

TYROSINASE-BASED CONJUGATE COMPOSITIONS AND METHODS OF USE**CROSS-REFERENCE**

[0001] This application claims the benefit of and priority to U.S. Provisional Patent Application No. 63/610,252, filed December 14, 2023, the disclosure of which is hereby incorporated by reference in their entirety for all purposes.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on December 12, 2024, is named CEB-105WO_SL.xml and is 86,114 bytes in size.

BACKGROUND

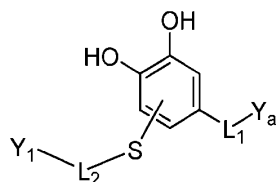
[0003] Coupling biomolecules (e.g., polypeptides such as viral proteins) to payloads to generate conjugates, while preserving the function of the biomolecule, has long been a goal of chemical biology and biopharmaceutical research. Examples of conjugates include protein-peptide conjugates for vaccine development, antibody-drug conjugates, and antibody-protein conjugates for immunotherapies.

[0004] While many techniques have been developed to allow for attachment of payloads to polypeptides, it has been challenging to develop methods for attaching payloads in a site-specific manner to any position on a polypeptide's surface. There is a need for improved conjugation procedures that can modify a polypeptide or biomolecule in a simple yet site specific manner.

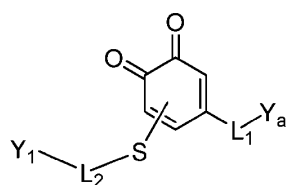
SUMMARY

[0005] The present disclosure provides for chemoselective modification of a polypeptide or biomolecule.

[0006] Described herein are conjugates of Formula E or Formula F:



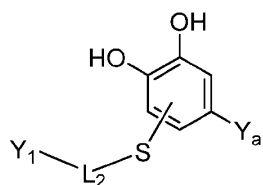
Formula E



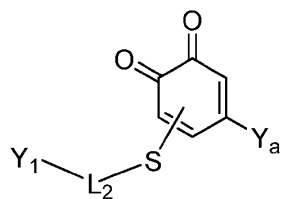
Formula F

wherein: Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof; L_1 is an optional first linker; L_2 is an optional second linker; and Y_1 is a payload.

[0007] In some embodiments, the conjugate comprises Formula E-A or Formula F-A:



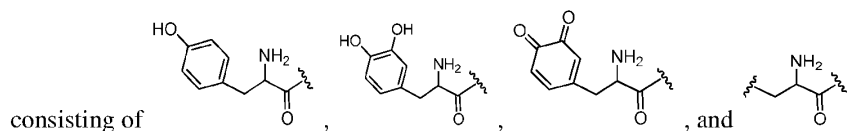
Formula E-A



Formula F-A

wherein: Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof; L_2 is an optional second linker; and Y_1 is a payload.

[0008] In some embodiments, the tyrosine or a portion thereof is selected from the group



[0009] In some embodiments, the non-terminal tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K).

[0010] In some embodiments, the non-terminal tag comprises $X_1X_2YX_4X_5$.

[0011] In some embodiments, X_1 and X_2 are proton-donating amino acids.

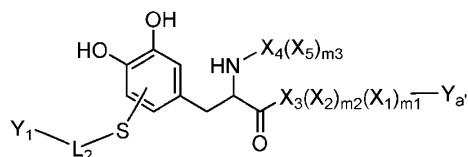
[0012] In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

[0013] In some embodiments, X_4 and X_5 are small, neutral amino acids.

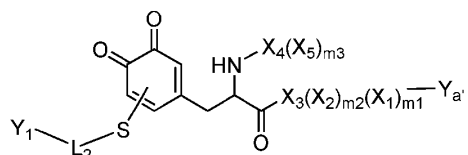
[0014] In some embodiments, X_4 is A or G.

[0015] In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L.

[0016] Described herein are conjugates of Formula E-I or Formula F-I:



Formula E-I



Formula F-I

wherein: $Y_{a'}$ is a polypeptide, each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a non-terminal tag in the polypeptide, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0; L_2 is an optional linker; and Y_1 is a payload

[0017] In some embodiments, X_2 and X_3 are proton-donating amino acids.

[0018] In some embodiments, X_2 and X_3 are each selected from the group consisting E, H, Q, D, and N.

[0019] In some embodiments, X_4 and X_5 are small, neutral amino acids.

[0020] In some embodiments, X_4 is A or G and X_5 is E, V, D, A, I, G, S, T, or L.

[0021] In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.

[0022] In some embodiments, the non-terminal tag comprises the sequence VEDYAIG, ANYAA, AEYAA, AQYAA or NYA.

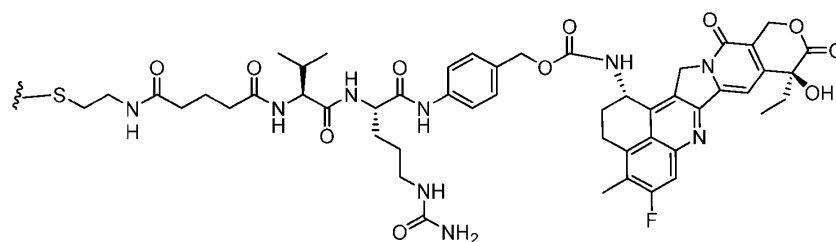
[0023] In some embodiments, the non-terminal tag comprises at least one of GGGGY, RGGGY, RGRGY, RRRGY, RRRRY, EGGGY, EGEGY, EEEGY, EEEEY, GGGWY, GGWGY, RRRWY, RRWRY, EEEWY, EEWY, DDDDY, SGGY, SGY, KKKKY, RRKKY, RRRKY, EDEDY, EDDDY, EEDDY, RRKKY, KKGGY, and DDGGY.

[0024] In some embodiments, the non-terminal tag comprises SGGY or SGY.

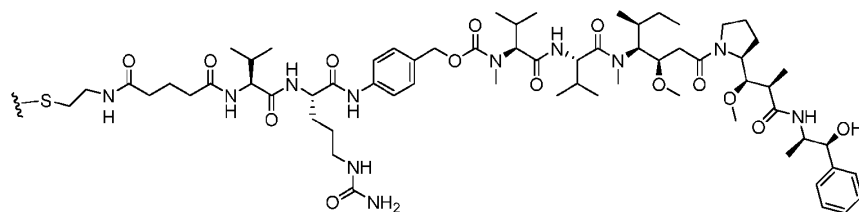
[0025] In some embodiments, the non-terminal tag comprises EEEY.

[0026] In some embodiments, Y_a is glycosylated.

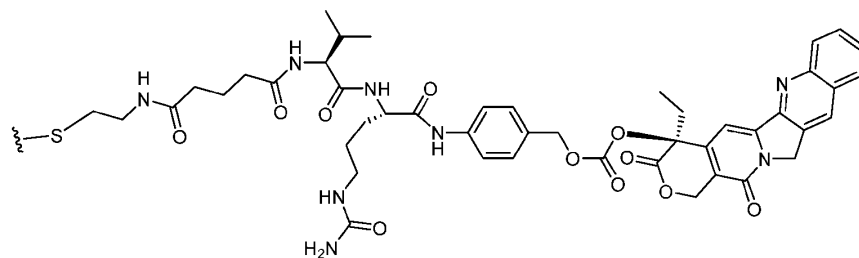
- [0027] In some embodiments, Y_a comprises a non-terminal tyrosine.
- [0028] In some embodiments, at least one of L₁ or L₂ is a cleavable linker.
- [0029] In some embodiments, the cleavable linker is an electrophilically cleavable linker, a nucleophilically cleavable linker, a photocleavable linker, a metal cleavable linker, an electrolytically-cleavable linker, an acid cleavable linker, or a proteolytically cleavable linker.
- [0030] In some embodiments, the cleavable linker further includes a pegylated group, a sugar group, or a modification that increases hydrophilicity.
- [0031] In some embodiments, the cleavable linker is cleavable under reductive and/or oxidative conditions.
- [0032] In some embodiments, the cleavable linker is cleavable under acidic conditions.
- [0033] In some embodiments, the cleavable linker comprises a disulfide bond.
- [0034] In some embodiments, the cleavable linker is a proteolytically cleavable linker and comprises a protease recognition sequence.
- [0035] In some embodiments, the protease recognition sequence is recognized by a protease selected from the group comprising a metalloprotease, cathepsin B, and tobacco etch virus (TEV).
- [0036] In some embodiments, the cleavable linker comprises a dipeptide or tetrapeptide.
- [0037] In some embodiments, the dipeptide is a valine-citrulline (Val-Cit) dipeptide, a valine-lysine dipeptide, and a valine-alanine dipeptide.
- [0038] In some embodiments, the tetrapeptide is a glycine-glycine-phenylalanine-glycine (GGFG) tetrapeptide.
- [0039] In some embodiments, the cleavable linker is selected from the group comprising PABC (p-aminobenzyl alcohol), glucuronide, and MABC (m-aminobenzyl alcohol).
- [0040] In some embodiments, the cleavable linker is Val-Cit-PABC.
- [0041] In some embodiments, “Y₁-L₂-S-” is selected from the group consisting of:



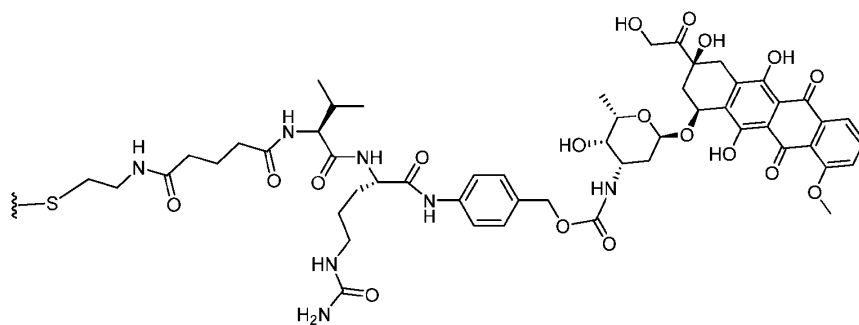
S-VC-PAB-DXD,



S-VC-PAB-MMAE,



S-VC-PAB-Camptothecin, and



S-VC-PAB-Doxorubicin.

[0042] In some embodiments, Y_a is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, a constant chain of an antibody or antibody fragment, a peptide, a fluorescent protein, or a cyclic peptide.

[0043] In some embodiments, Y_a is a viral protein.

[0044] In some embodiments, Y_a is a cell surface receptor.

[0045] In some embodiments, Y_a is a binding peptide.

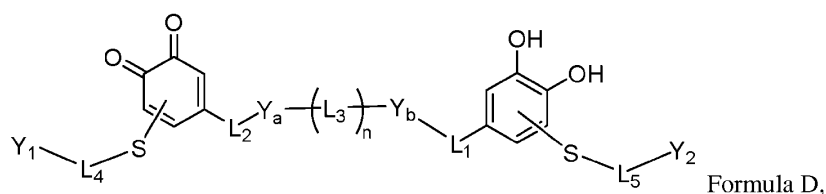
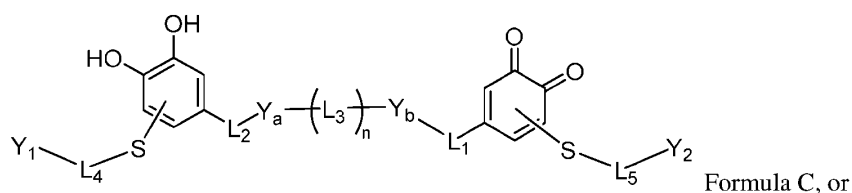
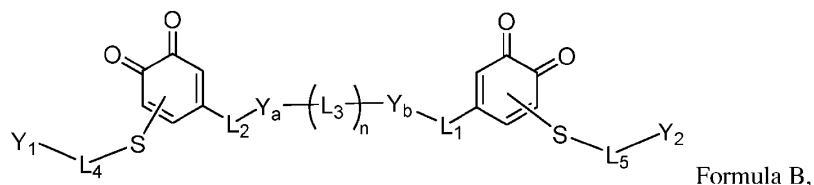
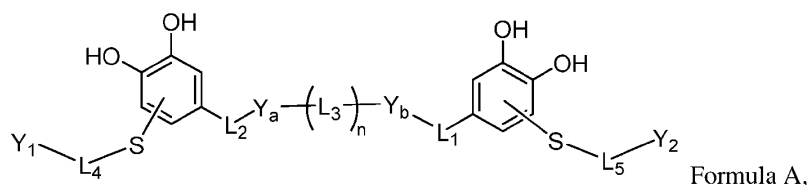
[0046] In some embodiments, Y_1 is a small molecule.

[0047] In some embodiments, the small molecule is selected from the group comprising deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan, lurototecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimmitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4),

pyrolodiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT.

[0048] In some embodiments, Y_1 comprises a nucleic acid, an immune agonist, a peptide, a cytokine, a protein, or a binding domain. In some embodiments, Y_1 comprises a protein. In some embodiments, Y_1 comprises a peptide. In some embodiments, Y_1 comprises a binding peptide.

[0049] Described herein are conjugates of Formula A, Formula B, Formula C, or Formula D:



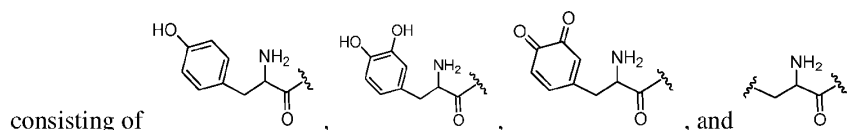
wherein: Y_a is a first polypeptide comprising a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

Y_b is a second polypeptide comprising a second tag different from the first tag, the second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} comprises a tyrosine or a portion

thereof, wherein m_2 is an integer greater or equal to 0, wherein Y_b is linked to "S" via the tyrosine or a portion thereof; n is an integer greater than 0; L_1 is an optional first linker; L_2 is an optional second linker; L_3 is an optional third linker; L_4 is an optional fourth linker; L_5 is an optional fifth linker;

Y_1 is a first payload; and Y_2 is a second payload; and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag.

[0050] In some embodiments, the tyrosine or a portion thereof is selected from the group



[0051] In some embodiments, the non-terminal tag comprises $X_1X_2YX_4X_5$.

[0052] In some embodiments, X_1 and X_2 are proton-donating amino acids. In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

[0053] In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G. In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L.

[0054] In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.

[0055] In some embodiments, the non-terminal tag comprises the sequence VEDYAIG, ANYAA, AEYAA, AQYAA or NYA.

[0056] In some embodiments, the first tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K).

[0057] In some embodiments, the second tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K).

[0058] In some embodiments, the first tag comprises at least one of GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGGGY (SEQ ID NO: 13), EEEGY (SEQ ID NO: 14), EEEEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWEY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 30), KKGGY (SEQ ID NO: 31), and DDGGY (SEQ ID NO: 32).

- [0059] In some embodiments, the first tag comprises SGGY (SEQ ID NO: 23) or SGY and the second tag comprises EEEY (SEQ ID NO: 33).
- [0060] In some embodiments, the first tag is a non-terminal tag. In some embodiments, the first tag is a terminal tag. In some embodiments, the second tag is a terminal tag. In some embodiments, the second tag is a non-terminal tag. In some embodiments, the first tag and the second tag are non-terminal tags.
- [0061] In some embodiments, Y_a, Y_b, or both Y_a and Y_b are glycosylated.
- [0062] In some embodiments, Y_a, Y_b, or both Y_a and Y_b comprise a non-terminal tyrosine.
- [0063] In some embodiments, at least one of L₁, L₂, L₃, L₄, or L₅ is a cleavable linker.
- [0064] In some embodiments, the cleavable linker is an electrophilically cleavable linker, a nucleophilically cleavable linker, a photocleavable linker, a metal cleavable linker, an electrolytically-cleavable linker, an acid cleavable linker, or a proteolytically cleavable linker.
- [0065] In some embodiments, the cleavable linker further includes a pegylated group, a sugar group, or a modification that increases hydrophilicity.
- [0066] In some embodiments, the cleavable linker is cleavable under reductive and/or oxidative conditions.
- [0067] In some embodiments, the cleavable linker is cleavable under acidic conditions.
- [0068] In some embodiments, the cleavable linker comprises a disulfide bond.
- [0069] In some embodiments, the cleavable linker is a proteolytically cleavable linker and comprises a protease recognition sequence.
- [0070] In some embodiments, the protease recognition sequence is recognized by a protease selected from the group comprising a metalloprotease, cathepsin B, and tobacco etch virus (TEV).
- [0071] In some embodiments, the cleavable linker comprises a dipeptide or tetrapeptide
- [0072] In some embodiments, the dipeptide is a valine-citrulline (Val-Cit) dipeptide, a valine-lysine dipeptide, or a valine-alanine dipeptide.
- [0073] In some embodiments, the tetrapeptide is a glycine-glycine-phenylalanine-glycine (GGFG) tetrapeptide (SEQ ID NO: 75).
- [0074] In some embodiments, the cleavable linker is selected from the group comprising PABC (p-aminobenzyl alcohol), glucuronide, and MABC (m-aminobenzyl alcohol).
- [0075] In some embodiments, Y_a is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, a constant chain of an antibody or antibody fragment, a peptide, or a cyclic peptide.

- [0076] In some embodiments, Y_b is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, or a constant chain of an antibody, antibody fragment, a peptide, or a cyclic peptide.
- [0077] In some embodiments, Y_a or Y_b is a viral protein.
- [0078] In some embodiments, Y_a and Y_b are viral proteins.
- [0079] In some embodiments, Y_a or Y_b is a cell surface receptor.
- [0080] In some embodiments, Y_a and Y_b are attached via L_3 .
- [0081] In some embodiments, L_3 comprises a peptide sequence, a dimerization and docking domain, a leucine zipper, or knobs-into-holes.
- [0082] In some embodiments, L_3 comprises a peptide bond, a disulfide bond, a maleimide bond, thioether bond, an azide-alkyne cycloaddition, a cysteinyl-dopa, or a hydrogen bond.
- [0083] In some embodiments, L_3 is a linker.
- [0084] In some embodiments, the linker comprises a sequence selected from the group consisting of (GS) n_3 (SEQ ID NO: 66), (G2S) n_3 (SEQ ID NO: 67), (G3S) n_3 (SEQ ID NO: 68), (G4S) n_3 (SEQ ID NO: 69), (G) n_3 (SEQ ID NO: 70), (GGSGGD) n_3 (SEQ ID NO: 71), (GGSGGE) n_3 (SEQ ID NO: 72), (GGGSGSGGGGS) n_3 (SEQ ID NO: 73), and (GGGGPGGGGP) n_3 (SEQ ID NO: 74) and wherein n_3 is an integer from 2 to 20.
- [0085] In some embodiments, L_3 comprises a terminal or non-terminal tyrosine.
- [0086] In some embodiments, n is 1, 2, 3, 4, or 5. In some embodiments, n is 1.
- [0087] In some embodiments, Y_1 and Y_2 are the same.
- [0088] In some embodiments, Y_1 and Y_2 are different.
- [0089] In some embodiments, Y_1 , Y_2 , or both Y_1 and Y_2 is a small molecule.
- [0090] In some embodiments, the small molecule is selected from the group comprising deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan, lurototecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimmitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4), pyrrolidiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT.
- [0091] In some embodiments, Y_1 and Y_2 , or both Y_1 and Y_2 comprises a nucleic acid, an immune agonist, a peptide, a cytokine, or a binding domain.
- [0092] In some embodiments, Y_1 , Y_2 , or both Y_1 and Y_2 comprises a nucleic acid.
- [0093] In some embodiments, Y_1 and Y_2 , or both Y_1 and Y_2 comprises a peptide.

[0094] Described herein are polypeptides comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0.

[0095] In some embodiments, the non-terminal tag comprises $X_1X_2YX_4X_5$.

[0096] In some embodiments, X_1 and X_2 are proton-donating amino acids.

[0097] In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

[0098] In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G. In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L.

[0099] In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.

[0100] In some embodiments, the non-terminal tag comprises the sequence VEDYAIG, ANYAA, AEYAA, AQYAA or NYAA.

[0101] Described herein are compositions of a first polypeptide and a second polypeptide, wherein the first polypeptide comprises a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof; and wherein $m1$ is an integer greater than or equal to 0, and the second polypeptide comprises a second non-terminal tag different from the first terminal tag, the second terminal tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine or a portion thereof, and wherein $m2$ is an integer greater than or equal to 0.

[0102] In some embodiments, the first polypeptide and/or the second polypeptide comprise a non-terminal tyrosine.

[0103] In some embodiments, the non-terminal tag comprises $X_1X_2YX_4X_5$.

[0104] In some embodiments, X_1 and X_2 are proton-donating amino acids. In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

[0105] In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G. In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L.

[0106] In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.

[0107] In some embodiments, the non-terminal tag comprises the sequence VEDYAIG, ANYAA, AEYAA, AQYAA or NYA.

[00108] Described herein are methods of covalently linking at least two polypeptides to at least two payloads, the method comprising:

a) contacting a first polypeptide of the at least two polypeptides and a first payload of the at least two payloads using a first tyrosinase, wherein the first polypeptide comprises a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0; and

b) contacting a second polypeptide of the at least two polypeptides and a second payload of the at least two payloads using a second tyrosinase, wherein the second polypeptide comprises a second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, wherein $m2$ is an integer greater or equal to 0, and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag.

[00109] In some embodiments, the first tag and the second tag are non-terminal tags.

[00110] In some embodiments, the first tyrosinase, the second tyrosinase, or both the first tyrosinase and the second tyrosinase is *Agricus bisporus* tyrosinase (abTYR).

[00111] In some embodiments, the first tyrosinase, the second tyrosinase, or both comprises a sequence at least 80% identity to any one of SEQ ID NOs: 1 or 43

[00112] In some embodiments, the first tyrosinase, the second tyrosinase, or both the first tyrosinase and the second tyrosinase is Catenase.

[00113] In some embodiments, the first tyrosinase, the second tyrosinase, or both comprises a sequence at least 90% identity any one of SEQ ID NOs: 2-6 or 53.

[00114] In some embodiments, the first tyrosinase is *Agricus bisporus* tyrosinase (abTYR) and the second tyrosinase is Catenase.

[00115] In some embodiments, the first tyrosinase and the second tyrosinase are provided at a same time.

[00116] In some embodiments, the first tyrosinase is provided first followed by the second tyrosinase.

[00117] Described herein are methods of covalently linking a polypeptide to a payload, the method comprising:

contacting a polypeptide and a payload using a tyrosinase, wherein the polypeptide comprises a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0.

[00118] In some embodiments, the tyrosinase is *Agricus bisporus* tyrosinase (abTYR).

[00119] In some embodiments, the tyrosinase comprises a sequence at least 80% identity to any one of SEQ ID NOs: 2-6 or 53.

[00120] In some embodiments, the tyrosinase is a Catenase.

[00121] In some embodiments, the tyrosinase comprises a sequence at least 90% identity to any one of SEQ ID NOs: 2-6 or 53. In some embodiments, the tyrosinase comprises a sequence at least 90% identity to SEQ ID NO: 53.

[00122] Described herein in an engineered tyrosinase comprising at least 80% identity to SEQ ID NO: 53. In some embodiments, the engineered tyrosinase comprises at least 90% identity to SEQ ID NO: 53. In some embodiments, the engineered tyrosinase comprises 100% identity to SEQ ID NO: 53.

[00123] In some embodiments, the engineered tyrosinase has at least an 11-fold increase in conjugation efficiency as compared to AbTyr (SEQ ID NO:1). In some embodiments, the engineered tyrosinase has at least a 6-fold greater Vmax activity as compared to AbTyr (SEQ ID NO:1).

BRIEF DESCRIPTION OF THE DRAWINGS

[00124] **Figure (FIG.) 1** is an illustration of the different reaction sites accessible through Catenase conjugation, cysteine residues (Site 1), heavy and light chain termini (Sites 2 and 4), and native loop tyrosines (site 3-Y*).

[00125] FIG. 2 shows a protein structure of nanoluciferase (PDB: 8AQI (SEQ ID NO: 85)). The loop that was selected for incorporation of a tyrosine or non-terminal tag is identified in a box. The non-terminal tag or tyrosine was inserted after the aspartate residue in the sequence identified by an arrow.

[00126] FIG. 3A shows an SDS-Page gel of the purification of nanoluciferase C166S with non-terminal tag VEDYAIG (SEQ ID NO: 76). Feed is the initial loaded sample; Fx is fractions.

[00127] FIG. 3B shows an FPLC trace of the purification of nanoluciferase C166S with non-terminal tag VEDYAIG(SEQ ID NO: 76).

[00128] FIG. 3C shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag VEDYAIG (SEQ ID NO: 76) to Txn with Catenase K12 (Cat K12).

[00129] FIG. 3D shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag VEDYAIG (SEQ ID NO: 76) to Txn with abTYR (mushroom tyrosinase).

[00130] FIG. 4A shows an SDS-Page gel of the purification of nanoluciferase C166S with non-terminal tag ANYAA (SEQ ID NO: 50). Feed is the initial loaded sample; Fx is fractions.

[00131] FIG. 4B shows an FPLC trace of the purification of nanoluciferase C166S with non-terminal tag ANYAA (SEQ ID NO: 50).

[00132] FIG. 4C shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag ANYAA (SEQ ID NO: 50) to Txn with Catenase K12 (Cat K12).

[00133] FIG. 4D shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag ANYAA (SEQ ID NO: 50) to Txn with abTYR (mushroom tyrosinase).

[00134] FIG. 5A shows an SDS-Page gel of the purification of nanoluciferase C166S with non-terminal tag AEYAA (SEQ ID NO: 51). Feed is the initial loaded sample; Fx is fractions.

[00135] FIG. 5B shows an FPLC trace of the purification of nanoluciferase C166S with non-terminal tag AEYAA (SEQ ID NO: 51).

[00136] FIG. 5C shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag AEYAA (SEQ ID NO: 51) to Txn with Catenase K12 (Cat K12).

[00137] FIG. 5D shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag AEYAA (SEQ ID NO: 51) to Txn with abTYR (mushroom tyrosinase).

[00138] FIG. 6A shows an SDS-Page gel of the purification of nanoluciferase C166S with non-terminal tag AQYAA (SEQ ID NO: 52). Feed is the initial loaded sample; Fx is fractions.

[00139] FIG. 6B shows an FPLC trace of the purification of nanoluciferase C166S with non-terminal tag AQYAA (SEQ ID NO: 52).

[00140] FIG. 6C shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag AQYAA (SEQ ID NO: 52) to Txn with Catenase K12 (Cat K12).

[00141] FIG. 6D shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag AQYAA (SEQ ID NO: 52) to Txn with abTYR (mushroom tyrosinase).

[00142] FIG. 7A shows an SDS-Page gel of the purification of nanoluciferase C166S with non-terminal tag NYA. Feed is the initial loaded sample; Fx is fractions.

[00143] FIG. 7B shows an FPLC trace of the purification of nanoluciferase C166S with non-terminal tag NYA.

[00144] FIG. 7C shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag NYA to Txn with Catenase K12 (Cat K12).

[00145] FIG. 7D shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with non-terminal tag NYA to Txn with abTYR (mushroom tyrosinase).

[00146] FIG. 8A shows an SDS-Page gel of the purification of nanoluciferase C166S with a tyrosine in a loop. Feed is the initial loaded sample; Fx is fractions.

[00147] FIG. 8B shows an FPLC trace of the purification of nanoluciferase C166S with a tyrosine in a loop.

[00148] FIG. 8C shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with a tyrosine in a loop to Txn with Catenase K12 (Cat K12).

[00149] FIG. 8D shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with a tyrosine in a loop to Txn with abTYR (mushroom tyrosinase).

[00150] FIG. 8E shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with a C-terminal SGGGGY (SEQ ID NO: 77) tag to Txn with Catenase K12 (Cat K12).

[00151] FIG. 8F shows an SDS-Page gel of the conjugation reaction of nanoluciferase (nLuc) C166S with a C-terminal SGGGGY (SEQ ID NO: 77) tag to Txn with abTYR (mushroom tyrosinase).

[00152] FIG. 9 shows the conjugation efficiency of nanoluciferase with the indicated non-terminal tags, a tyrosine in the loop, or a C-terminal tag SG4Y (SEQ ID NO: 77) to Txn with either CatK12 or abTYR. FIG. 9 discloses SEQ ID NOS 76, 50, 51, 52 and 77, respectively, in order of appearance.

[00153] FIG. 10 shows an SDS-Page gel of a conjugation reaction of nanoluciferase with a C-terminal tag SG4Y (SEQ ID NO: 77) to txn with CatK12 using txn that has been stored at 4°C for three months or freshly purified txn.

[00154] FIG. 11 shows the conjugation efficiency of the reactions of FIG. 10 compared to the conjugation reaction efficiencies of nanoluciferase with the indicated non-terminal tags. FIG. 11 discloses SEQ ID NOS 76, 50, 51, 52, 76, 50, 51, 52, 77 and 77, respectively, in order of appearance.

[00155] FIG. 12A shows the quantification of the gel purity of nanoluciferase with the indicated non-terminal tags in a non-reducing or reducing SDS-page gel. . FIG. 12A discloses SEQ ID NOS 76, 52, 50 and 51, respectively, in order of appearance.

[00156] FIGs. 13A and 13B show the kinetic activity of abTYR or CatK12 on the indicated peptides after 15 minutes of incubation.

[00157] FIGs. 13C and 13D show the Vmax calculated from the kinetic activity readings of abTYR or CatK12 on the indicated peptides after 15 minutes of incubation. FIG. 13D discloses SEQ ID NOS 78-82, 45-47, 83 and 76, respectively, in order of appearance.

[00158] FIG. 13E shows the final absorbance after 24 hours of incubation of the respective tyrosinases with the indicated peptides.

[00159] FIG. 14 shows an NMR spectra of the synthesized SH-VC-PAB-DXD compound.

[00160] FIG. 15 shows an NMR spectra of the synthesized SH-VC-PAB-MMAE compound.

DETAILED DESCRIPTION

[00161] Conjugates such as antibody-drug conjugates are one of the fastest growing drugs for diseases and disorders such as cancer. This approach generally comprises a polypeptide (e.g., an antibody) conjugated to a cytotoxic payload via a chemical linker. Conjugates are complex molecules that require careful attention to various components. Selection of an appropriate target, the polypeptide, the cytotoxic payload, and the manner in which the polypeptide is linked to the payload are key determinants of the safety and efficacy of such conjugates. There is a need for improved conjugation procedures that can modify a polypeptide or biomolecule in a simple yet site specific manner.

[00162] Provided herein are compositions and methods for conjugation of a polypeptide or biomolecule to one or more payloads.

Definitions

[00163] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this present disclosure pertains. Further, unless otherwise required by context, singular terms shall include the plural and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of, biochemistry, enzymology, molecular and cellular biology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art.

[00164] Throughout this specification and claims, the word “comprise” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

[00165] Unless otherwise indicated, and as an example for all sequences described herein under the general format “SEQ ID NO:”, “nucleic acid comprising SEQ ID NO: 1” refers to a nucleic acid, at least a portion of which has either (i) the sequence of SEQ ID NO: 1, or (ii) a sequence complementary to SEQ ID NO: 1. The choice between the two is dictated by the context. For instance, if the nucleic acid is used as a probe, the choice between the two is dictated by the requirement that the probe be complementary to the desired target.

[00166] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the disclosure. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure.

[00167] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present disclosure, the preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

[00168] It must be noted that as used herein and in the appended claims, the singular forms “a,” “an,” and “the” include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to “a thiol group” includes a plurality of such thiol groups and reference to “the thiol group” includes reference to one or more thiol groups and equivalents thereof known to those skilled in the art, and so forth. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as “solely,” “only” and the like in connection with the recitation of claim elements, or use of a “negative” limitation.

[00169] It is appreciated that certain features of the disclosure, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the disclosure, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination. All combinations of the embodiments pertaining to the disclosure

are specifically embraced by the present disclosure and are disclosed herein just as if each and every combination was individually and explicitly disclosed. In addition, all sub-combinations of the various embodiments and elements thereof are also specifically embraced by the present disclosure and are disclosed herein just as if each and every such sub-combination was individually and explicitly disclosed herein.

[00170] The term “fluorophore” refers to any molecular entity that is capable of absorbing energy of a first wavelength and re-emit energy at a different second wavelength. In certain embodiments, the subject biomolecule includes a fluorophore attached to one end of the biomolecule or at a central position. In some embodiments, the fluorophore may be attached to one end of the biomolecule. The fluorophore attached to the biomolecule need not be a single molecule, but may include multiple molecules.

[00171] The fluorophore may be synthetic or biological in nature, as known to those of skill in the art. More generally, any fluorophore can be used that is stable under coupling conditions and that can be sufficiently suppressed when in close proximity to the quencher such that a significant change in the intensity of fluorescence of the fluorophore is detectable in response to target specifically binding the probe. Examples of suitable fluorophores include, but are not limited to Oregon Green 488 dye, rhodamine and rhodamine derivatives, fluorescein isothiocyanate, fluorescein, 6-carboxyfluorescein (6-FAM), coumarin and coumarin derivatives, cyanine and cyanine derivatives, Alexa Fluors, DyLight Fluors, and the like.

[00172] In certain embodiments, the biomolecule includes a metal-chelating agent. A “chelate” as used herein in reference to a complex between a metal and a chelating ligand, refers to a combination of a metallic ion bonded to one or more ligands to form a heterocyclic ring structure. Chelate formation through neutralization of the positive charge(s) of the metal ion may be through the formation of ionic, covalent or coordinate covalent bonding. In certain embodiments, the metal chelating agent includes, but are not limited to, 1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetraacetic acid (also referred to as, DOTA, or tetraxetan).

[00173] The terms “polynucleotide” and “nucleic acid,” used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. Thus, this term includes, but is not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer comprising purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases.

[00174] The terms “polypeptide” and “protein”, used interchangeably herein, refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones. . A “polypeptide” encompasses a chemically modified polypeptide. For example, the term “polypeptide” can include a monovalent or divalent radical of i) a chemical structure of a polypeptide or ii) a chemical structure of a polypeptide substituted with a chemical functional group. In some embodiments, a polypeptide is a polypeptide that does not include a non-terminal tag. The term “fusion protein” or grammatical equivalents thereof is meant to include a protein composed of a plurality of polypeptide components, that while typically unjoined in their native state, typically are joined by their respective amino and carboxyl termini through a peptide linkage to form a single continuous polypeptide. Fusion proteins may be a combination of two, three or even four or more different proteins.

[00175] In general, polypeptides may be of any length, e.g., 2 or greater amino acids, greater than 4 amino acids, greater than about 10 amino acids, greater than about 20 amino acids, greater than about 50 amino acids, greater than about 100 amino acids, greater than about 300 amino acids, usually up to about 500 or 1000 or more amino acids. “Peptides” are generally 2 or greater amino acids in length, such as greater than 4 amino acids, greater than about 10 amino acids, greater than about 20 amino acids, usually up to about 50 amino acids. In some embodiments, peptides are between 2 and 30 amino acids in length.

[00176] The terms “antibodies” and “immunoglobulin” include antibodies or immunoglobulins of any isotype, fragments of antibodies that retain specific binding to antigen, including, but not limited to, Fab, Fv, scFv, and Fd fragments, chimeric antibodies, humanized antibodies, single-chain antibodies (scAb), single domain antibodies (dAb), single domain heavy chain antibodies, a single domain light chain antibodies, nanobodies, bi-specific antibodies, multi-specific antibodies, and fusion proteins comprising an antigen-binding (also referred to herein as antigen binding) portion of an antibody and a non-antibody protein. The antibodies can be detectably labeled, e.g., with a radioisotope, an enzyme that generates a detectable product, a fluorescent protein, and the like. The antibodies can be further conjugated to other moieties, such as members of specific binding pairs, e.g., biotin (member of biotin-avidin specific binding pair), and the like. The antibodies can also be bound to a solid support, including, but not limited to, polystyrene plates or beads, and the like. Also encompassed by the term are Fab', Fv, F(ab')₂, and or other antibody fragments that retain specific binding to antigen, and monoclonal antibodies. As used herein, a

monoclonal antibody is an antibody produced by a group of identical cells, all of which were produced from a single cell by repetitive cellular replication. That is, the clone of cells only produces a single antibody species. While a monoclonal antibody can be produced using hybridoma production technology, other production methods known to those skilled in the art can also be used (e.g., antibodies derived from antibody phage display libraries). An antibody can be monovalent or bivalent. An antibody can be an Ig monomer, which is a “Y-shaped” molecule that consists of four polypeptide chains: two heavy chains and two light chains connected by disulfide bonds.

[00177] The term “humanized immunoglobulin” as used herein refers to an immunoglobulin comprising portions of immunoglobulins of different origin, wherein at least one portion comprises amino acid sequences of human origin. For example, the humanized antibody can comprise portions derived from an immunoglobulin of nonhuman origin with the requisite specificity, such as a mouse, and from immunoglobulin sequences of human origin (e.g., chimeric immunoglobulin), joined together chemically by conventional techniques (e.g., synthetic) or prepared as a contiguous polypeptide using genetic engineering techniques (e.g., DNA encoding the protein portions of the chimeric antibody can be expressed to produce a contiguous polypeptide chain). Another example of a humanized immunoglobulin is an immunoglobulin containing one or more immunoglobulin chains comprising a complementarity-determining region (CDR) derived from an antibody of nonhuman origin and a framework region derived from a light and/or heavy chain of human origin (e.g., CDR-grafted antibodies with or without framework changes). Chimeric or CDR-grafted single chain antibodies are also encompassed by the term humanized immunoglobulin. See, e.g., Cabilly et al., U.S. Pat. No. 4,816,567; Cabilly et al., European Patent No. 0,125,023 B1; Boss et al., U.S. Pat. No. 4,816,397; Boss et al., European Patent No. 0,120,694 B1; Neuberger, M. S. et al., WO 86/01533; Neuberger, M. S. et al., European patent No. 0,194,276 B1; Winter, U.S. Pat. No. 5,225,539; Winter, European Patent No. 0,239,400 B1; Padlan, E. A. et al., European Patent Application No. 0,519,596 A1. See also, Ladner et al., U.S. Pat. No. 4,946,778; Huston, U.S. Pat. No. 5,476,786; and Bird, R. E. et al., *Science*, 242: 423- 426 (1988)), regarding single chain antibodies.

[00178] The term “nanobody” (Nb), as used herein, refers to the smallest antigen binding fragment or single variable domain (VHH) derived from naturally occurring heavy chain antibody and is known to the person skilled in the art. They are derived from heavy chain only antibodies, seen in camelids (Hamers-Casterman et al., (1993) *Nature* 363:446; Desmyter et al., (1996) *Nature Struct. Biol.* 3:803). In the family of “camelids”

immunoglobulins devoid of light polypeptide chains are found. "Camelids" comprise old world camelids (*Camelus bactrianus* and *Camelus dromedarius*) and new world camelids (for example, *Llama paccos*, *Llama glama*, *Llama guanicoe* and *Llama vicugna*). A single variable domain heavy chain antibody is referred to herein as a nanobody or a VHH antibody.

[00179] "Antibody fragments" comprise a portion of an intact antibody, for example, the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057- 1062 (1995)); domain antibodies (dAb; Holt et al. (2003) Trends Biotechnol. 21:484); single- chain antibody molecules; and multi-specific antibodies formed from antibody fragments. Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen combining sites and is still capable of cross-linking antigen.

[00180] As used herein, the term "Fv" is the minimum antibody fragment that contains a complete antigen-recognition and – binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H-V_L dimer.

[00181] Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

[00182] The "Fab" fragment also contains the constant domain of the light chain and the first constant domain (Chi) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxyl terminus of the heavy chain Chi domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

[00183] The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains. Depending on the amino acid sequence of

the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these classes can be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2. The subclasses can be further divided into types, e.g., IgG2a and IgG2b.

[00184] As used herein, “Single-chain Fv” or “sFv” or “scFv” antibody fragments comprise the V_H and V_L domains of antibody, wherein these domains are present in a single polypeptide chain. In some embodiments, the Fv polypeptide further comprises a polypeptide linker between the V_H and V_L domains, which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg and Moore eds., Springer -Verlag, New York, pp. 269-315 (1994).

[00185] As used herein, the term “diabodies” refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) in the same polypeptide chain (V_H - V_L). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:6444-6448.

[00186] As used herein, the term “affinity” refers to the equilibrium constant for the reversible binding of two agents (e.g., an antibody and an antigen) and is expressed as a dissociation constant (K_D). Affinity can be at least 1-fold greater, at least 2-fold greater, at least 3-fold greater, at least 4-fold greater, at least 5-fold greater, at least 6-fold greater, at least 7-fold greater, at least 8-fold greater, at least 9-fold greater, at least 10-fold greater, at least 20-fold greater, at least 30-fold greater, at least 40-fold greater, at least 50-fold greater, at least 60-fold greater, at least 70-fold greater, at least 80-fold greater, at least 90-fold greater, at least 100-fold greater, or at least 1,000-fold greater, or more, than the affinity of an antibody for unrelated amino acid sequences. Affinity of an antibody to a target protein can be, for example, from about 100 nanomolar (nM) to about 0.1 nM, from about 100 nM to about 1 picomolar (pM), or from about 100 nM to about 1 femtomolar (fM) or more. As used herein, the term “avidity” refers to the resistance of a complex of two or more agents to dissociation after dilution. The terms “immunoreactive” and “preferentially binds” are used interchangeably herein with respect to antibodies and/or antigen binding fragments.

[00187] As used herein, the term “binding” refers to a direct association between two molecules, due to, for example, covalent, electrostatic, hydrophobic, and ionic and/or hydrogen-bond interactions, including interactions such as salt bridges and water bridges. “Specific binding” refers to binding with an affinity of at least about 10^7 M or greater, e.g., 5×10^7 M, 10^8 M, 5×10^8 M, and greater. “Nonspecific binding” refers to binding with an affinity of less than about 10^7 M, e.g., binding with an affinity of 10^6 M, 10^5 M, 10^4 M, etc.

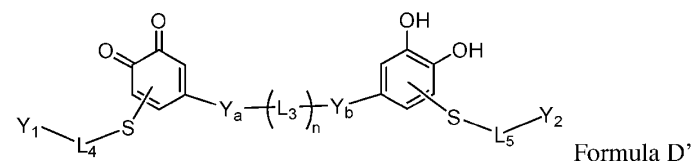
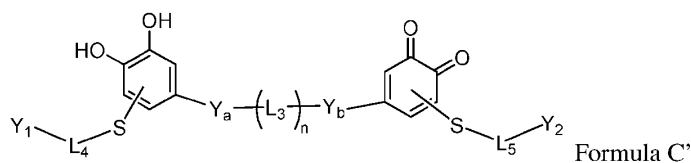
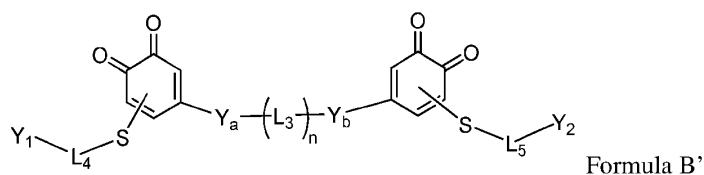
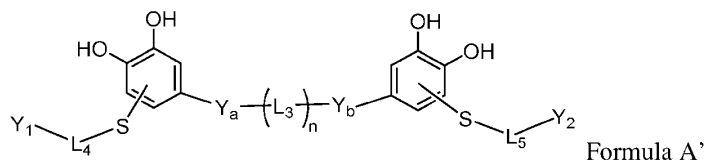
[00188] An “isolated” polypeptide is one that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In some embodiments, the polypeptide will be purified (1) to greater than 90%, greater than 95%, or greater than 98%, by weight of protein as determined by the Lowry method, for example, more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing or nonreducing conditions using Coomassie blue or silver stain. Isolated polypeptide includes the polypeptide in situ within recombinant cells since at least one component of the polypeptide’s natural environment will not be present. In some instances, isolated polypeptide will be prepared by at least one purification step.

[00189] As will be apparent to those of skill in the art upon reading this disclosure, each of the individual embodiments described and illustrated herein has discrete components and features which may be readily separated from or combined with the features of any of the other several embodiments without departing from the scope or spirit of the present disclosure. Any recited method can be carried out in the order of events recited or in any other order which is logically possible.

[00190] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the disclosure. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure.

Polypeptides of the Disclosure

[00191] Described herein, in certain embodiments, are conjugates of Formula (A), Formula (B), Formula (C), or Formula (D):



wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine or a portion thereof, wherein Y_a is linked to "S" through the tyrosine or a portion thereof ;

Y_b is a second polypeptide comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine or a portion thereof, and wherein $m2$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine or a portion thereof, wherein Y_b is linked to "S" through the tyrosine or a portion thereof;

n is an integer greater than 0;

L_3 is an optional third linker;

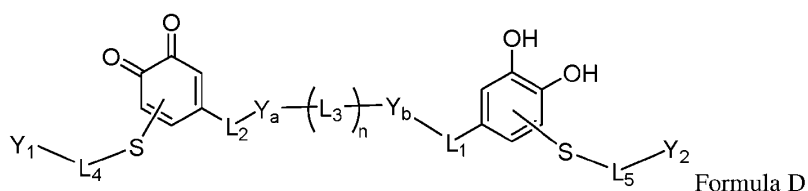
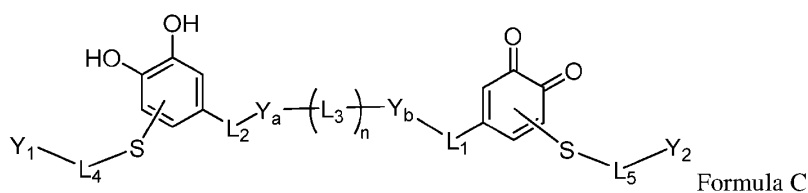
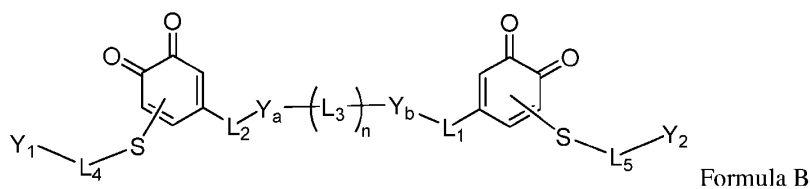
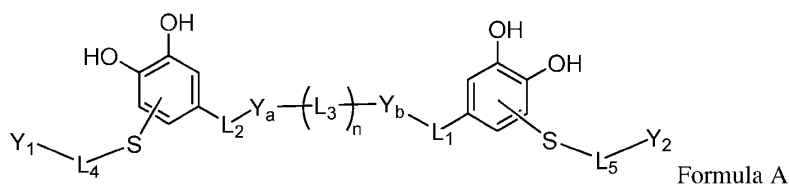
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

[00192] Described herein, in certain embodiments, are conjugates of Formula (A), Formula (B), Formula (C), or Formula (D):



wherein:

Y_a is a first polypeptide comprising a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0;

Y_b is a second polypeptide comprising a second tag different from the first tag, the second comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least

one amino acid of X_6 - X_{10} is a tyrosine, and wherein m_2 is an integer greater or equal to 0; and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag;

n is an integer greater than 0;

L_1 is an optional first linker;

L_2 is an optional second linker;

L_3 is an optional third linker;

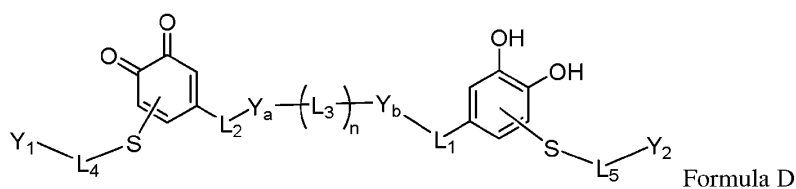
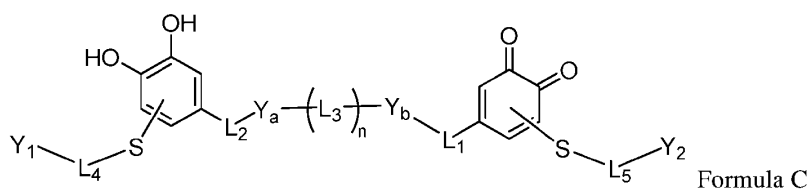
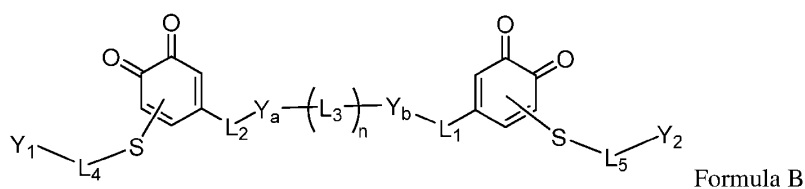
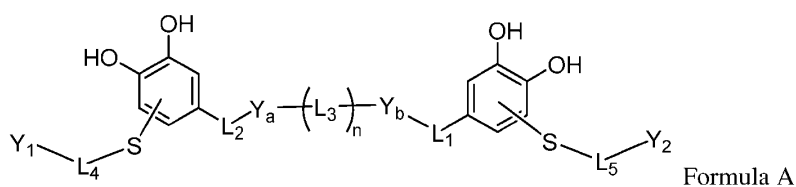
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

[00193] Described herein, in certain embodiments, are conjugates of Formula (A), Formula (B), Formula (C), or Formula (D):



wherein:

Y_a is a first biomolecule comprising a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0;

Y_b is a second biomolecule comprising a second tag different from the first tag, the second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, and wherein $m2$ is an integer greater or equal to 0 and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag;

n is an integer greater than 0;

L_1 is an optional first linker;

L_2 is an optional second linker;

L_3 is an optional third linker;

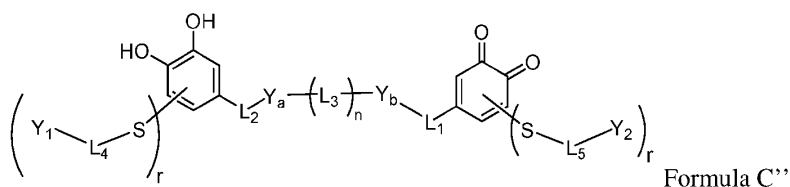
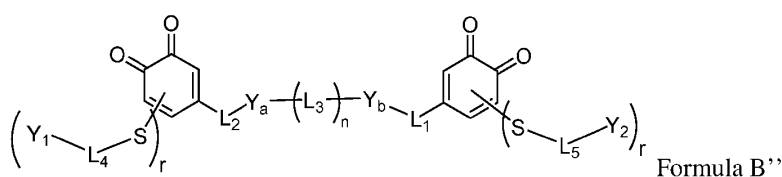
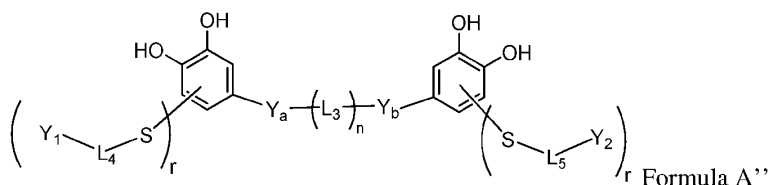
L_4 is an optional fourth linker;

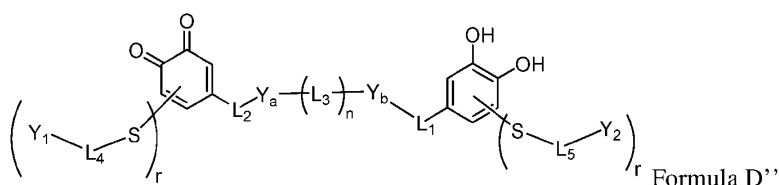
L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

[00194] Described herein, in certain embodiments, are conjugates of Formula A'', Formula B'', Formula C'', or Formula D'':





wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m_1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein m_1 is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine or a portion thereof, wherein Y_a is linked to "S" through the tyrosine or a portion thereof ;

Y_b is a second polypeptide comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m_2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine or a portion thereof, and wherein m_2 is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine or a portion thereof, wherein Y_b is linked to "S" through the tyrosine or a portion thereof;

r is 1, 2, or 3;

n is an integer greater than 0;

L_3 is an optional third linker;

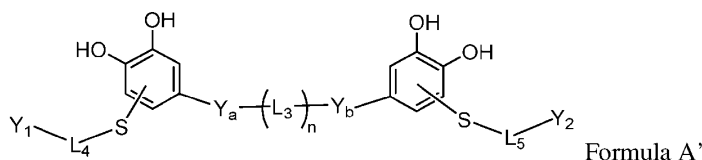
L_4 is an optional fourth linker;

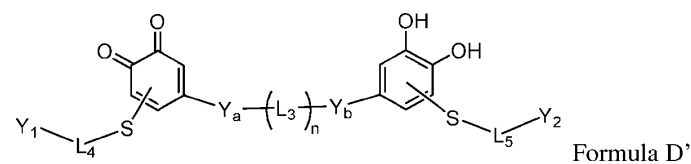
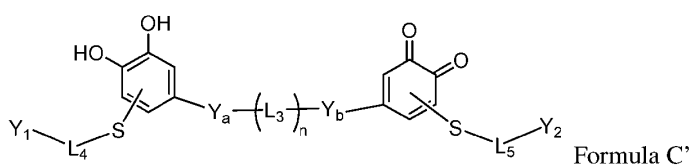
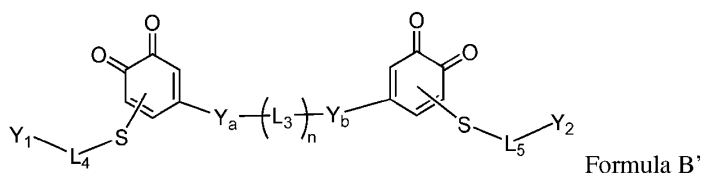
L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

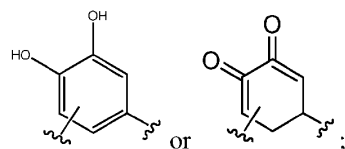
[00195] Described herein, in certain embodiments, are conjugates of Formula A', Formula B', Formula C', or Formula D':



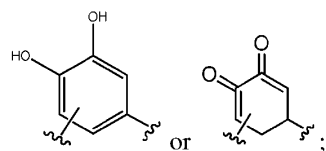


wherein:

Y_a is a group derived from a first biomolecule comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine, wherein the phenol of the tyrosine is oxidized to form:



Y_b is a second biomolecule comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, and wherein $m2$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine wherein the phenol of the tyrosine is oxidized to form:



n is an integer greater than 0;

L_3 is an optional first linker;

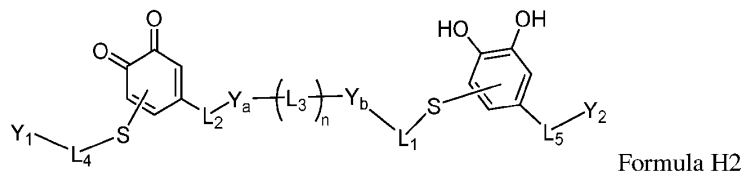
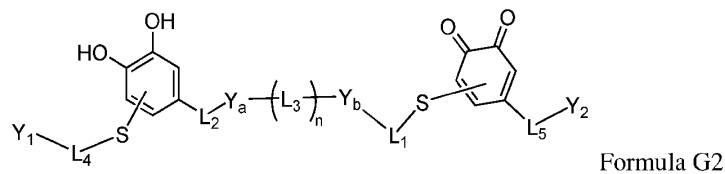
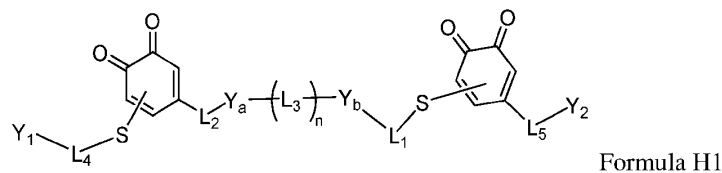
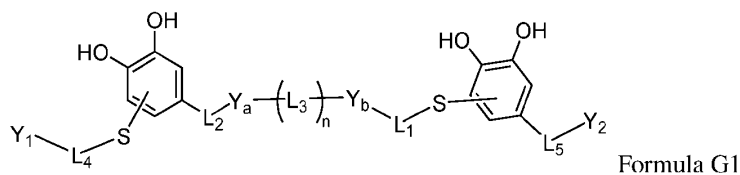
L_4 is an optional second linker;

L_5 is an optional third linker;

Y_1 is a first payload; and

Y_2 is a second payload.

[00196] Described herein, in certain embodiments, are conjugates of Formula G1, Formula H1, Formula G2, or Formula H2:



wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine, wherein Y_a is linked to "S" through the tyrosine or a portion thereof;

Y_b is a second polypeptide; n is an integer greater than 0;

L_1 is an optional first linker;

L_2 is an optional second linker;

L_3 is an optional third linker;

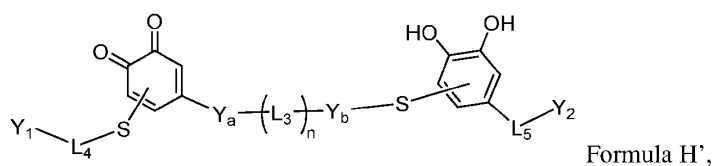
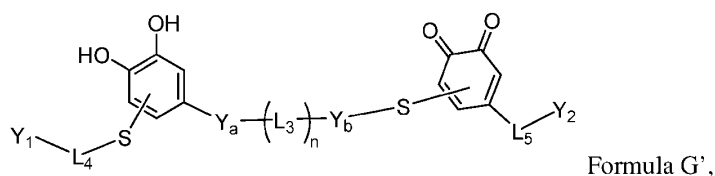
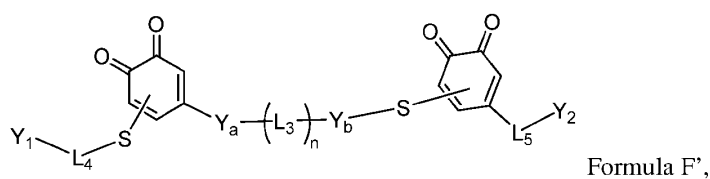
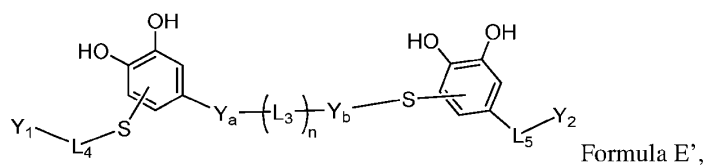
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

[00197] Described herein, in certain embodiments, are conjugates of Formula E', Formula F', Formula G', Formula H', Formula J', or Formula K':



wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine or a portion thereof, wherein Y_a is linked to "S" through the tyrosine or a portion thereof;

Y_b is a second polypeptide;

n is an integer greater than 0;

L_3 is an optional first linker;

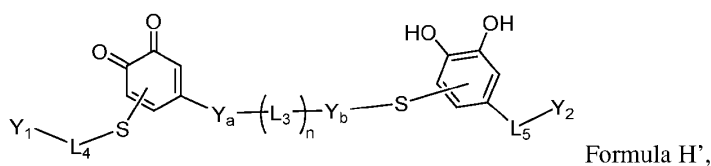
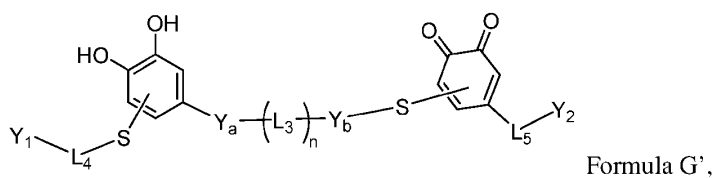
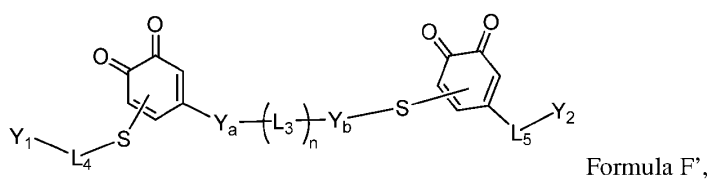
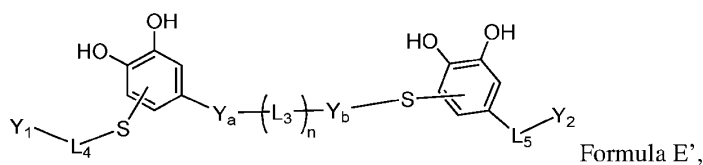
L_4 is an optional second linker;

L_5 is an optional third linker in Y_2 ;

Y_1 is a first payload; and

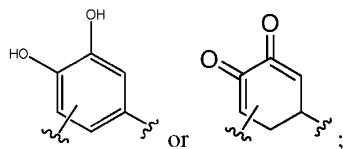
Y_2 is a group derived from a second payload comprising a phenol, wherein the phenol is linked to “S.”

[00198] Described herein, in certain embodiments, are conjugates of Formula E', Formula F', Formula G', Formula H', Formula J', or Formula K':



wherein:

Y_a is a group derived from a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5Y$, wherein X_1-X_5 is any amino acid provided that no more than two amino acids of X_2 , X_3 , X_4 , and X_5 are aspartate or glutamate and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1-X_3 is any amino acid provided that at least one amino acid of X_1-X_3 is a tyrosine, wherein the phenol of the tyrosine is oxidized to form:



Y_b is a second polypeptide;

n is an integer greater than 0;

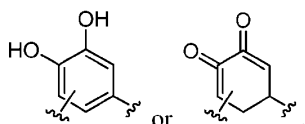
L_3 is an optional first linker;

L_4 is an optional second linker;

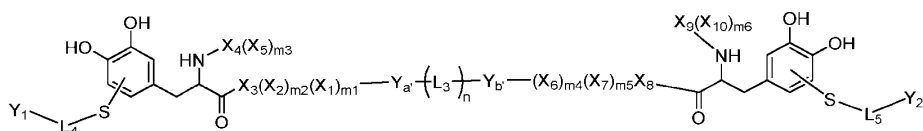
L_5 is an optional third linker in Y_2 ;

Y_1 is a first payload; and

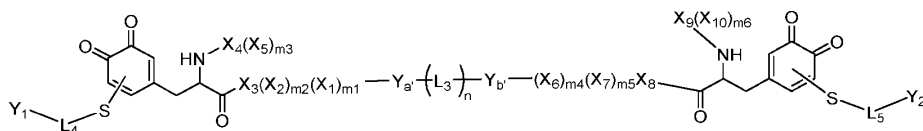
Y_2 is a group derived from a second payload comprising a phenol, wherein the phenol is oxidized to form:



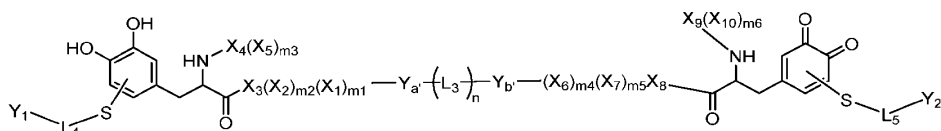
[00199] In another aspect, provided herein is a conjugate of Formula A-I, Formula B-I, Formula C-I, or Formula D-I:



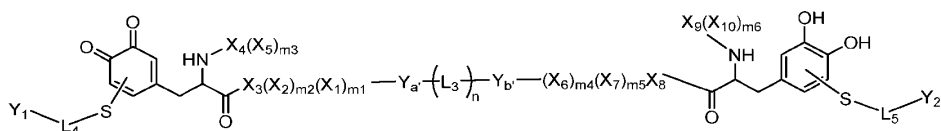
Formula A-I,



Formula B-I,



Formula C-I, or



Formula D-I;

wherein:

Y_a is a first polypeptide,

each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a first tag, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0;

Y_b is a second polypeptide,

each of X_6 , X_7 , X_8 , X_9 , and X_{10} is independently any amino acid in a second tag, wherein m_4 , m_5 , and m_6 are each an integer greater than or equal to 0, and wherein the $-(X_1)_{m_1}(X_2)_{m_2}X_3X_4-$ moiety is different from the $-(X_6)_{m_1}(X_7)_{m_2}X_8X_9-$ moiety;

n is an integer greater than 0;

L_3 is an optional third linker;

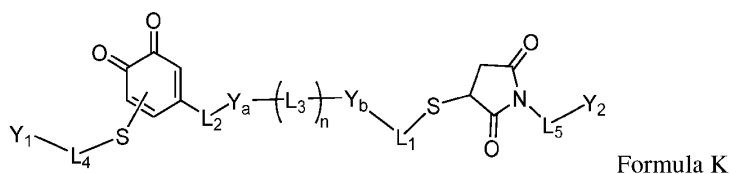
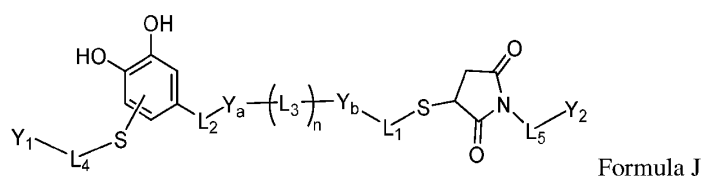
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload; wherein the first tag, the second tag, or both the first and second tag is a non-terminal tag.

[00200] Described herein, in certain embodiments, are conjugates of Formula (J) or Formula (K):



wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m_1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein m_1 is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine;

Y_b is a second polypeptide;

n is an integer greater than 0;

L_1 is an optional first linker;

L_2 is an optional second linker;

L_3 is an optional third linker;

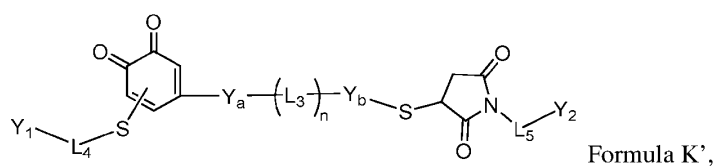
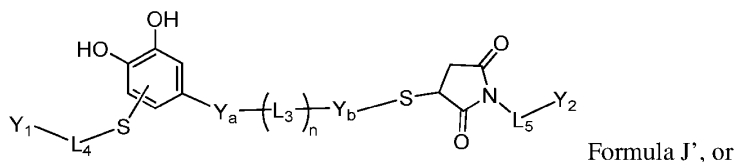
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

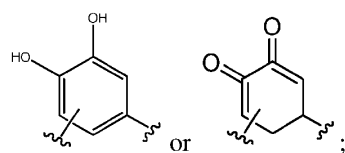
Y_2 is a second payload.

[00201] Described herein, in certain embodiments, are conjugates of Formula J' or Formula K':



wherein:

Y_a is a group derived from a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine, wherein the phenol of the tyrosine is oxidized to form:



Y_b is a second polypeptide;

n is an integer greater than 0;

L_3 is an optional first linker;

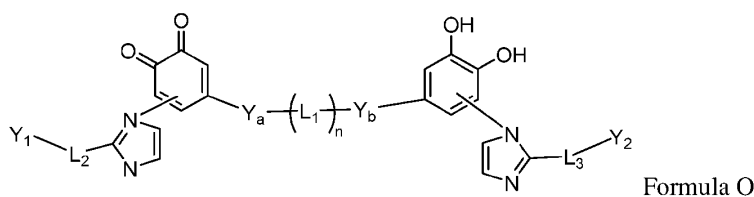
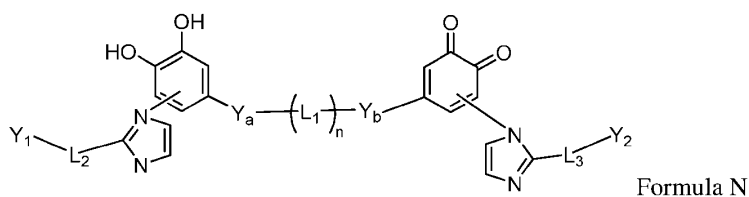
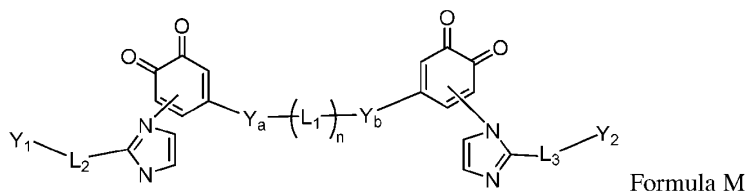
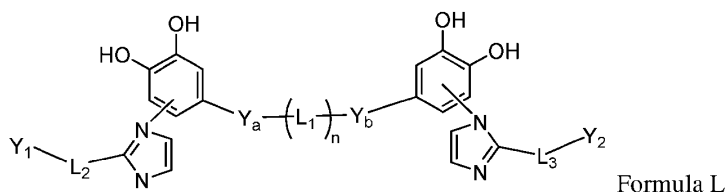
L_4 is an optional second linker;

L_5 is an optional third linker;

Y_1 is a first payload; and

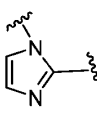
Y_2 is a second payload.

[00202] Described herein, in certain embodiments, are conjugates of Formula L, Formula M, Formula N, and Formula O:

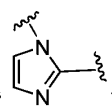


wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1-X_5 is any amino acid provided that at least one amino acid of X_1-X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1-X_3 is any amino acid provided that at least one amino acid of X_1-

X_3 is a tyrosine or a portion thereof, wherein Y_a is linked to “” through the tyrosine or a portion thereof;

Y_b is a second polypeptide comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6-X_{10} is any amino acid provided that at least one amino acid of X_6-X_{10} is a tyrosine or a portion thereof, and wherein $m2$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1-X_3 is any amino acid provided that at least one amino acid of X_1-X_3 is a tyrosine or a

portion thereof, wherein Y_b is linked to “” through the tyrosine or a portion thereof;

n is an integer greater than 0;

L_1 is an optional third linker;

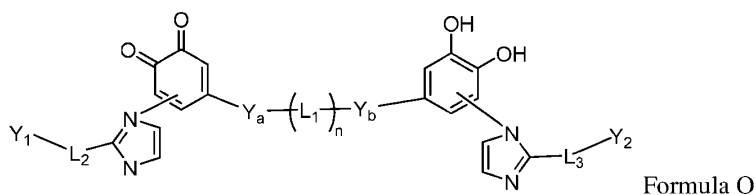
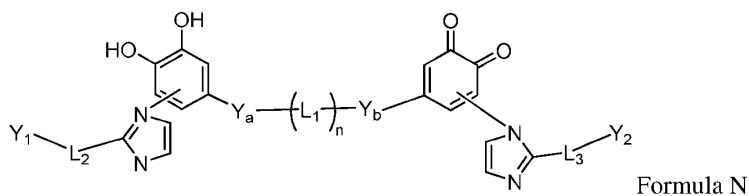
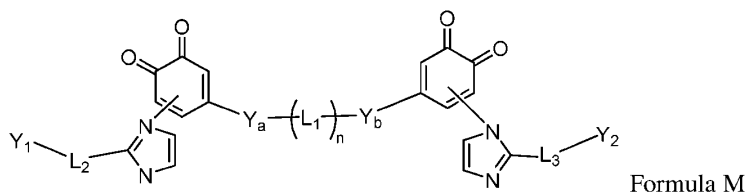
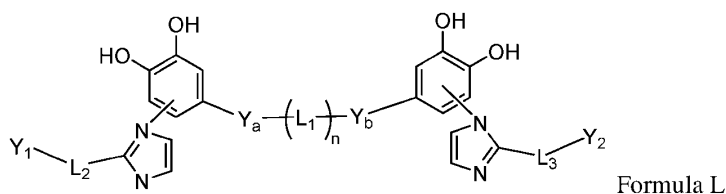
L_2 is an optional fourth linker;

L_3 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

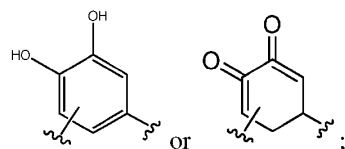
[00203] Described herein, in certain embodiments, are conjugates of Formula L, Formula M, Formula N, and Formula O:



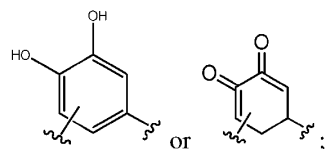
wherein:

Y_a is group derived from a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one

amino acid of X_1 - X_5 is a tyrosine, and wherein m_1 is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine, wherein the phenol of the tyrosine is oxidized to form:



Y_b is a group derived from a second polypeptide comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m_2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, and wherein m_2 is an integer greater than or equal to 0, wherein the phenol of the tyrosine is oxidized to form:



n is an integer greater than 0;

L_1 is an optional first linker;

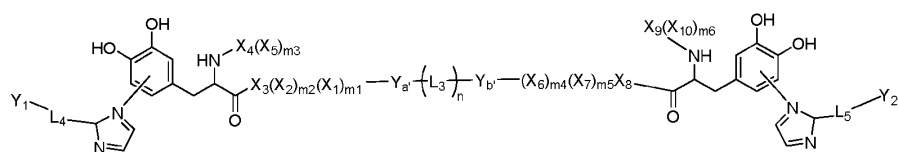
L_2 is an optional second linker;

L_3 is an optional third linker;

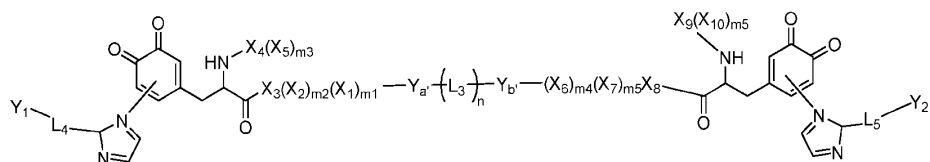
Y_1 is a first payload; and

Y_2 is a second payload.

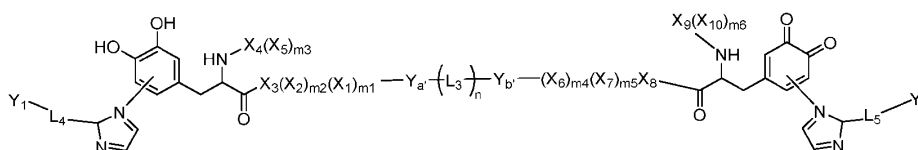
[00204] Described herein, in certain embodiments, are conjugates of Formula L-I, Formula M-I, Formula N-I, and Formula O-I:



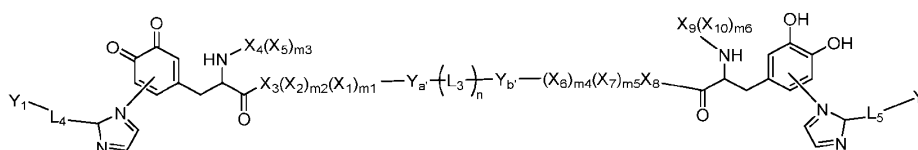
Formula L-I,



Formula M-I,



Formula N-I.



Formula O-I,

wherein:

Y_a' is a first polypeptide,

each of X₁, X₂, X₃, X₄, and X₅ is independently any amino acid in a non-terminal tag in the first polypeptide, wherein m₁, m₂, and m₃ are each an integer greater than or equal to 0;

Y_b' is a second polypeptide,

each of X₆, X₇, X₈, X₉, and X₁₀ is independently any amino acid in a non-terminal tag in the first polypeptide, wherein m₄, m₅, and m₆ are each an integer greater than or equal to 0, and wherein the -(X₁)_{m1}(X₂)_{m2}X₃X₄- moiety is different from the -(X₆)_{m1}(X₇)_{m2}X₈ X₉- moiety;

n is an integer greater than 0;

L₃ is an optional third linker;

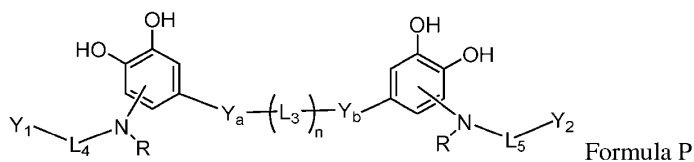
L₄ is an optional fourth linker;

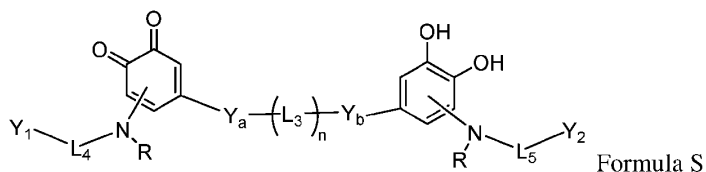
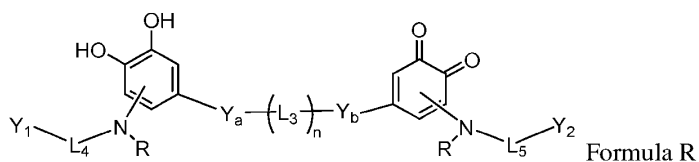
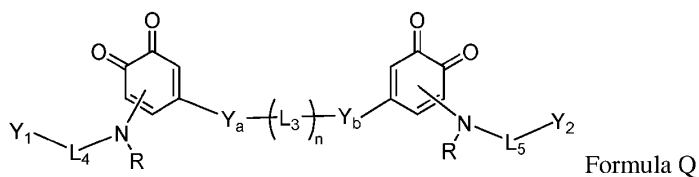
L₅ is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload.

[00205] Described herein, in certain embodiments, are conjugates of Formula P, Formula Q, Formula R, and Formula S:





wherein:

Y_a is a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1-X_5 is any amino acid provided that at least one amino acid of X_1-X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1-X_3 is any amino acid provided that at least one amino acid of X_1-X_3 is a tyrosine or a portion thereof, wherein Y_a is linked to “N” through the tyrosine or a portion thereof ;

Y_b is a second polypeptide comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6-X_{10} is any amino acid provided that at least one amino acid of X_6-X_{10} is a tyrosine or a portion thereof, and wherein $m2$ is an integer greater than or equal to 0, or $X_1X_2X_3$, wherein X_1-X_3 is any amino acid provided that at least one amino acid of X_1-X_3 is a tyrosine or a portion thereof, wherein Y_b is linked to “N” through the tyrosine or a portion thereof;

n is an integer greater than 0;

L_3 is an optional third linker;

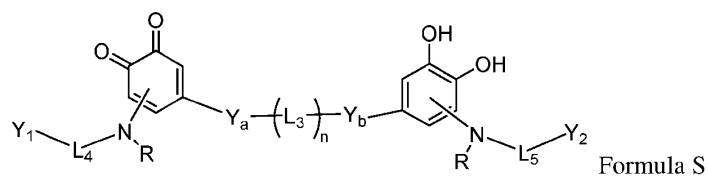
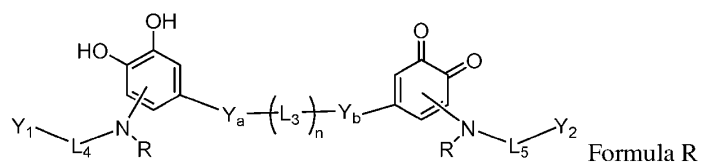
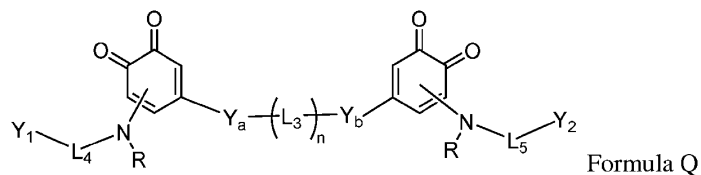
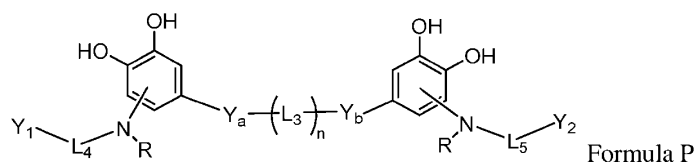
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

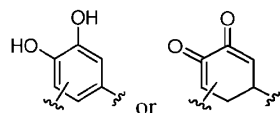
Y_2 is a second payload.

[00206] Described herein, in certain embodiments, are conjugates of Formula P, Formula Q, Formula R, and Formula S:

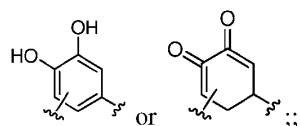


wherein:

Y_a is a group derived from a first polypeptide comprising a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater than or equal to 0; , or $X_1X_2X_3$, wherein X_1 - X_3 is any amino acid provided that at least one amino acid of X_1 - X_3 is a tyrosine, wherein the phenol of the tyrosine is oxidized to form:



Y_b is a group derived from a second polypeptide comprising a second non-terminal tag different from the first non-terminal tag, the second non-terminal tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, and wherein $m2$ is an integer greater than or equal to 0, wherein the phenol of the tyrosine is oxidized to form:



n is an integer greater than 0;

L_3 is an optional third linker;

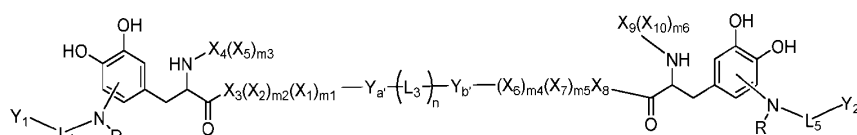
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

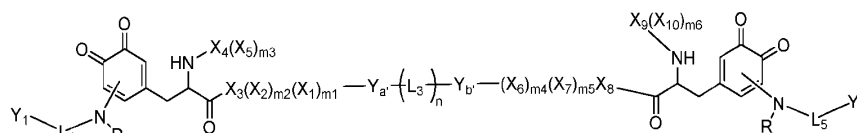
Y_1 is a first payload; and

Y_2 is a second payload.

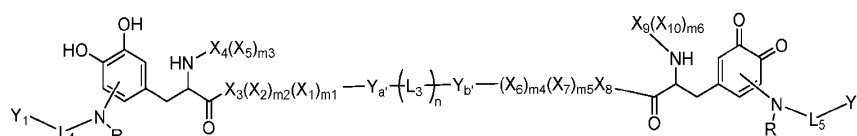
[00207] Described herein, in certain embodiments, are conjugates of Formula P-I, Formula Q-I, Formula R-I, and Formula S-I:



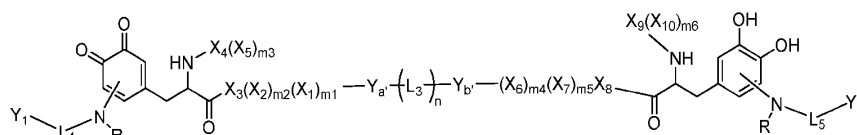
Formula P-I,



Formula Q-I,



Formula R-I,



Formula S-I,

wherein:

Y_a is a first polypeptide,

each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a non-terminal tag in the first polypeptide, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0;

Y_b is a second polypeptide,

each of X_6 , X_7 , X_8 , X_9 , and X_{10} is independently any amino acid in a non-terminal tag in a non-terminal tag in the second polypeptide, wherein m_4 , m_5 , and m_6 are each an integer

greater than or equal to 0, and wherein the $-(X_1)_{m1}(X_2)_{m2}X_3X_4-$ moiety is different from the $-(X_6)_{m1}(X_7)_{m2}X_8X_9-$ moiety;

n is an integer greater than 0;

L_3 is an optional third linker;

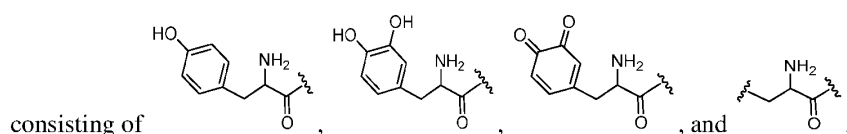
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

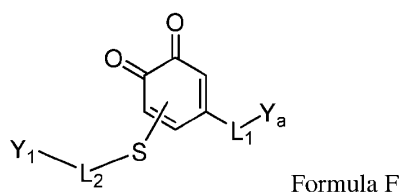
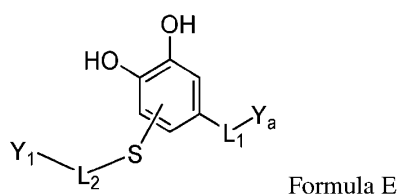
Y_1 is a first payload; and

Y_2 is a second payload.

[00208] In some embodiments, the tyrosine or a portion thereof is selected from the group



[00209] Described herein, in certain embodiments, are conjugates of Formula E of Formula F:



wherein:

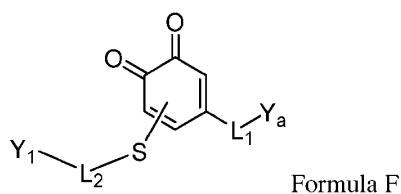
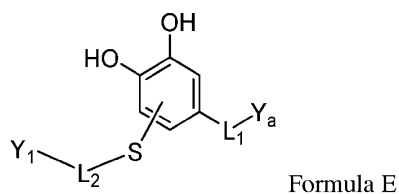
Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1-X_5 is any amino acid provided that at least one amino acid of X_1-X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to “S” via the tyrosine or a portion thereof;

L_1 is an optional first linker;

L_2 is an optional second linker; and

Y_1 is a payload.

[00210] Described herein, in certain embodiments, are conjugates of Formula E of Formula F:



wherein:

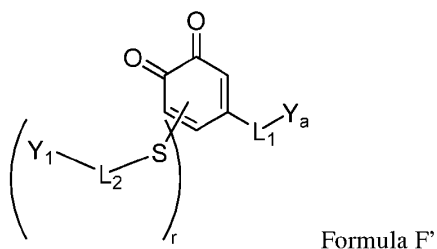
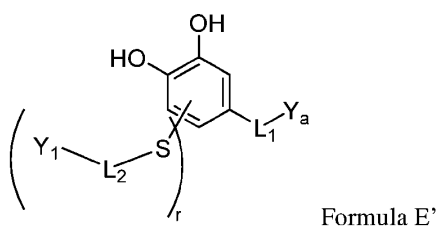
Y_a is a biomolecule comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

L_1 is an optional first linker;

L_2 is an optional second linker; and

Y_1 is a payload.

[00211] Described herein, in certain embodiments, are conjugates of Formula E' of Formula F':



wherein:

Y_a is a biomolecule comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a

tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

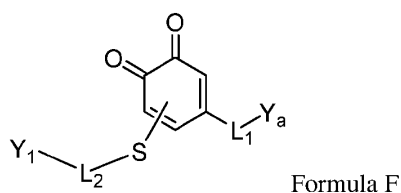
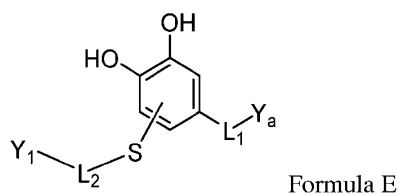
r is 1, 2, or 3;

L_1 is an optional first linker;

L_2 is an optional second linker; and

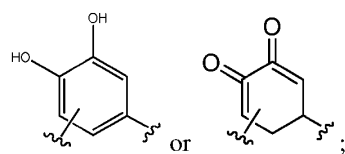
Y_1 is a payload.

[00212] Described herein, in certain embodiments, are conjugates of Formula E or Formula F:



wherein:

Y_a is a group derived from a biomolecule comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater than or equal to 0, wherein the phenol of the tyrosine is oxidized to form:

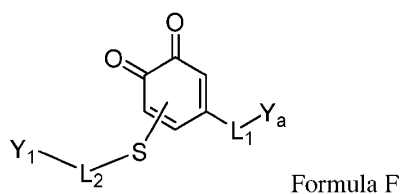
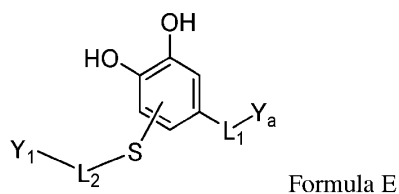


L_1 is an optional first linker;

L_2 is an optional second linker; and

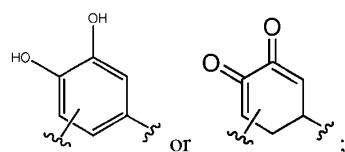
Y_1 is a payload.

[00213] Described herein, in certain embodiments, are conjugates of Formula E or Formula F:



wherein:

Y_a is a group derived from a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater than or equal to 0, wherein the phenol of the tyrosine is oxidized to form:

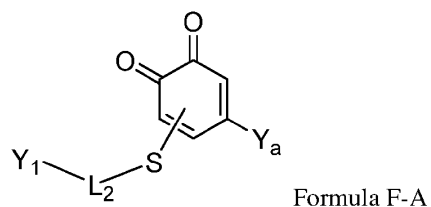
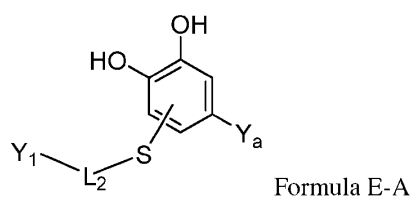


L_1 is an optional first linker;

L_2 is an optional second linker; and

Y_1 is a payload.

[00214] Described herein, in certain embodiments, are conjugates of Formula E-A or Formula F-A:



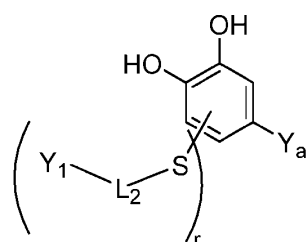
wherein:

Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

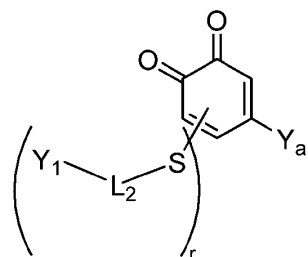
L_2 is an optional linker; and

Y_1 is a payload.

[00215] Described herein, in certain embodiments, are conjugates of Formula E'-A or Formula F'-A:



Formula E'-A



Formula F'-A

wherein:

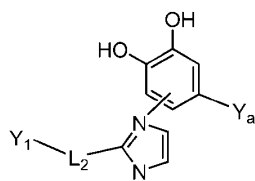
Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

r is 1, 2, or 3;

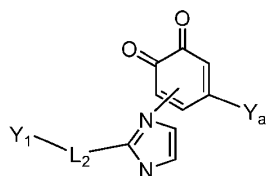
L_2 is an optional linker; and

Y_1 is a payload

[00216] Described herein, in certain embodiments, are conjugates of Formula T or Formula U:



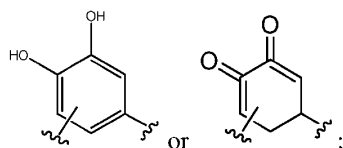
Formula T



Formula U

wherein:

Y_a is a group derived from a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater than or equal to 0, wherein the phenol of the tyrosine is oxidized to form:

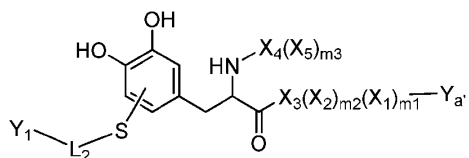


wherein:

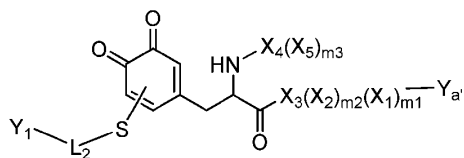
L_2 is an optional linker; and

Y_1 is a payload.

[00217] Described herein are, in certain embodiments, are conjugates of Formula E-I or Formula F-I:



Formula E-I



Formula F-I

wherein:

$Y_{a'}$ is a polypeptide,

each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a non-terminal tag in the polypeptide, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0;

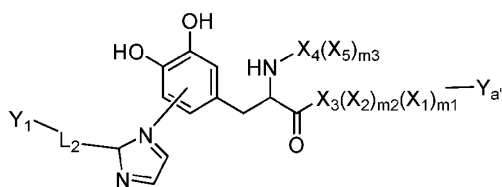
L_2 is an optional linker; and

Y_1 is a payload.

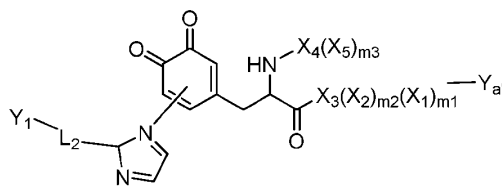
[00218] In some embodiments, X_2 and X_3 are proton-donating amino acids. In some embodiments, X_2 and X_3 are each selected from the group consisting of E, H, Q, D, and N.

[00219] In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G and X_5 is E, V, D, A, I, G, S, T, or L.

[00220] Described herein are, in certain embodiments, are conjugates of Formula T-I or Formula U-I:



Formula T-I



Formula U-I

wherein:

$Y_{a'}$ is a polypeptide,

each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a non-terminal tag in the polypeptide, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0;

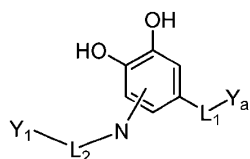
L_2 is an optional linker; and

Y_1 is a payload.

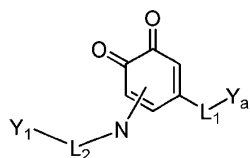
[00221] In some embodiments, X_2 and X_3 are proton-donating amino acids. In some embodiments, X_2 and X_3 are each selected from the group consisting of E, H, Q, D, and N.

[00222] In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G and X_5 is E, V, D, A, I, G, S, T, or L.

[00223] Described herein are, in certain embodiments, are conjugates of Formula V or Formula W:



Formula V



formula W

wherein:

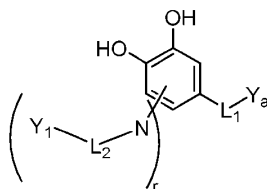
Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

L_1 is an optional first linker;

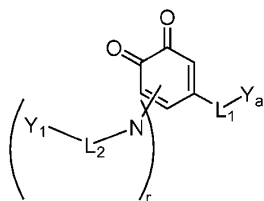
L_2 is an optional second linker; and

Y_1 is a payload.

[00224] Described herein, in certain embodiments, are conjugates of Formula V' or Formula W':



Formula V'



Formula W'

wherein:

Y_a is a biomolecule comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

r is 1, 2, or 3;

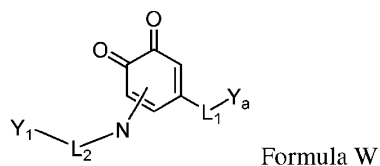
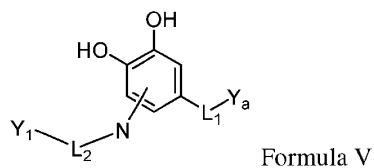
L_1 is an optional first linker;

L_2 is an optional second linker; and

Y_1 is a payload.

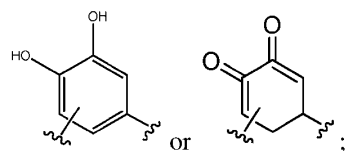
[00225] In some embodiments, Y_a is a polypeptide.

[00226] Described herein, in certain embodiments, are conjugates of Formula V or Formula W:



wherein:

Y_a is a group derived from a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater than or equal to 0, wherein the phenol of the tyrosine is oxidized to form:

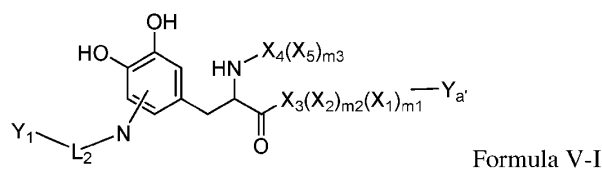


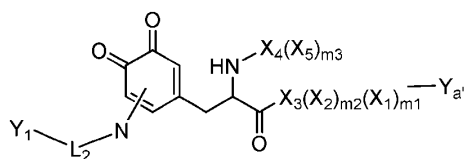
L_1 is an optional first linker;

L_2 is an optional second linker; and

Y_1 is a payload.

[00227] Described herein are, in certain embodiments, are conjugates of Formula V-I or Formula W-I:





Formula W-I

wherein: $Y_{a'}$ is a polypeptide,

each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a non-terminal tag in the polypeptide, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0;

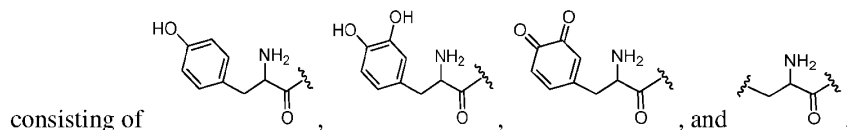
L_2 is an optional linker; and

Y_1 is a payload.

[00228] In some embodiments, X_2 and X_3 are proton-donating amino acids. In some embodiments, X_2 and X_3 are each selected from the group consisting of E, H, Q, D, and N.

[00229] In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G and X_5 is E, V, D, A, I, G, S, T, or L.

[00230] In some embodiments, the tyrosine or a portion thereof is selected from the group



[00231] In some embodiments, Y_a , Y_b , $Y_{a'}$, $Y_{b'}$, or Y_a and Y_b is a nanoparticle, a polymer, a nucleic acid sequence, a polypeptide, or an aptamer.

[00232] In some embodiments, X_1 is a tyrosine. In some embodiments, X_2 is a tyrosine. In some embodiments, X_3 is a tyrosine. In some embodiments, X_4 is a tyrosine. In some embodiments, X_5 is a tyrosine.

[00233] In some embodiments, X_6 is a tyrosine. In some embodiments, X_7 is a tyrosine. In some embodiments, X_8 is a tyrosine. In some embodiments, X_9 is a tyrosine. In some embodiments, X_{10} is a tyrosine.

[00234] In some embodiments, at least one of L_3 , L_4 , or L_5 is a cleavable linker. In some embodiments, the cleavable linker is an electrophilically cleavable linker, a nucleophilically cleavable linker, a photocleavable linker, a metal cleavable linker, an electrolytically-cleavable linker, an acid cleavable linker, or a proteolytically cleavable linker.

[00235] In some embodiments, the cleavable linker further include a pegylated group, a sugar group, or a modification that increases hydrophilicity. In some embodiments, the cleavable linker is cleavable under reductive and/or oxidative conditions. In some

embodiments, the cleavable linker is cleavable under acidic conditions. In some embodiments, the cleavable linker comprises a disulfide bond.

[00236] In some embodiments, the cleavable linker is a proteolytically cleavable linker and comprises a protease recognition sequence. In some embodiments, the protease recognition sequence is recognized by a protease selected from the group comprising a metalloprotease, cathepsin B, and tobacco etch virus (TEV).

[00237] In some embodiments, the cleavable linker comprises a dipeptide, tripeptide or tetrapeptide. In some embodiments, the dipeptide is a valine-citrulline (Val-Cit) dipeptide, a valine-lysine dipeptide, a valine-alanine dipeptide. In some embodiments, the tetrapeptide is a glycine-glycine-phenylalanine-glycine (GGFG) tetrapeptide (SEQ ID NO: 75).

[00238] In some embodiments, the cleavable linker is selected from the group comprising PABC (p-aminobenzyl alcohol), glucuronide, and MABC (m-aminobenzyl alcohol),

[00239] In some embodiments, L_3 comprises a peptide bond, a disulfide bond, a maleimide bond, thioether bond, an azide-alkyne cycloaddition, a cystinyl-dopa, or a hydrogen bond.

[00240] In some embodiments, L_3 is a linker. In some embodiments, the linker comprises a sequence selected from the group consisting of $(GS)_{n3}$ (SEQ ID NO: 66), $(G2S)_{n3}$ (SEQ ID NO: 67), $(G3S)_{n3}$ (SEQ ID NO: 68), $(G4S)_{n3}$ (SEQ ID NO: 69), $(G)_{n3}$ (SEQ ID NO: 70), $(GGSGGD)_{n3}$ (SEQ ID NO: 71), $(GGSGGE)_{n3}$ (SEQ ID NO: 72), $(GGGSGSGGGGS)_{n3}$ (SEQ ID NO: 73), and $(GGGGGPGGGGP)_{n3}$ (SEQ ID NO: 74) and wherein $n3$ is an integer from 2 to 20. In some embodiments, n is 1, 2, 3, 4, or 5. In some embodiments, n is 1.

[00241] In some embodiments, Y_1 and Y_2 are the same.

[00242] In some embodiments, Y_1 and Y_2 are different.

[00243] In some embodiments, Y_1 , Y_2 , or both Y_1 and Y_2 is a small molecule.

[00244] In some embodiments, the small molecule is selected from the group consisting of deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan, luritotecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimmitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4), pyrolodiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT.

[00245] In some embodiments, the small molecule is selected from the group of a topoisomerase inhibitor and a tubulin inhibitor.

[00246] In some embodiments, Y_1 , Y_2 , or both Y_1 and Y_2 comprises a nucleic acid, an immune agonist, a peptide, a cytokine, or a binding domain. In some embodiments, Y_1 , Y_2 ,

or both Y_1 and Y_2 comprises a nucleic acid. In some embodiments, Y_1 and Y_2 , or both Y_1 and Y_2 comprises an oligonucleotide.

[00247] In some embodiments, Y_1 and Y_2 , or both Y_1 and Y_2 comprises a peptide. In some embodiments, Y_1 and Y_2 , or both Y_1 and Y_2 comprises a binding peptide.

[00248] Described herein, in certain embodiments, are conjugates comprising a phenol moiety or a catechol moiety conjugated to one or more payloads comprising a thiol moiety.

[00249] In some embodiments, Y_a comprises a first tag. In some embodiments, the first tag comprises $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0. In some embodiments, the first tag comprises $(X_1)_{m1}X_2X_3X_4X_5Y$, wherein $m1$ is an integer greater or equal to 0. In some embodiments, the first tag comprises $Y(X_1)_{m1}X_2X_3X_4X_5$, wherein $m1$ is an integer greater or equal to 0.

[00250] In some embodiments, the first tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K). In some embodiments, the first tag comprises at least one of GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGGY (SEQ ID NO: 13), EEEGY (SEQ ID NO: 14), EEEEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 30), KKGYY (SEQ ID NO: 31), and DDGGY (SEQ ID NO: 32).

[00251] In some embodiments, the non-terminal tag comprises $X_1X_2YX_4X_5$. In some embodiments, X_1 and X_2 are proton-donating amino acids. In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N. In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G. In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L. In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7. In some embodiments, the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.

[00252] In some embodiments, Y_b comprises a second tag. In some embodiments, Y_b comprises a second tag different from the first tag. In some embodiments, the second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one

amino acid of X_6 - X_{10} is a tyrosine, and wherein m_2 is an integer greater or equal to 0. In some embodiments, the second tag comprises $(X_6)_{m_2}X_7X_8X_9X_{10}Y$, wherein m_2 is an integer greater or equal to 0. In some embodiments, the second tag comprises $Y(X_6)_{m_2}X_7X_8X_9X_{10}$, wherein m_2 is an integer greater or equal to 0.

[00253] In some embodiments, the second tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K). In some embodiments, the second tag comprises at least one of GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGGY (SEQ ID NO: 13), EEEGY (SEQ ID NO: 14), EEEEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWEY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 30), KKGYY (SEQ ID NO: 31), and DDGGY (SEQ ID NO: 32).

[00254] In some embodiments, the first tag comprises the first tag comprises SGGY (SEQ ID NO: 23) or SGY and the second tag comprises EEEY (SEQ ID NO: 33).

[00255] In some embodiments, Y_a , Y_b , or both Y_a and Y_b are glycosylated. In some embodiments, Y_a , Y_b , or both Y_a and Y_b comprise a non-terminal tyrosine.

[00256] In some embodiments, Y_a is a heavy chain variable region (VH) of an antibody or antibody fragment, a light chain variable region (VL) of an antibody or antibody fragment, a heavy chain of an antibody or antibody fragment, a light chain of an antibody or antibody fragment, a constant chain of an antibody or antibody fragment, a peptide, or a cyclic peptide.

[00257] In some embodiments, Y_b is a heavy chain variable region (VH) of an antibody or antibody fragment, a light chain variable region (VL) of an antibody or antibody fragment, a heavy chain of an antibody or antibody fragment, a light chain of an antibody or antibody fragment, a constant chain of an antibody or antibody fragment, a peptide, or a cyclic peptide.

[00258] In some embodiments, Y_a and Y_b are attached via L_3 . In some embodiments, L_3 comprises a peptide sequence, a dimerization and docking domain, a leucine zipper, or knobs-into-holes. In some embodiments, L_3 comprises a peptide bond, a disulfide bond, a maleimide bond, thioether bond, an azide-alkyne cycloaddition, a cysteinyl-dopa, or a hydrogen bond. In some embodiments, L_3 is a linker. In some embodiments, the linker comprises a sequence selected from the group consisting of $(GS)_n$ (SEQ ID NO: 66), $(G_2S)_n$ (SEQ ID NO: 67), $(G_3S)_n$ (SEQ ID NO: 68), $(G_4S)_n$ (SEQ ID NO: 69), $(G)_n$

(SEQ ID NO: 70), (GGSGGD)_{n3} (SEQ ID NO: 71), (GGSGGE)_{n3} (SEQ ID NO: 72), (GGGSGSGGGGS)_{n3} (SEQ ID NO: 73), and (GGGGPGGGGP)_{n3} (SEQ ID NO: 74) and wherein n₃ is an integer from 2 to 20.

[00259] In some embodiments, L₃ comprises a non-terminal tyrosine.

[00260] In some embodiments, n is 1, 2, 3, 4, or 5. In some embodiments, n is 1.

[00261] In some embodiments, Y₁ and Y₂ are the same payload. In some embodiments, Y₁ and Y₂ are different payloads. In some embodiments, Y₁ and Y₂, or both Y₁ and Y₂ is a small molecule.

[00262] In some embodiments, the small molecule is selected from the group comprising deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan, lurototecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimnitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4), pyrolodiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT.

[00263] In some embodiments, Y₁ and Y₂, or both Y₁ and Y₂ comprises a nucleic acid, an immune agonist, a peptide, a cytokine, or a binding domain.

[00264] In some embodiments, the non-terminal tag is part of a tyrosinase activatable loop. In some embodiments, the tyrosinase activatable loop comprises a tyrosine. In some embodiments, a polypeptide of interest is modified to contain a non-terminal tag as described herein to comprise an unstructured (e.g., non-alpha helical or non-beta sheet) loop that contains a tyrosine residue exposed to solvent in such a way as to be activated by tyrosinase enzyme. In some embodiments, the polypeptide is not modified. For example, a polypeptide would not need to be modified if the non-terminal tag as described herein is already surrounding a tyrosine in the polypeptide. In this example, the non-terminal tag sequence that includes the tyrosine would then allow for non-terminal activation with a tyrosinase enzyme and conjugation with a payload molecule containing a free thiol, such as a cytotoxic payload, a cytokine, a binding domain, or a nucleic acid, as described herein. In some embodiments, a tyrosinase activatable loop is the stretch of amino acids containing at least a non-terminal tag containing to a tyrosine residue.

[00265] In some embodiments, the tyrosinase activatable loop comprises 3, 4, 5, 6, 7, or 8 amino acids, wherein at least one of the amino acids in the tyrosinase activatable loop is a tyrosine residue. In some embodiments the tyrosine is adjacent to a proton donating residue (e.g., aspartate, glutamate, histidine, asparagine, glutamine, lysine). In some embodiments,

the tyrosinase activatable loop comprises 3 or more amino acids from a sequence selected from the group comprising EEQYNST (SEQ ID NO: 45), NSTYRVV (SEQ ID NO: 46), DRVYIHP (SEQ ID NO: 47), IDHYSWE (SEQ ID NO: 48), and VEDYAIGP (SEQ ID NO: 49). In some embodiments, the tyrosinase activatable loop comprises 3 or more amino acids in the sequence EEQYNST (SEQ ID NO: 45). In some embodiments, the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.

[00266] In some embodiments, the non-terminal tag includes a native tyrosine at the beginning or at the end. In some embodiments, the non-terminal tag includes a native tyrosine in an amino acid that is not at the beginning or the end of the non-terminal tag as described herein.

Linkers

[00267] Linker sequences can be used to separate different components of the conjugates described herein. In some embodiments, the conjugates described herein comprise linkers L₁, L₂, L₃, L₄, and/or L₅.

[00268] In some embodiments, any one of the linkers comprises at least 5 to about 50 amino acids. In some embodiments, the linkers comprise about 5 to about 50 amino acids, about 5 to about 45 amino acids, about 5 to about 40 amino acids, about 5 to about 35 amino acids, about 5 to about 30 amino acids, about 5 to about 25 amino acids, about 5 to about 20 amino acids, about 5 to about 15 amino acids, about 5 to about 10 amino acids, about 10 to about 50 amino acids, about 15 to about 50 amino acids, about 20 to about 50 amino acids, about 25 to about 50 amino acids, about 30 to about 50 amino acids, about 35 to about 50 amino acids, about 40 to about 50 amino acids, or about 45 to about 50 amino acids.

[00269] In some embodiments, any one of the linkers comprises a sequence selected from the group consisting of (GS)_n (SEQ ID NO: 66), (G2S)_n (SEQ ID NO: 67), (G3S)_n (SEQ ID NO: 68), (G4S)_n (SEQ ID NO: 69), and (G)_n (SEQ ID NO: 70), and wherein n is an integer from 2 to 20. In some embodiments, n is an integer from 2 to 18, from 2 to 16, from 2 to 14, from 2 to 12, from 2 to 10, from 2 to 8, from 2 to 6, from 2 to 4, from 4 to 20, from 6 to 20, from 8 to 20, from 10 to 20, from 12 to 20, from 14 to 20, from 16 to 20, or from 18 to 20.

[00270] In some embodiments, any one of the linkers comprises a sequence selected from the group consisting of (GGSGGD)_n (SEQ ID NO: 71) or (GGSGGE)_n (SEQ ID NO: 72), and wherein n is an integer from 2 to 6.

[00271] In some embodiments, any one of the linkers comprises a sequence selected from the group consisting of (GGGSGSGGGGS)_n and (GGGGGPGGGGP)_n, and wherein n is an integer from 1 to 3.

[00272] In some embodiments, any one of the linkers comprises a sequence selected from the group consisting of (GX)_n, (GGX)_n, (GGGX)_n, (GGGGX)_n, and (GzX)_n, wherein z is between 1 and 20, and wherein n is at least 8. In some embodiments, z is between 2 and 18, 2 and 16, 2 and 14, 2 and 12, 2 and 10, 2 and 8, 2 and 6, 2 and 4, 4 and 20, 6 and 20, 8 and 20, 10 and 20, 12 and 20, 14 and 20, 16 and 20, or 18 and 20. In some embodiments, X is serine, aspartic acid, glutamic acid, threonine, or proline.

[00273] In some embodiments, a linker in the conjugates disclosed herein is a cleavable linker. In some embodiments, at least one of L₁, L₂, L₃, L₄, or L₅ is a cleavable linker. In some embodiments, the cleavable linker is an electrophilically cleavable linker, a nucleophilically cleavable linker, a photocleavable linker, a metal cleavable linker, an electrolytically cleavable linker, an acid cleavable linker, or a proteolytically cleavable linker. In some embodiments, the cleavable linker is cleavable under reductive and/or oxidative conditions. In some embodiments, the cleavable linker is cleavable under acidic conditions. In some embodiments, the cleavable linker is cleaved by an enzyme. In certain cases, the cleavable linker is a linker that is cleaved under reducing conditions. In some embodiments, the cleavable linker is cleaved rapidly by glutathione reduction. In some embodiments, the cleavable linker comprises a disulfide bond. In some embodiments, the cleavable linker is cleaved by a physical stimulus. In some embodiments, the cleavable linker is photocleavable.

[00274] In some embodiments, L₁, L₂, L₃, L₄, or L₅ is an acid-labile linker. In some embodiments, the linker cleaves at a pH of 6 or less, such as, 6.0, 5.95, 5.9, 5.85, 5.8, 5.75, 5.7, 5.65, 5.6, 5.55, 5.5, 5.45, 5.4, 5.35, 5.3, 5.25, 5.2, 5.15, 5.1, 5.05, 5.0, 4.9, 4.85, 4.80, 4.75, 4.7, 4.65, 4.6, 4.55, 4.5 or even less.

[00275] In some embodiments, L₁, L₂, L₃, L₄, or L₅ is a photocleavable linker. Suitable photocleavable linkers include ortho-nitrobenzyl-based linkers, phenacyl linkers, alkoxybenzoin linkers, chromium arene complex linkers, NpSSMpact linkers and pivaloylglycol linkers, as described in Guillier et al. (Chem. Rev. 2000 1000:2091-2157).

[00276] In some embodiments, L₁, L₂, L₃, L₄ or L₅ is a proteolytically cleavable linker. The proteolytically cleavable linker can include a protease recognition sequence recognized by a protease selected from the group consisting of alanine carboxypeptidase, Armillaria mellea astacin, bacterial leucyl aminopeptidase, cancer procoagulant, cathepsin B, clostripain, cytosol alanyl aminopeptidase, elastase, endoproteinase Arg-C, enterokinase, gastricsin,

gelatinase, Gly-X carboxypeptidase, glycyl endopeptidase, human rhinovirus 3C protease, hypodermin C, IgA-specific serine endopeptidase, leucyl aminopeptidase, leucyl endopeptidase, lysC, lysosomal pro-X carboxypeptidase, lysyl aminopeptidase, methionyl aminopeptidase, myxobacter, nardilysin, pancreatic endopeptidase E, picornain 2A, picornain 3C, proendopeptidase, prolyl aminopeptidase, proprotein convertase I, proprotein convertase II, russellysin, saccharopepsin, semenogelase, T-plasminogen activator, thrombin, tissue kallikrein, tobacco etch virus (TEV), togavirin, tryptophanyl aminopeptidase, U-plasminogen activator, V8, venombin A, venombin AB, and Xaa-pro aminopeptidase.

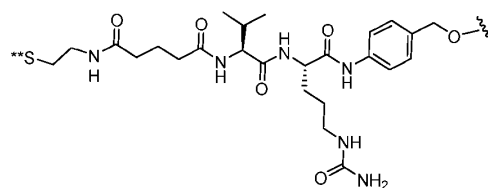
[00277] In some embodiments, the proteolytically cleavable linker comprises a matrix metalloproteinase cleavage site, e.g., a cleavage site for a MMP selected from collagenase-1, -2, and -3 (MMP-1, -8, and -13), gelatinase A and B (MMP-2 and -9), stromelysin 1, 2, and 3 (MMP-3, -10, and -11), matrilysin (MMP-7), and membrane metalloproteinases (MT1-MMP and MT2-MMP).

[00278] In some embodiments, the cleavable linker comprises a disulfide bond and is cleavable under reducing conditions, e.g., using b-mercaptoethanol, cysteine-HCl, Iris (2-carboxyethyl) phosphine hydrochloride, or another reducing agent.

[00279] In some embodiments, the cleavable linker comprises a dipeptide. In some embodiments, the dipeptide is a valine-citrulline (Val-Cit) dipeptide or a valine-lysine dipeptide. In some embodiments, the cleavable linker comprises a tetrapeptide. In some embodiments, the tetrapeptide is a glycine-glycine-phenylalanine-glycine tetrapeptide (GGFG).

[00280] In some embodiments, the cleavable linker is PABC (p-aminobenzyl alcohol), glucuronide, or MABC (m-aminobenzyl alcohol).

[00281] In some embodiments, the cleavable linker is a valine-citrulline (Val-Cit) PABC (p-aminobenzyl alcohol) (VC-PABC) linker. In some embodiments, the VC-PABC linker has

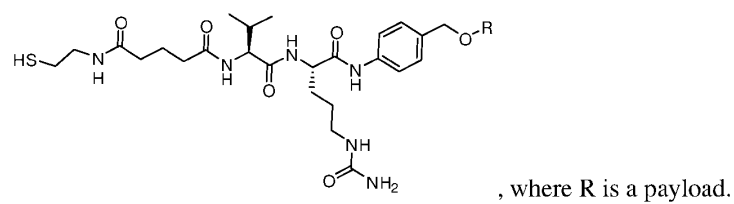


the structure:.,

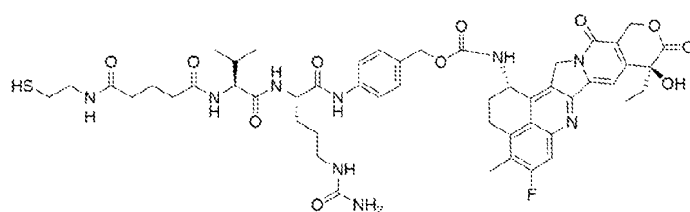
attachment of the payload is “.”

wherein the point of

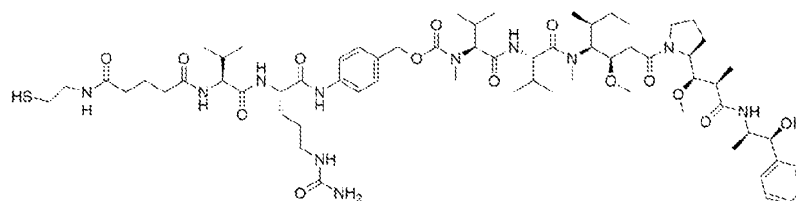
[00282] In some embodiments, the disclosure provides a compound of formula L-I,



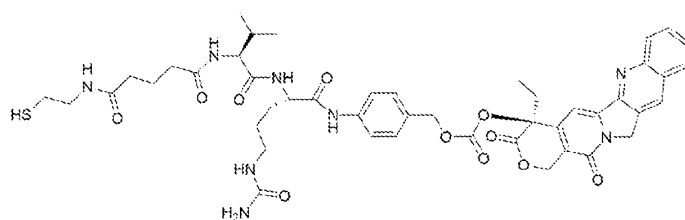
[00283] In some embodiments, the compound is selected from the compound is selected from the group consisting of:



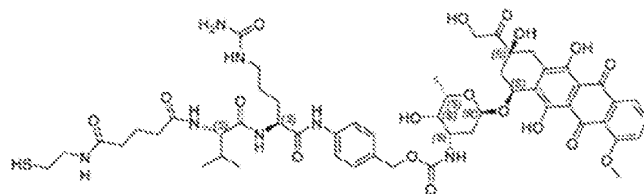
SH-VC-PAB-DXD,



SH-VC-PAB-MMAE,



SH-VC-PAB-Camptothecin, and



SH-VC-PAB-Doxorubicin.

[00284] In some embodiments, the cleavable linker further includes a polyethylene glycol group (i.e., a pegylated group), a sugar group, or a modification that increases hydrophilicity. In some embodiments, a cleavable linker including a sugar group is a glucuronide linker. In some embodiments, a modification that increases hydrophilicity include a longer peptide chain, inclusion of charged residues, additional PEG spacers, branched PEG substituents, peptide bonds, sulfonate bonds, glucuronide linkers, and methylation of tertiary amines to produce cations.

Polypeptides and Biomolecules

[00285] Described herein, in certain embodiments, are conjugates comprising one or more polypeptides or biomolecules.

[00286] In some embodiments, the polypeptide is an enzyme, an antibody, a structural polypeptide, a ligand for a receptor, a viral protein, or a receptor, such as a cell surface receptor.

[00287] In some embodiments, the polypeptide is an antibody. In some embodiments, the antibody is a single chain Fv (scFv). In some embodiments, the polypeptide is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, a constant chain of an antibody or antibody fragment. Other antibody-based recognition domains (cAb VF₁Fi (camelid antibody variable domains) and humanized versions, IgNAR VF₁ (shark antibody variable domains) and humanized versions, sdAb VF₁ (single domain antibody variable domains) and “camelized” antibody variable domains are suitable for use. In some embodiments, T- cell receptor (TCR) based recognition domains such as single chain TCR (scTv, single chain two-domain TCR containing nanobody) are also suitable for use.

[00288] Biomolecules that are suitable for use in a method or conjugate of the present disclosure include polypeptides, peptides, cyclic peptides, polynucleotides, nucleic acids, aptamers, glycoproteins, small molecules, carbohydrates, lipids, glycolipids, lipoproteins, fatty acids, lipopolysaccharides, sugars, amino acids, organic dyes, synthetic polymers, steroids, nanoparticles, polymers, purines, pyrimidines, derivatives, structural analogs thereof and combinations thereof.

[00289] In some embodiments, a biomolecule suitable for conjugating to a payload is a peptide. In some embodiments, the peptide is a peptide ligand or a binding peptide. In some embodiments, the binding peptide can be of different origins, e.g., synthetic, human, mouse, or rat. In some embodiments, a binding peptide may be or have been engineered to include one or more (e.g., two, three, four, or five) solvent-exposed cysteine or lysine residues, which

may provide a site for conjugation. In some embodiments, the binding peptides may include only naturally occurring amino acid residues, or may include one or more non-naturally occurring amino acid residues. In some embodiments, binding peptides may be linear or cyclic. In some embodiments, the peptide ligand or binding peptide is a bicyclic peptide. In some embodiments, a binding peptide can be a monospecific peptide or multi-specific peptide (e.g., a bispecific peptide or a trispecific peptide). Bispecific peptide ligands or dual specific peptide ligands may bind one target at a time or bind two targets simultaneously. In some embodiments, a binding peptide is a cysteine motif binding peptide.

[00290] As used herein, “binding peptide” refers to peptides having a potential capability of binding other compounds and/or structures (e.g., polypeptides and proteins).

[00291] Suitable lipids may include, e.g., 3-N-[(methoxypoly (ethylene glycol) 2000) carbamoyl]-1,2-dimyristyloxy-propylamine (PEG-C-DMA), 1,2-dilinoleyloxy-N,N-dimethyl-3-aminopropane (DLinDMA), 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), cholesterol, dipalmitoylphosphatidylcholine, 3-N-[(w-methoxy poly(ethylene glycol)2000)carbamoyl]-1,2-dimyrestyloxypropylamine, 1,2-dilinoleyloxy-3-N,Ndimethylaminopropane, 1,2-distearoyl-sn-glycero-3-phosphocholine, PEG-cDMA, 1,2-dilinoleyloxy-3-(N ;N-dimethyl)aminopropane (DLinDMA), 2,2-dilinoley1-4-dimethylaminoethyl-[1,3]-dioxolane (DLin-KC2-DMA), and the like.

[00292] Suitable biomolecules may include affinity moieties. Suitable affinity moieties may include His5 (HHHHH) (SEQ ID NO: 34); HisX6 (HHHHHH) (SEQ ID NO: 35); c-myc (EQKLISEEDL) (SEQ ID NO:36); Flag (DYKDDDDK) (SEQ ID NO:37); StrepTag (WSHPQFEK) (SEQ ID NO:38); hemagglutinin, e.g., HA Tag (YPYDVPDYA) (SEQ ID NO:39); glutathione-S-transferase (GST); thioredoxin; cellulose binding domain, RYIRS (SEQ ID NO:40); Phe-His- His-Thr (SEQ ID NO:41); chitin binding domain; S-peptide; T7 peptide; SH2 domain; C-end RNA tag, WEAAAREACCRECCARA (SEQ ID NO:42); metal binding domains, e.g., zinc binding domains or calcium binding domains such as those from calcium-binding proteins, e.g., calmodulin, troponin C, calcineurin B, myosin light chain, recoverin, S-modulin, visinin, VILIP, neurocalcin, hippocalcin, frequenin, caltractin, calpain large-subunit. SI 00 proteins parvalbumin, calbindin D9K, calbindin D28K, and calretinin; biotin; streptavidin; MyoD; leucine zipper polypeptides; and maltose binding protein. In some embodiments, a suitable biomolecule is biotin.

[00293] In some embodiments, a biomolecule suitable for conjugating to a payload is a dimerization domain. Non-limiting examples of suitable dimerization domains include polypeptides of the following dimerization pairs:

- a) FK506 binding protein (FKBP) and FKBP;
- b) FKBP and calcineurin catalytic subunit A (CnA);
- c) FKBP and cyclophilin;
- d) FKBP and FKBP-rapamycin associated protein (FRB);
- e) gyrase B (GyrB) and GyrB;
- f) dihydrofolate reductase (DHFR) and DHFR;
- g) DmrB and DmrB;
- h) PYL and ABI;
- i) Cry2 and CIB 1 ; and
- j) GAI and GID1.

[00294] In some cases, a biomolecule suitable for conjugating to a payload is a member of a specific binding pair. Specific binding pairs include, e.g.: i) antibody-antigen; ii) cell adhesion molecule-extracellular matrix; iii) ligand-receptor; iv) biotin-avidin; and the like.

[00295] Suitable synthetic polymers include, but are not limited to, polyalkylenes such as polyethylene and polypropylene and polyethyleneglycol (PEG); poly chloroprene; polyvinyl ethers such as poly(vinyl acetate); polyvinyl halides such as poly(vinyl chloride); polysiloxanes; polystyrenes; polyurethanes; polyacrylates such as poly(methyl (meth)acrylate), poly(ethyl (meth)acrylate), poly(n-butyl (meth)acrylate), poly(isobutyl (meth)acrylate), poly(tert-butyl (meth)acrylate), poly(hexyl (meth)acrylate), poly(isodecyl (meth)acrylate), poly(lauryl (meth)acrylate), poly(phenyl (meth)acrylate), poly(methyl acrylate), poly(isopropyl acrylate), poly (isobutyl acrylate), and poly(octadecyl acrylate); polyacrylamides such as poly (acrylamide), poly (methacrylamide), poly(ethyl acrylamide), poly (ethyl methacrylamide), poly(N-isopropyl acrylamide), poly(n, iso, and tert-butyl acrylamide); and copolymers and mixtures thereof.

[00296] In some embodiments, the biomolecule is a fluorescent protein. Suitable fluorescent proteins include, but are not limited to, green fluorescent protein (GFP) or variants thereof, blue fluorescent variant of GFP (BFP), cyan fluorescent variant of GFP (CFP), yellow fluorescent variant of GFP (YFP), enhanced GFP (EGFP), enhanced CFP (ECFP), enhanced YFP (EYFP), GFPS65T, Emerald, Topaz (TYFP), Venus, Citrine, mCitrine, GFPuv, destabilized EGFP (dEGFP), destabilized ECFP (dECFP), destabilized EYFP (dEYFP), mCFPm, Cerulean, T-Sapphire, CyPet, YPet, mKO, HcRed, t-HcRed, DsRed, DsRed2, DsRed- monomer, J-Red, dimer2, t-dimer2(12), mRFPI, pocilloporin, Renilla GFP, Monster GFP, paGFP, Kaede protein and kindling protein, Phycobiliproteins and Phycobiliprotein conjugates including B-Phycoerythrin, R-Phycoerythrin and

Allophycocyanin. Other examples of fluorescent proteins include mHoneydew, mBanana, mOrange, dTomato, tdTomato, mTangerine, mStrawberry, mCherry, mGrapel, mRaspberry, mGrape2, mPlum (Shaner et al. (2005) Nat. Methods 2:905-909), and the like. Any of a variety of fluorescent and colored proteins from Anthozoan species, as described in, e.g., Matz et al. (1999) Nature Biotechnol. 17:969-973, is suitable for use.

[00297] In some embodiments, the biomolecule is a nucleic acid. In some embodiments, the nucleic acid is a DNA molecule. In some embodiments, the nucleic acid is an RNA molecule. In some embodiments, the nucleic acid comprises both deoxyribonucleotides and ribonucleotides. In some embodiments, the nucleic acid is a single-stranded DNA molecule. In some embodiments, the nucleic acid is a double-stranded DNA molecule. In some embodiments, the nucleic acid is a single-stranded RNA molecule. Suitable nucleic acids include, e.g., a small interfering RNA (siRNA), a short hairpin RNA (shRNA), a microRNA (miRNA), a ribozyme, an aptamer, and the like. Suitable nucleic acids include nucleic acids that are or act as siRNAs or other RNA interference reagents (RNAi agents or iRNA agents), shRNA, antisense oligonucleotides, self-cleaving RNAs, ribozymes, fragment thereof and/or variants thereof (such as Peptidyl transferase 23S rRNA, RNase P, Group I and Group II introns, GIR1 branching ribozymes, Leadzyme, Hairpin ribozymes, Hammerhead ribozymes, HDV ribozymes, Mammalian CPEB3 ribozyme, VS ribozymes, glmS ribozymes, CoTC ribozyme, etc.), microRNAs, microRNA mimics, supermirs, aptamers, antimirs, antagomirs, U1 adaptors, triplex-forming oligonucleotides, RNA activators, long non-coding RNAs, short non-coding RNAs (e.g., piRNAs), immunomodulatory oligonucleotides (such as immunostimulatory oligonucleotides, immunoinhibitory oligonucleotides), GNA, LNA, ENA, PNA, TNA, HNA, TNA, XNA, HeNA, CeNA, morpholinos, G-quadruplex (RNA and DNA), antiviral oligonucleotides, and decoy oligonucleotides. Nucleic acids can be of any length, and can include one or more of a modified ribonucleotide base, a modified deoxyribonucleotide base, a modified deoxyribose, a modified ribose, and a modified backbone linkage (e.g., a phosphorothioate linkage).

[00298] In some embodiments, the biomolecule is a viral protein. In some embodiments, the viral protein is a capsid protein, an envelope protein, or a membrane protein. In some embodiments, the viral protein is a capsid protein.

[00299] The viral protein may be derived from any virus. Examples of viruses include a parvovirus (e.g., adeno-associated viruses (AAV)), retrovirus (e.g., Retroviridae family viral vector), adenovirus (e.g., Ad5, Ad26, Ad34, Ad35, and Ad48), coronavirus, negative strand RNA viruses such as orthomyxovirus (e.g., influenza virus), rhabdovirus (e.g., rabies and

vesicular stomatitis virus), paramyxovirus (e.g. measles and Sendai), positive strand RNA viruses, such as picornavirus and alphavirus, and double stranded DNA viruses including adenovirus, herpesvirus (e.g., Herpes Simplex virus types 1 and 2, Epstein-Barr virus, cytomegalovirus), and poxvirus (e.g., vaccinia, modified vaccinia Ankara (MVA), fowlpox and canarypox). Other viruses include Norwalk virus, togavirus, flavivirus, reoviruses, papovavirus, hepadnavirus, human papilloma virus, human foamy virus, and hepatitis virus, for example. Examples of retroviruses are avian leukosis-sarcoma, avian C-type viruses, mammalian C-type, B-type viruses, D-type viruses, oncoretroviruses, HTLV-BLV group, lentivirus, alpharetrovirus, gammaretrovirus, spumavirus (Coffin, J. M., *Retroviridae: The viruses and their replication*, Virology, Third Edition (Lippincott-Raven, Philadelphia, (1996))). Other examples are murine leukemia viruses, murine sarcoma viruses, murine mammary tumor virus, bovine leukemia virus, feline leukemia virus, feline sarcoma virus, avian leukemia virus, human T-cell leukemia virus, baboon endogenous virus, Gibbon ape leukemia virus, nucleic acid molecule Pfizer monkey virus, simian immunodeficiency virus, simian sarcoma virus, Simian virus 40 (SV40), Rous sarcoma virus and lentiviruses.

AAV Vectors

[00300] In some embodiments, the payload is a viral protein from a recombinant AAV (rAAV) vectors. rAAV vectors may have recombinant nucleic acid constructs that contain (1) a nucleic acid molecule and (2) nucleic acids that facilitate and expression of the heterologous genes. The viral nucleic acids may include those sequences of AAV that are required in cis for replication and packaging (e.g., functional ITRs) of the DNA into a virion. Such rAAV vectors may also contain marker or reporter genes.

[00301] rAAV vectors include those having one or more of the naturally-occurring AAV genes deleted in whole or in part, but retain functional flanking ITR sequences. The AAV ITRs may be of any serotype (e.g., derived from serotype 2 or 5) suitable for a particular application. Methods for using rAAV vectors are described, for example, in Tal et al., *J. Biomed. Sci.* 7:279-291 (2000), and Monahan and Samulski, *Gene Delivery* 7:24-30 (2000).

[00302] The capsid proteins of AAV compose the exterior, non-nucleic acid portion of the virion and are encoded by the AAV cap gene. The cap gene encodes three viral coat proteins, VP1, VP2 and VP3, which are required for virion assembly. The construction of rAAV virions has been described, for example, in US Patent Nos. 5,173,414; 5,139,941; 5,863,541; 5,869,305; 6,057,152; and 6,376,237; as well as in Rabinowitz et al., *J. Virol.* 76:791-801 (2002) and Bowles et al., *J. Virol.* 77:423-432 (2003), the disclosures of each of which are incorporated herein by reference as they pertain to AAV vectors for gene delivery.

[00303] rAAV virions described herein include those derived from a variety of AAV serotypes including AAV 1, 2, 3, 4, 5, 6, 7, 8 and 9. Construction and use of AAV vectors and AAV proteins of different serotypes are described, for example, in Chao et al., *Mol. Ther.* 2:619-623 (2000); Davidson et al., *Proc. Natl. Acad. Sci. USA* 97:3428-3432 (2000); Xiao et al., *J. Virol.* 72:2224-2232 (1998); Halbert et al., *J. Virol.* 74:1524-1532 (2000); Halbert et al., *J. Virol.* 75:6615-6624 (2001); and Auricchio et al., *Hum. Molec. Genet.* 10:3075-3081 (2001), the disclosures of each of which are incorporated herein by reference as they pertain to AAV vectors for gene delivery.

[00304] Also described herein are pseudotyped rAAV vectors. Pseudotyped vectors include AAV vectors of a given serotype pseudotyped with a capsid gene derived from a serotype other than the given serotype (e.g., AAV1, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, or AAV9, among others). For example, a representative pseudotyped vector is an AAV2 vector encoding a therapeutic protein pseudotyped with a capsid gene derived from AAV serotype 8 or AAV serotype 9. Techniques involving the construction and use of pseudotyped rAAV virions are known in the art and are described, for example, in Duan et al., *J. Virol.* 75:7662-7671 (2001); Halbert et al., *J. Virol.* 74:1524-1532 (2000); Zolotukhin et al., *Methods*, 28:158-167 (2002); and Auricchio et al., *Hum. Molec. Genet.*, 10:3075-3081 (2001).

[00305] AAV virions that have mutations within the virion capsid may be used to infect particular cell types more effectively than non-mutated capsid virions. For example, suitable AAV mutants may have ligand insertion mutations for the facilitation of targeting AAV to specific cell types. The construction and characterization of AAV capsid mutants including insertion mutants, alanine screening mutants, and epitope tag mutants is described in Wu et al., *J. Virol.* 74:8635-45 (2000).

[00306] Other rAAV virions include those capsid hybrids that are generated by molecular breeding of viruses as well as by exon shuffling. See, e.g., Soong et al., *Nat. Genet.*, 25:436-439 (2000) and Kolman and Stemmer, *Nat. Biotechnol.* 19:423-428 (2001).

[00307] In some embodiments, the AAV the viral protein is from includes AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV 11, AAV 12, Rh10, Rh74 and AAV-2i8, AAV3B, an AAV3B variant, AAVcy10, AAVrh10, AAVrh74, AAVdj, AAV-Anc80, as well as variants (e.g., capsid variants, such as amino acid insertions, additions, substitutions and deletions) thereof, for example, as set forth in WO 2013/158879 (International Application PCT/US2013/037170), WO 2015/013313 (International Application PCT/US2014/047670) and US 2013/0059732 (US Patent No. 9, 169,299,

discloses LK01, LK02, LK03, etc.). In some embodiments, the viral protein is a capsid protein. In some embodiments, the capsid protein is a VP1, VP2, or VP3 capsid.

[00308] In some embodiments, the biomolecule is an oligonucleotide. In some embodiments, an oligonucleotide can include an oligonucleotide complementary to a gene sense sequence, a pre- mRNA sense sequence, and/or mRNA sense sequence, or a portion thereof. In some embodiments, an oligonucleotide can include an oligonucleotide of a gene sense sequence, a pre- mRNA sense sequence, and/or mRNA sense sequence, or a portion thereof. In some embodiments, oligonucleotides described herein can also be nucleotide chemical analog-based compounds capable of binding to a gene sense sequence, a pre-mRNA sense sequence, and/or an mRNA sense sequence, or a portion thereof. In some embodiments, the oligonucleotide is a sense oligonucleotide. In some embodiments, the oligonucleotide is an antisense oligonucleotide. In some embodiments, the antisense oligonucleotide can be a single-stranded nucleic acid molecule.

[00309] In some embodiments, the oligonucleotide may be an oligonucleotide sequence of 5 to 100 nucleotides in length, for example, 10 to 40 nucleotides in length, for example, 14 to 40 nucleotides in length, for example, 10 to 30 nucleotides in length, for example, 14 to 30 nucleotides in length, for example, 14 to 25 nucleotides in length, for example, 15 to 22 nucleotides in length, for example, 16 to 40 nucleotides in length, for example, 18 to 24 nucleotides in length, for example 20 to 40 nucleotides in length, or for example, 20 to 24 nucleotides in length.. An oligonucleotide may comprise an oligonucleotide sequence complementary to one or more than one portion of an mRNA sequence.

[00310] In some embodiments, an oligonucleotide comprises one or more ribonucleotides, one or more deoxyribonucleotides, or a mixture of ribonucleotides and deoxyribonucleotides.

[00311] In some embodiments, an oligonucleotide comprises one or more modified nucleosides, for example, 5-methylcytidine, 5-methyl-2'-deoxycytidine, deoxycytidine, 5-methyl-2'-deoxycytidine 5'-monophosphate, or 5-methyl-2'-deoxycytidine-5'-monophosphorothioate. In certain embodiments, an oligonucleotide comprises one or more modified nucleosides, for example, 2'-O-methylcytidine, 2'-O-methylguanosine, 2'-O-methylthymidine, 2'-O-methyluridine, or 2'-O-methyladenosine. In some embodiments, an oligonucleotide comprises one or more modified nucleotide, for example, 5-methyl cytosine or 5-methylguanine. In some embodiments, an oligonucleotides comprises one or more modified nucleotides, for example, 2'-O- (2-methoxyethyl) nucleosides, 2'-deoxy-2'-fluoro nucleosides, or 2'-fluoro- β -D-arabinonucleosides.

[00312] In some embodiments, an oligonucleotide comprises bridged nucleic acids, locked nucleic acids (LNA), constrained ethyl (cET) nucleic acids, tricyclo-DNAs (tcDNA), 2'-O,4'-C-ethylene linked nucleic acids (ENA), or peptide nucleic acids (PNA).

[00313] In some embodiments, an oligonucleotide may have a modified linkage, such as a phosphorothioate linkage, a phosphorodithioate linkage, a phosphotriester linkage, an alkylphosphonate linkage, an aminoalkylphosphotriester linkage, an alkylene phosphonate linkage, a phosphinate linkage, a phosphoramidate linkage, a phosphoromorpholidate linkage, a phosphoropiperazidate linkage, and an aminoalkylphosphoramidate linkage, a thiophosphoramidate linkage, a thionoalkylphosphonate linkage, a thionoalkylphosphotriester linkage, a thiophosphate linkage, a selenophosphate linkage, and/or a boranophosphate linkage.

Cell Surface receptors

[00314] In some embodiments, the polypeptide is a cell surface receptor (CSR). A CSR is a protein that is expressed on the surface of a cell and typically includes a transmembrane domain or other moiety that anchors it to the surface of a cell. As a receptor it binds to ligands that mediate or participate in an activity of the cell surface receptor, such as signal transduction or ligand internalization. In some embodiments, a cell surface receptors include, but are not limited to, single transmembrane receptors and G-protein coupled receptors. Receptor tyrosine kinases, such as growth factor receptors, also are among such cell surface receptors.

[00315] In some embodiments, the CSR is a tyrosine kinase. In some embodiments, a receptor tyrosine kinase (RTK) refers to a protein, typically a glycoprotein, that is a member of the growth factor receptor family of proteins. Growth factor receptors are typically involved in cellular processes including cell growth, cell division, differentiation, metabolism and cell migration. RTKs also are known to be involved in cell proliferation, differentiation and determination of cell fate as well as tumor growth. RTKs have a conserved domain structure including an extracellular domain, a membrane-spanning (transmembrane) domain and an intracellular tyrosine kinase domain. Typically, the extracellular domain binds to a polypeptide growth factor or a cell membrane-associated molecule or other ligand. The tyrosine kinase domain is involved in positive and negative regulation of the receptor.

Payloads of the Disclosure

[00316] Described herein, in certain embodiments, are conjugates comprising one or more payloads.

[00317] In some embodiments, the conjugates comprise multiple payloads. In some embodiments, the conjugates comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more than 10 payloads. In some embodiments, the conjugates comprise 1 payload. In some embodiments, the conjugates comprise 2 payloads. In some embodiments, the conjugates comprise 3 payloads. In some embodiments, the conjugates comprise 4 payloads. In some embodiments, the conjugates comprise 5 payloads. In some embodiments, the conjugates comprise 6 payloads. In some embodiments, the conjugates comprise 7 payloads. In some embodiments, the conjugates comprise 8 payloads. In some embodiments, the conjugates comprise 9 payloads. In some embodiments, the conjugates comprise 10 payloads. In some embodiments, the conjugates comprise multiple payloads conjugated to different chains of an antibody (e.g., a light chain and a heavy chain).). In some embodiments, the conjugates comprise multiple payloads that are the same. In some embodiments, the conjugates comprise multiple payloads that are different. In some embodiments, the conjugates comprise multiple payloads that are different, wherein more than one of each of the different payloads are the same. For example, the conjugates comprise two tubulin inhibitor payloads and two topoisomerase inhibitor payloads.

[00318] The payloads can be naturally occurring, or may be synthetically or recombinantly produced, and may be isolated, substantially purified, or present within the native milieu of the unmodified molecule upon which the moiety-containing payloads is based (e.g., on a cell surface or within a cell, including within a host animal, e.g., a mammalian animal, such as a murine host (e.g., rat, mouse), hamster, canine, feline, bovine, swine, and the like). In some embodiments, the payloads are present *in vitro* in a cell-free reaction. In other embodiments, the payloads are present in a cell and/or displayed on the surface of a cell. In many embodiments of interest, the payloads are in a living cell; on the surface of a living cell; in a living organism, e.g., in a living multicellular organism. Suitable living cells include cells that are part of a living multicellular organism; cells isolated from a multicellular organism; immortalized cell lines; and the like.

[00319] The payloads may be composed of D-amino acids, L -amino acids, or both, and may be further modified, either naturally, synthetically, or recombinantly, to include other moieties. For example, the payloads may be lipoproteins, glycoproteins, or other such modified proteins.

[00320] In some embodiments, the payloads comprise at least one thiol moiety for reaction with at least one polypeptide comprising a reactive moiety, but may comprise 2 or more, 3 or more, 5 or more, 10 or more thiol moieties. The number of thiol moieties that may be present

in a target molecule will vary according to the intended application of the modified target molecule of the reaction, the nature of the target molecule itself, and other considerations which will be readily apparent to the ordinarily skilled artisan in practicing the methods as disclosed herein.

[00321] The payloads can be modified to comprise a thiol moiety at the point at which linkage to the polypeptides comprising a reactive moiety is desired. For example, when the payload is a peptide or a polypeptide, the payload may be modified to contain an N-terminal thiol moiety, thereby producing a subject target peptide or polypeptide comprising a thiol moiety. It will be understood that any convenient location on a peptide or a polypeptide substrate may be modified to contain a thiol moiety and thereby produce a target peptide or polypeptide for use in the subject methods.

[00322] In some embodiments, the payloads are small molecules.

[00323] The payloads comprising a reactive moiety will in some embodiments comprise a small molecule drug, toxin, or other molecule for delivery to a cell. The small molecule drug, toxin, or other molecule will in some embodiments provide for a pharmacological activity. The small molecule drug, toxin, or other molecule will in some embodiments serve as a target for delivery of other molecules.

[00324] Small molecule drugs may be small organic or inorganic compounds having a molecular weight of more than 50 Daltons and less than about 2,500 Daltons. Small molecule drugs may comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and may include at least an amine, carbonyl, hydroxyl or carboxyl group, and may contain at least two of the functional chemical groups. The drugs may comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Small molecule drugs are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof.

[00325] In some embodiments, the payloads are cancer chemotherapeutic agents, which may include small molecules as described herein.

[00326] Suitable cancer chemotherapeutic agents include, e.g., alkylating agents, such as nitrogen mustards (for example, chlorambucil, chlormethine, cyclophosphamide, ifosfamide, and melphalan); nitrosoureas (for example, carmustine, fotemustine, lomustine, and streptozocin); platinum compounds (for example, carboplatin, cisplatin, oxaliplatin, and BBR3464); busulfan; dacarbazine; mechlorethamine; procarbazine; temozolomide; thiopeta; uramustine; antimetabolites, such as folic acid (for example, methotrexate, pemetrexed, and

raltitrexed); purine (for example, cladribine, clofarabine, fludarabine, mercaptopurine, and tioguanine); pyrimidine (for example, capecitabine); cytarabine; fluorouracil; gemcitabine; plant alkaloids, such as podophyllum (for example, etoposide, and teniposide), taxane (for example, docetaxel and paclitaxel), vinca (for example, vinblastine, vincristine, vindesine, and vinorelbine); cytotoxic/antitumor antibiotics, such as anthracycline family members (for example, daunorubicin, doxorubicin, epirubicin, idarubicin, mitoxantrone, and valrubicin), bleomycin, rifampicin, hydroxyurea, and mitomycin; topoisomerase inhibitors, such as topotecan and irinotecan; photosensitizers, such as aminolevulinic acid, methyl aminolevulinate, porfimer sodium, and verteporfin; and other agents, such as alitretinoin, altretamine, amsacrine, anagrelide, arsenic trioxide, asparaginase, axitinib, bexarotene, bevacizumab, bortezomib, celecoxib, denileukin diftitox, erlotinib, estramustine, gefitinib, hydroxycarbamide, imatinib, lapatinib, pazopanib, pentostatin, masoprocol, mitotane, pegaspargase, tamoxifen, sorafenib, sunitinib, emurafenib, vandetanib, and tretinoin. In some embodiments, the target molecule is an antibody; and the biomolecule is a cancer chemotherapeutic agent. In some embodiments, the target molecule is an antibody; and the biomolecule is a cancer chemotherapeutic agent. In some embodiments, the target molecule is a topoisomerase inhibitor. In some embodiments, the cancer chemotherapeutic agent is a tubulin inhibitor. In some embodiments, the conjugates comprise two cancer chemotherapeutic agents. In some embodiments, the two cancer chemotherapeutic agents are a topoisomerase inhibitor and a tubulin inhibitor.

[00327] In some embodiments, a payload comprising a reactive moiety is selected from the group consisting of deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan, lurototecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimmitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4), pyrolodiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT. In some embodiments, a payload comprising a reactive moiety is selected from the group consisting of exatecan, auristatin (MMAE, MMAF, MMAG, MMAH), doxorubicin, and camptothecin (CPT). In some embodiments, a payload comprising a reactive moiety comprises an immune stimulating compound. In some embodiments, a payload comprising a reactive moiety comprises a DNA repair inhibitor.

[00328] [00258] In some embodiments, a payload comprising a reactive moiety is selected from the group consisting of STING agonists including cyclic di-GMP, diABZI and derivatives thereof, TLR7 agonists, TLR8 agonists, and other agonists of immune receptors.

[00329] In another embodiment, a payload comprising a reactive moiety comprises one of a pair of binding partners (e.g., a ligand; a ligand-binding portion of a receptor; an antibody; an antigen-binding fragment of an antibody; an antigen; a hapten; a lectin; a lectin binding carbohydrate). For example, the payload can comprise a polypeptide that serves as a viral receptor and, upon binding with a viral envelope protein or viral capsid protein, facilitates attachment of virus to the cell surface on which the biomolecule is displayed.

[00330] In some embodiments, the payload is a viral protein. In some embodiments, the viral protein is a capsid protein, an envelope protein, or a membrane protein. In some embodiments, the viral protein is a capsid protein.

[00331] In some embodiments, the payload is a nucleic acid. In some embodiments, the nucleic acid is a DNA molecule. In some embodiments, the nucleic acid is an RNA molecule. In some embodiments, the nucleic acid comprises both deoxyribonucleotides and ribonucleotides. In some embodiments, the nucleic acid is a single-stranded DNA molecule. In some embodiments, the nucleic acid is a double-stranded DNA molecule. In some embodiments, the nucleic acid is a single-stranded RNA molecule. Suitable nucleic acids include, e.g., a small interfering RNA (siRNA), a short hairpin RNA (shRNA), a microRNA (miRNA), a ribozyme, an aptamer, and the like. Suitable nucleic acids include nucleic acids that are or act as siRNAs or other RNA interference reagents (RNAi agents or iRNA agents), shRNA, antisense oligonucleotides, self-cleaving RNAs, ribozymes, fragment thereof and/or variants thereof (such as Peptidyl transferase 23S rRNA, RNase P, Group I and Group II introns, GIR1 branching ribozymes, Leadzyme, Hairpin ribozymes, Hammerhead ribozymes, HDV ribozymes, Mammalian CPEB3 ribozyme, VS ribozymes, glmS ribozymes, CoTC ribozyme, etc.), microRNAs, microRNA mimics, supermirs, aptamers, antimirs, antagomirs, U1 adaptors, triplex-forming oligonucleotides, RNA activators, long non-coding RNAs, short non-coding RNAs (e.g., piRNAs), immunomodulatory oligonucleotides (such as immunostimulatory oligonucleotides, immunoinhibitory oligonucleotides), GNA, LNA, ENA, PNA, TNA, HNA, TNA, XNA, HeNA, CeNA, morpholinos, G-quadruplex (RNA and DNA), antiviral oligonucleotides, and decoy oligonucleotides. Nucleic acids can be of any length, and can include one or more of a modified ribonucleotide base, a modified deoxyribonucleotide base, a modified deoxyribose, a modified ribose, and a modified backbone linkage (e.g., a phosphorothioate linkage).

[00332] In some embodiments, the payload is an oligonucleotide. In some embodiments, an oligonucleotide can include an oligonucleotide complementary to a gene sense sequence, a pre- mRNA sense sequence, and/or mRNA sense sequence, or a portion thereof. In some embodiments, an oligonucleotide can include an oligonucleotide of a gene sense sequence, a pre- mRNA sense sequence, and/or mRNA sense sequence, or a portion thereof. In some embodiments, oligonucleotides described herein can also be nucleotide chemical analog-based compounds capable of binding to a gene sense sequence, a pre-mRNA sense sequence, and/or an mRNA sense sequence, or a portion thereof. In some embodiments, the oligonucleotide is a sense oligonucleotide. In some embodiments, the oligonucleotide is an antisense oligonucleotide. In some embodiments, the antisense oligonucleotide can be a single-stranded nucleic acid molecule.

[00333] In some embodiments, the oligonucleotide may be an oligonucleotide sequence of 5 to 100 nucleotides in length, for example, 10 to 40 nucleotides in length, for example, 14 to 40 nucleotides in length, for example, 10 to 30 nucleotides in length, for example, 14 to 30 nucleotides in length, for example, 14 to 25 nucleotides in length, for example, 15 to 22 oligonucleotides in length, for example, 18 to 40 nucleotides in length, for example, 18 to 24 nucleotides in length, for example 20 to 40 nucleotides in length, or for example, 20 to 24 nucleotides in length. An oligonucleotide may comprise an oligonucleotide sequence complementary to one or more than one portion of an mRNA sequence.

[00334] In some embodiments, an oligonucleotide comprises one or more ribonucleotides, one or more deoxyribonucleotides, or a mixture of ribonucleotides and deoxyribonucleotides.

[00335] In some embodiments, an oligonucleotide comprises one or more modified nucleosides, for example, 5-methylcytidine, 5-methyl-2'-deoxycytidine, deoxycytidine, 5-methyl-2'-deoxycytidine 5'-monophosphate, or 5-methyl-2'-deoxycytidine-5'-monophosphorothioate. In certain embodiments, an oligonucleotide comprises one or more modified nucleosides, for example, 2'-O-methylcytidine, 2'-O-methylguanosine, 2'-O-methylthymidine, 2'-O-methyluridine, or 2'-O-methyladenosine. In some embodiments, an oligonucleotide comprises one or more modified nucleotide, for example, 5-methyl cytosine or 5-methylguanine. In some embodiments, an oligonucleotides comprises one or more modified nucleotides, for example, 2'-O- (2-methoxyethyl) nucleosides, 2'-deoxy-2'-fluoro nucleosides, or 2'-fluoro- β -D-arabinonucleosides.

[00336] In some embodiments, an oligonucleotide comprises bridged nucleic acids, locked nucleic acids (LNA), constrained ethyl (cET) nucleic acids, tricyclo-DNAs (tcDNA), 2'-O,4'-C-ethylene linked nucleic acids (ENA), or peptide nucleic acids (PNA).

[00337] In some embodiments, an oligonucleotide may have a modified linkage, such as a phosphorothioate linkage, a phosphorodithioate linkage, a phosphotriester linkage, an alkylphosphonate linkage, an aminoalkylphosphotriester linkage, an alkylene phosphonate linkage, a phosphinate linkage, a phosphoramidate linkage, a phosphoromorpholidate linkage, a phosphoropiperazidate linkage, and an aminoalkylphosphoramidate linkage, a thiophosphoramidate linkage, a thionoalkylphosphonate linkage, a thionoalkylphosphotriester linkage, a thiophosphate linkage, a selenophosphate linkage, and/or a boranophosphate linkage.

[00338] In some embodiments, a payload is a peptide. In some embodiments, the peptide is a peptide ligand or a binding peptide. In some embodiments, the binding peptide can be of different origins, e.g., synthetic, human, mouse, or rat. In some embodiments, a binding peptide may be or have been engineered to include one or more (e.g., two, three, four, or five) solvent-exposed cysteine or lysine residues, which may provide a site for conjugation. In some embodiments, the binding peptides may include only naturally occurring amino acid residues, or may include one or more non-naturally occurring amino acid residues. In some embodiments, binding peptides may be linear or cyclic. In some embodiments, the peptide ligand or binding peptide is a bicyclic peptide. In some embodiments, a binding peptide can be a monospecific peptide or multi-specific peptide (e.g., a bispecific peptide or a trispecific peptide). Bispecific peptide ligands or dual specific peptide ligands may bind one target at a time or bind two targets simultaneously. In some embodiments, a binding peptide is a cysteine motif binding peptide.

Tyrosinase Polypeptides

[00339] Tyrosinase polypeptides that are suitable for use in generating a reactive moiety (e.g., an orthoquinone) include a tyrosinase polypeptide having at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%, amino acid sequence identity to any one of the tyrosinase amino acid sequences set forth any one of in SEQ ID NOs: 1-6, 43-44, and 53. In some embodiments, the tyrosinase polypeptide is an *Agricus bisporus* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Bacillus megaterium* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Streptomyces castaneoglobisporus* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Citrobacter freundii* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Homo sapiens* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Malus domestica* tyrosinase polypeptide.

In some embodiments, the tyrosinase polypeptide is an *Aspergillus oryzae* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Solanum lycopersicum* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Burkholderia thailandensis* tyrosinase polypeptide. In some embodiments, the tyrosinase polypeptide is a *Juglans regia* tyrosinase polypeptide. See, e.g., Pretzler et al. Sci. Rep. 2017, 7 (1), 1810; Ren et al. BMC Biotechnol. 2013, 13, 18; Faccio et al. Process Biochem. 2012, 47 (12), 1749-1760; Fairhead et al. FEBS J. 2010, 277 (9), 2083-2095; Do et al. Sci. Rep. 2017, 7 (1), 17267; Elsayed and Danial J. Appl. Pharm. Sci. 2018, 8 (09), 93-101; Lopez-Tejedor and Palomo Protein Expr. Purif. 2018, 145, 64-70; and Fairhead et al. Nature Biotechnol. 2012, 29 (2), 183-191.

[00340] In some embodiments, the tyrosinase polypeptide selectively acts on (e.g., generates a reactive moiety such as an orthoquinone) a substrate (a biomolecule or polypeptide) comprising a phenol moiety (e.g., a tyrosine) or a catechol moiety, where the substrate is neutral or positively charged within 50 Å (e.g., within 50 Å, within 40 Å, within 30 Å, or within 20 Å) of the phenol or the catechol moiety. For example, a tyrosinase having at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%, amino acid sequence identity to any one of the tyrosinase amino acid sequences set forth in any one of SEQ ID NOs: 2 and 4-6 can selectively modify a phenol or catechol moiety on a substrate, where the substrate is neutral or positively charged within 50 Å (e.g., within 50 Å, within 40 Å, within 30 Å, or within 20 Å) of the phenol or the catechol moiety. In some embodiments, at least one of the at least two polypeptides of a conjugate comprises at least 2 neutral or positively charged amino acids within 10 amino acids of the phenol moiety (e.g., a tyrosine) or a catechol moiety. In some embodiments, at least one of the at least two polypeptides of a conjugate comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 neutral or positively charged amino acids within 10 amino acids of the phenol moiety (e.g., a tyrosine) or a catechol moiety. In some embodiments, at least one of the at least two polypeptides of a conjugate comprises the amino acid sequence GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGGGY (SEQ ID NO: 13), EEEGY (SEQ ID NO: 14), EEEYY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWEY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 30), KKGYY (SEQ ID

NO: 31), and DDGGY (SEQ ID NO: 32). In some embodiments, the polypeptide of a conjugate comprises at least 2 neutral or positively charged amino acids within 10 amino acids of the phenol moiety (e.g., a tyrosine) or a catechol moiety. In some embodiments, the polypeptide of a conjugate comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 neutral or positively charged amino acids within 10 amino acids of the phenol moiety (e.g., a tyrosine) or a catechol moiety. In some embodiments, the polypeptide of a conjugate comprises a non-terminal tag comprising $X_1X_2YX_4X_5$. In some embodiments, X_1 and X_2 are proton-donating amino acids. In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N. In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G. In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L. In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7. In some embodiments, the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.

[00341] In some embodiments, the tyrosinase polypeptide selectively acts on (e.g., generates a reactive moiety such as an orthoquinone) a substrate (a biomolecule) comprising a phenol moiety (e.g., a tyrosine) or a catechol moiety, where the substrate is negatively charged within 50 Å (e.g., within 50 Å, within 40 Å, within 30 Å, or within 20 Å) of the phenol or the catechol moiety. In some embodiments, a tyrosinase having at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or 100%, amino acid sequence identity to any one of the tyrosinase amino acid sequences set forth in any one of SEQ ID NOs: 2-3 can selectively modify a phenol or catechol moiety on a substrate, where the substrate is negatively charged within 50 Å (e.g., within 50 Å, within 40 Å, within 30 Å, or within 20 Å) of the phenol or the catechol moiety. In some embodiments, at least one of the at least two polypeptides of a conjugate comprises at least 2 negatively charged amino acids within 10 amino acids of the phenol moiety (e.g., a tyrosine) or a catechol moiety. In some embodiments, at least one of the at least two polypeptides of a conjugate comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 negatively charged amino acids within 10 amino acids of the phenol moiety (e.g., a tyrosine) or a catechol moiety. In some embodiments, at least one of the at least two polypeptides of a conjugate comprises the amino acid sequence GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGGY (SEQ ID NO: 13), EEEGY (SEQ ID NO: 14), EEEEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRWY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWY (SEQ ID NO: 21), DDDDY (SEQ ID NO:

22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 30), KKGGY (SEQ ID NO: 31), and DDGGY (SEQ ID NO: 32).

Methods

[00342] Described herein, in certain embodiments, are methods of covalently linking at least two polypeptides to at least two payloads. In some embodiments, the methods comprise contacting a first polypeptide of the at least two polypeptides and a first payload of the at least two payloads using a first tyrosinase, wherein the first polypeptide comprises a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0; and contacting a second polypeptide of the at least two polypeptides and a second payload of the at least two payloads using a second tyrosinase, wherein the second polypeptide comprises a second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, and wherein $m2$ is an integer greater or equal to 0, and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag. In some embodiments, the first tag is a non-terminal tag and the second tag is a terminal tag. In some embodiments, the second tag is a non-terminal tag and the first tag is a terminal tag. In some embodiments, the first tag and the second tag are non-terminal tags. The contacting is carried out under conditions sufficient for conjugation of the at least two payloads to the at least two polypeptides, thereby producing a conjugate.

[00343] In some embodiments, X_1 is a tyrosine. In some embodiments, X_2 is a tyrosine. In some embodiments, X_3 is a tyrosine. In some embodiments, X_4 is a tyrosine. In some embodiments, X_5 is a tyrosine.

[00344] In some embodiments, X_6 is a tyrosine. In some embodiments, X_7 is a tyrosine. In some embodiments, X_8 is a tyrosine. In some embodiments, X_9 is a tyrosine. In some embodiments, X_{10} is a tyrosine.

[00345] In some embodiments, the non-terminal tag comprises $X_1X_2YX_4X_5$. In some embodiments, X_1 and X_2 are proton-donating amino acids. In some embodiments, X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N. In some embodiments, X_4 and X_5 are small, neutral amino acids. In some embodiments, X_4 is A or G. In some embodiments, X_5 is E, V, D, A, I, G, S, T, or L. In some embodiments, the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7. In some embodiments, the non-terminal tag comprises the

sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.

[00346] In some embodiments, a subject method for chemoselective modification of a payload comprises contacting: i) a payload comprising a thiol moiety; ii) a polypeptide comprising a phenol moiety or a catechol moiety; and iii) an enzyme capable of oxidizing the phenol or catechol moiety; wherein the enzyme oxidizes the phenol or catechol moiety of the polypeptide to generate a reactive moiety, thereby generating a polypeptide comprising the reactive moiety, and wherein the reactive moiety reacts with the thiol moiety, thereby conjugating the payload and the polypeptide to one another, thereby producing a conjugate. In some embodiments, the payload comprises a single thiol moiety.

[00347] In some embodiments, a subject method for chemoselective modification of a polypeptide comprises contacting: i) a payload comprising a phenol moiety or a catechol moiety; ii) a polypeptide comprising a thiol moiety; and iii) an enzyme capable of oxidizing the phenol or catechol moiety; wherein the enzyme oxidizes the phenol or catechol moiety of the payload to generate a reactive moiety, thereby generating a payload comprising the reactive moiety, and wherein the reactive moiety reacts with the thiol moiety, thereby conjugating the payload and the polypeptide to one another, thereby producing a conjugate. In some embodiments, the polypeptide comprises a single thiol moiety. In some embodiments, the polypeptide comprises a two or more thiol moieties.

[00348] In some embodiments, a subject method for chemoselective modification of a polypeptide comprises contacting: i) a payload comprising a maleimide moiety; ii) a polypeptide comprising a thiol moiety; and iii) an enzyme capable of oxidizing the maleimide moiety; wherein the enzyme oxidizes the maleimide moiety of the payload to generate a reactive moiety, thereby generating a payload comprising the reactive moiety, and wherein the reactive moiety reacts with the thiol moiety, thereby conjugating the payload and the polypeptide to one another, thereby producing a conjugate. In some embodiments, the polypeptide comprises a single thiol moiety. In some embodiments, the polypeptide comprises a two or more thiol moieties.

[00349] In some embodiments, the present disclosure provides for attachment a payload comprising a thiol moiety to a polypeptide comprising a phenol or catechol moiety. In some embodiments, the present disclosure provides for attachment for attachment of a payload comprising a phenol or catechol moiety to a polypeptide comprising a thiol moiety. In some embodiments, the present disclosure provides for attachment for attachment of a payload comprising maleimide moiety to a polypeptide comprising a thiol moiety . In some

embodiments, the methods generally involve reacting thiol containing payloads with a polypeptide comprising a reactive moiety (e.g., an orthoquinone moiety). In some embodiments, the methods generally involve reacting thiol containing polypeptides with at least one payload comprising a reactive moiety (e.g., an orthoquinone moiety). In some embodiments, the methods generally involve reacting thiol containing polypeptides with at least one payload comprising a maleimide moiety.

[00350] The methods disclosed herein provide a simple coupling procedure that can attach payloads of interest in a site-specific manner to any position on the surface of polypeptides, thereby producing a conjugate of interest. The payload can be any of a variety of molecules (e.g., polypeptides; nucleic acids; small molecules; etc.). In some embodiments, the payload is a small molecule (e.g., a cancer chemotherapeutic agent).

[00351] Polypeptides of interest include antibodies. In some instances, the polypeptide of interest is an antibody fragment or binding derivative thereof. In some embodiments, the antibody fragment or binding derivative thereof is selected from the group consisting of a Fab fragment, a F(ab')₂ fragment, a single-chain Fv (scFv), a diabody, a nanobody, and a triabody. In some embodiments, the antibody fragment or binding derivative thereof is selected from the group consisting of a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, or a constant chain of an antibody.

[00352] In some embodiments, a polypeptide further comprises one or more moieties selected from a fluorophore, an active small molecule, an affinity tag, and a metal-chelating agent. In some embodiments, the polypeptide further comprises a fluorescent protein. In some embodiments, the fluorescent protein is a green fluorescent protein (GFP). In some embodiments, the polypeptide is an enzyme. In some embodiments, the polypeptide is a receptor.

[00353] Payloads of interest include, but are not limited to, small molecules, polypeptides, polynucleotides, nucleic acids, carbohydrates, lipids, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs thereof and combinations thereof. In some embodiments, the payload is a small molecule (e.g., a cancer chemotherapeutic agent); and the polypeptide is an antibody (e.g., a scFv; a nanobody; and the like).

[00354] In some embodiments, the enzyme capable of oxidizing the phenol moiety or the catechol moiety is a phenol oxidase or a catechol oxidase. In certain cases, the enzyme is a tyrosinase.

[00355] The term "tyrosinase" is used herein to refer to monophenol monooxygenase (EC 1.14.18.1; CAS number: 9002-10-2) or a derivative thereof, an enzyme that catalyzes the

oxidation of phenols (such as tyrosine). They are a copper-containing enzymes originally present in plant and animal tissues that catalyzes the production of melanin and other pigments from tyrosine by oxidation. Tyrosinases have in common a binuclear type 3 copper center within their active site. Here two copper atoms are each coordinated with three histidine residues. Matoba et al., "Crystallographic evidence that the dinuclear copper center of tyrosinase is flexible during catalysis," J Biol Chem. 2006 Mar 31 ;281(13):8981-90. Epub 2006 Jan 25, disclose a three dimensional model of a tyrosinase catalytic center.

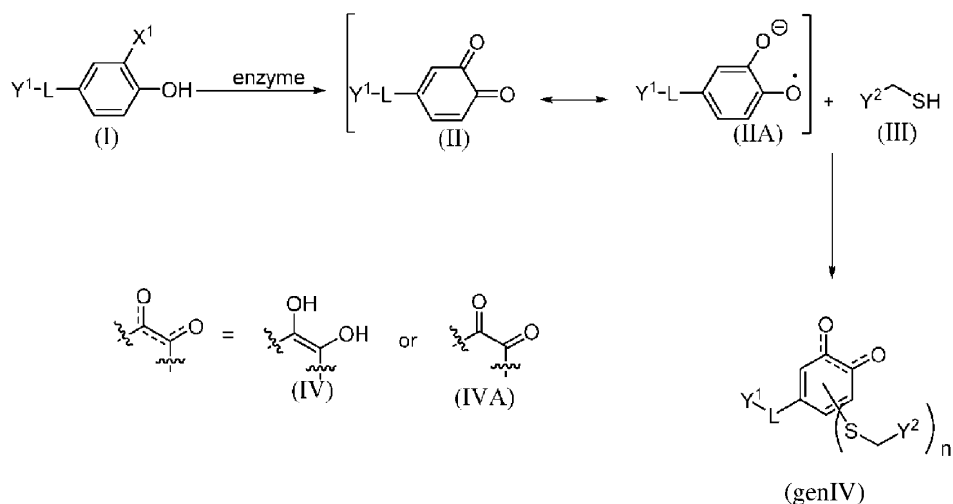
[00356] In some embodiments, the phenol moiety of a conjugate is present in a tyrosine residue. In some embodiments, the tyrosine residue is part of the polypeptide of interest. In some embodiments, the tyrosine residue is synthetically introduced into a polypeptide of a conjugate. In some embodiments, the tyrosine residue is linked to a polypeptide via a linker (e.g., as described herein). A tyrosine residue can be introduced using standard recombinant techniques, e.g., by modifying a nucleotide sequence encoding a polypeptide such that a tyrosine residue is introduced into the polypeptide.

[00357] In some embodiments, a phenol or catechol moiety is part of an unnatural (non-genetically encoded) amino acid that is introduced into a polypeptide of interest. For example, amber codon (TAG) suppression can be used to incorporate a non-genetically encoded amino acid residue that comprises a phenol moiety or a catechol moiety. See, e.g., Chin et al. (2002) J. Am. Chem. Soc. 124:9026; Chin and Schultz (2002) Chem. Biol. Chem. 3:1135; Chin et al. (2002) Proc. Natl. Acad. Sci. USA 99:11020; U.S. 2015/0240249; and US 2018/0171321. As another example, an orthogonal RNA synthetase and/or an orthogonal tRNA can be used for introducing a non- genetically encoded amino acid into a polypeptide, where the non-genetically encoded amino acid comprises a phenol moiety or a catechol moiety.

[00358] In some embodiments of the subject methods, the thiol moiety present in the payload is part of a cysteine residue. In certain cases, the cysteine residue is a native cysteine residue. In certain cases, the cysteine residue is a residue synthetically introduced into the target molecule.

[00359] In some embodiments, a polypeptide contains a thiol moiety and the payload contains a phenol or catechol moiety.

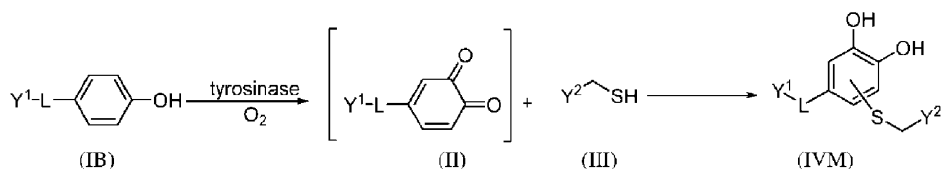
[00360] In certain embodiments, the reactive moiety is an orthoquinone or a semi-quinone radical, or a combination thereof. In certain embodiments, the subject methods provide a reaction between an orthoquinone reactive intermediate and a thiol moiety, as depicted in Scheme 1 below:



where Y^1 is a polypeptide; L is an optional linker (e.g., as described herein); X^1 is selected from hydrogen and hydroxyl; Y^2 is a payload; and n is an integer from 1 to 3.

[00361] As depicted in Scheme 1, in certain embodiments, a polypeptide comprising a phenol or catechol moiety (e.g., of formula (I)), undergoes activation with an enzyme capable of oxidizing the phenol or catechol moiety. In some embodiments, activation is achieved with a tyrosinase enzyme in the presence of oxygen to generate an intermediate comprising a reactive moiety (e.g., orthoquinone of formula (II) and/or semi-quinone radical of formula (IIA)), and the said reactive moiety reacts with a payload comprising a thiol based nucleophile (e.g., of formula (III)), to result in conjugation of the target molecule to the biomolecule, thereby producing a modified target molecule (e.g., of formula (genIV)). In certain embodiments, a payload of formula (III) may comprise any payload, e.g., as described herein. In some embodiments, Y^2 in formula (III) is a polypeptide. In some embodiments, the conjugate is described by the formula (IV). In some embodiments, the conjugate is described by the formula (IV A).

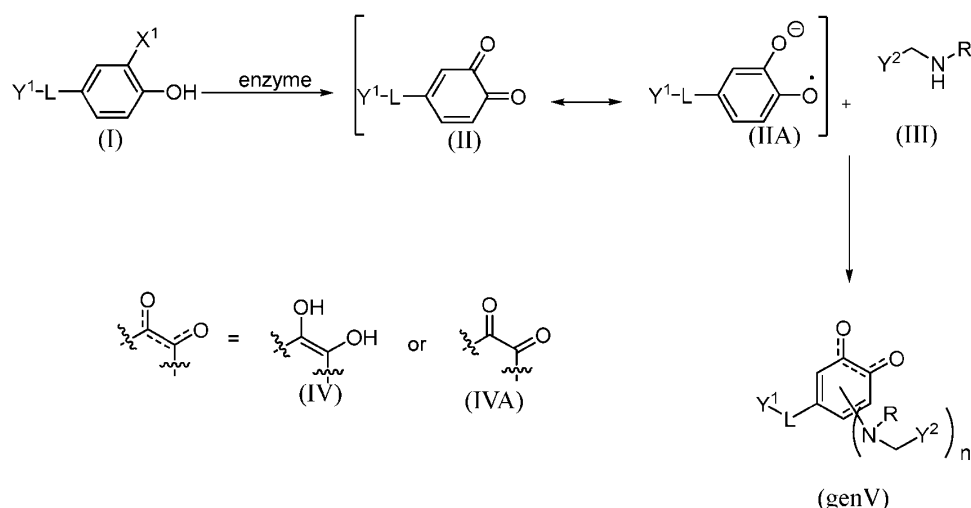
[00362] In certain embodiments, the subject methods provide a reaction between an orthoquinone reactive intermediate and a thiol moiety, as depicted in Scheme 2 below:



[00363] As depicted in Scheme 2, in certain embodiments, a polypeptide comprising a phenol moiety (e.g., of formula (IB)) undergoes activation with a tyrosinase enzyme in the

presence of oxygen to generate an intermediate comprising a reactive moiety (e.g., orthoquinone of formula (II)), and the said reactive moiety reacts with a payload comprising a thiol based nucleophile (e.g., of formula (III)), to result in conjugation of the target molecule to the biomolecule, thereby producing a modified target molecule (e.g., of formula (IVM)). In certain embodiments, a payload of formula (III) may comprise any payload, e.g., as described herein. In certain cases, Y^2 in formula (III) is a polypeptide. In certain cases of the conjugate of formula (IVM), the thiol group is at the 3-position of the catechol ring. In certain cases of the conjugate of formula (IVM), the thiol group is at the 5-position of the catechol ring. In certain cases of the conjugate of formula (IVM), the thiol group is at the 6-position of the catechol ring.

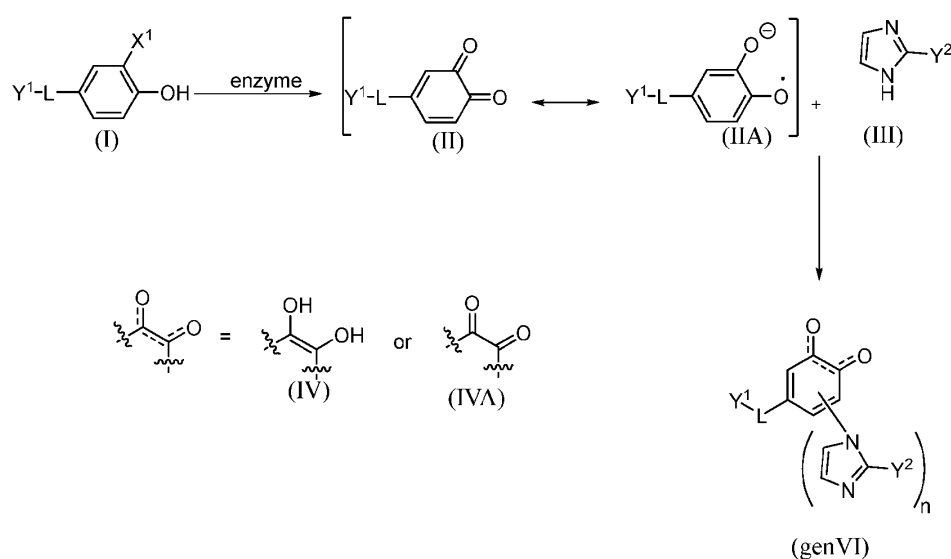
[00364] In certain embodiments, the subject methods provide a reaction between an orthoquinone reactive intermediate and an amine moiety (e.g., a lysine residue present in polypeptide), as depicted in Scheme 3 below:



where Y^1 is a polypeptide; L is an optional linker (e.g., as described herein); X^1 is selected from hydrogen and hydroxyl; Y^2 is a payload; R is selected from hydrogen, alkyl, substituted alkyl, heteroalkyl, substituted heteroalkyl, cycloalkyl, substituted cycloalkyl, heterocycloalkyl, substituted heterocycloalkyl, aryl, substituted aryl, arylalkyl, substituted arylalkyl, heteroaryl, substituted heteroaryl, heteroarylalkyl, and substituted heteroarylalkyl; and n is an integer from 1 to 3. In certain instances, R is an amino acid selected from alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, or valine.

[00365] As depicted in Scheme 3, in certain embodiments, a polypeptide comprising a phenol or catechol moiety (e.g., of formula (I)), undergoes activation with an enzyme capable of oxidizing the phenol or catechol moiety. In some embodiments, activation is achieved with a tyrosinase enzyme in the presence of oxygen to generate an intermediate comprising a reactive moiety (e.g., orthoquinone of formula (II) and/or semi-quinone radical of formula (IIA)), and the said reactive moiety reacts with a payload comprising an amine based nucleophile (e.g., of formula (III)), to result in conjugation of the target molecule to the biomolecule, thereby producing a modified target molecule (e.g., of formula (genIV)). In certain embodiments, a payload of formula (III) may comprise any payload, e.g., as described herein. In some embodiments, Y² in formula (III) is a polypeptide. In some embodiments, the conjugate is described by the formula (IV). In some embodiments, the conjugate is described by the formula (IV A).

[00366] In certain embodiments, the subject methods provide a reaction between an orthoquinone reactive intermediate and an imidazole moiety (e.g., a histidine residue present in a polypeptide), as depicted in Scheme 4 below:

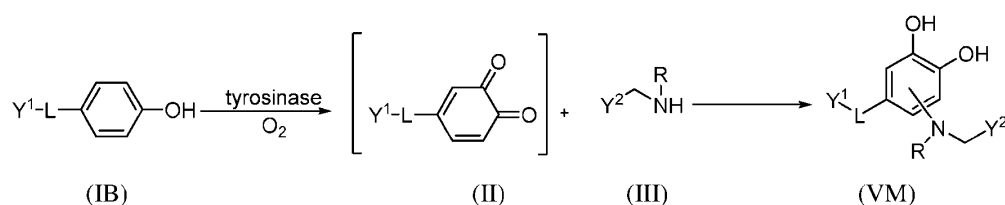


where Y¹ is a polypeptide; L is an optional linker (e.g., as described herein); X¹ is selected from hydrogen and hydroxyl; Y² is a payload; and n is an integer from 1 to 3.

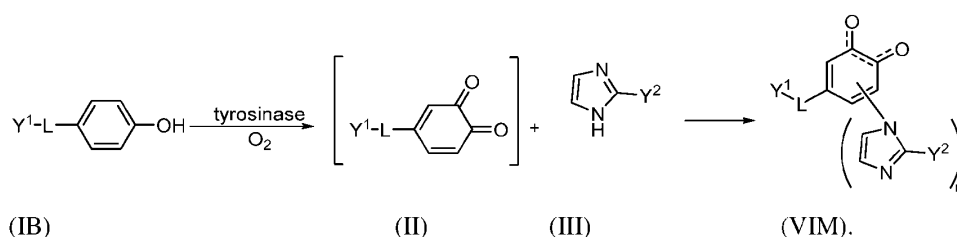
[00367] As depicted in Scheme 4, in certain embodiments, a polypeptide comprising a phenol or catechol moiety (e.g., of formula (I)), undergoes activation with an enzyme capable of oxidizing the phenol or catechol moiety. In some embodiments, activation is achieved with a tyrosinase enzyme in the presence of oxygen to generate an intermediate comprising a

reactive moiety (e.g., orthoquinone of formula (II) and/or semi-quinone radical of formula (IIA)), and the said reactive moiety reacts with a payload comprising an imidazole based nucleophile (e.g., of formula (III)), to result in conjugation of the target molecule to the biomolecule, thereby producing a modified target molecule (e.g., of formula (genIV)). In certain embodiments, a payload of formula (III) may comprise any payload, e.g., as described herein. In some embodiments, Y2 in formula (III) is a polypeptide. In some embodiments, the conjugate is described by the formula (IV). In some embodiments, the conjugate is described by the formula (IV A).

[00368] In certain embodiments, the subject methods provide a reaction between an orthoquinone reactive intermediate and an amine moiety, as depicted in Scheme 4 below:



[00369] In certain embodiments, the subject methods provide a reaction between an orthoquinone reactive intermediate and an imidazole moiety, as depicted in Scheme 5 below:



[00370] In some embodiments, the method is carried out at a pH from 4 to 9, such as 4.2, 4.5, 4.8, 5.0, 5.2, 5.5, 5.8, 6.0, 6.2, 6.5, 6.8, 7.0, 7.2, 7.5, 7.8, 8.0, 8.2, 8.5, 8.8 or 9. In some embodiments, the method is carried out at a pH of from 5 to 8, such as 5.2, 5.5, 5.8, 6.0, 6.2, 6.5, 6.8, 7.0, 7.2, 7.5, 7.8 or 8.0. In certain cases, the method is carried out at a pH of 6 to 7.5, such as 6.0, 6.3, 6.4, 6.5, 6.6, 6.8, 7.0, 7.2, 7.4, or 7.5. In some embodiments, the method is carried out at neutral pH. As used herein, the expression “neutral pH” means a pH of about 7.0 to about 7.4. The expression “neutral pH” includes pH values of about 7.0, 7.05, 7.1, 7.15, 7.2, 7.25, 7.3, 7.35, and 7.4.

[00371] In some embodiments, the methods may be carried out under physiological conditions. In some embodiments, the method is carried out on living cells in vitro. In other embodiments, the method is carried out on living cells ex vivo.

[00372] In some embodiments, the methods may be carried out in aqueous media in the presence of one or more buffers. Buffers of interest include, but are not limited to, a phosphate buffer, 2-amino-2-(hydroxymethyl)propane-1,3-diol (TRIS), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), and the like. In some embodiments, the methods may be carried out in an organic solvent. In some embodiments, the organic solvent is a water miscible solvent. In some embodiments, the organic solvent is a dipolar aprotic solvent. In some embodiments, the organic solvent is selected from acetonitrile, dimethyl formamide, methanol and acetone. In certain cases, the organic solvent is present in an amount from 1 to 20%, relative to water, such as 2%, 5%, 10%, 15% or 20%. In some embodiments, the subject method is carried out in from 1% to 20% acetonitrile, such as 5%, 10%, 15% or 20%. In some embodiments, the subject method is carried out in from 1% to 20% dimethyl formamide, such as 5%, 10%, 15%, or 20%. In some embodiments, the subject method is carried out in from 1% to 20% methanol, such as 5%, 10%, 15%, or 20%. In some embodiments, the subject method is carried out in from 1% to 20% acetone, such as 5%, 10%, 15%, or 20%.

[00373] In some embodiments, the buffer is 50 mM Phosphate pH 6.5, 150 mM NaCl, and 2 mM EDTA or 50 mM Acetate pH 5.5, 10% sucrose 1 mM EDTA. The phosphate in the first listed buffer can be replaced with HEPES, Tris, Acetate, and MES buffers, from pH 5 to pH 9. In some embodiments, sucrose and glucose can be added to the buffer up to 15% v/v, Trehalose can be added at up to 300 mM, and EDTA can be added at up to 5 mM. In some embodiments, the buffer is 100 mM Sodium Phosphate, 2 mM EDTA, pH 6.5. In some embodiments, the buffer is 100 mM Sodium Phosphate, pH 6.5.

[00374] In some embodiments, the conjugate is a product of double or triple conjugation (e.g., referring to formula (IV), when n is 2 or 3, referred to collectively herein as “multiple conjugation products”). In some embodiments, multiple conjugation products are present in less than 1 part in 10 by weight of one or more multiple conjugation products relative to the single conjugation product (e.g., referring to formula (IV), when n is 1), such as less than 1 part in 20, less than 1 part in 25, less than 1 part in 50, less than 1 part in 75, less than 1 part in 100, or even less. In some embodiments, no multiple conjugation products are observed.

[00375] In some embodiments, the conjugate is stable at a range of pH and temperature values and in the presence of a number of additional molecules. In some embodiments, the conjugate is stable from 0 °C to 50 °C, such as 4 °C to 40 °C, such as 4 °C to 37 °C. In certain cases, the conjugate is stable over a pH range of 4 to 9, such as at pH 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5 or 9. In some embodiments, the conjugate is stable in the presence of

biologically relevant molecules. In some embodiments, the conjugate is stable in the presence of molecules such as, the guanidinium group of an arginine residue, the primary amine of a lysine residue, and aniline moieties. In some embodiments, the conjugate is stable in physiological conditions; for example, In some embodiments, the conjugate is stable in human serum. In some embodiments, the conjugate is stable in human serum at 37 °C for a period of time of at least 2 days, at least 3 days, at least 4 days, at least 5 days, at least 6 days, at least 7 days, at least 10 days, or at least 14 days. In some embodiments, the conjugate is stable in human serum at 37 °C for a period of time from about 2 days to about 7 days, from about 7 days to about 10 days, or from about 10 days to about 14 days.

[00376] The present disclosure provides a method of linking at least two polypeptides to at least two payloads in a sequential manner. The method takes advantage of the substrate preferences of tyrosinase polypeptides, as described above. The method can be carried out on an insoluble substrate, i.e., an immobilized surface, such as a bead.

[00377] Thus, the present disclosure provides a method of linking at least two polypeptides to at least two payloads, the method comprising: a) contacting a first polypeptide of the at least two polypeptides and a first payload of the at least two payloads using a first tyrosinase, wherein the first polypeptide comprises a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1-X_5 is any amino acid provided that at least one amino acid of X_1-X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0; and b) contacting a second polypeptide of the at least two polypeptides and a second payload of the at least two payloads using a second tyrosinase, wherein the second polypeptide comprises a second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6-X_{10} is any amino acid provided that at least one amino acid of X_6-X_{10} is a tyrosine, and wherein $m2$ is an integer greater or equal to 0, and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag. In some embodiments, the first polypeptide comprises two or more negatively charged amino acids within ten amino acids of the phenol moiety or the catechol moiety and the second polypeptide comprises two or more neutral or positively charged within ten amino acids of the phenol moiety or the catechol moiety. In some embodiments, the first enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 75% amino acid sequence identity to any one of the amino acid sequences depicted in any one of SEQ ID NOs: 2 and 4-6. In some embodiments, the second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 75% amino acid sequence identity to any one of the amino acid sequences depicted in any one of SEQ ID NOs: 2-3.

[00378] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 1-6 or 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to any one of the amino acid of SEQ ID NOs: 1-6 or 53. In some embodiments, first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 1-6 or 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 1-6 or 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to any one of the amino acid sequences of SEQ ID NOs: 1-6 or 53.

[00379] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 2-6. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 2-6. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 2-6. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to any one of the amino acid sequences of SEQ ID NOs: 2-6. In some embodiments, the second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to any one of the amino acid sequences of SEQ ID NOs: 2-6.

[00380] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 2. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 2. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 2. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid

sequence identity to the amino acid sequence of SEQ ID NO: 2. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to amino acid of SEQ ID NO: 2.

[00381] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 3. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 3. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 3. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 3. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to amino acid of SEQ ID NO: 3.

[00382] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to amino acid of SEQ ID NO: 4.

[00383] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 5. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 5. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to the amino acid

sequence of SEQ ID NO: 5. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 5. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to amino acid of SEQ ID NO: 5.

[00384] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 6. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 6. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 6. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 4. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to amino acid of SEQ ID NO: 6.

[00385] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence having at least 99% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 53. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide comprising an amino acid sequence according to amino acid of SEQ ID NO: 53.

[00386] In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide encoded by a nucleotide sequence having at least 85% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 54. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide encoded by a nucleotide sequence having at least 90% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 54. In some

embodiments, the first and/or second enzyme is a tyrosinase polypeptide encoded by a nucleotide sequence having at least 95% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 54. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide encoded by a nucleotide sequence having at least 99% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 54. In some embodiments, the first and/or second enzyme is a tyrosinase polypeptide encoded by a nucleotide sequence having 100% amino acid sequence identity to the amino acid sequence of SEQ ID NO: 54.

[00387] In some embodiments, the phenol moiety present in the first polypeptide is present in a Tyrosine (Tyr, Y) residue. In some embodiments, the phenol moiety present in the second polypeptide is present in a Tyr residue.

[00388] By alternating use of: a) a tyrosinase enzyme that preferentially modifies a Tyr residue that is present in a negatively charged environment (e.g., where the polypeptide comprises two or more negatively charged within ten amino acids of the Tyr residue); and b) a tyrosinase enzyme that preferentially modifies a Tyr residue that is present in a neutral or positively charged environment ((e.g., where the polypeptide comprises two or more neutral or positively charged within ten amino acids of the Tyr residue), a first payload can be added to the first polypeptide and a second payload can be added to the second polypeptide in a sequential manner. In some embodiments, the first payload and the second payload are the same. In some embodiments, the first payload and the second payload are different.

[00389] In some embodiments, the tyrosinase enzyme is inactivated or removed between any two steps of the method and before adding a further tyrosinase enzyme. For example, between step (b) and step (c) of the method described above, the second enzyme is inactivated or removed.

[00390] Described herein, in certain embodiments, are methods for selectively conjugating at least two polypeptides to at least two payloads by modulating the temperature during the one or more tyrosinase reactions. In some embodiments, a first polypeptide (e.g, a heavy chain) is selectively modified through temperature control during a first tyrosinase reaction. In some embodiments, the temperature during the first tyrosinase reaction occurs at a temperature of about 0 °C to about 18 °C, such as about 0 °C to 15 °C, about 0 °C to 10 °C, about 4 °C to 15 °C, or about 4 °C to 10 °C. In some embodiments, the temperature during the first tyrosinase reaction occurs at a temperature of about 4 °C. In some embodiments, the temperature during the first tyrosinase reaction occurs at room temperature. In some embodiments, a second polypeptide (e.g, a light chain) is selectively modified through temperature control during a second tyrosinase reaction. In some embodiments, the

temperature during the second tyrosinase reaction occurs at a temperature of about 23 °C to about 50 °C, such as about 25 °C to 50 °C, about 25 °C to 40 °C or about 25 °C to 37 °C. In some embodiments, the temperature during the second tyrosinase reaction occurs at a temperature of about 30 °C. In some embodiments, the temperature during the second tyrosinase reaction occurs at a temperature of about 37 °C. In some embodiments, the temperature during the second tyrosinase reaction occurs at room temperature. In some embodiments, the same tyrosinase is used for the reaction at the different temperatures

[00391] Described herein, in certain embodiments, are methods for selectively conjugating at least two polypeptides to at least two payloads by deglycosylation of one or more polypeptides of the least two polypeptides. In some embodiments, one or more sugars of the one or more polypeptides is removed (i.e., deglycosylated). In some embodiments, at least about 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or more than 90% of the sugars of the one or more polypeptides is removed. In some embodiments, removal of the one or more sugars occurs at a temperature of about 0 °C to about 50 °C, such as about 4 °C to 40 °C or about 4 °C to 37 °C. In some embodiments, removal of the one or more sugars occurs at a temperature of about 4 °C. In some embodiments, removal of the one or more sugars occurs at a temperature of about 37 °C. In some embodiments, the deglycosylated antibodies are conjugated to one or more payloads at a temperature of about 4 °C. In some embodiments, the deglycosylated antibodies are conjugated to one or more payloads at a temperature of about 37 °C. In some embodiments, the deglycosylated antibodies are conjugated to one or more payloads at a temperature of about room temperature.

[00392] Described herein, in certain embodiments, are methods for selectively conjugating at least two polypeptides to at least two payloads, wherein conjugation occurs at less than about 10 hours, 9 hours, 8 hours, 7 hours, 6 hours, 5 hours, 4 hours, 3 hours, 2 hours, or 1 hour. In some embodiments, the conjugation occurs less than about 1 hour, 50 minutes, 40 minutes, 30 minutes, 20 minutes, 10 minutes, 5 minutes, 1 minute, or 30 seconds. In some embodiments, the conjugation occurs less than about 10 minutes.

[00393] Described herein, in certain embodiments, are methods for selectively conjugating a polypeptide to at least one payload, wherein conjugation occurs at less than about 10 hours, 9 hours, 8 hours, 7 hours, 6 hours, 5 hours, 4 hours, 3 hours, 2 hours, or 1 hour. In some embodiments, the conjugation occurs less than about 1 hour, 50 minutes, 40 minutes, 30 minutes, 20 minutes, 10 minutes, 5 minutes, 1 minute, or 30 seconds. In some embodiments, the conjugation occurs less than about 10 minutes.

EXAMPLES

EXAMPLE 1: Exemplary use a tyrosinase activatable loop

[00394] This Example describes conjugation of a polypeptide (e.g., a viral protein such as a capsid) by incorporating a tyrosinase activatable loop.

[00395] An exemplary polypeptide of interest (e.g., an antibody, a binding protein, a capsid, or a cell surface receptor) has naturally occurring tyrosine residue on a loop. A loop is a stretch of amino acids of the folded protein that is structurally accessible by, for example, an enzyme such as a tyrosinase. The polypeptide of interest is then modified to contain a non-terminal tag as described comprising an unstructured (non alpha helical, non beta sheet) loop that contains a tyrosine residue exposed to solvent in such a way as to be activated by tyrosinase enzyme. The polypeptide may not need to be modified if an amino acid sequence that satisfies the description of a non-terminal tag as described herein is already surrounding a tyrosine in the polypeptide. The inserted sequence would then allow for non-terminal activation with a tyrosinase enzyme and conjugation with a payload molecule containing a free thiol, such as a cytotoxic payload, a cytokine, a binding domain, or a nucleic acid, as described herein. This is a tyrosinase activatable loop, a stretch of amino acids containing at least a non-terminal tag containing a tyrosine residue.

[00396] In one example, the tyrosinase activatable loop is incorporated into a capsid protein or cell surface proteins to enable conjugation of binding and signaling domains with active thiols.

EXAMPLE 2: Exemplary polypeptide with a loop insert conjugated to a payload

[00397] This Example describes conjugation of a polypeptide (e.g., a fluorescent protein) to an exemplary payload by incorporating a tyrosinase activatable loop.

[00398] An exemplary polypeptide, nanoluciferase, had at least one loop that was solvent accessible, identified in FIG. 2. One loop was selected for incorporation of a tyrosine or a non-terminal tag that includes a tyrosine. The non-terminal tags tested were VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52), and NYA. The nanoluciferase constructs were cloned with routine methods known in the art and expressed using E. coli cells, such as BL21 (DE3). The expressed proteins were purified using a fast protein liquid chromatography (FPLC) instrument. The SDS-page gel and FPLC traces of the purification of each nanoluciferase construct is shown in FIGS. 3A-3B, 4A-4B, 5A-5B, 6A-6B, 7A-7B, and 8A-8B.

[00399] To perform the conjugation reaction, a 1:2 molar ratio of a nanoluciferase protein (SEQ ID NO: 55) with a non-terminal tag or a tyrosine in a loop was incubated with a cysteine-containing payload and Catenase-K12 for 30 min at 4 °C with shaking at 750 rpm. For example, 5 μ M of a nanoluciferase construct was used with 10 μ M of Ricin-A-Chain-H6 (Txn) (SEQ ID NO: 59). For the reactions of this example, Txn (SEQ ID NO: 59) was used as a cysteine containing protein. The reaction was then quenched with a tris(2-carboxyethyl)phosphine (TCEP): Tropolone reducing inhibitor mix and analyzed for conjugation efficiency, quantified using SDS-PAGE densitometry. The efficiency of the conjugation reactions can also be assessed by mass spectrometry. To calculate the efficiency using SDS-PAGE densitometry, the reactions were run in triplicate and the Catenase-K12 (CatK12), Txn, nanoluciferase, and conjugation product samples were run on an SDS-PAGE gel. The signal intensity of the bands of the conjugated product and the unreacted, starting nanoluciferase protein were averaged from the SDS-PAGE gel (FIGs. 3C-3D, 4C-4D, 5C-5D, 6C-6D, 7C-7D, and 8C-8D). The difference in signal was calculated as percent (%) conjugation efficiency. The conjugation efficiencies are summarized below in Table 1 and FIG. 9. The fold increase in conjugation efficiency of the CatK12 on a non-terminal tag was at least 11-fold greater compared to the conjugation efficiency of AbTyr on a non-terminal tag.

Table 1. Conjugation Efficiency of nanoluciferase to Txn using Catenase-K12

% Conjugation efficiency					
Non-terminal tag	CatK12		AbTyr		Fold Increase
	Average	StDev	Average	StDev	
VEDYAIG (SEQ ID NO: 76)	53.8	4.2	2.6	0.2	20.7
ANYAA (SEQ ID NO: 50)	50.2	2.3	2.1	1.1	23.6
AEYAA (SEQ ID NO: 51)	46.8	4.2	2.7	0.4	17.2
AQYAA (SEQ ID NO: 52)	48.6	1.0	4.1	0.45	11.9
NYA	47.7	2.6	3.4	0.5	13.9
Y	16.5	3.1	0	0	
(C-terminal tag) SGGGGY (SEQ ID NO: 77)	87.5	1.2	40.5	0.9	2.2

[00400] As shown in the SDS-page gels, only one major product was formed from each conjugation reaction, showing the reaction is a chemoselective reaction.

[00401] The conjugation reactions were then run as described above but using *Agricus bisporus* tyrosinase (abTyr), a mushroom tyrosinase. The results of the conjugation

efficiencies are summarized in Table 1 and FIGS. 3C-3D, 4C-4D, 5C-5D, 6C-6D, 7C-7D, and 8C-8D. The conjugation reactions with CatK12 and a nanoluciferase with a non-terminal tag were between 11X – 24X more efficient than the conjugation reactions with abTyr.

[00402] A nanoluciferase with a C-terminal tag containing a tyrosine, SGGGGY (SEQ ID NO: 77), was also generated and tested. The conjugation efficiency of nanoluciferase with a terminal tag, SGGGGY (SEQ ID NO: 77), and Txn using either CatK12 or abTyr is shown in FIGS. 8E-8F and summarized in Table 1. The conjugation efficiency with the C-terminal tag was greater than the non-terminal tags, demonstrating that solvent accessibility contributes to the efficiency of the reaction.

[00403] The purity of the nanoluciferase constructs with non-terminal tags was measured by running the purified proteins on a non-reducing and a reducing gel and calculating the purity using SDS-PAGE densitometry. The results are summarized in Table 2 and FIG. 12A, and the relevant gels are shown in FIGS. 12B-12C. The nanoluciferase constructs with non-terminal tags were about 100% pure as shown in both the non-reducing and reducing gel calculation.

Table 2. Gel Purity of nanoluciferase constructs with non-terminals

Non-terminal tag	% Gel Purity	
	Non-reducing	Reducing
VEDYAIG (SEQ ID NO: 76)	100	100
ANYAA (SEQ ID NO: 50)	96	97
AEYAA (SEQ ID NO: 51)	100	99
AQYAA (SEQ ID NO: 52)	100	100
NYA	100	99
Y	100	100

EXAMPLE 3: Kinetic activity of Catenase K12

[00404] This Example describes the kinetic activity of Catenase K12 (SEQ ID NO: 53) compared to *Agricus bisporus* tyrosinase (abTYR) (SEQ ID NO:1).

[00405] To measure the kinetic activity of Catenase K12 (CatK12), peptides were used that were previously used to report the activity of abTYR (see Casey et al., J. Am. Chem. Soc. 2021 Sep 1;143(34):13538-13547).

[00406] Procedure 1 was used to measure the activity of CatK12 or abTYR with peptides.

Procedure 1:

- 1) Reconstitute peptides in 1:1 DMF/Buffer A (100 mM Sodium Phosphate, 2 mM EDTA, pH 6.5) to a concentration of 20 mM.
- 2) Make 50 mM 3-Methyl-2-benzothiazolinone hydrazone (MBTH) with 1:1 DMF/Buffer B (100 mM Sodium Phosphate, pH 6.5).
- 3) Dilute 1 mg/mL Catenase 1:10 with Buffer B.
- 4) Dilute 2 mg/mL abTYR 1:10 with Buffer B.
- 5) Make reaction mix and aliquot into well of a 96-well plate.
 - i) Final conc. for reaction: 250 μ M peptide, 5 mM MBTH, 10 μ g/mL Catenase or 20 μ g/mL abTYR.
- 6) Mix peptides or blank into wells. Run in triplicate.
- 7) Add Catenase to reaction at a 1:10 ratio for a final dilution ratio of 1:100 in reaction. Total volume of each reaction: 200 μ L.
- 8) Read plate for 15 min, with a measurement taken about every 10 seconds, at a wavelength of 505 nm.
- 9) Incubate plate at 4°C for 24 hours.
- 10) Read plate at 505 nm.

[00407] The sequences and corresponding ID number of the tested peptides are summarized in Table 3.

Table 3. Cyclic peptides used in to measure kinetic activity of tyrosinases

ID Number	Peptide Sequence	SEQ ID NO:
PEP001	EEYEEA	78
PEP002	EAEYAAE	79
PEP003	EEAYAAE	80
PEP004	EANYAAE	81
PEP005	EHAYAAE	82
PEP006	EEQYNST	45
PEP007	NSTYRVV	46
PEP008	DRVYIHP	47

PEP009	ADHYSWE	83
PEP010	VEDYAIG	76

[00408] The enzyme kinetic curves from the 15 min kinetic readings are shown for abTYR and CatK12 in FIGs. 13A and 13B, respectively. The kinetic curves were used to calculate the Vmax of abTYR and CatK12, which are summarized in FIGs. 13C and 13D, respectively, and in Table 4. The Vmax of CatK12 was greater than the Vmax of abTYR for all tested peptides. The increase in Vmax of the CatK12 activity was at least 3-fold greater than the Vmax compared to AbTyr activity. The increase in Vmax of the CatK12 activity was between 3-fold to 43.5-fold greater than the Vmax compared to AbTyr activity.

Table 4. Vmax of AbTyr and CatK12 on peptides

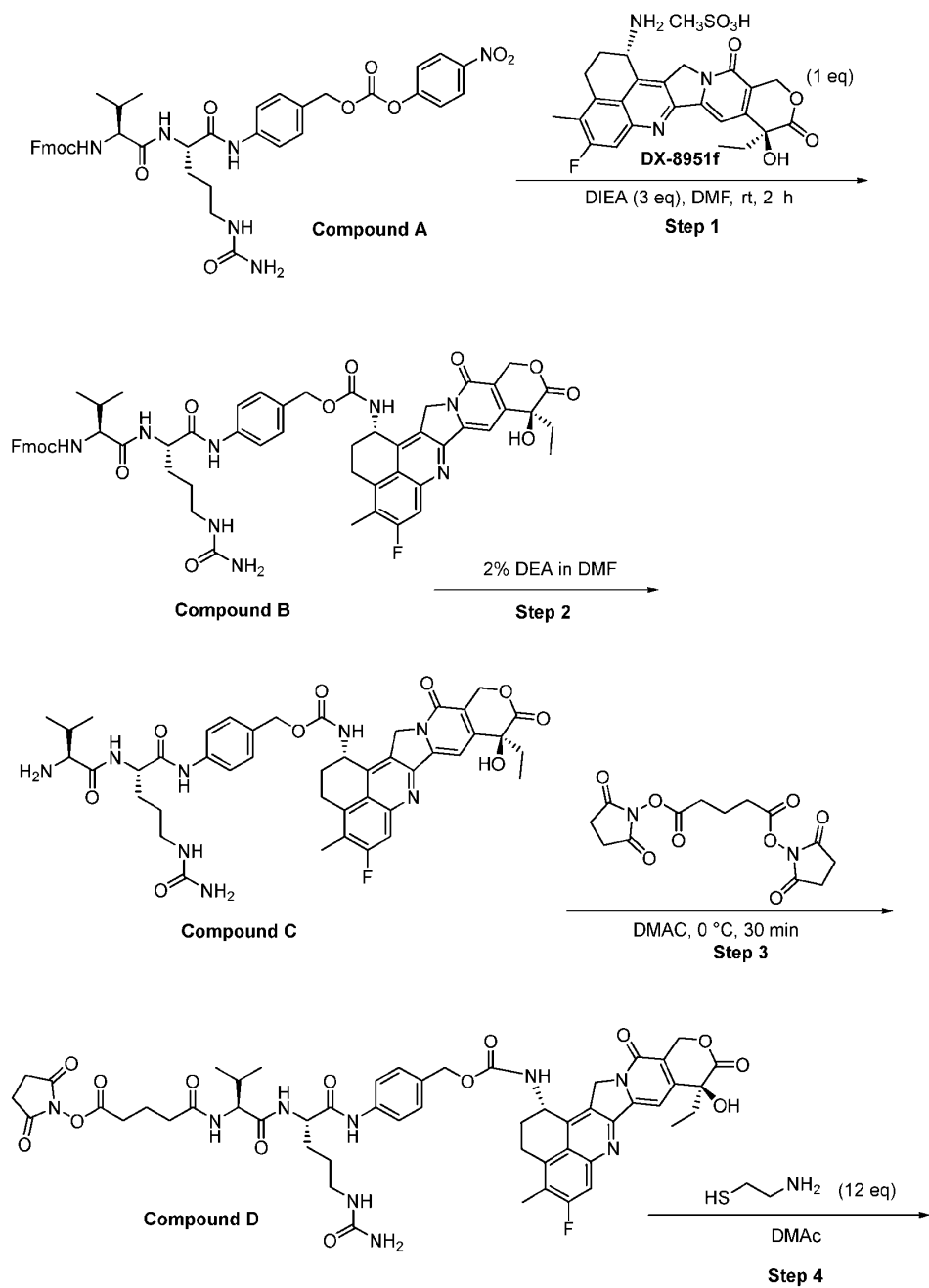
Sample ID	Vmax [Abs(505)/sec]				
	AbTyr		CatK12		Fold Increase
	Average	StDev	Average	StDev	
PEP-001	1.4	0.3	46.8	3.5	32.5
PEP-002	4.6	0.1	134.9	5.0	29.2
PEP-003	2.0	0.3	32.3	2.2	15.8
PEP-004	11.6	1.1	505.6	49.3	43.5
PEP-005	26.8	0.6	160.8	9.7	6.0
PEP-006	1.4	0.3	32.0	3.4	23.6
PEP-007	9.2	1.1	26.9	0.3	2.9
PEP-008	2.0	0.2	20.0	0.5	10.2
PEP-009	2.4	0.2	79.9	3.2	33.3
PEP-010	27.3	1.6	276.9	58.5	10.2
Blank	-1.0	0.2	1.5	0.2	

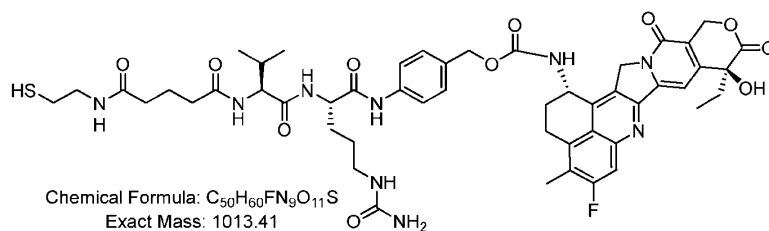
[00409] The absorbance after a 24 hour incubation was compared between CatK12 and abTRY (FIG. 13E). CatK12 showed greater activity than abTYR on most of the peptides after the 24 hour incubation.

EXAMPLE 4: Synthesis of Exatecan with a thiol-containing linker

[00410] This Example describes a synthesis of a payload, Exatecan, to a linker with a thiol. Exatecan is abbreviated as DXD in the scheme below.

Process 1:



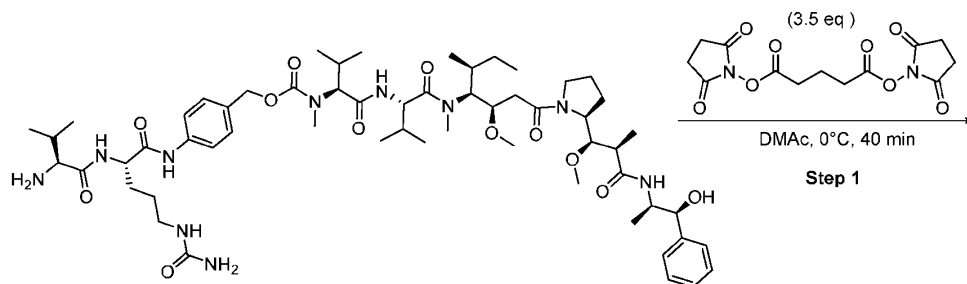
**SH-VC-PAB-DXD**

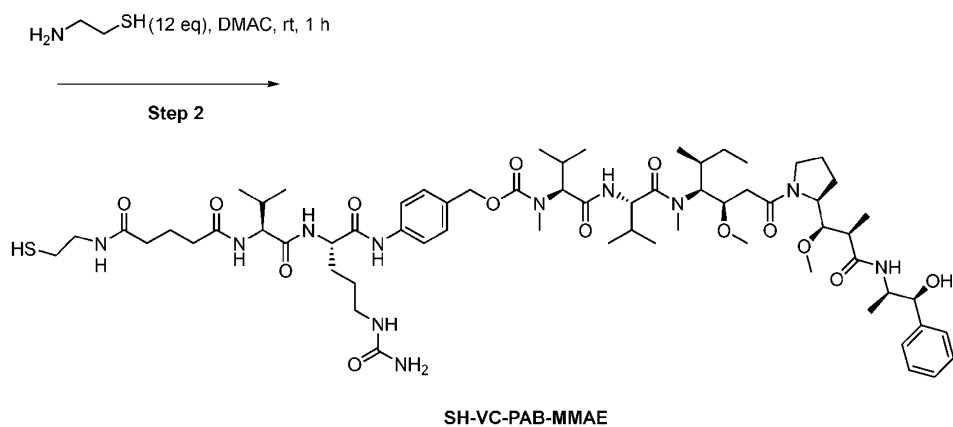
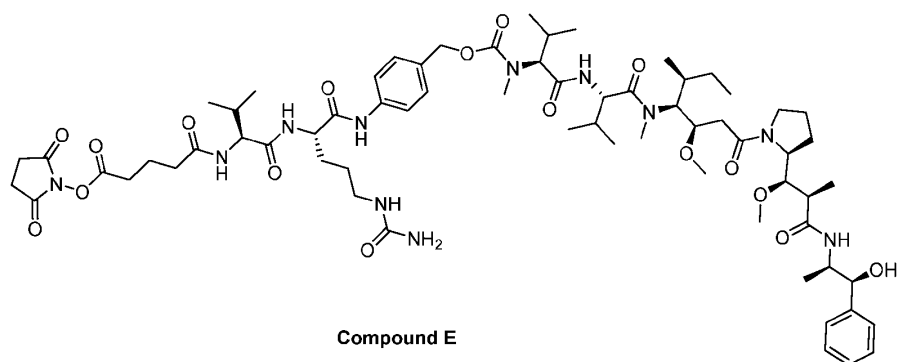
[00411] In Step 1, Compound A was reacted with DX-8951f in the presence of N,N-Diisopropylethylamine (DIEA) in dimethylformamide (DMF) at room temperature for two hours to provide Compound B. In the next step, Compound B was subjected to 2% Diethanolamine (DEA) in DMF to deprotect the Fmoc group and form Compound C. Then, Compound C was reacted with bis(2,5-dioxopyrrolidin-1-yl) glutarate in the presence of dimethylacetamide (DMAC) at 0°C, for 30 minutes to yield Compound D. Next, Compound D was reacted with 12 equivalents of 2-aminoethane-1-thiol in the presence of DMAC. The final SH-VC-PAB-DXD compound was purified, yielding 38 mg at 90 % purity. $[M+H]^+$ 1,014.35.

[00412] The NMR spectra of the final compound is provided in FIG. 14.

EXAMPLE 5: Synthesis of MMEA with a thiol-containing linker

[00413] This Example describes a synthesis of a payload, MMEA, to a linker with a thiol.

Process 2:



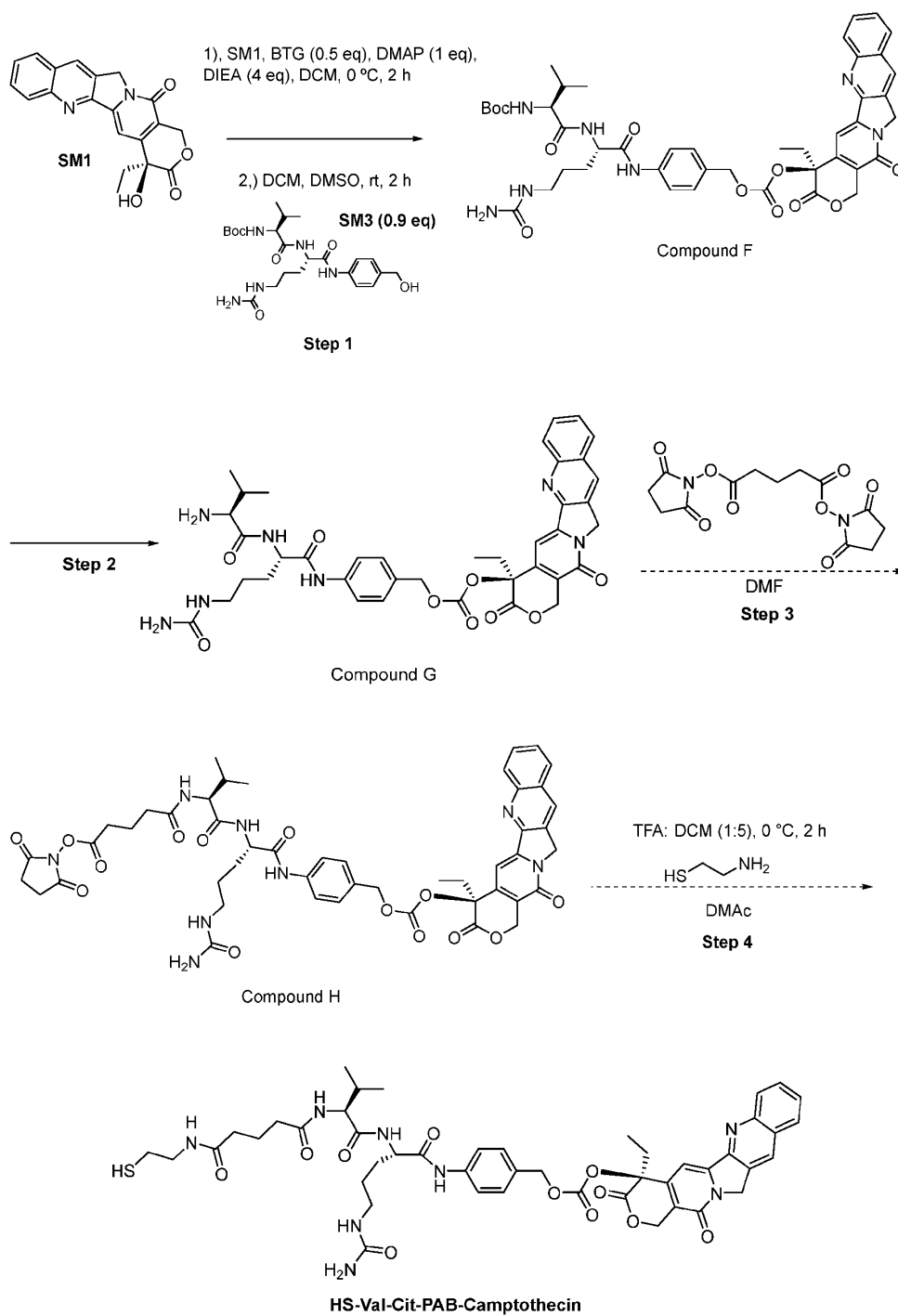
[00414] First, VC-PAB-MMAE was reacted with bis(2,5-dioxopyrrolidin-1-yl) glutarate in the presence of DMAC at 0 °C for 40 min. Next, Compound E was reacted with 12 equivalents of bis(2,5-dioxopyrrolidin-1-yl) glutarate in the presence of DMAC at room temperature for 1 hour. Finally, the SH-VC-PAB-MMAE compound was purified, yielding 59 mg. [M+H]⁺: 649.20.

[00415] The NMR spectra of the final compound is provided in FIG. 15.

EXAMPLE 6: Synthesis of Camptothecin with a thiol-containing linker

[00416] This Example describes a synthesis of a payload, Camptothecin, to a linker with a thiol.

Process 3:

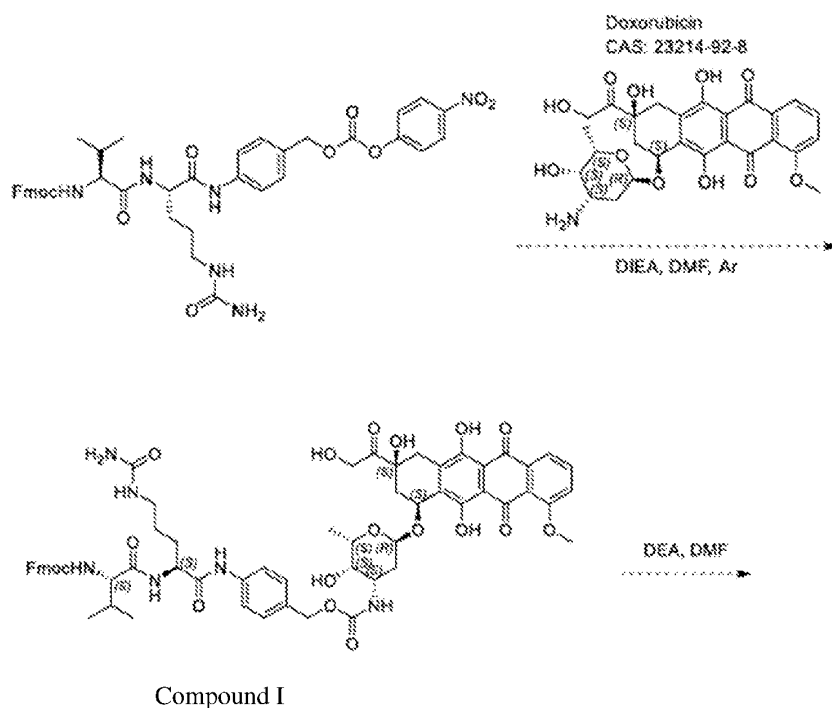


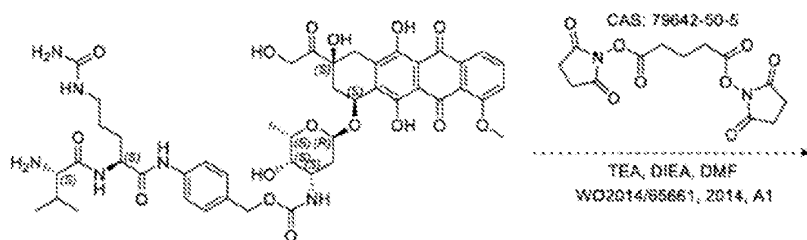
[00417] First, camptothecin was reacted with 0.5 equivalent of an agent that acts as a carbonyl source, 1 equivalent of 4-Dimethylaminopyridine (DMAP), and 4 equivalents of N,N-Diisopropylethylamine (DIPEA) in DCM at 0 °C for 2 hours. Next, the resulting compound was reacted with 0.9 equivalent of Boc-Val-Cit dipeptide in DCM and DMSO at room temperature for 2 hours. In step 2, Compound F was deprotected to form Compound G. In Step 3, Compound G was reacted with bis(2,5-dioxopyrrolidin-1-yl) glutarate in DMF to provide Compound H. In Step 4, Compound H was reacted with 2-aminoethane-1-thiol in trifluoroacetic acid (TFA) and DCM, in a 1:5 ratio, at 0 °C, for 2 hours. Finally, the compound was purified to give the SH-VC-PAB-Camptothecin compound.

EXAMPLE 7: Synthesis of Doxorubicin with a thiol-containing linker

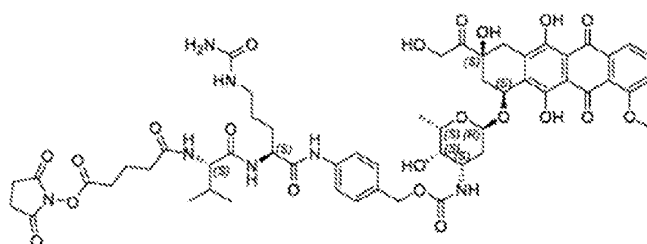
[00418] This Example describes a synthesis of a payload, Doxorubicin, to a linker with a thiol.

Process 4

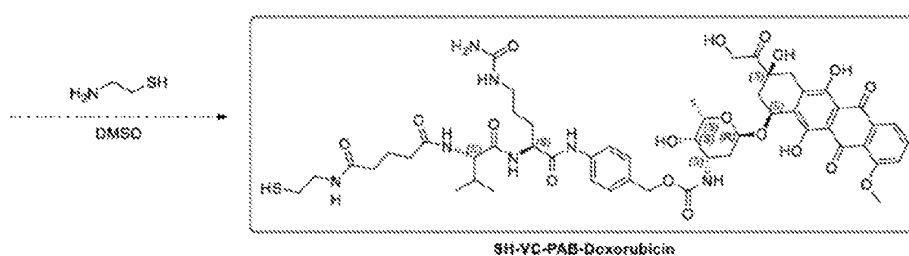




Compound J



Compound K



[00419] In Step 1, an Fmoc protected VC-PAB compound was reacted with doxorubicin in DIA and DMF to form Compound I. Next, Compound I was deprotected using DEA and DMF to form Compound J. Next, Compound J was reacted with bis(2,5-dioxopyrrolidin-1-yl) glutarate, TEA, DIEA, and DMA to provide Compound K. In the last step, Compound K was reacted with 2-aminoethane-1-thiol to provide the final compound, SH-VC-PAB-Doxorubicin.

[00420] The procedures disclosed herein can be conducted in a number of ways based on the teachings contained herein and synthetic procedures known in the art. In the description of the synthetic methods described above, it is to be understood that all proposed reaction conditions, including choice of solvent, reaction atmosphere, reaction temperature, duration of the experiment and workup procedures, can be chosen to be the conditions standard for that

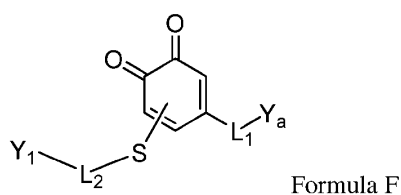
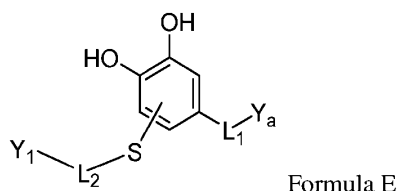
reaction, unless otherwise indicated. It is understood by one skilled in the art of organic synthesis that the functionality present on various portions of the molecule should be compatible with the reagents and reactions proposed. Substituents not compatible with the reaction conditions will be apparent to one skilled in the art, and alternate methods are therefore indicated. The starting materials for the examples are either commercially available or are readily prepared by standard methods from known materials.

[00421] It should be recognized that illustrated embodiments are only examples of the disclosed product and methods and should not be considered a limitation on the scope of the embodiments. Rather, the scope of the embodiments is defined by the claims. We therefore claim all that comes within the scope and spirit of these claims

CLAIMS

What is claimed is:

1. A conjugate of Formula E or Formula F:



wherein:

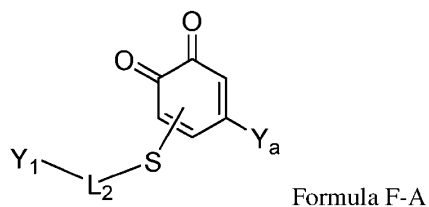
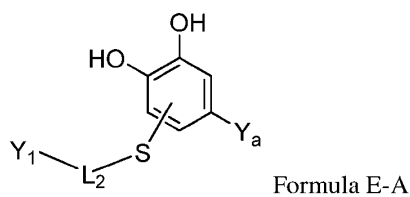
Ya is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Ya is linked to "S" via the tyrosine or a portion thereof;

L₁ is an optional first linker;

L₂ is an optional second linker; and

Y₁ is a payload.

2. The conjugate of claim 1, wherein the conjugate comprises Formula E-A or Formula F-A:



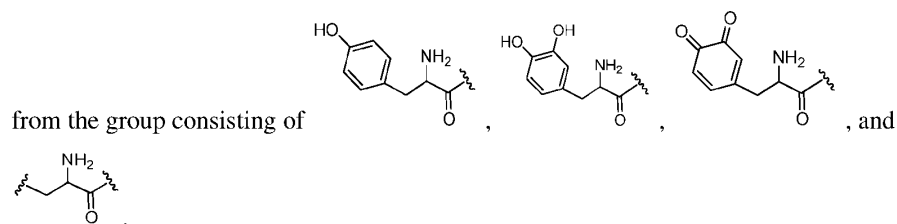
wherein:

Y_a is a polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

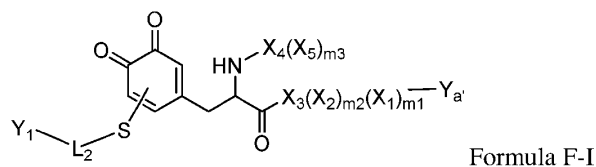
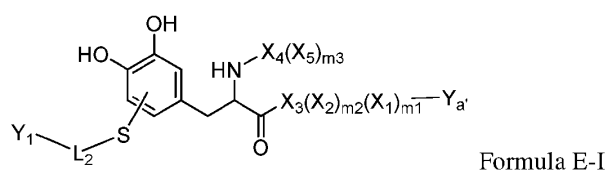
L_2 is an optional second linker; and

Y_1 is a payload.

3. The conjugate of claim 1 or 2, wherein the tyrosine or a portion thereof is selected



4. The conjugate of any one of claims 1 -3, wherein the non-terminal tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K).
5. The conjugate of any one of claims 1 -3, wherein the non-terminal tag comprises $X_1X_2YX_4X_5$.
6. The conjugate of claim 5, wherein X_1 and X_2 are proton-donating amino acids.
7. The conjugate of claim 5 or 6, wherein X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.
8. The conjugate of any one of claims 5-7, wherein X_4 and X_5 are small, neutral amino acids.
9. The conjugate of any one of claims 5-8, wherein X_4 is A or G.
10. The conjugate of any one of claims 5-9, wherein X_5 is E, V, D, A, I, G, S, T, or L.
11. A conjugate of Formula E-I or Formula F-I:



wherein:

Y_a is a polypeptide,

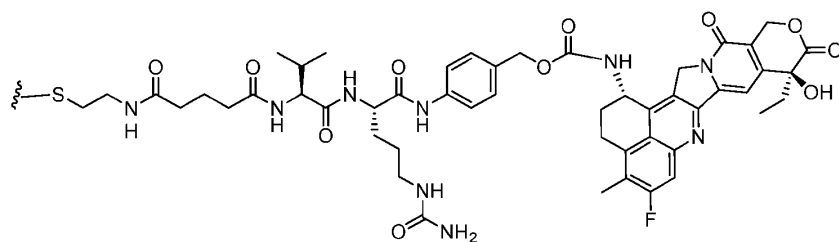
each of X_1 , X_2 , X_3 , X_4 , and X_5 is independently any amino acid in a non-terminal tag in the polypeptide, wherein m_1 , m_2 , and m_3 are each an integer greater than or equal to 0;

L_2 is an optional linker; and

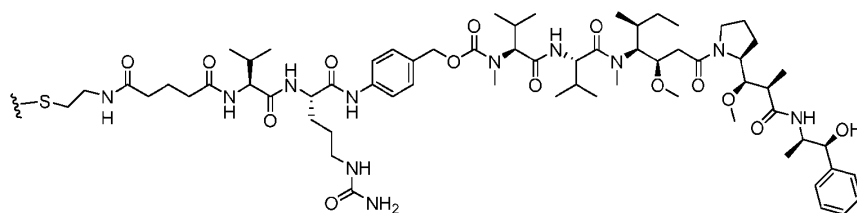
Y_1 is a payload

12. The conjugate of claim 11, wherein X_2 and X_3 are proton-donating amino acids.
13. The conjugate of claim 11 or 12, wherein X_2 and X_3 are each selected from the group consisting E, H, Q, D, and N.
14. The conjugate of any one of claims 11-13, wherein X_4 and X_5 are small, neutral amino acids.
15. The conjugate of any one of claims 11-14, wherein X_4 is A or G and X_5 is E, V, D, A, I, G, S, T, or L.
16. The conjugate of any one of claims 5-15, wherein the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.
17. The conjugate of any one of claims 5-16, wherein the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.
18. The conjugate of any one of claims 1-4, wherein the non-terminal tag comprises at least one of GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGEY (SEQ ID NO: 13), EEGY (SEQ ID NO: 14), EEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEWY (SEQ ID NO: 20), EEWEY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 25), KKGGY (SEQ ID NO: 30), and DDGGY (SEQ ID NO: 31).
19. The conjugate of any one of claims 1-4 or 18, wherein the non-terminal tag comprises SGGY (SEQ ID NO: 23) or SGY.
20. The conjugate of any one of claims 1-4 or 18, wherein the non-terminal tag comprises EEEY.
21. The conjugate of any one of claims 1-20, wherein Y_a is glycosylated.
22. The conjugate of any one of claims 1-21, wherein Y_a comprises a non-terminal tyrosine.

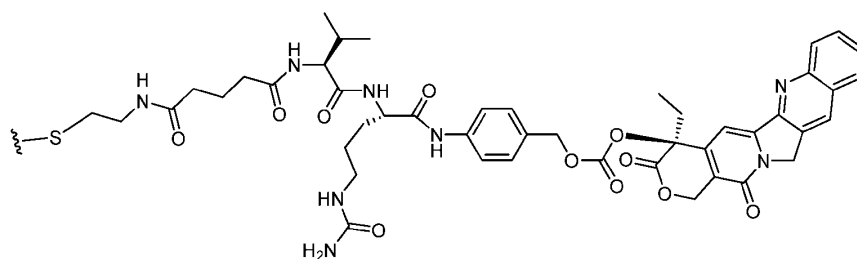
23. The conjugate of any one of claims 1-22, wherein at least one of L₁ or L₂ is a cleavable linker.
24. The conjugate of claim 23, wherein the cleavable linker is an electrophilically cleavable linker, a nucleophilically cleavable linker, a photocleavable linker, a metal cleavable linker, an electrolytically-cleavable linker, an acid cleavable linker, or a proteolytically cleavable linker.
25. The conjugate of claim 23, wherein the cleavable linker further includes a pegylated group, a sugar group, or a modification that increases hydrophilicity.
26. The conjugate of claim 23, wherein the cleavable linker is cleavable under reductive and/or oxidative conditions.
27. The conjugate of claim 23, wherein the cleavable linker is cleavable under acidic conditions.
28. The conjugate of claim 23, wherein the cleavable linker comprises a disulfide bond.
29. The conjugate of claim 23, wherein the cleavable linker is a proteolytically cleavable linker and comprises a protease recognition sequence.
30. The conjugate of claim 29, wherein the protease recognition sequence is recognized by a protease selected from the group comprising a metalloprotease, cathepsin B, and tobacco etch virus (TEV).
31. The conjugate of claim 23, wherein the cleavable linker comprises a dipeptide or tetrapeptide.
32. The conjugate of claim 31, wherein the dipeptide is a valine-citrulline (Val-Cit) dipeptide, a valine-lysine dipeptide, and a valine-alanine dipeptide.
33. The conjugate of claim 31, wherein the tetrapeptide is a glycine-glycine-phenylalanine-glycine (GGFG) tetrapeptide (SEQ ID NO: 75).
34. The conjugate of claim 23, wherein the cleavable linker is selected from the group comprising PABC (p-aminobenzyl alcohol), glucuronide, and MABC (m-aminobenzyl alcohol).
35. The conjugate of claim 23, wherein the cleavable linker is Val-Cit-PABC.
36. The conjugate of claim 1, wherein "Y₁-L₂-S-" is selected from the group consisting of:



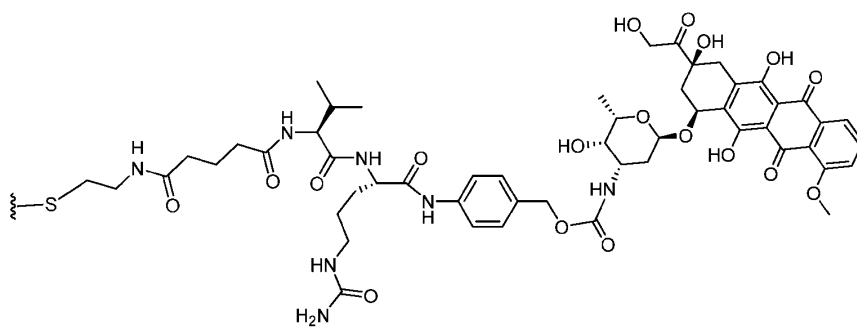
S-VC-PAB-DXD,



S-VC-PAB-MMAE,



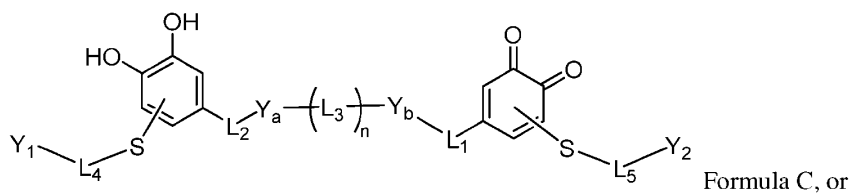
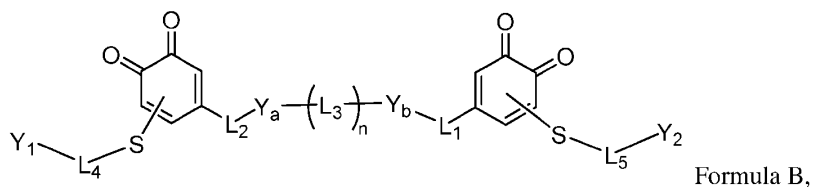
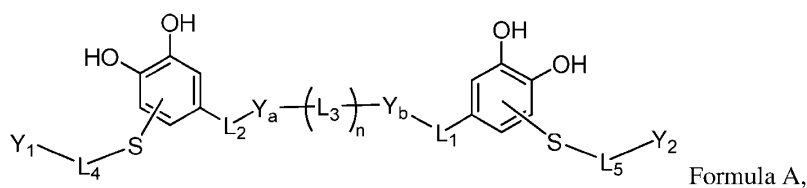
S-VC-PAB-Camptothecin, and

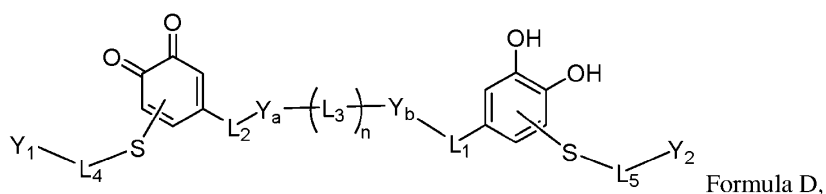


S-VC-PAB-Doxorubicin.

37. The conjugate of any one of claims 1-36, wherein Y_a is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, a constant chain of an antibody or antibody fragment, a peptide, a fluorescent protein, or a cyclic peptide.
38. The conjugate of any one of claims 1-36, wherein Y_a is a viral protein.
39. The conjugate of any one of claims 1-36, wherein Y_a is a cell surface receptor.

40. The conjugate of any one of claims 1-36, wherein Y_a is a binding peptide.
41. The conjugate of any one of claims 1-39, wherein Y_1 is a small molecule.
42. The conjugate of claim 41, wherein the small molecule is selected from the group comprising deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan, lurototecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimmitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4), pyrolodiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT.
43. The conjugate of any one of claims 1-39, wherein Y_1 comprises a nucleic acid, an immune agonist, a peptide, a cytokine, a protein, or a binding domain.
44. The conjugate of any one of claims 1-39, wherein Y_1 comprises a protein.
45. The conjugate of any one of claims 1-39, wherein Y_1 comprises a peptide.
46. The conjugate of any one of claims 1-39, wherein Y_1 comprises a binding peptide.
47. A conjugate of Formula A, Formula B, Formula C, or Formula D:





wherein:

Y_a is a first polypeptide comprising a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 comprises a tyrosine or a portion thereof, and wherein $m1$ is an integer greater or equal to 0, wherein Y_a is linked to "S" via the tyrosine or a portion thereof;

Y_b is a second polypeptide comprising a second tag different from the first tag, the second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} comprises a tyrosine or a portion thereof, wherein $m2$ is an integer greater or equal to 0, wherein Y_b is linked to "S" via the tyrosine or a portion thereof;

n is an integer greater than 0;

L_1 is an optional first linker;

L_2 is an optional second linker;

L_3 is an optional third linker;

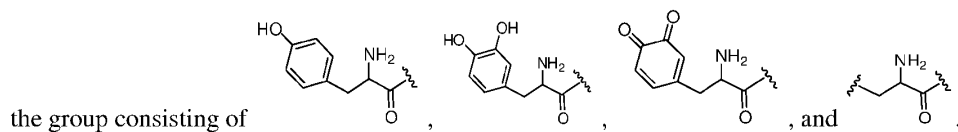
L_4 is an optional fourth linker;

L_5 is an optional fifth linker;

Y_1 is a first payload; and

Y_2 is a second payload; and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag.

48. The conjugate of claim 47, wherein the tyrosine or a portion thereof is selected from



49. The conjugate of claim 47, wherein the non-terminal tag comprises $X_1X_2YX_4X_5$.

50. The conjugate of claim 47, wherein X_1 and X_2 are proton-donating amino acids.

51. The conjugate of claim 47, wherein X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

52. The conjugate of any one of claims 47-51, wherein X_4 and X_5 are small, neutral amino acids.

53. The conjugate of any one of claims 47-52, wherein X_4 is A or G.
54. The conjugate of any one of claims 47-53, wherein X_5 is E, V, D, A, I, G, S, T, or L.
55. The conjugate of any one of claims 47-54, wherein the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.
56. The conjugate of any one of claims 47-55, wherein the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.
57. The conjugate of claim 47, wherein the first tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K).
58. The conjugate of claim 47 or 57, wherein the second tag comprises at least two amino acids of aspartate (D), glutamate (E), arginine (R), and lysine (K).
59. The conjugate of any one of claims 47 or 57-58, wherein the first tag comprises at least one of GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGEY (SEQ ID NO: 13), EEEY (SEQ ID NO: 14), EEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWEY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 25), KKGYY (SEQ ID NO: 31), and DDGGY (SEQ ID NO: 32).
60. The conjugate of any one of claims 47 or 57-59, wherein the second tag comprises at least one of GGGGY (SEQ ID NO: 7), RGGGY (SEQ ID NO: 8), RGRGY (SEQ ID NO: 9), RRRGY (SEQ ID NO: 10), RRRRY (SEQ ID NO: 11), EGGGY (SEQ ID NO: 12), EGEY (SEQ ID NO: 13), EEEY (SEQ ID NO: 14), EEEY (SEQ ID NO: 15), GGGWY (SEQ ID NO: 16), GGWGY (SEQ ID NO: 17), RRRWY (SEQ ID NO: 18), RRWRY (SEQ ID NO: 19), EEEWY (SEQ ID NO: 20), EEWEY (SEQ ID NO: 21), DDDDY (SEQ ID NO: 22), SGGY (SEQ ID NO: 23), SGY, KKKKY (SEQ ID NO: 24), RRKKY (SEQ ID NO: 25), RRRKY (SEQ ID NO: 26), EDEDY (SEQ ID NO: 27), EDDDY (SEQ ID NO: 28), EEDDY (SEQ ID NO: 29), RRKKY (SEQ ID NO: 25), KKGYY (SEQ ID NO: 31), and DDGGY (SEQ ID NO: 32).
61. The conjugate of any one of claims 47 or 57-59, wherein the first tag comprises SGGY (SEQ ID NO: 23) or SGY and the second tag comprises EEEY (SEQ ID NO: 33).
62. The conjugate of any one of claims 47-61, wherein the first tag is a non-terminal tag.

63. The conjugate of any one of claims 47-61, wherein the first tag is a terminal tag.
64. The conjugate of any one of claims 47-63, wherein the second tag is a terminal tag.
65. The conjugate of any one of claims 47-63, wherein the second tag is a non-terminal tag.
66. The conjugate of any one of claims 47-61, wherein the first tag and the second tag are non-terminal tags.
67. The conjugate of any one of claims 47-66, wherein Y_a , Y_b , or both Y_a and Y_b are glycosylated.
68. The conjugate of any one of claims 47-61, wherein Y_a , Y_b , or both Y_a and Y_b comprise a non-terminal tyrosine.
69. The conjugate of any one of claims 47-68, wherein at least one of L_1 , L_2 , L_3 , L_4 , or L_5 is a cleavable linker.
70. The conjugate of claim 69, wherein the cleavable linker is an electrophilically cleavable linker, a nucleophilically cleavable linker, a photocleavable linker, a metal cleavable linker, an electrolytically-cleavable linker, an acid cleavable linker, or a proteolytically cleavable linker.
71. The conjugate of claim 69, wherein the cleavable linker further includes a pegylated group, a sugar group, or a modification that increases hydrophilicity.
72. The conjugate of claim 69, wherein the cleavable linker is cleavable under reductive and/or oxidative conditions.
73. The conjugate of claim 69, wherein the cleavable linker is cleavable under acidic conditions.
74. The conjugate of claim 69, wherein the cleavable linker comprises a disulfide bond.
75. The conjugate of claim 69, wherein the cleavable linker is a proteolytically cleavable linker and comprises a protease recognition sequence.
76. The conjugate of claim 75, wherein the protease recognition sequence is recognized by a protease selected from the group comprising a metalloprotease, cathepsin B, and tobacco etch virus (TEV).
77. The conjugate of claim 69, wherein the cleavable linker comprises a dipeptide or tetrapeptide
78. The conjugate of claim 78, wherein the dipeptide is a valine-citrulline (Val-Cit) dipeptide, a valine-lysine dipeptide, or a valine-alanine dipeptide.
79. The conjugate of claim 78, wherein the tetrapeptide is a glycine-glycine-phenylalanine-glycine (GGFG) tetrapeptide (SEQ ID NO: 75).

80. The conjugate of claim 69, wherein the cleavable linker is selected from the group comprising PABC (p-aminobenzyl alcohol), glucuronide, and MABC (m-aminobenzyl alcohol).
81. The conjugate of any one of claims 47-80, wherein Y_a is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, a constant chain of an antibody or antibody fragment, a peptide, or a cyclic peptide.
82. The conjugate of any one of claims 47-81, wherein Y_b is a heavy chain variable region (VH), a light chain variable region (VL), a heavy chain, a light chain, or a constant chain of an antibody, antibody fragment, a peptide, or a cyclic peptide.
83. The conjugate of any one of claims 47-80, wherein Y_a or Y_b is a viral protein.
84. The conjugate of any one of claims 47-80, wherein Y_a and Y_b are viral proteins.
85. The conjugate of any one of claims 47-80, wherein Y_a or Y_b is a cell surface receptor.
86. The conjugate of any one of claims 47-81, wherein Y_a and Y_b are attached via L_3 .
87. The conjugate of any one of claims 47-86, wherein L_3 comprises a peptide sequence, a dimerization and docking domain, a leucine zipper, or knobs-into-holes.
88. The conjugate of any one of claims 47-86, wherein L_3 comprises a peptide bond, a disulfide bond, a maleimide bond, thioether bond, an azide-alkyne cycloaddition, a cysteinyl-dopa, or a hydrogen bond.
89. The conjugate of any one of claims 47-86, wherein L_3 is a linker.
90. The conjugate of claim 89, wherein the linker comprises a sequence selected from the group consisting of (GS) n_3 (SEQ ID NO: 66), (G2S) n_3 (SEQ ID NO: 67), (G3S) n_3 (SEQ ID NO: 68), (G4S) n_3 (SEQ ID NO: 69), (G) n_3 (SEQ ID NO: 70), (GGSGGD) n_3 (SEQ ID NO: 71), (GGSGGE) n_3 (SEQ ID NO: 72), (GGGSGSGGGG) n_3 (SEQ ID NO: 73), and (GGGGPGGGGP) n_3 (SEQ ID NO: 74), and wherein n_3 is an integer from 2 to 20.
91. The conjugate of any one of claims 47-86, wherein L_3 comprises a terminal or non-terminal tyrosine.
92. The conjugate of any one of claims 47-91, wherein n is 1, 2, 3, 4, or 5.
93. The conjugate of any one of claims 47-91, wherein n is 1.
94. The conjugate of any one of claims 47-91, wherein Y_1 and Y_2 are the same.
95. The conjugate of any one of claims 47-91, wherein Y_1 and Y_2 are different.
96. The conjugate of any one of claims 47-95, wherein Y_1 , Y_2 , or both Y_1 and Y_2 is a small molecule.
97. The conjugate of claim 96, wherein the small molecule is selected from the group comprising deruxtecan, exatecan, FL118, irinotecan, topotecan, SN-38, rubitecan, belotecan,

lurototecan, gimatecan, diflomotecan, karenitecan, silatecan, namitecan, elomotecan, DRF-1042, delimotecan, NSC606985, chimmitecan, ZBH-1205, auristatin (MMAE, MMAF, MMAG, MMAH) calicheamicin, doxorubicin, taxol and taxol derivatives, maytansinoids (DM1-4), pyrolodiazepines (PBDs), tubulysins, eribulin, anthramycin, duocarmycin, anthracycline, and camptothecin (CPT), including the lactone and carboxylate forms of CPT.

98. The conjugate of any one of claims 47-95, wherein Y_1 and Y_2 , or both Y_1 and Y_2 comprises a nucleic acid, an immune agonist, a peptide, a cytokine, or a binding domain.

99. The conjugate of any one of claims 47-95, wherein Y_1 , Y_2 , or both Y_1 and Y_2 comprises a nucleic acid.

100. The conjugate of any one of claims 47-95, wherein Y_1 and Y_2 , or both Y_1 and Y_2 comprises a peptide.

101. A polypeptide comprising a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof, and wherein $m1$ is an integer greater than or equal to 0.

102. The polypeptide of claim 101, wherein the non-terminal tag comprises $X_1X_2YX_4X_5$.

103. The polypeptide of claim 101 or 102, wherein X_1 and X_2 are proton-donating amino acids.

104. The polypeptide any one of claims 101-103 wherein X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

105. The polypeptide of any one of claims 101-104, wherein X_4 and X_5 are small, neutral amino acids.

106. The polypeptide of any one of claims 101-105, wherein X_4 is A or G.

107. The polypeptide of any one of claims 101-106, wherein X_5 is E, V, D, A, I, G, S, T, or L.

108. The polypeptide of any one of claims 101-107, wherein the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.

109. The polypeptide of any one of claims 101-108, wherein the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYAA (SEQ ID NO: 84).

110. A composition of a first polypeptide and a second polypeptide, wherein the first polypeptide comprises a first non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine or a portion thereof; and wherein $m1$ is an integer greater than or equal to 0, and the second polypeptide comprises a second non-terminal tag different from the first terminal tag, the second terminal

tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine or a portion thereof, and wherein $m2$ is an integer greater than or equal to 0.

111. The composition of claim 110, wherein the first polypeptide and/or the second polypeptide comprise a non-terminal tyrosine.

112. The composition of claim 110 or 111, wherein the non-terminal tag comprises $X_1X_2YX_4X_5$.

113. The composition of any one of claims 110-112, wherein X_1 and X_2 are proton-donating amino acids.

114. The composition of claims 110-113, wherein X_1 and X_2 are each selected from the group consisting E, H, Q, D, and N.

115. The composition of any one of claims 110-114, wherein X_4 and X_5 are small, neutral amino acids.

116. The composition of any one of claims 110-115, wherein X_4 is A or G.

117. The composition of any one of claims 110-116, wherein X_5 is E, V, D, A, I, G, S, T, or L.

118. The composition of any one of claims 110-117, wherein the non-terminal tag has a charge of +1, 0, -1, -2, or -3 at pH 7.

119. The composition of any one of claims 110-118, wherein the non-terminal tag comprises the sequence VEDYAIG (SEQ ID NO: 76), ANYAA (SEQ ID NO: 50), AEYAA (SEQ ID NO: 51), AQYAA (SEQ ID NO: 52) or NYA.

120. A method of covalently linking at least two polypeptides to at least two payloads, the method comprising:

a) contacting a first polypeptide of the at least two polypeptides and a first payload of the at least two payloads using a first tyrosinase, wherein the first polypeptide comprises a first tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1 - X_5 is any amino acid provided that at least one amino acid of X_1 - X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0; and

b) contacting a second polypeptide of the at least two polypeptides and a second payload of the at least two payloads using a second tyrosinase, wherein the second polypeptide comprises a second tag comprising $(X_6)_{m2}X_7X_8X_9X_{10}$, wherein X_6 - X_{10} is any amino acid provided that at least one amino acid of X_6 - X_{10} is a tyrosine, wherein $m2$ is an integer greater or equal to 0, and wherein the first tag, the second tag, or both the first tag and the second tag is a non-terminal tag.

121. The method of claim 120, wherein the first tag and the second tag are non-terminal tags.
122. The method of any one of claims 120-121, wherein the first tyrosinase, the second tyrosinase, or both the first tyrosinase and the second tyrosinase is *Agricus bisporus* tyrosinase (abTYR).
123. The method of any one of claims 120-121, wherein the first tyrosinase, the second tyrosinase, or both comprises a sequence at least 80% identity to any one of SEQ ID NOs: 1 or 43
124. The method of any one of claims 120-121, wherein the first tyrosinase, the second tyrosinase, or both the first tyrosinase and the second tyrosinase is Catenase.
125. The method of any one of claims 120-121 or 124, wherein the first tyrosinase, the second tyrosinase, or both comprises a sequence at least 90% identity to any one of SEQ ID NOs: 2-6 or 53.
126. The method of any one of claims 120-125, wherein the first tyrosinase is *Agricus bisporus* tyrosinase (abTYR) and the second tyrosinase is Catenase.
127. The method of any one of claims 120-126, wherein the first tyrosinase and the second tyrosinase are provided at a same time.
128. The method of any one of claims 120-126, wherein the first tyrosinase is provided first followed by the second tyrosinase.
129. A method of covalently linking a polypeptide to a payload, the method comprising:
contacting a polypeptide and a payload using a tyrosinase, wherein the polypeptide comprises a non-terminal tag comprising $(X_1)_{m1}X_2X_3X_4X_5$, wherein X_1-X_5 is any amino acid provided that at least one amino acid of X_1-X_5 is a tyrosine, and wherein $m1$ is an integer greater or equal to 0.
130. The method of claim 129, wherein the tyrosinase is *Agricus bisporus* tyrosinase (abTYR).
131. The method of claim 129, wherein the tyrosinase comprises a sequence at least 80% identity to any one of SEQ ID NOs: 2-6 or 53.
132. The method of claim 129 or 131, wherein the tyrosinase is a Catenase.
133. The method of any one of claims 129 or 131-132, wherein the tyrosinase comprises a sequence at least 90% identity to any one of SEQ ID NOs: 2-6 or 53.
134. The method of any one of claims 129 or 131-133, wherein the tyrosinase comprises a sequence at least 90% identity to SEQ ID NO: 53.
135. An engineered tyrosinase comprising at least 80% identity to SEQ ID NO: 53.

136. The engineered tyrosinase of claim 135, wherein the engineered tyrosinase comprises at least 90% identity to SEQ ID NO: 53.

137. The engineered tyrosinase of claim 135 or 136, wherein the engineered tyrosinase comprises 100% identity to SEQ ID NO: 53.

138. The engineered tyrosinase of any one of claims 135-137, wherein the engineered tyrosinase has at least an 11-fold increase in conjugation efficiency as compared to AbTyr (SEQ ID NO:1).

139. The engineered tyrosinase of any one of claims 135-137, wherein the engineered tyrosinase has at least a 3-fold greater Vmax activity as compared to AbTyr (SEQ ID NO:1).