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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
08 February 2024 (08.02.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/030858 A1

(51) International Patent Classification:
A61P 35/00 (2006.01) *C12N 15/62* (2006.01)
C12N 9/64 (2006.01)

(21) International Application Number:
PCT/US2023/071326

(22) International Filing Date:
31 July 2023 (31.07.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/370,023 01 August 2022 (01.08.2022) US

(71) Applicant: **CYTOMX THERAPEUTICS, INC.** [US/US]; 151 Oyster Point Boulevard, Suite 400, South San Francisco, California 94080 (US).

(72) Inventors: **VASILJEVA, Olga**; c/o CytomX Therapeutics, Inc., 151 Oyster Point Boulevard, Suite 400, South San Francisco, California 94080 (US). **WINTER, Michael B.**; C/O CytomX Therapeutics, Inc., 151 Oyster Point Blvd., Suite 400, South San Francisco, California 94080 (US).

(74) Agent: **HARSTON, Aydin H.** et al.; Rothwell, Figg, Ernst & Manbeck, P.C., 901 New York Avenue, NW, Suite 900 East, Washington, District of Columbia 20001 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,

(54) Title: PROTEASE-CLEAVABLE SUBSTRATES AND METHODS OF USE THEREOF

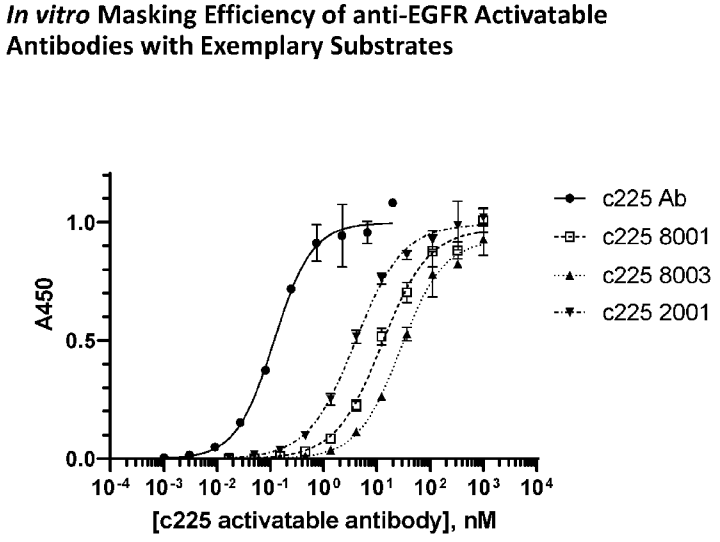

Fig. 1A

(57) Abstract: Isolated polypeptides that include a cleavable moiety that is a substrate for at least two proteases (e.g., MT-SP1 and an MMP) are disclosed. Activatable molecules including the isolated polypeptides are disclosed. Methods of making and using the isolated polypeptides and activatable molecules including the isolated polypeptides in a variety of therapeutic, diagnostic, and prophylactic applications are disclosed.

[Continued on next page]

WO 2024/030858 A1

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

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

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

PROTEASE-CLEAVABLE SUBSTRATES AND METHODS OF USE THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the priority benefit of U.S. provisional application no. 63/370,023 filed August 1, 2022, the contents of which are incorporated herein in their entireties by reference thereto.

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 July 27, 2023 is named CYTX-094-PCT_SL.xml and is 512,000 bytes in size.

TECHNICAL FIELD

[0003] The present disclosure generally relates to polypeptides that include a substrate comprising one or more cleavable moieties, each of which is cleavable by a protease, and to methods of making and using the polypeptides and activatable molecules in a variety of therapeutic, diagnostic, and prophylactic applications.

BACKGROUND

[0004] Proteases are enzymes that catalyze the hydrolysis of peptide bonds between amino acid residues. Some proteases are known to break specific peptide bonds based on the presence of a particular amino acid sequence within a protein. Proteases occur naturally in all organisms and are involved in a variety of physiological reactions from simple degradation to highly regulated pathways. Some proteases break specific peptide bonds based on the presence of a particular amino acid sequence within a protein while some amino acid sequences are resistant to cleavage by particular proteases.

[0005] Accordingly, there exists a need to identify new substrates for proteases and to use these substrates in a variety of therapeutic, diagnostic and prophylactic applications.

SUMMARY OF THE INVENTION

[0006] In one aspect, the present disclosure provides an isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the CM1 comprises the amino acid sequence of HQSRS (SEQ ID NO: 2).In another aspect, the present disclosure provides an isolated polypeptide comprising a substrate, wherein the

substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the CM1 comprises the amino acid sequence of HQSR (SEQ ID NO: 1), HQSK (SEQ ID NO: 3), or HQSRSA (SEQ ID NO: 4). In some embodiments, the CM1 comprises the amino acid sequence of DHQSRS (SEQ ID NO: 5). In some embodiments, the CM1 comprises the amino acid sequence of SDHQSRS (SEQ ID NO: 6). In some embodiments, the CM2 comprises the amino sequence of AIALYAD (SEQ ID NO: 7), AIALYA (SEQ ID NO: 8), or AIAL (SEQ ID NO: 9). In some embodiments, the CM2 comprises the amino sequence of GPWGLL (SEQ ID NO: 237) or SGLLSS (SEQ ID NO: 238). In some embodiments, the CM2 comprises the amino sequence of ALAHGLF (SEQ ID NO: 74), or ISSGLLSS (SEQ ID NO: 75).

[0007] According to some embodiments of the present disclosures, the isolated polypeptide is a molecule in which cleavage of the CM by a protease results in a part or component of the molecule being separated from the remainder of the molecule. In some aspects of the present disclosure, cleavage of the CM by a protease activates the molecule. In some aspects, the isolated polypeptide is a molecule in which multiple proteases cleave the CM. In some aspects, the isolated polypeptide is a molecule in which MMP2 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which MMP9 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which MMP14 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which MT-SP1 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which two or all of MMP2, MMP9, MMP14, MT-SP1 cleave the CM. In some aspects, the isolated polypeptide is a molecule in which the % cleavability of the CM is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, e.g., at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100% cleavable by any one of MMP2, MMP9, MMP14, MT-SP1 or any two of MMP2, MMP9, MMP14, MT-SP1 or each of MMP2, MMP9, MMP14, MT-SP1. In some aspects, the isolated polypeptide is a molecule in which the % cleavability of the CM is improved by 1.5x, 2x, 2.5x, 3x, 5x, 7x, 8x, or 10x or more over the % cleavability of SEQ ID NO: 78 (see, e.g., Example 2). According to some embodiments of the present disclosures, the isolated polypeptide is a molecule that has high *in vivo* stability such that it is not cleaved in plasma as demonstrated by less than 50%, less than 40%, or less than 25% in vivo activation following 7 days of administration *in vivo* (see, e.g., Example 3). According to embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^2

$M^{-1} s^{-1}$. According to some embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^3 $M^{-1} s^{-1}$. According to some embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^4 $M^{-1} s^{-1}$. According to some embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^5 $M^{-1} s^{-1}$.

[0008] In some embodiments, the CM1 comprises the amino acid sequence of DHQSRS (SEQ ID NO: 5). In some embodiments, the CM1 comprises the amino acid sequence of SDHQSRS (SEQ ID NO: 6). In some embodiments, the CM2 comprises the amino acid sequence of AIALYA (SEQ ID NO: 8). In some embodiments, the CM2 comprises the amino acid sequence of AIALYAD (SEQ ID NO: 7). In some embodiments, the substrate comprises the amino acid sequence of AIALYADHQSRS (SEQ ID NO: 10), DHQSRSAIALYAD (SEQ ID NO: 11), HQSRSAIALYAD (SEQ ID NO: 12), DHQSRSAIALYA (SEQ ID NO: 13), or HQSRSAIALYA (SEQ ID NO: 14). In some embodiments, the substrate comprises the amino acid sequence of DHQSRSAIALYA (SEQ ID NO: 13).

[0009] In some embodiments, the substrate comprises the amino acid sequence of HQSRSAIALY (SEQ ID NO: 636). In some embodiments, the substrate comprises the amino acid sequence of HQSRSAIAL (SEQ ID NO: 637). In some embodiments, the substrate comprises the amino acid sequence of AIALHQSRS (SEQ ID NO: 41). In some embodiments, the substrate comprises the amino acid sequence of AIALYAHQSRS (SEQ ID NO: 33) or AIALYHQSRS (SEQ ID NO: 37). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPRGLL (SEQ ID NO: 638). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPAGLL (SEQ ID NO: 639). In some embodiments, the substrate comprises the amino acid sequence of HQSRSPWG (SEQ ID NO: 640). In some embodiments, the substrate comprises the amino acid sequence of HQSRAGPWGLL (SEQ ID NO: 645). In some embodiments, the substrate comprises the amino acid sequence of HQSRSAPWGLL (SEQ ID NO: 646). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGAWGLL (SEQ ID NO: 647). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPWALL (SEQ ID NO: 648). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPWGAL (SEQ ID NO: 649). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPWGLA (SEQ ID NO: 650). In some embodiments, the substrate comprises the amino acid sequence of HQSRSAIALYA (SEQ ID NO: 14). In some

embodiments, the substrate comprises the amino acid sequence of HQSRSPWGL (SEQ ID NO: 126). In some embodiments, the substrate comprises the amino acid sequence of PWGLHQSRS (SEQ ID NO: 489). In some embodiments, the substrate comprises the amino acid sequence of PWGLLHQSRS (SEQ ID NO: 485). In some embodiments, the substrate comprises the amino acid sequence of PWGLHQSR (SEQ ID NO: 502). In some embodiments, the substrate comprises the amino acid sequence of PWGLHQSKS (SEQ ID NO: 490). In some embodiments, the substrate comprises the amino acid sequence of DHQSKSGPWGLL (SEQ ID NO: 178). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPWGL (SEQ ID NO: 122). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPYGLL (SEQ ID NO: 138). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPFGLL (SEQ ID NO: 158). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPRGLL (SEQ ID NO: 638). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPAGLL (SEQ ID NO: 639). In some embodiments, the substrate comprises the amino acid sequence of HQSRPWGL (SEQ ID NO: 134).

[0010] In some embodiments, the present disclosure includes an isolated polypeptide comprising a substrate, wherein the substrate comprises the amino acid sequence of AQSRSGPWGLL (SEQ ID NO: 641). In some embodiments, the substrate comprises the amino acid sequence of HASRSGPWGLL (SEQ ID NO: 642). In some embodiments, the substrate comprises the amino acid sequence of HQARSGPWGLL (SEQ ID NO: 643). In some embodiments, the substrate comprises the amino acid sequence of HQSASGPWGLL (SEQ ID NO: 644).

[0011] In some embodiments, the CM2 comprises the amino acid sequence GPX_1GLX_2 (SEQ ID NO: 651), wherein X_1 is W, Y, F, R, K, Q, A, or M and wherein X_2 is any amino acid or absent.

[0012] In some embodiments, the CM2 comprises the amino acid sequence of PXGL, wherein X is alanine (A), tyrosine (Y), arginine (R), phenylalanine (F), tryptophan (W), lysine (K), glutamine (Q), or methionine (M).

[0013] In certain aspects, the present disclosure may include substitution of any alanine in the disclosed CM sequences with a valine.

[0014] In some embodiments, the CM2 comprises the amino acid sequence of GPWGLL (SEQ ID NO: 237). In some embodiments, the substrate comprises the amino acid sequence of HQSRSGPWGLL (SEQ ID NO: 15), DHQSRSGPWGLL (SEQ ID NO: 16), or

GPWGLLDHQSRS (SEQ ID NO: 17). In some embodiments, the CM2 comprises the amino acid sequence of SGLLSS (SEQ ID NO: 238). In some embodiments, the CM2 comprises the amino acid sequence of ISSGLLSS (SEQ ID NO: 75). In some embodiments, the substrate comprises the amino acid sequence of ISSGLLSSDHQSRS (SEQ ID NO: 18), DHQSRSISSGLLSS (SEQ ID NO: 19), or DHQSRSSGLLSS (SEQ ID NO: 20). In some embodiments, the CM2 comprises the amino acid sequence of ALAHGLF (SEQ ID NO: 74). In some embodiments, the substrate comprises the amino acid sequence of ALAHGLFSDHQSRS (SEQ ID NO: 21) or SDHQSRSALAHGLF (SEQ ID NO: 22).

[0015] In another aspect, the present disclosure provides an isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the substrate comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650, optionally wherein the substrate comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, and 460-634.

[0016] In some embodiments, the CM1 and the CM2 has a structural arrangement from N-terminus to C-terminus as CM1-CM2. In some embodiments, the CM1 and the CM2 has a structural arrangement from N-terminus to C-terminus as CM2-CM1. In some embodiments, the CM1 and the CM2 are indirectly coupled via a linking peptide. In some embodiments, the CM1 and the CM2 are directly coupled to each other. In some embodiments, at least a portion of the CM1 overlaps with at least a portion of the CM2. In some embodiments, the substrate comprises one or more CMs in addition to the CM1 and the CM2. In some embodiments, the substrate comprises CM1 cleavable by a first protease, CM2 cleavable by a second protease, CM3 cleavable by a third protease, and CM4 cleavable by a fourth protease.

[0017] In some embodiments, the isolated polypeptide is an activatable molecule and further comprises an "active moiety" (AM) that specifically binds a target. In some embodiments, the AM is a therapeutic macromolecule. In some embodiments, the AM is an antibody or antigen binding fragment thereof. In some embodiments, the antibody is a full-length antibody. single-chain variable fragment (scFv), diabody (a noncovalent dimer of scFv), single chain antibody (scab), a VHH, a domain antibody (dAb) or single domain antibody (nanobody, e.g., single domain heavy chain antibody, single domain light chain antibody). In some embodiments, the antibody is a monoclonal antibody, single chain antibody, Fab fragment, $F(ab')_2$ fragment, single-chain variable fragment (scFv), diabody (a noncovalent dimer of scFv), single chain antibody (scab), a VHH, a domain antibody (dAb)

or single domain antibody (nanobody, e.g., single domain heavy chain antibody, single domain light chain antibody). According to some embodiments of the present disclosures, the isolated polypeptide is an activatable molecule that has high *in vivo* stability such that it is not cleaved in plasma as demonstrated by less than 50%, less than 40%, or less than 25% in vivo activation following 7 days of administration *in vivo* (e.g., as exemplified in Example 3). According to some embodiments of the present disclosures, the isolated polypeptide is an activatable molecule that has masking efficiency of 18x, 20x, 30x, 33x, 50x, 60x, 82x, 95x, 105x, 121x, 123x, 150x, 200x, 234x, 300x, or higher (e.g., as exemplified in Example 4). According to some embodiments of the present disclosures, the activatable molecule is activated by one, two, or all of MMP2, MMP9, MMP14 and MT-SP1. According to some embodiments of the present disclosures, the activatable molecule is activated to an extent of having a cleavability percentage of at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, e.g., at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100% cleavable by any one of MMP2, MMP9, MMP14 and MT-SP1 or any two of MMP2, MMP9, MMP14 and MT-SP1or each of MMP2, MMP9, MMP14 and MT-SP1. In some embodiments, the AM is a cytokine. In some embodiments, the AM is a chimeric antigen receptor. In some aspects, the AM is a drug or agent, e.g., a therapeutic, imaging, or diagnostic agent.

[0018] In some embodiments, the AM is coupled to the CM1. In some embodiments, the AM is directly coupled to the CM1. In some embodiments, the AM is indirectly coupled to the CM1 via a linking peptide. In some embodiments, the AM is coupled to the CM2. In some embodiments, the AM is directly coupled to the CM2. In some embodiments, the AM is indirectly coupled to the CM2 via a linking peptide. In some embodiments, the AM is coupled to the CM3. In some embodiments, the AM is directly coupled to the CM3. In some embodiments, the AM is indirectly coupled to the CM3 via a linking peptide. In some embodiments, the AM is indirectly coupled to the CM1 via one or more components of the activatable protein. In some embodiments, the AM is indirectly coupled to the CM2 via one or more components of the activatable protein. In some embodiments, the AM is indirectly coupled to the CM3 via one or more components of the activatable protein.

[0019] In some embodiments, the isolated polypeptide further comprises a masking moiety (MM). In some embodiments, the MM has a dissociation constant for binding to the AM that is greater than the dissociation constant of the AM for binding to the target. In some embodiments, the MM does not interfere or compete with the AM for binding to the target in

in the activated molecule (*i.e.*, following cleavage of the CM by a protease). In some embodiments, the MM is 2 to 40 amino acids in length. In some embodiments, the MM does not bind the AM, but still interferes with AM's binding to its binding partner through non-specific interactions. In some embodiments, the MM is a steric mask. In some embodiments, the MM is a protein. In some embodiments, the MM is linked to the substrate such that the isolated polypeptide comprises a structural arrangement from N-terminus to C-terminus as follows: MM-CM1-CM2-AM, MM-CM2-CM1-AM, AM-CM1-CM2-MM, AM-CM2-CM1-MM, MM-CM2-CM1-CM3-AM, MM-CM1-CM2-CM3-AM, MM-CM1-CM3-CM2-AM, MM-CM3-CM1-CM2-AM, or MM-CM3-CM2-CM1-AM. In some embodiments, the MM is directly coupled to the CM1. In some embodiments, the MM is directly coupled to the CM2. In some embodiments, the MM is directly coupled to the CM3.

[0020] In some embodiments, the MM is indirectly coupled to the CM1. In some embodiments, the MM is indirectly coupled to the CM2. In some embodiments, the MM is indirectly coupled to the CM3. In some embodiments, the MM is indirectly coupled to the CM1 via a linking peptide. In some embodiments, the MM is indirectly coupled to the CM2 via a linking peptide. In some embodiments, the MM is indirectly coupled to the CM3 via a linking peptide. In some embodiments, the isolated polypeptide comprises a linking peptide (LP) and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM1-CM2-AM, MM-CM1-CM2-LP-AM, MM-LP-CM2-CM1-AM, MM-CM2-CM1-LP-AM, MM-LP-CM2-CM1-CM3-AM, MM-LP-CM1-CM2-CM3-AM, MM-LP-CM1-CM3-CM2-AM, MM-LP-CM3-CM1-CM2-AM, MM-LP-CM3-CM2-CM1-AM, MM-CM2-CM1-CM3-LP-AM, MM-CM1-CM2-CM3-LP-AM, MM-CM1-CM3-CM2-LP-AM, MM-CM3-CM1-CM2-LP-AM, or MM-CM3-CM2-CM1-LP-AM. In some embodiments, the isolated polypeptide comprises a first linking peptide (LP1) and a second linking peptide (LP2), and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM- LP1-CM1-CM2-LP2-AM, MM-LP1-CM2-CM1-LP2-AM, AM-LP2-CM1-CM2-LP1-MM, or AM-LP2-CM2-CM1-LP1-MM, MM-LP1-CM2-CM1-CM3-LP2-AM, MM-LP1-CM1-CM2-CM3-LP2-AM, MM-LP1-CM1-CM3-CM2-LP2-AM, MM-LP1-CM3-CM1-CM2-LP2-AM, MM-LP1-CM3-CM2-CM1-LP2-AM, MM-LP2-CM2-CM1-CM3-LP1-AM, MM-LP2-CM1-CM2-CM3-LP1-AM, MM-LP2-CM1-CM3-CM2-LP1-AM, MM-LP2-CM3-CM1-CM2-LP1-AM, or MM-LP2-CM3-CM2-CM1-LP1-AM. In some embodiments, the LP1 and LP2 are not

identical to each other. In some embodiments, the LP1 and LP2 are identical to each other. In some embodiments, each of the LP1 and LP2 is a peptide of 1 to 20 amino acids in length.

[0021] In some embodiments, the isolated polypeptide comprises an additional linking peptide (LP3) and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM1-LP3-CM2-AM, MM-CM1-LP3-CM2-LP-AM, MM-LP-CM2-LP3-CM1-AM, MM-CM2-LP3-CM1-LP-AM, MM-LP-CM2-LP3-CM1-CM3-AM, MM-LP-CM1-LP3-CM2-CM3-AM, MM-LP-CM1-LP3-CM3-CM2-AM, MM-LP-CM3-LP3-CM1-CM2-AM, MM-LP-CM3-LP3-CM2-CM1-AM, MM-CM2-LP3-CM1-CM3-LP-AM, MM-CM1-LP3-CM2-CM3-LP-AM, MM-CM1-LP3-CM3-CM2-LP-AM, MM-CM3-LP3-CM1-CM2-LP-AM, MM-CM3-LP3-CM2-CM1-LP-AM, MM-LP-CM2-CM1-LP3-CM3-AM, MM-LP-CM1-CM2-LP3-CM3-AM, MM-LP-CM1-CM3-LP3-CM2-AM, MM-LP-CM3-CM1-LP3-CM2-AM, MM-LP-CM3-CM2-LP3-CM1-AM, MM-CM2-CM1-LP3-CM3-LP-AM, MM-CM1-CM2-LP3-CM3-LP-AM, MM-CM1-CM3-LP3-CM2-LP-AM, MM-CM3-CM1-LP3-CM2-LP-AM, MM-CM3-CM2-LP3-CM1-LP-AM,MM-LP1-CM1-LP3-CM2-LP2-AM, MM-LP1-CM2-LP3-CM1-LP2-AM, AM-LP1-CM1-LP3-CM2-LP2-MM, or AM-LP1-CM2-LP3-CM1-LP2-MM, MM-LP1-CM2-LP3-CM1-CM3-LP2-AM, MM-LP1-CM1-LP3-CM2-CM3-LP2-AM, MM-LP1-CM1-LP3-CM3-CM2-LP2-AM, MM-LP1-CM3-LP3-CM1-CM2-LP2-AM, MM-LP1-CM3-LP3-CM2-CM1-LP2-AM, MM-LP2-CM2-LP3-CM1-CM3-LP1-AM, MM-LP2-CM1-LP3-CM2-CM3-LP1-AM, MM-LP2-CM1-LP3-CM3-CM2-LP1-AM, MM-LP2-CM3-LP3-CM1-CM2-LP1-AM, or MM-LP2-CM3-LP3-CM2-CM1-LP1-AM. In some embodiments, the LP1, LP2 and LP3 are not identical to each other. In some embodiments, the LP1, LP2 and LP3 are identical to each other. In some embodiments, each of the LP1, LP2 and LP3 is a peptide of 1 to 20 amino acids in length.

[0022] In some embodiments, the isolated polypeptide comprises an additional linking peptide (LP4) and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM2-LP3-CM1-LP4-CM3-AM, MM-LP-CM1-LP3-CM2-LP4-CM3-AM, MM-LP-CM1-LP3-CM3-LP4-CM2-AM, MM-LP-CM3-LP3-CM1-LP4-CM2-AM, MM-LP-CM3-LP3-CM2-LP4-CM1-AM, MM-CM2-LP3-CM1-LP4-CM3-LP-AM, MM-CM1-LP3-CM2-LP4-CM3-LP-AM, MM-CM1-LP3-CM3-LP4-CM2-LP-AM, MM-CM3-LP3-CM1-LP4-CM2-LP-AM, MM-CM3-LP3-CM2-LP4-CM1-LP-AM, MM-LP-CM2-LP4-CM1-LP3-CM3-AM, MM-LP-CM1-LP4-CM2-LP3-CM3-AM, MM-LP-CM1-LP4-CM3-LP3-CM2-AM, MM-LP-CM3-LP4-CM1-LP3-CM2-AM, MM-LP-

CM3-LP4-CM2-LP3-CM1-AM, MM-CM2-LP4-CM1-LP3-CM3-LP-AM, MM-CM1-LP4-CM2-LP3-CM3-LP-AM, MM-CM1-LP4-CM3-LP3-CM2-LP-AM, MM-CM3-LP4-CM1-LP3-CM2-LP-AM, MM-CM3-LP4-CM2-LP3-CM1-LP-AM, MM-LP1-CM2-LP3-CM1-LP4-CM3-LP2-AM, MM-LP1-CM1-LP3-CM2-LP4-CM3-LP2-AM, MM-LP1-CM1-LP3-CM3-LP4-CM2-LP2-AM, MM-LP1-CM3-LP3-CM1-LP4-CM2-LP2-AM, MM-LP1-CM3-LP3-CM2-LP4-CM1-LP2-AM, MM-LP2-CM2-LP3-CM1-LP4-CM3-LP1-AM, MM-LP2-CM1-LP3-CM2-LP4-CM3-LP1-AM, MM-LP2-CM1-LP3-CM3-LP4-CM2-LP1-AM, MM-LP2-CM3-LP3-CM1-LP4-CM2-LP1-AM, or MM-LP2-CM3-LP3-CM2-LP4-CM1-LP1-AM. In some embodiments, the LP1, LP2, LP3 and LP4 are not identical to each other. In some embodiments, the LP1, LP2, LP3 and LP4 are identical to each other. In some embodiments, each of the LP1, LP2, LP3 and LP4 is a peptide of 1 to 20 amino acids in length.

[0023] In general, in each embodiment herein, unless otherwise stated, a polypeptide may comprise one or more optional linkers between each of the elements listed, and such linkers may be 1 to 30, 6 to 29, 7 to 28, 8 to 27, 9 to 26, 10 to 25, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or 27 amino acids in length.

[0024] In some embodiments, the first protease is a serine protease. In some embodiments, the serine protease is membrane type serine protease 1 (MT-SP1). In some embodiments, the k_{cat}/K_M of the substrate by MT-SP1 cleavage is at least 1×10^3 $M^{-1}s^{-1}$. In some embodiments, the k_{cat}/K_M of the substrate by MT-SP1 cleavage is at least 1×10^4 $M^{-1}s^{-1}$. In some embodiments, the second protease is a matrix metalloproteinase (MMP). In some embodiments, the MMP is MMP2, MMP9, or MMP14. In some embodiments, the k_{cat}/K_M of the substrate by MMP14 cleavage is at least 1×10^3 $M^{-1}s^{-1}$. In some embodiments, the k_{cat}/K_M of the substrate by MMP14 cleavage is at least 1×10^4 $M^{-1}s^{-1}$. In some embodiments, the k_{cat}/K_M of the substrate by MMP9 cleavage is at least 1×10^3 $M^{-1}s^{-1}$. In some embodiments, the k_{cat}/K_M of the substrate by MMP9 cleavage is at least 1×10^4 $M^{-1}s^{-1}$. In some embodiments, the k_{cat}/K_M of the substrate by MMP2 cleavage is at least 1×10^3 $M^{-1}s^{-1}$. In some embodiments, the k_{cat}/K_M of the substrate by MMP2 cleavage is at least 1×10^4 $M^{-1}s^{-1}$.

[0025] In some embodiments, the isolated polypeptide is resistant to cleavage *in situ* in human bone marrow. In some embodiments, the isolated polypeptide is resistant to cleavage *in vivo* in human bone marrow.

[0026] In another aspect, the present disclosure provides an isolated polypeptide comprising a cleavable moiety (CM) comprising an amino acid sequence with one-amino

acid or two-amino acid mutation(s) of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650, wherein the CM is a substrate for a protease. In another aspect, the present disclosure provides an isolated polypeptide comprising a cleavable moiety (CM) comprising an amino acid sequence with one-amino acid or two-amino acid mutation(s) of any one of SEQ ID NOs: 10-73, 83-216, and 460-634, wherein the CM is a substrate for a protease. For example, the mutations may include substitution between any one of lysine, arginine, and histidine residues. In certain aspects, the present disclosure may include substitution of any arginine in the disclosed sequences with a lysine. In other aspects, the present disclosure also includes substitution of any arginine in the disclosed sequences with an amino acid that is not lysine. In certain aspects, the present disclosure may include substitution of any alanine in the disclosed sequences with a valine. For example, the mutations may include substitution between any one of alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, and tryptophan residues. For example, the mutations may include substitution between any one of glycine, asparagine, glutamine, cysteine, serine, threonine, and tyrosine residues. For example, the mutations may include substitution between any one of arginine, asparagine, aspartate, glutamine, glutamate, histidine, lysine, serine, and threonine residues. For example, the mutations may include substitution between any one of alanine, cysteine, isoleucine, leucine, methionine, phenylalanine, proline, tryptophan, tyrosine and valine residues. For example, the mutations may include substitution between any one of serine and threonine residues. For example, the mutations may include substitution between any one of asparagine and glutamine residues. For example, the mutations may include substitution between any one of alanine, valine, leucine and isoleucine residues. For example, the mutations may include substitution between any one of phenylalanine, tryptophan, and tyrosine residues.

[0027] In another aspect, the present disclosure provides a polypeptide complex comprising one or more of the isolated polypeptides herein bound to a second isolated polypeptide. In some aspects, the complex comprises one or more of the isolated polypeptides of the present disclosure bound to a second isolated polypeptide, e.g., via protein-protein affinity interactions, hydrophobic interactions, disulfide linkage(s), cross-link(s), covalent bond(s), chemical linkage(s), or any other type of binding between two polypeptides.

[0028] In another aspect, the present disclosure provides a conjugated polypeptide comprising the isolated polypeptide or AM herein conjugated to an agent. In some embodiments, the agent is conjugated to the isolated polypeptide or AM via a conjugating linker. In some embodiments, the conjugating linker is cleavable. In some embodiments, the

conjugating linker is non-cleavable. In some embodiments, the conjugating linker comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650, optionally selected from SEQ ID NOs: 10-73, 83-216, and 460-634. In some embodiments, the agent is a toxin, a microtubule inhibitor, a nucleic acid damaging agent, a dolastatin, an auristatin, a maytansinoid, a duocarmycin, a calicheamicin, or a combination thereof.

[0029] In another aspect, the present disclosure provides a composition comprising the isolated polypeptide, the polypeptide complex, the conjugated polypeptide herein, and a carrier. In some embodiments, the carrier is a pharmaceutically acceptable carrier. In some embodiments, the composition comprises an additional agent. In some embodiments, the additional agent is a therapeutic, imaging, or diagnostic agent.

[0030] In each of the foregoing embodiments, and unless otherwise stated, the polypeptide may comprise, e.g., one or more optional linkers between each of the elements listed. In some embodiments, a linker is a peptide having a length of 5 to 30, 6 to 29, 7 to 28, 8 to 27, 9 to 26, 10 to 25, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, or 24 amino acids. In the disclosed structural arrangements in the foregoing paragraphs and throughout this disclosure, one or more linkers may optionally be present between the elements. Further, this disclosure also contemplates and includes activatable proteins in which any one or more of the disclosed elements optionally directly abut each other such that there are no linkers or other amino acid sequences between the elements.

[0031] In another aspect, the present disclosure provides an isolated nucleic acid molecule encoding the isolated polypeptide herein.

[0032] In another aspect, the present disclosure provides a vector comprising the isolated nucleic acid molecule herein.

[0033] In another aspect, the present disclosure provides a cell comprising the isolated nucleic acid molecule or the vector herein.

[0034] In another aspect, the present disclosure provides a method of manufacturing an activatable molecule that contains a tandem substrate, the method comprising expressing and recovering a polypeptide comprising the isolated polypeptide herein.

[0035] In another aspect, the present disclosure provides a method of treating, alleviating a symptom of, or delaying the progression of a disease or disorder in a subject, comprising administering a therapeutically effective amount of the isolated polypeptide, the polypeptide complex, the conjugated polypeptide, or the composition herein to the subject. In some

embodiments, the disease is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

[0036] In another aspect, the present disclosure provides a method of detecting or diagnosing a disease or health condition of a subject, comprising: contacting the isolated polypeptide, the polypeptide complex, the conjugated polypeptide, or the composition herein with a sample from the subject; and measuring a level of cleavage of the isolated polypeptide, thereby detecting or diagnosing the disease or health condition of the subject. In some embodiments, the disease is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

BRIEF DESCRIPTION OF THE DRAWINGS

[0037] An understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention may be utilized, and the accompanying drawings of which:

[0038] **FIGs. 1A-1C** are graphs showing the *in vitro* masking efficiency of exemplary anti-EGFR activatable antibodies of the present disclosure. These exemplary results showed that the substrates affected the masking efficiency of the prodomain of the activatable antibody.

[0039] **FIGs. 2A** and **2C-2E** show the effects of activatable antibodies with exemplary substrates on H292 xenograft tumors. **FIG. 2B** shows the effects of activatable antibodies with exemplary substrates on SUM149 xenograft tumors. **FIGs. 2F-2I** show results of an intra-tumoral activation assay using activatable antibodies with exemplary substrates.

[0040] **FIG. 3** shows the *in situ* stability of activatable antibodies with exemplary substrates in human bone marrow aspirates.

[0041] **FIGs. 4A-4C** show the activation of activatable antibodies with exemplary substrates with colorectal cancer tumor tissue (**FIG. 4A**), head and neck cancer tumor tissue (**FIG. 3B**), and triple negative breast cancer tumor tissue (**FIG. 4C**)

[0042] **FIGs. 5A-5B** show the activation of activatable antibodies with exemplary substrates with multiple myeloma cells (**FIG. 5A**) and acute myeloid leukemia cells (**FIG. 5B**).

DETAILED DESCRIPTION OF THE INVENTION

[0043] Proteases play a critical role in the homeostasis of healthy tissues but are known to be dysregulated within diseases, including cancer and autoimmune disorders (Vasiljeva et al. "The multifaceted roles of tumor-associated proteases and harnessing their activity for prodrug activation," Biol. Chem. 2019 Apr 22). This dysregulation of protease activity provides new opportunities for the development of protease-activatable therapeutic molecules, which are preferentially activated in the local tissue microenvironment. These therapeutics have demonstrated a greater therapeutic window and safety profile with less on-target toxicities occurring in healthy tissues. Hence, there is a need for identification of substrates that act as cleavage recognition sites for proteases that are found to be dysregulated in disease tissues. These substrates or cleavable moieties (CMs) may have multiple cleavage sites for leveraging the activities of multiple disease-associated proteases.

[0044] Understanding the substrate cleavage profile and using these substrates as tools for activation in a specific disease or cancer type will enable the development of new therapeutic protease-activatable molecules. Fine tuning the therapeutic-activatable molecules by using protease substrates with unique cleavage profiles will allow for treatment options for a broader spectrum of patients while offering an improved therapeutic index. For example, "omics" studies have demonstrated the distribution of numerous matrix metalloproteases (MMPs) across numerous cancer types and differences in the expression of the MMPs compared to normal tissues (Gobin et al. "A pan-cancer perspective of matrix metalloproteases (MMP) gene expression profile and their diagnostic/prognostic potential," BMC Cancer. 2019 Jun 14; 19(1):581), highlighting the need for appropriate cleavable moiety selection. Indeed, the first protease-activatable antibodies were designed using MMP substrates (Bleuez et al., "Exploiting protease activation for therapy," Drug Discovery Today, 2022 Jun;27(6):1743-1754). In addition, membrane type serine protease 1 (MT-SP1) shows great potential for protease-activatable antibody development (Howng, B. et al. "Novel Ex Vivo Zymography Approach for Assessment of Protease Activity in Tissues with Activatable Antibodies," Pharmaceutics 2021, 13(9), 1390).

[0045] The present disclosure provides polypeptides comprising a substrate that comprise a plurality of cleavage moieties (CMs), each of which is cleavable by a protease. For example, the substrate may comprise a first CM (CM1) cleavable by a MMP and a second CM (CM2) cleavable by a serine protease. In some aspects, the substrates herein are cleaved in a diseased tissue (e.g., tumor tissue) but less in a healthy tissue. These substrates are useful in a variety of therapeutic, diagnostic and prophylactic applications. In some embodiments, the substrate-

containing polypeptides are activatable molecules and further comprise an active moiety (AM) that specifically binds a target. For example, the AM may be a therapeutic protein, a therapeutic agent, an imaging agent, a diagnostic agent, an antibody or antigen-binding fragment, a cytokine, chimeric antigen receptor or other molecules used in therapeutic and diagnostic applications.

[0046] Also provided herein are related compositions, kits, nucleic acids, vectors, and recombinant cells, as well as related methods, including methods of using and methods of producing any of the substrate-containing polypeptides described herein.

DEFINITIONS

[0047] 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 invention belongs. Methods and materials are described herein for use in the present disclosure; other, suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

[0048] The terms "a" and "an" refer to one or more (i.e., at least one) of the grammatical object of the article. By way of example, "a cell" encompasses one or more cells.

[0049] As used herein, the terms "about" and "approximately," when used to modify an amount specified in a numeric value or range, indicate that the numeric value as well as reasonable deviations from the value known to the skilled person in the art. For example ± 20%, ± 10%, or ± 5%, are within the intended meaning of the recited value where appropriate.

[0050] Concentrations, amounts, and other numerical data may be expressed or presented herein in a range format. It is to be understood that such a range format is used merely for convenience and brevity and thus should be interpreted flexibly to include not only the numerical values explicitly recited as the limits of the range, but also to include all the individual numerical values or sub-ranges encompassed within that range as if each numerical value and sub-range is explicitly recited. As an illustration, a numerical range of "about 0.01 to 2.0" should be interpreted to include not only the explicitly recited values of about 0.01 to about 2.0, but also include individual values and sub-ranges within the indicated range. Thus, included in this numerical range are individual values such as 0.5, 0.7, and 1.5, and sub-ranges such as from 0.5 to 1.7, 0.7 to 1.5, and from 1.0 to 1.5, etc. Furthermore, such an interpretation

should apply regardless of the breadth of the range or the characteristics being described. Additionally, it is noted that all percentages are in weight, unless specified otherwise.

[0051] In understanding the scope of the present disclosure, the terms "including" or "comprising" and their derivatives, as used herein, are intended to be open ended terms that specify the presence of the stated features, elements, components, groups, integers, and/or steps, but do not exclude the presence of other unstated features, elements, components, groups, integers and/or steps. The foregoing also applies to words having similar meanings such as the terms "including", "having" and their derivatives. The term "consisting" and its derivatives, as used herein, are intended to be closed terms that specify the presence of the stated features, elements, components, groups, integers, and/or steps, but exclude the presence of other unstated features, elements, components, groups, integers and/or steps. The term "consisting essentially of," as used herein, is intended to specify the presence of the stated features, elements, components, groups, integers, and/or steps as well as those that do not materially affect the basic and novel characteristic(s) of features, elements, components, groups, integers, and/or steps. It is understood that reference to any one of these transition terms (i.e. "comprising," "consisting," or "consisting essentially") provides direct support for replacement to any of the other transition term not specifically used. For example, amending a term from "comprising" to "consisting essentially of" or "consisting of" would find direct support due to this definition for any elements disclosed throughout this disclosure. Based on this definition, any element disclosed herein or incorporated by reference may be included in or excluded from the claimed invention.

[0052] As used herein, a plurality of compounds, elements, or steps may be presented in a common list for convenience. However, these lists should be construed as though each member of the list is individually identified as a separate and unique member. Thus, no individual member of such list should be construed as a *de facto* equivalent of any other member of the same list solely based on their presentation in a common group without indications to the contrary.

[0053] The term "exemplary" is used herein to mean serving as an example, instance, or illustration. Any aspect or design described herein as "exemplary" is not necessarily to be construed as preferred or advantageous over other aspects or designs. Rather, use of the word exemplary is intended to present concepts in a more concrete fashion.

[0054] Furthermore, certain molecules, constructs, compositions, elements, moieties, excipients, disorders, conditions, properties, steps, or the like may be discussed in the context

of one specific embodiment or aspect or in a separate paragraph or section of this disclosure. It is understood that this is merely for convenience and brevity, and any such disclosure is equally applicable to and intended to be combined with any other embodiments or aspects found anywhere in the present disclosure and claims, which all form the application and claimed invention at the filing date. For example, a list of constructs, molecules, method steps, kits, or compositions described with respect to a construct, molecule, isolated polypeptide, activatable molecule, composition, or method is intended to and does find direct support for embodiments related to constructs, molecules, isolated polypeptides, activatable molecules, compositions, formulations, and methods described in any other part of this disclosure, even if those method steps, active agents, kits, or compositions are not re-listed in the context or section of that embodiment or aspect.

[0055] The term "isolated polynucleotide" as used herein shall mean a polynucleotide of genomic, cDNA, RNA, mRNA, or synthetic origin or some combination thereof, which by virtue of its origin the "isolated polynucleotide" (1) is not associated with all or a portion of a polynucleotide in which the "isolated polynucleotide" is found in nature, (2) is operably linked to a polynucleotide which it is not linked to in nature, and/or (3) does not occur in nature as part of a larger sequence. In some embodiments, polynucleotides include the nucleic acid molecules encoding heavy chain immunoglobulin molecules, and nucleic acid molecules encoding light chain immunoglobulin molecules. The term isolated polypeptide includes and provides support for activatable molecules including activatable macromolecules, activatable polypeptides, activatable antibodies, activatable cytokines, and the like. The term isolated polypeptide includes and provides support for activatable molecules in which cleavage of the CM activates the molecule.

[0056] The term "isolated polypeptide" as used herein refers a polypeptide that is present in a form other than that found in nature. An "isolated polypeptide" as used herein may be encoded by cDNA, recombinant RNA, recombinant DNA, messenger RNA, or a polynucleotide of synthetic origin or some combination thereof. By virtue of its origin, or source of derivation, the "isolated polypeptide" (1) is not in a naturally occurring organism (e.g., is not an endogenous polypeptide of a naturally occurring organism) and (2) is present in a form not found in nature. In some aspects, the "isolated polypeptide" is expressed by a cell from a different species. In some aspects, the "isolated polypeptide" is a therapeutic protein or a diagnostic protein and not a naturally occurring protein. For example, as used

herein, the "isolated polypeptide" is not a plant protein or a protein naturally occurring in bacteria or other natural organisms.

[0057] The term "polypeptide" is used herein as a generic term to refer to a native protein, fragments, or analogs of a polypeptide sequence. Hence, proteins, protein fragments, and analogs are species of the polypeptide genus. In some embodiments, polypeptides in accordance with the disclosure comprise the heavy chain immunoglobulin, and the light chain immunoglobulin molecules, as well as antibody molecules formed by combinations comprising the heavy chain immunoglobulin molecules with light chain immunoglobulin molecules, such as kappa light chain immunoglobulin molecules, and vice versa, as well as fragments and analogs thereof.

[0058] As discussed herein, minor variations in the amino acid sequences of polypeptides are contemplated as being encompassed by the present disclosure, providing that the variations in the amino acid sequence maintain at least 75%, in some embodiments, at least 80%, at least 90%, at least 95%, and in some embodiments, at least 99% identity to the amino acid sequence that is not varied. In particular, conservative amino acid substitutions are contemplated. Conservative substitutions include those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids are generally divided into families: (1) acidic amino acids are aspartate, glutamate; (2) basic amino acids are lysine, arginine, histidine; (3) non-polar amino acids are alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar amino acids are glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. The hydrophilic amino acids include arginine, asparagine, aspartate, glutamine, glutamate, histidine, lysine, serine, and threonine. The hydrophobic amino acids include alanine, cysteine, isoleucine, leucine, methionine, phenylalanine, proline, tryptophan, tyrosine and valine. Other families of amino acids include (i) serine and threonine, which are the aliphatic-hydroxy family; (ii) asparagine and glutamine, which are the amide containing family; (iii) alanine, valine, leucine and isoleucine, which are the aliphatic family; and (iv) phenylalanine, tryptophan, and tyrosine, which are the aromatic family. For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid will not have a major effect on the binding or properties of the resulting molecule, especially if the replacement does not involve an amino acid within a framework site. Whether an amino acid change results in a functional peptide can readily be

determined by assaying the specific activity of the polypeptide derivative. Assays are described in detail herein. Fragments or analogs of antibodies or immunoglobulin molecules can be readily prepared by those of ordinary skill in the art. Suitable amino- and carboxyl-termini of fragments or analogs occur near boundaries of functional domains. Structural and functional domains can be identified by comparison of the nucleotide and/or amino acid sequence data to public or proprietary sequence databases. In some embodiments, computerized comparison methods are used to identify sequence motifs or predicted protein conformation domains that occur in other proteins of known structure and/or function. Methods to identify protein sequences that fold into a known three-dimensional structure are known, e.g., as described in Bowie *et al.* Science 253:164 (1991). Thus, the foregoing examples demonstrate that those of skill in the art can recognize sequence motifs and structural conformations that may be used to define structural and functional domains in accordance with the disclosure.

[0059] Suitable amino acid substitutions include those that: (1) alter susceptibility to proteolysis, (2) reduce susceptibility to oxidation, (3) alter binding affinity for forming protein complexes, (4) alter binding affinities, and (5) confer or modify other physicochemical or functional properties of such analogs. Analogs can include various muteins of a sequence other than the naturally-occurring peptide sequence. For example, single or multiple amino acid substitutions (for example, conservative amino acid substitutions) may be made in the naturally- occurring sequence (for example, in the portion of the polypeptide outside the domain(s) forming intermolecular contacts. A conservative amino acid substitution should not substantially change the structural characteristics of the parent sequence (*e.g.*, a replacement amino acid should not tend to break a helix that occurs in the parent sequence, or disrupt other types of secondary structure that characterizes the parent sequence). Examples of art-recognized polypeptide secondary and tertiary structures are described in Proteins, Structures and Molecular Principles (Creighton, Ed., W. H. Freeman and Company, New York (1984)); Introduction to Protein Structure (C. Branden and J. Tooze, eds., Garland Publishing, New York, N.Y. (1991)); and Thornton et at. Nature 354:105 (1991).

[0060] The term "sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. Included within the usage of the term "sample," therefore, is blood and a fraction or component of blood including blood serum, blood plasma, or lymph.

[0061] The term "therapeutic macromolecule" refers to any protein or nucleic acid that may be administered to a subject and have a therapeutic effect. In some embodiments, the therapeutic macromolecule may be a therapeutic polynucleotide or therapeutic polypeptide, i.e., a polynucleotide or polynucleotide that may be used in therapy.

[0062] As generally provided herein, an activatable molecule may comprise MM-substrate construct(s), also referred to herein as a prodomain. Accordingly, as used herein, the term "prodomain" refers to a polypeptide domain comprising a masking moiety (MM) and a cleavable substrate. The substrate may comprise one, two, three, or more CMs. In some embodiments, the MM and the substrate are separated by a linker, referred to herein as LP1. In some embodiments, the prodomain comprises a linker (referred to herein as LP2) that links the substrate of the prodomain to the active moiety (AM) in an activatable molecule. In some embodiments, the prodomain comprises a linker between the MM and the substrate and a linker between the substrate and the AM. In some embodiments, the MM and the substrate are not separated by a linker. In certain embodiments, a prodomain comprises one of the following formulas (where the formulas below represent amino acid sequences in either N- to C-terminal direction or C- to N-terminal direction): MM-LP1-substrate, MM-substrate-LP2, MM-LP1-substrate-LP2, or MM-substrate. As used herein and unless otherwise stated, each dash (-) between the components of the activatable molecule represents either a direct linkage or indirect linkage via one or more linking peptides.

[0063] Proteases are involved in the control of numerous physiological processes, and their dysregulation has been identified in a number of pathologies, such as, for example, oncological, cardiovascular, autoimmune, and neurodegenerative diseases. *See, e.g.*, O. Vasiljeva, et al., "Monitoring protease activity in biological tissues using antibody prodrugs as sensing probes," *Scientific Reports*, 10, 5894 (2020); O. Erster, et al., "Site-specific targeting of antibody activity in vivo mediated by disease-associated proteases," *J. Control Release*, 161(3):804-812 (2012); L. Desnoyers, et al., "Tumor-specific activation of an EGFR-targeting probody enhances therapeutic index," *Science Translational Medicine*, 5(207):207ra144 (2013); and B. Turk "Targeting proteases: successes, failures and future prospects" Nature Reviews Drug Discovery, 5 (2006). Protease-activated antibodies have been described in the literature that are activated by native proteases which are more prevalently active in, for example, tumor tissue, and the like, when compared to normal tissue. *Id.* These prodrugs have incorporated within their structure, a protease substrate that releases active drug following exposure to the appropriate protease and its subsequent cleavage. What

appears evident, however, is that the profile of dysregulated protease activity in diseased tissue may differ from one type of disease tissue/disorder to another. Thus, it is desirable to have a collection of substrates that target a variety of different protease activity profiles.

[0064] In some aspects, the present disclosure provides a substrate that comprises a first cleavable moiety (CM1) that is cleavable by a first protease and a second cleavable moiety (CM2) that is cleavable by a second protease. In some embodiments, the first protease is an MMP (e.g., MMP2, MMP9, or MMP14). In some embodiments, the first protease is a serine protease (e.g., a membrane type serine protease 1 (MT-SP1)). In some embodiments, the first protease is a serine protease (e.g., MT-SP1) and the second protease is an MMP (e.g., MMP2, MMP9, or MMP14) and. In some embodiments, the CM1 is cleavable by a first set of proteases and the CM2 is cleavable by a second set of proteases. In some examples, the first and the second sets of proteases may overlap, i.e., one or more proteases in the first set are also in the second set. In some alternative examples, the first and second sets of proteases do not overlap, i.e., the first set does not include any protease in the second set.

[0065] In certain aspects, the substrates are selectively cleavable by certain proteases (e.g., MT-SP1 and/or MMP), but have reduced or no cleavability by another protease. For example, the substrates may be resistant or substantially resistant to cleavage in bone marrow tissue, e.g., bone marrow aspirate. In some aspects, resistance of substrates to protease cleavage in healthy tissue may reduce systemic toxicities by limiting binding of the activatable molecule to targets that also may be present in healthy tissues. Therefore, substrates with bone marrow tissue resistance have the potential to demonstrate a greater therapeutic window and safety profile with less on-target toxicities occurring in healthy tissues.

[0066] In a specific aspect, the present disclosure provides polypeptides (e.g., isolated polypeptides) comprising a substrates comprising a plurality of cleavable moieties (CMs). A CM is a polypeptide that comprises a substrate for a sequence-specific protease. In some aspects, the present disclosure provides polypeptides and polypeptide complexes comprising a substrate and an active moiety.

[0067] In some embodiments, the CM1 in the substrate comprises the amino acid sequence of SEQ ID NO: 1. In some embodiments, the CM1 in the substrate comprises the amino acid sequence of SEQ ID NO: 2. In some embodiments, the CM1 in the substrate comprises the amino acid sequence of SEQ ID NO: 3. In some embodiments, the CM1 in the substrate comprises the amino acid sequence of SEQ ID NO: 4. In some embodiments, the

CM1 in the substrate comprises the amino acid sequence of SEQ ID NO: 5. In some embodiments, the CM1 in the substrate comprises the amino acid sequence of SEQ ID NO: 6. In some embodiments, the CM1 in the substrates comprises the amino acid sequence of SEQ ID NO: 77.

[0068] In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 1. In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 2. In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 3. In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 4. In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 5. In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 6. In some embodiments, the CM1 in the substrate consists of the amino acid sequence of SEQ ID NO: 77.

[0069] In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 7. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 8. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 9. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 74. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 75. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 76. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 237. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 238. In some embodiments, the CM2 in the substrate comprises the amino acid sequence of SEQ ID NO: 635. The substrate may consist of an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650.

[0070] In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 7. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 8. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 9. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 74. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 75. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 76. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of

SEQ ID NO: 237. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 238. In some embodiments, the CM2 in the substrate consists of the amino acid sequence of SEQ ID NO: 635. The substrate may comprise an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. The substrate may comprise an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 10. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 11. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 12. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 13. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 14. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 15. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 16. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 17. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 18. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 19. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 20. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 21. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 22. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 23. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 24. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 25. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 26. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 27. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 28. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 29. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 30. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 31. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 32. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 33. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 34. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 35. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 36. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 37. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 38. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 39. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 40. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 41. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 42. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 43. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 44. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 45. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 46. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 47. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 48. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 49. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 50. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 51. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 52. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 53. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 54. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 55. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 56. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 57. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 58. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 59. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 60. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 61. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 62. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 63. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 64. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 65. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 66. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 67. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 68. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 69. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 70. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 71. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 72. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 73. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 83. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 84. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 85. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 86. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 87. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 88. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 89. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 90. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 91. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 92. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 93. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 94. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 95. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 96. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 97. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 98. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 99. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 100. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 101. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 102. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 103. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 104. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 105. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 106. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 107. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 108. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 109. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 110. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 111. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 112. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 113. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 114. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 115. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 116. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 117. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 118. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 119. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 120. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 121. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 122. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 123. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 124. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 125. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 126. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 127. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 128. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 129. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 130. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 131. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 132. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 133. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 134. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 135. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 136. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 137. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 138. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 139. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 140. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 141. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 142. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 143. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 144. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 145. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 146. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 147. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 148. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 149. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 150. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 151. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 152. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 153. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 154. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 155. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 156. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 157. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 158. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 159. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 160. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 161. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 162. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 163. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 164. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 165. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 166. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 167. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 168. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 169. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 170. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 171. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 172. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 173. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 174. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 175. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 176. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 177. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 178. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 179. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 180. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 181. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 182. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 183. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 184. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 185. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 186. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 187. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 188. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 189. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 190. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 191. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 192. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 193. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 194. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 195. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 196. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 197. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 198. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 199. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 200. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 201. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 202. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 203. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 204. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 205. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 206. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 207. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 208. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 209. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 210. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 211. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 212. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 213. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 214. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 215. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 216. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 460. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 461. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 462. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 463. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 464. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 465. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 466. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 467. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 468. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 469. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 470. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 471. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 472. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 473. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 474. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 475. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 476. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 477. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 478. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 479. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 480. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 481. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 482. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 483. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 484. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 485. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 486. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 487. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 488. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 489. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 490. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 491. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 492. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 493. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 494. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 495. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 496. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 497. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 498. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 499. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 500. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 501. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 502. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 503. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 504. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 505. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 506. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 507. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 508. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 509. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 510. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 511. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 512. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 513. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 514. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 515. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 516. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 517. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 518. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 519. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 520. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 521. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 522. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 523. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 524. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 525. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 526. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 527. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 528. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 529. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 530. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 531. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 532. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 533. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 534. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 535. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 536. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 537. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 538. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 539. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 540. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 541. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 542. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 543. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 544. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 545. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 546. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 547. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 548. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 549. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 550. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 551. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 552. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 553. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 554. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 555. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 556. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 557. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 558. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 559. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 560. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 561. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 562. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 563. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 564. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 565. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 566. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 567. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 568. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 569. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 570. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 571. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 572. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 573. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 574. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 575. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 576. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 577. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 578. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 579. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 580. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 581. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 582. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 583. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 584. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 585. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 586. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 587. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 588. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 589. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 590. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 591. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 592. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 593. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 594. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 595. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 596. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 597. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 598. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 599. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 600. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 601. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 602. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 603. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 604. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 605. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 606. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 607. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 608. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 609. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 610. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 611. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 612. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 613. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 614. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 615. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 616. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 617. In some embodiments, the substrate comprises the amino acid sequence of SEQ
ID NO: 618. In some embodiments, the substrate comprises the amino acid sequence of SEQ

ID NO: 619. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 620. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 621. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 622. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 623. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 624. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 625. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 626. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 627. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 628. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 629. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 630. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 631. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 632. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 633. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 634. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 636. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 637. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 638. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 639. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 640. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 641. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 642. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 643. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 644. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 645. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 646. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 647. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 648. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 649. In some embodiments, the substrate comprises the amino acid sequence of SEQ ID NO: 650.

[0071] The substrate may consist of an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. The substrate may consist of an amino acid sequence

selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 10. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 11. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 12. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 13. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 14. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 15. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 16. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 17. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 18. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 19. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 20. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 21. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 22. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 23. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 24. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 25. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 26. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 27. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 28. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 29. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 30. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 31. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 32. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 33. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 34. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 35. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 36. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 37. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 38. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 39. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 40. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 41. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 42. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 43. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 44. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 45. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 46. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 47. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 48. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 49. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 50. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 51. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 52. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 53. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 54. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 55. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 56. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 57. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 58. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 59. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 60. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 61. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 62. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 63. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 64. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 65. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 66. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 67. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 68. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 69. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 70. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 71. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 72. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 73. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 83. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 84. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 85. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 86. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 87. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 88. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 89. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 90. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 91. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 92. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 93. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 94. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 95. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 96. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 97. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 98. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 99. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 100. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 101. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 102. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 103. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 104. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 105. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 106. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 107. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 108. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 109. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 110. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 111. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 112. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 113. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 114. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 115. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 116. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 117. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 118. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 119. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 120. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 121. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 122. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 123. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 124. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 125. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 126. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 127. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 128. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 129. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 130. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 131. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 132. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 133. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 134. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 135. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 136. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 137. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 138. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 139. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 140. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 141. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 142. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 143. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 144. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 145. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 146. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 147. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 148. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 149. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 150. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 151. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 152. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 153. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 154. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 155. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 156. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 157. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 158. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 159. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 160. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 161. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 162. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 163. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 164. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 165. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 166. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 167. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 168. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 169. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 170. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 171. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 172. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 173. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 174. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 175. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 176. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 177. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 178. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 179. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 180. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 181. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 182. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 183. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 184. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 185. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 186. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 187. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 188. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 189. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 190. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 191. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 192. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 193. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 194. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 195. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 196. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 197. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 198. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 199. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 200. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 201. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 202. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 203. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 204. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 205. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 206. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 207. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 208. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 209. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 210. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 211. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 212. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 213. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 214. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 215. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 216. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 460. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 461. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 462. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 463. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 464. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 465. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 466. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 467. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 468. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 469. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 470. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 471. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 472. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 473. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 474. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 475. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 476. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 477. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 478. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 479. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 480. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 481. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 482. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 483. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 484. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 485. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 486. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 487. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 488. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 489. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 490. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 491. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 492. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 493. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 494. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 495. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 496. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 497. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 498. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 499. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 500. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 501. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 502. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 503. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 504. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 505. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 506. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 507. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 508. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 509. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 510. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 511. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 512. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 513. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 514. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 515. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 516. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 517. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 518. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 519. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 520. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 521. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 522. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 523. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 524. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 525. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 526. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 527. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 528. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 529. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 530. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 531. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 532. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 533. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 534. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 535. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 536. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 537. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 538. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 539. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 540. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 541. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 542. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 543. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 544. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 545. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 546. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 547. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 548. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 549. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 550. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 551. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 552. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 553. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 554. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 555. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 556. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 557. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 558. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 559. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 560. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 561. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 562. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 563. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 564. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 565. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 566. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 567. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 568. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 569. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 570. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 571. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 572. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 573. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 574. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 575. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 576. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 577. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 578. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 579. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 580. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 581. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 582. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 583. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 584. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 585. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 586. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 587. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 588. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 589. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 590. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 591. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 592. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 593. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 594. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 595. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 596. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 597. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 598. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 599. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 600. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 601. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 602. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 603. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 604. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 605. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 606. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 607. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 608. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 609. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 610. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 611. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 612. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 613. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 614. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 615. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 616. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 617. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 618. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 619. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 620. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 621. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 622. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 623. In some embodiments,

the substrate consists of the amino acid sequence of SEQ ID NO: 624. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 625. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 626. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 627. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 628. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 629. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 630. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 631. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 632. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 633. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 634. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 636. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 637. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 638. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 639. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 640. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 641. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 642. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 643. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 644. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 645. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 646. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 647. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 648. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 649. In some embodiments, the substrate consists of the amino acid sequence of SEQ ID NO: 650.

[0072] In some embodiments, the substrate comprises a combination, a C-terminal truncation variant, a C-terminal extension variant, an N-terminal truncation variant, or an N-terminal extension variant of the amino acid sequences of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some embodiments, the substrate comprises a combination, a C-terminal truncation variant, a C-terminal extension variant, an N-terminal truncation variant, or an N-terminal extension variant of the amino acid sequences of any one of SEQ ID NOs: 10-73, 83-216, and 460-634. Truncation variants of the aforementioned

amino acid sequences that are suitable for use in a substrate may be any that retain the recognition site for the corresponding protease. These include C-terminal and/or N-terminal truncation variants comprising at least 1, 2, 3, 4, 5, or more contiguous amino acids of the above-described amino acid sequences that retain a recognition site for a protease. In certain embodiments, the truncation variant comprises a C-terminal deletion and/or an N-terminal deletion of one amino acid residue from an amino acid sequence selected from the group consisting of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In certain embodiments, the truncation variant comprises a C-terminal deletion and/or an N-terminal deletion of one amino acid residue from an amino acid sequence selected from the group consisting of SEQ ID NOs: 10-73, 83-216, and 460-634. Extension variants of the aforementioned amino acid sequences that are suitable for use in a substrate may be any that have one or more (e.g., 1, 2, 3, 4, 5 or more) additional amino acids and retain the recognition site for the corresponding protease. In some examples, the additional amino acids are coupled to the C-terminus of the aforementioned amino acid sequences. In some examples, the additional amino acids are coupled to the N-terminus of the aforementioned amino acid sequences. In some examples, the extension variants may comprise additional amino acids coupled to both the C-terminus and the N-terminus of the aforementioned amino acid sequences. In some instances, the C-terminus or N-terminus extension variants can have a C-terminal glycine or an N-terminal serine amino acid.

[0073] In some embodiments, the substrate comprises one, two, three, four, five, six or more amino acids in addition to the amino acid sequence of any one of SEQ ID Nos: 10-73, 83-216, 460-634, and 636-650. In some embodiments, the substrate comprises one, two, three, four, five, six or more amino acids in addition to the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, and 460-634. In some examples, the substrate comprises one, two, three, four, five, six or more additional amino acids at the N-terminus of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some examples, the substrate comprises one, two, three, four, five, six or more additional amino acids at the N-terminus of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, and 460-634. In some examples, the substrate comprises one, two, three, four, five, six or more additional amino acids at the C-terminus of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some examples, the substrate comprises one, two, three, four, five, six or more additional amino acids at the C-terminus of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, and 460-634. In some

examples, the substrate comprises one, two, three, four, five, six or more additional amino acids at the N-terminus, and one, two, three, four, five, six or more additional amino acids at the C-terminus of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some examples, the substrate comprises one, two, three, four, five, six or more additional amino acids at the N-terminus, and one, two, three, four, five, six or more additional amino acids at the C-terminus of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, and 460-634.

[0074] In some embodiments, the substrate comprises a sequence with mutation(s) of one or more amino acid of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. In some embodiments, the substrate comprises a sequence with mutation(s) of one or more amino acid of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, and 460-634. For example, the substrate comprises a sequence with one-amino acid, two-amino acid, three-amino acid, four-amino acid, or five-amino acid mutation(s) of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216 460-634, and 636-650. For example, the substrate comprises a sequence with one-amino acid, two-amino acid, three-amino acid, four-amino acid, or five-amino acid mutation(s) of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216 and 460-634. In some embodiments, the substrate comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650 and having one conservative substitution. In some embodiments, the substrate comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 10-73, 83-216, and 460-634 and having one conservative substitution.

[0075] In some embodiments, the substrate consists of a sequence with mutation(s) of one or more amino acid of the amino acid sequence of any one of SEQ ID Nos: 10-73, 83-216, 460-634, and 636-650. In some embodiments, the substrate consists of a sequence with mutation(s) of one or more amino acid of the amino acid sequence of any one of SEQ ID Nos: 10-73, 83-216, and 460-634. For example, the substrate consists of a sequence with one-amino acid, two-amino acid, three-amino acid, four-amino acid, or five-amino acid mutation(s) of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650. For example, the substrate consists of a sequence with one-amino acid, two-amino acid, three-amino acid, four-amino acid, or five-amino acid mutation(s) of the amino acid sequence of any one of SEQ ID NOs: 10-73, 83-216, and 460-634.

[0076] In some embodiments, the substrate comprises a total of 3 amino acids to 25 amino acids. For example, the substrate may comprise a total of 3 to 25, 3 to 20, 3 to 15, 3 to 10, 3 to 5, 5 to 25, 5 to 20, 5 to 15, 5 to 10, 10 to 25, 10 to 20, 10 to 15, 15 to 25, 15 to 20, or 20 to 25 amino acids. In some embodiments, the substrate consists of a total of 3 amino acids to 25 amino acids. For example, the substrate may consist of a total of 3 to 25, 3 to 20, 3 to 15, 3 to 10, 3 to 5, 5 to 25, 5 to 20, 5 to 15, 5 to 10, 10 to 25, 10 to 20, 10 to 15, 15 to 25, 15 to 20, or 20 to 25 amino acids.

[0077] The substrate may be specifically cleaved by one or more protease (e.g., by MT-SP1 and/or MMP) at a desired rate. The rate may be measured as substrate cleavage kinetics (k_{cat}/K_M) as disclosed in WO2016118629, which is incorporated by reference in its entirety. In brief, k_{cat} is the turnover number and describes how many substrate molecules are transformed into products per unit time by a protease. The K_M value describes the affinity of the substrate to the active site of the protease. The k_{cat}/K_M ratio provides a measurement of cleavability of the substrate by the protease. In general, the greater the ratio, the higher the rate of cleavability is; conversely, the lower the ratio, the slower the rate of cleavability is. The k_{cat}/K_M values may be determined with the following equation

$$\frac{k_{cat}}{K_M} = -\ln(1 - C)/(t*p)$$

where C is product conversion, t is time (s), and p is protease concentration (M), which assumes that the substrate concentration is below the K_M and in excess of the protease concentration.

[0078] In some embodiments, the substrate is cleaved by MT-SP1 at a rate that has a k_{cat}/K_M value from 1×10 to 1×10^6 $M^{-1}s^{-1}$, e.g., from 1×10 to 5×10, from 5×10 to 1×10^2, from 1×10^2 to 5×10^2, from 5×10^2 to 1×10^3, from 1×10^3 to 5×10^3, from 5×10^3 to 1×10^4, from 1×10^4 to 5×10^4, from 5×10^4 to 1×10^5, from 1×10^5 to 5×10^5, or from 5×10^5 to 1×10^6 $M^{-1}s^{-1}$. In some embodiments, the substrate is cleaved by MT-SP1 at a rate that has a k_{cat}/K_M value of at least 1×10, at least 5×10, at least 1×10^2, at least 5×10^2, at least 1×10^3, 5×10^3, at least 1×10^4, at least 5×10^4, at least 1×10^5, at least 5×10^5, or at least 1×10^6 $M^{-1}s^{-1}$.

[0079] In some embodiments, the substrate is cleaved by an MMP at a rate that has a k_{cat}/K_M value from 1×10 to 1×10^6 $M^{-1}s^{-1}$, e.g., from 1×10 to 5×10, from 5×10 to 1×10^2, from 1×10^2 to 5×10^2, from 5×10^2 to 1×10^3, from 1×10^3 to 5×10^3, from 5×10^3 to 1×10^4, from 1×10^4 to 5×10^4, from 5×10^4 to 1×10^5, from 1×10^5 to 5×10^5, or from 5×10^5 to 1×10^6 $M^{-1}s^{-1}$. In some

embodiments, the substrate is cleaved by an MMP at a rate that has a k_{cat}/K_M value of at least 1×10, at least 5×10, at least 1×10^2, at least 5×10^2, at least 1×10^3, 5×10^3, at least 1×10^4, at least 5×10^4, at least 1×10^5, at least 5×10^5, or at least 1×10^6 $M^{-1}s^{-1}$.

[0080] In some embodiments, the substrate is cleaved by MMP2 at a rate that has a k_{cat}/K_M value from 1×10 to 1×10^6 $M^{-1}s^{-1}$, e.g., from 1×10 to 5×10, from 5×10 to 1×10^2, from 1×10^2 to 5×10^2, from 5×10^2 to 1×10^3, from 1×10^3 to 5×10^3, from 5×10^3 to 1×10^4, from 1×10^4 to 5×10^4, from 5×10^4 to 1×10^5, from 1×10^5 to 5×10^5, or from 5×10^5 to 1×10^6 $M^{-1}s^{-1}$. In some embodiments, the substrate is cleaved by MMP2 at a rate that has a k_{cat}/K_M value of at least 1×10, at least 5×10, at least 1×10^2, at least 5×10^2, at least 1×10^3, 5×10^3, at least 1×10^4, at least 5×10^4, at least 1×10^5, at least 5×10^5, or at least 1×10^6 $M^{-1}s^{-1}$.

[0081] In some embodiments, the substrate is cleaved by MMP9 at a rate that has a k_{cat}/K_M value from 1×10 to 1×10^6 $M^{-1}s^{-1}$, e.g., from 1×10 to 5×10, from 5×10 to 1×10^2, from 1×10^2 to 5×10^2, from 5×10^2 to 1×10^3, from 1×10^3 to 5×10^3, from 5×10^3 to 1×10^4, from 1×10^4 to 5×10^4, from 5×10^4 to 1×10^5, from 1×10^5 to 5×10^5, or from 5×10^5 to 1×10^6 $M^{-1}s^{-1}$. In some embodiments, the substrate is cleaved by MMP9 at a rate that has a k_{cat}/K_M value of at least 1×10, at least 5×10, at least 1×10^2, at least 5×10^2, at least 1×10^3, 5×10^3, at least 1×10^4, at least 5×10^4, at least 1×10^5, at least 5×10^5, or at least 1×10^6 $M^{-1}s^{-1}$.

[0082] In some embodiments, the substrate is cleaved by MMP14 at a rate that has a k_{cat}/K_M value from 1×10 to 1×10^6 $M^{-1}s^{-1}$, e.g., from 1×10 to 5×10, from 5×10 to 1×10^2, from 1×10^2 to 5×10^2, from 5×10^2 to 1×10^3, from 1×10^3 to 5×10^3, from 5×10^3 to 1×10^4, from 1×10^4 to 5×10^4, from 5×10^4 to 1×10^5, from 1×10^5 to 5×10^5, or from 5×10^5 to 1×10^6 $M^{-1}s^{-1}$. In some embodiments, the substrate is cleaved by MMP14 at a rate that has a k_{cat}/K_M value of at least 1×10, at least 5×10, at least 1×10^2, at least 5×10^2, at least 1×10^3, 5×10^3, at least 1×10^4, at least 5×10^4, at least 1×10^5, at least 5×10^5, or at least 1×10^6 $M^{-1}s^{-1}$.

[0083] In some embodiments, the substrate is cleaved by a protease in a bone marrow tissue, e.g., in a bone marrow aspirate, at a rate that has a k_{cat}/K_M value from 1×10 to 1×10^6 $M^{-1}s^{-1}$, e.g., from 1×10 to 5×10, from 5×10 to 1×10^2, from 1×10^2 to 5×10^2, from 5×10^2 to 1×10^3, from 1×10^3 to 5×10^3, from 5×10^3 to 1×10^4, from 1×10^4 to 5×10^4, from 5×10^4 to 1×10^5, from 1×10^5 to 5×10^5, or from 5×10^5 to 1×10^6 $M^{-1}s^{-1}$. In some embodiments, the substrate is cleaved by a protease in a bone marrow tissue, e.g. in a bone marrow aspirate at a rate that has a k_{cat}/K_M value of at least 1×10, at least 5×10, at least 1×10^2, at least 5×10^2, at least 1×10^3, 5×10^3, at least 1×10^4, at least 5×10^4, at least 1×10^5, at least 5×10^5, or at least 1×10^6 $M^{-1}s^{-1}$.

[0084] In some embodiments, the cleavability of the substrates are presented as the percentage of the fraction of cleaved substrates (or polypeptides comprising the substrates), e.g., as determined in a capillary electrophoresis assay described Example 2. In some examples, the cleavability of the substrate by a protease is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%. In some examples, the cleavability of the substrate by MT-SP1 is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, when 500nM activatable antibody c225 containing a prodomain with the substrate being tested was incubated with 10 nM of MT-SP1 for 1.5 hours at 37°C. In some examples, the cleavability of the substrate by an MMP is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100% when 500nM activatable antibody c225 containing a prodomain with the substrate being tested was incubated with 10 nM of MMP for 1.5 hours at 37°C. In some examples, the cleavability of the substrate by MMP2 is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, when 500nM activatable antibody c225 containing a prodomain with the substrate being tested was incubated with 10 nM of MMP2 for 1.5 hours at 37°C. In some examples, the cleavability of the substrate by MMP9 is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, when 500nM activatable antibody c225 containing a prodomain with the substrate being tested was incubated with 10 nM of MMP9 for 1.5 hours at 37°C. In some examples, the cleavability of the substrate by MMP14 is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, when 500nM activatable antibody c225 containing a prodomain with the substrate being tested was incubated with 10 nM of MMP14 for 1.5 hours at 37°C. In some examples, the cleavability of the substrate by a protease in a bone marrow tissue (e.g., in a bone marrow aspirate) is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, when incubated with a bone marrow tissue (e.g., bone marrow aspirate) for 1.5 hours or 4 hours at 37°C.

[0085] In some embodiments, for specific cleavage by an enzyme, contact between the enzyme and substrate is made. When a substrate-containing polypeptide (e.g., activatable molecule comprising an AM coupled to a MM and a substrate) is in the presence of target and sufficient protease activity, the substrate can be cleaved. Sufficient protease activity refers to the ability of the protease to access the substrate and effect cleavage.

[0086] In some embodiments, a substrate according to the present disclosure and a reference polypeptide can be cleaved by the same protease, but the substrate according to the

present disclosure has reduced cleavage or resistance to cleavage (e.g., by a different protease) in certain tissues *in situ* compared to a reference polypeptide. In some examples, a substrate according to the present disclosure and a reference polypeptide can be cleaved by MT-SP1 and/or MMP, but the substrate according to the present disclosure has reduced cleavage or resistance to cleavage (e.g., by a different protease(s) than MT-SP1 and/or MMP2, MMP9, and MMP14) in the bone marrow *in situ* compared to a reference polypeptide. For example, the cleavage (e.g., by a different protease than MT-SP1 and/or MMP) in the bone marrow *in situ* of the substrate may be less than 99%, less than 95%, less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% compared to the cleavage of the reference polypeptide. In some examples, such proteases different than MT-SP1 and/or an MMP may be other proteases in the bone marrow or other normal tissues, as well as other proteases involved in inflammation and wound healing. In some examples, the cleavage (e.g., by a different protease than MT-SP1 and/or MMP) in the bone marrow *in situ* may be measured by increased activity of an activatable molecule comprising the substrate or the reference polypeptide in the bone marrow, e.g., the method described in Example 6. A substrate that is resistant to cleavage by a protease, or a sample or tissue comprising a protease, refers to (i) a substrate in which no peptide bond is hydrolyzed by the protease, or no peptide bond is hydrolyzed when incubated in the sample or tissue comprising the protease, or (ii) a substrate in which a reduced level of peptide bond is hydrolyzed by the protease, or reduced level of peptide bond is hydrolyzed when incubated in the sample or tissue comprising the protease, compared to a reference substrate.

[0087] In some embodiments, the substrate is cleavable by more than two proteases. For example, the substrate may be cleaved by a serine protease (e.g., MT-SP1) and an MMP (e.g., MMP2, MMP9, and/or MMP14) and by one or more additional proteases. Examples of the additional protease could be any one or more of the following proteases: a disintegrin and metalloprotease (ADAM), an ADAM-like, or a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS, such as, for example, ADAM8, ADAM9, ADAM10, ADAM12, ADAM15, ADAM17/TACE, ADAMDEC1, ADAMTS1, ADAMTS4, ADAMTS5); an aspartate protease (such as, for example, BACE, Renin, and the like); an aspartic cathepsin (such as, for example, Cathepsin D, Cathepsin E, and the like); a caspase (such as, for example, Caspase 1, Caspase 2, Caspase 3, Caspase 4, Caspase 5, Caspase 6, Caspase 7, Caspase 8, Caspase 9, Caspase 10, Caspase 14, and the like); a cysteine cathepsin

(such as, for example, Cathepsin B, Cathepsin C, Cathepsin K, Cathepsin L, Cathepsin S, Cathepsin V/L2, Cathepsin X/Z/P); a cysteine proteinase (such as, for example, Cruzipain, Legumain, Otubain-2, and the like); a kallikrein-related peptidase (KLK) (such as, for example, KLK4, KLK5, KLK6, KLK7, KLK8, KLK10, KLK11, KLK13, KLK14, and the like); a metalloproteinase (such as, for example, Meprin, Neprilysin, prostate-specific membrane antigen (PSMA), bone morphogenetic protein 1 (BMP-1), and the like); a matrix metalloproteinase (MMP, such as, for example, MMP1, MMP2, MMP3, MMP7, MMP8, MMP9, MMP10, MMP11, MMP12, MMP13, MMP14, MMP15, MMP16, MMP17, MMP19, MMP20, MMP23, MMP24, MMP26, MMP27, and the like); a serine protease (such as, for example, activated protein C, Cathepsin A, Cathepsin G, Chymase, a coagulation factor protease (such as, for example, FVIIa, FIXa, FXa, FXIa, FXIIa, and the like); elastase, granzyme B, Guanidinobenzoatase, HtrA1, proteinase 3, neutrophil elastase, neutrophil serine protease 4 (NSP4), Lactoferrin, Marapsin, NS3/4A, PACE4, Plasmin, prostate-specific antigen (PSA), tissue plasminogen activator (tPA), Thrombin, Tryptase, urokinase-type plasminogen activator (uPA), a Type II transmembrane Serine Protease (TTSP) (such as, for example, DESC1, DPP-4, FAP, Hepsin, Matriptase-2, MT-SP1/Matriptase, TMPRSS2, TMPRSS3, TMPRSS4, TMPRSS5, TMPRSS6, TMPRSS7, TMPRSS8, TMPRSS9, TMPRSS10, TMPRSS11, and the like), and the like. Specific substrates are described, for example, in WO 2010/081173, WO 2015/048329, WO 2015/116933, and WO 2016/118629, each of which is incorporated herein by reference in its entirety.

[0088] In some embodiments, at least a portion of the CM1 in a substrate overlaps with at least a portion of the CM2 in the substrate, such that one or more amino acids in the substrate belongs to both CM1 and CM2. For example, a substrate with the sequence $X_1X_2X_3X_4X_5X_6$ (each X is an amino acid), may comprise overlapping CM1 and CM2, in which CM1 is $X_1X_2X_3X_4$ and CM2 is $X_3X_4X_5X_6$.

[0089] In some embodiments, the CM1 and CM2 in a substrate do not overlap in amino acid sequence such that no amino acid in the substrate belongs to both CM1 and CM2. For example, a substrate with the sequence $X_1X_2X_3X_4X_5X_6X_7X_8$ (each X is an amino acid) may comprise non-overlapping CM1 and CM2, in which CM1 is $X_1X_2X_3X_4$ and CM2 is $X_5X_6X_7X_8$. In some embodiments, the non-overlapping CM1 and CM2 are coupled directly. In some embodiments, the non-overlapping CM1 and CM2 are coupled indirectly (e.g., via a linking peptide).

[0090] In some embodiments, the CM1 and CM2 in a substrate have a structural arrangement from N-terminus to C-terminus as CM1-CM2. In some embodiments, the CM1 and the CM2 in a substrate have a structural arrangement from N-terminus to C-terminus as CM2-CM1. As used herein, the CM1 and CM2 in the formula CM1-CM2 or CM2-CM1 may be overlapping CM1 and CM2, non-overlapping CM1 and CM2 coupled directly, or non-overlapping CM1 and CM2 coupled indirectly (e.g., via a linking peptide).

ACTIVATABLE MOLECULES

[0091] In some embodiments, the polypeptide or polypeptide complex comprising a substrate is an activatable molecule. The activatable molecule may comprise an active moiety (AM) that specifically binds a target. The AM may be coupled to the substrate. In some embodiments, the activatable molecule comprises a masking moiety (MM) coupled with the AM via the substrate.

[0092] The coupling of two components in a polypeptide or polypeptide complex (e.g., an activatable molecule) may be direct or indirect. When the two components are coupled directly, the amino acid residue at the C-terminus of a component forms a peptide bond with the amino acid residue at the N-terminus of the other component. When the two components are coupled indirectly, there is a stretch of amino acids between the two components. In some examples, the two components of a polypeptide may be indirectly coupled via one or more other components in the polypeptide, i.e., the one or more other components are between the two coupled components. For indirectly coupling or linking via another component, the one or more other components may be a linker, AM(s), CM(s), MM(s), or any combination thereof.

[0093] As used herein, the term "activatable molecule" refers to a molecule that comprises at least one set of MM, substrate, and AM and which exhibits attenuated binding to a target as compared to the binding of a counterpart "activated" molecule comprising the same AM to the same target. The terms "activated molecule," and "cleaved activatable molecule," are used interchangeably herein to refer to the AM-containing cleavage product that is generated after exposure of the activatable molecule to a substrate-specific protease (i.e., after cleavage of the substrate by at least one protease). In some embodiments, a cleaved activatable molecule may lack a MM due to cleavage of the substrate (e.g., by a protease), resulting in release of the MM.

[0094] An AM may be any polypeptide that specifically binds a target. In some examples, the AM may be a therapeutic macromolecule. In some examples, the AM may be an antibody

or an antigen-binding fragment. In some examples, the AM may be an antineoplastic macromolecule. In some examples, the AM may be a cytokine. In some examples, the AM may be a chimeric antigen receptor.

[0095] In some examples, the AM may be a diagnostic macromolecule. For example, the diagnostic macromolecule may be a diagnostic polypeptide having 3 to 30, 5 to 25, 7 to 20, or 9 to 15 amino acids in length. Such diagnostic polypeptide may be used, in non-limiting aspects, e.g., for testing cleavage in tissues, and/or assessment of the tissue microenvironment.

[0096] As used herein, the terms "specific binding" and "specifically binds" refer to the non-covalent interactions of the type that occur between an AM and its target, e.g., an immunoglobulin molecule and an antigen or a cytokine and its receptor, for which the AM is specific. The strength or affinity of binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater affinity. Unless indicated otherwise, as used herein, "affinity" refers to intrinsic binding affinity, which reflects a 1:1 interaction between members of an AM and its target. Affinity can be measured by common methods known in the art, including those described herein. Affinity can be determined, for example, using surface plasmon resonance (SPR) technology (e.g., BIACORE®) or biolayer interferometry (e.g., FORTEBIO®). Additional methods for determining the affinity for an AM and its target are known in the art. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and geometric parameters that equally influence the rate in both directions. Thus, both the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. (*See* Nature 361:186-87 (1993)). The ratio of K_{off} /K_{on} enables the cancellation of all parameters not related to affinity and is equal to the dissociation constant K_d. (*See, generally*, Davies et al. (1990) Annual Rev Biochem 59:439-473). As used herein, a statement that an AM "specifically binds" to its target refers to an AM that binds its target with a dissociation constant (K_d) of less than 100 μM (e.g., less than 5 μM or 10 μM). In some examples, the AM specifically binds its target with a K_d of about 0.01 nM to about 500 nM. In some examples, an AM is said to specifically bind the target, when the equilibrium binding constant (K_d) is ≤1 μM, in some embodiments ≤ 100 nM, in some embodiments ≤ 10

nM, and in some embodiments ≤ 100 pM to about 1 pM, as measured by assays such as radioligand binding assays or similar assays known to those skilled in the art.

[0097] In general, an activatable molecule may be designed by selecting an AM of interest and constructing the remainder of the activatable molecule so that, when conformationally constrained, the MM provides for masking of the AM or reduction of binding of the AM to its target. Structural design criteria can be to be taken into account to provide for this functional feature.

[0098] Activatable molecules may be provided in a variety of structural configurations. Exemplary formulas for activatable molecules are provided below. It is contemplated that the N- to C-terminal order of the AM, MM and substrate may be reversed within an activatable molecule. For example, activatable molecules can be represented by the following formulas (in order from an amino (N) terminal region to carboxyl (C) terminal region):

MM-Substrate-AM,

AM- Substrate -MM

MM-CM1-CM2-AM,

MM-CM2-CM1-AM,

AM-CM1-CM2-MM, or

AM-CM2-CM1-MM

As used herein and unless otherwise stated, each dash (-) between the components of the activatable molecule represents either a direct linkage or indirect linkage via one or more linkers, or overlapping between the components. It should be noted that although MM and substrate are indicated as distinct components in the formulas above, in all exemplary embodiments (including formulae) disclosed herein it is contemplated that the amino acid sequences of the MM and the substrate may overlap, *e.g.*, such that the substrate is completely or partially contained within the MM. In addition, the formulas above provide for additional amino acid sequences that may be positioned N-terminal or C-terminal to the activatable molecules components. Examples include targeting moieties (e.g., a ligand for a receptor of a cell present in a target tissue) and half-life extending moieties.

[0099] In some embodiments, MM, substrate, and/or AM are coupled indirectly via one or more linkers (e.g., a linking peptide (LP)). For example, an activatable molecule may comprise one of the following formulae (in order from an amino (N) terminal region to carboxyl (C) terminal region) :

MM-LP-Substrate-AM

MM-Substrate-LP-AM

MM-LP1-Substrate-LP2-AM

MM-LP2-Substrate-LP1-AM

AM-LP-Substrate -MM

AM-Substrate-LP-MM

AM-LP2-Substrate-LP1-MM

AM-LP1-Substrate-LP2-MM

MM-LP-CM1-CM2-AM

MM-LP-CM2-CM1-AM

MM-CM1-CM2-LP-AM

MM-CM2-CM1-LP-AM

MM-LP1-CM1-CM2-LP2-AM

MM-LP1-CM2-CM1-LP2-AM

MM-LP2-CM1-CM2-LP1-AM

MM-LP2-CM2-CM1-LP1-AM

AM-LP-CM1-CM2-MM

AM-LP-CM2-CM1-MM

AM-CM1-CM2-LP-MM

AM-CM2-CM1-LP-MM

AM-LP2-CM1-CM2-LP1-MM

AM-LP2-CM2-CM1-LP1-MM

AM-LP1-CM1-CM2-LP2-MM

AM-LP1-CM2-CM1-LP2-MM

wherein LP1 and LP2 are two linking peptides. In some examples, the LP1 and LP2 are identical to each other. In some examples, the LP1 and LP2 are not identical to each other. As used herein and unless otherwise stated, each dash (-) between the components of the activatable molecule represents either a direct linkage or indirect linkage via one or more linkers, or overlapping between the components.

[0100] In some embodiments, the substrate comprise more than two CMs. For example, the substrate may comprise a CM1 cleavable by a first protease, a CM2 cleavable by a second protease, and a CM3 cleavable by a third protease. Examples of the additional CM(s), e.g., CM3, in the substrate that are not the CM comprising the sequence of any of SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650 include those described in WO 2010/081173,

WO2021207669, WO2021207657, WO2021142029, WO2021061867, WO2020252349, WO2020252358, WO2020236679, WO2020176672, WO2020118109, WO2020092881, WO2020086665, WO2019213444, WO2019183218, WO2019173771, WO2019165143, WO2019075405, WO2019046652, WO2019018828, WO2019014586, WO2018222949, WO2018165619, WO2018085555, WO2017011580, WO2016179335, WO2016179285, WO2016179257, WO2016149201, WO2016014974, which are incorporated herein by reference in their entireties for all purposes. In some examples, one or more of the additional CMs may be cleavable by legumain. In some examples, the CM cleavable by legumain may comprise a sequence of any of SEQ ID NO: 10-73, 83-216, 460-634, and 636-650 and an Asparagine (Asn) residue at the N-terminus or C-terminus. In some examples, the CM cleavable by legumain may comprise a sequence of any of SEQ ID NO: 10-73, 83-216, and 460-634 and an Asparagine (Asn) residue at the N-terminus or C-terminus.

[0101] In some embodiments, the activatable molecule comprises a structural arrangement from N-terminus to C-terminus as follows: MM-CM1-CM2-AM, MM-CM2-CM1-AM, AM-CM1-CM2-MM, or AM-CM2-CM1-MM, MM-CM2-CM1-CM3-AM, MM-CM1-CM2-CM3-AM, MM-CM1-CM3-CM2-AM, MM-CM3-CM1-CM2-AM, or MM-CM3-CM2-CM1-AM. Likewise, a CM4 may be inserted any position between the MM and AM.

[0102] In some embodiments, the activatable molecule comprises a linking peptide (LP) and wherein the activatable molecule has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM1-CM2-AM, MM-CM1-CM2-LP-AM, MM-LP-CM2-CM1-AM, MM-CM2-CM1-LP-AM, MM-LP-CM2-CM1-CM3-AM, MM-LP-CM1-CM2-CM3-AM, MM-LP-CM1-CM3-CM2-AM, MM-LP-CM3-CM1-CM2-AM, MM-LP-CM3-CM2-CM1-AM, MM-CM2-CM1-CM3-LP-AM, MM-CM1-CM2-CM3-LP-AM, MM-CM1-CM3-CM2-LP-AM, MM-CM3-CM1-CM2-LP-AM, or MM-CM3-CM2-CM1-LP-AM. Likewise, a CM4 may be inserted any position between the MM and AM.

[0103] In some embodiments, the activatable molecule comprises a first linking peptide (LP1) and a second linking peptide (LP2), and wherein the activatable molecule has a structural arrangement from N-terminus to C-terminus as follows: MM- LP1-CM1-CM2-LP2-AM, MM-LP1-CM2-CM1-LP2-AM, AM-LP2-CM1-CM2-LP1-MM, or AM-LP2-CM2-CM1-LP1-MM, MM-LP1-CM2-CM1-CM3-LP2-AM, MM-LP1-CM1-CM2-CM3-LP2-AM, MM-LP1-CM1-CM3-CM2-LP2-AM, MM-LP1-CM3-CM1-CM2-LP2-AM, MM-LP1-CM3-CM2-CM1-LP2-AM, MM-LP2-CM2-CM1-CM3-LP1-AM, MM-LP2-CM1-

CM2-CM3-LP1-AM, MM-LP2-CM1-CM3-CM2-LP1-AM, MM-LP2-CM3-CM1-CM2-LP1-AM, or MM-LP2-CM3-CM2-CM1-LP1-AM. Likewise, a CM4 may be inserted any position between the MM and AM.

[0104] In some embodiments, the activatable molecule comprises an additional linking peptide (LP3) and wherein the activatable molecule has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM1-LP3-CM2-AM, MM-CM1-LP3-CM2-LP-AM, MM-LP-CM2-LP3-CM1-AM, MM-CM2-LP3-CM1-LP-AM, MM-LP-CM2-LP3-CM1-CM3-AM, MM-LP-CM1-LP3-CM2-CM3-AM, MM-LP-CM1-LP3-CM3-CM2-AM, MM-LP-CM3-LP3-CM1-CM2-AM, MM-LP-CM3-LP3-CM2-CM1-AM, MM-CM2-LP3-CM1-CM3-LP-AM, MM-CM1-LP3-CM2-CM3-LP-AM, MM-CM1-LP3-CM3-CM2-LP-AM, MM-CM3-LP3-CM1-CM2-LP-AM, MM-CM3-LP3-CM2-CM1-LP-AM, MM-LP-CM2-CM1-LP3-CM3-AM, MM-LP-CM1-CM2-LP3-CM3-AM, MM-LP-CM1-CM3-LP3-CM2-AM, MM-LP-CM3-CM1-LP3-CM2-AM, MM-LP-CM3-CM2-LP3-CM1-AM, MM-CM2-CM1-LP3-CM3-LP-AM, MM-CM1-CM2-LP3-CM3-LP-AM, MM-CM1-CM3-LP3-CM2-LP-AM, MM-CM3-CM1-LP3-CM2-LP-AM, MM-CM3-CM2-LP3-CM1-LP-AM,MM-LP1-CM1-LP3-CM2-LP2-AM, MM-LP1-CM2-LP3-CM1-LP2-AM, AM-LP1-CM1-LP3-CM2-LP2-MM, or AM-LP1-CM2-LP3-CM1-LP2-MM, MM-LP1-CM2-LP3-CM1-CM3-LP2-AM, MM-LP1-CM1-LP3-CM2-CM3-LP2-AM, MM-LP1-CM1-LP3-CM3-CM2-LP2-AM, MM-LP1-CM3-LP3-CM1-CM2-LP2-AM, MM-LP1-CM3-LP3-CM2-CM1-LP2-AM, MM-LP2-CM2-LP3-CM1-CM3-LP1-AM, MM-LP2-CM1-LP3-CM2-CM3-LP1-AM, MM-LP2-CM1-LP3-CM3-CM2-LP1-AM, MM-LP2-CM3-LP3-CM1-CM2-LP1-AM, or MM-LP2-CM3-LP3-CM2-CM1-LP1-AM. Likewise, a CM4 may be inserted any position between the MM and AM.

[0105] In some embodiments, the activatable molecule has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM2-LP3-CM1-LP4-CM3-AM, MM-LP-CM1-LP3-CM2-LP4-CM3-AM, MM-LP-CM1-LP3-CM3-LP4-CM2-AM, MM-LP-CM3-LP3-CM1-LP4-CM2-AM, MM-LP-CM3-LP3-CM2-LP4-CM1-AM, MM-CM2-LP3-CM1-LP4-CM3-LP-AM, MM-CM1-LP3-CM2-LP4-CM3-LP-AM, MM-CM1-LP3-CM3-LP4-CM2-LP-AM, MM-CM3-LP3-CM1-LP4-CM2-LP-AM, MM-CM3-LP3-CM2-LP4-CM1-LP-AM, MM-LP-CM2-LP4-CM1-LP3-CM3-AM, MM-LP-CM1-LP4-CM2-LP3-CM3-AM, MM-LP-CM1-LP4-CM3-LP3-CM2-AM, MM-LP-CM3-LP4-CM1-LP3-CM2-AM, MM-LP-CM3-LP4-CM2-LP3-CM1-AM, MM-CM2-LP4-CM1-LP3-CM3-LP-AM, MM-CM1-LP4-CM2-LP3-CM3-LP-AM, MM-CM1-LP4-CM3-LP3-CM2-LP-AM, MM-CM3-LP4-

CM1-LP3-CM2-LP-AM, MM-CM3-LP4-CM2-LP3-CM1-LP-AM, MM-LP1-CM2-LP3-CM1-LP4-CM3-LP2-AM, MM-LP1-CM1-LP3-CM2-LP4-CM3-LP2-AM, MM-LP1-CM1-LP3-CM3-LP4-CM2-LP2-AM, MM-LP1-CM3-LP3-CM1-LP4-CM2-LP2-AM, MM-LP1-CM3-LP3-CM2-LP4-CM1-LP2-AM, MM-LP2-CM2-LP3-CM1-LP4-CM3-LP1-AM, MM-LP2-CM1-LP3-CM2-LP4-CM3-LP1-AM, MM-LP2-CM1-LP3-CM3-LP4-CM2-LP1-AM, MM-LP2-CM3-LP3-CM1-LP4-CM2-LP1-AM, or MM-LP2-CM3-LP3-CM2-LP4-CM1-LP1-AM. Likewise, a CM4 may be inserted any position between the MM and AM.

[0106] In some embodiments, two CMs, e.g., CM1 and CM2, in a substrate have a structural arrangement from N-terminus to C-terminus as CM1-CM2. In some embodiments, two CMs, e.g., CM1 and the CM2 in a substrate have a structural arrangement from N-terminus to C-terminus as CM2-CM1. As used herein, the CM1 and CM2 in the formula CM1-CM2 or CM2-CM1 may be overlapping CM1 and CM2, non-overlapping CM1 and CM2 coupled directly, or non-overlapping CM1 and CM2 coupled indirectly (e.g., via a linking peptide).

[0107] In some embodiments, two CMs, e.g., CM2 or CM4 and CM3 or CM4, in a substrate have a structural arrangement from N-terminus to C-terminus as CM2-CM3, CM2-CM4 or CM3-CM4. In some embodiments, two CMs, e.g., CM2 and the CM3 in a substrate have a structural arrangement from N-terminus to C-terminus as CM3-CM2 or CM4-CM2 or CM4-CM3. As used herein, the CM2 and CM3 in the formula CM2-CM3 or CM3-CM2 may be overlapping CM2 and CM3, non-overlapping CM2 and CM3 coupled directly, or non-overlapping CM2 and CM3 coupled indirectly (e.g., via a linking peptide). As used herein, the CM2 and CM4 in the formula CM2-CM4 or CM4-CM2 may be overlapping CM2 and CM4, non-overlapping CM2 and CM4 coupled directly, or non-overlapping CM2 and CM4 coupled indirectly (e.g., via a linking peptide). As used herein, the CM4 and CM3 in the formula CM4-CM3 or CM3-CM4 may be overlapping CM4 and CM3, non-overlapping CM4 and CM3 coupled directly, or non-overlapping CM4 and CM3 coupled indirectly (e.g., via a linking peptide).

Antibodies and Antigen-Binding Fragments

[0108] In some embodiments, the AM is an antibody or antigen-binding fragment thereof. The term "antibody" is used herein in its broadest sense and includes certain types of immunoglobulin molecules that include one or more target-binding domains that specifically bind an antigen or epitope. Examples of antibodies include intact antibodies (e.g., intact immunoglobulins), antibody fragments, bispecific, and multi-specific antibodies. One

example of a target-binding domain is formed by a V_H -V_L dimer. Additional examples of an antibody are described herein. Additional examples of an antibody are known in the art.

[0109] A "light chain" includes one variable domain (VL) and one constant domain (CL). There are two different light chains termed kappa or lambda. A "heavy chain" consists of one variable domain (VH) and three constant region domains (CH1, CH2, CH3). There are five main heavy-chain classes or isotypes, some of which have several subtypes, and these determine the functional activity of an antibody molecule. The five major classes of immunoglobulin are immunoglobulin M (IgM), immunoglobulin D (IgD), immunoglobulin G (IgG), immunoglobulin A (IgA), and immunoglobulin E (IgE). IgG is by far the most abundant immunoglobulin and has several subclasses (IgG1, IgG2, IgG3, and IgG4 in humans).

[0110] In some embodiments, the antigen-binding fragment is a Fab fragment, a $F(ab')_2$ fragment, a scFv, a scAb, a dAb, a single domain heavy chain antibody, or a single domain light chain antibody. Additional examples of the antigen-binding fragments include a VH domain, a VHH domain, a VNAR domain, and a single chain fragment variable (scFv), BiTE or a component thereof, a $(scFv)_2$, a NANOBODY®, a nanobody-HSA, VHH-scAb, a VHH-Fab, a Dual scFab, a $F(ab')_2$, a diabody, a CROSSMAB®, a DAF (two-in-one), a DAE (four-in-one), a DUTAMAB®, a DT- IgG, a knobs-in-holes common light chain, a knobs-in-holes assembly, a charge pair, a Fab-arm exchange, a SEEDbody, a LUZ-Y, a FcAb, a kl-body, an orthogonal Fab, a DVD-IgG, a IgG(H)-scFv, a scFv-(H)IgG, IgG(L)-scFv, scFv-(L)IgG, IgG(L,H)-Fv, IgG(H)-V, V(H)-IgG, IgG(L)-V, V(L)-IgG, KIH IgG-scFab, 2scFv-IgG, IgG-2scFv, scFv4-Ig, ZYBODY™, DVI-IgG, Diabody-CH3, a triple body, a miniantibody, a minibody, a TriBi minibody, scFv-CH3 KIH , Fab-scFv, a F(ab')2-scFv2, a scFv-KIH, a Fab-scFv-Fc, a tetravaient HCAb, a scDiabody-Fc, a Diabody-Fc, a tandem scFv-Fc, a VHH-Fc, a tandem VHH-Fc, a L'HH-Fc KiH, a Fab- VHH-Fc, an Intrabody, a dock and lock, an ImmTAC® (immune-mobilizing monoclonal TCRs (T cell receptors) against cancer), an IgG-IgG conjugate, a Cov-X-Body, a scFvl- PEG-scFv2, an Adnectin, a DARPin®, a fibronectin, an IgG, an IgM, an IgA, an IgE, an IgD, a DEP conjugate, TMEAbody™, SAFEbody®, TRITAC® , or SHIELD antibody.

[0111] A "fragment antigen binding" (Fab) includes a complete light chain paired with the VH domain and the CH1 domain of a heavy chain. A $F(ab')_2$ fragment is formed when an antibody is cleaved by pepsin (or otherwise truncated) below the hinge region, in which case the two fragment target-binding domains (Fabs) of the antibody molecule remain linked. A

$F(ab')_2$ fragment contains two complete light chains paired with the two VH and CH1 domains of the heavy chains joined together by the hinge region. A "fragment crystallizable" (Fc) fragment (also referred to herein as F_C domain) corresponds to the paired CH2 and CH3 domains and is the part of the antibody molecule that interacts with effector molecules and cells. The functional differences between heavy-chain isotypes lie mainly in the Fc fragment. A "single chain fragment variable" (scFv) contains only the variable domain of a light chain (VL) linked by a stretch of peptide to a variable domain of a heavy chain (VH). The name single-chain Fv is derived from Fragment variable. A "hinge region" or "interdomain" is flexible amino acid stretch that joins or links the Fab fragment to the Fc domain. A "synthetic hinge region" is an amino acid sequence that joins or links a Fab fragment to an Fc domain.

[0112] An "Fv" fragment includes a non-covalently-linked dimer of one heavy chain variable domain and one light chain variable domain. A "dual variable domain immunoglobulin G" or "DVD-IgG" refers to multivalent and multispecific target-binding proteins as described, e.g., in DiGiammarino et al., Methods Mol. Biol. 899:145-156, 2012, Jakob et al., MABs 5:358-363, 2013; and U.S, Patent Nos. 7,612,181; 8,258,268; 8,586,714; 8,716,450; 8,722,855; 8,735,546; and 8,822,645, each of which is incorporated by reference in its entirety. Examples of DARTs are described in, e.g., Garber, Nature Reviews Drug Discovery 13:799-801, 2014.

[0113] A VHH domain is a single monomeric variable antibody domain that can be found in camelids. A VNAR domain is a single monomeric variable antibody domain that can be found in cartilaginous fish. Non-limiting aspects of VHH domains and VNAR domains are described in, e.g., Cromie et al., Curr. Top. Med. Chem. 15:2543-2557, 2016; De Genst et al. Dev. Comp. Immunol. 30:187-198, 2006; De Meyer et al, Trends Biotechnol 32:263-270, 2014; Kijanka et al., Nanomedicine 10:161-174, 2015; Kovaleva et al., Expert. Opin. Biol. Ther. 14: 1527-1539, 2014; Krah et al., Immunopharmacol. Immunotoxicol. 38:21-28, 2016; Mujic-Delic et al., Trends Pharmacol. Sci. 35:247-255, 2014; Muyldermans, J. Biotechnol. 74:277-302, 2001, Muyldermans et al., Trends Biocheni. Sci. 26:230-235, 2001; Muyldermans, Ann. Rev. Biochem. 82:775-797, 2013; Rahbarizadeh et al., Immunol, invest. 40:299-338, 2011; Van Audenhove et al., EBioMedicine 8:40-48, 2016; Van Bockstaele et al., Curr. Opin. Investig. Drugs 10:1212-1224, 2009; Vincke et al. Methods Mol, Biol, 911:15-26, 2012; and Wesolowski et al. Med. Microbiol. Immunol. 198:157-174, 2009, each of which is incorporated by reference herein in its entirety.

[0114] In some embodiments, the AM may be a mouse, rat, rabbit, goat, camel, donkey, primate, human, or humanized or chimeric polypeptide. In one example, the AM may be a human polypeptide. In one example, the AM may be a humanized (e.g., fully humanized) polypeptide.

[0115] The term "humanized" refer to an AM having an amino acid sequence that includes VH and VL region sequences from a reference protein raised in a non-human species (e.g., a mouse), but also includes modifications in those sequences relative to the reference protein intended to render them more "human-like," i.e., more similar to human germline variable sequences. In some embodiments, a "humanized" AM is one that immunospecifically binds an antigen of interest and that has a framework (FR) region having substantially the amino acid sequence as that of a human protein, and a complementary determining region (CDR) having substantially the amino acid sequence as that of a non-human protein contains humanized VH and VL regions.

[0116] The term "human polypeptide" is intended to include AMs having variable and constant regions generated, assembled, or derived from human immunoglobulin sequences. In some embodiments, an AM may be considered to be "human" even though its amino acid sequence include residues or elements not encoded by human germline immunoglobulin sequences (e.g., include sequence variations, for example that may (originally) have been introduced by random or site-specific mutagenesis in vitro or by somatic mutation in vivo), e.g., in one or more CDRs.

[0117] Examples of antibodies and antigen-binding fragments include those binding to cell surface receptors and secreted binding proteins (*e.g.*, growth factors), soluble enzymes, structural proteins (*e.g.* collagen, fibronectin) and the like, or an extracellular target (e.g., an extracellular protein target). In some embodiments, antibodies and antigen-binding fragments are designed for cellular uptake and are activatable inside a cell.

[0118] Examples of antibodies and antigen-binding fragments include those in Example 1, e.g., those comprising a light chain comprising a sequence selected from SEQ ID NOs: 217, 219, 221, 222, and 224-235, and a heavy chain comprising a sequence selected from SEQ ID NOs: 218, 220, and 223. In some examples, the polypeptide (e.g., isolated polypeptide) herein comprises a sequence selected from SEQ ID NOs: 221, 222, and 224-235.

Multispecific Activatable Antibodies

[0119] In some embodiments, the activatable antibodies are multispecific activatable antibodies. In some examples, the multispecific activatable antibodies herein recognize two or more different antigens or epitopes and that include at least one masking moiety (MM) linked to at least one antigen- or epitope-binding domain of the multispecific antibody such that coupling of the MM reduces the ability of the antigen- or epitope-binding domain to bind its target. In some embodiments, the MM is coupled to the antigen- or epitope-binding domain of the multispecific antibody via a substrate cleavable by one or multiple proteases, e.g., MT-SP1 and/or an MMP. The activatable multispecific antibodies provided herein are stable in circulation, activated at intended sites of therapy and/or diagnosis but not in normal, i.e., healthy tissue, and, when activated, exhibit binding to a target that is at least comparable to the corresponding, unmodified multispecific antibody.

[0120] The multispecific activatable molecules may be used to target a first and a second target tissues. In one embodiment, the first and second target tissues are spatially separated, for example, at different sites in the organism. In one embodiment, the first and second target tissues are the same tissue temporally separated, for example the same tissue at two different points in time, for example the first time point is when the tissue is an early stage tumor, and the second time point is when the tissue is a late stage tumor.

[0121] In some embodiments, the multispecific activatable antibody includes a first antibody or antigen-binding fragment thereof (AB1) that binds a first target, where the AB1 is coupled to a masking moiety (MM1) such that coupling of the MM1 reduces the ability of the AB1 to bind the first target, and the multispecific activatable antibody includes a second antibody or antigen-binding fragment thereof (AB2) that binds a second target, where the AB2 is coupled to a masking moiety (MM2) such that coupling of the MM2 reduces the ability of the AB2 to bind the second target. In some embodiments, AB1 is coupled to MM1 via Substrate 1, and AB2 is coupled to MM2 via Substrate 2. In some embodiments, there are linking peptides between AB1 and Substrate 1, between Substrate 1 and MM1, between AB2 and Substrate 2, and/or between Substrate 2 and MM2. In some embodiments, AB1 is directly coupled to Substrate 1, Substrate 1 is directly coupled to MM1, AB2 is directly coupled to Substrate 2, and/or Substrate 2 is directly coupled to MM2.

[0122] For example, the multispecific activatable antibodies can be represented by the following formulas (in order from an amino (N) terminal region to carboxyl (C) terminal region) :

MM1-Substrate 1-AB1 : MM2-Substrate 2-AB2
AB1-Substrate 1-MM1 : MM2-Substrate 2-AB2
AB1-Substrate 1-MM1 : AB2-Substrate 2-MM2
MM1-Substrate 1-AB1 : AB2-Substrate 2-MM2

wherein ":" separates two polypeptides, which may be two independent polypeptides on two different molecules, or two polypeptides on the same molecule (e.g., two polypeptide chains of the same protein). As used herein and unless otherwise stated, each dash (-) between the components of the activatable molecule represents either a direct linkage or indirect linkage via one or more linking peptides, or overlapping between the components.

[0123] In some embodiments, the multispecific activatable antibodies are designed to engage immune effector cells, also referred to herein as immune-effector cell engaging multispecific activatable antibodies. In some embodiments, the multispecific activatable antibodies are designed to engage leukocytes, also referred to herein as leukocyte engaging multispecific activatable antibodies. In some embodiments, the multispecific activatable antibodies are designed to engage T cells, also referred to herein as T-cell engaging multispecific activatable antibodies. In some embodiments, the multispecific activatable antibodies engage a surface antigen on a leukocyte, such as on a T cell, on a natural killer (NK) cell, on a myeloid mononuclear cell, on a macrophage, and/or on another immune effector cell. In some embodiments, the immune effector cell is a leukocyte. In some embodiments, the immune effector cell is a T cell. In some embodiments, the immune effector cell is a NK cell. In some embodiments, the immune effector cell is a mononuclear cell, such as a myeloid mononuclear cell. In some embodiments, the multispecific activatable antibodies are designed to bind or otherwise interact with more than one target and/or more than one epitope, also referred to herein as multi-antigen targeting activatable antibodies. As used herein, the terms "target" and "antigen" are used interchangeably.

[0124] In some embodiments, immune effector cell engaging multispecific activatable antibodies of the disclosure include a targeting antibody or antigen-binding fragment thereof and an immune effector cell engaging antibody or antigen-binding portion thereof, where at least one of the targeting antibody or antigen-binding fragment thereof and/or the immune effector cell engaging antibody or antigen-binding portion thereof is masked.

[0125] In some embodiments, the non-immune effector cell engaging antibody is a cancer targeting antibody. In some embodiments the non-immune cell effector antibody is an IgG.

In some embodiments the immune effector cell engaging antibody is a scFv. In some embodiments the targeting antibody (e.g., non-immune cell effector antibody) is an IgG and the immune effector cell engaging antibody is a scFv. In some embodiments, the immune effector cell is a leukocyte. In some embodiments, the immune effector cell is a T cell. In some embodiments, the immune effector cell is a NK cell. In some embodiments, the immune effector cell is a myeloid mononuclear cell.

[0126] In some embodiments of an immune effector cell engaging multispecific activatable antibody, one antigen is typically an antigen present on the surface of a tumor cell or other cell type associated with disease, and another antigen is typically a stimulatory or inhibitory receptor present on the surface of a T-cell, natural killer (NK) cell, myeloid mononuclear cell, macrophage, and/or other immune effector cell.

[0127] One embodiment of the disclosure is a multispecific activatable antibody that is activatable in a cancer microenvironment and that includes an antibody, for example an IgG or scFv, directed to a tumor target and an agonist antibody, for example an IgG or scFv, directed to a co-stimulatory receptor expressed on the surface of an activated T cell or NK cell, wherein at least one of the cancer target antibody and/or agonist antibody is masked. In this embodiment, the multispecific activatable antibody, once activated by tumor-associated proteases, effectively crosslinks and activates the T cell or NK cell expressed co-stimulatory receptors in a tumor-dependent manner to enhance the activity of T cells that are responding to any tumor antigen via their endogenous T cell antigen or NK-activating receptors. The activation-dependent nature of these T cell or NK cell costimulatory receptors focuses the activity of the activated multispecific activatable antibody to tumor-specific T cells, without activating all T cells independent of their antigen specificity. In one embodiment, at least the co-stimulatory receptor antibody of the multispecific activatable antibody is masked to prevent activation of auto-reactive T cells that may be present in tissues that also express the antigen recognized by the tumor target-directed antibody in the multispecific activatable antibody, but whose activity is restricted by lack of co-receptor engagement.

[0128] One embodiment of the disclosure is a multispecific activatable antibody that is activatable in a disease characterized by T cell overstimulation, such as an autoimmune disease or inflammatory disease microenvironment. Such a multispecific activatable antibody includes an antibody, for example an IgG or scFv, directed to a target comprising a surface antigen expressed in a tissue targeted by a T cell in autoimmune or inflammatory disease and an antibody, for example an IgG or scFv, directed to an inhibitory receptor expressed on the

surface of a T cell or NK cell, wherein at least one of the disease tissue target antibody and/or T cell inhibitory receptor antibody is masked. Examples of a tissue antigen targeted by T cells in autoimmune disease include a surface antigen expressed on myelin or nerve cells in multiple sclerosis or a surface antigen expressed on pancreatic islet cells in Type 1 diabetes. In this embodiment, the multispecific activatable antibody when localized in the tissue under autoimmune attack or inflammation is activated and co-engages the T cell or NK cell inhibitory receptor to suppress the activity of autoreactive T cells responding to any disease tissue-targeted antigens via their endogenous TCR or activating receptors. In one embodiment, at least one or multiple antibodies are masked to prevent suppression of T cell responses in non-disease tissues where the target antigen may also be expressed.

[0129] In some embodiments, the multi-antigen targeting antibodies and/or multi-antigen targeting activatable antibodies include at least a first antibody or antigen-binding fragment thereof that binds a first target and/or first epitope and a second antibody or antigen-binding fragment thereof that binds a second target and/or a second epitope. In some embodiments, the multi-antigen targeting antibodies and/or multi-antigen targeting activatable antibodies bind two or more different targets. In some embodiments, the multi-antigen targeting antibodies and/or multi-antigen targeting activatable antibodies bind two or more different epitopes on the same target. In some embodiments, the multi-antigen targeting antibodies and/or multi-antigen targeting activatable antibodies bind a combination of two or more different targets and two or more different epitopes on the same target.

Masking Moieties (MMs)

[0130] The activatable molecules herein may comprise one or more masking moieties (MMs) capable of interfering with the binding of the AMs to the target. A masking moiety in an activatable molecule "masks" or reduces or otherwise inhibits the binding of the activatable molecule to its target. In some embodiments, the coupling of an AM (e.g., an antibody or fragment thereof, or other therapeutic or diagnostic protein) with an MM may inhibit the ability of the AM to specifically bind its target by means of inhibition known in the art (e.g., structural change, competition for antigen-binding domain, and the like). In some embodiments, the coupling of an AM with an MM may effect a structural change that reduces or inhibits the ability of the AM to specifically bind its target. In some embodiments, the coupling of a protein comprising an AM with an MM sterically blocks, reduces or inhibits the ability of the AM to specifically bind its target and or epitope. In some embodiments, when an activatable molecule is not activated, the MM prevents the AM from target binding;

but when the activatable molecule is activated (when the substrate is cleaved by one or more proteases), the MM does not substantially or significantly interfere with the AM's binding to the target.

[0131] An MM may be coupled to an AM (e.g., an antibody or fragment thereof, or other therapeutic or diagnostic protein) via the substrate described herein, either directly or indirectly (e.g., via one or more linkers described herein). Alternatively, an MM interfering with the target binding of an AM may be coupled, either directly or indirectly, to a component of the activatable molecule that is not the AM. For example, the MM may be coupled, either directly or indirectly, to a different AM. In another example, the MM may be coupled, either directly or indirectly, with a half-life extending moiety (EM). In either case, in the tertiary or quaternary structure of the activatable structure, the MM may be in a position (e.g., proximal to the AM to be masked) that allows the MM to mask the AM.

[0132] In some embodiments, an MM interacts with the AM, thus reducing or inhibiting the interaction between the AM and its binding partner. In some embodiments, the MM comprises at least a partial or complete amino acid sequence of a naturally occurring binding partner of the AM. The term "naturally occurring" as used herein as applied to an object refers to the fact that an object can be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including viruses or bacteria) that can be isolated from a source in nature and that has not been intentionally modified by man in the laboratory or otherwise is naturally occurring.

[0133] For example, the MM may be a fragment of a naturally occurring binding partner. The fragment may retain at least 95%, at least 90%, at least 80%, at least 75%, at least 70%, at least 60%, at least 50%, at least 40%, at least 30%, at least 25%, or at least 20% nucleic acid or amino acid sequence homology to the naturally occurring binding partner. In some embodiments, the MM is a cognate peptide of the AM. For example, the MM may comprise a sequence of the AM's epitope or a fragment thereof.

[0134] In some embodiments, the MM comprises an amino acid sequence that is not naturally occurring or does not contain the amino acid sequence of a naturally occurring binding partner or target protein. In certain embodiments, the MM is not a natural binding partner of the AM. In some embodiments, the MM does not comprise a subsequence of more than 4, 5, 6, 7, 8, 9 or 10 consecutive amino acid residues of a natural binding partner of the AM. The MM may be a modified binding partner for the AM which contains amino acid changes that decrease affinity and/or avidity of binding to the AM. In some embodiments, the

MM contains no or substantially no nucleic acid or amino acid homology to the AM's natural binding partner. In other embodiments the MM has no more than 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, or 80% homology to the natural binding partner of the AM.

[0135] In some embodiments, the MM is a polypeptide that binds to the AM. In some examples, the MM may be an antibody or antibody fragment (e.g., a Fab fragment, a $F(ab')_2$ fragment, a scFv, a scAb, a dAb, a single domain heavy chain antibody, and a single domain light chain antibody) that binds to the AM such that interrupts the AM's binding to its target. In some examples, the MM may be a ligand, a receptor, a fragment thereof (e.g., an extracellular domain of a receptor) of the AM that binds to the AM and interrupts the AM's binding to its target. In some examples, when the AM is an antibody or antibody fragment thereof, the MM may be an anti-idiotypic antibody or fragment thereof (e.g., scFv) that binds to the idiotype of the AM. In some examples, the MM may be a cytokine or a receptor for a cytokine. In some examples, the MM may have an amino acid sequence that is at least 85% identical to a cytokine or to a receptor for a cytokine.

[0136] In some embodiments, the MM does not bind the AM, but still interferes with AM's binding to its binding partner through non-specific interactions such as steric hindrance. For example, the MM may be positioned in the activatable molecule such that the tertiary or quaternary structure of the activatable molecule allows the MM to mask the AM through charge-based interaction, thereby holding the MM in place to interfere with binding partner access to the AM. Examples of such MMs include an albumin, e.g., human serum albumin (HSA), a fragment crystallizable (Fc) domain, an antibody constant domain (e.g., CH domains), a polymer (e.g., branched or multi-armed polyethylene glycol (PEG)), a latency associated protein (LAP), and any polypeptide or other moieties that sterically interfere AM-target interactions. In some examples, the MM may recruit a large protein binding partner that sterically interfere AM-target interactions. For example, the MM may be an antibody or a fragment thereof that binds to serum albumin.

[0137] Examples of suitable masking moieties include the full-length or a AM-binding fragment or mutein of a cognate receptor of the AM, and AM-binding antibodies and fragment thereof, e.g., a polyclonal antibody, a recombinant antibody, a human antibody, a humanized antibody a single chain variable fragment (scFv), single-domain antibody such as a heavy chain variable domain (VH), a light chain variable domain (VL), a variable domain of camelid-type nanobody (VHH), a dAb and the like. Other exemplary antigen-binding

domain that bind the AM can also be used as an MM include non-immunoglobulin proteins that mimic antibody binding and/or structure such as, anticalins, affilins, affibody molecules, affimers, affitins, alphabodies, avimers, DARPins, fynomers, kunitz domain peptides, monobodies, and binding domains based on other engineered scaffolds such as SpA, GroEL, fibronectin, lipocallin and CTLA4 scaffolds. As another example, a peptide that is modified by conjugation to a water-soluble polymer, such as PEG, can sterically inhibit or prevent binding of the cytokine to its receptor. Antibodies and antigen-binding domains that bind to, for example, a protein with a long serum half- life such as HSA, immunoglobulin or transferrin, or to a receptor that is recycled to the plasma membrane, such as FcRn or transferrin receptor, can also inhibit the cytokine, particularly when bound to their antigen. In some embodiments, the MMs (e.g., those sterically interfere with the AM-target interaction) can also function as half-life extending elements.

[0138] In some embodiments, the MM may have a dissociation constant for binding to the AM that is no more than the dissociation constant of the AM to the target. In some embodiments, the MM does not interfere or compete with the AM for binding to the target in in the activated molecule (i.e., following cleavage of the substrate by a protease).

[0139] The structural properties of the MMs may be selected according to factors such as the minimum amino acid sequence required for interference with the AM binding to target, the target protein-protein binding pair of interest, the size of the AM, the presence or absence of linkers, and the like.

[0140] In some embodiments, the MM may be unique for the coupled AM. Examples of MMs include MMs that were specifically screened to bind a binding domain of the AM or fragment thereof (e.g., affinity masks). Methods for screening MMs to obtain MMs unique for the AM and those that specifically and/or selectively bind a binding domain of a binding partner/target are provided herein and can include protein display methods.

[0141] As used herein, the term "masking efficiency" refers to the activity (e.g., EC_{50}) of the activatable molecule divided by the activity of a control molecule, wherein the control molecule may be either cleavage product of the activatable molecule (i.e., the activated molecule) or the AM used in the activatable molecule. An activatable molecule having a reduced level of an AM activity may have a masking efficiency that is greater than 10. In some embodiments, the activatable molecules described herein have a masking efficiency that is greater than 10, 100, 1000, or 5000.

[0142] In some embodiments, the MM is a polypeptide of about 2 to 50 amino acids in length. For example, the MM may be a polypeptide of from 2 to 40, from 2 to 30, from 2 to 20, from 2 to 10, from 5 to 15, from 10 to 20, from 15 to 25, from 20 to 30, from 25 to 35, from 30 to 40, from 35 to 45, from 40 to 50 amino acids in length. For example, the MM may be a polypeptide with 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length. In some examples, the MM may be a polypeptide of more than 50 amino acids in length, e.g., 100, 200, 300, 400, 500, 600, 700, 800, or more amino acids. In some embodiments, the MM is a steric mask.

[0143] In some embodiments, in an activatable molecule with an AM and an interfering MM, in the presence of the target of an AM, there is no binding or substantially no binding of the AM to the target, or no more than 0.001%, 0.01%, 0.1%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, or 50% binding of the AM to its target, as compared to the binding of an counterpart molecule without the interfering MM, for at least 0.1, 0.5, 1, 2, 4, 6, 8, 12, 28, 24, 30, 36, 48, 60, 72, 84, or 96 hours, or 5, 10, 15, 30, 45, 60, 90, 120, 150, or 180 days, or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months when measured in vitro immunoabsorbant assay, e.g., as described in US20200308243A1.

[0144] The binding affinity of the AM towards the target or binding partner with an interfering MM may be at least 5, 10, 25, 50, 100, 250, 500, 1,000, 2,500, 5,000, 10,000, 50,000, 100,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000 times lower than the binding affinity of the AM towards its binding partner without an interfering MM, or between 5-10, 10-100, 10-1,000, 10-10,000, 10-100,000, 10-1,000,000, 10-10,000,000, 100-1,000, 100-10,000, 100-100,000, 100-1,000,000, 100-10,000,000, 1,000-10,000, 1,000-100,000, 1,000-1,000,000, 1000-10,000,000, 10,000-100,000, 10,000-1,000,000, 10,000-10,000,000, 100,000-1,000,000, or 100,000-10,000,000 times lower than the binding affinity of the AM towards its binding partner when there is no interfering MM.

[0145] The dissociation constant (K_d) of the MM towards the AM it masks, may be greater than the dissociation constant of the AM towards the target. The dissociation constant of the MM towards the masked AM may be at least 5, 10, 25, 50, 100, 250, 500, 1,000, 2,500, 5,000, 10,000, 100,000, 1,000,000 or even 10,000,000 times greater than the dissociation constant of the AM towards the target. Conversely, the binding affinity of the MM towards the masked AM may be lower than the binding affinity of the AM towards the target. The binding affinity of MM towards the AM may be at least 5, 10, 25, 50, 100, 250, 500, 1,000,

2,500, 5,000, 10,000, 100,000, 1,000,000 or even 10,000,000 times lower than the binding affinity of the AM towards the target.

[0146] In some embodiments, the K_d of the activatable molecule comprising an MM and a substrate towards the AM's target is at least 5, 10, 25, 50, 100, 250, 500, 1,000, 2,500, 5,000, 10,000, 50,000, 100,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000 or greater, or between 5-10, 10-100, 10-1,000, 10-10,000, 10-100,000, 10-1,000,000, 10-10,000,000, 100-1,000, 100-10,000, 100-100,000, 100-1,000,000, 100-10,000,000, 1,000-10,000, 1,000-100,000, 1,000-1,000,000, 1000-10,000,000, 10,000-100,000, 10,000-1,000,000, 10,000-10,000,000, 100,000-1,000,000, or 100,000-10,000,000 times greater than the K_d of a counterpart molecule that is substantially the same as the activatable molecule but does not comprise the MM or substrate towards the AM's target. Conversely, the binding affinity of the activatable molecule comprising an MM and a substrate towards the AM's target is at least 5, 10, 25, 50, 100, 250, 500, 1,000, 2,500, 5,000, 10,000, 50,000, 100,000, 500,000, 1,000,000, 5,000,000, 10,000,000, 50,000,000 or greater, or between 5-10, 10-100, 10-1,000, 10-10,000, 10-100,000, 10-1,000,000, 10-10,000,000, 100-1,000, 100-10,000, 100-100,000, 100-1,000,000, 100-10,000,000, 1,000-10,000, 1,000-100,000, 1,000-1,000,000, 1000-10,000,000, 10,000-100,000, 10,000-1,000,000, 10,000-10,000,000, 100,000-1,000,000, or 100,000-10,000,000 times lower than the binding affinity of a counterpart molecule that is substantially the same as the activatable molecule but does not comprise the MM or substrate towards the AM's target.

[0147] In some embodiments, when the AM is coupled with an MM and is in the presence of the target, the specific binding of the AM to its target is reduced or inhibited, as compared to the specific binding of the AM not coupled with the MM. When compared to the binding of the AM not coupled with the MM to the target, the target-binding ability of the AM coupled with the MM may be reduced by at least 50%, 60%, 70%, 80%, 90%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% for at least 2, 4, 6, 8, 12, 28, 24, 30, 36, 48, 60, 72, 84, or 96 hours, or 5, 10, 15, 30, 45, 60, 90, 120, 150, or 180 days, or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months or more when measured *in vivo* or in an *in vitro* assay.

[0148] In some embodiments, the MM comprises a non-binding steric moiety (NB) that does not bind the AM but is able to interfere the binding between the AM and its target via steric hindrance. In some embodiments, the MM comprises a binding partner (BP) for a NB, where the BP recruits or otherwise attracts the NB to the activatable molecule.

[0149] In some embodiments, the MM contains genetically encoded or genetically non-encoded amino acid(s). Examples of genetically non-encoded amino acids include D-amino acids, β-amino acids, and γ-amino acids. In specific embodiments, the MMs contain no more than 50%, 40%, 30%, 20%, 15%, 10%, 5% or 1% of genetically non-encoded amino acids.

[0150] In some embodiments, once released from the activatable molecule and in a free state, the MM may have a biological activity or a therapeutic effect, such as binding capability. For example, the free peptide may bind with the same or a different binding partner. In certain embodiments, the free MM may exert a therapeutic effect, providing a secondary function to the compositions disclosed herein. In some embodiments, once uncoupled from the activatable molecule and in a free state, the MM may advantageously not exhibit biological activity. For example, in some embodiments the MM after cleavage from the activatable molecule does not elicit an immune response in the subject.

[0151] Suitable MMs may be identified and/or further optimized through a screening procedure from a library of candidate activatable molecule having variable MMs. For example, an AM and a substrate may be selected to provide for a desired enzyme/target combination, and the amino acid sequence of the MM can be identified by the screening procedure described below to identify an MM that provides for an activatable phenotype. For example, a random peptide library (e.g., of peptides comprising 2 to 40 amino acids or more) may be used in the screening methods disclosed herein to identify a suitable MM.

[0152] In some embodiments, MMs with specific binding affinity for an AM may be identified through a screening procedure that includes providing a library of peptide scaffolds comprising candidate MMs wherein each scaffold is made up of a transmembrane protein and the candidate MM. The library may then be contacted with an entire or portion of a protein such as a full length protein, a naturally occurring protein fragment, or a non-naturally occurring fragment containing a protein (also capable of binding the binding partner of interest), and identifying one or more candidate MMs having detectably bound protein. The screening may be performed by one more rounds of magnetic-activated sorting (MACS) or fluorescence-activated sorting (FACS), as well as determination of the binding affinity of MM towards the AM and subsequent determination of the masking efficiency, e.g., as described in WO2009025846 and US20200308243A1, which are incorporated herein by reference in their entireties.

[0153] Examples of suitable MMs are disclosed in WO2021207657, WO2021142029, WO2021061867, WO2020252349, WO2020252358, WO2020236679, WO2020176672,

WO2020118109, WO2020092881, WO2020086665, WO2019213444, WO2019183218, WO2019173771, WO2019165143, WO2019075405, WO2019046652, WO2019018828, WO2019014586, WO2018222949, WO2018165619, WO2018085555, WO2017011580, WO2016179335, WO2016179285, WO2016179257, WO2016149201, and WO2016014974, which are incorporated herein by reference in their entireties.

[0154] In some embodiments, the AM in an activatable molecule is an antibody or antigen-binding fragment that specifically binds EGFR. In some examples, such an activatable molecule comprises an MM that comprises the amino acid sequence of SEQ ID NO: 81. In some examples, such an activatable molecule comprises an MM that comprises the amino acid sequence of SEQ ID NO: 82. In some examples, such an activatable molecule comprises an MM that consists of the amino acid sequence of SEQ ID NO: 81. In some examples, such an activatable molecule comprises an MM that consists of the amino acid sequence of SEQ ID NO: 82.

[0155] In some aspects, the present disclosure includes an activatable antibody comprising an anti-EGFR antibody coupled directly or indirectly to a substrate, wherein the substrate is directly or indirectly coupled to an MM that comprises or consists of the amino acid sequence of SEQ ID NO: 82.

Linkers

[0156] The activatable molecules may comprise one or more linkers. The linkers may be linking peptides that comprise a stretch of amino acid sequence that link two components in the activatable molecule. The linkers may be non-cleavable by any protease. In some embodiments, one or more linkers may be introduced into the activatable molecule to provide flexibility at one or more of the junctions between domains, between moieties, between moieties and domains, or at any other junctions where a linker would be beneficial. In some embodiments, where the activatable molecule is provided as a conformationally constrained construct, a flexible linker may be inserted to facilitate formation and maintenance of a structure in the activatable molecule. Any of the linkers described herein may provide the desired flexibility to facilitate the inhibition of the binding of a target, or to facilitate cleavage of a substrate by a protease. In some embodiments, linkers included in the activatable molecule may be all or partially flexible, such that the linker can include a flexible linker as well as one or more portions that confer less flexible structure to provide for a desired activatable molecule. Some linkers may include cysteine residues, which may form disulfide bonds and reduce flexibility of the construct.

[0157] In some embodiments, a linker coupled to an MM may have a length that allows the MM to be in a position in the tertiary or quaternary to effectively mask an AM, (e.g., proximal to the AM to be masked) that allows the MM to mask the AM.

[0158] In most instances, the linker's length may be determined by counting, in a N- to C- direction, the number of amino acids from the N-terminus of the linker adjacent to the C-terminal amino acid of the preceding component, to the C-terminus of the linker adjacent to the N-terminal amino acid of the following component (i.e., where the linker length does not include either the C-terminal amino acid of the preceding component or the N-terminal amino acid of the following component).

[0159] In some embodiments, a linker may include a total of 1 to 50, 1 to 40, 1 to 30, 1 to 25 (e.g., 1 to 24, 1 to 22, 1 to 20, 1 to 18, 1 to 16, 1 to 15, 1 to 14, 1 to 12, 1 to 10, 1 to 8, 1 to 6, 1 to 5, 1 to 4, 1 to 3, 1 to 2, 2 to 25, 2 to 24, 2 to 22, 2 to 20, 2 to 18, 2 to 16, 2 to 15, 2 to 14, 2 to 12, 2 to 10, 2 to 8, 2 to 6, 2 to 5, 2 to 4, 2 to 3, 4 to 25, 4 to 24, 4 to 22, 4 to 20, 4 to 18, 4 to 16, 4 to 15, 4 to 14, 4 to 12, 4 to 10, 4 to 8, 4 to 6, 4 to 5, 5 to 25, 5 to 24, 5 to 22, 5 to 20, 5 to 18, 5 to 16, 5 to 15, 5 to 14, 5 to 12, 5 to 10, 5 to 8, 5 to 6, 6 to 25, 6 to 24, 6 to 22, 6 to 20, 6 to 18, 6 to 16, 6 to 15, 6 to 14, 6 to 12, 6 to 10, 6 to 8, 8 to 25, 8 to 24, 8 to 22, 8 to 20, 8 to 18, 8 to 16, 8 to 15, 8 to 14, 8 to 12, 8 to 10, 10 to 25, 10 to 24, 10 to 22, 10 to 20, 10 to 18, 10 to 16, 10 to 15, 10 to 14, 10 to 12, 12 to 25, 12 to 24, 12 to 22, 12 to 20, 12 to 18, 12 to 16, 12 to 15, 12 to 14, 14 to 25, 14 to 24, 14 to 22, 14 to 20, 14 to 18, 14 to 16, 14 to 15, 15 to 25, 15 to 24, 15 to 22, 15 to 20, 15 to 18, 15 to 16, 16 to 25, 16 to 24, 16 to 22, 16 to 20, 16 to 18, 18 to 25, 18 to 24, 18 to 22, 18 to 20, 20 to 25, 20 to 24, 20 to 22, 22 to 25, 22 to 24, or 24 to 25 amino acids). In some embodiments, the linker may include a total of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 amino acids.

[0160] In some embodiments, a linker may be rich in glycine (Gly or G) residues. In some embodiments, the linker may be rich in serine (Ser or S) residues. In some embodiments, the linker may be rich in glycine and serine residues. In some embodiments, the linker may have one or more glycine-serine residue pairs (GS) (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more GS pairs).

[0161] In some embodiments, the linker may have one or more Gly-Gly-Gly-Ser (GGGS) (SEQ ID NO: 402) sequences (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more GGGS sequences (SEQ ID NO: 402)). In some embodiments, the linker may have one or more Gly-Gly-Gly-Gly-Ser (GGGGS) (SEQ ID NO: 408) sequences (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more GGGGS sequences (SEQ ID NO: 408)). In some embodiments, the linker may have one or

more Gly-Gly-Ser-Gly (GGSG) (SEQ ID NO: 395) sequences (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more GGSG sequences (SEQ ID NO: 402)). Examples of the linkers may include glycine polymers (G)n, glycine-serine polymers (including, for example, (GS)n, (GGS)n, (GSGGS)n (SEQ ID NO: 652) and (GGGS)n (SEQ ID NO: 402), where n is an integer of at least one), glycine-alanine polymers, alanine-serine polymers, and other flexible linkers known in the art. Glycine and glycine-serine polymers may be relatively unstructured, and therefore may be able to serve as a neutral link between components. Glycine accesses significantly more phi-psi space than even alanine, and is much less restricted than residues with longer side chains (see Scheraga, Rev. Computational Chem. 11173-142 (1992)). Exemplary flexible linkers include one of or a combination of one or more of: GGSG (SEQ ID NO: 395), GGSGG (SEQ ID NO: 396), GSGSG (SEQ ID NO: 397), GSGGG (SEQ ID NO: 398), GGGSG (SEQ ID NO: 399), GSSSG (SEQ ID NO: 400), GSSGGSGGSGG (SEQ ID NO: 401), GGGS (SEQ ID NO: 402), GGGSGGGS (SEQ ID NO: 403), GGGSGGGSGGGS (SEQ ID NO: 404), GGGGSGGGGSGGGGS (SEQ ID NO: 405), GGGGSGGGGSGGGGSGGGGSGGGGS (SEQ ID NO: 406), GGGGSGGGGS (SEQ ID NO: 407), GGGGS (SEQ ID NO: 408), GS, GGGGSGS (SEQ ID NO: 409), GGGGSGGGGSGGGGSGS (SEQ ID NO: 410), GGSLDPKGGGGS (SEQ ID NO: 411), PKSCDKTHTCPPCPAPELLG (SEQ ID NO: 412), SKYGPPCPPCPAPEFLG (SEQ ID NO: 413), GKSSGSGSESKS (SEQ ID NO: 414), GSTSGSGKSSEGKG (SEQ ID NO: 415), GSTSGSGKSSEGSGSTKG (SEQ ID NO: 416), GSTSGSGKPGSGEGSTKG (SEQ ID NO: 417), GSTSGSGKPGSSEGST (SEQ ID NO: 418), GGGSSGGS (SEQ ID NO: 419), GGGGSGGGGSS (SEQ ID NO: 420), GGGSSGGSGGSSGGS (SEQ ID NO: 421), and GSTSGSGKPGSSEGST (SEQ ID NO: 422).

[0162] Examples of linkers may further include a sequence that is at least 70% identical (e.g., at least 72%, at least 74%, at least 75%, at least 76%, at least 78%, at least 80%, at least 82%, at least 84%, at least 85%, at least 86%, at least 88%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to the example linkers described herein. An ordinarily skilled artisan will recognize that design of an activatable molecules can include linkers that are all or partially flexible, such that the linker can include a flexible linker as well as one or more portions that confer less flexible structure to provide for a desired activatable molecules structure.

[0163] In some embodiments, an activatable molecule may include one, two, three, four, five, six, seven, eight, nine, or ten linker sequence(s) (e.g., the same or different linker

sequences of any of the exemplary linker sequences described herein or known in the art). In some embodiments, a linker may comprise sulfo-SIAB, SMPB, and sulfo-SMPB, wherein the linkers react with primary amines sulfhydryls.

Half-life extending moieties (EMs)

[0164] The activatable molecule may further comprise a half-life extending moiety (EM). In some examples, the half-life extending moiety may be a serum half-life extending moiety, i.e., capable of extending the serum half-life of the molecule attached to the EM.

[0165] In some examples, the EM may comprise a fragment crystallizable region (Fc domain) of an antibody. For example, the EM may be the Fc domain of an IgG (e.g., IgG1, IgG2, IgG3, or IgG4). In some examples, the EM may comprise a dimer formed by two Fc domains. The Fc domain may be a wild type peptide or a mutant. For example, the EM may comprise a dimer formed by two Fc domain mutants. In such cases, the two Fc domain mutants may be a Fc domain hole mutant and a Fc domain knob mutant. The knob and hole mutants may interact with each other to facilitate the dimerization of the two Fc domains. In some embodiments, the knob and hole mutants may comprise one or more amino acid modifications within the interface between two Fc domains (e.g., in the CH3 domain). In one example, the modifications comprise amino acid substitution T366W and optionally the amino acid substitution S354C in one IgG Fc domain and the amino acid substitutions T366S, L368A, Y407V and optionally Y349C in the other IgG Fc domain (numbering according to EU numbering system). Example of Fc mutants also include SEQ ID NOs: 423-424.

[0166] Examples of the Fc domain mutants also include those described in U.S. Pat. Nos. 7,695,936, which is incorporated herein by reference in its entirety. In one example, the modifications comprise amino acid substitution T366Y in one IgG Fc domain, and the amino acid substitutions Y407T in the other IgG Fc domain. In one example, the modifications comprise amino acid substitution T366W in one IgG Fc domain, and the amino acid substitutions Y407A in the other IgG Fc domain. In one example, the modifications comprise amino acid substitution F405A in one IgG Fc domain, and the amino acid substitutions T394W in the other IgG Fc domain. In one example, the modifications comprise amino acid substitution T366Y and F405A in one IgG Fc domain, and the amino acid substitutions T394W and Y407T in the other IgG Fc domain. In one example, the modifications comprise amino acid substitution T366W and F405W in one IgG Fc domain, and the amino acid substitutions T394S and Y407A in the other IgG Fc domain. In one example, the modifications comprise amino acid substitution F405W and Y407A in one IgG Fc domain,

and the amino acid substitutions T366W and T394S in the other IgG Fc domain. In one example, the modifications comprise amino acid substitution F405W in one IgG Fc domain, and the amino acid substitutions T394S in the other IgG Fc domain. The mutation positions in the Fc domains are numbered according to EU numbering system. The IgG Fc domain may be comprise a sequence of SEQ ID NOs: 425-428 (IgG1, IgG2, IgG3 or IgG4). In these sequences, amino acids 1-107 correspond to EU numbering 341-447.

[0167] In some examples, the Fc domains mutants may have reduced effector function. Examples of such Fc domains include those disclosed in in US20190135943, which incorporated herein by reference in its entirety.

[0168] Further examples of EMs include immunoglobulin (e.g., IgG), serum albumin (e.g., human serum albumin (HSA), hexa-hat GST (glutathione S-transferase) glutathione affinity, Calmodulin-binding peptide (CBP), Strep-tag, Cellulose Binding Domain, Maltose Binding Protein, S-Peptide Tag, Chitin Binding Tag, Immuno-reactive Epitopes, Epitope Tags, E2Tag, HA Epitope Tag, Myc Epitope, FLAG Epitope, AU1 and AU5 Epitopes, Glu-Glu Epitope, KT3 Epitope, IRS Epitope, Btag Epitope, Protein Kinase-C Epitope, and VSV Epitope.

[0169] In some embodiments, the serum half-life of an activatable molecule comprising an EM is longer than that of a counterpart molecule that is substantially the same as the activatable molecule but does not comprise the EM, e.g., the pK of the activatable molecule is longer than that of the reference molecule. In some examples, the activatable molecule with an EM may have a serum half-life that is at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 2-fold, 4-fold, 6-fold, 8-fold, 10-fold, 20-fold, 30-fold, 40-fold, 50-fold, 60-fold, 70-fold, 80-fold, 90-fold, 100-fold longer than the serum half-life of the reference counterpart molecule. In some embodiments, the serum half-life of the activatable molecule with an EM may be at least 15 days, 12 days, 11 days, 10 days, 9 days, 8 days, 7 days, 6 days, 5 days, 4 days, 3 days, 2 days, 1 day, 20 hours, 18 hours, 16 hours, 14 hours, 12 hours, 10 hours, 8 hours, 6 hours, 4 hours, 3 hours, 2 hours, or 1 hour when administered to an organism.

Conjugation Agents

[0170] In some aspects, the present disclosure provides conjugated polypeptides. In some embodiments, a conjugated polypeptide comprises a substrate-containing polypeptide herein conjugated to one or more agent, e.g., a targeting moiety to facilitate delivery to a cell or tissue of interest, a therapeutic agent (e.g., an antineoplastic agent such as chemotherapeutic or anti-neoplastic agent), a toxin, or a fragment thereof. The agents may be conjugated to a

component of the activatable molecules. In some embodiments, the conjugated polypeptide is an antibody-drug conjugate (ADC), which comprises an antibody or antigen-binding fragment thereof conjugated with a drug. In some examples, the antibody or antigen-binding fragment thereof may be conjugated with the drug via a substrate disclosed herein. In some examples, the antibody or antigen-binding fragment thereof may be an activatable antibody or antigen-binding fragment thereof (e.g., coupled with a MM via a substrate), which is further conjugated with a drug (e.g., via a cleavable or non-cleavable conjugating linker).

[0171] The term "agent" is used herein to denote a chemical compound, a mixture of chemical compounds, a biological macromolecule, or an extract made from biological materials. Examples of the agent include toxin, a microtubule inhibitor, a nucleic acid damaging agent, a dolastatin, an auristatin, a maytansinoid, a duocarmycin, a calicheamicin, or a combination thereof.

[0172] In some embodiments, the activatable molecule is conjugated to a cytotoxic agent, e.g., a toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof) or a radioactive isotope.

[0173] Examples of cytotoxic agents include that can be conjugated to the activatable molecules dolastatins and derivatives thereof (e.g., auristatin E, AFP, monomethyl auristatin D (MMAD), monomethyl auristatin F (MMAF), monomethyl auristatin E (MMAE), desmethyl auristatin E (DMAE), auristatin F, desmethyl auristatin F (DMAF), dolastatin 16 (DmJ), dolastatin 16 (Dpv), auristatin derivatives (e.g., auristatin tyramine, auristatin quinolone), maytansinoids (e.g., DM-1, DM-4), maytansinoid derivatives, duocarmycin, alpha-amanitin, turbostatin, phenstatin, hydroxyphenstatin, spongistatin 5, spongistatin 7, halistatin 1, halistatin 2, halistatin 3, halocomstatin, pyrrolobenzimidazoles (PBI), cibrostatin6, doxaliform, cemadotin analogue (CemCH2-SH), Pseudomonas toxin A (PES8) variant, Pseudomonase toxin A (ZZ-PE38) variant, ZJ-101, anthracycline, doxorubicin, daunorubicin, bryostatin, camptothecin, 7-substituted campothecin, 11-difluoromethylenedioxycamptothecin, combretastatins, debromoaplysiatoxin, KahaMide-F, discodermolide, and Ecteinascidins. In some embodiments, the agent is DM1 or DM4. In some embodiments, the agent is a duocarmycin or derivative thereof. In some embodiments, the agent is a calicheamicin or derivative thereof. In some embodiments, the agent is a pyrrolobenzodiazepine.

[0174] Examples of enzymatically active toxins that can be conjugated to the activatable molecules include diphtheria toxin, exotoxin A chain from Pseudomonas aeruginosa, ricin A

chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleuriies fordii proteins, dianfhin proteins, Phytolaca Americana proteins (e.g., PAPI, PAPII, and PAP-8), momordica charantia inhibitor, curcin, crotirs, sapaonaria officinalis inhibitor, geionin, mitogeliin, restrictocin, phenomycin, neomycin, and tricothecenes. A variety of radionuclides are available for the production of radioconjugated molecules. Examples of radionuclides include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

[0175] Examples of anti-neoplastics that can be conjugated to the activatable molecules include: adriamycin, cerubidine, bleomycin, alkeran, velban, oncovin, fluorouracil, methotrexate, thiotepa, bisantrene, novantrone, thioguanine, procarabizine, and cytarabine.

[0176] Examples of antivirals that can be conjugated to the activatable molecules include acyclovir, vira A, and symmetrel. Examples of antifungals that can be conjugated to the activatable molecules include: nystatin. Examples of detection reagents that can be conjugated to the activatable molecules include: fluorescein and derivatives thereof, fluorescein isothiocyanate (FITC). Examples of antibacterials that can be conjugated to the activatable molecules include: aminoglycosides, streptomycin, neomycin, kanamycin, amikacin, gentamicin, and tobramycin. Examples of 3beta,16beta,17alpha-trihydroxycholest-5-en-22-one 16-O-(2-O-4-methoxybenzoyl-beta-D-xylopyranosyl)-(1-->3)-(2-O-acetyl-alpha-L-arabinopyranoside) (OSW-1) that can be conjugated to the activatable molecules include: s-nitrobenzyloxycarbonyl derivatives of O6-benzylguanine, toposisomerase inhibitors, hemiasterlin, cephalotaxine, homoharringionine, pyrrol obenzodiazepine dimers (PBDs), functionalized pyrrolobenzodiazepenes, calcicheamicins, podophyiitoxins, taxanes, and vinca alkoids. Examples of radiopharmaceuticals that can be conjugated to the activatable molecules include: ^{123}I , ^{89}Zr, ^{125}I, ^{131}I, ^{201}T1, ^{62}Cu, ^{18}F, ^{68}Ga, ^{13}N, ^{15}O, ^{38}K, ^{82}Rb, ^{111}In, ^{133}Xe, ^{11}C, and ^{99m}Tc (Technetium). Examples of heavy metals that can be conjugated to the activatable molecules include: barium, gold, and platinum. Examples of anti-mycoplasmals that can be conjugated to the activatable molecules include: tylosine, spectinomycin, streptomycin B, ampicillin, sulfanilamide, polymyxin, and chloramphenicol.

[0177] In some embodiments, the agent is a nucleic acid damaging agent, such as a DNA alkylator or DNA intercalator, or other DNA damaging agent.

[0178] Additional examples of the agents that can be conjugated include those in Table 1 below.

Table 1: Exemplary Pharmaceutical Agents for Conjugation

CYTOTOXIC AGENTS	

Auristatins	Turbostatin
Auristatin E	Phenstatins
Monomethyl auristatin D (MMAD)	Hydroxyphenstatin
Monomethyl auristatin E (MMAE)	Spongistatin 5
Desmethyl auristatin E (DMAE)	Spongistatin 7
Auristatin F	Halistatin 1
Monomethyl auristatin F (MMAF)	Halistatin 2
Desmethyl auristatin F (DMAF)	Halistatin 3
Auristatin derivatives, *e.g.*, amides thereof	Modified Bryostatins
Auristatin tyramine	Halocomstatins
Auristatin quinoline	Pyrrolobenzimidazoles (PBI)
Dolastatins	Cibrostatin6
Dolastatin derivatives	Doxaliform
Dolastatin 16 DmJ	Anthracyclins analogues
Dolastatin 16 Dpv	
Maytansinoids, *e.g.* DM-1; DM-4	
Maytansinoid derivatives	Cemadotin analogue (CemCH2-SH)
Duocarmycin	Pseudomonas toxin A (PE38) variant
Duocarmycin derivatives	Pseudomonas toxin A (ZZ-PE38) variant
Alpha-amanitin	ZJ-101
Anthracyclines	OSW-1
Doxorubicin	4-Nitrobenzyloxycarbonyl Derivatives of O6-Benzylguanine
Daunorubicin	Topoisomerase inhibitors
Bryostatins	Hemiasterlin
Camptothecin	Cephalotaxine
Camptothecin derivatives	Homoharringtonine
7-substituted Camptothecin	Pyrrolobenzodiazepine dimers (PBDs)
10, 11-Difluoromethylenedioxycamptothecin	Functionalized pyrrolobenzodiazepenes
Combretastatins	Calicheamicins
Debromoaplysiatoxin	Podophyllotoxins
Kahalalide-F	Taxanes
Discodermolide	Vinca alkaloids
Ecteinascidins	
	CONJUGATABLE DETECTION REAGENTS
ANTIVIRALS	Fluorescein and derivatives thereof
Acyclovir	Fluorescein isothiocyanate (FITC)
Vira A	
Symmetrel	**RADIOPHARMACEUTICALS**
	^{125}I
ANTIFUNGALS	^{131}I
Nystatin	^{89}Zr
	^{111}In
ADDITIONAL ANTI-NEOPLASTICS	^{123}I
Adriamycin	^{131}I

Cerubidine	^{99}mTc (Technetium)
Bleomycin	^{201}Tl
Alkeran	^{133}Xe
Velban	^{11}C
Oncovin	^{62}Cu
Fluorouracil	^{18}F
Methotrexate	^{68}Ga
Thiotepa	^{13}N
Bisantrene	^{15}O
Novantrone	^{38}K
Thioguanine	^{82}Rb
Procarabizine	
Cytarabine	
	HEAVY METALS
ANTI-BACTERIALS	Barium
Aminoglycosides	Gold
Streptomycin	Platinum
Neomycin	
Kanamycin	**ANTI-MYCOPLASMALS**
Amikacin	Tylosine
Gentamicin	Spectinomycin
Tobramycin	
Streptomycin B	Nanoparticles
Spectinomycin	
Ampicillin	
Sulfanilamide	
Polymyxin	
Chloramphenicol	

[0179] In some embodiments, the activatable molecule comprises a signal peptide. If comprising multiple polypeptides, the activatable molecule may comprise multiple signal peptides, e.g., one signal peptide for each of the multiple polypeptides. A signal peptide may be a peptide (e.g., 10-30 amino acids long) present at a terminus (e.g., the N-terminus or C-terminus) of a newly synthesized proteins that are destined toward the secretory pathway. In some embodiments, the signal peptide may be conjugated to the activatable molecule via a spacer. In some embodiments, the spacer may be conjugated to the activatable molecule in the absence of a signal peptide.

[0180] Those of ordinary skill in the art will recognize that a large variety of possible agents may be conjugated to any of the activatable molecules described herein. The agents may be conjugated to another component of the activatable molecule by a conjugating linker. Conjugation may include any chemical reaction that binds the two molecules so long as the

activatable molecule and the other moiety retain their respective activities. Conjugation may include many chemical mechanisms, e.g., covalent binding, affinity binding, intercalation, coordinate binding, and complexation. In some embodiments, the binding may be covalent binding. Covalent binding may be achieved either by direct condensation of existing side chains or by the incorporation of external bridging molecules. Many bivalent or polyvalent linking agents may be useful in conjugating any of the activatable molecules described herein. For example, conjugation may include organic compounds, such as thioesters, carbodiimides, succinimide esters, glutaraldehyde, diazobenzenes, and hexamethylene diamines. In some embodiments, the activatable molecules may include, or otherwise introduce, one or more non-natural amino acid residues to provide suitable sites for conjugation.

[0181] In some embodiments, an agent is attached by disulfide bonds (e.g., disulfide bonds on a cysteine molecule) to the activatable molecule. Since many cancers naturally release high levels of glutathione, a reducing agent, glutathione present in the cancerous tissue microenvironment can reduce the disulfide bonds, and subsequently release the agent at the site of delivery.

[0182] In some embodiments, when the agent binds its target in the presence of complement within the target site (e.g., diseased tissue (e.g., cancerous tissue)), the amide or ester bond attaching the agent to the linker is cleaved, resulting in the release of the agent in its activated form. These agents when administered to a subject, may accomplish delivery and release of the agent at the target site (e.g., diseased tissue (e.g., cancerous tissue)). These agents may be effective for the in vivo delivery of any of the agents described herein.

[0183] In some embodiments, the one or more agents is conjugated to a component of the activatable molecule (e.g., AM) via a conjugating linker. The conjugating linker may be a peptide or chemical moiety linking the agent and the activatable molecule. In some examples, the conjugating linker may be cleavable (e.g., by an enzyme such as a protease). In some examples, the conjugating linker may be non-cleavable (e.g., cannot be cleaved by an enzyme such as a protease). In some embodiments, the conjugating linker may be non-cleavable by enzymes of the complement system. In some embodiments, two or more conjugating linkers are present. The two or more conjugating linkers may be the same, i.e., cleavable or non-cleavable. The two or more conjugating linkers may be different, i.e., at least one cleavable and at least one non-cleavable. For example, the agent may be released without complement activation since complement activation ultimately lyses the target cell. In such embodiments, the conjugate and/or agent is to be delivered to the target cell (e.g., hormones, enzymes,

corticosteroids, neurotransmitters, or genes). Furthermore, the conjugating linker may be mildly susceptible to cleavage by serum proteases, and the conjugate and/or agent is released slowly at the target site.

[0184] In some embodiments, the agent is conjugated to a component of the activatable molecule via a maleimide caproyl-valine-citrulline linker or a maleimide PEG-valine-citrulline linker. In some embodiments, the agent is conjugated to a component of the activatable molecule via a maleimide caproyl-valine-citrulline linker. In some embodiments, the agent is conjugated to a component of the activatable molecule via a maleimide PEG-valine-citrulline linker. In some embodiments, the agent is monomethyl auristatin D (MMAD) conjugated to a component of the activatable molecule via a maleimide PEG-valine-citrulline-para-aminobenzyloxycarbonyl linker, and this linker payload construct is vc-MMAD. In some embodiments, the agent is monomethyl auristatin E (MMAE) conjugated to a component of the activatable molecule via a maleimide PEG-valine-citrulline-para-aminobenzyloxycarbonyl linker, and this linker payload construct is vc-MMAE.

[0185] In some embodiments, the agent may be designed such that the agent is delivered to the target site (e.g., disease tissue (e.g., cancerous tissue)) but the conjugate and/or agent is not released.

[0186] In some embodiments, the agent may be attached to an AM either directly or via amino acids (e.g., D-amino acids), peptides, thiol-containing moieties, or other organic compounds that may be modified to include functional groups that can subsequently be utilized in attachment to AM by methods described herein.

[0187] In some embodiments, an activatable molecule includes at least one point of conjugation for an agent. In some embodiments, all possible points of conjugation are available for conjugation to an agent. In some embodiments, the one or more points of conjugation may include sulfur atoms involved in disulfide bonds, sulfur atoms involved in interchain disulfide bonds, sulfur atoms involved in interchain sulfide bonds but not sulfur atoms involved in intrachain disulfide bonds, and/or sulfur atoms of cysteine or other amino acid residues containing a sulfur atom. In such cases, residues may occur naturally in the protein construct structure or may be incorporated into the protein construct using methods including site-directed mutagenesis, chemical conversion, or mis-incorporation of non-natural amino acids.

[0188] The present disclosure also provides methods and materials for preparing an activatable molecule with one or more conjugated agents. In some embodiments, an

activatable molecule may be modified to include one or more interchain disulfide bonds. For example, disulfide bonds may undergo reduction following exposure to a reducing agent such as, without limitation, TCEP, DTT, or β-mercaptoethanol. In some cases, the reduction of the disulfide bonds may be only partial. As used herein, the term partial reduction refers to situations where an activatable molecule is contacted with a reducing agent and a fraction of all possible sites of conjugation undergo reduction (e.g., not all disulfide bonds are reduced). In some embodiments, an activatable molecule may be partially reduced following contact with a reducing agent if less than 99%, (e.g., less than 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10% or 5%) of all possible sites of conjugation are reduced. In some embodiments, the activatable molecule having a reduction in one or more interchain disulfide bonds may be conjugated to a drug reactive with free thiols.

[0189] The present disclosure also provides methods and materials for conjugating a therapeutic agent to a particular location on an activatable molecule. In some embodiments, an activatable molecule may be modified so that the therapeutic agents can be conjugated to the activatable molecule at particular locations on the activatable molecule. For example, an activatable molecule may be partially reduced in a manner that facilitates conjugation to the activatable molecule. In such cases, partial reduction of the activatable molecule may occur in a manner that conjugation sites in the activatable molecule are not reduced. In some embodiments, the conjugation site(s) on the activatable molecule may be selected to facilitate conjugation of an agent at a particular location on the protein construct. Various factors can influence the "level of reduction" of the activatable molecule upon treatment with a reducing agent. For example, without limitation, the ratio of reducing agent to activatable molecule, length of incubation, incubation temperature, and/or pH of the reducing reaction solution can require optimization in order to achieve partial reduction of the activatable molecule with the methods and materials described herein. Any appropriate combination of factors (e.g., ratio of reducing agent to activatable molecule, the length and temperature of incubation with reducing agent, and/or pH of reducing agent) may be used to achieve partial reduction of the activatable molecule (e.g., general reduction of possible conjugation sites or reduction at specific conjugation sites).

[0190] An effective ratio of reducing agent to activatable molecule can be any ratio that at least partially (i.e., partially or fully) reduces the activatable molecule in a manner that allows conjugation to an agent (e.g., general reduction of possible conjugation sites or

reduction at specific conjugation sites). In some embodiments, the ratio of reducing agent to activatable molecule may be in a range from about 20:1 to 1:1, from 10:1 to 1:1, from 9:1 to 1:1, from 8:1 to 1:1, from 7:1 to 1:1, from 6:1 to 1:1, from 5:1 to 1:1, from 4:1 to 1:1, from 3:1 to 1:1, from 2:1 to 1:1, from 20:1 to 1:1.5, from 10:1 to 1:1.5, from 9:1 to 1:1.5, from 8:1 to 1:1.5, from 7:1 to 1:1.5, from 6:1 to 1:1.5, from 5:1 to 1:1.5, from 4:1 to 1:1.5, from 3:1 to 1:1.5, from 2:1 to 1:1.5, from 1.5:1 to 1:1.5, or from 1:1 to 1:1.5.

[0191] An effective incubation time and temperature for treating an activatable molecule with a reducing agent may be any time and temperature that at least partially reduces the activatable molecule in a manner that allows conjugation of an agent to an activatable molecule (e.g., general reduction of possible conjugation sites or reduction at specific conjugation sites). In some embodiments, the incubation time and temperature for treating an activatable molecule may be in a range from about 1 hour at 37 °C to about 12 hours at 37 °C (or any subranges therein).

[0192] An effective pH for a reduction reaction for treating an activatable molecule with a reducing agent can be any pH that at least partially reduces the activatable molecule in a manner that allows conjugation of the activatable molecule to an agent (e.g., general reduction of possible conjugation sites or reduction at specific conjugation sites).

[0193] When a partially-reduced activatable molecule is contacted with an agent containing thiols, the agent may conjugate to the interchain thiols in the activatable molecule. An agent can be modified in a manner to include thiols using a thiol-containing reagent (e.g., cysteine or N-acetyl cysteine). For example, the activatable molecule can be partially reduced following incubation with reducing agent (e.g., TEPC) for about 1 hour at about 37 °C at a desired ratio of reducing agent to activatable molecule. An effective ratio of reducing agent to activatable molecule may be any ratio that partially reduces at least two interchain disulfide bonds located in the activatable molecule in a manner that allows conjugation of a thiol-containing agent (e.g., general reduction of possible conjugation sites or reduction at specific conjugation sites).

[0194] In some embodiments, an activatable molecule may be reduced by a reducing agent in a manner that avoids reducing any intrachain disulfide bonds. In some embodiments of, an activatable molecule may be reduced by a reducing agent in a manner that avoids reducing any intrachain disulfide bonds and reduces at least one interchain disulfide bond.

[0195] In some embodiments, the agent may be a detectable moiety such as, for example, a label or other marker. For example, the agent may be or include a radiolabeled amino acid,

one or more biotinyl moieties that can be detected by marked avidin (e.g., streptavidin containing a fluorescent marker or enzymatic activity that can be detected by optical or calorimetric methods), one or more radioisotopes or radionuclides, one or more fluorescent labels, one or more enzymatic labels, and/or one or more chemiluminescent agents. In some embodiments, detectable moieties may be attached by spacer molecules. In some embodiments, the detectable label may include an imaging agent, a contrasting agent, an enzyme, a fluorescent label, a chromophore, a dye, one or more metal ions, or a ligand-based label. In some embodiments, the imaging agent may comprise a radioisotope. In some embodiments, the radioisotope may be indium or technetium. In some embodiments, the contrasting agent may comprise iodine, gadolinium or iron oxide. In some embodiments, the enzyme may comprise horseradish peroxidase, alkaline phosphatase, or β-galactosidase. In some embodiments, the fluorescent label may comprise yellow fluorescent protein (YFP), cyan fluorescent protein (CFP), green fluorescent protein (GFP), modified red fluorescent protein (mRFP), red fluorescent protein tdimer2 (RFP tdimer2), HCRED, or a europium derivative. In some embodiments, the luminescent label may comprise an N- methylacrydium derivative. In some embodiments, the label may comprise an Alexa Fluor® label, such as Alex Fluor® 680 or Alexa Fluor® 750. In some embodiments, the ligand-based label may comprise biotin, avidin, streptavidin or one or more haptens.

[0196] Further examples of detectable labels also include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

[0197] In some embodiments, the agent may be conjugated to the activatable molecule using a carbohydrate moiety, sulfhydryl group, amino group, or carboxylate group. In some embodiments, the agent may be conjugated to the activatable molecule via a linker and/or a substrate described herein. In some embodiments, the agent may be conjugated to a cysteine

or a lysine in the activatable molecule. In some embodiments, the agent may be conjugated to another residue of the activatable molecule, such as those residues disclosed herein.

[0198] In some embodiments, a variety of bifunctional protein-coupling agents may be used to conjugate the agent to the activatable molecule including N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (e.g., dimethyl adipimidate HCL), active esters (e.g., disuccinimidyl suberate), aldehydes (e.g., glutareldehyde), bis-azido compounds (e.g., bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (e.g., bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (e.g., tolyene 2,6-diisocyanate), and bis-active fluorine compounds (e.g., 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science 238: 1098 (1987). In some embodiments, a carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) chelating agent can be used to conjugate a radionucleotide to the activatable molecule. (See, e.g., WO94/11026).

[0199] Suitable conjugating linkers also include those described in the literature. (See, for example, Ramakrishnan, S. et al., Cancer Res. 44:201-208 (1984) describing use of MBS (M-maleimidobenzoyl-N-hydroxysuccinimide ester). See also, U.S. Patent No. 5,030,719, describing use of halogenated acetyl hydrazide derivative coupled to an activatable molecule by way of an oligopeptide. In some embodiments, suitable conjugating linkers include: (i) EDC (1-ethyl-3-(3-dimethylamino-propyl) carbodiimide hydrochloride; (ii) SMPT (4-succinimidyloxycarbonyl-alpha-methyl-alpha-(2-pridyl-dithio)-toluene (Pierce Chem. Co., Cat. (21558G); (iii) SPDP (succinimidyl-6 [3-(2-pyridyldithio) propionamido] hexanoate (Pierce Chem. Co., Cat #21651G); (iv) Sulfo-LC-SPDP (sulfosuccinimidyl 6 [3-(2-pyridyldithio)-propianamide] hexanoate (Pierce Chem. Co. Cat. #2165-G); and (v) sulfo-NHS (N-hydroxysulfo-succinimide: Pierce Chem. Co., Cat. #24510) conjugated to EDC. Additional example agents include SMCC, sulfo-SMCC, SPDB, and sulfo-SPDB.

[0200] Exemplary conjugating linkers for attachment to reduced activatable molecules include those having certain reactive groups capable of reaction with a sulfhydryl group of a reduced antibody or fragment. Such reactive groups include reactive haloalkyl groups (including, for example, haloacetyl groups), p-mercuribenzoate groups and groups capable of Michael-type addition reactions (including, for example, maleimides and groups of the type described by Mitra and Lawton, 1979, J. Amer. Chem. Soc. 101: 3097-3110).

[0201] Exemplary conjugating linkers for attachment to neither oxidized nor reduced activatable molecules include those having certain functional groups capable of reaction with the primary amino groups present in unmodified lysine residues in the activatable molecules. Such reactive groups include NHS carboxylic or carbonic esters, sulfo-NHS carboxylic or carbonic esters, 4-nitrophenyl carboxylic or carbonic esters, pentafluorophenyl carboxylic or carbonic esters, acyl imidazoles, isocyanates, and isothiocyanates, and other dehydrating agents utilized for carboxamide formation. In these instances, the functional groups present in the suitable conjugating linkers include primary and secondary amines, hydrazines, hydroxylamines, and hydrazides.

[0202] The agent may be attached to the conjugating linker before or after the conjugating linker is attached to the activatable molecule. In certain applications it may be desirable to first produce an activatable molecule-conjugating linker intermediate in which the conjugating linker is free of an associated agent. Depending upon the particular application, a specific agent may then be covalently attached to the conjugating linker. In some embodiments, the AM is first attached to the MM, substrate and associated linking peptides and then attached to the conjugating linker for conjugation purposes.

[0203] In specific embodiments, branched conjugating linkers that have multiple sites for attachment of agents are utilized. For multiple site conjugating linkers, a single covalent attachment to an activatable molecule may result in an activatable molecule-linker intermediate capable of binding an agent at a number of sites. The sites may be aldehyde or sulfhydryl groups or any chemical site to which agents can be attached.

[0204] In some embodiments, higher specific activity (or higher ratio of agents to activatable molecule) can be achieved by attachment of a single site conjugating linker at a plurality of sites on the activatable molecule. This plurality of sites may be introduced into the activatable molecule by either of two methods. First, one may generate multiple aldehyde groups and/or sulfhydryl groups in the same activatable molecule. Second, one may attach to an aldehyde or sulfhydryl of the activatable molecule a branched conjugating linker having multiple functional sites for subsequent attachment to conjugating linkers. The functional sites of the branched conjugating linker or multiple site conjugating linker may be aldehyde or sulfhydryl groups, or may be any chemical site to which conjugating linkers may be attached. Still higher specific activities may be obtained by combining these two approaches, that is, attaching multiple site conjugating linkers at several sites on the activatable molecule.

[0205] Peptide conjugating linkers that are susceptible to cleavage by enzymes of the complement system, such as but not limited to u-plasminogen activator, tissue plasminogen activator, trypsin, plasmin, or another enzyme having proteolytic activity may be used in one embodiment of the present disclosure. According to one method of the present disclosure, an agent is attached via a conjugating linker susceptible to cleavage by complement. The antibody is selected from a class that can activate complement. The antibody-agent conjugate, thus, activates the complement cascade and releases the agent at the target site. According to another method of the present disclosure, an agent is attached via a conjugating linker susceptible to cleavage by enzymes having a proteolytic activity such as a u-plasminogen activator, a tissue plasminogen activator, plasmin, or trypsin. These cleavable conjugating linkers are useful in conjugated activatable molecules that include an extracellular toxin, *e.g.*, by way of non-limiting example, any of the extracellular toxins shown in Table 1.

[0206] Non-limiting examples of cleavable linker sequences include any cleavable sequence disclosed herein or incorporated herein by reference as well as the exemplary sequences provided in Table 2.

Table 2: Exemplary Conjugating Linker Sequences for Conjugation

Types of Cleavable Sequences	Amino Acid Sequence
Plasmin cleavable sequences	
Pro-urokinase	PRFKIIGG (SEQ ID NO: 429)
	PRFRIIGG (SEQ ID NO: 430)
TGFβ	SSRHRRALD (SEQ ID NO: 431)
Plasminogen	RKSSIIIRMRDVVL (SEQ ID NO: 432)
Staphylokinase	SSSFDKGKYKKGDDA (SEQ ID NO: 433)
	SSSFDKGKYKRGDDA (SEQ ID NO: 434)
Factor Xa cleavable sequences	IEGR (SEQ ID NO: 435)
	IDGR (SEQ ID NO: 436)
	GGSIDGR (SEQ ID NO: 437)
MMP cleavable sequences	
Gelatinase A	PLGLWA (SEQ ID NO: 438)
Collagenase cleavable sequences	
Calf skin collagen (α1(I) chain)	GPQGIAGQ (SEQ ID NO: 439)
Calf skin collagen (α2(I) chain)	GPQGLLGA (SEQ ID NO: 440)
Bovine cartilage collagen (α1(II) chain)	GIAGQ (SEQ ID NO: 441)
Human liver collagen (α1(III) chain)	GPLGIAGI (SEQ ID NO: 442)
Human α_2M	GPEGLRVG (SEQ ID NO: 443)
Human PZP	YGAGLGVV (SEQ ID NO:444)
	AGLGVVER (SEQ ID NO: 445)

	AGLGISST (SEQ ID NO: 446)
Rat α_1M	EPQALAMS (SEQ ID NO: 447)
	QALAMSAI (SEQ ID NO: 448)
Rat α_2M	AAYHLVSQ (SEQ ID NO: 449)
	MDAFLESS (SEQ ID NO: 450)
Rat α_1I_3(2J)	ESLPVVAV (SEQ ID NO: 451)
Rat α_1I_3(27J)	SAPAVESE (SEQ ID NO: 452)
Human fibroblast collagenase	DVAQFVLT (SEQ ID NO: 453)
(autolytic cleavages)	VAQFVLTE (SEQ ID NO: 454)
	AQFVLTEG (SEQ ID NO: 455)
	PVQPIGPQ (SEQ ID NO:456)

[0207] In addition, the agents may be attached via disulfide bonds (for example, the disulfide bonds on a cysteine molecule) to the activatable molecule. Since many tumors naturally release high levels of glutathione (a reducing agent) this can reduce the disulfide bonds with subsequent release of the agent at the site of delivery. In some embodiments, the reducing agent that would modify a substrate would also modify the conjugating linker of the conjugated activatable molecule.

[0208] In some embodiments, it may be necessary to construct the conjugating linker in such a way as to optimize the spacing between the agent and the activatable molecule. This may be accomplished by use of a conjugating linker of the general structure:

$$W - (CH_2)n - Q$$

wherein

W is either -NH-CH_2- or -CH_2-;

Q is an amino acid, a polypeptide having between 2 to 20 amino acids; and

n is an integer from 0 to 20.

[0209] In some embodiments, the conjugating linker may comprise a spacer element and a cleavable element. The spacer element serves to position the cleavable element away from the core of the activatable molecule such that the cleavable element is more accessible to the enzyme responsible for cleavage. Certain of the branched linkers described above may serve as spacer elements.

[0210] Throughout this discussion, it should be understood that attachment of the conjugating linker to the agent (or of spacer element to cleavable element, or cleavable element to agent) need not be by a particular mode of attachment or reaction. Any reaction providing a product of suitable stability and biological compatibility is acceptable.

[0211] In some embodiments, when release of an agent is desired, an activatable molecule that is an antibody of a class that can activate complement is used. The resulting conjugate retains both the ability to bind antigen and activate the complement cascade. Thus, according to this embodiment of the present disclosure, an agent is joined to one end of the cleavable conjugating linker or cleavable element and the other end of the conjugating linker group is attached to a specific site on the activatable molecule. For example, if the agent has a hydroxyl group or an amino group, it may be attached to the carboxyl terminus of a peptide, amino acid or other suitably chosen conjugating linker via an ester or amide bond, respectively. For example, such agents may be attached to the linker peptide via a carbodimide reaction. If the agent contains functional groups that would interfere with attachment to the conjugating linker, these interfering functional groups can be blocked before attachment and deblocked once the product conjugate or intermediate is made. The opposite or amino terminus of the linker is then used either directly or after further modification for binding to an activatable molecule that is capable of activating complement.

[0212] Conjugating linkers (or spacer elements of conjugating linkers) may be of any desired length, one end of which can be covalently attached to specific sites on the activatable molecule. The other end of the conjugating linker or spacer element may be attached to an amino acid or peptide conjugating linker.

[0213] Thus when these conjugates bind antigen in the presence of complement the amide or ester bond that attaches the agent to the linker will be cleaved, resulting in release of the agent in its active form. These conjugates, when administered to a subject, will accomplish delivery and release of the agent at the target site, and are particularly effective for the in vivo delivery of pharmaceutical agents, antibiotics, antimetabolites, antiproliferative agents and the like.

[0214] In some embodiments, release of the agent without complement activation is desired since activation of the complement cascade will ultimately lyse the target cell. Hence, this approach is useful when delivery and release of the agent should be accomplished without killing the target cell. Such is the goal when delivery of cell mediators such as hormones, enzymes, corticosteroids, neurotransmitters, genes or enzymes to target cells is desired. These conjugates may be prepared by attaching the agent to an activatable molecule that is not capable of activating complement via a linker that is mildly susceptible to cleavage by serum proteases. When this conjugate is administered to an individual, antigen-antibody complexes

will form quickly whereas cleavage of the agent will occur slowly, thus resulting in release of the compound at the target site.

[0215] In some embodiments, the activatable molecule may be conjugated to one or more therapeutic agents using certain biochemical cross-linkers. Cross-linking reagents form molecular bridges that tie together functional groups of two different molecules. To link two different proteins in a step-wise manner, hetero-bifunctional cross-linkers can be used that eliminate unwanted homopolymer formation.

[0216] Peptidyl conjugating linkers cleavable by lysosomal proteases are also useful, for example, Val-Cit, Val-Ala or other dipeptides. In addition, acid-labile conjugating linkers cleavable in the low-pH environment of the lysosome may be used, for example: bis-sialyl ether. Other suitable conjugating linkers include cathepsin-labile substrates, particularly those that show optimal function at an acidic pH.

[0217] Exemplary hetero-bifunctional cross-linkers are referenced in Table 3.

Table 3: Exemplary Hetero-Bifunctional Cross-Linkers

HETERO-BIFUNCTIONAL CROSS-LINKERS			
Linker	Reactive Toward	Advantages and Applications	Spacer Arm Length after cross-linking (Angstroms)
SMPT	Primary amines Sulfhydryls	Greater stability	11.2 Å
SPDP	Primary amines Sulfhydryls	Thiolation Cleavable cross-linking	6.8 Å
LC-SPDP	Primary amines Sulfhydryls	Extended spacer arm	15.6 Å
Sulfo-LC-SPDP	Primary amines Sulfhydryls	Extender spacer arm Water-soluble	15.6 Å
SMCC	Primary amines Sulfhydryls	Stable maleimide reactive group Enzyme-antibody conjugation Hapten-carrier protein conjugation	11.6 Å
Sulfo-SMCC	Primary amines Sulfhydryls	Stable maleimide reactive group Water-soluble Enzyme-antibody conjugation	11.6 Å

MBS	Primary amines Sulfhydryls	Enzyme-antibody conjugation Hapten-carrier protein conjugation	9.9 Å
Sulfo-MBS	Primary amines Sulfhydryls	Water-soluble	9.9 Å
SIAB	Primary amines Sulfhydryls	Enzyme-antibody conjugation	10.6 Å
Sulfo-SIAB	Primary amines Sulfhydryls	Water-soluble	10.6 Å
SMPB	Primary amines Sulfhydryls	Extended spacer arm Enzyme-antibody conjugation	14.5 Å
Sulfo-SMPB	Primary amines Sulfhydryls	Extended spacer arm Water-soluble	14.5 Å
EDE/Sulfo-NHS	Primary amines Carboxyl groups	Hapten-Carrier conjugation	0
ABH	Carbohydrates Nonselective	Reacts with sugar groups	11.9 Å

[0218] In some embodiments, the agent may be designed so that the agent is delivered to the target but not released. This may be accomplished by attaching an agent to an activatable molecule either directly or via a non-cleavable conjugating linker.

[0219] These non-cleavable conjugating linkers may include amino acids, peptides, D-amino acids or other organic compounds that may be modified to include functional groups that can subsequently be utilized in attachment to activatable molecules by the methods described herein.

[0220] In some embodiments, a compound may be attached to activatable molecules that do not activate complement. When using activatable molecules that are incapable of complement activation, this attachment may be accomplished using conjugating linkers that are susceptible to cleavage by activated complement or using linkers that are not susceptible to cleavage by activated complement.

[0221] The substrate-containing polypeptides disclosed herein can also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein et al., Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang et al., Proc. Natl Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556. Particularly useful liposomes can be generated by the reverse-phase evaporation

method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. A component of an activatable molecule can be conjugated to the liposomes as described in Martin et al., J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction.

[0222] The agents described above may contain components that have different attributes, thus leading to conjugates with differing physio-chemical properties. For example, sulfo-NHS esters of alkyl carboxylates are more stable than sulfo-NHS esters of aromatic carboxylates. NHS-ester containing linkers are less soluble than sulfo-NHS esters. Further, the SMPT contains a sterically-hindered disulfide bond, and can form conjugates with increased stability. Disulfide linkages, are in general, less stable than other linkages because the disulfide linkage is cleaved in vitro, resulting in less conjugate available. Sulfo-NHS, in particular, can enhance the stability of carbodimide couplings. Carbodimide couplings (such as EDC) when used in conjunction with sulfo-NHS, forms esters that are more resistant to hydrolysis than the carbodimide coupling reaction alone.

[0223] Those of ordinary skill in the art will recognize that a large variety of possible agents can be conjugated to the activatable molecule of the disclosure. (See, for example, "Conjugate Vaccines", Contributions to Microbiology and Immunology, J. M. Cruse and R. E. Lewis, Jr (eds), Carger Press, New York, (1989), the entire contents of which are incorporated herein by reference). In general, an effective conjugation of an agent (e.g., cytotoxic agent) to an activatable molecule can be accomplished by any chemical reaction that will bind the agent to the activatable molecule while also allowing the agent and the activatable molecule to retain functionality.

NUCLEIC ACIDS AND VECTORS

[0224] In some aspects, the present disclosure further provides nucleic acids comprising sequences that encode the substrate-containing polypeptides and polypeptide complexes (e.g., activatable molecules) herein, or components or fragment thereof. The nucleic acids may comprise coding sequences for the substrates. The nucleic acids may further comprise coding sequences for other components in an activatable molecule, e.g., the AMs, the MMs, the EM and/or the linker(s). In cases where the activatable molecule comprises multiple polypeptides, the nucleic acids may comprise coding sequences for the multiple polypeptides. In some examples, the coding sequence for one of the polypeptides is comprised in a nucleic acid molecule, and the coding sequence for another one of the polypeptides is comprised in another

nucleic acid molecule. In some examples, the coding sequences for two or more of the multiple polypeptides are comprised in the same nucleic acid molecule.

[0225] Unless otherwise specified, a "nucleic acid sequence encoding a protein" includes all nucleotide sequences that are degenerate versions of each other and thus encode the same amino acid sequence. The term "nucleic acid" refers to a deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), or a combination thereof, in either a single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses complementary sequences as well as the sequence explicitly indicated. In some embodiments, the nucleic acid is DNA. In some embodiments, the nucleic acid is RNA.

[0226] Modifications may be introduced into a nucleotide sequence by standard techniques known in the art, such as site-directed mutagenesis and polymerase chain reaction (PCR)-mediated mutagenesis. Conservative amino acid substitutions are ones in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include: amino acids with acidic side chains (e.g., aspartate and glutamate), amino acids with basic side chains (e.g., lysine, arginine, and histidine), non-polar amino acids (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, and tryptophan), uncharged polar amino acids (e.g., glycine, asparagine, glutamine, cysteine, serine, threonine and tyrosine), hydrophilic amino acids (e.g., arginine, asparagine, aspartate, glutamine, glutamate, histidine, lysine, serine, and threonine), hydrophobic amino acids (e.g., alanine, cysteine, isoleucine, leucine, methionine, phenylalanine, proline, tryptophan, tyrosine, and valine). Other families of amino acids include: aliphatic-hydroxy amino acids (e.g., serine and threonine), amide family (e.g., asparagine and glutamine), alphatic family (e.g., alanine, valine, leucine and isoleucine), and aromatic family (e.g., phenylalanine, tryptophan, and tyrosine).

[0227] The present disclosure further provides vectors and sets of vectors comprising any of the nucleic acids described herein. One skilled in the art will be capable of selecting suitable vectors or sets of vectors (e.g., expression vectors) for making any of the activatable molecules described herein, and using the vectors or sets of vectors to express any of the activatable molecules described herein. For example, in selecting a vector or a set of vectors, the type of cell may be selected such that the vector(s) may need to be able to integrate into a

chromosome of the cell and/or replicate in it. Example vectors that can be used to produce an activatable molecule are also described herein. As used herein, the term "vector" refers to a polynucleotide capable of inducing the expression of a recombinant protein (e.g., a first or second monomer) in a cell (e.g., any of the cells described herein). A "vector" is able to deliver nucleic acids and fragments thereof into a host cell, and includes regulatory sequences (e.g., promoter, enhancer, poly(A) signal). Exogenous polynucleotides may be inserted into the expression vector in order to be expressed. The term "vector" also includes artificial chromosomes, plasmids, retroviruses, and baculovirus vectors.

[0228] Methods for constructing suitable vectors that comprise any of the nucleic acids described herein, and suitable for transforming cells (e.g., mammalian cells) are well-known in the art. See, e.g., Sambrook et al., Eds. "Molecular Cloning: A Laboratory Manual," 2nd Ed., Cold Spring Harbor Press, 1989 and Ausubel et al., Eds. "Current Protocols in Molecular Biology," Current Protocols, 1993.

[0229] Examples of vectors include plasmids, transposons, cosmids, and viral vectors (e.g., any adenoviral vectors (e.g., pSV or pCMV vectors), adeno-associated virus (AAV) vectors, lentivirus vectors, and retroviral vectors), and any Gateway® vectors. A vector may, for example, include sufficient cis-acting elements for expression; other elements for expression may be supplied by the host mammalian cell or in an in vitro expression system. Skilled practitioners will be capable of selecting suitable vectors and mammalian cells for making any activatable molecule described herein.

[0230] In some embodiments, the substrate-containing polypeptides may be made biosynthetically using recombinant DNA technology and expression in eukaryotic or prokaryotic species.

CELLS

[0231] In some aspects, the present disclosure provides recombinant host cells comprising any of the vectors or nucleic acids described herein. The cells may be used to produce the substrate-containing polypeptides (e.g., activatable molecules) described herein. In some embodiments, the cell may be an animal cell, a mammalian cell (e.g., a human cell), a rodent cell (e.g., a mouse cell, a rat cell, a hamster cell, or a guinea pig cell), a non-human primate cell, an insect cell, a bacterial cell, a fungal cell, or a plant cell. In some embodiments, the cell may be a eukaryotic cell. As used herein, the term "eukaryotic cell" refers to a cell having a distinct, membrane-bound nucleus. Such cells may include, for example, mammalian (e.g., rodent, non-human primate, or human), insect, fungal, or plant cells. In

some embodiments, the eukaryotic cell is a yeast cell, such as Saccharomyces cerevisiae. In some embodiments, the eukaryotic cell is a higher eukaryote, such as mammalian, avian, plant, or insect cells. Non-limiting examples of mammalian cells include Chinese hamster ovary (CHO) cells and human embryonic kidney cells (e.g., HEK293 cells). In some embodiments, the cell may be a prokaryotic cell, e.g., an E coli cell.

[0232] Methods of introducing nucleic acids and vectors (e.g., any of the vectors or any of the sets of vectors described herein) into a cell are known in the art. Examples of methods that can be used to introducing a nucleic acid into a cell include: lipofection, transfection, calcium phosphate transfection, cationic polymer transfection, viral transduction (e.g., adenoviral transduction, lentiviral transduction), nanoparticle transfection, and electroporation.

[0233] In some embodiments, the introducing step includes introducing into a cell a vector (e.g., any of the vectors or sets of vectors described herein) including a nucleic acid encoding the monomers that make up any activatable molecule described herein.

COMPOSITIONS AND KITS

[0234] The present disclosure also provides compositions and kits comprising the substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides) described herein. The compositions and kits may further comprise one or more excipients, carriers, reagents, instructions needed for the use of the activatable molecules.

[0235] In some embodiments, the compositions may be pharmaceutical compositions, which comprise the substrate-containing polypeptides, derivatives, fragments, analogs and homologs thereof. The pharmaceutical compositions may comprise the substrate-containing and a pharmaceutically acceptable carrier. As used herein, the term "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Suitable examples of such carriers or diluents include water, saline, ringer's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use

thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

[0236] A pharmaceutical composition may be formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (e.g., topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application may include one or more of the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA); buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH may be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic. In some, any of the activatable molecules described herein are prepared with carriers that protect against rapid elimination from the body, e.g., sustained and controlled release formulations, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, e.g., ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, polylactic-co-glycolic acid, and polylactic acid. Methods for preparation of such pharmaceutical compositions and formulations are apparent to those skilled in the art. For example, the activatable molecules may be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions.

[0237] Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the substrate-containing polypeptides, which matrices are in the form of shaped articles, e.g., films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides, copolymers of L-glutamic acid and y ethyl-L-glutamate,

non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers (e.g., injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(−)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods.

[0238] In some embodiments, pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor® EL (CAS No. 61791-12-6) (BASF, Parsippany, N.J.), which is a mixture of polyoxyethylated triglycerides, by reacting castor oil with ethylene oxide in a molar ratio of 1 : 35, that acts as a nonionic surfactant, or phosphate buffered saline (PBS). The composition may be sterile and should be fluid and of a viscosity that facilitates easy syringeability. It may be stable under the conditions of manufacture and storage and preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier may be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. For dispersed particulate compositions, the proper fluidity can be maintained, for example, by the use of a coating on the particles such as lecithin, and by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. In some embodiments, the pharmaceutical compositions may further comprise one or more antibacterial and/or antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In some embodiments, isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, and the like, as well as salts, such as, for example, sodium chloride and the like may be included in the composition. Prolonged absorption of the injectable compositions may be brought about by including in the composition an agent that delays absorption, for example, aluminum monostearate and gelatin.

[0239] In some embodiments, the pharmaceutical composition may comprise a sterile injectable solution. Sterile injectable solutions may be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions may be prepared by incorporating the active compound into a sterile vehicle that

contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

[0240] In some embodiments, the pharmaceutical composition may comprise an oral composition. Oral compositions may include an inert diluent or an edible carrier. They may be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound may be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions may also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials may be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primojel®(sodium starch glycolate), or corn starch; a lubricant such as magnesium stearate; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

[0241] In some embodiments, the pharmaceutical composition may be formulized for administration by inhalation. For example, the compounds may be delivered in the form of an aerosol spray from pressured container or dispenser that contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

[0242] In some embodiments, the pharmaceutical composition may be formulized for systemic administration. For example, systemic administration may be by intravenous, as well by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated may be used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration may be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds may be formulated into ointments, salves, gels, or creams as generally known in the art.

[0243] In some embodiments, the pharmaceutical composition may be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

[0244] In one embodiment, the pharmaceutical composition may be prepared with carriers that protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers may be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, polylactic-co-glycolic acid and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art.

[0245] It may be advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the disclosure may be dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

[0246] In some embodiments, the compositions (e.g., pharmaceutical compositions) may be included in a container, vial, syringe, injector pen, pack, or dispenser, optionally together with instructions for administration.

[0247] Also provided herein are kits that include any of the substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides) described herein, any of the compositions that include any of the polypeptides described herein, or any of the pharmaceutical compositions that include any of the polypeptides described herein. Also provided are kits that include one or more second therapeutic agent(s) in addition to a polypeptide described herein. The second therapeutic agent(s) may be provided in a dosage administration form that is separate from the polypeptides herein. Alternatively, the second therapeutic agent(s) may be formulated together with the polypeptides herein.

[0248] Any of the kits described herein can include instructions for using any of the compositions (e.g., pharmaceutical compositions) and/or any of the substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides) described herein. In

some embodiments, the kits can include instructions for performing any of the methods described herein. In some embodiments, the kits can include at least one dose of any of the compositions (e.g., pharmaceutical compositions) described herein. In some embodiments, the kits can provide a syringe for administering any of the pharmaceutical compositions described herein.

[0249] Also provided herein are substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides) produced by any of the methods described herein. Also provided are compositions (e.g., pharmaceutical compositions) that comprise any of the polypeptides produced by any of the methods described herein. Also provided herein are kits that include at least one dose of any of the compositions (e.g., pharmaceutical compositions) described herein.

METHODS OF PRODUCING SUBSTRATE-CONTAINING POLYPEPTIDES

[0250] Provided herein are methods of producing the substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides) described herein that include: (a) culturing any of the recombinant host cells described herein in a liquid culture medium under conditions sufficient to produce the substrate-containing polypeptides; and (b) recovering the substrate-containing polypeptides from the host cell and/or the liquid culture medium.

[0251] Methods of culturing cells are well known in the art. In some embodiments, cells may be maintained in vitro under conditions that favor cell proliferation, cell differentiation and cell growth. For example, the recombinant cells may be cultured by contacting a cell (e.g., any of the cells described herein) with a cell culture medium that includes the necessary growth factors and supplements sufficient to support cell viability and growth.

[0252] In some embodiments, the method may further include isolating the recovered substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides). The isolation of the substrate-containing polypeptides may be performed using any separation or purification technique for separating protein species, e.g., affinity tag-based protein purification (e.g., polyhistidine (His) tag, glutathione-S-transferase tag, and the like), ammonium sulfate precipitation, polyethylene glycol precipitation, size exclusion chromatography, ligand-affinity chromatography (e.g., Protein A chromatography), ion-exchange chromatography (e.g., anion or cation), hydrophobic interaction chromatography, and the like.

[0253] Compositions and methods described herein may involve use of non-reducing or partially-reducing conditions that allow disulfide bonds to form between the MM and the AM of the activatable molecules.

[0254] In some embodiments, the method further includes formulating the isolated polypeptides into a pharmaceutical composition. Various formulations are known in the art and are described herein. Any isolated polypeptides described herein can be formulated for any route of administration (e.g., intravenous, intratumoral, subcutaneous, intradermal, oral (e.g., inhalation), transdermal (e.g., topical), transmucosal, or intramuscular).

METHODS OF USING SUBSTRATE-CONTAINING POLYPEPTIDES

[0255] In some aspects, the present disclosure further provides methods of using the substrate-containing polypeptides herein. In some embodiments, the present disclosure provides methods of the treating a disease (e.g., a cancer (e.g., any of the cancers described herein)) in a subject including administering a therapeutically effective amount of any of the polypeptides (e.g., activatable molecules or conjugated polypeptides) described herein to the subject. In some embodiments, the disclosure provides methods of preventing, delaying the progression of, treating, alleviating a symptom of, or otherwise ameliorating disease in a subject by administering a therapeutically effective amount of an polypeptides (e.g., activatable molecules or conjugated polypeptides) described herein to a subject in need thereof. The term "treatment" refers to ameliorating at least one symptom of a disorder. In some embodiments, the disorder being treated may be a cancer or autoimmune disease or to ameliorate at least one symptom of a cancer or autoimmune disease. As used herein, the term "subject" refers to any mammal. In some embodiments, the subject is a feline (e.g., a cat), a canine (e.g., a dog), an equine (e.g., a horse), a rabbit, a pig, a rodent (e.g., a mouse, a rat, a hamster or a guinea pig), a non-human primate (e.g., a simian (e.g., a monkey (e.g., a baboon, a marmoset), or an ape (e.g., a chimpanzee, a gorilla, an orangutan, or a gibbon)), or a human. In some embodiments, the subject is a human. The terms subject and patient are used interchangeably herein. In some embodiments, the subject has been previously identified or diagnosed as having the disease (e.g., cancer (e.g., any of the cancers described herein)).

[0256] A therapeutically effective amount of a substrate-containing polypeptide (e.g., activatable molecule or conjugated polypeptide) of the disclosure relates generally to the amount needed to achieve a therapeutic objective. As noted above, this may be a binding interaction between the AM and its target that, in certain cases, interferes with the functioning of the targets. The amount required to be administered will furthermore depend on the binding

affinity of the polypeptides for its specific target, and will also depend on the rate at which an administered polypeptide is depleted from the free volume other subject to which it is administered. Common ranges for therapeutically effective dosing of a polypeptides of the disclosure may be, by way of non-limiting example, from about 0.001, 0.01, 0.1, 0.3, 0.5, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, or 50 mg/kg body weight or higher. The structure of the polypeptides of the present disclosure makes it possible to reduce the dosage of the polypeptide that is administered to a subject compared to conventional activatable molecules and compared to conventional antibodies. For example, the administered dose on a unit dosage basis or total dosage over a dosage regimen period may be reduced by 10, 20, 30, 40, or 50% compared to the corresponding dose of a corresponding conventional therapeutic molecules.

[0257] Common dosing frequencies may range, for example, from once or twice daily, weekly, biweekly, or monthly.

[0258] Efficaciousness of treatment is determined in association with any known method for diagnosing or treating the particular disorder. Methods for the screening of polypeptides that possess the desired specificity include, but are not limited to, enzyme linked immunosorbent assay (ELISA) and other immunologically mediated techniques known within the art.

[0259] In another embodiment, a polypeptide directed two or more targets are used in methods known within the art relating to the localization and/or quantitation of the targets (e.g., for use in measuring levels of one or more of the targets within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, a polypeptide directed two or more targets, or a derivative, fragment, analog or homolog thereof, that contain the antibody derived antigen binding domain, are utilized as pharmacologically active compounds (referred to hereinafter as "Therapeutics").

[0260] The substrate-containing polypeptides used in any of the embodiments of these methods and uses may be administered at any stage of the disease. For example, such a polypeptide may be administered to a patient suffering cancer of any stage, from early to metastatic. In some embodiments, the substrate-containing polypeptides and formulations thereof may be administered to a subject suffering from or susceptible to a disease or disorder associated with aberrant target expression and/or activity.

[0261] A subject suffering from or susceptible to a disease or disorder associated with aberrant target expression and/or activity may be identified using any of a variety of methods

known in the art. For example, subjects suffering from cancer or other neoplastic condition may be identified using any of a variety of clinical and/or laboratory tests such as, physical examination and blood, urine and/or stool analysis to evaluate health status. For example, subjects suffering from inflammation and/or an inflammatory disorder may be identified using any of a variety of clinical and/or laboratory tests such as physical examination and/or bodily fluid analysis, e.g., blood, urine and/or stool analysis, to evaluate health status.

[0262] In some embodiments, administration of a polypeptide to a patient suffering from a disease or disorder associated with aberrant target expression and/or activity may be considered successful if any of a variety of laboratory or clinical objectives is achieved. For example, administration of a polypeptide to a patient suffering from a disease or disorder associated with aberrant target expression and/or activity may be considered successful if one or more of the symptoms associated with the disease or disorder is alleviated, reduced, inhibited or does not progress to a further, i.e., worse, state. Administration of a polypeptide to a patient suffering from a disease or disorder associated with aberrant target expression and/or activity may be considered successful if the disease or disorder enters remission or does not progress to a further, i.e., worse, state.

[0263] As used herein, the term "treat" includes reducing the severity, frequency or the number of one or more (e.g., 1, 2, 3, 4, or 5) symptoms or signs of a disease (e.g., a cancer (e.g., any of the cancers described herein)) in the subject (e.g., any of the subjects described herein). In some embodiments where the disease is cancer, treating results in reducing cancer growth, inhibiting cancer progression, inhibiting cancer metastasis, or reducing the risk of cancer recurrence in a subject having cancer.

[0264] In some embodiments, the substrate comprises a substrate for a protease that is active, *e.g.*, upregulated or otherwise unregulated, in a disease condition or diseased tissue. Exemplary disease conditions include, for example, a cancer (*e.g.*, where the diseased tissue is a tumor tissue) and an inflammatory or autoimmune condition (*e.g.*, where the diseased tissue is inflamed tissue). In some embodiments, the substrate comprises a substrate for an extracellular protease. In some embodiments, the substrate comprises a substrate for an intracellular protease. In some embodiments, the substrate is an intracellular protease and an extracellular protease. In some embodiments, the disease may be a cancer. In some embodiments, the subject may have been identified or diagnosed as having a cancer. Examples of cancer include: solid tumor, hematological tumor, sarcoma, a leukemia (e.g., hairy cell leukemia, chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML),

chronic myeloid leukemia (CML), acute lymphocytic leukemia (ALL)), , stomach cancer, urothelial carcinoma, lung cancer, renal cell carcinoma, gastric and esophageal cancer, pancreatic cancer, prostate cancer, brain cancer, colon cancer, bone cancer, lung cancer, breast cancer, colorectal cancer, ovarian cancer, non-small cell lung carcinoma (NSCLC), squamous cell head and neck carcinoma, endometrial cancer, bladder cancer, cervical cancer, and liver cancer. Metastases of the aforementioned cancers may also be treated or prevented in accordance with the methods described herein.

[0265] In some embodiments, the disease may be an autoimmune disease or condition. In some embodiments, the subject may have been identified or diagnosed as having an autoimmune disease or condition or is at heightened risk of developing an autoimmune disease or condition. Examples of autoimmune diseases include Type 1 diabetes, Rheumatoid arthritis (RA), Psoriasis/psoriatic arthritis, Multiple sclerosis, Systemic lupus erythematosus, Inflammatory bowel disease (e.g., Crohn's disease, ulcerative colitis), chronic inflammation, or transplant rejection (e.g., in kidney, liver, or heart transplantation), autoimmune diseases, infectious disease, chronic inflammation, or transplant rejection. In some embodiments, the disease is a cardiovascular disorder. In some embodiments, the disease is a neurodegenerative disorder.

[0266] In some embodiments, the methods herein may result in a reduction in the number, severity, or frequency of one or more symptoms of the cancer in the subject (e.g., as compared to the number, severity, or frequency of the one or more symptoms of the cancer in the subject prior to treatment).

[0267] The methods may further comprise administering to a subject one or more additional agents. In some embodiments, the substrate-containing polypeptides (e.g., activatable molecules or conjugated polypeptides) may be administered during and/or after treatment in combination with one or more additional agents. In some embodiments, the polypeptide may be formulated into a single therapeutic composition, and the polypeptide and additional agent(s) may be administered simultaneously. Alternatively, the polypeptide and additional agent(s) may be separate from each other, e.g., each is formulated into a separate therapeutic composition, and the polypeptide and the additional agent are administered simultaneously, or the polypeptide and the additional agent are administered at different times during a treatment regimen. For example, the polypeptide may be administered prior to the administration of the additional agent, subsequent to the administration of the

additional agent, or in an alternating fashion. The polypeptide and additional agent(s) may be administered in single doses or in multiple doses.

[0268] One of more of the polypeptides herein may be co-formulated with, and/or co-administered with, one or more anti-inflammatory drugs, immunosuppressants, or metabolic or enzymatic inhibitors. In some embodiments, one or more polypeptides herein may be combined with one or more polypeptides of other types.

[0269] The present disclosure also provides methods of detecting presence or absence of a cleaving agent and/or the target in a subject or a sample. Such methods may comprise (i) contacting a subject or biological sample with an activatable molecule, wherein the activatable molecule includes a detectable label that is positioned on a portion of the activatable molecule that is released following cleavage of the substrate and (ii) measuring a level of activated molecule in the subject or biological sample, wherein a detectable level of activated molecule in the subject or biological sample indicates that the cleaving agent, the target or both the cleaving agent and the target are absent and/or not sufficiently present in the subject or biological sample, such that the target binding and/or protease cleavage of the activatable molecule cannot be detected in the subject or biological sample, and wherein a reduced detectable level of activated molecule in the subject or biological sample indicates that the cleaving agent and the target are present in the subject or biological sample.

[0270] Such detection methods may be adapted to also provide for detection of the presence or absence of a target that is capable of binding the AM of the activatable molecules when cleaved. Thus, the assays can be adapted to assess the presence or absence of a cleaving agent and the presence or absence of a target of interest. The presence or absence of the cleaving agent can be detected by the presence of and/or an increase in detectable label of the activatable molecules as described above, and the presence or absence of the target can be detected by detection of a target-AM complex *e.g.*, by use of a detectably labeled anti-target antibody.

[0271] In some embodiments, activatable molecules are also useful in *in situ* imaging for the validation of activatable molecule activation, *e.g.*, by protease cleavage, and binding to a particular target. *In situ* imaging is a technique that enables localization of proteolytic activity and target in biological samples such as cell cultures or tissue sections. Using this technique, it is possible to confirm both binding to a given target and proteolytic activity based on the presence of a detectable label (e.g., a fluorescent label).

[0272] These techniques are useful with any frozen cells or tissue derived from a disease site (*e.g.* tumor tissue) or healthy tissues. These techniques are also useful with fresh cell or tissue samples.

[0273] In these techniques, an activatable molecule may be labeled with a detectable label. The detectable label may be a fluorescent dye, (e.g. a fluorophore, Fluorescein Isothiocyanate (FITC), Rhodamine Isothiocyanate (TRITC), an Alexa Fluor® label), a near infrared (NIR) dye (e.g., Qdot® nanocrystals), a colloidal metal, a hapten, a radioactive marker, biotin and an amplification reagent such as streptavidin, or an enzyme (e.g. horseradish peroxidase or alkaline phosphatase).

[0274] Detection of the label in a sample that has been incubated with the labeled, activatable molecule indicates that the sample contains the target and contains a protease that is specific for the substrate of the activatable molecule. In some embodiments, the presence of the protease can be confirmed using broad spectrum protease inhibitors such as those described herein, and/or by using an agent that is specific for the protease, for example, an antibody such as A11, which is specific for the protease matriptase and inhibits the proteolytic activity of matriptase; see e.g., International Publication Number WO 2010/129609, published 11 November 2010. The same approach of using broad spectrum protease inhibitors such as those described herein, and/or by using a more selective inhibitory agent can be used to identify a protease that is specific for the substrate of the activatable molecule. In some embodiments, the presence of the target can be confirmed using an agent that is specific for the target, *e.g.*, another antibody, or the detectable label can be competed with unlabeled target. In some embodiments, unlabeled activatable molecule may be used, with detection by a labeled secondary antibody or more complex detection system.

[0275] Similar techniques are also useful for *in vivo* imaging where detection of the fluorescent signal in a subject, *e.g.*, a mammal, including a human, indicates that the disease site contains the target and contains a protease that is specific for the substrate of the activatable molecule.

[0276] These techniques are also useful in kits and/or as reagents for the detection, identification or characterization of protease activity in a variety of cells, tissues, and organisms based on the protease-specific substrate in the activatable molecule.

[0277] A reduced level of detectable label may be, for example, a reduction of at least 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or a reduction of substantially 100%. In some embodiments, the

detectable label may be conjugated to a component of the polypeptide, e.g., the AM. In some embodiments, measuring the level of polypeptide in the subject or sample may be accomplished using a secondary reagent that specifically binds the activated protein, wherein the reagent comprises a detectable label. The secondary reagent may be an antibody comprising a detectable label.

[0278] In some embodiments, the substrate-containing polypeptides may also be useful in the detection of the target in patient samples and accordingly are useful as diagnostics. For example, the polypeptides may be used in in vitro assays, e.g., ELISA, to detect target levels in a patient sample. For example, a polypeptide may be immobilized on a solid support (e.g., the well(s) of a microtiter plate). The immobilized polypeptide may serve as a capture protein for any target that may be present in a test sample. Prior to contacting the immobilized polypeptide with a patient sample, the solid support may be rinsed and treated with a blocking agent such as milk protein or albumin to prevent nonspecific adsorption of the analyte.

[0279] In some embodiments, based on the results obtained using the polypeptides in an in vitro diagnostic assay, the stage of a disease in a subject may be determined based on expression levels of the target protein (e.g., antigen). For a given disease, samples of blood may be taken from subjects diagnosed as being at various stages in the progression of the disease, and/or at various points in the therapeutic treatment of the disease. Using a population of samples that provides statistically significant results for each stage of progression or therapy, a range of concentrations of the target protein (e.g., antigen) that may be considered characteristic of each stage is designated.

[0280] Polypeptides herein may also be used in diagnostic and/or imaging methods. In some embodiments, such methods may be in vitro methods. In some embodiments, such methods may be in vivo methods. In some embodiments, such methods may be in situ methods. In some embodiments, such methods may be ex vivo methods. For example, polypeptides having a substrate may be used to detect the presence or absence of an enzyme capable of cleaving the substrate. Such polypeptides may be used in diagnostics, which can include in vivo detection (e.g., qualitative or quantitative) of enzyme activity (or, in some embodiments, an environment of increased reduction potential such as that which can provide for reduction of a disulfide bond) through measured accumulation of activated antibodies (i.e., antibodies resulting from cleavage of a polypeptide) in a given cell or tissue of a given host organism. Such accumulation of activated proteins indicates not only that the tissue expresses enzymatic activity (or an increased reduction potential depending on the nature of the

substrate) but also that the tissue expresses target to which the activated protein binds. In some examples, the polypeptides may be used for detecting protease activity with an assay that does not rely on target binding, e.g., a quantitative ex vivo zymography (QZ) assay as described in Howng et al., "Novel Ex Vivo Zymography Approach for Assessment of Protease Activity in Tissues with Activatable Antibodies," Pharmaceutics. 2021 Sep 2; 13(9):1390, which is incorporated by reference herein in its entirety.

[0281] For example, the substrate may be selected to be a protease substrate for a protease found at the site of a tumor, at the site of a viral or bacterial infection at a biologically confined site (e.g., such as in an abscess, in an organ, and the like), and the like. The AM may be one that binds a target protein (e.g., antigen). Using methods familiar to one skilled in the art, a detectable label (e.g., a fluorescent label or radioactive label or radiotracer) may be conjugated to an AM or other region of a polypeptide. Suitable detectable labels may be discussed in the context of the above screening methods and additional specific examples are provided below. Using an AM specific to a protein or peptide of the disease state, along with a protease whose activity is elevated in the disease tissue of interest, polypeptides may exhibit an increased rate of binding to disease tissue relative to tissues where the substrate specific enzyme is not present at a detectable level or is present at a lower level than in disease tissue or is inactive (e.g., in zymogen form or in complex with an inhibitor). Since small proteins and peptides are rapidly cleared from the blood by the renal filtration system, and because the enzyme specific for the substrate is not present at a detectable level (or is present at lower levels in non-disease tissues or is present in inactive conformation), accumulation of activated protein in the disease tissue may be enhanced relative to non-disease tissues.

[0282] In some embodiments, the substrate-containing polypeptides may be useful for in vivo imaging where detection of the fluorescent signal in a subject, e.g., a mammal, including a human, indicates that the disease site contains the target and contains a protease that is specific for the substrate of the polypeptide. The in vivo imaging may be used to identify or otherwise refine a patient population suitable for treatment with a polypeptide of the disclosure. For example, patients that test positive for both the target and a protease that cleaves the substrate in the substrate of the polypeptide being tested (e.g., accumulate activated proteins at the disease site) are identified as suitable candidates for treatment with such a polypeptide comprising such a substrate. Likewise, patients that test negative may be identified as suitable candidates for another form of therapy (i.e., not suitable for treatment with the polypeptide being tested). In some embodiments, such patients that test negative with

respect to a first polypeptide can be tested with other polypeptides comprising different substrates until a suitable polypeptide for treatment is identified (e.g., a polypeptide comprising a substrate that is cleaved by the patient at the site of disease).

[0283] In some embodiments, in situ imaging may be useful in methods to identify which patients to treat. For example, in in situ imaging, the polypeptides may be used to screen patient samples to identify those patients having the appropriate protease(s) and target(s) at the appropriate location, e.g., at a tumor site. In some embodiments, in situ imaging is used to identify or otherwise refine a patient population suitable for treatment with a polypeptide of the disclosure. For example, patients that test positive for both the target and a protease that cleaves the substrate in the substrate of the polypeptide being tested (e.g., accumulate activated antibodies at the disease site) are identified as suitable candidates for treatment with such a polypeptide comprising such a substrate. Likewise, patients that test negative for either or both of the target and the protease that cleaves the substrate used in the polypeptide being tested using these methods are identified as suitable candidates for another form of therapy (i.e., not suitable for treatment with the polypeptide being tested). In some embodiments, such patients that test negative with respect to a first polypeptide can be tested with other polypeptides comprising different substrates until a suitable polypeptide for treatment is identified (e.g., a polypeptide comprising a substrate that is cleaved by the patient at the site of disease).

[0284] The present application also provides aspects and embodiments as set forth in the following numbered Statements:

[0285] Statement 1. An isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the CM1 comprises the amino acid sequence of HQSR (SEQ ID NO: 1) or HQSK (SEQ ID NO: 3). In certain aspects, the present disclosure may include substitution of any alanine in the disclosed CM sequences with a valine. In some aspects, the isolated polypeptide comprises a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the CM1 comprises the amino acid sequence of AQSRSGPWGLL (SEQ ID NO: 641), HASRSGPWGLL (SEQ ID NO: 642), HQARSGPWGLL (SEQ ID NO: 643), HQSASGPWGLL (SEQ ID NO: 644). According to some embodiments of the present disclosures, the isolated polypeptide is a molecule in which cleavage of the CM by a protease results in a part or component of the molecule being

separated from the remainder of the molecule. In some aspects of the present disclosure, cleavage of the CM by a protease activates the molecule. In some aspects, the isolated polypeptide is a molecule in which multiple proteases cleave the CM. In some aspects, the isolated polypeptide is a molecule in which MMP2 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which MMP9 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which MMP14 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which MT-SP1 cleaves the CM. In some aspects, the isolated polypeptide is a molecule in which two or all of MMP2, MMP9, MMP14, MT-SP1 cleave the CM. In some aspects, the isolated polypeptide is a molecule in which the % cleavability of the CM is at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100%, e.g., at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, or 100% cleavable by any one of MMP2, MMP9, MMP14, MT-SP1 or any two of MMP2, MMP9, MMP14, MT-SP1 or each of MMP2, MMP9, MMP14, MT-SP1. In some aspects, the isolated polypeptide is a molecule in which the % cleavability of the CM is improved by 1.5x, 2x, 2.5x, 3x, 5x, 7x, 8x, or 10x or more over the % cleavability of SEQ ID NO: 78 (see, e.g., Example 2). According to some embodiments of the present disclosures, the isolated polypeptide is a molecule that has high *in vivo* stability such that it is not cleaved in plasma as demonstrated by less than 50%, less than 40%, or less than 25% in vivo activation following 7 days of administration *in vivo* (see, e.g., Example 3). According to embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^2 M^{-1} s^{-1}. According to some embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^3 M^{-1} s^{-1}. According to some embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^4 M^{-1} s^{-1}. According to some embodiments of the present disclosures, the isolated polypeptide is a molecule comprising a CM that has a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^5 M^{-1} s^{-1}.

[0286] Statement 2. The isolated polypeptide of Statement 1, wherein the CM1 comprises the amino acid sequence of HQSRS (SEQ ID NO: 2).

[0287] Statement 3. The isolated polypeptide of Statement 1, wherein the CM1 comprises the amino acid sequence of HQSRSA (SEQ ID NO: 4).

[0288] Statement 4. The isolated polypeptide of Statement 1, wherein the CM1 comprises the amino acid sequence of DHQSRS (SEQ ID NO: 5).

[0289] Statement 5. The isolated polypeptide of Statement 2, wherein the CM1 comprises the amino acid sequence of SDHQSRS (SEQ ID NO: 6).

[0290] Statement 6. The isolated polypeptide of any one of Statements 1-5, wherein the CM2 comprises the amino acid sequence of AIAL (SEQ ID NO: 9).

[0291] Statement 7. The isolated polypeptide of Statement 1-5, wherein the CM2 comprises the amino acid sequence of AIALY (SEQ ID NO: 635).

[0292] Statement 8. The isolated polypeptide of any one of Statements 1-5, wherein the CM2 comprises the amino acid sequence of AIALYA (SEQ ID NO: 8).

[0293] Statement 9. The isolated polypeptide of Statement 8, wherein the CM2 comprises the amino acid sequence of AIALYAD (SEQ ID NO: 7).

[0294] Statement 10. The isolated polypeptide of Statement 1, wherein the substrate comprises the amino acid sequence of AIALYADHQSRS (SEQ ID NO: 10), DHQSRSAIALYAD (SEQ ID NO: 11), HQSRSAIALYAD (SEQ ID NO: 12), DHQSRSAIALYA (SEQ ID NO: 13), HQSRSAIALY (SEQ ID NO: 636), HQSRSAIAL (SEQ ID NO: 637), AIALYAHQSRS (SEQ ID NO: 33), AIALYHQSRS (SEQ ID NO: 37), HQSRSAIALY (SEQ ID NO: 636), HQSRSAIAL (SEQ ID NO: 637), AIALHQSRS (SEQ ID NO: 41), HQSRSGPRGLL (SEQ ID NO: 638), HQSRSGPAGLL (SEQ ID NO: 639), HQSRSPWG (SEQ ID NO: 640), HQSRAGPWGLL (SEQ ID NO: 645), HQSRSAPWGLL (SEQ ID NO: 646), HQSRSGAWGLL (SEQ ID NO: 647), HQSRSGPWALL (SEQ ID NO: 648), HQSRSGPWGAL (SEQ ID NO: 649), HQSRSGPWGLA (SEQ ID NO: 650), HQSRPWGL (SEQ ID NO: 134), or HQSRSAIALYA (SEQ ID NO: 14) , optionally wherein the substrate comprises the amino acid sequence of AIALYADHQSRS (SEQ ID NO: 10), DHQSRSAIALYAD (SEQ ID NO: 11), HQSRSAIALYAD (SEQ ID NO: 12), DHQSRSAIALYA (SEQ ID NO: 13), or HQSRSAIALYA (SEQ ID NO: 14).

[0295] Statement 11. The isolated polypeptide of Statement 1, wherein the substrate comprises the amino acid sequence of DHQSRSAIALYA (SEQ ID NO: 13).

[0296] Statement 12. The isolated polypeptide of any one of Statements 1-5, wherein the CM2 comprises the amino acid sequence of GPWGLL (SEQ ID NO: 237).

[0297] Statement 13. The isolated polypeptide of Statement 12, wherein the substrate comprises the amino acid sequence of HQSRSGPWGLL (SEQ ID NO: 15), DHQSRSGPWGLL (SEQ ID NO: 16), or GPWGLLDHQSRS (SEQ ID NO: 17).

[0298] Statement 14. The isolated polypeptide of any one of Statements 1-5, wherein the CM2 comprises the amino acid sequence of PXGL, wherein X is alanine (A), tyrosine (Y),

arginine (R), phenylalanine (F), tryptophan (W), lysine (K), glutamine (Q), or methionine (M) or the CM2 comprises the amino acid sequence GPX_1GLX_2, (SEQ ID NO: 651) wherein X_1 is W, Y, F, R, K, Q, A, or M and wherein X_2 is any amino acid or absent.

[0299] Statement 15. The isolated polypeptide of Statement 14, wherein the substrate comprises the amino acid sequence of HQSRSPWGL (SEQ ID NO: 126), PWGLHQSRS (SEQ ID NO: 489), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSR (SEQ ID NO: 502), PWGLHQSKS (SEQ ID NO: 490), DHQSKSGPWGLL (SEQ ID NO: 178), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), or HQSRSGPAGLL (SEQ ID NO: 639).

[0300] Statement 16. The isolated polypeptide of any one of Statements 1-5, wherein the CM2 comprises the amino acid sequence of SGLLSS (SEQ ID NO: 238).

[0301] Statement 17. The isolated polypeptide of Statement 16, wherein the CM2 comprises the amino acid sequence of ISSGLLSS (SEQ ID NO: 75).

[0302] Statement 18. The isolated polypeptide of Statement 16, wherein the substrate comprises the amino acid sequence of ISSGLLSSDHQSRS (SEQ ID NO: 18), DHQSRSISSGLLSS (SEQ ID NO: 19), or DHQSRSSGLLSS (SEQ ID NO: 20).

[0303] Statement 19. The isolated polypeptide of any one of Statements 1-5, wherein the CM2 comprises the amino acid sequence of ALAHGLF (SEQ ID NO: 74).

[0304] Statement 20. The isolated polypeptide of Statement 19, wherein the substrate comprises the amino acid sequence of ALAHGLFSDHQSRS (SEQ ID NO: 21) or SDHQSRSALAHGLF (SEQ ID NO: 22).

[0305] Statement 21. An isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the substrate comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650, optionally wherein the substrate comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, and 460-634.

[0306] Statement 22. The isolated polypeptide of any one or combination of Statements 1-9, 12, 14, 16, 17, and 19, wherein the CM1 and the CM2 has a structural arrangement from N-terminus to C-terminus as CM1-CM2.

[0307] Statement 23. The isolated polypeptide of any one or combination of Statements 1-9, 12, 14, 16, 17 and 19, wherein the CM1 and the CM2 has a structural arrangement from N-terminus to C-terminus as CM2-CM1.

[0308] Statement 24. The isolated polypeptide of Statement 22 or 23, wherein the CM1 and the CM2 are indirectly coupled via a linking peptide.

[0309] Statement 25. The isolated polypeptide of Statement 22 or 23, wherein the CM1 and the CM2 are directly coupled to each other.

[0310] Statement 26. The isolated polypeptide of any one or combination of Statements 1-9, 12, 14, 16, 17 and 19, wherein at least a portion of the CM1 overlaps with at least a portion of the CM2.

[0311] Statement 26. An isolated polypeptide comprising a cleavable moiety (CM) comprising an amino acid sequence with one-amino acid mutation in any one of SEQ ID NOs: 10-73, 83-216, 460-634 and 636-650, wherein the CM is a substrate for a protease.

[0312] Statement 28. An isolated polypeptide comprising a cleavable moiety (CM) comprising an amino acid sequence with two-amino acid mutations in any one of SEQ ID NOs: 10-73, 83-216, 460-634 and 636-650, wherein the CM is a substrate for a protease.

[0313] Statement 29. An isolated polypeptide comprising a cleavable moiety (CM) comprising an amino acid sequence with three-amino acid mutations in any one of SEQ ID NOs: 10-73, 83-216, 460-634 and 636-650, wherein the CM is a substrate for a protease.

[0314] Statement 30. An isolated polypeptide comprising a cleavable moiety (CM) comprising an amino acid sequence with four-amino acid mutations in any one of SEQ ID NOs: 10-73, 83-216, 460-634 and 636-650, wherein the CM is a substrate for a protease.

[0315] Statement 31. The isolated polypeptide of any one or combination of Statements 1-30, wherein the substrate further comprises one or more additional CMs, optionally wherein at least a portion of a first CM overlaps with at least a portion of a second CM in the substrate, such that one or more amino acids belong to both CMs.

[0316] Statement 32. The isolated polypeptide of any one or combination of Statements 1-31, wherein the isolated polypeptide is an activatable molecule and further comprises an active moiety (AM) that specifically binds a target. According to embodiments of the present disclosures, the isolated polypeptide is an activatable molecule that has high *in vivo* stability such that it is not cleaved in plasma as demonstrated by less than 50%, less than 40%, or less than 25% in vivo activation following 7 days of administration *in vivo*. According to some embodiments of the present disclosures, the isolated polypeptide is an activatable molecule

that has masking efficiency of 18x, 20x, 30x, 33x, 50x, 60x, 82x, 95x, 105x, 121x, 123x, 150x, 200x, 234x, 300x, or higher (e.g., as exemplified in Example 4). According to some embodiments of the present disclosures, the activatable molecule is activated by one, two, or all of MMP2, MMP9, MMP14 and MT-SP1. According to some embodiments of the present disclosures, the activatable molecule exhibits attenuated binding to a target as compared to the binding of a counterpart "activated" molecule comprising the same active moiety (AM) to the same target.

[0317] Statement 33. The isolated polypeptide of Statement 32, wherein the AM is an antibody or antigen binding fragment thereof.

[0318] Statement 34. The isolated polypeptide of Statement 33, wherein the antigen binding fragment thereof is selected from the group consisting of a Fab fragment, a $F(ab')_2$ fragment, a scFv, a scAb, a dAb, a single domain heavy chain antibody, and a single domain light chain antibody.

[0319] Statement 35. The isolated polypeptide of Statement 32, wherein the AM is a therapeutic macromolecule.

[0320] Statement 36. The isolated polypeptide of Statement 32, wherein the AM is a cytokine or a chimeric antigen receptor.

[0321] Statement 37. The isolated polypeptide of any one or combination of Statements 32-36, wherein the AM is coupled to the CM1.

[0322] Statement 38. The isolated polypeptide of Statement 37, wherein the AM is directly coupled to the CM1.

[0323] Statement 39. The isolated polypeptide of Statement 37, wherein the AM is indirectly coupled to the CM1 via a linking peptide.

[0324] Statement 40. The isolated polypeptide of any one or combination of Statements 32-39, wherein the AM is coupled to the CM2.

[0325] Statement 41. The isolated polypeptide of Statement 40, wherein the AM is directly coupled to the CM2.

[0326] Statement 42. The isolated polypeptide of Statement 40, wherein the AM is indirectly coupled to the CM2 via a linking peptide.

[0327] Statement 43. The isolated polypeptide of any one or combination of Statements 32-42, further comprising a masking moiety (MM).

[0328] Statement 44. The isolated polypeptide of Statement 43, wherein the MM has a dissociation constant for binding to the AM that is greater than the dissociation constant of the AM for binding to the target.

[0329] Statement 45. The isolated polypeptide of Statement 43 or 44, wherein the MM is 2 to 40 amino acids in length.

[0330] Statement 46. The isolated polypeptide of any one or combination of Statements 43-45, wherein the MM interferes with AM's binding to its binding partner through non-specific interactions such as steric hindrance, optionally wherein the MM is positioned in the activatable molecule such that the tertiary or quaternary structure of the activatable molecule allows the MM to mask the AM through charge-based interaction, optionally wherein the MM is an albumin, e.g., human serum albumin (HSA), a fragment crystallizable (Fc) domain, an antibody constant domain (e.g., CH domains), a polymer (e.g., branched or multi-armed polyethylene glycol (PEG)), a latency associated protein (LAP), and any polypeptide or other moieties that sterically interfere AM-target interactions, optionally wherein the MM may recruit a large protein binding partner that sterically interfere AM-target interactions, optionally wherein the MM is an antibody or a fragment thereof that binds to an albumin, optionally wherein the MM comprises a full-length or a AM-binding fragment or mutein of a cognate receptor of the AM, and AM-binding antibodies and fragment thereof, e.g., a polyclonal antibody, a recombinant antibody, a human antibody, a humanized antibody a single chain variable fragment (scFv), single-domain antibody such as a heavy chain variable domain (VH), a light chain variable domain (VL), a variable domain of camelid-type nanobody (VHH), or a dAb, optionally wherein the MM is a non-immunoglobulin proteins that mimic antibody binding and/or structure such as, anticalins, affilins, affibody molecules, affimers, affitins, alphabodies, avimers, DARPins, fynomers, kunitz domain peptides, monobodies, and binding domains based on other engineered scaffolds such as SpA, GroEL, fibronectin, lipocallin and CTLA4 scaffolds, optionally wherein the MM is a peptide that is modified by conjugation to a water-soluble polymer, such as a polyalkylene glycol, e.g., a polyethylene glycol (PEG), optionally wherein the MM is an antibody or antigen-binding domain that binds to a protein with a long serum half- life such as HSA, immunoglobulin or transferrin, or to a receptor that is recycled to the plasma membrane, such as FcRn or a transferrin receptor.

[0331] Statement 47. The isolated polypeptide of any one or combination of Statements 43-46, wherein the MM is linked to the substrate such that the isolated polypeptide comprises

a structural arrangement from N-terminus to C-terminus as follows: MM-CM1-CM2-AM, MM-CM2-CM1-AM, AM-CM1-CM2-MM, or AM-CM2-CM1-MM.

[0332] Statement 48. The isolated polypeptide of Statement 47, wherein the MM is directly coupled to the CM1.

[0333] Statement 49. The isolated polypeptide of Statement 50, wherein the MM is indirectly coupled to the CM1 via a linking peptide.

[0334] Statement 50. The isolated polypeptide of Statement 48, wherein the MM is directly coupled to the CM2.

[0335] Statement 51. The isolated polypeptide of Statement 50, wherein the MM is indirectly coupled to the CM2 via a linking peptide.

[0336] Statement 52. The isolated polypeptide of any one or combination of Statements 43-51, wherein the isolated polypeptide comprises a linking peptide (LP), and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM-LP-CM1-CM2-AM, MM-LP-CM2-CM1-AM, AM-CM1-CM2-LP-MM, or AM-CM2-CM1-LP-MM, MM-CM1-CM2-LP-AM, MM-CM2-CM1-LP-AM, AM-LP-CM1-CM2 - MM, or AM-LP-CM2-CM1-MM.

[0337] Statement 53. The isolated polypeptide of any one or combination of Statements 43-52, wherein the isolated polypeptide comprises a first linking peptide (LP1) and a second linking peptide (LP2), and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM-LP1-CM1-CM2-LP2-AM, MM-LP1-CM2-CM1-LP2-AM, AM-LP2-CM1-CM2-LP1-MM, AM-LP2-CM2-CM1-LP1-MM, MM- LP1-CM1-LP3-CM2-LP2-AM, MM-LP1-CM2-LP3-CM1-LP2-AM, AM-LP2-CM1-LP3-CM2-LP1-MM, or AM-LP2-CM2-LP3-CM1-LP1-MM.

[0338] Statement 54. The isolated polypeptide of Statement 53, wherein the LP1 and LP2 are not identical to each other.

[0339] Statement 55. The isolated polypeptide of Statement 53, wherein the LP1 and LP2 are identical to each other.

[0340] Statement 56. The isolated polypeptide of any one of Statements 53-55, wherein each of the LP1 and LP2 is a peptide of 1 to 20 amino acids in length.

[0341] Statement 57. The isolated polypeptide of any one or combination of Statements 1-56, wherein the first protease is a serine protease.

[0342] Statement 58. The isolated polypeptide of Statement 57, wherein the serine protease is membrane type serine protease 1 (MT-SP1).

[0343] Statement 59. The isolated polypeptide of Statement 58, wherein the k_{cat}/K_M of the substrate by MT-SP1 cleavage is at least 1×10^3 $M^{-1}s^{-1}$.

[0344] Statement 60. The isolated polypeptide of Statement 58, wherein the k_{cat}/K_M of the substrate by MT-SP1 cleavage is at least 1×10^4 $M^{-1}s^{-1}$.

[0345] Statement 61. The isolated polypeptide of any one or combination of Statements 1-60, wherein the second protease is a matrix metalloproteinase (MMP). The isolated polypeptide of any one or combination of Statements 1-32, wherein the CM is a substrate for a MT-SP1.

[0346] Statement 62. The isolated polypeptide of Statement 60, wherein the MMP is MMP2, MMP9, or MMP14. The isolated polypeptide of Statement 33, wherein the MMP is MMP2. The isolated polypeptide of Statement 33, wherein the MMP is MMP9. The isolated polypeptide of Statement 33, wherein the MMP is MMP14.

[0347] Statement 63. The isolated polypeptide of Statement 62, wherein the k_{cat}/K_M of the substrate by MMP2 cleavage is at least 1×10^3 $M^{-1}s^{-1}$, optionally at 37°C in 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35.

[0348] Statement 64. The isolated polypeptide of Statement 62, wherein the k_{cat}/K_M of the substrate by MMP2 cleavage is at least 1×10^4 $M^{-1}s^{-1}$, optionally at 37°C in 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35.

[0349] Statement 65. The isolated polypeptide of any one of Statements 62-64, wherein the k_{cat}/K_M of the substrate by MMP9 cleavage is at least 1×10^3 $M^{-1}s^{-1}$, optionally at 37°C in 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35.

[0350] Statement 66. The isolated polypeptide of any one of Statements 62-64, wherein the k_{cat}/K_M of the substrate by MMP9 cleavage is at least 1×10^4 $M^{-1}s^{-1}$, optionally at 37°C in 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35.

[0351] Statement 67. The isolated polypeptide of any one of Statements 62-66, wherein the k_{cat}/K_M of the substrate by MMP14 cleavage is at least 1×10^3 $M^{-1}s^{-1}$, optionally at 37°C in 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35.

[0352] Statement 68. The isolated polypeptide of any one of Statements 62-66, wherein the k_{cat}/K_M of the substrate by MMP14 cleavage is at least 1×10^4 $M^{-1}s^{-1}$, optionally at 37°C in 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35.

[0353] Statement 69. The isolated polypeptide of any one or combination of Statements 1-68, wherein the isolated polypeptide is resistant to cleavage *in situ* in human bone marrow, optionally resistant to cleavage *in situ* in human bone marrow compared to a CM of SEQ ID

NO: 78, optionally resistant to cleavage *in situ* in human bone marrow compared to a CM of SEQ ID NO: 79, or optionally resistant to cleavage *in situ* in human bone marrow compared to a CM of SEQ ID NO: 80.

[0354] Statement 70 The isolated polypeptide of any one or combination of Statements 1-68, wherein the substrate is resistant to cleavage *in vivo* in human bone marrow, optionally resistant to cleavage *in situ* in human bone marrow compared to a CM of SEQ ID NO: 78, optionally resistant to cleavage *in situ* in human bone marrow compared to a CM of SEQ ID NO: 79, or optionally resistant to cleavage *in situ* in human bone marrow compared to a CM of SEQ ID NO: 80.

[0355] Statement 71. The isolated polypeptide of any one or combination of Statements 32-70, wherein the AM is an antibody or antigen-binding fragment that binds EGFR and the MM comprises the amino acid sequence of SEQ ID NO: 82.

[0356] Statement 72. An isolated polypeptide comprising an antibody or antigen-binding fragment thereof that binds EGFR (AB), a masking moiety (MM) comprising the SEQ ID NO: 82, and a cleavable moiety (CM), wherein AB is coupled with the MM via the CM.

[0357] Statement 73. A polypeptide complex comprising one or more of the isolated polypeptides of any one or combination of Statements 1-72.

[0358] Statement 74. A conjugated polypeptide comprising the isolated polypeptide of any one or combination of Statements 1-72 conjugated to an agent.

[0359] Statement 75. The conjugated polypeptide of Statement 74, wherein the agent is conjugated to the isolated polypeptide via a conjugating linker.

[0360] Statement 76. The conjugated polypeptide of Statement 75, wherein the conjugating linker is cleavable.

[0361] Statement 77. The conjugated polypeptide of Statement 75, wherein the conjugating linker is non-cleavable.

[0362] Statement 78. The conjugated polypeptide of Statement 76, wherein the conjugating linker comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, and 636-650.

[0363] Statement 79. The conjugated polypeptide of any one or combination of Statements 74-78, wherein the agent is a toxin, a microtubule inhibitor, a nucleic acid damaging agent, a dolastatin, an auristatin, a maytansinoid, a duocarmycin, a calicheamicin, or a combination thereof.

[0364] Statement 80. A composition comprising the isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, and a carrier.

[0365] Statement 81. The composition of Statement 80, wherein the carrier is a pharmaceutically acceptable carrier.

[0366] Statement 82. The composition of Statements 81, comprising an additional agent.

[0367] Statement 83. The composition of Statement 82, wherein the additional agent is a therapeutic agent.

[0368] Statement 84. An isolated nucleic acid molecule encoding the isolated polypeptide of any one or combination of Statements 1-73.

[0369] Statement 85. A vector comprising the isolated nucleic acid molecule of Statement 84.

[0370] Statement 86. A cell comprising the isolated nucleic acid molecule of Statement 81 or the vector of Statement 85.

[0371] Statement 87. A method of manufacturing an activatable molecule that contains a tandem substrate, the method comprising expressing and recovering a polypeptide comprising the isolated polypeptide of any one or combination of Statements 1-73, optionally wherein the polypeptide is an activatable molecule.

[0372] Statement 88. A method of treating, alleviating a symptom of, or delaying the progression of a disease or disorder in a subject, comprising administering a therapeutically effective amount of the isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, or the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for use as a medicament or for use in therapy, optionally for treating a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder, optionally with an additional agent which is optionally a therapeutic agent. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination

of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for treating a cancer. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for treating an infection. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for treating an inflammatory disorder. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for treating a cardiovascular disorder. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for treating a neurodegenerative disorder. The isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, the composition of any one of Statements 80-83 to the subject, or the nucleic acid molecule of Statement 84, the vector of Statement 85, or the cell of Statement 86 for treating an autoimmune disorder.

[0373] Statement 89. The method of Statement 88, wherein the disease is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

[0374] Statement 90. A kit comprising the isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, or the composition of any one of Statements 08-83.

[0375] Statement 91. The use of the isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any

one or combination of Statements 75-79, or the composition of any one of Statements 80-83 for the manufacture of a medicament for the treatment of a disease or disorder.

[0376] Statement 92. The use of Statement 91, wherein the disease or disorder is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

[0377] Statement 93. A method of detecting or diagnosing a disease or health condition of a subject, comprising: contacting the isolated polypeptide of any one or combination of Statements 1-73, the polypeptide complex of Statement 74, the conjugated polypeptide of any one or combination of Statements 75-79, or the composition of any one of Statements 80-83 with a sample from the subject; and measuring a level of cleavage of the isolated polypeptide, thereby detecting or diagnosing the disease or health condition of the subject.

[0378] Statement 94. The method of Statement 93, wherein the disease is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

EXAMPLES

EXAMPLE 1: Exemplary Activatable Antibodies and Protease Cleavable Substrates

[0379] The studies provided herein describe exemplary substrates that include at least one cleavable moiety cleavable by a matrix metalloprotease (MMP), and at least one cleavable moiety cleavable by a matriptase (MT-SP1).

[0380] Exemplary activatable antibodies were constructed such that each one includes one of the substrates listed in Table 4. The exemplary activatable antibodies, the sequences of which are listed in Table 5, include an antibody or antigen binding fragment thereof (AB) that is based on a mouse/human chimeric monoclonal antibody that specifically binds to epidermal growth factor receptor (EGFR). The exemplary activatable antibodies also include a prodomain coupled to the N-terminus of the light chain of the AB. Each prodomain includes a masking moiety (MM) and a substrate, and the substrate includes at least one sequence of Table 4.

Table 4: Exemplary MMP / MT-SP1 Substrates

Name	Sequence	SEQ ID NO:
8000	AIALYADHQSRS	10
8001	DHQSRSAIALYAD	11
8002	HQSRSAIALYAD	12
8003	DHQSRSAIALYA	13

8004	HQSRSAIALYA	14
8005	HQSRSGPWGLL	15
8006	DHQSRSGPWGLL	16
8007	GPWGLLDHQSRS	17
8008	ISSGLLSSDHQSRS	18
8009	DHQSRSISSGLLSS	19
8010	DHQSRSSGLLSS	20
8011	ALAHGLFSDHQSRS	21
8012	SDHQSRSALAHGLF	22

Table 5: Activatable Antibody Sequences

Molecule Name	**Light chain EGFR Mask**	**Light chain substrate**	**Protein description (Light chain /Heavy chain)**
CX-122	CISPRGCPDGPYVMY (SEQ ID NO: 81)	2001 (SEQ ID NO: 78)	C225v5 3954 2001 (SEQ ID NO: 217 / SEQ ID NO: 218)
AA w/5007	CISPRGCPDGPYVMY (SEQ ID NO: 81)	5007 (SEQ ID NO: 80)	C225v5 3954 5007 (SEQ ID NO: 219 / SEQ ID NO: 220)
CX-229	CISPRGCLDGPYVMY (SEQ ID NO: 82)	3001 (SEQ ID NO: 79)	C225v5 3954 3001 (SEQ ID NO: 221 / SEQ ID NO: 218)
ProC070	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8000 (SEQ ID NO: 10)	C225v5 3954 8000 (SEQ ID NO: 222 / SEQ ID NO: 223)
ProC071	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8001 (SEQ ID NO: 11)	C225v5 3954 8001 (SEQ ID NO: 224 / SEQ ID NO: 223)
ProC1590	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8002 (SEQ ID NO: 12)	C225v5 3954 8002 (SEQ ID NO: 225 / SEQ ID NO: 218)
ProC1591	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8003 (SEQ ID NO: 13)	C225v5 3954 8003 (SEQ ID NO: 226 / SEQ ID NO: 218)
ProC1592	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8004 (SEQ ID NO: 14)	C225v5 3954 8004 (SEQ ID NO: 227 / SEQ ID NO: 218)
ProC1593	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8005 (SEQ ID NO: 15)	C225v5 3954 8005 (SEQ ID NO: 228 / SEQ ID NO: 218)
ProC905	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8006 (SEQ ID NO: 16)	C225v5 3954 8006 (SEQ ID NO: 229 / SEQ ID NO: 218)
ProC907	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8007 (SEQ ID NO: 17)	C225v5 3954 8007

			(SEQ ID NO: 230 / SEQ ID NO: 218)
ProC072	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8008 (SEQ ID NO: 18)	C225v5 3954 8008 (SEQ ID NO: 231 / SEQ ID NO: 223)
ProC073	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8009 (SEQ ID NO: 19)	C225v5 3954 8009 (SEQ ID NO: 232 / SEQ ID NO: 223)
ProC074	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8010 (SEQ ID NO: 20)	C225v5 3954 8010 (SEQ ID NO: 233 / SEQ ID NO: 223)
ProC075	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8011 (SEQ ID NO: 21)	C225v5 3954 8011 (SEQ ID NO: 234 / SEQ ID NO: 223)
ProC076	CISPRGCPDGPYVMY (SEQ ID NO: 81)	8012 (SEQ ID NO: 22)	C225v5 3954 8012 (SEQ ID NO: 235 / SEQ ID NO: 223)

EXAMPLE 2: *In Vitro* Cleavability of Activatable Antibodies with Exemplary Substrates

[0381] The studies provided herein evaluate the *in vitro* cleavability of activatable antibodies containing exemplary substrates that include at least one cleavable moiety cleavable by a matrix metalloprotease (MMP) and/or at least one cleavable moiety cleavable by matriptase (MT-SP1).

[0382] The stability (*e.g.*, cleavability) of the activatable antibodies were measured in the presence of the indicated recombinant proteases (MT-SP1, MMP2, MMP9, and MMP14). Each activatable antibody (500 nM) was incubated with 10 nM of the indicated protease for 1.5 hours at 37°C. Human recombinant proteases were purchased from R&D Systems: MMP2 (catalog No: 902-MP), MMP9 (catalog No: 911-MP), MMP14 (catalog No: 918-MP) and MT-SP1 (catalog No: 3946-SEB). MMPs were activated according to the manufacturer's instructions. Protease concentrations were determined by active site titration. Activity assays for MMP2 and MMP9 were performed in the following buffer: 50 mM Tris-HCl, 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35, pH 7.5. Activity assays were performed for MMP14 using 50 mM HEPES (pH 6.8), 10 mM $CaCl_2$, 0.5 mM $MgCl_2$, and for MT-SP1 activity assays used 50mM TRIS-HCl (pH 7.4), 150 mM NaCl, 0.05% Tween 20. Following incubation, the presence of cleavage product was determined by capillary electrophoresis for each protease enzyme using a LabChip GXII Touch system (Perkin Elmer), or by capillary electrophoresis immunoassay for each protease enzyme using the Wes™ Western Blot

protocol (Protein Simple). For capillary electrophoresis assays, the HT Protein Express 100 protocol (Perkin Elmer) was used. LabChip GXII Touch HT Chips (Perkin Elmer #760499) were set up using the protocol of the Protein Express Assay Reagent Kit (Perkin Elmer #CLS960008). The fraction of cleaved activatable antibody was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the LabChip GX Reviewer software (Perkin Elmer). For capillary electrophoresis immunoassays, the A110UK goat anti-human IgG antibody (American Qualex) and an anti-goat secondary antibody (Jackson ImmunoResearch) were used. The fraction of cleaved activatable antibody was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the Compass software (Protein Simple). The fraction of activatable antibody, and hence the CM that is cleaved by each particular protease, is presented as a "cleavability percentage" in Table 6A.

[0383] In addition, an exemplary study was performed to determine the cleavability kinetics (*i.e.*, k_{cat}/K_M (M^{-1} s^{-1})) of the indicated substrates with the indicated proteases. The exemplary results of this *in vitro* study are summarized in Table 6B.

[0384] These exemplary results show that the substrates showed a range of cleavability by the indicated proteases.

[0385] These exemplary results show that the substrates showed a range of cleavability by matriptase, MMP2, MMP9, and/or MMP14. These exemplary results also show a group of substrates in which the MT-SP1 is at least 25%. Such substrates include those having the amino acid sequence of 8000 (AIALYADHQSRS; SEQ ID NO: 10), 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8002 (HQSRSAIALYAD; SEQ ID NO: 12), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), 8004 (HQSRSAIALYA; SEQ ID NO: 14), 8005 (HQSRSGPWGLL; SEQ ID NO: 15), 8007 (GPWGLLDHQSRS; SEQ ID NO:17), 8008 (ISSGLLSSDHQSRS; SEQ ID NO: 18), 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19), 8010 (DHQSRSSGLLSS; SEQ ID NO: 20), 8011 (ALAHGLFSDHQSRS; SEQ ID NO: 21), or 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0386] These exemplary results also show a group of substrates in which the MT-SP1 cleavability is at least 60%. Such substrates include those having the amino acid sequence of 8002 (HQSRSAIALYAD; SEQ ID NO: 12), 8004 (HQSRSAIALYA; SEQ ID NO: 14), 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19), or 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0387] These exemplary results also show a group of substrates in which the MMP9 and MMP14 cleavability are both at least 50%. Such substrates include those having the amino acid sequence of 8005 (HQSRSGPWGLL; SEQ ID NO: 15) or 8006 (DHQSRSGPWGLL; SEQ ID NO: 16).

[0388] These exemplary results also show a group of substrates in which the MMP2 and MMP14 cleavability are both at least 60% and the matriptase cleavability is at least 25%. Such substrates include those having the amino acid sequence of 8011 (ALAHGLFSDHQSRS; SEQ ID NO: 21), or 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0389] These exemplary results also show a group of substrates in which the MMP2, MMP9 and MMP14 cleavability are all at least 30%. Such substrates include those having the amino acid sequence of 8006 (DHQSRSGPWGLL; SEQ ID NO: 16), or 8005 (HQSRSGPWGLL; SEQ ID NO: 15).

[0390] These exemplary results also show a group of substrates in which the MMP2 cleavability is at least 30%. Such substrates include those having the amino acid sequence of 8000 (AIALYADHQSRS; SEQ ID NO: 10), 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8002 (HQSRSAIALYAD; SEQ ID NO: 12), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), 8004 (HQSRSAIALYA; SEQ ID NO: 14), 8005 (HQSRSGPWGLL; SEQ ID NO: 15), 8006 (DHQSRSGPWGLL; SEQ ID NO: 16), 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19), 8011 (ALAHGLFSDHQSRS; SEQ ID NO: 21), or 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0391] These exemplary results also show a group of substrates in which the MMP2 cleavability is at least 30% and MT-SP1 cleavability is at least 60%. Such substrates include those having the amino acid sequence of 8002 (HQSRSAIALYAD; SEQ ID NO: 12), 8004 (HQSRSAIALYA; SEQ ID NO: 14), 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19) or 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0392] These exemplary results also show a group of substrates in which the MMP14 and MT-SP1 cleavability are both at least 30%. Such substrates include those having the amino acid sequence of 8010 (DHQSRSSGLLSS; SEQ ID NO: 20), 8008 (ISSGLLSSDHQSRS; SEQ ID NO: 18), 8005 (HQSRSGPWGLL; SEQ ID NO: 15), 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19), 8011 (ALAHGLFSDHQSRS; SEQ ID NO: 21), or 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0393] These exemplary results also show a group of substrates in which the MMP2, MMP14 and MT-SP1 cleavability are all at least 60%. Such substrates include those having the amino acid sequence of 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22).

[0394] These exemplary results also show a group of substrates in which the MMP2, MMP9, MMP14, and MT-SP1 cleavability are all at least 30%. Such substrates include those having the amino acid sequence of 8005 (HQSRSGPWGLL; SEQ ID NO: 15).

[0395] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^3 M^{-1} s^{-1} for *in vitro* cleavability with MT-SP1. Such substrates include those having the amino acid sequence of 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), 8006 (DHQSRSGPWGLL; SEQ ID NO: 16), or 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19).

[0396] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^4 $M^{-1}s^{-1}$ for *in vitro* cleavability with MT-SP1. Such substrates include those having the amino acid sequence of 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), or 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19).

[0397] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^3 M^{-1} $s^{-1}$for *in vitro* cleavability with MMP14. Such substrates include those having the amino acid sequence of 8003 (DHQSRSAIALYA; SEQ ID NO: 13), 8006 (DHQSRSGPWGLL; SEQ ID NO: 16), or 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19).

[0398] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^4 $M^{-1}s^{-1}$ for *in vitro* cleavability with MMP14. Such substrates include those having the amino acid sequence of 8006 (DHQSRSGPWGLL; SEQ ID NO: 16).

[0399] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^3 M^{-1} $s^{-1}$for *in vitro* cleavability with MMP9. Such substrates include those having the amino acid sequence of 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), or 8006 (DHQSRSGPWGLL; SEQ ID NO: 16).

[0400] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^4 $M^{-1}s^{-1}$ for *in vitro* cleavability with MMP9. Such substrates include those having the amino acid sequence of 8006 (DHQSRSGPWGLL; SEQ ID NO: 16).

[0401] These exemplary results of Table 6B also show that exemplary had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^3 M^{-1} $s^{-1}$for *in vitro* cleavability with MMP2. Such substrates include

those having the amino acid sequence of 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), 8006 (DHQSRSGPWGLL; SEQ ID NO: 16), or 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19).

[0402] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^4 $M^{-1}s^{-1}$ for *in vitro* cleavability with MMP2. Such substrates include those having the amino acid sequence of 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), or 8006 (DHQSRSGPWGLL; SEQ ID NO: 16).

[0403] These exemplary results of Table 6B also show that exemplary substrates had a k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^4 $M^{-1}s^{-1}$ for *in vitro* cleavability with MMP2, MMP9, and MMP14 and k_{cat}/K_M (M^{-1} s^{-1}) of greater than 1 x 10^3 $M^{-1}s^{-1}$ for *in vitro* cleavability with MT-SP1. Such substrates include those having the amino acid sequence of 8006 (DHQSRSGPWGLL; SEQ ID NO: 16).

Table 6A: *In Vitro* Activation of Activatable Antibodies with Exemplary Substrates

Exp No.	Substrate of Activatable Antibody	Substrate	Cleavability (%)			
			MT-SP1	MMP2	MMP9	MMP14
1	2001	ISSGLLSGRSDNH (SEQ ID NO: 78)	11	27	10	61
	8001	DHQSRSAIALYAD (SEQ ID NO: 11)	39	40	6	3
	8002	HQSRSAIALYAD (SEQ ID NO: 12)	73	53	13	8
	8003	DHQSRSAIALYA (SEQ ID NO: 13)	53	57	11	14
	8005	HQSRSGPWGLL (SEQ ID NO: 15)	45	50	78	64
	8006	DHQSRSGPWGLL (SEQ ID NO: 16)	19	42	69	64
	Substrate of Activatable Antibody	**Substrate**	**MT-SP1**	**MMP2**	**MMP9**	**MMP14**
2	2001	ISSGLLSGRSDNH (SEQ ID NO: 78)	12	33	-	56
	8001	DHQSRSAIALYAD (SEQ ID NO: 11)	52	53	13	9
	8004	HQSRSAIALYA (SEQ ID NO: 14)	80	49	18	16

	Substrate of Activatable Antibody	Substrate	MT-SP1	MMP2	MMP9	MMP14
3	2001	ISSGLLSGRSDNH (SEQ ID NO: 78)	11	25	-	67
	8000	AIALYADHQSRS (SEQ ID NO: 10)	36	75	9	8
	8001	DHQSRSAIALYAD (SEQ ID NO: 11)	71	61	7	3
	8008	ISSGLLSSDHQSRS (SEQ ID NO: 18)	39	25	-	88
	8009	DHQSRSISSGLLSS (SEQ ID NO: 19)	74	44	-	30
	8010	DHQSRSSGLLSS (SEQ ID NO: 20)	35	18	-	33
	8011	ALAHGLFSDHQSRS (SEQ ID NO: 21)	43	100	5	89
	8012	SDHQSRSALAHGLF (SEQ ID NO: 22)	89	89	8	73
	Substrate of Activatable Antibody	**Substrate**	**MT-SP1**	**MMP2**	**MMP9**	**MMP14**
4	2001	ISSGLLSGRSDNH (SEQ ID NO: 78)	9	25	-	58
	8006	DHQSRSGPWGLL (SEQ ID NO: 16)	21	67	89	66
	8007	GPWGLLDHQSRS (SEQ ID NO: 17)	27	26	24	70

Note: "-" indicates below limit of quantification.

Table 6B: *In Vitro* Activation of Activatable Antibodies with Exemplary Substrates

Substrate of Activatable Antibody	Substrate	k_{cat}/K_M ($M^{-1} s^{-1}$)				
		MT-SP1	uPA	MMP2	MMP9	MMP14
2001	ISSGLLSGRSDNH (SEQ ID NO: 78)	1.60 E+03	6.78 E+02	5.08 E+03	-	1.21 E+04
8001	DHQSRSAIALYAD (SEQ ID NO: 11)	1.68 E+04	-	1.35 E+04	1.03 E+03	9.16 E+02
8003	DHQSRSAIALYA (SEQ ID NO: 13)	1.94 E+04	-	1.79 E+04	1.55 E+03	1.98 E+03

8006	DHQSRSGPWGLL (SEQ ID NO: 16)	4.85 E+03	-	1.88 E+04	2.62 E+04	1.82 E+04
8009	DHQSRSISSGLLSS (SEQ ID NO: 19)	1.65 E+04	-	8.28 E+03	-	6.22 E+03

Note: "-" indicates below limit of quantification.

EXAMPLE 3: *In Vivo* Stability of Activatable Antibodies with Exemplary Substrates

[0404] The study provided herein evaluates the in vivo stability of activatable antibodies with the exemplary substrates.

[0405] This exemplary study measured the stability of activatable antibodies containing exemplary substrates by administering a dose of the activatable antibodies to mice, and then measuring the cleaved activatable antibody in the plasma by capillary electrophoresis immunoassay. The stability was compared to the activatable antibody with control substrates 3001 (AVGLLAPPGGLSGRSDNH; SEQ ID NO: 79) and 5007 (APRSALAHGLF; SEQ ID NO: 80).

[0406] In this study, nu/nu mice of about 7-8 weeks of age were administered intraperitoneally with the indicated test article at a dosage of 10 mg/kg. After 7 days following the administration, terminal blood was collected by cardiac puncture and processed to plasma within 1 hour of collection. The collected sample was diluted 1:50 in phosphate-buffered saline solution and denatured. The sample was analyzed using the Wes™ Western Blot protocol (Protein Simple) using the A110UK goat anti-human IgG antibody (American Qualex) and an anti-goat secondary antibody (Jackson ImmunoResearch). The fraction of cleaved activatable antibody was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the Compass software (Protein Simple). The results of these exemplary assays are summarized in Table 7.

[0407] These exemplary results showed that activatable antibodies with exemplary substrates 8000, 8001, 8003, 8005, 8006, 8007, 8008, 8009, 8010, 8011, and 8012 demonstrate a higher or comparable *in vivo* stability than activatable antibodies that have the control substrates 3001 (AVGLLAPPGGLSGRSDNH; SEQ ID NO: 79) and 5007 (APRSALAHGLF; SEQ ID NO: 80).

Table 7: *In Vivo* Stability of Activatable Antibodies with Exemplary Substrates

Substrate of Activatable Antibody	Substrate	*In Vivo* % Activation (nu/nu, day 7)

3001	AVGLLAPPGGLSGRSDNH (SEQ ID NO: 79)	63.0
5007	APRSALAHGLF (SEQ ID NO: 80)	56.0
8000	AIALYADHQSRS (SEQ ID NO: 10)	41.2
8001	DHQSRSAIALYAD (SEQ ID NO: 11)	50.5
8003	DHQSRSAIALYA (SEQ ID NO: 13)	50.1
8005	HQSRSGPWGLL (SEQ ID NO: 15)	50.9
8006	DHQSRSGPWGLL (SEQ ID NO: 16)	46.9
8007	GPWGLLDHQSRS (SEQ ID NO: 17)	41.4
8008	ISSGLLSSDHQSRS (SEQ ID NO: 18)	30.9
8009	DHQSRSISSGLLSS (SEQ ID NO: 19)	36.8
8010	DHQSRSSGLLSS (SEQ ID NO: 20)	28.4
8011	ALAHGLFSDHQSRS (SEQ ID NO: 21)	41.0
8012	SDHQSRSALAHGLF (SEQ ID NO: 22)	53.1

EXAMPLE 4: Masking Efficiency of Activatable Antibodies with Exemplary Substrates

[0408] The studies provided herein evaluate the *in vitro* masking efficiency of activatable antibodies that include exemplary substrates cleavable by matrix metalloprotease (MMP) and/or matriptase (MT-SP1).

[0409] In these studies, a solid-phase binding assay (ELISA) was used to demonstrate the binding affinity of anti-EGFR activatable antibodies that include substrates of the present disclosure to recombinant EGFR. The binding affinity to EGFR of the activatable antibodies with the indicated substrate of the present disclosure was measured and compared to the unmasked control c225v5 antibody (SEQ ID NOs: 457 and 458). A summary of these exemplary results is shown in Figs. 1A-1C and Table 8.

[0410] These exemplary results showed that the substrates had an effect by increasing the apparent masking efficiency of the masking moiety in the activatable antibody.

Table 8. ***In Vitro*** **Binding Activity and Masking Efficiency of Activatable Antibodies**

Test Article	K_D in nM	Masking Efficiency
Unmasked control C225v5 antibody	0.12	1x
C225v5-3954-2001	3.99	33x
C225v5-3954-8003	28.05	234x
C225v5-3954-8001	12.57	105x
Test Article	**K_D in nM**	**Masking Efficiency**
Unmasked control C225v5 antibody	0.18	1x
C225v5-3954-2001	5.44	30x
C225v5-3954-8001	17.04	95x
C225v5-3954-8009	14.80	82x
C225v5-3954-8012	10.73	60x
Test Article	**K_D in nM**	**Masking Efficiency**
Unmasked control C225v5 antibody	0.096	1x
C225v5-3954-2001	1.77	18x
C225v5-3954-8007	11.82	123x
C225v5-3954-8006	11.57	121x

EXAMPLE 5: ***In Vivo*** **Efficacy of Anti-EGFR Activatable Antibodies with Exemplary Substrates**

[0411] The studies provided herein evaluate the *in vivo* efficacy of activatable antibodies that include exemplary substrates cleavable by matrix metalloprotease (MMP) and/or matriptase (MT-SP1) using the H292 (human lung cancer cell line) xenograft model and using the SUM149 (human triple negative breast cancer cell line) xenograft model.

[0412] In these studies, H292 (human lung cancer-derived cell line) subcutaneous xenograft tumors in female nu/nu mice of 6-8 weeks of age were grown to an average volume of 89-279 mm^3. The H292 cell line is responsive to the anti-EGFR antibody cetuximab. The mice were then randomized into groups of 8 mice each and each group was dosed intraperitoneally on day 1 with 7.5 mg/kg or 9 mg/kg of the indicated test article as noted in Figs. 2A and 2C-2E . The mean tumor volume ± SEM was plotted for each time point following administration of the test article, as shown in Figs. 2A and 2C-2E. Each mouse was treated with activated antibodies with the indicated substrates, or with cetuximab or immunoglobulin (IVIG) control. The efficacy was determined with activatable antibodies that have the control substrate 2001 (ISSGLLSGRSDNH; SEQ ID NO: 78).

[0413] In addition, one study was preformed using SUM149 (human triple negative breast cancer-derived cell line) xenograft subcutaneous tumors in female nu/nu mice of 6-8 weeks of age grown to an average volume of 158-219 mm^3. The SUM149 cell line is responsive to the anti-EGFR antibody cetuximab. The mice were then randomized into groups of 8 mice each and each group was dosed intraperitoneally on day 1 with 5 mg/kg of the indicated test article. The mean tumor volume ± SEM was plotted for each time point following administration of the test article, as shown in Fig. 2B. Each mouse was treated with activated antibodies with the indicated substrates, or with cetuximab or immunoglobulin (IVIG) control. The efficacy was determined with activatable antibodies that having control substrates 5007 (APRSALAHGLF; SEQ ID NO: 80) and 3001 (AVGLLAPPGGLSGRSDNH; SEQ ID NO: 79).

[0414] In addition to the *in vivo* efficacy study, an intra-tumoral activation assay was performed using the indicated activatable antibodies as shown in Figs. 2F-2I. Tumors were collected from the mice at days 4, 7, or 8 as indicated after dosing. The tumor tissue was lysed with immunoprecipitation buffer (Pierce) containing HALT protease inhibitor cocktail (Thermo Fisher) and EDTA and lysed using the Barocycler (Pressure Bioscience). The sample was analyzed using the Wes™ Western Blot instrument (Protein Simple) using the A110UK goat anti-human IgG antibody (American Qualex) and an anti-goat secondary antibody (Jackson ImmunoResearch). The fraction of cleaved activatable antibody was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody. The results of these exemplary assays are summarized in Figs. 2F-2I.

[0415] As shown in Figs. 2A-2E, the activatable antibodies with exemplary substrates 8012 (SDHQSRSALAHGLF; SEQ ID NO: 22), 8009 (DHQSRSISSGLLSS; SEQ ID NO: 19), 8001 (DHQSRSAIALYAD; SEQ ID NO: 11), 8003 (DHQSRSAIALYA; SEQ ID NO: 13), 8007 (GPWGLLDHQSRS; SEQ ID NO: 17), and 8006 (DHQSRSGPWGLL; SEQ ID NO: 16) demonstrated an *in vivo* efficacy that was comparable with cetuximab, which lacks a prodomain.

[0416] As shown in Figs. 2F-2I, the activatable antibodies with substrates 8012, 8009, 8001, 8003, 8007, and 8006 demonstrated light chain activation in the H292 xenograft model that was comparable with the activatable antibody with the control substrate 2001 (ISSGLLSGRSDNH; SEQ ID NO: 78).

EXAMPLE 6: *In situ* Stability of Anti-EGFR Activatable Antibodies in Human Bone Marrow Aspirates

[0417] The study provided herein evaluates the *in situ* stability of activatable antibodies with the exemplary substrates 8006 (DHQSRSGPWGLL; SEQ ID NO: 16) and 8007 (GPWGLLDHQSRS; SEQ ID NO: 17) by human bone marrow aspirates. Fresh human bone marrow aspirates from healthy donors were purchased from Stemcell Technology Inc. (Catalog No. 70502) and AllCells Inc. and were processed to lyse red blood cells and washed 5 times with buffer or serum-free media. The cells were plated at a density of 250,000 cells per well in serum-free RPMI media and incubated for 30 min at room temperature with an equal volume of 80 μg/mL unmasked control c225v5 antibody prepared in serum-free RPMI media. An equal volume of AF647-labeled c225 antibodies prepared at 40 μg/mL in serum-free RPMI media were then added to form a mixture and incubated at a final concentration of 20 μ/mL at 37°C for 21 or 24 hours. Cells were pelleted through centrifugation for 5 min at 300 x *g*. Supernatants were collected from each incubated mixture and transferred into a well of a 96-well PCR plate for assay by capillary electrophoresis. Each supernatant sample was mixed with Pico Sample Buffer (Perkin Elmer) containing 2-beta-mercaptoethanol at four parts sample and one part of Pico Sample Buffer and then heated at 95°C for 10 minutes. Substrate cleavage was measured by capillary electrophoresis using a LabChip GXII Touch system (Perkin Elmer) with the HT Pico Protein Express 100 protocol (Perkin Elmer). Protein Express Assay LabChips (Perkin Elmer #760499) were set up using the protocol of the Protein Pico Assay Reagent Kit (Perkin Elmer #760498). The fraction of cleaved activatable antibody in bone marrow cell supernatants was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the LabChip GX Reviewer software (Perkin Elmer). Data was averaged from N=2-4 donors and normalized to the 2001 substrate cleavage measured by each donor.

[0418] As shown in Fig. 3, activatable antibodies with substrate 8006 (DHQSRSGPWGLL; SEQ ID NO: 16) and 8007 (GPWGLLDHQSRS; SEQ ID NO: 17) demonstrate a resistance to cleavage *in situ* in the bone marrow as compared to control substrates 2001 (ISSGLLSGRSDNH; SEQ ID NO: 78), 5007 (APRSALAHGLF; SEQ ID NO: 80) (WO2020118109) and 3001 (AVGLLAPPGGLSGRSDNH; SEQ ID NO: 79) (WO2016118629).

EXAMPLE 7: Evaluation of Protease Activity in Patient-Derived Tumor Samples

[0419] Protease activities in patient-derived tumor samples were evaluated using the quantitative zymography (QZ) assay. See Howng, B. et al. "Novel Ex Vivo Zymography Approach for Assessment of Protease Activity in Tissues with Activatable Antibodies," Pharmaceutics 2021, 13(9), 1390. Tumor tissue samples from colorectal cancer (CRC), head and neck cancer, and triple negative breast cancer (TNBC) patients were analyzed utilizing activatable antibodies of the present disclosure with the substrates 8000 (AIALYADHQSRS;SEQ ID NO: 10), 8001 (DHQSRSAIALYAD;SEQ ID NO: 11), 8008 (ISSGLLSSDHQSRS;SEQ ID NO: 18), 8009 (DHQSRSISSGLLSS ;SEQ ID NO: 19), 8010 (DHQSRSSGLLSS;SEQ ID NO: 20), 8011 (ALAHGLFSDHQSRS;SEQ ID NO: 21), and 8012 (SDHQSRSALAHGLF;SEQ ID NO: 22).

[0420] The protease activity in the tumor sections of 12 um thickness was assessed. A hydrophobic barrier was drawn around the tissue sample to maintain liquid on the tissue using an ImmEdge Hydrophobic Barrier Pen (Vector Laboratories), and the slides were then incubated with 40 μg/mL unmasked control c225v5 antibody in buffer consisting of 150 mM Tris HCl pH 7.4, 5 mM $CaCl_2$ 100 μM $ZnCl_2$, and 0.005% Tween-20 (QZ assay buffer) for 30 minutes at room temperature. An equal volume of AF647-labeled c225 activatable antibodies prepared at 40 μg/mL in QZ buffer were then added directly onto the tissue containing the buffer to form a mixture and incubated at a final concentration of 20 μg/mL in a humidified chamber at 37°C for 48 hours.

[0421] Following 48 hours of incubation, supernatants were collected from each incubated mixture and transferred into a well of a 96-well PCR plate for assay by capillary electrophoresis. Each supernatant sample was mixed with Pico Sample Buffer (Perkin Elmer) containing 2-beta-mercaptoethanol at four parts sample and one part of Pico Sample Buffer and then heated at 95°C for 10 minutes. The composition of each supernatant sample was then assessed using the LabChip GXII Touch (Perkin Elmer) with the HT Pico Protein Express 100 protocol (Perkin Elmer). Protein Express Assay LabChips (Perkin Elmer #760499) were set up using the protocol of the Protein Pico Assay Reagent Kit (Perkin Elmer #760498). The fraction of cleaved activatable antibody in the tumor tissue supernatants was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the LabChip GX Reviewer software (Perkin Elmer).

[0422] As shown in Figs. 4A-4C, increased LC activation is observed in activatable antibodies with substrate 8000, 8001, 8008-8012 of the present disclosure compared to

activatable antibody with the control substrate 2001 in the CRC tumor, the head and neck tumor, and the TNBC tumor tissues. Comparable levels of LC activation are observed for activatable antibodies with substrates 8008-8012 of the present disclosure and activatable antibodies with the control substrates 5007 and 3001 in CRC tumor tissues, whereas substrates 8000 and 8001 showed a higher LC activation. Comparable levels of LC activation are observed for activatable antibodies with substrates 8000, 8001, 8008-8012 of the present disclosure and activatable antibodies with the control substrates 5007 and 3001 in head and neck tumor tissues. Increased levels of LC activation are observed for activatable antibodies with substrates 8000, 8001, 8008-8012 of the present disclosure compared to activatable antibodies with the control substrate 5007 in TNBC tumor tissues. Comparable levels of LC activation are observed for activatable antibodies with substrates 8008-8010 and 8012 of the present disclosure and activatable antibodies with the control substrate 3001 in TNBC tumor tissues, whereas substrates 8000, 8001, and 8011 showed a higher LC activation.

EXAMPLE 8: *In situ* Activation of Anti-EGFR Activatable Antibodies in Patient-Derived "Liquid Tumor" Samples

[0423] The study provided herein evaluates the *in situ* cleavage of exemplary activatable antibodies of the present disclosure with the substrate 8003 (DHQSRSAIALYA; SEQ ID NO: 13) by human multiple myeloma and acute myeloid leukemia samples using the QZ assay. Frozen multiple myeloma and acute myeloid leukemia bone marrow mononuclear cells (BMMCs) were purchased from Discovery Life Sciences. The cells were plated at a density of 28,000 cells per well in serum-free RPMI media and incubated for 30 min at room temperature with an equal volume of 80 µg/mL unmasked control c225v5 antibody prepared in serum-free RPMI media. An equal volume of AF647-labeled c225 activatable antibodies prepared at 40 µg/mL in serum-free RPMI media were then added to form a mixture and incubated at a final concentration of 20 µg/mL at 37°C for 24 hours. Cells were pelleted through centrifugation for 5 min at 300 x *g*. Supernatants were collected from each incubated mixture and transferred into a well of a 96-well PCR plate for assay by capillary electrophoresis. Each supernatant sample was mixed with Pico Sample Buffer (Perkin Elmer) containing 2-beta-mercaptoethanol at four parts sample and one part of Pico Sample Buffer and then heated at 95°C for 10 minutes. Substrate cleavage was measured by capillary electrophoresis using a LabChip® GXII Touch™ system (Perkin Elmer) with the HT Pico Protein Express 100 protocol (Perkin Elmer). Protein Express Assay LabChips (Perkin Elmer #760499) were set up using the protocol of the Protein Pico Assay Reagent Kit (Perkin Elmer

#760498). The fraction of cleaved activatable antibody in multiple myeloma and acute myeloid leukemia BMMC supernatants was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the LabChip® GX Reviewer software (Perkin Elmer).

[0424] As shown in FIG. 5A, increased LC activation is observed in the activatable antibodies with CM 8003 compared to the activatable antibodies with the control substrates 2001 or 5007 in multiple myeloma BMMCs.

[0425] As shown in FIG. 5B, increased LC activation is observed in the activatable antibodies with CM 8003 compared to the activatable antibodies with the control substrates 2001 or 5007 in acute myeloid leukemia BMMCs.

EXAMPLE 9: *In Vitro* Cleavability of Exemplary CMs in a Peptide Probe Cleavage Assay

[0426] The study provided herein evaluated the cleavability kinetics (i.e., pM/s and k_{cat}/K_M ($M^{-1}s^{-1}$)) of CMs with membrane type serine protease 1 (MT-SP1), matrix metalloprotease (MMP) 2, MMP9, and MMP14. The CMs listed in Table 9 below were presented in an internally quenched peptide probe format, rather than included in an activatable antibody format. In the internally quenched probes, the CM sequence was positioned between a 7-methoxycoumarin-4-acetyl (MCA) fluorophore and a 2,4-dinitrophenyl (DNP) quencher such that cleavage of the CM sequence produced a fluorescence signal. The probes were of the following designs: (MCA)-His-Gln-Ser-Arg-Ser-Ala-Ile-Ala-Leu-X_1-X_2-Lys(DNP)-*D*-Arg (SEQ ID NO: 653) where X_1 is Y and X_2 is A; (MCA)-His-Gln-Ser-Arg-Ser-Ala-Ile-Ala-Leu-X_1-Gly-Lys(DNP)-*D*-Arg (SEQ ID NO: 654) where X_1 is Y; (MCA)-His-Gln-Ser-Arg-Ser-Ala-Ile-Ala-Leu-Gly-Ser-Lys(DNP)-*D*-Arg (SEQ ID NO: 655); (MCA)-His-Gln-Ser-Arg-Ser-Gly-Pro-Trp-Gly-Leu-Leu-Lys(DNP)-*D*-Arg (SEQ ID NO: 662); (MCA)-His-Gln-Ser-Arg-Ser-Pro-Trp-Gly-Leu-Gly-Lys(DNP)-*D*-Arg (SEQ ID NO: 665); (MCA)-Ser-Pro-Trp-Gly-Leu-His-Gln-Ser-Arg-Ser-Lys(DNP)-*D*-Arg (SEQ ID NO: 657); and *D*-Arg-Lys(DNP)-Ala-Ile-Ala-Leu-His-Gln-Ser-Arg-Ser-Lys(MCA)-*D*-Arg (SEQ ID NO: 658), as indicated in Table 9. The cleavage rates (pM/s) were measured using 20 μM internally quenched peptide probe and 20 nM MT-SP1, MMP2, MMP9, or MMP14. Cleavability kinetics (i.e., pM/s and k_{cat}/K_M ($M^{-1}s^{-1}$)) were determined in 96- or 384-well plate format at 37°C in the following buffers: 50 mM TRIS-HCl (pH 7.4), 150 mM NaCl, 0.05% Tween 20 for MT-SP1, 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35 for MMP2 and MMP9, and 50 mM HEPES (pH 6.8), 10 mM

$CaCl_2$, 0.5 mM $MgCl_2$, 0.05% (w/v) Brij-35 for MMP14. Cleavability kinetics were determined on an Infinite 200 PRO (Tecan) multimode plate reader using a fluorescence excitation wavelength of 320 nm and an emission wavelength of 405 nm.

Table 9. CMs tested in peptide format

Sequence	SEQ ID NO:
HQSRSGPWGLL	15
HQSRSPWGL	126
PWGLHQSRS	489
HQSRSAIALYA	14
HQSRSAIALY	636
HQSRSAIAL	637
AIALHQSRS	41

[0427] Table 10 provides exemplary probe cleavage rates (pM/s) of the exemplary CMs of Table 9 with MT-SP1, MMP2, MMP9, and MMP14. Table 11 provides exemplary k_{cat}/K_M (M^{-1} s^{-1}) values of the exemplary CMs of Table 9 with MT-SP1, MMP2, MMP9, and MMP14.

Table 10. *In Vitro* Activation of Peptide Probes with Exemplary CMs (pM/s)

Probe CM Sequence	Probe Cleavage (pM/s)			
	MT-SP1	MMP2	MMP9	MMP14
HQSRSGPWGLL (SEQ ID NO: 15)	3380	284	2210	1020
HQSRSPWGL (SEQ ID NO: 126)	1210	44.2	372	52.5
PWGLHQSRS (SEQ ID NO: 489)	1280	2550	5150	5880
HQSRSAIALYA (SEQ ID NO: 14)	1750	468	255	300
HQSRSAIALY (SEQ ID NO: 636)	1960	434	328	92.7
HQSRSAIAL	Above limit of quantitation	76.9	41.9	95.5

(SEQ ID NO: 637)				
AIALHQSRS (SEQ ID NO: 41)	11500	4220	2550	1280

Table 11. ***In Vitro*** **Activation of Peptide Probes with Exemplary CMs (k_{cat}/K_M)**

Probe CM Sequence	**Probe Cleavage (k_{cat}/K_M (M^{-1} s^{-1}))**			
	MT-SP1	**MMP2**	**MMP9**	**MMP14**
HQSRSGPWGLL (SEQ ID NO: 15)	1.2×10^5	3.5×10^3	6.5×10^4	9.4×10^3
HQSRSPWGL (SEQ ID NO: 126)	3.5×10^4	-	1.1×10^4	-
PWGLHQSRS (SEQ ID NO: 489)	2.8×10^4	5.4×10^4	1.2×10^5	6.0×10^5
HQSRSAIALYA (SEQ ID NO: 14)	6.3×10^5	2.9×10^4	2.4×10^4	1.3×10^4
HQSRSAIALY (SEQ ID NO: 636)	7.9×10^5	6.7×10^4	5.7×10^4	1.1×10^4
HQSRSAIAL (SEQ ID NO: 637)	5.1×10^5	4.9×10^3	-	1.9×10^3
AIALHQSRS (SEQ ID NO: 41)	5.5×10^4	1.7×10^4	-	3.9×10^3

Note: "-" indicates not determined.

[0428] These exemplary results show that CM HQSRSGPWGLL (SEQ ID No: 15), CM PWGLHQSRS (SEQ ID No: 489), CM HQSRSAIALYA (SEQ ID No: 14), and CM HQSRSAIALY (SEQ ID No: 636) are cleavable by the MMPs MMP2, MMP9, and MMP14 with a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^3 M^{-1} s^{-1}. These exemplary results show that that CM PWGLHQSRS (SEQ ID No: 489), CM HQSRSAIALYA (SEQ ID No: 14), and CM HQSRSAIALY (SEQ ID No: 636) are cleavable by the MMPs MMP2, MMP9, and MMP14 with a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^4 M^{-1} s^{-1}. These exemplary results show that CM HQSRSPWGL (SEQ ID No: 126) is cleavable by MMP9 with a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1×10^4 M^{-1} s^{-1}. These exemplary results show that CM HQSRSAIAL (SEQ ID No: 637)

and CM AIALHQSRS (SEQ ID No: 41) are cleavable by MMP2 and MMP14 with a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^3 M^{-1} s^{-1}.

[0429] These exemplary results show that CM HQSRSGPWGLL (SEQ ID No: 15), CM HQSRSPWGL (SEQ ID No: 126), CM PWGLHQSRS (SEQ ID No: 489), CM HQSRSAIALYA (SEQ ID No: 14), CM HQSRSAIALY (SEQ ID No: 636), CM HQSRSAIAL (SEQ ID No: 637), and CM AIALHQSRS (SEQ ID No: 41) are cleavable by MT-SP1 with a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^4 M^{-1} s^{-1}. These exemplary results show that CM HQSRSGPWGLL (SEQ ID No: 15), CM HQSRSAIALYA (SEQ ID No: 14), CM HQSRSAIALY (SEQ ID No: 636), and CM HQSRSAIAL (SEQ ID No: 637) are cleavable by MT-SP1 with a k_{cat}/K_M ($M^{-1}s^{-1}$) of greater than 1 x 10^5 M^{-1} s^{-1}.

EXAMPLE 10: *In Vitro* Cleavability of Additional Exemplary CMs in a Peptide Probe Cleavage Assay

[0430] The study provided herein evaluated the cleavability kinetics (i.e., pM/s) of CMs with membrane type serine protease 1 (MT-SP1), matrix metalloprotease (MMP) 2, MMP9, and MMP14. The CMs listed in Table 12 below were presented in an internally quenched peptide probe format, rather than included in an activatable antibody format. In the internally quenched probes, the CM sequence was positioned between a 7-methoxycoumarin-4-acetyl (MCA) fluorophore and a 2,4-dinitrophenyl (DNP) quencher such that cleavage of the CM sequence produced a fluorescence signal. The probes were of the following designs: (MCA)-His-Gln-Ser-Arg-Ser-Gly-Pro-Trp-Gly-Leu-X_1-Lys(DNP)-*D*-Arg (SEQ ID NO: 659) where X_1 is Leu or absent, (MCA)-His-Gln-Ser-Arg-Pro-Trp-Gly-Leu-Gly-Ser-Lys(DNP)-*D*-Arg (SEQ ID NO: 660); (MCA)-His-Gln-Ser-Arg-Ser-Pro-Trp-Gly-Gly-Ser-Lys(DNP)-*D*-Arg (SEQ ID NO: 661); (MCA)-His-Gln-Ser-Arg-Ser-Gly-Pro-Trp-Gly-Leu-Leu-Lys(DNP)-*D*-Arg (SEQ ID NO: 662) where any single amino acid in the sequence His-Gln-Ser-Arg-Ser-Gly-Pro-Trp-Gly-Leu-Leu (SEQ ID NO: 15) is replaced with Ala; (MCA)-Ser-Pro-Trp-Gly-Leu-His-Gln-Ser-Arg-Gly-Lys(DNP)-*D*-Arg (SEQ ID NO: 663); and (MCA)-Ser-Pro-Trp-Gly-Leu-His-Gln-Ser-Lys-Ser-Lys(DNP)-*D*-Arg (SEQ ID NO: 664). The cleavage rates (pM/s) were measured using 20 μM internally quenched peptide probe and 20 nM MT-SP1, MMP2, MMP9, or MMP14. Cleavability kinetics (i.e., pM/s) were determined in 96- or 384-well plate format at 37°C in the following buffers: 50 mM TRIS-HCl (pH 7.4), 150 mM NaCl, 0.05% Tween 20 for MT-SP1, 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35 for MMP2 and MMP9, and 50 mM HEPES (pH 6.8), 10 mM $CaCl_2$, 0.5 mM $MgCl_2$, 0.05% (w/v) Brij-35 for MMP14. Cleavability kinetics were determined on an

Infinite 200 PRO (Tecan) multimode plate reader using a fluorescence excitation wavelength of 320 nm and an emission wavelength of 405 nm.

Table 12. CMs tested in peptide format

Sequence	SEQ ID NO:
HQSRSGPWGLL	15
HQSRSGPWGL	122
HQSRPWGL	134
HQSRSPWG	640
AQSRSGPWGLL	641
HASRSGPWGLL	642
HQARSGPWGLL	643
HQSASGPWGLL	644
HQSRAGPWGLL	645
HQSRSAPWGLL	646
HQSRSGAWGLL	647
HQSRSGPAGLL	639
HQSRSGPWALL	648
HQSRSGPWGAL	649
HQSRSGPWGLA	650
PWGLHQSR	502
PWGLHQSKS	490

[0431] Table 13 provides exemplary probe cleavage rates (pM/s) of the exemplary CMs of Table 9 with MT-SP1, MMP2, MMP9, and MMP14.

Table 13. *In Vitro* Activation of Peptide Probes with Exemplary CMs (pM/s)

Probe CM Sequence	Probe Cleavage (pM/s)			
	MT-SP1	MMP2	MMP9	MMP14
HQSRSGPWGLL (SEQ ID NO: 15)	3380	284	2210	1020
HQSRSGPWGL	9930	94.8	705	27.3

(SEQ ID NO: 122)				
HQSRPWGL (SEQ ID NO: 134)	32.0	4020	9600	548
HQSRSPWG (SEQ ID NO: 640)	3470	61.6	86.5	0.00
AQSRSGPWGLL (SEQ ID NO: 641)	Above limit of quantitation	726	7380	437
HASRSGPWGLL (SEQ ID NO: 642)	1860	692	6470	424
HQARSGPWGLL (SEQ ID NO: 643)	11300	623	7430	344
HQSASGPWGLL (SEQ ID NO: 644)	8.00	696	5410	335
HQSRAGPWGLL (SEQ ID NO: 645)	7330	594	7440	379
HQSRSAPWGLL (SEQ ID NO: 646)	Above limit of quantitation	1390	10700	660
HQSRSGAWGLL (SEQ ID NO: 647)	Above limit of quantitation	65.4	436	36.4
HQSRSGPAGLL (SEQ ID NO: 639)	10800	17600	13900	1300
HQSRSGPWALL (SEQ ID NO: 648)	4120	488	7500	475
HQSRSGPWGAL (SEQ ID NO: 649)	7780	61.1	5.9	0.00
HQSRSGPWGLA (SEQ ID NO: 650)	7920	410	6810	345
PWGLHQSR (SEQ ID NO: 502)	540	6840	18600	1200
PWGLHQSKS (SEQ ID NO: 490)	1630	6640	19300	1190

[0432] These exemplary results show that CM HQSRSGPWGLL (SEQ ID NO: 15), CM HQSRPWGL (SEQ ID NO: 134), CM AQSRSGPWGLL (SEQ ID NO: 641), CM HASRSGPWGLL (SEQ ID NO: 642), CM HQARSGPWGLL (SEQ ID NO: 643), CM HQSASGPWGLL (SEQ ID NO: 644), CM HQSRAGPWGLL (SEQ ID NO: 645), CM HQSRSAPWGLL (SEQ ID NO: 646), CM HQSRSGPAGLL (SEQ ID NO: 639), CM HQSRSGPWALL (SEQ ID NO: 648), CM HQSRSGPWGLA (SEQ ID NO: 650), CM PWGLHQSR, (SEQ ID NO: 502), and CM PWGLHQSKS (SEQ ID NO: 490) have high MMP9 cleavage with a cleavage rate of greater than 2000 pM/s. The exemplary results show that CM HQSRPWGL (SEQ ID NO: 134), CM HQSRSAPWGLL (SEQ ID NO: 646), CM HQSRSGPAGLL (SEQ ID NO: 639), CM PWGLHQSR (SEQ ID NO: 502), and CM PWGLHQSKS (SEQ ID NO: 490) have high MMP2 and MMP14 cleavage with a cleavage rate of greater than 1000 pM/s for MMP2 and 500 pM/s for MMP14.

[0433] These exemplary results show that CM HQSRSGPWGLL (SEQ ID NO: 15), HQSRSGPWGL (SEQ ID NO: 122), HQSRSPWG (SEQ ID NO: 640), CM AQSRSGPWGLL (SEQ ID NO: 641), CM HASRSGPWGLL (SEQ ID NO: 642), CM HQARSGPWGLL (SEQ ID NO: 643), CM HQSRAGPWGLL (SEQ ID NO: 645), CM HQSRSAPWGLL (SEQ ID NO: 646), CM HQSRSGAWGLL (SEQ ID NO: 647), HQSRSGPAGLL (SEQ ID NO: 639), CM HQSRSGPWALL (SEQ ID NO: 648), CM HQSRSGPWGAL (SEQ ID NO: 649), CM HQSRSGPWGLA (SEQ ID NO: 650), CM PWGLHQSR, (SEQ ID NO: 502), and CM PWGLHQSKS (SEQ ID NO: 490) have high MT-SP1 cleavage with a cleavage rate of greater than 500 pM/s or above the limit of quantitation.

EXAMPLE 11: Additional Exemplary Activatable Antibodies and Protease Cleavable Substrates

[0434] The studies provided herein describe additional exemplary substrates that include at least one cleavable moiety cleavable by a matrix metalloprotease (MMP), and at least one cleavable moiety cleavable by a matriptase (MT-SP1).

[0435] Exemplary activatable antibodies were constructed such that each one includes one of the substrates listed in Table 14. The exemplary activatable antibodies, the sequences of which are listed in Table 15, include an antibody or antigen binding fragment thereof (AB) that is based on a mouse/human chimeric monoclonal antibody that specifically binds to epidermal growth factor receptor (EGFR). The exemplary activatable antibodies also include a prodomain coupled to the N-terminus of the light chain of the AB. Each prodomain includes

a masking moiety (MM) and a substrate, and the substrate includes at least one sequence of Table 14.

Table 14: Exemplary MMP / MT-SP1 Substrates

Sequence	SEQ ID NO:
DHQSRSGPWGLL	16
PWGLLHQSRS	485
PWGLHQSRS	489
HQSRSPWGL	126
DHQSKSGPWGLL	178
HQSRSGPWGL	122
HQSRSGPYGLL	138
HQSRSGPFGLL	158
HQSRSGPRGLL	638
HQSRSGPAGLL	639

Table 15: Activatable Antibody Sequences

Molecule Name	Light chain EGFR Mask	Light chain substrate	Protein description (Light chain /Heavy chain)
ProC905	CISPRGCPDGPYVMY (SEQ ID NO: 81)	DHQSRSGPWGLL (SEQ ID NO: 16)	C225v5 3954 9820 (SEQ ID NO: 229 / SEQ ID NO: 218)
ProC3633	CISPRGCPDGPYVMY (SEQ ID NO: 81)	PWGLLHQSRS (SEQ ID NO: 485)	C225v5 3954 2198 (SEQ ID NO: 240 / SEQ ID NO: 218)
ProC3634	CISPRGCPDGPYVMY (SEQ ID NO: 81)	PWGLHQSRS (SEQ ID NO: 489)	C225v5 3954 2298 (SEQ ID NO: 241 / SEQ ID NO: 218)
ProC3635	CISPRGCPDGPYVMY (SEQ ID NO: 81)	HQSRSPWGL (SEQ ID NO: 126)	C225v5 3954 9822 (SEQ ID NO: 242 / SEQ ID NO: 218)
ProC3636	CISPRGCPDGPYVMY (SEQ ID NO: 81)	DHQSKSGPWGLL (SEQ ID NO: 178)	C225v5 3954 9823 (SEQ ID NO: 243 / SEQ ID NO: 218)
ProC3637	CISPRGCPDGPYVMY (SEQ ID NO: 81)	HQSRSGPWGL (SEQ ID NO: 122)	C225v5 3954 9824 (SEQ ID NO: 244 / SEQ ID NO: 218)
ProC3638	CISPRGCPDGPYVMY (SEQ ID NO: 81)	HQSRSGPYGLL (SEQ ID NO: 138)	C225v5 3954 9827 (SEQ ID NO: 245 / SEQ ID NO: 218)
ProC3639	CISPRGCPDGPYVMY (SEQ ID NO: 81)	HQSRSGPFGLL (SEQ ID NO: 158)	C225v5 3954 9828 (SEQ ID NO: 246 / SEQ ID NO: 218)
ProC3640	CISPRGCPDGPYVMY	HQSRSGPRGLL	C225v5 3954 9829

	(SEQ ID NO: 81)	(SEQ ID NO: 638)	(SEQ ID NO: 247 / SEQ ID NO: 218)
ProC3641	CISPRGCPDGPYVMY (SEQ ID NO: 81)	HQSRSGPAGLL (SEQ ID NO: 639)	C225v5 3954 9837 (SEQ ID NO: 248 / SEQ ID NO: 218)

EXAMPLE 12: *In Vitro* Cleavability of Activatable Antibodies with Additional Exemplary Substrates

[0436] The studies provided herein evaluate the *in vitro* cleavability of activatable antibodies containing exemplary substrates that include at least one cleavable moiety cleavable by a matrix metalloprotease (MMP) and/or at least one cleavable moiety cleavable by matriptase (MT-SP1).

[0437] The cleavability of the activatable antibodies was measured in the presence of the indicated recombinant proteases (MT-SP1, MMP2, MMP9, and MMP14). Each activatable antibody (500 nM) was incubated with 10 nM of the indicated protease for 4 hours at 37°C. Human recombinant proteases were purchased from R&D Systems: MMP2 (catalog No: 902-MP), MMP9 (catalog No: 911-MP), MMP14 (catalog No: 918-MP) and MT-SP1 (catalog No: 3946-SEB). MMPs were activated according to the manufacturer's instructions. Protease concentrations were determined by active site titration. Activity assays for MMP2 and MMP9 were performed in the following buffer: 50 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 150 mM NaCl, 0.05% (w/v) Brij-35. Activity assays were performed for MMP14 using 50 mM HEPES (pH 6.8), 10 mM $CaCl_2$, 0.5 mM $MgCl_2$, and for MT-SP1 activity assays used 50 mM TRIS-HCl (pH 7.4), 150 mM NaCl, 0.05% Tween 20. Following incubation, the presence of cleavage product was determined at 4 hours by capillary electrophoresis for each protease enzyme using a LabChip GXII Touch system (Perkin Elmer). For capillary electrophoresis assays, the HT Protein Express 100 protocol (Perkin Elmer) was used. LabChip GXII Touch HT Chips (Perkin Elmer #760499) were set up using the protocol of the Protein Express Assay Reagent Kit (Perkin Elmer #CLS960008). The fraction of cleaved activatable antibody was determined by quantifying the fraction of the higher mobility polypeptide corresponding to the cleaved activatable antibody using the LabChip GX Reviewer software (Perkin Elmer). The fraction of activatable antibody, and hence the CM that is cleaved by each particular protease, is presented as a "cleavability percentage" in Table 16.

[0438] These exemplary results show that the substrates showed a range of cleavability by the indicated proteases. These exemplary results show that the substrates showed a range of cleavability by MT-SP1, MMP2, MMP9, and/or MMP14.

Table 16. ***In Vitro*** **Activation of Activatable Antibodies with Exemplary Substrates**

Substrate	Cleavability (%)			
	MT-SP1	**MMP2**	**MMP9**	**MMP14**
DHQSRSGPWGLL (SEQ ID NO: 16)	39.5	86.4	100	69.5
PWGLLHQSRS (SEQ ID NO: 485)	100	29.0	57.4	72.9
PWGLHQSRS (SEQ ID NO: 489)	100	59.2	93.3	59.9
HQSRSPWGL (SEQ ID NO: 126)	27.4	37.1	57.5	4.4
DHQSKSGPWGLL (SEQ ID NO: 178)	18.4	84.6	100	59.0
HQSRSGPWGL (SEQ ID NO: 122)	81.8	53.2	93.5	31.2
HQSRSGPYGLL (SEQ ID NO: 138)	69.1	100	100	67.3
HQSRSGPFGLL (SEQ ID NO: 158)	83.0	96.2	100	58.7
HQSRSGPRGLL (SEQ ID NO: 638)	100	100	100	87.9
HQSRSGPAGLL (SEQ ID NO: 639)	83.3	100	100	86.9

[0439] These exemplary results show a group of substrates in which cleavage by MT-SP1 is at least 25%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSRS (SEQ ID NO: 489), HQSRSPWGL (SEQ ID NO: 126), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0440] These exemplary results also show a group of substrates in which cleavage by both MMP9 and MMP14 is at least 50%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSRS (SEQ ID NO: 489), DHQSKSGPWGLL (SEQ ID NO: 178),

HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0441] These exemplary results also show a group of substrates in which cleavage by both MMP2 and MMP14 is at least 50% and cleavage by matriptase is at least 25%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLHQSRS (SEQ ID NO: 489), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0442] These exemplary results also show a group of substrates in which cleavage by each of MMP2, MMP9 and MMP14 is at least 25%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSRS (SEQ ID NO: 489), DHQSKSGPWGLL (SEQ ID NO: 178), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0443] These exemplary results also show a group of substrates in which cleavage by MMP2 cleavability is at least 50%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLHQSRS (SEQ ID NO: 489), DHQSKSGPWGLL (SEQ ID NO: 178), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0444] These exemplary results also show a group of substrates in which cleavage by MMP2 is at least 50% and cleavage by MT-SP1 is at least 25%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLHQSRS (SEQ ID NO: 489), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0445] These exemplary results also show a group of substrates in which cleavage by both MMP14 and MT-SP1 is at least 25%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSRS (SEQ ID NO: 489), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0446] These exemplary results also show a group of substrates in which cleavage by each of MMP2, MMP14 and MT-SP1 is at least 50%. Such substrates include those having the amino acid sequence of PWGLHQSRS (SEQ ID NO: 489), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0447] These exemplary results also show a group of substrates in which cleavage by each of MMP2, MMP9, MMP14, and MT-SP1 is at least 25%. Such substrates include those having the amino acid sequence of DHQSRSGPWGLL (SEQ ID NO: 16), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSRS (SEQ ID NO: 489), HQSRSGPWGL (SEQ ID NO: 122), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

[0448] These exemplary results also show a group of substrates in which cleavage by each of MMP9, MMP14 and MT-SP1 is at least 50%. Such substrates include those having the amino acid sequence of PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSRS (SEQ ID NO: 489), HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), and HQSRSGPAGLL (SEQ ID NO: 639).

EXAMPLE 13: *In situ* Stability of Additional Exemplary CMs in a Peptide Probe Format in Supernatant from Human Bone Marrow Aspirates

[0449] The study provided herein evaluates the *in situ* stability of peptide probes with the exemplary substrates HQSRSPWGL (SEQ ID NO: 126), PWGLHQSRS (SEQ ID NO: 489), and PAGLSGRS (SEQ ID NO: 666) in human bone marrow aspirates. The probes were of the following designs: (MCA)-His-Gln-Ser-Arg-Ser-Pro-Trp-Gly-Leu-Gly-Lys(DNP)-*D*-Arg (SEQ ID NO: 665), (MCA)-Ser-Pro-Trp-Gly-Leu-His-Gln-Ser-Arg-Ser-Lys(DNP)-*D*-Arg (SEQ ID NO: 657), and (MCA)-Ser-Pro-Ala-Gly-Leu-Ser-Gly-Arg-Ser-Gly-Lys(DNP)-*D*-Arg (SEQ ID NO: 667). Fresh human bone marrow aspirates from healthy donors were purchased from Stemcell Technology Inc. (Catalog No. 70502) and were processed to lyse red blood cells and washed 5 times with serum-free media. The cells were plated at a density of 250,000 cells per well in serum-free RPMI media and incubated at 37°C for 24 hours. Cells were pelleted through centrifugation for 5 min at 300 x *g* and bone marrow supernatant was collected. Collected supernatant was cleared of any cellular debris with centrifugation for 10 min at 15,000 x g. Stability assays were performed by mixing the probe in serum-free media with an equal volume of bone marrow supernatant to a final probe concentration of 20 μM. Cleavability kinetics (i.e., pM/s) were determined in 384-well plate format at 37°C on an

Infinite 200 PRO (Tecan) multimode plate reader using a fluorescence excitation wavelength of 320 nm and an emission wavelength of 405 nm.The substrates HQSRSPWGL (SEQ ID NO: 126) and PWGLHQSRS (SEQ ID NO: 489) were stable in supernatant from human bone marrow aspirates, with cleavage rates of 0 and 302 pM/s, respectively. The substrate PAGLSGRS (SEQ ID NO: 666) had reduced stability in supernatant from human bone marrow aspirates, with a cleavage rate of 2191 pM/s.

Table 17. Exemplary sequences

SEQ ID NO	Notes	Sequences
1	6000	HQSR
2	6001	HQSRS
3	6062	HQSK
4	6004	HQSRSA
5	6005	DHQSRS
6	6006	SDHQSRS
7	7000	AIALYAD
8	7001	AIALYA
635	7004	AIALY
636		HQSRSAIALY
637		HQSRSAIAL
638		HQSRSGPRGLL
639		HQSRSGPAGLL
640		HQSRSPWG
641		AQSRSGPWGLL
642		HASRSGPWGLL
643		HQARSGPWGLL
644		HQSASGPWGLL
645		HQSRAGPWGLL
646		HQSRSAPWGLL
647		HQSRSGAWGLL
648		HQSRSGPWALL
649		HQSRSGPWGAL
650		HQSRSGPWGLA
9	7005	AIAL
10	8000	AIALYADHQSRS
11	8001	DHQSRSAIALYAD
12	8002	HQSRSAIALYAD

13	8003	DHQSRSAIALYA
14	8004	HQSRSAIALYA
15	8005	HQSRSGPWGLL
16	8006	DHQSRSGPWGLL
17	8007	GPWGLLDHQSRS
18	8008	ISSGLLSSDHQSRS
19	8009	DHQSRSISSGLLSS
20	8010	DHQSRSSGLLSS
21	8011	ALAHGLFSDHQSRS
22	8012	SDHQSRSALAHGLF
23	8013	AIALYADHQSR
24	8014	AIALYADHQSK
25	8015	AIALYDHQSRS
26	8016	AIALYDHQSKS
27	8017	AIALYDHQSR
28	8018	AIALYDHQSK
29	8019	AIALDHQSRS
30	8020	AIALDHQSKS
31	8021	AIALDHQSR
32	8022	AIALDHQSK
33	8023	AIALYAHQSRS
34	8024	AIALYAHQSKS
35	8025	AIALYAHQSR
36	8026	AIALYAHQSK
37	8027	AIALYHQSRS
38	8028	AIALYHQSKS
39	8029	AIALYHQSR
40	8030	AIALYHQSK
41	8031	AIALHQSRS
42	8032	AIALHQSKS
43	8033	AIALHQSR
44	8034	AIALHQSK
45	8035	AIALYADHQSRSG
46	8036	AIALYADHQSRSA
47	8037	AIALYADHQSKS
48	8038	SDHQSRSAIALYAD
49	8039	DHQSKSAIALYAD
50	8040	HDHQSRSAIALYAD
51	8041	LDHQSRSAIALYAD
52	8042	LLDHQSRSAIALYAD
53	8043	SSDHQSRSAIALYAD

54	8044	FSDHQSRSAIALYAD
55	8045	SDHQSRSAIALYADG
56	8046	DHQSKSAIALYADG
57	8047	HDHQSRSAIALYADG
58	8048	KDHQSRSAIALYADG
59	8049	LDHQSRSAIALYADG
60	8070	LLDHQSRSAIALYADG
61	8071	SSDHQSRSAIALYADG
62	8072	FSDHQSRSAIALYADG
63	8073	SHQSRSAIALYAD
64	8074	HQSKSAIALYAD
65	8075	SDHQSRSAIALYA
66	8076	DHQSKSAIALYA
67	8077	HDHQSRSAIALYA
68	8078	KDHQSRSAIALYA
69	8079	LDHQSRSAIALYA
70	8080	LLDHQSRSAIALYA
71	8081	SSDHQSRSAIALYA
72	8082	FSDHQSRSAIALYA
73	8083	SDHQSRSAIALYAG
83	8084	DHQSKSAIALYAG
84	8085	HDHQSRSAIALYAG
85	8086	KDHQSRSAIALYAG
86	8087	LDHQSRSAIALYAG
87	8088	LLDHQSRSAIALYAG
88	8089	SSDHQSRSAIALYAG
89	8090	FSDHQSRSAIALYAG
90	8091	SDHQSRSAIALY
91	8092	DHQSRSAIALY
92	8093	DHQSKSAIALY
93	8094	HDHQSRSAIALY
94	8095	KDHQSRSAIALY
95	8096	LDHQSRSAIALY
96	8097	LLDHQSRSAIALY
97	8098	SSDHQSRSAIALY
98	8099	FSDHQSRSAIALY
99	9000	SDHQSRSAIAL
74	1401	ALAHGLF
75	1001	ISSGLLSS
76	1004	AVGLLAPP

77	0001	LSGRSDNH
78	2001	ISSGLLSGRSDNH
79	3001	AVGLLAPPGGLSGRSDNH
80	5007	APRSALAHGLF
81	EGFR Mask	CISPRGCPDGPYVMY
82	EGFR Mask	CISPRGCLDGPYVMY
100	9001	DHQSRSAIAL
101	9002	DHQSKSAIAL
102	9003	HDHQSRSAIAL
103	9004	KDHQSRSAIAL
104	9005	LDHQSRSAIAL
105	9006	LLDHQSRSAIAL
106	9007	SSDHQSRSAIAL
107	9008	FSDHQSRSAIAL
108	9009	SHQSRSAIALYA
109	9010	HQSKSAIALYA
110	9011	SHQSRSAIALY
111	9012	SHQSRSAIAL
112	9013	HQSRSAIALY
113	9014	HQSKSAIALY
114	9015	HQSRSAIAL
115	9016	HQSKSAIAL
116	9017	HQSRAIAL
117	9018	HQSKAIAL
118	9019	SHQSRSGPWGLL
119	9020	HQSKSGPWGLL
120	9021	HQSRSGPWGLG
121	9022	SHQSRSGPWGLG
122	9023	HQSRSGPWGL
123	9024	HQSKSGPWGL
124	9025	HQSRSPWGLL
125	9026	HQSKSPWGLL
126	9027	HQSRSPWGL
127	9028	HQSKSPWGL
128	9029	HQSRGPWGLL
129	9030	HQSKGPWGLL
130	9031	HQSRGPWGL
131	9032	HQSKGPWGL
132	9033	HQSRPWGLL
133	9034	HQSKPWGLL

134	9035	HQSRPWGL
135	9036	HQSKPWGL
136	9037	SHQSRSGPWGL
137	9038	SHQSRSGPYGLL
138	9039	HQSRSGPYGLL
139	9040	HQSKSGPYGLL
140	9041	HQSRSGPYGLG
141	9042	SHQSRSGPYGLG
142	9043	HQSRSGPYGL
143	9044	HQSKSGPYGL
144	9045	HQSRSPYGLL
145	9046	HQSKSPYGLL
146	9047	HQSRSPYGL
147	9048	HQSKSPYGL
148	9049	HQSRGPYGLL
149	9050	HQSKGPYGLL
150	9051	HQSRGPYGL
151	9052	HQSKGPYGL
152	9053	HQSRPYGLL
153	9054	HQSKPYGLL
154	9055	HQSRPYGL
155	9056	HQSKPYGL
156	9057	SHQSRSGPYGL
157	9058	SHQSRSGPFGLL
158	9059	HQSRSGPFGLL
159	9060	HQSKSGPFGLL
160	9061	HQSRSGPFGLG
161	9062	SHQSRSGPFGLG
162	9063	HQSRSGPFGL
163	9064	HQSKSGPFGL
164	9065	HQSRSPFGLL
165	9066	HQSKSPFGLL
166	9067	HQSRSPFGL
167	9068	HQSKSPFGL
168	9069	HQSRGPFGLL
169	9070	HQSKGPFGLL
170	9071	HQSRGPFGL
171	9072	HQSKGPFGL
172	9073	HQSRPFGLL

173	9074	HQSKPFGLL
174	9075	HQSRPFGL
175	9076	HQSKPFGL
176	9077	SHQSRSGPFGL
177	9078	SDHQSRSGPWGLL
178	9079	DHQSKSGPWGLL
179	9080	DHQSRSPWGLL
180	9081	DHQSKSPWGLL
181	9082	DHQSRSPWGL
182	9083	DHQSKSPWGL
183	9084	DHQSRGPWGLL
184	9085	DHQSKGPWGLL
185	9086	DHQSRGPWGL
186	9087	DHQSKGPWGL
187	9088	DHQSRPWGLL
188	9089	DHQSKPWGLL
189	9090	DHQSRPWGL
190	9091	DHQSKPWGL
191	9092	DHQSRSGPWGLG
192	9093	SDHQSRSGPWGLG
193	9094	DHQSRSGPWGL
194	9095	SDHQSRSGPWGL
195	9096	DHQSKSGPWGL
196	9097	SDHQSRSGPYGLL
197	9098	DHQSRSGPYGLL
198	9099	DHQSKSGPYGLL
199	10000	DHQSRSPYGLL
200	10001	DHQSKSPYGLL
201	10002	DHQSRSPYGL
202	10003	DHQSKSPYGL
203	10004	DHQSRGPYGLL
204	10005	DHQSKGPYGLL
205	10006	DHQSRGPYGL
206	10007	DHQSKGPYGL
207	10008	DHQSRPYGLL
208	10009	DHQSKPYGLL
209	10010	DHQSRPYGL
210	10011	DHQSKPYGL
211	10012	DHQSRSGPYGLG

212	10013	SDHQSRSGPYGLG
213	10014	DHQSRSGPYGL
214	10015	DHQSKSGPYGL
215	10016	SDHQSRSGPYGL
216	10017	SDHQSRSGPFGLL
460	10018	DHQSRSGPFGLL
461	10019	DHQSKSGPFGLL
462	10020	DHQSRSPFGLL
463	10021	DHQSKSPFGLL
464	10022	DHQSRSPFGL
465	10023	DHQSKSPFGL
466	10024	DHQSRGPFGLL
467	10025	DHQSKGPFGLL
468	10026	DHQSRGPFGL
469	10027	DHQSKGPFGL
470	10028	DHQSRPFGLL
471	10029	DHQSKPFGLL
472	10030	DHQSRPFGL
473	10031	DHQSKPFGL
474	10032	DHQSRSGPFGLG
475	10033	SDHQSRSGPFGLG
476	10034	DHQSRSGPFGL
477	10035	DHQSKSGPFGL
478	10036	SDHQSRSGPFGL
479	10037	GPWGLGDHQSRS
480	10038	GPWGLLDHQSKS
481	10039	GPWGLLHQSRS
482	10040	GPWGLLHQSKS
483	10041	PWGLLDHQSRS
484	10042	PWGLLDHQSKS
485	10043	PWGLLHQSRS
486	10044	PWGLLHQSKS
487	10045	PWGLDHQSRS
488	10046	PWGLDHQSKS
489	10047	PWGLHQSRS
490	10048	PWGLHQSKS
491	10049	GPWGLDHQSRS
492	10050	GPWGLLEDHQSRS
493	10051	GPWGLSDHQSRS

494	10052	GPWGLGDHQSR
495	10053	GPWGLLDHQSK
496	10054	PWGLLDHQSR
497	10055	PWGLLDHQSK
498	10056	PWGLLHQSR
499	10057	PWGLLHQSK
500	10058	PWGLDHQSR
501	10059	PWGLDHQSK
502	10060	PWGLHQSR
503	10061	PWGLHQSK
504	10062	GPWGLDHQSR
505	10063	GPWGLLEDHQSR
506	10064	GPWGLSDHQSR
507	10065	GPWGLGDHQSRSG
508	10066	GPWGLDHQSRSG
509	10067	GPWGLLEDHQSRSG
510	10068	GPWGLSDHQSRSG
511	10069	GPWGLGDHQSRSA
512	10070	GPWGLDHQSRSA
513	10071	GPWGLLEDHQSRSA
514	10072	GPWGLSDHQSRSA
515	10073	GPWGLGDHQSKS
516	10074	GPWGLDHQSKS
517	10075	GPWGLLEDHQSKS
518	10076	GPWGLSDHQSKS
519	10077	ISSGLLSSDHQSRSR
520	10078	ISSGLLSSDHQSRSG
521	10079	ISSGLLSSDHQSRSA
522	10080	ISSGLLSSDHQSR
523	10081	ISSGLLSSDHQSKS
524	10082	ISSGLLSSDHQSK
525	10083	ISSGLLSSHQSRS
526	10084	ISSGLLSSHQSKS
527	10085	ISSGLLSSHQSR
528	10086	ISSGLLSSHQSK
529	10087	ISSGLLSHQSRS
530	10088	ISSGLLSHQSKS
531	10089	ISSGLLSHQSR
532	10090	ISSGLLSHQSK

533	10091	ISSGLLHQSRS
534	10092	ISSGLLHQSKS
535	10093	ISSGLLHQSR
536	10094	ISSGLLHQSK
537	10095	ISSGLDHQSRS
538	10096	ISSGLDHQSKS
539	10097	ISSGLDHQSR
540	10098	ISSGLDHQSK
541	10099	ISSGLHQSRS
542	10100	ISSGLHQSKS
543	10101	ISSGLHQSR
544	10102	ISSGLHQSK
545	10103	ISSGLLSDHQSRS
546	10104	ISSGLLSDHQSKS
547	10105	ISSGLLSDHQSR
548	10106	ISSGLLSDHQSK
549	10107	ISSGLSSDHQSRS
550	10108	ISSGLSSDHQSKS
551	10109	ISSGLSSDHQSR
552	10110	ISSGLSSDHQSK
553	10111	ISSGLSDHQSRS
554	10112	ISSGLSDHQSKS
555	10113	ISSGLSDHQSR
556	10114	ISSGLSDHQSK
557	10115	ISSGLLDHQSRS
558	10116	ISSGLLDHQSKS
559	10117	ISSGLLDHQSR
560	10118	ISSGLLDHQSK
561	10119	HQSRSISSGLLSS
562	10120	HQSKSISSGLLSS
563	10121	DHQSRSISSGLLS
564	10122	DHQSKSISSGLLS
565	10123	DHQSRSISSGLL
566	10124	DHQSKSISSGLL
567	10125	HQSRSISSGLL
568	10126	HQSKSISSGLL
569	10127	DHQSRSISSGL
570	10128	DHQSKSISSGL
571	10129	HQSRSISSGL

572	10130	HQSKSISSGL
573	10131	HQSRISSGL
574	10132	HQSKISSGL
575	10133	DHQSRSSGLLS
576	10134	DHQSRSSGLL
577	10135	DHQSRSSGL
578	10136	HQSRSSGL
579	10137	DHQSKSSGLLS
580	10138	DHQSKSSGLL
581	10139	DHQSKSSGL
582	10140	HQSKSSGL
583	10141	ALAHGLFSDHQSRSR
584	10142	ALAHGLFSDHQSRSG
585	10143	ALAHGLFSDHQSRSA
586	10144	ALAHGLFSDHQSR
587	10145	ALAHGLFSDHQSKS
588	10146	ALAHGLFSDHQSK
589	10147	ALAHGLFSHQSRS
590	10148	ALAHGLFSHQSKS
591	10149	ALAHGLFSHQSR
592	10150	ALAHGLFSHQSK
593	10151	ALAHGLFDHQSRS
594	10152	ALAHGLFDHQSKS
595	10153	ALAHGLFDHQSR
596	10154	ALAHGLFDHQSK
597	10155	ALAHGLSDHQSRS
598	10156	ALAHGLSDHQSKS
599	10157	ALAHGLSDHQSR
600	10158	ALAHGLSDHQSK
601	10159	ALAHGLFHQSRS
602	10160	ALAHGLFHQSKS
603	10161	ALAHGLFHQSR
604	10162	ALAHGLFHQSK
605	10163	ALAHGLDHQSRS
606	10164	ALAHGLDHQSKS
607	10165	ALAHGLDHQSR
608	10166	ALAHGLDHQSK
609	10167	ALAHGLHQSRS
610	10168	ALAHGLHQSKS

611	10169	ALAHGLHQSR
612	10170	ALAHGLHQSK
613	10171	LAHGLHQSRS
614	10172	LAHGLHQSKS
615	10173	LAHGLHQSR
616	10174	LAHGLHQSK
617	10175	SDHQSKSALAHGLF
618	10176	SDHQSRSALAHGL
619	10177	SDHQSRSLAHGLF
620	10178	DHQSRSALAHGLF
621	10179	DHQSKSALAHGLF
622	10180	HQSRSALAHGLF
623	10181	HQSKSALAHGLF
624	10182	DHQSRSALAHGL
625	10183	DHQSKSALAHGL
626	10184	HQSRSALAHGL
627	10185	HQSKSALAHGL
628	10186	HQSRSLAHGL
629	10187	HQSKSLAHGL
630	10188	HQSRALAHGL
631	10189	HQSKALAHGL
632	10190	HQSRLAHGL
633	10191	HQSKLAHGL
634	10192	DHQSRSLAHGLF
217	CX-122 Light Chain	QGQSGQ**CISPRGCPDGPYVMY**GSSGGSGGSGGSG**ISS GLLSGRSDNH**GSSGTQILLTQSPVILSVSPGERVSFSCR ASQSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSG SGSGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGT KLELKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFY PREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSS TLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
218	CX-122 Heavy Chain	QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWV RQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNS KSQVFFKMNSLQSQDTAIYYCARALTYYDYEFAYWG QGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLV KDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSS VVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE KTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKG FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK

		LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK
219	AA w/5007 Light Chain	QGQSGQ**CISPRGCPDGPYVMY**GGGSSGGS**APRSALA HGLF**GGGSQILLTQSPVILSVSPGERVSFSCRASQSIGT NIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTD FTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKV QWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK ADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
220	AA w/5007 Heavy Chain	QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWV RQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNS KSQVFFKMNSLQSQDTAIYYCARALTYYDYEFAYWG QGTLVTVS<u>A</u>ASTKGPSVFPLAPSSKSTSGGTAALGCLV KDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSS VVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE KTISKAKGQPREPQVYTLPPSR<u>D</u>ELTKNQVSLTCLVKG FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK
221	CX-229 Light Chain	QGQSGQ**CISPRGC<u>L</u>DGPYVMY**GSSGGSGGSGGSG**AV GLLAPPGGLSGRSDNH**GSSGTQILLTQSPVILSVSPGE RVSFSCRASQSIGTNIHWYQQRTNGSPRLLIKYASESIS GIPSRFSGSGSGTDFTLSINSVESEDIADYYCQQNNNWP TTFGAGTKLELKRTVAAPSVFIFPPSDEQLKSGTASVV CLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVT KSFNRGEC
	CX-229 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
222	ProC070 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**AIALYA DHQSRS**GGGSQILLTQSPVILSVSPGERVSFSCRASQSI GTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSG TDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLEL KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

223	ProC070 Heavy Chain	QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWV RQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNS KSQVFFKMNSLQSQDTAIYYCARALTYYDYEFAYWG QGTLVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLV KDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSS VVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKG FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK
224	ProC071 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**DHQSRS AIALYAD**GGGSQILLTQSPVILSVSPGERVSFSCRASQSI GTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSG TDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLEL KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC071 Heavy Chain	Same as SEQ ID NO:223
225	ProC1590 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**HQSRSA IALYAD**GGGSQILLTQSPVILSVSPGERVSFSCRASQSIG TNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGT DFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC1590 Heavy Chain	Same as SEQ ID NO:218
226	ProC1591 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**DHQSRS AIALYA**GGGSQILLTQSPVILSVSPGERVSFSCRASQSIG TNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGT DFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC1591 Heavy Chain	Same as SEQ ID NO:218
227	ProC1592 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**HQSRSA IALYA**GGGSQILLTQSPVILSVSPGERVSFSCRASQSIGT NIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTD FTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKV QWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK ADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

	ProC1592 Heavy Chain	Same as SEQ ID NO:218
228	ProC1593 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**HQSRSG PWGLL**GGGSQILLTQSPVILSVSPGERVSFSCRASQSIG TNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGT DFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC1593 Heavy Chain	Same as SEQ ID NO:218
229	ProC905 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**DHQSRS GPWGLL**GGGSQILLTQSPVILSVSPGERVSFSCRASQSI GTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSG TDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLEL KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC905 Heavy Chain	Same as SEQ ID NO:218
230	ProC907 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**GPWGL LDHQSRS**GGGSQILLTQSPVILSVSPGERVSFSCRASQS IGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSG TDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLEL KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC907	Same as SEQ ID NO:218
231	ProC072 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**ISSGLLS SDHQSRS**GGGSQILLTQSPVILSVSPGERVSFSCRASQSI GTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSG TDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLEL KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC072 Heavy Chain	Same as SEQ ID NO:223
232	ProC073 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**DHQSRS ISSGLLSS**GGGSQILLTQSPVILSVSPGERVSFSCRASQS IGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSG TDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLEL KRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREA KVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTL SKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC073 Heavy Chain	Same as SEQ ID NO:223

233	ProC074 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**DHQSRS SGLLSS**GGGSQILLTQSPVILSVSPGERVSFSCRASQSIG TNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGT DFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC074 Heavy Chain	Same as SEQ ID NO:223
234	ProC075 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**ALAHG LFSDHQSRS**GGGSQILLTQSPVILSVSPGERVSFSCRAS QSIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSG SGTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKL ELKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPRE AKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC075 Heavy Chain	Same as SEQ ID NO:223
235	ProC076 Light Chain	QGQSGQG**CISPRGCPDGPYVMY**GGGSSGGS**SDHQSR SALAHGLF**GGGSQILLTQSPVILSVSPGERVSFSCRASQ SIGTNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGS GTDFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLE LKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPRE AKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLT LSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC076 Heavy Chain	Same as SEQ ID NO:223
237		GPWGLL
238		SGLLSS
240	ProC3633 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSPWGLLH QSRSGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGTN IHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFT LSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRT VAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQ WKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKA DYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3633 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
241	ProC3634 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSPWGLHQS RSGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGTNIH WYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTL SINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRTV AAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQW KVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD YEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3634 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above

242	ProC3635 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSHQSRSPW GLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGTNIH WYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFTL SINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRTV AAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQW KVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKAD YEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3635 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
243	ProC3636 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSDHQSKSG PWGLLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIG TNIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGT DFTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAK VQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLS KADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3636 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
244	ProC3637 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSHQSRSGP WGLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGTNI HWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFT LSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRT VAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQ WKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKA DYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3637 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
245	ProC3638 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSHQSRSGP YGLLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGT NIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTD FTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKV QWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK ADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3638 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
246	ProC3639 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSHQSRSGP FGLLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGTN IHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFT LSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRT VAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQ WKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKA DYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3639 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
247	ProC3640 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSHQSRSGP RGLLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGTN IHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTDFT

		LSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKRT VAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQ WKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKA DYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3640 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
248	ProC3641 Light Chain	QGQSGQGCISPRGCPDGPYVMYGGGSSGGSHQSRSGP AGLLGGGSQILLTQSPVILSVSPGERVSFSCRASQSIGT NIHWYQQRTNGSPRLLIKYASESISGIPSRFSGSGSGTD FTLSINSVESEDIADYYCQQNNNWPTTFGAGTKLELKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKV QWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSK ADYEKHKVYACEVTHQGLSSPVTKSFNRGEC
	ProC3641 Heavy Chain	Same as CX-122 HC SEQ ID NO:218 above
249-394	Intentionally left blank	
395	Linking peptide	GGSG
396	Linking peptide	GGSGG
397	Linking peptide	GSGSG
398	Linking peptide	GSGGG
399	Linking peptide	GGGSG
400	Linking peptide	GSSSG
401	Linking peptide	GSSGGSGGSGG
402	Linking peptide	GGGS
403	Linking peptide	GGGSGGGS
404	Linking peptide	GGGSGGGSGGGS
405	Linking peptide	GGGGSGGGGSGGGGS
406	Linking peptide	GGGGSGGGGSGGGGSGGGGSGGGGS
407	Linking peptide	GGGGSGGGGS
408	Linking peptide	GGGGS
409	Linking peptide	GGGGSGS
410	Linking peptide	GGGGSGGGGSGGGGSGS
411	Linking peptide	GGSLDPKGGGGS
412	Linking peptide	PKSCDKTHTCPPCPAPELLG
413	Linking peptide	SKYGPPCPPCPAPEFLG
414	Linking peptide	GKSSGSGSESKS
415	Linking peptide	GSTSGSGKSSEGKG
416	Linking peptide	GSTSGSGKSSEGSGSTKG
417	Linking peptide	GSTSGSGKPGSGEGSTKG
418	Linking peptide	GSTSGSGKPGSSEGST
419	Linking peptide	GGGSSGGS

420	Linking peptide	GGGGSGGGGSS
421	Linking peptide	GGGSSGGSGGSSGGS
422	Linking peptide	GSTSGSGKPGSSEGST
423	Fc domain hole mutant	ESKYGPPCPPCPAPEFEGGPSVFLFPPKPKDTLMISRTPE VTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPREE QFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLPS SIEKTISKAKGQPREPQVCTLPPSQEEMTKNQVSLSCA VKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL VSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSLS LSLG *Note: the Fc may further comprise a lysine residue (K) at the C-terminus.
424	Fc domain knob mutant	GSSKYGPPCPPCPAPEFEGGPSVFLFPPKPKDTLMISRTP EVTCVVVDVSQEDPEVQFNWYVDGVEVHNAKTKPRE EQFNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKGLP SSIEKTISKAKGQPREPQVYTLPPCQEEMTKNQVSLWC LVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF LYSRLTVDKSRWQEGNVFSCSVMHEALHNHYTQKSL SLSLG* *Note: the Fc may further comprise a lysine residue (K) at the C-terminus.
425	Human IgG1	ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVT VSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSL GTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCP APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSH EDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG QPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAV EWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSR WQQGNVFSCSVMHEALHNHYTQKSLSLSPG* *Note: the Fc may further comprise a lysine residue (K) at the C-terminus.
426	Human IgG2	ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVTV SWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSNFG TQTYTCNVDHKPSNTKVDKTVERKCCVECPPCPAPPV AGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV QFNWYVDGVEVHNAKTKPREEQFNSTFRVVSVLTVV HQDWLNGKEYKCKVSNKGLPAPIEKTISKTKGQPREP QVYTLPPSREEMTKNQVSLTCLVKGFYPSDISVEWESN GQPENNYKTTPPMLDSDGSFFLYSKLTVDKSRWQQGN VFSCSVMHEALHNHYTQKSLSLSPG* *Note: the Fc may further comprise a lysine residue (K) at the C-terminus.
427	Human IgG3	ASTKGPSVFPLAPCSRSTSGGTAALGCLVKDYFPEPVT VSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSL GTQTYTCNVNHKPSNTKVDKRVELKTPLGDTTHTCPR CPEPKSCDTPPPCPRCPEPKSCDTPPPCPRCPEPKSCDTP PPCPRCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCV

		VVDVSHEDPEVQFKWYVDGVEVHNAKTKPREEQYNS TFRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT ISKTKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFY PSDIAVEWESSGQPENNYNTTPPMLDSDGSFFLYSKLT VDKSRWQQGNIFSCSVMHEALHNRFTQKSLSLSPG* *Note: the Fc may further comprise a lysine residue (K) at the C-terminus.
428	Human IgG4	ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPVT VSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSL GTKTYTCNVDHKPSNTKVDKRVESKYGPPCPSCPAPE FLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSQEDP EVQFNWYVDGVEVHNAKTKPREEQFNSTYRVVSVLT VLHQDWLNGKEYKCKVSNKGLPSSIEKTISKAKGQPR EPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIAVEW ESNGQPENNYKTTPPVLDSDGSFFLYSRLTVDKSRWQ EGNVFSCSVMHEALHNHYTQKSLSLSLG* *Note: the Fc may further comprise a lysine residue (K) at the C-terminus.
429		PRFKIIGG
430		PRFRIIGG
431		SSRHRRALD
432		RKSSIIIRMRDVVL
433		SSSFDKGKYKKGDDA
434		SSSFDKGKYKRGDDA
435		IEGR
436		IDGR
437		GGSIDGR
438		PLGLWA
439		GPQGIAGQ
440		GPQGLLGA
441		GIAGQ
442		GPLGIAGI
443		GPEGLRVG
444		YGAGLGVV
445		AGLGVVER
446		AGLGISST
447		EPQALAMS
448		QALAMSAI
449		AAYHLVSQ
450		MDAFLESS
451		ESLPVVAV
452		SAPAVESE
453		DVAQFVLT

454		VAQFVLTE
455		AQFVLTEG
456		PVQPIGPQ
457	Unmasked control c225v5 antibody heavy chain	QVQLKQSGPGLVQPSQSLSITCTVSGFSLTNYGVHWV RQSPGKGLEWLGVIWSGGNTDYNTPFTSRLSINKDNS KSQVFFKMNSLQSQDTAIYYCARALTYYDYEFAYWG QGTLVTVSAASTKGPSVFPLAPSSKSTSGGTAALGCLV KDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSS VVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVT CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQY NSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKG FYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSK LTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSP GK
458	Unmasked control c225v5 antibody light chain	QILLTQSPVILSVSPGERVSFSCRASQSIGTNIHWYQQR TNGSPRLLIKYASESISGIPSRFSGSGSGTDFTLSINSVES EDIADYYCQQNNNWPTTFGAGTKLELKRTVAAPSVFI FPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNAL QSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKV YACEVTHQGLSSPVTKSFNRGEC

OTHER EMBODIMENTS

[0450] It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

[0451] All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition section headings, the materials, methods, and examples are illustrative only and not intended to be limiting.

WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, wherein the CM1 comprises the amino acid sequence of HQSR (SEQ ID NO: 1), or HQSK (SEQ ID NO: 3).

2. The isolated polypeptide of claim 1, wherein the CM1 comprises the amino acid sequence of HQSRS (SEQ ID NO: 2).

3. The isolated polypeptide of claim 1, wherein the CM1 comprises the amino acid sequence of HQSRSA (SEQ ID NO: 4).

4. The isolated polypeptide of claim 1, wherein the CM1 comprises the amino acid sequence of DHQSRS (SEQ ID NO: 5).

5. The isolated polypeptide of claim 4, wherein the CM1 comprises the amino acid sequence of SDHQSRS (SEQ ID NO: 6).

6. The isolated polypeptide of any one of claims 1-5, wherein the CM2 comprises the amino acid sequence of AIAL (SEQ ID NO: 9).

7. The isolated polypeptide of any one of claims 1-6, wherein the CM2 comprises the amino acid sequence of AIALY (SEQ ID NO: 635).

8. The isolated polypeptide of any one of claims 1-6, wherein the CM2 comprises the amino acid sequence of AIALYA (SEQ ID NO: 8).

9. The isolated polypeptide of any one of claims 1-6, wherein the CM2 comprises the amino acid sequence of AIALYAD (SEQ ID NO: 7).

10. The isolated polypeptide of claim 1, wherein the substrate comprises the amino acid sequence of AIALYADHQSRS (SEQ ID NO: 10), DHQSRSAIALYAD (SEQ ID NO: 11), HQSRSAIALYAD (SEQ ID NO: 12), DHQSRSAIALYA (SEQ ID NO: 13), HQSRSAIALY (SEQ ID NO: 636), HQSRSAIAL (SEQ ID NO: 637),

AIALHQSRS (SEQ ID NO: 41), AIALYAHQSRS (SEQ ID NO: 33), AIALYHQSRS (SEQ ID NO: 37), HQSRSGPRGLL (SEQ ID NO: 638), HQSRSGPAGLL (SEQ ID NO: 639), HQSRPWGL (SEQ ID NO: 134), HQSRSPWG (SEQ ID NO: 640), HQSRAGPWGLL (SEQ ID NO: 645), HQSRSAPWGLL (SEQ ID NO: 646), HQSRSGAWGLL (SEQ ID NO: 647), HQSRSGPWALL (SEQ ID NO: 648), HQSRSGPWGAL (SEQ ID NO: 649), HQSRSGPWGLA(SEQ ID NO: 650), or HQSRSAIALYA (SEQ ID NO: 14), optionally wherein the substrate comprises the amino acid sequence of AIALYADHQSRS (SEQ ID NO: 10), DHQSRSAIALYAD (SEQ ID NO: 11), HQSRSAIALYAD (SEQ ID NO: 12), DHQSRSAIALYA (SEQ ID NO: 13), or HQSRSAIALYA (SEQ ID NO: 14).

11. The isolated polypeptide of claim 1, wherein the substrate comprises the amino acid sequence of DHQSRSAIALYA (SEQ ID NO: 13).

12. The isolated polypeptide of any one of claims 1-5, wherein the CM2 comprises the amino acid sequence of GPWGLL (SEQ ID NO: 237).

13. The isolated polypeptide of claim 12, wherein the substrate comprises the amino acid sequence of HQSRSGPWGLL (SEQ ID NO: 15), DHQSRSGPWGLL (SEQ ID NO: 16), or GPWGLLDHQSRS (SEQ ID NO: 17).

14. The isolated polypeptide of any one of claims 1-5, wherein the CM2 comprises the amino acid sequence of PXGL, wherein X is alanine (A), tyrosine (Y), arginine (R), phenylalanine (F), tryptophan (W), lysine (K), glutamine (Q), or methionine (M) or the CM2 comprises the amino acid sequence GPX_1GLX_2 (SEQ ID NO: 651), wherein X_1 is W, Y, F, R, K, Q, A, or M and wherein X_2 is any amino acid or absent.

15. The isolated polypeptide of claim 14, wherein the substrate comprises the amino acid sequence of HQSRSPWGL (SEQ ID NO: 126), PWGLHQSRS (SEQ ID NO: 489), PWGLLHQSRS (SEQ ID NO: 485), PWGLHQSR (SEQ ID NO: 502), PWGLHQSKS (SEQ ID NO: 490), HQSRPWGL (SEQ ID NO: 134), DHQSKSGPWGLL (SEQ ID NO: 178), HQSRSGPWGL (SEQ ID NO: 122),

HQSRSGPYGLL (SEQ ID NO: 138), HQSRSGPFGLL (SEQ ID NO: 158), HQSRSGPRGLL (SEQ ID NO: 638), or HQSRSGPAGLL (SEQ ID NO: 639).

16. The isolated polypeptide of any one of claims 1-5, wherein the CM2 comprises the amino acid sequence of SGLLSS (SEQ ID NO: 238).

17. The isolated polypeptide of claim 16, wherein the CM2 comprises the amino acid sequence of ISSGLLSS (SEQ ID NO: 75).

18. The isolated polypeptide of claim 16, wherein the substrate comprises the amino acid sequence of ISSGLLSSDHQSRS (SEQ ID NO: 18), DHQSRSISSGLLSS (SEQ ID NO: 19), or DHQSRSSGLLSS (SEQ ID NO: 20).

19. The isolated polypeptide of any one of claims 1-5, wherein the CM2 comprises the amino acid sequence of ALAHGLF (SEQ ID NO: 74).

20. The isolated polypeptide of claim 19, wherein the substrate comprises the amino acid sequence of ALAHGLFSDHQSRS (SEQ ID NO: 21) or SDHQSRSALAHGLF (SEQ ID NO: 22).

21. An isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease or set of proteases and a second cleavable moiety (CM2) cleavable by a second protease or set of proteases, wherein the substrate comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, 636-650, optionally SEQ ID NOs: 10-73, 83-216, and 460-634.

22. The isolated polypeptide of any one of claims 1-9, 12, 14, 16, 17, and 19, wherein the CM1 and the CM2 has a structural arrangement from N-terminus to C-terminus as CM1-CM2.

23. The isolated polypeptide of any one of claims 1-9, 12, 14, 16, 17, and 19, wherein the CM1 and the CM2 has a structural arrangement from N-terminus to C-terminus as CM2-CM1.

24. The isolated polypeptide of claim 22 or 23, wherein the CM1 and the CM2 are indirectly coupled via a linking peptide.

25. The isolated polypeptide of claim 22 or 23, wherein the CM1 and the CM2 are directly coupled to each other.

26. The isolated polypeptide of any one of claims 1-9, 12, 14, 16, 17, and 19, wherein at least a portion of the CM1 overlaps with at least a portion of the CM2.

27. An isolated polypeptide comprising a substrate, wherein the substrate comprises a first cleavable moiety (CM1) cleavable by a first protease and a second cleavable moiety (CM2) cleavable by a second protease, and wherein the substrate comprises an amino acid sequence with one-amino acid or two-amino acid mutation(s) in any one of SEQ ID NOs: 10-73, 83-216, 460-634, 636-650, optionally SEQ ID NOs: 10-73, 83-216, and 460-634.

28. The isolated polypeptide of any one of claims 1-27, wherein the isolated polypeptide is an activatable molecule and further comprises an active moiety (AM) that specifically binds a target.

29. The isolated polypeptide of claim 28, wherein the AM is an antibody or antigen binding fragment thereof.

30. The isolated polypeptide of claim 29, wherein the antigen binding fragment thereof is selected from the group consisting of a Fab fragment, a F(ab')$_2$ fragment, a scFv, a scAb, a dAb, a single domain heavy chain antibody, and a single domain light chain antibody.

31. The isolated polypeptide of claim 28, wherein the AM is a therapeutic macromolecule.

32. The isolated polypeptide of claim 28, wherein the AM is a cytokine or a chimeric antigen receptor.

33. The isolated polypeptide of any one of claims 28-32, wherein the AM is coupled to the CM1.

34. The isolated polypeptide of claim 33, wherein the AM is directly coupled to the CM1.

35. The isolated polypeptide of claim 33, wherein the AM is indirectly coupled to the CM1 via a linking peptide.

36. The isolated polypeptide of any one of claims 28-35, wherein the AM is coupled to the CM2.

37. The isolated polypeptide of claim 36, wherein the AM is directly coupled to the CM2.

38. The isolated polypeptide of claim 36, wherein the AM is indirectly coupled to the CM2 via a linking peptide.

39. The isolated polypeptide of any one of claims 27-38, further comprising a masking moiety (MM).

40. The isolated polypeptide of claim 39, wherein the MM has a dissociation constant for binding to the AM that is greater than the dissociation constant of the AM for binding to the target.

41. The isolated polypeptide of claim 39 or 40, wherein the MM is 2 to 40 amino acids in length.

42. The isolated polypeptide of any one of claims 39-41, wherein the MM is linked to the substrate such that the isolated polypeptide comprises a structural arrangement from N-terminus to C-terminus as follows: MM-CM1-CM2-AM, MM-CM2-CM1-AM, AM-CM1-CM2-MM, or AM-CM2-CM1-MM.

43. The isolated polypeptide of claim 42, wherein the MM is directly coupled to the CM1.

44. The isolated polypeptide of claim 42, wherein the MM is indirectly coupled to the CM1 via a linking peptide.

45. The isolated polypeptide of claim 42, wherein the MM is directly coupled to the CM2.

46. The isolated polypeptide of claim 42, wherein the MM is indirectly coupled to the CM2 via a linking peptide.

47. The isolated polypeptide of any one of claims 39-46, wherein the isolated polypeptide comprises a first linking peptide (LP1) and a second linking peptide (LP2), and wherein the isolated polypeptide has a structural arrangement from N-terminus to C-terminus as follows: MM- LP1-CM1-CM2-LP2-AM, MM-LP1-CM2-CM1-LP2-AM, AM-LP2-CM1-CM2-LP1-MM, or AM-LP2-CM2-CM1-LP1-MM.

48. The isolated polypeptide of claim 47, wherein the LP1 and LP2 are not identical to each other.

49. The isolated polypeptide of claim 47, wherein the LP1 and LP2 are identical to each other.

50. The isolated polypeptide of any one of claims 47-49, wherein each of the LP1 and LP2 is a peptide of 1 to 20 amino acids in length.

51. The isolated polypeptide of any one of claims 1-50, wherein the first protease is a serine protease.

52. The isolated polypeptide of claim 51, wherein the serine protease is membrane type serine protease 1 (MT-SP1).

53. The isolated polypeptide of claim 52, wherein the k_{cat}/K_M of the substrate by MT-SP1 cleavage is at least 1×10^3 $M^{-1}s^{-1}$.

54. The isolated polypeptide of claim 52, wherein the k_{cat}/K_M of the substrate by MT-SP1 cleavage is at least 1×10^4 $M^{-1}s^{-1}$.

55. The isolated polypeptide of any one of claims 1-54, wherein the second protease is a matrix metalloproteinase (MMP).

56. The isolated polypeptide of claim 55, wherein the MMP is MMP2, MMP9, or MMP14.

57. The isolated polypeptide of claim 56, wherein the k_{cat}/K_M of the substrate by MMP2 cleavage is at least 1×10^3 $M^{-1}s^{-1}$.

58. The isolated polypeptide of claim 56, wherein the k_{cat}/K_M of the substrate by MMP2 cleavage is at least 1×10^4 $M^{-1}s^{-1}$.

59. The isolated polypeptide of claim 56, wherein the k_{cat}/K_M of the substrate by MMP9 cleavage is at least 1×10^3 $M^{-1}s^{-1}$.

60. The isolated polypeptide of claim 56, wherein the k_{cat}/K_M of the substrate by MMP9 cleavage is at least 1×10^4 $M^{-1}s^{-1}$.

61. The isolated polypeptide of claim 54, wherein the k_{cat}/K_M of the substrate by MMP14 cleavage is at least 1×10^3 $M^{-1}s^{-1}$.

62. The isolated polypeptide of claim 56, wherein the k_{cat}/K_M of the substrate by MMP14 cleavage is at least 1×10^4 $M^{-1}s^{-1}$.

63. The isolated polypeptide of any one of claims 1-62, wherein the substrate is resistant to cleavage *in situ* in human bone marrow.

64. The isolated polypeptide of any one of claims 1-62, wherein the substrate is resistant to cleavage *in vivo* in human bone marrow.

65. The isolated polypeptide of any one of claims 1-64, further comprising one or more additional cleavable moieties, optionally wherein at least a portion of the substrate overlaps with at least a portion of a second cleavable moiety.

66. A polypeptide complex comprising one or more of the isolated polypeptides of any one of claims 1-65 bound to a second isolated polypeptide.

67. A conjugated polypeptide comprising the isolated polypeptide of any one of claims 1-65 conjugated to an agent.

68. The conjugated polypeptide of claim 67, wherein the agent is conjugated to the AM via a conjugating linker.

69. The conjugated polypeptide of claim 68, wherein the conjugating linker is cleavable.

70. The conjugated polypeptide of claim 68, wherein the conjugating linker is non-cleavable.

71. The conjugated polypeptide of claim 69, wherein the conjugating linker comprises an amino acid sequence selected from SEQ ID NOs: 10-73, 83-216, 460-634, 636-650.

72. The conjugated polypeptide of any one of claims 67-71, wherein the agent is a toxin, a microtubule inhibitor, a nucleic acid damaging agent, a dolastatin, an auristatin, a maytansinoid, a duocarmycin, a calicheamicin, or a combination thereof.

73. A composition comprising the isolated polypeptide of any one of claims 1-65, the polypeptide complex of claim 66, the conjugated polypeptide of any one of claims 67-72, and a carrier.

74. The composition of claim 73, wherein the carrier is a pharmaceutically acceptable carrier.

75. The composition of claim 74, comprising an additional agent.

76. The composition of claim 75, wherein the additional agent is a therapeutic agent.

77. An isolated nucleic acid molecule encoding the isolated polypeptide of any one of claims 1-65.

78. A vector comprising the isolated nucleic acid molecule of claim 77.

79. A cell comprising the isolated nucleic acid molecule of claim 77 or the vector of claim 78.

80. A method of manufacturing an activatable molecule that contains a tandem substrate, the method comprising expressing and recovering an activatable molecule comprising the isolated polypeptide of any one of claims 1-65.

81. A method of treating, alleviating a symptom of, or delaying the progression of a disease or disorder in a subject, comprising administering a therapeutically effective amount of the isolated polypeptide of any one of claims 1-65, the polypeptide complex of claim 66, the conjugated polypeptide of any one of claims 67-72, or the composition of any one of claims 73-76 to the subject.

82. The method of claim 81, wherein the disease is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

83. The isolated polypeptide of any one of claims 1-65, the polypeptide complex of claim 66, or the conjugated polypeptide of any one of claims 67-72, the composition of any one of claims 73-76, the nucleic acid molecule of claim 77, the vector of claim 78, or the cell of claim 79 for use as a medicament or for use in therapy, optionally for treating a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder, optionally with an additional agent which is optionally a therapeutic agent.

84. A method of detecting or diagnosing a disease or health condition of a subject, comprising:

contacting the isolated polypeptide of any one of claims 1-65, the polypeptide complex of claim 66, the conjugated polypeptide of any one of claims 67-72, or the composition of any one of claims 73-76 with a sample from the subject; and

measuring a level of cleavage of the isolated polypeptide, thereby detecting or diagnosing the disease or health condition of the subject.

85. The method of claim 84, wherein the disease is a cancer, an infection, an inflammatory disorder, a cardiovascular disorder, a neurodegenerative disorder, or an autoimmune disorder.

In vitro Masking Efficiency of anti-EGFR Activatable Antibodies with Exemplary Substrates

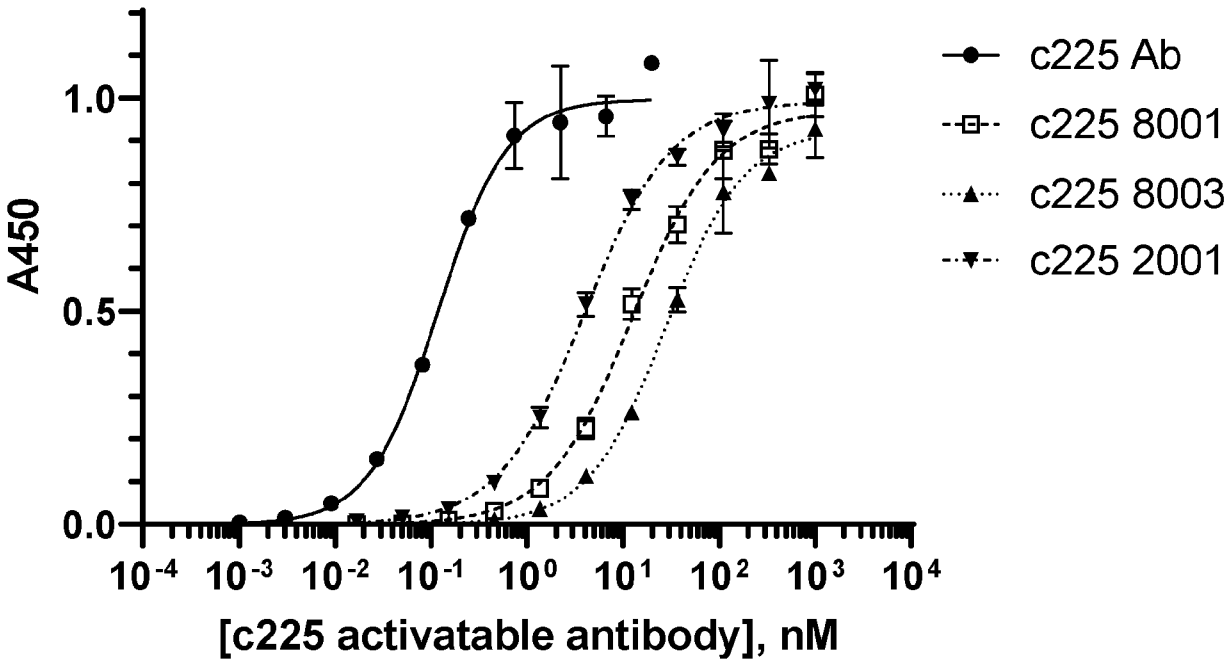

Fig. 1A

In vitro Masking Efficiency of anti-EGFR Activatable Antibodies with Exemplary Substrates

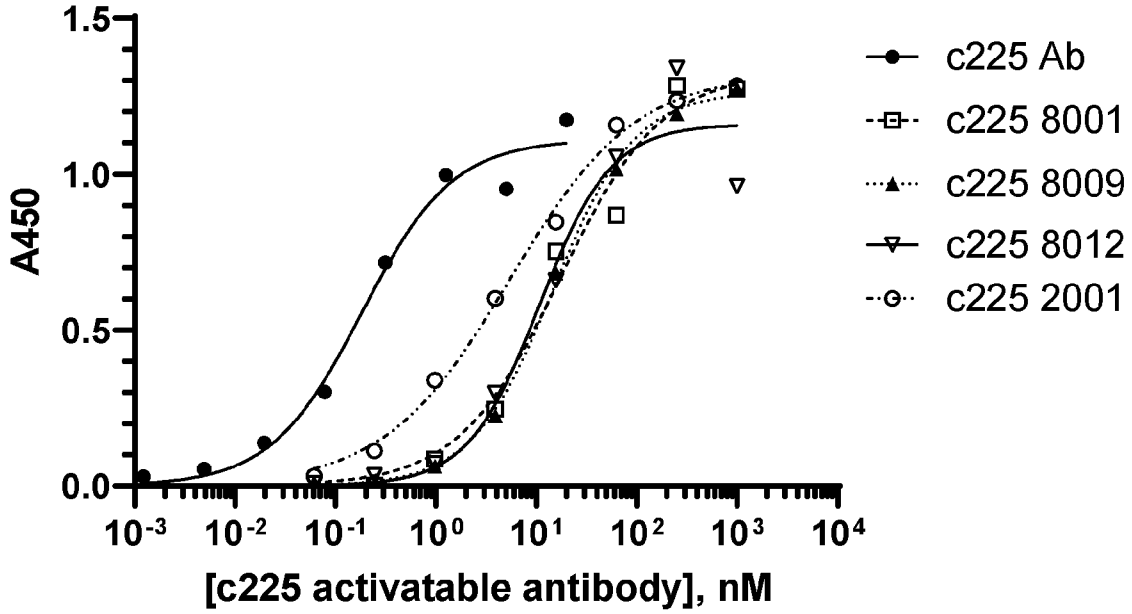

Fig. 1B

In vitro Masking Efficiency of anti-EGFR Activatable Antibodies with Exemplary Substrates

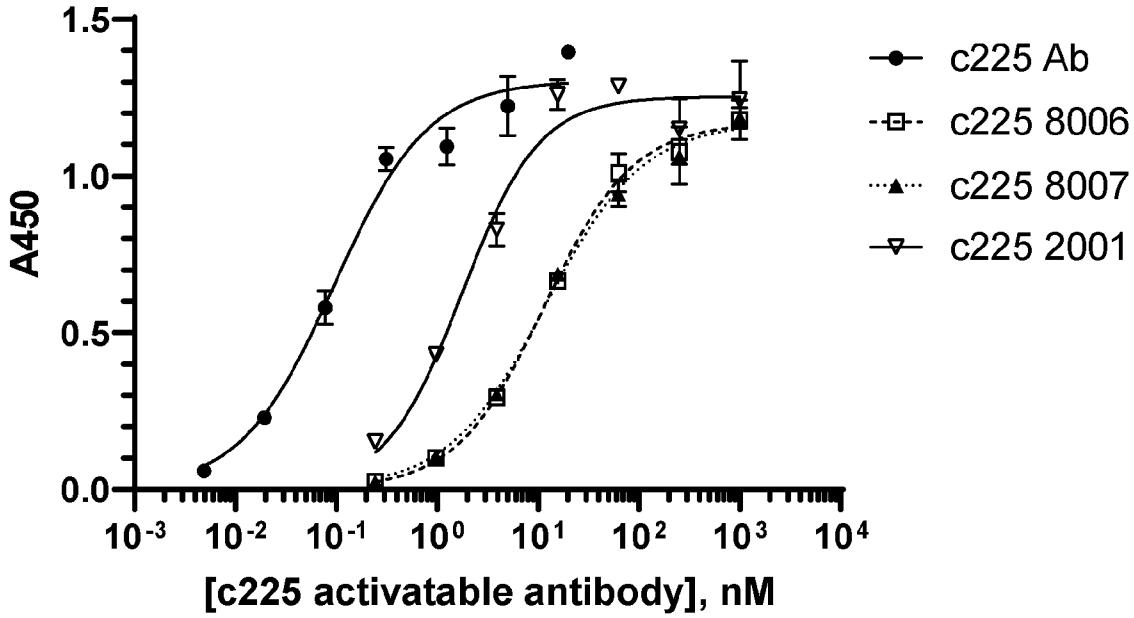

Fig. 1C

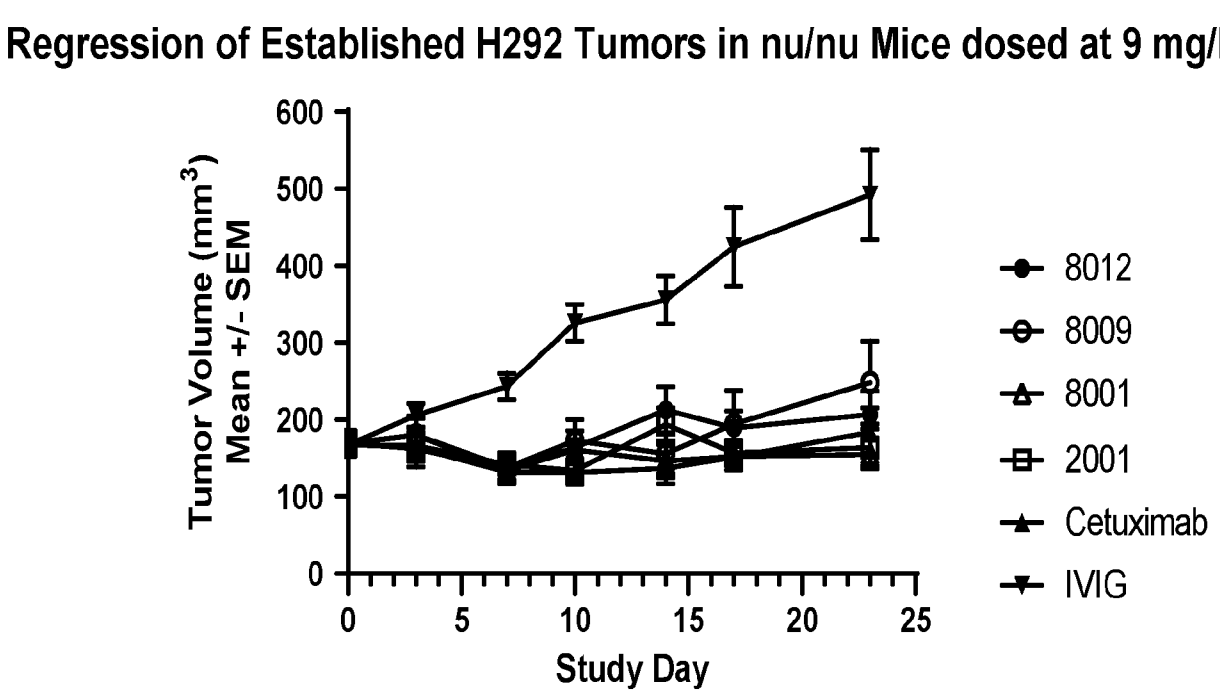
Regression of Established H292 Tumors in nu/nu Mice dosed at 9 mg/kg
Tumor Volume (mm3)
Mean +/- SEM
600
500
400
300
200
100
0
0
5
10
15
20
25
Study Day
8012
8009
8001
2001
Cetuximab
IVIG

FIG. 2A

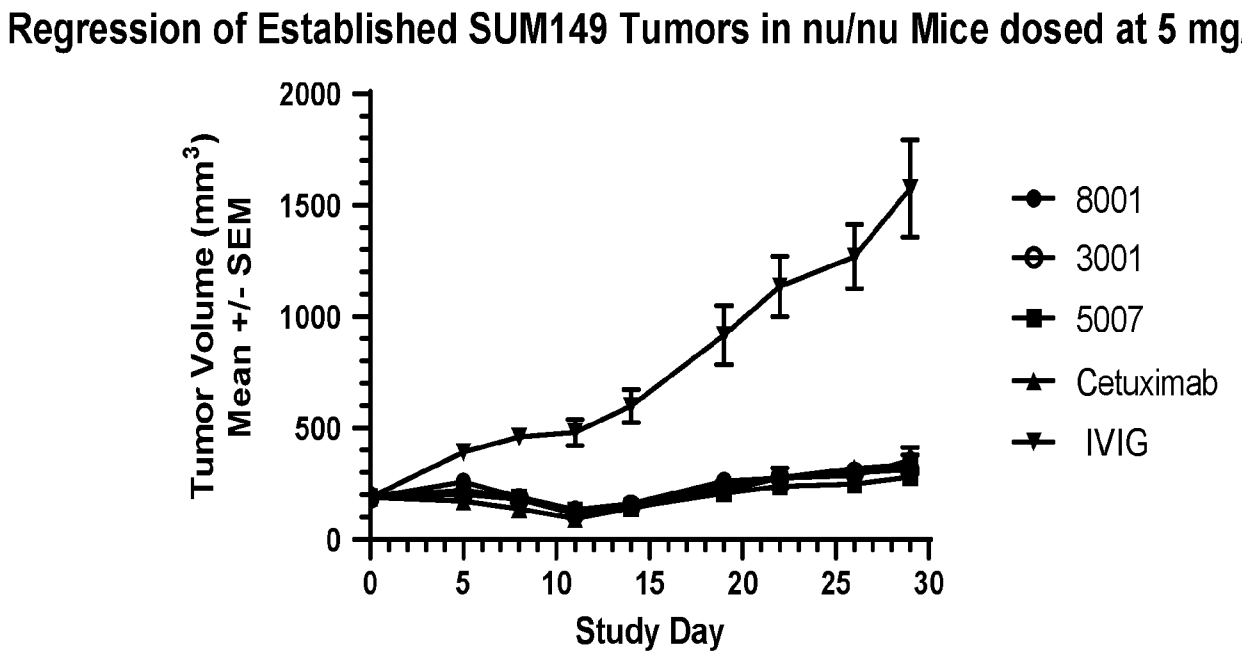
Regression of Established SUM149 Tumors in nu/nu Mice dosed at 5 mg/kg
Tumor Volume (mm3)
Mean +/- SEM
2000
1500
1000
500
0
0
5
10
15
20
25
30
Study Day
8001
3001
5007
Cetuximab
IVIG

FIG. 2B

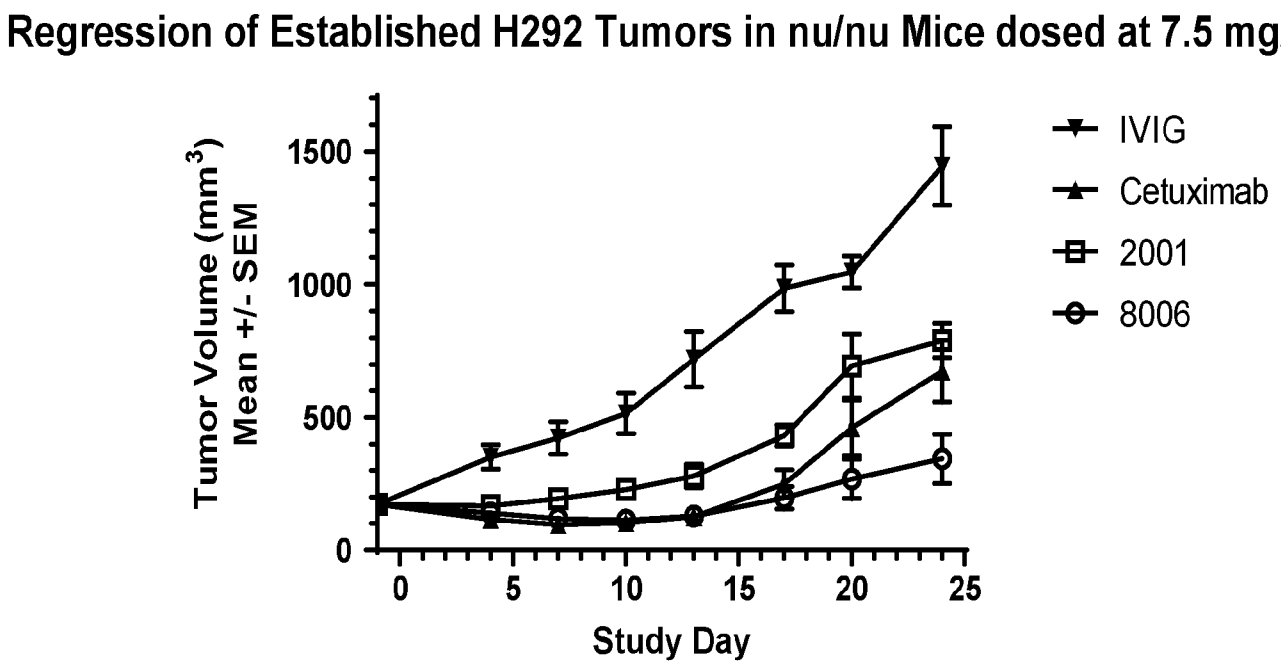
Regression of Established H292 Tumors in nu/nu Mice dosed at 7.5 mg/kg
Tumor Volume (mm3)
Mean +/- SEM
1500
1000
500
0
0
5
10
15
20
25
Study Day
IVIG
Cetuximab
2001
8006

FIG. 2C

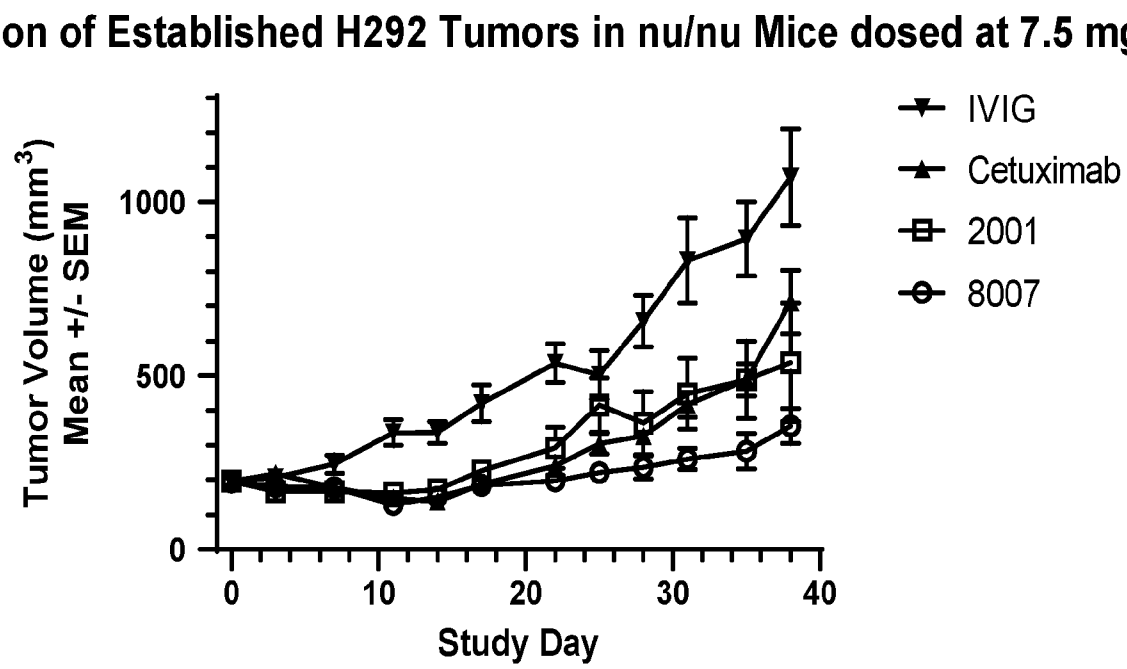
Regression of Established H292 Tumors in nu/nu Mice dosed at 7.5 mg/kg
Tumor Volume (mm3)
Mean +/- SEM
1000
500
0
0
10
20
30
40
Study Day
IVIG
Cetuximab
2001
8007

FIG. 2D

Regression of Established H292 Tumors in nu/nu Mice dosed at 7.5 mg/kg

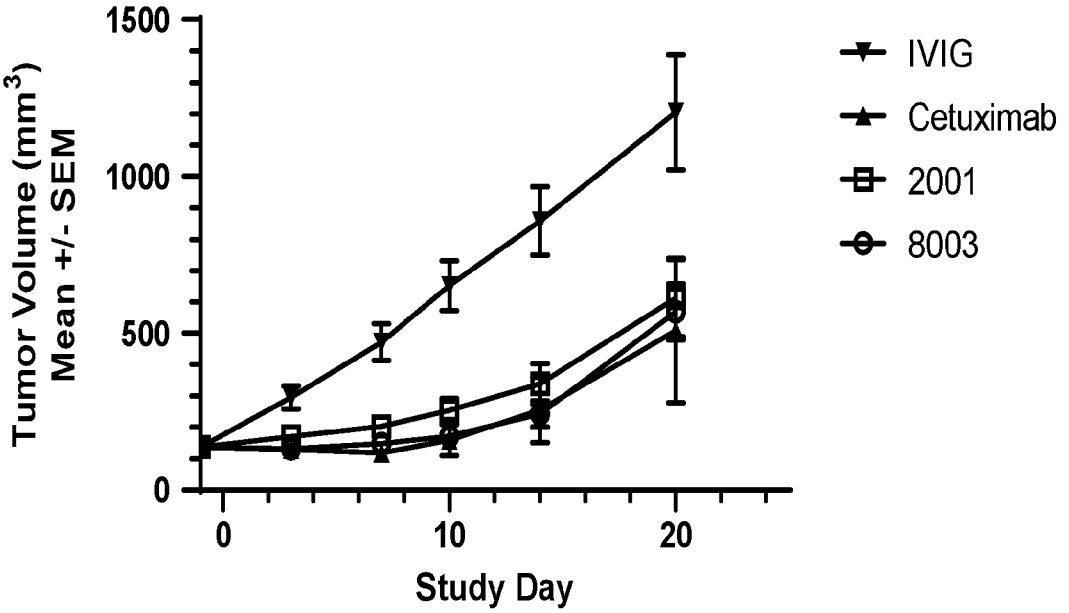

FIG. 2E

In Vivo Intratumoral Activation of c225 Activatable Antibodies in Established H292 Tumors at Day 8

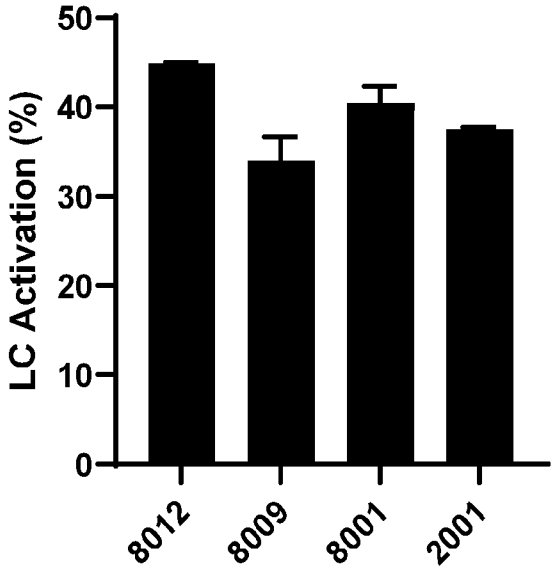

FIG. 2F

In Vivo Intratumoral Activation of c225 Activatable Antibodies in Established H292 Tumors at Day 7

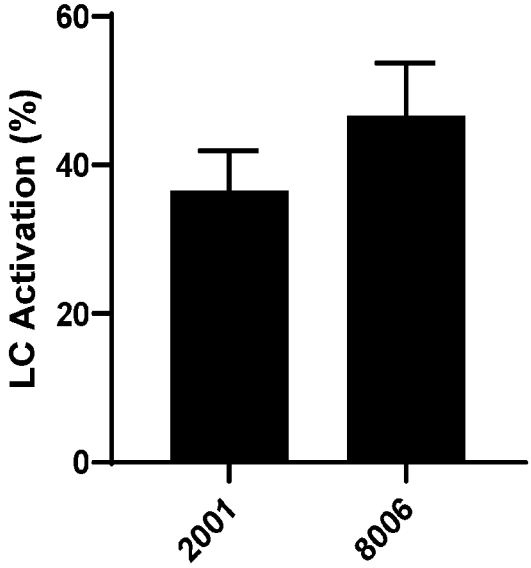

FIG. 2G

In Vivo Intratumoral Activation of c225 Activatable Antibodies in Established H292 Tumors at Day 8

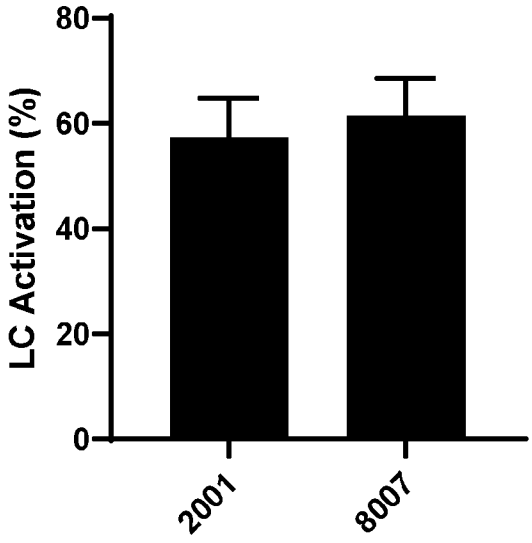

FIG. 2H

In Vivo Intratumoral Activation of c225 Activatable Antibodies in Established H292 Tumors at Day 4

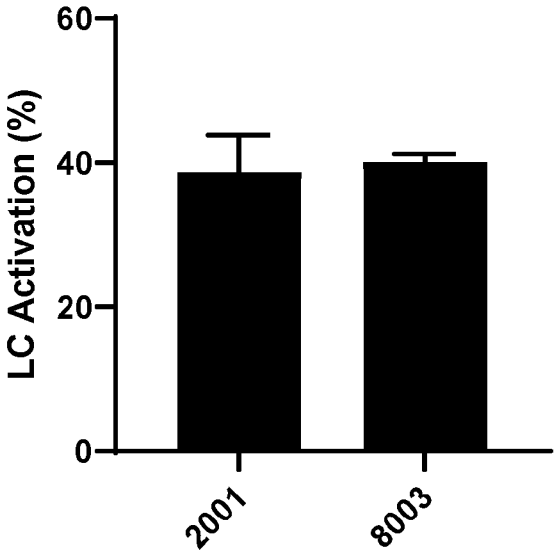

FIG. 2I

In Situ Stability of anti-EGFR Activatable Antibodies in Human Bone Marrow Aspirates

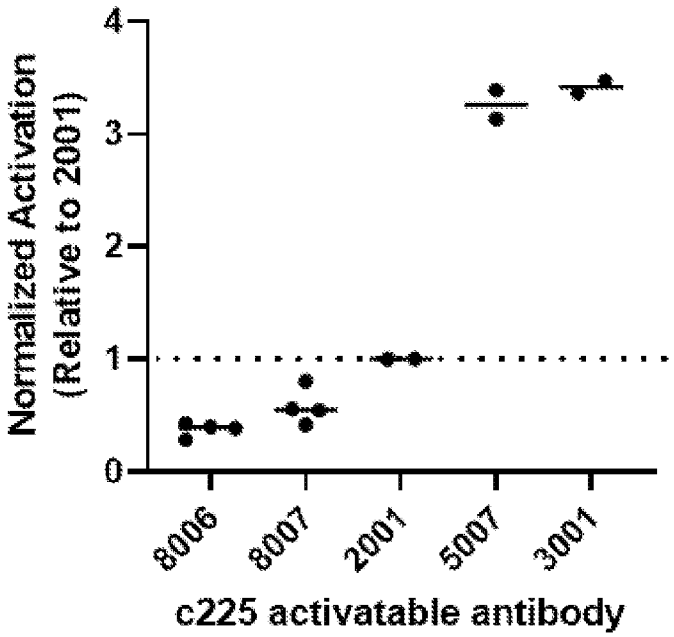

FIG. 3

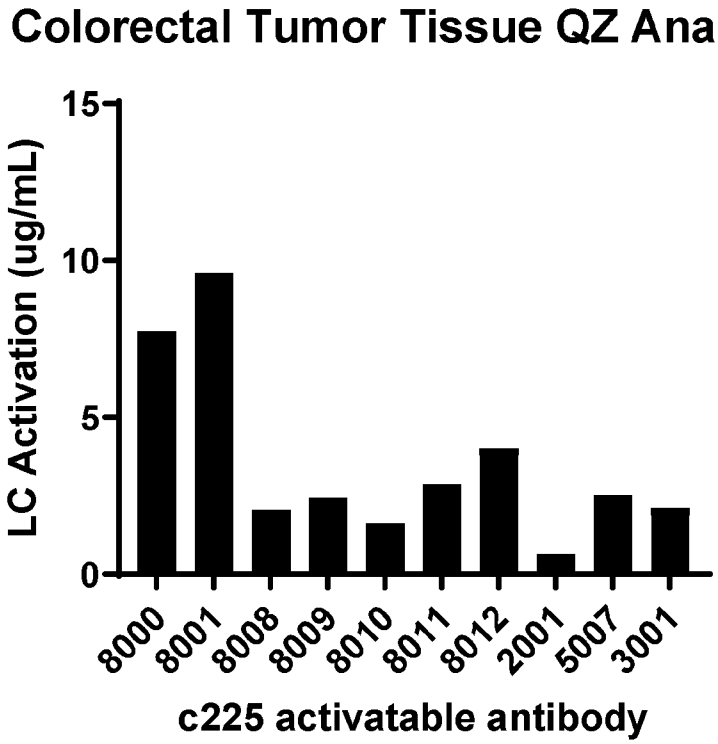
Colorectal Tumor Tissue QZ Analysis
LC Activation (ug/mL)
15
10
5
0
8000
8001
8008
8009
8010
8011
8012
2001
5007
3001
c225 activatable antibody

FIG. 4A

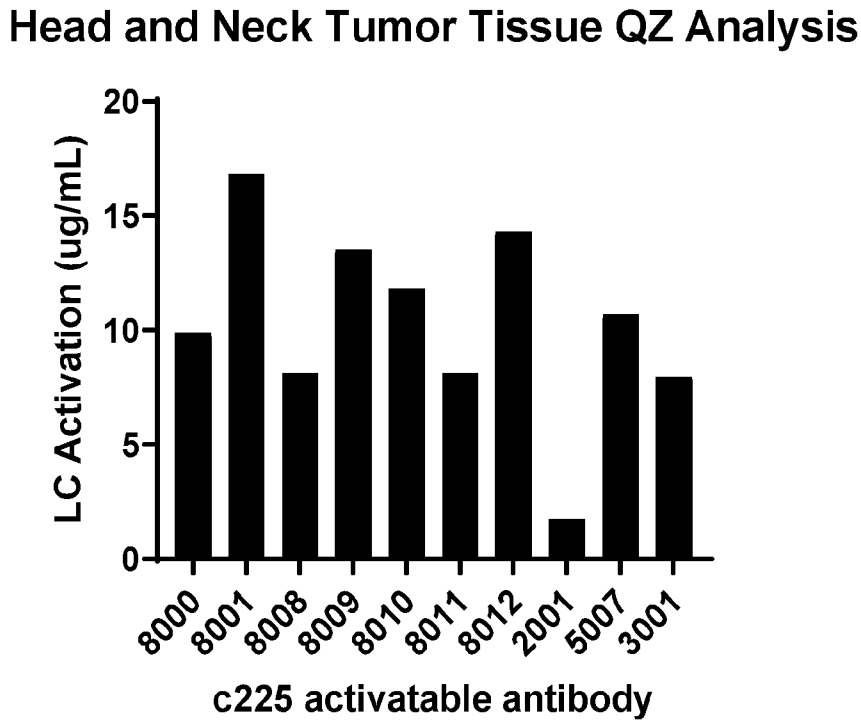
Head and Neck Tumor Tissue QZ Analysis
LC Activation (ug/mL)
20
15
10
5
0
8000
8001
8008
8009
8010
8011
8012
2001
5007
3001
c225 activatable antibody

FIG. 4B

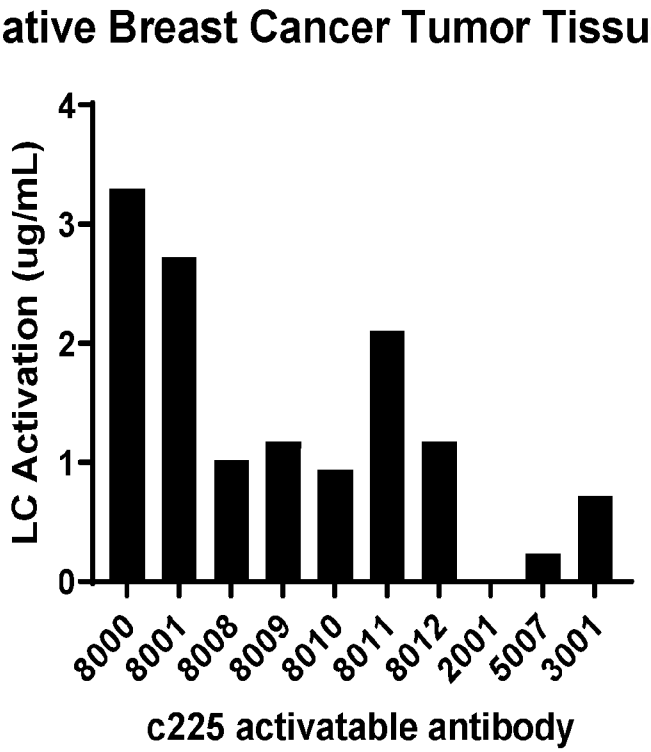
Triple Negative Breast Cancer Tumor Tissue QZ Analysis
LC Activation (ug/mL)
0
1
2
3
4
8000
8001
8008
8009
8010
8011
8012
2001
5007
3001
c225 activatable antibody

FIG. 4C

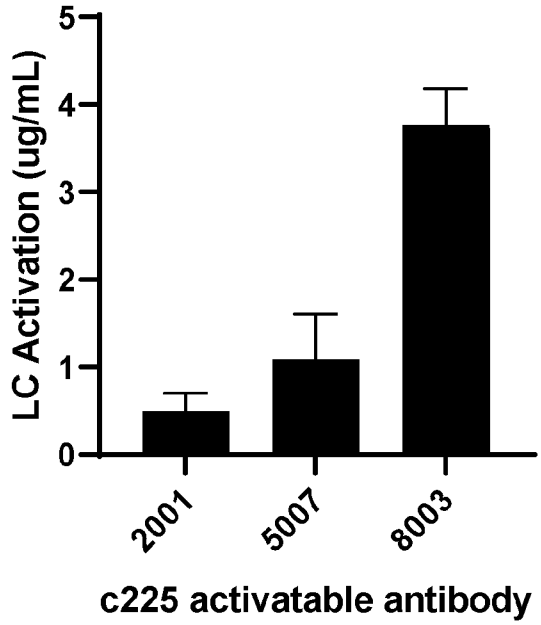
Multiple Myeloma QZ Analysis
LC Activation (ug/mL)
0
1
2
3
4
5
2001
5007
8003
c225 activatable antibody

FIG. 5A

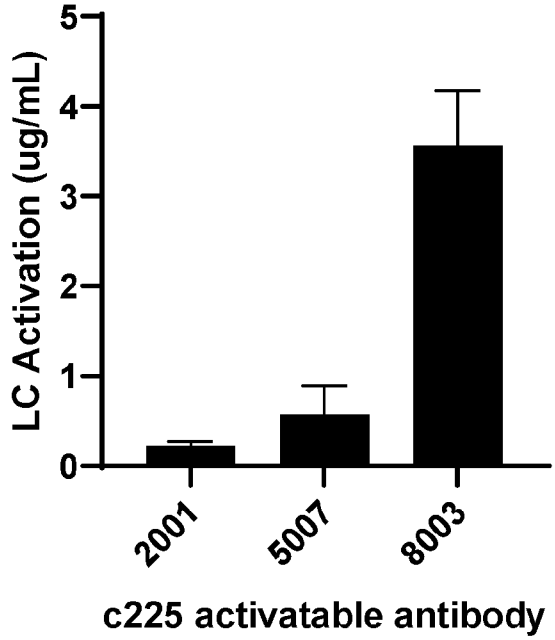
Acute Myeloid Leukemia QZ Analysis
LC Activation (ug/mL)
0
1
2
3
4
5
2001
5007
8003
c225 activatable antibody

FIG. 5B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/071326

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61P35/00 C12N9/64 C12N15/62
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61P C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data, EMBL, BIOSIS, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE UniProt [Online] 1 October 2014 (2014-10-01), "SubName: Full=Putative aromatic di-alanine and TPR containing protein {ECO:0000313|EMBL:KEP46916.1};", XP93095538, retrieved from EBI accession no. UNIPROT:A0A074RII3 Database accession no. A0A074RII3 sequence ----- -/--	1,8,22, 51-64, 77-79

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 October 2023	13/11/2023

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Kools, Patrick

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/071326

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE UniProt [Online] 11 May 2016 (2016-05-11), "SubName: Full=ATP-dependent bile acid permease {ECO:0000313|EMBL:KXN92134.1};", XP93095548, retrieved from EBI accession no. UNIPROT:A0A137QY42 Database accession no. A0A137QY42 sequence ----- DATABASE UniProt [Online] 5 October 2016 (2016-10-05), "SubName: Full=Uncharacterized protein {ECO:0000313|EMBL:OAQ32481.1};", XP93095423, retrieved from EBI accession no. UNIPROT:A0A197K6Z3 Database accession no. A0A197K6Z3 sequence ----- DATABASE UniProt [Online] 29 September 2021 (2021-09-29), "RecName: Full=thioredoxin-dependent peroxiredoxin {ECO:0000256|ARBA:ARBA00013017}; EC=1.11.1.24 {ECO:0000256|ARBA:ARBA00013017};", XP93095433, retrieved from EBI accession no. UNIPROT:A0A836AHN9 Database accession no. A0A836AHN9 sequence -----	1,22, 51-64, 77-79 1,10,22, 51-64, 77-79 1,14,22, 51-64, 77-79
Y	WO 2015/116933 A2 (CYTOMX THERAPEUTICS INC [US]) 6 August 2015 (2015-08-06) cited in the application Whole document especially table 8a, 9a and 0147-148, 150, 152, 153 -----	1-85
Y	WO 2015/048329 A2 (CYTOMX THERAPEUTICS INC [US]) 2 April 2015 (2015-04-02) cited in the application Whole document, especially tables 8A-8M. ----- -/--	1-85

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DARRAGH MOLLY R ET AL: "MT-SP1 proteolysis and regulation of cell-microenvironment interactions", FRONTIERS IN BIOSCIENCE, FRONTIERS IN BIOSCIENCE, ALBERTSON, NY, US, vol. 13, 1 January 2008 (2008-01-01), pages 528-539, XP002786599, ISSN: 1093-9946 the whole document	1-85
A	LEVESQUE JEAN-PIERRE ET AL: "Characterization of hematopoietic progenitor mobilization in protease-deficient mice", BLOOD, vol. 104, no. 1, 1 July 2004 (2004-07-01), pages 65-72, XP093095092, US ISSN: 0006-4971, DOI: 10.1182/blood-2003-05-1589 the whole document	1-85
A	ELTER ADRIAN ET AL: "Protease-Activation of Fc-Masked Therapeutic Antibodies to Alleviate Off-Tumor Cytotoxicity", FRONTIERS IN IMMUNOLOGY, vol. 12, 3 August 2021 (2021-08-03), XP093048541, DOI: 10.3389/fimmu.2021.715719 the whole document	1-85
A	LUCCHI ROBERTA ET AL: "The Masking Game: Design of Activatable Antibodies and Mimetics for Selective Therapeutics and Cell Control", ACS CENTRAL SCIENCE, vol. 7, no. 5, 26 April 2021 (2021-04-26), pages 724-738, XP055827094, ISSN: 2374-7943, DOI: 10.1021/acscentsci.0c01448 the whole document	1-85

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/071326

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2023/071326

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2015116933 A2	06-08-2015	AU 2015210862 A1	11-08-2016
		AU 2020205249 A1	06-08-2020
		BR 112016017649 A2	10-10-2017
		CA 2937370 A1	06-08-2015
		CN 106459153 A	22-02-2017
		CN 114106099 A	01-03-2022
		DK 3628328 T3	05-12-2022
		EP 3099319 A2	07-12-2016
		EP 3628328 A1	01-04-2020
		EP 4176896 A1	10-05-2023
		ES 2932777 T3	25-01-2023
		IL 270562 A	30-01-2020
		IL 294782 A	01-09-2022
		IL 302642 A	01-07-2023
		JP 6835586 B2	24-02-2021
		JP 7337864 B2	04-09-2023
		JP 2017509594 A	06-04-2017
		JP 2021072848 A	13-05-2021
		KR 20160113717 A	30-09-2016
		NZ 722401 A	30-06-2023
		RU 2016135239 A	12-03-2018
		US 2015218217 A1	06-08-2015
		US 2017204139 A1	20-07-2017
		US 2019135864 A1	09-05-2019
		US 2022411470 A1	29-12-2022
		WO 2015116933 A2	06-08-2015
WO 2015048329 A2	02-04-2015	AU 2014324884 A1	21-04-2016
		AU 2020203910 A1	02-07-2020
		BR 112016006665 A2	12-09-2017
		CA 2925106 A1	02-04-2015
		CA 3214529 A1	02-04-2015
		CN 106163556 A	23-11-2016
		EP 3049111 A2	03-08-2016
		IL 291329 A	01-05-2022
		JP 6915987 B2	11-08-2021
		JP 2016538240 A	08-12-2016
		JP 2020015749 A	30-01-2020
		JP 2021138707 A	16-09-2021
		KR 20160074510 A	28-06-2016
		KR 20230011497 A	20-01-2023
		NZ 718283 A	27-05-2022
		NZ 756892 A	27-05-2022
		RU 2016115542 A	30-10-2017
		RU 2020106752 A	11-03-2020
		US 2015087810 A1	26-03-2015
		US 2019284283 A1	19-09-2019
		WO 2015048329 A2	02-04-2015