

Cu(I)-catalysed N–H insertion in water: a new tool for chemical biology†

Cite this: *Chem. Sci.*, 2013, **4**, 4401

Kiril Tishinov, Na Fei and Dennis Gillingham*

We demonstrate for the first time that Cu(I)-catalysis can deliver N–H insertion (NHI) with α -diazocarbonyl compounds in aqueous media. Despite being carried out in water only trace amounts of O–H insertion are seen, indicating the catalyst's overwhelming preference for NHI. Our optimized conditions for NHI converged with those used for the Cu-catalyzed azide–alkyne cycloaddition (CuAAC) for bioconjugation, spurring us to develop an auto-tandem catalytic process where both catalytic cycles operate simultaneously in one-pot. We explore the scope of the method with both small molecule and complex DNA and RNA substrates.

Received 15th May 2013

Accepted 22nd August 2013

DOI: 10.1039/c3sc51363g

www.rsc.org/chemicalscience

Introduction

Synthetic economy is determined by the brevity of a synthesis¹ and by practical considerations such as solvent use, choice of reagents, and choice of catalysts.² While new methodology developments underpin improvements in synthetic strategy, the precise ingredients for optimizing practicality are more nebulous and often case-specific. One generalization that can be made is that achieving multiple synthetic tasks in a single vessel is economical.³ For catalytic reactions developments in this direction include tandem, domino, or one-pot-sequential processes.⁴ Despite these developments, none of the known one-pot tandem catalytic reactions are amenable to the needs of chemical biologists because they are unproven in aqueous buffers or ineffective with biomolecules. We describe here our surprising discovery that Cu-catalyzed N–H insertion (NHI) is efficient and selective in aqueous media with both simple (Table 2) and complex (Tables 1, 3 and 4) aniline-type substrates. The NHI process is fully compatible with the Cu-catalyzed azide–alkyne cycloaddition (CuAAC),^{5,6} and indeed both processes can be run simultaneously with the same catalyst (formally this is an auto-tandem catalytic reaction).⁷ Importantly, the reaction is not only efficient with small molecules, but can also be used to directly modify complex biomolecular structures such as DNA and RNA (Scheme 1).

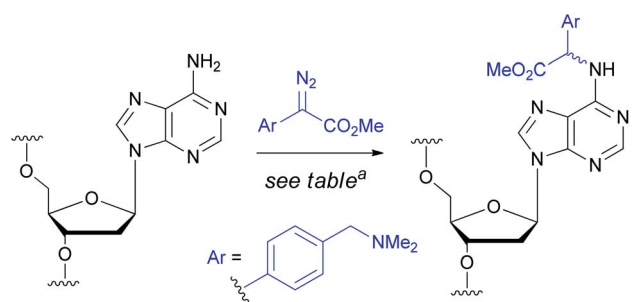
Methods for DNA and RNA tailoring are crucial for understanding their biological role as well as adapting them for diagnostic or therapeutic use. In this sense the CuAAC is a mainstay of nucleic acid (NA) chemical biology as it enables the

efficient and straight-forward introduction of a variety of functionally important tags and reporter groups into NAs under mild conditions.^{8–10} However, a major obstacle in applying the CuAAC with NAs is the need to introduce the alkyne or azide moiety in large complex structures. This is usually done with modified monomer units for solid-phase synthesis,¹¹ modified deoxynucleoside triphosphates for PCR,¹² or by enzyme-catalyzed 3'-terminal transfer of an azido-derivatized nucleotide.¹³ We have recently reported an alternative way to tackle this problem through direct catalytic modification of native NAs.¹⁴ In contrast to other approaches, our method relies on an N–H insertion (NHI) reaction of a rhodium(II)-stabilized carbenoid directly into the exocyclic amino groups of the nucleobases. Combining our Rh-catalysed method with the CuAAC reaction, however, still required separate catalysts and reaction vessels. We therefore considered the possibility of a simultaneous CuAAC/NHI process, but were surprised to find that Cu(I)-catalyzed NHI in aqueous medium had never been reported.

NHI with Cu(I)-carbenoids derived from α -diazo carbonyl compounds in organic solvents has been known for decades.¹⁵ Nevertheless, the dominance of rhodium-based catalysts has left Cu(I) systems underdeveloped.¹⁶ The works of Perez,¹⁷ Jørgensen,¹⁸ Fu,¹⁹ and Zhou²⁰ have spurred a renaissance in Cu(I)-carbenoid chemistry, but studies of its potential in aqueous media are lacking. There are two examples from the sixties where large excesses of Cu(II) salts were used to effect protein modification with diazo peptides.^{21,22} Interpreted through the lens of the past fifty years of developments, these reactions were likely a result of *in situ* reduction to Cu(I) followed by Cu-carbenoid transfer.²³ Indeed a very recent report has shown that intramolecular Cu-catalysed cyclopropanation can be carried out in aqueous media using DNA as a chiral control element.²⁴ These observations hint at unrealized potential for catalytic Cu(I)-carbenoid chemistry in water.

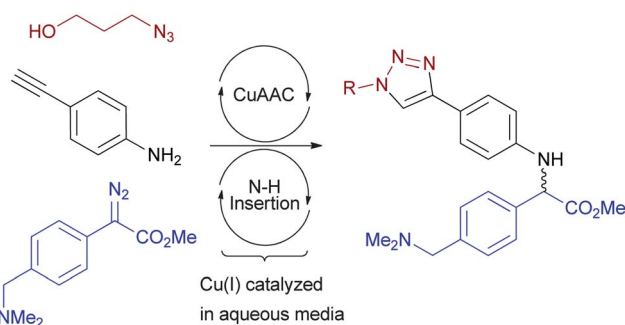
Department of Chemistry, University of Basel, St. Johanns-Ring 19, CH-4056, Basel, Switzerland. E-mail: dennis.gillingham@unibas.ch

† Electronic supplementary information (ESI) available: See DOI: 10.1039/c3sc51363g

Table 1 Optimization of d(ATGC) alkylation^a


Entry	Cu salt	Ligand	[Ascorbate] (mM) ^b	d(ATGC) conv. (%)
1	CuSO ₄	—	0	14
2	Cu(OTf) ₂	—	0	13
3	CuSO ₄	2,2'-Bipyridyl	0	6
4	CuSO ₄	1,10-Phenanthroline	0	6
5	CuSO ₄	THPTA	0	18
6 ^c	CuSO ₄	—	50	47
7	CuSO ₄	2,2'-Bipyridyl	50	30
8	CuSO ₄	1,10-Phenanthroline	50	36
9	CuSO ₄	THPTA	50	52

^a Reaction conditions: 5 mM d(ATGC), 50 mM diazo substrate, 1 mM Cu salt, 1 mM ligand, 100 mM MES buffer, pH 6, 24 h, room temperature. Although A is shown, other major products are monoalkylations of G and C nucleobases as determined by MS/MS data (see ESI† for details). ^b Ascorbate reduces Cu(II) to Cu(I). ^c In the absence of THPTA ligand monoalkylation is the major product, but other unidentified products and extensive oxidative decomposition of d(ATGC) are also observed.

**Scheme 1** Overview of the auto-tandem catalytic process with Cu(I).

Results and discussion

Given our interest in catalytic methods for NA alkylation, our initial examination of the catalytic NHI started with the short oligonucleotide d(ATGC) (see Table 1) in aqueous MES buffer at pH 6. A screening of transition metals at different oxidation states identified copper as the best NHI catalyst (see ESI, Table S1†). Further refinement of the reaction conditions, including testing different ligands and copper sources, led to a convergence with the CuAAC conditions for bioconjugation developed

by Finn (Table 1).⁹ The key components of the Finn conditions are sodium ascorbate and the tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) ligand. Ascorbate insures the Cu(I) oxidation state is maintained even in the presence of oxygen (diazo compounds can also mediate the Cu(II) → Cu(I) reduction, but the process is much slower²³). A consequence of Cu(I) in the presence of oxygen, however, is the formation of reactive oxygen species (ROS). The role of the THPTA ligand seems to be as a sacrificial substrate to scavenge ROS before they wreak havoc on other reaction components. Consistent with this hypothesis, NHI with d(ATGC) in the absence of THPTA leads to low mass recovery and extensive decomposition of the starting d(ATGC) (cf. entries 6 and 9, Table 1).

Considering the remarkable facility with which Cu(I) can catalyze NHI with a complex substrate like d(ATGC), it is surprising that this type of catalysis has never been studied in water. We therefore proceeded to outline the scope of the method with other simpler nucleophiles like anilines, alkyl amines, alcohols, and thiols (see Table 2). Indeed using the optimized conditions from Table 1, anilines were transformed with uniformly high efficiency (entries 1–9, Table 2). The reactions proceeded smoothly at as low as 1 mol% catalyst loading, giving complete conversion of the starting material and isolated yields between 51 and 85% of the monoalkylation products. Little dependence on the electronic nature of the aniline was seen, since comparable results were obtained with electronically dissimilar anilines (entries 3–7, Table 2). A detailed analysis of the minor side-products seen in the crude reaction mixtures revealed the formation of a small amount of doubly alkylated aniline (<6%, see Table S2 in the ESI†), as well as trace amounts of O–H insertion products with water and ascorbate. The observation of O–H insertion products prompted us to test 1-butanol (entry 17, Table 2) but it proved completely unreactive. The reactivity observed with water is likely a consequence of its enormous excess as the reaction solvent. For ascorbate the presence of an acidic O–H group at C-3 (pK_a = 4.25)²⁵ likely accelerates the carbenoid insertion.

In contrast to Cu(I) catalysis in organic solvents,^{17,26} other types of nucleophiles were unreactive (entries 10–17, Table 2). In the case of the nitrogen-based nucleophiles a complete selection for anilines over aliphatic amines and hydroxylamines was observed. At least in the case of anilines vs. alkyl amines this astonishing selectivity is likely attributable to the different protonation states of the substrates due to the differences in pK_a of the respective conjugate acids. From a practical standpoint this solvent-dependent reactivity offers a strategy to select the site of reaction in molecules that contain both aryl and alkyl amines.

Amino acids, peptides, and proteins proved poorly reactive with the copper-carbenoids (entries 18–20, Table 2). This intriguing result suggests that while rhodium(II) targets N–H bonds in both proteins and NAs,^{14,27,28} copper catalysis may offer a way to selectively target NAs in complex mixtures. In this vein, although peptides and proteins were unreactive, in general they do not inhibit the copper-catalysed NA labelling. For example the alkylation of d(TAT) proceeds identically in the presence or absence of lysozyme or streptavidin (see the ESI, Table S3†). The

Table 2 NHI reactions with various nucleophiles

Entry	RX-H	Conv. (%)	Time (min)	Yield (%)
1		>99	15	85
2		>99	15	82
3		>99	60	75
4		>99	30	77
5		>99	15	70
6		>99	15	74
7		>99	15	75
8		>99	20 h	51
9		>99	20 h	76
10		<2	24 h	—
11		<2	24 h	—
12		<2	24 h	—
13	H ₂ NOH	<2	24 h	—
14	H ₂ NOMe	<2	24 h	—
15		<2	24 h	—
16		<2	24 h	—
17		<2	24 h	—
18	H-Trp-OMe	<2	24 h	—
19	Streptavidin	<2	24 h	—
20	Lysozyme	<2	24 h	—

only protein we have tested that has a deleterious effect on the reaction is BSA; we attribute this effect to the propensity of BSA to sequester hydrophobic molecules such as the THPTA-copper complex. We are currently further exploring this observation and synthesizing new ligands for copper that could interrupt the inhibitory interaction with BSA.

A kinetic analysis of the aniline NHI did not display Michaelis–Menten saturation kinetics. Instead the decrease in the initial velocities of diazo substrate consumption at high

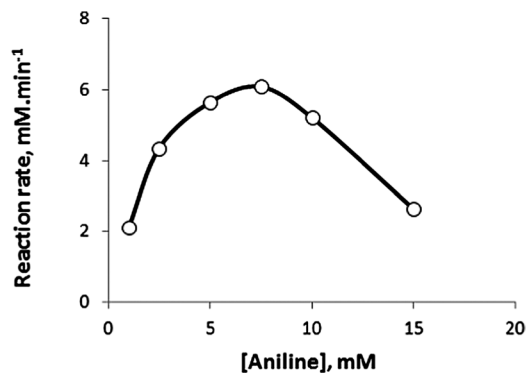


Fig. 1 Dependence of the initial rate of diazo substrate disappearance on the aniline concentration. Data points are an average of two measurements. *Reaction conditions:* 0.025 mM CuSO₄, 12.5 mM diazo substrate (see Table 2 for structure), 20 mM sodium ascorbate, 200 mM MES buffer, pH 6.0.

aniline concentrations (Fig. 1) suggests substrate inhibition of the catalyst. This finding is consistent with the lack of reactivity observed with strongly Lewis-basic substrates such as imidazole, thiols, and hydroxyl amines (entries 13–16, Table 2). Furthermore, conversion of the two pyridyl-aniline substrates (entries 8 and 9, Table 2) was substantially slower than for the other anilines.

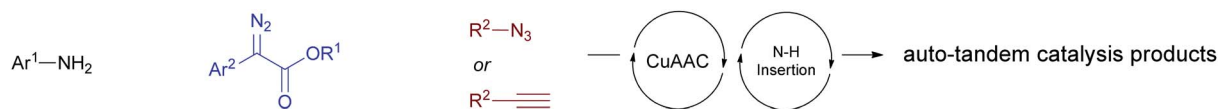
Armed with a better understanding of the key reaction parameters, we focused on delineating the reaction scope with NAs (Table 3). The alkylation products of the three trimers,

Table 3 Scope of NA alkylation^a

Entry	Sequence	Number of alkylation products ^b			NA conv. (%) ^c
		(+1)	(+2)	(+3)	
1	d(ATGC)	7	4	1	70
2 ^d	d(TAT)	2	—	—	37
3 ^d	d(TGT)	4	1	—	46
4 ^d	d(TCT)	4	—	—	27
5	d(TTTT)	7	1	—	19
6	d(ATG)	3	2	2	61
7	r(ACUGCUC)	>11	—	—	67
8 ^e	T T GCAAGC T ••••• T T CGTTCG	2	—	—	9
9 ^e	T T GCAAGC T ••••• T T CGTTCGA	4	—	—	16
10 ^e	T T GCAAGC A ••••• T T CGTTCG	6	—	—	16

^a *Reaction conditions:* 5 mM oligonucleotide, 0.5 mM CuSO₄, 2.5 mM THPTA, 50 mM diazo substrate (see Table 2 for structure), 10 mM sodium ascorbate, 100 mM MES, pH 6, 24 h, room temperature.

^b Determined by HPLC-analysis of the reaction mixture. ^c Amount of the modified species as a percentage of the total amount of NA. ^d In all three cases the alkylation site was in the central nucleobase as determined by MS/MS. ^e Reaction is run for 48 h.

Table 4 Auto-tandem catalytic CuAAC/NHI

- alkyne can be incorporated at Ar^1 , Ar^2 , or R^1 , or R^2
- azide can be incorporated at Ar^1 , Ar^2 , or R^1 , or R^2

Entry	Amine	Diazo	Azide or alkyne	Product	Yield (%)
1 ^a					62
2 ^b					57
3 ^b					70
4 ^b					63
5 ^b					53
6 ^a					70
7 ^{c,d}	d(ATGC)				82% conv.
8 ^{c,d}	d(GGAGGC)				59% conv.
9 ^{c,d}	d(ATGC)				27% conv.

^a 0.5 mM CuSO₄, 50 mM aniline, 50 mM diazo substrate, 55 mM azide, 35% (v/v) *t*-BuOH, 24 h, room temperature. ^b These reactions were performed with 5 mol% CuSO₄ and 50% (v/v) *t*-BuOH. ^c 5 mM oligonucleotide, 0.5 mM CuSO₄, 2.5 mM THPTA, 50 mM diazo substrate, 55 mM azide, 24 h, room temperature. ^d Multiply alkylated (+2, +3, +4) species were observed with DNA substrates at high conversion. Conversion represents consumption of the starting DNA.

d(TAT), d(TGT), and d(TCT) have previously been characterized by NMR.¹⁴ This unambiguous characterization affords us a set of MS and HPLC standards to allow assignment of the sites of reactivity in the present study (see the ESI, Fig. S1–S9†). Based on these comparisons we can identify those reaction products from entries 2–4 of Table 3 that result from NHI with the exocyclic amino groups of the A, G, or C bases (comparison by ESI-MS, MS/MS, and HPLC, see ESI, Fig. S4–S20†). For d(TGT) and d(TCT) a second set of products was observed; these likely come from the modification of the endocyclic nitrogen, but we could not completely characterize these products due to insufficient material. In the case of d(TGT) we also discovered a peak corresponding to double-alkylation of the G nucleobase – a result indicative of the potency of the Cu(I) catalyst system.

The tetramer d(TTTT) reacted sluggishly delivering only 19% conversion, an observation consistent with our previous studies on rhodium-carbenoids where nucleobases lacking readily available polar X–H bonds (T and U) were poorly alkylated.¹⁴ The mixture of minor products was not characterized but likely arises from targeting the phosphate oxygens or the terminal 3'- or 5'-hydroxyl groups in the absence of other more potent nucleophiles.

RNA also proved to be a viable substrate for carbene insertion (Table 3, entry 7). Its extensive alkylation produced a number of product species highlighting the potential of the Cu(I)-system for RNA-tagging and dense functionalization.

The propensity of double-stranded DNA motifs to react was tested on three hairpin structures (entries 8–10, Table 3). All three hairpins were significantly less reactive than the single-stranded NAs tested, delivering only modest yields of modified DNA at extended reaction times. As expected the hairpin containing only Ts in the loop (entry 8, Table 3) proved to be the least reactive with only 9% conversion. In comparison the remaining two hairpins (entries 9 and 10, Table 3) exhibit a two-fold increase in reactivity. These contain an unpaired A as either a 3'-overhang (entry 9) or in the loop region (entry 10), indicating that copper preferentially targets nitrogens that are not involved in Watson–Crick base-pairing.

With the scope outlined, the culmination of our study was to combine the CuAAC with NHI in a single-catalyst process. The concept was tested with a collection of different amines, alkynes, diazo compounds, and azides as shown in Table 4. Entries 1–6 illustrate our initial experiments with small molecule substrates and in all cases the *N*-aryl triazole/NHI products were obtained in good yields (53–70%). Furthermore, changing the position of the participating functional groups had little impact on the reaction. This multi-component catalytic process should prove useful for medicinal and combinatorial chemists since it provides a robust strategy to quickly assemble complex molecular scaffolds.

Guided by the results with small molecules we attempted the simultaneous CuAAC/NHI with more challenging DNA-based substrates. With DNA we employed a bifunctional propargylated α -diazocarbonyl compound to serve as a connecting element between the NA and the azide-bearing tag (see entries 7–9, Table 4). A number of alkylation products were observed for each oligonucleotide with moderate (27%) to good (59–82%)

conversion of the DNA starting material. Tandem MS analysis of the d(ATG) modification products revealed that the guanine base is the primary alkylation site, undergoing two consecutive modifications before the adenine is modified. The efficacy of the CuAAC reaction assured that all modified NAs bore the triazole tag; we have never observed NHI products that have not also undergone a CuAAC reaction.

The potency of the Cu(I) system facilitates the introduction of multiple alkylations on a single NA when the reaction is run to high conversions of the starting NA. Therefore, in its present form the auto-tandem catalytic approach to NA alkylation is best suited to applications in labelling, detection, or dense functionalization of NAs. Multiply-alkylated NAs are important in a variety of applications. For example, fluorescence *in situ* hybridization (FISH) detection probes for DNA and RNA are far more sensitive when multiple fluorophores are incorporated.²⁹ In the field of DNA nanotechnology NAs densely functionalized with aldehydes have been employed to control the deposition of silver nanowires, a strategy that may facilitate the construction of molecular-scale circuitry.^{11,12} Direct unselective modification simplifies the covalent introduction of biotin into NAs (as in entry 9 of Table 4), and could be used to label complex NAs isolated from natural sources.³⁰ The presence of biotin opens the door to the entire molecular biology toolbox based on the biotin–streptavidin interaction, such as in affinity purification or highly sensitive detection by catalytic amplification with the horseradish peroxidase/streptavidin fusion protein.³¹ Biotin is traditionally introduced through costly and complicated enzymatic protocols, solid-phase synthesis, or by destructive UV-light promoted modification.^{31–33} The direct NA modification method we report here is orthogonal and complementary to these established approaches, and is mild enough to tolerate both DNA and RNA.

Conclusions

Our work outlines a novel approach for molecular conjugation based on the discovery that Cu(I)-carbenoid chemistry is viable in water. The efficacy of Cu(I) for catalysing both CuAAC and arylamine NHI drove us to combine both reactions in a one-pot operationally simple process. The substrate range includes simple arylamines as well as the nucleobases in DNA and RNA. The next step is to identify ligands for copper that facilitate site-specific labelling of long NAs. Already in its present form, however, the method should open new avenues to the synthesis of diverse small molecule scaffolds, and in the labelling and dense functionalization of NAs.

Acknowledgements

The University of Basel is gratefully acknowledged for support of this work. Mr Julian Ruoss and Prof. Thomas Ward are thanked for a gift of wild-type Streptavidin.

Notes and references

- 1 T. Gaich and P. S. Baran, *J. Org. Chem.*, 2010, 75, 4657–4673.

- 2 N. Anderson, *Practical Process Research and Development*, Elsevier Academic Press, Waltham, MA, 2012.
- 3 Y. Hayashi and S. Umeyama, *Angew. Chem., Int. Ed.*, 2013, **52**, 3450–3452.
- 4 N. T. Patil, V. S. Shinde and B. Gajula, *Org. Biomol. Chem.*, 2012, **10**, 211–224.
- 5 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2002, **41**, 2596–2599.
- 6 C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057–3064.
- 7 D. E. Fogg and E. N. dos Santos, *Coord. Chem. Rev.*, 2004, **248**, 2365–2379.
- 8 A. H. El-Sagheer and T. Brown, *Chem. Soc. Rev.*, 2010, **39**, 1388–1405.
- 9 V. Hong, S. I. Presolski, C. Ma and M. G. Finn, *Angew. Chem., Int. Ed.*, 2009, **48**, 9879–9883.
- 10 C. Y. Jao and A. Salic, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15779–15784.
- 11 J. Gierlich, G. A. Burley, P. M. E. Gramlich, D. M. Hammond and T. Carell, *Org. Lett.*, 2006, **8**, 3639–3642.
- 12 G. A. Burley, J. Gierlich, M. R. Mofid, H. Nir, S. Tal, Y. Eichen and T. Carell, *J. Am. Chem. Soc.*, 2006, **128**, 1398–1399.
- 13 M.-L. Winz, A. Samanta, D. Benzinger and A. Jäschke, *Nucleic Acids Res.*, 2012, **40**, e78.
- 14 K. Tishinov, K. Schmidt, D. Häussinger and D. G. Gillingham, *Angew. Chem., Int. Ed.*, 2012, **51**, 12000–12004.
- 15 P. Yates, *J. Am. Chem. Soc.*, 1952, **74**, 5376–5381.
- 16 D. Gillingham and N. Fei, *Chem. Soc. Rev.*, 2013, **42**, 4918–4931.
- 17 M. E. Morilla, M. M. Diaz-Requejo, T. R. Belderrain, M. C. Nicasio, S. Trofimenko and P. J. Perez, *Chem. Commun.*, 2002, 2998–2999.
- 18 S. Bachmann, D. Fielenbach and K. A. Jørgensen, *Org. Biomol. Chem.*, 2004, **2**, 3044–3049.
- 19 T. C. Maier and G. C. Fu, *J. Am. Chem. Soc.*, 2006, **128**, 4594–4595.
- 20 S.-F. Zhu and Q.-L. Zhou, *Acc. Chem. Res.*, 2012, **45**, 1365–1377.
- 21 T. G. Rajagopalan, W. H. Stein and S. Moore, *J. Biol. Chem.*, 1966, **241**, 4295–4297.
- 22 G. R. Delpierre and J. S. Fruton, *Proc. Natl. Acad. Sci. U. S. A.*, 1966, **56**, 1817–1822.
- 23 R. G. Salomon and J. K. Kochi, *J. Am. Chem. Soc.*, 1973, **95**, 3300–3310.
- 24 J. Oelerich and G. Roelfes, *Chem. Sci.*, 2013, **4**, 2013–2017.
- 25 B. Zümreoglu-Karan, *Coord. Chem. Rev.*, 2006, **250**, 2295–2307.
- 26 M. R. Fructos, T. R. Belderrain, M. C. Nicasio, S. P. Nolan, H. Kaur, M. M. Díaz-Requejo and P. J. Pérez, *J. Am. Chem. Soc.*, 2004, **126**, 10846–10847.
- 27 J. M. Antos and M. B. Francis, *J. Am. Chem. Soc.*, 2004, **126**, 10256–10257.
- 28 B. V. Popp and Z. T. Ball, *J. Am. Chem. Soc.*, 2010, **132**, 6660–6662.
- 29 A. Pernthaler, C. M. Preston, J. Pernthaler, E. F. DeLong and R. Amann, *Appl. Environ. Microbiol.*, 2002, **68**, 661–667.
- 30 E. P. Diamandis and T. K. Christopoulos, *Clin. Chem.*, 1991, **37**, 625–636.
- 31 G. T. Hermanson, *Bioconjugate techniques*, Elsevier/Academic Press, Amsterdam, 2008.
- 32 S. H. Weisbrod and A. Marx, *Chem. Commun.*, 2008, 5675–5685.
- 33 E. T. Kool, *Curr. Opin. Chem. Biol.*, 2000, **4**, 602–608.