ChemComm
Chemical Communications

## Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: R. D. Vargas, Y. Ding, H. O. Trial, R. Qian and Z. T. Ball, *Chem. Commun.*, 2023, DOI: 10.1039/D3CC03371F.




ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemcomm

View Article Online
DOI: 10.1039/D3CC03371F

COMMUNICATION

# Polyol recognition in catalysis: Toward selective modification of glycosylated polypeptides with boronic acid — rhodium(II) catalysts

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Reyner D. Vargas,[a] Yuxuan Ding, [a] Hallie O. Trial, [a] Rouyu Qian,[a] Zachary T. Ball[a*]



**Proximity-induced methodologies for peptide and protein modification have been developed using recognition elements like inhibitors, antibodies, or affinity tags on amino acids. However, the recognition a of saccharides for chemical modification remains widely unexplored. Recent studies exploring boronic acids and their derivatives have shown their alluring capabilities as selective molecular recognition elements for saccharides, providing the first insight into a recognition methodology for this moiety. Herein is described the discovery of catalytic proximity-induced rhodium(II) methodology for covalent modification of glycopeptides using boronic acids as a saccharide recognition element.**

Chemoselective manipulation of biomacromolecules remains a daunting and fundamental reactivity challenge. The complex polyfunctional and aqueous environment of polynucleotides, proteins, and saccharides limits the chemical tools that can be deployed and requires unique approaches to selectivity challenges. Among these biopolymer types, polysaccharides remain one of the most challenging and little-studied motifs for chemoselective bioconjugation, in spite of the essential and diverse roles that glycosylation plays in biological systems. Glycosylation is a common post-translational modification (PTM)[1] of proteins, and influences cellular tracking, adhesion and recognition[2–4], structural properties such as folding and protein stability[5,6], and disease virulence.[7,8]

We previously explored noncovalent molecular recognition as a tool for selective modification of peptides and proteins with a recognition element conjugated to a rhodium(II) catalyst capable of modifying natural amino acids via a metallocarbene intermediate. A catalytic approach to proximity–driven modification has unique advantages, including the use of recognition elements at lower, biologically relevant, concentrations, and the production of "traceless" modifications without a recognition element.

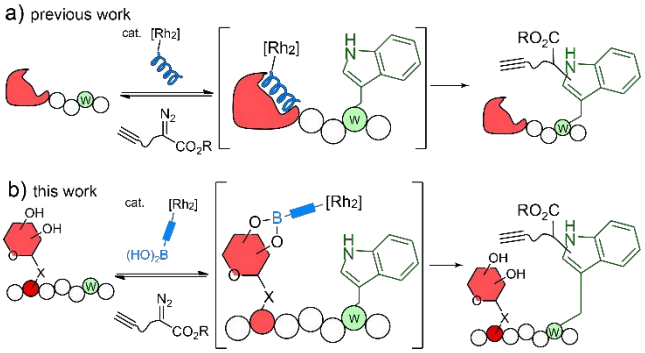



***Figure 1***. Proximity-driven modification of tryptophan residues using rhodium catalysts (a) previous work using peptide molecular recognition of peptides and proteins. [9–11] (b) this work using boronic acids as a sugar recognition element.

Manipulating glycosylation sites is challenging. The heterogeneous ensembles of glycosylation found in natural systems are not genetically encoded, saccharide molecules do not provide much in the way of unique functional groups for the design of selectively reactive reagents, and chemical tools to probe or alter glycan structures are quite limited.[12,13] Unnatural sugars can be used to incorporate unique reactive handles in living systems. Unnatural amino acids have been developed for bioorthogonal posttranslational glycosylation, and chemoenzymatic methodologies can sometimes be used for residue-selective glycosylation.[2–14] For some target structures, total chemical synthesis is a suitable tool. We became interested in exploring dynamic covalent molecular recognition, as an alternative to noncovalent ligand docking, to control selectivity in the modification of glycosylated polypeptides.

While saccharide-selective small molecule ligands are rare, boronic acids have a rich chemistry of transient boronate ester formation with proximal diols. These properties have been utilized in the development of biological probes[14], sugar sensors[15], and affinity media.[16] In thinking of a glycosylation site as a handle for docking a catalyst via dynamic covalent molecular recognition, we hoped to explore a possible new tool for direct selectivity in modification of complex polyol-containing structures.

[a] *Department of Chemistry, Rice University, Houston, Texas 77005.*
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

To explore the potential for dynamic covalent chemistry of boronic acids to facilitate catalytic metallocarbene chemistry at glycopeptide structures, we designed a series of heteroleptic rhodium–boronic acid conjugates (Figure 2). While the potential for boronate–rhodium interactions presented some concerns, gratifyingly, we found that installing an unprotected boronic acid within rhodium(II) tetracarboxylate framework was readily achieved by displacement of a trifluoroacetate ligand from the heteroleptic precursor, $Rh_2(OAc)_3(tfa)$ (Figure 2a). This synthetic approach relies on a synthetic method we developed for preparing complex rhodium(II)–peptide conjugates,[17] and has subsequently been shown to be quite general for the preparation of other rhodium(II) structures with complex, polyfunctional ligands[9,10,18–24]. Using this approach, we accessed a set of conjugates (**1-7**) with different lengths and electronical properties (Figure 2b).

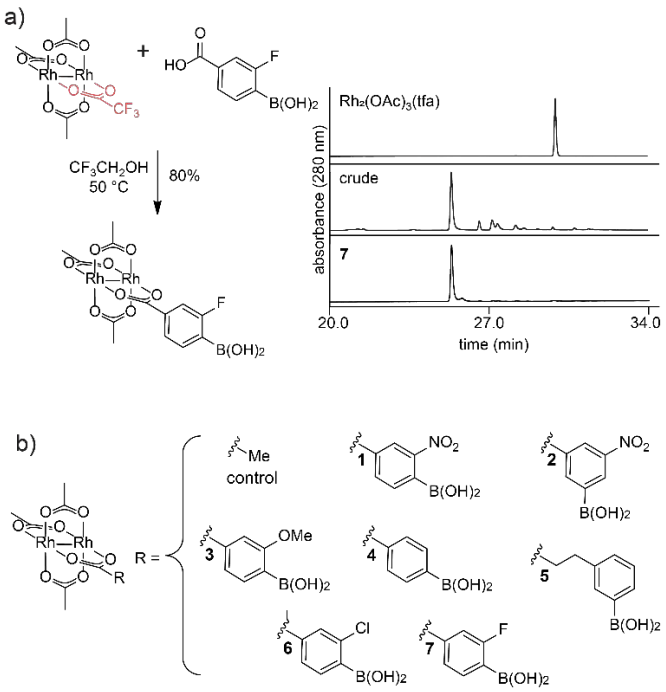



***Figure 2***. (a) Synthesized rhodium-boronic acid conjugates (b) Synthesis example, formation of catalyst 7, by ligand exchange. HPLC traces of (top) $Rh_2(OAc)_3(tfa)$ (middle) crude reaction (bottom) purified catalyst **7**.

We examined catalytic performance of the boronic acid–rhodium conjugates for the functionalization of a model substrate containing a polyol (sorbitol) and a tryptophan-containing hexapeptide (**gpep1**, Table 1 and Figure 3). The choice of sorbitol as our model polyol reflects its success in preliminary reactivity studies as well as ease of synthesis. In previous experiments, tryptophan had demonstrated the highest reactivity in rhodium metallocarbene reactions.[25,26] We tested reactivity with an alkyne-containing styryldiazo **8**.[27] Our initial screen tested modification of the peptide **gpep1** (500 μM) with diazo **8** in *N-tert*-butylhydroxylamine buffer, which was previously found optimal for diazo–tryptophan coupling.[19,26] We confirmed that boronate ester formation occurs in both phosphate buffer and in *N*-tert-butylhydroxylamine buffer in a mass spec assay with a model arylboronic acid (Figure S62). The desired diazo modification of **gpep1** (Figure 3) could be analysed by both mass spectrometry (MS) and HPLC. No modification of the tryptophan indole was observed with a simple control catalyst, $Rh_2(OAc)_4$, while catalyst **7**, bearing ortho fluoro substitution, exhibited significant conversion to tryptophan modification products. Subsequent confirmation and quantification by HPLC indicated that conjugate **7** afforded modified glycopeptide in 40% overall yield.

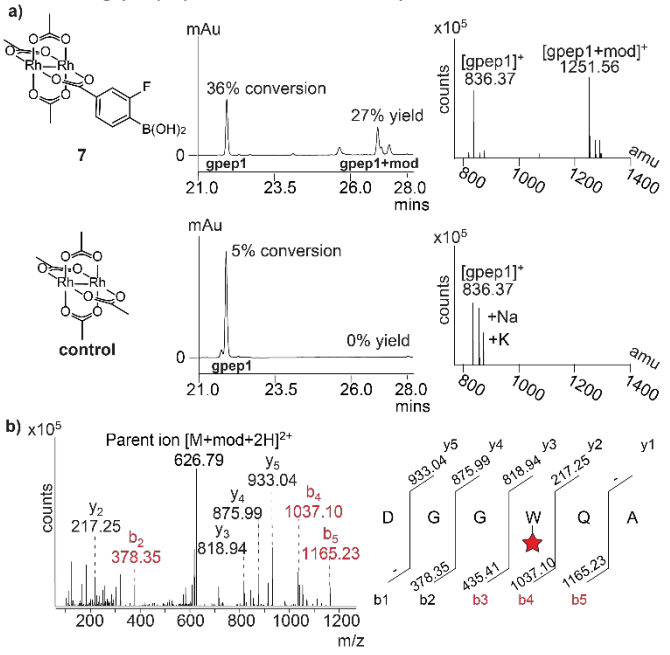


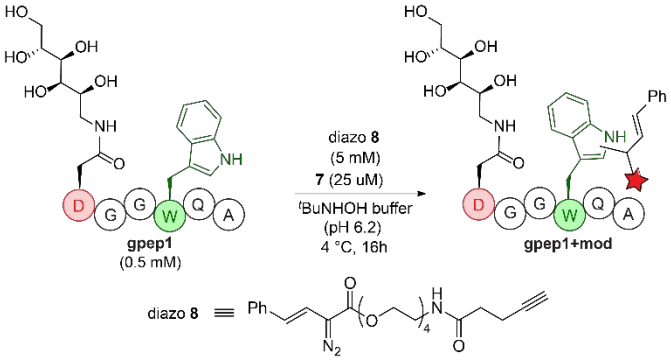

***Figure 3*** (a) HPLC trace showing the SM conversion and calculated product yield along the deconvoluted EIC-MS spectrum of the peptidic peaks in the modification of **gpep1** with diazo compound **8** under optimized conditions using both **7** as catalyst and **$Rh_2OAc_4$** as a control (b) MSMS spectrum and fragmentation ladder of **gpep1 + mod** showing the selective tryptophan modification.

Reaction optimization demonstrated that lowering the loading of rhodium conjugate **7** to 1 mol% (5 μM, entry f) still resulted in product formation, and loading as low as 5 mol% (25 μM, entry e) exhibited conversions similar to stoichiometric rhodium reactions (entry c). Arylboronic acids typically bind linear sugars, such as sorbitol, with $K_D$ values in the range of 1–20 mM.[28,29] The success of dynamic covalent chemistry to template catalytic modification at substantially lower catalyst concentrations, as low as 5 μM, indicates that the boronate ester adducts needed for catalysis are both transient and in low abundance under the reaction conditions. Modification was not observed in the presence of free D-sorbitol (entries l,m), and parent polypeptides without any polyol were unreactive, consistent with templated catalysis. Although *N*-tert-butylhydroxylamine buffer leads to improved catalysis, it is not required; reaction was also observed in phosphate buffer (PBS, entry g). Notably, we were unable to observe appreciable modification with the control $Rh_2(OAc)_4$ catalyst, even at higher concentrations of reagents and/or longer reaction times.

The modification product of **gpep1** with conjugate 7 was characterized by MS/MS, which confirmed exclusive modification of the tryptophan residue (Fig. 3b). HPLC analysis did indicate more than one product of diazo incorporation,

consistent with a mixture of heteroarene N-H and C-H insertion products observed in indole modification reactions previously, although products were not isolated.[31,32] Relative to previous noncovalent binding efforts[10,19,26], the reaction is more sensitive to p*H* and temperature. Elevated temperatures and elevated p*H* considerably reduced product formation, possibly reflecting boronic ester formation kinetics and/or thermodynamics[30] (Table 1).

Table 1. Catalyst scope, gpep1 reaction optimization conditions and competition studies.



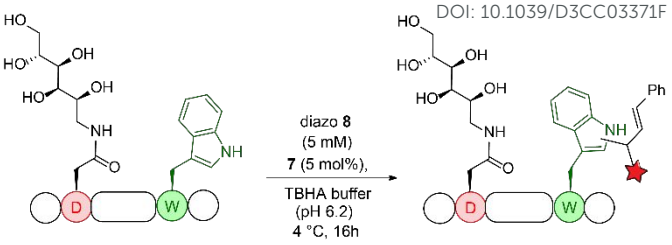

| entry | [8](mM) | [cat](μM) | variation | convn (%) |
|---|---|---|---|---|
| **a** | 12.5 | 1,000 | cat **1-5** | <5 |
| b | 12.5 | 1,000 | cat **6** | 16 |
| c | 25 | 1,000 | | 42 |
| d | 5 | 250 | | 39 |
| **e** | **5** | **25** (5 mol%) | | **36** |
| f | | 5 | | 15 |
| g | | | PBS buffer | 21 |
| h | | | pH 7.2 | <5 |
| i | | | pH 8.2 | <5 |
| j | | | 22 °C | 20 |
| k | | | 35 °C | <5 |
| l | | | 5 mM sorbitol | <5 |
| m | | | 50 mM sorbitol | <5 |

*Conversion was determined measuring **gpep1** concentration before and after the reaction.

With reliable reaction conditions and characterization methodology in hand, we began exploring the glycopeptide structure and reactivity relationships. We first examined polypeptides in which the intervening glycine amino acids were substituted with more hindered alanine (**gpep4**) or with proline (**gpep6**), which puts significant conformational constraints on the structure. The increased conformational constraints of **gpep1** and **gpep2** did not interfere with proximity-driven catalysis, and similar or marginally higher yields were observed (49% and 51%, respectively).

Next, variation in peptide sequence was explored, including the role of the distance between the polyol and the tryptophan. Significant reactivity was observed for a variety of spacings (Table 2 and Figure 4), from one to three intervening residues. The shortest chain (**gpep2**) with the smallest distance to the tryptophan showed reduced reactivity, conceivably reflecting spacing and conformational limitations of the rather rigid catalyst **7**.

**Table 2.** Variation of peptide sequence.




| Gpep | sequence | yield (%) |
|---|---|---|
| 1 | $^{Ac}$DGGWQA$^{NH2}$ | 27 |
| 2 | $^{Ac}$DGWQA$^{NH2}$ | 16 |
| 3 | $^{Ac}$DAAAWQA$^{NH2}$ | 28 |
| 4 | $^{Ac}$DAAWQA$^{NH2}$ | 51 |
| 5 | $^{Ac}$DPPPWQA$^{NH2}$ | 17 |
| 6 | $^{Ac}$DPPWQA$^{NH2}$ | 49 |
| 7 | $^{Ac}$TLDAAWSV$^{NH2}$ | 45 |

*Yield was determined by measuring product concentration.

Further polyol and residue effects were explored with the development of **gpep8-9** and **gpep10-11**, respectively (Figure 4), containing more biologically common cyclic sugar structures.[33] Therefore, we synthesized two other polyol models (**gpep8-9**) and found that the reducing sugar in **gpep8** was competent to direct tryptophan modification. However, the nonreducing methyl glycoside analogue **gpep9** provided lower reaction efficiency, potentially due to restricted conformations in this substrate. The amino-acid selectivity of the reactivity was high. Unlike with previous noncovalent recognition studies,[25,34] no cross-reactivity with other aromatic residues was observed. Peptides **gpep10** and **gpep11** (Figure 4) were unreactive under typical conditions.

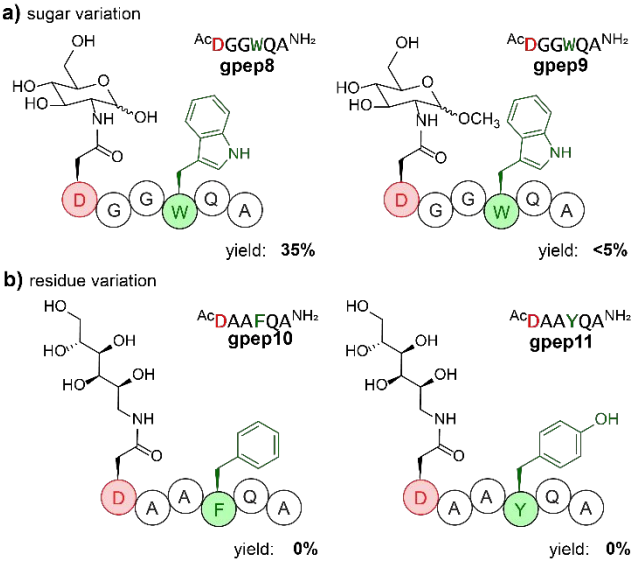



***Figure 4*** (a) assessment of two different sugars (b) phenylalanine **gpep10** and tyrosine **gpep11** studied models.

With the development of a series of rhodium-boronic acid conjugates, we demonstrate that boronic acid recognition of polyol motifs can template catalytic modification of nearby tryptophan residues. The synthesis of these heteroleptic rhodium conjugates is achieved in a straightforward manner

demonstrating chemoselectivity in the presence of free boronic acids groups. This work provides a further example of bifunctional rhodium(II) complexes and demonstrates that boronic acid dynamic covalent linkages can effectively control bioconjugation chemistry.[35,36] These experiments demonstrate that dynamic covalent chemistry of boronic acids with diols can be used to control proximity-induced catalytic modification of a model polyol-containing peptide. The methodology proved to be robust enough to succeed at quite low catalytic concentrations of rhodium under aqueous buffer conditions. The results indicate that the catalysis is quite selective with regard to the structure of the polyol unit and of the rhodium catalyst, indicating that discrimination among natural structures with subtle structural differences may be possible.

R.D.V., Y.D., H.O.T., and Z.T.B. designed the experiments, R.D.V., H.O.T., Y.D. and R.Q. conducted the experiments. R.D.V., Y.D., H.O.T., and Z.T.B. analyzed the data. R.D.V. and Z.T.B. wrote the initial draft, which was reviewed and edited by all authors.

The authors acknowledge support from the Robert A. Welch Foundation Research Grant C-1680 (Z.T.B.) and the National Science Foundation under grant numbers CHE-1904865 and CHE-2203948. The authors also thank Christopher Pennington for the assistance with MS experiments.

## Conflicts of interest

There are no conflicts to declare.

## Notes and references


1 R. Apweiler, *Biochim. Biophys. Acta BBA - Gen. Subj.*, 1999, **1473**, 4–8.
2 P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek, *Science*, 2001, **291**, 2370–2376.
3 A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, and others, .
4 J.-P. Zanetta, S. Kuchler, S. Lehmann, A. Badache, S. Maschke, D. Thomas, P. Dufourcq and G. Vincendon, *Histochem. J.*, 1992, **24**, 791–804.
5 K. Mariño, J. Bones, J. J. Kattla and P. M. Rudd, *Nat. Chem. Biol.*, 2010, **6**, 713–723.
6 R. Raman, S. Raguram, G. Venkataraman, J. C. Paulson and R. Sasisekharan, *Nat. Methods*, 2005, **2**, 817–824.
7 S. Wittrock, T. Becker and H. Kunz, *Angew. Chem. Int. Ed.*, 2007, **46**, 5226–5230.
8 M. Lajko, A. F. Haddad, C. A. Robinson and S. A. Connolly, *J. Virol. Methods*, 2015, **221**, 81–89.
9 B. V. Popp and Z. T. Ball, *J. Am. Chem. Soc.*, 2010, **132**, 6660–6662.
10 Z. Chen, B. V. Popp, C. L. Bovet and Z. T. Ball, *ACS Chem. Biol.*, 2011, **6**, 920–925.
11 F. Vohidov, J. M. Coughlin and Z. T. Ball, *Angew. Chem. Int. Ed.*, 2015, **54**, 4587–4591.
12 C. R. Bertozzi and and L. L. Kiessling, *Science*, 2001, **291**, 2357–2364.
13 J. A. Prescher and C. R. Bertozzi, *Cell*, 2006, **126**, 851–854.
14 B. Akgun and D. G. Hall, *Angew. Chem. Int. Ed.*, 2018, **57**, 13028–13044.
15 G. T. Williams, J. L. Kedge and J. S. Fossey, *ACS Sens.*, 2021, **6**, 1508–1528.
16 Z. Liu and H. He, *Acc. Chem. Res.*, 2017, **50**, 2185–2193.
17 V. L. Pecoraro, Ed., *Methods in enzymology: peptide, protein and enzyme design*, Academic Press, Amsterdam, Netherlands, 2016.
18 N. Van Den Heuvel, S. M. Mason, B. Q. Mercado and S. J. Miller, *J. Am. Chem. Soc.*, 2023, **145**, 12377–12385.
19 J. Ohata and Z. T. Ball, *J. Am. Chem. Soc.*, 2017, **139**, 12617–12622.
20 M. B. Minus, M. K. Kang, S. E. Knudsen, W. Liu, M. J. Krueger, M. L. Smith, M. S. Redell and Z. T. Ball, *Chem. Commun.*, 2016, **52**, 11685–11688.
21 Z. Chen, F. Vohidov, J. M. Coughlin, L. J. Stagg, S. T. Arold, J. E. Ladbury and Z. T. Ball, *J. Am. Chem. Soc.*, 2012, **134**, 10138–10145.
22 M. B. Minus, H. Wang, J. O. Munoz, A. M. Stevens, A. E. Mangubat-Medina, M. J. Krueger, W. Liu, M. M. Kasembeli, J. C. Cooper, M. I. Kolosov, D. J. Tweardy, M. S. Redell and Z. T. Ball, *Org. Biomol. Chem.*, 2020, **18**, 3288–3296.
23 F.-M. Siu, I. W.-S. Lin, K. Yan, C.-N. Lok, K.-H. Low, T. Y.-C. Leung, T.-L. Lam and C.-M. Che, *Chem. Sci.*, 2012, **3**, 1785.
24 R. P. Reddy and H. M. L. Davies, *Org. Lett.*, 2006, **8**, 5013–5016.
25 B. V. Popp and Z. T. Ball, *Chem. Sci.*, 2011, **2**, 690–695.
26 J. M. Antos, J. M. McFarland, A. T. Iavarone and M. B. Francis, *J. Am. Chem. Soc.*, 2009, **131**, 6301–6308.
27 S. C. Martin, F. Vohidov, H. Wang, S. E. Knudsen, A. A. Marzec and Z. T. Ball, *Bioconjug. Chem.*, 2017, **28**, 659–665.
28 G. Springsteen and B. Wang, *Tetrahedron*, 2002, **58**, 5291–5300.
29 W. L. A. Brooks, C. C. Deng and B. S. Sumerlin, *ACS Omega*, 2018, **3**, 17863–17870.
30 J. Yan, G. Springsteen, S. Deeter and B. Wang, *Tetrahedron*, 2004, **60**, 11205–11209.
31 J. M. Antos and M. B. Francis, *J. Am. Chem. Soc.*, 2004, **126**, 10256–10257.
32 R. Gibe and M. A. Kerr, *J. Org. Chem.*, 2002, **67**, 6247–6249.
33 P. Stanley, K. W. Moremen, N. E. Lewis, N. Taniguchi and M. Aebi, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, D. Mohnen, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar and P. H. Seeberger, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 4th edn., 2022.
34 B. V. Popp and Z. T. Ball, *J. Am. Chem. Soc.*, 2010, **132**, 6660–6662.
35 M. J. Swierczynski, Y. Ding and Z. T. Ball, *Bioconjug. Chem.*, 2022, **33**, 2307–2313.
36 J. P. M. António, R. Russo, C. P. Carvalho, P. M. S. D. Cal and P. M. P. Gois, *Chem. Soc. Rev.*, 2019, **48**, 3513–3536.


View Article Online
DOI: 10.1039/D3CC03371F