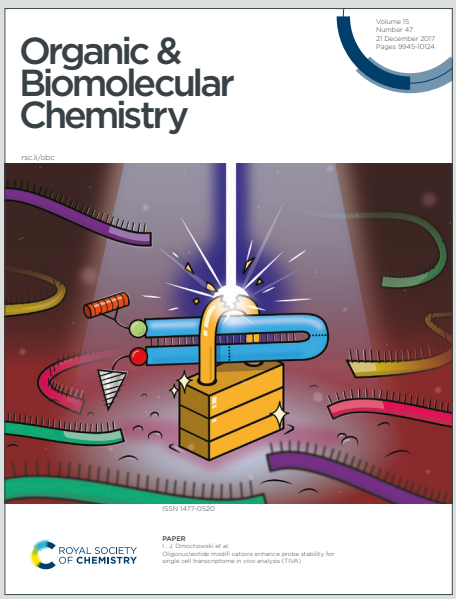
Organic & Biomolecular Chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: H. Chen, S. A. Nasseri, P. Rahfeld, J. F. Wardman, M. Kohsiek and S. G. Withers, *Org. Biomol. Chem.*, 2021, DOI: 10.1039/D0OB02484H.






ROYAL SOCIETY OF CHEMISTRY

rsc.li/obc

View Article Online
DOI: 10.1039/D0OB02484H

COMMUNICATION

# Synthesis and evaluation of sensitive coumarin-based fluorogenic substrates for discovery of α-*N*-acetyl galactosaminidases through droplet-based screening

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

DOI: 10.1039/x0xx00000x

Hong-Ming Chen†, Seyed Amirhossein Nasseri†, Peter Rahfeld, Jacob F. Wardman, Maurits Kohsiek and Stephen G. Withers*

**As part of a search for a substrate for droplet-based microfluidic screening assay of α-N-acetylgalactosaminidases, spectral and physical characteristics of a series of coumarin derivatives were measured. From among these a new coumarin-based fluorophore, Jericho Blue, was selected as having optimal characteristics for our screen. A reliable method for the challenging synthesis of coumarin glycosides of α-GalNAc was then developed and demonstrated with nine examples. The α-GalNAc glycoside of Jericho Blue prepared in this way was shown to function well under screening conditions.**

*N*-Acetyl-galactosamine (GalNAc) linked as an α-glycoside plays two especially important roles in nature. In one instance, it is found linked to a serine or threonine residue as the first sugar attached to proteins within mucin-type O-glycans, where it is typically decorated with galactose, *N*-acetylglucosamine, or another GalNAc to form the core structures of mammalian O-glycans. Its other common occurrence is as the terminal sugar on the A-antigen found on red blood cells and tissues of individuals with A blood type. In this case it is attached α-1,3 to the galactose of the H-antigen, fucose-α-1,2-galactose, which sits at the end of glycan chains on membrane-bound glycolipids or glycoproteins[1,2].

Enzymatic cleavage of terminal α-GalNAc residues is accomplished by exo-acting α-*N*-acetyl-galactosaminidases (α-GalNAcase (EC 3.2.1.49)) that are found primarily in five of the sequence-based CAZy families: GH27, GH36, GH109, GH129 and most recently GH31[3,4]. Access to an expanded range of α-GalNAcases with different specificity could be of value in several applications. One of these is in the enzymatic conversion of A-antigens to H-antigens on red blood cells as a means of creating universal donor O-type red blood cells; a topic that has attracted considerable interest[4–6]. The other would be as an analytical tool for the analysis of O-glycan structures and in determining their attachment points to proteins. It is quite probable that distinctly different enzyme classes will be required for these two applications.

As part of our quest to identify such enzymes in a high-throughput manner from within large metagenomic libraries, we required access to suitable fluorogenic α-*N*-acetylgalactosaminide substrates to be used in screening via 'water-in-oil' droplet methodologies[7]. While we have previously utilized plate-based approaches for such screening efforts, droplet-based microfluidics enables higher-throughput screening compared to plate-based assays (on the approximate scales of >$10^5$ clones per day compared to <$10^4$ respectively). Indeed, the capacity of microfluidic generated droplets for high-throughput screenings has enabled rapid and expansive directed evolution[8,9] and enzyme discovery campaigns[10,11].

Our past experiences in plate-based screening for other enzymatic activities have shown that coumarin glycosides, or more specifically the glycosides of 7-hydroxycoumarin derivatives (known as umbelliferones), are useful screening substrates.[12,13] Of these the 4-methylumbelliferones are the best known. Umbelliferone derivatives have many characteristics that make them attractive fluorophores for screening purposes. Their aromatic moieties are readily synthesised and, once attached to the anomeric centres of sugars via the 7-hydroxyl, form compact substrates that are virtually non-fluorescent. However, once the glycosidic linkage has been hydrolyzed, they are highly fluorescent in their deprotonated, anionic form. For our screening purposes, we thus needed to incorporate hydroxycoumarins with $pK_a$ values low enough that they would exist primarily in their fluorescent, anionic form when released in screening assays that are typically performed at neutral pH values.

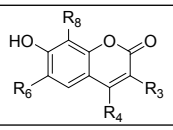
The second consideration in the design of substrates is specific for droplet-based screening methods, which involve the creation of small (<1 nL) sized 'droplets' of water within an oil matrix. Each droplet functions as a small reaction vessel containing enzyme and the fluorogenic glycoside. It is therefore

*Department of Chemistry, University of British Columbia*
*2016 Main Mall, Vancouver, B.C., Canada*
*E-mail: withers@chem.ubc.ca.*
† These authors have contributed equally to this work.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

View Article Online
DOI: 10.1039/D0OB02484U



Table **1**: Spectroscopic and chemical properties of coumarin derivatives



| # | Compound | $R_3$ | $R_4$ | $R_6$ | $R_8$ | $\lambda_{ex}/\lambda_{em}$ [a] | $pK_a$ [b] | Relative fluor. brightness [c] |
|---|---|---|---|---|---|---|---|---|
| 1 | Umbelliferone | H | H | H | H | 365/455 | 7.6 | 0.082 |
| 2 | MU | H | Me | H | H | 365/450 | 7.7 | 0.077 |
| 3 | 6-FMU | H | Me | F | H | 360/450 | 6.6 | 0.12 |
| 4 | 6,8-$F_2$MU | H | Me | F | F | 360/450 | 4.9 | 0.10 |
| 5 | 6-ClMU | H | Me | Cl | H | 365/450 | 6.1 | 0.14 |
| 6 | TFMU | H | $CF_3$ | H | H | 390/505 | 7.1 | 0.049 |
| 7 | 6,8-$F_2$TFMU | H | $CF_3$ | F | F | 390/505 | 4.3 | 0.036 |
| 8 | 3-CU | $CO_2H$ | H | H | H | 390/450 | 7.3 | 0.26 |
| 9 | 3-CUOMe | $CO_2Me$ | H | H | H | 400/450 | 6.5 | 0.68 |
| 10 | 6-Cl-3-CU | $CO_2H$ | H | Cl | H | 395/450 | 5.7 | 0.35 |
| 11 | 6-Cl-3-CUOMe | $CO_2Me$ | H | Cl | H | 410/450 | 4.8 | 0.58 |
| 12 | 8-Cl-3-CU | $CO_2H$ | H | H | Cl | 390/450 | 5.7 | 0.32 |
| 13 | 8-Cl-3-CUOMe | $CO_2Me$ | H | H | Cl | 410/450 | 4.9 | 0.49 |
| 14 | JB | $CO_2H$ | H | F | H | 390/450 | 6.0 | 0.26 |
| 15 | JBOMe | $CO_2Me$ | H | F | H | 405/450 | 5.2 | 0.45 |
| 16 | PB | $CO_2H$ | H | F | F | 390/450 | 4.7 | 0.31 |
| 17 | PBOMe | $CO_2Me$ | H | F | F | 405/450 | 3.7 | 0.54 |
| 18 | Fluorescein | - | - | - | - | 484/519 | 6.2 | 1 |

[a] Measured for 100 μM solutions of each compounds in PBS buffer containing 5% DMSO.
[b] pKa of the phenol; calculated based on measured absorbance of 50 μM solutions of the compounds in buffers with various pH values, each containing 5% DMSO.
[c] The slope of the graph of fluorescence vs concentration of compound. Measured at maximum wavelengths of each compound in PBS buffer. Values are reported relative to Fluorescein. An estimated 10% error should be considered for these values due to different purity grades of compounds from different sources.

crucial that the hydrolysed fluorophore should not be able to diffuse to the oil layer or, even worse, exchange between droplets. The mechanism behind droplet-droplet exchange is thought to be driven by encapsulation and transport of the fluorophore between droplets in surfactant formed micelles[14] and increasing the polarity of the fluorophore has been shown to limit both partitioning into the oil and micellar transport (albeit the mechanistic underpinnings of this have not been fully established).[15] As demonstrated within this text, the partial negative charge present on the phenolic oxygen of coumarins of low pKa is not sufficient to retain the molecule in the water droplets. This is further corroborated by other studies using similar fluorophores in various oil systems[14]. Our approach to this problem was to use hydroxycoumarins containing a charged carboxylate moiety that will both block entry into the oil and limit micellar transport.

Thus, in order to select a suitable fluorophore, we surveyed the many different coumarins available from commercial sources or reported in the literature. However, we found that a systematic study of their properties has not been done, resulting in some contradictory values being found in different sources[12,16–19]. We therefore examined, and report here, the spectroscopic and chemical characteristics of a collection of coumarin-based fluorophores that are relevant in the selection of an optimal substrate for screening purposes (Table 1). These include their excitation and emission wavelengths, their phenolic $pK_a$ values and their relative fluorescence brightness. These data will be very valuable for any similar study in deciding which one of the many coumarin derivatives best matches the specific assay being developed.

The most important characteristics for our purposes are the $pK_a$ of the 7-hydroxyl as well as the fluorescence intensity of the molecule under assay conditions. The fluorescence intensity is itself determined by the inherent fluorescence intensity of the molecule along with its $pK_a$. Importantly, the $pK_a$ of the phenol also correlates to the leaving group ability of the coumarin derivative and thus the reactivity/sensitivity of the substrate. High reactivity increases the probability that enzymes from metagenomic sources, which are often expressed at low levels, may be detected. However, making the substrate too reactive runs the risk of high background fluorescence in the screen from spontaneous hydrolysis of the substrates, masking the activity of these minimally expressed enzymes. Therefore, the phenolic pKa should be as low as possible but just high enough that the glycoside is stable.

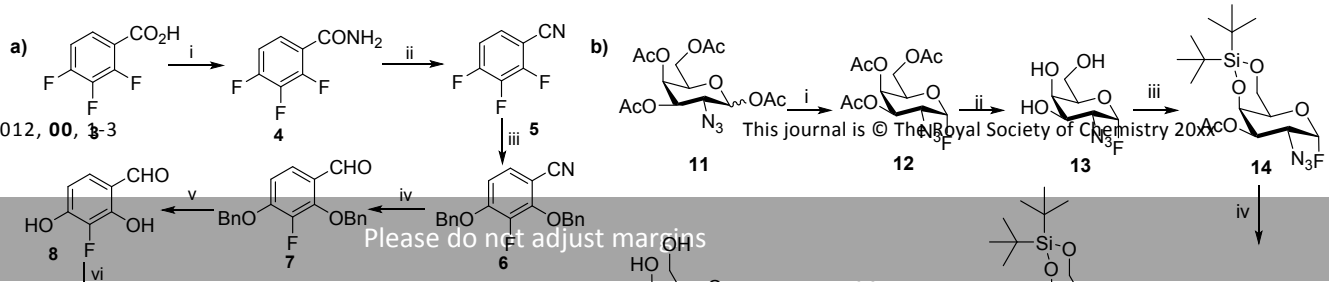
As noted in Table 1, the derivatives with a carboxylic acid group as $R_3$ are generally more fluorescent than their 4-methyl counterparts. This is particularly so if the carboxyl group is esterified. However, as shown below, since the charged carboxyl group is needed to keep the fluorophores within the droplet, the esterified versions cannot be used for our purposes. Nevertheless, for purposes other than droplet-based assays, the esterified versions have an advantage.

Finally, as these fluorophores need to fit into the enzyme active site, the smaller they are the better. Thus, even though chlorinated derivatives generally have a lower pKa and are more fluorescent compared to their fluorinated analogues (for instance see entries 3 vs 5, and 12 vs 14), the latter are preferred. Indeed, we have previously observed some families of glycosidases that do not hydrolyze 6-chloromethylumbelliferone analogues (unpublished data).

Based upon the factors discussed above, we chose to initially synthesise the α-GalNAc glycoside of the coumarin derivative known as Pacific Blue (**1**) shown in Figure 1. However, as discussed in more detail in the SI, during the course of synthesis we found out that the protected glycoside of Pacific Blue is too unstable to withstand the reaction conditions required to access the final product **1**. Not only did this make the current synthetic route infeasible but also it indicated that the final product would be too unstable to function as a useful screening reagent for assays that typically run overnight. Presumably the phenol, with its low pKa of 3.7 (as the methyl ester), is just too good a leaving group.

In light of this result we reconsidered our strategy. Clearly, a phenol of higher pKa was needed. The conceptually simplest way to reduce the leaving group ability of the phenol was to

**Fig. 1** Pacific Blue α-D-GalNAc (**1**) and Jericho Blue α-D-GalNAc (**2**)



View Article Online
DOI: 10.1039/D0OB02484H

# COMMUNICATION

**Scheme 1.** Reagents and conditions: a) i. DCM, oxalyl chloride, DMF, rt; $CHCl_3$, $NH_4OH$, 0°C. ii. $POCl_3$, 80°C. iii. DMF, BnOH, $K_2CO_3$, 105°C. iv. DCM, DIBAL in cyclohexane (1 M), -78°C. v. MeOH, THF, $H_2$, Pd/C, rt. vi. $H_2O$, Meldrum's acid, $NH_4OAc$, rt. vii. MeOH, $H_2SO_4$, refluxed, 90%. b) i). DCM, HF/Pyridine, 0°C, 93%. ii). MeOH, $NH_3$, 0°C, 95%. iii). Pyridine, di-*tert*-butylsilylbistriflate, 0°C-rt; $Ac_2O$, rt, 81%. iv). DCM, Jericho Blue.TBA salt (19c), 4Å Molecular sieves, $BF_3{\cdot}Et_2O$, 0°C. v) THF, AcOH, TBAF, rt; Pyridine, $Ac_2O$, rt, 65%; THF, $H_2O$, $Ph_3P$, silica gel, 50°C; Pyridine, $Ac_2O$, rt, 72%; MeOH, Na, rt; THF, $H_2O$, LiOH, 0°C, 69%

reduce the number of ortho fluorine atoms from two to one, as embodied in compound **2** (Figure 1). However, this fluorophore has not been reported previously and we had to first synthesize the requisite monofluorocoumarin derivative. Our approach to the synthesis of 8-fluoro-7-hydroxycoumarin-3-carboxylic acid started from the inexpensive, commercially available 2,3,4-trifluorobenzoic acid (**3**),[20] as shown in Scheme 1a. Treatment with oxalyl chloride in dichloromethane containing a catalytic amount of dimethylformamide gave the aryl chloride which, upon treatment with ammonium hydroxide, afforded amide **4**. This amide was then dehydrated by reaction with phosphorus oxychloride to afford the electron deficient nitrile **5** which, on treatment with benzyl alcohol and base underwent nucleophilic aromatic substitution at the ortho and para positions to yield bis-benzyl-substituted arene **6**. Conversion of the nitrile to the aldehyde was achieved by reduction with DIBAL at -78°C and subsequent hydrolysis of the resultant imine to give aldehyde **7**. The two benzyl protecting groups were removed by hydrogenation (Pd/C), to yield the diol **8** in good (82%) yield, along with a small amount of the by-product formed from reduction of the aldehyde, which was removed by flash column chromatography. Reaction of compound **8** with Meldrum's acid and ammonium acetate in water afforded the desired Pacific Blue analogue, monofluorocoumarin **9**, which we have named Jericho Blue in honour of Jericho beach adjacent to the UBC campus in Vancouver. The overall yield (from **3** to **9**) was a respectable 57% with only one intermediate requiring purification by flash column chromatography, the rest being accomplished by extraction and precipitation. We then proceeded to characterize this novel fluorophore as reported in Table 1. The 1.3-unit pKa difference between Jericho Blue and Pacific Blue was thought to be sufficient to make the glycoside stable enough for use in screens.

Next, we had to develop a reliable method for the α-glycosylation reaction of GalNAc with the phenols. Such syntheses of alpha-configured GalNAc derivatives are much more difficult than the synthesis of the corresponding beta glycosides, which benefit from neighbouring group participation from the 2-acetamide. However, several methods to effect stereo-controlled 1,2-cis glycosylations of this kind have been developed over the years, with varying levels of success. Such approaches have been nicely reviewed by Demchenko[21].

We initially based our approach on the method pioneered by Kiso in which a di-tert-butylsilylene protecting group installed across the 4 and 6 positions strongly directs α-glycosylation.[22,23] This is thought to be the outcome of steric blocking of the top (beta) face of the galactoside by the bulky silyl protecting group. For the glycosylation step, we first followed the method described by Kiso using the Mitsunobu reaction.[24] However, we were unable to achieve the described yields, quite likely due to the poor nucleophilicity of the low pKa coumarin derivatives as well as the insolubility of the coumarin derivatives in the appropriate solvents. After testing a number of approaches and conditions, we shifted our strategy to one involving use of a glycosyl fluoride donor in which coupling was performed under catalysis from $BF_3{\cdot}Et_2O$. Further, to improve the solubility of the coumarins we elected to convert them to their tetrabutylammonium salts (see Scheme S3). The route followed is shown in Scheme 1b and allowed for successful formation of the Jericho Blue GalNAc glycoside. Further couplings were attempted with a range of coumarin aglycones with yields typically around 70%, though two of the more electron-deficient phenols yielded only around 45%. A detailed discussion of this synthesis can be found in the SI.

Having the substrates at hand, we moved on to demonstrate that the newly synthesized fluorophore would be useful in screening metagenomic libraries in nano-droplets. First, we tested to see if the free fluorophore would leak out of the water-in-fluorinated oil droplets and exchange with adjacent droplets. As depicted in Figure S1, only fluorophores that contain a charged carboxylate are retained within the droplets upon extended incubation. The retention of the fluorescence

View Article Online
DOI: 10.1039/D0OB02484H

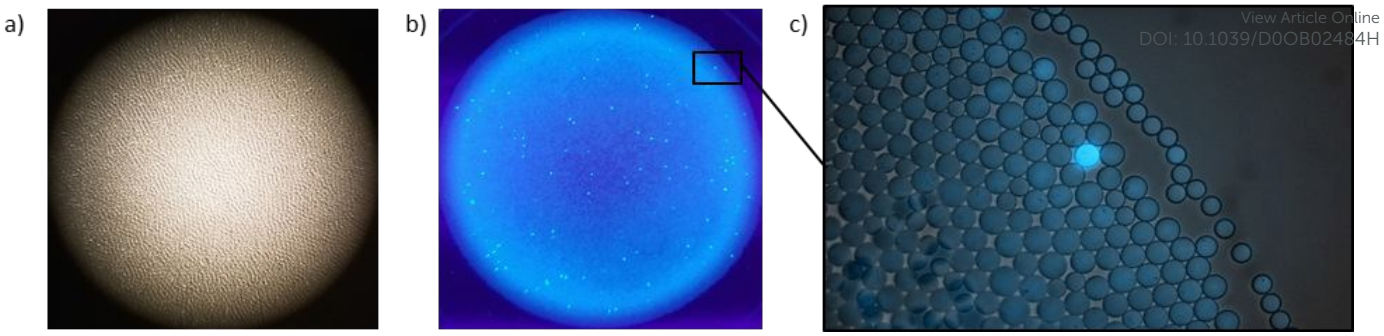



**Fig 2**. Droplets after incubation overnight (16 h) at 37°C, **a)** under bright light and **b)** UV light (365 nm). The droplets were loaded with enzyme-active and inactive *E. coli* in a 1:1000 ratio. Using a fluorescence microscope **c)** a single droplet containing cleaved substrate can be identified, showing clear separation from the neighboring droplets

and lack of crosstalk between droplets thus maintains the genotype-phenotype linkage required for successful screening. These leakage results are largely consistent with those previously observed for similar 7-aminocoumarin derivatives in the same perfluorinated oil system[15]. Next we performed a set of mock screens where two types of *E. coli* cells were encapsulated within aqueous droplets: those bearing an empty plasmid and those with the plasmid for expression of an α-GalNAcase at a ratio of 1000:1 respectively. The aqueous droplets, suspended in perfluorinated oil, contained the appropriate growth medium for *E. coli* cells along with 50 μM of the substrate **2**. As shown in Figure 2, we observed that approximately 1 in each 1000 droplets were fluorescent, and that the fluorescence remained stable over time for at least three days with no observable leakage of fluorophore to the neighbouring droplets. This experiment was repeated using the less expensive and more readily available mineral oil and similar results were observed (Figure S2). Therefore, substrates that are based upon Jericho Blue appear to be ideal for droplet-based screening of metagenomic libraries

In conclusion, we have synthesized a novel 3-carboxycoumarin, Jericho Blue, that is tailored for water-in-oil droplet-based screening of glycosidases. We have also devised a reliable synthetic method for the challenging stereoselective α-glycosylation of 2-acetamido sugars with these coumarin derivatives. Finally, we demonstrated the applicability of Jericho Blue glycosides for droplet-based screening approaches.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgments

We thank the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes for Health Research for financial support. We also thank Dr Ryan Sweeney for early suggestions on the synthesis, Soroush Nasseri for his kind gift of the microdroplet chips, and Dan Seale and Dr. Sachia Traving for assistance in acquiring the microscope images.

## Notes and references

[1] M. E. Taylor, K. Drickamer, *Introduction to Glycobiology*, Oxford University Press., **2011**.

[2] A. Varki, *Essentials of Glycobiology*, Old Spring Harbor Laboratory Press, **2017**.

[3] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **2013**, *42*, D490–D495.

[4] P. Rahfeld, J. F. Wardman, K. Mehr, D. Huff, C. Morgan-Lang, H.-M. Chen, S. J. Hallam, S. G. Withers, *J. Biol. Chem.* **2019**, *294*, 16400–16415.

[5] Q. P. Liu, G. Sulzenbacher, H. Yuan, E. P. Bennett, G. Pietz, K. Saunders, J. Spence, E. Nudelman, S. B. Levery, T. White, et al., *Nat. Biotechnol.* **2007**, *25*, 454–464.

[6] P. Rahfeld, L. Sim, H. Moon, I. Constantinescu, C. Morgan-Lang, S. J. Hallam, J. N. Kizhakkedathu, S. G. Withers, *Nat. Microbiol.* **2019**, *4*, 1475–1485.

[7] A. Zinchenko, S. R. A. Devenish, B. Kintses, P.-Y. Colin, M. Fischlechner, F. Hollfelder, *Anal. Chem.* **2014**, *86*, 2526–2533.

[8] A. Debon, M. Pott, R. Obexer, A. P. Green, L. Friedrich, A. D. Griffiths, D. Hilvert, *Nat. Catal.* **2019**, *2*, 740–747.

[9] F. Gielen, R. Hours, S. Emond, M. Fischlechner, U. Schell, F. Hollfelder, *Proc. Natl. Acad. Sci.* **2016**, *113*, E7383–E7389.

[10] P.-Y. Colin, B. Kintses, F. Gielen, C. M. Miton, G. Fischer, M. F. Mohamed, M. Hyvönen, D. P. Morgavi, D. B. Janssen, F. Hollfelder, *Nat. Commun.* **2015**, *6*, 10008.

[11] M. Hosokawa, Y. Hoshino, Y. Nishikawa, T. Hirose, D. H. Yoon, T. Mori, T. Sekiguchi, S. Shoji, H. Takeyama, *Biosens. Bioelectron.* **2015**, *67*, 379–385.

[12] H.-M. Chen, Z. Armstrong, S. J. Hallam, S. G. Withers, *Carbohydr. Res.* **2016**, *421*, 33–39.

[13] S. A. Nasseri, L. Betschart, D. Opaleva, P. Rahfeld, S. G. Withers, *Angew. Chemie Int. Ed.* **2018**, *57*, 11359–11364.

[14] A. Autour, M. Ryckelynck, *Micromachines* **2017**, *8*, 128.

[15] G. Woronoff, A. El Harrak, E. Mayot, O. Schicke, O. J. Miller, P. Soumillion, A. D. Griffiths, M. Ryckelynck, *Anal. Chem.* **2011**, *83*, 2852–2857.

[16] C. Briciu-Burghina, B. Heery, F. Regan, *Analyst* **2015**, *140*, 5953–5964.

[17] W. C. Sun, K. R. Gee, R. P. Haugland, *Bioorganic Med. Chem. Lett.* **1998**, *8*, 3107–3110.

View Article Online
DOI: 10.1039/D0OB02484H

[18] B. Abrams, Z. Diwu, O. Guryev, S. Aleshkov, R. Hingorani, M. Edinger, R. Lee, J. Link, T. Dubrovsky, *Anal. Biochem.* **2009**, *386*, 262–269.
[19] C. Dubiella, H. Cui, M. Groll, *Angew. Chemie - Int. Ed.* **2016**, *55*, 13330–13334.
[20] M. M. Lee, B. R. Peterson, *ACS Omega* **2016**, *1*, 1266–1276.
[21] S. S. Nigudkar, A. V Demchenko, *Chem. Sci.* **2015**, *6*, 2687–2704.
[22] A. Imamura, H. Ando, S. Korogi, G. Tanabe, O. Muraoka, H. Ishida, M. Kiso, *Tetrahedron Lett.* **2003**, *44*, 6725–6728.
[23] A. Imamura, A. Kimura, H. Ando, H. Ishida, M. Kiso, *Chem. - A Eur. J.* **2006**, *12*, 8862–8870.
[24] A. Imamura, H. Ando, H. Ishida, M. Kiso, *Org. Lett.* **2005**, *7*, 4415–4418.

View Article Online
DOI: 10.1039/D0OB02481



Synthesis of sensitive coumarin α-GalNAc glycosides as substrates for droplet-based screening of GalNAcases