Selective Derivatization of *N*-Terminal Cysteines Using Cyclopentenediones

Omar Brun,[†] Jordi Agramunt,[†] Lluis Raich,[†,‡] Carme Rovira,[†,‡,||] Enrique Pedroso,[†,§] and Anna Grandas*,[†,§]

[†]Departament de Química Inorgànica i Orgànica, [‡]secció Q. Orgànica, IQTCUB, and [§]secció Q. Orgànica, IBUB), Facultat de Química, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain

[||]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluis Companys 23, 08010 Barcelona, Spain

S *Supporting Information*
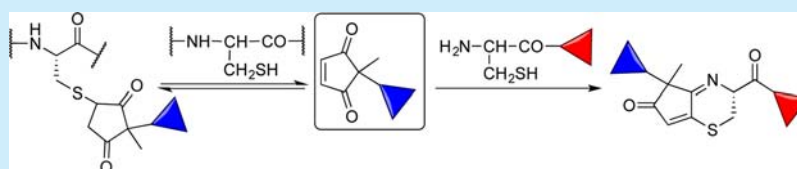
ABSTRACT: The outcome of the Michael-type reaction between thiols and 2,2-disubstituted cyclopentenediones varies depending on the thiol. Stable compounds with two fused rings were formed upon reaction with 1,2-aminothiols (such as *N*-terminal cysteines in peptides). Other thiols gave reversibly Michael-type adducts that were in equilibrium with the starting materials. This differential reactivity allows differently placed cysteines to be distinguished and has been exploited to prepare bioconjugates incorporating two or three different moieties.

The Michael-type addition of a thiol to a maleimide is the thiol-involving click reaction most frequently exploited for bioconjugation.[1] It is water compatible, quick, and clean and takes place in high yield without the need of any catalyst. However, it has recently been found that the resulting Michael-type adducts (MTAs) may undergo thiol exchange and hydrolysis to an extent depending on both the nature of the maleimide and the environment.[2] Thiol exchange may cause premature loss of payload from therapeutically relevant compounds such as antibody−drug conjugates,[3] and as a result of hydrolysis, conjugates may lose their structural integrity and become a mixture of products.

Since thiols are natural components of peptides and proteins, several groups are actively searching for new thiol-involving conjugation methods providing stable linkages.[4]

2,2-Disubstituted cyclopent-4-ene-1,3-diones (CPDs), which formally differ from *N*-substituted maleimides in that the *N*-alkyl group is replaced by a dialkylated carbon, hold promise for conjugation purposes because the MTAs resulting from reaction with thiols cannot be hydrolyzed. Although the homologue of acid **2** (see below) and *N*-acetylcysteine reacted to give a stable MTA as a mixture of four stereoisomers,[5] 2,2,4,5-tetramethyl-cyclopent-4-ene-1,3-dione failed to form thiol adducts with *N*-acetylcysteamine.[6]

Here, we investigate the reaction of CPDs with cysteine-containing compounds and their potential in the preparation of bioconjugates. We have found that the outcome of the Cys−CPD reaction varies depending on the position of the cysteine residue within peptide (or, in general, polyamide) chains.

Four CPDs differing in one of the C2 substituents (**1−4**, Figure 1) were prepared (see the Supporting Information, section 2).
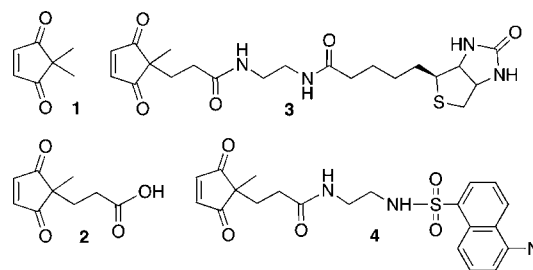

**Figure 1.** Synthesized CPDs.

First experiments aimed at examining the effect of the cysteine position within peptide chains on the Cys−CPD reaction. Therefore, peptides H-WGRGC-NH₂ (**5**), H-KYAYCG-NH₂ (**6**), and H-CYG-NH₂ (**7**) were reacted with CPDs **1** or **2** (water, 37 °C, Scheme 1 and Supporting Information, sections 3 and 4), and the reaction progress was monitored by HPLC.

In reactions involving **5** and **6**, it was observed that peptides were not fully consumed in spite of the presence of an excess of CPD (5 equiv) in the mixture and that the amount of the MTAs in the crude decreased as the unreacted peptide underwent oxidation to the disulfide dimer. Isolated MTAs

Scheme 1. Reaction of CPDs with Different Thiols
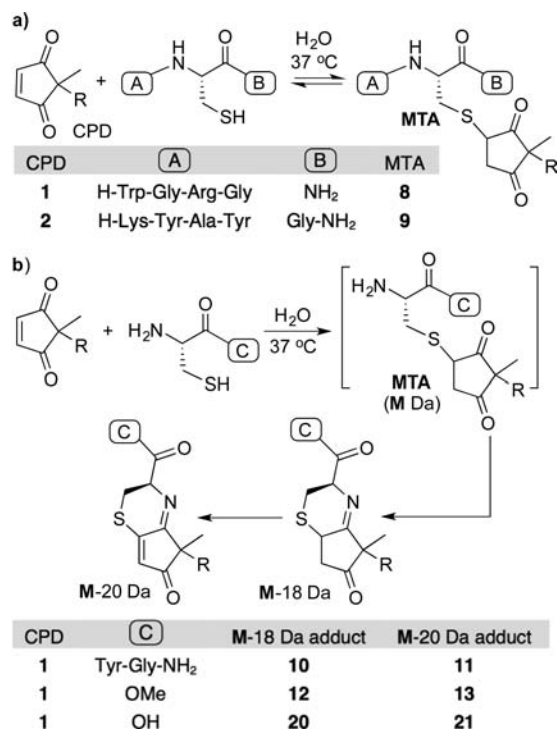
(**8** and **9**, respectively) were found to revert to the parent CPD and peptide after 60 min incubation in water. These results differ from previously described ones[5] and from formation of the MTA upon reacting *N*-acetylcysteine and 2,4,5-tetramethylcyclopent-4-ene-1,3-dione (though in 15% yield[6]). In contrast to their maleimide counterparts,[7] Michael additions to cyclopent-4-ene-1,3-diones may be reversible.

DFT calculations showed that maleimide−thiol MTAs have a greater stability than CPD−thiol MTAs (Supporting Information, sections 12 and 13), which is in agreement with the equilibria observed in the experiments reported here.

The mass corresponding to the Michael-type adduct of **7** + **1** (**M** Da, Scheme 1b) was not observed. Instead, when the reaction was monitored by HPLC/MS it was observed that two new products with masses of 18 and 20 Da lower were formed and that the one with the **M**−18 Da mass (**10**) evolved into the final, stable compound whose mass was **M**−20 Da (**11**). We thus surmised that the Michael-type adduct very quickly underwent intramolecular imine formation to generate a 6-membered ring, which would account for the presence of the **M**−18 Da compound **10** in the reaction mixture. Subsequent loss of two H atoms would render the oxidized **M**−20 Da adduct **11**.[8]

Intramolecular cyclization to form an imine-containing six-membered ring is not unknown and occurs when cysteine reacts with α-bromo ketones.[9] In our case, the presence of the carbonyl group in the five-membered ring likely favors the unprecedented subsequent oxidation, which yields the CO-C=C-C≡N conjugated system. The UV−vis spectrum of the **M**−20 Da adducts exhibits a maximum around 330 nm, which is consistent with such a conjugated structure.

To confirm the structure of the **M**−20 Da adduct, CPD **1** was reacted with methyl cysteinate (which mimicked a peptide containing an *N*-terminal cysteine), and the final adduct (**13**, Scheme 1b) was isolated and characterized by NMR (¹H, ¹³C, gHSQC) and electrospray HRMS. This proved that **13** is a

4,7,7-trisubstituted 2-thia-5-azabicyclo[4.3.0]nona-1(9),5-diene-8-one (see the Supporting Information, section 9.3, for an assessment of the stability of these adducts).

Methyl cysteinate was also reacted with CPDs **2−4** (Supporting Information, section 5). The evolution of the reaction was the same as with CPD **1**. Initially formed **M**−18 Da adducts (**14**, **16**, and **18**) evolved into the **M**−20 Da ones (**15**, **17**, and **19**, respectively), which proves that the reaction outcome is independent of the CPD substituents at the 2 position.
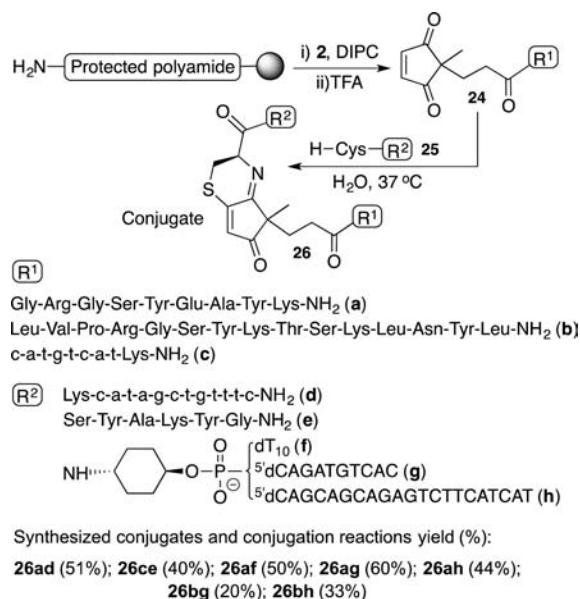
In all previous experiments, we had perceived that the **M**−18 Da to **M**−20 Da adduct oxidation was the rate-limiting step. Upon reacting methyl cysteinate and **2** under different conditions (Supporting Information, section 6), it was found that the **14** to **15** conversion rate increased with temperature and decreased when water was replaced by 1:1 mixtures of water and an organic cosolvent (the conversion rate was higher using dimethyl sulfoxide than using methanol or acetonitrile). It was also found that use of an excess of CPD increased the **14** to **15** conversion rate, but not the use of an excess of the thiol component.

In the case of the **1** + cysteine reaction, the **M**−20 Da adduct (**21**) underwent reversible hydrolysis of the imine to yield an **M**−2 Da zwitterionic compound (**22**, see Supporting Information, section 7).

Homocysteine (Hcy) was also included in this study because of its biological relevance in cardiovascular and neurological pathologies.[10] It has been reported that Cys and Hcy can react with the same compound, generally at different rates, to generate homologous structures differing in only one methylene.[4a,11] However, as in reactions involving *N*-acylated 1,2-aminothiols, the **M**−20 Da and **M**−18 Da adducts were not detected, and only the MTA (**23**) was formed (Supporting Information, section 8). The isolated MTA reverted to the parent compounds. DFT calculations (Supporting Information, section 12) confirmed that the energy barriers of the first two steps leading to imine formation (7-membered ring) are higher for Hcy than for H-Cys-OH or H-Cys-OMe. This discrimination provides a basis on which separation and/or detection systems could be devised.

At this point, it was time to assess whether the Cys−CPD reaction could be of use to prepare conjugates. It was first verified that CPD **2** can be attached to resin-linked peptide and peptide nucleic acid (PNA) chains and the corresponding CPD-polyamides **24a−c** be obtained after the trifluoroacetic acid mediated deprotection reaction (Scheme 2). Then, peptide **24a** was reacted with the cysteine-derivatized PNA **25d** and PNA **24c** with peptide **25e**. Peptide−oligonucleotide conjugates were also synthesized by reacting peptides **24a** and **24b** with cysteine-derivatized oligonucleotides (**25f−h**) prepared following described procedures.[13] In all cases, the target conjugates **26** were satisfactorily obtained (2−5 h reaction times).

Since the outcome of the reaction between CPDs and cysteines differs depending on whether the amine is free or acylated, we decided to investigate the possibility of selectively labeling *N*-terminal cysteines in the presence of differently placed cysteines, which is not straightforward. Aldehydes[14] and 2-cyanobenzothiazoles[4a] are among the few reagents that react selectively with *N*-terminal cysteines. In the other cases, even though it has been shown that local environment may affect cysteine reactivity and render them chemically distinguish-

Scheme 2. Synthesis of Cys−CPD Conjugates[12]

Synthesized conjugates and conjugation reactions yield (%):
26ad (51%); **26ce** (40%); **26af** (50%); **26ag** (60%); **26ah** (44%);
26bg (20%); **26bh** (33%)

able,[15] regioselectivity is typically difficult to achieve unless cysteines are orthogonally protected.

In a preliminary experiment, CPD **2** was incubated with a 1:1 mixture of peptides **6** (H-KYAYCG-NH$_2$) and H-CSYAKYG-NH$_2$ (**25e**) (Supporting Information, section 10). As expected from previous experiments, HPLC/MS analysis of the reaction mixture showed that **25e** had reacted completely (yielding **M**−20 Da adducts **27**), whereas **6** had not at all. This result confirmed that *N*-terminal cysteines could be selectively labeled in the presence of peptides with cysteines at internal positions.

Next, peptide H-CSYACKYG-NH$_2$ (**28**) was reacted with two model alkylating agents, namely CPD **2** and 3-maleimidopropanoic acid (Scheme 3). The peptide was first

Scheme 3. Double-Derivatization Experiments

incubated with **2** (1.1 equiv). After 90 min reaction time, 3-maleimidopropanoic acid (5 equiv) was added. HPLC/MS analysis of the crude 1 h later showed the presence of two new peaks (see Figure S38), whose mass fit with that of the doubly derivatized peptide (**29a**). A similar procedure was used to prepare conjugates **29b** and **29c** (Scheme 3), assembling three different moieties, namely a peptide (H-CWGRGC-NH$_2$, **30**), a PNA (**24c** or **31**), and a tagging agent (either biotin or a fluorophore). These results, as well as those summarized in Scheme 2, prove that formation of the **M**−20 Da adduct takes place irrespective of the moiety linked to the cysteine carboxyl group.

To sum up, this work has revealed that CPDs react differently from maleimides, even though CPDs can be formally considered maleimide mimics. Strikingly, the outcome of the reaction between CPDs and cysteines whose amine is free (1,2-aminothiols) differs from that of the reactions involving *N*-acylated cysteines and Hcy. Only the former react irreversibly with CPDs to give water-stable bicyclic nonadienone adducts that can be exploited for bioconjugation. In the other cases, the Michael-type adducts (MTAs) revert to the starting materials. The different behavior of maleimides and cyclopentenediones is reproduced by DFT calculations.

CPDs also distinguish *N*-terminal cysteines from cysteines placed at different positions, thus pointing to new alternatives for the regioselective derivatization of peptides and proteins, to which *N*-terminal cysteines can be added for conjugation purposes.[4a,16] The reaction between CPDs and cysteines with a free amine has been successfully exploited to prepare a variety of conjugates, both incorporating two and three different components.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.6b02301.

> Experimental procedures, compound characterization data, spectra, HPLC profiles, quantum chemical calculations and orbital analysis, and Cartesian coordinates (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: anna.grandas@ub.edu.

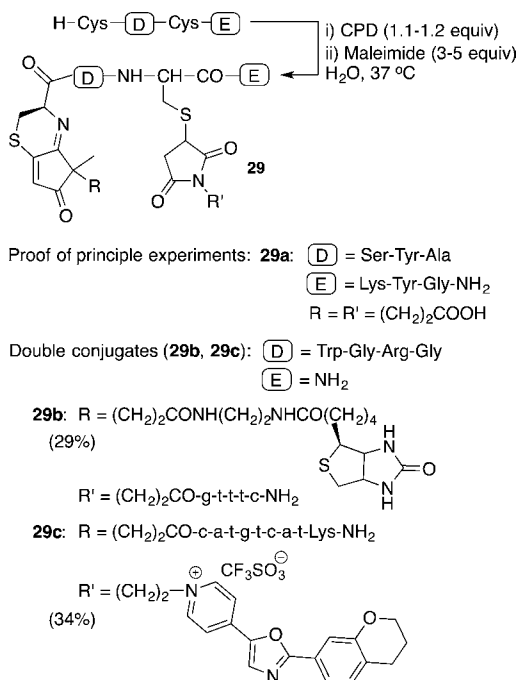**Notes**

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) (a) Goodchild, J. *Bioconjugate Chem.* **1990**, *1*, 165−187. (b) Venkatesan, N.; Kim, B. H. *Chem. Rev.* **2006**, *106*, 3712−3761. (c) Kalia, J.; Raines, R. T. *Curr. Org. Chem.* **2010**, *14*, 138−147. (d) Lu,

K.; Duan, Q.-P.; Ma, L.; Zhao, D.-X. *Bioconjugate Chem.* **2010**, *21*, 187−202. (e) Singh, Y.; Murat, P.; Defrancq, E. *Chem. Soc. Rev.* **2010**, *39*, 2054−2070. (f) Juliano, R. L.; Ming, X.; Nakagawa, O. *Acc. Chem. Res.* **2012**, *45*, 1067−1076. (g) Tang, W.; Becker, M. L. *Chem. Soc. Rev.* **2014**, *43*, 7013−7039.

(2) (a) Lewis, M. R.; Shively, J. E. *Bioconjugate Chem.* **1998**, *9*, 72−86. (b) Alley, S. C.; Benjamin, D. R.; Jeffrey, S. C.; Okeley, N. M.; Meyer, D. L.; Sanderson, R. J.; Senter, P. D. *Bioconjugate Chem.* **2008**, *19*, 759−765. (c) Ryan, C. P.; Smith, M. E. B.; Schumacher, F. F.; Grohmann, D.; Papaioannou, D.; Waksman, G.; Werner, F.; Baker, J. R.; Caddick, S. *Chem. Commun.* **2011**, *47*, 5452−5454. (d) Baldwin, A. D.; Kiick, K. L. *Bioconjugate Chem.* **2011**, *22*, 1946−1953. (e) Shen, B.-Q.; et al. *Nat. Biotechnol.* **2012**, *30*, 184−189. (f) Strop, P.; Liu, S.-H.; Dorywalska, M.; Delaria, K.; Dushin, R. G.; Tran, T.-T.; Ho, W.-H.; Farias, S.; Casas, M. G.; Abdiche, Y.; Zhou, D.; Chandrasekaran, R.; Samain, C.; Loo, C.; Rossi, A.; Rickert, M.; Krimm, S.; Wong, T.; Chin, S. M.; Yu, J.; Dilley, J.; Chaparro-Riggers, J.; Filzen, G. F.; O'Donnell, C. J.; Wang, F.; Myers, J. S.; Pons, J.; Shelton, D. L.; Rajpal, A. *Chem. Biol.* **2013**, *20*, 161−167.

(3) (a) Ornes, S. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 13695−13695. (b) Cal, P. M. S. D.; Bernardes, G. J. L.; Gois, P. M. P. *Angew. Chem., Int. Ed.* **2014**, *53*, 10585−10587.

(4) (a) Ren, H.; Xiao, F.; Zhan, K.; Kim, Y.-P.; Xie, H.; Xia, Z.; Rao, J. *Angew. Chem., Int. Ed.* **2009**, *48*, 9658−9662. (b) Spokoyny, A. M.; Zou, Y.; Ling, J. J.; Yu, H.; Lin, Y.-S.; Pentelute, B. L. *J. Am. Chem. Soc.* **2013**, *135*, 5946−5949. (c) Toda, N.; Asano, S.; Barbas, C. F., III *Angew. Chem., Int. Ed.* **2013**, *52*, 12592−12596. (d) Zhang, C.; Spokoyny, A. M.; Zou, Y.; Simon, M. D.; Pentelute, B. L. *Angew. Chem., Int. Ed.* **2013**, *52*, 14001−14005. (e) Badescu, G.; et al. *Bioconjugate Chem.* **2014**, *25*, 460−469. (f) Abbas, A.; Xing, B.; Loh, T.-P. *Angew. Chem., Int. Ed.* **2014**, *53*, 7491−7494. (g) Maruani, A.; Alom, S.; Canavelli, P.; Lee, M. T. W.; Morgan, R. E.; Chudasama, V.; Caddick, S. *Chem. Commun.* **2015**, *51*, 5279−5282. (h) Moody, P.; Chudasama, V.; Nathani, R. I.; Maruani, A.; Martin, S.; Smith, M. E. B.; Caddick, S. *Chem. Commun.* **2014**, *50*, 4898−4900. (i) Brown, S. P.; Smith, A. B., III *J. Am. Chem. Soc.* **2015**, *137*, 4034−4037. (j) Kalia, D.; Malekar, P. V.; Parthasarathy, M. *Angew. Chem., Int. Ed.* **2016**, *55*, 1432−1435. (k) Bandyopadhyay, A.; Cambray, S.; Gao, J. *Chem. Sci.* **2016**, *7*, 4589−4593.

(5) Billington, S.; Mann, J.; Quazi, P.; Alexander, R.; Eaton, M. A. W.; Millar, K.; Millican, A. *Tetrahedron* **1991**, *47*, 5231−5236.

(6) McComas, C. C.; Perales, J. B.; Van Vranken, D. L. *Org. Lett.* **2002**, *4*, 2337−2340.

(7) (a) Hoyle, C. E.; Lowe, A. B.; Bowman, C. N. *Chem. Soc. Rev.* **2010**, *39*, 1355−1387. (b) Nguyen, L.-T. T.; Gokmen, M. T.; Du Prez, F. E. *Polym. Chem.* **2013**, *4*, 5527−5536.

(8) Formation of **M**−18 Da adducts was confirmed by ESI MS analysis. When the **M**−18 Da HPLC peaks were collected and the products analyzed by MALDI-TOF MS, the mass found was **M**−20 Da. This suggests that oxidation had taken place either during the isolation process or upon MALDI-TOF MS analysis, or both.

(9) He, L.; Xu, Q.; Liu, Y.; Wei, H.; Tang, Y.; Lin, W. *ACS Appl. Mater. Interfaces* **2015**, *7*, 12809−12813.

(10) (a) Refsum, H.; Ueland, P. M.; Nygård, O.; Vollset, S. E. *Annu. Rev. Med.* **1998**, *49*, 31−62. (b) Seshadri, S.; Beiser, A.; Selhub, J.; Jacques, P. F.; Rosenberg, I. H.; D'Agostino, R. B.; Wilson, P. W. F.; Wolf, P. A. *N. Engl. J. Med.* **2002**, *346*, 476−483.

(11) (a) Rusin, O.; St. Luce, N. N.; Agbaria, R. A.; Escobedo, J. O.; Jiang, S.; Warner, I. M.; Dawan, F. B.; Lian, K.; Strongin, R. M. *J. Am. Chem. Soc.* **2004**, *126*, 438−439. (b) Zhang, X.; Ren, X.; Xu, Q.-H.; Loh, K. P.; Chen, Z.-K. *Org. Lett.* **2009**, *11*, 1257−1260. (c) Xiong, L.; Zhao, Q.; Chen, H.; Wu, Y.; Dong, Z.; Zhou, Z.; Li, F. *Inorg. Chem.* **2010**, *49*, 6402−6408. (d) Zhang, R.; Yu, X.; Ye, Z.; Wang, G.; Zhang, W.; Yuan, J. *Inorg. Chem.* **2010**, *49*, 7898−7903. (e) Niu, L.-Y.; Guan, Y.-S.; Chen, Y.-Z.; Wu, L.-Z.; Tung, C.-H.; Yang, Q.-Z. *J. Am. Chem. Soc.* **2012**, *134*, 18928−18931. (f) Jung, H. S.; Chen, X.; Kim, J. S.; Yoon, J. *Chem. Soc. Rev.* **2013**, *42*, 6019−6031. (g) Liu, J.; Sun, Y.-Q.; Huo, Y.; Zhang, H.; Wang, L.; Zhang, P.; Song, D.; Shi, Y.; Guo, W. *J. Am. Chem. Soc.* **2014**, *136*, 574−577. (h) Liu, Y.; Lv, X.; Liu, J.; Sun, Y.-Q.; Guo, W. *Chem. - Eur. J.* **2015**, *21*, 4747−4754. (i) Zhang, H.; Liu, R.; Liu, J.; Li, L.; Wang, P.; Yao, S. Q.; Xu, Z.; Sun, H. *Chem. Sci.* **2016**, *7*, 256−260.

(12) Lower case letters a, c, g, and t refer to PNA monomers. Since in this scheme amino acids are depicted by the three-letter code, here A, C, G, and T mean 2′-deoxyadenosine, 2′-deoxycytidine, 2′-deoxyguanosine, and 2′-deoxythymidine.

(13) Stetsenko, D. A.; Gait, M. J. *J. Org. Chem.* **2000**, *65*, 4900−4908.

(14) (a) Zhang, L.; Tam, J. P. *Anal. Biochem.* **1996**, *233*, 87−93. (b) Botti, P.; Pallin, T. D.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 10018−10024.

(15) (a) Nathani, R. I.; Moody, P.; Chudasama, V.; Smith, M. E. B.; Fitzmaurice, R. J.; Caddick, S. *Chem. Sci.* **2013**, *4*, 3455−3458. (b) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. *Nat. Chem.* **2016**, *8*, 120−128.

(16) (a) Tolbert, T. J.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2002**, *41*, 2171−2174. (b) Gentle, I. E.; De Souza, D. P.; Baca, M. *Bioconjugate Chem.* **2004**, *15*, 658−663. (c) Casi, G.; Huguenin-Dezot, N.; Zuberbühler, K.; Scheuermann, J.; Neri, D. *J. Am. Chem. Soc.* **2012**, *134*, 5887−5892. (d) Wang, H.-C.; Yu, C.-C.; Liang, C.-F.; Huang, L.-D.; Hwu, J.-R.; Lin, C.-C. *ChemBioChem* **2014**, *15*, 829−835.