Organic LETTERS

Letter

pubs.acs.org/OrgLett

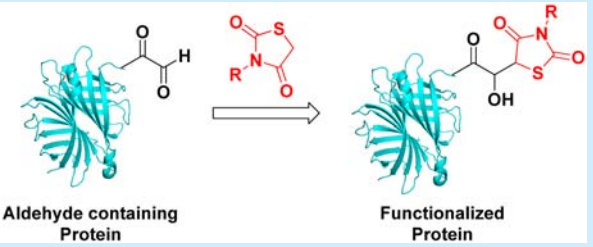
Site-Specific Chemical Modification of Peptide and Protein by Thiazolidinediones

Peng Wang, Shumei Zhang, Qiuyue Meng, Ying Liu, Luqing Shang,* and Zheng Yin*

College of Pharmacy & State Key Laboratory of Elemento-Organic Chemistry, Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Nankai University, Tianjin 300071, China

Supporting Information

ABSTRACT: A direct aldol reaction employing 2,4-thiazolidinediones as nucleophilic donors was performed to modify peptides and protein under mild conditions. Various functional groups could be readily introduced into protein without conformation change.



Chemical modification of proteins is an important tool for monitoring, modulating, and tracking the behavior of proteins in living systems.[1] Thus, it is widely used in biochemical assays,[2] diagnosis,[3] imaging in vivo,[4] PEGylation,[5] pharmaceutical research,[6] etc. To this end, a number of bioorthogonal reactions have been developed over the last few decades. The chemical functionalization of protein at an early stage is an example of a classic nucleophilic/electrophilic modification of the natural amino acids, such as lysine,[7] cysteine,[8] tryptophan,[9] etc. Selectivity is the major challenge because most proteins display multiple copies of the targeted residue on their surface.[10] To address this issue, some unnatural amino acids containing azide, alkene or halogen were introduced into protein to achieve site-specific modification.[10,11] Development of site-specific methods for chemical modification of protein have attracted significant amount of attention.

Regardless of azide and alkyne groups, aldehyde and ketone are preferable choices as a chemical handle because of their unique reactivity as mild electrophiles.[12] Moreover, aldehyde could be easily introduced into protein by either a genetic method[13] or an oxidative biomimetic transamination reaction.[14] Subsequently, aldehyde could be further modified with various reagents to form the structures of oxime,[15] imine, hydrazine,[16] etc. However, such ligations normally may suffer from slow kinetics and could only be accelerated by catalysts that may limit their further applications in a biological environment.[17] More modifications could be achieved by the Pictet−Spengler reaction,[18] Wittig reaction,[19] and Mukaiyama aldol reaction.[20] However, the Pictet−Spengler reaction could attain high conversion only under acidic conditions, and the Wittig reaction or Mukaiyama aldol reaction may have limitations on the substrates.[18b,19,20] The goal of our research is to develop a novel bioorthogonal reaction with aldehyde as a chemical handle using a substrate that could be easily diversified. The aldol reaction is one of the most prevalently investigated and applied to C−C bond formation in organic synthesis. To the best of our knowledge, there is no publication of enabling a direct aldol reaction for the site-specific modification of protein. The reason could be that direct aldol reaction under biological conditions, such as the addition of large amount of water or aqueous buffer as solvent, typically results in low yield as well as slow kinetics,[21] which is not really appropriate for the modification of protein. Recently, it was reported that 2,4-thiazolidinediones could serve as an aldol nucleophilic donor, and the reaction proceeded rapidly "on water" under moderate conditions.[22] We envisaged that 2,4-thiazolidinediones could be employed for the site-specific modification of protein in biological conditions, and structure diversity could be readily introduced into the protein through easy modification on the nitrogen of 2,4-thiazolidinediones. Herein we report a direct aldol reaction using thiazolidinediones as the donor for site-specific modification of peptide and protein.

To investigate the feasibility of this direct aldol reaction on protein modification, an initial model study was undertaken on the dipeptide **1** (Scheme 1). The N-terminal serine was oxidized into aldehyde by $NaIO_4$ under moderate conditions to afford **2**. The aldol reactions between **2** and 3-methyl-2,4-thiazolidinedione (**3a**) in various organic solvents with different polarity were investigated (entries 1−5, Table 1). Surprisingly, this reaction did not proceed in all tested organic solvents under the described conditions. When water was applied as the solvent (entry 6, Table 1), no product was detected as well. We speculated that the "salting out effect" could promote this reaction in aqueous media.[23] However, no reaction took place in either brine or 4 M LiCl aqueous solution (entries 7 and 8, Table 1). Eventually, we were delighted to find that this aldol reaction proceeded very well in phosphate buffer solution (25 mM, pH = 7.0) and gave 90% yield of the desired product.

Received: January 2, 2015
Published: March 4, 2015

DOI: 10.1021/acs.orglett.5b00005
Org. Lett. 2015, 17, 1361−1364

Scheme 1. Model Study by Using Dipeptide 1

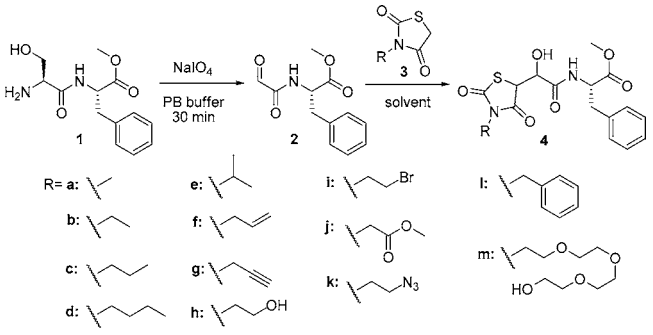


Table 1. Optimization of the Aldol Reaction of Dipeptide Aldehyde 2 and 3a[a]

entry	catalyst	solvent[b]	yield[c] (%)	ratio[d]
1	NO	DMSO	ND[e]	
2	NO	DMF	ND	
3	NO	CH_3OH	ND	
4	NO	1,4-dioxane	ND	
5	NO	DCM	ND	
6	NO	H_2O	ND	
7	NO	brine	ND	
8	NO	H_2O (4 M LiCl)	ND	
9	NO	PB buffer	90	57:43
10	NO	citric acid–NaOH buffer	82	61:39
11	NO	Tris–HCl buffer	30	65:35
12	L-proline (20 mol %)	PB buffer	91	53:47
13[f]	NO	PB buffer	91	55:45

[a]Unless specified, a mixture of **2** (23 mg, 0.1 mmol) and **3a** (20 mg, 0.15 mmol) in a solvent (0.5 mL) was stirred at for 6 h. [b]Unless specified, the concentration of buffer was 25 mM and the pH value was 7.0. [c]Isolated yield. [d]Isomeric ratio was determined by ^{13}C NMR analysis of the crude reaction mixture. [e]Not detected by TLC. [f]3 equiv of **3a** was used.

Citric acid–NaOH buffer gave results comparable to those for the phosphate buffer (entry 10, Table 1), while a lower yield was observed in Tris–HCl buffer (entry 11, Table 1) at the same pH value. The intriguing solvent effect on the reaction still remains to be explained.

The diastereoselectivity of the reaction was poor as shown in Table 1. L-Proline was reported to promote stereoselective aldol reaction in the presence of water.[24] However, it had no significant effect on the reaction (entry 12, Table 1). It is likely that the ionic interactions and/or hydrogen bonds that are critical for stabilizing the transition state of reaction were interrupted by the large amount of water, consequently altering the stereoselectivity.[25] Nevertheless, the development of the bioorthogonal reaction placed emphasis primarily on efficiency. Therefore, the optimization efforts were focused little on the stereochemistry of this reaction.

Further optimization of reaction conditions demonstrated that the addition of L-proline as catalyst and the extra addition of **3a** had no significant influence on yield (entry 13, Table 1). As a result, 1.5 equiv of **3a** in the presence of no catalyst in phosphate buffer was chosen as the reaction condition for further study.

Next, the pH-dependence of the reaction was screened, noting a higher yield at neutral to basic pH (Table 2). The pH was fixed at 7.0 because it not only gave a high yield in the aldol reaction but also was close to the physiological environment.

Table 2. pH-Dependence Study of the Aldol Reaction of Dipeptide Aldehyde 2 and 3a[a]

entry	pH	yield[b] (%)	ratio[c]
1	PB buffer, pH = 2.3	ND[d]	
2	PB buffer, pH = 4.8	trace	
3	PB buffer, pH = 6.0	42	60:40
4	PB buffer, pH = 6.5	79	56:44
5	PB buffer, pH = 7.0	90	57:43
6	PB buffer, pH = 8.0	89	62:38
7	Na_2CO_3–$NaHCO_3$ buffer, pH = 9.1	90	65:35

[a]Unless specified, see footnote *a* in Table 1. [b]Isolated yield. [c]Isomeric ratio was determined by ^{13}C NMR analysis of the crude reaction mixture. [d]Not detected by TLC.

The solubility of substrate is an important factor for aqueous organic reactions. Accordingly, **3a–d** (entries 1–4, Table 3) were synthesized to study the effect of solubility on the aldol ligation. While the solubility of **3a–d** decreased with increasing length of the carbon chain, the rate of reaction declined along with the solubility. It may suggest that a shorter carbon chain (no more than two carbons) or hydrophilic chain could be a better candidate for the linkage. Isopropyl-substituted 2,4-thiazolidinediones (**3e**) were synthesized to explore the steric effect on the reaction. Compared to **3c**, **3e** gave much lower yield even though the two substrates possessed similar solubility. This may indicate that bulky substituent groups slow the reaction.

Table 3. Solubility and Steric Effect Study of the Aldol Reaction of Dipeptide Aldehyde 2 and 3[a]

entry	R	time (h)	yield[b] (%)	ratio[c]	solubility of **3** (mg/mL)[d]
1	**a**	6	90	57:43	≥50
2	**b**	24	84	60:40	43
3	**c**	24	65	61:39	18
4	**d**	24	40	54:46	4
5	**e**	24	37	65:35	14

[a]Unless specified, a mixture of **2** (23 mg, 0.1 mmol) and **3** (0.15 mmol) in PB buffer (0.5 mL, 25 mM, pH = 7.0) was stirred for a specified time period. [b]Isolated yield. [c]The isomeric ratio was determined by ^{13}C NMR analysis of the crude reaction mixture. [d]30 mg of **3** was added to 0.6 mL of PB buffer (25 mM, pH = 7.0). The system was incubated in a shaker (100 rpm, 37 °C) for 48 h. The solubility was obtained by HPLC analysis.

Structure diversity could be readily introduced to 2,4-thiazolidinedione by a one-step substitution reaction with bromides. A number of selected functional N-substituted 2,4-thiazolidinediones were investigated to explore the scope of this direct aldol reaction. It was inspiring that a variety of substrates performed well to afford the aldol products with yields ranging from 52 to 91% (Table 4). Different functional groups including olefin, alkyne, ester, azide, bromide, and a PEG chain were introduced to the peptide model. It was noteworthy that the reaction with olefin and alkyne substrates proceeded rapidly with high yield. The results indicated that this method could be applicable for functionalization of protein.

Thioglycol (200 mM) was used for a thiol competition test, and it had no effect on the efficiency of this reaction. To further

DOI: 10.1021/acs.orglett.5b00005
Org. Lett. 2015, 17, 1361–1364

Table 4. Substrate Study of the Aldol Reaction of Dipeptide Aldehyde 2 and 3[a]

entry	R	time (h)	yield[b] (%)	ratio[c]
1	**f**	6	91	69:31
2	**g**	6	88	65:35
3	**h**	6	89	60:40
4	**i**	48	70	63:37
5	**j**	48	82	65:35
6	**k**	48	73	57:43
7	**l**	48	52	62:38
8	**m**	24	72	58:42

[a]Unless specified, see footnote *a* in Table 3. [b]Isolated yield. [c]Isomeric ratio was determined by ^{13}C NMR analysis of the crude reaction mixture.

test this aldol ligation reaction toward the standards of bioorthogonality in biological environments of increasing complexity, **2** and **3a** were subjected to cell media (DMEM +5% FBS) and an aqueous solution containing 10% untreated lysate (Table 5).[26] The reactions were carried out at room

Table 5. Tolerance Test of the Aldol Reaction of Dipeptide Aldehyde 2 and 3a in Cell Media and Cell Lysate[a]

entry	solvent	yield[b] (%)	ratio[c]
1	PB buffer/DMEM (+5% FBS) = 1:1	83	64:36
2	PB buffer/lysate =9:1	85	62:38

[a]Unless specified, a mixture of **2** (23 mg, 0.1 mmol) and **3a** (20 mg, 0.15 mmol) in a solvent (0.5 mL) was stirred for 12 h. [b]Isolated yield. [c]Isomeric ratio was determined by ^{13}C NMR analysis of the crude reaction mixture.

temperature for 12 h and followed by standard workup and purification. The results suggested that the aldol reaction proceeded efficiently in cell media and cell lysate. A kinetic study was conducted in PB buffer (25% of DMSO-d_6) by NMR. The second-order rate constant for **2** and **3a** was 0.0078 M^{-1} s^{-1} (Figure S1, Supporting Information), which was faster than the Staudinger reaction (0.0025 M^{-1} s^{-1}).[27] Overall, the results were encouraging for the modification of protein by this direct aldol reaction.

Subsequently, we sought to apply this reaction to the modification of protein. Myoglobin was selected to be the protein model because its N-terminal residue glycine could be conveniently oxidized to aldehyde by pyridoxal-5′-phosphate (PLP).[25] Compound **3a** was chosen for the protein modification because of its high reactivity. In the same PB buffer which was used in the PLP oxidation (25 mM, pH = 6.5), the aldehyde containing myoglobin was subjected to **3a** in 1.5 mL plastic tube with final protein concentration 37.5 μM and incubated at 37 °C for 3 h in a shaker (100 rpm). As expected, the desired modified myoglobin was observed by MALDI-TOF analysis (Figure 1, a and b). A conversion rate of 83% was quantitatively determined by LC−ESI-Orbitrap MS analysis (Table 6 and Figure S3, Supporting Information).[28] Unmodified myoglobin that having no aldehyde functionality was employed as control and no product was observed. The site-specific modification was further confirmed by the trypsin digestion. LC−MS/MS analysis of the resulting peptide fragments showed that the aldol reaction occurred site-specifically with aldehyde at the N-terminal of myoglobin (Figure S4, Supporting Information).

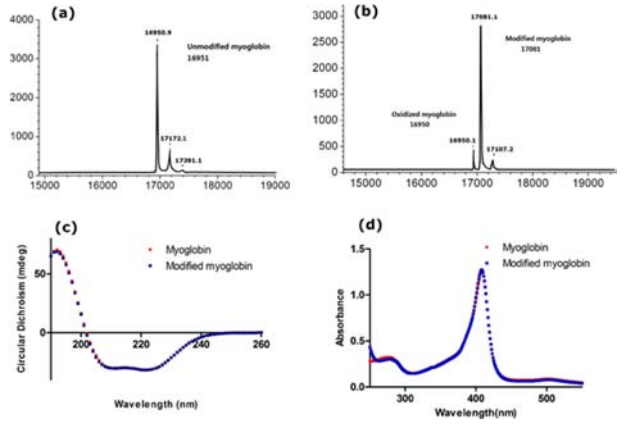
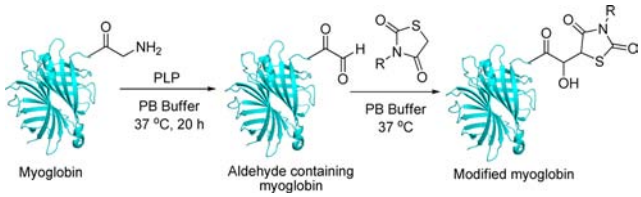


Figure 1. MALDI-TOF analysis of unmodified and modified myoglobin: (a) unmodified myoglobin; (b) modified myoglobin by **3a**; (c and d) CD and UV/vis spectra of unmodified and modified myoglobin.

Table 6. Site-Specific Modification of Aldehyde Containing Myoglobin



entry	R	conv[a] (%)
1	**a**	83
2	**g**	75
3	**i**	81
4	**k**	80
5	**m**	66

[a]Calculated by LC−ESI-Orbitrap MS.

To expand the scope of substrates on protein modification, 2,4-thiazolinediones (**3g**,**i**,**k**,**m**) carrying an alkyne, azide, bromide, and short polyethylene glycol chain, respectively, were subjected to the aldehyde containing myoglobin following the above method. The solvent was changed to PB buffer/*t*-BuOH (5:1) to improve the solubility of **3**.[20] All functionalized myoglobins were detected by MALDI-TOF analysis (Figure S2, Supporting Information), and the conversions were quantitatively determined by the LC−ESI-Orbitrap MS analysis (Table 6 and Figure S3, Supporting Information). The results showed that the aldehyde-containing myoglobins were chemically modified with yields ranging from 66 to 81% under the mild conditions. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis provided further confirmation of the modification (Figure S5, Supporting Information).

The comparison of circular dichroism (CD) and UV/vis spectra of the native and functionalized myoglobin with **3a** by the direct aldol reaction showed little change, which provided evidence that the structure/conformation of protein was not changed in the process of modification (Figure 1c,d). Moreover, the protein function of storing and releasing oxygen was tested under both the oxidation and reduction conditions by the visible spectra (450−700 nm) (Figure S6, Supporting Information). The results indicated that its function was still retained with the modification by the aldol reaction. In

DOI: 10.1021/acs.orglett.5b00005
Org. Lett. 2015, 17, 1361−1364

addition, 2,4-thiazolinedione (**3a**) was proved to be relatively safe to 3T3 and LO2 cells even at high concentration (Table S2, Supporting Information), which indicated that the aldol ligation has potential for application inside the cells.

In conclusion, we have described the first direct aldol reaction that could be applied for site-specific chemical modification of an aldehyde-containing protein. N-Substituted 2,4-thiazolidinediones could selectively form a stable C−C bond with an aldehyde-containing protein under moderate conditions and thereby introduce diverse functional groups including azide, olefin, and bromide into protein without changing its native conformation. This direct aldol ligation may offer a potential bioorthogonal reaction for functionalization of protein.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details and characterization data for all new compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: shanglq@nankai.edu.cn.
*E-mail: zheng_yin@nankai.edu.cn.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China (973 program, Grant No. 2013CB911104), the National Natural Science Foundation of China (Grant No. 21202087), the Tianjin Science and Technology Program (Grant No. 13JCYBJC24300, 13JCQNJC13100), and the "111" Project of the Ministry of Education of China (Project No. B06005). We sincerely thank Ms. Zhongyuan Sun of Protein Chemistry Facility, Center of Biomedical Analysis, Tsinghua University, for the service of the trypsin digestion analysis and Dr. Yaxin Lu, Ms. Man Wang of Nankai University, and Dr. Peiming Gu of Thermo Fisher Scientific for the service of the LC−ESI-Orbitrap MS analysis.

■ REFERENCES

(1) (a) Keppler, A.; Pick, H.; Arrivoli, C.; Vogel, H.; Johnsson, K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9955−9959. (b) Yin, J.; Liu, F.; Li, X.; Walsh, C. T. *J. Am. Chem. Soc.* **2004**, *126*, 7754−7755. (c) Borrmann, A.; van Hest, J. C. M. *Chem. Sci.* **2014**, *5*, 2123−2134.

(2) Sigala, P. A.; Fafarman, A. T.; Bogard, P. E.; Boxer, S. G.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129*, 12104−12105.

(3) Heller, A.; Feldman, B. *Chem. Rev.* **2008**, *108*, 2482−505.

(4) (a) Parker, D. *Chem. Soc. Rev.* **1990**, *19*, 271−291. (b) Wang, Z.; Ding, X.; Li, S.; Shi, J.; Li, Y. *RSC Adv.* **2014**, *4*, 7235.

(5) (a) Hinds, K. D.; Kim, S. W. *Adv. Drug Delivery Rev.* **2002**, *54*, 505−530. (b) Li, Y.; Yang, M.; Huang, Y.; Song, X.; Liu, L.; Chen, P. R. *Chem. Sci.* **2012**, *3*, 2766−2770.

(6) Astronomo, R. D.; Lee, H.-K.; Scanlan, C. N.; Pantophlet, R.; Huang, C.-Y.; Wilson, I. A.; Blixt, O.; Dwek, R. A.; Wong, C.-H.; Burton, D. R. *J. Virol.* **2008**, *82*, 6359−6368.

(7) (a) Jentoft, N.; Dearborn, D. G. *J. Biol. Chem.* **1979**, 254. (b) McFarland, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2005**, *127* (39), 13490−13491. (c) van Maarseveen, J. H.; Reek, J. N. H.; Back, J. W. *Angew. Chem., Int. Ed.* **2006**, *45*, 1841−1843.

(8) (a) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052−5053. (b) Dondoni, A. *Angew. Chem.* **2008**, *120*, 9133−9135. (c) Kim, Y.; Ho, S. O.; Gassman, N. R.; Korlann, Y.; Landorf, E. V.; Collart, F. R.; Weiss, S. *Bioconjugate Chem.* **2008**, *19*, 786−791.

(9) (a) Schlick, T. L.; Ding, Z.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2005**, *127*, 3718−3723. (b) Tilley, S. D.; Francis, M. B. *J. Am. Chem. Soc.* **2006**, *128*, 1080−1081. (c) Kodadek, T.; Duroux-Richard, I.; Bonnafous, J.-C. *Trends Pharmacol. Sci.* **2005**, *26*, 210−217.

(10) Gauthier, M. A.; Klok, H.-A. *Chem. Commun.* **2008**, 2591−2611.

(11) Humenik, M.; Huang, Y.; Wang, Y.; Sprinzl, M. *ChemBioChem.* **2007**, *8*, 1103−1106.

(12) Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. *ACS Chem. Biol.* **2014**, *9*, 592−605.

(13) (a) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. *Nat. Chem. Biol.* **2007**, *3*, 321−2. (b) Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3000−3005.

(14) Gilmore, J. M.; Scheck, R. A.; Esser-Kahn, A. P.; Joshi, N. S.; Francis, M. B. *Angew. Chem., Int. Ed.* **2006**, *45*, 5307−5311.

(15) (a) Esser-Kahn, A. P.; Francis, M. B. *Angew. Chem., Int. Ed.* **2008**, *47*, 3751−3754. (b) Scheck, R. A.; Dedeo, M. T.; Iavarone, A. T.; Francis, M. B. *J. Am. Chem. Soc.* **2008**, *130*, 11762−11770.

(16) (a) Gaertner, H. F.; Rose, K.; Cotton, R.; Timms, D.; Camble, R.; Offord, R. E. *Bioconjugate Chem.* **1992**, *3*, 262−268. (b) Mikolajczyk, S. D.; Meyer, D. L.; Starling, J. J.; Law, K. L.; Rose, K.; Dufour, B.; Offord, R. E. *Bioconjugate Chem.* **1994**, *5*, 636−646.

(17) (a) Kalia, J.; Raines, R. T. *Angew. Chem., Int. Ed.* **2008**, *47*, 7523−7526. (b) Crisalli, P.; Kool, E. T. *J. Org. Chem.* **2013**, *78*, 1184−1189. (c) Rashidian, M.; Mahmoodi, M. M.; Shah, R.; Dozier, J. K.; Wagner, C. R.; Distefano, M. D. *Bioconjugate Chem.* **2013**, *24*, 333−342. (d) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7581−7584.

(18) (a) Sasaki, T.; Kodama, K.; Suzuki, H.; Fukuzawa, S.; Tachibana, K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4550−4553. (b) Agarwal, P.; van der Weijden, J.; Sletten, E. M.; Rabuka, D.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 46−51.

(19) Han, M.-J.; Xiong, D.-C.; Ye, X.-S. *Chem. Commun.* **2012**, *48*, 11079−11081.

(20) Alam, J.; Keller, T. H.; Loh, T. P. *J. Am. Chem. Soc.* **2010**, *132*, 9546−9548.

(21) Mase, N.; Nakai, Y.; Ohara, N.; Yoda, H.; Takabe, K.; Tanaka, F.; Barbas, C. F. *J. Am. Chem. Soc.* **2006**, *128*, 734−735.

(22) Paladhi, S.; Chauhan, A.; Dhara, K.; Tiwari, A. K.; Dash, J. *Green Chem.* **2012**, *14*, 2990−2995.

(23) Breslow, R.; Rizzo, C. J. *J. Am. Chem. Soc.* **1991**, *113*, 4340−4341.

(24) Hayashi, Y.; Aratake, S.; Itoh, T.; Okano, T.; Sumiya, T.; Shoji, M. *Chem. Commun. (Cambridge)* **2007**, 957−959.

(25) Alam, J.; Keller, T. H.; Loh, T.-P. *J. Am. Chem. Soc.* **2010**, *132*, 9546−9548.

(26) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518−13519.

(27) Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 18638−18643.

(28) Ruan, Q.; Ji, Q.; Arnold, M. E.; Humphreys, W. G.; Zhu, M. *Anal. Chem.* **2011**, *83*, 8937−8944.

DOI: 10.1021/acs.orglett.5b00005
Org. Lett. 2015, 17, 1361−1364