Modification of N-Terminal α-Amino Groups of Peptides and Proteins Using Ketenes

Anna On-Yee Chan,[†] Chi-Ming Ho,[†] Hiu-Chi Chong,[‡] Yun-Chung Leung,[‡] Jie-Sheng Huang,[†] Man-Kin Wong,*,[‡] and Chi-Ming Che*,[†]

[†]Department of Chemistry, State Key Laboratory of Synthetic Chemistry, and Open Laboratory of Chemical Biology of the Institute of Molecular Technology for Drug Discovery and Synthesis, The University of Hong Kong, Pokfulam Road, Hong Kong, China

[‡]Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

ⓢ *Supporting Information*

ABSTRACT: A method of highly selective N-terminal modification of proteins as well as peptides by an isolated ketene was developed. Modification of a library of unprotected peptides XSKFR (X varies over 20 natural amino acids) by an alkyne-functionalized ketene (**1**) at room temperature at pH 6.3 resulted in excellent N-terminal selectivity (modified α-amino group/modified ε-amino group = >99:1) for 13 out of the 20 peptides and moderate-to-high N-terminal selectivity (4:1 to 48:1) for 6 of the 7 remaining peptides. Using an alkyne-functionalized *N*-hydroxysuccinimide (NHS) ester (**2**) instead of **1**, the modification of peptides XSKFR gave internal lysine-modified peptides for 5 out of the 20 peptides and moderate-to-low N-terminal selectivity (5:1 to 1:4) for 13 out of the 20 peptides. Proteins including insulin, lysozyme, RNaseA, and a therapeutic protein BCArg were selectively N-terminally modified at room temperature using ketene **1**, in contrast to the formation of significant or major amounts of di-, tri-, or tetra-modified proteins in the modification by NHS ester **2**. The **1**-modified proteins were further functionalized by a dansyl azide compound through click chemistry without the need for prior treatment.
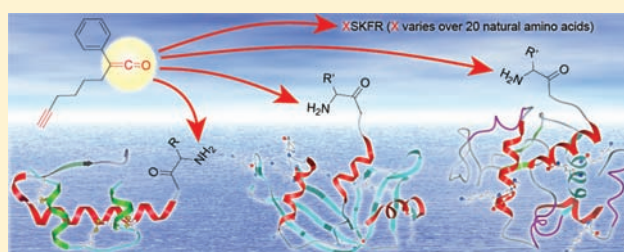
INTRODUCTION

Covalent incorporation of small molecule(s) of unique function(s) into proteins is of paramount importance in chemical biology.[1,2] These sophisticated chemical reactions for protein modification need to be conducted in aqueous media at low concentrations (typically <100 μM of protein) under mild conditions and with high selectivity. Well-documented strategies include the following: (i) direct modification of natural amino acid residue(s) such as lysine,[3a−e] cysteine,[3f−k,4b,c,e,f] tryptophan,[3l,m,4a,d] and serine,[3n,4a] including those at the N-terminus,[4] in preintroduced sequence,[3f,g,i,n] or in paired peptide chains;[5] (ii) installation of unnatural functional group(s)[6,7] at the natural amino acid residue(s) including the N-terminal residue,[7] followed by reactions such as condensation,[6a−c,7] Staudinger ligation,[6d−g] olefin metathesis,[6h] and 1,3-dipolar cycloaddition.[6i−n] Protein modification based on these strategies usually relies on the presence of one or few particular types of amino acid residues. Notably, the amino acid residue to be modified can vary over a range of amino acids and yet can retain high site-selectivity in peptide modification at N-terminal residue by a biomimetic transamination with subsequent oxime formation[7] and in peptide modification at the side chain of an internal amino acid residue by a molecular recognition combined with metal-catalyzed carbenoid transfer reaction.[5] The two methods feature preremoval of the N-terminal α-

amino group to generate an α-keto group[7] or preformation of peptide chain pair in close proximity.[5] One of the challenges is to develop a relatively simple and general method for the incorporation of a functional molecule into a natural amino acid residue of peptides or proteins with high site-selectivity.

In this endeavor, a potentially useful approach is through acylation of the N-terminal α-amino group, considering its considerably lower pK_a value (∼8) than that of the ε-amino group of lysine (pK_a value ∼10).[1] In the literature, trans-amination methods are applicable for selectively removing the α-amino group of any N-terminal residue to generate an α-keto group,[2k,8,9] using pyridoxal at 100 °C[8] or using glyoxylate in the presence of divalent metal ion and base at room temperature.[9] The biomimetic transamination under mild conditions with subsequent oxime formation[7] can efficiently modify various N-terminal residues except histidine, tryptophan, lysine, proline, cysteine, serine, and glutamine. Common reagents for amine acylation, such as *N*-hydroxysuccinimide (NHS) esters, usually afford heterogeneous bioconjugates due to facile modification of −OH group(s) or internal lysine residue(s),[10,11] although at pH 6.5 a moderate N-terminal selectivity of 3:1 (ratio of α-amino- to ε-amino-modified peptide) has been reported for

modification of peptide IQKVAGTWYSLA with NHS−LC−biotin.[11b,12] It thus would be of interest to seek an amine acylation reagent suitable for modification of peptides and proteins with high N-terminal selectivity.

In our efforts to develop new bioconjugation reactions,[13,14] we have reported a "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H$_2$O$_2$" protocol[14] [H$_2$(2,6-Cl$_2$TPP) = *meso*-tetrakis(2,6-dichlorophenyl)porphyrin] for N-terminal α-amino group ligation of peptides through oxidative amide bond formation.[15,16] Using this protocol, acylation of the N-terminal amino group of six peptides was achieved with their internal lysines residues remaining intact;[14] however, oxidation at cysteine and methionine residues was found. Such oxidation may alter the structure of proteins, thus hindering the application of the protocol in protein modification. Mechanistic studies suggested that ketenes (RR′C=C=O) are the key intermediates accounting for the selective N-terminal modification of peptides by the "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H$_2$O$_2$" protocol.[14] A ketene in situ generated from the photolysis of diazo compound PhC(O)CHN$_2$ could selectively acylate the N-terminal amino group of three peptides.[14] These findings prompted us to examine the possibility of selective N-terminal amino group acylation of peptides and proteins using an isolated ketene.

Ketenes have long been employed as versatile synthetic intermediates in organic synthesis.[17,18] However, reports on their application in protein modification are sparse. H$_2$C=C=O, generated by thermal decomposition of acetone vapor, was applied for extensive acetylation of amino groups of pepsin, insulin, and ovalbumin.[19] Nonspecific labeling of proteins at lysine or histidine residue(s) using ketenes generated in situ from photochemical decomposition of diazo compounds was also reported.[20] The nonspecific labeling might be due to the generation of reactive carbene intermediates in the photochemical reactions. A small amount of specific labeling of lysozyme upon photolysis of diazo N-acetylglucosamine (NAG) was observed, but the precise site of modification was not well characterized.[20a] To the best of our knowledge, highly selective N-terminal modification of proteins using ketene compounds has not been reported.

In the present paper, an alkyne-functionalized ketene (1) was synthesized and used for selective N-terminal modification of peptides and proteins (Scheme 1). The proteins modified

Scheme 1

peptide or protein

N-terminal modified peptide or protein

include insulin, lysozyme, RNaseA, and a therapeutic protein, BCArg. By utilizing the alkyne group in 1, the 1-incorporated proteins were further functionalized in situ via a known type of

copper-catalyzed click reaction with an organic azide[21] (Scheme 1). Remarkably, in contrast to an alkyne-functionalized NHS ester (2), ketene 1 exhibited good-to-excellent N-terminal selectivity in modifying almost all of a library of unprotected peptides XSKFR with X varying over 20 natural amino acids.

■ EXPERIMENTAL SECTION

Preparation of Ketene 1. Lithium bis(trimethylsilyl)amide (1 M in hexane, 20 mL) was added dropwise to a solution of phenylacetic acid methyl ester (2.73 g, 18.2 mmol) in dried tetrahydrofuran (40 mL) at −78 °C. After 1 h, the reaction mixture was warmed to 0 °C, and 6-iodo-hex-1-yne (4.16 g, 20 mmol) in dried tetrahydrofuran (5 mL) was added dropwise to the solution. After stirring at 0 °C for 1.5 h, the reaction mixture was quenched with water, washed with a saturated ammonium chloride solution, and extracted with ether twice. The organic layers were combined and dried over anhydrous magnesium sulfate to give 4.17 g (76%) of intermediate **A**. Colorless oil, analytical TLC (silica gel 60) (10% EtOAc in *n*-hexane), R$_f$ = 0.33. ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.34−7.25 (m, 5H), 3.65 (s, 3H), 3.54 (t, J = 7.5 Hz, 1H), 2.16 (dt, J = 7.0, 2.5 Hz, 2H), 2.10−2.07 (m, 1H), 1.92 (t, J = 2.5 Hz, 1H), 1.80−1.75 (m, 1H), 1.57−1.49 (m, 2H), 1.42−1.32 (m, 2H). ^{13}C NMR (100.62 MHz, CDCl$_3$): δ 174.7, 139.3, 128.9, 128.2, 127.5, 84.5, 68.6, 52.2, 51.7, 33.2, 28.4, 26.9, 18.5. IR (KBr, neat, cm^{-1}): 2117, 1736 cm^{-1}. EIMS: *m/z* 230 (M$^+$). HRMS (EI) for C$_{15}$H$_{18}$O$_2$: calcd 230.1307, found 230.1304. A solution of intermediate **A** (4.17 g, 18.1 mmol) in methanol (100 mL) and H$_2$O (2 mL) was treated with potassium hydroxide pellets (1.5 g, 27 mmol) and refluxed overnight. The reaction mixture was evaporated to dryness using a rotary evaporator. Water was added to the reaction mixture, and the reaction mixture was washed with diethyl ether. The aqueous layer was collected and acidified with hydrochloric acid and then extracted with ether twice. The organic layers were combined, dried over anhydrous magnesium sulfate, and evaporated to give intermediate **B** (3.32 g, 87%) as a colorless oil: analytical TLC (silica gel 60) (20% EtOAc in *n*-hexane), R$_f$ = 0.25. ^{1}H NMR (300 MHz, CDCl$_3$): δ 7.31−7.24 (m, 5H), 3.54 (t, J = 7.5 Hz, 1H), 2.14 (dt, J = 7.0, 2.5 Hz, 2H), 2.09−2.04 (m, 1H), 1.90 (t, J = 2.5 Hz, 1H), 1.83−1.74 (m, 1H), 1.56−1.48 (m, 2H), 1.43−1.33 (m, 2H). ^{13}C NMR (75.48 MHz, CDCl$_3$): δ 180.3, 138.3, 128.7, 128.0, 127.5, 84.2, 68.4, 51.4, 32.4, 28.1, 26.5, 18.1. IR (KBr, neat, cm^{-1}): 2117, 1703 cm^{-1}. EIMS: *m/z* 216 (M$^+$). HRMS (EI) for C$_{14}$H$_{16}$O$_2$: calcd 216.1150, found 216.1148. Oxalyl chloride (2 M in dichloromethane, 5 mL) was added to a solution of intermediate **B** (1.08 g, 5 mmol) in dried dichloromethane (5 mL) at rt, and the reaction mixture was stirred for 2 h. The solvent was distilled under nitrogen atmosphere to give a light yellow oil. The light yellow oil was dissolved in dried tetrahydrofuran (10 mL), and dried triethylamine (6 mL, 20 mmol) was added dropwise to the solution at 0 °C. The resulting mixture was stirred at 0 °C for 2 h. The salt formed was filtered under nitrogen atmosphere, and the filtrate was distilled at 110 °C (1 mmHg) to give 1 (0.26 g, 35%) as a bright yellow oil. ^{1}H NMR (400 MHz, CDCl$_3$): δ 7.33−7.25 (m 3H), 7.07−7.00 (m, 2H), 2.41 (t, J = 7.0 Hz, 2H), 2.22 (dt, J = 7.0, 2.5 Hz, 2H), 1.94 (t, J = 2.5 Hz, 1H), 1.72−1.60 (m, 4H). ^{13}C NMR (100.62 MHz, CDCl$_3$): δ 204.6, 132.5, 129.0, 124.3, 124.1, 84.0, 68.6, 28.0, 27.1, 23.1, 18.1. IR (KBr, neat, cm^{-1}): 2116, 2096 cm^{-1}. ESI-MS for C$_{14}$H$_{14}$O: calcd 198.1045, found 198.1264.

Modification of Peptides XSKFR Using 1. In a 1.0 mL eppendorf tube, XSKFR solution in H$_2$O (1 μmol/mL, 10 μL), 1 (10 equiv, 10 μL of a 10 μmol/mL stock solution of 1 in dried THF), and phosphate buffer (pH 6.3, 80 μL) were mixed. The reaction mixture was kept at rt for 15 min. The conversions of the 20 XSKFR peptides were determined from TIC (total ion count) of LC−MS/MS (liquid chromatography−tandem mass spectrometry) analysis of the reaction mixtures. Using EIC (extracted ion chromatogram) analysis, we determined the N-terminal selectivity.

Scale-up Modification of Peptide YTSSSKNVVR Using 1. A solution of peptide YTSSSKNVVR (4 mg) and 1 (8 mg, 10 equiv) in pH 6.3 phosphate buffer (40 mL) was prepared and placed into a 50 mL centrifugal tube. The reaction mixture was kept at rt overnight.
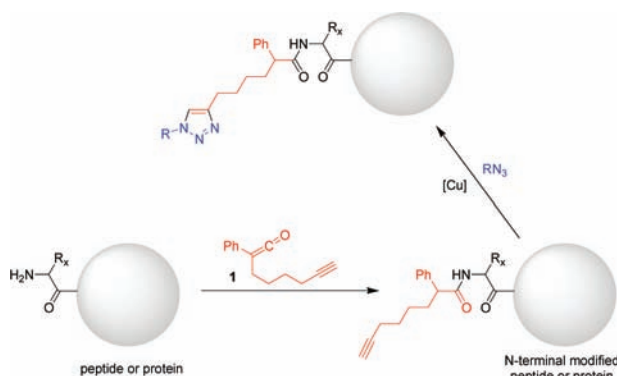
2590

dx.doi.org/10.1021/ja208009r | *J. Am. Chem. Soc.* 2012, 134, 2589−2598

The aqueous phase was freeze-dried, and the residue was purified by preparative reverse-phase HPLC equipped with a C_{18} column using $CH_3CN/H_2O/TFA$ as the solvent system. The **1**-modified YTSSSKNVVR was isolated in 40% yield (3.2 mg). The sequence of the **1**-modified peptide was confirmed by LC−MS/MS analyses.

Modification of Insulin Using 1. In a 1.0 mL eppendorf tube, insulin solution in H_2O (1 μmol/mL, 10 μL), **1** (10 equiv, 10 μL of a 10 μmol/mL stock solution of **1** in dried THF), and phosphate buffer (pH 6.3, 80 μL) were mixed at rt for 15 min. The conversion was determined from TIC of LC−MS/MS analysis of the reaction mixture. To reduce the disulfide bond between chain A and chain B, dithiothreitol (10 equiv, 1 μL of a 10 μmol/100 μL stock solution of dithiothreitol in H_2O) was added to the **1**-modified insulin solution (100 μL). The reaction mixture was incubated at 56 °C for 3 h. The N-terminal modification of chain B of insulin by **1** was confirmed by LC−MS/MS analysis.

Modification of Lysozyme Using 1. In a 1.0 mL eppendorf tube, lysozyme solution in H_2O (1 μmol/mL, 10 μL) was mixed with phosphate buffer (pH 9.2, 84 μL) at 37 °C. Ketene **1** (6 equiv, 6 μL of a 10 μmol/mL stock solution of **1** in dried THF) was added in three equal portions at 15 min interval. The mixture was kept at 37 °C overnight. The conversion was determined from TIC of LC−MS/MS analysis of the reaction mixture.

Trypsin Digestion of 1-Modified Lysozyme. In a 1.0 mL eppendorf tube, **1**-modified lysozyme mixture (100 μL) was mixed with 1:1 butanol/H_2O (100 μL) at 65 °C for 15 min. The reaction mixture was diluted 4-fold by ammonium bicarbonate solution (50 mM, 600 μL). Trypsin solution [1 mg/mL, 2.86 μL; ratio of trypsin to lysozyme = 1:50 (w/w)] was added to the eppendorf tube at 0 °C. The reaction mixture was incubated at 37 °C overnight, and the trypsin digested mixture was analyzed using LC−MS/MS analysis.

Modification of RNaseA Using 1. This was performed using RNaseA and **1** (15 equiv) by a procedure similar to that for the modification of lysozyme (see the Supporting Information).

Trypsin Digestion of 1-Modified RNaseA. This was performed using **1**-modified RNaseA mixture and trypsin solution (1 mg/mL, 2.74 μL) by a procedure similar to that for the trypsin digestion of **1**-modified lysozyme (see the Supporting Information).

Modification of BCArg Using 1. In a 1.0 mL eppendorf tube, **1** (500 equiv, 13 μL of a 100 μmol/10 μL stock solution of **1** in dried THF) was added to BCArg solution in H_2O (0.5 μmol/mL, 500 μL). The reaction mixture was kept at 37 °C overnight. The modified protein was subjected to ESI−MS analysis.

Trypsin Digestion of 1-Modified BCArg. In a 1.0 mL eppendorf tube, **1**-modified protein mixture was mixed with trypsin [BCArg/trypsin 1:20 (w/w)] and incubated at 56 °C. After incubation for 2 h, trypsin was added to the reaction mixture for replenishment, and the reaction mixture was incubated at 56 °C for 1 h. The insoluble protein fraction was collected by centrifugation at 10 000 × g for 3 min. The pellet was solubilized in 0.1% aqueous TFA with 30% CH_3CN. The reaction mixture was concentrated using C_{18} ZipTip and applied to MALDI−TOF (matrix-assisted laser desorption/ionization−time-of-flight) MS/MS analysis.

Azide−Alkyne Cycloaddition of 1-Modified Proteins (Insulin, Lysozyme, or RNaseA). In a 1.0 mL eppendorf, the **1**-modified protein mixture (10 μL), a solution of N-(3-azidopropyl)-5-(dimethylamino)-1-naphthalenesulfonamide (**3**)[22] in DMSO (10 μmol/mL, 1 μL), a solution of tris(2-carboxyethyl)phosphine hydrochloride (TCEP) in H_2O (10 μmol/mL, 10 μL), a solution of tris(benzyltriazolylmethyl)amine (TBTA) in 1:4 DMSO-t-butanol (13.5 μmol/mL, 10 μL), a $CuSO_4$ solution in H_2O (15.7 μmol/mL, 10 μL), and PBS buffer (pH 8.4, 59 μL) were mixed. The reaction mixture was kept at rt for 2 h. The dansyl-tagged proteins were characterized by LC−MS analysis.

■ RESULTS

Before starting to synthesize ketene **1**, we examined the applicability of the "[Mn(2,6-Cl_2TPP)Cl]/alkyne/H_2O_2" protocol, which would generate ketene RR′C=C=O in situ, for

modifying the N-terminus of a library of 20 unprotected peptides, XSKFR [Scheme 2 and Table 1, X = residues of Asp

Scheme 2
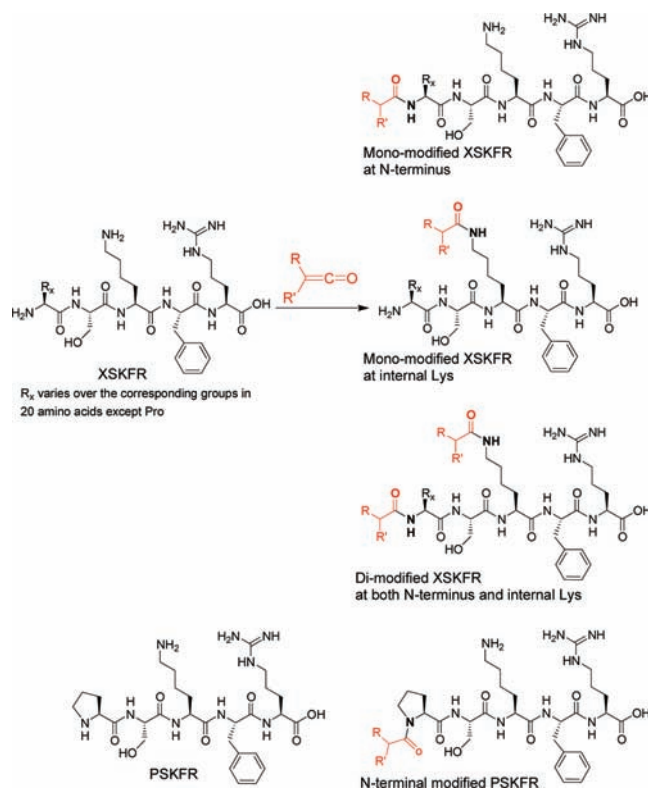
(D), Glu (E), Gly (G), Ala (A), Pro (P), Ser (S), Thr (T), Tyr (Y), Trp (W), Met (M), Gln (Q), Cys (C), Lys (K), His (H), Asn (N), Val (V), Phe (F), Leu (L), Ile (I), or Arg (R)]. Previous studies[14] demonstrated modification of N-terminal residues glycine, serine, threonine, tyrosine, and histidine of six other peptides using this protocol. The modification of XSKFR (100 μM) was conducted with the alkyne phenylacetylene (2 equiv, Table 1; 10 equiv, Table S1 in the Supporting Information). The conversions are in the range of 8−100% (Table 1), as revealed by LC-MS/MS analysis. Mono-modified peptides (Scheme 2, R = Ph, R′ = H) were obtained for all the 20 peptides (for KSKFR and RSKFR, 3% and 2%, respectively, of di-modified products depicted in Scheme 2 were also found). Excellent N-terminal selectivity (>99:1, determined by LC−MS/MS analysis) was achieved for 14 of the 20 peptides XSKFR[23] (Table 1, entries 1−14), and moderate-to-high N-terminal selectivities (3:1 to 15:1) were obtained for the others (Table 1, entries 15−20). Oxidation at residues tryptophan, methionine, and cysteine was observed (Table 1, entries 9, 10, 12).

Synthesis of an Alkyne-Functionalized Ketene 1 and NHS Ester 2. In view of the results described above, we envisaged that isolated ketenes should be promising reagents for N-terminal modification of peptides and proteins. Ketene **1** was synthesized. The synthetic route (Scheme 3) includes alkylation of phenylacetic acid methyl ester with 6-iodo-hex-1-yne to give alkyne-functionalized ester **A** followed by its hydrolysis to afford carboxylic acid **B**. Reaction of **B** with oxalyl chloride followed by treatment with triethylamine afforded **1**, which could be isolated in reasonably pure form in 35% yield by

Table 1. Modification of Peptides XSKFR Using Ketene 1 Compared with That Using NHS Ester 2 and the "[Mn(2,6-Cl$_2$TPP)Cl]/Phenylacetylene/H$_2$O$_2$" Method (I)[a]

entry	peptide sequence	conversion (%)[b]			N-terminal selectivity of mono-modified XSKFR[c]		
		I	1	2	I	1	2
1	DSKFR	43	33	78[d]	>99:1	>99:1	1:1
2	ESKFR	100	62	93	>99:1	>99:1	2:1
3	GSKFR	100	33	81	>99:1	>99:1	5:1
4	ASKFR	14	9	84	>99:1	>99:1	<1:99
5	PSKFR	39	44	64	>99:1	>99:1	<1:99
6	SSKFR	24	44	81	>99:1	>99:1	1:1
7	TSKFR	23	37	72	>99:1	>99:1	<1:99
8	YSKFR	100	47	73	>99:1	>99:1	<1:99
9	WSKFR	17[e]	19	21	>99:1	>99:1	>99:1
10	MSKFR	8[f]	29	58	>99:1	>99:1	30:1
11	QSKFR	13	94	100	>99:1	>99:1	4:1
12	CSKFR	74[g]	10	25	>99:1	8:1	1:1
13	KSKFR	34[h]	48[i]	55[j]	>99:1	6:1	1:1
14	HSKFR	35	52[k]	79[l]	>99:1	1:1	1:4
15	NSKFR	19	17	48	4:1	>99:1	2:1
16	VSKFR	19	23	72	13:1	>99:1	<1:99
17	FSKFR	15	29	41	15:1	29:1	1:1
18	LSKFR	25	23	71	8:1	29:1	2:1
19	ISKFR	37	37	67	5:1	48:1	2:1
20	RSKFR	11[m]	28	19[n]	3:1	4:1	1:1.5

[a]Conditions: XSKFR (0.01 μmol) and 1 or 2 (0.1 μmol) in phosphate buffer of pH 6.3 (100 μL), 15 min, rt, or XSKFR (0.01 μmol), phenylacetylene (0.02 μmol, 2 equiv), [Mn(2,6-Cl$_2$TPP)Cl] (2 nmol), H$_2$O$_2$ (1.18 μmol), NaHCO$_3$ (0.1 μmol) in CH$_3$CN−H$_2$O (3:2 v/v, 100 μL), 6 h, rt. [b]Determined by total ion count (TIC) of LC−MS analysis. [c]Ratio of N-terminal α-amino group modified peptide to lysine ε-amino group modified peptide; this ratio was determined by extracted ion chromatogram (EIC) of LC−MS analysis. [d]Di-modified DSKFR accounted for 10%. [e]The indole group of tryptophan was oxidized and presumably gave 1,3-dihydro-indol-2-one. [f]Methionine was oxidized to its corresponding sulfoxide. [g]Cysteine was oxidized to give disulfide-bonded peptide. [h]Di-modified KSKFR accounted for 3%. [i]Di-modified KSKFR accounted for 8%. [j]Di-modified KSKFR accounted for 4%. [k]Di-modified HSKFR accounted for 19%. [l]Di-modified HSKFR accounted for 8%. [m]Di-modified RSKFR accounted for 2%. [n]Di-modified RSKFR accounted for 7%.

Scheme 3
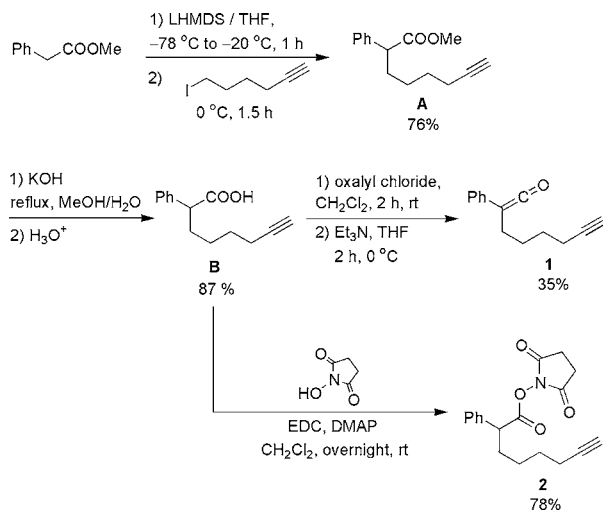
vacuum distillation (110 °C, 1 mmHg). The isolated ketene 1 can be stored at −20 °C for 1 week without significant decomposition. For comparing the N-terminal selectivity of 1 with that of a common type of reagent used for amino group modification of peptides and proteins, NHS ester 2 was synthesized in 78% yield by the coupling reaction of **B** with *N*-hydroxysuccinimide (Scheme 3).

Modification of Peptides XSKFR Using Ketene 1. The N-terminal selectivity of 1 was examined using the foregoing library of 20 peptides XSKFR. Modification of XSKFR (100 μM) with 1 (1 mM; 10 equiv) was conducted in 100 μL of pH 6.3 phosphate buffer/THF (9:1) for 15 min at room temperature. The results are listed in Table 1.

As depicted in Table 1, the conversions range from 9% (ASKFR) to 94% (QSKFR). Mono-modified peptides (Scheme 2, R = Ph, R′ = (CH$_2$)$_4$C≡CH) were obtained for all 20 peptides, of which KSKFR and HSKFR also gave 8% and 19%, respectively, of di-modified peptides (Scheme 2).

Peptides with N-terminal Asp, Glu, Gly, Ala, Pro, Ser, Thr, Tyr, Trp, Met, Gln, Asn, or Val gave excellent N-terminal selectivity (>99:1; Table 1, entries 1−11, 15, 16). Moderate-to-high N-terminal selectivities (4:1 to 48:1) were obtained for the N-terminal Cys, Lys, Phe, Leu, Ile, and Arg peptides (entries 12, 13, 17−20). However, a low N-terminal selectivity of 1:1 was observed for HSKFR having N-terminal histidine residue (Table 1, entry 14).

We have also studied the effect of the amount of 1 on the modification of peptide LSKL (Table S4 in the Supporting Information). With 10 equiv of 1, the ratio of unreacted LSKL to modified LSKL was 48:52 with 28:1 N-terminal selectivity, and no di-modified LSKL was observed. Upon increase of the amount of 1 from 10 to 50 equiv, the ratio of LSKL to mono-modified LSKL to di-modified LSKL was found to be 4:87:9. Despite the formation of di-modified LSKL, the N-terminal selectivity for mono-modified LSKL increased to 68:1. In addition, N-terminal modification of LSKL in buffers of different pH was also conducted. The N-terminal selectivity was 15:1 at pH 7.4 and 14:1 at pH 9.2 (Table S4 in the Supporting Information).

Modification of Peptides XSKFR Using NHS Ester 2. For comparison, the library of 20 peptides XSKFR were modified using NHS ester 2 (Scheme 2 and Scheme S1 in the Supporting Information) under the same conditions as those for the modification with ketene 1.

Poor N-terminal selectivity was obtained for most of the peptides when the NHS ester 2 was used (Table 1). For the N-terminal alanine, proline, threonine, tyrosine, and valine peptides, the product detected was the lysine-modified peptide, instead of the corresponding N-terminally modified peptide (Table 1, entries 4, 5, 7, 8, 16). Unexpectedly, the N-terminal selectivity for N-terminal tryptophan and methionine peptides is exceptionally high (>99:1 and 30:1, respectively; Table 1, entries 9, 10).

Modification of Cysteine-Containing Peptide STSSSCNLSK Using Ketene 1 or NHS Ester 2. In previous work,[14] exclusive N-terminal ligation of a cysteine-containing peptide STSSSCNLSK (N-terminal serine), without cysteine modification, was achieved using the "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H$_2$O$_2$" protocol. This is attributed to the oxidation of the cysteine thiol group to disulfide under oxidizing conditions. In the present work, the ligation of STSSSCNLSK (100 μM) using 1 (1 mM, 10 equiv) in 100 μL of phosphate buffer/THF (9:1) at pH 6.3 for 1 h at room temperature afforded mono-

modified STSSSCNLSK. The ratio of unmodified STSSSCNLSK to mono-modified STSSSCNLSK was 80:20. The ratio of N-terminally modified STSSSCNLSK to cysteine-modified STSSSCNLSK was about 1:1, and no lysine-modified peptide was detected. The STSSSCNLSK with thioester linkage (10 μM) resulting from cysteine modification can be hydrolyzed by hydroxylamine (500 μM, 50 equiv) in 200 μL of phosphate buffer of pH 9.2 at 37 °C overnight to afford disulfide-bonded N-terminally modified STSSSCNLSK under alkaline reaction conditions.

In contrast, no N-terminal modification was observed in the modification of STSSSCNLSK using NHS ester **2** under the same conditions as those using ketene **1**. The modification using **2** gave a 54:46 ratio of unmodified STSSSCNLSK to mono-modified STSSSCNLSK, the latter being a cysteine-modified product.

Scale-Up Modification of Peptide YTSSSKNVVR Using Ketene 1. Modification of peptides with **1** could be scaled up. An 8 mg scale one-pot modification of peptide YTSSSKNVVR with **1** gave 3.2 mg of the N-terminally modified YTSSSKNVVR (40% isolated yield, Scheme 4) after purification by preparative reversed-phase HPLC.
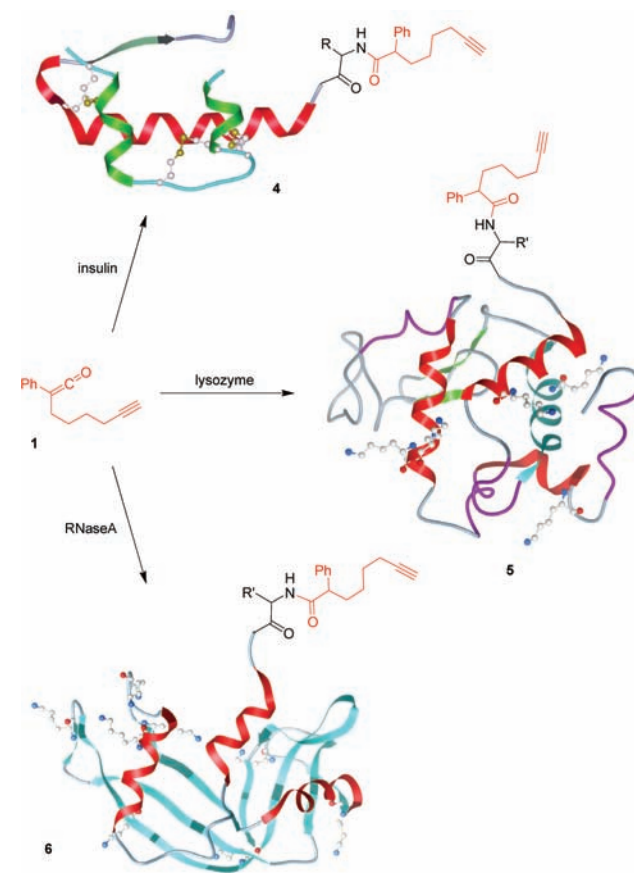
Scheme 4
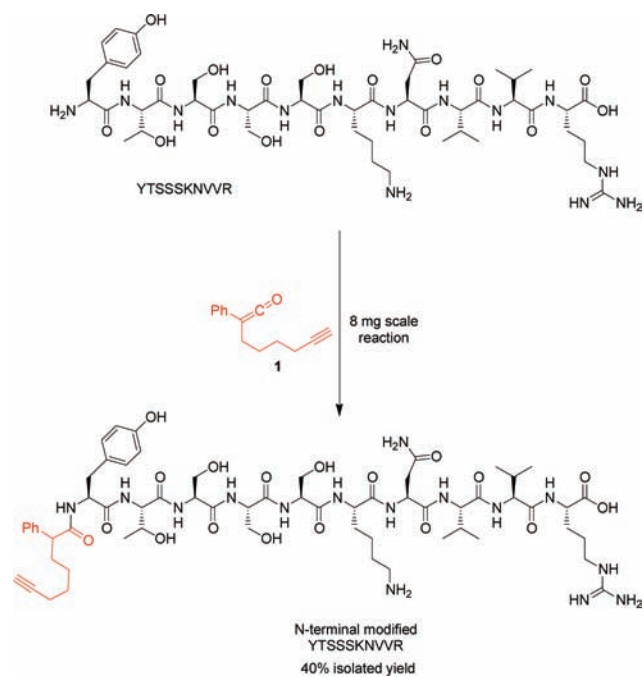
Modification of Insulin, Lysozyme, RNaseA, and BCArg Using Ketene 1. To demonstrate the applicability of ketene **1** in protein modification, we examined the reactivity of **1** toward the following proteins (for their sequences, see Chart S1 in the Supporting Information): insulin (carrying N-terminal glycine on chain A and N-terminal phenylalanine and one internal lysine on chain B), lysozyme (carrying N-terminal lysine and five internal lysines), RNaseA (carrying N-terminal lysine and nine internal lysines), and BCArg (carrying N-terminal methionine and 14 internal lysines). Strikingly, these proteins were modified by **1** selectively at their N-terminus (Scheme 5).
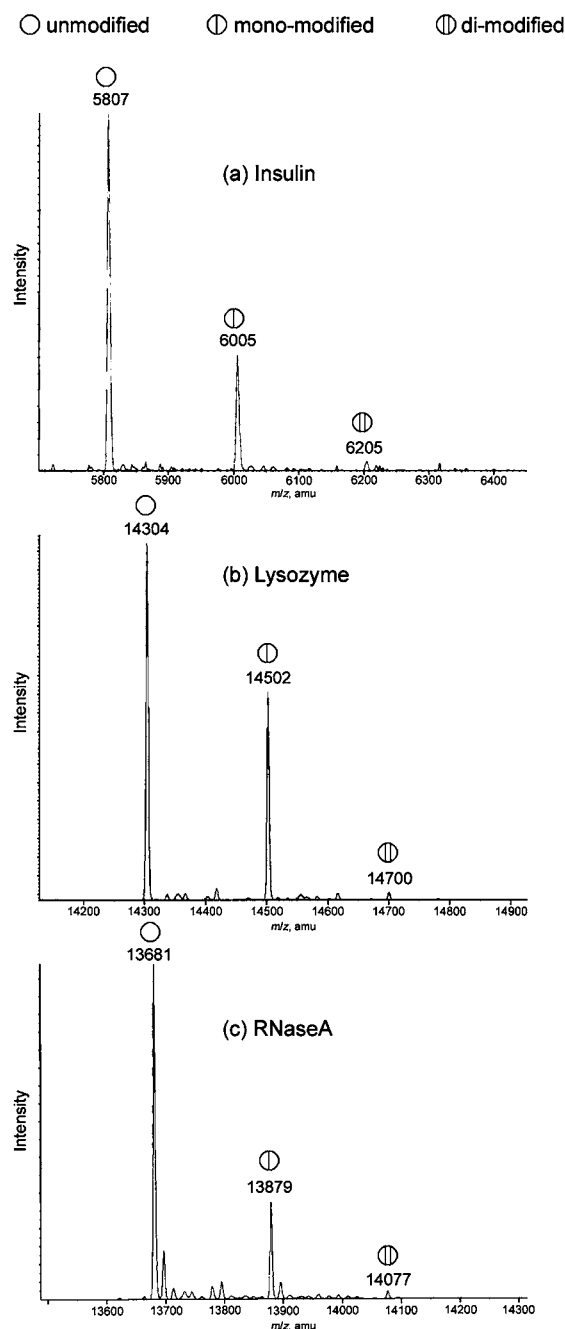
The modification of insulin (100 μM) using **1** (10 equiv) was conducted in 100 μL of phosphate buffer of pH 6.3 at room

Scheme 5

temperature for 15 min. Product analysis by LC−MS (Figure 1a) revealed a peak at 6005 Da, which was assigned to mono-modified insulin (**4**, Scheme 5) on the basis of the mass shift from that (5807 Da) of unmodified insulin (unmodified insulin/mono-modified insulin = 2:1). Only a trace amount of di-modified insulin was detected in the reaction mixture. LC−MS/MS analysis of the reaction mixture, after treatment with dithiothreitol to reduce the internal disulfide bonds between chain A and chain B of the **1**-modified insulin, revealed that the modification of insulin by **1** selectively occurred at the N-terminal phenylalanine of chain B. For comparison, modification of insulin using NHS ester **2** was conducted in phosphate buffer of pH 6.3. Although **2** also resulted in the formation of mono-modified insulin as the major product, a significant amount of di-modified and trace amount of tri-modified insulin were found by LC−MS analysis (see Table S13 in the Supporting Information).

For lysozyme and RNaseA, modification was performed in 100 μL of phosphate buffer of pH 9.2 at 37 °C overnight. The amount of **1** used is 6 equiv for lysozyme (100 μM) and 15 equiv for RNaseA (100 μM); in both cases, **1** was added in three equal portions at 15 min intervals. LC−MS analysis of the reaction mixtures showed peaks at 14502 Da (Figure 1b) and 13879 Da (Figure 1c), which were assigned to the mono-modified lysozyme (**5**, Scheme 5) and mono-modified RNaseA (**6**, Scheme 5), respectively (unmodified lysozyme/mono-modified lysozyme = 62:38; unmodified RNaseA/mono-modified RNaseA = 77:23). A trace amount of di-modified product was also detected. Upon trypsin digestion, the modification by **1** was found to selectively occur at the N-terminus, as revealed by LC−MS/MS analysis (see, for

Figure 1. Mass reconstruction spectrum in LC−MS analysis of the reaction mixture in the modification of (a) insulin, (b) lysozyme, and (c) RNaseA using **1**.

example, Figure 2a). At pH 9.2, modification of lysozyme or RNaseA using 10 equiv of ketene **1** at room temperature for 15 min produced mono-modified product only (see Tables S14 and S16 in the Supporting Information), while that using NHS ester **2** mainly gave tri- or tetra-modified product (see Tables S15 and S17 in the Supporting Information).

BCArg, carrying 14 internal lysine residues, is a therapeutic protein for cancer treatment.[24] The coupling reaction between BCArg (2.6 mM) and **1** (500 equiv) was performed in 100 μL of H_2O at 37 °C overnight. In MALDI−TOF MS analysis of the reaction mixture, the peaks at 33256 and 33455 Da are assigned to unmodified BCArg and mono-modified BCArg, respectively. By trypsin digestion, the site of modification was

found at the N-terminal methionine by MALDI−TOF MS/MS analysis (Figure 2b).

Without further purification, the crude **1**-modified proteins (insulin, lysozyme, and RNaseA, 10 μM) containing an alkyne moiety were further modified with a dansyl azide **3**[22] (100 μM in DMSO) in the presence of the triazolyl ligand TBTA (100 μM in 1:4 DMSO/*t*-butanol), tris-carboxyethylphosphine (TCEP) (1 mM), and $CuSO_4$ (1 mM) in 100 μL of PBS buffer of pH 8.4 (Scheme 6) in a one-pot reaction (see the Experimental Section). The resulting **3**-modified proteins (**7**− **9**, Scheme 6) were characterized by LC−MS analysis.

■ DISCUSSION

Peptide Modification by Ketene 1. Bioconjugation reaction using ketene conducted in aqueous media has to face competitive hydration of ketene, which affords carboxylic acid as a side product.[17] Since the hydration rate of ketene is slower than the amination rate,[17] amination of ketenes in the presence of water has previously been reported in the literature and used for amide bond formation in organic synthesis.[17]

To synthesize a relatively stable ketene with functional diversity for modification of peptides and proteins, we designed the alkyne-linked ketene **1**. The key features of **1** include the following: (1) a terminal aliphatic alkyne group is installed for further modification with different biophysical probes through bio-orthogonal click chemistry;[6k−n] (2) an aryl ring is installed for enhancing the stability of ketene for vacuum distillation. To the best of our knowledge, **1** is the first isolated alkyne-functionalized ketene.[25]

The library of 20 unprotected peptides XSKFR was chosen mainly for examining the generality of the N-terminal modification of peptides by ketenes. The serine and lysine residues in the sequence of XSKFR are for studying the chemo- and site-selectivity in the presence of both −OH group (from serine) and amino groups (from lysine and the N-terminus) (see Scheme 2). The aromatic phenylalanine residue was incorporated for assisting HPLC analysis, and the C-terminal residue, arginine, was incorporated for obtaining better signal in MS analysis.

Amination of almost all of the peptides XSKFR by **1** preferentially occurred at the N-terminal α-amino group despite the presence of nucleophilic −OH group and more basic lysine ε-amino groups. At pH 6.3, excellent N-terminal selectivity (>99:1) was obtained in 13 out of the 20 XSKFR peptides, and moderate (4:1) to high (48:1) N-terminal selectivity was obtained in the modification of the remaining 6 XSKFR peptides (Table 1). No oxidation of amino acid residues was observed, in contrast to the oxidation at the cysteine, methionine, and tryptophan side chains by the "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H_2O_2" protocol (Table 1). Except for the oxidation problem, the "[Mn(2,6-Cl$_2$TPP)Cl]/ alkyne/H_2O_2" protocol, which would generate ketenes in situ, resulted in moderate-to-excellent N-terminal selectivity for the 20 peptides XSKFR (Table 1).

No −OH modification on serine residue was observed for the modification of peptides XSKFR using **1** or the "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H_2O_2" protocol (Table 1). The absence of −OH modification might be due to the lower nucleophilicity of the −OH group than that of the α- or ε-amino groups of the N-terminus and lysine residue.

In the modification of cysteine-containing peptide STSSSNCLSK using **1**, both cysteine-modified and N-terminally modified peptides were obtained. This is presumably
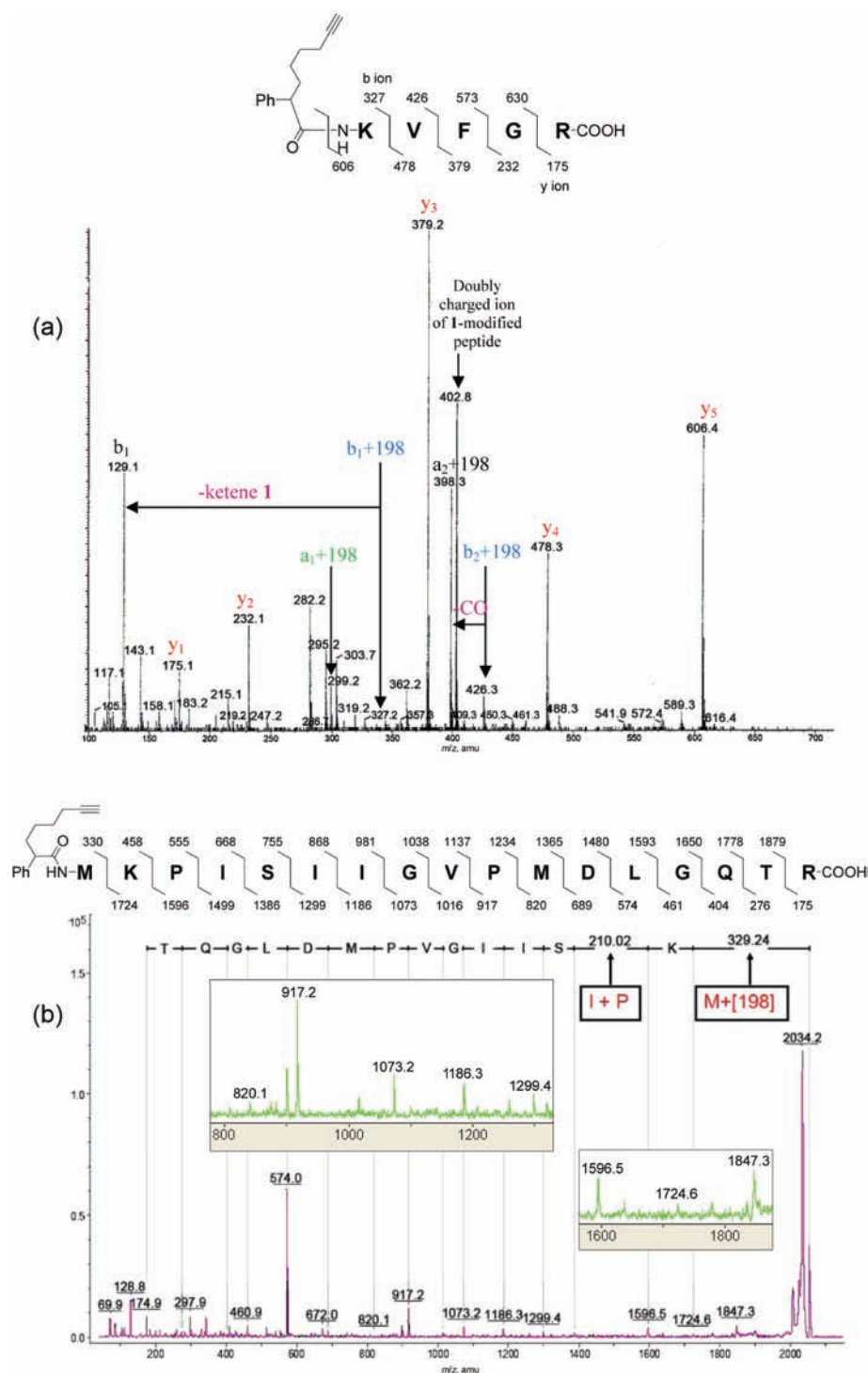
**Figure 2.** The MS/MS spectrum of the N-terminal **1**-modified fragment of (a) lysozyme and (b) BCArg.

due to the higher nucleophilicity of the cysteine thiol group, which results in the formation of thioester linkage. The thioester linkage can be reduced back to a free cysteine thiol group by treatment with hydroxylamine under basic conditions.

Comparison of Peptide Modification by Ketenes with Literature Methods. The high N-terminal selectivity of ketenes in bioconjugation can be revealed by comparison with the result of modification of peptides XSKFR using NHS ester **2**. As depicted in Table 1, **2** gave moderate-to-low N-terminal selectivity (i.e., from 5:1 to <1:99) in modifying 18 out of the 20 peptides XSKFR. The observed modification of

STSSSCNLSK by **2** at the cysteine residue but not at the N-terminus further indicates that ketene **1** is more selective toward N-terminal α-amino group than NHS ester **2**.

Several methods have been reported for selective N-terminal modification of peptides or proteins. Apart from the trans-amination methods described in the Introduction section,[7−9] other methods each suitable for modifying one or two types of N-terminal residues have been developed, such as modification of serine or threonine residues by periodate oxidation,[4a,9b] modification of tryptophan residues by Pictet−Spengler reaction,[4d] and modification of cysteine residues by native

Scheme 6
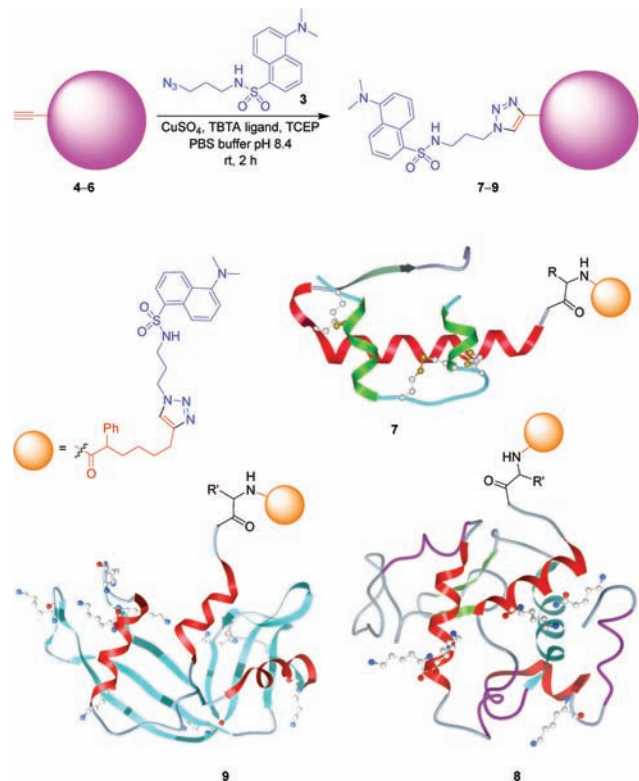
chemical ligation.[4b,c,e,f] Transamination with subsequent Pictet−Spengler reaction can selectively modify N-terminal glycine residues.[26] The N-terminal modification using ketene **1** not only exhibits impressive generality but also directly incorporates a functional molecule into the N-terminal α-amino group through a relatively simple acylation reaction, without preferentially modifying the side chain(s) of the N-terminus.

Protein Modification by Ketene 1. In the modification of peptide LSKL using **1**, the conversion of the modification reaction increased with increasing amount of **1** used (from 10 to 50 equiv). Also, increasing pH of the reaction mixture resulted in higher reaction conversion, while the N-terminal selectivity remained high (14:1 at pH 9.2). On the other hand, lowering the pH value of the reaction mixture could increase the N-terminal selectivity but could slightly decrease the reaction conversion. Increase in reaction conversion may be compromised by multisite modification; the latter could be minimized by adding the ketene in portions.

Accordingly, the reaction conversion for the modification of proteins was raised by choosing the right amount of ketene and the right pH value of the reaction mixture while attaining excellent N-terminal selectivity. Proteins including insulin, lysozyme, RNaseA, and BCArg were N-terminally modified using **1**, as confirmed by trypsin digestion and LC−MS/MS analysis.[27] Selective N-terminal modification on chain B of insulin was found in the modification of insulin using **1**. In this modification, the N-terminal glycine of chain A of insulin remained intact. One rationale for this finding is the higher solvent-accessible area of the N-terminal phenylalanine of chain B, which was calculated to be 31.19 Å^2 [13b] by GETAREA 1.4. The solvent-accessible area of N-terminal glycine of chain A was found to be 18.31 Å^2,[13b] rendering the N-terminus of chain A less exposed for modification using **1**.

Modification of proteins RNaseA and lysozyme were conducted in phosphate buffer at pH 9.2 and at 37 °C in order to have higher conversions. Although most of the N-terminal α-amino groups should have been protonated under basic conditions, excellent N-terminal selectivity was still observed in the modification of RNaseA and lysozyme using ketene **1**. No internal-lysine-modified protein was observed in either case. The modification of RNaseA and lysozyme using NHS ester **2** gave multiple site-modified proteins (up to +7 modified RNaseA was formed using NHS ester **2**, as depicted in Table S17 in the Supporting Information). These results further reveal the high selectivity of ketenes toward the N-terminal α-amino group.

The N-terminal selectivity in the modification of insulin (N-terminal phenylalanine), lysozyme (N-terminal lysine), and RNaseA (N-terminal lysine) using ketene **1** is higher than that for similar reactions with peptides carrying the same N-terminal amino acid residues. The peptides XSKFR with N-terminal lysine and phenylalanine residues gave N-terminal selectivity of 6:1 and 29:1, respectively (Table 1, entries 13 and 17). One possibility is that the tertiary structure of proteins leads to microenvironment that reduces the accessibility of ketene to the internal lysine residues, thus resulting in an increase of N-terminal selectivity.

The alkyne functional group on ketene **1** allows further modification of the **1**-modified proteins under *in vitro* or *in vivo* conditions through bio-orthogonal reactions in click chemistry,[6k−n] such as a copper-catalyzed [3 + 2] cycloaddition of azide with alkyne.[21] In this work, the copper-catalyzed [3 + 2] cycloaddition reaction of dansyl azide **3** with **1**-modified proteins was achieved without prior purification of the **1**-modified proteins.

BCArg is a protein that displays anticancer activity. The excellent N-terminal selectivity obtained in the modification of BCArg by ketene **1** demonstrates the applicability of alkyne-linked ketene in the modification of pharmaceutical proteins.[28] Therapeutic proteins are of increasing global demand. However, therapeutic proteins suffer from poor stability, low solubility, and short circulating half-life. PEGylation is the most widely used method to improve the pharmacological profiles of protein drugs.[29] The alkyne-incorporated BCArg would allow further ligation with PEG moiety through click chemistry. The high N-terminal selectivity accomplished by **1** would allow the bioconjugation of pharmaceutical proteins to occur at a precise location resulting in homogeneous pharmaceutical proteins with higher consistency in clinical performance.

There is a growing need to develop methods for modifying proteins at two distinct sites for fluorescence resonance energy transfer (FRET) studies.[30] The high selectivity of ketenes to the N-terminus of proteins opens a new way for site-selective incorporation of one functional group into proteins. The bioconjugation reaction conducted using ketenes is expected to be compatible with most of the literature-reported bioconjugation reactions.[31] Thus, the second functional group could be installed on proteins through modification of cysteine, tyrosine, or tryptophan without the need for prior treatment of the reaction mixture of the ketene-modified proteins.[32]

■ CONCLUSION

We have observed highly selective N-terminal modification of peptides and proteins by an isolated, alkyne-functionalized ketene **1**. This method can directly incorporate a functional molecule into the N-terminus of peptides XSKFR under mild

conditions (pH 6.3, room temperature), with excellent terminal selectivity for the majority of the 20 natural amino acid residues (X) and with moderate-to-high terminal selectivity for the remaining X's except X = His. Studies on peptides HSKFR and STSSSCNLSK reveal that the selectivities for N-terminal histidine and serine are comparable to those for the coexistent internal lysine and cysteine, respectively. Proteins including insulin, lysozyme, RNaseA, and therapeutic protein BCArg have been modified by **1** at their N-terminus under conditions of pH 6.3 (room temperature) or 9.2 (37 °C). Further functionalization of the **1**-incorporated proteins with a dansyl azide **3** has been achieved through click chemistry. Ketene **1** can circumvent the oxidation problem encountered using the "[Mn(2,6-Cl$_2$TPP)Cl/alkyne/H$_2$O$_2$" protocol[14] for peptide modification. The present work lends credence to practical application of ketenes as a highly site-selective acylation reagent for N-terminal modification of peptides and proteins.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental section including general, biological mass spectrometry, and procedures not included in the text, NMR spectra of **1**−**3**, results of modification of peptide LSKL using **1**, **2**, and the "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H$_2$O$_2$" protocol, and characterization of modified peptides and proteins (such as MS/MS spectra). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

cmche@hku.hk; bcmkwong@inet.polyu.edu.hk

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) (a) Aslam, M.; Dent, A. *Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences*; Macmillan Reference Ltd.: London, 1998. (b) Lundblad, R. L. *Chemical Reagents for Protein Modification*, 3rd ed.; CRC Press: Boca Raton, FL, 2005. (c) Walsh, C. T. *Post-translational Modification of Proteins: Expanding Nature's Inventory*; Roberts and Company Publishers: Greenwood Village, Colorado, 2006. (d) Francis, M. B. In *Chemical Biology: From Small Molecules to Systems Biology and Drug Design*; Schreiber, S. L., Kapoor, T. M., Wess, G., Eds.; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2007; Vol. 2, p 593. (e) Hermanson, G. T. *Bioconjugate Techniques*, 2nd ed.; Academic Press: San Diego, CA, 2008.

(2) For recent reviews, see: (a) Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13. (b) van Swieten, P. F.; Leeuwenburgh, M. A.; Kessler, B. M.; Overkleeft, H. S. *Org. Biomol. Chem.* **2005**, *3*, 20. (c) Antos, J. M.; Francis, M. B. *Curr. Opin. Chem. Biol.* **2006**, *10*, 253. (d) van Maarseveen, J. H.; Reek, J. N. H.; Back, J. W. *Angew. Chem., Int. Ed.* **2006**, *45*, 1841. (e) Carrico, I. S. *Chem. Soc. Rev.* **2008**, *37*, 1423. (f) Gauthier, M. A.; Klok, H.-A. *Chem. Commun.* **2008**, 2591. (g) Hackenberger, C. P. R.; Schwarzer, D. *Angew. Chem., Int. Ed.* **2008**,

47, 10030. (h) Bernardes, G. J. L.; Castagner, B.; Seeberger, P. H. *ACS Chem. Biol.* **2009**, *4*, 703. (i) Kurpiers, T.; Mootz, H. D. *Angew. Chem., Int. Ed.* **2009**, *48*, 1729. (j) Lin, Y. A.; Chalker, J. M.; Davis, B. G. *ChemBioChem* **2009**, *10*, 959. (k) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 6974. (l) Afagh, N. A.; Yudin, A. K. *Angew. Chem., Int. Ed.* **2010**, *49*, 262. (m) Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213. (n) Canalle, L. A.; Löwik, D. W. P. M.; van Hest, J. C. M. *Chem. Soc. Rev.* **2010**, *39*, 329. (o) Kalia, J.; Raines, R. T. *Curr. Org. Chem.* **2010**, *14*, 138.

(3) For examples, see: (a) Reference 1a. (b) Reference 1b. (c) Jentoft, N.; Dearborn, D. G. *J. Biol. Chem.* **1979**, *254*, 4359. (d) Tanaka, K.; Kamatani, M.; Mori, H.; Fujii, S.; Ikeda, K.; Hisada, M.; Itagaki, Y.; Katsumura, S. *Tetrahedron* **1999**, *55*, 1657. (e) Tanaka, K.; Fujii, Y.; Fukase, K. *ChemBioChem* **2008**, *9*, 2392. (f) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269. (g) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *J. Am. Chem. Soc.* **2002**, *124*, 6063. (h) Rademann, J. *Angew. Chem., Int. Ed.* **2004**, *43*, 4554. (i) Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. *J. Am. Chem. Soc.* **2007**, *129*, 8672. (j) Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S. *Bioconjugate Chem.* **2008**, *19*, 786. (k) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chem.—Asian J.* **2009**, *4*, 630. (l) Antos, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 10256. (m) Antos, J. M.; McFarland, J. M.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2009**, *131*, 6301. (n) Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. *J. Am. Chem. Soc.* **2009**, *131*, 438.

(4) For examples, see: (a) Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992**, *3*, 138. (b) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776. (c) Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923. (d) Li, X.; Zhang, L.; Hall, S. E.; Tam, J. P. *Tetrahedron Lett.* **2000**, *41*, 4069. (e) Yeo, D. S. Y.; Srinivasan, R.; Chen, G. Y. J.; Yao, S. Q. *Chem.—Eur. J.* **2004**, *10*, 4664. (f) Flavell, R. R.; Muir, T. W. *Acc. Chem. Res.* **2009**, *42*, 107.

(5) (a) Zaykov, A. N.; MacKenzie, K. R.; Ball, Z. T. *Chem.—Eur. J.* **2009**, *15*, 8961. (b) Popp, B. V.; Ball, Z. T. *J. Am. Chem. Soc.* **2010**, *132*, 6660. (c) Popp, B. V.; Ball, Z. T. *Chem. Sci.* **2011**, *2*, 690.

(6) For examples, see: (a) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 8150. (b) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7581. (c) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nat. Chem. Biol.* **2007**, *3*, 321. (d) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007. (e) Lemieux, G. A.; de Graffenried, C. L.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2003**, *125*, 4708. (f) Soellner, M. B.; Nilsson, B. L.; Raines, R. T. *J. Am. Chem. Soc.* **2006**, *128*, 8820. (g) Hangauer, M. J.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2008**, *47*, 2394. (h) Lin, Y. A.; Chalker, J. M.; Davis, B. G. *ChemBioChem* **2009**, *10*, 959. (i) Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. *Angew. Chem., Int. Ed.* **2008**, *47*, 2832. (j) Song, W.; Wang, Y.; Qu, J.; Lin, Q. *J. Am. Chem. Soc.* **2008**, *130*, 9654. (k) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644. (l) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16793. (m) Codelli, J. A.; Baskin, J. M.; Agard, N. J.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2008**, *130*, 11486. (n) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664.

(7) (a) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307. (b) Scheck, R. A.; Francis, M. B. *ACS Chem. Biol.* **2007**, *2*, 247. (c) Scheck, R. A.; Dedeo, M. T.; Lavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 11762.

(8) Cennamo, C.; Carafoli, B.; Bonetti, E. P. *J. Am. Chem. Soc.* **1956**, *78*, 3523.

(9) (a) Dixon, H. B. F. *Biochem. J.* **1964**, *92*, 661. (b) Dixon, H. B. F. *J. Protein Chem.* **1984**, *3*, 99.

(10) Abello, N.; Kerstjens, H. A. M.; Postma, D. S.; Bischoff, R. *J. Proteome Res.* **2007**, *6*, 4770.

(11) (a) Gaudriault, G.; Vincent, J.-P. *Peptides* **1992**, *13*, 1187. (b) Sélo, I.; Négroni, L.; Créminon, C.; Grassi, J.; Wal, J. M. *J. Immunol. Methods* **1996**, *199*, 127.

(12) For peptide YGGFL, which bears no internal lysine residue, selective modification at the N-terminal α-amino group has been obtained.[11a]

(13) (a) Shiu, H.-Y.; Chan, T.-C.; Ho, C.-M.; Liu, Y.; Wong, M.-K.; Che, C.-M. *Chem.—Eur. J.* **2009**, *15*, 3839. (b) Ho, C.-M.; Zhang, J.-L.; Zhou, C.-Y.; Chan, O.-Y.; Yan, J. J.; Zhang, F.-Y.; Huang, J.-S.; Che, C.-M. *J. Am. Chem. Soc.* **2010**, *132*, 1886.

(14) Chan, W.-K.; Ho, C.-M.; Wong, M.-K.; Che, C.-M. *J. Am. Chem. Soc.* **2006**, *128*, 14796.

(15) For a review on amide bond formation, see: Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61*, 10827.

(16) For examples of amide bond formation, see: (a) Gunanathan, C.; Ben-David, Y.; Milstein, D. *Science* **2007**, *317*, 790. (b) Shen, B.; Makley, D. M.; Johnston, J. N. *Nature* **2010**, *465*, 1027.

(17) Tidwell, T. T. *Ketenes*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, 2006.

(18) For reviews, see: (a) Tidwell, T. T. *Angew. Chem., Int. Ed.* **2005**, *44*, 5778. (b) Paull, D. H.; Weatherwax, A.; Lectka, T. *Tetrahedron* **2009**, *65*, 6771.

(19) (a) Herriott, R. M.; Northrop, J. H. *J. Gen. Physiol.* **1934**, *18*, 35. (b) Stern, K. G.; White, A. *J. Biol. Chem.* **1938**, *122*, 371. (c) Neuhaus, O. W.; Miller, L. *Arch. Biochem. Biophys.* **1957**, *71*, 162.

(20) (a) Burkhardt, A. E.; Russo, S. O.; Rinehardt, C. G.; Loudon, G. M. *Biochemistry* **1975**, *14*, 5465. (b) Henderson, B. S.; Larsen, B. S.; Schwab, J. M. *J. Am. Chem. Soc.* **1994**, *116*, 5025. (c) Banzon, J. A.; Kuo, J. M.; Miles, B. W.; Fischer, D. R.; Stang, P. J.; Raushel, F. M. *Biochemistry* **1995**, *34*, 743. (d) Fleming, S. A. *Tetrahedron* **1995**, *51*, 12479. (e) Pedone, E.; Brocchini, S. *React. Funct. Polym.* **2006**, *66*, 167. (f) Givens, R. S.; Yousef, A. L.; Yang, S.; Timberlake, G. T. *Photochem. Photobiol.* **2008**, *84*, 185.

(21) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.

(22) Compound **3** was prepared according to the literature method: Díaz, D. D.; Rajagopal, K.; Strable, E.; Schneider, J.; Finn, M. G. *J. Am. Chem. Soc.* **2006**, *128*, 6056.

(23) For KSKFR, the exact ligation site at the N-terminal lysine (α-amino or ε-amino group) could not be identified using MS/MS alone. Note that the pK_a value of the N-terminal α-amino group is considerably lower than that of the ε-amino group of lysine (see the Introduction section).

(24) (a) Cheng, P. N. M.; Leung, Y. C.; Lo, W. H.; Tsui, S. M.; Lam, K. C. *Cancer Lett.* **2005**, *224*, 67. (b) Cheng, P. N.-M.; Lam, T.-L.; Lam, W.-M.; Tsui, S.-M.; Cheng, A. W.-M.; Lo, W.-H.; Leung, Y.-C. *Cancer Res.* **2007**, *67*, 309. (c) Lam, T. L.; Wong, G. K. Y.; Chong, H. C.; Cheng, P. N. M.; Choi, S. C.; Chow, T. L.; Kwok, S. Y.; Poon, R. T. P.; Wheatley, D. N.; Lo, W. H.; Leung, Y. C. *Cancer Lett.* **2009**, *277*, 91. (d) Tsui, S.-M.; Lam, W.-M.; Lam, T.-L.; Chong, H.-C.; So, P.-K.; Kwok, S.-Y.; Arnold, S.; Cheng, P. N.-M.; Wheatley, D. N.; Lo, W.-H.; Leung, Y.-C. *Cancer Cell Int.* **2009**, *9*, 9.

(25) For examples of other types of isolated ketenes, see: (a) Baigrie, L. M.; Seiklay, H. R.; Tidwell, T. T. *J. Am. Chem. Soc.* **1985**, *107*, 5391. (b) Hodous, B. L.; Fu, G. C. *J. Am. Chem. Soc.* **2002**, *124*, 1578. (c) Hodous, B. L.; Fu, G. C. *J. Am. Chem. Soc.* **2002**, *124*, 10006. (d) Wiskur, S. L.; Fu, G. C. *J. Am. Chem. Soc.* **2005**, *127*, 6176.

(26) Sasaki, T.; Kodama, K.; Suzuki, H.; Fukuzawa, S.; Tachibana, K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4550.

(27) The N-terminal selectivity in protein modification by **1** is consistent with (i) the mono-modification revealed by Figure 1, (ii) the presence of readily accessible reactive α-NH$_2$ group at the N-terminus of proteins (compared with internal amino acid residues), (iii) the moderate-to-excellent N-terminal selectivity observed in modification of the library of XSKFR peptides (X $\neq$ His) with **1** (Table 1), and (iv) the detection of N-terminal modified fragment (Figure 2). We have made many efforts to look for fragments corresponding to protein modification at internal amino acid residues, but these fragments have not been detected, though the possibility of their existence in undetectable amounts could not be excluded. In view of the low N-terminal selectivity of 1:1 for **1**-mono-modified peptide HSKFR (which bears histidine) and the occurrence of cysteine modification in ~50% of the **1**-modified peptide STSSSCNLSK, the unobserved modification of histidine or cysteine in **1**-labeled proteins could be rationalized as follows: (i) The moderate-to-excellent N-terminal selectivity for **1**-modified peptides XSKFR (X $\neq$ His) suggests a low reactivity of internal lysine (relative to the N-terminal X residues) toward **1**. Because the mono-modification of HSKFR with **1** gave a 1:1 mixture of N-terminal histidine and internal lysine modified HSKFR, we believe that N-terminal histidine is relatively unreactive toward labeling with ketene **1**. On going from N-terminal histidine to internal histidine, its reactivity toward **1** should be further reduced. Indeed, upon trypsin digestion of the **1**-modified protein BCArg, the histidine-containing fragments (such as LMHHHHHH) detected by LC−MS/MS were not found to be labeled by **1**. (ii) STSSSCNLSK contains a reactive free thiol group in the cysteine side chain, but the protein BCArg is devoid of cysteine, and in the structures of insulin, lysozyme, and RNaseA, all the cysteine residues are covalently paired through disulfide bridge. In previous work,[14] we observed exclusive N-terminal modification of STSSSCNLSK, without cysteine modification, by in situ generated ketene using the "[Mn(2,6-Cl$_2$TPP)Cl]/alkyne/H$_2$O$_2$" protocol, a phenomenon that can be attributed to the oxidation of the cysteine thiol group to disulfide under oxidizing conditions, as mentioned above in the Results section.

(28) The unreacted ketene is likely to be removed by ultrafiltration using, for example, commercially available PALL Nanosep centrifugal devices (which have different molecular cut off such as 3K according to manufacturing instruction).

(29) Harris, J. M.; Chess, R. B. *Nat. Rev. Drug Discovery* **2003**, *2*, 214.

(30) For reviews, see: (a) Truong, K.; Ikura, M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 573. (b) Heyduk, T. *Curr. Opin. Biotechnol.* **2002**, *13*, 292. (c) Roy, R.; Hohng, S.; Ha, T. *Nat. Methods* **2008**, *5*, 507.

(31) The bioactivity of protein before and after treatment with ketene **1** was examined using lysozyme as an example by lysozyme assay purchased from Invitrogen. No significant influence on the lysozyme activity by labeling with **1** was observed (see Figure S111 in the Supporting Information). Note that THF (which may cause protein denaturation) is not a solvent required for protein modification with **1**; it was used in this work for preparing a stock solution of **1**, and the reaction mixture contained only 2.5−15% v/v THF with water being the predominant solvent.

(32) Lundblad, R. L. *Techniques in Protein Modification*; CRC Press: Boca Raton, FL, 1995.