JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Rapid Chemoselective Bioconjugation through Oxidative Coupling of Anilines and Aminophenols

Christopher R. Behrens,[†] Jacob M. Hooker,[‡] Allie C. Obermeyer,[†] Dante W. Romanini,[†] Elan M. Katz,[†] and Matthew B. Francis*[,†,§]

[†]Department of Chemistry, University of California, Berkeley, California 94720-1460, United States

[‡]Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts 02129, United States

[§]Materials Sciences Division, Lawrence Berkeley National Laboratories, Berkeley, California 94720-1460, United States

Supporting Information

ABSTRACT: A highly efficient protein bioconjugation method is described involving addition of anilines to *o*-aminophenols in the presence of sodium periodate. The reaction takes place in aqueous buffer at pH 6.5 and can reach high conversion in 2−5 min. The major product was characterized using X-ray crystallography, which revealed that an unprecedented oxidative ring contraction occurs after the coupling step. The compatibility of the reaction with protein substrates has been demonstrated through attachment of small molecules, polymer chains, and peptides to *p*-aminophenylalanine residues introduced into viral capsids through amber stop codon suppression. Coupling of anilines to *o*-aminophenol groups derived from tyrosine residues is also described. The compatibility of this method with thiol modification chemistry is shown through attachment of a near-IR fluorescent chromophore to cysteine residues inside the viral capsid shells, followed by attachment of integrin-targeting RGD peptides to anilines on the exterior surface.

Chemical modification of proteins[1] is crucial for the study of biochemical function[2] and scalable synthesis of biomolecular materials,[3] targeted imaging agents,[4] and protein−drug conjugates.[5] As applications of bioconjugates become ever more complex, an expanded set of chemical strategies is needed to add new functionality to specific locations with high chemoselectivity and yield. In most cases to date, site-specific protein labeling has been achieved by targeting cysteine residues,[1] taking advantage of the highly nucleophilic character of thiolate anions and their particularly low abundance on the surfaces of most proteins. Useful as this chemistry is, however, many circumstances still require additional chemical reactions to install a second set of functional groups or to avoid alkylating native cysteine groups that are required for protein function. We have previously used cysteine-based strategies to install drug[6] and imaging cargo[7] inside genome-free viral capsids for applications in targeted delivery. More difficult, however, is the installation of peptides[8] and nucleic acid aptamers[9] on the external surface that can bind to specific receptors on tumor tissue. We and others[10] have found that "bioorthogonal" protein modification reactions[11] are particularly useful in these situations, targeting ketone,[12] azide,[13] alkyne,[14] and aniline[15] groups specifically while ignoring native protein functionality. The functional groups required for these reactions can be introduced as unnatural amino acids[16] or installed using more traditional bioconjugation reactions. However, many existing reactions proceed at low coupling rates or require high

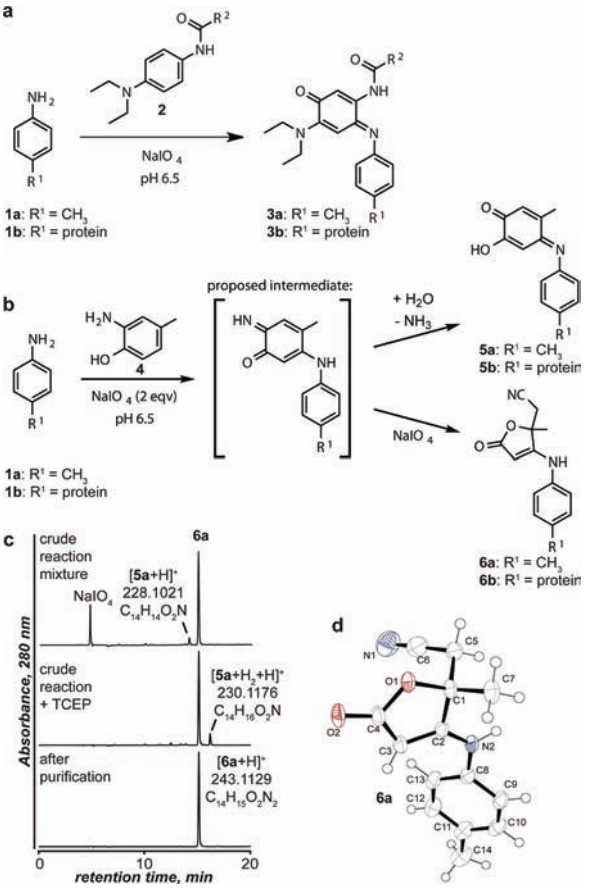

Figure 1. Oxidative coupling reactions involving aniline groups on proteins. (a) The previously reported coupling between anilines and phenylene diamines occurs within 30−60 min. (b) The coupling between anilines and aminophenols occurs much faster, reaching high levels of conversion in <2 min. (c) An unequal mixture of two products (**5a** and **6a**) is obtained, as indicated by reversed-phase HPLC analysis. (d) Product **6a** was characterized using X-ray diffraction.

Received: April 21, 2011
Published: September 15, 2011

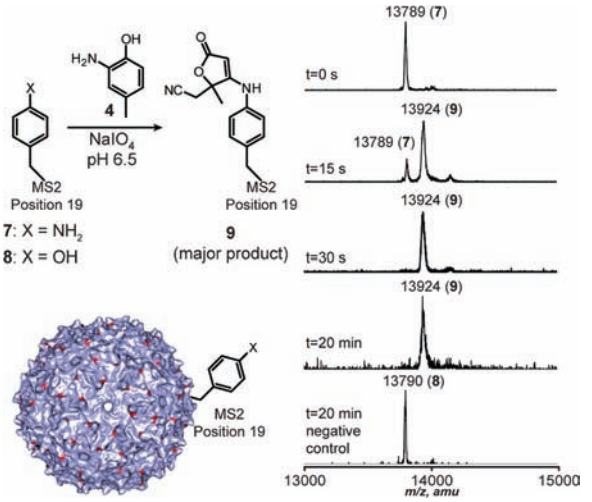

Figure 2. Targeting an artificial amino acid through oxidative coupling. T19*p*AF MS2 (30 μM) was reacted with 2-amino-4-methylphenol (100 μM) in the presence of sodium periodate (1 mM). Aliquots were removed at various time points and quenched by the addition of tris(2-carboxyethyl)phosphine. Single protein modification was complete by 30 s. High-resolution ESI-MS analysis revealed that the major product was **9**, in addition to a minor amount of the addition product corresponding to **5b**. These signals do not resolve using MALDI-MS (see SI Figure S10 for ESI spectra). No additional modification was detected after 20 min. A negative control reaction using T19Y MS2 showed no modification under otherwise identical conditions.

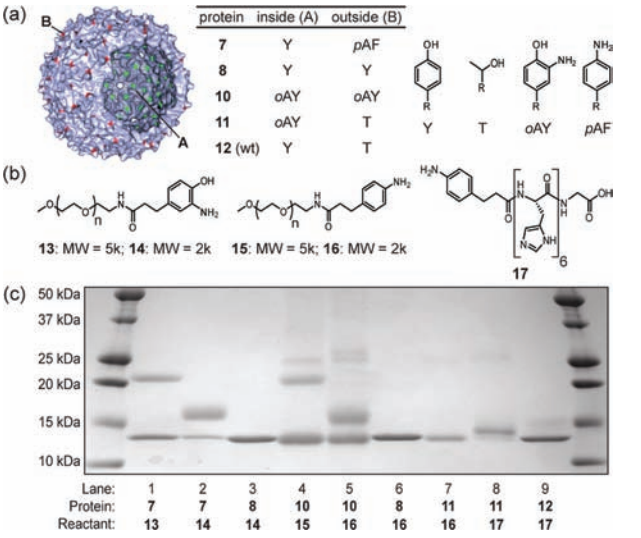

Figure 3. A panel of reactants was exposed to the MS2 capsids listed in (a), which presented phenol, *o*-aminophenol (*o*AY), or aniline (*p*AF) groups in varying combinations. Each reaction contained 30 μM protein, a coupling partner selected from (b) at 200−500 μM, and $NaIO_4$ (1 mM for lanes 1−7, 5 mM for lanes 8 and 9). Each reaction time was 2 min, after which the samples were quenched with loading buffer and analyzed via SDS-PAGE (c).

concentrations of reactants that are often in short supply. The growing complexity of modern protein bioconjugates creates a need for additional reactions that can proceed with increased efficiency and yield.

To provide new opportunities for site-selective protein bioconjugation with highly functionalized substrates, we previously explored oxidative couplings between aniline side chains and *N*-acylphenylene diamine groups in the presence of sodium periodate, Figure 1a.[17] This method yields hydrolytically stable "A+B" products with excellent chemoselectivity and has been used to attach peptides,[8] nucleic acids,[9,6b] and porphyrins[18] (as R^2 groups in **2**) to *p*-aminophenylalanine (*p*AF) side chains (**1b**) introduced using amber stop codon suppression.[15] After completing a series of optimization studies, we now report a substantially more efficient version of this reaction involving addition of anilines to *o*-aminophenols under similar oxidizing conditions, Figure 1b. This new reaction proceeds to very high levels of conversion in well under 2 min, yielding stable protein bioconjugates with equivalent chemoselectivity. We introduce the functional groups required for this reaction either by modifying native tyrosine residues or through direct incorporation of artificial amino acids. The high speed of this coupling reaction makes it well-suited for assembling multicomponent structures from building blocks at low concentrations, or for achieving protein modification in situations that are particularly time-sensitive, such as radiolabeling.

The appropriate reaction conditions and product structure of this new coupling method were determined through small-molecule studies. It was found that 2-amino-4-methylphenol (**4**) reacted with 1 equiv of *p*-toluidine (**1a**) in the presence of 10 equiv of sodium periodate[19] in pH 6.5 phosphate buffer, Figure 1b. In <5 min, a major product was obtained in 40% isolated yield and a second product in lesser amounts, Figure 1c. The major product was identified as **6a** using NMR and X-ray diffraction analysis of an obtained crystal. The ^{13}C NMR spectrum showed a clear nitrile signal at 115.6 ppm, and IR analysis showed an absorption band at 2218 cm^{-1}. The X-ray data matched structure **6a** with a goodness-of-fit of 1.05, leaving no doubt as to the nature of the product. We presume that this structure was formed through oxidation of **4** to the *o*-iminoquinone species, although single-electron coupling pathways can also be considered as plausible mechanisms.

Addition of aniline to the 5-position of the iminoquinone, followed by oxidation of this species by a second equivalent of periodate, would yield the proposed intermediate shown in Figure 1b. The formation of **6a** can be rationalized by water addition to the carbonyl, followed by cleavage of the resulting imino alcohol by the periodate. This breaks the six-membered ring and forms both a carboxylic acid and an acrylonitrile group, which react to form the butenolide ring (see Supporting Information (SI) Figure S1 for our current mechanistic hypotheses). The minor reaction product (which cannot be converted into **6a** with additional periodate) was assigned as **5a**, resulting from the competitive hydrolysis of the proposed imine intermediate in Figure 1b. This species can be reduced to the corresponding hydroquinone using tris(2-carboxyethyl)phosphine (TCEP). Although several unidentified dyestuffs accompany the reaction product in trace amounts, the combined "A+B" reaction products (**6a** and **5a**) are obtained with surprisingly high efficiency for such complex reaction pathways. The low yield of isolated product was attributed to incomplete extraction of the compound from the dilute aqueous solution in which it was generated.

The ability of this reaction to reach high conversion under mild conditions and at sub-mM concentrations in aqueous buffer suggested that it would perform similarly well using protein substrates. To evaluate this potential, we introduced aniline groups on the external surface of genome-free MS2 capsids.[8] To do this, we introduced *p*-aminophenylalanine (*p*AF) into position 19 of each protein monomer using an amber stop codon suppressor ^{t}RNA/aminoacyl ^{t}RNA synthetase pair developed in the Schultz laboratory.[15] The small-molecule model reaction in

Figure 1b was then repeated with T19*p*AF MS2 (**7**) as the aniline component. **4** (100 μM) was combined with **7** (30 μM in monomer, 167 nM in capsid) in pH 6.5 phosphate buffer at room temperature. A 50 mM solution of freshly prepared sodium periodate was added to achieve a final concentration of 1 mM. Aliquots were removed at various time points and quenched by addition of TCEP. Subsequent analysis using MALDI-TOF MS (following capsid disassembly during the sample preparation process) indicated complete conversion in 30 s, with no additional modification or degradation occurring during the next 20 min of reaction time, Figure 2. As a negative control, the reaction was repeated using an MS2 mutant bearing a tyrosine at position 19 (T19Y MS2, **8**). This single atom change resulted in no observed product after 20 min of reaction. High-resolution ESI-MS analysis of the product mixture indicated that two adducts with a mass difference of 15 Da were formed (these peaks did not resolve in the MALDI spectra). The mass change corresponding to the major product matched that of **6a**, and the minor product corresponded to **5a** (see SI Figure S10). Thus, similar coupling efficiency and chemoselectivity were observed on the complex protein substrate.

To generate more useful bioconjugates, we used this oxidative coupling reaction to attach larger molecules to MS2 capsids with similar effectiveness. As one example, *o*-aminophenol-terminated poly(ethylene glycol) chains (**13** and **14**) coupled readily to MS2 capsids bearing external *p*AF groups in 2 min, but not to proteins bearing only tyrosines in the same positions (Figure 3c, lanes 1−3). According to optical densitometry of the Coomassie-stained SDS-PAGE gel, ~65% of the *p*AF-containing monomers were modified when reacted with **13**, and 75% were modified by **14**. This corresponded to the installation of 117−135 copies of the PEG chains on each capsid. The decrease in modification percentage with the larger PEG chain is likely due to increased steric congestion as high levels of modification are approached. As an initial test of this reaction on another protein substrate, we have found that aminophenol polymer **14** couples with similar efficiency to an aniline-labeled sample of lysozyme (see SI Figure S8 for details).

Coupling of macromolecules to *o*-aminophenol-containing MS2 was also investigated. In a previous study,[20] we found that aminophenol groups introduced inside MS2 using a diazonium coupling/dithionite reduction procedure could be oxidized in the presence of sodium periodate to yield a species that coupled to *p*-aminophenyl acrylamide. Using the information and protein characterization methods available to us at that time, we originally assigned this product as the hetero-Diels−Alder cycloaddition product between the oxidized aminophenol group and the acrylamide dienophile.[21] It now appears that the reaction involved unanticipated addition of the aniline group to the aminophenol as described herein. However, the mass change reported in the previous paper corresponded more closely to an analogue of **5a**, rather than **6a**. To verify this, the previously reported conditions (100 μM sodium periodate, pH 6.5, 2 h) were repeated on a peptide substrate using *p*-toluidine as the coupling partner (see SI for synthetic details). The mass difference of the product was indeed similar to that of proposed adduct **5a** (SI Figure S4f) but corresponded to **6a** using the 1−5 mM periodate conditions reported in this work (Figure S4e).

Capsids bearing aminophenol groups on the internal (**11**) or internal and external (**10**) surfaces were generated as described in the SI. For these experiments, 5 kDa and 2 kDa monomethoxy PEG anilines **15** and **16** were synthesized, as well as peptide aniline **17**. Under identical reaction conditions, the peptide aniline coupled to the internal surface of MS2-*o*-aminophenol-

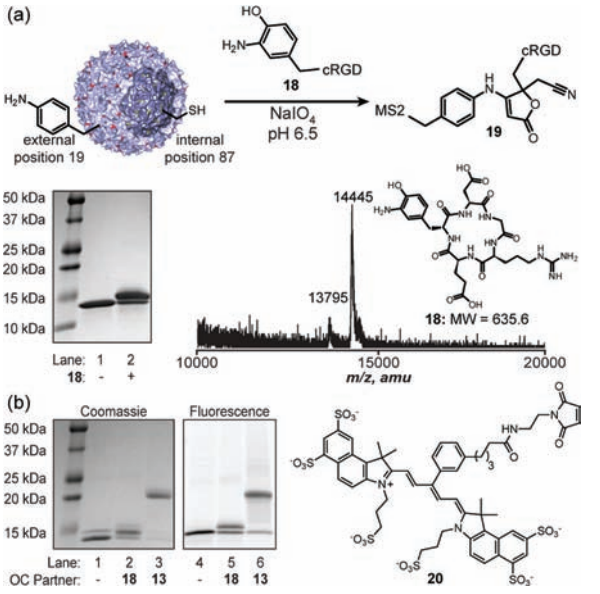

Figure 4. Attachment of a cyclic RGD peptide to T19*p*AF MS2. (a) Compound **18** (200 μM) was exposed toT19*p*AFMS2 (30 μM) in the presence of sodium periodate (5 mM) for 2 min. MS2 capsids were disassembled into monomers for analysis via SDS-PAGE and MALDI-TOF. Expected masses are *m*/*z* 13 795 forT19*p*AFMS2 and *m*/*z* 14 443 for **19**. (b) MS2 capsids bearing internal cysteines (N87C, yellow) and external *p*AF19 groups (red) were labeled with NIR dye **20**. Following disassembly, the samples were analyzed by SDS-PAGE (lanes 1 and 4). Coomassie staining is shown on the left, and fluorescence imaging of **20** is shown on the right. Portions of the capsids were then externally labeled with **18** (290 μM, lanes 2 and 5) or **13** (220 μM, lanes 3 and 6) upon exposure to 5 mM periodate for 5 min.

85 capsids (**11**) in high yield, as determined by MALDI-MS and SDS-PAGE (Figure 3c, lane 8, and SI Figure S2). Interestingly, neither of the PEG anilines coupled to **11** (shown for 2k-PEG substrate **16** in lane 7), likely because they were too large to pass through the 1.8 nm pores[22] and access the *o*-aminophenol at position 85 on the inner surface of the capsids. However, coupling of PEG-aniline to the bis(*o*-aminophenol)-MS2 capsids (**10**) showed significant conversion to the single polymer conjugate (lanes 4 and 5). Optical densitometry of the Coomassie-stained gel indicated a conversion of 40% with **15** and 50% with **16**. These numbers were lower than the above case, presumably because chemical modification of tyrosine to *o*-aminophenol does not proceed to completion, while the genetically introduced *p*-aminophenylalanine groups are present in all monomers. It was further observed that small amounts of MS2 dimers were obtained when the aminophenol groups were present on the protein component (see faint upper bands in lanes 4, 5, and 8). This minor pathway could occur through addition of adjacent protein nucleophiles under the high local concentration condition, or dimerization of the aminophenol groups. Though this corresponds to <10% of the protein monomers, it still seems more practical to include the aniline group on the protein component.

We have also used this reaction to create a protein conjugate with practical applications. Cyclic RGD is a widely used peptide for the binding of integrin $\alpha_V\beta_3$.[23] A five amino acid version of cyclic RGD was converted to the corresponding *o*-aminophenol (**18**) via reaction with tetranitromethane and subsequent reduction with sodium dithionite (see SI Scheme S1). **18** (120 μM)

was then coupled to T19*p*AF MS2 (7, 30 μM in monomer) in the presence of 5 mM sodium periodate at pH 6.5. SDS-PAGE and MALDI-TOF analysis of the MS2 coat protein monomers showed excellent conversion to the peptide conjugate (Figure 4a), with densitometry analysis indicating that 65% of the monomers had been modified.

We believe the chemoselectivity of this reaction stems from the lack of potent nucleophiles at neutral pH on most proteins, in combination with the short reaction times that are required. Free thiols would be an exception to this, and therefore we recommend first labeling cysteine residues with a desired functional group before carrying out the oxidative coupling step. As with any reaction involving sodium periodate, there is a possibility of N-terminal serine cleavage, carbohydrate oxidation, and methionine oxidation.

As an example of the compatibility of this chemistry with cysteine labeling, we have prepared MS2 capsids bearing near-IR dyes on the internal surface and cRGD peptides or PEG on the exterior. MS2 capsids bearing cysteine residues in position 87 and *p*AF residues in position 19 were prepared using the amber codon suppression technique and exposed to maleimide chromophore **20** in pH 6.5 phosphate buffer.[9] SDS-PAGE analysis indicated that 20% of the interior cysteine residues had been modified with the bulky group (previous studies in our laboratory have shown that *p*AF19 MS2 capsids lacking the cys87 groups are unreactive toward maleimide reagents under these conditions[9]). Separate samples of the chromophore-labeled capsids were then exposed to cRGD-aminophenol **18** and 5k-PEG-aminophenol **13** for 5 min in the presence of periodate. SDS-PAGE and fluorescence imaging clearly indicated the presence of monomers labeled with both the chromophores and the external groups, Figure 4b. Analysis of the resulting capsids using SEC indicated intact capsids (see SI Figure S5).

Through these studies we have developed a rapid, chemoselective reaction for the coupling of anilines to aminophenols. We demonstrated that the reaction conditions are mild enough to modify proteins under non-denaturing conditions and showed its capability to attach cancer-targeting groups to viral capsids that contain optical imaging cargo. The efficiency of this coupling technique has great promise for building a number of complex protein-based materials, and it is currently being explored for this purpose. We are also working to elucidate the mechanism of this unprecedented coupling reaction.

■ ASSOCIATED CONTENT

Supporting Information. Experimental procedures and characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
mbfrancis@berkeley.edu

■ ACKNOWLEDGMENT

These studies were supported by the DOD Breast Cancer Research Program (BC061995). C.R.B. was supported by DOE California Alliance for Radiotracer Education grant DESC0002061, and A.C.O. was supported by an NSF graduate research fellowship. The UC Berkeley Chemical Biology Graduate Program (Training Grant 1 T32 GMO66698) is also acknowledged for their support of C.R.B., D.W.R., and A.C.O.

■ REFERENCES

(1) Hermanson, G. T. *Bioconjugate Techniques*, 2nd ed.; Academic Press: San Diego, 2008.

(2) For some interesting lead examples, see: (a) Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T. *Annu. Rev. Biochem.* **2008**, *77*, 51–76. (b) Evans, M. J.; Cravatt, B. F. *Chem. Rev.* **2006**, *106*, 3279–3301. (c) Cecconi, C.; Shank, E. A.; Bustamante, C.; Marqusee, S. *Science* **2005**, *309*, 2057–2060. (d) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705–6710. (e) Valiyaveetil, F. I.; Sekedat, M.; MacKinnon, R.; Muir, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17045–17049. (f) Cooper, J. A.; Walker, S. B.; Pollard, T. D. *J. Mus. Res. Cell. Motil.* **1983**, *4*, 253–262.

(3) (a) Wang, P.; Sergeeva, M. V.; Lim, L.; Dordick, J. S. *Nat. Biotechnol.* **1997**, *15*, 789–793. (b) Nam, Y. S.; Magyar, A. P.; Lee, D.; Kim, J.; Yun, D. S.; Park, H.; Pollom, T. S.; Weitz, D. A.; Belcher, A. M. *Nat. Nanotech.* **2010**, *5*, 340–344.

(4) Day, J. J.; Marquez, B. V.; Beck, H. E.; Aweda, T. A.; Gawande, P. D.; Meares, C. F. *Curr. Opin. Chem. Biol.* **2010**, *14*, 1–7.

(5) Alley, S. C.; Okeley, N. M.; Senter, P. D. *Curr. Opin. Chem. Biol.* **2010**, *14*, 529–537.

(6) (a) Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *Angew. Chem., Int. Ed.* **2009**, *48*, 9493–9497. (b) Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. *ACS Nano* **2010**, *4*, 6014–6020.

(7) (a) Hooker, J. M.; Datta, A.; Botta, M.; Raymond, K. N.; Francis, M. B. *Nano Lett.* **2007**, *7*, 2207–2210. (b) Meldrum, T.; Seim, K. L.; Bajaj, V. S.; Palaniappan, K.; Wu, W.; Francis, M. B.; Wemmer, D. E.; Pines, A. *J. Am. Chem. Soc.* **2010**, *132*, 5936–5937.

(8) Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. *Chem. Commun.* **2008**, 1205–1207.

(9) Tong, G. J.; Hsiao, S. C.; Carrico, Z. M. *J. Am. Chem. Soc.* **2009**, *131*, 11174–11178.

(10) Strable, E.; Prasuhn, D. E., Jr.; Udit, A. K.; Brown, S.; Link, A. J.; Ngo, J. T.; Lander, G.; Quispe, J.; Potter, C. S.; Carragher, B.; Tirrell, D. A.; Finn, M. G. *Bioconjugate Chem.* **2008**, *19*, 866–875.

(11) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem., Int. Ed.* **2009**, *48*, 6974–6998.

(12) Cornish, V. W.; Hahn, K. M.; Schultz, P. G. *J. Am. Chem. Soc.* **1996**, *118*, 8150–8151.

(13) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 11782–11783.

(14) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164–11165.

(15) Mehl, R. A.; Anderson, J. C.; Santoro, S. W.; Wang, L.; Martin, A. B.; King, D. S.; Horn, D. M.; Schultz, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 935–939.

(16) (a) Wang, L.; Schultz, P. G. *Chem. Commun.* **2002**, 1–11. (b) Kiick, K. L.; Tirrell, D. A. *Tetrahedron* **2000**, *56*, 9487–9493.

(17) Hooker, J. M.; Esser-Kahn, A. P.; Francis, M. B. *J. Am. Chem. Soc.* **2006**, *128*, 15558–15559.

(18) Stephanopoulos, N.; Carrico, Z. M.; Francis, M. B. *Angew. Chem., Int. Ed.* **2009**, *48*, 9498–9502.

(19) For examples of other bioconjugation reactions involving sodium periodate, see: (a) Geoghegan, K. F.; Stroh, J. G. *Bioconjugate Chem.* **1992**, *3*, 138–146. (b) Liu, B.; Burdine, L.; Kodadek, T. *J. Am. Chem. Soc.* **2006**, *128*, 15228–15235.

(20) Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 3718–3719.

(21) Nair, V.; Mathew, B. *Tetrahedron* **1999**, *55*, 11017.

(22) Valegard, K.; Liljas, L.; Fridborg, K.; Unge, T. *Nature* **1990**, *345*, 36–41.

(23) For other examples of viral capsids bearing tissue-specific targeting groups, see: (a) Suci, P. A.; Varpness, Z.; Gillitzer, E.; Douglas, T.; Young, M. *Langmuir* **2007**, *23*, 12280–12286. (b) Banerjee, D.; Liu, A.; Voss, N. R.; Schmid, S.; Finn, M. G. *ChemBioChem* **2010**, *11*, 1273–1279. (c) Mastico, R. A.; Talbot, S. J.; Stockley, P. G. *J. Gen. Virol.* **1993**, *74*, 541–548.