Organic LETTERS

Letter

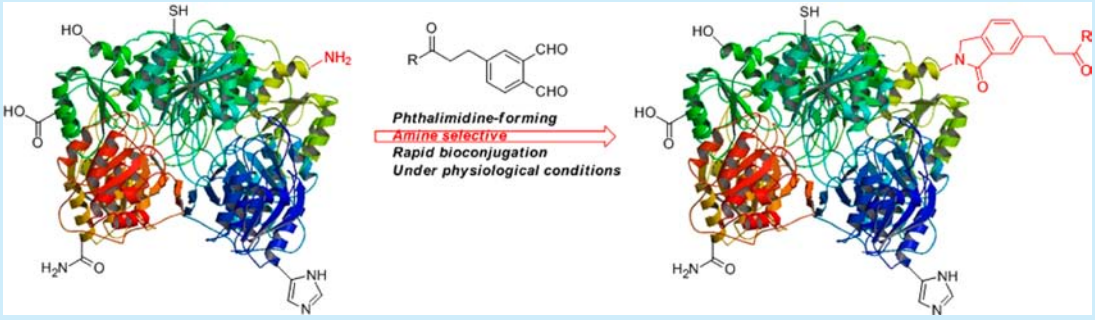
pubs.acs.org/OrgLett

Traceless and Chemoselective Amine Bioconjugation via Phthalimidine Formation in Native Protein Modification

Chun Ling Tung, Clarence T. T. Wong, Eva Yi Man Fung, and Xuechen Li*

Department of Chemistry, The State Key Laboratory of Synthetic Chemistry, The University of Hong Kong, Hong Kong, China

Supporting Information



ABSTRACT: *ortho*-Phthalaldehyde (OPA) and its derivatives are found to react chemoselectively with amino groups on peptides and proteins rapidly and tracelessly under the physiological condition via formation of phthalimidines, which provides a novel and promising approach when performing bioconjugation on native proteins. The notable advantages of this method over the existing native protein lysine-labeling approaches include a traceless process, a self-reacting, specific and fast reaction, ease of operation, and the ability to use nonhydrolyzable reagents. Its applications have been effectively demonstrated including conjugation of peptides and proteins, and generation of an active PEGlyated L-asparaginase.

Native protein modification has found many applications in studying the functions of the intracellular proteins such as immunostaining and immunolocalization and developing therapeutic biologics such as polyethylene glycol (PEG)−protein conjugate drugs, antibody−drug conjugates (ADC), carbohydrate−protein conjugate vaccines, and so on.[1] Thus, new approaches and strategies for effective protein bioconjugation are continually being developed.[2]

Native protein modification often involves the conjugation chemistry to take place at the nucleophilic side chain of the amino acids present within a protein under the physiological conditions, among which lysine and cysteine are the most common labeling sites.[3] As the thiol group of cysteine is a superior nucleophile, there are a handful of functional groups capable of chemoselectively reacting with the N-terminal or internal cysteine residues.[2,4] Nevertheless, most cysteine residues on native proteins are engaged in forming disulfides with other cysteine residues that are important for function, protein folding, and stability. Conjugation at the ε-amine group of the lysine residue and the N-terminus involves using acyl donors as the labeling reagent.[3,5]

One of most common methods is to use *N*-hydroxysuccinimide (NHS) esters, which were introduced 30 years ago.[6] Various homobifunctional NHS esters have been used to conjugate dyes, biotin, PEG, and drugs onto native proteins.[3] However, studies have revealed that the reaction of NHS esters is not completely chemoselective toward amino groups and NHS esters could also react with histidine, serine, and tyrosine.[7,8] Furthermore, extensive washing and purification steps are needed to remove the NHS residue before performing biological assays. Some traceless acyl donors are being used, such as isothiocyanates and isocyanates, which however, require a higher pH (>9) condition.[9] Generally speaking of the methods for native protein labeling at amino group sites, one often encounters the problem of instability or low reactivity of the acyl donor reagents or generation of byproducts. In addition, the reaction condition is also critical to the effectiveness of protein bioconjugation, wherein pH, solvents, and temperature vary for different labeling reagents.[3]

During the course of our previous work in the development of a three-component reaction of *ortho*-phthalaldehyde (OPA), amines, and isonitriles to generate 1-carboxamide isoindoles, before an optimal condition (2 equiv $NaHSO_3$ as additive in $DMSO/H_2O$) was found, the N-substituted isoindolin-1-ones (phthalimidines) were observed as the major or sole product in all the organic solvents tested.[10] Indeed, such a reaction of OPA and amines to form phthalimidines was well documented, first reported by Thiele in 1909,[11a] and it has been used to construct small molecules with the isoindolinone skeleton under the Lewis or Brønsted acid catalyzed condition generally with low to modest yields.[11] The reaction proceeds likely via formation of an iminoaldehyde followed by an energetically favorable [1,5]-H sigmatropic rearrangement of the aldehydic proton as the key

Received: April 5, 2016
Published: May 18, 2016

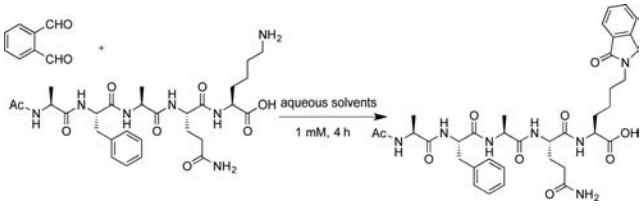
DOI: 10.1021/acs.orglett.6b00983
Org. Lett. 2016, 18, 2600−2603

step of the reaction pathway.[12] Nevertheless, the reaction between OPA and amines in the absence of external reagents under physiological conditions with which to label and modify native proteins has not been explored previously.

At the time, we were pondering whether the reaction of OPA and amines could well proceed under the aqueous condition, enabling phthalimidine-mediated native protein labeling. To this end, the key criteria for biocompatible chemical transformations have to be met, including aqueous solvent ideally with a neutral pH, ambient temperature, low reactant concentrations, stability of the reactant under aqueous conditions, and chemoselective reaction, all of which make the development of effective strategies challenging. Herein, we report that OPA and its derivatives could react chemoselectively with amino groups on native proteins rapidly under the physiological condition via formation of phthalimidines, providing an alternative and promising tool for native protein bioconjugation/labeling. The notable advantages of this method over the existing protein lysine-labeling approaches include traceless process, fast and specific reaction, ease of operation, and using nonhydrolyzable reagents.

Our studies started with the reaction of OPA with an N-terminal acetylated pentapeptide (*Ac*-AFAQK-*OH*) under aqueous conditions. The reaction could proceed in water, but in a very low conversion (12%) after 4 h, as analyzed by RP-LCMS (Table 1, entry 1). Addition of salts (e.g., NaCl, KCl) did

Table 1. Screening of the Reaction Condition



entry	aqueous solutions	conversion (%)
1	H_2O	12
2	100 mM NaCl	10
3	100 mM KCl	19
4	sodium acetate buffer (pH 3.0)	38
5	PBS (pH 7.4)	100
6	borate buffer (pH 10.0)	100
7	10 mM Na_2HPO_4	100
8	10 mM NaH_2PO_4	5
9	HEPES buffer (pH 6.8)	100
10	10 mM tricine buffer (pH 7.4)	100

not help improve the conversion (Table 1, entries 2 and 3). As acidic conditions would be beneficial for the phthalimidine formation in organic medium,[11] sodium acetate buffer (pH = 3.0) was tested but found to be ineffective (Table 1, entry 4). To our delight, when the PBS (pH 7.4) and borate buffer (pH 10.0) were tested, a full conversion of the starting peptide to a new product was observed within 4 h (Table 1, entries 5 and 6). Interestingly, sodium hydrogen phosphate alone could also facilitate the reaction (Table 1, entry 7). Other neutral buffers, such as and HEPES and tricine buffer (Table 1, entries 9 and 10), also worked well for such a reaction. The fact that the reaction in neutral buffer proceeds much faster than that in water indicates that the acid−base catalysis is likely involved.

The product of OPA and amines generated under the aqueous condition (Table 1, entry 5) was isolated and characterized with ^1H and ^{13}C NMR spectra as a phthalimidine derivative. Since phthalimidines could be formed under the PBS buffer at a neutral pH from the reaction of OPA with primary amines, this would be an ideal condition for performing peptides and proteins modification. Therefore, we chose this condition as our main focus to study its bioconjugation potentials. To study the concentration effect on the rate of OPA and peptide conjugation, time course LC−MS studies of the pentapeptide at different substrate concentrations were conducted (Figure S2). Even the pentapeptide was at a diluted concentration of 100 μM, a full conversion to the desired product was still observed at 16 min. Not only OPA showed high reactivity, it also had great chemoselectivity toward primary amines. Four peptides (*Ac*-RMF-*OH*, *Ac*-DTHC-*OH*, *Ac*-WYQ-*OH*, and *H*-PFASPLPIP-*OH*) containing all the amino acid side chain functional groups except the lysine amine were treated with OPA for 24 h and no reaction was observed. It is worth mentioning that the internal cysteine residue did not react with OPA under this condition and the N-terminal proline was also inert. These results showed that OPA reacted exclusively with primary amines in PBS (pH 7.4). It is noted that the N-terminal amine, except N-terminal Pro, also reacted under this condition. Several proteins including cytochrome c, lysozyme, ribonuclease A, and BSA have thus been modified and confirmed by MALDI-TOF (Figures S18−S23 in the Supporting Information).

In comparison with the widely used NHS chemistry, both OPA and *N*-hydroxysuccinimide (NHS) activated ester **1** were reacted with a model peptide (*Ac*-SHTHYGAK-*OH*) for 6 h (OPA or NHS ester/peptide: 1 mol/1 mol; 50 μM). Along with a large quantity of the unmodified peptide, 2 singly (MW 1022.4) and 2 doubly (MW 1102.5) modified peptides were observed under the NHS ester conditions, which probably arise from nonspecific reactions with other nucleophilic residues like serine, histidine, and tyrosine in the peptide.[7] The reaction with OPA resulted in a singly modified product (MW 1058.4) with only trace amount of the unmodified peptide (Figure 1). Next,

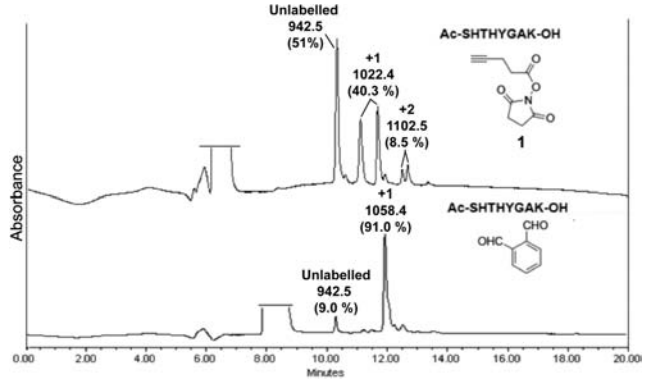


Figure 1. LC−MS chromatogram of crude peptide modifications with NHS ester (1.0 equiv) and OPA (1.0 equiv). The NHS ester resulted in single and double modifications (top); OPA resulted in a single modification (bottom). On the basis of a full PDA wavelength 190−400 nm, the ratio of the peaks was calculated by Waters' Empower system.

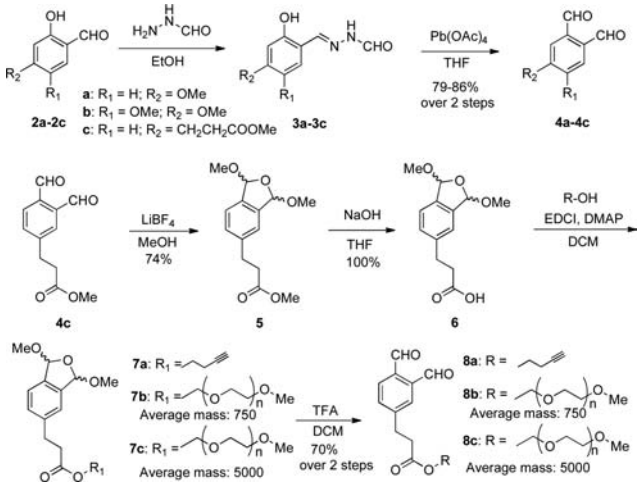
the OPA-conjugated product was isolated and analyzed by HCD MS/MS, which showed the modification to be at the lysine site forming a phthalimidine (Figure S10). These results indicate that OPA has a higher degree of selectivity and efficiency toward primary amines modification compared to the NHS ester.

To allow for attachment of various functional groups onto proteins, OPA derivatives were synthesized, which included

DOI: 10.1021/acs.orglett.6b00983
Org. Lett. 2016, 18, 2600−2603

different sizes of PEGs, an alkyne, and a carboxylic acid group (Scheme 1). The substituent effects on the rate of reactions

Scheme 1. Synthesis of Functionalized OPA Derivatives

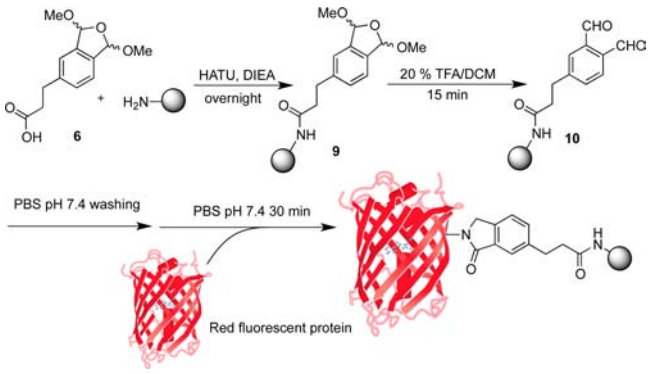


were investigated. The synthesis of OPA derivatives started with salicylaldehyde derivatives **2a−2c** to be refluxed with formic hydrazide to form **3a−3c**, which was then treated with lead(IV) acetate to yield OPA derivatives **4a−4c**. Different functions groups were attached by simple EDCI-mediated coupling of compound **6** and thus **8a−8c** were synthesized. Then all the synthetic derivatives were tested with the pentapeptide (*Ac*-AFAQK-*OH*) in the PBS buffer. The reactions of all the derivatives showed to achieve more than 80% conversion at 16 min (Figure S3 in the Supporting Information). In comparison, an OPA derivative bearing an alkyne group **8a** could be conjugated to the peptide completely after 32 min, while an NHS-alkyne **1** only gave a 67% conversion within the same amount of time. These results showed that OPA has a faster conversion reacting with amines, and thus, the second order reaction rate constants of OPA and a model peptide at different pH conditions were determined. The rate constant of NHS-alkyne **1** was also compared with OPA-alkyne **8a** under the same condition (Table S2 in the Supporting Information). These results showed that the conjugation rate of OPA-alkyne (7.49 M^{-1} s^{-1}) was 5.9-fold faster than that of NHS-alkyne (1.27 M^{-1} s^{-1}) and the reaction became even faster under the basic conditions by the second order rate constant. Although the reaction products are not the same, the results indicate that the OPA-amine reaction proceeds also faster than other conjugation and ligation methods.[14]

Protein immobilization onto a solid support such as resin, glasses, metal, and nanomaterials have been widely applied for drug delivery, microarray, and ELISA, etc.[13] As such, we explored the use of this phthalimidine-forming bioconjugation method to perform protein immobilization. Compound **6** (Scheme 2) was coupled onto Rink amide resin. Red fluorescent protein (RFP) in PBS solution was allowed to react with the resin **10** for 30 min. After washing, as seen under the fluorescent microscope, the RFP was successfully immobilized onto the solid support in its native state (Figure 2).

PEGylated pharmaceutical proteins exhibit improved clinical properties including physical and thermal stability, protection against enzymatic degradation, better solubility, prolonged blood circulating half-life, reduced immunogenicity, and reduced toxicity.[15] Up to date, nine PEGylated protein drugs

Scheme 2. Protein Immobilization



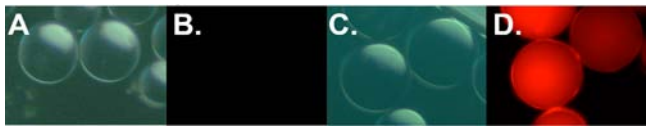


Figure 2. Fluorescent microscopic images of protein immobilization using an OPA derivative. (A) Bright-field image of rink amide resin coupled with OPA derivative. (B) Fluorescent image of rink amide resin coupled with OPA derivative. (C) Bright-field image of rink amide resin coupled with OPA derivative after conjugated with RFP. (D) Fluorescent image of rink amide resin coupled with OPA derivative after conjugated with RFP.

have been approved by the FDA, for six of which PEGylations are realized using the NHS ester method.[16] However, the NHS ester method has some potential issues, including the instability of the NHS ester, generation of NHS byproduct which requires careful purification, and nonspecific reactions at lysine, serine, tyrosine, and histidine sites.[16] In this regard, development of an improved PEGylation method would be of great value.

To investigate the potential of our newly developed bioconjugation method in PEGylating native proteins, we synthesized an OPA derivative **8b** bearing monomethyl ethylene glycol with average Mn 750 (mPEG750) and reacted it with cytochrome c as a model protein (Figure S25). The PEGylation of cytochrome c was performed at a 1:2 lysine to OPA-mPEG750 ratio, with a final protein concentration of 50 μM in PBS (pH 7.4). The degree of PEGylation was monitored by SDS-PAGE and MALDI-TOF at different time points. The results showed that all the accessible 19 lysine amines on cytochrome c were conjugated with mPEG750. The mass of the conjugating protein was increasing due to incomplete PEGylation from 0.5 to 2 h. After 4 h, a constant mass was observed as the fully PEGylated cytochrome c product.

Pegaspargase (Oncaspar), a PEGylated form of *E. coli* L-asparaginase, is used to treat childhood acute lymphoblastic leukemia.[17] By depleting the asparagine level in blood, the malignant lymphoid cells that rely on extracellular asparagine thus undergo apoptosis.[18] PEGylation of asparaginase is usually performed randomly on lysine residues with PEG having an average mass of 5000 Da using the NHS ester method.[16] It was noted that the PEGylation also took place at serine, tyrosine, and histidine sites.[16] To explore the utility of the OPA-amine bioconjugation developed here, an OPA derivative equipped with mPEG (**8c**) with an average Mn of 5000 was reacted with *E. coli* L-asparaginase for 4 h under the same conditions as above. The PEGylated protein was then purified by size exclusion column chromatography. The activity of thus-obtained

DOI: 10.1021/acs.orglett.6b00983
Org. Lett. 2016, 18, 2600−2603

PEGlyated asparaginase was determined by the asparaginase activity kit (Sigma-Aldrich). In contrast to the native L-asparaginase, the obtained PEGylated asparaginase retained 46.7% of the initial activity (Figure 3). Next, the OPA-

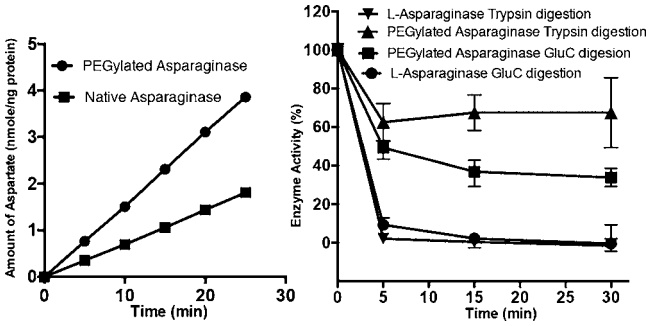


Figure 3. Enzymatic activity (left) and proteolysis stability of the PEGylated and native L-asparaginase (right).

PEGylated L-asparaginase was able to be resistant to proteolytic degradation by enzymes including trypsin and Glu-C and retained 70.0% and 40% of activities respectively; while non-PEGylated asparaginase lost its activity completely after enzymatic proteolysis (Figure 3).

In summary, we have demonstrated that the OPA-amine reaction to form a phthalimidine is an effective biocompatible chemical transformation, enabling phthalimidine-mediated native protein bioconjugation. Under the physiological buffer condition (e.g., PBS buffer, pH 7.4), OPA and its derivatives efficiently and chemoselectively react with the amino group of native proteins, which provides a user-friendly and promising method for native protein modification. This method has the characteristics of operational simplicity, chemoselective reaction, low reactant concentration, traceless process, and readily assessable and nonhydrolyzable reagents. Its potential applications have been demonstrated in functionalization, immobilization, and PEGylation of native proteins.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.6b00983.

Experimental procedures and characterization data (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: xuechenl@hku.hk.

Author Contributions

C.L.T. and C.T.T.W. contributed equally. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by the National Basic Research Program of China (Grant 2013CB836900), the Area of Excellence Scheme of the University Grants Committee (Grant AoE/M-12/06), Strategic Research Theme on Drug of the University of Hong Kong, Shenzhen Basic Research Grant (Grant JCYJ20140903112959961), and the CAS Interdisciplinary Innovation Team.

■ REFERENCES

(1) (a) Biju, V. *Chem. Soc. Rev.* **2014**, *43*, 744. (b) Takaoka, Y.; Ojida, A.; Hamachi, I. *Angew. Chem., Int. Ed.* **2013**, *52*, 4088. (c) Boutureira, O.; Bernardes, G. J. L. *Chem. Rev.* **2015**, *115*, 2174.

(2) (a) MacDonald, J. I.; Munch, H. K.; Moore, T.; Francis, M. B. *Nat. Chem. Biol.* **2015**, *11*, 326. (b) Chan, A. O.-Y.; Ho, C.-M.; Chong, H.-C.; Leung, Y.-C.; Huang, J.-S.; Wong, M.-K.; Che, C.-M. *J. Am. Chem. Soc.* **2012**, *134*, 2589. (c) Abbas, A.; Xing, B.; Loh, T.-P. *Angew. Chem., Int. Ed.* **2014**, *53*, 7491.

(3) (a) Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213. (b) Spicer, C. D.; Davis, B. G. *Nat. Commun.* **2014**, *5*, 4740.

(4) (a) Kurpiers, T.; Mootz, H. D. *Angew. Chem., Int. Ed.* **2007**, *46*, 5234. (b) Chudasama, V.; Smith, M. E.; Schumacher, F. F.; Papaioannou, D.; Waksman, G.; Baker, J. R.; Caddick, S. *Chem. Commun.* **2011**, *47*, 8781. (c) Abad, S.; Nolis, P.; Gispert, J. D.; Spengler, J.; Albericio, F.; Rojas, S.; Herance, J. R. *Chem. Commun.* **2012**, *48*, 6118. (d) Dhar, T.; Kurpiers, T.; Mootz, H. D. *Methods Mol. Biol.* **2011**, *751*, 131. (e) Sejwal, P.; Han, Y.; Shah, A.; Luk, Y. *Org. Lett.* **2007**, *9*, 4897.

(5) (a) Stephanopoulos, N.; Francis, M. B. *Nat. Chem. Biol.* **2011**, *7*, 876. (b) Chen, X.; Muthoosamy, K.; Pfisterer, A.; Neumann, B.; Weil, T. *Bioconjugate Chem.* **2012**, *23*, 500. (c) Fernandez-Moreira, V.; Alegre-Requena, J. V.; Herrera, R. P.; Marzo, I.; Gimeno, M. C. *RSC Adv.* **2016**, *6*, 14171.

(6) Lomant, A. J.; Fairbanks, G. *J. Mol. Biol.* **1976**, *104*, 243.

(7) Kalkhof, S.; Sinz, A. *Anal. Bioanal. Chem.* **2008**, *392*, 305.

(8) Swaim, C. L.; Smith, J. B.; Smith, D. L. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 736.

(9) (a) Morales, A. R.; Schafer-Hales, K. J.; Marcus, A. I.; Belfield, K. D. *Bioconjugate Chem.* **2008**, *19*, 2559. (b) Varadarajan, A.; Sharkey, R. M.; Goldenberg, D. M.; Hawthorne, M. F. *Bioconjugate Chem.* **1991**, *2*, 102.

(10) Zhang, Y. F.; Lee, C. L.; Liu, H.; Li, X. *Org. Lett.* **2012**, *14*, 5146.

(11) (a) Thiele, J.; Schneider, J. *Liebigs Annalen* **1909**, *369*, 287. (b) Takahashi, I.; Hatanaka, M. *Heterocycles* **1997**, *45*, 2475. (c) Kulla, E.; Zuman, P. *Org. Biomol. Chem.* **2008**, *6*, 3771. (d) Cronin, J.R.; Pizzarello, S.; Gandy, W. E. *Anal. Biochem.* **1979**, *93*, 174. (e) Pérard-Viret, J.; Prangé, T.; Tomas, A.; Royer, J. *Tetrahedron* **2002**, *58*, 5103.

(12) Alajarin, M.; Sanchez-Andrada, P.; Lopez-Leonardo, C.; Alvarez, A. *J. Org. Chem.* **2005**, *70*, 7617.

(13) (a) Camarero, J. A. *Biopolymers* **2008**, *90*, 450. (b) Wang, T.-H.; Lee, W.-C. *Biotechnol. Bioprocess Eng.* **2003**, *8*, 263. (c) Taylor, R. H.; Fournier, S. M.; Simons, B. L.; Kaplan, H.; Hefford, M. A. *J. Biotechnol.* **2005**, *118*, 265. (d) Besselink, T.; Liu, M.; Ottens, M.; van Beckhoven, R.; Janssen, A. E.; Boom, R. M. *J. Sep. Sci.* **2013**, *36*, 1185.

(14) Saito, F.; Noda, H.; Bode, J. W. *ACS Chem. Biol.* **2015**, *10*, 1026.

(15) (a) Abuchowski, A.; van Es, T.; Palczuk, N. C.; Davis, F. F. *J. Biol. Chem.* **1977**, *252*, 3578. (b) Harris, J. M.; Chess, R. B. *Nat. Rev. Drug Discovery* **2003**, *2*, 214.

(16) Alconcel, S. N. S.; Baas, A. S.; Maynard, H. D. *Polym. Chem.* **2011**, *2*, 1442.

(17) (a) Graham, M. L. *Adv. Drug Delivery Rev.* **2003**, *55*, 1293. (b) Ortega, J. A.; Nesbit, M. E., Jr.; Donaldson, M. H.; Hittle, R. E.; Weiner, J.; Karon, M.; Hammond, D. *Cancer Res.* **1977**, *37*, 535.

(18) Masurekar, A.; Fong, C.; Hussein, A.; Revesz, T.; Hoogerbrugge, P. M.; Love, S.; Ciria, C.; Parker, C.; Krishnan, S.; Saha, V. *Blood Cancer J.* **2014**, *4*, e203−1.

DOI: 10.1021/acs.orglett.6b00983
Org. Lett. 2016, 18, 2600−2603