Claims

[Claim 1] A modified protein, wherein the modified protein includes a protein moiety and a modifying moiety, the protein moiety is optionally linked to PAS at the N-terminus, the modifying moiety is connected directly or via a connector to a cysteine residue in the protein moiety or a cysteine residue in the PAS, and the cysteine residue is either naturally occurring or introduced through mutation.

[Claim 2] The modified protein of claim 1, wherein the protein is fibroblast growth factor 21.

[Claim 3] The modified protein of claim 1, wherein the modifying moiety is an albumin-binding agent;

preferably, the structure of the albumin-binding agent is shown in the following formula:

$$R_1\text{-}(X_1)_m\text{-}(X_2)_n\text{-}K(\varepsilon\text{-}R_2);$$

wherein K is a lysine;

R_1 is a substituted or unsubstituted $C_{1\text{-}20}$ acyl;

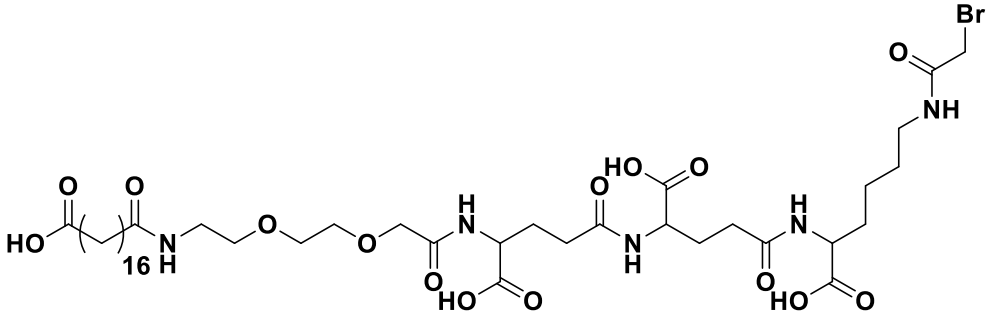
m is an integer from 0 to 5;

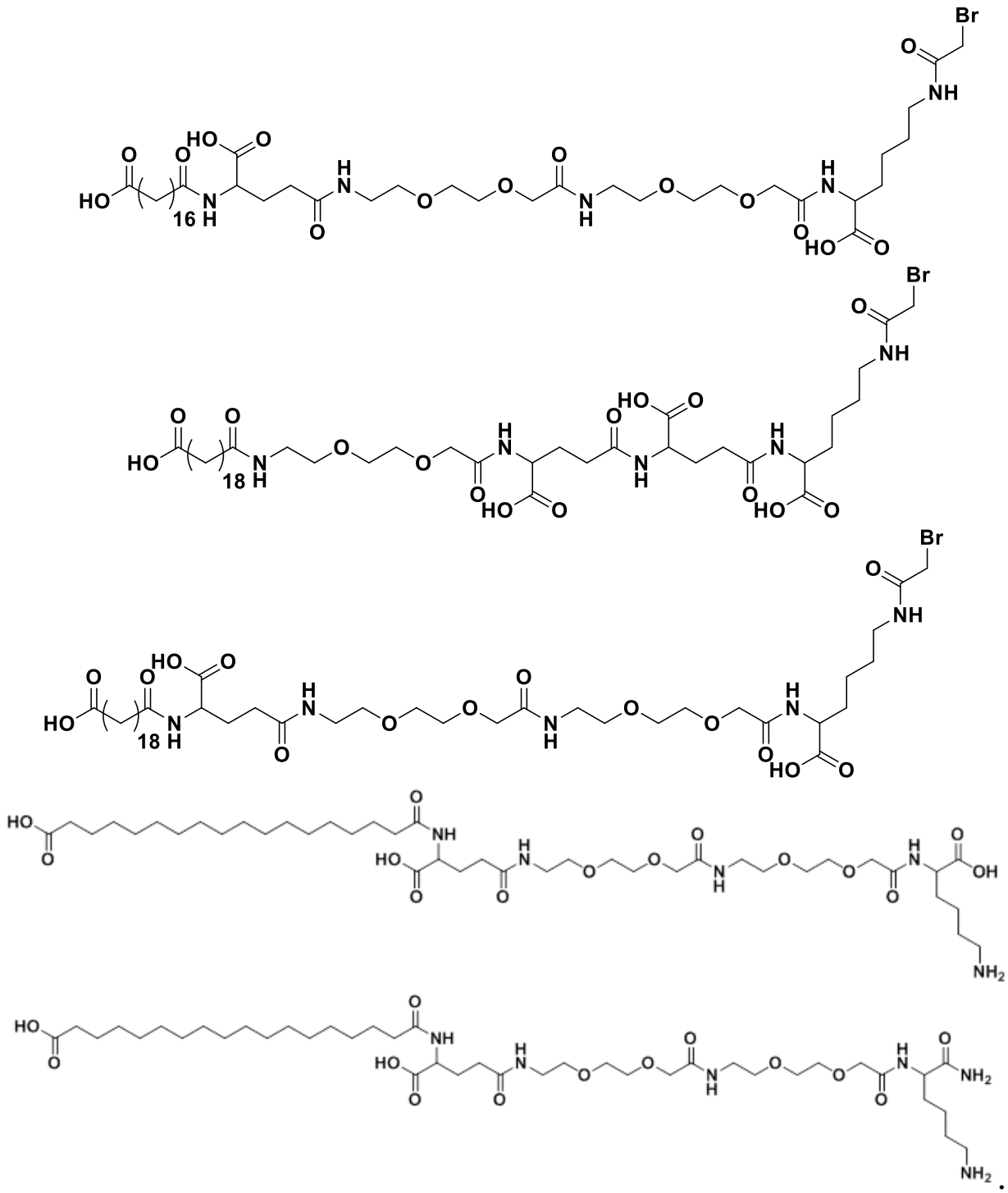
n is an integer from 0 to 5;

X_1 and X_2 are independently selected from the following group: alanine (Ala), D-alanine (D-Ala), β-alanine (β-Ala), 4-aminobutyric acid (GABA), 2-aminoisobutyric acid (Aib), 2-aminobutyric acid (Abu), arginine (Arg), aspartic acid (Asp), asparagine (Asn), cysteine (Cys), glutamic acid (Glu), D-glutamic acid (D-Glu), γ-glutamic acid (γ-Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), proline (Pro), phenylalanine (Phe), serine (Ser), tyrosine (Tyr), threonine (Thr), tryptophan (Trp), valine (Val), methionine (Met), tranexamic acid (Trx), AEEA, PEG;

R_2 is selected from a halogen-substituted $C_{1\text{-}6}$ acyl or H;

more preferably, the albumin-binding agent is selected from the group consisting of:



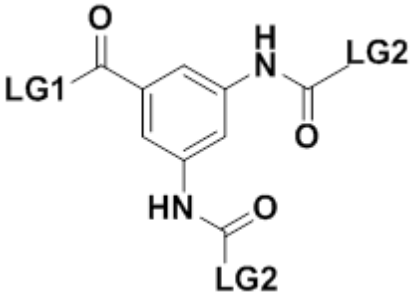


.

[Claim 4] The modified protein of claim 1, wherein the PAS is PAS_{15}-PAS_{200}; preferably PAS_{15}-PAS_{100}; for example, PAS_{50}.

[Claim 5] The modified protein of claim 1, wherein the connector is a multivalent connector, thereby connecting one or more albumin-binding agents and one or more protein moieties;

preferably, the structure of the multivalent connector is shown in the following formula:



wherein LG1 is a leaving group reactive to primary amino groups; and

LG2 is a leaving group reactive to sulfhydryl groups.

[Claim 6] The modified protein of any one of claims 1-5, wherein the modified protein is selected from the compounds with the following numbers: 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 7.2, 7.3, 7.4, 7.5, 12.1, 12.2, 12.3, 12.4, 12.5, 12.6.

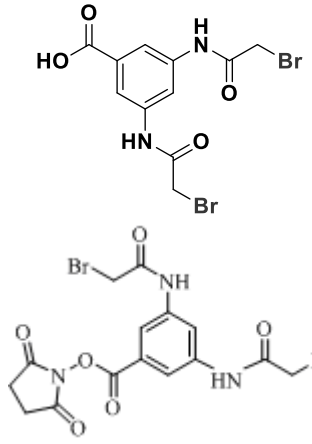
[Claim 7] A protein polymer comprising a multivalent connector and multiple protein monomers, wherein the protein monomers are optionally linked to PAS at their N-terminals, the multivalent connector is connected to the cysteine residues in the protein monomers or the cysteine residues in the PAS, and the cysteine residues are either naturally occurring or introduced through mutation;

preferably,

the PAS is PAS_{15}-PAS_{200}; preferably PAS_{15}-PAS_{100}; for example, PAS_{50}, or

the protein is fibroblast growth factor 21, or

the structure of the multivalent connector is shown in the following formula:



, or

the protein polymer is selected from the compounds with the following numbers: 10.1, 10.2, 10.3, 10.4, 10.5.

[Claim 8] A pharmaceutical composition, comprising the modified protein of any one of claims 1-6 or the protein polymer of claim 7, and a pharmaceutically acceptable excipient.

[Claim 9] Use of the modified protein of any one of claims 1-6 or the protein polymer of claim 7 in the preparation of a medicament.

[Claim 10] An albumin-binding agent, and the structure of the albumin-binding agent is

shown in the following formula:

R_1-$(X_1)_m$-$(X_2)_n$-K(ε-R_2);

wherein K is a lysine;

R_1 is a substituted or unsubstituted C_{1-20} acyl;

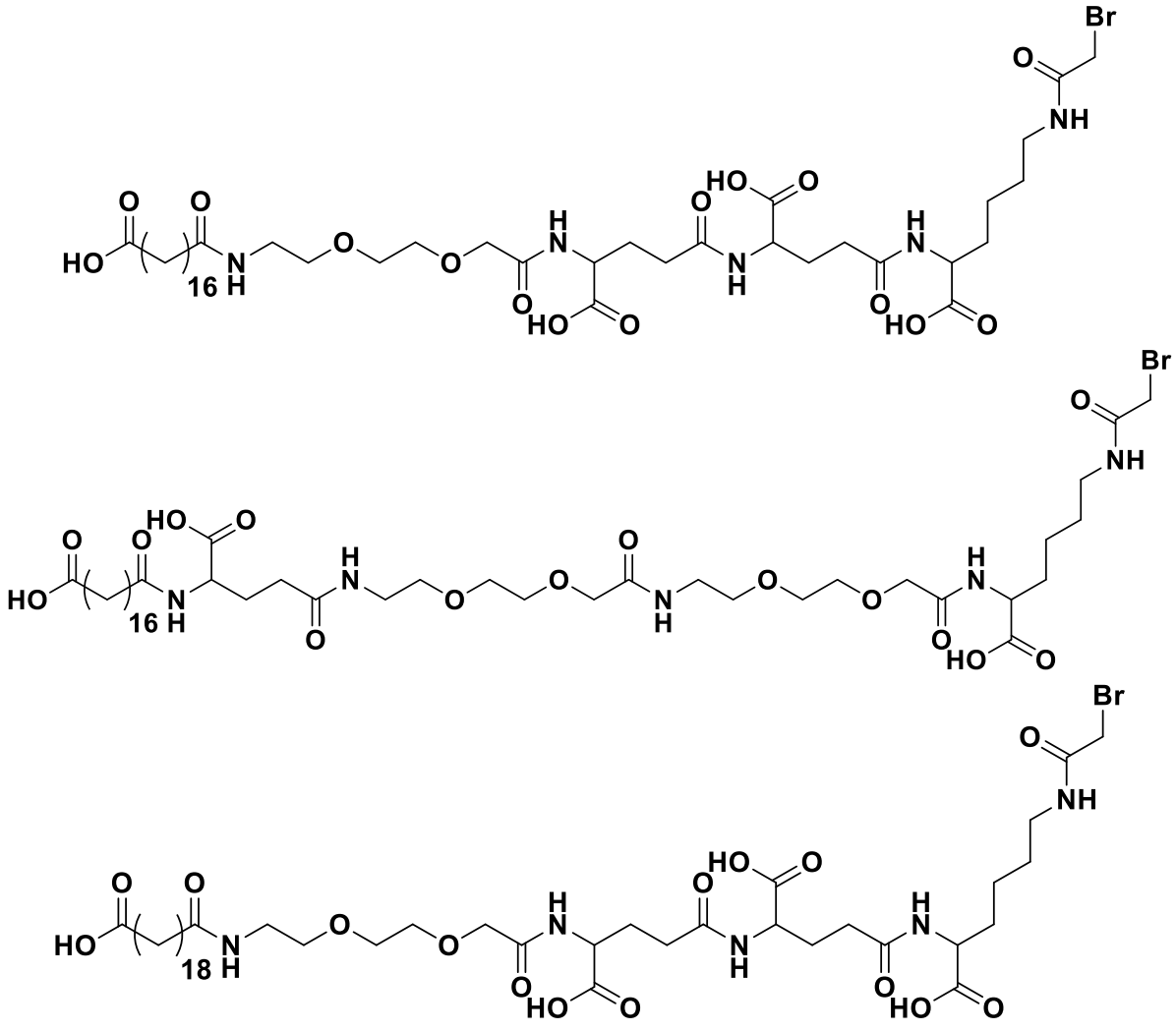
m is an integer from 0 to 5;

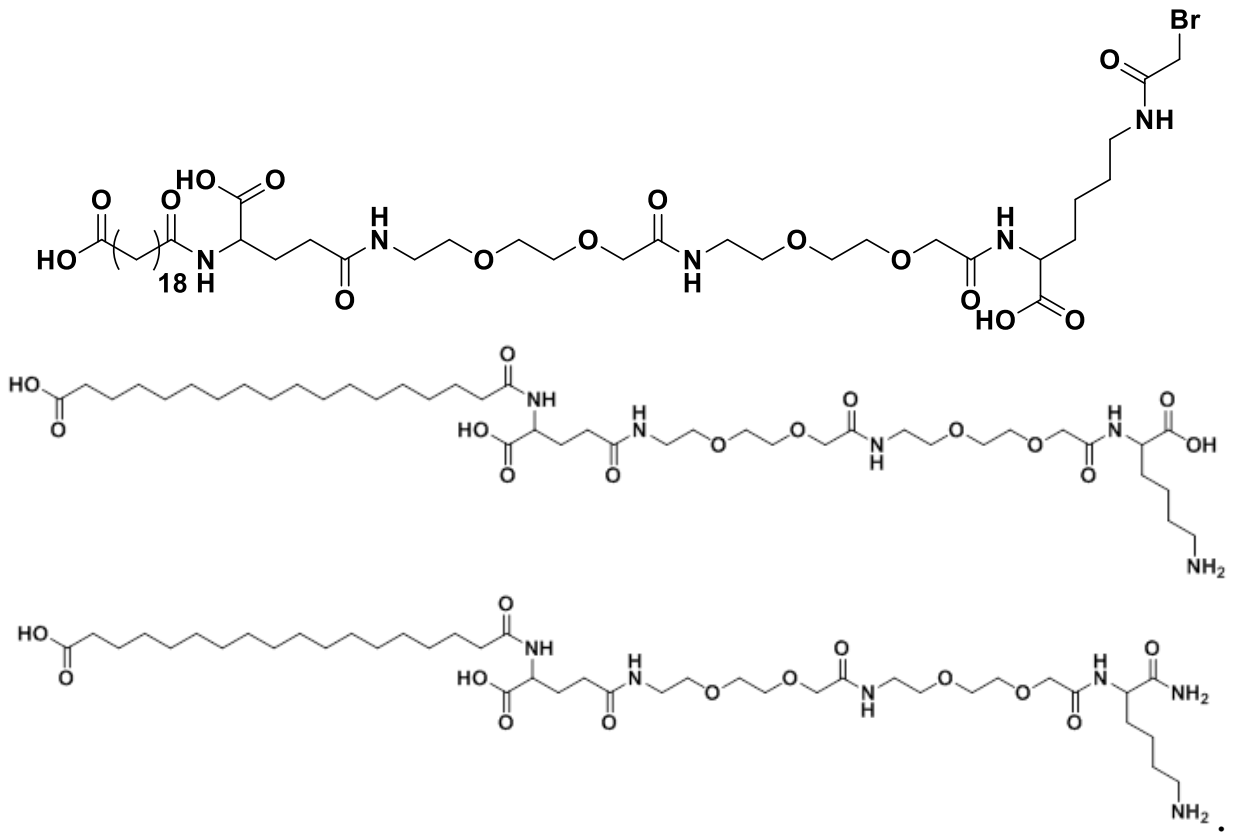
n is an integer from 0 to 5;

X_1 and X_2 are independently selected from the following group: alanine (Ala), D-alanine (D-Ala), β-alanine (β-Ala), 4-aminobutyric acid (GABA), 2-aminoisobutyric acid (Aib), 2-aminobutyric acid (Abu), arginine (Arg), aspartic acid (Asp), asparagine (Asn), cysteine (Cys), glutamic acid (Glu), D-glutamic acid (D-Glu), γ-glutamic acid (γ-Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), proline (Pro), phenylalanine (Phe), serine (Ser), tyrosine (Tyr), threonine (Thr), tryptophan (Trp), valine (Val), methionine (Met), tranexamic acid (Trx), AEEA, PEG;

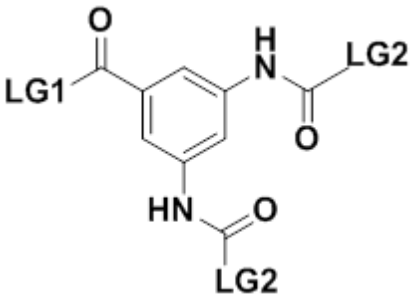
R_2 is selected from a halogen-substituted C_{1-6} acyl or H;

preferably, the albumin-binding agent is selected from the group consisting of:





[Claim 11] A multivalent connector, and the structure of the multivalent connector is shown in the following formula:



wherein LG1 is a leaving group reactive to primary amino groups; and

LG2 is a leaving group reactive to sulfhydryl groups.

[Claim 12] A polypeptide, comprising the amino acid sequence shown in SEQ ID NO: 11; or

the amino acid sequence of which is shown in SEQ ID NO: 11, but with a cysteine residue inserted at position 3 of the N-terminus.

[Claim 13] A polypeptide, comprising an amino acid sequence shown in any one of SEQ ID NO: 4-9; preferably an amino acid sequence shown in any one of SEQ ID NO: 5-9.

[Claim 14] An isolated nucleic acid molecule, encoding the polypeptide of claim 12 or 13.

[Claim 15] An expression vector, comprising the isolated nucleic acid molecule of claim 14.

[Claim 16] A host cell, comprising the expression vector of claim 15, or having the

nucleic acid molecule of claim 14 integrated into the genome of the host cell.

[Claim 17] Use of the polypeptide of claim 12 or 13 in the preparation of the modified protein of any one of claims 1-6, the protein polymer of claim 7, or the pharmaceutical composition of claim 8.