UNIVERSITY OF CALIFORNIA,
IRVINE

Molecular Reporters for Mammalian Sugars via Hydrogen Bonding Auxiliaries

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Chemistry

by

Yuwen Wang

Thesis Committee:
Associate Professor Seunghyun Sim, Chair
Associate Professor Joe P. Patterson
Professor Zhibin Guan

2026

DEDICATION

To

Everyone who helped on the way

TABLE OF CONTENTS

LIST OF FIGURES

Page

ACKNOWLEDGMENTS

Prof. Seunghyun Sim directed and supervised research which forms the basis for the thesis. Thank you for always being very supportive throughout the past three years.

I am grateful to my thesis committee members, Prof. Patterson and Pro. Guan, for their valuable feedback, thoughtful suggestions, and time devoted to reviewing this work.

I would also like to thank the past and present members of the Sim lab for their assistance, insightful discussions, and for creating a supportive research environment.

Thank my coauthors: Chad Moorman, Dipankar Barpuzary, Melody Sun, Duyen Ta, Matthew E. Griffin, and Seunghyun Sim for their contributions in this study.

Financial support was provided by the University of California, Irvine. NIH Grant R35GM150770 and DP2AT013749. Thank Dr. Dmitry Fishman for his help with fluorescence characterization of our PBA probes. Thank the support of the Chao Family Comprehensive Cancer Center Optical Biology Core Shared Resource, supported by the National Cancer Institute of the National Institutes of Health under award number P30CA062203. Thank Professor Jennifer Prescher for the access to the HPLC instrument and constructive comments and discussion.

I would also like to thank my girlfriend, Celine Feng, for supporting me through my graduate study. My parents, Yidong Wang and Xin Rong, for their unconditional love, encouragement, and unwavering support throughout my academic journey. My cats, Lili and Hihi, for the emotional support and companionship throughout the years.

ABSTRACT OF THE THESIS

Molecular Reporters for Mammalian Sugars via Hydrogen Bonding Auxiliaries

by
Yuwen Wang
Master of Science in Chemistry
University of California, Irvine, 2026
Associate Professor Seunghyun Sim, Chair

Glycans are essential transmitters of biological information. However, their complex and dynamic structures remain difficult to characterize, posing a major barrier to fully understanding their biological functions. Here, we demonstrate that hydrogen bonding auxiliaries positioned adjacent to the boron site of turn-on fluorogenic phenylboronic acid reporters serve as key modulators of sensitivity to mammalian sugars, providing a new design principle analogous to the hydrogen bonding networks found in lectin binding pockets. We show that these auxiliaries tune sensitivity both globally and for individual mammalian sugars and reveal a striking differential turn-on response between the anomeric forms of monosaccharides. The versatility of these selective and reversible molecular reporters is demonstrated through in-gel glycoprotein imaging, flow cytometry, and cell-surface visualization, establishing them as a generalizable chemical platform for interrogating mammalian glycans.

INTRODUCTION

Glycans are information-relaying molecules for cellular signaling and communication across all domains of life.[1-5] They are also often covalently attached to other structural biomolecules, such as proteins and lipids.6 Importantly, glycan structures are not directly templated by the genome; instead, their synthesis and modification are controlled by a network of highly regulated enzymes that carry out non-stoichiometric reactions, resulting in variable modifications across substrates.[7, 8] Therefore, the glycophenotype of a cell is highly context-dependent and can change with cellular state and microenvironment. [5, 7-9] For example, cancer cells often display changes in their glycosylation patterns, such as heightened levels of sialylation and highly branched N-linked glycans.[9] Despite their functional importance, the molecular codes of glycans for conveying biological information are still not well understood due to difficulties associated with characterizing their structure and dynamic changes.[3]

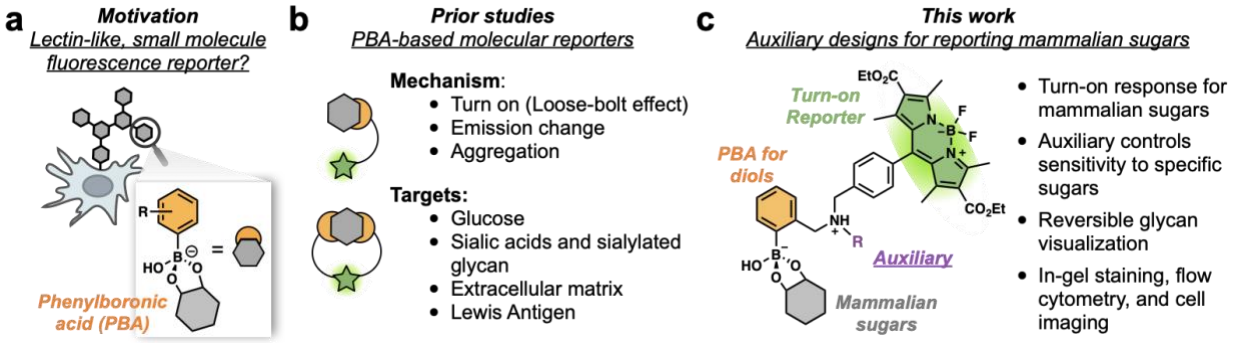


Figure 1. Schematic illustration of the motivation, relevant prior work, and summary of this project. (**a**) Motivation of this and other studies in developing phenylboronic acid (PBA)-based fluorogenic molecular probes for glycans. (**b**) Prior studies in this area. (**c**) Auxiliary designs for reporting mammalian sugars.

Profiling of glycan structures by imaging is critical for mechanistic studies of cellular processes and diseases. Nature utilizes many types of glycan-binding proteins, including lectins, to decipher the molecular code of glycans.[10, 11] These proteins exhibit varying specificity and

affinity to structural motifs in glycan and have become indispensable tools for glycan imaging and profiling.[12, 13] However, challenges with fluorescently labeled lectins include cross-reactivity, a limited repertoire of commercially available and fully characterized lectins, and batch-to-batch variability.[14] Expanding reporting tools for carbohydrates is critical to complement existing lectin-based technologies and accelerate new discoveries in glycobiology.

An appealing strategy for selectively labeling glycans is to use reversible small-molecule recognition with phenylboronic acid (PBA), which forms a dynamic covalent bond with available diols on carbohydrates (Figure 1a).[15] Low pKa PBAs, such as Wulff-type ortho-aminomethyl PBAs, bind to a range of diol substrates at neutral pH and have thus been explored for their utility in reporting glucose and other biological sugars (Figure 1b).[16-24] Fluorogenic PBA reporters with turn-on response (i.e., increased emission upon binding) are especially attractive for high-contrast imaging. The seminal work by James, Shinkai, and others on turn-on ortho-aminomethyl PBA fluorophores responsive to glucose[16-20] provided the conceptual foundation for many sugar-responsive fluorogenic PBA sensors.[24] A mechanistic study by Anslyn, James, and Sun has formulated a unifying theory that -OH on PBA functions as the energy acceptor, and arresting this "loose-bolt" by diol binding is the reason for the observed turn-on response of ortho-aminomethyl PBA upon binding to its diol substrate.[25, 26] The utility of PBA-based fluorogenic reporters has been demonstrated in increasingly demanding biological contexts, such as reversible glucose imaging in live cells and zebrafish,[22] broad glycan-binding dyes that enable wash-free extracellular matrix visualization,[27] and fluorescence imaging of sialylated glycans in cancer tissue.[28] However, our understanding of the relationship between the molecular structure of reporters and their sensitivity to specific carbohydrates remains incomplete.

In this thesis, we report that hydrogen bonding auxiliaries in turn-on fluorogenic PBA reporters control the sensitivity for neutral mammalian sugars (Figure 1c). The selectivity of lectin-carbohydrate interactions is controlled by the specific hydrogen bonding patterns formed within the lectin's binding pocket.[29] We hypothesized that controlling the chemical environment adjacent to the boron site of PBA can provide adjustable sensitivity to certain sugars through hydrogen bonding. After identifying an appropriate base scaffold for turn-on response based on its photophysical properties and synthetic scalability, we showed that auxiliaries adjacent to PBA control sensitivity globally and for individual mammalian sugars. We also observed a striking difference in the turn-on response between the different anomeric forms of monosaccharides. The utility of these selective and reversible molecular reporters for in-gel glycoprotein imaging, flow cytometry, and cell-surface visualization showcases them as a generalizable chemical tool for interrogating mammalian glycans.

CHAPTER 1: PBA-based probe design and characterization

1. Base scaffolds and their intrinsic property

We designed and synthesized a series of ortho-aminomethyl PBA appended to boron dipyrromethene (BODIPY) as base scaffolds (Figure 2a, Figures S2.1–2.3). Wulff-type ortho-aminomethyl PBAs are known to possess low pKa values, making them well suited for diol recognition under physiological conditions.[15] Using the base scaffolds **A–C**, we assessed their intrinsic photophysical properties by incubating the compounds in aqueous solution (50% DMSO in phosphate-buffered saline (PBS)) with 0.1 M sorbitol, a strong binder of general PBA, for 1 h. We note that the high concentration of DMSO in this initial photophysical characterization is to avoid any effect from aggregation. All probes exhibited similar photophysical characteristics, with absorption maxima around 500 nm and emission maxima near 515 nm (Figures S3). Although A displayed a slightly red-shifted spectrum relative to **B** and **C**, it did not show a fluorescence turn-on response (Figure S3) and exhibited low quantum yield in the presence of sorbitol (Figure 2a). In contrast, **B** and **C** displayed pronounced fluorescence turn-on responses upon incubation with sorbitol (Figure 2b), with relatively high quantum yields in the presence of sorbitol (Figure 2a). Titration of **C** with sorbitol (1% DMSO in PBS) revealed a concentration-dependent increase in fluorescence, reaching a plateau (Figure 2c). We note that the fluorescence turn-on response was stronger in a more aqueous solution (1% DMSO in PBS) than in a mixture containing a higher DMSO concentration, suggesting a solubility effect. The observed pKa values from absorbance measurements for **A–C** were all below 6 (Figures 2a), indicating reversible diol binding at physiological pH. **B** and **C** were generally comparable in photophysical properties as well as turn-on response with sorbitol, making them suitable as a base scaffold for our study.

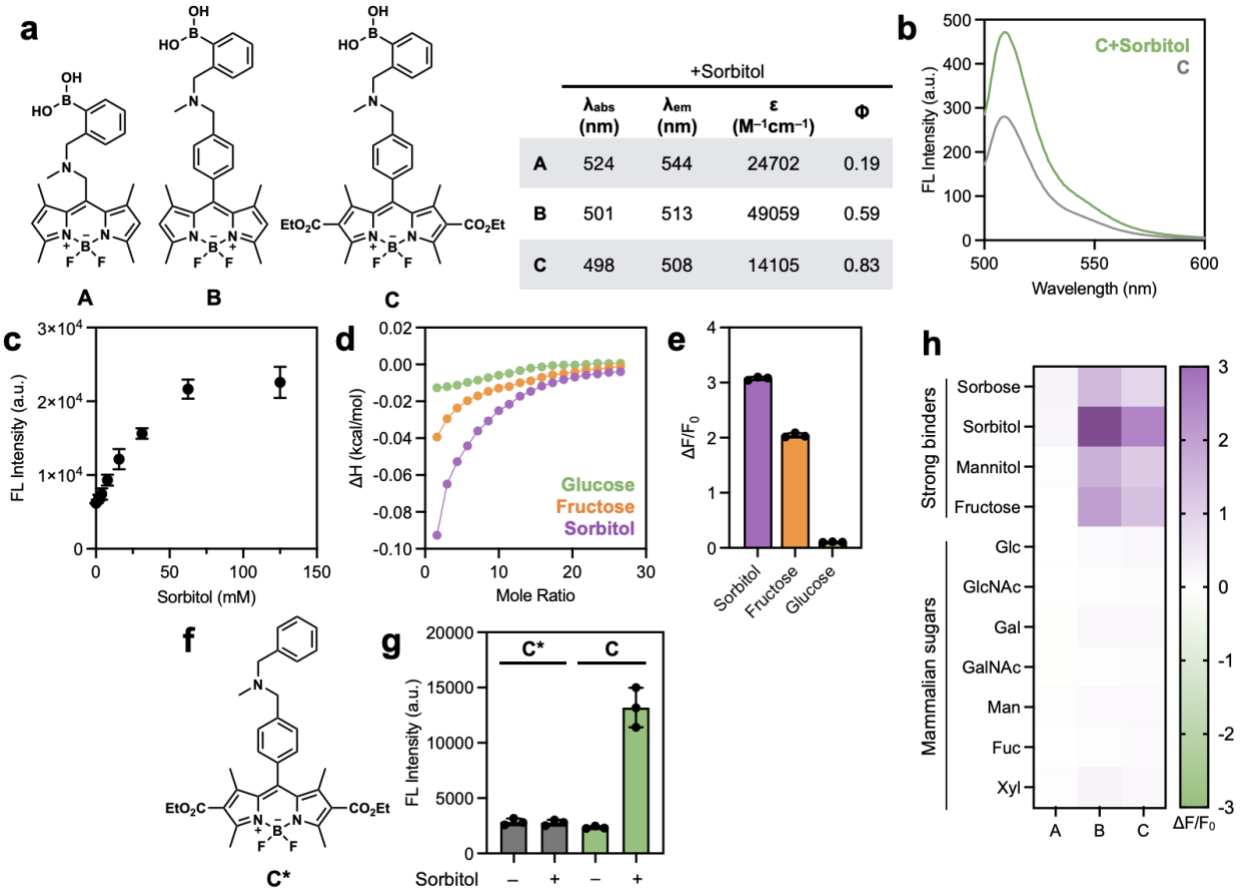


Figure 2. Base scaffolds and their intrinsic properties. (**a**) Structures of the three base scaffolds (**A–C**) based on ortho-aminomethyl PBA and their intrinsic photophysical property in 1:1 DMSO/PBS containing 0.1 M D-sorbitol. (**b**) Fluorescent emission spectrum of compound **C** in 1:1 DMSO/PBS incubated with (green) or without (grey) 0.1 M D-sorbitol. (**c**) Fluorescent response of **C** (20 μM) in PBS (1% DMSO) containing different concentrations of D-sorbitol. (**d**) Isothermal titration calorimetry binding curve obtained from 0.5 mM **B** with 70 mM of diol species. (**e**) fluorescent turn-on response ($\Delta F/F_0$) of **B** (20 μM) with diol species (100 mM) in PBS (1% DMSO). (**f**) Structure of the control fluorescent molecule **C***. (**g**) Fluorescent turn-on response in 100 mM D-sorbitol with **C*** and **C**. (**h**) A heat map summarizing monosaccharide library and fluorescent turn-on response of **A–C** in PBS (1% DMSO). The error bars represent standard deviation in three technical replicates.

To investigate whether the observed fluorescence enhancement correlates with PBA-diol binding, we performed an isothermal titration calorimetry (ITC) assay using **B** with representative, well-characterized diol analytes (sorbitol, fructose, and glucose) (Figure 2d, Figure S6). We note that B (10% DMSO in PBS) was employed in this study because of the high solubility required for reliable ITC measurements. The measured association constants (K_a) were 239.8 M^{-1} for sorbitol, 172.1 M^{-1} for fructose, and 51.5 M^{-1} for glucose. The observed trend is consistent with

previously reported values.[15, 30] The trend observed in the increase in fluorescence emission using a plate reader ($\Delta F/F_0$; $\Delta F = F - F_0$, where F_0 is the fluorescence intensity of the probe without diol and F is the fluorescence intensity with diol) (Figure 2e) supports a positive correlation between PBA-diol binding and fluorescence turn-on response.

To further validate that the observed fluorescence response arises from binding of the phenylboronic acid (PBA) moiety to diol species, we synthesized a control dye (**C***) that retains the same fluorogenic BODIPY scaffold as C but lacks the phenylboronic acid handle (Figure 2f, Figure S2.4). **C** and **C*** (20 μM) were incubated with sorbitol (100 mM) in PBS (1% DMSO, pH 7.4) for 20 min at 37 °C (Figure 2g). **C*** showed no significant change in fluorescence, in contrast to **C**, supporting that the fluorescence enhancement arises from PBA–diol complex formation. Incubation of **C** with sorbitol in HEPES (pH 7.2) and Tris buffer (pH 8.2) produced comparable fluorescence enhancements, indicating that the observed turn-on response is robust across commonly used buffer systems (Figure S3). We also observed an immediate fluorescence turn-on response upon mixing (Figure S5).

We next quantified the ability of **A–C** to report on well-established strong binders as well as seven neutral mammalian monosaccharides – glucose (Glc), N-acetylglucosamine (GlcNAc), galactose (Gal), N-acetylgalactosamine (GalNAc), mannose (Man), fucose (Fuc), and xylose (Xyl) – via fluorescence turn-on response. Well-established strong binders of PBA (sorbitol, mannitol, fructose, and sorbose) were included as positive controls. These assays were conducted with **C** (20 μM) and diol substrate (100 mM) in PBS (1% DMSO, pH 7.4). Fluorescence was measured using a plate reader after 20 min of incubation at 37 °C (Figure 2h, Figure S5). **B** and **C** exhibited robust fluorescence turn-on responses in the presence of known strong binders, whereas **A** showed no significant fluorescence changes across all monosaccharides in the library. The relative

fluorescence increment for the mammalian sugars was similar between **B** and **C**. While both **B** and **C** generated a turn-on response upon binding to diol species, we concluded that the relatively high overall isolation yield of **C** (13.9%) compared to **B** (3.1%) made it more suitable for further derivatization.

2. Hydrogen bonding auxiliaries control the detection sensitivity globally and for individual sugars

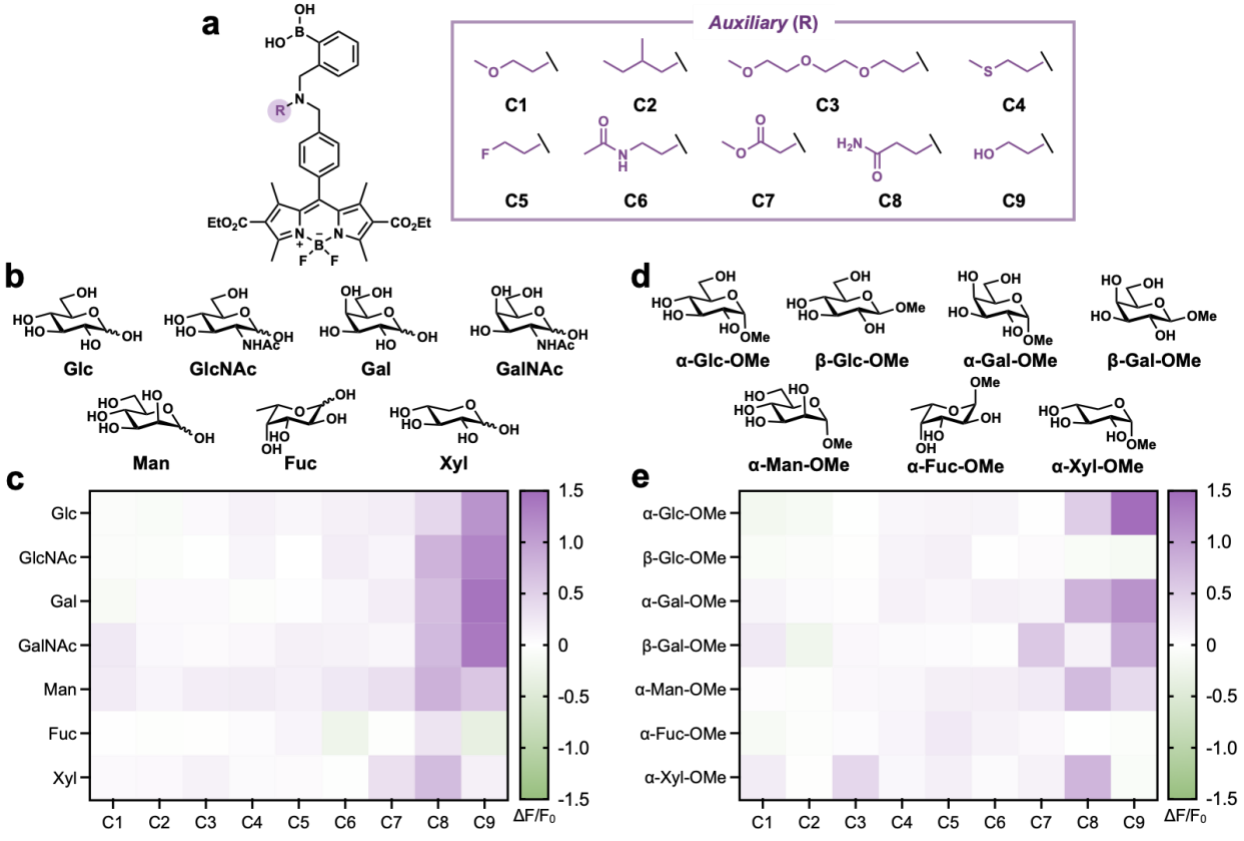


Figure 3. Hydrogen bonding auxiliaries control the detection sensitivity globally and for individual sugars. (**a**) Structures of derivatives **C1**–**C9** through auxiliary modifications on base scaffold **C**. (**b**) Structures of seven neutral mammalian monosaccharides involved in screening. All of these sugars are in D-form except for Fucose. (**c**) A heat map summarizing the fluorescent turn-on response of **C1**–**C9** (20 μM) with monosaccharides (100 mM) in PBS (1% DMSO). (**d**) Structures of methylated pyranoside derivatives involved in screening. (**e**) A heat map summarizing the fluorescent turn-on response of **C1**–**C9** (20 μM) with 100 mM methylated pyranosides (100 mM) in PBS (1% DMSO).

Lectin-carbohydrate binding is controlled by multiple noncovalent interactions within the binding pocket. Inspired by this, we sought to incorporate a range of auxiliaries to the base scaffold

C and synthesized **C1–C9** (Figure 3a, Figures S2.5–2.6 and S4). These auxiliaries include motifs derived from amino acid residues commonly found in lectin-binding pockets and represent various noncovalent interactions. We hypothesized that these auxiliaries can tune the steric, hydrogen bonding, and polar environments proximal to the boronic acid, thereby influencing the fluorescence readout sensitivity to a range of mammalian sugars.

We first tested the turn-on response of **C1–C9** upon binding to seven neutral mammalian monosaccharides (Figures 3b and 3c, Figures S7). We excluded acidic substrates, such as glucuronic acid (GlcA), to avoid the effects of negative charge at physiological pH. We incubated our probes (**C1–C9**, 20 μM) with these monosaccharides (100 mM) in PBS (1% DMSO) for 20 minutes at 37 °C and measured the fluorescence with a plate reader. Among the nine probes, **C1–C7** exhibited weak or near-baseline fluorescence responses for all substrates. Notably, **C8** with an amide substituent and **C9** with a hydroxyl substituent exhibited globally increased sensitivity to most of the monosaccharide substrates compared to **C** and **C1–C7** (Figure 3c, Figure S7). Compared to the base scaffold **C** (ΔF/F0 = 0.16), **C9** exhibited significant fluorescence enhancement (ΔF/F0 = 1.47) in the presence of D-galactose, an approximately 9-fold improvement in sensitivity upon installing a relatively simple functionality with a hydroxyl group. It is also notable that **C1** and **C9** differ by only one methyl group, yet this small structural variation leads to a remarkable difference in their fluorescence turn-on performance. While the relative sensitivity of **B** and **C**'s turn-on response for each substrate was generally similar (Figure 2h), **C8** and **C9** showed markedly different responses for individual monosaccharides. For example, while **C9** shows robust turn-on for Gal and GalNAc, it shows a weak response to Xyl. On the other hand, **C8** shows a moderately strong response to many sugars in the library, including Man and Xyl. These results suggest that the auxiliaries in these probes control the fluorescence turn-on response

globally and to specific sugars. We suspect the amino acid-like, hydrogen bonding functionalities in **C8** and **C9** may facilitate hydrogen bonding interactions with nearby hydroxyl groups of saccharides or with ether oxygens in glycosidic linkages, potentially stabilizing the boronated ester complex. Such hydrogen bonding interactions are commonly observed in lectin–carbohydrate recognition, which contribute to the affinity and specificity of sugar binding.[29] Changes in the photophysical properties of the bound complex, including its fluorescence quantum yield, may also contribute to the observed enhancement. Further thermodynamic and photophysical studies will be required to distinguish these effects. We also tested N-Acetyl-β-neuraminic acid methyl ester with our library and observed a modest turn-on response (Figure S7.10).

Reducing monosaccharides exist in an equilibrium mixture of α- and β-anomers, pyranose and furanose, and minor open-chain forms in aqueous solution. In contrast, monomers in glycans are linked via glycosidic bonds at the anomeric carbon, which locks connected sugar residues in a fixed α or β diastereomer. To understand the auxiliary effect on diastereomerically locked substrates, we employed methylated pyranoside derivatives (Figure 3d). The methyl substitution to the OH on the anomeric carbon prevents mutarotation and fixes the monosaccharide in a defined α- or β-pyranose form. Strikingly, both **C8** and **C9** showed a preference for the α-anomer of Glc over its β-anomer (Figure 3e, Figures S7). **C8** also showed α-anomer preference for Gal, whereas such a difference was not observed in **C9** (Figure 3e, Figures S7). Further investigation will be necessary to establish the mechanistic basis of this anomeric selectivity. We note that employing computational approaches to rationalize changes in fluorescence turn-on response sensitivity is particularly challenging in highly flexible, solvated systems with multiple potential interaction sites, as comprehensively sampling the full ensemble of thermally accessible conformations is computationally demanding, and each conformation can differentially influence the fluorescence

response. Together, these results showcase the striking effect of hydrogen bonding auxiliaries in Wulff-type PBA on the sensitivity of the fluorescence turn-on response to certain sugars.

3. Recognition of complex polysaccharides

Encouraged by the performance of **C9**, we next assessed its ability to recognize complex carbohydrate structures, including polysaccharides and a mixture of glycoproteins. Dextran is a branched polysaccharide composed of a glucose backbone linked primarily through α-1,6-glycosidic bonds with α-1,3 branches. Different molecular weights of Dextran (10 kDa and 40 kDa) were prepared at 5% and 10% (w/v) in PBS and incubated with **C9** (20 μM) for 20 min at 37 °C, after which fluorescence measurements were recorded (Figure 4a and Figure S8). Similar to the results with Glc, or the α-anomer of Glc (Figures 3c and 3e), we observed a robust fluorescence turn-on response of **C9** with Dextran. A higher fold change was observed with a larger Dextran (40 kDa) at the same w/v concentration, suggesting a potential multivalent effect. **C9** exhibited a fluorescence turn-on response in mannan solutions, whereas no appreciable increase in fluorescence intensity was observed with heparan sulfate or chondroitin sulfate (Figure S8).

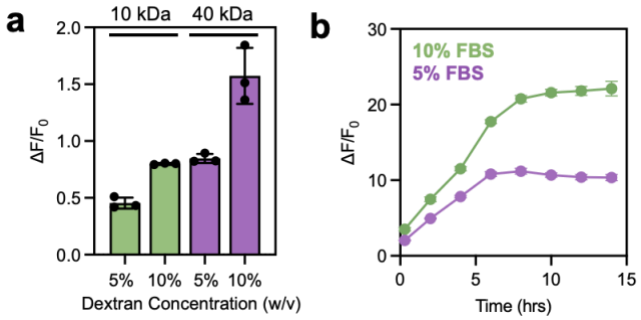


Figure 4. Recognition of complex polysaccharides. (**a**) Fluorescent turn-on response of C9 (20 μM) incubated with dextran (10 kDa or 40 kDa) in PBS (1% DMSO). (**b**) Fluorescent turn-on response over time of C9 (20 μM) with fetal bovine serum (FBS, 5 or 10%) in DMEM. The error bars represent standard deviation in three technical replicates.

These results prompted us to investigate whether **C9** could operate with a mixture of glycoproteins in a common medium. We prepared Dulbecco's Modified Eagle Medium (DMEM)

supplemented with 5% or 10% fetal bovine serum (FBS). FBS contains several glycoproteins, such as fetuin, which possess multiple glycosylation sites for both N- and O-linked glycans. DMEM containing 5% or 10% FBS was incubated with **C9** (20 μM), and fluorescence intensity was monitored over time (Figure 4b and Figure S9). In DMEM alone, the background **C9** fluorescence decreased gradually (Figure S4). In contrast, DMEM supplemented with FBS remained strongly fluorescent ($\Delta F/F_0$ up to 22), with the signal reaching a plateau approximately around 6 h (Figure 4b). These data show that **C9** can report complex carbohydrates in a polymeric and glycoprotein form.

CHAPTER 2: Application of PBA-based probes

1. Selective and reversible in-gel detection of glycoproteins

The robust fluorescence turn-on response of **C9** (Figures 3 and 4) prompted us to investigate its utility for in-gel glycoprotein detection. Current methods (e.g., Pro-Q™ reagents) rely on periodate-mediated oxidation of cis-diols to generate aldehyde functionalities. This oxidation step chemically modifies native glycans, thereby altering their natural structures. Furthermore, the subsequent dye conjugation step is typically irreversible, precluding downstream characterization of glycoproteins. We envisioned that the reversible PBA-diol binding could be further leveraged to enable complementary analyses without sacrificing sample integrity after destaining, such as total protein quantification and western blotting.

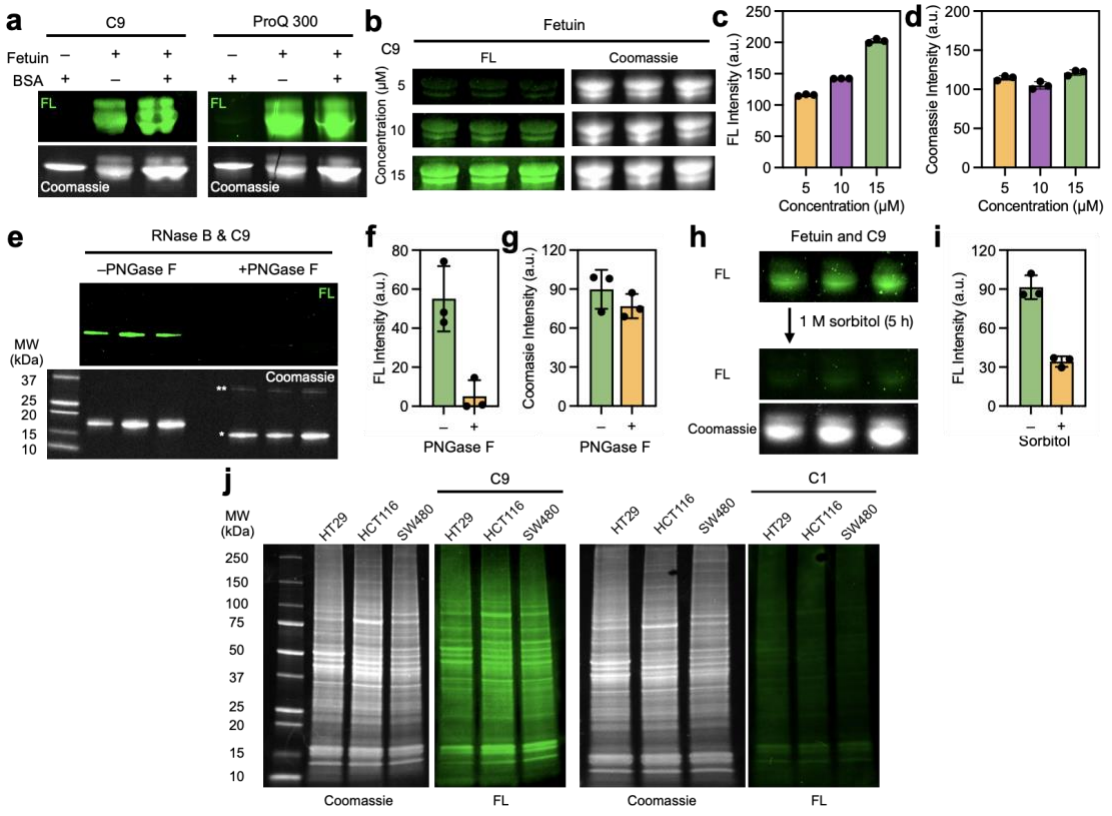


Figure 5. Selective and reversible in-gel detection of glycoproteins. (**a**) Fluorescent and Coomassie-stained images of fetuin (3 μg) and BSA (3 μg) on SDS-PAGE gel with **C9** (10 μM) or ProQ 300. (**b**) Fluorescent and Coomassie-stained images of fetuin (3 μg) on SDS-GAGE gel with various concentrations

of **C9**. (**c**) Quantification of fluorescence intensity per lane on SDS-GAGE gel with fetuin (3 μg) and various concentrations of **C9**. (**d**) Quantification of Coomassie intensity per lane on SDS-GAGE gel with fetuin (3 μg) and various concentrations of **C9**. (**e**) Fluorescent and Coomassie-stained images of RNase B with or without PNGase F treatment on SDS-PAGE gel with **C9** (10 μM). *Deglycosylated RNase B and **PNGase F. (**f**) Quantification of fluorescence intensity from RNase B per lane on SDS-GAGE gel with or without PNGase F treatment. (**g**) Quantification of Coomassie intensity from RNase B per lane on SDS-PAGE gel with or without PNGase F treatment. (**h**) **C9**-labeled fetuin gel treated with sorbitol (1 M and 5 h) exhibited a significant loss in fluorescent intensity. (**i**) Quantification of fluorescence intensity from fetuin on SDS-PAGE gel per lane before and after sorbitol treatment. (**j**) Fluorescent and Coomassie-stained images of the cell lysates from HT29, HCT116, and SW480 cancer cells (15 μg of total protein per lane) using **C9** and **C1** (10 μM). The error bars represent standard deviation in three technical replicates. All uncropped gel images with a molecular weight ladder are provided in the supporting information.

Fetuin was chosen as the model glycoprotein for this study. It contains multiple glycosylation sites (N-linked and O-linked) and is typically a heterogeneous mixture of distinct glycoforms.[31] We selected bovine serum albumin (BSA), which is generally considered a non-glycosylated protein, as a negative control. Following sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS–PAGE) separation and gel fixation, the gel was first equilibrated in PBS (pH 7.4) to restore the appropriate pH, and subsequently incubated with **C9** (10 μM in PBS) for fluorescence staining. Fetuin exhibited clear labeling by **C9**, whereas no fluorescence signal was observed with BSA (Figure 5a and Figure S10). A similar staining pattern was observed with a commercial glycoprotein detection reagent, the Pro-Q Emerald 300 staining kit (Figure 5b and Figure S10), validating our data with **C9**. A C9 concentration-dependent increase in fluorescence intensity on the gel (Figures 5b–d and Figure S10) further supports binding and fluorescence activation of **C9** in the presence of glycoproteins.

To verify that **C9** binds specifically to glycans rather than nonspecifically associating with other regions of the glycoprotein, we employed RNase B, a well-established high-mannose glycoprotein containing a single N-linked glycosylation site at Asn34.[32] Treatment of this protein with PNGase F, an enzyme that selectively cleaves N-linked glycans, will remove the glycan structure altogether and abolish fluorescent labeling. Consistent with this hypothesis, treatment of RNase B with PNGase F resulted in loss of fluorescence signal and shift of protein band in

Coomassie-stained gel (Figures 5e–g and Figure S10). These results indicate that **C9** specifically stains glycan structures.

To assess the reversibility of **C9**-based glycoprotein labeling, **C9**-stained fetuin gels were incubated with 1 M sorbitol in PBS for 5 h, after which a significant loss of fluorescence signal was observed (Figures 5h and 5i, Figure S11). In contrast, control samples incubated in PBS alone exhibited no appreciable decrease in fluorescence (Figure S11), confirming that the observed signal loss was attributable to the competing diol rather than passive destaining. This fluorescence reduction is attributable to the competitive displacement of **C9** from the glycoprotein by excess sorbitol, consistent with the dynamic covalent nature of the PBA-diol bond. Collectively, these results demonstrate that **C9**-based glycoprotein labeling is reversible by introducing exogenous diol competitors.

We also evaluated the utility of **C9** using cell lysates derived from three cancer cell lines (HT29, HCT116, and SW480), all of which are known to express abundant glycosylated proteins (Figure 5j, Figure S12). Consistent with trends observed in monosaccharide substrate screening (Figure 3), **C9** produced substantially stronger fluorescence signals with a higher signal-to-noise ratio compared to **C1**. **C9** also showed a staining pattern comparable to that with the Pro-Q Emerald 300 staining kit in the cell lysate gels, and incubating **C9**-stained cell lysate gels in 1 M sorbitol (5 h) resulted in loss of fluorescence signals (Figure S12). Together, our probe **C9** selectively and reversibly detects glycoproteins in gels.

2. Utility of C9 for flow cytometry and cell imaging

The robust performance of **C9** for in-gel imaging of glycoproteins prompted us to explore its utility for labeling glycosylated cell surfaces. In particular, given the dynamic nature of PBA-diol binding, validating their performance under flow cytometry conditions is essential to confirm that

C9 can reliably and stably label glycosylated cell surfaces. Flow cytometry analysis (Figures 6a and 6b, Figures S13–S14) of live THP-1 cells labeled with **C*** (control), **C1**, **C8**, and **C9** revealed a similar trend as the monosaccharide study (Figure 3), with all dyes consistently exhibiting higher fluorescence intensity compared to the control probe at all concentrations. To confirm labeling stability over time, live THP-1 cells labeled with each probe were then incubated for 0–90 min in complete medium without probe (Figure S14). Here, we observed slight decreases in labeling for **C1** and **C8** (Figure S14), consistent with the reversible nature of binding. Nevertheless, cell labeling remained distinguishable across the entire time course of the wash-out experiment. Consistent with its stronger fluorescence intensity in prior experiments, **C9** showed no loss of signal with an apparent slight increase observed in later timepoints, which may be due to endocytosis of stably labeled glycoproteins and endosomal acidification.

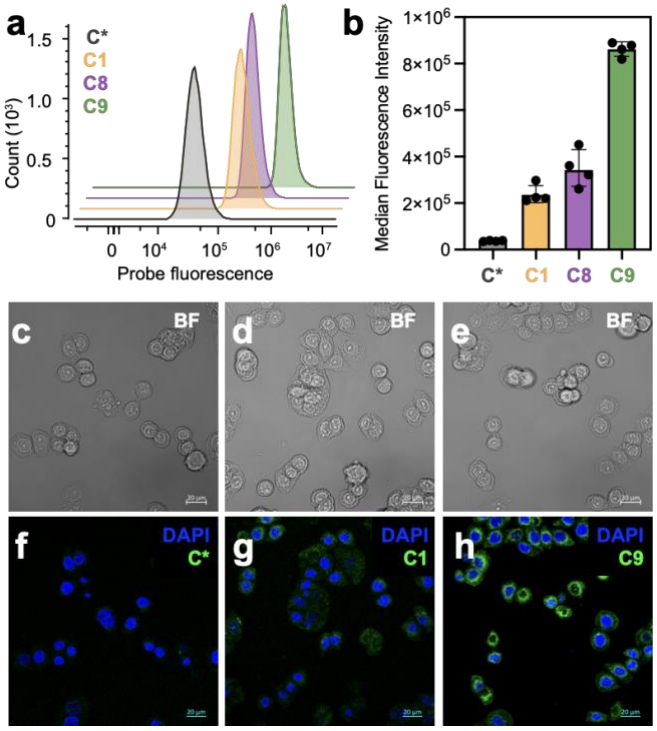


Figure 6. Utility of C9 for flow cytometry and cell imaging. (**a**) Representative stacked histograms from the flow cytometry using live THP-1 cells (5×10^5 cells/mL) labeled with C*, C1, C8, and C9 (100 nM). (**b**) Flow cytometric quantification using median fluorescence intensity and plotted using geometric mean and geometric standard deviations in four technical replicates. (**c–e**) Representative brightfield images from

the confocal microscopy imaging of fixed HT29 cells with (**c**) C*, (**d**) C1, and (**e**) C9 (100 nM). (**f**–**g**) Representative false-colored fluorescence images (blue: DAPI, and green: PBA probes) from the confocal microscopy imaging of fixed HT29 cells with (**f**) C*, (**g**) C1, and (**h**) C9 (100 nM). Scale bar = 20 μm.

Fixed-cell confocal microscopy imaging of HT29 cells (Figures 6c–h, Figure S15.1) further confirmed differential labeling by our probes. **C9** produced the most intense and distinct fluorescence (Figure 6h, Figure S15.1) compared to **C1** or **C***. When these fixed cells labeled with **C9** were washed with sorbitol (200 mM), a significant reduction in fluorescence signal was observed under microscopy (Figure S15.2). We note that, unlike in-gel or fixed-cell experiments, a sortibol wash (150 mM) did not result in a meaningful reduction in labeling of live cells (Figure S14). The robust performance of **C9** in both flow cytometry and imaging showcase its potential utility for broader cellular analyses.

CHAPTER 3: Summary and Outlook

This study demonstrates that hydrogen bonding auxiliaries in turn-on fluorogenic PBA reporters serve as a powerful strategy for tuning the sensitivity for detecting neutral mammalian sugars. Drawing inspiration from the hydrogen bonding network that controls the affinity and selectivity for natural lectin-carbohydrate interactions, we established that auxiliaries adjacent to the boron site of PBA can substantially change the turn-on response to individual mammalian sugars. Notably, a significant difference in turn-on response was observed between the anomeric forms of monosaccharides, highlighting the unique capabilities of these probes to distinguish sugar diastereomers and potentially linkages. Furthermore, the successful demonstration of these reporters across multiple biological analysis platforms, including in-gel glycoprotein imaging, flow cytometry, and cellular microscopy, establishes their broad utility as chemical tools for studying mammalian glycans. Together, these findings lay a foundation for the rational design of next-generation fluorogenic PBA reporters with programmable selectivity, advancing our ability to interrogate complex glycans in biological systems.

While the current platform does not afford absolute selectivity for a single sugar or a specific anomeric form, this limitation does not diminish its broader utility for glycobiology. Upon expanding the PBA library and establishing a robust quantitative structure-property relationship between the molecular structures and their fluorogenic turn-on responses to biological sugars, this platform can be synergistically integrated with existing technologies, such as glycan microarrays. Particularly noteworthy is the reversible nature of boronic acid-diol binding, which presents a unique and compelling advantage in the context of glycan microarray applications. The reversibility of these PBA probes will allow for repeated use of a single microarray with multiple structurally distinct PBA probes. This capability facilitates the generation of large datasets from a

single microarray, which is especially valuable for training machine learning algorithms. Collectively, we believe this work will establish a new molecular platform for next-generation glycan profiling.

REFERENCES

1. A. Varki, *Glycobiology*, 2017, **27**, 3-49.
2. N. Haines and K. D. Irvine, *Nature Reviews Molecular Cell Biology*, 2003, **4**, 786-797.
3. N. J. Agard and C. R. Bertozzi, *Accounts of Chemical Research*, 2009, **42**, 788-797.
4. L. L. Kiessling and R. A. Splain, *Annual Review of Biochemistry*, 2010, **79**, 619-653.
5. A. Varki, *Glycobiology*, 1993, **3**, 97-130.
6. S. C. Purcell and K. Godula, *Interface Focus*, 2019, **9**.
7. P. Gagneux, V. Panin, T. Hennet, M. Aebi and A. Varki, 2022.
8. M. E. Griffin and L. C. Hsieh-Wilson, *Cell*, 2022, **185**, 2657-2677.
9. D. H. Dube and C. R. Bertozzi, *Nature reviews Drug discovery*, 2005, **4**, 477-488.
10. N. Sharon and H. Lis, *Glycobiology*, 2004, **14**, 53R-62R.
11. O. Hendrickson and A. Zherdev, *Critical Reviews in Analytical Chemistry*, 2018, **48**, 279-292.
12. D. Bojar, L. Meche, G. Meng, W. Eng, D. F. Smith, R. D. Cummings and L. K. Mahal, *ACS Chemical Biology*, 2022, **17**, 2993-3012.
13. Y. Chen, L. Ding and H. Ju, *Accounts of Chemical Research*, 2018, **51**, 890-899.
14. H. Li, L. Wei, P. Fang and P. Yang, *Analytical Methods*, 2014, **6**, 2003-2014.
15. W. L. A. Brooks, C. C. Deng and B. S. Sumerlin, *ACS Omega*, 2018, **3**, 17863-17870.
16. T. D. James, K. R. A. S. Sandanayake, R. Iguchi and S. Shinkai, *Journal of the American Chemical Society*, 1995, **117**, 8982-8987.
17. T. D. James, K. R. A. S. Sandanayake and S. Shinkai, *Angewandte Chemie International Edition in English*, 1996, **35**, 1910-1922.
18. T. D. James, K. R. A. S. Sandanayake and S. Shinkai, *Journal of the Chemical Society, Chemical Communications*, 1994, DOI: 10.1039/C39940000477, 477-478.
19. K. R. A. S. Sandanayake and S. Shinkai, *Journal of the Chemical Society, Chemical Communications*, 1994, DOI: 10.1039/C39940001083, 1083-1084.
20. S. Arimori, M. L. Bell, C. S. Oh, K. A. Frimat and T. D. James, *Chemical Communications*, 2001, DOI: 10.1039/B105994G, 1836-1837.
21. W. Zhai, L. Male and J. S. Fossey, *Chemical Communications*, 2017, **53**, 2218-2221.
22. K. Wang, R. Zhang, X. Zhao, Y. Ma, L. Ren, Y. Ren, G. Chen, D. Ye, J. Wu, X. Hu, Y. Guo, R. Xi, M. Meng, Q. Yao, P. Li, Q. Chen and T. D. James, *Journal of the American Chemical Society*, 2023, **145**, 8408-8416.
23. Y. e. Wang, R. Rong, H. Chen, M. Zhu, B. Wang and X. Li, *Bioorganic & Medicinal Chemistry Letters*, 2017, **27**, 1983-1988.
24. G. T. Williams, J. L. Kedge and J. S. Fossey, *ACS Sensors*, 2021, **6**, 1508-1528.
25. X. Sun, T. D. James and E. V. Anslyn, *Journal of the American Chemical Society*, 2018, **140**, 2348-2354.
26. X. Sun, B. M. Chapin, P. Metola, B. Collins, B. Wang, T. D. James and E. V. Anslyn, *Nature Chemistry*, 2019, **11**, 768-778.
27. A. Fiore, G. Yu, J. J. Northey, R. Patel, T. A. Ravenscroft, R. Ikegami, W. Kolkman, P. Kumar, T. L. Dilan, V. M. S. Ruetten, M. B. Ahrens, H. Shroff, S. Wang, V. M. Weaver and K. Pedram, *Nature Methods*, 2025, **22**, 1070-1080.
28. K. Ko, M. Gao, S. Sarkar, H.-Y. Kwon and Y.-T. Chang, *Journal of the American Chemical Society*, 2025, **147**, 19718-19726.
29. W. I. Weis and K. Drickamer, *Annual Review of Biochemistry*, 1996, **65**, 441-473.

30. G. Springsteen and B. Wang, *Tetrahedron*, 2002, **58**, 5291-5300.
31. E. Chekol Abebe, Z. Tilahun Muche, A. Behaile T/Mariam, T. Mengie Ayele, M. Mekonnen Agidew, M. Teshome Azezew, E. Abebe Zewde, T. Asmamaw Dejenie and M. Asmamaw Mengstie, *Frontiers in Cell and Developmental Biology*, 2022, **Volume 10 - 2022**.
32. G. Gotte, M. Libonati and D. V. Laurents, *Journal of Biological Chemistry*, 2003, **278**, 46241-46251.

APPENDIX 1: Supplementary Information

S1. Materials, Instrumentations, and Methods

S1.1. Materials: Reagents and solvents were used as received from commercial sources without further purification. Anhydrous dichloromethane (DCM), 2,4-dimethylpyrrole, 2-chloroacetylchloride, boron trifluoride diethyl etherate, 2-(bromomethyl)phenylboronic acid, potassium carbonate, benzyl bromide, N-(2-aminoethyl)acetamide, methyl glycinate, 3-amino propanamide, 2-aminoethan-1-ol, D-(+)-galactose, D-(+)-xylose, methyl-β-D-glucopyranoside, methyl-α-D-galactopyranoside, methyl-α-D-mannopyranoside, methyl-α-D-fucopyranoside, methyl-α-D-xylopyranoside, D(-)-fructose, D-sorbitol, and poly-D-lysine were purchased from Sigma-Aldrich (MO, USA). 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) and methyl-α-D-glucopyranoside were purchased from Chem-Impex International (IL, USA). DCM, ethyl acetate, methanol (MeOH), acetonitrile (MeCN), triethylamine, and acetic acid were purchased from Fisher Chemical (NH, USA). Sodium bicarbonate ($NaHCO_3$) was purchased from Electron Microscopy Sciences (PA, USA). Chloroform ($CHCl_3$) was purchased from VWR chemicals (PA, USA). Methylamine, 2-methoxyethylamine, 2-methylbutylamine, and methyl-β-D-galactopyranoside were purchased from TCI America (OR, USA). $CDCl_3$-d and DMSO-d6, and D_2O were purchased from Cambridge Isotopes Laboratories (MA, USA). Neutral aluminum oxide, hexane, pentane, dimethyl sulfoxide (DMSO), D-(+)-glucose, N-acetyl-D-glucosamine, N-acetyl galactosamine, D-(+)-mannose, L-(-)-fucose, and Formaldehyde in DPBS were purchased from ThermoFisher Scientific (MA, USA). 4-(chloromethyl)benzaldehyde, ethyl 2,4-dimethyl-1H-pyrrole-3-carboxylate, and 2-(2-(2-methoxyethoxy)ethoxy)ethan-1-amine were purchased from Ambeed (IL, USA). 2,2,2-trifluoroacetic acid was purchased from Acros Organics (Belgium). 2-(methylthio)ethan-1-amine was purchased from Matrix Scientific (SC, USA). 2-fluoroethan-1-amine was purchased from ChemScene (NJ, USA). Phosphate-buffered saline (PBS) premix powder was purchased from Apex Bioresearch Products (CA, USA). Fetal bovine serum (FBS), Dulbecco's Modified Eagle Medium (DMEM), and Dulbecco's Phosphate Buffered Saline (DPBS) were purchased from Gibco (MA, USA). Sodium dodecyl sulfate (SDS), Precision Plus Protein All Blue Standards, PageRuler Prestained Protein Ladder (10-180 kDa), loading buffer, protein

gels, Coomassie stain, were purchased from BioRad (CA, USA). PNGase F, RNase B, and Fetuin were purchased from New England Biolab (MA, USA). Zombie NIR was purchased from BioLegend (CA, USA). THP-1 cells, HT29 cells, SW480 cells, and HCT116 cells were purchased from ATCC (CA, USA).

S1.2. Instrumentations: A Horiba type LAQUAtwin-pH-22B compact pH meter was used for pH measurements. Absorbance and fluorescence spectra were recorded on Cary-60 absorption spectrometer and Cary Eclipse fluorimeter from the Laser Lab at UCI. Fluorescence measurements were recorded on a BioTek™ Synergy™ H1 Hybrid Multi-Mode microplate reader. Organic solvents were concentrated under reduced pressure using a rotary evaporator. All silica column chromatography were performed using Buchi Pure C-815 Flash equipped with FlashPure EcoFlex columns. Gels were performed in Bio-Rad mini-PROTEAN® tetra cell with mini- PROTEAN® TGX™ precast gels (4-20% gradient). Bio-Rad ChemiDoc™ MP imaging system were used to visualized and capture gel images. Malvern Microcal PEAQ-ITC instrument was used to obtain isothermal titration calorimetry (ITC) measurements. Eppendorf cellxpert incubator were used to maintain cells for this study. Invitrogen countess II FL were used to count cells for this study. All cell work were done in Nuair class II biosafety cabinet. High-resolution mass spectrometry spectra were obtained at the University of California Irvine Mass Spectrometry Facility. Cells were analyzed using a three-laser Agilent Novocyte flow cytometer. Confocal imaging was performed on a Zeiss LSM 980 confocal laser scanning microscope using a 40× immersion objective placed underneath the glass coverslip. ^{1}H and ^{13}C{^{1}H} NMR spectra were obtained using Bruker instruments AV400, AE500, and AVANCE600 equipped with a cryoprobe. ^{1}H NMR spectra were acquired at 400 MHz, 500 MHz, or 600 MHz, ^{13}C{^{1}H} NMR spectra were acquired at 126 MHz, and all spectra were referenced to residual solvent signals (7.26 ppm for ^{1}H and 77.16 ± 0.06 ppm for ^{3}C{^{1}H} $CDCl_3$ peaks, and 2.09 ppm for ^{1}H acetone residual peaks). Solvent peaks, such as methanol as a quartet at 4.01 ppm and a triplet at 3.16 ppm, water as a broad singlet at 3.33 ppm, and DMSO as a septet at 2.50 ppm, were annotated. Peaks labeled with an asterisk (*) on ^{13}C{^{1}H} NMR spectra in DMSO-d_6 are DMSO at 39.52 ppm.

S1.3. Methods

Quantum Yield determination: Fluorescence quantum yield (Φ_{fl}) of the probes was measured using Fluorescein isothiocyanate isomer I (FITC, MilliporeSigma, F4274, Φ_{fl} = 0.92 in PBS) as standard at an excitation wavelength of 496 nm and recorded emission spectra from 500 to 600 nm. The fluorescence quantum yield (Φ_{fl}) was calculated according to the following equation:

$$\Phi_{probe} = \Phi_{std} \frac{F_{probe}}{F_{std}} \frac{\varepsilon_{std}[std]}{\varepsilon_{probe}[probe]}$$

Φ_{probe} and Φ_{std} are the quantum yield of the standard and probe. F_{probe} and F_{std} are the calculated areas under the fluorescence emission profiles of the probe and standard. ε_{probe} and ε_{std} are the molar absorption coefficients of the probe and standard. [probe] and [std] are the concentrations of the probe and standard. Solvent reflective index (η) = 1 in all cases due to the same solvent in probe and standard.

pKa determination (A, B, and C): A 10% solution of MeCN in 0.1 M sodium phosphate was prepared before adjusting the solution pH to 2 using 5 M HCl. Probe stock solutions were prepared in DMSO, of which 60 μL was added to achieve a final concentration of 10 μM. A 1.0 M sodium hydroxide (NaOH) solution was added dropwise to gradually raise the pH of the probe solution. The absorbance of the solution at λ_{max} was taken after every few drops using a Nanodrop spectrophotometer. The absorbance measurements were then plotted against solution pH to obtain a titration curve, from which the probe's pKa can be determined.

Intrinsic photophysical property determination: A 0.1 M sorbitol solution was prepared in a 1:1 (v/v) DMSO/PBS mixture. Probe stock solutions were prepared in DMSO and added to 3.0 mL of the sorbitol solution to afford a final probe concentration of 5 μM. The resulting solutions were incubated at 37 °C for 20 min, after which absorption and fluorescence spectra were recorded using a Cary 60 UV-Vis spectrophotometer and a Cary Eclipse fluorescence spectrophotometer.

Plate reader fluorescence determination of monosaccaride: In a 96-well plate, solutions containing individual diol derivatives were prepared in PBS (pH 7.4). Probe stock solutions were prepared in DMSO and added to give a final probe concentration of 20 μM in the presence of 100 mM diol (total volume, 100 μL; final solvent composition, 99% PBS by volume). The samples were incubated at 37 °C for 20 min, and fluorescence intensity at the emission maximum was measured using a plate reader. The fluorescence responses obtained for the various diol species were then compared with that of the probe in PBS alone.

ITC method to determine binding constant: The binding experiments were performed in an ITC instrument (Malvern Microcal PEAQ-ITC) at 25 °C. 70 mM monosaccharides (glucose, fructose and sorbitol) and B were suspended in buffer and the reference ampoule was filled with tris buffer (pH 8.1). The concentration of B was kept constant at 0.5 mM for all experiments and inserted into the sample cell. The syringe of the ITC device was purged three times with the monosaccharide solution and then filled with the same solution. Then the sample cell was equipped with the syringe and the injection started after the thermal equilibrium was reached. Each experiment involved a series of 18 injections of 2 μL each after the first injection of 0.4 μL. The heat of a ligand dilution was measured in the absence of receptor molecules in separate experiments and subtracted. The resulting titration isotherm was fitted with analysis software of Micro Cal PEAQ-ITC system using the fitting model "One set of sites".

Detection of polysaccharide species in physiological condition: Fresh polysaccharide solutions of the indicated concentrations (in % w/v) were prepared in PBS (pH 7.4) immediately prior to use. In a 96-well plate, probe stock solutions prepared in DMSO were added to the polysaccharide solutions to afford a final probe concentration of 20 μM (total volume, 100 μL; final solvent composition, 99% PBS by volume). The samples were incubated at 37 °C for 20 min, and fluorescence intensity at the emission maximum was measured using a plate reader. The fluorescence responses observed for the various polysaccharides were then compared with that of the probe in PBS alone.

Detection of serum in biological medium: Fresh 5% and 10% (v/v) fetal bovine serum (FBS) solutions in Dulbecco's modified Eagle medium (DMEM) were prepared immediately prior to use. In a 96-well plate, probe stock solutions prepared in DMSO were added to afford a final probe concentration of 20 μM in

DMEM alone, 5% FBS in DMEM, or 10% FBS in DMEM. Fluorescence intensity at the emission maximum was measured using a plate reader over time in triplicate. The fluorescence responses obtained under each condition were compared with those of the probe in DMEM alone. As additional controls, the emission at λ_{max} of DMEM alone, 5% FBS in DMEM, and 10% FBS in DMEM was also recorded in the absence of probe.

In-gel glycoprotein staining: SDS–PAGE was performed in all cases. Protein samples were first denatured in 3% SDS and heated at 95 °C for 10 min. Samples were then loaded to afford 3 μg of protein per lane, and electrophoresis was carried out at 150 V. Upon completion of electrophoresis, gels were fixed in a solution of 50% methanol, 10% acetic acid, and 40% distilled water for 1 h on a rocker. The fixed gels were then washed three times with distilled water (5 min per wash), followed by three washes with PBS (pH 7.4, 5 min per wash). After washing, the gel pH was confirmed to be above 7 to ensure conditions suitable for phenylboronic acid labeling. Probe solutions of varying concentrations were freshly prepared in PBS (30 mL total volume) and incubated with the gels for 1 h in the dark. Following staining, the gels were destained twice with 20% methanol in PBS (10 min per wash) to reduce background fluorescence, followed by three additional washes with PBS (pH 7.4, 10 min per wash). The gels were then imaged directly using a Bio-Rad ChemiDoc™ imaging system with the Alexa Fluor 488 channel. Coomassie staining was subsequently performed to visualize total protein.

PNGaseF treatment on RNase B: PNGaseF and RNase B standards were acquired from New England Biolab directly and used without further purification. Denature and PNGaseF cleaving conditions were followed according to the protocol provided by New England Biolab. Samples were then loaded to afford 3 μg of protein per lane, and electrophoresis was carried out at 150 V. Upon completion of electrophoresis, gels were fixed in a solution of 50% methanol, 10% acetic acid, and 40% distilled water for 1 h on a rocker. The fixed gels were then washed three times with distilled water (5 min per wash), followed by three washes with PBS (pH 7.4, 5 min per wash). After washing, the gel pH was confirmed to be above 7 to ensure conditions suitable for phenylboronic acid labeling. 10 μM probe solutions were freshly prepared in PBS (30 mL total volume) and incubated with the gels for 1 h in the dark. Following staining, the gels were

destained twice with 20% methanol in PBS (10 min per wash) to reduce background fluorescence, followed by three additional washes with PBS (pH 7.4, 10 min per wash). The gels were then imaged directly on-gel using a Bio-Rad ChemiDoc™ imaging system with the Alexa Fluor 488 channel. Coomassie staining was subsequently performed to visualize total protein.

Cell lysate preparation and in-gel lysate staining/destaining: For adherent cells, medium was aspirated from 10-cm plates of HCT-116, HT-29, and SW480 at 80% confluency, and cells were rinsed twice with 1x phosphate buffered saline (PBS) at room temperature. To each plate, 500 µL of cell lysis buffer (4% SDS, 1x PBS, 1x complete protease inhibitor complex without EDTA (PIC-)) was directly added. Cells were then scraped from the plate, and the lysate was transferred to a 1.5-mL microcentrifuge tube. Lysates were sonicated (30% intensity, 2 s on / 2 s off, 10-20 s total on) until the fluid was no longer viscous. SDS–PAGE was performed. Samples were first denatured at 95 °C for 10 min. Samples were then loaded to afford 15 µg of total protein per lane, and electrophoresis was carried out at 120V for the first 10 minutes and 150V for the rest of the run. Upon completion of electrophoresis, gels were fixed in a solution of 50% methanol, 10% acetic acid, and 40% distilled water for 1 h on a rocker. The fixed gels were then washed three times with distilled water (5 min per wash), followed by three washes with PBS (pH 7.4, 5 min per wash). After washing, the gel pH was confirmed to be above 7 to ensure conditions suitable for phenylboronic acid labeling. 10 µM of probe solutions were freshly prepared in PBS (30 mL total volume) and incubated with the gels for 1 h in the dark. Following staining, the gels were destained twice with 20% methanol in PBS (10 min per wash) to reduce background fluorescence, followed by three additional washes with PBS (pH 7.4, 10 min per wash). The gels were then imaged directly using a Bio-Rad ChemiDoc™ imaging system with the Alexa Fluor 488 channel. Coomassie staining was subsequently performed to visualize total protein.

Sorbitol wash: Following fluorescence imaging of the protein and cell lysate samples using a Bio-Rad ChemiDoc™ imaging system with the Alexa Fluor 488 channel, the gel was incubated in 50 mL of PBS containing 1 M sorbitol for 5 h on a rocking platform. After incubation, the gel was reimaged using identical acquisition settings, and the fluorescence signals obtained before and after sorbitol treatment were

compared side by side. The gel was subsequently stained with Coomassie Brilliant Blue to visualize total protein.

Flow cytometry visualization of THP-1 cells:

Dose-dependent labeling: Aliquots of THP-1 cells (180 µL at $5x10^5$ cells mL^{-1} in complete RPMI) were added to individual tubes of sterile 8-strip tubes (Fisher Scientific, 072-003-22). To each tube, 20 µL of complete RPMI containing the indicated concentration of each probe at 10x concentration was then added (final concentration range of probes: 0-1000 nM; final concentration of DMSO in each well: 0.1%), and cell suspensions were gently mixed. Cells were incubated for 1 h at 37 °C with 5% CO_2. After incubation, 800 µL of Dulbecco's Phosphate Buffered Saline (DPBS) at 4 °C was added to each tube for a final volume of 1 mL. Cells were then centrifuged at 500 x *g* for 3 minutes at 4 °C. Supernatant was removed, leaving 50 µL, and 950 µL of DPBS was added to each tube. Cells were again centrifuged as described above, and rinse buffer was removed, leaving 40 µL. Zombie NIR (BioLegend, 423105) was diluted 1:200 in DPBS at room temperature, and 10 µL was added to each tube (final concentration of Zombie NIR: 1:4,000). Cells were incubated for 15 min at room temperature protected from light. Live/dead staining reagent was then removed by rinsing with DPBS as described above, leaving 50 µL of cell suspension. Cells were diluted with 150 µL of DPBS and transferred to a 96-well V bottom plate (VWR, 07-200-695). Cells were then analyzed using a three-laser Agilent Novocyte flow cytometer with excitation and emission measured using the B530-A and R780-A channels.

Wash-out: Aliquots of THP-1 cells (360 µL at $5x10^5$ cells mL^{-1} in complete RPMI) were added to individual tubes of sterile Corning 8-strip tubes. To each tube, 40 µL of complete RPMI containing the indicated concentration of each probe at 10x concentration was then added (final concentration range of probes: 0-1000 nM; final concentration of DMSO in each well: 0.1%), and cell suspensions were gently mixed. Cells were incubated for 1 h at 37 °C with 5% CO_2. After incubation, 600 µL of complete RPMI pre-warmed to 37 °C was added to each tube for a final volume of 1 mL. Cells were centrifuged at 500 x *g* for 3 minutes at room temperature. Supernatant was removed leaving 50 µL, and 350 µL of pre-warmed complete RPMI was added to each tube. Immediately, a 70 µL aliquot was removed and immediately analyzed by flow

cytometry as described above as the 0 min time point. After each 30 min interval, 70 µL aliquots were removed and immediately analyzed by flow cytometry as described above.

Sorbitol wash: Aliquots of THP-1 cells (360 µL at $5x10^5$ cells mL-1 in complete RPMI) were added to individual tubes of sterile Corning 8-strip tubes. To each tube, 40 µL of DPBS containing 1000 nM of probe was added, and cell suspensions were gently mixed. Cells were incubated for 1 h at 37 °C with 5% CO_2. After incubation, 600 µL of room temperature DPBS with 0.2% FBS (referred to as flow buffer for the remainder of this experiment) was added to each tube for a final volume of 1 mL. Cells were then centrifuged at 500 x g for 3 minutes. Supernatant was removed, leaving 50 µL, and 950 µL of flow buffer was added to each tube. Cells were again centrifuged as described above, supernatant was removed, leaving 50 µL. Cells were then washed as described above using flow buffer with or without 150 mM D-sorbitol (Thermo Scientific Chemicals, 132730010). This was performed four times with 5-minute incubations between each wash. After the final sorbitol wash, all cells were rinsed twice as described above with DPBS only. After the second rinse, the supernatant was removed, leaving 40 µL. Zombie NIR was diluted 1:200 in DPBS at room temperature, and 10 µL was added to each tube. (final concentration of Zombie NIR: 1:4,000). Cells were incubated for 15 min at room temperature protected from light. Live/dead staining reagent was then removed by rinsing twice with flow buffer as described above, leaving 50 µL of cell suspension. Cells were diluted with 150 µL of flow buffer and transferred to a 96-well V bottom plate. Cells were then analyzed using a three-laser Agilent Novocyte flow cytometer with excitation and emission measured using the B530-A and R780-A channels.

Fixed HT29 cell preparation and confocal imaging: For cell imaging, cells were resuspended in the appropriate cell culture medium at 500,000 cells mL^{-1}, and 30 µL of cell suspension was added to each channel of a polymer ibiTreat µ-Slide VI channel slide (ibidi, 80606). Cells were incubated for 2 h at 37 °C with 5% CO2 to allow them to adhere to the surface, after which 60 µL of pre-warmed cell culture medium was added to either end of the channel. Cells were then incubated for 36 h at 37 °C with 5% CO_2. Cell culture medium was then removed by gentle aspiration, and cells were rinsed once with DPBS. Cells were then fixed using 4% formaldehyde in DPBS (freshly diluted from 16% formaldehyde; Thermo Fisher

Scientific, 28908) for 20 min at room temperature. Fixed cells were then rinsed three times with DPBS and then stored at 4 °C until ready for probe incubation. 100 nM probe solution in DPBS were incubated with fixed cells at room temperature for 1 hour and washed three times with 100 μL of DPBS. The cells were subsequently incubated with 300 nM of DAPI solution in DPBS and washed three times with 100 μL of DPBS. Confocal imaging was performed. Confocal laser scanning microscopy (CLSM) analysis: CLSM imaging was performed on a Zeiss LSM 980 confocal laser scanning microscope using a 40× immersion objective placed underneath the glass coverslip. All images were acquired over an effective imaging area of 42 μm × 42 μm at a resolution of 4096 × 4096 pixels at 16-bits with a frame time of 5.03 s. Non-fluorescent images were collected in brightfield contrast mode using a PMT detector at a detection range of 300–900 nm. Fluorescence from DAPI and BODIPY was recorded in separate channels using excitation wavelengths of 405 nm and 488 nm, with emission collected over 411–490 nm and 508–658 nm, respectively.

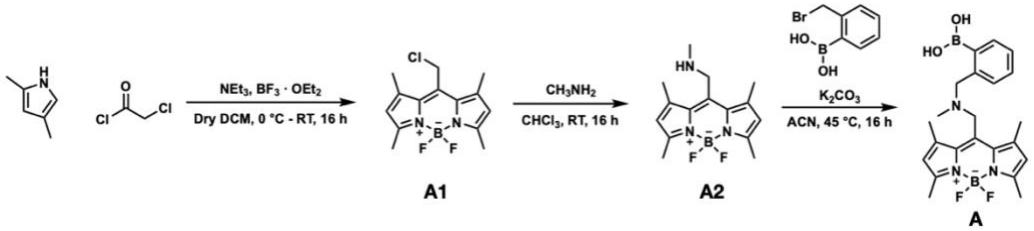
Sorbitol wash: 100 nM probe solution in DPBS were incubated with fixed cells at room temperature for 1 hour and washed (3x 5 minutes) with either 100 μL of DPBS or 100 μL of 200 mM sorbitol in PBS. The cells were subsequently incubated with 300 nM of DAPI solution in DPBS and washed three times with 100 μL of DPBS. Confocal imaging was performed with the same settings.

Data visualization, analysis, and statistics: Unless otherwise noted, data were compiled, graphed, and analyzed using Prism (GraphPad, v.10.4.2 or v.11.0.0). For all bar graphs, individual data points and error bars corresponding to standard deviation are graphed. Normalized fluorescence values ($\Delta F/F_0$) were calculated by dividing the difference in fluorescence between experimental and control values by the background fluorescence. Quantum yield values were calculated as described above. ITC binding data were fitted into MicroCal PEAQ-ITC software (v. 1.52). Flow cytometric data were analyzed using FlowJo (Waters Bioscience, v.10.10.0 or v.11.1) to calculate median fluorescent intensities for gated features. Statistical analyses of flow cytometric data were performed using one-way lognormal ANOVA with Dunnett's post-hoc tests. The only exception is for the sorbitol wash experiment, in which case a lognormal Welch's t-test was performed. Where appropriate, samples were compared using a two-way one-sided

ANOVA with Dunnett's post-hoc test for individual comparisons. *P* values <0.05 were considered statistically significant, and stars to denote statistical significance correspond to the following values: ns = *P*>0.05, **P*<0.05, ***P*<0.01, ****P*<0.001, and *****P*<0.0001. Images from in-gel fluorescence experiments were prepared using Image Lab Touch Software (v3.0.1) and quantification analysis were performed using ImageJ (v. 1.54J). Images from confocal microscopy experiments were prepared using Zen core v. 3.13 from Zeiss. NMR spectra were graphed and analyzed using MNova (Mestrelab, v12.0.3 - 21384).

S2. Synthesis and Characterization of Chemical Compounds

S2.1. Synthesis and characterization of compound A



S2.1.1. Synthesis and characterization of A1

To an oven dried round-bottom flask, 2,4-dimethylpyrrole (1000 µL, 9.7 mmol, 2.1 equiv.) and 2-chloroacetyl chloride (367 µL, 4.6 mmol, 1 equiv.) was dissolved in 30 mL of anhydrous DCM. The reaction mixture was allowed to stir for 16 hours and used without further purification. The reaction mixture was then cooled to 0 ºC, triethylamine (1923 µL, 13.8 mmol, 3 equiv.) was added to the reaction mixture dropwise at 0 ºC and allow to stir for 15 min. Boron trifluoride diethyl etherate (2838 µL, 23 mmol, 5 equiv.) was then added to the reaction mixture dropwise at 0 ºC and stirred at room temperature for 16 hours. Upon completion, the reaction mixture was quenched with 1M HCl and extracted with DCM. The organic extracts were combined, and crude residue was further purified by column chromatography using a mixture of DCM and Hexane (70: 30). to get intermediate product 10-(chloromethyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2] diazaborinine **A1** as a red powder (518.37 mg). This compound was isolated with 38% yield.

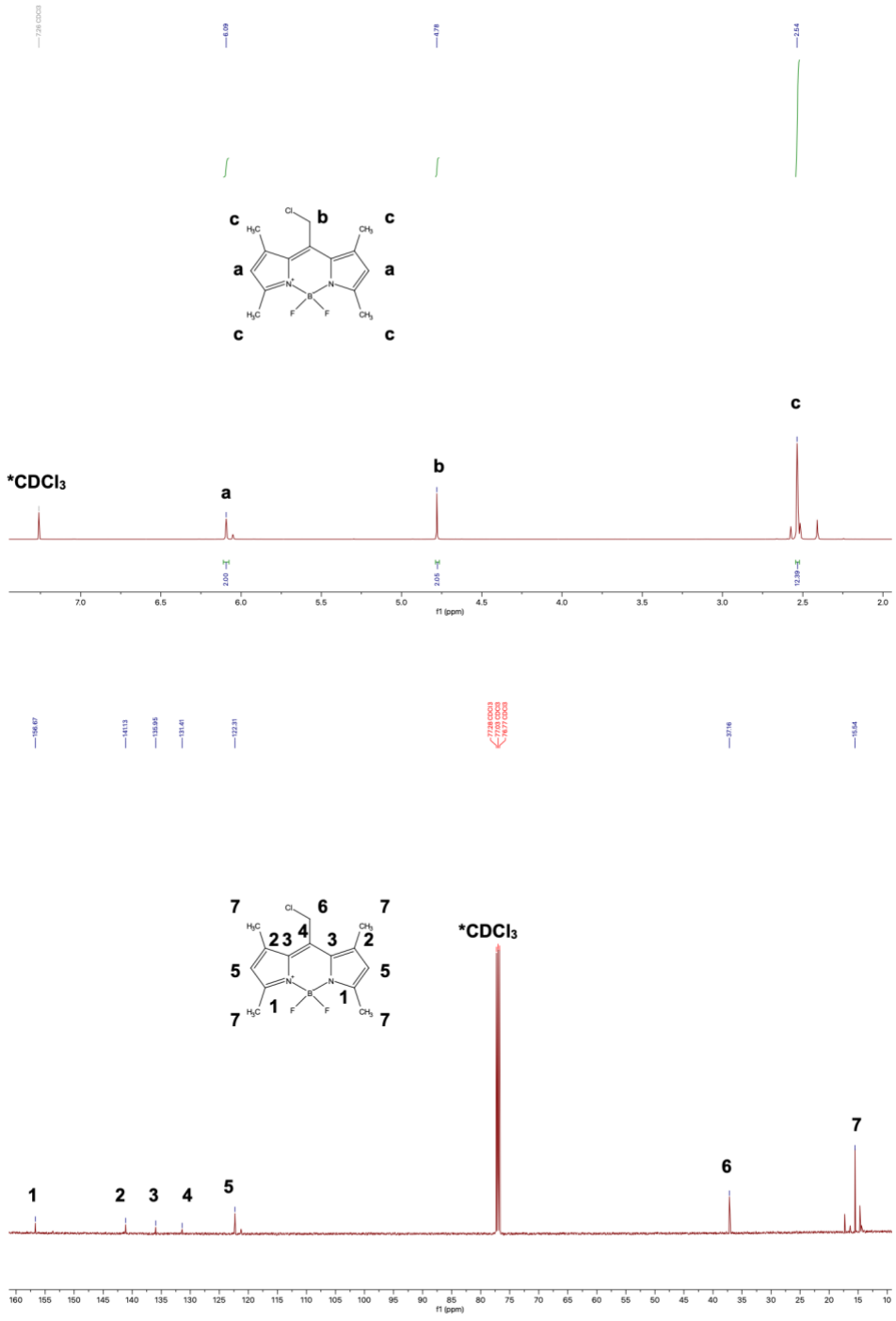


Figure S2.1.1. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 6.09 (s, 2H), 4.78 (s, 2H), 2.54 (s, 12H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 156.67, 141.13, 135.95, 131.41, 122.31, 37.16, 15.54.

S2.1.2. Synthesis and characterization of A2

10-(chloromethyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-$^{4}\lambda$4,5λ^{4}-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2] diazaborinine **A1** (518 mg, 1.7 mmol, 1 equiv.) was dissolved in 20 mL of chloroform and methylamine (40%, w/v, 10 mL) was added to the reaction mixture and stirred for 16 hours at room temperature. After the reaction was complete, volatiles were removed under reduced pressure, and crude residue was further purified by column chromatography using a mixture of ethyl acetate and methanol (9: 1). to get 1-(5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^{4},5λ^{4}-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)-*N*-methylmethanamine **A2** as a red powder (137.41 mg). This compound was isolated with 27% yield.

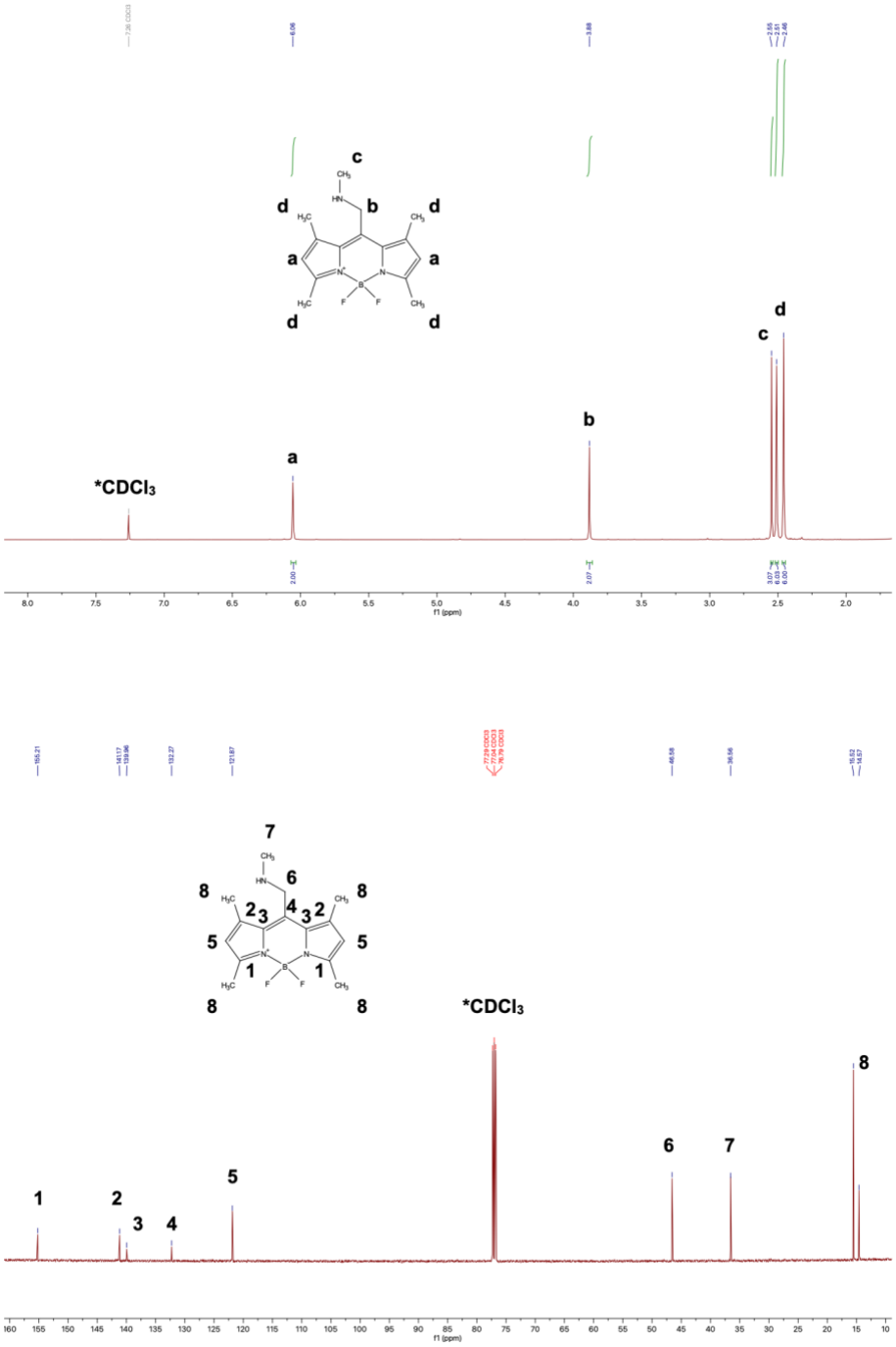


Figure S2.1.2. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 6.06 (s, 2H), 3.88 (s, 2H), 2.55 (s, 3H), 2.51 (s, 6H), 2.46 (s, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 155.21, 141.17, 139.96, 132.27, 121.87, 46.58, 36.56, 15.52, 14.57.

S2.1.3. Synthesis and characterization of A

1-(5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)-*N*-methylmethanamine **A2** (60 mg, 0.2 mmol, 1 equiv.), 2-(bromomethyl)phenylboronic acid (66 mg, 0.3 mmol, 1.5 equiv.), and potassium carbonate (55 mg, 0.4 mmol, 2 equiv.) were dissolved in 5 mL of acetonitrile. The reaction mixture was left to stir at 60 ºC for 16 hours. volatiles were removed under reduced pressure and crude residue was further purified by neutral alumina column chromatography using a mixture of ethyl acetate and methanol (95: 5) to get (2-((((5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)methyl)(methyl)amino)methyl)phenyl)boronic acid **A** as a red powder (23 mg). This compound was isolated with 11.5% yield.

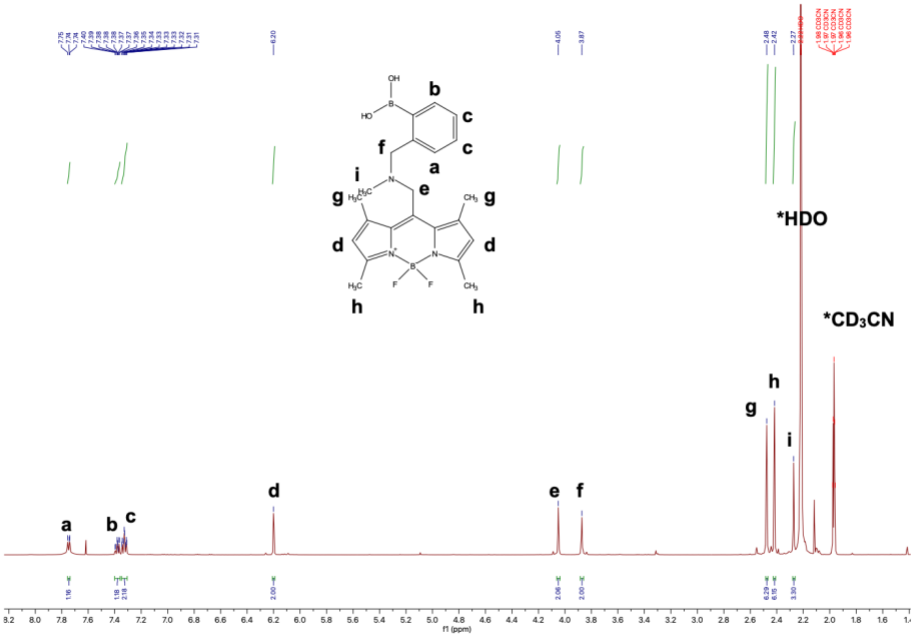
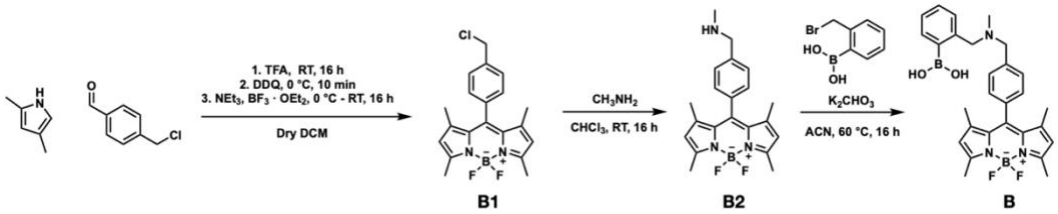


Figure S2.1.3. ^{1}H NMR (500 MHz, Acetonitrile-d_3) δ 7.76 – 7.74 (m, 1H), 7.40 – 7.36 (m, 1H), 7.33 (td, *J* = 8.4, 1.6 Hz, 2H), 6.20 (s, 2H), 4.05 (s, 2H), 3.87 (s, 2H), 2.48 (s, 6H), 2.42 (s, 6H), 2.27 (s, 3H).

S2.2. Synthesis of compound B



S2.2.1. Synthesis and characterization of B1

To an oven dried round-bottom flask, 2,4-dimethylpyrrole (387 mg, 4.1 mmol, 2.1 equiv.) and 4-(chloromethyl)benzaldehyde (300 mg, 2.0 mmol, 1 equiv.) was dissolved in 20 mL of anhydrous DCM. 2,2,2-Trifluoroacetic acid (15 µL, 0.1 mmol, 0.1 equiv.) was added to the reaction mixture. The reaction mixture was allowed to stir for 16 hours. Without further purification, the reaction mixture was then cooled to 0 ºC. DDQ (357 mg, 1.6 mmol, 1.1 equiv.) was added to the reaction mixture and stirred at 0 ºC for 30 minutes. Then, triethylamine (990 µL, 7.1 mmol, 5 equiv.) was added to the reaction mixture dropwise at 0 ºC and allow to stir for 1 hour. Boron trifluoride diethyl etherate (2020 mg, 14.2 mmol, 10 equiv.) was then added to the reaction mixture dropwise at 0 ºC and stirred at room temperature for 16 hours. Upon completion, the reaction mixture was quenched with dilute $NaHCO_3$ and extracted with DCM. The organic extracts were combined, and crude residue was further purified by column chromatography using a mixture of hexane and DCM (4: 6) to get intermediate product 10-(4-(chloromethyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine **B1** as an orange powder (63.5 mg). This compound was isolated with 12% yield.

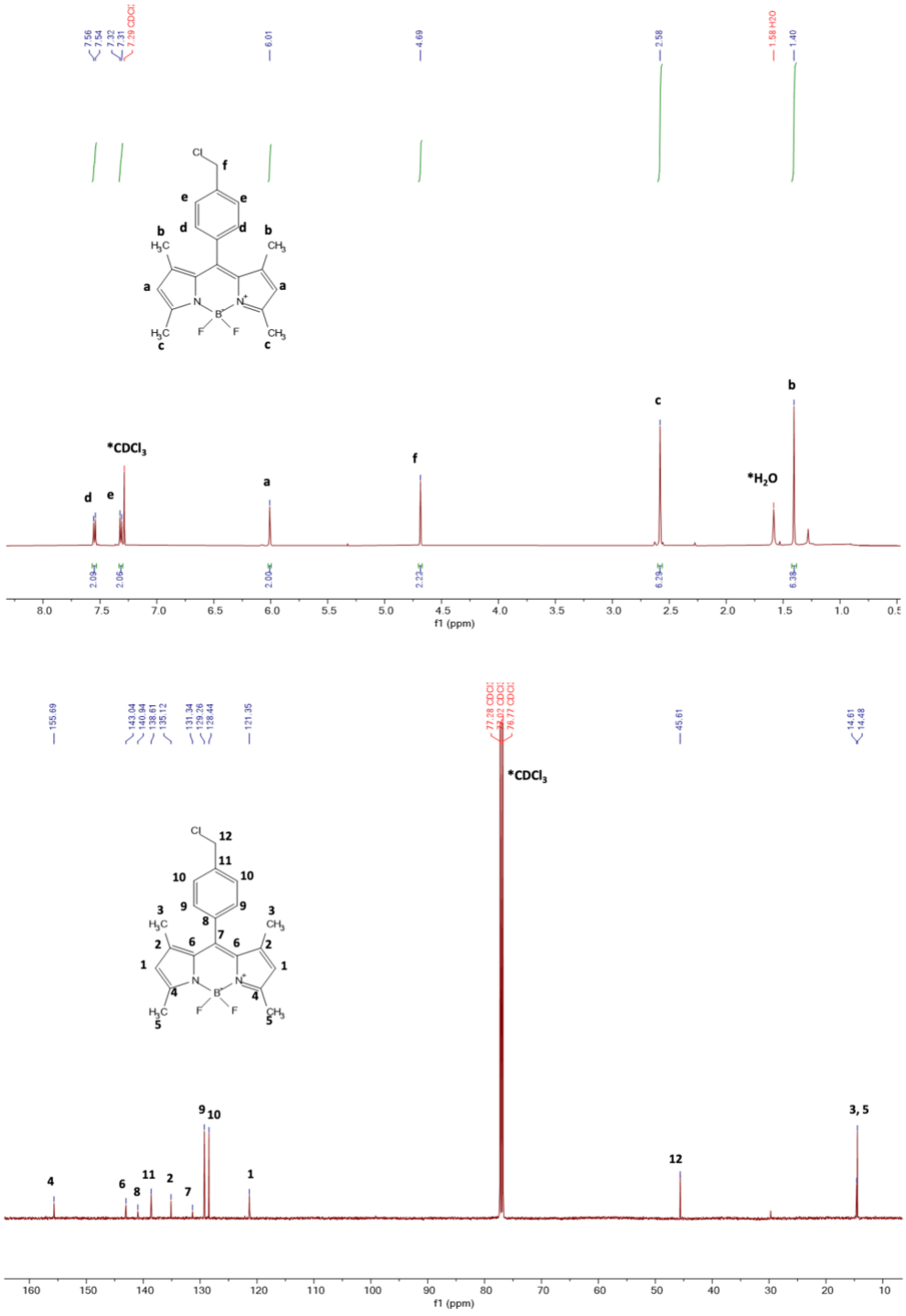


Figure S2.2.1. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55 (d, J = 8.0 Hz, 2H), 7.32 (d, J = 8.1 Hz, 2H), 6.01 (s, 2H), 4.69 (s, 2H), 2.58 (s, 6H), 1.40 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.69, 143.04, 140.94, 138.61, 135.12, 131.34, 129.26, 128.44, 121.35, 45.61, 14.61, 14.48.

S2.2.2. Synthesis and characterization of B2

10-(4-(chloromethyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine **B1** (63.5 mg, 0.2 mmol, 1 equiv.) was dissolved in 20 mL of chloroform. Methylamine (40%, w/v, 5 mL) was added to the reaction mixture, which was then stirred for 16 hours at room temperature. After the reaction was complete, volatiles were removed under reduced pressure and crude residue was further purified by column chromatography using a mixture of ethyl acetate and methanol (8:2) to get 1-(4-(5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)phenyl)-*N*-methylmethanamine **B2** as an orange powder (36 mg). This compound was isolated with 57.7% yield.

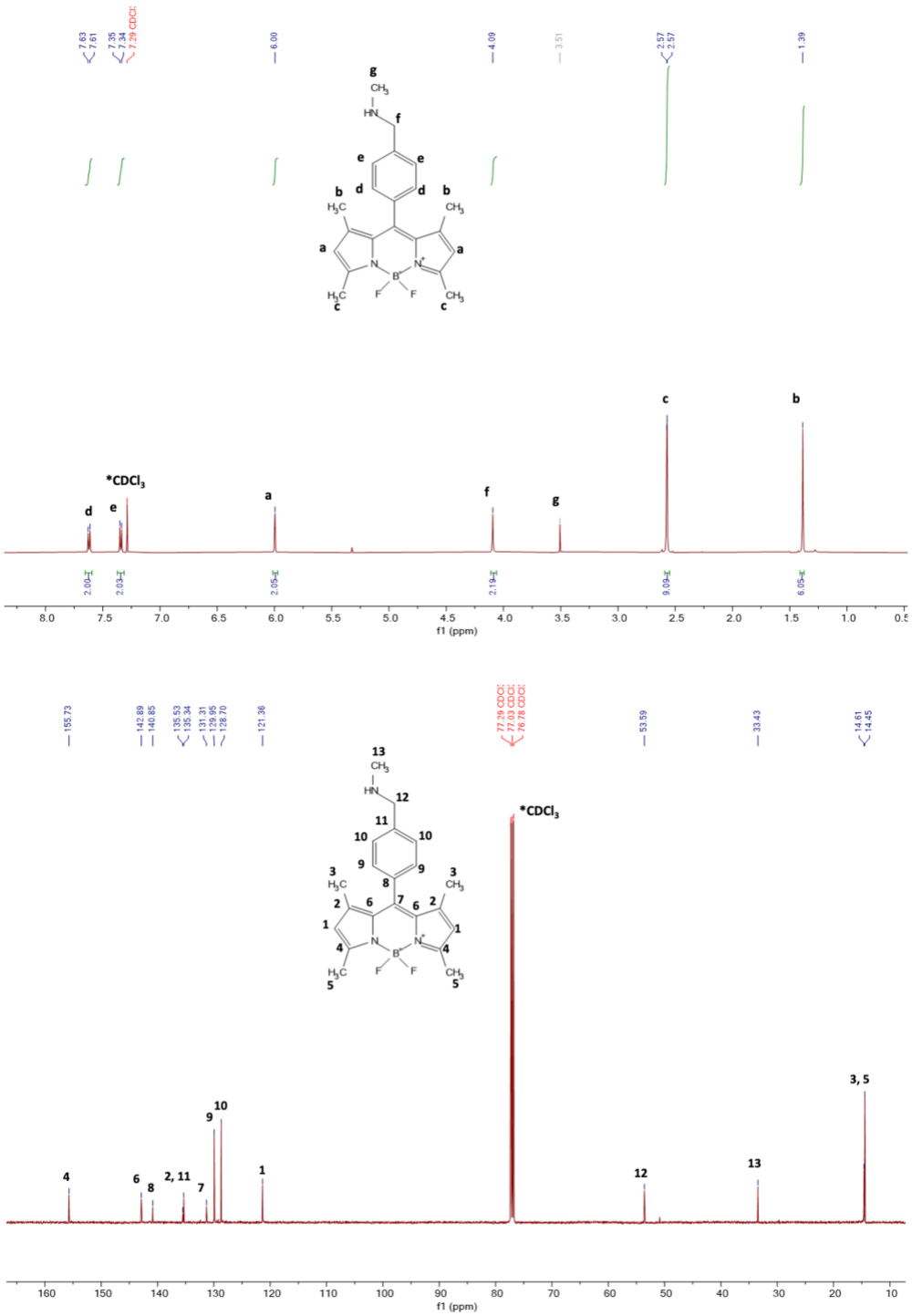


Figure S2.2.2. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.62 (d, J = 8.0 Hz, 2H), 7.34 (d, J = 8.1 Hz, 2H), 6.00 (s, 2H), 4.09 (s, 2H), 2.57 (d, J = 1.9 Hz, 9H), 1.39 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.73, 142.89, 140.85, 135.53, 135.34, 131.31, 129.95, 128.70, 121.36, 53.59, 33.43, 14.61, 14.45.

S2.2.3. Synthesis and characterization of B

1-(4-(5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)phenyl)-*N*-methylmethanamine **B2** (30 mg, 0.1 mmol, 1 equiv.), 2-(bromomethyl)phenylboronic acid (26 mg, 0.1 mmol, 1.5 equiv.), and potassium carbonate (22 mg, 0.2 mmol, 2 equiv.) were dissolved in 5 mL of acetonitrile. The reaction mixture was left to stir at 60 ºC for 16 hours. Volatiles were removed under reduced pressure and crude residue was further purified by neutral alumina column chromatography using a mixture of ethyl acetate and methanol (95: 5) to get (2-(((4-(5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(methyl)amino)methyl)phenyl)boronic acid **B** as an orange powder (18 mg). This compound was isolated with 44.9% yield.

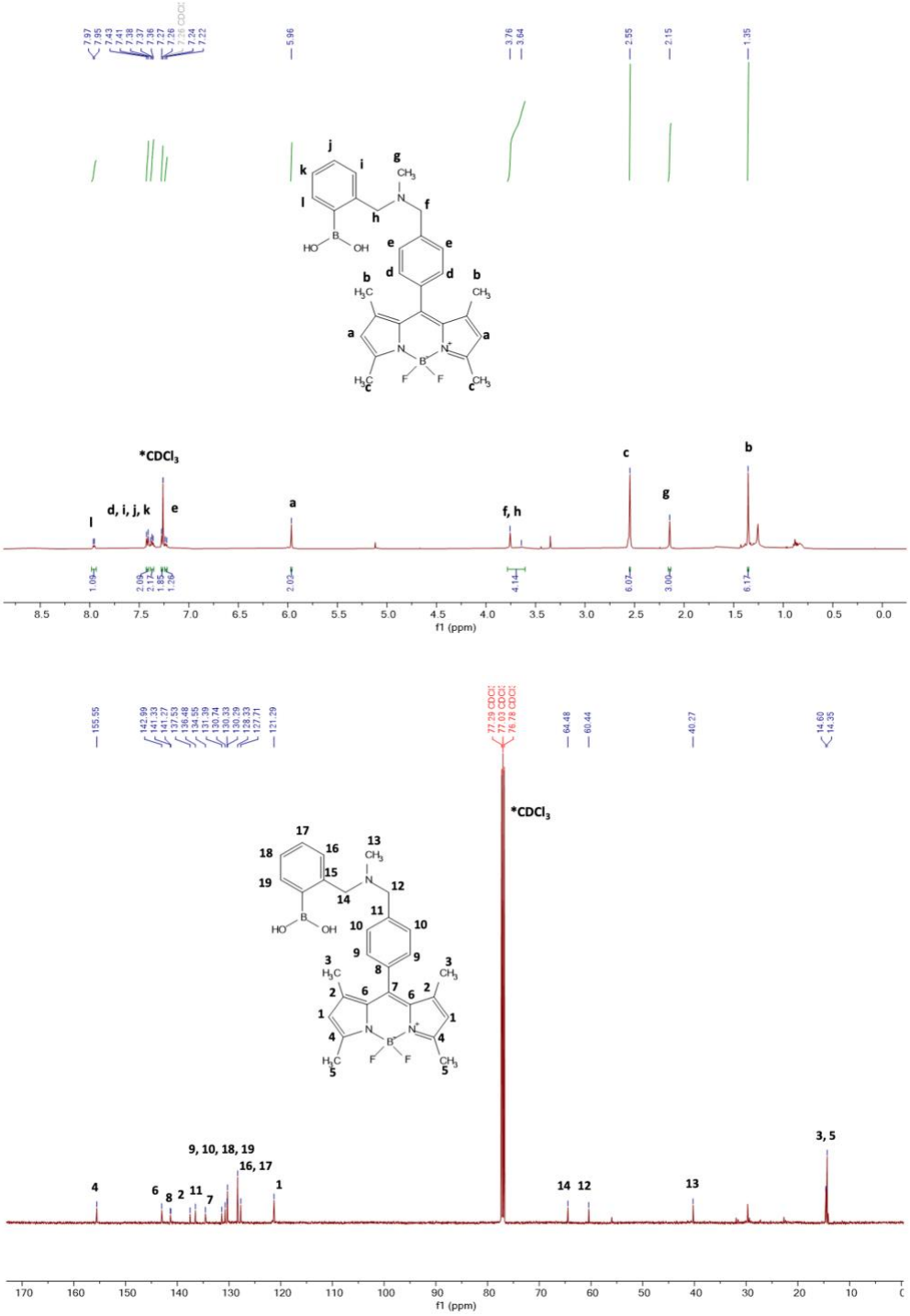


Figure S2.2.3. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.96 (d, J = 6.5 Hz, 1H), 7.42 (d, J = 8.0 Hz, 2H), 7.39 – 7.35 (m, 2H), 7.27 (s, 2H), 7.23 (d, J = 7.7 Hz, 1H), 5.96 (s, 2H), 3.76 (s, 4H), 2.55 (s, 6H), 2.15 (s, 3H), 1.35 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.55, 142.99, 141.33, 141.27, 137.53, 136.48, 134.55, 131.39, 130.74, 130.33, 130.29, 128.33, 127.71, 121.29, 64.48, 60.44, 40.27, 14.60, 14.35.

S2.3. Synthesis of compound C

S2.3.1. Synthesis of 1C

To an oven-dried round-bottom flask, ethyl 2,4-dimethyl-1*H*-pyrrole-3-carboxylate (2272 mg, 13.6 mmol, 2.1 equiv.) and 4-(chloromethyl)benzaldehyde (1000 mg, 6.5 mmol, 1 equiv.) was dissolved in 50 mL of anhydrous DCM. 2,2,2-trifluoroacetic acid (50 µL, 0.6 mmol, 0.1 equiv.) was added to the reaction mixture. The reaction mixture was allowed to stir for 16 hours. Upon completion, the reaction mixture was quenched with water and extracted with DCM. Volatiles were removed under reduced pressure and crude residue was further purified by column chromatography using a mixture of hexane and ethyl acetate (80: 20) to get intermediate product diethyl 5,5'-((4-(chloromethyl)phenyl)methylene)bis(2,4-dimethyl-1*H*-pyrrole-3-carboxylate) **1C** as an orange solid (2971.15 mg). This compound was isolated with 97.5% yield.

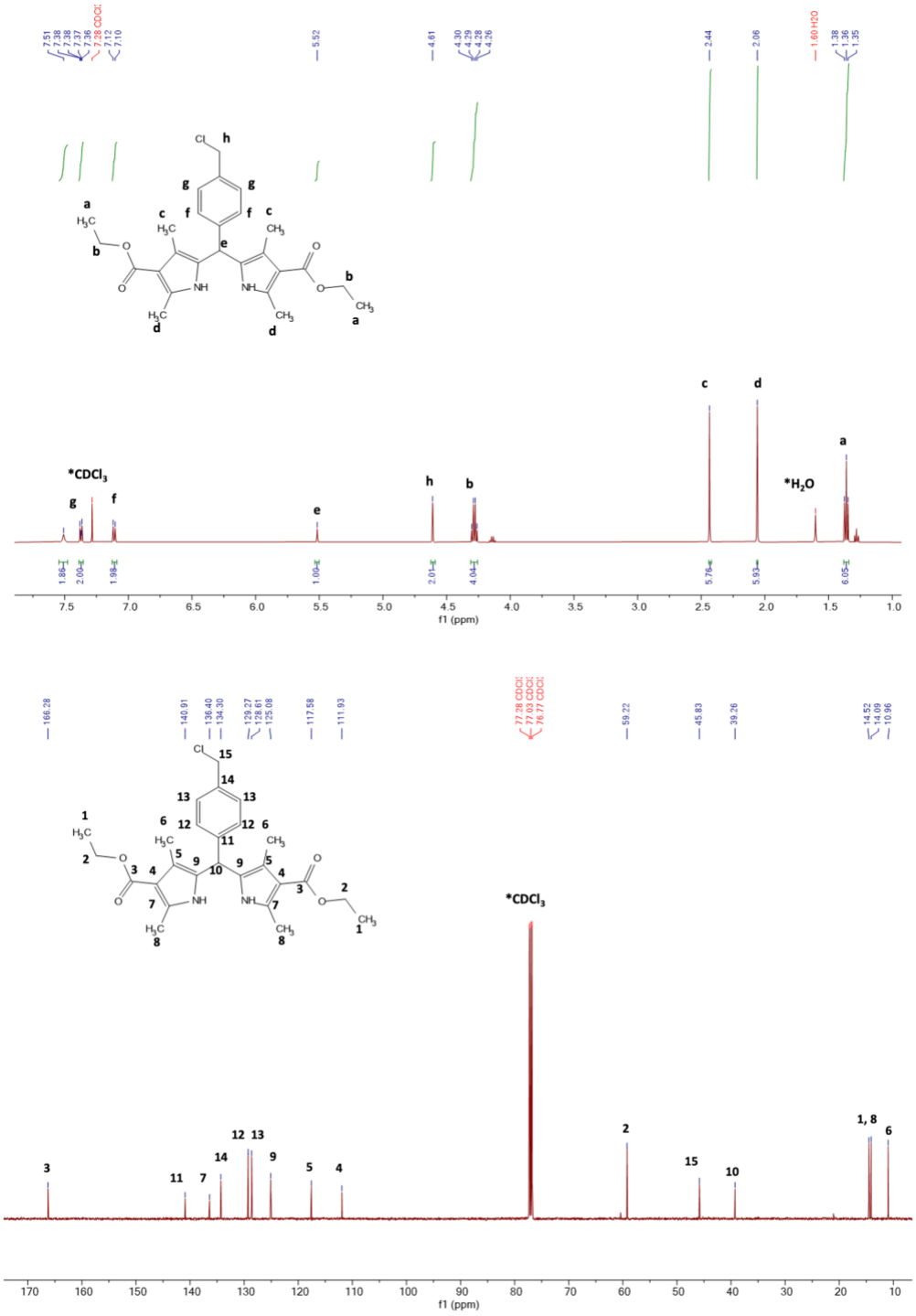


Figure S2.3.1. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.51 (s, 2H), 7.39 – 7.35 (m, 2H), 7.11 (d, J = 8.1 Hz, 2H), 5.52 (s, 1H), 4.61 (s, 2H), 4.28 (q, J = 7.1 Hz, 4H), 2.44 (s, 6H), 2.06 (s, 6H), 1.36 (t, J = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.28, 140.91, 136.40, 134.30, 129.27, 128.61, 125.08, 117.58, 111.93, 59.22, 45.83, 39.26, 14.52, 14.09, 10.96.

S2.3.2. Synthesis of 2C

To an oven-dried round-bottom flask, diethyl 5,5'-((4-(chloromethyl)phenyl)methylene)bis(2,4-dimethyl-1*H*-pyrrole-3-carboxylate) **1C** was dissolved in 40 mL of anhydrous DCM and cooled to 0 ºC. DDQ (1576 mg, 7.0 mmol, 1.1 equiv.) was then added to the reaction mixture and stirred at 0 ºC for 30 minutes. Then, triethylamine (4390 µL, 31.5 mmol, 5 equiv.) was added dropwise at 0 ºC and allowed to stir for 1 hour. Boron trifluoride diethyl etherate (8942 mg, 63 mmol, 10 equiv.) was added to the reaction mixture dropwise at 0 ºC, then stirred at room temperature for 16 hours. Upon completion, the reaction mixture was quenched with dilute $NaHCO_3$ and extracted with DCM. The organic extracts were combined, and crude residue was further purified by column chromatography using a mixture of hexane and DCM (40: 60) to get intermediate product diethyl 10-(4-(chloromethyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **2C** as an orange powder (1888 mg). This compound was isolated with 58% yield.

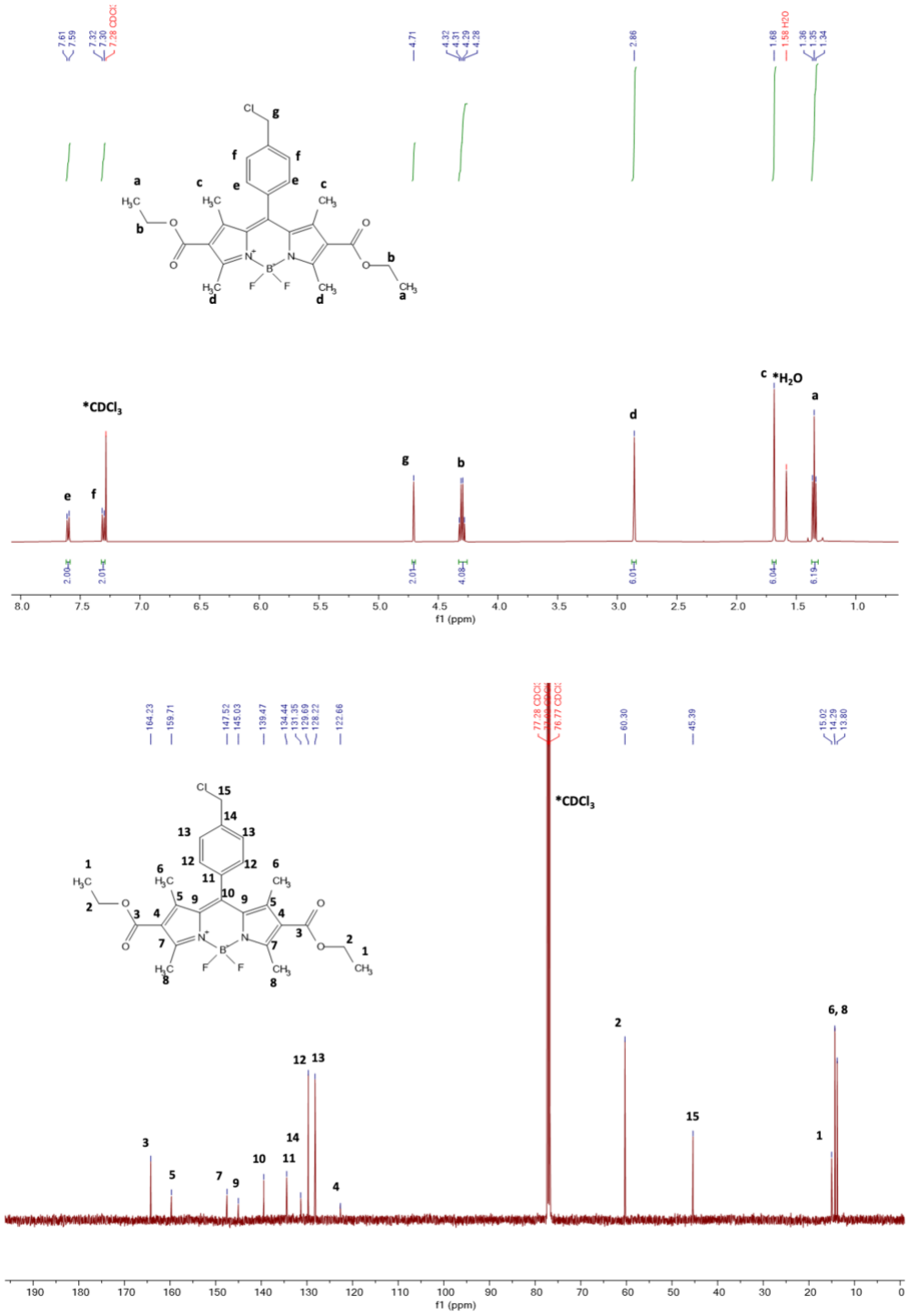


Figure S2.3.2. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.60 (d, J = 8.2 Hz, 2H), 7.31 (d, J = 8.1 Hz, 2H), 4.71 (s, 2H), 4.30 (q, J = 7.1 Hz, 4H), 2.86 (s, 6H), 1.68 (s, 6H), 1.35 (t, J = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.23, 159.71, 147.52, 145.03, 139.47, 134.44, 131.35, 129.69, 128.22, 122.66, 60.30, 45.39, 15.02, 14.29, 13.80.

S2.3.3. Synthesis of 3C

Diethyl 10-(4-(chloromethyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **2C** (200 mg, 0.4 mmol, 1 equiv.) was dissolved in 30 mL of chloroform. Methylamine (40%, w/v, 10 mL) was added to the reaction mixture before being stirred for 16 hours at room temperature. After the reaction was complete, volatiles were removed under reduced pressure and the crude residue was further purified by column chromatography using a mixture of ethyl acetate and methanol (8:2) to get diethyl 5,5-difluoro-1,3,7,9-tetramethyl-10-(4-((methylamino)methyl)phenyl)-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **3C** as an orange powder (189 mg). This compound was isolated with 95.5% yield.

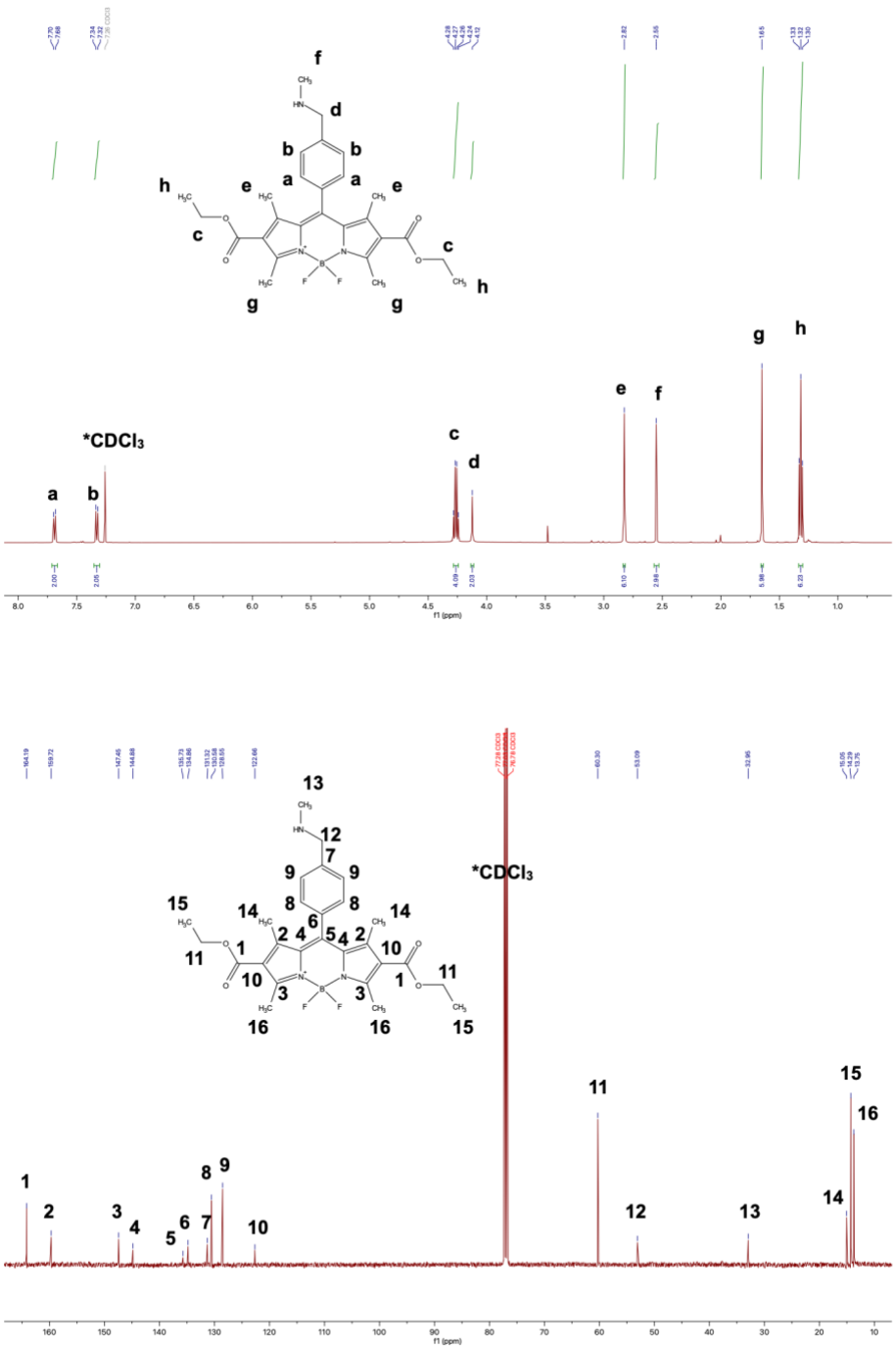


Figure S2.3.3. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.69 (d, *J* = 8.1 Hz, 2H), 7.33 (d, *J* = 8.1 Hz, 2H), 4.27 (t, *J* = 7.1 Hz, 4H), 4.12 (s, 2H), 2.82 (s, 6H), 2.55 (s, 3H), 1.65 (s, 6H), 1.32 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.19, 159.72, 147.45, 144.88, 135.73, 134.86, 131.32, 130.58, 128.55, 122.66, 60.30, 53.09, 32.95, 15.05, 14.29, 13.75.

S2.3.4. Synthesis of C

Diethyl 5,5-difluoro-1,3,7,9-tetramethyl-10-(4-((methylamino)methyl)phenyl)-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **3C** (33 mg, 0.1 mmol, 1 equiv.), 2-(bromomethyl)phenylboronic acid (21 mg, 0.1 mmol, 1.5 equiv.), and potassium carbonate (18 mg, 0.1 mmol, 2 equiv.) were dissolved in 5 mL of acetonitrile. The reaction mixture was left to stir at 60 ºC for 16 hours. Volatiles were removed under reduced pressure. The crude residue was further purified by neutral alumina column chromatography using a mixture of ethyl acetate and hexane (95: 5) to get product 10-(4-(((2-boronobenzyl)(methyl)amino)methyl)phenyl)-2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium **C** as an orange powder (15 mg). This compound was isolated with 25.8% yield.

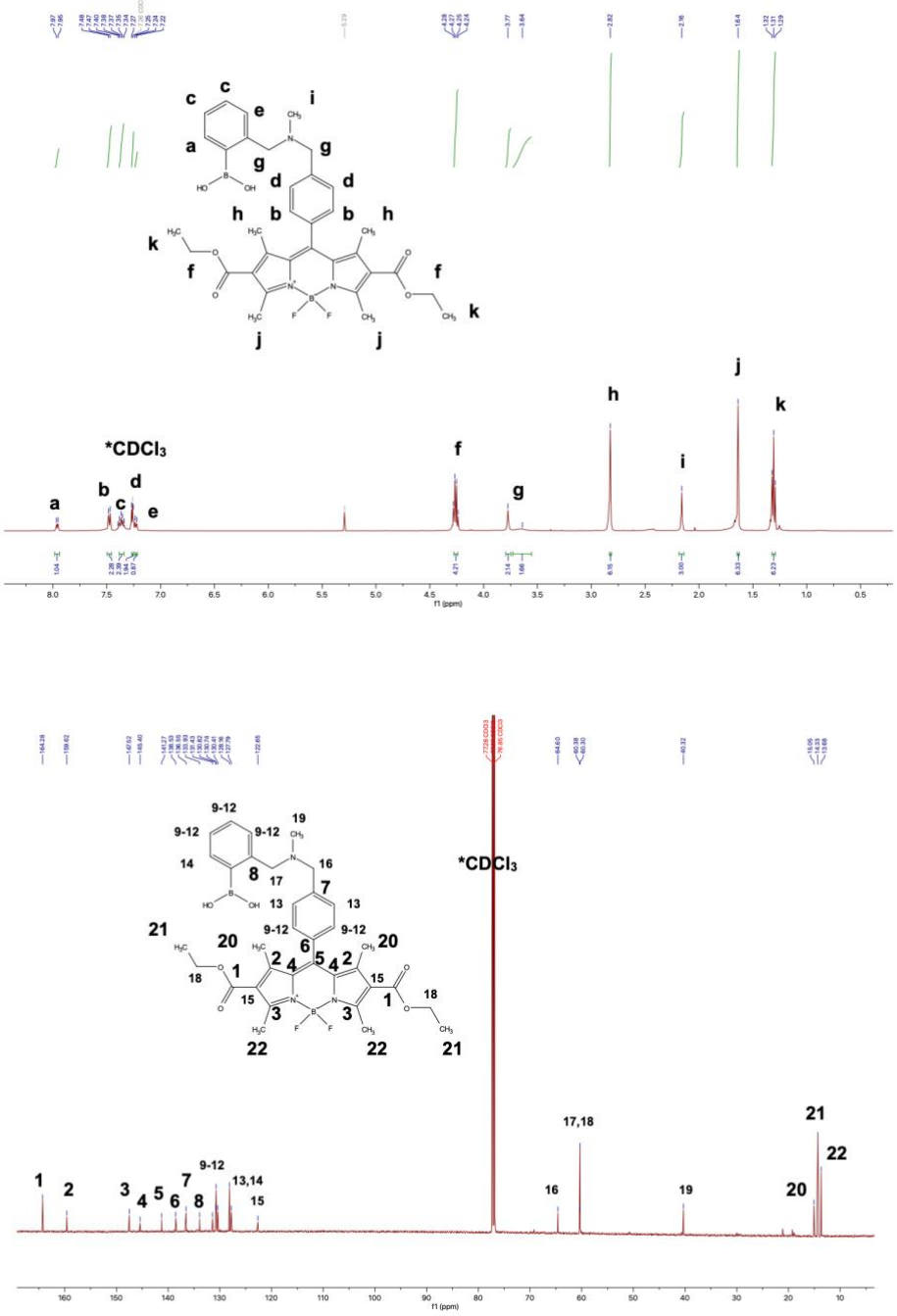
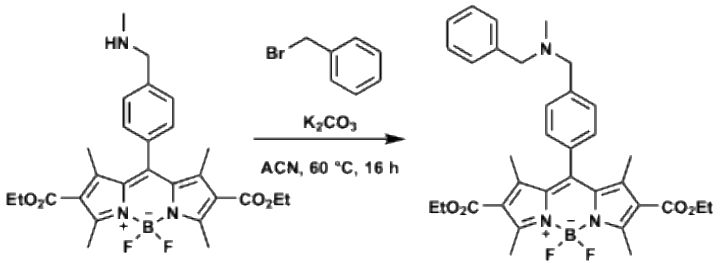


Figure S2.3.4. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.96 (d, *J* = 8.1 Hz, 1H), 7.48 (d, *J* = 7.9 Hz, 2H), 7.38 – 7.34 (m, 2H), 7.25 (s, 2H), 7.23 (d, *J* = 7.0 Hz, 1H), 4.26 (d, *J* = 7.1 Hz, 4H), 3.77 (s, 2H), 3.64 (s, 2H), 2.82 (s, 6H), 2.16 (s, 3H), 1.64 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (151 MHz, Chloroform-*d*)

δ 164.28, 159.62, 147.52, 145.40, 141.27, 138.53, 136.55, 133.93, 131.43, 130.82, 130.74, 130.41, 128.16, 127.79, 122.65, 64.60, 60.38, 60.30, 40.32, 15.05, 14.33, 13.68.

S2.4. Synthesis of C*



Diethyl 5,5-difluoro-1,3,7,9-tetramethyl-10-(4-((methylamino)methyl)phenyl)-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **3C** (100 mg, 0.2 mmol, 1 equiv.), benzyl bromide (34 µL, 0.3 mmol, 1.5 equiv.), and potassium carbonate (52 mg, 0.4 mmol, 2 equiv.) were dissolved in 5 mL of acetonitrile. The reaction mixture was left to stir at 60 ºC for 16 hours. Volatiles were removed under reduced pressure, and the crude residue was further purified by column chromatography using a mixture of hexane and ethyl acetate (9:1) to get product diethyl 10-(4-((benzyl(methyl)amino)methyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **C*** as an orange solid (36 mg). This compound was isolated with 31.5% yield.

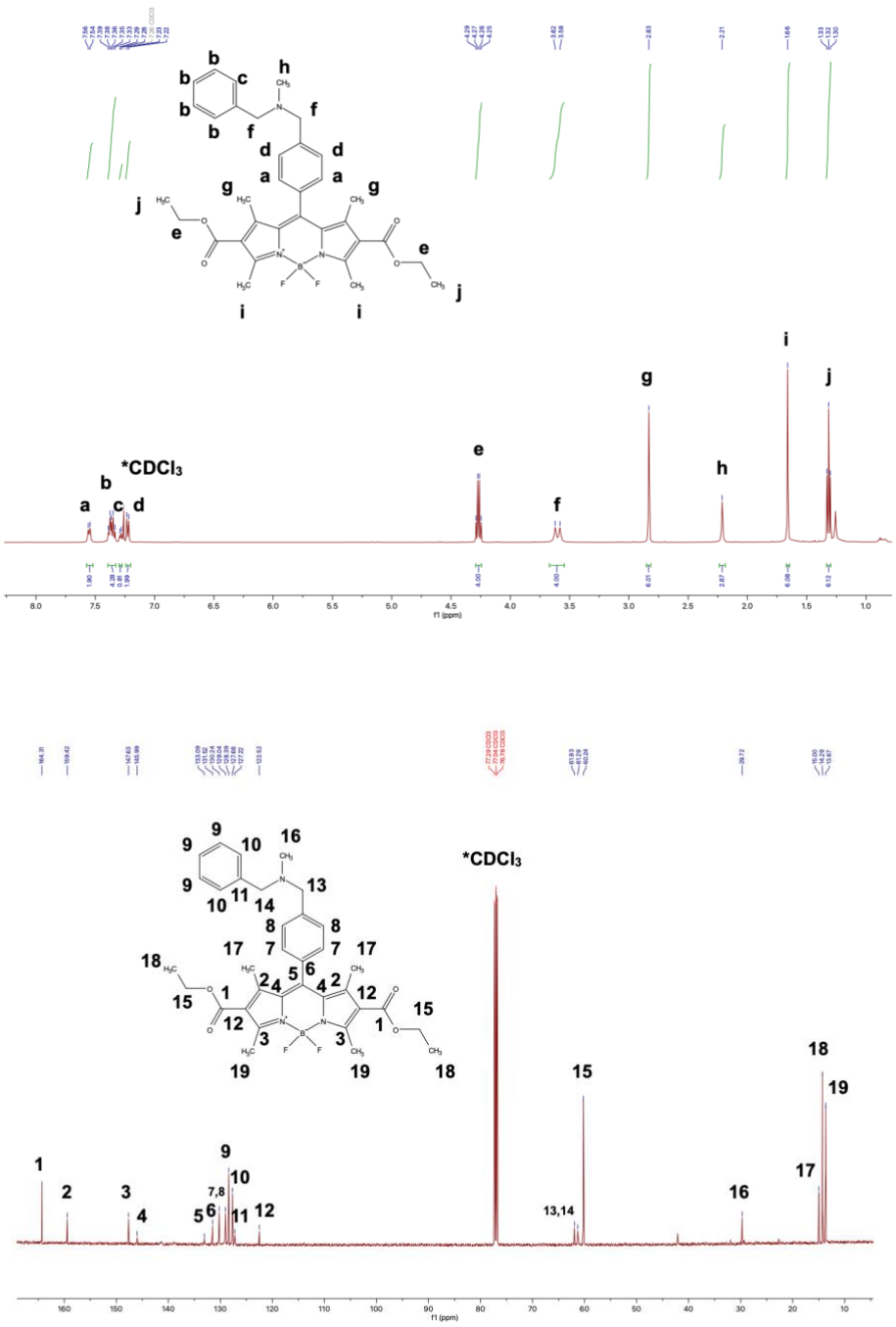
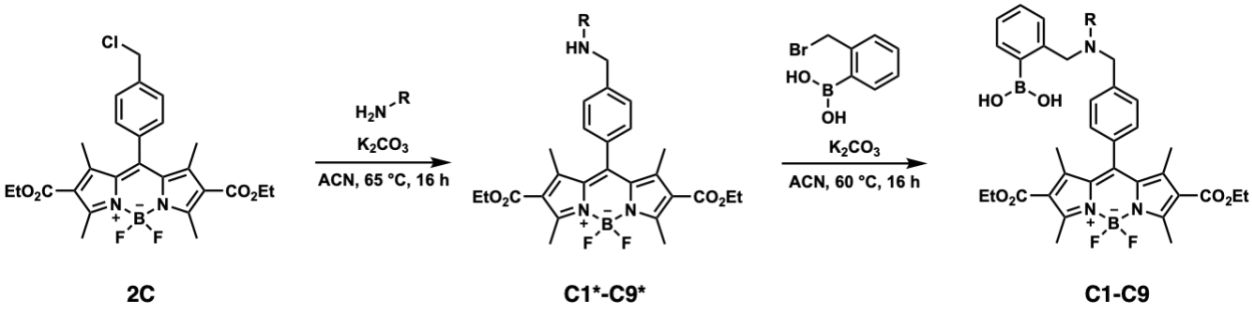


Figure S2.4. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.55 (d, *J* = 7.6 Hz, 2H), 7.36 (dt, *J* = 14.9, 7.4 Hz, 4H), 7.29 (d, *J* = 7.0 Hz, 1H), 7.23 (d, *J* = 7.9 Hz, 2H), 4.27 (q, *J* = 7.1 Hz, 4H), 3.60 (d, *J* = 19.6 Hz, 4H), 2.83 (s, 6H), 2.21 (s, 3H), 1.66 (s, 6H), 1.32 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*)

δ 164.31, 159.42, 147.63, 145.99, 133.09, 131.52, 130.24, 129.04, 128.39, 127.68, 127.22, 122.52, 61.93, 61.29, 60.24, 29.72, 15.00, 14.29, 13.67.

S2.5. General Synthesis of C1*–C9*



Diethyl 10-(4-(chloromethyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate **2C** (1 equiv.), amine derivatives (2.5 equiv.), and potassium carbonate (2 equiv.) were dissolved in 5 mL of acetonitrile. The reaction mixture was left to stir at 60 ºC for 16 hours. Once completed, the reaction mixture was filtered and volatiles were removed under reduced pressure. The crude residue was further purified by column chromatography to afford intermediates C1*–C9*. See the table for yield and characteristics.

Label	2C used (mg)	Derivatives	Eluent	Characteristics	Isolated yield (%)
C1*	150	2-methoxyethan-1-amine	8:2 (EA: MeOH)	Orange powder	26.7
C2*	150	2-methylbutan-1-amine	9:1 (EA: MeOH)	Orange powder	21.3
C3*	142	2-(2-(2-methoxyethoxy)ethoxy)ethan-1-amine	8:2 (EA: MeOH)	Orange powder	43.2
C4*	200	2-(methylthio)ethan-1-amine	8:2 (EA: MeOH)	Orange powder	36.8
C5*	150	2-fluoroethan-1-amine	9:1 (EA: MeOH)	Orange powder	35.5
C6*	200	*N*-(2-aminoethyl)acetamide	7:3 (EA: MeOH)	Orange powder	34.8
C7*	200	methyl glycinate	8:2 (EA: MeOH)	Orange powder	47.6
C8*	200	3-aminopropanamide	7:3 (DCM: MeOH)	Orange powder	25.9
C9*	200	2-aminoethan-1-ol	7:3 (EA: MeOH)	Orange powder	43.3

S2.5.1. Characterization of C1*

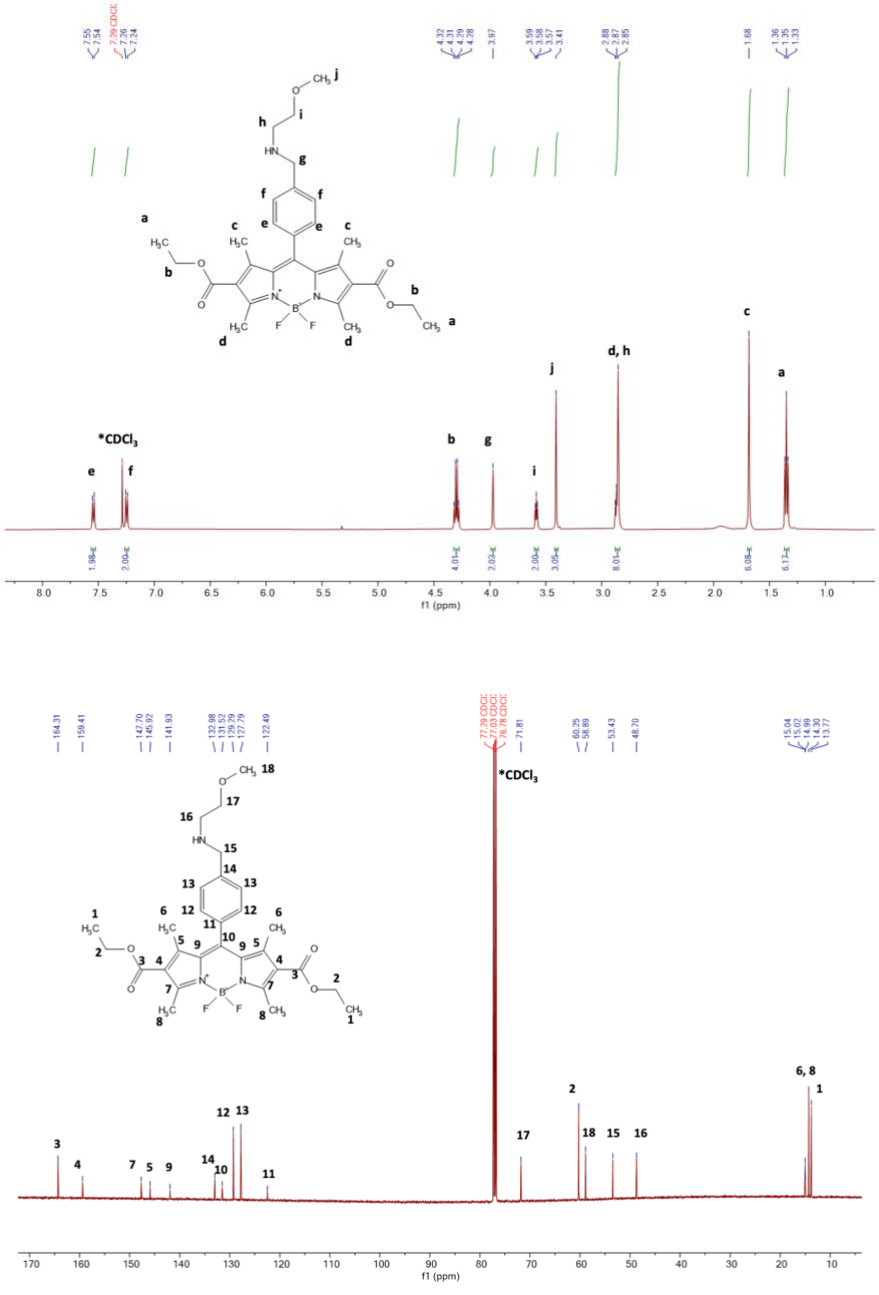


Figure S2.5.1. Diethyl 5,5-difluoro-10-(4-(((2-methoxyethyl)amino)methyl)phenyl)-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate, ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.54 (d, *J* = 7.7 Hz, 2H), 7.25 (d, *J* = 7.9 Hz, 2H), 4.30 (q, *J* = 7.1 Hz, 4H), 3.97 (s, 2H), 3.58 (t, *J* = 5.0 Hz, 2H), 3.41 (s, 3H), 2.86 (d, *J* = 7.2 Hz, 8H), 1.68 (s, 6H), 1.35 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz,

$CDCl_3$) δ 164.31, 159.41, 147.70, 145.92, 141.93, 132.98, 131.52, 129.29, 127.79, 122.49, 77.24, 71.81, 60.25, 58.89, 53.43, 48.70, 15.04, 15.02, 14.99, 14.30, 13.77.

S2.5.2. Characterization of C2*

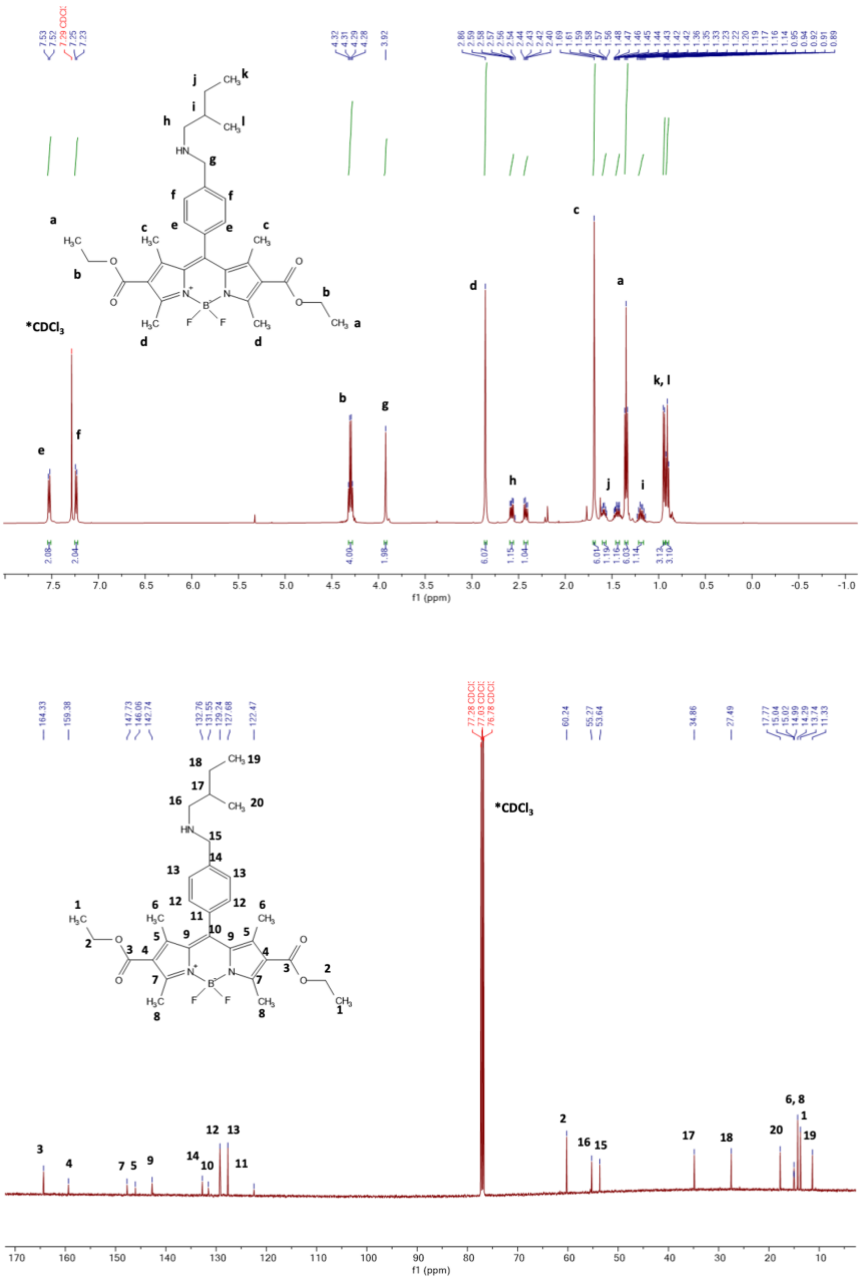


Figure S2.5.2. Diethyl 5,5-difluoro-1,3,7,9-tetramethyl-10-(4-(((2-methylbutyl)amino)methyl)phenyl)-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate, ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.53 (d, *J* = 7.8 Hz, 2H), 7.24 (d, *J* = 8.0 Hz, 2H), 4.30 (q, *J* = 7.1 Hz, 4H), 3.92 (s, 2H), 2.86 (s, 6H), 2.57

(dd, *J* = 11.4, 5.9 Hz, 1H), 2.42 (dd, *J* = 11.5, 7.2 Hz, 1H), 1.69 (s, 6H), 1.60 – 1.56 (m, 1H), 1.44 (dd, *J* = 7.8, 5.3 Hz, 1H), 1.35 (t, *J* = 7.1 Hz, 6H), 1.19 (dd, *J* = 14.3, 6.8 Hz, 1H), 0.94 (d, *J* = 6.6 Hz, 3H), 0.91 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.33, 159.38, 147.73, 146.06, 142.74, 132.76, 131.55, 129.24, 127.68, 122.47, 60.24, 55.27, 53.64, 34.86, 27.49, 17.77, 15.04, 15.02, 14.99, 14.29, 13.74, 11.33.

S2.5.3. Characterization of C3*

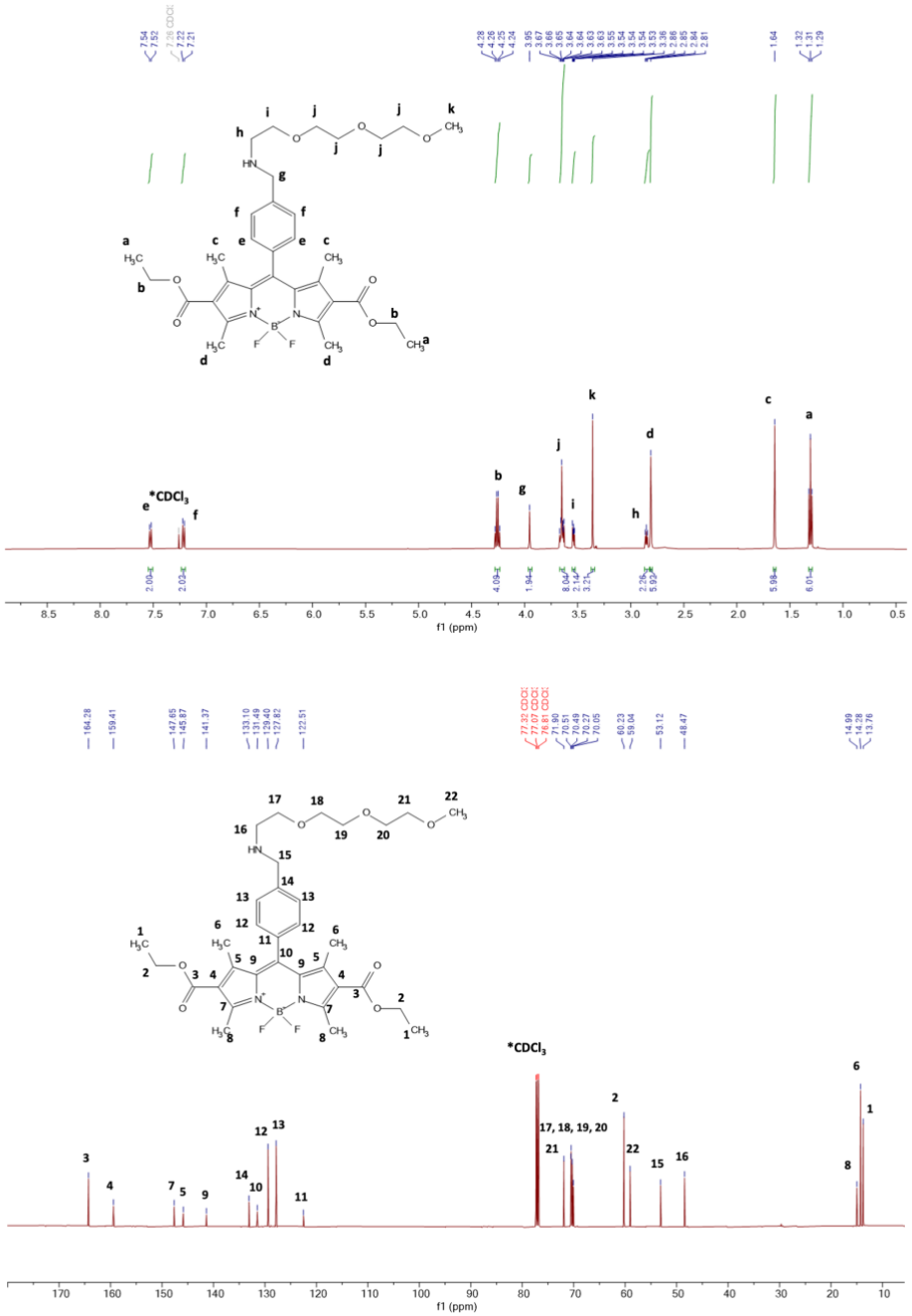
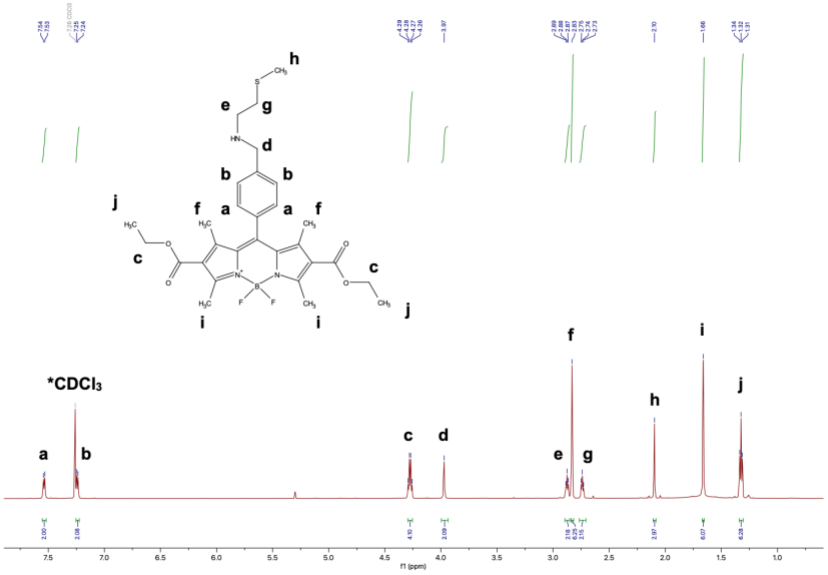


Figure S2.5.3. Diethyl 10-(4-(5,8,11-trioxa-2-azadodecyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate, ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.53 (d, *J* = 8.1 Hz, 2H), 7.22 (d, *J* = 8.1 Hz, 2H), 4.26 (q, *J* = 7.1 Hz, 4H), 3.95 (s, 2H), 3.67 – 3.62 (m, 8H), 3.55 – 3.52 (m, 2H), 3.36 (s, 3H), 2.87 – 2.82 (m, 2H), 2.81 (s, 6H), 1.64 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.28, 159.41, 147.65, 145.87, 141.37, 133.10, 131.49, 129.40, 127.82, 122.51, 71.90, 70.56, 70.51, 70.49, 70.27, 70.05, 69.29, 60.23, 59.04, 53.12, 48.47, 14.99, 14.28, 13.76.

S2.5.4. Characterization of C4*



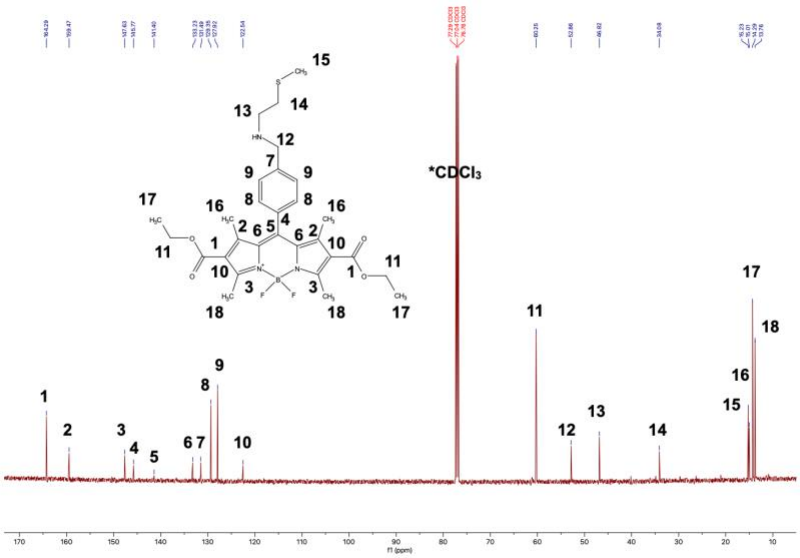
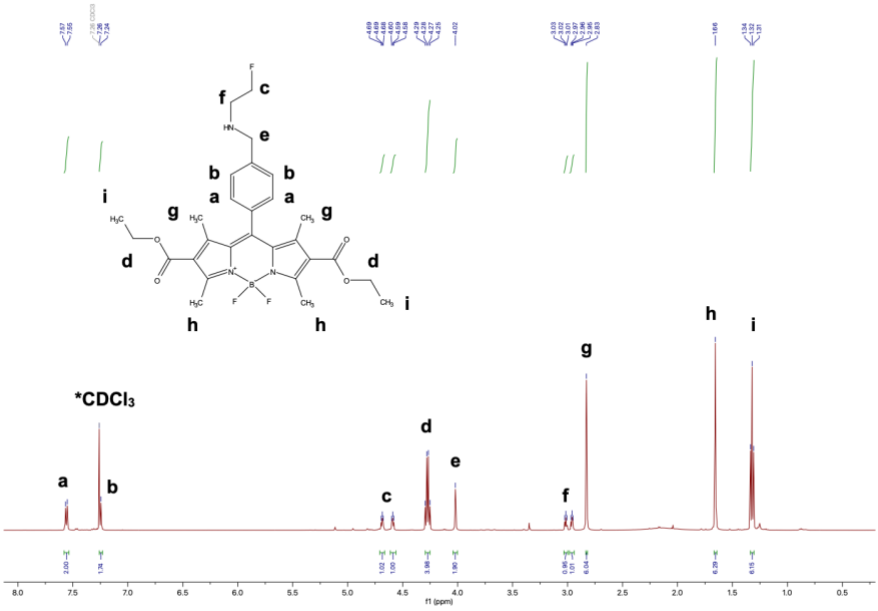


Figure S2.5.4. Diethyl 5,5-difluoro-1,3,7,9-tetramethyl-10-(4-(((2-(methylthio)ethyl)amino)methyl)phenyl)-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinine-2,8-dicarboxylate, ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.54 (d, *J* = 7.6 Hz, 2H), 7.24 (d, *J* = 6.3 Hz, 2H), 4.27 (q, *J* = 6.8 Hz, 4H), 3.97 (s, 2H), 2.88 (t, *J* = 6.1 Hz, 2H), 2.83 (s, 6H), 2.74 (t, *J* = 6.2 Hz, 2H), 2.10 (s, 3H), 1.66 (s, 6H), 1.32 (t, *J* = 6.9 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.29, 159.47, 147.63, 145.77, 141.40, 133.23, 131.49, 129.35, 127.92, 122.54, 60.25, 52.86, 46.82, 34.08, 15.23, 15.01, 14.29, 13.76.

S2.5.5. Characterization of C5*



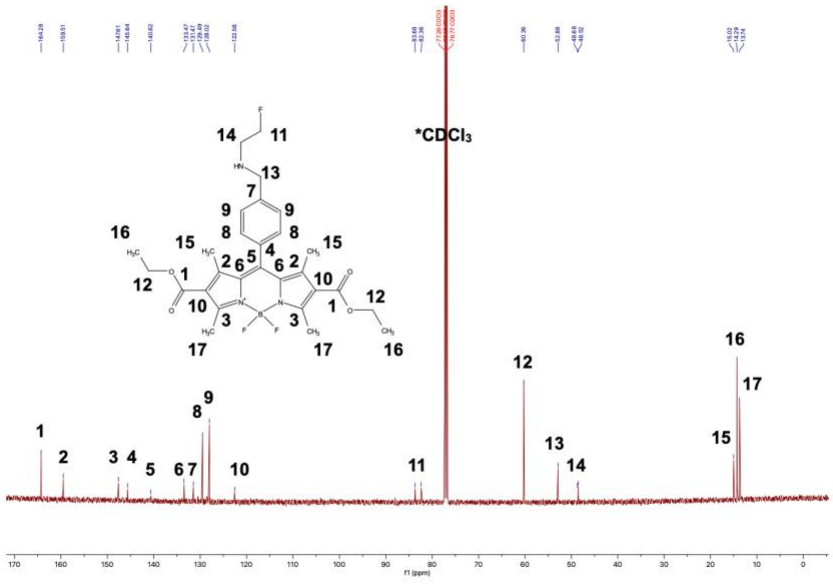


Figure S2.5.5. 2,8-bis(ethoxycarbonyl)-5,5-difluoro-10-(4-(((2-fluoroethyl)amino)methyl)phenyl)-1,3,7,9-tetramethyl-5*H*-5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.56 (d, *J* = 7.9 Hz, 2H), 7.24 (s, 2H), 4.69 (t, *J* = 4.7 Hz, 1H), 4.59 (t, *J* = 4.7 Hz, 1H), 4.27 (d, *J* = 7.1 Hz, 4H), 4.02 (s, 2H), 3.03 – 3.00 (m, 1H), 2.96 (t, *J* = 4.7 Hz, 1H), 2.83 (s, 6H), 1.66 (s, 6H), 1.32 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.28, 159.51, 147.61, 145.64, 140.62, 133.47, 131.47, 129.49, 128.02, 122.56, 83.68, 82.36, 60.26, 52.88, 48.68, 48.52, 15.02, 14.29, 13.74.

S2.5.6. Characterization of C6*

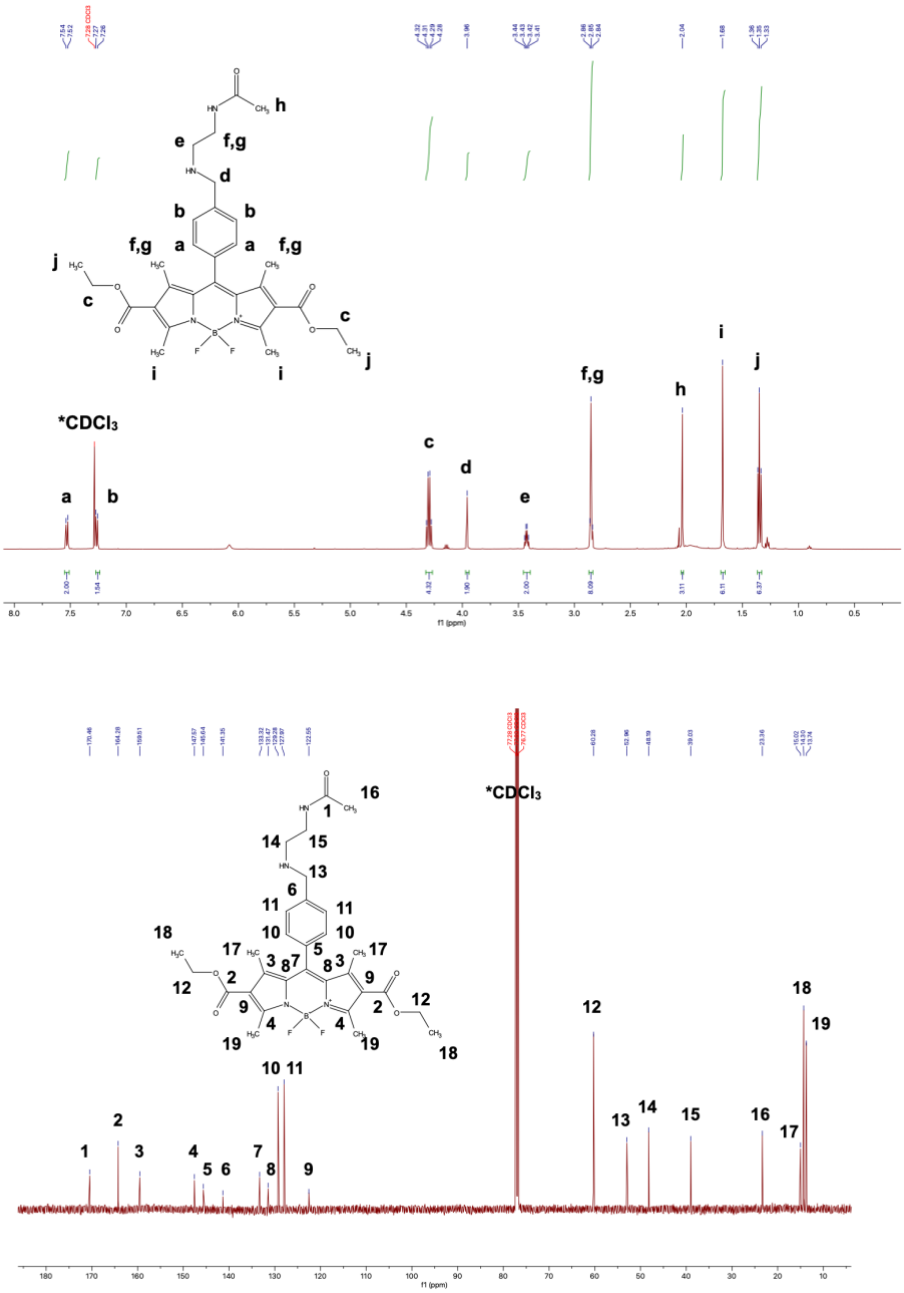


Figure S2.5.6. 10-(4-(((2-acetamidoethyl)amino)methyl)phenyl)-2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.53 (d, *J* = 8.0 Hz, 2H), 7.26 (s, 2H), 4.30 (q, *J* = 7.1 Hz, 4H), 3.96 (s, 2H), 3.43 (q, *J* = 5.7 Hz, 2H), 2.85 (s, 8H), 2.04 (s, 3H), 1.68 (s, 6H), 1.35 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz,

Chloroform-*d*) δ 170.46, 164.28, 159.51, 147.57, 145.64, 141.35, 133.32, 131.47, 129.28, 127.97, 122.55, 60.28, 52.96, 48.19, 39.03, 23.36, 15.02, 14.30, 13.74.

S2.5.7. Characterization of C7*

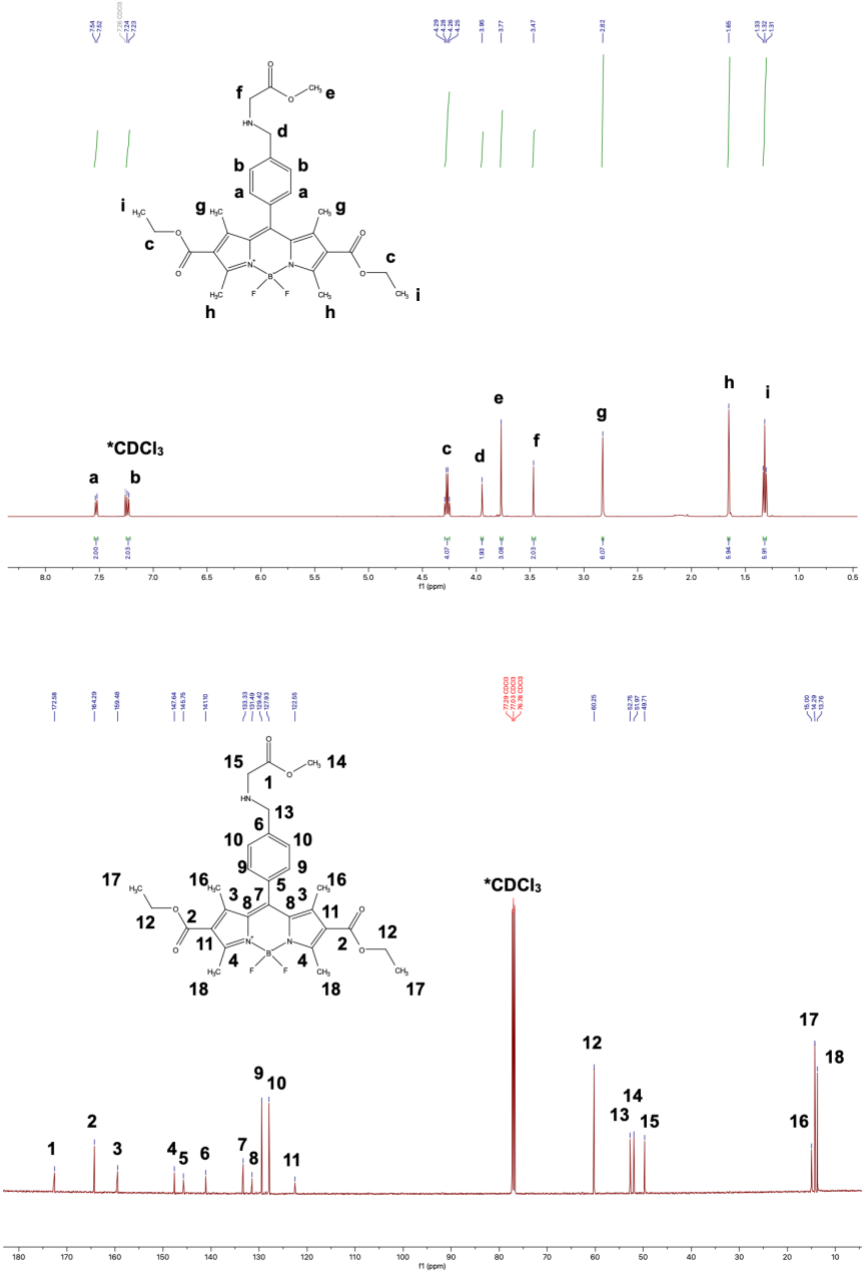
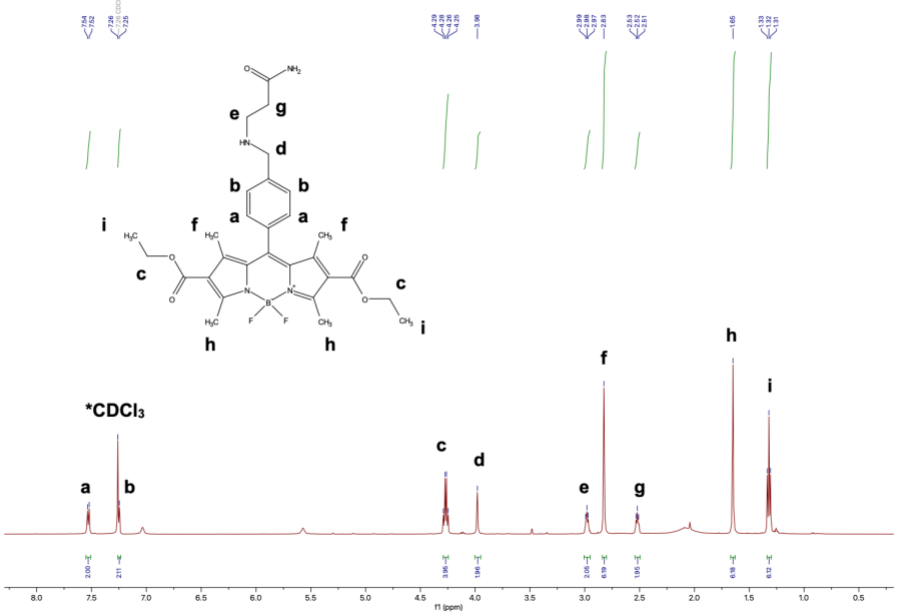


Figure S2.5.7. 2,8-bis(ethoxycarbonyl)-5,5-difluoro-10-(4-(((2-methoxy-2-oxoethyl)amino)methyl)phenyl)-1,3,7,9-tetramethyl-5*H*-5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.53 (d, J = 8.1 Hz, 2H), 7.24 (d, J = 8.1 Hz, 2H), 4.27 (q, J =

7.1 Hz, 4H), 3.95 (s, 2H), 3.77 (s, 3H), 3.47 (s, 2H), 2.82 (s, 6H), 1.65 (s, 6H), 1.32 (t, J = 7.1 Hz, 6H).

^{13}C NMR (126 MHz, Chloroform-*d*) δ 172.58, 164.29, 159.48, 147.64, 145.75, 141.10, 133.33, 131.49, 129.42, 127.93, 122.55, 60.25, 52.75, 51.97, 49.71, 15.00, 14.29, 13.76.

S2.5.8. Characterization of C8*



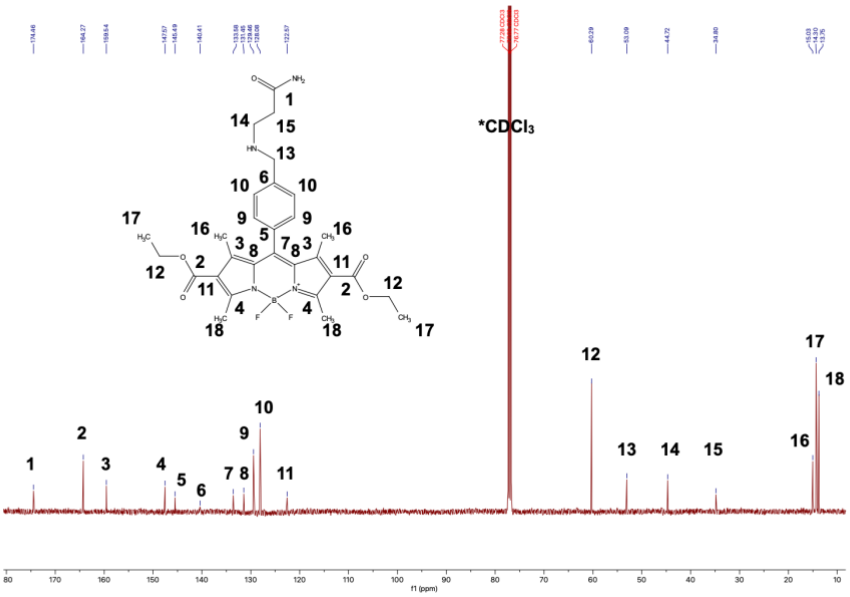
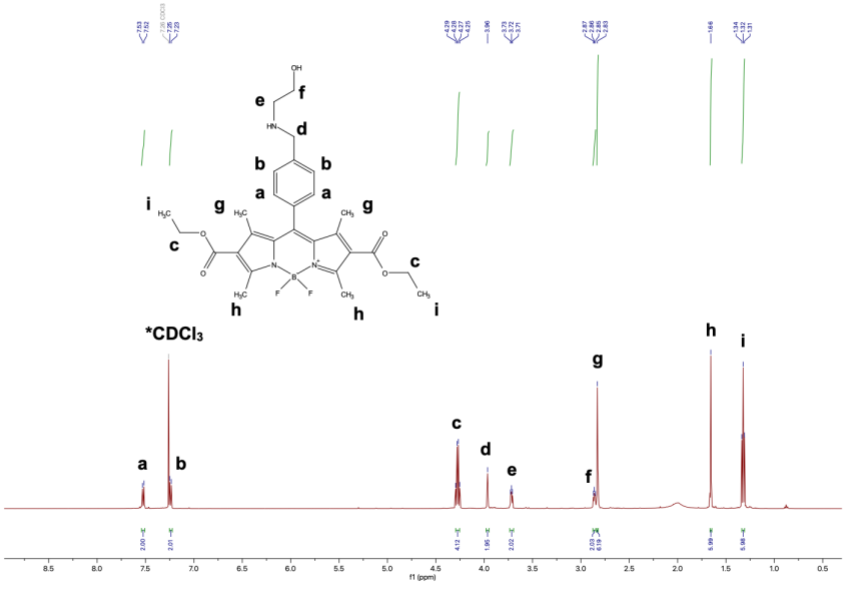


Figure S2.5.8. 10-(4-(((3-amino-3-oxopropyl)amino)methyl)phenyl)-2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.53 (d, *J* = 7.7 Hz, 2H), 7.25 (s, 2H), 4.27 (q, *J* = 7.1 Hz, 4H), 3.98 (s, 2H), 2.98 (t, *J* = 5.9 Hz, 2H), 2.83 (s, 6H), 2.52 (t, *J* = 5.9 Hz, 2H), 1.65 (s, 6H), 1.32 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 174.46, 164.27, 159.54, 147.57, 145.49, 140.41, 133.58, 131.45, 129.46, 128.08, 122.57, 60.29, 53.09, 44.72, 34.80, 15.03, 14.30, 13.75.

S2.5.9. Characterization of C9*



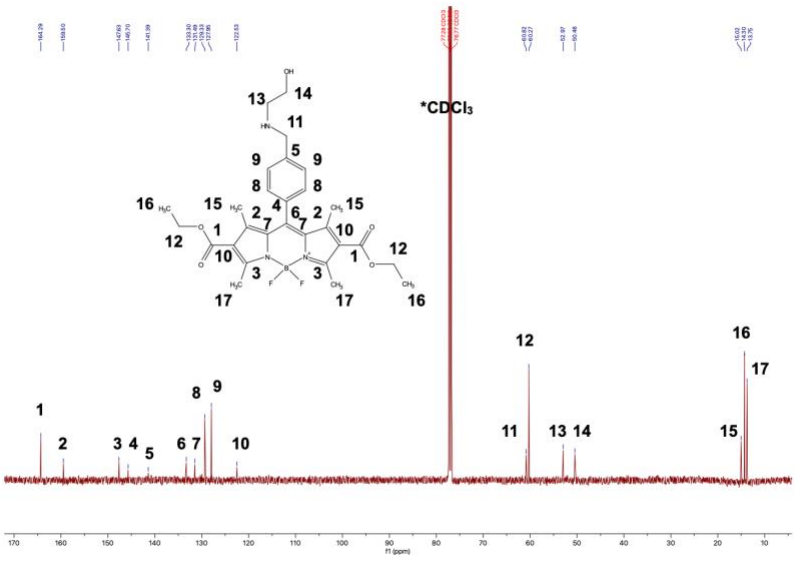
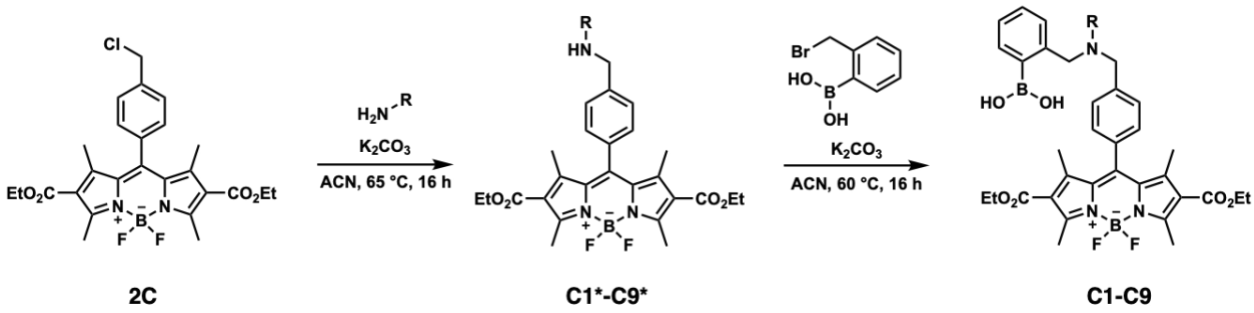


Figure S2.5.9. 2,8-bis(ethoxycarbonyl)-5,5-difluoro-10-(4-(((2-hydroxyethyl)amino)methyl)phenyl)-1,3,7,9-tetramethyl-5*H*-5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.52 (d, *J* = 7.9 Hz, 2H), 7.24 (d, *J* = 8.1 Hz, 2H), 4.29 – 4.25 (m, 4H), 3.96 (s, 2H), 3.74 – 3.70 (m, 2H), 2.86 (t, *J* = 5.1 Hz, 2H), 2.83 (s, 6H), 1.66 (s, 6H), 1.32 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.29, 159.50, 147.63, 145.70, 141.39, 133.30, 131.49, 129.33, 127.95, 122.53, 60.82, 60.27, 52.97, 50.46, 15.02, 14.30, 13.75.

S2.6. Synthesis of C1–C9



C1*–C9* derivatives (1 equiv.), 2-(bromomethyl)phenylboronic acid (1.5 equiv.), and potassium carbonate (2 equiv.) were dissolved in 5 mL of acetonitrile. The reaction mixture was left to stir at 60 ºC for 16 hours. The reaction mixture was then filtered, and volatiles were removed under reduced pressure. The crude residue was further purified by neutral alumina column chromatography to afford product C1–C9. To minimize grease, purified products were subsequently dissolved in acetonitrile and extracted with 3 times 10 mL pentane. C8 was eluted in HPLC instead of neutral alumina column due to polarity. See the table for yield and characteristics.

Label	CX* used (mg)	Amine Derivatives	Eluent	Characteristics	Isolated yield (%)
C1	34	C1*	95:5 (EA: MeOH)	Orange powder	79.8
C2	29	C2*	95:5 (EA: MeOH)	Orange powder	71.2
C3	37	C3*	95:5 (EA: MeOH)	Orange powder	45.1
C4	80	C4*	95:5 (EA: MeOH)	Orange powder	54.7
C5	56	C5*	95:5 (EA: MeOH)	Orange powder	44.4
C6	68	C6*	95:5 (EA: MeOH)	Orange powder	58.1
C7	92	C7*	95:5 (EA: MeOH)	Orange powder	60.4
C8	70	C8*	20%ACN to 80% ACN in H_2O over 20 minutes	Orange powder	35.4 (conversion calculated on HPLC)
C9	80	C9*	95:5 (EA: MeOH)	Orange powder	37.2

S2.6.1. Characterization of C1

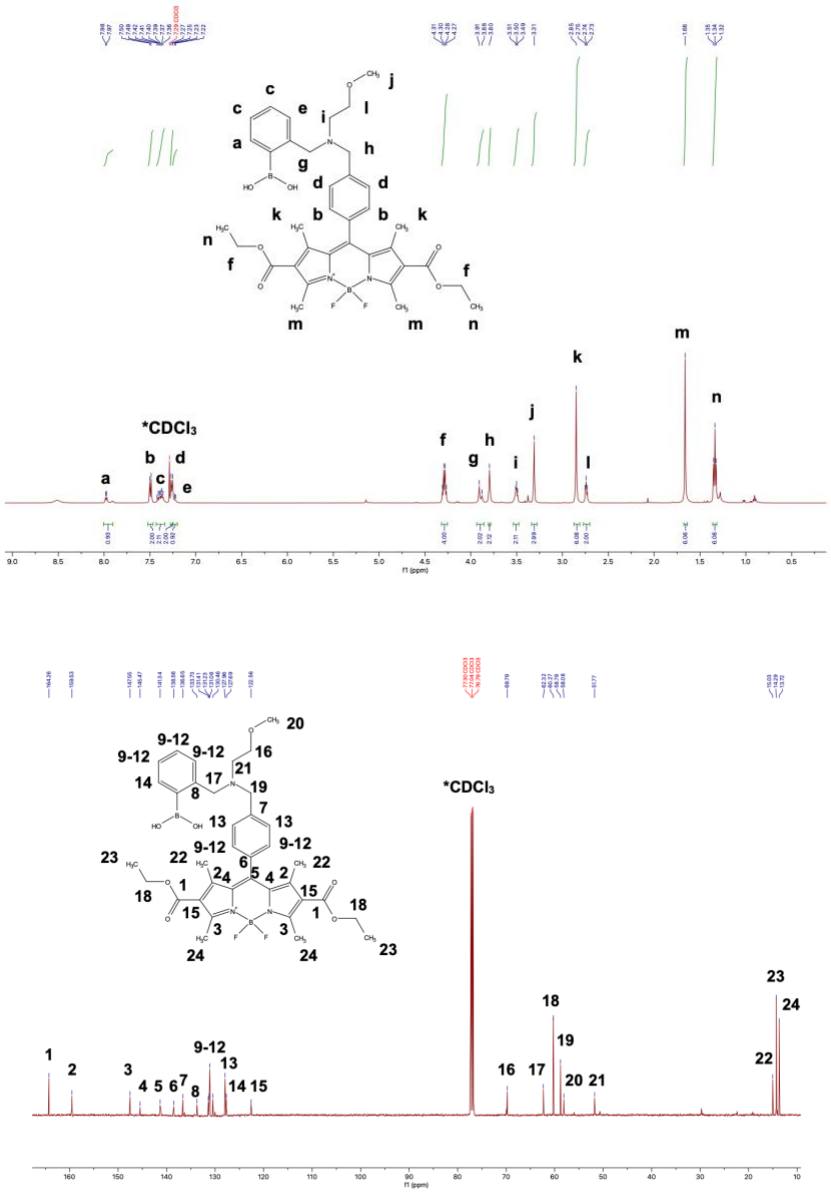


Figure S2.6.1. (2-(((4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4, 5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(2-methoxyethyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.98 (d, *J* = 5.3 Hz, 1H), 7.49 (d, *J* = 7.6 Hz, 2H), 7.39 (dt, *J* = 19.3, 6.9 Hz, 2H), 7.26 (d, *J* = 7.7 Hz, 2H), 7.22 (d, *J* = 5.6 Hz, 1H), 4.29 (q, *J* = 7.1 Hz, 4H), 3.89 (d, *J* = 14.6 Hz, 2H), 3.80 (s, 2H), 3.50 (t, *J* = 5.2 Hz, 2H), 3.31 (s, 3H), 2.85 (s, 6H), 2.74 (t, *J* = 5.6 Hz, 2H), 1.66 (s, 6H), 1.34 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.26, 159.53, 147.55, 145.47,

141.34, 138.56, 136.65, 133.73, 131.41, 131.23, 131.08, 130.46, 127.96, 127.69, 122.56, 69.78, 62.32, 60.27, 58.78, 58.08, 51.77, 15.03, 14.29, 13.72.

S2.6.2. Characterization of C2

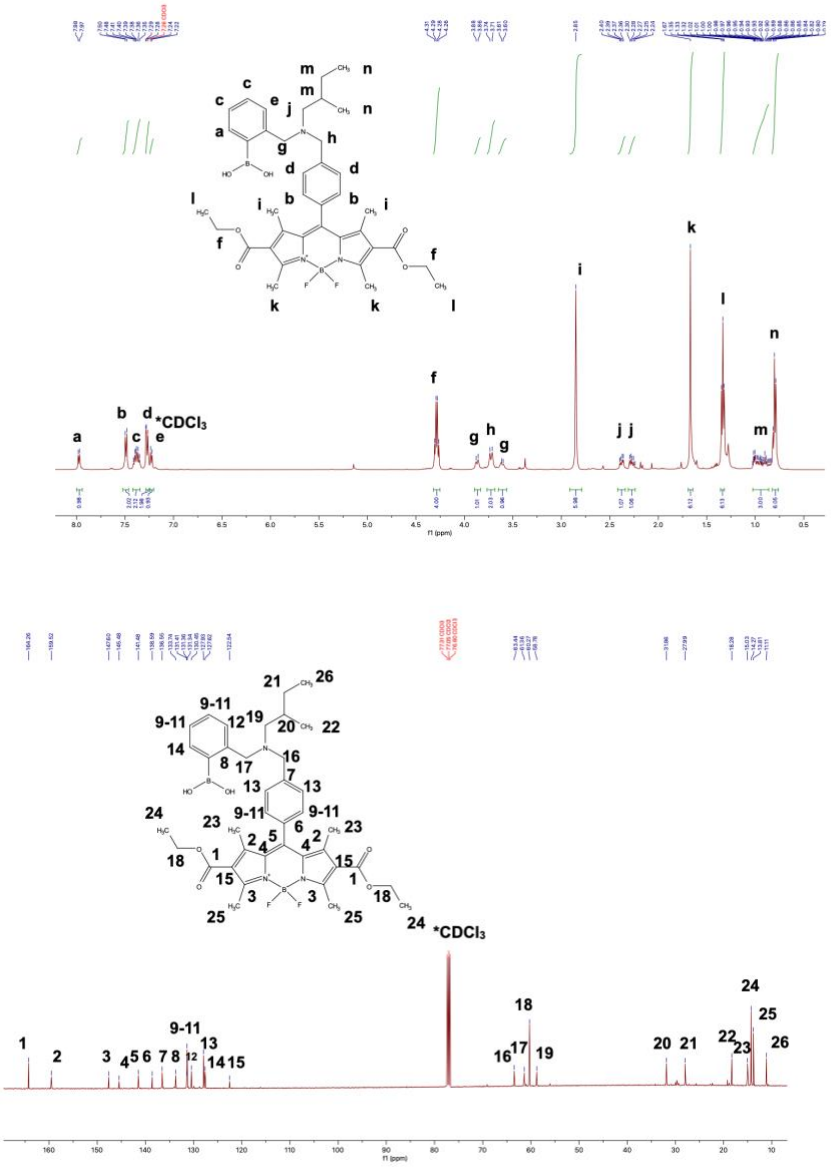
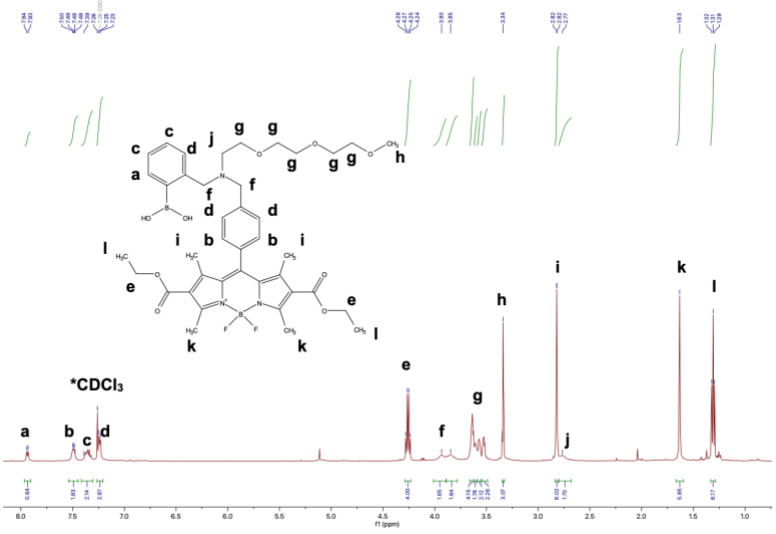


Figure S2.6.2. (2-(((4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(2-methylbutyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.98 (d, *J* = 6.1 Hz, 1H), 7.49 (d, *J* = 7.9 Hz, 2H), 7.38 (dq, *J* = 15.4, 7.0 Hz, 2H), 7.28 (s, 2H), 7.23 (d, *J* = 7.1 Hz, 1H), 4.29 (q, *J* = 7.1 Hz, 4H), 3.87 (d, *J* = 10.0 Hz, 1H),

3.72 (d, *J* = 11.6 Hz, 2H), 3.61 (d, *J* = 7.0 Hz, 1H), 2.85 (s, 6H), 2.38 (dd, *J* = 12.8, 5.6 Hz, 1H), 2.28 (dd, *J* = 12.9, 7.6 Hz, 1H), 1.67 (s, 6H), 1.33 (t, *J* = 7.1 Hz, 6H), 1.03 – 0.86 (m, 3H), 0.80 (t, *J* = 7.3 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.26, 159.52, 147.60, 145.48, 141.48, 138.59, 136.55, 133.74, 131.41, 131.36, 131.34, 130.45, 127.93, 127.62, 122.54, 63.44, 61.36, 60.27, 58.78, 31.86, 27.99, 18.28, 15.03, 14.27, 13.81, 11.11.

S2.6.3. Characterization of C3



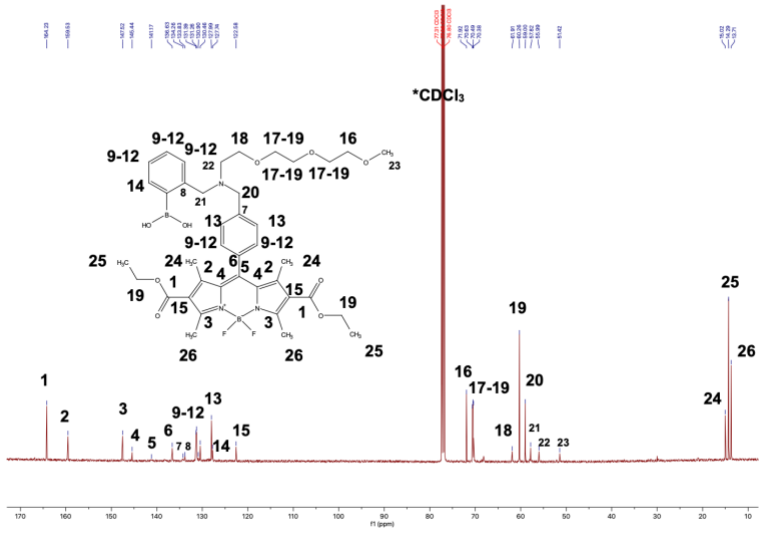
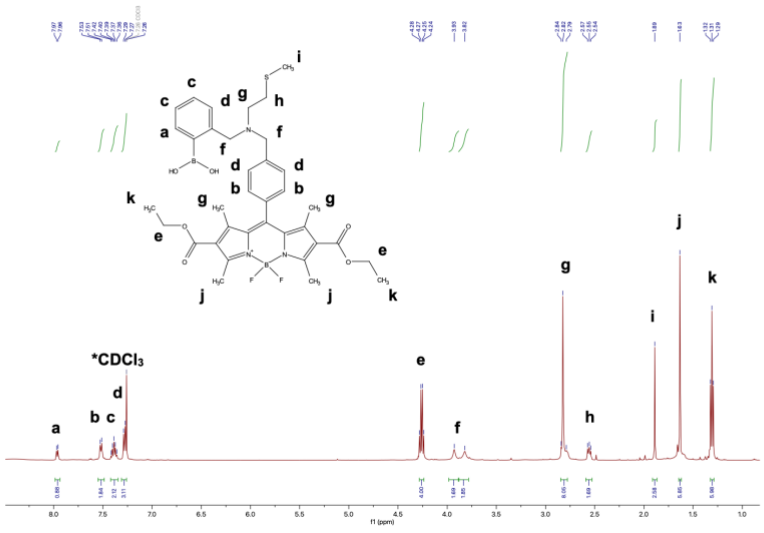


Figure S2.6.3. (2-(2-(4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)-5,8,11-trioxa-2-azadodecyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 7.1 Hz, 1H), 7.54 – 7.45 (m, 2H), 7.39 (s, 2H), 7.26 – 7.21 (m, 3H), 4.26 (q, *J* = 7.1 Hz, 4H), 3.93 (s, 2H), 3.85 (s, 2H), 3.64 (q, *J* = 4.7 Hz, 4H), 3.60 (d, *J* = 3.4 Hz, 2H), 3.59 – 3.55 (m, 2H), 3.53 (dd, *J* = 5.8, 3.5 Hz, 2H), 3.34 (s, 3H), 2.84 – 2.80 (m, 6H), 2.77 (s, 2H), 1.63 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.23, 159.53, 147.52, 145.44, 141.17, 136.63, 134.26, 133.83, 131.39, 131.26, 130.90 130.46, 127.99, 127.74, 122.58, 71.92, 70.63, 70.49, 70.38, 61.91, 60.26, 59.00, 57.82, 55.99, 51.42, 15.02, 14.29, 13.71.

S2.6.4. Characterization of C4



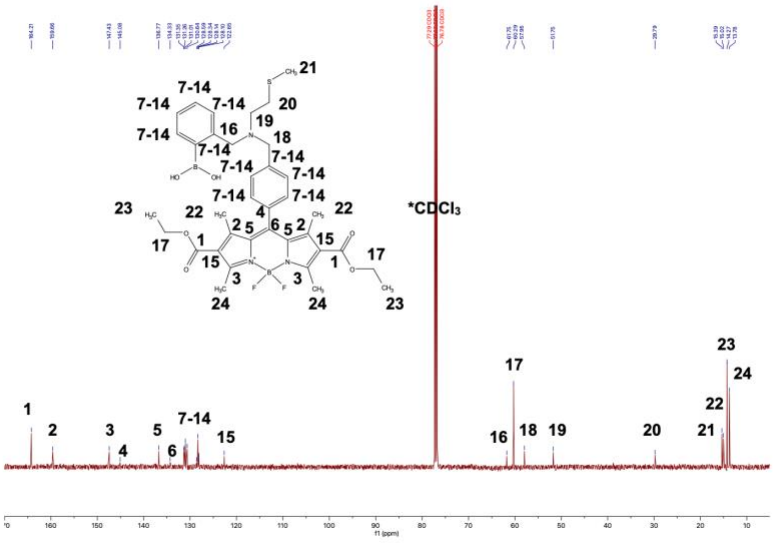


Figure S2.6.4. (2-(((4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(2-(methylthio)ethyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.96 (d, *J* = 7.1 Hz, 1H), 7.52 (d, *J* = 7.7 Hz, 2H), 7.39 (p, *J* = 7.2 Hz, 2H), 7.31 – 7.26 (m, 3H), 4.26 (q, *J* = 7.1 Hz, 4H), 3.93 (s, 2H), 3.82 (s, 2H), 2.82 (s, 8H), 2.59 – 2.53 (m, 2H), 1.89 (s, 3H), 1.63 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.21, 159.66, 147.43, 145.08, 136.77, 134.33, 131.35, 131.26, 131.01, 130.64, 128.59, 128.34, 128.14, 128.10, 122.65, 61.75, 60.29, 57.95, 51.75, 29.79, 15.39, 15.02, 14.27, 13.78.

S2.6.5. Characterization of C5

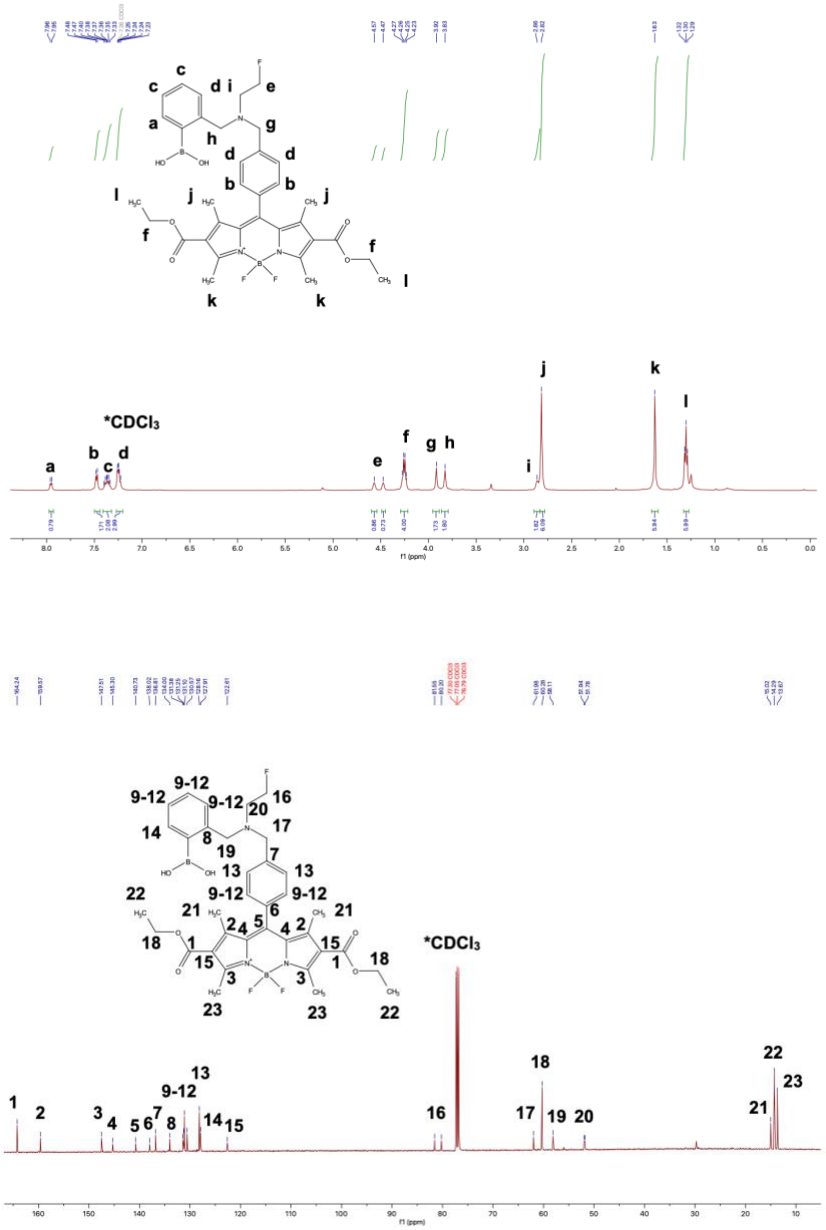


Figure S2.6.5. (2-(((4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(2-fluoroethyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.95 (d, *J* = 7.1 Hz, 1H), 7.48 (d, *J* = 7.8 Hz, 2H), 7.36 (dt, *J* = 14.5, 7.1 Hz, 2H), 7.24 (dd, *J* = 8.5, 5.7 Hz, 3H), 4.57 (s, 1H), 4.47 (s, 1H), 4.25 (q, *J* = 6.9 Hz, 4H), 3.92 (s, 2H), 3.83 (s, 2H), 2.86 (s, 2H), 2.82 (s, 6H), 1.63 (s, 6H), 1.30 (t, *J* = 7.0 Hz, 6H). ^{13}C NMR (126 MHz,

Chloroform-*d*) δ 164.24, 159.57, 147.51, 145.30, 140.73, 138.02, 136.81, 134.00, 131.38, 131.25, 131.10, 130.57, 128.16, 127.91, 122.61, 81.55, 80.20, 61.98, 60.28, 58.11, 51.94, 51.78, 15.02, 14.29, 13.67.

S2.6.6. Characterization of C6

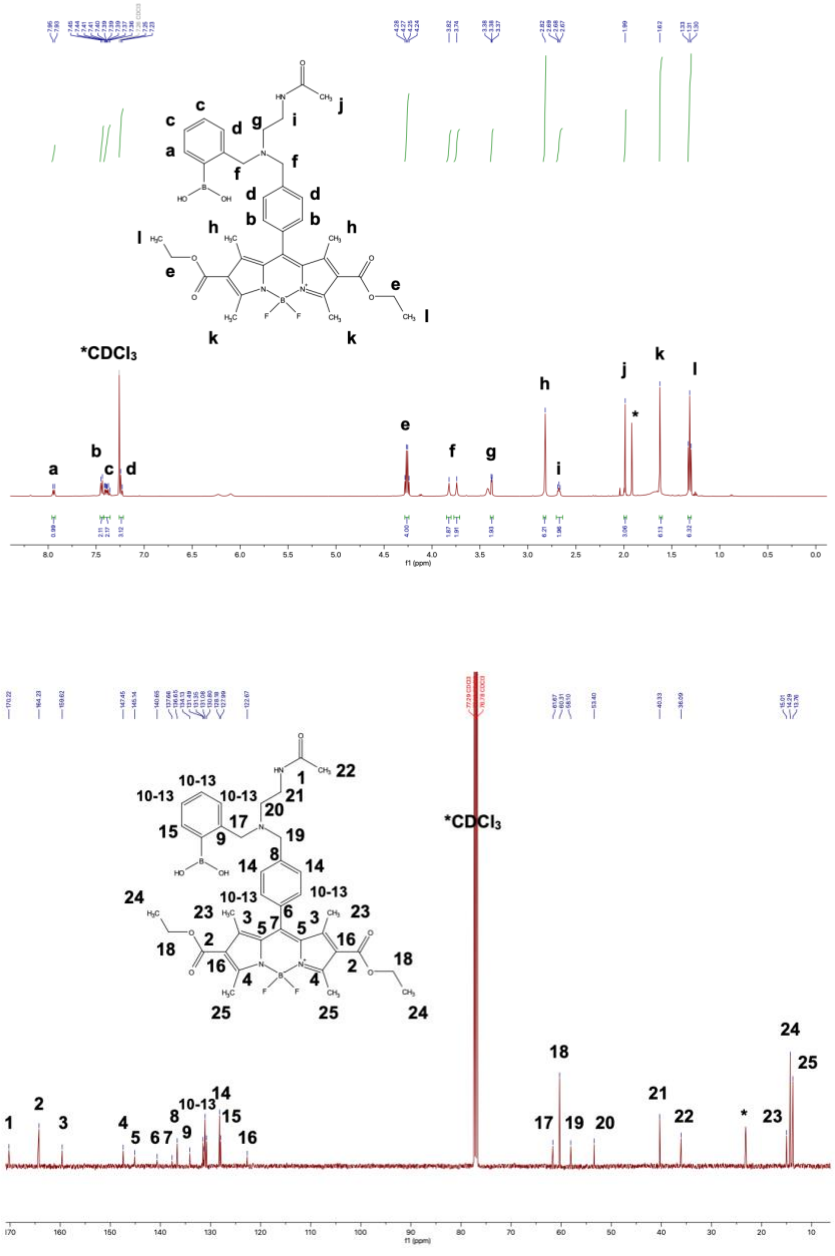


Figure S2.6.6. (2-(((2-acetamidoethyl)(4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 8.7 Hz, 1H), 7.44 (d, *J* = 8.0 Hz, 2H), 7.42 – 7.35 (m, 2H),

7.24 (d, *J* = 8.1 Hz, 3H), 4.27 (t, *J* = 7.1 Hz, 4H), 3.82 (s, 2H), 3.74 (s, 2H), 3.39 – 3.36 (m, 2H), 2.82 (s, 6H), 2.68 (t, *J* = 5.5 Hz, 2H), 1.99 (s, 3H), 1.62 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 170.22, 164.23, 159.62, 147.45, 145.14, 140.65, 137.66, 136.65, 134.13, 131.49, 131.35, 131.08, 130.80, 128.18, 127.99, 122.67, 61.67, 60.31, 58.10, 53.40, 40.33, 36.09, 15.01, 14.29, 13.76.

S2.6.7. Characterization of C7

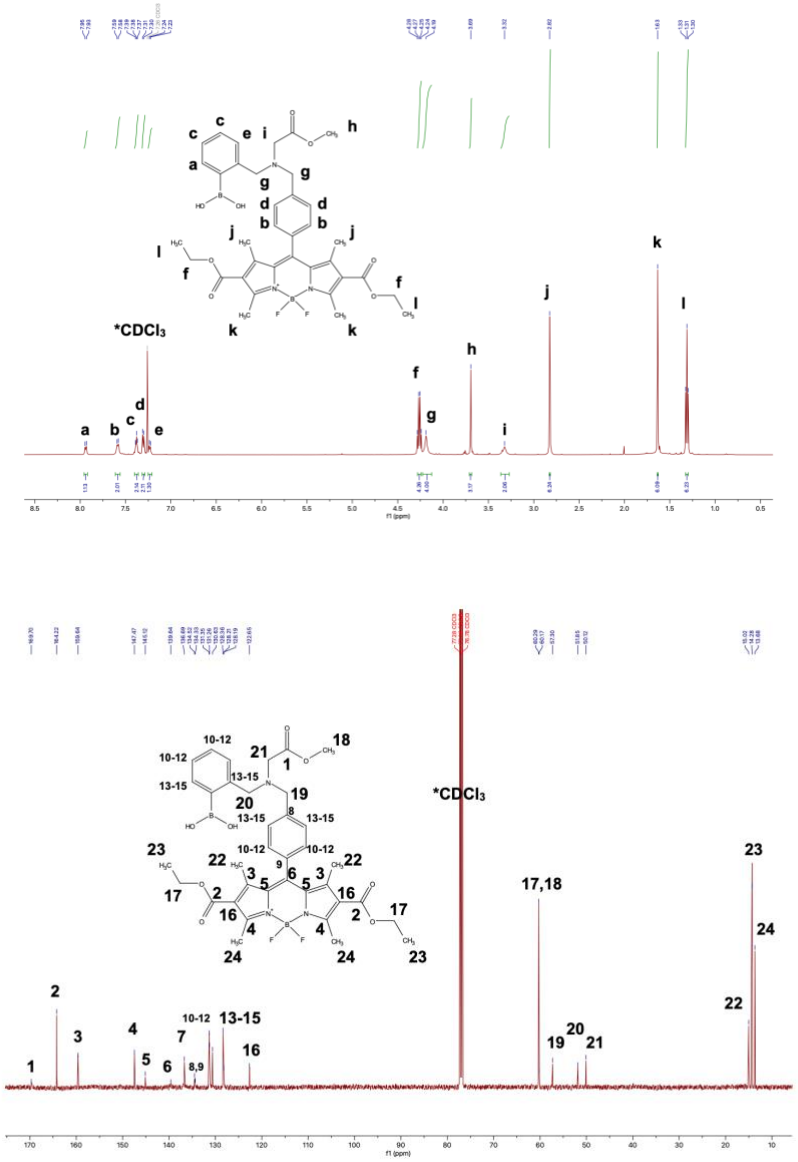


Figure S2.6.7. (2-(((4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(2-methoxy-2-

oxoethyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 8.4 Hz, 1H), 7.59 (d, *J* = 7.0 Hz, 2H), 7.38 (t, *J* = 5.4 Hz, 2H), 7.30 (d, *J* = 7.7 Hz, 2H), 7.23 (d, *J* = 6.4 Hz, 1H), 4.26 (d, *J* = 7.1 Hz, 4H), 4.19 (s, 4H), 3.69 (s, 3H), 3.32 (s, 2H), 2.82 (s, 6H), 1.63 (s, 6H), 1.31 (t, *J* = 7.1 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 169.70, 164.22, 159.64, 147.47, 145.12, 136.69, 134.52, 134.33, 131.35, 131.26, 130.63, 130.36, 128.36, 128.21, 128.19, 122.65, 60.29, 60.17, 57.30, 51.85, 50.12, 15.02, 14.28, 13.68.

S2.6.8. Characterization of C8

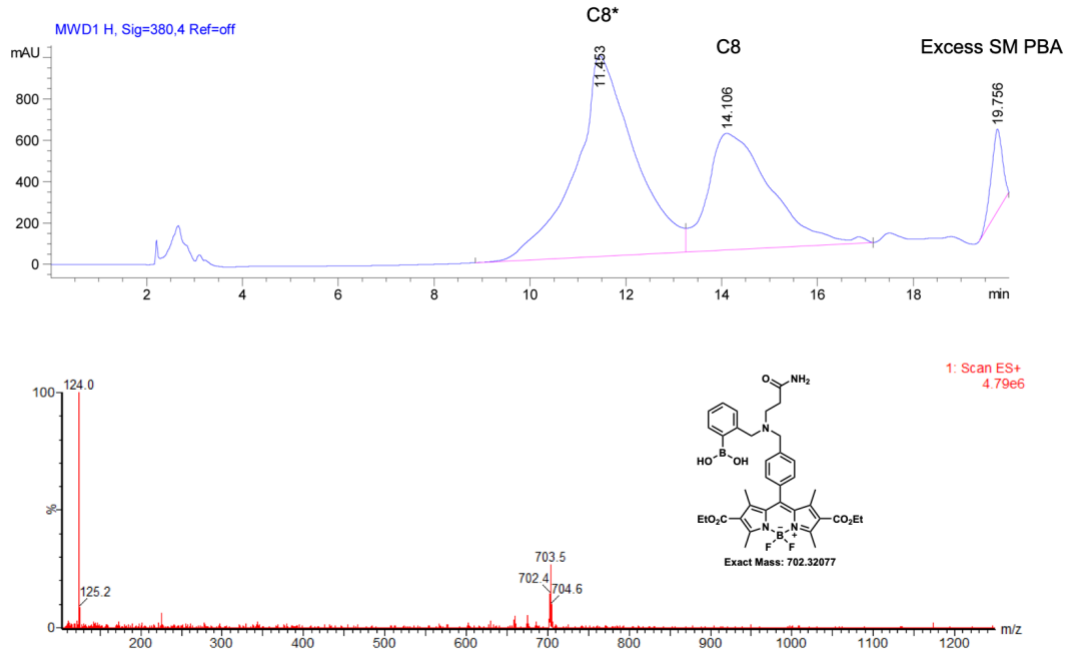


Figure S2.6.8. (2-(((3-amino-3-oxopropyl)(4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^{4},5λ^{4}-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)amino)methyl)phenyl)boronic acid. This product was isolated at ~0.7 mg, and therefore, no NMR experiment could be performed with good resolution. However, HRMS trace was obtained on the isolated peak.

S2.6.9. Characterization of C9

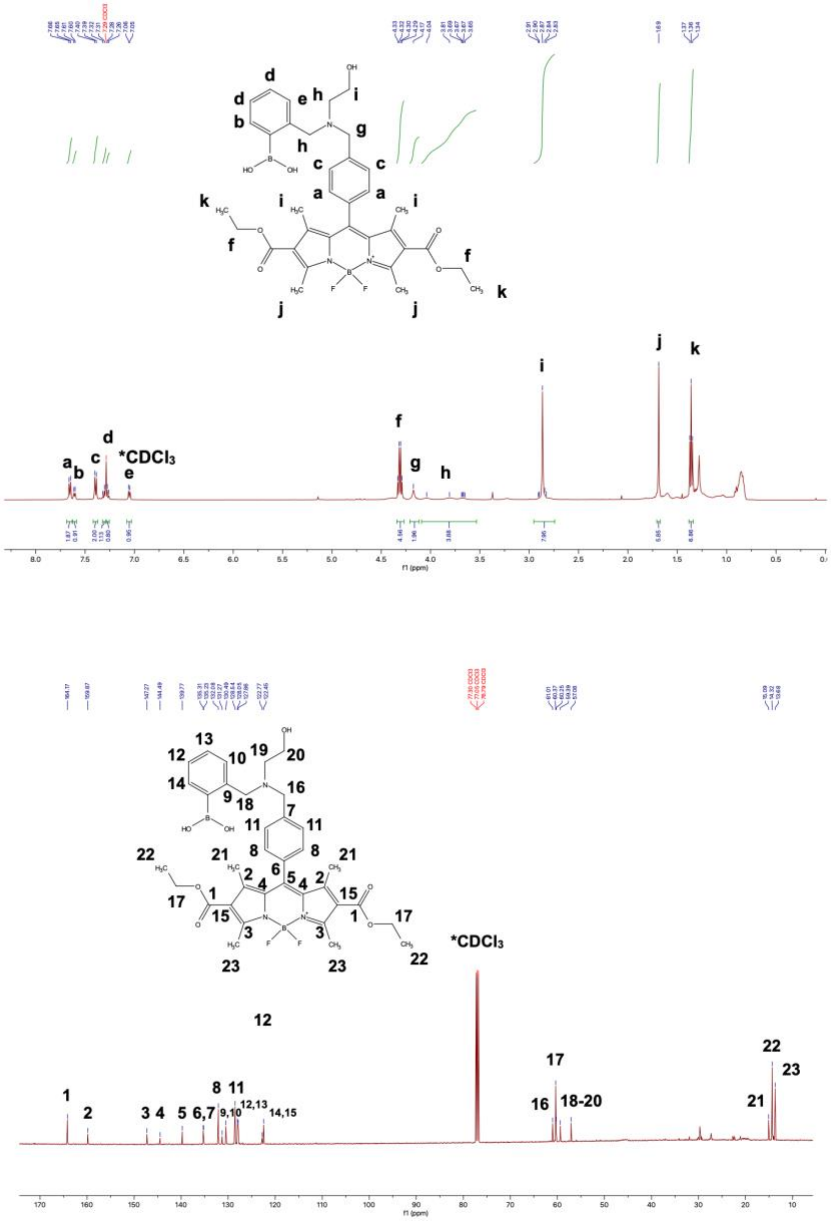


Figure S2.6.9. (2-(((4-(2,8-bis(ethoxycarbonyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-4λ^4,5λ^4-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-10-yl)benzyl)(2-hydroxyethyl)amino)methyl)phenyl)boronic acid, ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.65 (d, *J* = 7.8 Hz, 2H), 7.61 (d, *J* = 6.7 Hz, 1H), 7.39 (d, *J* = 8.1 Hz, 2H), 7.31 (d, *J* = 8.3 Hz, 1H), 7.27 (d, *J* = 7.3 Hz, 1H), 7.05 (d, *J* = 7.1 Hz, 1H), 4.31 (q, *J* = 7.1 Hz, 4H), 4.17 (s, 2H), 4.09 – 3.54 (m, 4H), 2.87 (s, 8H), 1.69 (s, 6H), 1.36 (t, *J* = 7.1 Hz, 7H). ^{13}C NMR

(126 MHz, Chloroform-*d*) δ 164.17, 159.87, 147.27, 144.49, 139.77, 135.31, 135.23, 132.08, 131.27, 130.49, 128.54, 128.05, 127.86, 122.77, 122.45, 61.01, 60.37, 60.25, 59.39, 57.08, 15.09, 14.32, 13.68.

S3. Intrinsic Physical Properties of Base Scaffolds.

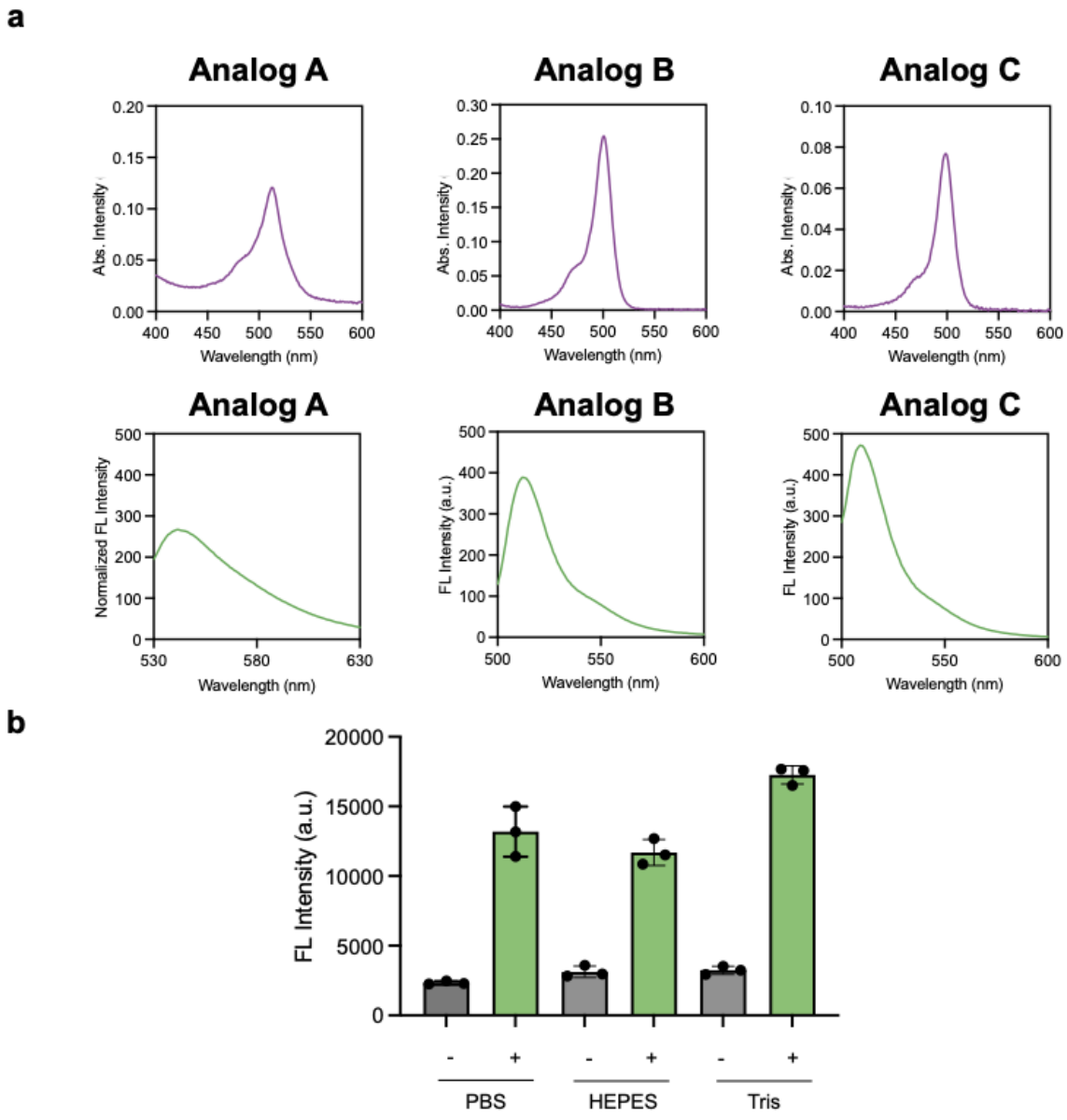


Figure S3. a. Absorbance (top) and emission spectrum (bottom) of 5 μM analogs A (left), B (middle), and C (right) in a 1:1 (v/v) DMSO/PBS mixture containing 0.1 M sorbitol. **b.** Technical triplicate of raw fluorescent readout of 20 μM analog B in different buffer systems containing with and without 0.1 M sorbitol (99% buffer).

S4. Intrinsic Physical Properties of C1–C9

S4.1. UV-Vis Spectra of C1–C9

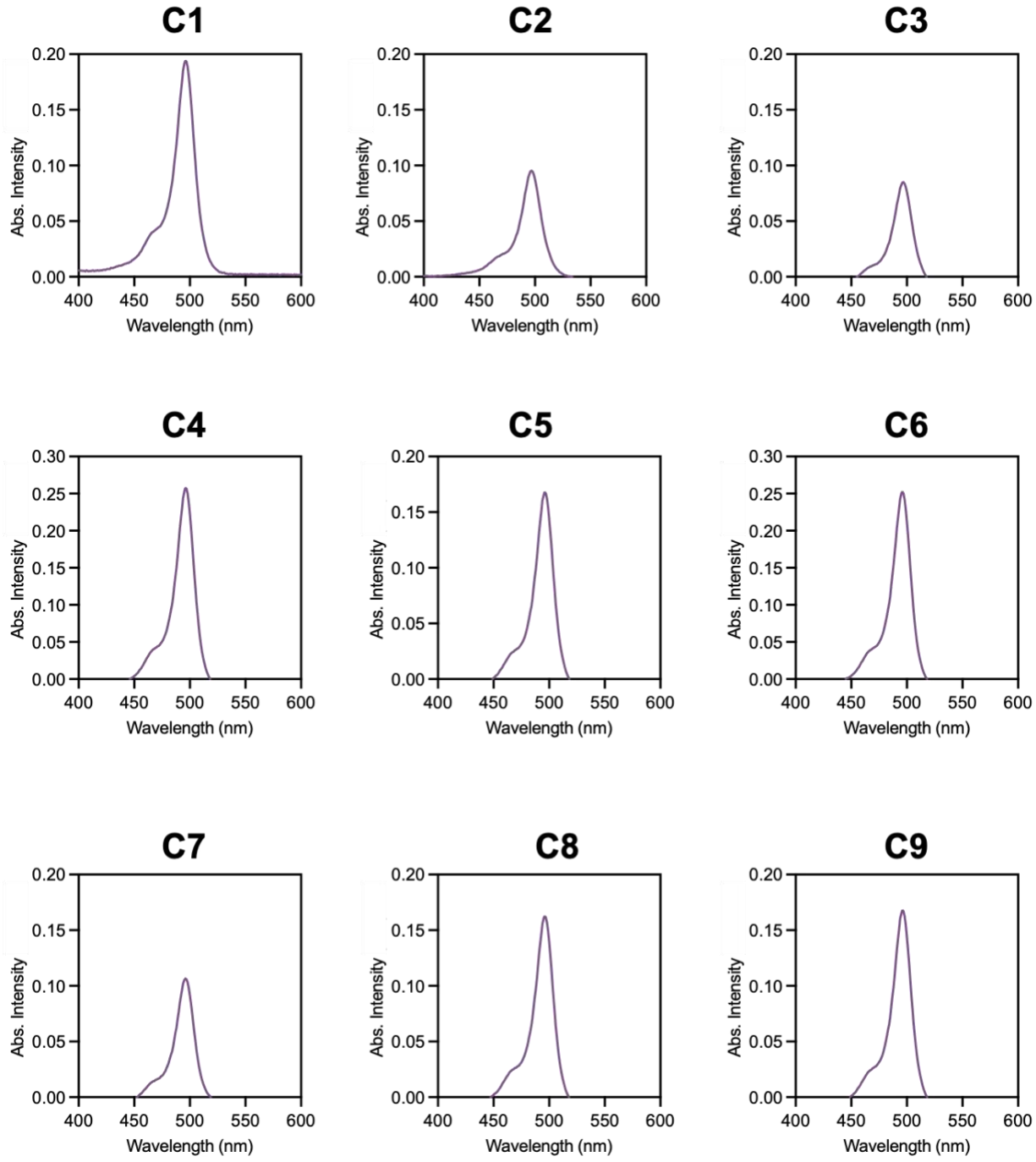


Figure S4.1. Absorbance of analogs C1–C9 (5 μM) in a 1:1 (v/v) DMSO/PBS mixture containing 0.1 M sorbitol.

S4.2. Fluorescence Emission Spectra of C1–C9

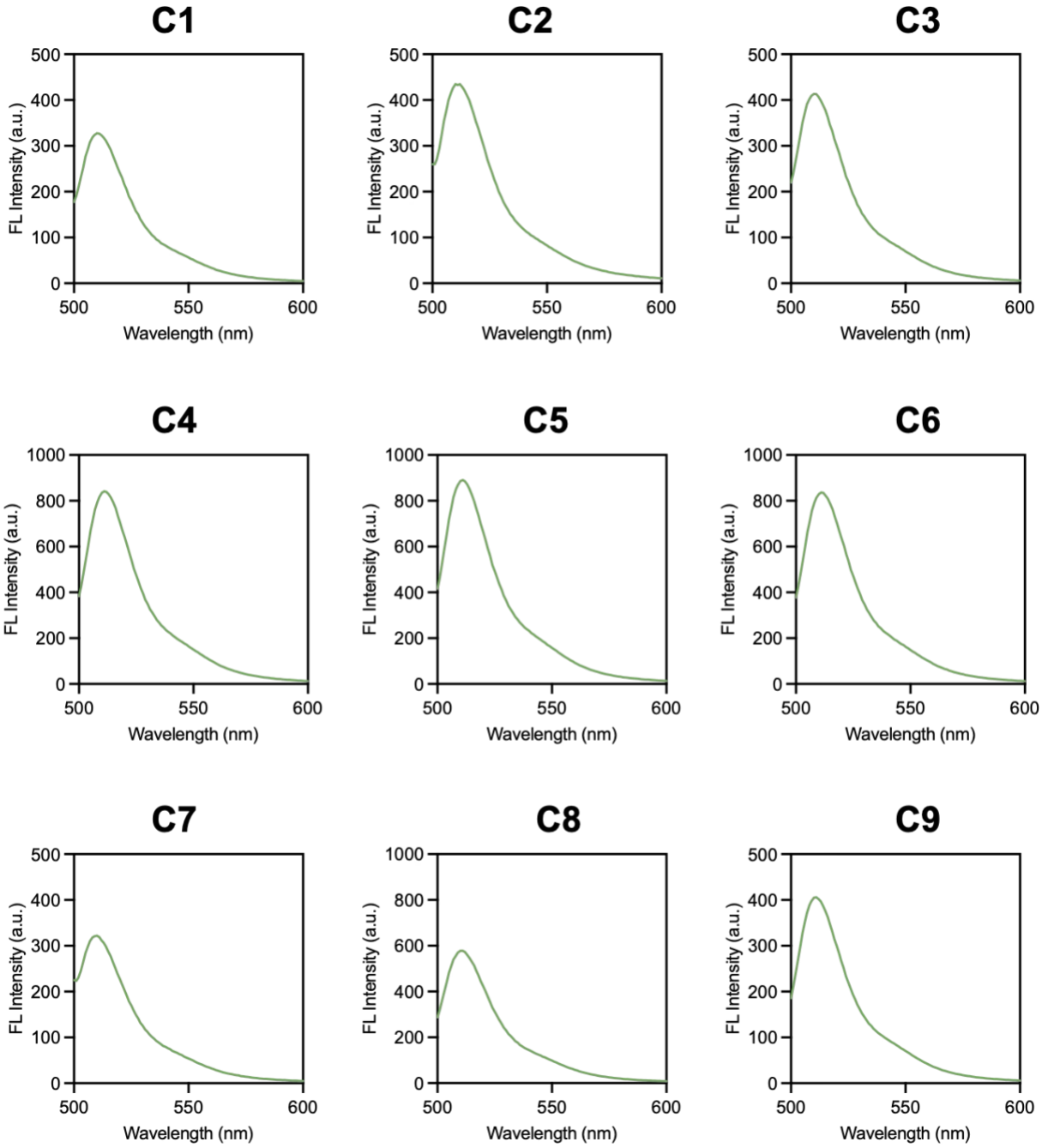


Figure S4.2. Emission spectrum bottom of analogs C1–C9 (5 μM) in a 1:1 (v/v) DMSO/PBS mixture containing 0.1 M sorbitol.

	λ_{abs} (nm)	λ_{em} (nm)	ε ($M^{-1}cm^{-1}$)	Φ
C1	496	510	29930	0.28
C2	497	510	14468	0.33
C3	497	510	21906	0.48
C4	496	511	42052	0.52
C5	496	511	30998	0.74
C6	497	511	48490	0.44
C7	496	510	23703	0.34
C8	496	511	30553	0.48
C9	496	511	42465	0.68

Table S4.1. Photophysical intrinsic property of analogs C1-9 in a 1:1 (v/v) DMSO/PBS mixture containing 0.1 M sorbitol.

S5. Fluorescence Turn-on Behavior of Base Scaffolds in PBS with Various Diol Species

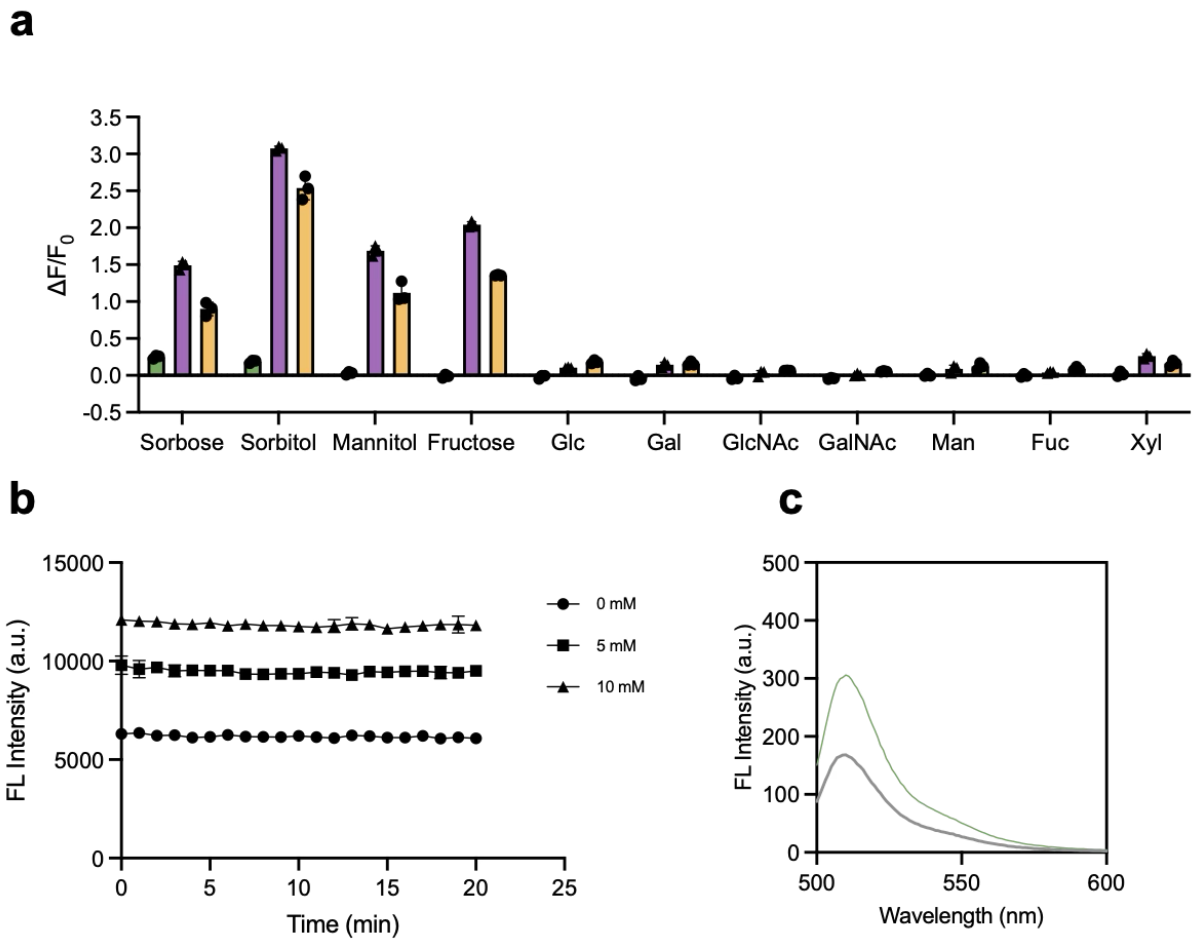


Figure S5. a. $\Delta F/F_0$ of 20 μM analogs A (green), B (purple), and C (orange), in PBS containing 0.1 M of different diol derivatives. All experiments were performed in technical triplicate. b. 20 μM of C9 with various concentrations of sorbitol for a period of 20 minutes. c. fluorescent emission spectrum of compound C9 in 1:1 DMSO/PBS incubated with (green) or without (grey) 0.1 M D-sorbitol.

S6. Isothermal Titration Calorimetry of B with Various Diol Species

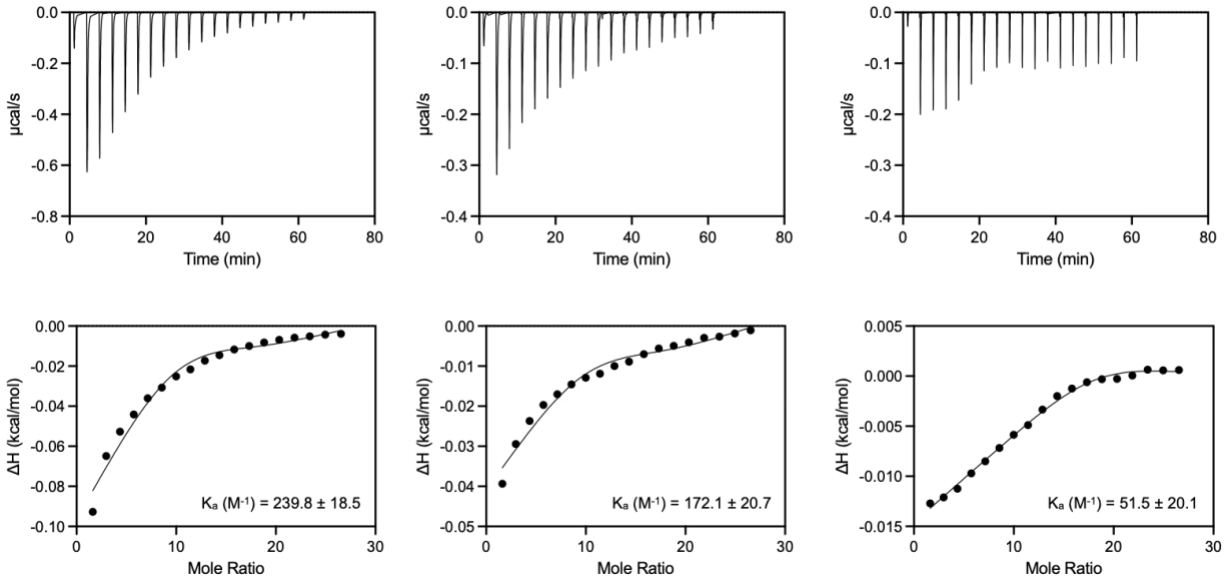


Figure S6. Titration curves (top) for sorbitol (left), fructose (middle), and glucose (right) binding to analog B. Heat plots (bottom) for sorbitol (left), fructose (middle), and glucose (right) binding to analog B in 10% DMSO in Tris buffer (pH 8.1). Titration condition: 0.5 mM of analog B in the cell and 70 mM of diol derivatives in the titration syringe.

S7. Fluorescence Turn-on Behavior of C1–C9 with Various Monosaccharides

S7.1. C1

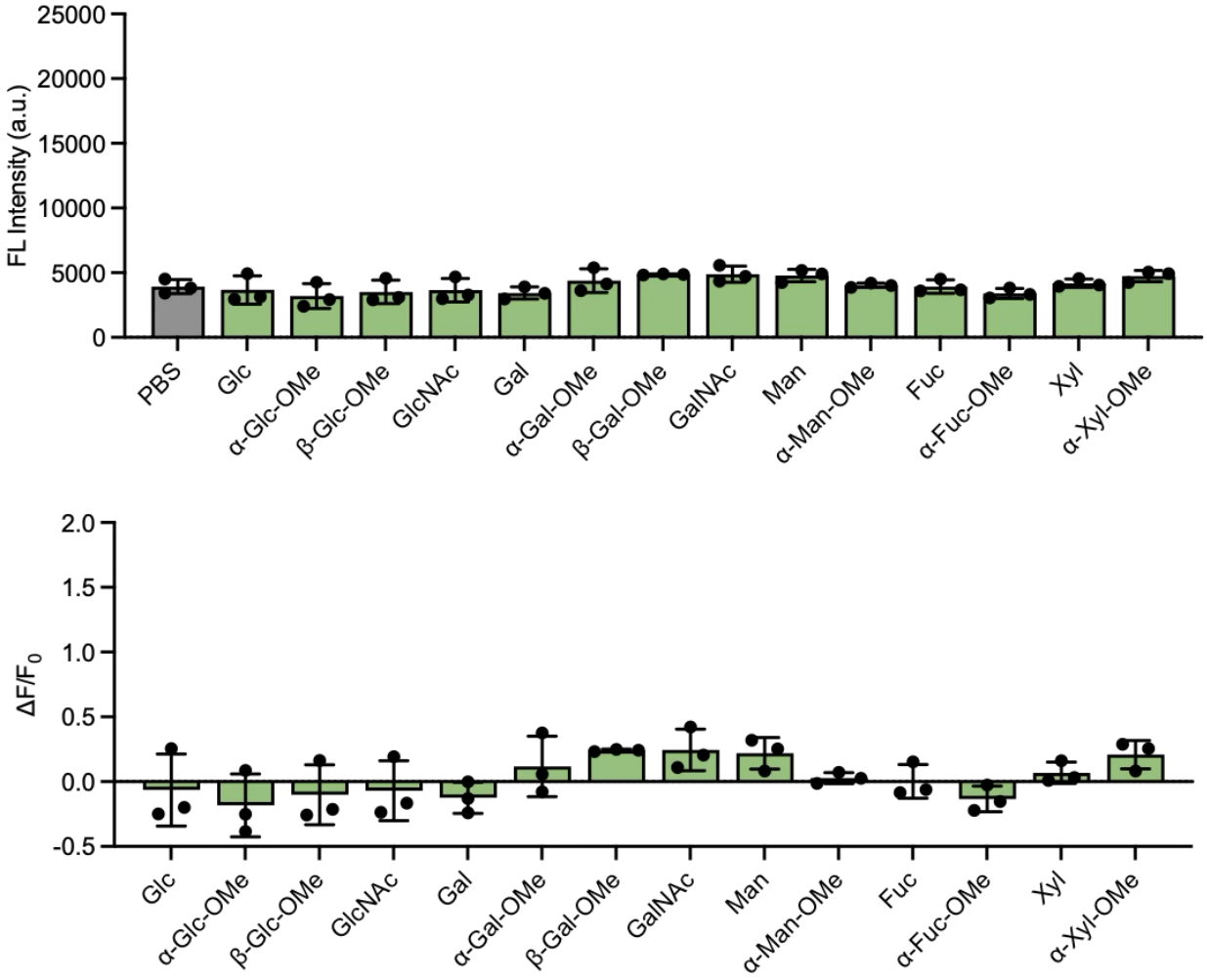


Figure S7.1. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C1. All experiments were performed in technical triplicate.

S7.2. C2

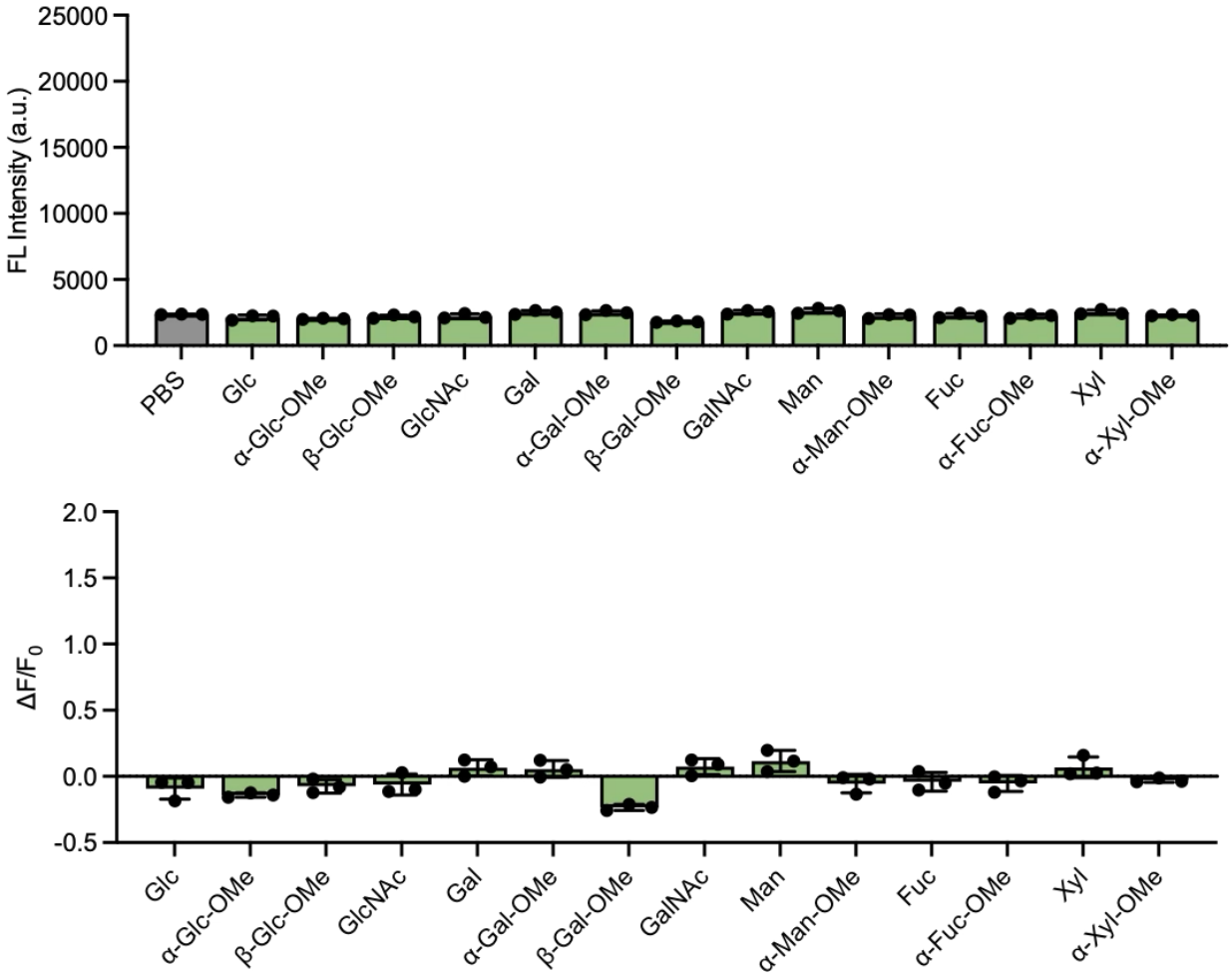


Figure S7.2. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C2. All experiments were performed in technical triplicate.

S7.3. C3

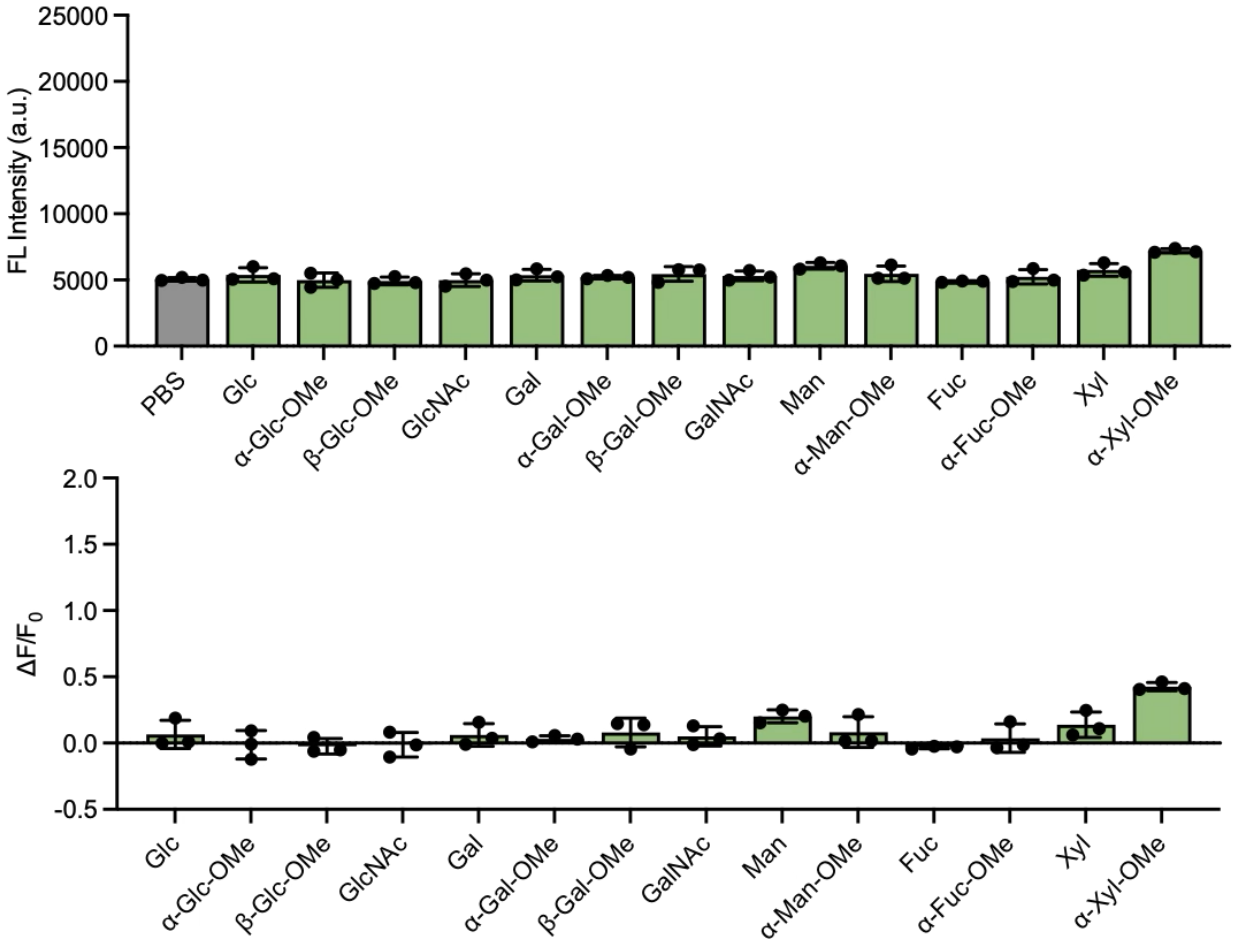


Figure S7.3. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C3. All experiments were performed in technical triplicate.

S7.4. C4

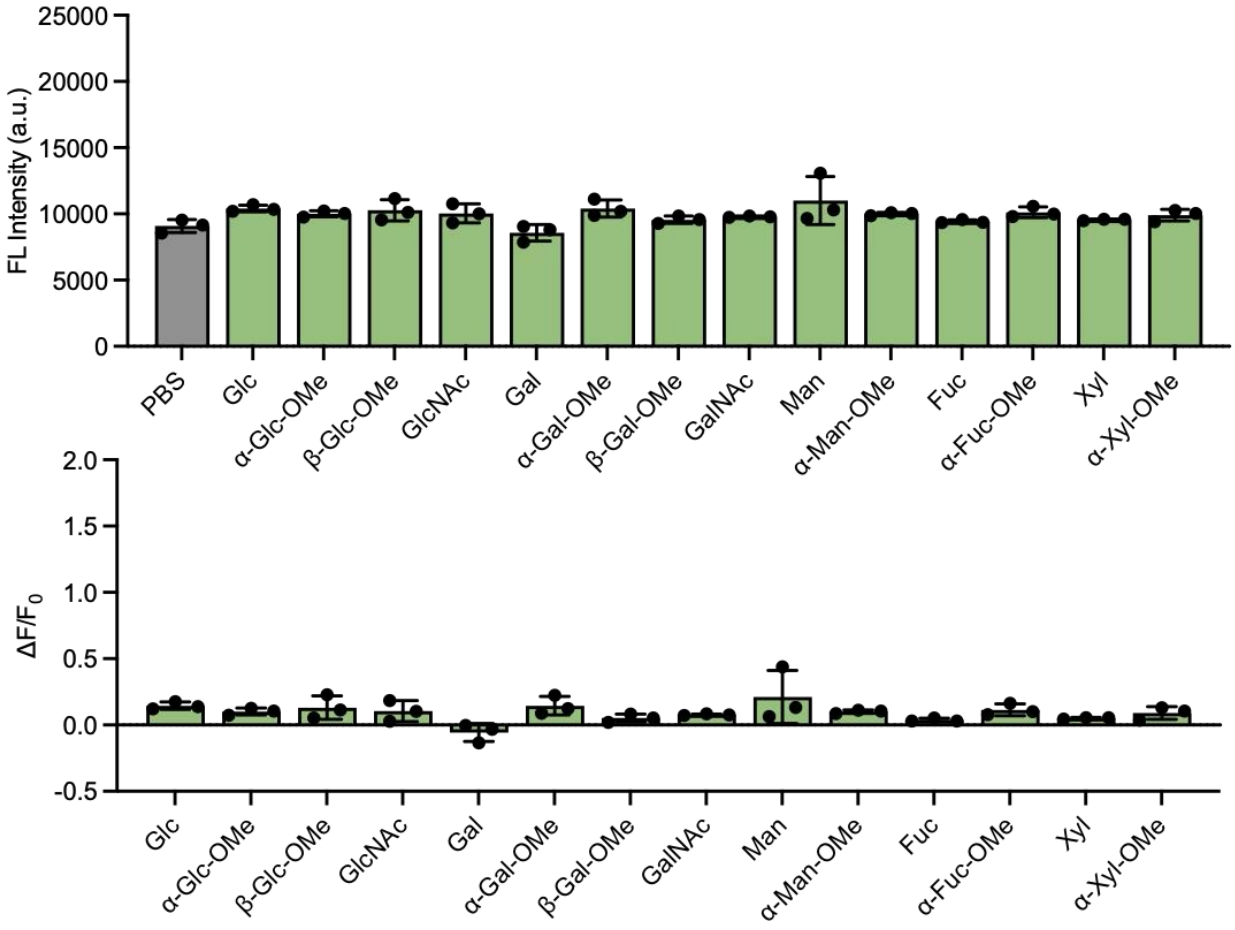


Figure S7.4. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C4. All experiments were performed in technical triplicate.

S7.5. C5

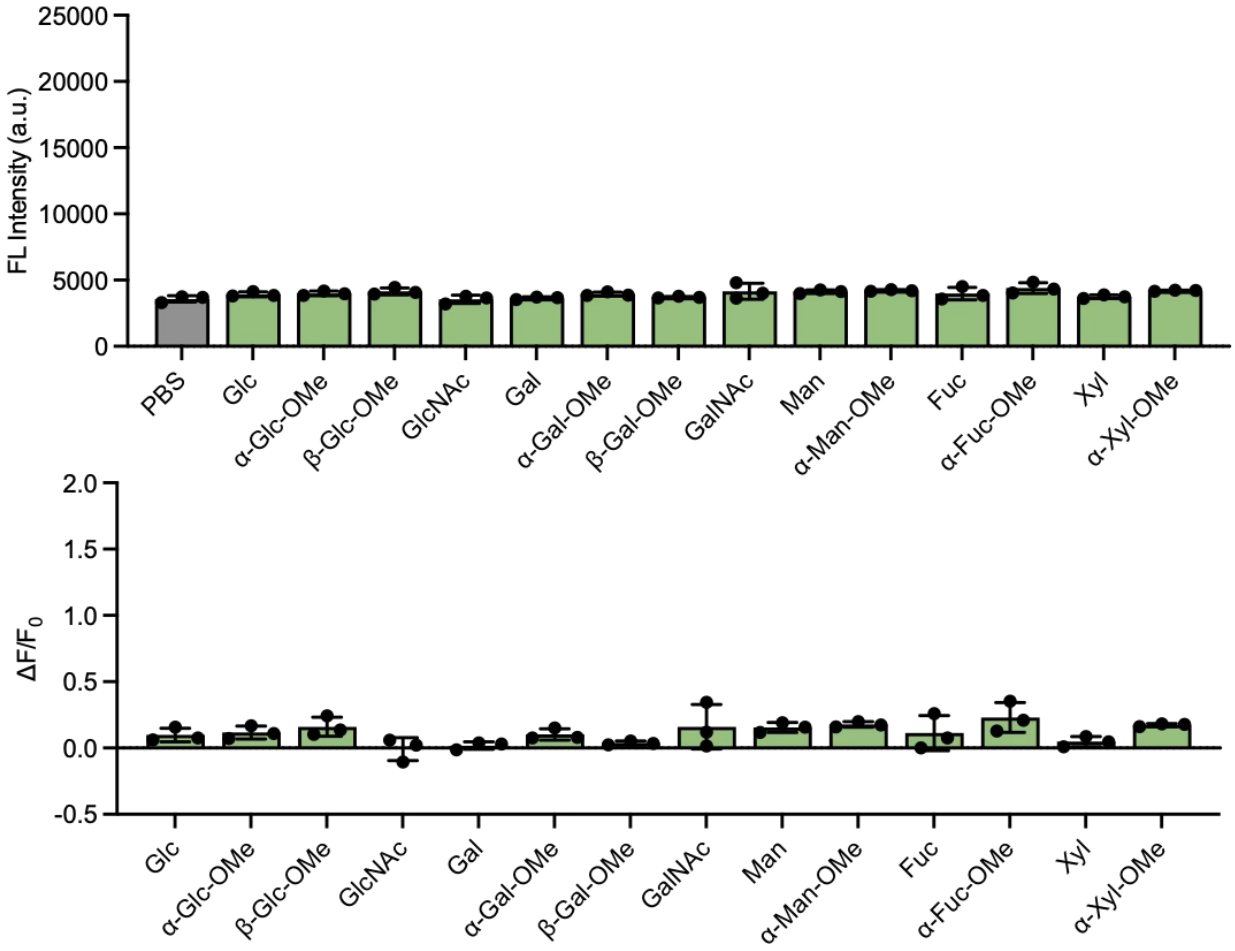


Figure S7.5. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C5. All experiments were performed in technical triplicate.

S7.6. C6

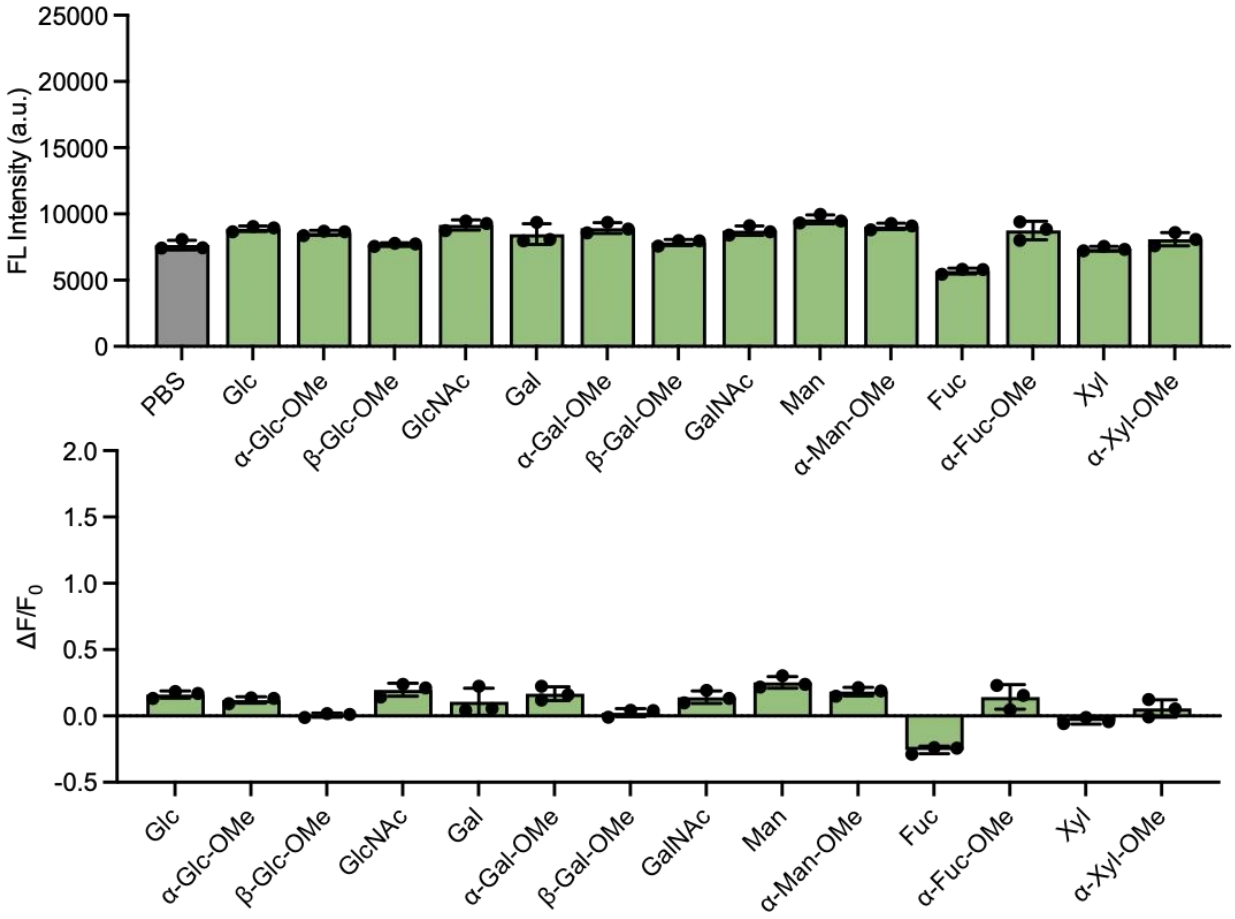


Figure S7.6. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C6. All experiments were performed in technical triplicate.

S7.7. C7

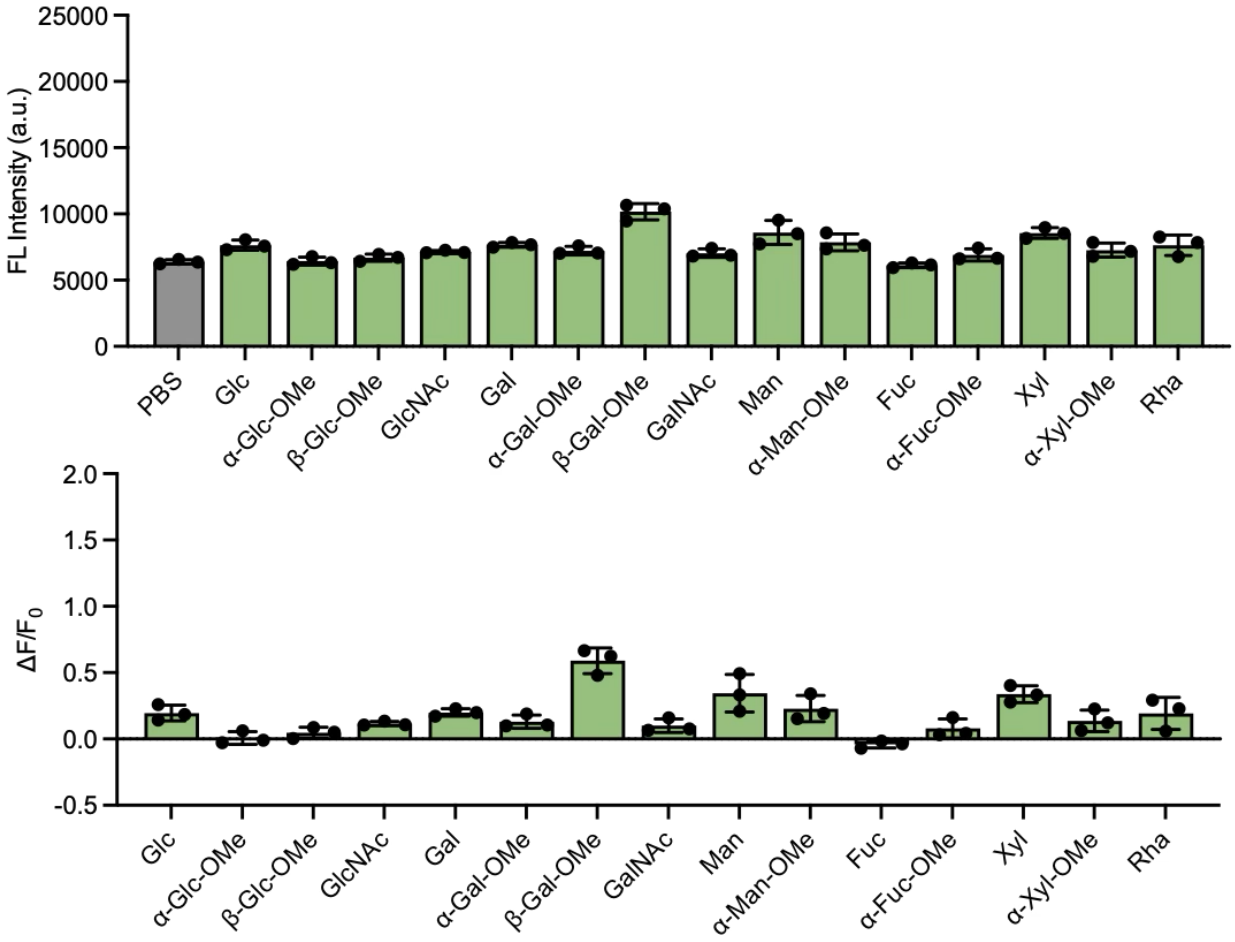


Figure S7.7. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C7. All experiments were performed in technical triplicate.

S7.8. C8

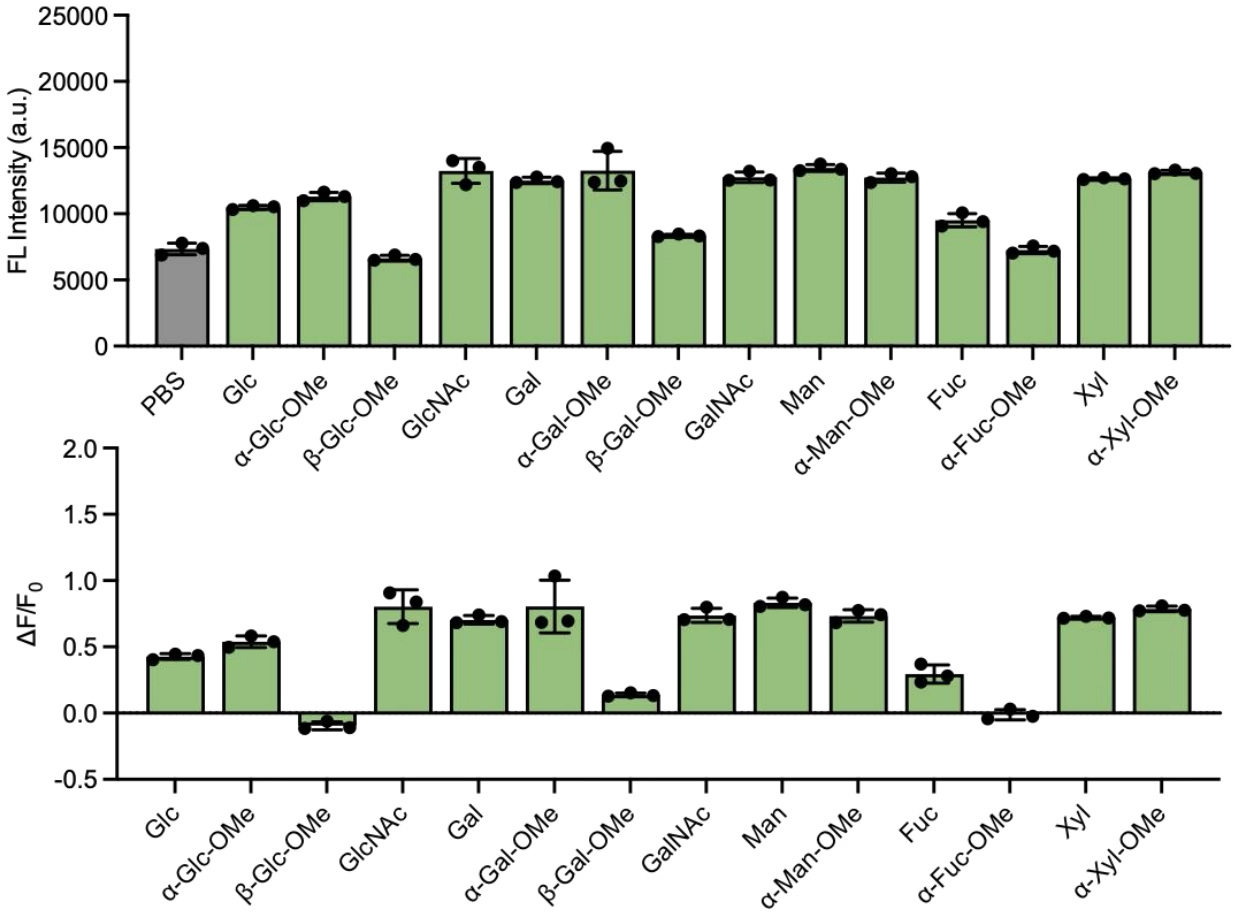


Figure S7.8. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C8. All experiments were performed in technical triplicate.

S7.9. C9

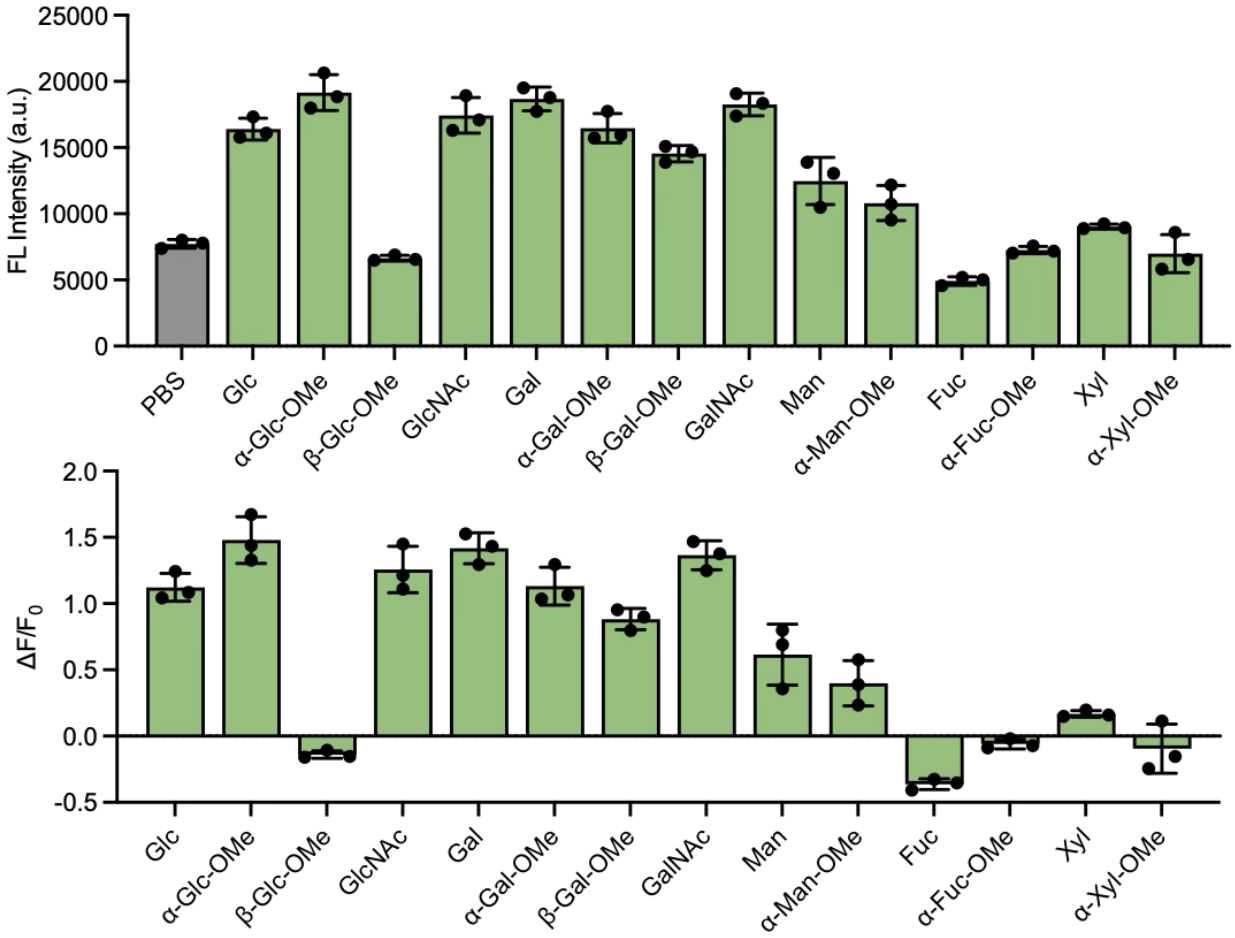


Figure S7.9. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in PBS containing 0.1 M of different diol derivatives with 20 μM of C9. All experiments were performed in technical triplicate.

S7.10. library with N-Acetyl-β-neuraminic acid methyl ester

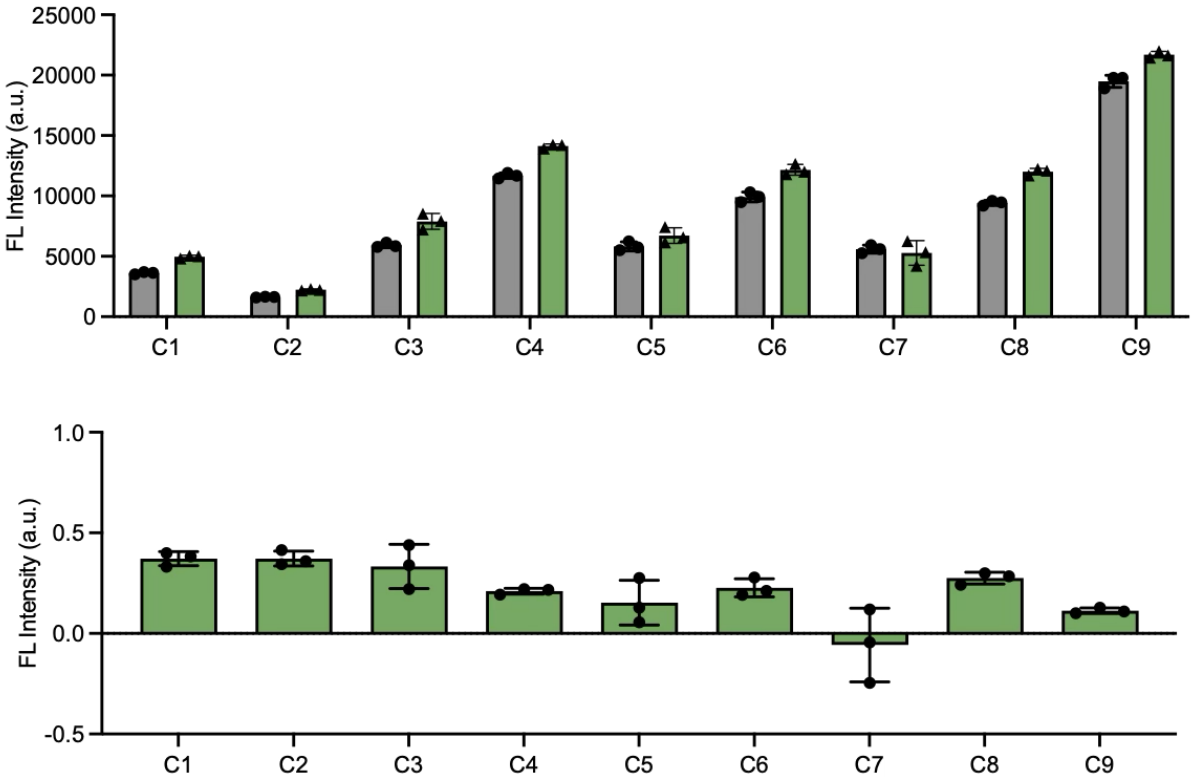


Figure S7.10. Raw fluorescence read-out (top) and $\Delta F/F_0$ (bottom) in 10% methanol in PBS containing 0.25% (w/v) of N-acetyl-β-neuraminic acid methyl ester with 20 μM of C1 to C9. All experiments were performed in technical triplicate.

S8. Fluorescence Turn-on Behavior of C9 with Various Polysaccharides

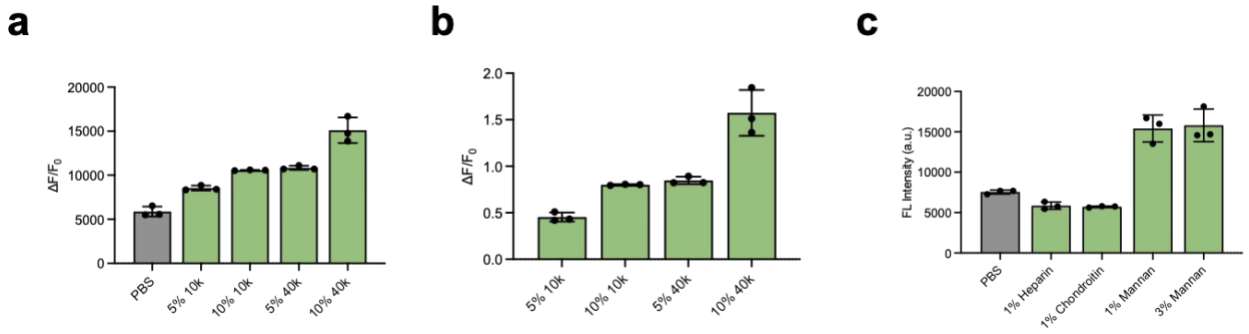


Figure S8. a. Raw fluorescence read-out containing various %(w/v) dextran with 20 μM of C9. **b.** $\Delta F/F_0$ in PBS containing various %(w/v) dextran with 20 μM of C9. **c.** Raw fluorescence read-out containing various %(w/v) of heparin sulfate, chondroitin sulfate, and mannan in PBS with 20 μM of C9. All experiments were performed in technical triplicate.

S9. Fluorescence Turn-on Behavior of C9 in Serum-containing Media

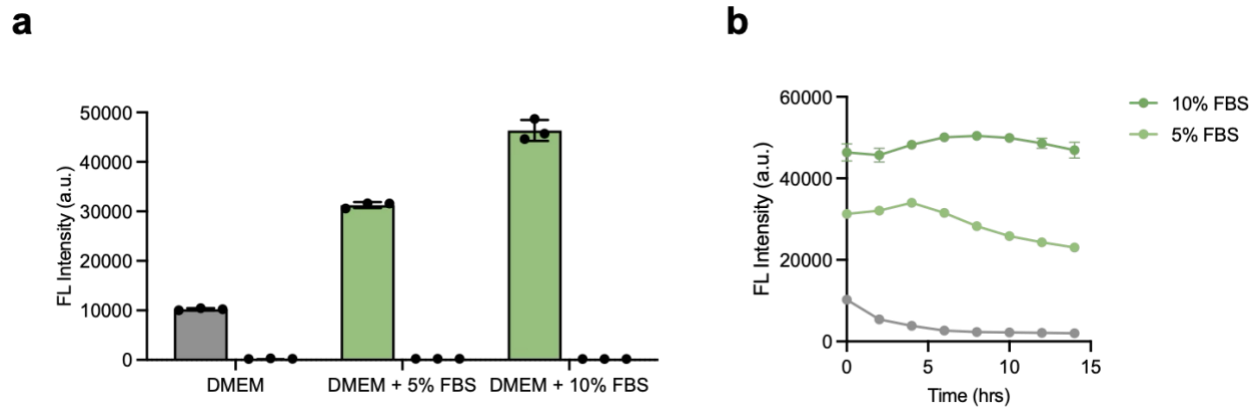


Figure S9. **a.** Raw fluorescence read-out of 20 μM of C (in DMEM, DMEM containing 5% FBS, and DMEM containing 10% FBS (%v/v)). The right bars indicate fluorescent read-out in various media containing no C9. **b.** Raw fluorescence read-out in various serum-containing media over a period of 14 hours.

S10. Gel Imaging of Glycoproteins with C9

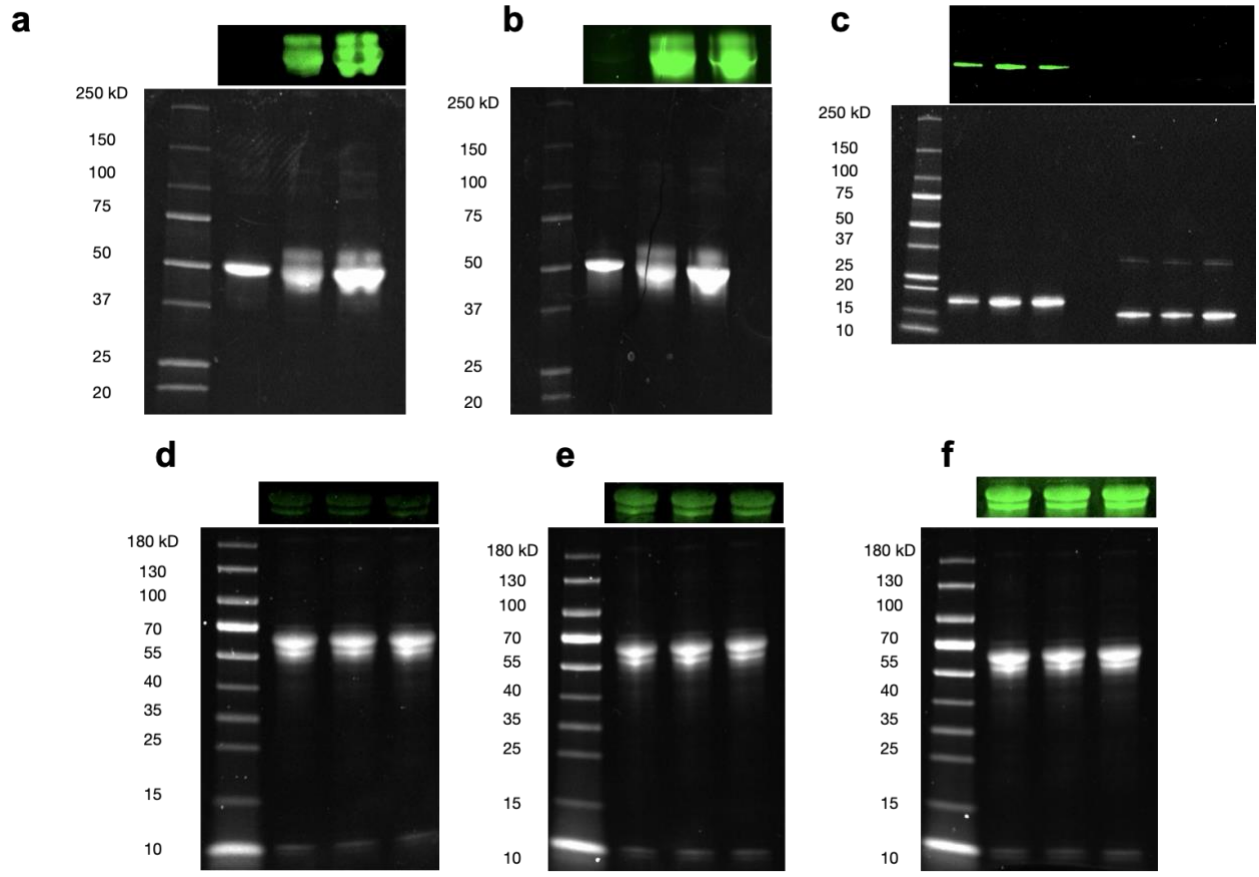


Figure S10. a. Single protein fluorescent and Coomassie gel imaging of BSA (left), fetuin (middle), and BSA + fetuin (right) upon incubation with 10 μM C9. **b.** Single protein fluorescent and Coomassie gel imaging of BSA (left), fetuin (middle), and BSA + fetuin (right) upon incubation with Pro-Q Emerald 300. **c.** Single protein fluorescent and Coomassie gel imaging of RNase B treated with (left three lanes) and without PNGaseF (right three lanes) upon incubation with 10 μM C9. **d.** Single protein fluorescent and Coomassie gel imaging of fetuin upon incubation with 5 μM C9 in triplicate. **e.** Single protein fluorescent and Coomassie gel imaging of fetuin upon incubation with 10 μM C9 in triplicate. **f.** Single protein fluorescent and Coomassie gel imaging of fetuin upon incubation with 15 μM C9 in triplicate.

S11. Reversibility of C9 Staining

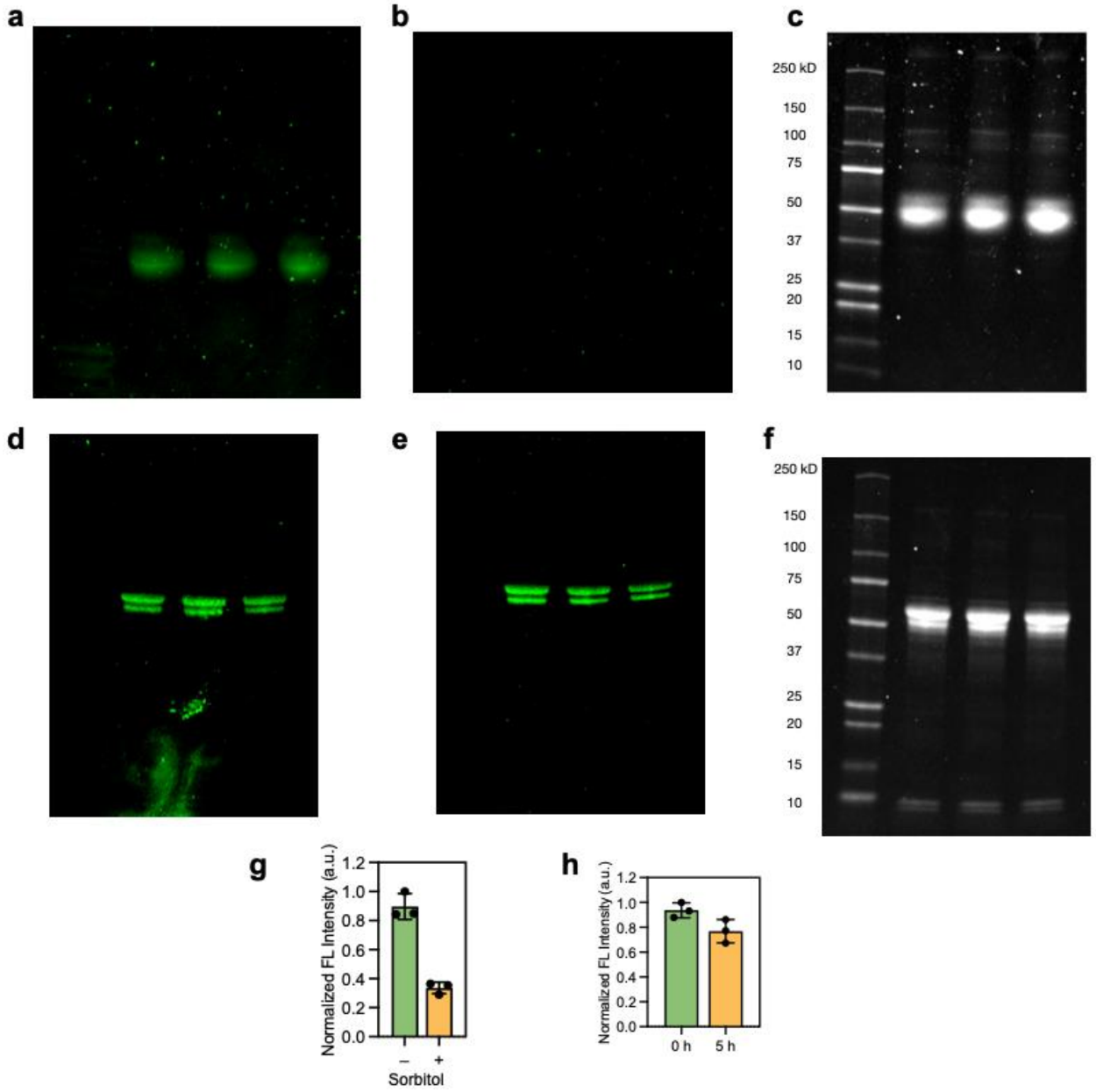


Figure S11. **a.** Single protein fluorescent gel imaging of fetuin in triplicate upon incubation with 10 µM C9. **b.** Single protein fluorescent gel imaging of fetuin in triplicate from after incubation with 1 M sorbitol in PBS for 5 h. **c.** Coomassie gel imaging of the gel shown in (b). **d.** Single protein fluorescent gel imaging of fetuin in triplicate upon incubation with 10 µM C9. **e.** Single protein fluorescent gel imaging of fetuin in triplicate (d) after incubation with 1x PBS for 5 h. **f.** Coomassie gel imaging of the gel shown in (e). **g.**

Quantification of bands on (a) and (b) using ImageJ. **h.** Quantification of bands on (d) and (e) using ImageJ. The error bars represent standard deviation in three technical replicates.

S12. Gel Imaging of Cell Lysates with C9

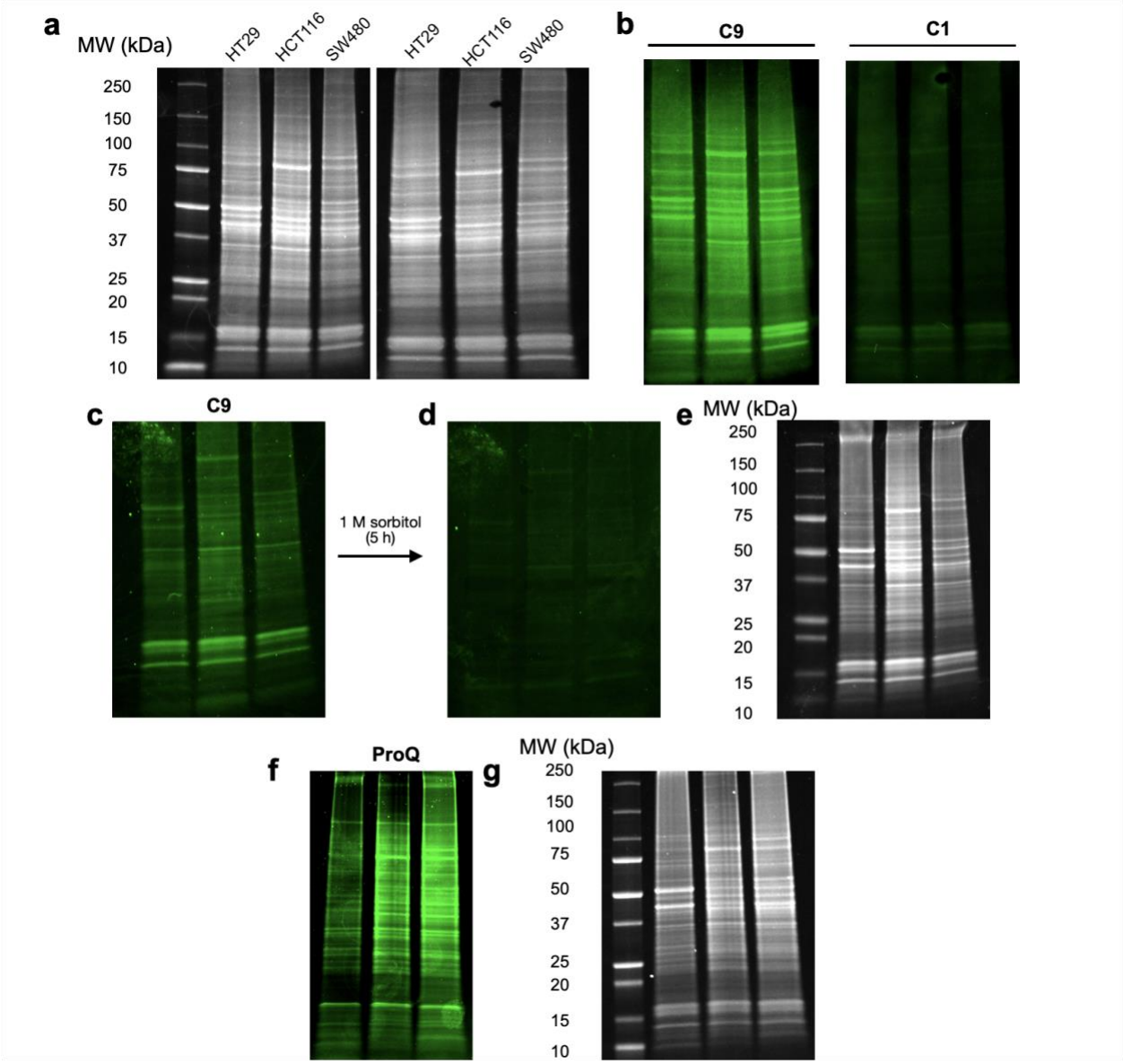


Figure S12. a. Coomassie imaging of HT29, HCT116, and SW480 cell lysate **b.** Fluorescent gel imaging of HT29, HCT116, and SW480 cell lysate upon incubation with 10 μM C9 in PBS for 1 h. **c.** Fluorescent gel imaging of HT29, HCT116, and SW480 cell lysate upon incubation with 10 μM C9 in PBS for 1 h before incubation with 1 M sorbitol in PBS. **d.** Fluorescent gel imaging of HT29, HCT116, and SW480 cell

lysate upon incubation with 10 μM C9 in PBS for 1 h after incubation with 1 M sorbitol in PBS. **e.** Coomassie imaging of HT29, HCT116, and SW480 cell lysate after incubation with 1M sorbitol in PBS. **f.** ProQ 300 staining of HT29, HCT116, and SW480 cell lysate **g.** Coomassie imaging of HT29, HCT116, and SW480 cell lysate after ProQ 300 staining.

S13. Gating Strategy for MFI Quantification of Staining

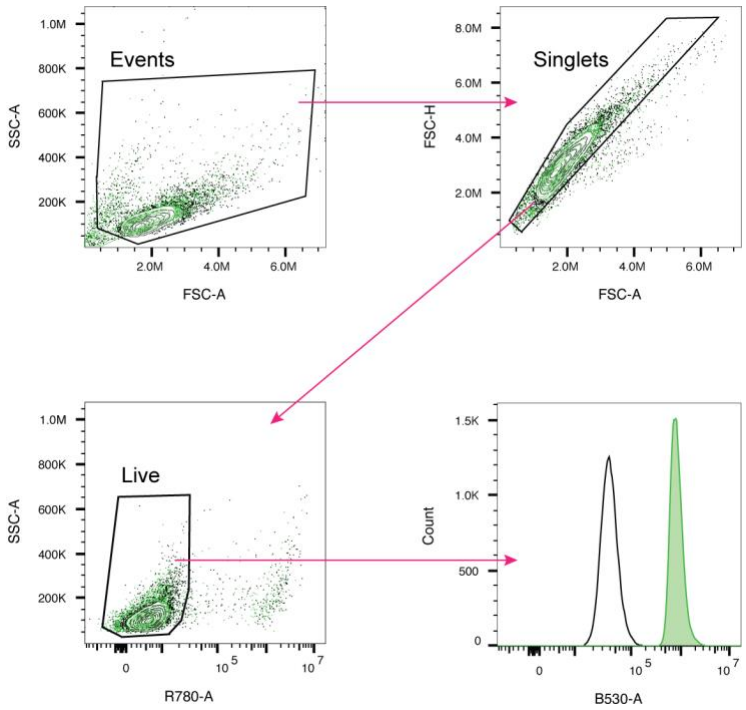


Figure S13. Representative gating strategy for median fluorescence intensity (MFI) quantification of boronic acid probe staining. Events were first gated to remove debris, followed by selection of forward scatter singlets and live cells as distinguished by Zombie NIR staining in the R780-A channel (640 nm excitation, 780/60 filter). Fluorescence of live cells was then measured in the B530-A channel (488 nm excitation, 530/30 filter). Samples shown are stained with 100 nM of **C*** (black) or **C9** (green). SSC-A (side scatter area), FSC-A (forward scatter area).

S14. Flow Cytometry on THP-1 Cells

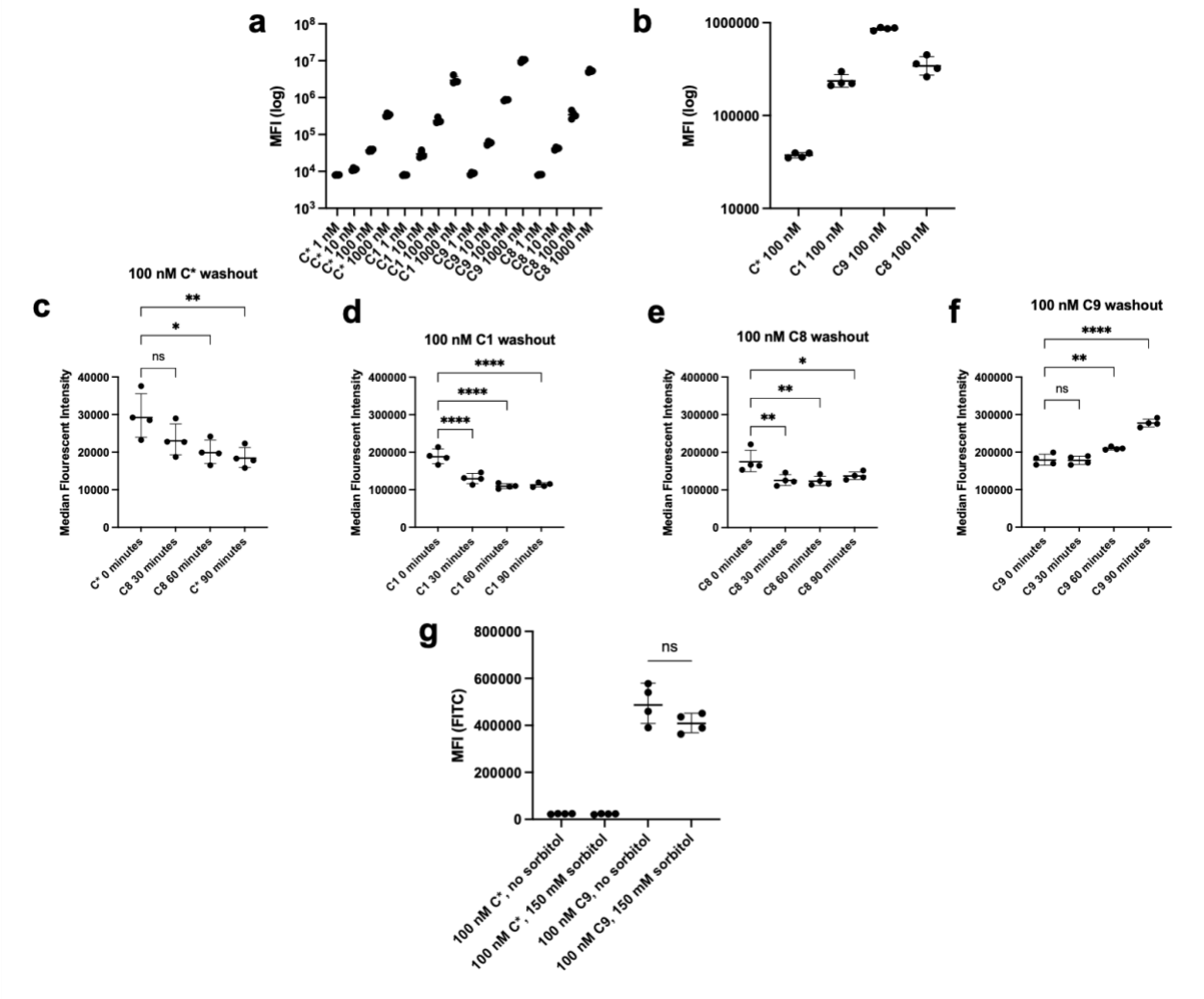


Figure S14. a. median fluorescence intensity (MFI) in response to various concentrations of C*, C1, C8, and C9. **b.** MFI (log) of 100 nM of C*, C1, C8, and C9. **c-f.** washout staining over ninety minutes upon incubation of 100 nM of C*, C1, C8, and C9, MFI was taken at 0, 30, 60, and 90 minutes. **g.** C* or C9-stained cells washed with buffer (no sorbitol) or 150 mM sorbitol. Data were analyzed by a two-way one-sided ANOVA with Dunnett's post-hoc test. ns = P>0.05, *P<0.05, **P<0.01, ***P<0.001, and ****P<0.0001. All error bars represent geometric mean and geometric standard deviations in four technical replicates.

S15. Confocal Imaging of Fixed-HT29 Cells

S15.1. Confocal Imaging of Fixed-HT29 Cells with C*, C1, and C9

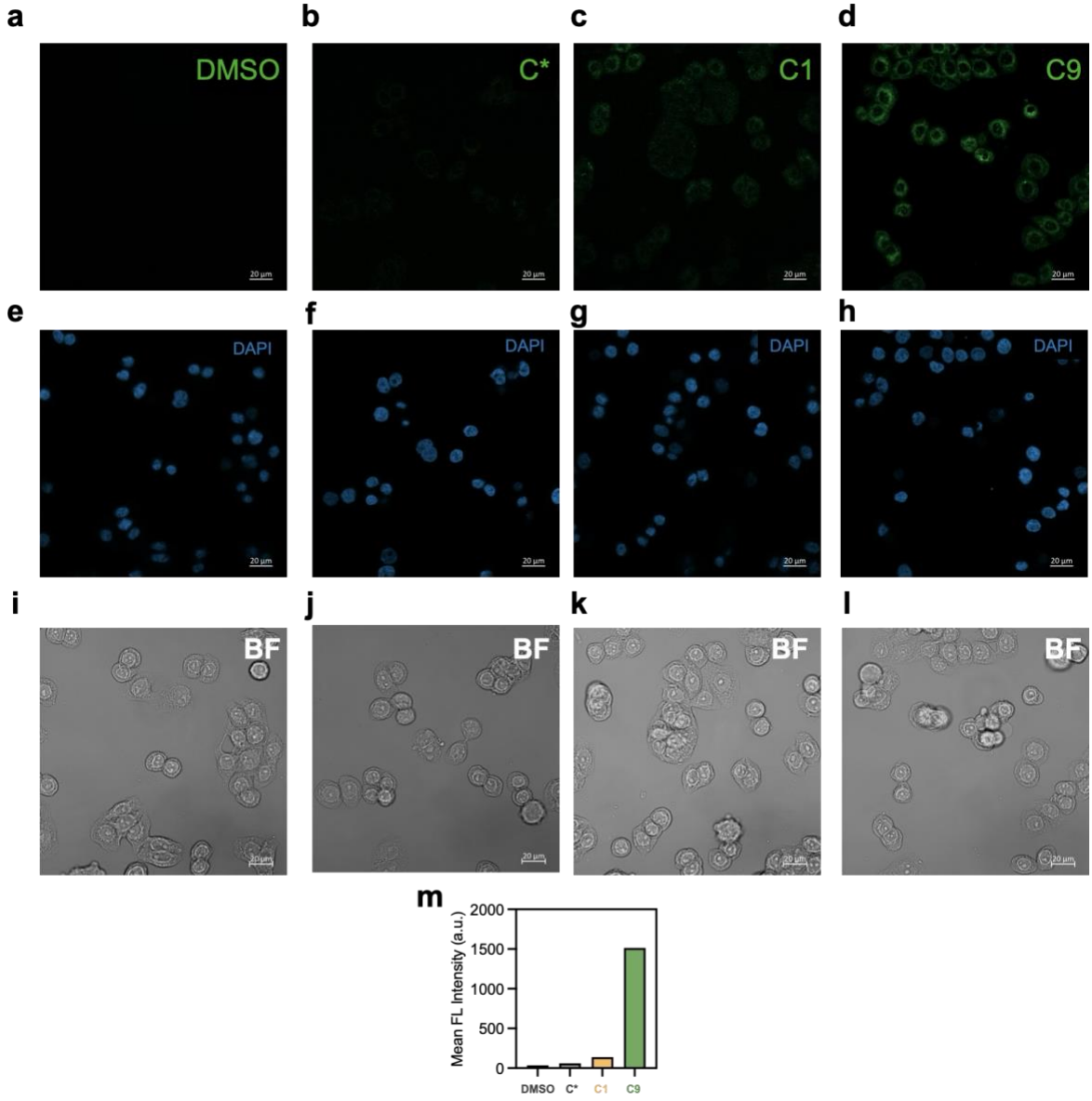


Figure S15.1. a-d. Confocal fluorescent imaging of HT29 cells upon 1 hour incubation with 1% DMSO, 100 nM C* with 1% DMSO, 100 nM C1 with 1% DMSO, or 100 nMC9 with 1% DMSO. An excitation wavelength of 488 nm was used, with emission collected over 508–658 nm. **e-h.** Confocal fluorescent imaging of HT29 cells in (a-d) upon 1 hour incubation with 300 nM DAPI. An excitation wavelength of 405 nm was used, with emission collected over 411–490 nm. **i-l.** Microscopy images of HT29 cells were

collected in brightfield contrast mode using a PMT detector at a detection range of 300–900 nm. **m.** Quantification of mean green fluorescent intensity in images a-d. Scale bar = 20 μm.

S15.2. Confocal Imaging of Fixed-HT29 Cells with or without Sorbitol Wash

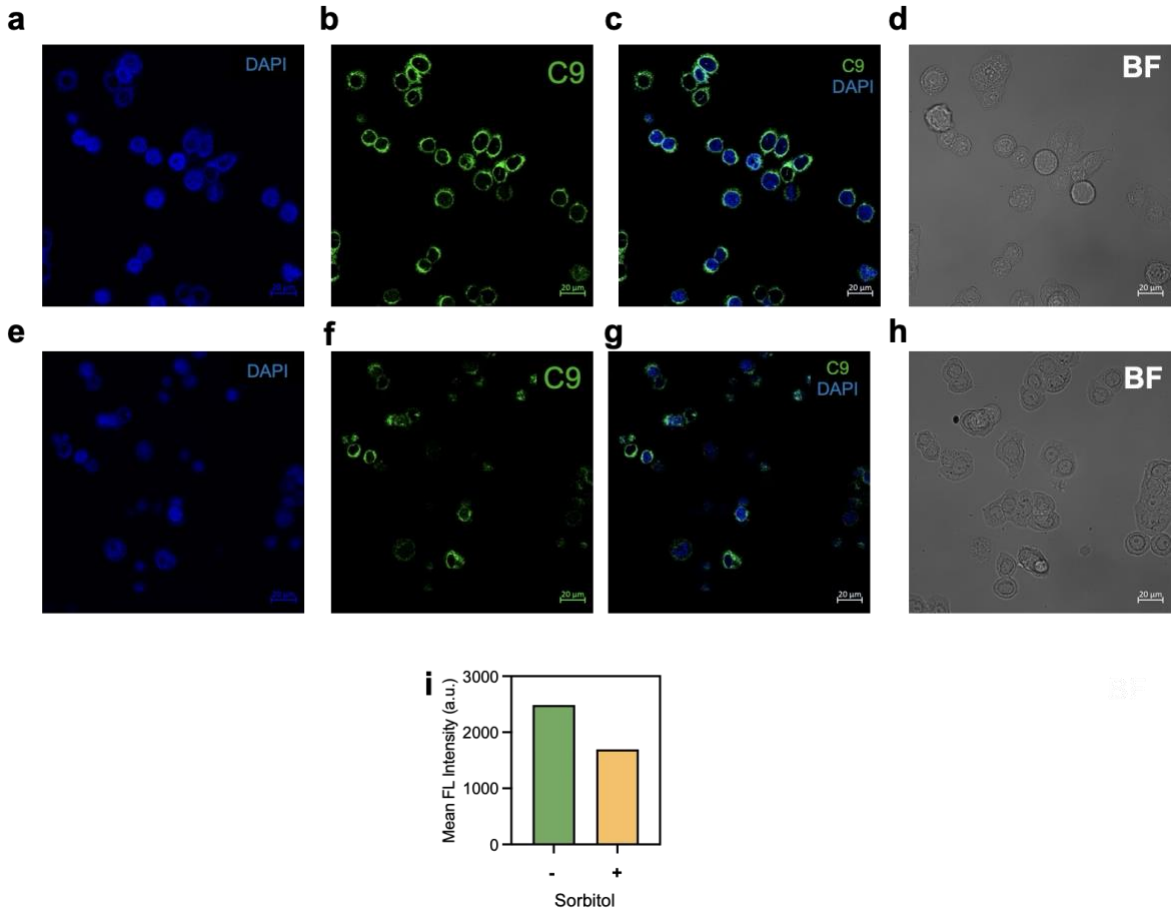


Figure S15.2. a-d. Confocal fluorescent imaging of HT29 cells upon 1 hour incubation with 300 nM of DAPI and 100 nM C9, followed by 3 washes with DPBS. For C9, an excitation wavelength of 488 nm was used, with emission collected over 508–658 nm. For DAPI, an excitation wavelength of 405 nm was used, with emission collected over 411–490 nm. **e-h.** Confocal fluorescent imaging of HT29 cells upon 1 hour incubation with 300 nM of DAPI and 100 nM C9, followed by 3 washes with 200 mM sorbitol in DPBS. For C9, an excitation wavelength of 488 nm was used, with emission collected over 508–658 nm. For DAPI, an excitation wavelength of 405 nm was used, with emission collected over 411–490 nm. **i.** Quantification of mean green fluorescent intensity in images b and f. Scale bar = 20 μm.