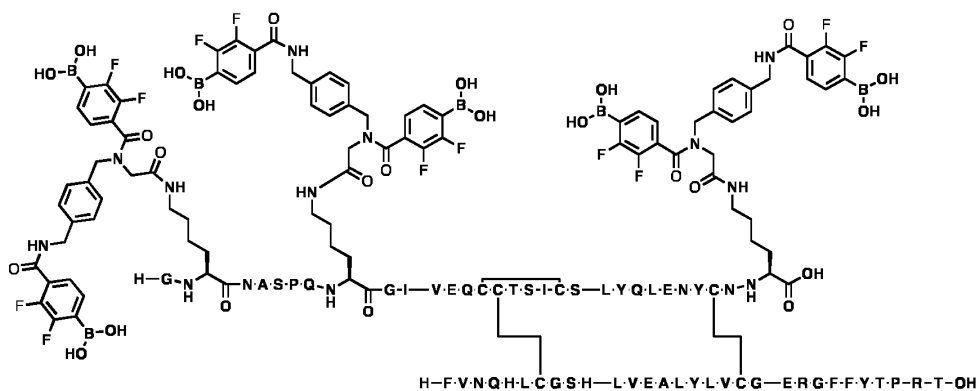
Example 622:



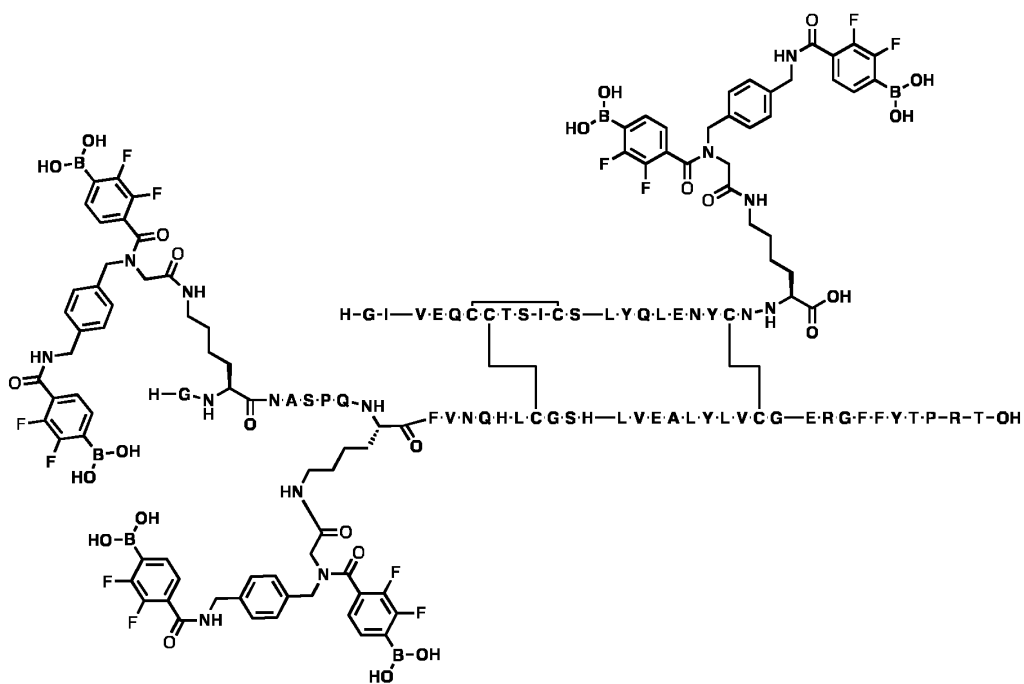
Example 623:

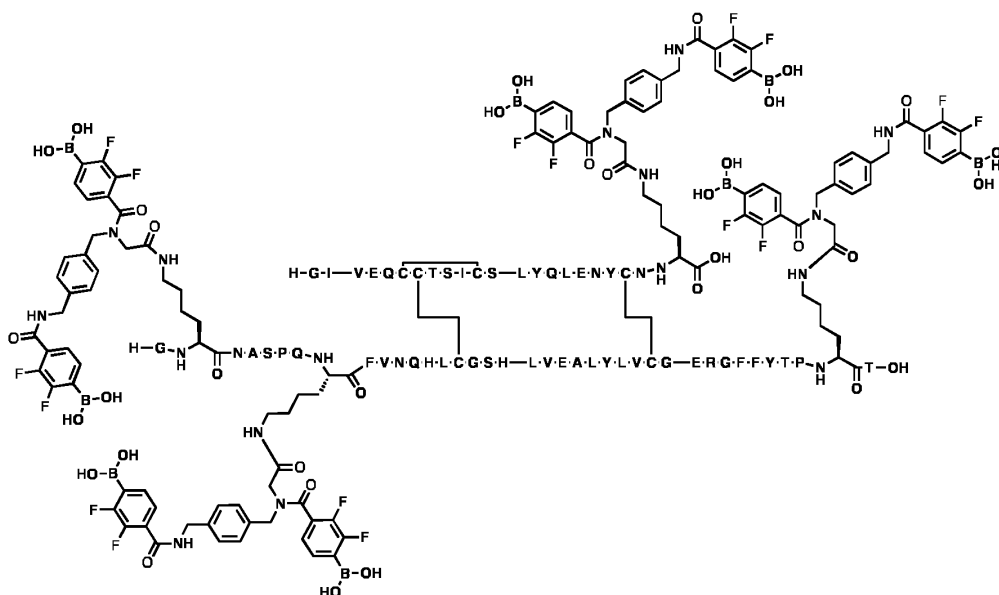


Example 624:

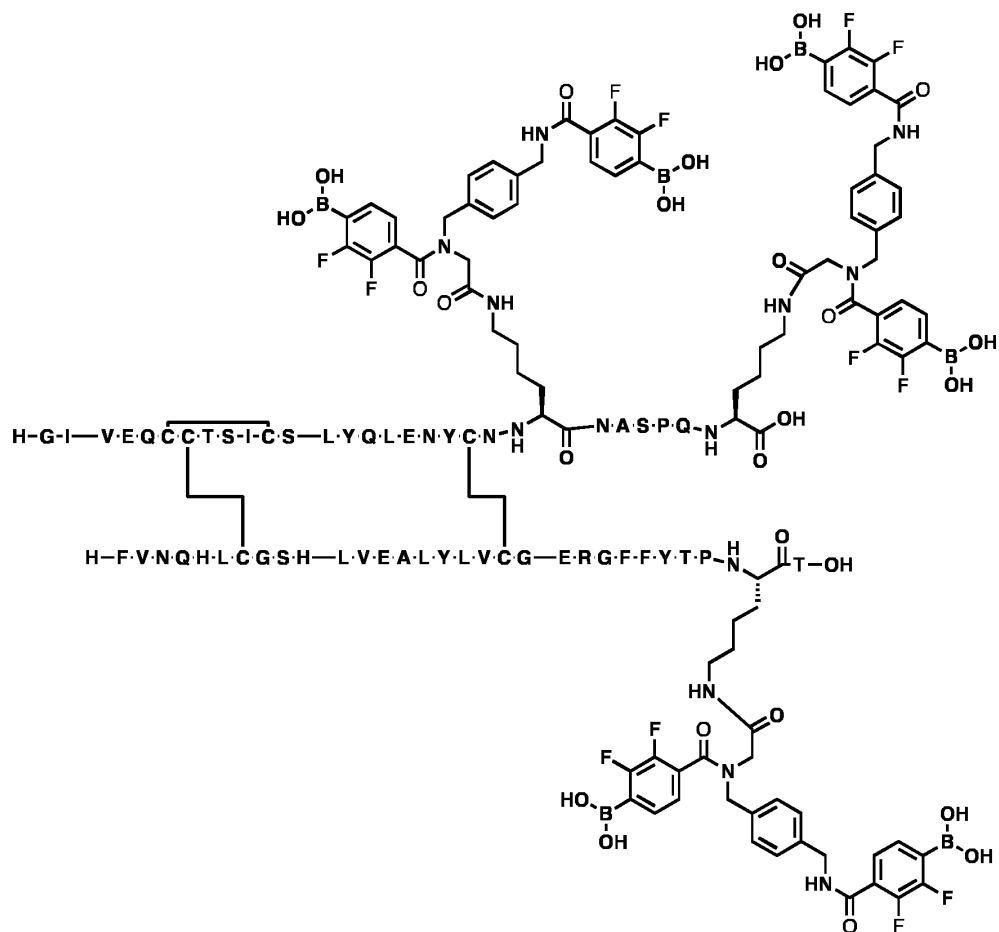


Example 625:

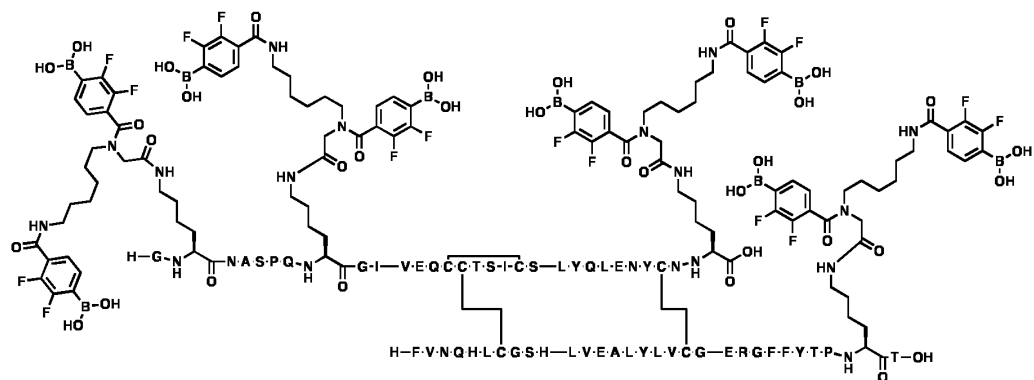


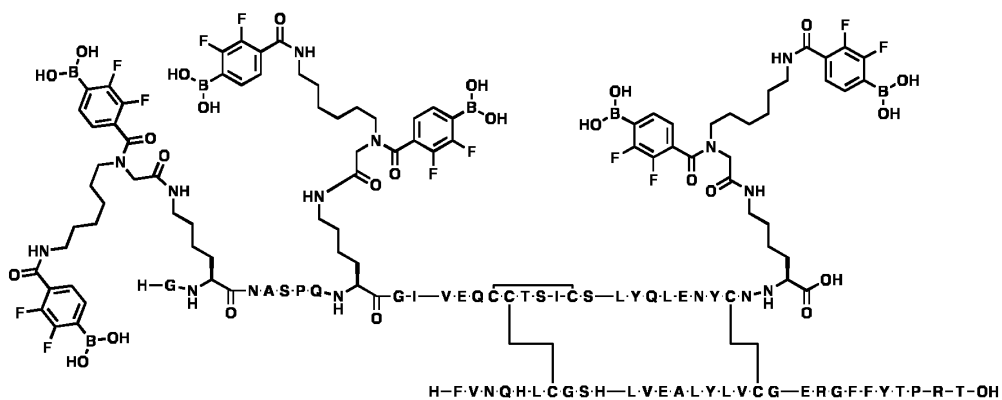
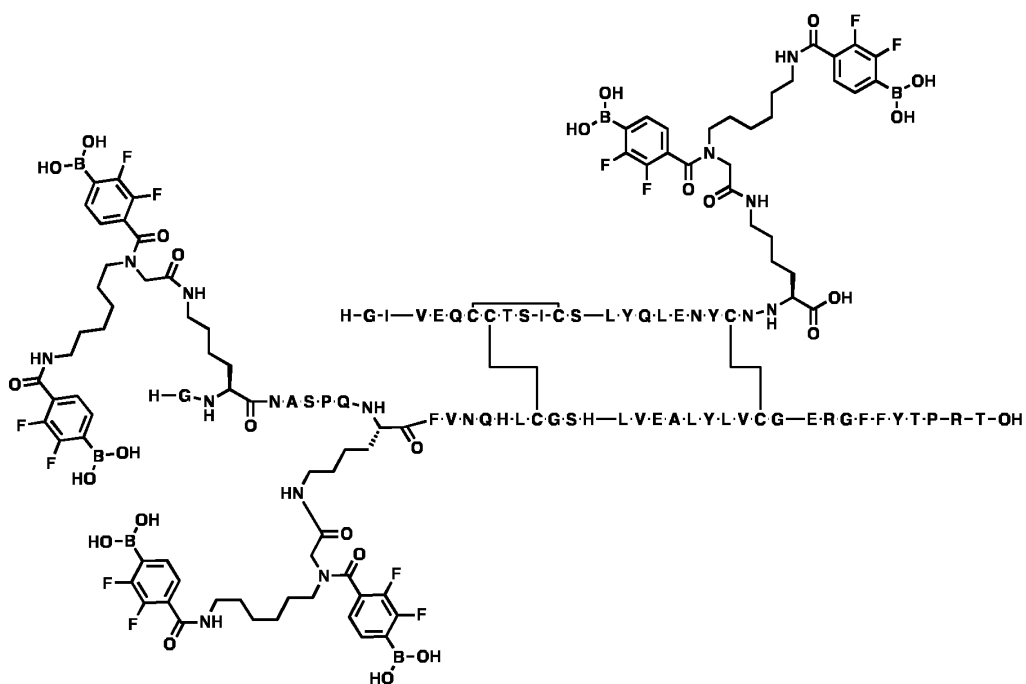
Example 626:

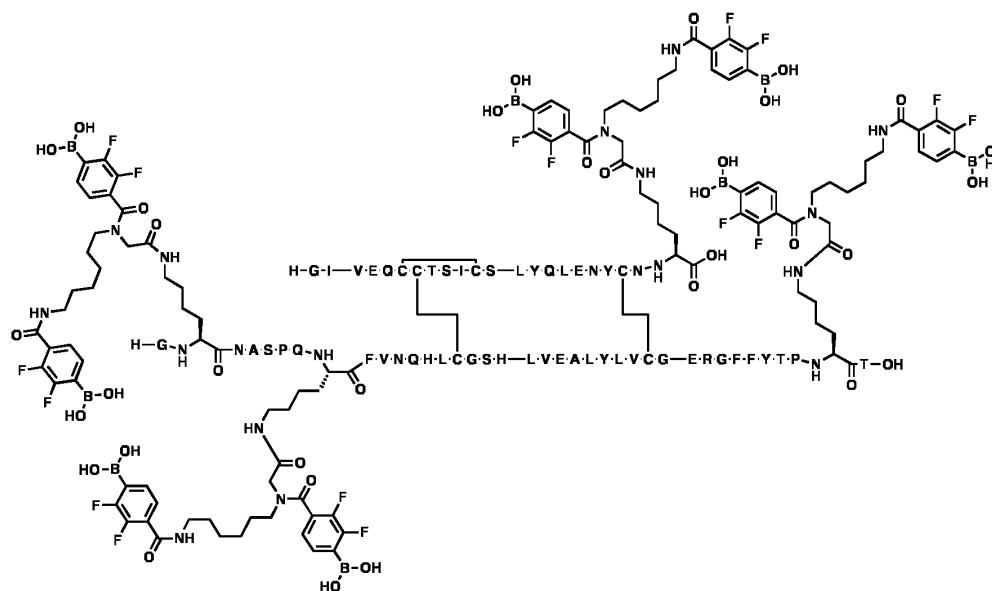
Example 627:



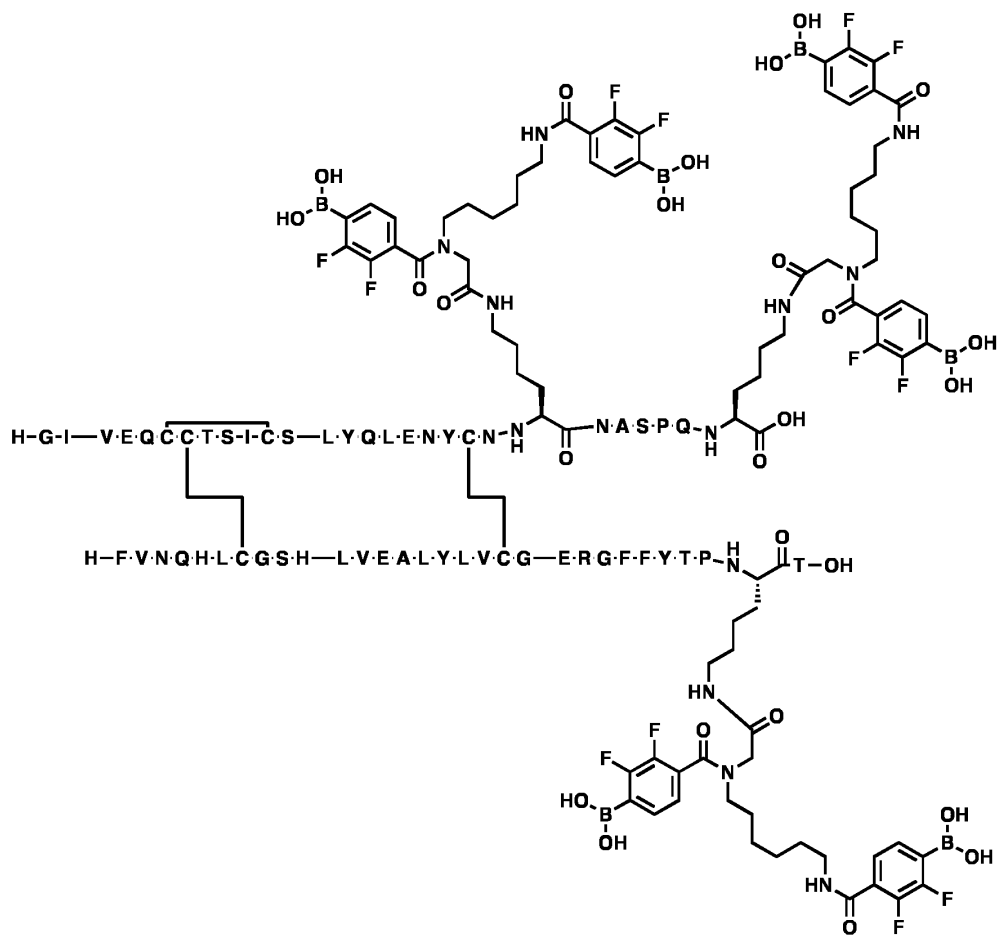
Example 628:



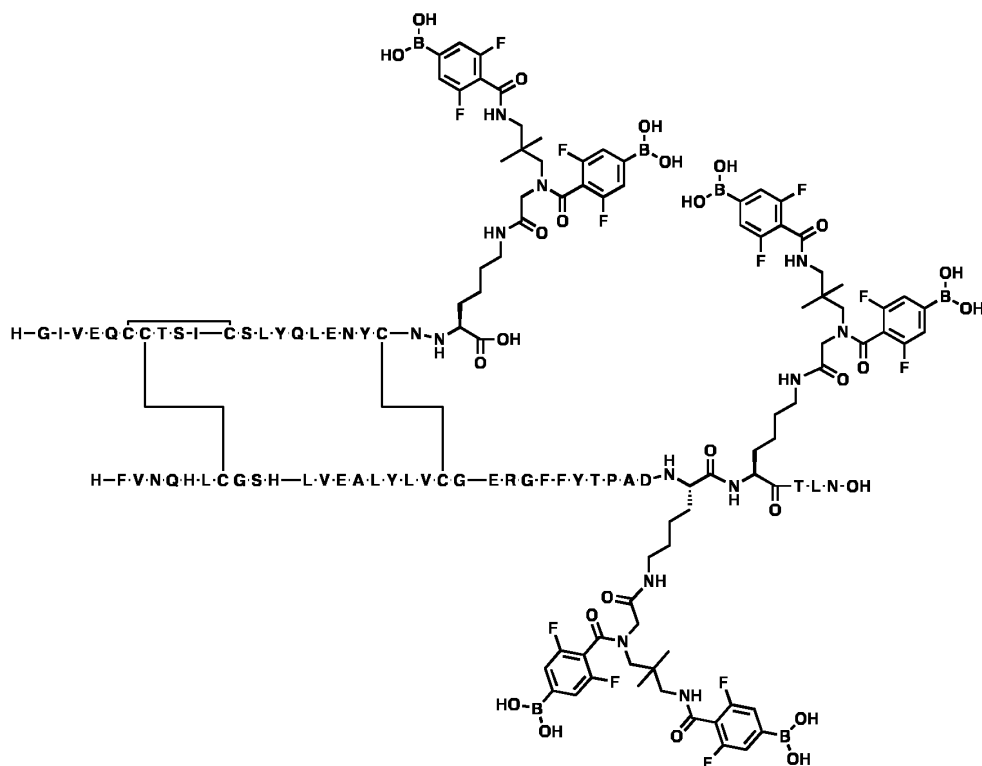
Example 629:**Example 630:**

Example 631:

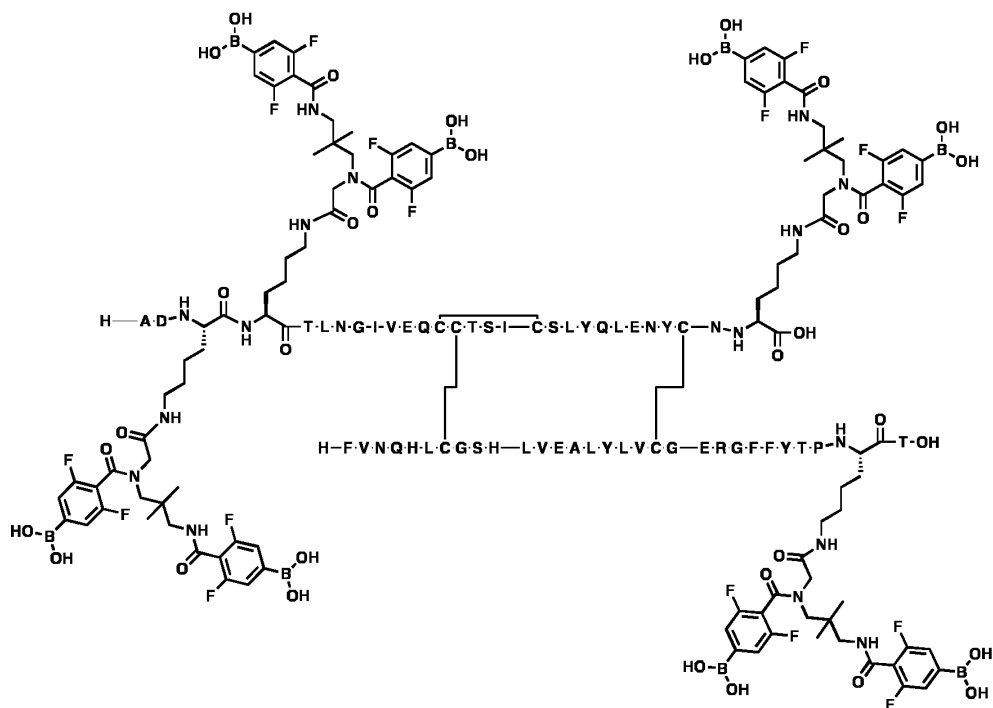
Example 632:



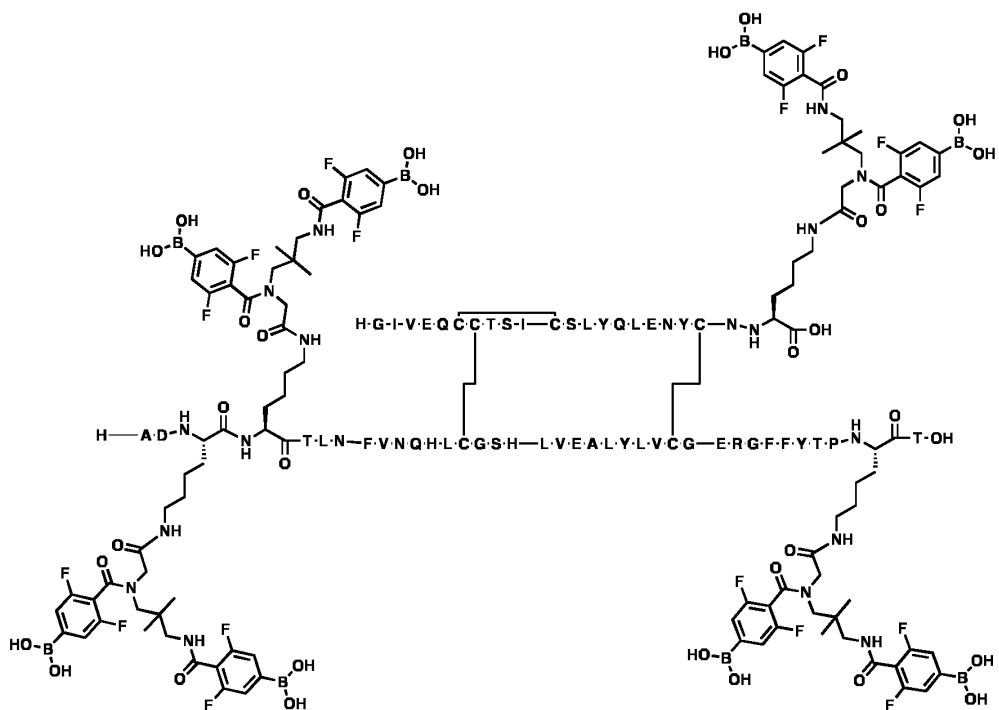
Example 633:



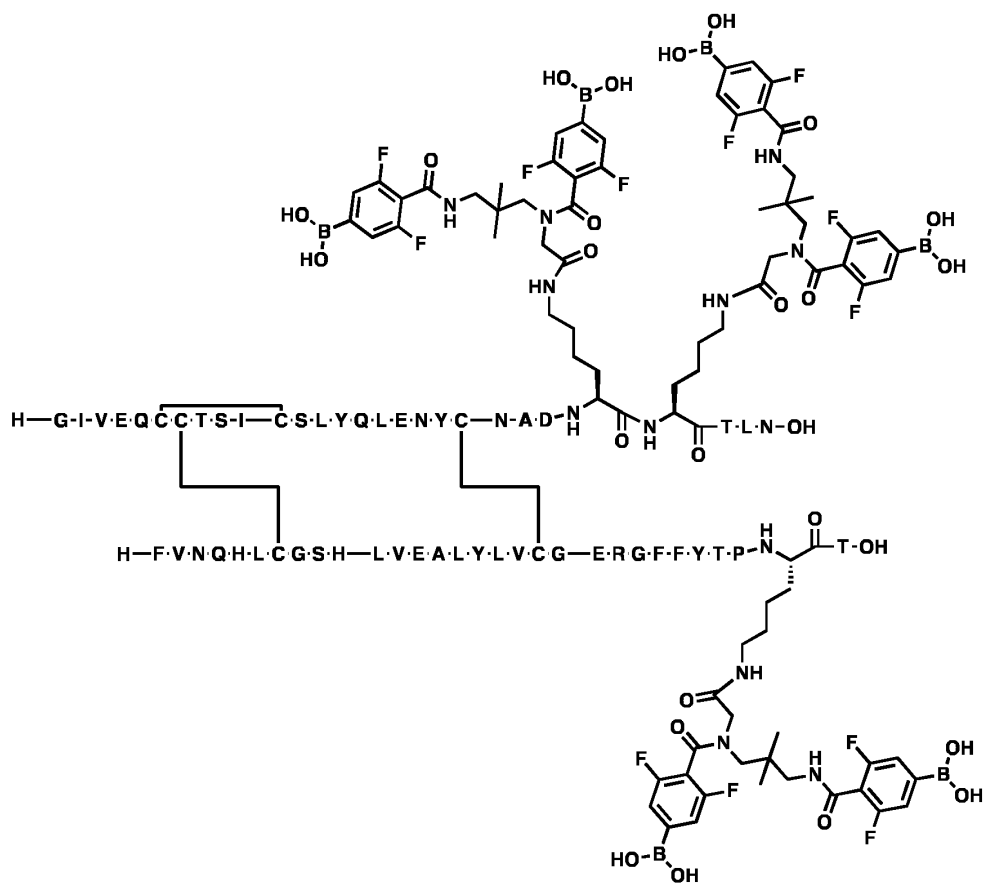
Example 634:

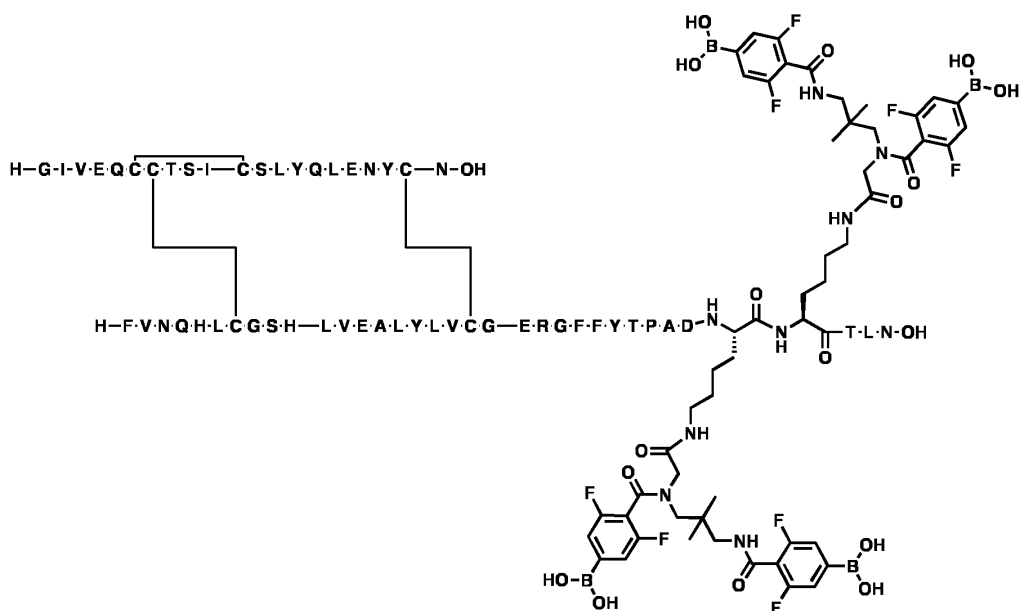


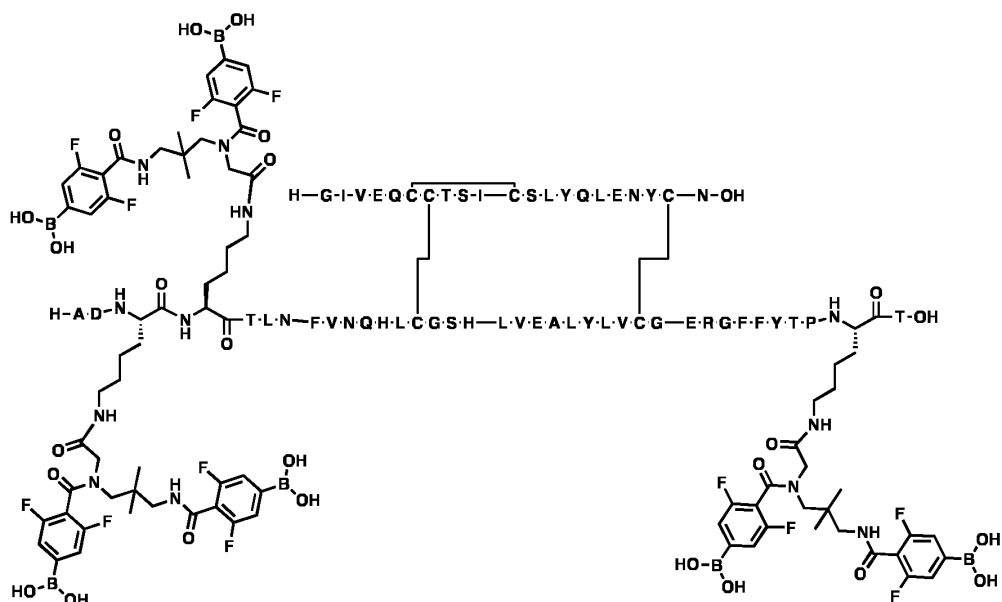
Example 635:

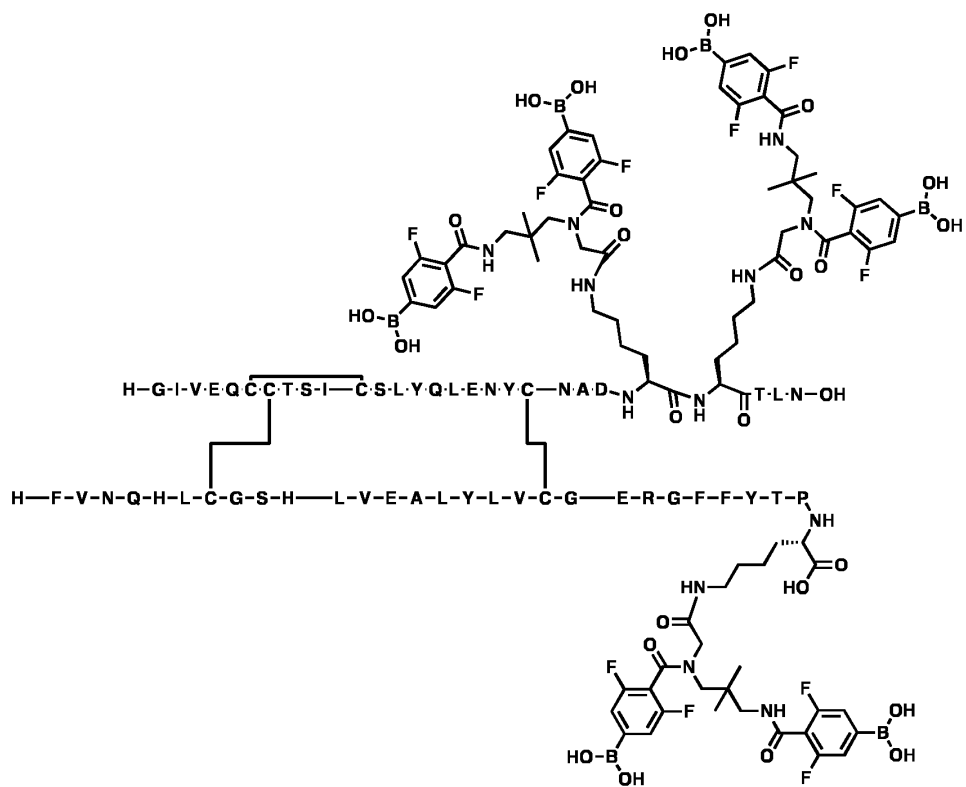


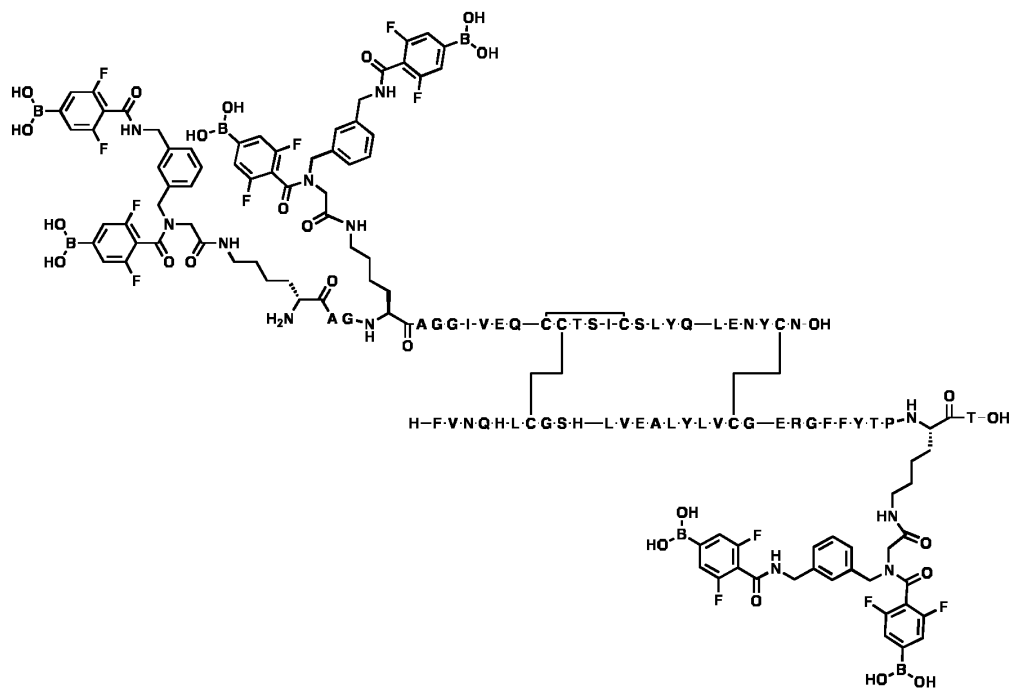
Example 636:

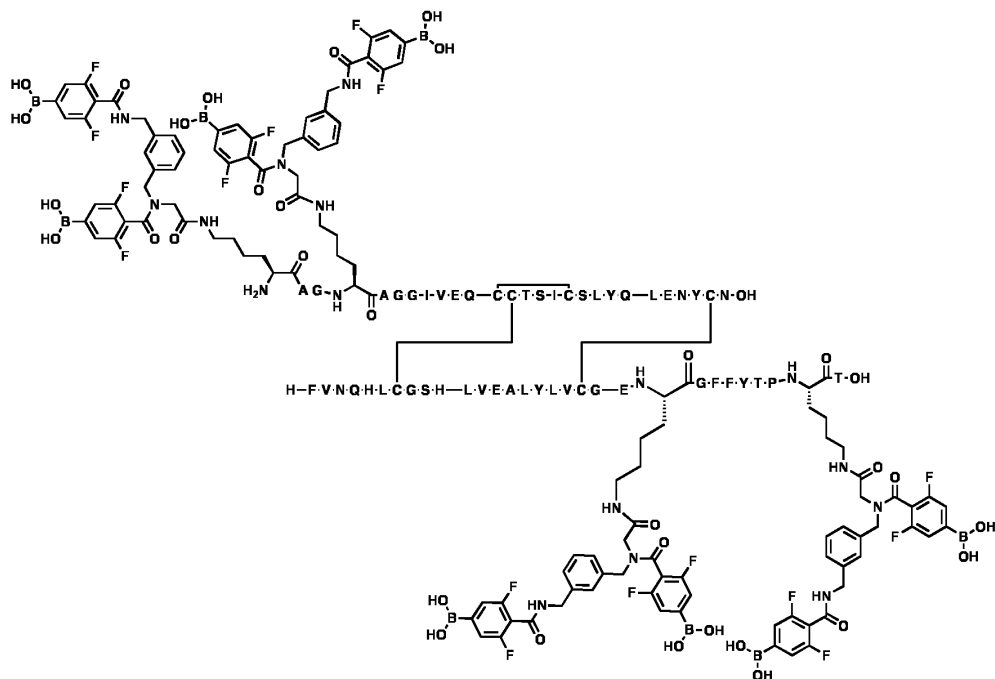
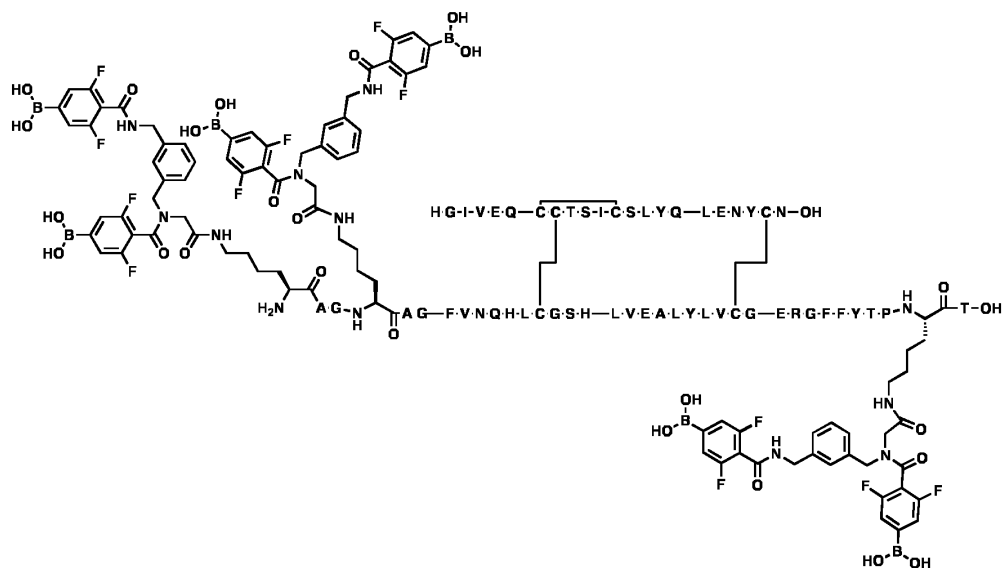


[illegible]

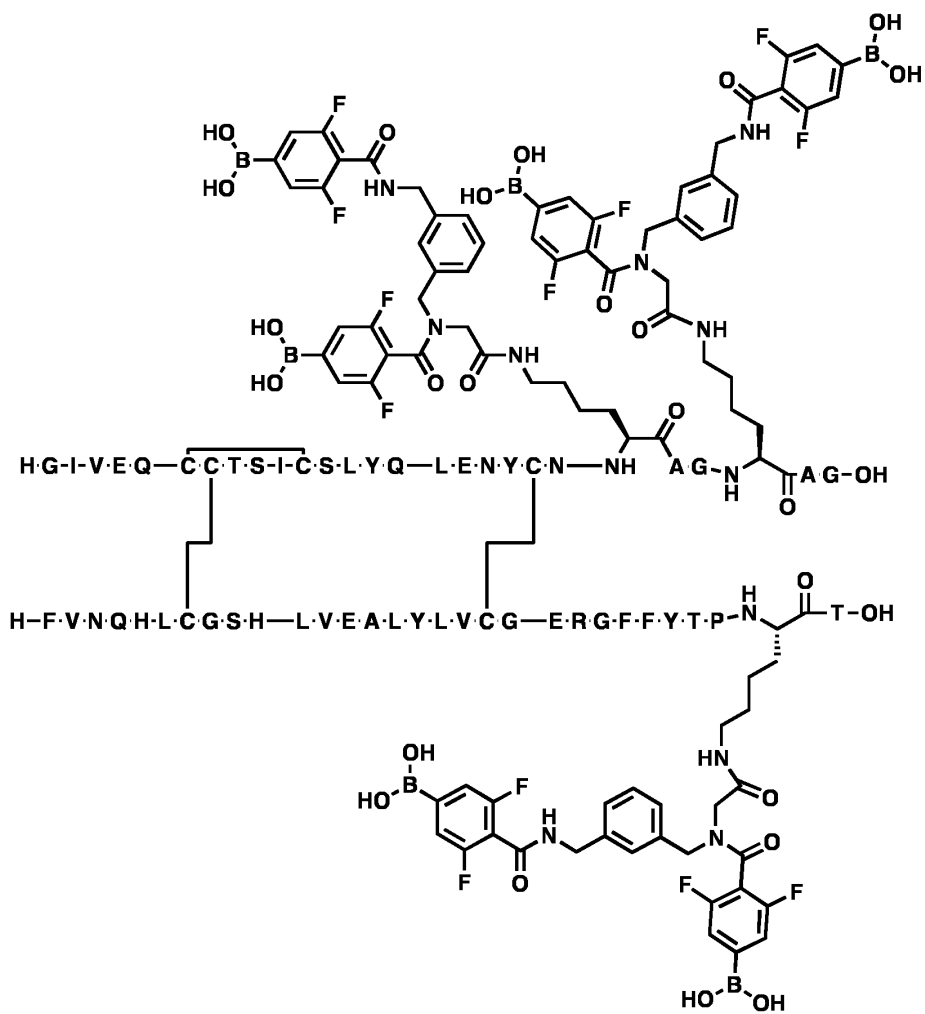
Example 639:

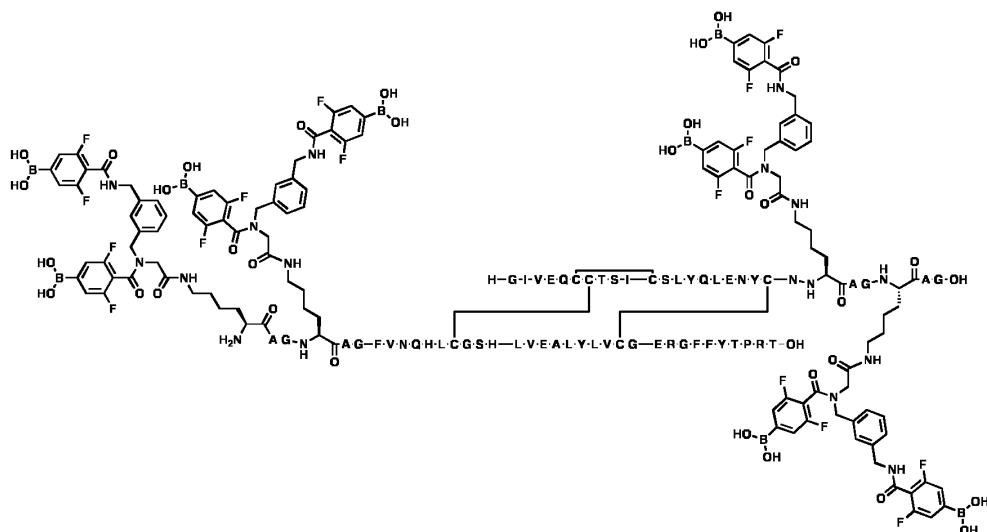
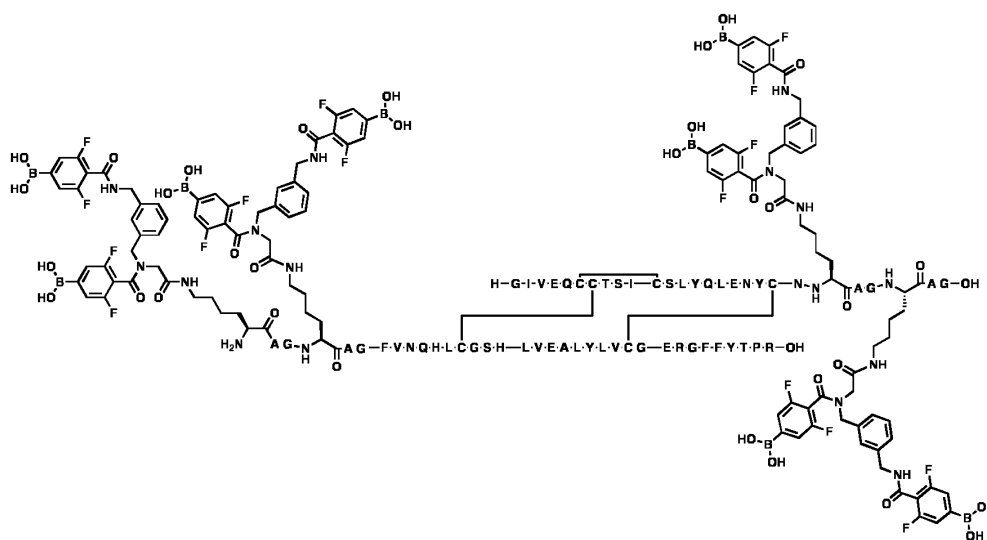
Example 640:

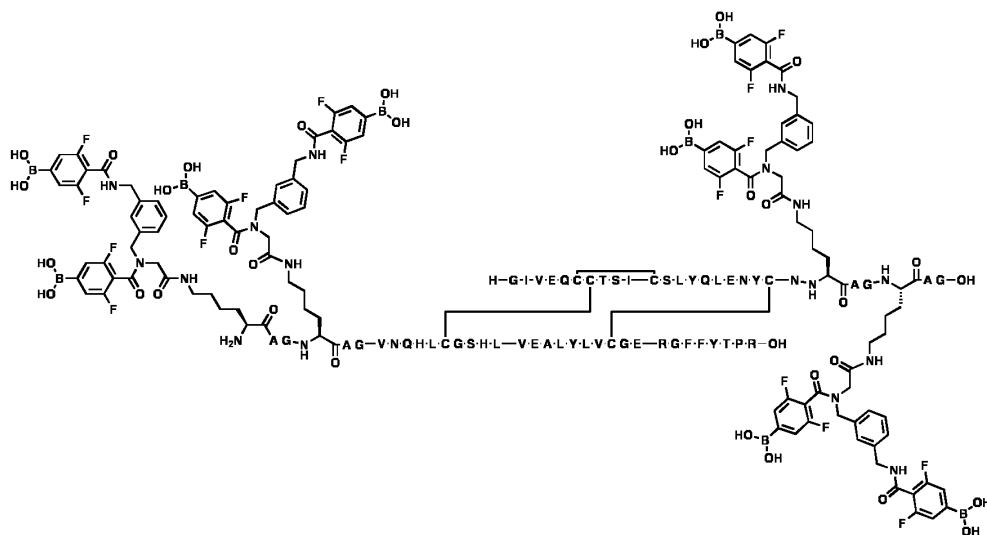
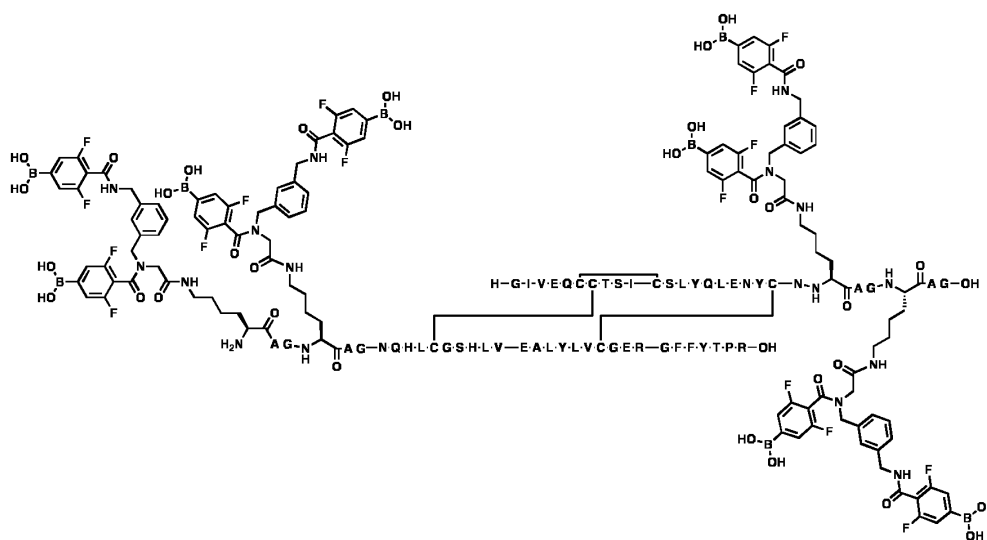
Example 641:

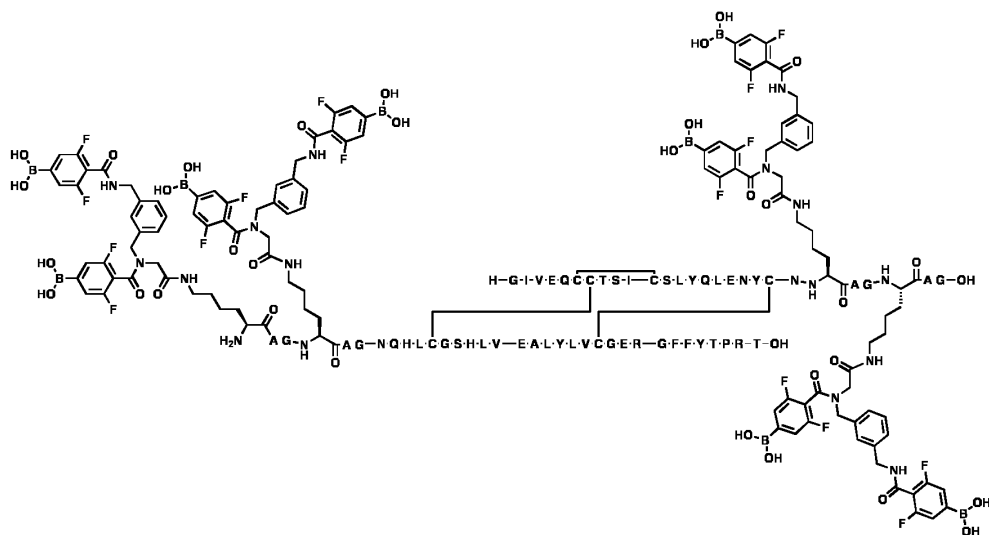
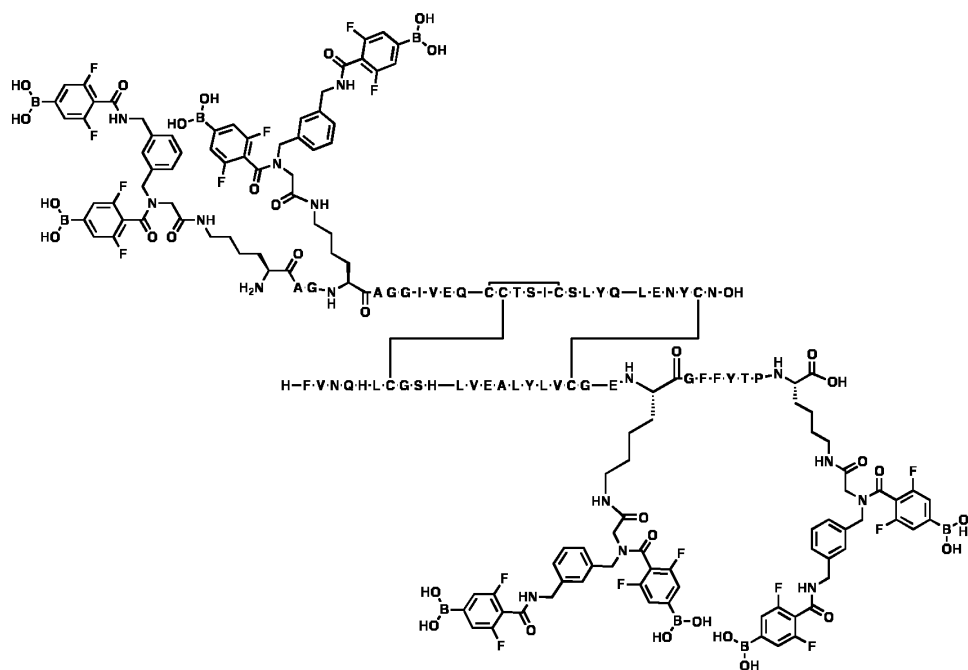
Example 642:**Example 643:**

Example 644:

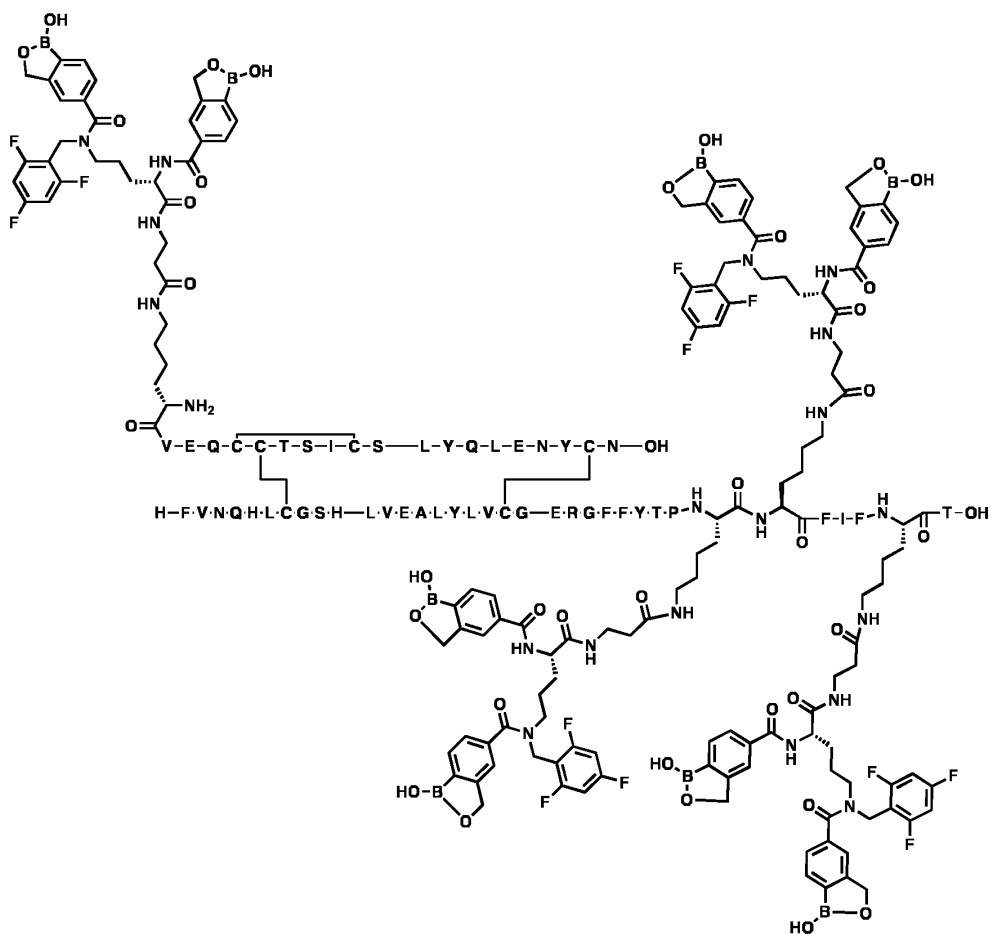


Example 645:**Example 646:**

Example 647:**Example 648:**

Example 649:**Example 650:**

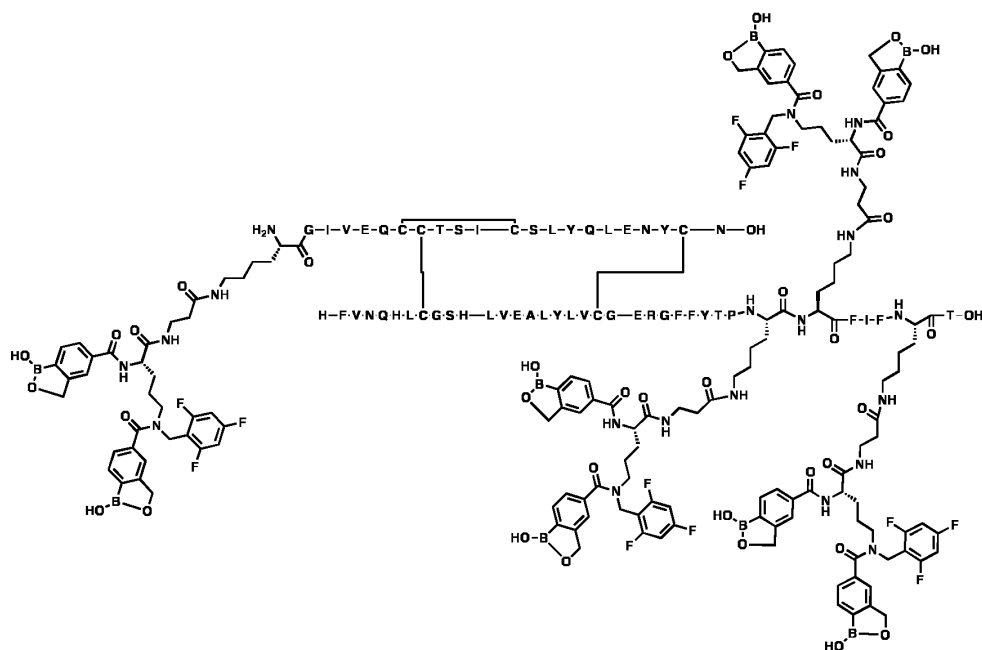
Example 651:



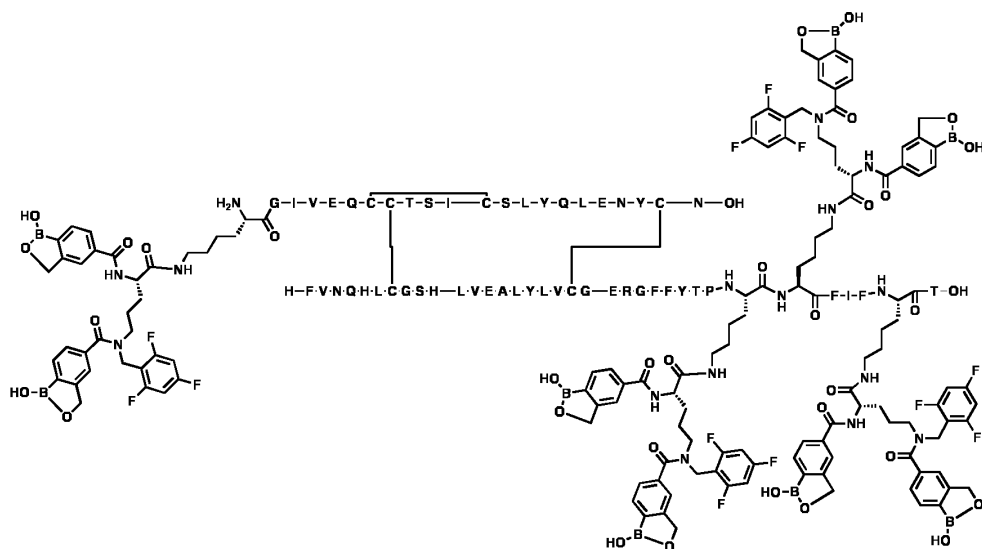
The chemical structure represents the 120 kDa subunit of the 20S proteasome core particle. It is a large, complex polypeptide chain with multiple domains. The structure is shown in a ribbon representation with various side chains and functional groups highlighted. Key features include:

- Fluorinated Aromatic Rings:** Several benzene rings with fluorine substituents are present, particularly in the upper and lower regions of the structure.
- Boron-containing Groups:** Multiple boron atoms are shown, often as part of a boronate ester or boronic acid group (B(OH)₂).
- Central Amino Acid Sequence:** A specific segment of the polypeptide chain is highlighted with a red box and labeled with the sequence: G-V-E-Q-C-C-T-S-I-C-S-L-Y-Q-L-E-N-Y-C-N-OH.
- Side Chains and Functional Groups:** Various side chains are shown, including amide groups, hydroxyl groups, and fluorinated aromatic rings.

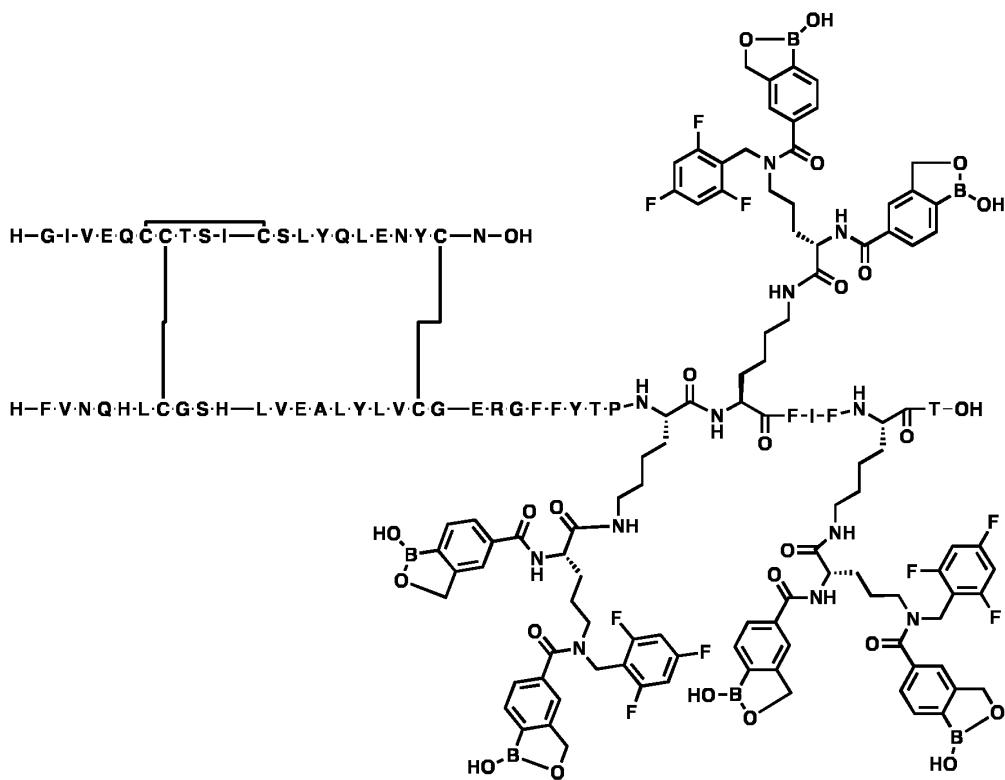
Example 653:

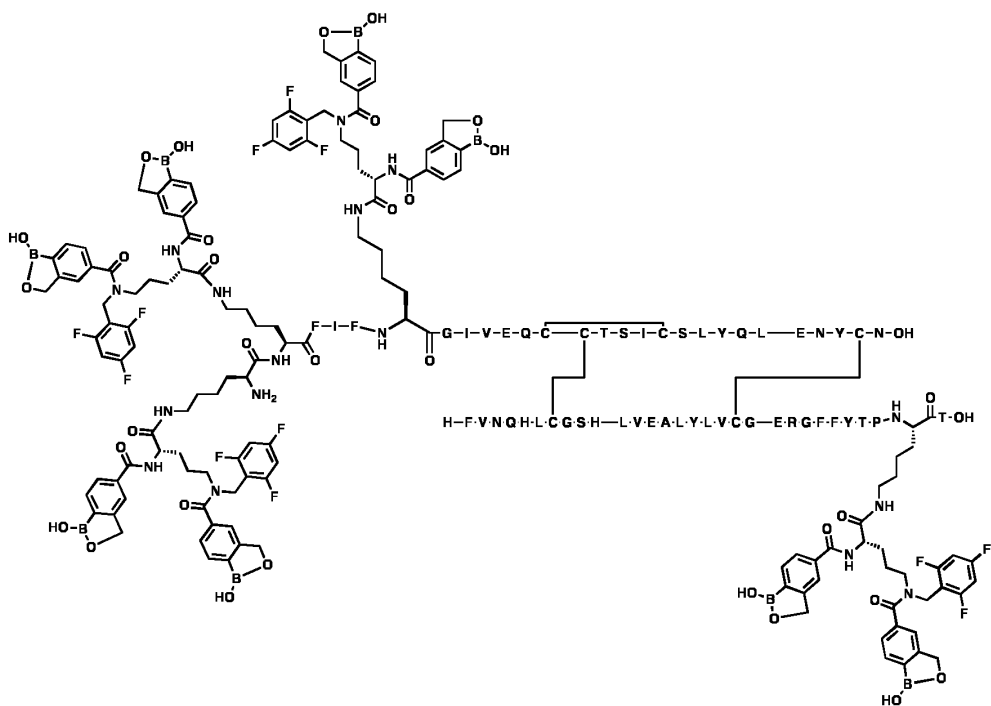


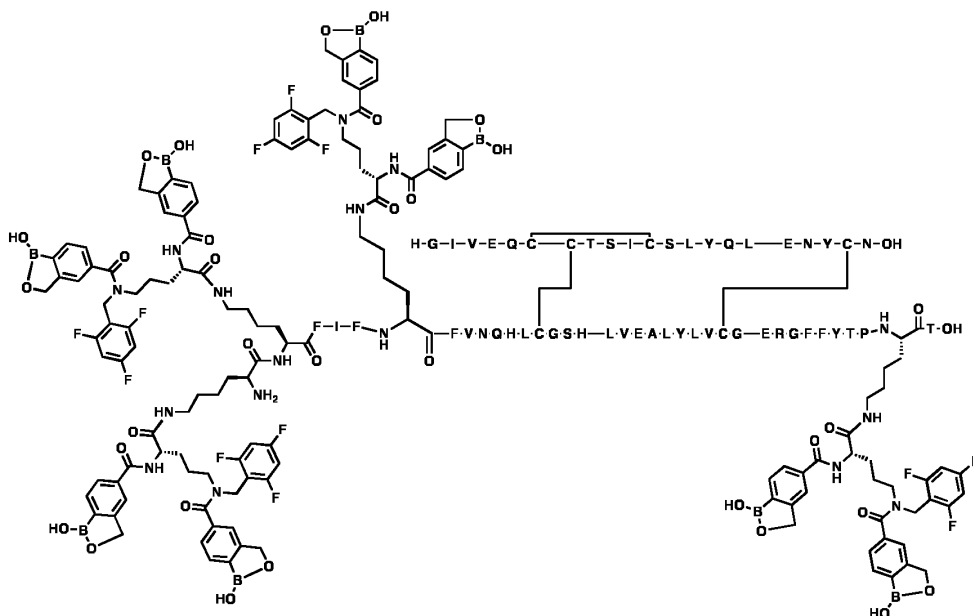
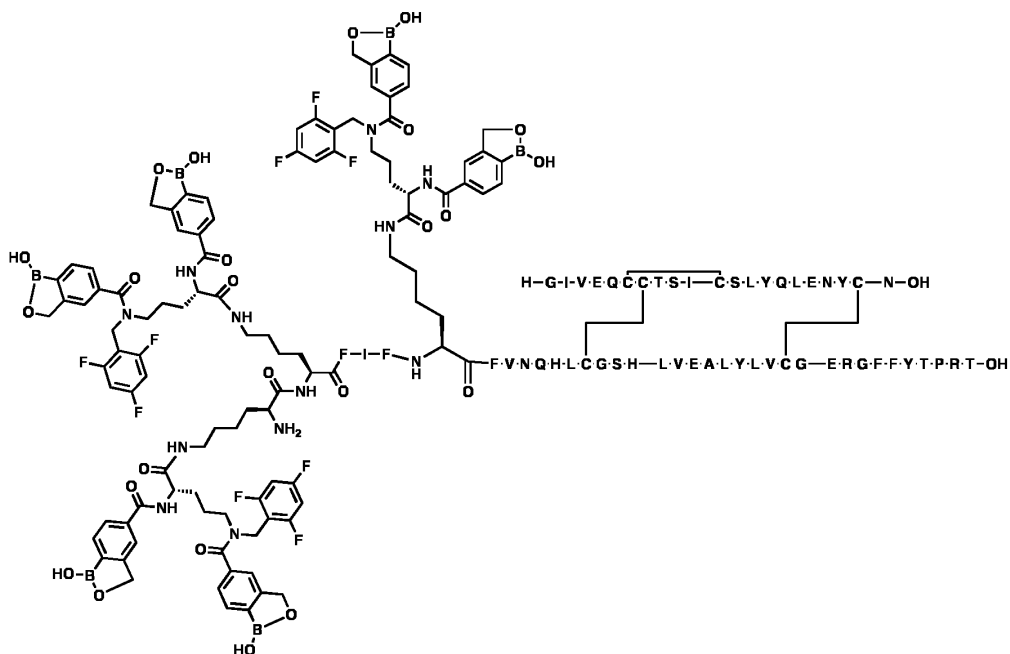
Example 654:

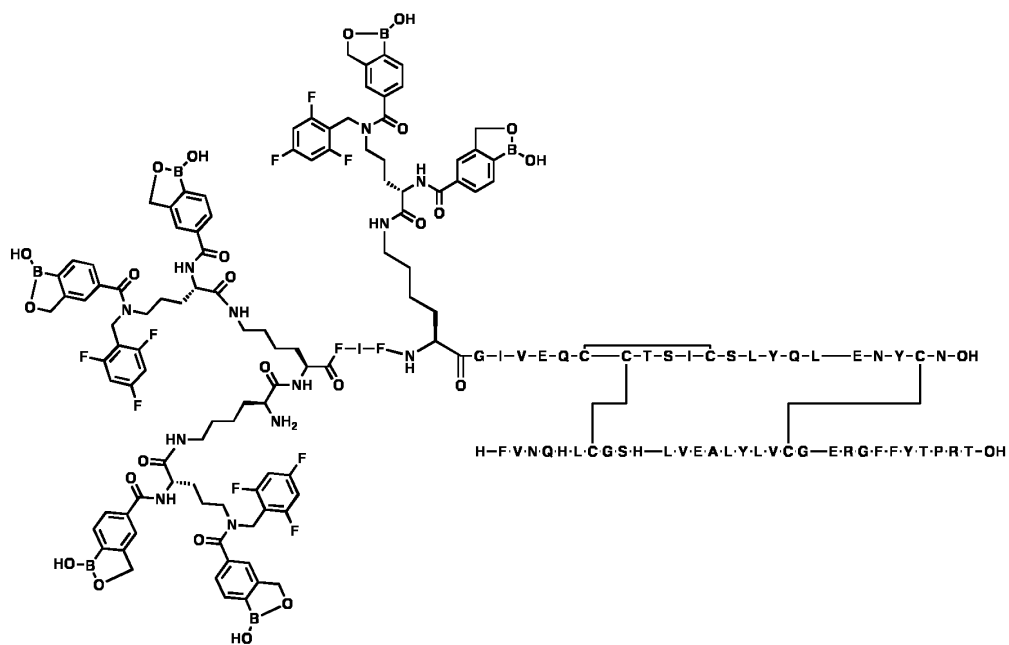


Example 655:

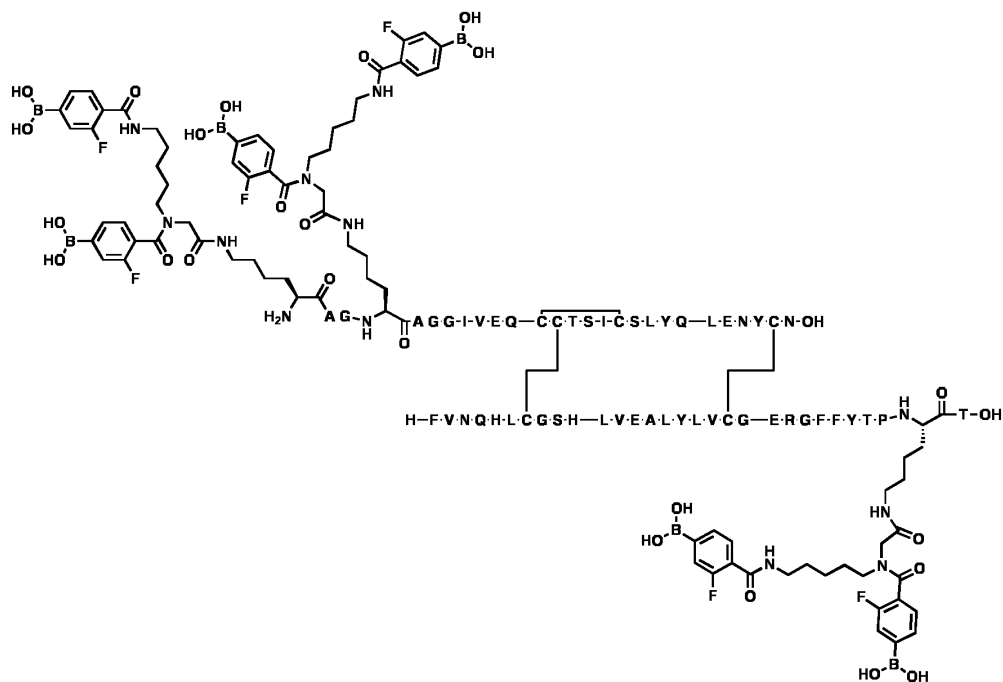


Example 656:

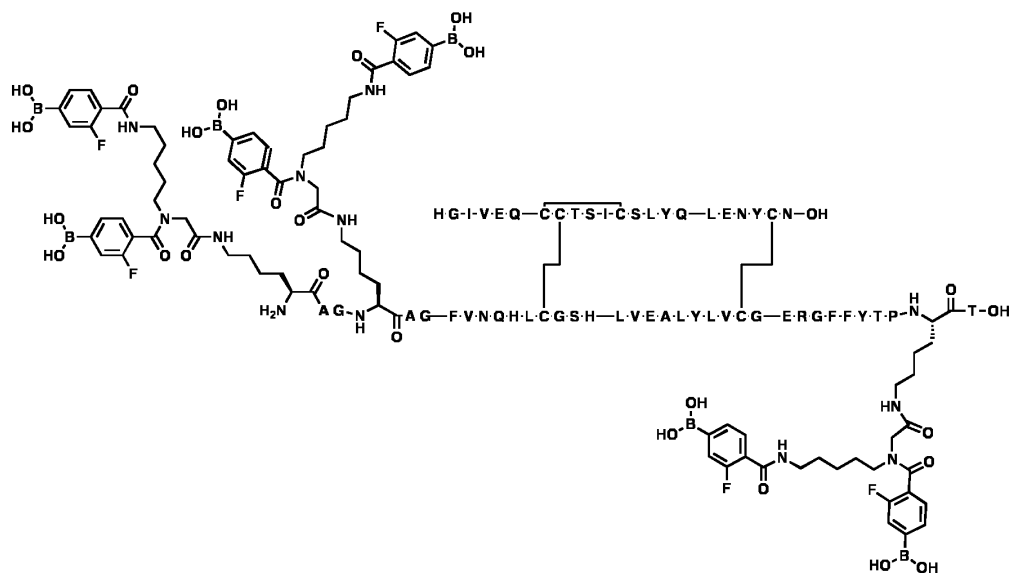
Example 657:**Example 658:**

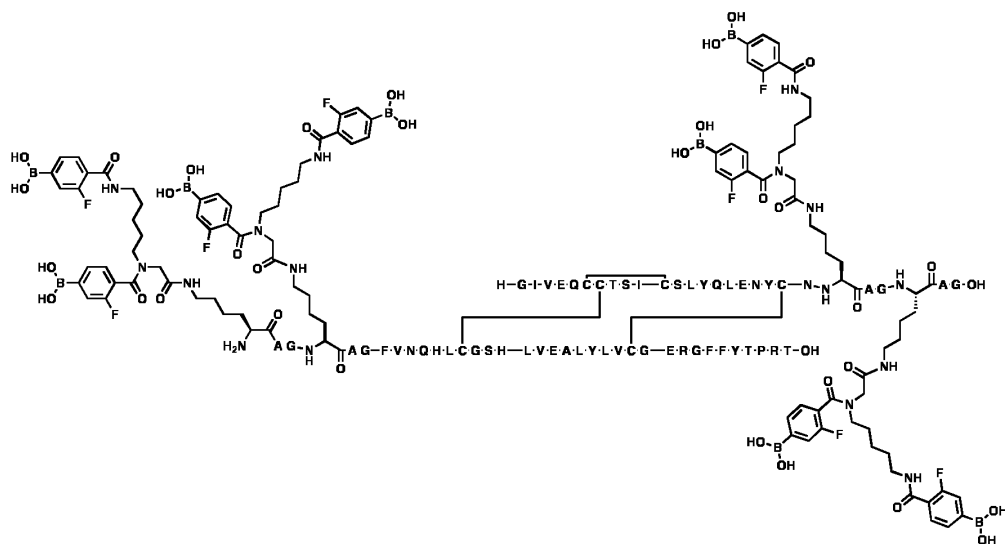
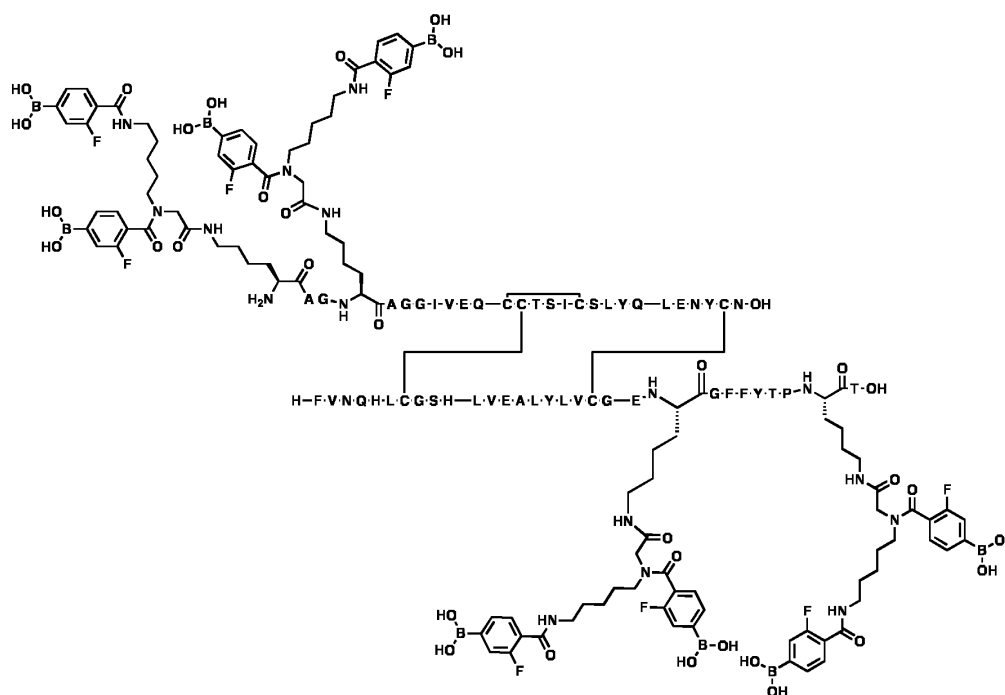
Example 659:

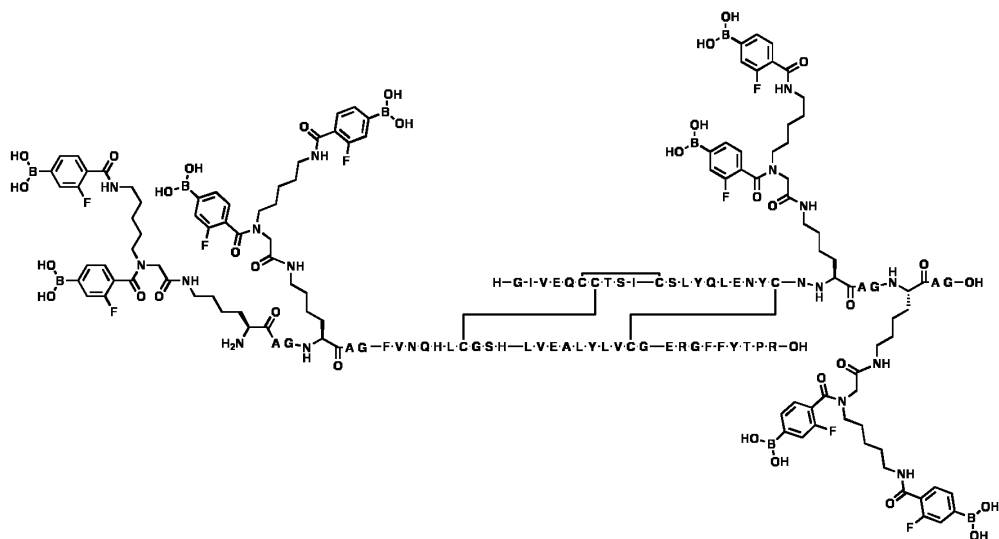
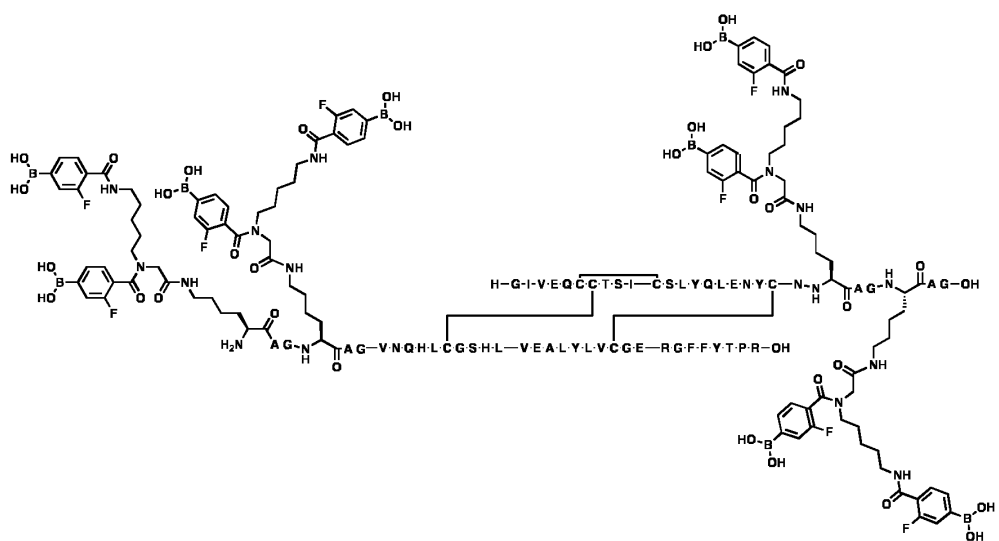
Example 660:



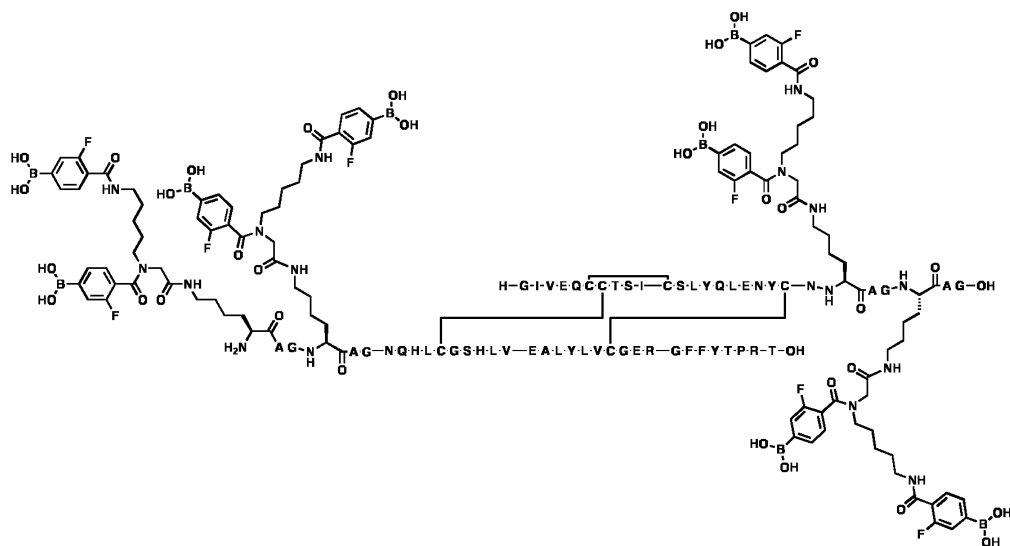
Example 661:



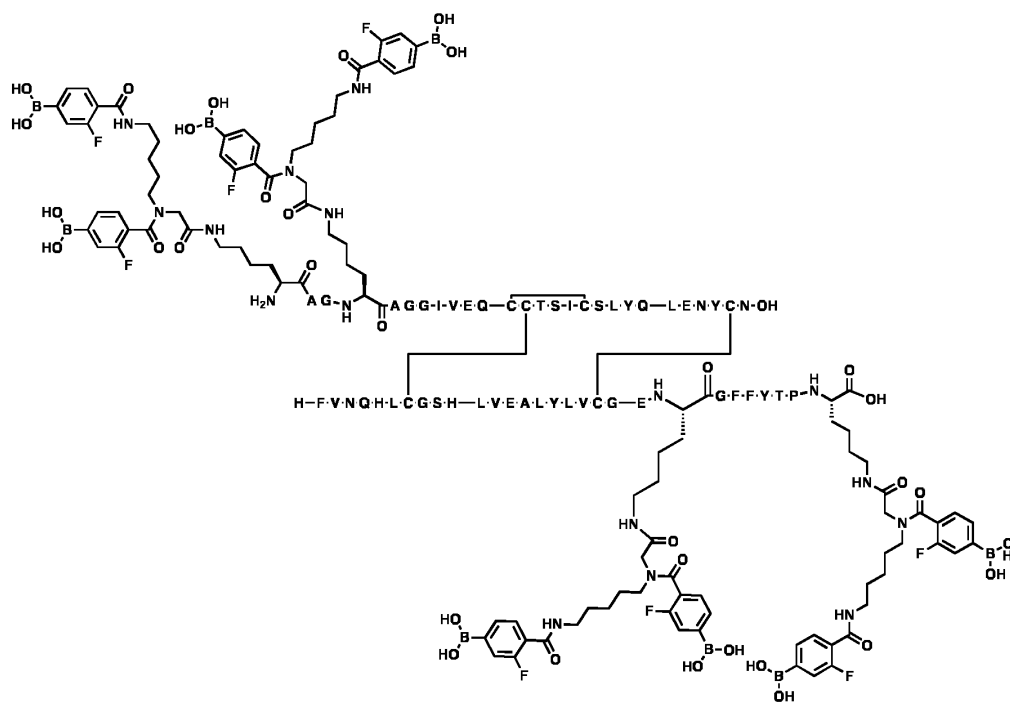
Example 662:**Example 663:**

Example 664:**Example 665:**

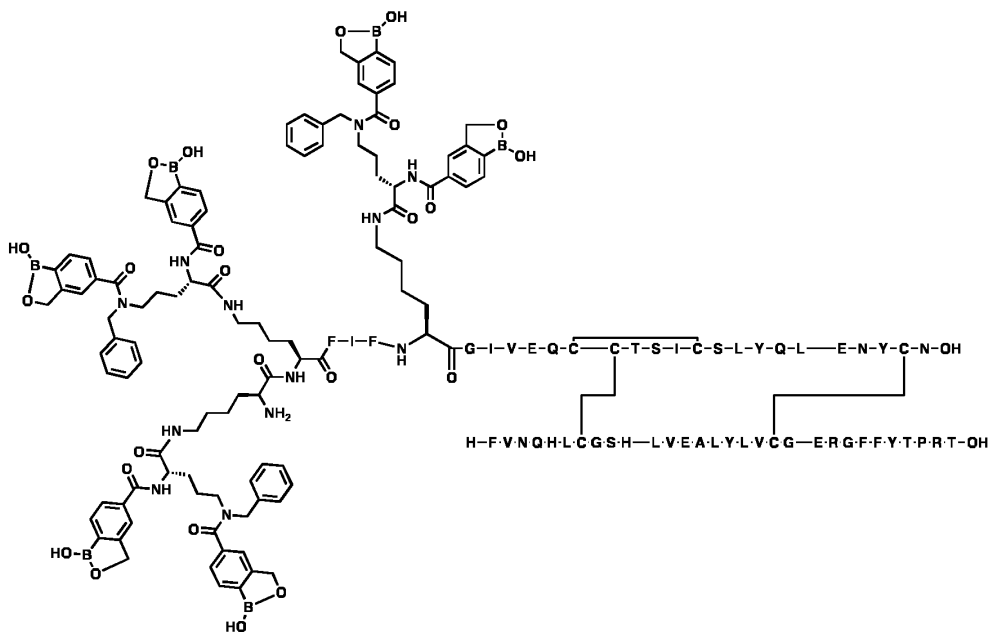
Example 666:



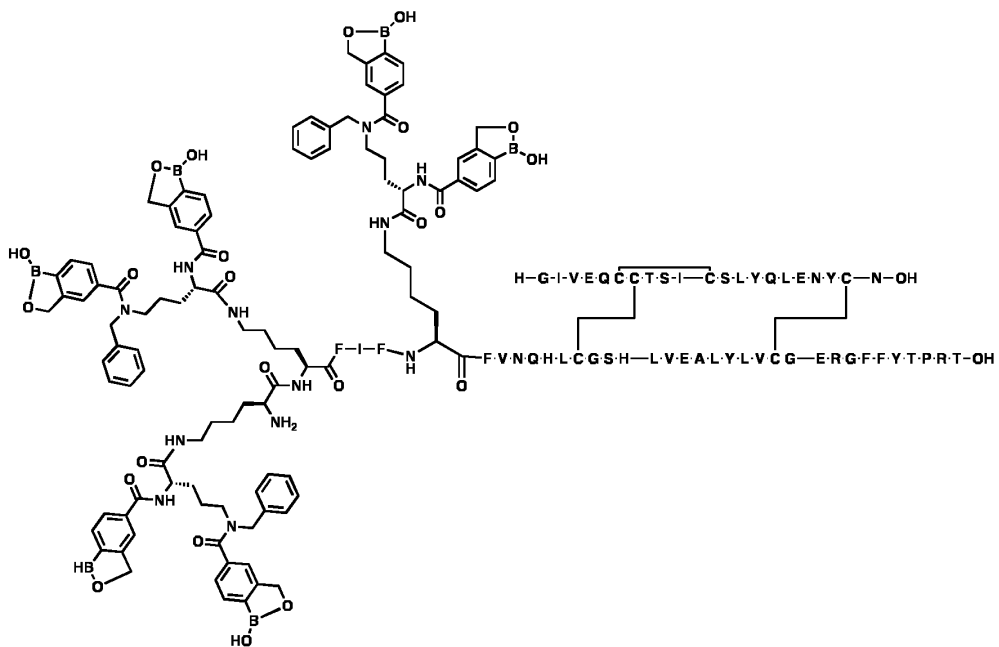
Example 667:

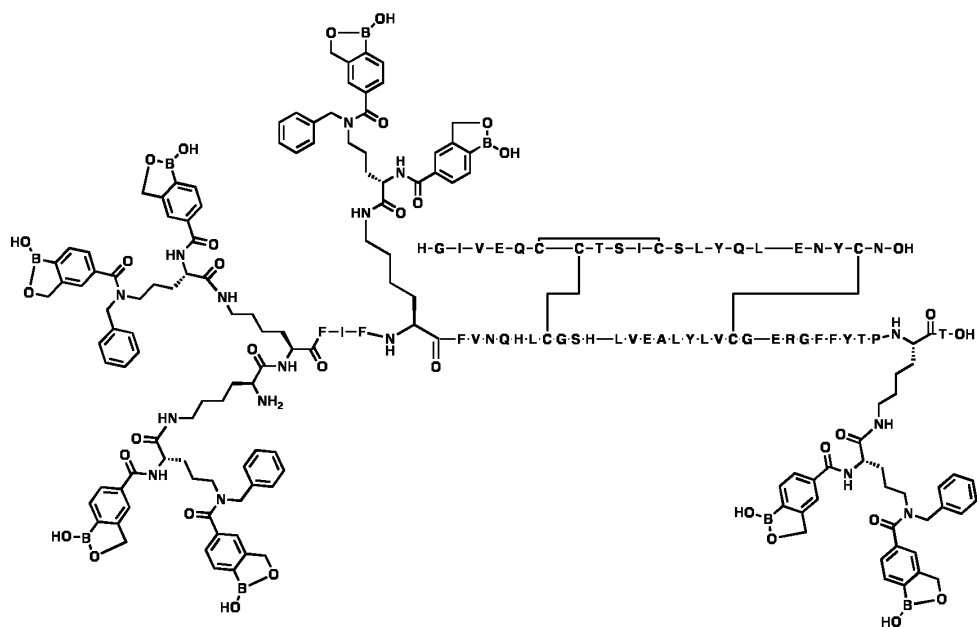


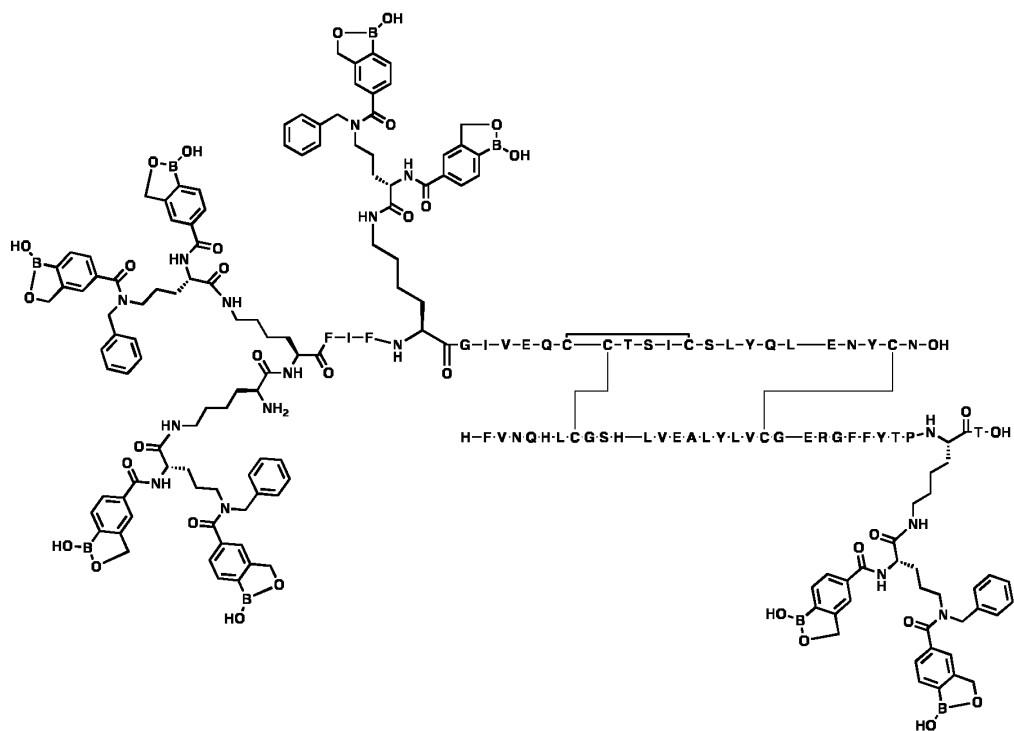
Example 668:



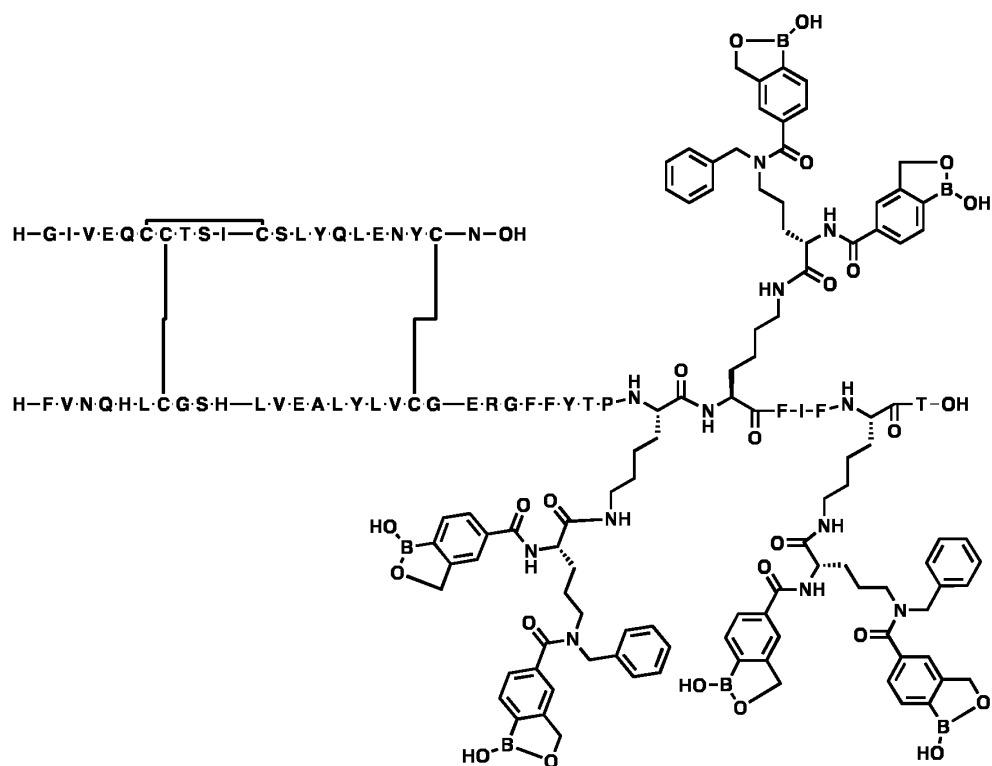
Example 669:

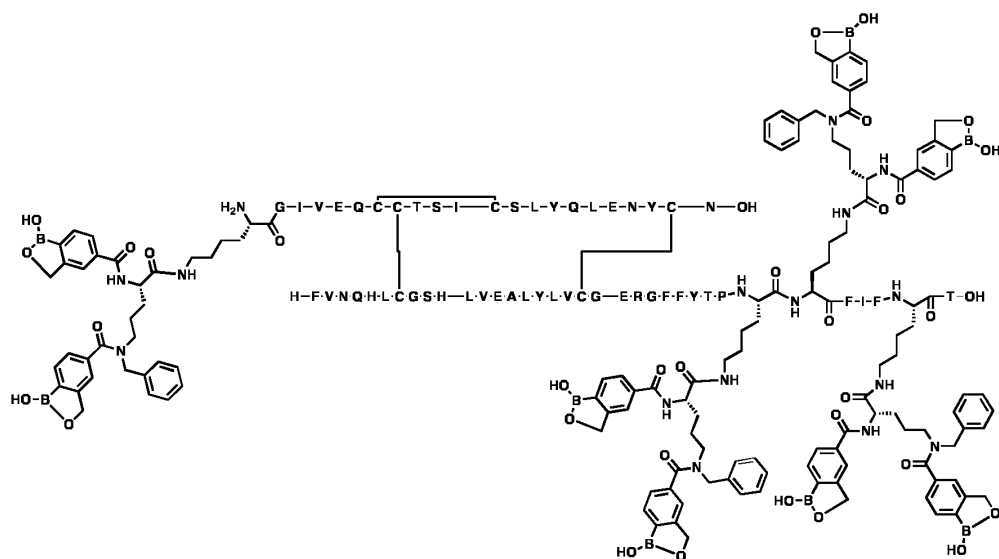
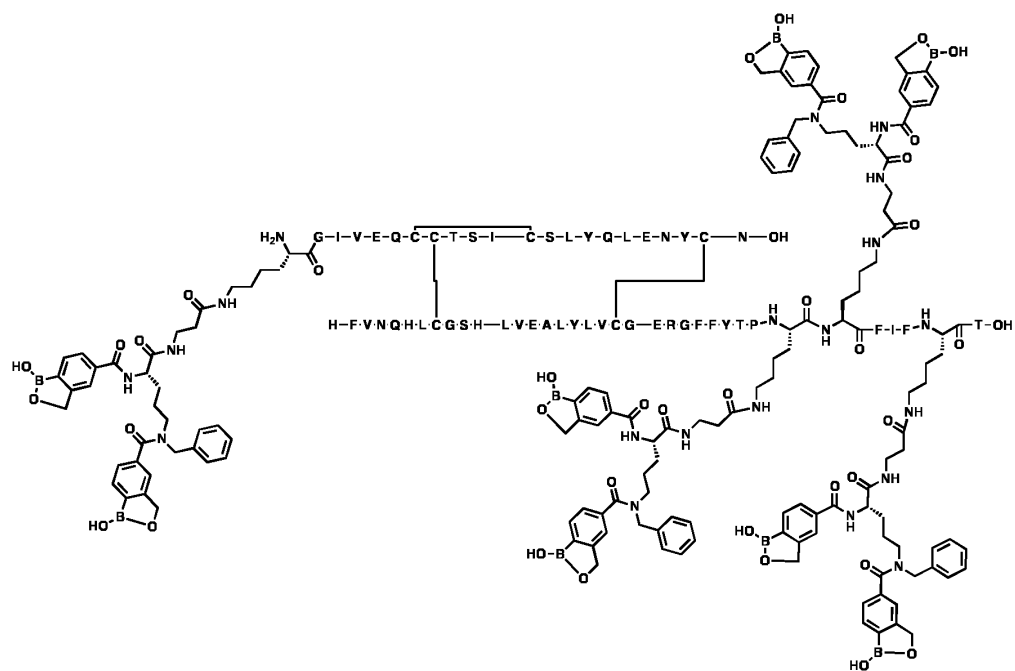


Example 670:

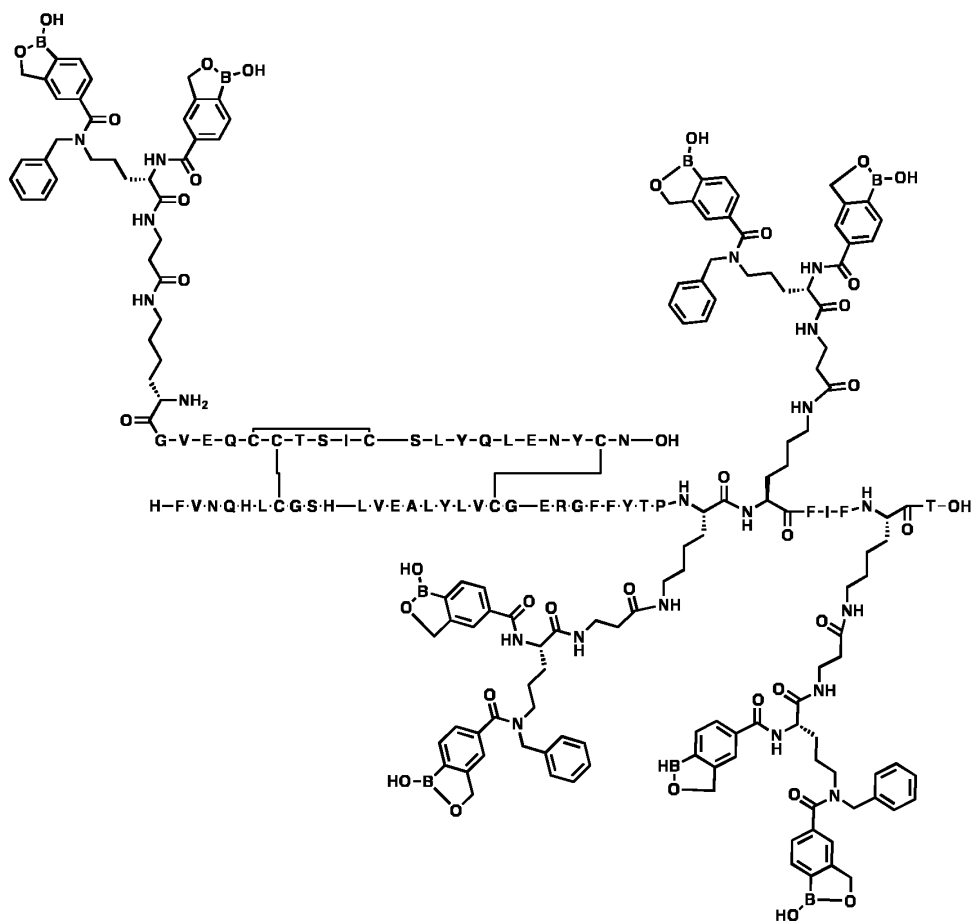
Example 671:

Example 672:

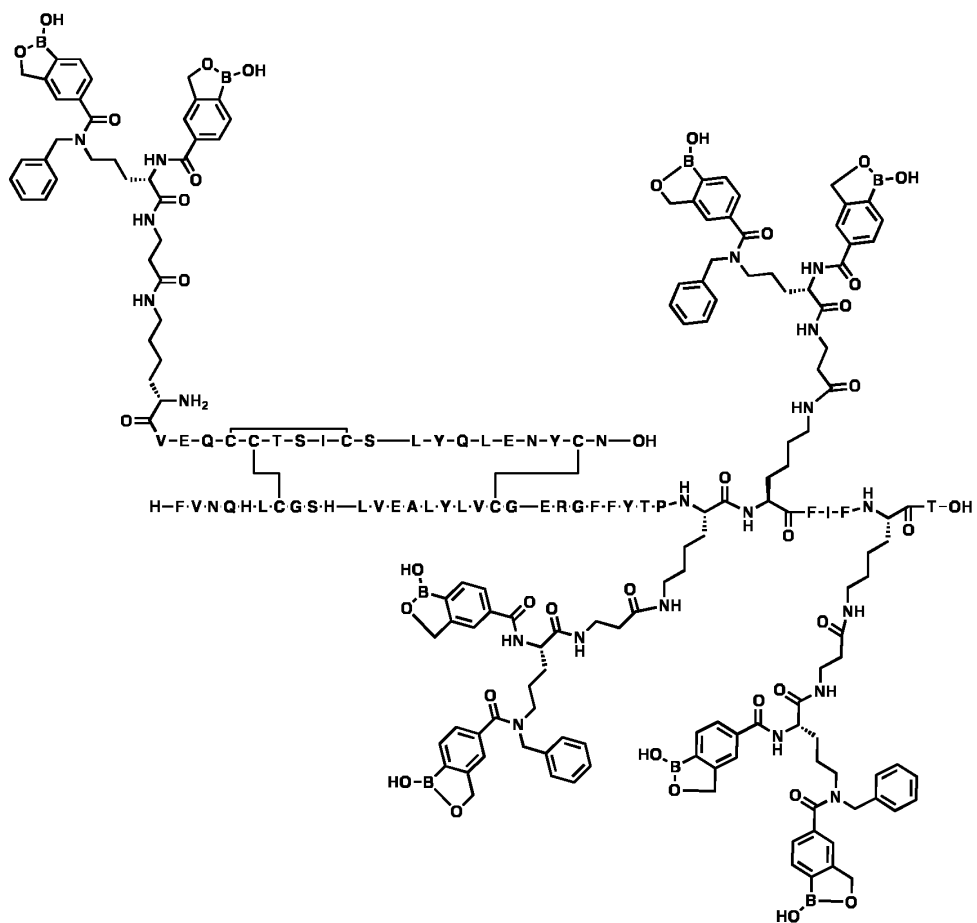


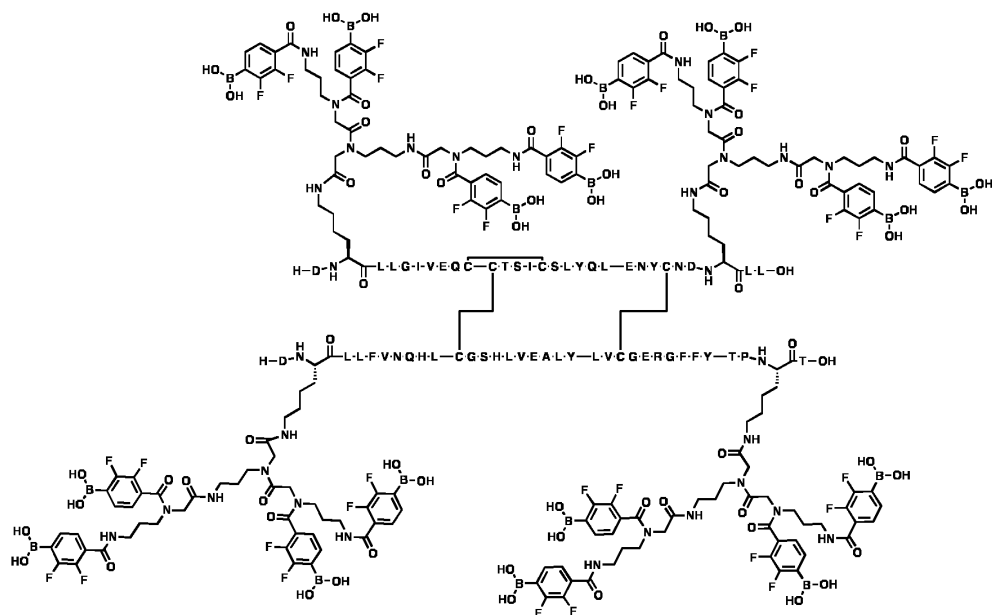
Example 673:**Example 674:**

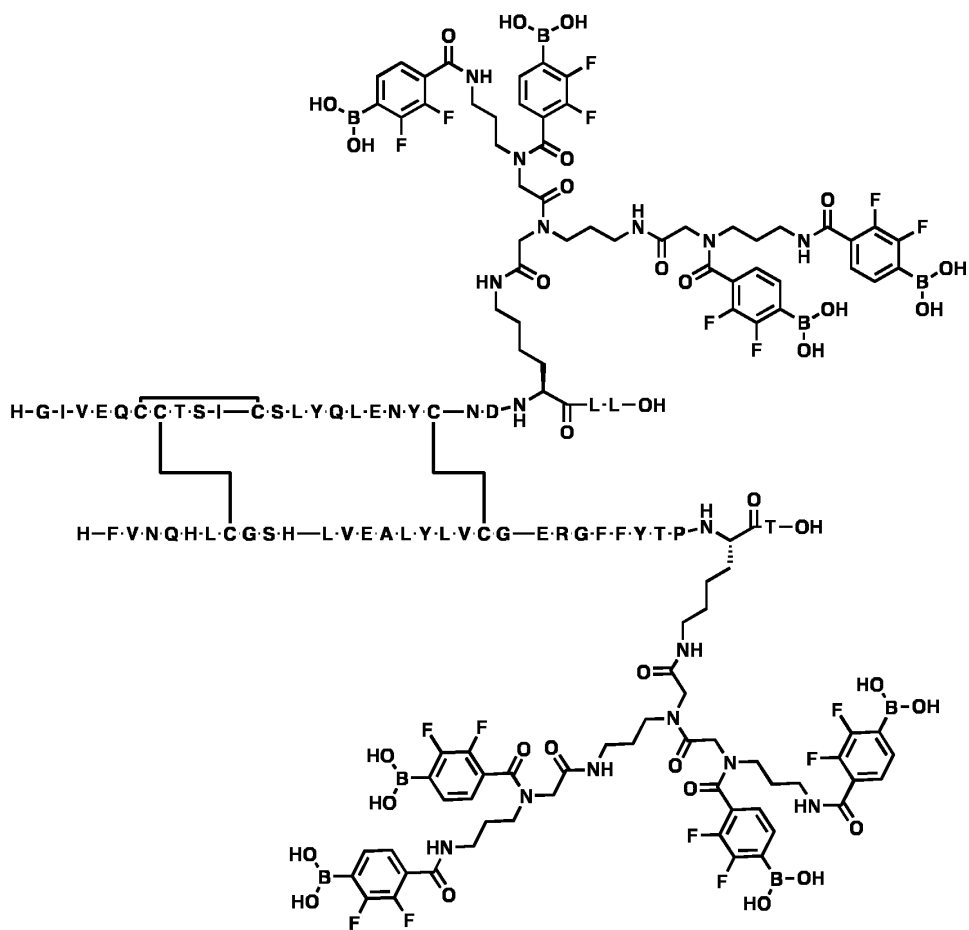
Example 675:



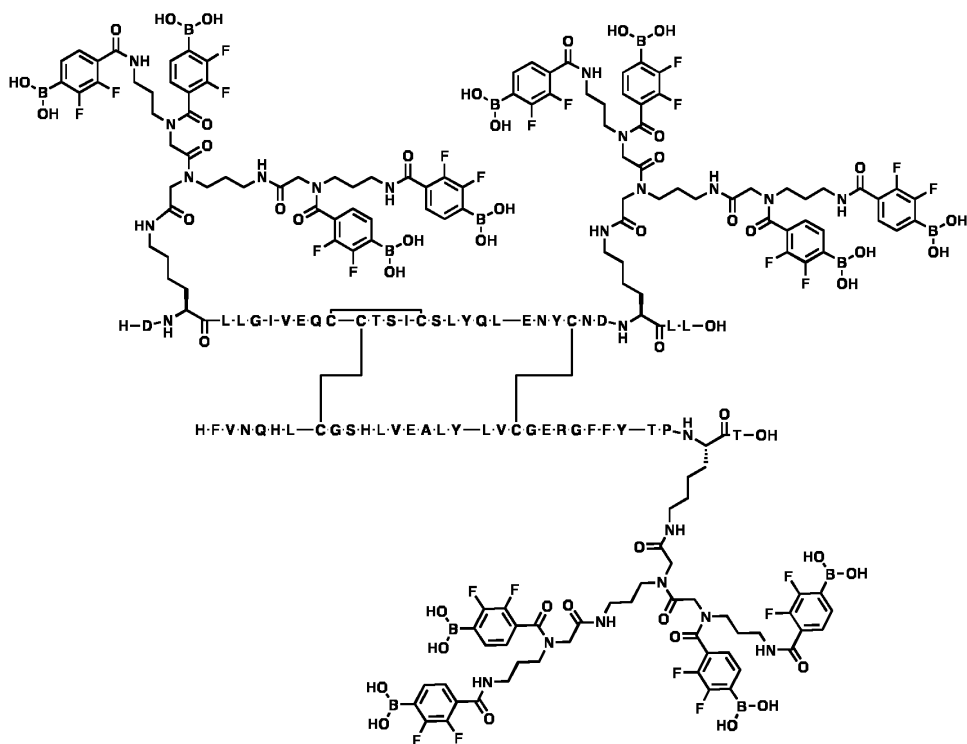
Example 676:

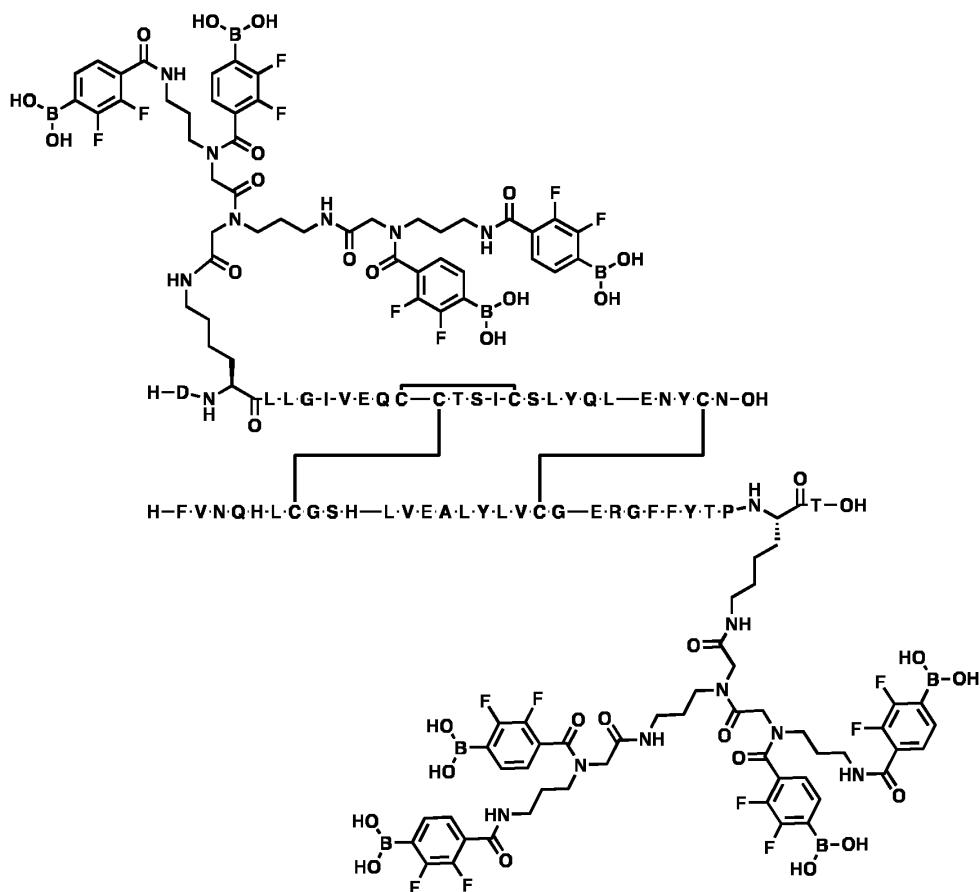


Example 677:

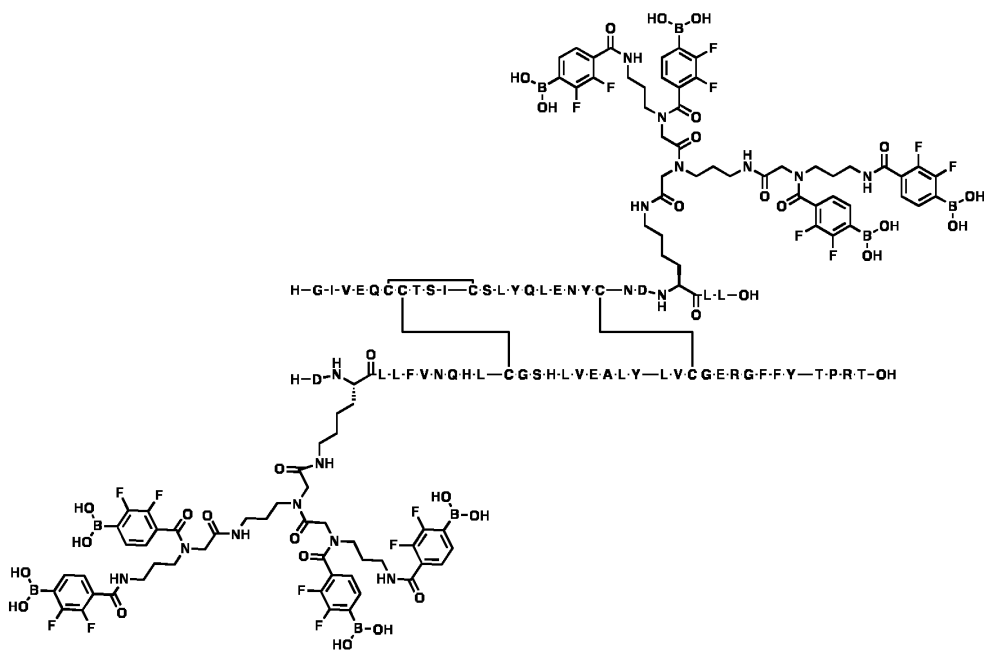


Example 679:

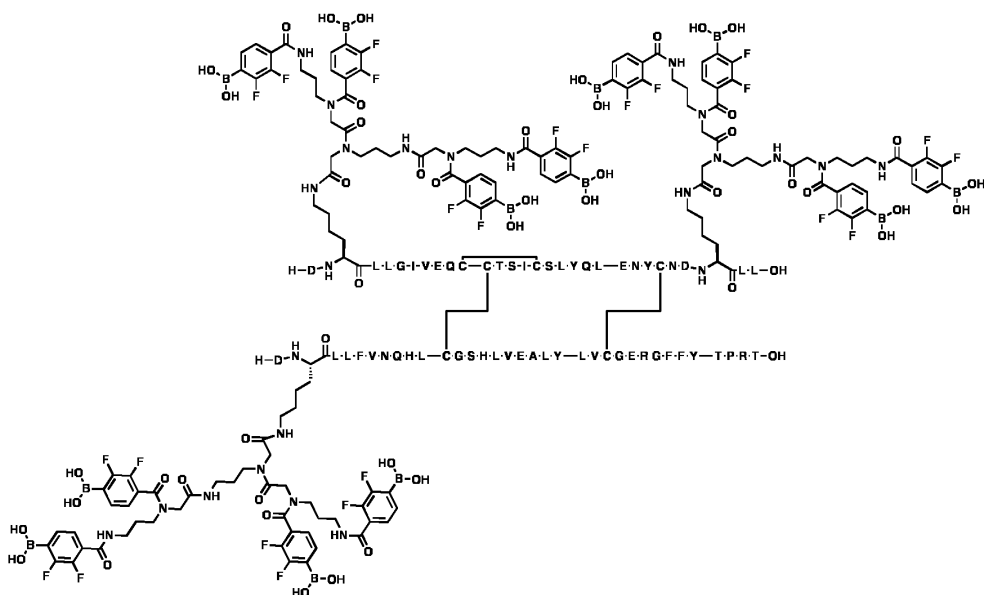




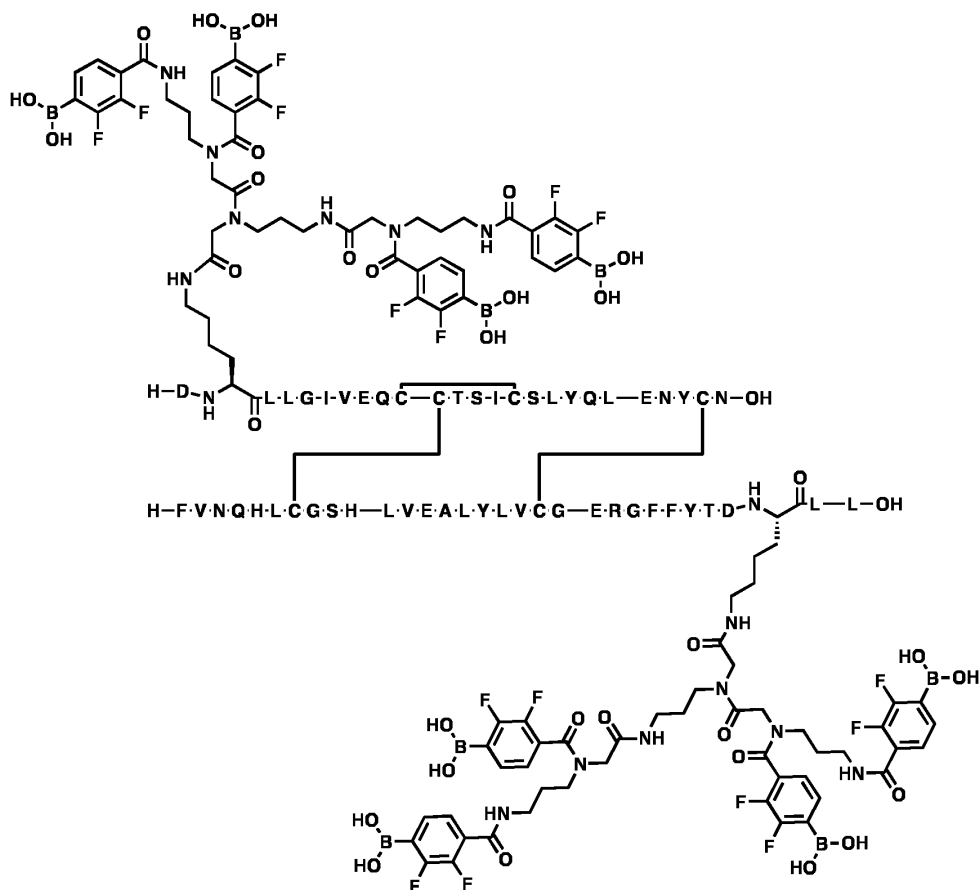
Example 681:



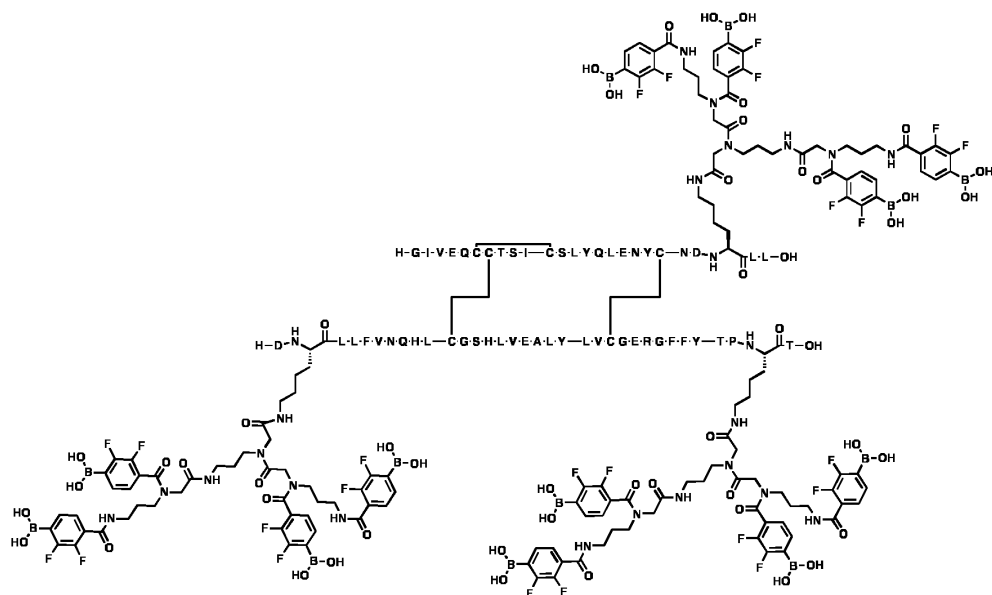
Example 682:



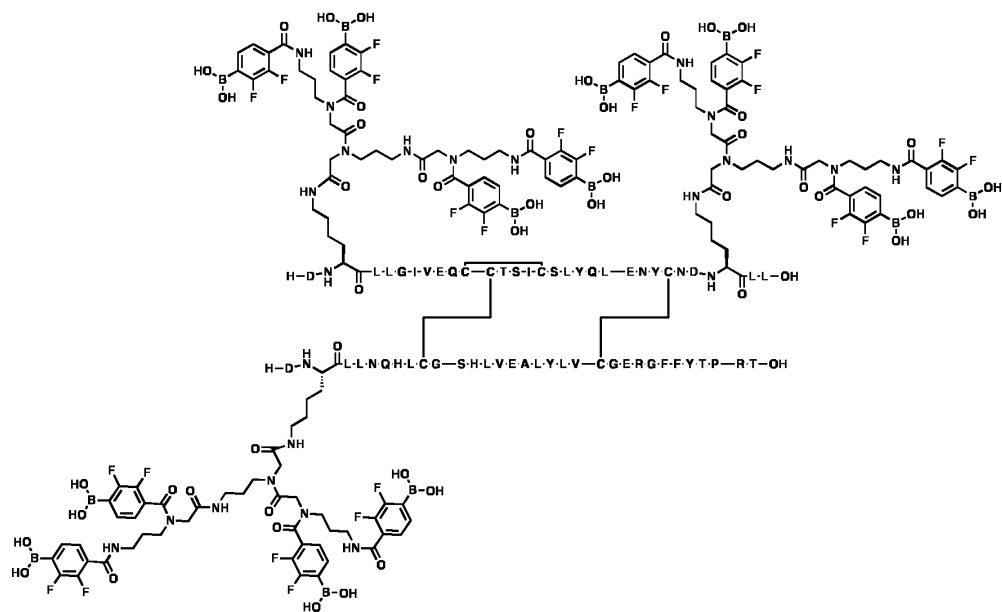
Example 683:



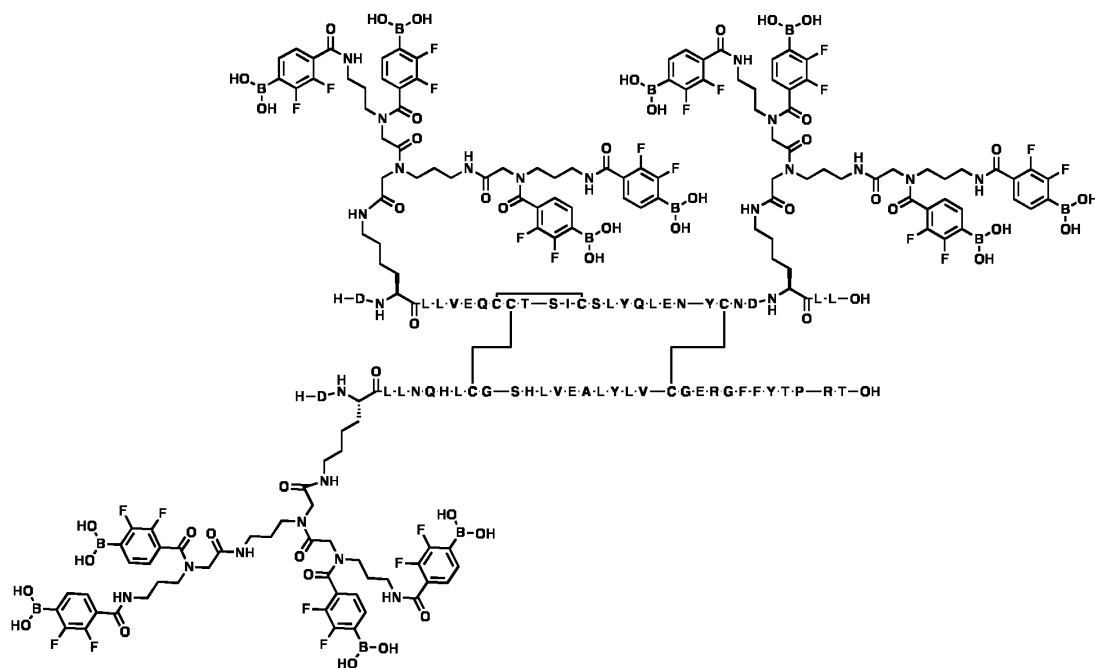
Example 684:



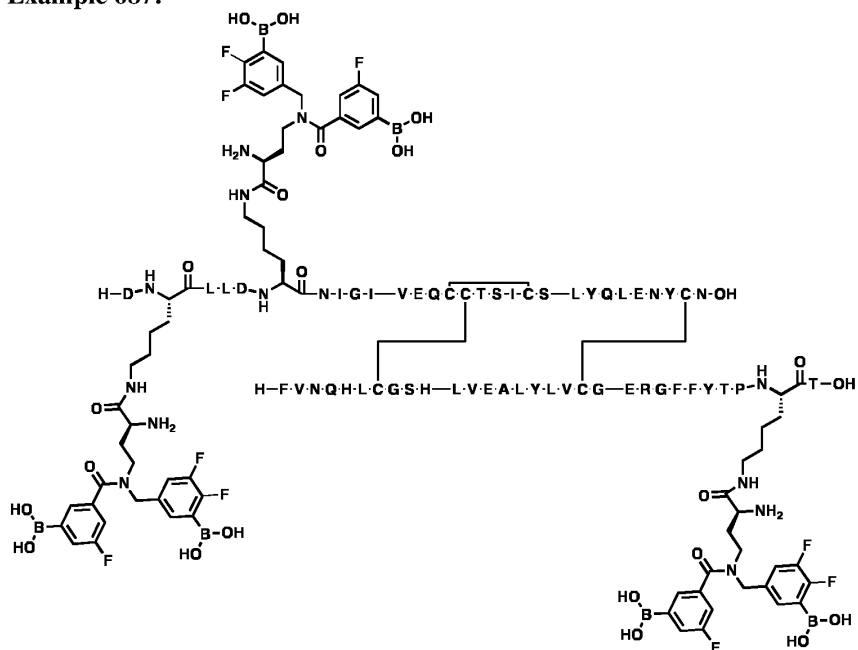
Example 685:



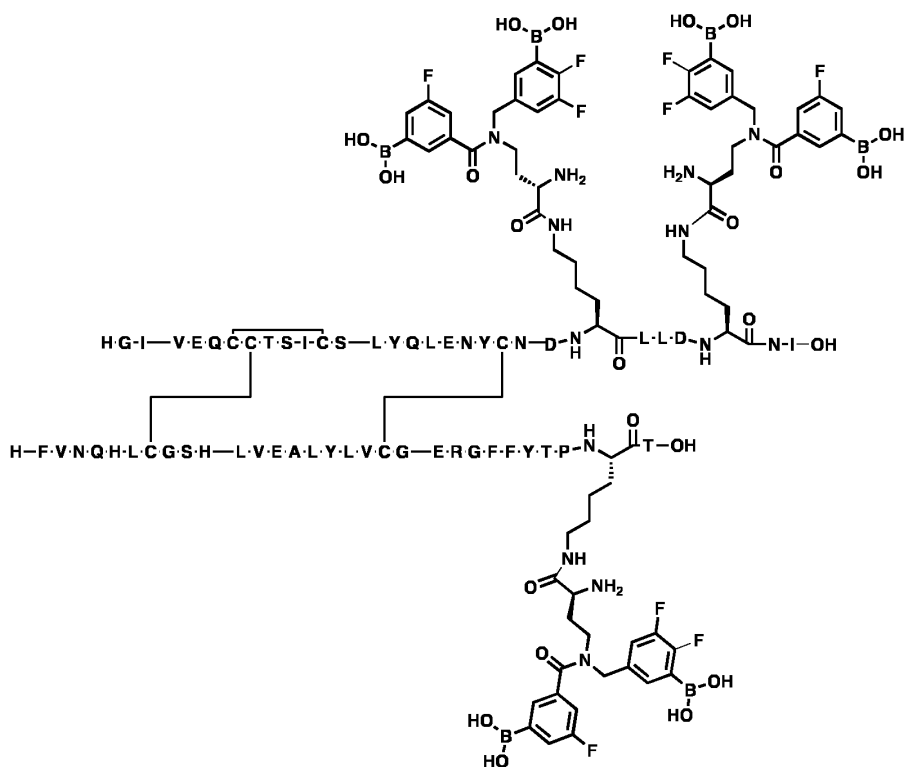
Example 686:



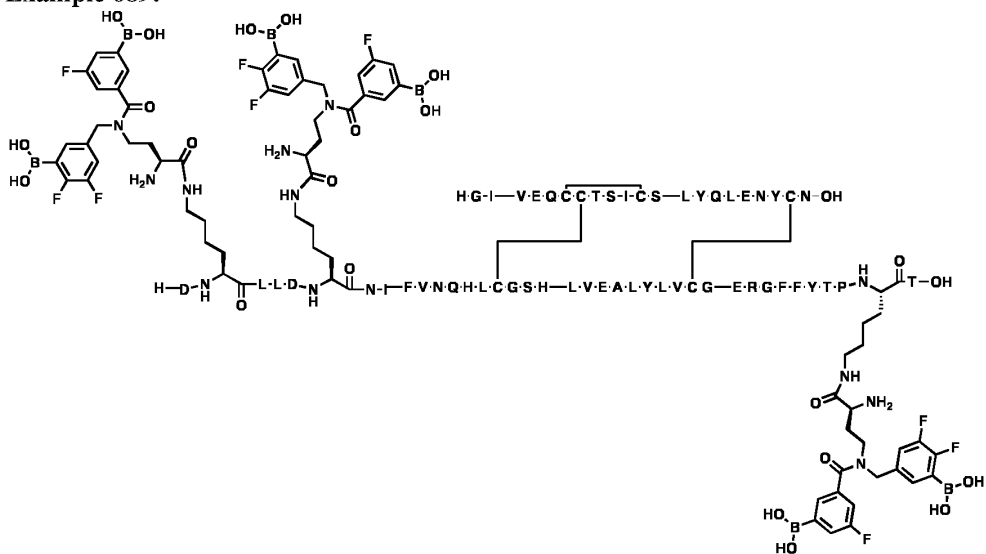
Example 687:

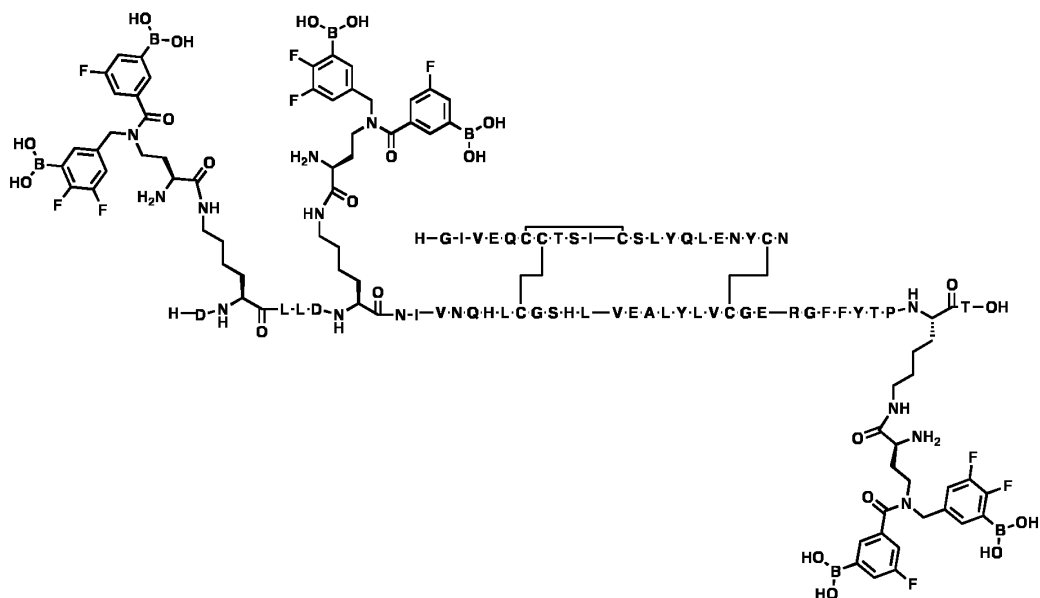
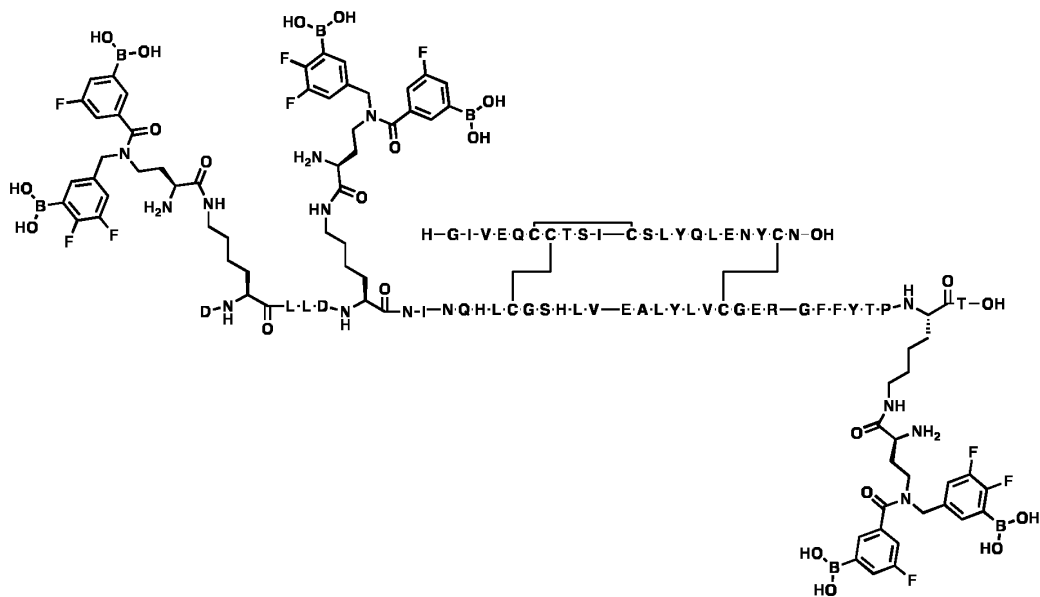


Example 688:

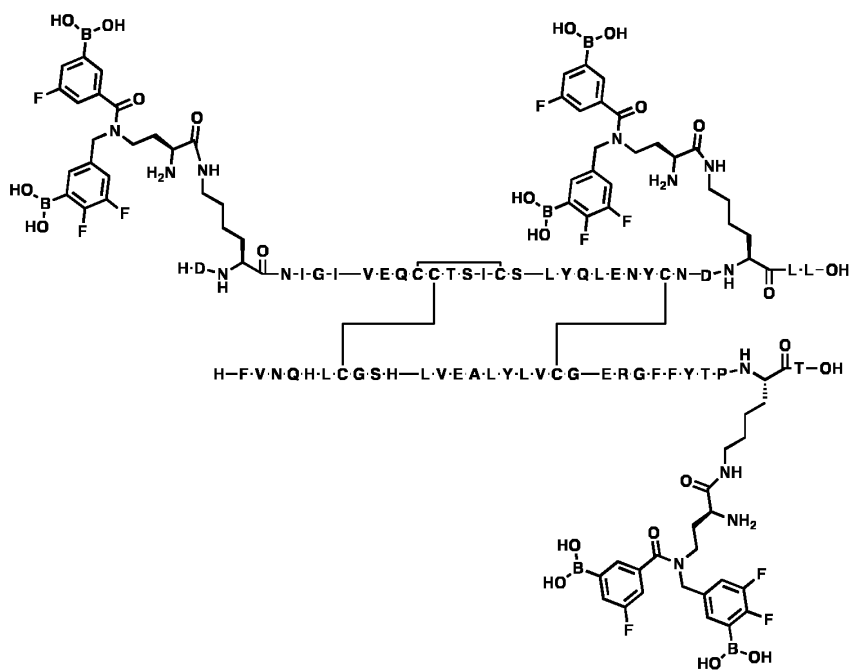


5 Example 689:



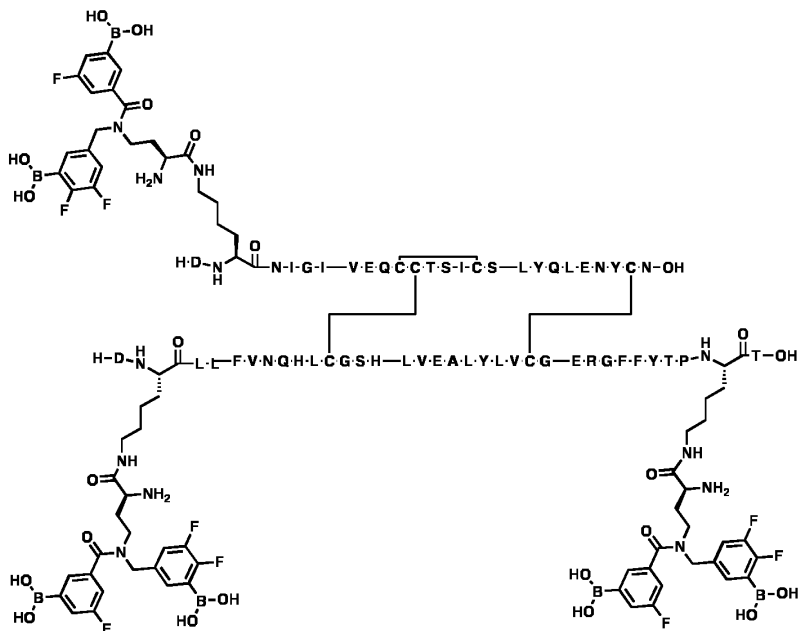
Example 690:**5 Example 691:**

Example 692:

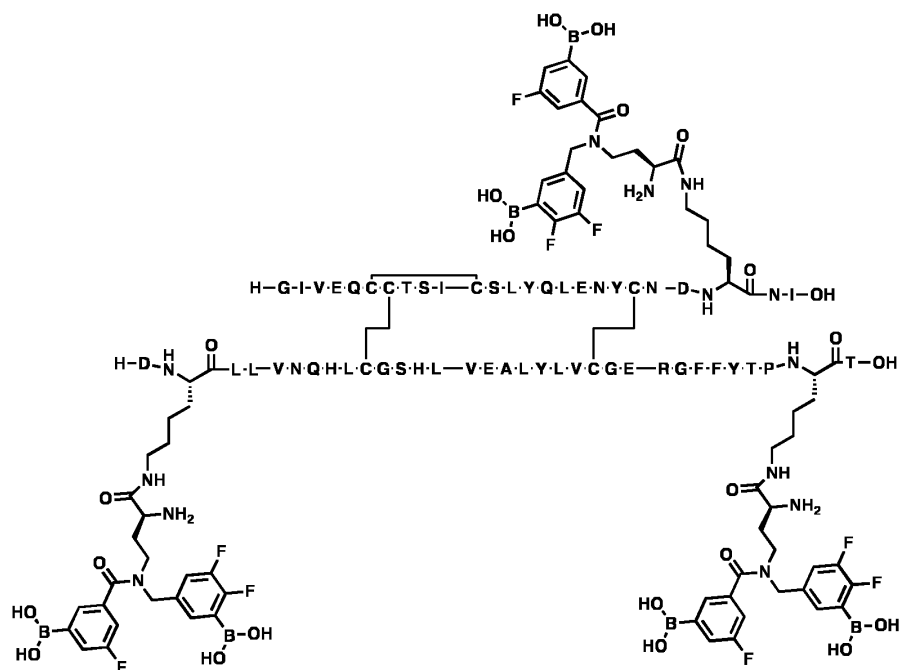


Example 693:

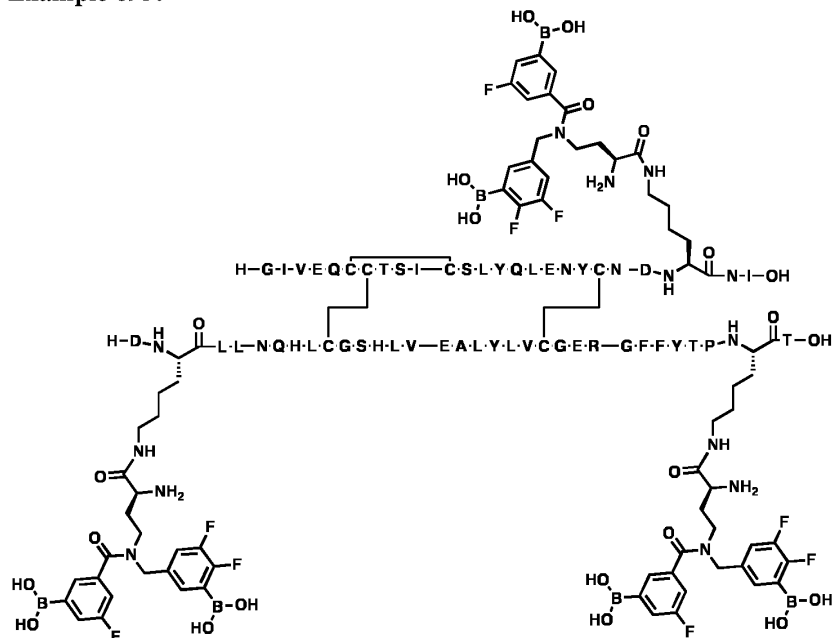
5



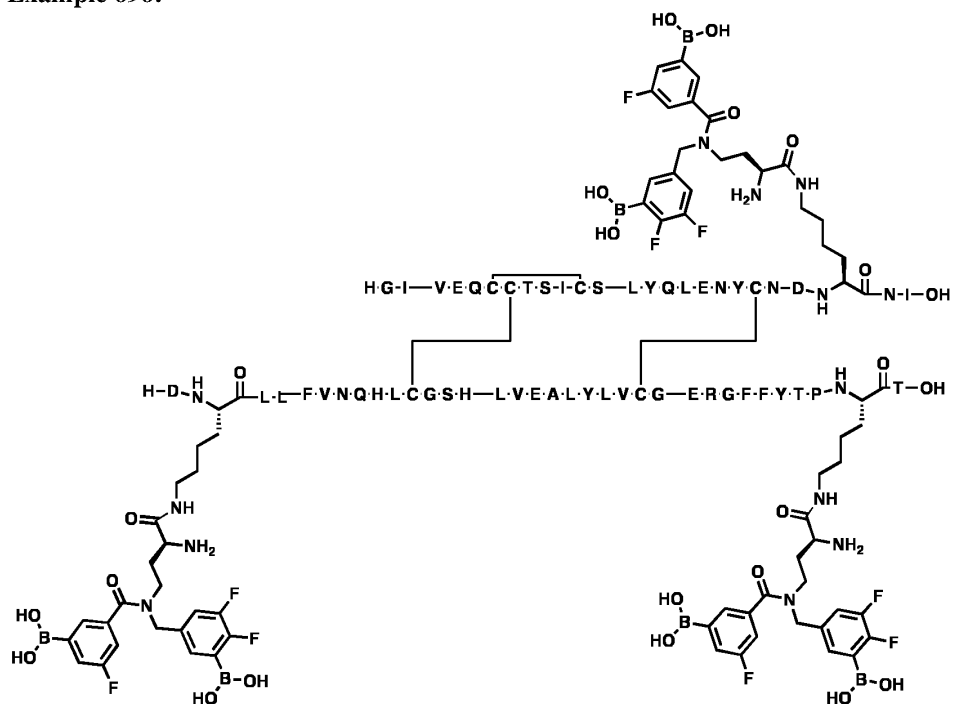
Example 694:



5 Example 695:

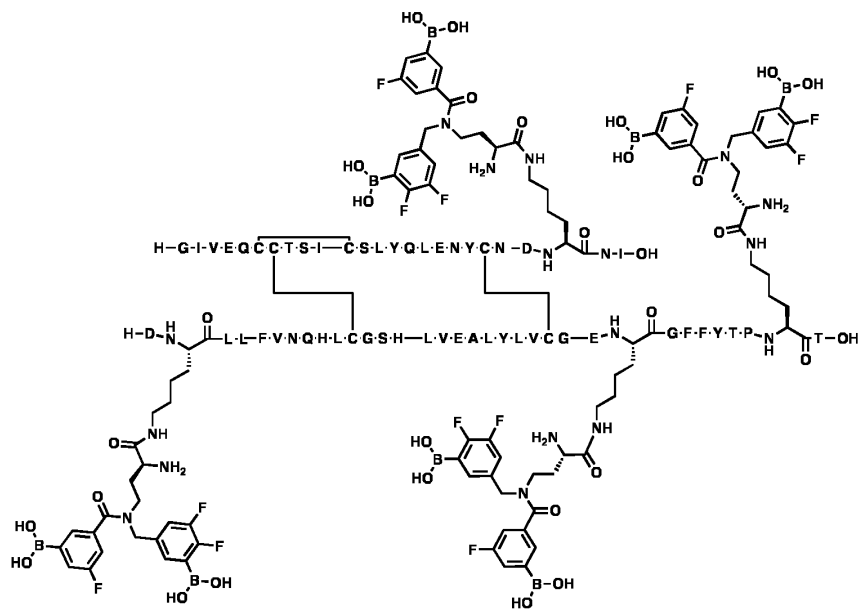


Example 696:

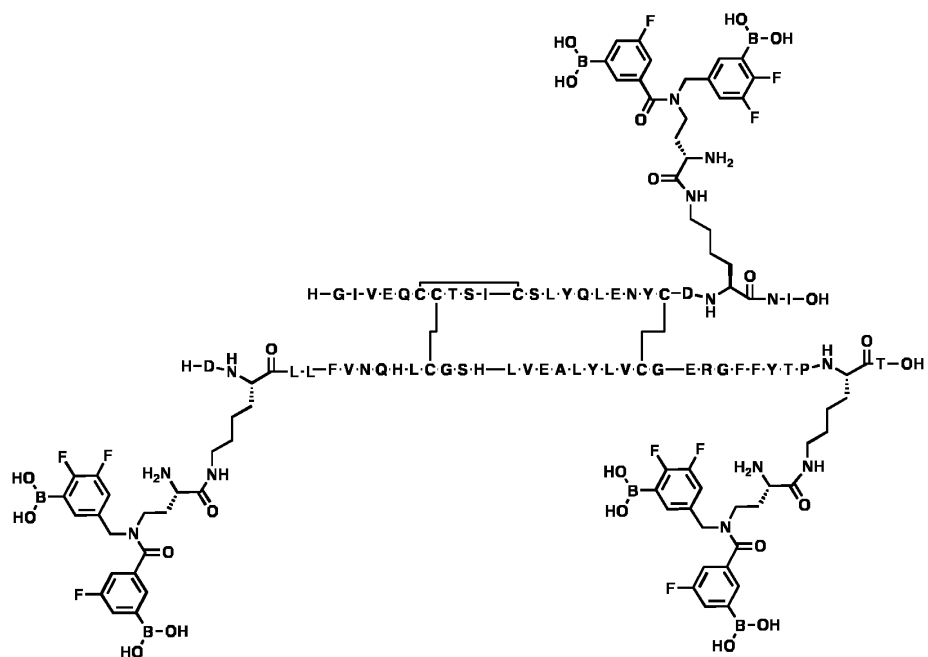


Example 697:

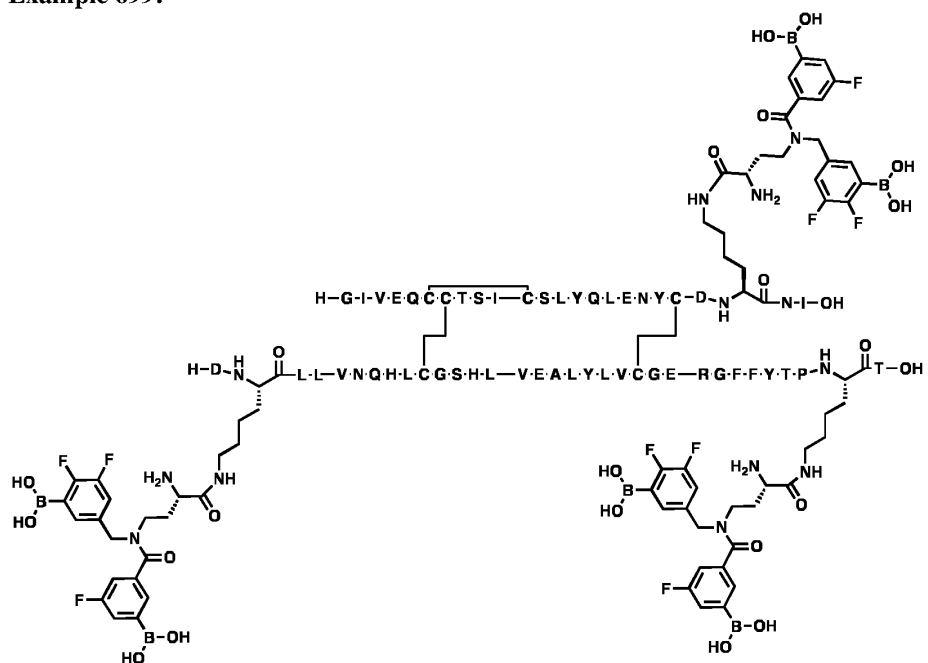
5

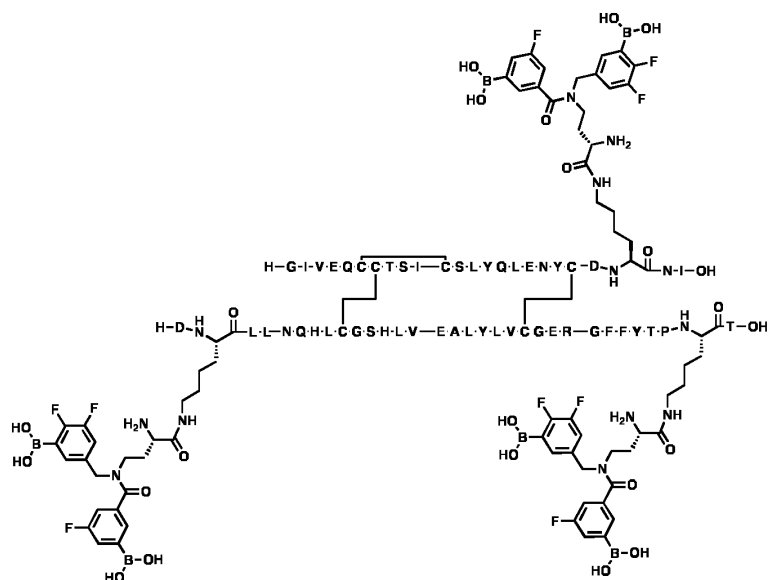
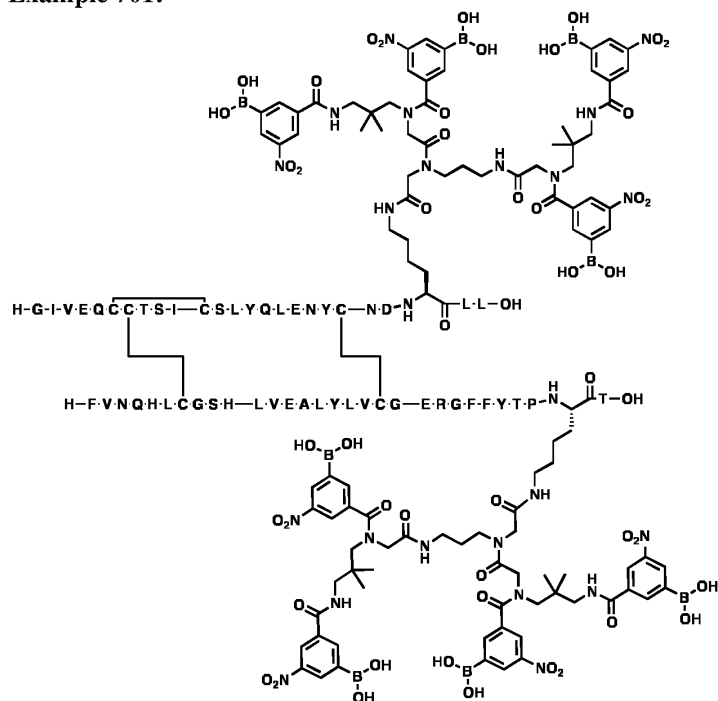


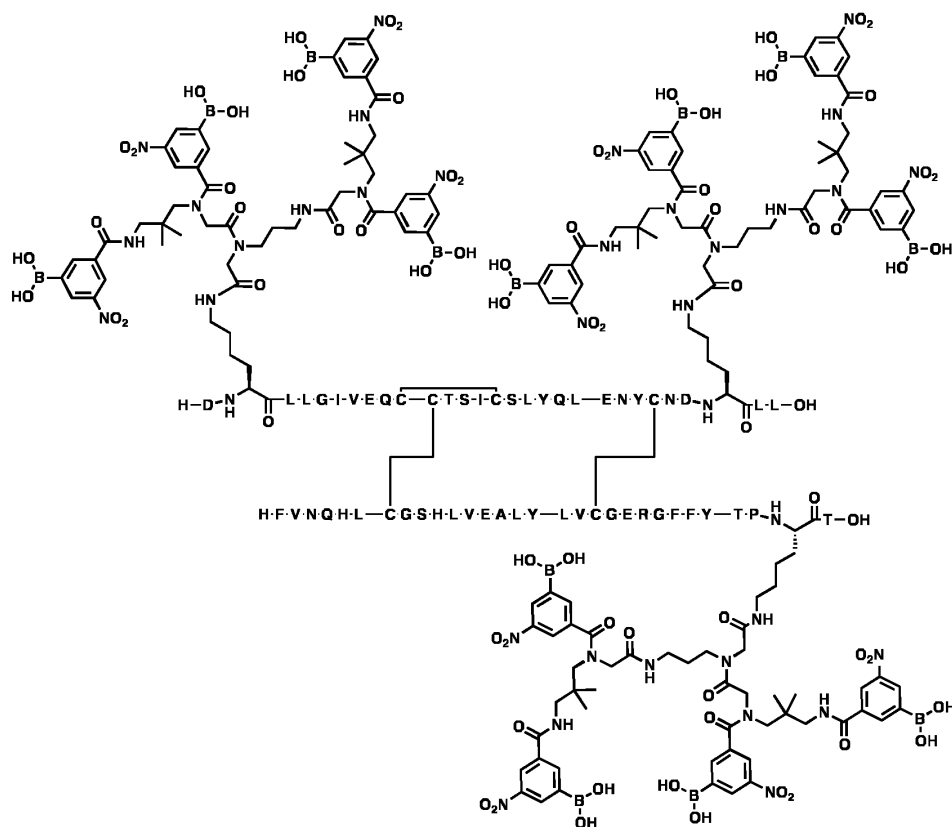
Example 698:



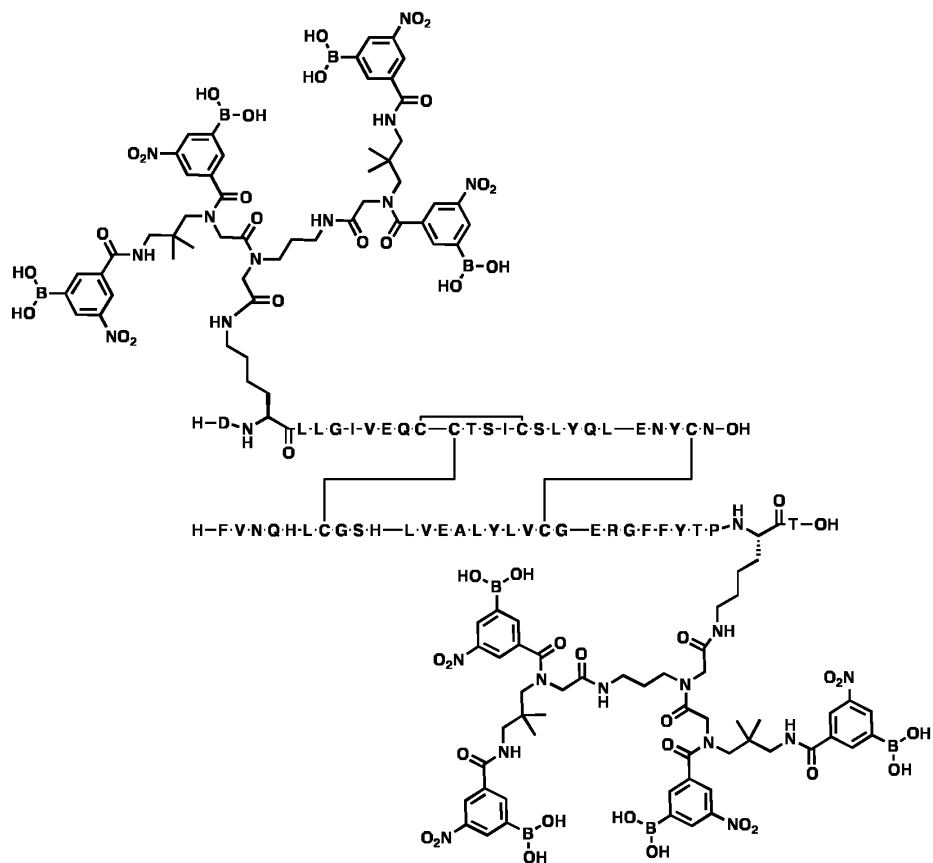
5 Example 699:



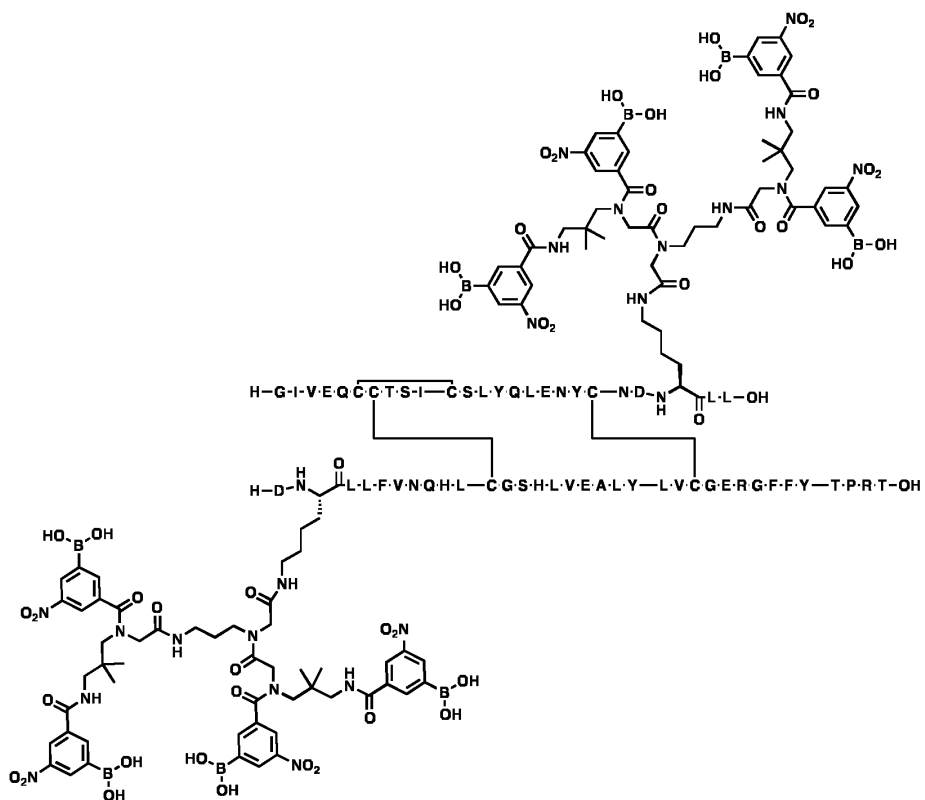
Example 700:**5 Example 701:**

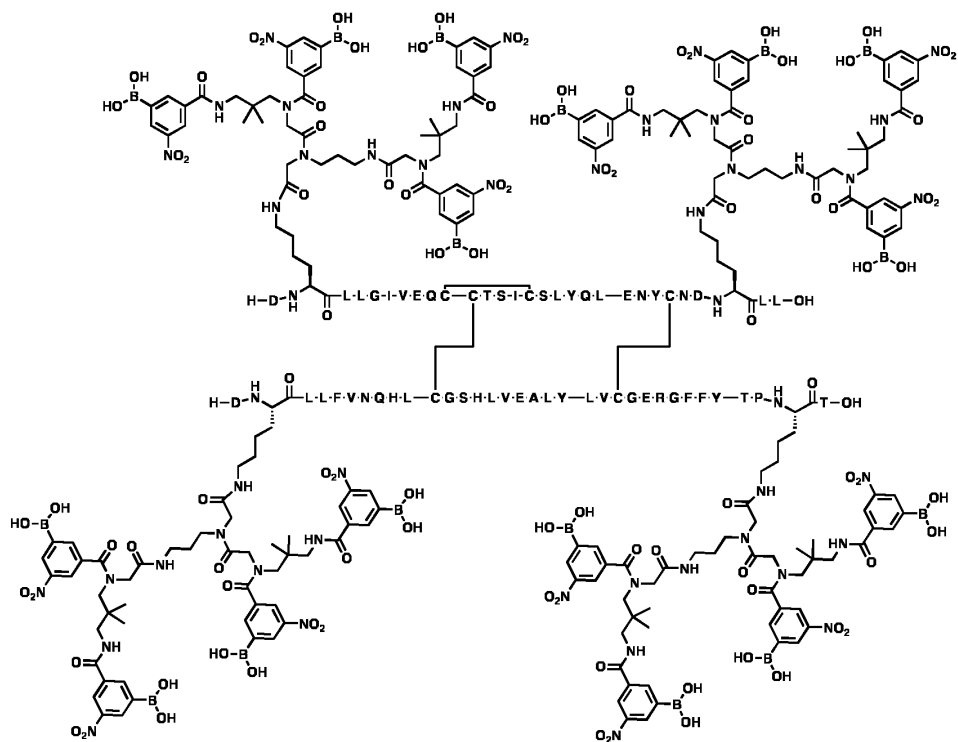


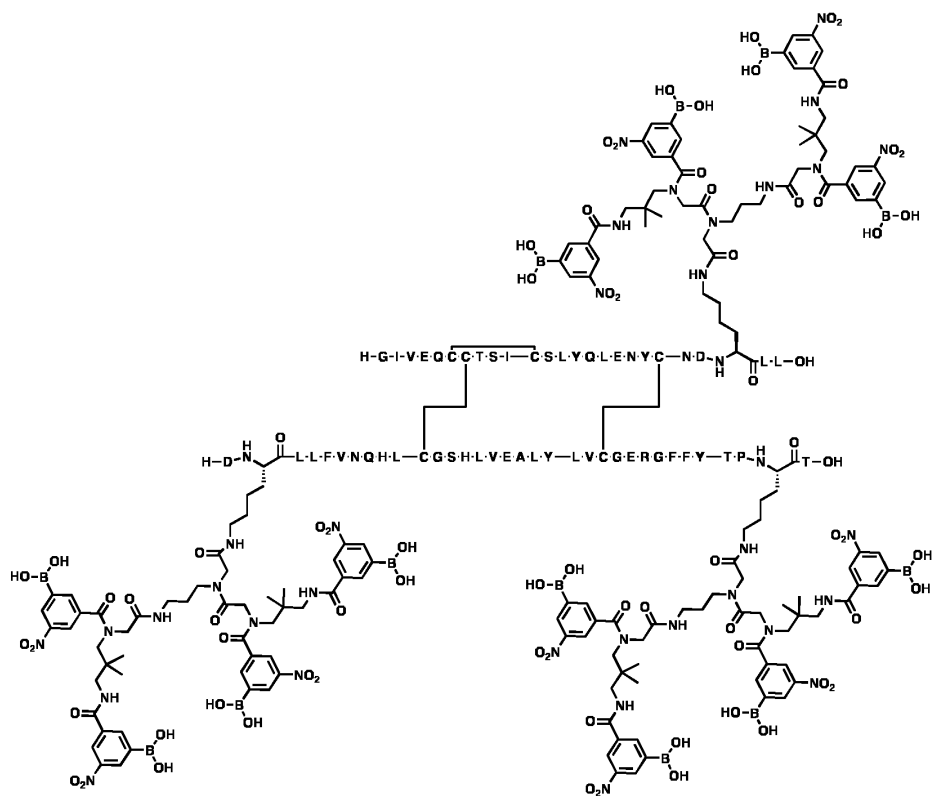
Example 703:

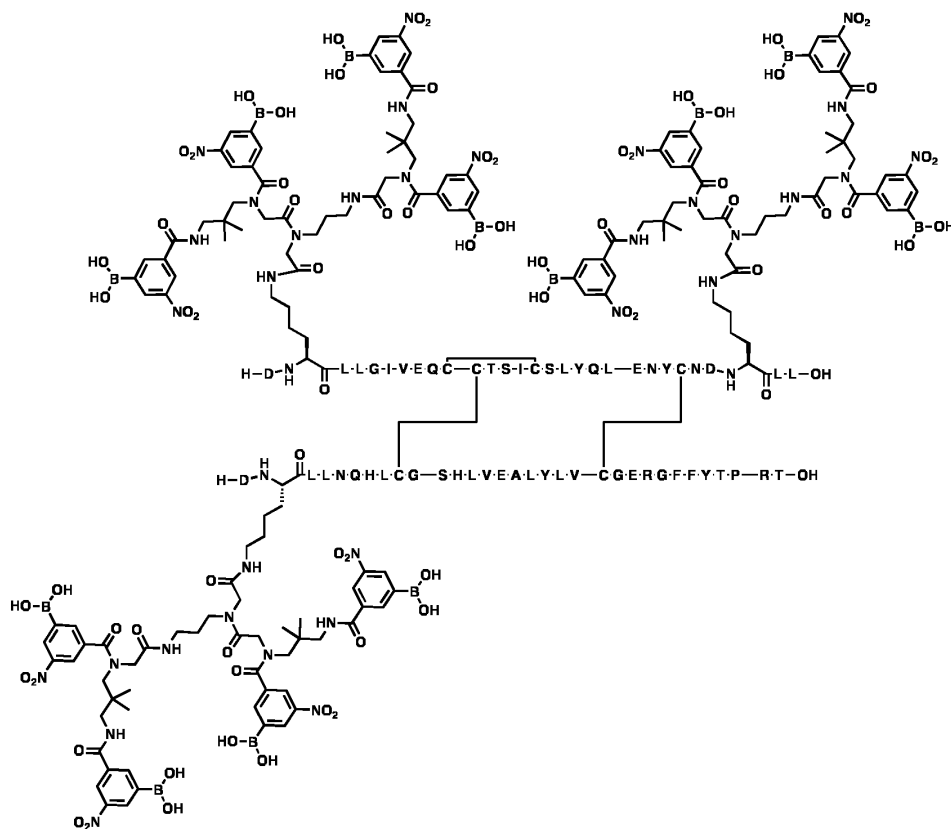


Example 704:

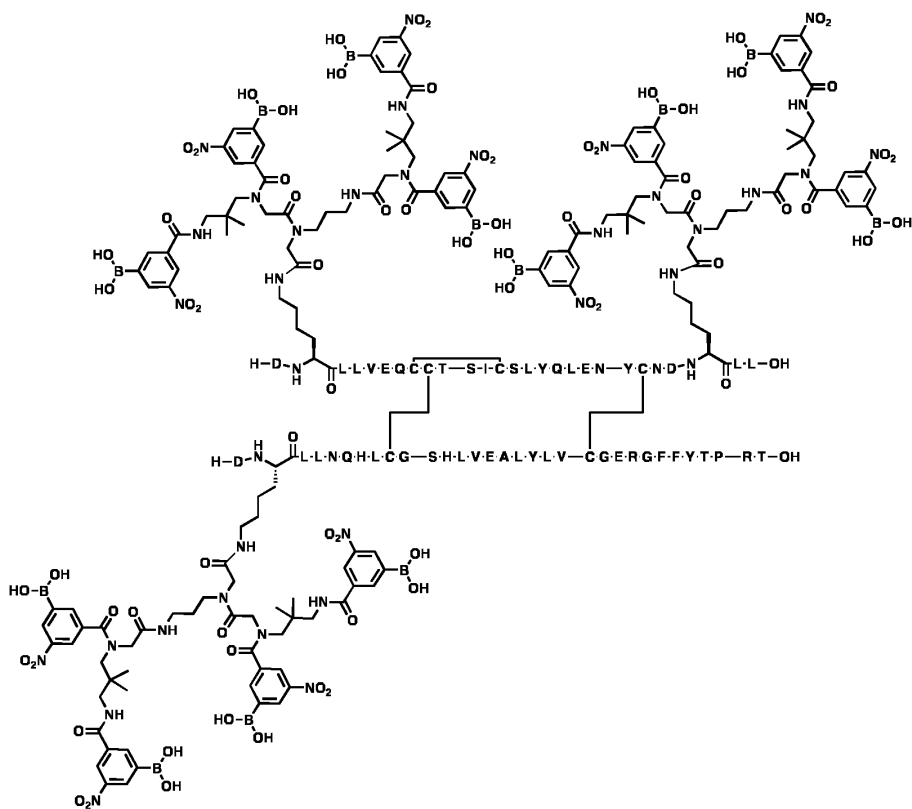




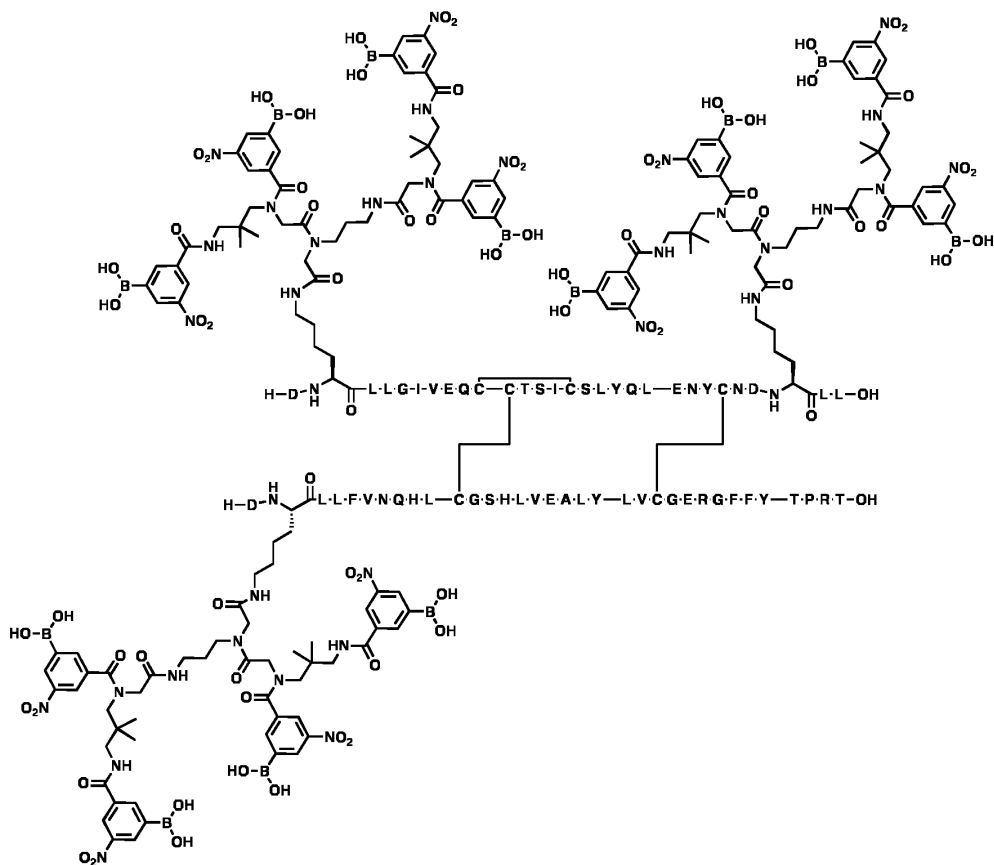
Example 706:

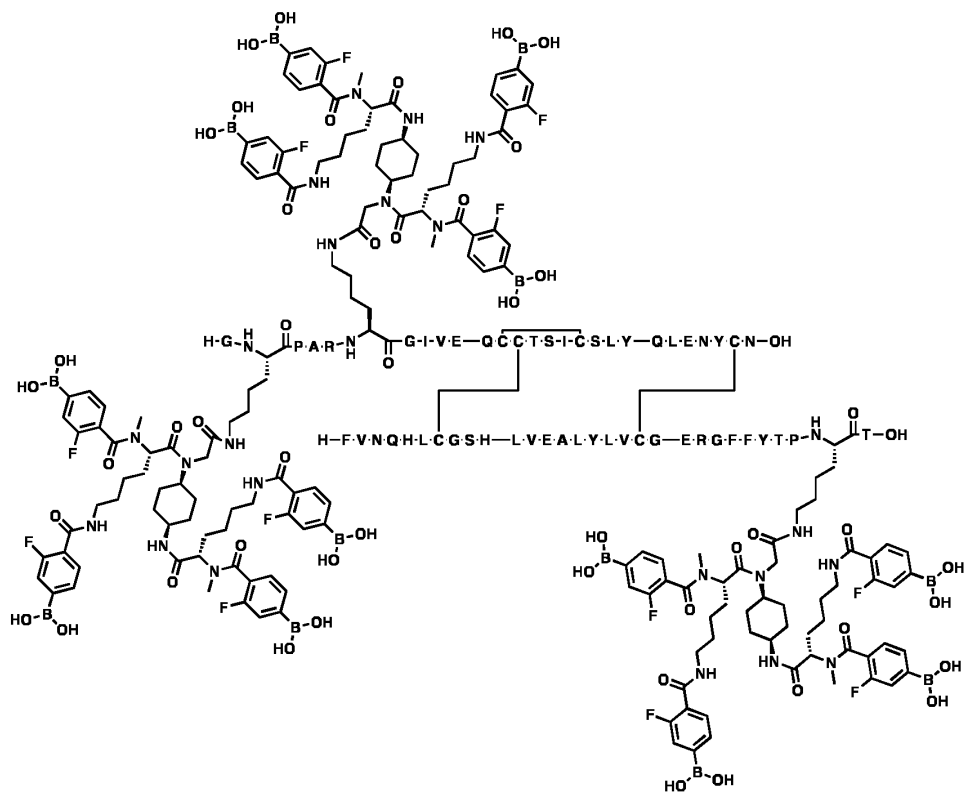
Example 707:

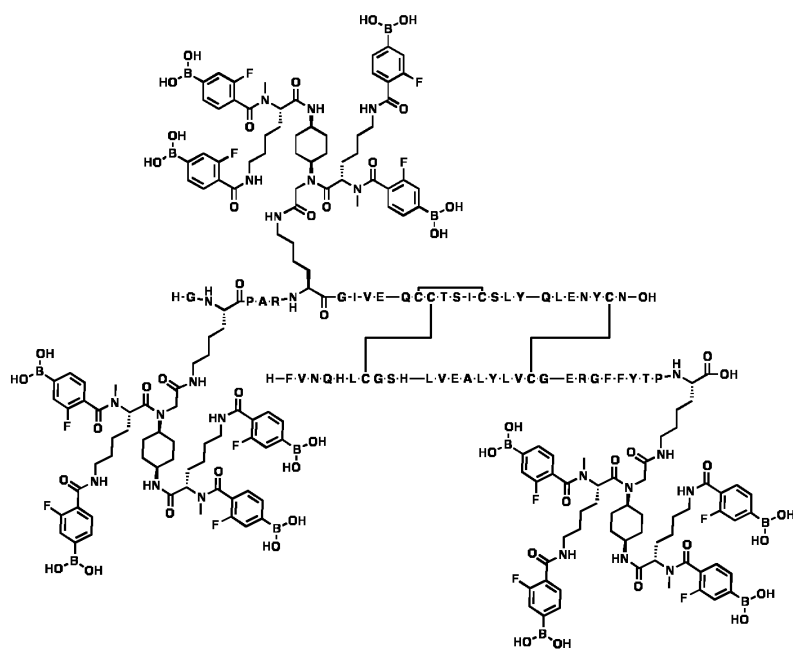
Example 708:



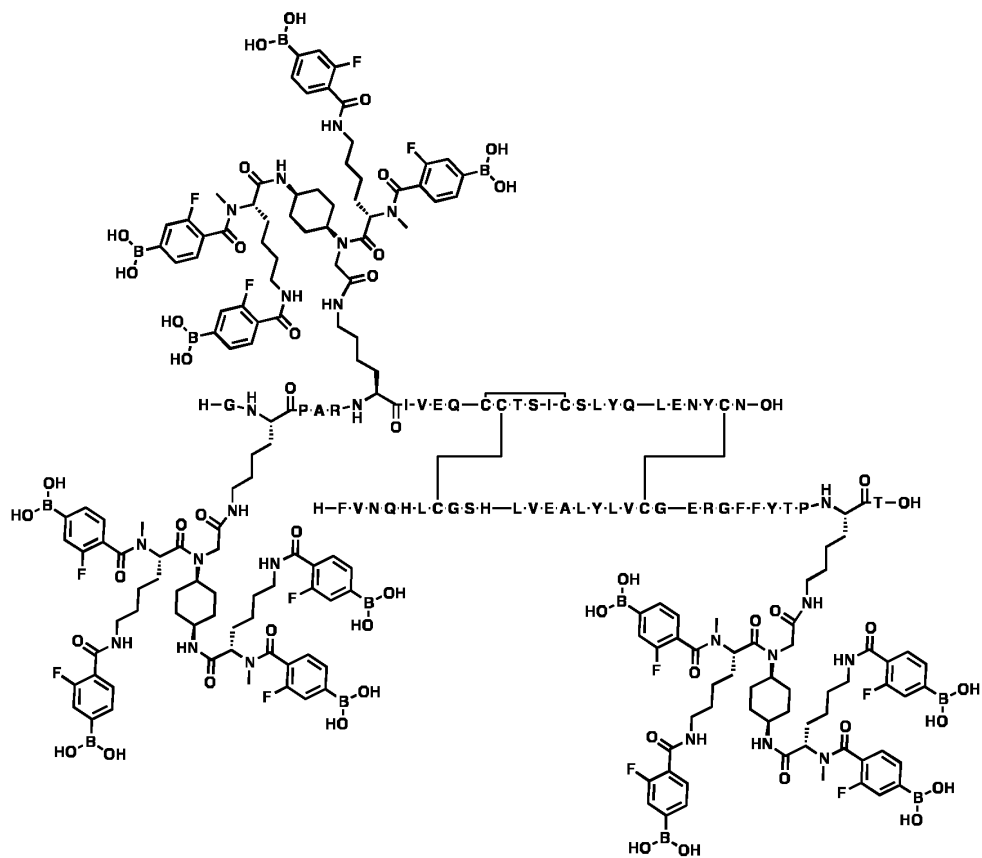
Example 709:



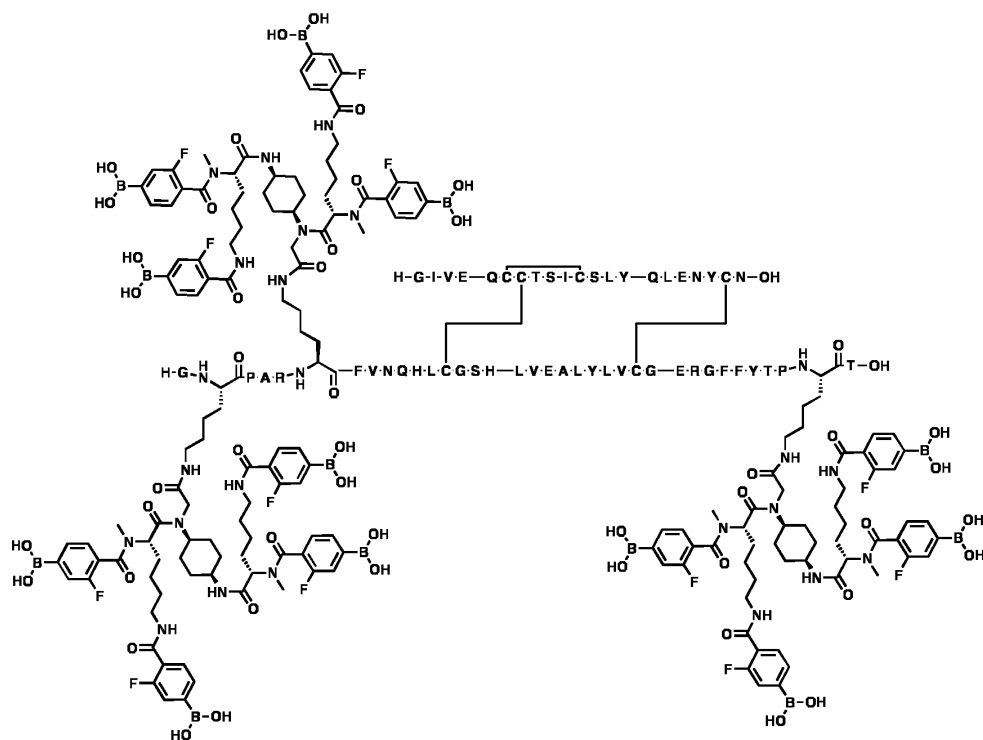
Example 710:

Example 711:

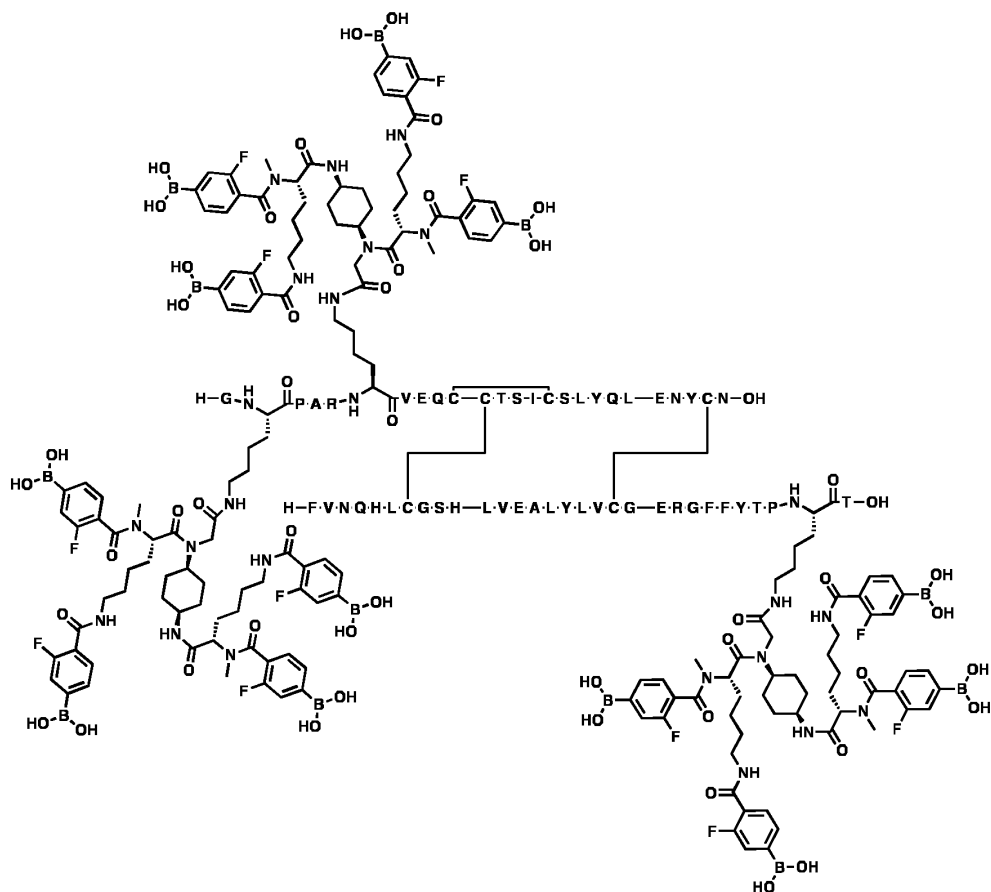
Example 712:



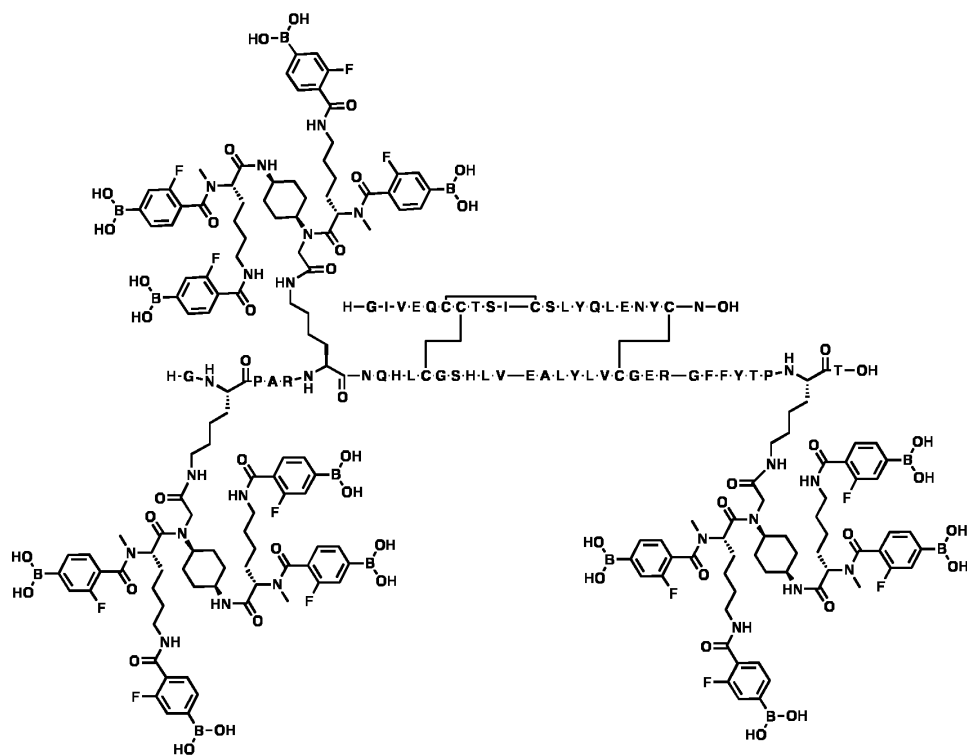
Example 713:

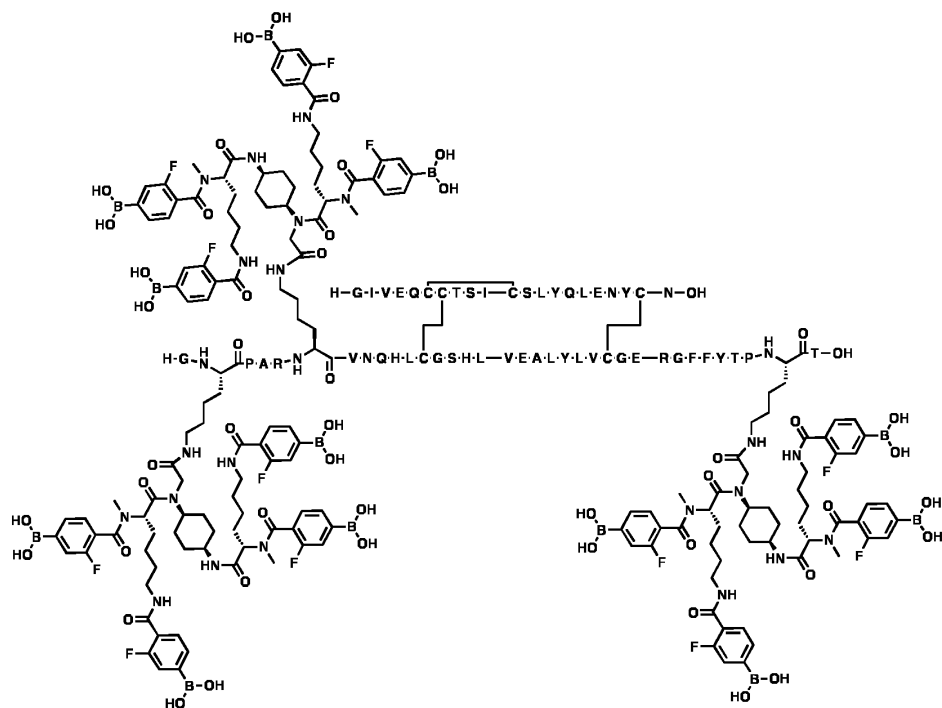


Example 714:

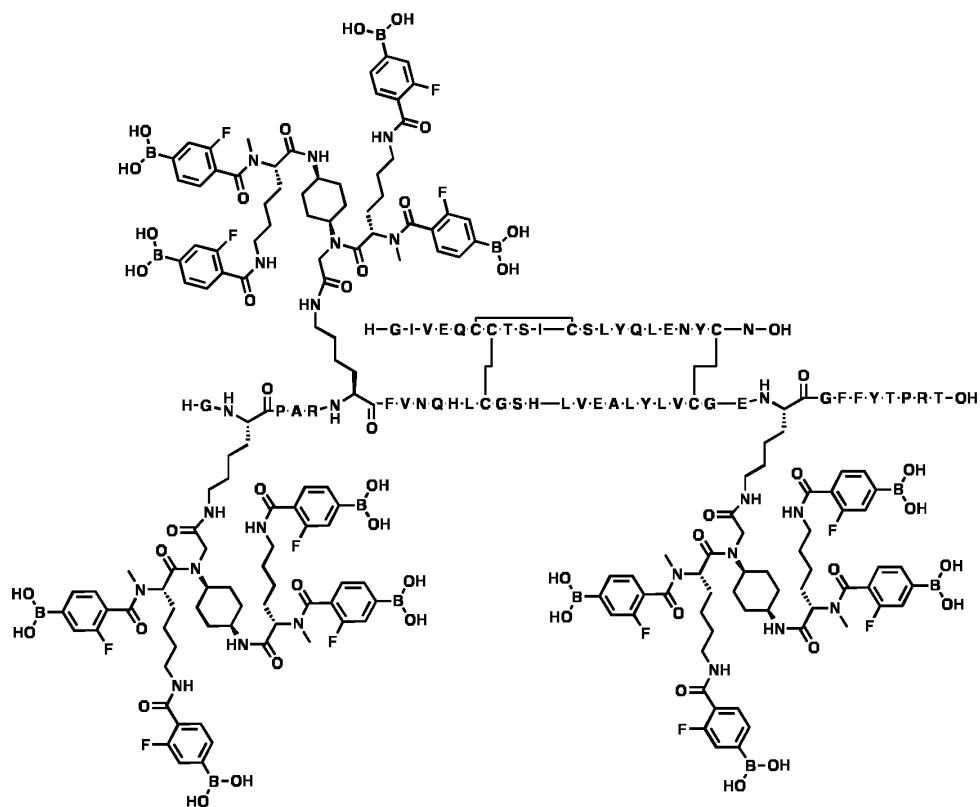


Example 715:

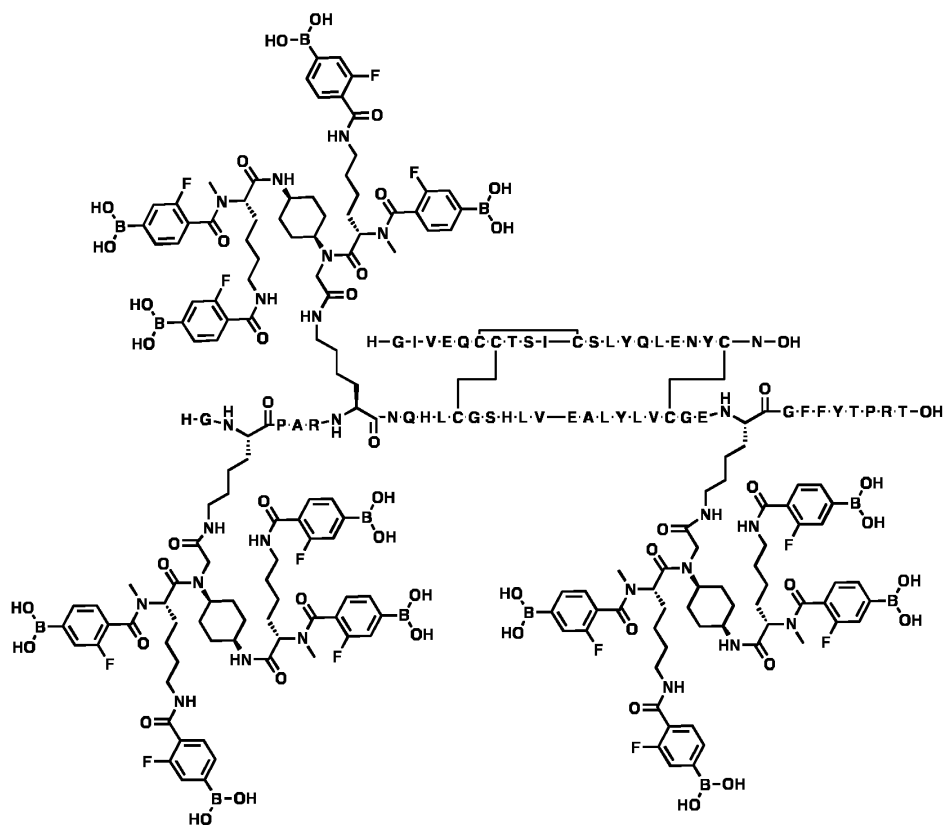


Example 716:

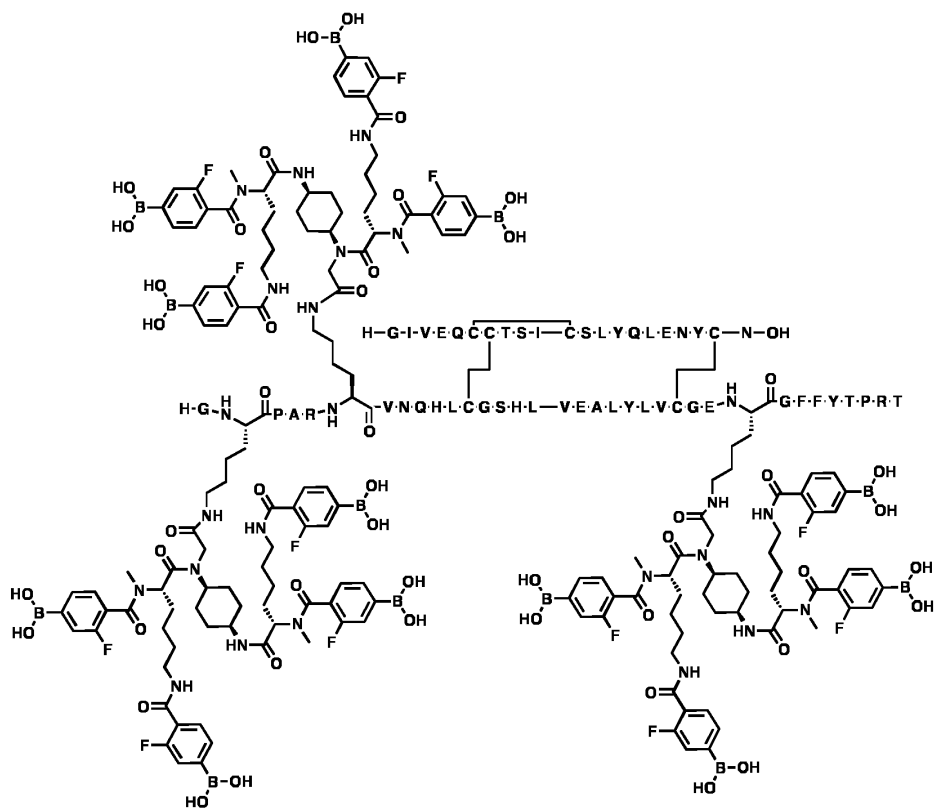
Example 717:

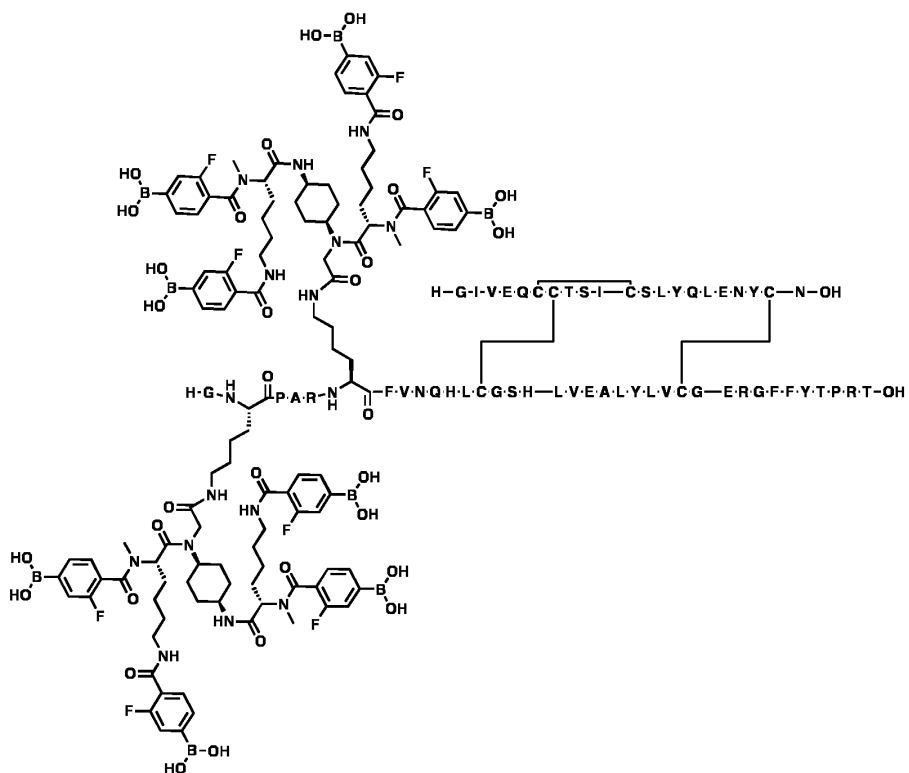
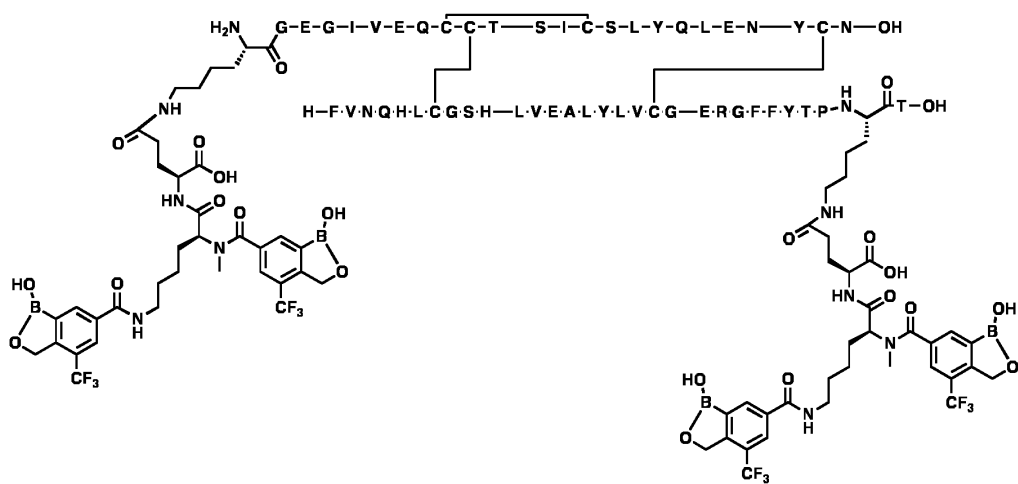


Example 718:

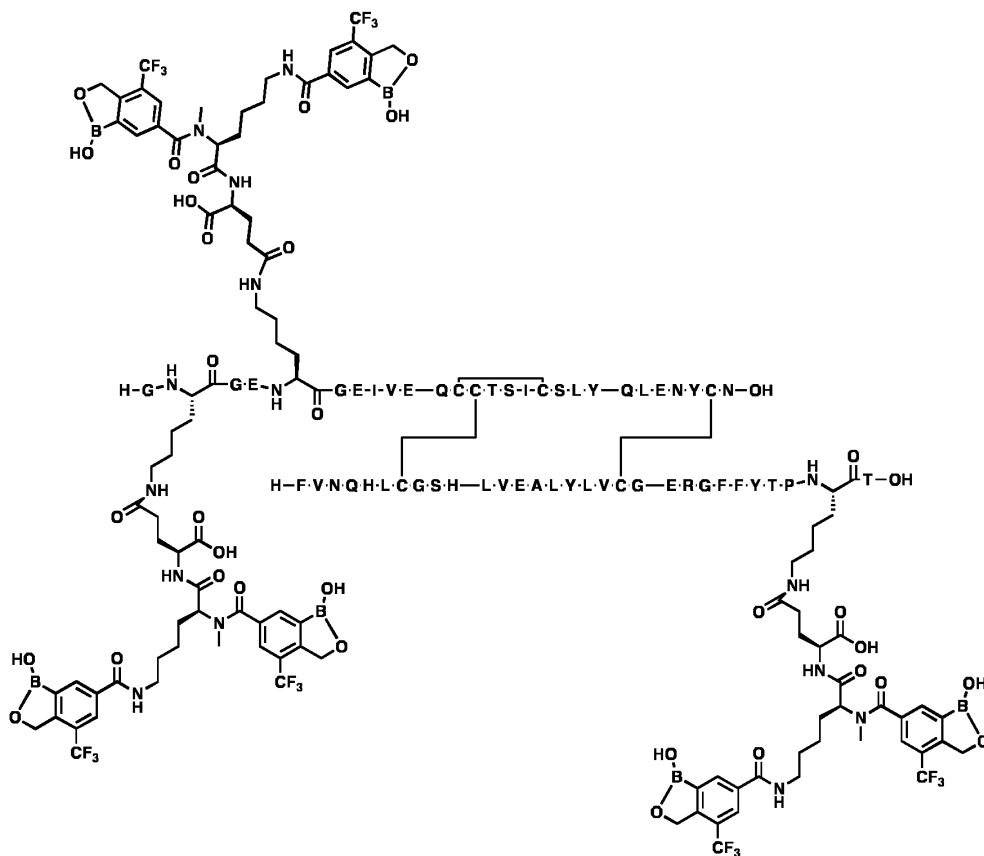


Example 719:

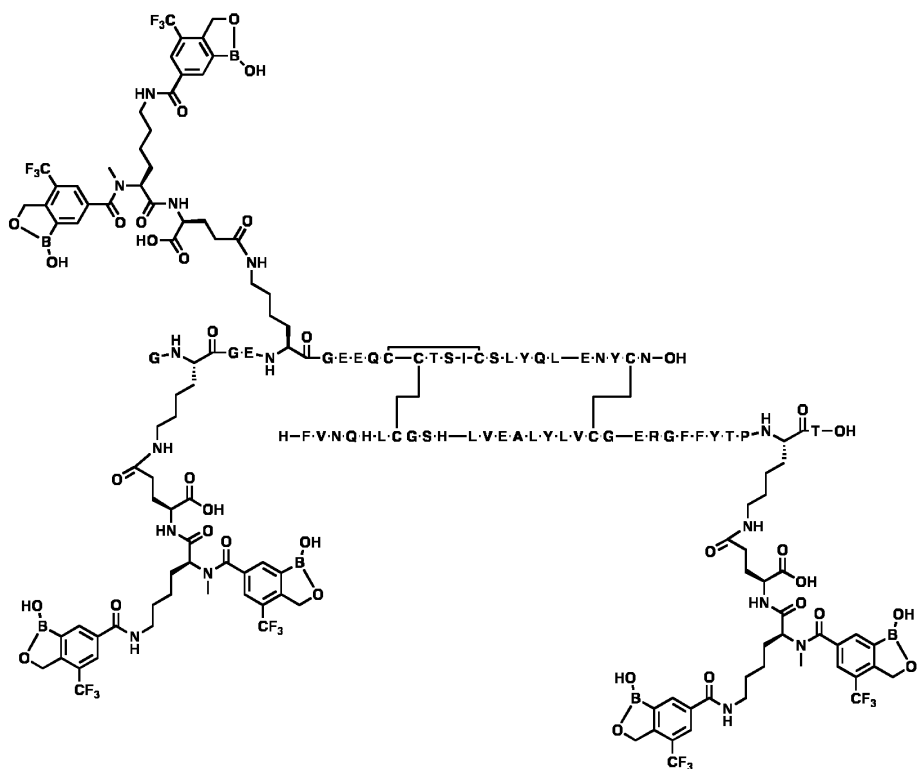


Example 720:**Example 721:**

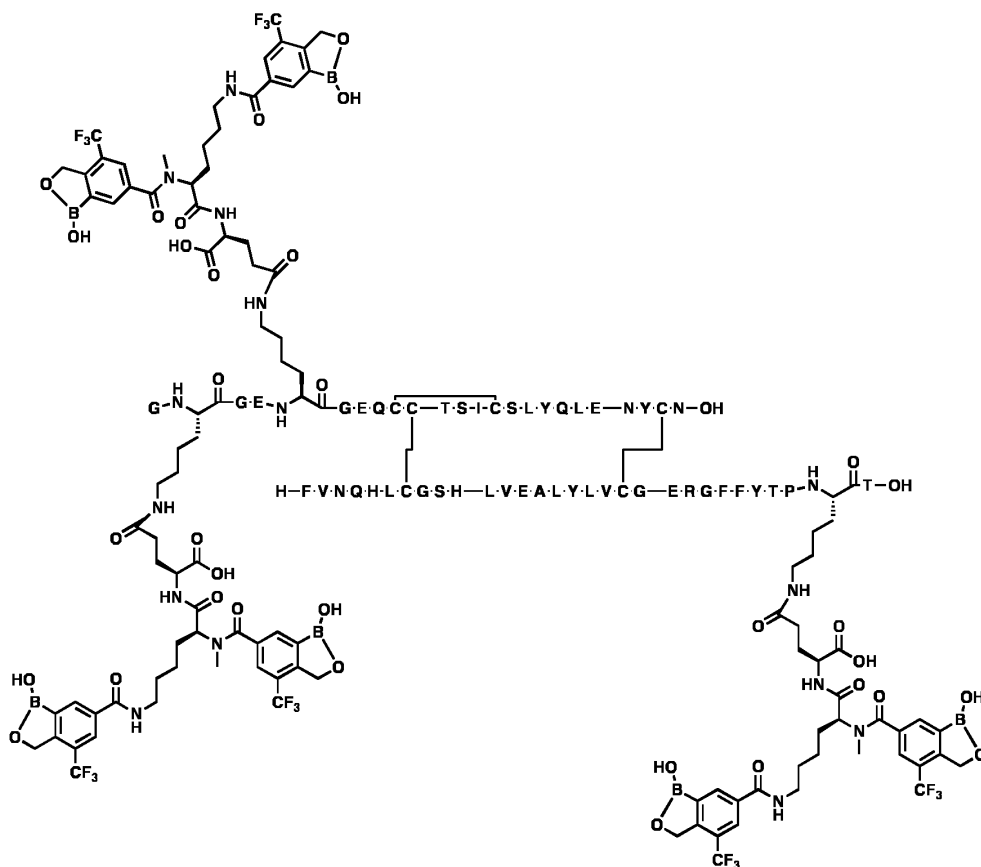
Example 722:



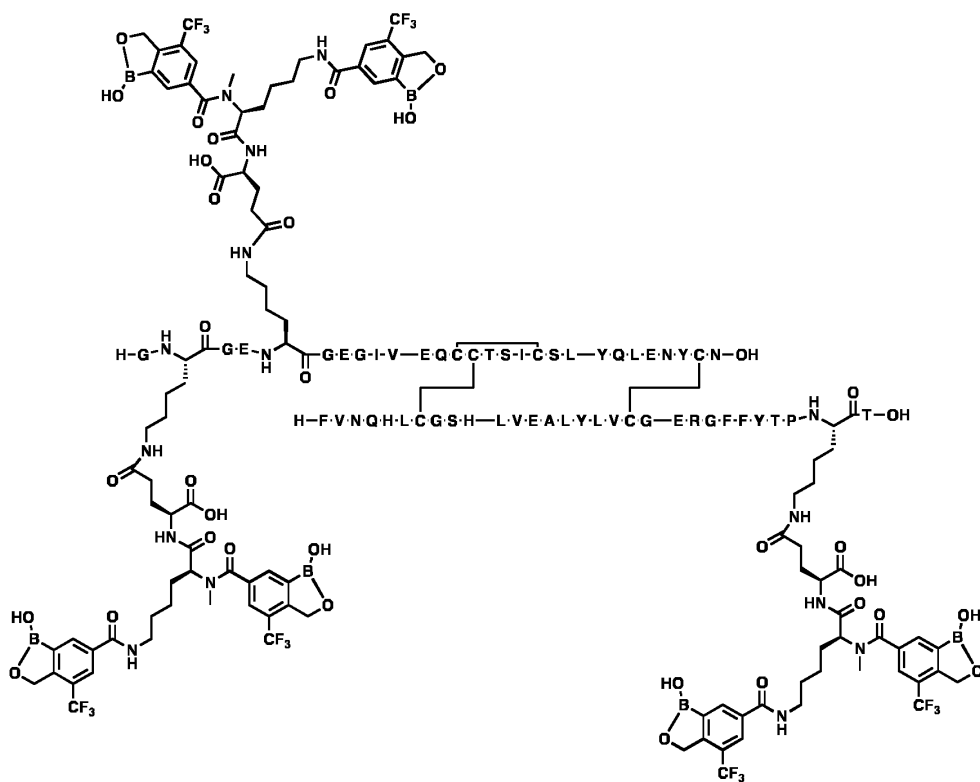
Example 723:



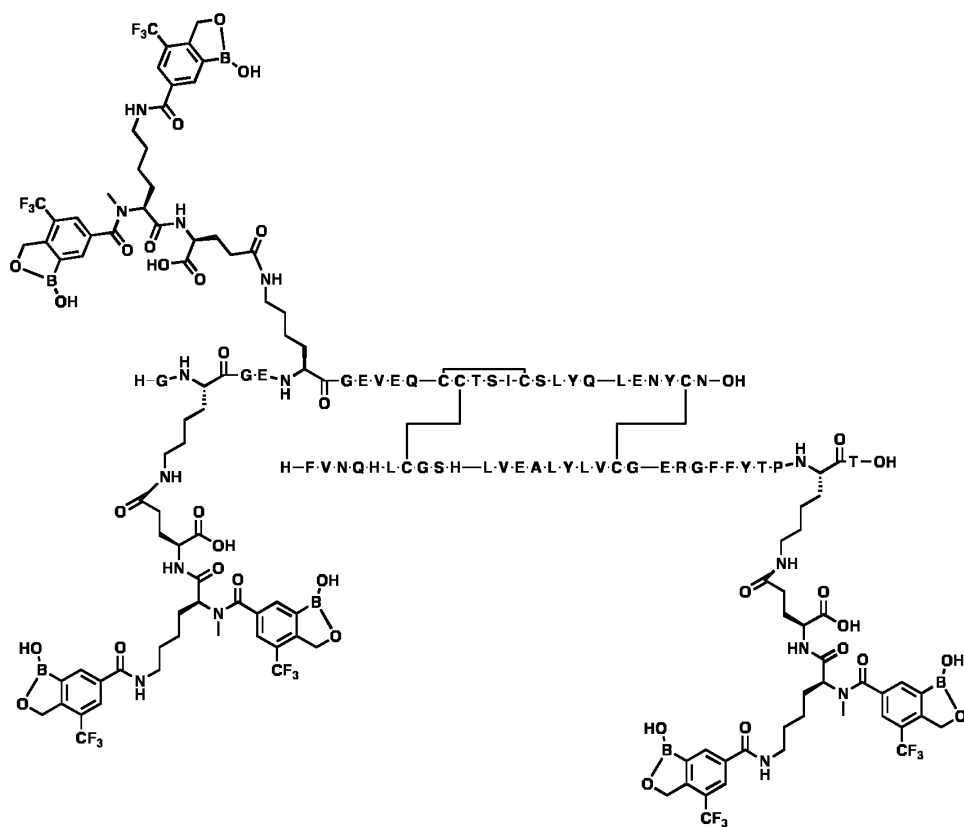
Example 724:



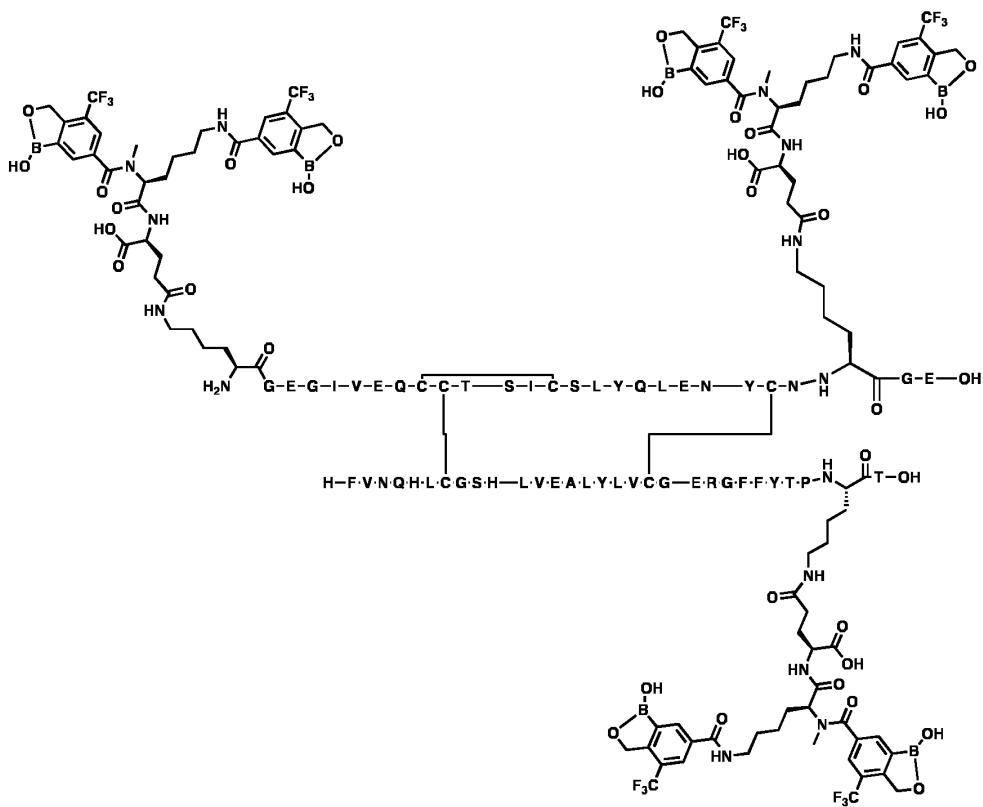
Example 725:



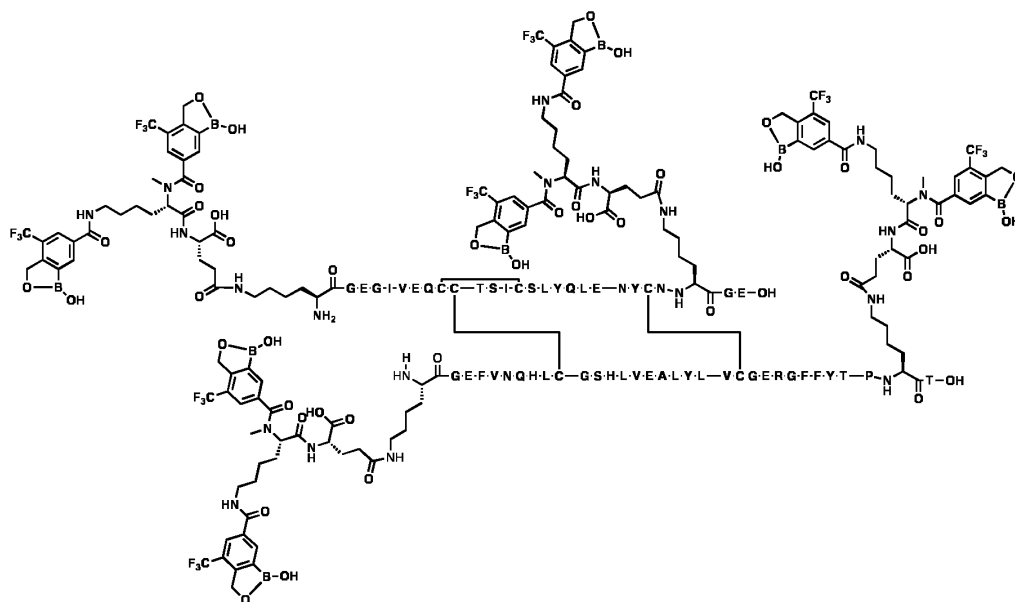
Example 726:



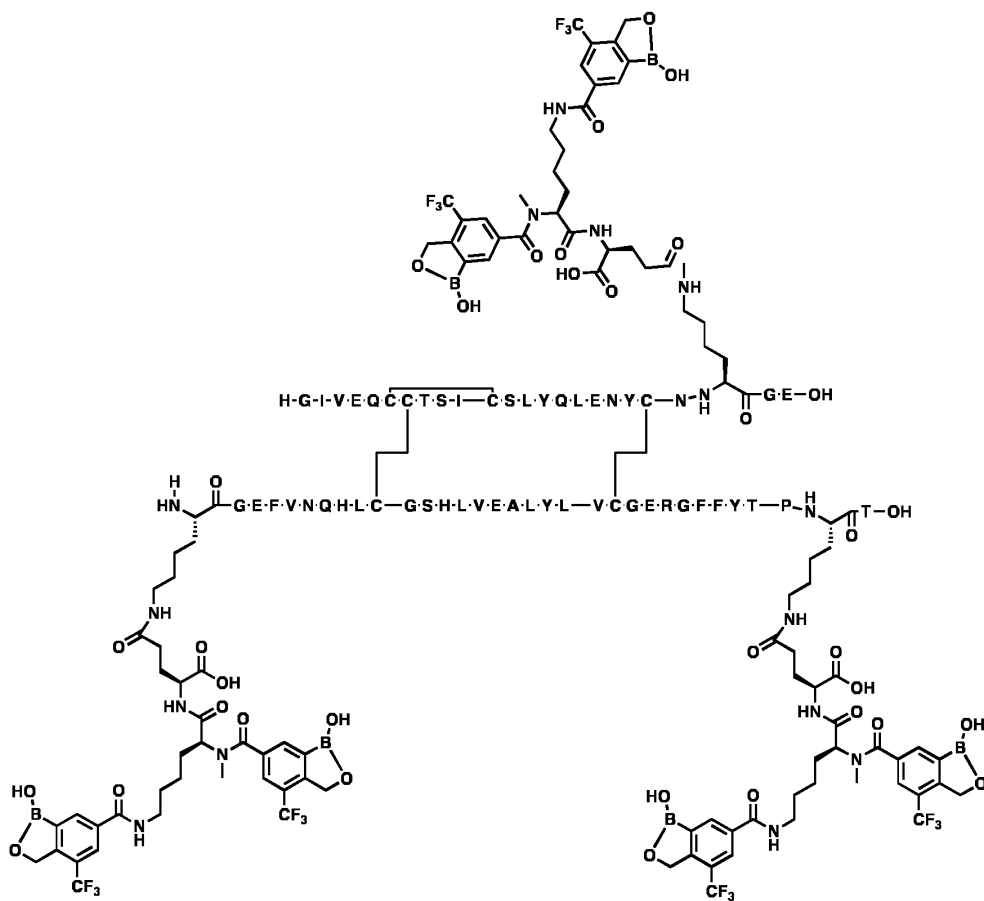
Example 727:



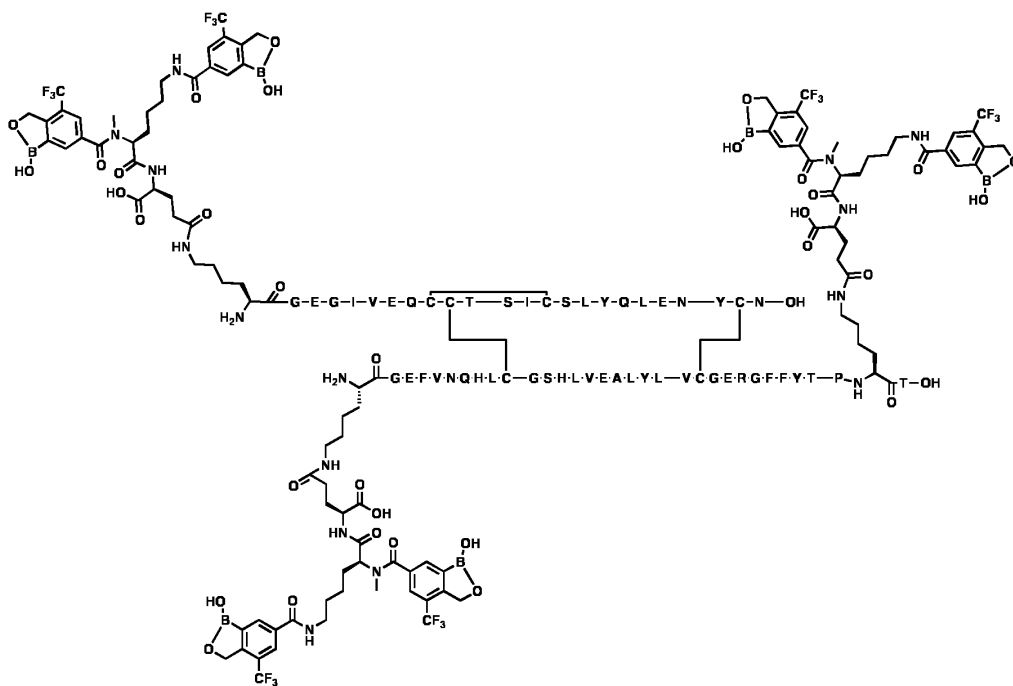
Example 728:

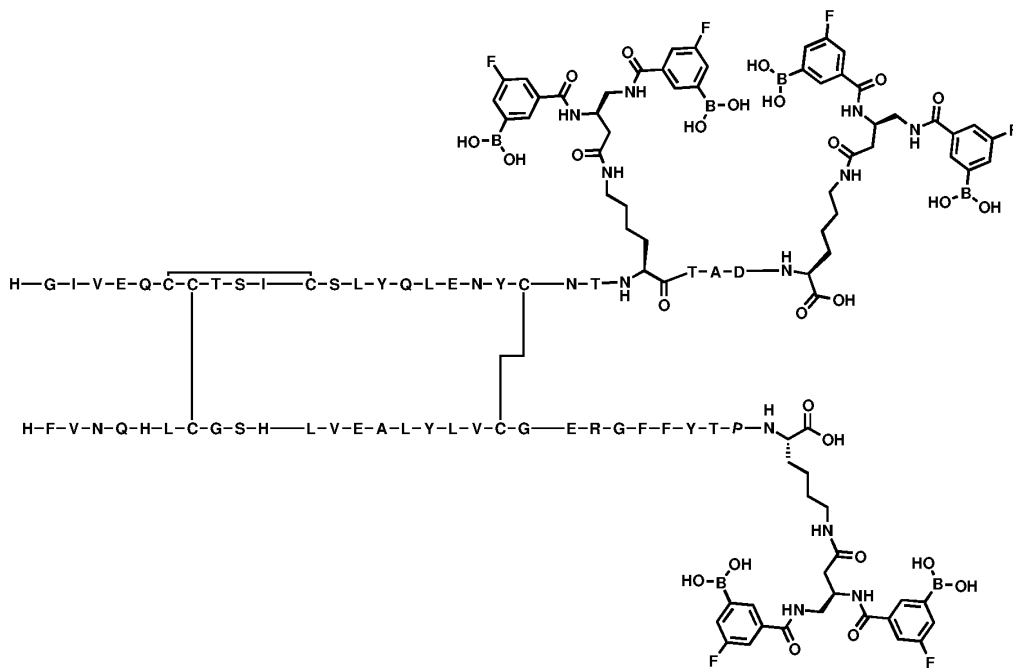
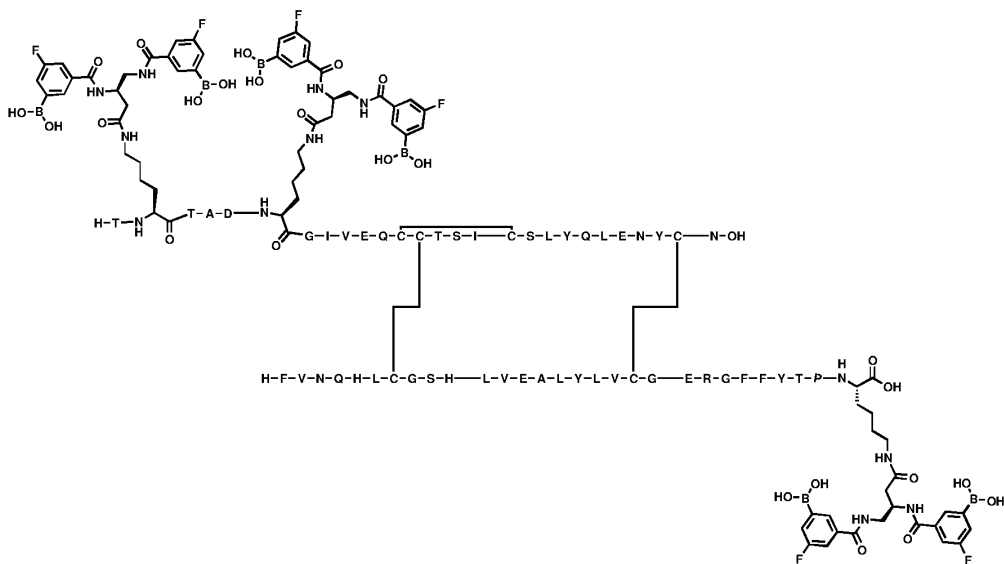


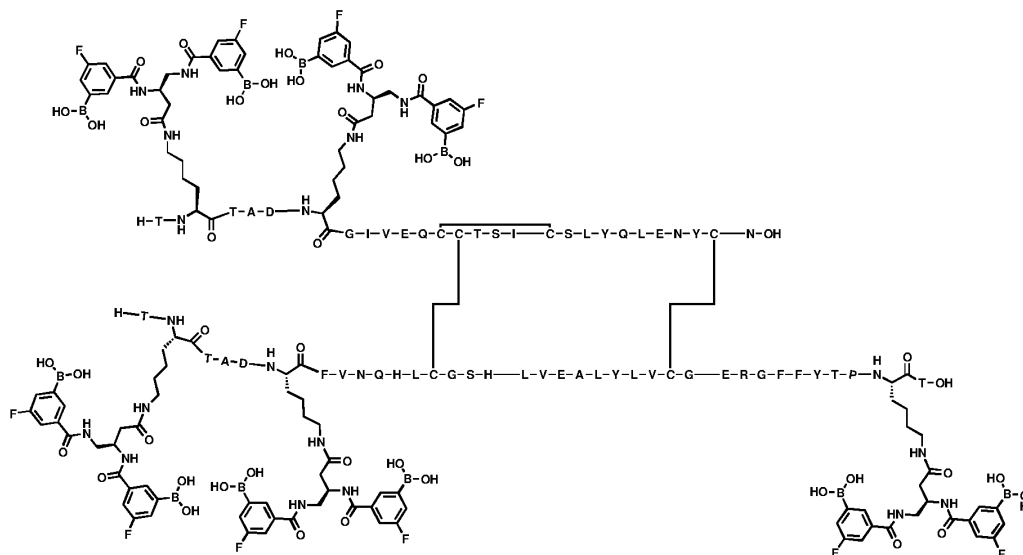
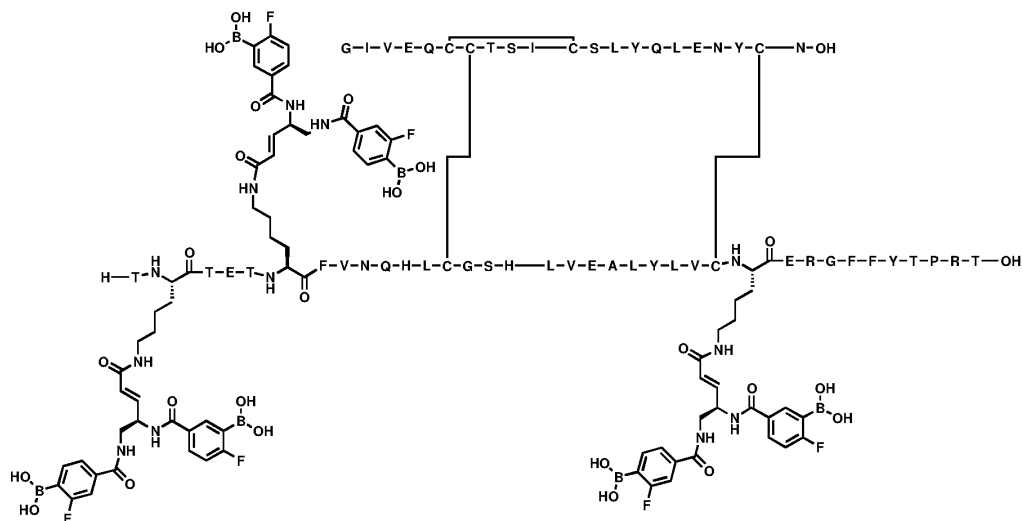
Example 729:

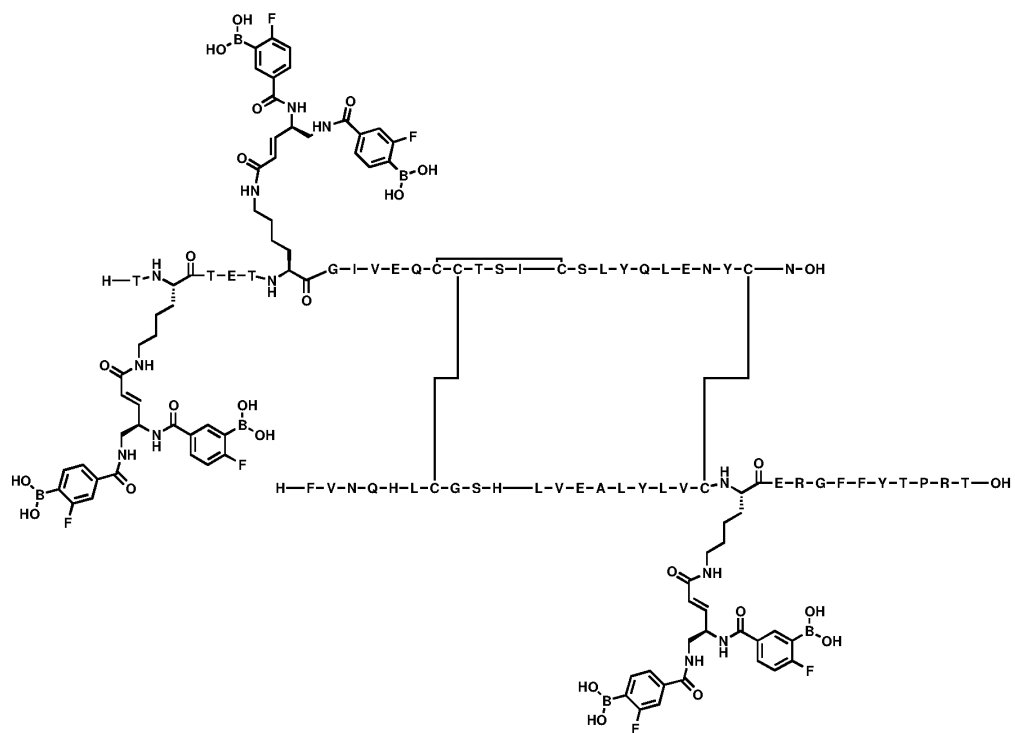


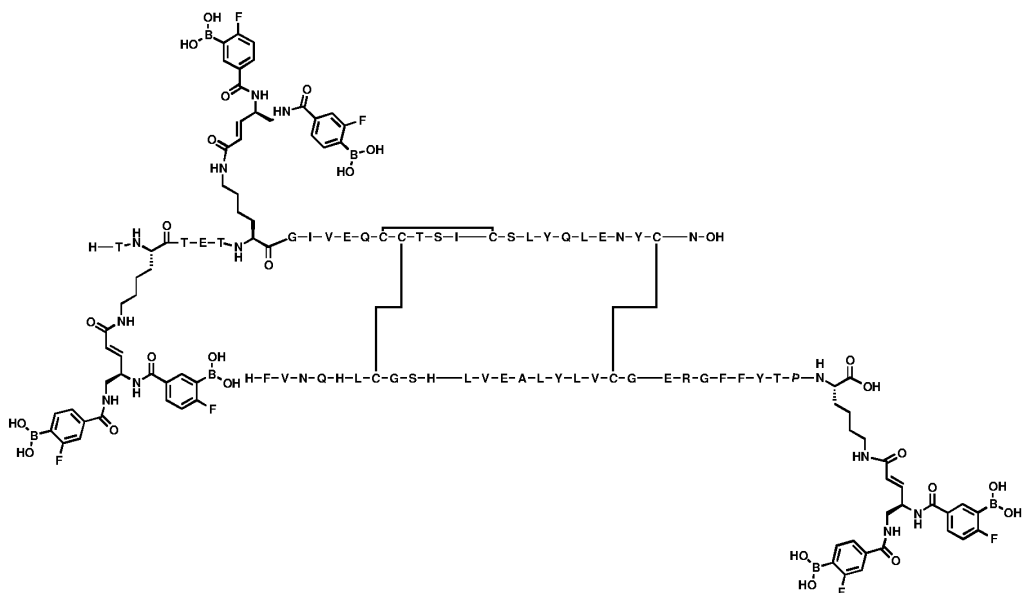
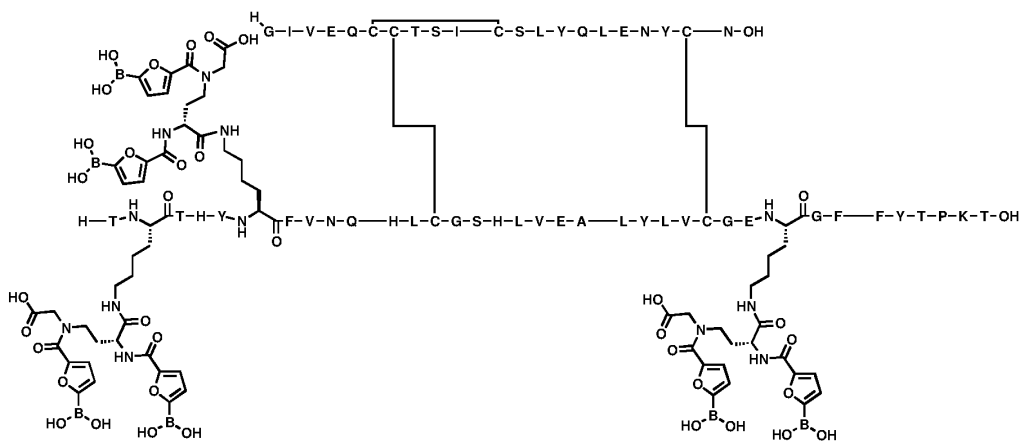
Example 730:

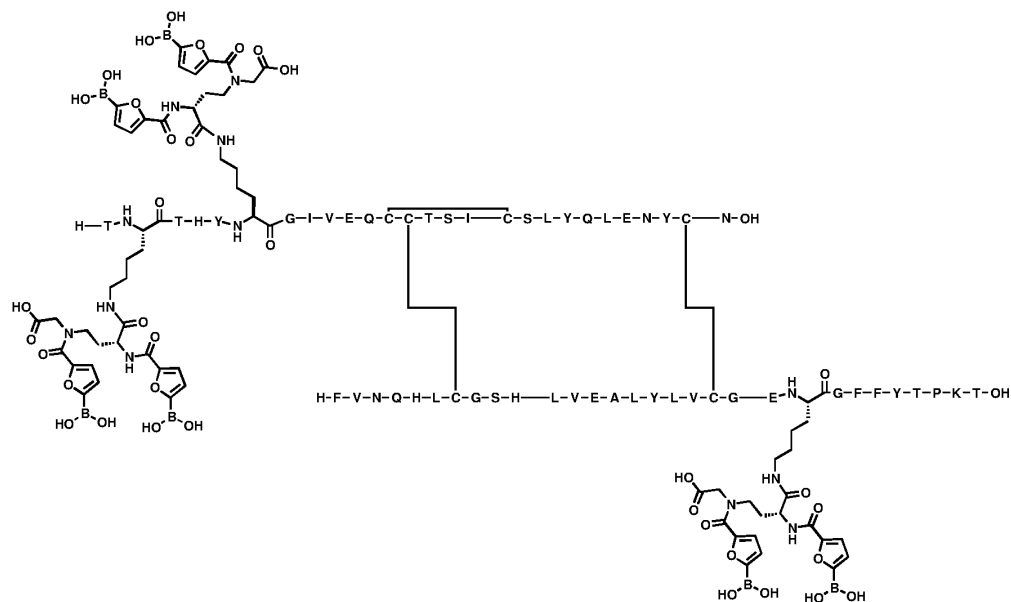


Example 731:**Example 732:**

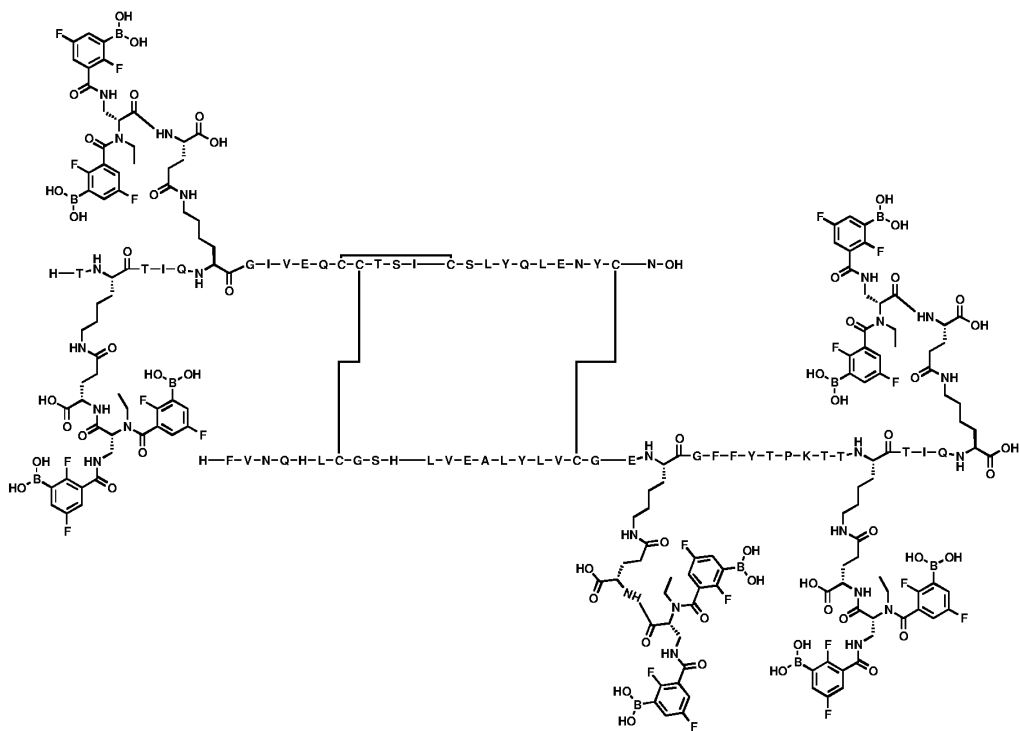
Example 733:**Example 734:**

Example 735:

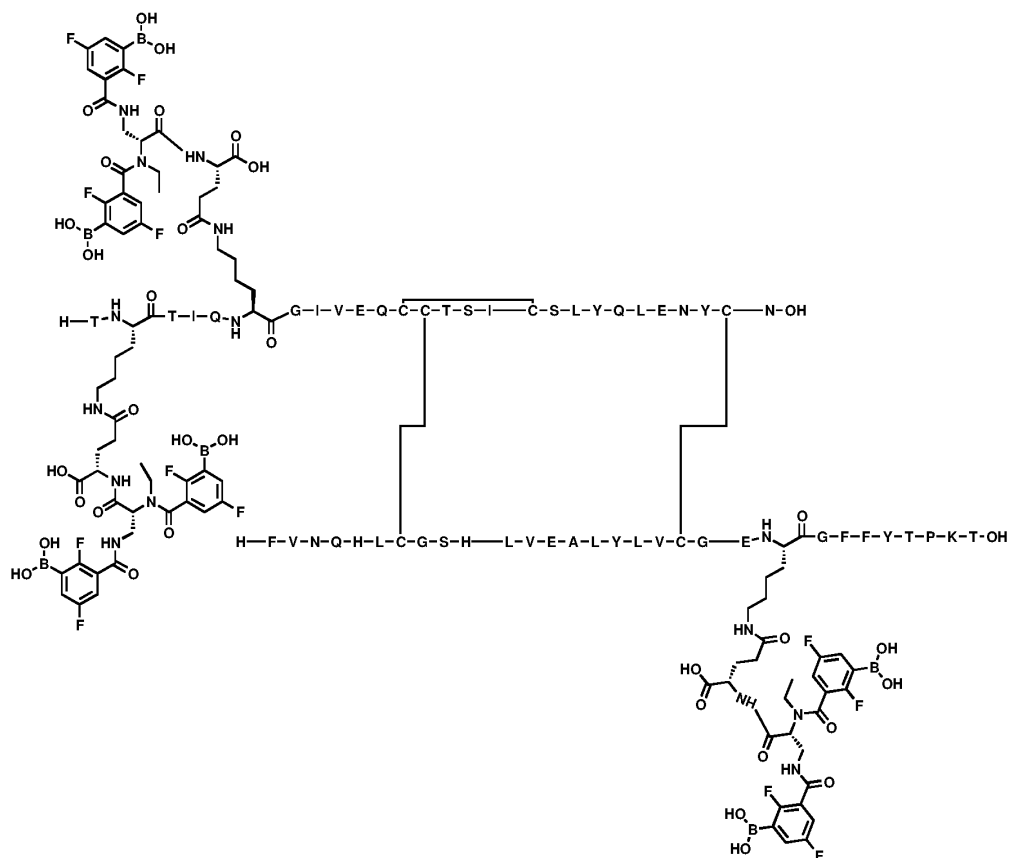
Example 736:**Example 737:**

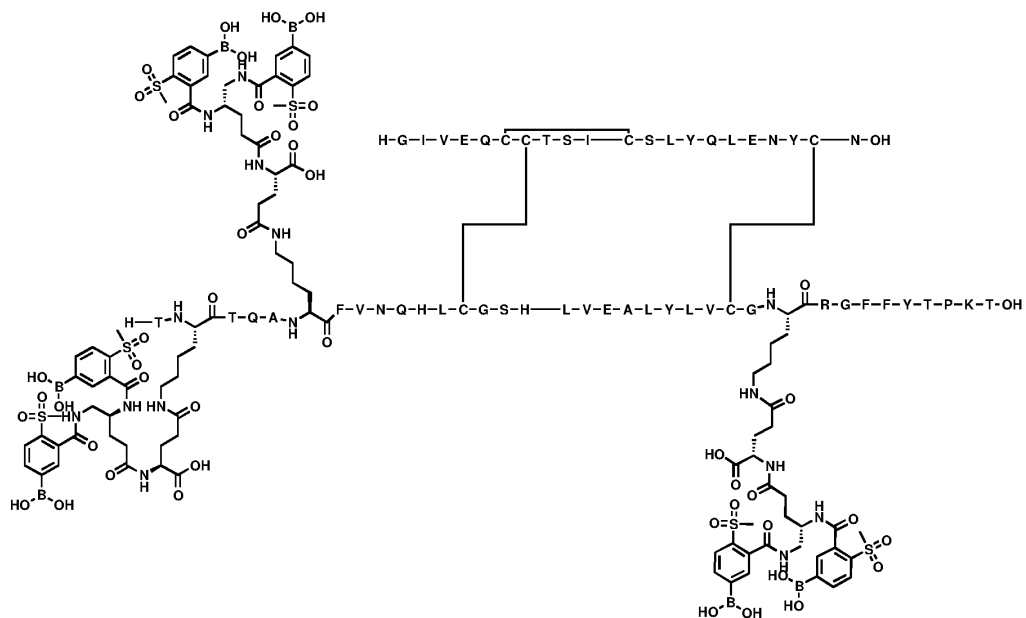
[illegible]

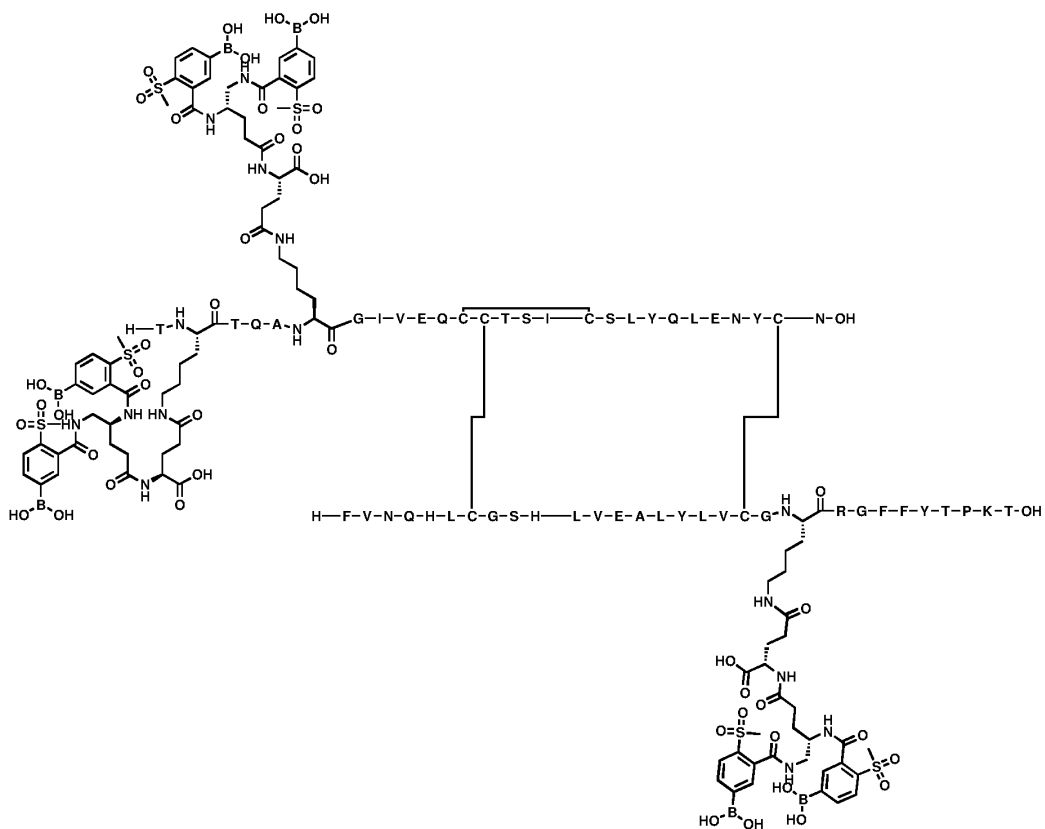
The chemical structure of the 120-120 protein is shown, a 120-residue protein. The structure is a complex fold, featuring a central alpha-helical bundle. The protein is composed of several domains, including a large N-terminal domain, a central alpha-helical bundle, and a C-terminal domain. The structure is highly symmetrical, with multiple fluorinated boronic acid groups (F-B(OH)2) attached to the protein backbone. The structure is shown in a ribbon representation, with the main chain in black and the side chains in gray. The structure is highly symmetrical, with multiple fluorinated boronic acid groups (F-B(OH)2) attached to the protein backbone. The structure is shown in a ribbon representation, with the main chain in black and the side chains in gray. The structure is highly symmetrical, with multiple fluorinated boronic acid groups (F-B(OH)2) attached to the protein backbone.

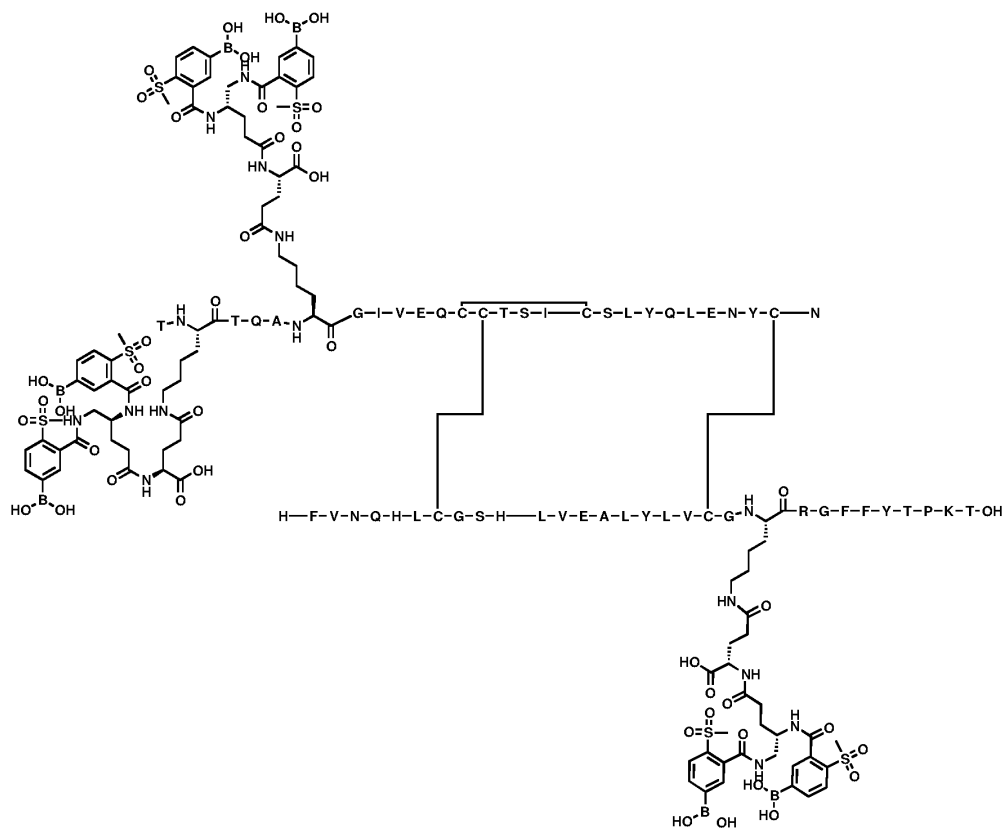
Example 741:

Example 742:

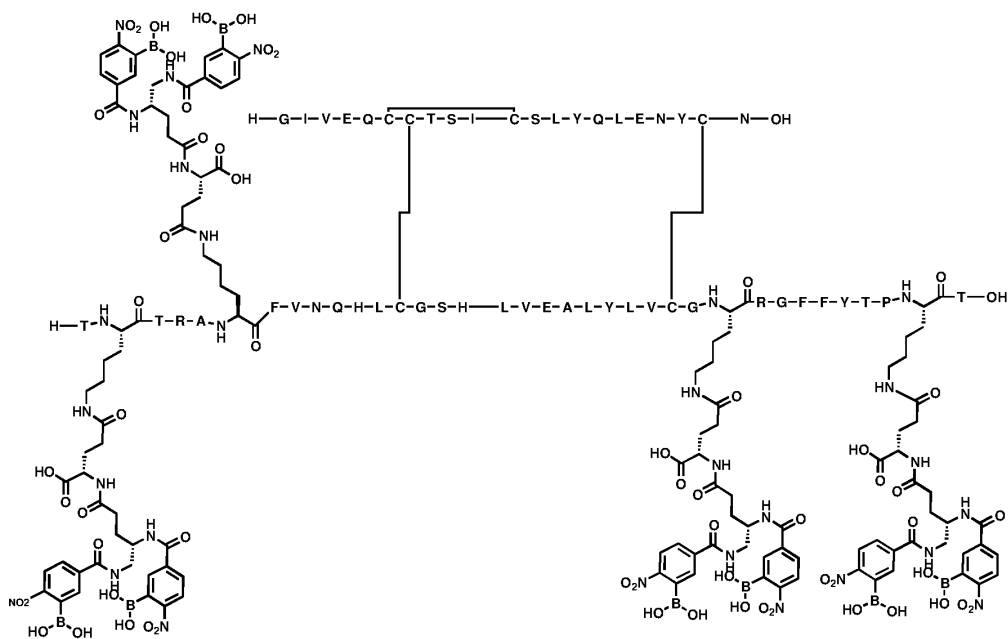


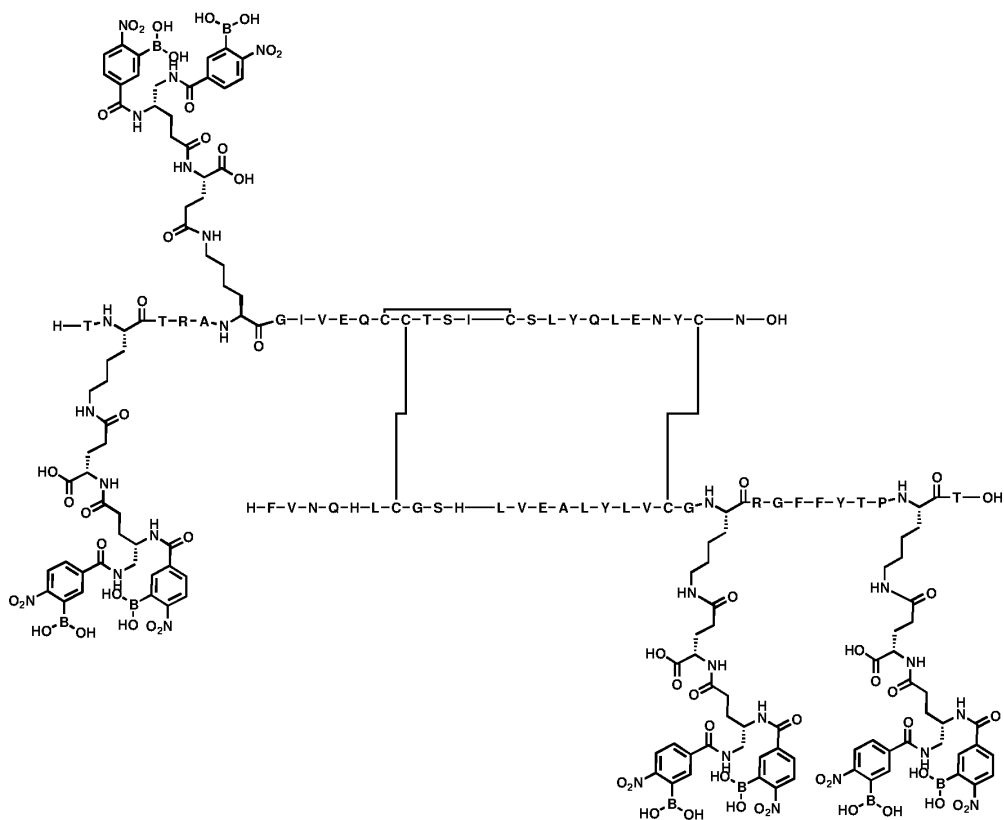
Example 743:

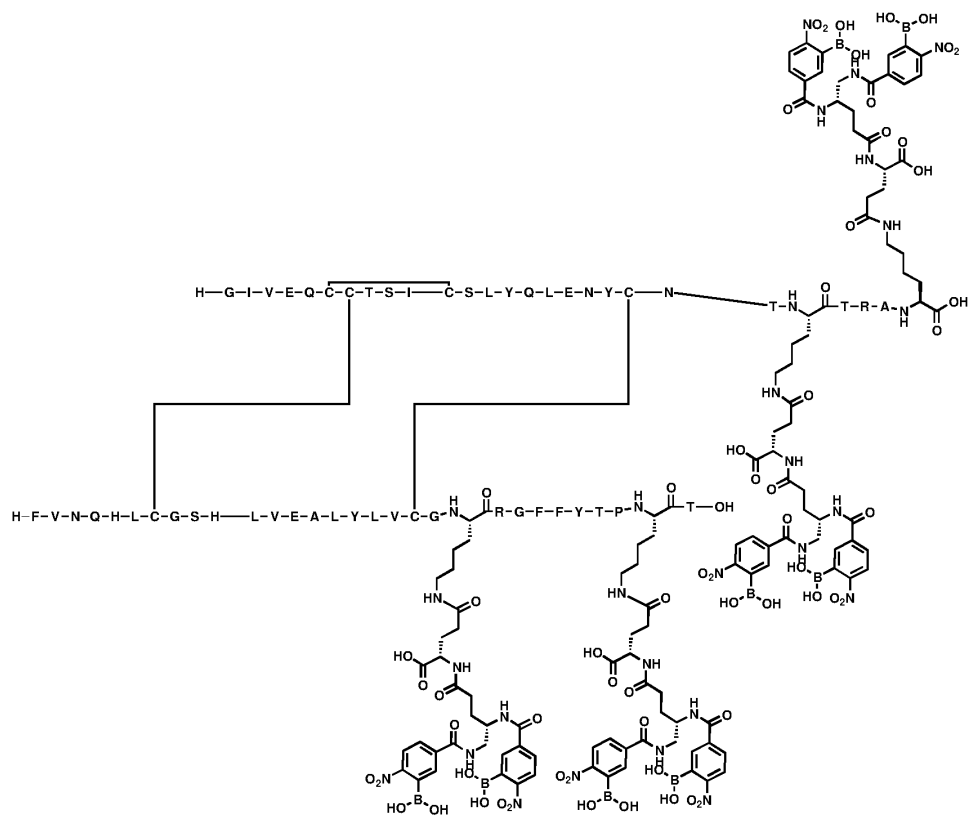
Example 744:

Example 745:

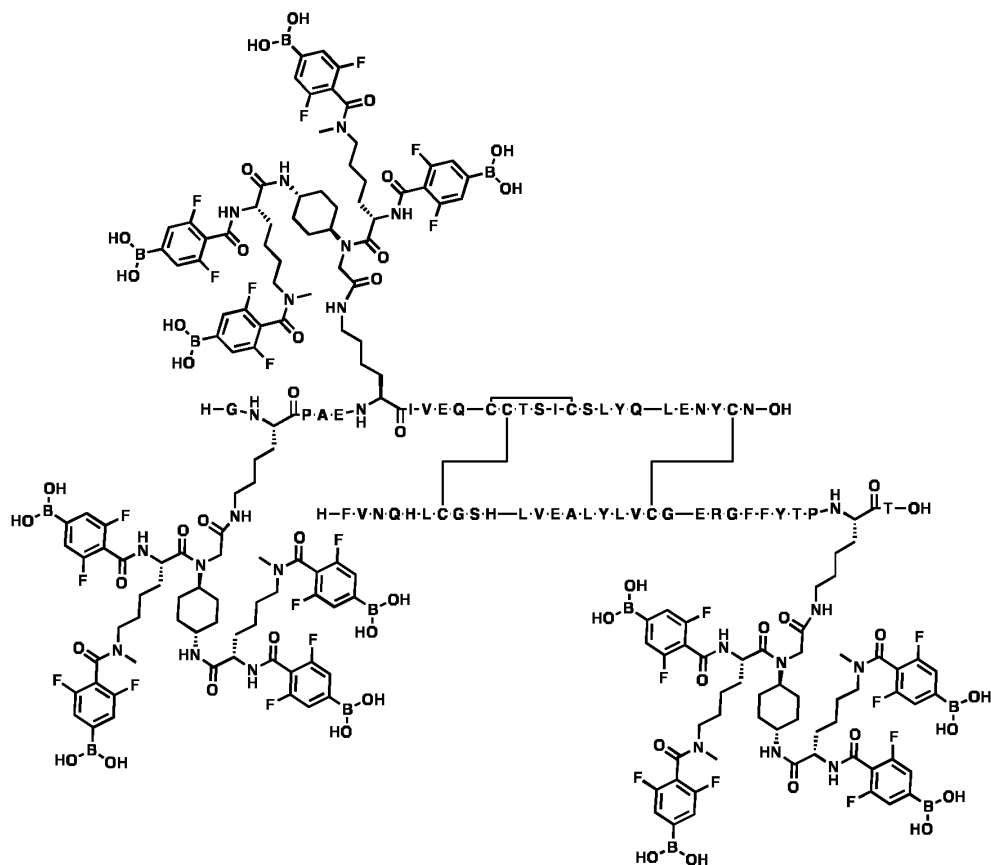
Example 746:



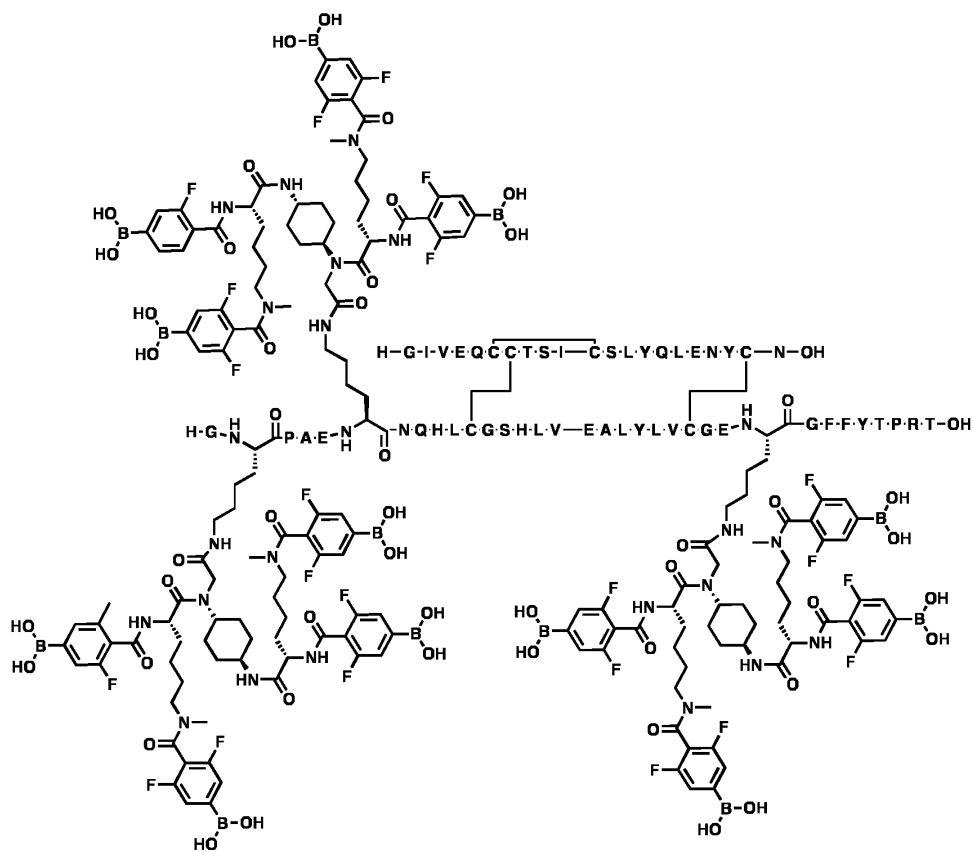
Example 747:

Example 748:

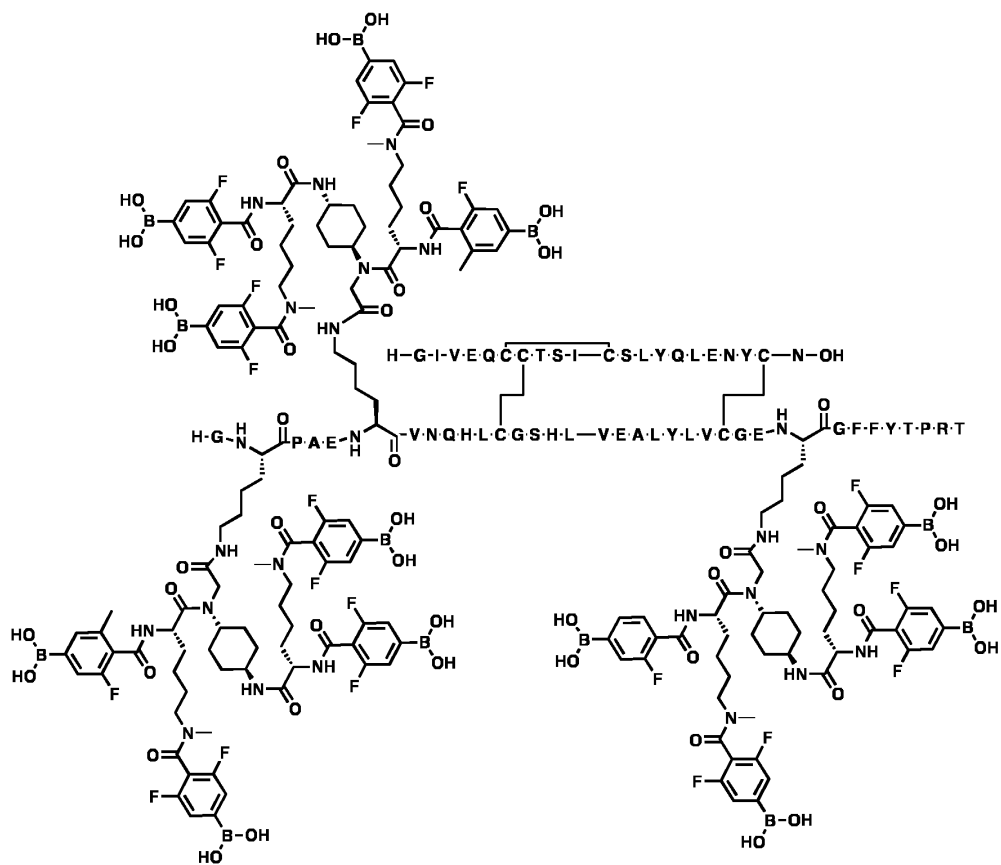
Example 749:



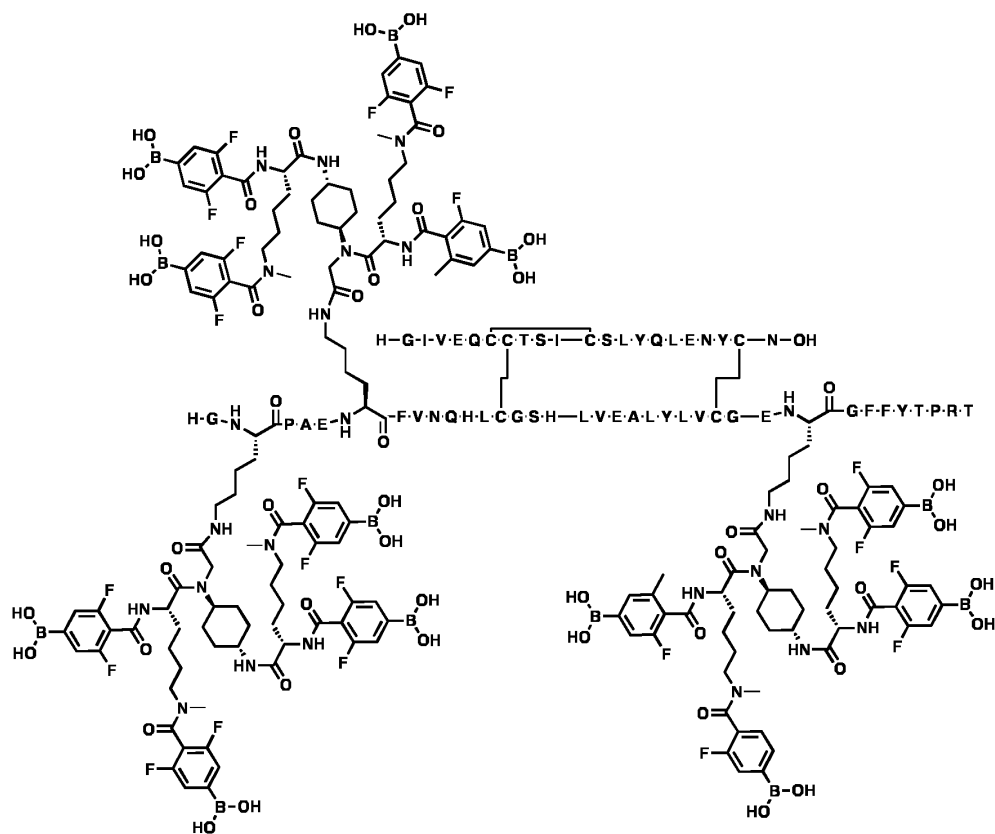
Example 750:



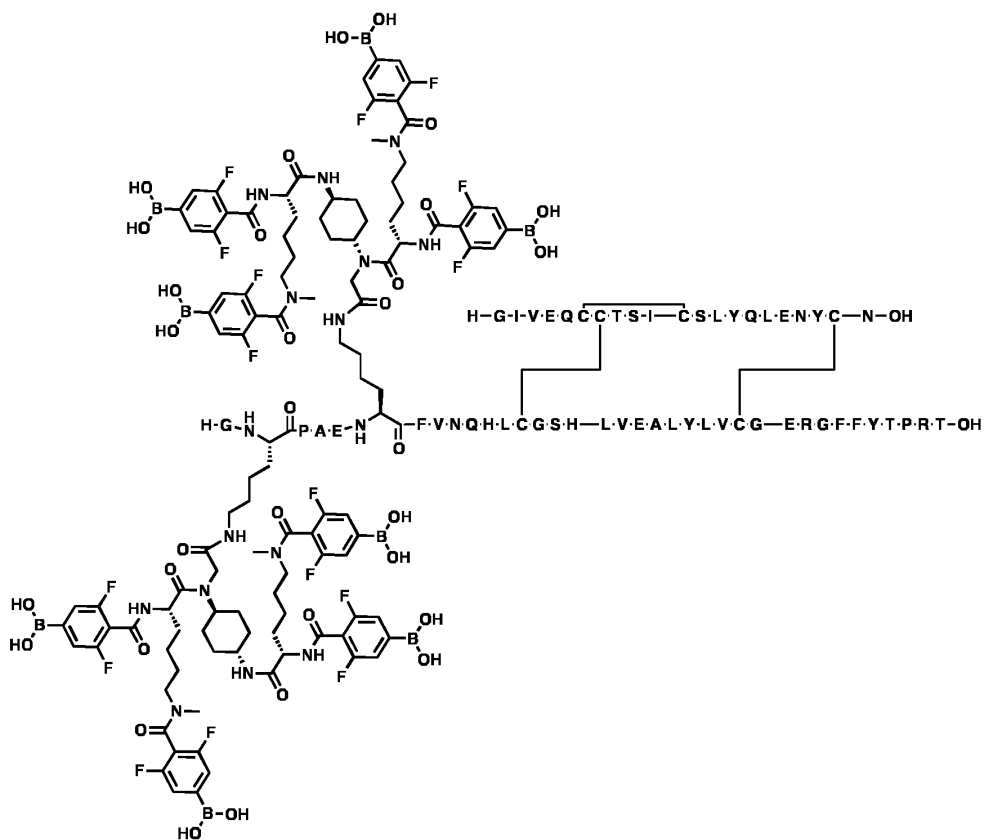
Example 751:



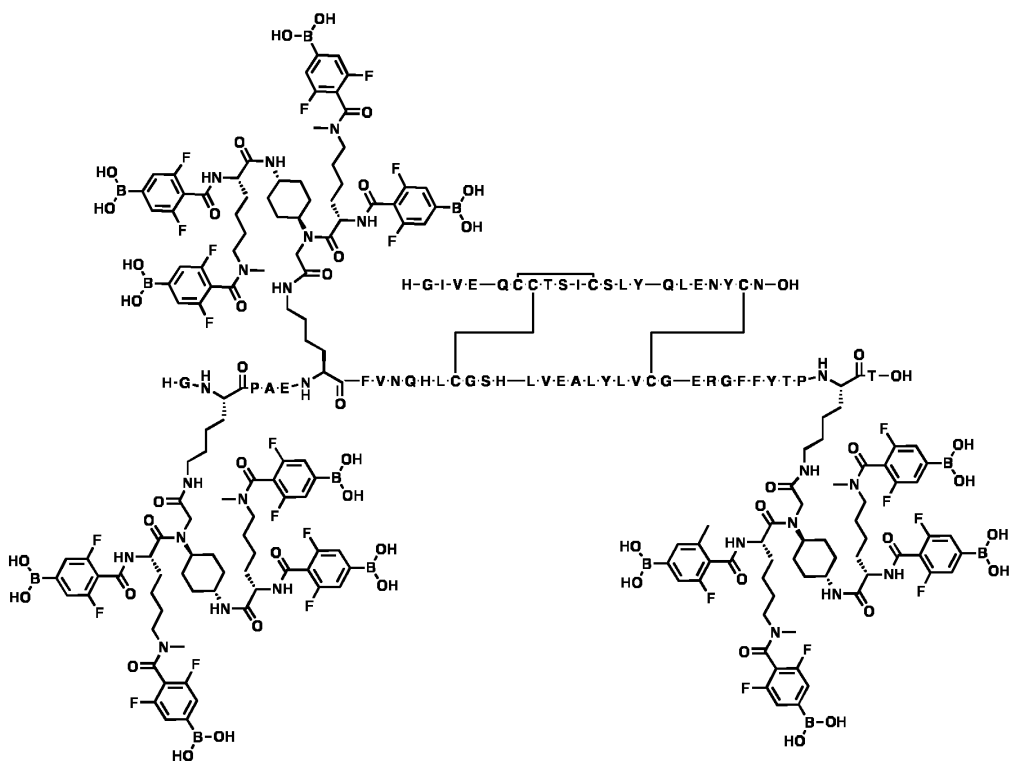
Example 752:



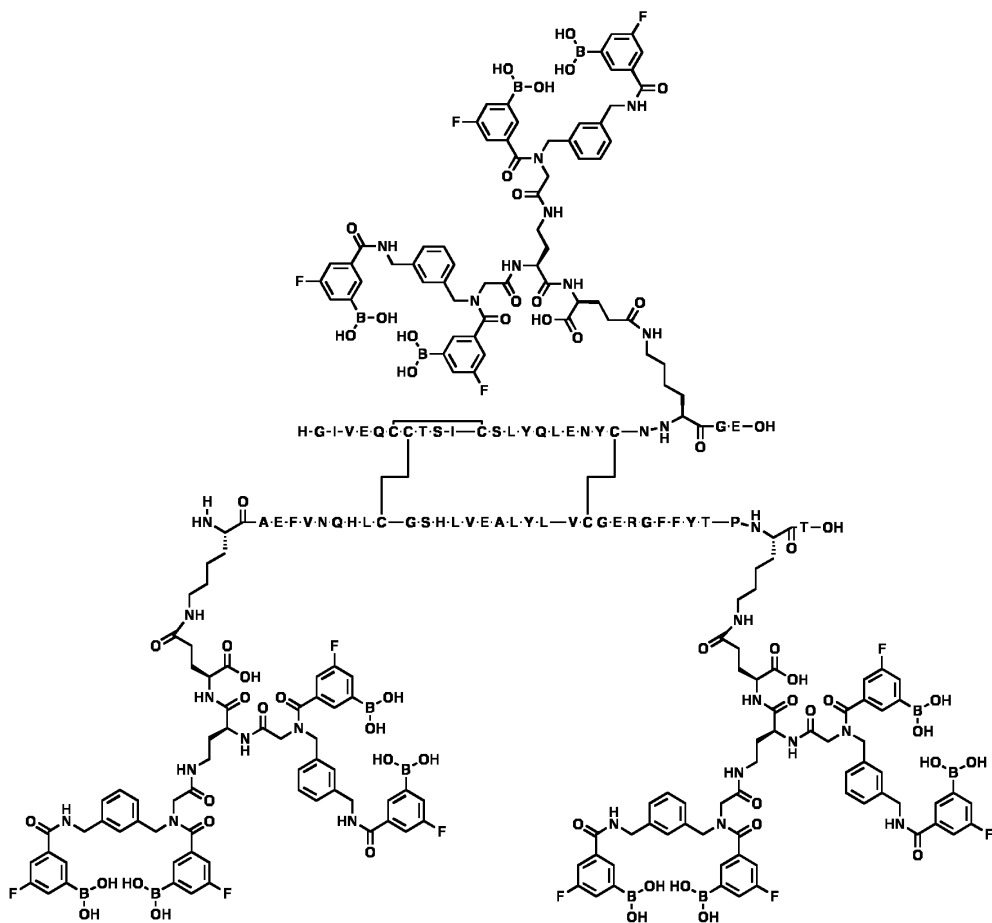
Example 753:



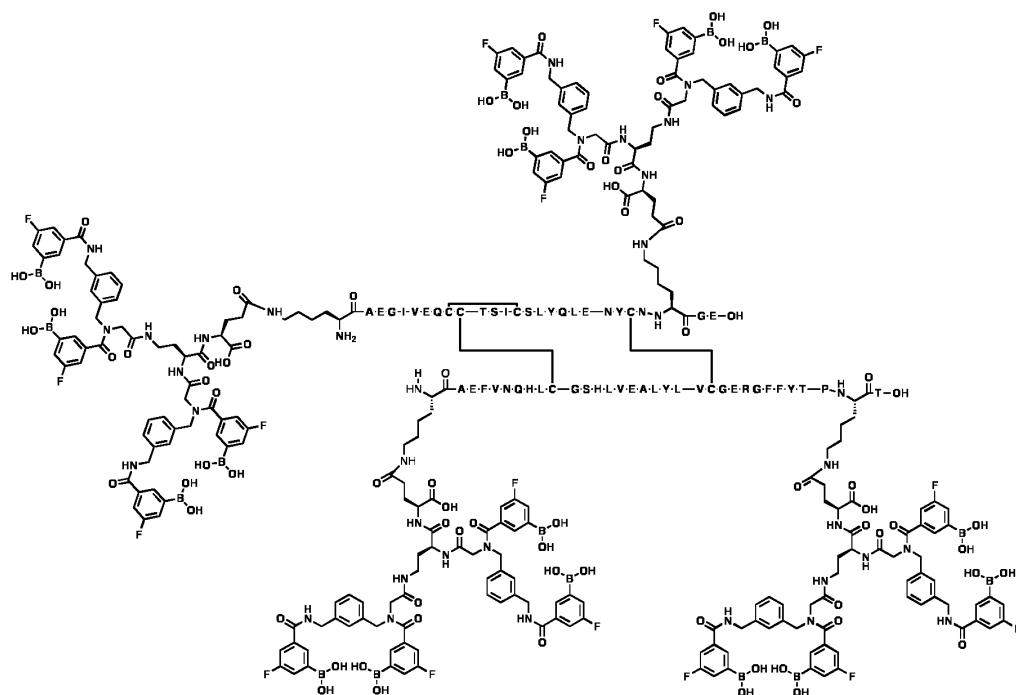
Example 754:



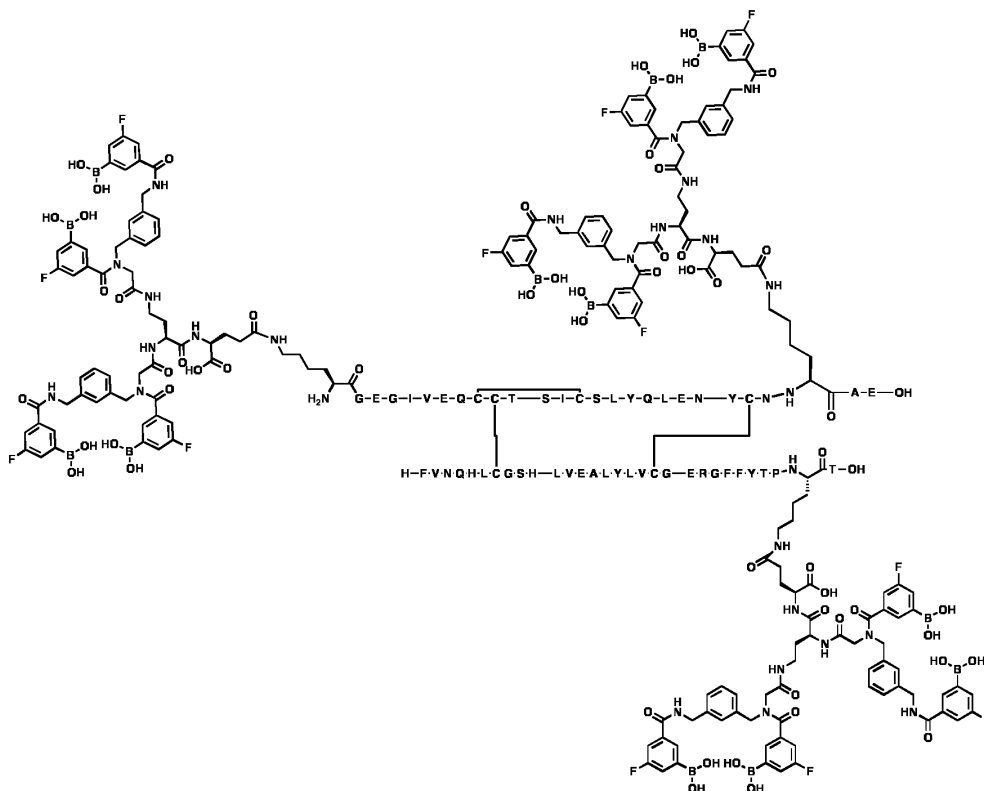
Example 755:



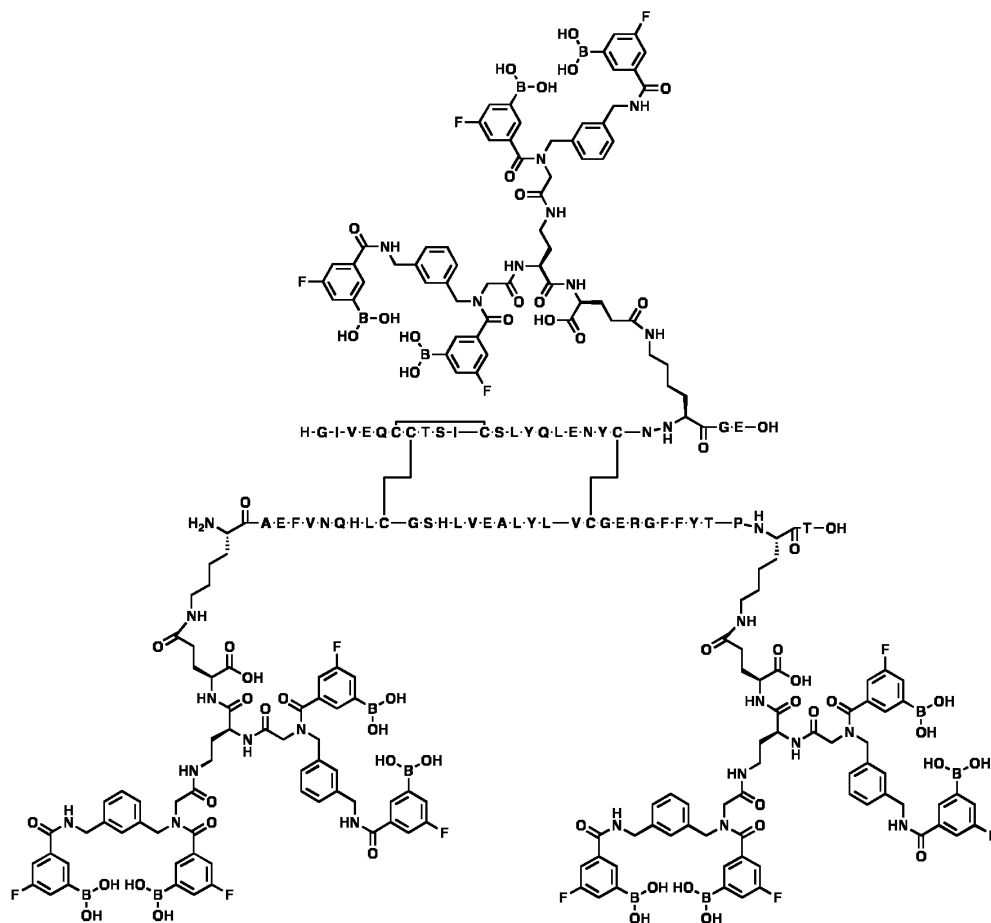
Example 756:



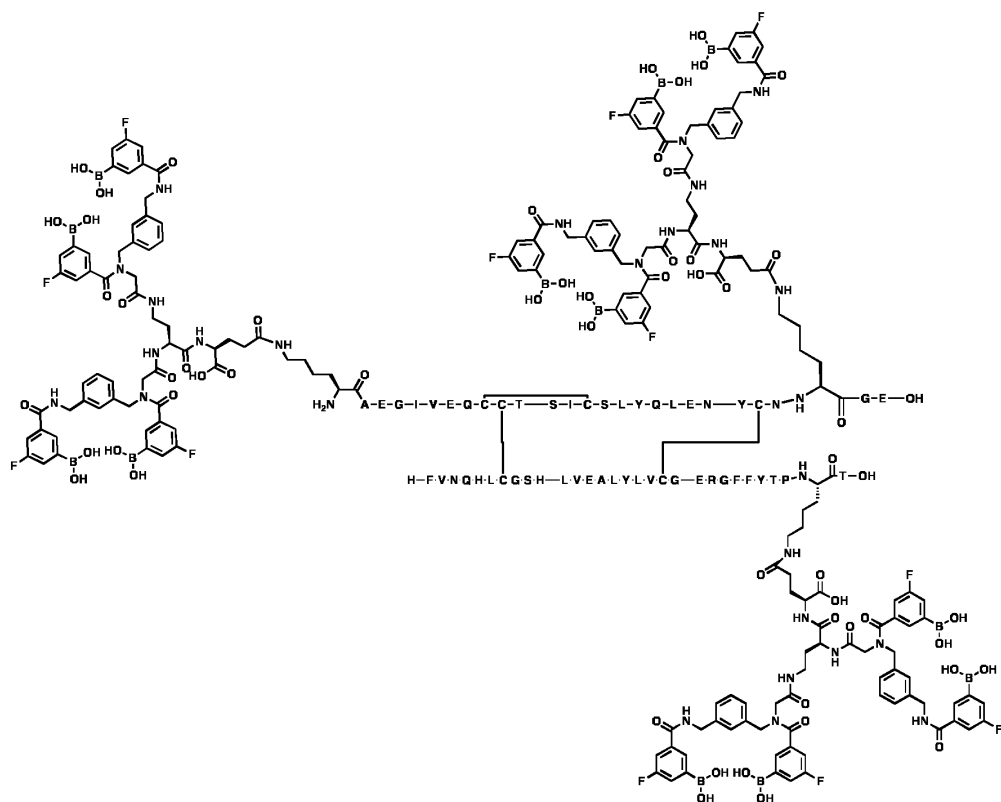
Example 757:



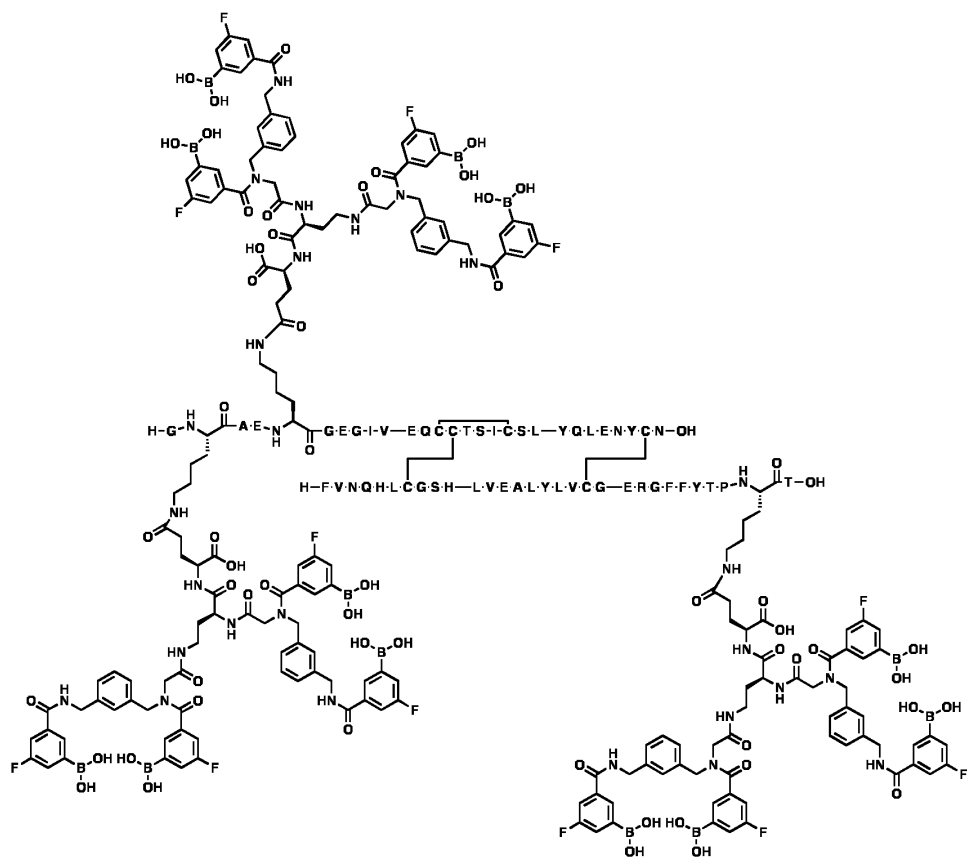
Example 758:



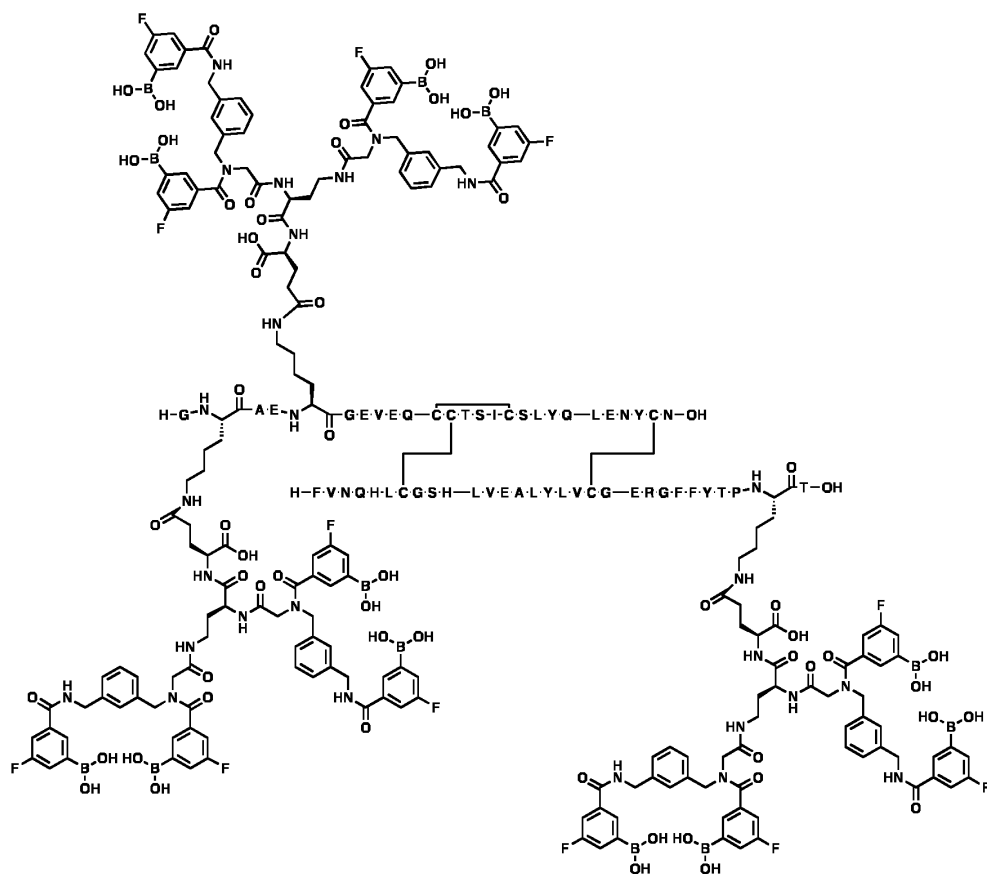
Example 759:

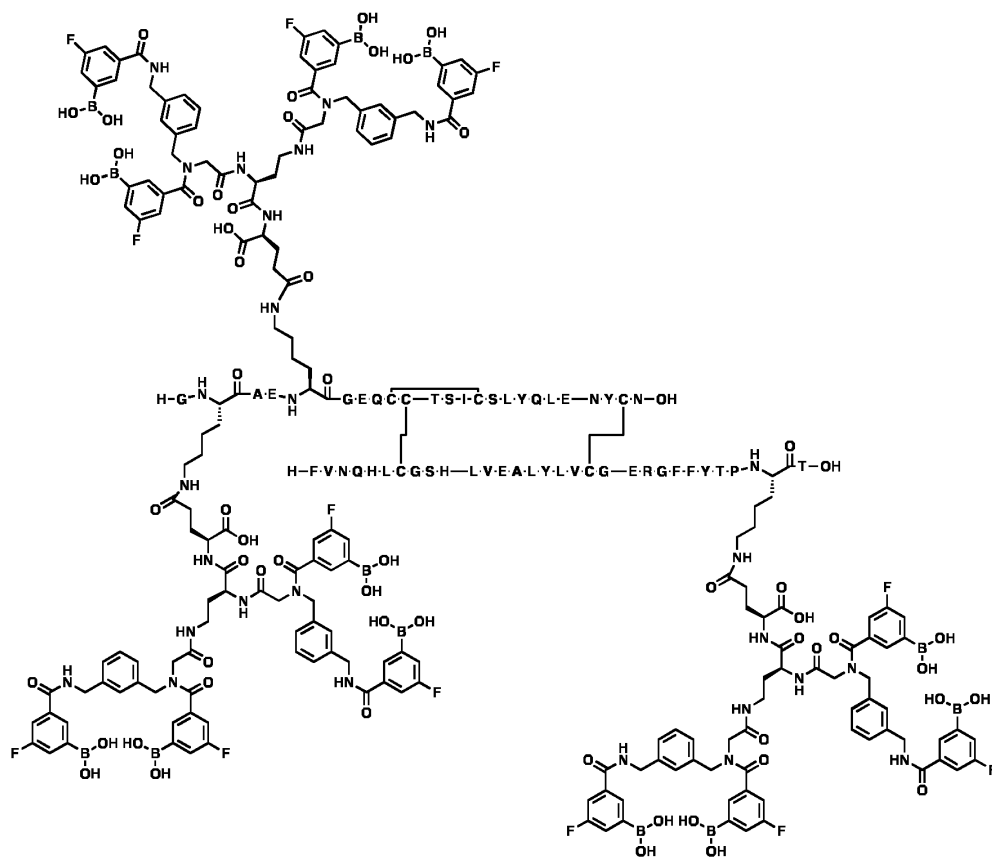


Example 760:

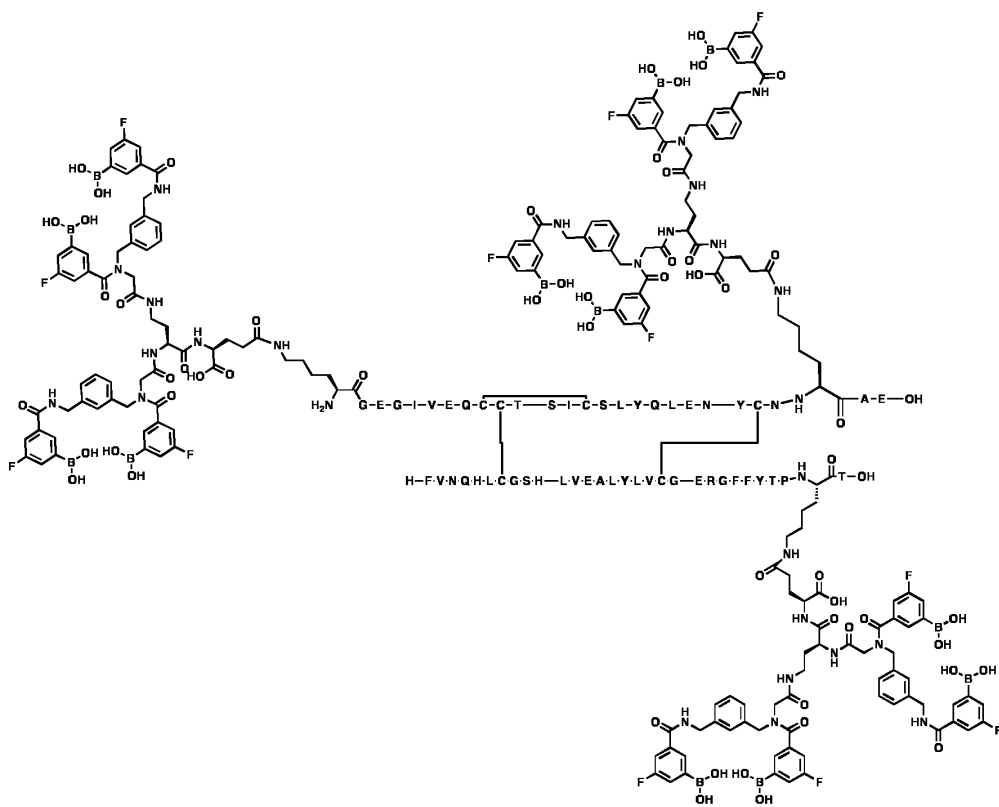


Example 761:

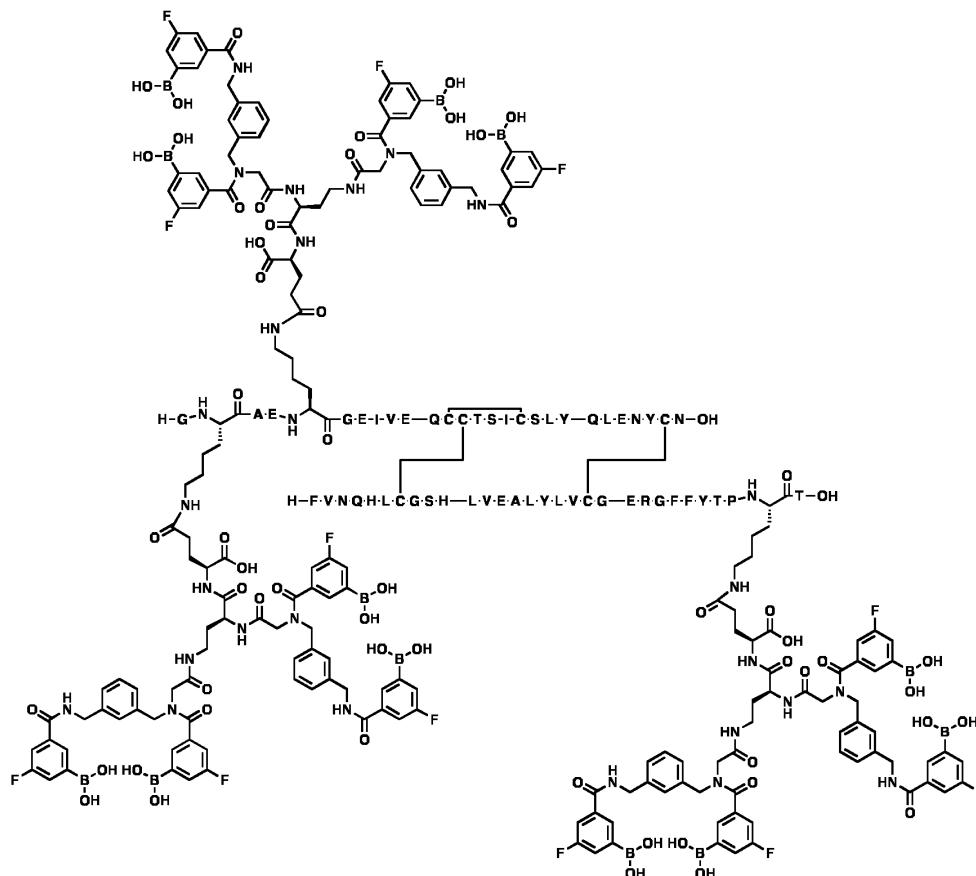


Example 762:

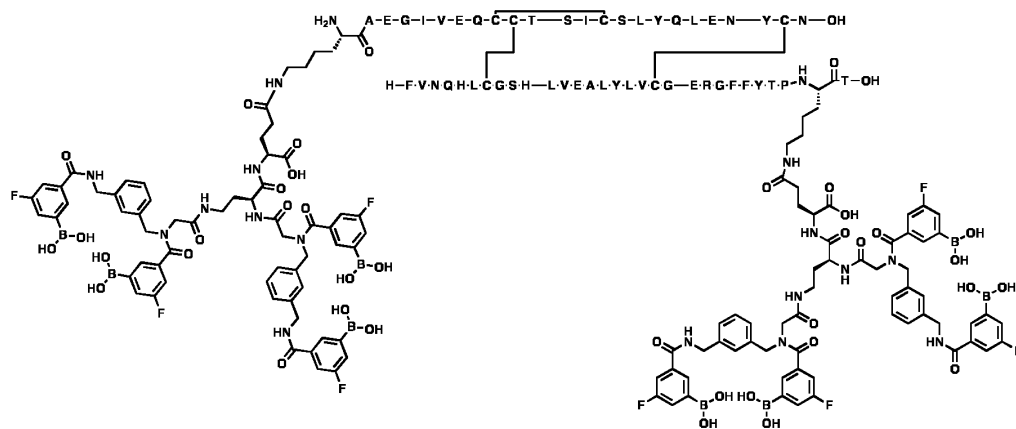
Example 763:



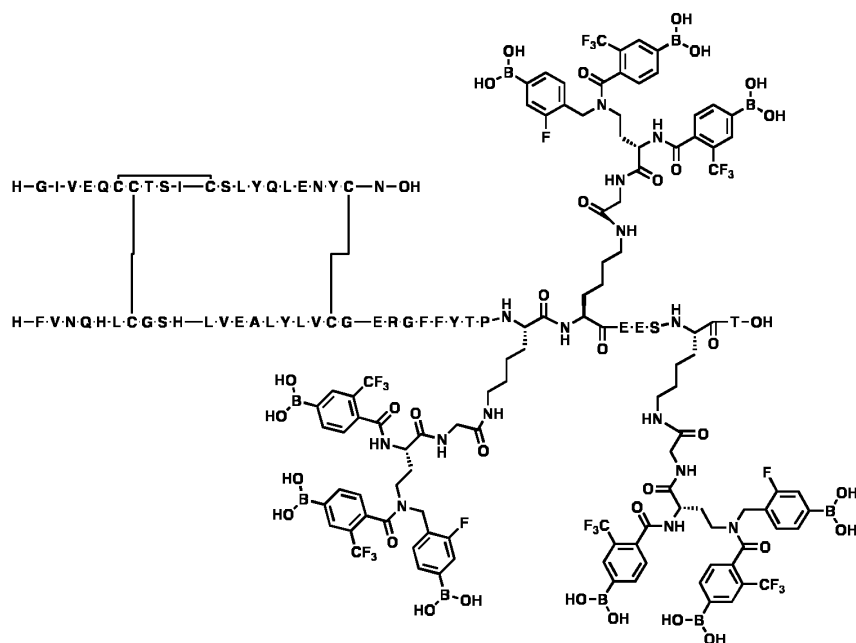
Example 764:



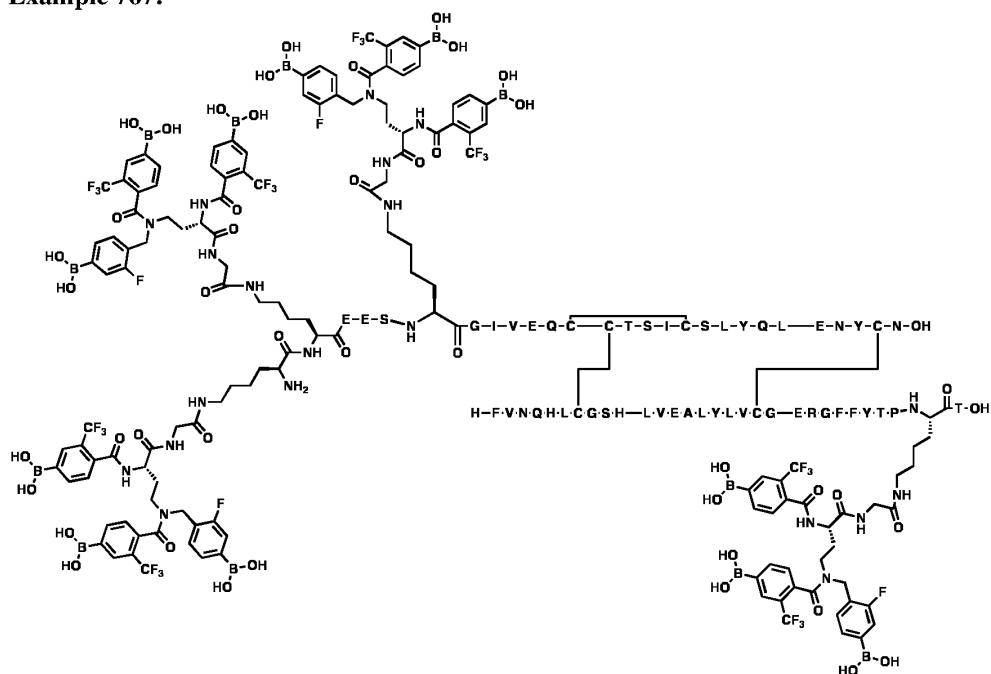
Example 765:



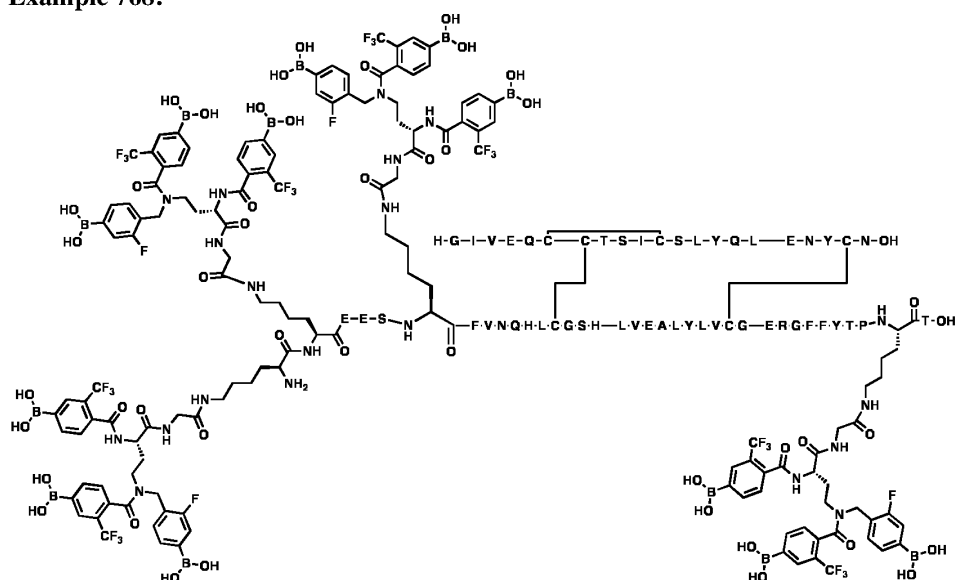
Example 766:



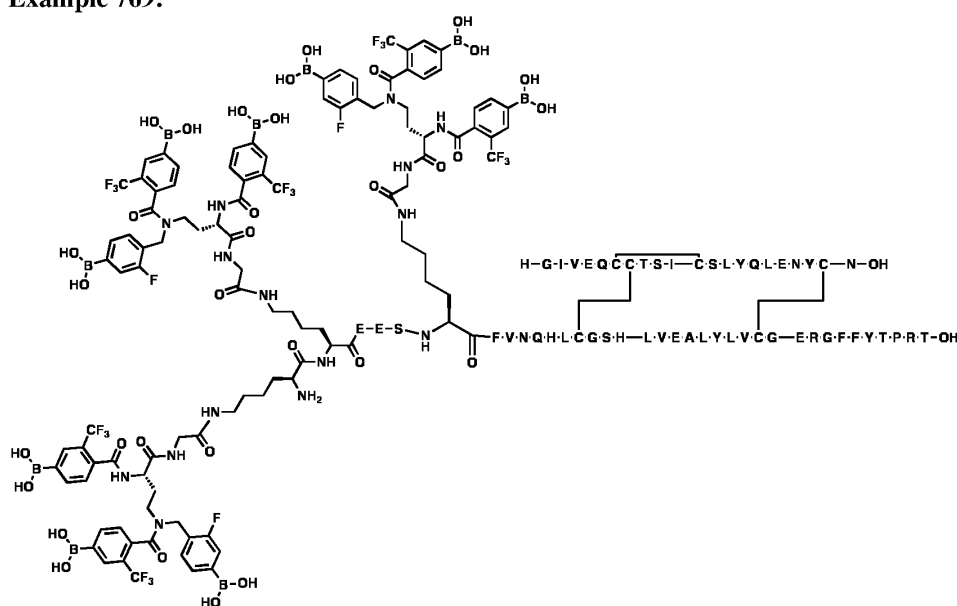
5 Example 767:



Example 768:



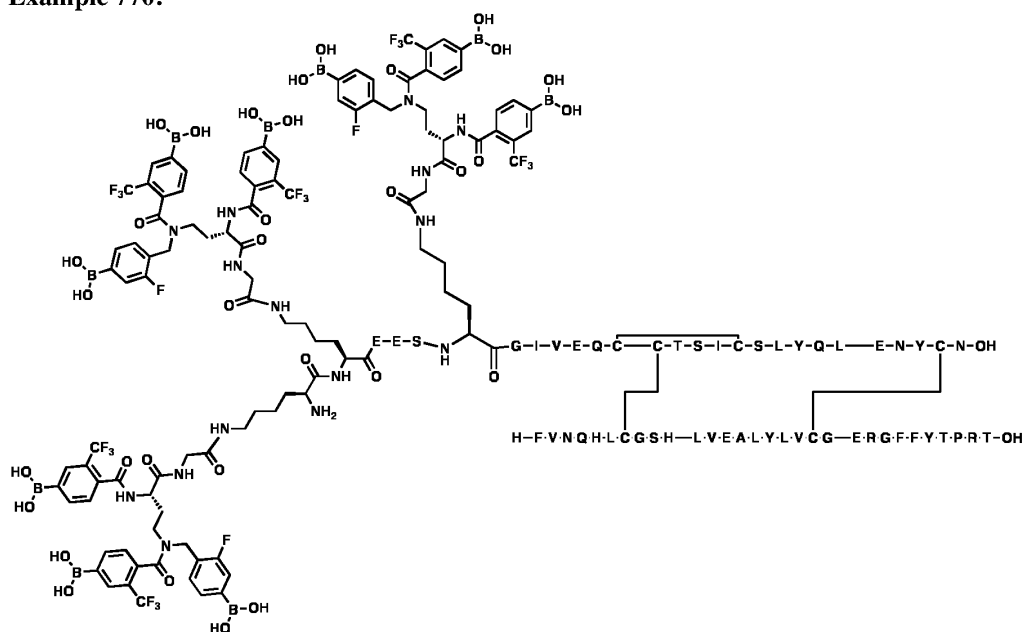
Example 769:



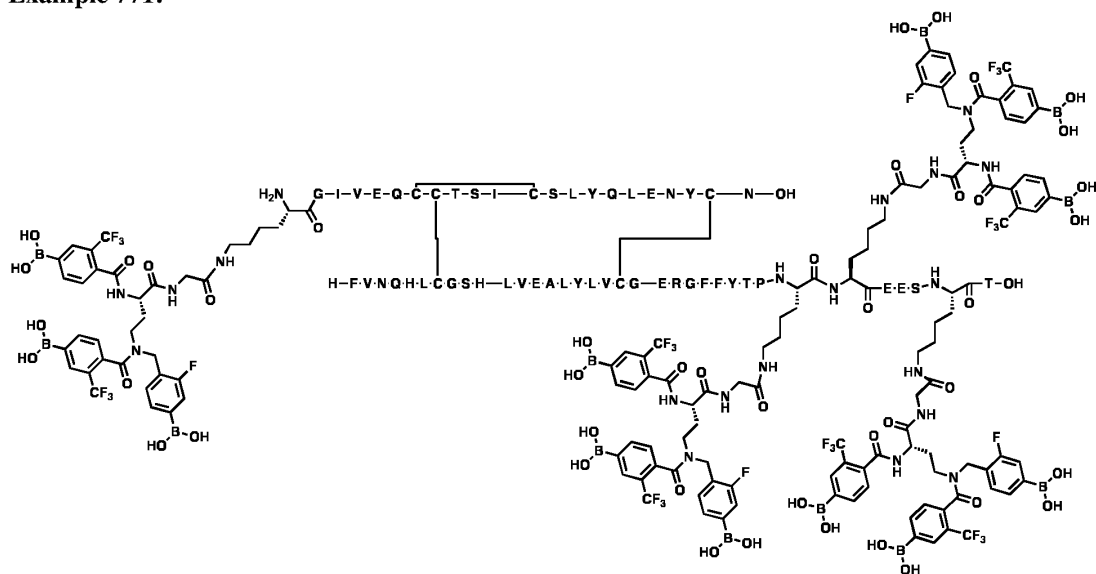
5

10

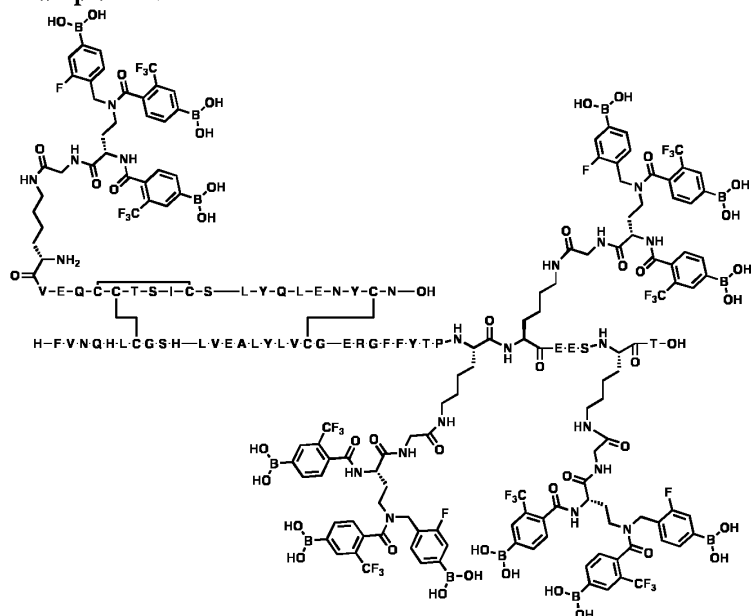
Example 770:



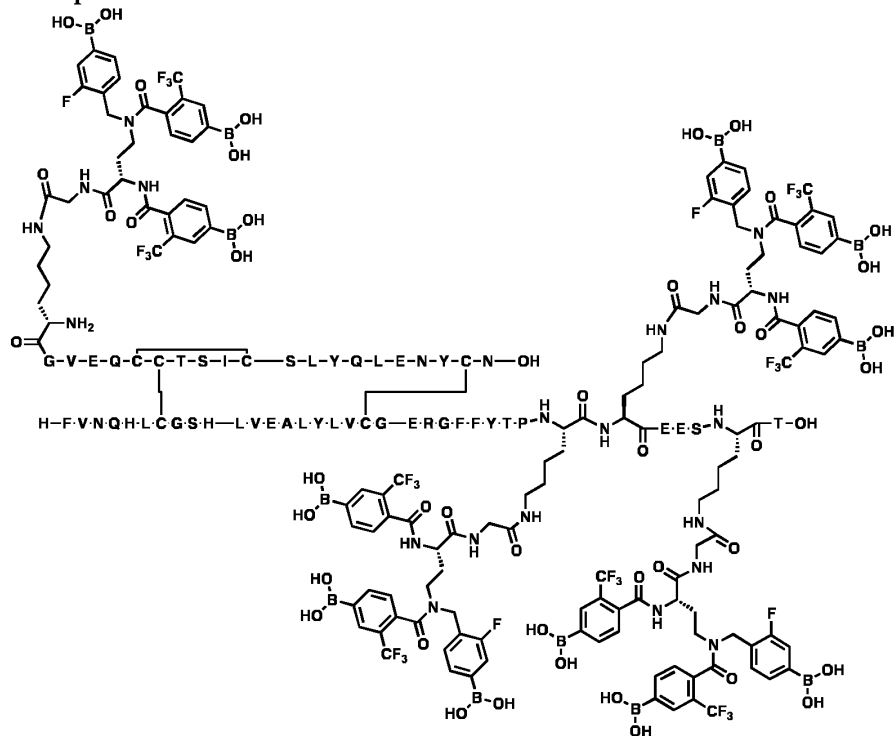
Example 771:



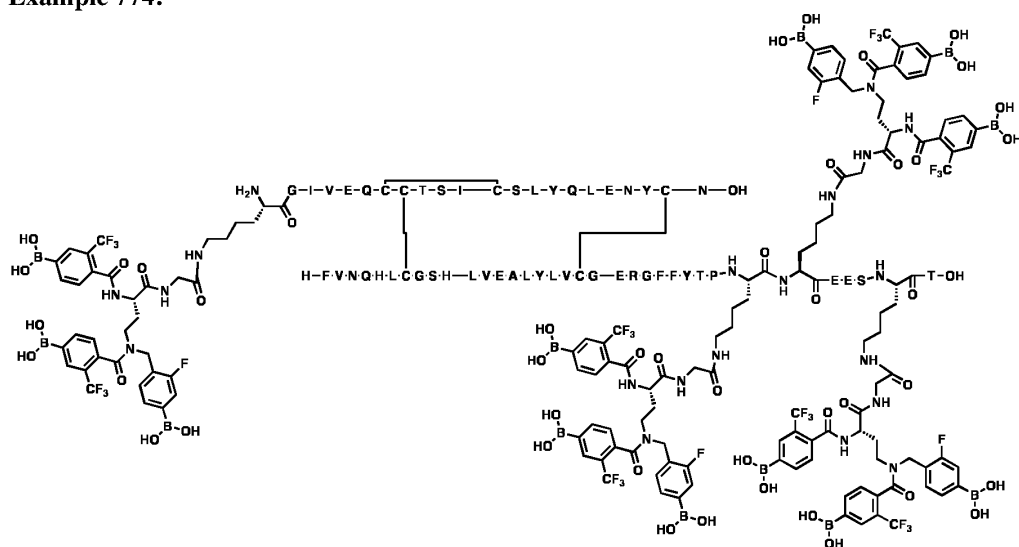
Example 772:



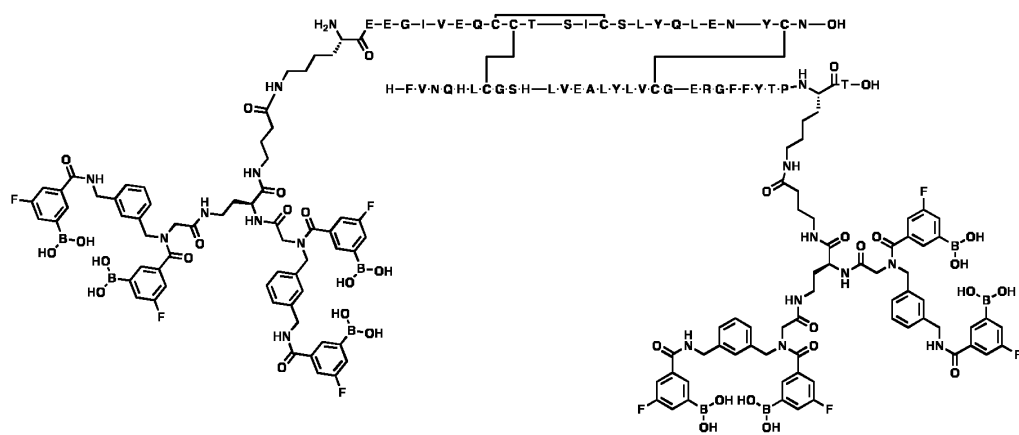
Example 773:



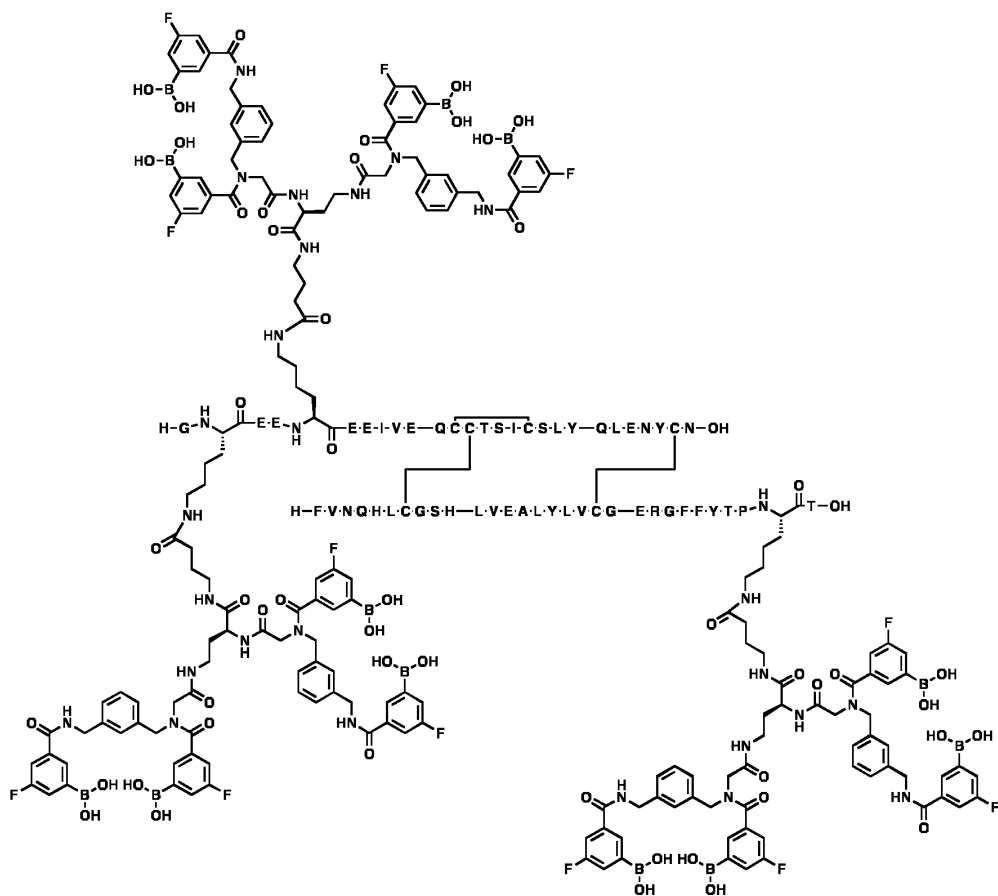
Example 774:



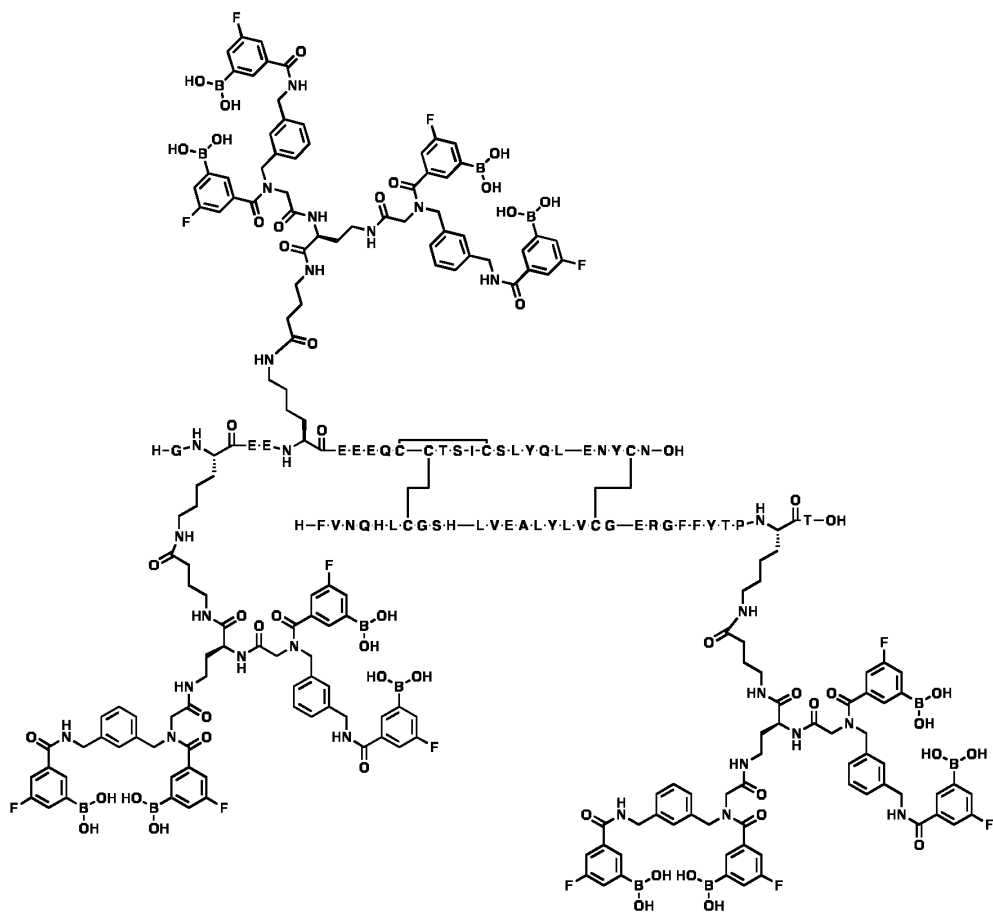
Example 775:



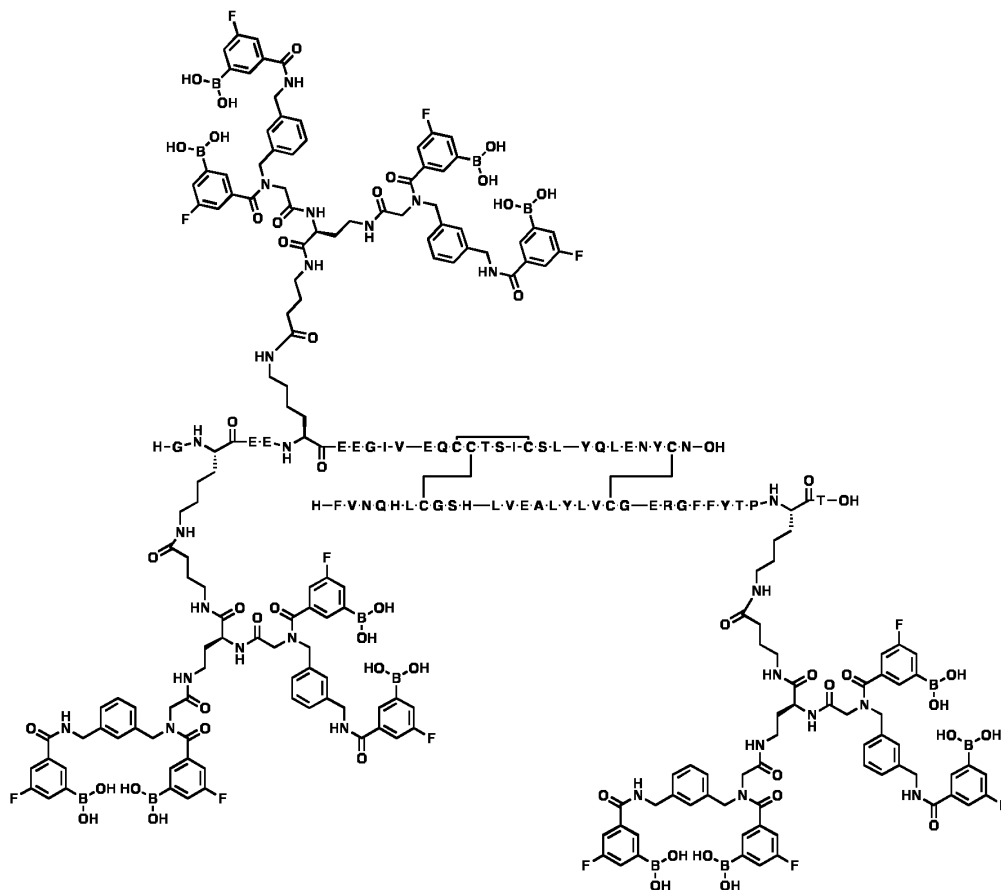
Example 776:

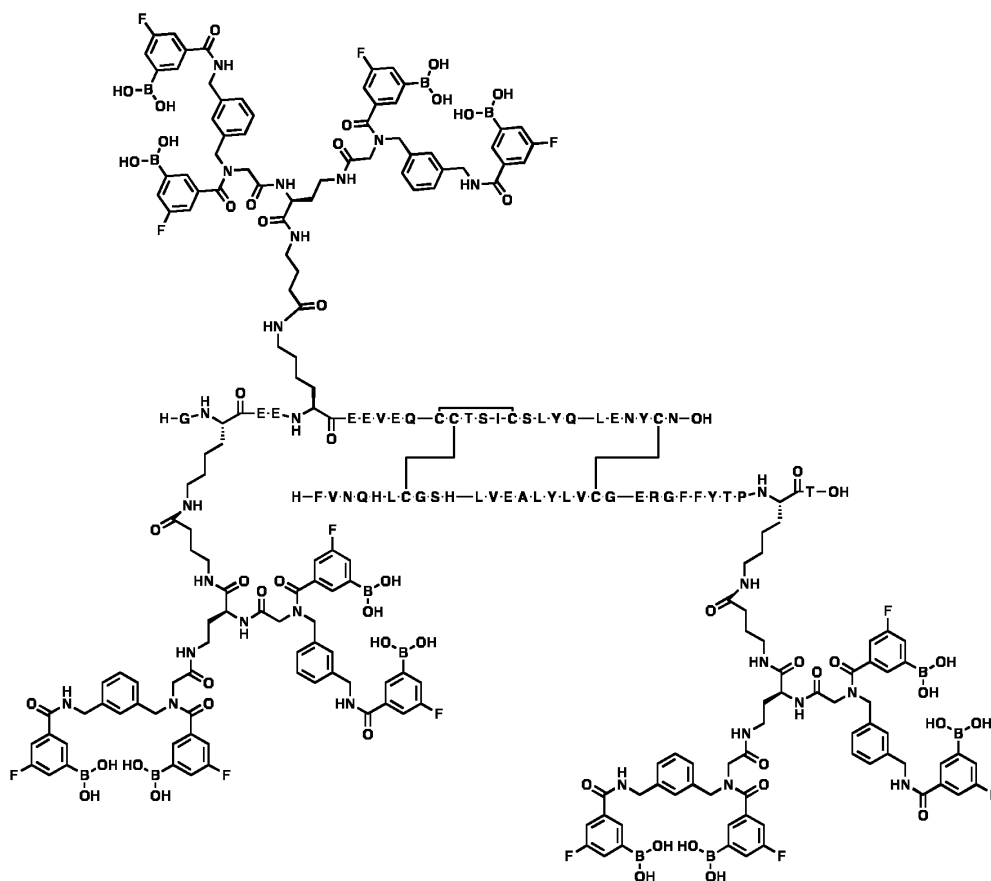


Example 777:

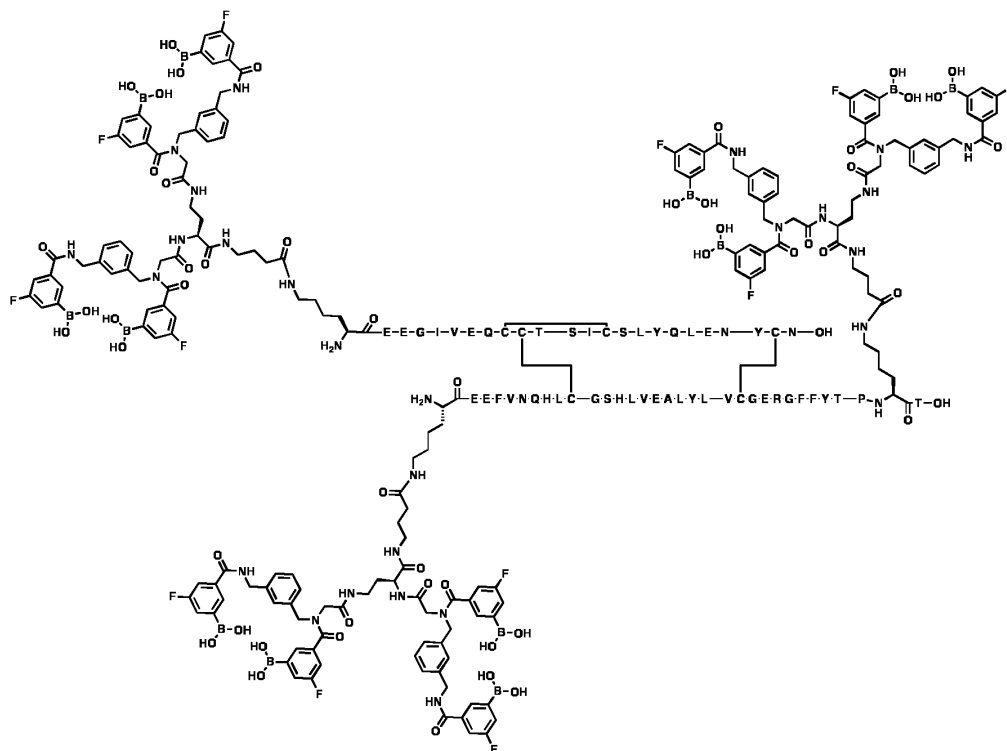


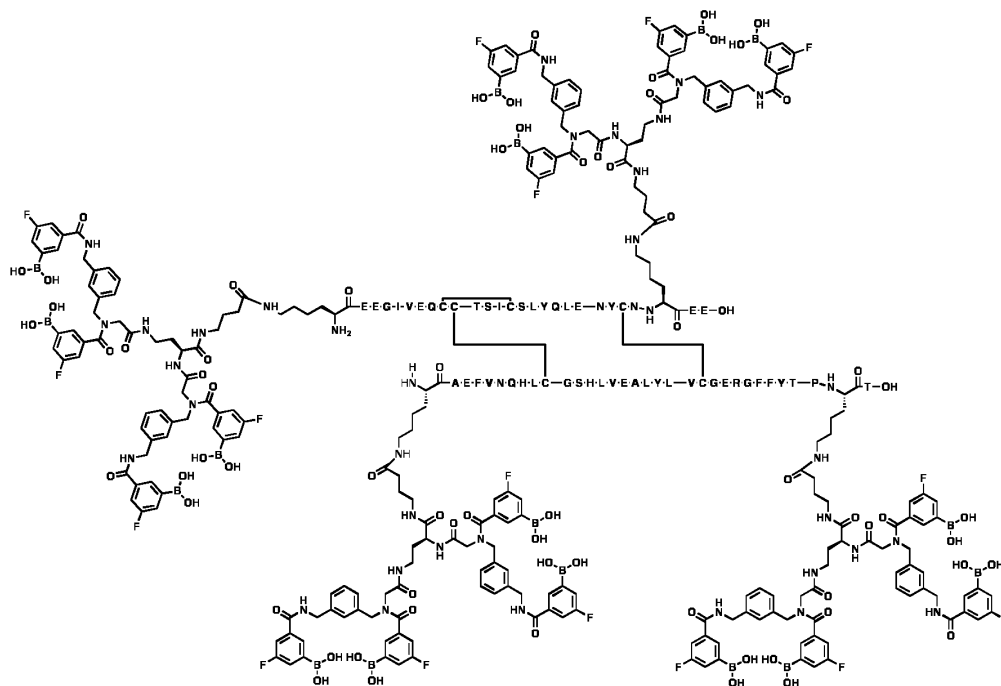
Example 778:



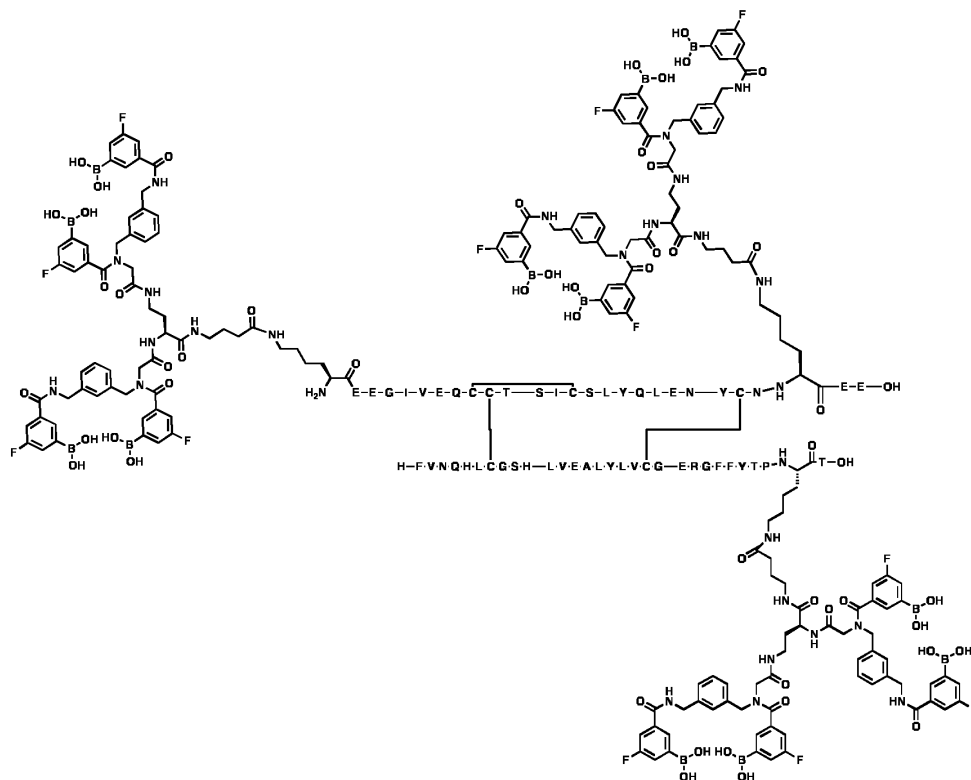
Example 779:

Example 780:

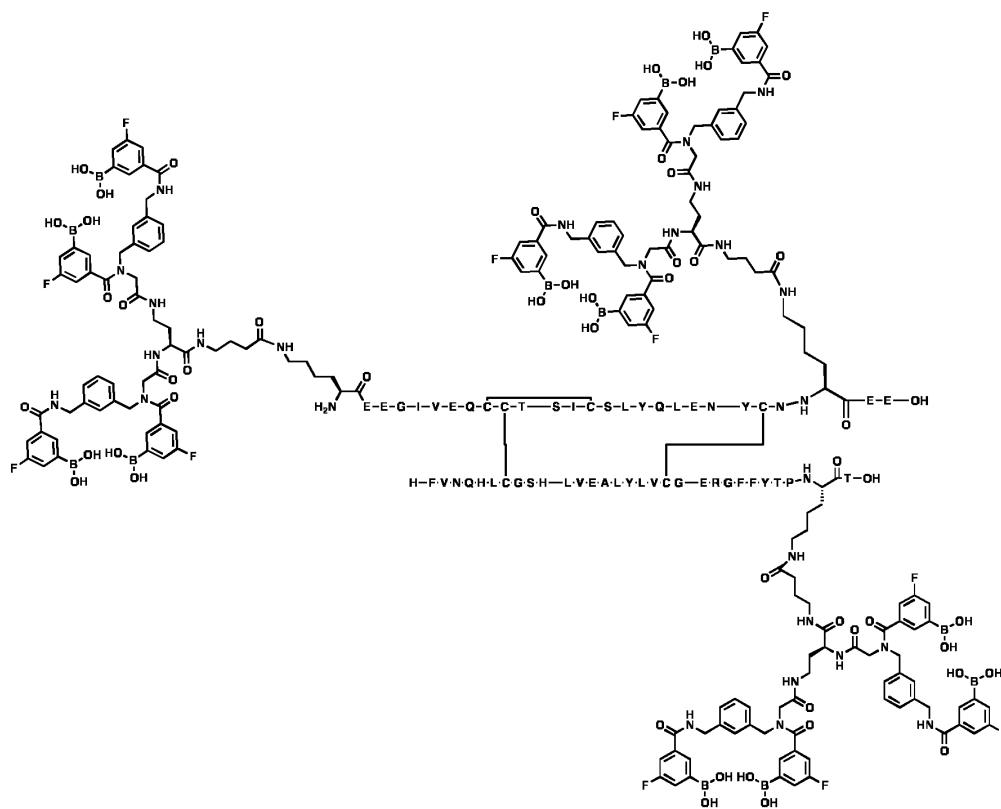


Example 781:

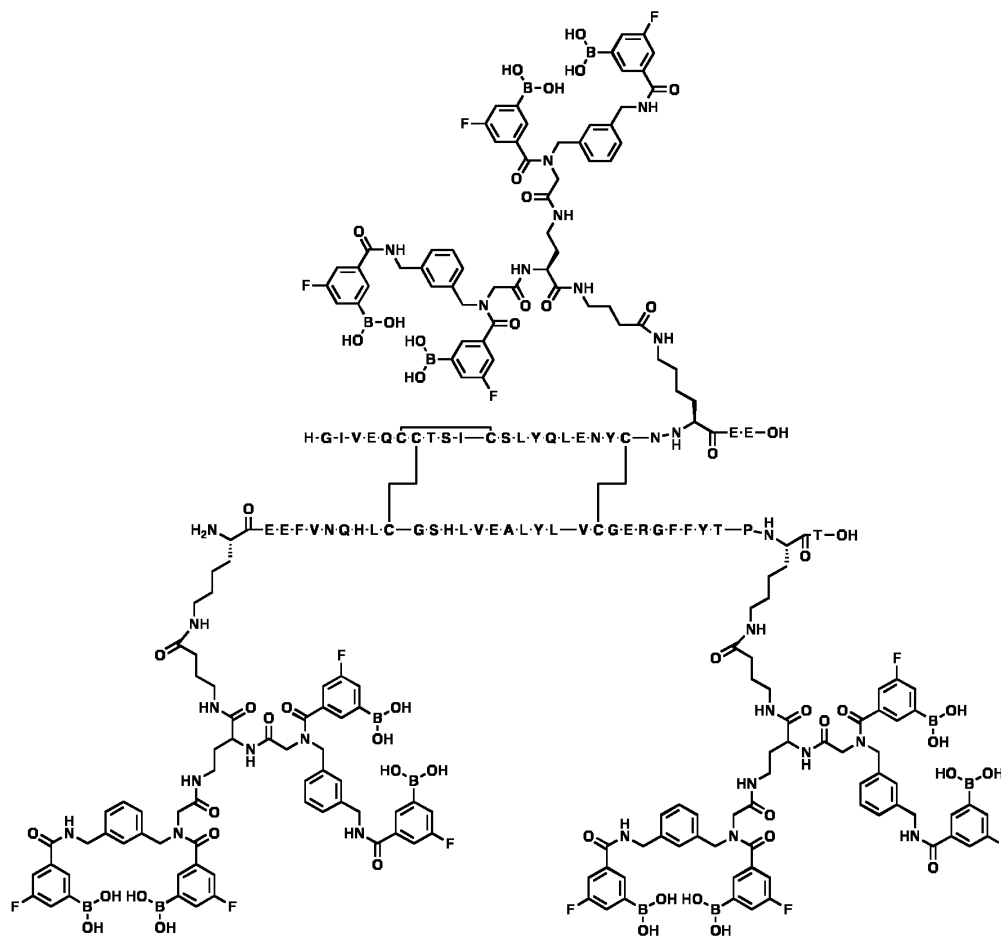
Example 782:

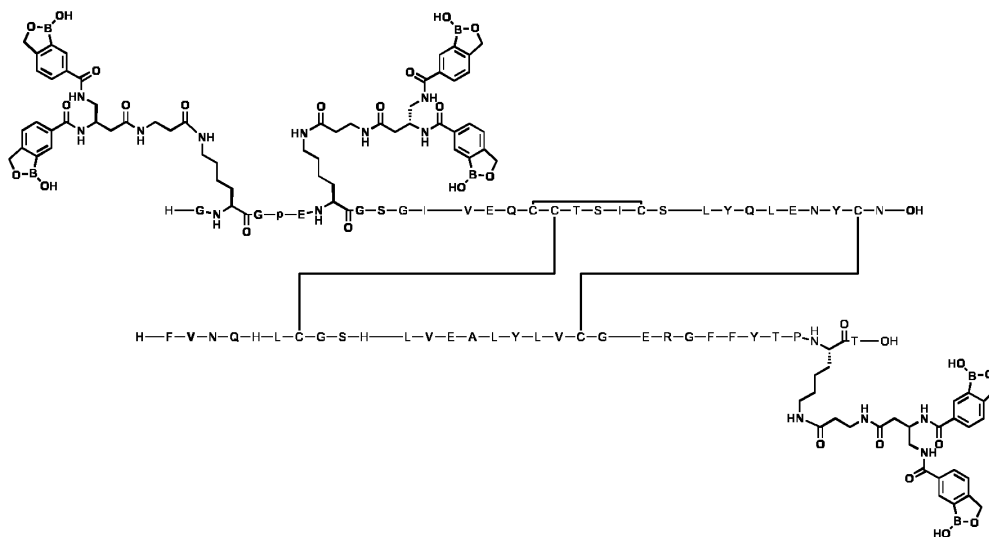
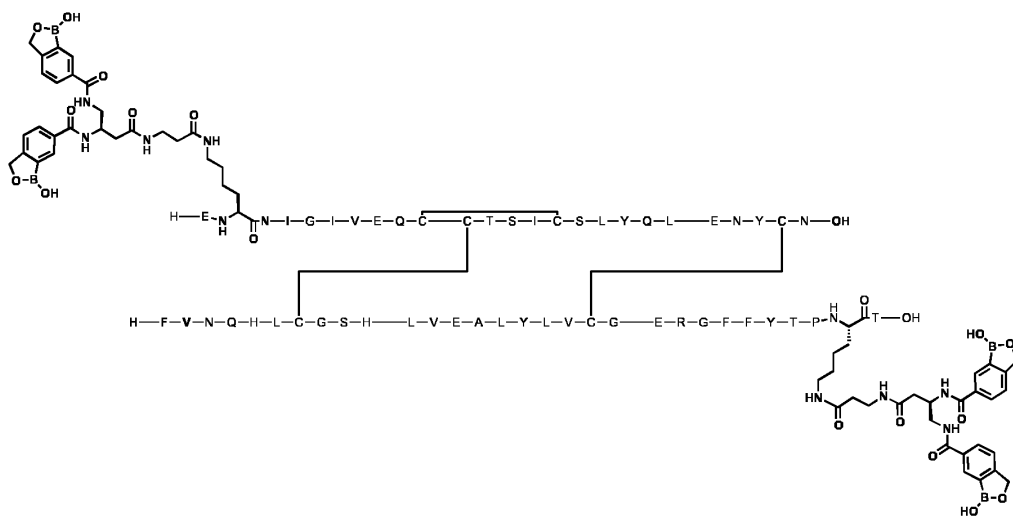


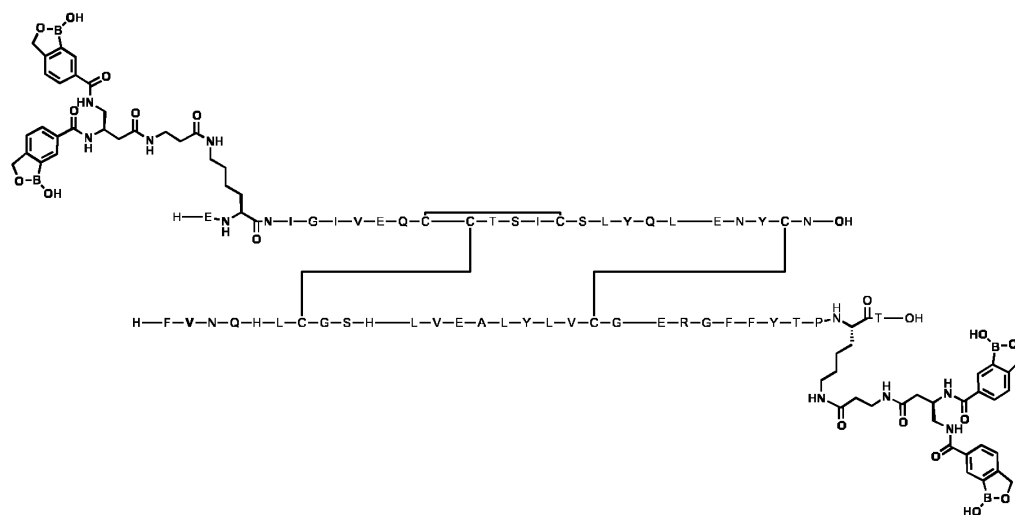
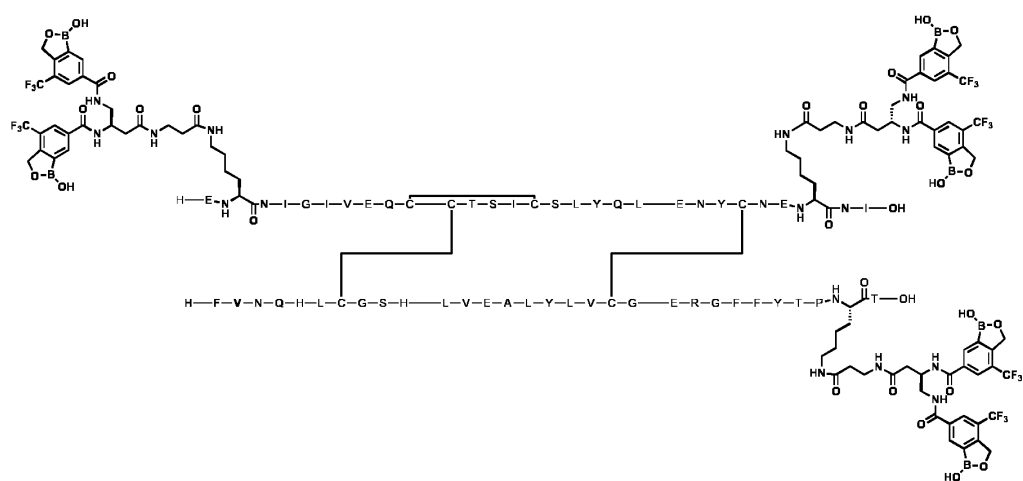
Example 783:

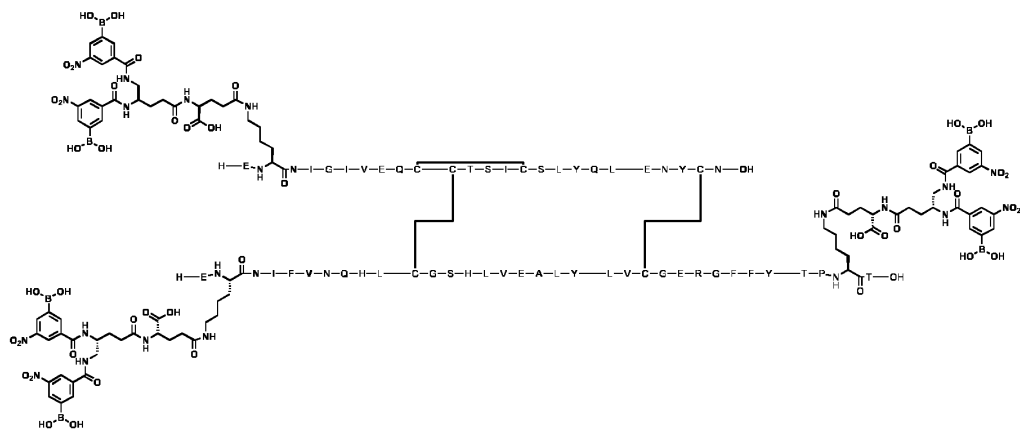
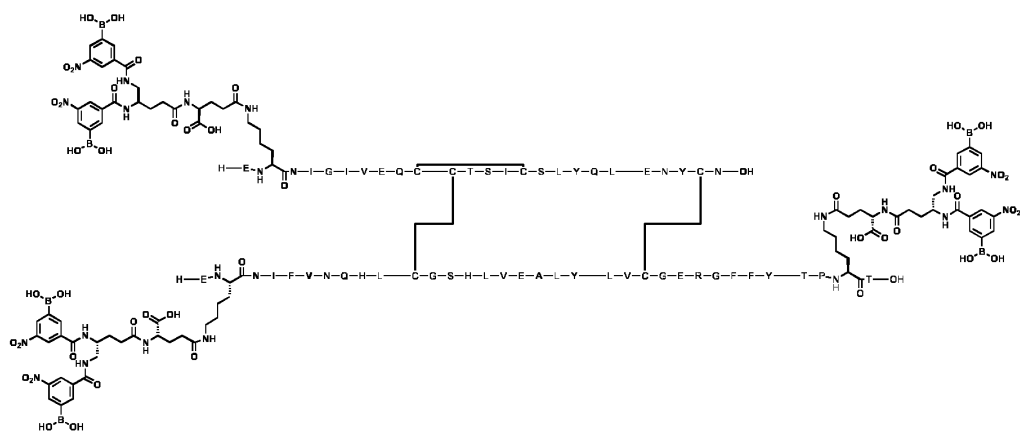


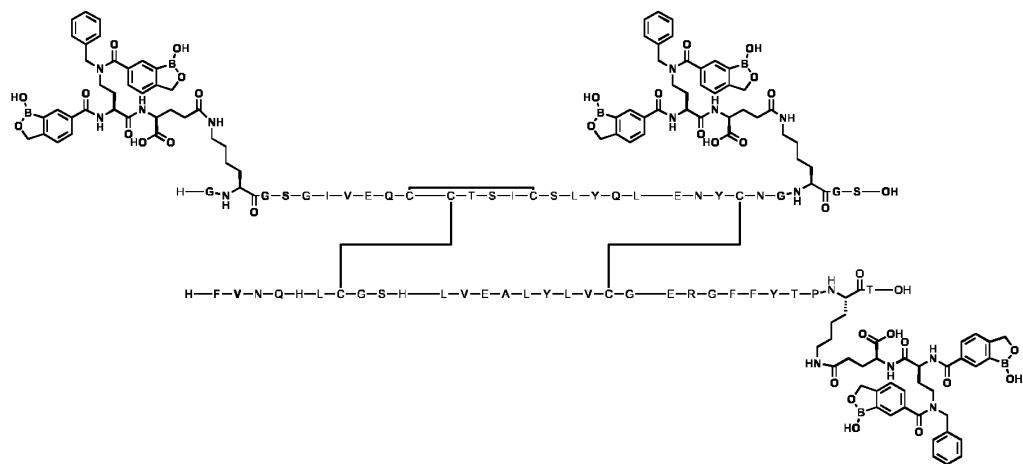
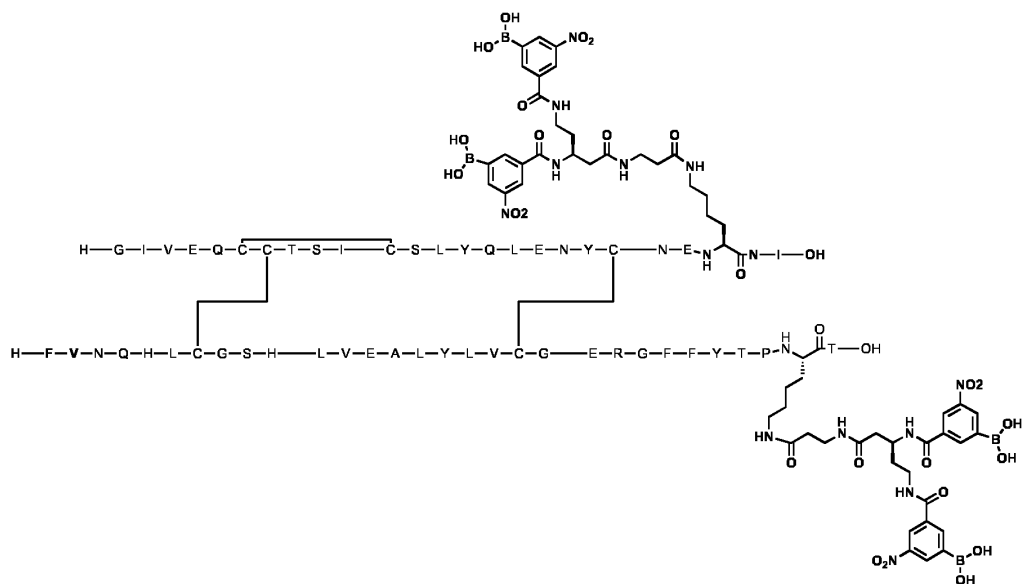
Example 784:

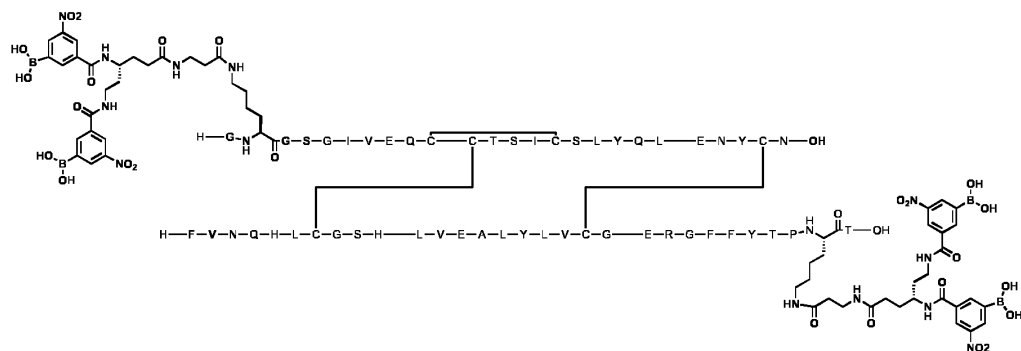
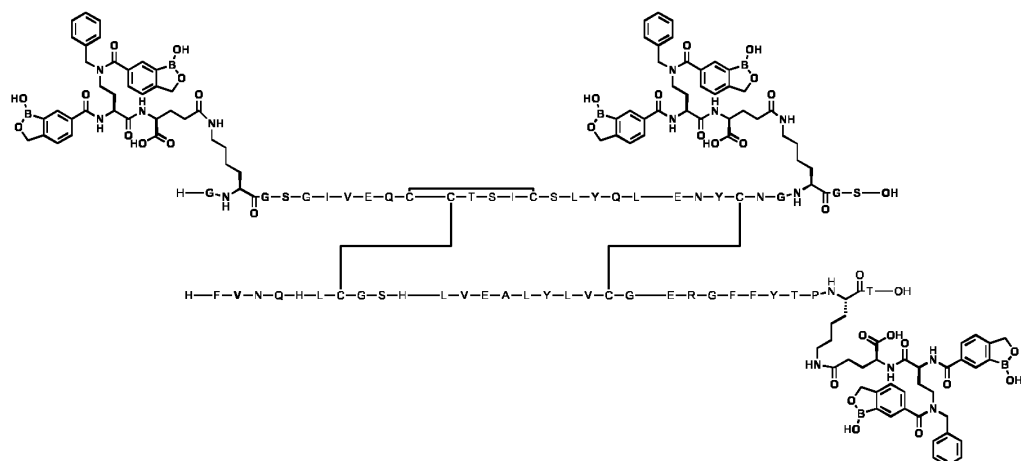


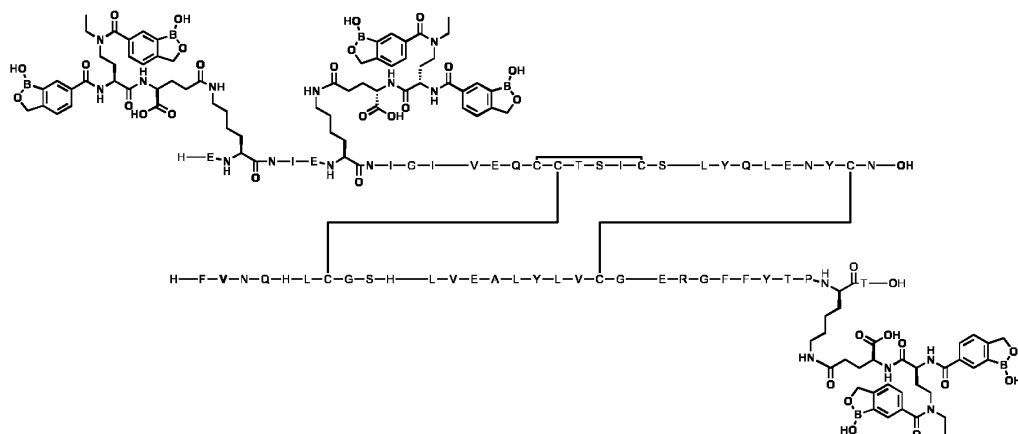
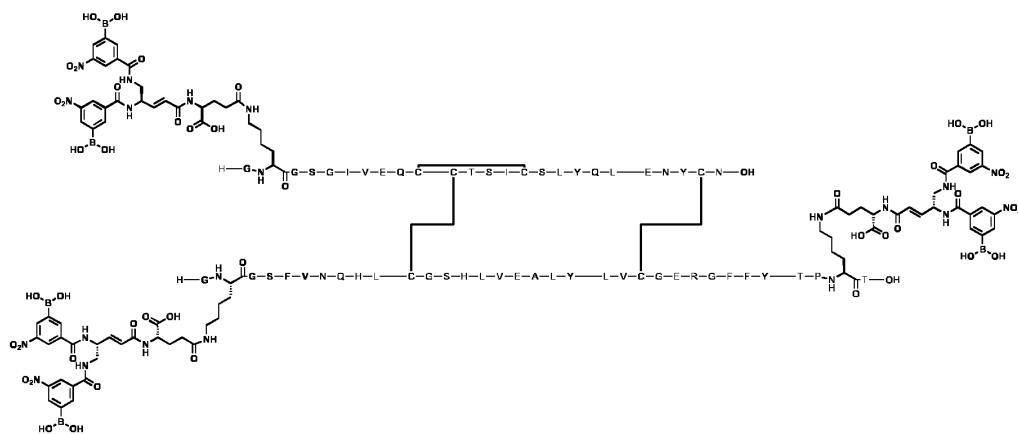
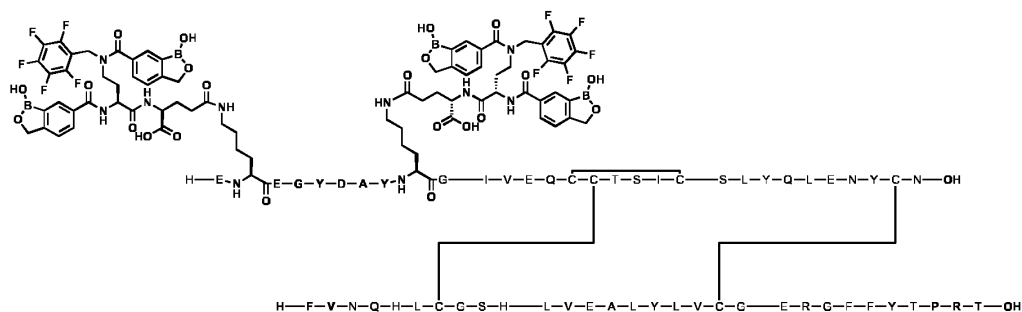
Example 785:**Example 786:**

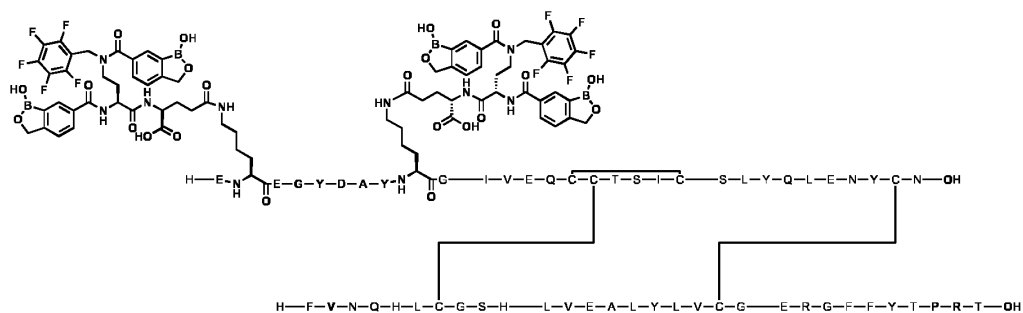
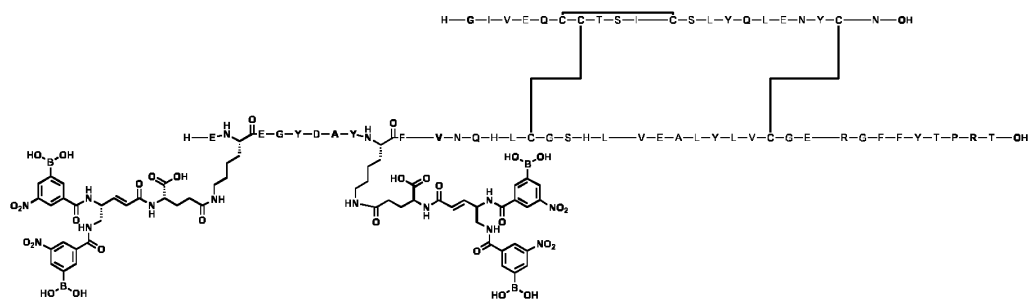
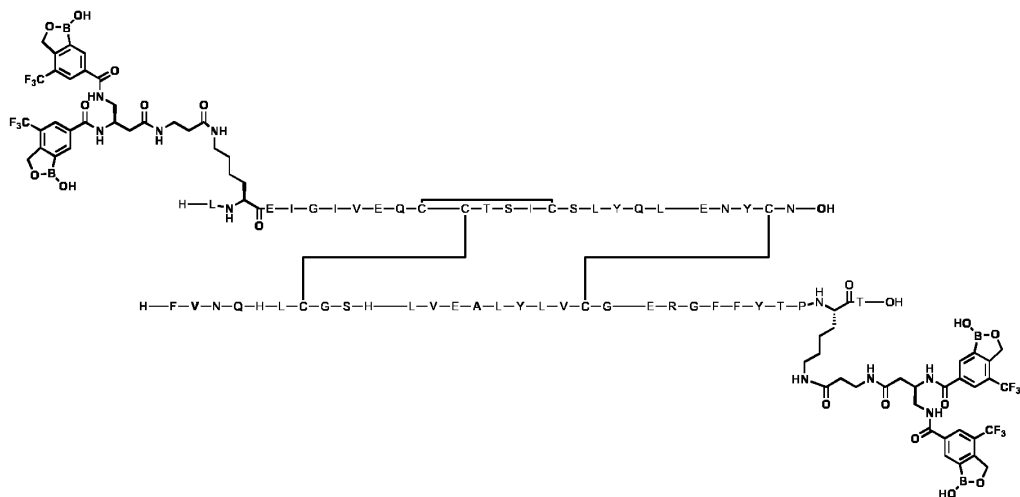
Example 787:**Example 788:**

Example 789:**Example 790:**

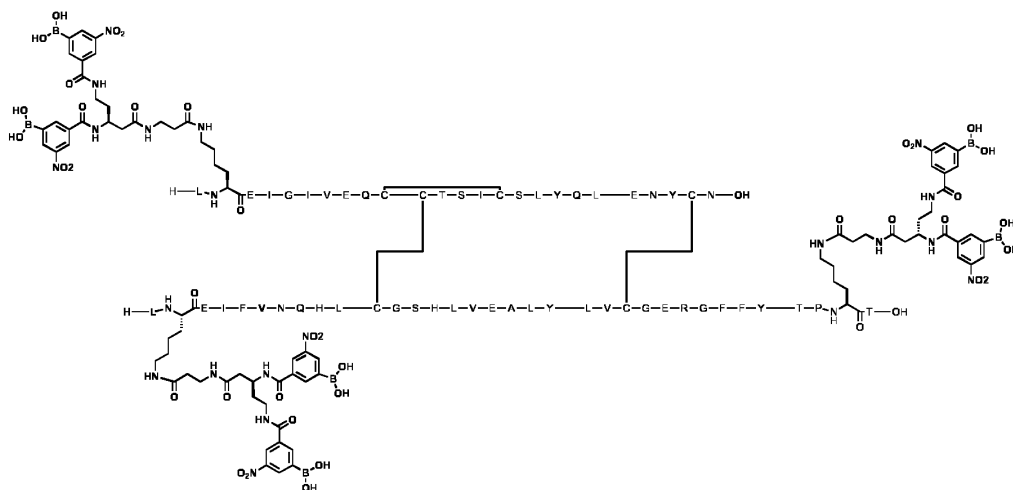


Example 793:**Example 794:**

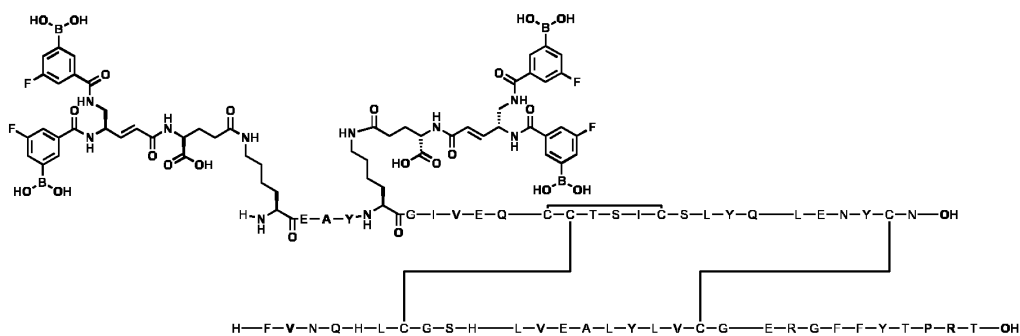
Example 795:**Example 796:****5 Example 797:**

Example 798:**Example 799:****5 Example 800:**

Example 801:

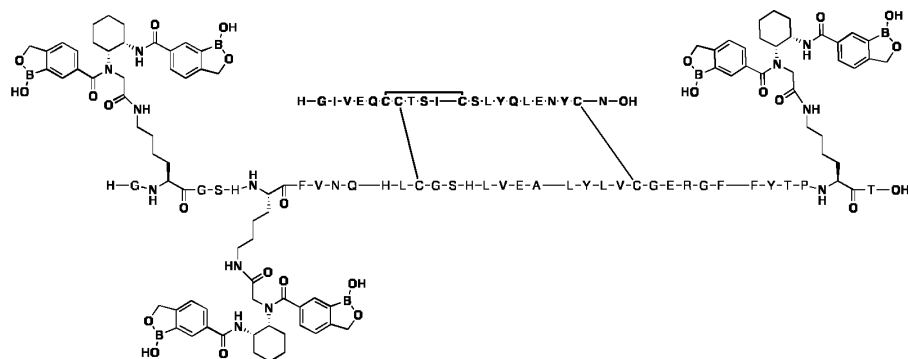


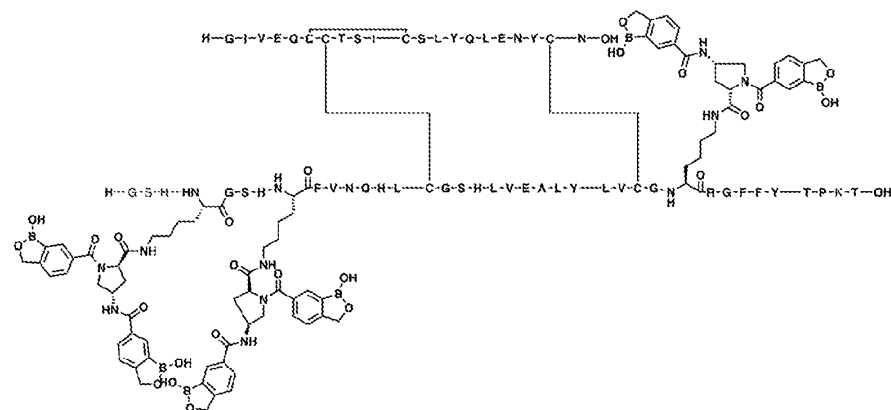
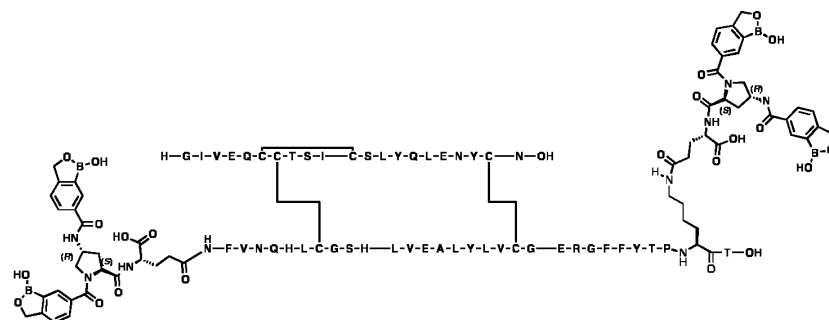
Example 802:



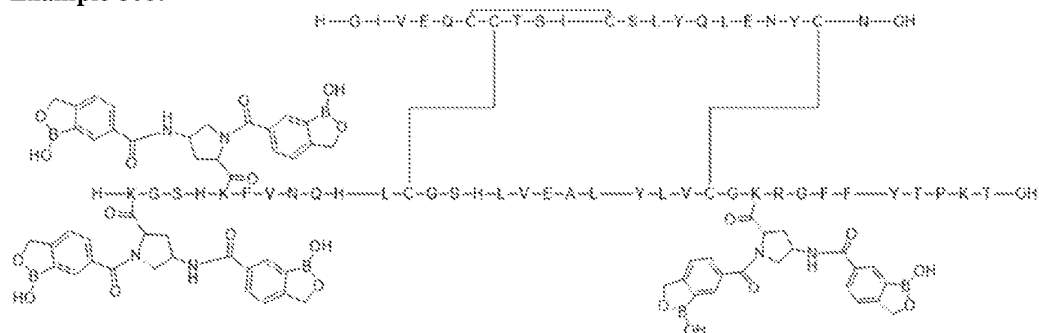
5

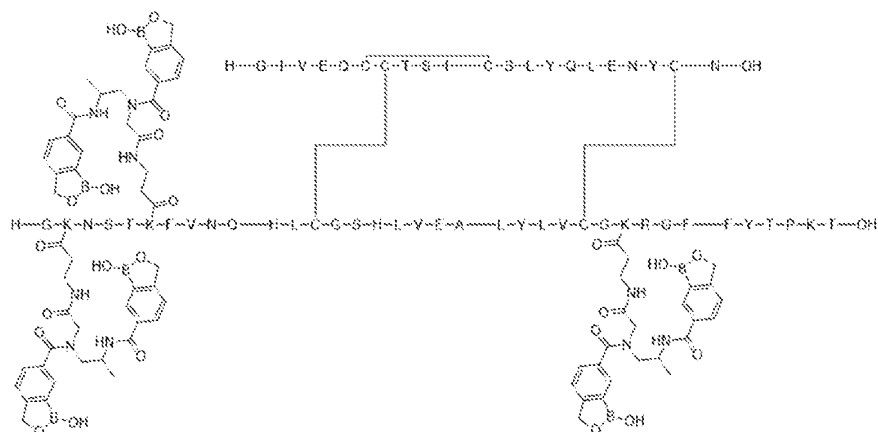
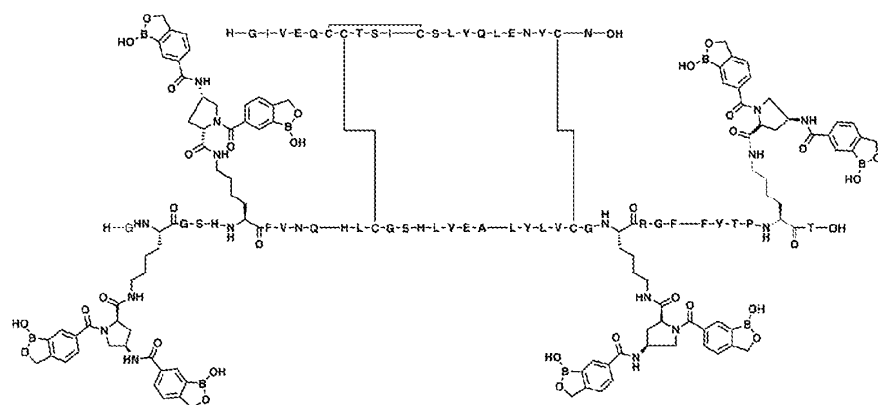
Example 803:

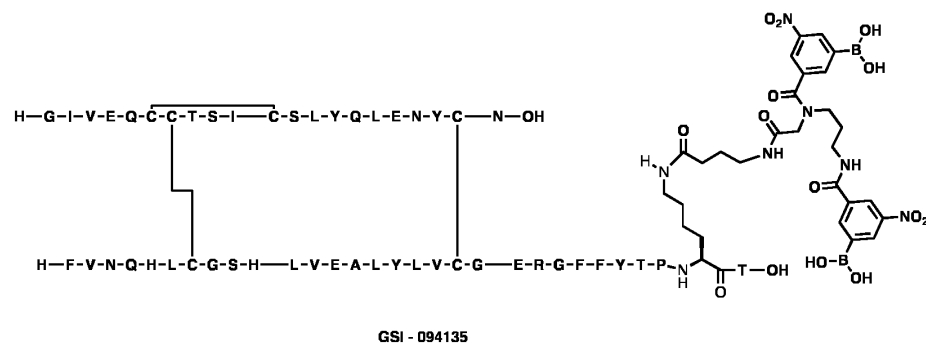
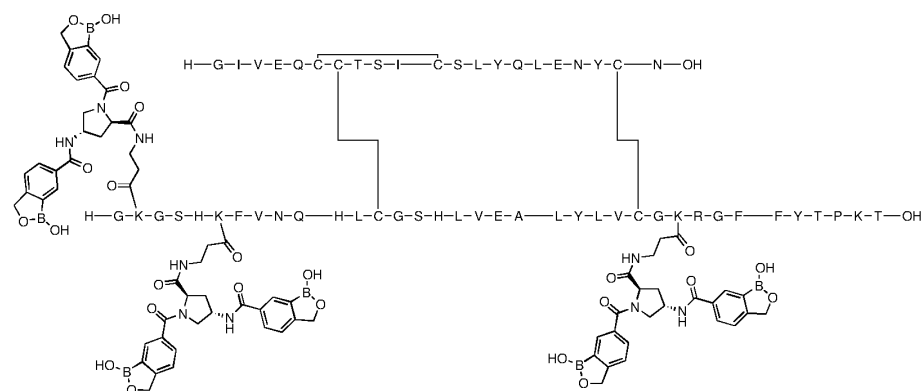


Example 804:**Example 805:**

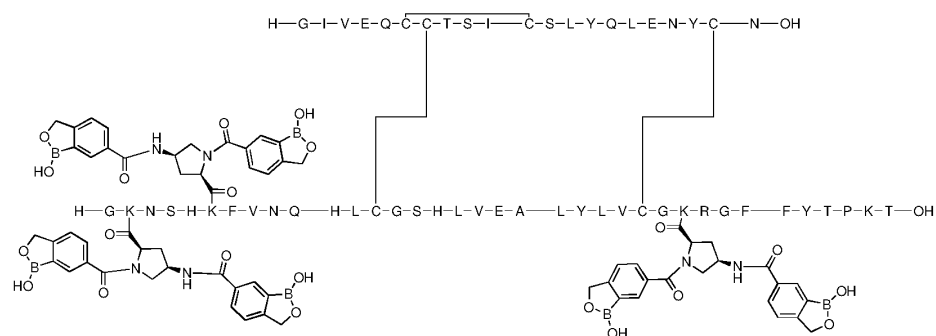
5

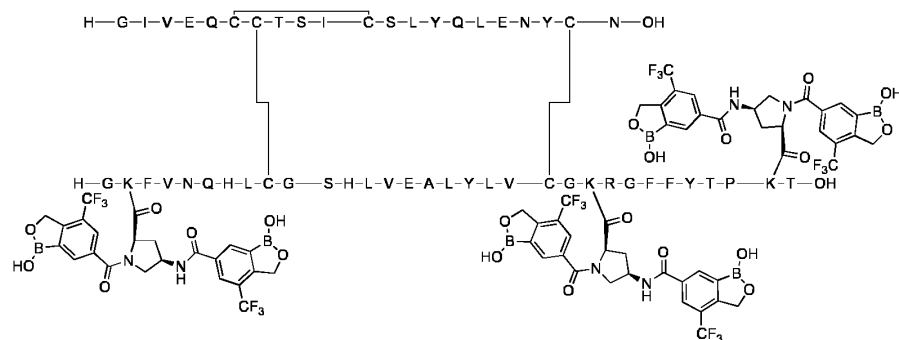
Example 806:

Example 807:**Example 808:**

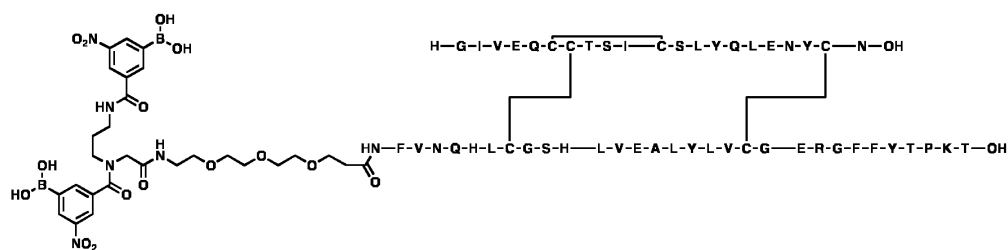
Example 809:**Example 810:**

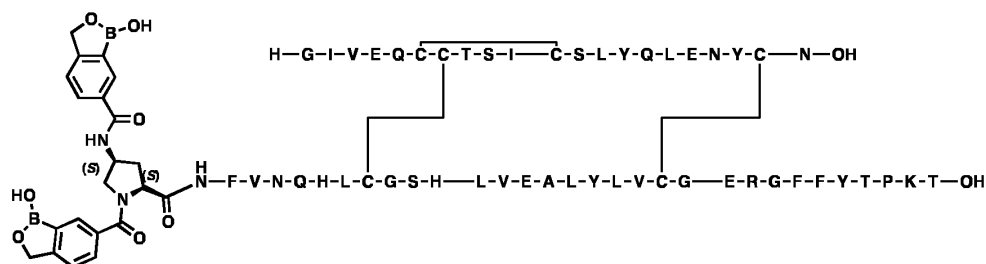
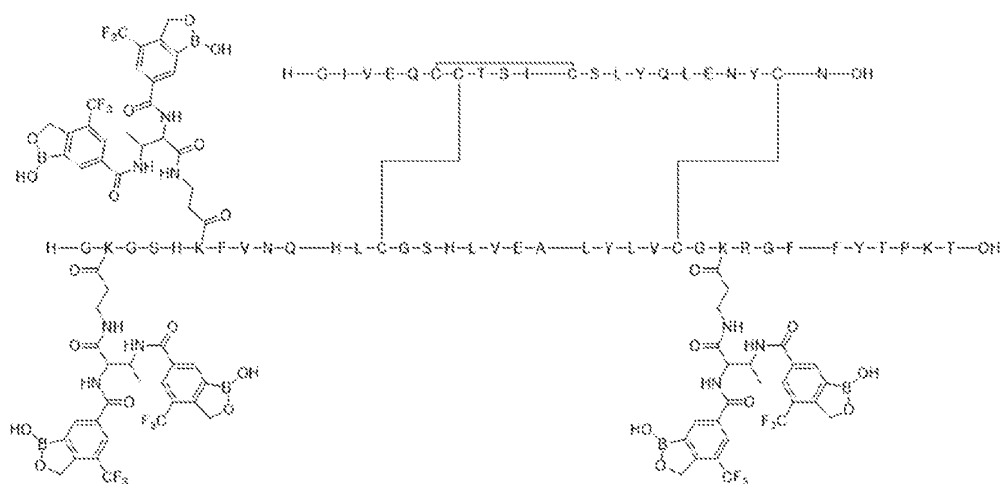
5

Example 811:

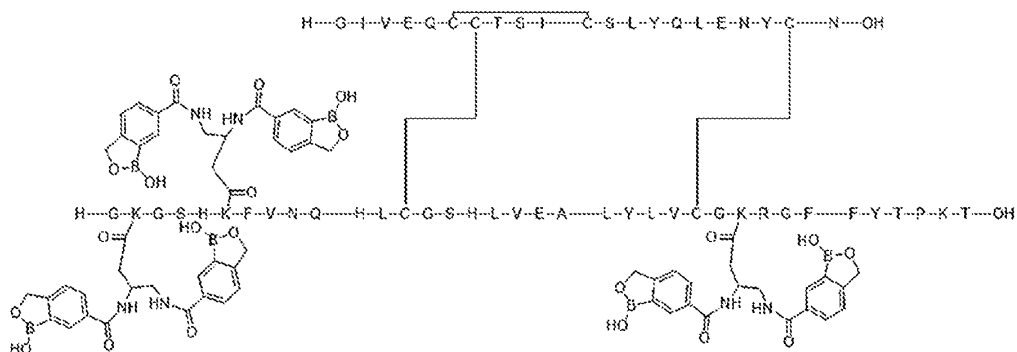
[illegible]

Example 814:

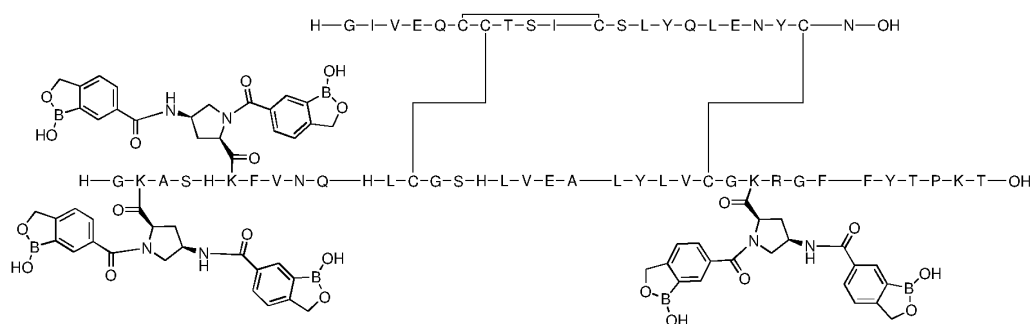


Example 815:**Example 816:**

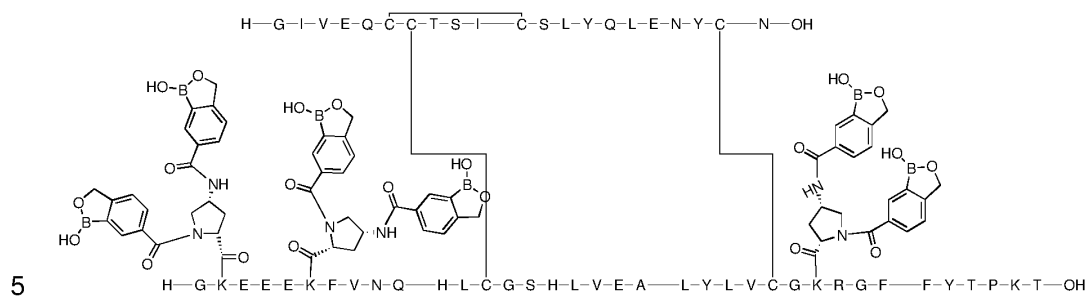
5

Example 817:

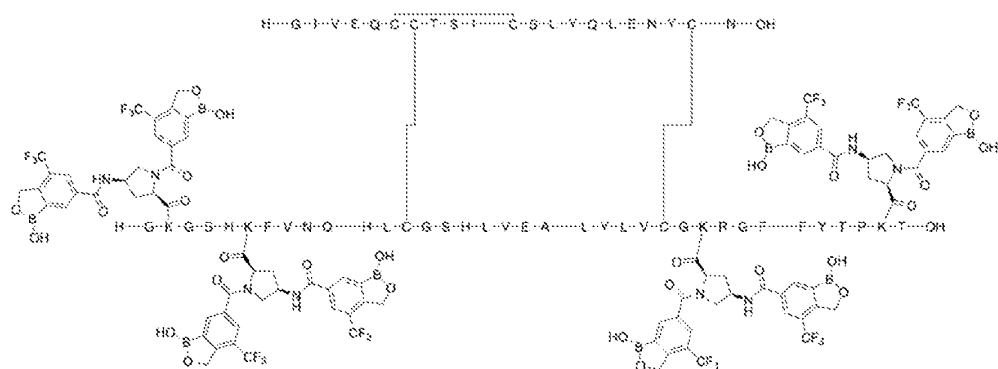
Example 818:

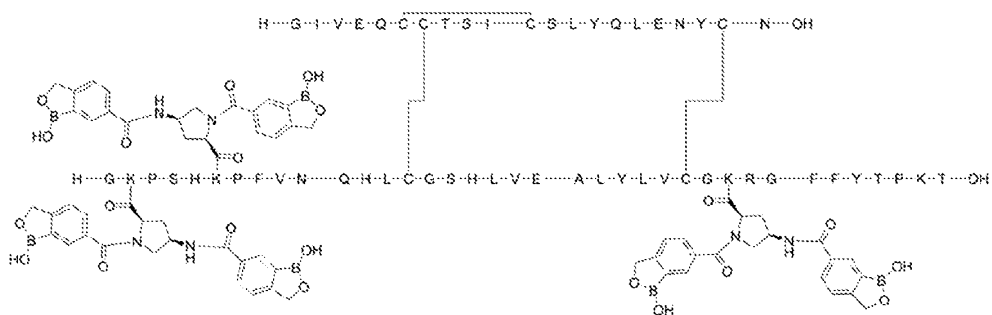
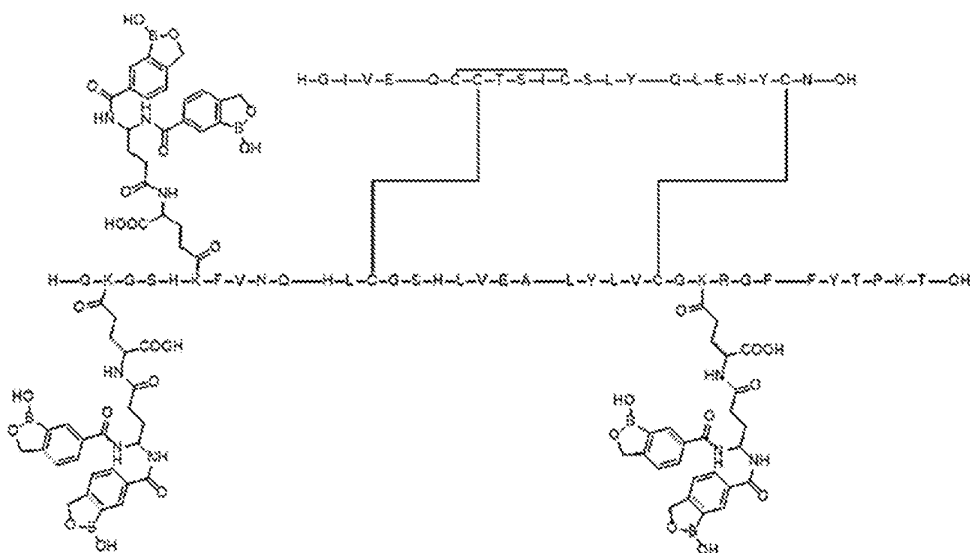


Example 819:

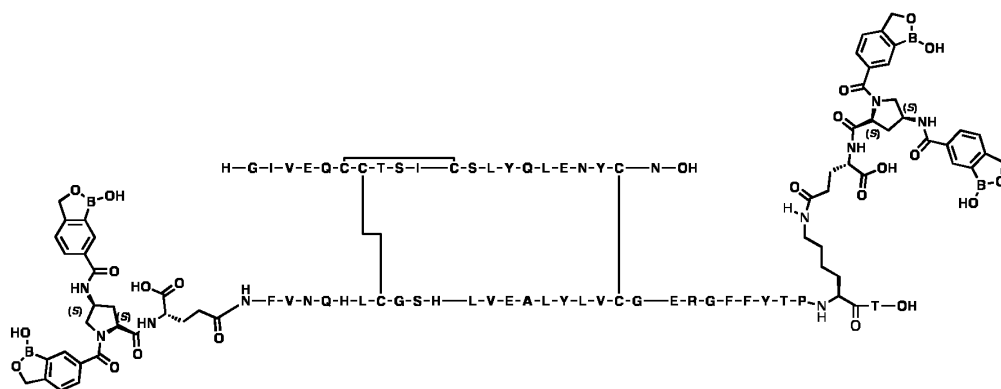


Example 820:

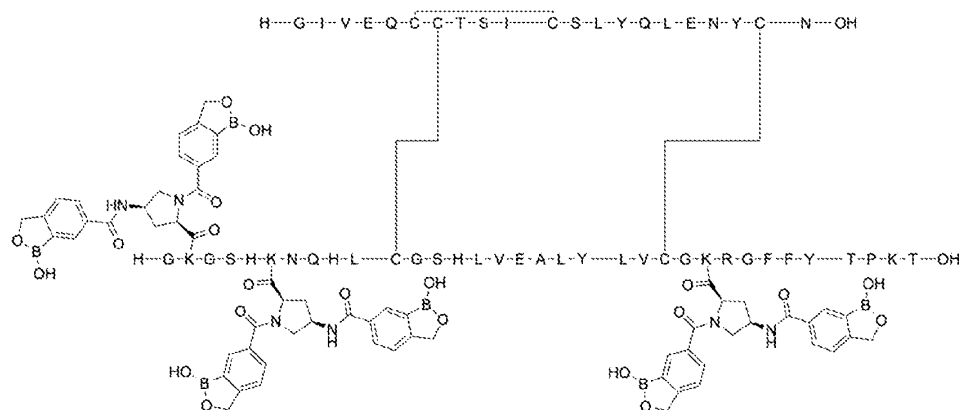


Example 821:**Example 822:**

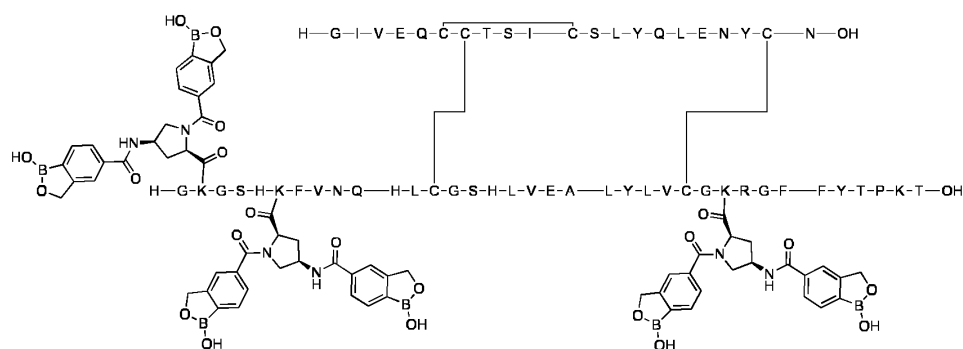
Example 823:

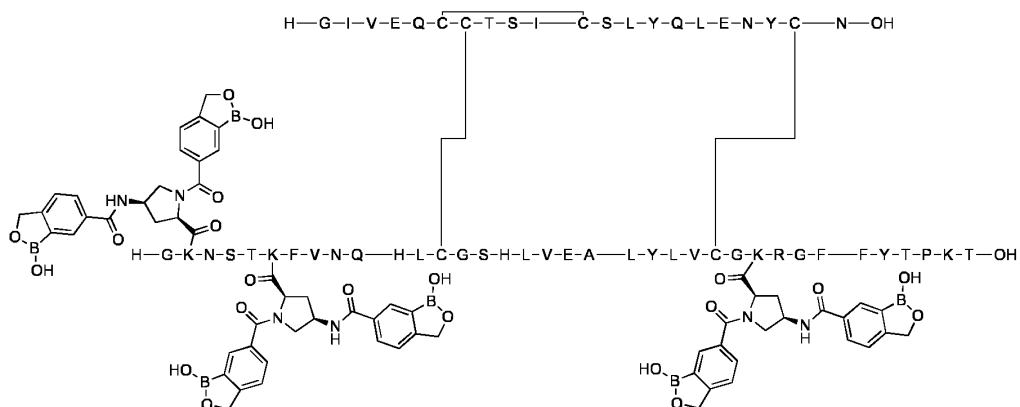
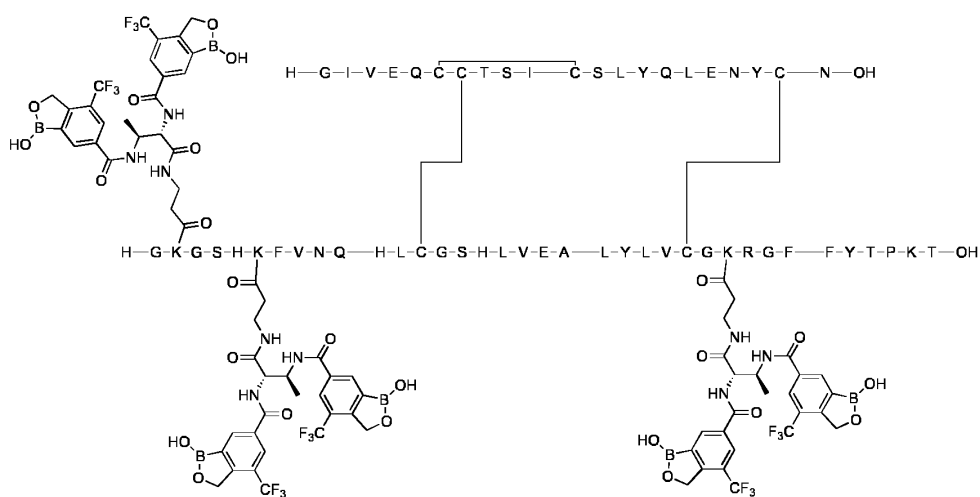


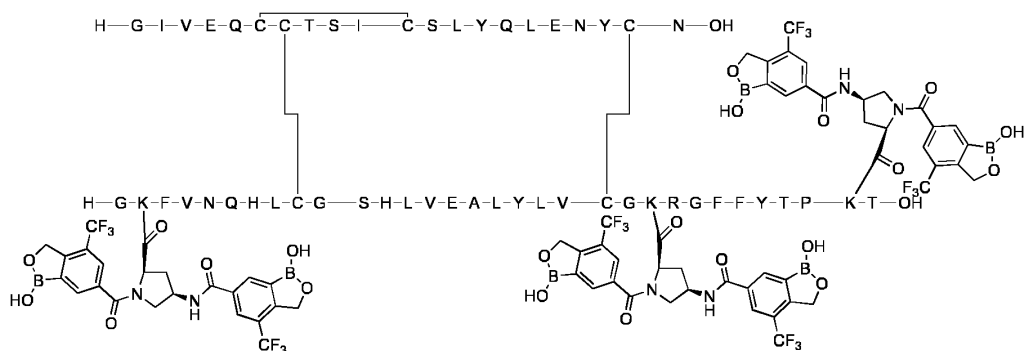
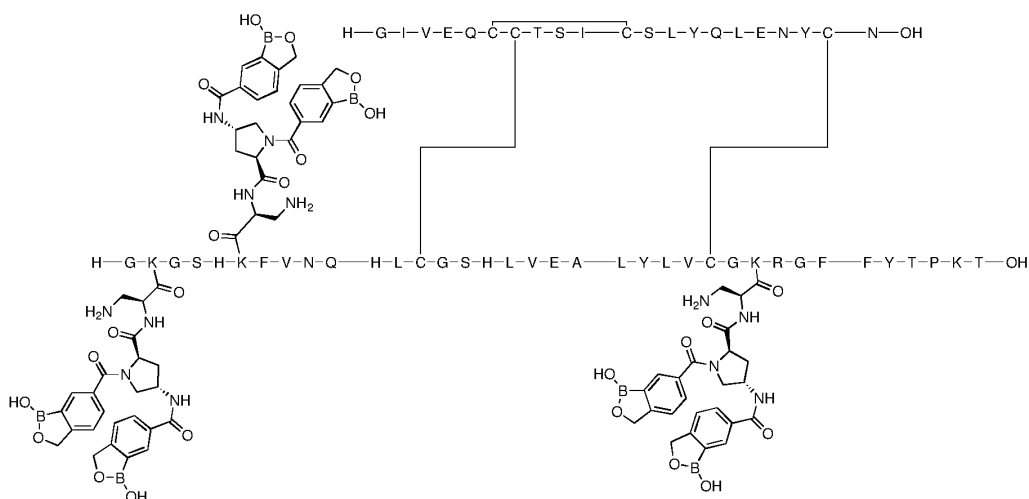
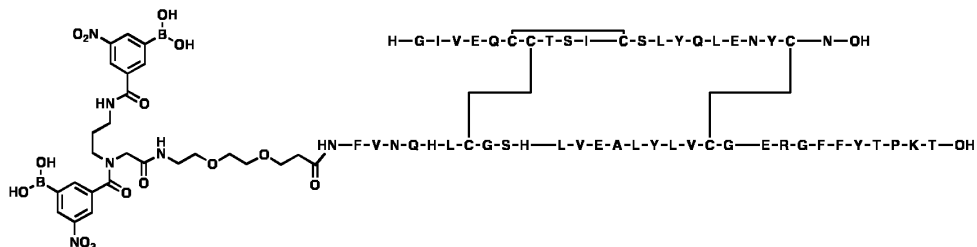
Example 824:

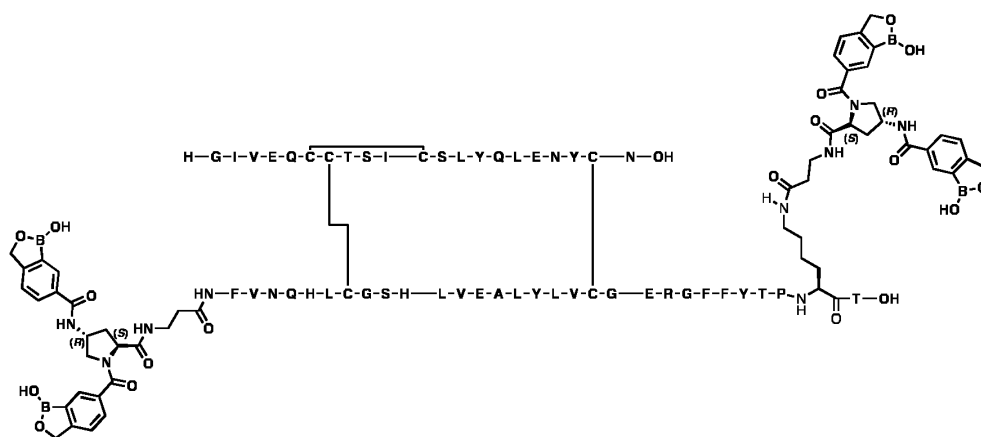
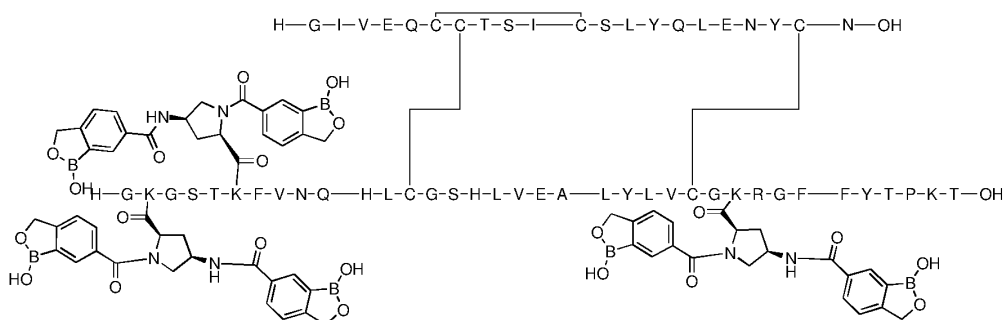


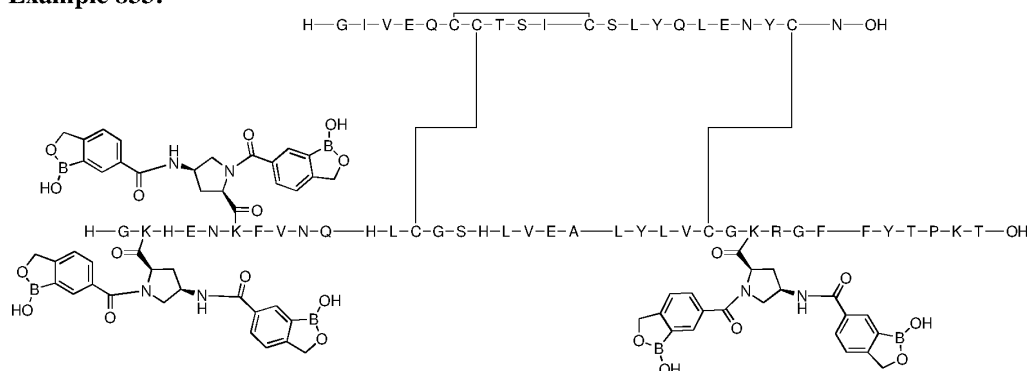
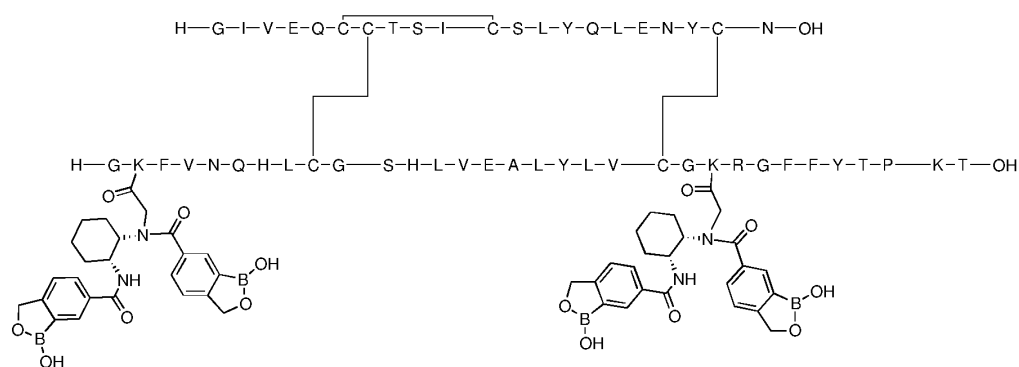
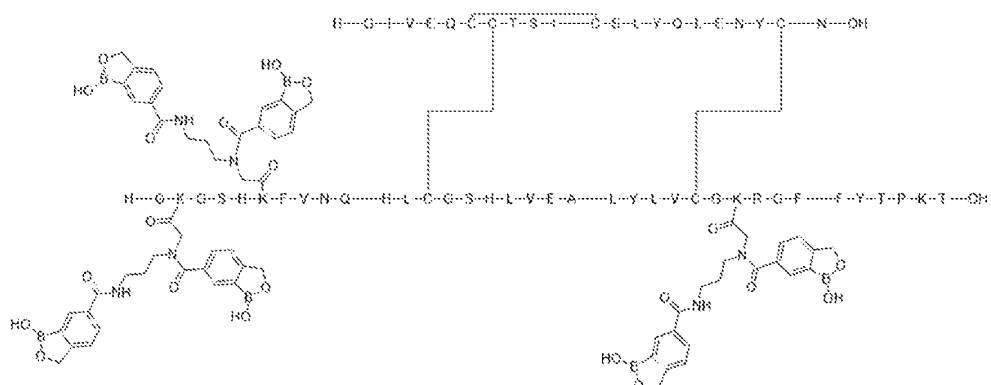
5 Example 825:

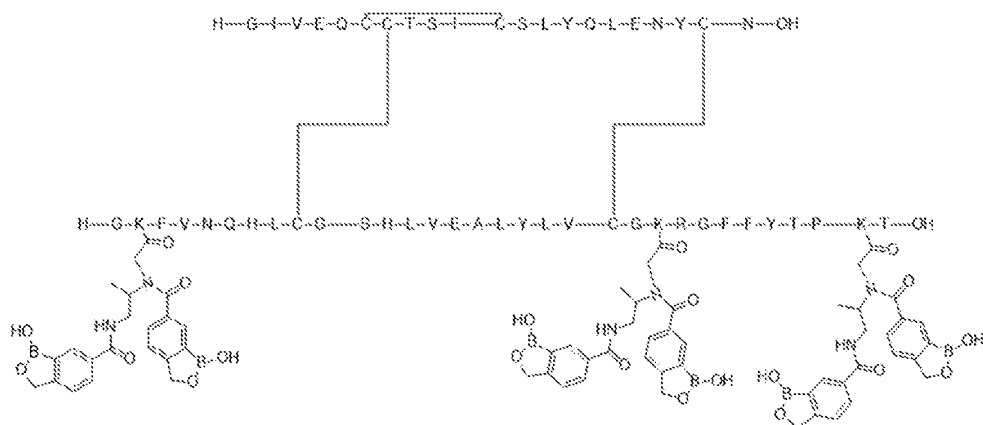
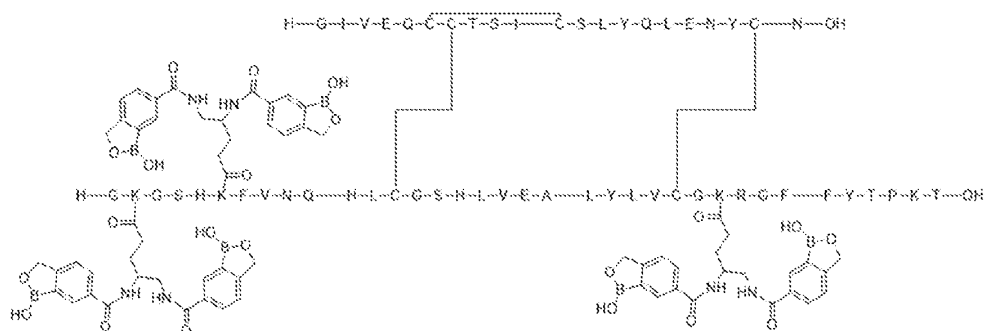
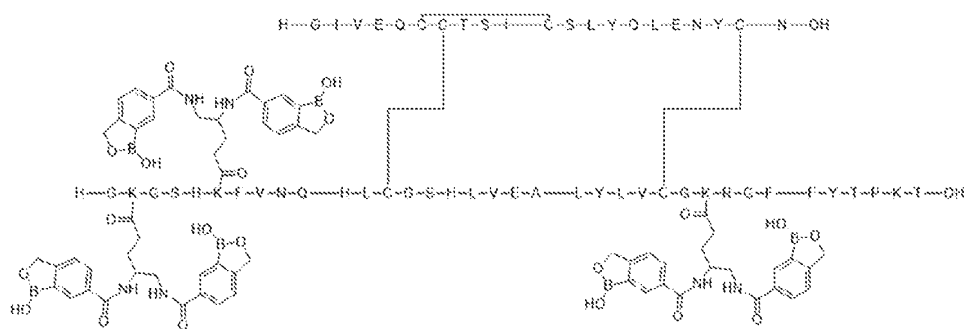


Example 826:**Example 827:**

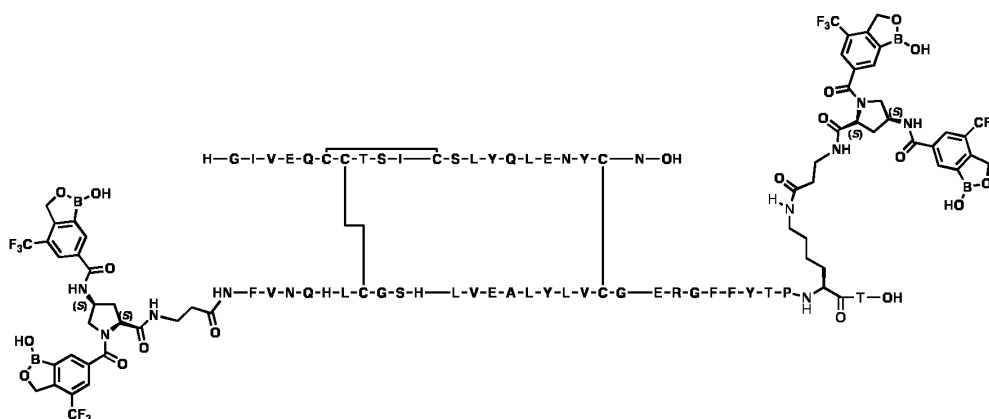
Example 828:**Example 829:****5 Example 830:**

Example 831:**Example 832:**

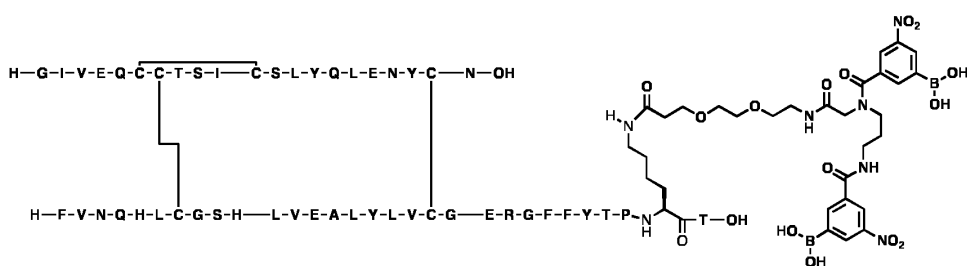
Example 833:**Example 834:****5 Example 835:**

Example 836:**Example 837:****5 Example 838:**

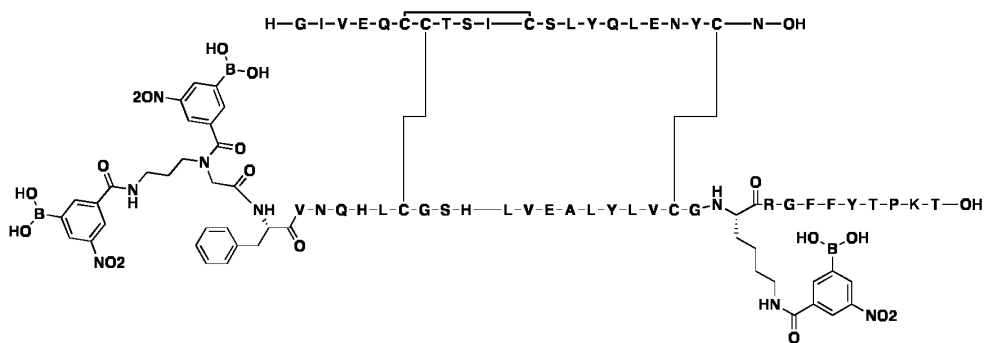
Example 839:

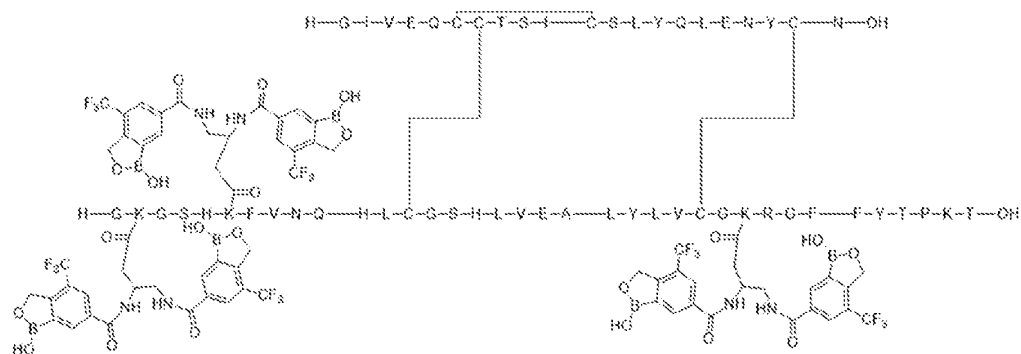
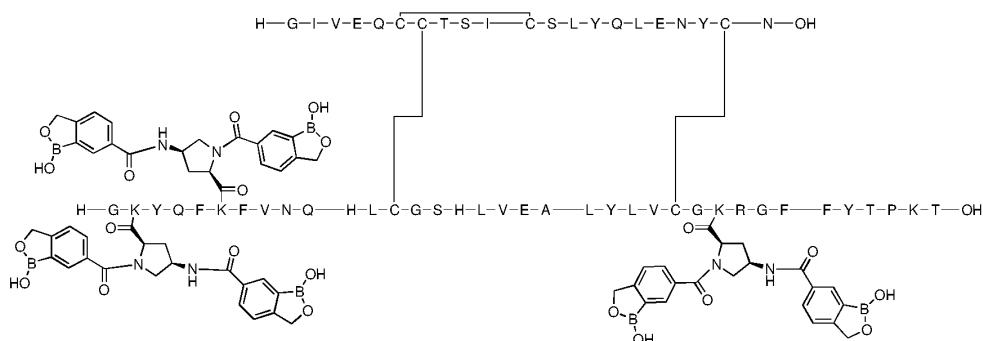
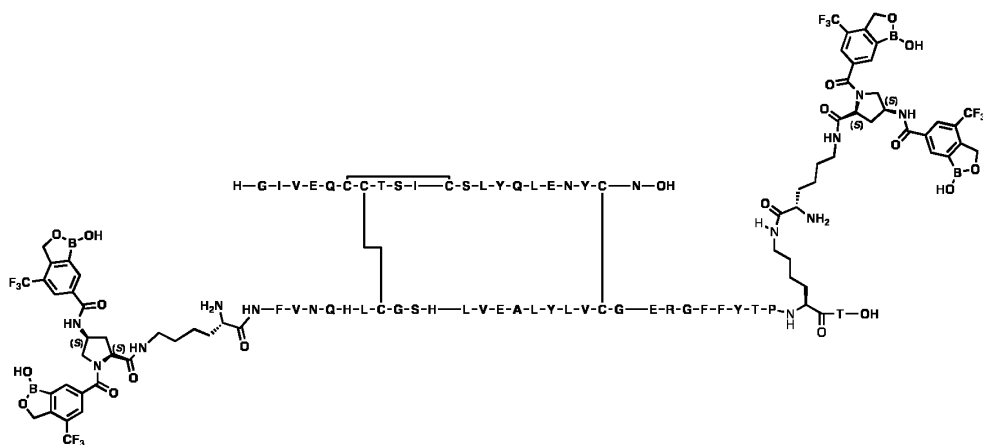


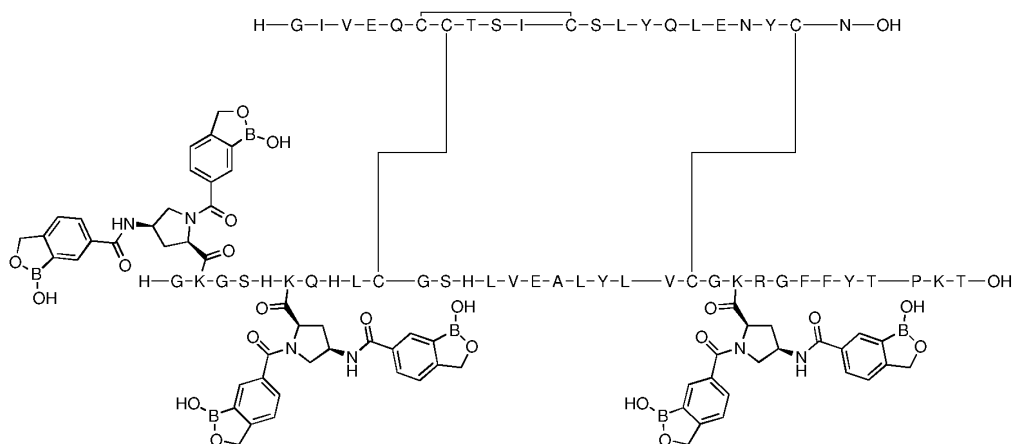
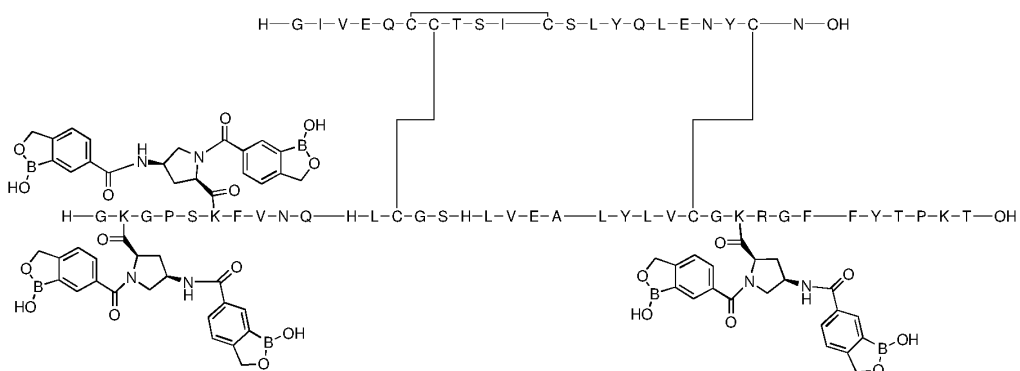
Example 840:



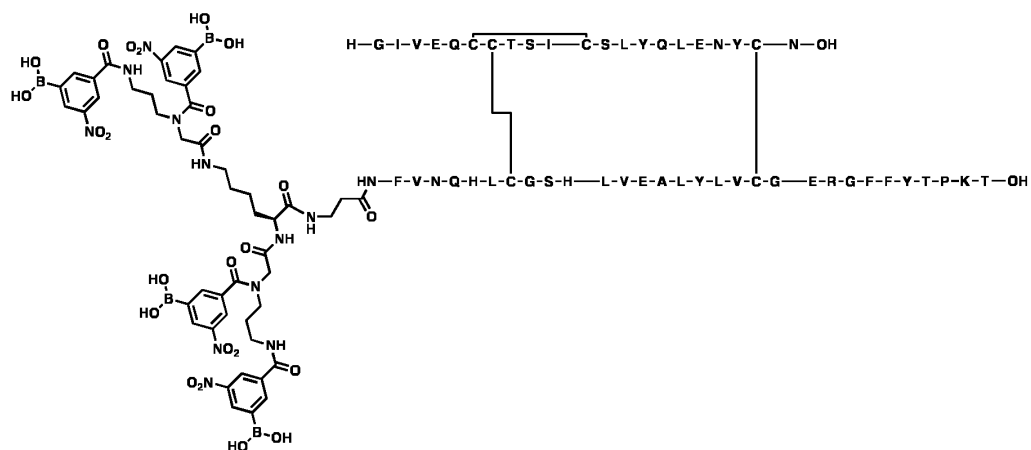
5 Example 841:



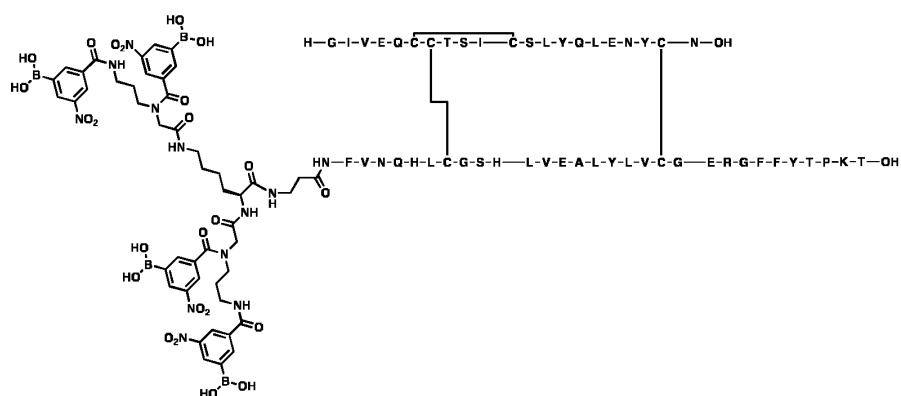
Example 842:**Example 843:****5 Example 844:**

Example 845:**Example 846:**

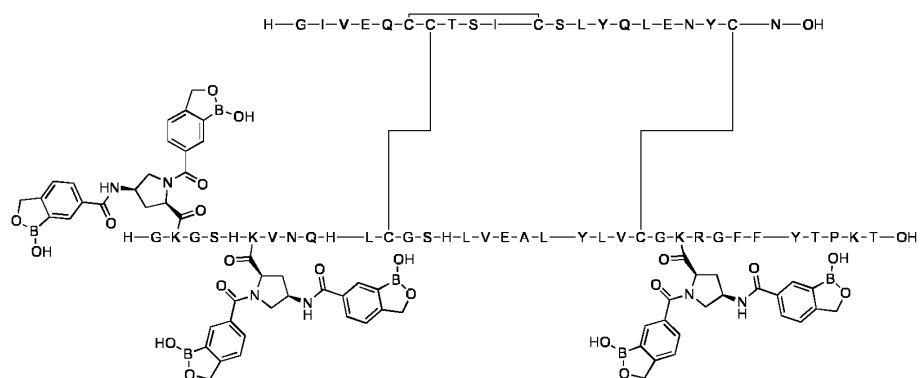
Example 847:

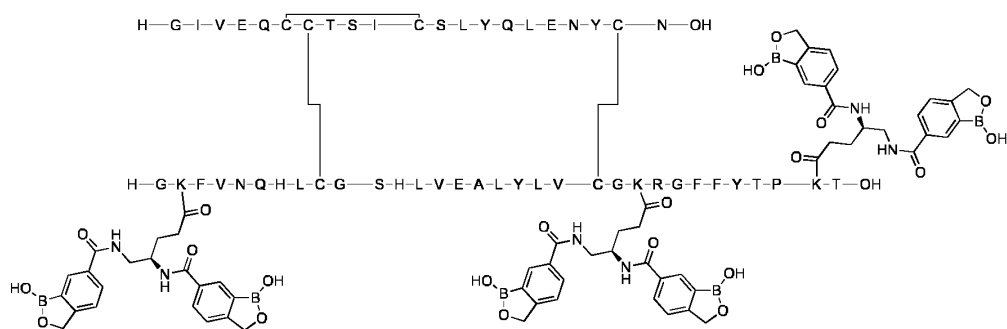
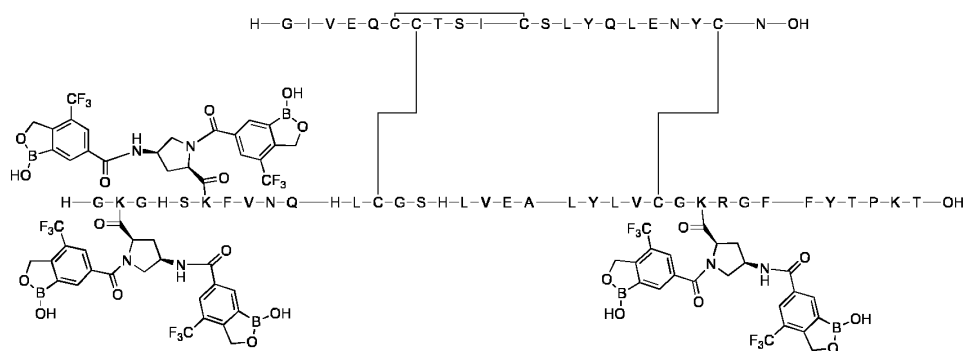


Example 848:

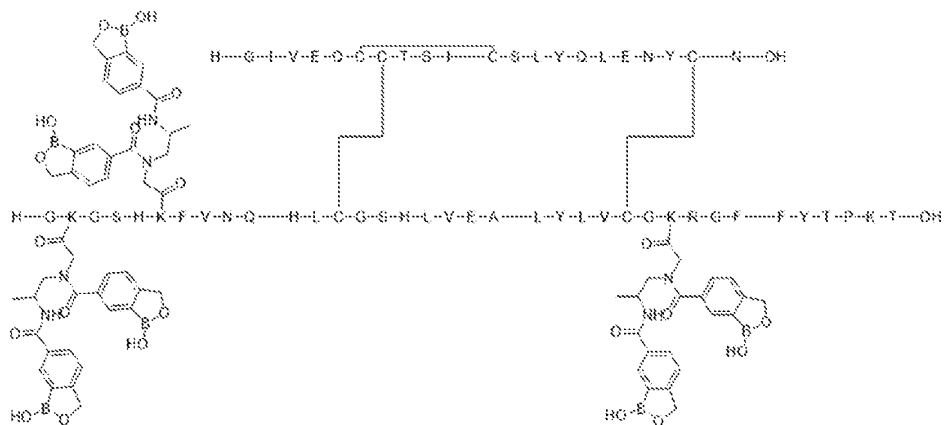


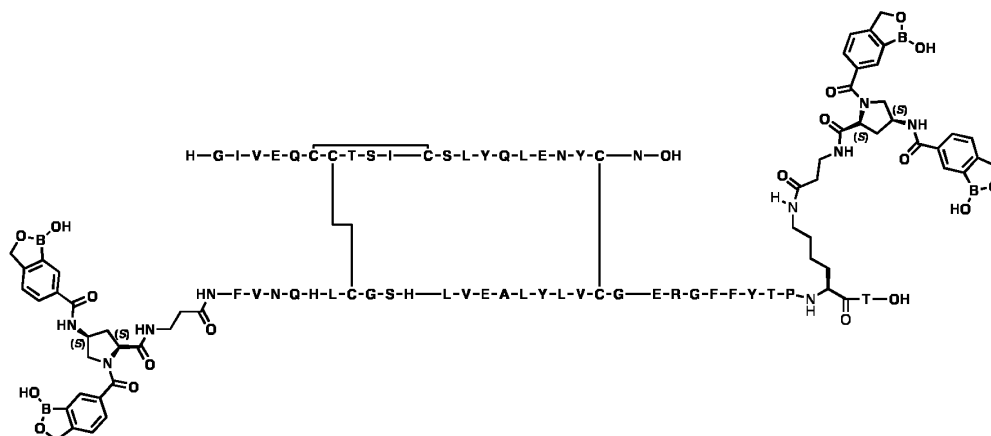
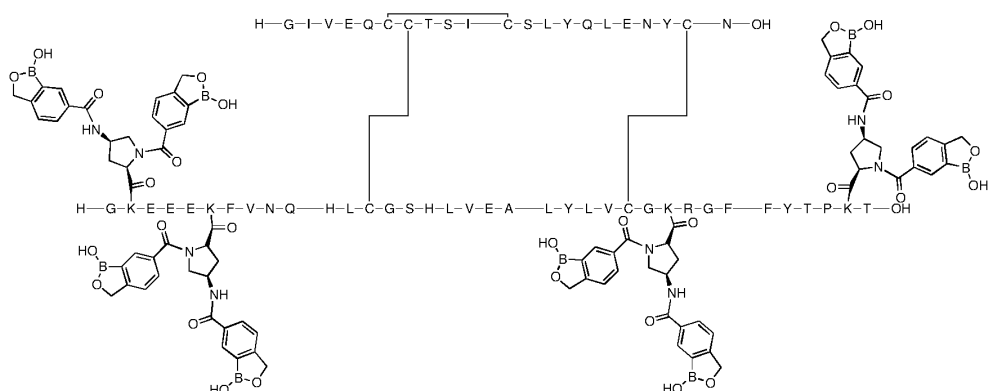
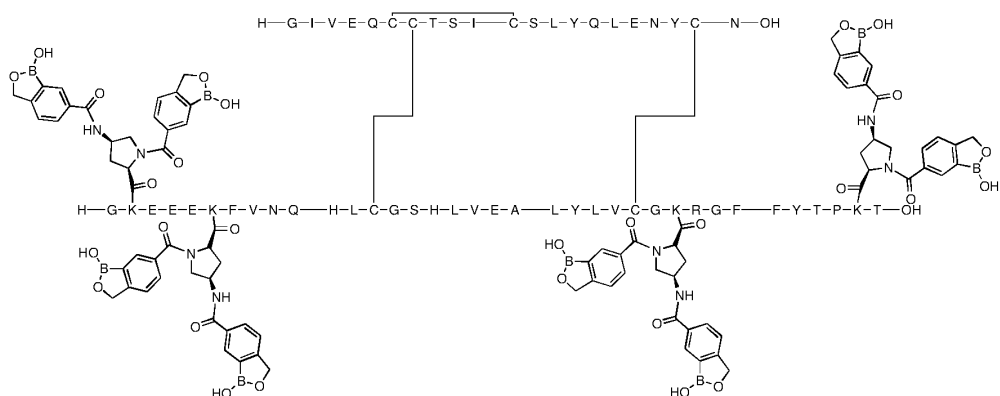
5 Example 849:

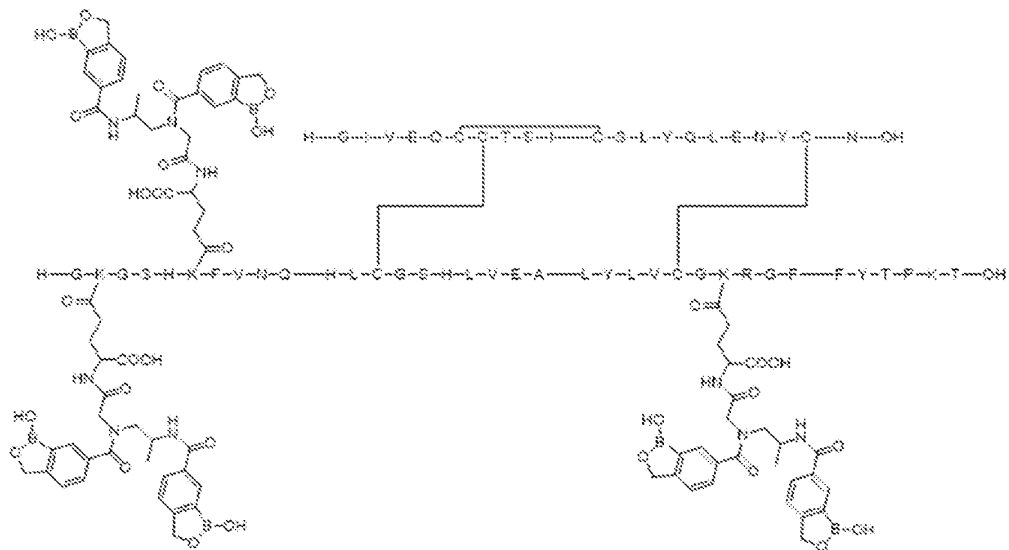
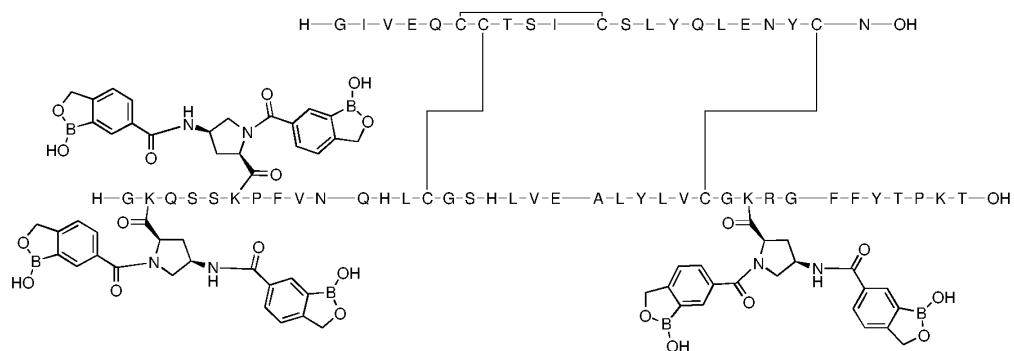


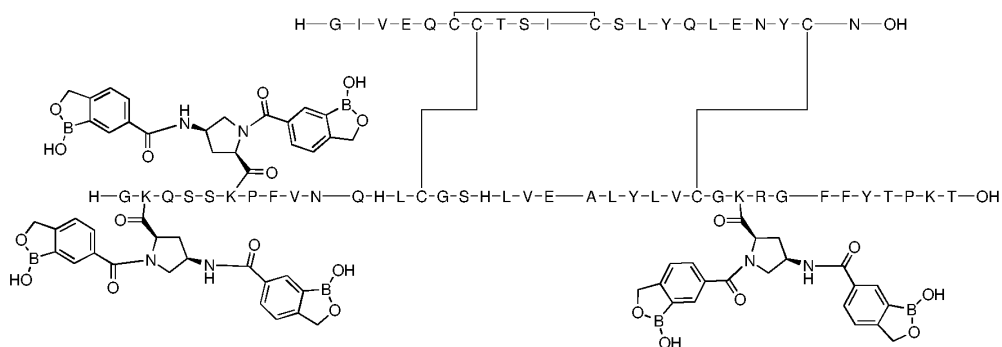
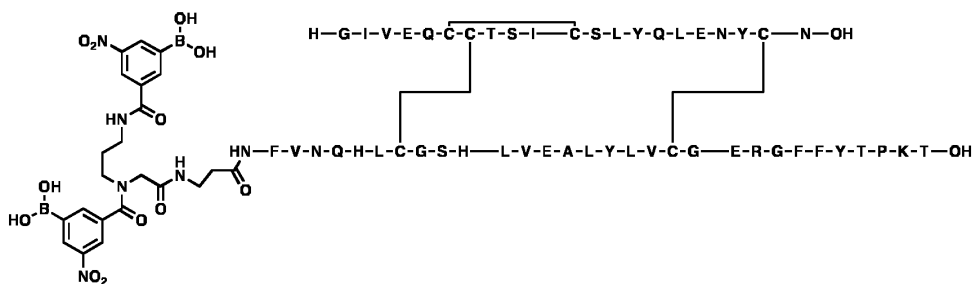
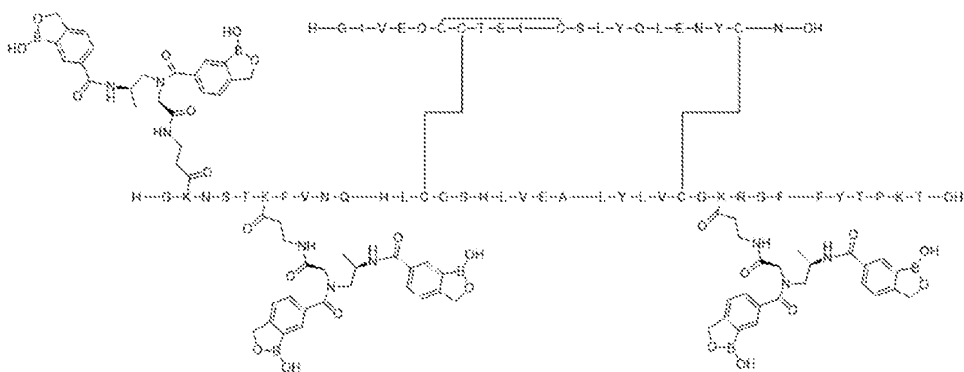
Example 850:**Example 851:**

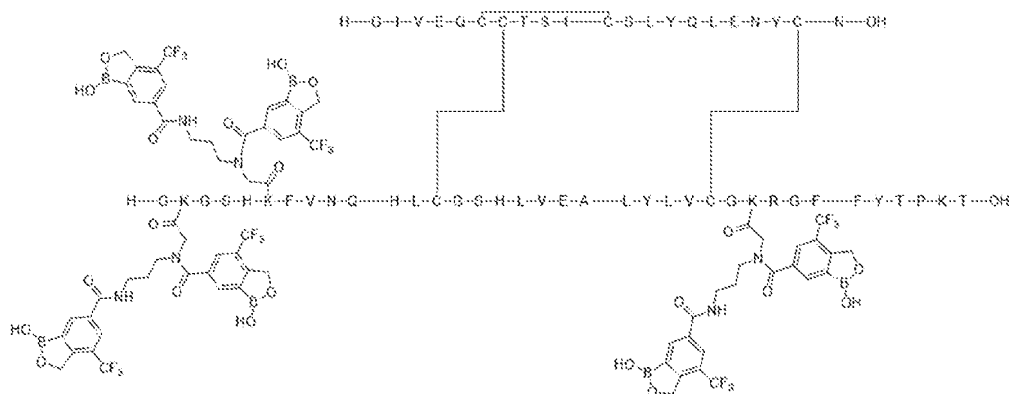
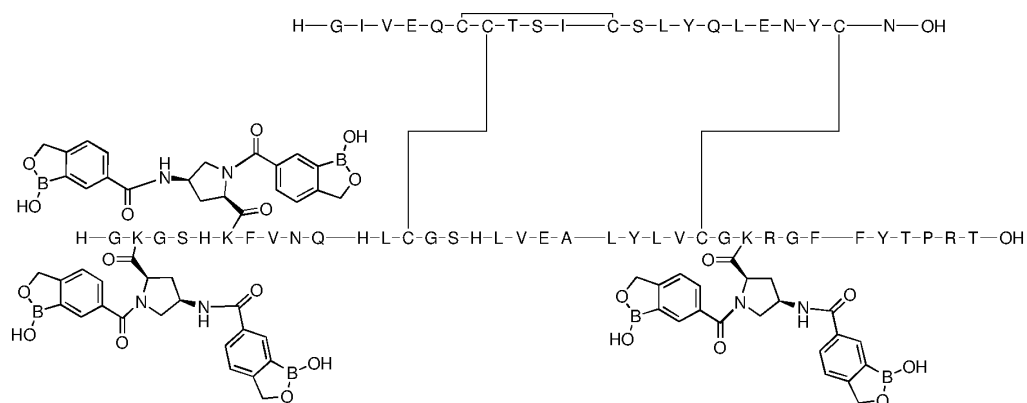
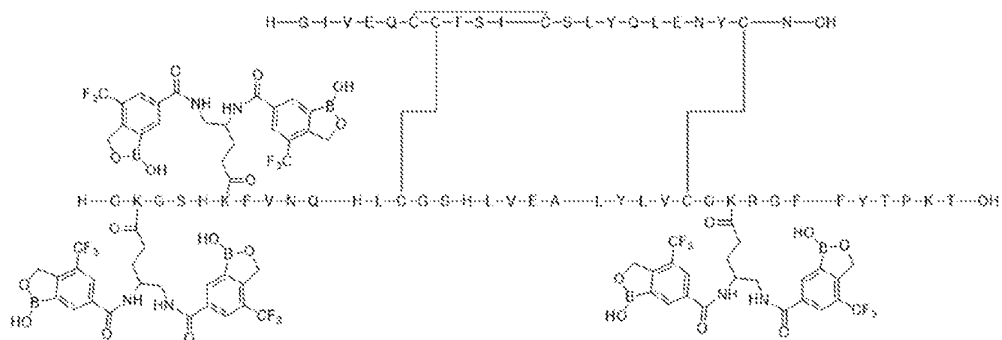
5

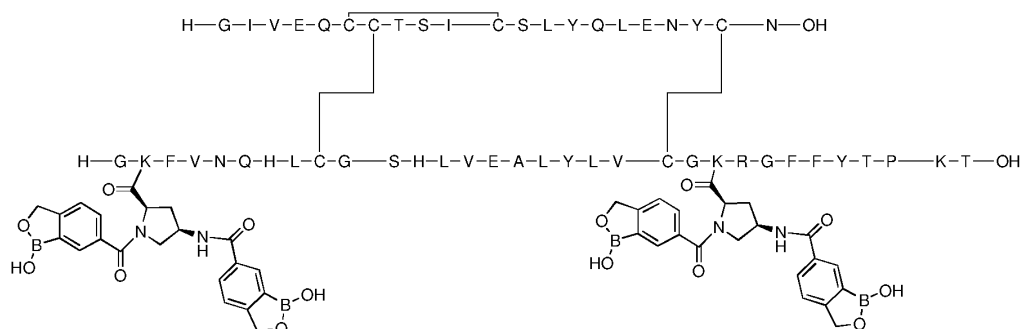
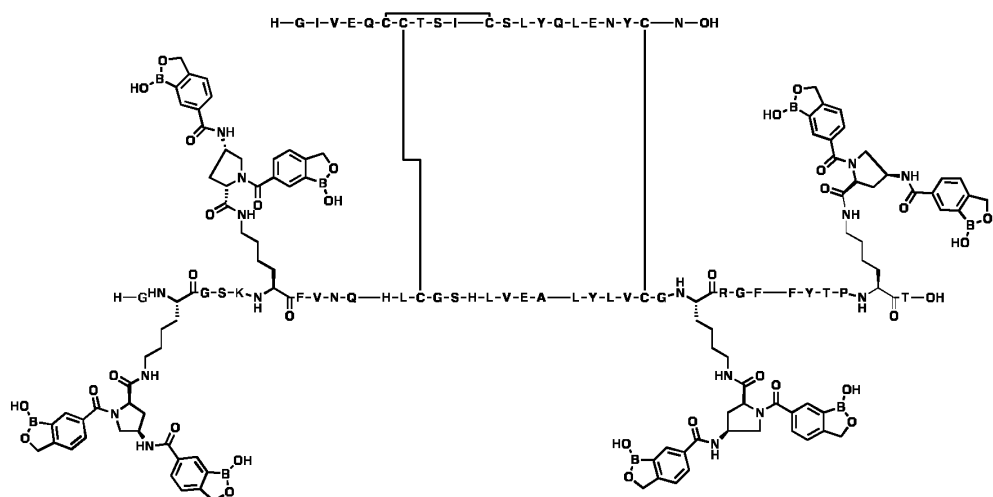
Example 852:

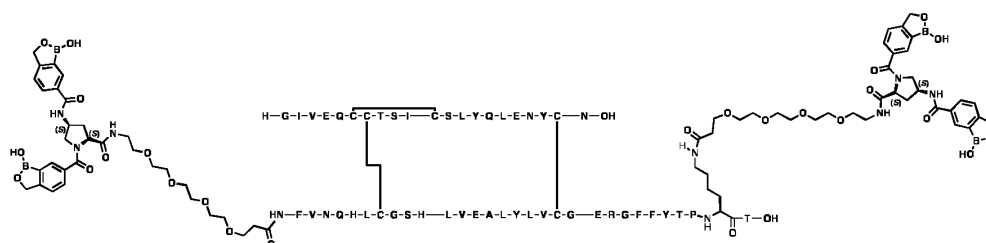
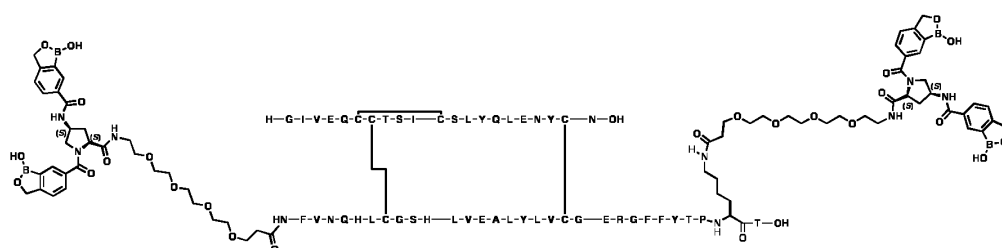
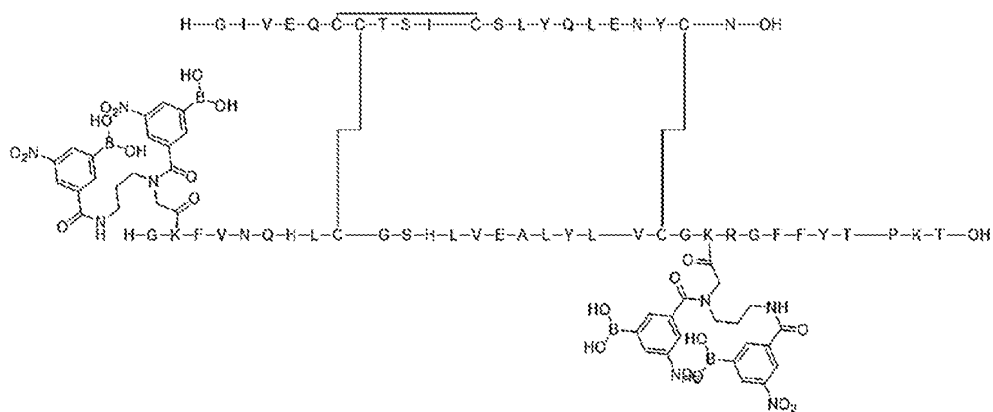
Example 853:**Example 854:****5 Example 855:**

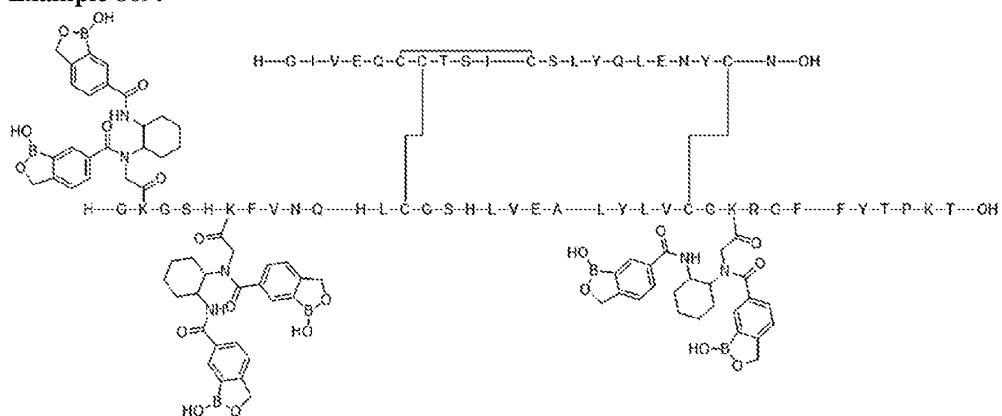
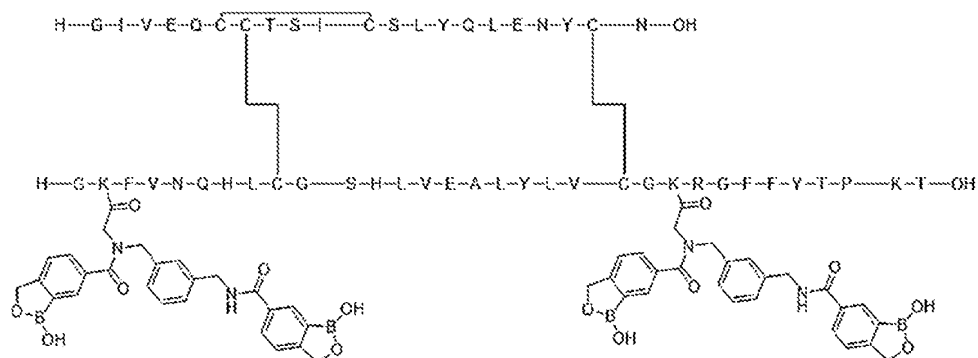
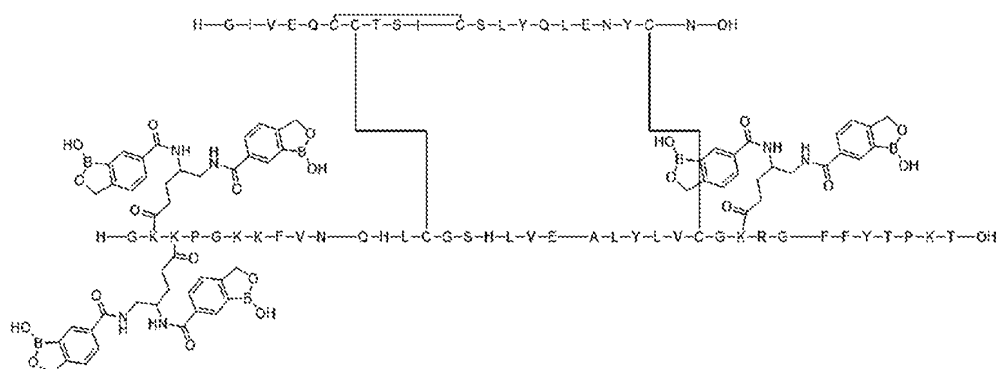
Example 856:**Example 857:**

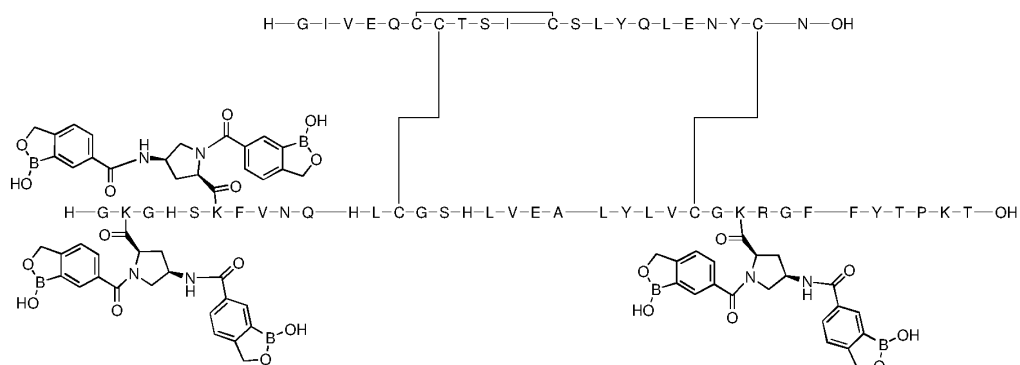
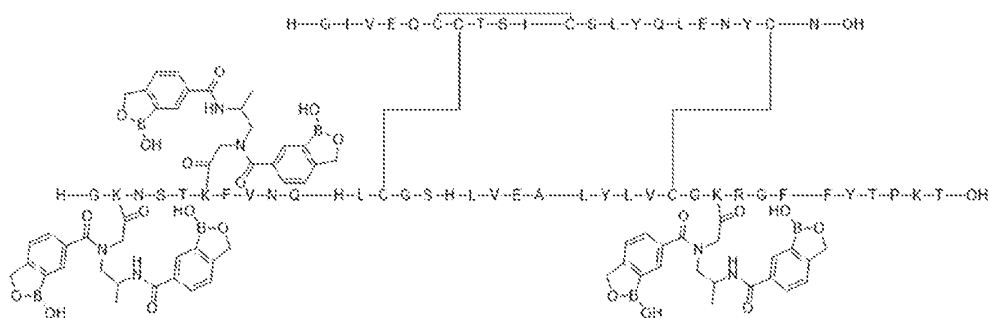
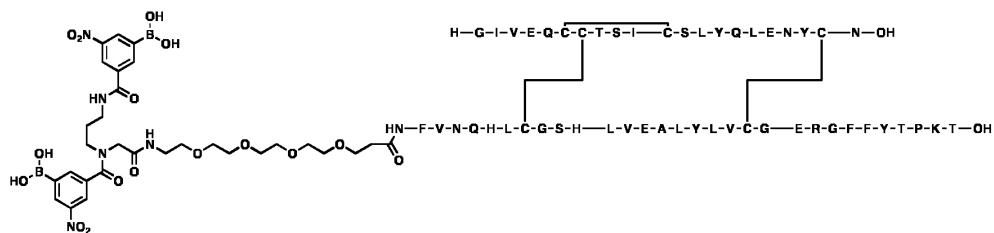
Example 858:**Example 859:****5 Example 860:**

Example 861:**Example 862:****5 Example 863:**

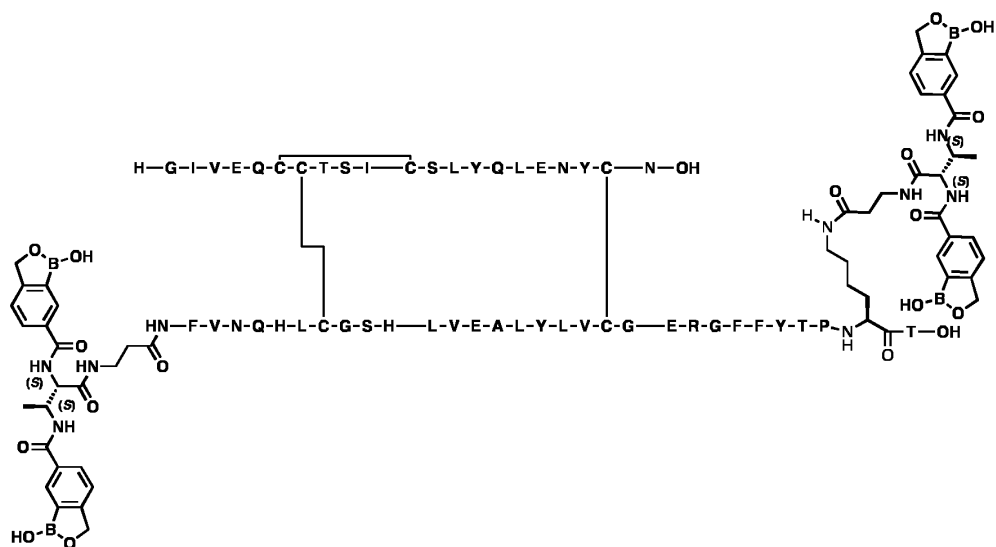
Example 864:**Example 865:**

Example 866:**Example 867:****5 Example 868:**

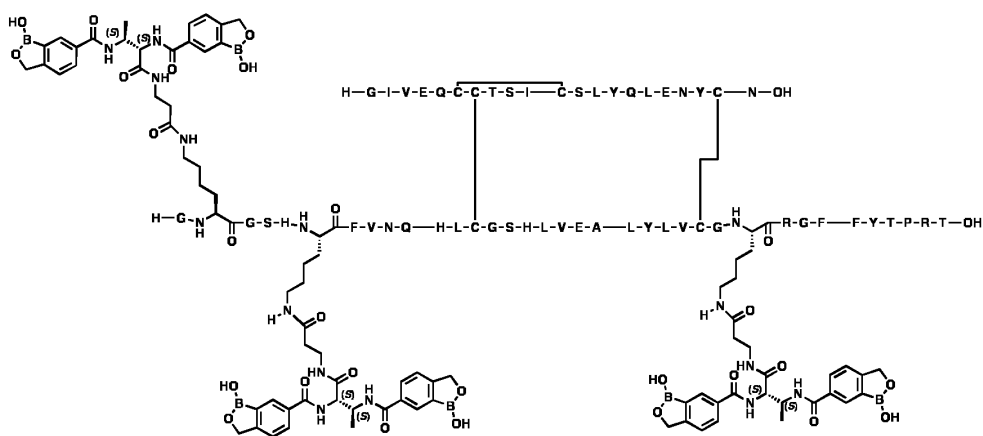
Example 869:**Example 870:****5 Example 871:**

Example 872:**Example 873:****5 Example 874:**

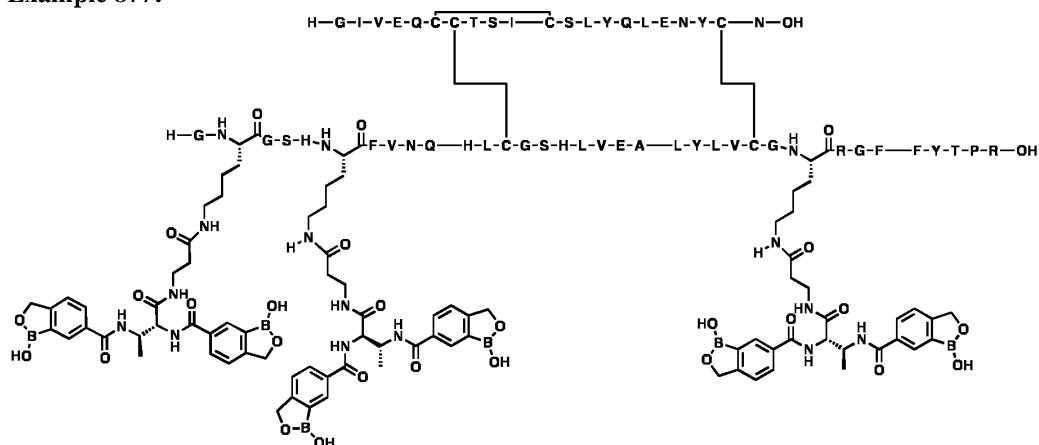
Example 875:



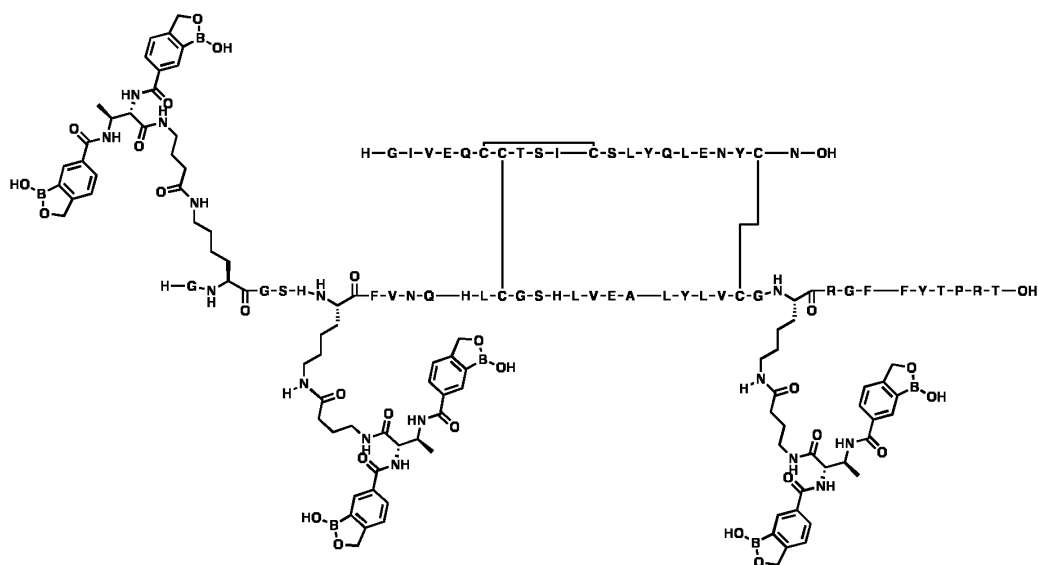
Example 876:



Example 877:

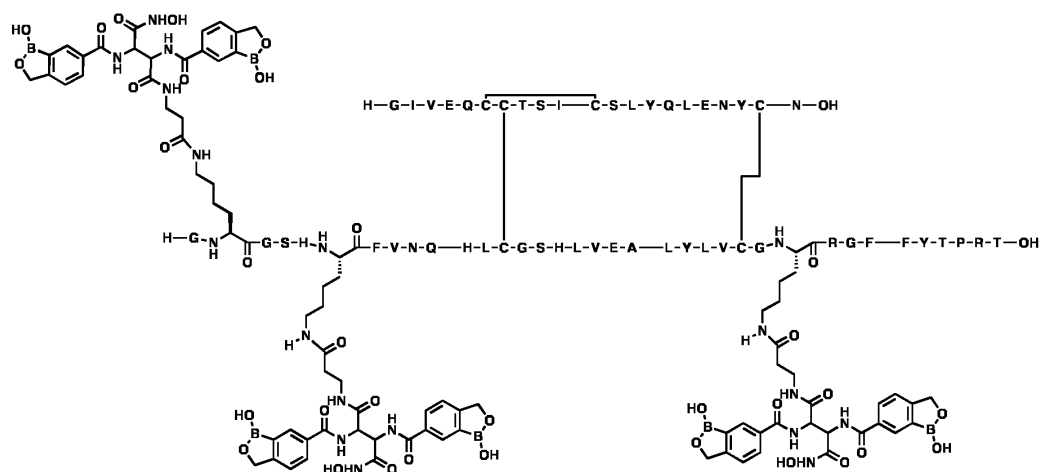


5 Example 878:

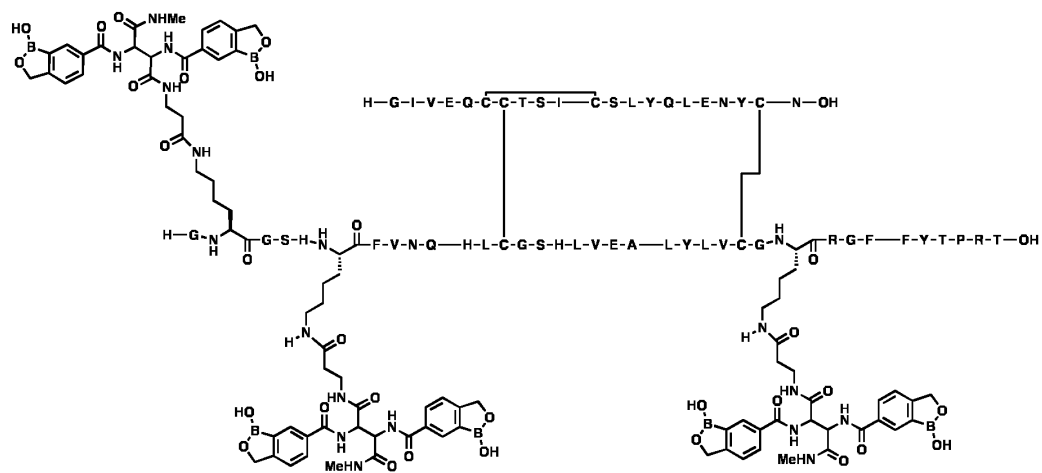


10

15



5 Example 880:



5 WHAT IS CLAIMED IS:

1. A compound comprising X1 and one or more Z1c, or a tautomer, stereoisomer or a mixture of stereoisomers, or a pharmaceutically acceptable salt, or hydrate, or isotopic derivative thereof,

wherein:

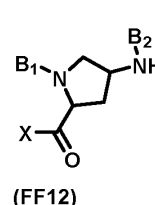
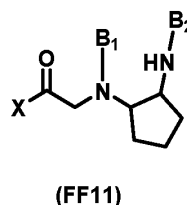
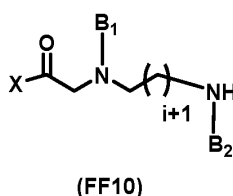
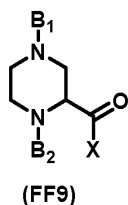
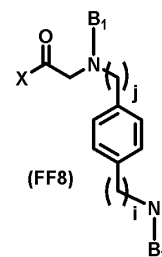
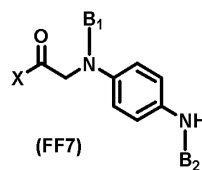
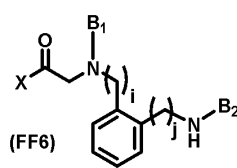
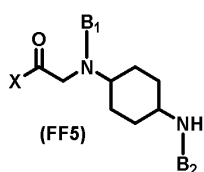
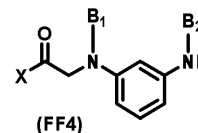
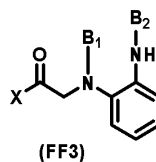
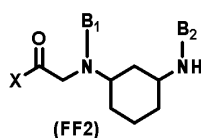
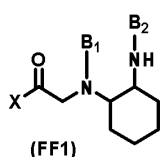
10 X1 comprises:

- (i) NH_2 or OH ;
- (ii) a drug substance comprising an amine;
- (iii) a drug substance that is covalently conjugated to an amine containing linker; or
- 15 (iv) an amine configured to be covalently conjugated to a drug substance;

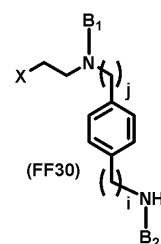
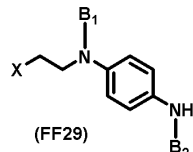
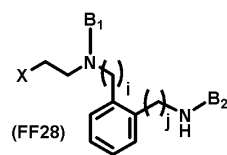
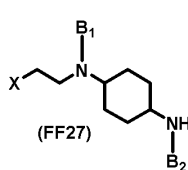
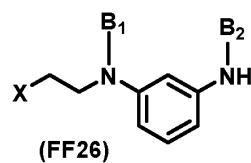
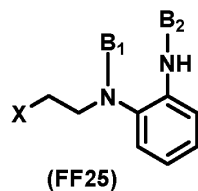
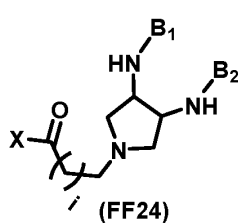
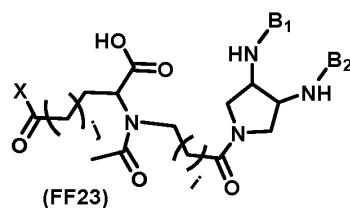
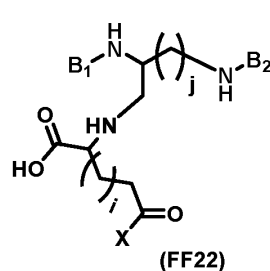
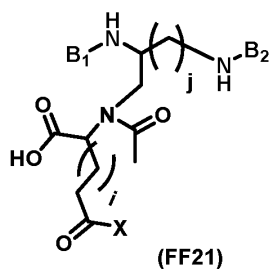
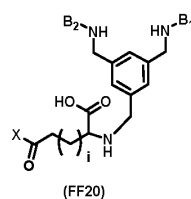
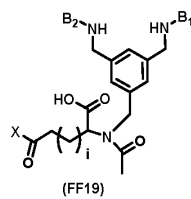
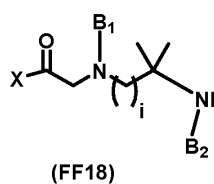
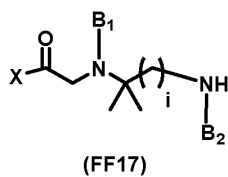
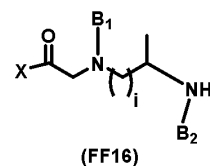
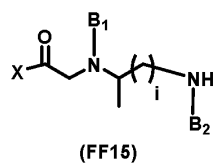
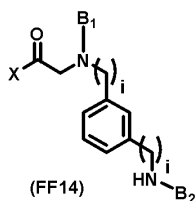
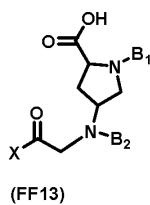
wherein each Z1c is independently selected from Formulae FF1-FF224;

wherein each Z1c is covalently conjugated, directly or indirectly, to an amine in X1 or to OH when X1 is OH , and

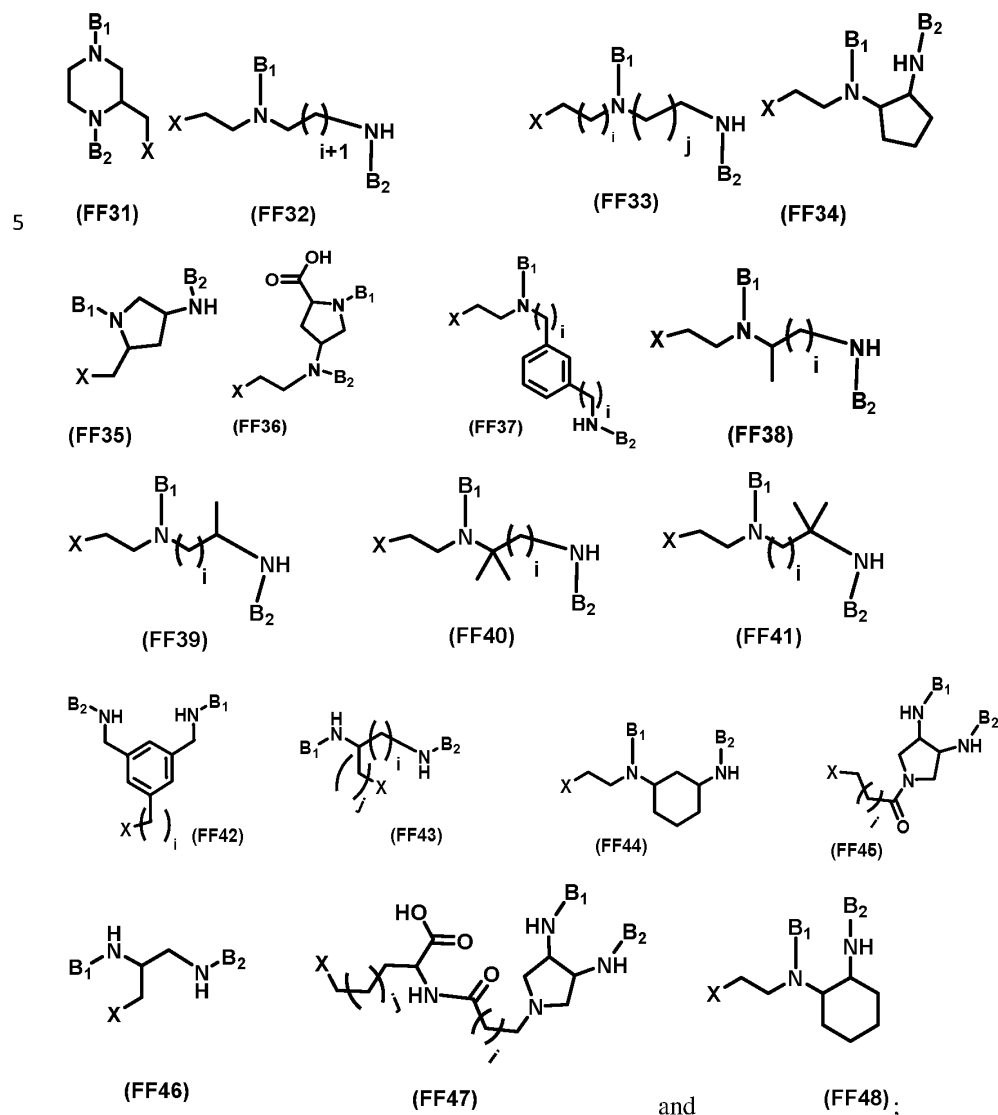
20 wherein Formulae FF1-FF48 are:



5



10



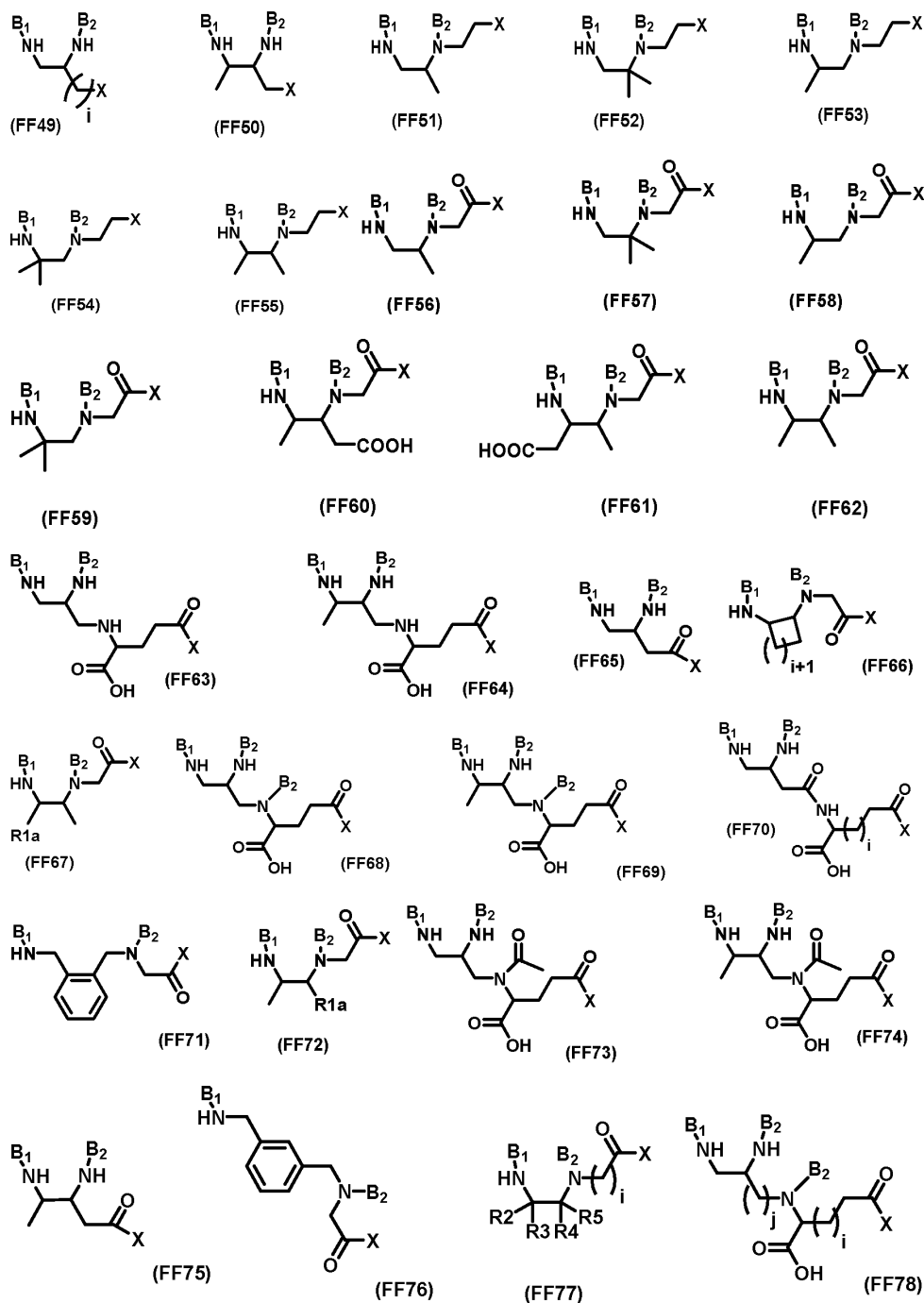
10 wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

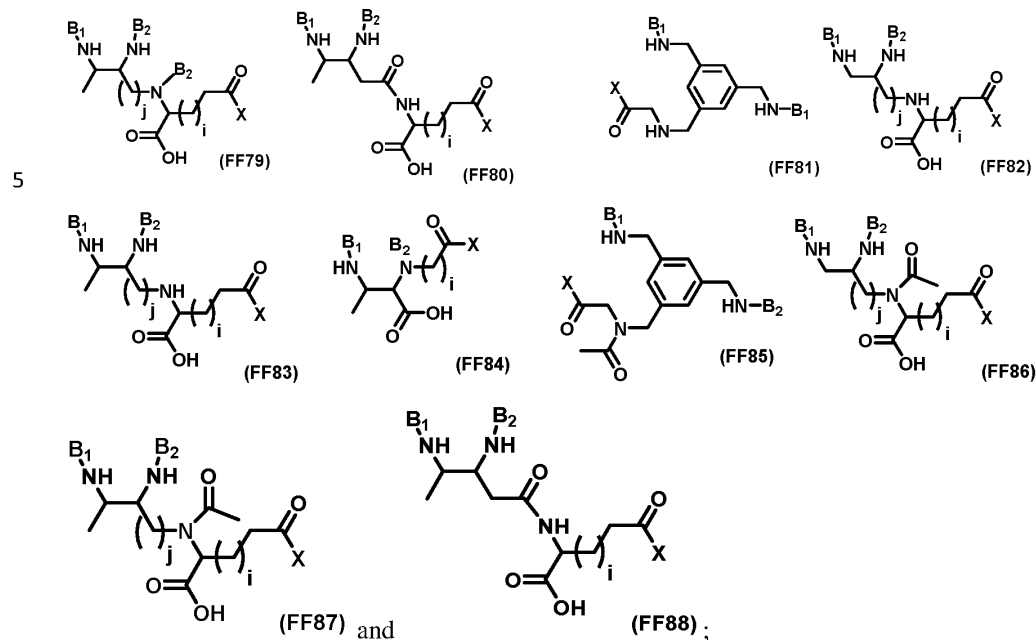
i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7; and

B₁ and B₂, which may be identical or different, each independently represents an
15 aromatic boron-containing group;

5 wherein Formulae FF49-FF88 are:





wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

10 i is 1, 2, 3, 4, 5, 6, or 7;

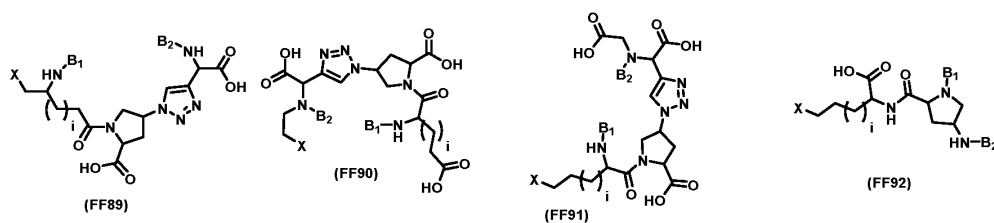
j is 1, 2, 3, 4, 5, 6, or 7;

R1a is selected from COOH, CH₃, H, and OH;

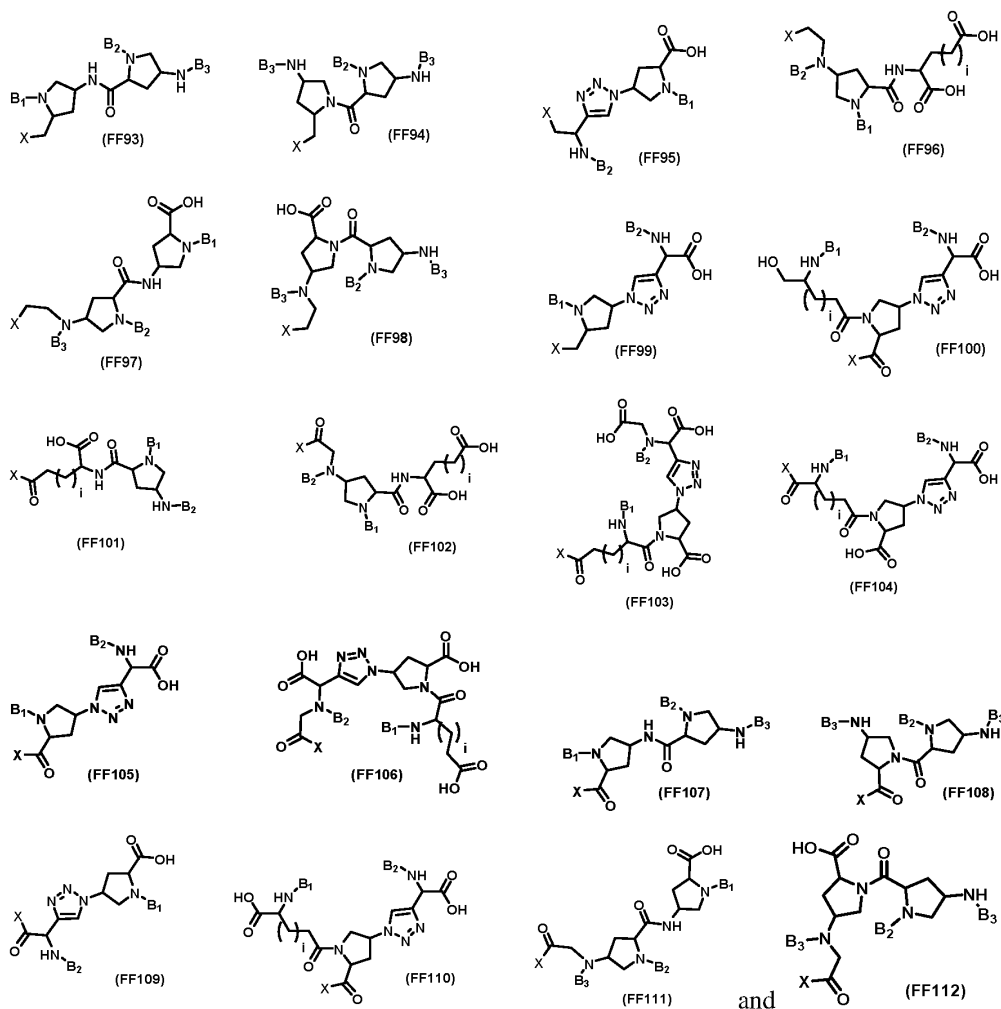
R2, R3, R4 and R5 are each independently selected from CH₃, H, OH, and COOH, and at least one of R2, R3, R4 and R5 is CH₃ or OH; and

15 B₁ and B₂, which may be identical or different, are each independently an aromatic boron-containing group;

wherein Formulae FF89-FF112 are:



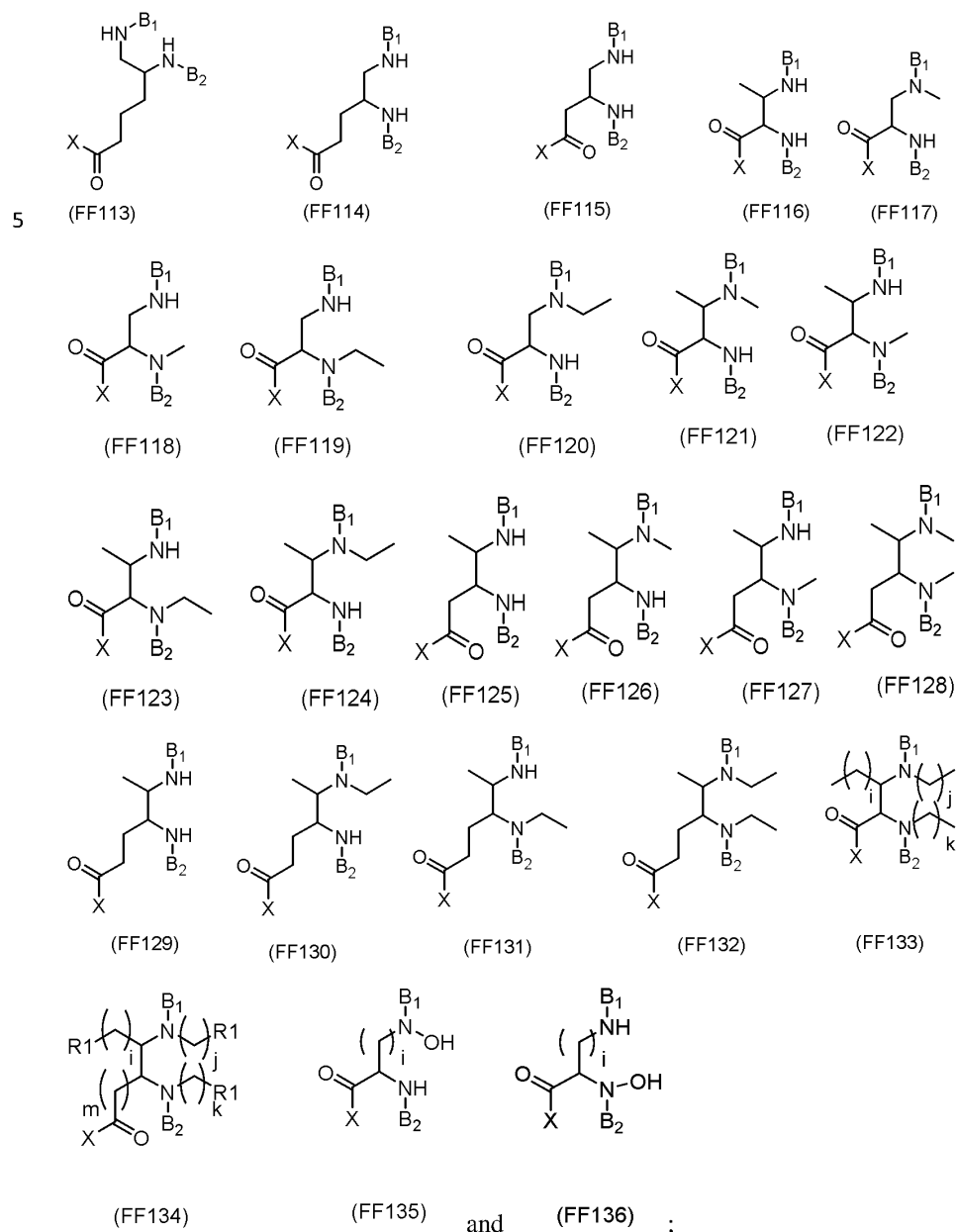
5



10 wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH; i is 1, 2, 3, 4, 5, 6, or 7; and

B₁, B₂ and B₃, which may be identical or different, are each independently an aromatic boron-containing group, a carboxylic acid derivative, or a H, wherein in each FF89-FF112
 15 structure containing B₁, B₂ and B₃ groups, at least two of the B₁, B₂ and B₃ groups are independently an aromatic boron-containing group;

wherein Formulae FF113-FF136 are:



10 wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7;

5 k is 1, 2, 3, 4, 5, 6, or 7;

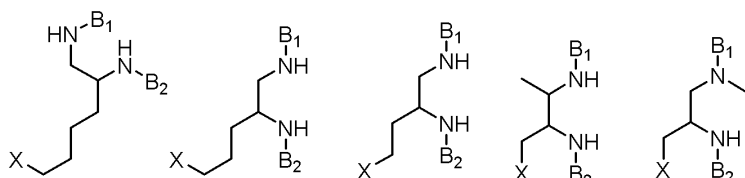
m is 1, 2, 3, 4, 5, 6, or 7;

each R₁ is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R₁ optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl, or aryl

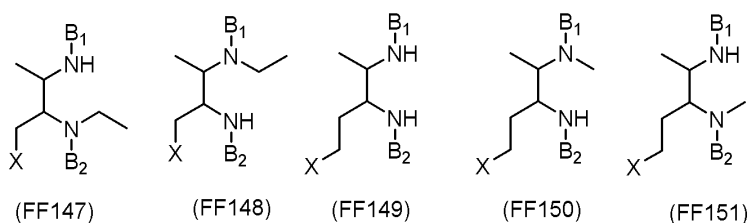
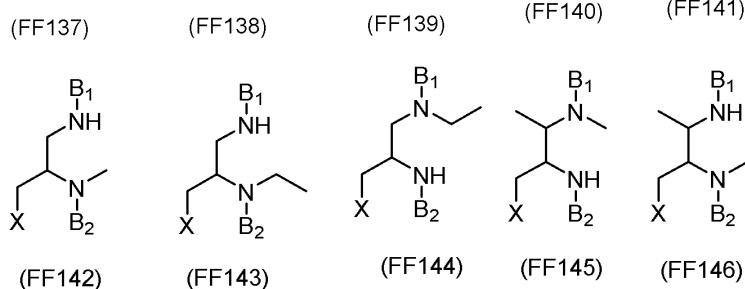
10 groups; and

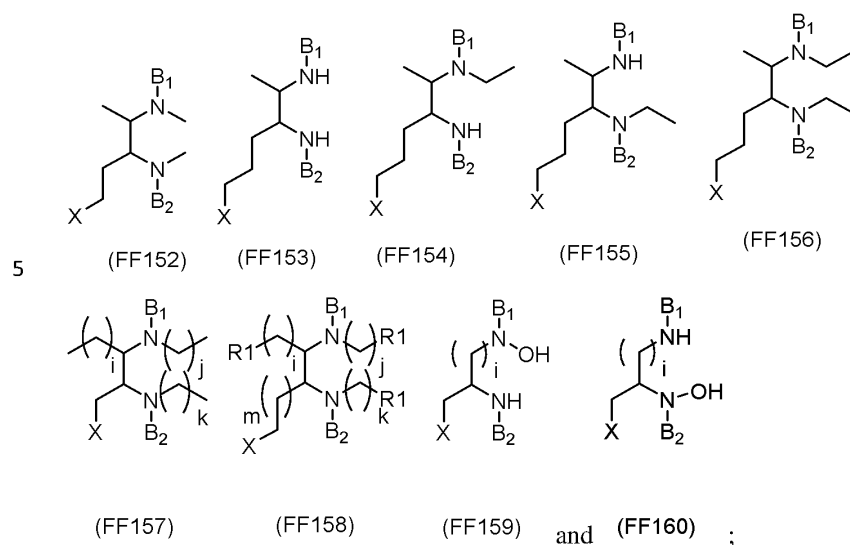
B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group;

wherein Formulae FF137-FF160 are:



15





wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

10 i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7;

k is 1, 2, 3, 4, 5, 6, or 7;

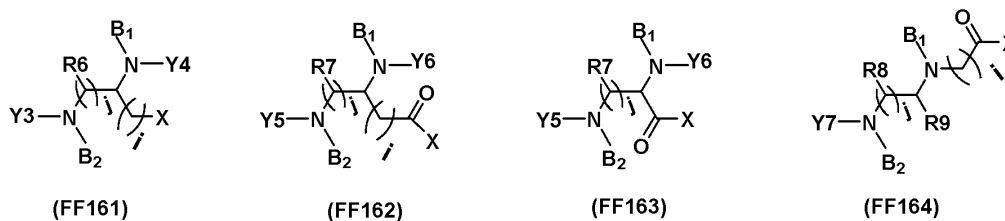
m is 1, 2, 3, 4, 5, 6, or 7;

each R1 is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl, or aryl groups; and

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group;

20

wherein Formulae FF161-FF164 are:



5

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

i is 1, 2, 3, 4, or 5;

j is 1, 2, 3, 4, or 5;

10 each R6, R7, R8, and R9 for different values of j is independently selected from H, CF₃, CH₃, CHF₂, and (CH₂)_mCH₃, wherein m is 1, 2, 3, 4, or 5;

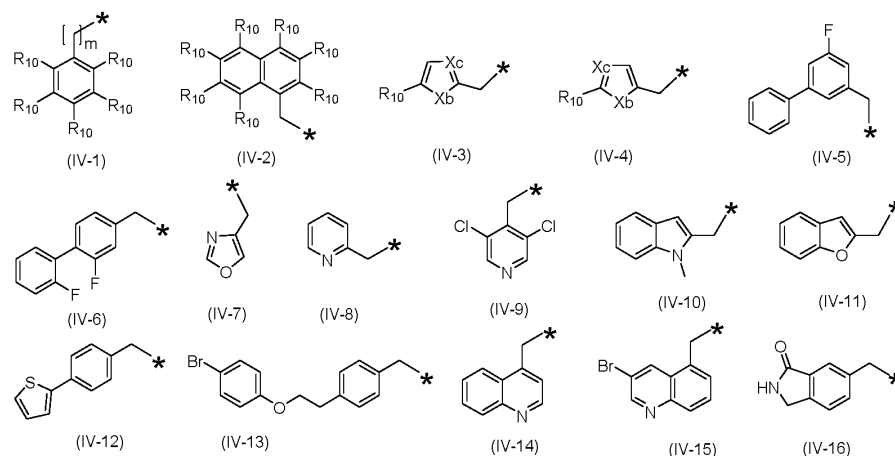
Y3, Y4, Y5, Y6 and Y7 are each independently selected from H, CH₂—X4, and Formulae IV-1 to IV-135;

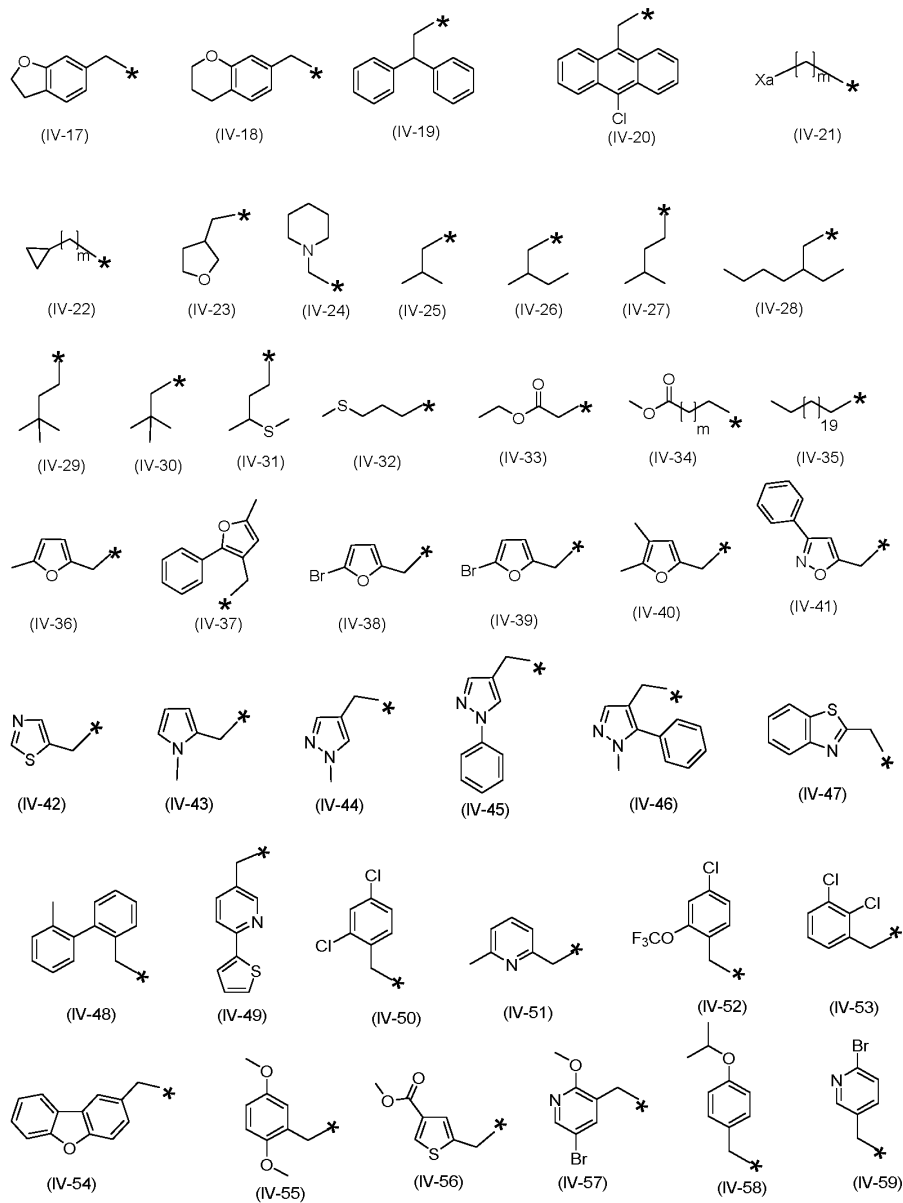
15 wherein X4 is selected from -COOH, -(CH₂)_mCOOH, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each X4 optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; wherein m is 1, 2, 3, 4, or 5;

wherein at least one of Y5, Y6 and Y7 in Formulae FF162 and FF163 is not H and at least one of Y7, R8 and R9 in FF164 is not H; and

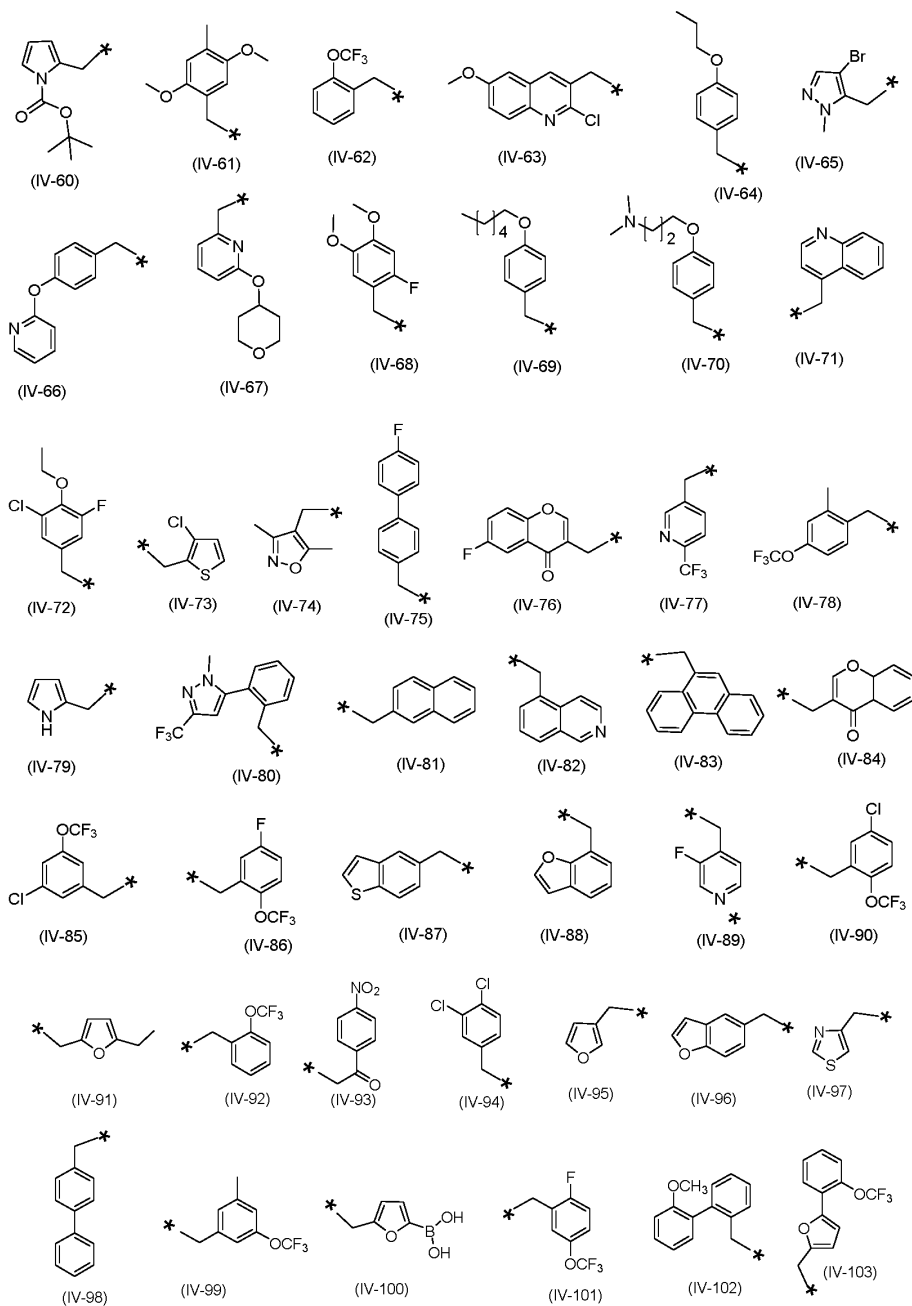
20

wherein Formulae IV-1 to IV-135 are:

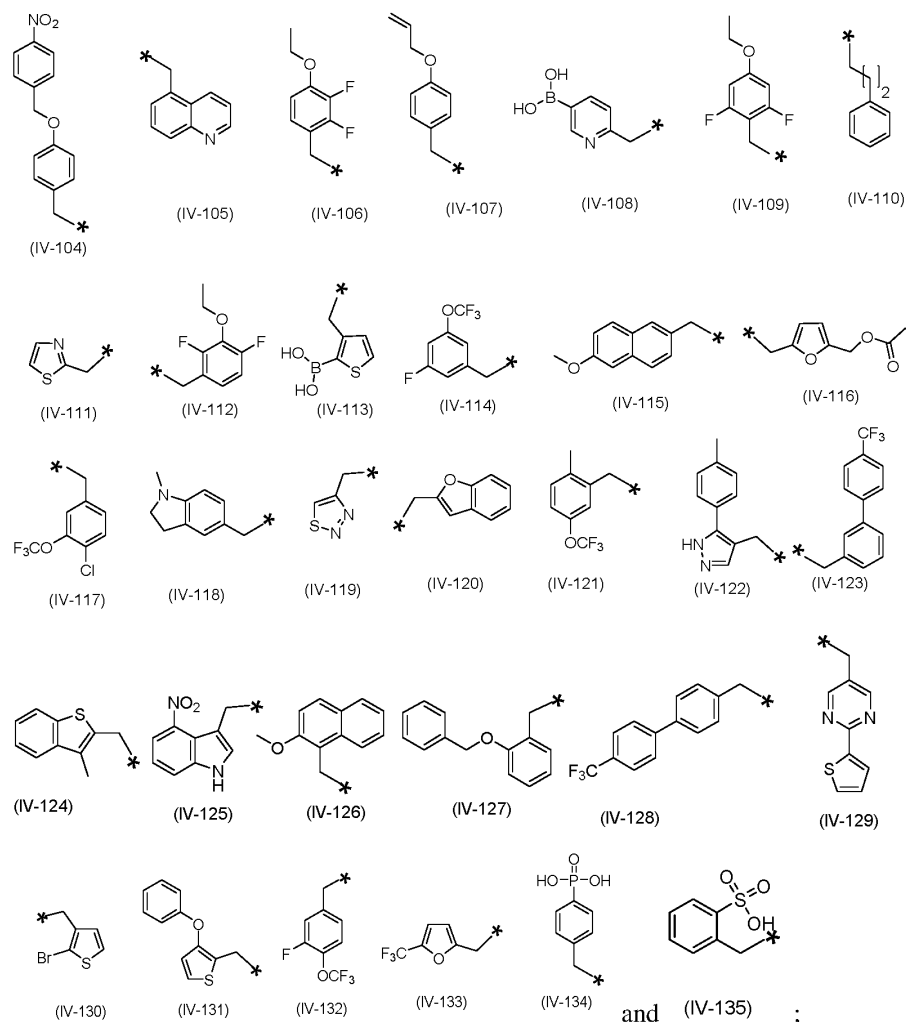




5



5



wherein Xa represents CH=O, CHF₂, CF₃, CH₂SH, COOH, CH₂OH, CH₂NO₂,

10 CH₂NH₂, CH₃, C(CH₃)₃, CH(CH₃)₂, CH((CH₂)₃-CH₃)₂, or CH(CH₂-CH₃)₂;

Xb represents O, NH, CH₂, or S;

Xc represents CH or N;

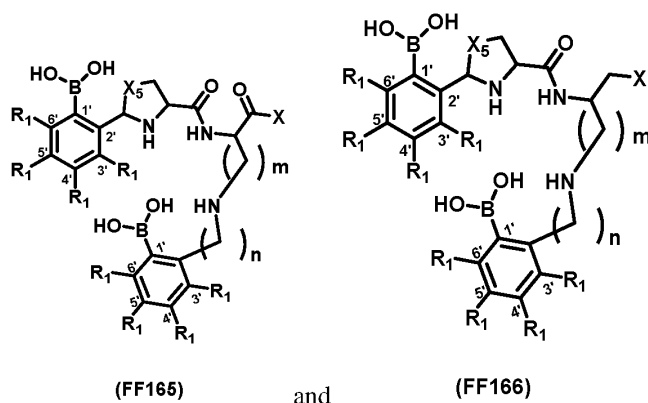
each R₁₀ is independently selected from H, F, Cl, Br, CH₃, CF₃, CH=O, OH, COOH, and (CH₂)_nCH₃,

15 m is 1, 2, 3, 4, or 5; and n is 1, 2, 3, 4, or 5;

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group; and

5 * in Formulae IV-1 to IV-135 represents a point of attachment to corresponding
Formulae FF161-164;

wherein Formulae FF165-FF166 are:



10

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

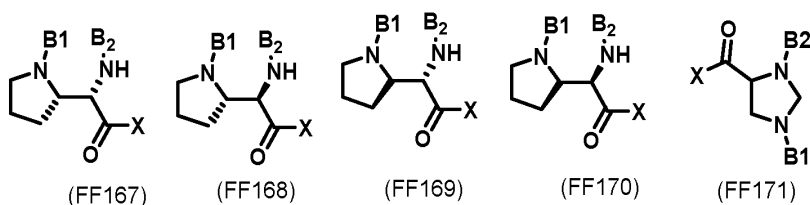
m is 1, 2, 3, 4, 5, 6, or 7;

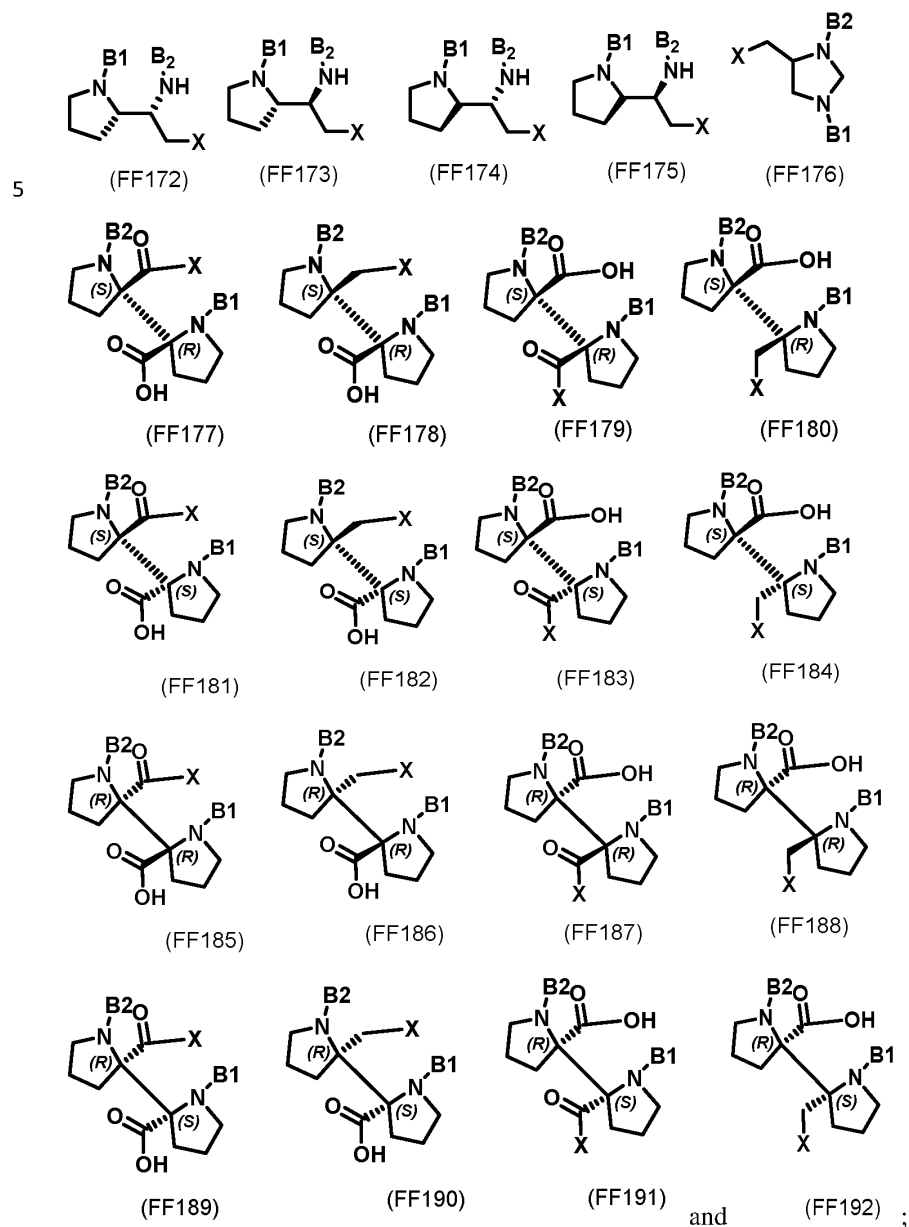
n is 1, 2, 3, 4, 5, 6, or 7;

15 X5 is S, O, or NH; and

each R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH-CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

20 wherein Formulae FF167-FF192 are:

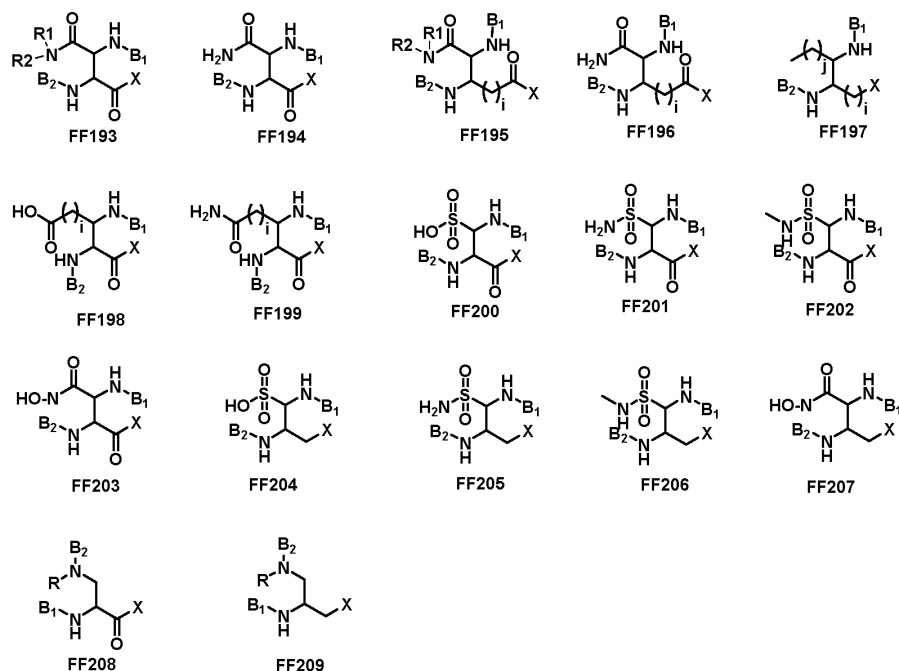




10 wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group;

5 wherein Formulae FF193-FF209 are:



wherein R in FF208 and FF209 is an alkyl, aryl or halide that is covalently conjugated through at least one CH₂ group to the amino group in the side chain of FF208 or FF209,

R1 and R2 are independently selected from H, CH₃, alkyl, and formulae IV-1 to IV-

10 135;

i is 1, 2, 3, 4, or 5;

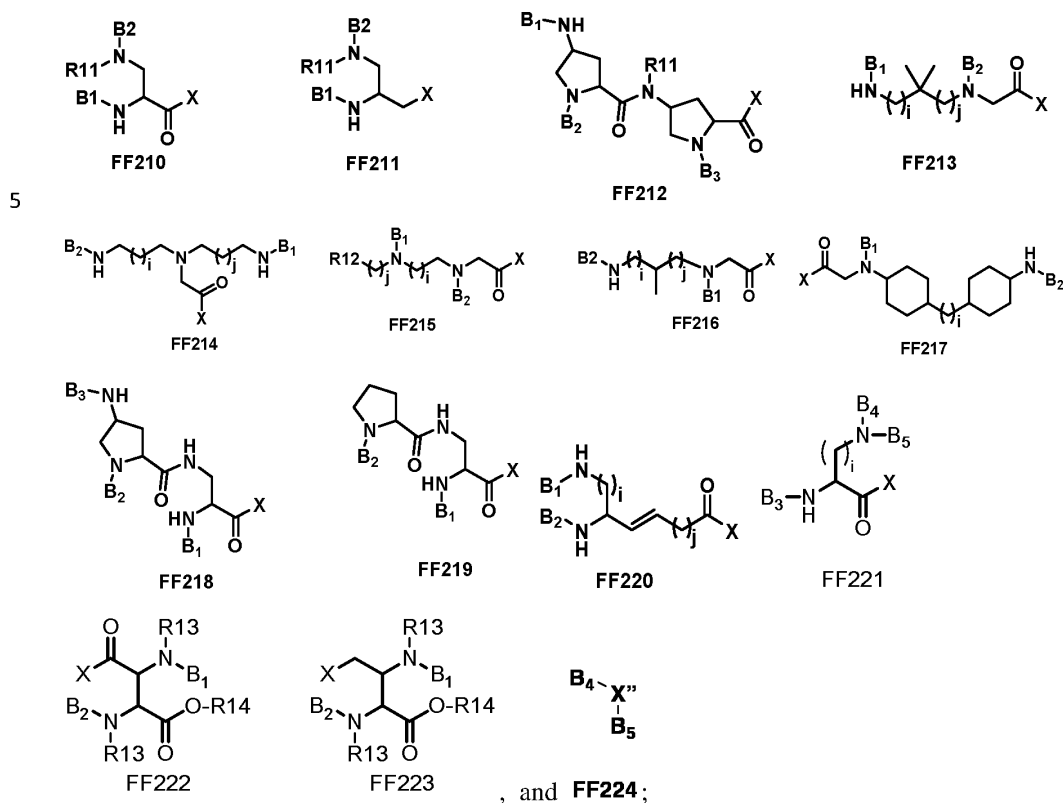
j is 1, 2, 3, 4, or 5; and

wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

15 and

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group, and

wherein Formulae FF210-FF224 are:



10 wherein R11 in FF210 to FF212 is selected from Formulae IV-1 to IV-135 and R12 is selected from an amine, a hydroxyl, an alkyl, and a halide group;

wherein each R13 is independently selected from H, CH₃, alkyl, aryl and Formulae IV-1 to IV-135; R14 is selected from H, CH₃, alkyl, aryl and heteroaryl;

15 wherein X represents a point of covalent attachment either directly to an amine in X1 or to an amine that is covalently conjugated directly or indirectly to X1, or to OH when X1 is OH;

X'' represents a point of covalent attachment to an amine --N in the compound, wherein -- represents a single covalent bond to a CH₂ or CH group in the compound;

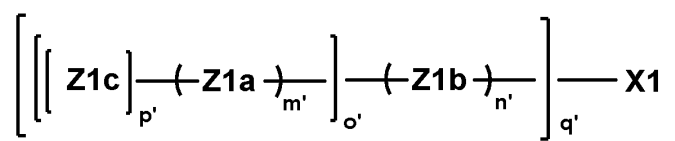
i is 1, 2, 3, 4, or 5;

j is 1, 2, 3, 4, or 5; and

20 wherein at least one primary or secondary amine in FF1-FF223 is optionally covalently conjugated to B₆; and

5 B₁, B₂, B₃, B₄, B₅, and B₆ each independently represents an aromatic boron-containing group, wherein in each FF structure containing B₁, B₂ and B₃ groups, at least two of the B₁, B₂ and B₃ groups are independently an aromatic boron-containing group.

2. The compound of claim 1, wherein the compound is a molecular conjugate represented by Formula I, or a stereoisomer or a mixture of stereoisomers, or pharmaceutically acceptable salt thereof:



(Formula I)

wherein

X1 comprises:

- 15 (i) NH₂ or OH,
 (ii) a polypeptide drug substance comprising an amine,
 (iii) a polypeptide drug substance that is covalently conjugated to an amine containing linker, or
 (iv) an amine configured to be covalently conjugated to a polypeptide drug substance;

20 each Z1c is independently selected from Formulae FF1-FF224 and covalently conjugated either directly, or via Z1a and/or Z1b, to X1;

each Z1a independently comprises 1 to 50 amino acids connected together using amide or peptide bonds;

25 each Z1b is independently a small-molecule linker;

each m' is independently 0 or 1;

each n' is independently 0 or a positive integer;

each o' is independently an integer greater than or equal to 1;

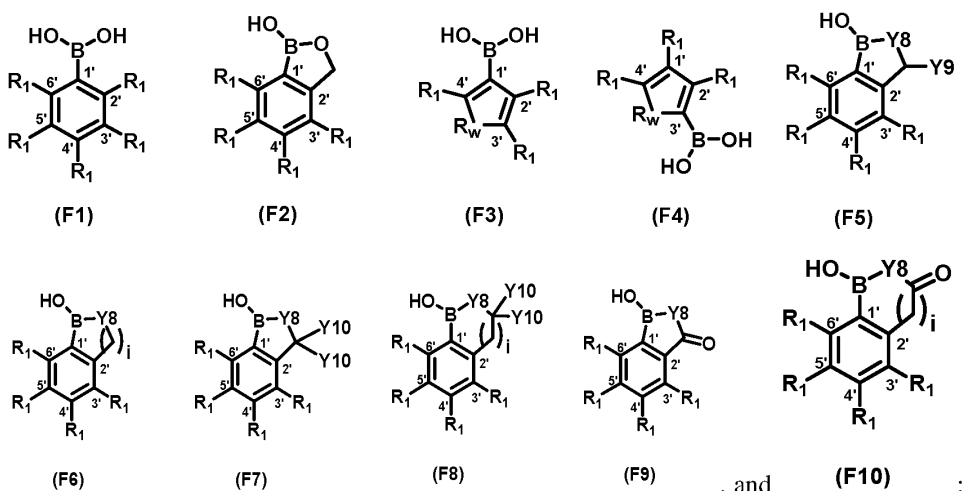
each p' is a positive integer; and

30 q' is a positive integer of at least 1 and not more than two times the total number of amine groups in X1, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different;

wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1; and

5 wherein optionally the molecular conjugate may comprise one or more isotopes at any position of the molecular conjugate of Formula I.

3. The compound of claim 1 or 2, wherein the compound comprises at least one of B₁, B₂ and B₃ independently selected from Formulae F1-F12 or wherein the compound comprises at least one of B₄, B₅ and B₆ independently selected from Formulae F1-F10, wherein Formulae F1-F10 are:



wherein for B₁, B₂, and B₃:

one R₁ represents (C=O)---*, S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein
15 ---* represents the attachment point to the rest of Z1c, and m is 1, 2, 3, 4, 5, 6, or 7; and

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH-CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

wherein for B₄ and B₅:

20 one R₁ for B₄ represents (CH₂)_m---ø, wherein ---ø represents an attachment point to the rest of Z1c and one R₁ for B₅ represents (C=O)---*, S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein ---* represents an attachment point to the rest of Z1c, and m is 1, 2, 3, 4, 5, 6, or 7; and

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH-CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

wherein for B₆:

5 one R_1 for B_6 represents $(CH_2)_m---$, wherein $---$ represents an attachment point to the rest of the compound, and m is 1, 2, 3, 4, 5, 6, or 7; and

each remaining R_1 is independently selected from H, F, Cl, Br, OH, CH_2-NH_2 , NH_2 , $(C=O)-NH_2$, $CH=O$, SO_2CH_3 , SO_2CF_3 , CF_3 , CHF_2 , NO_2 , CH_3 , OCH_3 , $O(CH_2)_mCH_3$, $-(SO_2)NH-CH_3$, $-(SO_2)NH(CH_2)_mCH_3$, and OCF_3 , wherein m is 1, 2, 3, 4, 5, 6, or 7;

10

wherein,

for Formulae F3-F4:

R_w is O or S;

for Formulae F5-F10:

15 Y_8 is selected from O, N, and NR, wherein R is an alkyl group or H;

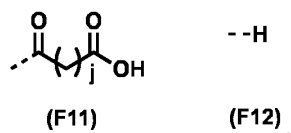
Y_9 is H, CH_3 , or an alkyl group, provided that when Y_8 is O, Y_9 is a CH_3 or an alkyl group;

each Y_{10} is independently selected from H, CH_3 , F, CF_3 , and OCH_3 ; and

i is 1, 2, or 3; and

20

wherein Formulae F11-F12 are:

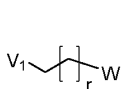


j is 1, 2, 3, 4, 5, 6, or 7; and

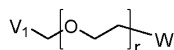
25 $---$ represents an attachment point to the rest of Z1c.

4. The compound of claim 2, wherein the compound comprises at least one Z1b selected from Formulae IIa-IIai and Formulae IIIa-IIIai,

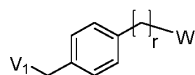
wherein Formulae IIa-IIai are:



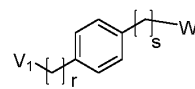
Formula IIa



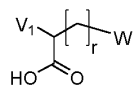
Formula IIb



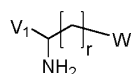
Formula IIc



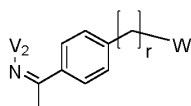
Formula IId



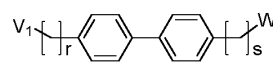
Formula IIe



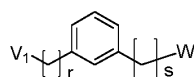
Formula IIff



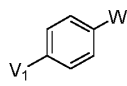
Formula IIg



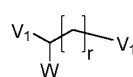
Formula IIh



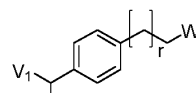
Formula Iii



Formula IIj

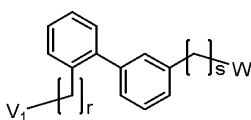


Formula IIk

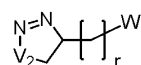


Formula Iil

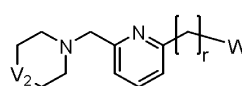
5



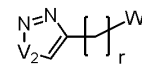
Formula IIm



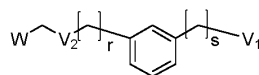
Formula IIln



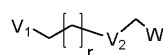
Formula IIo



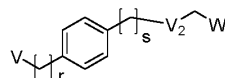
Formula IIp



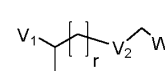
Formula IIq



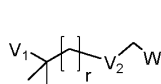
Formula IIr



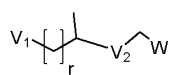
Formula IIs



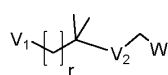
Formula IIt



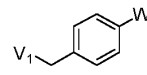
Formula IIu



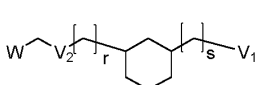
Formula IIv



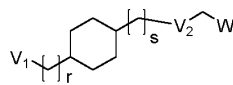
Formula IIw



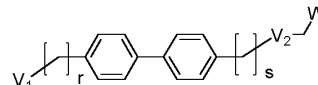
Formula IIx



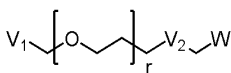
Formula IIy



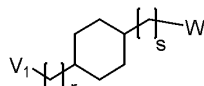
Formula IIz



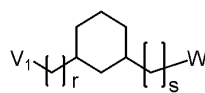
Formula IIaa



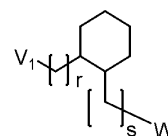
Formula IIab



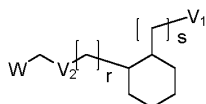
Formula IIac



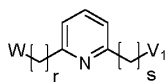
Formula IIad



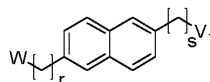
Formula IIae



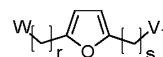
Formula IIaf



Formula IIag



Formula IIah



Formula IIai

and ;

wherein:

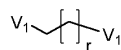
r is 0, 1, 2, 3, 4, or 5;

s is 0, 1, 2, 3, 4, or 5;

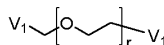
W represents $\text{CH}_2\text{---}$ or $(\text{C}=\text{O})\text{---}$, wherein --- is a covalent linkage to X1; and

each V_1 is independently selected from NH--- , $\text{CH}_2\text{---}$, and $(\text{C}=\text{O})\text{---}$ and each V_2 is N--- , wherein --- is a covalent linkage towards successive Z1b, Z1a or Z1c, provided that V_1 is NH--- when connected to Z1c; and the covalent linkages between Z1a and Z1b units each independently comprise an amine linkage or an amide linkage; and when $n'=0$ and $m'=1$, Z1a is directly conjugated to X1 by an amine linkage or amide linkage, and

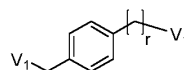
wherein Formulae IIIa-IIIai are:



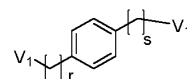
Formula IIIa



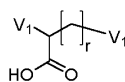
Formula IIIb



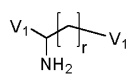
Formula IIIc



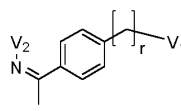
Formula IIId



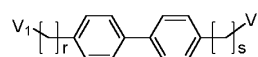
Formula IIIe



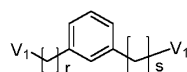
Formula IIIf



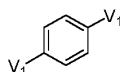
Formula IIIg



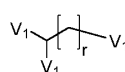
Formula IIIh



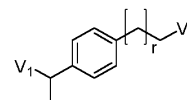
Formula IIIi



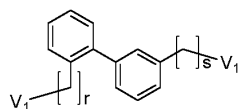
Formula IIIj



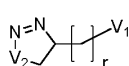
Formula IIIk



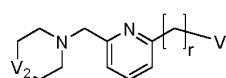
Formula IIIl



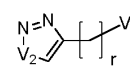
Formula IIIIm



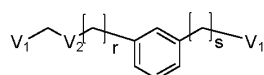
Formula IIIIn



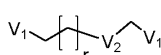
Formula IIIIo



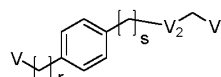
Formula IIIIp



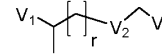
Formula IIIIq



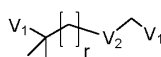
Formula IIIIr



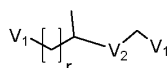
Formula IIIIs



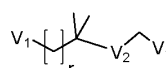
Formula IIIIt



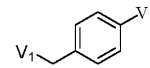
Formula IIIIu



Formula IIIIv

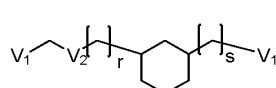


Formula IIIIw

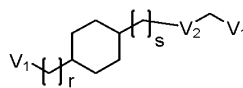


Formula IIIIx

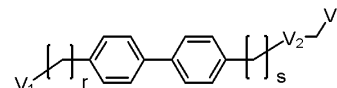
5



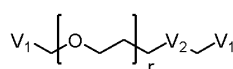
Formula IIIIy



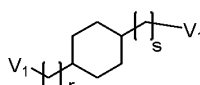
Formula IIIIz



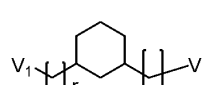
Formula IIIIaa



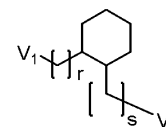
Formula IIIIab



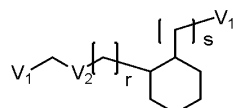
Formula IIIIac



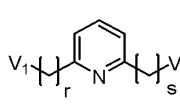
Formula IIIIad



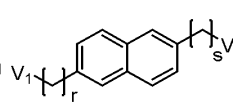
Formula IIIIae



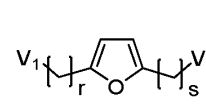
Formula IIIIaf



Formula IIIIag



Formula IIIIah



Formula IIIIai

and ;

wherein:

10

r is 1, 2, 3, 4, or 5;

s is 1, 2, 3, 4, or 5; and

each V_1 is independently selected from $\text{NH---}\dagger$, $\text{CH}_2\text{---}\dagger$, and $(\text{C=O})\text{---}\dagger$ andeach V_2 is $\text{N---}\dagger$, wherein $\text{---}\dagger$ is a covalent linkage towards successive Z1b, Z1a or Z1c,provided that V_1 is $\text{NH---}\dagger$ when connected to Z1c; and the covalent linkages between

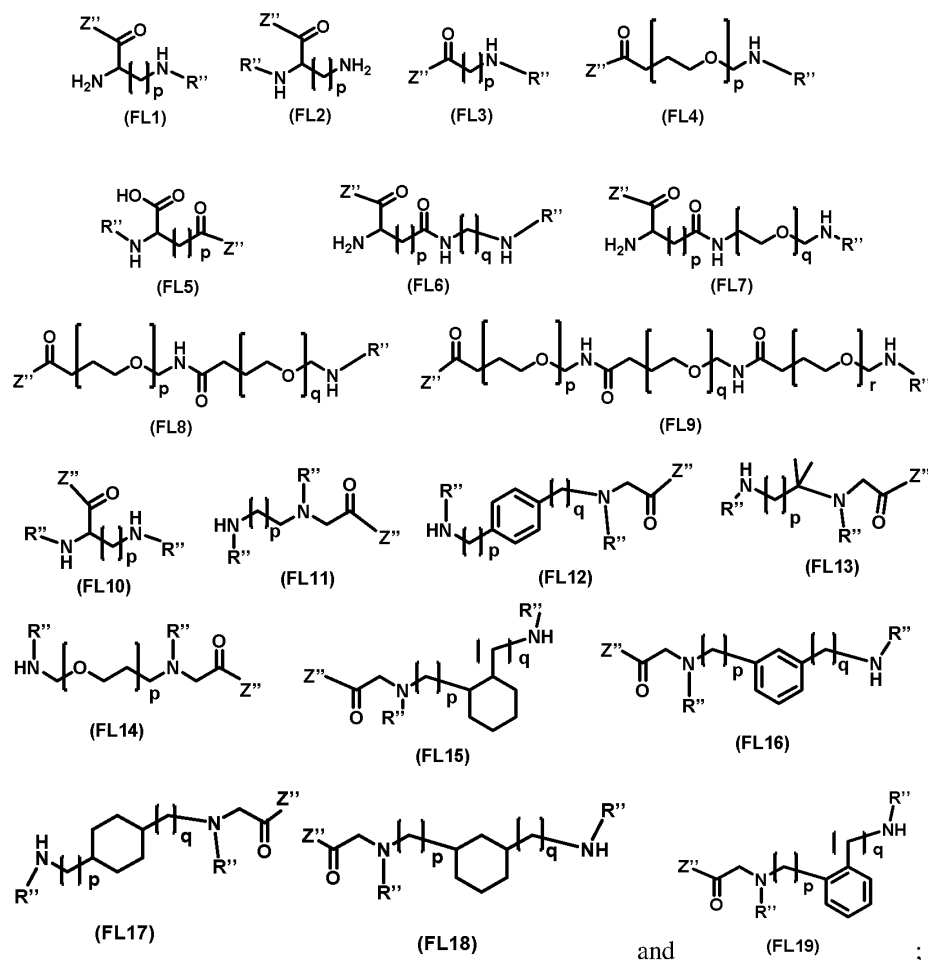
15

Z1a and Z1b units each independently comprise an amine linkage or an amide linkage;

and when $n'=0$ and $m'=1$, Z1a is directly conjugated to X1 by an amine linkage or

amide linkage.

5. The compound of claim 1 or 2, wherein the at least one Z1c is covalently conjugated indirectly via a linker selected from (i) Formulae FL1-FL19:



10 wherein, in Formulae FL1 to FL19:

Z'' represents an attachment point toward X1;

R'' represents an attachment point toward Z1c;

p is 1, 2, 3, 4, or 5,

q is 1, 2, 3, 4, or 5,

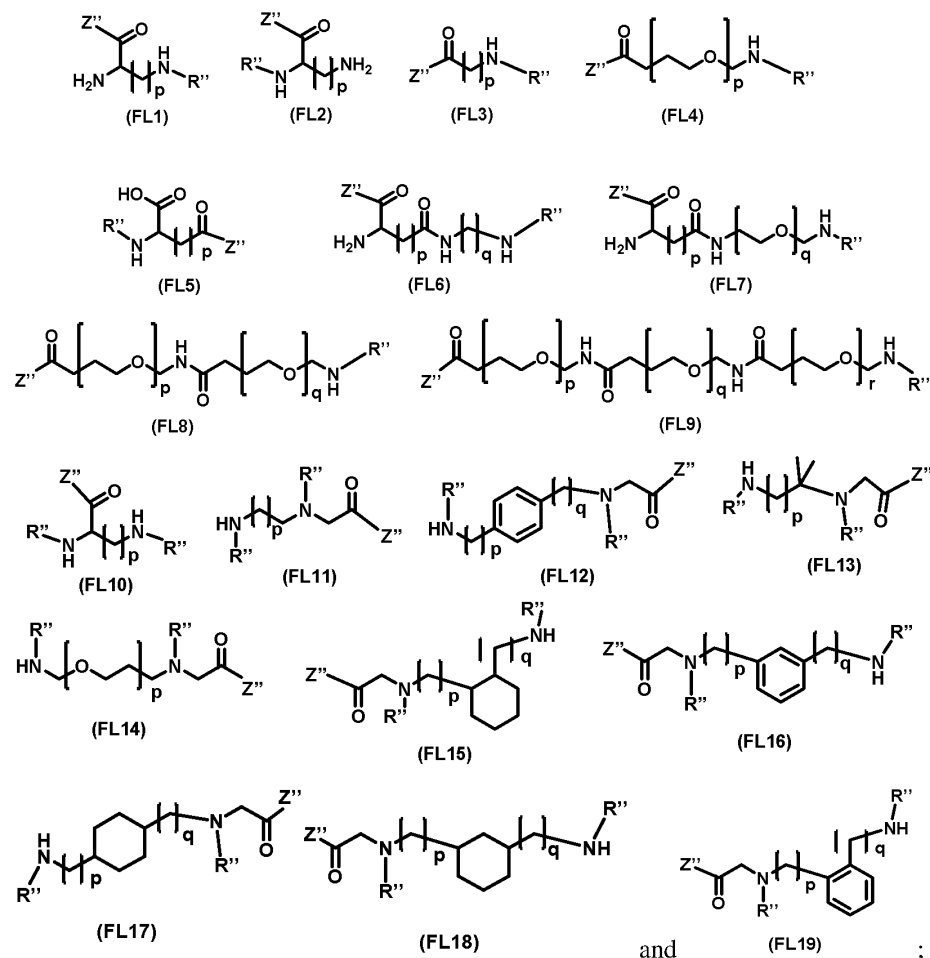
15 r is 1, 2, 3, 4, or 5; and

any primary amine is optionally acetylated or alkylated; and

(ii) an L- or D-amino acid comprising at least one amine group directly conjugated to Z1c, wherein an acid functional group of the amino acid is conjugated toward X1 in Formula I.

5

6. The compound of claim 2, wherein n' is 1 and each of the Z1b is independently selected from (i) Formulae FL1-FL19:



10

wherein, in Formulae FL1 to FL19:

Z'' represents an attachment point toward X1;

R'' represents an attachment point toward Z1c;

p is 1, 2, 3, 4, or 5,

15

q is 1, 2, 3, 4, or 5,

r is 1, 2, 3, 4, or 5; and

any primary amine is optionally acetylated or alkylated; and

- 5 (ii) an L- or D-amino acid comprising at least one amine group directly conjugated to Z1c, wherein an acid functional group of the amino acid is conjugated toward X1 in Formula I.
7. The compound of claim 1 or 2, wherein the compound comprises a drug substance comprising a human polypeptide hormone of the human pancreas, insulin, glucagon, GLP-1, a somatostatin, a gastric inhibitory polypeptide, a glucose-dependent insulinotropic polypeptide, a hybrid peptide comprising sequences from two or more human polypeptide hormones, or an analogue thereof.
- 10 8. The compound of claim 5, wherein:
- X1 comprises human insulin or a human insulin analogue comprising an A-chain and a B-chain, wherein the A-chain comprises a sequence selected from SEQ ID NOs 1 and 3 to 33, and the B-chain comprises a sequence selected from SEQ ID NOs 2 and 34 to 74, 24047, and 24048;
- 15 each Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, FF163, FF193, FF194, FF203, and FF221-FF224 and covalently conjugated either directly, or indirectly via the linker, to Z1a and/or Z1b, or to X1;
- 20 each Z1a is independently absent or independently comprises a sequence selected from K, GK, KGSH (SEQ ID NO:24049), KGSHK (SEQ ID NO:4238), KNSTK (SEQ ID NO:5085), GKASHK (SEQ ID NO:12414), GKKEEK (SEQ ID NO:12677), GKKEHK (SEQ ID NO:12680), GKGHSK (SEQ ID NO:13120), GKGSH (SEQ ID NO:24050), GKGSHK (SEQ ID NO:13198), GKGSTK (SEQ ID NO:13205), GKHENK (SEQ ID NO:13271),
- 25 GKNSHK (SEQ ID NO:13982), GKNSTK (SEQ ID NO:13989), GKQSSK (SEQ ID NO:14380), GKYQFK (SEQ ID NO:15128), GKGSKK (SEQ ID NO:24045), GKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGSBK (SEQ ID NO:24042);
- each said linker is selected from FL1, FL3, FL4, and FL5;
- 30 each m' is independently 0 or 1;
- each n' is independently 0, 1, 2, or 3;
- each o' is independently 1, 2, 3, 4, or 5;
- each p' is 1, 2, 3, 4, or 5; and
- q' is 1, 2, 3, or 4, wherein when any of n', o', p', or q' is 2 or more, the corresponding
- 35 groups Z1a, Z1b, and Z1c are independently selected and may be the same or different; and

5 wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1.

9. The compound of claim 5 or 8, wherein:

 X1 comprises the human insulin or human insulin analogue comprising an A-chain and a B-chain, wherein the A-chain comprises SEQ ID NO:1; and the B-chain is selected from
10 SEQ ID NOs 2, 36, 24047, and 24048;

 each Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, FF193, FF194, FF203, and FF221-FF224 and covalently conjugated either directly, or indirectly via the linker, to Z1a and/or Z1b, or to X1;

 each Z1a independently comprises a sequence selected from K, GK, KGSH (SEQ ID
15 NO:24049), KGSHK (SEQ ID NO:4238), KNSTK (SEQ ID NO:5085), GKASHK (SEQ ID NO:12414), GKEEEEK (SEQ ID NO:12677), GKEEHK (SEQ ID NO:12680), GKGHSK (SEQ ID NO:13120), GKGSH (SEQ ID NO:24050), GKGSHK (SEQ ID NO:13198), GKGSTK (SEQ ID NO:13205), GKHENK (SEQ ID NO:13271), GKNSHK (SEQ ID NO:13982), GKNSTK (SEQ ID NO:13989), GKQSSK (SEQ ID NO:14380), GKYQFK (SEQ ID
20 NO:15128), GKGSKK (SEQ ID NO:24045), GKKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGSBK (SEQ ID NO:24042);

 each said linker is independently absent or independently selected from FL3 and FL5;

 each m' is independently 0 or 1;

25 each n' is independently 0 or 2;

 each o' is independently 1, 2, or 3;

 each p' is 1, 2, or 3; and

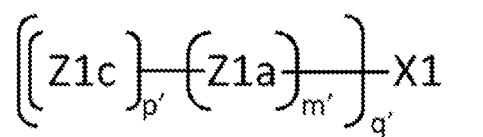
 q' is 1, 2, or 3, wherein when any of n', o', p', or q' is 2 or more, the corresponding groups Z1a, Z1b, and Z1c are independently selected and may be the same or different;

30 wherein each Z1c is independently covalently conjugated, directly or indirectly, to an amine of Z1a, to an amine of Z1b, or to X1.

10. The compound of any one of claims 2-9, wherein each of the Z1a is independently absent or independently comprises a sequence selected from K, GK, KGSH

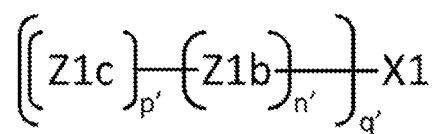
- 5 (SEQ ID NO:24049), GKGS₂H (SEQ ID NO:24050), KGSHK (SEQ ID NO:4238), and GKGS₂HK (SEQ ID NO:13198).
11. The compound of any one of claims 2-10, wherein each of the Z1c is independently selected from FF1, FF10, FF12, FF14, FF15, FF114, FF115, FF116, and FF221-FF224, and wherein the B₁ and the B₂ are independently selected from Formulae F1 and F2.
- 10 12. The compound of claim 11, wherein the B₁ and the B₂ are F2.
13. The compound of claim 12, wherein at least one R₁ in B₁ or B₂ is F or CF₃.
14. The compound of any one of claims 2-4 and 6-11, wherein Z1b is independently absent, FL3, or FL5.
- 15 15. The compound of any one of claims 1-14, wherein each of the Z1c is independently selected from FF10, FF12, FF116, FF221, FF222, and FF224.
16. The compound of any one of claims 2-4 and 6-7, wherein:
each B₁ and B₂ is F2 and is covalently conjugated to Z1c using an amide linkage,
each Z1b is independently absent; FL3 wherein p is 1, 2, or 3; or FL5 wherein p is 2, 3,
or 4;
- 20 each FF is independently selected from FF10, FF12, FF116, FF134, FF163, FF193, FF203, FF221, FF222 and FF224; wherein each FF12 and FF222 has either (S,R) or (S,S) stereochemistry;
- each Z1c is conjugated either directly or indirectly through FL3 or FL5 to the amine group in one or more lysine side chain in X1 or the N-terminus in X1; and
- 25 X1 is a polypeptide drug substance and/or an insulin optionally having from 0 to 4 residues replaced, inserted, or mutated to lysines, and wherein the lysines are each conjugated directly or indirectly to a Z1c.
17. The compound of any one of claims 1-2, 8 and 9, wherein Z1c is FF224, n' is 0, and Z1a is an amine containing amino acid.
- 30 18. The compound of any one of claims 2-17, wherein the compound is selected from:

5



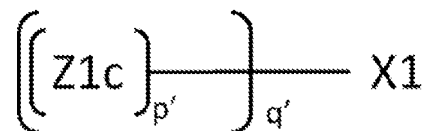
;

IA



; and

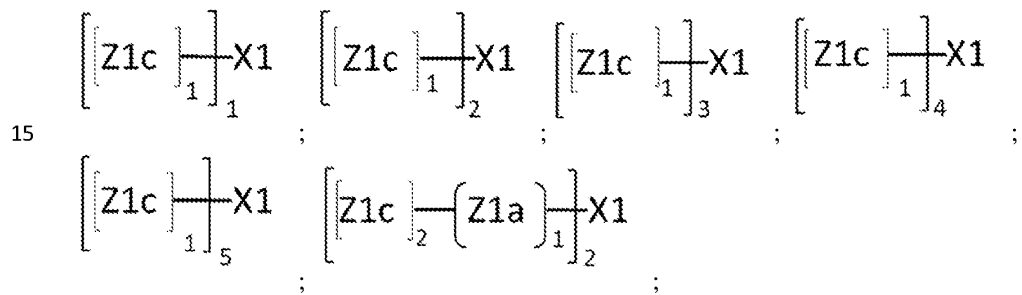
IB

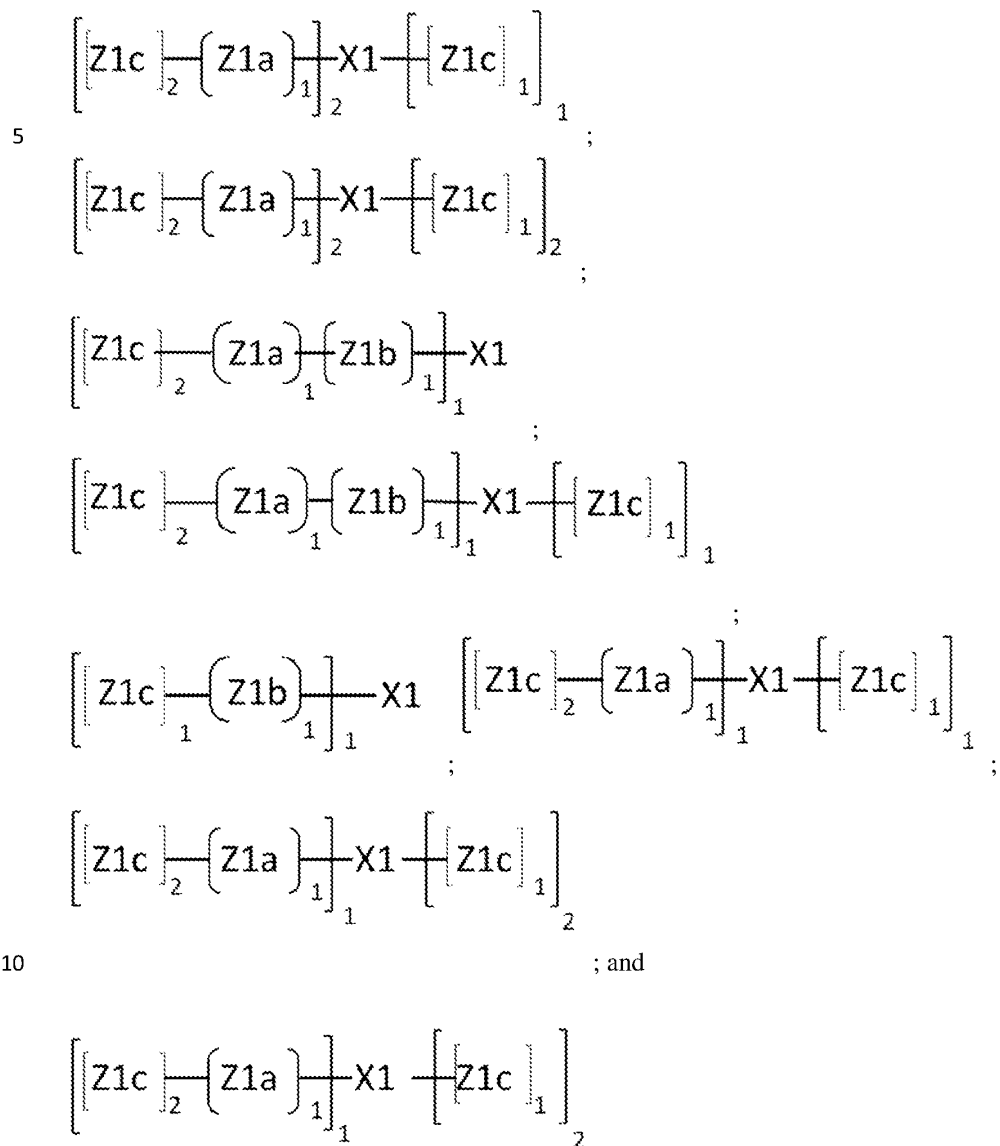


10

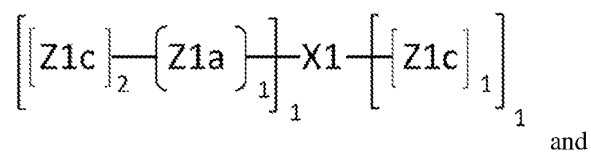
IC

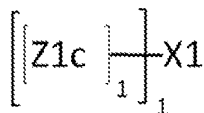
19. The compound of any one of claims 2-18, wherein the compound is selected from:





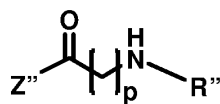
20. The compound of claim 19, wherein the compound is selected from





5

21. The compound of any one of claims 1-2 and 8-9, wherein Z1c is covalently conjugated directly to X1 via a linker, and wherein the linker is independently selected from gamma-glutamic acid, beta-alanine, and

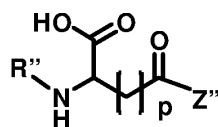


(FL3)

10

Formula FL3,

wherein p is 1, 2, or 3; and



(FL5)

15

Formula FL5,

wherein p is 2, 3, or 4.

22. The compound of any one of claims 1-21, wherein X1 is OH or NH₂, and the compound further comprises a drug substance covalently conjugated directly or indirectly to the compound.

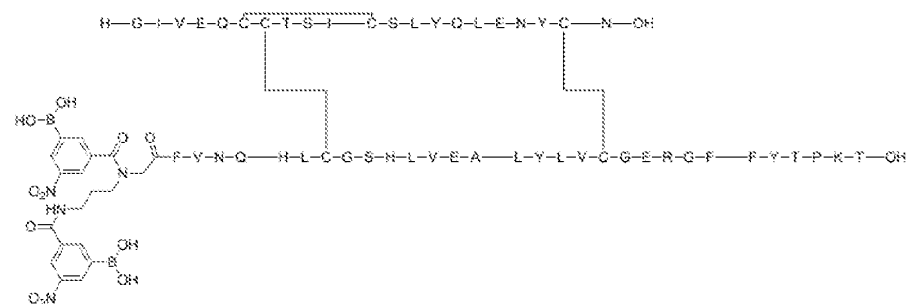
20

23. The compound of any one of claims 1-22, wherein the compound is selected from:

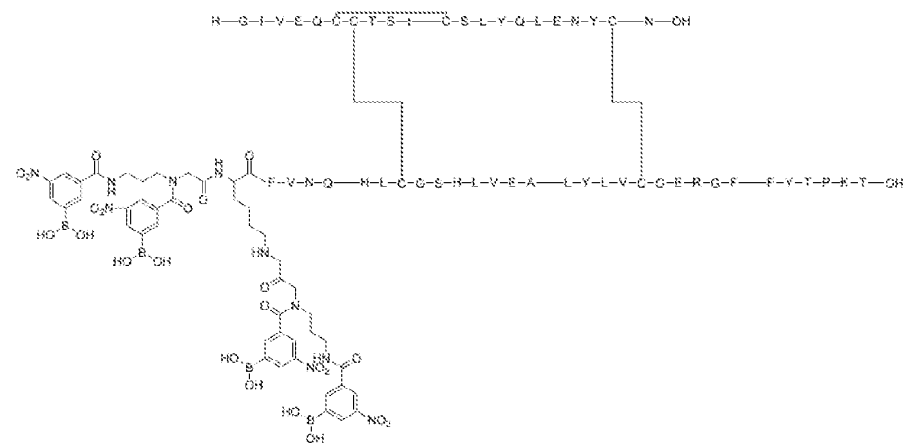
25

5

Example 315:

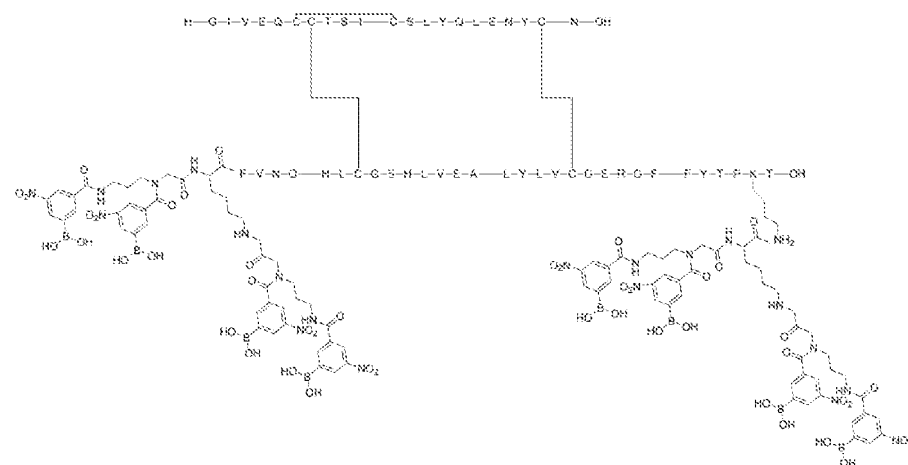


Example 318:

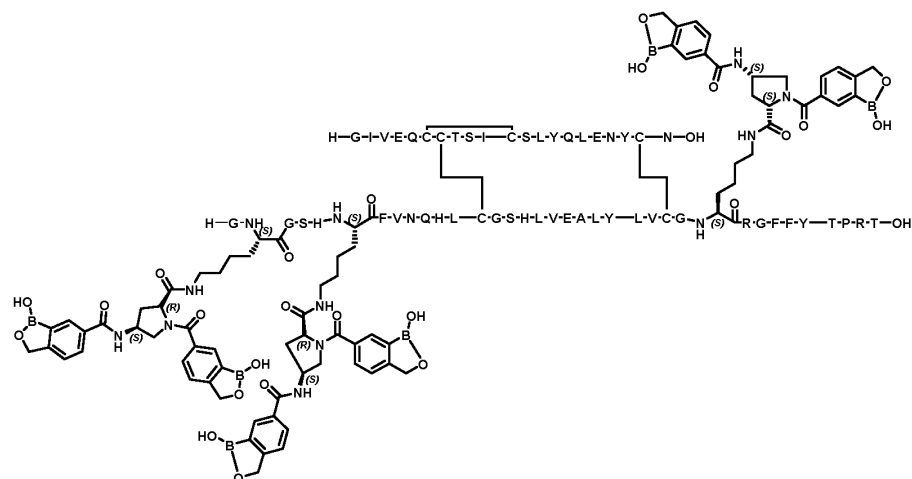


10

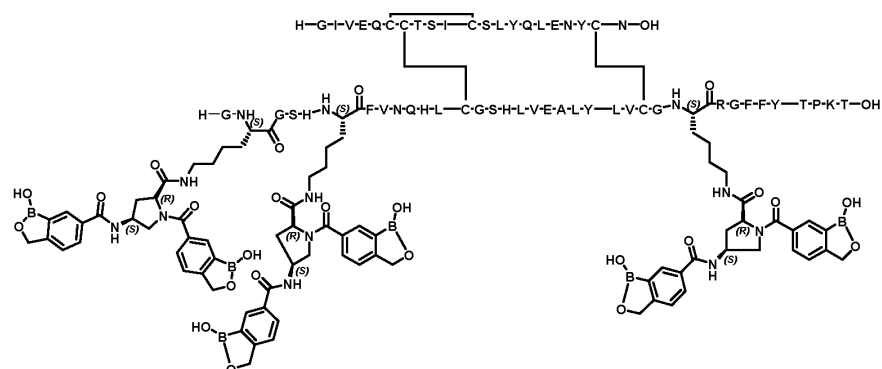
Example 320:



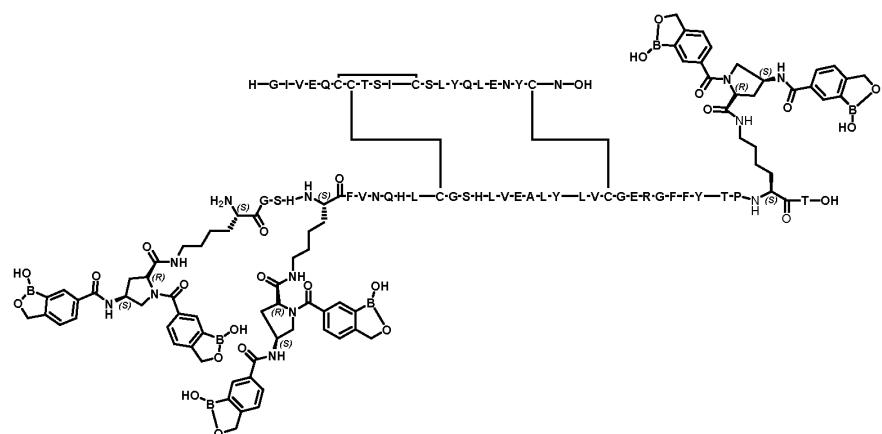
5 Example 605:



Example 606:

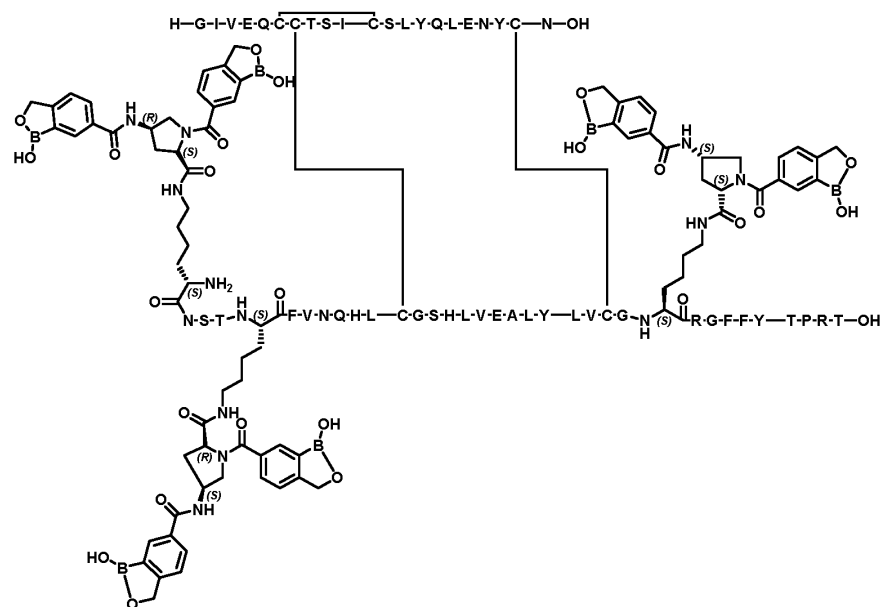


Example 607:

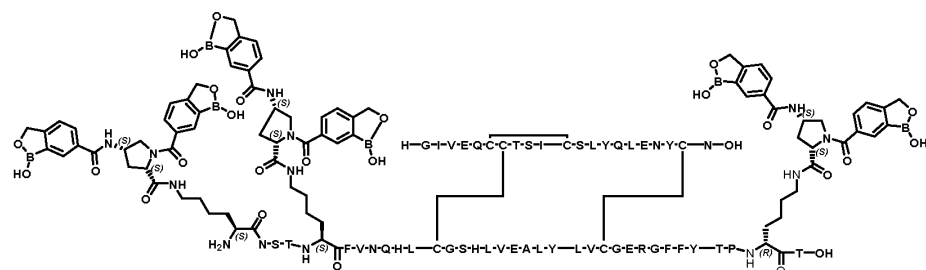


5

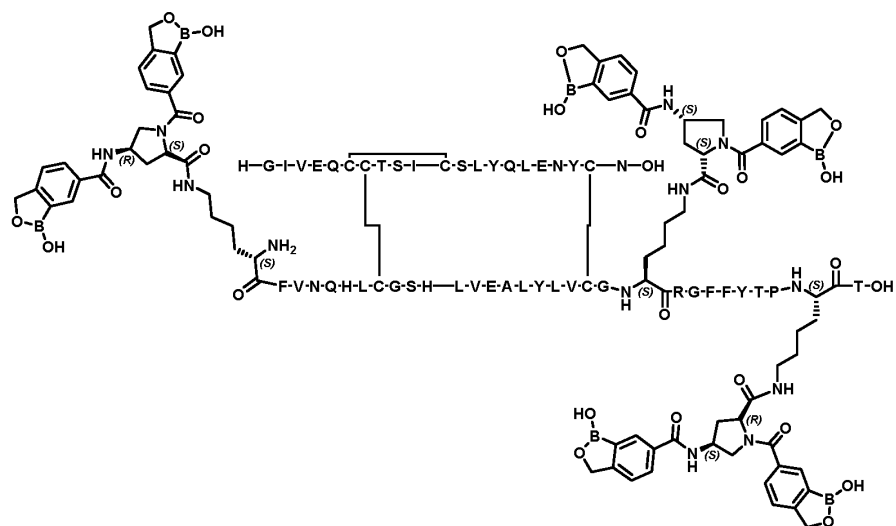
Example 608:



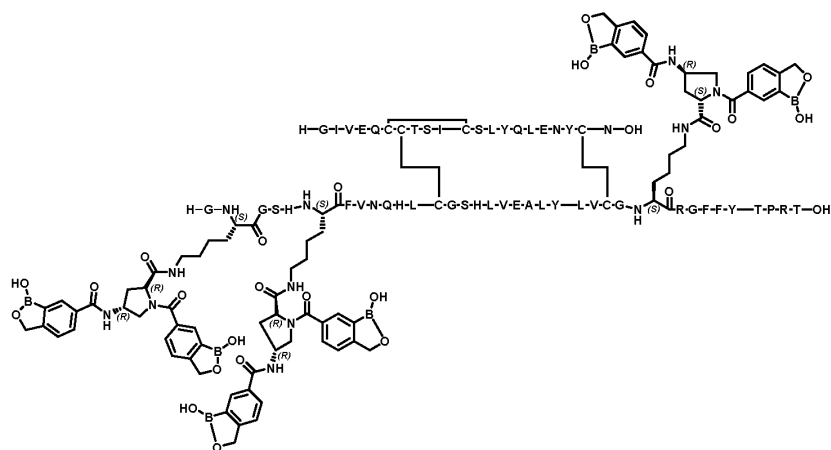
Example 610:



5 Example 611:



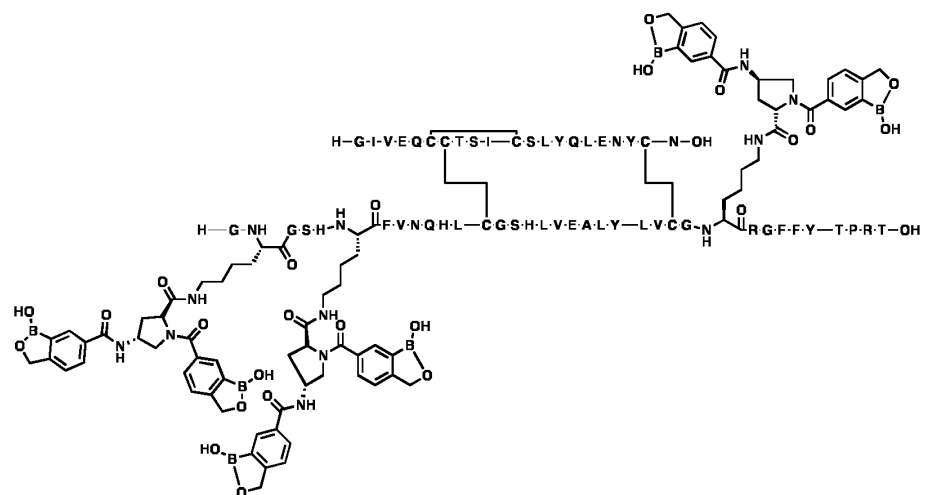
Example 612:



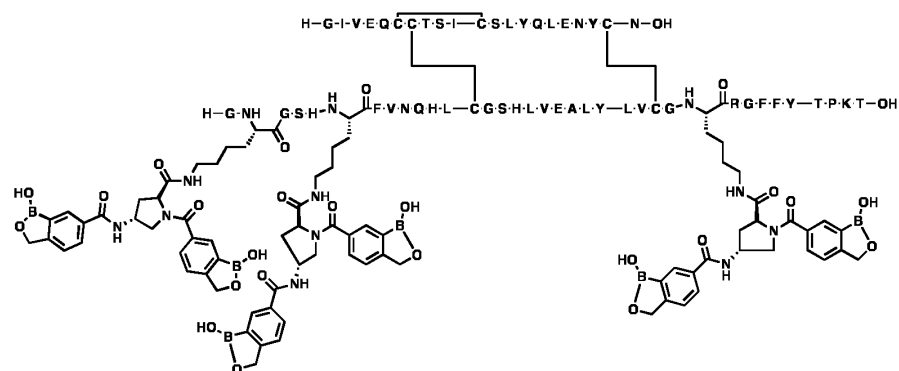
10

15

5 Example 589:



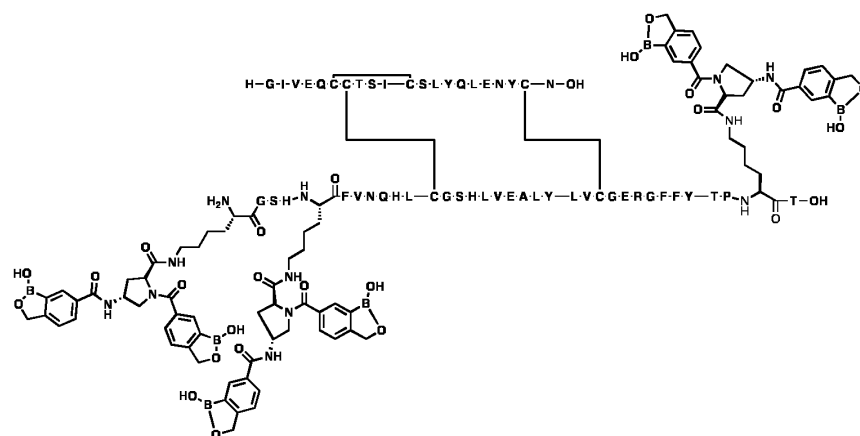
Example 590:



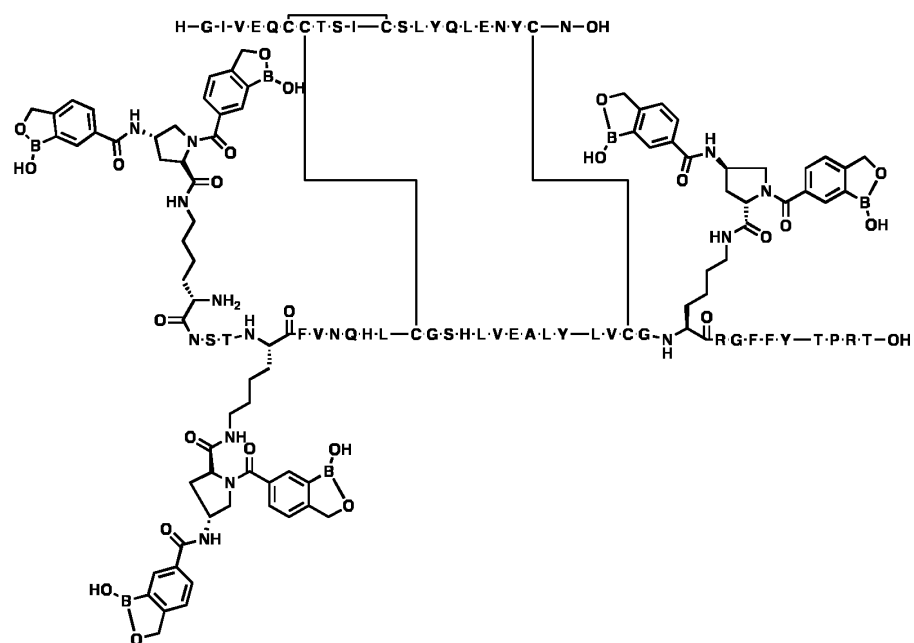
10

15

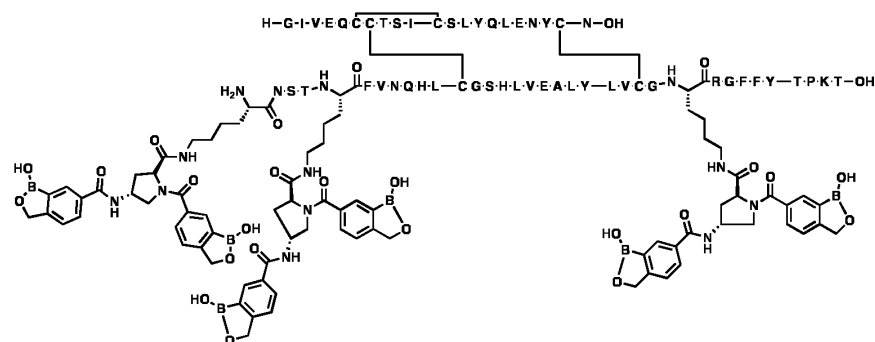
5 Example 591:



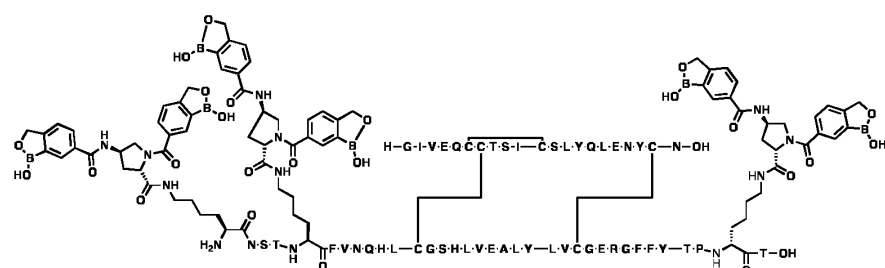
Example 592:



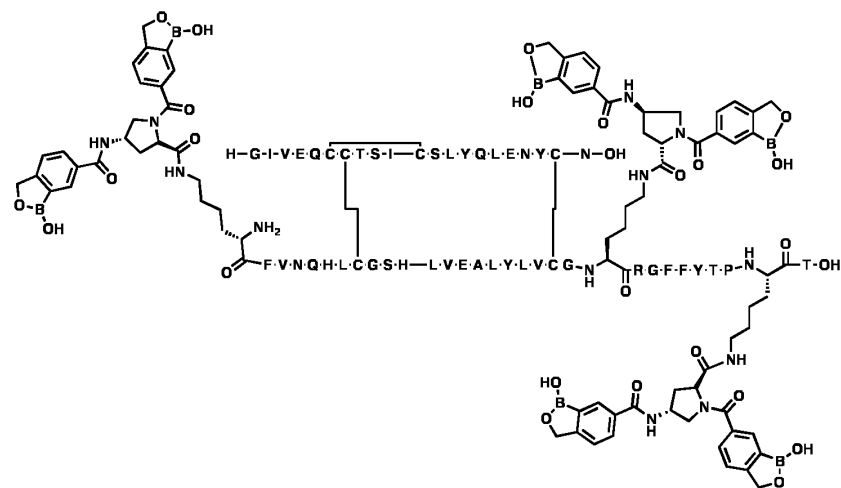
5 Example 593:



Example 594:

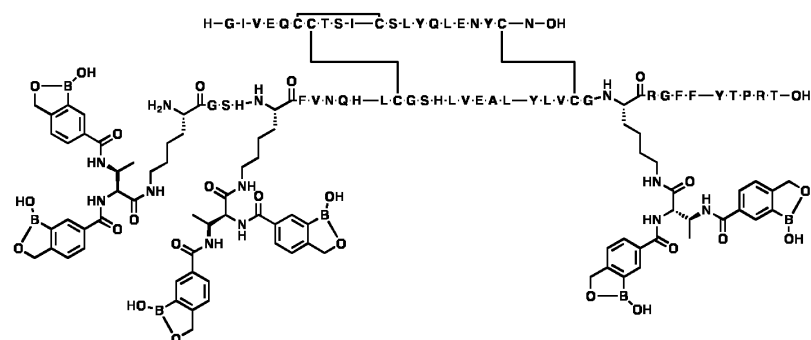


10 Example 595:

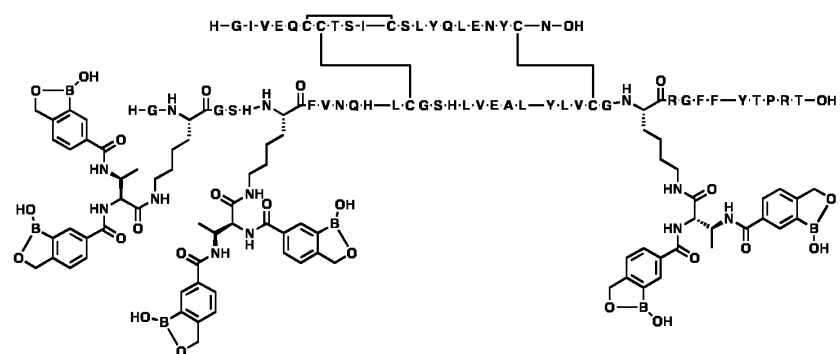


5

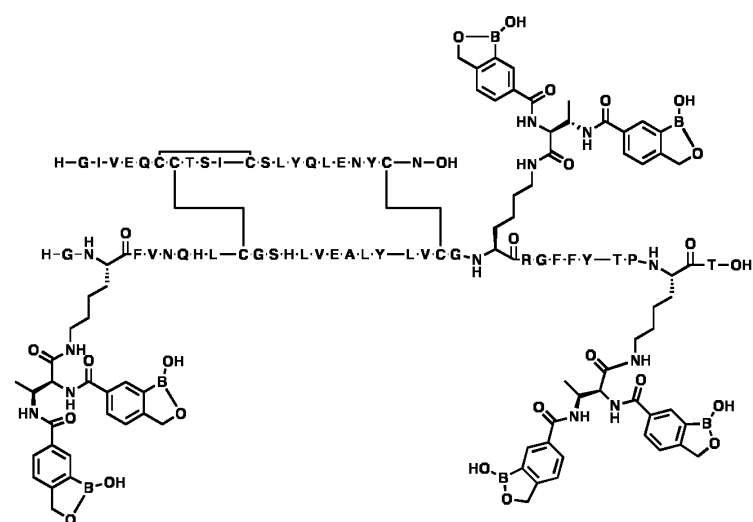
Example 562:



Example 563:

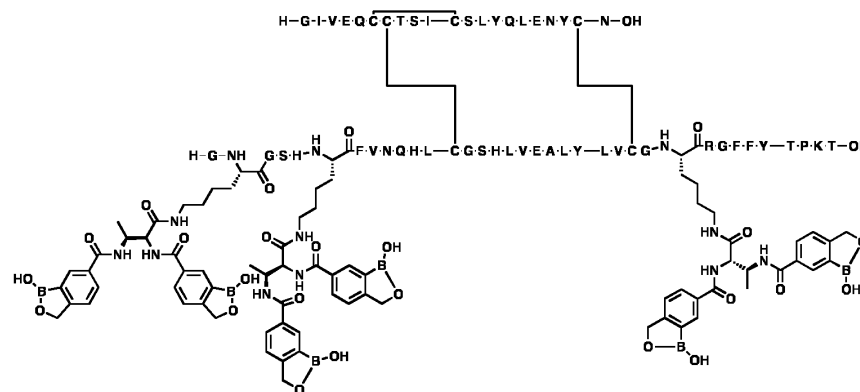


Example 564:

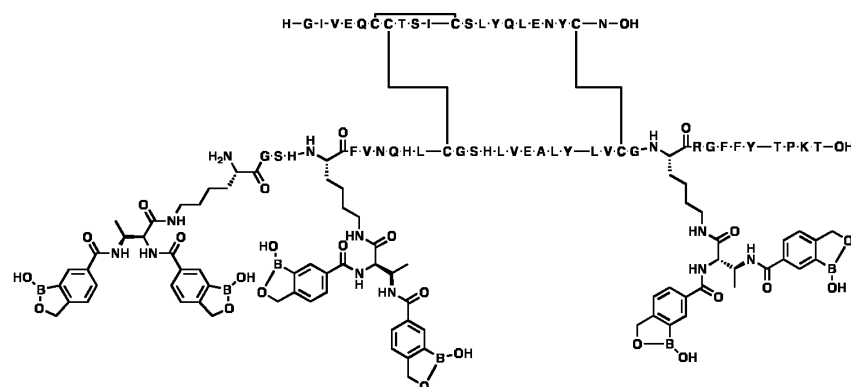


10

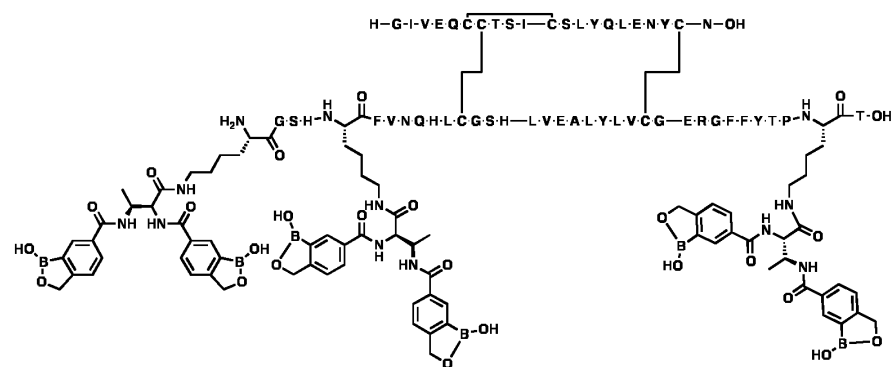
5 Example 565:



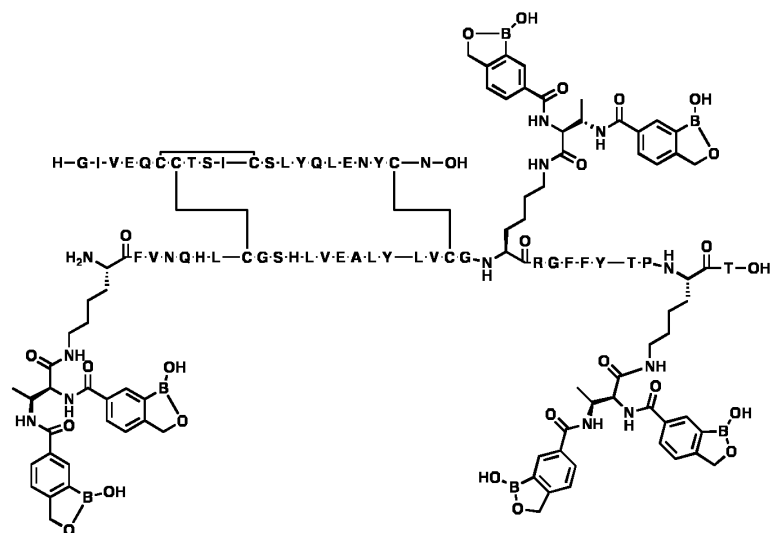
Example 566:



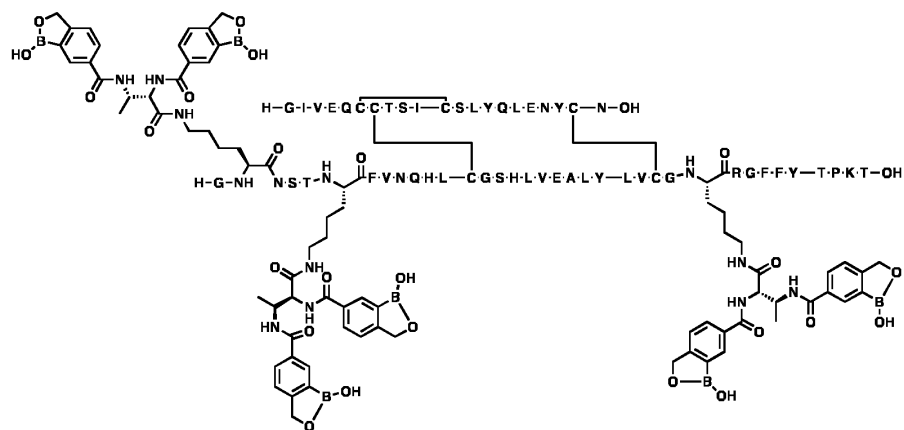
Example 567:



5 Example 568:



Example 569:

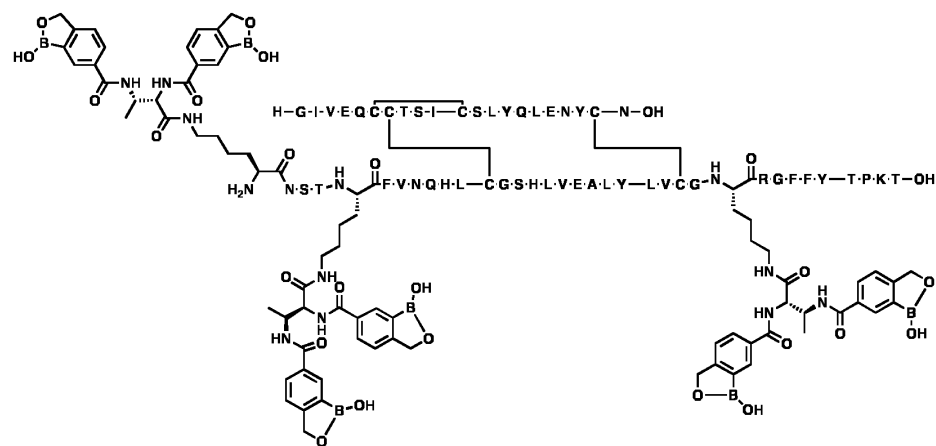


10

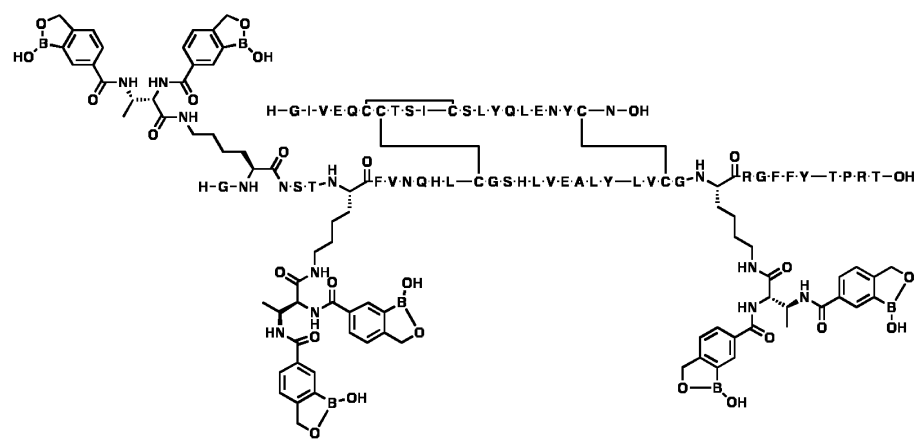
15

5

Example 570:



Example 571:

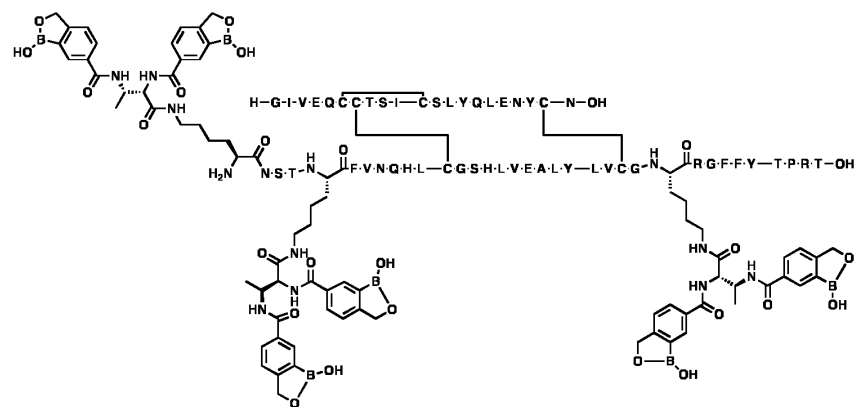


10

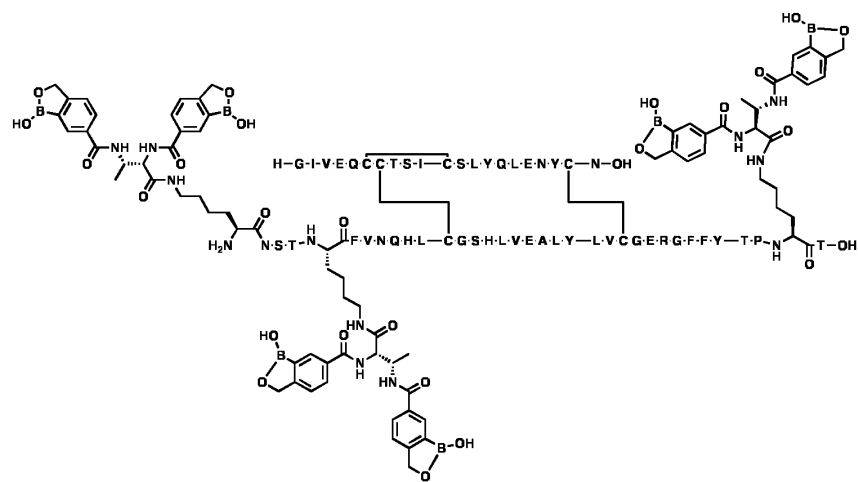
15

5

Example 572:



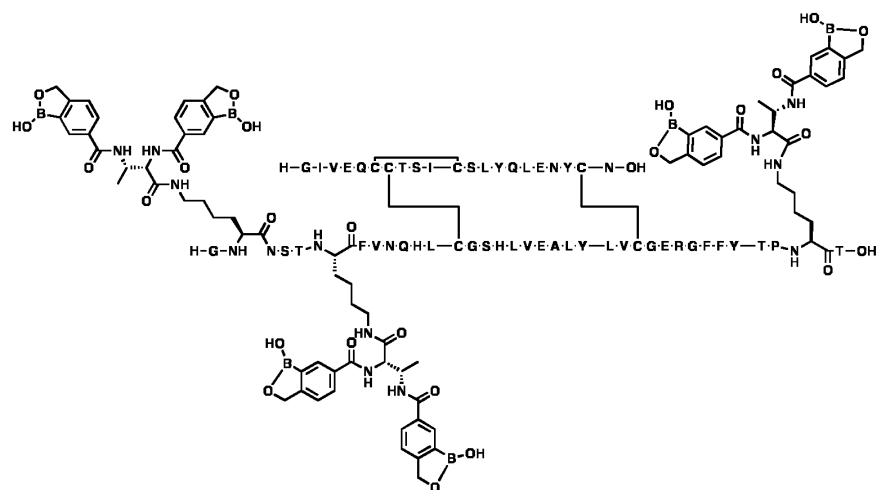
Example 573:



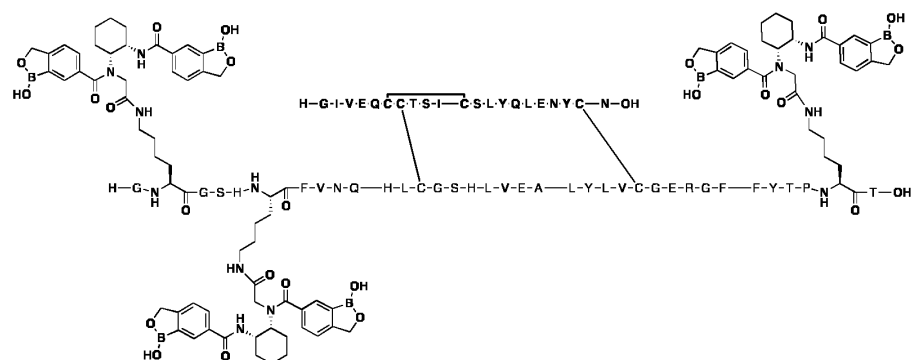
10

15

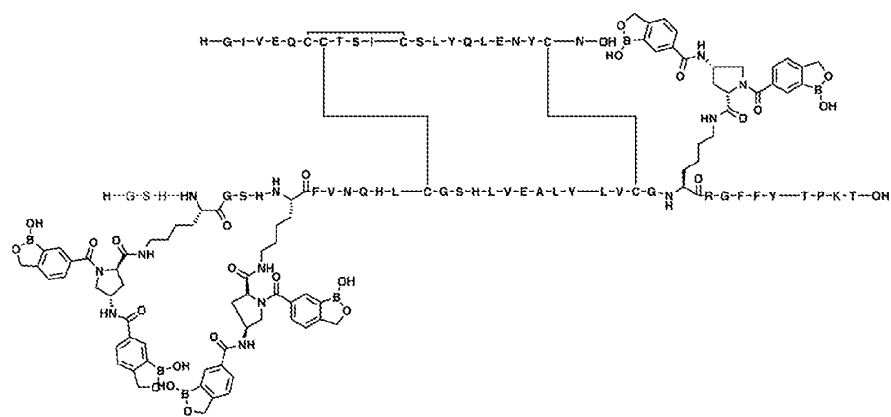
5 Example 574:



Example 803:

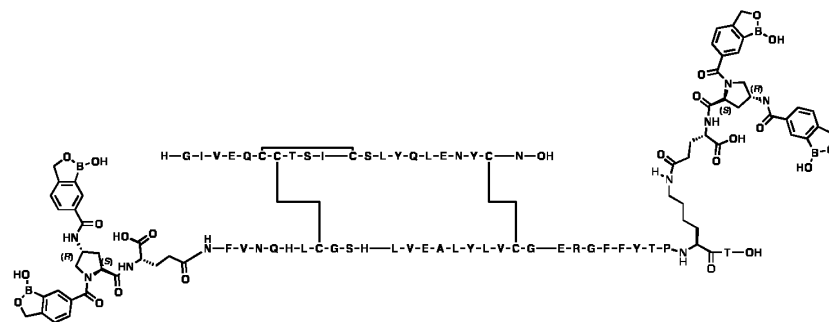


Example 804:

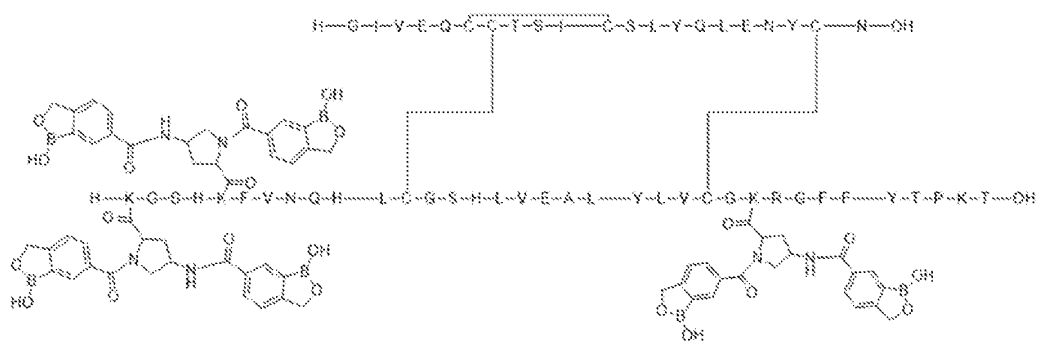


5

Example 805:

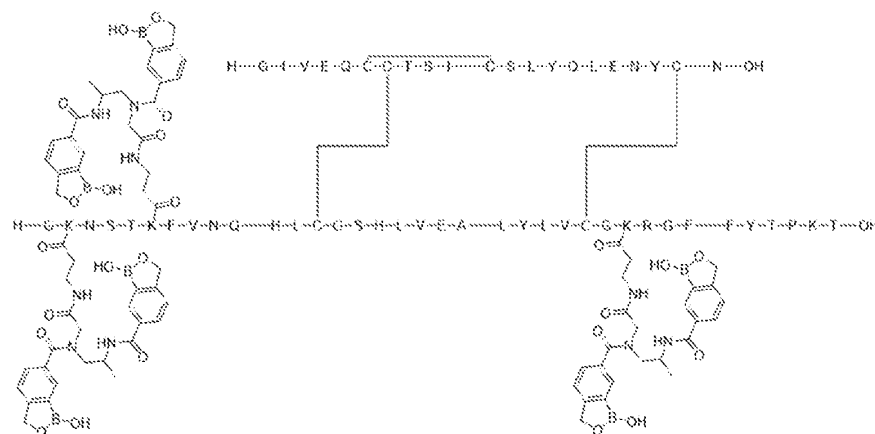


Example 806:



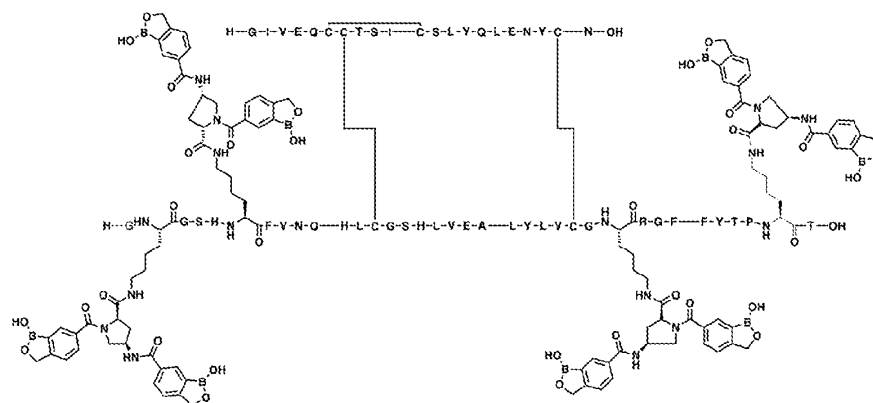
10

Example 807:

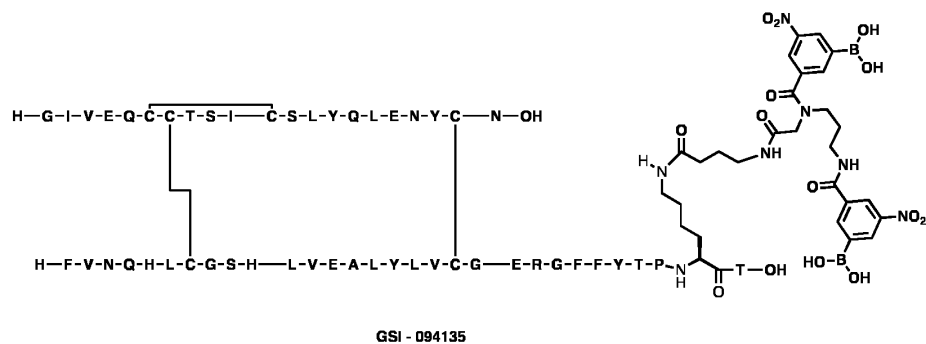


5

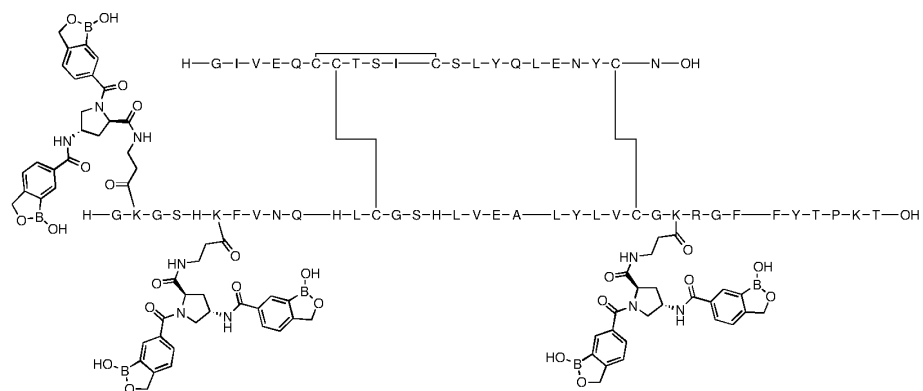
Example 808:



Example 809:

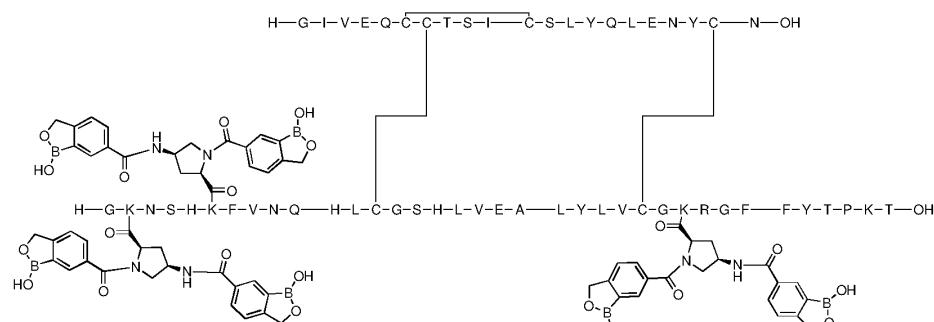


Example 810:

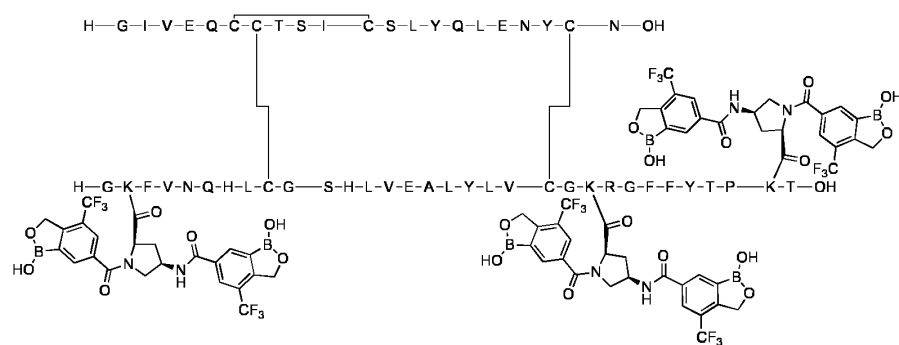


10

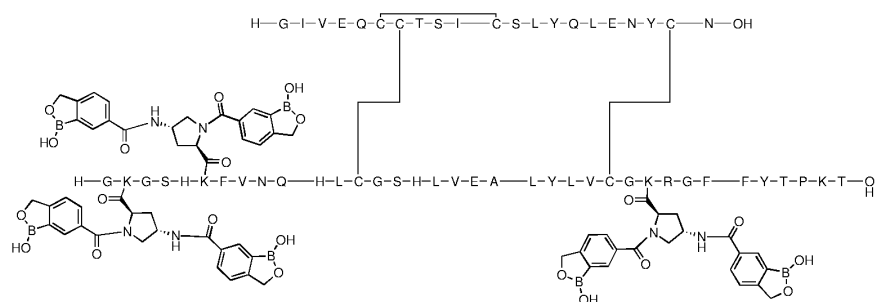
5 Example 811:



Example 812:

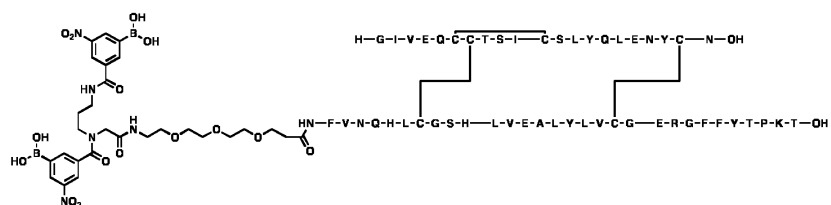


Example 813:



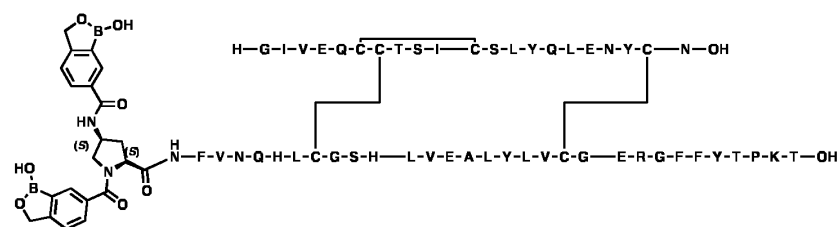
10

Example 814:

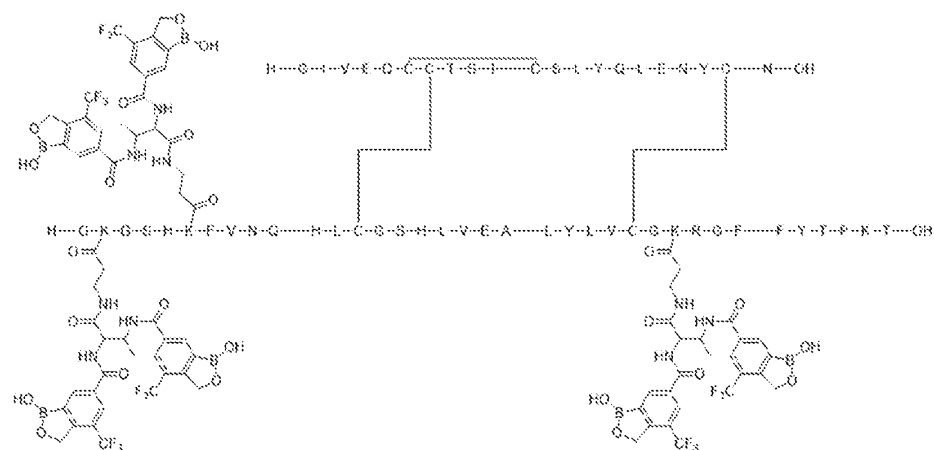


5

Example 815:

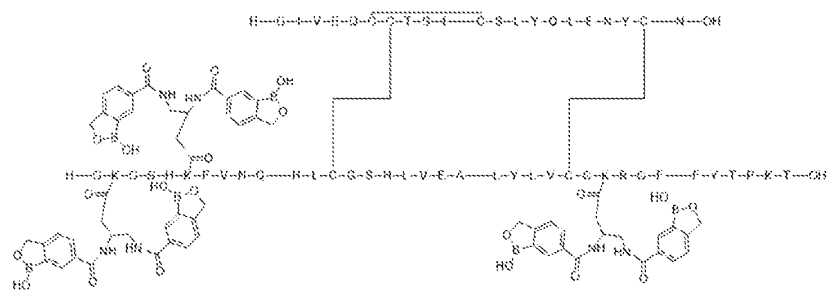


Example 816:



10

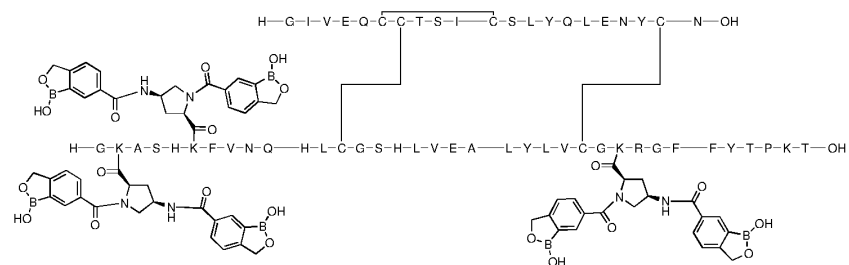
Example 817:



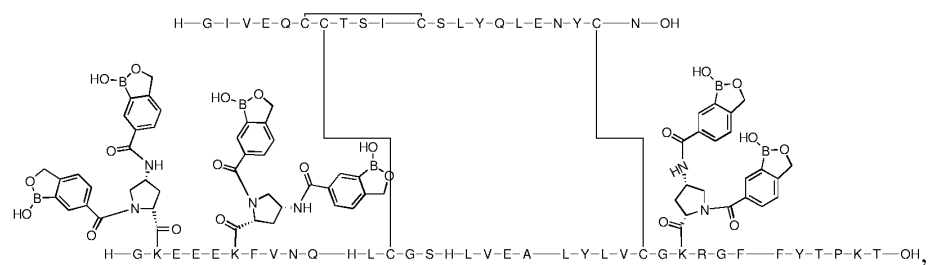
15

5

Example 818:

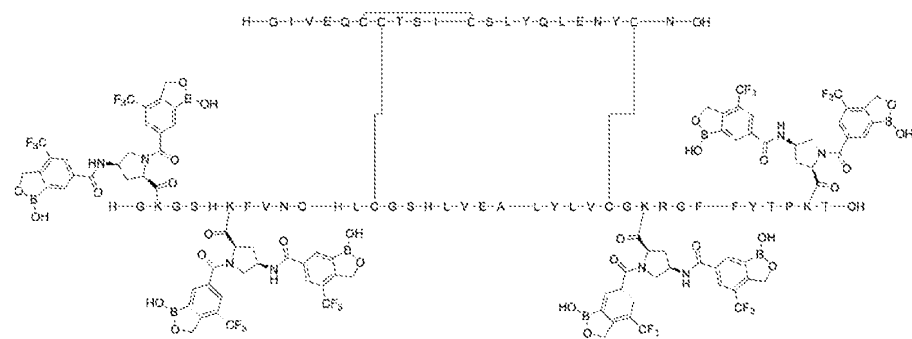


Example 819:

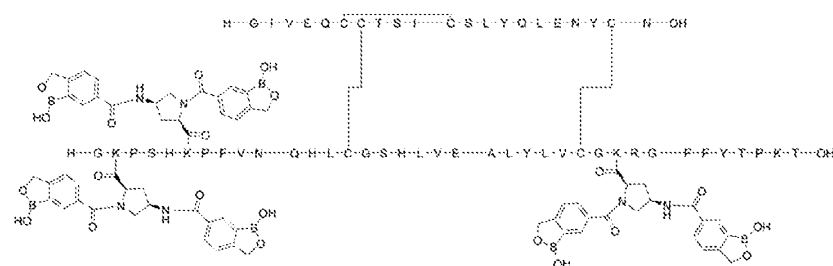


10

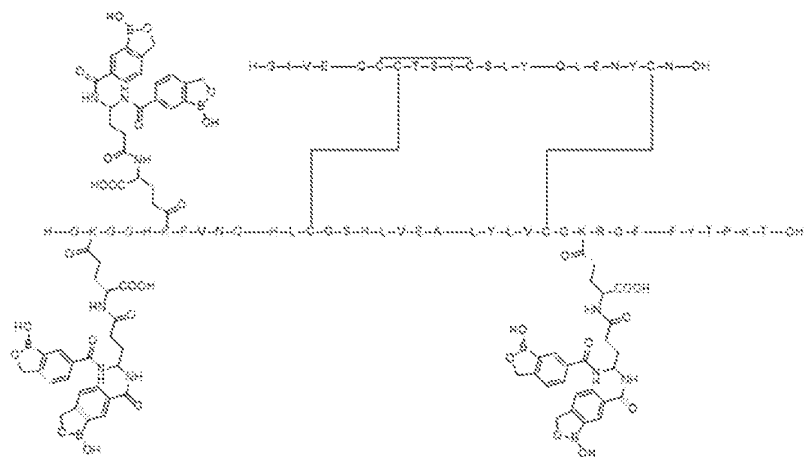
Example 820:



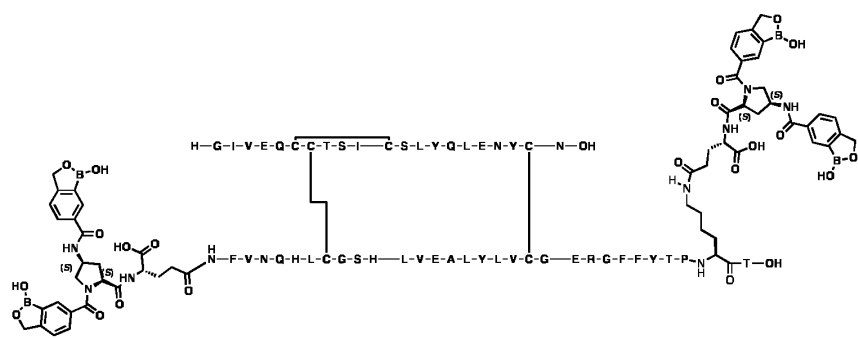
Example 821:



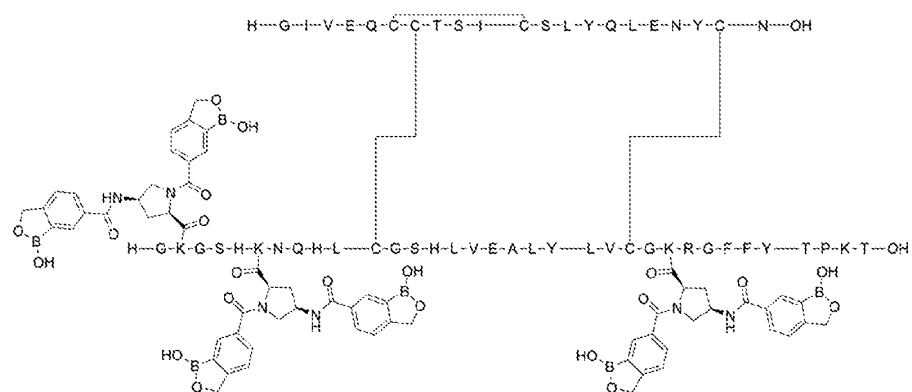
5 Example 822:



Example 823:

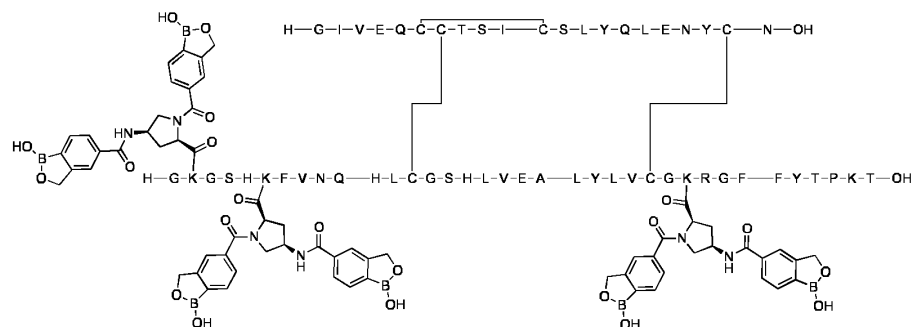


10 Example 824:

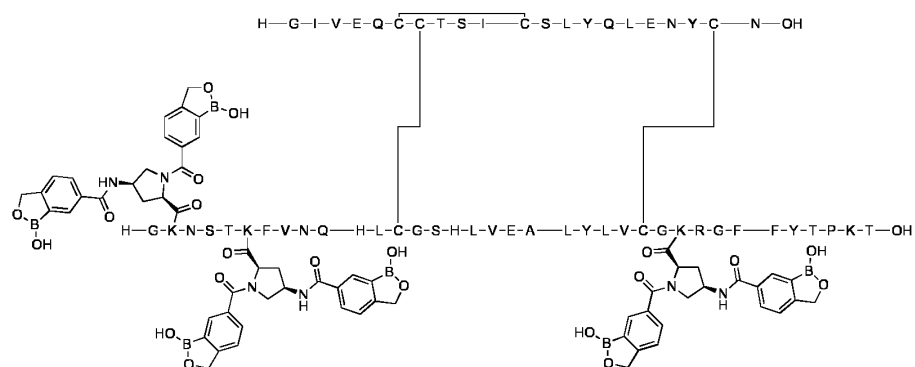


5

Example 825:

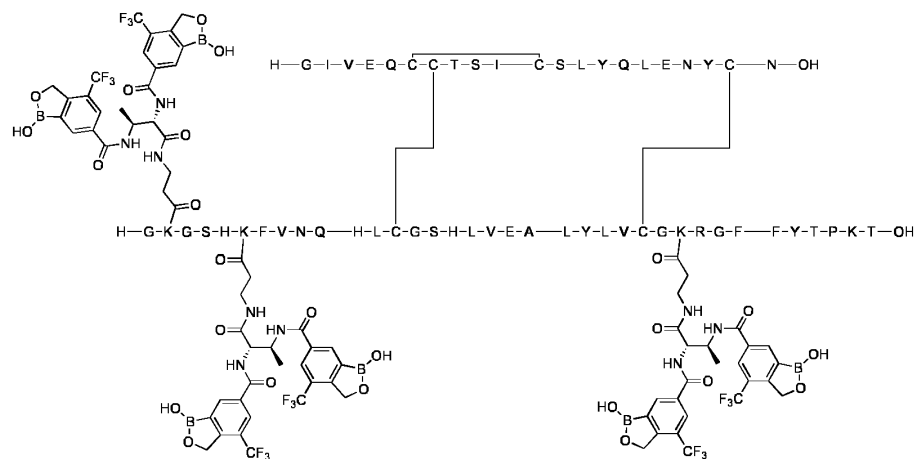


Example 826:



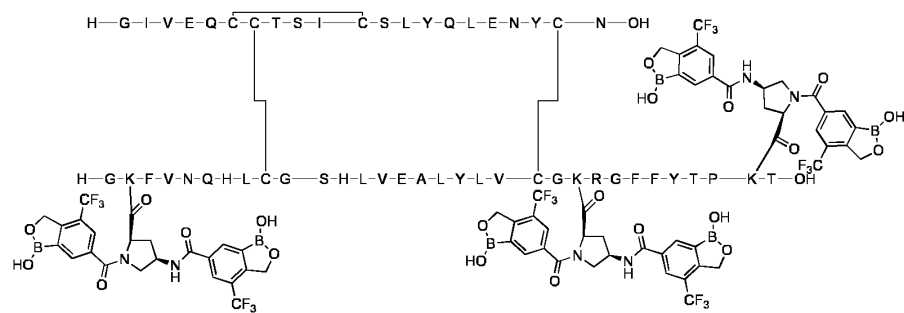
10

Example 827:

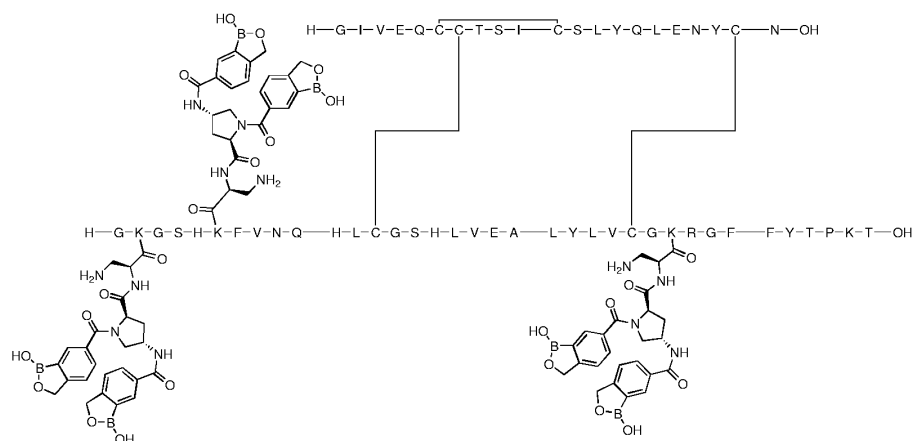


5

Example 828:

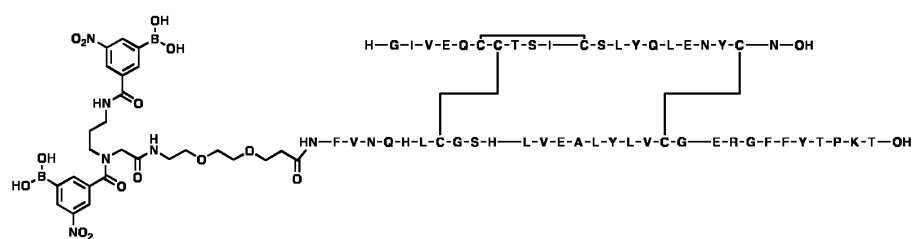


Example 829:

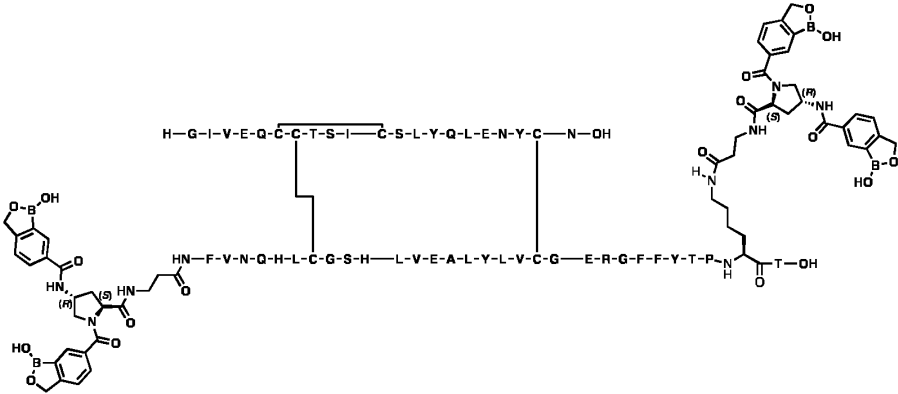


10

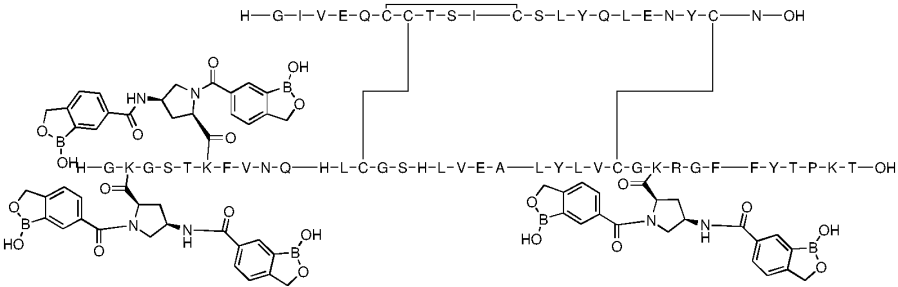
Example 830:



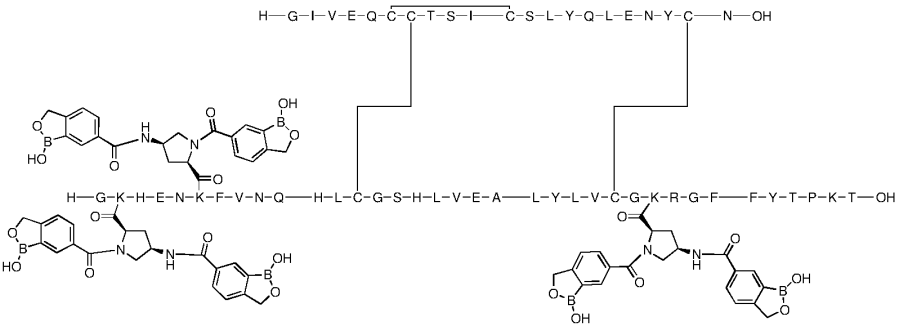
5 Example 831:



Example 832:



Example 833:

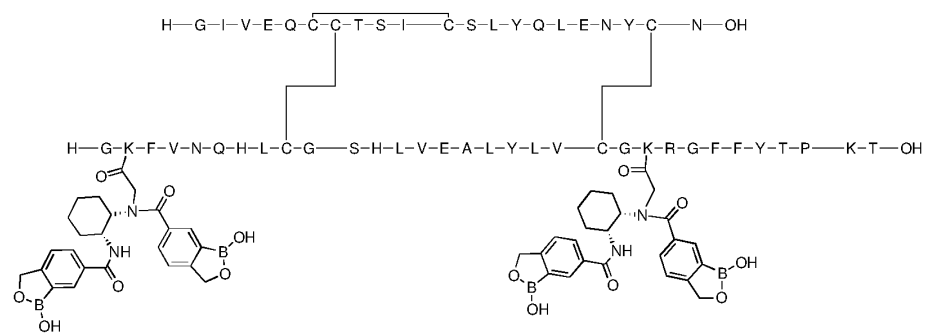


10

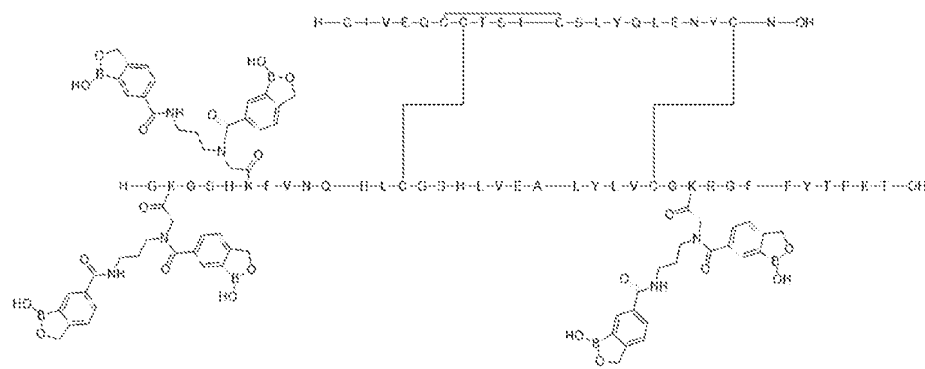
15

5

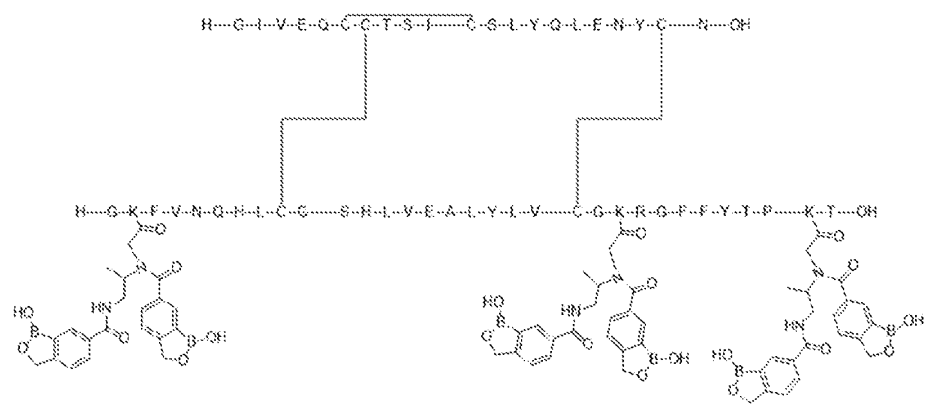
Example 834:



Example 835:

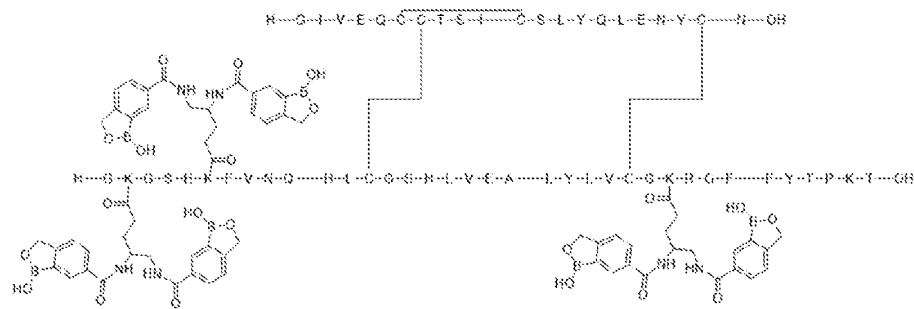


Example 836:

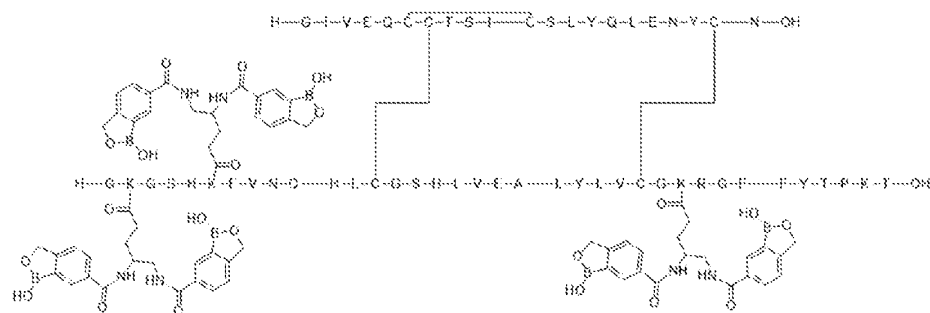


10

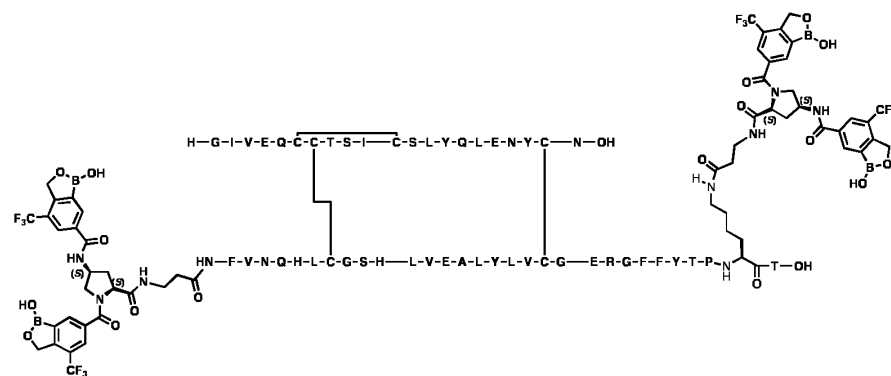
5 Example 837:



Example 838:



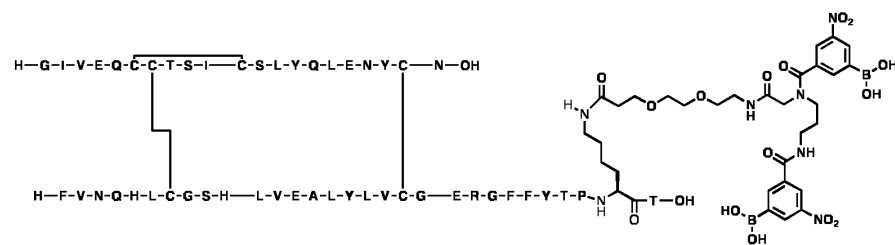
Example 839:



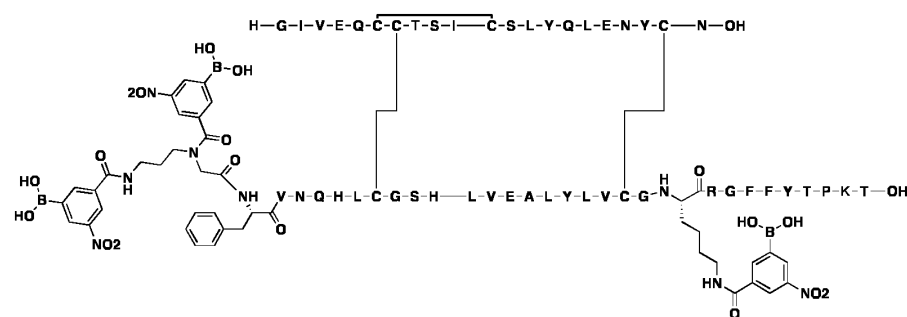
10

15

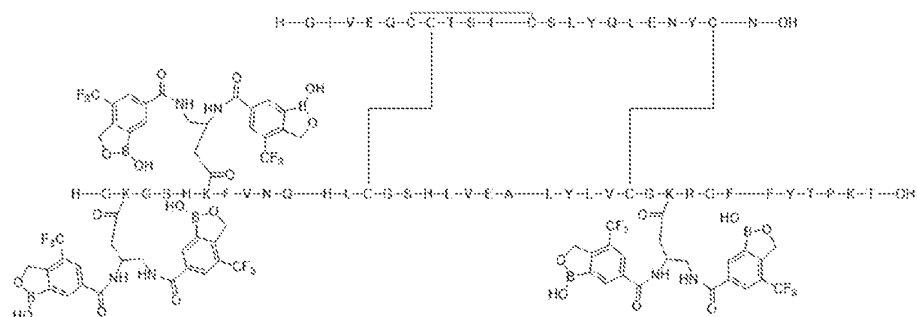
5 Example 840:



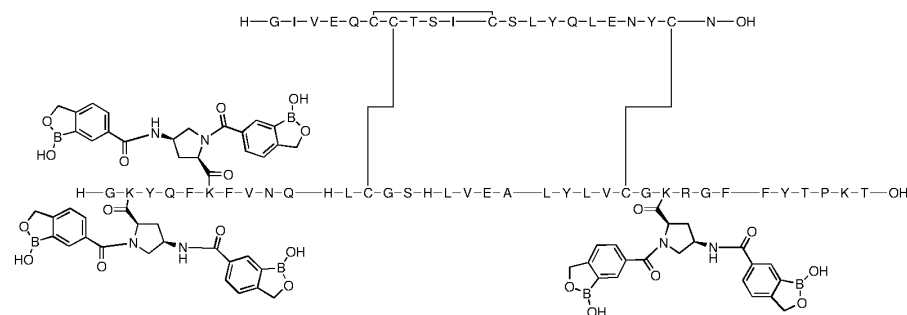
Example 841:



Example 842:

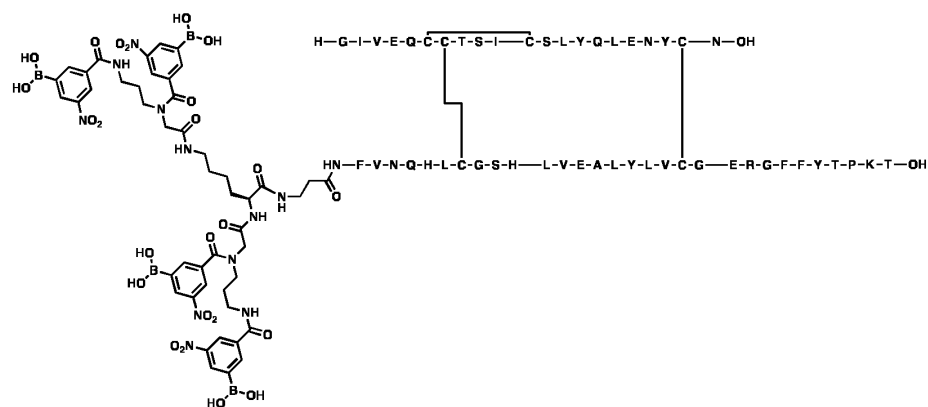


Example 843:

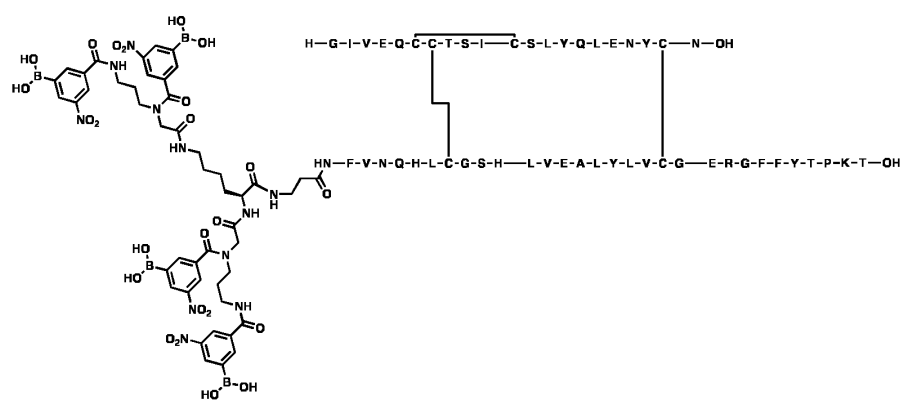


5

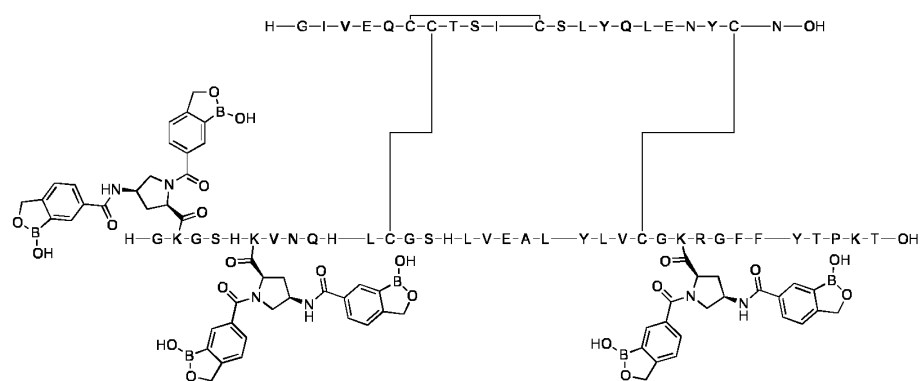
Example 847:



Example 848:



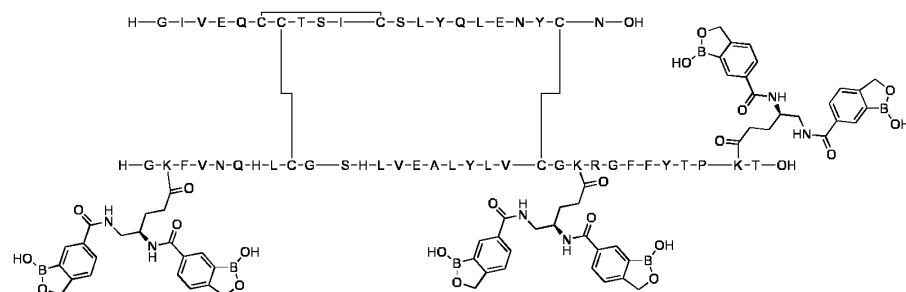
Example 849:



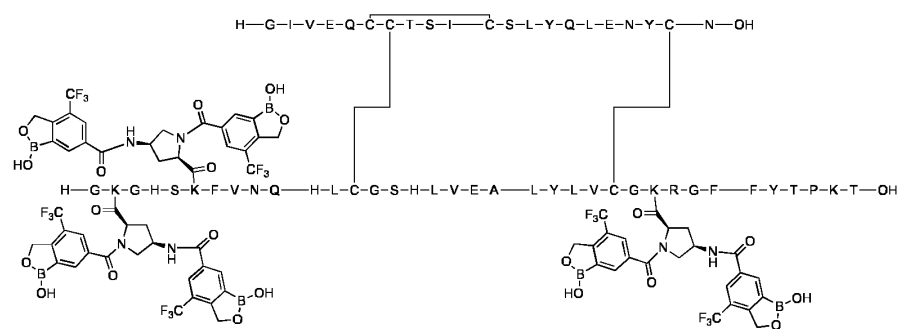
10

5

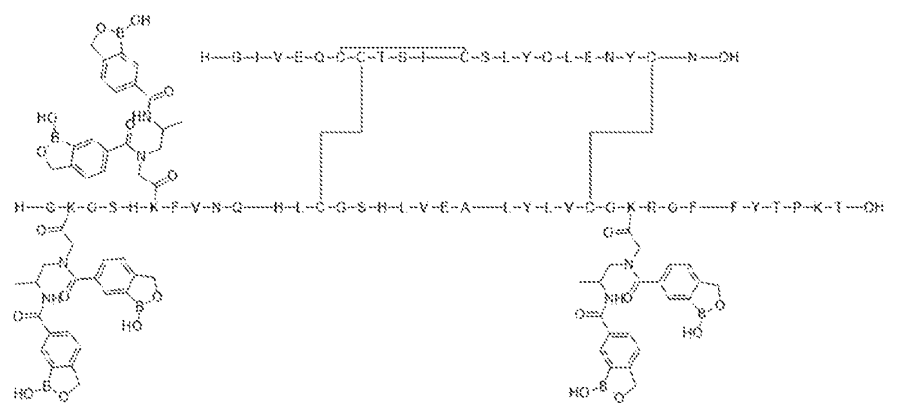
Example 850:



Example 851:



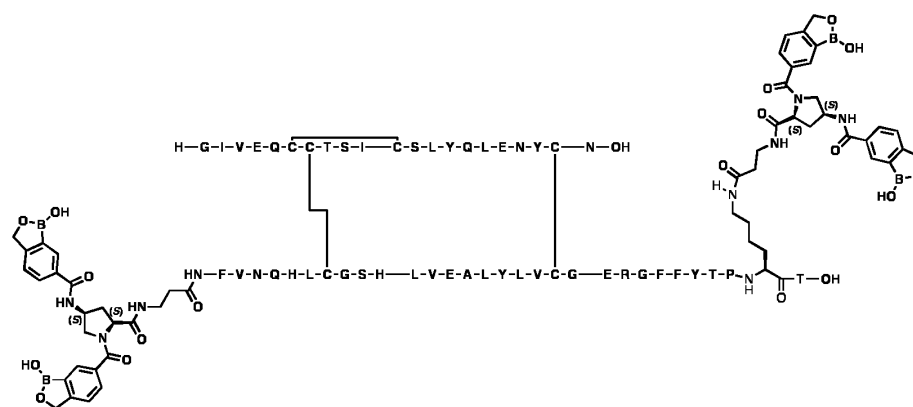
Example 852:



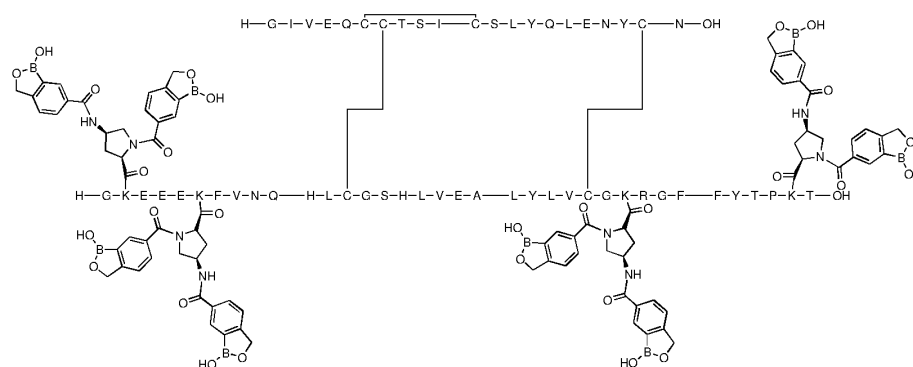
10

15

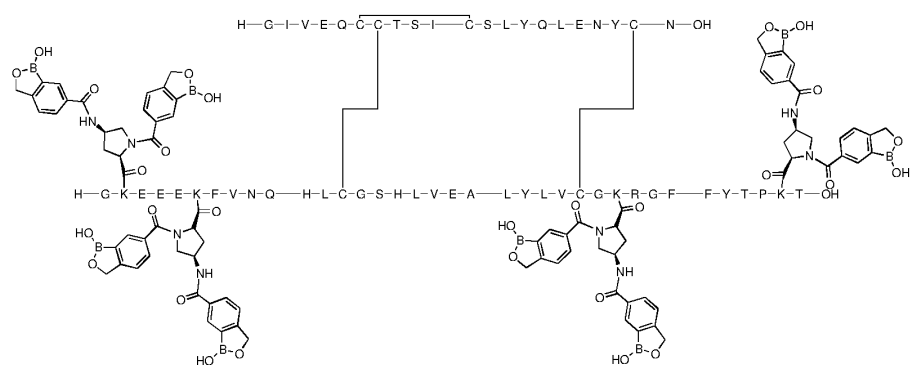
5 Example 853:



Example 854:

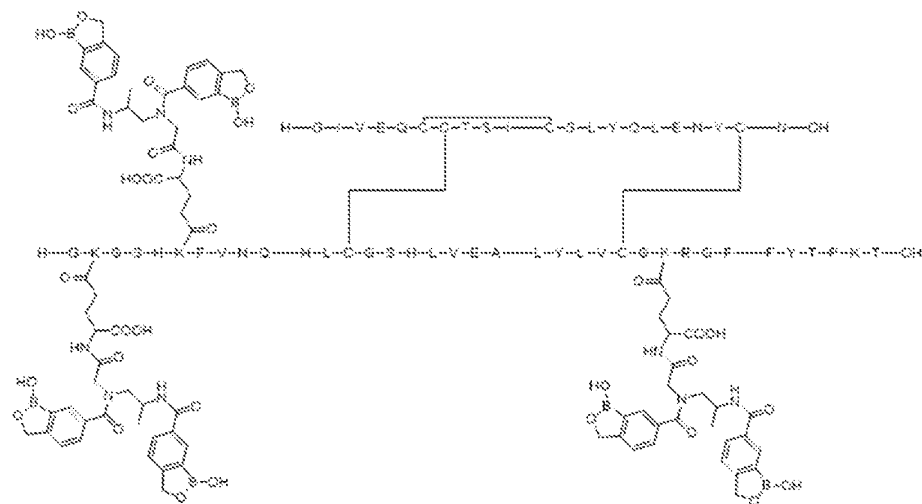


Example 855:

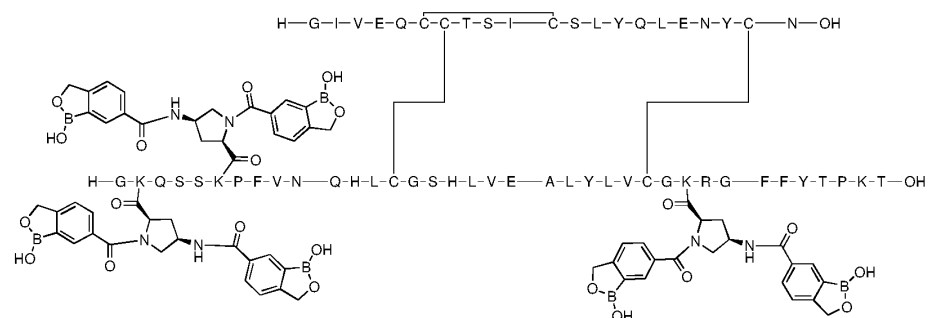


5

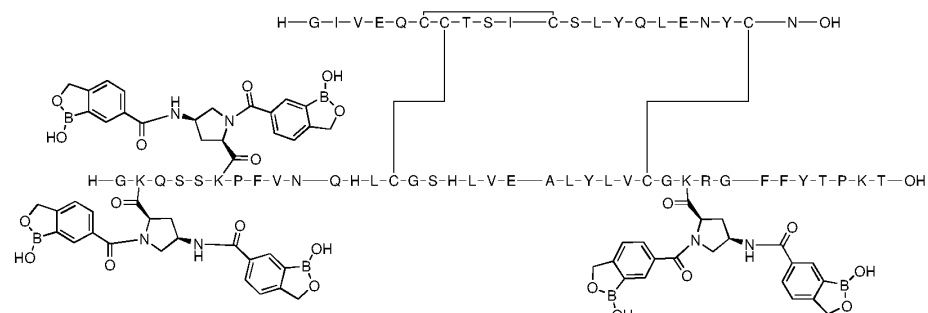
Example 856:



Example 857:



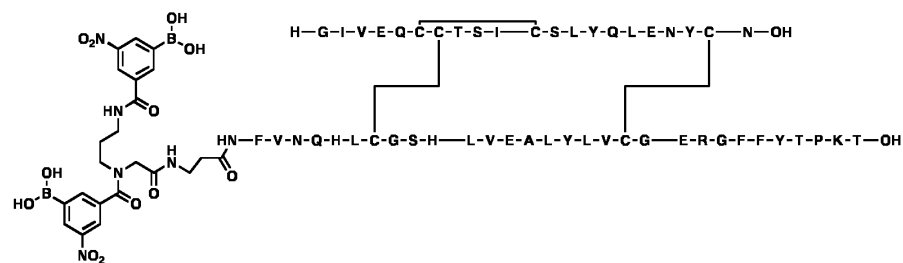
Example 858:



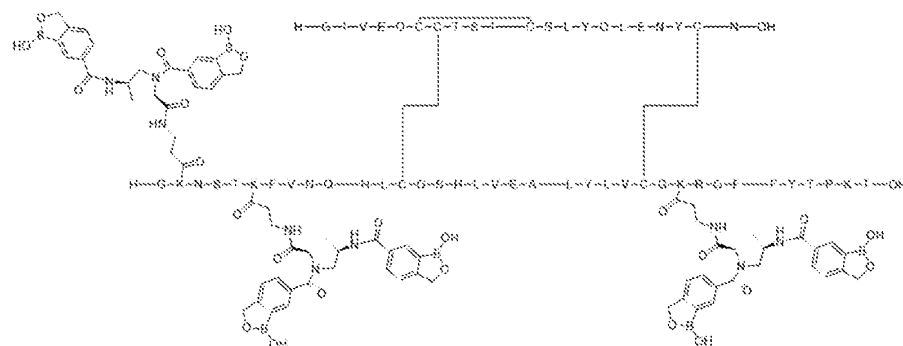
10

5

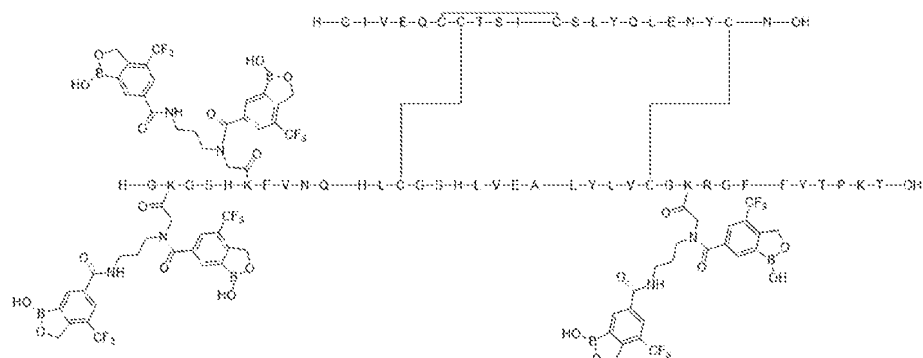
Example 859:



Example 860:



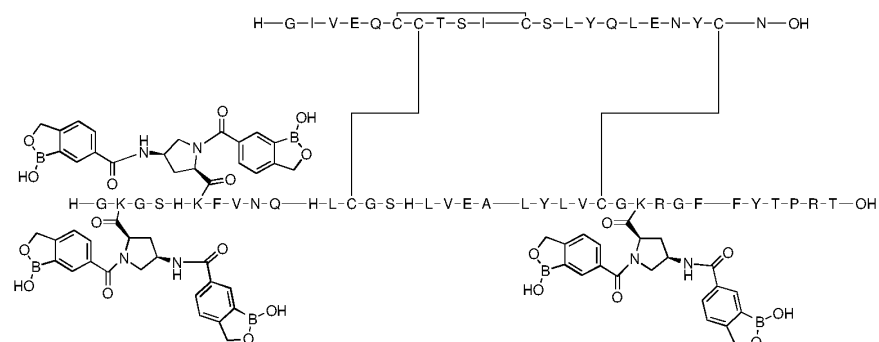
Example 861:



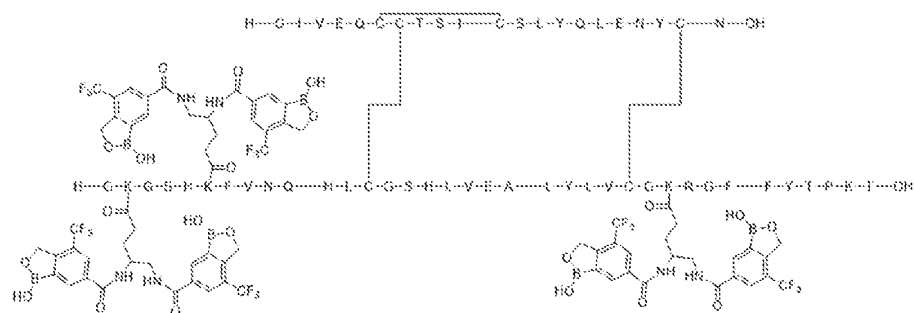
10

5

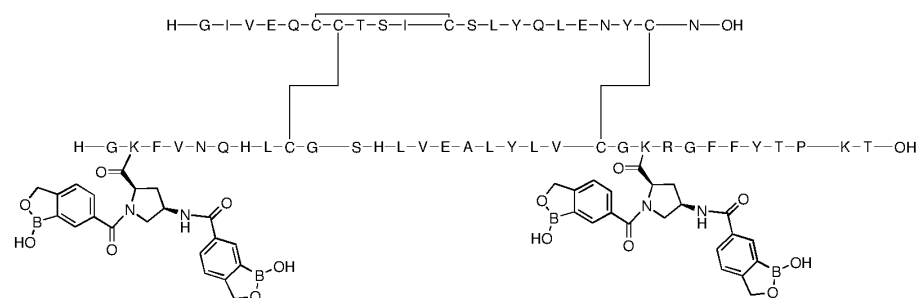
Example 862:



Example 863:

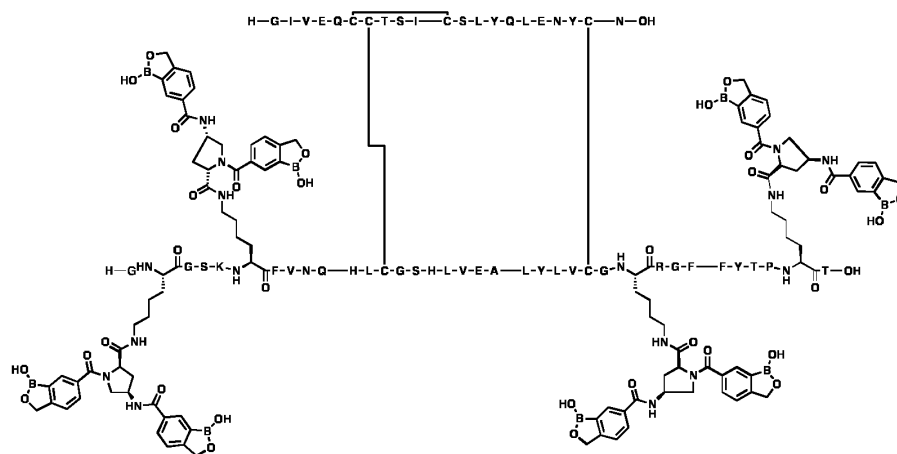


Example 864:

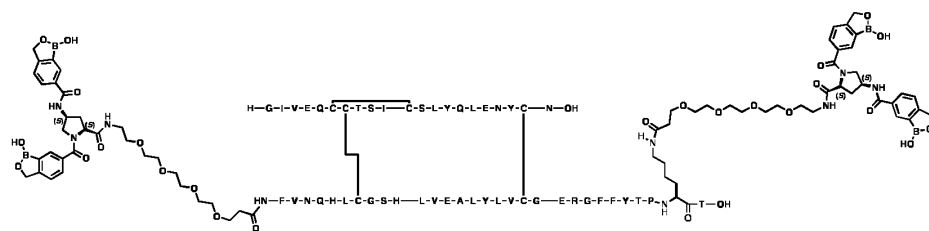


10

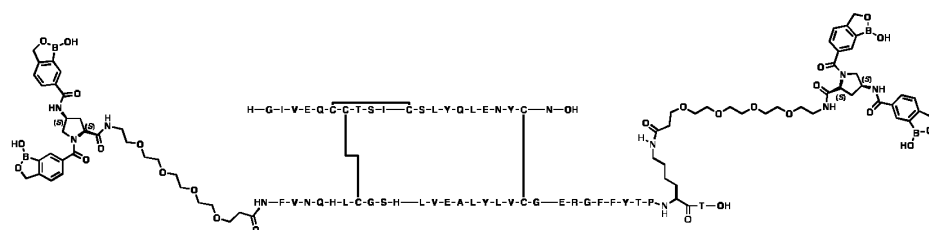
5 Example 865:



Example 866:

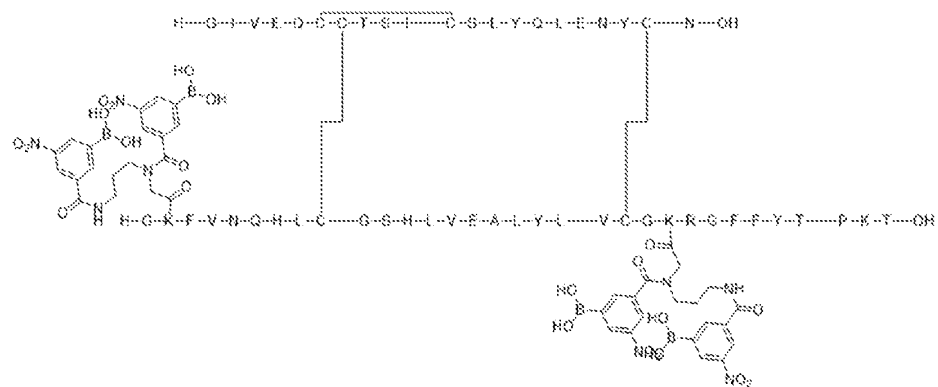


Example 867:

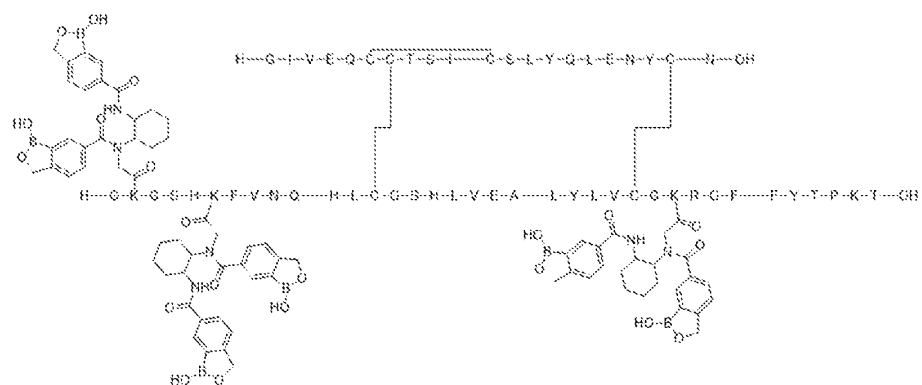


5

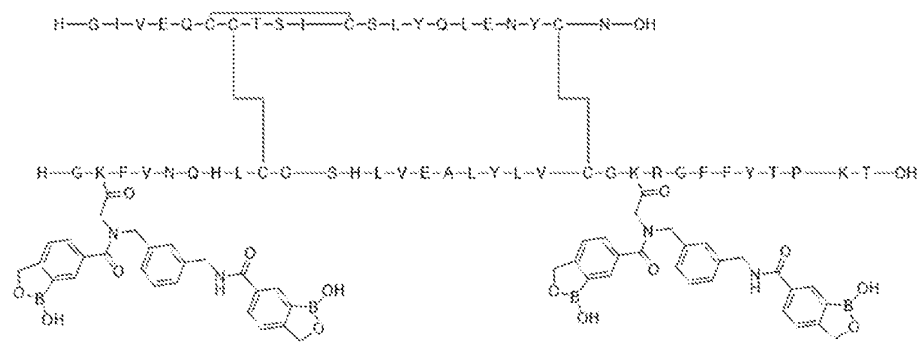
Example 868:



Example 869:

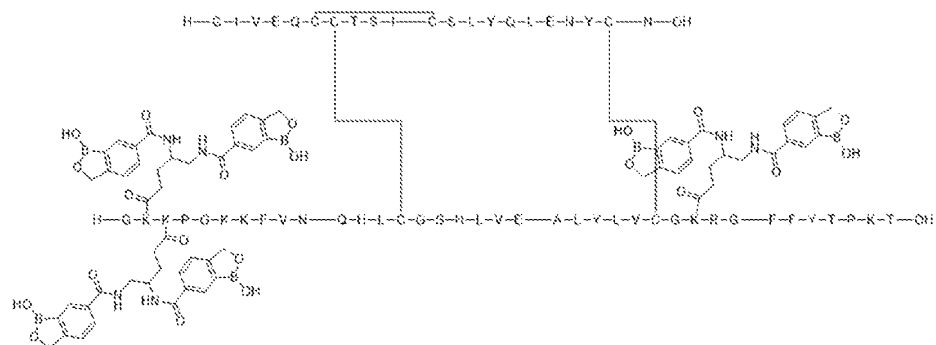


Example 870:

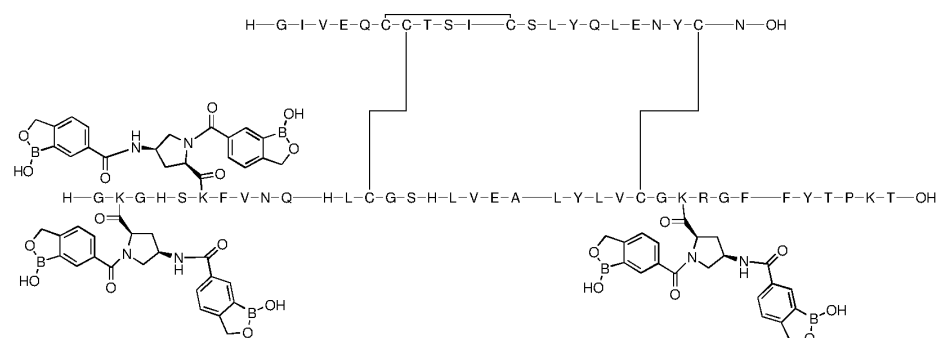


10

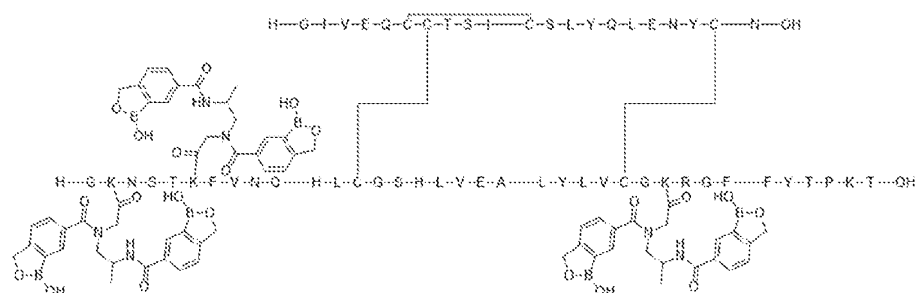
5 Example 871:



Example 872:

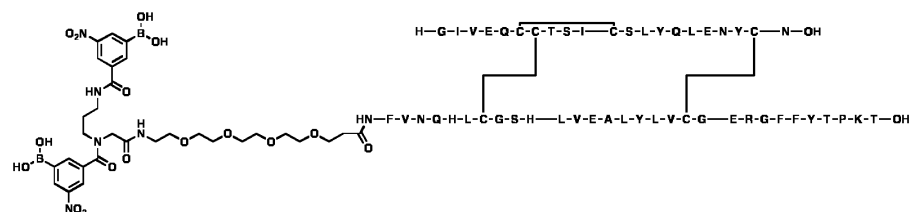


Example 873:

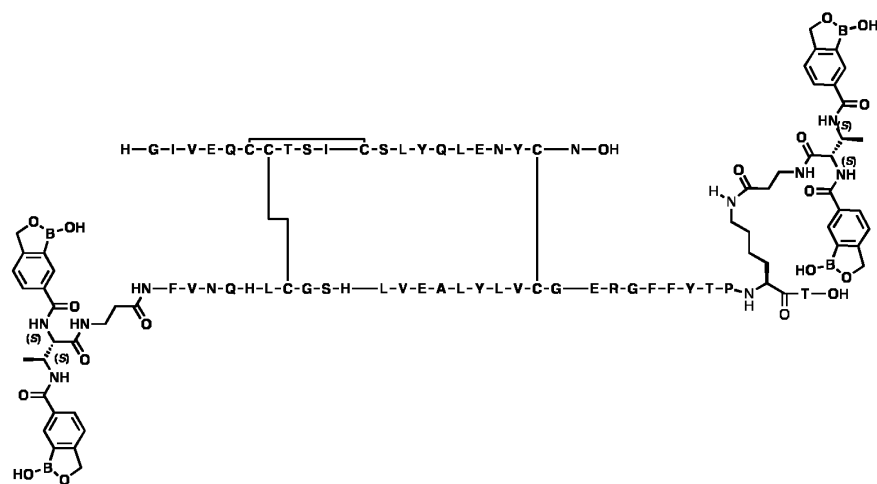


10

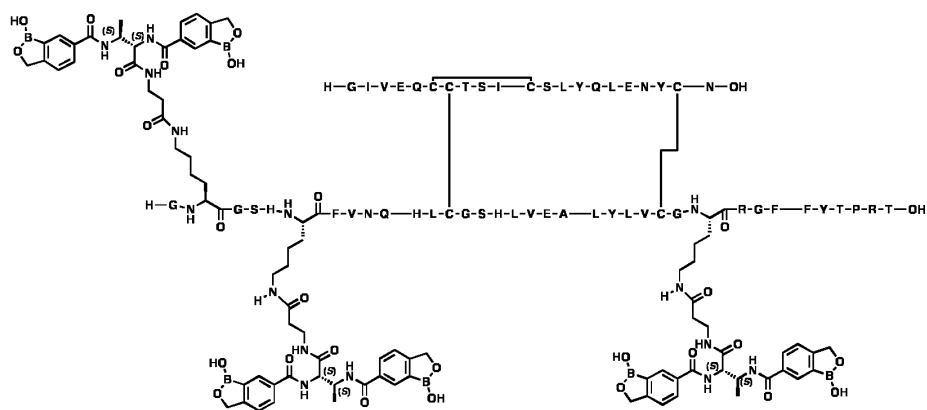
Example 874:



5 Example 875:



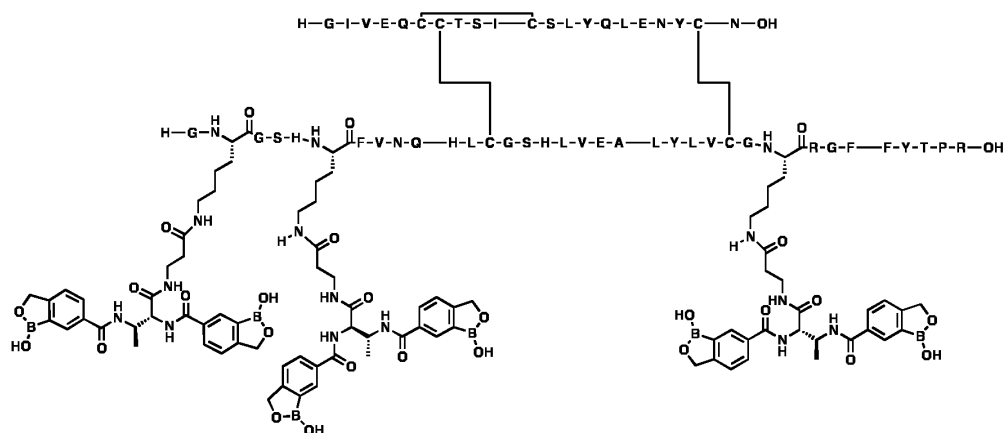
Example 876:



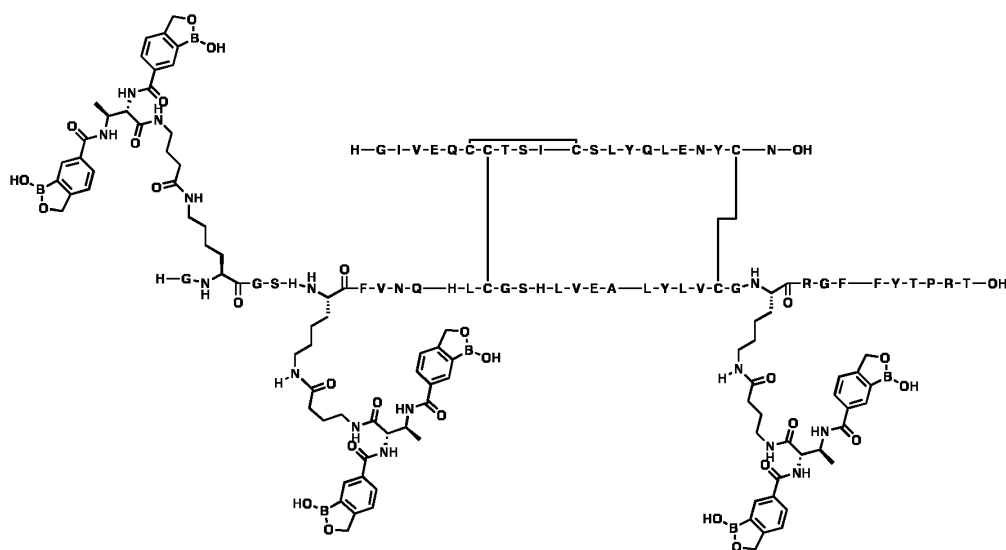
10

15

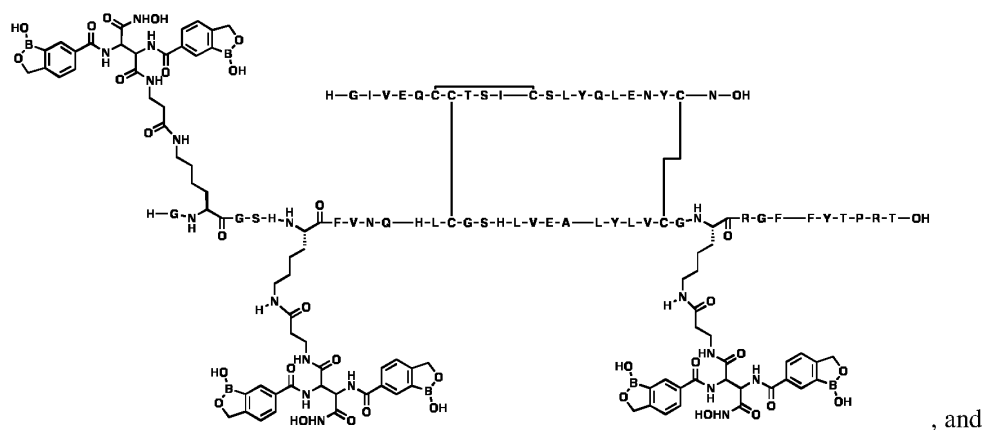
5 Example 877:



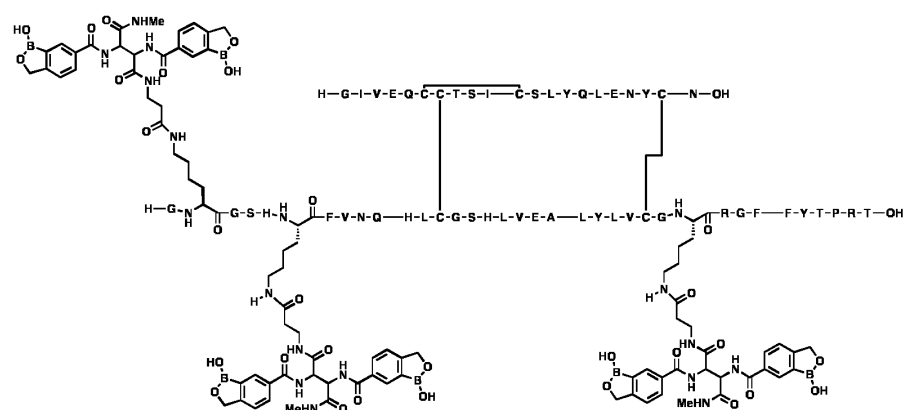
Example 878:



5 Example 879:



Example 880:



- 10 24. The compound of claim 1 or 2, wherein X1 is a polypeptide drug substance and/or an insulin optionally having from 0 to 4 residues replaced, inserted, or mutated to lysines, and wherein the lysines are each conjugated to a Z1c.
25. The compound of claim 1 or 2, wherein one or more amines are each independently acetylated and/or independently alkylated.
- 15 26. The compound of claim 1 or 2, wherein X1 comprises a polypeptide drug substance and the covalent conjugation to X1 is to amino group(s) in one or more lysine residues and/or to the N-terminal amino groups in X1.
27. The compound of claim 1 or 2, wherein each R1 is independently selected from a C₁-C₂₂ alkyl group, a C₁-C₂₂ acyl group, a (C₃-C₈)cycloalkyl group, a C₁-C₂₂ haloalkyl group,

25. The compound of claim 1 or 2, wherein one or more amines are each independently acetylated and/or independently alkylated.

- 15 26. The compound of claim 1 or 2, wherein X1 comprises a polypeptide drug substance and the covalent conjugation to X1 is to amino group(s) in one or more lysine residues and/or to the N-terminal amino groups in X1.

27. The compound of claim 1 or 2, wherein each R1 is independently selected from a C₁-C₂₂ alkyl group, a C₁-C₂₂ acyl group, a (C₃-C₈)cycloalkyl group, a C₁-C₂₂ haloalkyl group,

5 an aryl group, and a heteroaryl group, each R1 optionally comprises one or more C₁-C₂₂ alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, C₁-C₂₂ alkyl, or aryl groups.

28. The compound of claim 1 or 2, wherein X₄ is selected from -COOH, -
(CH₂)_mCOOH, a C₁-C₂₂ alkyl group, a C₁-C₂₂acyl group, a (C₃-C₈)cycloalkyl group, a C₁-C₂₂
haloalkyl group, an aryl group, and a heteroaryl group, each X₄ optionally comprises one or
10 more C₁-C₂₂ alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, C₁-C₂₂ alkyl, or
aryl groups; wherein m is 1, 2, 3, 4, or 5.

29. The compound of claim 3, wherein the alkyl group of Y₉ is a C₁-C₂₂ alkyl.

30. The compound of claim 29, wherein Y₉ is CH₃.

31. The compound of claim 1, wherein the at least one primary or secondary amine
15 in FF1-FF223 is covalently conjugated to B₆.

32. The compound of claim 1 or 2, wherein an amine in the compound is conjugated
via an amide linkage to an aromatic boron-containing group.

33. The compound of claim 32, wherein the aromatic boron-containing group is
selected from a phenylboronic acid, boroxole, and phenylboronate.

34. The compound of any one of claims 1-33, wherein the compound is formulated
20 in a solution comprising one or more of a buffer, stabilizer, vasodilator, preservative,
surfactant, salt, sugar, or compounds containing one or more hydroxyls, alcohols, diols, or
phenols.

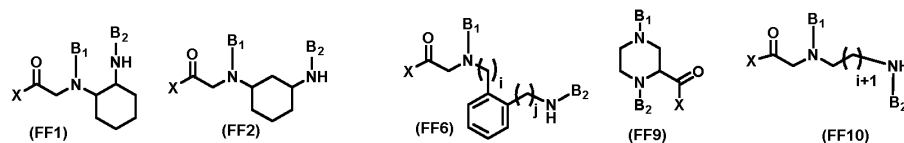
35. The compound of claim 34, wherein the solution comprises one or more of
25 citrate, zinc, and/or cresol.

36. The compound of claim 1 or 2, wherein Z1c is conjugated to a cysteine.

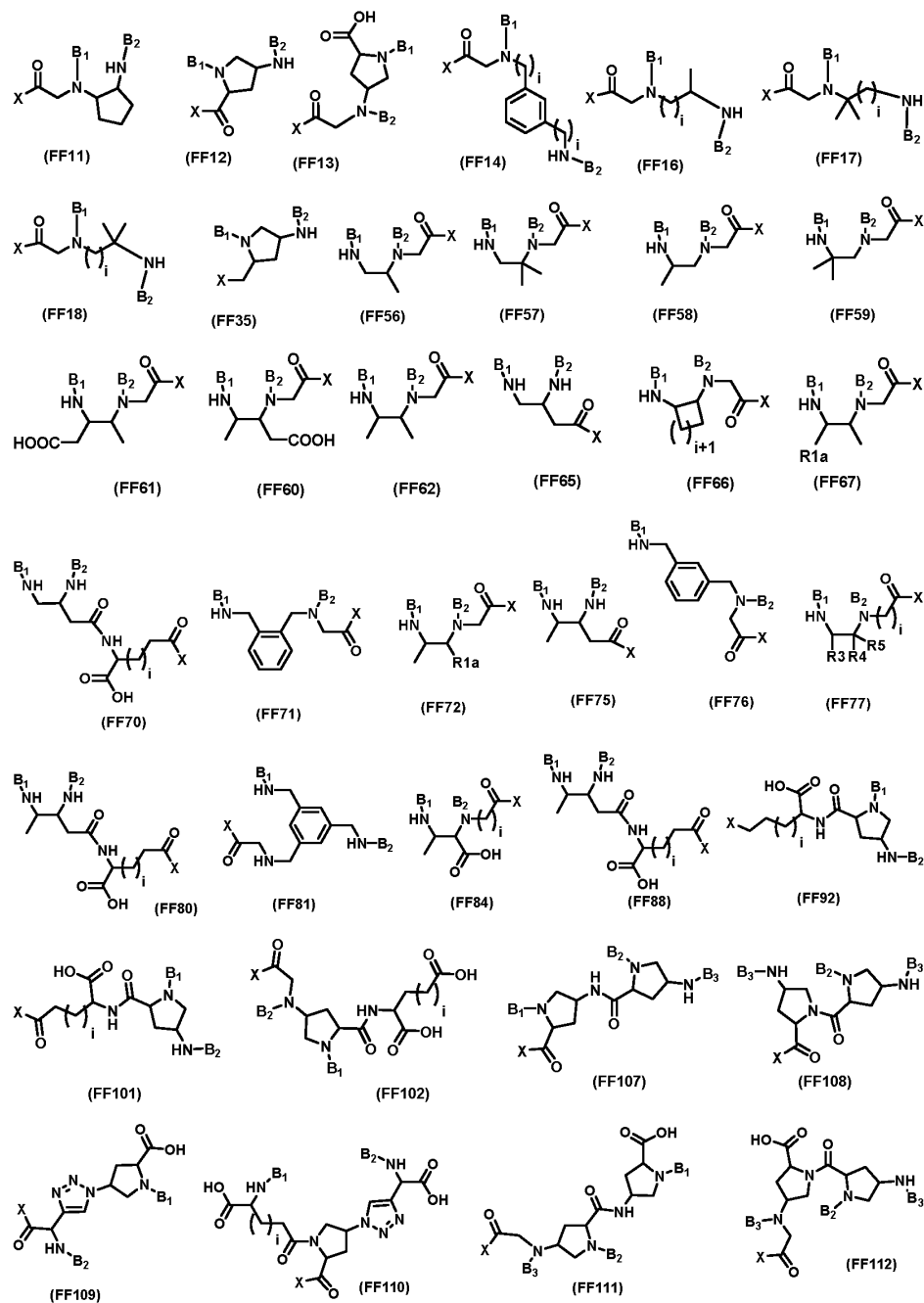
37. The compound of claim 1 or 2, wherein the compound is covalently conjugated
either directly or through a linker to a diol, sugar, carbohydrate or a diol containing molecule.

38. The compound of claim 1 or 2, wherein the compound is covalently conjugated
30 to an antibody, albumin or a fragment thereof, or covalently conjugated either directly or
through a linker to a molecule that can bind to at least one protein present in human plasma.

39. The compound of any one of claims 1-7, wherein the compound comprises at
least one Z1c selected from:

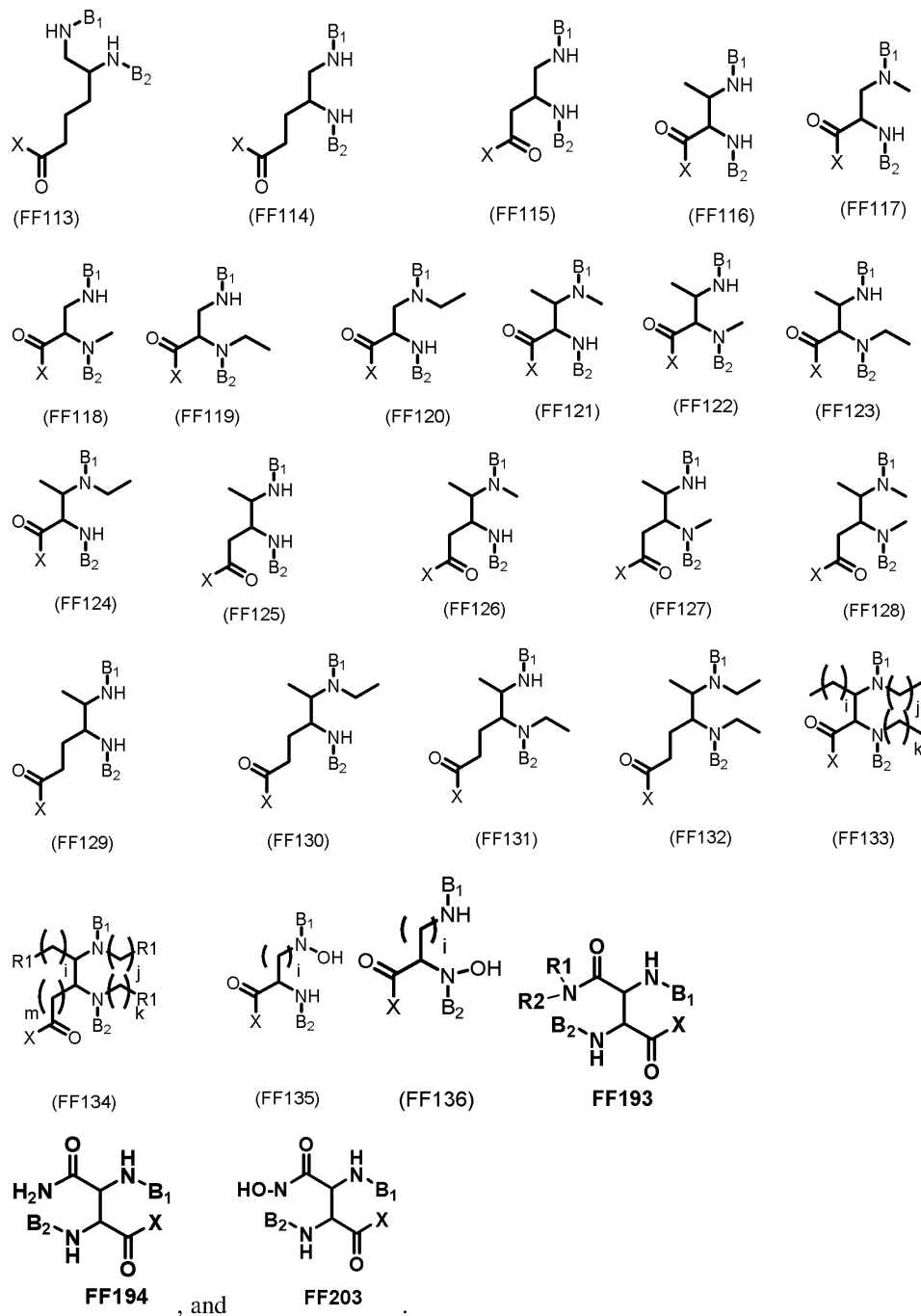


5



10

5



10

40. The compound of claim 1 or 2, wherein the compound comprises at least one Z1c having at least one chiral center and selected from FF1, FF2, FF5, FF9, FF11-FF13, FF15-

5 FF24, FF27, FF31, FF34-FF36, FF38, FF39, FF43-FF58, FF60-FF70, FF72-FF75, FF77-FF80, FF82-FF84, FF86-FF212, FF216-FF220, FF222, FF223, and combinations thereof.

41. The compound of claim 40, wherein the compound comprises at least one FF12 and/or FF116; and

10 wherein the stereochemistry of FF12 and FF116 is independently selected from (*S,S*); (*S,R*); (*R,R*); and (*R,S*).

42. The compound of claim 1 or 2, wherein X1 comprises human insulin or a human insulin analogue comprising an A-chain and a B-chain, wherein the C-terminus of the A-chain of the human insulin analogue is optionally extended with a polypeptide of up to 20 residues, and/or the N-terminus of the B-chain of the human insulin analogue is optionally extended with
15 a polypeptide of up to 10 residues.

43. The compound of claim 42, wherein X1 comprises at least one lysine having an amine side chain, and Z1c is covalently conjugated directly to the amine side chain.

44. The compound of claim 1 or 2, wherein X1 comprises a drug substance covalently conjugated to at least one Z1c through an acid containing linker.

20 45. A composition or a mixture comprising at least one compound of any one of claims 1-44, for use as a medicament for the treatment of diabetes, for control of blood sugar levels, or to control the release of a drug based on physiological levels of diol containing small molecules or sugars.

46. A method of administering the compound of any one of claims 1-44 to a human
25 subject as a therapeutic or prophylactic agent.

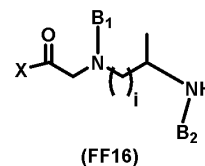
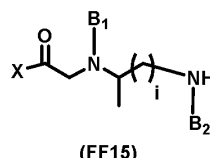
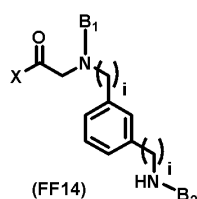
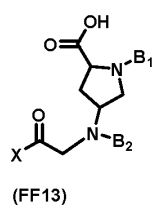
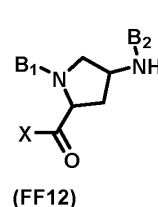
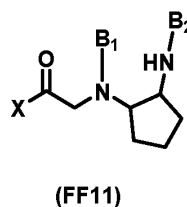
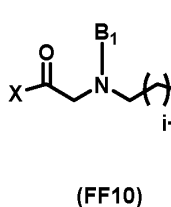
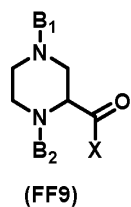
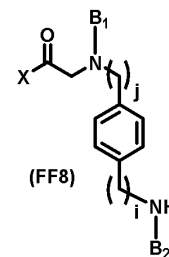
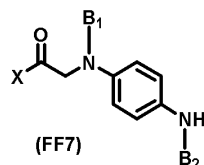
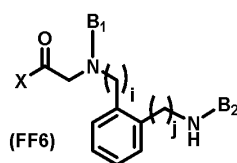
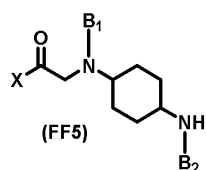
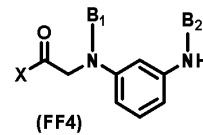
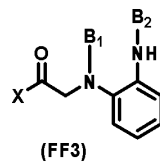
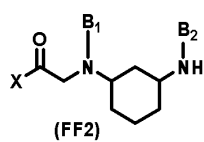
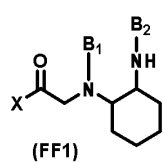
47. A method of making a compound of any one of claims 1-44, wherein the method comprises at least one alkylation and/or amidation step.

48. A method of treating a subject by administering a device or formulation comprising a compound of any one of claims 1-44 and Examples 1-880.

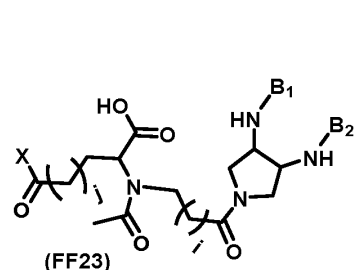
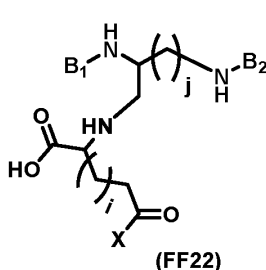
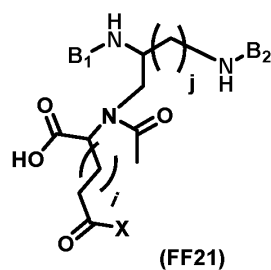
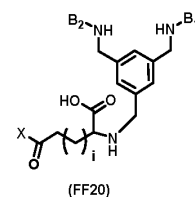
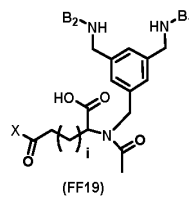
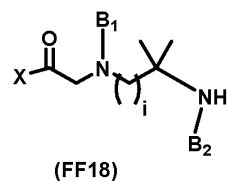
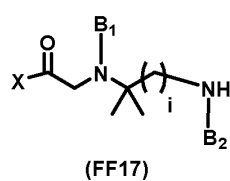
30 49. A method of treatment or prevention of diabetes, impaired glucose tolerance, hyperglycemia, or metabolic syndrome, wherein the method comprises administering to a subject in need thereof a therapeutically effective amount of the compound of any one of claims 1-44 or the composition or the mixture of claim 45.

50. A compound selected from Formulae FF1-FF224:
35 wherein Formulae FF1-FF48 are:

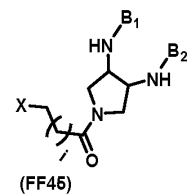
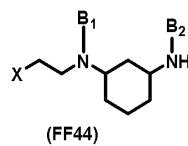
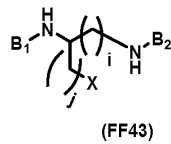
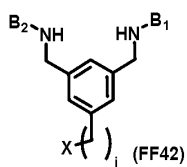
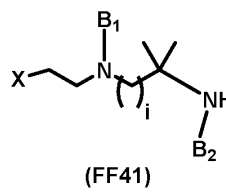
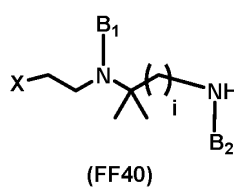
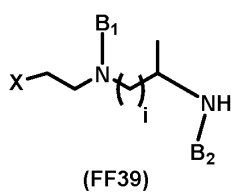
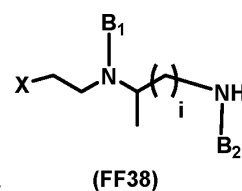
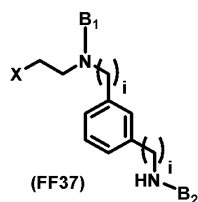
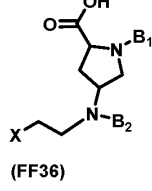
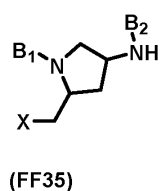
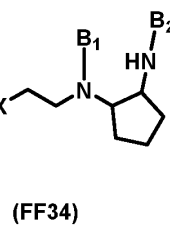
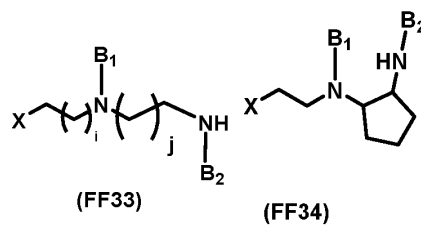
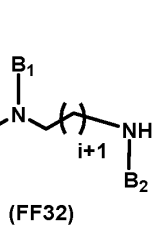
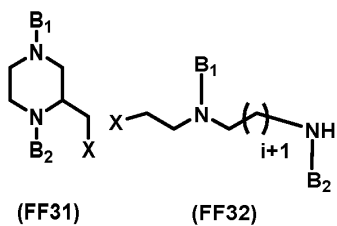
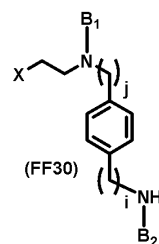
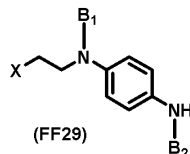
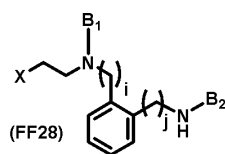
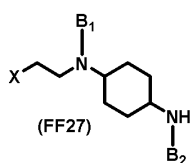
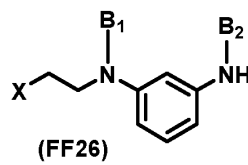
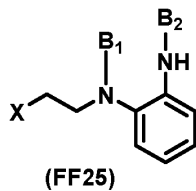
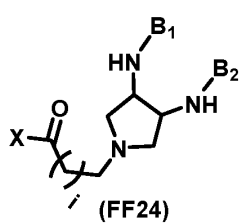
5



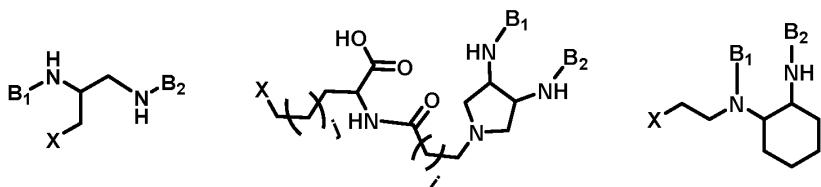
10



5



10



5 (FF46) (FF47) and (FF48);

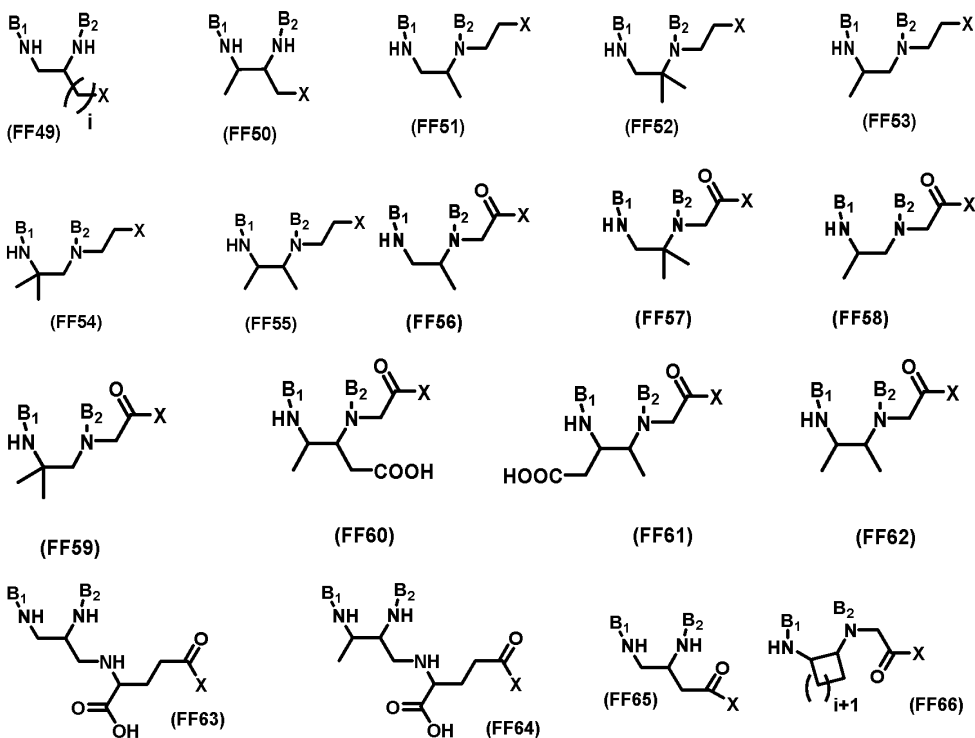
wherein X is selected from an amine, OH, and halogen; and

i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7; and

B₁ and B₂, which may be identical or different, each independently represents an
10 aromatic boron-containing group; and

wherein Formulae FF49-FF88 are:





j is 1, 2, 3, 4, 5, 6, or 7;

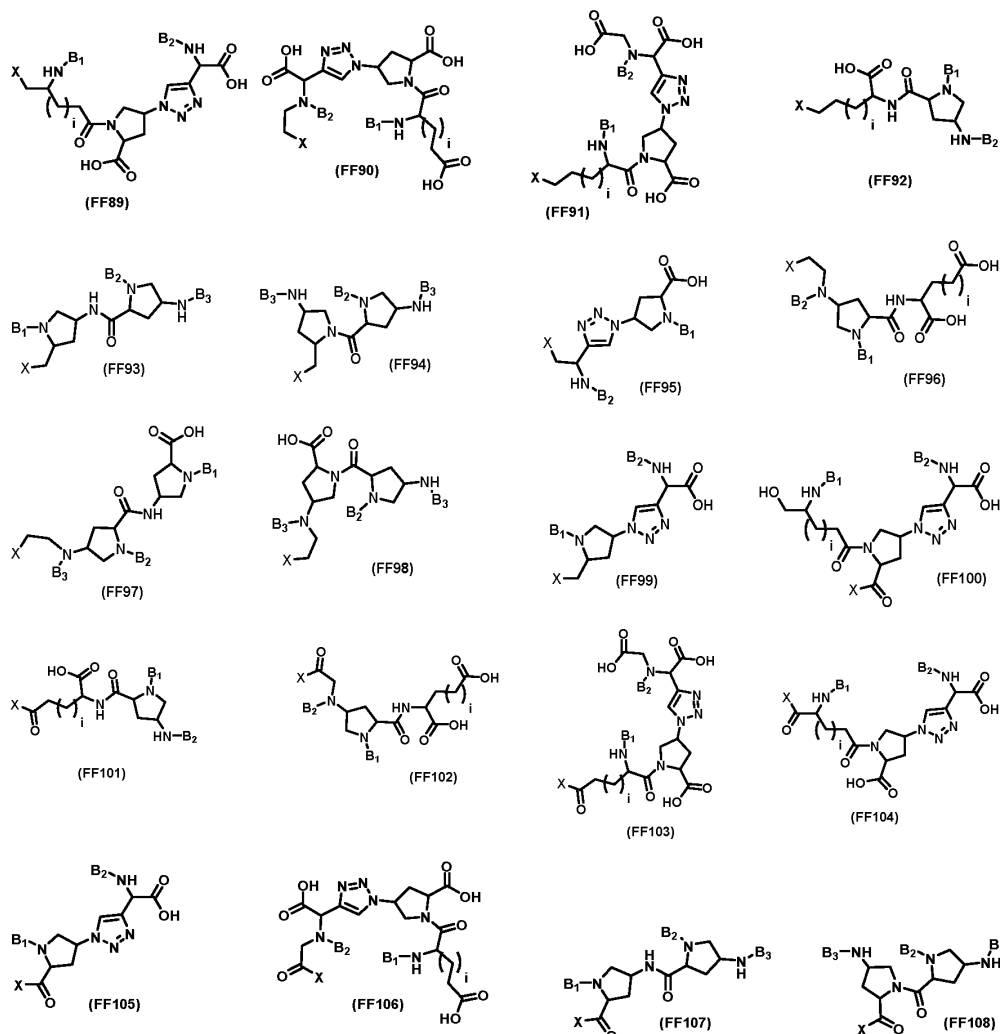
5 R1a is selected from COOH, CH₃, H, and OH;

R2, R3, R4 and R5 is each independently selected from CH₃, H, OH, and COOH, and at least one of R2, R3, R4 and R5 is CH₃ or OH; and

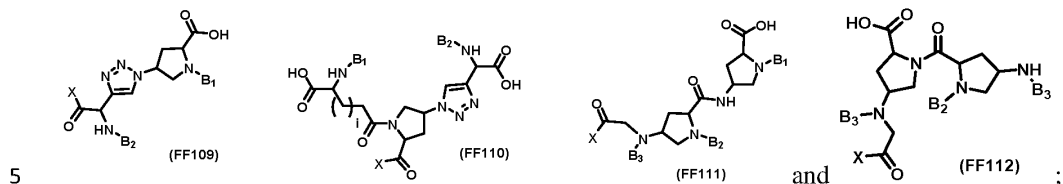
B₁ and B₂, which may be identical or different, are each independently an aromatic boron-containing group; and

10

wherein Formulae FF89-FF112 are:



15

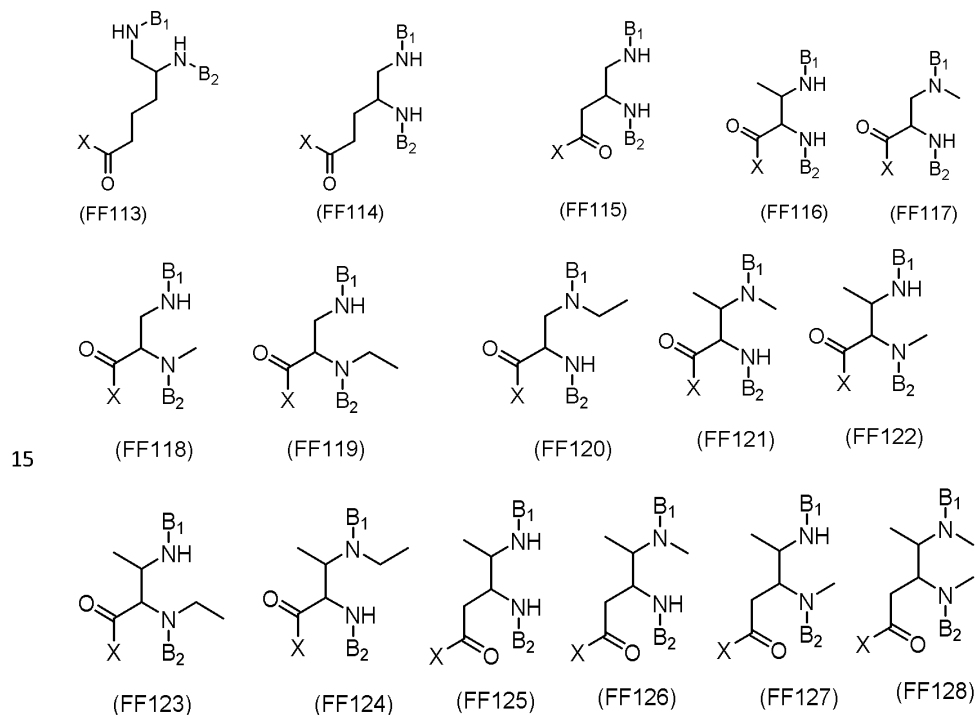


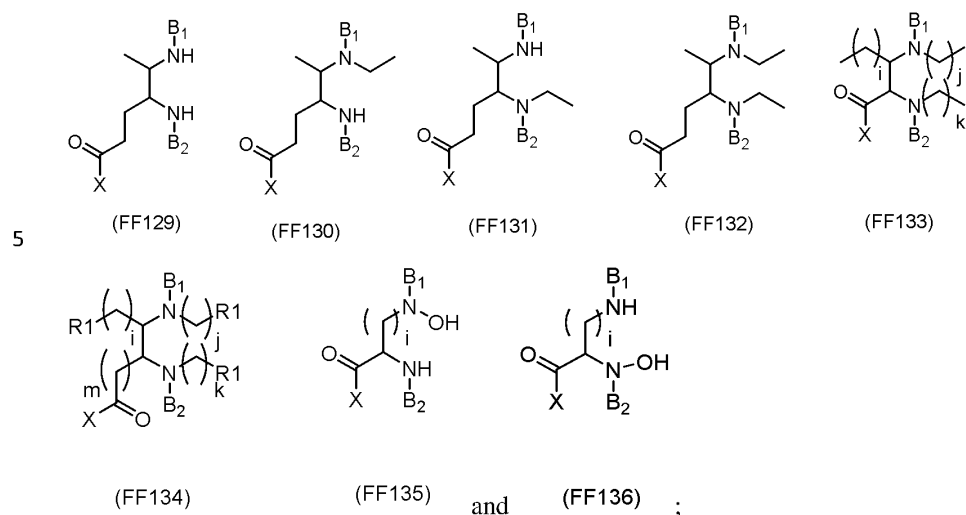
wherein X is selected from an amine, OH, and halogen;

i is 1, 2, 3, 4, 5, 6, or 7; and

B₁, B₂ and B₃, which may be identical or different, each independently represents an
 10 aromatic boron-containing group, a carboxylic acid derivative, or a H, wherein at least two of
 B₁, B₂ and B₃ in each FF structure are independently an aromatic boron-containing group; and

wherein Formulae FF113-FF136 are:





wherein X is selected from an amine, OH, and halogen;

i is 1, 2, 3, 4, 5, 6, or 7;

j is 1, 2, 3, 4, 5, 6, or 7;

10 k is 1, 2, 3, 4, 5, 6, or 7;

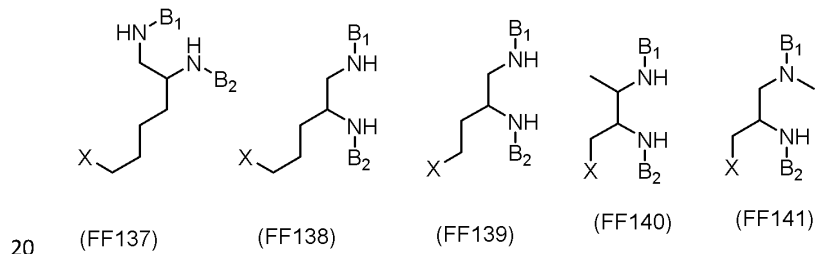
m is 1, 2, 3, 4, 5, 6, or 7;

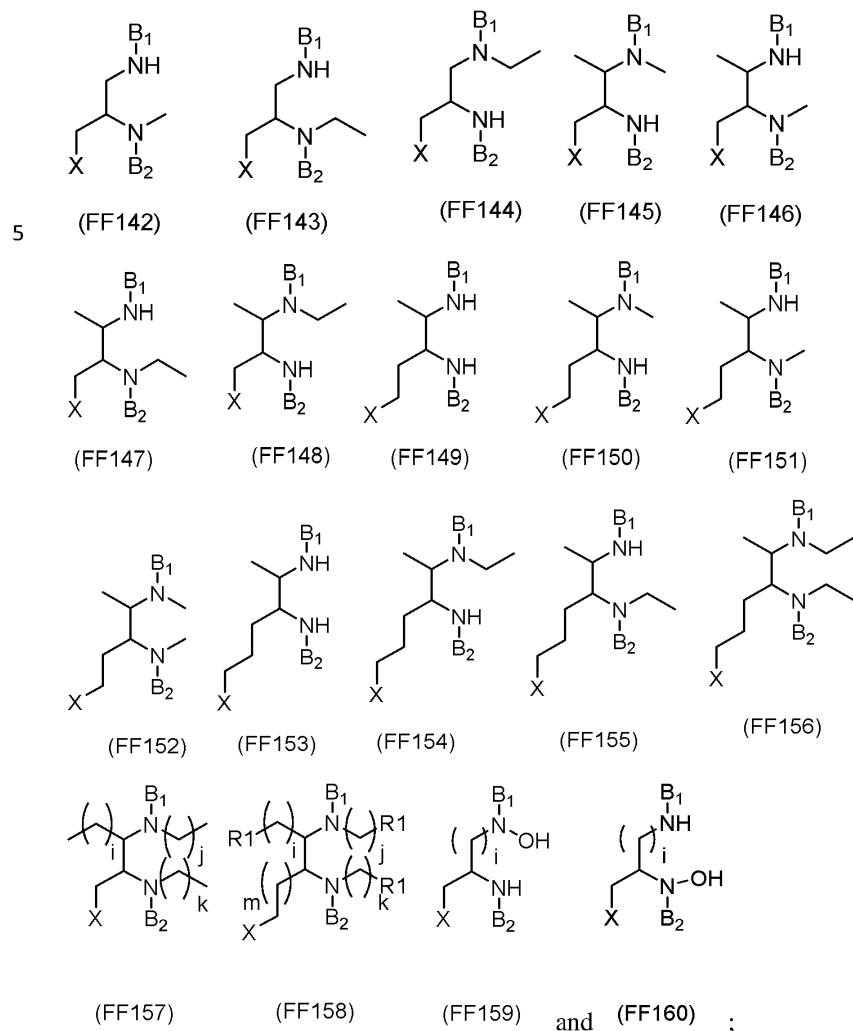
each R1 is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl

15 groups; and

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group; and

wherein Formulae FF137-FF160 are:



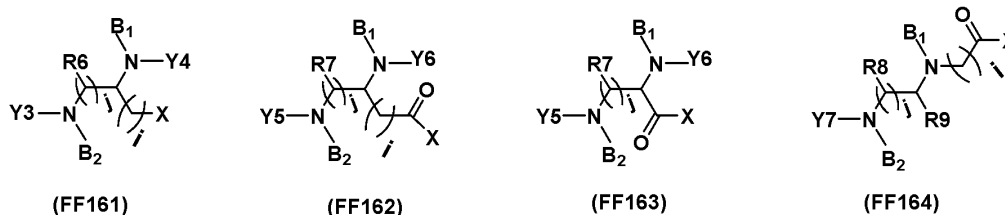


- 10 wherein X is selected from an amine, OH, and halogen;
 i is 1, 2, 3, 4, 5, 6, or 7;
 j is 1, 2, 3, 4, 5, 6, or 7;
 k is 1, 2, 3, 4, 5, 6, or 7;
 m is 1, 2, 3, 4, 5, 6, or 7;
- 15 each R1 is independently selected from H, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each R1 optionally comprises

- 5 one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; and

B_1 and B_2 , which may be identical or different, each independently represents an aromatic boron-containing group; and

- 10 wherein Formulae FF161-FF164 are:



wherein X is selected from an amine, OH, and halogen;

i is 1, 2, 3, 4, or 5;

- 15 j is 1, 2, 3, 4, or 5;

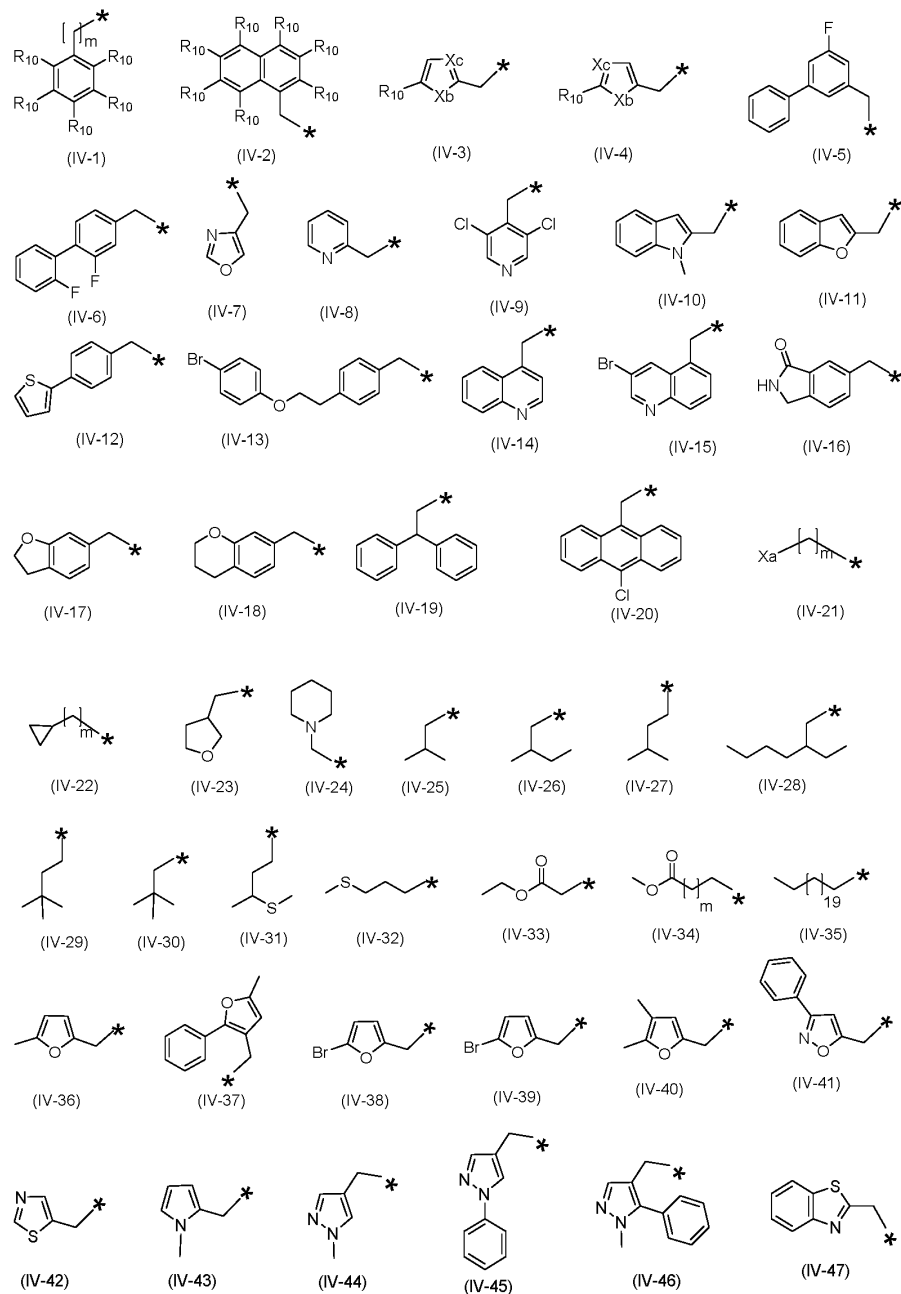
each R6, R7, R8, and R9 for different values of j is independently selected from H, CF₃, CH₃, CHF₂, and (CH₂)_mCH₃, wherein m is 1, 2, 3, 4, or 5;

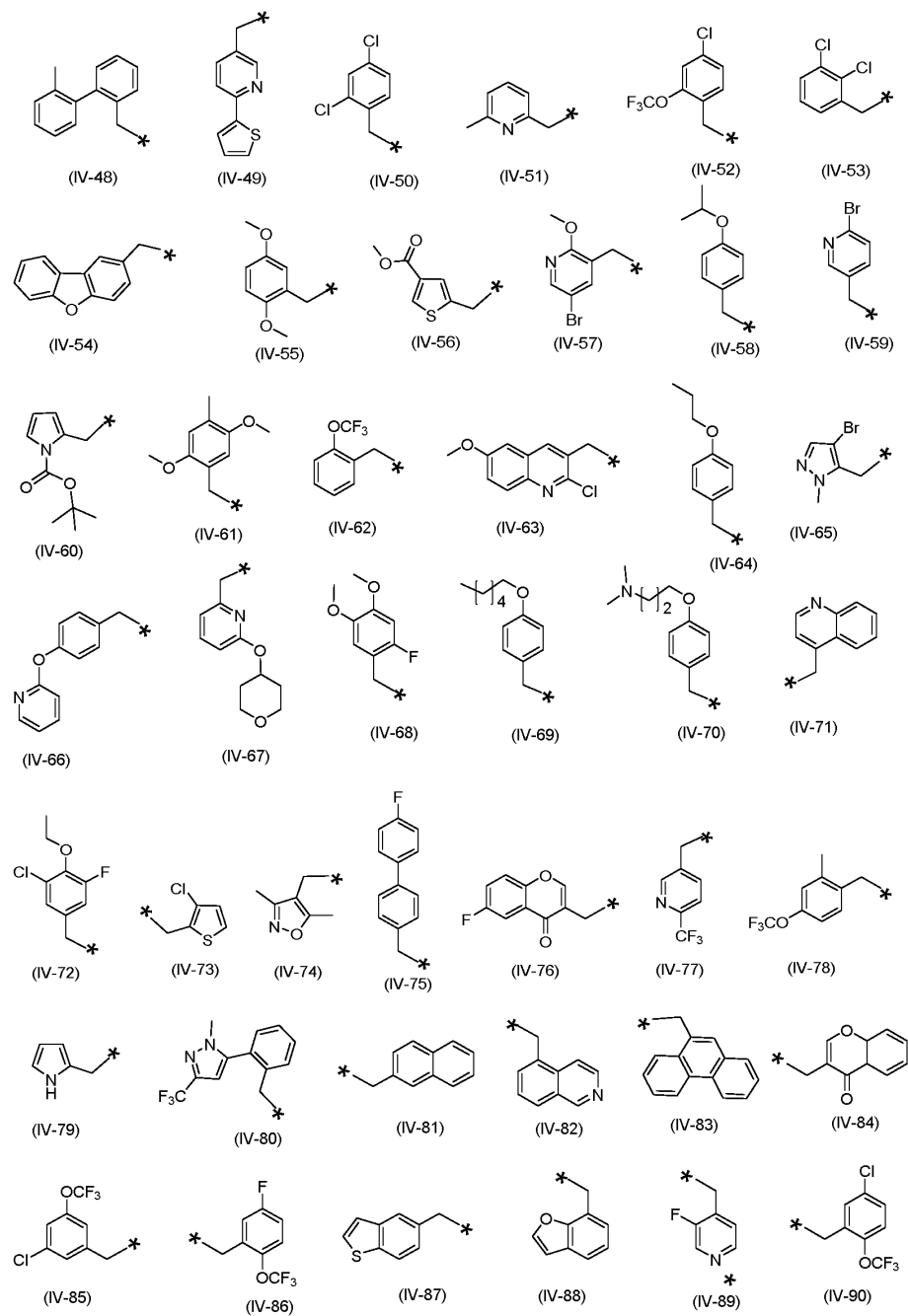
Y3, Y4, Y5, Y6 and Y7 are each independently selected from H, CH₂—X4, and Formulae IV-1 to IV-135;

- 20 wherein X4 is selected from -COOH, -(CH₂)_mCOOH, an alkyl group, an acyl group, a cycloalkyl group, a haloalkyl group, an aryl group, and a heteroaryl group, each optionally comprising one or more alkyl-halide, halide, sulfhydryl, aldehyde, amine, acid, hydroxyl, alkyl or aryl groups; wherein m is 1, 2, 3, 4, or 5;

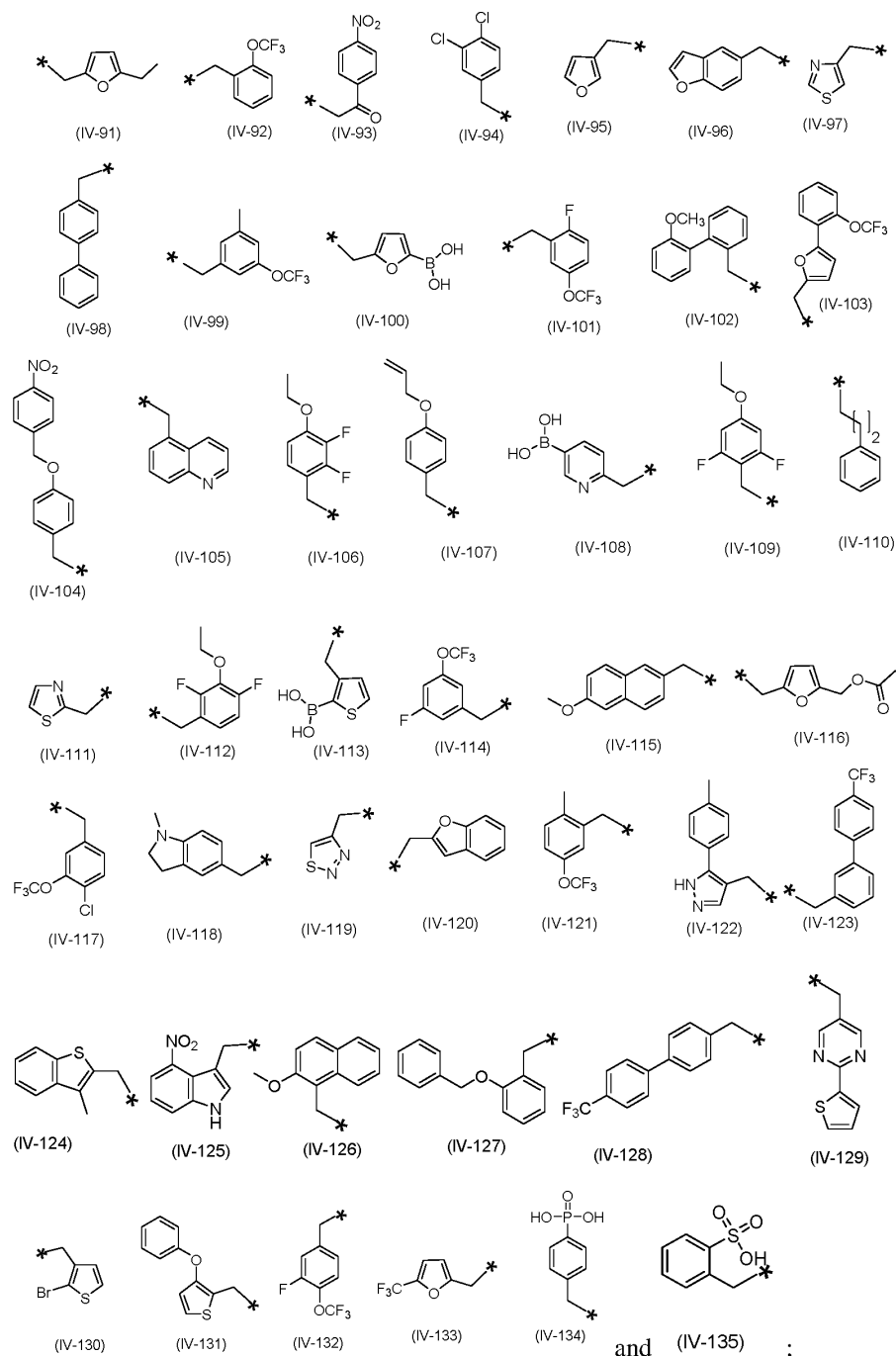
- wherein at least one of Y5, Y6, and Y7 in Formulae FF162 and FF163 is not H and at
25 least one of Y7, R8 and R9 in FF164 is not H; and

wherein Formulae IV-1 to IV-135 are:





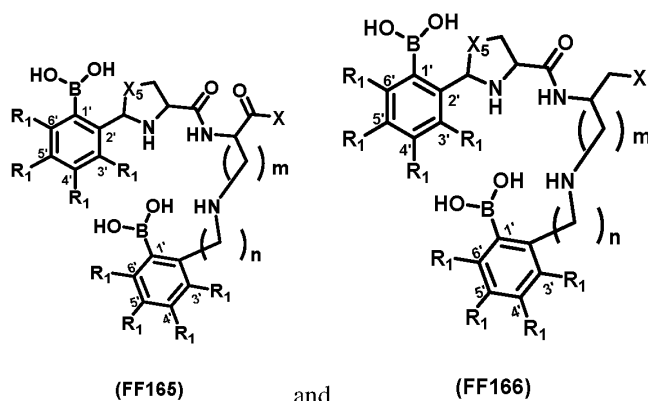
5



10 wherein:

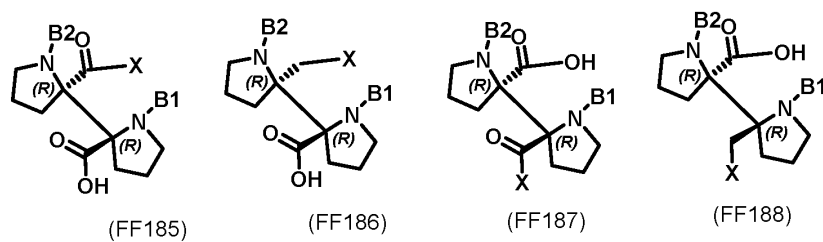
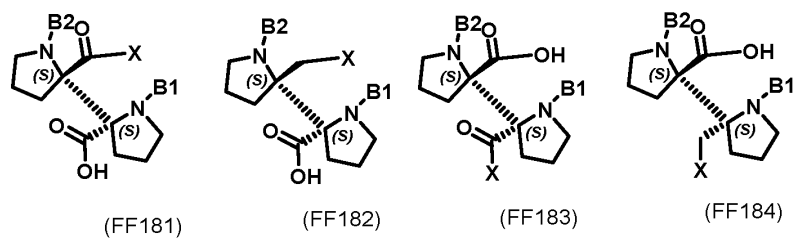
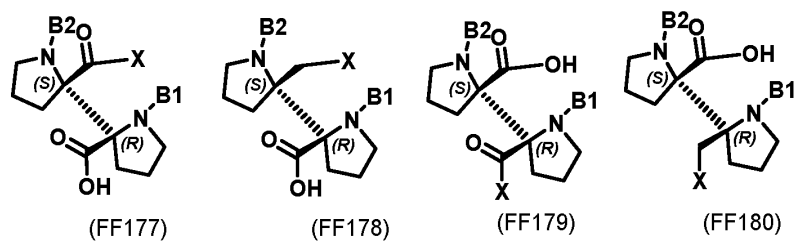
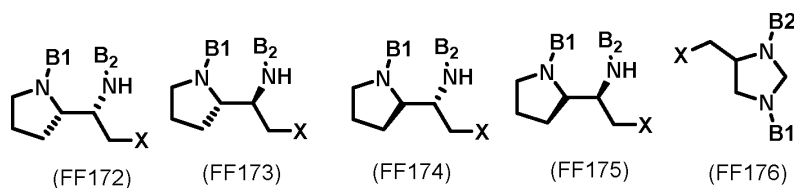
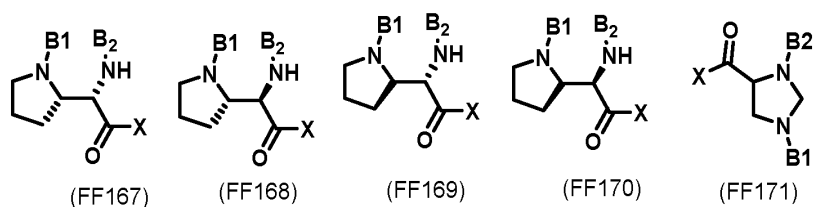
- 5 Xa represents CH=O, CHF₂, CF₃, CH₂SH, COOH, CH₂OH, CH₂NO₂, CH₂NH₂, CH₃, C(CH₃)₃, CH(CH₃)₂, CH((CH₂)₃ CH₃)₂, or CH(CH₂ CH₃)₂;
 Xb represents O, NH, CH₂, or S;
 Xc represents CH or N;
 each R₁₀ is independently selected from H, F, Cl, Br, CH₃, CF₃, CH=O, OH, COOH,
 10 and (CH₂)_nCH₃,
 m is 1, 2, 3, 4, or 5; and n is 1, 2, 3, 4, or 5;
 B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group; and
 * in Formulae IV-1 to IV-135 represents the point of attachment to corresponding
 15 Formulae FF161-164; and

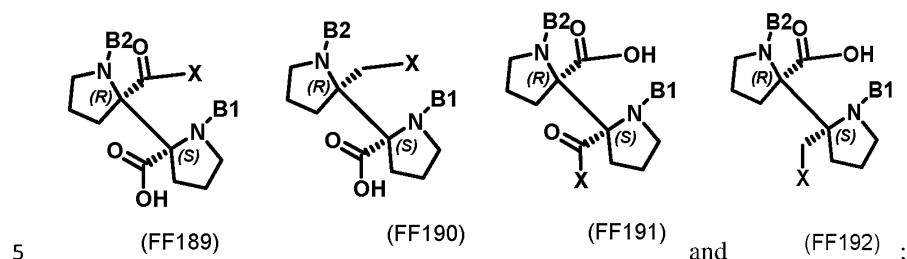
wherein Formulae FF165-FF166 are:



- 20 wherein X is selected from an amine, OH, and halogen;
 m is 1, 2, 3, 4, 5, 6, or 7;
 n is 1, 2, 3, 4, 5, 6, or 7;
 X5 is S, O, or NH; and
 25 each R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7; and

5 wherein Formulae FF167-FF192 are:

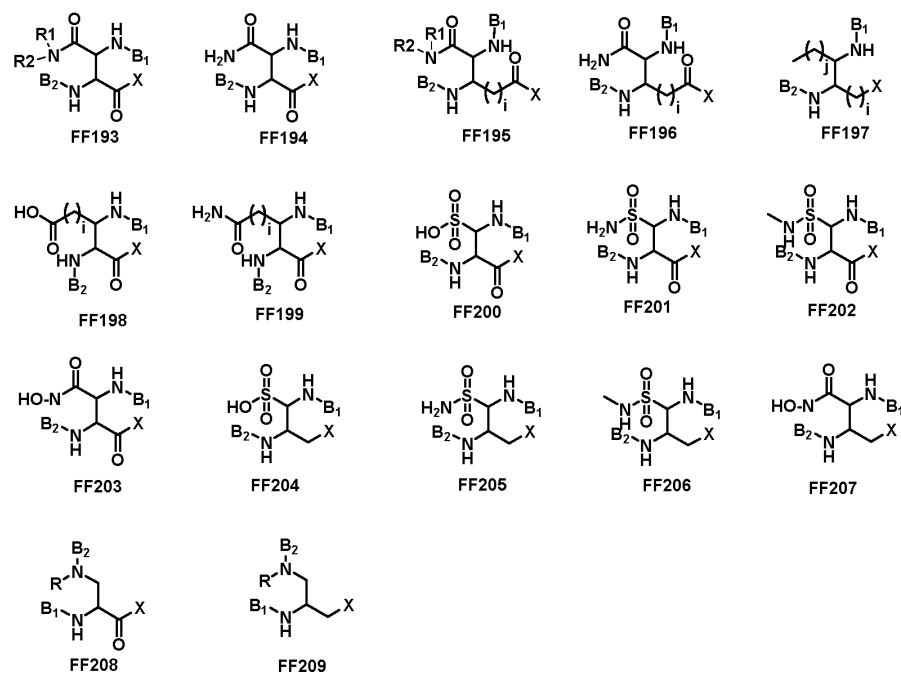




wherein X is selected from an amine, OH, and halogen;

B₁ and B₂, which may be identical or different, each independently represents an aromatic boron-containing group, and

10 wherein Formulae FF193-FF209 are:



wherein R in FF208 and FF209 is an alkyl, aryl or halide that is covalently conjugated through at least one CH₂ group to the amino group in the side chain of FF208 or FF209;

R₁ and R₂ are independently selected from H, CH₃, alkyl, and formulae IV-1 to IV-

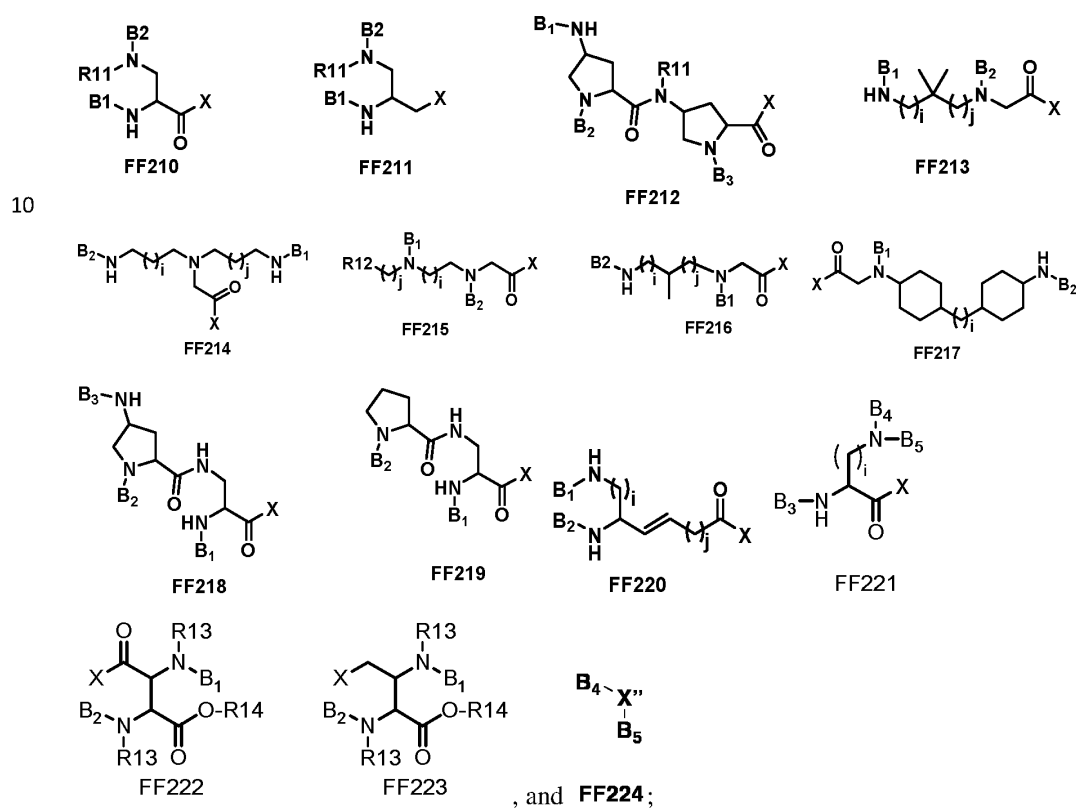
15 135;

i is 1, 2, 3, 4, or 5;

j is 1, 2, 3, 4, or 5; and

- 5 wherein X is selected from an amine, OH, and halogen; and
 B₁ and B₂, which may be identical or different, each independently represents an
 aromatic boron-containing group; and

wherein Formulae FF210-FF224 are:



- wherein R₁₁ in FF210 to FF212 is independently selected from Formulae IV-1 to IV-135 and
- 15 R₁₂ is selected from an amine, a hydroxyl, an alkyl, and a halide group;
- wherein each R₁₃ is independently selected from H, CH₃, alkyl, aryl, and formulae IV-1 to IV-135; R₁₄ is selected from H, CH₃, alkyl, aryl, and heteroaryl;
- wherein X is independently selected from an amine, OH, and halogen;
- X'' is an amine;
- 20 i is 1, 2, 3, 4, or 5;
- j is 1, 2, 3, 4, or 5; and

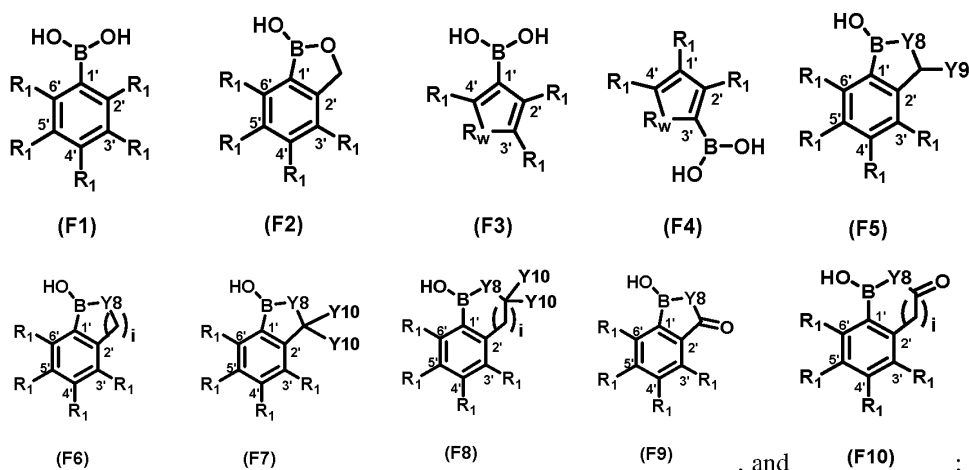
5 wherein at least one primary or secondary amine in FF1-FF223 is optionally covalently conjugated to B₆; and

B₁, B₂, B₃, B₄, B₅ and B₆ each independently represents an aromatic boron-containing group, wherein in each FF structure containing B₁, B₂ and B₃ groups, at least two of the B₁, B₂ and B₃ groups are independently an aromatic boron-containing group; and

10 when X is an amine in any one of Formulae FF1 to FF223, X is optionally acetylated or alkylated.

51. The compound of claim 50, wherein the compound comprises at least one of B₁, B₂ and B₃ independently selected from Formulae F1-F12 or wherein the compound comprises
15 at least one of B₄, B₅ and B₆ independently selected from Formulae F1-F10,

wherein Formulae F1-F10 are:



wherein for B₁, B₂, B₃:

20 one R₁ represents (C=O)---*, S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein ---* represents an attachment point to the rest of Z1c, and m is 1, 2, 3, 4, 5, 6, or 7;

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

25 wherein for B₄, B₅:

one R₁ for B₄ represents (CH₂)_m---ø, wherein ---ø represents the attachment point (representing a covalent bond) to an amine in X1 and one R₁ for B₅ represents (C=O)---*,

- 5 S(=O)(=O)---*, (CH₂)_m(C=O)---*, or (CH₂)_m---*, wherein ---* represents the attachment point to the same amine in X1, and m is 1, 2, 3, 4, 5, 6, or 7;

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —(SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

- 10 wherein for B₆:

one R₁ for B₆ represents (CH₂)_m---ø, wherein ---ø represents the attachment point (representing a covalent bond) to the rest of the compound, and m is 1, 2, 3, 4, 5, 6, or 7;

each remaining R₁ is independently selected from H, F, Cl, Br, OH, CH₂-NH₂, NH₂, (C=O)-NH₂, CH=O, SO₂CH₃, SO₂CF₃, CF₃, CHF₂, NO₂, CH₃, OCH₃, O(CH₂)_mCH₃, —

- 15 (SO₂)NH CH₃, —(SO₂)NH(CH₂)_mCH₃, and OCF₃, wherein m is 1, 2, 3, 4, 5, 6, or 7;

for Formulae F3-F4:

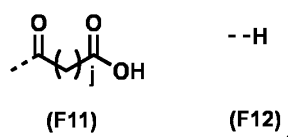
R_w is O or S;

for Formulae F5-F10:

- 20 Y8 is selected from O, N, and NR, wherein R is an alkyl group or H;
Y9 is H, CH₃, or an alkyl group, provided that when Y8 is O, Y9 is a CH₃ or higher alkyl group;
each Y10 is independently selected from H, CH₃, F, CF₃, and OCH₃; and
i is 1, 2, or 3; and

25

wherein Formulae F11-F12 are:



j is 1, 2, 3, 4, 5, 6, or 7; and

- 30 --- represents the attachment point to the rest of Z1c.

52. The compound of claim 50 or 51, wherein the compound is selected from:

N-(3-(3-borono-5-nitrobenzamido)propyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS01);

- 5 N-(4-((4-(3-borono-5-nitrobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS02);
N-(4-((3-borono-5-nitrobenzamido)methyl)benzyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS03);
N-(3-((3-borono-5-nitrobenzamido)methyl)benzyl)-N-(3-borono-5-nitrobenzoyl)glycine
10 (DS04);
N-(4-(3-borono-5-nitrobenzamido)butyl)-N-(3-borono-5-nitrobenzoyl)glycine (DS05);
N-(3-(3-borono-5-fluorobenzamido)propyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS06);
N-(3-(3-borono-5-fluorobenzamido)-2,2-dimethylpropyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS07);
15 bis(3-(3-borono-5-fluorobenzamido)propyl)glycine (DS08);
N-(4-((3-borono-5-fluorobenzamido)methyl)benzyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS09);
N-(3-((3-borono-5-fluorobenzamido)methyl)benzyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS10);
20 N-(2-(3-borono-5-fluorobenzamido)cyclohexyl)-N-(3-borono-5-fluorobenzoyl)glycine (DS11);
N-(3-(3-borono-4-fluorobenzamido)propyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS12);
N-(4-((4-(3-borono-4-fluorobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS13);
N-(3-(3-borono-4-fluorobenzamido)-2,2-dimethylpropyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS14);
25 fluorobenzoyl)glycine (DS14);
N-(4-((3-borono-4-fluorobenzamido)methyl)benzyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS15);
N-(3-((3-borono-4-fluorobenzamido)methyl)benzyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS16);
30 N-((1S,2R)-2-(3-borono-4-fluorobenzamido)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS17);
N-((1S,2S)-2-(3-borono-4-fluorobenzamido)cyclohexyl)-N-(3-borono-4-fluorobenzoyl)glycine (DS18);
N-(3-(3-borono-5-bromobenzamido)propyl)-N-(3-borono-5-bromobenzoyl)glycine (DS19);
35 N-(4-((4-(3-borono-5-bromobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(3-borono-5-bromobenzoyl)glycine (DS20);

- 5 bis(3-(3-borono-5-bromobenzamido)propyl)glycine (DS21);
N-(4-((3-borono-5-bromobenzamido)methyl)benzyl)-N-(3-borono-5-bromobenzoyl)glycine (DS22);
N-(3-((3-borono-5-bromobenzamido)methyl)benzyl)-N-(3-borono-5-bromobenzoyl)glycine (DS23);
- 10 N-(2-(3-borono-5-bromobenzamido)cyclohexyl)-N-(3-borono-5-bromobenzoyl)glycine (DS24);
N-(3-(4-borono-3-fluorobenzamido)propyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS25);
N-(4-((4-borono-3-fluorobenzamido)cyclohexyl)methyl)cyclohexyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS26);
- 15 N-(3-(4-borono-3-fluorobenzamido)-2,2-dimethylpropyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS27);
bis(3-(4-borono-3-fluorobenzamido)propyl)glycine (DS28);
N-(4-((4-borono-3-fluorobenzamido)methyl)benzyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS29);
- 20 N-(3-((4-borono-3-fluorobenzamido)methyl)benzyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS30);
N-((1S,2R)-2-(4-borono-3-fluorobenzamido)cyclohexyl)-N-(4-borono-3-fluorobenzoyl)glycine (DS31);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-(1-hydroxy-1,3-
- 25 dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS32);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pentyl)glycine (DS33);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-2,2-dimethylpropyl)glycine (DS34);
- 30 bis(3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS35);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(3-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)benzyl)glycine (DS36);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-((1S,2R)-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine (DS37);
- 35 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butyl)glycine (DS38);

- 5 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-((1S,2S)-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine (DS39);
(R)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS40);
(S)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)propyl)glycine (DS41);
- 10 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)glycine (DS42);
N-(3-(4-borono-3,5-difluorobenzamido)propyl)-N-(4-borono-3,5-difluorobenzoyl)glycine (DS43);
- 15 N-(3-(4-borono-2-fluorobenzamido)propyl)-N-(4-borono-2-fluorobenzoyl)glycine (DS44);
N-(2-(N-ethyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)ethyl)-N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)glycine (DS45);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(2-(1-hydroxy-N-(2-hydroxyethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)ethyl)glycine (DS46);
- 20 N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)hexyl)glycine (DS47);
N-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-N-(4-((4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)cyclohexyl)methyl)cyclohexyl)glycine (DS48);
((2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS49);
- 25 ((2S,4S)-4-(3-borono-4-fluorobenzamido)-1-(3-borono-4-fluorobenzoyl)pyrrolidine-2-carbonyl)glycine (DS50);
((2S,4S)-4-(3-borono-5-nitrobenzamido)-1-(3-borono-5-nitrobenzoyl)pyrrolidine-2-carbonyl)glycine (DS51);
- 30 ((2S,4S)-4-(5-borono-2-fluorobenzamido)-1-(5-borono-2-fluorobenzoyl)pyrrolidine-2-carbonyl)glycine (DS52);
(S)-(1,4-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)piperazine-2-carbonyl)glycine (DS53);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-benzyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS54);
- 35

- 5 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(4-(trifluoromethyl)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS55);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-N-ethyl-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS56);
- 10 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-propyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS57);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-isobutyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS58);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-((5-(thiophen-2-yl)pyridin-2-yl)methyl)-1,3-
- 15 dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS59);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-isopentyl-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS60);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(quinolin-5-ylmethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-
- 20 carboxamide (DS61);
(S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(2-(trifluoromethoxy)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS62);
- 25 (S)-N-(3-amino-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-3-oxopropyl)-1-hydroxy-N-(4-(methylsulfonyl)benzyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS63);
(3-((2S,4S)-4-(5-borono-2-(methylsulfonyl)benzamido)-2-carbamoylpyrrolidine-1-carbonyl)-4-(methylsulfonyl)phenyl)boronic acid (DS64);
- 30 (4-(((3S,5S)-1-(4-borono-2,6-difluorobenzoyl)-5-carbamoylpyrrolidin-3-yl)carbamoyl)-3,5-difluorophenyl)boronic acid (DS65);
(R,E)-4,5-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pent-2-enoic acid (DS66);
(2S,4S)-1-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-
- 35 (1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxamide (DS67);

- 5 N,N'-((2S,3S)-1-amino-1-oxobutane-2,3-diyl)bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide) (DS68);
(R)-3,4-bis(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid (DS69);
3-((2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxamido)propanoic acid
10 (DS70);
(S)-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid (DS71);
(R)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pentanoic acid (DS72);
15 (2S,4R)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS73);
(2S,4R)-1-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-
20 2-carboxylic acid (DS74);
(2S,3S)-3-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-2-(1-hydroxy-7-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)butanoic acid (DS75);
(R)-5-(1-hydroxy-4-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-(1-hydroxy-7-(trifluoromethyl)-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)pentanoic
25 acid (DS76);
((2S,4S)-1-(5-borono-2-nitrobenzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS77);
((2S,4S)-1-(5-borono-2-(methylsulfonyl)benzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS78);
30 ((2S,4S)-1-(3-borono-2,6-difluorobenzoyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carbonyl)glycine (DS79);
(S)-3-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS80);
35 (S)-3-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS81);

- 5 (S)-(3-((3-boronobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS82);
(S)-(3-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS83);
(S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-nitrophenyl)boronic acid (DS84);
- 10 (S)-(5-((3-borono-N-(5,6-diamino-6-oxohexyl)-4-fluorobenzamido)methyl)-2-fluorophenyl)boronic acid (DS85);
(S)-(5-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS86);
- 15 (S)-(3-((3-borono-N-(5,6-diamino-6-oxohexyl)-4-fluorobenzamido)methyl) phenyl)boronic acid (DS87);
(S)-(5-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS88);
(S)-(5-((4-borono-3-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS89);
- 20 (S)-(4-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS90);
(S)-(4-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS91);
- 25 (S)-(4-((3-boronobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS92);
(S)-(4-((4-borono-2-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS93);
(S)-(4-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-2-fluorophenyl)boronic acid (DS94);
- 30 (S)-(5-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)-2-fluorophenyl)boronic acid (DS95);
(S)-(3-((4-borono-3,5-difluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-bromophenyl)boronic acid (DS96);
- 35 (S)-(3-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl) phenyl)boronic acid (DS97);

- 5 (S)-(3-((3-borono-5-bromo-N-(5,6-diamino-6-oxohexyl)benzamido)methyl)-5-methoxyphenyl)boronic acid (DS98);
(S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-bromophenyl)boronic acid (DS99);
(S)-(3-((3-borono-4-fluorobenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic
10 acid (DS100);
(S)-(3-((4-borono-3-methoxybenzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic acid (DS101);
(S)-(3-((4-borono-2-(trifluoromethyl)benzyl)(5,6-diamino-6-oxohexyl)carbamoyl)-5-fluorophenyl)boronic acid (DS102);
15 (S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-2-fluorophenyl)boronic acid (DS103);
(S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-2,6-difluorophenyl)boronic acid (DS104);
(S)-(3-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-
20 carboxamido)methyl)phenyl)boronic acid (DS105);
(S)-(4-((N-(5,6-diamino-6-oxohexyl)-1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)methyl)-3-methoxyphenyl)boronic acid (DS106);
(S)-N-(5,6-diamino-6-oxohexyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS107);
25 (S)-N-(4-amino-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-4-oxobutyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS108);
(S)-N-(6-amino-5-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)-6-oxohexyl)-1-hydroxy-N-((1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborol-6-yl)methyl)-1,3-
30 dihydrobenzo[c][1,2]oxaborole-6-carboxamide (DS109);
(2S,4S)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS110);
(2S,3S)-2-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-5-carboxamido)-3-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)butanoic acid (DS111); and
35 (2S,4R)-1-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carbonyl)-4-(1-hydroxy-1,3-dihydrobenzo[c][1,2]oxaborole-6-carboxamido)pyrrolidine-2-carboxylic acid (DS112).

53. The compound of any one of claims 1-44 and 50-52, wherein the compound is used as an intermediate in the manufacture of a drug substance or a therapeutic of a prophylactic compound.
54. A human insulin analog, comprising an A-chain and a B-chain, wherein the sequence of the A-chain comprises:
- 10 $X_{aa}'X_{bb}'X_{cc}'X_{dd}'X_{ee}'X_{ff}'X_{gg}'VEQCCX_{hh}'X_{ii}'ICSLYQLENYCNX_{jj}'X_{kk}'X_{ll}'X_{mm}'X_{nn}'X_{oo}'X_{pp}'$ (SEQ ID NO:24015); and
- wherein the sequence of the B-chain comprises:
- (i)
- $X_{aa}'X_{bb}'X_{cc}'X_{dd}'KX_{ee}'X_{ff}'X_{gg}'X_{hh}'X_{ii}'X_{jj}'KX_{kk}'X_{ll}'X_{mm}'X_{nn}'QHLGSHLVEALYLVCX_{oo}'X_{pp}'X_{qq}'GFFYT$
- 15 $X_{rr}'X_{ss}'X_{tt}'X_{uu}'X_{vv}'X_{ww}$ (SEQ ID NO:24016),
- wherein X_{aa}' , X_{bb}' , X_{cc}' , X_{dd}' , X_{ee}' , X_{ff}' , X_{gg}' , X_{hh}' , X_{ii}' , X_{jj}' , X_{kk}' , X_{ll}' , X_{mm}' , X_{nn}' , X_{oo}' , X_{pp}' , X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y and W,
- 20 (ii)
- $X_{aa}'X_{bb}'X_{cc}'X_{dd}'KPX_{ee}'X_{ff}'X_{gg}'X_{hh}'X_{ii}'X_{jj}'X_{kk}'X_{ll}'X_{mm}'X_{nn}'QHLGSHLVEALYLVCX_{oo}'X_{pp}'X_{qq}'GFFYT$
- $X_{rr}'X_{ss}'X_{tt}'X_{uu}'X_{vv}'X_{ww}$ (SEQ ID NO:24017),
- wherein X_{aa}' , X_{bb}' , X_{cc}' , X_{dd}' , X_{ee}' , X_{ff}' , X_{gg}' , X_{hh}' , X_{ii}' , X_{jj}' , X_{kk}' , X_{ll}' , X_{mm}' , X_{nn}' , X_{oo}' , X_{pp}' , X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y and W, and wherein X_{ee} is selected from amino acid residues A, E, F, H, I, K, L, N, P, Q, R, S, T, V, Y and W,
- 25 (iii)
- $X_{aa}'X_{bb}'X_{cc}'X_{dd}'KX_{ee}'X_{ff}'X_{gg}'X_{hh}'X_{ii}'X_{jj}'KX_{kk}'X_{ll}'X_{mm}'X_{nn}'QHLGSHLVEALYLVCX_{oo}'X_{pp}'X_{qq}'GFFYT$
- 30 $X_{rr}'X_{ss}'X_{tt}'X_{uu}'X_{vv}'X_{ww}$ (SEQ ID NO:24018),
- wherein X_{aa}' , X_{bb}' , X_{cc}' , X_{dd}' , X_{ee}' , X_{ff}' , X_{gg}' , X_{hh}' , X_{ii}' , X_{jj}' , X_{kk}' , X_{ll}' , X_{mm}' , X_{nn}' , X_{oo}' , X_{pp}' , X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} , X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is
- 35 present and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is G,

5 (iv)

$X_{aa}X_{bb}X_{cc}X_{dd}KX_{ee}X_{ff}X_{gg}X_{hh}X_{ii}X_{jj}KX_{kk}X_{ll}X_{mm}X_{nn}QHLGSHLVEALYLVCX_{oo}X_{pp}X_{qq}GFFYT$
 $X_{rr}X_{ss}X_{tt}X_{uu}X_{vv}X_{ww}$ (SEQ ID NO:24019),

wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$,

$X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} ,

10 X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is present and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is S, or

(v)

$X_{aa}X_{bb}X_{cc}X_{dd}KX_{ee}X_{ff}X_{gg}X_{hh}X_{ii}X_{jj}KX_{kk}X_{ll}X_{mm}X_{nn}QHLGSHLVEALYLVCX_{oo}X_{pp}X_{qq}GFFYT$

15 $X_{rr}X_{ss}X_{tt}X_{uu}X_{vv}X_{ww}$ (SEQ ID NO:24020),

wherein $X_{aa'}$, $X_{bb'}$, $X_{cc'}$, $X_{dd'}$, $X_{ee'}$, $X_{ff'}$, $X_{gg'}$, $X_{hh'}$, $X_{ii'}$, $X_{jj'}$, $X_{kk'}$, $X_{ll'}$, $X_{mm'}$, $X_{nn'}$, $X_{oo'}$,

$X_{pp'}$, X_{aa} , X_{bb} , X_{cc} , X_{dd} , X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} , X_{kk} , X_{ll} , X_{mm} , X_{nn} , X_{oo} , X_{pp} , X_{qq} , X_{rr} , X_{ss} , X_{tt} ,

X_{uu} , X_{vv} , and X_{ww} are each independently either absent or selected from amino acid residues A, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, Y, W and at least two of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} are

20 present and at least one of X_{ee} , X_{ff} , X_{gg} , X_{hh} , X_{ii} , X_{jj} is S, and another is G.

55. The insulin of claim 54, wherein the A-chain comprises a sequence selected from SEQ ID NOs 1 and 3 to 33, and is optionally appended at the N-terminus and/or at the C-

terminus by at least one selected from KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ,

KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR,

25 KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR,

KDS, KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER,

KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS,

KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS,

KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS,

30 KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY,

KL A, KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY,

KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY,

KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA,

KQD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA,

35 KRD, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA,

KSD, KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD,

- 5 KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75 to 24014, KGSH (SEQ ID NO:24049), GKGS (SEQ ID NO:24050), GKGSKK (SEQ ID NO:24045), GKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP (SEQ ID NO:24043), and GSHKGS (SEQ ID NO:24042); and
- 10 wherein the B-chain comprises a sequence selected from SEQ ID NOs 2 and 34 to 74, 24047, and 24048, and is optionally appended at the N-terminus and/or at the C-terminus by at least one selected from KA, KD, KE, KF, KG, KH, KI, KL, KN, KP, KQ, KR, KS, KT, KY, KAA, KAD, KAE, KAF, KAG, KAH, KAI, KAL, KAN, KAQ, KAR, KAS, KAT, KAY, KDA, KDD, KDE, KDF, KDG, KDH, KDI, KDL, KDN, KDQ, KDR, KDS,
- 15 KDT, KDY, KEA, KED, KEE, KEF, KEG, KEH, KEI, KEL, KEN, KEQ, KER, KES, KET, KEY, KFA, KFD, KFE, KFF, KFG, KFH, KFI, KFL, KFN, KFQ, KFR, KFS, KFT, KFY, KGA, KGD, KGE, KGF, KGG, KGH, KGI, KGL, KGN, KGQ, KGR, KGS, KGT, KGY, KHA, KHD, KHE, KHF, KHG, KHH, KHI, KHL, KHN, KHQ, KHR, KHS, KHT, KHY, KIA, KID, KIE, KIF, KIG, KIH, KII, KIL, KIN, KIQ, KIR, KIS, KIT, KIY, KLA,
- 20 KLD, KLE, KLF, KLG, KLH, KLI, KLL, KLN, KLQ, KLR, KLS, KLT, KLY, KNA, KND, KNE, KNF, KNG, KNH, KNI, KNL, KNN, KNQ, KNR, KNS, KNT, KNY, KPA, KPD, KPE, KPF, KPG, KPH, KPI, KPL, KPN, KPQ, KPR, KPS, KPT, KPY, KQA, QKD, KQE, KQF, KQG, KQH, KQI, KQL, KQN, KQQ, KQR, KQS, KQT, KQY, KRA, KRQ, KRE, KRF, KRG, KRH, KRI, KRL, KRN, KRQ, KRR, KRS, KRT, KRY, KSA, KSD,
- 25 KSE, KSF, KSG, KSH, KSI, KSL, KSN, KSQ, KSR, KSS, KST, KSY, KTA, KTD, KTE, KTF, KTG, KTH, KTI, KTL, KTN, KTQ, KTR, KTS, KTT, KTY, KYA, KYD, KYE, KYF, KYG, KYH, KYI, KYL, KYN, KYQ, KYR, KYS, KYT, KYY, SEQ ID NOs 75 to 24014, KGSH (SEQ ID NO:24049), GKGS (SEQ ID NO:24050), GKGSKK (SEQ ID NO:24045), GKPGKK (SEQ ID NO:24046), GKGPSK (SEQ ID NO:24044), GKPSHKP
- 30 (SEQ ID NO:24043), and GSHKGS (SEQ ID NO:24042).
56. The insulin of claim 54, wherein no more than 4 residues are added or deleted from the A-chain and/or the B-chain.
57. The insulin of claim 54, wherein a K residue is present at the N-terminus of the A-chain and/or the B-chain, and/or
- 35 wherein no more than three K residues are present at the N-terminus of the A-chain and/or the B-chain, and/or

5 wherein (i) the tyrosine at A14 is replaced with glutamic acid, and/or (ii) the tyrosine at B16 is replaced with histidine, and/or (iii) the phenylalanine at B25 is replaced with a histidine, and/or

 wherein one to three residues selected from residues B20, B21, and B22-B29 of the B-chain, residues A4 or A8 of the A-chain, and residues of an optionally extended
10 polypeptide, are lysine residues, and/or

 wherein only one K residue is present within 10 residues of the N-terminus of B-chain.
58. The compound of any one of claims 1-44 and 50-52, wherein X1 comprises the insulin of any one of claims 54-57.

59. The insulin of any one of claims 54-57, wherein an amino group of the side chain(s) of
15 one to four lysine residues is each independently covalently conjugated as described by Formula I of claim 2.

60. The insulin of any one of claims 54-57, wherein the insulin is covalently conjugated as described by Formula I of claim 2,

 n'=0 and the C-terminus of Z1a is directly conjugated to the N-terminus of the B-chain
20 of insulin through a peptide bond;

 Z1a comprises at least one amino acid selected from K, P, E, G, S, T, A, and R, such that the sequence comprises at least one lysine, at least one proline, and at least one amino acid selected from H, R, A and T; and

 the amino group of least one lysine side chain in Z1a is covalently conjugated as
25 described by Formula I.

61. The insulin of any one of claims 54-57, wherein the insulin is covalently conjugated as described by Formula I of claim 2,

 Z1a comprises a polypeptide comprising the sequence (XA₁A₂A₃X)_m (SEQ ID NO:24022), wherein:

30 A₁, A₂, and A₃ are each independently an L- or D-amino acid;

 m is an integer in the range of 1 to 4;

 each X is K or KP; and

 the epsilon amine group of at least one lysine side chain in Z1a is covalently conjugated as described by Formula I.

35 62. The insulin of any one of claims 54-57, wherein the insulin is covalently conjugated as described by Formula I of claim 2,

- 5 Z1a comprises a polypeptide comprising a sequence selected from
(XA1X)_m(GGGGS)_n (SEQ ID NO:24023), (XA1A2X)_m (GGGGS)_n (SEQ ID NO:24024),
(XA1A2A3X)_m(GGGGS)_n (SEQ ID NO:24025), (XA1X)_m(GGGGS)_n (XA2X)_o (SEQ ID
NO:24026), and (XA1A2X)_m(GGGGS)_n (XA3A4X)_o (SEQ ID NO:24027), wherein:
- A₁, A₂, A₃, and A₄ are each independently an L- or D-amino acid;
- 10 m is an integer in the range of 1 to 4;
 n is an integer in the range of 1 to 4;
 o is an integer in the range of 1 to 4;
 each X is K or KP; and
 the epsilon amine group of each lysine side chain of at least one lysine side chain in Z1a
15 is further covalently conjugated as described by Formula I.
63. The insulin of any one of claims 54-57, wherein the insulin is covalently conjugated as
described by Formula I of claim 2,
- Z1a comprises a polypeptide comprising the sequence (GX)_m, wherein:
- X is KV;
- 20 m is an integer in the range of 1 to 4, and
 the epsilon amine group of at least one lysine side chain in Z1a is further
covalently conjugated as described by Formula I.
64. The insulin of any one of claims 54-57, wherein the insulin is covalently conjugated as
described by Formula I of claim 2,
- 25 Z1a comprises a polypeptide comprising a sequence selected from:
(GXA1KGEA2XT)_m(GGSGSSS)_n (GXGXA3GSSSGSSSXT)_o (SEQ ID NO:24028),
(GXA1ESA2LYL)_m (SEQ ID NO:24029), (TXEX)_m(GPGS)_n (SEQ ID NO:24030),
(GXESA1VA)_m (KA2K)_n (SEQ ID NO:24031), (GXEA1A2)_m(GGS)_n (TYA3XXT)_o (SEQ
ID NO:24032), and (TXAXYT)_m(TSSS)_n (SEQ ID NO:24033), wherein:
- 30 each X is KV or KP;
 A₁, A₂, A₃ are each independently an L- or D-amino acid;
 m is an integer in the range of 1 to 4;
 n is an integer in the range of 1 to 4; and
 o is an integer in the range of 1 to 4; and

- 5 the epsilon amine group of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.
65. The insulin of any one of claims 54-57, wherein the insulin is covalently conjugated as described by Formula I of claim 2,
- Z1a comprises a polypeptide comprising a sequence selected from
- 10 (TKPYA1KEVETA2GSGS)_m (GGGS)_n (SEQ ID NO:24034),
(YTPLEA1KPYSTSYKPYSEA1L)_m(GKPTSLEA2FLVEA2LYTKP)_n (SEQ ID NO:24035),
and
(GKEALYLTPLESALYKP)_m(TKPLEALYLKPEILSLKPESLA)_n(GKPGSSSKPDTSSSGTP
KTAAGS)_o (SEQ ID NO:24036), wherein:
- 15 A₁ and A₂ are each independently an L- or D-amino acid;
 m is an integer in the range of 1 to 4;
 n is an integer in the range of 1 to 4; and
 the epsilon amine group of at least one lysine side chain in Z1a is further covalently conjugated as described by Formula I.
- 20 66. The insulin of claim 54, wherein (i) the A- and/or B-chain sequence of the insulin is appended at the N-terminus or C-terminus by KX'K, KX', or X'K wherein X' represents a continuous sequence of 2, 3, 4, or 5 residues selected from within wild-type A-chain (SEQ ID NO:1) and wild-type B-chain (SEQ ID NO:2), or
- (ii) wherein X' is a polypeptide of up to 30 residues with amino acids independently
- 25 selected from: K, G, S, E, H, E, N, Q, D, A, P, R and C, and
- wherein in (i) and (ii) each K residue is optionally and independently covalently conjugated as described by Formula I of claim 2.