

UNNATURAL AMINO ACID, AND USE THEREOF, RECOMBINANT PROTEIN CONTAINING SAME, AND RECOMBINANT PROTEIN CONJUGATE

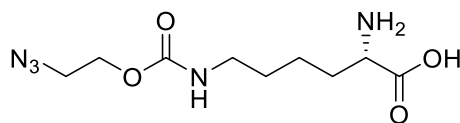
TECHNICAL FIELD

5 The present invention relates to the field of biopharmaceuticals, and particularly to an unnatural amino acid, a recombinant protein containing such unnatural amino acid and a conjugate formed with the recombinant protein.

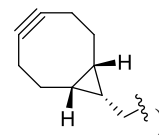
BACKGROUND

10 The development of various scientific research and product applications can be realized by introducing unnatural amino acids containing special groups into proteins. For example, the introduction of photosensitive unnatural amino acids into the proteins or the special labeling to the unnatural amino acids is convenient for studying the interaction between the proteins. Another example is the unnatural amino acids which are used for the directional evolution of
15 enzymes to improve the activity or stability of the enzymes, or to facilitate the efficient immobilization of the enzymes. For another example, safe live bacterium or live virus vaccines may be prepared by using the characteristic that the insertion of the unnatural amino acids which cannot be realized in conventional hosts. An important application of using codon expansion technology in introducing the unnatural amino acids into the proteins lies in the site-specific
20 modification of the proteins, which could improve the function, stability and half-life of the proteins, and may be used for the development of innovative biological drugs. At present, a series of achievements have been made, such as the preparation of long-acting recombinant proteins from PEG site-specific coupled recombinant human growth hormones, the development of antibody-drug conjugates from small molecular toxin site-specific coupled monoclonal antibodies,
25 and the like. Thus, it can be seen that the unnatural amino acids have a very important role and a very wide range of applications.

The prior arts (such as CN 102838671B, CN 106146663A, and *J. Am. Chem. Soc.* **2009**, 131, 8720) disclosed an unnatural amino acid Lys-azido, with a structural formula as follows:



30 An azide structure (-N₃) at the end of the Lys-azido can be chemically linked with a carrier



drug (such as PEG) modified with an alkyne-containing structure (such as BCN, i.e.

to obtain a conjugate (for example, Chinese patent CN 103153927B), which has very high specific selectivity. However, alkyne structures with high costs need to be introduced in this coupling method and chemical modification method, and an acceptable drug-antibody coupling ratio can be obtained only when the alkyne structures are used in larger equivalent, which increases corresponding production costs, and has a complicated technological process and harsh technological conditions.

Therefore, it is necessary to develop an unnatural amino acid with a novel structure, simple preparation and a low cost, so as to expand the types and applications of amino acids.

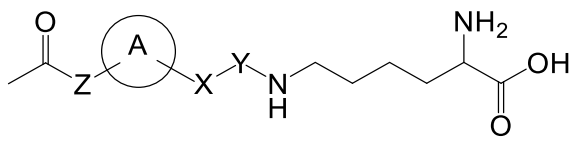
SUMMARY

In order to overcome the defects in the prior art, one object of the present invention is to provide an unnatural amino acid.

Another object of the present invention is to provide use of the unnatural amino acid.

Yet another object of the present invention is to provide a recombinant protein and a recombinant protein conjugate.

The unnatural amino acid provided by the present invention is a compound with a structure as shown in formula (I) or an enantiomer thereof,



I

wherein X and Z each independently represent substituted or unsubstituted C0-C20 linear or branched alkylene, one or more -CH₂- are optionally substituted by one or more of -O-, -S-, -NH-, -C(O)- and -S(O)-, Y represents -C(O)-, -S(O)- or -CH₂-, and A represents substituted or unsubstituted C6-C20 aryl; and

when X, Z and A each independently represent a substituent, the substituent may be selected from one or more of hydroxyl, sulfhydryl, halogen, nitro, cyano, alkyl, alkenyl, alkynyl, alkoxy, acyl, acylamino, carboxyl, ester, amino, sulfonyl, sulfinyl, cycloalkyl, heterocyclyl, aryl and heteroaryl.

As shown in Example 11, the inventor of the present invention found that, besides the high cost and the complicated process, the azide structure (-N₃) at the end of Lys-azido was easily reduced into an amino structure (-NH₂) when the azide structure was inserted into a protein, thus

losing coupling activity, so that the reduction reaction reduces a yield in a technology preparation process.



Carbonyl is introduced at the end of the unnatural amino acid provided by the present invention as an active reaction group, so that the unnatural amino acid is not only novel in structure, simple and convenient to prepare, but also mild in coupling conditions and low in production cost, and is not easy to lose reaction activity due to a structural change when being inserted into a protein sequence. The unnatural amino acid provided by the present invention further contains aryl connected with the terminal carbonyl, the introduction of the aryl can further improve the stability of the obtained conjugate, and the conjugate is not easy to decompose even under a lower pH condition. In addition, the unnatural amino acid provided by the present invention further contains alkylene with a certain chain length, so that the compound has better flexibility and is easier to form various conjugates.

In the unnatural amino acid provided by the present invention, "C0-Cn" includes C0-C1, C0-C2, ..., C0-Cn, and C0 indicates that the group does not exist, and C atoms at two ends of the group are directly connected to form a bond. For example, the "C0-C6" group refers to that there are 0 to 6 carbon atoms in this part, which means that the group does not exist, the group contains 1 carbon atom, 2 carbon atoms, 3 carbon atoms, 4 carbon atoms, 5 carbon atoms or 6 carbon atoms. The "C6-C10" group refers to that there are 6 to 10 carbon atoms in this part, which means that the group contains 6 carbon atoms, 7 carbon atoms, 8 carbon atoms, 9 carbon atoms or 10 carbon atoms.

In the unnatural amino acid provided by the present invention, the "aryl" refers to a carbocyclic aromatic system containing one or two rings, wherein the rings may be connected together in a fused manner. The "aryl" includes monocyclic or bicyclic aryl, such as aromatic groups of phenyl, naphthyl and tetrahydronaphthyl. Preferably, the aryl is C6-C10 aryl, more preferably, the aryl is the phenyl and the naphthyl, and most preferably, the aryl is the phenyl.

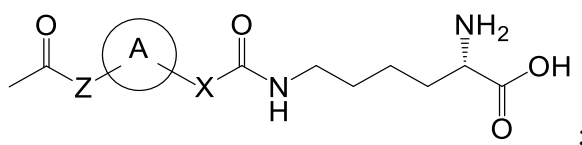
In some preferred embodiments of the present invention, the substituent may be selected from one or more of hydroxyl, sulfhydryl, halogen, nitro, cyano, C1-C6 alkyl, C1-C6 alkoxy, acyl, acylamino, carboxyl, ester, amino, sulfonyl, sulfinyl, C3-C8 cycloalkyl, C3-C8 heterocyclyl, C6-C20 aryl and C4-C10 heteroaryl.

In some preferred embodiments of the present invention, X and Z may each independently represent C0-C10 linear or branched alkylene, wherein one or more -CH₂- may be optionally substituted by one or more of -O-, -S- and -NH-. In some more preferred embodiments of the present invention, X and Z may each independently represent C0-C6 linear alkylene, wherein one

or more -CH₂- may be optionally substituted by one or more of -O-, -S- and -NH-. In some more preferred embodiments of the present invention, X and Z are not C0 alkylene simultaneously, which means that the X and Z groups cannot exist simultaneously.

In some preferred embodiments of the present invention, A may represent substituted or unsubstituted C₆-C₁₀ aryl, and more preferably, A may represent substituted or unsubstituted phenyl or naphthyl.

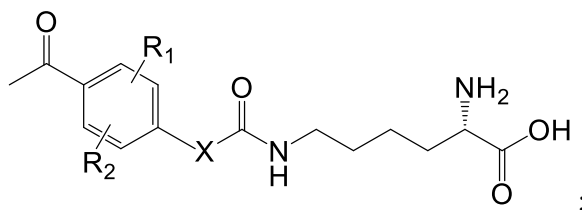
In some preferred embodiments of the present invention, the unnatural amino acid may be a compound with a structure as shown in formula (I-1),



10 I-1

wherein X, Y and A are each independently as defined in any one of the above technical solutions.

In some preferred embodiments of the present invention, the unnatural amino acid may be a compound with a structure as shown in formula (I-2),

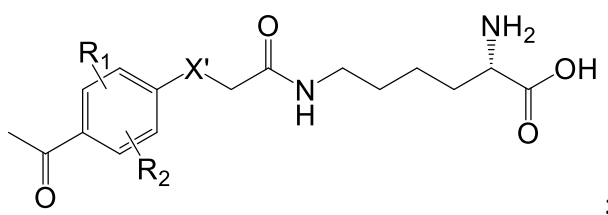


15 I-2

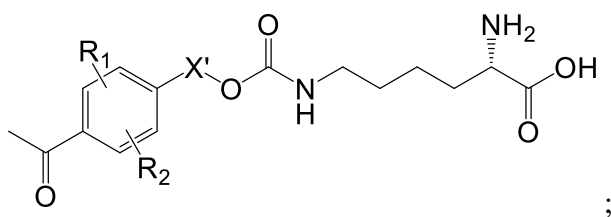
wherein X is as defined in any one of the above technical solutions; and

R₁ and R₂ each independently represent hydrogen, hydroxyl, sulfhydryl, halogen, nitro, cyano, C₁-C₆ alkyl, C₁-C₆ alkoxy, acyl, acylamino, carboxyl, ester, amino, sulfonyl, sulfinyl, C₃-C₈ cycloalkyl, C₃-C₈ heterocyclyl, C₆-C₂₀ aryl or C₄-C₁₀ heteroaryl.

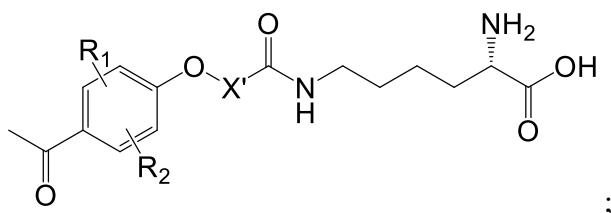
In some preferred embodiments of the present invention, the unnatural amino acid is a compound with a structure as shown in formula (I-3), formula (I-4), formula (I-5) or formula (I-6),



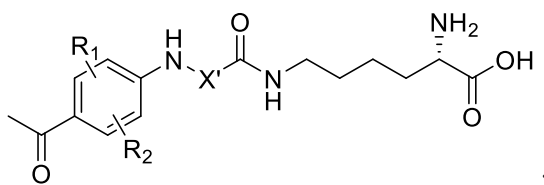
25 I-3



I-4



I-5



I-6

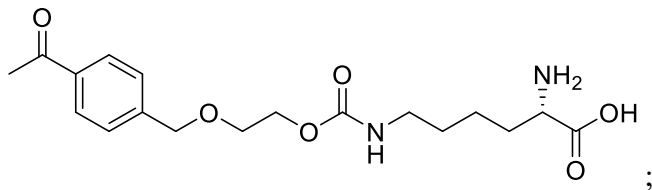
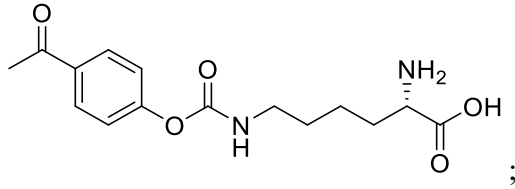
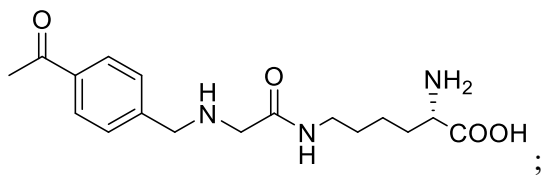
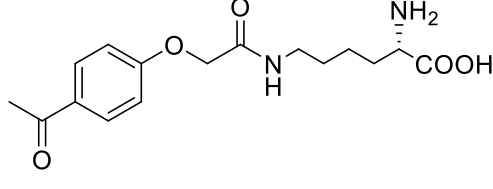
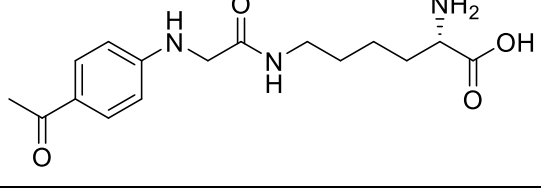
wherein X' represents C0-C6 linear alkylene, and more preferably represents C0-C4 linear alkylene, and one or more -CH₂- are optionally substituted by -O- and/or -NH-; and

R₁ and R₂ are each independently as defined in any one of the above technical solutions.

The unnatural amino acid provided by the present invention includes optically pure enantiomer and racemate.

In some more preferred embodiments of the present invention, the unnatural amino acid provided by the present invention is a compound with one of the structures as follows:

(1) NBOK	
(2) NPAK	

(3) NBPK	
(4) NPOK	
(5) NBGK	
(6) NPOK-2	 or
(7) NBGK-2	

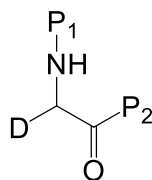
The present invention further provides use of the unnatural amino acid according to any one of the above technical solutions in preparation of a recombinant protein or a recombinant protein conjugate.

- 5 In the use of the present invention, the recombinant protein may be a recombinant protein obtained by inserting the unnatural amino acid according to any one of the above technical solutions into any kind of protein common in the art at any site in any quantity. The recombinant protein conjugate may be a conjugate obtained by coupling any recombinant protein obtained with a coupling moiety common in the art, wherein the coupling moiety may include, but is not
- 10 limited to, one or more of a polymers (such as polyethylene glycol with any molecular weight), proteins, polypeptides or small molecule drugs.

In some preferred embodiments of the present invention, the recombinant protein may be a recombinant human growth hormone, and the recombinant protein conjugate may be a recombinant human growth hormone-polyethylene glycol conjugate.

The present invention further provides a recombinant protein, wherein at least one site in the amino acid sequence of the recombinant protein is the unnatural amino acid according to any one of the above technical solutions.

Further, the recombinant protein may have a structure as shown in formula (II),



II

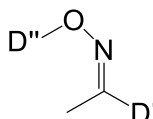
in the formula (II), D represents a residue of the unnatural amino acid according to any one of the above technical solutions with an aminocarboxylic acid part removed, and P₁ and P₂ respectively represent connecting parts of the amino and the carboxyl of the unnatural amino acid in the amino acid sequence.

The recombinant protein provided by the present invention may be prepared by a preparation method common in the art, including realizing cloning and expression of the recombinant protein containing the unnatural amino acid by a gene codon expansion technology.

The recombinant protein provided by the present invention may be a recombinant protein obtained by inserting the unnatural amino acid according to any one of the above technical solutions into any kind of protein common in the art at any site in any quantity, such as a recombinant human growth hormone.

The present invention further provides a recombinant protein conjugate, wherein the recombinant protein conjugate is formed by forming an oxime bond with a terminal carbonyl of the unnatural amino acid in the recombinant protein according to any one of the above technical solutions and a coupling moiety containing an "NH₂-O-" terminal group.

Further, the recombinant protein conjugate may have a structure as shown in formula (III),



III

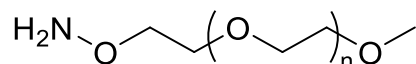
in the formula (III), D' represents a residue of the recombinant protein according to any one of the above technical solutions with the terminal carbonyl part of the unnatural amino acid removed, and D'' represents a residue of the coupling moiety with the "NH₂-O-" terminal group removed.

In the recombinant protein conjugate provided by the present invention, the coupling moiety may include one or more of polymers (such as polyethylene glycol with any molecular weight),

proteins, polypeptides or small molecule drugs. Different kinds of coupling moieties may be separately coupled with the recombinant protein, or different kinds of coupling moieties may be connected first and then coupled with the recombinant protein.

In some preferred embodiments of the present invention, the recombinant protein may be a recombinant human growth hormone, and the coupling moiety containing the "NH₂-O-" terminal group may be polyethylene glycol containing an "NH₂-O-" terminal group.

In some more preferred embodiments of the present invention, the polyethylene glycol containing the "NH₂-O-" terminal group has a structural formula as follows:



wherein the molecular weight of the polyethylene glycol containing the "NH₂-O-" terminal group may be 10 KD to 100 KD, including, but being not limited to, molecular weight values of about 10 KD, 20 KD, 30 KD, 40 KD, 50 KD, 60 KD, 70 KD, 80 KD, 90 KD and 100 KD, or a molecular weight range of any combination. Preferably, the molecular weight of the polyethylene glycol containing the "NH₂-O-" terminal group may be 20 KD to 50 KD.

In some more preferred embodiments of the present invention, the amino acid sequence of the recombinant human growth hormone is as shown in SEQ ID NO: 1; and further preferably, a corresponding 107th site in the amino acid sequence of SEQ ID NO: 1 is the unnatural amino acid according to any one of the above technical solutions.

When the recombinant protein is the recombinant human growth hormone, the present invention further provides use of the recombinant protein conjugate according to any one of the above technical solutions in preparation of a drug for treating growth and development disorders caused by endogenous growth hormone secretion insufficiency, growth and development disorders caused by Turner syndrome or adult growth hormone deficiency.

The technical solutions provided by the present invention have the following advantages.

(1) The terminal carbonyl and the aryl connected with the terminal carbonyl are introduced into the structure of the unnatural amino acid provided by the present invention, and compared with the current terminal azido unnatural amino acid (such as Lys-azido), the unnatural amino acid is simple and convenient to prepare, good in safety, not easy to inactivate when being inserted into a protein, and high in binding rate with a coupling moiety, and the obtained conjugate is higher in stability.

(2) As an amino acid derivative, the unnatural amino acid provided by the present invention may have the properties of amino acid, thus expanding potential types of amino acids, so that the unnatural amino acid may be applied to numerous fields as the amino acid derivative, especially in preparation of a recombinant protein or a recombinant protein conjugate.

(3) The unnatural amino acid provided by the present invention may be successfully recognized in a prokaryotic expression system and an eukaryotic expression system and inserted

into a protein, thus producing a protein containing the unnatural amino acid at a specific site, so that recombinant proteins with different physical and chemical properties and biochemical activities may be formed, and types and potential application scopes of the proteins are expanded. In addition, the unnatural amino acid of the present invention has higher expression efficiency in protein and better practicability.

(4) Because the recombinant protein provided by the present invention contains the unnatural amino acid of the present invention, the terminal carbonyl active group contained in the unnatural amino acid may easily form a protein conjugate (or a conjugate), such as a polyethylene glycol conjugate and a polyethylene glycol-active drug conjugate. These protein conjugates may have many improved biological activities (such as anti-tumor activity) due to property modification.

(5) The present invention further provides a novel protein conjugate platform, and multiple types of proteins and multiple types of coupling moieties may be bound through the novel unnatural amino acid contained in protein and the coupling moiety connected.

DESCRIPTION OF THE DRAWINGS

FIG. 1 is a profile of expression plasmid pET21-rhGH107 for a recombinant human growth hormone in Example 8.

FIG. 2 is an SDS-PAGE electrophoretogram of fermentation products obtained by adding different types of unnatural amino acids in Example 8, wherein various lanes respectively represent the followings: Lane 1, protein molecular weight Marker; Lane 2, wild-type recombinant human growth hormone; Lane 3, expression product of recombinant human growth hormone added with NBOK; Lane 4, expression product of recombinant human growth hormone added with NPAK; Lane 5, expression product of recombinant human growth hormone added with NBPK; Lane 6, expression product of recombinant human growth hormone added with NBGK; Lane 7, expression product of recombinant human growth hormone added with NPOK; Lane 8, expression product of recombinant human growth hormone added with NPOK-2; Lane 9, expression product of recombinant human growth hormone added with NBGK-2; and Lane 10, expression product of recombinant human growth hormone not added with any unnatural amino acid.

FIG. 3 is an SDS-PAGE electrophoretogram of coupling products of recombinant human growth hormone containing an unnatural amino acid coupled with PEG in Example 8, wherein various lanes respectively represent the followings: Lane 1, molecular weight Marker; Lane 2, wild-type recombinant human growth hormone; Lane 3, coupling product of recombinant human growth hormone containing NBOK coupled with PEG; Lane 4, coupling product of recombinant human growth hormone containing NPAK coupled with PEG; Lane 5, coupling product of

recombinant human growth hormone containing NBPK coupled with PEG; Lane 6, coupling product of recombinant human growth hormone containing NBGK coupled with PEG; Lane 7, coupling product of recombinant human growth hormone containing NPOK coupled with PEG; and Lane 8, coupling product of recombinant human growth hormone containing NPOK-2 coupled with PEG.

FIG. 4 is an SDS-PAGE electrophoretogram of purified PEG-coupled recombinant human growth hormones in Example 8, wherein various lanes respectively represent the followings: Lane 1, molecular weight Marker; Lane 2, standard product of recombinant human growth hormone; Lane 3, coupling product of recombinant human growth hormone containing NPOK coupled with PEG; and Lane 4, coupling product of recombinant human growth hormone containing NBOK coupled with PEG.

FIG. 5A to FIG. 5G are cell activity curves in Example 8, wherein FIG. 5A is a cell activity curve of rhGH standard product of National Institutes for Food and Drug Control; FIG. 5B is a cell activity curve of Jintropin®; FIG. 5C is a cell activity curve of Polyethylene Glycol Recombinant Human Somatropin Injection®; FIG. 5D is a cell activity curve of rhGH (NPOK); FIG. 5E is a cell activity curve of PEG-rhGH (NPOK); FIG. 5F is a cell activity curve of rhGH (NBOK); and FIG. 5G is a cell activity curve of PEG-rhGH (NBOK).

FIG. 6 is a fluorescence micrograph of CHO cells transiently expressing inserted unnatural amino acids in Example 9.

FIG. 7 is a profile of an expression plasmid pCDNA3.1-Trastuzumab-UAG142 in Example 10.

FIG. 8A to FIG. 8C are respectively HIC-HPLC spectra of trastuzumab containing NBPK, trastuzumab containing NBOK and trastuzumab containing NPOK-2 coupled with a toxin in Example 10.

FIG. 9 is an inhibitory effect graph of trastuzumab containing NBOK and DM1-modified trastuzumab containing NBOK on BT-474 cells respectively in Example 10.

FIG. 10A and FIG. 10B are respectively mass spectra of rhGH with a mutation into Lys-azido at 140th site and rhGH with a mutation into NBOK at 140th site in Example 11.

FIG. 11A and FIG. 11B are respectively SDS-PAGE electrophoretograms of coupling processes of rhGH with a mutation into Lys-azido at 140th site and rhGH with a mutation into NBOK at 140th site in Example 11 with 30KD PEG, wherein various lanes in FIG. 11A respectively represent the followings: Lane 1, molecular weight Marker; Lane 2, wild-type recombinant human growth hormone; Lane 3, rhGH-Lys-azido-140; Lane 4, product of coupling reaction for 72 hours when rhGH-Lys-azido-140: 30K BCN-PEG is 1:15 (molar ratio, the same below); and Lane 5, product of coupling reaction for 72 hours when rhGH-Lys-azido-140: 30K BCN-PEG is 1:25; various lanes in FIG. 11B respectively represent the followings: Lane 1,

molecular weight Marker; Lane 2, rhGH-NBOK-140; Lane 3, product of coupling reaction for 6 hours when rhGH-NBOK-140: 30K BCN-PEG is 1:15; Lane 4, product of coupling reaction for 9 hours when rhGH-NBOK-140: 30K BCN-PEG is 1:15; Lane 5, product of coupling reaction for 12 hours when rhGH-NBOK-140: 30K BCN-PEG is 1:15; Lane 6, product of coupling reaction for 24 hours when rhGH-NBOK-140: 30K BCN-PEG is 1:15; and Lane 7, product of coupling reaction for 48 hours when rhGH-NBOK-140: 30K BCN-PEG is 1:15.

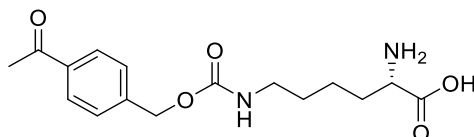
EMBODIMENT

The technical solutions of the present invention will be further described in detail hereinafter with reference to the specific examples.

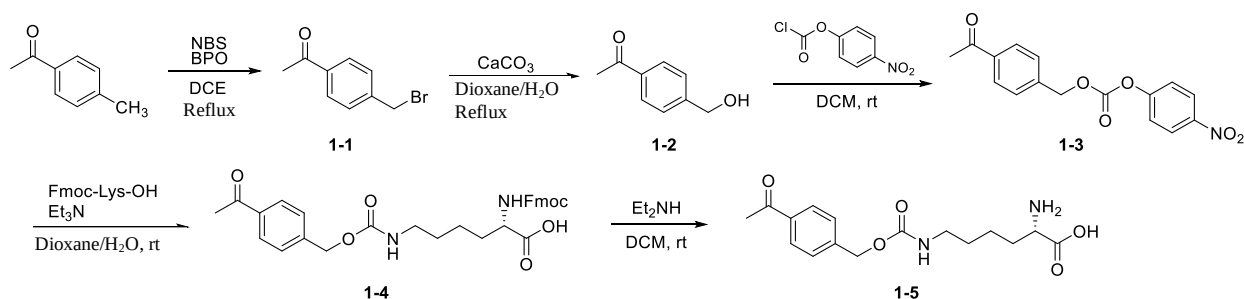
Unless otherwise specified, the reagents or raw materials used in the examples of the present invention are all commercially available.

Example 1 Preparation of unnatural amino acid NBOK

The structural formula of NBOK was as follows:



The reaction process was shown in the following figure:



The preparation process included the following steps.

a) *P*-methyl acetophenone (4.0 mL, 30.0 mmol), a solvent DCE (50.0 mL), NBS (6.41 g, 36.0 mmol) and BPO (0.05 g, 0.3 mmol) were added into a reaction flask. The mixture was refluxed at 80°C for 24 hours, then the container was cooled in ice water to precipitate a solid, and the solid was removed by filtration. The solution was washed with saturated Na₂CO₃ for 3 times, and then extracted with DCM for 3 times. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a crude product **1-1** (5.46 g, yield 85%), which was directly used in next step without purification.

b) The product **1-1** (2.73 g, 12.80 mmol), a solvent dioxane (40 mL) and water (40 mL) were added into a reaction flask, and then added with calcium carbonate (7.68 g, 76.8 mmol). The mixture was refluxed at 105°C for 24 hours, cooled to room temperature, filtered to remove the

solid, and then extracted with DCM for 3 times. Organic phases were combined, concentrated under a reduced pressure, and purified by column chromatography (eluent: PE:EA = 3:1) to obtain a product **1-2** (1.80 g, yield 94%).

c) *P*-nitrophenyl chloroformate (2.90 g, 14.4 mmol) and a solvent DCM (10.0 mL) were added into a two-neck reaction flask, cooled to 0°C, added with the product **1-2** (1.80 g, 12.0 mmol) and pyridine (1.2 mL, 14.4 mmol), stirred at room temperature for 18 hours, and then the reaction solution was added with a saturated sodium carbonate solution (10 mL), and extracted with DCM (50 mL) for 3 times. Organic phases were combined, washed with water twice, dried with anhydrous sodium sulfate, filtered, concentrated under a reduced pressure, and purified by column chromatography (eluent: PE:EA = 5:1) to obtain a product **1-3** (3.14 g, yield 83%).

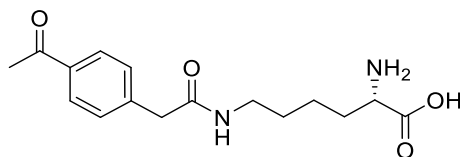
d) The product **1-3** (1.26 g, 4.0 mmol) and Fmoc-Lys-OH hydrochloride (1.40 g, 3.33 mmol), a solvent dioxane (15 mL) and water (5 mL) were added into a reaction flask, and then added with triethylamine (1.2 mL, 8.3 mmol). The mixture was reacted at room temperature for 24 hours, and then was added with an appropriate amount of 1 M HCl solution, extracted with DCM, and concentrated under a reduced pressure to obtain a crude product **1-4**, which was directly used in next step.

e) The product **1-4** (1.10 g, 0.19 mmol) was dissolved in DCM (10 mL) in a reaction flask, and added with diethylamine (5.0 mL) to react at room temperature for 6 hours to precipitate a product. The product was filtered, and then pulped with DCM for 3 times to obtain the target product NBOK (**1-5**, 817 mg, two-step yield 63%).

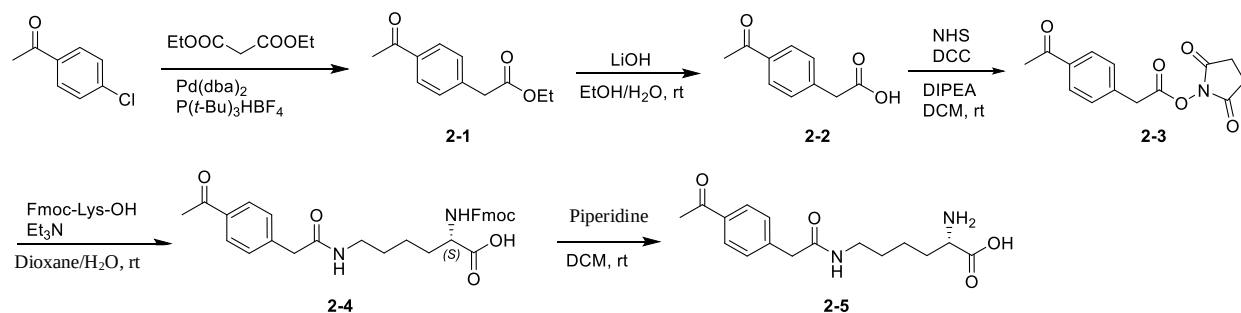
¹H-NMR (400 MHz, heavy water) δ 8.04 (d, *J* = 8.4 Hz, 2H), 7.55 (d, *J* = 8.0 Hz, 2H), 5.21 (s, 2H), 3.74 (t, *J* = 6.0 Hz, 1H), 3.17 (t, *J* = 6.4 Hz, 2H), 2.70 (s, 3H), 1.95 – 1.83 (m, 2H), 1.62 – 1.52 (m, 2H), 1.47 – 1.35 (m, 2H).

Example 2 Preparation of unnatural amino acid NPAK

The structural formula of NPAK was as follows:



The reaction process was shown in the following figure:



The preparation process included the following steps.

a) *P*-chloroacetophenone (1.00 g, 6.47 mmol) was added into a reaction flask, diethyl malonate (6.84 g, 47.70 mmol), KHCO₃ (0.97 g, 9.70 mmol) and K₂CO₃ (1.34 g, 9.70 mmol) were added under a nitrogen atmosphere, and then Pd(dba)₂ (0.019 g, 0.030 mmol) and
5 P(*t*-Bu)₃HBF₄ (0.021 g, 0.071 mmol) were added. Nitrogen was replaced for protection after addition, and the mixture was heated to 160°C to react for 40 hours. After the completion of reaction was detected by TLC, the reaction solution was added with water (30 mL), extracted with EA for 3 times. Organic phases were combined, washed with water twice, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure at 0°C to 5°C to
10 obtain a colorless and transparent liquid, and the liquid was purified by column chromatography (eluent: PE:EA= 10:) to obtain a product **2-1** (0.80 g, yield 60%).

b) LiOH (0.30 g, 11.64 mmol) and water (5.0 mL), ethanol (10 mL), and the product **2-1** (0.80 g, 3.88 mmol) were added into a reaction flask, and stirred at room temperature for 2 hours. After the completion of reaction was detected by TLC, the reaction solution was added with 2 M
15 HCl solution to adjust a pH value to be 1 to 2, and then extracted with EA for 3 times. Organic phases were combined, washed with water twice, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a product **2-2** (0.5 g, yield 72%).

c) The product **2-2** (0.20 g, 1.12 mmol) was added into a reaction flask, and N-hydroxysuccinimide (NHS, 0.19 g, 1.68 mmol), DIPEA (0.07 g, 0.56 mmol) and DCM (2.0
20 mL) were added in sequence. The mixture was cooled to 0°C to 5°C, added with a solution of DCC (0.23 g, 1.12 mmol) and DCM (2.0 mL) to react in a thermally insulated manner for 2 hours, and heated to room temperature to stir overnight. After the completion of reaction was detected by TLC, the reaction solution was filtered, and washed with DCM, and the mother solution was concentrated under a reduced pressure, and purified by column chromatography (eluent: PE:EA =
25 5:1) to obtain a product **2-3** (0.19 g, yield 62%).

d) The product **2-3** (0.10 g, 0.36 mmol) was added into a reaction flask, and triethylamine (0.04 g, 0.36 mmol), Fmoc-Lys-OH hydrochloride (0.13 g, 0.36 mmol), dioxane (2.0 mL) and water (2.0 mL) were added in sequence. The mixture was stirred at room temperature to react for
30 18 hours. After the completion of reaction was detected by TLC, the reaction solution was concentrated under a reduced pressure, extracted with EA for 3 times, dried with anhydrous sodium sulfate, concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM:MeOH = 15:1) to obtain a product **2-4** (0.03 g, yield 61%).

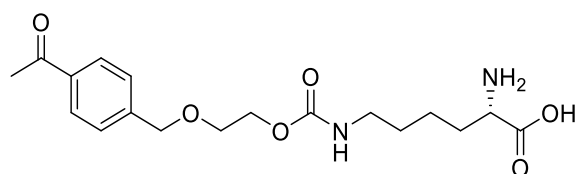
e) The product **2-4** (0.08 g, 0.15 mmol), DCM (1.0 mL) and piperidine (0.04 g, 0.47 mmol) were added into a reaction flask, and stirred at room temperature for 3 hours. After the
35 completion of reaction was detected by TLC, the reaction solution was concentrated under a reduced pressure, pulped with petroleum ether (5 mL) for 1 hour, and filtered. The filter cake was

pulped with petroleum ether (5 mL) for 1 hour again, and filtered, and then the filter cake was repeatedly pulped with ethanol for 4 times to remove residual piperidine to finally obtain a white-like solid **2-5** (0.02 g, yield 43%).

¹H-NMR (400 MHz, heavy water) δ 7.85 (d, J = 8.2 Hz, 2H), 7.33 (d, J = 8.2 Hz, 2H), 3.94 (t, J = 6.3 Hz, 1H), 3.56 (s, 2H), 3.12 (t, J = 6.8 Hz, 2H), 2.54 (s, 3H), 1.80 – 1.70 (m, 2H), 1.54 – 1.45 (m, 2H), 1.40 - 1.224 (m, 2H).

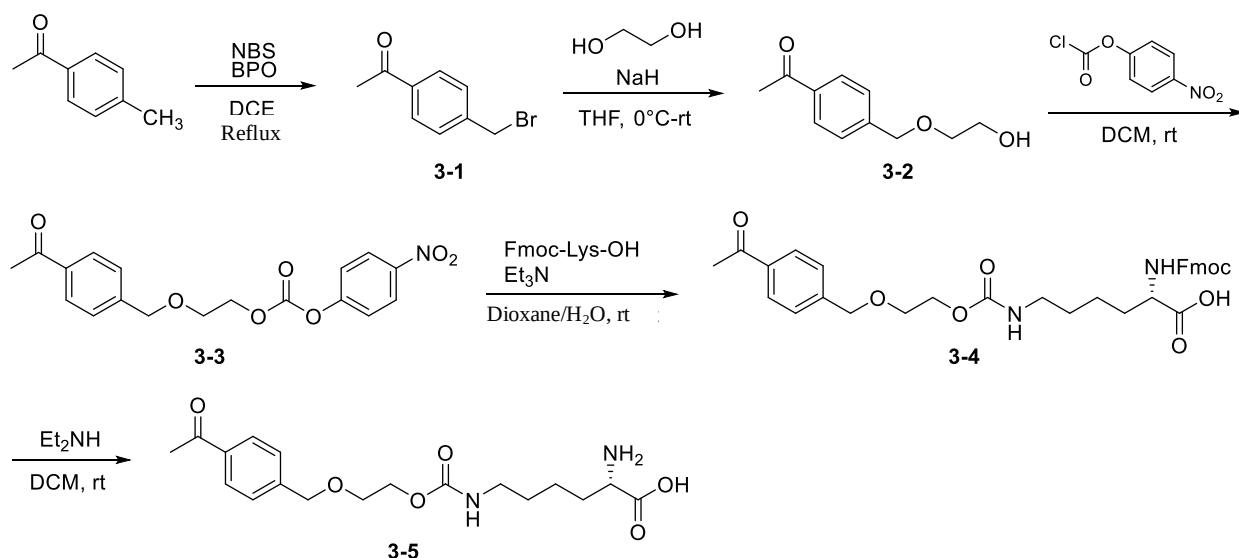
Example 3 Preparation of unnatural amino acid NBPK

The structural formula of NBPK was as follows:



10

The reaction process was shown in the following figure:



The preparation process included the following steps.

a) *P*-methyl acetophenone (4.0 mL, 30.0 mmol), DCE (50.0 mL), NBS (6.41 g, 36.0 mmol) and BPO (0.05 g, 0.3 mmol) were added into a reaction flask, and the mixture was refluxed at 80°C for 24 hours. After the completion of reaction was detected by TLC, the container was cooled in ice water to precipitate a solid, and the solid was removed by filtration. The solution was washed with saturated Na₂CO₃ for 3 times, and extracted with DCM for 3 times. Organic phases were combined, added with anhydrous sodium sulfate for drying, filtered, and concentrated under a reduced pressure to obtain a crude product **3-1** (5.46 g, yield 85%), which was directly used in next step without purification.

b) NaH (0.58 g, 14.64 mmol, 60%) and a dry solvent THF (20 mL) were added into a reaction flask, and ethylene glycol (6.7 mL, 122.0 mmol) was slowly added under cooling in an ice bath. The mixture was stirred at room temperature for 1 hour. Subsequently, the product **3-1** (2.60 g, 12.2 mmol) was added, and heated and refluxed at 70°C for 48 hours until the reaction was completed. Saturated NH₄Cl was slowly and dropwise added under cooling in an ice bath to quench NaH, and then the solution was washed with water, and extracted with EtOAc for 3 times. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, concentrated under a reduced pressure, and purified by column chromatography (eluent: PE:EA = 2:1) to obtain a product **3-2** (1.39 g, yield 59%).

c) The product **3-2** (1.39 g, 7.2 mmol) and a solvent DCM (10 mL) were added into a reaction flask, and p-nitrophenyl chloroformate (1.74 g, 8.64 mmol) and pyridine (0.7 mL, 8.64 mmol) were added under cooling in an ice bath. The mixture was stirred at room temperature for 18 hours. After the completion of reaction was detected by TLC, the reaction solution was added with water for washing, and extracted with EtOAc for 3 times. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, concentrated under a reduced pressure, and then purified by column chromatography (eluent: PE:EA = 3:1) to obtain a product **3-3** (2.27 g, yield 88%).

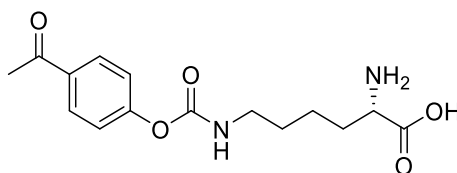
d) The product **3-3** (2.27 g, 6.32 mmol), a solvent dioxane (16 mL), water (4 mL), and Fmoc-Lys-OH hydrochloride (2.13 g, 5.27 mmol) were added into a reaction flask, and then triethylamine (1.85 mL, 13.2 mmol) was added. The mixture was stirred at room temperature for 18 hours until the reaction was completed. The solution was added with an appropriate amount of 1 M HCl to adjust a pH value to be about 2, and extracted with ethyl acetate. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a crude product **3-4**, which was directly used in next step.

e) The product **3-4** in the previous step was dissolved in DCM (10 mL) in a reaction flask, and added with diethylamine (5 mL) to react at room temperature for 6 hours. After the completion of reaction was detected by TLC, the reaction solution was concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM: MeOH: H₂O = 40:10:1) to obtain a product of a white solid (**3-5**, 0.95 g, two-step yield 49%).

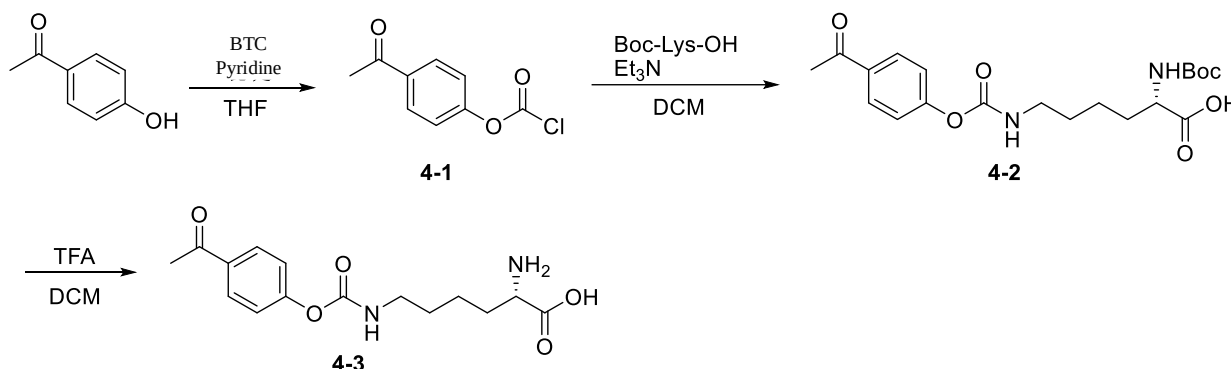
¹H-NMR (400 MHz, heavy water) δ 8.04 (d, *J* = 8.0 Hz, 2H), 7.56 (d, *J* = 8.0 Hz, 2H), 4.72 (s, 2H), 4.25 (s, 2H), 3.81 (s, 2H), 3.74 (t, *J* = 6.0 Hz, 1H), 3.16 – 3.08 (m, 2H), 2.70 (s, 3H), 1.97 – 1.79 (m, 2H), 1.60 – 1.48 (m, 2H), 1.48 – 1.35 (m, 2H).

Example 4 Preparation of unnatural amino acid NPOK

The structural formula of NPOK was as follows:



The reaction process was shown in the following figure:



5

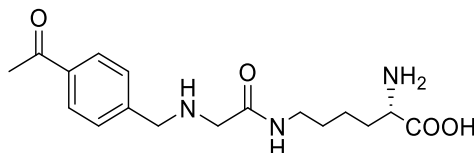
The preparation process included the following steps.

- a) Triphosgene (BTC, 2.18 g, 7.35 mmol) and a solvent THF (10.0 mL) were added into a reaction flask, and *p*-hydroxyacetophenone (2.0 g, 14.7 mmol) and pyridine (1.5 mL, 17.64 mmol) were added under cooling in an ice bath. The mixture was reacted at room temperature for 24 hours. After the completion of reaction was detected by TLC, the reaction solution was added with an appropriate amount of water, and extracted with EtOAc for 3 times. Organic phases were combined, dried with a hydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a crude product **4-1** (1.20 g), which was directly used in next step.
- b) Boc-lysine (1.1 g, 5.0 mmol), a solvent DCM (10.0 mL), the product **4-1** (1.20 g) and triethylamine (2 mL, 15 mmol) were added into a reaction flask. The mixture was stirred at room temperature for 24 hours. After the completion of reaction was detected by TLC, the reaction solution was added with an appropriate amount of 1 M HCl to adjust pH to be faintly acid, and extracted with DCM for 3 times. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM:MeOH = 5:1) to obtain a product **4-2** (1.70 g, yield 84%).
- c) The product **4-2** (1.70 g, 4.2 mmol), a solvent DCM (5 mL), and a trifluoroacetic acid (5 mL) were added into a reaction flask. The mixture was reacted at room temperature for 1 hour. After the completion of reaction was detected by TLC, the reaction solution was directly concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM: MeOH: H₂O = 40:10:1) to obtain a product **4-3** (1.19 g, yield 92%).

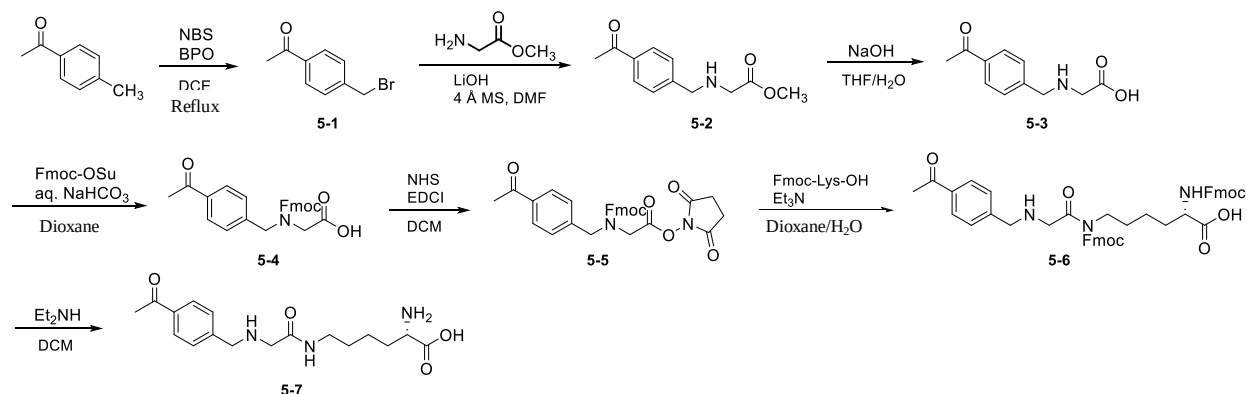
$^1\text{H-NMR}$ (400 MHz, heavy water) δ 8.09 (d, J = 8.6 Hz, 2H), 7.31 (d, J = 8.6 Hz, 2H), 3.78 (t, J = 6.0 Hz, 1H), 3.27 (t, J = 6.8 Hz, 2H), 2.70 (s, 3H), 1.98 – 1.87 (m, 2H), 1.72 – 1.60 (m, 2H), 1.55 – 1.43 (m, 2H).

Example 5 Preparation of unnatural amino acid NBGK

5 The structural formula of NBGK was as follows:



The reaction process was shown in the following figure:



The preparation process included the following steps.

10 a) *P*-methyl acetophenone (8.0 mL, 60.0 mmol), a solvent DCE (80.0 mL), NBS (12.82 g, 72.0 mmol) and BPO (145 mg, 0.6 mmol) were added into a reaction flask, and the mixture was refluxed at 90°C for 24 hours. After the completion of reaction was detected by TLC, the container was cooled in ice water to precipitate a solid, and the solid was removed by filtration. The solution was washed with saturated Na_2CO_3 for 3 times, and extracted with DCM for 3 times.

15 Organic phases were combined, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a crude product **5-1** (11.12 g, yield 87%), which was directly used in next step without purification.

b) 4 Å MS (14 g) and LiOH (1.45 g, 34.54 mmol) were dissolved in DMF (70 mL) in a reaction flask and stirred at room temperature for 20 minutes, then added with glycine methylester hydrochloride (2.0 g, 15.7 mmol) and stirred for 45 minutes, and then added with the product **5-1** (4.0 g, 18.8 mmol) and stirred at room temperature for 18 hours. After the completion of reaction was detected by TLC, the solid was removed by filtration, the filter cake was washed with EA, and the filtrate was washed with water twice, dried with anhydrous sodium sulfate, and concentrated under a reduced pressure to obtain a crude product **5-2**, which was directly used in

25 next step.

c) The product **5-2** in the previous step was dissolved in dioxane (20 mL) in a reaction flask, and slowly and dropwise added with 1 M NaOH to react for 2 hours, and then the

completion of hydrolysis reaction was monitored by TLC to obtain a product **5-3**. The product **5-3** was added with 20 mL of saturated NaHCO₃, slowly added with Fmoc-OSu dissolved in dioxane (10 mL), and stirred at room temperature overnight. After the completion of reaction was monitored by TLC, the reaction solution was adjusted to be faintly acid with 1 M HCl, extracted with EA, dried with anhydrous sodium sulfate, concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM:MeOH = 10:1) to obtain a product **5-4** (3.60 g, yield 89%).

d) The product **5-3** (3.60 g, 8.0 mmol), NBS (1.10 g, 9.6 mmol), EDCI (1.85 g, 9.6 mmol), and a solvent DCM (50 mL) were added into a reaction flask. The mixture was reacted at room temperature for 18 hours. After the completion of reaction was monitored by TLC, the reaction solution was washed with water for 3 times, dried with anhydrous sodium sulfate, and concentrated under a reduced pressure to obtain a product **5-5** (3.20 g, yield 75%).

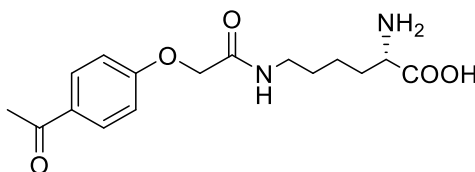
e) The product **5-5** (3.20 g, 6.0 mmol), a solvent dioxane (40 mL), water (10 mL), and Fmoc-Lys-OH hydrochloride (3.0 g, 7.2 mmol) were added into a reaction flask, and then triethylamine (2.0 mL, 15.0 mmol) was added. The mixture was stirred at room temperature for 18 hours until the reaction was completed. The solution was added with an appropriate amount of 1 M HCl to adjust a pH value to be about 2, and extracted with EA. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM: MeOH: AcOH = 20:1:0.5) to obtain a product **5-5** (3.50 g, yield 75%).

f) The product **5-5** was dissolved in DCM (20 mL) in a reaction flask, and added with diethylamine (20 mL) to react at room temperature for 6 hours. After the completion of reaction was monitored by TLC, the reaction solution was concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM: MeOH: H₂O=30:10:1) to obtain a final product **5-7** of white powder (0.55 g, yield 37%).

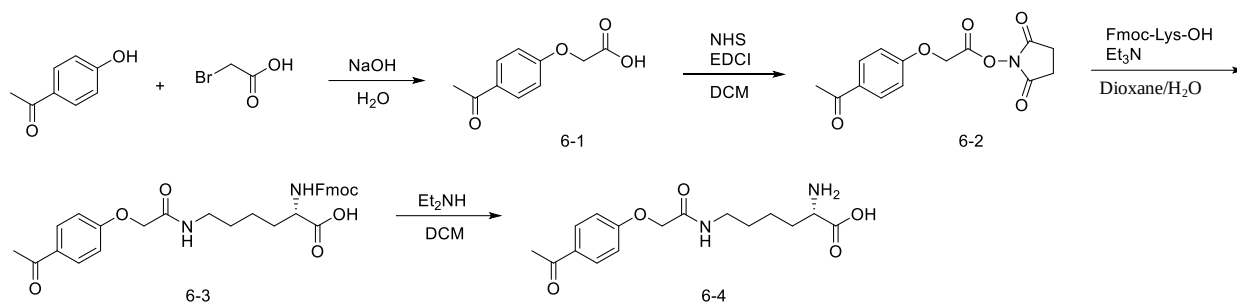
¹H-NMR (400 MHz, heavy water) δ 7.98 (d, *J* = 8.2 Hz, 2H), 7.50 (d, *J* = 8.2 Hz, 2H), 3.84 (s, 2H), 3.71 (s, 1H), 3.31 (s, 2H), 3.17 (t, *J* = 6.9 Hz, 2H), 2.67 (s, 3H), 1.97 – 1.73 (m, 2H), 1.58 – 1.45 (m, 2H), 1.44- 1.27 (m, 2H).

Example 6 Preparation of unnatural amino acid NPOK-2

The structural formula of NPOK-2 was as follows:



The reaction process was shown in the following figure:



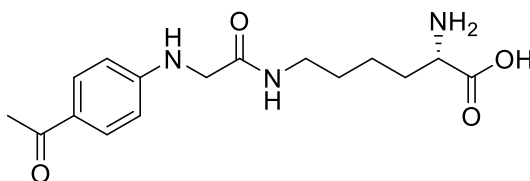
The preparation process included the following steps.

- P*-acetylphenol (2.05 g, 15.0 mmol) and a bromoacetic acid (2.50 g, 18.0 mmol) were added into a reaction flask, and then an aqueous solution (6 mL) of NaOH (1.20 g, 30 mmol) was added. The mixture was refluxed at 100°C for 24 hours until the reaction was completed. The reaction solution was cooled to room temperature, adjusted to be acidic by adding 1 M hydrochloric acid to precipitate a solid, and filtered to obtain a white crude product **6-1** (3.32 g, yield 113%), which was directly used in next step without purification.
- The crude product **6-1** (3.32 g, 17.0 mmol) in the previous step was dissolved in DCM (50 mL) in a reaction flask, and added with NHS (2.35 g, 20.4 mmol) and EDCI (3.90 g, 20.4 mmol). The mixture was stirred at room temperature for 18 hours until the reaction was completed. The reaction solution was extracted with DCM. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM: MeOH: AcOH = 20:1:0.5) to obtain a product **6-2** (1.67 g, yield 38%).
- The product **6-2** (1.67 g, 5.7 mmol), a solvent dioxane (20 mL) and water (50 mL), and Fmoc-Lys-OH hydrochloride (1.9 g, 4.8 mmol) were added into a reaction flask, and then triethylamine (1.7 mL, 12.0 mmol) was added. The mixture was stirred at room temperature for 18 hours until the reaction was completed. The solution was added with an appropriate amount of 1 M HCl to adjust a pH value to be about 2, and extracted with EA. Organic phases were combined, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure. The crude product **6-3** obtained was directly dissolved in DCM (10 mL), and then added with diethylamine (5 mL). The mixture was stirred at room temperature for 18 hours until the reaction was completed. The reaction solution was concentrated under a reduced pressure, and purified by column chromatography (eluent: DCM: MeOH: H₂O = 40:10:1) to obtain a final product **6-4** (709 mg, two-step yield 39%).

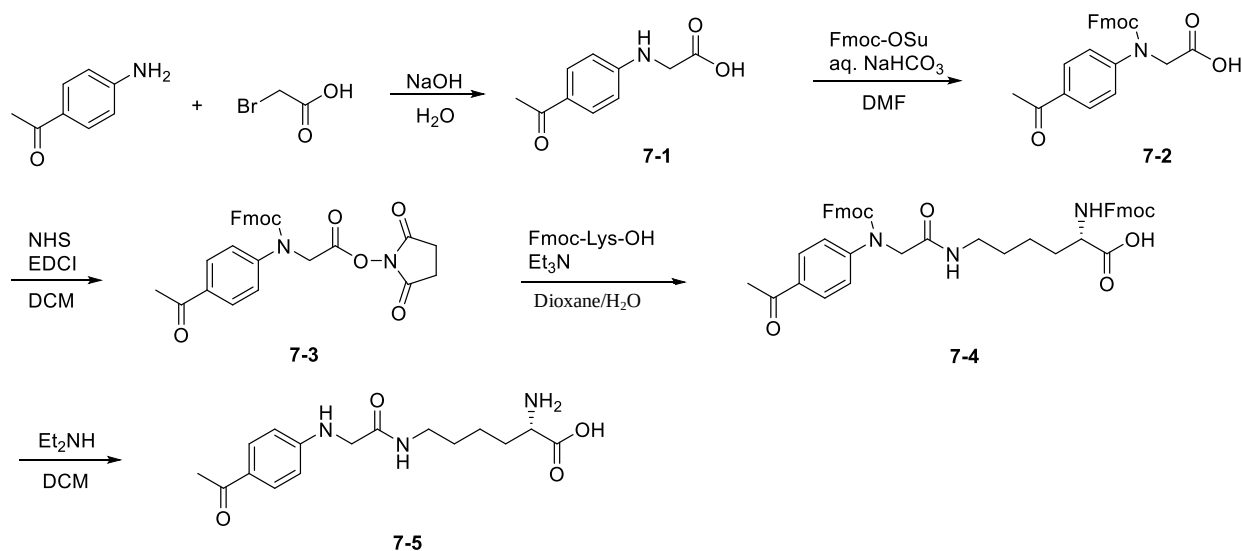
¹H-NMR (400 MHz, heavy water) δ 7.91 (d, J = 8.8 Hz, 2H), 6.98 (d, J = 8.8 Hz, 2H), 4.60 (s, 2H), 3.59 (t, J = 6.4 Hz, 1H), 3.19 (t, J = 6.8 Hz, 2H), 2.52 (s, 3H), 1.87 – 1.65 (m, 2H), 1.56 – 1.40 (m, 2H), 1.35 – 1.15 (m, 2H).

Example 7 Preparation of unnatural amino acid NBGK-2

The structural formula of NBGK-2 was as follows:



The reaction process was shown in the following figure:



5

The preparation process included the following steps.

a) A bromoacetic acid (2.10 g, 15.0 mmol) and an aqueous solution (10 mL) of NaOH (0.80 g, 20 mmol) were added into a reaction flask and stirred for 10 minutes, and then *p*-acetylaniline (1.40 g, 10.0 mmol) was added. The mixture was refluxed at 100°C for 18 hours until the reaction was completed. The reaction solution was cooled to room temperature, filtered, and washed with water to obtain a white crude product **7-1** (1.30 g, yield 67%), which was directly used in next step without purification.

b) The product **7-1** (1.30 g, 6.7 mmol) and an aqueous solution (20 mL) of NaHCO₃ (1.70 g, 20.1 mmol) were added into a reaction flask, and then Fmoc-OSu (2.80 g, 8.1 mmol) and DMF (20 mL) were added. The mixture was stirred at 60°C for 18 hours until the reaction was completed. The reaction solution was cooled to room temperature, and extracted with EA. An aqueous phase was retained and adjusted to have a pH value of about 2 with 1 M hydrochloric acid. Subsequently, the reaction solution was extracted with EA to obtain organic phases, which were dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a product **7-2**, which was directly used in next step without purification.

c) The product **7-2** (about 6.7 mmol) in the previous step, NHS (0.90 g, 8.0 mmol) and EDCI (1.50 g, 8.0 mmol) were dissolved in DMF (50 mL) in a reaction flask. The reaction mixture was stirred at room temperature for 24 hours until the reaction was completed. The

reaction solution was added with water, and extracted with DCM to obtain organic phases, which were dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a product **7-3**, which was directly used in next step without purification.

d) The product **7-3** (about 6.7 mmol) in the previous step, Fmoc-Lys-OH hydrochloride (2.30 g, 5.6 mmol) and triethylamine (2.0 mL, 14.0 mmol) were added into a reaction flask. The reaction mixture was stirred at room temperature for 3 hours until the reaction was completed. Subsequently, the reaction solution was adjusted to have a pH value of about 2 with 1 M hydrochloric acid, extracted with EA, dried with anhydrous sodium sulfate, filtered, and concentrated under a reduced pressure to obtain a product **7-4**, which was directly used in next step without purification.

e) The product **7-4** in the previous step, a solvent DCM (20 mL) and diethylamine (10 mL) were added into a reaction flask. The reaction mixture was stirred at room temperature for 12 hours until the reaction was completed. The reaction solution was concentrated under a reduced pressure first, added with acetonitrile (50 mL) to dissolve again, and then concentrated under a reduced pressure, and the operation was repeated for 3 times to remove excess diethylamine. The reaction solution was added with DCM for pulping twice to obtain a final product **7-5** (1.65 g, total yield 51%).

¹H-NMR (400 MHz, heavy water) δ 7.71 (d, J = 8.8 Hz, 2H), 6.51 (d, J = 8.8 Hz, 2H), 3.80 (s, 2H), 3.53 (t, J = 6.8 Hz, 1H), 3.09 (t, J = 6.8 Hz, 2H), 2.39 (s, 3H), 1.73 – 1.60 (m, 2H), 1.42 – 1.33 (m, 2H), 1.25 – 1.14 (m, 2H).

Example 8

Recombinant human growth hormones were prepared in a prokaryotic expression system with NPOK, NPAK, NBOK, NBPK, NBGK, NPOK-2 and NBGK-2 prepared in Examples 1 to 7, and PEG site-specific conjugates were prepared.

(1) Acquisition of helper plasmid

A helper plasmid pSupAR-MbPylRS was purchased from the plasmid preservation organization Addgene (catalog number: #91705). The plasmid could express a tRNA specifically recognizing an unnatural amino acid derived from pyrrolysine in *Escherichia coli* and a tRNA synthetase. The helper plasmid was extracted after shake-flask culture in an LB medium added with 37.5 mg/L chloramphenicol.

(2) Construction of expression plasmid of recombinant human growth hormone containing termination codons in reading frame

An mRNA sequence of a coding gene of a Homo sapiens growth hormone (with an amino acid sequence as shown in SEQ ID NO: 1) was obtained from the database of the National Center for Biotechnology Information of the United States, a purification tag consisting of six histidines was added to the C-terminal of the translated protein thereof. Meanwhile, the codons of the amino acid at the 107th site in SEQ ID NO: 1 was changed to amber codons (TAG), and then the

complete DNA sequence was synthesized by whole gene synthesis to obtain a gene sequence (SEQ ID NO: 2) of the recombinant human growth hormone.

The amino acid sequence (SEQ ID NO: 1) of the Homo sapiens growth hormone was as follows:

5 FPTIPLSRLFDNAMLRAHRLHQLAFDITYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP
TPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEG
IQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDALLKNYGLLYCFRKDMDKVETFLRIV
QCRSVEGSCGF

10 The gene sequence (SEQ ID NO: 2) of the recombinant human growth hormone was as follows:

ATGTTTCCTACTATAACCACTATCTCGTCTATTCGATAACGCTATGCTTCGGGCCCCAT
CGTCTTCATCAGCTGGCCTTTGACACCTACCAGGAGTTTGAAGAAGCCTATATCCCAA
AGGAACAGAAGTATTCATTCCTGCAGAACCCCCAGACCTCCCTCTGTTTCTCAGAGTC
TATTCCGACACCCTCCAACAGGGAGGAAACACAACAGAAATCCAACCTAGAGCTGCT
15 CCGCATCTCCCTGCTGCTCATCCAGTCGTGGCTGGAGCCCGTGCAGTTCCTCAGGAGT
GTCTTCGCCAACAGCCTGGTGTACGGCGCCTCTTAGAGCAACGTCTATGACCTCCTAA
AGGACCTAGAGGAAGGCATCCAAACGCTGATGGGGAGGCTGGAAGATGGCAGCCCC
CGGACTGGGCAGATCTTCAAGCAGACCTACAGCAAGTTCGACACAACTCACACAAC
GATGACGCACTACTCAAGAACTACGGGCTGCTCTACTGCTTCAGGAAGGACATGGAC
20 AAGGTCGAGACATTCCTGCGCATCGTGCAGTGCCGCTCTGTGGAGGGCAGCTGTGGC
TTCCTCGAGCACCACCACCACCACCTAA

The gene sequence (SEQ ID NO: 2) of the recombinant human growth hormone was cloned between two enzyme cutting sites *Nde* I and *Xho* I of pET21a (Novagen, catalog number: #69740-3) through one-step subcloning to obtain an expression plasmid pET21-rhGH107, which
25 was confirmed to be consistent with the expected sequence by sequencing. The pET21-rhGH107 could be used to express the recombinant human growth hormone with the codons of the amino acid at the 107th site changed to the amber codons, and the C-terminal of the protein contained the purification tag consisting of six histidines. The profile of the expression plasmid pET21-rhGH107 of the recombinant human growth hormone was shown in FIG. 1.

30 (3) Acquisition of target expression strain

The helper plasmid pSupAR-MbPylRS and the expression plasmid pET21-rhGH107 above were co-transformed into competent cells of *Escherichia coli* OrigamiB (DE3) (Novagen, catalog number: #70911-3), and a double-resistant strain was obtained by screening with an LB medium containing 100 mg/L ampicillin and 37.5 mg/L chloramphenicol, i.e., an expression strain of the
35 recombinant human growth hormone.

(4) Expression of recombinant protein inserted with unnatural amino acids

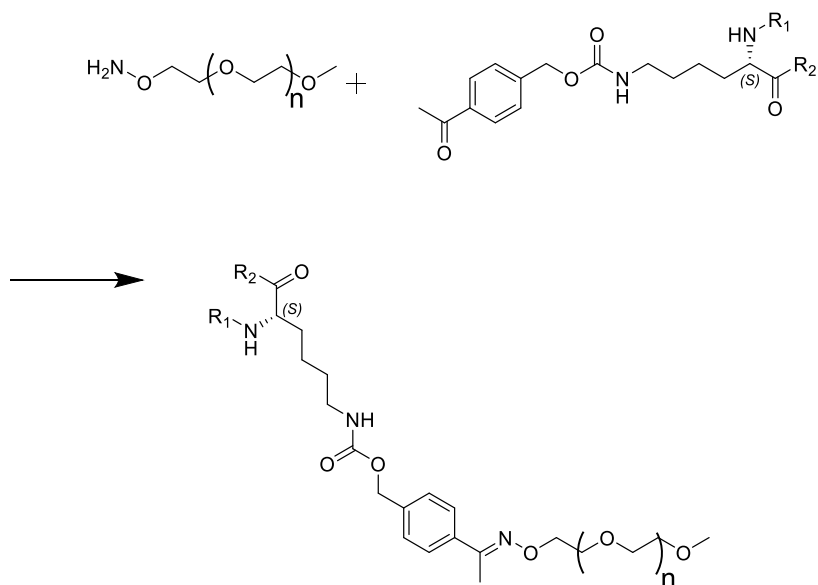
The screened expression strain of the recombinant human growth hormone was inoculated into eight parts of 2×YT medium (16 g/L yeast extract, 10 g/L tryptone and 5 g/L NaCl, containing 100 mg/L ampicillin and 37.5 mg/L chloramphenicol), and cultured at 37°C until
40 OD600 of the bacterial liquid was 2.0 ± 0.2. Eight aliquots were respectively added with IPTG (with a final concentration of 1 mM) and arabinose (with a final concentration of 0.2%), the first

aliquot was added with NBOK (with a final concentration of 1 mM), the second aliquot was added with NPAK (with a final concentration of 1 mM), the third aliquot was added with NBPK (with a final concentration of 1 mM), the fourth aliquot was added with NBGK (with a final concentration of 1 mM), the fifth aliquot was added with NPOK (with a final concentration of 1 mM), the sixth aliquot was added with NPOK-2 (with a final concentration of 1 mM), the seventh aliquot was added with NBGK-2 (with a final concentration of 1 mM), and the eighth aliquot was not added with the unnatural amino acid as a negative control. After the eight aliquots were cultured at 37°C to induce expression for 5 hours to 6 hours, 1 ml of each culture solution was taken, centrifuged at 10,000 rpm for 1 minute, and resuspended with PBS until OD600 was 10, and each aliquot was taken for SDS-PAGE electrophoresis. The SDS-PAGE electrophoretogram of each strain referred to FIG. 2. Results in FIG. 2 showed that the expression strains respectively added with the seven unnatural amino acids could express the target proteins.

(5) Purification of recombinant protein inserted with unnatural amino acid

Each aliquot after the induced expression above was centrifuged at 10,000 rpm for 5 minutes, a precipitate was taken and added with an NTA Buffer (20 mM Tris-HCl at pH 7.9, 0.5 M NaCl and 10% glycerol) with 1/20 cell growth volume and PMSF with a final concentration of 1 mM, and the cells were homogenized and disrupted by ultrasonic wave. The solution was centrifuged at 10,000 rpm for 20 minutes, a supernatant was taken, and a target protein with a His tag was adsorbed by a Ni-NTA chromatography column, and then eluted with an eluent NTA Buffer (20 mM Tris-HCl at pH 7.9, 0.5 M NaCl, 10% glycerol and 250 mM imidazole) to obtain a target protein sample (i.e., the recombinant protein inserted with the unnatural amino acid) with a purity of about 90%. Each sample was placed in a 10 K nitrocellulose dialysis bag, and dialyzed in a PBS buffer at pH 7.0 (200 ml of dialysate per 1 ml of protein solution) overnight, and the dialysate was changed once. The protein solution after dialysis was collected, and the protein concentration was detected by SDS-PAGE.

(6) Coupling reaction between recombinant protein inserted with unnatural amino acid and PEG



As shown in the synthetic route above (wherein the direction from of R₁ to R₂ was the direction from the N-terminal to the C-terminal of the amino acid sequence).

Taking an oximation reaction of 30 KD aminoxy PEG coupled with the recombinant protein inserted with the unnatural amino acid as an example, the coupling reaction was operated as follows: before the coupling reaction, the target protein obtained above was adjusted to be 0.5 mg/ml with a PBS buffer at pH 7.0, added with a 30 KD aminoxy PEG solid (from Beijing Jenkem Technology Co., Ltd.) according to 1:15 (a molar ratio of the recombinant protein to the aminoxy PEG), and fully shaken for dissolution to obtain a clear and transparent solution. Subsequently, the reaction solution was sealed, and shaken in a constant temperature shaker (25°C, 100 rpm). The coupling situation was analyzed by SDS-PAGE 48 hours later, as shown in FIG. 3. Results in FIG. 3 showed that the six target proteins all achieved 30 KD PEG coupling, and further indicated that the seven unnatural amino acids above were inserted into the target proteins. Coupling products of the recombinant protein inserted with the unnatural amino acid coupled with PEG were abbreviated as "PEG-rhGH".

(7) Purification of PEG-rhGH

The PEG-rhGH after coupling was separated and purified by a Butyl HP hydrophobic chromatography column (loading buffer: 50 mM NaH₂PO₄ · 2H₂O, and 0.8 M (NH₄)₂SO₄ at pH 8.5; elution buffer: 20 mM Tris-HCl at pH 8.5, and 40 times column volume gradient elution) to obtain pure PEG-rhGH of an electrophoresis pure grade (>95%). Taking PEG-rhGH containing NPOK and PEG-rhGH containing NBOK as examples, the electrophoregram was shown in FIG. 4.

Other PEG-rhGH containing the unnatural amino acid of the present invention also reached the electrophoresis pure grade after purification.

(8) Activity analysis of PEG-rhGH

The specific process was as follows: HEK293-GHR cells (cell strains referred to Chinese patent: 2020112649827) were cultured with a DMEM (Gibco, catalog number: 11995040) medium containing 0.1 mg/mL hygromycin B (Sangon Biotech (Shanghai) Co., Ltd., catalog number: A600230-0001), 0.75 mg/mL G418 (Sangon Biotech (Shanghai) Co., Ltd., catalog number: A600958-0005) and 10% fetal bovine serum (Gibco, catalog number: 10099141) at 37°C in 5% CO₂ to a sufficient amount, and 9×10⁴ HEK293-GHR cells were inoculated into each well of a black 96-well cell culture plate (Corning, catalog number: 3904), with 90 μL/well, and cultured in a starved manner at 37°C in 5% CO₂ for 16 hours. A rhGH standard product (purchased from National Institutes for Food and Drug Control, referred to as "NIFDC"), and samples of rhGH containing NPOK (i.e., rhGH (NPOK)), PEG-rhGH of PEG-modified NPOK (i.e., PEG-rhGH (NPOK)), PEG-rhGH containing NBOK (i.e., rhGH (NBOK)), PEG-rhGH of PEG-modified NBOK (i.e., PEG-rhGH (NBOK)), a commercially available rhGH drug Jintropin® (Changchun GeneScience Pharmaceutical Co., Ltd.) and a commercially available PEG-rhGH drug Polyethylene Glycol Recombinant Human Somatropin Injection® (Changchun GeneScience Pharmaceutical Co., Ltd.) were diluted in gradient, with a total of nine concentrations, and cultured at 37°C in 5% CO₂ for 4 hours. 50 μL of ONE-Glo™ (Promega, catalog number: E6120) detection reagent was added into each well, and the chemiluminescence

value was read by microplate reader. The curve was fitted according to the sample concentration and the reading value of the microplate reader, and EC50 was calculated. Results were shown in FIG. 5A to FIG. 5G and Table 1. The results showed that the cell activity of each site before coupling was equivalent to that of the rhGH standard product from NIFDC. The cell activity after coupling was reduced, which was about 25% to 50% of that of the standard product, and the change rule of biological activity was consistent with that of PEG-rhGH on the market.

The cell activity of other PEG-rhGH containing the unnatural amino acid of the present invention after coupling was also reduced to 25% to 50% of that of the standard product, and a change rule of biological activity was also consistent with that of PEG-rhGH on the market.

Table 1

Name of sample	EC50 (ng/mL)	Relative activity (%)
rhGH standard product of NIFDC	12.0	100%
rhGH (NPOK)	14.3	84%
PEG-rhGH (NPOK)	44.3	27%
rhGH (NBOK)	13.0	92%
PEG-rhGH (NBOK)	25.3	47%
Jintropin [®]	12.6	95%
Polyethylene Glycol Recombinant Human Somatropin Injection [®]	68.6	18%

Example 9

Recombinant GFP proteins were prepared in a eukaryotic expression system with unnatural amino acids NPOK, NPAK, NBOK, NBPK, NBGK, NPOK-2 and NBGK-2 prepared in Examples 1 to 7.

(1) Acquisition of helper plasmid

A helper plasmid pCMV-MbPylRS was purchased from the plasmid preservation organization addgene (catalog number: #91706), and the plasmid encoded an aminoacyl tRNA synthetase specifically recognizing an unnatural amino acid derived from pyrrolysine in mammalian cells and a corresponding tRNA (recognizing the amber codons UAG).

(2) Construction of expression vector of green fluorescent protein containing amber codons in gene reading frame

An expression vector expressing a wild-type green fluorescent protein (with a gene coding sequence shown in SEQ ID NO: 3) containing a purification tag was subjected to point mutation

to obtain an expression plasmid pEGFP40 with the codons of the amino acid at the 40th site in the reading frame mutated into amber codons, and the complete sequence of the expression plasmid pEGFP40 was shown in SEQ ID NO: 4.

5 The gene sequence (SEQ ID NO: 3) of the wild-type green fluorescent protein containing the purification tag was as follows:

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTTCGAG
CTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGA
TGCCACCTAGGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGT
GCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTAC
10 CCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTC
CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTG
AAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG
GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACAGCCACAACGT
CTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCA
15 CAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCAT
CGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCT
GAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGC
CGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGAAGTACGAAGATTACAAGGA
TGACGACGATAAGTAA

20 The gene sequence (SEQ ID NO: 4) of the expression plasmid pEGFP40 was as follows:

GTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTG
ATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATG
GAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACC
CCCGCCCATGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTC
25 CATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAG
TGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGG
CATTATGCCCAGTACATGACCTTATGGGACTTTTCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATGACGTCAATGGGAGTTTGT
30 TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC
AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAA
CCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGG
GACCGATCCAGCCTCCGGACTCTAGAGGATCCCCGCCACCATGGTGAGCAAGGGCGA
GGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGGCGACGTAAACGG
35 CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAGGGCAAGCTGA
CCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCCTCGTGA

CCACCCTGACCTACGGCGTGCA GTGCTTCAGCCGCTACCCCGACCACATGAAGCAGC
 ACGACTTCTTCAAGTCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTT
 CAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCC
 TGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGG
 5 GGCACAAGCTGGAGTACA ACTACAACAGCCACAACGTCTATATCATGGCCGACAAGC
 AGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGC
 GTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG
 CTGCCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG
 AAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCA
 10 TGGACGAGCTGTACAAGAAGTACGAAGATTACAAGGATGACGACGATAAGTAAGAAT
 TCCGGAAGGGTTCGATCCCTACCGGTTAGTAATGAGTTTAAACGGGGGAGGCTAACTG
 AAACACGGAAGGAGACAATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGA
 CAGAATAAAACGCACGGGTGTTGGGTTCGTTTGTTCATAAACGCGGGGTTCGGTCCCA
 GGGCTGGCACTCTGTCGATACCCACCGAGACCCCATTTGGGGCCAATACGCCCCGCGTT
 15 TCTTCCTTTTCCCCACCCCAACCCCAAGTTCGGGTGAAGGCCAGGGCTCGCAGCC
 AACGTGCGGGCGGCAGGCCCTGCCATAGCAGATCTGCGCAGCTGGGGCTCTAGGGGG
 TATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTGGTTACGCGCA
 GCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCC
 TTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGCATCCCTTTAGG
 20 GTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGT
 TCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCA
 CGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAACACTCAACCCTATCTCGGTC
 TATTCTTTTGATTTATAAGGGATTTTGGGGATTTCGGCCTATTGGTTAAAAAATGAGCTG
 ATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGG
 25 AAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTC
 AGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCAT
 GCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCCATCCCGCCCCCTAA
 CTCCGCCCAGTTCCGCCCATCTCCGCCCATGGCTGACTAATTTTTTTTATTTATGCAG
 AGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTTGG
 30 AGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATC
 AAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTT
 CTCCGGCCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCG
 GCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGT
 CAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATC
 35 GTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGC
 GGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCA

CCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACG
 CTTGATCCGGCTACCTGCCCATTGACCACCAAGCGAAACATCGCATCGAGCGAGCAC
 GTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGG
 GGCTCGCGCCAGCCGAACGTTCGCCAGGCTCAAGGCGCGCATGCCCCGACGGCGAG
 5 GATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCC
 GCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACAT
 AGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTC
 CTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGCAGCGCATCGCCTTCTATCGCCTTCT
 TGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGCGAAATGACCGACCAAGCGACGC
 10 CCAACCTGCCATCACGAGATTTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTT
 CGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTG
 GAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAA
 TAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGT
 CCAAACCTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGG
 15 CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCACAC
 AACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAA
 CTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCAGTCGGGAAACCTGTCGTGCC
 AGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCT
 CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCGAGCGGT
 20 ATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGA
 AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
 GCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCA
 AGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGA
 AGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTT
 25 TCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCG
 GTGTAGGTCGTTTCGTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGAC
 CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
 GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
 TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGT
 30 ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
 GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGC
 GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
 GTGGAACGAAAACTCACGTAAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTC
 ACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAA
 35 ACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCT
 ATTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGG

GCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC
 AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGC
 AACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT
 CGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACG
 5 CTCGTCGTTTGGTATGGCTTCATTAGCTCCGGTTCCCAACGATCAAGGCGAGTTACAT
 GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
 AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTA
 CTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTC
 TGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATA
 10 CCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG
 AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
 CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG
 GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCA
 TACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGAT
 15 ACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGGTTCCGCGCACATTTCCCCGA
 AAAGTGCCACCTGACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTTCGACTCT
 CAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGT
 TGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGA
 CCGACAATTGCATGAAGAATCTGCTTAGG

20 (3) Expression of recombinant GFP protein inserted with unnatural amino acid

Chinese hamster ovary cells CHO-K1 (catalog number: #CCL-61-ATC) were purchased from American Type Culture Collection (ATCC), and cultured in an adherent manner with an RPMI1640 medium containing 10% fetal bovine serum. The helper plasmid pCMV-MbPylRS and the expression plasmid pEGFP40 of the green fluorescent protein in the above steps were
 25 extracted, and transiently transfected with a lipo2000 transfection reagent (Invitrogen company, catalog number: #12566014) according to the instructions (cells were inoculated into a 24-well plate as 50,000 cells/well, with a total of eight wells, and were transfected at 24 hours after inoculation, with 500 ng of plasmid per well). During transfection, the first well was added with NPOK (with a final concentration of 1 mM), the second well was added with NPAK (with a final
 30 concentration of 1 mM), the third well was added with NBOK (with a final concentration of 1 mM), the fourth well was added with NBPK (with a final concentration of 1 mM), the fifth well was added with NBGK (with a final concentration of 1 mM), the sixth well was added with NPOK-2 (with a final concentration of 1 mM), the seventh well was added with NBGK-2 (with a final concentration of 1 mM), and the eighth well was not added with the unnatural amino acid as
 35 a negative control. The cells were observed with a fluorescence microscope after 48 hours culturing in a CO₂ incubator, and results showed that there was obvious green fluorescence in the

first well to the seventh well (referring to FIG. 6), which proved that the unnatural amino acid could be inserted into the green fluorescent protein to obtain a complete green fluorescent protein without affecting the function of the fluorescent protein.

Example 10

Anti-HER2 monoclonal antibodies containing unnatural amino acids were prepared in a eukaryotic expression system with NPOK, NPAK, NBOK, NBPK, NBGK, NPOK-2 and NBGK-2 prepared in Examples 1 to 7, and conjugates with toxin were prepared.

(1) Acquisition of helper plasmid

A helper plasmid pCMV-MbPyIRS was purchased from the plasmid preservation organization addgene (catalog number: #91706), and the plasmid encoded an aminoacyl tRNA synthetase specifically recognizing an unnatural amino acid derived from pyrrolysine in mammalian cells and a corresponding tRNA (recognizing the amber codons UAG).

(2) Construction of expression vector of anti-HER2 antibody (trastuzumab) containing amber codons in gene reading frame

Heavy-chain and light-chain DNAs (corresponding amino acid sequences were SEQ ID NO: 5 and SEQ ID NO: 6 respectively) encoding trastuzumab were synthesized by whole gene synthesis, and subcloned into a eukaryotic expression vector pCDNA3.1+, then the obtained expression vector was subjected to point mutation to obtain an expression plasmid pCDNA3.1-Trastuzumab-UAG142 with the codons of the amino acid at the 142nd site of the heavy-chain reading frame mutated into amber codons. The profile of the plasmid was shown in FIG. 7, and the complete sequence of the plasmid was shown in SEQ ID NO: 7.

The heavy-chain amino acid sequence (SEQ ID NO: 5) of trastuzumab was as follows:

EVQLVESGGGLVQPGGSLRLSCAASG FNIKDTYIHVVRQAPGKGLEWVARIYPTNG
YTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWGGDGFYAMDYWGQGT
LVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPA
VLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE
LLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR
EEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL
PPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTV
DKSRWQQGNV FSCSV MHEALHNHYTQKSLSLSPGK

The light-chain amino acid sequence (SEQ ID NO: 6) of trastuzumab was as follows:

DIQMTQSPSSLSASVGDRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFLYS
G VPSRFGSGRSGTDFTLTISLQPEDFATYYCQQHYTTPPTFGQGTKVEIKRTVAAPSVFIFPP
SDEQLKSGTASVVCLLNFPYPREAKVQWKVDNALQSGNSQESVTEQDSKDYSLSSSTL
TLISKADYEKHKVYACEVTHQGLSPVTKSFNRGEC

The gene sequence (SEQ ID NO: 7) of the expression plasmid

pCDNA3.1-Trastuzumab-UAG142 was as follows:

GACGGATCGGGAGATCTCCCGATCCCCCTATGGTGCACCTCTCAGTACAATCTGCTCT
GATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGT
AGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGA
5 AGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACG
CGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCAT
AGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGG
10 CAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAA
TGGCCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTA
CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGG
GCGTGGATAGCGGTTTGA CTCACGGGGATTTC AAGTCTCCACCCCATTGACGTCAAT
GGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA ACTCCG
15 CCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGC
TCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACT
ATAGGGAGACCCAAGCTGGCTAGCGTTTAACTTAAGCTTATGTACCGGATGCAGCTG
CTGAGCTGTATCGCCCTGTCTCTGGCCCTCGTGACCAACAGCGAAGTGCAGCTGGTG
GAAAGCGGCGGAGGACTGGTGCAGCCTGGCGGATCTCTGAGACTGAGCTGTGCCGC
20 CAGCGGCTTCAACATCAAGGACACCTACATCCACTGGGTGCGCCAGGCCCTGGCAA
GGGACTGGAATGGGTGGCCAGAATCTACCCACCAACGGCTACACCAGATACGCCGA
CAGCGTGAAGGGCCGGTTCACCATCAGCGCCGACACCAGCAAGAACACCGCCTACCT
GCAGATGAACAGCCTGCGGGCCGAGGACACCGCCGTGTACTACTGTAGTAGATGGGG
AGGCGACGGCTTCTACGCCATGGACTATTGGGGCCAGGGCACCCCTCGTGACAGTGTCT
25 AGTGCGTAGACCAAGGGGCCCTCGGTCTTCCCCCTGGCACCCCTCCTCCAAGAGCACC
TCTGGGGGCACAGCGGCCCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCGGTG
ACGGTGTCGTGGA ACTCAGGCGCCCTGACCAGCGGCGTGACACCTTCCCGGCTGTC
CTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGGTGACCGTGCCCTCCAGCAGCT
TGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACACCAAGGTGG
30 ACAAGAAAGTTGAGCCCAAATCTTGTGACAAA ACTCACACATGCCCACCGTGCCCAG
CACCTGAACTCCTGGGGGGACCGTCAGTCTTCCTCTTCCCCC AAAACCCAAGGACA
CCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACG
AAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCA
AGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTC
35 ACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAAC
AAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGA

GAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTC
 AGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAG
 AGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGGACTCCGAC
 GGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGG
 5 AACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGA
 GCCTCTCCCTGTCTCCCGGCAAGTGATAAGGCCGGCCCCCTCTCCCTCCCCCCCCCCTA
 ACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGGTGTGCGTTTGTCTATATGTTATTT
 TCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCCCCGGAAACCTGGCCCTGTCTTCTT
 GACGAGCATTCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAAT
 10 GTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCG
 ACCCTTTGCAGGCAGCGGAACCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAG
 CCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCAGTGCCACGTTGTGAGT
 TGGATAGTTGTGGAAGAGTCAAATGGCTCTCCTCAAGCGTATTCAACAAGGGGCTG
 AAGGATGCCCAGAAGGTACCCCATTTGTATGGGATCTGATCTGGGGCCTCGGTACACAT
 15 GCTTTACATGTGTTTAGTCGAGGTTAAAAAACGTCTAGGCCCCCGAACCACGGGG
 ACGTGGTTTTCTTTGAAAAACACGATGATAATATGGCCACAACCATGTACCGCATGC
 AACTCCTGTCTTGCATTGCACTAAGTCTTGCACTTGTCAAAACAGTGATATCCAGATG
 ACCCAGAGCCCCAGCAGCCTGTCTGCCAGCGTGGGCGACAGAGTGACCATCACCTGT
 AGAGCCAGCCAGGACGTGAACACCGCCGTGGCCTGGTATCAGCAGAAGCCTGGCAA
 20 GGCCCCCAAGCTGCTGATCTACAGCGCCAGCTTCCTGTACAGCGGCGTGCCCAGCAG
 ATTCAGCGGCAGCAGATCCGGCACCGACTTCACCCTGACCATCAGCTCCCTGCAGCCC
 GAGGACTTCGCCACCTACTACTGCCAGCAGCACTACACCACCCCCCCCCACATTTGGCC
 AGGGCACCAAGGTGGAAATCAAGCGTACGGTGGCTGCACCATCTGTCTTCATCTTCCC
 GCCATCTGATGAGCAGTTGAAATCTGGAACCTGCCTCTGTTGTGTGCCTGCTGAATAAC
 25 TTCTACCCCAGAGAAGCCAAAGTGCAGTGGAAGGTGGACAACGCCCTGCAGAGCGG
 AAACAGCCAGGAAAGCGTGACAGAGCAGGATTCCAAGGATTCCACATACAGCCTGAG
 CAGCACACTGACACTGTCCAAGGCCGACTACGAGAAGCACAAGGTGTACGCCTGCG
 AAGTGACACACCAGGGACTGTCCTCCCCTGTGACAAAGAGCTTCAACAGAGGAGAA
 TGCTGATGATCTAGAGGGGCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTA
 30 GTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGC
 CACTCCCCTGTCTTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGT
 GTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAG
 ACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAA
 CCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGG
 35 CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCG
 CTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCT

CTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCA
AAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTT
TCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGA
ACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTATAAGGGATTTTGCCGATTTCG
5 GCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGG
AATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATG
CAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCA
GCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAAGTCCCGCCCC
TAACTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGC
10 TGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCA
GAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGC
TTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTG
AACAAGATGGATTGCACGCAGGTCTCCGGCCGCTTGGGTGGAGAGGCTATTTCGGCTA
TGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGC
15 GCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACT
GCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCCTTGCGCAGC
TGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCC
GGGGCAGGATCTCCTGTCTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT
GATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTTCGACCACCAAG
20 CGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGG
ATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGCCAGGCTCA
AGGCGCGCATGCCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGC
CGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGG
TGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTT
25 GGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCG
AGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTC
GAAATGACCGACCAAGCGACGCCAACCTGCCATCACGAGATTCGATTCCACCGCC
GCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCC
TCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGC
30 TTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTT
CACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATAC
CGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCT
GGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTT
35 CCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAG
AGGCGGTTTGCGTATTGGGCGCTCTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCG

GTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
 CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
 AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGAC
 GAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
 5 AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
 CGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGC
 TCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGC
 ACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTC
 CAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAG
 10 CAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGG
 CTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA
 AAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACACCGCTGGTAGCGGTTTTTTTG
 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTT
 TTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTAAAGGGATTTTGGTCATG
 15 AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAATAAATGAAGTTTAAATC
 AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG
 CACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGT
 AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCG
 AGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
 20 CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
 GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGC
 TACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCC
 AACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTT
 CGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGG
 25 CAGCACTGCATAATTCTCTTACTGTTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGT
 GAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
 CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT
 TGGAACACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATCCAGT
 TCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGCGT
 30 TTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGA
 CACGGAAATGTTGAATACTCATACTCTTCCTTTTTTCAATATTATTGAAGCATTATCAGG
 GTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGG
 GTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

(3) Insertion of unnatural amino acid

35 HEK293 cells adopted to suspension were inoculated into a Wayne293™ medium (Quacell Biotechnology, catalog number: A21501) at a density of 0.3×10^5 /mL, and cultured by shaking in

a 1 L shake flask with 240 mL of medium under conditions of 120 rpm, 5% CO₂ and 80% humidity. The cells were transfected when the cell density reached about 1×10⁶/mL, the helper plasmid pCMV-MbPylRS and the expression plasmid pCDNA3.1-Trastuzumab-UAG142 described in the steps (1) and (2) were extracted, and the endotoxin was removed for later use.

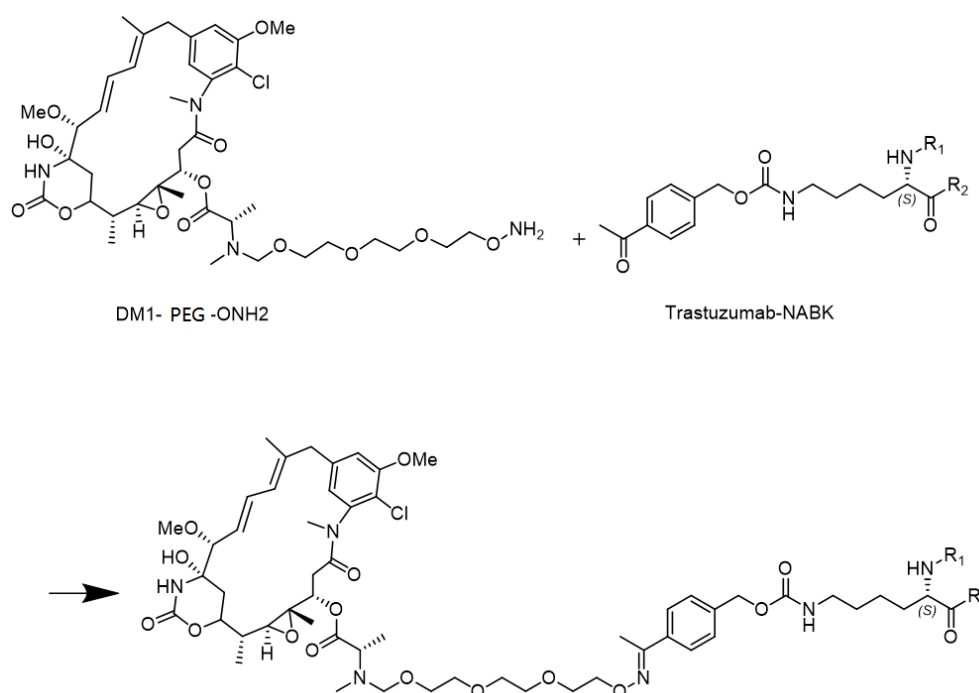
- 5 During transfection of each shake flask, 120 µg of expression plasmid and 120 µg of helper plasmid were used and added into a centrifuge tube and diluted to 7.2 mL with 1×PBS buffer, and then 720 µg of PEI transfection reagent (Polyethyleneimine) was added into another centrifuge tube, and added with 1×PBS buffer to dilute to 7.2 mL. The two centrifuge tubes were respectively left to stand for 5 minutes after mixing evenly, and then liquids in the two centrifuge
- 10 tubes were gently and evenly mixed and allowed to stand for 10 minutes, slowly and dropwise added into 240 mL of cell suspension in the 1 L shake flask by a total of 14.4 mL, continuously and gently shaken during dropwise adding, cultured by shaking under conditions of 120 rpm, 5% CO₂ and 80% humidity for 4 hours, then added with the unnatural amino acid with the final concentration of 1 mM, and continuously cultured under conditions of 120 rpm, 5% CO₂ and
- 15 80% humidity for 5 days, and a cell culture supernatant was collected for antibody purification.

(4) Purification

The cell culture supernatant was purified with 1 ml of HiTrap Protein A prepacked column, and eluted with an elution buffer of 100 mmol/L glycine and 200 mmol/L acetate at pH 3.5 to obtain the purified trastuzumab inserted with the unnatural amino acid.

20 (5) Coupling

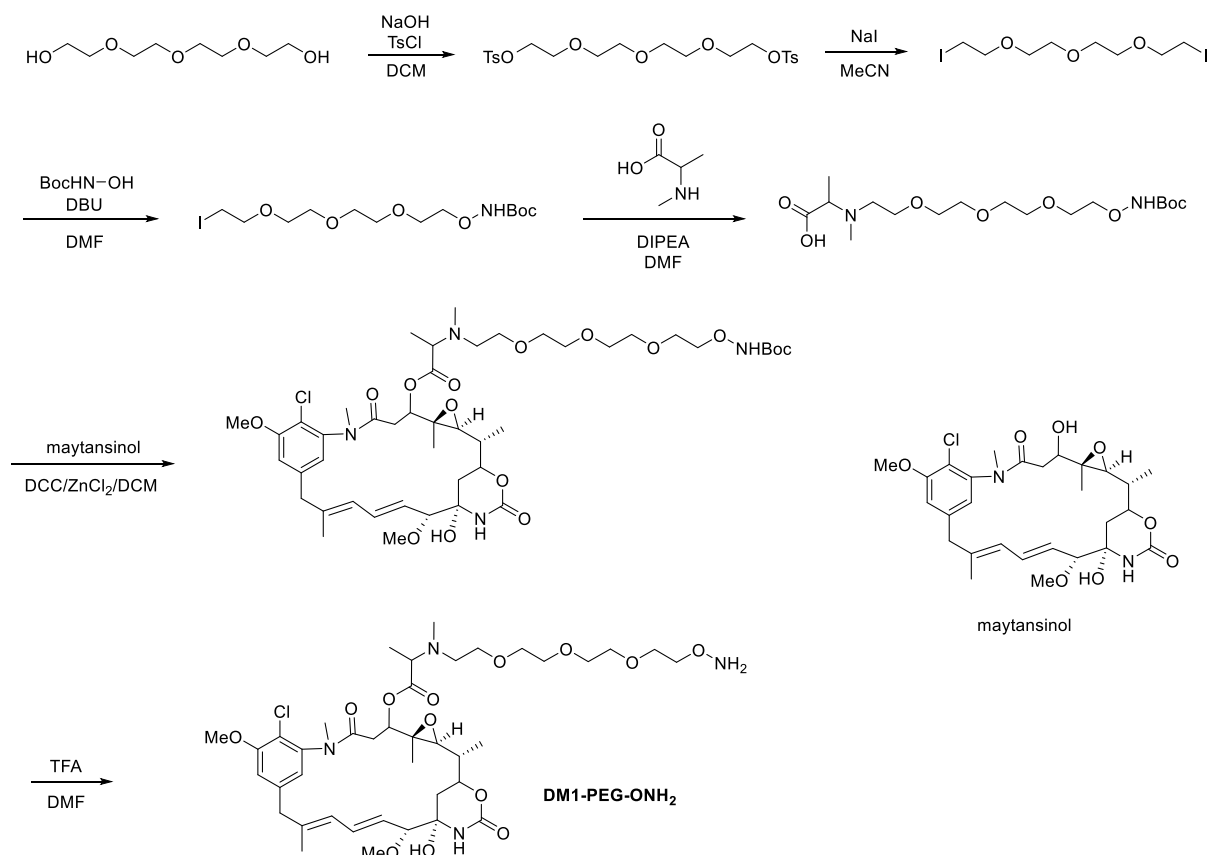
A synthetic route was as follows (taking NBOK as an example):



A toxin containing an amino-oxygen terminal group (DM1-PEG-OH₂) and the purified

trastuzumab inserted with the unnatural amino acid (wherein the direction from R₁ to R₂ was the direction from the N-terminal to the C-terminal of the amino acid sequence) were coupled in a site-specific manner through an oximation reaction to obtain a monoclonal antibody-toxin conjugate.

5 The preparation process of DM1-PEG-ONH₂ was as follows:



Specific operations of the coupling process were as follows: the purified trastuzumab inserted with the unnatural amino acid was mixed and dissolved with the DM1-PEG-ONH₂ in a molar ratio of 1:15, then the pH value was adjusted to be 4.0 with 10 M acetic acid, the mixture was shaken on a shaker (25°C, 200 rpm), and sampled 48 hours later, and the reaction situation of trastuzumab was detected by HPLC (HIC-HPLC) based on the hydrophobic chromatography principle. Results showed that 93.0% of trastuzumab containing NBPK was modified by the toxin, 93.0% of trastuzumab containing NBOK was modified by the toxin, and 87.4% of trastuzumab containing NPOK-2 was modified by the toxin, as shown in FIG. 8A to FIG. 8C. Coupling ratio of trastuzumab = 100% - ratio of remaining uncoupled raw materials.

The coupled monoclonal antibody-toxin conjugate was subjected to buffer change through a 50 kDa ultrafiltration centrifuge tube to remove unreacted toxin raw materials, and the changed buffer was 20 mM histidine buffer (pH 6.5).

HIC-HPLC analysis conditions were as follows:

20 mobile phase A (2 M ammonium sulfate, 75 mM K₂HPO₄, pH 7.2±0.2); and
mobile phase B (75 mM K₂HPO₄, 25% isopropanol, pH 7.2±0.2).

Chromatographic column	TSK gel Butyl-NPR column, 4.6 × 100 mm, 2.5 μm		
Detection wavelength	280 nm		
Temperature of column thermerstat	40°C		
Temperature of sample injector	5°C		
Flow velocity	0.5 ml/min		
Sample volume	10 μl		
Gradient elution	Time (min)	A (%)	B (%)
	0	80	20
	30	40	60
	35	25	75
	40	25	75
	45	80	20
	60	80	20

(6) Activity analysis

BT-474 cell strains (ATCC, catalog number: HTB-20) were used to detect the inhibitory
5 effect of the sample on cell proliferation.

The specific process was as follows: BT-474 cells were cultured with an ATCC Hybri-Care
Medium (ATCC, catalog number: 46-X) containing 10% fetal bovine serum (Gibco, catalog
number: 10099141) at the condition of 37°C and 5% CO₂ to a sufficient amount, 1.5×10⁴ BT-474
cells were inoculated into each well of a 96-well cell culture plate (Corning, catalog number:
10 3599) (including a vehicle control well), and cell-free medium was added into a blank control
well, with 80 μL/well. The trastuzumab containing NBOK and the DM1-modified trastuzumab
containing NBOK were gradiently diluted from 17 μg/mL to 0.13 μg/mL, with a total of eight
concentrations, and two repeated wells for each concentration. The diluted sample was
transferred to the culture plate inoculated with the BT-474 cells with 10 μL/well, and 10 μL of
15 medium was added into the vehicle control well and the blank control well respectively, and
cultured at the condition of 37°C and 5% CO₂. On the fifth day, 10 μL of CCK-8 (Beyotime,
catalog number: C0043) detection reagent was added into each well, and developed at the
condition of 37°C and 5% CO₂ for 6 hours, and the optical density was read by a microplate
reader at a wavelength of 450 nm. The growth inhibition of the detected sample was calculated
20 by the following formula: growth inhibition (%) = (OD compound – OD blank control) / (OD
vehicle control – OD blank control) ×100%. Growth inhibitions of compounds with different

concentrations were calculated in Excel, and then the growth inhibition curve was made by GraphPad Prism7 software and IC50 was calculated. Results were shown in FIG. 9. The results showed that the IC50 of the DM1-modified trastuzumab containing NBOK was about three times lower than that of the trastuzumab containing NBOK, so that the tumor killing effect was significantly improved.

(7) Stability examination of DM1-modified trastuzumab containing unnatural amino acid

The coupled monoclonal antibody-toxin conjugate was subjected to buffer exchange through a 50 kDa ultrafiltration centrifuge tube (Millipore, catalog number: UFC905024#) to remove unreacted toxin raw materials. The exchanged buffer was 20 mM histidine buffer (pH 6.0). The protein solution was divided into three parts, adjusted to have pH 4.0, pH 6.0 and pH 8.0 with 1 M citric acid solution or 1 M Tris solution respectively, placed in a water bath at 25°C, sampled 24 hours, 48 hours, 72 hours, 96 hours and 120 hours later respectively, and analyzed by HIC-HPLC (the detection conditions are the same as above). Results were shown in Table 2.

Table 2 Stability of DM1-modified trastuzumab containing NBOK at 25°C under different pH values

Sampling time	Content of uncoupled trastuzumab containing NOPK/%		
	pH 4.0	pH 6.0	pH 8.0
0 hour	7.0%		
24 hours	7.3%	7.50%	7.61%
48 hours	9.2%	7.50%	7.50%
72 hours	11.5%	7.62%	7.63%
96 hours	13.8%	7.68%	7.71%
120 hours	15.2%	7.50%	7.61%

Results in Table 2 showed that the oxime bond formed in the DM1-modified trastuzumab containing the unnatural amino acid provided by the present invention was still relatively stable at a low pH.

Example 11 Reduction of active group of Lys-azido

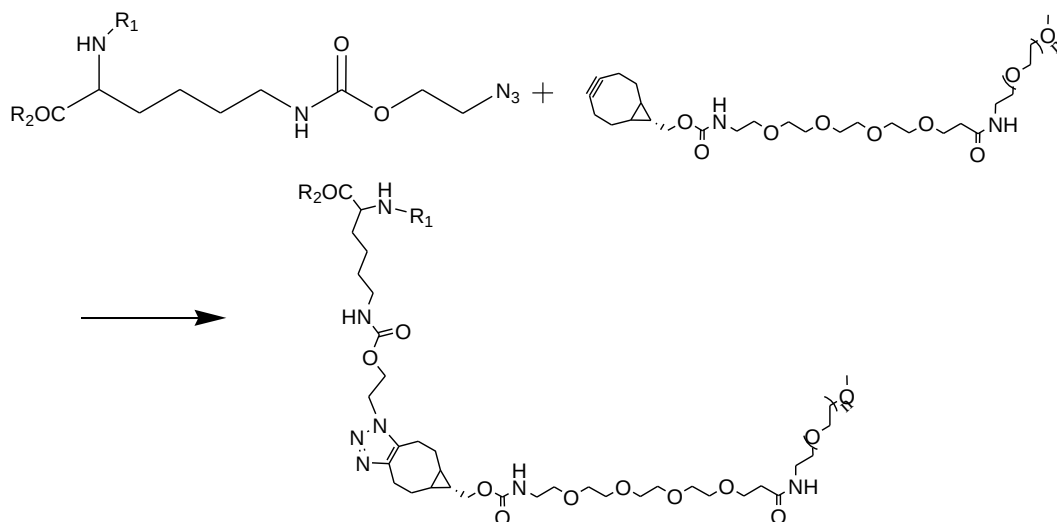
(1) Analysis of molecular weight of product by high-resolution mass spectrometry

According to the expression method of the secretory growth hormone in a reference (Lihua Song, et al., Journal of Biology, 16 (1): 6-8, 1999), an OmpA signal peptide coding sequence was added to the N-terminal of rhGH on the expression vector according to the signal peptide sequence provided in the reference, and in combination with the helper plasmid pSupAR-MbPylRS in Example 8, an OrigamiB (DE3) expression strain of rhGH with the K140

codons mutated into amber codons (TAG) was constructed. Natural rhGH with a mutation into Lys-azido at the 140th site was expressed by adding Lys-azido during fermentation, which was named rhGH-Lys-azido-140. Natural rhGH with a mutation into NBOK at the 140th site was expressed by adding NBOK during fermentation, which was named rhGH-NBOK-140. Moreover, purification was carried out by corresponding purification means in the reference above. Complete molecular weights of the rhGH-Lys-azido-140 and the rhGH-NBOK-140 were analyzed through liquid chromatography-mass spectrometry (high-resolution mass spectrometer: XevoG2-XS Q-Tof, Waters Corporation; Ultra-high performance liquid chromatography: UPLC (Acquity UPLC I-Class), Waters Corporation).

As shown in FIG. 10A, a component with a molecular weight that was 26 Da less than the theoretical molecular weight (22,265 Da) appeared in the rhGH-Lys-azido-140 sample, and the component was inferred to be a product of reduction of the azide structure (-N₃) at the terminal of Lys-azido into (-NH₂), which indicated that the azide group was unstable when the commonly used unnatural amino acid Lys-azido was inserted into a protein, and easily reduced to form a reduction product, thus losing coupling activity. As shown in FIG. 10B, the rhGH-NBOK-140 sample did not have a high proportion of impurities, which indicated that the recombinant protein product prepared by NBOK was more stable than that prepared by Lys-azido.

(2) Comparison of coupling efficiency



As shown in the above reaction route, 30 KD BCN-PEG (home-made according to the Chinese patent CN112279906A) was coupled with rhGH-Lys-azido-140 in a site-specific manner by Click reaction (wherein the direction from R₁ to R₂ was the direction from the N-terminal to the C-terminal of the amino acid sequence). The target protein obtained above was adjusted to be 0.5 mg/ml with a PBS buffer at pH 7.0, BCN-PEG powder was added into the rhGH-Lys-azido solution according to 1:15 and 1:25 (molar ratios of recombinant protein to 30 KD BCN-PEG) respectively, and fully shaken for dissolution to obtain a clear and transparent solution, and then the reaction solution was sealed, and shaken to react in a constant temperature shaker (25°C, 70

rpm). Sampling was carried out and reaction results were detected by SDS-PAGE at regular intervals, and the reaction was stopped 72 hours later, with a conversion rate of about 50% to 70%. Reaction results were shown in FIG. 11A.

5 With reference to Example 8, the 30 KD aminoxy PEG was coupled with the rhGH-NBOK-140 in a site-specific manner by the oximation reaction. Sampling was carried out and reaction results were detected by SDS-PAGE at regular intervals, and the reaction was stopped 48 hours later, with a conversion rate of about 90%. Reaction results were shown in FIG. 11B.

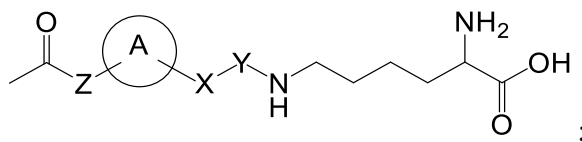
10 Combined with molecular weight detection results of high-resolution mass spectrometry, it could be judged that the reduction of Lys-azido led to the incapability of coupling the product with BCN-PEG, thus reducing the coupling rate. However, when the recombinant protein containing the unnatural amino acid of the present invention was coupled with PEG, the conversion rate was obviously higher than that of the recombinant protein containing the Lys-azido, and the reaction time was also obviously shortened, thus obviously improving reaction
15 efficiency.

Unless otherwise specified, the terms used in the present invention have the meanings commonly understood by those skilled in the art.

The described embodiments of the present invention are for exemplary purposes only, and are not intended to limit the scope of protection of the present invention. Those skilled in the art
20 can make various other substitutions, changes and improvements within the scope of the present invention. Therefore, the present invention is not limited to the above embodiments, but only limited by the claims.

Claims

1. An unnatural amino acid, wherein the unnatural amino acid is a compound with a structure as shown in formula (I) or an enantiomer thereof,



5

wherein X and Z each independently represent substituted or unsubstituted C0-C20 linear or branched alkylene, one or more -CH₂- are optionally substituted by one or more of -O-, -S-, -NH-, -C(O)- and -S(O)-, Y represents -C(O)-, -S(O)- or -CH₂-, and A represents substituted or unsubstituted C6-C20 aryl; and

when X, Z and A each independently represent a substituent, and the substituent is selected from one or more of hydroxyl, sulfhydryl, halogen, nitro, cyano, alkyl, alkenyl, alkynyl, alkoxy, acyl, acylamino, carboxyl, ester, amino, sulfonyl, sulfinyl, cycloalkyl, heterocyclyl, aryl and heteroaryl.

15

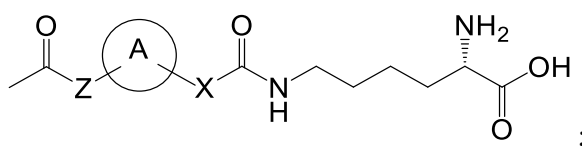
2. The unnatural amino acid according to claim 1, wherein the substituent is selected from one or more of hydroxyl, sulfhydryl, halogen, nitro, cyano, C1-C6 alkyl, C1-C6 alkoxy, acyl, acylamino, carboxyl, ester, amino, sulfonyl, sulfinyl, C3-C8 cycloalkyl, C3-C8 heterocyclyl, C6-C20 aryl and C4-C10 heteroaryl;

preferably, X and Z each independently represent C0-C10 linear or branched alkylene, and preferably represent C0-C6 linear alkylene, and one or more -CH₂- are optionally substituted by one or more of -O-, -S- and -NH-; and more preferably, X and Z are not C0 alkylene simultaneously; and/or

A represents substituted or unsubstituted C6-C10 aryl, and more preferably, A represents substituted or unsubstituted phenyl or naphthyl.

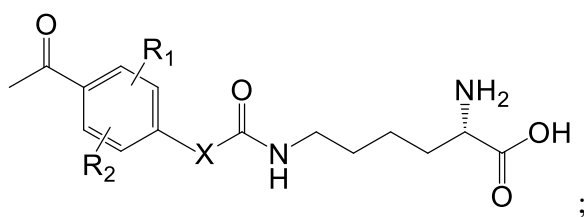
25

3. The unnatural amino acid according to claim 1 or 2, wherein the unnatural amino acid is a compound with a structure as shown in formula (I-1),



30

wherein X, Y and A are each independently as defined in claim 1 or 2;
preferably, the unnatural amino acid is a compound with a structure as shown in formula (I-2),

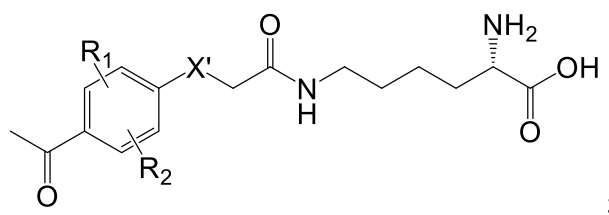


I-2

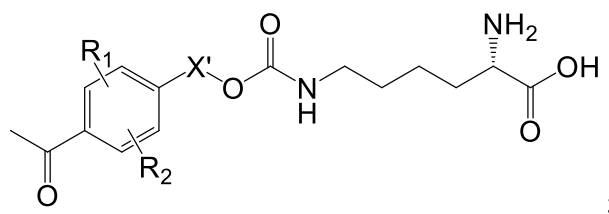
wherein X is defined according to claim 1 or 2; and

R₁ and R₂ each independently represent hydrogen, hydroxyl, sulfhydryl, halogen, nitro, cyano, C1-C6 alkyl, C1-C6 alkoxy, acyl, acylamino, carboxyl, ester, amino, sulfonyl, sulfinyl, C3-C8 cycloalkyl, C3-C8 heterocyclyl, C6-C20 aryl or C4-C10 heteroaryl.

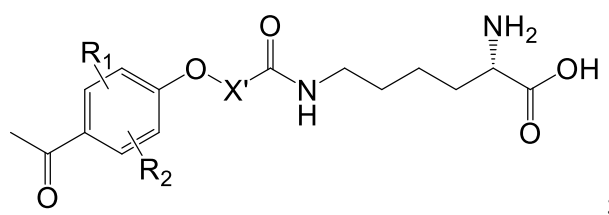
4. The unnatural amino acid according to claim 3, wherein the unnatural amino acid is a compound with a structure as shown in formula (I-3), formula (I-4), formula (I-5) or formula (I-6),



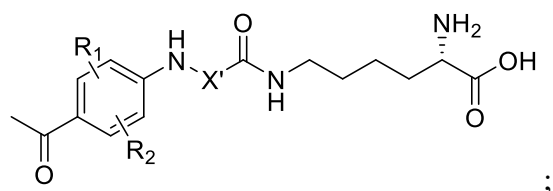
I-3



I-4



I-5



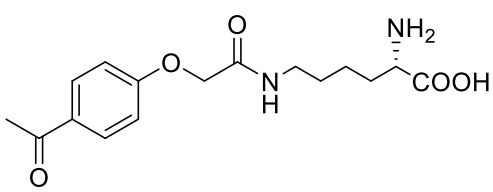
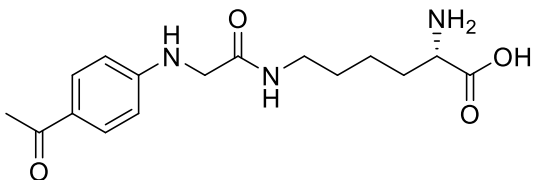
I-6

wherein X' represents C0-C6 linear alkylene, and preferably represents C0-C4 linear
 5 alkylene, and one or more -CH₂- are optionally substituted by -O- and/or -NH-; and
 R₁ and R₂ are each independently as defined in claim 3.

5. The unnatural amino acid according to any one of claims 1-4, wherein the unnatural
 amino acid is a compound with one of the structures as follows,

10

(1)	
(2)	
(3)	
(4)	
(5)	

(6)	 or
(7)	

6. Use of the unnatural amino acid according to any one of claims 1-5 in preparation of a recombinant protein or a recombinant protein conjugate; preferably, the recombinant protein is a recombinant human growth hormone; or, preferably, the recombinant protein conjugate is a recombinant human growth hormone-polyethylene glycol conjugate.

7. A recombinant protein, wherein at least one site in the amino acid sequence of the recombinant protein is the unnatural amino acid according to any one of claims 1-5.

8. A recombinant protein conjugate, wherein the recombinant protein conjugate is formed by forming an oxime bond with a terminal carbonyl of the unnatural amino acid in the recombinant protein according to claim 7 and a coupling moiety containing an "NH₂-O-" terminal group.

9. The recombinant protein conjugate according to claim 8, wherein the recombinant protein is a recombinant human growth hormone, the coupling moiety containing the "NH₂-O-" terminal group is polyethylene glycol containing an "NH₂-O-" terminal group, and is preferably polyethylene glycol with a molecular weight of 10 KD to 100 KD; more preferably, an amino acid sequence of the recombinant human growth hormone is as shown in SEQ ID NO: 1; further preferably, a corresponding 107th site in the amino acid sequence of SEQ ID NO: 1 contains the unnatural amino acid according to any one of claims 1-5.

10. Use of the recombinant protein conjugate according to claim 9 in preparation of a drug for treating growth and development disorders caused by endogenous growth hormone secretion insufficiency, growth and development disorders caused by Turner syndrome or adult growth hormone deficiency.