

Boronic Esters as Bioorthogonal Probes in Site-Selective Labeling of Proteins

by

Burcin Akgun

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry  
University of Alberta

© Burcin Akgun, 2018

## Abstract

The development of rapid and bioorthogonal chemical reactions has expanded greatly the utility of bioconjugation chemistry in the service of site-selective protein labeling, even allowing molecular imaging in live cells or animals. ‘Click’ chemistry is particularly desirable due to its fast reactivity in an aqueous environment at low concentrations ( $< 100 \mu\text{M}$ ), with high yields and selectivity without any side products. Although a number of click reactions were developed for this purpose, many are associated with drawbacks and limitations that justify the development of alternative systems for both single- or dual-labeling applications. To address these challenges, this thesis presents novel bioorthogonal tools based on a ‘click’ boronic ester formation, which is attractive due to the synthetic accessibility of boronic acids and diols, their low toxicity and fast kinetics in catalyst-free conditions.

Chapter 1 summarizes progress made in the use of boronic acids in bioorthogonal chemistry to enable site-selective labeling of proteins and compares these bioorthogonal reactions with the most commonly applied bioorthogonal reactions. Chapter 2 describes a new ‘click’ bioorthogonal reaction system, which was devised to enable the fast ligation of conjugatable derivatives of the rigid cyclic diol (nopoldiol) and a carefully optimized boronic acid partner (2-methyl-5-carboxymethyl-phenylboronic acid) with a rate constant of  $8 \text{ M}^{-1}\text{s}^{-1}$  as measured by NMR spectroscopy. Using NMR and fluorescence spectroscopy studies, the resulting boronates were found to form reversibly within minutes at low micromolar concentration, providing submicromolar equilibrium dissociation constants. Efficient protein conjugation under physiological conditions was successfully demonstrated with model proteins,

thioredoxin (Trx) and albumin, and characterized using mass spectrometry and gel electrophoresis.

Boronic ester formation is a fast dehydrative process; however, it is intrinsically reversible in an aqueous medium. Also, fluorogenic reactions have been invaluable tools in bioorthogonal chemistry since these reactions allow the visualization of biomolecules without the need to remove the excess unreacted probes. In Chapter 3, early efforts toward designing an irreversible and fluorogenic boronic ester formation are summarized. Even though attempts towards a fluorogenic boronic ester system were unsuccessful, an irreversible synergic system based on two bifunctional reagents, a thiosemicarbazide-functionalized nopoldiol and an 2-acetylarylboronic acid was developed. Both reagents were shown to be chemically stable and non-toxic to HEK293T cells at concentrations as high as 50  $\mu\text{M}$ . The resulting boronate/thiosemicarbazone adduct is a medium sized ring that forms rapidly and irreversibly without any catalyst at low  $\mu\text{M}$  concentrations, in neutral buffer, with a rate constant of  $9 \text{ M}^{-1}\text{s}^{-1}$  as measured by NMR spectroscopy. Control experiments in the presence of competing boronic acids showed no cross-over side-products and confirmed the stability and lack of reversibility of the boronate/thiosemicarbazone conjugates. Moreover, formation of the conjugates is not affected by the presence of biological diols like fructose, glucose and catechol, and the thiosemicarbazide-functionalized nopoldiol is inert to aldehyde electrophiles of the sort found on protein-bound glyoxylyl units. The suitability of this system in the cell-surface labeling of live cells was demonstrated using a SNAP-tag approach to install the boronic acid reagent onto the extracellular domain of Beta-2 adrenergic receptor in HEK293T cells, followed by incubation with the optimal thiosemicarbazide-functionalized nopoldiol reagent labeled with a fluorescein dye. Successful visualization by fluorescence microscopy was possible with a reagent

concentration as low as 10  $\mu\text{M}$ , thus confirming the potential of this system in biological applications.

In Chapter 4, efforts to discover a reactive peptide tag toward 2-acetylarylboronic acid *via* both imine and boronate formation, are summarized. Replacing one of the bioorthogonal handles could be achieved by expressing and installing it as a short, non-invasive peptide tag on a protein of interest (POI). In order to discover the reactive peptide, the phage display platform with a serine-terminated library of over  $10^8$  heptapeptides was employed in collaboration with the Laboratory of Prof. Ratmir Derda. Some peptide hits were selected and synthesized; however, using ESI-MS and NMR analyses, those conjugation studies of peptides with 2-acetylbenzeneboronic acid confirmed that the desired ligation was not observed.

## Preface

Chapter 1 of this thesis has been submitted to *Angewandte Chemie International Edition* for publication as Akgun, B.; Hall, D. G. “Boronic Acids as Bioorthogonal Probes in Site-Selective Labeling of Proteins”. Prof. D. G. Hall as the supervisory author and I were responsible for conception and manuscript composition. **Scheme 1-7** is adapted with permission from Stress, C. J.; Schmidt, P. J.; Gillingham, D. G. “Comparison of Boron-assisted Oxime and Hydrazone Formations Leads to the Discovery of a Fluorogenic Variant,” *Org. Biomol. Chem.* **2016**, *14*, 5529–5533. Copyright © 2016 Royal Society of Chemistry. **Scheme 1-11** is adapted with permission from Andersen, K. A.; Smith, T. P.; Lomax, J. E.; Raines, R. T. “Boronic Acid for the Traceless Delivery of Proteins Into Cells,” *ACS Chem. Biol.* **2016**, *11*, 319–323. Copyright © 2016 American Chemical Society. **Scheme 1-12** is adapted with permission from Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. “Selective Recognition of Protein Tetraserine Motifs With a Cell-permeable, Pro-fluorescent Bis-boronic acid,” *J. Am. Chem. Soc.* **2009**, *131*, 438–439. Copyright © 2009 American Chemical Society. **Figure 4-4** is adapted with permission from the research presentation of Prof. Ratmir Derda.

Chapter 2 of this thesis has been published as Akgun, B.; Hall, D. G. “Fast and Tight Boronate Formation for ‘Click’ Bioorthogonal Conjugation,” *Angew. Chem. Int. Ed.* **2016**, *55*, 3909–3913. As the sole experimentalist, I was responsible for the system optimization, synthesis of the reagents, design of kinetic and stability studies, design of the fluorescence quenching experiment, compatibility of the system with biological polyols, bioconjugation studies on model proteins, albumin (BSA) and thioredoxin (Trx). Also, I wrote the manuscript with assistance from Prof. D. G. Hall, who was the supervisory author and was involved with conception and project creation.

Chapter 3 of this thesis has been published as Akgun, B.; Li, C.; Hao, Y.; Lambkin, G.; Derda, R.; Hall, D. G. “Synergic ‘Click’ Boronate/Thiosemicarbazone System for Fast and Irreversible Bioorthogonal Conjugation in Live Cells,” *J. Am. Chem. Soc.* **2017**, *139*, 14285–14291. I was responsible for the system optimization, synthesis of the reagents, design of kinetic and stability studies, compatibility of the system with biological polyols and a model aldehyde. C. Li was responsible with live cell imaging under fluorescence microscope. Y. Hao was partly

responsible with the design of live cell labeling protocol. G. Lambkin was partly responsible for the design of live cell imaging, taking care of live cells and plasmid amplification. I wrote the manuscript with assistance from Prof. D. G. Hall. Prof. D. G. Hall, Prof. R. Derda as the supervisory authors and I were involved with conception, manuscript composition and project creation.

Chapter 4 of this thesis is based on non-published work initiated by Prof. D. G. Hall and Prof. R. Derda. I was responsible for the synthesis of biotin derivatized reagents, selection protocol, synthesis of peptides and conjugation studies. From the Derda Laboratory, a graduate student, V. T. Guzman, guided me in the process of selection protocol and synthesis of peptides, and a post-doctoral fellow, N. Bennet, designed and provided the blocking phage.

## Acknowledgment

I would like to start by thanking my family for always being open and supportive of my decisions. My mother, my father, and my sister express limitless love to me under all circumstances, and I feel the power of their presence even from miles away.

I would like to express my deepest gratitude to Prof. Dennis G. Hall for making me part of his group. His unwavering enthusiasm for organic chemistry, his willingness to share all his knowledge clearly, and his wise presence has always been inspiring. His patience and trust in my work have created a marvelous space where I have been motivated to be more creative, to work harder, to learn more, and to trust my own scientific and professional judgment.

I am also thankful to Prof. Ratmir Derda and all the Derda Laboratory members for allowing and helping me to use their great technology. Moreover, I am very grateful for his suggestions and contributions to my work.

I would like to thank my Supervisory Committee and Ph. D. Examination Committee members: Prof. Ratmir Derda, Prof. Michael J. Serpe, Prof. Frederick West, Prof. Rylan Lundgren, and Prof. Jianmin Gao. I am grateful for the comments, suggestions, and the guidance that they have provided me over the years.

I wish to acknowledge the administrative and research services at the University of Alberta. Specifically, I am thankful for the help provided by Lynne Lechelt, Anita Weiler, Ryan Lewis, Laura Pham, Scott Stelck, Bonnie Gover, Lin Ferguson, Matthew Kingston, and Bernie Hippel. In regards to the Scientific Services Staff, I am particularly thankful to Wayne Moffat (spectroscopy), Mark Miskolzie (NMR), Nupur Dabral (NMR), Ed Fu (HPLC), Gareth Lambkin (biology), Béla Reiz (Mass), and Dr. Angelina Morales-Izquierdo (Mass) for their expertise in providing advice and assistance to my scientific research. Furthermore, I am thankful to other staff members of the NMR, Mass Spectrometry, and Analytical labs, as well as the machine shop, the storeroom, and the glass shop.

I am also thankful to all my lab mates, past and present. I am very grateful for having the pleasure to work with such talented and knowledgeable individuals who have helped and guided me during my study. I am specifically thankful to Dr. Sylvain Bernard, Dr. Sergey Vshyvenko,

Negar Mokhtarihaj, Dr. Michele Boghi, Helen Clement and Xiangyu Li for their friendship and support. Most of all, I am grateful to Dr. Taras Rybak, who became my closest friend in the department. He was very supportive, honest, patient, and kind with me. I will definitely miss our Wednesday wings nights, followed by salsa dancing. I am thankful to those who have provided advice while I was writing my dissertation: Dr. Vanessa Incani, Helen Clement, Dr. Michele Boghi, Dr. Taras Rybak, Dr. Deniz Meneksedag, Vivian T. Guzman, and Dr. Anna Jordan. I would also like to thank Dr. Hayley Wan for being a supportive lab coordinator.

Definitely, this journey would not be easier without friends. I would like to express my deepest gratitude to all my friends for all their support and advice. In my first year, Jong Myoung became my closest friend in the department, and classes would not be enjoyable without him. My Turkish friends, Deniz Meneksedag, Fatih Erol and Sefa Burak Ozguluk, helped me to feel close to home. I have also had joyful and great experiences with my group of friends, “Aliens”, which include Nada Djokic, Dr. Yazeed Nabeeh Al Hamarneh, Dr. Maica and Dr. Leonidas Perez, Dr. Elena and Dr. Nikolaus Klamerth, Dr. Cristina Hernandez, Sefa Burak Ozguluk, Daniel Hernandez and Dr. Vanessa Incani.

Vivian T. Guzman with her patience and unwavering friendship has been very inspiring, and she became one of my closest friends in the last year of my study. I am grateful for her pure and openhearted goodwill. Also, I am thankful for all her assistance, support and guidance to my research and mental health. Laurie Bernes, my Karma team partner, made my Saturday mornings fun and special while we were cleaning the studio (Sattva School of Yoga). A cleaning job could not be easier and so much enjoyable without her. Also, I am grateful to our teaching experience (yoga) in the department of chemistry. Moreover, I would like to thank Sattva School of Yoga for providing an incredible space where I have had the pleasure to relax and boost up my energy.

I wish to express my appreciation for the support from Alberta Innovates Technology Futures, and I would like to acknowledge awards given by Gilead Alberta and the University of Alberta. With these recognitions, the journey became more pleasant and motivating.

Last but not least, I would like to thank my best friend and husband, Sefa Burak Ozguluk, for believing and supporting me at all times with unconditional love. I am truly grateful for his positive vibes and presence, which often become an inspiration for me to transcend my difficult moments.

## Table of Contents

<b>Chapter 1. Boronic Acids as Bioorthogonal Probes in Site-Selective Labeling of Proteins...</b>	<b>1</b>
1.1 Introduction.....	1
1.1.1 Targeting single natural amino acids.....	1
1.1.2 Targeting peptide tags .....	2
1.1.2.1 Targeting large peptide tags .....	2
1.1.2.2 Targeting small peptide tags.....	3
1.1.3 Bioorthogonal ligation reactions .....	3
1.1.3.1 Aldehyde/ketone condensation.....	4
1.1.3.2 Staudinger ligation.....	5
1.1.3.3 Copper catalyzed azide alkyne cycloaddition (CuAAC).....	5
1.1.3.4 Strain promoted azide alkyne cycloaddition (SPAAC).....	5
1.1.3.5 Inverse electron demand Diels-Alder reaction (IEDDA) .....	6
1.2 Boronic acids as bioorthogonal probes in the service of site-selective protein labeling .....	7
1.2.1 Iminoboronates in bioorthogonal chemistry.....	9
1.2.2 Boronic esters in bioorthogonal chemistry.....	19
1.2.3 Boronic acid as a transient group in bioorthogonal chemistry .....	22
1.3 Summary and future directions.....	30
1.4 Thesis objectives.....	35
1.5 References.....	36
<b>Chapter 2. Fast and Tight Boronate Formation for ‘Click’ Bioorthogonal Conjugation....</b>	<b>43</b>
2.1 Introduction.....	43
2.1.1 Pinanediol boronate stability compared to other 1,2- <i>cis</i> diols .....	43
2.2 Results and discussion .....	45
2.2.1 Optimization of boronate formation – hydrolytic stability & kinetic studies .....	45
2.2.1.1 Initial attempts – stopped flow technique & RP-HPLC .....	45
2.2.1.2 <sup>1</sup> H NMR spectroscopy .....	47
2.2.1.3 Overall results and discussion of nopol-boronate optimization .....	61

2.2.1.4. Fluorescence quenching kinetic study – Förster Resonance Energy Transfer (FRET) quenching .....	75
2.2.2 Competitive effect of biological polyols on boronate formation <i>via</i> RP-HPLC .....	83
2.2.3 Proof of concept on model proteins.....	87
2.2.3.1 Labeling of boronic acid-BSA conjugate with fluorescein-derivatized diol <i>via</i> SDS PAGE.....	87
2.2.3.2 Boronate formation on thioredoxin (Trx) analyzed <i>via</i> HPLC-MS .....	89
2.2.4 Cytotoxicity study .....	97
2.3 Conclusions.....	99
2.4 Experimental.....	99
2.4.1 General information.....	99
2.4.2 Chemical synthesis and analytical data .....	100
2.4.3. Determining $K_{eq}$ <i>via</i> NMR.....	149
2.4.4 Kinetic study.....	151
2.4.4.1 NMR study .....	151
2.4.4.2 Fluorescence quenching study – FRET .....	152
2.4.5 Biological competition assay <i>via</i> RP-HPLC .....	154
2.4.6 SDS-PAGE study .....	155
2.4.7 Trx study.....	156
2.4.8 Cytotoxicity assay .....	159
2.5 References.....	159
<b>Chapter 3. Synergic ‘Click’ Boronate/Thiosemicarbazone System for Fast and Irreversible Bioorthogonal Conjugation in Live Cells.....</b>	<b>163</b>
3.1 Introduction.....	163
3.2 Results and discussion .....	166
3.2.1 Optimization of the synergic system .....	166
3.2.1.1 Design and synthesis of nopoldiol derivatives with an amine/ <i>O</i> -hydroxylamine unit .....	166
3.2.1.2 Conjugation studies – $^1\text{H}$ NMR analysis .....	172
3.2.1.3 Conjugation studies – HPLC-MS analysis .....	174
3.2.1.4 Stability study of <b>3-3aa</b> <i>via</i> a crossover experiment .....	178

3.2.2 Design of hydrazine-functionalized fluorogenic nopoldiol derivatives .....	180
3.2.2.1 Conjugation studies of diols <b>3-1h/3-1i</b> with boronic acid <b>3-2c</b> .....	187
3.2.2.2 Stability assessment of <b>3-3hc/3-3ic</b> <i>via</i> a cross-over study .....	190
3.2.2.3 Study of the fluorogenicity of the reaction between diol <b>3-1i</b> with boronic acid <b>3-2c</b> .....	193
3.2.3 Design of hydrazine-functionalized fluorescently labeled nopoldiol derivatives .....	195
3.2.3.1 Conjugation study of diol <b>3-1j</b> and <b>3-2c</b> and stability test of conjugate <b>3-3jc</b> .....	196
3.2.4 Design of fluorescein conjugated nopoldiol thiosemicarbazide.....	204
3.2.4.1 Conjugation studies of diol <b>3-1k</b> and boronic acid <b>3-2c</b> .....	206
3.2.4.2 Stability studies of the conjugate <b>3-3kc</b> .....	216
3.2.4.3 Mechanistic study – kinetic studies .....	218
3.2.4.4 Studies to assess the bioorthogonality of the boronate/thiosemicarbazone system .....	230
3.2.4.5 Cytotoxicity and live cell labeling studies.....	237
3.3 Conclusions.....	243
3.4 Experimental.....	244
3.4.1 Experimental details and general information.....	244
3.4.2 Chemical synthesis and analytical data .....	245
3.4.3 Optimization of the synergic system .....	270
3.4.3.1 Conjugation and stability study – HPLC-MS method.....	270
3.4.4 Design of fluorogenic nopoldiol derivatives with hydrazine .....	271
3.4.4.1 Conjugation and stability study – HPLC-MS method.....	271
3.4.5 Design of fluorescently labeled nopoldiol derivative with hydrazine.....	271
3.4.5.1 Conjugation and stability study – HPLC-MS method.....	271
3.4.5.2 Elemental composition analysis of the reaction <b>3-1j</b> and <b>3-2c</b> <i>via</i> high resolution- ESI .....	271
3.4.6 Design of fluorescein conjugated nopoldiol thiosemicarbazide.....	271
3.4.6.1 Conjugation studies – <sup>1</sup> H-NMR analysis.....	271
3.4.6.2 Conjugation studies – HPLC-MS analysis .....	272
3.4.6.3 Elemental composition analysis of <b>3-3kc/3-3lc/3-3mc</b> formations <i>via</i> high resolution-ESI.....	273

3.4.6.4 Stability studies of conjugate <b>3-3kc</b> .....	273
3.4.6.5 Mechanistic study – kinetic studies .....	274
3.4.6.6 Biological competition assay .....	275
3.4.6.7 Study to confirm the inertness and bioorthogonality of diol <b>3-1k</b> towards model aldehyde <b>3-2l</b> .....	276
3.4.6.8 Cytotoxicity assay .....	277
3.4.6.9 Cell culture and labeling procedures .....	278
3.5 References .....	280

## **Chapter 4. Investigation of a Reactive Peptide Tag Towards 2-Acetylarylboronic Acid via**

<b>Phage-Display Screening .....</b>	<b>284</b>
4.1 Introduction .....	284
4.2 Results and discussion .....	290
4.2.1 Overview and design of phage display protocol .....	290
4.2.2 Design and synthesis of biotinylated reagents used in the selection .....	293
4.2.3 Using phage display to select a 2-acetylarylboronic acid-binding peptide .....	298
4.2.3.1 First attempt .....	298
4.2.3.2 Second attempt .....	301
4.2.4 Characterizing the selected peptides – screening and interpretation of data .....	302
4.2.5 Analyzing the reactivity of peptides toward 2-acetylbenzeneboronic acid (2-ABBA) .....	305
4.3 Conclusions .....	311
4.4 Experimental .....	312
4.4.1 General information .....	312
4.4.2 Chemical synthesis and analytical data .....	313
4.4.3 Phage amplification and preparation for sequencing .....	321
4.4.4 Synthesis of peptides on solid support .....	322
4.4.5 Low-resolution ESI-MS flow injection analysis of the reactions between peptides <b>1-12</b> and 2-ABBA/acetophenone/phenylboronic acid .....	323
4.4.6 High-resolution ESI-MS analysis of the reactions between peptides <b>7-12</b> and 2- ABBA/acetophenone/phenylboronic acid .....	323
4.4.7 <sup>1</sup> H NMR analysis of the reactions between peptides <b>5/6/11</b> and 2-ABBA .....	324
4.5 References .....	324

<b>Chapter 5. Conclusions</b> .....	<b>327</b>
5.1 Conclusions and future perspectives.....	327
5.2 References.....	331
<b>Bibliography</b> .....	<b>333</b>
<b>Appendices</b> .....	<b>345</b>
Appendix 1: Selected copies of NMR spectra .....	345
Appendix 2: MatLab script .....	390
Appendix 3: HRMS and HPLC-MS chromatograms of peptides <b>1–12</b> .....	394

## List of Tables

<b>Table 1-1:</b> Most commonly applied bioorthogonal reactions. ....	7
<b>Table 1-2:</b> Summary of the applications of boronic acids as bioorthogonal reagents in the service of site-selective protein labeling. ....	33
<b>Table 2-1:</b> Hydrolytic stability and pKa values of selected arylboronic acids with diol <b>2-2a</b> . ....	61
<b>Table 2-2:</b> Results of hydrolytic stability study, forward rate constant ( $k_{ON}$ , $M^{-1}s^{-1}$ ) and association constant measurements ( $K_{eq}$ , $M^{-1}$ ). ....	64
<b>Table 2-3:</b> Selected results of hydrolytic stability study, forward rate constant ( $k_{ON}$ , $M^{-1}s^{-1}$ ) and association constant measurement ( $K_{eq}$ , $M^{-1}$ ). ....	74
<b>Table 2-4:</b> Competitive effect of biological 1,2- <i>cis</i> diols on boronate formation. ....	85
<b>Table 3-1:</b> Results of initial conjugation attempts of diol derivatives ( <b>3-1a</b> – <b>3-1f</b> ) with selected boronic acids ( <b>3-2a</b> – <b>3-2e</b> ). ....	175
<b>Table 3-2:</b> Results of the aqueous conjugation reaction between <b>3-2j</b> and <b>3-2c</b> , and the stability test of conjugate <b>3-3jc•H<sub>2</sub>O</b> and <b>3-3jc</b> via HPLC-MS analysis. ....	198
<b>Table 3-3:</b> Results of the aqueous conjugation reaction of <b>3-1k</b> , and negative controls <b>3-1l</b> and <b>3-1m</b> , with <b>3-2c</b> . ....	208
<b>Table 3-4:</b> Examination of the reactivity of <b>3-1n</b> , <b>3-1o</b> and control <b>3-1f</b> with <b>3-2c</b> to identify the rate determining step. ....	222
<b>Table 4-1:</b> Conjugation results of peptides <b>1–12</b> with 2-ABBA via low-resolution ESI-MS. ...	308

## List of Figures

- Figure 2-1:** a) Pinanediol forms the most stable boronates with arylboronic acids in the presence of various competing diols. b) This work involves the design of a bioorthogonal ‘click’ boronate conjugation using conjugatable nopoldiol derivatives and optimal arylboronic acids in neutral aqueous media. .... 45
- Figure 2-2:** Representative  $^1\text{H}$  NMR spectra for hydrolytic stability study of boronate **2-4a**. A) 5 mM of boronate **2-4a**. B) 1 mM of boronate **2-4a**..... 52
- Figure 2-3:** Representative  $^1\text{H}$  NMR spectrum for hydrolytic stability study of 1 mM of boronate **2-5a**..... 53
- Figure 2-4:** Representative  $^1\text{H}$  NMR spectrum for  $K_{\text{eq}}$  measurement of boronate **2-5k**. .... 55
- Figure 2-5:** Representative time dependent  $^1\text{H}$  NMR analysis for the forward reaction between diol **2-2a** and boronic acid **2-3a**. A)  $^1\text{H}$ -NMR spectrum of **2-2a** in the designated buffer system. B) Full spectrum. C) Diol **2-2a** and boronate **2-4a**  $-\text{OCH}_3$  region. D) Diol **2-2a**, boronate **2-4a**  $-\text{OCH}_3$  peaks and internal standard DMSO region. .... 57
- Figure 2-6:** Representative forward rate constant measurements, second order plot of kinetics **2-2a** and **2-3a**. .... 58
- Figure 2-7:** Representative forward rate constant measurements, second order plot of kinetics **2-2b** and **2-3d**..... 58
- Figure 2-8:** Representative time dependent  $^1\text{H}$  NMR analysis for backward reaction of boronate **2-5d**..... 59
- Figure 2-9:** Backward reaction and  $K_{\text{eq}}$  vs time graph of **2-5d**..... 60
- Figure 2-10:** Backward reaction and  $K_{\text{eq}}$  vs time graph of **2-5h**..... 60
- Figure 2-11:** Linear correlation of hydrolytic stability vs pKa values of arylboronic acids..... 63
- Figure 2-12:** A) Förster Resonance Energy Transfer (FRET) quenching mechanism between a donor fluorophore and a dark quencher. B) Donor emission and dark quencher absorption spectral overlap..... 76

<b>Figure 2-13:</b> Second order plots for kinetic data of diol <b>2-1</b> and dabcyyl-tagged boronate <b>2-9a</b> at 10 and 20 $\mu\text{M}$ . ( $k_{\text{ON}} = 340 \pm 40 \text{ M}^{-1}\text{s}^{-1}$ ).....	81
<b>Figure 2-14:</b> Second order plots for kinetic data of diol <b>2-1</b> and dabcyyl-tagged boronate <b>2-9b</b> at 10 $\mu\text{M}$ . ( $k_{\text{ON}} = 1200 \pm 300 \text{ M}^{-1}\text{s}^{-1}$ ).....	82
<b>Figure 2-15:</b> Estimation of $K_{\text{eq}}$ from backward reaction of <b>2-10a</b> , and forward reaction of diol <b>2-1</b> and boronate <b>2-9a</b> .....	82
<b>Figure 2-16:</b> Estimation of $K_{\text{eq}}$ from backward reaction of <b>2-10b</b> , and forward reaction of diol <b>2-1</b> and boronate <b>2-9b</b> .....	83
<b>Figure 2-17:</b> HPLC chromatograms of diol <b>2-1</b> and <b>2-3d</b> (10 $\mu\text{M}$ ) in the absence of 1,2- <i>cis</i> diols. ....	86
<b>Figure 2-18:</b> HPLC chromatograms of diol <b>2-1</b> and <b>2-3d</b> (10 $\mu\text{M}$ ) in the presence of 1,2- <i>cis</i> diols. ....	86
<b>Figure 2-19:</b> HPLC chromatograms of diol <b>2-1</b> and <b>2-3d</b> (20 $\mu\text{M}$ ) in the absence of 1,2- <i>cis</i> diols. ....	86
<b>Figure 2-20:</b> HPLC chromatograms of diol <b>2-1</b> and <b>2-3d</b> (20 $\mu\text{M}$ ) in the presence of 1,2- <i>cis</i> diols. ....	87
<b>Figure 2-21:</b> Labeling of boronic acid modified BSA with nopoldiol-fluorescein-diol ( <b>2-14</b> ) and gel results. A) Modification of free cysteine with <b>2-12</b> or <b>2-13</b> and labeling with <b>2-14</b> . B) Time-dependent and C) dose-dependent fluorescent labeling of BSA on gel containing phosphate buffer. Protein loading was determined by Coomassie blue staining. Compound <b>2-13</b> is a negative control probe devoid of a boronic acid. D) Fluorescent labeling of BSA on gel containing tris buffer (a triol), inducing breakdown of conjugates. ....	88
<b>Figure 2-22:</b> Boronate formation on Trx analyzed by HPLC-MS. A) Modification of Trx with <b>2-12</b> and subsequent labeling with <b>2-2b</b> . B) HPLC result without <b>2-2b</b> . C) HPLC result in the presence of <b>2-2b</b> . ....	90

<b>Figure 2-23:</b> A) HPLC chromatograms for Trx, boronic acid <b>2-12</b> and diol <b>2-2b</b> in the absence of serum. B) LC-MS trace of boronic acid-Trx. C) LC-MS trace of Bpin-Trx. D) LC-MS trace of boronate-Trx. ....	91
<b>Figure 2-24:</b> A) HPLC chromatograms for Trx, boronic acid <b>2-12</b> and diol <b>2-2b</b> in the presence of serum. B) LC-MS trace of free Trx. C) LC-MS trace of boronate-Trx (12333 Da).....	92
<b>Figure 2-25:</b> Boronate formation on Trx in complete cell media analyzed by fluorescence HPLC-MS. A) Modification of Trx with <b>2-12</b> and subsequent labeling with <b>2-1</b> . B) Fluorescence HPLC result without <b>2-12</b> . C) Fluorescence HPLC result in the presence of <b>2-1</b> . ....	93
<b>Figure 2-26:</b> A) HPLC-UV data for Trx, compound <b>2-12</b> and diol <b>2-1</b> in complete cell media. B) Fluorescence HPLC chromatograms for Trx, compound <b>2-12</b> and diol <b>2-1</b> in complete cell media. C) LC-MS trace for diol <b>2-1</b> . D) LC-MS trace free boronate formed from diol <b>2-1</b> and boronic acid <b>2-12</b> . E) LC-MS trace for boronate-Trx. F) LC-MS trace for double boronate-Trx. ....	95
<b>Figure 2-27:</b> A) HPLC-UV data for Trx and diol <b>2-1</b> in complete cell media. B) Fluorescence HPLC data. C) LC-MS trace for diol <b>2-1</b> . D) LC-MS trace for reduced Trx. E) LC-MS trace for unreduced Trx.....	97
<b>Figure 2-28:</b> Cytotoxicity results of bioorthogonal handles <b>2-15</b> and <b>2-3d</b> on HEK293 cells. Blank was DMSO, and Triton X-100 was used as a positive control. ....	98
<b>Figure 2-29:</b> The principle and mechanism of the WST-1 assay. ....	98
<b>Figure 3-1:</b> A) Previously reported ‘click’ boronate formation system. B) Enhancement in the rate and stability of the imine/hydrazone/oxime ligation in the presence of a boronyl unit in <i>ortho</i> -position. C) Positive cooperativity and higher stability of a three-component amine/boronic acid/catechol system. D) New synergic design based on the optimal placement of a hydrazine or (thio)semicarbazide on the nopoldiol reagent, along with a carbonyl handle ( <i>ortho</i> or <i>meta</i> ) on the arylboronic acid. ....	165

<b>Figure 3-2:</b> Representative <sup>1</sup> H NMR spectra of optimization studies. A) Only diol <b>3-1a</b> . B) Formation of conjugate <b>3-3aa</b> in 2 h. C) Formation of conjugate <b>3-3aa</b> formation in 24 h. ....	174
<b>Figure 3-3:</b> Representative HPLC-MS chromatogram for optimization studies. A) Representative HPLC chromatogram for boronate and oxime formation between <b>3-1e</b> and <b>3-2a</b> . B) LC-MS trace of conjugate <b>3-3ea</b> . ....	175
<b>Figure 3-4:</b> HPLC-MS chromatogram of reaction of <b>3-1a</b> and <b>3-2a</b> after 24 h. A) HPLC chromatogram of conjugate <b>3-3aa</b> . B) LC-MS trace of <b>3-3aa</b> . ....	179
<b>Figure 3-5:</b> HPLC-MS chromatogram of reaction between <b>3-1a</b> and <b>3-2a</b> , followed by the addition of <b>3-2h</b> after 24 h. A) HPLC chromatogram of the reaction mixture after 24 h. B) LC-MS traces of the reaction mixture after 24 h. ....	180
<b>Figure 3-6:</b> Schematic example of a fluorogenic bioorthogonal biolabeling reaction. ....	181
<b>Figure 3-7:</b> A) HPLC chromatogram of diol <b>3-1h</b> and <b>3-2c</b> reaction after 2 h. B) (+)-LC-MS trace of <b>3-3hc</b> . C) (+)-LC-MS trace of <b>3-3hc•H<sub>2</sub>O</b> . ....	189
<b>Figure 3-8:</b> A) HPLC chromatogram of diol <b>3-1i</b> and <b>3-2c</b> reaction after 2 h. B) (+)-LC-MS trace of <b>3-1i</b> . C) (+)-LC-MS trace of <b>3-3ic</b> . D) (+)-LC-MS trace of <b>3-3ic•H<sub>2</sub>O</b> ... ..	190
<b>Figure 3-9:</b> A) HPLC chromatogram of cross-over experiment of <b>3-3hc</b> with <b>3-2i</b> after 1 h. B) HPLC chromatogram of cross-over experiment of <b>3-3hc</b> with <b>3-2i</b> after 24 h. C) (+)-LC-MS trace of <b>3-3hc</b> D) (+)-LC-MS trace of <b>3-3hi</b> . ....	192
<b>Figure 3-10:</b> A) HPLC chromatogram of cross-over experiment of <b>3-3ic</b> with <b>3-2i</b> after 1 h. B) HPLC chromatogram of cross-over experiment of <b>3-3ic</b> with <b>3-2i</b> after 24 h. C) (+)-LC-MS trace of <b>3-3ic</b> . ....	193
<b>Figure 3-11:</b> Fluorogenicity study of the reaction between diol <b>3-1i</b> and boronic acid <b>3-2c</b> . A) Monitoring fluorescence intensity change of the reaction <b>3-1i</b> with <b>3-2c</b> (50:50 μM) at excitation 440 nM and at emission 560 nM. B) Fluorescence intensity comparison ( $I/I_0$ ) vs time: fluorescence intensity of the reaction mixture (I) over fluorescence intensity of <b>3-1i</b> ( $I_0$ ). Results are the average values of triplicate testing. ....	195

<b>Figure 3-12:</b> UV and Fluorescence HPLC-MS analyses of reaction of <b>3-1j</b> and <b>3-2c</b> (50/50 $\mu\text{M}$ ) in 10 mM PBS buffer. ....	201
<b>Figure 3-13:</b> UV and Fluorescence HPLC analyses of the reaction of <b>3-1j</b> and <b>3-2c</b> (250/250 $\mu\text{M}$ ) in ACN/10 mM PBS buffer (35/65 w%) before and after dilution/cross-over study. ....	202
<b>Figure 3-14:</b> High resolution ESI-MS results for the reaction between <b>3-1j</b> with <b>3-2c</b> . ....	204
<b>Figure 3-15:</b> Stability of diol <b>3-1k</b> . A) $^1\text{H}$ NMR spectrum of diol <b>3-1k</b> in 50 mM $\text{D}_2\text{O}$ phosphate buffer/ACN- $\text{d}_3$ (65/35 w%) in one day. B) $^1\text{H}$ NMR spectrum of diol <b>3-1k</b> in one week. ....	206
<b>Figure 3-16:</b> Representative $^1\text{H}$ NMR spectra for the reaction of <b>3-1k/3-1l/3-1m</b> with <b>3-2c</b> in 10 mM PBS buffer/ $\text{D}_2\text{O}$ (9/1 v/v). A) $^1\text{H}$ NMR spectrum of boronic acid <b>3-2c</b> . B) $^1\text{H}$ NMR spectrum of only diol <b>3-1k</b> . C) $^1\text{H}$ NMR expanded spectrum of boronate <b>3-3kc</b> formation over time. D) $^1\text{H}$ NMR spectrum of boronate <b>3-3kc</b> after 24 h. E) $^1\text{H}$ NMR spectrum of only <b>3-1l</b> . F) $^1\text{H}$ NMR spectrum of <b>3-3lc</b> formation after 24 h. G) $^1\text{H}$ NMR spectrum of only <b>3-1m</b> . H) $^1\text{H}$ NMR spectrum of <b>3-3mc</b> formation after 24 h.....	212
<b>Figure 3-17:</b> UV and Fluorescence HPLC analyses of the reaction <b>3-1k</b> (50 $\mu\text{M}$ ) and <b>3-2c</b> (50 $\mu\text{M}$ ) in 10 mM PBS buffer. ....	214
<b>Figure 3-18:</b> High resolution ESI-MS results for the reaction between <b>3-1k</b> and <b>3-2c</b> in 10 mM ammonium acetate.....	216
<b>Figure 3-19:</b> Studies on the stability of <b>3-3kc</b> <i>via</i> HPLC-MS analysis. ....	217
<b>Figure 3-20:</b> HPLC-MS chromatograms of stability studies of the conjugate <b>3-3kc</b> in 10 mM PBS buffer. ....	218
<b>Figure 3-21:</b> Representative time dependent $^1\text{H}$ NMR analysis for the reaction of diol <b>3-1k</b> and <b>3-2c</b> . A) $^1\text{H}$ NMR spectrum of <b>3-1k</b> in corresponding buffer system. B) Full spectrum. C) Diol <b>3-1k</b> and boronate <b>3-3kc</b> methyl region. ....	220
<b>Figure 3-22:</b> Rate constant measurements, second order kinetics plot for <b>3-1k</b> and <b>3-2c</b> .....	221

<b>Figure 3-23:</b> Representative <sup>1</sup> H NMR spectra for mechanistic studies. A) <sup>1</sup> H NMR spectrum of boronic acid <b>3-2c</b> . B) <sup>1</sup> H NMR spectrum of diol <b>3-1n</b> . C) <sup>1</sup> H NMR spectrum of boronate <b>3-3nc</b> in 10 min. D) <sup>1</sup> H NMR spectrum of boronate <b>3-3nc</b> in 120 min. E) <sup>1</sup> H NMR spectrum of diol <b>3-1o</b> . F) <sup>1</sup> H NMR spectrum of boronate <b>3-3oc</b> in 30 min. G) <sup>1</sup> H NMR spectrum of boronate <b>3-3oc</b> in 24 h. H) <sup>1</sup> H NMR spectrum of diol <b>3-1f</b> . I) <sup>1</sup> H NMR spectrum of boronate <b>3-3fc</b> in 24 h. ....	227
<b>Figure 3-24:</b> Rate constant measurements, second order kinetics plot for <b>3-1k</b> and <b>3-2k</b> . ....	229
<b>Figure 3-25:</b> Competitive effect of biological polyols on the formation of conjugates <b>3-3kc/3-3kk</b> . ....	230
<b>Figure 3-26:</b> Representative HPLC-MS chromatogram of reaction of diol <b>3-1k</b> with <b>3-2c</b> (10 μM) in the presence of 1,2- <i>cis</i> diols. ....	232
<b>Figure 3-27:</b> Study to confirm the inertness and bioorthogonality of diol <b>3-1k</b> towards model aldehyde <b>3-2l</b> , a mimic of protein-bound glyoxylyl unit. ....	234
<b>Figure 3-28:</b> HPLC-MS chromatograms from a study to confirm the inertness and bioorthogonality of diol <b>3-1k</b> towards model aldehyde <b>3-2l</b> , a mimic of protein-bound glyoxylyl unit. ....	237
<b>Figure 3-29:</b> Cytotoxicity results of bioorthogonal reagents <b>3-1o</b> , <b>3-2c</b> and <b>3-2k</b> on HEK293T cells. Blank is DMSO, and Triton X-100 was used as a positive control. ....	238
<b>Figure 3-30:</b> The SNAP-tag approach exhibiting the expression of the SNAP <sub>F</sub> -Beta-2 adrenergic membrane receptor on the cell membrane. ....	240
<b>Figure 3-31:</b> Imaging of boronic acid <b>3-2m</b> surface labeled HEK293T live cells with diol <b>3-1k</b> . A) Scheme of installation of boronic acid. B) Scheme of labeling. C) Imaging results; i) HEK293T live cells not incubated with <b>3-2m</b> , but only incubated with <b>3-1k</b> (30 μM) in 2.5 h. ii) HEK293T live cells incubated with both <b>3-2m</b> (10 μM) and <b>3-1l</b> (30 μM) in 2.5 h. iii) HEK293T live cells incubated with both <b>3-2m</b> (10 μM) and <b>3-1m</b> (30 μM) in 2.5 h. iv) HEK293T live cells incubated with both <b>3-2m</b> (10 μM) and <b>3-1k</b> (30 μM) in 0.5 h. v) HEK293T live cells incubated with	

both **3-2m** (10  $\mu$ M) and **3-1k** (10  $\mu$ M) in 2.5 h. Scale bar corresponds to 10  $\mu$ m. Hoechst 33342 was used for live cell nuclear staining. .... 242

**Figure 4-1:** A) Labeling proteins with chelating peptide tags. B) Labeling proteins with enzyme substrate peptide tags. Common methods to site-selectively label proteins in live cells. a) Transfer of recombinant DNA of the POI-tag fusion into the cell. b) Addition of a metal- or metalloid-containing compound into the cells and the resulting emission of a signal. c) Installation of a chemical handle into the tag on the POI *via* an enzyme. d) Direct installation of a fluorophore into the tag on the POI *via* an enzyme. e) Bioorthogonal labeling of the chemical handle with its bioorthogonal reagent pair. .... 286

**Figure 4-2:** Reported short peptide tags that undergo enzyme-free, highly selective bioconjugation reactions with biocompatible chemical reagents. A) Labeling lysine amino group embedded in a unique 21-mer peptide onto the POI with 1,3-diketone derivatives at neutral pH. B) Labeling *N*-terminal histidine of a specific 11-mer peptide onto the POI with hydrazide derivatives. C) Labeling cysteine residue of a unique 4-mer peptide onto the POI with perfluoroaromatic reagents. D) Labeling cysteine residue of a specific 11-mer peptide onto the POI with 2-cyanobenzothiazole derivatives. .... 288

**Figure 4-3:** A) Synergic, fast and irreversible thiosemicarbazone/boronate bioorthogonal system. B) Positive cooperativity and higher stability of a three-component amine/boronic acid/catechol system. C) Labeling proteins with a short and reactive peptide tag containing the right geometry of alcohol and amino groups that favor a stable ligation with 2-acetylarylboronic acid through imine and boronate formation. .... 289

**Figure 4-4:** A common screening protocol using a phage display library with M13 phage. .... 291

**Figure 4-5:** Using a phage display to select for 2-acetylarylboronic acid-reactive peptides. .... 292

**Figure 4-6:** First attempt to select 2-acetylarylboronic acid reactive peptides. .... 300

**Figure 4-7:** A) Second attempt to select 2-acetylarylboronic acid reactive peptides. B) Expected cleavage of the conjugation under acidic elution buffer from beads. C) Definition of test and control sets. .... 302

**Figure 4-8:** Post-selection analysis on Ser-X7 against the reagent **4-1a**. A) Heat map describing the peptide hits from the intersection of A/B and A/C. B) Volcano plots showing sequences that are enriched ( $R \geq 3$ ,  $p\text{-value} \leq 0.1$ ) using set B and C as controls. The sets of these sequences were abbreviated as A/B and A/C. The composition of the enriched sequences was indicated as red on the top right of volcano plots. The Figure 4-8 was created by *MakeFigure4-8.m* MatLab script, which is available in Appendix 2. .... 304

**Figure 4-9:** A representative result of low-resolution ESI-MS analysis of the reaction between peptide **1** and 2-ABBA/phenylboronic acid/acetophenone. .... 309

**Figure 4-10:** A representative result of  $^1\text{H}$  NMR analysis of the reaction between peptide **5** and 2-ABBA..... 311

**Figure 5-1:** Application of the ‘click’ boronic ester system in fragment-based drug discovery. .... 330

## List of Schemes

<b>Scheme 1-1:</b> General scheme of site-selective protein labeling.....	2
<b>Scheme 1-2:</b> Formation of tetrahedral (sp <sup>3</sup> ) adducts of boronic acids. ....	8
<b>Scheme 1-3:</b> Four different strategies to exploit boronic acids in bioorthogonal chemistry. ....	9
<b>Scheme 1-4:</b> A) Reversible protein modification <i>via</i> iminoboronate formation. B) Receptor-mediated internalization <i>via</i> fluorescent-folic acid iminoboronate conjugate. ....	10
<b>Scheme 1-5:</b> Rapid oxime condensation between equimolar concentration of 2-FBBA and <i>N</i> -benzylhydroxylamine under neutral pH.....	12
<b>Scheme 1-6:</b> A) Mechanistic proposal of iminoboronate formation that leads to fast and reversible conjugation chemistry of $\alpha$ -nucleophiles at neutral pH. B) Application of hydrazone formation on a model protein, VHP35.....	13
<b>Scheme 1-7:</b> A) Rapid formation of an irreversible boron heterocycle DAB, <b>B</b> , through iminoboronate formation, and SDS-PAGE results of labeling BSA or BSA-hydrazine with coumarin-modified 2-formyl-4 hydroxyphenylboronic acid. B) The discovery of a fluorogenic variant of BIQ formation. ....	15
<b>Scheme 1-8:</b> A) Fast and irreversible labeling of bacterial pathogens <i>via</i> diazaborine formation. B) Side reaction of 2-FBBA: fast and selective labeling of <i>N</i> -terminal cysteines <i>via</i> TzB formation. ....	16
<b>Scheme 1-9:</b> Chemoselective modification of <i>N</i> -terminal cysteine and internal cysteine of calcitonin <i>via</i> reversible TzB formation on demand. ....	18
<b>Scheme 1-10:</b> A) Higher stability of a three-component amine/2-FBBA/catechol system. B) Three-component, boronic acid-mediated couplings of catechols and <i>N</i> -hydroxylamines. ....	19
<b>Scheme 1-11:</b> Boronic acid for the traceless delivery of proteins into cells.....	20

<b>Scheme 1-12:</b> A) Cell-permeable turn-on fluorescent rhodamine-derived bis-boronic acid (RhoBo) for selective recognition of protein tetraserine motifs. B) Labeling SSPGSS motif-incorporated Hela cells with RhoBo. ....	21
<b>Scheme 1-13:</b> Highly stable, rapid bioorthogonal boronic ester formation used to assemble a monomeric weak agonist into a dimeric strong agonist at a cellular level. ....	22
<b>Scheme 1-14:</b> Application of Suzuki-Miyaura cross-coupling on a chemically modified model protein SBL-156ArI. ....	23
<b>Scheme 1-15:</b> A) <i>E. coli</i> surface labeling with Suzuki-Miyaura cross-coupling. B) <i>E. coli</i> surface labeling with novel glucose-, galactose- and mannose-based boronic acids followed by a modulation of interactions with fluorescein-lectin conjugates. ....	25
<b>Scheme 1-16:</b> A) Site-selective <sup>18</sup> F-protein-labeling. B) Site-selective PEGylation with Suzuki-Miyaura cross-coupling. ....	27
<b>Scheme 1-17:</b> Oxidative Heck reaction as a bioorthogonal reaction. ....	28
<b>Scheme 1-18:</b> The use of alkenylboronic acid IEDDA/Carboni-Lindsey reaction to enhance the reactivity of non-strained alkenes. ....	29
<b>Scheme 1-19:</b> Fluorogenic bond-cleaving transformation between <i>N</i> -oxide and boron reagent. ....	30
<b>Scheme 2-1:</b> Initial approaches for kinetic studies and observation of boric ester <b>2-1</b> from the HPLC analysis. ....	47
<b>Scheme 2-2:</b> Synthesis of (–)-nopol-PEG-diol-1 <b>2-2a</b> and its corresponding boronic esters <b>2-4</b> . ....	48
<b>Scheme 2-3:</b> Synthesis of (–)-nopol-PEG-diol-2 <b>2-2b</b> and its corresponding boronic esters <b>2-5</b> . ....	49
<b>Scheme 2-4:</b> Attempts to synthesize (–)-nopol-aryl-diol <b>2-2c</b> . ....	50
<b>Scheme 2-5:</b> Representative scheme of hydrolytic stability study and association constant measurement ( $K_{eq}$ , $M^{-1}$ ). ....	51
<b>Scheme 2-6:</b> Alizarin Red S (ARS) assay system to determine the $K_{eq}$ of an arylboronic acid <b>2-3a</b> towards diol <b>2-2b</b> . ....	54
<b>Scheme 2-7:</b> Representative scheme of forward rate constant measurements ( $k_{ON}$ , $M^{-1}s^{-1}$ ). ....	56
<b>Scheme 2-8:</b> Equation for the monitoring of a backward reaction. ....	60

<b>Scheme 2-9:</b> Proposed mechanism for the boronate formation. ....	64
<b>Scheme 2-10:</b> Synthesis of boronic acid <b>2-3k</b> .....	73
<b>Scheme 2-11:</b> Synthesis of dansyl-tagged diol <b>2-1</b> and nopol-dansyl without 1,2- <i>cis</i> diol <b>2-11</b> . 78	
<b>Scheme 2-12:</b> Synthesis of dabcyI derivatives of the corresponding neopentyl boronates <b>2-9a/2-9b</b> .....	79
<b>Scheme 2-13:</b> Kinetic analysis using a fluorescence quenching experiment to measure the initial rate of boronate formation.....	80
<b>Scheme 2-14:</b> Competitive effect of biological polyols on boronate formation.....	84
<b>Scheme 3-1:</b> Desired conjugation test between nopoldiol amine derivatives ( <b>3-1a – 3-1f</b> ) and selected boronic acids with aldehyde/acetylketone/phenylketone ( <b>3-2a – 3-2e</b> ) in <i>ortho</i> - or <i>meta</i> - position.....	167
<b>Scheme 3-2:</b> Synthesis of nopoldiol amine derivatives <b>3-1a</b> and <b>3-1b</b> .....	168
<b>Scheme 3-3:</b> Synthesis of nopoldiol phenylamine <b>3-1c</b> and nopoldiol benzylamine <b>3-1d</b> .....	169
<b>Scheme 3-4:</b> Possible resonance forms of oximes. ....	169
<b>Scheme 3-5:</b> Synthesis of nopoldiol hydroxylamine <b>3-1e</b> . ....	170
<b>Scheme 3-6:</b> Scheme of the desired conjugation test between nopoldiol aldehyde <b>3-1g</b> and boronic acids with amine in the <i>ortho</i> - or <i>meta</i> - position <b>3-2f</b> and <b>3-2g</b> . ....	170
<b>Scheme 3-7:</b> Attempts for the synthesis of nopoldiol aldehyde <b>3-1g</b> . ....	171
<b>Scheme 3-8:</b> Another strategy for the synthesis of nopoldiol aldehyde <b>3-1g</b> . ....	171
<b>Scheme 3-9:</b> Attempted Kornblum oxidation for the synthesis of nopoldiol aldehyde <b>3-1g</b> .....	172
<b>Scheme 3-10:</b> Stability test of <b>3-3aa</b> <i>via</i> a cross-over experiment.....	179
<b>Scheme 3-11:</b> Design of fluorogenic or fluorescently labeled diol derivatives towards faster and tighter boronate formation.....	181
<b>Scheme 3-12:</b> Fluorogenic boronate formation. A) Selected examples of fluorogenic boronic acid sensors. B) Fluorogenicity study of diol <b>2-2b</b> and 4-sulfonaphthalic anhydride complex. ....	183

<b>Scheme 3-13:</b> Fluorogenic hydrazone ligation. A) Fluorogenic bodipy hydrazone formation. B) Fluorogenic NBD hydrazine derivatives.....	184
<b>Scheme 3-14:</b> The conjugation of nopoldiol NBD hydrazine designs <b>3-1h</b> and <b>3-1i</b> with boronic acids <b>3-2c</b> .....	185
<b>Scheme 3-15:</b> Synthesis trials of nopoldiol NBD hydrazine design <b>3-1h</b> .....	186
<b>Scheme 3-16:</b> Synthesis of nopoldiol NBD hydrazine designs <b>3-1h</b> and <b>3-1i</b> . A) Synthesis of diol <b>3-1h</b> . B) Synthesis of diol <b>3-1i</b> .....	187
<b>Scheme 3-17:</b> The conjugation of diols <b>3-1h/3-1i</b> (2.0 mM) with <b>3-2c</b> (2.5 mM).....	188
<b>Scheme 3-18:</b> Stability test of <b>3-3hc/3-3ic</b> via a cross-over experiment. ....	191
<b>Scheme 3-19:</b> Fluorogenicity study of diol <b>3-1i</b> and <b>3-2c</b> .....	194
<b>Scheme 3-20:</b> Synthesis of NBD-benzyl bromide <b>3-15</b> .....	196
<b>Scheme 3-21:</b> Synthesis of nopoldiol hydrazine-functionalized NBD <b>3-1j</b> .....	196
<b>Scheme 3-22:</b> Syntheses of nopoldiol thiosemicarbazide <b>3-1k</b> . ....	205
<b>Scheme 3-23:</b> Synthesis of the conjugatable <i>o</i> -acetylphenylpinacol boronate with a <i>p</i> -amide group <b>3-2k</b> . ....	228
<b>Scheme 3-24:</b> Rate constant for the reaction between <b>3-1k</b> and <b>3-2k</b> as measured by <sup>1</sup> H NMR spectroscopy. ....	229
<b>Scheme 4-1:</b> Initial design of biotin derivatized 2-formylarylboronic acid reagent <b>4-2a</b> and control reagents <b>4-2b</b> and <b>4-2c</b> .....	294
<b>Scheme 4-2:</b> Attempts to synthesize biotin derivatized 2-formylarylboronic acid reagent <b>4-2a</b> . ....	295
<b>Scheme 4-3:</b> Second design of biotin derivatized 2-formylarylboronic acid reagent <b>4-3a</b> and control reagents <b>4-3b</b> and <b>4-3c</b> .....	296
<b>Scheme 4-4:</b> Another attempt to synthesize biotin derivatized 2-formylarylboronic acid reagent <b>4-3a</b> .....	296
<b>Scheme 4-5:</b> Design of biotin derivatized 2-acetylarylboronic acid reagent <b>4-1a</b> and control reagents <b>4-1b</b> and <b>4-1c</b> . ....	297

<b>Scheme 4-6:</b> Synthesis of biotin derivatized 2-acetylarylboronic acid reagent <b>4-1a</b> and control reagents <b>4-1b</b> and <b>4-1c</b> . .....	298
<b>Scheme 4-7:</b> Reactions of peptides <b>1–12</b> with 2-ABBA/negative controls and their analysis <i>via</i> low-resolution ESI-MS. ....	307
<b>Scheme 4-8:</b> Reactions of peptides <b>7–12</b> with 2-ABBA/negative controls and their analysis <i>via</i> high-resolution ESI-MS. ....	310
<b>Scheme 4-9:</b> Reactions of peptides <b>5/6/11</b> with 2-ABBA and their analysis <i>via</i> <sup>1</sup> H NMR analysis.....	310
<b>Scheme 5-1:</b> Fluorogenic hydrazone ligation. A) Fluorogenic bodipy hydrazone formation. B) A proposed fluorogenic design with the nopoldiol bodipy hydrazine for future investigations.....	329
<b>Scheme 5-2:</b> A HPLC method with spectrofluorimetric detection using the ‘click’ boronate system to detect, analyze and quantify trace amounts of boronic acids in pharmaceuticals.....	331

## List of Abbreviations

2-ABBA	2-Acetylbenzeneboronic acid
2-FBBA	2-Formylbenzeneboronic acid
9-BBN	9-Borabicyclo(3.3.1)nonane
AA	Acetic acid
ABDH	4-4-Aminosulphonyl-7-hydrazino-2,1,3-benzoxadiazole
Ac	Acetyl
ACN	Acetonitrile
ACP	Acyl carrier protein
ACS	American Chemical Society
ADHP	2-Amino-4,6-dihoxypyrimidine
ADR $\beta$ 2	$\beta$ 2 Adrenergic receptor
AIBN	2,2'-Azobis(2-methylpropionitrile)
app t	Apparent triplet
ARS	Alizarin Red S
BIAN	Bis(aryl)acenaphthequinonediimine
BIQ	4,3-Borazaroisoquinoline
BirA	Biotin ligase
Boc	<i>tert</i> -Butyloxycarbonyl
bodipy	4,4-Difluoro-4-bora-3a,4a-diaza-s-indacene
br	Broad
BSA	Bovine serum albumin
<i>n</i> -Bu	Normal butyl
<i>t</i> -Bu	<i>tert</i> -Butyl
c-MpI	Myeloproliferative leukemia virus
calcd	Calculated
cm <sup>-1</sup>	Wavenumbers
comp m	Complex multiplet
CuAAC	Copper catalyzed azide alkyne cycloaddition

Cy	Cyanine
DAB	1,2-Dihydro-1-hydroxy-2,3,1-benzodiazaborine
DBDH	4-( <i>N,N</i> -Dimethylaminosulphonyl)-7-hydrazino-benz-2,1,3-oxadiazole
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
dd	Doublet of doublets
ddd	Doublet of doublet of doublets
dq	Doublet of quartets
dt	Doublet of triplets
DFT	Density functional theory
DIPEA	<i>N,N</i> -Diisopropylethylamine
DMAP	4-(Dimethylamino)pyridine
DMEM	Dulbecco's modified eagle's medium
DMF	<i>N,N</i> -Dimethylformamide
DMP	Dess-Martin periodinane
DMPU	1,3-Dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribose nucleoside triphosphate
dppf	1,1'-Bis(diphenylphosphino)ferrocene
dsDNA	Double strand deoxyribonucleic acid
ECFP	Enhanced cyan fluorescent protein
<i>E. Coli</i>	Escherichia Coli
EDTA	Ethylenediaminetetraacetic acid
EI	Electron impact
eq	Equation
equiv	Equivalent
ESI	Electrospray ionization
Et	Ethyl
EtOH	Ethanol
FBS	Fetal bovine serum

FlAsH	Fluorescein arsenical hairpin
FRET	Förster Resonance Energy Transfer
FP	Fluorescent protein
GFP	Green fluorescent protein
GSH	Glutathione
h	Hour
HATU	1-[Bis(dimethylamino)methylene]-1 <i>H</i> -1,2,3-triazolo[4,5 <i>b</i> ]pyridinium 3-oxid hexafluorophosphate
HBTU	<i>N,N,N',N'</i> -Tetramethyl- <i>O</i> -(1 <i>H</i> -benzotriazol-1-yl)uranium hexafluorophosphate
HEK	Human embryonic kidney
HPLC	High performance liquid chromatography
HR-ESI	High resolution electrospray ionization
HRMS	High resolution mass spectrometry
IEDDA	Inverse electron demand Diels-Alder reaction
IPA	Isopropanol
IPTG	Isopropyl- $\beta$ -D-thiogalactoside
IR	Infrared spectroscopy
$K_d$	Dissociation constant
$K_{eq}$	Equilibrium constant
lacZ	Bacterial $\beta$ -galactosidase gene
LC-MS	Liquid chromatography–mass spectrometry
LiHMDS	Lithium hexamethyldisilazide
m	Multiplet
MBP	Maltose-binding protein
MBq	Megabecquerel
Me	Methyl
MEF	Mouse embryonic fibroblast
MeOH	Methanol
$\mu$ L	Microliter
$\mu$ M	Micromolar

min	Minute
mL	Milliliter
mmol	Millimole
mol	Mole
mp	Melting point
Ms	Methanesulfonyl
NBDH	7-Hydrazino-4-nitrobenzo-2-oxa-1,3,-diazole
NBS	<i>N</i> -Bromosuccinimide
NMO	<i>N</i> -Methylmorpholine <i>N</i> -oxide
NMR	Nuclear magnetic resonance
Nu	Nucleophile
Omp	Outer membrane protein
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PET	Photo-induced electron transfer
pfu	Plaque forming unit
pin	Pinacolato
Ph	Phenyl
POI	Protein of interest
q	Quartet
qd	Quartet of doublets
ReAsH	Resorufin arsenical hairpin
RhoBo	Rhodamine-derived bisboronic acid
RNase	Ribonuclease
rt	Room temperature
SBL	Subtilisin Bacillus lentus
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
S <sub>N</sub> 2	Bimolecular nucleophilic substitution
S <sub>N</sub> Ar	Nucleophilic aromatic substitution
SPAAC	Strain promoted azide alkyne cycloaddition

ssDNA	Single strand deoxyribonucleic acid
t	Triplet
TBAI	Tetrabutylammonium iodide
TEMPO	2,2,6,6-Tetramethylpiperidinyloxy
Tf	Trifluoromethanesulfonate
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
THPTA	Tris(3-hydroxypropyltriazolylmethyl)amine
TIRF	Total internal reflection fluorescence
TLC	Thin-layer chromatography
TML	Trimethyl lock
TMS	Trimethylsilyl
TPAP	Tetrapropylammonium perruthenate
Tris	Trisaminomethane
Trx	Thioredoxin
Ts	Tosyl
TzB	Thiazolidine boronate
UV/Vis	Ultraviolet/visible
VHP	Villin headpiece
WST	Water-soluble tetrazolium
X-gal	5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside

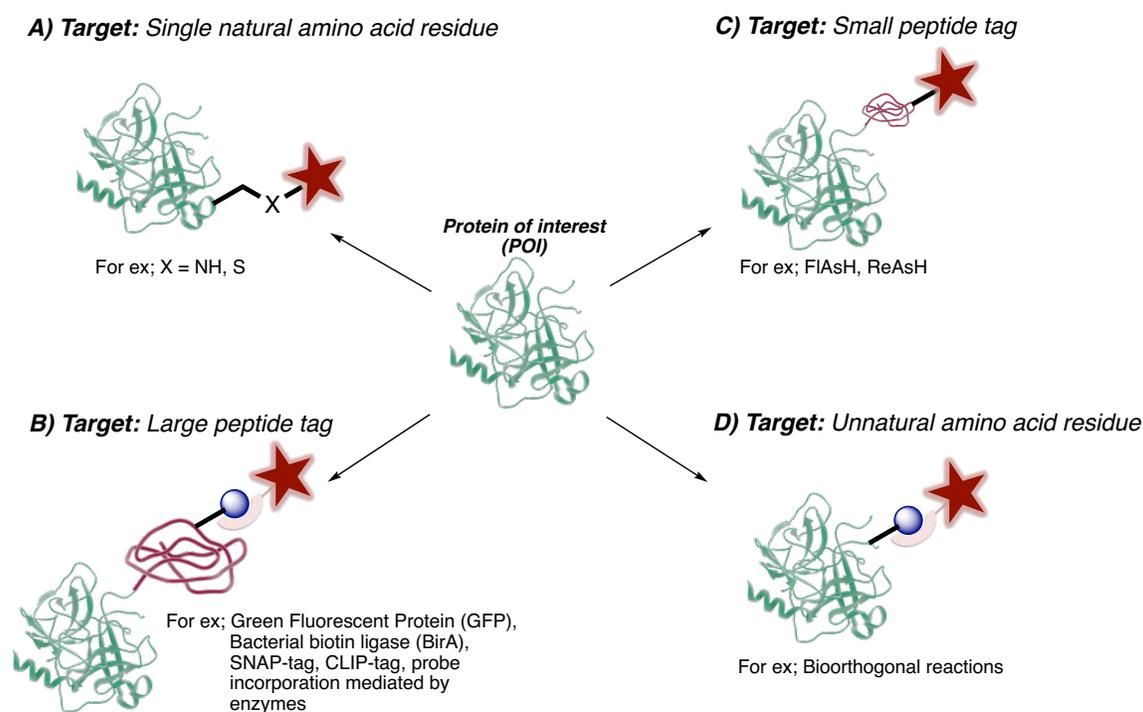
# 1 Chapter 1. Boronic Acids as Bioorthogonal Probes in Site-Selective Labeling of Proteins

## 1.1 Introduction

Nature effortlessly and continuously creates various post-translational chemical modifications of proteins at the cellular level. These subtle changes can dramatically influence the proteins' dynamics, localization, and interactions with other biomolecules. Deciphering the proteins' roles is essential to generate new fundamental biology knowledge and to enable therapeutic applications. Mimicking nature's ability is, however, challenging due to the need for a highly site-selective modification of the desired protein. This particular chemical modification must occur chemo- and regio-selectively under ambient, neutral aqueous conditions, as well as in the presence of numerous reactive functional groups, such as amines, alcohols and carboxylic acids. Various sets of reactions have been developed in the past two decades to address this challenge.<sup>1-6</sup> In many examples, these methods have been utilized in the labeling and tracking of proteins with fluorescent dyes in live cells,<sup>7,8</sup> for installing post-translational modifications of proteins<sup>9,10</sup> and for constructing antibody drug conjugates.<sup>11,12</sup>

### 1.1.1 Targeting single natural amino acids

Reactions with natural amino acids can occur at a single residue with low abundant amino acids such as cysteine and lysine (Scheme 1-1A).<sup>2,3,13</sup> These reactions are limited to the modification of isolated proteins *in vitro*. For instance,  $\alpha$ -halocarbonyl electrophiles such as iodoacetamides are used to modify cysteine residues within proteins.<sup>14</sup> Yet cysteine may be present in more than one copy in the protein of interest (POI), as well as in other proteins in the biological system being studied.



**Scheme 1-1:** General scheme of site-selective protein labeling.

## 1.1.2 Targeting peptide tags

### 1.1.2.1 Targeting large peptide tags

Another approach, which targets peptide tags, resolves the selectivity problem and also can be applied *in vivo*.<sup>15</sup> The motivation to design peptide tags emanated with the idea of using Green Fluorescent Protein (GFP) as a label to track proteins in 1994.<sup>16</sup> Later, Roger Tsien showed that GFP can be used as a labeling tool to understand the role of proteins.<sup>17,18</sup> Other naturally occurring proteins and biochemically derived fluorescent proteins (FPs) were also developed for the same purpose.<sup>18,19</sup> One major limitation of this method is the large size of FPs, generally 240 amino acids,<sup>20</sup> which may alter the native protein's structure and its role in cells. As a replacement to this bulky protein, peptide tags have been widely investigated.<sup>1,15,21</sup> They serve as specific sites designed or evolved for covalent binding to chemical probes (Scheme 1-1B). In many cases, the attachment of chemical groups requires enzymes. For example, in the ACP-tag system, acyl carrier protein (ACP) is incorporated into the POI; thus an essential serine residue of ACP can be individually labeled with a 4'-phosphopantetheine-linked probe from coenzyme A, in the presence of an external enzyme, a phosphopantetheine transferase.<sup>22</sup> Another example is the use of biotin ligase, which leads to the covalent labeling of a 15-aa peptide tag

with ketone-modified biotin molecules.<sup>23</sup> Next, this modified peptide may readily react with hydrazide or hydroxylamine fluorescent dyes. Many other enzymatically-linked peptide tags have been extensively reviewed.<sup>1,15</sup> However, the main limitation of these methods is the necessity for an external enzyme. Other methods, such as Halo-tag and SNAP-tag, do not require an enzyme, but they still involve the incorporation of large sized peptide into the POI.<sup>1,15</sup>

### 1.1.2.2 Targeting small peptide tags

A potentially better method to sustain the native function of POI is to target small peptide tags (Scheme 1-1C). For example, the cell-permeable fluorogenic biarsenical reagents FIAsh,<sup>24</sup> ReAsH<sup>25</sup> and Cy3As<sup>26</sup> have been developed to label a small tetra-cysteine motif at the C-terminus of enhanced cyan fluorescent protein (ECFP). The main drawback of these probes are the cytotoxicity of the metal and background labeling.<sup>27</sup> Ideally, what would be more appealing is to site-specifically functionalize a small natural peptide sequence with a benign small probe without the need for an external enzyme. Only a handful of examples of this strategy have been described in the literature.<sup>28</sup> For instance, Pentelute and co-workers recently reported an elegant example where a four-amino-acid  $\pi$ -clamp unit (Phe-Cys-Pro-Phe) was functionalized site-specifically through cysteine with a perfluoroaryl compound even in the presence of other cysteine residues on the same POI.<sup>28d</sup> The peptide sequence is usually located at the *N*- or *C*-terminus of the POI as an internal part of the protein is difficult to modify with a peptide sequence without interfering in its native structures and functions. Conceivably, a single residue modification would be an optimal way to perform an internal modification of proteins, and would only minimally disrupt their structure and function.

### 1.1.3 Bioorthogonal ligation reactions

Bioorthogonal ligation reactions utilize a biocompatible unnatural functionality introduced into pre-determined sites on proteins or with genetic code expansion techniques, followed by the modification of the protein with another biocompatible synthetic probe (Scheme 1-1D). These ligation methods constitute one of the most preferred tools to achieve single residue site-specific modification of proteins<sup>1-5</sup> as they are fast, selective and compatible even in complex protein mixtures, cells and living organisms.<sup>6,29</sup> Consequently, these reactions have been widely applied in various concepts such as the post-translational modification of proteins,<sup>9,10</sup> antibody-drug conjugate formation,<sup>11,12</sup> live-cell imaging *in vitro* and *in vivo*<sup>7,8</sup> or

site-selective PEGylation of proteins.<sup>30-34</sup> To achieve these applications, an ideal bioorthogonal reaction must be simple and readily implementable, and fulfill some key properties;

- Large reaction rate constant (ideally within a few minutes at  $\mu\text{M}/\text{nM}$  concentration)
- High selectivity and conversion
- Low reagent concentration
- Easily synthesized/accessible reagents
- Small, stable and benign reagents
- Stable product bonds
- Cell-permeable reagents to penetrate cells and compartments within the cell, and to get into specific tissues within animals
- Fluorogenic bond formation for bioimaging to eliminate the need to wash out excess labeling reagent
- Cross-selectivity with existing bioorthogonal reactions for dual labeling
- Facile methods to site-selectively incorporate the corresponding unnatural amino acid into the POI

During the last two decades, numerous bioorthogonal reactions<sup>4,5,29</sup> were developed yet the most commonly applied bioorthogonal reactions are limited to aldehyde/ketone condensation, Staudinger ligation, copper catalyzed azide alkyne cycloaddition (CuAAC), strain promoted azide alkyne cycloaddition (SPAAC), and inverse electron demand Diels Alder reaction (IEDDA).

### 1.1.3.1 Aldehyde/ketone condensation

Aldehyde/ketone condensation, one of the first bioorthogonal reactions, is a reversible ligation of reactive carbonyl groups of aldehyde/ketones with  $\alpha$ -nucleophiles such as hydrazines and alkoxyamines (Table 1-1A).<sup>4,36</sup> This dehydrative reaction often requires acidic conditions (pH 4 – 6) in the absence of external catalysts e.g., aniline.<sup>37</sup> Carbonyl compounds with neighboring acid/base groups or hydrazines with an adjacent carboxylic acid were also found to accelerate the rate of aldehyde/ketone condensation at neutral pH.<sup>38</sup> Initially, reported rate constants were mostly in the range of  $10^{-4} - 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$  under uncatalyzed conditions but later the same reactions reached rates up to  $20 \text{ M}^{-1} \text{ s}^{-1}$  with some modifications of substrates and optimized

pH conditions.<sup>4</sup> Aldehyde/ketone condensation is mostly suitable for *in vitro* or cell-surface labeling. One of the reasons is that  $\alpha$ -nucleophiles might react with keto and aldehydic metabolites present within cells and biological systems.<sup>39,40</sup> Acidic conditions, which are especially required for ketone substrates, are also a major limitation in intracellular labeling.<sup>2,4</sup> Using an excess amount of labeling reagent may compensate for the low rates; however, it might lead to toxicity and background signals.

### 1.1.3.2 Staudinger ligation

Organic azides, which are small, stable and inert in the biological environment are at the basis of a significant class of bioorthogonal chemistry.<sup>2,4,35</sup> One of the early bioorthogonal reactions is the Staudinger ligation,<sup>41</sup> which involves the reaction of azides with phosphines, followed by an amide bond formation. A slight modification of the substrate may lead to a product without a phosphine oxide unit; this reaction is coined a traceless Staudinger ligation (Table 1-1B).<sup>42</sup> The Staudinger ligation has applications in the labeling of living cells and even animals but exhibits very slow kinetics ( $10^{-3} \text{ M}^{-1} \text{ s}^{-1}$ ). Furthermore, the phosphine reagent is prone to oxidation *via* air or metabolic enzymes.<sup>3,41,43,44</sup>

### 1.1.3.3 Copper catalyzed azide alkyne cycloaddition (CuAAC)

The reactivity of azides is not limited only to phosphines; azides exhibit high reactivity towards terminal alkynes through [3+2] cycloadditions in the presence of copper (I) salts (Table 1-1C).<sup>35</sup> CuAAC is faster than the Staudinger ligation but it requires a cytotoxic copper (I) catalyst and the reaction rate is highly dependent on the amount of catalyst.<sup>29,45</sup> To be applicable *in vivo*, various copper (I) ligands have been developed.<sup>46,47</sup> For instance, the use of copper chelating azides such as picolyl azides with designed biocompatible ligands may diminish cell toxicity as well as improve the reaction rates.<sup>48</sup> With that, CuAAC has become more important for *in vivo* and live cell labeling studies. The requirement of a catalyst and ligand is, however, inconvenient, especially for animals.

### 1.1.3.4 Strain promoted azide alkyne cycloaddition (SPAAC)

Not long after the advent of CuAAC, SPAAC was developed to eliminate the need for a copper catalyst (Table 1-1D).<sup>3,49</sup> The reactivity of the alkyne partner towards azides was enhanced by exploiting ring strain, leading to various cyclooctyne-based probes.<sup>5,29</sup> These probes

have been applied in the labeling of proteins in live mammalian cells and animals.<sup>44,50-53</sup> The main drawback of SPAAC is its slow kinetics ( $10^{-2} - 1 \text{ M}^{-1}\text{s}^{-1}$ ). Furthermore, a few unwanted side products can be observed due to alkynes' reactivity with biological thiols.<sup>53,54</sup> The synthesis of these strained alkyne reagents may also be challenging and lengthy.

### **1.1.3.5 Inverse electron demand Diels-Alder reaction (IEDDA)**

IEDDA, which occurs between tetrazines and strained alkene or alkyne derivatives, is one of the most favorable bioorthogonal reactions (Table 1-1E) because of the extremely fast rates reaching upwards of  $10^7 \text{ M}^{-1}\text{s}^{-1}$  in some cases.<sup>55,56</sup> Moreover, fluorogenic designs have been reported and have been applied in bacterial and mammalian live cell protein labeling.<sup>56-58</sup> Unnatural amino acids with various reactive alkene or tetrazine groups have been genetically encoded.<sup>6,56,59,60</sup> Nevertheless, some of the highly reactive trans-cyclooctenes might isomerize into their unreactive *cis*- form or react with cellular thiols.<sup>29,61</sup> Furthermore, some tetrazine derivatives are susceptible to hydrolysis.<sup>61</sup>

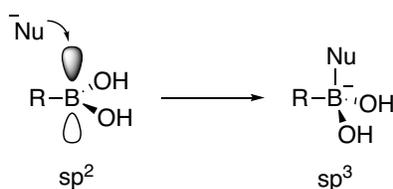
**Table 1-1:** Most commonly applied bioorthogonal reactions.

Reaction type	Reaction scheme	Approximate rate constant ( $M^{-1}s^{-1}$ )	Comments
A) Aldehyde/ketone condensation		$10^{-4} - 20$	<ul style="list-style-type: none"> <li>*Low rates</li> <li>*Cell surface labeling</li> <li>*Slightly acidic pH required</li> <li>*Adducts prone to hydrolysis</li> <li>*Aniline catalyst to enhance the rate</li> </ul>
B) Traceless Staudinger ligation		$10^{-3}$	<ul style="list-style-type: none"> <li>*Low rates</li> <li>*Phosphines susceptible to hydrolysis <i>In vivo</i></li> </ul>
C) CuAAC		$10 - 200$	<ul style="list-style-type: none"> <li>*Fast rates</li> <li>*Small reagents</li> <li>*Copper catalyst and ligands are required</li> </ul>
D) SPAAC		$10^{-2} - 2$	<ul style="list-style-type: none"> <li>*Low rates</li> <li>*No metal required</li> <li>*Applicable <i>in vivo</i></li> <li>*Some cyclooctynes susceptible to thiol attack</li> <li>*Challenging and low yielding alkyne synthesis</li> </ul>
E) IEDDA		$1 - 10^7$	<ul style="list-style-type: none"> <li>*Low to very high rates</li> <li>*Cell surface and intracellular labeling</li> <li>*Fluorogenic designs</li> <li>*Some trans-cyclooctenes' isomerization and side reaction with cellular thiols</li> <li>*Tetrazines might be prone to hydrolysis</li> </ul>

## 1.2 Boronic acids as bioorthogonal probes in the service of site-selective protein labeling

A boronic acid is a derivative of boric acid where one OH group is substituted with an aryl or alkyl group. Boronic acids have a vacant p-orbital, which enables them to coordinate basic molecules and become a mild class of organic Lewis acids. Tetrahedral adducts with a formal and negative charge on the boron atom result (Scheme 1-2). In 1959, Lorand and Edwards reported that the trivalent, neutral form of boronic acids is in equilibrium with the

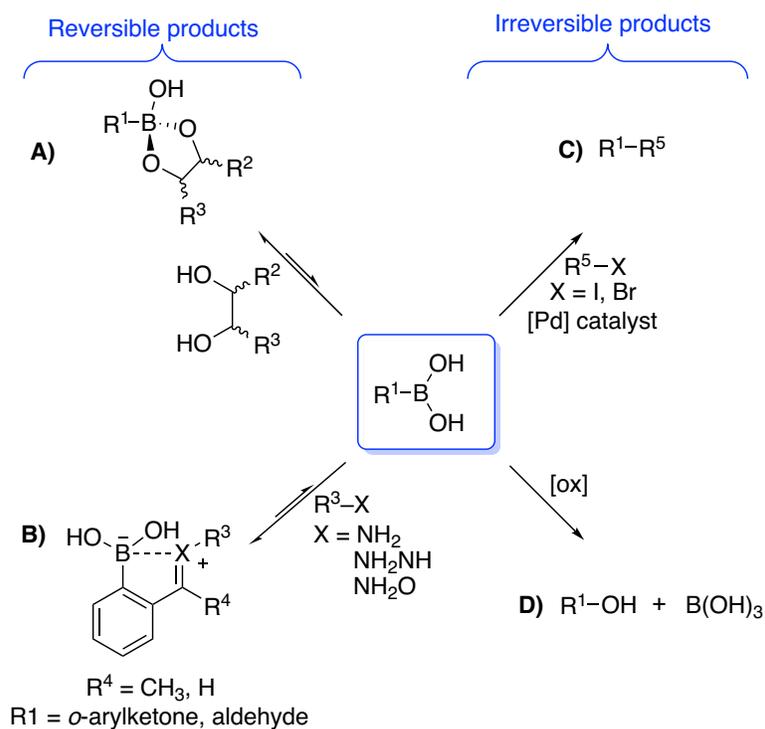
tetrahedral, anionic form.<sup>62</sup> Due to this unique property, boronic acids have been used in diverse fields of applications. For instance, in organic chemistry, boronic acids are significant building blocks in the popular Suzuki-Miyaura cross-coupling reaction,<sup>63</sup> Diels-Alder reactions,<sup>64</sup> carboxylic acid activation<sup>65</sup> and many other organic reactions.<sup>66</sup> With 1,2- or 1,3-*cis* diols, amino alcohols, or with carbohydrates, boronic acids quickly form a covalent boronic ester in neutral aqueous conditions. For this reason, boronic acids were applied as sensors for carbohydrates<sup>67</sup> and amino alcohols,<sup>68</sup> to recognize carbohydrate biomarkers on the cell surface,<sup>69</sup> and also to purify glycoproteins of interest.<sup>70,71</sup> Although it is hard to foresee their use as bioorthogonal probes due to their apparent formation of covalent bonds with biological polyols, boronic acids have been recently introduced in bioorthogonal chemistry. In this chapter, I will discuss recent reports that describe the use of boronic acids as bioorthogonal probes and compare these systems with the most commonly applied bioorthogonal reactions employed in site-selective protein labeling.



**Scheme 1-2:** Formation of tetrahedral ( $\text{sp}^3$ ) adducts of boronic acids.

Before these reports are discussed, it is essential to review the properties that make boronic acids attractive reagents in bioorthogonal chemistry. Even though boronic acids readily form boronic esters with biological polyols at a neutral pH, these covalent bonds are reversible, and their dissociation constants ( $K_d$ ) are in the range of  $200 - 10^5 \mu\text{M}$ .<sup>72</sup> In other words, these boronic esters are unfavorable at low  $\mu\text{M}/\text{nM}$  concentrations that are the norm in bioorthogonal chemistry, thus eliminating unwanted competition from these biological polyols. Moreover, ingenious design of a synthetic 1,2- or 1,3-*cis* diol might lead to a benign and tighter boronic ester (low  $\mu\text{M}/\text{nM}$   $K_d$ ) with a very rapid rate constant ( $\geq 10^3 \text{ M}^{-1}\text{s}^{-1}$ ) (Scheme 1-3A). In other cases, a boronic acid unit might be crucial in favoring the ligation as a directing group. For instance, arylboronic acids can increase the rate of imine/hydrazone/oxime formation almost three-fold due to a dative N-B bond that is made possible when the boronyl substituent is located at the *ortho* position to the aldehyde/ketone functionality (Scheme 1-3B).<sup>73-75</sup> In other cases, such

as in the Suzuki-Miyaura cross-coupling, boronic acids are consumed as a transient group and do not remain in the product. Thus this Nobel Prize-winning chemistry can also be utilized as a bioorthogonal reaction (Scheme 1-3C). Ultimately, arylboronic acids may be susceptible to oxidation, and this reactivity can also be tuned into a bioorthogonal reaction (Scheme 1-3D). In this regard, Schultz and co-workers reported a site-directed unnatural amino acid mutagenesis of boronic acids into GFP,<sup>76</sup> which provides a convenient avenue to further advance and apply these bioorthogonal reactions.

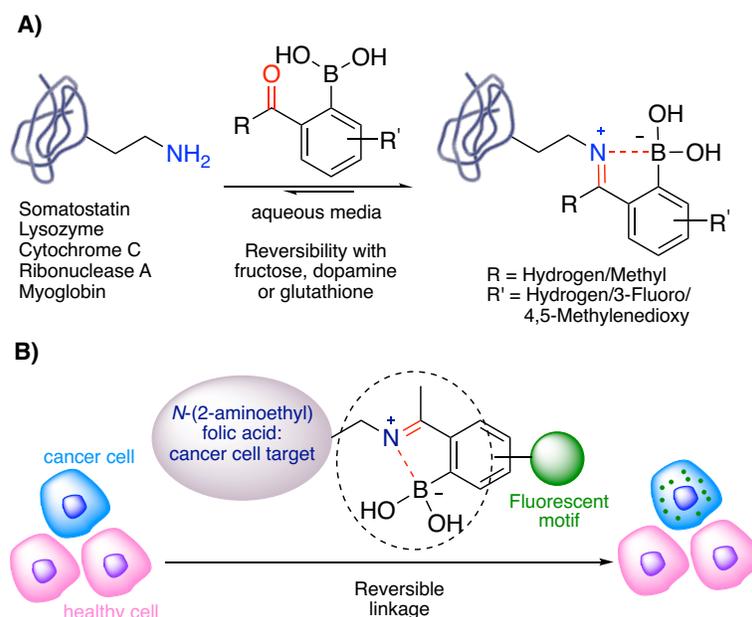


**Scheme 1-3:** Four different strategies to exploit boronic acids in bioorthogonal chemistry.

### 1.2.1 Iminoboronates in bioorthogonal chemistry

The development of iminoboronate ligation has been primarily encouraged by a study reported in 2012.<sup>73</sup> In this study, Gois and co-workers focused on one of the simplest bioconjugation methods, imine formation, because a protein's *N*-terminal and the  $\epsilon$ -amino group of lysine are both attractive targets in bioconjugation chemistry.<sup>4,5</sup> Yet, due to its reversibility in aqueous media, an imine bond often requires a reduction step to be turned into a stable and irreversible amine unit. Gois and co-workers envisioned another way to improve an imine bond's stability through the use of a well-known dative N-B bond (Scheme 1-4A).<sup>73</sup> Their initial results

showed that 2-formylbenzeneboronic acid (2-FBBA) and 2-acetylbenzeneboronic acid (2-ABBA) (0.16 M) formed iminoboronates with lysine (0.16 M) in aqueous media with 61 and 71% yields, respectively. The same boronic acids (50 mM) were capable of fully conjugating somatostatin (10  $\mu$ M), a short-lived hormonal neuropeptide, within five minutes. This fast and efficient iminoboronate formation was also successfully demonstrated on various proteins such as lysozyme, cytochrome c, ribonuclease A and myoglobin. Additionally, the stimuli-responsive reversibility of iminoboronate formation, which is a promising tool for delivery systems such as antibody-drug conjugates, was achieved upon the addition of glutathione (GSH), dopamine or fructose. A detailed density functional theory (DFT) study confirmed the role of the boronic acid: a dative B–N bond exists between the boron atom and the imine nitrogen lone pair. The B–N bond in the final product induces a 7.4 kcal/mol energy gain when compared with the intermediate lacking a B–N bond. Moreover, the B–O bond promotes the nucleophilic attack of the amine and leads to a smaller energy barrier for the formation of the iminoboronate.

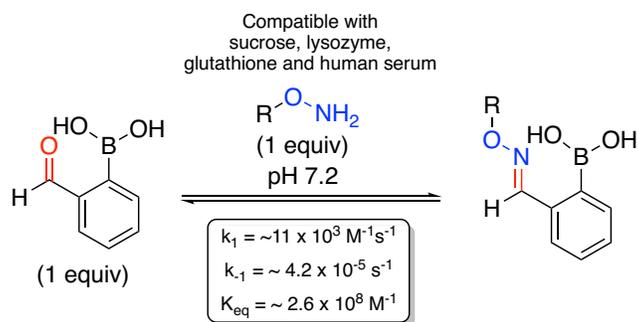


**Scheme 1-4:** A) Reversible protein modification *via* iminoboronate formation. B) Receptor-mediated internalization *via* fluorescent-folic acid iminoboronate conjugate.

Gois and co-workers utilized this iminoboronate system to assemble a fluorescent motif and *N*-(2-aminoethyl) folic acid, which is a vitamin used for cell function (Scheme 1-4B).<sup>77</sup> Receptors for folic acid are overexpressed in cancer cells compared to normal cells. Therefore, 2-

ABBA, which led to the most stable iminoboronate in buffer solutions, was derivatized and conjugated to *N*-(2-aminoethyl) folic acid to help visualize cancer cells. This vitamin-fluorescent motif assembly (20  $\mu\text{M}$ ) specifically entered into cancer cells through a receptor-mediated internalization. In further studies, 2-ABBA was functionalized with azide, alkyne, PEG or a cytotoxic drug, paclitaxel, thus demonstrating its utility in various other ways.<sup>78</sup> Moreover, 2-ABBA derivatives with azide or alkyne were modified *via* SPAAC and CuAAC reactions. Additionally, 2-ABBA-paclitaxel (anti-cancer drug conjugated with 2-ABBA) and folic acid were assembled *via* iminoboronate formation against an NCI-H460 cancer cell line.<sup>78</sup> The construct disassembles in the presence of GSH upon internalization into cancer cells, where GSH is highly abundant. Whilst the designed drug conjugate was cytotoxic towards the NCI-H460 cancer cell line ( $\text{IC}_{50}$  value of 20.7 nM), it was benign against a healthy human-skin fibroblast cell line.

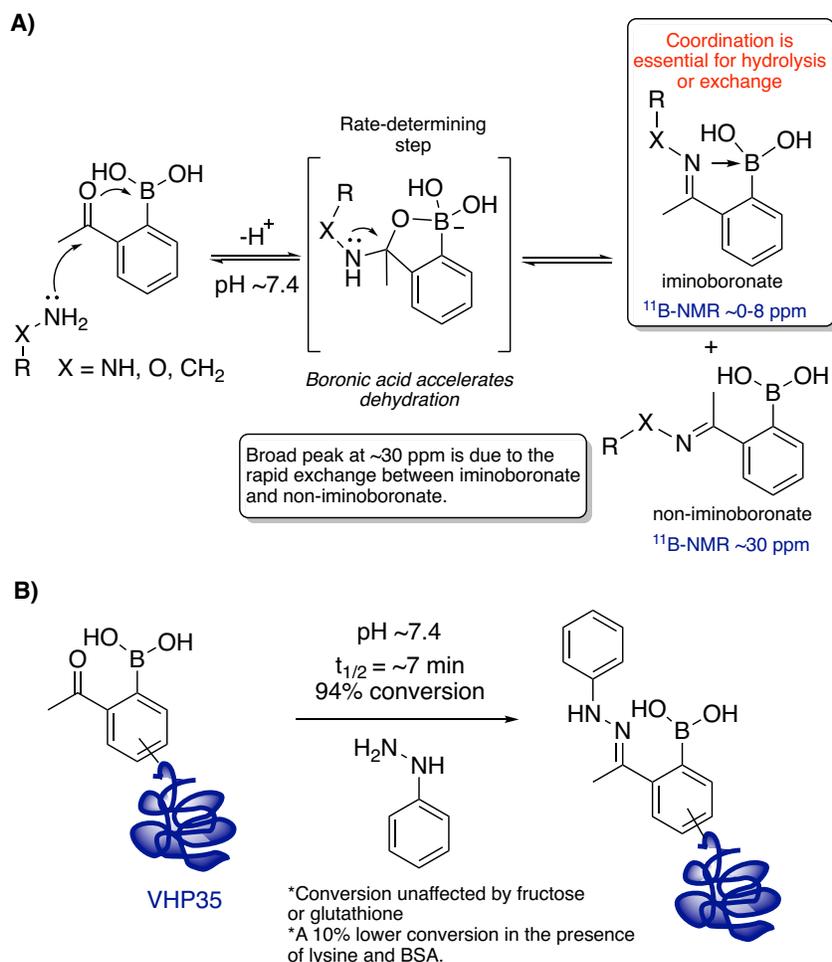
Iminoboronate chemistry is efficient due to its stability and reversibility but its reactivity towards abundant amine groups in biological systems makes it unsuitable for site-selective protein labeling, especially for *in vivo* studies. This chemistry, however, led to the discovery of new bioorthogonal reactions that are presumably applicable in site-selective protein labeling *in vitro* and *in vivo*. In 2015, Gillingham and co-workers reported a rapid alkoxyiminoboronate formation between an equimolar concentration of 2-FBBA and *N*-benzylhydroxylamine at neutral pH (Scheme 1-5).<sup>79</sup> The presence of a boronyl unit removed the requirements for a catalyst and a low pH. The reaction's rate and equilibrium constants ( $k_1 = \sim 11 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ ,  $k_{-1} = \sim 4.2 \times 10^{-5} \text{ s}^{-1}$ ,  $K_{\text{eq}} = \sim 2.6 \times 10^8 \text{ M}^{-1}$ ) were studied with HPLC, NMR and fluorescence quenching. The ligation became reversible when an excess of *O*-methylhydroxylamine was added to alkoxyiminoboronate, and it reached equilibrium in three days. The alkoxyiminoboronate formation was also tested in complex environments with potentially interfering molecules (sugars, biothiols, proteins, human serum) and monitored with reverse phase HPLC, which confirmed the compatibility of the conjugation with those biocompetitors.<sup>79</sup>



**Scheme 1-5:** Rapid oxime condensation between equimolar concentration of 2-FBBA and *N*-benzylhydroxylamine under neutral pH.

Gao and co-workers investigated the effect of *ortho*-substituted arylboronic acids in the conjugation of  $\alpha$ -nucleophiles with aryl ketones.<sup>74</sup> UV/Vis analysis data provided an approximate dissociation constant  $K_d = \sim 14 \mu\text{M}$  for oxime ligation between 6-aminoxy hexanoic acid and 2-ABBA, whereas hydrazone formation of phenylhydrazine with the same boronic acid afforded a more stable adduct  $K_d = \sim 0.07 \mu\text{M}$ . The authors' mechanistic proposal emphasizes the importance of the *ortho* boronyl group in a nucleophilic attack as well as in the dehydration step, which is rate-limiting for imine/oxime/hydrazone formation (Scheme 1-6A). The iminoboronate unit kinetically assists rapid hydrolysis or exchange with water and other nucleophiles (Scheme 1-6A). <sup>11</sup>B-NMR analysis easily distinguished the B–N coordinated iminoboronate product due to an upfield shift from 30 ppm to 8–0 ppm. The authors also observed a broad peak for the B–N non-coordinated iminoboronate product, in the  $\sim 30$ – $25$  ppm range, with a small upfield shift compared to the peak monitored for 2-ABBA. This broad peak corresponds to a rapid exchange between two forms of the products, B–N coordinated iminoboronate and non-coordinated iminoboronate (Scheme 1-6A). This rapid exchange is, however, sufficient for hydrolysis. The more stable conjugate of 2-ABBA and phenylhydrazine exhibited a broad peak ( $\sim 30$  ppm) with a less upfield shift in comparison to the less stable conjugate of 2-ABBA and 6-aminoxy hexanoic acid ( $\sim 25$  ppm). Using the favorable conjugate, a low concentration of phenylhydrazine ( $1 \mu\text{M}$ ) is sufficient to label  $\sim 90\%$  of a model peptide conjugated with 2-ABBA ( $0.2 \mu\text{M}$ ). Its forward rate constant ( $k_1$ ) was measured as  $\sim 1.7 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$  with a reverse rate constant of ( $k_{-1}$ )  $1.2 \times 10^{-4} \text{ s}^{-1}$ . The potential of hydrazone ligation was also demonstrated on a model protein, villin headpiece subdomain (VHP35) (Scheme 1-6B). The half-life of the labeling process was  $\sim 7$  minutes, and a 94% conversion was recorded. Fortunately, hydrazone formation was unaffected

by fructose and GSH, but a 10% lower conversion was monitored in the presence of lysine and bovine serum albumin (BSA).

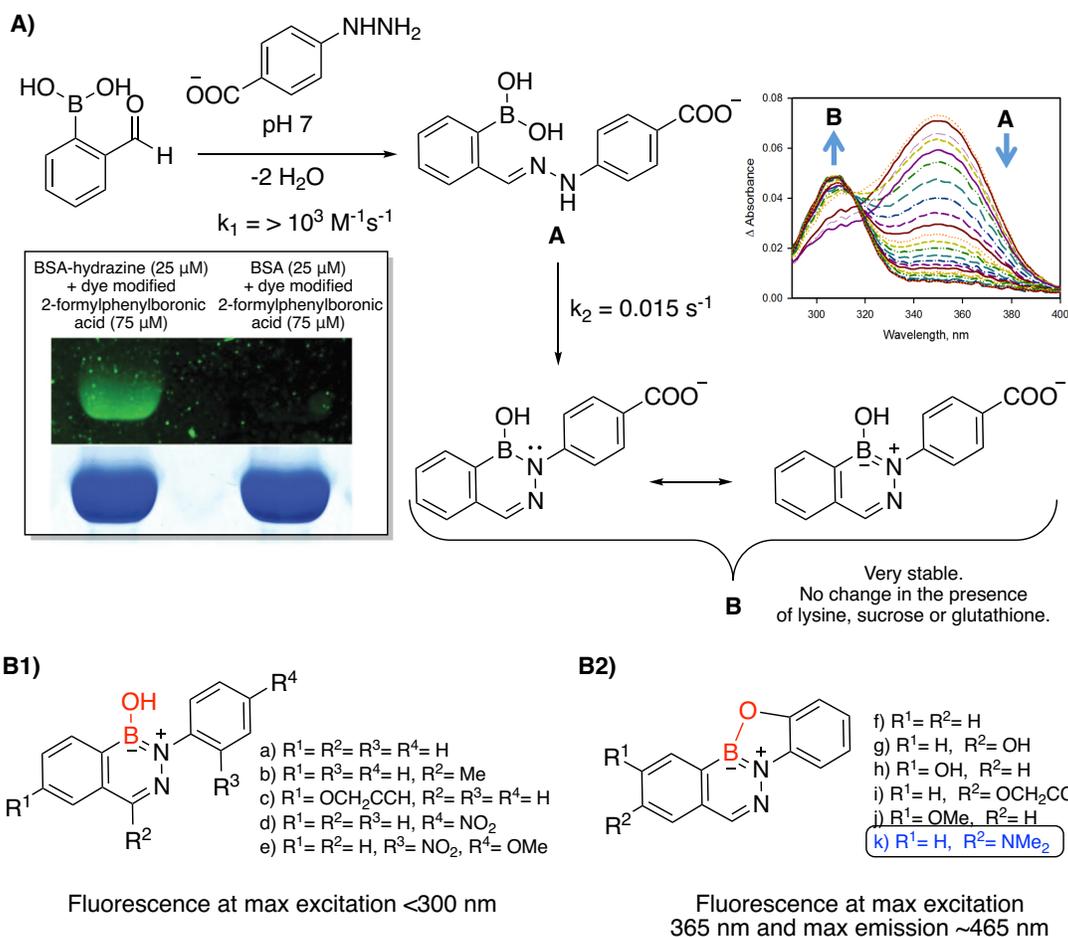


**Scheme 1-6:** A) Mechanistic proposal of iminoboronate formation that leads to fast and reversible conjugation chemistry of  $\alpha$ -nucleophiles at neutral pH. B) Application of hydrazone formation on a model protein, VHP35.

Apart from 2-ABBA, the more reactive 2-FBBA was also tested towards phenylhydrazine. Bane and co-workers demonstrated that a low equimolar concentration of 2-FBBA and 4-hydrazinylbenzoic acid rapidly led to an irreversible product, 1,2-dihydro-1-hydroxy-2,3,1-benzodiazaborine (DAB) (Scheme 1-7A).<sup>80</sup> This boron heterocycle, **B**, formed through a bimolecular iminoboronate, **A**, ( $k_1 = >10^3 \text{ M}^{-1}\text{s}^{-1}$ ) which was monitored by UV spectroscopy at  $\sim 350 \text{ nm}$ , followed by a unimolecular reaction ( $k_2 = 0.015 \text{ s}^{-1}$ ) giving appeared as a new product peak at  $\sim 305 \text{ nm}$ . DAB formation was compatible with the presence of a

millimolar concentration of sucrose, GSH and lysine. Moreover, a coumarin-modified 2-formyl-4-hydroxyphenylboronic acid (75  $\mu\text{M}$ ) was conjugated to a hydrazine-functionalized BSA protein (25  $\mu\text{M}$ ) and the robust ligation was characterized with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Scheme 1-7A).

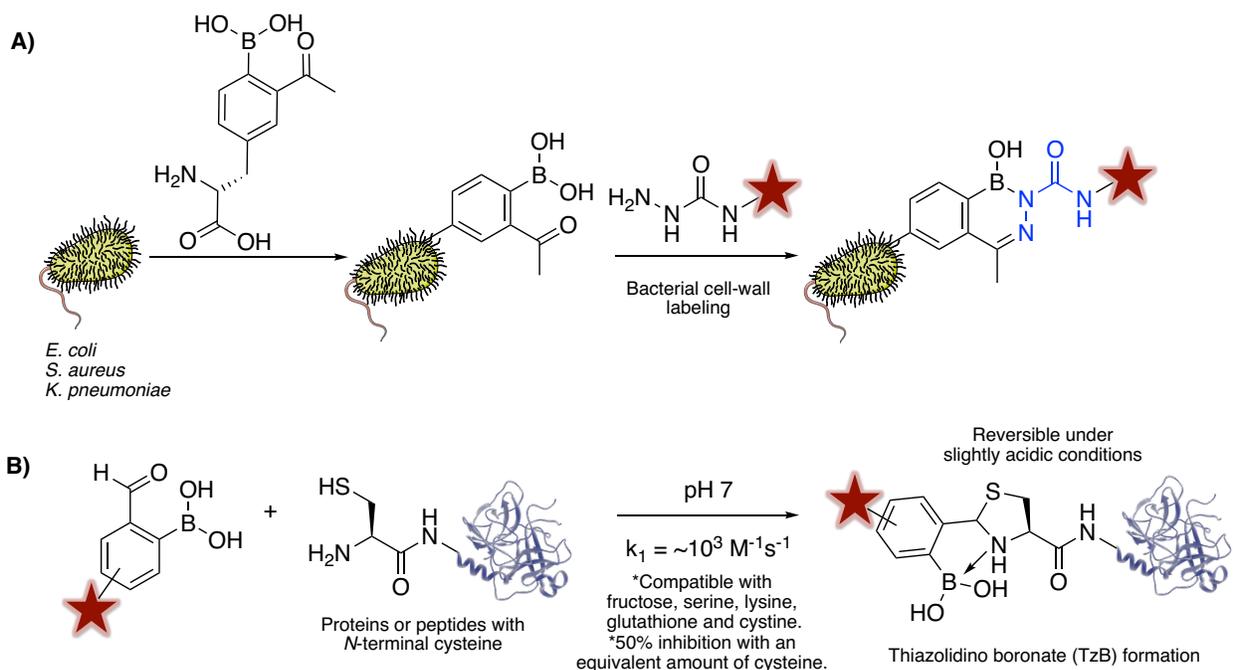
In a related study, Gillingham and co-workers showed that aldoximes have an improved stability over ketoximes by adding 2-FBBA into a ketoxime-containing solution. The observed equilibrium shift, favouring the aldoxime product, was likely as a result of the formyl Schiff base being more kinetically favored (lower transition state energy) over its acetyl equivalent. The authors claimed that this energetic difference was likely due to the propensity of the more basic (because of the extra  $\sigma$ -donation of methyl group) ketoximine nitrogen to remain in the B–N coordinated iminoboronate form, thus favoring its hydrolysis.<sup>81</sup> The aldoxime system might be convenient in cases where slow reversibility or release is desired. Although slow or fast reversibility of oxime/hydrazone is possibly advantageous for some applications,<sup>2</sup> an irreversible bond-forming process is ideal for most applications in bioorthogonal chemistry. Similar to the findings of Bane et. al.,<sup>80</sup> Gillingham and co-workers also observed that arylhydrazines afford highly stable aromatic 4,3-borazaroisoquinolines (BIQs) with 2-FBBA/2-ABBA derivatives *via* a secondary intramolecular reaction.<sup>81</sup> They further designed new substrates and developed fluorogenic systems (Scheme 1-7B). Five BIQs (Scheme 1-7B1) exhibited insufficient fluorescence at maximum excitation wavelengths  $<300$  nm. Through masking the B-OH group (fluorescence quencher) *via* an O-substituted five-membered ring formation, the resulting BIQs became better fluorogenic conjugates (Scheme 1-7B2). From six examples, an electron-rich dimethylamino substituent in the boronic acid gave the best blue fluorophore with a five-fold increase in fluorescence and a bathochromic shift maximum excitation at 365 nm, with a maximum emission at  $\sim 450$  nm.



**Scheme 1-7:** A) Rapid formation of an irreversible boron heterocycle DAB, **B**, through iminoboronate formation, and SDS-PAGE results of labeling BSA or BSA-hydrazine with coumarin-modified 2-formyl-4 hydroxyphenylboronic acid. B) The discovery of a fluorogenic variant of BIQ formation.

Even though phenylhydrazine formed a fast and irreversible conjugate with 2-FBBA derivatives, it is cytotoxic and oxidatively unstable, which limits its use in living systems. Recent work by Gao and co-workers addresses this problem by using a benign and more stable hydrazine derivative, semicarbazide, which enabled a stable conjugate with 2-FBBA or 2-ABBA.<sup>82</sup> This new reaction proceeds through a semicarbazone formation, which rapidly converts into a stable diazaborine (Scheme 1-8A). Although 2-ABBA led to a lower conversion (60% conversion for 50:50  $\mu\text{M}$ ) compared to 2-FBBA (95%), it suffered less interference from biomolecules in blood serum or cell lysates (Scheme 1-8B). Gao and co-workers also applied this bioorthogonal chemistry in the labeling of bacterial cell-wall. A synthetic amino acid

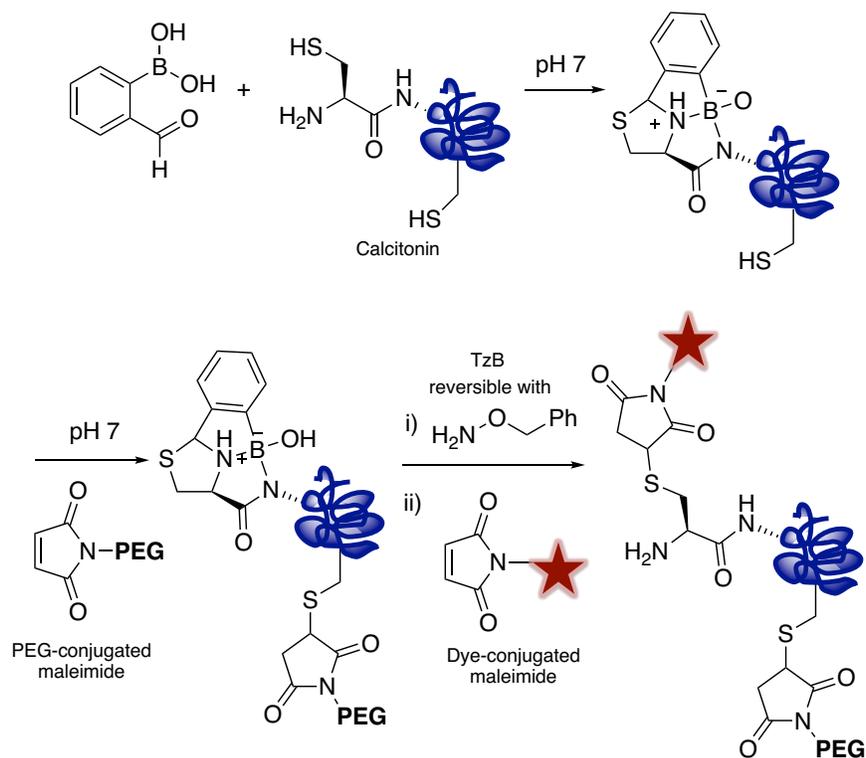
derivative of 2-ABBA was incubated with various bacterial species and incorporated into their cell membrane peptidoglycans (Scheme 1-8A). Next, a fluorophore-labeled semicarbazide was conjugated with the boronic acid on the cell wall of each pathogen. *E. coli* cell-wall labeling was significantly higher than other bacteria's labeling in blood serum. According to the authors, this amino acid may easily cross the outer membrane of Gram-positive *E. coli*, which is free of peptidoglycan. Another explanation might be due to *E. coli* transpeptidase, which plays a role in peptidoglycan remodeling and prefers aromatic amino acids. This system was described as a robust detection method for bacterial pathogens in blood serum. As confirmed in this study, 2-ABBA was the ideal candidate for bioorthogonal chemistry because of its inertness to biomolecules in comparison to 2-FBBA. In the same way, recent reports<sup>83,84</sup> showed that 2-FBBA selectively and rapidly forms thiazolidine boronate (TzB) with the *N*-terminal cysteine of peptides or proteins at a neutral pH (Scheme 1-8B). Gao and co-workers validated this reaction for diazaborine formation between 2-FBBA and semicarbazide in the presence of *S. aureus* cell lysates where *N*-terminal cysteines are abundant.<sup>82</sup>



**Scheme 1-8:** A) Fast and irreversible labeling of bacterial pathogens *via* diazaborine formation. B) Side reaction of 2-FBBA: fast and selective labeling of *N*-terminal cysteines *via* TzB formation.

TzB formation is advantageous towards labeling natural proteinogenic amino acids site-selectively without the need to incorporate an unnatural amino acid into the POI. Cysteine with a reactive sulfhydryl side chain has been an invaluable target for selective labeling due to its low abundance (<2%).<sup>8,85</sup> Even though there are numerous reports featuring this concept, with generally unselective methods,<sup>2</sup> one of the most selective ways is to target *N*-terminal cysteine residue, such as in native chemical ligation,<sup>3</sup> or with aromatic cyanides<sup>4</sup> or aldehydes.<sup>5</sup> Gao and co-workers also used *N*-terminal cysteine as a site-specific labeling residue for proteins and peptides.<sup>83</sup> Their innovative approach was based on the selective formation of thiazolidines between *N*-terminal cysteine and aldehydes,<sup>86</sup> however it suffers from a low rate constant and occurs only at a low pH. The authors quickly envisioned that 2-FBBA was an excellent substitute to rapidly form TzB with *N*-terminal cysteine at a neutral pH.<sup>73,74,79</sup> A significant acceleration in rate ( $\sim 10^3 \text{ M}^{-1} \text{ s}^{-1}$ ) was observed due to the iminoboronate intermediate (Scheme 1-8B).<sup>83</sup> TzB conjugation was selective and stable even in the presence of competitive biomolecules. Moreover, TzB conjugate exhibited rapid reversibility upon mild acidification, which might be practical for the development of drug–protein conjugates that release a small molecule drug in acidic endosomes. However, due to its targeting of *N*-terminal cysteine, this system is not truly considered to be a bioorthogonal reaction and its application is limited to *in vitro* analysis.

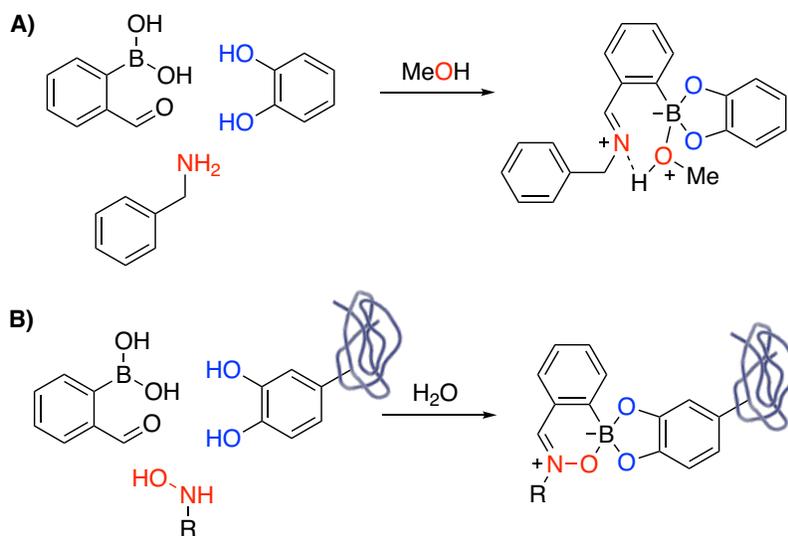
Gois and co-workers explored the orthogonality and reversibility of the TzB system in the dual selective labeling of peptides (Scheme 1-9).<sup>84</sup> The TzB product exhibited high stability at a neutral pH and slightly basic conditions, but was completely reversible in the presence of excess benzylhydroxylamine (20 equiv). The TzB conjugate could selectively mask the *N*-terminal cysteine of calcitonin, thus allowing a PEG-conjugated maleimide to react only with internal cysteines of calcitonin (Scheme 1-9). Afterwards, the TzB conjugate was replaced with another dye-conjugated maleimide resulting in site-selective labeling of *N*-terminal cysteine. Detailed DFT calculations also highlighted the important role of the boronic acid in activating amines *via* a dative B–N bond. The boron plays a dual role in this system; first, electrophilic activation of the resulting imine promotes the double cyclization from the thiol nucleophile. Secondly, the resulting product is further stabilized by coordination of the amide nitrogen, leading to an overall 5-membered tricyclic core.



**Scheme 1-9:** Chemoselective modification of *N*-terminal cysteine and internal cysteine of calcitonin *via* reversible TzB formation on demand.

Recently, Anslyn and James described another way to achieve an irreversible adduct through iminoboronate chemistry.<sup>87a</sup> The authors performed a structural and thermodynamic analysis of *ortho*-iminophenylboronate ester between 2-FBBA, catechol and benzylamine. This study confirmed that binding of catechol and benzylamine to 2-FBBA is cooperative through a solvent insertion in protic media (Scheme 1-10A).<sup>87b</sup> In other words, initial boronate formation between catechol and 2-FBBA enhances the complexation between benzylamine and the formyl unit on the resulting boronic ester. Furthermore, initial imine formation between benzylamine and 2-FBBA improves the stability of the boronate formation between catechol and boronic acid of the imine complex. Using this chemistry, they reported a bioorthogonal reaction involving three components: 2-FBBA, a catechol derivative and a *N*-hydroxylamine derivative. The hydroxyl group of *N*-hydroxylamine takes the role and the same position as the oxygen of an inserted solvent; therefore this boronic acid-mediated coupling of catechol and *N*-hydroxylamines leads to an irreversible conjugation in aqueous media (Scheme 1-10B). This was

deemed compatible with biological functional groups, however the apparent reactivity of 2-FBBA towards *N*-terminal cysteines restricts its application to *in vitro* studies.

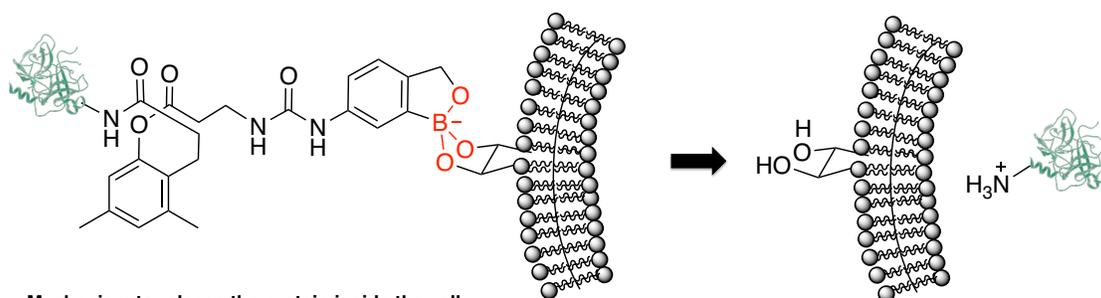


**Scheme 1-10:** A) Higher stability of a three-component amine/2-FBBA/catechol system. B) Three-component, boronic acid-mediated couplings of catechols and *N*-hydroxylamines.

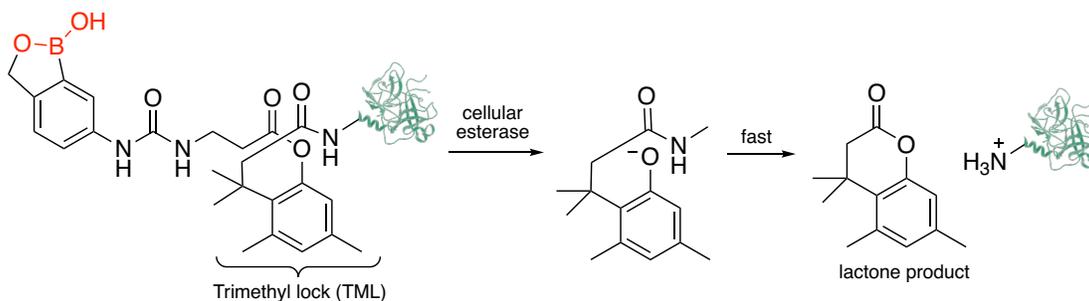
### 1.2.2 Boronic esters in bioorthogonal chemistry

Boronate formation has been used to identify, sense, deliver and purify proteins.<sup>67-71</sup> Most of these applications exploit boronate formation with biological polyols or saccharides units on glycoproteins.<sup>88</sup> For example, Raines and co-workers recently took advantage of the hemiboronic acid, benzoxaborole, to form boronate adducts with the glycocalyx on the cell membrane (Scheme 1-11).<sup>89,90</sup> They showed that the cellular delivery of a model protein, RNase A, was enhanced by the benzoxaborole-conjugated RNase A. Benzoxaborole is attached to the protein with an amide linkage and a trimethyl lock (TML) unit that readily hydrolyzes in the presence of cellular esterases, in turn releasing a lactone product (Scheme 1-11).<sup>90</sup> Cellular esterases are abundant inside human cells. Thus, RNase A is released from the benzoxaborole in a controlled manner only inside the cell and exerts its biological properties at full capacity.

### Boronic acid for the traceless delivery of proteins into cells



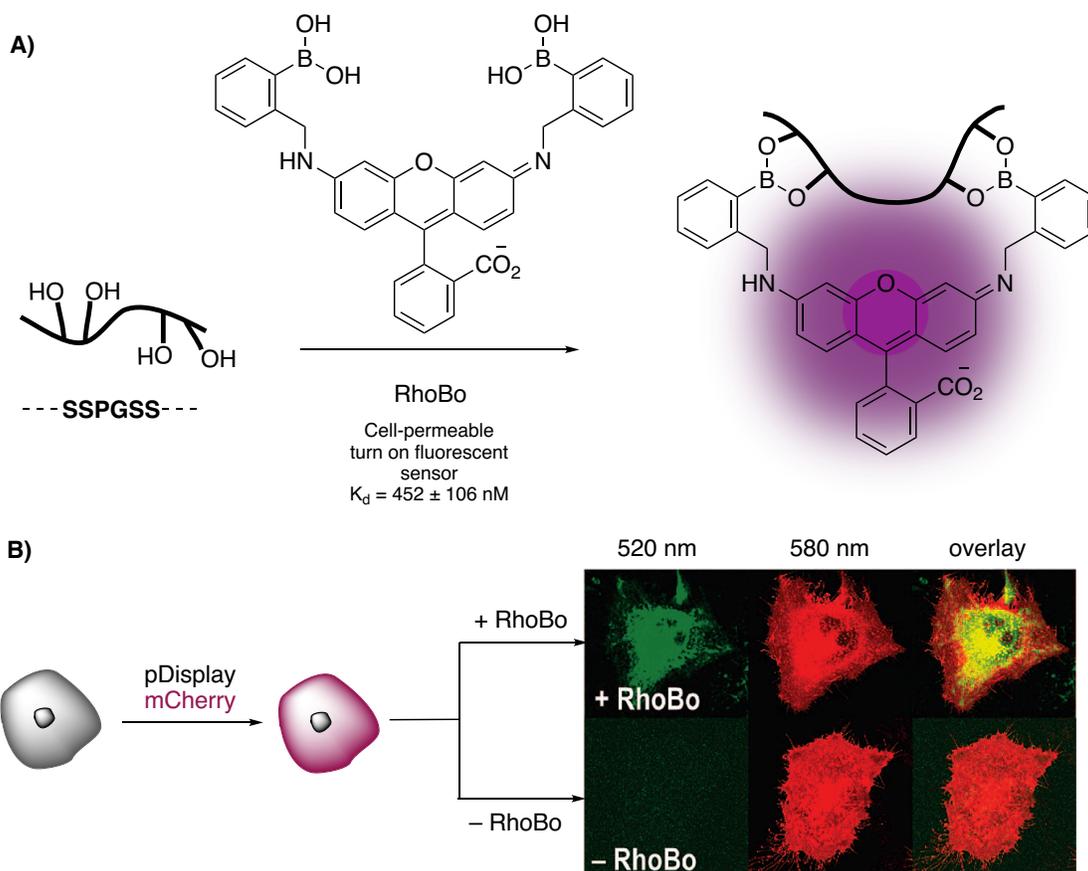
Mechanism to release the protein inside the cell:



**Scheme 1-11:** Boronic acid for the traceless delivery of proteins into cells.

These reports are, however, inapplicable in the site-selective labeling of proteins because boronate formation usually displays poor selectivity and stability. Few examples have been reported on bioorthogonal boronate formation. One early example, by Schepartz and co-workers, employed a nontoxic and redox-insensitive rhodamine-derived bisboronic acid (RhoBo) as a cell-permeable turn-on fluorescent sensor for proteins with an *N*-terminal SSPGSS tag (Scheme 1-12A).<sup>91</sup> To determine the optimal peptide tag, the authors designed and synthesized nine short peptides (8-9 amino acids) containing 2–4 serine/threonine residues, which were separated by 1-2 amino acids, such as glycine, proline and lysine. After the incubation of each peptide (various concentrations, 0–80  $\mu$ M) with RhoBo (17.1  $\mu$ M) in buffer at 37  $^{\circ}$ C, the fluorescence emission at 580 nm confirmed that an optimal tetraserine motif (SSPGSS) exhibited significantly tightest binding towards RhoBo ( $K_d = 452 \pm 106$  nM). The SSPGSS tag also exhibited higher affinity towards RhoBo compared to simple monosaccharides, such as sialic acid and *N*-acetylglucosamine. A low concentration of RhoBo ( $\sim 10$   $\mu$ M) was sufficient for selective labeling of an *N*-terminal SSPGSS tag, while a high amount of RhoBo ( $\sim 0.1$ – $1.0$  mM) was required for sialic acid and *N*-acetylglucosamine. Moreover, live pDisplay-mCherry transfected HeLa cells, which carry an SSPGSS tag, were imaged in the presence or absence of a RhoBo sensor (Scheme 1-12B). As expected, high fluorescent intensity was observed in the cytoplasmic region, while

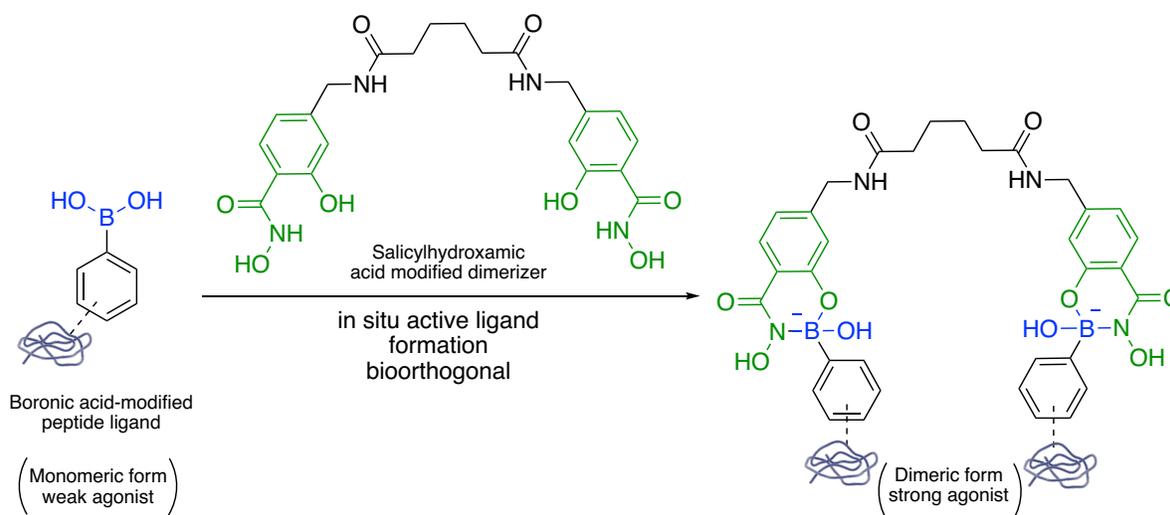
the nucleus and outer plasma membrane exhibited relatively lower fluorescence intensity. Although this example is applicable even in live mammalian cells, it might lack selectivity in *in vivo* studies since an SSPGSS sequence exists in more than 100 human proteins.<sup>91</sup> Furthermore, it targets a short peptide tag instead of a single amino acid as often desired in bioorthogonal chemistry.



**Scheme 1-12:** A) Cell-permeable turn-on fluorescent rhodamine-derived bis-boronic acid (RhoBo) for selective recognition of protein tetraserine motifs. B) Labeling SSPGSS motif-incorporated HeLa cells with RhoBo.

Another tight boronate system, which was initially developed for protein immobilization, forms between phenylboronic acid and salicylhydroxamic acid (Scheme 1-13).<sup>92</sup> Jaffrey and co-workers adapted this conjugation system into a bioorthogonal reaction, which can rapidly ( $7.01 \pm 2.04 \times 10^6 \text{ M}^{-2} \text{ s}^{-1}$ ) form a highly stable boronic ester ( $17,800 \text{ M}^{-1}$ ) at pH 7.4 (Scheme 1-13).<sup>93</sup> The reaction was found to be orthogonal and the resulting boronic ester was shown to be stable in the presence of typical biological nucleophiles such as cysteine, lysine, serine and also

Dulbecco's Modified Eagle's Medium (DMEM), which contains a high amount of glucose. The utility of the system was demonstrated on *in situ* assemblies of a dimeric c-Mpl (myeloproliferative leukemia virus) agonist between a monomeric boronic acid-modified peptide ligand and a salicylhydroxamic-modified dimerizer. The activity of the dimeric form was studied on BaF<sub>3</sub>-cMpl cells (which overexpress c-Mpl) and compared with the activity of the monomeric form. Dimeric potent agonists exhibited two-fold p42/44 MAPK phosphorylation activity compared to monomeric weak agonists. This rapid and highly stable bioorthogonal system allows the formation of high molecular weight compounds at a cellular level. Consequently, it might be helpful for transferring large peptides/proteins with low membrane permeability across the cell membrane and for improving proteins' pharmacokinetics. This new bioorthogonal reaction, however, is presumably limited to *in vitro* analysis due to the inconvenience in installing two boronic acid unit into the POI and the large size of the salicylhydroxamic dimer.

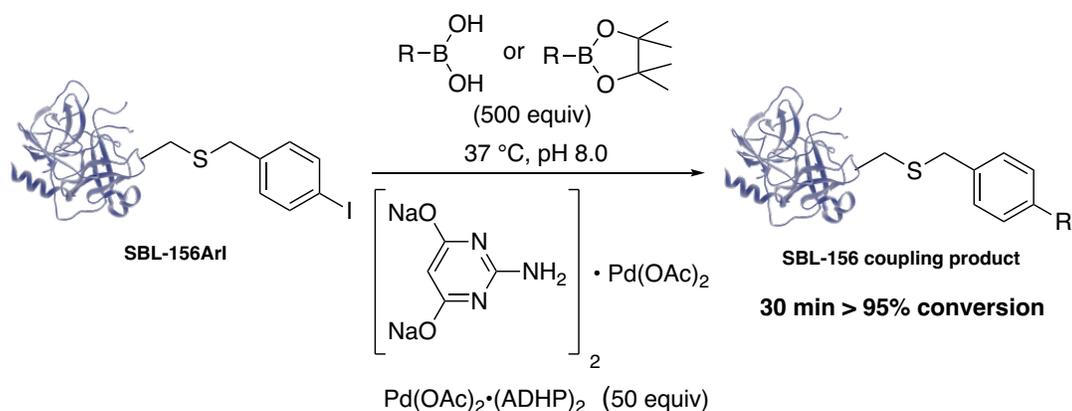


**Scheme 1-13:** Highly stable, rapid bioorthogonal boronic ester formation used to assemble a monomeric weak agonist into a dimeric strong agonist at a cellular level.

### 1.2.3 Boronic acid as a transient group in bioorthogonal chemistry

Bioorthogonal reactions involving (imino)boronates preserve the boronyl unit in the bioconjugation product, which may be in equilibrium with the starting reagents. Another approach uses boronic acid as a transient group to enable an irreversible transformation, which is usually a desirable feature in a bioorthogonal reaction. In this regard, the palladium-catalyzed  $sp^2$ - $sp^2$  coupling between aryl/alkenyl halide and arylboronic acids (Suzuki-Miyaura cross-

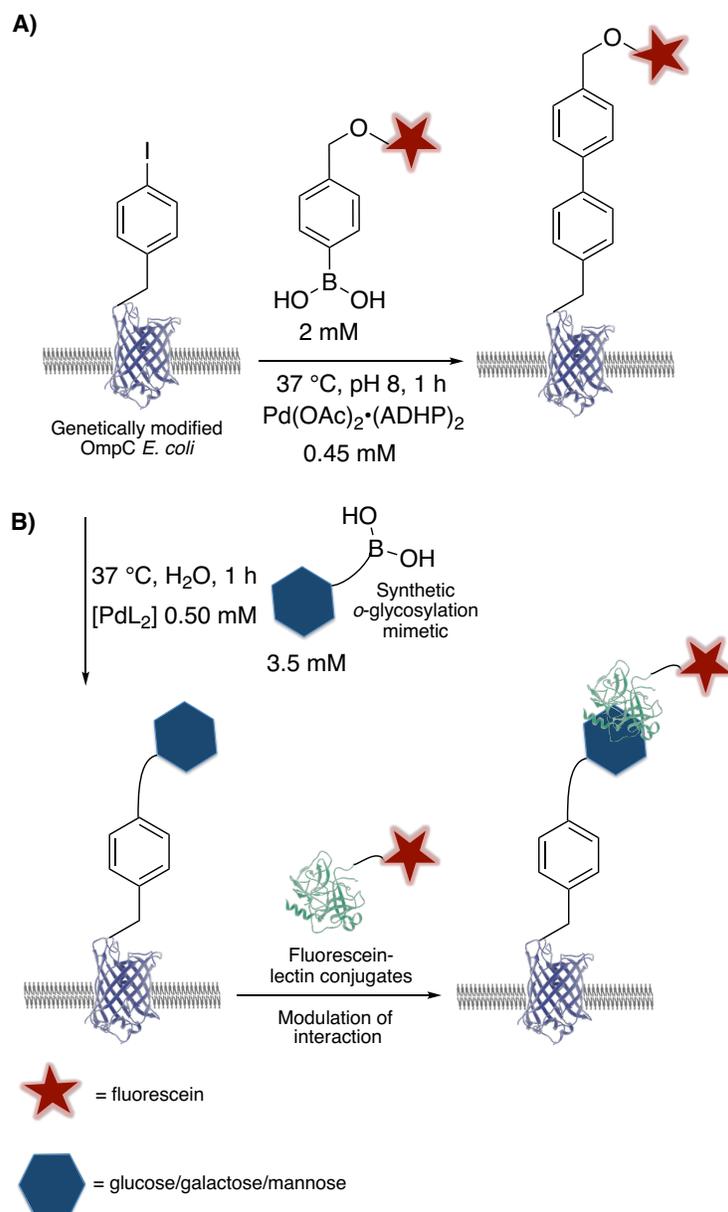
coupling) is one way to achieve a stable bioconjugate.<sup>63</sup> Turning this venerable reaction into a bioorthogonal reaction requires considerable modifications of the reaction conditions to address several issues, such as high temperature, the presence of organic solvent, tolerance of palladium towards the biological environment, and cytotoxicity of the catalyst. After Hamachi and co-workers<sup>94</sup> reported the first Suzuki-Miyaura coupling on synthetic peptides in aqueous solutions, Schultz and co-workers implemented the coupling on proteins with *p*-boronophenylalanine; however the reaction suffered from low yields (30%) and protein denaturation due to high temperatures (70 °C).<sup>76</sup> In 2009, Davis and co-workers ultimately solved this problem with the discovery of a water- and air-stable ligand, 2-amino-4,6-dihydropyrimidine (ADHP) that enabled an efficient Suzuki-Miyaura cross-coupling at 37 °C in water at pH 8 (Scheme 1-14).<sup>95</sup> A *p*-iodoarene was chemically installed on the single cysteine of a model protein, subtilisin Bacillus lentus (SBL) mutant S156C. A large amount of boronic acid (500 equiv) and catalyst (50 equiv) was loaded to achieve the desired product in a short time (Scheme 1-14).<sup>95</sup> This reaction functions with a variety of aryl and alkenyl boronic acids. Moreover, due to the hydrophilic nature of arylboronic acids, this system can be used to modify the POI with water-soluble boronic acids containing hydrophobic aryl moieties. Modification of POI with hydrophobic moieties is unlikely to be readily achievable with most of the existing bioorthogonal reactions.



**Scheme 1-14:** Application of Suzuki-Miyaura cross-coupling on a chemically modified model protein SBL-156ArI.

In another demonstration, Davis and co-workers achieved site-specific genetic incorporation of the unnatural amino acid, *p*-iodophenylalanine, into a model maltose-binding

protein (MBP-E13) using UAG stop-codon suppression.<sup>96</sup> After the successful installation of the synthetic amino acid, they performed Suzuki-Miyaura cross-coupling on MBP-E13. Full conversion was achieved within two hours, however with an ultra high loading of boronic acid (680 equiv) and Pd(OAc)<sub>2</sub>•(ADHP)<sub>2</sub> catalyst (50 equiv). Shortly after, Davis and co-workers applied this system to the labeling of *E. coli* cells.<sup>97</sup> Their target was the outer membrane protein C (OmpC) homotrimer, which is abundantly present on the surface of *E. coli* and plays a significant role during phage infection as a site for receptor binding. They identified and modified four accessible “tag” sites on OmpC without disturbing the cell, protein structure or function. These sites were replaced with *p*-iodophenylalanine through site-directed mutagenesis. Next, live cells were treated with a fluorescent boronic acid (1.6 mM) and Pd(OAc)<sub>2</sub>•(ADHP)<sub>2</sub> (0.35 mM) for one hour at 37 °C in pH 8 phosphate buffer, and the desired labeling was observed on the cell surface (Scheme 1-15A). The reaction was compatible with a high level of D-Glucose (3.2 mM), but its efficiency dropped significantly when the temperature was brought down to 30 °C. The boronic acid concentration (2 mM) and Pd levels (0.45 mM) were optimized and negligible toxicity (<3%) was reported with these concentrations. The toxic effects, however, were only considered in the context of cell surface studies. Lack of toxicity in the intracellular context remains to be confirmed.



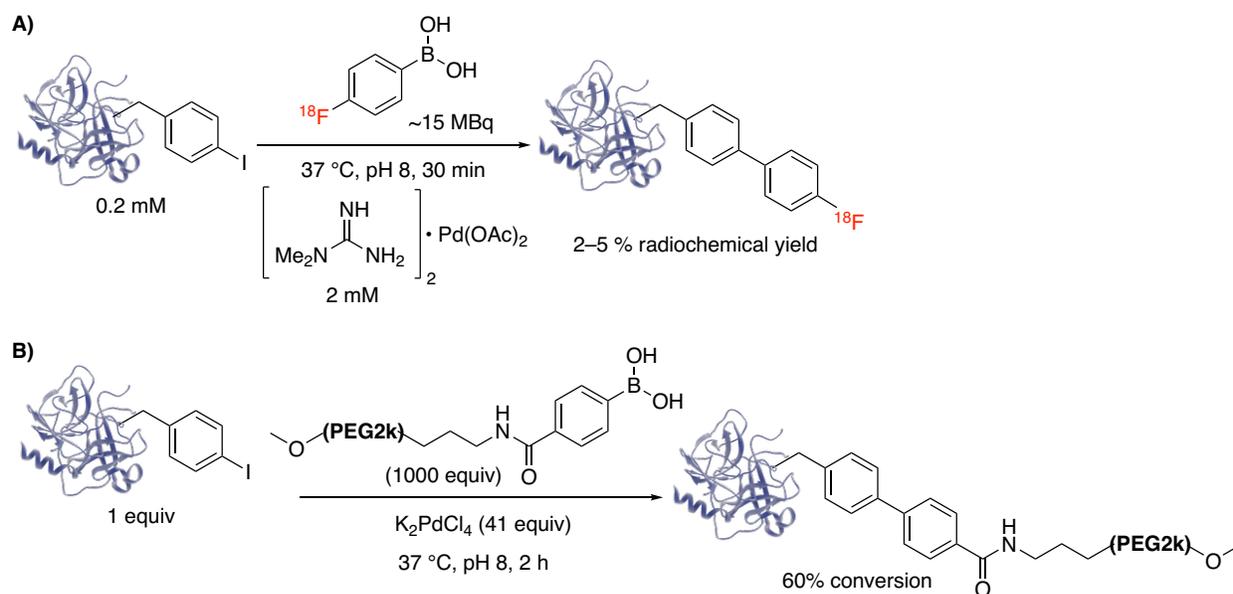
**Scheme 1-15:** A) *E. coli* surface labeling with Suzuki-Miyaura cross-coupling. B) *E. coli* surface labeling with novel glucose-, galactose- and mannose-based boronic acids followed by a modulation of interactions with fluorescein-lectin conjugates.

Suzuki-Miyaura cross-coupling was demonstrated on the site-selective surface labeling of *E. coli* with novel glucose-, galactose- and mannose-based boronic acids (Scheme 1-15B).<sup>98</sup> The system was envisioned to be applicable toward elucidating prokaryotic glycobiology since the newly conjugated carbohydrate may be modulated with biomolecular partners. As envisaged, the modulation of the cell surface with either mannose- or galactose-based boronic acids resulted in

a high interaction with the *Lens culinai*s agglutinin or *Griffonia simplicifolia* lectin I, respectively. In contrast, glucose-modified cells showed no interaction with these lectins. These interactions were detected through the selective binding of fluorescein-lectin conjugates on the modified cell surface. The affinity of these lectins towards carbohydrates was retained even in a complex biological context.

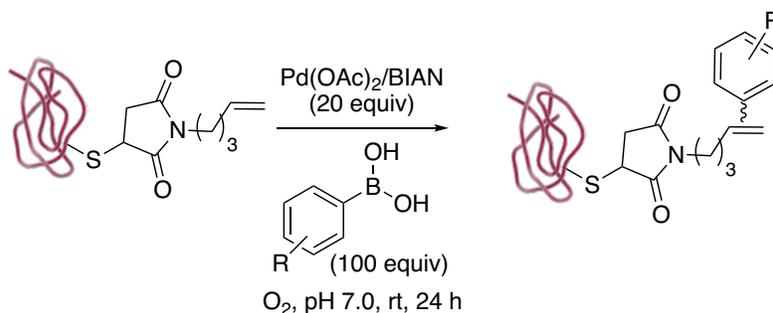
Suzuki-Miyaura cross-coupling was also applied for a challenging site-specific  $^{18}\text{F}$ -protein-labeling, which often needs high temperatures, lengthy multiple steps or organic solvent.<sup>99</sup> Because the nucleus  $^{18}\text{F}$  has a short half-life (109 min), it is best if the  $^{18}\text{F}$ -protein-labeling is achieved within a short time (<1 h). Furthermore, the labeling process must be efficient and rapid even at low concentrations of radioactive material. The above described Suzuki-Miyaura cross-coupling with  $\text{Pd}(\text{OAc})_2 \cdot (\text{ADHP})_2$  catalyst was inadequate because it requires an excess amount of boronic acid substrate. Fortunately, a water-soluble and non-toxic dimethylguanidine ligand enhanced the reaction dramatically, and a 75% conversion was achieved with a 1 mM solution of the boronic acid substrate (Scheme 1-16A) while the previously reported ADHP ligand did not provide the desired coupling product. A radiosynthesis of [ $^{18}\text{F}$ ]4-fluorophenylboronic acid (~15 megabecquerel, MBq) in two steps (40 min) was feasible, and a direct  $^{18}\text{F}$ -protein labeling (0.2 mM) afforded a low 2-5 % radiochemical yield in 30 min (Scheme 1-16A). Despite its seemingly low yield on a model protein, this system was efficient as it could couple small molecules and peptides (0.1 – 0.2 mM) in radiochemical yields of up to 87%.

Another application of the Suzuki-Miyaura cross-coupling is in protein PEGylation,<sup>100</sup> a process that enhances the stability and pharmacokinetics of protein drugs. PEGylation should be site-selective in order to preserve the biological activity of the corresponding protein drug. Davis and co-workers reported the first example of a self-liganded palladium-catalyzed coupling on a model protein. The PEG-boronic acid derivative (1000 equiv) and a single Pd catalyst (41 equiv) without a ligand afforded a 60% conversion with a genetically modified model protein, Npb-69pIPhe (1.0 equiv) (Scheme 1-16B). In other words, the presence of the PEG group removes the need to use a ligand. Thus, this system encourages further development of a unified ligand-reagent-metal system, which is likely more suitable for *in vivo* applications.



**Scheme 1-16:** A) Site-selective  $^{18}\text{F}$ -protein-labeling. B) Site-selective PEGylation with Suzuki-Miyaura cross-coupling.

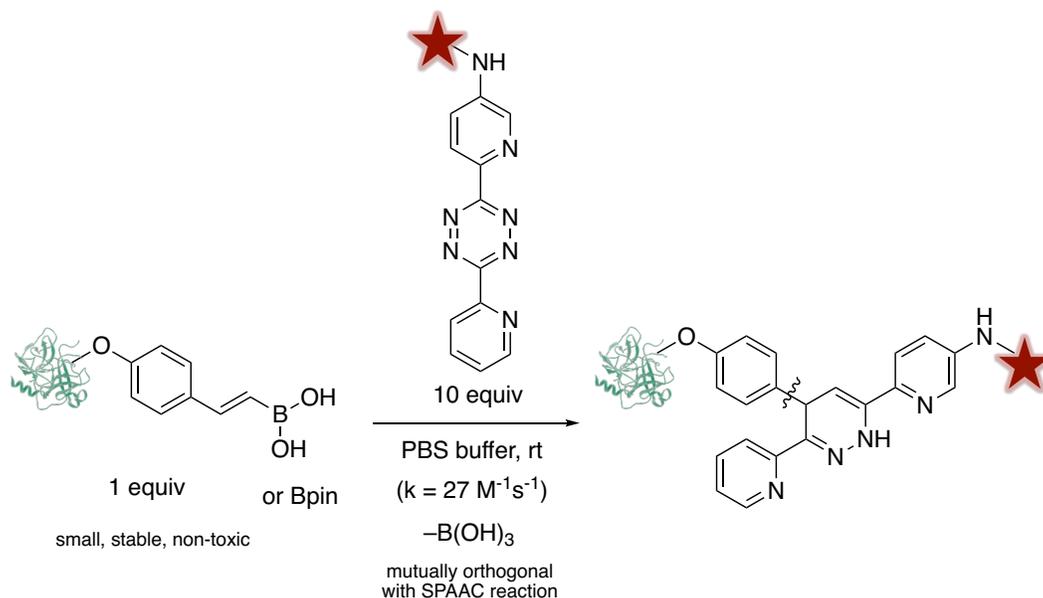
Apart from Suzuki-Miyaura cross-coupling, Dekker and co-workers adapted the oxidative Heck reaction,<sup>101</sup> a process occurring under aerobic conditions between alkenes and arylboronic acids, into a bioorthogonal reaction (Scheme 1-17). The  $\text{Pd}(\text{OAc})_2$ /ligand bis(aryl)acenaphthequinonediimine (BIAN) catalyst had a low water solubility, so a mixture of buffer and DMF (6:1) was needed. A model protein-bound alkene fully converted into its arylated product with the optimal use of a  $\text{Pd}(\text{OAc})_2$ /BIAN catalyst (20 equiv) and arylboronic acid (100 equiv) in 24 hours. The reaction proceeded with *cis* alkenes but not with *trans* alkenes. This selectivity is helpful to differentiate *cis*-unsaturated fatty acids from *trans*-unsaturated fatty acids in biological systems, but at the same time, it might lack site-selectivity for *in vivo* applications where an alkene functionalized POI is targeted. Finally, a model protein-bound alkene was ligated efficiently and site-selectively with the fluorescent 3-(dansylamino)phenylboronic acid even in the soluble fraction of lysates from RAW264.7 enriched macrophages.



**Scheme 1-17:** Oxidative Heck reaction as a bioorthogonal reaction.

Boronic acids also play a role in the inverse electron-demand Diels Alder reaction (IEDDA/Carbonyl-Lindsey reaction). Bonger and co-workers recently reported a bioorthogonal reaction involving a boronic acid unit, which presence on a non-strained alkene surprisingly enhanced the reactivity of the unsaturation towards 3,6-dipyridyl-s-tetrazines (Scheme 1-18).<sup>102</sup> The authors focused on alkenylboronic acids as they are water soluble, small, readily synthesized and non-toxic. It was found that (*E*)-*p*-methoxyphenylvinylboronic acid ( $\sim 27 \text{ M}^{-1}\text{s}^{-1}$ ) reacts with 3,6-dipyridyl-s-tetrazines faster than  $\beta$ -methoxystyrene ( $\sim <0.0025 \text{ M}^{-1}\text{s}^{-1}$ ), *p*-methoxystyrene ( $\sim 0.13 \text{ M}^{-1}\text{s}^{-1}$ ) and commonly used norbornene ( $\sim 2.2 \text{ M}^{-1}\text{s}^{-1}$ ). However, (*E*)-*p*-methoxyphenylethyleneboronic reacted slowly with other tetrazines such as 3-phenyl-s-tetrazine and 3-phenyl-6-methyl-s-tetrazine. Although a putative boronate anion, as an electron-donating group, was considered as the cause for the large rate improvement, other factors such as coordination between boronic acid and a nitrogen atom on the dipyridyl residue of tetrazines cannot be ruled out. The conjugation product lacked a boronic acid unit, which can be explained by proteolytic deboronation after the IEDDA cycloaddition. After the cycloaddition, nitrogen gas was ejected through a retro-Diels-Alder reaction. Alkenylboronic acids and tetrazine were found to be stable and biocompatible in cell lysates. Moreover, both alkenylboronic acid and boric acid, the side product, were confirmed to be non-toxic with up to 100  $\mu\text{M}$  concentrations. In the *in vitro* application, the desired product was observed within five minutes on a chemically modified model protein, human serum albumin-vinylboronic acid (50  $\mu\text{M}$ ) modified with a fluorophore-conjugated tetrazine (500  $\mu\text{M}$ ). Fortunately, the reaction also worked with pinacol-protected alkenylboronates, which were easier to synthesize than free alkenylboronic acids. The reaction was mutually orthogonal with the SPAAC reaction but incompatible with the CuAAC reaction. Although its rate constant was inferior to the rates provided by the cycloaddition of tetrazines

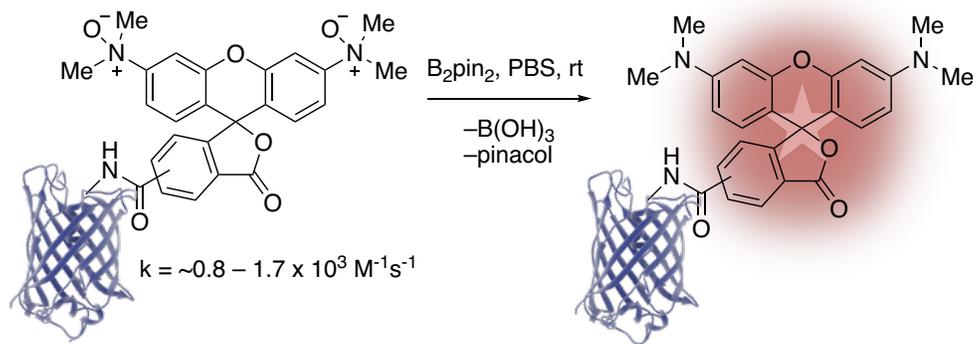
with trans-cyclooctene, the small size of alkenylboronic acids along with their water solubility, stability, mutual orthogonality with SPAAC, and easy accessibility might render it more useful in bioorthogonal chemistry.



**Scheme 1-18:** The use of alkenylboronic acid IEDDA/Carbonyl-Lindsey reaction to enhance the reactivity of non-strained alkenes.

The field of bioorthogonal chemistry has primarily been envisioned, designed and implemented as a toolbox of bond-forming processes. A small number of reports have examined bioorthogonal bond-cleaving or uncaging reactions.<sup>103-105</sup> Bertozzi and co-workers recently discovered a bond-cleaving transformation consisting of a boron reagent and a biocompatible *N*-oxide, which are present in some organisms such as deep-sea fish.<sup>106</sup> Fortunately, *N*-oxides are inert and biocompatible with their congener biomolecules. Moreover, they do not exist in most other organisms including humans. Trimethylamine *N*-oxide can undergo a functional-group tolerant and high yielding hydroxydeboration reaction with alkyl boranes releasing triethylamine and boric acid as products. Initial kinetic studies of the reaction between trimethylamine *N*-oxide and *p*-nitrophenylboronic acid, which released trimethylamine, boric acid and *p*-nitrophenol, showed a very low second-order rate constant of  $\sim 2.83 \times 10^{-6} \text{ M}^{-1}\text{s}^{-1}$ . The reaction rate was enhanced three-fold ( $\sim 1.28 \times 10^{-3} \text{ M}^{-1}\text{s}^{-1}$ ) by the fluorogenic *N,N,N*-dialkylaryl *N*-oxide. This rate was still insufficient; therefore the focus shifted from C–B to a much weaker B–B bond

cleavage, which eventually formed a strong B–O bond. As expected, kinetics improved five-fold ( $\sim 8.05 \times 10^2 \text{ M}^{-1}\text{s}^{-1}$ ) by using *N,N,N*-dialkylaryl *N*-oxide and bis(pinacolato)diboron ( $\text{B}_2\text{pin}_2$ ) (Scheme 1-19). A much faster rate constant ( $\sim 1.71 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ ) was detected with the designated HaloTag linker-bound profluorophore for cell-labeling studies. The reaction was efficient and robust in mammalian (Jurkat) cell lysates with five equivalents of  $\text{B}_2\text{pin}_2$  (5  $\mu\text{M}$ ) and *N,N,N*-dialkylaryl *N*-oxide (1  $\mu\text{M}$ ) for 30 minutes. However, this system was not applicable for *E. coli* since it expresses *N*-oxide reductase, which can compete with  $\text{B}_2\text{pin}_2$ . While no toxicity was observed for HEK293T, Hela and MEF cells up to 1 mM of  $\text{B}_2\text{pin}_2$ , low toxicity was recorded for MEF cells exposed to  $>1 \text{ mM}$   $\text{B}_2\text{pin}_2$ . This system was also found to be mutually orthogonal with commonly applied bioorthogonal reactions such as oxime ligation, SPAAC, and tetrazine–cyclopropene ligation. Transiently transfected HEK293T cells with a cytosolic GFP-HaloTag fusion construct were incubated with a HaloTag linker-bound profluorophore (100  $\mu\text{M}$ ). The cells were then treated with  $\text{B}_2\text{pin}_2$  (0, 10, 100 and 1000  $\mu\text{M}$ ) for 45 minutes. Only 10  $\mu\text{M}$  concentration of  $\text{B}_2\text{pin}_2$  was sufficient to observe vivid cell imaging, thus confirming the cell permeability of  $\text{B}_2\text{pin}_2$ .



**Scheme 1-19:** Fluorogenic bond-cleaving transformation between *N*-oxide and boron reagent.

### 1.3 Summary and future directions

Interest in the use of boronic acids in bioorthogonal conjugation has grown substantially in the past 5 years. One variant, iminoboronate chemistry stands as an attractive alternative for aldehyde/ketone condensation because its rate constant ( $\sim 10^3 \text{ M}^{-1}\text{s}^{-1}$ ) is high, it requires a low/stoichiometric amount of substrates (high  $K_{\text{eq}}$ ), it functions in neutral pH, and it results in

potentially stable diazaborine or thiazolidine boronate adducts (Table 1-2A-D). The 2-ABBA reagent is conceivably a better choice than the 2-FBBA, which is not truly a bioorthogonal reagent because of its side reactivity towards *N*-terminal cysteines (Table 1-2C). To implement iminoboronate chemistry in live cells or animals, it would be useful to employ methods allowing the metabolic or genetic incorporation of 2-ABBA. Further studies with 2-ABBA and stable semicarbazide derivatives for improving conversion, finding a cell-permeable semicarbazide derivative and designing its fluorogenic version are significant objectives to advance this chemistry for *in vivo* applications. This chemistry, however, may lack compatibility with protein-bound electrophiles, which are abundant due to pyruvoyl (ketone) and glyoxylyl (aldehyde) electrophilic post-translational modifications in human cells.<sup>39,40</sup> Yet, iminoboronate chemistry operates with a lower amount of reagents and in neutral pH compared to aldehyde/ketone condensation; therefore considering these factors, hydrazine/hydroxylamine/semicarbazide reagents may be relatively inert to these protein-bound electrophiles.

On the other hand, fast and reversible but highly stable boronates might be advantageous in cases where linkages are cleaved on demand (Table 1-2E-F). In particular, Rho-Bo chemistry is applicable for live cells or animals because of its fluorogenic properties (Table 1-2F). This class of boronic acid-containing bioorthogonal reactions is most likely to be mutually orthogonal with other bioconjugation chemistries like SPAAC, IEDDA, Staudinger ligation and aldehyde/ketone condensation. However, further experimental evidence is essential.

To achieve more stable bioorthogonal reaction products, transient boronic acids have been applied in Suzuki-Miyaura cross-coupling, oxidative Heck reaction and IEDDA (Table 1-2G-I). To this end, effective methodologies to incorporate *p*-iodophenylalanine into eukaryotic systems<sup>107</sup> and to use amber codon suppression<sup>76</sup> for the site-specific incorporation of *p*-boronophenylalanine into proteins directly in *E. coli*, are crucial for the *in vivo* labeling of boronic acid- or iodo-containing proteins. An unusual bond-cleaving reaction that uses the popular diboronyl B<sub>2</sub>pin<sub>2</sub> reagent instead of boronic acid provides a fast, irreversible and fluorogenic bioorthogonal reaction (Table 1-2J). More importantly, most of the bioorthogonal reactions described in this chapter involve benign, and easily accessible synthetic probes compared to CuAAC, SPAAC and IEDDA. The rate constants for these reactions reach speeds up to 10<sup>3</sup> M<sup>-1</sup>s<sup>-1</sup>. Fortunately, both bond-cleaving reactions and alkenyl boronic acid-IEDDA were confirmed to be mutually orthogonal with SPAAC or aldehyde/ketone condensation

reactions. Although Suzuki-Miyaura cross-coupling and oxidative Heck reactions are also likely to be orthogonal with SPAAC and IEDDA, they must be further evaluated to prove their compatibility for dual labeling applications.

It is important to mention that since 2010 some arylboronic acids and B<sub>2</sub>pin<sub>2</sub> have been labeled as a novel class of bacterial mutagens due to their positive response in the Ames test<sup>108</sup> though only a few arylboronic acids were positive at high concentrations (>1 mM) in mammalian *in vitro* genotoxicity tests.<sup>109</sup> Furthermore, the only arylboronic acid that was tested *in vivo* did not display any genotoxic properties.<sup>110</sup> Although these results are reassuring, further evidence is needed to show that arylboronic acids, especially those related to the reported bioorthogonal reactions, are benign.

Ultimately, despite their promising properties, additional studies of boronic acid-bioorthogonal reactions must be performed to find more stable conjugates and reagents, to design fluorogenic ligations, to improve cell permeability of reagents, to develop methods for metabolic and genetic encoding of bioorthogonal groups, and to assess the genotoxicity of boronic acids. It is, however, exciting to witness all these advances and their applications, which can potentially shed light on complicated biological processes and the roles of proteins and, it is hoped, pave the way to cure dreadful human diseases.

**Table 1-2:** Summary of the applications of boronic acids as bioorthogonal reagents in the service of site-selective protein labeling.

Reaction type	Reaction scheme	~ Rate constant ( $M^{-1}s^{-1}$ )	Comments
A) Iminoboronate formation	<p>Reversible conjugation</p> <p>Irreversible conjugation</p>	$10^3$	<ul style="list-style-type: none"> <li>*Reversible but high <math>K_{eq}</math></li> <li>*Fast</li> <li>*Compatible with sucrose, lysozyme, GSH and human serum</li> <li>*Irreversible with 2-FBBA and phenylhydrazine but unstable and cytotoxic phenylhydrazine</li> <li>*Reactive 2-FBBA towards <i>N</i>-terminal cysteine</li> </ul>
B) Irreversible diazaborine formation		$\sim 10^3$	<ul style="list-style-type: none"> <li>*Irreversible, fast, stable and benign semicarbazide</li> <li>*Moderate conversion (60%)</li> <li>*Bacterial labeling</li> </ul>
C) Thiazolidine boronate formation	<p>Reversible under slightly acidic conditions</p>	$\sim 10^3$	<ul style="list-style-type: none"> <li>*Fast, stable</li> <li>*<i>N</i>-terminal cysteine selective at neutral pH</li> <li>*Reversible under slightly acidic conditions</li> <li>*Likely limited to <i>in vitro</i></li> </ul>
D) Boronic acid mediated couplings of catechol and <i>N</i> -hydroxylamines.		NA	<ul style="list-style-type: none"> <li>*Irreversible</li> <li>*Likely limited to <i>in vitro</i></li> </ul>

Reaction type	Reaction scheme	~ Rate constant (M <sup>-1</sup> s <sup>-1</sup> )	Comments
E) Highly stable boronate ester synthesis (HiSBE)		~10 <sup>7</sup> M <sup>-2</sup> s <sup>-1</sup>	<ul style="list-style-type: none"> <li>*Fast, reversible but high Keq</li> <li>*Compatible with biological nucleophiles such as cysteine, lysine, serine and DMEM that contains high amount of glucose</li> <li>*Likely limited to <i>in vitro</i></li> </ul>
F) Fluorogenic RhoBo boronate formation		NA	<ul style="list-style-type: none"> <li>*Highly selective, benign and cell permeable RhoBo</li> <li>*Rate not reported</li> <li>*Tight binding, low RhoBo concentration sufficient</li> <li>*Fluorescent turn-on</li> <li>*N-terminal SSPGSS tag on and within living cells</li> <li>*Intracellular cell labeling</li> </ul>
G) Pd-catalyzed cross-coupling		NA	<ul style="list-style-type: none"> <li>*Use of large excess of palladium catalyst and boronic acid</li> <li>*Rate not reported</li> <li>*Bacterial cell surface labeling and modification</li> <li>*Nontoxic catalyst at least on cell surface labeling</li> <li>*Installation of hydrophobic moieties onto protein</li> </ul>
H) Aqueous oxidative Heck reaction		NA	<ul style="list-style-type: none"> <li>*Use of large excess palladium catalyst and boronic acid</li> <li>*Rate not reported</li> <li>*Compatible with aerobic conditions</li> <li>*Small alkene moieties</li> <li>*Reactive towards <i>cis</i>-unsaturated fatty acids, possible interference for <i>in vivo</i> applications</li> <li>*Installation of hydrophobic moieties onto protein</li> </ul>

Reaction type	Reaction scheme	~ Rate constant (M <sup>-1</sup> s <sup>-1</sup> )	Comments
I) IEDDA – (Alkenylboronic acids and 3,6-dipyridyl-s-tetrazines)		27	<ul style="list-style-type: none"> <li>*Synthetically accessible, stable, small alkenylboronic acids or boronates</li> <li>*Reactive towards only 3,6-dipyridyl-s-tetrazines</li> <li>*Moderate rate</li> <li>*Mutually orthogonal with SPAAC</li> </ul>
J) Bond-cleaving reaction (N-Oxide and B <sub>2</sub> pin <sub>2</sub> )		~10 <sup>3</sup>	<ul style="list-style-type: none"> <li>*A bond-cleaving transformation</li> <li>*Very fast</li> <li>*Incompatible with <i>E. coli</i> but compatible with mammalian cells HEK293T, HeLa and MEF cells</li> <li>*Benign and cell permeable B<sub>2</sub>pin<sub>2</sub></li> <li>*Mutual orthogonality with aminoxy-aldehyde condensation, SPAAC and the tetrazine-cyclopropene ligation</li> <li>*Intracellular cell imaging</li> </ul>

## 1.4 Thesis objectives

The goal of the work described in this thesis is to develop novel bioorthogonal tools in the service of site-selective protein labeling. Boronic ester formation was targeted due to the synthetic accessibility of boronic acids and diols, their low toxicity, and fast kinetics in catalyst-free conditions. (–)-Nopoldiol, a pinanediol derivative, was chosen as the 1,2-*cis* diol because it is well established that hindered and preorganized vicinal diols afford some of the most hydrolytically robust boronic esters.<sup>111,112</sup> Chapter 2 describes efforts made in developing and optimizing a novel boronate bioorthogonal reaction system, which enables a fast and tight ligation. The compatibility of the optimized system is assessed in the presence of biocompetitors such as glucose, fructose and catechol. Also, its bioorthogonality and efficiency are studied on model proteins, BSA and thioredoxin (Trx).

Although the optimized nopol-boronate system is fast and tight in aqueous medium, it is inherently reversible. This can likely limit its application in live cell imaging since the boronate adduct might undergo unwanted cleavage during the washing operations for excess reagents. In this regard, Chapter 3 of this thesis introduces a novel nopol-boronate system with improved stability. The optimal synergic system is based on two bifunctional reagents, a thiosemicarbazide-functionalized nopoldiol and an 2-acetylarylboronic acid to provide a fast and irreversible boronate/thiosemicarbazone system. The stability of boronate/thiosemicarbazone conjugates (50  $\mu\text{M}$ ) is assessed by four separate methods; 1) diluting the adduct solution from 50  $\mu\text{M}$  to 10  $\mu\text{M}$ , 2) introducing an excess amount of reactive 2-FBBA competitor (100  $\mu\text{M}$ ) to the existing conjugate, 3) acidifying the media to pH 3 and 4) basifying the media to pH 9. Moreover, the system's compatibility with biological polyols and protein bound electrophiles, and live cell labeling studies are described in detail in this chapter.

The robust and efficient bioorthogonal systems described above contain two abiotic components that are orthogonal to other functional groups on the POI. However, installation of one bioorthogonal handle into the POI might be challenging and would likely require a site-directed unnatural amino acid mutagenesis, which is a difficult and tedious technique. Thus, it would be more advantageous if one of the two components, the polyol, could be expressed as a short, non-invasive peptide tag. In this regard, the thiosemicarbazide-functionalized nopoldiol was aimed to be replaced with a short and reactive peptide tag, which could also exhibit a high affinity toward 2-acetylarylboronic acid *via* both imine and boronate formation. It was envisioned that 2-acetylarylboronic acid could form a tight complex with a short peptide embedding as many as three hydroxyl side chains (serine, threonine) and a lysine residue or a terminal amine. To discover the optimum peptide, we collaborated with the Laboratory of Prof. Ratmir Derda, who is an expert of phage display selections.<sup>113</sup> Chapter 4 presents the work performed towards the selection of peptide tags *via* phage display technology with a serine-terminated library of over  $10^8$  heptapeptides and the synthesis and conjugation of selected peptides with 2-acetylarylboronic acids.

## 1.5 References

1. Chen, X.; Wu, Y.-W. *Org. Biomol. Chem.* **2016**, *14*, 5417–5439.
2. Boutureira, O.; Bernardes, G. J. L. *Chem. Rev.* **2015**, *115*, 2174–2195.

3. Spicer, C. D.; Davis, B. G. *Nat. Commun.* **2014**, *5*, 4740.
4. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
5. Ramil, C. P.; Lin, Q. *Chem. Commun. (Camb.)* **2013**, *49*, 11007–11022.
6. Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764–4806.
7. a) Zhang, G.; Zheng, S.; Liu, H.; Chen, P. R. *Chem. Soc. Rev.* **2015**, *44*, 3405–3417. b) Hang, H. C.; Wilson, J. P.; Charron, G. *Acc. Chem. Res.* **2011**, *44*, 699–708.
8. a) Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F. P. J. T.; van Hest, J. C. M.; Lefeber, D. J.; Friedl, P.; van Delft, F. L. *Angew. Chem. Int. Ed.* **2010**, *49*, 9422–9425. b) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. *Nat. Chem.* **2012**, *4*, 2012, 298–304. c) Peng, T.; Hang, H. C. *J. Am. Chem. Soc.* **2016**, *138*, 14423–14433.
9. a) Baskin, J. M.; Bertozzi, C. R. *QSAR Comb. Sci.* **2007**, *26*, 1211–1219. b) Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. *Nat. Chem.* **2016**, *8*, 103–113.
10. a) Yang, Y-Y.; Ascano, J. M.; Hang, H. C. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641. b) Heal, W. P.; Jovanovic, B.; Bessin, S.; Wright, M. H.; Magee, A. I.; Tate, E. W. *Chem. Commun.* **2011**, *47*, 4081–4083.
11. a) Beck, A.; Goetsch, L.; Dumontet, C.; Corvaia, N. *Nat. Rev. Drug Discov.* **2017**, *16*, 315–337. b) van Vught, R.; Pieters, R. J.; Breukink, E. *Comput. Struct. Biotechnol. J.* **2014**, *9*, 1–13.
12. a) Axupa, J. Y.; Bajjuri, K. M.; Ritland, M.; Hutchins, B. M.; Kim, C. H.; Kazane, S. A.; Halder, R.; Forsyth, J. S.; Santidrian, A. F.; Stafin, K.; Lu, Y.; Tran, H.; Seller, A. J.; Biroc, S. L.; Szydlak, A.; Pinkstaff, J. K.; Tian, F.; Sinha, S. C.; Felding-Habermann, B.; Smider, V. V.; Schultz, P. G. *PNAS* **2012**, *109*, 16101–16106. b) Zimmerman, E. S.; Heibeck, T. H.; Gill, A.; Li, X.; Murray, C. J.; Madlansacay, M. R.; Tran, C.; Uter, N. T.; Yin, G.; Rivers, P. J.; Yam, A. Y.; Wang, W. D.; Steiner, A. R.; Bajad, S. U.; Penta, K.; Yang, W.; Hallam, T. J.; Thanos, C. D.; Sato, A. K. *Bioconjugate Chem.* **2014**, *25*, 351–361.
13. Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213–227.
14. Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chem. Asian J.* **2009**, *4*, 630–640.
15. Lotze, J.; Reinhardt, U.; Seitz, O.; Beck-Sickinger, A. G. *Mol. BioSyst.* **2016**, *12*, 1731–

- 1745.
16. Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802–805.
  17. Heim, R.; Cubitt, A. B.; Tsien, R. Y. *Nature* **1995**, *373*, 663–664.
  18. Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. *Physiol. Rev.* **2010**, *90*, 1103–1163.
  19. Rizzo, M. A.; Davidson, M. W.; Piston, D. W. *Cold Spring Harb. Protoc.* **2009**, *12*, 1–22.
  20. Jing, C.; Cornish, V. W. *Acc. Chem. Res.* **2011**, *44*, 784–792.
  21. Adumeau, P.; Sharma, S. K.; Brent, C.; Zeglis, B. M. *Mol Imaging Biol.* **2016**, *18*, 153–165.
  22. George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.* **2004**, *126*, 8896–8897.
  23. Chen, I.; Howarth, M.; Lin, W. Y.; Ting, A. Y. *Nature Methods* **2005**, *2*, 99–104.
  24. Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269–272.
  25. Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *J. Am. Chem. Soc.* **2002**, *124*, 6063–6076.
  26. Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. *J. Am. Chem. Soc.* **2007**, *129*, 8672–8673.
  27. Jung, D.; Min, K.; Jung, J.; Jang, W.; Kwon, Y. *Mol. BioSyst.* **2013**, *9*, 862–872.
  28. a) Tanaka, F.; Fuller, R.; Asawapornmongkol, L.; Warsinke, A.; Gobuty, S. *Bioconjugate Chem.* **2007**, *18*, 1318–1324. b) Eldridge, G. M.; Weiss, G. A. *Bioconjugate Chem.* **2011**, *22*, 2143–2153. c) Ramil, C. P.; An, P.; Yu, Z.; Lin, Q. *J. Am. Chem. Soc.* **2016**, *138*, 5499–5502. d) Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. *Nat. Chem.* **2016**, *8*, 120–128.
  29. Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16–20.
  30. Zheng, M.; Zheng, L.; Zhang, P.; Li, J.; Zhang, Y. *Molecules* **2015**, *20*, 3190–3205.
  31. a) Fan, X.; Li, J.; Chen, P. R. *Natl. Sci. Rev.* **2017**, *0*, 1–3. b) Chen, Y.-X.; Triola, G.; Waldmann, H. *Acc. Chem. Res.* **2011**, *44*, 762–773.
  32. Dozier, J. K.; Distefano, M. D. *Int. J. Mol. Sci.* **2015**, *16*, 25831–25864.
  33. Gong, Y.; Pan, L. *Tetrahedron Lett.* **2015**, *56*, 2123–2132.
  34. Kim, C. H.; Axup, J. Y.; Schultz, P. G. *Curr. Opin. Chem. Biol.* **2013**, *17*, 412–419.

35. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
36. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125–1128.
37. Rashidian, M.; Mahmoodi, M. M.; Shah, R.; Dozier, J. K.; Wagner, C. R.; Distefano, M. D. *Bioconjugate Chem.* **2013**, *24*, 333–342.
38. Kool, E. T.; Park, D.-H.; Crisalli, P. *J. Am. Chem. Soc.* **2013**, *135*, 17663–17666.
39. Prescher, J. A., Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21.
40. Matthews, M. L.; He, L.; Horning, B. D.; Olson, E. J.; Correia, B. E.; Yates, J. R.; Dawson, P. E.; Cravatt, B. F. *Nat. Chem.* **2017**, *9*, 234–243.
41. van Berkel, S. S.; van Eldijk, M. B.; van Hest, J. C. M. *Angew. Chem. Int. Ed.* **2011**, *50*, 8806–8827.
42. Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. *Org. Lett.* **2000**, *2*, 2141–2143.
43. Shah, L.; Laughlin, S. T.; Carrico, I. S. *J. Am. Chem. Soc.* **2016**, *138*, 5186–5189.
44. Chang, P. V.; Prescher, J. A.; Sletten, E. M.; Baskin, J. M.; Miller, I. A.; Agard, N. J.; Lo, A.; Bertozzi, C. R. *PNAS* **2010**, *107*, 1821–1826.
45. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644–648.
46. Yang, M.; Li, J.; Chen, P. R. *Chem. Soc. Rev.* **2014**, *43*, 6511–6526.
47. Aioub, A. G.; Dahora, L.; Gamble, K.; Finn, M. G. *Bioconjugate Chem.* **2017**, *28*, 1693–1701.
48. Uttamapinant, C.; Tangpeerachaikul, A.; Grecian, S.; Clarke, S.; Singh, U.; Slade, P.; Gee, K. R.; Ting, A. Y. *Angew. Chem. Int. Ed.* **2012**, *51*, 5852–5856.
49. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
50. Dieterich, D. C.; Hodas, J. J. L.; Gouzer, G.; Shadrin, I. Y.; Ngo, J. T.; Triller, A.; Tirrell, D. A.; Schuman, E. M. *Nat. Neurosci.* **2010**, *13*, 897–905.
51. Beatty, K. E.; Fisk, J. D.; Smart, B. P.; Lu, Y. Y.; Szychowski, J.; Hangauer, M. J.; Baskin, J. M.; Bertozzi, C. R.; Tirrell, D. A. *ChemBioChem* **2010**, *11*, 2092–2095.
52. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotechnol.* **2007**, *25*, 1483–1487.
53. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *PNAS* **2007**, *104*, 16793–16797.
54. Conte, Lo, M.; Staderini, S.; Marra, A.; Sanchez-Navarro, M.; Davis, B. G.; Dondoni, A.

- Chem. Commun.* **2011**, *47*, 11086–11088.
55. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
  56. Wu, H.; Devaraj, N. K. *Top. Curr. Chem.* **2016**, *374*, 1–22.
  57. Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R.; Weissleder, R. *Angew. Chem. Int. Ed.* **2010**, *49*, 2869–2872.
  58. Carlson, J. C. T.; Meimetis, L. G.; Hilderbrand, S. A.; Weissleder, R. *Angew. Chem. Int. Ed.* **2013**, *52*, 6917–6920.
  59. Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *J. Am. Chem. Soc.* **2012**, *134*, 10317–10320.
  60. Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y. *J. Am. Chem. Soc.* **2012**, *134*, 792–795.
  61. Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. *ACS Chem. Biol.* **2014**, *9*, 592–605.
  62. Lorand, J. P.; Edwards, J. O. *J. Org. Chem.* **1959**, *24*, 769–774.
  63. Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, *95*, 2457–2483.
  64. Zheng, H.; Lejkowski, M.; Hall, D.G. *Tetrahedron Lett.* **2013**, *54*, 91–94
  65. Gernigon, N.; Al-Zoubi, R. M.; Hall, D. G. *J. Org. Chem.* **2012**, *77*, 8386–8400.
  66. Hall, D. G. *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials (Volume 1 and 2), Second Edition* (Ed: Hall, D. G.), Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, p. 213–590.
  67. Sun, X.; Zhai, W.; Fossey, J. S.; James, T. D. *Chem. Commun.* **2016**, *52*, 3456–3469.
  68. Jabbour, A.; Steinberg, D.; Dembitsky, V. M.; Moussaieff, A.; Zaks, A. B.; Srebnik, M. *J. Med. Chem.* **2004**, *47*, 2409–2410.
  69. Chu, Y.; Wang, D. Z.; Wang, K.; Liu, Z. R.; Weston, B.; Wang, B. H. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 6307–6309.
  70. Wang, X.; Xia, N.; Liu, L. *Int. J. Mol. Sci.* **2013**, *14*, 20890–20912.
  71. Lacina, K.; Skládal, P. S.; James, T. D. *Chem. Cent. J.* **2014**, *8*, 1–17.
  72. Yan, J.; Springsteen, G.; Deeter, S.; Wang, B. *Tetrahedron* **2004**, *60*, 11205–11209.
  73. Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.
  74. Bandyopadhyay, A.; Gao, J. *Chem. Eur. J.* **2015**, *21*, 14748–14752.
  75. Gillingham, D. *Org. Biomol. Chem.* **2016**, *14*, 7606–7609.

76. Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220–8223.
77. Cal, P. M. S. D.; Frade, R. F. M.; Chudasama, V.; Cordeiro, C.; Caddick, S.; Gois, P. M. P. *Chem. Commun. (Camb.)* **2014**, *50*, 5261–5263.
78. Cal, P. M. S. D.; Frade, R. F. M.; Cordeiro, C.; Gois, P. M. P. *Chem. Eur. J.* **2015**, *21*, 8182–8187.
79. Schmidt, P.; Stress, C.; Gillingham, D. *Chem. Sci.* **2015**, *6*, 3329–3333.
80. Dilek, O.; Lei, Z.; Mukherjee, K.; Bane, S. *Chem. Commun.* **2015**, *51*, 16992–16995.
81. Stress, C. J.; Schmidt, P. J.; Gillingham, D. G. *Org. Biomol. Chem.* **2016**, *14*, 5529–5533.
82. Bandyopadhyay, A.; Cambray, S.; Gao, J. *J. Am. Chem. Soc.* **2017**, *139*, 871–878.
83. Bandyopadhyay, A.; Cambray, S.; Gao, J. *Chem. Sci.* **2016**, *7*, 4589–4593.
84. Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P. *Chem. Sci.* **2016**, *7*, 5052–5058.
85. Ratner, V.; Kahana, E.; Eichler, M.; Haas, E. *Bioconjugate Chem.* **2002**.
86. Botti, P.; Pallin, T. D.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 10018–10024.
87. a) Meadows, M. K.; Roesner, E. K.; Lynch, V. M.; James, T. D.; Anslyn, E. V. *Org. Lett.* **2017**, *19*, 3179–3182. b) Chapin, B. M.; Metola, P.; Lynch, V. M.; Stanton, J. F.; James, T. D.; Anslyn, E. V. *J. Org. Chem.* **2016**, *81*, 8319–8330.
88. Dowlut, M.; Hall, D. G. *J. Am. Chem. Soc.* **2006**, *128*, 4226–4227.
89. Ellis, G. A.; Palte, M. J.; Raines, R. T. *J. Am. Chem. Soc.* **2012**, *134*, 3631–3634.
90. Andersen, K. A.; Smith, T. P.; Lomax, J. E.; Raines, R. T. *ACS Chem. Biol.* **2016**, *11*, 319–323.
91. Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. *J. Am. Chem. Soc.* **2009**, *131*, 438–439.
92. Stolowitz, M. L.; Ahlem, C.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Li, G.; Lund, K. P.; Torkelson, S. M.; Wiley, J. P. *Bioconjugate Chem.* **2001**, *12*, 229–239.
93. Bin Y Shin, S.; Almeida, R. D.; Gerona-Navarro, G.; Bracken, C.; Jaffrey, S. R. *Chem. Biol.* **2010**, *17*, 1171–1176.
94. Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. *Tetrahedron Lett.* **2005**, *46*, 3301–3305.
95. Chalker, J. M.; Wood, C. S. C.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, *131*, 16346–16347.
96. Spicer, C. D.; Davis, B. G. *Chem. Commun. (Camb.)* **2011**, *47*, 1698–1700.

97. Spicer, C. D.; Triemer, T.; Davis, B. G. *J. Am. Chem. Soc.* **2012**, *134*, 800–803.
98. Spicer, C. D.; Davis, B. G. *Chem. Commun. (Camb.)* **2013**, *49*, 2747–2749.
99. Gao, Z.; Gouverneur, V.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, *135*, 13612–13615.
100. Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G. *Angew. Chem. Int. Ed.* **2013**, *52*, 3916–3921.
101. Ourailidou, M. E.; van der Meer, J. Y.; Baas, B. J.; Jeronimus Stratingh, M.; Gottumukkala, A. L.; Poelarends, G. J.; Minnaard, A. J.; Dekker, F. J. *ChemBioChem* **2014**, *15*, 209–212.
102. Eising, S.; Lelivelt, F.; Bongers, K. M. *Angew. Chem. Int. Ed.* **2016**, *55*, 12243–12247.
103. Li, J.; Jia, S.; Chen, P. R. *Nat. Chem. Biol.* **2014**, *10*, 1003–1005.
104. Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; Chen, P. R. *Nat. Chem.* **2014**, *6*, 352–361.
105. Völker, T.; Meggers, E. *Curr. Opin. Chem. Biol.* **2015**, *25*, 48–54.
106. Kim, J.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2015**, *54*, 15777–15781.
107. Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. *J. Mol. Biol.* **2010**, *395*, 361–374.
108. O'Donovan, M. R.; Mee, C. D.; Fenner, S.; Teasdale, A.; Phillips, D. H. *Mutat. Res.* **2011**, *724*, 1–6.
109. Walmsley, R. M. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2015**, *777*, 68–72.
110. Ciaravino, V.; Plattner, J.; Chanda, S. *Environ. Mol. Mutagen.* **2013**, *54*, 338–346.
111. Matteson, D. S.; Man, H.-W. *J. Org. Chem.* **1996**, *61*, 6047–6051.
112. a) Roy, C. D.; Brown, H. C. *Monatsh. Chem.* **2007**, *138*, 747–753. b) Roy, C. D.; Brown, H. C. *Monatsh. Chem.* **2007**, *138*, 879–887.
113. Ng, S.; Jafari, M. R.; Derda, R. *ACS Chem. Biol.* **2012**, *7*, 123–138.

## 2. Chapter 2. Fast and Tight Boronate Formation for ‘Click’ Bioorthogonal Conjugation

This chapter was modified from: Akgun, B.; Hall, D. G. Fast and Tight Boronate Formation for ‘Click’ Bioorthogonal Conjugation. *Angew. Chem. Int. Ed.* **2016**, *55*, 3909–3913.

### 2.1 Introduction

Bioorthogonal chemistry allows the investigation of the roles of proteins of interest (POIs) that are challenging to reveal using conventional methods.<sup>1</sup> A general protocol of bioorthogonal chemistry starts with introducing a biocompatible unnatural functionality into pre-determined sites on proteins in live cells or animals followed by modifying the protein with another biocompatible synthetic probe. Especially, the development of rapid bioorthogonal chemical reactions has been a focal point in the service of protein imaging in live cells or animals.<sup>1</sup> ‘Click’ reactions, which is a term coined by K. B. Sharpless in 2001, are spontaneous bond-forming processes that ideally provide a fast bioorthogonal reaction in aqueous media.<sup>2</sup> ‘Click’ chemistry involves reactions exhibiting fast reactivity, high yield, high selectivity, wide in scope and insensitivity to oxygen and water.<sup>2b</sup> Additionally, these reactions release benign by-products and use no solvent or benign solvents, such as water.<sup>2b</sup> Unfortunately, the repertoire of ‘click’ reactions between abiotic and biocompatible functional groups is relatively small, and some of the most popular and efficient reactions present inconveniences that may limit some of their applications in living systems. For example, copper-free azide-alkyne cycloadditions tend to display rates that are slower than desired,<sup>1c,3</sup> and tetrazine cycloadditions, while very fast, may be susceptible to hydrolysis and side-reactions with endogenous thiols.<sup>4</sup> Furthermore, many common bioorthogonal reactions are not mutually orthogonal due to possible cross-reactions,<sup>5</sup> thus new systems would provide more options for interrogating cellular processes in tandem.

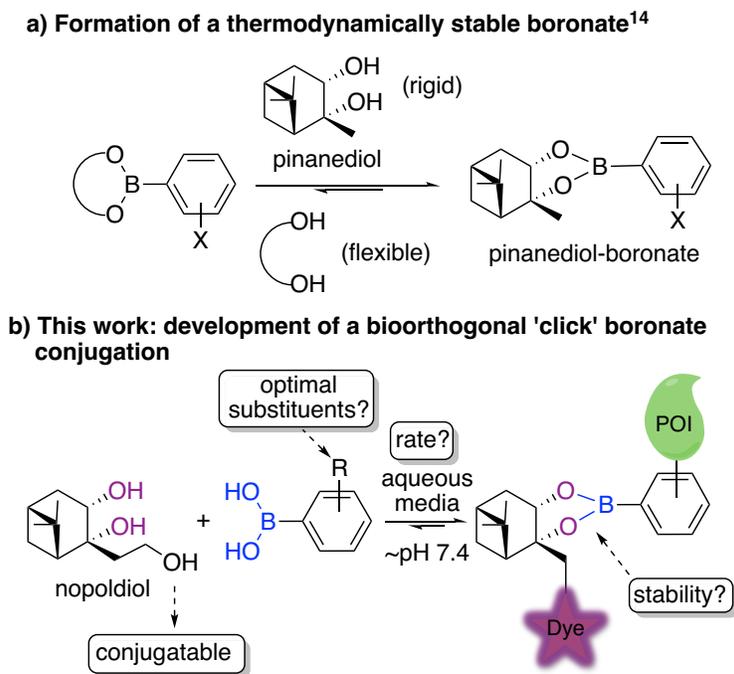
#### 2.1.1 Pinanediol boronate stability compared to other 1,2-*cis* diols

In this context, boronic ester formation, a process better known in the field of carbohydrate sensing,<sup>6</sup> has remained an underexplored option in bioconjugation despite recent applications in surface conjugation,<sup>7</sup> detection of reactive oxygen/nitrogen species<sup>8</sup> and protein modification.<sup>9</sup> Significant attributes of boronate bioconjugation are its simplicity (no catalyst

required), bioorthogonality (boronic acids complex biological diols with low affinities), fast on-rates, and potential reversibility. In this regard, Jaffrey and co-workers recently reported a phenylboronic acid–salicylhydroxamic acid bioorthogonal conjugation system, which can rapidly form a stable boronic ester at neutral pH.<sup>10</sup> The reaction was shown to be orthogonal and stable in the presence of typical biological nucleophiles such as cysteine, lysine, serine and also Dulbecco's Modified Eagle's Medium (DMEM), which contains a high amount of glucose. Although this system shows promise, it displays moderate binding affinity ( $K_{\text{eq}} \approx 17,800 \text{ M}^{-1}$ ). The design of a tight boronate conjugation system is hampered by mechanistic ambiguities<sup>11</sup> and a dearth of comprehensive kinetic data ( $k_{\text{ON}}$ ,  $k_{\text{OFF}}$ ,  $K_{\text{eq}}$ ) on boronic ester formation in water.<sup>12</sup> For instance, studies and debates are still ongoing to investigate the reactive species of boronic acid (trigonal boronic acid vs tetrahedral boronate ion) towards diols.<sup>11,12</sup> The type of group attached to the boronic acid and pH of the medium can affect the reactivity and stability of boronic ester; however it is experimentally challenging to perform reliable reaction mechanism and kinetic measurements of boronic ester formation in aqueous media due to the proton ambiguity.<sup>12</sup> However, this information would enable a systematic optimization of both diol and boronic acid partners. As a condensation reaction, boronic ester formation is intrinsically unfavored in water owing to Le Châtelier's Principle. Thermodynamically, because the balance of bond enthalpy is somewhat neutral (two O–H and two B–O bonds are broken to form four similar bonds), the process is driven solely by the favorable entropy of ring formation.

It is also well established that hindered and pre-organized vicinal diols mitigate the loss of entropy in the diol substrate, and are thus more favorable.<sup>13</sup> In 2007, Brown and co-workers reported a quantitative transesterification analysis of phenylboronic esters (five representative chiral and achiral boronic esters) with twenty different cyclic, acyclic, sterically hindered diols under neutral conditions to investigate the relative stability of these arylboronic esters and understand the factors that affect this stability.<sup>14</sup> Unsurprisingly, they found that pinanediol affords some of the most hydrolytically robust boronic esters (Figure 1a).<sup>14</sup> The high hydrolytic stability of pinanediol boronic esters is simply due to the orientation of hydroxyl groups and the rigidity of the free diol.<sup>13</sup> Meanwhile, the use of nopoldiol, a pinanediol derivative, and arylboronic acid for bioconjugation purposes was described in a patent by Baggio and co-workers.<sup>15</sup> However, in this work, there is no information about the rates of nopoldiol boronate formation in aqueous solvents, the question of its reversibility, and more importantly, the

structural requirements (sterics and electronics) for an optimal boronic acid partner. In this chapter, I aim to describe the design and proof-of-concept of a fast and tight-binding ‘click’ boronate bioconjugation system using the conjugatable nopoldiol (Figure 1b).<sup>15</sup>



**Figure 2-1:** a) Pinanediol forms the most stable boronates with arylboronic acids in the presence of various competing diols. b) This work involves the design of a bioorthogonal ‘click’ boronate conjugation using conjugatable nopoldiol derivatives and optimal arylboronic acids in neutral aqueous media.

## 2.2 Results and discussion

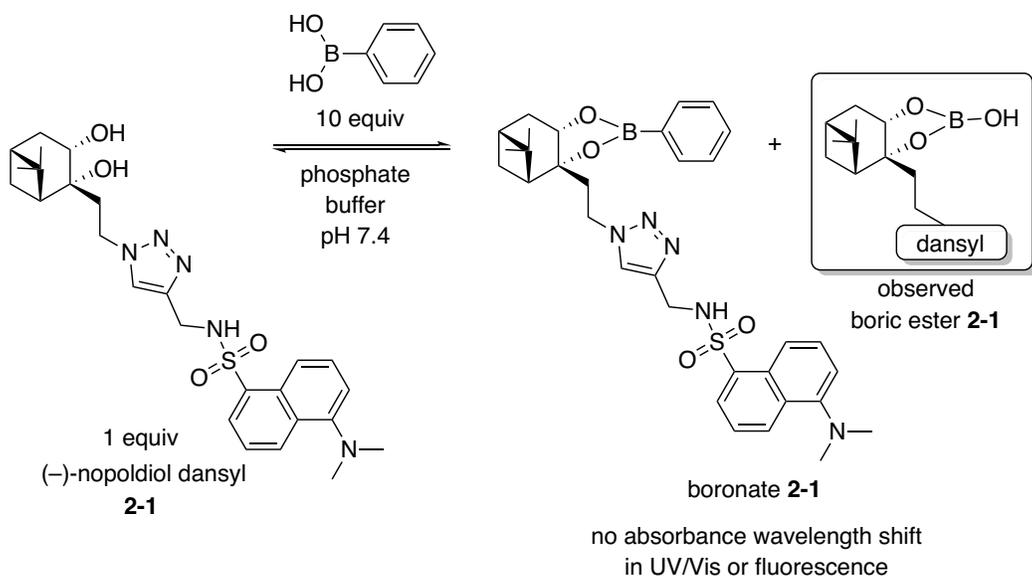
### 2.2.1 Optimization of boronate formation – hydrolytic stability & kinetic studies

#### 2.2.1.1 Initial attempts – stopped flow technique & RP-HPLC

To identify fast-forming boronates in aqueous medium, I first focused on exploring a method to measure the initial rate constants of boronate formation. Initial trials showed that the reaction was very fast, complete within a few minutes. One way to measure the initial rates of second order reactions, such as boronate formation, uses excess amount (10 equiv) of one reactant, thus allowing the reaction to become a pseudo first order process. Measured pseudo first order rate constants of the reaction may then be employed to calculate its second order rate

constant.<sup>16</sup> However, pseudo first order reaction conditions further increased the rate of boronate formation, which became unfeasible to measure by <sup>1</sup>H NMR spectroscopy; therefore low reactant concentrations ( $\mu$ M) were required to monitor the initial rates. Even then, it was challenging to obtain data within a few seconds. Therefore, a common method named stopped-flow technique was considered. In this approach, two reactant solutions in two separate syringes are rapidly driven into a mixing chamber to initiate the reaction. This approach was initially considered to work for fast-forming boronate systems (3 to 60 s).<sup>17</sup> Monitoring of the reaction is usually accomplished with photometric detectors, such as UV-Vis and fluorescence. This detection method, however, depends upon a shift in absorbance/fluorescence between the reactants and the product to be able to determine its rate constant. This technique was unfortunately determined to be unsuitable since a desired shift in absorbance or fluorescence excitation wavelength was unseen upon boronate formation between fluorescently labeled nopoldiol dansyl (**2-1**) and phenylboronic acid (Scheme 2-1). The diol derivative **2-1** was synthesized *via* copper catalyzed azide alkyne cycloaddition (CuAAC) between easily accessible (-)-nopoldiol azide and a terminal alkyne derivatized dansyl fluorophore.

Using reverse phase-HPLC with a fluorescence detector, which could efficiently separate and detect the peaks of diol **2-1** and its boronate, was also considered. This method, however, was found to be inefficient and cumbersome because each detection period lasted 10 minutes and it was not possible to monitor only one reaction in real time with short intervals (every 10 seconds). In short, this method required preparation of many reaction solutions to detect the conversion of a single boronate formation within short time intervals. Evidently, screening various arylboronic acids would be overly time-consuming with this HPLC method. Interestingly, during this study it was found that a small amount of boric ester **2-1** formed between diol **2-1** and boric acid (Scheme 2-1). I confirmed that the boric acid source in the system originated from the glass bottle used to store water for the HPLC mobile phase. Boric acid is likely leached from the glass bottle into the water over time. This information led us to keep the water in a plastic bottle for the following studies so that a boric acid-free HPLC system was obtained.



**Scheme 2-1:** Initial approaches for kinetic studies and observation of boric ester **2-1** from the HPLC analysis.

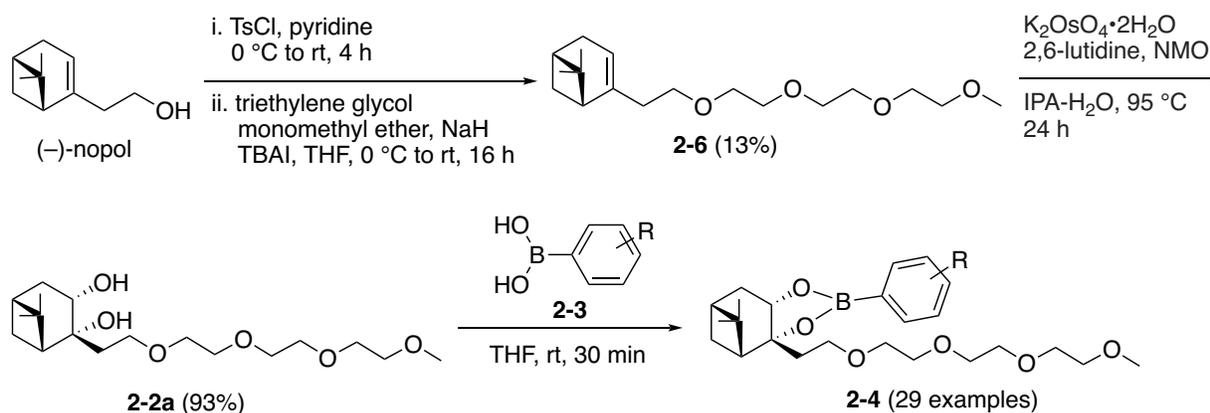
### 2.2.1.2 $^1\text{H}$ NMR spectroscopy

Next, I decided to directly measure the second order rate constants using equimolar concentrations of nopoldiol and arylboronic acid. Thus,  $^1\text{H}$  NMR spectroscopy was selected as the desired instrumentation owing to its convenience, even though its capability to monitor fast reactions is limited.<sup>18</sup> The low intrinsic sensitivity of NMR makes it difficult to detect very low concentrations ( $< 100 \mu\text{M}$ ), therefore it requires higher reaction concentration at which fast processes ( $> 50 \text{ M}^{-1}\text{s}^{-1}$ ) are not readily monitored. Moreover, early rates ( $< 40 \text{ s}$ ) cannot be captured owing to the essential gradient shimming process. Thus, because the formation of boronic ester is likely reversible, the forward rate data obtained from NMR kinetics provide an underestimation of rates, which is however sufficient for a preliminary comparison and identification of promising reagents. In this way, NMR analysis allowed us to rapidly obtain relative hydrolytic stability data and forward rate constants ( $k_{\text{ON}}, \text{M}^{-1}\text{s}^{-1}$ ) for a large set of arylboronic acids.

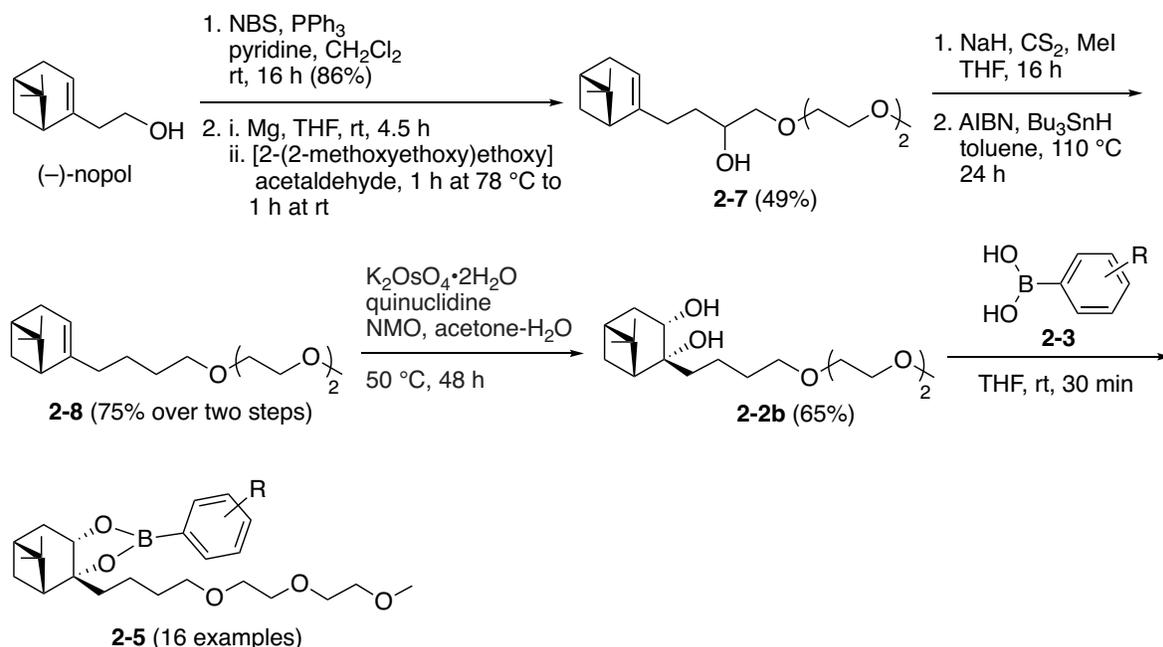
#### 2.2.1.2.1 Design and synthesis of model water soluble nopol-PEG-derivatives

To identify the optimum boronates in aqueous medium, model water-soluble nopol-PEG-diols **2-2a/2-2b** and their boronates **2-4/2-5** were designed and synthesized as described in Scheme 2-2 and Scheme 2-3. Diol **2-2a** was synthesized from commercially available (-)-nopol,

of which the hydroxyl group was initially converted into a good leaving group (tosylate) followed by etherification of crude (–)-nopol-tosylate with triethylene glycol monomethyl ether in the presence of sodium hydride (NaH) to obtain (–)-nopol-PEG **2-6** (Scheme 2-2). Then, compound **2-6** was subjected to Upjohn dihydroxylation to obtain diol **2-2a** with 93% conversion. Another diol derivative, **2-2b**, without a proximal oxygen (replaced with a methylene group) was designed to study the effect of this oxygen atom on the hydrolytic stability and rate constant. Diol **2-2b** was also synthesized from (–)-nopol, which was converted into (–)-nopol bromide **2-7a** using the Appel reaction followed by Grignard reagent formation and nucleophilic addition to freshly prepared [2-(2-methoxyethoxy)ethoxy]acetaldehyde<sup>19</sup> to obtain (–)-nopol-PEG-hydroxyl **2-7** (Scheme 2-3). Next, the hydroxyl group was removed with a Barton-McCombie deoxygenation, which provided (–)-nopol-PEG derivative **2-8** without the proximal oxygen. Dihydroxylation of **2-8** was also performed under Upjohn dihydroxylation conditions in the presence of quinuclidine, which was critical to provide a relatively better yield of 65%. The corresponding boronic esters with various arylboronic acids **2-4** and **2-5** were synthesized and easily isolated by flash chromatography to be employed in hydrolytic stability and backward kinetic studies.

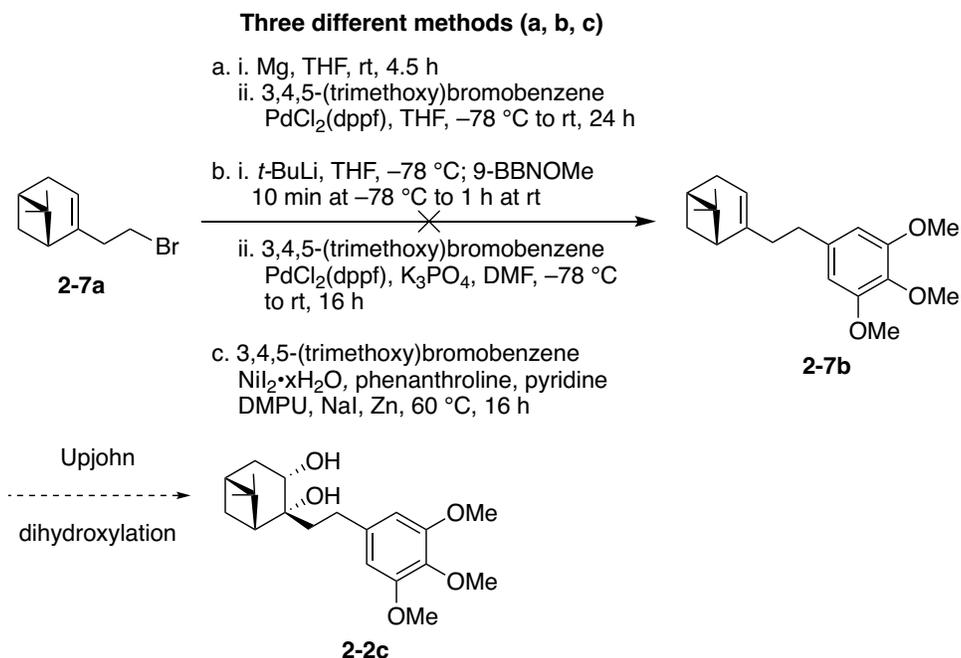


**Scheme 2-2:** Synthesis of (–)-nopol-PEG-diol-1 **2-2a** and its corresponding boronic esters **2-4**.



**Scheme 2-3:** Synthesis of (-)-nopol-PEG-diol-2 **2-2b** and its corresponding boronic esters **2-5**.

In addition to the proximal oxygen effect, I also aimed to study the effect of the aromatic group installed on (-)-nopol-aryl-diol **2-2c**. As shown in Scheme 2-4, 3,4,5-(trimethoxy)bromobenzene was chosen as a substrate in this reaction due to its three methoxy substituents, which could increase the aqueous solubility of diol **2-2c**. To obtain **2-2c**, three different cross-coupling conditions were tried, however I could not synthesize the desired product. The first coupling conditions involved the formation of the Grignard reagent from **2-7a**, which was then added to a solution of 3,4,5-(trimethoxy)bromobenzene in the presence of catalyst PdCl<sub>2</sub>(dppf) (Scheme 2-4, reaction condition a).<sup>20</sup> The second attempt was alkyl Suzuki-Miyaura cross-coupling reaction. The first step was *in-situ* formation of (-)-nopol-BBN that was then allowed to react with 3,4,5-(trimethoxy)bromobenzene in the presence of catalyst PdCl<sub>2</sub>(dppf) and K<sub>3</sub>PO<sub>4</sub> (Scheme 2-4, reaction condition b).<sup>21</sup> However, this reaction also did not proceed efficiently and the desired product **2-7b** could not be isolated. Lastly, conditions from a recently reported nickel-catalyzed direct reductive alkylation of aryl bromides were applied to synthesize (-)-nopol-aryl **2-7b** (Scheme 2-4, reaction condition c).<sup>22</sup> This attempt was also unsuccessful so finally it was decided to continue with the easily accessible model water-soluble nopol-PEG-diols **2-2a** and **2-2b**.



**Scheme 2-4:** Attempts to synthesize (-)-nopol-aryl-diol **2-2c**.

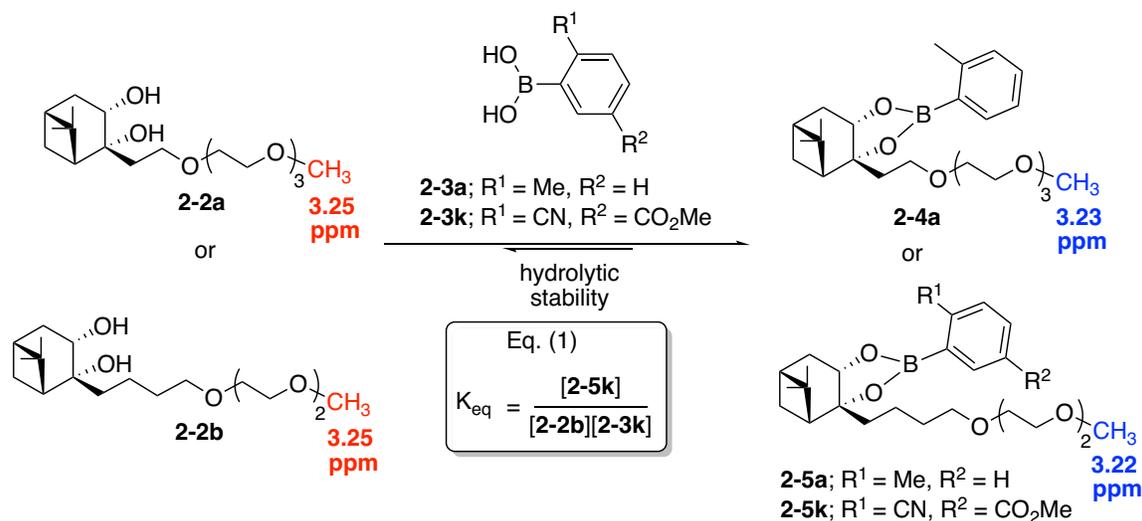
#### 2.2.1.2.2 Hydrolytic stability and association constant ( $K_{eq}$ ) measurement

Before discussing the entire results of optimization studies, I would like to clarify the methodology for hydrolytic stability, association constant ( $K_{eq}$ ) and forward/backward reaction rate constant measurements, and to present their outcomes. Later, all these studies will be discussed and summarized to determine the optimized nopol-boronate system.

Even though PEG conjugated diols **2-2a** and **2-2b** were designed to be water soluble, the solubility of their corresponding boronates was intrinsically low at high concentrations (0.5 – 5 mM). Therefore, several co-solvents such as 2-propanol-*d*<sub>8</sub>, methanol-*d*<sub>4</sub>, dimethyl sulfoxide-*d*<sub>6</sub>, acetone-*d*<sub>6</sub> or acetonitrile-*d*<sub>3</sub> (ACN-*d*<sub>3</sub>) were tested to dissolve boronates **2-4/2-5**. <sup>1</sup>H NMR studies confirmed that a mixture of D<sub>2</sub>O and ACN-*d*<sub>3</sub> (65/35 w%) was the ideal solution to dissolve 10 mM concentration of boronates **2-4/2-5** and provide a clear chemical shift difference between -OCH<sub>3</sub> of **2-2a/2-2b** and -OCH<sub>3</sub> of **2-4/2-5**. This chemical shift difference was critical to allow a proper integration of peaks and measure the proportions between **2-2a/2-2b** and **2-4/2-5** during hydrolytic stability and kinetic studies.

The hydrolytic stability of boronates **2-4** or **2-5** was investigated in 0.1 M of D<sub>2</sub>O potassium phosphate buffer:ACN-*d*<sub>3</sub> (65:35 w%, pD 7.4). This solvent system was determined

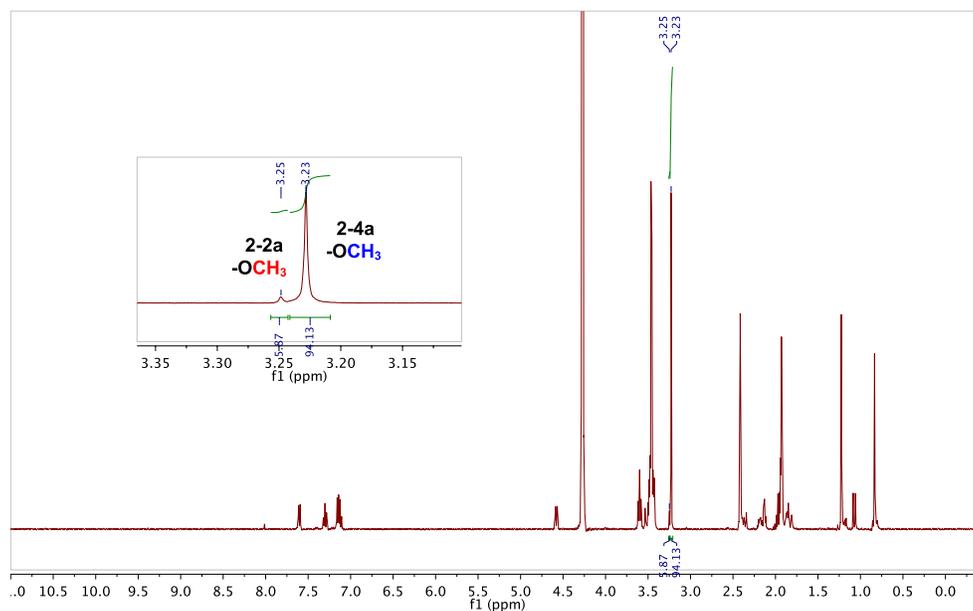
according to the solubility of boronate **2-4a** (at 10 mM) and the same solvent system was used for all other boronates. Boronates **2-4** or **2-5** were dissolved at various concentrations (10, 5, 1 or 0.1 mM) in the designated solution. When the  $^1\text{H}$  NMR spectra of free **2-2a** or **2-2b** were compared with  $^1\text{H}$  NMR spectra of **2-4** or **2-5**, an apparent chemical shift was observed for  $-\text{OCH}_3$  from 3.25 to 3.22 or 3.23 ppm. After boronate solutions were incubated for 24 h at room temperature, their hydrolytic stability (**2-4/2-2a** or **2-5/2-2b**) was monitored *via* integral ratio of both methyls (Scheme 2-5, Figure 2-2 and 2-3). Not surprisingly, owing to Le Châtelier's Principle, higher amounts of free diols **2-2a/2-2b** were present at lower concentration of boronates **2-4/2-5**. For example, Figure 2-2A shows the hydrolytic equilibrium as 94:6 (**2-4a:2-2a**) at 5 mM concentration and Figure 2-2B confirms a lower hydrolytic equilibrium as 87:13 (**2-4a:2-2a**) at 1 mM concentration.



**Scheme 2-5:** Representative scheme of hydrolytic stability study and association constant measurement ( $K_{\text{eq}}$ ,  $\text{M}^{-1}$ ).

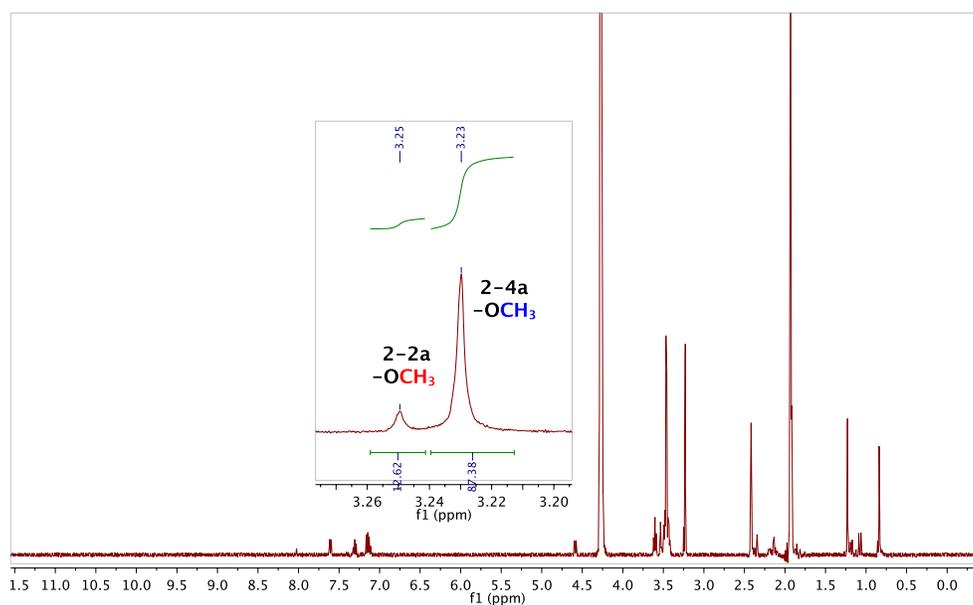
A)

Boronate **2-4a**; Hydrolytic stability study (5 mM);  
400.394 MHz H1 PRESAT in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm), temp 27.0 C -> actual temp = 27.0 C, m400gz probe



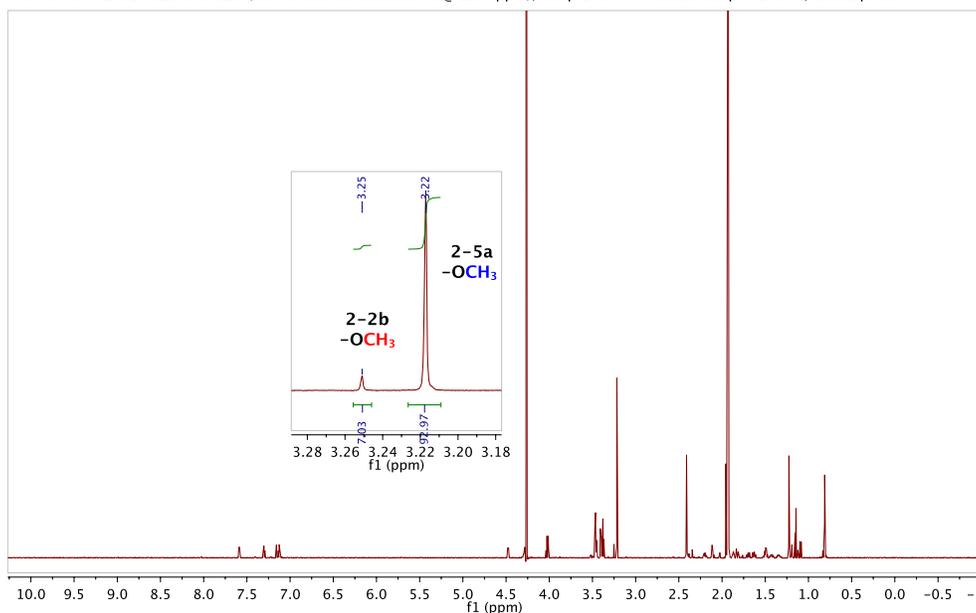
B)

Boronate **2-4a**; Hydrolytic stability study (1 mM);  
400.393 MHz H1 PRESAT in d<sub>2</sub>o (ref. to external acetonitrile @ 1.93 ppm), temp 27.0 C -> actual temp = 27.0 C, m400gz probe



**Figure 2-2:** Representative <sup>1</sup>H NMR spectra for hydrolytic stability study of boronate **2-4a**. A) 5 mM of boronate **2-4a**. B) 1 mM of boronate **2-4a**.

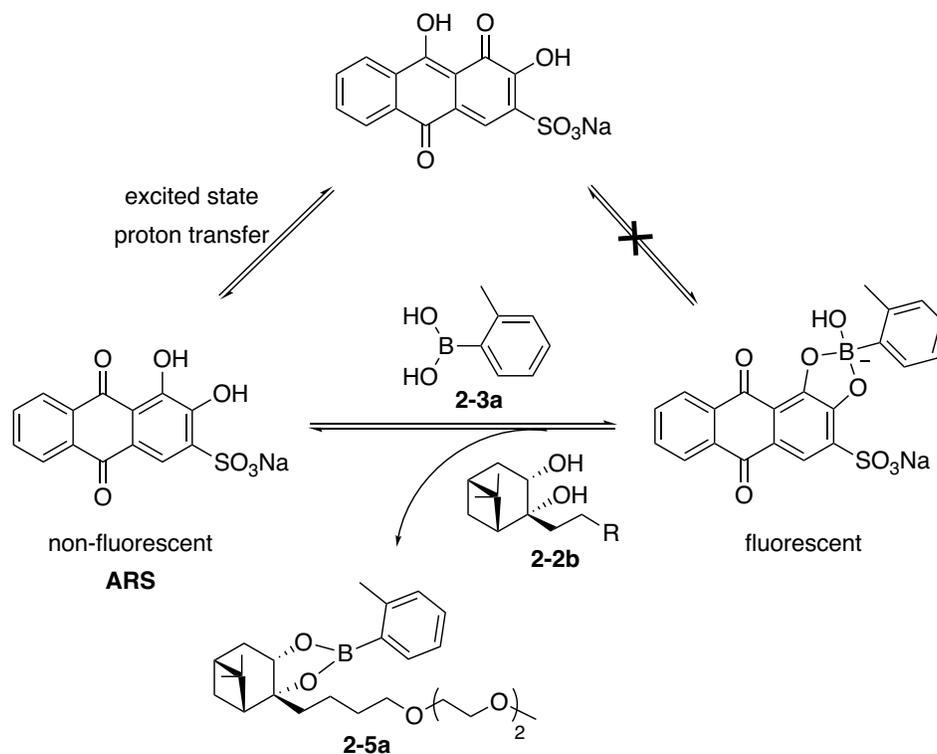
Boronate **2-5a**; Hydrolytic stability study (1 mM);  
699.770 MHz H1 PRESAT in d2o (ref. to external acetonitrile @ 1.93 ppm), temp 27.5 C -> actual temp = 27.0 C, coldid probe



**Figure 2-3:** Representative <sup>1</sup>H NMR spectrum for hydrolytic stability study of 1 mM of boronate **2-5a**.

Apart from hydrolytic stability studies, I sought to determine the  $K_{eq}$  values of boronate formation. A well-known method to investigate  $K_{eq}$  of a boronic acid towards a 1,2-diol is the Alizarin Red S (ARS) competitive dye displacement assay, described by Wang and co-workers.<sup>23</sup> This assay has been widely applied in carbohydrate-boronic acid interactions, both quantitatively and qualitatively. Non-fluorescent ARS (due to its excited state) becomes fluorescent when it binds to a boronic acid to form its boronic ester, thus removing the quenching effect of catechol dihydroxyl active hydrogens (Scheme 2-6). This process may lead to significant changes in the fluorescent intensity (quantitative) and color (qualitative) from pink to yellow. However, when a competing 1,2-diol, such as diol **2-2b**, is introduced into the solution, the ARS-boronic acid equilibrium may shift towards non-fluorescent ARS formation because of the displacement of the boronic acid. This change in the ARS-boronic acid equilibrium generates a difference in fluorescent intensity, which is applied to measure the  $K_{eq}$  of boronate formation. Yet, diol **2-2b** shows a significantly higher binding affinity towards arylboronic acids ( $10^5$ – $10^6$  M<sup>-1</sup>) compared to ARS ( $\sim 1300$  M<sup>-1</sup>), thus the addition of even small amount of diol **2-2b** (20  $\mu$ M) to the ARS-boronic acid complex resulted in a complete cleavage of the ARS-boronic acid, forming only boronate **2-5a** (Scheme 2-6). In other words, with the ARS assay, it was not feasible to detect a

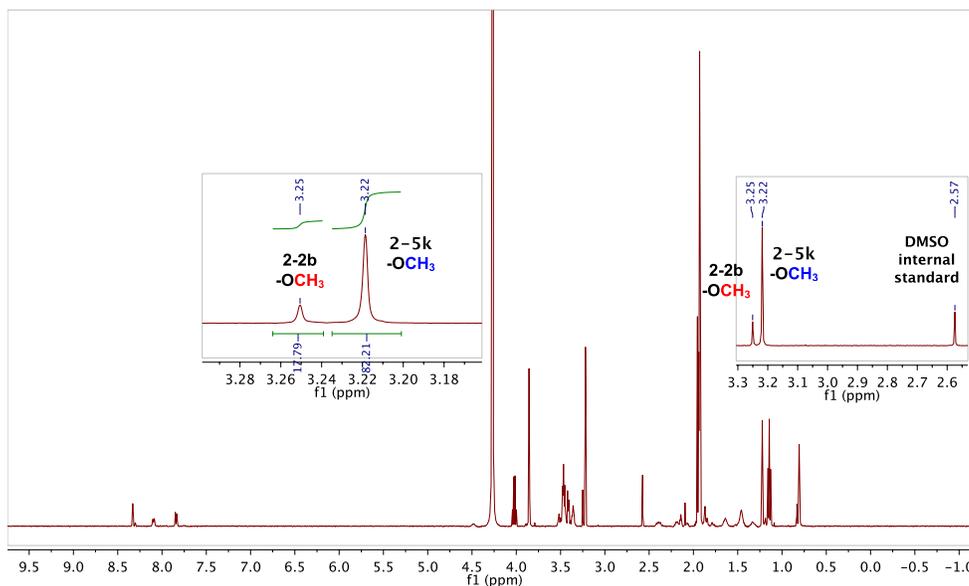
gradual fluorescent change upon addition of various concentrations of **2-2b** to the ARS-boronic acid complex and to determine the  $K_{eq}$  of boronate formation.



**Scheme 2-6:** Alizarin Red S (ARS) assay system to determine the  $K_{eq}$  of an arylboronic acid **2-3a** towards diol **2-2b**.

On the other hand,  $^1\text{H}$  NMR spectroscopy enabled an apparent  $-\text{OCH}_3$  peak separation (Figure 2-4), which was suitable to quantify the concentrations of both free diol **2-2b** and boronate **2-5** in equilibrium at several dilutions of **2-5** in the presence of an internal standard (DMSO). These data were then used to calculate  $K_{eq}$  with the formula indicated in Scheme 2-5.

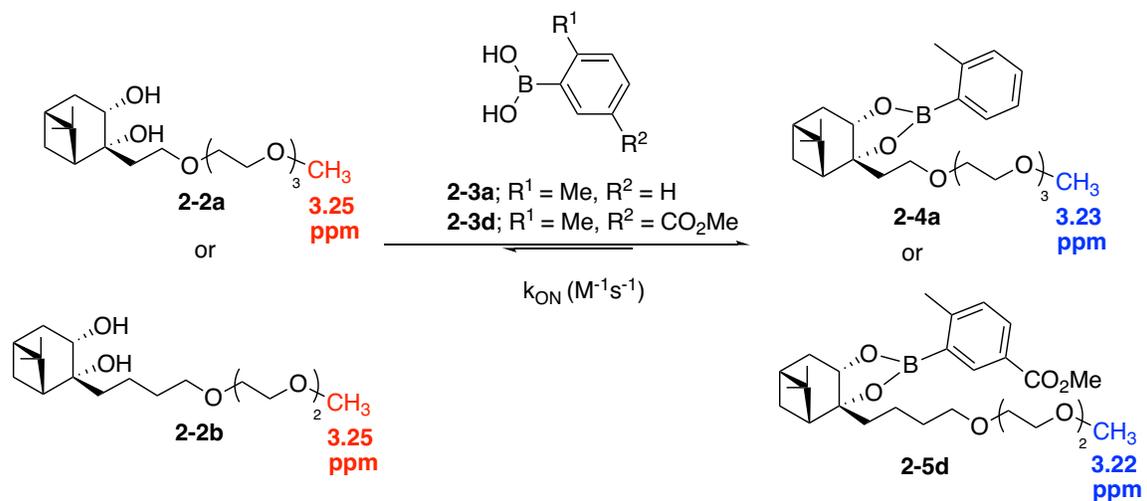
Boronate **2-5k**;  $K_{eq}$  measurement NMR study;  
498.120 MHz H1 1D in dmsd (ref. to acetonitrile @ 1.93 ppm), temp 26.4 C -> actual temp = 27.0 C, autoxdb probe



**Figure 2-4:** Representative  $^1\text{H}$  NMR spectrum for  $K_{eq}$  measurement of boronate **2-5k**.

### 2.2.1.2.3 Forward rate constant measurements ( $k_{ON}$ , $\text{M}^{-1}\text{s}^{-1}$ )

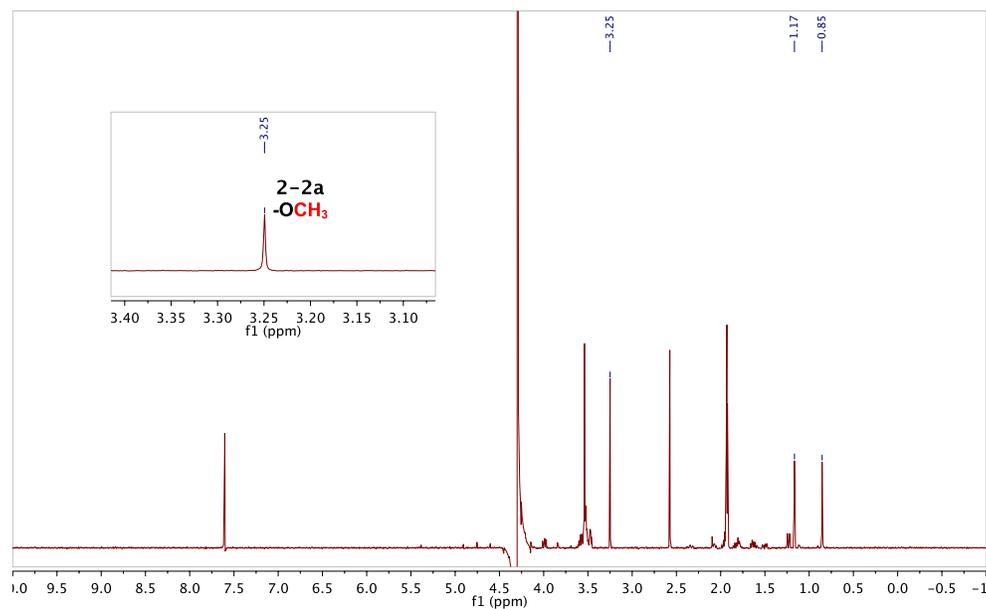
The forward reaction between diols **2-2a/2-2b** and arylboronic acids **2-3** was monitored at  $27^\circ\text{C}$  or  $0^\circ\text{C}$  in 0.05 M of  $\text{D}_2\text{O}$  potassium phosphate buffer:ACN- $\text{d}_3$  (70:30 w%, pD 7.4). A small change of solvent system (from 65:35 to 70:30 w%) was performed for the kinetic study because the measurements were collected at lower concentrations (0.5–2.5 mM) compared to the hydrolytic stability study, minimizing the risk of limited water solubility issues. Thus I decided to increase the amount of  $\text{D}_2\text{O}$  to improve the biological relevance of the solvent system. All boronates were found to be soluble in this solvent system and these conditions were kept constant. ACN- $\text{d}_3$  was locked to monitor the reactions and DMSO was chosen as an internal standard, appearing as a singlet at 2.58 ppm without overlapping peaks of diols **2-2a/2-2b**. The disappearance of the  $-\text{OCH}_3$  signal of **2-2a/2-2b** (3.25 ppm) was monitored at various time intervals (14 s to 60 s) (Figure 2-5, Scheme 2-7). These data points were converted into diol **2-2a/2-2b** concentrations by exploiting a calibration curve (various concentrations of diol **2-2a/2-2b** vs integral ratio of  $\text{CH}_3/\text{DMSO}$ ). The second order rate constant was determined by plotting  $1/[\mathbf{2-2a}$  or  $\mathbf{2-2b}]$  vs time (Figure 2-6 and Figure 2-7).



**Scheme 2-7:** Representative scheme of forward rate constant measurements ( $k_{\text{ON}}, \text{M}^{-1}\text{s}^{-1}$ ).

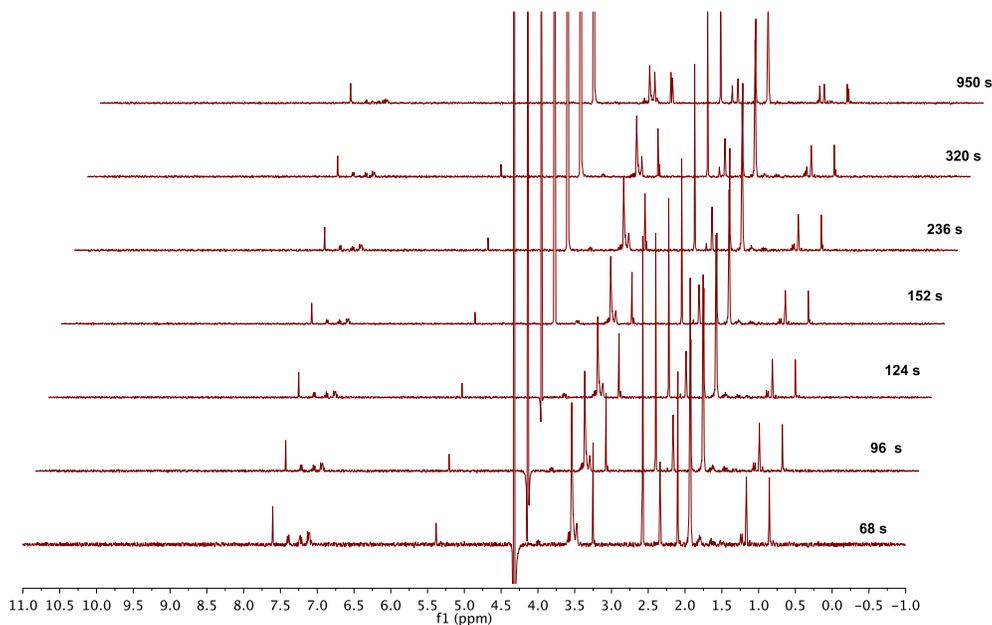
A)

$^1\text{H-NMR}$  of **2-2a** in buffer system; temp 27.0C --> actual temp = 27.0, sw400 probe



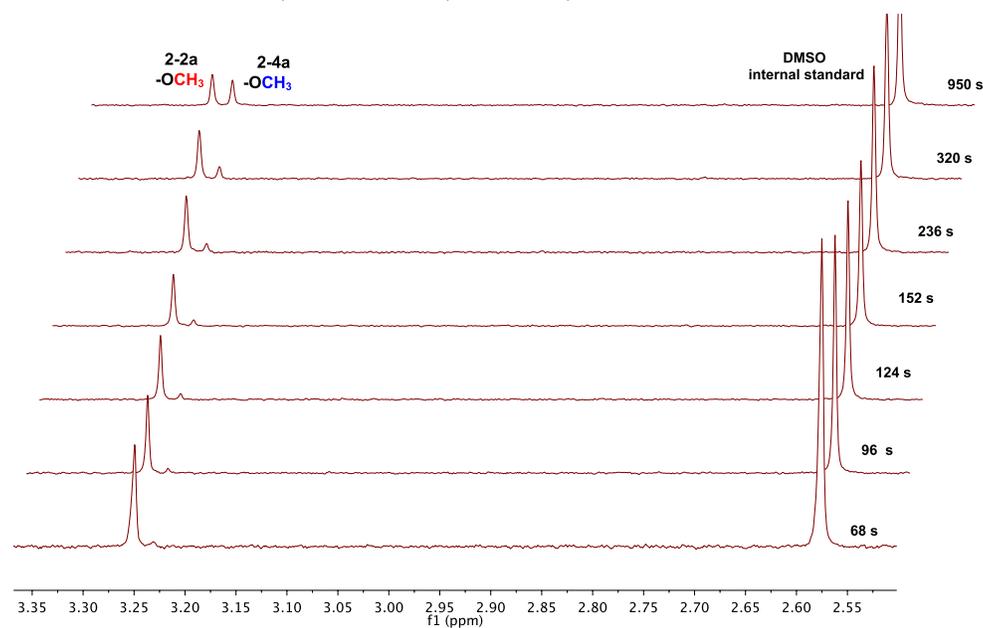
B)

2-2a and 2-3a; NMR kinetic study;  
399.953 MHz H1 PRESAT in cd3cn temp 27.0C --> actual temp = 27.0, sw400 probe



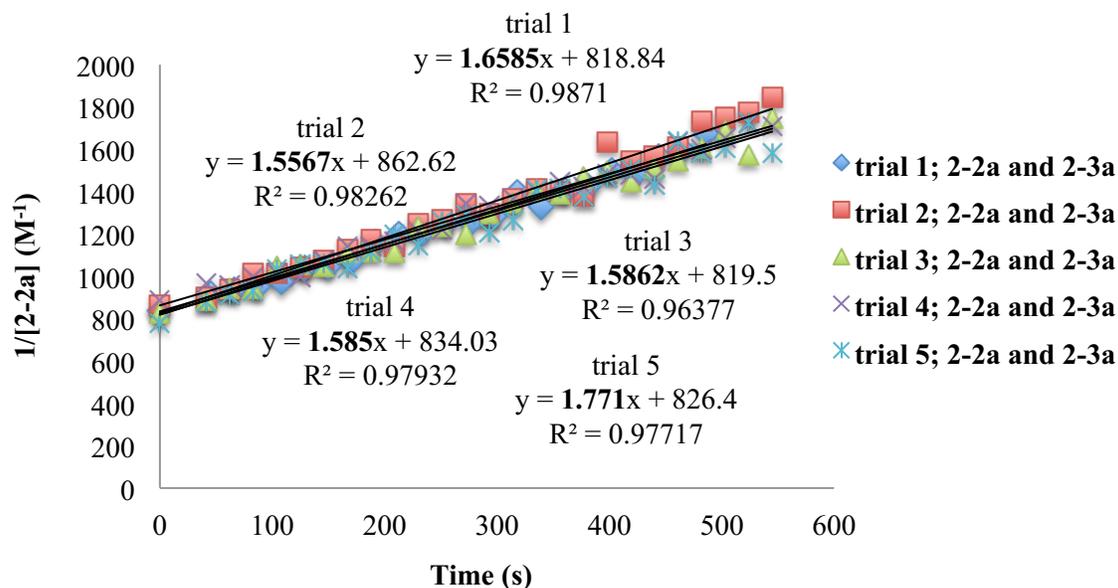
C)

2-2a and 2-3a; NMR kinetic study;  
399.953 MHz H1 PRESAT in cd3cn temp 27.0C --> actual temp = 27.0, sw400 probe

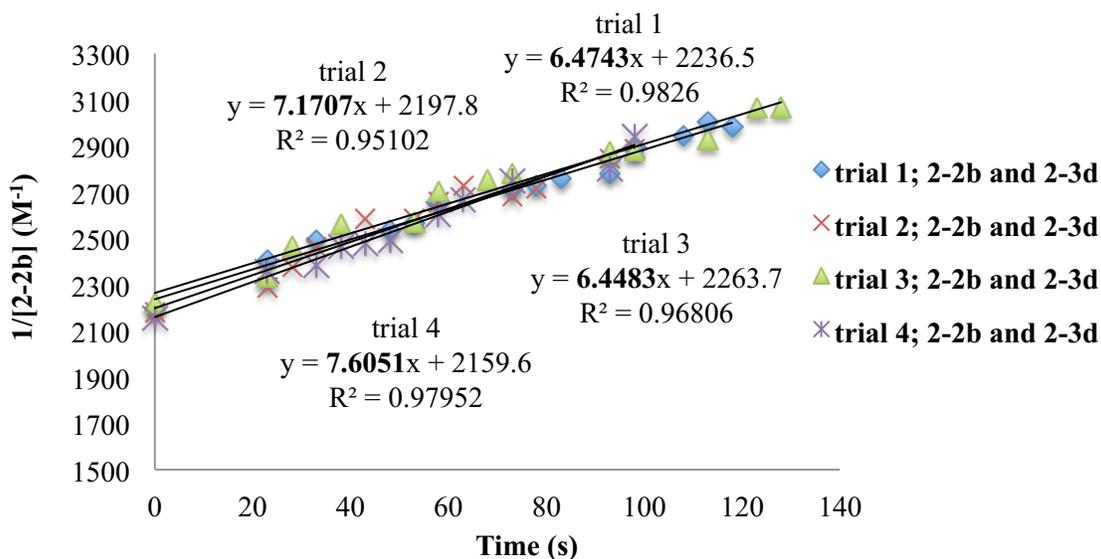


**Figure 2-5:** Representative time dependent <sup>1</sup>H NMR analysis for the forward reaction between diol **2-2a** and boronic acid **2-3a**. A) <sup>1</sup>H-NMR spectrum of **2-2a** in the designated buffer system.

B) Full spectrum. C) Diol **2-2a** and boronate **2-4a** –OCH<sub>3</sub> region. D) Diol **2-2a**, boronate **2-4a** – OCH<sub>3</sub> peaks and internal standard DMSO region.



**Figure 2-6:** Representative forward rate constant measurements, second order plot of kinetics **2-2a** and **2-3a**.

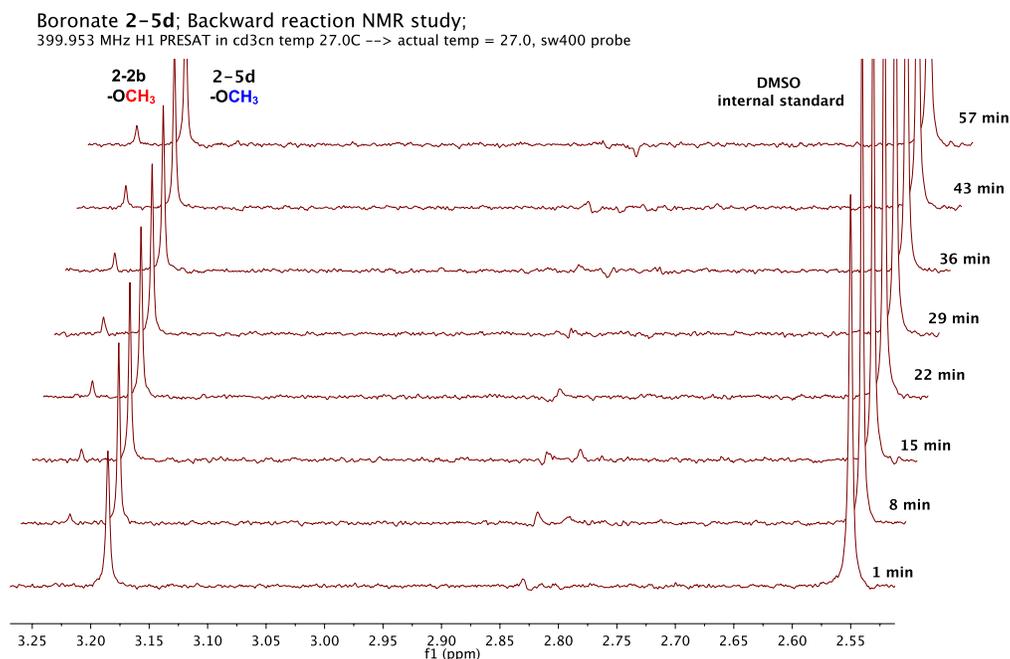


**Figure 2-7:** Representative forward rate constant measurements, second order plot of kinetics **2-2b** and **2-3d**.

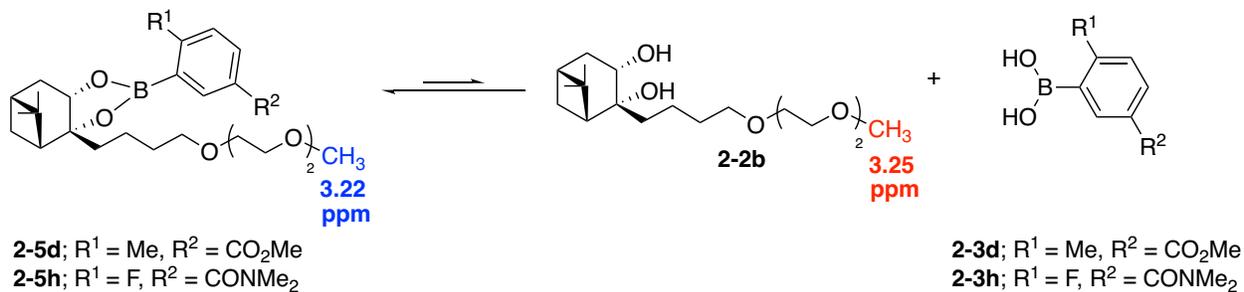
#### 2.2.1.2.4 Monitoring the backward reaction

The backward reaction of boronates **2-5d** and **2-5h** was monitored at 27 °C in 0.1 M of D<sub>2</sub>O potassium phosphate buffer:ACN-d<sub>3</sub> (65:35 w%, pD 7.4). The appearance of the **-OCH<sub>3</sub>** peak of **2-2b** at 3.25 ppm was monitored and converted into a diol concentration by using a calibration curve.

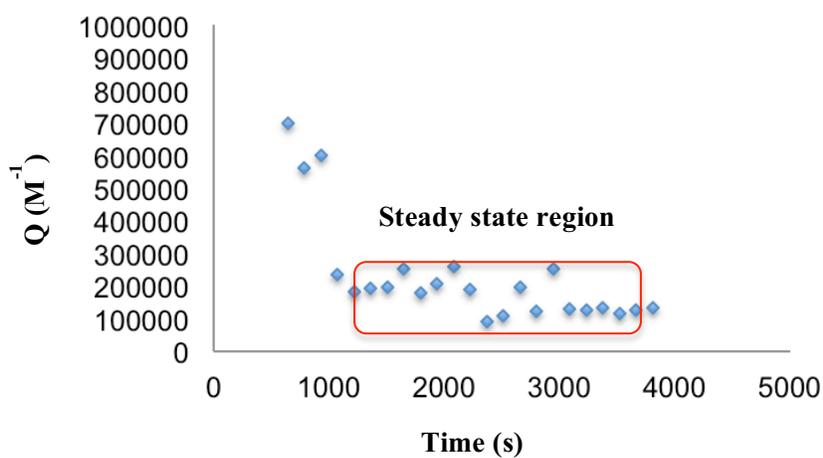
More specifically, the backward reaction of **2-5d** was monitored over 60 min. Figure 2-8 shows a representative example of two trials, and the equilibrium reached its steady state ( $K_{eq} = 13 \times 10^4 \text{ M}^{-1}$ ) in ~25 min in both trials (Figure 2-9, Scheme 2-8). On the other hand, the backward reaction of **2-5h** was monitored over 5 min. Figure 2-10 displays the equilibrium of boronate **2-5h**, which reached its steady state ( $K_{eq} = 12 \times 10^3 \text{ M}^{-1}$ ) in ~1 min according to the results of two trials.



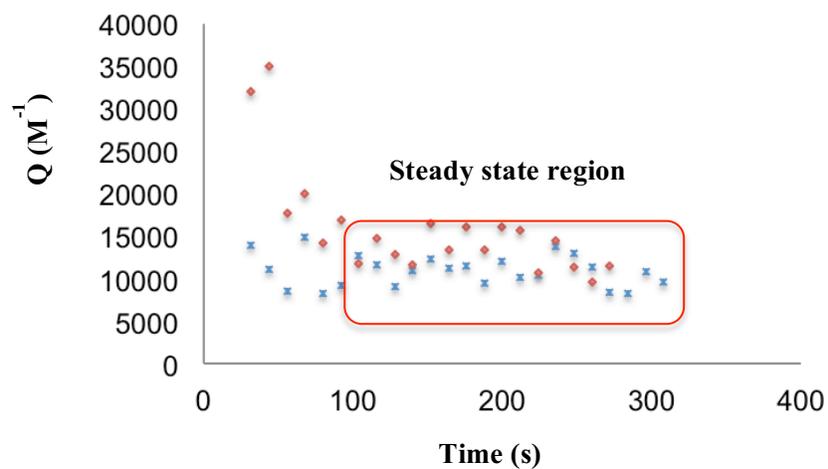
**Figure 2-8:** Representative time dependent <sup>1</sup>H NMR analysis for backward reaction of boronate **2-5d**.



**Scheme 2-8:** Equation for the monitoring of a backward reaction.



**Figure 2-9:** Backward reaction and Q vs time graph of **2-5d**.

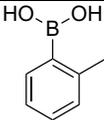
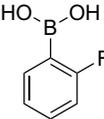
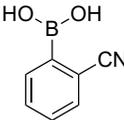


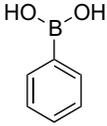
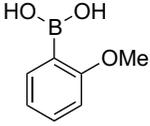
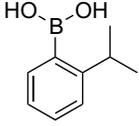
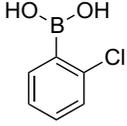
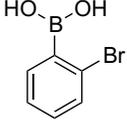
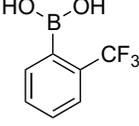
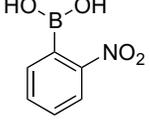
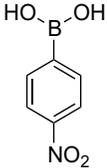
**Figure 2-10:** Backward reaction and Q vs time graph of **2-5h**.

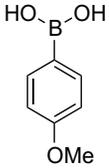
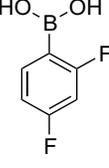
### 2.2.1.3 Overall results and discussion of nopol-boronate optimization

After extensive optimization studies, it was found that *ortho*-methyl or isopropyl groups improved the hydrolytic stability of boronates likely because of their bulkiness, while halogens and other electron- withdrawing groups decreased boronate stability, which is probably due to the increased acidity of these arylboronic acids (Table 2-1 and 2-2). Hydrolytic stability showed a near-linear correlation with the acidity of arylboronic acids (Figure 2-11).<sup>11d</sup> This observation is indeed consistent with a recent report suggesting that more acidic arylboronic acids tend to form the  $sp^3$  trihydroxyborate, a “side-product” shown to be less reactive compared to a neutral ( $sp^2$ ) boronic acid, thereby shifting the equilibrium to the left and lowering boronate stability (Scheme 2-9).<sup>11g</sup> Moreover, the *ortho*-methyl substituent slowed down boronate formation (lower  $k_{ON}$ ) while *ortho*-halogens and especially a cyano group accelerated it to a significant measure (Table 2-2, entry 1–3). The proposed reaction mechanism starts with the exchange of one of the two hydroxyl groups of the boronic acid, which is likely the rate-determining step, followed by a fast cyclization to form the Lewis acidic boronic ester. Moreover,  $^{11}B$  NMR analysis of boronates **2-5a** and **2-5b** (1 mM) in 0.1 M of  $D_2O$  potassium phosphate buffer:ACN (65:35 w%, pD 7.4) afforded the resonances at 0.06 and  $-0.93$  ppm, respectively. This result indicates that boronate products rest in the form of tetrahedral hydroxyboronate ions (Scheme 2-9).

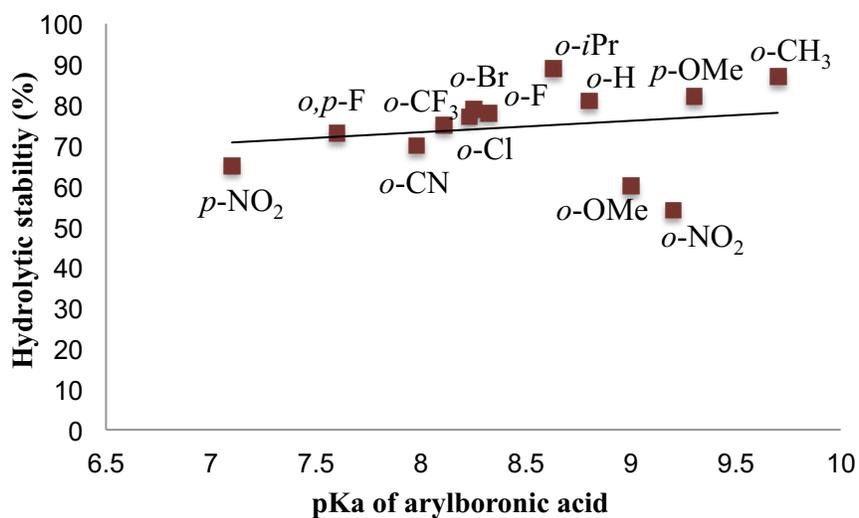
**Table 2-1:** Hydrolytic stability and pKa values of selected arylboronic acids with diol **2-2a**.

Entry	Arylboronic acid	Hydrolytic stability (1 mM) <b>2-4/2-2a</b>	pKa values
1		<b>2-4a/2-2a:</b> 87/13	9.7 <sup>[a]</sup>
2		<b>2-4b/2-2a:</b> 78/22	8.32 ± 0.58 <sup>[b]</sup>
3		<b>2-4c/2-2a:</b> 70/30	7.98 ± 0.58 <sup>[b]</sup>

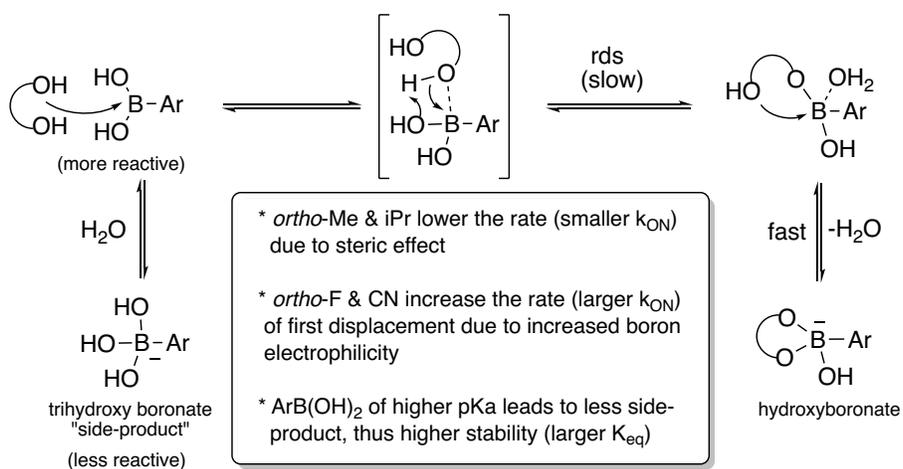
Entry	Arylboronic acid	Hydrolytic stability (1 mM) <b>2-4/2-2a</b>	pKa values
4		<b>2-4d/2-2a:</b> 81/19	8.8 <sup>[a]</sup>
5		<b>2-4e/2-2a:</b> 60/40	9.0 <sup>[a]</sup>
6		<b>2-4f/2-2a:</b> 89/11	8.63 ± 0.58 <sup>[b]</sup>
7		<b>2-4g/2-2a:</b> 77/23	8.23 ± 0.58 <sup>[b]</sup>
8		<b>2-4h/2-2a:</b> 79/21	8.25 ± 0.58 <sup>[b]</sup>
9		<b>2-4i/2-2a:</b> 75/25	8.11 ± 0.58 <sup>[b]</sup>
10		<b>2-4j/2-2a:</b> 75/25	9.2 <sup>[a]</sup>
11		<b>2-4k/2-2a:</b> 65/35	7.1 <sup>[a]</sup>

Entry	Arylboronic acid	Hydrolytic stability (1 mM) <b>2-4/2-2a</b>	pKa values
12		<b>2-4l/2-2a</b> : 82/18	9.3 <sup>[a]</sup>
13		<b>2-4m/2-2a</b> : 73/27	7.6 <sup>[c]</sup>

[a] pKa values of arylboronic acids were obtained from the literature.<sup>24</sup> [b] Estimated pKa values of arylboronic acids were obtained from Advanced Chemistry Development (ACD/Labs) Software calculations accessed *via* SciFinder. [c] pKa values of arylboronic acids were obtained from a previous report.<sup>11d</sup>



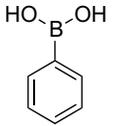
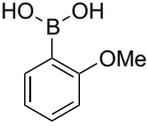
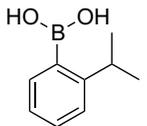
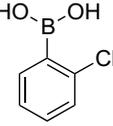
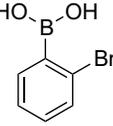
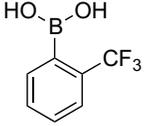
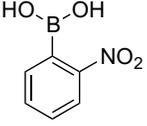
**Figure 2-11:** Linear correlation of hydrolytic stability vs pKa values of arylboronic acids.

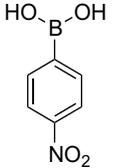
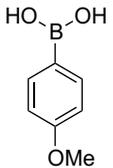
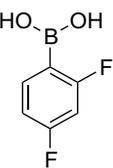
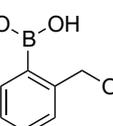
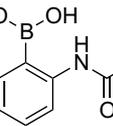
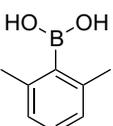


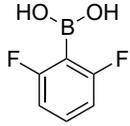
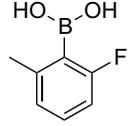
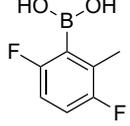
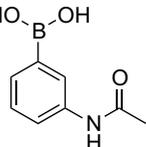
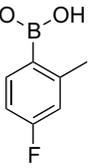
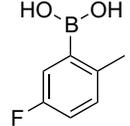
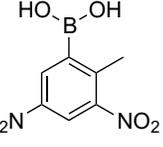
**Scheme 2-9:** Proposed mechanism for the boronate formation.

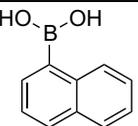
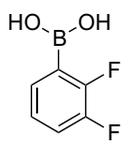
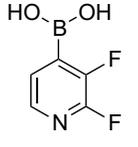
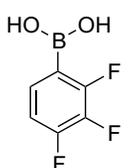
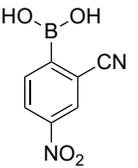
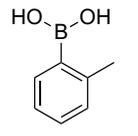
**Table 2-2:** Results of hydrolytic stability study, forward rate constant ( $k_{ON}$ ,  $M^{-1}s^{-1}$ ) and association constant measurements ( $K_{eq}$ ,  $M^{-1}$ ).

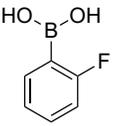
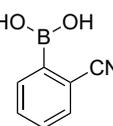
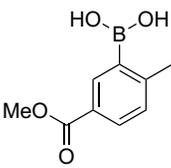
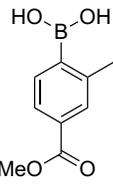
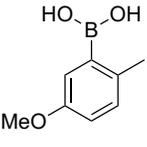
Entry	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM)	Hydrolytic stability (1 mM)	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} = M^{-1}s^{-1}$ )
1 (2-2a)		5 mM 2-4a/2-2a: 94/6	2-4a/2-2a: 87/13	—	$1.6 \pm 0.1$
2 (2-2a)		5 mM 2-4b/2-2a: 89/11	2-4b/2-2a: 78/22	—	$18 \pm 4$
3 (2-2a)		5 mM 2-4c/2-2a: 87/13	2-4c/2-2a: 70/30	—	$> 50^{[a]}$

Entry (2-2a/ 2-2b)	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM) 2-4/2-2a or 2-5/2-2b	Hydrolytic stability (1 mM) 2-4/2-2a or 2-5/2-2b	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} =$ $M^{-1}s^{-1}$ )
4 (2-2a)		10 mM 2-4d/2-2a: 94/6	2-4d/2-2a: 81/19	—	$2.0 \pm 0.1$
5 (2-2a)		10 mM 2-4e/2-2a: 86/14	2-4e/2-2a: 60/40	—	$2.0 \pm 0.8$
6 (2-2a)		10 mM 2-4f/2-2a: 96/4	2-4f/2-2a: 89/11	—	Very slow (It was not plotted)
7 (2-2a)		10 mM 2-4g/2-2a: 93/7	2-4g/2-2a: 77/23	—	$11 \pm 3$
8 (2-2a)		5 mM 2-4h/2-2a: 91/9	2-4h/2-2a: 79/21	—	$9.0 \pm 0.1$
9 (2-2a)		5 mM 2-4i/2-2a: 90/10	2-4i/2-2a: 75/25	—	Very slow
10 (2-2a)		10 mM 2-4j/2-2a: 82/18	2-4j/2-2a: 54/46	—	Very slow

Entry	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM)	Hydrolytic stability (1 mM)	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} = M^{-1}s^{-1}$ )
(2-2a/ 2-2b)		2-4/2-2a or 2-5/2-2b	2-4/2-2a or 2-5/2-2b		
11 (2-2a)		10 mM 2-4k/2-2a: 88/12	2-4k/2-2a: 65/35	—	—
12 (2-2a)		10 mM 2-4l/2-2a: 94/6	2-4l/2-2a: 82/18	—	$4.0 \pm 0.2$
13 (2-2a)		10 mM 2-4m/2-2a: 92/8	2-4m/2-2a: 73/27	—	$23 \pm 4$
14 (2-2a)		10 mM 2-4n/2-2a: 92/8	2-4n/2-2a: 76/24	—	—
15 (2-2a)		5 mM 2-4o/2-2a: 84/16	2-4o/2-2a: 67/33	—	Very slow
16 (2-2a)		5 mM 2-4p/2-2a: 100/0	2-4p/2-2a: 100/0	—	Very slow

Entry	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM)	Hydrolytic stability (1 mM)	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} = M^{-1}s^{-1}$ )
(2-2a/ 2-2b)		2-4/2-2a or 2-5/2-2b	2-4/2-2a or 2-5/2-2b		
17 (2-2a)		5 mM 2-4q/2-2a: 77/23	2-4q/2-2a: 52/48	—	—
18 (2-2a)		10 mM 2-4r/2-2a: 91/9	2-4r/2-2a: 74/26	—	$0.3 \pm 0.1$
19 (2-2a)		5 mM 2-4s/2-2a: 86/14	2-4s/2-2a: 69/31	—	$0.7 \pm 0.1$
20 (2-2a)		10 mM 2-4t/2-2a: 89/11	2-4t/2-2a: 74/26	—	$12 \pm 5$
21 (2-2a)		5 mM 2-4u/2-2a: 90/10	2-4u/2-2a: 85/15	—	$2.0 \pm 0.8$
22 (2-2a)		5 mM 2-4v/2-2a: 88/12	2-4v/2-2a: 82/18	—	$4 \pm 1$
23 (2-2a)		5 mM 2-4w/2-2a: 82/18	2-4w/2-2a: 68/32	—	$36 \pm 12$

Entry (2-2a/ 2-2b)	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM) 2-4/2-2a or 2-5/2-2b	Hydrolytic stability (1 mM) 2-4/2-2a or 2-5/2-2b	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} =$ $M^{-1}s^{-1}$ )
24 (2-2a)		5 mM 2-4x/2-2a: 93/7	2-4x/2-2a: 87/13	–	$4.0 \pm 0.1$
25 (2-2a)		5 mM 2-4w/2-2a: 91/9	2-4w/2-2a: 78/22	–	$36 \pm 2$ (0 °C)
26 (2-2a)		5 mM 2-4z/2-2a: 63/37	2-4z/2-2a: 52/48	–	–
27 (2-2a)		10 mM 2-4aa/2-2a: 88/12	2-4aa/2-2a: 70/30	–	$> 50^{[b]}$
28 (2-2a)		10 mM 2-4ab/2-2a: 87/13	2-4ab/2-2a: 64/36	–	$> 50^{[b]}$
29 (2-2a)		5 mM 2-4ac/2-2a: 80/20	2-4ac/2-2a: 52/48	–	$21 \pm 2$
30 (2-2b)		0.1 mM 2-5a/2-2b: 81/19	2-5a/2-2b: 93/7	$180 \pm$ $28 \times$ $10^3$	$2.3 \pm 0.2$

Entry	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM)	Hydrolytic stability (1 mM)	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} = M^{-1}s^{-1}$ )
(2-2a/ 2-2b)		2-4/2-2a or 2-5/2-2b	2-4/2-2a or 2-5/2-2b		
31 (2-2b)		—  5 mM	2-5b/2-2b: 84/16	27 ±  $6 \times 10^3$	33 ± 2
32 (2-2b)		2-5c/2-2b: 92/8 0.1 mM 2-5c/2-2b: 54/46	2-5c/2-2b: 83/17	25 ±  $3 \times 10^3$	> 50 <sup>[b]</sup>
33 (2-2b)		0.1 mM 2-5d/2-2b: 75/25	2-5d/2-2b: 91/9	120 ±  $6 \times 10^3$	6.9 ± 0.6 for 1 mM; at ~10 min it reaches equilibrium, 85%
34 (2-2b)		—	2-5e/2-2b: 90/10	91 ±  $6 \times 10^3$	7.8 ± 0.7
35 (2-2b)		—	2-5f/2-2b: 92/8	130 ±  $10 \times 10^3$	3.3 ± 0.6

Entry	Arylboric acid	Hydrolytic stability (10, 5 or 0.1 mM)	Hydrolytic stability (1 mM)	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} = M^{-1}s^{-1}$ )
(2-2a/ 2-2b)		2-4/2-2a or 2-5/2-2b	2-4/2-2a or 2-5/2-2b		
36 (2-2b)		—	2-5g/2-2b: 94/6	$330 \pm 50 \times 10^3$	$1.0 \pm 0.2$  $> 50^{[b]}$ for 0.5 mM.
37 (2-2b)		—	2-5h/2-2b: 75/25	$15 \pm 1 \times 10^3$	It reaches equilibrium at ~1 min (65% conversion).
38 (2-2b)		—	2-5i/2-2b: 82/18	$25 \pm 2 \times 10^3$	$> 50^{[b]}$
39 (2-2b)		0.1 mM 2-5j/2-2b: 61/39	2-5j/2-2b: 85/15	$40 \pm 2 \times 10^3$	$18 \pm 1$
40 (2-2b)		—	2-5k/2-2b: 78/22	$15 \pm 1 \times 10^3$	$> 50^{[b]}$

Entry ( <b>2-2a</b> / <b>2-2b</b> )	Arylboronic acid	Hydrolytic stability (10, 5 or 0.1 mM) <b>2-4/2-2a</b> or <b>2-5/2-2b</b>	Hydrolytic stability (1 mM) <b>2-4/2-2a</b> or <b>2-5/2-2b</b>	$K_{eq}$ ( $M^{-1}$ )	Forward rate constant ( $k_{ON} =$ $M^{-1}s^{-1}$ )
41 ( <b>2-2b</b> )		0.1 mM <b>2-5l/2-2b</b> : 70/30	<b>2-5l/2-2b</b> : 87/13	$56 \pm 7$ $\times 10^3$	$2.4 \pm 0.9$
42 ( <b>2-2b</b> )		0.1 mM <b>2-5m/2-2b</b> : 86/14	<b>2-5m/2-2b</b> : 92/8	–	$0.20 \pm 0.03$
43 ( <b>2-2b</b> )		0.1 mM <b>2-5n/2-2b</b> : 55/45	<b>2-5n/2-2b</b> : 81/19	–	$18 \pm 1$
44 ( <b>2-2b</b> )		0.1 mM <b>2-5o/2-2b</b> : 63/37	<b>2-5o/2-2b</b> : 83/17	$42 \pm 3$ $\times 10^3$	$1.4 \pm 0.2$
45 ( <b>2-2b</b> )		0.1 mM <b>2-5p/2-2b</b> : 84/16	<b>2-5p/2-2b</b> : 93/17	$270 \pm$ $18 \times$ $10^3$	$0.8 \pm 0.2$

[a] This reaction showed 50% conversion in 30 seconds. It slowed down at 0 °C and provided  $49 \pm 2 M^{-1}s^{-1}$ . [b] This reaction could not be monitored due to its very fast conversion.

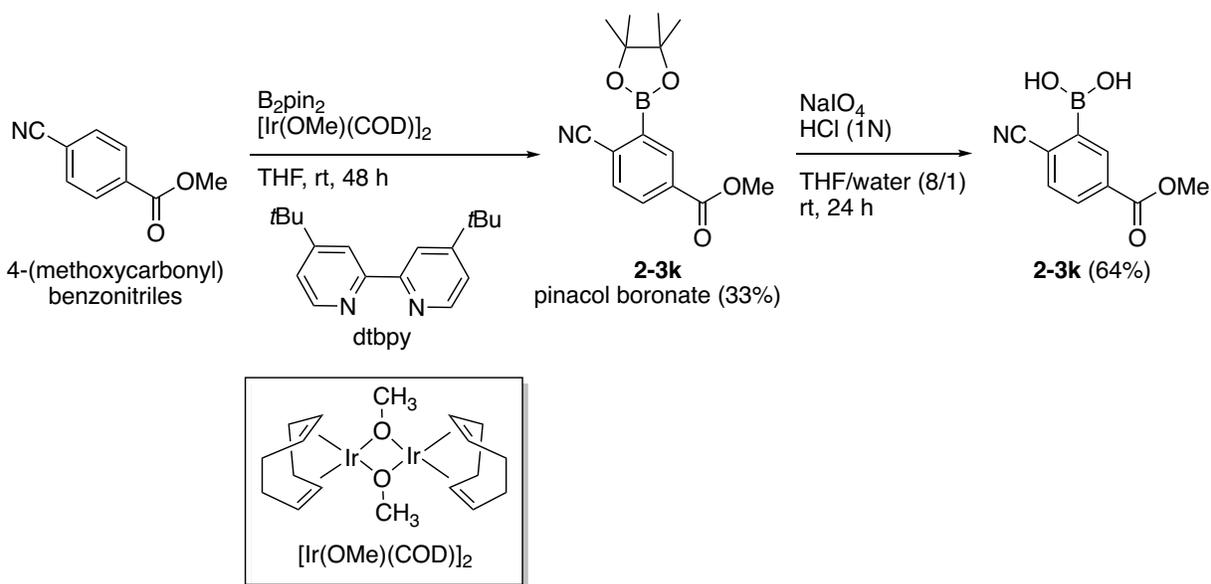
Our initial NMR-based screening with diol **2-2a** led to a selection of three promising boronic acids (data indicated in Table 2-3, entries 1–3). These lead reagents are *ortho*-tolylboronic acid **2-3a**, which provided better boronate stability than phenylboronic acid (Table 2-1 and 2-2, entries 1 and 4), *ortho*-fluorophenylboronic acid **2-3b** and *ortho*-

cyanophenylboronic acid **2-3c**, which both showed excellent  $k_{\text{ON}}$  values. The reaction of diol **2-2a** and boronic acid **2-3c** showed ~50% conversion in 30 seconds; however it slowed down at 0 °C and provided  $k_{\text{ON}} = 49 \pm 2 \text{ M}^{-1}\text{s}^{-1}$ . Also, the forward rate constant  $> 50 \text{ M}^{-1}\text{s}^{-1}$  indicates that these reactions could not be monitored by the NMR method due to their very fast conversion.

Brown and co-worker investigated the effect of a number of nopoldiol derivatives having coordinating sites, such as -OH, -NH<sub>2</sub> and -OMe, on the stability of their resulting boronic esters.<sup>14</sup> These coordinating groups are placed in the same position as that occupied by the closest PEG group oxygen atom on diol **2-2a**. This study showed that these nopoldiol derivatives afforded less stable boronic esters compared to nopoldiols having non-coordinating moieties (alkyl, aryl, chloride).<sup>14</sup> Based on this observation, I suspected that the proximal oxygen of **2-2a** might be detrimental by assisting hydrolysis. Thus, another nopol-PEG-diol **2-2b** was synthesized, as described earlier (c.f. Scheme 2-3), to examine this potential issue on the stability and  $k_{\text{ON}}$ . Boronic acids **2-3a** – **2-3c** were then tested with diol **2-2b**. To our satisfaction, their hydrolytic stability and  $k_{\text{ON}}$  both improved (Table 2-3, entries 4–6). Other arylboronic acids were also evaluated with diol **2-2b**, but **2-3a** – **2-3c** remained superior (Table 2-2). I then proceeded to design, synthesize and evaluate conjugatable derivatives of **2-3a** – **2-3c** (Table 2-3, entries 7–14). The synthesis of boronic acids **2-3d-g,i,j** was achieved by lithiation-borylation of their brominated precursors as described in the experimental section.<sup>25</sup> Boronic acid **2-3k** was synthesized *via* iridium-catalyzed *ortho*-directed C–H borylation of 4-(methoxycarbonyl)benzotrile, a method reported by Smith and co-workers, followed by deprotection of the resulting pinacol boronate **2-3k** in the presence of NaIO<sub>4</sub> (Scheme 2-10).<sup>26</sup>  $K_{\text{eq}}$  of the new boronic acids and diol **2-2b** were also determined *via* <sup>1</sup>H NMR along with their forward rate constants ( $k_{\text{ON}}$ ). Consistent with our initial screening results, the presence of an electron-withdrawing group (CO<sub>2</sub>Me or CONMe<sub>2</sub>) at either *meta*- or *para*- positions enhanced the reactivity compared to boronic acids **2-3a** – **2-3c**, and a decrease in  $K_{\text{eq}}$  was noted (Table 2-3, entry 4 vs 7,8; 5 vs 11).

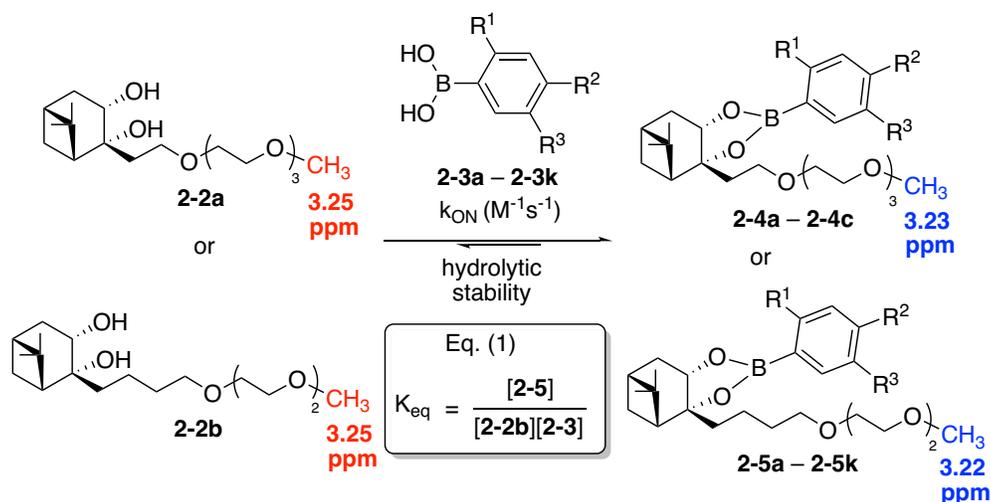
On the other hand, whereas an electron-donating group (OMe) impeded the reactivity (lower  $k_{\text{ON}}$ ), it improved  $K_{\text{eq}}$  specifically when placed in the *para* position (Table 2-3, compare entry 4 vs 9,10; 5 vs 12,13). Even though *ortho*-fluoro derivatives **2-3h** – **2-3j** enabled high reactivity, their lower hydrolytic stability caused concerns towards their use under the conditions

of bioorthogonal conjugation at micromolar concentrations. Boronic acid **2-3k** was eliminated because of its lower  $K_{eq}$  and its susceptibility to deboronation. In the end, *ortho*-tolylboronic acid derivative, specifically boronic acid **2-3d**, was selected on the grounds of its high reactivity and the hydrolytic stability of the resulting boronate. Yet because the reaction of *ortho*-fluorophenylboronic acid derivative **2-3h** was too fast to be monitored by NMR, I decided to design a more sensitive fluorescence quenching assay in order to compare the rate constants of **2-3d** and **2-3h**.



**Scheme 2-10:** Synthesis of boronic acid **2-3k**.<sup>26</sup>

**Table 2-3:** Selected results of hydrolytic stability study, forward rate constant ( $k_{\text{ON}}$ ,  $\text{M}^{-1}\text{s}^{-1}$ ) and association constant measurement ( $K_{\text{eq}}$ ,  $\text{M}^{-1}$ ).



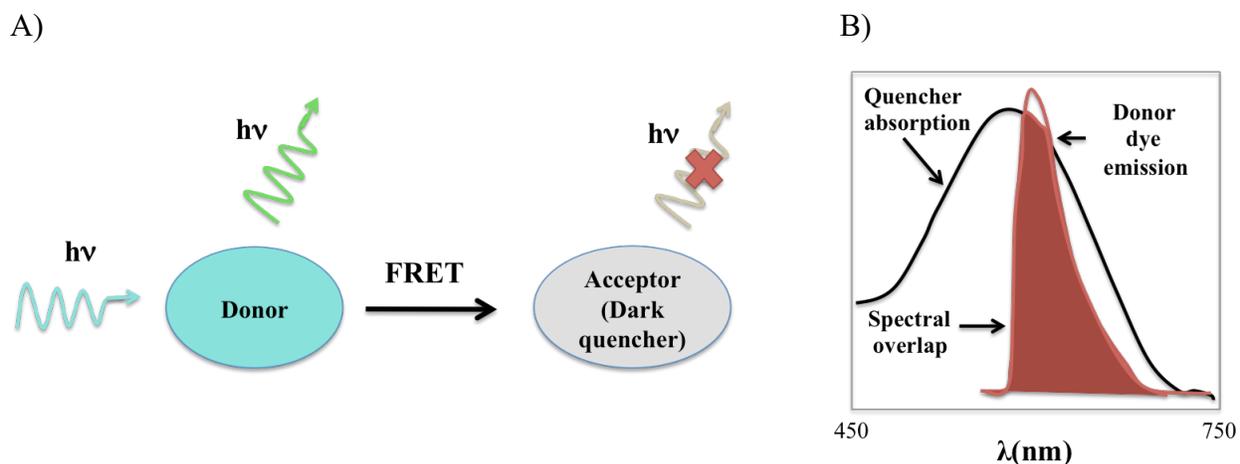
Entry	Arylboronic acid	Hydrolytic stability <sup>[a]</sup>	$K_{\text{eq}} \times 10^3$ ( $\text{M}^{-1}$ )	$k_{\text{ON}}$ ( $\text{M}^{-1}\text{s}^{-1}$ )
(2-2a/2-2b)	2-3a – 2-3k: $\text{R}^1, \text{R}^2, \text{R}^3$	(2-4/2-2a or 2-5/2-2b)		
1 (2-2a)	2-3a: Me, H, H	2-4a/2-2a: 87/13	–	$1.6 \pm 0.1$
2 (2-2a)	2-3b: F, H, H	2-4b/2-2a: 78/22	–	$18 \pm 4$
3 (2-2a)	2-3c: CN, H, H	2-4c/2-2a: 70/30	–	$> 50$
4 (2-2b)	2-3a: Me, H, H	2-5a/2-2b: 93/7	180	$2.3 \pm 0.2$
5 (2-2b)	2-3b: F, H, H	2-5b/2-2b: 84/16	27	$33 \pm 2$
6 (2-2b)	2-3c: CN, H, H	2-5c/2-2b: 83/17	25	$> 50$
7 (2-2b)	2-3d: Me, H, $\text{CO}_2\text{Me}$	2-5d/2-2b: 91/9	120	$7.7 \pm 0.5$
8 (2-2b)	2-3e: Me, $\text{CO}_2\text{Me}$ , H	2-5e/2-2b: 90/10	91	$7.8 \pm 0.7$
9 (2-2b)	2-3f: Me, H, OMe	2-5f/2-2b: 92/8	130	$3.3 \pm 0.6$
10 (2-2b)	2-3g: Me, OMe, H	2-5g/2-2b: 94/6	330	$1.0 \pm 0.2$

Entry	Arylboronic acid	Hydrolytic stability <sup>[a]</sup>	$K_{eq} \times 10^3$	$k_{ON}$
(2-2a/2-2b)	2-3a – 2-3k: R <sup>1</sup> , R <sup>2</sup> , R <sup>3</sup>	(2-4/2-2a or 2-5/2-2b)	(M <sup>-1</sup> ) <sup>[c]</sup>	(M <sup>-1</sup> s <sup>-1</sup> )
11 (2-2b)	2-3h: F, H, CONMe <sub>2</sub>	2-5h/2-2b: 75/25	15	> 50
12 (2-2b)	2-3i: F, H, OMe	2-5i/2-2b: 82/18	25	> 50
13 (2-2b)	2-3j: F, OMe, H	2-5j/2-2b: 85/15	40	18 ± 1
14 (2-2b)	2-3k: CN, H, CO <sub>2</sub> Me	2-5k/2-2b: 78/22	12	> 50

[a] Hydrolytic stability of boronates 2-4/2-5 was studied at 1 mM concentration *via* <sup>1</sup>H NMR in 0.1 M D<sub>2</sub>O phosphate buffer (pD 7.4): ACN-d<sub>3</sub> (65:35). The integral ratio of CH<sub>3</sub>/CH<sub>3</sub> was monitored after 24 h and recorded as 2-4/2-2a or 2-5/2-2b.

#### 2.2.1.4 Fluorescence quenching kinetic study – Förster Resonance Energy Transfer (FRET) quenching

Fluorescence quenching involves any process that reduces the fluorescence intensity of a fluorophore and provides the basis for Förster Resonance Energy Transfer (FRET).<sup>27</sup> The mechanism of energy transfer between two light-sensitive molecules (donor and acceptor) is described as FRET. In the absence of an acceptor molecule, a typical donor fluorophore, which is excited at a particular wavelength, emits light while returning from its excited state to ground state. However, in the presence of an acceptor molecule, the excited donor fluorophore may transfer its energy to the acceptor molecule through non-radiative dipole-dipole coupling without the emission of light. Therefore, the acceptor fluorophore reaches its excited state and emits light to return to its ground state. The acceptor molecule can also be an excited dark quencher, which comes back to the ground state without the emission of light.<sup>28</sup> This is also described as a FRET mechanism, however dark quenchers cannot emit light and seemingly decrease the fluorescent intensity of the donor fluorophore (Figure 2-12A). In order for FRET quenching to occur efficiently, the emission spectrum of a donor fluorophore must overlap with the excitation spectrum of a dark quencher (Figure 2-12B).<sup>27</sup> Moreover, this sensitive process is highly dependent on the distance (~10–100 Å) between donor and dark quencher, and inversely proportional to the sixth power of this distance (equation 1).



**Figure 2-12:** A) Förster Resonance Energy Transfer (FRET) quenching mechanism between a donor fluorophore and a dark quencher. B) Donor emission and dark quencher absorption spectral overlap.

$$E = \frac{1}{1 + \frac{R^6}{R_o^6}} \quad (1)$$

\* E is FRET efficiency.

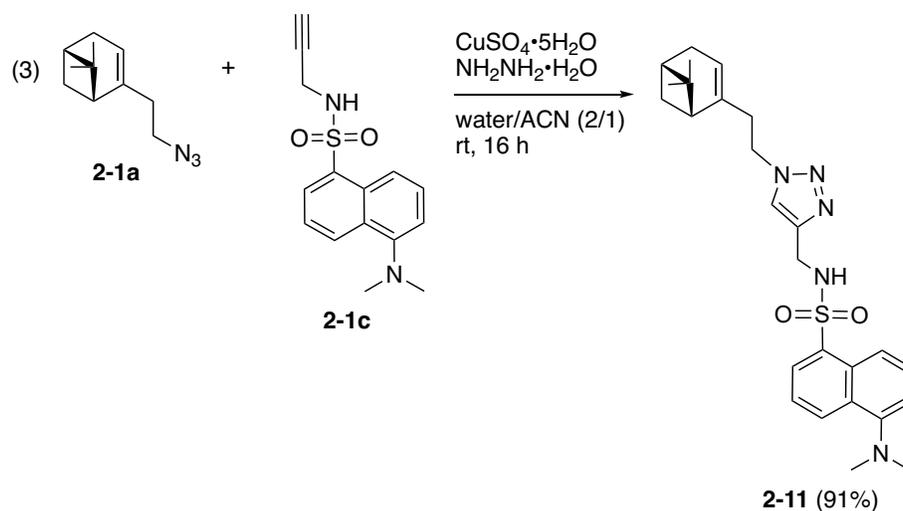
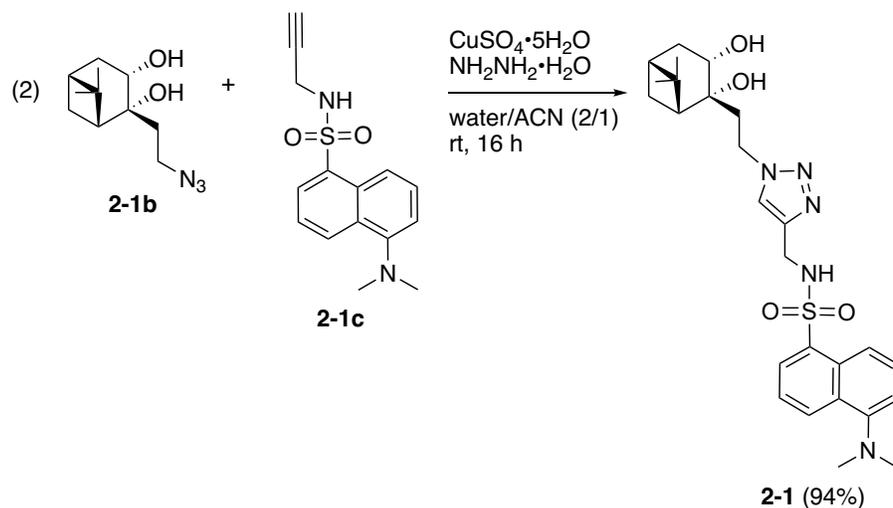
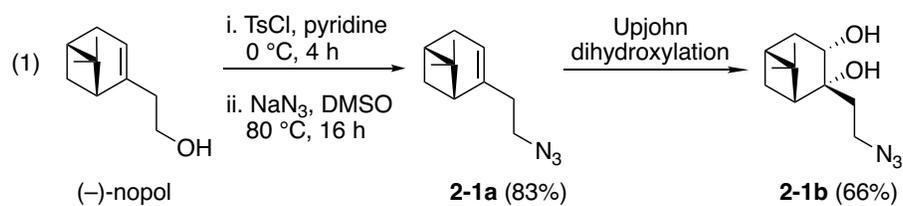
\*  $R_o$  is the **Förster critical distance** at which 50% of the excitation energy is transferred to the acceptor.

\*  $R$  is the experimental distance between donor and dark quencher.

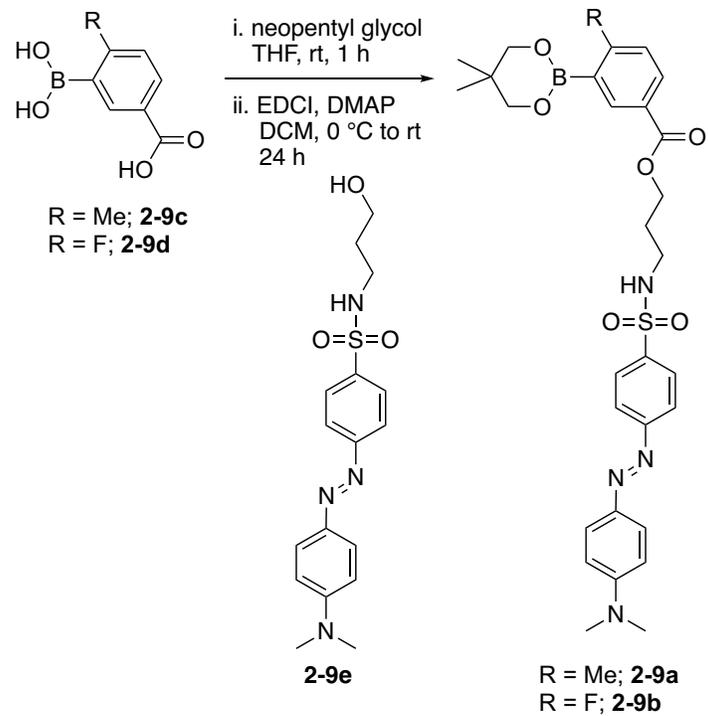
The FRET quenching method can be used to monitor a reaction and to determine its rate constant/ $K_{eq}$ . For instance, a donor fluorophore and a dark quencher may separately be conjugated to two reactants, and FRET quenching may take place efficiently upon the conjugation of these reactants, thereby reducing the fluorescence intensity of the donor fluorophore. This information can be used to determine the rate constant/ $K_{eq}$  of the reaction. In this regard, readily available dansyl fluorophore (donor) and a dabcyll dark quencher (acceptor) were chosen to monitor the boronate formation due to the spectral overlap of dansyl emission (maximum wavelength 520 nm) and dabcyll absorption (maximum wavelength 475 nm).<sup>28</sup> With this, a dansyl-tagged diol **2-1** and dabcyll derivatives of the corresponding neopentyl boronates **2-9a/2-9b** were synthesized as shown in Scheme 2-11 and 2-12, respectively. Dansyl-tagged diol

**2-1** was prepared *via* CuAAC of propargyl dansyl derivative **2-1c** and nopoldiol azide **2-1b**, which was synthesized from Upjohn dihydroxylation of nopol azide **2-1a** (Scheme 2-11, equations 1 and 2). Dabcyl derivatives of the corresponding neopentyl boronates **2-9a/2-9b** were obtained by the protection of boronic acids **2-9c/2-9d** with neopentyl glycol and esterification of the carboxyl group of the resulting boronates with dabcyl hydroxyl derivative **2-9e** (Scheme 2-12). Briefly, FRET quenching was expected to occur only upon the formation of boronates **2-10a-b** (Scheme 2-13A). Additionally nopol-dansyl without 1,2-*cis* diol **2-11** was designed as a negative control, which was not supposed to promote FRET quenching due to the absence of boronate formation (Scheme 2-11, equation 3 and Scheme 2-13B).

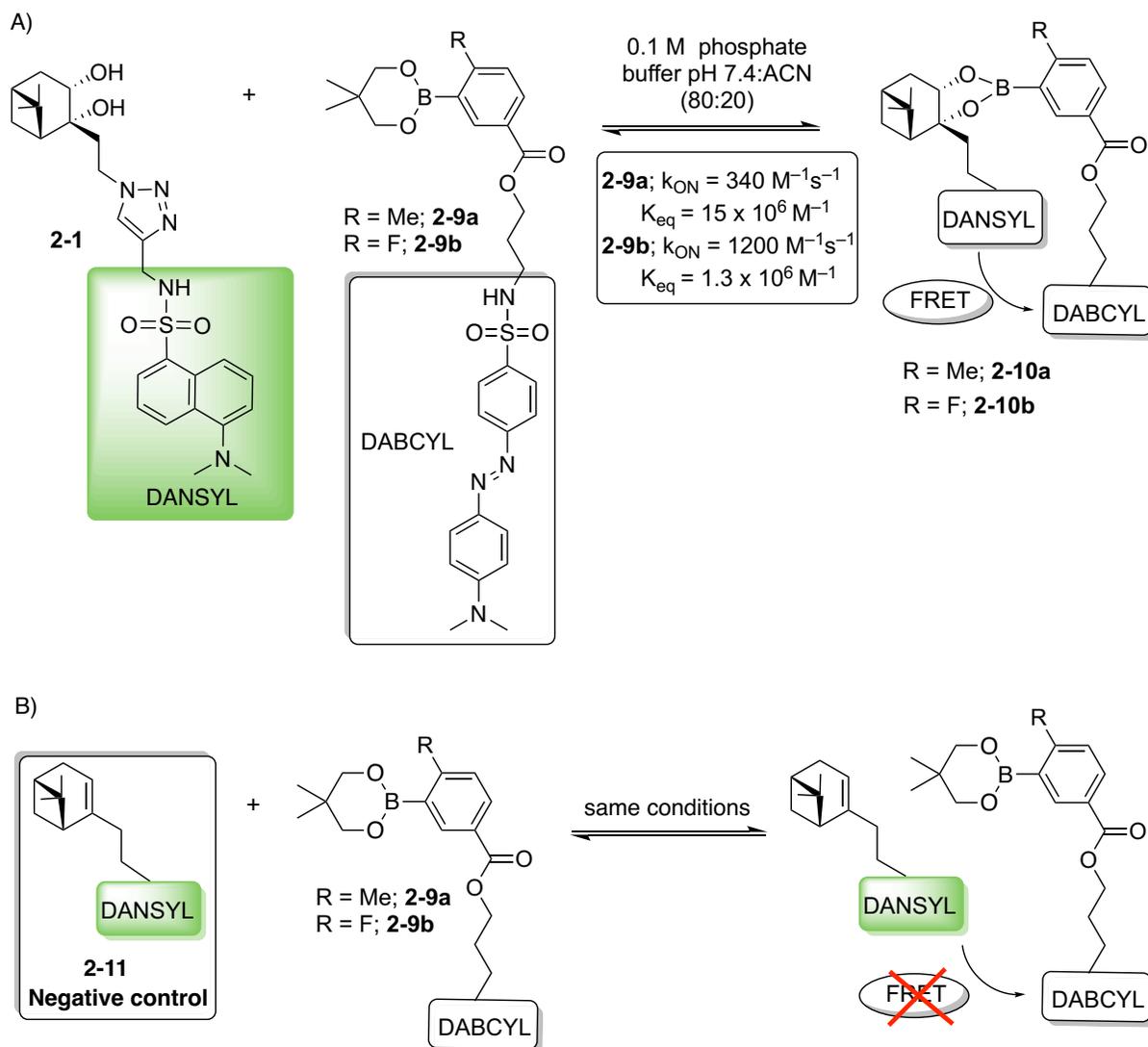
The fluorescence quenching experiments were performed using a TECAN fluorescence plate reader. The fluorescence was monitored at the maximum absorption of diol **2-1** (335 nm) in potassium phosphate buffer (0.1 M, pH 7.4)/ACN solution 80:20 w%. The quenching was monitored at 20 or 10  $\mu$ M over 10 min. Around 6% bleaching effect for diol **2-1** was observed at the concentration of 10  $\mu$ M. Since this bleaching effect was almost negligible, the data was not adjusted for bleaching. Moreover, as expected, negative control **2-11** preserved its fluorescent intensity in the presence of **2-9a/2-9b** due to its incapacity to form a boronic ester **2-10a/2-10b**.



**Scheme 2-11:** Synthesis of dansyl-tagged diol **2-1** and nopol-dansyl without 1,2-*cis* diol **2-11**.



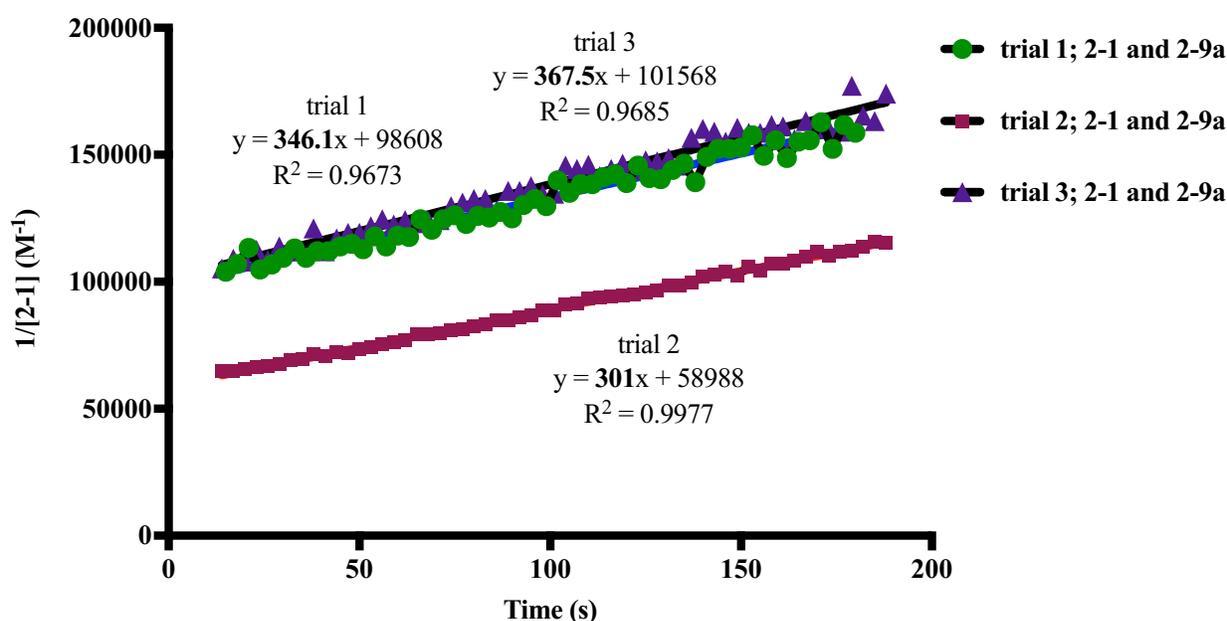
**Scheme 2-12:** Synthesis of dabcy1 derivatives of the corresponding neopentyl boronates **2-9a/2-9b**.



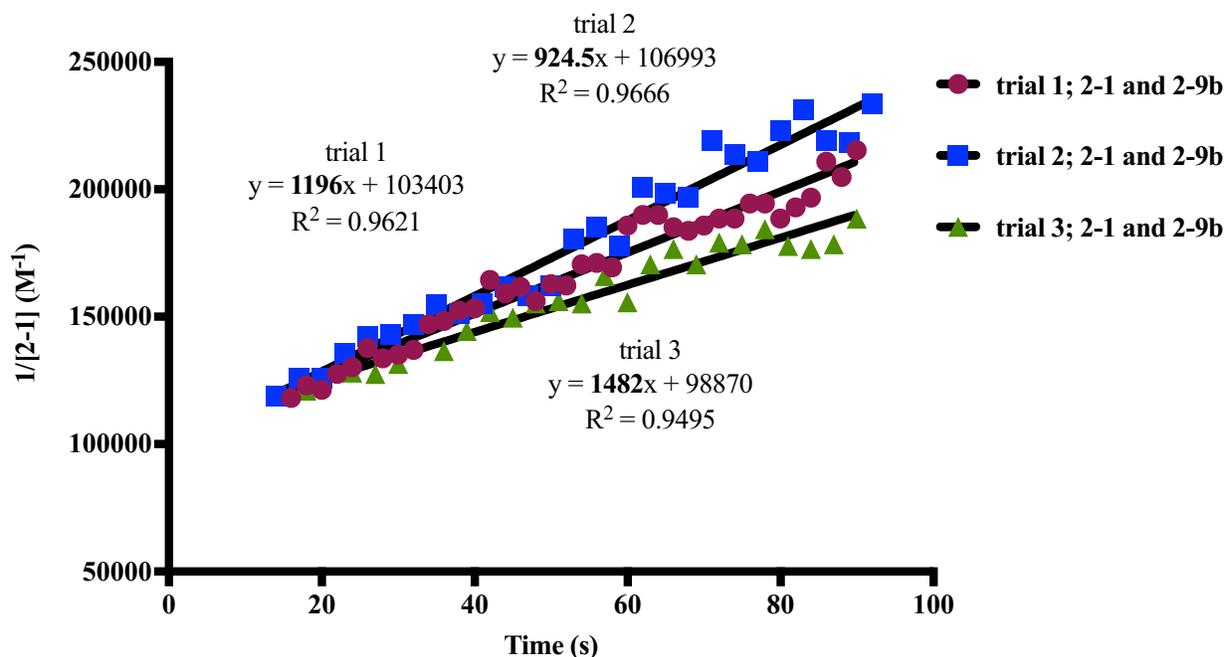
**Scheme 2-13:** Kinetic analysis using a fluorescence quenching experiment to measure the initial rate of boronate formation. Reported  $k_{ON}$  and  $K_{eq}$  are the average of at least 3 measurements.

Initial forward rate constants ( $k_{ON}$ ) were determined at 10 or 20  $\mu\text{M}$  concentration (Figure 2-13 and 2-14). The results (**2-9a**;  $338 \pm 34 \text{ M}^{-1}\text{s}^{-1}$  and **2-9b**;  $1201 \pm 279 \text{ M}^{-1}\text{s}^{-1}$ ) are significantly higher than that obtained *via*  $^1\text{H}$  NMR. The low intrinsic sensitivity of NMR makes it difficult to detect very low concentrations ( $< 100 \mu\text{M}$ ), therefore it requires higher reaction concentration at which fast processes ( $> 50 \text{ M}^{-1}\text{s}^{-1}$ ) are not readily monitored. Moreover, early rates ( $< 40 \text{ s}$ ) cannot be captured owing to the essential gradient shimming process. Thus, it is likely that the NMR method underestimated the actual rates of these fast reactions. Their rough  $K_{eq}$  were also determined by allowing the corresponding boronate products (**2-10a**, **2-10b**) to reach equilibrium

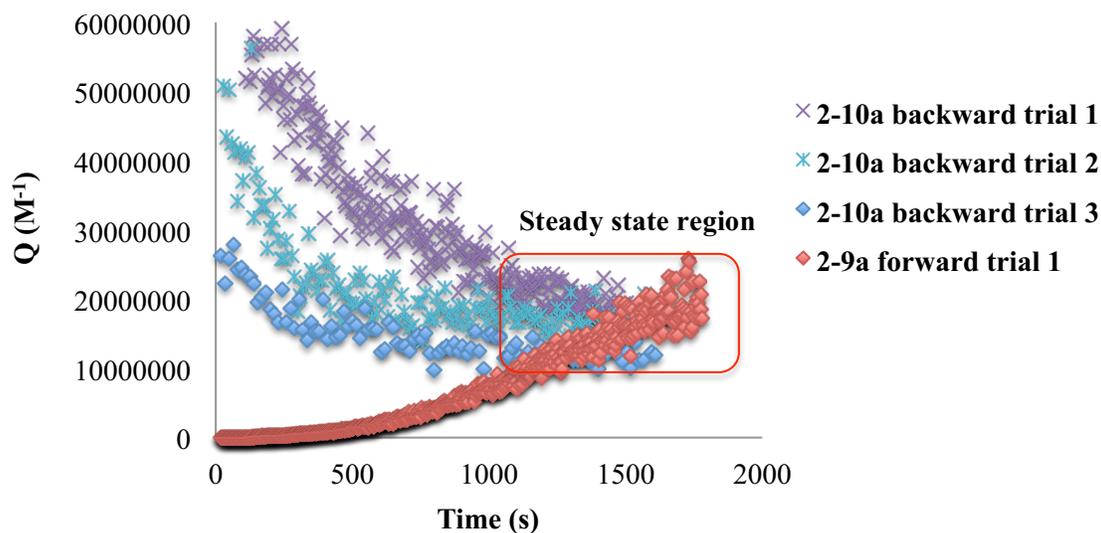
in the fluorescence quenching assay, giving  $K_{eq}$  values of  $\sim 15 \times 10^6$  and  $1.3 \times 10^6 \text{ M}^{-1}$  for *ortho*-methyl and *ortho*-fluoro derivatives, respectively (Figure 2-15 and 2-16). High  $K_{eq}$  values for substrates **2-1** and **2-9a/2-9b** relative to  $^1\text{H}$  NMR values (for **2-3d** and **2-3h**) may be due to hydrophobic and  $\pi$  interactions between fluorescent moieties, dansyl and dabcyI. Nonetheless, equilibration times obtained from both methods are consistent, with **2-9a/2-10a** taking about 20 min ( $k_{off} = k_{ON}/K_{eq}$ ;  $k_{off} = \sim 2 \times 10^{-5} \text{ s}^{-1}$ ) to reach a steady state with  $\sim 85\%$  of **2-10a** (from both directions, at 10  $\mu\text{M}$  concentration), whereas **2-9b/2-10b** needed only 2 min ( $k_{off} = \sim 90 \times 10^{-5} \text{ s}^{-1}$ ) to reach 60% of **2-10b** (Figure 2-15 and 2-16).



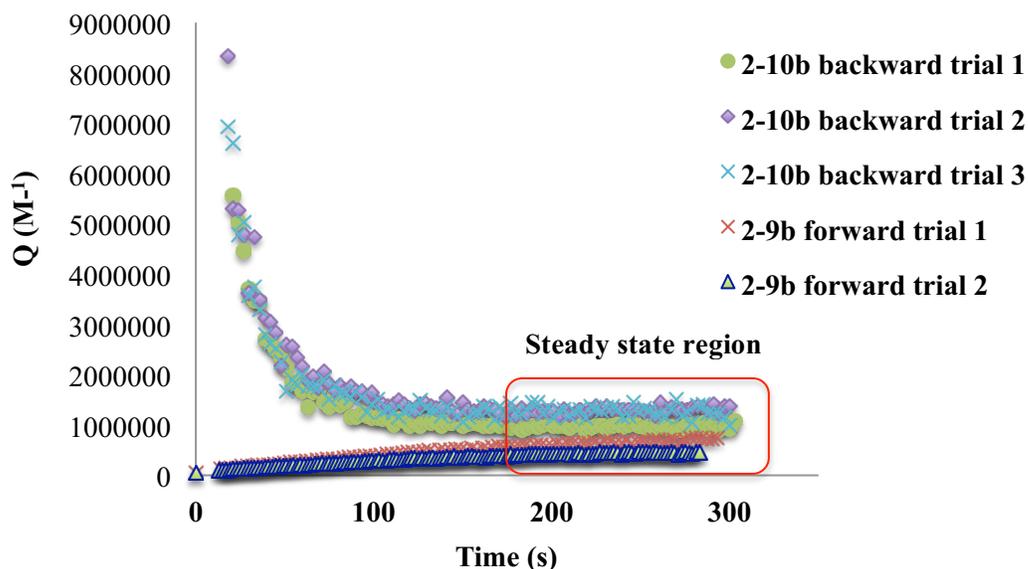
**Figure 2-13:** Second order plots for kinetic data of diol **2-1** and dabcyI-tagged boronate **2-9a** at 10 and 20  $\mu\text{M}$ . ( $k_{ON} = 338 \pm 34 \text{ M}^{-1}\text{s}^{-1}$ ). The straight line for each trial is a linear fit to the data using GraphPad Prism7 version 7.0d software.



**Figure 2-14:** Second order plots for kinetic data of diol **2-1** and dabcytl-tagged boronate **2-9b** at 10  $\mu\text{M}$ . ( $k_{\text{ON}} = 1201 \pm 279 \text{ M}^{-1}\text{s}^{-1}$ ). The straight line for each trial is a linear fit to the data using GraphPad Prism7 version 7.0d software



**Figure 2-15:** Estimation of  $K_{\text{eq}}$  from backward reaction of **2-10a**, and forward reaction of diol **2-1** and boronate **2-9a**.

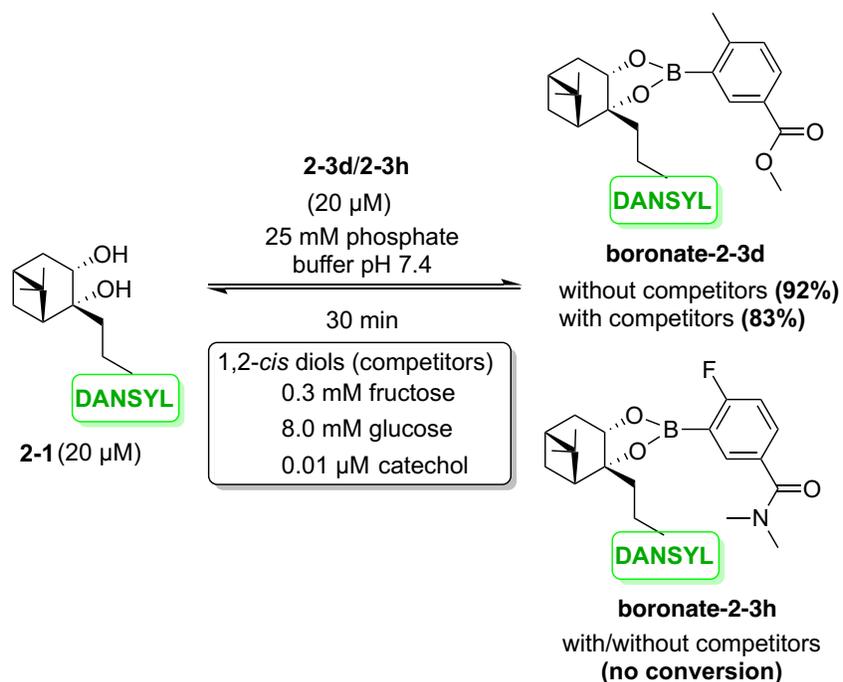


**Figure 2-16:** Estimation of  $K_{eq}$  from backward reaction of **2-10b**, and forward reaction of diol **2-1** and boronate **2-9b**.

### 2.2.2 Competitive effect of biological polyols on boronate formation *via* RP-HPLC

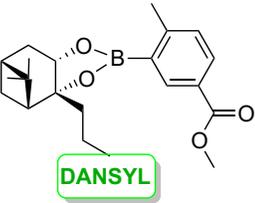
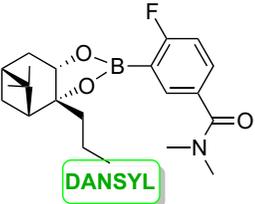
A biological competition assay was designed to determine whether the optimal nopoldiol boronates compete favorably with biological polyols such as glucose (4–7 mM), fructose (8  $\mu$ M) or catecholamines ( $\sim$ 0.0014  $\mu$ M) found in the blood stream.<sup>6a,29</sup> Thus, diol **2-1** was allowed to form boronates with **2-3d** and **2-3h** in the presence of a mixture of these biocompetitors used at concentrations higher than that found in the blood (Scheme 2-14, Table 2-4). This experiment was analyzed by HPLC (Figure 2-17 to 2-20). Fortunately, **2-3d** preserved its high affinity towards diol **2-1** in the presence of the biocompetitors (Table 2-4, entries 1–4). In contrast, **2-3h** showed none or little conversion even in the absence of the biocompetitors (Table 2-4, entries 5–8). Even though **2-3h** showed promise at higher concentrations, its conversion to boronate product is too low at micromolar concentrations to be useful in bioconjugation. This information clearly indicates that the boronate-**2-3h** exhibits lower  $K_{eq}$  compared to the boronate **2-10b**, which can provide a moderate conversion at 20  $\mu$ M concentration. As already observed during the optimization study, the  $K_{eq}$  is highly affected by the substituents present on arylboronic acids. In this case, the dabcyll group of boronic acid **2-9b** probably enhances the stability of boronate **2-**

**10b** via its hydrophobic and  $\pi$ - $\pi$  interactions with the dansyl moiety; however the boronate-**2-3h** cannot display such interactions. Furthermore, the formation of boronate **2-10b** was achieved in aqueous buffer:ACN (80:20) solution, whereas the competitive effect of biological polyols assay was performed in only aqueous buffer solution that might also induce further the hydrolysis of boronate-**2-3h**. Also, the backward reaction of boronate-**2-3h** is expected to be ~50 times faster than of boronate-**2-3d** based on fluorescence quenching studies (cf. page 81); therefore, boronate-**2-3h** might escape detection by undergoing a fast breakdown into its corresponding starting materials **2-1** and **2-3h** under the time scale of HPLC conditions.

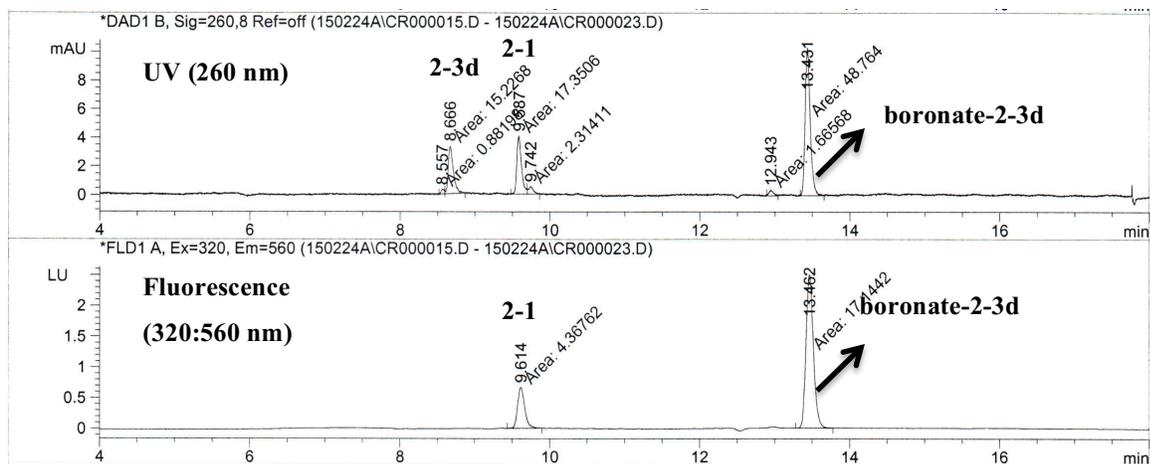


**Scheme 2-14:** Competitive effect of biological polyols on boronate formation. Boronate conversions were immediately monitored and determined by HPLC under both UV and fluorescence after diol **2-1** and **2-3d/2-3h** were allowed to react for 30 minutes.

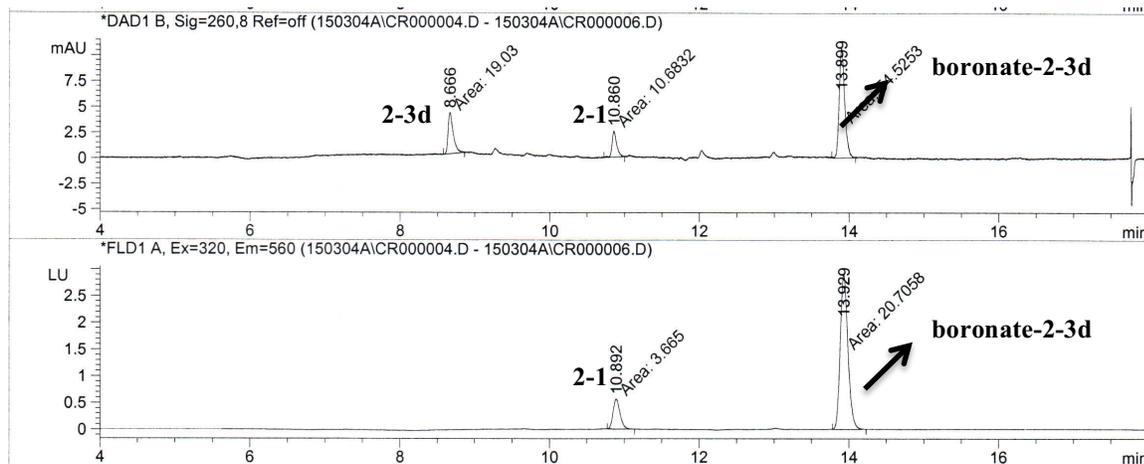
**Table 2-4:** Competitive effect of biological 1,2-*cis* diols on boronate formation.

Entry	Final concentration (μM)	1,2- <i>cis</i> diols <sup>[a]</sup>	Product	UV (260 nm) Conversion (%) <sup>[b]</sup>	Fluo (320 nm, 560 nm) Conversion (%) <sup>[b]</sup>
1	10	–		78	82
2	10	+		61	63
3	20	–		91	93
4	20	+		83	83
5	10	–			0 %
6	10	+	0 %		0 %
7	20	–	0 %		0 %
8	20	+	0 %		0 %
9	20 (2-11)	–	No boronate formation	0 %	0 %

[a] Glucose (8 mM), fructose (0.3 μM) and catechol (0.01 μM) were used as 1,2-*cis* diols. [b] Boronate conversions were immediately monitored and determined with HPLC under both UV and fluorescence after diol **2-1** and **2-3d/2-3h** were allowed to react for 30 min. Recorded conversions are the average value of duplicates. Corresponding peaks were identified with LC-MS traces.



**Figure 2-19:** HPLC chromatograms of diol **2-1** and **2-3d** (20  $\mu$ M) in the absence of 1,2-*cis* diols.



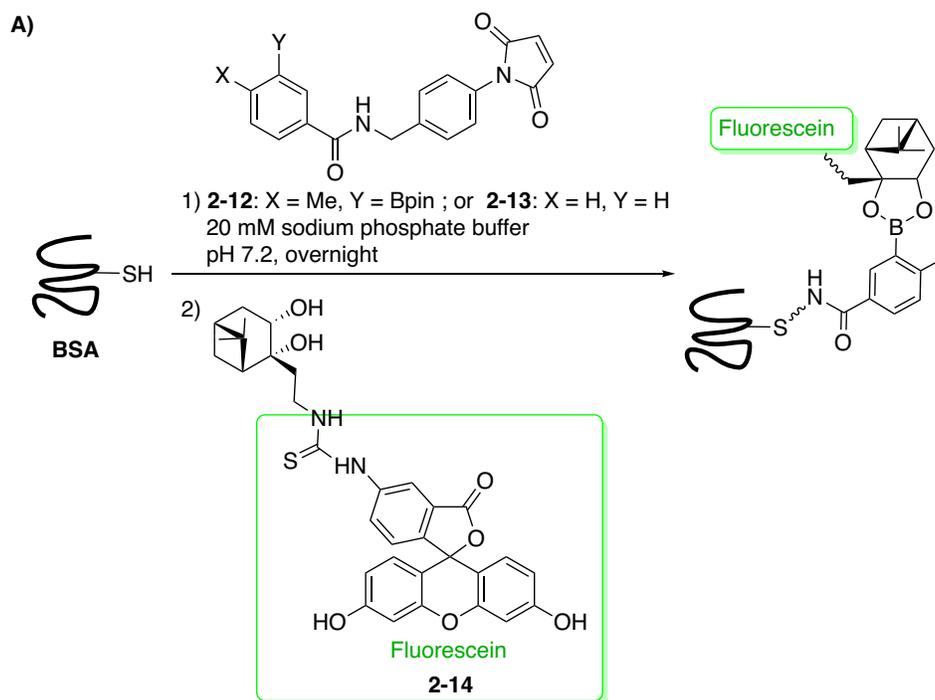
**Figure 2-20:** HPLC chromatograms of diol **2-1** and **2-3d** (20  $\mu$ M) in the presence of 1,2-*cis* diols.

### 2.2.3 Proof of concept on model proteins

#### 2.2.3.1 Labeling of boronic acid-BSA conjugate with fluorescein-derivatized diol *via* SDS PAGE

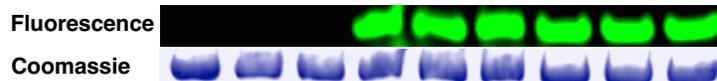
The optimal boronic acid derivative **2-3d**, which fulfills the criteria of bioorthogonality, reactivity and stability, was used to demonstrate boronate ligation on proteins. Towards this end, bovine serum albumin (BSA), which contains one free cysteine, was allowed to react with the boronyl-containing maleimide **2-12** to form a boronic acid-BSA adduct (Figure 2-21A). Then, the ‘click’ conjugation with fluorescein-derivatized diol **2-14** was conducted in both time and dose dependent fashion and monitored *via* gel electrophoresis and gel fluorescence imaging using 10% polyacrylamide gels (Figure 2-21B and 2-21C). It is important to mention that phosphate buffer was applied in gel preparation since the commonly used tris buffer system interfered with boronate formation and induced the breakdown of boronate conjugate (Figure 2-21D). Tris buffer is composed of tris(hydroxymethyl)aminomethane, which has 1,3-*cis* triol and also 1,2-*cis* alcohol amine; therefore the presence of these competing groups leads to the cleavage of boronate conjugate. Thus, using phosphate buffer mostly eliminated this interfering effect. Fluorescence was measured at 494 nm excitation and 512 nm emission wavelengths. Total protein loading was confirmed by subsequent staining with Coomassie blue. As depicted in

Figure 2-21B, the reaction between **2-14** and the boronic acid-BSA conjugate was complete within 5 minutes when 100  $\mu\text{M}$  of **2-14** was employed. Even at lower concentrations of **2-12** and **2-14** (20 and 40  $\mu\text{M}$ ), a significant level of conjugation was detected (Figure 2-21C). Expectedly, when a maleimide adduct **2-13** devoid of a boronic acid head was used, no labeling was detected.



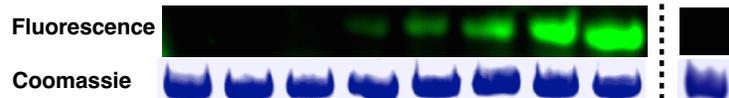
B) Time dependent

Reaction time (min)	120	120	120	5	15	30	60	90	120
<b>2-12</b> (100 $\mu\text{M}$ )	-	+	-	+	+	+	+	+	+
<b>2-14</b> (100 $\mu\text{M}$ )	+	-	+	+	+	+	+	+	+
<b>2-13</b> (100 $\mu\text{M}$ )	-	-	+	-	-	-	-	-	-



C) Dose dependent

<b>2-12</b> ( $\mu\text{M}$ )	-	100	-	20	40	60	80	100	100
<b>2-14</b> ( $\mu\text{M}$ )	100	-	100	20	40	60	80	100	200
<b>2-13</b> ( $\mu\text{M}$ )	-	-	100	-	-	-	-	-	-



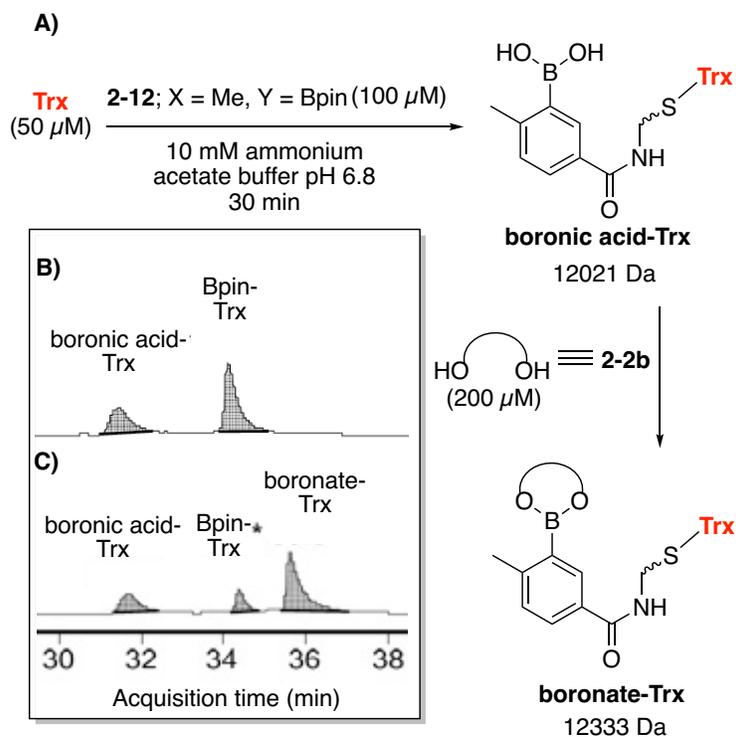
**Figure 2-21:** Labeling of boronic acid modified BSA with nopoldiol-fluorescein-diol (**2-14**) and gel results. A) Modification of free cysteine with **2-12** or **2-13** and labeling with **2-14**. B) Time-

dependent and C) dose-dependent fluorescent labeling of BSA on gel containing phosphate buffer. Protein loading was determined by Coomassie blue staining. Compound **2-13** is a negative control probe devoid of a boronic acid. D) Fluorescent labeling of BSA on gel containing tris buffer (a triol), inducing breakdown of conjugates.

### **2.2.3.2 Boronate formation on thioredoxin (Trx) analyzed *via* HPLC-MS**

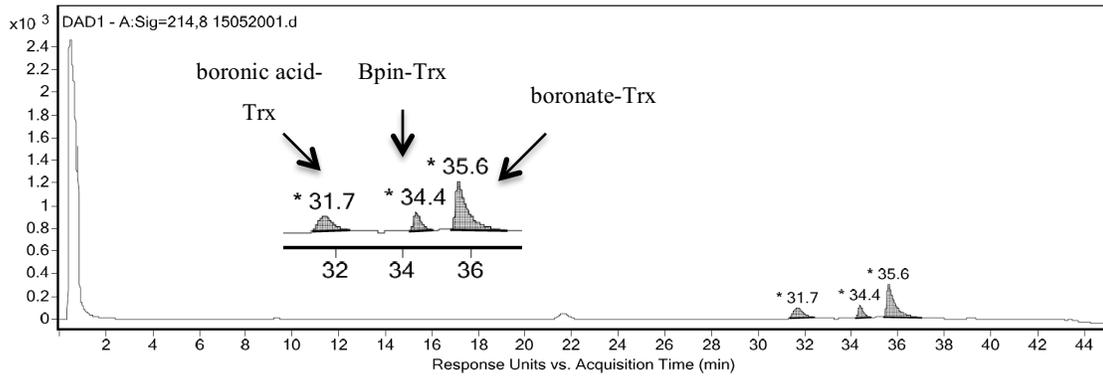
With this promising result in hand, another selective protein labeling experiment was performed. This time, thioredoxin (Trx, ~11.7 kDa) (50  $\mu$ M), a protein with a single disulfide, was reduced with TCEP (1 mM) overnight followed by functionalization with boronyl-containing maleimide **2-12** (100  $\mu$ M) in ammonium acetate buffer for 30 minutes (Figure 2-23A and 2-23B). Trx was completely converted into boronic acid-Trx (12021 Da) and also Bpin-Trx (12122 Da) conjugates, and analyzed with HPLC and LC-MS (Figure 2-22B and 2-23). The molecular ions indicate that pinacol boronate was partially hydrolyzed into its boronic acid, and also the resulting boronic acid-Trx likely reacted intramolecularly with a proximal –OH or –NH side chain, thus forming a boronate based on a lack of water molecule from the expected molecular weight (12039 Da). Following the addition of **2-2b** (200  $\mu$ M) to the mixture, HPLC and LC-MS were used to monitor and detect the boronate-Trx conjugate after a 30 min reaction time (Figure 2-22). As shown in Figure 2-22C and 2-23, the desired bioconjugate (12333 Da) was formed with > 50% conversion. Additionally, boronic acid-Trx was reacted with **2-2b** in the presence of fetal bovine serum. Even though the exact conversion could not be determined due to spectral overlap with serum compounds, a large amount of boronate-Trx was detected (Figure 2-24). As a negative control, a maleimide adduct **2-13** devoid of a boronic acid head was also reacted with Trx followed by the addition of diol **2-2b** into the **2-13**-Trx adduct. As expected, boronate formation was not detected. Moreover, to address the orthogonality of the reaction in complex protein mixtures, this experiment was repeated by mixing the boronic acid-Trx adduct with an equivalent volume of complete cell media (DMEM + 10% fetal bovine serum + 1% ampicillin + 1% glutamax) prior to addition of dansyl conjugated diol **2-1**. Remarkably, boronate-Trx (12517 Da, 83%) was detected with HPLC-Fluo and LC-MS (Figure 2-25 and 2-26). A small amount of double boronate-Trx was also observed as shown in Figure 2-25 and 2-26. This double labeling is probably due to the addition of two molecules of **2-12** onto the reduced Trx *via* Michael additions of two reactive thiols obtained by the reduction of disulphide.

As expected, when only diol **2-1** was incubated with Trx, no other new peaks from unselective labeling were observed (Figure 2-27).



**Figure 2-22:** Boronate formation on Trx analyzed by HPLC-MS. A) Modification of Trx with **2-12** and subsequent labeling with **2-2b**. B) HPLC result without **2-2b**. C) HPLC result in the presence of **2-2b**.

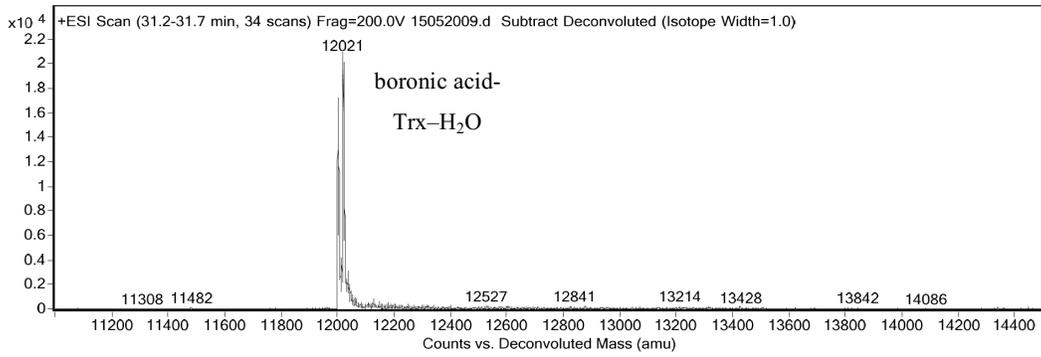
A)



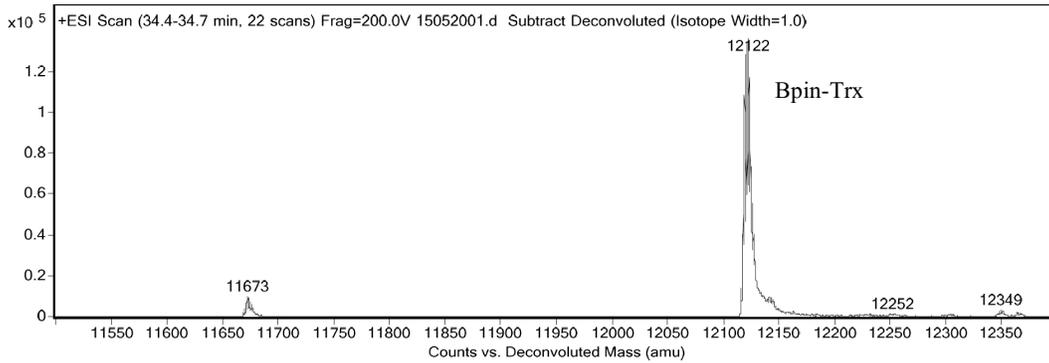
Integration Peak List

Peak	Start	RT	End	Height	Area	Area %
1	31.3	31.7	32.4	89.72	2487.87	35.82
2	34.2	34.4	34.9	105.63	1673.63	24.09
3	35.4	35.6	37	284.94	6946.13	100

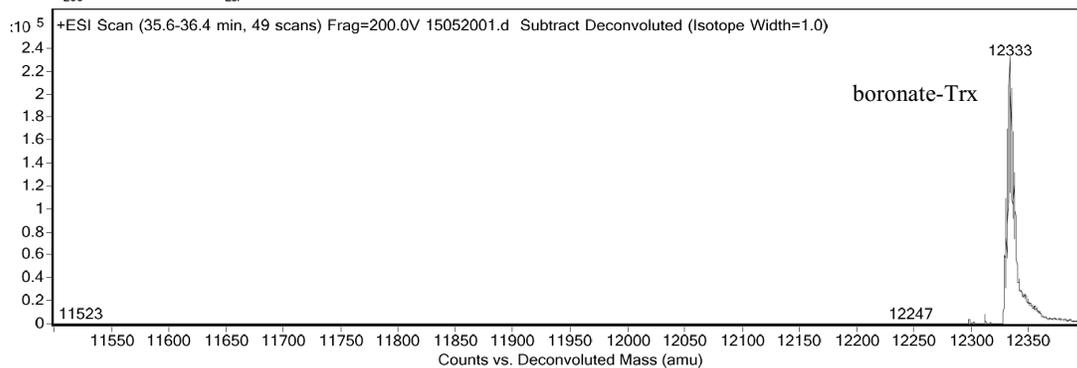
B)



C)

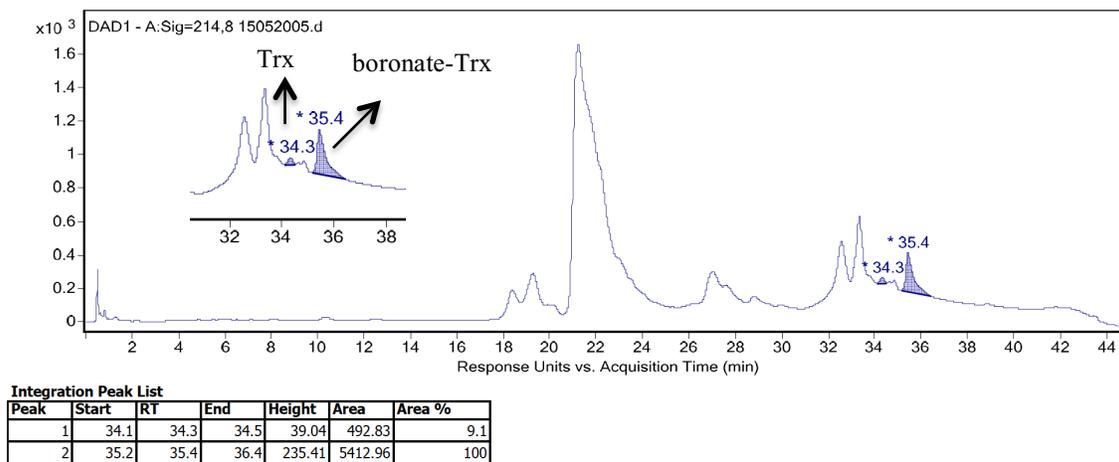


D)

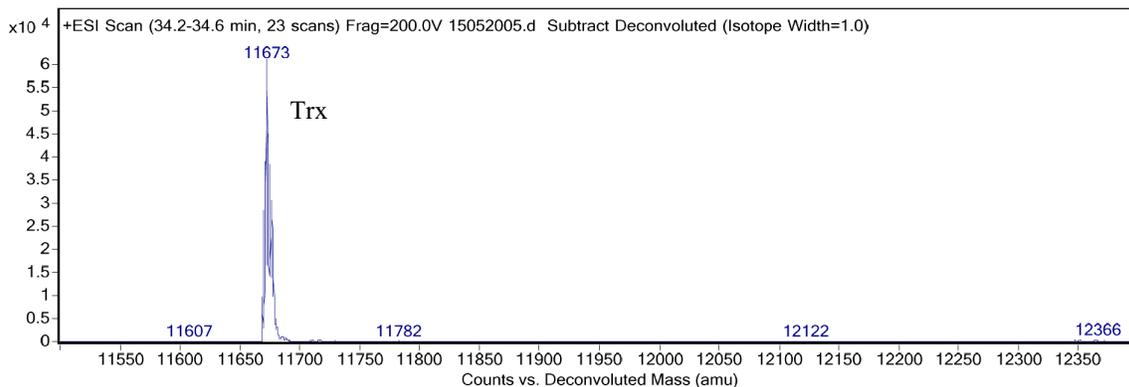


**Figure 2-23:** A) HPLC chromatograms for Trx, boronic acid **2-12** and diol **2-2b** in the absence of serum. B) MS trace of boronic acid-Trx. C) MS trace of Bpin-Trx. D) MS trace of boronate-Trx.

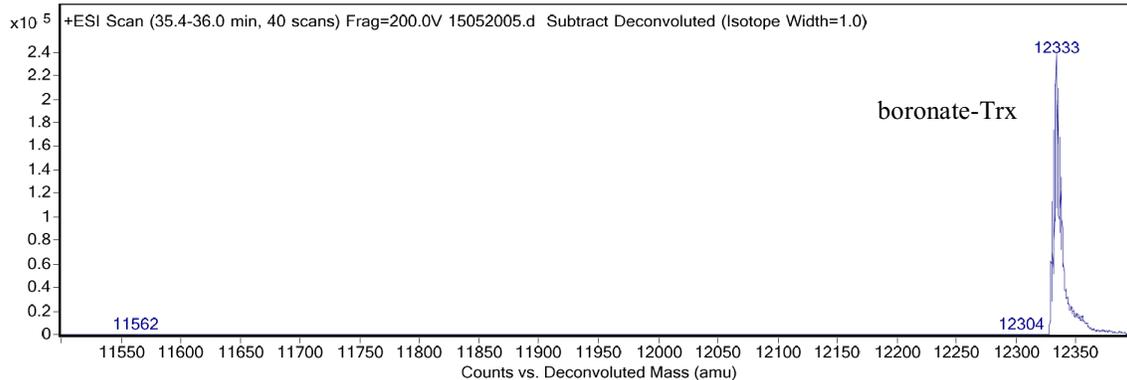
A)



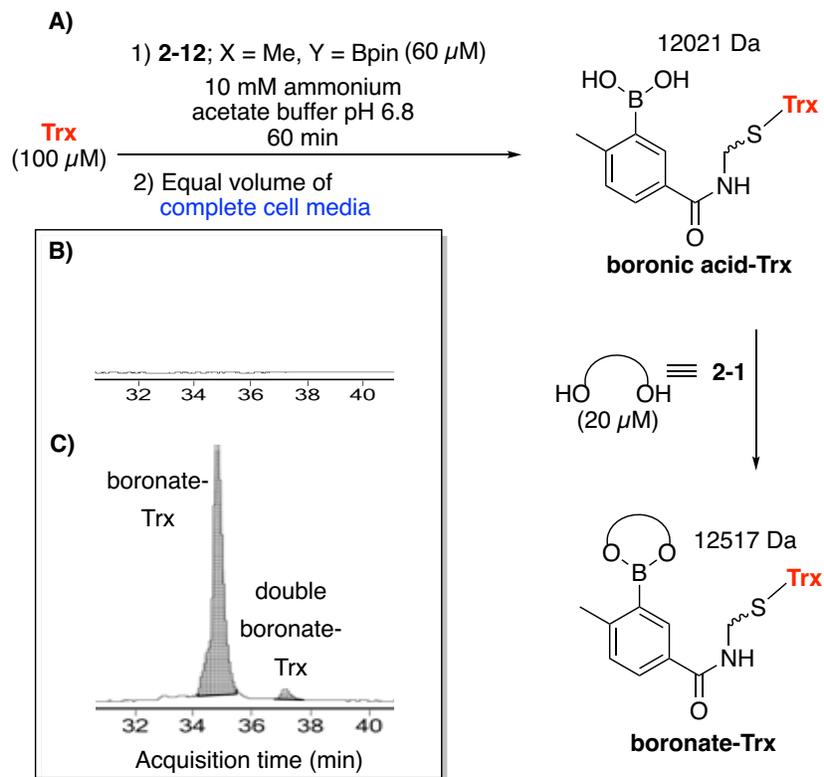
B)



C)

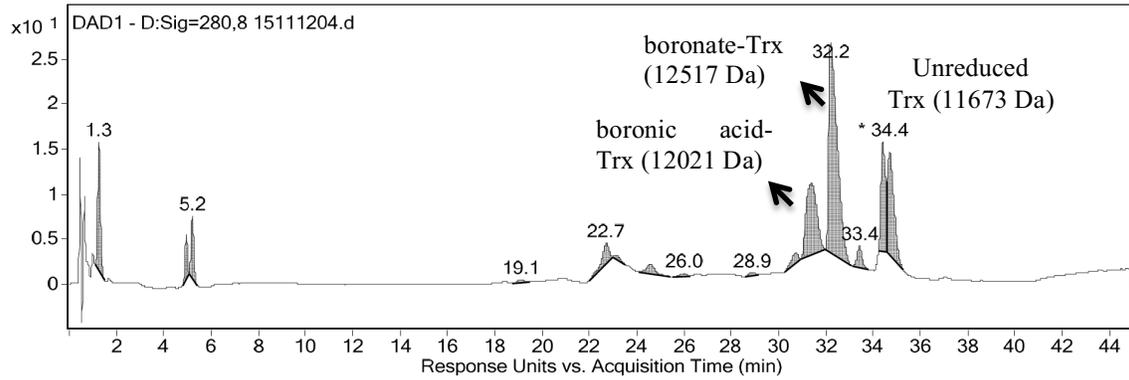


**Figure 2-24:** A) HPLC chromatograms for Trx, boronic acid **2-12** and diol **2-2b** in the presence of serum. B) MS trace of free Trx. C) MS trace of boronate-Trx (12333 Da).

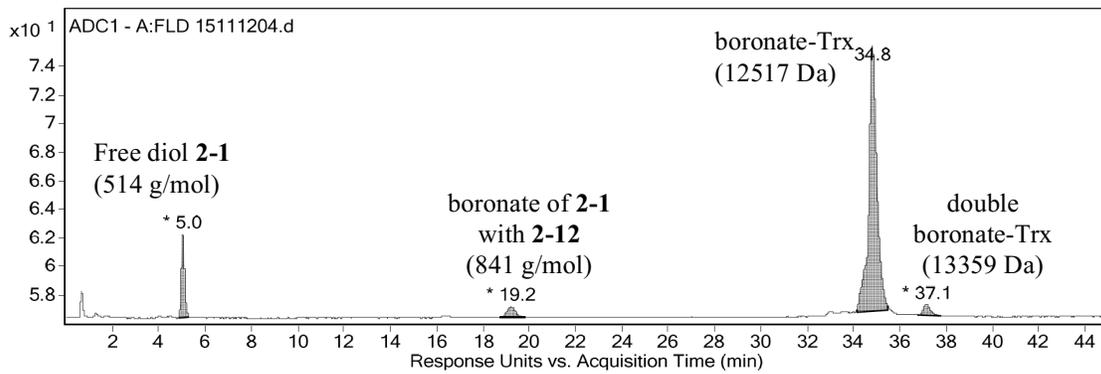


**Figure 2-25:** Boronate formation on Trx in complete cell media analyzed by fluorescence HPLC-MS. A) Modification of Trx with **2-12** and subsequent labeling with **2-1**. B) Fluorescence HPLC result without **2-12**. C) Fluorescence HPLC result in the presence of **2-1**.

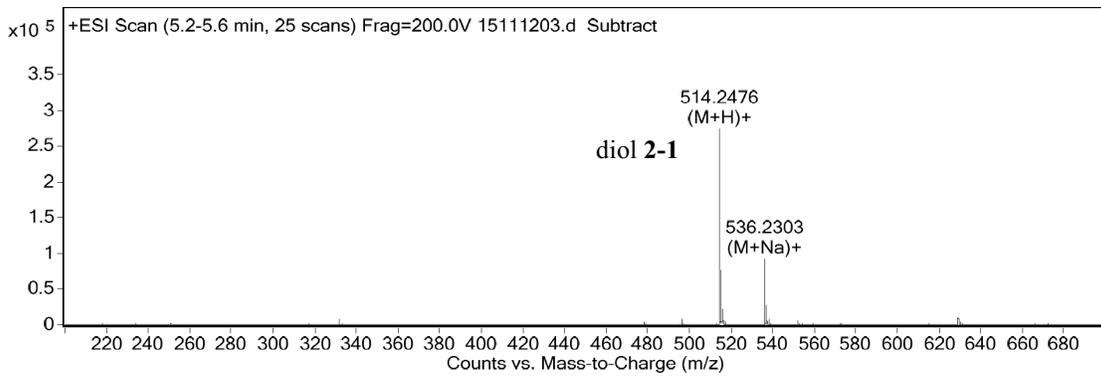
A)



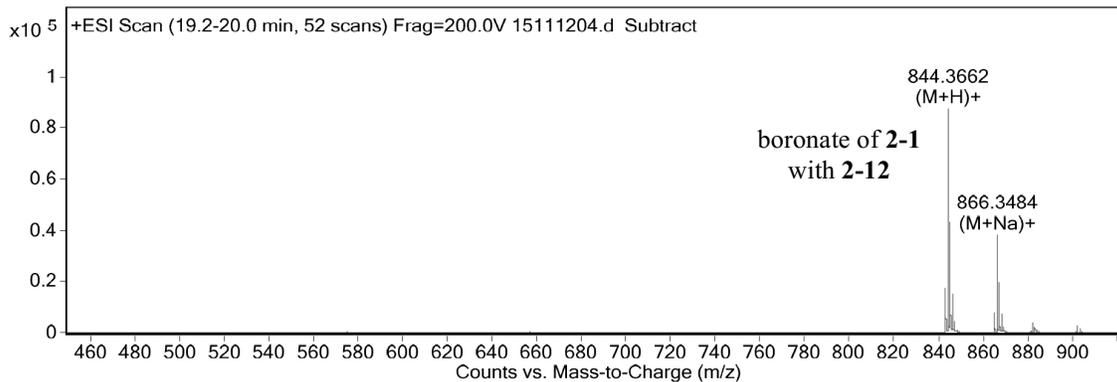
B)



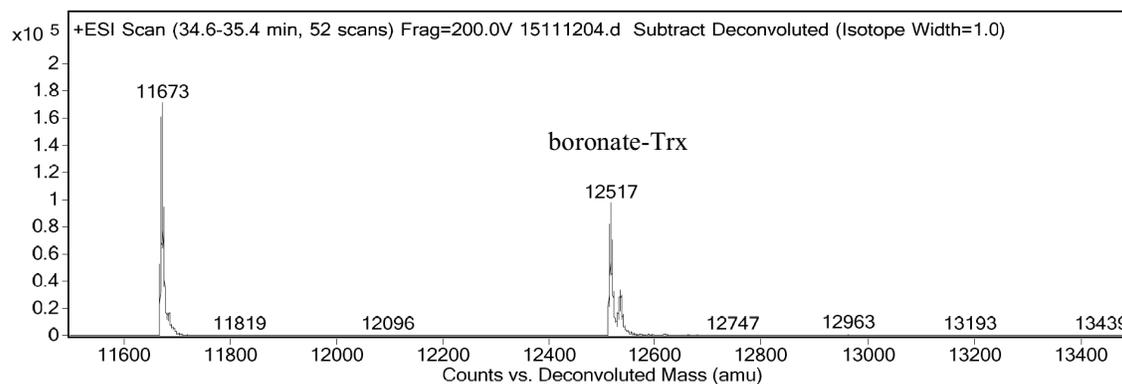
C)



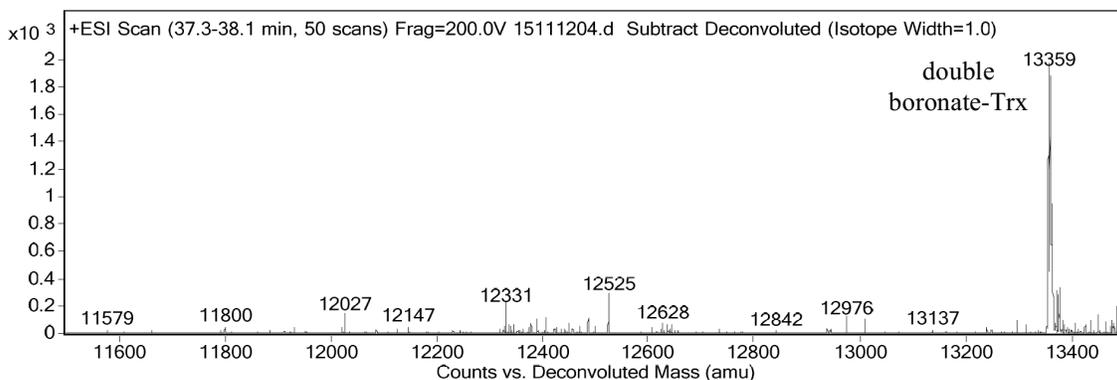
D)



E)

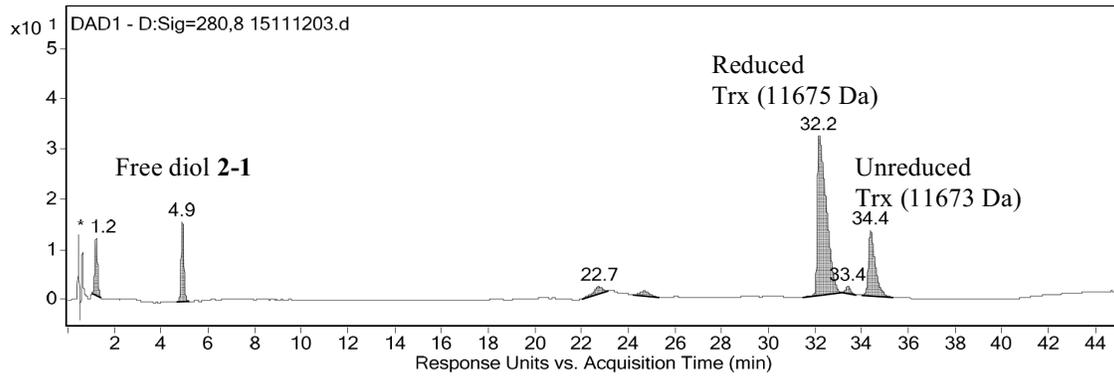


F)

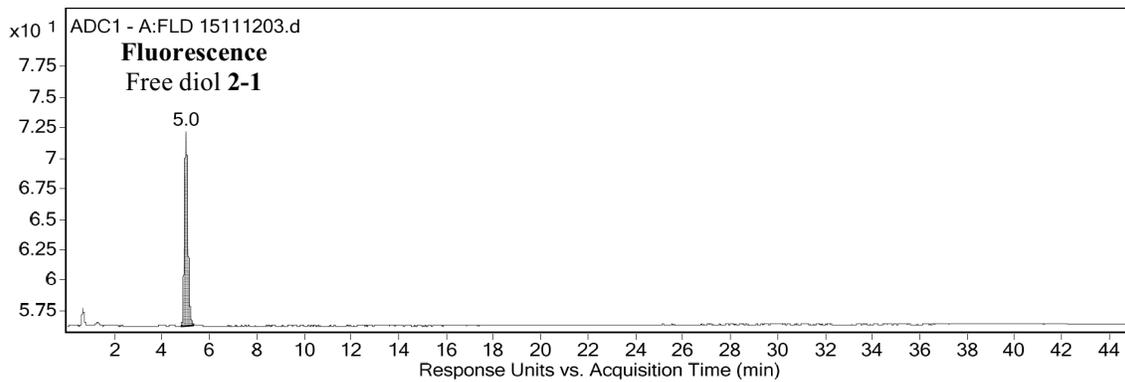


**Figure 2-26:** A) HPLC-UV data for Trx, compound **2-12** and diol **2-1** in complete cell media. B) Fluorescence HPLC chromatograms for Trx, compound **2-12** and diol **2-1** in complete cell media. C) MS trace for diol **2-1**. D) MS trace free boronate formed from diol **2-1** and boronic acid **2-12**. E) MS trace for boronate-Trx. F) MS trace for double boronate-Trx.

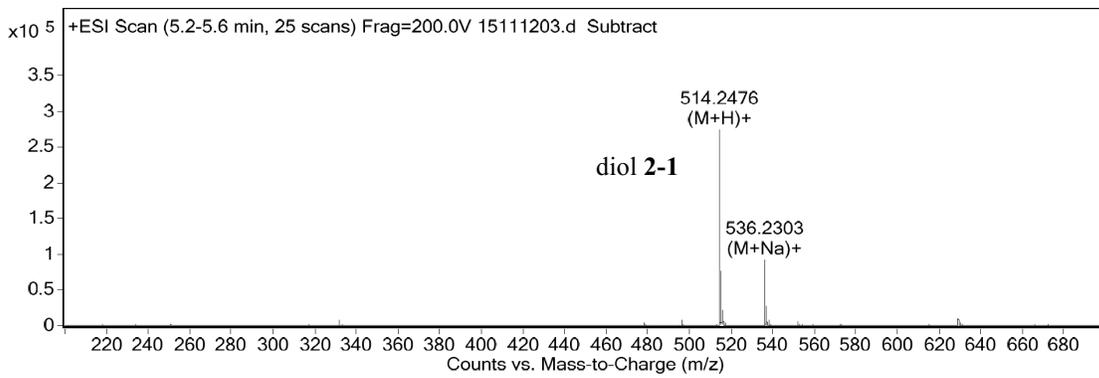
A)



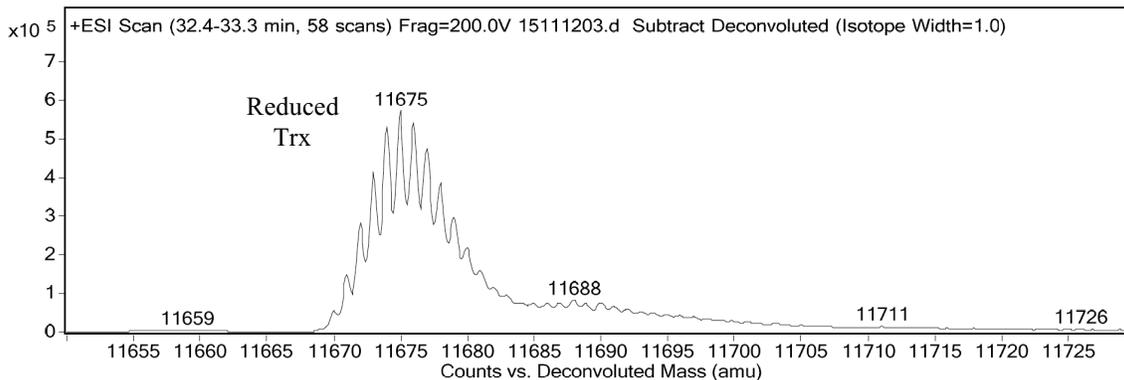
B)



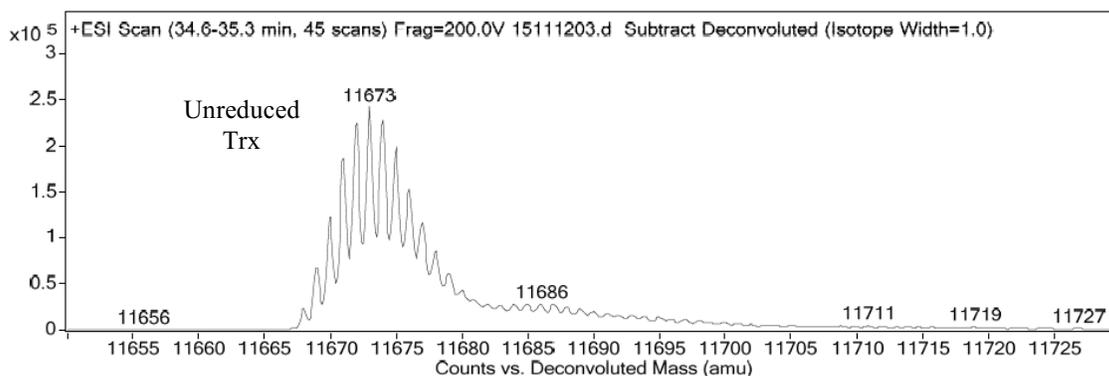
C)



D)



E)

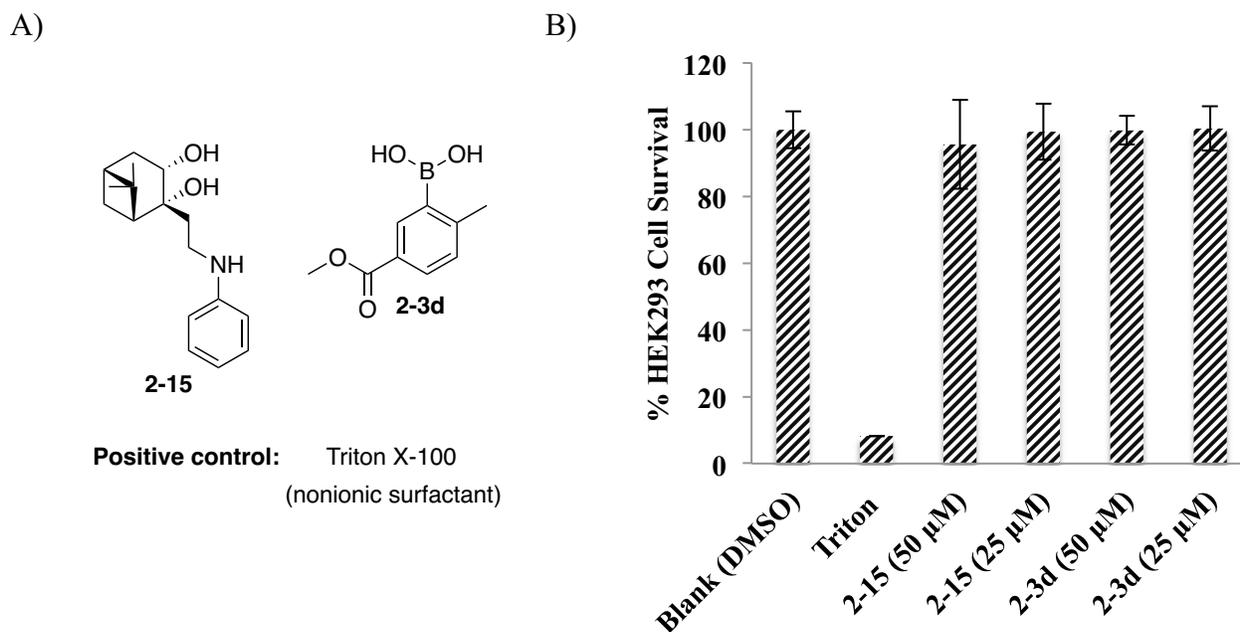


**Figure 2-27:** A) HPLC-UV data for Trx and diol **2-1** in complete cell media. B) Fluorescence HPLC chromatograms. C) MS trace for diol **2-1**. D) MS trace for reduced Trx. E) MS trace for unreduced Trx.

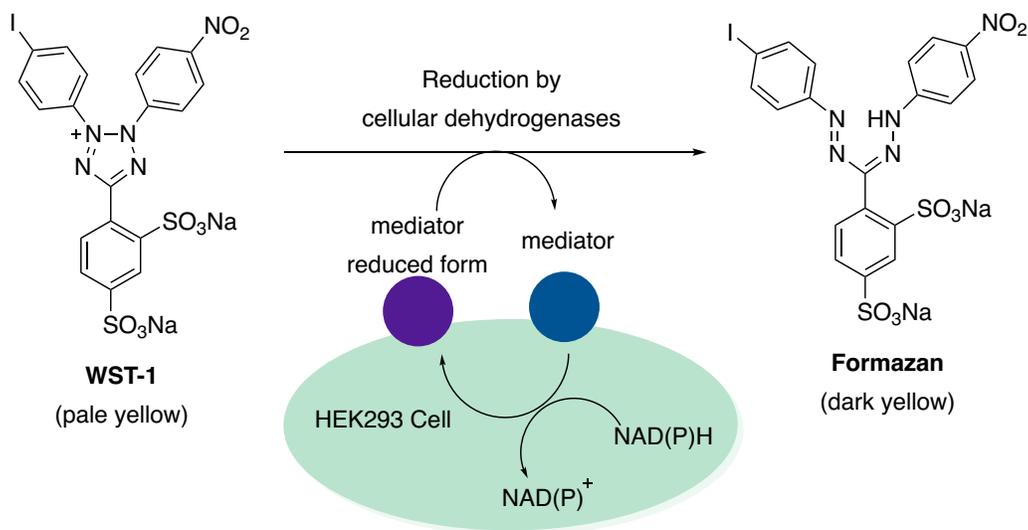
## 2.2.4 Cytotoxicity study

With these promising results in hand, this preliminary study was completed by examining the cytotoxicity of the optimized bioorthogonal handles. To perform this study, I designed and synthesized nopoldiol phenyl **2-15** (Figure 2-28A), which does not interfere with the cytotoxicity test measurement recorded at 450 nm and 600 nm. As a positive control, easily accessible triton X-100 (1  $\mu$ l) was chosen to cause cell death.<sup>30</sup> Triton X-100 was used directly without its further dilution in order to ensure complete cell death. Cell proliferation reagent WST-1 (water-soluble tetrazolium salt) was employed to study the cytotoxicity of diol **2-15** and optimal boronic acid **2-3d** towards mammalian HEK293 cells. This water-soluble reagent is susceptible to metabolically

active cells, and the assay operates at cell surface through reduction of WST-1 (pale yellow) and formation of water-soluble formazan (dark yellow) by cellular dehydrogenases (Figure 2-29).<sup>31a</sup> The presence of mediator (dehydrogenase) reduced form, which is responsible for the formazan formation, is highly determined by the production of NAD(P)H. This NAD(P)H is abundant in viable cells only and reduces the oxidized dehydrogenase (Figure 2-29). Thus, dark yellow formazan formation confirms cell viability. The WST-1 assay was chosen over commonly used MTT assay. The WST-1 assay generates a water-soluble formazan while MTT leads to an insoluble formazan; therefore the WST-1 assay can be followed more directly without a necessity for solubilizing its formazan.<sup>31b</sup> This assay occurs at the outside of viable cells since WST-1 is cell impermeable, unlike cell permeable MTT. Thus, WST-1 is a safer option because both WST-1 and its soluble formazan may not cross the cell membrane and cause toxicity due to their accumulation inside the cell.<sup>31b</sup> Finally, the cytotoxicity of bioorthogonal handles **2-15** and **2-3d** were examined using the WST-1 assay towards HEK293 cells. Fortunately, the reagents were found to be benign over 18 hours at 50 and 25  $\mu\text{M}$  concentrations (Figure 2-28B).



**Figure 2-28:** Cytotoxicity results of bioorthogonal handles **2-15** and **2-3d** on HEK293 cells. Blank was DMSO, and Triton X-100 was used as a positive control. The results are the average of three replicates and error bars are the standard deviations of these replicates.



**Figure 2-29:** The principle and mechanism of the WST-1 assay.

## 2.3 Conclusions

In conclusion, I developed a ‘click’ bioorthogonal reaction system enabling the fast ligation ( $\sim 7.7 \text{ M}^{-1}\text{s}^{-1}$ ) of easily synthesized, conjugatable derivatives of nopoldiol and 2-methyl-5-carboxyphenylboronic acid **2-3d** to form tightly bound boronates with submicromolar equilibrium dissociation constant. For the first time, boronate formation with a rigid diol in water was studied thoroughly with an optimization of both reaction partners, thus providing useful knowledge of this important condensation process. This efficient and robust boronate conjugation was successfully demonstrated on model proteins (Trx and BSA). The bioorthogonal reagents **2-15** and **2-3d** were confirmed to be benign towards HEK293 cells at 50 and 25  $\mu\text{M}$  concentrations. Measured  $k_{\text{ON}}$  rates are faster than many existing bioconjugation systems, and a high  $K_{\text{eq}}$  enables high conversion under physiological conditions. Coupled with the ability to site-specifically encode boronic acid containing aminoacids in proteins,<sup>[9b,c]</sup> reversibility can be an additional asset when target turnover or time-based profiling are required. This preliminary account also suggests applications in affinity purification, surface immobilization, and materials chemistry.

## 2.4 Experimental

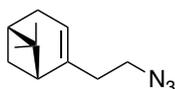
### 2.4.1 General information

Unless otherwise stated, all reactions were performed under a nitrogen atmosphere using

flame-dried glassware. THF, dichloromethane, DMF and methanol were obtained from a MBraun MB SPS\* solvent system prior to use. Most of the reagents were purchased from Sigma-Aldrich and used as received. Arylboronic acids were purchased either from Sigma-Aldrich or Combi-Blocks. Thin layer chromatography (TLC) was performed on Merck Silica Gel 60 F254 plates and was visualized with UV light, curcumin and KMnO<sub>4</sub> stain. NMR spectra were recorded on INOVA-400, INOVA-500, INOVA-600 or INOVA-700 MHz instruments. The residual solvent protons (<sup>1</sup>H) of CDCl<sub>3</sub> (7.26 ppm), ACN-d<sub>3</sub> (1.94 ppm), DMSO-d<sub>6</sub> (2.50 ppm), acetone-d<sub>6</sub> (2.05 ppm), CD<sub>3</sub>OD (3.31 ppm) and the solvent carbons (<sup>13</sup>C) of CDCl<sub>3</sub> (77.06 ppm), ACN-d<sub>3</sub> (1.32 and 118.26 ppm), DMSO-d<sub>6</sub> (39.52 ppm), acetone-d<sub>6</sub> (29.84 and 206.26 ppm), CD<sub>3</sub>OD (49 ppm) were used as internal standards. <sup>1</sup>H NMR data is presented as follows: chemical shift in ppm ( $\delta$ ) downfield from tetramethylsilane (multiplicity, coupling constant, integration). The following abbreviations are used in reporting NMR data: s, singlet; br s, broad singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; ddd, doublet of doublet of doublets; dddd, doublet of doublet of doublet of doublets; app s, apparent singlet; app ddt, apparent doublet of doublet of triplets; m, multiplet; comp m, complex multiplet. Accuracy of reported coupling constant (J) is  $\pm$  0.3 Hz. In <sup>13</sup>C NMR spectroscopy, The quaternary carbon bound to the boron atom is often missing due to the quadrupolar relaxation of boron. This effect was observed in each boron-containing compound. High-resolution mass spectra were recorded by the University of Alberta mass spectrometry services laboratory using either electron impact (EI) or electrospray ionization (ESI) techniques. Optical rotations were measured using a 1 mL cell with a 1 dm length on a P.E. 241 polarimeter. Melting points were determined in a capillary tube using a Gallenkamp melting point apparatus and are uncorrected. HPLC, LC-MS methods and biological instruments were described specifically in the corresponding text below.

## 2.4.2 Chemical synthesis and analytical data

### Synthesis of (-)-nopoldiol dansyl (2-1)



**(1*R*,5*S*)-2-(2-Azidoethyl)-6,6-dimethylbicyclo[3.1.1]hept-2-ene (2-1a):** To (-)-nopol (1.0 equiv, 1.2 g, 7.2 mmol) and pyridine (1.9 ml) solution, *p*-toluene sulfonyl chloride (1.2 equiv, 1.6

g, 8.6 mmol) was added at 0 °C under nitrogen balloon. The mixture was stirred for 1 h under ice-water bath, after which time the ice-water bath was removed and the solution was kept in fridge for 3 h. The completion of reaction was monitored with TLC. To the reaction mixture, the solution of concentrated HCl (2.0 ml) and distilled water (4.0 ml) was added. The aqueous mixture was extracted with Et<sub>2</sub>O (3 × 25 ml). The combined organic layer was mixed with K<sub>2</sub>CO<sub>3</sub> and MgSO<sub>4</sub>, and the mixture was stirred for 20 min, filtered and concentrated in vacuo. (–)-Nopol-tosyl (crude yellow oily product) was obtained and used immediately for the next step without further purification. Then, (–)-nopol-tosyl was immediately used to synthesize nopol-azide (**2-1a**). To a solution of (–)-nopol-tosyl (1.0 equiv, 16.9 mmol) in DMSO (20 ml), sodium azide (2.0 equiv, 2.2 g, 33.8 mmol) was added at 0 °C under nitrogen. Then the reaction mixture was brought to 80 °C and stirred for 16 h. Distilled water (50 ml) was added and the aqueous part was extracted with EtOAc (3 × 30 ml). The combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (hexanes to EtOAc/hexanes 5:95) to obtain (–)-nopol-azide (**2-1a**) (2.7 g, 83% yield).

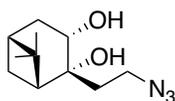
Yellow liquid  $R_f = 0.83$  (5:95, EtOAc/hexanes).

<sup>1</sup>H NMR δ/ppm: (700 MHz, CDCl<sub>3</sub>) 5.34 – 5.32 (m, 1H), 3.47 – 3.07 (m, 2H), 2.40 – 2.36 (m, 1H), 2.31 – 2.16 (comp m, 4H), 2.11 – 2.08 (m, 1H), 2.02 (app td,  $J = 5.6, 1.6$  Hz, 1H), 1.28 (s, 3H), 1.17 (d,  $J = 8.6$  Hz, 1H), 0.84 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 144.3, 119.0, 49.4, 45.7, 40.8, 38.1, 36.0, 31.7, 31.4, 26.3, 21.2.

IR (Microscope, cm<sup>-1</sup>): 3030, 2985, 2918, 2834, 2098, 1454, 1264.

HRMS (GC-CIMS): for C<sub>11</sub>H<sub>18</sub>N<sub>3</sub> (M + H)<sup>+</sup>: *calcd.*: 192.2; *found*: 192.5.



**(1R,2R,3S,5R)-2-(2-Azidoethyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (2-1b):** *N*-Methylmorpholine *N*-oxide (NMO) 50 wt% in H<sub>2</sub>O (1.3 equiv, 2.1 ml, 10.2 mmol) and pyridine (1.2 equiv, 0.76 ml, 9.4 mmol) were added to the solution of (–)-nopol-azide (**2-1a**) (1.0 equiv, 1.5 g, 7.8 mmol) in acetone:water (12:0.9 ml). Lastly, K<sub>2</sub>OsO<sub>4</sub>•2H<sub>2</sub>O (2 mol%, 31 mg, 0.016 mmol) was added. The mixture was stirred and refluxed at 65 °C for 24 h under nitrogen balloon.

Then, the reaction mixture was concentrated in vacuo and mixed with EtOAc (50 ml). The organic phase was washed with HCl (1 × 5.0 ml, 1 N), distilled water (1 × 5.0 ml) and brine (1 × 5.0 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (4:6, EtOAc/hexanes) to obtain (–)-nopoldiol azide (**2-1b**) (1.2 g, 66% yield).

Pale-white solid  $R_f = 0.54$  (4:6, EtOAc/hexanes).

Mp = 30.1 – 32.2 °C.

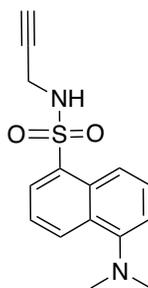
$[\alpha]_D^{20} : -5.0$  (c 0.98, CHCl<sub>3</sub>).

<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 4.09 (ddd,  $J = 9.4, 5.7, 5.6$  Hz, 1H), 3.64 – 3.48 (m, 2H), 3.24 (s, 1H), 2.82 (d,  $J = 6.1$  Hz, 1H), 2.48 (dddd,  $J = 14.0, 9.4, 3.7, 2.5$  Hz, 1H), 2.24 – 2.18 (m, 1H), 2.06 (app t,  $J = 5.8$  Hz, 1H), 1.99 – 1.86 (comp m, 2H), 1.71 (app dt,  $J = 14.3, 6.4$  Hz, 1H), 1.65 (ddd,  $J = 14.0, 5.4, 2.5$  Hz, 1H), 1.39 (d,  $J = 10.4$  Hz, 1H), 1.28 (s, 3H), 0.95 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 74.6, 68.1, 52.9, 47.5, 40.8, 40.6, 38.8, 38.0, 27.93, 27.88, 24.3.

IR (Microscope, cm<sup>-1</sup>): 3383, 2987, 2924, 2872, 2096, 1475, 1453, 1299, 1262.

HRMS (ESI-TOF): for C<sub>11</sub>H<sub>19</sub>N<sub>3</sub>NaO<sub>2</sub> (M + Na)<sup>+</sup>: *calcd.*: 248.1369; *found*: 248.1367.



**5-(Dimethylamino)-N-(prop-2-yn-1-yl)naphthalene-1-sulfonamide (2-1c):** To dansyl chloride (1.0 equiv, 0.11 g, 0.40 mmol) in DCM (4 ml), propargylamine (1.1 equiv, 28 μl, 0.44 mmol) and DIPEA (2.0 equiv, 140 μl, 0.80 mmol) were added at 0 °C. The solution was stirred at 0 °C for 5 min, after which time the ice-water bath was removed and the reaction mixture was stirred for 16 h under ambient atmosphere. Then DCM was evaporated under vacuo and the crude residue was purified by flash chromatography (3:7, EtOAc/hexanes) to obtain pure propargyl dansyl (**2-1c**) (92 mg, 80% yield).

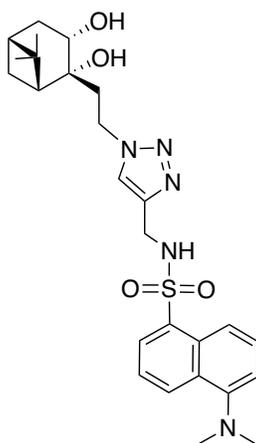
Yellow-green viscous oil,  $R_f = 0.42$  (3:7, EtOAc/hexanes).

**<sup>1</sup>H NMR** δ/ppm: (700 MHz, CDCl<sub>3</sub>) 8.56 (app dt, *J* = 8.5, 1.1 Hz, 1H), 8.28 (dd, *J* = 7.3, 1.3 Hz, 1H), 8.26 (app dt, *J* = 8.7, 0.9 Hz, 1H), 7.58 (dd, *J* = 8.6, 7.5 Hz, 1H), 7.53 (dd, *J* = 8.5, 7.3 Hz, 1H), 7.20 (dd, *J* = 7.6, 0.9 Hz, 1H), 4.80 (t, *J* = 6.1 Hz, 1H), 3.77 (dd, *J* = 6.0, 2.5 Hz, 2H), 2.89 (s, 6H), 1.92 (t, *J* = 2.5 Hz, 1H).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CDCl<sub>3</sub>) 152.2, 134.2, 130.9, 130.00, 129.96, 129.8, 128.6, 123.2, 118.6, 115.3, 77.8, 72.8, 45.5, 33.1.

**IR** (Microscope, cm<sup>-1</sup>): 3290, 2924, 2853, 2790, 1660, 1588, 1574, 1456, 1324, 1162, 1145, 1324, 1162, 1145.

**HRMS** (ESI-TOF): for C<sub>15</sub>H<sub>17</sub>N<sub>2</sub>O<sub>2</sub>S (M + H)<sup>+</sup>: *calcd.*: 289.1005; *found*: 289.1006; for C<sub>15</sub>H<sub>16</sub>N<sub>2</sub>NaO<sub>2</sub>S (M + Na)<sup>+</sup>: *calcd.*: 311.0825; *found*: 311.0824.



***N*-((1-(2-((1*R*,2*R*,3*S*,5*R*)-2,3-dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl)-1*H*-1,2,3-triazol-4-yl)methyl)-5-(dimethylamino)naphthalene-1-sulfonamide (2-1)**: To the solution of nopoldiol azide **2-1b** (1.0 equiv, 27 mg, 0.12 mmol) and compound **2-1c** (1.2 equiv, 40 mg, 0.14 mmol) in water:ACN (2:1 ml), CuSO<sub>4</sub>•5H<sub>2</sub>O (0.10 equiv, 3.0 mg, 0.012 mmol) was added under ambient atmosphere at room temperature. Hydrazine monohydrate (NH<sub>2</sub>NH<sub>2</sub>•H<sub>2</sub>O) (0.10 equiv, 6.2 μl, 0.012 mmol) was added dropwise to the reaction mixture under vigorous stirring. The reaction mixture was stirred for 16 h at room temperature and it turned to pale yellow or light green color. Then, water (5 ml) was added to the mixture and the product was extracted with EtOAc (3 × 20 ml). The combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (EtOAc) to obtain (-)-nopoldiol dansyl (**2-1**) (58 mg, 94% yield).  
Green powder, *R<sub>f</sub>* = 0.39 (EtOAc).

Mp = 62.5 – 64.3 °C.

<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 8.52 (d, *J* = 8.5 Hz, 1H), 8.27 (dd, *J* = 7.2, 1.3 Hz, 1H), 7.55 – 7.47 (comp m, 2H), 7.43 (s, 1H), 7.16 (d, *J* = 7.5 Hz, 1H), 6.20 (t, *J* = 6.2 Hz, 1H), 4.49 – 4.27 (comp m, 2H), 4.26 – 4.13 (comp m, 3H), 4.01 – 3.85 (comp m, 2H), 2.87 (s, 6H), 2.46 – 2.38 (m, 1H), 2.22 – 2.13 (m, 1H), 2.05 – 1.82 (comp m, 4H), 1.69 – 1.61 (m, 1H), 1.38 (d, *J* = 10.5 Hz, 1H), 1.23 (s, 3H), 0.88 (s, 3H).

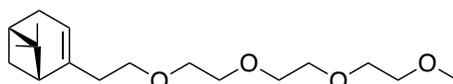
<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 152.0, 143.9, 134.7, 130.7, 129.8, 129.7, 129.57, 129.54, 128.7, 123.2, 122.6, 118.7, 115.4, 74.2, 67.6, 52.2, 46.0, 45.5, 42.6, 40.5, 38.8, 38.1, 27.9, 27.8, 24.3.

IR (Microscope, cm<sup>-1</sup>): 3294, 3143, 2925, 2869, 2790, 2249, 1613, 1588, 1575, 1454, 1323, 1161, 1144, 1060.

HRMS (ESI-TOF): for C<sub>26</sub>H<sub>36</sub>N<sub>5</sub>O<sub>4</sub>S (M + H)<sup>+</sup>: *calcd.*: 514.2483; *found*: 514.2480; for C<sub>26</sub>H<sub>35</sub>N<sub>5</sub>NaO<sub>4</sub>S (M + Na)<sup>+</sup>: *calcd.*: 536.2302; *found*: 536.2297.

### **Synthesis of (-)-nopol-PEG-diol-1 (2-2a)**

#### **(-)-Nopol-PEG-1 (2-6):**



**13-((1*R*,5*S*)-6,6-Dimethylbicyclo[3.1.1]hept-2-en-2-yl)-2,5,8,11-tetraoxatridecane (2-6):** (-)-Nopol-tosyl was synthesized from commercially available (-)-nopol according to the method described for the synthesis of compound **2-1a**. To tri(ethyleneglycol)monomethyl ether (1.3 equiv, 1.5 g, 9.3 mmol) in dry THF (4.0 ml), sodium hydride 60% (w/w) dispersion in mineral oil (1.2 equiv, 0.34 g, 8.6 mmol) was added in small portions over 5 min at 0 °C under nitrogen balloon and the mixture was stirred at room temperature until bubbles disappeared. Then, TBAI (0.10 equiv, 0.26 g, 0.72 mmol) was added to the reaction mixture. The crude (-)-nopol-tosyl in dry THF (4 ml) was added dropwise to the mixture over 10 min at 0 °C after which time the ice-water bath was removed and the solution was stirred for 24 h at room temperature. Then, the reaction mixture was concentrated in vacuo and mixed with EtOAc (50 ml). The organic phase was washed with HCl (2 × 5.0 ml, 1 N), distilled water (1 × 5.0 ml) and brine (1 × 5.0 ml), dried

(MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:1, EtOAc/hexane) to provide (–)-nopol-PEG-1 (**2-6**) (0.37 g, 17% yield).

Yellowish oil:  $R_f = 0.60$  (1:1, EtOAc/hexane).

$[\alpha]_D^{20} : -3.0$  (*c* 0.93, CHCl<sub>3</sub>).

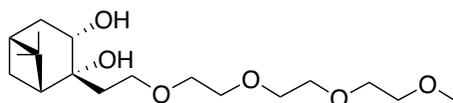
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 5.26 – 5.23 (m, 1H), 3.69 – 3.60 (comp m, 8H), 3.59 – 3.52 (comp m, 4H), 3.49 – 3.43 (m, 2H), 3.38 (s, 3H), 2.36 – 2.32 (m, 1H), 2.28 – 2.12 (comp m, 4H), 2.09 – 2.04 (m, 1H), 2.04 – 2.00 (m, 1H), 1.26 (s, 3H), 1.14 (dd, *J* = 8.5, 1.0 Hz, 1H), 0.81 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 144.8, 117.6, 71.8, 70.43, 70.42, 70.40, 70.3, 69.8, 69.6, 58.8, 45.7, 40.6, 37.8, 36.9, 31.4, 31.1, 26.1, 21.0.

IR (Microscope, cm<sup>-1</sup>): 2983, 2914, 2875, 1467, 1452, 1114.

HRMS (ESI-TOF): for C<sub>18</sub>H<sub>32</sub>NaO<sub>4</sub> (M + Na)<sup>+</sup>: *calcd.*: 335.2193; *found*: 335.2190.

(–)-Nopol-PEG-diol-1 (**2-2a**):



**(1R,2R,3S,5R)-6,6-Dimethyl-2-(2,5,8,11-tetraoxatridecan-13-yl)bicyclo[3.1.1]heptane-2,3-diol (2-2a)**: *N*-Methylmorpholine *N*-oxide (NMO) (50%) (1.3 equiv, 0.33 ml, 1.6 mmol) and 2,6-lutidine (1.2 equiv, 0.17 ml, 1.4 mmol) were added to compound **2-6** (1.0 equiv, 0.38 g, 1.2 mmol) in isopropanol (4 ml). Lastly, to the reaction mixture, K<sub>2</sub>OsO<sub>4</sub>•H<sub>2</sub>O purchased from strem chemicals (2 mol%, 8.7 mg, 0.024 mmol) was added. The mixture was stirred and refluxed at 95 °C for 16 h under ambient atmosphere. Then, the reaction mixture was concentrated in vacuo and mixed with EtOAc (50 ml). The organic phase was washed with HCl (1 × 5.0 ml, 1 N), distilled water (1 × 5.0 ml) and brine (1 × 5.0 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:3, acetone/hexanes) to provide (–)-nopol-PEG-diol-1 (**2-2a**) (0.38 g, 93% yield).

Yellowish oil:  $R_f = 0.29$  (1:3, acetone/hexanes).

$[\alpha]_D^{20} : -5.5$  (*c* 0.74, CHCl<sub>3</sub>).

<sup>1</sup>H NMR δ/ppm: (400 MHz, CDCl<sub>3</sub>) 4.36 (bs, 1H), 4.19 (app t, *J* = 7.7 Hz, 1H), 4.08 (bs, 1H), 3.86 (td, *J* = 9.6, 2.6 Hz, 1H), 3.73 – 3.47 (comp m, 13H), 3.38 (s, 3H), 2.39 (dddd, *J* = 13.6, 9.6,

3.8, 2.5 Hz, 1H), 2.15 (app dtd,  $J = 10.2, 6.1, 2.4$  Hz, 1H), 2.05 – 1.92 (comp m, 2H), 1.90 – 1.81 (m, 1H), 1.63 (ddd,  $J = 13.9, 5.2, 2.6$  Hz, 1H), 1.53 (ddd,  $J = 14.8, 5.2, 2.6$  Hz, 1H), 1.46 (d,  $J = 10.2$  Hz, 1H), 1.26 (s, 3H), 0.94 (s, 3H).

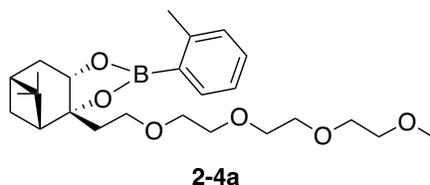
$^{13}\text{C}$  NMR  $\delta$ /ppm: (101 MHz,  $\text{CDCl}_3$ ) 77.6, 74.7, 72.3, 70.9, 70.44, 70.41, 68.1, 67.8, 59.4, 54.7, 42.1, 40.9, 39.0, 37.8, 28.5, 28.4, 24.7 (One missing carbon peak might overlap with solvent peak).

IR (Microscope,  $\text{cm}^{-1}$ ): 3445, 2910, 2872, 1472, 1453, 1108.

HRMS (ESI-TOF): for  $\text{C}_{18}\text{H}_{34}\text{NaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 369.2248; *found*: 369.2245; for  $\text{C}_{18}\text{H}_{35}\text{O}_6$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 347.2428; *found*: 347.2430.

### General Synthesis of (–)-Nopol-PEG-boronates-1 (2-4a – 2-4ac)

Compound **2-2a** (1.0 equiv, 0.1 mmol) was mixed with various arylboronic acids (1.0 equiv, 0.1 mmol) in THF (1ml). The reaction mixture was stirred for 1 h at room temperature under ambient atmosphere. Then, toluene (0.2 ml) was added to the mixture and it was concentrated in vacuo. The crude residue was purified by flash chromatography.



### **(3a*R*,4*R*,6*R*,7a*S*)-5,5-Dimethyl-2-(*o*-tolyl)-3a-(2,5,8,11-tetraoxatridecan-13-yl)hexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborole (2-4a)**

Colorless oil (38 mg, 93% yield):  $R_f = 0.47$  (1:1, EtOAc/hexanes).

$^1\text{H}$  NMR  $\delta$ /ppm: (400 MHz,  $\text{CDCl}_3$ ) 7.78 (dd,  $J = 7.8, 1.6$  Hz, 1H), 7.33 (app td,  $J = 7.5, 1.6$  Hz, 1H), 7.22 – 7.12 (comp m, 2H), 4.60 (dd,  $J = 8.7, 1.8$  Hz, 1H), 3.67 (app t,  $J = 6.8$ , 2H), 3.65 – 3.51 (comp m, 12H), 3.37 (s, 3H), 2.53 (s, 3H), 2.49 – 2.35 (m, 1H), 2.24 – 2.21 (comp m, 2H), 2.05 (app td,  $J = 6.8, 2.1$  Hz, 2H), 2.01 – 1.90 (comp m, 2H), 1.31 (s, 3H), 1.24 (d,  $J = 10.3$  Hz, 1H), 0.92 (s, 3H).

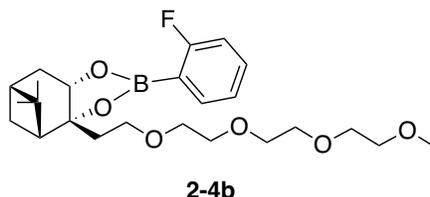
$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 144.9, 136.1, 130.9, 129.9, 124.8, 87.0, 76.6, 72.0, 70.65, 70.63, 70.57, 70.4, 67.3, 59.1, 50.2, 40.4, 39.8, 38.2, 35.8, 27.1, 26.2, 24.1, 23.0, 22.4; The quaternary carbon bound to the boron atom is often missing due to the quadrupolar relaxation of

boron. This effect was observed in each boron-containing compound.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (128 MHz,  $\text{CDCl}_3$ ) 31.2.

IR (Microscope,  $\text{cm}^{-1}$ ): 2924, 2872, 1601, 1444, 1368, 1122.

HRMS (ESI-TOF): for  $\text{C}_{25}\text{H}_{39}\text{BNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 469.2732; *found*: 469.2733.



**(3a*R*,4*R*,6*R*,7a*S*)-2-(2-Fluorophenyl)-5,5-dimethyl-3a-(2,5,8,11-tetraoxatridecan-13-yl)hexahydro-4,6-methanobenzo[d][1,3,2]dioxaborole (2-4b)**

Colorless oil (35 mg, 77% yield):  $R_f$  = 0.38 (1:1, EtOAc/hexanes).

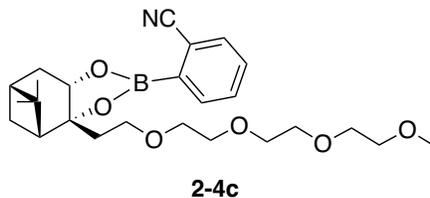
$^1\text{H}$  NMR  $\delta$ /ppm: (400 MHz,  $\text{CDCl}_3$ ) 7.76 – 7.72 (m, 1H), 7.50 – 7.37 (m, 1H), 7.14 (app td,  $J$  = 7.4, 0.9 Hz, 1H), 7.04 (app t,  $J$  = 8.9 Hz, 1H), 4.66 (dd,  $J$  = 8.8, 1.9 Hz, 1H), 3.67 (app td,  $J$  = 6.6, 1.7 Hz, 2H), 3.64 – 3.51 (comp m, 12H), 3.37 (s, 3H), 2.47 – 2.38 (m, 1H), 2.28 – 2.20 (comp m, 2H), 2.14 – 1.89 (comp m, 4H), 1.31 (s, 3H), 1.23 (d,  $J$  = 10.4 Hz, 1H), 0.91 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 167.3 (d,  $J_{\text{CF}}$  = 250.7 Hz), 136.9 (d,  $J_{\text{CF}}$  = 7.9 Hz), 133.3 (d,  $J_{\text{CF}}$  = 8.7 Hz), 123.7 (d,  $J_{\text{CF}}$  = 3.2 Hz), 115.3 (d,  $J_{\text{CF}}$  = 23.9 Hz), 87.3, 72.0, 70.65, 70.62, 70.57, 70.60, 70.4, 67.2, 59.1, 54.3, 50.2, 40.2, 39.7, 38.3, 35.7, 27.1, 26.2, 24.1.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (128 MHz,  $\text{CDCl}_3$ ) 32.9.

IR (Microscope,  $\text{cm}^{-1}$ ): 2923, 2872, 1615, 1489, 1448, 1362, 1116.

HRMS (ESI-TOF): for  $\text{C}_{24}\text{H}_{36}\text{BFNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 473.2481; *found*: 473.2483.



**2-((3a*R*,4*R*,6*R*,7a*S*)-5,5-Dimethyl-3a-(2,5,8,11-tetraoxatridecan-13-yl)hexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl)benzonitrile (2-4c)**

Colorless oil (33 mg, 72% yield):  $R_f$  = 0.26 (1:1, EtOAc/hexanes).

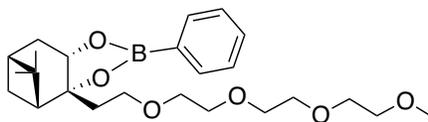
$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 7.90 (dd,  $J = 7.4, 0.6$  Hz, 1H), 7.70 (dd,  $J = 7.5, 0.6$  Hz, 1H), 7.61 – 7.51 (comp m, 2H), 4.69 (dd,  $J = 8.8, 1.9$  Hz, 1H), 3.78 – 3.66 (m, 2H), 3.65 – 3.50 (comp m, 12H), 3.36 (s, 3H), 2.47 – 2.40 (m, 1H), 2.32 – 2.17 (comp m, 2H), 2.14 – 2.00 (comp m, 3H), 1.97 – 1.94 (m, 1H), 1.30 (s, 3H), 1.21 (d,  $J = 9.8$  Hz, 1H), 0.91 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 136.1, 133.5, 131.6, 131.2, 119.1, 117.4, 88.2, 85.0, 77.6, 72.0, 70.62, 70.59, 70.57, 70.5, 70.4, 67.1, 59.1, 49.9, 40.2, 39.6, 38.3, 35.5, 27.1, 26.2, 24.0.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 29.4.

IR (Microscope,  $\text{cm}^{-1}$ ): 2923, 2872, 2226, 1595, 1388, 1361, 1121.

HRMS (ESI-TOF): for  $\text{C}_{25}\text{H}_{36}\text{BNNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 480.2528; *found*: 480.2531.



**2-4d**

**(3a*R*,4*R*,6*R*,7a*S*)-5,5-Dimethyl-2-phenyl-3a-(2,5,8,11-tetraoxatridecan-13-yl)hexahydro-4,6-methanobenzo[d][1,3,2]dioxaborole (2-4d)**

Colorless oil (38 mg, 87% yield):  $R_f = 0.44$  (1:1, EtOAc/hexanes).

$^1\text{H NMR}$   $\delta$ /ppm: (400 MHz,  $\text{CDCl}_3$ ) 7.90 – 7.65 (m, 2H), 7.53 – 7.41 (m, 1H), 7.40 – 7.33 (m, 2H), 4.62 (dd,  $J = 8.8, 1.9$  Hz, 1H), 3.71 – 3.48 (comp m, 14H), 3.37 (s, 3H), 2.48 – 2.35 (m, 1H), 2.29 – 2.16 (comp m, 2H), 2.05 (app td,  $J = 6.9, 4.8$  Hz, 2H), 2.01 – 1.88 (comp m, 2H), 1.30 (s, 3H), 1.22 (d,  $J = 10.4$  Hz, 1H), 0.91 (s, 3H).

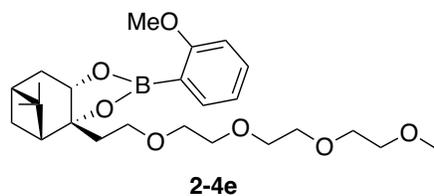
$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ )  $\delta$  134.9, 131.3, 127.8, 87.3, 72.0, 70.63, 70.60, 70.55, 70.4, 67.2, 59.1, 54.4, 50.2, 40.3, 39.7, 38.3, 35.7, 27.1, 26.2, 24.0.

$^{11}\text{B NMR}$   $\delta$ /ppm: (128 MHz,  $\text{CDCl}_3$ ) 30.6.

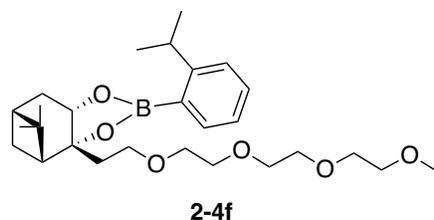
IR (Microscope,  $\text{cm}^{-1}$ ): 3079, 3054, 2920, 2871, 1603, 1499, 1440, 1360, 1121, 1095.

HRMS (ESI-TOF): for  $\text{C}_{24}\text{H}_{37}\text{BNNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 455.2575; *found*: 455.2578.

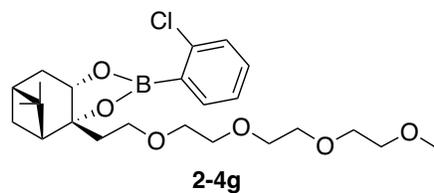
The following boronates **2-4e** – **2-4ac** were also characterized as **2-4a** – **2-4d** but due to the large volume of data, their full characterization data were not provided.



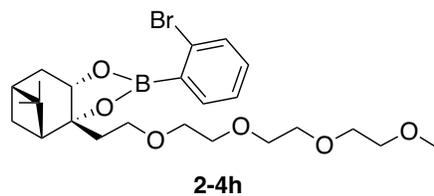
**HRMS (ESI-TOF):** for C<sub>25</sub>H<sub>39</sub>BNaO<sub>7</sub> (M + Na)<sup>+</sup>: *calcd.*: 485.2681; *found*: 485.2681.



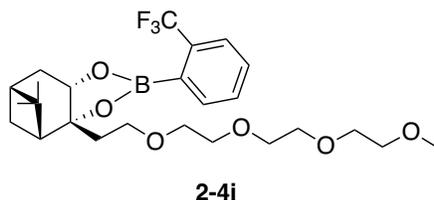
**HRMS (ESI-TOF):** for C<sub>27</sub>H<sub>43</sub>BNaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 497.3045; *found*: 497.3046.



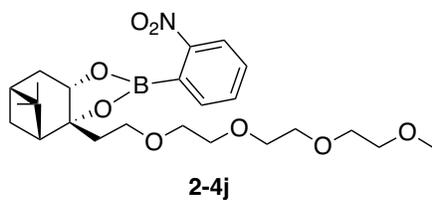
**HRMS (ESI-TOF):** for C<sub>24</sub>H<sub>36</sub>BClNaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 489.2186; *found*: 489.2187.



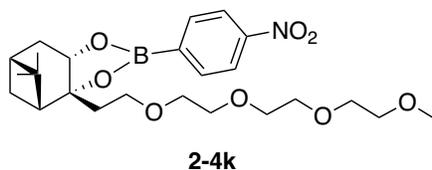
**HRMS (ESI-TOF):** for C<sub>24</sub>H<sub>36</sub>BBrNaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 535.17; *found*: 535.1658.



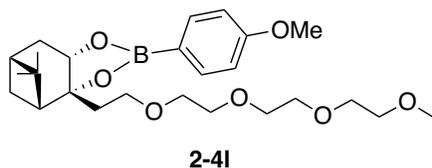
**HRMS (ESI-TOF):** for C<sub>25</sub>H<sub>36</sub>BF<sub>3</sub>NaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 523.2449; *found*: 523.2449.



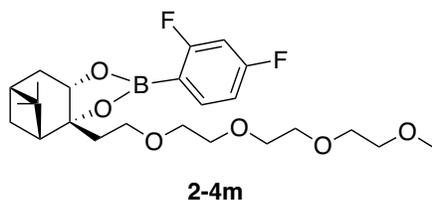
**HRMS** (ESI-TOF): for  $C_{24}H_{36}BNNaO_8$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 500.2426; *found*: 500.2428.



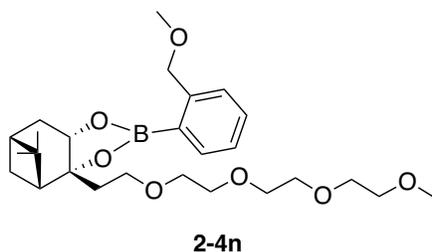
**HRMS** (ESI-TOF): for  $C_{24}H_{36}BNNaO_8$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 500.2426; *found*: 500.2428.



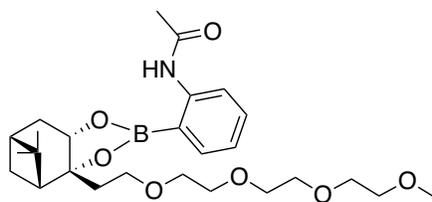
**HRMS** (ESI-TOF): for  $C_{25}H_{39}BNaO_7$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 485.2681; *found*: 485.2684.



**HRMS** (ESI-TOF): for  $C_{24}H_{35}BF_2NaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 491.2387; *found*: 491.2389.

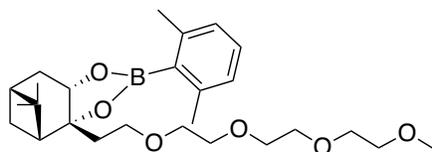


**HRMS** (ESI-TOF): for  $C_{26}H_{41}BNaO_7$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 499.2838; *found*: 499.2838.



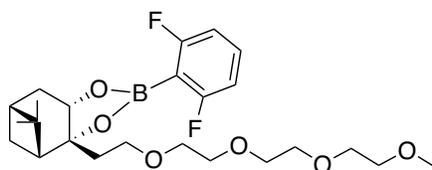
**2-4o**

**HRMS** (ESI-TOF): Not available.



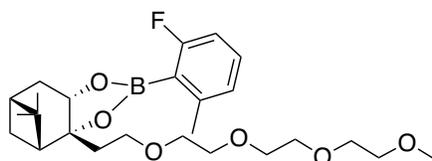
**2-4p**

**HRMS** (ESI-TOF): for C<sub>26</sub>H<sub>41</sub>BNaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 483.2888; *found*: 483.2893.



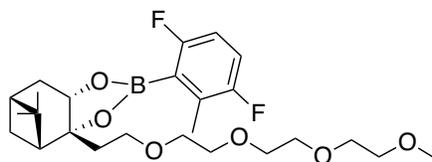
**2-4q**

**HRMS** (ESI-TOF): for C<sub>24</sub>H<sub>35</sub>BF<sub>2</sub>NaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 491.2387; *found*: 491.2387.



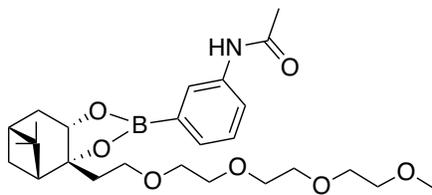
**2-4r**

**HRMS** (ESI-TOF): for C<sub>25</sub>H<sub>37</sub>BFNaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 487.2638; *found*: 487.2635.



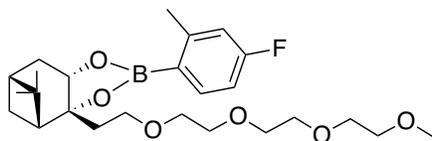
**2-4s**

**HRMS** (ESI-TOF): for C<sub>25</sub>H<sub>37</sub>BF<sub>2</sub>NaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 505.2543; *found*: 505.2545.



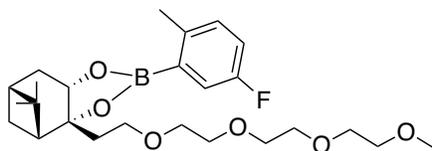
**2-4t**

**HRMS** (ESI-TOF): for  $C_{26}H_{40}BNNaO_7$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 512.279; *found*: 512.2789.



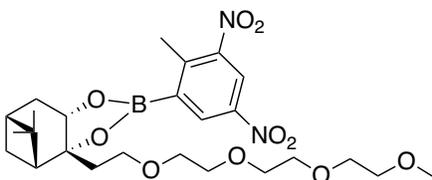
**2-4u**

**HRMS** (ESI-TOF): for  $C_{25}H_{38}BFNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 487.2638; *found*: 487.2639.



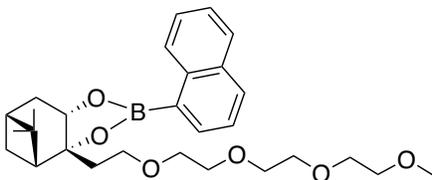
**2-4v**

**HRMS** (ESI-TOF): for  $C_{25}H_{38}BFNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 487.2638; *found*: 487.2637.



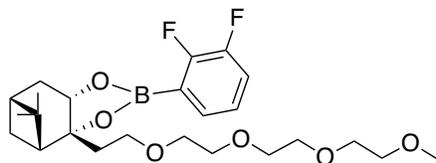
**2-4w**

**HRMS** (ESI-TOF): for  $C_{25}H_{37}BN_2NaO_{10}$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 559.2433; *found*: 559.2426.



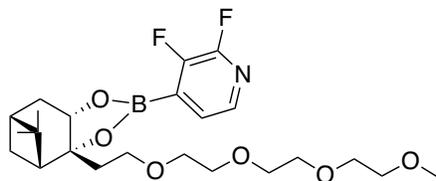
**2-4x**

**HRMS** (ESI-TOF): for  $C_{28}H_{39}BNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 505.2732; *found*: 505.2731.



**2-4y**

**HRMS** (ESI-TOF): for  $C_{24}H_{35}BF_2NaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 491.2387; *found*: 491.2385.



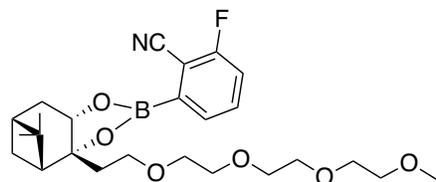
**2-4z**

**HRMS** (ESI-TOF): for  $C_{23}H_{34}BF_2NNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 492.2339; *found*: 492.2342.



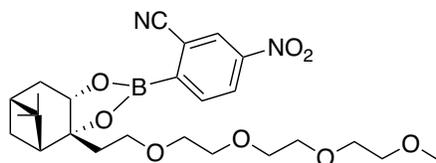
**2-4aa**

**HRMS** (ESI-TOF): for  $C_{24}H_{34}BF_3NaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 509.2293; *found*: 509.2292.



**2-4ab**

**HRMS** (ESI-TOF): for  $C_{25}H_{35}BFNNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 498.2434; *found*: 498.2432.

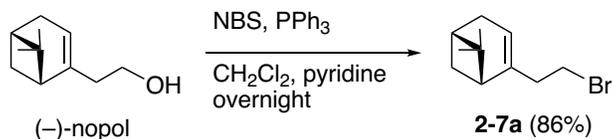


**2-4ac**

**HRMS** (ESI-TOF): for  $C_{25}H_{35}BN_2NaO_8$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 525.2379; *found*: 525.238.

## Synthesis of (-)-nopol-PEG-diol-2 (2-2b)

### (-)-Nopol-bromide (2-7a):



**(1*R*,5*S*)-2-(2-Bromoethyl)-6,6-dimethylbicyclo[3.1.1]hept-2-ene (2-7a):** Triphenylphosphine (2.0 equiv, 6.1 g, 23 mmol) was dissolved in dry DCM (23 ml) under nitrogen. To this solution, *N*-bromosuccinimide (NBS) (2.0 equiv, 4.2 g, 23 mmol) was added in small portions over 5 min under ice-water bath. Then, the ice-water bath was removed and the resulting deep red mixture was stirred at room temperature for 30 min. Then pyridine (1 ml) was added and the color turned to reddish-brown. (-)-Nopol (1.0 equiv, 2.0 ml, 12 mmol) was added to the mixture dropwise over 10 min. The reaction mixture was stirred overnight at room temperature under nitrogen balloon. Then, the mixture was diluted with hexanes (40 ml) and filtered through a plug of silica gel. In order to obtain high yield, the reaction flask was stirred three times with EtOAc:hexanes (25:25 ml) around 5 min and filtered through the silica gel plug. Then, it was concentrated in vacuo and the remaining off-white solid powder was stirred with hexanes (100 ml) for 15 min. Then it was filtered and concentrated in vacuo. The crude residue (colorless oil) was purified by flash chromatography (hexanes) to obtain (-)-nopol-bromide (**2-7a**) (2.3 g, 86% yield).

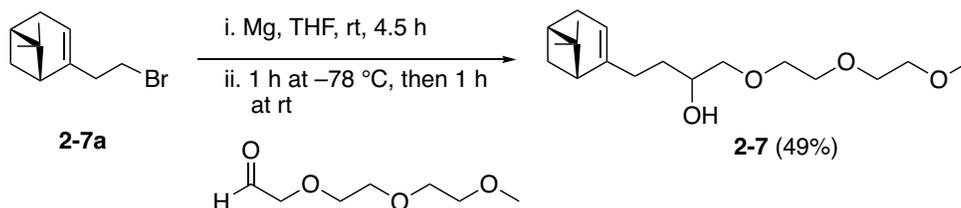
**<sup>1</sup>H NMR** δ/ppm: (400 MHz, CDCl<sub>3</sub>) 5.34 – 5.30 (m, 1H), 3.41 – 3.30 (m, 2H), 2.56 – 2.47 (comp m, 2H), 2.37 (app dt, *J* = 8.6, 5.6 Hz, 1H), 2.31 – 2.15 (comp m, 2H), 2.12 – 2.06 (m, 1H), 2.01 (app td, *J* = 5.6, 1.6 Hz, 1H), 1.28 (s, 3H), 1.17 (d, *J* = 8.6 Hz, 1H), 0.84 (s, 3H).

**<sup>13</sup>C NMR** δ/ppm: (101 MHz, CDCl<sub>3</sub>) 145.2, 119.2, 45.5, 40.7, 40.5, 38.1, 31.7, 31.4, 30.9, 26.3, 21.3.

**IR** (Microscope, cm<sup>-1</sup>): 3027, 2986, 2968, 2916, 1468, 1444, 1266.

**HRMS** (EI): for C<sub>11</sub>H<sub>17</sub>Br (M)<sup>+</sup>: *calcd.*: 228.05136; *found*: 228.05084.

**(-)-Nopol-PEG-OH (2-7):**



**4-((1R,5S)-6,6-Dimethylbicyclo[3.1.1]hept-2-en-2-yl)-1-(2-(2-methoxyethoxy)ethoxy)butan-2-ol (2-7)**

**Preparation of Grignard reagent:** A solution of **2-7a** (1.0 equiv, 1.8 g, 8.0 mmol) in dry THF (4 ml) was added dropwise to a suspension of pre-activated magnesium metal (2.0 equiv, 0.39 g, 16 mmol) in dry THF (1 ml) in a flamed dried three-necked flask coupled to a reflux condenser. The reaction mixture was stirred at room temperature for 4.5 h.

**Addition of Grignard Reagent to the Aldehyde:** Grignard reagent solution (2 ml) was added slowly over 1 h to a solution of the freshly prepared crude [2-(2-methoxyethoxy)ethoxy]acetaldehyde in THF (10 ml) at -78 °C. It was prepared from 11.8 mmol triethylene glycol monomethyl ether and the crude aldehyde was used without distillation.<sup>19</sup> <sup>1</sup>H NMR of crude aldehyde showed that it contained ~4.0 mmol of desired aldehyde. Then the reaction mixture was brought to rt, stirred for 1 h, and quenched with saturated aqueous solution of NH<sub>4</sub>Cl (5 ml). The organic phase was dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The resulting residue (yellowish oil) was purified by flash chromatography (EtOAc) to obtain an inseparable mixture of diastereomers of (-)-nopol-PEG-OH (**2-7**) (0.61 g, 49% yield).

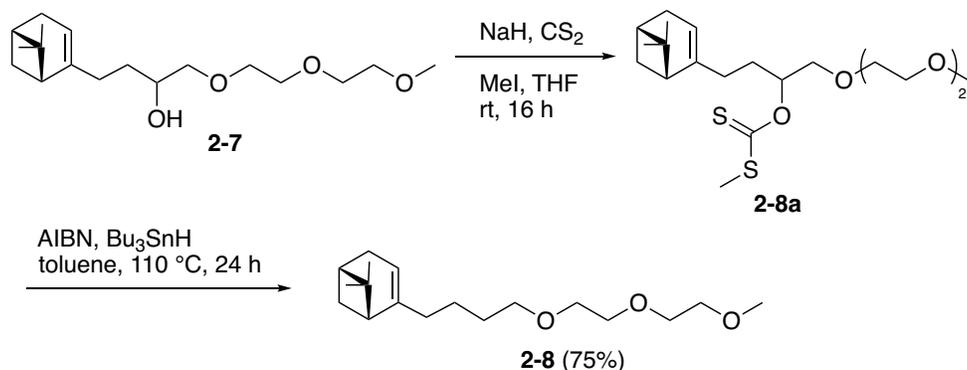
Yellowish oil:  $R_f = 0.48$  (EtOAc);

<sup>1</sup>H NMR  $\delta$ /ppm: (400 MHz, CDCl<sub>3</sub>) 5.23 – 5.17 (m, 1H), 3.87 – 3.73 (m, 1H), 3.73 – 3.58 (comp m, 6H), 3.58 – 3.44 (comp m, 3H), 3.38 (s, 3H), 3.31 (app ddd,  $J = 9.8, 8.2, 1.5$  Hz, 1H), 2.65 (bs, 1H), 2.34 (app dtd,  $J = 8.3, 5.6, 2.6$  Hz, 1H), 2.28 – 1.90 (comp m, 6H), 1.63 – 1.33 (comp m, 2H), 1.26 (s, 3H), 1.13 (dd,  $J = 8.5, 5.6$  Hz, 1H), 0.81 (d,  $J = 4.4$  Hz, 3H).

<sup>13</sup>C NMR  $\delta$ /ppm: (126 MHz, CDCl<sub>3</sub>) 147.99, 147.87, 116.16, 116.11, 75.8, 75.7, 72.0, 70.64, 70.60, 70.58, 70.1, 70.0, 59.1, 45.9, 45.8, 40.93, 40.91, 38.01, 37.96, 32.83, 32.82, 31.72, 31.68, 31.3, 30.62, 30.58, 26.38, 26.37, 21.3, 21.2.

IR (Microscope, cm<sup>-1</sup>): 3466, 2917, 2875, 1453, 1109.

HRMS (ESI-TOF): for C<sub>18</sub>H<sub>32</sub>NaO<sub>4</sub> (M + Na)<sup>+</sup>: *calcd.*: 335.2193; *found*: 335.2193.



***O*-(4-((1*R*,5*S*)-6,6-Dimethylbicyclo[3.1.1]hept-2-en-2-yl)-1-(2-(2-**

**methoxyethoxy)ethoxy)butan-2-yl) *S*-methyl carbonodithioate (2-8a):** To a solution of (–)-nopol-PEG-OH (**2-7**) (1.0 equiv, 0.26 g, 0.82 mmol) in THF (4.5 ml), sodium hydride (60% (w/w) dispersion in mineral oil (5.0 equiv, 0.16 g, 4.1 mmol) was added in small portions over 5 min at 0 °C under nitrogen. The reaction mixture was stirred for 10 min till the bubbles disappeared. Then carbon disulfide (15 equiv, 0.74 ml, 12 mmol) was added dropwise to the reaction mixture at room temperature. After the reaction mixture was stirred for 10 min, iodo methane (5.0 equiv, 0.26 ml, 4.1 mmol) was added dropwise at room temperature and the reaction mixture was stirred for 16 h. The reaction mixture was directly purified by flash chromatography (dry loading) (1:4, EtOAc/hexanes) to obtain (–)-nopol-xanthate (**2-8a**) (0.27 g, 81% yield).

Yellow oil:  $R_f = 0.44$  (1:4, EtOAc/hexanes).

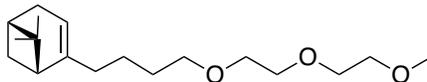
**<sup>1</sup>H NMR**  $\delta$ /ppm: (498 MHz, CDCl<sub>3</sub>) 6.00 – 5.56 (m, 1H), 5.23 – 5.18 (m, 1H), 3.76 – 3.58 (comp m, 8H), 3.57 – 3.50 (m, 2H), 3.37 (s, 3H), 2.55 (s,  $J = 0.8$  Hz, 3H), 2.40 – 2.29 (m, 1H), 2.29 – 2.11 (comp m, 2H), 2.11 – 1.93 (comp m, 4H), 1.93 – 1.71 (comp m, 2H), 1.26 (s, 3H), 1.14 (dd,  $J = 8.5, 6.9$  Hz, 1H), 0.81 (s, 3H).

**<sup>13</sup>C NMR**  $\delta$ /ppm: (126 MHz, CDCl<sub>3</sub>) 215.9, 147.0, 146.9, 116.7, 82.5, 82.4, 72.0, 71.5, 71.01, 71.00, 70.68, 70.66, 59.1, 45.78, 45.72, 40.9, 40.8, 38.02, 37.98, 32.4, 32.3, 31.73, 31.66, 31.3, 28.5, 28.3, 26.3, 21.3, 21.2, 19.0.

**IR** (Microscope, cm<sup>-1</sup>): 2983, 2918, 2832, 1449, 1381, 1219, 1110, 1055.

**HRMS** (ESI-TOF): for C<sub>20</sub>H<sub>38</sub>NO<sub>4</sub>S<sub>2</sub> (M + NH<sub>4</sub>)<sup>+</sup>: *calcd.*: 420.2237; *found*: 420.2238; for C<sub>20</sub>H<sub>34</sub>NaO<sub>4</sub>S<sub>2</sub> (M + Na)<sup>+</sup>: *calcd.*: 425.1791; *found*: 425.1794.

**(-)-Nopol-PEG-2 (2-8):**



**(1R,5S)-2-(4-(2-(2-Methoxyethoxy)ethoxy)butyl)-6,6-dimethylbicyclo[3.1.1]hept-2-ene (2-8):**

To a dry and deoxygenated solution of **2-8a** (1.0 equiv, 0.27 g, 0.67 mmol) in toluene (8 ml, dry, degassed and purged with argon before usage) in a flame dried three-necked flask, a mixture of tributyltin hydride (4.0 equiv, 0.72 ml, 2.7 mmol) and catalytic AIBN in toluene (3 ml, dry, degassed and purged with argon before usage) was added dropwise *via* a syringe over 4 h under argon at 110 °C reflux. Then the reaction mixture was stirred for 16 h at reflux. The mixture was concentrated in vacuo and the crude residue (yellowish oil) was purified by flash chromatography (3.5:6.5, EtOAc/hexanes) to obtain (-)-nopol-PEG-2 (**2-8**) (0.18 g, 91% yield, 75% yield over two steps).

Yellowish oil:  $R_f = 0.62$  (3.5:6.5, EtOAc/hexanes).

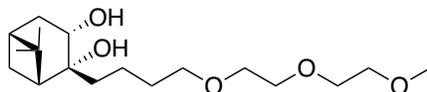
$[\alpha]_{\text{D}}^{20} : -5.0$  ( $c$  0.53,  $\text{CHCl}_3$ ).

$^1\text{H NMR}$   $\delta$ /ppm: (400 MHz,  $\text{CDCl}_3$ ) 5.18 – 5.14 (m, 1H), 3.69 – 3.62 (comp m, 4H), 3.62 – 3.52 (comp m, 4H), 3.45 (app t,  $J = 6.8$  Hz, 2H), 3.38 (s, 3H), 2.33 (app dt,  $J = 8.4, 5.6$  Hz, 1H), 2.29 – 2.11 (comp m, 2H), 2.09 – 2.03 (m, 1H), 1.99 (app td,  $J = 5.6, 1.5$  Hz, 1H), 1.97 – 1.90 (comp m, 2H), 1.62 – 1.52 (comp m, 2H), 1.49 – 1.28 (comp m, 2H), 1.26 (s, 3H), 1.13 (d,  $J = 8.4$  Hz, 1H), 0.81 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 148.3, 115.9, 72.0, 71.4, 70.7, 70.6, 70.1, 59.1, 45.8, 41.0, 38.0, 36.8, 31.7, 31.3, 29.5, 26.4, 23.7, 21.2.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3024, 2983, 2915, 2870, 2834, 1454, 1351, 1364, 1113.

**HRMS** (ESI-TOF): for  $\text{C}_{18}\text{H}_{36}\text{NO}_3$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 314.2690; *found*: 314.2688; for  $\text{C}_{18}\text{H}_{32}\text{NaO}_3$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 319.2244; *found*: 319.2243.



**(1R,2R,3S,5R)-2-(4-(2-(2-Methoxyethoxy)ethoxy)butyl)-6,6-dimethylbicyclo[3.1.1]heptane-**

**2,3-diol (2-2b):** *N*-Methylmorpholine *N*-oxide (NMO) as a solid (1.2 equiv, 0.15 g, 1.3 mmol) and quinuclidine (1.2 equiv, 0.14 g, 1.3 mmol) were added to compound **2-8** (1.0 equiv, 0.32 g,

1.1 mmol) in acetone:water (10:2 ml). Lastly,  $K_2OsO_4 \cdot 2H_2O$  purchased from strem chemicals (5 mol%, 22 mg, 0.06 mmol) was added to the reaction mixture. The mixture was stirred and refluxed at 50 °C for 48 h under nitrogen balloon. Then, the reaction mixture was concentrated in vacuo and mixed with EtOAc (50 ml). The organic phase was washed with HCl (1 × 5.0 ml, 1 N), distilled water (1 × 5.0 ml) and brine (1 × 5.0 ml), dried ( $MgSO_4$ ), filtered and concentrated in vacuo. The crude residue (yellowish oil) was purified by flash chromatography (1:3, acetone/hexanes) to obtain (–)-nopol-PEG-diol-2 (**2-2b**) (0.24 g, 65% yield).

Colorless oil:  $R_f = 0.22$  (1:3, acetone/hexanes)

$[\alpha]_D^{20}$  : –5.0 (*c* 0.53,  $CHCl_3$ ).

$^1H$  NMR  $\delta$ /ppm: (400 MHz,  $CDCl_3$ ) 3.99 (dd,  $J = 9.3, 4.9$  Hz, 1H), 3.68 – 3.53 (comp m, 9H), 3.49 (app td,  $J = 6.2, 0.9$  Hz, 2H), 3.38 (s, 3H), 2.47 (dddd,  $J = 14.0, 9.3, 3.6, 2.4$  Hz, 1H), 2.19 (app dtd,  $J = 10.3, 6.1, 2.4$  Hz, 1H), 2.07 (app t,  $J = 5.8$  Hz, 1H), 1.93 – 1.88 (m, 1H), 1.73 – 1.45 (comp m, 8H), 1.38 (d,  $J = 10.3$  Hz, 1H), 1.26 (s, 3H), 0.92 (s, 3H).

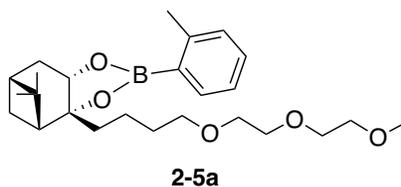
$^{13}C$  NMR  $\delta$ /ppm: (101 MHz,  $CDCl_3$ ) 75.6, 72.04, 71.3, 70.7, 70.6, 70.2, 68.3, 59.1, 52.0, 41.8, 40.7, 39.0, 38.6, 30.0, 27.9, 27.8, 24.3, 19.8.

IR (Microscope,  $cm^{-1}$ ): 3407, 2924, 2868, 1453, 1113.

HRMS (ESI-TOF): for  $C_{18}H_{34}NaO_5$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 353.2298; *found*: 353.2297.

### **General Synthesis of (–)-Nopol-PEG-boronates-2 (2-5a – 2-5p)**

Compound **2-1b** (1 equiv, 0.1 mmol) was mixed with arylboronic acid (1 equiv, 0.1 mmol) in THF (1ml). The reaction mixture was stirred for 1 h at room temperature under ambient atmosphere. Then, toluene (0.2 ml) was added to the mixture and it was concentrated in vacuo. The crude residue was purified by flash chromatography.



### **(3aR,4R,6R,7aS)-3a-(4-(2-(2-Methoxyethoxy)ethoxy)butyl)-5,5-dimethyl-2-(*o*-tolyl)hexahydro-4,6-methanobenzo[d][1,3,2]dioxaborole (2-5a)**

Colorless oil (36 mg, 85% yield):  $R_f = 0.65$  (1:3, acetone/hexanes).

$[\alpha]_{\text{D}}^{20}$  : -11.5 (*c* 0.48, CHCl<sub>3</sub>).

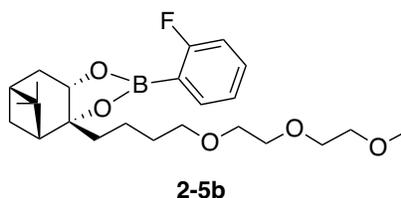
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.78 (dd, *J* = 7.7, 1.6 Hz, 1H), 7.32 (app td, *J* = 7.5, 1.5 Hz, 1H), 7.19 – 7.14 (comp m, 2H), 4.47 (dd, *J* = 8.9, 1.9 Hz, 1H), 3.65 – 3.59 (comp m, 4H), 3.59 – 3.55 (comp m, 2H), 3.55 – 3.50 (comp m, 2H), 3.48 – 3.42 (m, 2H), 3.37 (s, 3H), 2.54 (s, 3H), 2.43 (app ddt, *J* = 13.9, 8.9, 2.3 Hz, 1H), 2.31 – 2.13 (comp m, 2H), 2.04 – 1.88 (comp m, 2H), 1.72 (app t, *J* = 9.1, 2H), 1.66 – 1.57 (comp m, 2H), 1.57 – 1.41 (comp m, 2H), 1.31 (s, 3H), 1.25 (d, *J* = 9.8 Hz, 1H), 0.88 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 144.9, 136.1, 130.8, 129.9, 124.8, 88.0, 77.3, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 49.8, 40.5, 39.9, 38.2, 36.0, 30.0, 27.2, 26.4, 24.1, 22.4, 19.9;

<sup>11</sup>B NMR δ/ppm: (128 MHz, CDCl<sub>3</sub>) 31.3.

IR (Microscope, cm<sup>-1</sup>): 2924, 2854, 1601, 1445, 1348, 1124.

HRMS (ESI-TOF): for C<sub>25</sub>H<sub>43</sub>BNO<sub>5</sub> (M + NH<sub>4</sub>)<sup>+</sup>: *calcd.*: 448.3229; *found*: 448.3237; for C<sub>25</sub>H<sub>39</sub>BNaO<sub>5</sub> (M + Na)<sup>+</sup>: *calcd.*: 453.2783; *found*: 453.2790.



**(3a*R*,4*R*,6*R*,7a*S*)-2-(2-Fluorophenyl)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborole (2-5b)**

Colorless oil (18 mg, 70% yield): *R<sub>f</sub>* = 0.47 (1:3, acetone/hexanes).

$[\alpha]_{\text{D}}^{20}$  : -11.8 (*c* 0.26, CHCl<sub>3</sub>).

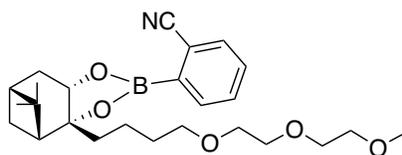
<sup>1</sup>H NMR δ/ppm: (700 MHz, CDCl<sub>3</sub>) 7.74 (ddd, *J* = 7.7, 6.0, 1.9 Hz, 1H), 7.44 (dddd, *J* = 8.2, 7.4, 5.5, 1.9 Hz, 1H), 7.14 (app t, *J* = 7.5, 1H), 7.03 (app t, *J* = 9.2, 1H), 4.51 (dd, *J* = 8.9, 2.0 Hz, 1H), 3.64 – 3.59 (comp m, 4H), 3.59 – 3.54 (comp m, 2H), 3.54 – 3.50 (comp m, 2H), 3.48 – 3.42 (comp m, 2H), 3.36 (s, 3H), 2.43 (app ddt, *J* = 14.6, 8.9, 2.4 Hz, 1H), 2.29 – 2.18 (comp m, 2H), 2.00 (ddd, *J* = 14.6, 3.4, 2.0 Hz, 1H), 1.96 – 1.93 (m, 1H), 1.76 – 1.69 (comp m, 2H), 1.63 – 1.57 (comp m, 2H), 1.56 – 1.49 (m, 1H), 1.49 – 1.39 (m, 1H), 1.31 (s, 3H), 1.24 (d, *J* = 10.4 Hz, 1H), 0.88 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 167.3 (d,  $J_{\text{CF}} = 250.7$  Hz), 136.9 (d,  $J_{\text{CF}} = 8.0$  Hz), 133.2 (d,  $J_{\text{CF}} = 8.7$  Hz), 123.7 (d,  $J_{\text{CF}} = 3.2$  Hz), 115.3 (d,  $J_{\text{CF}} = 23.9$  Hz), 88.3, 76.5, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 49.8, 40.4, 39.8, 38.2, 35.8, 30.0, 27.2, 26.4, 24.1, 19.7.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 29.5.

IR (Microscope,  $\text{cm}^{-1}$ ): 2927, 2870, 1615, 1489, 1448, 1364, 1115.

HRMS (ESI-TOF): for  $\text{C}_{24}\text{H}_{40}\text{BFNO}_5$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 452.2978; *found*: 452.2979; for  $\text{C}_{24}\text{H}_{36}\text{BFNaO}_5$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 457.2532; *found*: 457.2535.



**2-5c**

**2-((3aR,4R,6R,7aS)-3a-(4-(2-(2-Hydroxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl)benzonitrile (2-5c)**

Colorless oil (23 mg, 83% yield):  $R_f = 0.53$  (1:1, EtOAc/hexanes).

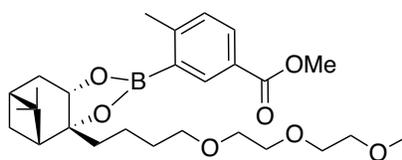
$^1\text{H}$  NMR  $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 7.90 (dd,  $J = 7.5, 1.3$  Hz, 1H), 7.70 (d,  $J = 7.5$ , 1H), 7.58 (app td,  $J = 7.5, 1.3$  Hz, 1H), 7.53 (app td,  $J = 7.6, 1.5$  Hz, 1H), 4.56 (dd,  $J = 8.8, 1.9$  Hz, 1H), 3.64 – 3.60 (comp m, 4H), 3.60 – 3.55 (comp m, 2H), 3.55 – 3.51 (comp m, 2H), 3.49 – 3.44 (comp m, 2H), 3.36 (s, 3H), 2.44 (app ddt,  $J = 14.6, 8.9, 2.3$  Hz, 1H), 2.30 – 2.21 (comp m, 2H), 2.05 (ddd,  $J = 14.8, 3.4, 1.9$  Hz, 1H), 1.97 – 1.94 (m, 1H), 1.79 – 1.71 (comp m, 2H), 1.65 – 1.53 (comp m, 3H), 1.50 – 1.43 (m, 1H), 1.31 (s, 3H), 1.23 (d,  $J = 9.9$  Hz, 1H), 0.88 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 136.0, 133.5, 131.6, 131.1, 119.1, 117.4, 89.2, 77.3, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 49.5, 40.3, 39.8, 38.2, 35.6, 29.9, 27.1, 26.4, 24.1, 19.8.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 29.5.

IR (Microscope,  $\text{cm}^{-1}$ ): 2925, 2870, 2227, 1595, 1488, 1443, 1388, 1362, 1121.

HRMS (ESI-TOF): for  $\text{C}_{25}\text{H}_{40}\text{BN}_2\text{O}_5$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 459.3025; *found*: 459.3036; for  $\text{C}_{25}\text{H}_{36}\text{BNNaO}_5$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 464.2579; *found*: 464.2582.



**2-5d**

**Methyl 3-((3a*R*,4*R*,6*R*,7a*S*)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborol-2-yl)-4-methylbenzoate (2-5d)**

Colorless oil (18 mg, 75% yield):  $R_f = 0.65$  (1:1, EtOAc/hexanes).

$[\alpha]_D^{20} : -6.1$  ( $c$  0.49,  $\text{CHCl}_3$ ).

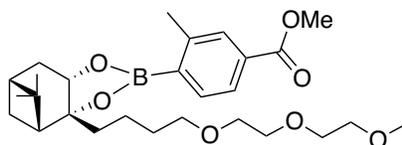
$^1\text{H NMR}$   $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 8.42 (d,  $J = 2.0$  Hz, 1H), 7.98 (dd,  $J = 8.0, 2.0$  Hz, 1H), 7.24 (d,  $J = 7.9$ , 1H), 4.48 (dd,  $J = 8.8, 2.0$  Hz, 1H), 3.90 (s, 3H), 3.64 – 3.60 (comp m, 4H), 3.58 – 3.56 (comp m, 2H), 3.54 – 3.50 (comp m, 2H), 3.46 (app td,  $J = 6.6, 2.8$  Hz, 2H), 3.36 (s, 3H), 2.58 (s, 3H), 2.43 (app ddt,  $J = 14.6, 8.8, 2.4$  Hz, 1H), 2.27 – 2.21 (comp m, 2H), 1.99 (ddd,  $J = 14.6, 3.4, 2.0$  Hz, 1H), 1.96 – 1.93 (m, 1H), 1.77 – 1.69 (comp m, 2H), 1.64 – 1.58 (comp m, 2H), 1.55 – 1.49 (m, 1H), 1.49 – 1.40 (m, 1H), 1.31 (s, 3H), 1.23 (d,  $J = 10.4$  Hz, 1H), 0.88 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 167.3, 150.5, 137.3, 131.9, 130.1, 126.9, 88.4, 76.4, 72.0, 71.3, 70.1, 70.6, 70.2, 59.1, 51.9, 49.8, 40.5, 39.9, 38.2, 35.9, 30.0, 27.2, 26.4, 24.1, 22.6, 19.8.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 31.1

**IR** (Microscope,  $\text{cm}^{-1}$ ): 2928, 2869, 1722, 1606, 1436, 1410, 1346, 1282, 1256, 1117.

**HRMS** (ESI-TOF): for  $\text{C}_{27}\text{H}_{45}\text{BNO}_7$  ( $\text{M} + \text{NH}_4$ ) $^+$  *calcd.*: 506.3284 *found*: 506.3284; for  $\text{C}_{27}\text{H}_{41}\text{BNaO}_7$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 511.2838; *found*: 511.2839.



**2-5e**

**Methyl 4-((3a*R*,4*R*,6*R*,7a*S*)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborol-2-yl)-3-methylbenzoate (2-5e)**

Colorless oil (20 mg, 81% yield):  $R_f = 0.69$  (1:1, EtOAc/hexanes).

$^1\text{H NMR}$   $\delta$ /ppm: (498 MHz,  $\text{CDCl}_3$ ) 7.84 – 7.78 (comp m, 3H), 4.48 (dd,  $J = 8.8, 1.9$  Hz, 1H), 3.91 (s, 3H), 3.64 – 3.60 (comp m, 4H), 3.59 – 3.54 (comp m, 2H), 3.54 – 3.49 (comp m, 2H),

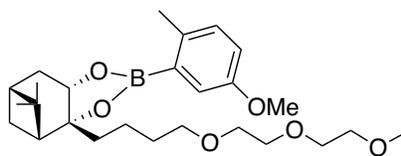
3.46 (app td,  $J = 6.5, 1.2$  Hz, 2H), 3.36 (s, 3H), 2.57 (s, 3H), 2.48 – 2.32 (m, 1H), 2.30 – 2.14 (comp m, 2H), 2.05 – 1.89 (comp m, 2H), 1.78 – 1.67 (comp m, 2H), 1.65 – 1.57 (comp m, 2H), 1.55 – 1.37 (comp m, 2H), 1.31 (s, 3H), 1.23 (d,  $J = 10.4$  Hz, 1H), 0.88 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 167.3, 145.0, 136.0, 131.9, 130.5, 125.6, 88.4, 76.5, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 52.1, 49.7, 40.5, 39.9, 38.2, 35.9, 30.0, 27.2, 26.4, 24.1, 22.3, 19.9.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 31.0.

IR (Microscope,  $\text{cm}^{-1}$ ): 2928, 2869, 1725, 1561, 1498, 1437, 1348, 1296, 1198, 1115.

HRMS (ESI-TOF): for  $\text{C}_{27}\text{H}_{45}\text{BNO}_7$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 506.3288; *found*: 506.3284; for  $\text{C}_{27}\text{H}_{41}\text{BNaO}_7$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 511.2842; *found*: 511.2842.



**2-5f**

**(3aR,4R,6R,7aS)-2-(5-Methoxy-2-methylphenyl)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborole (2-5f)**

Colorless oil (16 mg, 94% yield):  $R_f = 0.38$  (3:7, EtOAc/hexanes).

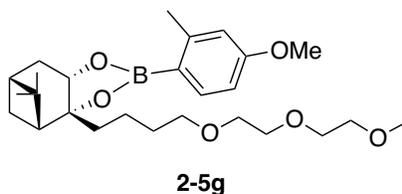
$^1\text{H}$  NMR  $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 7.31 (d,  $J = 2.9$  Hz, 1H), 7.09 (d,  $J = 8.3$  Hz, 1H), 6.88 (dd,  $J = 8.3, 3.0$  Hz, 1H), 4.46 (dd,  $J = 8.8, 2.0$  Hz, 1H), 3.80 (s, 3H), 3.66 – 3.60 (comp m, 4H), 3.60 – 3.55 (comp m, 2H), 3.55 – 3.50 (comp m, 2H), 3.48 – 3.41 (comp m, 2H), 3.37 (s, 3H), 2.46 (s, 3H), 2.42 (app ddt,  $J = 14.1, 8.3, 2.2$  Hz, 1H), 2.27 – 2.16 (comp m, 2H), 2.01 – 1.88 (comp m, 2H), 1.76 – 1.67 (comp m, 2H), 1.65 – 1.59 (comp m, 2H), 1.57 – 1.50 (m, 1H), 1.50 – 1.41 (m, 1H), 1.31 (s, 3H), 1.24 (d,  $J = 10.5$  Hz, 1H), 0.88 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 156.9, 136.9, 131.0, 120.4, 116.9, 88.1, 76.2, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 55.4, 49.8, 40.5, 39.9, 38.2, 36.0, 30.0, 27.2, 26.4, 24.1, 21.3, 19.8.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 30.4.

IR (Microscope,  $\text{cm}^{-1}$ ): 2932, 2868, 1570, 1497, 1446, 1415, 1343, 1286, 1237, 1122, 1067.

HRMS (ESI-TOF): for  $\text{C}_{26}\text{H}_{45}\text{BNO}_6$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 478.3339; *found*: 478.3339; for  $\text{C}_{26}\text{H}_{41}\text{BNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 483.2893; *found*: 483.2894.



**(3aR,4R,6R,7aS)-2-(4-Methoxy-2-methylphenyl)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborole (2-5g)**

Colorless oil (20 mg, 58% yield):  $R_f = 0.69$  (1:1, EtOAc/hexanes).

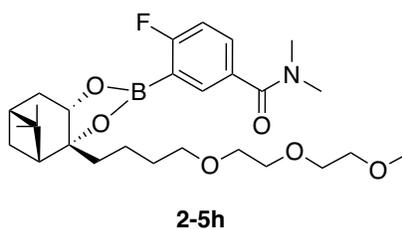
$^1\text{H NMR}$   $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 7.88 – 7.58 (m, 1H), 6.72 – 6.70 (comp m, 2H), 4.44 (dd,  $J = 8.8, 2.0$  Hz, 1H), 3.81 (s, 3H), 3.64 – 3.60 (comp m, 4H), 3.59 – 3.54 (comp m, 2H), 3.54 – 3.50 (comp m, 2H), 3.50 – 3.40 (comp m, 2H), 3.37 (s, 3H), 2.52 (s, 3H), 2.42 (app ddt,  $J = 14.5, 8.8, 2.4$  Hz, 1H), 2.27 – 2.16 (comp m, 2H), 2.03 – 1.85 (comp m, 2H), 1.76 – 1.65 (comp m, 2H), 1.63 – 1.57 (comp m, 2H), 1.56 – 1.49 (m, 1H), 1.48 – 1.40 (m, 1H), 1.30 (s, 3H), 1.24 (d,  $J = 10.3$  Hz, 1H), 0.87 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 161.7, 147.3, 138.0, 115.6, 110.2, 87.7, 76.1, 72.0, 71.4, 70.7, 70.6, 70.2, 59.1, 55.1, 49.8, 40.6, 39.9, 38.2, 36.0, 30.0, 27.2, 26.4, 24.1, 22.6, 19.8.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 30.7.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 2928, 2869, 1602, 1566, 1452, 1346, 1294, 1240, 1126, 1044.

**HRMS** (ESI-TOF): for  $\text{C}_{26}\text{H}_{45}\text{BNO}_6$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 478.3339; *found*: 478.3341; for  $\text{C}_{26}\text{H}_{41}\text{BNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 483.2893; *found*: 483.2894.



**4-Fluoro-3-((3aR,4R,6R,7aS)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-2-yl)-N,N-dimethylbenzamide (2-5h)**

Colorless oil (27 mg, 53% yield):  $R_f = 0.19$  (1:3, acetone/hexanes).

$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 7.82 (dd,  $J = 5.6, 2.3$  Hz, 1H), 7.51 (ddd,  $J = 8.4, 5.2, 2.4$  Hz, 1H), 7.06 (app t,  $J = 8.7$  Hz, 1H), 4.51 (dd,  $J = 8.9, 1.9$  Hz, 1H), 3.66 – 3.60 (comp m, 4H),

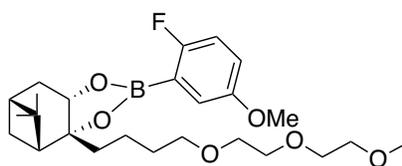
3.60 – 3.55 (comp m, 2H), 3.55 – 3.51 (comp m, 2H), 3.48 – 3.41 (comp m, 2H), 3.36 (s, 3H), 3.09 (s, 3H), 2.98 (s, 3H), 2.49 – 2.35 (m, 1H), 2.28 – 2.19 (comp m, 2H), 2.04 – 1.88 (comp m, 2H), 1.72 (app dd,  $J = 9.1, 7.1$  Hz, 2H), 1.64 – 1.56 (comp m, 2H), 1.55 – 1.47 (m, 1H), 1.47 – 1.38 (m, 1H), 1.30 (s, 3H), 1.20 (d,  $J = 10.4$  Hz, 1H), 0.87 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 170.8, 167.6 (d,  $J_{\text{CF}} = 254.3$  Hz), 136.1 (d,  $J_{\text{CF}} = 8.7$  Hz), 132.3 (d,  $J_{\text{CF}} = 9.3$  Hz), 132.0 (d,  $J_{\text{CF}} = 3.5$  Hz), 115.4 (d,  $J_{\text{CF}} = 24.7$  Hz), 88.5, 76.7, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 49.7, 40.3, 39.8, 38.2, 35.7, 35.4, 30.0, 27.2, 26.4, 24.1, 19.7.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 28.9.

IR (Microscope,  $\text{cm}^{-1}$ ): 2925, 2870, 1641, 1610, 1484, 1454, 1391, 1345, 1124.

HRMS (ESI-TOF): for  $\text{C}_{27}\text{H}_{42}\text{BFNO}_6$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 506.3089; *found*: 506.3088; for  $\text{C}_{27}\text{H}_{41}\text{BFNNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 528.2908; *found*: 528.2908.



2-5i

**2-(2-(4-((3aR,4R,6R,7aS)-2-(2-Fluoro-5-methoxyphenyl)-5,5-dimethyltetrahydro-4,6-methanobenzo[d][1,3,2]dioxaborol-3a(4H)-yl)butoxy)ethoxy)ethan-1-ol (2-5i)**

Colorless oil (19 mg, 68% yield):  $R_f = 0.41$  (1:3, acetone/hexanes).

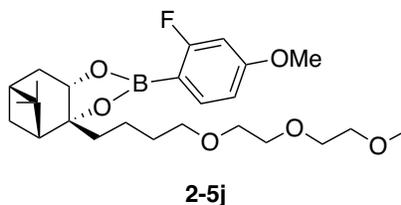
$^1\text{H}$  NMR  $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 7.21 – 7.19 (m, 1H), 6.99 – 6.91 (comp m, 2H), 4.51 (dd,  $J = 8.8, 2.0$  Hz, 1H), 3.80 (s, 3H), 3.63 – 3.60 (comp m, 4H), 3.59 – 3.55 (comp m, 2H), 3.54 – 3.50 (comp m, 2H), 3.48 – 3.41 (comp m, 2H), 3.37 (s, 3H), 2.43 (app ddt,  $J = 14.6, 8.9, 2.5$  Hz, 1H), 2.30 – 2.19 (comp m, 2H), 2.00 (ddd,  $J = 14.7, 3.4, 2.0$  Hz, 1H), 1.96 – 1.93 (m, 1H), 1.77 – 1.67 (comp m, 2H), 1.65 – 1.56 (comp m, 2H), 1.56 – 1.49 (m, 1H), 1.48 – 1.40 (m, 1H), 1.31 (s, 3H), 1.24 (d,  $J = 10.6$  Hz, 1H), 0.87 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 161.7 (d,  $J_{\text{CF}} = 243.4$  Hz), 155.3 (d,  $J_{\text{CF}} = 2.1$  Hz), 120.1 (d,  $J_{\text{CF}} = 8.2$  Hz), 119.1 (d,  $J_{\text{CF}} = 8.8$  Hz), 116.1 (d,  $J_{\text{CF}} = 26.3$  Hz), 88.4, 76.5, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 55.9, 49.8, 40.3, 39.8, 38.2, 35.8, 30.0, 27.2, 26.4, 24.1, 19.7.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 30.0.

IR (Microscope,  $\text{cm}^{-1}$ ): 2932, 2870, 1614, 1583, 1493, 1450, 1420, 1348, 1293, 1208, 1123, 1067.

**HRMS** (ESI-TOF): for  $C_{25}H_{42}BFNO_6$  ( $M + NH_4$ )<sup>+</sup>: *calcd.*: 482.3088; *found*: 482.3088; for  $C_{25}H_{38}BFNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 487.2643; *found*: 487.2642.



**(3a*R*,4*R*,6*R*,7a*S*)-2-(2-Fluoro-4-methoxyphenyl)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborole (2-5j)**

Colorless oil (17 mg, 95% yield):  $R_f = 0.47$  (1:3, acetone/hexanes).

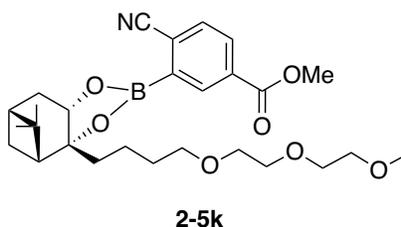
**<sup>1</sup>H NMR**  $\delta$ /ppm: (500 MHz,  $CDCl_3$ ) 7.65 (dd,  $J = 8.3, 7.2$  Hz, 1H), 6.70 (dd,  $J = 8.4, 2.3$  Hz, 1H), 6.58 (dd,  $J = 11.3, 2.3$  Hz, 1H), 4.49 (dd,  $J = 8.8, 2.0$  Hz, 1H), 3.82 (s, 3H), 3.65 – 3.59 (comp m, 4H), 3.59 – 3.55 (comp m, 2H), 3.55 – 3.51 (comp m, 2H), 3.48 – 3.41 (comp m, 2H), 3.37 (s, 3H), 2.42 (app ddt,  $J = 13.5, 8.8, 2.3$  Hz, 1H), 2.28 – 2.18 (comp m, 2H), 2.03 – 1.90 (comp m, 2H), 1.76 – 1.65 (comp m, 2H), 1.65 – 1.55 (comp m, 2H), 1.55 – 1.48 (m, 1H), 1.47 – 1.39 (m, 1H), 1.30 (s, 3H), 1.23 (d,  $J = 9.9$  Hz, 1H), 0.87 (s, 3H).

**<sup>13</sup>C NMR**  $\delta$ /ppm: (126 MHz,  $CDCl_3$ ) 168.6 (d,  $J_{CF} = 250.6$  Hz), 163.9 (d,  $J_{CF} = 11.7$  Hz), 137.8 (d,  $J_{CF} = 10.4$  Hz), 110.0 (d,  $J_{CF} = 2.8$  Hz), 101.2 (d,  $J_{CF} = 27.8$  Hz), 88.0, 76.3, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 55.5, 49.9, 40.4, 39.8, 38.2, 35.9, 30.0, 27.2, 26.4, 24.1, 19.7.

**<sup>11</sup>B NMR**  $\delta$ /ppm: (160 MHz,  $CDCl_3$ ) 28.9.

**IR** (Microscope,  $cm^{-1}$ ): 2933, 2868, 1619, 1570, 1359, 1286, 1129.

**HRMS** (ESI-TOF): for  $C_{25}H_{42}BFNO_6$  ( $M + NH_4$ )<sup>+</sup>: *calcd.*: 482.3088; *found*: 482.3088; for  $C_{25}H_{38}BFNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 487.2674; *found*: 487.2677.



**Methyl 4-cyano-3-((3a*R*,4*R*,6*R*,7a*S*)-3a-(4-(2-(2-methoxyethoxy)ethoxy)butyl)-5,5-dimethylhexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborol-2-yl)benzoate (2-5k)**

Colorless oil (22 mg, 51% yield):  $R_f = 0.73$  (3.5:6.5, acetone/hexanes).

$^1\text{H NMR}$   $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 8.53 (dd,  $J = 1.8, 0.6$  Hz, 1H), 8.18 (dd,  $J = 8.0, 1.8$  Hz, 1H), 7.79 (dd,  $J = 8.1, 0.6$  Hz, 1H), 4.58 (dd,  $J = 8.9, 2.0$  Hz, 1H), 3.96 (s, 3H), 3.64 – 3.61 (comp m, 4H), 3.60 – 3.56 (comp m, 2H), 3.55 – 3.51 (comp m, 2H), 3.48 – 3.44 (comp m, 2H), 3.37 (s, 3H), 2.45 (app ddt,  $J = 14.7, 8.9, 2.3$  Hz, 1H), 2.31 – 2.25 (comp m, 2H), 2.06 (ddd,  $J = 14.8, 3.5, 1.9$  Hz, 1H), 1.98 – 1.94 (m, 1H), 1.83 – 1.70 (comp m, 2H), 1.65 – 1.60 (comp m, 2H), 1.58 – 1.54 (m, 1H), 1.51 – 1.42 (m, 1H), 1.32 (s, 3H), 1.23 – 1.17 (m, 1H), 0.89 (s, 3H).

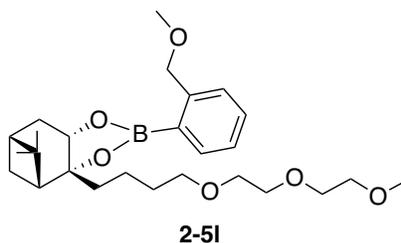
$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 165.7, 136.9, 133.6, 132.8, 131.9, 121.3, 118.3, 89.6, 77.4, 72.0, 71.3, 70.7, 70.6, 70.2, 59.1, 52.7, 49.5, 40.3, 39.8, 38.2, 35.5, 29.9, 27.1, 26.4, 24.1, 19.8.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 29.5.

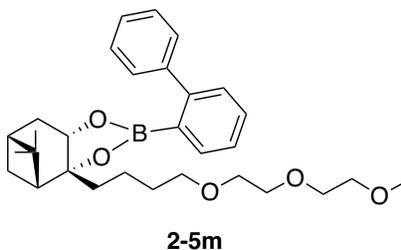
IR (Microscope,  $\text{cm}^{-1}$ ): 2925, 2869, 2228, 1730, 1604, 1415, 1276, 1115.

HRMS (ESI-TOF): for  $\text{C}_{27}\text{H}_{42}\text{BN}_2\text{O}_7$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 517.308; *found*: 517.3081; for  $\text{C}_{27}\text{H}_{38}\text{BNNaO}_7$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 522.2634; *found*: 522.2636.

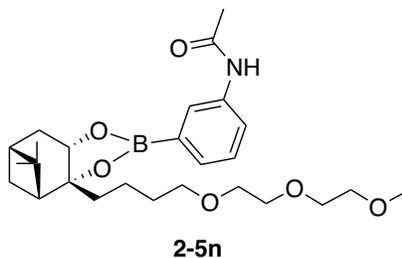
The following boronates **2-5l** – **2-5p** were also characterized as **2-5a** – **2-5k** but due to the large volume of data, their full characterization data were not provided in supporting information.



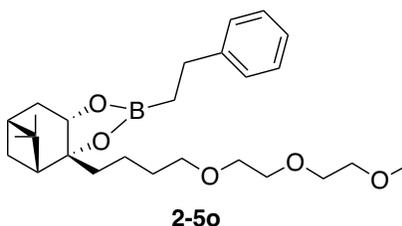
HRMS (ESI-TOF): for  $\text{C}_{26}\text{H}_{45}\text{BNO}_6$  ( $\text{M} + \text{NH}_4$ ) $^+$ : *calcd.*: 478.3334; *found*: 478.3334; for  $\text{C}_{26}\text{H}_{41}\text{BNaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 483.2888; *found*: 483.2888.



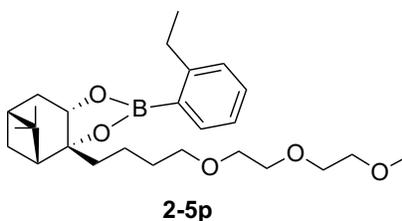
**HRMS** (ESI-TOF): for  $C_{30}H_{45}BNO_5$  ( $M + NH_4$ )<sup>+</sup>: *calcd.*: 510.3391; *found*: 510.339; for  $C_{30}H_{41}BNaO_5$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 515.2945; *found*: 515.2945.



**HRMS** (ESI-TOF): for  $C_{26}H_{44}BN_2O_6$  ( $M + NH_4$ )<sup>+</sup>: *calcd.*: 490.3323; *found*: 490.3329; for  $C_{26}H_{40}BNNaO_6$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 495.2877; *found*: 495.2883.



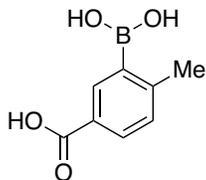
**HRMS** (ESI-TOF): for  $C_{26}H_{45}BNO_5$  ( $M + NH_4$ )<sup>+</sup>: *calcd.*: 461.3422; *found*: 461.343; for  $C_{26}H_{41}BNaO_5$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 466.2976; *found*: 466.2975.



**HRMS** (ESI-TOF): for  $C_{26}H_{45}BNO_5$  ( $M + NH_4$ )<sup>+</sup>: *calcd.*: 462.3385; *found*: 462.3389; for  $C_{26}H_{41}BNaO_5$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 467.2939; *found*: 467.2949.

#### **Synthesis of 4/5-carboxy-2-methylphenylboronic acid (2-9c – 2-16):**

Compounds 5-carboxy-2-methylphenylboronic acid (**2-9c**) and 4-carboxy-2-methylphenylboronic acid (**2-16**) were synthesized based on the described procedure by our group.<sup>25</sup>



### 3-Borono-4-methylbenzoic acid (2-9c)

White solid (0.21 g, 50% yield).

Mp = 172 – 174 °C.

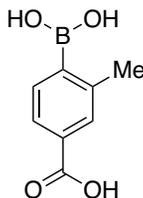
$^1\text{H NMR}$   $\delta$ /ppm: (400 MHz, acetone- $d_6$ ) 8.26 (d,  $J$  = 1.9 Hz, 1H), 7.86 (dd,  $J$  = 8.0, 2.0 Hz, 1H), 7.24 (d,  $J$  = 7.9 Hz, 1H), 2.54 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz, acetone- $d_6$ ) 167.2, 147.4, 134.6, 129.6, 128.9, 126.4, 21.19.

$^{11}\text{B NMR}$   $\delta$ /ppm: (128 MHz, acetone- $d_6$ ) 29.3.

IR (Microscope,  $\text{cm}^{-1}$ ): 3500-2500 (bp), 2965, 2927, 1698, 1607, 1406, 1362, 1296.

HRMS (ESI-TOF): for  $\text{C}_8\text{H}_8\text{BO}_4(\text{M} - \text{H})^-$ : *calcd.*: 179.0521; *found*: 179.0523; for  $\text{C}_8\text{H}_9\text{BClO}_4(\text{M} + \text{Cl})^-$ : *calcd.*: 215.0288; *found*: 215.0292.



### 4-Borono-3-methylbenzoic acid (2-16)

Pale-white solid (0.13 g, 30% yield).

Mp = 215 – 217 °C.

$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz, ACN- $d_3$ ) 7.75 – 7.74 (m, 1H), 7.73 – 7.70 (m, 1H), 7.54 (d,  $J$  = 7.6 Hz, 1H), 2.45 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz, ACN- $d_3$ ) 168.9, 142.6, 133.8, 131.5, 130.6, 126.17, 22.0.

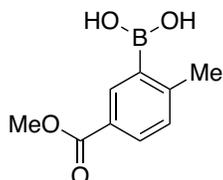
$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz, ACN- $d_3$ ) 29.5.

IR (Microscope,  $\text{cm}^{-1}$ ): 3500-2500 (bp), 3298, 2964, 1684, 1558, 1402, 1349, 1303, 1265.

HRMS (ESI-TOF): for  $\text{C}_8\text{H}_8\text{BO}_4(\text{M} - \text{H})^-$ : *calcd.*: 179.0521; *found*: 179.0522; for  $\text{C}_8\text{H}_9\text{BClO}_4(\text{M} + \text{Cl})^-$ : *calcd.*: 215.0288; *found*: 215.0290.

### **Methyl ester derivatives of compounds 2-9c/2-16 (2-3d/2-3e):**

To a solution of compound **2-9c/2-16** (81 mg, 0.45 mmol) in MeOH (2 ml), H<sub>2</sub>SO<sub>4</sub> (0.3 ml, 12 N) was added and the reaction mixture was stirred at reflux (65 °C) for 16 h. Then it was allowed to cool down to rt. The solvent was evaporated in vacuo. The crude residue was purified by flash chromatography (1:1, EtOAc/hexanes) to obtain methyl ester derivatives (**2-3d/2-3e**).



#### **(5-(Methoxycarbonyl)-2-methylphenyl)boronic acid (2-3d)**

White solid (86 mg, 98% yield)  $R_f = 0.47$  (1:1, EtOAc/hexanes).

Mp = 200 – 203 °C

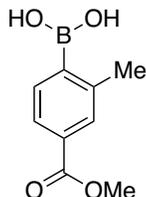
<sup>1</sup>H NMR δ/ppm: (500 MHz, ACN-*d*<sub>3</sub>) 8.12 (d,  $J = 2.0$  Hz, 1H), 7.84 (dd,  $J = 8.0, 2.0$  Hz, 1H), 7.25 (d,  $J = 8.0$  Hz, 1H), 3.82 (s, 3H), 2.48 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, ACN-*d*<sub>3</sub>) 168.0, 148.8, 135.1, 130.7, 130.4, 127.1, 52.25, 22.34.

<sup>11</sup>B NMR δ/ppm: (160 MHz, ACN-*d*<sub>3</sub>) 29.5.

IR (Microscope, cm<sup>-1</sup>): 3427, 2957, 1726, 1607, 1417, 1354, 1288, 1238, 1118.

HRMS (ESI-TOF): for C<sub>9</sub>H<sub>10</sub>BO<sub>4</sub> (M – H)<sup>-</sup>: *calcd.*: 193.0678; *found*: 193.0674.



#### **(4-(Methoxycarbonyl)-2-methylphenyl)boronic acid (2-3e)**

White solid (77 mg, 88% yield)  $R_f = 0.55$  (1:1, EtOAc/hexanes).

Mp = 165 – 167 °C.

<sup>1</sup>H NMR δ/ppm: (500 MHz, acetone-*d*<sub>6</sub>) 7.74 – 7.68 (comp m, 2H), 7.64 (d,  $J = 7.6$  Hz, 1H), 3.85 (s, 3H), 2.50 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, acetone-*d*<sub>6</sub>) 166.1, 141.5, 132.8, 129.5, 128.9, 124.4, 50.6, 20.7.

<sup>11</sup>B NMR δ/ppm: (128 MHz, acetone-*d*<sub>6</sub>) 29.3.

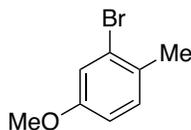
**IR** (Microscope,  $\text{cm}^{-1}$ ): 3428, 3308, 3020, 2958, 1706, 1560, 1435, 1415, 1351, 1305, 1260.

**HRMS** (ESI-TOF): for  $\text{C}_9\text{H}_{10}\text{BO}_4 (\text{M} - \text{H})^-$ : *calcd.*: 193.0678; *found*: 193.0677.

**Synthesis of 4/5-methoxy derivative of 2-fluoro/methyl phenylboronic acid (2-3f, 2-3i, 2-3j):**

**General synthesis of compounds 2-17 – 2-19:**

DMSO (5 ml) and potassium hydroxide (4.0 equiv, 0.54 g, 9.6 mmol) were mixed and stirred under nitrogen for 5 min at room temperature. Then phenol derivative (1.0 equiv, 2.4 mmol) was added followed by methyl iodide (2.0 equiv, 0.30 ml, 4.8 mmol). The reaction mixture was stirred for 4 h at room temperature. Then water (50 ml) was added and the product was extracted with  $\text{Et}_2\text{O}$  ( $3 \times 25$  ml). The combined organic layers were washed with water ( $2 \times 10$  ml), brine ( $1 \times 10$  ml), dried ( $\text{Na}_2\text{SO}_4$ ), filtered and concentrated in vacuo. The products were obtained pure as brown-yellowish liquid without further purification. The products are quite volatile so keeping them under vacuum would decrease the yield dramatically.



**2-Bromo-4-methoxy-1-methylbenzene (2-17)**

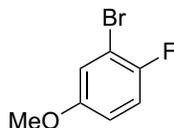
Red-brown liquid (0.29 g, 60% yield).

**$^1\text{H}$  NMR**  $\delta$ /ppm: (700 MHz,  $\text{CDCl}_3$ ) 7.12 (d,  $J = 8.4$ , 1H), 7.10 (d,  $J = 2.6$  Hz, 1H), 6.77 (dd,  $J = 8.4, 2.6$  Hz, 1H), 3.77 (s, 3H), 2.33 (s, 3H).

**$^{13}\text{C}$  NMR**  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 158.2, 131.0, 129.7, 124.9, 117.6, 113.4, 55.6, 21.8.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3067, 3002, 2925, 2835, 1606, 1569, 1494, 1290, 1240.

**HRMS** (EI): for  $\text{C}_8\text{H}_9\text{BrO} (\text{M})^+$ : *calcd.*: 201.98163 and 199.98367; *found*: 201.98136 and 199.98351.



**2-Bromo-1-fluoro-4-methoxybenzene (2-18)**

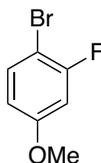
Yellow liquid (0.20 g, 40% yield).

**<sup>1</sup>H NMR** δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.07 (dd, *J* = 5.5, 3.0 Hz, 1H), 7.03 (dd, *J* = 9.1, 8.1 Hz, 1H), 6.80 (ddd, *J* = 9.1, 3.8, 3.0 Hz, 1H), 3.78 (s, 3H).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CDCl<sub>3</sub>) 156.1 (d, *J*<sub>CF</sub> = 2.4 Hz), 153.8 (d, *J*<sub>CF</sub> = 239.5 Hz), 118.2, 116.6 (d, *J*<sub>CF</sub> = 23.9 Hz), 114.4 (d, *J*<sub>CF</sub> = 6.8 Hz), 109.0 (d, *J*<sub>CF</sub> = 22.4 Hz), 56.01.

**IR** (Microscope, cm<sup>-1</sup>): 3105, 3007, 2943, 2838, 1600, 1588, 1495, 1259, 1217, 1207, 1044.

**HRMS** (EI): for C<sub>7</sub>H<sub>6</sub>BrFO (M)<sup>+</sup>: *calcd.*: 205.95656 and 203.95860; *found*: 205.95671 and 203.95863.



### 1-Bromo-2-fluoro-4-methoxybenzene (2-19)

Yellow liquid (0.15 g, 30% yield).

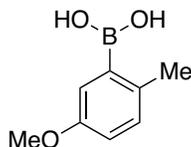
**<sup>1</sup>H NMR** δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.40 (dd, *J* = 8.9, 8.0 Hz, 1H), 6.69 (dd, *J* = 10.4, 2.8 Hz, 1H), 6.61 (ddd, *J* = 8.8, 2.8, 1.1 Hz, 1H), 3.79 (s, 3H).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CDCl<sub>3</sub>) 160.4 (d, *J*<sub>CF</sub> = 28.4 Hz), 159.4 (d, *J*<sub>CF</sub> = 208.0 Hz), 133.4 (d, *J*<sub>CF</sub> = 2.1 Hz), 111.4 (d, *J*<sub>CF</sub> = 3.1 Hz), 102.9 (d, *J*<sub>CF</sub> = 25.6 Hz), 99.3 (d, *J*<sub>CF</sub> = 21.2 Hz), 55.8.

**IR** (Microscope, cm<sup>-1</sup>): 3008, 2964, 2941, 2839, 1606, 1587, 1490, 1322, 1292, 1198, 1161.

**HRMS** (EI): for C<sub>7</sub>H<sub>6</sub>BrFO (M)<sup>+</sup>: *calcd.*: 205.95656 and 203.95860; *found*: 205.95671 and 203.95863.

### Synthesis of boronic acids 2-3f, 2-3i and 2-3j:



**(5-Methoxy-2-methylphenyl)boronic acid (2-3f)**: Compound **2-17** (1.0 equiv, 0.24 g, 1.2 mmol) in dry THF (4 ml) was added to the flame-dried flask under nitrogen. A solution of *n*-BuLi (2.5 M, 0.60 ml, 1.5 mmol) was added to the reaction mixture at -78 °C dropwise *via* a syringe over 5 min. After the mixture was stirred for 30 min at -78 °C, trimethyl borate (1.6 equiv, 0.22 ml, 2.0 mmol) was added into the solution. The reaction mixture was brought to rt

over 2 h, then quenched by HCl (1 ml, 1 N) and concentrated in vacuo. Then, HCl (30 ml, 1 N) was added and aqueous part was extracted with EtOAc (3 × 20 ml). The combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (4:6, EtOAc/hexanes) to obtain 2-methyl-5-methoxyphenylboronic acid (**2-3f**) as a pale-white solid (0.14 g, 70% yield).

Pale-white solid, *R<sub>f</sub>* = 0.35 (4:6, EtOAc/hexanes).

Mp = 136 – 138 °C.

<sup>1</sup>H NMR δ/ppm: (700 MHz, ACN-*d*<sub>3</sub>) 7.04 (d, *J* = 8.3 Hz, 1H), 7.02 (d, *J* = 2.9 Hz, 1H), 6.79 (dd, *J* = 8.3, 2.9 Hz, 1H), 3.72 (s, 3H), 2.33 (s, 3H).

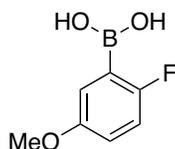
<sup>13</sup>C NMR δ/ppm: (126 MHz, ACN-*d*<sub>3</sub>) 157.3, 134.4, 131.1, 119.0, 115.3, 55.3, 21.0.

<sup>11</sup>B NMR δ/ppm: (128 MHz, ACN-*d*<sub>3</sub>) 29.6.

IR (Microscope, cm<sup>-1</sup>): 3283, 3142, 2984, 2916, 2832, 2788, 1588, 1575, 1454, 1329, 1161, 1145.

HRMS (ESI-TOF): for C<sub>8</sub>H<sub>10</sub>BO<sub>3</sub> (M – H)<sup>-</sup>: *calcd.*: 165.0728; *found*: 165.0726; for C<sub>8</sub>H<sub>11</sub>BClO<sub>3</sub> (M + Cl)<sup>-</sup>: *calcd.*: 201.0495; *found*: 201.0493.

Compounds **2-18** or **2-19** (1.0 equiv, 94 mg, 0.46 mmol) in dry THF (4 ml) was added to the flame-dried flask under nitrogen. A solution of *n*-BuLi (2.5 M, 0.23 ml, 0.55 mmol) was added to the reaction mixture at –100 °C dropwise *via* a syringe over 5 min. After the mixture was stirred for 30 min at –100 °C, trimethyl borate (1.6 equiv, 82 μl, 0.74 mmol) was added into the solution. The reaction mixture was brought to rt over 2 h. Then the reaction was quenched by HCl (1 ml, 1 N) and was concentrated in vacuo. Then, HCl (15 ml, 1 N) was added and aqueous part was extracted with EtOAc (3 × 10 ml). The combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography to obtain pale-white solid boronic acids (**2-3i**, **2-3j**).



**(2-Fluoro-5-methoxyphenyl)boronic acid (2-3i)**

Pale-white solid (56 mg, 72% yield), *R<sub>f</sub>* = 0.59 (1:1, EtOAc/hexanes).

Mp = 175 – 178 °C.

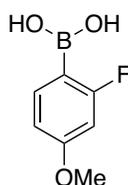
<sup>1</sup>H NMR δ/ppm: (500 MHz, ACN-*d*<sub>3</sub>) 7.13 (dd, *J* = 4.8, 2.9 Hz, 1H), 7.00 – 6.91 (comp m, 2H), 3.74 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, ACN-*d*<sub>3</sub>) 161.7 (d, *J*<sub>CF</sub> = 236.1 Hz), 156.1, 120.2 (d, *J*<sub>CF</sub> = 8.7 Hz), 118.5 (d, *J*<sub>CF</sub> = 5.3 Hz), 116.3 (d, *J*<sub>CF</sub> = 27.4 Hz), 56.0.

<sup>11</sup>B NMR δ/ppm: (160 MHz, ACN-*d*<sub>3</sub>) 28.0.

IR (Microscope, cm<sup>-1</sup>): 3240, 2964, 2938, 1612, 1492, 1420, 1349, 1153, 1030.

HRMS (EI): for C<sub>7</sub>H<sub>8</sub>BFO<sub>3</sub> (M)<sup>+</sup>: *calcd.*: 170.05505; *found*: 170.05509.



### **(2-Fluoro-4-methoxyphenyl)boronic acid (2-3j)**

Pale-white solid (33 mg, 43% yield): *R*<sub>f</sub> = 0.46 (3.5:6.5, acetone/hexanes).

Mp = 136 – 140 °C.

<sup>1</sup>H NMR δ/ppm: (700 MHz, ACN-*d*<sub>3</sub>) 7.59 (app t, *J* = 8.1 Hz, 1H), 6.73 (dd, *J* = 8.4, 2.3 Hz, 1H), 6.62 (dd, *J* = 12.2, 2.3 Hz, 1H), 3.78 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, ACN-*d*<sub>3</sub>) 168.9 (d, *J*<sub>CF</sub> = 244.1 Hz), 164.0 (d, *J*<sub>CF</sub> = 12.3 Hz), 137.8 (d, *J*<sub>CF</sub> = 10.8 Hz), 110.8 (d, *J*<sub>CF</sub> = 2.8 Hz), 101.5 (d, *J*<sub>CF</sub> = 29.2 Hz), 56.0.

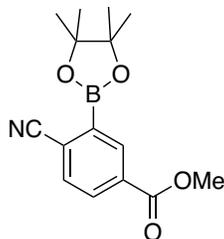
<sup>11</sup>B NMR δ/ppm: (128 MHz, acetone-*d*<sub>6</sub>) 27.9.

IR (Microscope, cm<sup>-1</sup>): 3530, 3363, 2978, 2846, 1620, 1566, 1413, 1389, 1326, 1310, 1272, 1153, 1134.

HRMS (EI): for C<sub>7</sub>H<sub>8</sub>BFO<sub>3</sub> (M)<sup>+</sup>: *calcd.*: 170.05505; *found*: 170.05546.

### **Synthesis of 2-cyano-5-methyl ester phenylboronic acid pinacol ester (2-3k pinacol boronate) and its boronic acid (2-3k):**

The compound **2-3k** pinacol boronate was synthesized based on the procedure described earlier by Smith and coworkers and its NMR data was confirmed with the literature values.<sup>26</sup>



**Methyl 4-cyano-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzoate (2-3k pinacol boronate)**

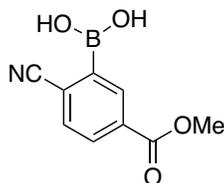
White solid (0.24 g, 33% yield).

Mp = 96.6 – 99.0 °C.

<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 8.52 (s, 1H), 8.31 – 8.07 (m, 1H), 7.77 (d, *J* = 8.1 Hz, 1H), 3.96 (s, 3H), 1.40 (s, 12H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 165.7, 138.2, 136.7, 133.5, 132.8, 131.9, 121.2, 118.2, 85.2, 52.7, 24.9.

HRMS (ESI-TOF): for C<sub>15</sub>H<sub>22</sub>BN<sub>2</sub>O<sub>4</sub> (M + NH<sub>4</sub>)<sup>+</sup>: *calcd.*: 305.1670; *found*: 305.1672.



**(2-Cyano-5-(methoxycarbonyl)phenyl)boronic acid (2-3k)**

Compound **2-3k** pinacol boronate (1.0 equiv, 0.24 g, 0.84 mmol) was dissolved in THF:water (8:1 ml) (4:1 v/v). Then sodium periodate (3.0 equiv, 0.54 g, 2.5 mmol) was added to the solution and stirred at room temperature for 30 min under ambient atmosphere. Lastly, HCl (0.65 ml, 1 N) was added to the reaction mixture, which was stirred for 24 h at room temperature. Sorbitol work-up<sup>32</sup> was performed to obtain pure (**2-3k**) (0.11 g, 64%).

**Sorbitol work-up:** D-sorbitol (1 M) was dissolved in aqueous solution of sodium carbonate (1 M). Meantime, reaction mixture was concentrated in vacuo. The crude residue was dissolved in EtOAc (50 ml) and washed with water (3 × 10 ml). Then, the organic part was washed with sorbitol solution (50 ml). This step was essential to allow boronic acid to transfer in basic sorbitol solution. Then, sorbitol part was washed with fresh EtOAc (3 × 40 ml) to remove organic impurities. Sorbitol part was separated and its pH was adjusted to pH 1 with

concentrated HCl solution under ice-water bath. Then sorbitol part was extracted with EtOAc (3 × 75 ml) and organic fractions were combined, dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo.

Pinkish-white powder.

<sup>1</sup>H NMR δ/ppm: (498 MHz, acetone-*d*<sub>6</sub>) 8.47 (d, *J* = 1.5 Hz, 1H), 8.12 (dd, *J* = 8.0, 1.9 Hz, 1H), 7.87 (dd, *J* = 8.1, 0.6 Hz, 1H), 3.91 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, acetone-*d*<sub>6</sub>) 163.6, 133.1, 131.4, 130.2, 128.3, 117.8, 116.5, 49.9.

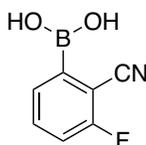
<sup>11</sup>B NMR δ/ppm: (128 MHz, acetone-*d*<sub>6</sub>) 27.5.

IR (Microscope, cm<sup>-1</sup>): 3350, 3113, 3052, 2958, 2485, 2237, 1726, 1606, 1278, 1258.

HRMS (ESI-TOF): for C<sub>9</sub>H<sub>7</sub>BNO<sub>4</sub> (M – H)<sup>-</sup>: *calcd.*: 204.0474; *found*: 204.0473; for C<sub>9</sub>H<sub>8</sub>BCINO<sub>4</sub> (M + Cl)<sup>-</sup>: *calcd.*: 240.0240; *found*: 240.0244; for C<sub>11</sub>H<sub>8</sub>BF<sub>3</sub>NO<sub>6</sub> (M + CF<sub>3</sub>COO)<sup>-</sup>: *calcd.*: 318.0402; *found*: 318.0397.

### **General synthesis of 2-cyanophenylboronic acid derivatives (2-20 – 2-21)**

2-Bromo-6-fluorobenzonitrile or 2-bromo-5-nitrobenzonitrile (1.0 equiv, 2.0 mmol), bis(pinacolato)diboron (1.2 equiv, 0.61 g, 2.4 mmol) and potassium acetate (3.0 equiv, 0.59 g, 6.0 mmol) were added to the flame dried round bottom flask and degassed followed by nitrogen purge. Then, dioxane (10 ml) was added under nitrogen. Lastly, [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium(II) complex with dichloromethane (Pd(dppf)Cl<sub>2</sub>) (5 mol%, 82 mg, 0.10 mmol) was added and the reaction mixture was stirred at 80 °C for 16 h. After the reaction mixture reached to rt, it was filtered over celite, washed with EtOAc (50 ml) and concentrated in vacuo. The residue was partially purified with a flash chromatography (4:6, EtOAc/hexanes) to obtain still impure 2-cyano-3-fluorophenylboronic acid pinacol ester or 2-cyano-4-nitrophenylboronic acid pinacol ester, which was Then used for the deprotection of pinacol. Pinacol boronates (1.0 equiv, 2.0 mmol) were dissolved in THF:water (16:4 ml) (4:1 v/v). Then sodium periodate (3.0 equiv, 1.3 g, 6.0 mmol) was added to the solution and stirred at room temperature for 30 min under ambient atmosphere. Lastly, HCl (1.5 ml, 1 N) was added to the reaction mixture and stirred for 48 h. Sorbitol work-up<sup>32</sup> was performed in order to obtain pure 2-cyano-3-fluorophenylboronic acid (**2-20**) as pale-white powder (0.28 g, 83%) and 2-cyano-4-nitrophenylboronic acid (**2-21**) as yellow powder (0.18 g, 47%).



**(2-Cyano-3-fluorophenyl)boronic acid (2-20)**

Pale-white solid (0.28 g, 83%).

Mp = 177 – 180 °C.

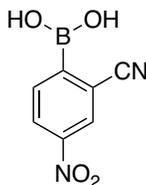
$^1\text{H NMR}$   $\delta$ /ppm: (498 MHz, acetone- $d_6$ ) 7.73 – 7.65 (comp m, 2H), 7.41 – 7.35 (m, 1H);  $^{13}\text{C}$

$\text{NMR}$   $\delta$ /ppm: (126 MHz, acetone- $d_6$ ) 164.7 (d,  $J_{\text{CF}} = 255.3$  Hz), 135.2 (d,  $J_{\text{CF}} = 8.3$  Hz), 131.6 (d,  $J_{\text{CF}} = 3.4$  Hz), 118.0 (d,  $J_{\text{CF}} = 20.3$  Hz), 115.0, 105.2 (d,  $J_{\text{CF}} = 12.9$  Hz).

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz, acetone- $d_6$ ) 27.4.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3212 (bs), 2260, 2240, 1602, 1460, 1193.

**HRMS** (ESI-TOF): for  $\text{C}_7\text{H}_6\text{BFNO}_2$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 166.047; *found*: 166.047.



**(2-Cyano-4-nitrophenyl)boronic acid (2-21)**

Yellow solid (0.18 g, 47%).

$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz, acetone- $d_6$ ) 8.51 (d,  $J = 2.2$  Hz, 1H), 8.42 (dd,  $J = 8.3, 2.3$  Hz, 1H), 8.10 (d,  $J = 8.3$  Hz, 1H).

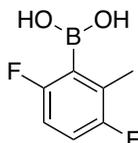
$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz, acetone- $d_6$ ) 149.2, 136.9, 128.3, 126.7, 118.4, 117.9.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz, acetone- $d_6$ ) 27.4.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3113, 2252, 1604, 1522, 1427, 1295, 1260.

**HRMS** (ESI-TOF): for  $\text{C}_7\text{H}_4\text{BN}_2\text{O}_4$  ( $\text{M} - \text{H}$ ) $^-$ : *calcd.*: 191.027; *found*: 191.0275; for  $\text{C}_7\text{H}_5\text{BClN}_2\text{O}_4$  ( $\text{M} + \text{Cl}$ ) $^-$ : *calcd.*: 227.0029; *found*: 227.0029.

**Synthesis of 2,5-difluoro-6-methylphenylboronic acid (2-22):**



**(3,6-Difluoro-2-methylphenyl)boronic acid (2-22):** 3,6-Difluoro-2-methylphenylboronic acid pinacol ester (1.0 equiv, 52 mg, 0.20 mmol) was dissolved in THF:water (4:1 ml) (4:1 v/v). Then sodium periodate (3.0 equiv, 128 mg, 0.60 mmol) was added to the solution and stirred at room temperature for 30 min under ambient atmosphere. Lastly, HCl (0.2 ml, 1 N) was added to the reaction mixture, which was stirred for 24 h at room temperature. The reaction mixture was concentrated in vacuo. Then it was dissolved in EtOAc (30 ml) and washed with water (1 × 8 ml), brine (1 × 8 ml). The organic fraction was dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo to obtain pure 3,6-difluoro-2-methylphenylboronic acid (**2-22**).

White powder (30 mg, 86%).

Mp = 98 – 100 °C.

<sup>1</sup>H NMR δ/ppm: (498 MHz, ACN-*d*<sub>3</sub>) 7.04 – 6.97 (m, 1H), 6.88 – 6.82 (m, 1H), 2.24 (d, *J* = 2.6 Hz, 3H).

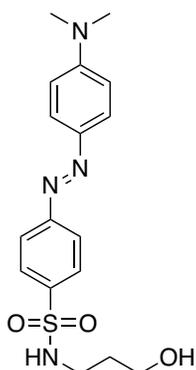
<sup>13</sup>C NMR δ/ppm: (126 MHz, ACN-*d*<sub>3</sub>) 161.0 (dd, *J*<sub>CF</sub> = 233.8, 1.9 Hz), 158.1 (d, *J*<sub>CF</sub> = 237.5 Hz), 129.0 (dd, *J*<sub>CF</sub> = 17.9, 10.7 Hz), 117.1 (dd, *J*<sub>CF</sub> = 26.2, 9.7 Hz), 113.9 (dd, *J*<sub>CF</sub> = 28.1, 8.4 Hz), 14.5 (d, *J*<sub>CF</sub> = 4.2 Hz).

<sup>11</sup>B NMR δ/ppm: (160 MHz, ACN-*d*<sub>3</sub>) 28.8.

HRMS (EI) for C<sub>7</sub>H<sub>7</sub>BF<sub>2</sub>O<sub>2</sub> (M)<sup>+</sup>: *calcd.*: 172.05072; *found*: 172.05070.

### Synthesis of dabcyI derivatized neopentyl boronates (2-9a – 2-9b) and their nopol-dansyl-diols boronates (2-10a – 2-10b)

#### Synthesis of dabcyI alcohol (2-9e)



**(E)-4-((4-(Dimethylamino)phenyl)diazenyl)-N-(3-hydroxypropyl)benzenesulfonamide (2-9e):** Excess amount of phosphoryl chloride (3 ml, 32 mmol) was added to methyl orange compound (0.33 g, 1 mmol) placed in a flamed dried flask at room temperature under nitrogen. Then nitrogen was removed and the reaction mixture was heated and stirred at 110 °C overnight. After toluene (2 ml) was added to the reaction mixture, the purple solution was concentrated in vacuo. This step was repeated three times until the crude dabcyyl chloride was dry. Then it was kept under vacuum for 5 h. The crude residue was immediately used to synthesize dabcyyl alcohol (**2-9e**). To the solution of crude dabcyyl chloride in DCM (7 ml), the mixture of pyridine (3 ml), 3-aminopropyl alcohol (0.43 ml, 5.6 mmol) and isopropanol (1 ml) was added dropwise under nitrogen at 0 °C over 30 min, after which time the ice-water bath was removed and the reaction mixture was stirred for 48 h. The reaction mixture was directly purified by flash chromatography (dry loading) (6:3:1, DCM/hexanes/IPA) to obtain dabcyyl alcohol (**2-9e**) (0.21 g, 58% yield).

Orange solid,  $R_f = 0.74$  (6:3:1, DCM/hexanes/IPA).

Mp = 134 – 136 °C.

**<sup>1</sup>H NMR**  $\delta$ /ppm: (400 MHz, CDCl<sub>3</sub>) 7.99 – 7.88 (comp m, 6H), 6.79 – 6.74 (m, 2H), 4.98 (t,  $J = 6.1$  Hz, 1H), 3.75 (t,  $J = 5.6$  Hz, 2H), 3.17 (app q,  $J = 6.1$  Hz, 2H), 3.12 (s, 6H), 1.76 – 1.68 (m, 2H).

**<sup>13</sup>C NMR**  $\delta$ /ppm: (126 MHz, CDCl<sub>3</sub>) 155.7, 153.2, 143.7, 139.4, 128.1, 125.8, 122.8, 111.5, 61.0, 41.4, 40.3, 31.4.

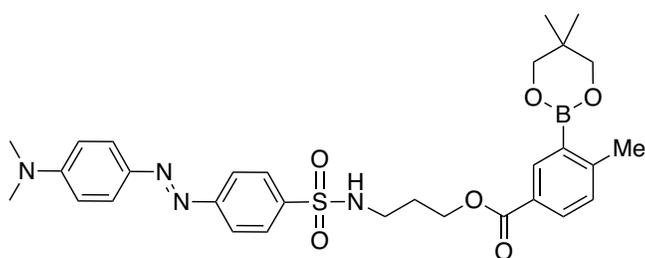
**IR** (Microscope, cm<sup>-1</sup>): 3481, 3174, 2943, 2891, 1607, 1520, 1390, 1370, 1313, 1303, 1158, 1139.

**HRMS** (ESI-TOF): for C<sub>17</sub>H<sub>23</sub>N<sub>4</sub>O<sub>3</sub>S (M + H)<sup>+</sup>: *calcd.*: 363.1485; *found*: 363.1485; for C<sub>17</sub>H<sub>22</sub>N<sub>4</sub>NaO<sub>3</sub>S (M + Na)<sup>+</sup>: *calcd.*: 385.1305; *found*: 385.1301.

### Synthesis of dabcyyl derivatized neopentyl boronates (**2-9a – 2-9b**)

To 2-methylphenylboronic acid **2-9c** or commercially available 2-fluorophenylboronic acid **2-9d** (1.00 equiv, 0.24 mmol) in THF (1 ml), neopentyl glycol (0.95 equiv, 24 mg, 0.23 mmol) was added and stirred at room temperature for 1 h under ambient atmosphere. Toluene (0.1 ml) was added to the reaction mixture that was then concentrated in vacuo, and dried under vacuum for 2 h. To the crude residue in dry DCM (3 ml), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) (1.2 equiv, 56 mg, 0.29 mmol) and DMAP (0.5 equiv, 15 mg, 0.12 mmol) were added

under nitrogen balloon at 0 °C and stirred for 10 min after which time compound **2-9e** (0.66 equiv, 58 mg, 0.16 mmol) was added to the reaction mixture. Orange color reaction mixture was stirred at 0 °C for 10 min. Then, ice-water bath was removed and the reaction mixture was stirred for 24 h at room temperature under nitrogen balloon. Then, the reaction mixture was mixed with DCM (20 ml). The organic phase was washed with distilled water (3 × 5.0 ml), saturated sodium bicarbonate (1 × 5.0 ml) and brine (1 × 5.0 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:1, EtOAc/hexanes) to obtain dabacyl derivatized neopentyl boronates (**2-9a – 2-9b**).



**(E)-3-(((4-((4-(Dimethylamino)phenyl)diazenyl)phenyl)sulfonamido)propyl 3-(5,5-dimethyl-1,3,2-dioxaborinan-2-yl)-4-methylbenzoate (2-9a)**

Orange-red amorphous solid (56 mg, 59% yield),  $R_f = 0.25$  (1:9, EtOAc/DCM).

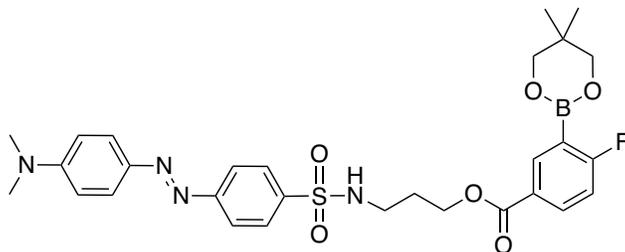
<sup>1</sup>H NMR δ/ppm: (498 MHz, CDCl<sub>3</sub>) 8.35 (s, 1H), 7.98 – 7.86 (comp m, 6H), 7.84 (dd,  $J = 7.9$ , 2.0, 1H), 7.17 (d,  $J = 8.0$  Hz, 1H), 6.77 – 6.73 (m, 2H), 5.02 (t,  $J = 6.5$  Hz, 1H), 4.35 (t,  $J = 5.9$  Hz, 2H), 3.78 (s, 4H), 3.15 – 3.18 (comp m, 8H), 2.54 (s, 3H), 1.93 (app p,  $J = 6.2$  Hz, 2H), 1.03 (s, 6H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 167.2, 155.6, 153.1, 150.1, 143.6, 139.5, 136.3, 131.1, 130.2, 128.0, 126.2, 125.7, 122.7, 111.5, 72.4, 61.5, 40.3, 40.2, 31.7, 29.3, 22.7, 21.9.

<sup>11</sup>B NMR δ/ppm: (160 MHz, CDCl<sub>3</sub>) 27.0.

IR (Microscope, cm<sup>-1</sup>): 3278, 2962, 1714, 1604, 1521, 1421, 1365, 1256, 1136.

HRMS (ESI-TOF): for C<sub>25</sub>H<sub>28</sub>BN<sub>4</sub>O<sub>6</sub>S (M – C<sub>5</sub>H<sub>8</sub> – H)<sup>-</sup>: *calcd.*: 523.2; *found*: 523.2; for C<sub>25</sub>H<sub>29</sub>BClN<sub>4</sub>O<sub>6</sub>S (M – C<sub>5</sub>H<sub>8</sub> + Cl)<sup>-</sup>: *calcd.*: 559.2; *found*: 559.2.



**(E)-3-((4-((4-(Dimethylamino)phenyl)diazenyl)phenyl)sulfonamido)propyl 3-(5,5-dimethyl-1,3,2-dioxaborinan-2-yl)-4-fluorobenzoate (2-9b)**

Orange-red amorphous solid (70 mg, 74% yield),  $R_f = 0.07$  (1:9, EtOAc/DCM).

$^1\text{H NMR}$   $\delta$ /ppm: (400 MHz,  $\text{CDCl}_3$ ) 8.39 – 8.34 (m, 1H), 8.06 – 7.81 (comp m, 7H), 7.05 – 6.97 (m, 1H), 6.78 – 6.70 (m, 2H), 5.10 (t,  $J = 5.9$  Hz, 1H), 4.35 (t,  $J = 5.3$  Hz, 2H), 3.80 (s, 4H), 3.17 – 3.06 (comp m, 8H), 1.98 – 1.89 (m, 2H), 1.02 (s, 6H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 170.1 (d,  $J_{\text{CF}} = 258.5$  Hz), 166.0, 155.6, 153.1, 143.6, 139.3, 138.5 (d,  $J_{\text{CF}} = 9.8$  Hz), 134.3 (d,  $J_{\text{CF}} = 10.5$  Hz), 128.0, 125.7, 125.4 (d,  $J_{\text{CF}} = 3.0$  Hz), 122.6, 115.8 (d,  $J_{\text{CF}} = 25.7$  Hz), 111.4, 72.5, 61.8, 40.3, 40.0, 31.8, 29.2, 21.9.

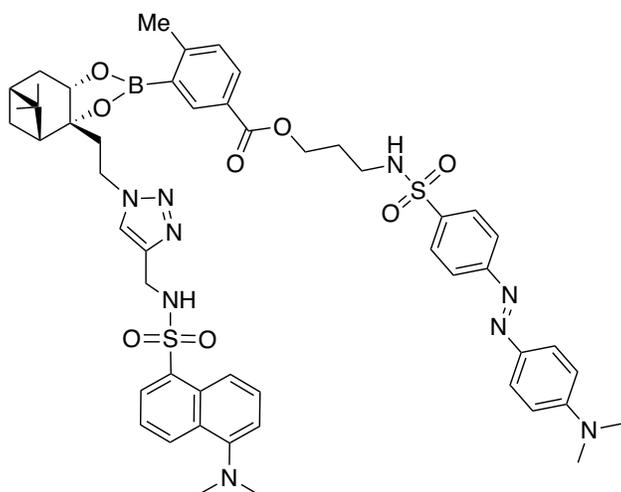
$^{11}\text{B NMR}$   $\delta$ /ppm: (128 MHz,  $\text{CDCl}_3$ ) 26.0.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3283, 2960, 2929, 1719, 1604, 1521, 1488, 1366, 1264, 1137.

**HRMS** (ESI-TOF): for  $\text{C}_{24}\text{H}_{25}\text{BFN}_4\text{O}_6\text{S}$  ( $\text{M} - \text{C}_5\text{H}_8 - \text{H}$ ) $^-$ : *calcd.*: 527.2; *found*: 527.2; for  $\text{C}_{24}\text{H}_{26}\text{BCIFN}_4\text{O}_6\text{S}$  ( $\text{M} - \text{C}_5\text{H}_8 + \text{Cl}$ ) $^-$ : *calcd.*: 563.1; *found*: 563.1.

**Synthesis of their nopol-dansyl-diols boronates (2-10a – 2-10b)**

Compound **2-9a/2-9b** (1 equiv, 0.033 mmol) was mixed with diol **2-1** (1.2 equiv, 0.039 mmol) in THF:water (1 ml/0.1 ml). The reaction mixture was stirred for 3 h at room temperature under ambient atmosphere. Then, toluene (0.2 ml) was added to the mixture and it was concentrated in vacuo. The crude residue was purified by flash chromatography.



**3-((4-((*E*)-(4-(Dimethylamino)phenyl)diazenyl)phenyl)sulfonamido)propyl 3-((3*aR*,4*R*,6*R*,7*aS*)-3*a*-(2-(4-(((5-(dimethylamino)naphthalene)-1-sulfonamido)methyl)-1*H*-1,2,3-triazol-1-yl)ethyl)-5,5-dimethylhexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborol-2-yl)-4-methylbenzoate (2-10a)**

Orange-red amorphous solid (26 mg, 79% yield),  $R_f = 0.27$  (3:7, EtOAc/DCM).

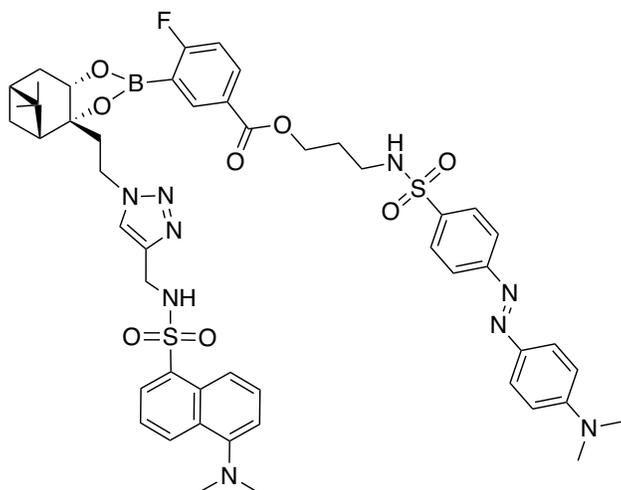
$^1\text{H NMR}$   $\delta$ /ppm: (498 MHz,  $\text{CDCl}_3$ ) 8.50 (d,  $J = 8.6$  Hz, 1H), 8.33 (d,  $J = 2.0$  Hz, 1H), 8.25 – 8.20 (m, 2H), 7.95 – 7.83 (comp m, 7H), 7.54 – 7.46 (comp m, 2H), 7.32 (s, 1H), 7.24 (d,  $J = 8.1$  Hz, 1H), 7.14 (d,  $J = 7.5$  Hz, 1H), 6.76 – 6.72 (m, 2H), 5.61 – 5.51 (comp m, 2H), 4.50 – 4.29 (comp m, 5H), 4.15 (d,  $J = 6.2$  Hz, 2H), 3.19 – 3.07 (comp m, 8H), 2.85 (s, 6H), 2.56 (s, 3H), 2.48 – 2.30 (m, 2H), 2.30 – 2.14 (comp m, 3H), 2.08 – 1.87 (comp m, 4H), 1.28 (s, 3H), 1.20 (d,  $J = 11.0$  Hz, 1H), 0.88 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 166.7, 155.6, 153.1, 152.0, 150.8, 143.95, 143.60, 143.6, 139.5, 137.3, 134.6, 132.2, 130.6, 130.3, 129.8, 129.6, 128.6, 128.0, 126.6, 125.7, 123.1, 122.6, 122.5, 118.7, 115.3, 111.5, 77.2, 86.6, 61.9, 49.0, 45.8, 45.4, 40.5, 40.3, 40.2, 39.6, 38.8, 38.1, 35.6, 29.2, 26.9, 26.1, 24.0, 22.7.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) 0.30.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3283, 2925, 2855, 1718, 1604, 1520, 1365, 1164, 1137.

**HRMS** (ESI-TOF): for  $\text{C}_{51}\text{H}_{61}\text{BN}_9\text{O}_8\text{S}_2$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 1002.4172; *found*: 1002.4194; for  $\text{C}_{51}\text{H}_{60}\text{BN}_9\text{NaO}_8\text{S}_2$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 1024.3992; *found*: 1024.4000.



**3-((4-((*E*)-(4-(Dimethylamino)phenyl)diazenyl)phenyl)sulfonamido)propyl 3-((3*aR*,4*R*,6*R*,7*aS*)-3*a*-(2-(4-(((5-(dimethylamino)naphthalene)-1-sulfonamido)methyl)-1*H*-1,2,3-triazol-1-yl)ethyl)-5,5-dimethylhexahydro-4,6-methanobenzo[*d*][1,3,2]dioxaborol-2-yl)-4-fluorobenzoate (2-10b)**

Orange-red amorphous solid (22 mg, 66% yield),  $R_f = 0.20$  (3:7, EtOAc/DCM).

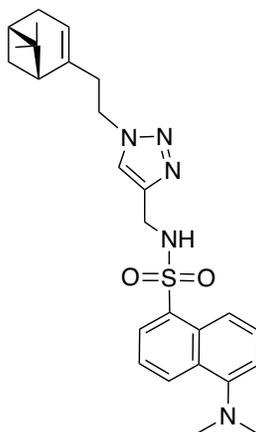
$^1\text{H NMR}$   $\delta$ /ppm: (498 MHz,  $\text{CDCl}_3$ ) 8.50 (d,  $J = 8.5$  Hz, 1H), 8.31 (dd,  $J = 5.8, 2.3$  Hz, 1H), 8.27 – 8.18 (m, 2H), 8.08 (ddd,  $J = 8.0, 5.2, 2.3$  Hz, 1H), 7.97 – 7.82 (comp m, 6H), 7.55 – 7.45 (comp m, 2H), 7.32 (s, 1H), 7.14 (d,  $J = 7.6$  Hz, 1H), 7.09 (app t,  $J = 8.6$  Hz, 1H), 6.78 – 6.70 (m, 2H), 5.57 (t,  $J = 6.3$  Hz, 1H), 5.52 (t,  $J = 6.3$  Hz, 1H), 4.52 – 4.31 (comp m, 5H), 4.17 (d,  $J = 6.2$  Hz, 2H), 3.18 – 3.07 (m, 9H), 2.85 (s, 6H), 2.48 – 2.31 (comp m, 2H), 2.31 – 2.14 (comp m, 3H), 2.11 – 1.89 (comp m, 4H), 1.28 (s, 3H), 1.19 (d,  $J = 11.1$  Hz, 1H), 0.87 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 170.1 (d,  $J_{\text{CF}} = 259.2$  Hz), 165.7, 155.6, 153.1, 152.0, 143.9, 143.6, 139.43, 139.41, 139.0 (d,  $J_{\text{CF}} = 9.4$  Hz), 135.5 (d,  $J_{\text{CF}} = 10.3$  Hz), 134.6, 130.6, 129.8, 129.6, 128.6, 128.0, 125.9 (d,  $J_{\text{CF}} = 2.7$  Hz), 125.7, 123.1, 122.62, 122.56, 118.7, 115.9 (d,  $J_{\text{CF}} = 24.9$  Hz), 111.5, 87.1, 77.6, 62.1, 48.8, 45.7, 45.4, 40.4, 40.3, 40.0, 39.5, 38.8, 38.2, 35.3, 29.2, 27.0, 26.1, 23.9.

$^{11}\text{B NMR}$   $\delta$ /ppm: (160 MHz,  $\text{CDCl}_3$ ) –0.32.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3287, 2958, 2925, 2854, 1728, 1604, 1465, 1366, 1274, 1164, 1138.

**HRMS** (ESI-TOF): for  $\text{C}_{50}\text{H}_{58}\text{BFN}_9\text{O}_8\text{S}_2$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 1006.3921; *found*: 1006.3931; for  $\text{C}_{50}\text{H}_{57}\text{BFN}_9\text{NaO}_8\text{S}_2$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 1028.3741; *found*: 1028.3749.



**5-(Dimethylamino)-N-((1-(2-((1R,5S)-6,6-dimethylbicyclo[3.1.1]hept-2-en-2-yl)ethyl)-1H-1,2,3-triazol-4-yl)methyl)naphthalene-1-sulfonamide (2-11):** To the solution of nopol-azide **1a** (1.0 equiv, 23 mg, 0.12 mmol) and compound **2-1c** (1.2 equiv, 40 mg, 0.14 mmol) in water:ACN (2:1 ml),  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$  (0.10 equiv, 3.0 mg, 0.012 mmol) was added at room temperature. Hydrazine monohydrate (0.10 equiv, 6.2  $\mu\text{l}$ , 0.012 mmol) was added dropwise to the reaction mixture under vigorous stirring. The reaction mixture was stirred for 16 h at room temperature and it turned to pale yellow or light green color. Then water (5 ml) was added to the mixture and the product was extracted with EtOAc (3  $\times$  20 ml). The combined organic layers were dried ( $\text{Na}_2\text{SO}_4$ ), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:1 EtOAc:hexanes) to obtain nopol-dansyl (**2-11**) (53 mg, 91% yield) Green powder,  $R_f = 0.26$  (1:1 EtOAc:hexanes).

Mp = 44.1 – 47.0  $^\circ\text{C}$ .

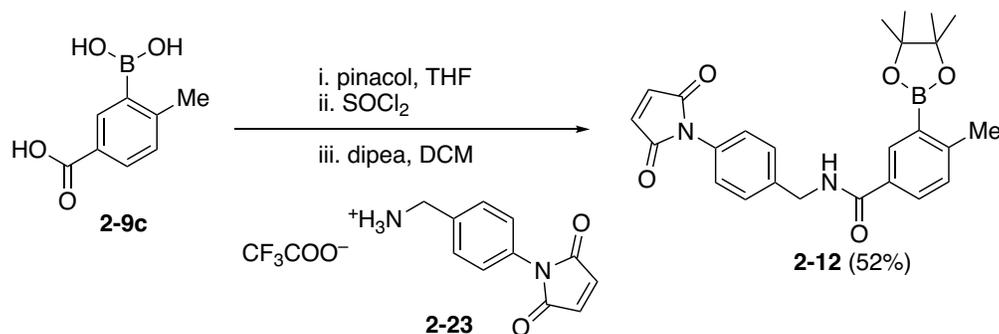
**$^1\text{H}$  NMR**  $\delta/\text{ppm}$ : (498 MHz,  $\text{CDCl}_3$ ) 8.56 – 8.53 (app dt,  $J = 8.5, 1.1$  Hz, 1H), 8.27 (dd,  $J = 7.3, 1.3$  Hz, 1H), 8.25 – 8.22 (app dt,  $J = 8.6, 1.1$  Hz, 1H), 7.58 – 7.50 (comp m, 2H), 7.21 – 7.17 (comp m, 2H), 5.26 – 5.23 (m, 1H), 5.15 (t,  $J = 6.2$  Hz, 1H), 4.24 – 4.17 (comp m, 4H), 2.89 (s, 6H), 2.46 – 2.34 (comp m, 3H), 2.26 – 2.13 (comp m, 2H), 2.11 – 2.05 (m, 1H), 2.01 (app td,  $J = 5.6, 1.5$  Hz, 1H), 1.28 (s, 3H), 1.08 (d,  $J = 8.6$  Hz, 1H), 0.78 (s, 3H).

**$^{13}\text{C}$  NMR**  $\delta/\text{ppm}$ : (126 MHz,  $\text{CDCl}_3$ ) 152.1, 143.5, 143.3, 134.6, 130.7, 130.0, 129.7, 129.6, 128.7, 123.2, 121.8, 119.8, 118.6, 115.4, 48.5, 45.6, 45.5, 40.6, 39.0, 38.1, 37.4, 31.7, 31.3, 26.2, 21.2.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3283, 3142, 2984, 2916, 2832, 2788, 1588, 1575, 1454, 1329, 1161, 1145.

**HRMS** (ESI-TOF): for  $C_{26}H_{34}N_5O_2S$  ( $M + H$ )<sup>+</sup>: *calcd.*: 480.2428; *found*: 480.2426; for  $C_{26}H_{33}N_5NaO_2S$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 502.2247; *found*: 502.2247.

### **Synthesis of maleimide derivatized boronate (2-12)**



***N*-(4-(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)benzyl)-4-methyl-3-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzamide (2-12)**: A mixture of compound 2-9c (1.0 equiv, 45 mg, 0.25 mmol), pinacol (1.1 equiv, 32 mg, 0.28 mmol) and THF (2 ml) was stirred at room temperature for 2 h under ambient atmosphere. Then, toluene (0.2 ml) was added to the mixture and it was concentrated in vacuo. After the product pinacol boronate was kept under vacuum for 2 h, it was used without further purification to obtain acyl chloride intermediate. Thionyl chloride (1 ml) was added to the crude residue at 0 °C under nitrogen balloon. Then, nitrogen balloon and the ice-water bath were removed and the reaction mixture was stirred for 1 h at room temperature. Excess thionyl chloride was removed under vacuo and the crude residue was kept under vacuum for 2 h. Then to the crude acyl chloride product in DCM (2.5 ml), maleimide derivatized benzyl amine compound 2-23 (1.0 equiv, 79 mg, 0.25 mmol) was added at 0 °C under nitrogen balloon. Lastly, DIPEA (1.0 equiv, 44  $\mu$ l, 0.25 mmol) was added dropwise at 0 °C and the reaction mixture was stirred for 10 min, after which time the ice-water bath was removed and the mixture was stirred for 7 h under nitrogen balloon. The reaction mixture was directly purified by flash chromatography (1:1, EtOAc:hexanes) to obtain maleimide derivatized pinacol boronate (2-12) (58 mg, 52% yield).

White solid,  $R_f$  = 0.51 (1:1, EtOAc:hexanes).

Mp = 184 – 186 °C.

**<sup>1</sup>H NMR** δ/ppm: (498 MHz, CDCl<sub>3</sub>) 8.06 (d, *J* = 2.2 Hz, 1H), 7.85 (dd, *J* = 7.9, 2.2 Hz, 1H), 7.50 – 7.45 (m, 2H), 7.36 – 7.31 (m, 2H), 6.85 (s, 2H), 6.48 (t, *J* = 5.8 Hz, 1H), 4.69 (d, *J* = 5.8 Hz, 2H), 2.58 (s, 3H), 1.35 (s, 12H).

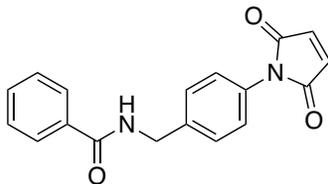
**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CDCl<sub>3</sub>) 169.5, 167.3, 149.2, 138.4, 134.3, 133.5, 130.6, 130.5, 130.4, 130.1, 128.9, 126.4, 83.9, 43.6, 25.0, 22.3.

**<sup>11</sup>B NMR** δ/ppm: (160 MHz, CDCl<sub>3</sub>) 30.9.

**IR** (Microscope, cm<sup>-1</sup>): 3335, 2979, 2931, 1717, 1644, 1517, 1401, 1144.

**HRMS** (ESI-TOF): for C<sub>25</sub>H<sub>28</sub>BN<sub>2</sub>O<sub>5</sub> (M + H)<sup>+</sup>: *calcd.*: 447.2086; *found*: 447.2093; for C<sub>25</sub>H<sub>27</sub>BN<sub>2</sub>NaO<sub>5</sub> (M + Na)<sup>+</sup>: *calcd.*: 469.1905; *found*: 469.1910.

### **Synthesis of benzoyl maleimide (2-13)**



***N*-(4-(2,5-Dioxo-2,5-dihydro-1*H*-pyrrol-1-yl)benzyl)benzamide (2-13)**: To benzoyl chloride (1.0 equiv, 15 mg, 0.1 mmol) in DCM (1.0 ml), compound **2-23** (1.0 equiv, 32 mg, 0.1 mmol) was added at 0 °C under nitrogen balloon. Lastly, DIPEA (1.0 equiv, 18 μl, 0.1 mmol) was added dropwise at 0 °C and the reaction mixture was stirred for 15 min, after which time the ice-water bath was removed and the mixture was stirred for 4 h under nitrogen balloon. The reaction mixture was directly purified by flash chromatography (1:1, EtOAc:hexanes) to obtain maleimide derivatized benzyl compound (**2-13**) (18 mg, 50% yield).

Pale-white solid, *R<sub>f</sub>* = 0.38 (1:1, EtOAc:hexanes).

*Mp* = 164 – 166 °C.

**<sup>1</sup>H NMR** δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.81 – 7.78 (m, 2H), 7.55 – 7.41 (comp m, 5H), 7.37 – 7.31 (m, 2H), 6.86 (s, 2H), 6.40 (bs, 1H), 4.69 (d, *J* = 5.7 Hz, 2H).

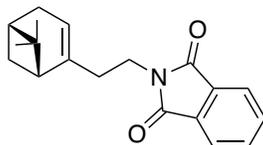
**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CDCl<sub>3</sub>) 169.4, 167.4, 138.2, 134.3, 134.3, 131.7, 130.7, 128.9, 128.7, 127.0, 126.5, 43.7.

**IR** (Microscope, cm<sup>-1</sup>) 3464, 3286, 3083, 2906, 1703, 1636, 1535.

**HRMS** (ESI-TOF) for  $C_{18}H_{15}N_2O_3$  ( $M + H$ )<sup>+</sup>: *calcd.*: 307.1077; *found*: 307.1082; for  $C_{18}H_{14}N_2NaO_3$  ( $M + Na$ )<sup>+</sup>: *calcd.*: 329.0897; *found*: 329.0900.

### **Synthesis of (–)-nopol-fluorescein-diol (2-14)**

#### **(–)-Nopol-phthalimide (2-24):**



#### **2-(2-((1*R*,5*S*)-6,6-Dimethylbicyclo[3.1.1]hept-2-en-2-yl)ethyl)isoindoline-1,3-dione (2-24):**

(–)-Nopol-tosyl was synthesized according to the described method earlier for the synthesis of compound **2-1a**. Then, (–)-nopol-tosyl was immediately used to synthesize (–)-nopol-phthalimide (**2-24**). To a solution of (–)-nopol-tosyl (1.0 equiv, 29.6 mmol) in dry DMF (11 ml), potassium phthalimide (1.1 equiv, 6.0 g, 32.5) was added under nitrogen at room temperature. Then, the reaction mixture was stirred at 90 °C for 16 h. The reaction mixture was diluted with EtOAc (100 ml). The organic phase was washed with distilled water (3 × 35 ml) and brine (1 × 20 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1.5:8.5, EtOAc/hexanes) to provide (–)-nopol-phthalimide (**2-24**) (5.8 g, 60% yield).

Yellow colorless oil,  $R_f = 0.53$  (1.5:8.5, EtOAc/hexanes).

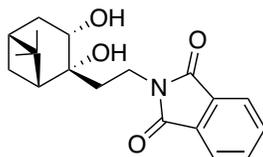
**<sup>1</sup>H NMR**  $\delta$ /ppm: (500 MHz, CDCl<sub>3</sub>) 7.85 – 7.81 (m, 2H), 7.72 – 7.68 (m, 2H), 5.34 – 5.30 (m, 1H), 3.78 – 3.63 (comp m, 2H), 2.42 – 2.35 (comp m, 2H), 2.34 – 2.26 (m, 1H), 2.25 – 2.12 (comp m, 3H), 2.09 – 2.04 (m, 1H), 1.29 (s, 3H), 1.11 (d,  $J = 8.6$  Hz, 1H), 0.82 (s, 3H).

**<sup>13</sup>C NMR**  $\delta$ /ppm: (126 MHz, CDCl<sub>3</sub>) 168.3, 144.6, 133.9, 132.3, 123.2, 119.0, 45.6, 40.7, 38.1, 36.5, 35.5, 31.8, 31.4, 26.3, 21.2.

**IR** (Microscope, cm<sup>-1</sup>) 2983, 2918, 2833, 1773, 1713, 1468, 1433, 1396, 1364.

**HRMS** (EI) for  $C_{19}H_{21}NO_2$  ( $M$ )<sup>+</sup>: *calcd.*: 295.15723; *found*: 295.15760.

**(-)-Nopol-phthalimide-diol (2-25):**



**2-(2-((1R,2R,3S,5R)-2,3-Dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl)isoindoline-1,3-dione (2-25):** (-)-Nopol-phthalimide-diol (**2-25**) was synthesized according to the procedure described earlier for the synthesis of compound **2-2a**. The crude residue was purified by flash chromatography (4:6, EtOAc/hexanes) to provide compound **2-25** (1.7 g, 60% yield).

White solid,  $R_f = 0.27$  (4:6, EtOAc/hexanes).

Mp = 91.0 – 93.4 °C.

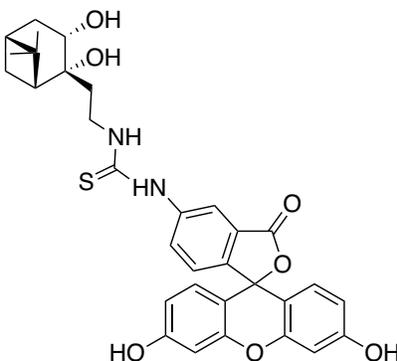
$[\alpha]_D^{20} : -5.4$  (*c* 0.65, CHCl<sub>3</sub>).

**<sup>1</sup>H NMR** δ/ppm: (498 MHz, CDCl<sub>3</sub>) 7.85 – 7.81 (m, 2H), 7.72 – 7.68 (m, 2H), 4.11 (ddd, *J* = 9.4, 5.4, 4.6 Hz, 1H), 3.88 (dd, *J* = 7.5, 6.5 Hz, 2H), 3.20 (d, *J* = 5.3 Hz, 1H), 3.10 (s, 1H), 2.50 (dddd, *J* = 14.2, 9.4, 3.5, 2.4 Hz, 1H), 2.23 – 2.16 (m, 1H), 2.08 (app t, *J* = 5.8 Hz, 1H), 2.00 – 1.86 (comp m, 3H), 1.68 (ddd, *J* = 14.1, 4.7, 2.7 Hz, 1H), 1.33 (d, *J* = 10.4 Hz, 1H), 1.26 (s, 3H), 0.93 (s, 3H).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CDCl<sub>3</sub>) 168.8, 134.0, 132.2, 123.3, 75.0, 67.6, 52.6, 40.5, 39.9, 38.8, 38.2, 33.5, 27.7, 27.4, 24.3.

**IR** (Microscope, cm<sup>-1</sup>) 3461, 2987, 2924, 2870, 1770, 1712, 1467, 1448, 1402, 1372.

**HRMS** (ESI-TOF) for C<sub>19</sub>H<sub>23</sub>NNaO<sub>4</sub> (M + Na)<sup>+</sup>: *calcd.*: 352.1519; *found*: 352.1522.



**1-(3',6'-Dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-yl)-3-(2-((1R,2R,3S,5R)-2,3-dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl)thiourea (2-14):** Compound **2-25**

(1.0 equiv, 0.13 g, 0.39 mmol) was dissolved in anhydrous ethanol (1.2 ml). Methylamine solution 40 wt. % in water (3.0 equiv, 0.1 ml, 1.2 mmol) was added to the reaction mixture, which was then heated at 50 °C for 2 h under ambient atmosphere. Then toluene (1 ml) was added to the reaction mixture that was then concentrated in vacuo. This step was repeated three times. The remaining crude residue (white solid) was kept under vacuum for 2 h and then it was used to synthesize (–)-nopoldiol fluorescein (**2-14**) without any purification. To the crude residue in anhydrous ethanol (3 ml), fluorescein isothiocyanate (0.66 equiv, 0.10 g, 0.26 mmol) was added under nitrogen balloon and stirred for 16 h. Then, the reaction mixture was concentrated in vacuo and purified by flash chromatography (1:1, acetone/hexanes) to obtain diol **2-14**.

Yellow-orange solid (76 mg, 51% yield),  $R_f = 0.26$  (1:1, hexanes/acetone).

Mp = 188 – 190 °C.

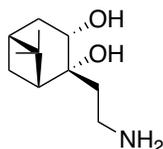
$^1\text{H NMR}$   $\delta$ /ppm: (498 MHz,  $\text{CD}_3\text{OD}$ ) 8.09 (d,  $J = 1.8$  Hz, 1H), 7.73 (d,  $J = 8.2$  Hz, 1H), 7.16 (dd,  $J = 8.7, 0.7$  Hz, 1H), 6.71 – 6.65 (comp m, 4H), 6.54 (ddd,  $J = 8.7, 2.4, 0.8$  Hz, 2H), 4.08 (dd,  $J = 9.4, 5.1$  Hz, 1H), 3.78 (bs, 2H), 2.52 – 2.43 (m, 1H), 2.23 – 2.16 (m, 1H), 2.13 (app t,  $J = 5.8$  Hz, 1H), 2.03 – 1.95 (m, 1H), 1.92 – 1.86 (m, 1H), 1.84 – 1.75 (m, 1H), 1.68 (ddd,  $J = 13.9, 5.1, 2.5$  Hz, 1H), 1.44 (d,  $J = 10.1$  Hz, 1H), 1.29 (s, 3H), 1.00 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CD}_3\text{OD}$ ) 182.3, 171.1, 161.5, 154.2, 142.2, 132.1, 130.3, 129.2, 125.8, 120.3, 113.7, 111.5, 103.5, 76.2, 68.7, 52.8, 49.6, 49.4, 49.3, 41.9, 41.7, 39.8, 39.2, 28.7, 28.4, 24.6.

IR (Microscope,  $\text{cm}^{-1}$ ): 3290, 2921, 1730, 1636, 1592.

HRMS (ESI-TOF): for  $\text{C}_{32}\text{H}_{31}\text{N}_2\text{O}_7\text{S}$  ( $\text{M} - \text{H}$ ) $^-$ : *calcd.*: 587.1857; *found*: 587.1850; for  $\text{C}_{32}\text{H}_{31}\text{ClN}_2\text{O}_7\text{S}$  ( $\text{M} + \text{Cl}$ ) $^-$ : *calcd.*: 623.1624; *found*: 623.1632.

### Synthesis of (–)-nopoldiol amine (**2-26**)



**(1R,2R,3S,5R)-2-(2-Aminoethyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (2-26):** (–)-Nopoldiol azide **2-1b** (0.19 g, 0.83 mmol) was dissolved in methanol (3.0 ml) under nitrogen in a single-neck round bottom flask. Adapter with both vacuum and nitrogen inlet was placed on the flask. The reaction mixture was degassed for 1 min and then purged with nitrogen for another 1

min. This was repeated three times. Then, 10 wt% Palladium on carbon (10 mg) was added to the reaction mixture. Degassing and purging with nitrogen was performed 3 more times. Finally, nitrogen was replaced with hydrogen filled balloon and the reaction mixture was stirred for 3 h at room temperature. Hydrogen gas was removed under vacuum before the adapter was removed. The reaction mixture was filtered through celite and the filtrate was concentrated in vacuo to obtain pure (–)-nopoldiol amine (**2-26**) (0.16 g, 97% yield).

Brown-yellowish oil.

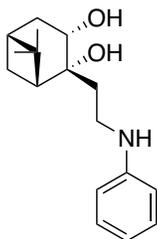
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 3.99 (dd, *J* = 9.3, 5.6 Hz, 1H), 3.10 – 3.03 (m, 2H), 2.48 – 2.37 (m, 1H), 2.22 – 2.12 (m, 1H), 2.07 (app t, *J* = 5.8 Hz, 1H), 1.92 – 1.86 (m, 1H), 1.78 – 1.71 (m, 1H), 1.71 – 1.62 (m, 1H), 1.60 – 1.52 (m, 1H), 1.47 (d, *J* = 10.2 Hz, 1H), 1.26 (s, 3H), 0.93 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 76.3, 67.9, 53.8, 41.8, 40.6, 38.8, 38.0, 37.7, 28.1, 27.8, 24.3.

IR (cast film, cm<sup>-1</sup>): 3520, 3429, 3354, 3319, 2986, 2921, 2867, 1098, 1053.

HRMS (ESI-TOF): for C<sub>11</sub>H<sub>22</sub>NO<sub>2</sub> (M + H)<sup>+</sup>: *calcd.*: 200.1645; *found*: 200.1641.

### Synthesis of (–)-nopoldiol phenylamine (**2-15**)



#### (1*R*,2*R*,3*S*,5*R*)-6,6-Dimethyl-2-(2-(phenylamino)ethyl)bicyclo[3.1.1]heptane-2,3-diol (**2-15**):

The procedure was adapted from the reported method by Buchwald and coworkers.<sup>33</sup> CuI (0.05 equiv, 3.3 mg, 0.017 mmol), K<sub>3</sub>PO<sub>4</sub> (2.0 equiv, 0.14 g, 0.66 mmol) and compound **2-26** (1.2 equiv, 80 mg, 0.40 mmol) were added to a microwave vial. The tube was evacuated and back-filled with nitrogen three times. 2-Propanol (0.33 ml), ethylene glycol (2.0 equiv, 37 μl, 0.66 mmol) and iodobenzene (1.0 equiv, 37 μl, 0.33 mmol) were added by syringe at room temperature under nitrogen. The reaction mixture was heated to 80 °C to furnish a pink-brown suspension and stirred for 16 h. Then, the reaction mixture was cooled to rt. Distilled water (2.0 ml) was added, and the mixture was extracted with diethyl ether (4 x 10 ml). The combined

organic phases were washed with brine (1 x 5.0 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:3, EtOAc/hexanes) to provide compound **2-15** as a pink powder (63 mg, 69% yield).

Melting point: 93.0 – 95.8 °C.

Pink powder: R<sub>f</sub> = 0.62 (1:1, EtOAc/hexanes)

<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.20 (dd, *J* = 8.6, 7.4 Hz, 2H), 6.79 (tt, *J* = 7.3, 1.1 Hz, 1H), 6.73 (dd, *J* = 8.6, 1.1 Hz, 2H), 4.09 (dd, *J* = 9.4, 5.3 Hz, 1H), 3.45 – 3.32 (comp m, 2H), 2.49 (dddd, *J* = 14.0, 9.4, 3.7, 2.4 Hz, 1H), 2.25 – 2.16 (m, 1H), 2.13 (app t, *J* = 5.8 Hz, 1H), 2.01 – 1.90 (comp m, 2H), 1.81 (ddd, *J* = 14.5, 7.5, 5.1 Hz, 1H), 1.67 (ddd, *J* = 13.9, 5.3, 2.4 Hz, 1H), 1.41 (d, *J* = 10.3 Hz, 1H), 1.28 (s, 3H), 0.95 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) δ 147.0, 128.3, 118.0, 113.5, 76.3, 76.1, 75.8, 75.0, 67.2, 51.9, 39.9, 39.6, 39.6, 37.9, 37.2, 26.9, 26.7, 23.3.

IR (cast film, cm<sup>-1</sup>): 3482, 3364, 3256, 3023, 3057, 2993, 2941, 2912, 2868, 2840, 1606, 1509, 1470, 1324, 1004.

HRMS (ESI-TOF): for C<sub>17</sub>H<sub>26</sub>NO<sub>2</sub> (M + H)<sup>+</sup>: *calcd.*: 276.1958; *found*: 276.1954.

### 2.4.3 Determining K<sub>eq</sub> via NMR

Nopol boronic ester **2-5** were synthesized and purified according to the procedure explained above. Then, boronates **2-5** were kept under vacuum overnight to make sure that they were dry. 0.1 M of D<sub>2</sub>O potassium phosphate buffer pD 7.4 was freshly prepared and mixed with ACN-d<sub>3</sub> as 65:35 w%. In this solvent system, 5.0 mM of **2-5** was prepared as stock solutions. The stock solutions were used to prepare several dilutions (2.5, 1.0, 0.8, 0.6, 0.4, 0.2, 0.1 mM, as duplicates) that were then allowed to reach equilibrium for 2 h. DMSO (4 mM) in the same solvent system was used as an internal standard. Each nopol boronic ester dilution (600 μl) and internal standard solution (50 μl) were mixed and shaken in an NMR tube. Free diol **2-2b** concentration in the dilutions was determined from the calibration curve as explained below. Ratio of [**2-5**]/[**2-3**] was obtained by the integral ratio of CH<sub>3</sub>/CH<sub>3</sub> as diol **2-2b** concentration was considered equal to boronic acid **2-3** concentration. Finally, K<sub>eq</sub> of each dilution was calculated with the formula as shown in Scheme 2-5. At least 14 different K<sub>eq</sub> values were obtained for each boronates **2-5**. The data was recorded as their average of at least n = 14 measurements ± expanded uncertainty ([standard deviation] × (t-value at 95% confidence interval when degrees of

freedom =  $n - 1$ ) /  $\sqrt{n}$  as shown in Table 2-2. T-distribution two-tailed table was used to obtain  $t$  values at 95% confidence interval.

For the calibration curve, five different dilutions (0.5, 0.4, 0.3, 0.2, 0.1 mM) of diol **2-2b** were prepared as duplicates. Then, 600  $\mu\text{l}$  of each **2-2b** dilution and 50  $\mu\text{l}$  internal standard solution were mixed and shaken in an NMR tube. The calibration curve was plotted as final **2-2b** concentration in an NMR tube vs integral ratio of  $\text{CH}_3/\text{DMSO}$ .

#### 2.4.4 Kinetic study

##### 2.4.4.1 NMR study

###### 2.4.4.1.1 Forward rate constant measurement

Diol **2-2a** or **2-2b** was dissolved in 0.05 M of  $\text{D}_2\text{O}$  potassium phosphate buffer:ACN- $\text{d}_3$  at 1 mM concentration. Initial reaction concentration was determined according to its rate. For instance, the concentration of fast reactions was dropped to 0.5 mM in order to effectively monitor their rates. Internal standard, DMSO, was mixed with 0.05 M of  $\text{D}_2\text{O}$  potassium phosphate buffer:ACN- $\text{d}_3$  at 9 mM. Then, this internal standard solution was used to dissolve arylboronic acids (12 mM). To 600  $\mu\text{l}$  of diol **2-2a** or **2-2b** solution (1.0 mM), 50  $\mu\text{l}$  of arylboronic acid solution (12 mM) was added, shaken quickly, and followed by monitoring the reaction at least every 10 s over 10 to 60 min. A calibration curve was used to calculate diol concentration over time. The second order rate constant was determined by plotting  $1/[\mathbf{2-2a}$  or  $\mathbf{2-2b}]$  vs time. The slope was recorded as the second order rate constant. Each reaction was performed at least three times. The average of at least 3 measurements  $\pm$  standard deviation was recorded as in Table 2-2.

###### 2.4.4.1.2 Monitoring the backward reaction

Boronates **2-5d** and **2-5h** (32.5 mM) were dissolved in ACN, since they are not susceptible to hydrolysis in ACN. Internal standard, DMSO (9 mM), was mixed with 0.1 M of  $\text{D}_2\text{O}$  potassium phosphate buffer:ACN- $\text{d}_3$  (65:35 w%, pD 7.4). To 580  $\mu\text{l}$  of 0.1 M of  $\text{D}_2\text{O}$  potassium phosphate buffer:ACN- $\text{d}_3$  (65:35 w%, pD 7.4) in an NMR tube, 50  $\mu\text{l}$  of internal standard solution and 20  $\mu\text{l}$  of boronates **2-5d** or **2-5h** (final concentration; 1 mM) were added, shaken quickly, and followed by monitoring the reaction at least every 14 s over 5 to 60 min. For boronate **2-5d**, even though every 14-second point was detected, not all data was plotted due to

its slow hydrolysis. A calibration curve was used to calculate diol concentration over time. [2-5]/[2-3] ratio was obtained by the integral ratio of  $\text{CH}_3/\text{CH}_3$  as diol **2-2b** concentration was considered equal to boronic acid **2-3** concentration. Therefore,  $Q$  ( $\text{M}^{-1}$ ) was calculated by using the formula indicated in Scheme 2-5. The backward reaction was plotted as time vs Quotient ( $Q$ ). Backward rate constant ( $k_{\text{OFF}}$ ) was not calculated but time required to reach the equilibrium was noted for each nopol boronic esters.

#### **2.4.4.2 Fluorescence quenching study – FRET**

##### **2.4.4.2.1 Forward rate constant determination**

The quenching experiment involves the following steps;

1. The stock solutions of diol **2-1** (500  $\mu\text{M}$ ), nopol-danysl **2-11** (500  $\mu\text{M}$ ) and dabcytl tagged boronates **2-9a** and **2-9b** (500  $\mu\text{M}$ ) were prepared in ACN.
2. Different concentrations of diol **2-1** (20, 17.5, 15, 12.5, 10, 7.5, 5 and 2.5  $\mu\text{M}$ ) were prepared with potassium phosphate buffer (0.1 M, pH 7.4)/ACN solution 80:20 w% for calibration curve. Fluorescence unit vs [diol **2-1**] graph was plotted and used to calculate the concentration of diol **2-1** during the quenching experiment.
3. For final concentration (20  $\mu\text{M}$ ), quenching experiment was performed by adding stock diol **2-1** (8  $\mu\text{l}$ ) and stock boronates **2-9a** or **2-9b** (8  $\mu\text{l}$ ) into the buffer solution (184  $\mu\text{l}$ ). Fluorescence was monitored at 335 nm every 2 or 3 seconds over 10 min. The delay between mixing time and obtaining first data was counted as around 14 seconds.
4. For final concentration (10  $\mu\text{M}$ ), quenching experiment was performed by adding stock diol **2-1** (4  $\mu\text{l}$ ) and stock boronates **2-9a** or **2-9b** (4  $\mu\text{l}$ ) into the buffer solution (192  $\mu\text{l}$ ). Fluorescence was monitored at 335 nm every 2 or 3 seconds over 10 min. Second order kinetic data was obtained and the results are for **2-9a**;  $338 \pm 34 \text{ M}^{-1}\text{s}^{-1}$  and for **2-9b**:  $1201 \pm 279 \text{ M}^{-1}\text{s}^{-1}$ .

##### **2.4.4.2.2 Backward reaction and determining $K_{\text{eq}}$**

The fluorescence was monitored at the maximum absorption of diol **2-1** (335 nm) in potassium phosphate buffer (0.1 M, pH 7.4)/ACN solution 80:20 w%. The quenching was monitored at 30, 20, 15 or 10  $\mu\text{M}$  over 5 to 30 min. The concentration of diol **2-1** was calculated

from the calibration curve and  $K_{eq}$  was estimated from this value by using the following formula:

$$\text{Quotient (Q)} = \frac{[\mathbf{2-10a} \text{ or } \mathbf{2-10b}]}{[\text{diol } \mathbf{2-1}] * [\mathbf{2-9a} \text{ or } \mathbf{2-9b}]} = \frac{([\text{Initial concentration } \mathbf{2-10a} \text{ or } \mathbf{2-10b}] - [\text{diol } \mathbf{2-1}])}{[\text{diol } \mathbf{2-1}]^2} \text{ using } [\text{diol } \mathbf{2-1}] = [\mathbf{2-9a} \text{ or } \mathbf{2-9b}] \text{ in mol/L}$$

All Q values were plotted against time (s) and steady state region was used to estimate  $K_{eq}$ .

Monitoring the backward reaction involves the following steps;

1. The stock solutions of diol **2-1** (500  $\mu\text{M}$ ), boronates **2-10a** and **2-10b** (500  $\mu\text{M}$ ), dabcy1 tagged boronates **2-9a** and **2-9b** (500  $\mu\text{M}$ ) were prepared in ACN.
2. Different concentrations of diol **2-1** (40, 30, 20, 15, 10, 7.5, 5 and 2.5  $\mu\text{M}$ ) were prepared with potassium phosphate buffer (0.1 M, pH 7.4)/ACN solution 80:20 w% for calibration curve. Fluorescence unit vs [diol **2-1**] graph was plotted and used to calculate the concentration of diol **2-1** during the quenching experiment.
3. For final concentration (30  $\mu\text{M}$ ), quenching experiment was performed by adding boronate **2-10a** or **2-10b** (12  $\mu\text{l}$ ) into the buffer solution (188  $\mu\text{l}$ ). For final concentration (20  $\mu\text{M}$ ), quenching experiment was performed by adding boronate **2-10a** or **2-10b** (8  $\mu\text{l}$ ) into the buffer solution (192  $\mu\text{l}$ ). For final concentration (10  $\mu\text{M}$ ), quenching experiment was performed by adding boronate **2-10a** or **2-10b** (4  $\mu\text{l}$ ) into the buffer solution (196  $\mu\text{l}$ ). Fluorescence was monitored at 335 nm every 3, 5 or 10 seconds over 5 to 30 min. The delay between mixing time and first data was determined as  $\sim 14$  seconds. The concentration of diol **2-1** was used to estimate and plot Q vs time graph.

The forward reaction was also repeated to validate that both backward and forward reactions provided the similar  $K_{eq}$ .

Monitoring the forward reaction involves the following steps;

1. For final concentration (10  $\mu\text{M}$ ), quenching experiment was performed by adding stock diol **2-1** (4  $\mu\text{l}$ ) and stock boronates **2-9a** or **2-9b** (4  $\mu\text{l}$ ) into the buffer solution (192  $\mu\text{l}$ ). Fluorescence was monitored at 335 nm every 2 or 3 seconds over 10 min. The delay between mixing time and first data was determined as  $\sim 14$  seconds. The concentration of diol **2-1** was used to estimate and plot Q vs time graph.

#### 2.4.5 Biological competition assay *via* RP-HPLC

The HPLC chromatograms were acquired from an Agilent 1100 Series Instrument. Chromatographic separation was achieved on Agilent ZOBAX SB – C18 Column (250 x 46 mm, 5  $\mu$ m; 1ml/min, 40 °C). Mass analyzer is Agilent G2908BA LC/MSD Quad. The mobile phase consisted of 0.1% AA in H<sub>2</sub>O (solvent A) and ACN (solvent B). A stepwise gradient was used for sample elution: 90% solvent A to 90% solvent B over a 20-min period. The sample injections were 10  $\mu$ l or 20  $\mu$ l.

Boronate conversions were determined *via* RP-PHLC using either UV (260 nm) or fluorescence (excitation 320 nm, emission 560 nm) after 30 min. In order to perform the experiment, first stock solutions of D-glucose (16 mM), D-fructose (5 mM) and catechol (10  $\mu$ M) were separately prepared in 25 mM phosphate buffer solution (pH 7.4). Besides, stock solutions of diol **2-1** (2.5 mM), nopol-dansyl **2-11** (2.5 mM), boronic acids **2-3d** and **2-3h** (2.5 mM) were prepared in a solution of 25 mM phosphate buffer (pH 7.4):ACN (50/50 v/v%).

Reaction mixture was prepared according to final concentrations of diol **2-1** and **2-3d/2-3h** (20 or 10  $\mu$ M), D-glucose (8 mM), D-fructose (0.3 mM) and catechol (0.01  $\mu$ M). Each solution was prepared as duplicate.

##### In the presence of biocompetitors;

For diol **2-1** and **2-3d/2-3h** (20  $\mu$ M), the following mixture was prepared from stock solutions:

Diol **2-1** (4  $\mu$ l), **2-3d/2-3h** (4  $\mu$ l), D-glucose (250  $\mu$ l), D-fructose (30  $\mu$ l), catechol (0.5  $\mu$ l) were mixed in 25 mM phosphate buffer solution (211.5  $\mu$ l).

For diol **2-1** and **2-3d/2-3h** (10  $\mu$ M):

Diol **2-1** (2  $\mu$ l), **2-3d/2-3h** (2  $\mu$ l), D-glucose (250  $\mu$ l), D-fructose (30  $\mu$ l), catechol (0.5  $\mu$ l) were mixed in 25 mM phosphate buffer solution (215.5  $\mu$ l).

##### In the absence of biocompetitors;

For diol **2-1** and **2-3d/2-3h** (20  $\mu$ M), the following mixture was prepared from stock solutions.

Diol **2-1** (4  $\mu$ l) and **2-3d/2-3h** (4  $\mu$ l) were mixed in 25 mM phosphate buffer solution (492  $\mu$ l).

For diol **2-1** and **2-3d/2-3h** (10  $\mu$ M):

Diol **2-1** (2  $\mu$ l) and **2-3d/2-3h** (2  $\mu$ l) were mixed in 25 mM phosphate buffer solution (496  $\mu$ l).

#### **2.4.6 SDS-PAGE study**

##### **SDS-PAGE protocol:**

##### **Protein labeling:**

Bovine serum albumin conjugates were prepared by the treatment of BSA (5 mg/mL) with maleimide conjugated boronic acid **2-12** (100  $\mu$ M) for overnight in 20 mM phosphate buffer pH 7.2. Then, conjugated BSA was reacted with nopoldiol fluorescein **2-14** at room temperature for 5 min to 120 min.

##### **In-gel fluorescence analysis of BSA conjugates:**

Boronic acid-labeled BSA samples were treated with compound **2-14** (20 – 100  $\mu$ M) or DMSO. After 5 min to 120 min, the modified BSA samples were analyzed by gel electrophoresis using 10% polyacrylamide gels. Phosphate buffer was used in gel preparation. Samples were heated at 95 °C for 2.5 min before they were loaded on gels. Running time was 20 min at 175 V. Gels were washed with distilled water and analyzed by in-gel fluorescence measurements on a GE ImageQuant RT ECL Imager. Fluorescence was measured at a 494 nm excitation wavelength and 512 nm emissions. Total protein loading was confirmed by subsequent staining with Coomassie brilliant blue.

##### **SDS-PAGE gel preparation with phosphate buffer**

Resolving Gel (prepared freshly each time):

- 4 ml of Solution A (30% acrylamide)
- 2.5 ml of Solution B (75 ml of 0.2 M sodium phosphate buffer (pH 8.8), 4 ml of 10% SDS, 21 ml H<sub>2</sub>O)
- 3.5 ml of H<sub>2</sub>O
- 50  $\mu$ l of 10% ammonium persulfate
- 5  $\mu$ l of TEMED (initiator)

Stacking gel (prepared freshly each time):

- 2.3 ml of H<sub>2</sub>O
- 0.67 ml of Solution A (30% acrylamide)
- 1.0 ml of Solution C (50 ml of 0.1 M sodium phosphate buffer (pH 6.8), 4 ml 10% SDS, 46 ml H<sub>2</sub>O)
- 30 µl of 10% ammonium persulfate
- 5 µl of TEMED

Sample buffer stock:

- 0.6 ml of 0.1 M sodium phosphate buffer (pH 6.8)
- 5 ml of 50% glycerol
- 2 ml of 10% SDS
- 1ml of 1% bromophenol blue
- 1.4 ml of H<sub>2</sub>O

Sample preparation (before loading to the gel):

- 6 µl of H<sub>2</sub>O
- 2 µl of sample buffer
- 4 µl of sample

Running buffer:

- 3.9 g of NaH<sub>2</sub>PO<sub>4</sub>•H<sub>2</sub>O
- 19.3 g of Na<sub>2</sub>HPO<sub>4</sub>•7H<sub>2</sub>O
- 0.5 g of SDS
- 2 liter of H<sub>2</sub>O

## 2.4.7 Trx study

### HPLC-Trx explanation:

RP-HPLC-UV-MS was performed using an Agilent 1200 SL HPLC System with a Jupiter C18 column, 5 $\mu$ m particle size, 300Å pore size, 50x2.0 mm (Phenomenex, Torrance, USA), thermostated at 40°C, with a buffer gradient system composed 0.1% trifluoroacetic acid in water as mobile phase A and 0.1% trifluoroacetic acid in ACN as mobile phase B. Samples were loaded onto the column at a flow rate of 0.50 mLmin<sup>-1</sup> and an initial buffer composition of 80% mobile phase A and 20% mobile phase B. After injection, the column was washed using the initial loading conditions for 2.0 min to remove the salt, followed by elution of the analytes by using a linear gradient from 20% to 30% mobile phase B over a period of 3 min, 30% to 45% mobile phase B over a period of 25 min, 45% to 65% mobile phase B over a period of 10 min, 65% to 95% mobile phase B over a period of 4 min and 95% to 20% mobile phase B over a period of 1 min. UV absorbance was monitored at 214 nm. Mass spectra were acquired in positive mode of ionization using an Agilent 6220 Accurate-Mass orthogonal TOF HPLC/MS system (Santa Clara, CA, USA) equipped with a dual sprayer electrospray ionization source with the second sprayer providing a reference mass solution. Mass spectrometric conditions were drying gas 10 L/min at 325°C, nebulizer 25 psi, mass range 100-3200 Da, acquisition rate of ~1.03 spectra/sec, fragmentor 200V, skimmer 65V, capillary 4000V, instrument state 4GHz High Resolution. Mass correction was performed for every individual spectrum using peaks at m/z 121.0509 and 922.0098 from the reference solution. Data acquisition was performed using the Mass Hunter software package (ver. B.04.00.). Analysis of the HPLC-UV-MS data was achieved using the Agilent Mass Hunter Qualitative Analysis software (ver. B.03.01 SP3).

### C18 Protein/Peptide Purification

C18 Micro SpinColumn<sup>TM</sup> purchased from HARVARD Apparatus (Cat. No. 74-4601; 96pcs, 25-75  $\mu$ l sample).

#### C18 purification:

1) To the spin column, add 300  $\mu$ l of water and let it hydrate for around 10 minutes. Centrifuge for 2-3 min at 500 RCF (relative centrifugal force)

- 2) Wash the column 3x with 330  $\mu$ l of 0.1% TFA ACN/0.1% TFA H<sub>2</sub>O (70/30%). Centrifuge at 650 RCF for 1.30 minutes.
- 3) Wash columns 3x with 330  $\mu$ l of 0.1% TFA ACN/0.1% TFA H<sub>2</sub>O (5/95%) to equilibrate the column. Centrifuge at 750 RCF for 2 minutes.
- 4) Load your sample on column. Centrifuge at 750 RCF for 2 minutes. Repeat it 3 times. Then the liquid was stored.
- 5) Wash the column 3 times with 100  $\mu$ l of 0.1% TFA ACN/0.1% TFA H<sub>2</sub>O (5/95%). The eluent was still stored with the previous liquid. The compound was expected to be in the column and some undesired impurities was separated from desired proteins with this washing step.
- 6) Proteins or peptides can be eluted from the C18 microspin column using 2x100  $\mu$ l of 0.1% TFA ACN/ 0.1% TFA H<sub>2</sub>O (60/40%).
- 7) Finally the protein will be dried with speed vacuum in mass facility.

This study was performed by HPLC and the corresponding peaks were determined and assigned with UV (215 nm) and LC-MS. The concentrations in brackets are the final concentrations. Trx purchased from Sigma-Aldrich (~11.7 kDa, 50  $\mu$ M) was reduced with TCEP (1 mM) for overnight at room temperature. Then, it was mixed with maleimide derivatized boronic acid **2-12** (100  $\mu$ M) for 30 min to give boronic acid-Trx. Complete conversion to boronic acid-Trx (12021 Da) and also bpin-Trx (12122 Da) was observed in the absence of serum. After fetal bovine serum was added to boronic acid-Trx solution, exact conversion could not be deducted due to spectral overlap with serum compounds. Then, diol **2-2b** (200  $\mu$ M) was added and kept for 30 min in conjugated Trx solutions. In the absence of serum, boronate-Trx (12333 Da, >50% conversion) was observed. In the presence of serum, a high amount of boronate-Trx was detected. As a negative control, compound **2-13** (maleimide derivative without boron) was applied and reacted with Trx followed by the addition of diol **2-2b** into the compound **2-13**-Trx adduct. As expected, boronate formation was not detected. Serum containing solutions were purified with C18 Micro SpinColumn before injected to HPLC.

Similar study was performed with fluorescent diol **2-1** and the corresponding peaks were determined and assigned with UV (215 nm) or fluorescence (excitation 320 nm, emission 560 nm) and LC-MS. Trx (11673 Da, 200  $\mu$ M) was reduced with TCEP (1 mM) for overnight at

room temperature to obtain reduced Trx (11675 Da). Then, it was mixed with boronic acid **2-12** (120  $\mu$ M) for 60 min to give boronic acid-Trx (12021 Da). Then complete cell media (DMEM + 10% fetal bovine serum + 1% ampicillin + 1% glutamax) was added to boronic acid-Trx solution. Lastly, diol **2-1** (20  $\mu$ M) was added and kept for 60 min in conjugated Trx solutions. Boronate-Trx (12517 Da, 83% conversion) was observed.

#### 2.4.8 Cytotoxicity assay

HEK 293 cells (grown in complete DMEM media with 10% fetal bovine serum (FBS), 2 mM glutamine) were grown to about 80% confluency before removal from the surface of a 20 cm petri-dish with 25% Trypsin with EDTA (5 minutes, 37 °C, 5% CO<sub>2</sub>). Cells were pelleted (9000 rpm, 5 minutes, 27 °C), diluted in 10 mL fresh complete DMEM media to generate aliquots of 30,000 cells per well (100  $\mu$ L), then incubated for 6 hours before addition of diol **2-15** (50 and 25  $\mu$ M), boronic acid **2-3d** (50 and 25  $\mu$ M), Triton X-100 (1  $\mu$ L) and DMSO (1  $\mu$ L). Stock solutions of diol and boronic acid were prepared fresh in DMSO (2.5 and 5 mM), and diluted to the working concentration as listed above in DMEM so that the total volume of DMSO added was kept constant at 1% (v/v). For each reagent, samples were set up as three replicates with one time point (18 hours). Control wells were loaded with DMSO (1  $\mu$ L) or Triton X-100 (1  $\mu$ L) (used as positive control for cytotoxicity). After incubation for 18 hours, WST-1 reagent (10  $\mu$ L) to each well was added and the wells were incubated for another 4 hours. Absorbance at 450 nm and 600 nm was recorded with a plate reader. The results of the averages and standard deviations of the wells were determined after normalization to the DMSO control.

#### 2.5 References

1. Reviews: a) Zheng, M.; Zheng, L.; Zhang, P.; Li, J.; Zhang, Y. *Molecules* **2015**, *20*, 3190–3205. b) Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16–20; c) Ramil, C. P.; Lin, Q. *Chem. Commun.* **2013**, *49*, 11007–11022. d) Sletten, E. M.; Bertozzi, C. R. *Acc. Chem. Res* **2011**, *44*, 666–676. e) Chen, Y. – X.; Triola, G.; Waldmann, H. *Acc. Chem. Res* **2011**, *44*, 762–773. f) Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010**, *39*, 1272–1279. g) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
2. Reviews: a) McKay, C. S.; Finn, M. G. *Chem. Biol.* **2014**, *21*, 1075–1101. b) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
3. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.

4. a) Šečkutė, J.; Devaraj, N. K. *Curr. Opin. Chem. Biol.* **2013**, *17*, 761–767. b) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519. c) Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y.; *J. Am. Chem. Soc.* **2012**, *134*, 792–795.
5. Shih, H-W.; Kamber, D. N.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2014**, *21*, 103–111.
6. Reviews: a) Wu, X.; Li, Z.; Chen, X. – X.; Fossey, J. S.; James, T. D.; Jiang, Y.-B. *Chem. Soc. Rev.* **2013**, *42*, 8032–8048. b) Jin, S.; Cheng, Y.; Reid, S.; Li, M.; Wang, B. *Med. Res. Rev.* **2010**, *30*, 171–257. c) James, T. D.; Shinkai, S. *Topics Curr. Chem.* **2002**, *218*, 159–200. d) Wang, W.; Gao, X.; Wang, B. *Curr. Org. Chem.* **2002**, *6*, 1285–1317.
7. a) Liu, H.; Li, Y.; Sun, K.; Fan, J.; Zhang, P.; Meng, J.; Wang, S.; Jiang, L. *J. Am. Chem. Soc.* **2013**, *135*, 7603–7609. b) Pan, G.; Guo, B.; Ma, Y.; Cui, W.; He, F.; Li, B.; Yang, H.; Shea, K. J. *J. Am. Chem. Soc.* **2014**, *136*, 6203–6206. Reviews: c) Bull, S. D.; Davidson, M. G.; Van Den Elsen, J. M. H.; Fossey, J. S.; Jenkins, A. T. A.; Jiang, Y-B.; Kubo, Y.; Marken, F.; Sakurai, K.; Zhao, J.; James, T. D. *Acc. Chem. Res.* **2013**, *46*, 312–326.
8. a) Lippert, A. R.; Van De Bittner, G. C.; Chang, C. J. *Acc. Chem. Res.* **2011**, *44*, 793–804. b) Wang, F.; Niu, W.; Guo, J.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 10132–10135. c) Lin, V.S.; Dickinson, B.C.; Chang, C. J. *Methods Enzymol.* **2013**, *526*, 19–43. d) Chen, Z-J.; Ren, W.; Wright, Q. E.; Ai, H-W. *J. Am. Chem. Soc.* **2013**, *135*, 14940–14943. e) Sun, X.; Xu, Q.; Kim, G.; Flower, S. E.; Lowe, J. P.; Yoon, J.; Fossey, J. S.; Qian, X.; Bull, S. D.; James, T. D. *Chem. Sci.* **2014**, *5*, 3368–3373.
9. a) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220–8223. b) Liu, C. C.; Mack, A. V.; Brustad, E. M.; Mills, J. H.; Groff, D.; Smider, V. V.; Schultz, P. G. *J. Am. Chem. Soc.* **2009**, *131*, 9616–9617. c) Ellis, G. A.; Palte, M. J.; Raines, R. T. *J. Am. Chem. Soc.* **2012**, *134*, 3631–3634.
10. Shin, S. B. Y.; Almeida, R. D.; Gerona-Navarro, G.; Bracken, C.; Jaffrey, S. R. *Chem. Biol.* **2010**, *17*, 1171–1176.
11. a) Lorand, J. P.; Edwards, J. O. *J. Org. Chem.* **1959**, *24*, 769–774. b) Babcock, L.; Pizer, R. *Inorg. Chem.* **1980**, *19*, 56–61. c) Pizer, R.; Tihal, C. *Inorg. Chem.* **1992**, *31*, 3243–3247. d) Yan, J.; Springsteen, G.; Deeter, S.; Wang, B. *Tetrahedron* **2004**, *60*, 11205–11209. e) Watanabe, E.; Miyamoto, C.; Tanaka, A.; Iizuka, K.; Iwatsuki, S.; Inamo, M.; Takagi, H. D.; Ishihara, K. *Dalton Trans.* **2013**, *42*, 8446–8453. f) Okamoto, T.; Tanaka, A.;

- Watanabe, E.; Miyazaki, T.; Sugaya, T.; Iwatsuki, S.; Inamo, M.; Takagi, H. D.; Odani, A.; Ishihara, K. *Eur. J. Inorg. Chem.* **2014**, 2389–2395. g) Furikado, Y.; Nagahata, T.; Okamoto, T.; Sugaya, T.; Iwatsuki, S.; Inamo, M.; Takagi, H. D.; Odani, A.; Ishihara, K. *Chem. Eur. J.* **2014**, *20*, 13194–13202. Review: h) Peters, J. A. *Coord. Chem. Rev.* **2014**, *268*, 1–22.
12. a) Ni, N.; Laughlin, S.; Wang, Y.; Feng, Y.; Zheng, Y.; Wang, B. *Bioorg. Med. Chem.* **2012**, *20*, 2957–2961. b) Tomsho, J. W.; Benkovic, S. J. *J. Org. Chem.* **2012**, *77*, 2098–2106. c) Monajemi, H.; Cheah, M. H.; Lee, V. S.; Mohd. Zain, S.; Wan Abdullah, W. A. T. *RSC Adv.* **2014**, *4*, 10505–10513.
  13. Matteson, D. S.; Man, H.-W. *J. Org. Chem.* **1996**, *61*, 6047–6051.
  14. a) Roy, C. D.; Brown, H. C. *Monatsh. Chem.* **2007**, *138*, 747–753. b) Roy, C. D.; Brown, H. C. *Monatsh. Chem.* **2007**, *138*, 879–887.
  15. Baggio, R. F.; Sparks, A. L.; Juo, R.-R.; Arenas, J. E. US Patent 0277143 A1 December 15, 2005.
  16. a) Ahmad, M.; Hamer, J. *J. Chem. Educ.* **1964**, *41*, 249–250. b) Pollard, T. D.; De La Cruz, E. M. *Mol. Biol. Cell.* **2013**, *24*, 1103–1110.
  17. Crouch, S. R.; Holwr, F. J.; Notz, P. K.; Beckwith, P. M. *Appl. Spectrosc. Rev.* **1977**, *13*, 156–259.
  18. Thordarson, P. *Chem. Soc. Rev.* **2011**, *40*, 1305–1323.
  19. Carlise, J. R.; Kriegel, R. M.; Rees, W. S.; Weck, M. *J. Org. Chem.* **2005**, *70*, 5550–5560.
  20. a) Tius, M. A.; Gu, X-q; Truesdell, J. W.; Savariae, S.; Crooker, P.P. *Synthesis* **1988**, *1*, 36–40. b) Hayashi, T.; Konishi, M.; Kobori, Y.; Kumada, M.; Higuchi, T.; Hirotsu, K. *J. Am. Chem. Soc.* **1984**, *106*, 158–163.
  21. Marshall, J. A.; Johns, B. A. *J. Org. Chem.* **1998**, *63*, 7885–7892.
  22. Everson, D. A.; Jones, B. A.; Weix, D. J. *J. Am. Chem. Soc.* **2012**, *134*, 6146–6159.
  23. Springsteen, G; Wang, B. *Tetrahedron* **2002**, *58*, 5291–5300.
  24. Hall, D. G. *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials (Volume 1 and 2), Second Edition* (Ed: Hall, D. G.), Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, p. 10.
  25. Pal, A.; Bérubé, M.; Hall, D. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 1492–1495.
  26. Chotana, G. A.; Rak, M. A.; Smith, M. R. *J. Am. Chem. Soc.* **2005**, *127*, 10539–10544.

27. a) Volkhard, H. "Fluorescence Resonance Energy Transfer." *Principles of Computational Cell Biology*. Wiley-VCH Verlag GmbH & Co. KGaA, **2008**, p. 202. b) Clegg, R. M. "Förster resonance energy transfer—FRET: what is it, why do it, and how it's done." *FRET and FLIM Techniques. Laboratory Techniques in Biochemistry and Molecular Biology* (Ed: Gadella; T. W. J.), Elsevier, **2009**, Volume 33, p. 1–57.
28. Johansson, M. K. "Choosing Reporter-Quencher Pairs for Efficient Quenching Through Formation of Intramolecular Dimers." *Methods in Molecular Biology* (Ed: Didenko, V. V.), Humana Press: Totowa, NJ, **2006**, p. 17–29.
29. a) de Champlain, J.; Farley, L.; Cousineau, D.; van Ameringen, M. R. *Circ. Res.* **1976**, *38*, 109–114. b) Kawasaki, T.; Akanuma, H.; Yamanouchi, T. *Diabetes Care* **2002**, *25*, 353–357.
30. Borner, M. M.; Schneider, E.; Pirnia, F.; Sartor, O.; Trepel, J. B.; Myers, C. E. *FEBS Lett.* **1994**, *353*, 129–132.
31. a) Ngamwongsatit, P.; Banada, P. P.; Panbangred, W.; Bhunia, A. K. *J. Microbiol. Methods.* **2008**, *73*, 211–215. b) Kim, K-M.; Lee, S-B.; Lee, S-H.; Lee, Y-K.; Kim, K-N. *Key Eng. Mater.* **2005**, *11*, 284–286.
32. Mothana, S.; Grassot, J-M.; Hall, D. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 2883–2887.
33. Kwong, F. Y.; Klapars, A.; Buchwald, S. L. *Org. Lett.* **2002**, *4*, 581–584.

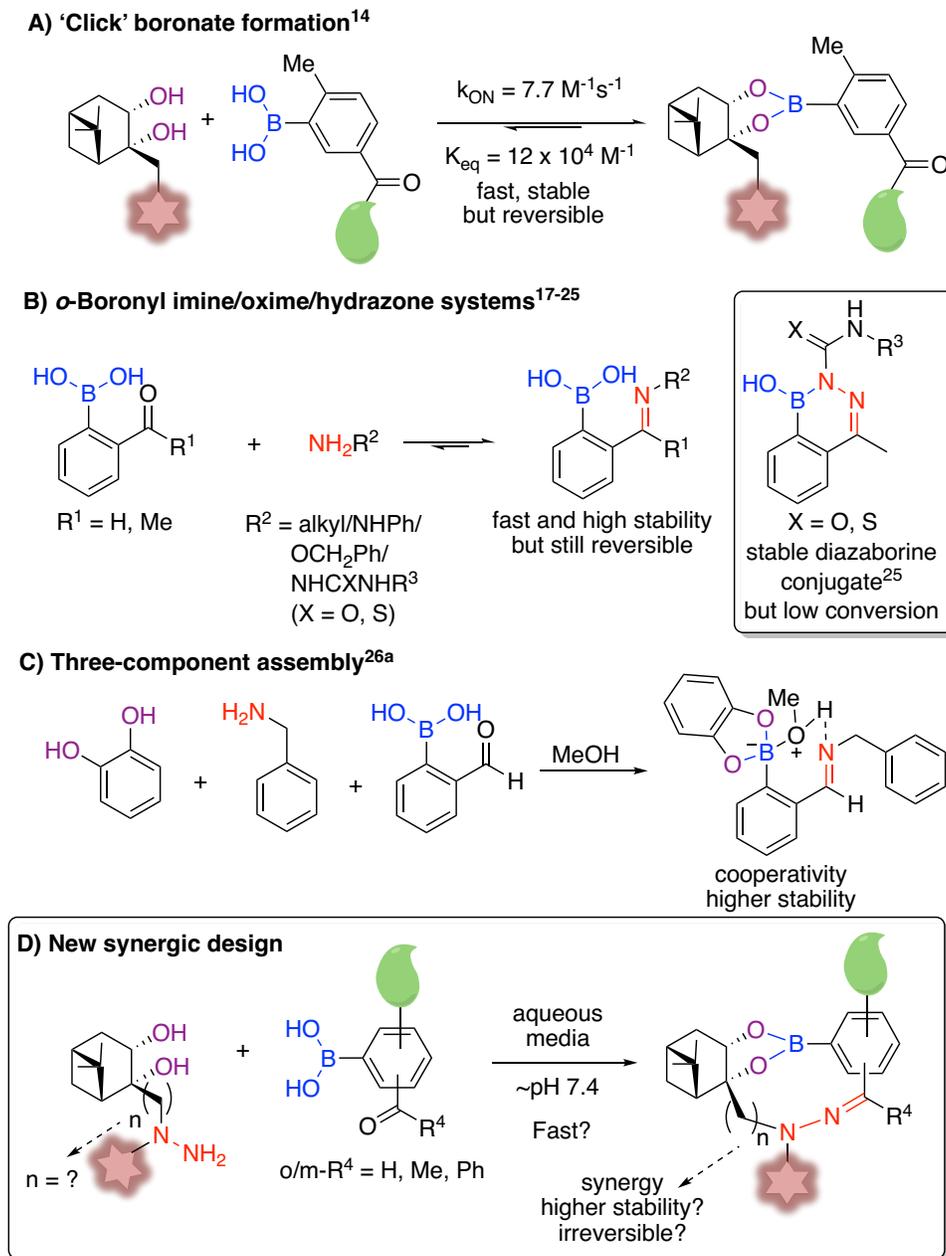
### 3 Chapter 3. Synergic ‘Click’ Boronate/Thiosemicarbazone System for Fast and Irreversible Bioorthogonal Conjugation in Live Cells

This chapter was modified from: Akgun, B.; Li, C.; Hao, Y.; Lambkin, G.; Derda, R.; Hall, D. G. Synergic ‘Click’ Boronate/Thiosemicarbazone System for Fast and Irreversible Bioorthogonal Conjugation in Live Cells. *J. Am. Chem. Soc.* **2017**, *139*, 14285–14291.

#### 3.1 Introduction

By exploiting reactions that are inert to native functional groups, bioorthogonal chemistry has become a powerful tool to chemically label proteins in living cells and help understand the roles of these proteins within their native environment.<sup>1-6</sup> Bioorthogonal chemistry has also been used in other applications such as installing post-translational modifications of proteins<sup>7</sup> or constructing antibody drug conjugates.<sup>8</sup> ‘Click’ chemistry<sup>9,10</sup> is central to the development of bioorthogonal reactions due to the following reasons: 1) it exhibits fast reactivity in an aqueous environment, 2) experiments can be performed at low concentrations (< 100  $\mu\text{M}$ ), thereby requiring a minimal amount of substrates and reducing any potential interference within living cells, 3) due to its high chemoselectivity, the reaction’s product is afforded in high yields without any side products.<sup>1,11,12</sup> Moreover, in view of dual-labeling applications, it is desirable for a click reaction to be orthogonal with other existing bioorthogonal reactions.<sup>3,13</sup> Even the most common bioorthogonal reactions display limitations such as slow rates, use of complex reagents, side-reactions or lack of mutual orthogonality due to possible cross reactivity.<sup>2,4,13</sup> In order to address these challenges, as explained in Chapter 2, we reported a novel bioorthogonal reaction based on a ‘click’ boronate formation, which enables fast dynamic ligation and high stability ( $k_{\text{ON}} = 7.7 \pm 0.5 \text{ M}^{-1}\text{s}^{-1}$ ,  $K_{\text{eq}} = 12 \times 10^4 \text{ M}^{-1}$ ) between nopoldiol derivatives and 2-methyl-5-carboxymethyl-phenylboronic acid (Figure 3-1A).<sup>14</sup> Although the stability of the resulting boronic ester is high, its reversibility may be detrimental in live cell imaging since washing steps to remove excess reagents might lead to undesirable cleavage. Therefore, our ongoing goal is to improve the stability of this system and ideally achieve a fast and irreversible boronate click reaction. It was hypothesized that such an advance could be achieved by exploiting a second, synergic interaction.

The most suitable additional covalent interaction must exhibit bioorthogonality and fast reactivity. Furthermore, this second interaction must involve reactive functional groups that must be small, non-cytotoxic, stable and easy to install. To this end, imine, hydrazone and oxime ligations attracted our attention. Hydrazone and oxime ligations have found many applications in bioorthogonal chemistry. Those methods, however, tend to suffer from slow rates at neutral pH and can also exhibit reversibility.<sup>1,15,16</sup> Fortunately, recent reports indicate that the presence of a boronic acid *ortho* to an aromatic aldehyde or acetylketone improves both the ligation rate and stability of the resulting imine, hydrazone and oxime product due to the existence of a dative B–N bond (Figure 3-1B).<sup>17-20</sup> For example, Gois and co-workers initially reported an iminoboronate formation between 2-formylbenzeneboronic acid (2-FBBA) or 2-acetylbenzeneboronic acid (2-ABBA) and lysine in aqueous media with 61 and 71% yield, respectively ( $K_{eq} = 10$  mM) (Scheme 1-4).<sup>17</sup> Gillingham and co-workers designed a rapid alkoxyiminoboronate condensation between 2-FBBA and *O*-benzylhydroxylamine at neutral pH ( $k_1 = \sim 11 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ ,  $k_{-1} = \sim 4.2 \times 10^{-5} \text{ s}^{-1}$ ,  $K_{eq} = \sim 2.6 \times 10^8 \text{ M}^{-1}$ ) (Scheme 1-5).<sup>19</sup> Meanwhile, Gao and co-workers demonstrated a rapid hydrazone ligation between phenyl hydrazine and 2-ABBA (Scheme 1-6).<sup>18</sup> Unfortunately, the reversibility of these reactions may limit their application in living cells. Bane and Gillingham recently disclosed the formation of 1,2-dihydro-1-hydroxy-2,3,1-benzodiazaborine as a stable product of click condensation between 2-FBBA and phenyl hydrazine (Scheme 1-7).<sup>21,22</sup> Phenyl hydrazine is, however, susceptible to oxidation, and it exhibits cytotoxicity. Furthermore, the bioorthogonality of 2-FBBA is problematic due to its high reactivity towards *N*-terminal cysteines (Scheme 1-8B and 1-9).<sup>23,24</sup> In this regard, Gao and co-workers demonstrated a new bioorthogonal process where a benign and stable (thio)semicarbazide forms a stable diazaborine conjugate with 2-ABBA (10 minutes, maximum 60% conversion at 50  $\mu\text{M}$  concentration for semicarbazide, maximum 27% conversion for thiosemicarbazide) (Scheme 1-8A).<sup>25</sup> According to the authors, this reaction stalls before completion likely due to the interaction of the product with the reactants. Remarkably, Anslyn and James reported a stable three-component assembly with 2-FBBA, catechol and benzylamine<sup>26a</sup> or *N*-hydroxylamines.<sup>26b</sup> The authors performed a structural and thermodynamic analysis of *ortho*-iminophenylboronate ester formation in a protic solvent, and they found that initial binding of either benzylamine or catechol to 2-FBBA improves the complexation of catechol and benzylamine, respectively (Figure 3-1C).<sup>26a</sup> In other words, binding of these two molecules to 2-FBBA becomes cooperative in a protic solvent.



**Figure 3-1:** A) Previously reported 'click' boronate formation system. B) Enhancement in the rate and stability of the imine/hydrazone/oxime ligation in the presence of a boronyl unit in *ortho*-position. C) Positive cooperativity and higher stability of a three-component amine/boronic acid/catechol system. D) New synergic design based on the optimal placement of a hydrazine or (thio)semicarbazide on the nopoldiol reagent, along with a carbonyl handle (*ortho* or *meta*) on the arylboronic acid.

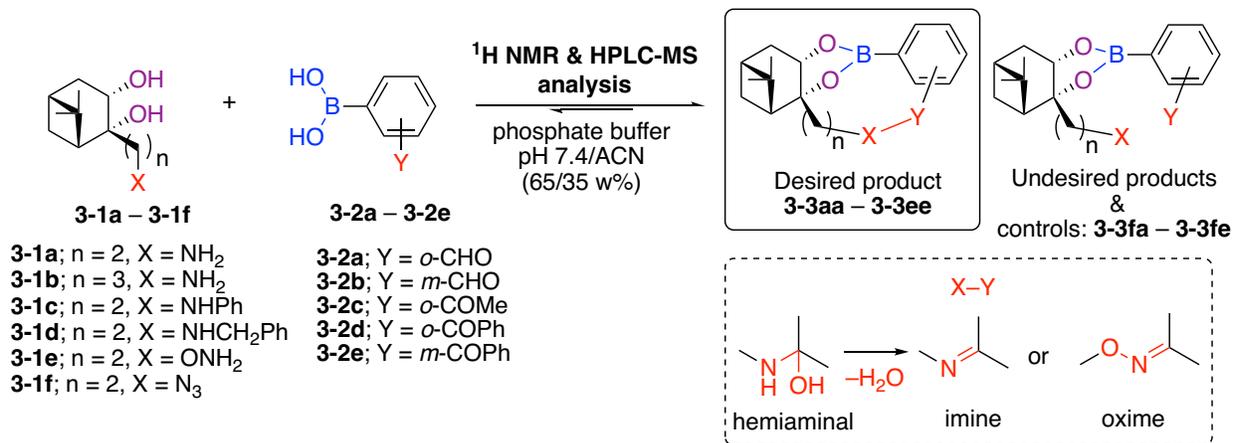
Altogether, despite these advances, the current systems suffer from issues such as stability and toxicity of phenyl hydrazine, incomplete conversion to the diazaborine product or the lack of bioorthogonality of 2-FBBA. Inspired by the iminoboronate concept of Anslyn and James, synergy was sought in our design based on the judicious placement of a hydrazine or (thio)semicarbazide unit on the nopoldiol reagent, along with a carbonyl handle (*ortho* or *meta*) on the arylboronic acid (Figure 3-1D). Therefore, this design was put to the test as a potentially bioorthogonal, fast and irreversible click reaction system with stable and benign reagents.

## **3.2 Results and discussion**

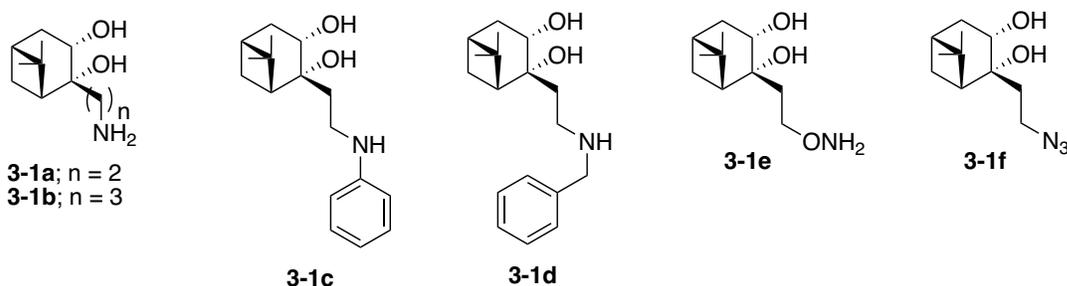
### **3.2.1 Optimization of the synergic system**

#### **3.2.1.1 Design and synthesis of nopoldiol derivatives with an amine/*O*-hydroxylamine unit**

In order to test the desired conjugation with selected boronic acid derivatives with a aldehyde/acetylketone/phenylketone group in *ortho*- or *meta*- position (**3-2a** – **3-2e**), nopoldiol derivatives (**3-1a** – **3-1f** with two different spacers ( $n = 2, 3$ ) and functional groups (amine/*O*-hydroxylamine) were designed and synthesized (Scheme 3-1).



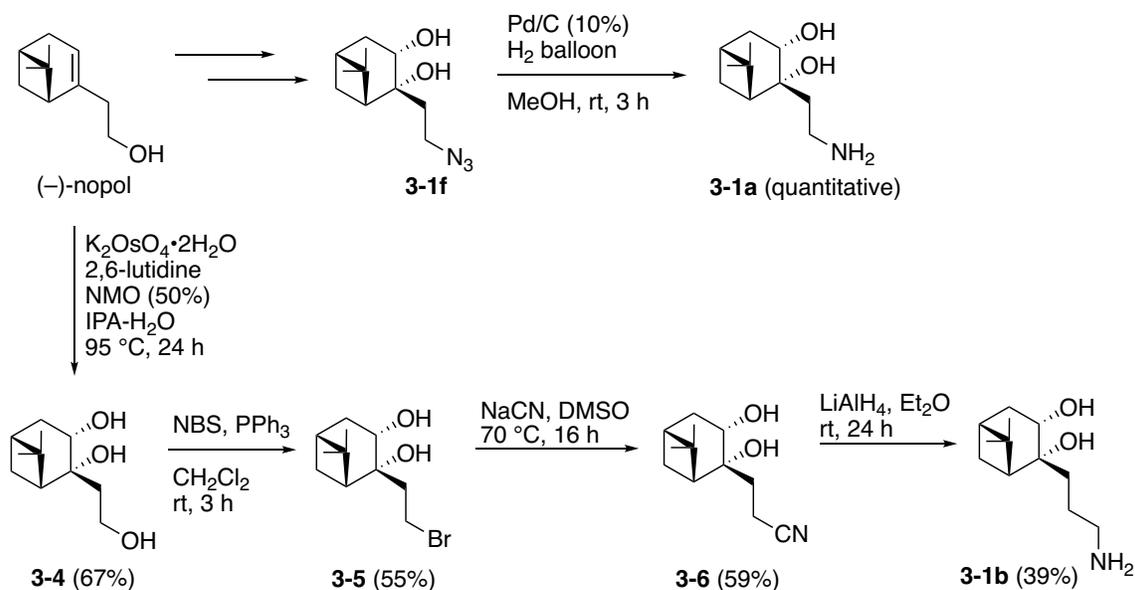
#### Diol derivatives:



**Scheme 3-1:** Desired conjugation test between nopoldiol amine derivatives (**3-1a – 3-1f**) and selected boronic acids with aldehyde/acetylketone/phenylketone (**3-2a – 3-2e**) in *ortho*- or *meta*-position.

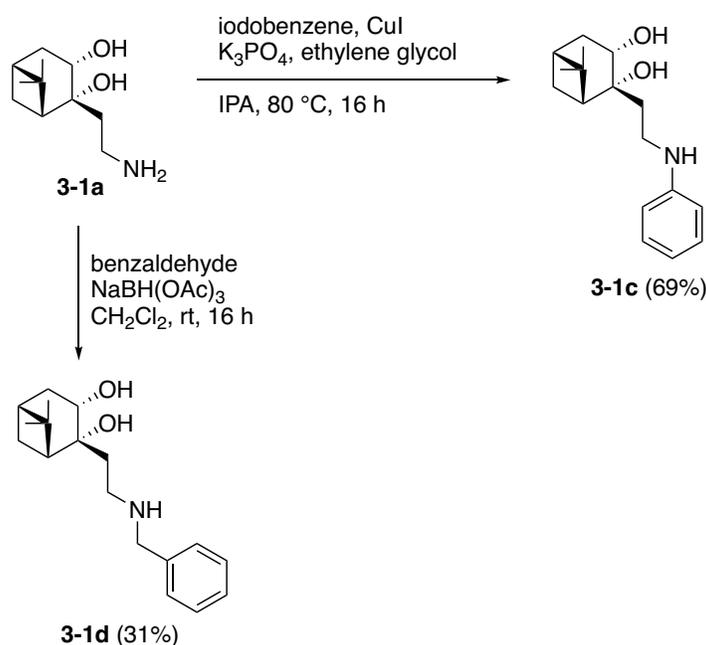
To begin with, nopoldiol amine derivative **3-1a** (n = 2) was synthesized from commercially available (–)-nopol, which was converted into nopol-azide by installing an azide unit, followed by an Upjohn dihydroxylation reaction to obtain nopoldiol azide **3-1f**. Diol **3-1f** was then subjected to hydrogenolysis conditions with palladium on carbon and hydrogen gas to provide diol **3-1a** (Scheme 3-2). To elongate the spacer from n = 2 to n = 3 and to synthesize nopoldiol amine derivative **3-1b** (n = 3), the nopoldiol cyano precursor **3-6** was initially accessed *via* (–)-nopol, which was converted into nopoldiol **3-4**, followed by bromination through an Appel reaction and cyanation of nopoldiol bromide **3-5** with sodium cyanide. Diol **3-5** could also be synthesized by bromination of (–)-nopol, followed by an Upjohn dihydroxylation; however the latter step was inefficient and numerous side products were formed that were not identified. Thus, I chose to follow the sequence of reactions described in Scheme 3-2. Next, the cyano

group of diol **3-6** was reduced to a primary amine group with lithium aluminum hydride to isolate diol **3-1b** ( $n = 3$ ).



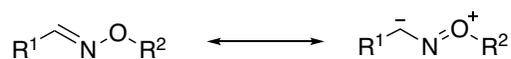
**Scheme 3-2:** Synthesis of nopoldiol amine derivatives **3-1a** and **3-1b**.

The synthesis of secondary amine nopoldiol derivatives **3-1c** and **3-1d** was performed from the diol precursor **3-1a** as shown in Scheme 3-3. Nopoldiol phenylamine **3-1c** was obtained with an Ullmann-type copper-catalyzed coupling of nopoldiol amine and phenyl iodide, a procedure described by Buchwald and co-workers.<sup>27</sup> Meanwhile, nopoldiol benzylamine **3-1d** was synthesized using a one-pot reductive amination procedure *via* the formation of nopoldiol iminium ion with benzaldehyde followed by reduction using the mild reducing agent, sodium triacetoxyborohydride.

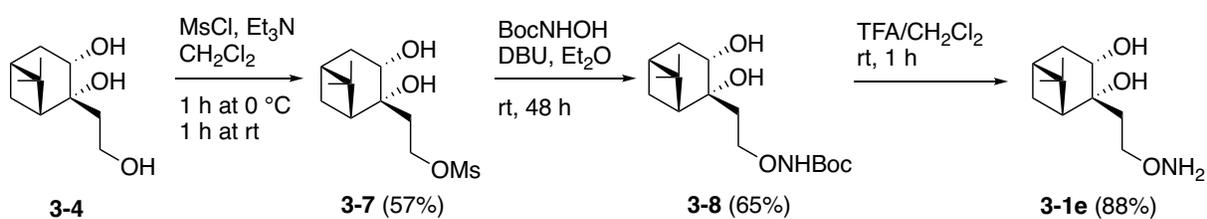


**Scheme 3-3:** Synthesis of nopoldiol phenylamine **3-1c** and nopoldiol benzylamine **3-1d**.

Apart from primary and secondary amine nopoldiol derivatives, the nopoldiol hydroxylamine derivative **3-1e** was also designed because *N*-hydroxylamine is expected to lead to a faster formed and more stable oxime conjugate compared to the corresponding primary and secondary amines.<sup>28</sup> *N*-hydroxylamines are more nucleophilic than amines due to the  $\alpha$ -effect, thus rendering *N*-hydroxylamines more reactive towards carbonyls. The greater stability of oximes was correlated with a possible resonance contribution of oxygen, which induces a higher negative charge on the carbon adjacent to nitrogen, forming a less electrophilic carbon (Scheme 3-4). Therefore, this less electrophilic and less reactive carbon likely maintains the oxime conjugate intact.<sup>28</sup> Another reason might be due to the repulsion of the lone pairs of nitrogen and oxygen, which is alleviated in the oxime conjugate.<sup>28</sup> *Tert*-butyloxycarbonyl (Boc) protected nopoldiol hydroxylamine **3-8** was prepared from nucleophilic addition of *N*-Boc-hydroxylamine to nopoldiol mesylate **3-7**, which was synthesized from diol **3-4** (Scheme 3-5). Then, the Boc group was removed in the presence of trifluoroacetic acid to obtain the desired diol **3-1e**.

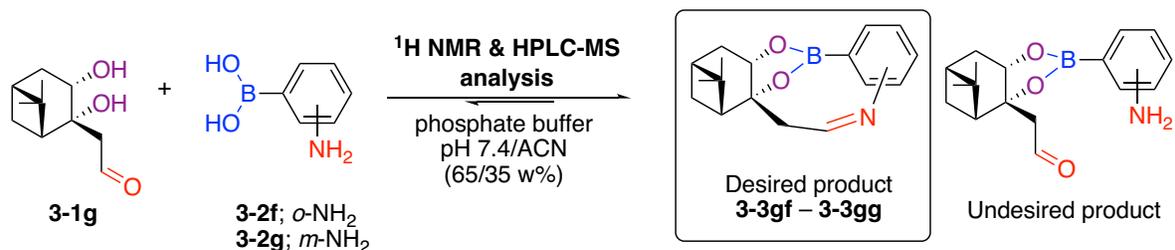


**Scheme 3-4:** Possible resonance forms of oximes.



**Scheme 3-5:** Synthesis of nopoldiol hydroxylamine **3-1e**.

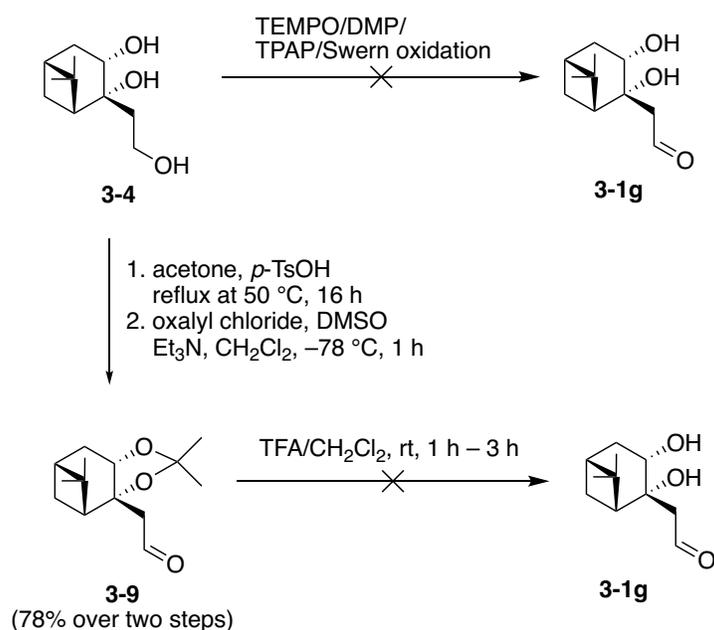
In this chapter, I focused mainly on the reactivity of nopoldiol amine/hydroxylamine (**3-1a** – **3-1e**) with aldehyde/acetylketone/phenylketone in the *ortho*- or *meta*- positions of the arylboronic acid (**3-2a** – **3-2e**). However, I also considered an alternative path where nopoldiol with an aldehyde unit **3-1g** could undergo an imine formation with an amine in the *ortho*- or *meta*- position to an arylboronic acid **3-2f** and **3-2g** (Scheme 3-6).



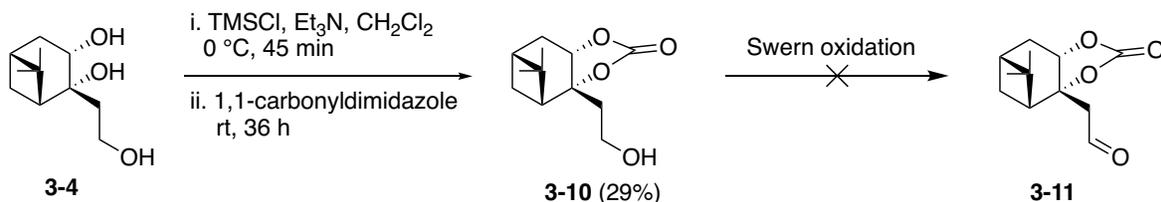
**Scheme 3-6:** Scheme of the desired conjugation test between nopoldiol aldehyde **3-1g** and boronic acids with amine in the *ortho*- or *meta*- position **3-2f** and **3-2g**.

In order to test this hypothesis, I made several attempts to synthesize diol **3-1g**; however, all these attempts were unsuccessful. At first, oxidation of nopoldiol's primary alcohol to an aldehyde was tested using common oxidation conditions, such as TEMPO, DMP, TPAP, and Swern oxidation (Scheme 3-7). All these conditions resulted in decomposition of the diol substrate **3-4** since various side products were observed. I thought that the 1,2-*cis* diol unit might interfere with the oxidation and cause many side products. Then, the protection of the 1,2-*cis* diols was performed with acetone and a catalytic amount of *p*-TsOH, affording a cyclic acetal nopol alcohol, which was directly and efficiently converted into nopol-acetal-aldehyde **3-9** via Swern oxidation (Scheme 3-7). Unfortunately, the deprotection step of the acetal unit under acidic conditions (TFA/CH<sub>2</sub>Cl<sub>2</sub>) to obtain the desired aldehyde **3-1g** resulted in the decomposition of **3-9** and did not lead to the desired product (Scheme 3-7). It was speculated that acidic deprotection conditions likely induce a rearrangement step;<sup>29</sup> therefore, basic deprotection

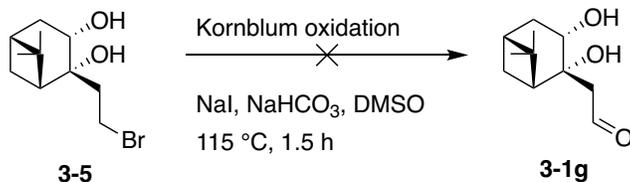
conditions would be ideal to access the desired product. Subsequently, carbonate formation of 1,2-*cis* diols, which requires basic conditions (KOH/MeOH) for its cleavage into its precursor diols, was performed to obtain nopol-carbonate **3-10** (Scheme 3-8). Yet, Swern oxidation did not afford the desired aldehyde formation. At last, an attempt was made through the Kornblum oxidation, which is an oxidation reaction between primary halides and DMSO forming an alkoxysulfonium ion that is then eliminated using a base to form an aldehyde. Diol **3-5** was subjected to Kornblum reaction conditions but the desired product was not observed either (Scheme 3-9). It is also possible that the desired product **3-1g** is unstable under even slightly acidic conditions, such as silicon dioxide, due to the possible cationic rearrangement products as the acid catalyzed transformation of pinane derivatives is also described in the literature.<sup>29</sup>



**Scheme 3-7:** Attempts for the synthesis of nopoldiol aldehyde **3-1g**.



**Scheme 3-8:** Another strategy for the synthesis of nopoldiol aldehyde **3-1g**.



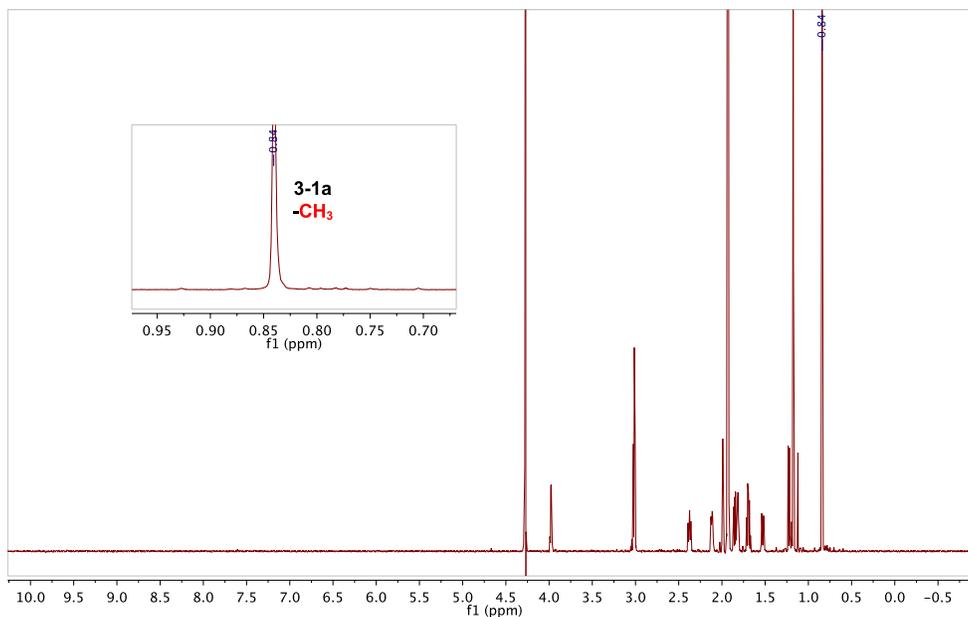
**Scheme 3-9:** Attempted Kornblum oxidation for the synthesis of nopoldiol aldehyde **3-1g**.

### 3.2.1.2 Conjugation studies – $^1\text{H}$ NMR analysis

Boronate formation and the occurrence of synergic interaction (imine/oxime conjugation) were investigated *via*  $^1\text{H}$  NMR analysis performed in 0.05 M of  $\text{D}_2\text{O}$  potassium phosphate buffer/ACN- $\text{d}_3$  (65/35 w%, pD 7.4). Individual solutions of diol derivatives (**3-1**, 2 mM or 1 mM) and selected boronic acids (**3-2**, 2 mM or 1 mM) were prepared in this solvent system. Equivalent volumes of diol and boronic acid solutions were mixed in a NMR tube to reach a final concentration of 1 mM or 0.5 mM. After the NMR samples were incubated for 2 h and 24 h at room temperature, their % conversions (**3-3/3-1**) were monitored *via* integral ratio of  $-\text{CH}_3$  product/ $-\text{CH}_3$  substrate (the gem dimethyl protons) (Figure 3-2 and Table 3-1). When the  $^1\text{H}$  NMR spectrum of free diol **3-1** was compared with the spectrum of conjugate **3-3**, a chemical shift was observed for the  $-\text{CH}_3$  of diol **3-1**. This chemical shift value was different for each diol. For instance, the  $-\text{CH}_3$  of diol **3-1a** appeared at 0.84 ppm and the  $-\text{CH}_3$  of conjugate **3-3aa** showed two distinct chemical shifts: 0.79 ppm (**3-3aa**• $\text{H}_2\text{O}$ ) and 0.89 ppm (**3-3aa**). Moreover, the  $-\text{CH}=\text{O}$  resonance of **3-2a** appeared at 9.94 ppm, and this peak shifted to 5.91 ppm (**3-3aa**• $\text{H}_2\text{O}$ ) and 8.28 ppm (**3-3aa**). Thus, the  $^1\text{H}$  NMR method could confirm imine formation for aldehyde **3-2a**. In case of boronic acids **3-2c** or **3-2d**, imine/oxime formation was further investigated and confirmed by HPLC-MS analysis.

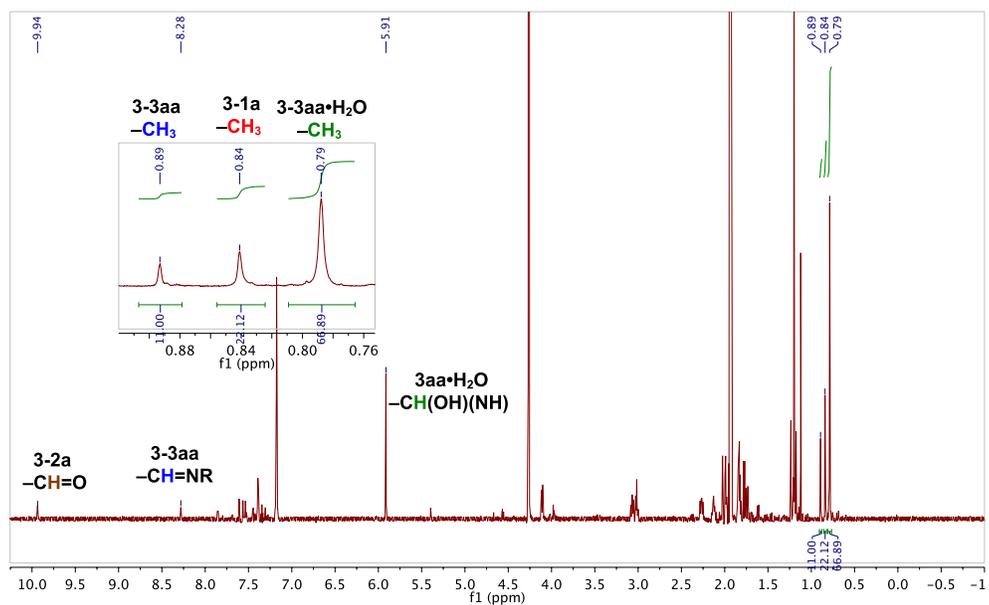
A)

Only diol **3-1a**-<sup>1</sup>H-NMR study of boronate and imine formation  
 699.764 MHz H1 PRESAT in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm), temp 27.5 C -> actual temp = 27.0 C, coldid probe



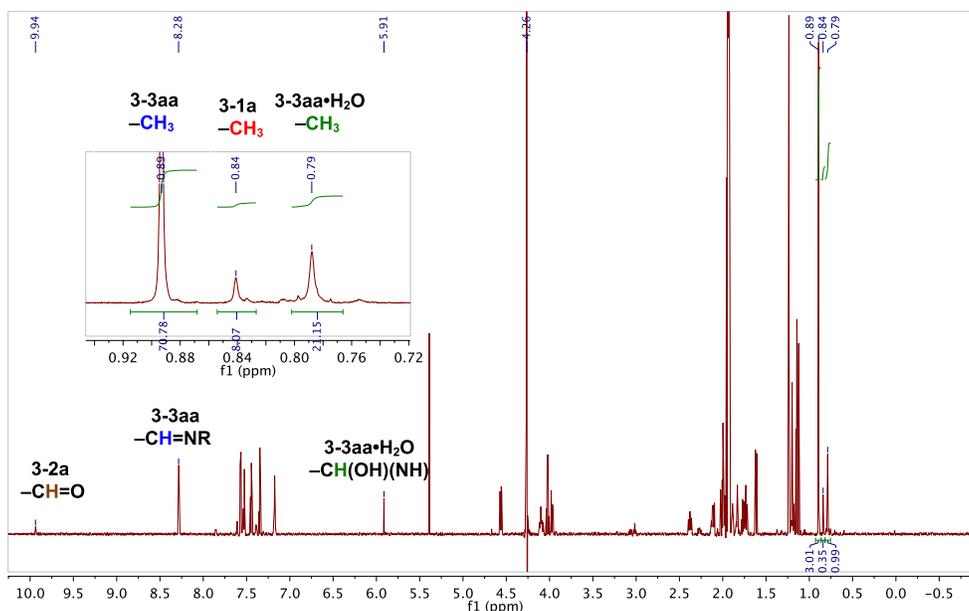
B)

Boronate **3-3aa-2** h-<sup>1</sup>H-NMR study of boronate and imine formation  
 699.764 MHz H1 PRESAT in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm), temp 27.5 C -> actual temp = 27.0 C, coldid probe



C)

Boronate **3-3aa**-24 h-<sup>1</sup>H-NMR study of boronate and imine formation  
699.764 MHz H1 PRESAT in d<sub>2</sub>O (ref. to external acetone @ 2.225 ppm), temp 27.5 C -> actual temp = 27.0 C, coldid probe

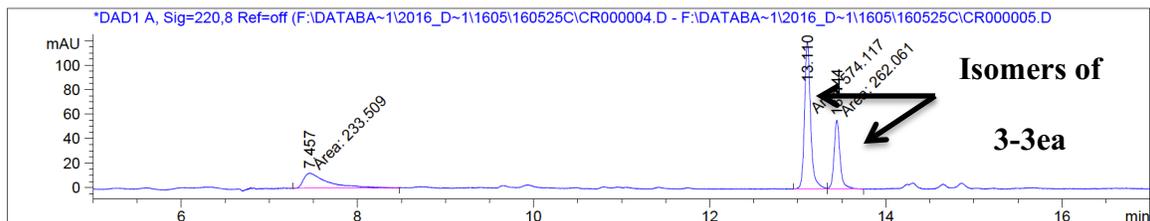


**Figure 3-2:** Representative <sup>1</sup>H NMR spectra of optimization studies. A) Only diol **3-1a**. B) Formation of conjugate **3-3aa** in 2 h. C) Formation of conjugate **3-3aa** formation in 24 h.

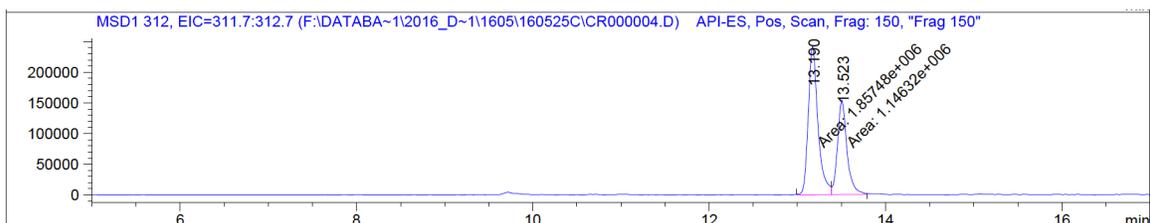
### 3.2.1.3 Conjugation studies – HPLC-MS analysis

The HPLC-MS study was performed only for a few representative reactions in order to confirm that the imine/oxime formation occurred. The boronate formation and synergic interaction with the imine or oxime group were also investigated by HPLC-MS in 0.025 M of potassium phosphate buffer/ACN (65/35 w%, pH 7.4); individual solutions of diol derivatives (**3-1a** and **3-1e**, 1 mM) and selected boronic acids (**3-2a**, **3-2c** and **3-2d**, 1 mM) were prepared in this solvent system. Equivalent volumes of diol **3-1** and boronic acids **3-2** were mixed in a HPLC vial to reach a final concentration of 0.5 mM. After the mixtures were incubated for 24 h at room temperature, they were monitored at UV 220 nm by HPLC-MS analysis (Figure 3-3 and Table 3-1).

A)

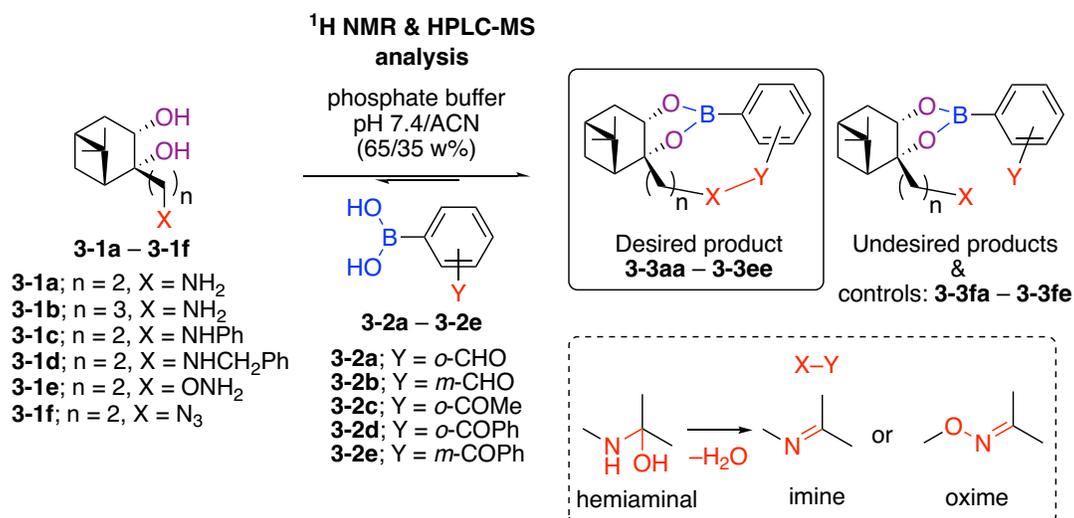


B)



**Figure 3-3:** Representative HPLC-MS chromatogram for optimization studies. A) Representative HPLC chromatogram for boronate and oxime formation between **3-1e** and **3-2a**. B) LC-MS trace of conjugate **3-3ea**.

**Table 3-1:** Results of initial conjugation attempts of diol derivatives (**3-1a – 3-1f**) with selected boronic acids (**3-2a – 3-2e**).



Entry (3-1a – 3-1f)	Boronic acid (3-2a – 3-2e)	Product and ratio <sup>[a]</sup> (2 h) (1 mM) 3-3/3-1	Product and ratio <sup>[a]</sup> (24 h) (1 mM) 3-3/3-1	Imine/ oxime formation?
1 (3-1a)	3-2a	<b>3-3aa/3-3aa•H<sub>2</sub>O/3-1a:</b> 11/67/22	71/21/8	Yes <sup>[b]</sup>
2 (3-1a)	3-2b	<b>3-3ab/3-1a:</b> 79/21	79/21	No
3 (3-1a)	3-2c	<b>3-3ac/3-1a:</b> 43/57	43/57	Yes <sup>[b]</sup>
4 (3-1a)	3-2d	<b>3-3ad/3-1a:</b> -	68/22	-
5 (3-1a)	3-2e	<b>3-3ae/3-1a:</b> 82/18	83/17	-
6 (3-1b)	3-2a	<b>3-3ba/3-1b:</b> 78/22	78/22	No
7 (3-1b)	3-2b	<b>3-3bb/3-1b:</b> 81/19	81/19	No
8 (3-1b)	3-2c	<b>3-3bc/3-1b:</b> 80/20	80/20	-
9 (3-1b)	3-2d	<b>3-3bd/3-1b:</b> 93/7	93/7	-
10 (3-1b)	3-2e	<b>3-3be/3-1b:</b> 84/16	81/19	-
11 (3-1c)	3-2a	<b>3-3ca/3-1c:</b> 68/32	68/32	No
12 (3-1c)	3-2b	<b>3-3cb/3-1c:</b> 74/26	73/27	No
13 (3-1c)	3-2c	<b>3-3cc/3-1c:</b> 67/33	67/33	-
14 (3-1c)	3-2d	<b>3-3cd/3-1c:</b> 89/11	84/16	-
15 (3-1c)	3-2e	<b>3-3ce/3-1c:</b> 74/26	74/26	-
16 (3-1d)	3-2a	<b>3-3da/3-1d:</b> 86/14	86/14	No
17 (3-1e)	3-2a	<b>3-3ea/3-1e:</b> 100/0	100/0	Yes <sup>[b]</sup>
18 (3-1e)	3-2b	<b>3-3eb/3-1e:</b> 78/22	78/22	No
19 (3-1e)	3-2c	<b>3-3ec/3-1e:</b> 100/0	100/0	Yes <sup>[b]</sup>
20 (3-1e)	3-2d	<b>3-3ed/3-1e:</b> 100/0	100/0	Yes <sup>[b]</sup>

Entry ( <b>3-1a</b> – <b>3-1f</b> )	Boronic acid ( <b>3-2a</b> – <b>3-2e</b> )	Product and ratio <sup>[a]</sup> (2 h) (1 mM)	Product and ratio <sup>[a]</sup> (24 h) (1 mM)	Imine/ oxime formation?
		<b>3-3/3-1</b>	<b>3-3/3-1</b>	
21 ( <b>3-1e</b> )	<b>3-2e</b>	<b>3-3ee/3-1e</b> : 81/19	81/19	-
22 ( <b>3-1f</b> )	<b>3-2a</b>	<b>3-3fa/3-1f</b> : 73/27	73/27	No
23 ( <b>3-1f</b> )	<b>3-2b</b>	<b>3-3fb/3-1f</b> : 75/25	75/25	No
24 ( <b>3-1f</b> )	<b>3-2c</b>	<b>3-3fc/3-1f</b> : 74/26	74/26	No
25 ( <b>3-1f</b> )	<b>3-2d</b>	<b>3-3fd/3-1f</b> : 86/14	86/14	No
26 ( <b>3-1f</b> )	<b>3-2e</b>	<b>3-3fe/3-1f</b> : 78/22	79/21	No

[a] The % product (boronate/imine/oxime) conversion was determined by <sup>1</sup>H NMR analysis. [b] Imine/oxime formation was also confirmed by HPLC-MS analysis.

Not surprisingly, it was found that *ortho*-substituted aryl boronic acids (**3-2a**, **3-2c**, **3-2d**) were more suitable to form the desired product. Notably, when the conjugation of **3-1a** and **3-2a** is compared with the conjugation of **3-1a** and **3-2b**, formation of an imine along with an enhancement in the product conversion from 79 to 92% was clearly observed for the reaction of **3-1a** and **3-2a** (Table 3-1, entry 1 and/vs 2). However, *m*-substituted aryl boronic acids **3-2b** and **3-2e** could only form the boronate product in moderate yields, without any imine/oxime being formed (Table 3-1, entry 2, 5, 18 and 21).

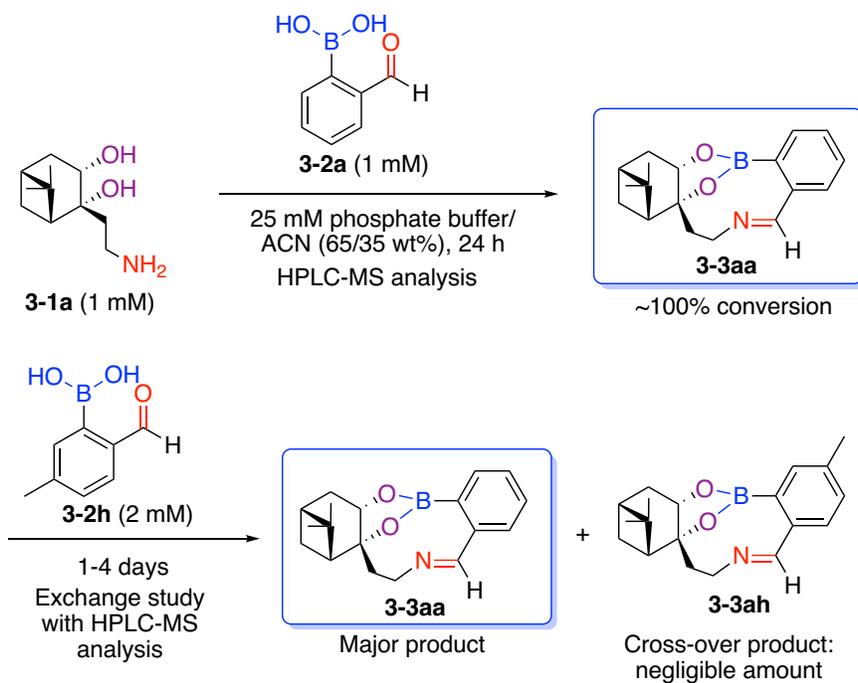
Boronic acid **3-2a** and diol **3-1a** could initially form the medium-sized ring hemiaminal **3-3aa•H<sub>2</sub>O**, which, according to the diagnostic –HC=N– resonance in <sup>1</sup>H NMR, slowly converted into imine **3-3aa** over 24 h in 1.0 mM concentration at room temperature (Table 3-1, entry 1). A comparison of **3-3aa** with **3-3fa** indicated that the presence of a second interaction (i.e., imine formation) clearly improved the conversion from 73 to 92% (Table 3-1, entry 1 and/vs 22). Moreover, boronic acid **3-2c** also formed an imine with **3-1a**, albeit with a lower conversion (43%) even after 24 h (Table 3-1, entry 3). On the other hand, **3-1b** with an additional –CH<sub>2</sub> spacer unit underwent only boronate formation with **3-2a** with no imine or hemiaminal formation observed (Table 3-1, entry 6). This result clearly indicated that **3-1a**, unlike **3-1b**, featured the right geometry and size that favor both boronate and imine formation. Diol **3-1a**

contains a simple primary amine, which could be further functionalized to include a reporter tag, but the resulting secondary amines would render iminium formation more difficult. On the other hand, hemiaminal formation may still occur, and hopefully, be sufficient to form a stable conjugate. Thus, **3-1c** and **3-1d** were tested, however, no hemiaminal formation was observed as evidenced by the lack of the corresponding  $-RCH(OH)(NHR')$  signal at  $\sim 6$  ppm (Table 3-1, entry 11 and 16) and, instead, only boronate formation was observed. Next, *O*-hydroxylamine **3-1e** was designed and tested with **3-2a**, **3-2c** and **3-2d**. All of these boronic acids formed the desired boronate-oxime conjugate, in a mixture of two oxime E/Z isomers, with full conversion in a very short time (Table 3-1, entry 17, 19, 20).

#### 3.2.1.4 Stability study of **3-3aa** via a crossover experiment

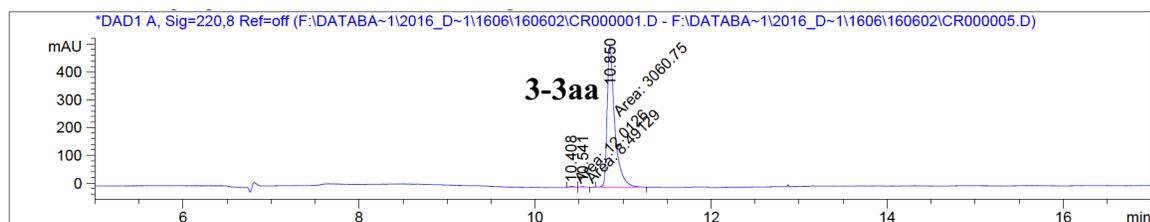
Before moving forward on the analysis of the bioconjugate chemistry, I tested the stability of the conjugate **3-3aa** via HPLC-MS analysis. Individual solutions of diol **3-1a** (2.0 mM) and boronic acid **3-2a** (2.0 mM) were prepared in 0.025 M of potassium phosphate buffer/ACN (65/35 w%, pH 7.4). Equivalent volumes of diol **3-1a** and boronic acid **3-2a** were mixed in a HPLC vial to reach a final concentration of 1.0 mM (Scheme 3-10). After they were incubated for 24 h at room temperature, the observed peaks were analyzed and characterized by HPLC-MS (Figure 3-4). Then, the boronic acid **3-2h** (80 mM, 5  $\mu$ l) was introduced into the reaction mixture at a final concentration of 2.0 mM. Afterwards, the samples were incubated for 1 d and 4 d at room temperature; the corresponding peaks were analyzed and characterized by HPLC-MS as shown in Figure 3-5.

Satisfactorily, the major product **3-3aa** was preserved even after 4 d and a negligible proportion of the exchange product **3-3ah** was observed (Scheme 3-10, Figure 3-4 and 3-5). This small amount of **3-3ah** may have formed due to a small amount of unreacted **3-1a** in the reaction mixture, rather than exchange of 'parts' released from the hydrolysis of **3-3aa**.

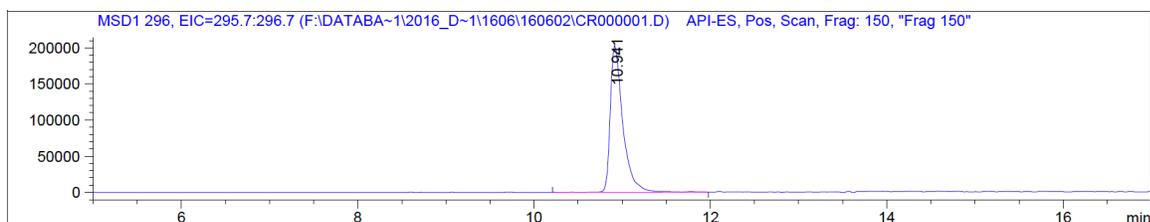


**Scheme 3-10:** Stability test of **3-3aa** via a cross-over experiment.

A)

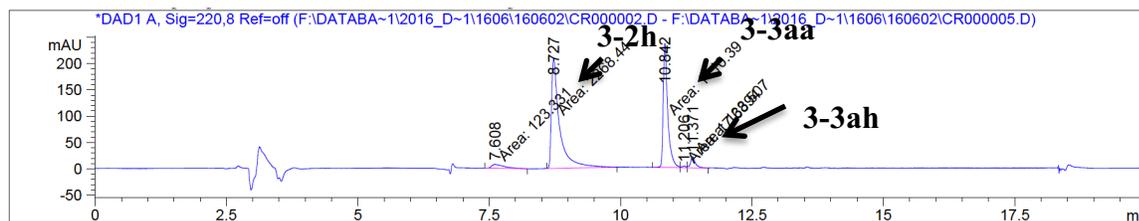


B)

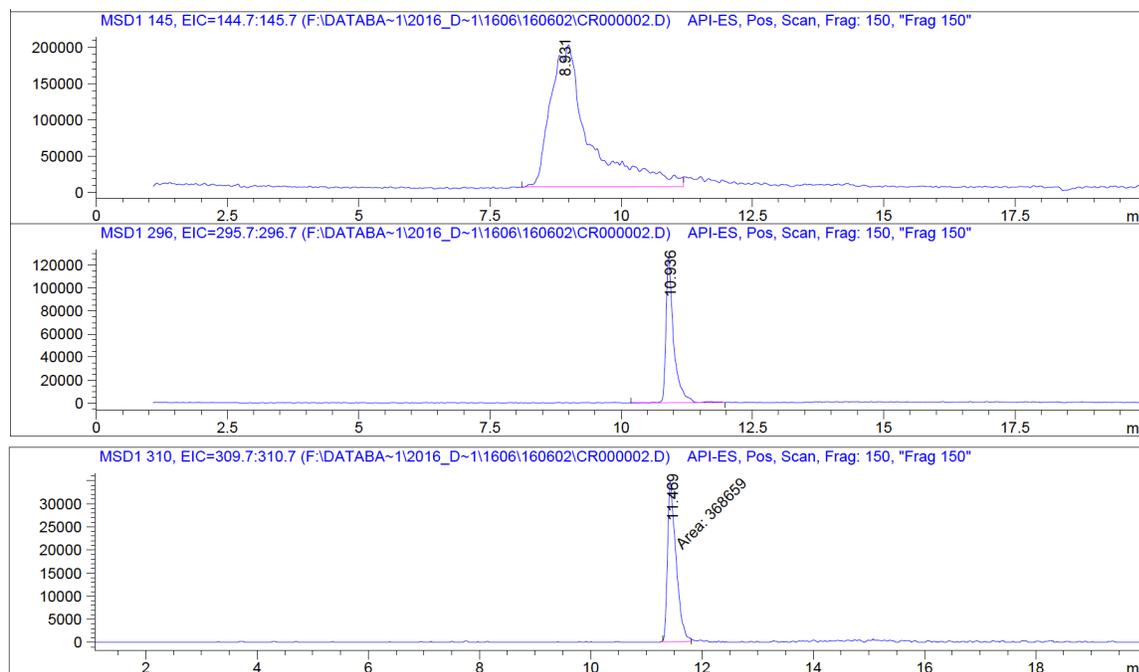


**Figure 3-4:** HPLC-MS chromatogram of reaction of **3-1a** and **3-2a** after 24 h. A) HPLC chromatogram of conjugate **3-3aa**. B) LC-MS trace of **3-3aa**.

A)



B)

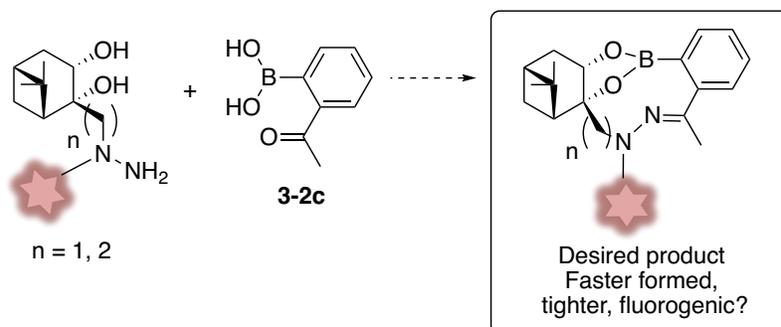


**Figure 3-5:** HPLC-MS chromatogram of reaction between **3-1a** and **3-2a**, followed by the addition of **3-2h** after 24 h. A) HPLC chromatogram of the reaction mixture after 24 h. B) LC-MS traces of the reaction mixture after 24 h.

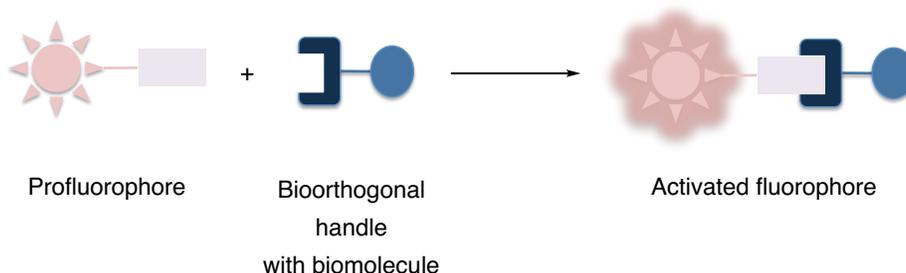
### 3.2.2 Design of hydrazine-functionalized fluorogenic nopoldiol derivatives

Based on the successful formation and stability of **3-3aa**, these initial studies led us to choose the most suitable spacer length for the ideal diol as  $n = 1$  and 2 (Scheme 3-11). Regarding the choice of boronic acid, *o*-formylated **3-2a** was no longer considered due to its reported side reactivity with *N*-terminal cysteines.<sup>23,24</sup> Moreover, because **3-2c** exhibits lower hydrophobicity than **3-2d**, **3-2c** was selected as the optimal boronic acid (Scheme 3-11). Next, our focus was to synthesize such a diol derivative with a disubstituted hydrazine unit as shown in Scheme 3-11. Hydrazine units are intrinsically more reactive towards carbonyl groups forming a more stable

hydrazone ligation compared to the imine product between an amine and a carbonyl unit. The substituted hydrazine unit was ideally aimed to be a fluorogenic dye or a fluorophore. Fluorogenic reactions have been invaluable tools in bioorthogonal chemistry, since these reactions allow the visualization of biomolecules without the need to remove the excess unreacted probes (Figure 3-6).<sup>30,31</sup> Thus, these probes prevent background labeling, which is generally an issue when excess fluorescent dye is required.



**Scheme 3-11:** Design of fluorogenic or fluorescently labeled diol derivatives towards faster and tighter boronate formation.

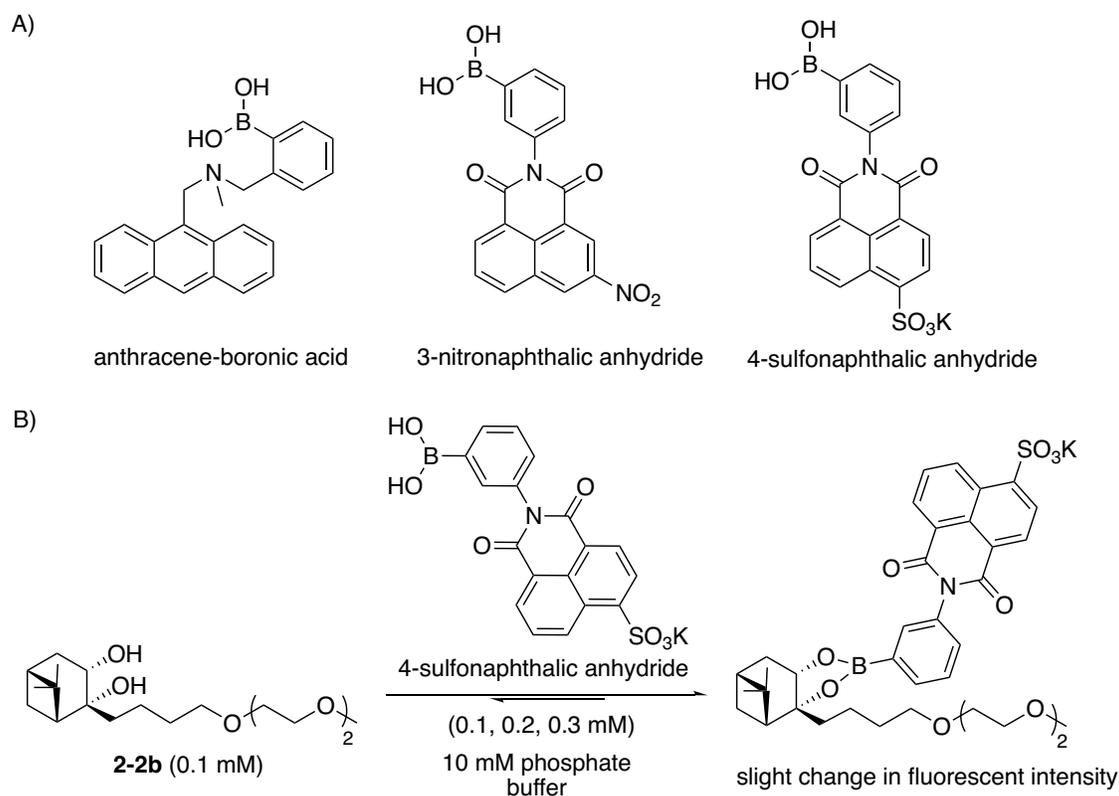


**Figure 3-6:** Schematic example of a fluorogenic bioorthogonal biolabeling reaction.

Various ways to design a fluorogenic system have been described in the literature and the general methods are based upon the loss of a quencher, induced electronic changes, or extension of conjugation.<sup>30-32</sup> Even though the loss of a quencher, which means the bond cleavage between a quencher and a fluorophore, is one of the most successful approaches, bioorthogonal chemistry cannot usually benefit from it due to its intrinsic use in bond formation instead of bond cleavage. Induced electron changes and extension of conjugation strategies are more suitable in the context of bond forming bioorthogonal reactions. For example, induced electronic change approaches are very efficient for ‘push-pull’ fluorescence molecules, which carry electron donating and electron

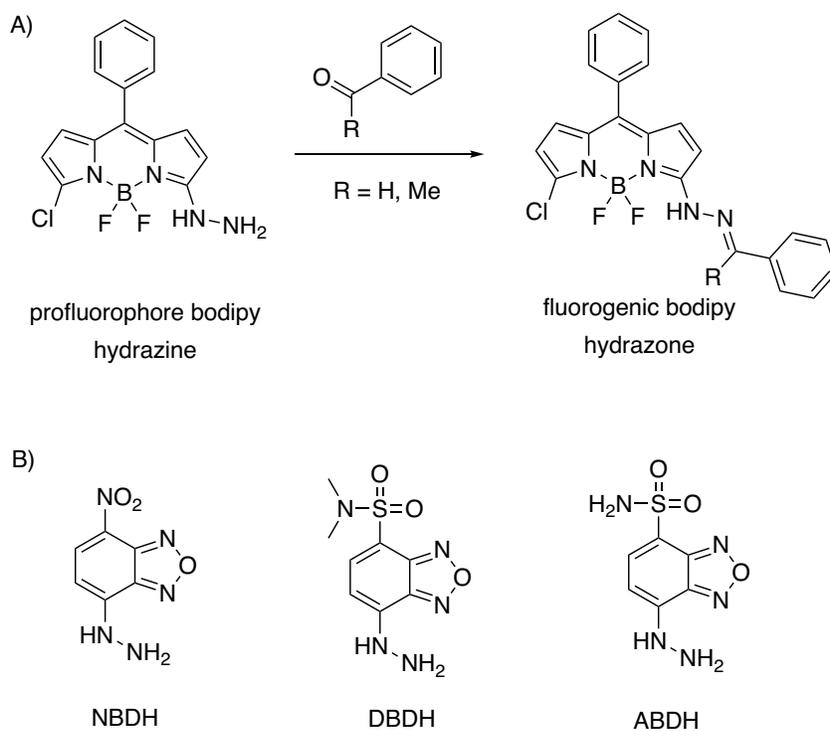
withdrawing groups. Even slight electronic changes of these molecules upon bioorthogonal ligation might lead to significant shifts on the excited state energy levels and also on the fluorescence properties. Furthermore, by adjusting the conjugation of the fluorophore, novel fluorogenic systems can be designed from weak or non-fluorescent bioorthogonal components upon bioorthogonal conjugation.

Considering our bioorthogonal method based on the hydrazone/boronate conjugation between nopoldiol hydrazine and 2-ABBA derivatives, either boronate or hydrazone formation can be used to introduce a fluorogenic property into the system. At first, a fluorogenic boronate was attempted. Several research groups have been active on the design of fluorogenic boronates mostly towards a saccharide recognition.<sup>33</sup> Boronic acid sensors, such as anthracene-boronic acid and naphthalic anhydride derivatives attracted our attention among many examples due to their relatively small size and synthetic accessibility (Scheme 3-12A).<sup>34</sup> Particularly, the 4-sulfonaphtalic acid sensor was preferred due to its high water solubility.<sup>34c</sup> Heagy and co-workers observed an increase in fluorescence intensity upon boronic acid-glucose complexation with dual emission response at 400 nm and 474 nm.<sup>34c</sup> The conjugation of this sensor with previously synthesized nopoldiol **2-2b** was monitored *via* a fluorescence plate reader at various equivalents of nopoldiol (1–3 equiv) (Scheme 3-12B). However, only a slight increase (maximum ~1.8 fold at 300:100  $\mu\text{M}$  concentration of nopoldiol:boronic acid) in fluorescence intensity of the corresponding boronate conjugate was detected at 360 nm (maximum excitation wavelength) and 470 nm (maximum emission wavelength) compared to the fluorescence intensity of the boronic acid solution alone (100  $\mu\text{M}$ ). As explained by Heagy and co-workers, the binding affinity is not directly proportional to the detected optical change, and the conformational restriction in boronate complex might have an impact on the excited state, thus leading to an insignificant change in fluorescence intensity.<sup>33c</sup>



**Scheme 3-12:** Fluorogenic boronate formation. A) Selected examples of fluorogenic boronic acid sensors. B) Fluorogenicity study of diol **2-2b** and 4-sulfonaphthalic anhydride complex.

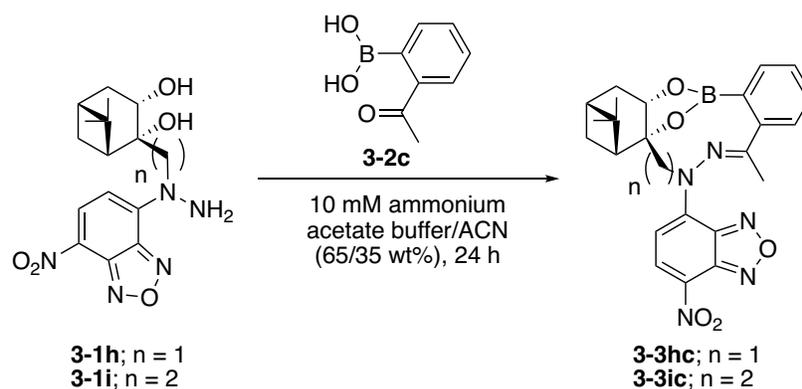
Next, the fluorogenic hydrazone strategy was investigated. Literature precedents suggest a number of available fluorogenic molecules, which have been developed for labeling of aldehydes and ketones, especially ones with a boron-dipyrromethene or nitrobenzoxadiazole backbone (Scheme 3-13).<sup>31,35</sup> Bane and co-workers reported many fluorogenic 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (bodipy) hydrazine compounds, which show enhancements in quantum yield (49–54 fold) and higher absorption/emission wavelengths (~37 nm shift) upon hydrazone ligation with aromatic carbonyls (Scheme 3-13A).<sup>35a</sup> The fluorogenic bodipy hydrazones were also applied in the labeling of the protein  $\alpha$ -tubulin, which was modified with the unnatural amino acid, 3-formyltyrosine.<sup>35b</sup>



**Scheme 3-13:** Fluorogenic hydrazone ligation. A) Fluorogenic bodipy hydrazone formation. B) Fluorogenic NBD hydrazine derivatives.

Even though bodipy hydrazines are great designs, fluorogenic nitrobenzoxadiazole derivatives are more attractive due to their small size and synthetic accessibility (Scheme 3-13B). Additionally, these molecules exhibit high fluorescence emission quantum yields, and long excitation and emission wavelengths.<sup>36</sup> These profluorophore nitrobenzoxadiazole derivatives have nucleophilic hydrazines or amino-oxy groups, thus resulting in hydrazones or oximes with enhanced fluorescence emission. Reported fluorogenic hydrazino-NBD derivatives include: 7-hydrazino-4-nitrobenzo-2-oxa-1,3,-diazole (NBDH),<sup>37</sup> 4-(N,N-dimethylaminosulphonyl)-7-hydrazino-benz-2,1,3-oxadiazole (DBDH),<sup>38</sup> 4-4-aminosulphonyl-7-hydrazino-2,1,3-benzoxadiazole (ABDH) (Scheme 3-13B).<sup>38</sup> Particularly, NBDH caught our attention due to its relatively simple synthetic accessibility and its reported applications. Frei and co-workers reported that NBDH exhibits higher fluorescence intensity upon hydrazone formation with aldehyde/ketone derivatives.<sup>37</sup> The authors reported a detailed method to quantify carbonyl compounds with NBDH *via* HPLC fluorescent detection.<sup>37</sup> This dye and its other derivatives have found applications in labeling of carbonyl containing compounds and more interestingly in the labeling of biomolecules.<sup>31,37,38</sup> Cairo and co-workers recently reported NBDH fluorophore

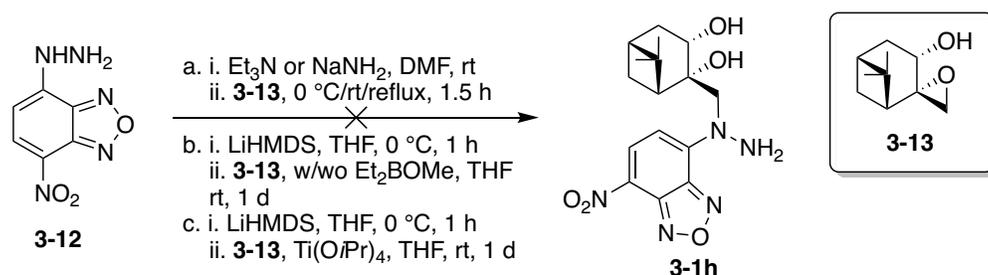
for the selective labeling of sialic acid for both *in vitro* and *in vivo* studies.<sup>35c</sup> The NBDH dye was shown to be fluorogenic with a 2.0-fold increase upon hydrazone formation over background. This observed increase in fluorescence intensity for an aliphatic aldehyde derivative (oxidized sialic acid) is not very promising. Aromatic arylboronic acids, however, might induce a higher fluorescence intensity compared to aliphatic carbonyls, due to the extended conjugation of aromatic groups. Therefore, I decided to synthesize nopoldiol NBD hydrazone derivatives ( $n = 1, 2$ ) to test this hypothesis (Scheme 3-14).



**Scheme 3-14:** The conjugation of nopoldiol NBD hydrazine designs **3-1h** and **3-1i** with boronic acids **3-2c**.

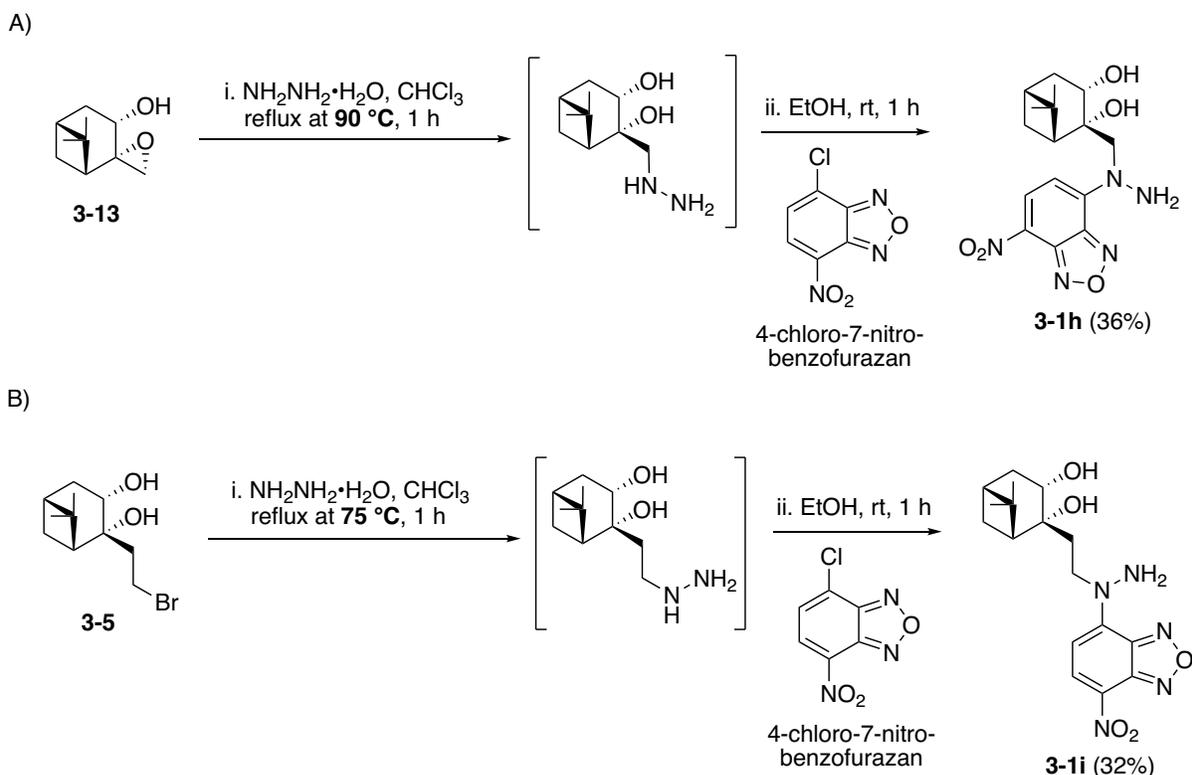
Our first approach was a nucleophilic substitution between NBDH hydrazine **3-12** and an epoxide **3-13** (Scheme 3-15). In order to synthesize target molecules **3-1h**, the nopoldiol precursor, epoxy alcohol **3-13** was synthesized based on literature examples.<sup>39</sup> The NBDH hydrazine **3-12** was easily synthesized *via* an  $S_NAr$  of hydrazine monohydrate with 4-chloro-7-nitrobenzofurazan.<sup>35c</sup> Taking the pKa value of phenylhydrazine (28.8) as a reference, the NBDH hydrazine must be intrinsically more acidic due to the presence of an electron withdrawing nitro unit at the *para* position as well as the extended conjugation. Therefore, an abstraction of NBDH acidic proton was attempted in either mildly basic  $\text{Et}_3\text{N}$  or reactive  $\text{NaNH}_2$  at room temperature. As part of a one-pot reaction, the epoxy alcohol **3-13** was then introduced into the reaction mixture to obtain the desired molecule **3-1h**. However, the reaction did not proceed even at a high temperature (90 °C) and the starting materials appeared unreacted. Afterwards, I considered to increase the reactivity of the epoxide using a Lewis-Acid promoted strategy through the coordination of epoxide oxygen with the adjacent secondary alcohol.<sup>40</sup> Both  $\text{Et}_2\text{BOMe}$  and

Ti(O*i*Pr)<sub>4</sub> were separately tested in the second step along with the epoxy alcohol **3-13**. Due to the high affinity of nopoldiol towards boronyl compounds, the reaction performed with Et<sub>2</sub>BOMe reagent provided the ethyl boronate derivative of the desired target. Unfortunately, the reaction efficiency was very low and deprotection of this boronyl unit was not straightforward. Next, the protection of NBDH hydrazine with acetone followed by S<sub>N</sub>2 chemistry was also performed, but, once more, the reaction did not proceed. This likely resulted due to the steric hindrance of the protected NBDH.



**Scheme 3-15:** Synthesis trials of nopoldiol NBD hydrazine design **3-1h**.

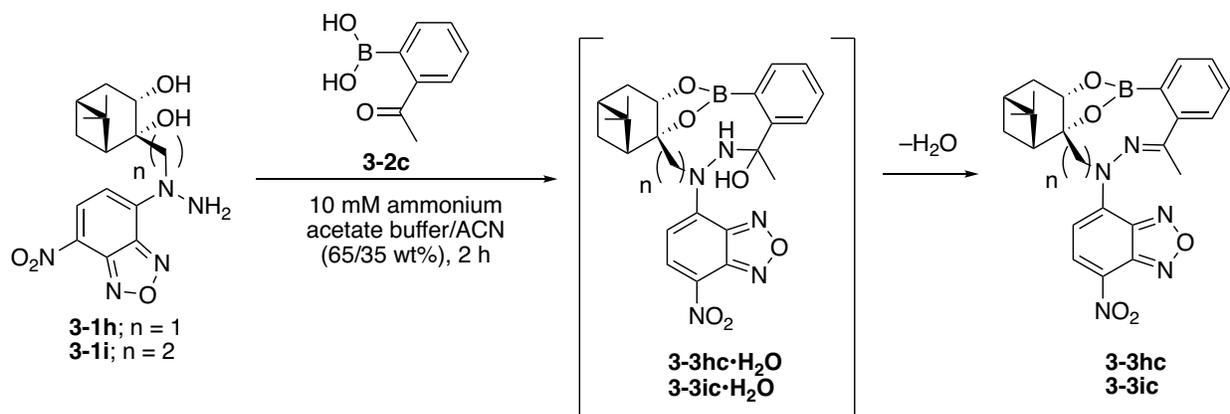
After several failed attempts, I followed another strategy where the hydrazine unit was first installed onto the epoxy alcohol **3-13** and diol **3-5** via a S<sub>N</sub>2 reaction under reflux conditions, followed by a S<sub>N</sub>Ar reaction of the most nucleophilic nitrogen (RNHNH<sub>2</sub>) with 4-chloro-7-nitro-benzofurazan (Scheme 3-16A and 3-16B). This approach was inspired by the simple S<sub>N</sub>Ar reaction of hydrazine monohydrate with 4-chloro-7-nitro-benzofurazan.<sup>35c</sup> Fortunately, this method afforded the desired homologues, **3-1h** (n = 1) and **3-1i** (n = 2) in moderate yields. In short, ring opening of epoxy alcohol **3-13** with excess hydrazine monohydrate was performed at 90 °C reflux for 1 h while more reactive diol **3-5** could be converted into its monosubstituted hydrazine derivative at 75 °C reflux. Then, it was critical to remove the excess remaining hydrazine under high vacuum for at least 2 h before 4-chloro-7-nitro-benzofurazan was introduced into the reaction vessel. It is noteworthy that purification of intermediate nopoldiol hydrazine derivatives was problematic; therefore, I decided to use these intermediates immediately in the second step without further purification.



**Scheme 3-16:** Synthesis of nopoldiol NBD hydrazine designs **3-1h** and **3-1i**. A) Synthesis of diol **3-1h**. B) Synthesis of diol **3-1i**.

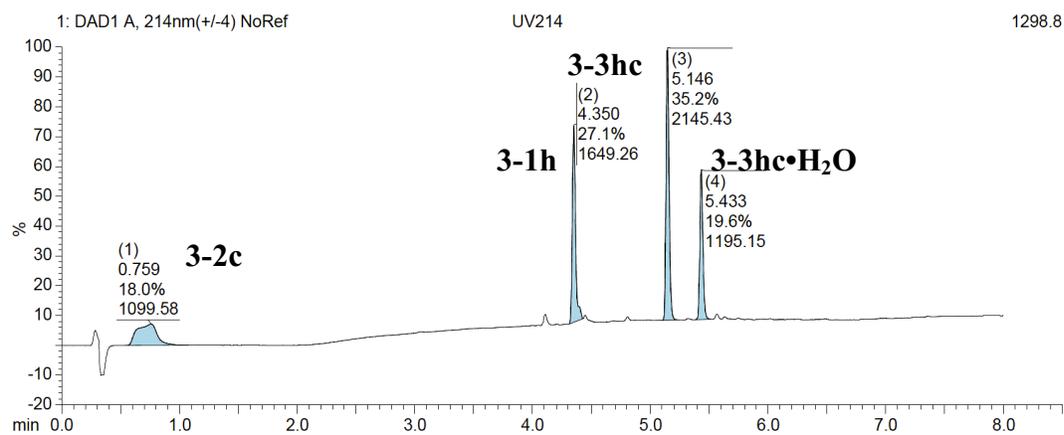
### 3.2.2.1 Conjugation studies of diols **3-1h/3-1i** with boronic acid **3-2c**

At first, conjugation of designed diols **3-1h/3-1i** and boronic acid **3-2c** was monitored by HPLC-MS in 10 mM of ammonium acetate buffer/ACN (65/35 w%, pH ~7) (Scheme 3-17). Equivalent volumes of boronic acid **3-2c** solution (5.0 mM, 2.5 mM final concentration) and diol **3-1h/3-1i** solution (4.0 mM, 2.0 mM final concentration) were mixed. After 2 h, the reaction mixture was monitored and analyzed (Figure 3-7 and 3-8). The results indicated that the reaction proceeded first *via* hemiaminal **3-3hc·H<sub>2</sub>O/3-3ic·H<sub>2</sub>O** conjugates, which were then slowly converted into the desired hydrazone **3-3hc/3-3ic** (Scheme 3-17). With these promising results in hand, next, I tested the stability of the conjugates **3-3hc** and **3-3ic**.

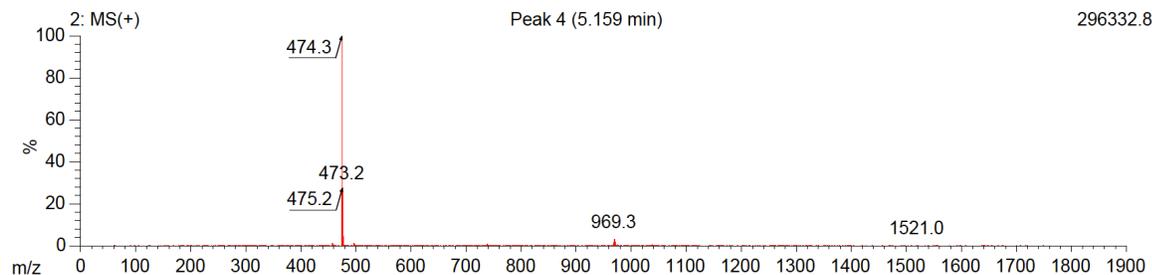


**Scheme 3-17:** The conjugation of diols **3-1h/3-1i** (2.0 mM) with **3-2c** (2.5 mM).

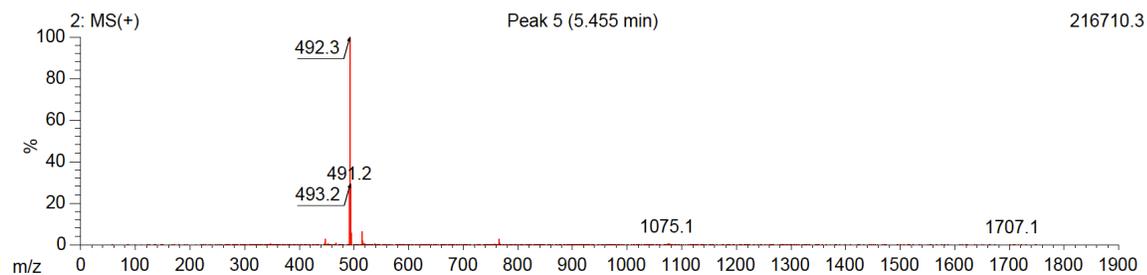
A)



B)

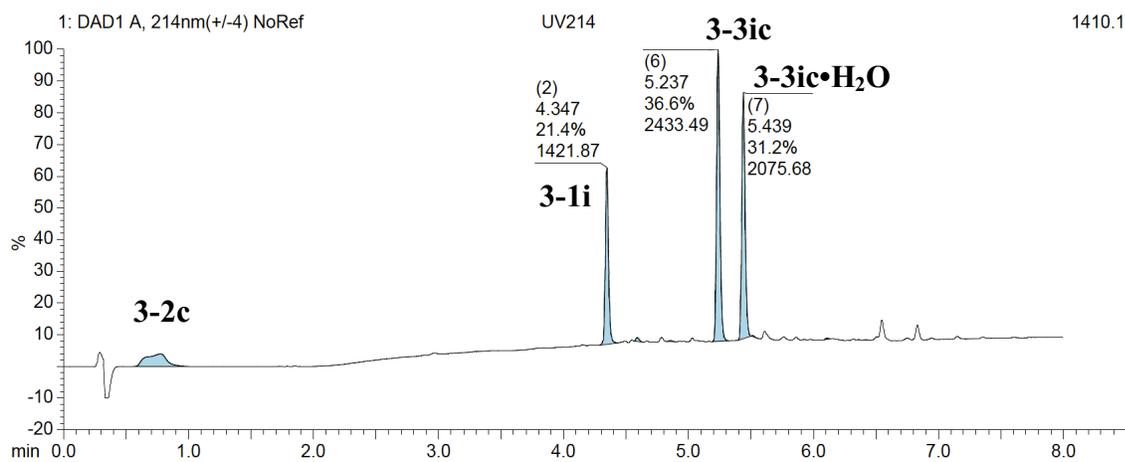


C)

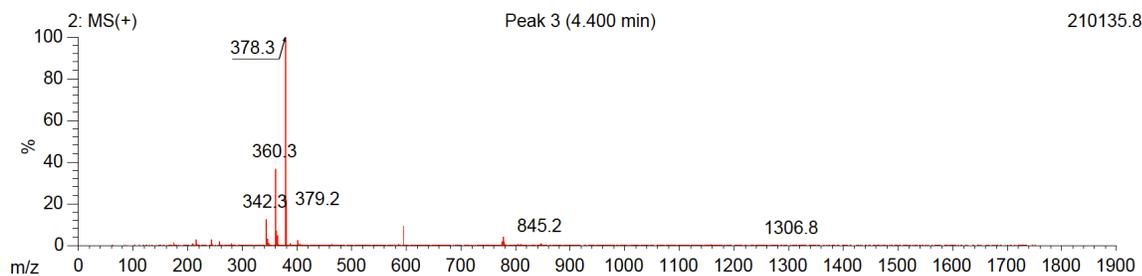


**Figure 3-7:** A) HPLC chromatogram of diol **3-1h** and **3-2c** reaction after 2 h. B) (+)-MS trace of **3-3hc**. C) (+)-MS trace of **3-3hc•H<sub>2</sub>O**.

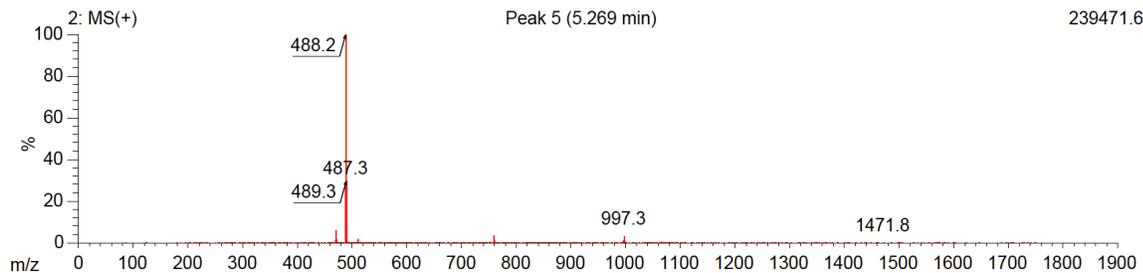
A)



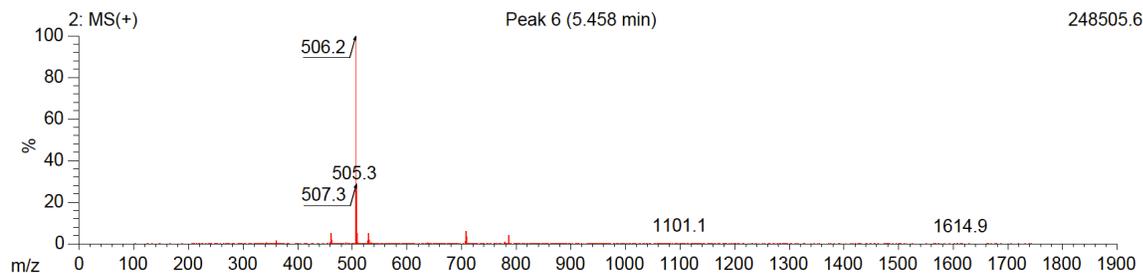
B)



C)



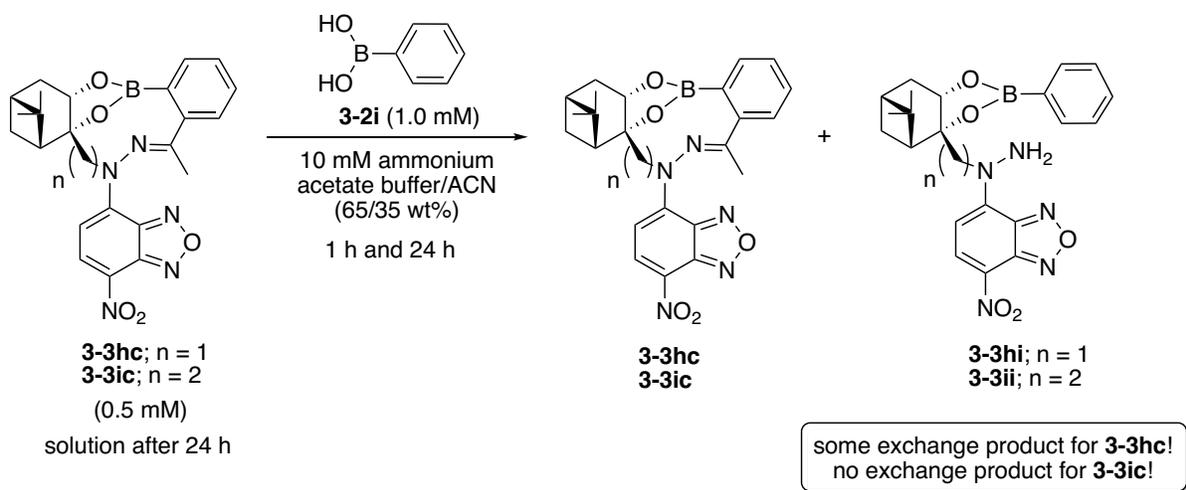
D)



**Figure 3-8:** A) HPLC chromatogram of diol **3-1i** and **3-2c** reaction after 2 h. B) (+)-MS trace of **3-1i**. C) (+)-MS trace of **3-3ic**. D) (+)-MS trace of **3-3ic•H<sub>2</sub>O**.

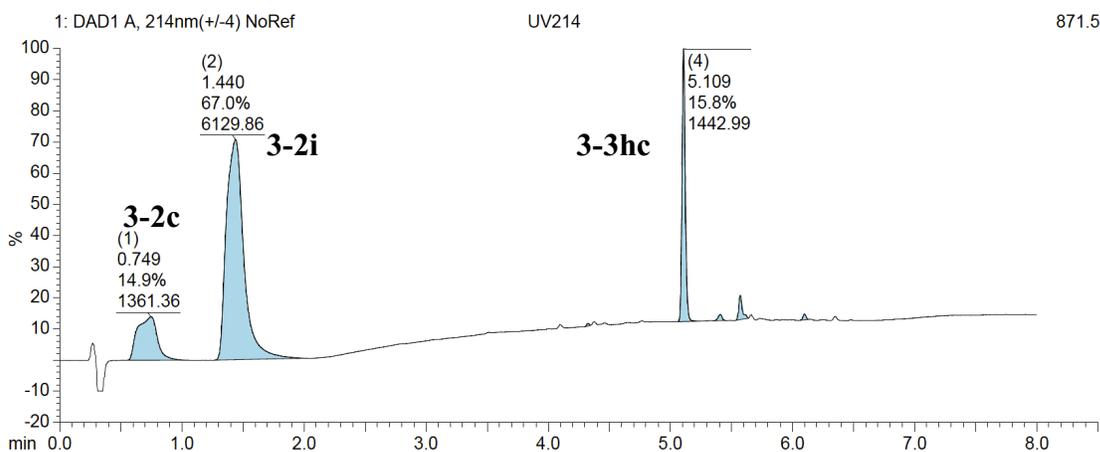
### 3.2.2.2 Stability assessment of **3-3hc/3-3ic** via a cross-over study

After the desired conjugations **3-3hc/3-3ic** were confirmed, their stability was studied *via* a cross-over experiment and characterized *via* HPLC-MS analysis (Scheme 3-18). Equivalent volumes of diol **3-1h/3-1i** (1.0 mM, 0.5 mM final concentration) and boronic acid **3-2c** (1.5 mM, 0.75 mM final concentration) were mixed. After the desired conjugation **3-3hc/3-3ic** was completely formed in 24 h, boronic acid **3-2i** (40 mM, final concentration ~1.0 mM, 2.0 equiv) was introduced into the reaction mixture. Then, the final mixture was incubated for 1 h and 24 h at room temperature; the corresponding peaks formed were analyzed and characterized by HPLC-MS (Figure 3-9 and 3-10). The results showed that some exchange product **3-3hi** was observed for **3-3hc** within 24 h (Figure 3-9B) while the conjugate **3-3ic** did not undergo any exchange with **3-2i** and remained intact even after 24 h (Figure 3-10B).

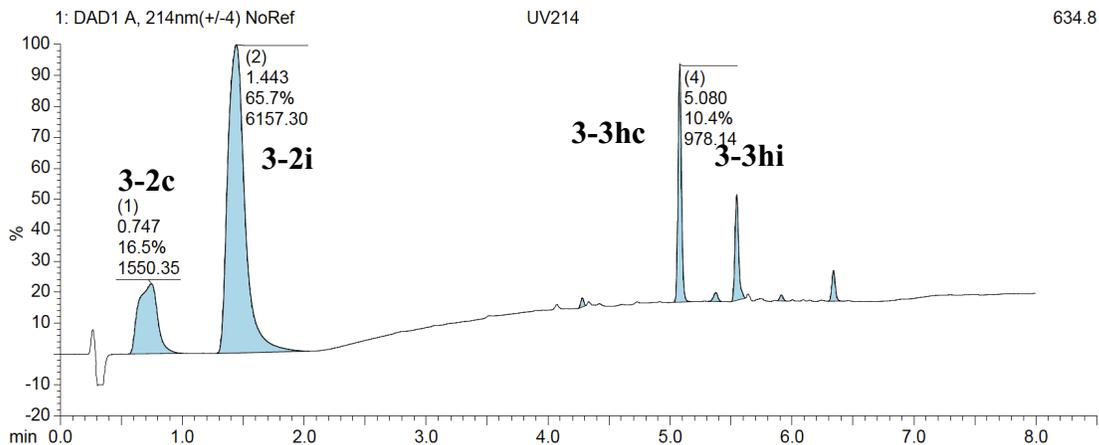


**Scheme 3-18:** Stability test of **3-3hc/3-3ic** via a cross-over experiment.

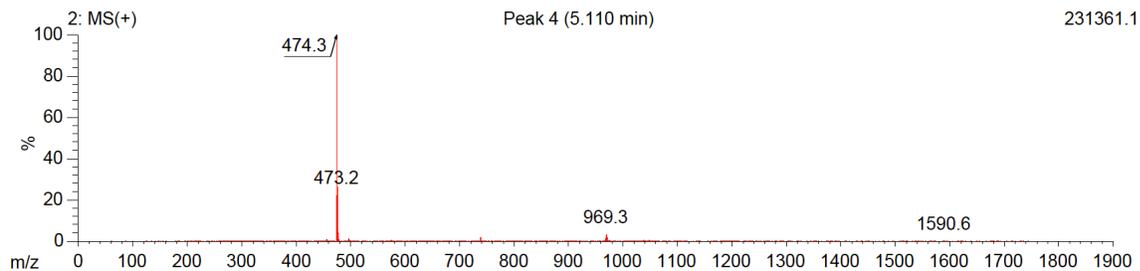
A)



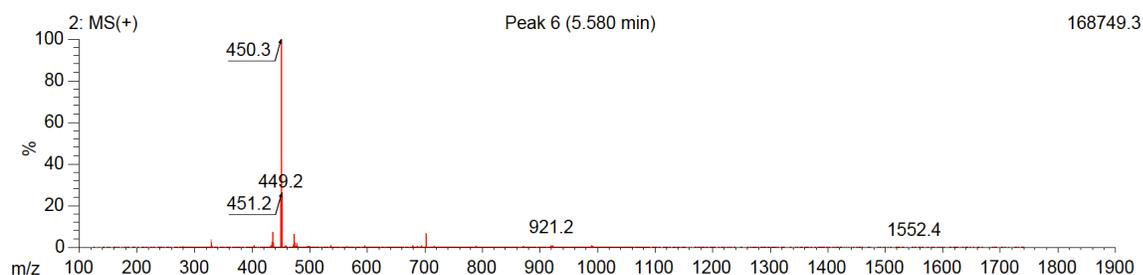
B)



C)

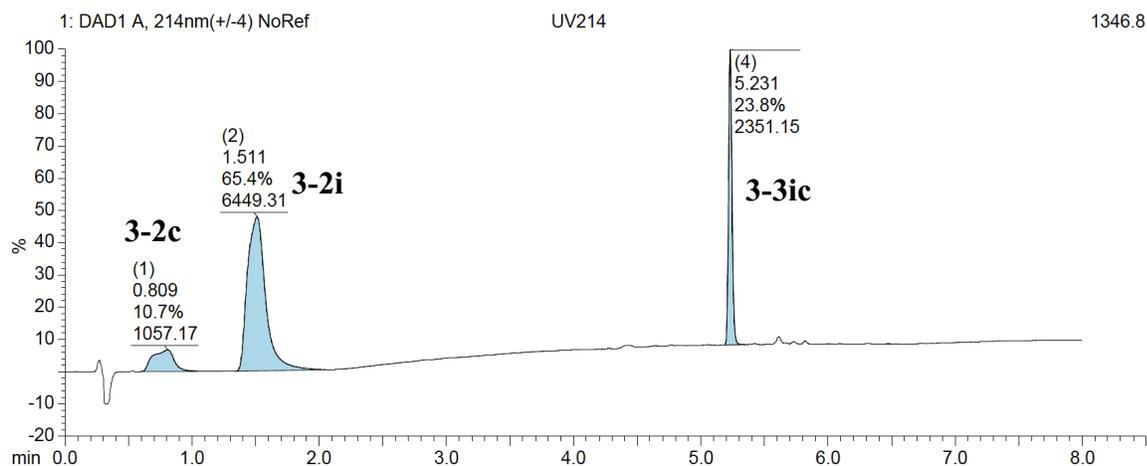


D)

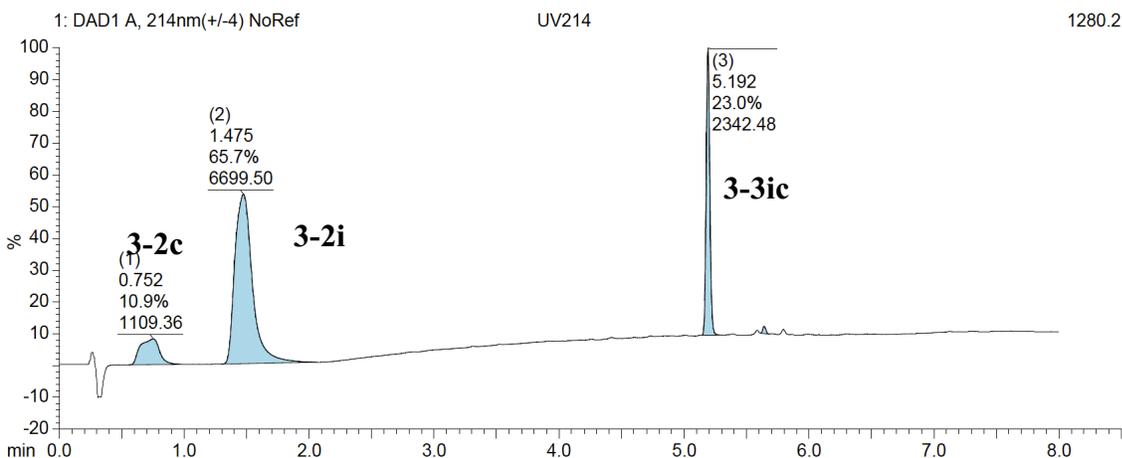


**Figure 3-9:** A) HPLC chromatogram of cross-over experiment of **3-3hc** with **3-2i** after 1 h. B) HPLC chromatogram of cross-over experiment of **3-3hc** with **3-2i** after 24 h. C) (+)-MS trace of **3-3hc** D) (+)-MS trace of **3-3hi**.

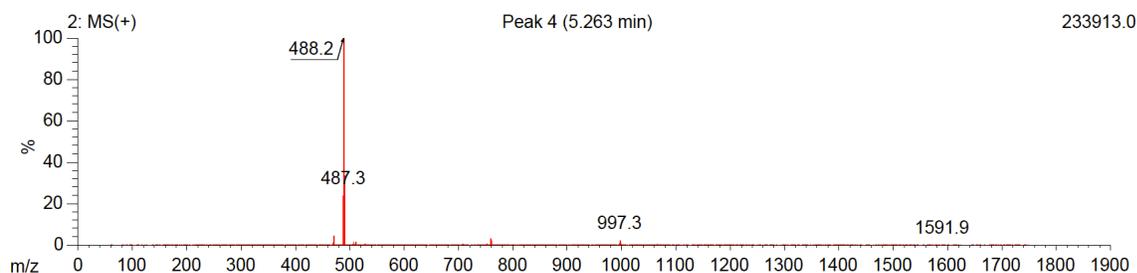
A)



B)



C)

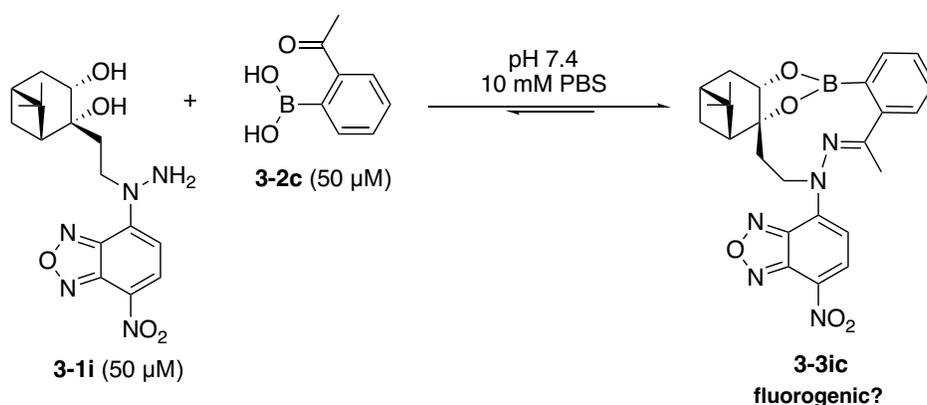


**Figure 3-10:** A) HPLC chromatogram of cross-over experiment of **3-3ic** with **3-2i** after 1 h. B) HPLC chromatogram of cross-over experiment of **3-3ic** with **3-2i** after 24 h. C) (+)-MS trace of **3-3ic**.

### 3.2.2.3 Study of the fluorogenicity of the reaction between diol **3-1i** with boronic acid **3-2c**

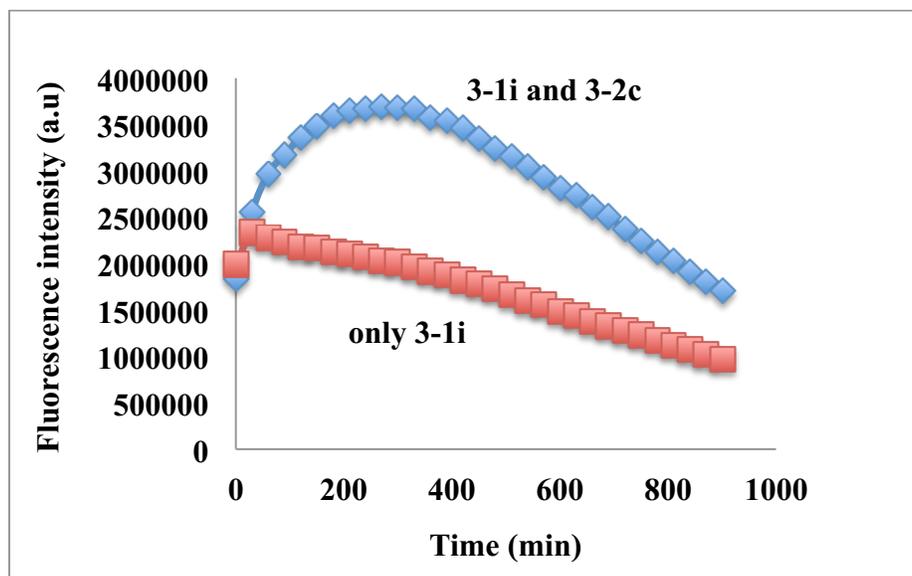
Based on the above stability study, the diol **3-1i** appeared superior to diol **3-1h**. Therefore, I studied the fluorogenicity of the model reaction between diol **3-1i** and boronic acid **3-2c** using a fluorescence plate reader (Scheme 3-19). When boronic acid **3-2c** (50  $\mu$ M final concentration) was introduced into the solution of diol **3-1i** (50  $\mu$ M final concentration) in 10 mM PBS buffer, it was observed that the maximum excitation shifted from 470 nm to 440 nm, and the maximum emission was detected as 560 nm. Thus, the reaction between diol **3-1i** and boronic acid **3-2c** (50  $\mu$ M/50  $\mu$ M) was monitored at excitation 440 nm and emission 560 nm over 15 h. The results depicted in Figure 3-11A indicated that fluorescence intensity gradually

increased and reached its maximum over approximately 4 h compared to the diol **3-1i** alone (50  $\mu\text{M}$ ), but later it started to decline likely due to photo bleaching. Notably, when the ratio of reaction fluorescence intensity ( $I$ ) to fluorescence intensity of diol **3-1i** ( $I_0$ ) was examined over time, only  $\sim 1.3$  fold change in 1 h and  $\sim 1.8$  fold in 4 h were observed (Figure 3-11B). These results demonstrated that the system slowly rendered only a slight increase in fluorescence intensity. Moreover, it is inadequate due to its very low fluorescence intensity. Furthermore, as also observed in the conjugation study, formation of hydrazone **3-3ic** occurs slowly from its hemiaminal precursor **3-3ic** $\cdot\text{H}_2\text{O}$ , which was not expected to be fluorescent.

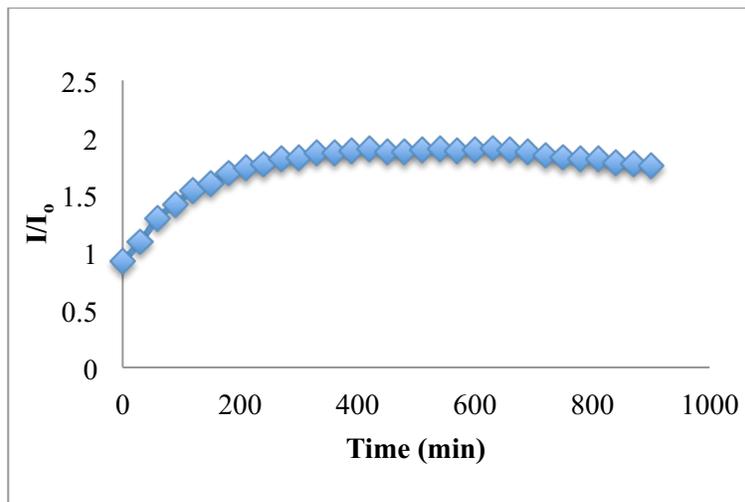


**Scheme 3-19:** Fluorogenicity study of diol **3-1i** and **3-2c**.

A)



B)



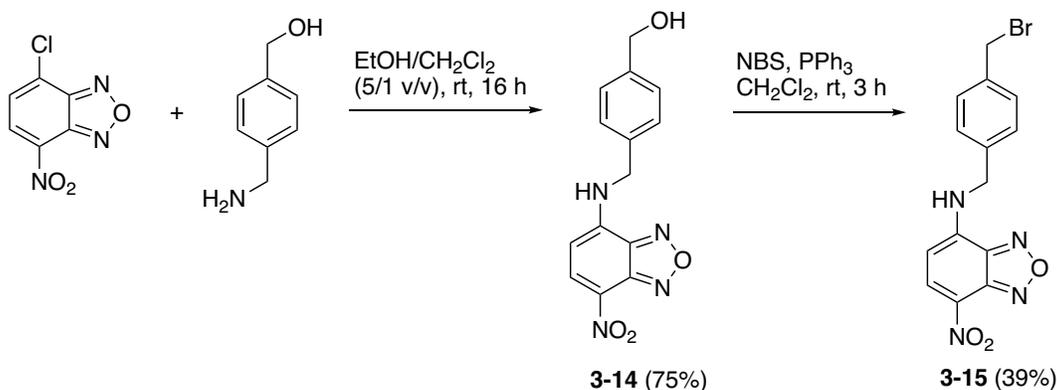
**Figure 3-11:** Fluorogenicity study of the reaction between diol **3-1i** and boronic acid **3-2c**. A) Monitoring fluorescence intensity change of the reaction **3-1i** with **3-2c** (50:50  $\mu\text{M}$ ) at excitation 440 nM and at emission 560 nM. B) Fluorescence intensity comparison ( $I/I_0$ ) vs time: fluorescence intensity of the reaction mixture ( $I$ ) over fluorescence intensity of **3-1i** ( $I_0$ ). Results are the average values of triplicate testing.

In summary, diol **3-1i** and the resulting conjugates **3-3ic** showed low fluorescence intensity in neutral aqueous media. Yet, these efforts were useful to investigate the potential of this reaction and determine the correct spacer length ( $n$ ). The reactivity of both **3-1h** and **3-1i** towards **3-2c** was inferior compared to **3-1e**. Also, the desired conjugates were formed first through boronate formation followed by a slow conversion into the cyclic hydrazone. Formation and stability of **3-3ic** were confirmed in the presence of excess phenylboronic acid while **3-3hc** exhibited a slight cross-over product. Both **3-1h** and **3-1i**, however, demonstrated a similar reactivity, and therefore diol **3-1i** was chosen for further studies due to its stability as well as its relatively simpler synthesis.

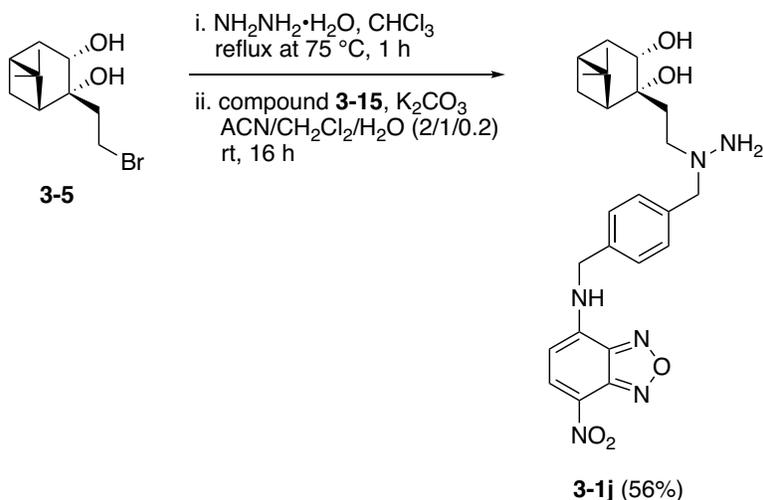
### 3.2.3 Design of hydrazine-functionalized fluorescently labeled nopoldiol derivatives

The conjugation and stability studies of fluorogenic designs led us to continue with spacer unit  $n = 2$  and to synthesize a nopoldiol derivative containing a disubstituted hydrazine unit as well as a fluorescent molecule such as NBD. In order to preserve the fluorescence intensity of NBD, there must be a space linker between hydrazine unit and NBD. Thus, a

nopoldiol derivative **3-1j** with a NBD dye was designed and synthesized (Scheme 3-20 and 3-21). A benzylic linker was employed for the NBD dye since the reactivity of a benzyl bromide **3-15** towards the *in situ* generated nopoldiol hydrazine is higher than that of alkyl bromides. Even though the reaction provided high conversion and diol **3-1j** preserved the fluorophore's intensity, it exhibited low stability and its purification led to partial decomposition. Nonetheless, a small amount of **3-1j** could still be isolated to test its conjugation with **3-2c** via HPLC-MS analysis.



**Scheme 3-20:** Synthesis of NBD-benzyl bromide **3-15**.



**Scheme 3-21:** Synthesis of nopoldiol hydrazine-functionalized NBD **3-1j**.

### 3.2.3.1 Conjugation study of diol **3-1j** and **3-2c** and stability test of conjugate **3-3jc**

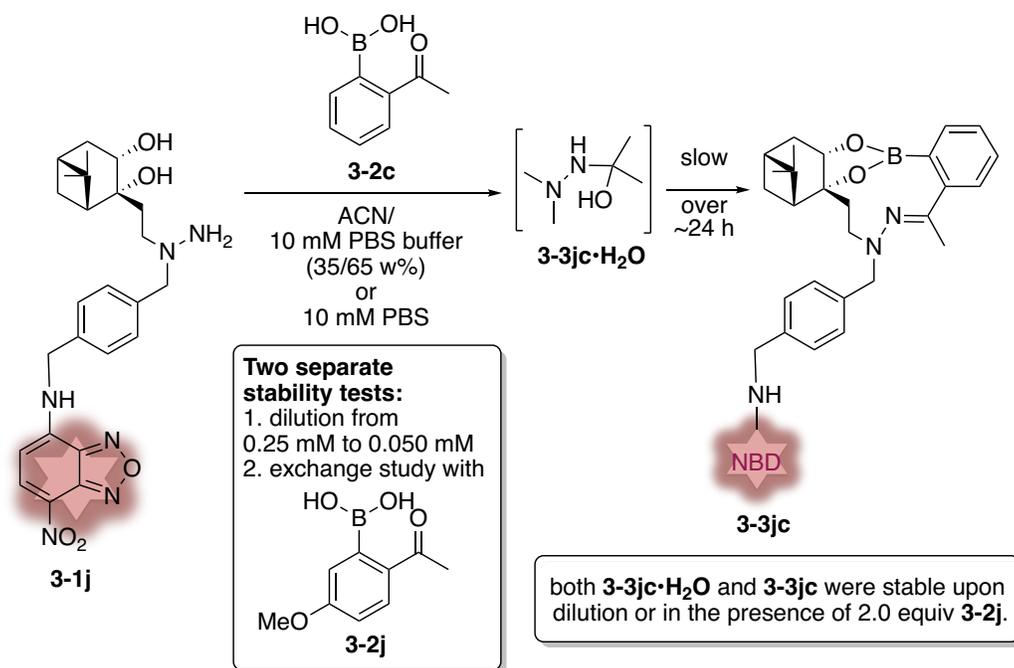
Conjugation of diol **3-1j** and boronic acid **3-2c** was analyzed in different solvent systems such as 10 mM PBS (pH 7.4), 10 mM PBS/ACN (65/35 w%, pH ~7), and 10 mM ammonium acetate buffer (Table 3-2). This study was monitored and characterized *via* UV and fluorescence

HPLC-MS at 10, 40, 70 min and 24 h (Figure 3-12 and 3-13). Percent conversions were calculated based on the integrated areas of **3-1j**, **3-3jc•H<sub>2</sub>O**, **3-3jc** observed in the fluorescence HPLC chromatogram.

On the other hand, dilution and cross-over experiments of conjugate **3-3jc** were performed to examine its stability (Table 3-2). Equivalent volumes of diol **3-1j** (0.5 mM in 10 mM PBS/ACN 65/35 w%, pH ~7, 0.25 mM final concentration) and boronic acid **3-2c** (0.5 mM, 0.25 mM final concentration) were mixed. After the reaction mixture was incubated at room temperature for 4 h, its UV (220 nm) and fluorescence spectra from the HPLC chromatograms (excitation at 465 nm and emission at 535 nm) were monitored. Then, the reaction mixture was diluted to 50  $\mu$ M with 10 mM PBS buffer. Separately, for a cross-over experiment, boronic acid **3-2j** (40 mM, 1.0 mM final concentration) was introduced into the reaction mixture, which was incubated for 24 h. These two studies were then monitored and analyzed *via* UV and fluorescence HPLC-MS after 1 h and 24 h (Figure 3-12 and 3-13).

The conjugation reaction again proceeded initially through the formation of hemiaminal boronate **3-3jc•H<sub>2</sub>O**, followed by slow conversion into the hydrazone **3-3jc** with two E/Z isomers based on the two product peaks with the same molecular weight observed in the HPLC-MS chromatogram (Table 3-2, entry 1-4, Figure 3-12). Moreover, elemental composition analysis of this reaction mixture was monitored *via* high resolution-ESI, which also confirmed that **3-3jc•H<sub>2</sub>O** ( $[M]^+ = 625$  m/z) slowly converts into **3-3jc** ( $[M]^+ = 607$  m/z) as shown in Figure 3-14. After 24 h, a precipitate in PBS buffer was observed likely due to the low solubility of **3-3jc** (Table 3-2, entry 1 and 2). Satisfactorily, both **3-3jc•H<sub>2</sub>O** and **3-3jc** preserved their stability under these conditions even after 24 h (Table 3-2 and Figure 3-13). Intermediate **3-3jc•H<sub>2</sub>O** was found to be a stable conjugate that transforms slowly into **3-3jc**. The approximate rate constant was estimated at  $\sim 8$  M<sup>-1</sup>s<sup>-1</sup> for the formation of **3-3jc•H<sub>2</sub>O**, based on the reaction's half-life as monitored by HPLC-MS analysis (Table 3-2).

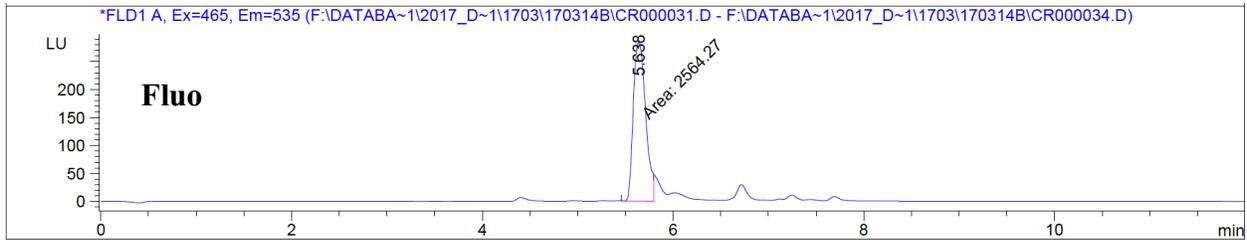
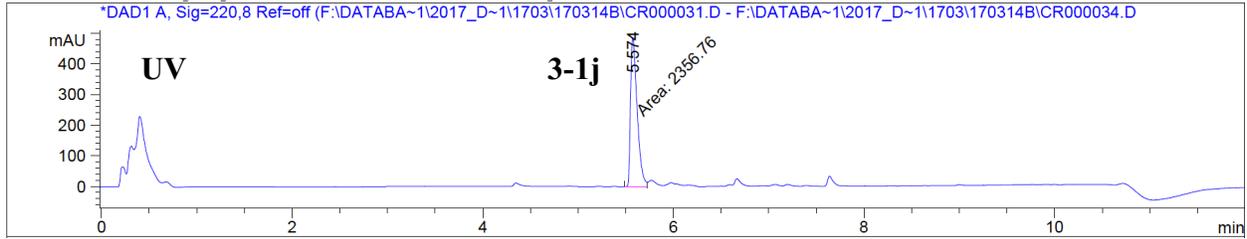
**Table 3-2:** Results of the aqueous conjugation reaction between **3-2j** and **3-2c**, and the stability test of conjugate **3-3jc•H<sub>2</sub>O** and **3-3jc** via HPLC-MS analysis.



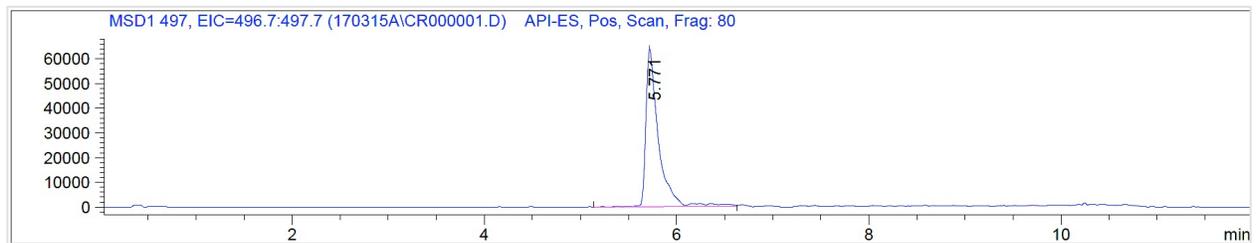
Entry	Conc. of reagents (mM)	% Proportions of <b>3-1j</b> / <b>3-3jc•H<sub>2</sub>O</b> / <b>3-3jc</b>			
		10 min	40 min	70 min	24 h
1	0.025 <sup>[a]</sup>	83/12/5	72/18/10	65/22/13	precipitate
2	0.050 <sup>[a]</sup>	71/23/6	54/36/10	49/40/11	precipitate
3	0.050 <sup>[b]</sup>	62/31/7	45/45/10	41/53/12	-
4	0.25 <sup>[c]</sup>	37/37/26	30/22/48	25/15/60	8/4/88

[a] 10 mM PBS buffer was used. [b] 10 mM ammonium acetate buffer was used. [c] ACN/10 mM PBS buffer (35/65 w%) was used.

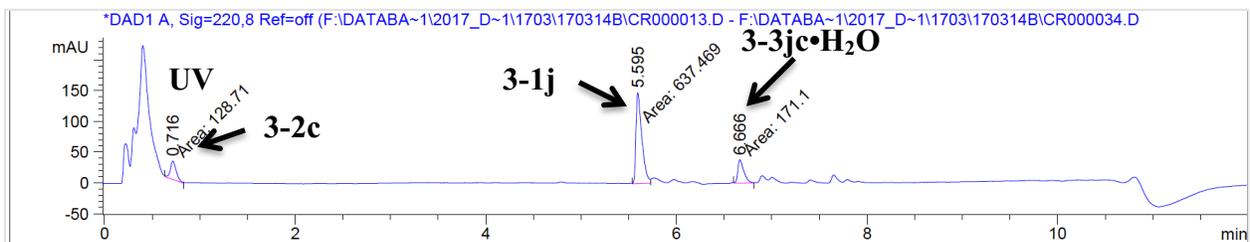
A) Only 3-1j

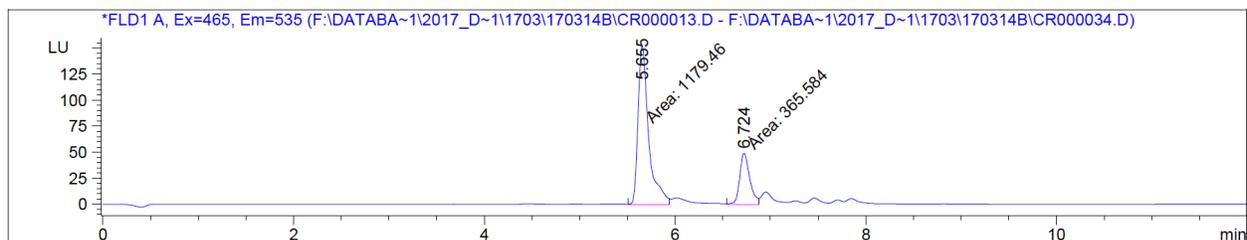


B) (+)-LC-MS trace of 3-1j

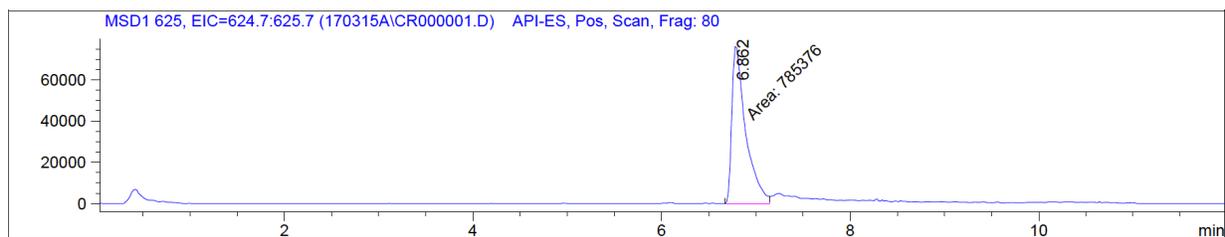


C) Reaction of 3-1j and 3-2c at 50  $\mu$ M in 10 min

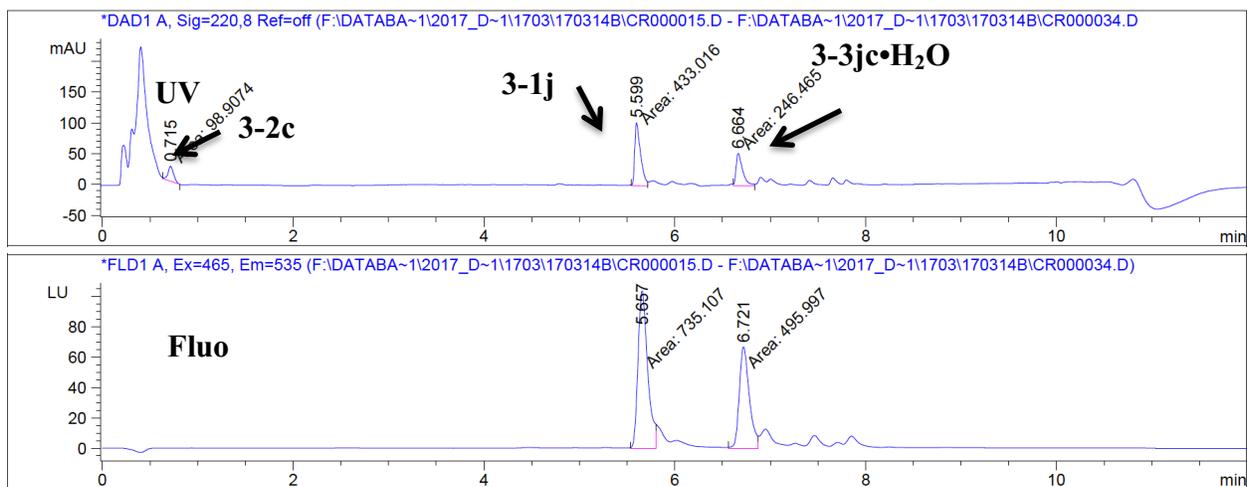




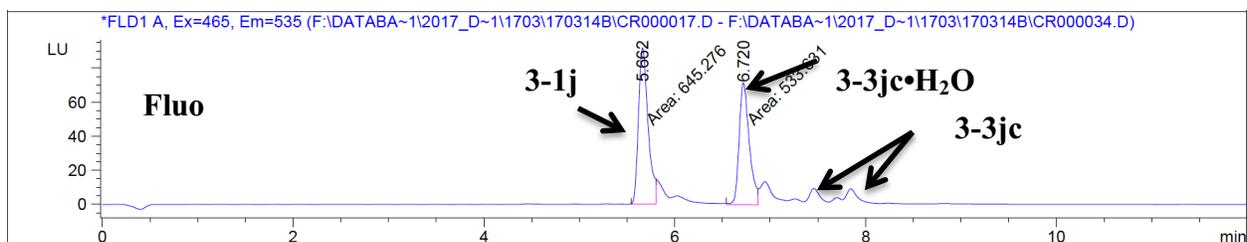
D) (+)-LC-MS trace of **3-3jc•H<sub>2</sub>O**



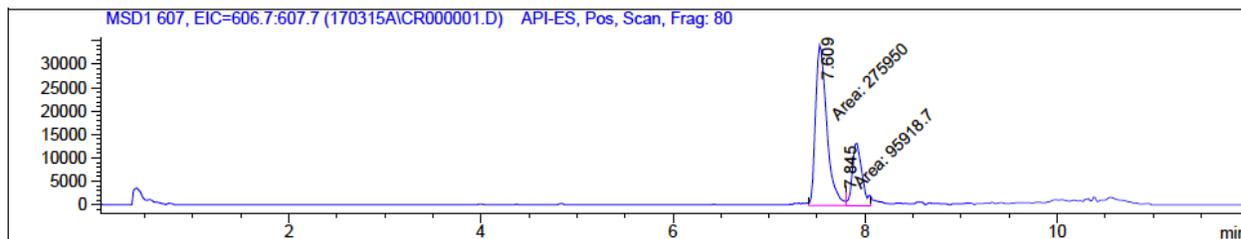
E) Reaction of **3-1j** and **3-2c** at 50  $\mu$ M in 40 min



F) Reaction of **3-1j** and **3-2c** at 50  $\mu$ M in 70 min

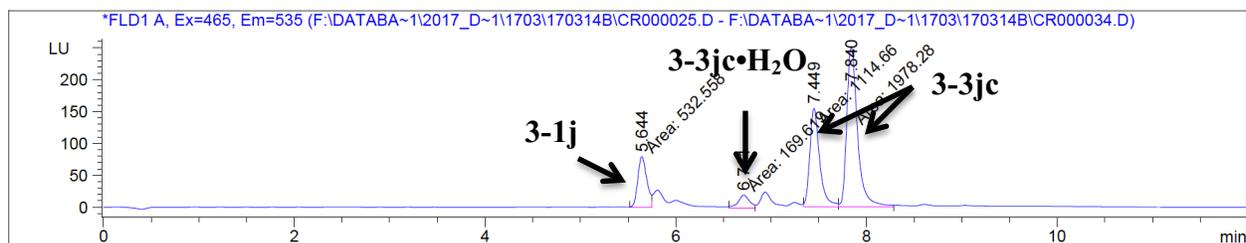


H) (+)-LC-MS trace of **3-3jc**

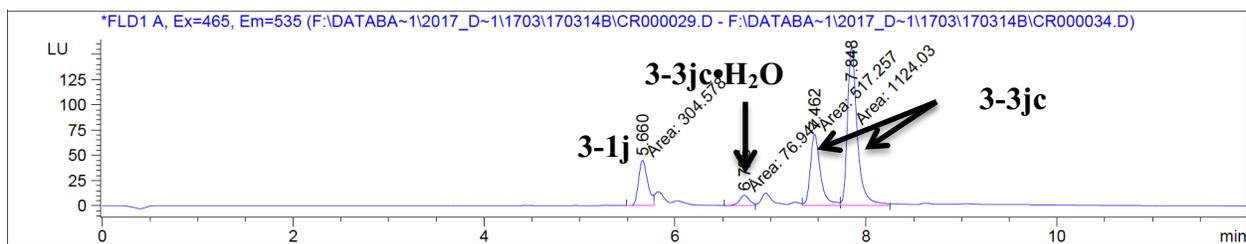


**Figure 3-12:** UV and Fluorescence HPLC-MS analyses of reaction of **3-1j** and **3-2c** (50/50  $\mu\text{M}$ ) in 10 mM PBS buffer.

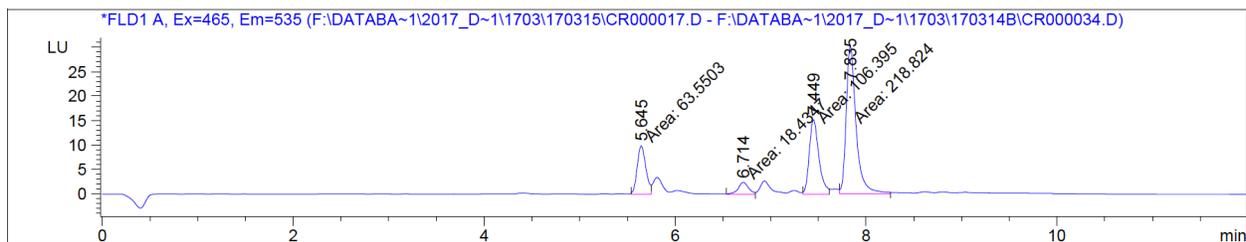
A) In 4 h before dilution and exchange



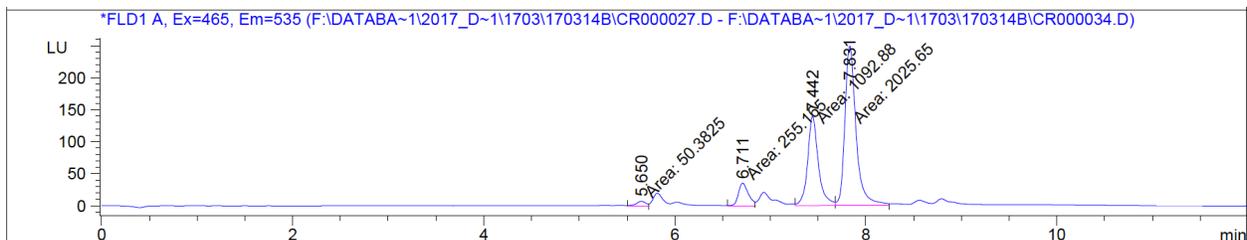
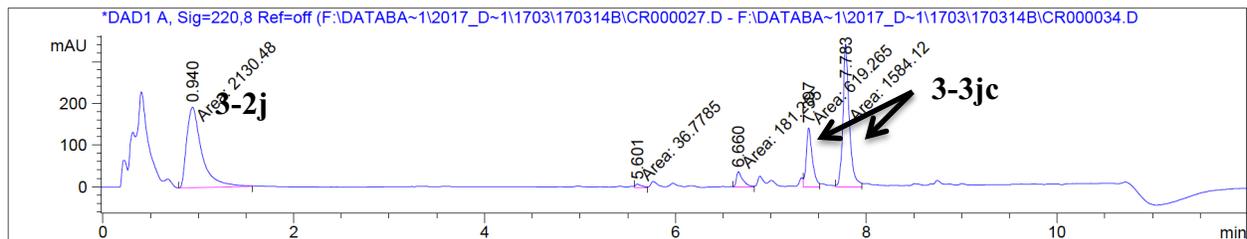
B) After dilution in 1 h



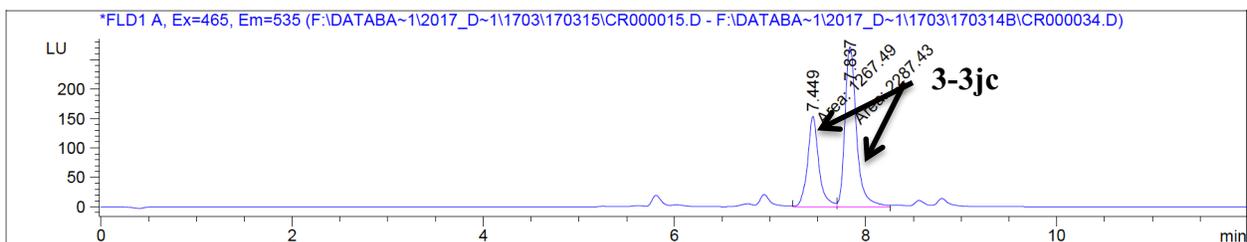
C) After dilution in 24 h



D) After exchange in 1 h

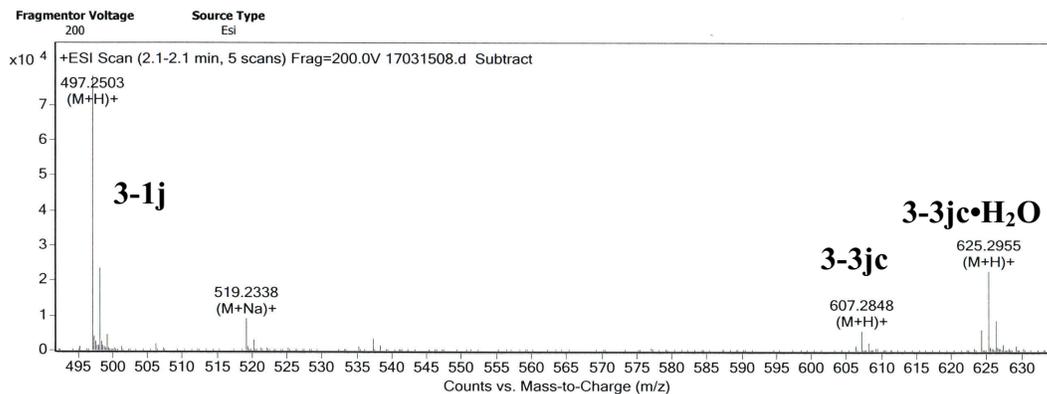


E) After exchange in 24 h



**Figure 3-13:** UV and Fluorescence HPLC analyses of the reaction of **3-1j** and **3-2c** (250/250  $\mu\text{M}$ ) in ACN/10 mM PBS buffer (35/65 w%) before and after dilution/cross-over study.

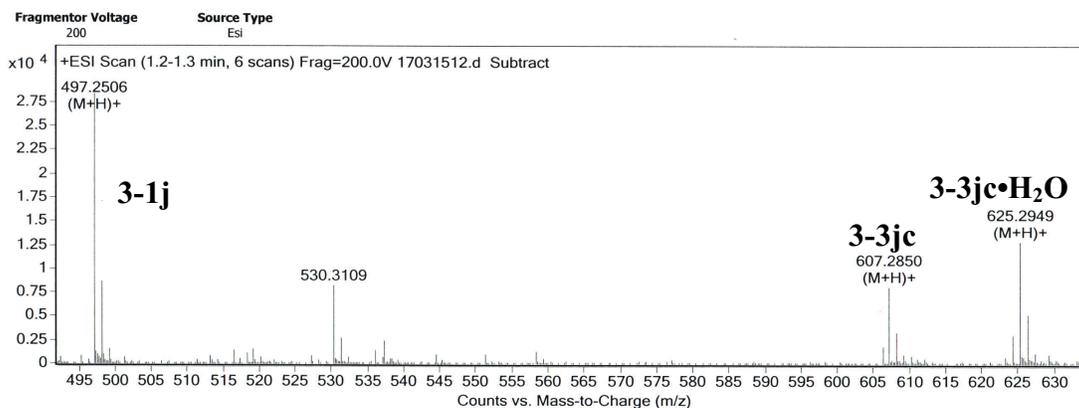
### A) In 30 min



#### Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C25 H32 N6 O5	C25 H33 N6 O5	496.2431	496.2434	497.2503	497.2507	0.36	0.73	13	(M+H)+	97.92
C25 H32 N6 O5	C25 H32 N6 Na O5	496.2445	496.2434	519.2338	519.2326	-1.1	-2.21	13	(M+Na)+	73.65
C33 H37 [11B] N6 O6	C33 H38 [11B] N6 O6	624.2882	624.2868	625.2955	625.294	-1.42	-2.28	19	(M+H)+	89.65
C33 H35 [11B] N6 O5	C33 H36 [11B] N6 O5	606.2773	606.2762	607.2848	607.2835	-1.1	-1.81	20	(M+H)+	75.64

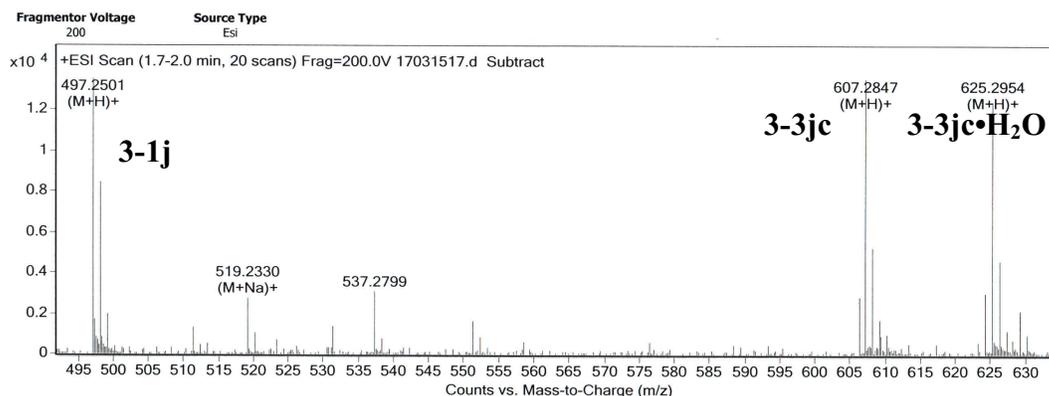
### B) In 2 h



#### Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C25 H32 N6 O5	C25 H33 N6 O5	496.2433	496.2434	497.2506	497.2507	0.15	0.3	13	(M+H)+	99.57
C33 H35 [11B] N6 O5	C33 H36 [11B] N6 O5	606.2777	606.2762	607.285	607.2835	-1.48	-2.43	20	(M+H)+	83.18
C33 H37 [11B] N6 O6	C33 H38 [11B] N6 O6	624.2876	624.2868	625.2949	625.294	-0.84	-1.35	19	(M+H)+	91.94

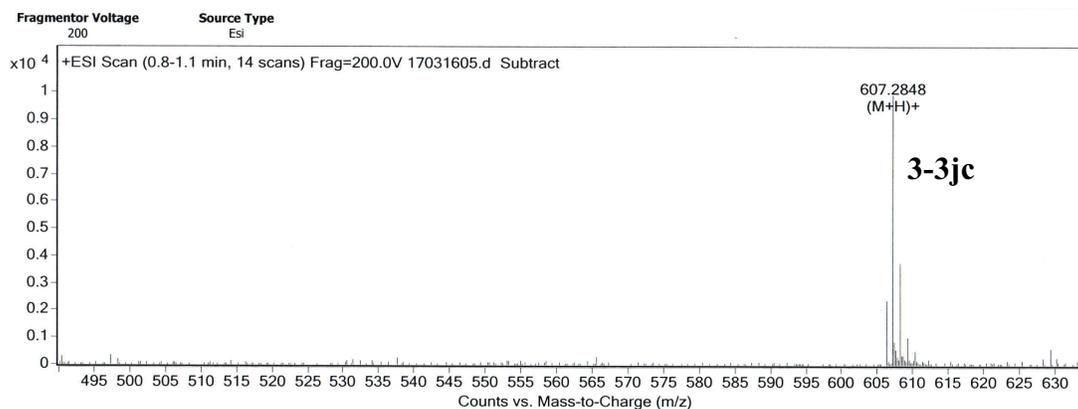
C) In 3 h



Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C25 H32 N6 O5	C25 H33 N6 O5	496.2429	496.2434	497.2501	497.2507	0.57	1.14	13	(M+H)+	89.68
C25 H32 N6 O5	C25 H32 N6 Na O5	496.2432	496.2434	519.233	519.2326	0.23	0.46	13	(M+Na)+	57.46
C33 H35 [11B] N6 O5	C33 H36 [11B] N6 O5	606.2775	606.2762	607.2847	607.2835	-1.32	-2.18	20	(M+H)+	86.45
C33 H37 [11B] N6 O6	C33 H38 [11B] N6 O6	624.2881	624.2868	625.2954	625.294	-1.35	-2.16	19	(M+H)+	91.68

D) In 24 h



Formula Calculator Results

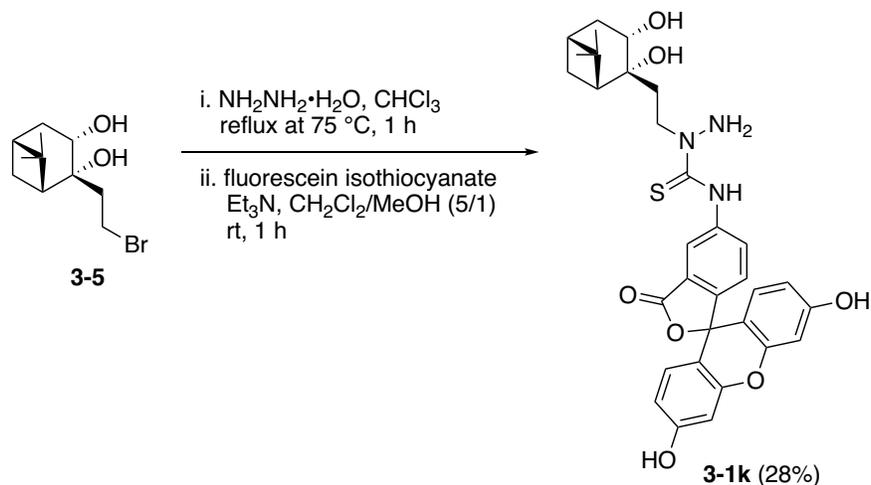
Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C33 H35 [11B] N6 O5	C33 H36 [11B] N6 O5	606.2775	606.2762	607.2848	607.2835	-1.33	-2.19	20	(M+H)+	86.59

Figure 3-14: High resolution ESI-MS results for the reaction between **3-1j** with **3-2c**.

### 3.2.4 Design of fluorescein conjugated nopoldiol thiosemicarbazide

In order to improve the chemical stability of the diol reagent, a (thio)semicarbazide unit, which was recently shown to be more stable and safer than phenyl hydrazine,<sup>25</sup> was chosen in the design and synthesis of diol reagent **3-1k**. This reagent was expected to be more stable, more reactive towards keto boronic acids, like **3-2c**. It should afford high conversion<sup>25</sup> because the expected product, **3-3kc**, would not exhibit any inhibitory interaction with the starting materials that could stall the reaction. Moreover, **3-1k** could be synthesized easily with commercially

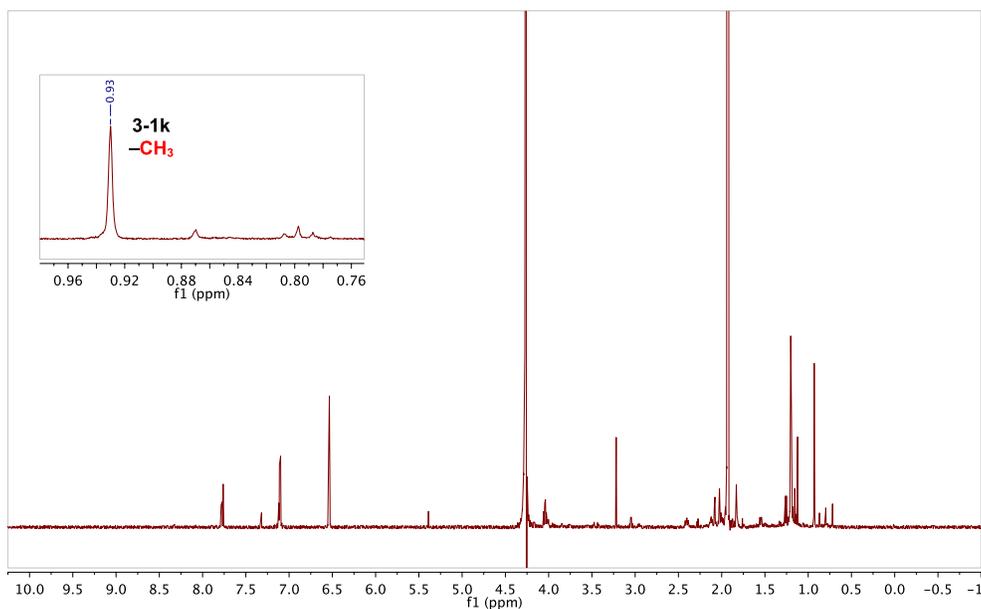
available fluorescein isothiocyanate and nopoldiol hydrazine, which was easily prepared *in situ* from diol **3-5** and hydrazine monohydrate, as the precursors (Scheme 3-22). The yield of this reaction was found to be low (~28%) due to possible side reactions, such as double/triple alkylation in the first step and low regioselectivity between two hydrazine nitrogens (NH<sub>2</sub> vs NH). As expected, the integrity of diol **3-1k**, in the designated buffer solution, remained intact even after one week according to <sup>1</sup>H-NMR analysis (Figure 3-15).



**Scheme 3-22:** Synthesis of nopoldiol thiosemicarbazide **3-1k**.

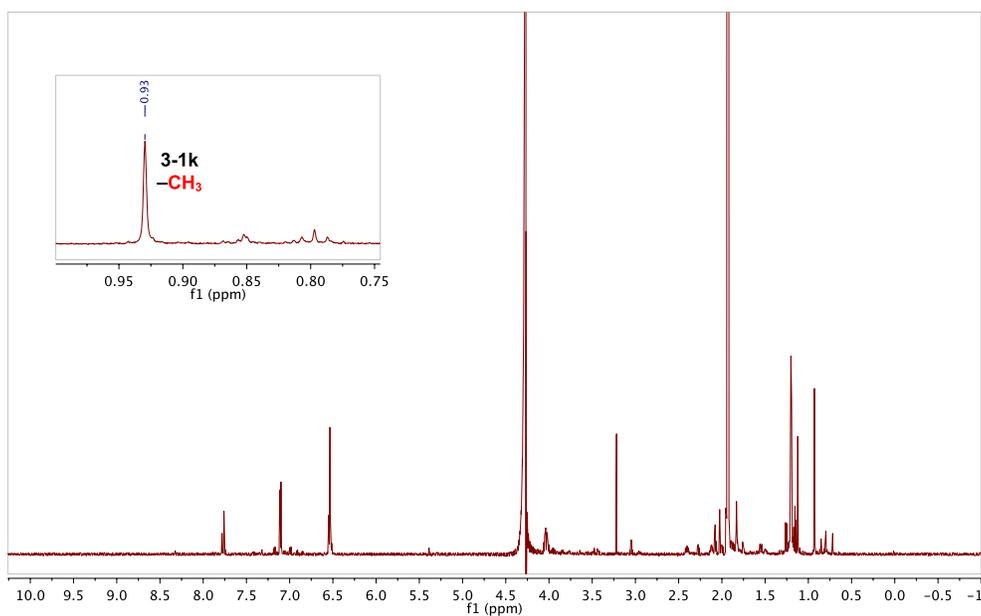
A)

Only **3-1k** - 1 day - (0.25 mM) in 50 mM D<sub>2</sub>O phosphate buffer/ACN-d<sub>3</sub>  
699.764 MHz H1 1D in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



B)

Only **3-1k** - 7 days - (0.25 mM) in 50 mM D<sub>2</sub>O phosphate buffer/ACN-d<sub>3</sub>  
699.764 MHz H1 1D in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



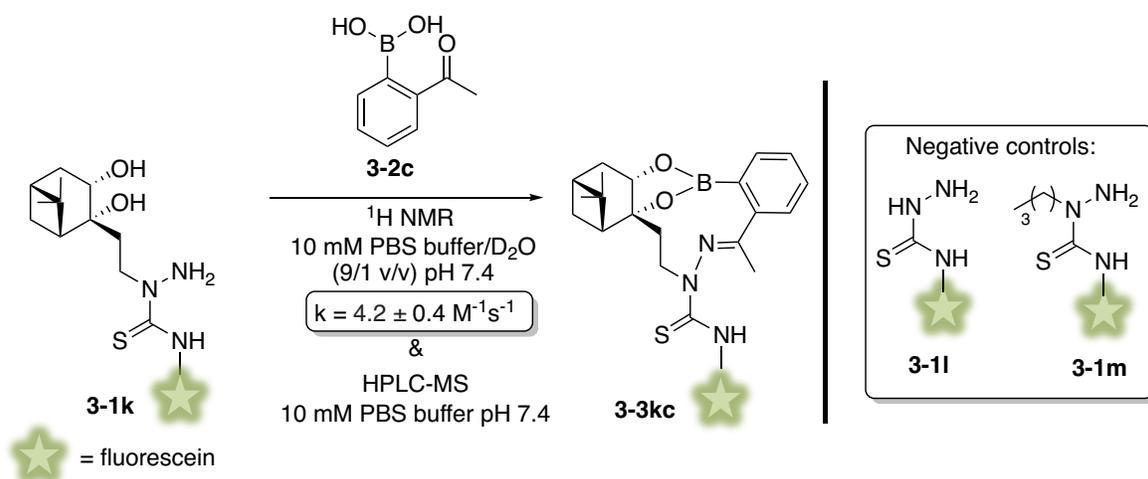
**Figure 3-15: Stability of diol 3-1k.** A) <sup>1</sup>H NMR spectrum of diol **3-1k** in 50 mM D<sub>2</sub>O phosphate buffer/ACN-d<sub>3</sub> (65/35 w%) in one day. B) <sup>1</sup>H NMR spectrum of diol **3-1k** in one week.

### 3.2.4.1 Conjugation studies of diol **3-1k** and boronic acid **3-2c**

With the more stable diol reagent **3-1k** in hand, the conjugation reaction was examined *via*  $^1\text{H}$  NMR, HPLC-MS and high resolution ESI-MS analyses (Figure 3-16, 3-17, 3-18). Boronate and thiosemicarbazone formation were investigated with  $^1\text{H}$  NMR analysis using 0.05 M of  $\text{D}_2\text{O}$  potassium phosphate buffer/ACN- $\text{d}_3$  (65/35 w%, pD 7.4) or 10 mM PBS buffer/ $\text{D}_2\text{O}$  (9/1 v/v). As negative controls, commercially available nopoldiol-free fluorescein-labeled thiosemicarbazide **3-1l** and synthetic derivative **3-1m** were used in order to understand the role of boronate formation. The reaction mixtures were monitored at several time intervals (10 min, 30 min, 60 min, 120 min, 3.5 h and 24 h) at room temperature, and percent conversion of **3-1/3-3** or **3-2c/3-3** were monitored *via* the integral ratios of  $-\text{CH}_3/-\text{CH}_3$  (gem dimethyl protons or acetyl protons) as indicated in Figure 3-16.  $^1\text{H}$  NMR spectra of free diol **3-1k** and the control compounds **3-1l** and **3-1m** were compared with  $^1\text{H}$  NMR spectra of their reactions with **3-2c**. A chemical shift was observed for the  $-\text{CH}_3$  group of diol **3-1k** at 0.98 ppm to 0.86 ppm upon conjugate formation **3-3kc** (Figure 3-16C). The percent conversion for control **3-1l** was monitored *via* the integral ratio of the methyl group on the benzylacetyl boronate between **3-2c** and **3-3lc** as shown in Figure 3-16F. A chemical shift was observed for the  $-\text{CH}_3$  of boronic acid **3-2c** from 2.65 ppm to 2.46 ppm upon conjugate **3-3lc** formation. Furthermore, **3-3mc** formation was not observed probably due to the steric effect caused by the butyl group since there was no new peak appearing for this reaction (Figure 3-16H). Moreover, boronate and thiosemicarbazone formation were monitored with HPLC-MS and high-resolution ESI-MS analyses. The results were found to be consistent with  $^1\text{H}$  NMR analysis (Figure 3-17, 3-18). The reaction appeared to proceed through concomitant formation of the boronate and thiosemicarbazone, with full conversion achieved within 3.5 hours at 50  $\mu\text{M}$ , while negative controls **3-1l** and **3-1m** provided diazaborine at a 27% yield and no thiosemicarbazone ligation, respectively (Table 3-3, entry 1-3). These results support the expectation that the initial boronate formation is critical to promote the thiosemicarbazone formation, which occurred second, by intramolecular fashion. It is important to note that the reaction rate was lower in higher buffer concentration (50 mM), which is likely due to the competitive effect of salts in the boronate formation (Table 3-3, entry 1 vs 4).<sup>41</sup> Moreover, a single isomer of the conjugate **3-3kc** was observed with  $^1\text{H}$  NMR and HPLC-MS analyses, yet the geometry (E or Z) of the thiosemicarbazone moiety could not be determined because all attempts to obtain a crystal structure for X-ray analysis were

unsuccessful. Another method is required to identify the geometry. A nOe experiment (expected nOe's: CH<sub>2</sub>-CH<sub>3</sub>, NH-CH<sub>3</sub>) might be helpful to determine the right geometry of the conjugate **3-3kc**.

**Table 3-3:** Results of the aqueous conjugation reaction of **3-1k**, and negative controls **3-1l** and **3-1m**, with **3-2c**.

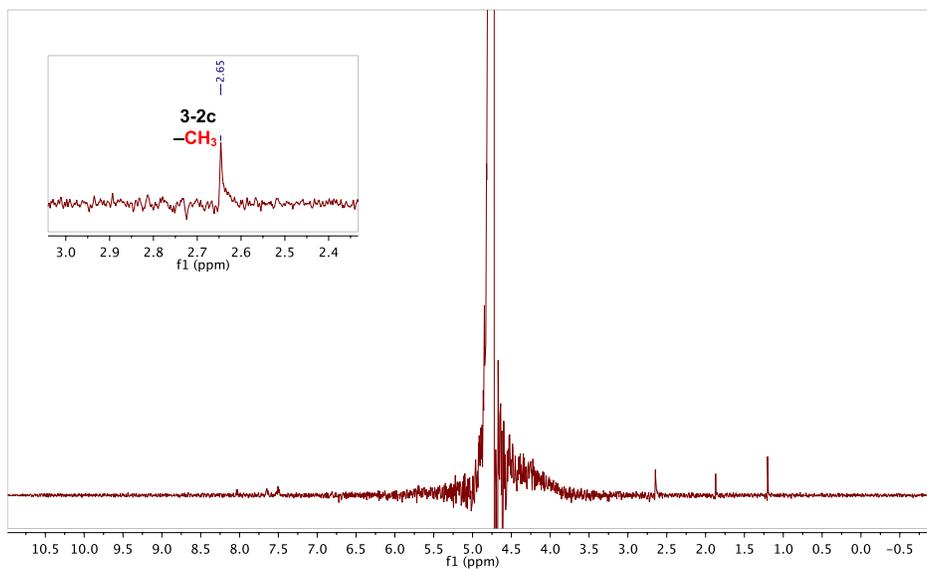


Entry	Conc. of reagents (mM)	% Proportions of <b>3-1/3-3</b>					
		10 min	30 min	60 min	120 min	3.5 h	24 h
1 ( <b>3-1k</b> )	0.050 <sup>[a]</sup>	90/10	86/14	64/36	35/65	0/100	0/100
2 ( <b>3-1l</b> )	0.050 <sup>[a]</sup>	73/27 (maximal conversion)					
3 ( <b>3-1m</b> )	0.050 <sup>[a]</sup>	100/0 (no product)					
4 ( <b>3-1k</b> )	0.25 <sup>[b]</sup>	89/11	81/19	62/38	47/53	-	0/100
5 ( <b>3-1l</b> )	0.25 <sup>[b]</sup>	72/28	64/36	-	64/36	-	64/36

[a] 10 mM PBS buffer/D<sub>2</sub>O (9/1 v/v) was used. [b] ACN-d<sub>3</sub>/50 mM D<sub>2</sub>O phosphate buffer (35/65 w%) was used.

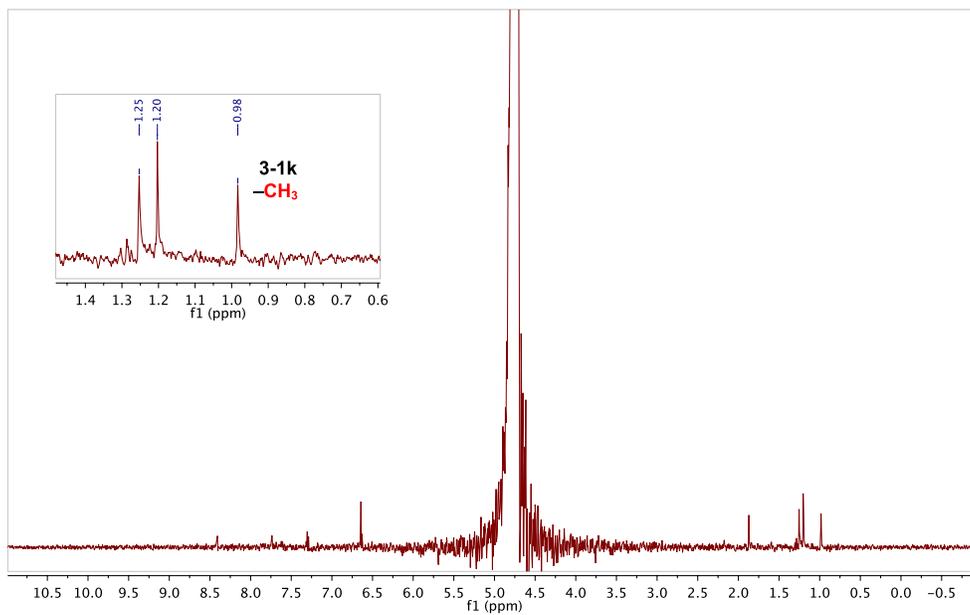
A)

Only boronic acid **3-2c** - in 10 mM PBS/D<sub>2</sub>O (9/1 v/v)  
599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



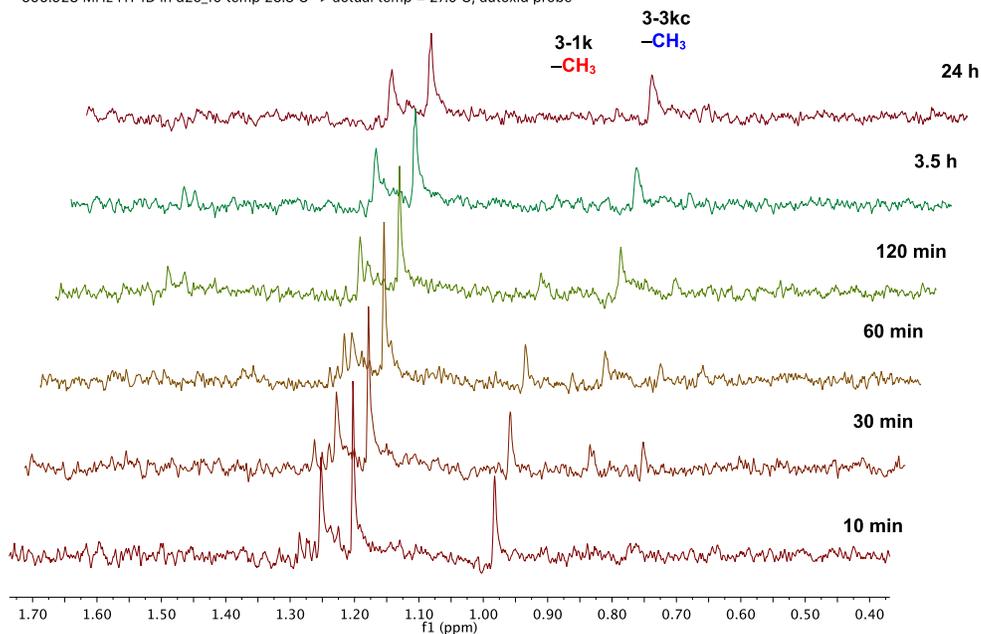
B)

Only diol **3-1k** - <sup>1</sup>H-NMR study of reaction of **3-1k** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v)  
599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



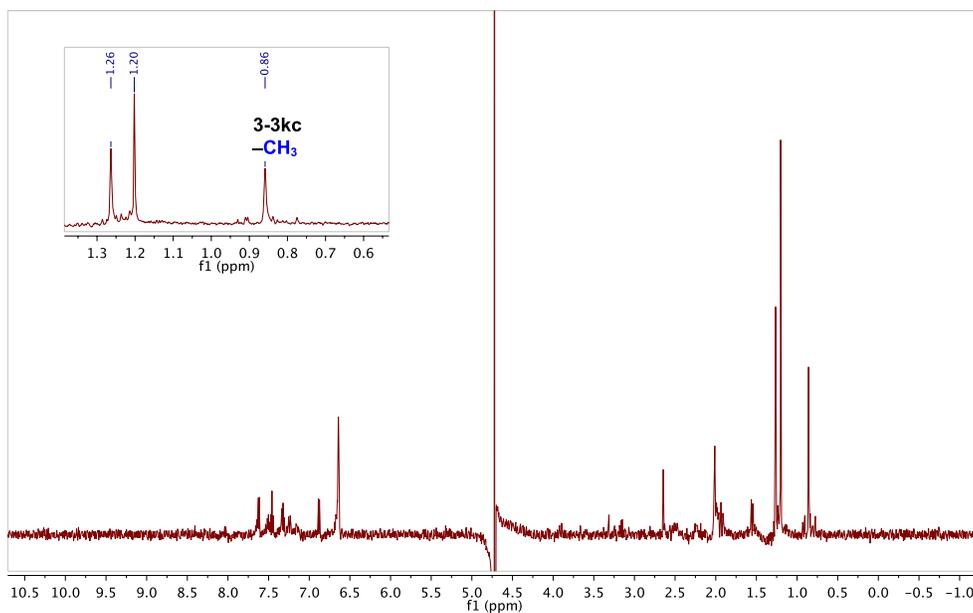
C)

Boronate **3-3kc** -  $^1\text{H}$ -NMR study of the reaction of **3-1k** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v) 599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



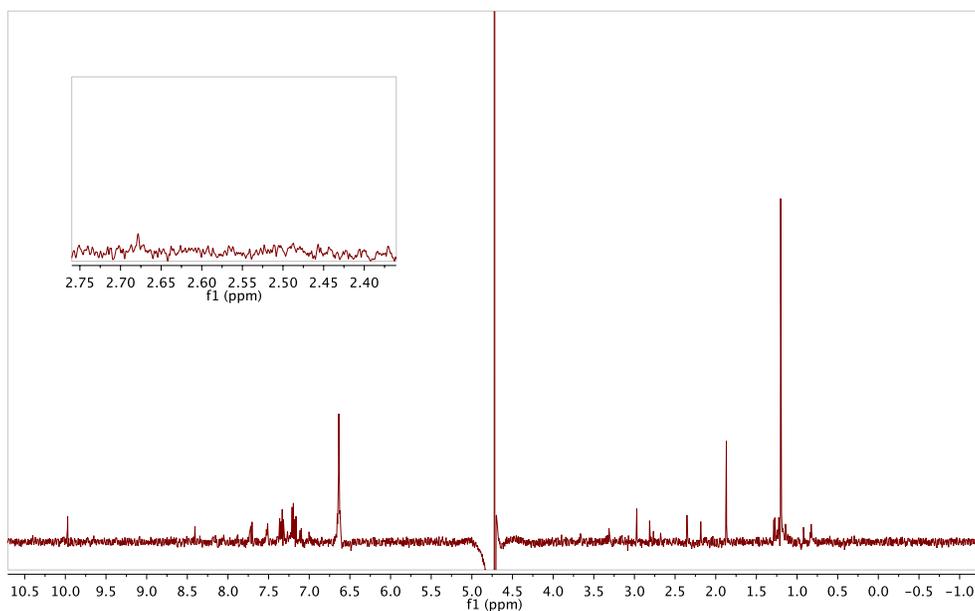
D)

Boronate **3-3kc** - 24 h -  $^1\text{H}$ -NMR study of the reaction of **3-1k** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v) 599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



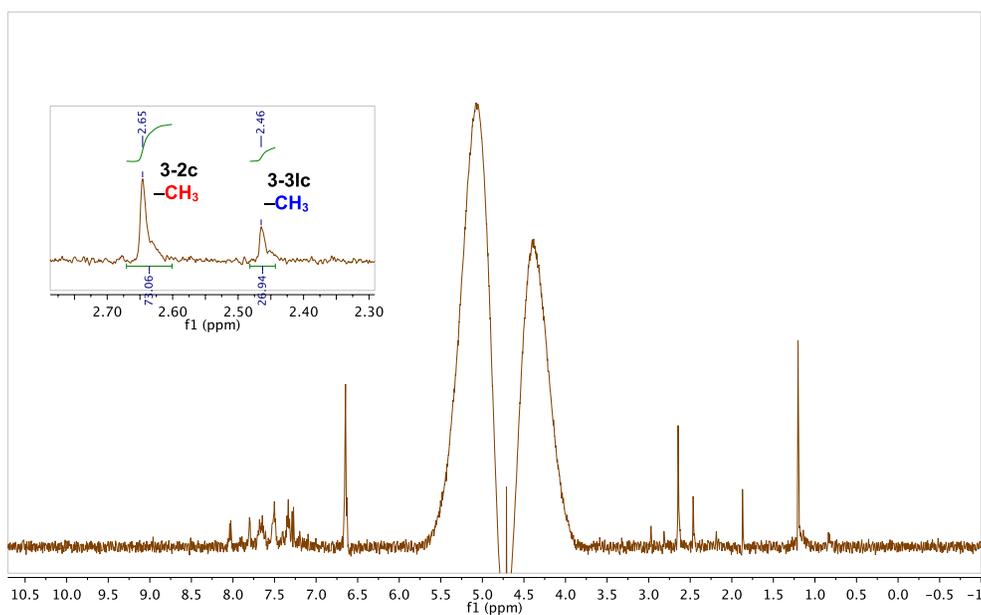
E)

Only **3-1l** -  $^1\text{H-NMR}$  study of reaction of **3-1l** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v)  
599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



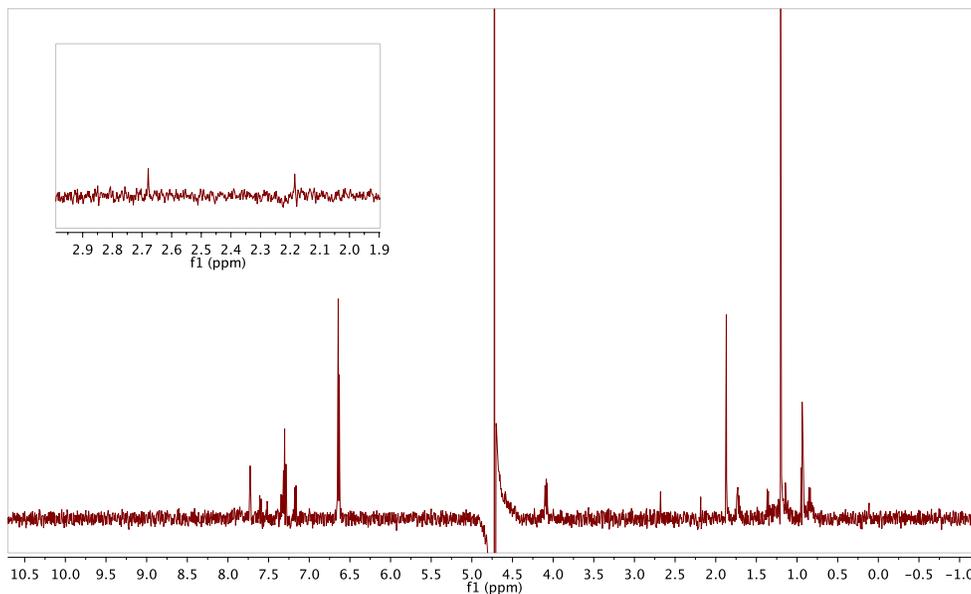
F)

Conjugate **3-3lc** - 24 h -  $^1\text{H-NMR}$  study of reaction of **3-1l** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v)  
599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



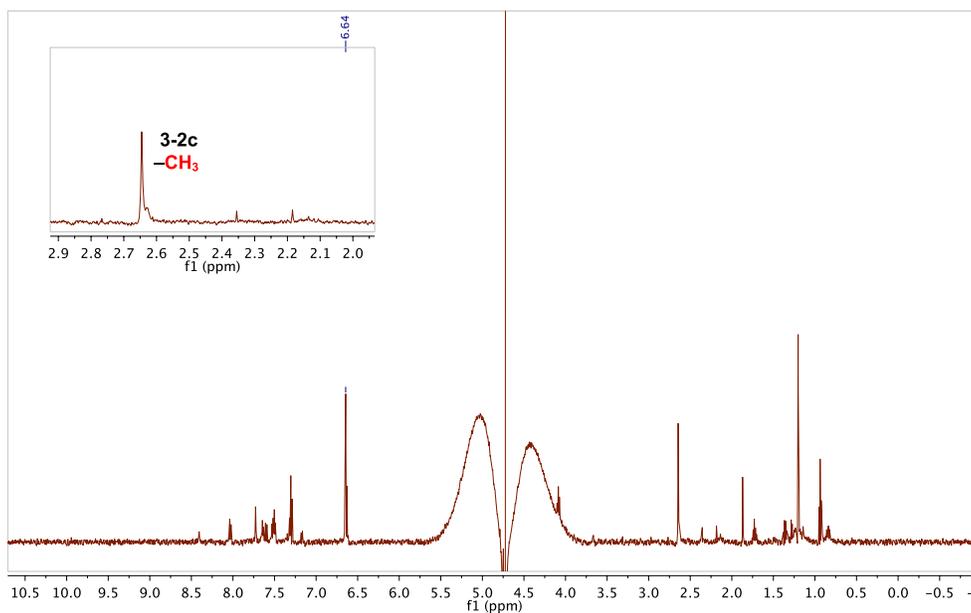
G)

Only **3-1m** -  $^1\text{H}$ -NMR study of the reaction of **3-1m** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v)  
599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe



H)

Conjugate **3-3mc** - 24 h -  $^1\text{H}$ -NMR study of reaction of **3-1m** (0.05 mM) and **3-2c** (0.05 mM) in 10 mM PBS/D<sub>2</sub>O (9/1 v/v)  
599.928 MHz H1 1D in d2o\_10 temp 25.8 C -> actual temp = 27.0 C, autoxid probe

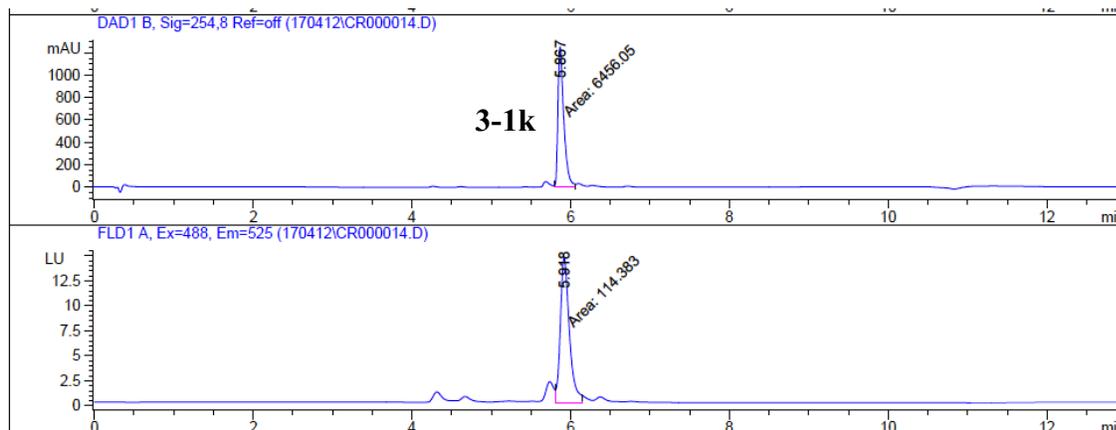


**Figure 3-16:** Representative  $^1\text{H}$  NMR spectra for the reaction of **3-1k/3-1l/3-1m** with **3-2c** in 10 mM PBS buffer/D<sub>2</sub>O (9/1 v/v). A)  $^1\text{H}$  NMR spectrum of boronic acid **3-2c**. B)  $^1\text{H}$  NMR spectrum of only diol **3-1k**. C)  $^1\text{H}$  NMR expanded spectrum of boronate **3-3kc** formation over time. D)  $^1\text{H}$

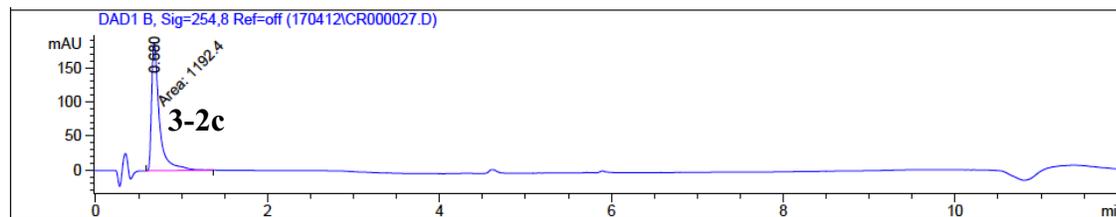
212

NMR spectrum of boronate **3-3kc** after 24 h. E)  $^1\text{H}$  NMR spectrum of only **3-1l**. F)  $^1\text{H}$  NMR spectrum of **3-3lc** formation after 24 h. G)  $^1\text{H}$  NMR spectrum of only **3-1m**. H)  $^1\text{H}$  NMR spectrum of **3-3mc** formation after 24 h.

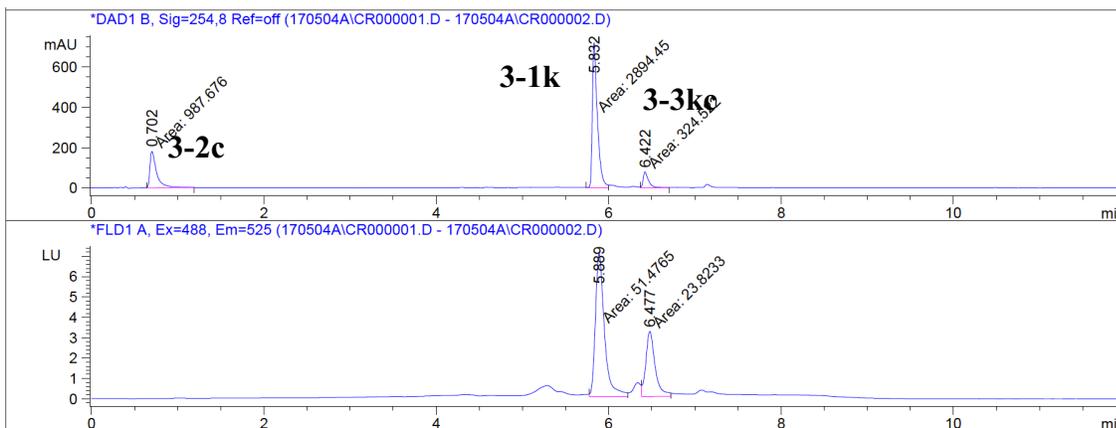
A) Only **3-1k** (50  $\mu\text{M}$ )



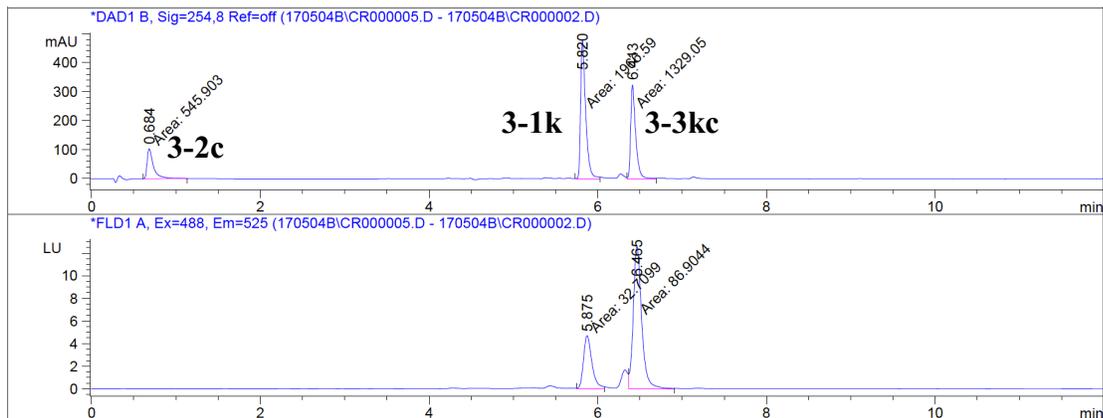
B) Only **3-2c**



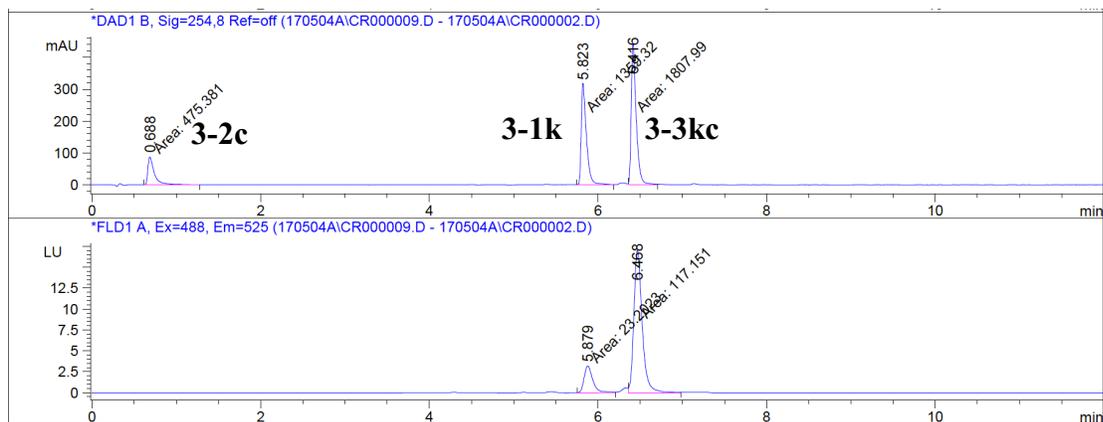
C) The reaction of **3-1k** and **3-2c** in 10 min



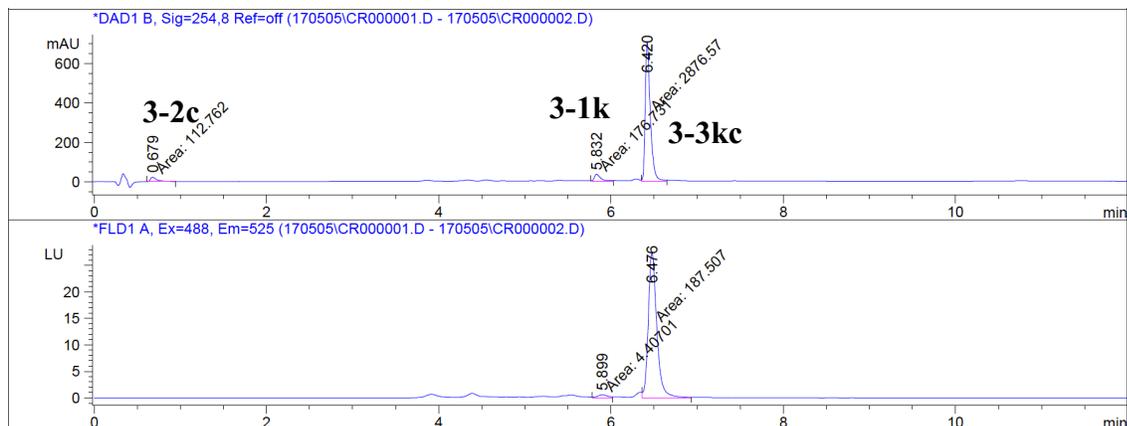
D) The reaction of **3-1k** and **3-2c** in 60 min



E) The reaction of **3-1k** and **3-2c** in 110 min

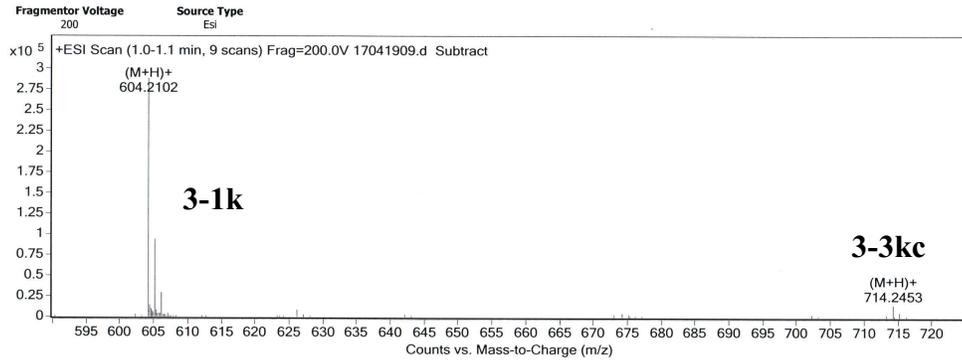


F) The reaction of **3-1k** and **3-2c** in 24 h



**Figure 3-17:** UV and Fluorescence HPLC analyses of the reaction **3-1k** (50  $\mu$ M) and **3-2c** (50  $\mu$ M) in 10 mM PBS buffer.

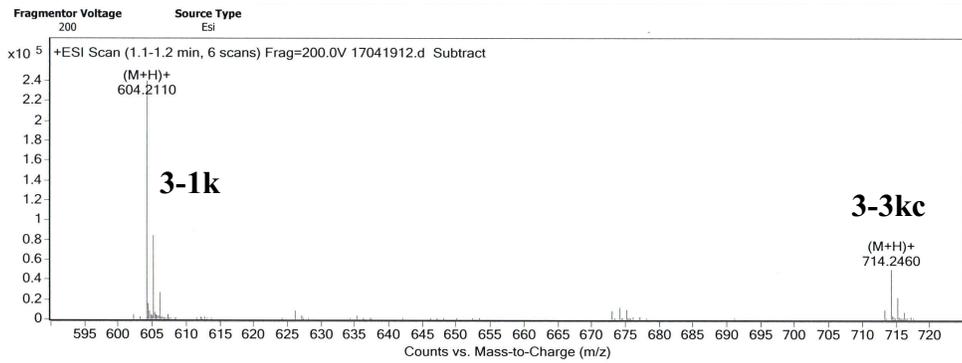
A) 10 min



Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C32 H33 N3 O7 S	C32 H34 N3 O7 S	603.203	603.2039	604.2102	604.2112	0.97	1.6	18	(M+H)+	89
C40 H36 [11B] N3 O7 S	C40 H37 [11B] N3 O7 S	713.238	713.2367	714.2453	714.244	-1.29	-1.81	25	(M+H)+	35.92

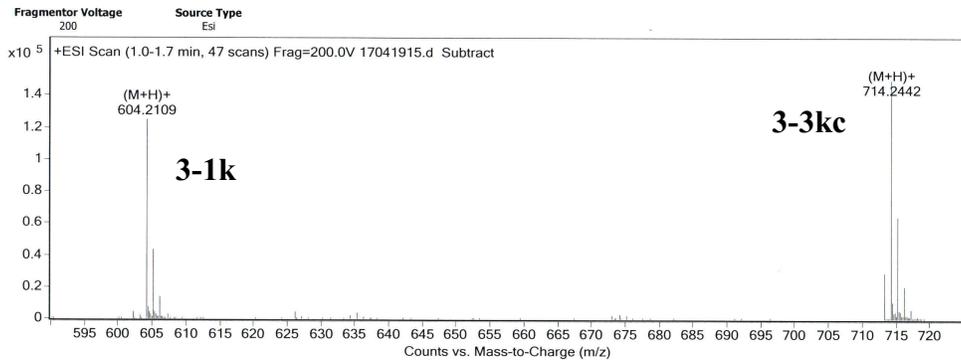
B) 30 min



Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C32 H33 N3 O7 S	C32 H34 N3 O7 S	603.2037	603.2039	604.211	604.2112	0.17	0.29	18	(M+H)+	92.84
C40 H36 [11B] N3 O7 S	C40 H37 [11B] N3 O7 S	713.2387	713.2367	714.246	714.244	-1.96	-2.75	25	(M+H)+	87.6

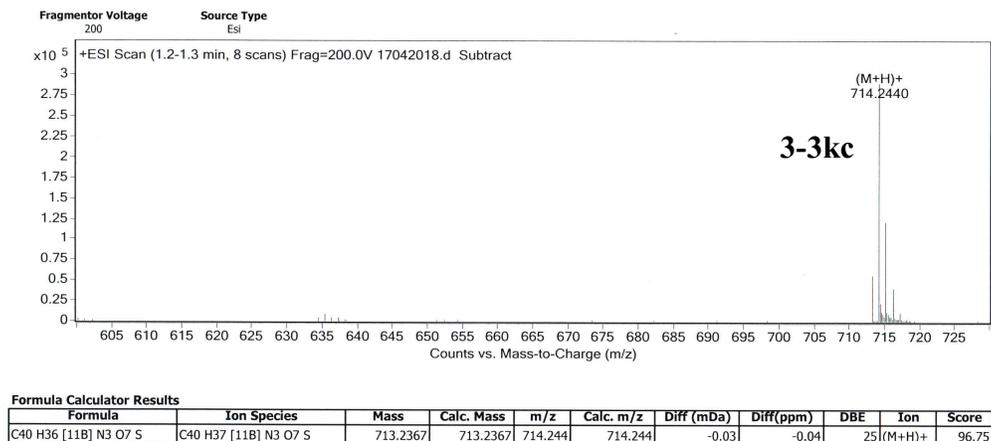
C) 90 min



Formula Calculator Results

Formula	Ion Species	Mass	Calc. Mass	m/z	Calc. m/z	Diff (mDa)	Diff(ppm)	DBE	Ion	Score
C32 H33 N3 O7 S	C32 H34 N3 O7 S	603.2037	603.2039	604.2109	604.2112	0.22	0.37	18	(M+H)+	95.49
C40 H36 [11B] N3 O7 S	C40 H37 [11B] N3 O7 S	713.237	713.2367	714.2442	714.244	-0.27	-0.37	25	(M+H)+	97.25

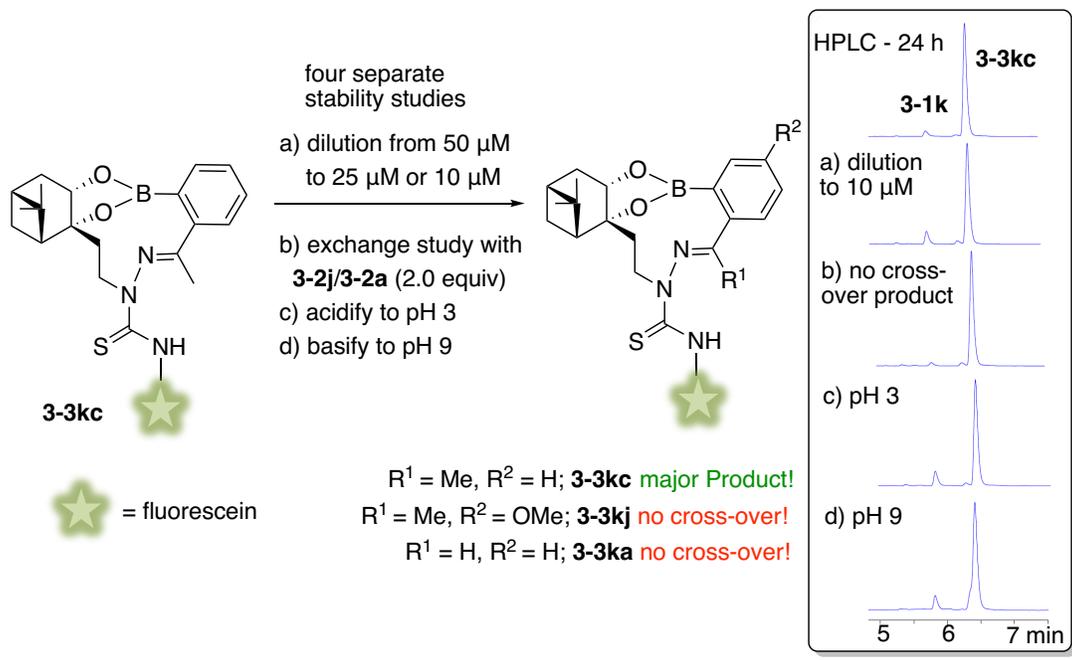
D) 24 h



**Figure 3-18:** High resolution ESI-MS results for the reaction between **3-1k** and **3-2c** in 10 mM ammonium acetate.

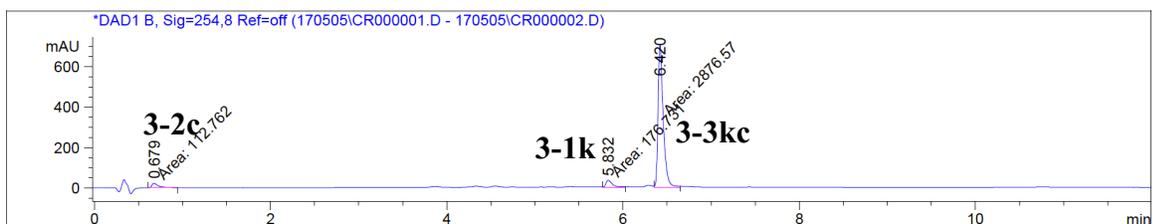
### 3.2.4.2 Stability studies of the conjugate **3-3kc**

In further experimentation, the stability of conjugate **3-3kc** was tested under various conditions *via* HPLC-MS analysis (Figure 3-19). Firstly, equimolar concentrations of **3-1k** and **3-2c** were incubated for one day, and the reaction mixture was subjected to dilutions. Conjugate **3-3kc** stayed intact even after 24 hours (Figure 3-19, 3-20A–C). Subsequently, another boronic acid such as **3-2j** or even the more reactive **3-2a** was introduced. No breakdown of **3-3kc** occurred since no cross-over products of these competitors were observed (Figure 3-19, 3-20D–E). Finally, subjecting **3-3kc** to either acidic (pH 3) or basic (pH 9) solutions did not significantly affect its integrity (Figure 3-19, 3-20F–G).

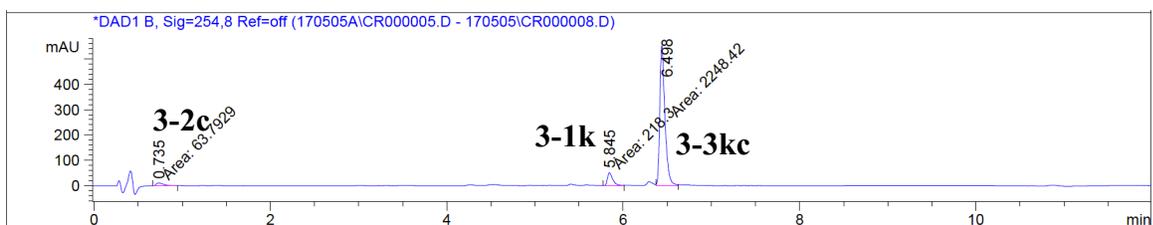


**Figure 3-19:** Studies on the stability of **3-3kc** via HPLC-MS analysis.

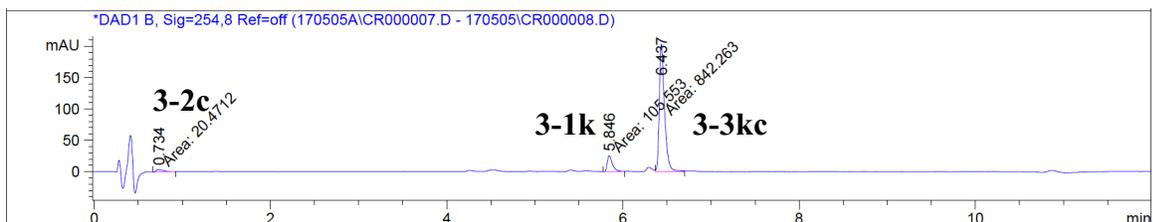
A) The conjugate **3-3kc** before the stability studies in 24 h (50  $\mu\text{M}$ )



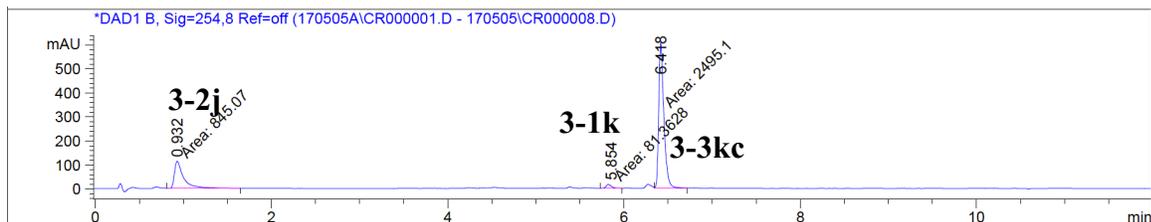
B) Dilution to 50  $\mu\text{M}$  from 250  $\mu\text{M}$  of conjugate **3-3kc**



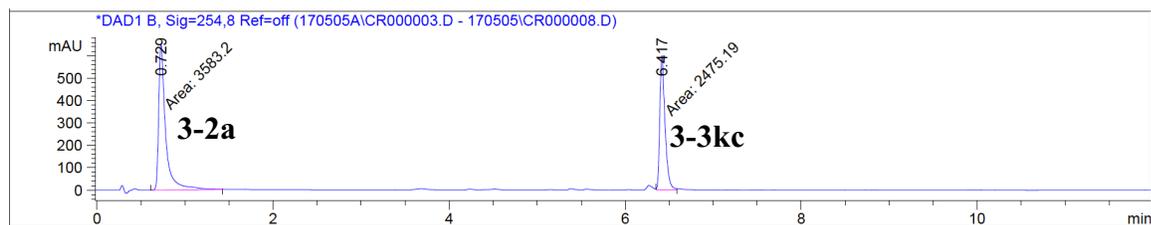
C) Dilution to 10  $\mu\text{M}$



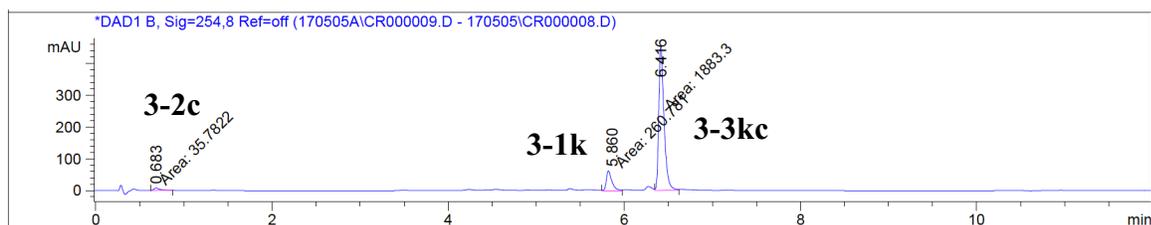
#### D) Exchange study with 3-2j



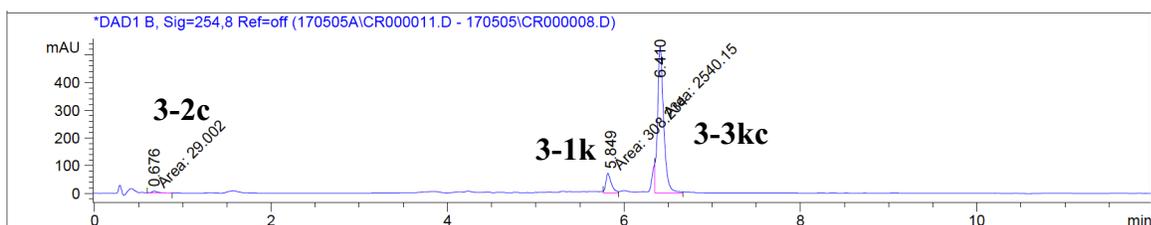
#### E) Exchange study with 3-2a



#### F) Acidification to pH 3



#### G) Basification to pH 9



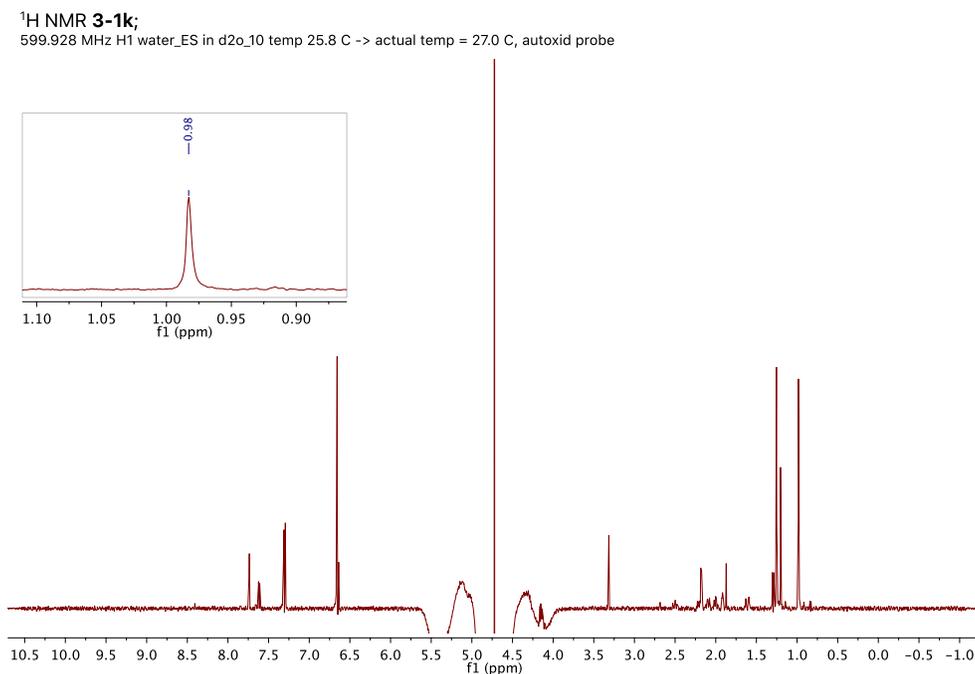
**Figure 3-20:** HPLC-MS chromatograms of stability studies of the conjugate **3-3kc** in 10 mM PBS buffer.

#### 3.2.4.3 Mechanistic study – kinetic studies

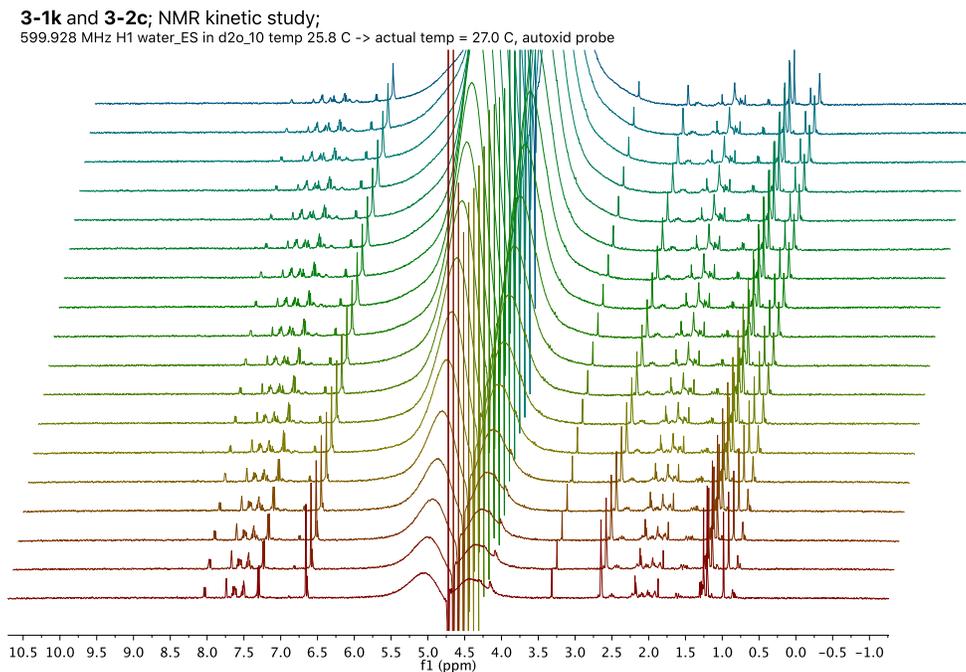
Afterwards, the kinetics of synergic boronate and thiosemicarbazone formation between diol **3-1k** and **3-2c** was studied through  $^1\text{H}$  NMR analysis at  $25^\circ\text{C}$  in 10 mM PBS buffer/ $\text{D}_2\text{O}$

(9/1 v/v). The disappearance of the  $-\text{CH}_3$  group of diol **3-1k** at 0.98 ppm and the appearance of the  $-\text{CH}_3$  group of **3-3kc** at 0.86 ppm were monitored at every 320 second intervals over 84 min and converted into the concentration of diol **3-1k** (Figure 3-21). The second order rate constant was determined by plotting  $1/[\mathbf{3-1k}]$  vs time and the slope of this graph was recorded as the second order rate constant (Figure 3-22). The overall rate constant for the formation of **3-3kc** in 10 mM PBS buffer was measured as  $4.2 \pm 0.4 \text{ M}^{-1}\text{s}^{-1}$ , which is in agreement with our previous study on rate constant measurements of nopoldiol boronate formation.<sup>14</sup> Moreover, negative control **3-1m** without a 1,2-*cis* diol gave no conversion (Table 3-3, entry 3). Thus boronate formation is essential in this system, acting as the overall rate-determining step preceding a fast intramolecular thiosemicarbazone formation.

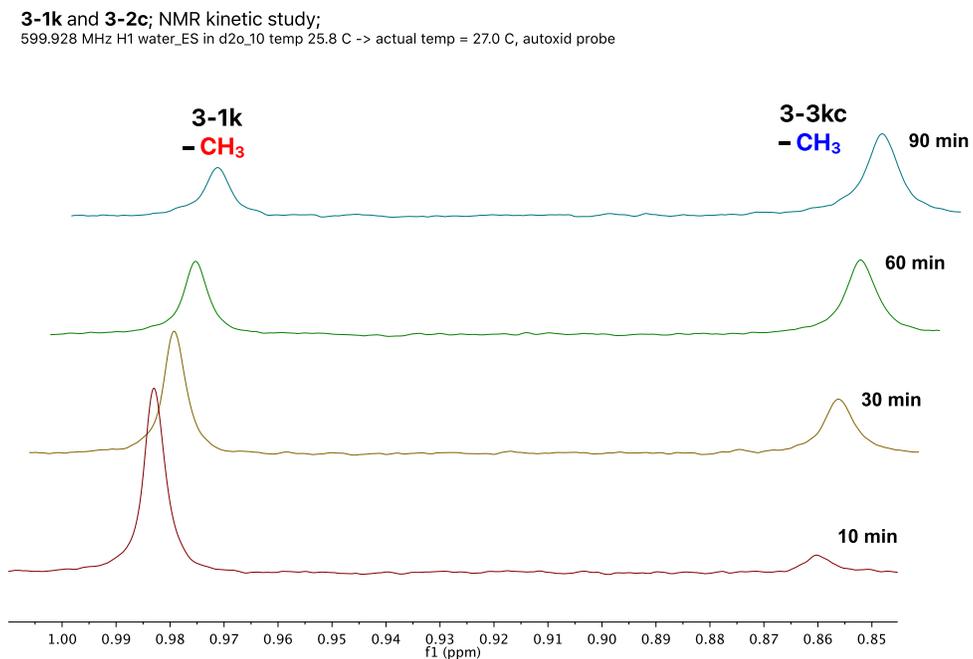
A)



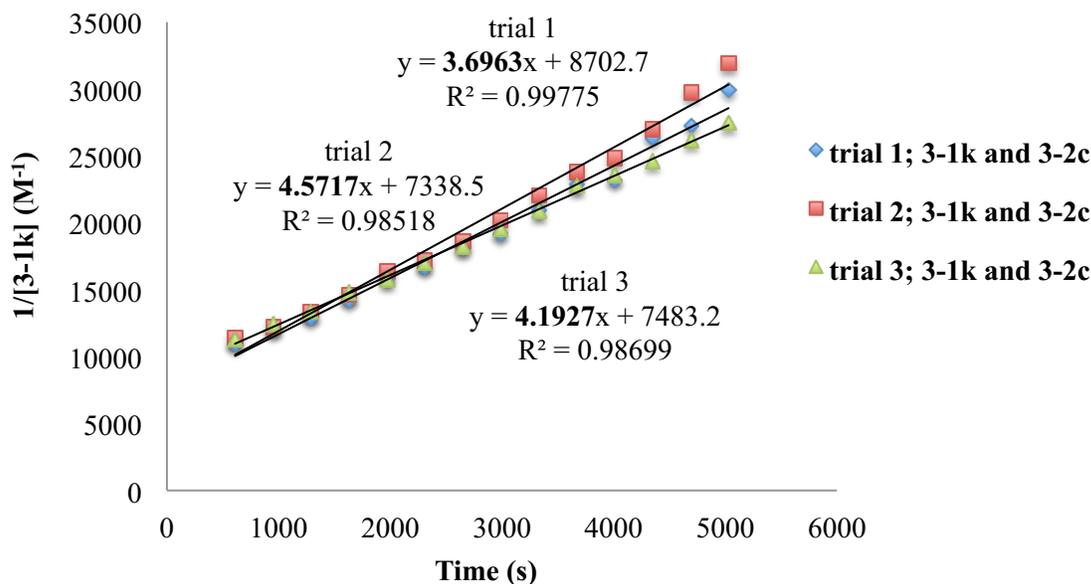
B)



C)



**Figure 3-21:** Representative time dependent  $^1\text{H}$  NMR analysis for the reaction of diol **3-1k** and **3-2c**. A)  $^1\text{H}$  NMR spectrum of **3-1k** in corresponding buffer system. B) Full spectrum. C) Diol **3-1k** and boronate **3-3kc** methyl region.

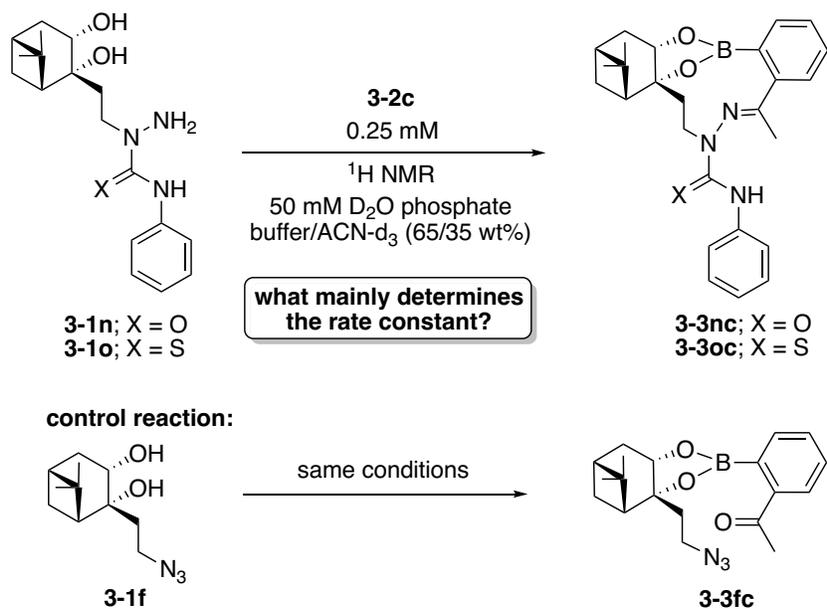


**Figure 3-22:** Rate constant measurements, second order kinetics plot for **3-1k** and **3-2c**.

To understand the reaction mechanism further, the semicarbazide diol **3-1n** and the thiosemicarbazide diol **3-1o** were designed and synthesized by nucleophilic addition of *in-situ* prepared nopoldiol hydrazine to phenylisocyanate and phenylisothiocyanate, respectively. Nopoldiol azide **3-1f**, which can only afford boronate formation, was used as a control. Using <sup>1</sup>H NMR analysis, boronate and thiosemicarbazone formation of these diols (0.25 mM) with **3-2c** (0.25 mM) were investigated in 0.05 M of D<sub>2</sub>O potassium phosphate buffer/ACN-d<sub>3</sub> (65/35 w%, pD 7.4). It was found that diol **3-1n** and **3-1o** initially formed hemiaminals **3-3nc•H<sub>2</sub>O** and **3-3oc•H<sub>2</sub>O** followed by (thio)semicarbazone formation **3-3nc** and **3-3oc** (Figure 3-23). When the <sup>1</sup>H NMR spectrum of boronic acid **3-2c** (Figure 3-23A) was compared with <sup>1</sup>H NMR spectra of conjugate **3-3nc** and **3-3oc** (Figure 3-23C, D, F and G), a change of chemical shift was observed for aryl -CH of **3-2c** at 7.92 ppm (doublet) to 7.96 ppm (doublet, **3-3nc•H<sub>2</sub>O**) or 7.97 ppm (doublet, **3-3oc•H<sub>2</sub>O**). Then, these peaks disappeared and a new peak appeared at 6.68 ppm (doublet, **3-3nc**) or 6.52 ppm (doublet, **3-3oc**). Therefore, percent proportions were determined as the integral ratio of aryl -CH/-CH, **3-2c**/(**3-3nc•H<sub>2</sub>O** + **3-3nc**) and **3-2c**/(**3-3oc•H<sub>2</sub>O** + **3-3oc**) (Table 3-4, Figure 3-23). It was also observed that diol **3-1o** provided faster thiosemicarbazone formation than diol **3-1n** (Table 3-4, entry 1 vs 2). In addition, percent proportions for **3-1f**/**3-3fc** were monitored *via* integral ratio of -CH<sub>3</sub>/-CH<sub>3</sub> as indicated in Figure 3-23I. The results indicate

that both **3-1n** and **3-1o** behave similarly and reach 100% conversion with the help of the synergic (thio)semicarbazone interaction. In contrast, control diol **3-1f** afforded a conversion similar to **3-1n** and **3-1o** in the first 60 min and then remained in equilibrium at a 53% conversion as expected (Table 3-4, entry 1-3). Hence, the reaction rate and conversion in this system is determined mainly by boronate formation.

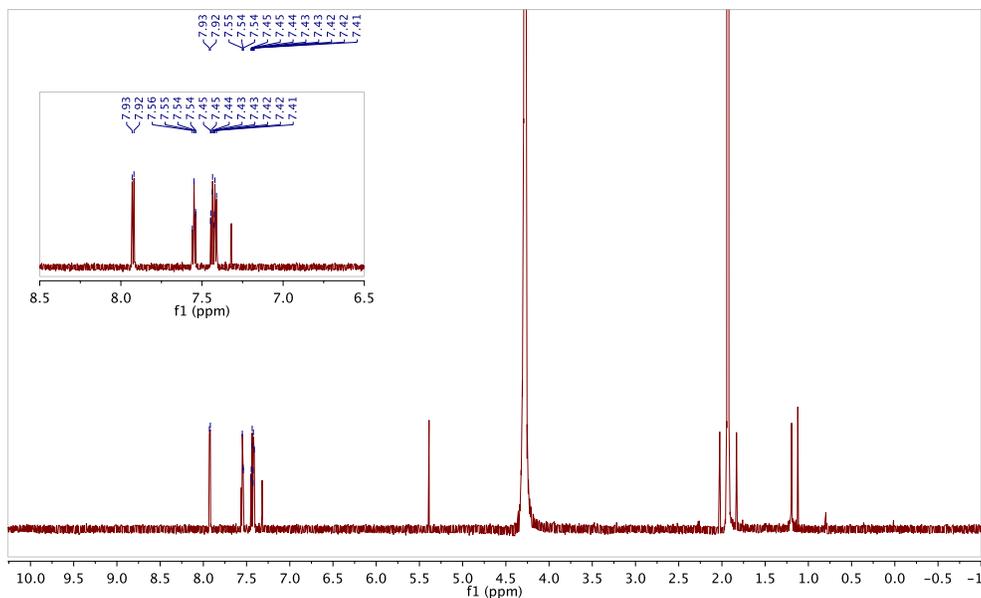
**Table 3-4:** Examination of the reactivity of **3-1n**, **3-1o** and control **3-1f** with **3-2c** to identify the rate determining step.



Entry	% Proportions of <b>3-1/3-3</b>							
	10 min	30 min	60 min	90 min	120 min	3.0 h	3.5 h	24 h
1 ( <b>3-1n</b> )	82/18	70/30	55/45	49/51	47/53	-	38/62	0/100
2 ( <b>3-1o</b> )	90/10	71/29	53/47	44/56	38/62	30/70	-	0/100
3 ( <b>3-1f</b> )	82/18	67/33	58/42	54/46	52/48	49/51	-	47/53

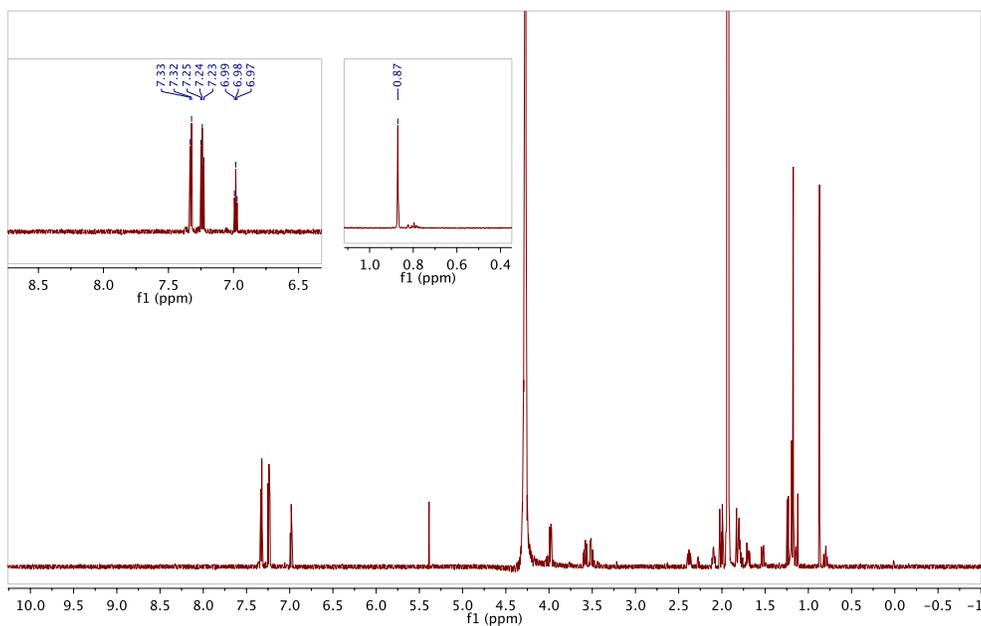
A)

$^1\text{H}$  NMR spectrum of boronic acid **3-2c**;  
699.764 MHz H1 1D in  $\text{d}_2\text{o}$  (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



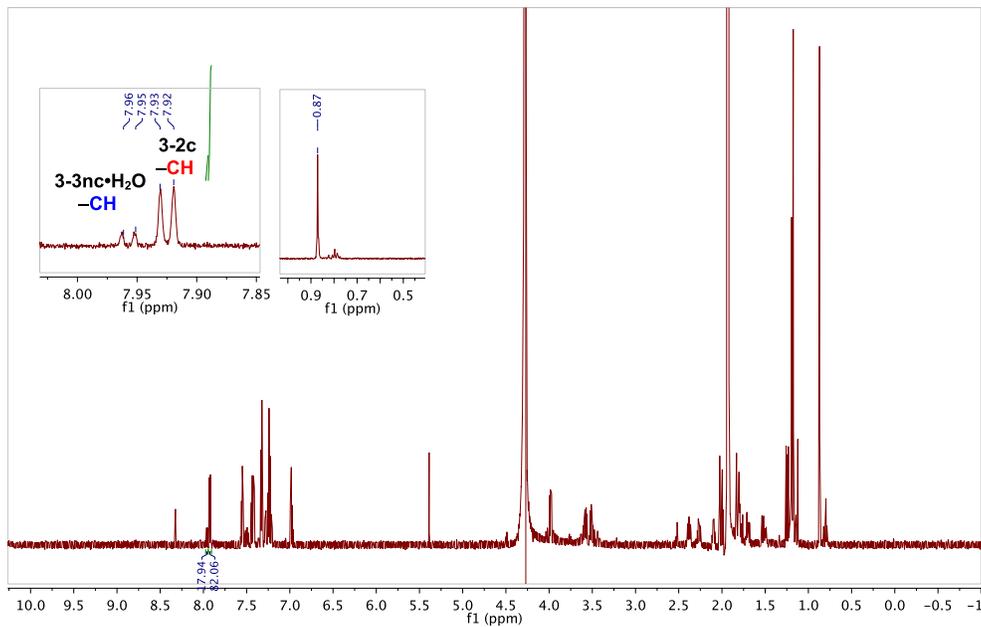
B)

$^1\text{H}$  NMR spectrum of diol **3-1n**;  
699.764 MHz H1 1D in  $\text{d}_2\text{o}$  (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



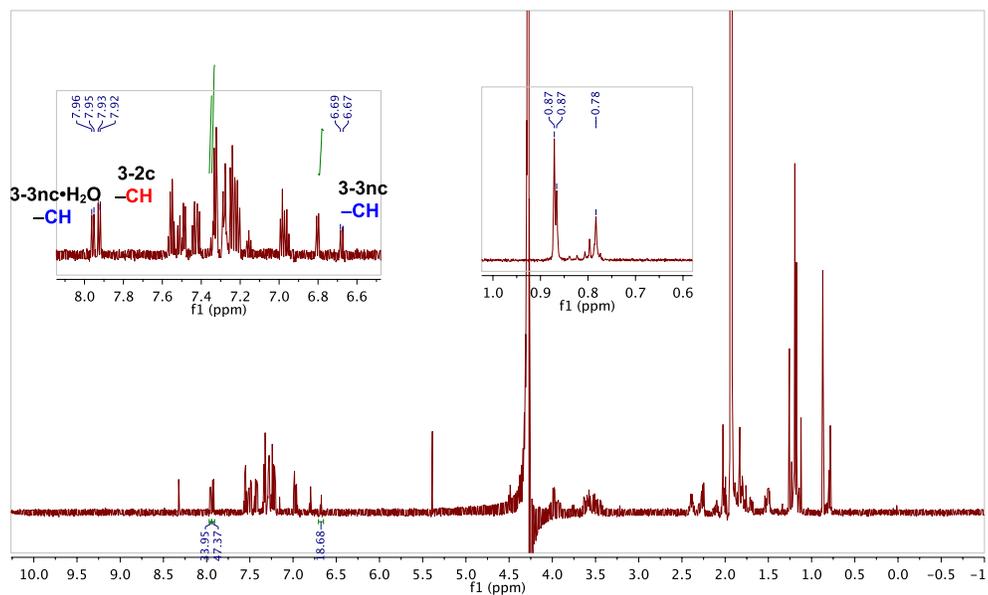
C)

Boronate **3-3nc** – 10 min –  $^1\text{H}$  NMR study of reaction of **3-1n** and **3-2c**  
 699.764 MHz H1 1D in  $\text{d}_2\text{o}$  (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



D)

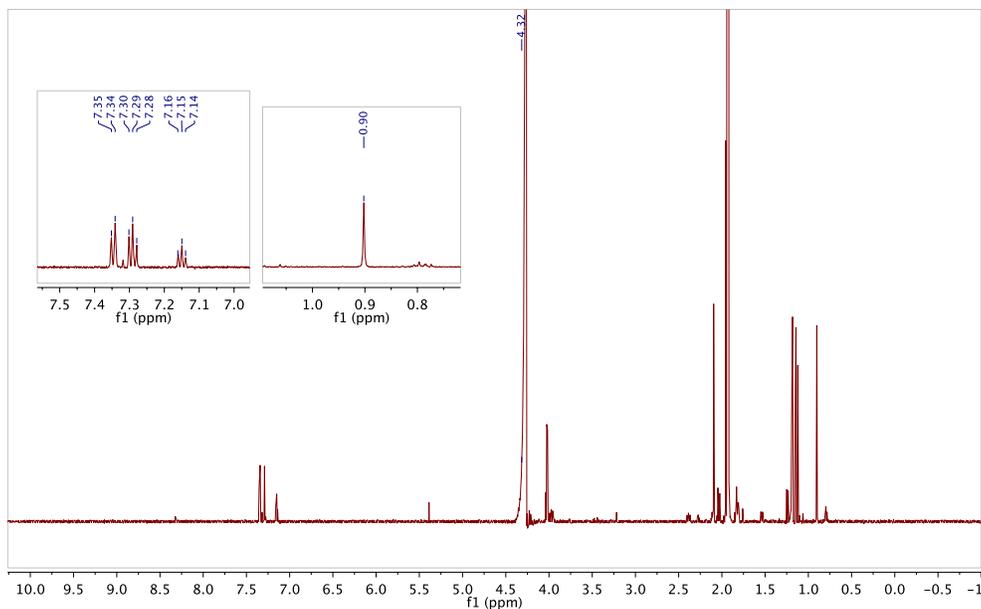
Boronate **3-3nc** – 120 min –  $^1\text{H}$  NMR study of reaction of **3-1n** and **3-2c**  
 699.764 MHz H1 1D in  $\text{d}_2\text{o}$  (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



E)

$^1\text{H}$  NMR spectrum of diol **3-1o**;

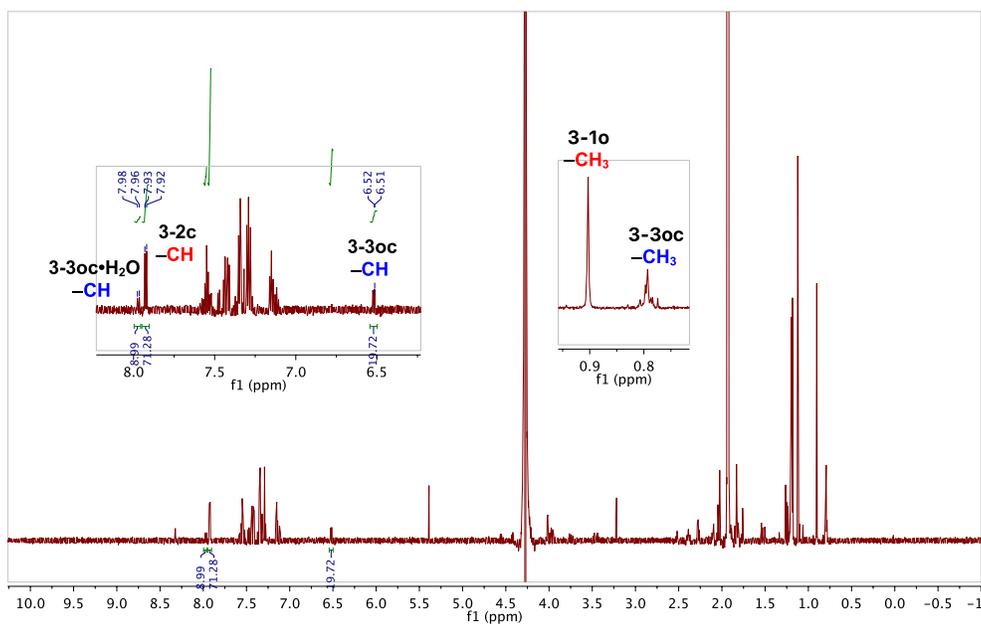
699.764 MHz H1 1D in  $\text{d}_2\text{o}$  (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



F)

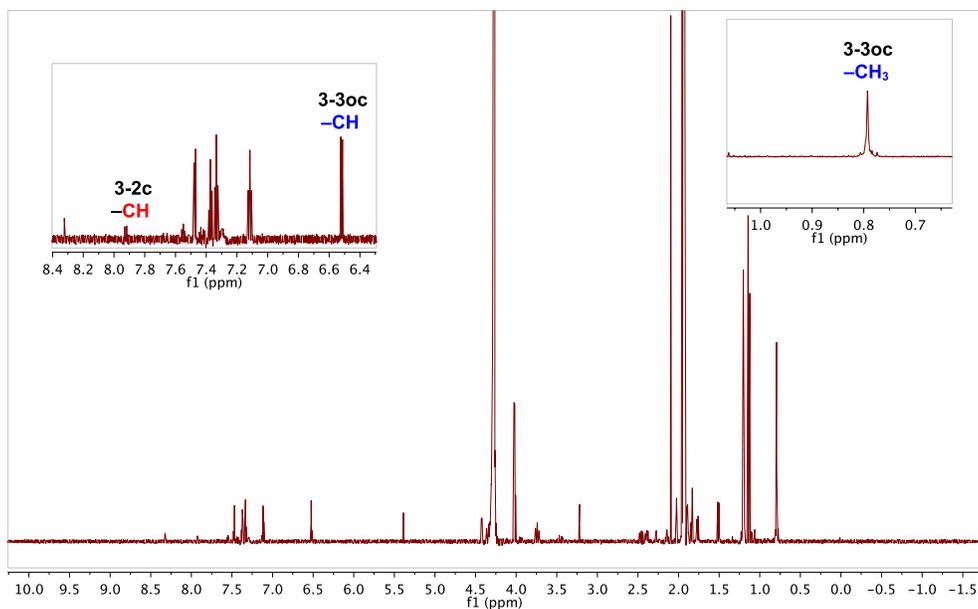
Boronate **3-3oc** – 30 min –  $^1\text{H}$  NMR study of reaction of **3-1o** and **3-2c**

699.764 MHz H1 1D in  $\text{d}_2\text{o}$  (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



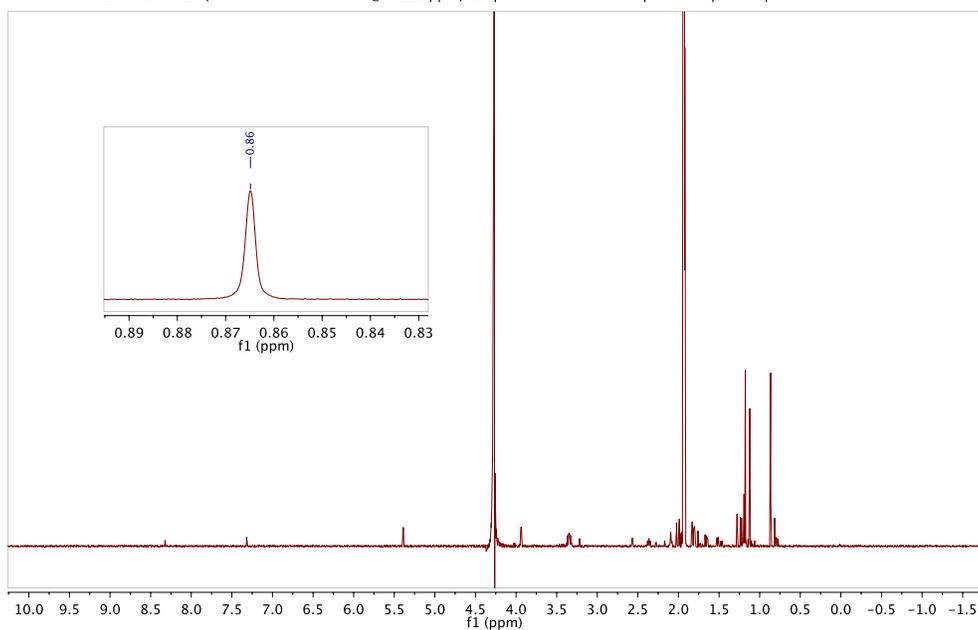
G)

Boronate **3-3oc** – 24 h –  $^1\text{H}$  NMR study of reaction of **3-1o** and **3-2c**  
699.764 MHz H1 1D in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



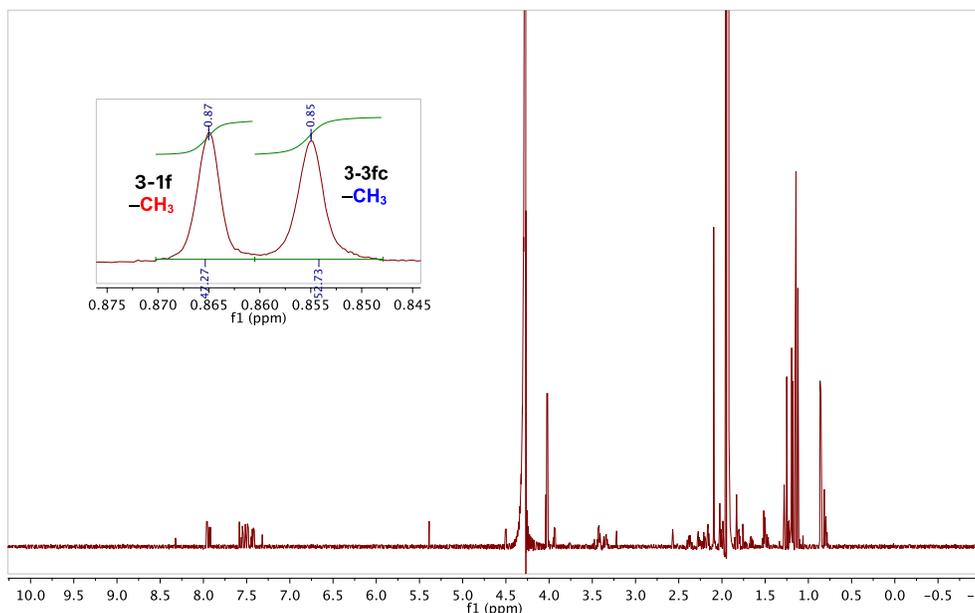
H)

$^1\text{H}$  NMR spectrum of diol **3-1f**;  
699.764 MHz H1 1D in d<sub>2</sub>o (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



I)

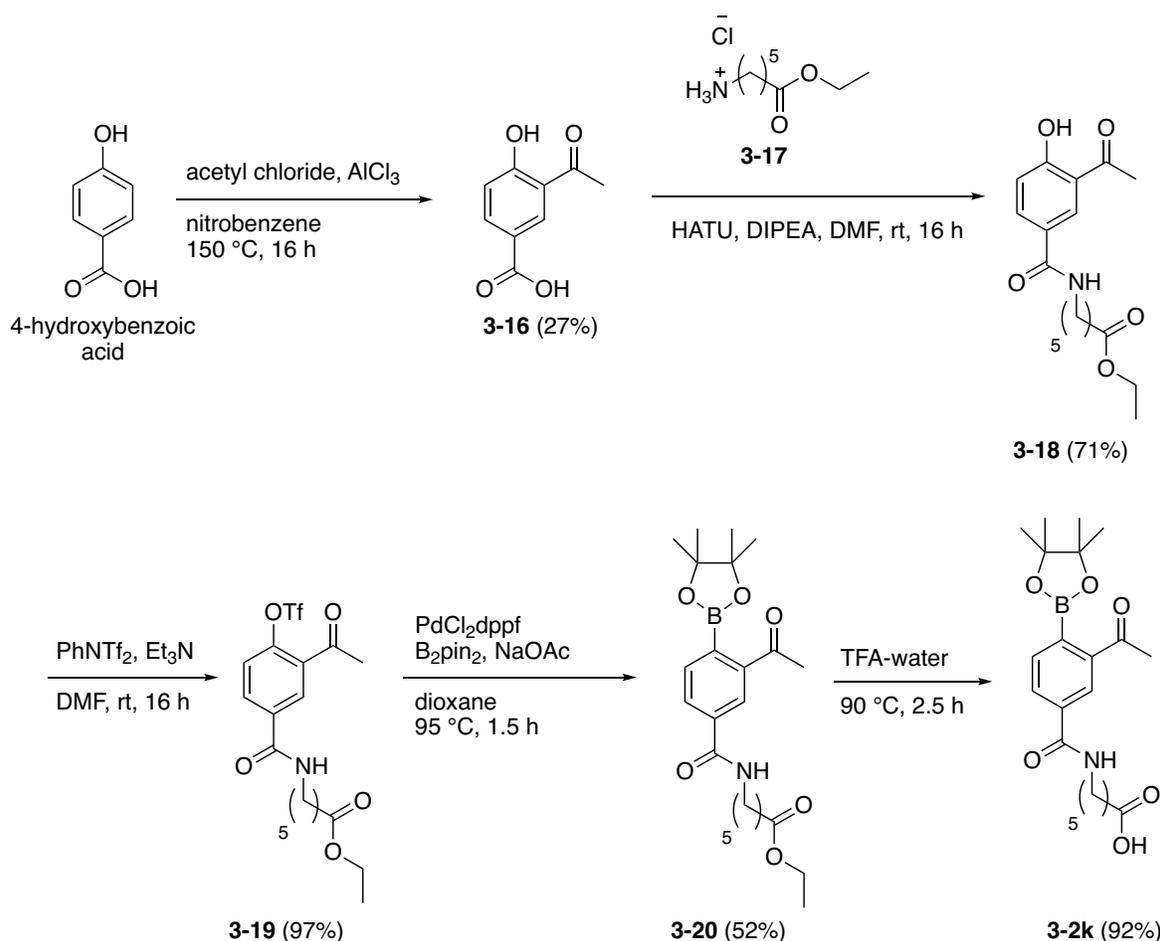
Boronate **3-3fc** – 24 h –  $^1\text{H}$  NMR study of reaction of **3-1f** and **3-2c**  
699.764 MHz H1 1D in d2o (ref. to external acetone @ 2.225 ppm) temp 27.5 C -> actual temp = 27.0 C, coldid probe



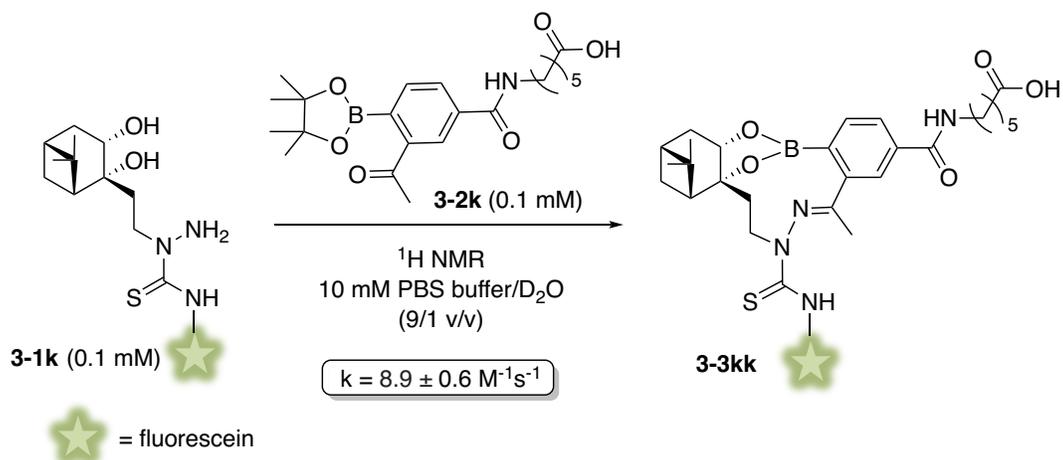
**Figure 3-23:** Representative  $^1\text{H}$  NMR spectra for mechanistic studies. A)  $^1\text{H}$  NMR spectrum of boronic acid **3-2c**. B)  $^1\text{H}$  NMR spectrum of diol **3-1n**. C)  $^1\text{H}$  NMR spectrum of boronate **3-3nc** in 10 min. D)  $^1\text{H}$  NMR spectrum of boronate **3-3nc** in 120 min. E)  $^1\text{H}$  NMR spectrum of diol **3-1o**. F)  $^1\text{H}$  NMR spectrum of boronate **3-3oc** in 30 min. G)  $^1\text{H}$  NMR spectrum of boronate **3-3oc** in 24 h. H)  $^1\text{H}$  NMR spectrum of diol **3-1f**. I)  $^1\text{H}$  NMR spectrum of boronate **3-3fc** in 24 h.

According to our previous study on nopoldiol boronate formation,<sup>14</sup> the reactivity of this synergic conjugation system could be improved *via* the placement of electron-withdrawing substituents on **3-2c**. Therefore, a conjugatable *o*-acetylphenylpinacol boronate with a *p*-amide group, **3-2k**, was designed and synthesized as shown in Scheme 3-23. Compounds **3-16** and **3-17** were easily synthesized based on the procedures described in the literature.<sup>42</sup> Then, an amide coupling of the carboxyl group of **3-16** and amine unit of **3-17** was performed to synthesize the product **3-18**, which was then converted into the triflate derivative **3-19**. Next, intermediate **3-19** was subjected to Suzuki-Miyaura borylation conditions to isolate the borylated product **3-20**.<sup>43</sup> The conjugatable *o*-acetylphenylpinacol boronate with a *p*-amide group, **3-2k** was finally obtained by a saponification of the ethyl ester moiety of **3-20**. Using  $^1\text{H}$  NMR analysis, the reaction between diol **3-1k** and **3-2k** was also monitored at 25 °C in 10 mM PBS buffer/D<sub>2</sub>O (9/1

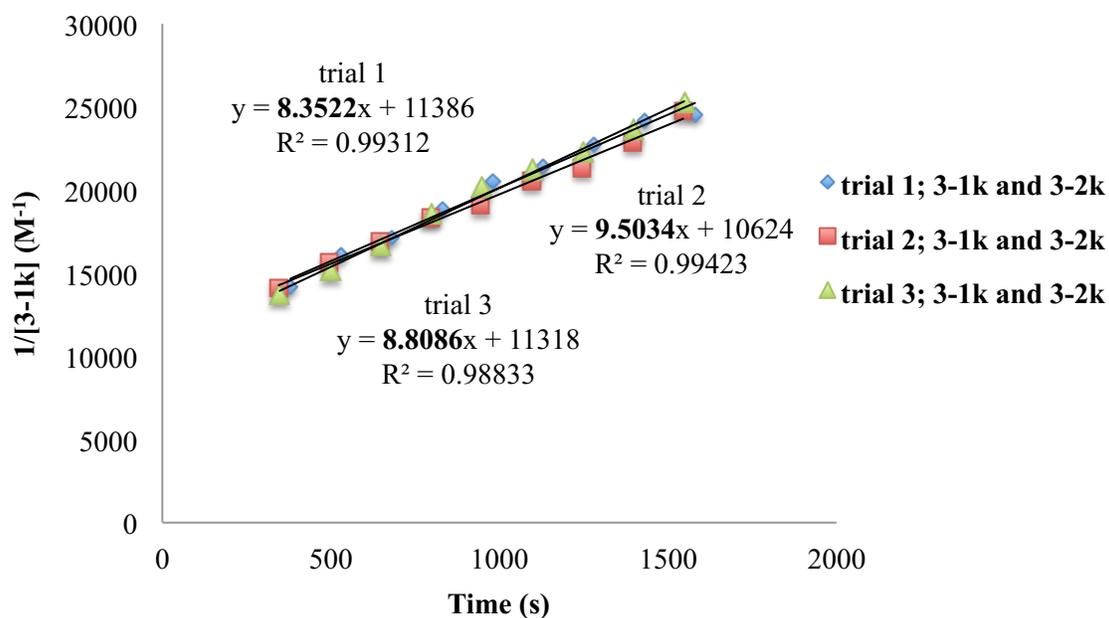
v/v). The disappearance of  $-\text{CH}_3$  of **3-1k** at 0.98 ppm and the appearance of  $-\text{CH}_3$  of **3-3kk** at 0.86 ppm were monitored at every 150 seconds over 45 min. The data were converted into the concentration of diol **3-1k** in the reaction. The second order rate constant was determined by plotting  $1/[\mathbf{3-1k}]$  vs time. As predicted, the rate constant increased to  $8.9 \pm 0.6 \text{ M}^{-1}\text{s}^{-1}$  (Scheme 3-24, Figure 3-24). Not surprisingly, the presence of a pinacol protecting group in **3-3k** did not impede the formation of conjugate **3-3kk** because arylboronic acid pinacol esters are known to hydrolyze partially in aqueous environment.<sup>26,44</sup>



**Scheme 3-23:** Synthesis of the conjugatable *o*-acetylphenylpinacol boronate with a *p*-amide group **3-2k**.



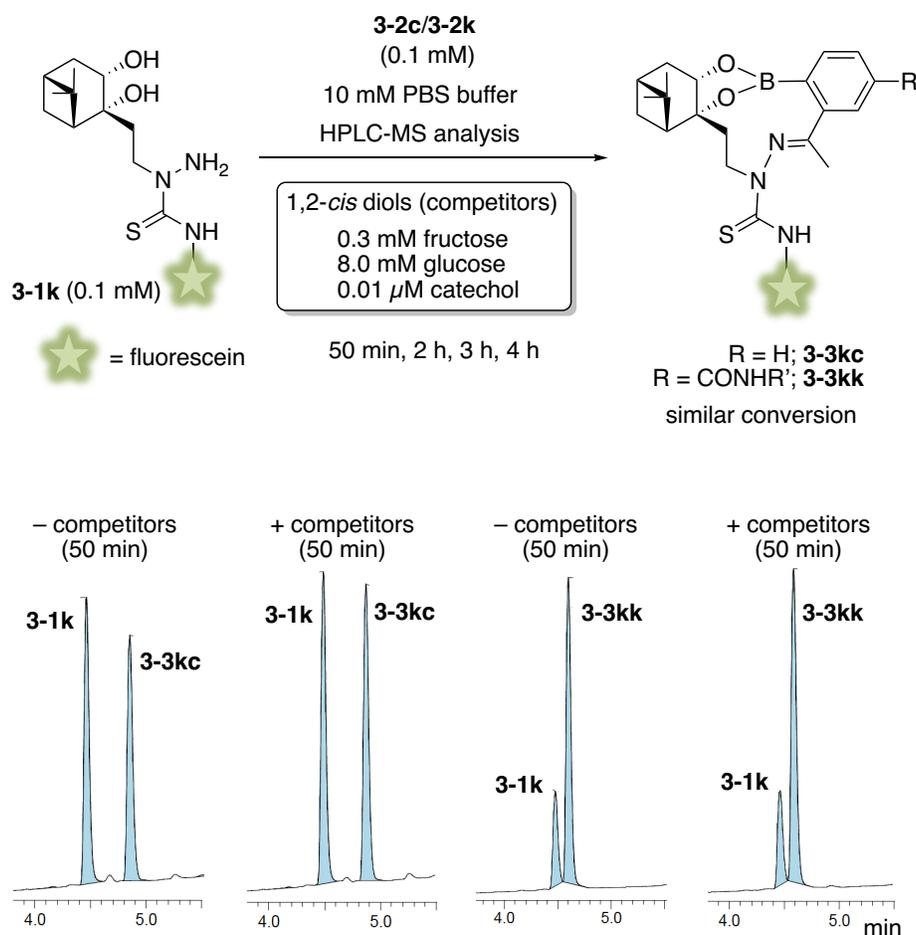
**Scheme 3-24:** Rate constant for the reaction between **3-1k** and **3-2k** as measured by  $^1\text{H}$  NMR spectroscopy.



**Figure 3-24:** Rate constant measurements, second order kinetics plot for **3-1k** and **3-2k**.

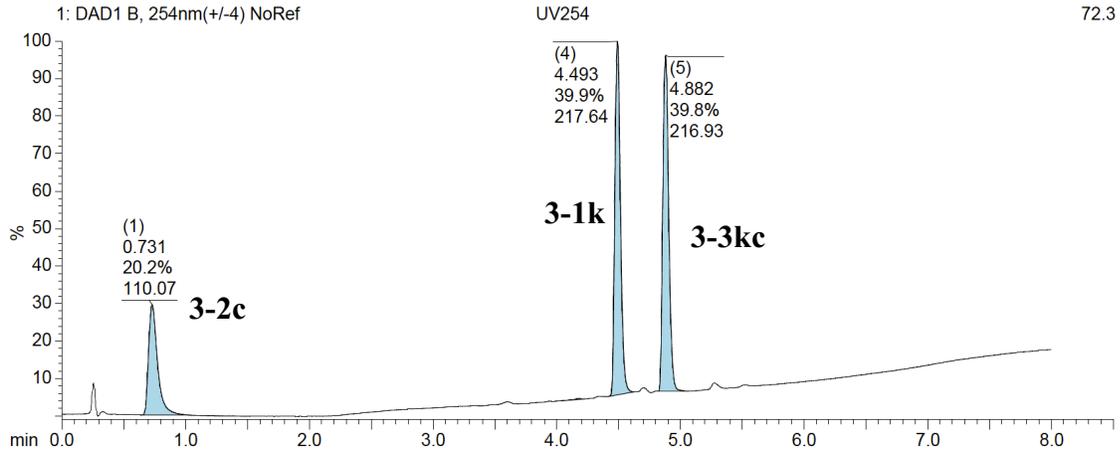
### 3.2.4.4 Studies to assess the bioorthogonality of the boronate/thiosemicarbazone system

True bioorthogonal reactions are minimally affected by the biomolecules present in biological milieu. The effect of biological polyols on the reaction progress was assessed using competition assays monitored by HPLC-MS analysis. Diol **3-1k** was allowed to form the irreversible conjugate **3-3kc/3-3kk** with **3-2c/3-2k** in the presence of a mixture of biocompetitors such as glucose (8 mM), fructose (300  $\mu$ M) or catechol (0.01  $\mu$ M) at concentrations higher than those found in the blood stream (Figure 3-25).<sup>45,46</sup> Remarkably, the formation of the desired conjugate **3-3kc/3-3kk** was not affected as similar conversions were obtained both in the absence and the presence of biological polyols (Figure 3-25, 3-26). As expected, higher conversion was observed for **3-3kk** when compared to **3-3kc** due to the higher reactivity of boronic acid **3-2k** (Figure 3-25).

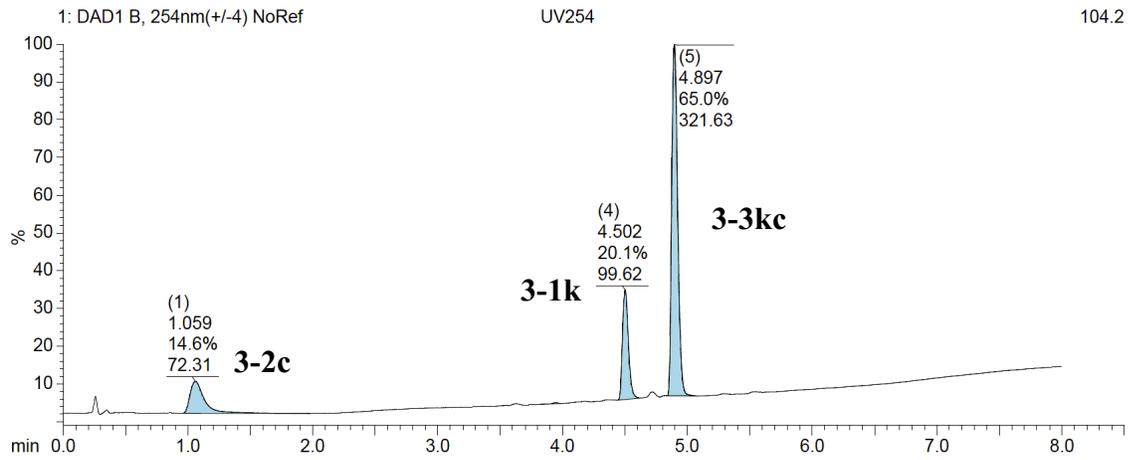


**Figure 3-25:** Competitive effect of biological polyols on the formation of conjugates **3-3kc/3-3kk**.

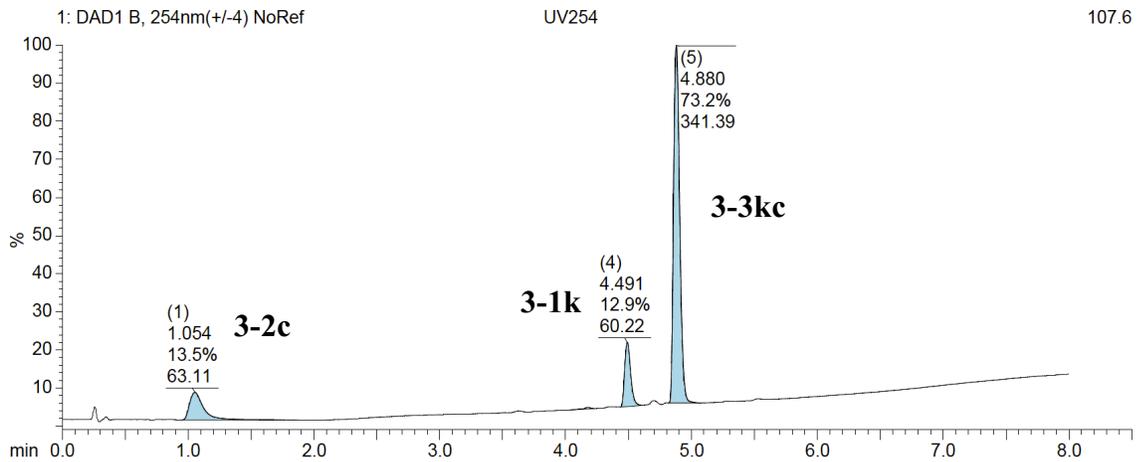
A) 50 min



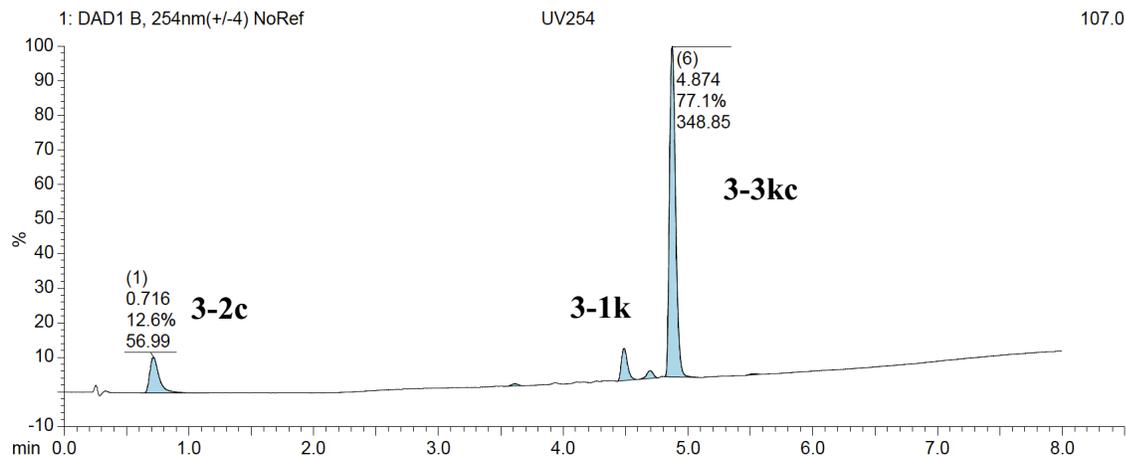
B) 2 h



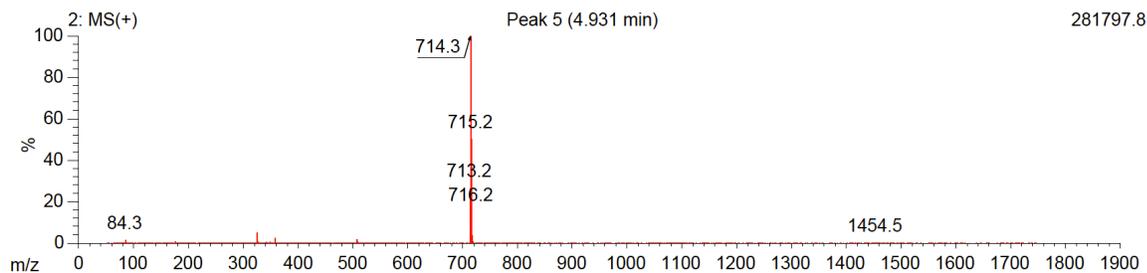
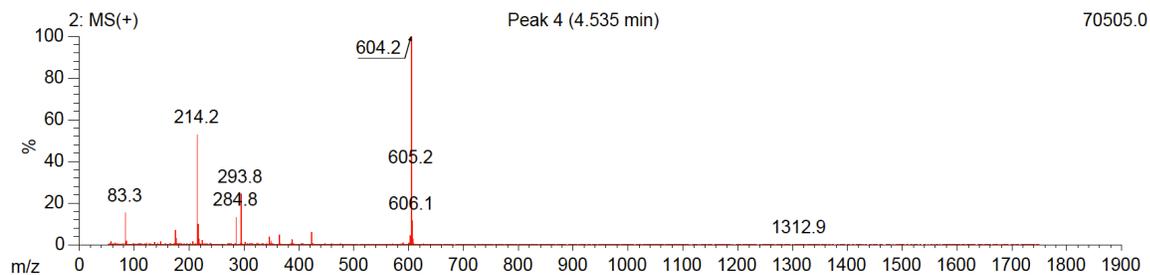
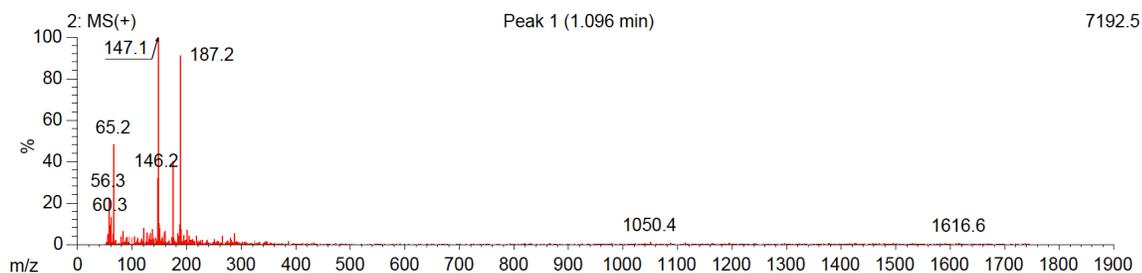
C) 3 h



D) 4 h

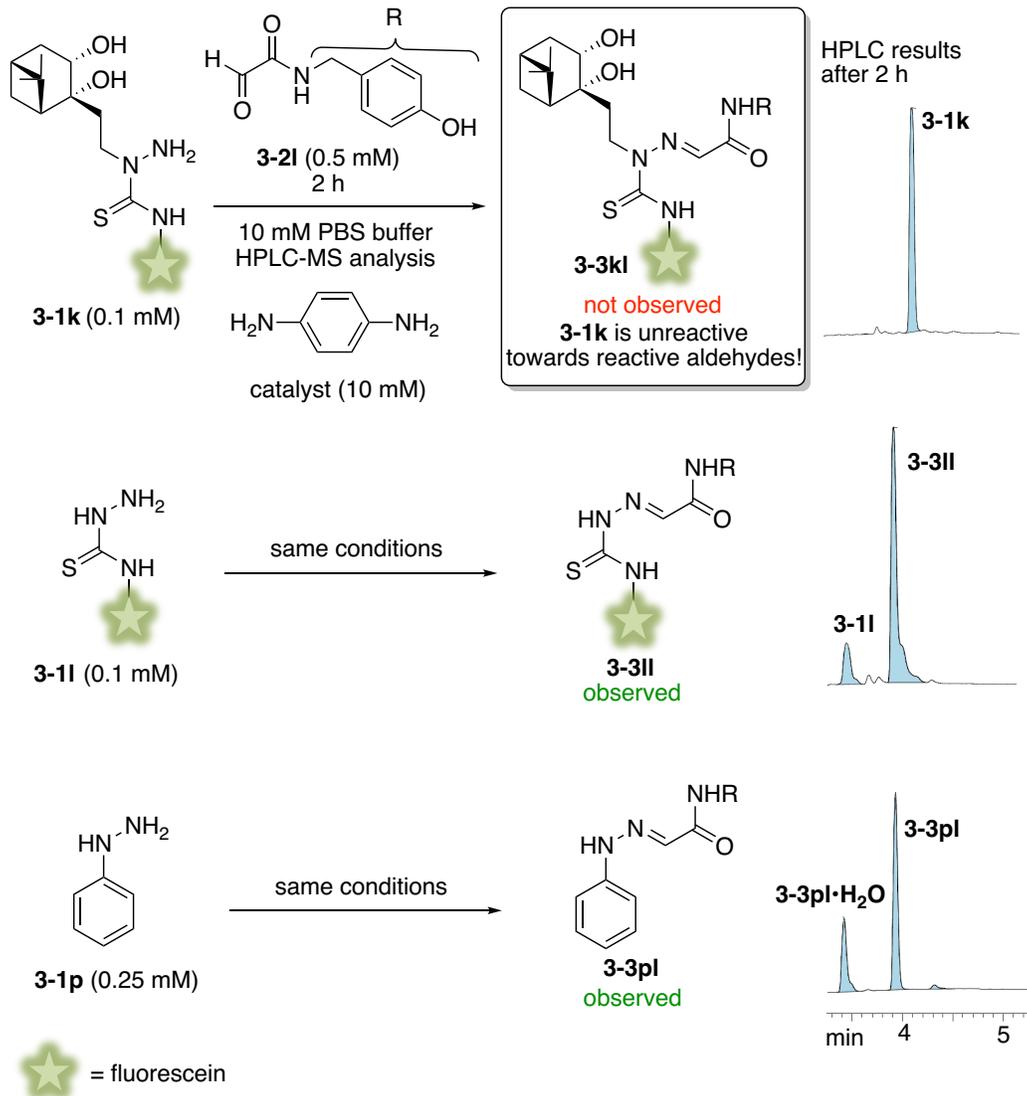


E) (+)-MS traces



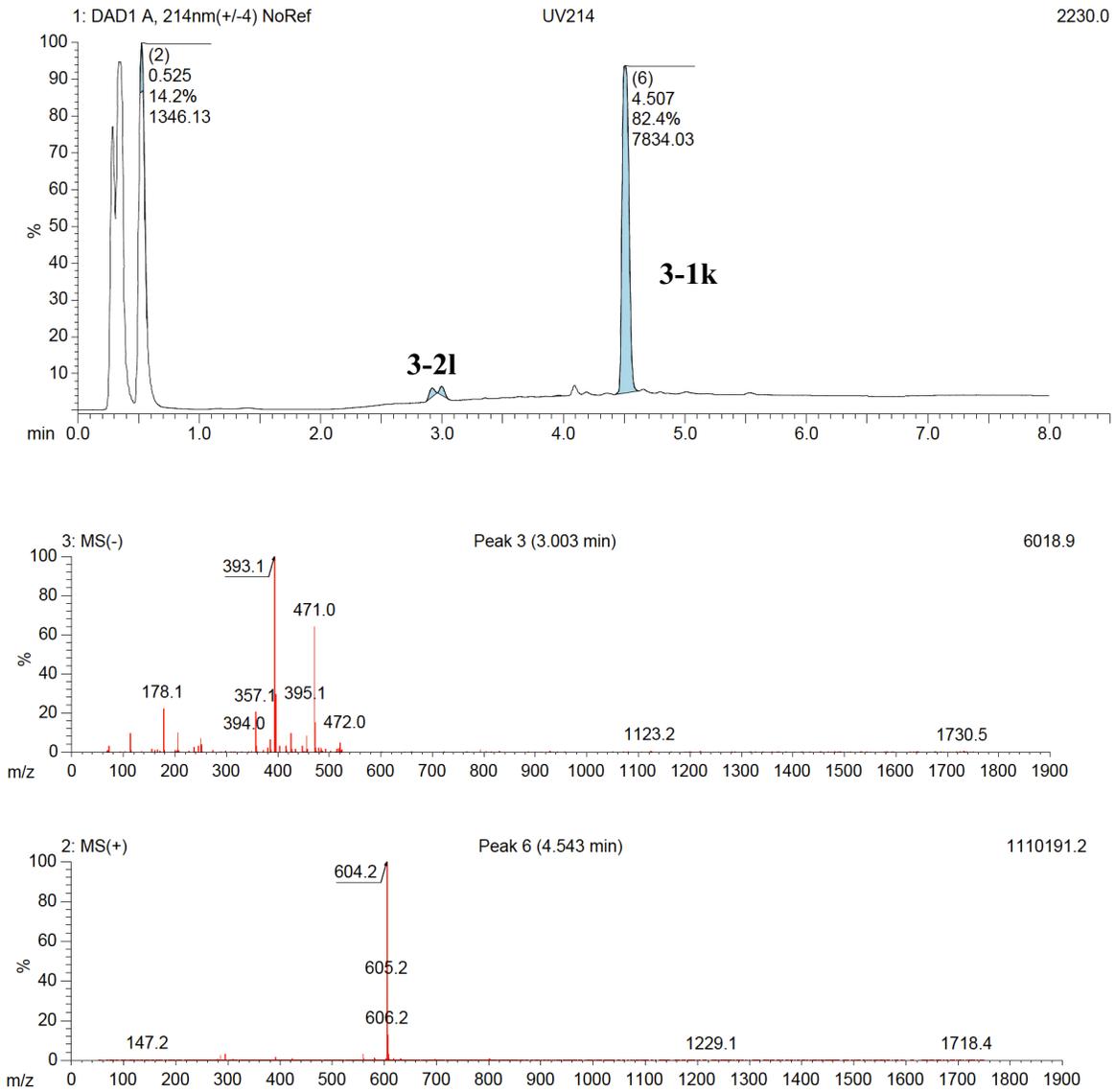
**Figure 3-26:** Representative HPLC-MS chromatogram of reaction of diol **3-1k** with **3-2c** (10  $\mu$ M) in the presence of 1,2-*cis* diols.

Apart from biological polyols, a recent report<sup>47,48</sup> highlighted the presence of protein-bound electrophiles caused by pyruvoyl (ketone) and glyoxylyl (aldehyde) electrophilic post-translational modifications in human (HEK293T cells and the human breast cancer cell line MDA-MB-231). Thus, in order to confirm the inertness and bioorthogonality of reagent **3-1k** and its thiosemicarbazide unit towards these potential electrophiles, especially the reactive glyoxylyl aldehydes, a readily available model aldehyde **3-2l**<sup>49</sup> (0.50 mM) was allowed to react with **3-1k/3-1l/3-1p** (0.10/0.10/0.25 mM) in the presence of a catalyst, *p*-phenylenediamine<sup>50</sup> (10 mM) (Figure 3-27). As the reactions were performed at neutral pH, the catalyst *p*-phenylenediamine<sup>50</sup> was required. These reactions were monitored by HPLC-MS and the analysis showed that **3-1k** is unreactive towards **3-2l** even in the presence of the catalyst while both **3-1l** and **3-1p** afforded the expected thiosemicarbazone and hydrazone adducts, **3-3ll** and **3-3pl**, respectively (Figure 3-27 and 3-28). The inertness of **3-1k** towards **3-2l** is likely due to the steric effect resulting from the nopol diol unit.

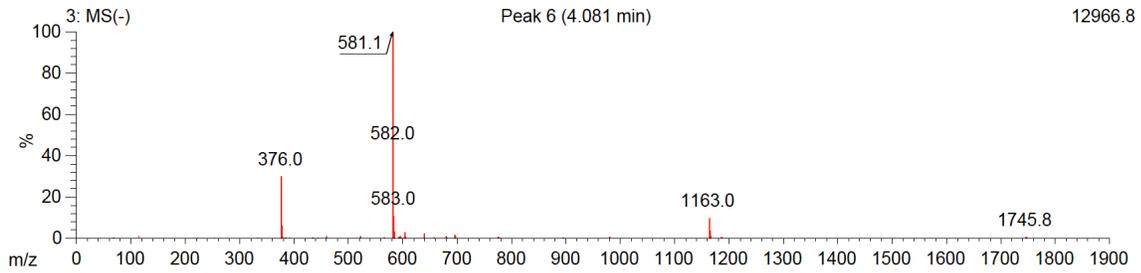
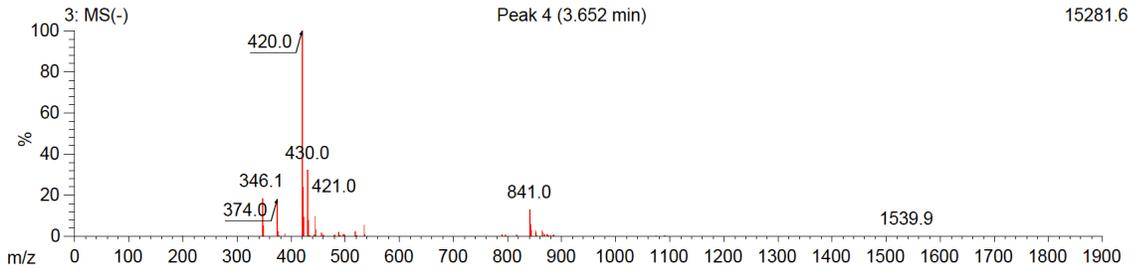
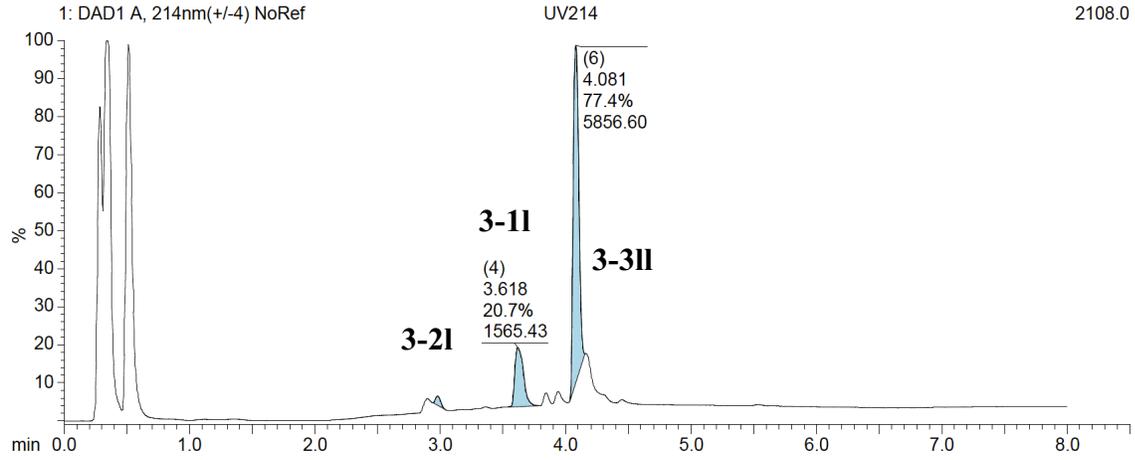


**Figure 3-27:** Study to confirm the inertness and bioorthogonality of diol **3-1k** towards model aldehyde **3-2l**, a mimic of protein-bound glyoxylyl unit.

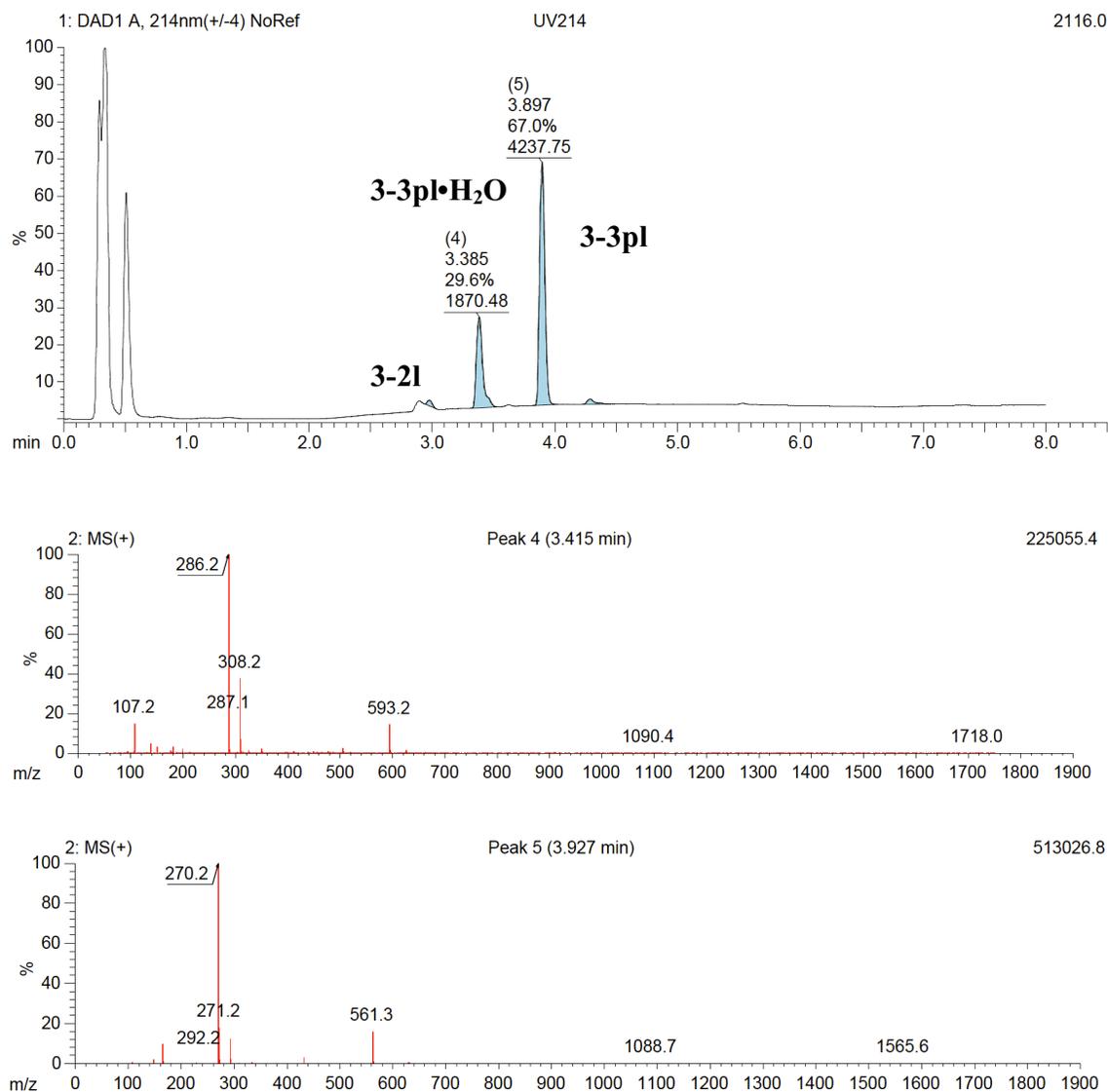
### A) The reaction between 3-1k and 3-2l



B) The reaction between **3-11** and **3-21**



### C) The reaction between **3-1p** and **3-2l**

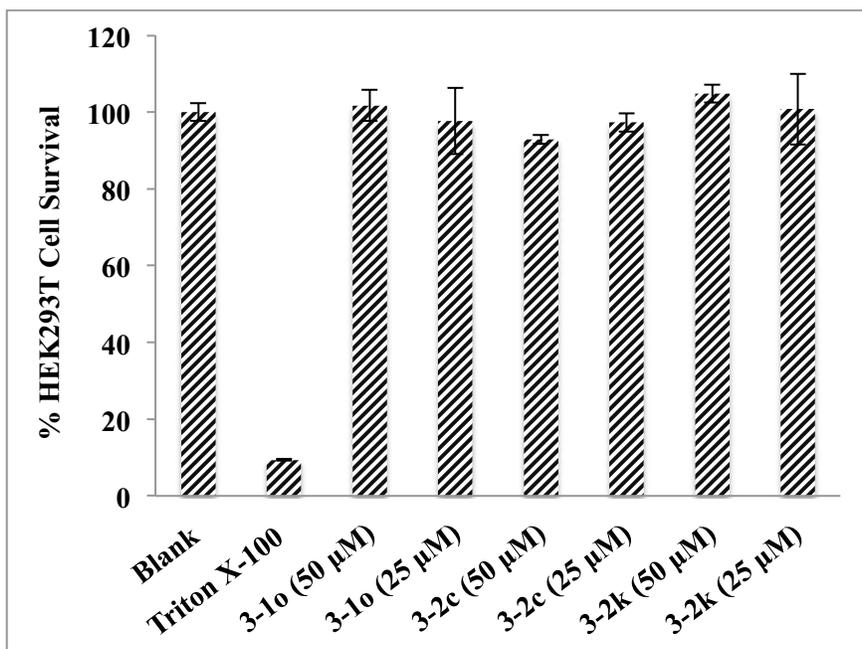


**Figure 3-28:** HPLC-MS chromatograms from a study to confirm the inertness and bioorthogonality of diol **3-1k** towards model aldehyde **3-2l**, a mimic of protein-bound glyoxylal unit.

#### 3.2.4.5 Cytotoxicity and live cell labeling studies

With these promising results in hand, the cytotoxicity of bioorthogonal reagents **3-1o**, **3-2c** and **3-2k** was examined using the WST-1 assay. The easily accessible triton X-100 (1  $\mu$ l) was chosen as a positive control that can cause cell death.<sup>51</sup> As already explained in Chapter 2, cell proliferation reagent WST-1 (water-soluble tetrazolium salt) is susceptible to metabolically

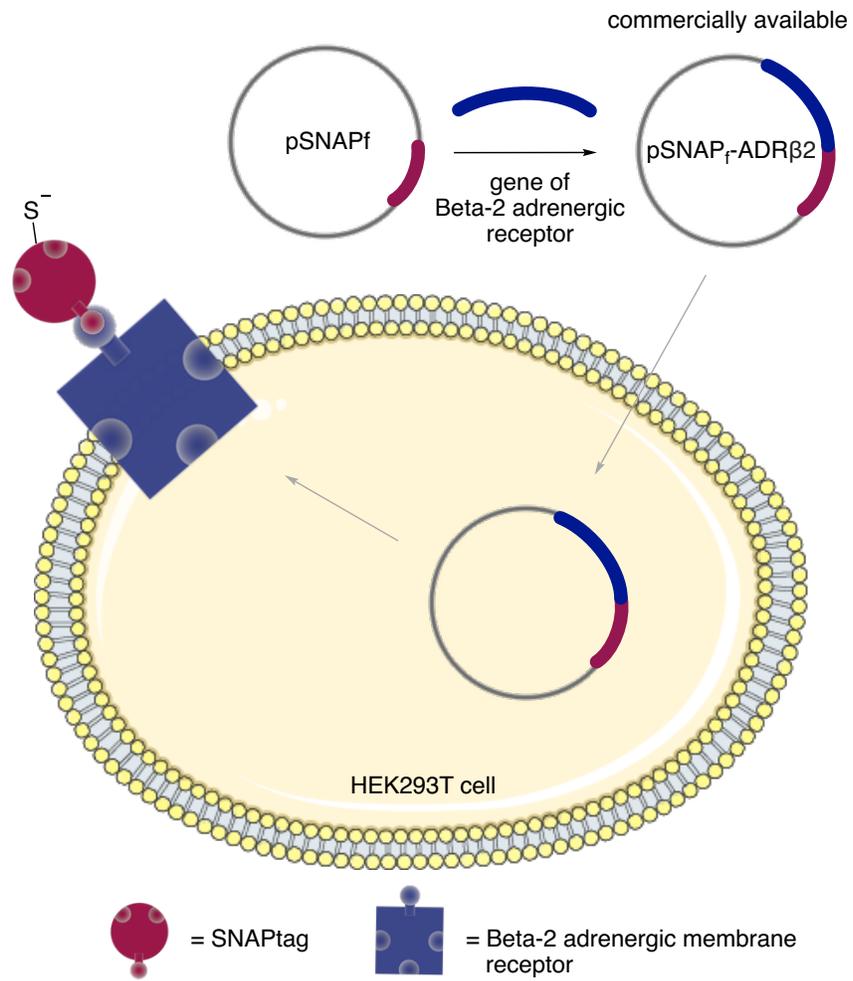
active cells, and the assay operates at cell surface through reduction of WST-1 (pale yellow) and formation of water-soluble formazan (dark yellow) by cellular dehydrogenases (Figure 2-29).<sup>52</sup> These reagents were found to be benign towards HEK293T cells over 18 hours at 50 and 25  $\mu\text{M}$  concentrations (Figure 3-29), an outcome consistent with Gao's recent report.<sup>25</sup>



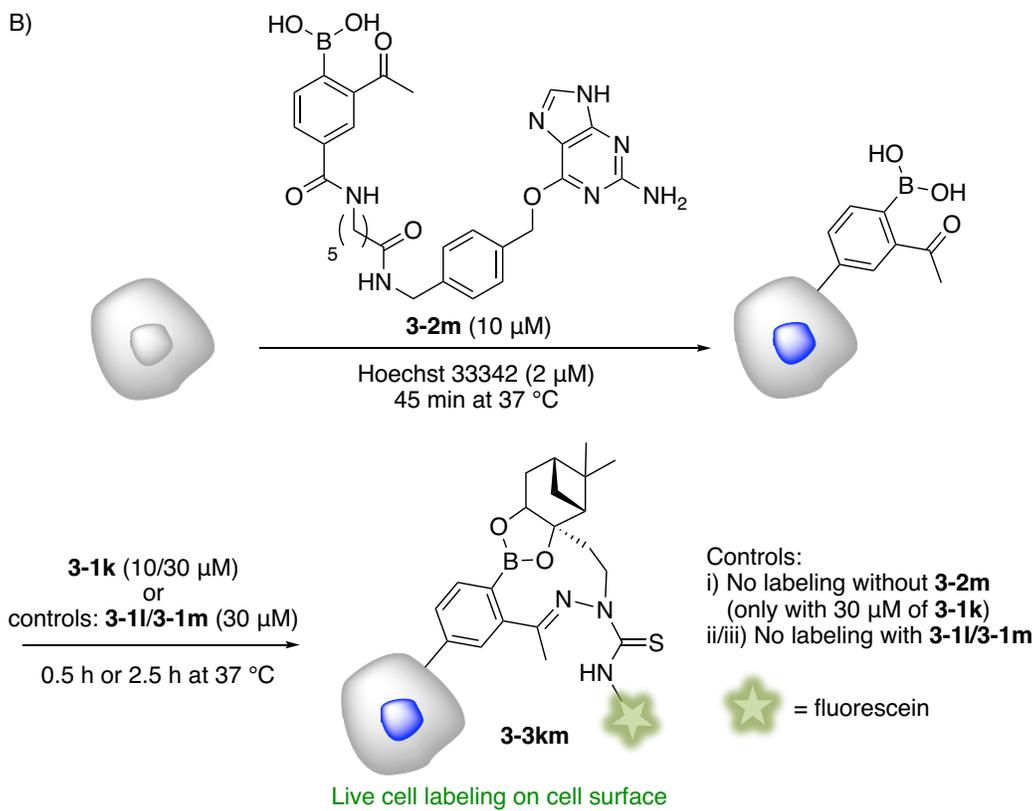
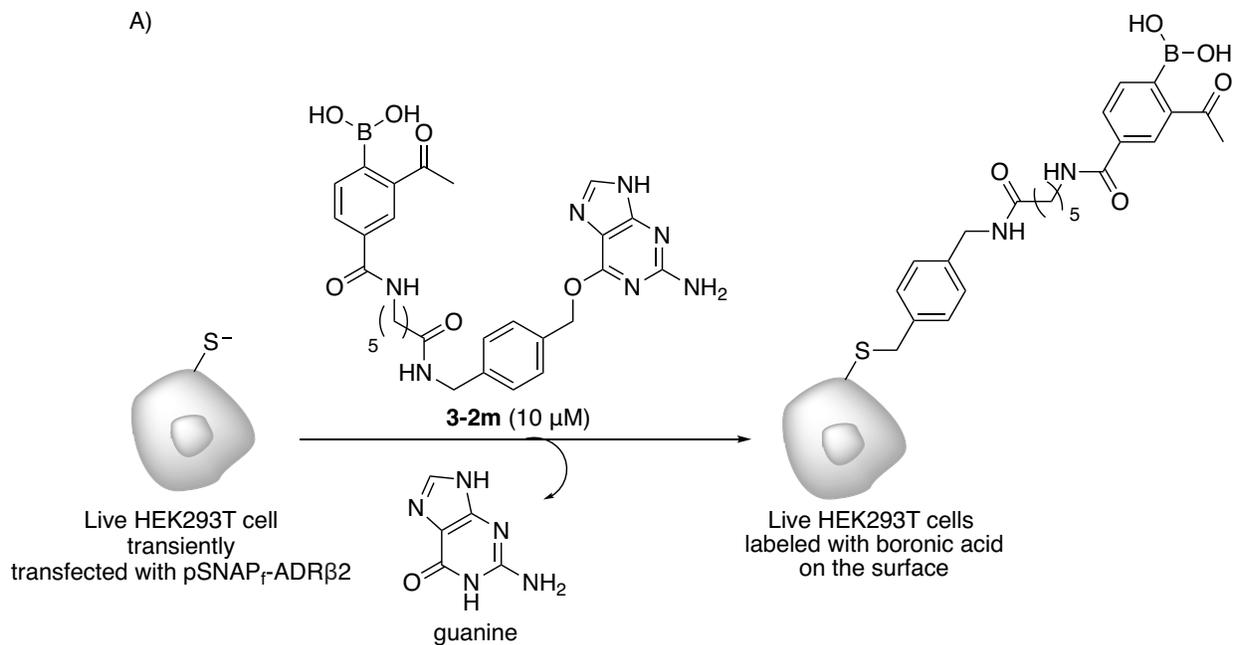
**Figure 3-29:** Cytotoxicity results of bioorthogonal reagents **3-1o**, **3-2c** and **3-2k** on HEK293T cells. Blank is DMSO, and Triton X-100 was used as a positive control. The results are the average of three replicates and error bars are the standard deviations of these replicates.

Finally, in order to demonstrate the viability of this irreversible boronate/thiosemicarbazone system in live-cell imaging, a readily available SNAP-tag approach was employed. The SNAP-tag (20 kDa polypeptide) can easily be introduced into any protein of interest, followed by a specific and irreversible labeling with a proper substrate of SNAP-tag, such as a fluorescently labeled ligand.<sup>53,54</sup> Therefore, this strategy along with fluorescence microscopy can enable the imaging and tracking of the protein of interest in live cells. The SNAP-tag strategy was exploited to install arylboronic acid **3-2m** on the cell membrane of transiently transfected HEK293T cells with the pSNAP<sub>T</sub>-ADRB<sub>2</sub> plasmid (Figure 3-30 and 3-31A). This plasmid carries a gene, which encodes the Beta-2 adrenergic membrane receptor fused to the C-terminus of a SNAP-tag, a mutant of the DNA repair protein O<sup>6</sup>-alkylguanine-DNA alkyltransferase (Figure 3-30). The SNAP<sub>T</sub>-Beta-2 adrenergic membrane receptor was

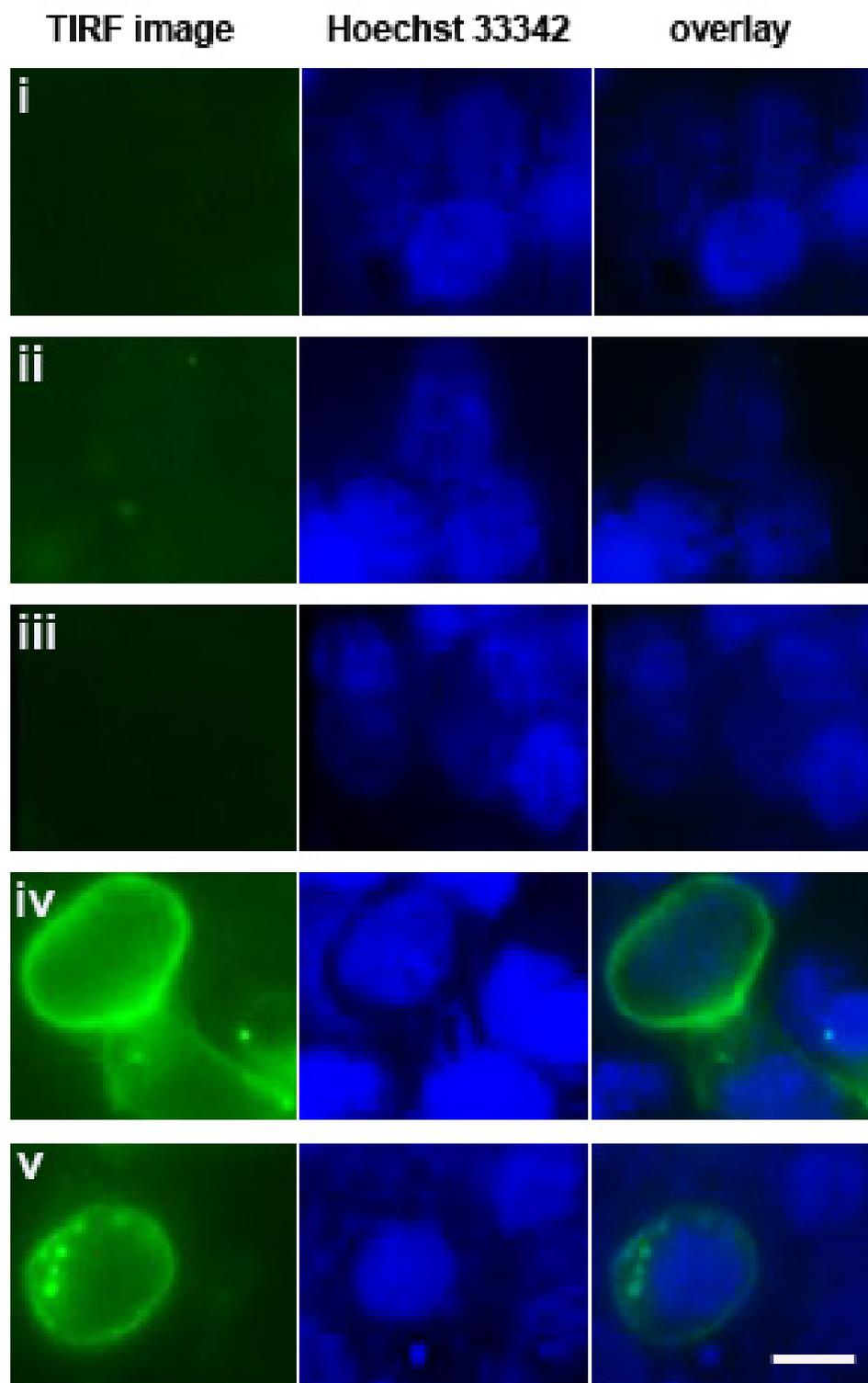
expressed in the cell membrane of HEK293T cells. This receptor fused with SNAP-tag contains a thiolate anion, which can specifically and covalently react with the benzylguanine-derivatized arylboronic acid **3-2m** (10  $\mu\text{M}$ ), thus securing the arylboronic acid on the cell membrane. In other words, the fused SNAP-tag provides an irreversible installation of the arylboronic acid onto the SNAP $\beta$ -Beta-2 adrenergic membrane receptor through its specific reactivity towards **3-2m** releasing a guanine by-product (Figure 3-31A). Diol **3-1k** (10 and 30  $\mu\text{M}$ ) or controls **3-1l/3-1m** (30  $\mu\text{M}$ ) were then introduced into the cell media and incubated for 0.5 h or 2.5 h at 37  $^{\circ}\text{C}$  (Figure 3-31B). In parallel, cells without **3-2m** were incubated with diol **3-1k** (30  $\mu\text{M}$ ) as a negative control. The cells were washed and imaged on a total internal reflection fluorescence (TIRF) microscope. TIRF is a unique technique, which exploits an evanescent wave (near-field wave) in fluorescence microscopy.<sup>55</sup> In short, incident light is reflected at the interface of a coverslip and a film of aqueous solution between the coverslip and adherent HEK293T cells; thus, the intensity of the evanescent field becomes strong at the interface and gets weaker with distance from the interface. This reduces the background fluorescence because non-bound fluorophores to the coverslip are not excited. Therefore, TIRF microscopy would allow for detecting fluorescence from the fluorescein labelled HEK293T cells that are bound to the coverslip. Whilst controls showed no labeling after 2.5 h (Figure 3-10C i, ii, iii), a clear cell surface imaging was observed for the HEK293T cells incubated with both diol **3-1k** and **3-2m** even after 0.5 h (Figure 3-31C iv). Moreover, a vivid image was possible even when using a low 10  $\mu\text{M}$  concentration of **3-1k** (Figure 3-31C v).



**Figure 3-30:** The SNAP-tag approach exhibiting the expression of the  $SNAP_f$ -Beta-2 adrenergic membrane receptor on the cell membrane.



c)



**Figure 3-31:** Imaging of boronic acid **3-2m** surface labeled HEK293T live cells with diol **3-1k**. A) Scheme of biochemical installation of boronic acid. B) Scheme of labeling. C) Imaging

results; i) HEK293T live cells not incubated with **3-2m**, but only incubated with **3-1k** (30  $\mu\text{M}$ ) in 2.5 h. ii) HEK293T live cells incubated with both **3-2m** (10  $\mu\text{M}$ ) and **3-1l** (30  $\mu\text{M}$ ) in 2.5 h. iii) HEK293T live cells incubated with both **3-2m** (10  $\mu\text{M}$ ) and **3-1m** (30  $\mu\text{M}$ ) in 2.5 h. iv) HEK293T live cells incubated with both **3-2m** (10  $\mu\text{M}$ ) and **3-1k** (30  $\mu\text{M}$ ) in 0.5 h. v) HEK293T live cells incubated with both **3-2m** (10  $\mu\text{M}$ ) and **3-1k** (10  $\mu\text{M}$ ) in 2.5 h. Scale bar corresponds to 10  $\mu\text{m}$ . Hoechst 33342 was used for live cell nuclear staining.

### 3.3 Conclusions

In conclusion, I have successfully developed an irreversible nopoldiol boronate bioconjugation system *via* a synergic thiosemicarbazone interaction that provides a stable medium-sized ring product of double condensation. A synergic system based on two bifunctional reagents, a thiosemicarbazide-functionalized nopoldiol and an *ortho*-acetylarylboronic acid, was designed and optimized. Both reagents were shown to be chemically stable and non-toxic to HEK293T cells at concentrations as high as 50  $\mu\text{M}$ . The resulting boronate/thiosemicarbazone adduct is a medium sized ring that forms rapidly and irreversibly without any catalyst at low  $\mu\text{M}$  concentrations, in neutral buffer, with a rate constant of  $9 \text{ M}^{-1}\text{s}^{-1}$  as measured by NMR analysis. Control experiments in the presence of competing boronic acids showed no cross-over side-products and confirmed the stability and lack of reversibility of the boronate/thiosemicarbazone conjugates. Formation of the conjugates was not affected by the presence of biological diols like fructose, glucose and catechol, and the thiosemicarbazide-functionalized nopoldiol was found to be inert to aldehyde electrophiles of the sort found on protein-bound glyoxylyl units. The suitability of this system in the cell-surface labeling of live cells was demonstrated using a SNAP-tag approach to install the boronic acid reagent onto the extracellular domain of Beta-2 adrenergic receptor in HEK293T cells, followed by incubation with the optimal thiosemicarbazide-functionalized nopoldiol reagent labeled with fluorescein dye. Successful visualization by fluorescence microscopy was possible with a reagent concentration as low as 10  $\mu\text{M}$ , thus confirming the potential of this system in biological applications. Therefore, this irreversible boronate/thiosemicarbazone synergic system is a valuable addition to the bioorthogonal chemistry toolbox to enable the study of complex biological processes.

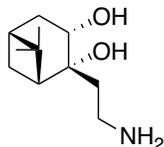
## 3.4 Experimental

### 3.4.1 Experimental details and general information

Unless otherwise stated, all reactions were performed under a nitrogen atmosphere. THF, dichloromethane, DMF and methanol were obtained from a MBraun MB SPS\* solvent system prior to use. Most of the reagents were purchased from Sigma-Aldrich and used as received. Arylboronic acids were purchased either from Sigma-Aldrich or Combi-Blocks. Thin layer chromatography (TLC) was performed on Merck Silica Gel 60 F254 plates and was visualized with UV light, curcumin and  $\text{KMnO}_4$  stain. NMR spectra were recorded on INOVA-400, INOVA-500, INOVA-600 or INOVA-700 MHz instruments. The residual solvent protons ( $^1\text{H}$ ) of  $\text{CDCl}_3$  (7.26 ppm),  $\text{ACN-d}_3$  (1.94 ppm),  $\text{DMSO-d}_6$  (2.50 ppm), acetone- $\text{d}_6$  (2.05 ppm),  $\text{CD}_3\text{OD}$  (3.31 ppm) and the solvent carbons ( $^{13}\text{C}$ ) of  $\text{CDCl}_3$  (77.06 ppm),  $\text{ACN-d}_3$  (1.32 and 118.26 ppm),  $\text{DMSO-d}_6$  (39.52 ppm), acetone- $\text{d}_6$  (29.84 and 206.26 ppm),  $\text{CD}_3\text{OD}$  (49 ppm) were used as internal standards.  $^1\text{H}$  NMR data is presented as follows: chemical shift in ppm ( $\delta$ ) downfield from tetramethylsilane (multiplicity, coupling constant, integration). The following abbreviations are used in reporting NMR data: s, singlet; br s, broad singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; ddd, doublet of doublet of doublets; dddd, doublet of doublet of doublet of doublets; app s, apparent singlet; app ddt, apparent doublet of doublet of triplets; m, multiplet; comp m, complex multiplet. The quaternary carbon bound to the boron atom is often missing due to the quadrupolar relaxation of boron. This effect was observed in each boron-containing compound. High-resolution mass spectra were recorded by the University of Alberta mass spectrometry services laboratory using either electron impact (EI) or electrospray ionization (ESI) techniques. Optical rotations were measured using a 1 mL cell with a 1 dm length on a P.E. 241 polarimeter. Melting points were determined in a capillary tube using a Gallenkamp melting point apparatus and are uncorrected. HPLC, LC-MS methods and biological instruments were described specifically in the corresponding text below.

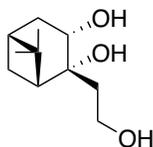
### 3.4.2 Chemical synthesis and analytical data

#### Synthesis of (–)-nopoldiol amine 1 (3-1a)



The synthesis of diol **3-1a** is described in Chapter 2. The diol **3-1a** was labeled as **2-26**.

#### Synthesis of (–)-nopoldiol alcohol (3-4)



**(1*R*,2*R*,3*S*,5*R*)-2-(2-Hydroxyethyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (3-4):** *N*-Methylmorpholine-*N*-oxide (50 wt% in H<sub>2</sub>O) (1.3 equiv, 2.4 ml, 12 mmol) and 2,6-lutidine (1.2 equiv, 1.2 ml, 11 mmol) were added to a solution containing (–)-nopol (1.0 equiv, 1.5 g, 8.9 mmol) in isopropanol (20 ml) at room temperature. Lastly, K<sub>2</sub>OsO<sub>4</sub>•H<sub>2</sub>O (1.7 mol%, 55 mg, 0.15 mmol) was added to the reaction mixture, which was then stirred and refluxed at 95 °C for 24 h under ambient atmosphere. Then, the reaction mixture was concentrated in vacuo and mixed with EtOAc (100 ml). The organic phase was washed with distilled water (1 × 10 ml), HCl (1 × 10 ml, 1 N), and brine (1 × 10 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:3, acetone/hexanes) to provide diol **3-4** as a white-transparent solid (1.2 g, 67% yield).

White-transparent solid:  $R_f = 0.31$  (35:65, acetone/hexanes).

Melting point: 30.2 – 31.6 °C.

$[\alpha]_{\text{D}}^{20} : -12.4$  ( $c$  0.59, CHCl<sub>3</sub>).

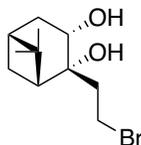
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 4.12 (dd,  $J = 9.5, 5.6$  Hz, 1H), 4.01 – 3.9.4 (m, A part of ABMM', 1H), 3.93 – 3.85 (m, B part of ABMM', 1H), 3.25 (bs, 3H), 2.50 – 2.41 (m, 1H), 2.24 – 2.16 (m, 1H), 2.11 (app t,  $J = 5.8$  Hz, 1H), 1.97 – 1.88 (comp m, 2H), 1.71 – 1.62 (comp m, 2H), 1.43 (d,  $J = 10.3$  Hz, 1H), 1.27 (s, 3H), 0.93 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 75.0, 67.1, 58.6, 52.1, 41.9, 39.6, 37.8, 36.8, 27.0, 26.8, 23.2.

IR (cast film,  $\text{cm}^{-1}$ ): 3350, 2922, 2870, 1058, 1019.

HRMS (ESI-TOF): for  $\text{C}_{11}\text{H}_{20}\text{NaO}_3$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 223.1305; *found*: 223.1306.

### Synthesis of (–)-nopoldiol bromide (3-5)



**(1R,2R,3S,5R)-2-(2-Bromoethyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (3-5):** To a mixture of compound **3-4** (1.0 equiv, 0.96 g, 4.8 mmol) and triphenylphosphine (1.1 equiv, 1.4 g, 5.3 mmol) in  $\text{CH}_2\text{Cl}_2$  (16 ml) was added NBS (1.1 equiv, 0.94 g, 5.3 mmol) partially under ice bath and inert atmosphere. The reaction mixture was then brought to rt and stirred for 2 h after which the reaction mixture was concentrated in vacuo. The crude residue was purified by flash chromatography (15:85, acetone/hexanes) to provide compound **3-5** as a white solid (0.70 g, 55% yield).

White solid:  $R_f = 0.46$  (25:75, acetone/hexanes).

Melting point: 59.5 – 62.3  $^\circ\text{C}$ .

$[\alpha]_{\text{D}}^{20}$  : –10.6 (*c* 0.46,  $\text{CHCl}_3$ ).

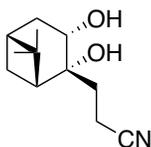
$^1\text{H}$  NMR  $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 4.10 (ddd,  $J = 9.4, 6.0, 5.1$  Hz, 1H), 3.63 – 3.57 (m, A part of ABMM', 1H), 3.57 – 3.51 (m, B part of ABMM', 1H), 2.97 (s, 1H), 2.54 (d,  $J = 6.0$  Hz, 1H), 2.53 – 2.47 (m, 1H), 2.28 – 2.20 (comp m, 2H), 2.13 – 2.06 (comp m, 2H), 1.97 – 1.91 (m, 1H), 1.65 (ddd,  $J = 14.0, 5.1, 2.5$  Hz, 1H), 1.36 (d,  $J = 10.5$  Hz, 1H), 1.28 (s, 3H), 0.96 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ )  $\delta$  74.6, 66.9, 51.0, 44.8, 39.6, 37.9, 37.3, 27.4, 26.8, 26.8, 23.2.

IR (cast film,  $\text{cm}^{-1}$ ): 3362, 2924, 1386, 1368, 1051, 1027.

HRMS (ESI-TOF): for  $\text{C}_{11}\text{H}_{16}\text{Br}$  ( $\text{M} - 2\text{H}_2\text{O} + \text{H}$ ) $^+$ : *calcd.*: 227.043; *found*: 227.0432; for  $\text{C}_{11}\text{H}_{20}\text{BrO}_2$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 263.0641; *found*: 263.0662.

### Synthesis of (–)-nopoldiol cyano (3-6)



#### **3-((1R,2R,3S,5R)-2,3-Dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)propanenitrile (3-6):**

Sodium cyanide (1.4 equiv, 62 mg, 1.3 mmol) was added to a solution containing **3-5** (1.0 equiv, 0.24 g, 0.91 mmol) in DMSO (1.0 ml). The mixture was stirred and heated at 70 °C for 5 h under inert atmosphere. Then, the reaction mixture was cooled down at room temperature and diluted with distilled water (10 ml). The aqueous layer was extracted with diethyl ether (4 × 50 ml) and organic fractions were combined, dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (4:6, EtOAc/hexanes) to provide compound **3-6** as a yellowish oil (0.11 g, 59% yield).

Yellowish oil:  $R_f = 0.14$  (1:4, acetone/hexanes).

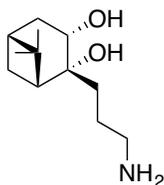
<sup>1</sup>H NMR  $\delta$ /ppm: (500 MHz, CDCl<sub>3</sub>) 4.05 (dd,  $J = 9.4, 5.0$  Hz, 1H), 3.13 (bs, 1H), 2.67 – 2.43 (comp m, 4H), 2.25 (app dtd,  $J = 10.5, 6.1, 2.5$  Hz, 1H), 2.08 (app t,  $J = 5.8$  Hz, 1H), 2.04 – 1.93 (comp m, 2H), 1.83 (ddd,  $J = 14.0, 8.9, 6.4$  Hz, 1H), 1.66 (ddd,  $J = 14.0, 5.0, 2.5$  Hz, 1H), 1.36 (d,  $J = 10.6$  Hz, 1H), 1.29 (s, 3H), 0.95 (s, 3H).

<sup>13</sup>C NMR  $\delta$ /ppm: (126 MHz, CDCl<sub>3</sub>) 119.9, 73.3, 66.9, 50.4, 39.6, 37.9, 37.5, 36.7, 26.8, 26.6, 23.2, 10.6.

IR (cast film, cm<sup>-1</sup>): 3444, 2926, 2247, 1114, 1066.

HRMS (ESI-TOF): for C<sub>12</sub>H<sub>19</sub>NNaO<sub>2</sub> (M + Na)<sup>+</sup>: *calcd.*: 232.1308; *found*: 232.1309.

### Synthesis of (–)-nopoldiol amine 2 (3-1b)



**(1R,2R,3S,5R)-2-(3-Aminopropyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (3-1b):** To a suspension of LiAlH<sub>4</sub> (2.0 equiv, 54 mg, 1.4 mmol) in freshly distilled diethyl ether (1.5 ml) under inert atmosphere was added compound **3-6** (1.0 equiv, 0.15 g, 0.72 mmol) as a solution in

freshly distilled diethyl ether (4.0 ml). The reaction mixture was stirred for 1 d at room temperature. The reaction was quenched by slow addition of 20% aqueous KOH solution (4.0 ml) at 0 °C. The solid was filtered and washed with diethyl ether. The combined filtrates were washed with distilled water, dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (95:5, EtOAc/Et<sub>3</sub>N) to provide diol **3-1b** as a yellow viscous oil (60 mg, 39% yield).

Yellowish oil: R<sub>f</sub> = smear product close to the baseline (95:5, EtOAc/Et<sub>3</sub>N).

[α]<sub>D</sub><sup>20</sup> : -11.1 (c 0.48, CHCl<sub>3</sub>).

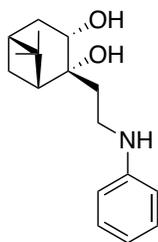
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 3.99 (dd, *J* = 9.3, 4.9 Hz, 1H), 3.75 – 3.60 (m, 2H), 2.49 (dddd, *J* = 14.0, 9.3, 3.5, 2.4 Hz, 1H), 2.31 (bs, 4H), 2.25 – 2.17 (m, 1H), 2.11 (app t, *J* = 5.8 Hz, 1H), 1.92 (m, 1H), 1.86 – 1.70 (comp m, 3H), 1.66 (ddd, *J* = 14.0, 4.9, 2.6 Hz, 1H), 1.63 – 1.56 (m, 1H), 1.38 (d, *J* = 10.4 Hz, 1H), 1.27 (s, 3H), 0.93 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 75.2, 68.4, 63.5, 52.0, 40.6, 39.1, 39.0, 38.6, 27.8, 27.7, 26.2, 24.3.

IR (cast film, cm<sup>-1</sup>): 3346, 2924, 2870, 1050, 1005.

HRMS (ESI-TOF): for C<sub>12</sub>H<sub>24</sub>NO<sub>2</sub> (M + H)<sup>+</sup>: *calcd.*: 214.1802; *found*: 214.1804.

### Synthesis of (-)-nopoldiol phenylamine (**3-1c**)



#### **(1*R*,2*R*,3*S*,5*R*)-6,6-Dimethyl-2-(2-(phenylamino)ethyl)bicyclo[3.1.1]heptane-2,3-diol (3-1c):**

The procedure was adapted from the reported method by Buchwald and coworkers.<sup>27</sup> CuI (0.05 equiv, 3.3 mg, 0.017 mmol), K<sub>3</sub>PO<sub>4</sub> (2.0 equiv, 0.14 g, 0.66 mmol) and compound **3-1a** (1.2 equiv, 80 mg, 0.40 mmol) were added to a microwave vial. The tube was evacuated and back-filled with nitrogen three times. 2-propanol (0.33 ml), ethylene glycol (2.0 equiv, 37 μl, 0.66 mmol) and iodobenzene (1.0 equiv, 37 μl, 0.33 mmol) were added by syringe at room temperature under nitrogen. The reaction mixture was heated to 80 °C to furnish a pink-brown suspension and stirred for 16 h. Then, the reaction mixture was cooled to rt. Distilled water (2.0

ml) was added, and the mixture was extracted with diethyl ether (4 x 10 ml). The combined organic phases were washed with brine (1 x 5.0 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:3, EtOAc/hexanes) to provide diol **3-1c** as a pink powder (63 mg, 69% yield).

Melting point: 93.0 – 95.8 °C.

Pink powder: R<sub>f</sub> = 0.62 (1:1, EtOAc/hexanes)

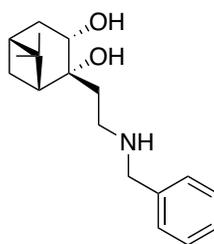
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.20 (dd, *J* = 8.6, 7.4 Hz, 2H), 6.79 (tt, *J* = 7.3, 1.1 Hz, 1H), 6.73 (dd, *J* = 8.6, 1.1 Hz, 2H), 4.09 (dd, *J* = 9.4, 5.3 Hz, 1H), 3.45 – 3.32 (comp m, 2H), 2.49 (dddd, *J* = 14.0, 9.4, 3.7, 2.4 Hz, 1H), 2.25 – 2.16 (m, 1H), 2.13 (app t, *J* = 5.8 Hz, 1H), 2.01 – 1.90 (comp m, 2H), 1.81 (ddd, *J* = 14.5, 7.5, 5.1 Hz, 1H), 1.67 (ddd, *J* = 13.9, 5.3, 2.4 Hz, 1H), 1.41 (d, *J* = 10.3 Hz, 1H), 1.28 (s, 3H), 0.95 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) δ 147.0, 128.3, 118.0, 113.5, 76.3, 76.1, 75.8, 75.0, 67.2, 51.9, 39.9, 39.6, 39.6, 37.9, 37.2, 26.9, 26.7, 23.3.

IR (cast film, cm<sup>-1</sup>): 3482, 3364, 3256, 3023, 3057, 2993, 2941, 2912, 2868, 2840, 1606, 1509, 1470, 1324, 1004.

HRMS (ESI-TOF): for C<sub>17</sub>H<sub>26</sub>NO<sub>2</sub> (M + H)<sup>+</sup>: *calcd.*: 276.1958; *found*: 276.1954.

### Synthesis of (–)-nopoldiol benzylamine (**3-1d**)



#### (1*R*,2*R*,3*S*,5*R*)-2-(2-(Benzylamino)ethyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (**3-1d**):

Sodium triacetoxyborohydride (1.4 equiv, 0.20 g, 0.92 mmol) was added to the solution of compound **3-1a** (1.0 equiv, 0.13 g, 0.66 mmol) and benzaldehyde (1.2 equiv, 0.084 g, 0.79 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (1.4 ml) under inert atmosphere. The reaction mixture was then stirred overnight at room temperature. Then, methanol (2.0 ml) was added and the mixture was concentrated in vacuo. The crude residue was purified by flash chromatography (10:90:0.5, methanol/CH<sub>2</sub>Cl<sub>2</sub>/ammonium hydroxide) to provide diol **3-1d** as a yellow oil (59 mg, 31% yield).

Yellow oil:  $R_f = 0.42$  (10:90:0.5, methanol/ $\text{CH}_2\text{Cl}_2$ /ammonium hydroxide).

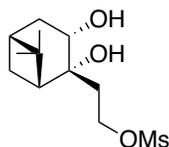
$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.36 – 7.27 (comp m, 5H), 3.97 (dd,  $J = 9.3, 5.5$  Hz, 1H), 3.79 (s, 2H), 3.04 – 2.93 (comp m, 2H), 2.43 (dddd,  $J = 13.3, 9.3, 3.8, 2.4$  Hz, 1H), 2.21 – 2.13 (m, 1H), 2.05 (app t,  $J = 5.8$  Hz, 1H), 1.92 – 1.86 (m, 1H), 1.82 (ddd,  $J = 14.8, 7.3, 3.5$  Hz, 1H), 1.72 – 1.60 (comp m, 2H), 1.46 (d,  $J = 10.2$  Hz, 1H), 1.25 (s, 3H), 0.91 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 137.2, 127.7, 127.4, 126.6, 75.3, 66.9, 52.6, 52.6, 44.0, 39.6, 38.4, 37.8, 37.1, 27.0, 26.8, 23.3.

**IR** (cast film,  $\text{cm}^{-1}$ ): 3439, 3289, 3064, 3029, 2986, 2919, 2867, 1496, 1473, 1099, 1056, 1026.

**HRMS (ESI-TOF)**: for  $\text{C}_{18}\text{H}_{28}\text{NO}_2$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 290.2115; *found*: 290.2111.

### Synthesis of (–)-nopoldiol mesylate (3-7)



#### 2-((1R,2R,3S,5R)-2,3-Dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl

**methanesulfonate (3-7)**: Methanesulfonyl chloride (1.1 equiv, 0.16 ml, 2.0 mmol) was added dropwise to a stirred solution of diol **3-4** (1.0 equiv, 0.36 g, 1.8 mmol) in  $\text{CH}_2\text{Cl}_2$  (3.0 ml) and  $\text{Et}_3\text{N}$  (2.3 equiv, 0.58 ml, 4.1 mmol) under ice bath and inert atmosphere. The reaction mixture was stirred for 1 h at 0 °C and 1 h at room temperature. Saturated ammonium chloride (2.0 ml) was then added; the mixture was extracted with ethyl acetate (3 x 10 ml). The organic phases were combined and washed with brine (1 x 5.0 ml), dried ( $\text{MgSO}_4$ ), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (3:7, acetone/hexanes) to provide diol **3-7** as a yellow viscous oil (0.29 g, 57% yield).

Yellowish viscous oil:  $R_f = 0.27$  (3:7, acetone/hexanes).

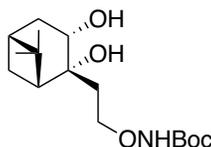
$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 4.57 – 4.49 (m, A part of ABMM', 1H), 4.45 – 4.37 (m, B part of ABMM', 1H), 4.12 (dd,  $J = 9.4, 5.0$  Hz, 1H), 3.02 (s and bs overlap, 5H), 2.49 (dddd,  $J = 14.0, 9.4, 3.6, 2.5$  Hz, 1H), 2.25 – 2.17 (m, 1H), 2.15 – 2.04 (comp m, 2H), 1.96 – 1.86 (comp m, 2H), 1.65 (ddd,  $J = 14.0, 5.0, 2.5$  Hz, 1H), 1.36 (d,  $J = 10.5$  Hz, 1H), 1.27 (s, 3H), 0.95 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 73.2, 66.7, 66.3, 51.7, 40.3, 39.5, 37.8, 37.1, 36.3, 26.8, 26.7, 23.2.

**IR** (cast film,  $\text{cm}^{-1}$ ): 3490, 2923, 2871, 1474, 1351, 1173, 973, 949.

**HRMS (ESI-TOF)**: for  $\text{C}_{12}\text{H}_{22}\text{NaO}_5\text{S}$  ( $\text{M} + \text{Na}$ )<sup>+</sup>: *calcd.*: 301.108; *found*: 301.1083; for  $\text{C}_{12}\text{H}_{26}\text{NO}_5\text{S}$  ( $\text{M} + \text{NH}_4$ )<sup>+</sup>: *calcd.*: 296.1526; *found*: 296.1528.

### Synthesis of (–)-nopoldiol *N*-Boc-hydroxylamine (3-8)



**Tert-butyl (2-((1*R*,2*R*,3*S*,5*R*)-2,3-dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethoxy)carbamate (3-8)**: To a solution of *N*-boc-hydroxylamine (1.4 equiv, 62 mg, 0.46 mmol) in anhydrous diethyl ether (1.0 ml) and DBU (1.5 equiv, 74  $\mu\text{l}$ , 0.50 mmol) was added compound **3-7** (1.0 equiv, 92 mg, 0.33 mmol) as a solution in anhydrous diethyl ether (2.0 ml) under inert atmosphere. The reaction mixture was stirred for 2 days at room temperature. Then, diethyl ether (20 ml) was added to the reaction mixture and it was washed with saturated ammonium chloride (2 x 5.0 ml), brine (1 x 5.0 ml), dried ( $\text{MgSO}_4$ ), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:4, acetone/hexanes) to provide compound **3-8** as a white solid (68 mg, 65% yield).

White solid:  $R_f = 0.51$  (1:3, acetone/hexanes).

Melting point: 111.0 – 113.2  $^{\circ}\text{C}$ .

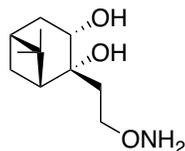
**$^1\text{H}$  NMR**  $\delta$ /ppm: (498 MHz,  $\text{CDCl}_3$ ) 7.41 (s, 1H), 4.43 – 4.36 (m, 1H), 4.32 (ddd,  $J = 10.5, 8.7, 1.5$  Hz, 1H), 3.86 (ddd,  $J = 8.7, 5.3, 2.5$  Hz, 1H), 2.42 (dddd,  $J = 13.7, 9.8, 3.8, 2.4$  Hz, 1H), 2.19 – 2.11 (m, 1H), 2.00 – 1.91 (comp m, 2H), 1.88 (m, 1H), 1.65 (ddd,  $J = 13.9, 5.8, 2.3$  Hz, 1H), 1.53 – 1.43 (m, 10H), 1.25 (s, 3H), 0.94 (s, 3H).

**$^{13}\text{C}$  NMR**  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 156.7, 82.0, 73.2, 72.5, 65.5, 53.5, 40.2, 39.6, 37.6, 36.0, 27.4, 27.2, 27.1, 23.4.

**IR** (cast film,  $\text{cm}^{-1}$ ): 3497, 3318, 2998, 2974, 2954, 2931, 2905, 1742, 1703, 1494.

**HRMS (ESI-TOF)**: for  $\text{C}_{11}\text{H}_{21}\text{NNaO}_3$  ( $\text{M} + \text{Na}$ )<sup>+</sup>: *calcd.*: 238.1414; *found*: 238.1416; for  $\text{C}_{11}\text{H}_{22}\text{NO}_3$  ( $\text{M} + \text{H}$ )<sup>+</sup>: *calcd.*: 216.1594; *found*: 216.1591.

### Synthesis of (–)-nopoldiol hydroxylamine (3-1e)



**(1R,2R,3S,5R)-2-(2-(Aminoxy)ethyl)-6,6-dimethylbicyclo[3.1.1]heptane-2,3-diol (3-1e):** To a solution of compound **3-8** (1.0 equiv, 90 mg, 0.28 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (0.70 ml) was added TFA (15 equiv, 0.32 ml, 4.2 mmol) under ice bath and inert atmosphere. The reaction mixture was then stirred at room temperature for 1 h. It was then concentrated in vacuo. Distilled water (1.0 ml) was added to the reaction mixture, which was then lyophilized to remove excess TFA. The crude residue was purified by flash chromatography (5:95:0.5, methanol/CH<sub>2</sub>Cl<sub>2</sub>/ammonium hydroxide) to provide diol **3-1e** as a white powder (54 mg, 88% yield).

White powder:  $R_f = 0.45$  (5:95:0.5, methanol/CH<sub>2</sub>Cl<sub>2</sub>/ammonium hydroxide).

Melting point: 55.9 – 58.0 °C.

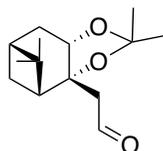
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 4.07 – 3.97 (comp m, 2H), 3.96 – 3.89 (m, 1H), 2.45 (dddd,  $J = 14.0, 9.4, 3.8, 2.5$  Hz, 1H), 2.23 – 2.16 (m, 1H), 2.09 – 1.98 (comp m, 2H), 1.93 – 1.88 (m, 1H), 1.73 (ddd,  $J = 15.3, 6.8, 3.4$  Hz, 1H), 1.66 (ddd,  $J = 13.9, 5.5, 2.4$  Hz, 1H), 1.44 (d,  $J = 10.3$  Hz, 1H), 1.27 (s, 3H), 0.94 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 73.8, 71.5, 67.0, 52.3, 39.9, 39.5, 37.8, 37.0, 27.0, 26.9, 23.2.

IR (cast film, cm<sup>-1</sup>): 3424, 3313, 2923, 2869, 1122, 1057, 1021.

HRMS (ESI-TOF): for C<sub>11</sub>H<sub>22</sub>NO<sub>3</sub> (M + H)<sup>+</sup>: *calcd.*: 216.1594; *found*: 216.1593.

### Synthesis of (–)-nopol-aldehyde-acetal (3-9)



**2-((3aR,4R,6R,7aS)-2,2,5,5-Tetramethyltetrahydro-4,6-methanobenzo[d][1,3]dioxol-3a(4H)-yl)acetaldehyde (3-9):** To a solution of diol **3-4** (1.0 equiv, 88 mg, 0.44 mmol) in acetone (5.0 ml) was added *p*-TsOH monohydrate (0.05 equiv, 4.2 mg, 0.022 mmol) under ambient

atmosphere at room temperature. The reaction mixture was refluxed at 50 °C for 16 h and then brought to rt. Na<sub>2</sub>SO<sub>4</sub> and K<sub>2</sub>CO<sub>3</sub> were added to the reaction mixture, which was then stirred for 15 min, filtered and concentrated in vacuo. The crude mixture was dissolved in hexanes:EtOAc (5.0 ml: 5.0 ml), filtered and concentrated in vacuo in order to remove Na<sub>2</sub>SO<sub>4</sub> that was partially soluble in acetone. Without further purification, the crude residue was directly used in the following oxidation step. To a solution of oxalyl chloride (1.2 equiv, 45 μl, 0.53 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (0.9 ml) at 78 °C under nitrogen was added freshly distilled DMSO (94 μl). The mixture was stirred for 30 min, and then the solution of crude protected nopol alcohol in CH<sub>2</sub>Cl<sub>2</sub> (1.5 ml) was added to the reaction mixture. The mixture was stirred for an additional 30 min, triethylamine (0.38 ml) was then added drop-wise. The reaction mixture was stirred for 15 min at 78 °C and then cooled to rt. Then, it was diluted with CH<sub>2</sub>Cl<sub>2</sub> (25 ml). The organic phase was washed with distilled water (3 × 3.0 ml), dried (Na<sub>2</sub>SO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:4, acetone/hexanes) to provide compound **3-9** as a yellowish oil (82 mg, 78% yield over two steps).

Yellowish oil: R<sub>f</sub> = 0.73 (1:4, acetone/hexanes)

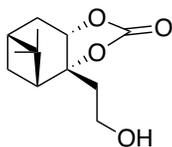
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 9.82 (t, *J* = 3.0 Hz, 1H), 4.17 (d, *J* = 7.2 Hz, 1H), 2.80 (d, *J* = 3.0 Hz, 2H), 2.28 – 2.23 (m, 1H), 2.23 – 2.14 (comp m, 2H), 2.02 – 1.96 (m, 1H), 1.96 – 1.92 (m, 1H), 1.68 (d, *J* = 10.6 Hz, 1H), 1.50 (s, 3H), 1.40 (s, 3H), 1.30 (s, 3H), 0.87 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 200.7, 107.6, 83.4, 76.1, 53.6, 48.5, 38.7, 37.1, 31.3, 26.3, 26.2, 25.4, 24.0, 23.1.

IR (Microscope, cm<sup>-1</sup>): 2986, 2923, 2872, 2738, 1724, 1208.

HRMS (EI) for C<sub>13</sub>H<sub>19</sub>O<sub>3</sub> (M + CH<sub>3</sub>)<sup>+</sup>: *calcd.*: 223.13342; *found*: 223.13305.

### Synthesis of (–)-nopol-alcohol-carbonate (**3-10**)



#### (3aR,4R,6R,7aS)-3a-(2-Hydroxyethyl)-5,5-dimethylhexahydro-4,6-

methanobenzo[d][1,3]dioxol-2-one (**3-10**): To a solution of diol **3-4** (1.0 equiv, 0.19 g, 0.93 mmol) and Et<sub>3</sub>N (1.0 equiv, 0.13 ml, 0.93 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (3.5 ml) was added TMSCl (1.0

equiv, 0.12 ml, 0.93 mmol) under inert atmosphere and ice bath. The reaction mixture was stirred for 30 min at room temperature. Then, 1,1'-carbonyldiimidazole (1.5 equiv, 0.23 g, 1.40 mmol) and 4-DMAP (0.05 equiv, 5.7 mg, 0.046 mmol) were added to the reaction mixture, which was then stirred at room temperature for 24 h. The reaction mixture was concentrated in vacuo and the crude residue was purified by flash chromatography (3:7, acetone/hexanes) to provide compound **3-10** as a colorless oil (78 mg, 37% yield).

Colorless oil:  $R_f = 0.22$  (3:7, acetone/hexanes)

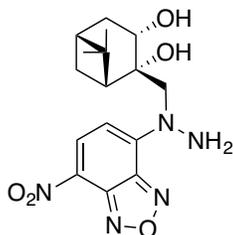
$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 4.88 (dd,  $J = 8.3, 1.7$  Hz, 1H), 3.97 – 3.81 (comp m, 2H), 2.43 – 2.29 (comp m, 3H), 2.20 – 1.97 (comp m, 4H), 1.34 (s, 3H), 1.30 (d,  $J = 11.4$  Hz, 1H), 0.89 (s, 3H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 153.6, 87.3, 75.0, 57.1, 48.0, 40.1, 37.8, 37.8, 32.4, 25.8, 24.6, 22.8.

IR (Microscope,  $\text{cm}^{-1}$ ): 3465, 2934, 1790, 1389, 1051.

HRMS (ESI-TOF) for  $\text{C}_{12}\text{H}_{18}\text{NaO}_4$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 249.1097; *found*: 249.1099.

### Synthesis of (–)-nopoldiol NBD **1** (**3-1h**)



**(1R,2S,3S,5R)-6,6-Dimethyl-2-((1-(7-nitrobenzo[c][1,2,5]oxadiazol-4-yl)hydrazinyl)methyl)bicyclo[3.1.1]heptane-2,3-diol (**3-1h**):** To a solution of epoxide **3-13** (1.0 equiv, 14 mg, 0.076 mmol) in  $\text{CHCl}_3$  (0.5 ml) was added  $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$  (20 equiv, 74  $\mu\text{l}$ , 1.5 mmol) under inert atmosphere at room temperature. The reaction mixture was refluxed for 1 h at 90 °C and then concentrated in vacuo. The crude residue was kept under vacuum for 2 h and used in the next step without further purification. Ethanol (6.0 ml) and  $\text{CH}_2\text{Cl}_2$  (1.0 ml) were added to the crude residue. Last, 4-chloro-7-nitrobenzofurazan (1.3 equiv, 20 mg, 0.099 mmol) was added to the reaction mixture, which was stirred for 1 h at room temperature. The dark red reaction mixture was concentrated in vacuo and the crude residue was purified by flash

chromatography (2:98 to 5:95, methanol/CH<sub>2</sub>Cl<sub>2</sub>) to provide diol **3-1h** as a reddish powder (10 mg, 36% yield).

Red powder:  $R_f = 0.17$  (5:95, methanol/CH<sub>2</sub>Cl<sub>2</sub>)

Melting point: decomposition (turning into black) was observed around 150 °C.

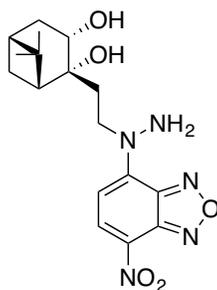
<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 8.47 (d,  $J = 9.3$  Hz, 1H), 6.91 (d,  $J = 9.3$  Hz, 1H), 4.48, 4.46 (ABq,  $J_{AB} = 15$  Hz, 2H), 4.34 (dd,  $J = 9.4, 5.6$  Hz, 1H), 2.50 (dddd,  $J = 13.5, 9.5, 3.8, 1.9$  Hz, 1H), 2.27 – 2.16 (comp m, 2H), 1.88 – 1.95 (m,  $J = 5.8, 3.7, 2.3$  Hz, 1H), 1.68 (ddd,  $J = 13.8, 5.7, 2.3$  Hz, 1H), 1.54 – 1.43 (m, 1H), 1.30 (s, 3H), 1.16 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 149.3, 146.3, 145.3, 137.0, 122.1, 103.9, 78.4, 67.2, 66.1, 52.8, 41.9, 39.9, 38.9, 28.9, 28.3, 24.6.

IR (cast film, cm<sup>-1</sup>): 3509, 3334, 3254, 3114, 2924, 2888, 1605, 1545, 1314, 1279, 1249.

HRMS (ESI-TOF): for C<sub>16</sub>H<sub>21</sub>N<sub>5</sub>NaO<sub>5</sub> (M + Na)<sup>+</sup>: *calcd.*: 386.1435; *found*: 386.1434.

### Synthesis of (–)-nopoldiol NBD 2 (**3-1i**)



#### (1*R*,2*R*,3*S*,5*R*)-6,6-Dimethyl-2-(2-(1-(7-nitrobenzo[*c*][1,2,5]oxadiazol-4-

yl)hydrazinyl)ethyl)bicyclo[3.1.1]heptane-2,3-diol (**3-1i**): To a solution of nopoldiol bromide **3-5** (1.0 equiv, 20 mg, 0.076 mmol) in CHCl<sub>3</sub> (0.5 ml) was added NH<sub>2</sub>NH<sub>2</sub>•H<sub>2</sub>O (20 equiv, 74 μl, 1.5 mmol) under inert atmosphere at room temperature. The reaction mixture was refluxed for 1 h at 75 °C and then concentrated in vacuo. The crude residue was kept under vacuum for 2 h and used in the next step without further purification. Ethanol (6.0 ml) and CH<sub>2</sub>Cl<sub>2</sub> (1.0 ml) were added to the crude residue. Last, 4-chloro-7-nitrobenzofurazan (1.3 equiv, 20 mg, 0.099 mmol) was added to the reaction mixture, which was stirred for 1 h at room temperature. The dark red reaction mixture was concentrated in vacuo and the crude residue was purified by flash chromatography (2:98 to 5:95, methanol/CH<sub>2</sub>Cl<sub>2</sub>) to provide diol **3-1i** as a reddish powder (8.5 mg, 32% yield).

Melting point: decomposition (turning into black) was observed around 150 °C.

Reddish powder:  $R_f = 0.30$  (5:95, methanol/ $\text{CH}_2\text{Cl}_2$ )

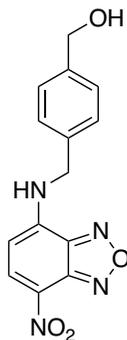
$^1\text{H NMR}$   $\delta/\text{ppm}$ : (500 MHz,  $\text{CD}_3\text{OD}$ ) 8.44 (d,  $J = 9.3$  Hz, 1H), 6.78 (bs, 1H), 4.62 – 4.29 (m, 2H), 4.16 (dd,  $J = 9.4, 5.3$  Hz, 1H), 2.55 – 2.46 (m, 1H), 2.26 – 2.10 (comp m, 3H), 2.03 – 1.88 (comp m, 2H), 1.69 (ddd,  $J = 13.9, 5.3, 2.5$  Hz, 1H), 1.46 (d,  $J = 10.1$  Hz, 1H), 1.32 (s, 3H), 0.94 (s, 3H). Note:  $\text{CD}_3\text{OD}/\text{CDCl}_3$  (0.50 ml/0.10 ml) solvent mixture was used to monitor both its  $^1\text{H NMR}$  and  $^{13}\text{C NMR}$ .

$^{13}\text{C NMR}$   $\delta/\text{ppm}$ : (126 MHz,  $\text{CD}_3\text{OD}$ ) 147.0, 146.1, 144.6, 137.0, 120.9, 102.9, 75.4, 68.2, 54.7, 52.6, 41.6, 39.7, 39.6, 38.9, 28.6, 28.2, 24.5.

**IR** (cast film,  $\text{cm}^{-1}$ ): 3513, 3297, 3218, 2990, 2980, 2942, 2911, 2868, 1647, 1599, 1553, 1307, 1278.

**HRMS (ESI-TOF)**: for  $\text{C}_{17}\text{H}_{23}\text{N}_5\text{NaO}_5$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 400.1591; *found*: 400.1591.

#### Synthesis of NBD-benzyl alcohol (3-14)



**(4-(((7-Nitrobenzo[*c*][1,2,5]oxadiazol-4-yl)amino)methyl)phenyl)methanol (3-14)**: To a solution of 4-aminomethylbenzyl alcohol (2.5 equiv, 0.52 g, 3.8 mmol) in ethanol/ $\text{CH}_2\text{Cl}_2$  (10 ml/2 ml) was added 4-chloro-7-nitrobenzofurazan (1.0 equiv, 0.30 g, 1.5 mmol) at room temperature under inert atmosphere. The reaction mixture was stirred for 2 h at room temperature and then concentrated in vacuo. The crude residue was purified by flash chromatography (dry loading) (3:7 to 7:3, ethyl acetate/hexanes) to provide compound **3-14** as a red-orange powder (0.34 g, 75% yield).

Red-orange powder:  $R_f = 0.44$  (7:3, EtOAc/hexanes).

Melting point: 177.7 – 180.1 °C.

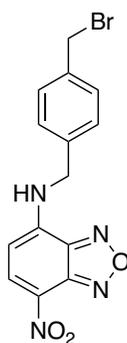
**<sup>1</sup>H NMR** δ/ppm: (500 MHz, DMSO-*d*<sub>6</sub>) 9.97 (s, 1H), 8.47 (d, *J* = 8.9 Hz, 1H), 7.36 (d, *J* = 8.1 Hz, 2H), 7.29 (d, *J* = 8.0 Hz, 2H), 6.34 (d, *J* = 8.9 Hz, 1H), 5.14 (app t, *J* = 5.7 Hz, 1H), 4.70 (app bs, 2H), 4.46 (d, *J* = 5.6 Hz, 2H).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, DMSO-*d*<sub>6</sub>) 144.8, 144.5, 144.1, 141.8, 137.7, 135.1, 127.0, 126.7, 121.3, 99.7, 62.6, 46.1.

**IR** (cast film, cm<sup>-1</sup>): 3579, 3504, 3204, 3145, 3056, 2983, 2915, 2872, 1614, 1595, 1529, 1454, 1324, 1307, 1233, 1021.

**HRMS (ESI-TOF)**: for C<sub>14</sub>H<sub>11</sub>N<sub>4</sub>O<sub>4</sub> (M – H)<sup>-</sup>: *calcd.*: 300.0856; *found*: 300.0859.

### NBD-benzyl bromide (3-15)



***N*-(4-(Bromomethyl)benzyl)-7-nitrobenzo[*c*][1,2,5]oxadiazol-4-amine (3-15)**: To a mixture of compound **3-14** (1.0 equiv, 0.33 g, 1.1 mmol) and triphenylphosphine (1.5 equiv, 0.43 g, 1.6 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (12 ml) was added NBS (1.5 equiv, 0.29 g, 1.6 mmol) partially under ice bath and inert atmosphere. The reaction mixture was then brought to rt and stirred for 16 h. Then, the reaction mixture was concentrated in vacuo. The crude residue was purified by flash chromatography (dry loading) (15:85 to 100:0, EtOAc/hexanes) to provide compound **3-15** as an orange powder (0.16 g, 39% yield).

Orange powder: *R<sub>f</sub>* = 0.55 (1:1, EtOAc/Hexanes)

Melting point: 142.0 – 143.9 °C.

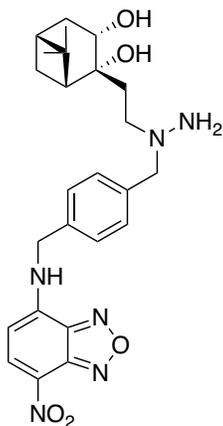
**<sup>1</sup>H NMR** δ/ppm: (500 MHz, DMSO-*d*<sub>6</sub>) 9.95 (d, *J* = 6.7 Hz, 1H), 8.47 (d, *J* = 8.9 Hz, 1H), 7.41 (app q, *J* = 8.2 Hz, 4H), 6.34 (d, *J* = 8.9 Hz, 1H), 4.72 (s, 2H), 4.68 (s, 2H).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, DMSO-*d*<sub>6</sub>) 144.7, 144.3, 143.9, 137.6, 137.0, 136.9, 129.4, 127.4, 121.3, 99.5, 54.7, 45.8, 34.0.

**IR** (cast film,  $\text{cm}^{-1}$ ): 3312, 3076, 3062, 3032, 2979, 1617, 1574, 1486, 1447, 1343, 1305, 1271, 1220, 1184.

**HRMS (ESI-TOF)**: for  $\text{C}_{14}\text{H}_{10}\text{BrN}_4\text{O}_3$  ( $\text{M} - \text{H}$ )<sup>-</sup>: *calcd.*: 360.9942; *found*: 360.9937.

### Synthesis of (–)-nopoldiol NBD 3 (3-1j)



**(1*R*,2*R*,3*S*,5*R*)-6,6-Dimethyl-2-(2-(1-(4-(((7-nitrobenzo[*c*][1,2,5]oxadiazol-4-yl)amino)methyl)benzyl)hydrazinyl)ethyl)bicyclo[3.1.1]heptane-2,3-diol (3-1j)**: To a solution of compound **3-5** (1.0 equiv, 40 mg, 0.15 mmol) in  $\text{CHCl}_3$  (1.0 ml) was added  $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$  (20 equiv, 150  $\mu\text{l}$ , 3.0 mmol) under inert atmosphere at room temperature. The reaction mixture was refluxed for 1 h at 75 °C and then concentrated in vacuo. The crude residue was kept under vacuum for 2 h and used in the next step without further purification. Compound **3-15** (1.1 equiv, 61 mg, 0.17 mmol) and  $\text{K}_2\text{CO}_3$  (1.5 equiv, 31 mg, 0.23 mmol) were added to the crude residue in  $\text{ACN}/\text{CH}_2\text{Cl}_2/\text{water}$  (0.50 ml/0.25 ml/0.050 ml) at room temperature under inert atmosphere. Then, the reaction mixture was stirred for 16 h at room temperature and concentrated in vacuo. The crude residue was purified by flash chromatography (0.25:9.75,  $\text{MeOH}/\text{CH}_2\text{Cl}_2$  in 0.05% ammonium hydroxide solution) to provide compound **3-1j** as an orange powder (43 mg, 56% yield). Its  $R_f$  value could not be reported. Note: The compound **3-1j** is unstable in the column so a rapid isolation might be helpful.

Melting point: decomposition (turning into black) was observed around 150 °C.

**$^1\text{H}$  NMR**  $\delta/\text{ppm}$ : (500 MHz,  $\text{CDCl}_3$ ) 8.47 (d,  $J = 7.7$  Hz, 1H), 7.38, 7.35 (ABq,  $J_{AB} = 5.0$  Hz, 4H), 6.63 (bs, 1H), 6.21 (d,  $J = 8.5$  Hz, 1H), 4.67 (d,  $J = 4.4$  Hz, 2H), 4.03 (dd,  $J = 9.3, 5.5$  Hz, 1H), 3.76, 3.71 (ABq,  $J_{AB} = 10$  Hz, 2H), 2.98 – 2.88 (m, A part of ABMM', 1H), 2.87 – 2.77 (m,

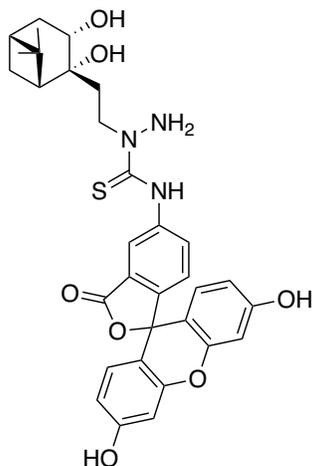
B part of ABMM', 1H), 2.50 – 2.38 (m, 1H), 2.19 – 2.12 (m, 1H), 2.03 (app t,  $J = 5.9$  Hz, 1H), 1.99 – 1.74 (comp m, 4H), 1.70 – 1.56 (comp m, 2H), 1.43 (d,  $J = 10.2$  Hz, 1H), 1.28 – 1.22 (m, 3H), 0.93 (s, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 144.5, 144.0, 143.4, 137.2, 136.3, 134.8, 130.5, 128.2, 125.0, 99.6, 76.4, 67.8, 66.5, 56.5, 54.1, 47.9, 40.6, 38.9, 38.3, 37.8, 28.1, 27.8, 24.4.

IR (cast film,  $\text{cm}^{-1}$ ): 3330, 3240, 3155, 3069, 2979, 2920, 2866, 1621, 1584, 1497, 1443, 1300, 1276.

HRMS (ESI-TOF): for  $\text{C}_{25}\text{H}_{33}\text{N}_6\text{O}_5$  ( $\text{M} + \text{H}^+$ ): *calcd.*: 497.2507; *found*: 497.25; for  $\text{C}_{25}\text{H}_{31}\text{N}_6\text{O}_5$  ( $\text{M} - \text{H}^-$ ): *calcd.*: 495.2361; *found*: 495.2355.

### Synthesis of (–)-nopoldiol fluorescein thiosemicarbazide (3-1k)



***N*-(3',6'-Dihydroxy-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthen]-5-yl)-1-(2-((1*R*,2*R*,3*S*,5*R*)-2,3-dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl)hydrazine-1-carbothioamide (3-1k):** To a solution of compound 3-5 (1.0 equiv, 30 mg, 0.11 mmol) in  $\text{CHCl}_3$  (1.0 ml) was added  $\text{NH}_2\text{NH}_2 \cdot \text{H}_2\text{O}$  (20 equiv, 110  $\mu\text{l}$ , 2.3 mmol) under inert atmosphere at room temperature. The reaction mixture was refluxed for 75 min at 75  $^\circ\text{C}$  and then concentrated in vacuo. The crude residue was kept under vacuum for 2 h and used in the next step without further purification. Then,  $\text{Et}_3\text{N}$  (1.2 equiv, 21  $\mu\text{l}$ , 0.15 mmol) was added to the crude residue in DMF/ $\text{CH}_2\text{Cl}_2$ /methanol (0.45 ml/0.45 ml/0.050 ml) at room temperature under inert atmosphere. Finally, fluorescein isothiocyanate (0.9 equiv, 40 mg, 0.10 mmol) was added to the reaction mixture, which was then stirred for 16 h at room temperature. Then, the reaction mixture was concentrated in vacuo and purified using reverse phase HPLC flash chromatography (C8

column) with water and ACN under neutral conditions to provide **3-1k** as an orange powder (18 mg, 28% yield). It is important to note that nopoldiol derivatives were observed to form a small amount of complex boronate with boric acid from water if and only if the water was kept in a glass bottle for a long time. Simply, boric acid released from a glass bottle into water was found as the cause of this resulting boronate formation. Therefore, water was kept in a plastic bottle, and used for HPLC analysis and purifications to prevent this complex formation.

Melting point: decomposition (turning into black) was observed around 150 °C.

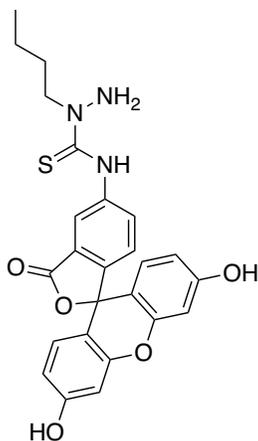
<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 8.31 (s, 1H), 7.86 (d, *J* = 8.2 Hz, 1H), 7.12 (d, *J* = 8.2 Hz, 1H), 6.75 – 6.62 (comp m, 4H), 6.54 (dd, *J* = 8.7, 2.4 Hz, 2H), 4.31 (app t, *J* = 7.7 Hz, 2H), 4.15 (dd, *J* = 9.4, 5.0 Hz, 1H), 2.55 – 2.42 (m, 1H), 2.25 – 2.07 (comp m, 3H), 2.02 – 1.92 (m, 1H), 1.90 (m, 1H), 1.69 (ddd, *J* = 13.9, 5.1, 2.6 Hz, 1H), 1.46 (d, *J* = 10.0 Hz, 1H), 1.30 (s, 3H), 1.03 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 180.7, 171.4, 161.7, 154.3, 143.0, 132.8, 130.4, 128.8, 125.1, 121.0, 113.8, 111.7, 103.5, 75.9, 68.6, 53.0, 51.5, 49.8, 41.9, 39.8, 30.8, 39.32, 39.26, 28.6, 28.4, 24.7.

IR (cast film, cm<sup>-1</sup>): 3248, 2918, 2850, 1735, 1606, 1506, 1451, 1319, 1209, 1178, 1111.

HRMS (ESI-TOF): for C<sub>32</sub>H<sub>34</sub>N<sub>3</sub>O<sub>7</sub>S (M + H)<sup>+</sup>: *calcd.*: 604.2112; *found*: 604.2117.

### Synthesis of butyl-thiosemicarbazide-FITC (**3-1m**)



**1-Butyl-N-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-yl)hydrazine-1-carbothioamide (3-1m):** To butylhydrazine•HCl (1.0 equiv, 14 mg, 0.11 mmol) and Et<sub>3</sub>N (1.2 equiv, 21 μl, 0.15 mmol) in DCM/MeOH (1.0 ml/0.3 ml) was added fluorescein isothiocyanate

(0.90 equiv, 40 mg, 0.10 mmol) at room temperature under inert atmosphere. The reaction mixture was then stirred for 1 h after which it was concentrated in vacuo and the crude residue was purified by flash chromatography (dry loading) (5:95, MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide compound **3-1m** as an orange powder (22 mg, 45% yield).

Orange powder:  $R_f = 0.32$  (10:90, MeOH/ CH<sub>2</sub>Cl<sub>2</sub>)

Melting point: decomposition (turning into black) was observed around 150 °C.

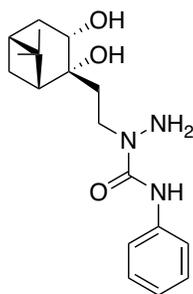
<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 8.30 (d,  $J = 1.8$  Hz, 1H), 7.86 (dd,  $J = 8.2, 1.9$  Hz, 1H), 7.13 (d,  $J = 8.2$  Hz, 1H), 6.75 – 6.63 (comp m, 4H), 6.54 (dd,  $J = 8.7, 2.4$  Hz, 2H), 4.17 (t,  $J = 7.6$  Hz, 2H), 1.78 (p,  $J = 7.6$  Hz, 2H), 1.42 (h,  $J = 7.4$  Hz, 2H), 1.02 (t,  $J = 7.4$  Hz, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 181.2, 171.4, 161.6, 154.3, 143.1, 133.0, 130.4, 128.8, 125.1, 121.2, 113.7, 111.7, 103.5, 54.8, 29.2, 20.9, 14.3. (One missing carbon must overlap with one of the existing peaks.)

IR (cast film, cm<sup>-1</sup>): 3256, 3050, 2959, 2928, 2871, 1732, 1600, 1507, 1453, 1383, 1316, 1174, 1104, 1081.

HRMS (ESI-TOF): for C<sub>25</sub>H<sub>24</sub>N<sub>3</sub>O<sub>5</sub>S (M + H)<sup>+</sup>: *calcd.*: 478.1431; *found*: 478.1424.

### Synthesis of (–)-nopoldiol benzyl semicarbazide (**3-1n**)



**1-(2-((1R,2R,3S,5R)-2,3-Dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl)-N-phenylhydrazine-1-carboxamide (3-1n):** To a solution of diol **3-5** (1.0 equiv, 30 mg, 0.11 mmol) in CHCl<sub>3</sub> (1.0 ml) was added NH<sub>2</sub>NH<sub>2</sub>•H<sub>2</sub>O (20 equiv, 110 μl, 2.3 mmol) under inert atmosphere at room temperature. The reaction mixture was refluxed for 75 min at 75 °C and then concentrated in vacuo. The crude residue was kept under vacuum for 2 h and used in the next step without further purification. Then, Et<sub>3</sub>N (1.2 equiv, 21 μl, 0.15 mmol) was added to the crude residue in CH<sub>2</sub>Cl<sub>2</sub>/methanol (1.0 ml/0.1 ml) at room temperature under inert atmosphere. Finally, phenyl isocyanate (0.90 equiv, 11 μl, 0.10 mmol) was added to the reaction mixture,

which was then stirred for 1 h at room temperature. The reaction mixture was concentrated in vacuo and the crude residue was purified by flash chromatography (dry loading) (2:98, MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide diol **3-1n** as a white solid (19 mg, 55% yield).

White solid:  $R_f = 0.14$  (2:98, MeOH/CH<sub>2</sub>Cl<sub>2</sub>)

Melting point: 85.5 – 88.0 °C.

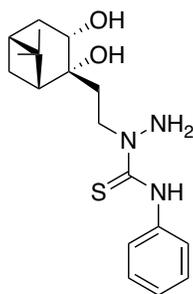
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 8.67 (s, 1H), 7.44 (d,  $J = 7.3$  Hz, 2H), 7.32 – 7.21 (m, 2H), 7.01 (t,  $J = 7.4$  Hz, 1H), 4.82 (bs, 1H), 4.45 (ddd,  $J = 15.1, 11.8, 3.3$  Hz, 1H), 4.24 (dd,  $J = 9.9, 4.8$  Hz, 1H), 4.10 (bs, 1H), 3.88 (bs, 2H), 3.27 (app dt,  $J = 15.3, 4.2$  Hz, 1H), 2.47 (app ddt,  $J = 16.2, 9.7, 3.0$  Hz, 1H), 2.19–2.11 (m, 1H), 1.99 (app t,  $J = 5.8$  Hz, 1H), 1.96 – 1.85 (comp m, 2H), 1.74–1.62 (comp m, 2H), 1.34 (d,  $J = 10.4$  Hz, 1H), 1.26 (s, 3H), 0.93 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 157.8, 138.8, 129.0, 123.0, 119.2, 75.1, 65.3, 54.2, 44.3, 40.5, 38.7, 37.5, 36.9, 27.9, 27.6, 24.6.

IR (cast film, cm<sup>-1</sup>): 3368, 3338, 3214, 3058, 2991, 2974, 2941, 2919, 1652, 1592, 1521, 1449.

HRMS (ESI-TOF): for C<sub>18</sub>H<sub>27</sub>N<sub>3</sub>NaO<sub>3</sub> (M + Na)<sup>+</sup>: *calcd.*: 356.1945; *found*: 356.1947; for C<sub>18</sub>H<sub>28</sub>N<sub>3</sub>O<sub>3</sub> (M + H)<sup>+</sup>: *calcd.*: 334.2125; *found*: 334.2123.

### Synthesis of (–)-nopoldiol benzyl thiosemicarbazide (**3-1o**)



**1-(2-((1R,2R,3S,5R)-2,3-Dihydroxy-6,6-dimethylbicyclo[3.1.1]heptan-2-yl)ethyl)-N-phenylhydrazine-1-carbothioamide (**3-1o**):** To a solution of diol **3-5** (1.0 equiv, 30 mg, 0.11 mmol) in CHCl<sub>3</sub> (1.0 ml) was added NH<sub>2</sub>NH<sub>2</sub>•H<sub>2</sub>O (20 equiv, 110 μl, 2.3 mmol) under inert atmosphere at room temperature. The reaction mixture was refluxed for 75 min at 75 °C and then concentrated in vacuo. The crude residue was kept under vacuum for 2 h and used in the next step without further purification. Then, Et<sub>3</sub>N (1.2 equiv, 21 μl, 0.15 mmol) was added to the crude residue in CH<sub>2</sub>Cl<sub>2</sub>/methanol (1.0 ml/0.10 ml) at room temperature under inert atmosphere. Finally, phenyl isothiocyanate (0.90 equiv, 12 μl, 0.10 mmol) was added to the reaction mixture,

which was then stirred for 1 h. The reaction mixture was then concentrated in vacuo and the crude residue was purified by flash chromatography (dry loading) (1:99, MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide diol **3-10** as a white solid (21 mg, 58% yield).

White solid:  $R_f = 0.17$  (0.25:9.75, MeOH/CH<sub>2</sub>Cl<sub>2</sub>)

Melting point: 115.1 – 117.6 °C.

<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 9.84 (s, 1H), 7.48 (d,  $J = 7.2$  Hz, 2H), 7.35 (app t,  $J = 7.9$  Hz, 2H), 7.20 (t,  $J = 7.4$  Hz, 1H), 5.70 – 5.61 (m, 1H), 4.59 (d,  $J = 5.4$  Hz, 1H), 4.36 – 4.25 (comp m, 2H), 4.15 (s, 2H), 3.40 (app dt,  $J = 15.0, 4.1$  Hz, 1H), 2.54 – 2.46 (m, 1H), 2.23 – 2.14 (m, 1H), 2.13 – 2.04 (m, 1H), 2.02 (app t,  $J = 5.8$  Hz, 1H), 1.91 (app tt,  $J = 5.9, 3.0$  Hz, 1H), 1.85 (app dt,  $J = 15.3, 3.8$  Hz, 1H), 1.69 (ddd,  $J = 14.2, 4.8, 2.6$  Hz, 1H), 1.36 (d,  $J = 10.4$  Hz, 1H), 1.28 (s, 3H), 0.94 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 180.2, 139.0, 128.7, 126.0, 125.4, 75.2, 64.6, 53.8, 48.5, 40.5, 38.7, 37.9, 37.7, 27.8, 27.6, 24.6.

IR (cast film, cm<sup>-1</sup>): 3530, 3292, 3194, 3048, 2990, 2972, 2920, 2864, 1641, 1598, 1517, 1307, 1046.

HRMS (ESI-TOF): for C<sub>18</sub>H<sub>27</sub>N<sub>3</sub>NaO<sub>2</sub>S (M + Na)<sup>+</sup>: *calcd.*: 372.1716; *found*: 372.1717; for C<sub>18</sub>H<sub>28</sub>N<sub>3</sub>O<sub>2</sub>S (M + H)<sup>+</sup>: *calcd.*: 350.1897; *found*: 350.1908.

### Synthesis of 2-acetyl-5-methoxy bromobenzene (**3-21**)



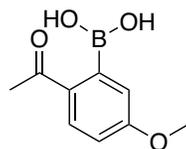
**1-(2-Bromo-4-methoxyphenyl)ethan-1-one (**3-21**):** The compound **3-21** was synthesized based on the procedure described earlier by Wang and coworkers and its NMR data was confirmed with the literature values.<sup>56</sup>

Yellowish liquid. (1.6 g, 65% yield)  $R_f = 0.36$  (15:85, EtOAc/Hexanes)

<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 7.58 (d,  $J = 8.7$  Hz, 1H), 7.14 (d,  $J = 2.5$  Hz, 1H), 6.86 (dd,  $J = 8.7, 2.5$  Hz, 1H), 3.83 (s, 3H), 2.61 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) δ 198.1, 160.9, 131.6, 130.7, 120.3, 118.5, 112.1, 54.7, 29.0.

## Synthesis of 2-acetyl-5-methoxy-phenylboronic acid (3-2j)



**(2-Acetyl-5-methoxyphenyl)boronic acid (3-2j):** Compound **3-21** (1.0 equiv, 0.23 g, 1.0 mmol), bis(pinacolato)diboron (1.1 equiv, 0.28 g, 1.1 mmol) and potassium acetate (3.0 equiv, 0.30 g, 3.0 mmol) were added to a flame dried round bottom flask, degassed and purged with nitrogen. Then, freshly distilled dioxane (4 ml) was added under nitrogen. Lastly, [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium(II) complex with dichloromethane (Pd(dppf)Cl<sub>2</sub>) (3 mol%, 24 mg, 0.030 mmol) was added and the reaction mixture was stirred at 90 °C for 16 h. After the reaction mixture reached to rt, CH<sub>2</sub>Cl<sub>2</sub> (20 ml) was added to the reaction mixture. The organic phase was washed with distilled water (1 x 5 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. After the crude residue was exposed to a short silica plug and washed with EtOAc, it was concentrated in vacuo. Without further purification, the crude residue and phenylboronic acid (5.0 equiv, 0.61 g, 5.0 mmol) were mixed in ACN/HCl (1M) (14 ml/1.5 ml) at room temperature under ambient atmosphere, and the reaction mixture was stirred for 24 h. Then, it was concentrated in vacuo and the crude residue was purified by flash chromatography (5:95, MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide compound **3-2j** as a yellow solid (42 mg, 22% yield over two steps).

Yellow solid: R<sub>f</sub> = 0.16 (5:95, MeOH/CH<sub>2</sub>Cl<sub>2</sub>), monitored with curcumin dye.

**<sup>1</sup>H NMR** δ/ppm: (500 MHz, CD<sub>3</sub>OD) 8.04, 7.99, 7.96 (d, *J* = 8.6 Hz, total 1H, mixture of three isomers), 7.06 – 6.90 (comp m, total 2H), 3.90, 3.883, 3.878 (s, total 3H, a mixture of three isomers), 2.62, 2.59, 2.57 (s, total 3H). Solvent mixture CD<sub>3</sub>OD/D<sub>2</sub>O (0.5 ml/0.1 ml) was used for both <sup>1</sup>H NMR and <sup>13</sup>C NMR. The compound was observed as a mixture of three different isomers due to the presence of acetyl ketone (39/47/12% ratios according to their integral values).

**<sup>13</sup>C NMR** δ/ppm: (126 MHz, CD<sub>3</sub>OD) 202.8, 165.9, 165.3, 164.9, 133.8, 133.7, 132.9, 132.8, 132.7, 117.7, 117.4, 117.1, 114.6, 114.2, 113.9, 56.3, 56.2, 56.1, 24.76, 24.68, 24.17. Three different peaks were observed for most of the carbons due to the presence of acetyl ketone. The

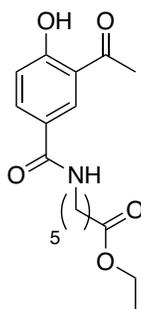
12% isomer was barely observed for quaternary carbons, however it was clearly observed for the corresponding  $-\text{CH}_3$  peaks.

$^{11}\text{B}$  NMR  $\delta$ /ppm: (128 MHz,  $\text{CD}_3\text{OD}$ ) 27.8.

IR (cast film,  $\text{cm}^{-1}$ ): 3393, 3070, 3005, 2961, 2939, 1654, 1593, 1559, 1359, 1319, 1270, 1227, 1033.

HRMS (ESI-TOF): for  $\text{C}_9\text{H}_{10}[\text{11B}]\text{O}_4$  ( $\text{M} - \text{H}$ ) $^-$ : *calcd.*: 193.0678; *found*: 193.0681.

### Synthesis of boronate 3-2k



**Ethyl 6-(3-acetyl-4-hydroxybenzamido)hexanoate (3-18)**: Compound **3-16** was synthesized according to the procedure described earlier by Wirz and coworkers.<sup>42</sup> Compound **3-16** (1.0 equiv, 0.41 g, 2.3 mmol) and HATU coupling reagent (1.0 equiv, 0.86 g, 2.3 mmol) were dissolved in dry DMF (20 ml) at room temperature under argon balloon. Then, compound **3-17** (1.0 equiv, 0.44 g, 2.3 mmol) and then DIPEA (3.2 equiv, 1.3 ml, 7.3 mmol) were added. The reaction mixture was then stirred for 16 h at room temperature, concentrated in vacuo and mixed with  $\text{CH}_2\text{Cl}_2$  (75 ml). The organic phase was washed with HCl (1  $\times$  10 ml, 1 N), and brine (1  $\times$  10 ml), dried ( $\text{MgSO}_4$ ), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:9 to 5:5, EtOAc/hexanes) to provide compound **3-18** as a dark purple viscous product (0.52 g, 71% yield).

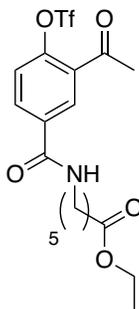
Dark purple viscous product:  $R_f = 0.29$  (1:1, EtOAc/hexanes)

$^1\text{H}$  NMR  $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 8.33 (d,  $J = 2.2$  Hz, 1H), 7.80 (dd,  $J = 8.7, 2.2$  Hz, 1H), 6.97 (d,  $J = 8.7$  Hz, 1H), 6.37 (t,  $J = 5.9$  Hz, 1H), 4.10 (q,  $J = 7.1$  Hz, 2H), 3.46 (td,  $J = 7.2, 5.9$  Hz, 2H), 2.68 (s, 3H), 2.31 (t,  $J = 7.2$  Hz, 2H), 1.70 – 1.59 (comp m, 4H), 1.48 – 1.34 (m, 2H), 1.23 (t,  $J = 7.1$  Hz, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 204.9, 173.9, 166.1, 164.7, 133.9, 131.2, 125.5, 119.5, 118.5, 60.5, 39.9, 34.1, 29.2, 26.9, 26.4, 24.3, 14.4.

IR (cast film,  $\text{cm}^{-1}$ ): 3313, 3074, 2936, 2865, 1734, 1643, 1545, 1483, 1369, 1289.

HRMS (ESI-TOF): for  $\text{C}_{17}\text{H}_{23}\text{NNaO}_5$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 344.1468; *found*: 344.147; for  $\text{C}_{17}\text{H}_{24}\text{NO}_5$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 322.1649; *found*: 322.1658.



**Ethyl 6-(3-acetyl-4-((trifluoromethyl)sulfonyl)oxy)benzamido)hexanoate (3-19):** *N*-Phenylbis(trifluoromethanesulfonimide) (1.3 equiv, 0.76 g, 2.1 mmol) was added portion wise to a solution of compound **3-18** (1.0 equiv, 0.53 g, 1.6 mmol) and  $\text{Et}_3\text{N}$  (4.0 equiv, 0.92 ml, 6.6 mmol) in dry DMF (5.0 ml) at room temperature under argon balloon, and the reaction mixture was stirred for 16 h. Then, the reaction mixture was concentrated in vacuo and purified by flash chromatography (1:3 to 1:1, EtOAc/hexanes) to provide compound **3-19** as a dark red viscous product (0.72 g, 97% yield).

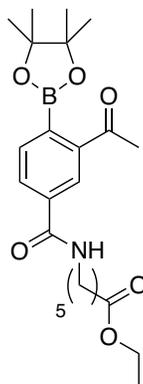
Dark red viscous product:  $R_f = 0.45$  (1:1, EtOAc/hexanes)

$^1\text{H}$  NMR  $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 8.23 (d,  $J = 2.3$  Hz, 1H), 8.01 (dd,  $J = 8.5, 2.3$  Hz, 1H), 7.40 (d,  $J = 8.5$  Hz, 1H), 6.49 (t,  $J = 5.7$  Hz, 1H), 4.11 (q,  $J = 7.1$  Hz, 2H), 3.50 (td,  $J = 7.0, 5.7$  Hz, 2H), 2.67 (s, 3H), 2.33 (t,  $J = 7.2$  Hz, 2H), 1.71 – 1.60 (comp m, 4H), 1.43 (tt,  $J = 9.9, 6.3$  Hz, 2H), 1.24 (t,  $J = 7.1$  Hz, 3H).

$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 196.2, 173.9, 165.0, 148.5, 135.2, 132.4, 132.2, 129.7, 127.4, 123.6, 123.1, 122.5, 120.0, 117.44, 114.8, 60.6, 40.0, 34.1, 29.8, 28.9, 26.3, 24.1, 14.4.

IR (cast film,  $\text{cm}^{-1}$ ): 3322, 3083, 2938, 2867, 1733, 1703, 1645, 1545, 1429, 1250, 1213, 1164, 1140.

HRMS (ESI-TOF): for  $\text{C}_{18}\text{H}_{22}\text{F}_3\text{NNaO}_7\text{S}$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 476.0961; *found*: 476.0967; for  $\text{C}_{18}\text{H}_{23}\text{F}_3\text{NO}_7\text{S}$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 454.1142 ; *found*: 454.1149.



**Ethyl-6-(3-acetyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzamido)hexanoate (3-20):** The procedure was adapted from the previous report.<sup>43</sup> Compound **3-19** (1.0 equiv, 0.34 g, 0.75 mmol), bis(pinacolato)diboron (2.0 equiv, 0.38 g, 1.5 mmol), sodium acetate (3.0 equiv, 0.18 g, 2.2 mmol) and [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium(II) complex with dichloromethane (Pd(dppf)Cl<sub>2</sub>) (10 mol%, 55 mg, 0.075 mmol) were added to a flame dried round bottom flask, and flushed 3 times with argon. Then, freshly distilled dioxane (6 ml) was added and degassed for 10 min with bubbling argon. The reaction mixture was stirred and heated at 95 °C for 90 min, after which it was concentrated in vacuo. The crude residue was purified by flash chromatography (3:7, EtOAc/CH<sub>2</sub>Cl<sub>2</sub>) to provide compound **3-20** as a colorless viscous product (0.17 g, 52% yield).

Colorless viscous product:  $R_f = 0.37$  (4:6, EtOAc/CH<sub>2</sub>Cl<sub>2</sub>)

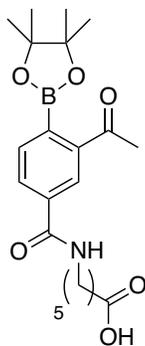
<sup>1</sup>H NMR δ/ppm: (500 MHz, CDCl<sub>3</sub>) 8.31 (s, 1H), 7.83 (d,  $J = 7.6$  Hz, 1H), 7.55 (d,  $J = 7.6$  Hz, 1H), 6.37 (t,  $J = 5.9$  Hz, 1H), 4.12 (q,  $J = 7.1$  Hz, 2H), 3.48 (app q,  $J = 6.8$  Hz, 2H), 2.66 (s, 3H), 2.33 (t,  $J = 7.3$  Hz, 2H), 1.72 – 1.62 (m, 4H), 1.47 – 1.37 (comp m, 14H), 1.24 (t,  $J = 7.1$  Hz, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) δ 199.6, 173.8, 166.5, 141.3, 135.4, 132.5, 129.8, 127.6, 84.2, 60.5, 40.0, 34.2, 29.2, 26.4, 25.8, 25.0, 24.4, 14.4.

<sup>11</sup>B NMR δ/ppm: (128 MHz, CDCl<sub>3</sub>) 31.2.

IR (cast film, cm<sup>-1</sup>): 3316, 3070, 2977, 2934, 2866, 1734, 1681, 1640, 1540, 1372, 1344, 1302, 1145.

HRMS (ESI-TOF): for C<sub>23</sub>H<sub>34</sub>[<sup>11</sup>B]NNaO<sub>6</sub> (M + Na)<sup>+</sup>: *calcd.*: 454.2371; *found*: 454.237.



**6-(3-Acetyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzamido)hexanoic acid (3-2k):**

Compound **3-20** (88 mg, 0.20 mmol) was mixed in a solution of TFA (0.1 ml) and distilled water (0.95 ml) under argon balloon, and the reaction mixture was stirred and heated at 90 °C for 2.5 h. Then, the reaction mixture was concentrated in vacuo and the crude residue was purified by flash chromatography (EtOAc) to provide compound **3-2k** as a colorless viscous product (76 mg, 92% yield).

Colorless viscous product:  $R_f$  = it is a smear spot close to baseline monitored by curcumin dye.

**$^1\text{H NMR}$**   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 8.29 (s, 1H), 7.82 (d,  $J = 7.6$  Hz, 1H), 7.55 (d,  $J = 7.6$  Hz, 1H), 6.47 (t,  $J = 5.3$  Hz, 1H), 3.45 (app q,  $J = 6.7$ , 2H), 2.65 (s, 3H), 2.37 (t,  $J = 7.1$  Hz, 2H), 1.73–1.59 (comp m, 4H), 1.48–1.38 (comp m, 14H).

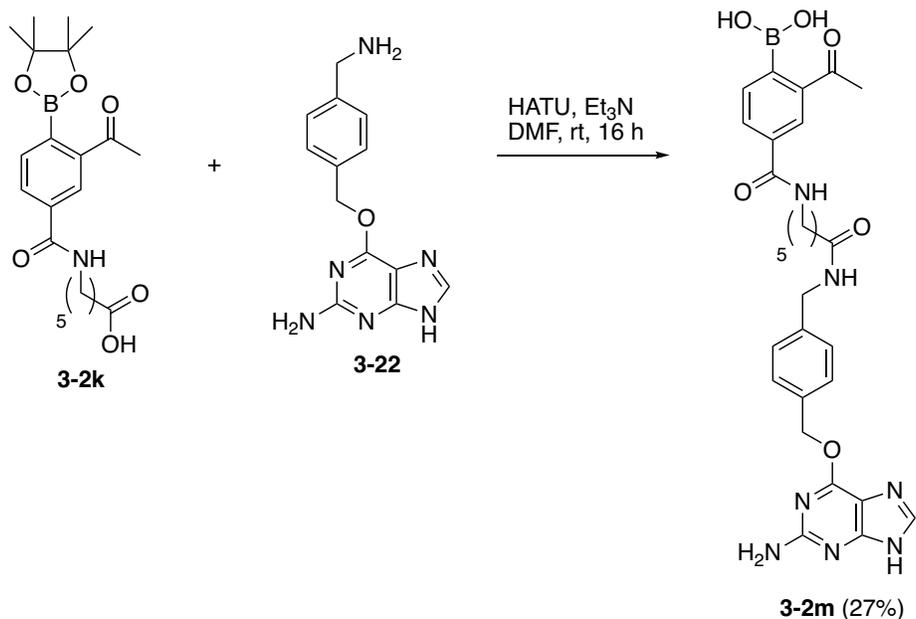
**$^{13}\text{C NMR}$**   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 199.7, 178.0, 166.8, 141.2, 135.2, 132.5, 129.9, 127.6, 84.2, 40.0, 33.8, 29.1, 26.3, 25.8, 25.0, 24.2.

**$^{11}\text{B NMR}$**   $\delta$ /ppm: (128 MHz,  $\text{CDCl}_3$ ) 31.6.

**IR** (cast film,  $\text{cm}^{-1}$ ): 3341, 2978, 2936, 1711, 1679, 1638, 1543, 1372, 1345, 1298, 1256, 1144.

**HRMS (ESI-TOF)**: for  $\text{C}_{21}\text{H}_{29}[\text{11B}]\text{NO}_6 (\text{M} - \text{H})^-$ : *calcd.*: 402.2093; *found*: 402.2096.

### Synthesis of snap-tag boronic acid reagent (3-2m)



**3-Acetyl-N-(6-((4-(((2-Amino-9H-purin-6-yl)oxy)methyl)benzyl)amino)-6-oxohexyl)-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzamide (3-2m):** Compound 3-2k (1.0 equiv, 32 mg, 0.079 mmol) and HATU coupling reagent (1.0 equiv, 30 mg, 0.079 mmol) were dissolved in dry DMF (0.80 ml) at room temperature under argon balloon. Compound 3-22 was synthesized according to the procedure described by Johnsson and coworkers.<sup>53</sup> Then, compound 3-22 (1.0 equiv, 21 mg, 0.079 mmol) and Et<sub>3</sub>N (3.2 equiv, 35  $\mu$ l, 0.25 mmol) were added. The reaction mixture was then stirred for 16 h, after which it was concentrated in vacuo and mixed. The crude residue was purified by reverse phase HPLC flash chromatography using C8 column with water and ACN under neutral conditions to provide 3-2m as a white powder (12 mg, 27% yield).

**<sup>1</sup>H NMR**  $\delta$ /ppm: (500 MHz, CD<sub>3</sub>OD) 8.46 (s, 1H), 8.39 (t,  $J$  = 6.0 Hz, 1H), 8.05 (dd,  $J$  = 7.8, 1.6 Hz, 1H), 7.83 (s, 1H), 7.50 (d,  $J$  = 7.5 Hz, 1H), 7.46 (d,  $J$  = 8.1 Hz, 2H), 7.28 (d,  $J$  = 8.3 Hz, 2H), 5.51 (s, 2H), 4.35 (s, 2H), 3.41 (t,  $J$  = 7.1 Hz, 2H), 2.66 (app s, 4H (acetyl -CH<sub>3</sub> and probably one of the labile protons), 2.27 (t,  $J$  = 7.4 Hz, 2H), 1.74 – 1.62 (comp m, 4H), 1.47 – 1.38 (m, 2H). Note: D<sub>2</sub>O was not added to this sample.

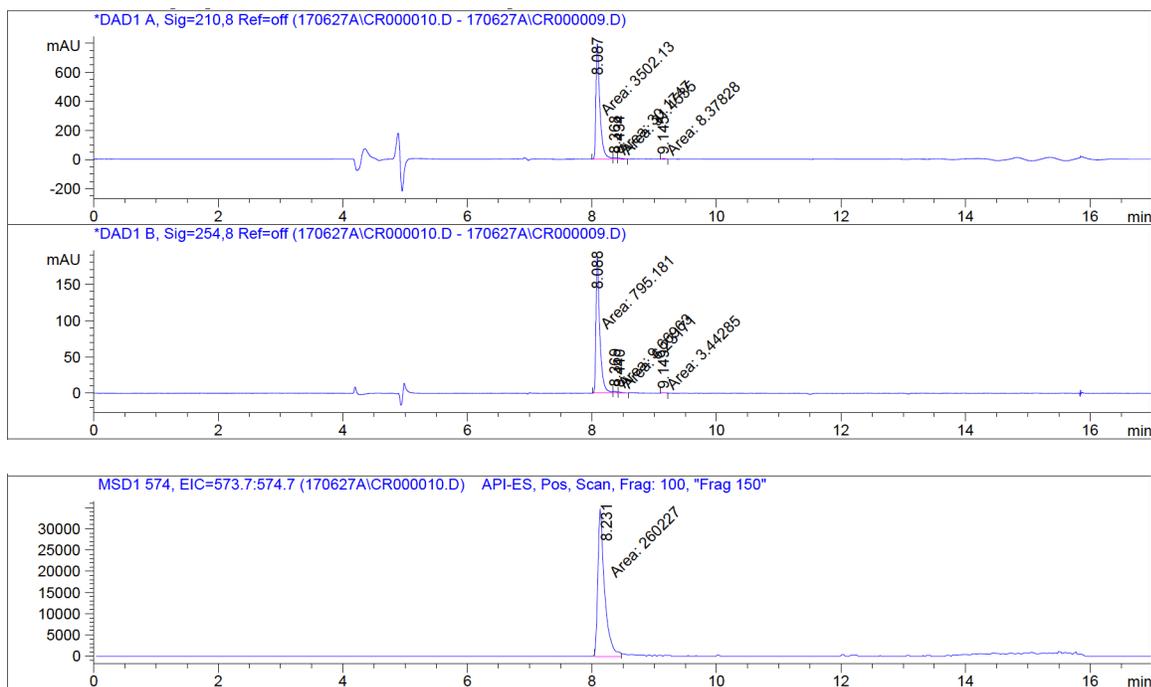
$^{13}\text{C}$  NMR  $\delta$ /ppm: (126 MHz,  $\text{CD}_3\text{OD}$ ) 202.2, 176.04, 175.96, 169.0, 161.7, 141.2, 140.1, 136.9, 136.2, 132.7, 132.1, 129.7, 128.8, 128.6, 68.6, 43.9, 43.8, 41.0, 40.4, 37.0, 36.9, 30.1, 27.6, 26.7.  
(Missing carbons must overlap with the existing peaks in aromatic region)

$^{11}\text{B}$  NMR  $\delta$ /ppm: 29.3.

IR (cast film,  $\text{cm}^{-1}$ ): 3324 (br), 2934, 2862, 1626, 1585, 1355, 1330, 1280.

HRMS (ESI-TOF): for  $\text{C}_{28}\text{H}_{32}[\text{11B}]\text{N}_7\text{NaO}_6$  ( $\text{M} + \text{Na}$ ) $^+$ : *calcd.*: 596.2; *found*: 596.2; for  $\text{C}_{28}\text{H}_{33}[\text{11B}]\text{N}_7\text{O}_6$  ( $\text{M} + \text{H}$ ) $^+$ : *calcd.*: 574.3; *found*: 574.3.

### HPLC-MS chromatogram:



### 3.4.3 Optimization of the synergic system

#### 3.4.3.1 Conjugation and stability study – HPLC-MS method

The HPLC data was acquired from an Agilent 1100 Series Instrument. Chromatographic separation was achieved on Agilent ZORBAX SB – C18 Column (250 x 46 mm, 5  $\mu\text{m}$ ; 1 ml/min, 40  $^\circ\text{C}$ ). The mobile phase consisted of 0.1% AA in  $\text{H}_2\text{O}$  (solvent A) and ACN (solvent B). A stepwise gradient was used for sample elution: 90% solvent A to 90% solvent B over a 20.6-min period. The sample injections were 2.0  $\mu\text{l}$ .

### **3.4.4 Design of fluorogenic nopoldiol derivatives with hydrazine**

#### **3.4.4.1 Conjugation and stability study – HPLC-MS method**

The HPLC data was acquired from an Agilent 1100 Series Instrument. Chromatographic separation was achieved on Agilent ZORBAX SB – C18 Column (150 x 4.6 mm, 5  $\mu$ m; 0.6 ml/min, 40 °C). The mobile phase consisted of 0.1% TFA in H<sub>2</sub>O (solvent A) and 0.1% AA ACN (solvent B). A stepwise gradient was used for sample elution: 90% solvent A to 90% solvent B over a 12-min period. The sample injections were 5.0 or 15  $\mu$ l.

### **3.4.5 Design of fluorescently labeled nopoldiol derivative with hydrazine**

#### **3.4.5.1 Conjugation and stability study – HPLC-MS method**

The HPLC data was acquired from an Agilent 1100 Series Instrument. Chromatographic separation was achieved on Agilent ZOBAX SB – C18 Column (150 x 4.6 mm, 5  $\mu$ m; 0.6 ml/min, 40 °C). Mass analyzer is Agilent G2908BA LC/MSD Quad. The mobile phase consisted of 0.1% TFA in H<sub>2</sub>O (solvent A) and 0.1% AA ACN (solvent B). A stepwise gradient was used for sample elution: 90% solvent A to 90% solvent B over a 12-min period. The sample injections were 5.0 or 15  $\mu$ l.

#### **3.4.5.2 Elemental composition analysis of the reaction 3-1j and 3-2c via high resolution-ESI**

HR-ESI positive ion mode was performed with the Agilent Technologies 6220 Accurate Mass oaTOF Dual ESI using ACN/water (1:1) as the carrier solvent and introducing the samples as direct flow injections. Data acquisition and processing were performed with Agilent Mass Hunter Workstation Data Acquisition and Mass Hunter Qualitative Data Analysis.

Equivalent volumes of diol **3-1j** (100  $\mu$ l, 0.5 mM in 10 mM ammonium acetate/ACN (65/35 w%, pH ~7) and boronic acid **3-2c** (100  $\mu$ l, 0.5 mM) were mixed in an eppendorf tube to reach the final concentration 0.25 mM. After its incubation for 30 min, 2 h, 3 h and 24 h at room temperature, the reaction mixture was monitored with high resolution ESI to detect and monitor **3-3jc•H<sub>2</sub>O** and **3-3jc** formations.

### 3.4.6 Design of fluorescein conjugated nopoldiol thiosemicarbazide

#### 3.4.6.1 Conjugation studies – <sup>1</sup>H-NMR analysis

Boronate and thiosemicarbazone formation were investigated by <sup>1</sup>H NMR in 0.05 M of D<sub>2</sub>O potassium phosphate buffer/ACN-d<sub>3</sub> (65/35 w%, pD 7.4). Individual stock solutions of diol **3-1k/3-1l** (0.5 mM) and boronic acid **3-2c** (0.5 mM) were prepared in this solvent system. Equivalent volumes of diol **3-1k/3-1l** (300 μl) and boronic acid **3-2c** (300 μl) were mixed in a NMR tube to reach the final concentration as 0.25 mM. After they were incubated for 10 min, 30 min, 60 min, 120 min and 24 h at room temperature, their % conversion for **3-1k/3-3kc** was monitored as explained previously.

Boronate and thiosemicarbazone formations were also investigated by <sup>1</sup>H NMR in 10 mM PBS buffer/D<sub>2</sub>O (9/1 v/v). <sup>1</sup>H NMR spectra were acquired on a 600 MHz four channel Agilent VNMRS spectrometer, equipped with a z-gradient HCN probe and using VNMJRJ 4.2A as the acquisition software. Samples were dissolved/recorded in 10 mM PBS buffer/D<sub>2</sub>O (9/1 v/v) at a concentration of 50 μM and a suppression of the H<sub>2</sub>O signal was performed using excitation sculpting.<sup>57,58</sup> All data were referenced to the water, set to be 4.72 ppm.

Individual stock solutions of diol **3-1k/controls 3-1l/3-1m** (0.1 mM) and boronic acid **3-2c** (0.1 mM) were prepared in this solvent system. Equivalent volumes of diol **3-1k/controls 3-1l/3-1m** (300 μl) and boronic acids **3-2c** (300 μl) were mixed in a NMR tube to reach the final concentration as 0.05 mM. After they were incubated for 10 min, 30 min, 60 min, 120 min, 3.5 h and 24 h at room temperature, their % conversion (**3-1k/3-3kc** or **3-2c/3-3lc**) was monitored as explained previously.

#### 3.4.6.2 Conjugation studies – HPLC-MS analysis

The HPLC data was acquired from an Agilent 1100 Series Instrument. Chromatographic separation was achieved on Agilent ZORBAX SB – C18 Column (150 x 4.6 mm, 5 μm; 0.6 ml/min, 40 °C). Mass analyzer is Agilent G2908BA LC/MSD Quad. The mobile phase consisted of 0.1% TFA in H<sub>2</sub>O (solvent A) and 0.1% AA ACN (solvent B). A stepwise gradient was used for sample elution: 90% solvent A to 90% solvent B over a 12-min period. The sample injections were 5.0, 15 or 25 μl.

Boronate and thiosemicarbazone formations between **3-1k** and **3-2c** were also investigated *via* HPLC-MS analysis in 10 mM PBS buffer at 50  $\mu$ M. Individual stock solutions of diol **3-1k** (100  $\mu$ M) and boronic acid **3-2c** (100  $\mu$ M) were prepared in this solvent system. Equivalent volumes of diol **3-1k** (100  $\mu$ l) and boronic acid **3-2c** (100  $\mu$ l) were mixed in a HPLC vial with a plastic insert to reach the final concentration as 50  $\mu$ M. After they were incubated for 10 min, 35 min, 60 min, 85 min, 110 min, 135 min and 24 h at room temperature, they were monitored *via* HPLC-MS (UV and fluorescence). The experiments were run in duplicate.

#### **3.4.6.3 Elemental composition analysis of 3-3kc/3-3lc/3-3mc formations *via* high resolution-ESI**

High resolution-ESI positive ion mode was performed with the Agilent Technologies 6220 Accurate Mass oaTOF Dual ESI using ACN/water (1:1) as the carrier solvent and introducing the samples as direct flow injections. Data acquisition and processing were performed with Agilent Mass Hunter Workstation Data Acquisition and Mass Hunter Qualitative Data Analysis.

Equivalent volumes of diol **3-1k/3-1l/3-1m** (100  $\mu$ l, 0.2 mM in 10 mM ammonium acetate/ACN (65/35 w%, pH ~7)) and boronic acid **3-2c** (100  $\mu$ l, 0.2 mM) were mixed in an eppendorf tube to reach the final concentration as 0.1 mM. After their incubation for 10 min, 30 min, 90 min and 24 h at room temperature, the reaction mixtures were monitored with high resolution-ESI.

#### **3.4.6.4 Stability studies of conjugate 3-3kc**

The HPLC data was acquired from an Agilent 1100 Series Instrument. Chromatographic separation was achieved on Agilent ZORBAX SB – C18 Column (150 x 4.6 mm, 5  $\mu$ m; 0.6 ml/min, 40  $^{\circ}$ C). The mobile phase consisted of 0.1% TFA in H<sub>2</sub>O (solvent A) and 0.1% AA ACN (solvent B). A stepwise gradient was used for sample elution: 90% solvent A to 90% solvent B over a 12-min period. The sample injections were 5.0, 15 or 25  $\mu$ l.

#### **Dilution study:**

Equivalent volumes of diol **3-1k** (100  $\mu$ l, 0.5 mM in 10 mM PBS buffer/ACN (65/35 w%, pH 7.4)) and boronic acid **3-2c** (100  $\mu$ l, 0.5 mM) were mixed in an eppendorf tube to reach

the final concentration as 0.25 mM. After the reaction mixture was incubated at room temperature for 24 h, it was diluted to 50 and 10  $\mu\text{M}$  with 10 mM PBS buffer (pH 7.4). After the diluted reaction mixtures were incubated for 4 h at room temperature, they were monitored and characterized with HPLC-MS. The experiments were run in duplicate.

#### **Exchange study:**

Equivalent volumes of diol **3-1k** (100  $\mu\text{l}$ , 100  $\mu\text{M}$  in 10 mM PBS buffer pH 7.4) and boronic acid **3-2c** (100  $\mu\text{l}$ , 100  $\mu\text{M}$  in 10 mM PBS buffer pH 7.4) were mixed in an eppendorf tube to reach the final concentration as 50  $\mu\text{M}$ . After the reaction mixture was incubated at room temperature for 24 h, boronic acid **3-2j** or **3-2a** (25  $\mu\text{l}$ , 12 mM) was introduced into the reaction mixture to reach the final concentration as  $\sim$ 100  $\mu\text{M}$ . Then, they were incubated for 4 h at room temperature, and the reaction mixtures were monitored and characterized with HPLC-MS. The experiments were run in duplicate.

#### **Acidification study to pH 3:**

Equivalent volumes of diol **3-1k** (100  $\mu\text{l}$ , 100  $\mu\text{M}$  in 10 mM PBS buffer pH 7.4) and boronic acid **3-2c** (100  $\mu\text{l}$ , 100  $\mu\text{M}$  in 10 mM PBS buffer pH 7.4) were mixed in an eppendorf tube to reach the final concentration 50  $\mu\text{M}$ . After the reaction mixture was incubated at room temperature for 24 h, it was acidified with HCl (2 M) to pH 3 followed by its incubation for 4 h at room temperature. The reaction mixture was monitored and characterized with HPLC-MS. The experiments were run in duplicate.

#### **Basification study to pH 9:**

Equivalent volume of diol **3-1k** (100  $\mu\text{l}$ , 100  $\mu\text{M}$  in 10 mM PBS buffer pH 7.4) and boronic acid **3-2c** (100  $\mu\text{l}$ , 100  $\mu\text{M}$  in 10 mM PBS buffer pH 7.4) were mixed in an eppendorf tube to reach the final concentration 50  $\mu\text{M}$ . After the reaction mixture was incubated at room temperature for 24 h, it was basified with NaOH (1 M) to pH 9 followed by incubation for 4 h at room temperature. The reaction mixture was monitored and characterized with HPLC-MS. The experiments were run in duplicate.

### 3.4.6.5 Mechanistic study – kinetic studies

$^1\text{H}$  NMR spectra were acquired on a 600 MHz four channel Agilent VNMRJ spectrometer, equipped with a z gradient HCN probe and using VNMRJ 4.2A as the acquisition software. Samples were dissolved/recorded in 10 mM PBS buffer/ $\text{D}_2\text{O}$  (9/1 v/v) at a concentration of 100  $\mu\text{M}$  and suppression of the  $\text{H}_2\text{O}$  signal was performed using excitation sculpting.<sup>57,58</sup> All data were referenced to the water, set to be 4.72 ppm.

The reaction between diol **3-1k** and **3-2c/3-2k** was monitored by  $^1\text{H}$  NMR at 25 °C in 10 mM PBS buffer/ $\text{D}_2\text{O}$  (9/1 v/v).  $\text{D}_2\text{O}$  (10%) was locked to monitor the reactions. The disappearance of  $-\text{CH}_3$  of **3-1k** at 0.98 ppm and the appearance of  $-\text{CH}_3$  of **3-3kc/3-3kk** at 0.86 ppm were monitored and converted into the concentration of diol **3-1k**. Diol **3-1k** and boronic acid **3-2c/3-2k** were separately dissolved in 10 mM PBS buffer/ $\text{D}_2\text{O}$  (9/1 v/v) at 0.2 mM. To 300  $\mu\text{l}$  of 0.2 mM diol **3-1k**, 300  $\mu\text{l}$  of 0.2 mM **3-2c/3-2k** solution was added, shaken quickly, and followed by monitoring the reaction at every 320 or 150 seconds over 84 or 45 min. The sum of integral of  $-\text{CH}_3$  of **3-3kc/3-3kk** and  $-\text{CH}_3$  of **3-1k** was accepted as 0.1 mM, thus the integral values of  $-\text{CH}_3$  of **3-1k** was converted into concentration (M). The second order rate constant was determined by plotting  $1/[\mathbf{3-1k}]$  vs time. The slope was recorded as the second order rate constant. Each reaction was performed at least three times. Their average of at least 3 measurements  $\pm$  standard deviation was recorded as  $4.2 \pm 0.4 \text{ M}^{-1}\text{s}^{-1}$  for **3-3kc** and  $8.9 \pm 0.6 \text{ M}^{-1}\text{s}^{-1}$  for **3-3kk**.

### 3.4.6.6 Biological competition assay

RP-HPLC-UV-MS was performed using an Agilent 1260 Infinity HPLC System with a Kinetex C8 reverse-phase analytical column (2.1x50 mm) (Phenomenex, Torrance, CA, USA), thermostated at 50°C followed by UV and mass spectrometric detection using a single quadrupole Agilent 6130 MSD. The buffer gradient system was composed of 0.1% formic acid in water as mobile phase A and 0.1% formic acid in ACN as mobile phase B.

For the separation of analytes, an aliquot was loaded onto the column at a flow rate of 0.50  $\text{mLmin}^{-1}$  and an initial buffer composition of 98% mobile phase A and 2% mobile phase B. Elution of the analytes was done by using a linear gradient from 2% to 95% mobile phase B over

a period of 5 minutes and held at 95% mobile phase B for 3 minutes to remove all analytes from the column.

UV absorbance was monitored at 254 nm. Mass spectra were acquired using positive/negative switching mode. Mass spectrometric conditions were drying gas 13 L/min at 350°C, nebulizer 30 psi, mass range 100-1500 Da, 1.14 sec/cycle, variable fragmentor voltage, capillary 4000V in positive and 3500V in negative mode. Data acquisition and analysis was performed using the OpenLab software package (Rev. C 01.07 [27]) and the Mass Hunter Analytical Studio Reviewer (Version B.02.01).

Boronate and thiosemicarbazone formations were determined *via* HPLC-MS using UV (260 nm) after 50 min, 2 h, 3 h and 4 h at room temperature. In order to perform the experiment, first stock solutions of D-glucose (16 mM), D-fructose (5 mM) and catechol (10 µM) were separately prepared in 10 mM PBS buffer (pH 7.4). Besides, stock solutions of diol **3-1k** (2.5 mM) and boronic acid **3-2c/3-2k** (2.5 mM) were prepared in a solution of 10 mM PBS buffer (pH 7.4)/ACN (50/50 v/v%).

Reaction mixture was prepared according to final concentrations of diol **3-1k** (100 µM) and **3-2c/3-2k** (100 µM), D-glucose (8 mM), D-fructose (0.3 mM) and catechol (0.01 µM). Each solution was prepared in duplicate.

In the presence of biocompetitors;

For diol **3-1k** (100 µM) and **3-2c/3-2k** (100 µM), the following mixture was prepared from stock solutions:

Diol **3-1k** (20 µl), **3-2c/3-2k** (20 µl), D-glucose (250 µl), D-fructose (30 µl), catechol (0.5 µl) were mixed in 10 mM PBS buffer (179.5 µl).

In the absence of biocompetitors;

For diol **3-1k** (100 µM) and **3-2c/3-2k** (100 µM), the following mixture was prepared from stock solutions:

Diol **3-1k** (20 µl), **3-2c/3-2k** (20 µl) were mixed 10 mM PBS buffer (460 µl).

### 3.4.6.7 Study to confirm the inertness and bioorthogonality of diol **3-1k** towards model aldehyde **3-2l**

RP-HPLC-UV-MS was performed using an Agilent 1260 Infinity HPLC System with a Kinetex C8 reverse-phase analytical column (2.1x50 mm) (Phenomenex, Torrance, CA, USA), thermostated at 50°C followed by UV and mass spectrometric detection using a single quadrupole Agilent 6130 MSD. The buffer gradient system was composed of 0.1% formic acid in water as mobile phase A and 0.1% formic acid in ACN as mobile phase B.

For the separation of analytes, an aliquot was loaded onto the column at a flow rate of 0.50 mLmin<sup>-1</sup> and an initial buffer composition of 98% mobile phase A and 2% mobile phase B. Elution of the analytes was done by using a linear gradient from 2% to 95% mobile phase B over a period of 5 minutes and held at 95% mobile phase B for 3 minutes to remove all analytes from the column.

UV absorbance was monitored at 214 nm. Mass spectra were acquired using positive/negative switching mode Mass spectrometric conditions were drying gas 13 L/min at 350°C, nebulizer 30 psi, mass range 100-1500 Da, 1.14 sec/cycle, variable fragmentor voltage, capillary 4000V in positive and 3500V in negative mode. Data acquisition and analysis was performed using the OpenLab software package (Rev. C 01.07 [27]) and the Mass Hunter Analytical Studio Reviewer (Version B.02.01).

Thiosemicarbazone/hydrazone formations between **3-1k/3-1l/3-1p** and model aldehyde **3-2l** were investigated *via* HPLC-MS analysis in 10 mM PBS buffer (pH 7.4) with 10 mM *p*-phenylenediamine as a catalyst. Model aldehyde **3-2l** was provided by Derda Research group, University of Alberta, Department of Chemistry. Individual stock solutions of **3-1k/3-1l** (200 μM), **3-1p** (500 μM) and **3-2l** (1 mM) were prepared in this solvent system. Equivalent volumes of **3-1k/3-1l/3-1p** (100 μl) and **3-2l** (100 μl) were mixed in a HPLC vial with a plastic insert to reach the final concentration 100 μM for **3-1k** and **3-1l**, 250 μM for **3-1p** and 500 μM for **3-2l**. After they were incubated for 2 h at room temperature, they were monitored *via* HPLC-MS (UV = 214 nm). The experiments were run in duplicate.

### 3.4.6.8 Cytotoxicity assay

HEK293T cells (grown in complete DMEM media with 10% fetal bovine serum (FBS), 2 mM glutamine) were grown to about 80% confluency before removal from the surface of a 20 cm petri-dish with 25% Trypsin with EDTA (5 min, 37 °C, 5% CO<sub>2</sub>). Cells were pelleted (9000 rpm, 5 min, 25 °C), diluted in 10 mL fresh complete DMEM media to generate aliquots of 30,000 cells per well (100 µL), then incubated for 6 h before addition of diol **3-1o** (50 and 25 µM), boronic acid **3-2c/3-2k** (50 and 25 µM), Triton X-100 (1 µl) and DMSO (1 µl). Stock solutions of diol and boronic acids were prepared fresh in DMSO (2.5 and 5 mM), and diluted to the working concentration as listed above in DMEM so that the total volume of DMSO added was kept constant at 1% (v/v). For each reagent, samples were set up as three replicates with one time point (18 h). Control wells were loaded with 1 µL DMSO or 1 µL of Triton X-100 (used as positive control for cytotoxicity). After incubation for 18 h, WST-1 reagent (10 µl) to each well was added and the wells were incubated for another 4 h. Absorbances at 450 nm and 600 nm were recorded with a plate reader. The results of the averages and standard deviations of the wells were determined after normalization to the DMSO control.

#### **3.4.6.9 Cell culture and labeling procedures**

HEK 293T cells were cultured in Gibco™ DMEM, High Glucose, GlutaMAX™ Supplement, Pyruvate (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) in T25 flask. Transient expression of SNAP-tagged protein (beta-2 adrenergic receptor) on the cell membrane with a plasmid pSNAP<sub>F</sub>-ADRβ<sub>2</sub> (purchased from New England BioLabs Inc.) was performed by following the standard protocol of Fugene 6 transfection reagent in T25 flask at 80% confluency. After 24 h, the washed and suspended cells were gently transferred into seven of 15 ml falcon tubes (each including 300 µl of the cell media). Then, the cells were centrifuged (900 rpm, 3.0 min) and the supernatant was gently removed. To five of these falcon tubes, boronic acid SNAP-tag reagent **3-2m** (500 µl, 10 µM freshly prepared solution in DMEM complete media from a 5.0 mM stock of **3-2m** in DMSO) and hoechst 33342 (3 µM in 500 µl of the same DMEM complete media, 16.0 mM stock in sterilized millipore H<sub>2</sub>O) was introduced and they were incubated at 37 °C for 45 min while the sixth tube was incubated with only hoechst 33342 in DMEM complete media (3 µM, 500 µl). After incubation, the cells were centrifuged (900 rpm, 3.0 min), the supernatant was gently removed and the cells were washed with DMEM complete media (1000 µl). This washing step was repeated two times. After the last

wash, the supernatant was removed, replaced with the following solutions (i-vi) and incubated at 37 °C for 0.5 h or 2.5 h.

- i) To the falcon tube previously incubated with only DMEM complete media and no **3-2m**, was added diol **3-1k** (500  $\mu$ l, 30  $\mu$ M solution in DMEM complete media from a 5.0 mM freshly prepared stock of **3-1k** in DMSO). Incubation time was 2.5 h.
- ii) To the one of the falcon tube previously incubated with **3-2m** was added **3-1l** (500  $\mu$ l, 30  $\mu$ M solution in DMEM complete media from a 5.0 mM freshly prepared stock of **3-1l** in DMSO). Incubation time was 2.5 h.
- iii) To the one of the falcon tube previously incubated with **3-2m** was added **3-1m** (500  $\mu$ l, 30  $\mu$ M solution in DMEM complete media from a 5.0 mM freshly prepared stock of **3-1m** in DMSO). Incubation time was 2.5 h.
- iv) To the one of the falcon tube previously incubated with **3-2m** was added **3-1k** (500  $\mu$ l, 30  $\mu$ M). Incubation time was 0.5 h.
- v) To the one of the falcon tube previously incubated with **3-2m** was added **3-1k** (500  $\mu$ l, 10  $\mu$ M). Incubation time was 2.5 h.
- vi) To the one of the falcon tube previously incubated with **3-2m** was added **3-1k** (500  $\mu$ l, 30  $\mu$ M). Incubation time was 2.5 h.

After 0.5 h or 2.5 h, the cells were centrifuged (900 rpm, 3.0 min), the supernatant was gently removed and the cells were washed with DMEM complete media (1000  $\mu$ l). This washing step was repeated two times. Finally, the cells in each falcon tube with an addition of 2.0 ml DMEM complete media were separately transferred into a 24 well glass bottom plate black plate with lid (Cellvis, # 1.5 High Performance Cover Glass (0.17 $\pm$ 0.005 mm)) and allowed to settle down at 37 °C for 14 h before fluorescence imaging. All TIRF fluorescence images were acquired in DMEM complete media.

Since the cells were exposed to extensive centrifuge and washing steps, they were mostly damaged, which made the imaging process more difficult. In order to confirm the corresponding damage is not due to the presence of the reagents, the same protocol was performed in the absence of the reagents. The cells' condition was observed similar to the ones performed in the presence of the reagents.

## Total internal reflection fluorescence microscope (TIRF) Imaging

NIKON Ti TIRF microscopy with 488 nm laser excitation and a  $525 \pm 50$  nm filter with a 60 x TIRF oil objective (NA 1.49) (imaging resolution is  $0.36 \mu\text{m pixel}^{-1}$ ) was used. Image analysis was performed in ImageJ (National Institute of Health) and Adobe Photoshop (CS5).

### 3.5 References

1. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
2. Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16–20.
3. Patterson, D. M.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2015**, *28*, 141–149.
4. Ramil, C. P.; Lin, Q. *Chem. Commun. (Camb.)* **2013**, *49*, 11007–11022.
5. Zheng, M.; Zheng, L.; Zhang, P.; Li, J.; Zhang, Y. *Molecules* **2015**, *20*, 3190–3205.
6. Fan, X.; Li, J.; Chen, P. R. *Natl. Sci. Rev.* **2017**, 1–3.
7. a) Yang, Y-Y.; Ascano, J. M.; Hang, H. C. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641. b) Heal, W. P.; Jovanovic, B.; Bessin, S.; Wright, M. H.; Magee, A. I.; Tate, E. W. *Chem. Commun.* **2011**, *47*, 4081–4083.
8. Selected examples: a) Axupa, J. Y.; Bajjuri, K. M.; Ritland, M.; Hutchins, B. M.; Kim, C. H.; Kazane, S. A.; Halder, R.; Forsyth, J. S.; Santidrian, A. F.; Stafin, K.; Lu, Y.; Tran, H.; Seller, A. J.; Biroc, S. L.; Szydluk, A.; Pinkstaff, J. K.; Tian, F.; Sinha, S. C.; Felding-Habermann, B.; Smider, V. V.; Schultz, P. G. *PNAS* **2012**, *109*, 16101–16106. b) Zimmerman, E. S.; Heibeck, T. H.; Gill, A.; Li, X.; Murray, C. J.; Madlansacay, M. R.; Tran, C.; Uter, N. T.; Yin, G.; Rivers, P. J.; Yam, A. Y.; Wang, W. D.; Steiner, A. R.; Bajad, S. U.; Penta, K.; Yang, W.; Hallam, T. J.; Thanos, C. D.; Sato, A. K. *Bioconjugate Chem.* **2014**, *25*, 351–361.
9. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
10. Baskin, J. M.; Bertozzi, C. R. *QSAR Comb. Sci.* **2007**, *26*, 1211–1219.
11. Best, M. D. *Biochemistry* **2009**, *48*, 6571–6584.
12. McKay, C. S.; Finn, M. G. *Chem. Biol.* **2014**, *21*, 1075–1101.
13. Shih, H.-W.; Kamber, D. N.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2014**, *21*, 103–111.
14. Akgun, B.; Hall, D. G. *Angew. Chem. Int. Ed.* **2016**, *55*, 3909–3913.
15. Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764–4806.
16. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125–1128.

17. Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.
18. Bandyopadhyay, A.; Gao, J. *Chem. Eur. J.* **2015**, *21*, 14748–14752.
19. Schmidt, P.; Stress, C.; Gillingham, D. *Chem. Sci.* **2015**, *6*, 3329–3333.
20. Gillingham, D. *Org. Biomol. Chem.* **2016**, *14*, 7606–7609.
21. Dilek, O.; Lei, Z.; Mukherjee, K.; Bane, S. *Chem. Commun.* **2015**, *51*, 16992–16995.
22. Stress, C. J.; Schmidt, P. J.; Gillingham, D. G. *Org. Biomol. Chem.* **2016**, *14*, 5529–5533.
23. Bandyopadhyay, A.; Cambray, S.; Gao, J. *Chem. Sci.* **2016**, *7*, 4589–4593.
24. Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P. *Chem. Sci.* **2016**, *7*, 5052–5058.
25. Bandyopadhyay, A.; Cambray, S.; Gao, J. *J. Am. Chem. Soc.* **2017**, *139*, 871–878.
26. a) Chapin, B. M.; Metola, P.; Lynch, V. M.; Stanton, J. F.; James, T. D.; Anslyn, E. V. *J. Org. Chem.* **2016**, *81*, 8319–8330. b) Meadows, M. K.; Roesner, E. K.; Lynch, V. M.; James, T. D.; Anslyn, E. V. *Org. Lett.* **2017**, *19*, 3179–3182.
27. Kwong, F. Y.; Klapars, A.; Buchwald, S. L. *Org. Lett.* **2002**, *4*, 581–584.
28. a) Kalia, J.; Raines, R. T. *Angew. Chem. Int. Ed.* **2008**, *47*, 7523–7526. b) Sander, E. G.; Jencks, W. P. *J. Am. Chem. Soc.* **1968**, *90*, 6154–6162.
29. Il'ina, I. V.; Volcho, K.P.; Salakhutdinov, N. F. *Russ. J. Organ. Chem.* **2008**, *44*, 1–23.
30. Chen, Y.; Tsao, K.; Keillor, J. W. *Can. J. Chem.* **2015**, *93*, 389–398.
31. Shieh, P.; Bertozzi, C. R. *Org. Biomol. Chem.* **2014**, *12*, 9307–9320.
32. Nadler, A.; Schultz, C. *Angew. Chem. Int. Ed.* **2013**, *52*, 2408–2410.
33. a) Jun, M. E.; Roy, B.; Ahn, K. H. *Chem. Commun.* **2011**, *47*, 7583–7601. b) Sun, X.; Zhai, W.; Fossey, J. S.; James, T. D. *Chem. Commun.* **2016**, *52*, 3456–3469. c) Wu, X.; Li, Z.; Chen, X-X.; Fossey, J. S.; James, T. D.; Jiang, Y-B. *Chem. Soc. Rev.* **2013**, *42*, 8032–8048. d) Lacina, K.; Skládal, P.; James, T. D. *Chem. Cent. J.* **2014**, *8*, 1–17.
34. a) Sun, X. L.; Xu, S. Y.; Flower, S. E.; Fossey, J. S.; Qian, X. H.; James, T. D. *Chem. Commun.* **2013**, *49*, 8311–8313. b) Cao, H.; Diaz, D. I.; DiCesare, N.; Lakowicz, J. R.; Heagy, M. D. *Org. Lett.* **2002**, *4*, 1503–1505. c) Cao, Z.; Nandhikonda, P.; Heagy, M. D. *J. Org. Chem.* **2009**, *74*, 3544–3546.

35. a) Dilek, O.; Bane, S. L. *Tetrahedron Lett.* **2008**, *49*, 1413–1416. b) Dilek, O.; Bane, S. L. *J. Fluoresc.* **2011**, *21*, 347–354. c) Key, J. A.; Li, C.; Cairo, C. W. *Bioconjugate Chem.* **2012**, *23*, 363–371.
36. Key, J. A.; Koh, S.; Timerghazin, Q. K.; Brown, A.; Cairo, C. W. *Dyes Pigm.* **2009**, *82*, 192–203.
37. Gubitz, G.; Wintersteiger, R.; Frei, R. W. *J. Liq. Chromatogr.* **1984**, *7*, 839–854.
38. Uzu, S.; Kanda, S.; Nakashima, K.; Akiyama, S. *Analyst* **1990**, *115*, 1477–1482.
39. a) Crandall, J. K.; Crawley, L. C. *Org. Synth.* **1973**, *53*, 17–5. b) Lakshmi, R.; Bateman, T. D.; McIntosh, M. C. *J. Org. Chem.* **2005**, *70*, 5313–5315.
40. Caron, M.; Sharpless, K. B. *J. Org. Chem.* **1985**, *50*, 1557–1560.
41. Bosch, L. I.; Fyles, T. M.; James, T. D. *Tetrahedron* **2004**, *60*, 11175–11190.
42. a) Givens, R. S.; Stensrud, K.; Conrad, P. G.; Yousef, A. L.; Perera, C.; Senadheera, S. N.; Heger, D.; Wirz, J. *Can. J. Chem.* **2011**, *89*, 364–384. b) Elsinghorst, P. W.; Tanarro, C. M. G.; Gütschow, M. *J. Med. Chem.* **2006**, *49*, 7540–7544.
43. Cal, P. M. S. D.; Frade, R. F. M.; Chudasama, V.; Cordeiro, C.; Caddick, S.; Gois, P. M. P. *Chem. Commun. (Camb.)* **2014**, *50*, 5261–5263.
44. a) James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Nature* **1995**, *374*, 345–347. b) Sun, X.; James, T. D. *Chem. Rev.* **2015**, *115*, 8001–8037.
45. Wu, X.; Li, Z.; Chen, X.-X.; Fossey, J. S.; James, T. D.; Jiang, Y.-B. *Chem. Soc. Rev.* **2013**, *42*, 8032–8048.
46. a) Champlain, J. de; Farley, L.; Cousineau, D.; Ameringen, M. R. van. *Circ. Res.* **1976**, *38*, 109–114. b) Kawasaki, T.; Akanuma, H.; Yamanouchi, T. *Diabetes Care* **2002**, *25*, 353–357.
47. Matthews, M. L.; He, L.; Horning, B. D.; Olson, E. J.; Correia, B. E.; Yates, J. R.; Dawson, P. E.; Cravatt, B. F. *Nat. Chem.* **2017**, *9*, 234–243.
48. Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21.
49. Kitov, P. I.; Vinals, D. F.; Ng, S.; Tjhung, K. F.; Derda, R. *J. Am. Chem. Soc.* **2014**, *136*, 8149–8152.
50. Rashidian, M.; Mahmoodi, M. M.; Shah, R.; Dozier, J. K.; Wagner, C. R.; Distefano, M. D. *Bioconjugate Chem.* **2013**, *24*, 333–342.

51. Borner, M. M.; Schneider, E.; Pirnia, F.; Sartor, O.; Trepel, J. B.; Myers, C. E. *FEBS Lett.* **1994**, *353*, 129–132.
52. Ngamwongsatit, P.; Banada, P. P.; Panbangred, W.; Bhunia, A. K. *J. Microbiol. Methods.* **2008**, *73*, 211–215.
53. Keppler, A.; Kindermann, M.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. *Methods* **2004**, *32*, 437–444.
54. Cole, N. B. *Curr. Protoc. Protein Sci.* **2013**, *73*, 30.1.1–30.1.16.
55. Mattheyses, A. L.; Simon, S. M.; Rappoport, J. Z. *J. Cell Sci.* **2010**, *123*, 3621–3628.
56. Hu, B.; Xing, S.; Ren, J.; Wang, Z. *Tetrahedron* **2010**, *66*, 5671–5674.
57. Hwang, T. L.; Shaka, A. J. *J. Magn. Reson.* **1995**, *A112*, 275–279.
58. Dalvit, C. *J. Biol. NMR*, **1998**, *11*, 437–444.

## 4. Chapter 4. Investigation of a Reactive Peptide Tag Towards 2-Acetylarylboronic Acid *via* Phage-Display Screening

### 4.1 Introduction

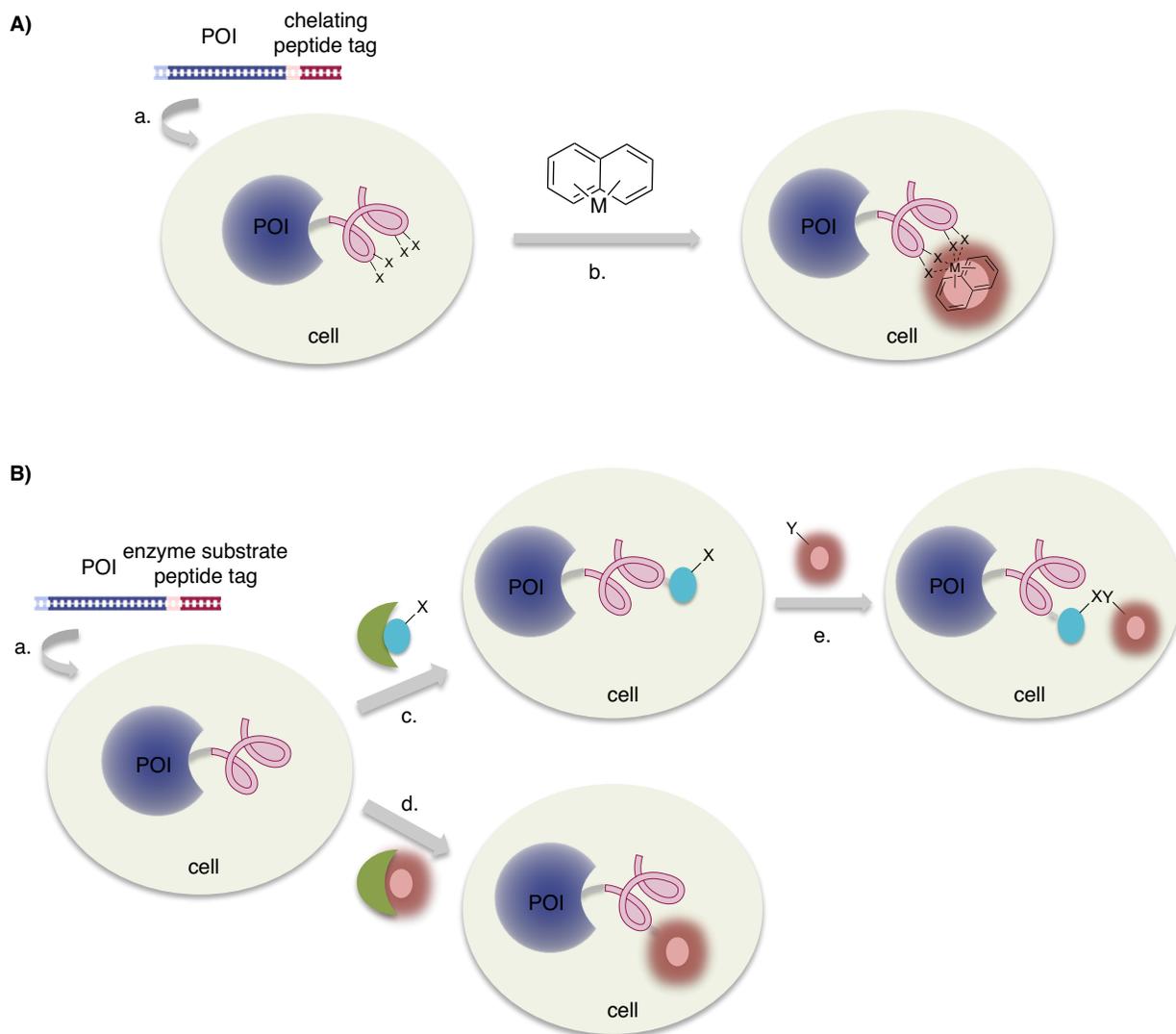
Bioorthogonal reactions often exploit a non-natural chemical probe, which is incorporated into a protein of interest (POI) *via* chemically or engineered biosynthetic pathways, such as amber stop codon suppression.<sup>1,2</sup> The installation of the chemical reporter into the POI is often challenging, and incorporation efficiency depends highly on the type of chemical reporters. To achieve an easier incorporation process into the POI and a more efficient bioorthogonal reaction, sequence-specific bioorthogonal reactions containing a natural peptide tag with highly specific chemical reactivity towards a non-natural probe would present a powerful avenue. Early examples of such bioorthogonal reactions involve short peptide tags (< 20 natural amino acids) that do not interfere with the structure and integrity of the POI being investigated.<sup>3</sup> For example, a tetracysteine motif (CCPGCC) can react selectively with fluorescein arsenical hairpin binder (FLAsH) or resorufin arsenical hairpin binder (ReAsH), and the resulting bioconjugate becomes fluorescent (Figure 4-1A).<sup>4,5</sup> Another tetracysteine-containing peptide, which is orthogonal to the one that binds FLAsH and ReAsH, was found to be reactive towards a biarsenical cyanine probe (Cy3As).<sup>6,7</sup> In addition to the tetracysteine motif, a tetraserine-containing peptide (SSPGSS) was designed to react selectively with rhodamine-derived bisboronic acid (RhoBo), and the resulting bisboronate conjugate also becomes fluorescent.<sup>8</sup> Other peptide tags, such as hexahistidine or tetraaspartate that can chelate with nickel- or zinc-containing compounds, have been reported.<sup>9-12</sup> Although all these peptide tags are advantageous due to their small size, fast reactivity and fluorogenicity, arsenic and nickel can be cytotoxic. Furthermore, tetracysteine motifs often need the addition of chelators to prevent high levels of background labeling, therefore, making the system more complex.

An alternative route to the chelating peptide tags is to employ enzyme substrate peptide tags (Figure 4-1B). For example, the biotin ligase (BirA) peptide tag (15 amino acids long), a substrate for BirA, can be genetically introduced into a POI.<sup>13</sup> After the peptide-POI fusion is expressed, BirA was employed to attach a biotin molecule site-selectively at the peptide tag site, thus allowing the modified POI to be captured by a quantum dot-modified streptavidin.<sup>13</sup> This

process is selective and compatible with mammalian cells because the BirA is acquired from a bacterial species, and also, the system is free of cytotoxic metals or reagents. Later, several alternatives to the BirA system, such as enzyme-linked tags that can introduce bioorthogonal handles (ketones, aldehydes and azides) into the peptide-POI, have been developed for further bioorthogonal modification (Figure 4-1B, pathway c).<sup>14-16</sup> Also, direct visualization of the peptide-POI was shown to be possible with fluorophores (Figure 4-1B, pathway d).<sup>17-19</sup> All these strategies need the genetic fusion of a peptide to a POI, either endogenous co-expression of a peptide modifying enzyme or exogenous addition of the enzyme and addition of the small molecule for labeling. Even though the direct modification of the peptide-POI fusion is attainable, bioorthogonal conjugation must be performed after the enzymatic process for some systems (Figure 4-1B, pathway e). This lengthy step can be time consuming and lead to experimental errors in the system. Furthermore, enzymatic strategies often require high substrate concentrations (which may reach 1.0 mM in some cases) and long incubation times, and the enzyme stability may also be problematic.

An ideal peptide tag for site-selective labeling of a POI must show a high tag specificity and reactivity to its non-cytotoxic bioorthogonal reaction pair without the need for an enzyme. The natural peptide tag must be small in size, and the bioorthogonal reaction must be fast and biologically compatible. Moreover, the system should provide low background and stable bond formation and allow the flexibility to choose diverse labeling groups. Such systems with short peptides and non-cytotoxic reagents, which exhibit enzyme-free, selective bioconjugation reactions, are scarce. For instance, Barbas and co-workers demonstrated that a reactive amino group placed in a unique 21-mer peptide could form an enaminone with 1,3-diketone derivatives (Figure 4-2A).<sup>20</sup> This system was applied for the labeling of maltose binding protein. In 2011, Weiss and co-workers found that an amino group of *N*-terminal histidine of a specific 11-mer peptide can react with biotin hydrazide and rhodamine B hydrazide at neutral pH, affording a loss of hydrazine and a ketone conjugate (Figure 4-2B).<sup>21</sup> Apart from reactive amino groups, cysteine containing peptide sequences have been reported. In this regard, Pentelute and co-workers recently reported a four-amino-acid  $\pi$ -clamp unit (Phe-Cys-Pro-Phe) that can be functionalized with a perfluoroaryl compound even in the presence of other cysteine residues on the same POI (Figure 4-2C).<sup>22</sup> They challenged this strategy to modify proteins involving several endogenous cysteine residues, such as antibodies and cysteine-based enzymes. Recently, Lin and

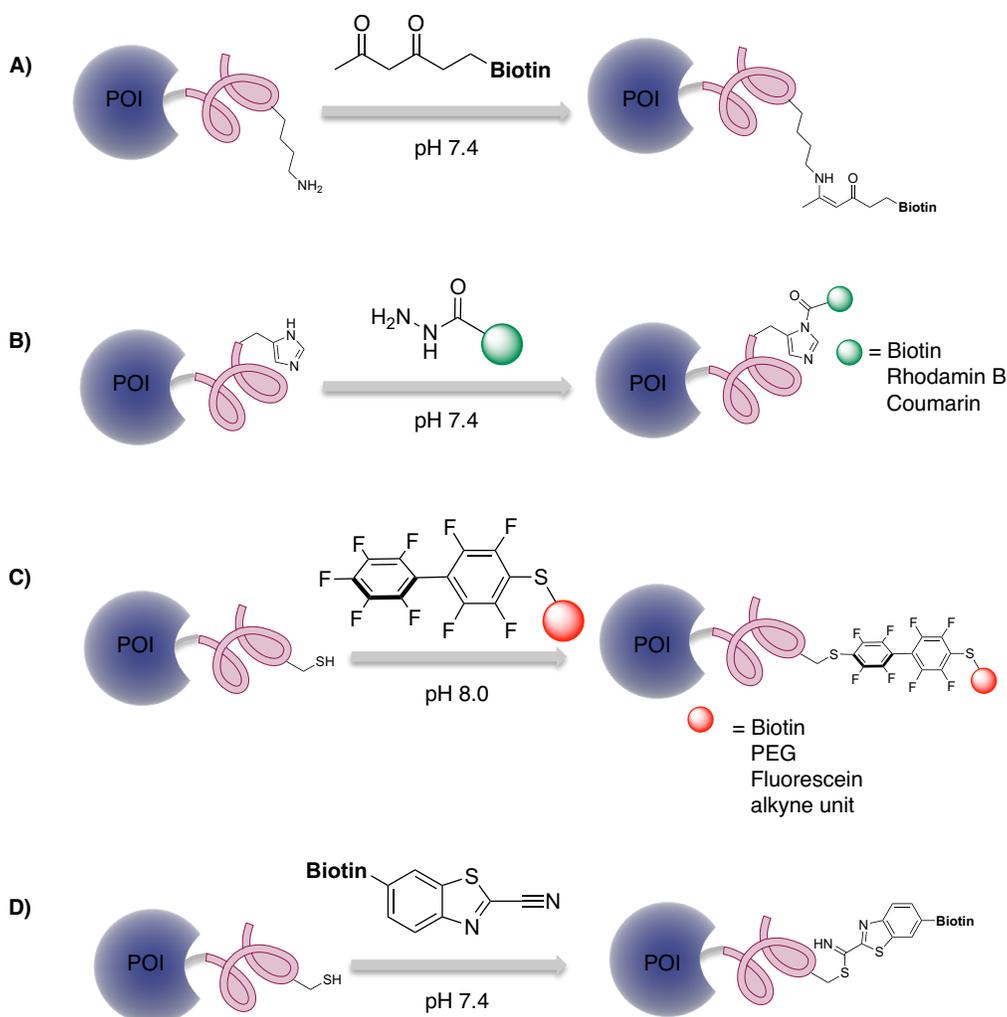
co-workers also demonstrated a cysteine containing peptide sequence (11 amino acid long), which can ligate with 2-cyanobenzothiazole derivatives (Figure 4-2D).<sup>23</sup> The system was employed in the surface protein labeling of *E. Coli* cells. Despite these great advances, all these systems often require a high concentration of labeling reagents (~500  $\mu$ M) due to their moderate reactivity. Therefore, what would be more appealing is to develop such a bioorthogonal system consisting of a short, reactive peptide and a non-cytotoxic reagent, which exhibits high reactivity and requires only a submicromolar concentration of labeling reagents.



**Figure 4-1:** A) Labeling proteins with chelating peptide tags. B) Labeling proteins with enzyme substrate peptide tags. Common methods to site-selectively label proteins in live cells. a) Transfer of recombinant DNA of the POI-tag fusion into the cell. b) Addition of a metal- or

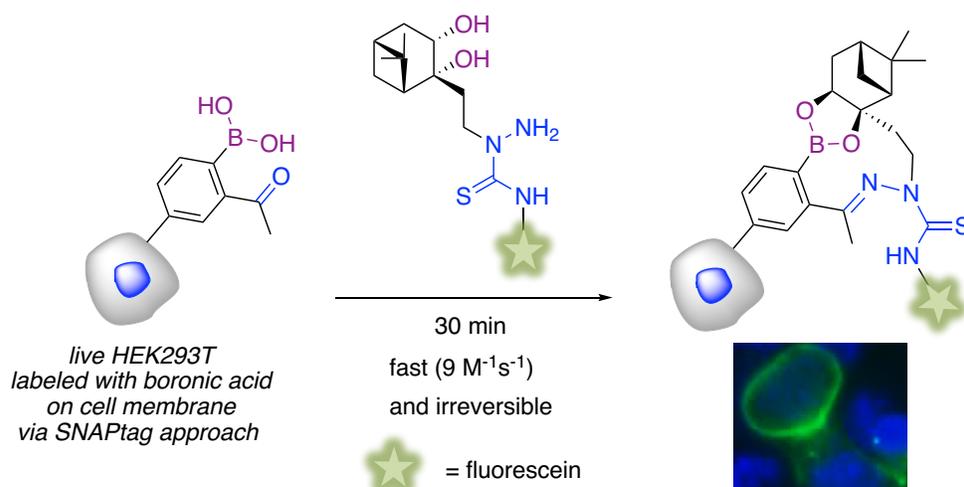
metalloid-containing compound into the cells and the resulting emission of a signal. c) Installation of a chemical handle into the tag on the POI *via* an enzyme. d) Direct installation of a fluorophore into the tag on the POI *via* an enzyme. e) Bioorthogonal labeling of the chemical handle with its bioorthogonal reagent pair.

In this regard, our synergic thiosemicarbazone/boronate system,<sup>24</sup> which affords fast and irreversible bioconjugation between 2-acetylarylboronic acid and thiosemicarbazide-functionalized nopoldiol (Figure 4-3A), inspired us towards the investigation of a short and highly reactive naturally-occurring peptide tag towards 2-acetylarylboronic acid, thus replacing the thiosemicarbazide-functionalized nopoldiol with a short peptide sequence. It was hoped that this reactive peptide, which must be ideally rare in *in vivo*, would preserve its fast and high affinity towards 2-acetylarylboronic acid *via* both imine and boronate formation. Anslyn and James also described a stable three-component assembly with 2-formylbenzeneboronic acid (2-FBBA), catechol and benzylamine (Figure 4-3B).<sup>25</sup> The authors confirmed that initial binding of either benzylamine or catechol to 2-FBBA improves the complexation of catechol and benzylamine, respectively. Binding of these two molecules to 2-FBBA becomes cooperative in a protic solvent. Thus, it was hypothesized that 2-acetylarylboronic acid might lead to a tight complex with a short peptide embedding as many as three hydroxyl side chains (serine, threonine) and a lysine residue or a terminal amine (Figure 4-3C). It is important to mention that this hypothetical system is more challenging compared to the reported systems<sup>24,25</sup> because both the expected boronate and imine conjugations are intrinsically less stable than the nopolboronate and thiosemicarbazone ligations. Furthermore, 2-acetylarylboronic acid is less reactive than 2-formylarylboronic acid. However, it was considered that the aromatic structure of boronic acid might also interact with aromatic amino acid side chains to enhance the desired bioconjugate formation further. The ideal rearrangement of these amino acids in a short flexible peptide could not be designed; therefore, to discover the reactive peptides, a phage display selection platform was exploited. This effort was undertaken in collaboration with the laboratory of Prof. Ratmir Derda, who is an expert on phage display selections.<sup>26,27</sup> The phage display platform with an serine-terminated library of over 10<sup>8</sup> heptapeptides was used in order to test this hypothesis. It was considered that serine-terminated peptides might be an additional asset because of a possible boronate formation with *N*-terminal serine and the boronic acid unit along with an imine conjugation between the acetyl group and the closest amine group within the peptide sequence.

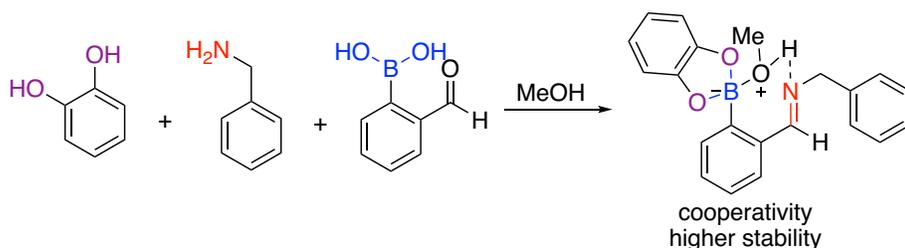


**Figure 4-2:** Reported short peptide tags that undergo enzyme-free, highly selective bioconjugation reactions with biocompatible chemical reagents. A) Labeling lysine amino group embedded in a unique 21-mer peptide onto the POI with 1,3-diketone derivatives at neutral pH. B) Labeling *N*-terminal histidine of a specific 11-mer peptide onto the POI with hydrazide derivatives. C) Labeling cysteine residue of a unique 4-mer peptide onto the POI with perfluoroaromatic reagents. D) Labeling cysteine residue of a specific 11-mer peptide onto the POI with 2-cyanobenzothiazole derivatives.

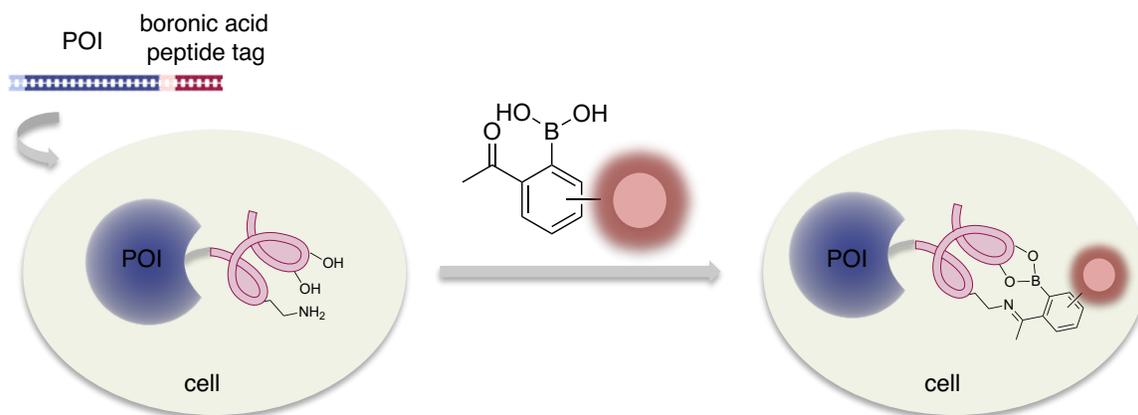
### A) Synergic design<sup>24</sup>



### B) Three-component assembly<sup>25</sup>



### C) This work: investigation of boronic acid peptide tag



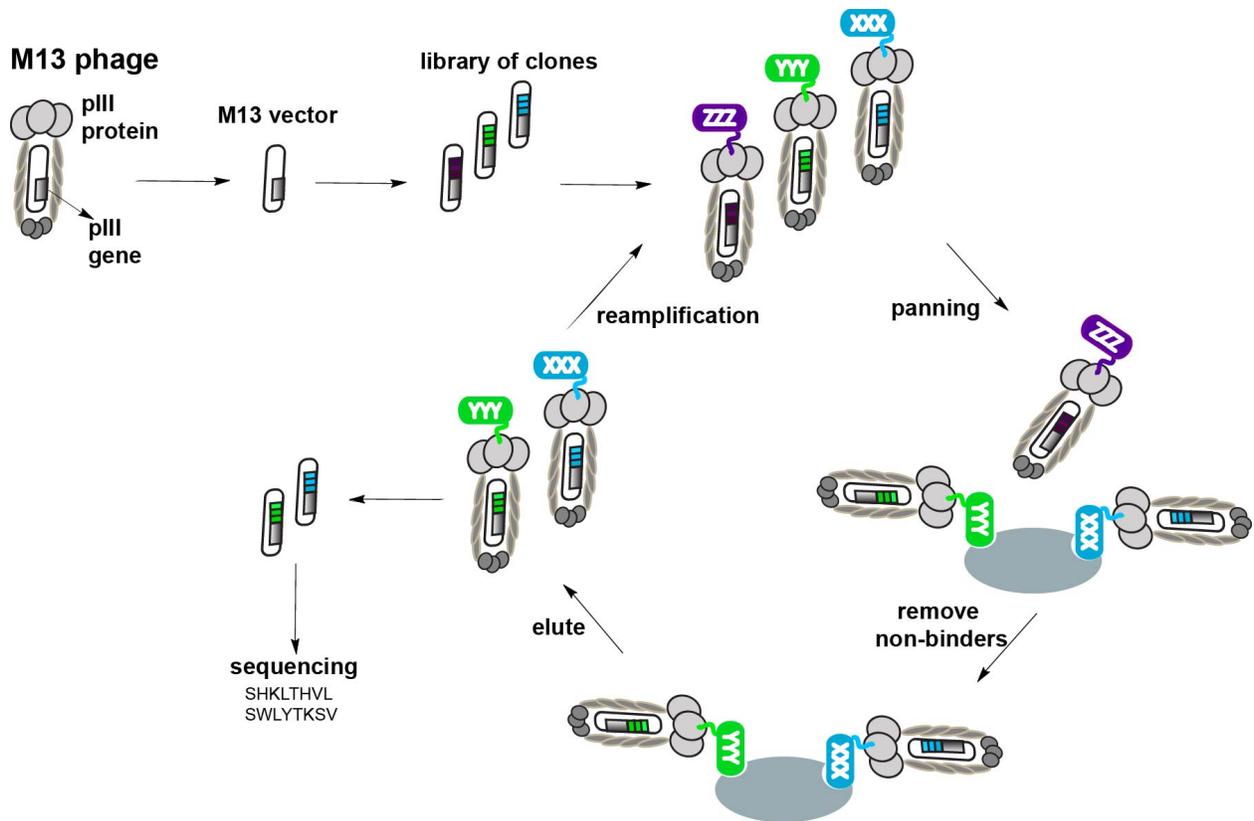
**Figure 4-3:** A) Synergic, fast and irreversible thiosemicarbazone/boronate bioorthogonal system. B) Positive cooperativity and higher stability of a three-component amine/boronic acid/catechol system. C) Labeling proteins with a short and reactive peptide tag containing the right geometry of alcohol and amino groups that favor a stable ligation with 2-acetylarylboronic acid through imine and boronate formation.

## 4.2 Results and discussion

### 4.2.1 Overview and design of phage display protocol

Phage display<sup>28</sup> as a high-throughput screening method exploits bacteriophages mostly to investigate protein-protein, protein-peptide and protein-DNA interactions. Also, this platform has been used for peptide substrate profiling, which afforded a fast and sequence-specific bioorthogonal Sonagashira cross-coupling reaction.<sup>29</sup> It was considered that this technology could be applied to find an optimum peptide-boronic acid bioorthogonal conjugation by screening libraries of peptides displayed on phage. Each bacteriophage displays a unique peptide sequence of gene encoding a peptide of interest, which can be traced thanks to the genetic code inside the phage capsid.<sup>28</sup> A library of bacteriophages (accessing high diversities of  $10^8$  different clones) can be prepared by inserting a library of DNA sequences into the bacteriophages' genome. As a common screening strategy, phage libraries are subjected to a surface or bead coated with a target molecule to pull-down the functional peptides that exhibit high affinity toward the target (Figure 4-4, panning).<sup>28</sup> Next, the phages that show no binding are removed by washing. The bound phages are finally eluted and re-amplified in bacteria to perform a second round of panning. Typically, selection protocols use 2–3 rounds to obtain binders with high specificity and affinity.<sup>28</sup> At each round, the ssDNA of eluted phages can be extracted and sequenced to determine the corresponding peptide sequences (Figure 4-4).

In this work, a custom-made M13 phage library (Ser-X<sub>7</sub>), which contains  $3 \times 10^8$  different phage virions displaying a unique, 8-amino acid peptide on the pIII protein, with serine at its *N*-terminus, was used. M13 phage contains a single-stranded DNA molecule wrapped inside a tube-shape structure composed of thousands of copies of a pVIII protein, the major coat protein (Figure 4-4).<sup>28</sup> The minor coat proteins cover the edges of the tube, and the interaction of one of the minor coat proteins (pIII) with a receptor of bacterium *E. Coli* results in the infection and reproduction of virions.



**Figure 4-4:** A common screening protocol using a phage display library with M13 phage.

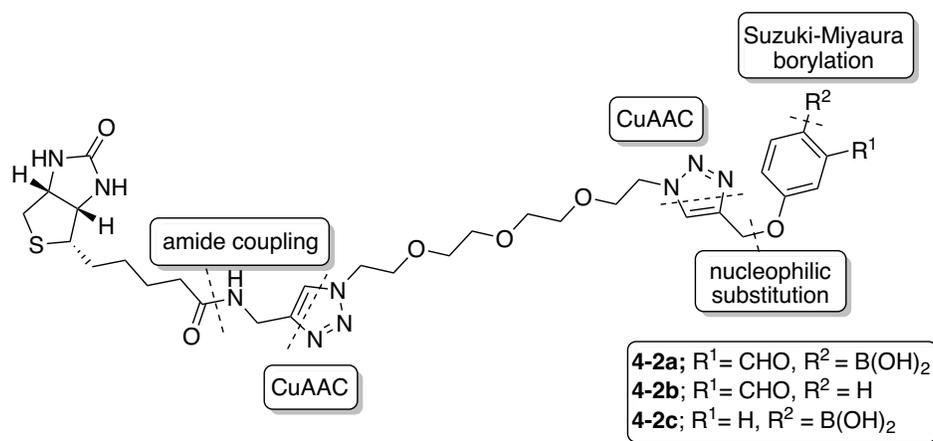


contain both boronate and imine ligation, are shown in Figure 4-5. Then, the reaction mixtures are purified with dialysis, which removes any remaining reagents **4-1a/4-1b/4-1c** in the reaction mixtures that might interfere with the pull-down assays (capturing biotinylated phages). The biotinylated phages can be captured with high capacity agarose streptavidin beads, and the non-biotinylated phages are removed from the beads with extensive washing. It is necessary to mention that sequences with affinity towards other components of the pull-down assay (species that bind non-specifically to streptavidin, agarose, BSA, plastic, etc.) are also selected, but they should be present in the control populations and thus, eliminated after the differential enrichment analysis. Later, the DNA of the captured phages, which have high affinity towards boronic acid **4-1a**, can be extracted with basic buffer solution and hexane, subjected to PCR and sequenced to analyze the peptide hits.

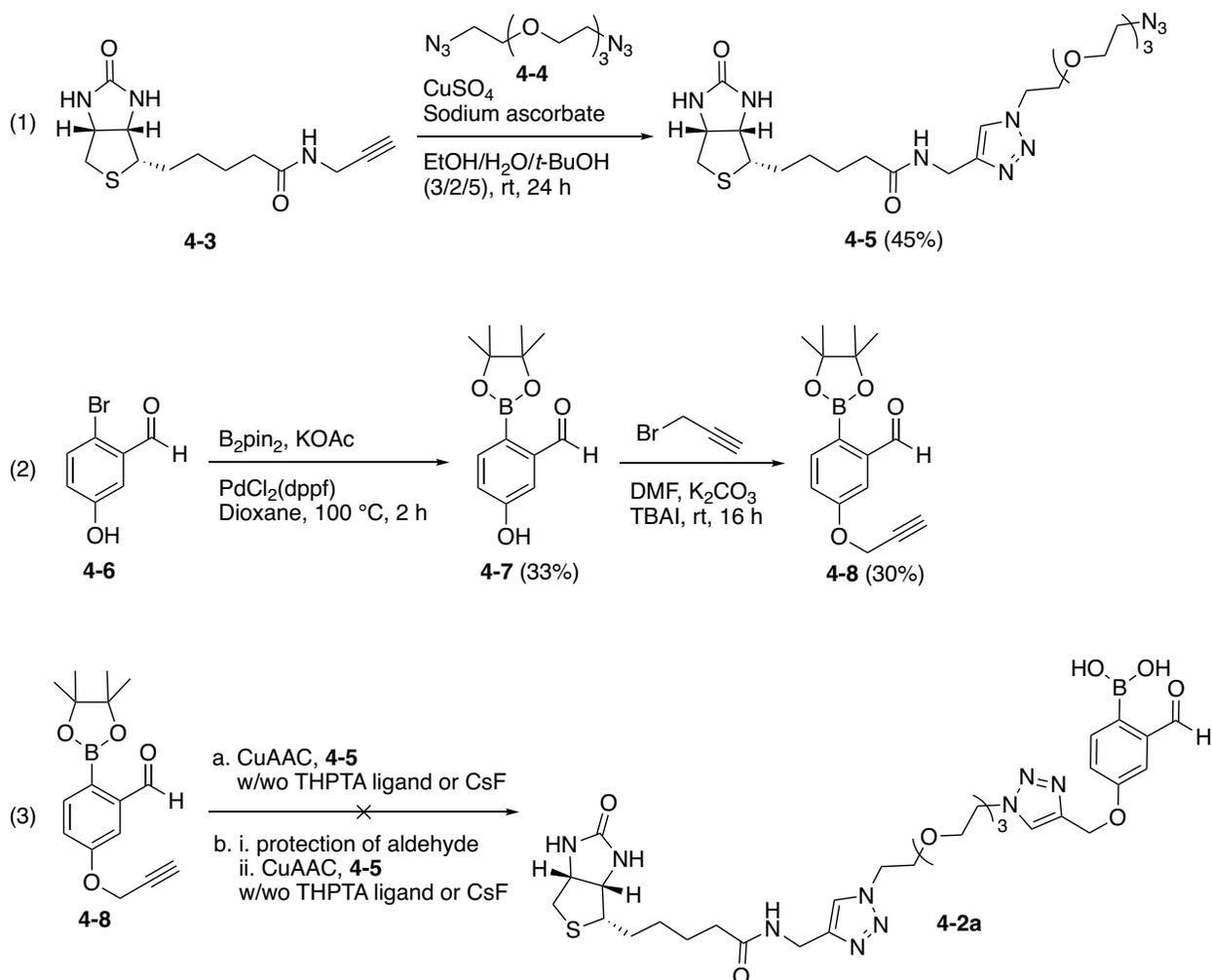
#### 4.2.2 Design and synthesis of biotinylated reagents used in the selection

Before discussing the selection protocol and results, the design and synthesis of biotinylated reagents will be mentioned. This project was initiated in January 2016, and at that time, the reactivity of 2-FBBA with *N*-terminal cysteine had not been reported. Therefore, the initial goal was to search for highly reactive peptides towards 2-formylarylboronic acid. A biotin derivatized 2-formylarylboronic acid reagent **4-2a** and its controls **4-2b** and **4-2c** were designed (Scheme 4-1). A triazole ring and PEG groups were chosen as linkers to conjugate aryl units and biotin because these groups can increase the water solubility of the final products. Biotin propargyl **4-3**, which was obtained easily from an amide coupling of biotin and propargyl amine,<sup>30</sup> was subjected to CuAAC conditions in the presence of PEG-diazide reagent **4-4**<sup>31</sup> to synthesize biotin-PEG-azide **4-5** (Scheme 4-2, equation 1). Propargyl derivatized boronate reagent **4-8** was synthesized from a Suzuki-Miyaura borylation<sup>32</sup> of 2-formyl-4-hydroxybromobenzene **4-6**, followed by a S<sub>N</sub>2 etherification<sup>33</sup> of boronate intermediate **4-7** and propargyl bromide (Scheme 4-2, equation 2). Later, CuAAC of biotin-PEG-azide **4-5** and propargyl boronate **4-8** was attempted under several conditions, such as with/without ligand tris(3-hydroxypropyltriazolymethyl)amine (THPTA) and with/without CsF (Scheme 4-2, equation 3). The ligand THPTA can accelerate the rate of triazole ring formation and have a role as a reductant, thus it was hoped that the ligand would prevent the copper(I)-mediated decomposition of the boronyl unit.<sup>34</sup> In addition, the additive CsF was found to prevent this

decomposition under CuAAC conditions by hampering copper insertion into the B–C bond because of the anionic tetrahedral formation of boron with fluorides.<sup>35</sup> However, all these attempts were unsuccessful. Also, the same conditions were performed after an acetal protection of the aldehyde unit of propargyl boronate **4-8**, but these trials were also unfruitful.



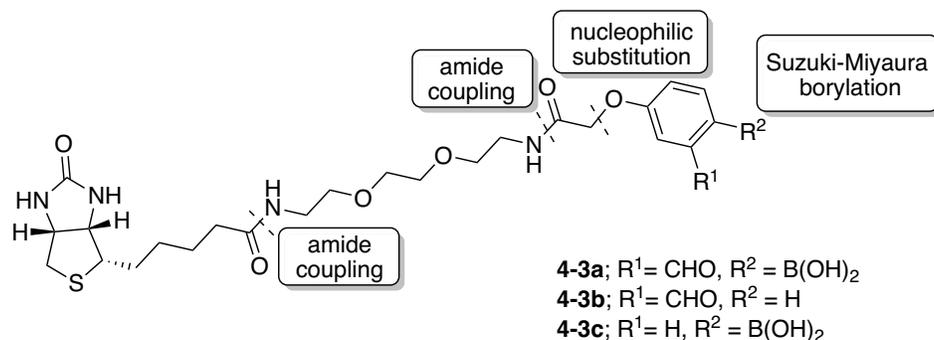
**Scheme 4-1:** Initial design of biotin derivatized 2-formylarylboronic acid reagent **4-2a** and control reagents **4-2b** and **4-2c**.



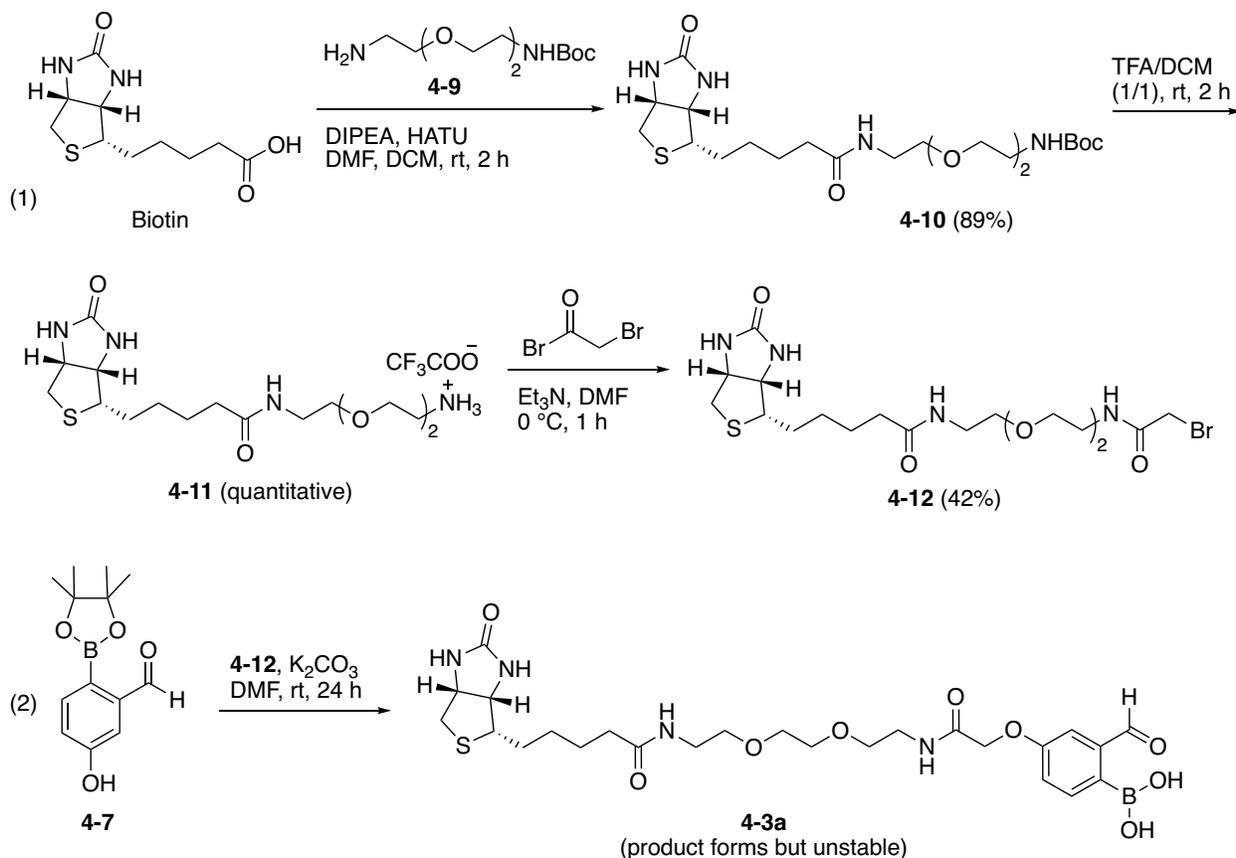
**Scheme 4-2:** Attempts to synthesize biotin derivatized 2-formylarylboronic acid reagent **4-2a**.

Next, another design, where CuAAC was replaced with a milder  $S_N2$  reaction, was considered to synthesize biotin derivatized 2-formylarylboronic acid reagent **4-3a** and control reagents **4-3b** and **4-3c** (Scheme 4-3). A more efficient route was selected to synthesize biotin bromide reagent **4-12**. First, an amide coupling of biotin and PEG-NHBoc amine **4-9** provided biotin-PEG-NHBoc compound **4-10**,<sup>36</sup> which was deprotected and finally acylated with  $\alpha$ -bromo acetyl bromide to synthesize biotin bromide **4-12** (Scheme 4-4, equation 1). Then, boronate **4-7** was reacted with biotin bromide **4-12** in the presence of anhydrous  $K_2CO_3$  to obtain the desired biotin derivatized 2-formylarylboronic acid reagent **4-3a** (Scheme 4-4, equation 2). HPLC-MS analysis of the crude reaction mixture confirmed the formation of the desired product; however, it was very unstable, and several purification attempts with C18 column led to product

decomposition. These unsuccessful attempts, together with the side-reactivity of 2-FBBA with an *N*-terminal cysteine, were factors in deciding to pursue with 2-acetylarylboronic acid and investigate its selective reactivity towards short peptide sequences.

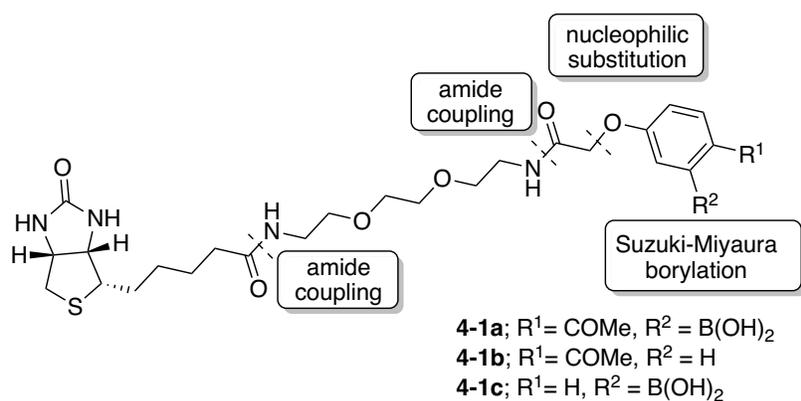


**Scheme 4-3:** Second design of biotin derivatized 2-formylarylboronic acid reagent **4-3a** and control reagents **4-3b** and **4-3c**.

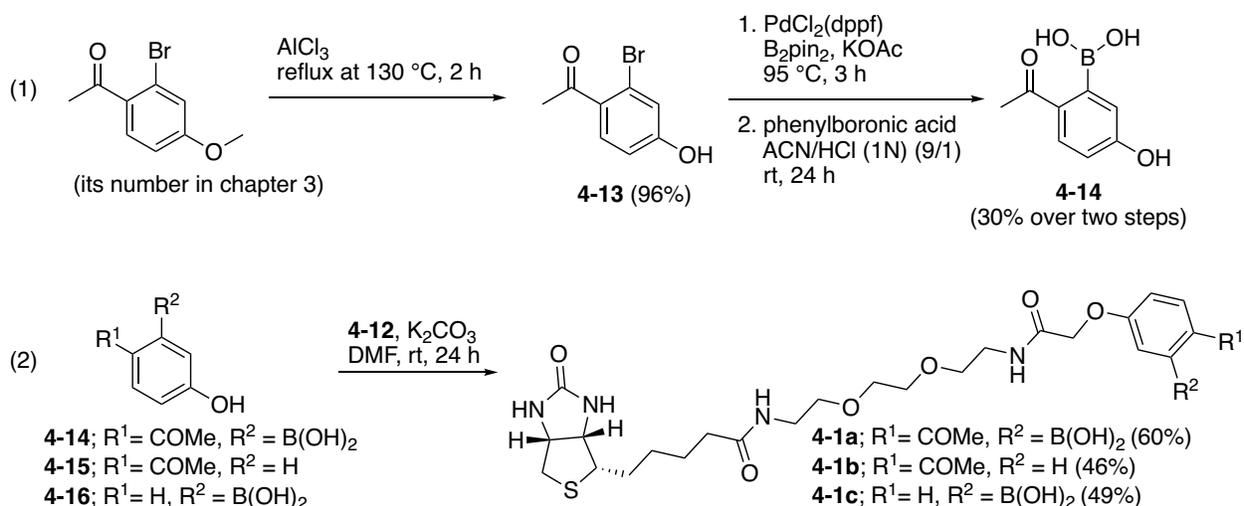


**Scheme 4-4:** Another attempt to synthesize biotin derivatized 2-formylarylboronic acid reagent **4-3a**.

Biotin derivatized 2-acetylarylboronic acid reagent **4-1a** and its control reagents **4-1b** and **4-1c** were designed and synthesized, as shown in Scheme 4-5 and 4-6. First, 2-acetyl-4-hydroxybromo benzene **4-13** was prepared by the demethylation of methoxy on 2-acetyl-4-methoxybromo benzene **4-12**. Then, boronic acid the phenolic **4-14** was obtained by the Suzuki-Miyaura borylation of the reagent **4-13** and removal of the pinacol unit through trans esterification in the presence of excess phenylboronic acid (Scheme 4-5, equation 1). Next, the desired reagent **4-1a** was synthesized easily with a  $S_N2$  reaction between the phenolic boronic acid **4-14** and biotin bromide **4-12** (Scheme 4-5, equation 2). In a similar fashion, control reagents **4-1b** and **4-1c** were prepared from the commercially available precursors 4-hydroxyacetophenone **4-15** and 3-hydroxyphenylboronic acid **4-16**, respectively.



**Scheme 4-5:** Design of biotin derivatized 2-acetylarylboronic acid reagent **4-1a** and control reagents **4-1b** and **4-1c**.



**Scheme 4-6:** Synthesis of biotin derivatized 2-acetylarylboronic acid reagent **4-1a** and control reagents **4-1b** and **4-1c**.

## 4.2.3 Using phage display to select a 2-acetylarylboronic acid-binding peptide

### 4.2.3.1 First attempt

The selection protocol started with the reactions of biotin derivatized 2-acetylarylboronic acid reagent **4-1a** or controls **4-1b/4-1c** (4 mM, 200  $\mu$ l) with  $2 \times 10^{12}$  plaque forming unit (pfu) total of phage Ser-X<sub>7</sub> library (200  $\mu$ l). These three separate reactions were set up in parallel and performed in PBS buffer (10 mM, pH 7.4) at room temperature for 1 h, after which a dialysis process to remove excess reagents **4-1a/4-1b/4-1c** was set up using Thermo Scientific 10K membrane's molecular weight cut-off, 0.5 ml capacity slide-a-lyzer dialysis cassettes at 4 °C for 2 d (6 buffer exchanges) (Figure 4-6). After the dialysis, the total amount of phages in the reactions was quantified as pfu because each phage generates a plaque in an agar overlay. The phage-displaying library contains a lacZ reporter gene, which leads to a blue plaque formation in a bacterial colony on an IPTG/Xgal plate *via* the activation of  $\beta$ -galactosidase enzyme. Thus, the reaction with the reagent **4-1a** afforded a total of  $1.4 \times 10^{12}$  pfu (253  $\mu$ l), and the control reactions with **4-1b** and **4-1c** provided a total of  $1.1 \times 10^{12}$  (245  $\mu$ l) and  $1.3 \times 10^{12}$  pfu (300  $\mu$ l), respectively.

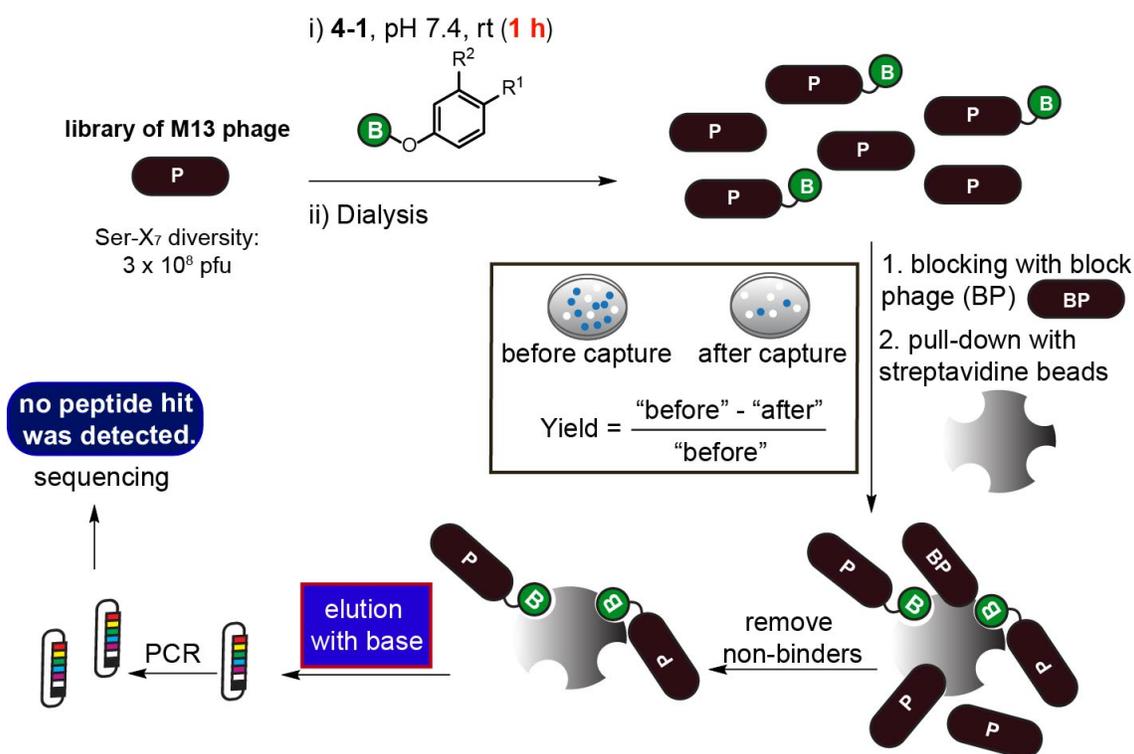
Next, to capture the entire population of biotinylated phages, each reaction mixture was split into three replicates, which all contained a total of  $4.5 \times 10^9$  pfu of phages. Each replicate was mixed with blocking phage (total  $2 \times 10^{10}$  pfu) and a solution of bovine serum albumin (2%

BSA), adjusting the final volume to 1 ml. Blocking phage and reagent BSA are used to reduce non-specific interactions between the beads and biotinylated phages. A post-doctoral fellow, N. Bennett in the Derda research laboratory, constructed the blocking phage. Thanks to its design, the DNA of this blocking phage cannot be amplified under the PCR conditions that are required to obtain double stranded DNA (dsDNA) of the captured phages; therefore, it does not interfere with sequencing and analysis. Then, the blocking process was performed at room temperature for 1.5 h on a Labquake tube rotator (Barnstead/Thermolyne). The high capacity agarose streptavidin beads (200  $\mu$ l of slurry beads solution) were washed with PBS buffer (10 mM, pH 7.4, 3 x 1 ml) for 10 min and the supernatant was removed after micro-centrifugation for 5 min at 2000 rpm. After a third wash, blocking of the beads was performed for 1.5 h with a solution of BSA (2%, 1 ml) and a blocking phage (total  $2 \times 10^{10}$  pfu). Next, the beads were centrifuged for 5 min at 2000 rpm, and the supernatant was gently removed. Preblocked biotinylated phage samples were transferred onto the blocked beads. Capturing of biotinylated phages was performed at room temperature for 1 h on a Labquake tube rotator (Barnstead/Thermolyne). Then, each mixture was centrifuged for 5 min at 2000 rpm, and the supernatant was removed and stored in an eppendorf tube labeled as wash zero. The biotinylated clones were quantified *via* a pull-down assay developed by Derda and co-workers.<sup>26</sup> This assay is based on the ability of a plaque formation by each phage on an agar overlay. As mentioned earlier, the phage-displaying library generates blue plaques while the blocking phage can form white plaques so that they can be differentiated on an agar overlay. In short, the ratio of quantified pfus before and after capture (wash zero) led to the percent capturing yield of the modification (Figure 4-6). This yield was found to be similar (20–35%) for each reaction, and no significant difference was observed between the reaction with **4-1a** and controls **4-1b/4-1c**.

However, this information was insufficient to decide if the selection protocol was unsuccessful due to the fact that the expected reactive population was not high, therefore, a sequencing analysis would be essential to reach a more reliable conclusion. The beads were washed ten times with tween in PBS buffer (0.1%, 1 ml, 10 min in rotator), followed by centrifugation of the mixture and removal of the supernatant. After wash ten, the resulting beads containing biotinylated phages were resuspended with a basic elution buffer (30  $\mu$ l of 10 mM Tris, pH 8.5, 30  $\mu$ l of hexanes), and the resulting mixture was shaken at 3000 rpm for 10 min to extract phage ssDNA. The organic phase (hexanes) was evaporated for 10 min on a heat block

set at 70 °C, and the extracted ssDNAs were subjected immediately to PCR to convert them into Illumina compatible dsDNAs and submit them for Illumina sequencing.

Next, to determine the enriched sequences for the test reagent **4-1a**, deep-sequencing and differential enrichment analysis (volcano analysis) were performed using a procedure similar to the one described by Derda and co-workers,<sup>27,37a</sup> and MatLab script was applied to process the data.<sup>38</sup> However, this attempt did not provide any peptide hits based on the same analysis, which is described in detailed for the second attempt in the subsequent section. A possible explanation for the unsuccessful screening might be due to the short reaction time (1 h), which was likely insufficient to allow the desired conjugation to occur. Furthermore, it was challenging for this conjugation to remain intact under the long dialysis and extensive washing steps; therefore, the targeted phage population was most likely very small, and the differential enrichment in the positive compared to control sets might be difficult to visualize in a reliable, reproducible fashion.



**Figure 4-6:** First attempt to select 2-acetylarylboronic acid reactive peptides.

#### 4.2.3.2 Second attempt

In the second trial, the screening procedure was changed slightly (Figure 4-7A). The reactions between the phage library and the biotinylated reagents (**4-1a/4-1b/4-1c**) were kept longer (3 h). The dialysis step and capturing procedures were not changed; however, the basic elution buffer solution, which was used for ssDNA extraction in the first attempt, was replaced with an acidic elution buffer (100  $\mu$ l, 0.2 M glycine-HCl, pH 2.2 + 0.1% (w/v) BSA, 10 min). This acidic elution buffer did not extract the ssDNA of phages, but a fast-cleavage of the boronic ester and imine conjugations into their boronic acid and ketone units was expected, thus exclusively releasing the phages from the streptavidin beads that were attached *via* the desired conjugation (Figure 4-7B).<sup>39</sup> Next, a neutralization buffer (25  $\mu$ l, 1 M Tris-HCl, pH 9.1) was added to the phage solution. Then, these phage samples were amplified because it was considered that differential enrichment in the positive compared to negative control sets could be visualized much easier due to the resulting higher copy number of captured/eluted phages; in addition, another round of selection could be performed after amplification. However, for this attempt, a single round of selection was performed, but several controls along with the reagents **4-1b** and **4-1c** were used to enable a more reliable analysis (Figure 4-7C). These controls involved the non-amplified samples before capture (sets G, H, I), amplified samples before capture (sets D, E, F) and amplified samples after capture (sets A, B, C). The samples before and after capture were compared in order to determine whether certain sequences were enriched or not. Moreover, non-amplified and amplified samples were analyzed to differentiate the truly reactive peptides because some clones have a tendency to be amplified more efficiently than others.<sup>37b</sup> Therefore, the comparison of amplified and non-amplified phages would be helpful in eliminating fast growing or parasite sequences. Lastly, ssDNA extraction of all these phage samples and their immediate PCR, followed by Illumina sequencing afforded nineteen peptide hits; these results are explained in the subsequent section.



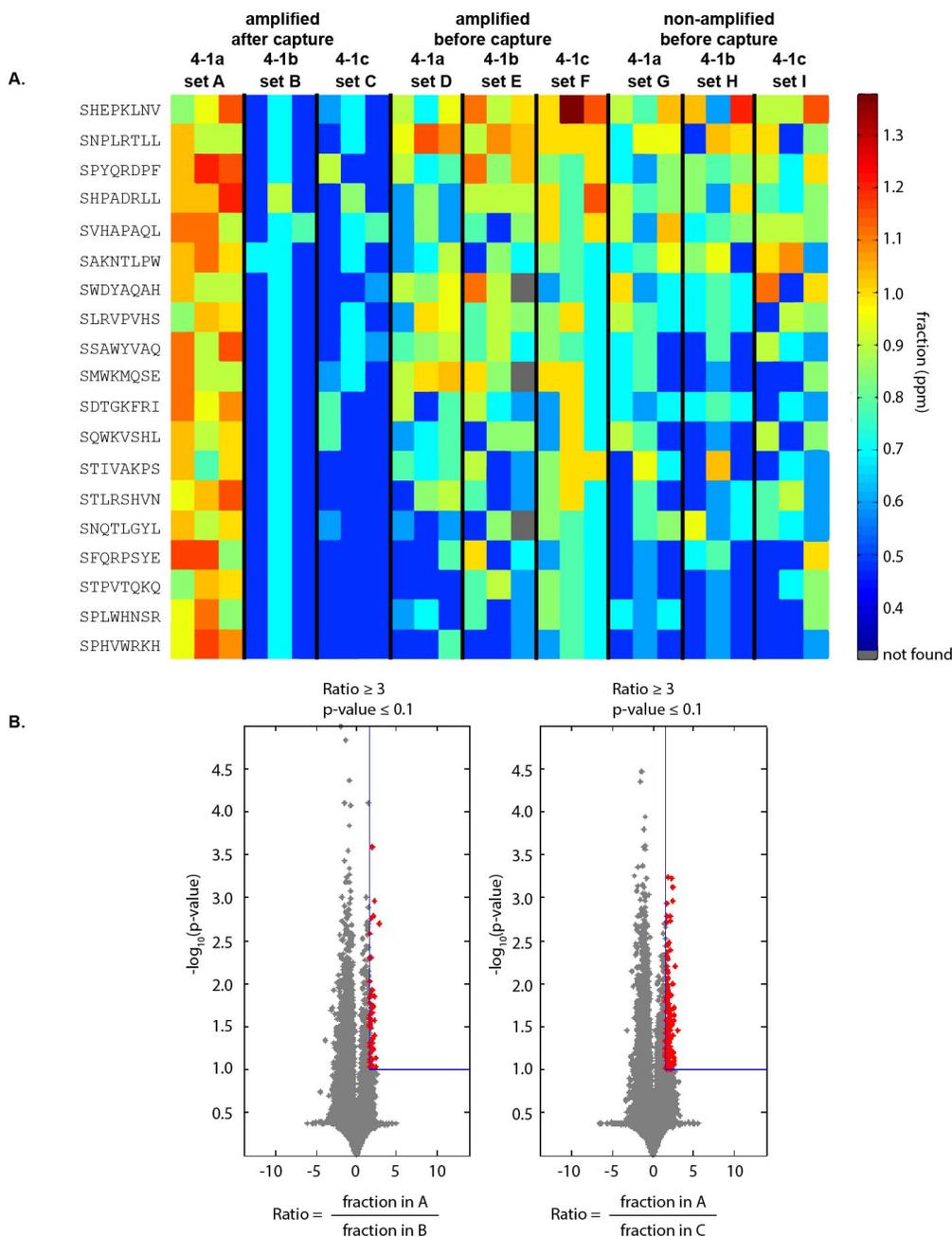
cleavage of the conjugation under acidic elution buffer from beads. C) Definition of test and control sets.

#### 4.2.4 Characterizing the selected peptides – screening and interpretation of data

In order to determine the enriched sequences for the test reagent **4-1a**, deep-sequencing and differential enrichment analysis were performed according to the described procedure by Derda and co-workers,<sup>27, 37a</sup> and similarly MatLab script was applied to process the data (MatLab script entitled with *MakeFigure4-8.m* available in Appendix 2).<sup>38</sup> To identify the peptide hits, set A (population reacted with **4-1a**) was compared with control sets B–I (Figure 4-7C and 4-8A). Sets A–C correspond to sets of amplified populations after capture, sets D–F are amplified populations before capture and sets G–I consist of non-amplified populations before capture. Sets A, D and G correspond to populations incubated with the test reagent **4-1a** while sets B, E, H and sets C, F, I were populations reacted with controls **4-1b** and **4-1c**, respectively. The average normalized frequency for each sequence was determined at each set, and then the ratio was calculated as an average frequency in a test set A divided by an average normalized frequency in the control experiments sets B or C. Also, the p-value between the replicates of set A and replicates of set B or C was calculated using a two-sided unequal variance t-test. Peptide hit sequences were identified using  $R \geq 3$  and  $p\text{-value} \leq 0.1$  for all control experiments, and results were presented as a heat map and a volcano plot, shown in Figure 4-8A and 4-8B, respectively; for example, A/B is the ratio between set A and set B. From the results of the different sets A/B and A/C, nineteen peptide hits were identified. Also, set A was compared with sets D–I, showing an enrichment of most of these peptides after capture as a color difference from blue (less abundant) to red (more abundant). These corresponding clones were not recognized as fast growers because the amplified samples sets D–F and non-amplified sets G–I provided a similar color pattern in the heat map. In other words, most of these peptides were observed in a similar fraction in both sets D–F and G–I.

Next, the identified peptide sequences were analyzed in order to search for a chemical motif that might induce reactivity with the 2-acetylarylboronic acid. Unfortunately, no such motif and chemical similarity in peptides were noticed. Thus, twelve of these nineteen peptide sequences (SDTGKFRI, SQWKVSHL, STLRSHVN, SNQTLGYL, STPVTQKQ, SPHVWRKH, SNPLRTLL, SPYQRDPF, SVHAPAQL, SAKNTLPW, SLRVPVHS,

SSAWYVAQ) were selected randomly in order to test their affinity towards 2-acetylarylboronic acid.



**Figure 4-8:** Post-selection analysis on Ser-X<sub>7</sub> against the reagent 4-1a. A) Heat map describing the peptide hits from the intersection of A/B and A/C. B) Volcano plots showing sequences that are enriched ( $R \geq 3$ ,  $p\text{-value} \leq 0.1$ ) using set B and C as controls. The sets of these sequences were abbreviated as A/B and A/C. The composition of the enriched sequences was indicated as

red on the top right of volcano plots. The Figure 4-8 was created by *MakeFigure4-8.m* MatLab script, which is available in Appendix 2.

#### 4.2.5 Analyzing the reactivity of peptides toward 2-acetylbenzeneboronic acid (2-ABBA)

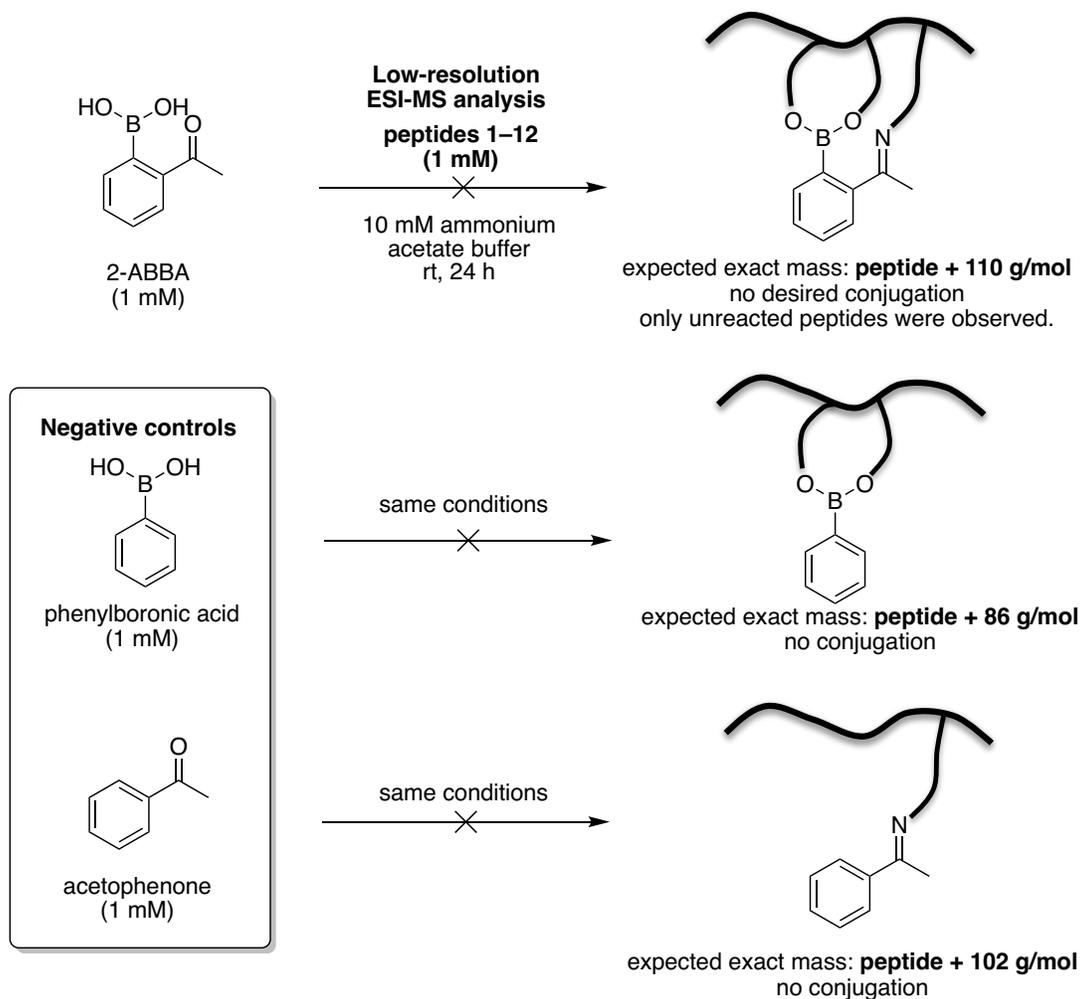
Despite the lack of clear structural motifs on peptide hits that would potentially react with 2-acetylarylboronic acids, twelve peptides SDTGKFRI, SQWKVSHL, STLRSHVN, SNQTLGYL, STPVTQKQ, SPHVWRKH, SNPLRTLL, SPYQRDPF, SVHAPAQL, SAKNTLPW, SLRVPVHS, SSAWYVAQ were selected to test their conjugation with 2-ABBA. These peptides were synthesized using solid phase peptide synthesis (Prelude X Instrument) and purified by preparative HPLC.

ESI-MS analysis was chosen to test the conjugation of peptides with 2-ABBA. This method was considered more suitable compared to HPLC-MS analysis, which might lead to a cleavage of boronic ester and imine in the column. However, ESI-MS analysis, a column-free methodology, would be more direct and accurate in analyzing the conjugations. As a reaction solvent, 10 mM ammonium acetate buffer (an ESI friendly solvent) was used, and as negative controls, parallel reactions of these peptides with phenylboronic acid and acetophenone were performed. The reactions of equimolar concentrations of each peptide and 2-ABBA/phenylboronic acid/acetophenone (1 mM) were prepared in 10 mM ammonium acetate buffer and incubated at room temperature for 24 h. Then, the reactions were analyzed *via* low-resolution ESI-MS analysis (Scheme 4-7). The exact mass of the desired conjugation products was expected to be 110 g/mol more. Negative controls were not supposed to undergo any reaction with the peptides at a concentration of 1 mM, but if they somehow reacted, an additional 86 and 102 g/mol to the peptide exact mass was expected for phenylboronic acid and acetophenone, respectively. However, no desired conjugation was observed between peptides and 2-ABBA (Table 4-1 and Figure 4-9). As expected, negative controls also did not provide any conjugation, and only unreacted peptides were detected *via* (+)- and (-)-ESI-MS analysis. In addition, some of these reactions were performed under high-resolution ESI-MS analysis, but the same results were obtained (Scheme 4-8).

Apart from mass analysis, the reactions of some peptides (**5**, **6** and **11**) with 2-ABBA were analyzed *via*  $^1\text{H}$  NMR spectroscopy in a phosphate buffer. The reactions of equimolar concentrations of peptide and 2-ABBA (2.5 mM) were prepared in 25 mM  $\text{D}_2\text{O}$  phosphate buffer

and incubated at room temperature for 24 h. Then, the reactions were monitored *via*  $^1\text{H}$  NMR analysis (Scheme 4-10). A clear chemical shift was expected upon a desired conjugation; however, no such chemical shift was observed (Figure 4-10). All these results support the fact that these peptides cannot undergo a desired and stable bioconjugation with 2-ABBA under the conditions used.

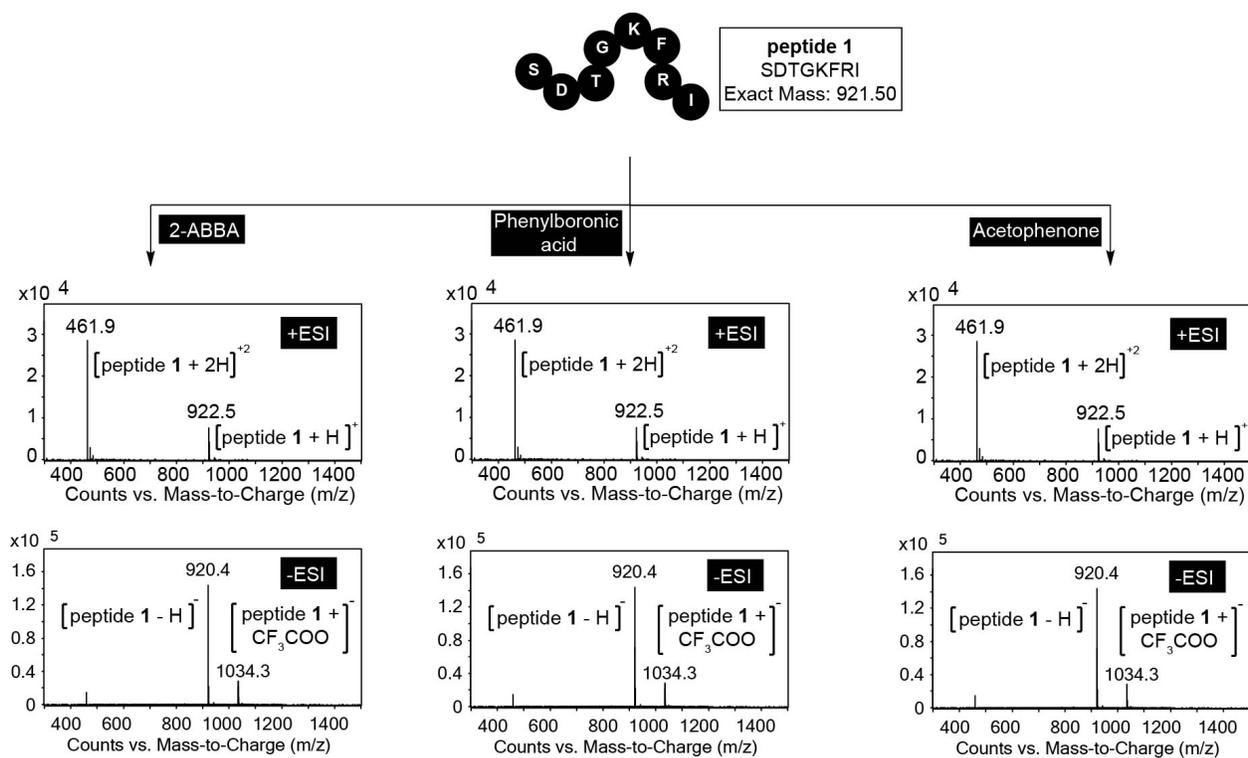
As mentioned above, this hypothetical conjugation is more challenging compared to the reported systems<sup>24,25</sup> because both expected boronate and imine conjugations are intrinsically less stable than nopolboronate and thiosemicarbazone ligations. Furthermore, 2-acetylarylboronic acid is less reactive than 2-formylarylboronic acid. The next steps toward the discovery of reactive peptides may involve repeating the selection over two or three rounds, with more extensive washing conditions, also with using different phage libraries enriched with lysine, serine and threonine. Furthermore, the reactivity of reagent **4-1a** can be improved by using a more electron deficient boronic acid derivative, such as 2-acetyl-4-carboxyphenylboronic acid, which was shown in Chapter 3 to exhibit a higher reactivity toward thiosemicarbazide-functionalized nopolydiol.<sup>24</sup>



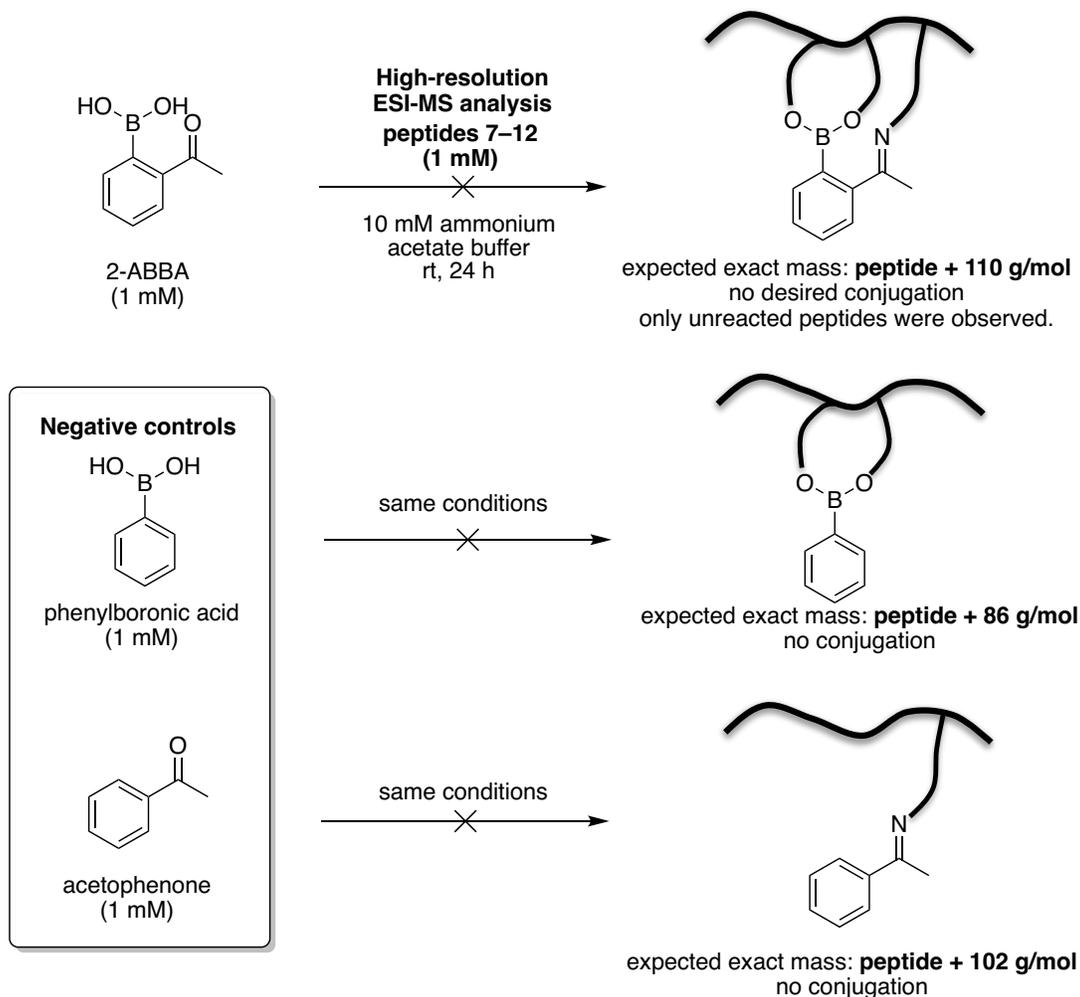
**Scheme 4-7:** Reactions of peptides 1–12 with 2-ABBA/negative controls and their analysis *via* low-resolution ESI-MS.

**Table 4-1:** Conjugation results of peptides **1–12** with 2-ABBA *via* low-resolution ESI-MS.

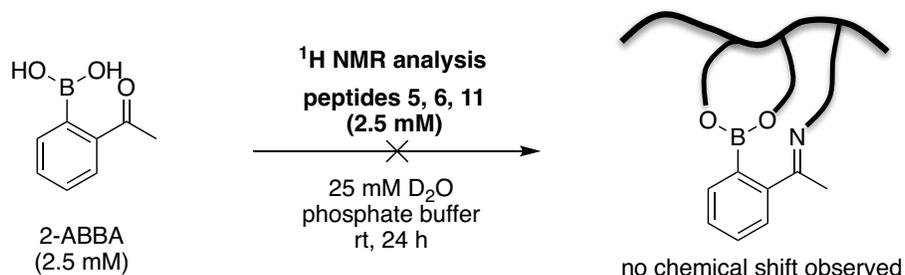
<b>Peptides (1 mM)</b>	<b>Exact mass</b>	<b>Desired conjugation with 2-ABBA (1 mM)?</b>
<b>1:</b> SDTGKFRI	921.50	No conjugation was detected.
<b>2:</b> SQWKVSHL	982.53	
<b>3:</b> STLRSHVN	911.49	
<b>4:</b> SNQTLGYL	893.46	
<b>5:</b> STPVTQKQ	886.49	
<b>6:</b> SPHVWRKH	1044.57	
<b>7:</b> SNPLRTLL	911.56	
<b>8:</b> SPYQRDPF	1007.48	
<b>9:</b> SVHAPAQL	820.46	
<b>10:</b> SAKNTLPW	914.50	
<b>11:</b> SLRVPVHS	892.52	
<b>12:</b> SSAWYVAQ	909.43	



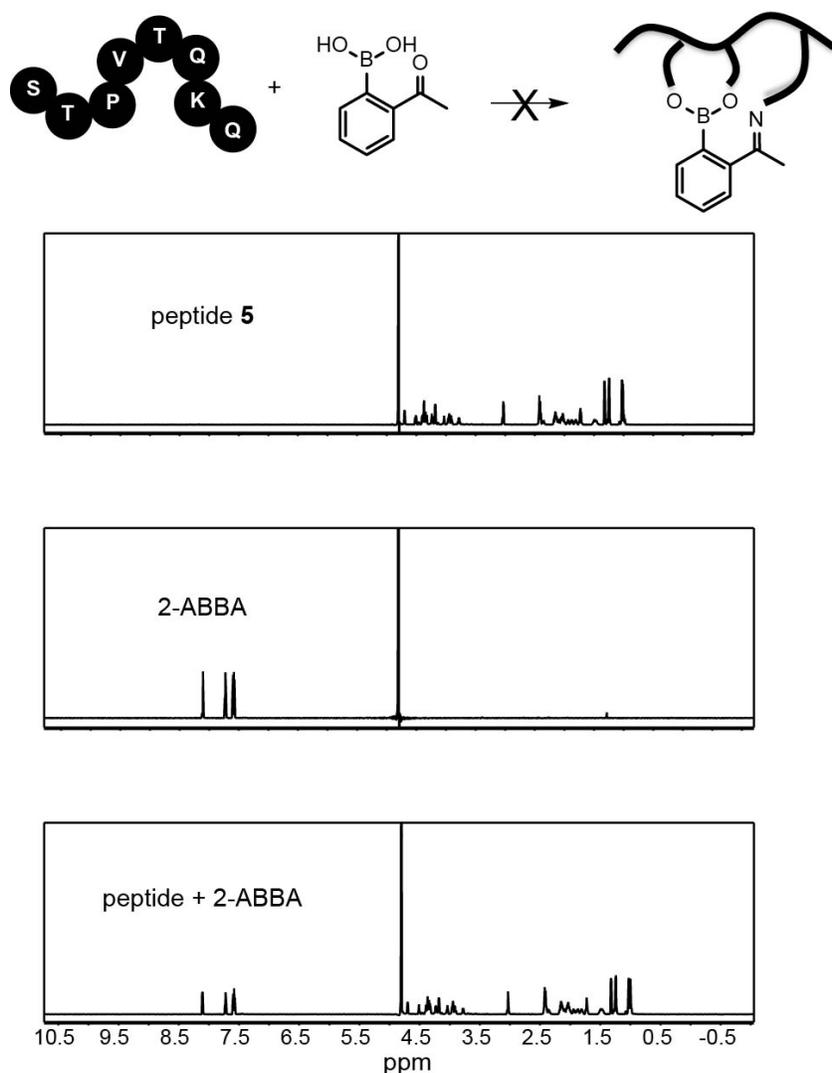
**Figure 4-9:** A representative result of low-resolution ESI-MS analysis of the reaction between peptide 1 and 2-ABBA/phenylboronic acid/acetophenone.



**Scheme 4-8:** Reactions of peptides 7–12 with 2-ABBA/negative controls and their analysis via high-resolution ESI-MS.



**Scheme 4-9:** Reactions of peptides 5/6/11 with 2-ABBA and their analysis via <sup>1</sup>H NMR analysis.



**Figure 4-10:** A representative result of  $^1\text{H}$  NMR analysis of the reaction between peptide 5 and 2-ABBA.

### 4.3 Conclusions

As a conclusion, Chapter 4 presented work performed towards the selection of peptide tags *via* phage display technology, followed by the synthesis and validation of selected peptides with 2-acetylarylboronic acid. It was hoped to replace the thiosemicarbazide-functionalized nopoldiol described in Chapter 3 with a short and reactive peptide tag, which also could exhibit a high reactivity towards 2-acetylarylboronic acid *via* both imine and boronate formation. It was envisioned that 2-acetylarylboronic acid could form a tight complex with a short peptide embedding as many as three hydroxyl side chains (serine, threonine) and a lysine residue or a

terminal amine. To investigate the reactive peptides, the phage display platform developed by Derda and co-workers was used. An *N*-terminal serine library of over  $10^8$  heptapeptides displayed on phage was screened over one-round of selection. Deep-sequencing and differential enrichment analyses were used to identify the peptide-hit sequences. Even though these peptide sequences lacked clear structural motifs that could react with 2-acetylarylboronic acids, twelve peptides SDTGKFRI, SQWKVSHL, STLRSHVN, SNQTLGYL, STPVTQKQ, SPHVWRKH, SNPLRTLL, SPYQRDPF, SVHAPAQL, SAKNTLPW, SLRVPVHS, SSAWYVAQ were selected randomly out of nineteen hits to test their conjugation with 2-ABBA. The conjugation studies were performed with ESI-MS and  $^1\text{H}$  NMR analyses, which unfortunately confirmed the lack of desired tight ligation.

## 4.4 Experimental

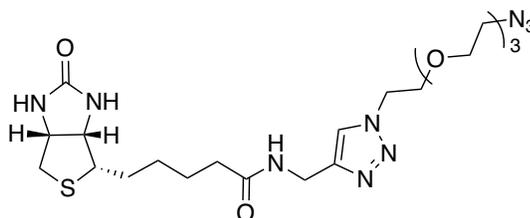
### 4.4.1 General information

Unless otherwise stated, all reactions were performed under a nitrogen atmosphere using flame-dried glassware. THF, dichloromethane, DMF and methanol were obtained from a MBraun MB SPS\* solvent system prior to use. Most of the reagents were purchased from Sigma-Aldrich and used as received. Arylboronic acids were purchased either from Sigma-Aldrich or Combi-Blocks. Thin layer chromatography (TLC) was performed on Merck Silica Gel 60 F254 plates and was visualized with UV light, curcumin and  $\text{KMnO}_4$  stain. NMR spectra were recorded on INOVA-400, INOVA-500, INOVA-600 or INOVA-700 MHz instruments. The residual solvent protons ( $^1\text{H}$ ) of  $\text{CDCl}_3$  (7.26 ppm),  $\text{ACN-d}_3$  (1.94 ppm),  $\text{DMSO-d}_6$  (2.50 ppm), acetone- $\text{d}_6$  (2.05 ppm),  $\text{CD}_3\text{OD}$  (3.31 ppm) and the solvent carbons ( $^{13}\text{C}$ ) of  $\text{CDCl}_3$  (77.06 ppm),  $\text{ACN-d}_3$  (1.32 and 118.26 ppm),  $\text{DMSO-d}_6$  (39.52 ppm), acetone- $\text{d}_6$  (29.84 and 206.26 ppm),  $\text{CD}_3\text{OD}$  (49 ppm) were used as internal standards.  $^1\text{H}$  NMR data is presented as follows: chemical shift in ppm ( $\delta$ ) downfield from tetramethylsilane (multiplicity, coupling constant, integration). The following abbreviations are used in reporting NMR data: s, singlet; br s, broad singlet; d, doublet; t, triplet; q, quartet; dd, doublet of doublets; ddd, doublet of doublet of doublets; dddd, doublet of doublet of doublet of doublets; app s, apparent singlet; app ddt, apparent doublet of doublet of triplets; m, multiplet; comp m, complex multiplet. The accuracy of the reported coupling constant ( $J$ ) is  $\pm 0.3$  Hz. In  $^{13}\text{C}$  NMR spectroscopy, the quaternary carbon bound to the boron atom is often missing due to the quadrupolar relaxation of boron. This

effect was observed in each boron-containing compound. High-resolution mass spectra were recorded by the University of Alberta mass spectrometry services laboratory using either electron impact (EI) or electrospray ionization (ESI) techniques. Optical rotations were measured using a 1 mL cell with a 1 dm length on a P.E. 241 polarimeter. Melting points were determined in a capillary tube using a Gallenkamp melting point apparatus and are uncorrected.

#### 4.4.2 Chemical synthesis and analytical data

##### Synthesis of biotin-PEG-azide (4-5)



***N*-((1-(2-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)ethyl)-1*H*-1,2,3-triazol-4-yl)methyl)-5-((3*aS*,4*S*,6*aR*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamide (4-5):** Biotin-PEG-alkyne (4-3) and PEG-diazide (4-4) were synthesized based on the described procedures by Yu and co-workers and Hawthorne and co-workers, respectively.<sup>30,31</sup> To the solution of biotin-PEG-alkyne 4-3 (1.0 equiv, 0.75 g, 2.7 mmol) and PEG-diazide 4-4 (1.5 equiv, 0.98 g, 4.0 mmol) in EtOH:water:*t*-BuOH:ACN (6:4:10:5 ml), CuSO<sub>4</sub>•5H<sub>2</sub>O (0.46 ml, 0.54 M in water, freshly prepared) was added under ambient atmosphere at room temperature. Sodium ascorbate (0.46 ml, 1.12 M in water, freshly prepared) was added to the reaction mixture under vigorous stirring. The reaction mixture was stirred for 24 h at room temperature, filtered through a plug of celite and concentrated in vacuo to obtain a dark green crude oil. The crude residue was purified by flash chromatography (dry loading, 10:90 to 15:85 MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to obtain biotin-PEG-azide (4-5) as a colorless viscous oil (0.63 g, 45% yield).

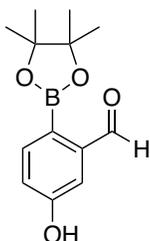
<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 7.90 (s, 1H), 4.56 (app t, 2H), 4.49 (ddd, *J* = 7.9, 5.0, 1.0 Hz, 1H), 4.43 (s, 2H), 4.30 (dd, *J* = 7.9, 4.4 Hz, 1H), 3.88 (app t, 2H), 3.69 – 3.58 (comp m, 10H), 3.37 (app t, 2H), 3.20 (ddd, *J* = 8.9, 5.9, 4.4 Hz, 1H), 2.93 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.71 (d, *J* = 12.7 Hz, 1H), 2.24 (t, *J* = 7.3 Hz, 2H), 1.77 – 1.54 (comp m, 4H), 1.48 – 1.37 (m, 2H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 175.9, 166.1, 146.1, 125.0, 71.6, 71.59, 71.56, 71.5, 71.1, 70.4, 63.3, 61.6, 57.0, 51.8, 51.4, 41.0, 36.6, 35.6, 29.7, 29.4, 26.7.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3294, 3075, 2930, 2865, 2109, 1701, 1644, 1547, 1462, 1120.

**HRMS** (ESI-TOF):  $\text{C}_{21}\text{H}_{35}\text{N}_9\text{NaO}_5\text{S}$  ( $\text{M} + \text{Na}$ )<sup>+</sup>: *calcd.*: 548.2374; *found*: 548.2366; for  $\text{C}_{21}\text{H}_{36}\text{N}_9\text{O}_5\text{S}$  ( $\text{M} + \text{H}$ )<sup>+</sup>: *calcd.*: 526.2555; *found*: 526.2558.

### Synthesis of 2-acetyl-4-hydroxyphenylboronic acid pinacol ester (4-7)



**5-Hydroxy-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzaldehyde (4-7)**: 2-Formyl-4-hydroxy bromobenzene (**4-6**) was synthesized based on the described procedure by Schmaiz and co-workers.<sup>40</sup> The procedure reported by Shabat and co-workers was adapted to prepare 2-acetyl-4-hydroxyphenylboronic acid pinacol ester (**4-7**).<sup>32</sup> Compound **4-6** (1.0 equiv, 0.25 g, 1.2 mmol), bis(pinacolato)diboron (1.1 equiv, 0.35 g, 1.4 mmol) and potassium acetate (3.0 equiv, 0.37 g, 3.7 mmol) were added to a flame dried round-bottom flask, degassed and purged with argon. Then, freshly distilled dioxane (5 ml) was added. Lastly, [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium(II) complex with dichloromethane ( $\text{Pd}(\text{dppf})\text{Cl}_2$ ) (3 mol%, 33 mg, 0.04 mmol) was added, and the reaction mixture was stirred at 100 °C for 2 h. After the reaction mixture reached rt,  $\text{CH}_2\text{Cl}_2$  (20 ml) was added to the reaction mixture. The organic phase was washed with distilled water (1 x 5 ml), dried ( $\text{MgSO}_4$ ), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:4, EtOAc/hexanes) to provide compound **4-7** as a white solid (0.10 g, 33%).

White solid:  $R_f = 0.22$  (1:4, EtOAc/hexanes).

**$^1\text{H}$  NMR**  $\delta/\text{ppm}$ : (500 MHz,  $\text{CDCl}_3$ ) 10.65 (s, 1H), 7.85 (d,  $J = 8.1$  Hz, 1H), 7.51 (d,  $J = 2.6$  Hz, 1H), 7.10 (dd,  $J = 8.2, 2.6$  Hz, 1H), 6.16 (bs, 1H), 1.36 (s, 12H).

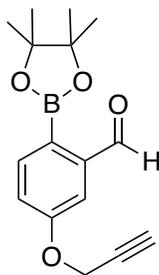
**$^{13}\text{C}$  NMR**  $\delta/\text{ppm}$ : (126 MHz,  $\text{CDCl}_3$ ) 195.3, 158.5, 143.6, 138.5, 120.5, 113.3, 84.3, 24.9.

**$^{11}\text{B}$  NMR**  $\delta/\text{ppm}$ : (128 MHz,  $\text{CDCl}_3$ ) 30.8.

**IR** (Microscope,  $\text{cm}^{-1}$ ): 3431, 2976, 2931, 1680, 1600, 1346, 1294, 1143.

**HRMS** (ESI-TOF): for  $\text{C}_{13}\text{H}_{16}[^{11}\text{B}]\text{O}_4$  ( $\text{M} - \text{H}$ )<sup>-</sup>: *calcd.*: 247.1147; *found*: 247.1144.

### Synthesis of 2-acetyl-4-propargylphenylboronic acid pinacol ester (4-8)



### 5-(Prop-2-yn-1-yloxy)-2-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)benzaldehyde (4-8):

The procedure described by Gillingham and co-workers was used to prepare 2-acetyl-4-propargylphenylboronic acid pinacol ester (4-8).<sup>33</sup>

White solid:  $R_f = 0.46$  (1:4, EtOAc/hexanes).

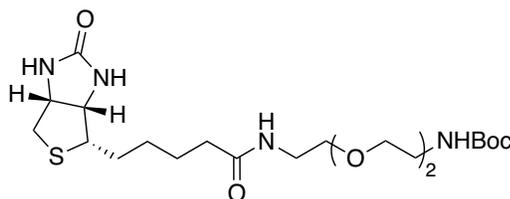
$^1\text{H NMR}$   $\delta$ /ppm: (500 MHz,  $\text{CDCl}_3$ ) 10.66 (s, 1H), 7.89 (d,  $J = 8.3$  Hz, 1H), 7.57 (d,  $J = 2.6$  Hz, 1H), 7.19 (dd,  $J = 8.3, 2.7$  Hz, 1H), 4.77 (d,  $J = 2.4$  Hz, 2H), 2.53 (t,  $J = 2.4$  Hz, 1H), 1.37 (s, 12H).

$^{13}\text{C NMR}$   $\delta$ /ppm: (126 MHz,  $\text{CDCl}_3$ ) 194.6, 159.8, 143.6, 138.0, 120.6, 111.6, 84.3, 77.8, 76.1, 55.9, 24.9.

$^{11}\text{B NMR}$   $\delta$ /ppm: (128 MHz,  $\text{CDCl}_3$ ) 30.9.

HRMS (EI): for  $\text{C}_{16}\text{H}_{19}[^{11}\text{B}]\text{O}_4$  ( $\text{M} - \text{H}$ )<sup>-</sup>: *calcd.*: 286.13763; *found*: 286.13764.

### Synthesis of biotin-PEG-NHBoc (4-10)



### *Tert*-butyl-(2-(2-(2-(5-((4*S*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-

yl)pentanamido)ethoxy)ethoxy)ethyl)carbamate (4-10): PEG-NHBoc was synthesized based on the described procedure by Tsukiji and co-workers.<sup>41</sup> D-(+)-biotin (1.2 equiv, 0.58 g, 2.4 mmol) and HATU coupling reagent (1.2 equiv, 0.91 g, 2.4 mmol) were dissolved in dry DMF (38 ml) at room temperature under an argon balloon. The PEG-NHBoc (1.0 equiv, 0.50 g, 2.0 mmol) was dissolved separately in  $\text{CH}_2\text{Cl}_2$  (19 ml), and this solution was added to the reaction

mixture. Lastly, dipea (2.4 equiv, 0.84 ml, 4.8 mmol) was added. The reaction mixture was stirred for 2 h at room temperature, concentrated in vacuo and mixed with CH<sub>2</sub>Cl<sub>2</sub> (50 ml). The organic phase was washed with HCl (1 × 10 ml, 1 N), and brine (1 × 10 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. The crude residue was purified by flash chromatography (1:9, MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide compound **4-10** as a white powder (0.84 g, 89% yield). Compound **4-10** characterization was confirmed with the literature example.<sup>36</sup>

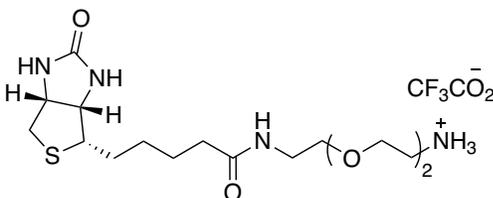
White powder:  $R_f = 0.36$  (1:9, MeOH/CH<sub>2</sub>Cl<sub>2</sub>).

<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 4.49 (ddd,  $J = 7.8, 5.0, 0.9$  Hz, 1H), 4.30 (dd,  $J = 7.9, 4.5$  Hz, 1H), 3.62 (s, 4H), 3.53 (app dt,  $J = 15.5, 5.6$  Hz, 4H), 3.37 (app q,  $J = 5.5$  Hz, 2H), 3.25 – 3.18 (comp m, 3H), 2.93 (dd,  $J = 12.7, 5.0$  Hz, 1H), 2.71 (d,  $J = 12.7$  Hz, 1H), 2.22 (t,  $J = 7.4$  Hz, 2H), 1.80 – 1.53 (comp m, 5H), 1.44 (comp m, 11H).

IR (Microscope, cm<sup>-1</sup>): 3299, 2972, 2931, 2868, 1706, 1645, 1530, 1170.

HRMS (ESI-TOF): for C<sub>21</sub>H<sub>38</sub>N<sub>4</sub>NaO<sub>6</sub>S (M + Na)<sup>+</sup>: *calcd.*: 497.2404; *found*: 497.2399; for C<sub>21</sub>H<sub>39</sub>N<sub>4</sub>O<sub>6</sub>S (M + H)<sup>+</sup>: *calcd.*: 475.2585; *found*: 475.259.

### Synthesis of biotin-PEG-amine (4-11)



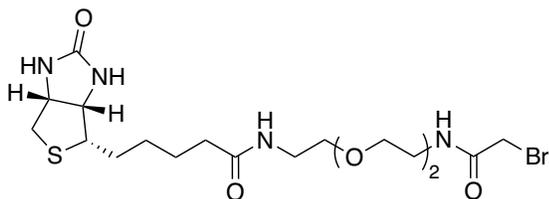
### 2-(2-(2-(5-((4S)-2-Oxohexahydro-1H-thieno[3,4-d]imidazol-4-

yl)pentanamido)ethoxy)ethoxy)ethan-1-aminium 2,2,2-trifluoroacetate (**4-11**): Biotin-PEG-amine synthesis was performed based on the literature described by Muir and co-workers.<sup>36</sup> Biotin-PEG-NHBoc **4-10** (1.0 equiv, 0.76 g, 1.6 mmol) was dissolved in TFA/CH<sub>2</sub>Cl<sub>2</sub> (2.5 ml/2.5 ml) mixture (1:1) and was stirred at room temperature for 2 h. The reaction mixture was concentrated under vacuo. The crude residue was dissolved in water (15 ml) and washed with toluene (2 × 3 ml). The aqueous solution was lyophilized to give biotin-PEG-amine **4-11** as a viscous oil (quantitative yield), which was used without further purification for the next step.

IR (Microscope, cm<sup>-1</sup>): 3286, 3091, 2932, 2876, 1686, 1550, 1202, 1140.

HRMS (ESI-TOF): for C<sub>16</sub>H<sub>31</sub>N<sub>4</sub>O<sub>4</sub>S (M)<sup>+</sup>: *calcd.*: 375.2061; *found*: 375.2062.

## Synthesis of biotin-PEG-bromide (4-12)



***N*-(2-(2-(2-(2-Bromoacetamido)ethoxy)ethoxy)ethyl)-5-((4*S*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamide (4-12)**: Biotin-PEG-amine **4-11** (1.0 equiv, 0.50 g, 1.0 mmol) and Et<sub>3</sub>N (2.0 equiv, 0.39 ml, 2.0 mmol) were dissolved in dry DMF (3.5 ml) under an inert atmosphere. Bromoacetyl bromide (1.5 equiv, 0.13 ml, 1.5 mmol) was added dropwise to the reaction mixture under ice bath. Then, the reaction mixture was stirred under ice bath for 1 h, after which it was concentrated under vacuo. The crude residue was purified by flash chromatography (5:95 to 20:80 MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide biotin-PEG-bromide **4-12** as a viscous brown oil (0.21 g, 42% yield).

Viscous brown oil: *R<sub>f</sub>* = 0.32 (15:85, MeOH/CH<sub>2</sub>Cl<sub>2</sub>).

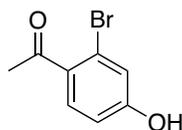
<sup>1</sup>H NMR δ/ppm: (500 MHz, DMSO-*d*<sub>6</sub>) 7.86 – 7.76 (comp m, *J* = 7.5 Hz, 3H), 6.41 (s, 2H), 4.31 (ddd, *J* = 7.7, 5.1, 1.0 Hz, 1H), 4.13 (dd, *J* = 7.8, 4.4 Hz, 1H), 3.63 – 3.49 (comp m, 6H), 3.40 (t, *J* = 6.1 Hz, 2H), 3.19 (app q, *J* = 6.0 Hz, 2H), 3.09 (ddd, *J* = 8.6, 6.1, 4.4 Hz, 1H), 2.97 (app h, *J* = 5.6 Hz, 2H), 2.82 (dd, *J* = 12.5, 5.1 Hz, 1H), 2.58 (d, *J* = 12.8 Hz, 1H), 2.07 (t, *J* = 7.4 Hz, 2H), 1.61 (ddt, *J* = 13.5, 9.7, 6.1 Hz, 1H), 1.55 – 1.38 (comp m, 3H), 1.37 – 1.19 (comp m, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, DMSO-*d*<sub>6</sub>) 172.1, 166.1, 162.7, 69.6, 69.5, 69.2, 68.7, 61.0, 59.2, 55.4, 45.7, 38.4, 35.1, 29.4, 28.2, 28.0, 25.2, 8.6.

IR (Microscope, cm<sup>-1</sup>): 3281, 3078, 2921, 2869, 1699, 1676, 1600, 1548, 1455, 1253.

HRMS (ESI-TOF): for C<sub>18</sub>H<sub>31</sub>BrN<sub>4</sub>NaO<sub>5</sub>S (M + Na)<sup>+</sup>: *calcd.*: 517.1091; *found*: 517.1088.

## Synthesis of 2-acetyl-5-hydroxybromobenzene (4-13)



**1-(2-Bromo-4-hydroxyphenyl)ethan-1-one (4-13)**: AlCl<sub>3</sub> (2.5 equiv, 0.84 g, 6.3 mmol) was added slowly to a solution of 2-acetyl-5-methoxy bromobenzene (1.0 equiv, 0.58 g, 2.5 mmol) in

toluene (17 ml) under an inert atmosphere, and the mixture was refluxed at 130 °C for 2 h. The reaction mixture was brought to rt and poured into HCl (3.0 M). The aqueous solution was extracted with EtOAc (3 x 25 ml), and the combined organic phase was washed with brine (10 ml), dried over MgSO<sub>4</sub> and concentrated in vacuo. The crude residue was purified by flash chromatography (1:9, EtOAc/hexane) to provide 2-acetyl-5-hydroxy bromobenzene **4-13** as a brown solid (0.48 g, 96% yield).

Brown solid:  $R_f = 0.35$  (3:7, EtOAc/hexanes).

Melting point: 68.1–71.3 °C.

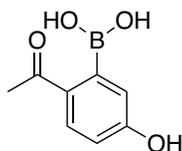
<sup>1</sup>H NMR δ/ppm: (700 MHz, CDCl<sub>3</sub>) 7.56 (d,  $J = 8.4$  Hz, 1H), 7.13 (d,  $J = 2.5$  Hz, 1H), 6.83 (dd,  $J = 8.5, 2.5$  Hz, 1H), 5.28 (s, 1H), 2.63 (s, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CDCl<sub>3</sub>) 200.1, 159.0, 132.6, 132.2, 121.6, 121.5, 114.7, 30.2.

IR (Microscope, cm<sup>-1</sup>): 3204, 1659, 1600, 1561, 1437, 1221.

HRMS (EI): for C<sub>8</sub>H<sub>7</sub>O<sub>2</sub>[<sup>81</sup>Br] (M)<sup>+</sup>: *calcd.*: 215.96089; *found*: 215.96046; for C<sub>7</sub>H<sub>4</sub>O<sub>2</sub>[<sup>81</sup>Br] (M – CH<sub>3</sub>)<sup>+</sup>: *calcd.*: 200.93742; *found*: 200.93699.

### Synthesis of 2-acetyl-5-hydroxyphenylboronic acid (**4-14**)



**(2-Acetyl-5-hydroxyphenyl)boronic acid (4-14)**: Compound **4-13** (1.0 equiv, 0.22 g, 1.0 mmol), bis(pinacolato)diboron (1.1 equiv, 0.28 g, 1.1 mmol) and potassium acetate (3.0 equiv, 0.30 g, 3.0 mmol) were added to a flame dried round-bottom flask, degassed and purged with nitrogen. Then, freshly distilled dioxane (4 ml) was added to the reaction mixture under nitrogen. Lastly, [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium(II) complex with dichloromethane (Pd(dppf)Cl<sub>2</sub>) (3 mol%, 24 mg, 0.030 mmol) was added, and the reaction mixture was stirred at 95 °C for 3 h. After the reaction mixture reached rt, CH<sub>2</sub>Cl<sub>2</sub> (20 ml) was added. The organic phase was washed with distilled water (1 x 5 ml), dried (MgSO<sub>4</sub>), filtered and concentrated in vacuo. After the crude residue was exposed to a short silica plug and washed with EtOAc, it was concentrated in vacuo. Without further purification, the crude residue and phenylboronic acid (5.0 equiv, 0.61 g, 5.0 mmol) were mixed in ACN/HCl (1M) (14 ml/1.5 ml)

at room temperature under ambient atmosphere, and the reaction mixture was stirred for 24 h. Then, it was concentrated in vacuo, and the crude residue was purified by flash chromatography (5:95, MeOH/CH<sub>2</sub>Cl<sub>2</sub>) to provide 2-acetyl-5-hydroxyphenylboronic acid as a yellow solid (77 mg, 30% yield over two steps).

Yellow solid:  $R_f = 0.37$  (5:95, MeOH/CH<sub>2</sub>Cl<sub>2</sub>).

<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 7.98 (s, 1H), 7.91 (d,  $J = 8.5$  Hz, 1H), 6.83 (dd,  $J = 8.5, 2.5$  Hz, 1H), 6.78 (d,  $J = 2.4$  Hz, 1H), 2.99 (s, 3H), 2.86 (s, 3H). It is important to note that this NMR sample lacks a drop of D<sub>2</sub>O, which is generally required for boronic acid samples to break down a boroxine formation. However, when D<sub>2</sub>O was added to this sample, its various isomers were observed due to the presence of the acetyl group. Moreover, when only CD<sub>3</sub>OD was used, more clear peaks and acetal formation were observed, likely due to the addition of methanol to the acetyl group, thus leading to the extra methyl protons at 2.99 ppm.

<sup>13</sup>C NMR δ/ppm: <sup>13</sup>C NMR (126 MHz, CD<sub>3</sub>OD) 202.1, 164.8, 164.6, 132.9, 118.3, 116.0, 36.9, 31.6.

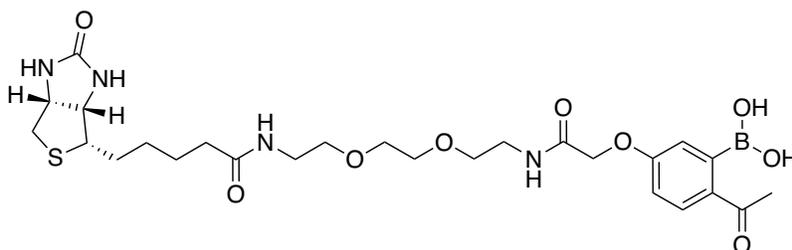
<sup>11</sup>B NMR δ/ppm: (128 MHz, CD<sub>3</sub>OD) 27.4.

IR (Microscope, cm<sup>-1</sup>): 3251, 2948, 1642, 1596, 1555, 1457, 1368, 1305.

HRMS (ESI-TOF) for C<sub>8</sub>H<sub>8</sub>[<sup>11</sup>B]O<sub>4</sub>(M - H)<sup>-</sup>: *calcd.*: 179.0521; *found*: 179.0519.

### General synthesis of biotin-PEG derivatized 2-acetylarylboronic acid derivative (4-1a) and controls (4-1b and 4-1c)

Compound **4-14/4-15/4-16** (1.3 equiv, 0.052 mmol), biotin-PEG-bromide **4-12** (1.0 equiv, 20 mg, 0.040 mmol) and K<sub>2</sub>CO<sub>3</sub> (1.5 equiv, 8.4 mg, 0.061 mmol) were mixed in dry DMF (0.3 ml) under an argon balloon at room temperature. The reaction mixture was stirred for 24 h, after which it was concentrated under vacuo. The crude residue was purified using reverse phase HPLC flash chromatography (C18 column) with water and ACN under neutral conditions to provide biotin derivatized compounds **4-1a/4-1b/4-1c** as transparent viscous oils.



**(2-Acetyl-5-((2,13-dioxo-17-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-6,9-dioxo-3,12-diazaheptadecyl)oxy)phenyl)boronic acid (4-1a):**

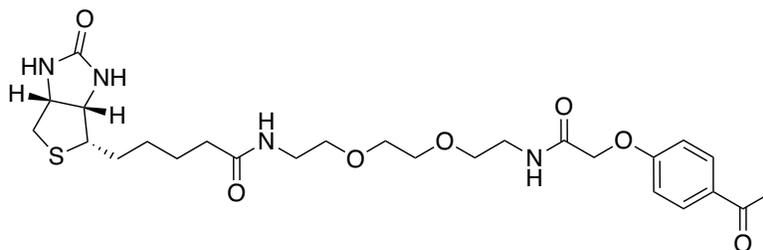
<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 8.04 (d, *J* = 8.6 Hz, 1H), 7.07 (dd, *J* = 8.6, 2.6 Hz, 1H), 7.02 (d, *J* = 2.5 Hz, 1H), 4.48 (ddd, *J* = 7.9, 5.0, 0.9 Hz, 1H), 4.29 (dd, *J* = 7.9, 4.5 Hz, 1H), 3.64 – 3.57 (comp m, 6H), 3.54 (t, *J* = 5.5 Hz, 2H), 3.48 (t, *J* = 5.6 Hz, 2H), 3.38 – 3.33 (m, 2H), 3.24 – 3.17 (comp m, 4H), 2.92 (dd, *J* = 12.7, 5.0 Hz, 1H), 2.69 (d, *J* = 12.7 Hz, 1H), 2.61 (s, 1H), 2.60 (dd, *J* = 2.7, 1.7 Hz, 1H), 2.20 (t, *J* = 7.4 Hz, 2H), 1.77 – 1.53 (comp m, 4H), 1.48 – 1.38 (m, 2H), 1.31 (app t, *J* = 7.3 Hz, 4H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 201.8, 176.1, 170.5, 166.1, 163.4, 134.9, 132.7, 118.1, 115.2, 71.3, 70.6, 70.4, 68.2, 63.4, 61.6, 57.0, 47.9, 41.0, 40.3, 40.0, 36.7, 29.8, 29.5, 26.8, 9.2.

<sup>11</sup>B NMR δ/ppm: (128 MHz, CD<sub>3</sub>OD) 28.6.

IR (Microscope, cm<sup>-1</sup>): 3304, 3090, 2930, 2872, 1680, 1598, 1561, 1203, 1132.

HRMS (ESI-TOF): for C<sub>26</sub>H<sub>38</sub>[<sup>11</sup>B]N<sub>4</sub>O<sub>9</sub>S (M – H)<sup>-</sup>: *calcd.*: 593.2458; *found*: 593.247; for C<sub>26</sub>H<sub>39</sub>[<sup>11</sup>B]N<sub>4</sub>NaO<sub>9</sub>S (M + Na)<sup>+</sup>: *calcd.*: 617.2423; *found*: 617.2412.



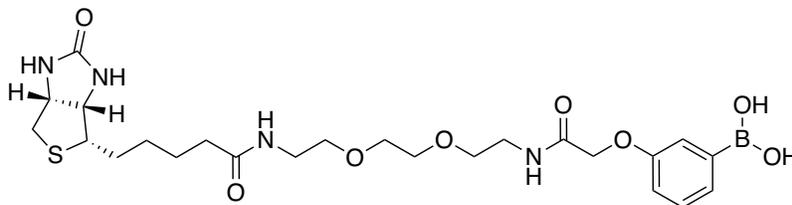
**N-(2-(2-(2-(2-(4-Acetylphenoxy)acetamido)ethoxy)ethoxy)ethyl)-5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (4-1b):**

<sup>1</sup>H NMR δ/ppm: 500 MHz, CD<sub>3</sub>OD) 8.00 (d, *J* = 8.8 Hz, 2H), 7.09 (d, *J* = 8.9 Hz, 2H), 4.51 – 4.46 (m, 1H), 4.30 (td, *J* = 7.4, 4.4 Hz, 1H), 3.64 – 3.57 (comp m, 7H), 3.48 (t, *J* = 5.6 Hz, 2H), 3.38 – 3.33 (comp m, 3H), 3.23 – 3.16 (m, 1H), 2.92 (dt, *J* = 12.7, 5.1 Hz, 1H), 2.70 (dd, *J* = 12.8, 4.5 Hz, 1H), 2.57 (s, 2H), 2.25 – 2.16 (comp m, 3H), 1.77 – 1.54 (comp m, 5H), 1.48 – 1.38 (comp m, 3H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 199.3, 176.1, 170.5, 166.1, 163.2, 132.3, 131.9, 115.8, 71.3, 70.6, 70.5, 69.8, 68.2, 63.4, 61.6, 57.0, 41.0, 40.8, 40.3, 40.0, 36.7, 29.8, 29.5, 26.8, 26.5.

IR (Microscope, cm<sup>-1</sup>): 3297, 3080, 2928, 2867, 1696, 1676, 1600, 1547, 1254.

HRMS (ESI-TOF): for C<sub>26</sub>H<sub>38</sub>N<sub>4</sub>NaO<sub>7</sub>S (M + Na)<sup>+</sup>: *calcd.*: 573.2353; *found*: 573.2347.



**(3-((2,13-Dioxo-17-((4*S*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)-6,9-dioxa-3,12-diazahepta-2,6-dienyl)oxy)phenyl)boronic acid (4-1c):**

<sup>1</sup>H NMR δ/ppm: (500 MHz, CD<sub>3</sub>OD) 7.45 – 7.34 (comp m, 2H), 7.30 (app t, *J* = 7.7 Hz, 1H), 7.05 (d, *J* = 8.3 Hz, 1H), 4.54 (s, 2H), 4.49 (dd, *J* = 8.0, 4.7 Hz, 1H), 4.29 (dd, *J* = 7.9, 4.4 Hz, 1H), 3.63 – 3.56 (comp m, 6H), 3.53 (t, *J* = 5.5 Hz, 2H), 3.49 (t, *J* = 5.5 Hz, 2H), 3.37 – 3.34 (m, 2H), 3.22 – 3.15 (m, 1H), 2.92 (dd, *J* = 12.8, 5.0 Hz, 1H), 2.70 (d, *J* = 12.8 Hz, 1H), 2.20 (t, *J* = 7.4 Hz, 2H), 1.76 – 1.52 (comp m, 4H), 1.41 (p, *J* = 7.5 Hz, 2H).

<sup>13</sup>C NMR δ/ppm: (126 MHz, CD<sub>3</sub>OD) 176.4, 171.6, 166.1, 158.5, 130.1, 128.5, 120.6, 118.1, 71.21, 71.17, 70.5, 70.4, 68.2, 63.3, 61.6, 56.9, 41.0, 40.2, 39.9, 36.7, 29.7, 29.4, 26.8.

<sup>11</sup>B NMR δ/ppm: (128 MHz, CD<sub>3</sub>OD) 28.1.

IR (Microscope, cm<sup>-1</sup>): 3307, 3083, 2927, 2866, 1693, 1551, 1429.

HRMS (ESI-TOF): for C<sub>24</sub>H<sub>36</sub>[<sup>11</sup>B]N<sub>4</sub>O<sub>8</sub>S (M – H)<sup>-</sup>: *calcd.*: 551.2352; *found*: 551.2338.

#### 4.4.3 Phage amplification and preparation for sequencing

The eluted phage samples and their corresponding phage samples before capture were amplified separately by adding the eluates (230 μL) into 6 ml of ER2738 culture (1:100 dilution of overnight culture). The phage and bacterial mixtures were incubated for 6.5 h at 37 °C with vigorous shaking. The cultures were centrifuged (15 min, 4700 rpm) at 4 °C to pellet the bacterial cells. An aliquot (~1.5 ml) of the amplified phage from selection of Ser-X<sub>7</sub> library against the reagents **4-1a/4-1b/4-1c** was transferred into a fresh tube, and the phage ssDNA was isolated using QIAprep spin M13 kit (Qiagen, #27704) according to manufacturer's instructions. Then, the ssDNA was converted to Illumina compatible dsDNA amplicon by PCR. Briefly, the ssDNA (~50–100 ng) was combined with 1x Phusion® buffer, 200 μM dNTPs (each), 0.5 μM forward and reverse primers, and one unit Phusion® High-Fidelity DNA Polymerase in a total volume of 50 μL. The temperature cycling protocol for PCR was as follows: 1) hold at 65 °C; 2) 98 °C for 3 min; 3) 98 °C for 10 s; 4) 50 °C for 20 s; 5) 72 °C for 30 s; 6) repeat 3-5 for 10 times; 7) 98 °C for 10 s; 8) 72 °C for 30 s; 9) repeat 7-8 for 20 times; 10) 72 °C for 5 min; 11) hold at 4

°C. The resulting amplicons were characterized by gel electrophoresis using 2% agarose gel, pooled together (20 ng per sample). Sequencing was performed using the Illumina NextSeq platform (Molecular Biology Service Unit, University of Alberta).

#### 4.4.4 Synthesis of peptides on solid support

Peptides were synthesized using Prelude X Instrument (Protein Technologies, Inc.) Briefly, Rink Amide AM resin (200 mg, 0.91 mmol/g, 0.18 mmol) was transferred into a reaction vessel, CH<sub>2</sub>Cl<sub>2</sub> (5 ml) was added to the dried resin for swelling, and after 10 min, the solvent was drained. The resin was washed with DMF (2 x 5 ml) and then deprotected with 20% (v/v) piperidine in DMF (5 ml) for 3 min. The deprotection was repeated for another 3 min using fresh 20% (v/v) piperidine in DMF (5 ml). The resin was washed with DMF (4 ml) for 1 min, followed by a CH<sub>2</sub>Cl<sub>2</sub> wash (4 x 4 ml). Fmoc-protected amino acid (1 mmol) dissolved in DMF (5 ml) and HBTU (1 mmol) dissolved in DMF (3 ml) were added to the resin, followed by dipea (2 ml, 2 mmol). After 15 min of agitation with nitrogen and heating at 50 °C, the reagents were removed, and the resin was washed with DMF (6 x 4 ml). The Fmoc-deprotection, amide coupling and washing steps were repeated consecutively to elongate the peptide sequence. After the final Fmoc-deprotection, the resin was washed with DMF (6 x 4 ml), followed by addition of CH<sub>2</sub>Cl<sub>2</sub> (5 x 4 ml). The resin was drained and dried in a reaction vessel and then transferred to Poly-Prep® chromatography column. A cleavage cocktail (2 ml) containing TFA/H<sub>2</sub>O/phenol/TIPS [85/5/5/5 (v/v/w/v)] was added to the dried resin. The column was left on a rocker for 4 h to cleave the peptide. The flow-through from the column was collected, and the resin was rinsed with TFA (1 ml). The combined cleavage mixture was concentrated *via* a nitrogen purge and added dropwise to cold diethyl ether (20 ml) in a centrifuge tube. The mixture was incubated on ice for 10 min. The precipitates were centrifuged for 5 min at 3000 rpm. The supernatant was decanted, and the precipitates were washed with cold diethyl ether (10 ml). The centrifugation and washing steps were repeated for another two cycles. The precipitates were air-dried, and a typical yield was 50-100 mg (crude). Crude peptide (50 mg) was dissolved in MQ water (0.5 ml), and if the peptide did not dissolve, ACN was added in small aliquots (0.05 ml) until the solution appeared clear. The solution was injected into a semi-preparative RP-HPLC system with a C18 column using MQ water (0.1% (v/v) TFA) and ACN (0.1% (v/v) TFA) at a flow rate of 12 ml/min. The fractions corresponding to the main peak were collected and the ACN was removed by evaporation under reduced pressure. The aqueous solution was

lyophilized to yield the peptide as a white fluffy powder (10–50 mg).

RP-HPLC were performed on a Waters HPLC system equipped with a Waters 1525 EF binary pump, a Waters FlexInject manual injector (dual mode) and a Waters 2489 tunable UV detector. A SymmetryPrep™ C18 semi-preparative column (19 x 50 mm, particle size 5 µm, pore size 100 Å) and XBridge BEH Amide OBD Prep column (19 x 250 mm, particle size 5 µm, pore size 130 Å) were used for purifications of peptides and arabinose glycopeptides at a typical flow rate of 12 ml/min. Characterization of peptides were performed by UPLC-MS and high resolution ESI-MS.

#### **4.4.5 Low-resolution ESI-MS flow injection analysis of the reactions between peptides 1–12 and 2-ABBA/acetophenone/phenylboronic acid**

ESI-MS spectra were acquired by flow injection analysis using a single quadrupole Agilent 1100 MSD (Santa Clara, CA, USA), equipped with an ESI source and an Agilent 1100 Series isocratic pump. The carrier solvent used for flow injection analysis consisted of H<sub>2</sub>O:ACN containing 10 mM ammonium acetate buffer (20:80 v/v%). Mass spectra were acquired using a fast polarity switching mode between 100–1500 Da. Data analysis was performed using the Agilent MassHunter Qualitative Analysis software package version B.07.00 SP1.

Equivalent volumes of peptide 1–12 (50 µl, 2.0 mM in 10 mM ammonium acetate pH ~7) and reagents 2-ABBA/acetophenone/phenylboronic acid (50 µl, 2.0 mM in 10 mM ammonium acetate pH ~7) were mixed in an HPLC vial to reach a final concentration of 1.0 mM. After incubation for 24 h at room temperature, the reaction mixtures were monitored with low resolution ESI-MS flow injection to detect and monitor the desired conjugation.

#### **4.4.6 High-resolution ESI-MS analysis of the reactions between peptides 7–12 and 2-ABBA/acetophenone/phenylboronic acid**

HR-ESI positive ion mode was performed with the Agilent Technologies 6220 Accurate Mass oaTOF Dual ESI using ACN/water (1:1) as the carrier solvent and introducing the samples as direct flow injections. Data acquisition and processing were performed with Agilent Mass Hunter Workstation Data Acquisition and Mass Hunter Qualitative Data Analysis.

Equivalent volumes of peptide 7–12 (50 µl, 2.0 mM in 10 mM ammonium acetate pH ~7) and reagents 2-ABBA/acetophenone/phenylboronic acid (50 µl, 2.0 mM in 10 mM ammonium

acetate pH ~7) were mixed in an eppendorf tube to reach a final concentration of 1.0 mM. After incubation for 24 h at room temperature, the reaction mixtures were monitored with high resolution ESI-MS to detect and monitor the desired conjugation.

#### 4.4.7 <sup>1</sup>H NMR analysis of the reactions between peptides 5/6/11 and 2-ABBA

Equivalent volumes of peptide 5/6/11 (300 µl, 5.0 mM in 25 mM D<sub>2</sub>O phosphate buffer pH ~7) and 2-ABBA (300 µl, 5.0 mM in 25 mM D<sub>2</sub>O phosphate buffer pH ~7) were mixed in a NMR tube to reach a final concentration of 2.5 mM. After incubation for 24 h at room temperature, the reaction mixtures were monitored using <sup>1</sup>H NMR analysis to detect and monitor the desired conjugation.

#### 4.5 References

1. Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764–4806.
2. Chen, X.; Wu, Y–W. *Org. Biomol. Chem.* **2016**, *14*, 5417–5439.
3. Lotze, J.; Reinhardt, U.; Seitz, O.; Beck-Sickinger, A. G. *Mol. BioSyst.* **2016**, *12*, 1731–1745.
4. Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269–272.
5. a) Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *J. Am. Chem. Soc.* **2002**, *124*, 6063–6076. b) Hoffmann, C.; Gaietta, G.; Zürn, A.; Adams, S. R.; Terrillon, S.; Ellisman, M. H.; Tsien, R. Y.; Lohse, M. *J. Nat. Protoc.* **2010**, *5*, 1666–1677.
6. Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. *J. Am. Chem. Soc.* **2007**, *129*, 8672–8673.
7. Fu, N.; Xiong, Y.; Squier, T. C. *Bioconjugate chem.* **2013**, *24*, 251–259.
8. Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. *J. Am. Chem. Soc.* **2009**, *131*, 438–439.
9. Hauser, C. T.; Tsien, R. Y. *PNAS* **2007**, *104*, 3693–3697.
10. Ojida, A.; Honda, K.; Shinmi, D.; Kiyonaka, S.; Mori, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2006**, *128*, 10452–10459.
11. Nonaka, H.; Tsukiji, S.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2007**, *129*, 15777–15779.

12. Uchinomiya, S.; Nonaka, H.; Wakayama, S.; Ojida, A.; Hamachi, I. *Chem. commun*, **2013**, *49*, 5022–5024.
13. Howarth, M.; Takao, K.; Hayashi, Y.; Ting, A. Y. *PNAS* **2005**, *102*, 7583–7588.
14. Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99–104.
15. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotechnol.* **2007**, *25*, 1483–1487.
16. Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. *PNAS* **2009**, *106*, 3000–3005.
17. Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708.
18. Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernández-Suárez, M.; Puthenveetil, S.; Ting, A. Y. *PNAS* **2010**, *107*, 10914–10919.
19. Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. *PNAS* **2005**, *102*, 15815–15820.
20. Tanaka, F.; Fuller, R.; Asawapornmongkol, L.; Warsinke, A.; Gobuty, S. *Bioconjugate Chem.* **2007**, *18*, 1318–1324.
21. Eldridge, G. M.; Weiss, G. A. *Bioconjugate Chem.* **2011**, *22*, 2143–2153.
22. Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. *Nat. Chem.* **2015**, 1–9.
23. Ramil, C. P.; An, P.; Yu, Z.; Lin, Q. *J. Am. Chem. Soc.* **2016**, *138*, 5499–5502.
24. Akgun, B.; Li, C.; Hao, Y.; Lambkin, G.; Derda, R.; Hall, D. G. *J. Am. Chem. Soc.* **2017**, *139*, 14285–14291.
25. Chapin, B. M.; Metola, P.; Lynch, V. M.; Stanton, J. F.; James, T. D.; Anslyn, E. V. *J. Org. Chem.* **2016**, *81*, 8319–8330.
26. Ng, S.; Jafari, M. R.; Derda, R. *ACS Chem. Biol.* **2012**, *7*, 123–138.
27. Ng, S.; Lin, E.; Kitov, P. I.; Tjhung, K. F.; Gerlits, O. O.; Deng, L.; asper, B.; Sood, A.; Paschal, B. M.; Zhang, P.; Ling, C-C.; Klassen, J. S.; Noren, C. J.; Mahal, L. K.; Woods, R. J.; Coates, L.; Derda, R. *J. Am. Chem. Soc.* **2015**, *137*, 5248–5251.
28. a) Smith, G. P. *Science*. **1985**, *228*, 1315–1317. b) Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391–410. c) Rentero Rebollo, I.; Heinis, C. *Methods* **2013**, *60*, 46–54. d) Heinis, C.; Winter, G. *Curr. Opin. Chem. Biol.* **2015**, *26*, 89–98.

29. Lim, R. K. V.; Li, N.; Ramil, C. P.; Lin, Q. *ACS Chem. Biol.* **2014**, *9*, 2139–2148.
30. Yang, J.; Li, K.; Hou, J-T.; Li, L-L.; Lu, C-Y.; Xie, Y-M.; Wang, X.; Yu, X-Q. *ACS Sens.* **2016**, *1*, 166–172.
31. Goswami, L. N.; Houston, Z. H.; Sarma, S. J.; Jalisatgi, S. S.; Frederick Hawthorne, M. F. *Org. Biomol. Chem.* **2013**, *11*, 1116–1126.
32. Sella, E.; Shabat, D. *Org. Biomol. Chem.* **2013**, *11*, 5074–5078.
33. Stress, C. J.; Schmidt, P. J.; Gillingham, D. G. *Org. Biomol. Chem.* **2016**, *14*, 5529–5533.
34. Presolski, S. I.; Hong, V. P.; Finn, M. G. *Curr. Protoc. Chem. Biol.* **2011**, *3*, 153–162.
35. Jin, S.; Choudhary, G.; Cheng, Y.; Dai, C.; Li, M.; Wang, B. *Chem. Commun.* **2009**, *0*, 5251–5253.
36. Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyo, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125*, 2416–2425.
37. a) Matochko, W. L.; Chu, K.; Jin, B.; Lee, S. W.; Whitesides, G. M.; Derda, R. *Methods.* **2012**, *58*, 47–55. b) Matochko, W.; Li, S. C.; Tang, S. K. Y.; Derda, R. *Nuc. Acid. Res.* **2014**, *42*, 1784–1798.
38. Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R. *J. Am. Chem. Soc.* **2016**, *138*, 32-35.
39. a) Hall, D. G. *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials (Volume 1 and 2), Second Edition* (Ed: Hall, D. G.), Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, p. 19–20. b) Stolowitz, M. L.; Ahlem, C.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Li, G.; Lund, K. P.; Torkelson, S. M. Wiley, J. P. *Bioconjugate Chem.* **2001**, *12*, 229–239. c) Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.
40. Kaiser, F.; Schwink, L.; Velder, J.; Schmalz, H-G. *J. Org. Chem.* **2002**, *67*, 9248–9256.
41. Otremba, T.; Ravoo, B. J. *Tetrahedron* **2017**, *73*, 4972–4978.

## 5 Chapter 5. Conclusions

### 5.1 Conclusions and future perspectives

In the last two decades, bioorthogonal chemistry has become a powerful tool to investigate the roles of proteins of interest (POIs), to install post-translational modifications of proteins and to construct antibody drug conjugates.<sup>1-8</sup> ‘Click’ chemistry,<sup>9,10</sup> a class of spontaneous bond-forming processes, is central to the development of bioorthogonal reactions because it exhibits fast reactivity in aqueous environment at low concentrations (< 100  $\mu$ M), with high yields and selectivity without any side products.<sup>1,11,12</sup> Even the most common ‘click’ bioorthogonal reactions, however, display limitations such as slow rates, use of complex reagents, side-reactions or lack of mutual orthogonality due to possible cross reactivity.<sup>2,4,13</sup> To this end, this thesis described new avenues in developing and applying novel bioorthogonal tools based on a ‘click’ boronic ester formation, which is attractive due to the synthetic accessibility of boronic acids and diols, their low toxicity, and fast kinetics under catalyst-free conditions.

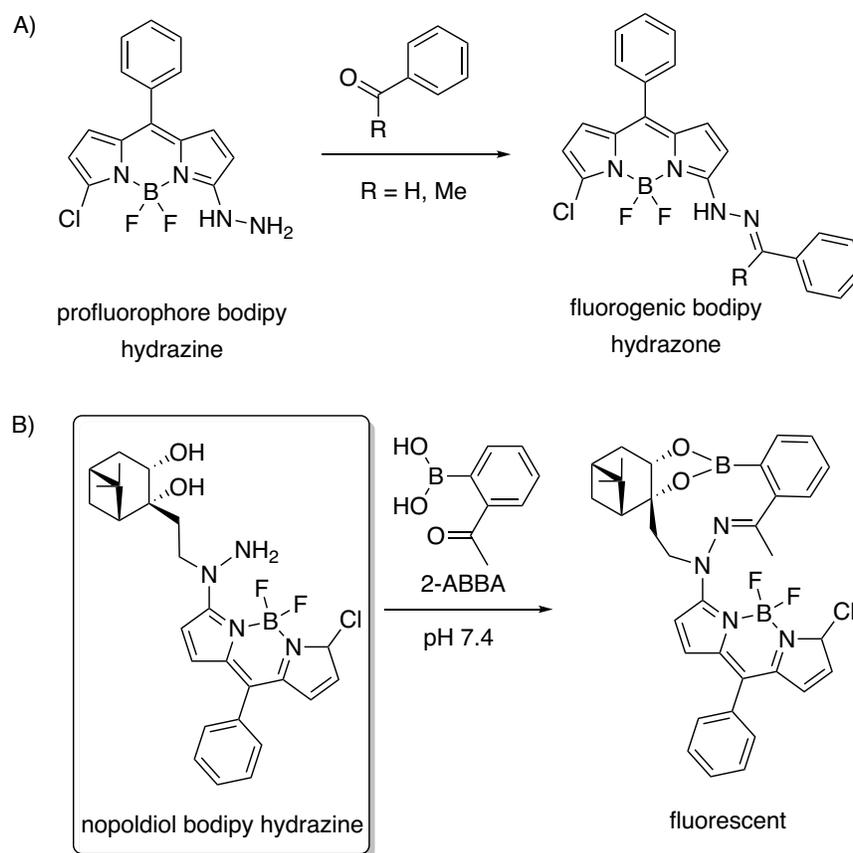
Chapter 1 summarized progress made in the use of boronic acids in bioorthogonal chemistry to enable site-selective labeling of proteins and compared these bioorthogonal reactions with the most popular bioorthogonal reactions. Chapter 2 focused on the development and optimization of a novel boronate bioorthogonal reaction system, which enables a fast and tight ligation of easily synthesized, conjugatable derivatives of nopoldiol and 2-methyl-5-carboxyphenylboronic acid with a submicromolar equilibrium constant  $K_d$ .<sup>14</sup> The kinetics and stability of this system were studied extensively with optimization of both reaction partners. The optimal boronate system was assessed in the presence of biological competitors, such as glucose, fructose and catechol. The system was shown to be compatible with only a 9% decrease in conversion. Also, this robust and efficient boronate conjugation was demonstrated successfully on model proteins, BSA and thioredoxin (Trx). The optimized reagents, nopoldiol and 2-methyl-5-carboxyphenylboronic acid, were shown to be benign towards HEK293 cells.

Even though the reversibility of the described bioorthogonal system in Chapter 2 can be an additional asset when target turnover or time-based profiling are required, this can limit its application in live cell imaging because the boronate adduct might undergo unwanted cleavage during the washing operations for excess reagents. In this regard, Chapter 3 introduced a novel

nopol-boronate system with an improved stability. The optimal synergic system is based on two bifunctional reagents, a thiosemicarbazide-functionalized nopoldiol and an 2-acetylarylboronic acid to provide a fast and irreversible boronate/thiosemicarbazone system.<sup>15</sup> The stability of the resulting boronate/thiosemicarbazone conjugate (50  $\mu\text{M}$ ) was assessed under chemical stress by four separate experiments: 1) diluting the adduct solution from 50  $\mu\text{M}$  to 10  $\mu\text{M}$ ; 2) introducing an excess amount of reactive 2-FBBA competitor (100  $\mu\text{M}$ ) to the existing conjugate; 3) acidifying it to pH 3; or 4) basifying it to pH 9. The boronate/thiosemicarbazone adduct was confirmed to be irreversible under all these different conditions. Not surprisingly, the adduct preserves its integrity in the presence of biological competitors such as fructose, glucose and catechol. Moreover, the thiosemicarbazide-functionalized nopoldiol was found to be unreactive towards a model aldehyde electrophile, structurally similar to protein-bound aldehydes present in human cells; therefore, the system is likely compatible with these cells. After both the thiosemicarbazide-functionalized nopoldiol and 2-acetylarylboronic acids were shown to be benign towards HEK293T cells, this robust and irreversible system was applied in the cell-surface protein labeling of live HEK293T cells. The SNAP-tag approach was employed to biochemically functionalize the extracellular domain of Beta-2 adrenergic receptor with the boronic acid reagent (10  $\mu\text{M}$ ). Then, the thiosemicarbazide-functionalized nopoldiol reagent labeled with fluorescein dye (10–30  $\mu\text{M}$ ) was introduced into the cell culture medium and successful visualization of the cell-surface protein labeling of live HEK293T cells was observed by fluorescence microscopy. Although preliminary attempts to design a fluorogenic system were unsuccessful, further detailed investigation can be worthwhile to advance this synergic system. For instance, fluorogenic bodipy hydrazine compounds were shown to exhibit enhancements in quantum yield (49–54 fold) and higher absorption/emission wavelengths ( $\sim 37$  nm shift) upon hydrazone ligation with aromatic carbonyls<sup>16</sup> (Scheme 5-1A); therefore, a nopoldiol bodipy hydrazine can be synthesized based on the method described in Chapter 3, and its reactivity and fluorogenicity upon ‘double-click’ boronate/hydrazone conjugation with 2-ABBA can then be tested (Scheme 5-1B).

The boronic acid-containing bioorthogonal reactions described in this thesis are most likely to be mutually orthogonal with other bioconjugation chemistries like SPAAC, IEDDA and Staudinger ligation; however, further experimental evidence is essential. Furthermore,

genotoxicity tests of boronic acid reagents used in the described bioorthogonal reactions must be performed in mammalian cells in order to confirm their benign nature.

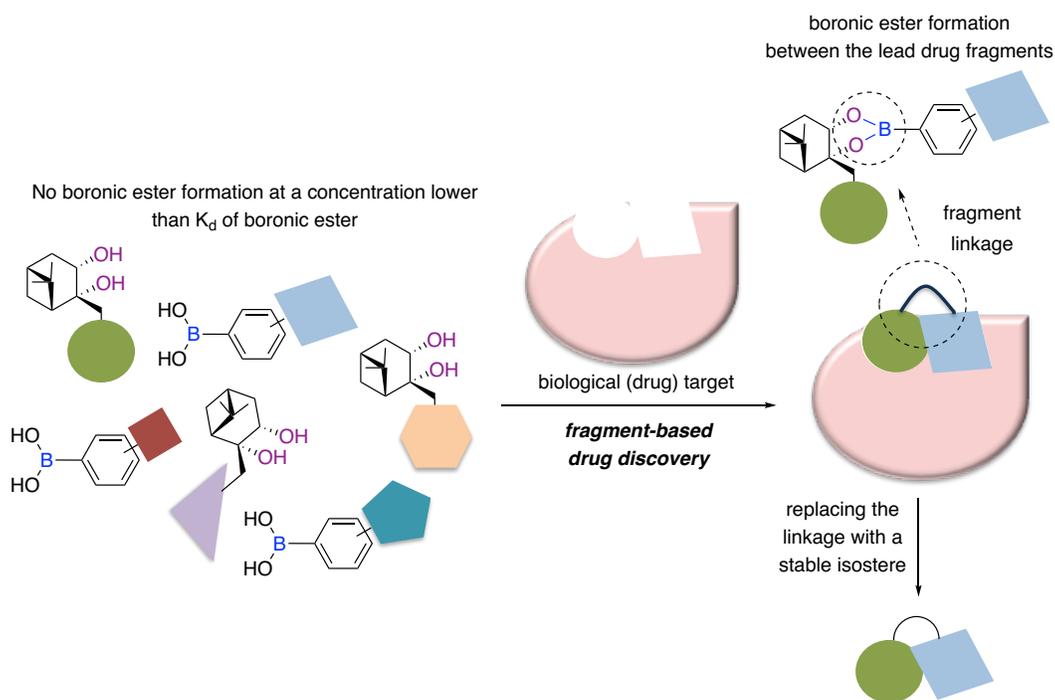


**Scheme 5-1:** Fluorogenic hydrazone ligation. A) Fluorogenic bodipy hydrazone formation. B) A proposed fluorogenic design with the nopoldiol bodipy hydrazine for future investigations.

Chapter 4 described efforts to search for a reactive peptide tag that could also exhibit a high affinity as the thiosemicarbazide-functionalized nopoldiol toward 2-acetylarylboronic acid *via* both imine and boronate formation. Such a system would be more advantageous because one of the two components, a polyol, could be expressed as a short, non-invasive peptide tag. To discover the optimum peptide, a phage display platform with a serine-terminated library of over  $10^8$  heptapeptides was exploited. Selected peptide hits were synthesized; however, their conjugation tests with 2-acetylarylboronic acid resulted in no desired ligation. For further work, even though this hypothetical system is very challenging, the selection over two or three rounds could be repeated using more extensive washing conditions and different phage libraries enriched with lysine, serine and threonine in order to discover such a peptide tag. In addition, the

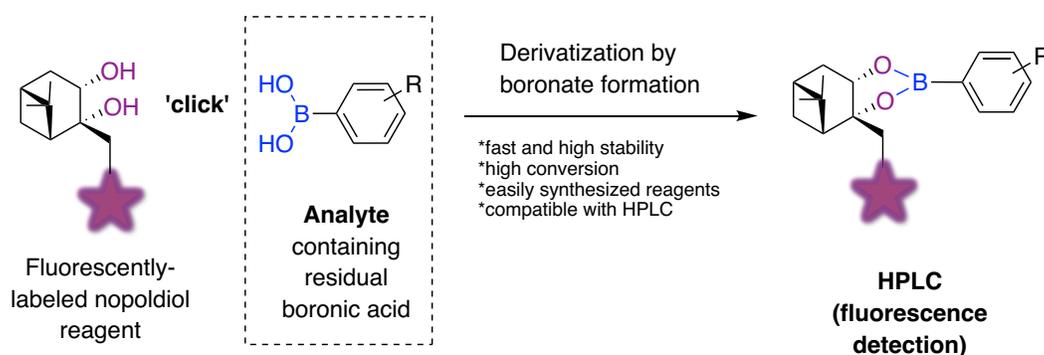
selection could be repeated with a more reactive biotinylated 2-acetylarylboronic acid, such as 2-acetyl-4-carboxyphenylboronic acid derivative, which was shown in Chapter 3, to exhibit a higher reactivity toward thiosemicarbazide-functionalized nopoldiol.

Apart from exploiting the ‘click’ boronic ester system described in Chapter 2 as a bioorthogonal reaction, one could also envision its use in fragment-based drug discovery, which is a strategy for discovering lead compounds.<sup>17</sup> Small chemical fragments derivatized with a nopoldiol unit or an arylboronic acid, which alone exhibit low binding affinities towards the biological target, can produce in the enzyme active site a lead compound that contains a boronic ester conjugation as a fragment linkage with a stronger affinity (Figure 5-1). Its reversibility is an additional asset because when this strategy is performed at a concentration slightly lower than the  $K_d$  of the corresponding boronic ester, the boronic ester formation would be expected only between two fragments that show the highest affinity towards the biological target. In other words, only the lead compounds can form a boronic ester in the presence of a biological target, which serves as a template to ‘click’ two optimal fragments together. Next, the lead compounds can be purified and identified with the techniques such as Western blot and Mass, and then the reversible boronate linkage can be replaced with a stable isostere.



**Figure 5-1:** Application of the ‘click’ boronic ester system in fragment-based drug discovery.

This ‘click’ boronic ester system can also be exploited to analyze and even remove traces of residual, potentially genotoxic boronic acids from pharmaceuticals made with boronic acid reagents in late-stage Suzuki-Miyaura cross-coupling or other reactions. The recent reports<sup>18,19</sup> highlight the potential genotoxicity of boronic acids, and this new information has created a concern in the production of pharmaceuticals. The nopoldiol based ‘click’ boronic ester system might be helpful in detecting, analyzing and quantifying trace amounts of boronic acids (< 100  $\mu$ M), which is currently challenging with existing methods. In this regard, an easy and commonly applied HPLC method with spectrofluorimetric detection can be developed using our system; thus, quantifying even small amount of boronic acids can be possible (Scheme 5-2).



**Scheme 5-2:** A HPLC method with spectrofluorimetric detection using the ‘click’ boronate system to detect, analyze and quantify trace amounts of boronic acids in pharmaceuticals.

Overall, further explorations of the utility of these boronic-acid containing bioorthogonal reactions would hopefully present new avenues in the service of site-selective protein labeling, especially in understanding complicated biological processes and the roles of POIs. Such understanding is essential in discovering cures for dreadful human diseases.

## 5.2 References

1. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
2. Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16–20.
3. Patterson, D. M.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2015**, *28*, 141–149.
4. Ramil, C. P.; Lin, Q. *Chem. Commun. (Camb.)* **2013**, *49*, 11007–11022.
5. Zheng, M.; Zheng, L.; Zhang, P.; Li, J.; Zhang, Y. *Molecules* **2015**, *20*, 3190–3205.
6. Fan, X.; Li, J.; Chen, P. R. *Natl. Sci. Rev.* **2017**, 1–3.

7. a) Yang, Y.-Y.; Ascano, J. M.; Hang, H. C. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641. b) Heal, W. P.; Jovanovic, B.; Bessin, S.; Wright, M. H.; Magee, A. I.; Tate, E. W. *Chem. Commun.* **2011**, *47*, 4081–4083.
8. Selected examples: a) Axupa, J. Y.; Bajjuri, K. M.; Ritland, M.; Hutchins, B. M.; Kim, C. H.; Kazane, S. A.; Halder, R.; Forsyth, J. S.; Santidrian, A. F.; Stafin, K.; Lu, Y.; Tran, H.; Seller, A. J.; Biroc, S. L.; Szydluk, A.; Pinkstaff, J. K.; Tian, F.; Sinha, S. C.; Felding-Habermann, B.; Smider, V. V.; Schultz, P. G. *PNAS* **2012**, *109*, 16101–16106. b) Zimmerman, E. S.; Heibeck, T. H.; Gill, A.; Li, X.; Murray, C. J.; Madlansacay, M. R.; Tran, C.; Uter, N. T.; Yin, G.; Rivers, P. J.; Yam, A. Y.; Wang, W. D.; Steiner, A. R.; Bajad, S. U.; Penta, K.; Yang, W.; Hallam, T. J.; Thanos, C. D.; Sato, A. K. *Bioconjugate Chem.* **2014**, *25*, 351–361.
9. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
10. Baskin, J. M.; Bertozzi, C. R. *QSAR Comb. Sci.* **2007**, *26*, 1211–1219.
11. Best, M. D. *Biochemistry* **2009**, *48*, 6571–6584.
12. McKay, C. S.; Finn, M. G. *Chem. Biol.* **2014**, *21*, 1075–1101.
13. Shih, H.-W.; Kamber, D. N.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2014**, *21*, 103–111.
14. Akgun, B.; Hall, D. G. *Angew. Chem. Int. Ed.* **2016**, *55*, 3909–3913.
15. Akgun, B.; Li, C.; Hao, Y.; Lambkin, G.; Derda, R.; Hall, D. G. *J. Am. Chem. Soc.* **2017**, *139*, 14285–14291.
16. a) Dilek, O.; Bane, S. L. *Tetrahedron Lett.* **2008**, *49*, 1413–1416. b) Dilek, O.; Bane, S. L. *J. Fluoresc.* **2011**, *21*, 347–354. c) Key, J. A.; Li, C.; Cairo, C. W. *Bioconjugate Chem.* **2012**, *23*, 363–371.
17. Wanner, J.; Romashko, D.; Werner, D. S.; May, E. W.; Peng, Y.; Schulz, R.; Foreman, K. W.; Russo, S.; Arnold, L. D.; Pingle, M.; Bergstrom, D. E.; Barany, F.; Thomson, S. *PLoS One*, **2015**, *10*, e0121793
18. O'Donovan, M. R.; Mee, C. D.; Fenner, S.; Teasdale, A.; Phillips, D. H. *Mutat. Res.* **2011**, *724*, 1–6.
19. Walmsley, R. M. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2015**, *777*, 68–72.

## Bibliography

1. Chen, X.; Wu, Y.-W. *Org. Biomol. Chem.* **2016**, *14*, 5417–5439.
2. Boutureira, O.; Bernardes, G. J. L. *Chem. Rev.* **2015**, *115*, 2174–2195.
3. Spicer, C. D.; Davis, B. G. *Nat. Commun.* **2014**, *5*, 4740.
4. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
5. Ramil, C. P.; Lin, Q. *Chem. Commun. (Camb.)* **2013**, *49*, 11007–11022.
6. Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764–4806.
7. Zhang, G.; Zheng, S.; Liu, H.; Chen, P. R. *Chem. Soc. Rev.* **2015**, *44*, 3405–3417.
8. Hang, H. C.; Wilson, J. P.; Charron, G. *Acc. Chem. Res.* **2011**, *44*, 699–708.
9. Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F. P. J. T.; van Hest, J. C. M.; Lefeber, D. J.; Friedl, P.; van Delft, F. L. *Angew. Chem. Int. Ed.* **2010**, *49*, 9422–9425.
10. Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. *Nat. Chem.* **2012**, *4*, 2012, 298–304.
11. Peng, T.; Hang, H. C. *J. Am. Chem. Soc.* **2016**, *138*, 14423–14433.
12. Baskin, J. M.; Bertozzi, C. R. *QSAR Comb. Sci.* **2007**, *26*, 1211–1219.
13. Krall, N.; da Cruz, F. P.; Boutureira, O.; Bernardes, G. J. L. *Nat. Chem.* **2016**, *8*, 103–113.
14. Yang, Y.-Y.; Ascano, J. M.; Hang, H. C. *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641.
15. Heal, W. P.; Jovanovic, B.; Bessin, S.; Wright, M. H.; Magee, A. I.; Tate, E. W. *Chem. Commun.* **2011**, *47*, 4081–4083.
16. Beck, A.; Goetsch, L.; Dumontet, C.; Corvaia, N. *NPG* **2017**, *16*, 315–337.
17. van Vught, R.; Pieters, R. J.; Breukink, E. *Comput. Struct. Biotechnol. J.* **2014**, *9*, 1–13.
18. Axupa, J. Y.; Bajjuri, K. M.; Ritland, M.; Hutchins, B. M.; Kim, C. H.; Kazane, S. A.; Halder, R.; Forsyth, J. S.; Santidrian, A. F.; Stafin, K.; Lu, Y.; Tran, H.; Seller, A. J.; Biroc, S. L.; Szydluk, A.; Pinkstaff, J. K.; Tian, F.; Sinha, S. C.; Felding-Habermann, B.; Smider, V. V.; Schultz, P. G. *PNAS* **2012**, *109*, 16101–16106.
19. Zimmerman, E. S.; Heibeck, T. H.; Gill, A.; Li, X.; Murray, C. J.; Madlansacay, M. R.; Tran, C.; Uter, N. T.; Yin, G.; Rivers, P. J.; Yam, A. Y.; Wang, W. D.; Steiner, A. R.; Bajad, S. U.; Penta, K.; Yang, W.; Hallam, T. J.; Thanos, C. D.; Sato, A. K. *Bioconjugate*

- Chem.* **2014**, *25*, 351–361.
20. Baslé, E.; Joubert, N.; Pucheault, M. *Chem. Biol.* **2010**, *17*, 213–227.
  21. Malins, L. R.; Payne, R. J. *Curr. Opin. Chem. Biol.* **2014**, *22*, 70–78.
  22. Lotze, J.; Reinhardt, U.; Seitz, O.; Beck-Sickinger, A. G. *Mol. BioSyst.* **2016**, *12*, 1731–1745.
  23. Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802–805.
  24. Heim, R.; Cubitt, A. B.; Tsien, R. Y. *Nature* **1995**, *373*, 663–664.
  25. Chudakov, D. M.; Matz, M. V.; Lukyanov, S.; Lukyanov, K. A. *Physiol. Rev.* **2010**, *90*, 1103–1163.
  26. Rizzo, M. A.; Davidson, M. W.; Piston, D. W. *Cold Spring Harb. Protoc.* **2009**, *12*, 1–22.
  27. Jing, C.; Cornish, V. W. *Acc. Chem. Res.* **2011**, *44*, 784–792.
  28. Adumeau, P.; Sharma, S. K.; Brent, C.; Zeglis, B. M. *Mol Imaging Biol.* **2016**, *18*, 153–165.
  29. George, N.; Pick, H.; Vogel, H.; Johnsson, N.; Johnsson, K. *J. Am. Chem. Soc.* **2004**, *126*, 8896–8897.
  30. Chen, I.; Howarth, M.; Lin, W. Y.; Ting, A. Y. *Nature Methods* **2005**, *2*, 99–104.
  31. Griffin, B. A.; Adams, S. R.; Tsien, R. Y. *Science* **1998**, *281*, 269–272.
  32. Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y. *J. Am. Chem. Soc.* **2002**, *124*, 6063–6076.
  33. Cao, H.; Xiong, Y.; Wang, T.; Chen, B.; Squier, T. C.; Mayer, M. U. *J. Am. Chem. Soc.* **2007**, *129*, 8672–8673.
  34. Jung, D.; Min, K.; Jung, J.; Jang, W.; Kwon, Y. *Mol. BioSyst.* **2013**, *9*, 862–872.
  35. Tanaka, F.; Fuller, R.; Asawapornmongkol, L.; Warsinke, A.; Gobuty, S. *Bioconjugate Chem.* **2007**, *18*, 1318–1324.
  36. Eldridge, G. M.; Weiss, G. A. *Bioconjugate Chem.* **2011**, *22*, 2143–2153.
  37. Ramil, C. P.; An, P.; Yu, Z.; Lin, Q. *J. Am. Chem. Soc.* **2016**, *138*, 5499–5502.
  38. Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. *Nat. Chem.* **2016**, *8*, 120–128.
  39. Lang, K.; Chin, J. W. *ACS Chem. Biol.* **2014**, *9*, 16–20.
  40. Zheng, M.; Zheng, L.; Zhang, P.; Li, J.; Zhang, Y. *Molecules* **2015**, *20*, 3190–3205.

41. Fan, X.; Li, J.; Chen, P. R. *Natl. Sci. Rev.* **2017**, *0*, 1–3.
42. Chen, Y.-X.; Triola, G.; Waldmann, H. *Acc. Chem. Res.* **2011**, *44*, 762–773.
43. Dozier, J. K.; Distefano, M. D. *Int. J. Mol. Sci.* **2015**, *16*, 25831–25864.
44. Gong, Y.; Pan, L. *Tetrahedron Lett.* **2015**, *56*, 2123–2132.
45. Kim, C. H.; Axup, J. Y.; Schultz, P. G. *Curr. Opin. Chem. Biol.* **2013**, *17*, 412–419.
46. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
47. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125–1128.
48. Rashidian, M.; Mahmoodi, M. M.; Shah, R.; Dozier, J. K.; Wagner, C. R.; Distefano, M. D. *Bioconjugate Chem.* **2013**, *24*, 333–342.
49. Kool, E. T.; Park, D.-H.; Crisalli, P. *J. Am. Chem. Soc.* **2013**, *135*, 17663–17666.
50. Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21.
51. Matthews, M. L.; He, L.; Horning, B. D.; Olson, E. J.; Correia, B. E.; Yates, J. R.; Dawson, P. E.; Cravatt, B. F. *Nat. Chem.* **2017**, *9*, 234–243.
52. van Berkel, S. S.; van Eldijk, M. B.; van Hest, J. C. M. *Angew. Chem. Int. Ed.* **2011**, *50*, 8806–8827.
53. Saxon, E.; Armstrong, J. I.; Bertozzi, C. R. *Org. Lett.* **2000**, *2*, 2141–2143.
54. Shah, L.; Laughlin, S. T.; Carrico, I. S. *J. Am. Chem. Soc.* **2016**, *138*, 5186–5189.
55. Chang, P. V.; Prescher, J. A.; Sletten, E. M.; Baskin, J. M.; Miller, I. A.; Agard, N. J.; Lo, A.; Bertozzi, C. R. *PNAS* **2010**, *107*, 1821–1826.
56. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644–648.
57. Yang, M.; Li, J.; Chen, P. R. *Chem. Soc. Rev.* **2014**, *43*, 6511–6526.
58. Aioub, A. G.; Dahora, L.; Gamble, K.; Finn, M. G. *Bioconjugate Chem.* **2017**, *28*, 1693–1701.
59. Uttamapinant, C.; Tangpeerachaikul, A.; Grecian, S.; Clarke, S.; Singh, U.; Slade, P.; Gee, K. R.; Ting, A. Y. *Angew. Chem. Int. Ed.* **2012**, *51*, 5852–5856.
60. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
61. Dieterich, D. C.; Hodas, J. J. L.; Gouzer, G.; Shadrin, I. Y.; Ngo, J. T.; Triller, A.; Tirrell, D. A.; Schuman, E. M. *Nat. Neurosci.* **2010**, *13*, 897–905.
62. Beatty, K. E.; Fisk, J. D.; Smart, B. P.; Lu, Y. Y.; Szychowski, J.; Hangauer, M. J.; Baskin, J. M.; Bertozzi, C. R.; Tirrell, D. A. *ChemBioChem* **2010**, *11*, 2092–2095.

63. Fernández-Suárez, M.; Baruah, H.; Martínez-Hernández, L.; Xie, K. T.; Baskin, J. M.; Bertozzi, C. R.; Ting, A. Y. *Nat. Biotechnol.* **2007**, *25*, 1483–1487.
64. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *PNAS* **2007**, *104*, 16793–16797.
65. Conte, Lo, M.; Staderini, S.; Marra, A.; Sanchez-Navarro, M.; Davis, B. G.; Dondoni, A. *Chem. Commun.* **2011**, *47*, 11086–11088.
66. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
67. Wu, H.; Devaraj, N. K. *Top. Curr. Chem.* **2015**, *374*, 1–22.
68. Devaraj, N. K.; Hilderbrand, S.; Upadhyay, R.; Mazitschek, R.; Weissleder, R. *Angew. Chem. Int. Ed.* **2010**, *49*, 2869–2872.
69. Carlson, J. C. T.; Meimetis, L. G.; Hilderbrand, S. A.; Weissleder, R. *Angew. Chem. Int. Ed.* **2013**, *52*, 6917–6920.
70. Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *J. Am. Chem. Soc.* **2012**, *134*, 10317–10320.
71. Liu, D. S.; Tangpeerachaikul, A.; Selvaraj, R.; Taylor, M. T.; Fox, J. M.; Ting, A. Y. *J. Am. Chem. Soc.* **2012**, *134*, 792–795.
72. Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. *ACS Chem. Biol.* **2014**, *9*, 592–605.
73. Lorand, J. P.; Edwards, J. O. *J. Org. Chem.* **1959**, *24*, 769–774.
74. Miyaura, N.; Suzuki, A. *Chem. Rev.* **1995**, *95*, 2457–2483.
75. Zheng, H.; Lejkowski, M.; Hall, D.G. *Tetrahedron Lett.* **2013**, *54*, 91–94
76. Gernigon, N.; Al-Zoubi, R. M.; Hall, D. G. *J. Org. Chem.* **2012**, *77*, 8386–8400.
77. Hall, D. G. *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials (Volume 1 and 2), Second Edition* (Ed: Hall, D. G.), Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, p. 213–590.
78. Sun, X.; Zhai, W.; Fossey, J. S.; James, T. D. *Chem. Commun.* **2016**, *52*, 3456–3469.
79. Jabbour, A.; Steinberg, D.; Dembitsky, V. M.; Moussaieff, A.; Zaks, A. B.; Srebnik, M. *J. Med. Chem.* **2004**, *47*, 2409–2410.
80. Chu, Y.; Wang, D. Z.; Wang, K.; Liu, Z. R.; Weston, B.; Wang, B. H. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 6307–6309.
81. Wang, X.; Xia, N.; Liu, L. *Int. J. Mol. Sci.* **2013**, *14*, 20890–20912.
82. Lacina, K.; Skládal, P. S.; James, T. D. *Chem. Cent. J.* **2014**, *8*, 1–17.

83. Yan, J.; Springsteen, G.; Deeter, S.; Wang, B. *Tetrahedron* **2004**, *60*, 11205–11209.
84. Cal, P. M. S. D.; Vicente, J. B.; Pires, E.; Coelho, A. V.; Veiros, L. F.; Cordeiro, C.; Gois, P. M. P. *J. Am. Chem. Soc.* **2012**, *134*, 10299–10305.
85. Bandyopadhyay, A.; Gao, J. *Chem. Eur. J.* **2015**, *21*, 14748–14752.
86. Gillingham, D. *Org. Biomol. Chem.* **2016**, *14*, 7606–7609.
87. Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220–8223.
88. Cal, P. M. S. D.; Frade, R. F. M.; Chudasama, V.; Cordeiro, C.; Caddick, S.; Gois, P. M. P. *Chem. Commun. (Camb.)* **2014**, *50*, 5261–5263.
89. Cal, P. M. S. D.; Frade, R. F. M.; Cordeiro, C.; Gois, P. M. P. *Chem. Eur. J.* **2015**, *21*, 8182–8187.
90. Schmidt, P.; Stress, C.; Gillingham, D. *Chem. Sci.* **2015**, *6*, 3329–3333.
91. Dilek, O.; Lei, Z.; Mukherjee, K.; Bane, S. *Chem. Commun.* **2015**, *51*, 16992–16995.
92. Stress, C. J.; Schmidt, P. J.; Gillingham, D. G. *Org. Biomol. Chem.* **2016**, *14*, 5529–5533.
93. Bandyopadhyay, A.; Cambray, S.; Gao, J. *J. Am. Chem. Soc.* **2017**, *139*, 871–878.
94. Bandyopadhyay, A.; Cambray, S.; Gao, J. *Chem. Sci.* **2016**, *7*, 4589–4593.
95. Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P. *Chem. Sci.* **2016**, *7*, 5052–5058.
96. Ratner, V.; Kahana, E.; Eichler, M.; Haas, E. *Bioconjugate Chem.* **2002**.
97. Botti, P.; Pallin, T. D.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 10018–10024.
98. Meadows, M. K.; Roesner, E. K.; Lynch, V. M.; James, T. D.; Anslyn, E. V. *Org. Lett.* **2017**, *19*, 3179–3182.
99. Chapin, B. M.; Metola, P.; Lynch, V. M.; Stanton, J. F.; James, T. D.; Anslyn, E. V. *J. Org. Chem.* **2016**, *81*, 8319–8330.
100. Dowlut, M.; Hall, D. G. *J. Am. Chem. Soc.* **2006**, *128*, 4226–4227.
101. Ellis, G. A.; Palte, M. J.; Raines, R. T. *J. Am. Chem. Soc.* **2012**, *134*, 3631–3634.
102. Andersen, K. A.; Smith, T. P.; Lomax, J. E.; Raines, R. T. *ACS Chem. Biol.* **2016**, *11*, 319–323.
103. Halo, T. L.; Appelbaum, J.; Hobert, E. M.; Balkin, D. M.; Schepartz, A. *J. Am. Chem. Soc.* **2009**, *131*, 438–439.
104. Stolowitz, M. L.; Ahlem, C.; Hughes, K. A.; Kaiser, R. J.; Kesicki, E. A.; Li, G.; Lund, K.

- P.; Torkelson, S. M.; Wiley, J. P. *Bioconjugate Chem.* **2001**, *12*, 229–239.
105. Bin Y Shin, S.; Almeida, R. D.; Gerona-Navarro, G.; Bracken, C.; Jaffrey, S. R. *Chem. Biol.* **2010**, *17*, 1171–1176.
106. Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. *Tetrahedron Lett.* **2005**, *46*, 3301–3305.
107. Chalker, J. M.; Wood, C. S. C.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, *131*, 16346–16347.
108. Spicer, C. D.; Davis, B. G. *Chem. Commun. (Camb.)* **2011**, *47*, 1698–1700.
109. Spicer, C. D.; Triemer, T.; Davis, B. G. *J. Am. Chem. Soc.* **2012**, *134*, 800–803.
110. Spicer, C. D.; Davis, B. G. *Chem. Commun. (Camb.)* **2013**, *49*, 2747–2749.
111. Gao, Z.; Gouverneur, V.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, *135*, 13612–13615.
112. Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G. *Angew. Chem. Int. Ed.* **2013**, *52*, 3916–3921.
113. Ourailidou, M. E.; van der Meer, J. Y.; Baas, B. J.; Jeronimus Stratingh, M.; Gottumukkala, A. L.; Poelarends, G. J.; Minnaard, A. J.; Dekker, F. J. *ChemBioChem* **2014**, *15*, 209–212.
114. Eising, S.; Lelivelt, F.; Bongers, K. M. *Angew. Chem. Int. Ed.* **2016**, *55*, 12243–12247.
115. Li, J.; Jia, S.; Chen, P. R. *Nat. Chem. Biol.* **2014**, *10*, 1003–1005.
116. Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; Chen, P. R. *Nat. Chem.* **2014**, *6*, 352–361.
117. Völker, T.; Meggers, E. *Curr. Opin. Chem. Biol.* **2015**, *25*, 48–54.
118. Kim, J.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2015**, *54*, 15777–15781.
119. Young, T. S.; Ahmad, I.; Yin, J. A.; Schultz, P. G. *J. Mol. Biol.* **2010**, *395*, 361–374.
120. O'Donovan, M. R.; Mee, C. D.; Fenner, S.; Teasdale, A.; Phillips, D. H. *Mutat. Res.* **2011**, *724*, 1–6.
121. Walmsley, R. M. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2015**, *777*, 68–72.
122. Ciaravino, V.; Plattner, J.; Chanda, S. *Environ. Mol. Mutagen.* **2013**, *54*, 338–346.
123. Matteson, D. S.; Man, H.-W. *J. Org. Chem.* **1996**, *61*, 6047–6051.
124. Roy, C. D.; Brown, H. C. *Monatsh. Chem.* **2007**, *138*, 747–753.
125. Roy, C. D.; Brown, H. C. *Monatsh. Chem.* **2007**, *138*, 879–887.
126. Ng, S.; Jafari, M. R.; Derda, R. *ACS Chem. Biol.* **2012**, *7*, 123–138.
127. Sletten, E. M.; Bertozzi, C. R. *Acc. Chem. Res.* **2011**, *44*, 666–676.
128. Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010**, *39*, 1272–1279.

129. McKay, C. S.; Finn, M. G. *Chem. Biol.* **2014**, *21*, 1075–1101.
130. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
131. Šečkutė, J.; Devaraj, N. K. *Curr. Opin. Chem. Biol.* **2013**, *17*, 761–767.
132. Shih, H-W.; Kamber, D. N.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2014**, *21*, 103–111.
133. Wu, X.; Li, Z.; Chen, X. – X.; Fossey, J. S.; James, T. D.; Jiang, Y.-B. *Chem. Soc. Rev.* **2013**, *42*, 8032–8048.
134. Jin, S.; Cheng, Y.; Reid, S.; Li, M.; Wang, B. *Med. Res. Rev* **2010**, *30*, 171–257.
135. James, T. D.; Shinkai, S. *Topics Curr. Chem.* **2002**, *218*, 159–200.
136. Wang, W.; Gao, X.; Wang, B. *Curr. Org. Chem.* **2002**, *6*, 1285–1317.
137. Liu, H.; Li, Y.; Sun, K.; Fan, J.; Zhang, P.; Meng, J.; Wang, S.; Jiang, L. *J. Am. Chem. Soc.* **2013**, *135*, 7603–7609.
138. Pan, G.; Guo, B.; Ma, Y.; Cui, W.; He, F.; Li, B.; Yang, H.; Shea, K. J. *J. Am. Chem. Soc.* **2014**, *136*, 6203–6206.
139. Bull, S. D.; Davidson, M. G.; Van Den Elsen, J. M. H.; Fossey, J. S.; Jenkins, A. T. A.; Jiang, Y-B.; Kubo, Y.; Marken, F.; Sakurai, K.; Zhao, J.; James, T. D. *Acc. Chem. Res.* **2013**, *46*, 312–326.
140. Lippert, A. R.; Van De Bittner, G. C.; Chang, C. J. *Acc. Chem. Res.* **2011**, *44*, 793–804.
141. Wang, F.; Niu, W.; Guo, J.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 10132–10135.
142. Lin, V.S.; Dickinson, B.C.; Chang, C. J. *Methods Enzymol.* **2013**, *526*, 19–43.
143. Chen, Z-J.; Ren, W.; Wright, Q. E.; Ai, H-W. *J. Am. Chem. Soc.* **2013**, *135*, 14940–14943.
144. Sun, X.; Xu, Q.; Kim, G.; Flower, S. E.; Lowe, J. P.; Yoon, J.; Fossey, J. S.; Qian, X.; Bull, S. D.; James, T. D. *Chem. Sci.* **2014**, *5*, 3368–3373.
145. Liu, C. C.; Mack, A. V.; Brustad, E. M.; Mills, J. H.; Groff, D.; Smider, V. V.; Schultz, P. G. *J. Am. Chem. Soc.* **2009**, *131*, 9616–9617.
146. Babcock, L.; Pizer, R. *Inorg. Chem.* **1980**, *19*, 56–61.
147. Pizer, R.; Tihal, C. *Inorg. Chem.* **1992**, *31*, 3243–3247.
148. Watanabe, E.; Miyamoto, C.; Tanaka, A.; Iizuka, K.; Iwatsuki, S.; Inamo, M.; Takagi, H. D.; Ishihara, K. *Dalton Trans.* **2013**, *42*, 8446–8453.
149. Okamoto, T.; Tanaka, A.; Watanabe, E.; Miyazaki, T.; Sugaya, T.; Iwatsuki, S.; Inamo, M.; Takagi, H. D.; Odani, A.; Ishihara, K. *Eur. J. Inorg. Chem.* **2014**, 2389–2395.

150. Furikado, Y.; Nagahata, T.; Okamoto, T.; Sugaya, T.; Iwatsuki, S.; Inamo, M.; Takagi, H. D.; Odani, A.; Ishihara, K. *Chem. Eur. J.* **2014**, *20*, 13194–13202.
151. Peters, J. A. *Coord. Chem. Rev.* **2014**, *268*, 1–22.
152. Ni, N.; Laughlin, S.; Wang, Y.; Feng, Y.; Zheng, Y.; Wang, B. *Bioorg. Med. Chem.* **2012**, *20*, 2957–2961.
153. Tomsho, J. W.; Benkovic, S. J. *J. Org. Chem.* **2012**, *77*, 2098–2106.
154. Monajemi, H.; Cheah, M. H.; Lee, V. S.; Mohd. Zain, S.; Wan Abdullah, W. A. T. *RSC Adv.* **2014**, *4*, 10505–10513.
155. Baggio, R. F.; Sparks, A. L.; Juo, R.-R.; Arenas, J. E. US Patent 0277143 A1 December 15, 2005.
156. Ahmad, M.; Hamer, J. *J. Chem. Educ.* **1964**, *41*, 249–250.
157. Pollard, T. D.; De La Cruz, E. M. *Mol. Biol. Cell.* **2013**, *24*, 1103–1110.
158. Crouch, S. R.; Holwr, F. J.; Notz, P. K.; Beckwith, P. M. *Appl. Spectrosc. Rev.* **1977**, *13*, 156–259.
159. Thordarson, P. *Chem. Soc. Rev.* **2011**, *40*, 1305–1323.
160. Carlise, J. R.; Kriegel, R. M.; Rees, W. S.; Weck, M. *J. Org. Chem.* **2005**, *70*, 5550–5560.
161. Tius, M. A.; Gu, X-q; Truesdell, J. W.; Savariae, S.; Crooker, P.P. *Synthesis* **1988**, *1*, 36–40.
162. Hayashi, T.; Konishi, M.; Kobori, Y.; Kumada, M.; Higuchi, T.; Hirotsu, K. *J. Am. Chem. Soc.* **1984**, *106*, 158–163.
163. Marshall, J. A.; Johns, B. A. *J. Org. Chem.* **1998**, *63*, 7885–7892.
164. Everson, D. A.; Jones, B. A.; Weix, D. J. *J. Am. Chem. Soc.* **2012**, *134*, 6146–6159.
165. Springsteen, G; Wang, B. *Tetrahedron* **2002**, *58*, 5291–5300.
166. Hall, D. G. *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials (Volume 1 and 2), Second Edition* (Ed: Hall, D. G.), Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, p. 10.
167. Pal, A.; Bérubé, M.; Hall, D. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 1492–1495.
168. Chotana, G. A.; Rak, M. A.; Smith, M. R. *J. Am. Chem. Soc.* **2005**, *127*, 10539–10544.
169. Volkhard, H. “Fluorescence Resonance Energy Transfer.” *Principles of Computational Cell Biology*. Wiley-VCH Verlag GmbH & Co. KGaA, **2008**, p. 202.

170. Clegg, R. M. "Förster resonance energy transfer—FRET: what is it, why do it, and how it's done." *FRET and FLIM Techniques. Laboratory Techniques in Biochemistry and Molecular Biology* (Ed: Gadella; T. W. J.), Elsevier, **2009**, Volume 33, p. 1–57.
171. Johansson, M. K. "Choosing Reporter-Quencher Pairs for Efficient Quenching Through Formation of Intramolecular Dimers." *Methods in Molecular Biology* (Ed: Didenko, V. V.), Humana Press: Totowa, NJ, **2006**, p. 17–29.
172. de Champlain, J.; Farley, L.; Cousineau, D.; van Ameringen, M. R. *Circ. Res.* **1976**, *38*, 109–114.
173. Kawasaki, T.; Akanuma, H.; Yamanouchi, T. *Diabetes Care* **2002**, *25*, 353–357.
174. Borner, M. M.; Schneider, E.; Pirnia, F.; Sartor, O.; Trepel, J. B.; Myers, C. E. *FEBS Lett.* **1994**, *353*, 129–132.
175. Ngamwongsatit, P.; Banada, P. P.; Panbangred, W.; Bhunia, A. K. *J. Microbiol. Methods.* **2008**, *73*, 211–215.
176. Kim, K-M.; Lee, S-B.; Lee, S-H.; Lee, Y-K.; Kim, K-N. *Key Eng. Mater.* **2005**, *11*, 284–286.
177. Mothana, S.; Grassot, J-M.; Hall, D. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 2883–2887.
178. Kwong, F. Y.; Klapars, A.; Buchwald, S. L. *Org. Lett.* **2002**, *4*, 581–584.
179. Patterson, D. M.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2015**, *28*, 141–149.
180. Best, M. D. *Biochemistry* **2009**, *48*, 6571–6584.
181. Akgun, B.; Hall, D. G. *Angew. Chem. Int. Ed.* **2016**, *55*, 3909–3913.
182. Kalia, J.; Raines, R. T. *Angew. Chem. Int. Ed.* **2008**, *47*, 7523–7526.
183. Sander, E. G.; Jencks, W. P. *J. Am. Chem. Soc.* **1968**, *90*, 6154–6162.
184. Il'ina, I. V.; Volcho, K.P.; Salakhutdinov, N. F. *Russ. J. Organ. Chem.* **2008**, *44*, 1–23.
185. Chen, Y.; Tsao, K.; Keillor, J. W. *Can. J. Chem.* **2015**, *93*, 389–398.
186. Shieh, P.; Bertozzi, C. R. *Org. Biomol. Chem.* **2014**, *12*, 9307–9320.
187. Nadler, A.; Schultz, C. *Angew. Chem. Int. Ed.* **2013**, *52*, 2408–2410.
188. Jun, M. E.; Roy, B.; Ahn, K. H. *Chem. Commun.* **2011**, *47*, 7583–7601.
189. Sun, X. L.; Xu, S. Y.; Flower, S. E.; Fossey, J. S.; Qian, X. H.; James, T. D. *Chem. Commun.* **2013**, *49*, 8311–8313.
190. Cao, H.; Diaz, D. I.; DiCesare, N.; Lakowicz, J. R.; Heagy, M. D. *Org. Lett.* **2002**, *4*, 1503–1505.

191. Cao, Z.; Nandhikonda, P.; Heagy, M. D. *J. Org. Chem.* **2009**, *74*, 3544–3546.
192. Dilek, O.; Bane, S. L. *Tetrahedron Lett.* **2008**, *49*, 1413–1416.
193. Dilek, O.; Bane, S. L. *J. Fluoresc.* **2011**, *21*, 347–354.
194. Key, J. A.; Li, C.; Cairo, C. W. *Bioconjugate Chem.* **2012**, *23*, 363–371.
195. Key, J. A.; Koh, S.; Timerghazin, Q. K.; Brown, A.; Cairo, C. W. *Dyes Pigm.* **2009**, *82*, 192–203.
196. Gubitz, G.; Wintersteiger, R.; Frei, R. W. *J. Liq. Chromatogr.* **1984**, *7*, 839–854.
197. Uzu, S.; Kanda, S.; Nakashima, K.; Akiyama, S. *Analyst* **1990**, *115*, 1477–1482.
198. Crandall, J. K.; Crawley, L. C. *Org. Synth.* **1973**, *53*, 17–5.
199. Lakshmi, R.; Bateman, T. D.; McIntosh, M. C. *J. Org. Chem.* **2005**, *70*, 5313–5315.
200. Caron, M.; Sharpless, K. B. *J. Org. Chem.* **1985**, *50*, 1557–1560.
201. Bosch, L. I.; Fyles, T. M.; James, T. D. *Tetrahedron* **2004**, *60*, 11175–11190.
202. Givens, R. S.; Stensrud, K.; Conrad, P. G.; Yousef, A. L.; Perera, C.; Senadheera, S. N.; Heger, D.; Wirz, J. *Can. J. Chem.* **2011**, *89*, 364–384.
203. Elsinghorst, P. W.; Tanarro, C. M. G.; Gütschow, M. *J. Med. Chem.* **2006**, *49*, 7540–7544.
204. James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Nature* **1995**, *374*, 345–347.
205. Sun, X.; James, T. D. *Chem. Rev.* **2015**, *115*, 8001–8037.
206. Kitov, P. I.; Vinals, D. F.; Ng, S.; Tjhung, K. F.; Derda, R. *J. Am. Chem. Soc.* **2014**, *136*, 8149–8152.
207. Keppler, A.; Kindermann, M.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. *Methods* **2004**, *32*, 437–444.
208. Cole, N. B. *Curr. Protoc. Protein Sci.* **2013**, *73*, 30.1.1–30.1.16.
209. Mattheyses, A. L.; Simon, S. M.; Rappoport, J. Z. *J. Cell Sci.* **2010**, *123*, 3621–3628.
210. Hu, B.; Xing, S.; Ren, J.; Wang, Z. *Tetrahedron* **2010**, *66*, 5671–5674.
211. Hwang, T. L.; Shaka, A. J. *J. Magn. Reson.* **1995**, *A112*, 275–279.
212. Dalvit, C. *J. Biol. NMR*, **1998**, *11*, 437–444.
213. Hoffmann, C.; Gaietta, G.; Zürn, A.; Adams, S. R.; Terrillon, S.; Ellisman, M. H.; Tsien, R. Y.; Lohse, M. J. *Nat. Protoc.* **2010**, *5*, 1666–1677.
214. Fu, N.; Xiong, Y.; Squiler, T. C. *Bioconjugate chem.* **2013**, *24*, 251–259.
215. Hauser, C. T.; Tsien, R. Y. *PNAS* **2007**, *104*, 3693–3697.
216. Ojida, A.; Honda, K.; Shinmi, D.; Kiyonaka, S.; Mori, Y.; Hamachi, I. *J. Am. Chem. Soc.*

- 2006**, *128*, 10452–10459.
217. Nonaka, H.; Tsukiji, S.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2007**, *129*, 15777–15779.
218. Uchinomiya, S.; Nonaka, H.; Wakayama, S.; Ojida, A.; Hamachi, I. *Chem. commun*, **2013**, *49*, 5022–5024.
219. Howarth, M.; Takao, K.; Hayashi, Y.; Ting, A. Y. *PNAS* **2005**, *102*, 7583–7588.
220. Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y. *Nat. Methods* **2005**, *2*, 99–104.
221. Wu, P.; Shui, W.; Carlson, B. L.; Hu, N.; Rabuka, D.; Lee, J.; Bertozzi, C. R. *PNAS* **2009**, *106*, 3000–3005.
222. Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707–708.
223. Uttamapinant, C.; White, K. A.; Baruah, H.; Thompson, S.; Fernández-Suárez, M.; Puthenveetil, S.; Ting, A. Y. *PNAS* **2010**, *107*, 10914–10919.
224. Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. *PNAS* **2005**, *102*, 15815–15820.
225. Zhang, C.; Welborn, M.; Zhu, T.; Yang, N. J.; Santos, M. S.; Van Voorhis, T.; Pentelute, B. L. *Nat. Chem.* **2015**, 1–9.
226. Akgun, B.; Li, C.; Hao, Y.; Lambkin, G.; Derda, R.; Hall, D. G. *J. Am. Chem. Soc.* **2017**, *139*, 14285–14291.
227. Ng, S.; Lin, E.; Kitov, P. I.; Tjhung, K. F.; Gerlits, O. O.; Deng, L.; Kasper, B.; Sood, A.; Paschal, B. M.; Zhang, P.; Ling, C-C.; Klassen, J. S.; Noren, C. J.; Mahal, L. K.; Woods, R. J.; Coates, L.; Derda, R. *J. Am. Chem. Soc.* **2015**, *137*, 5248–5251.
228. Smith, G. P. *Science*. **1985**, *228*, 1315–1317.
229. Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391–410.
230. Rentero Rebollo, I.; Heinis, C. *Methods* **2013**, *60*, 46–54.
231. Heinis, C.; Winter, G. *Curr. Opin. Chem. Biol.* **2015**, *26*, 89–98.
232. Lim, R. K. V.; Li, N.; Ramil, C. P.; Lin, Q. *ACS Chem. Biol.* **2014**, *9*, 2139–2148.
233. Yang, J.; Li, K.; Hou, J-T.; Li, L-L.; Lu, C-Y.; Xie, Y-M.; Wang, X.; Yu, X-Q. *ACS Sens.* **2016**, *1*, 166–172.
234. Goswami, L. N.; Houston, Z. H.; Sarma, S. J.; Jalisatgi, S. S.; Frederick Hawthorne, M. F. *Org. Biomol. Chem.* **2013**, *11*, 1116–1126.

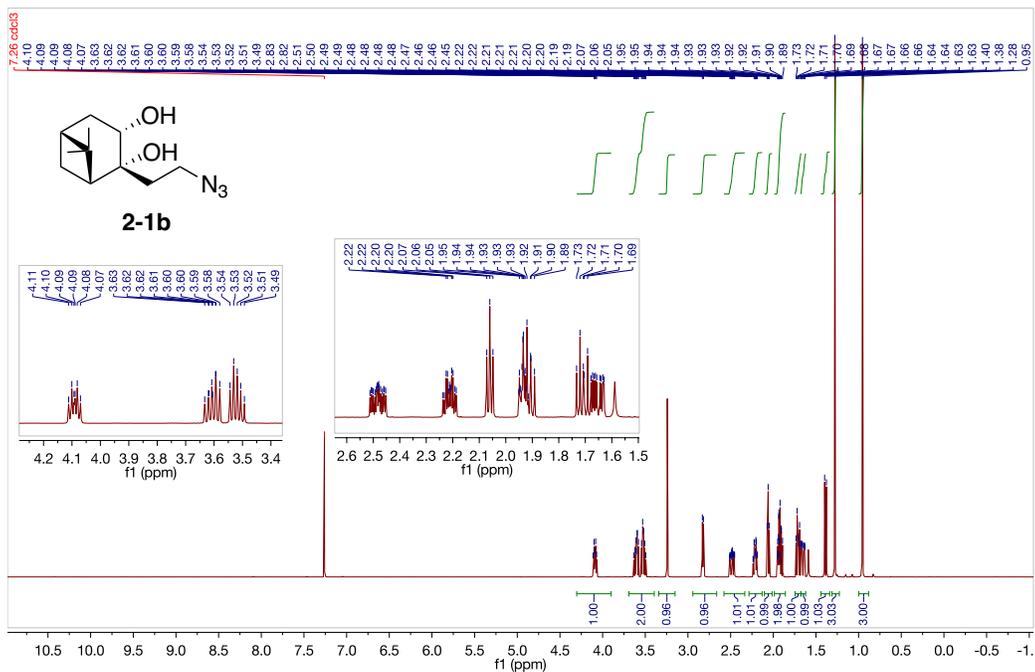
235. Sella, E.; Shabat, D. *Org. Biomol. Chem.* **2013**, *11*, 5074–5078.
236. Presolski, S. I.; Hong, V. P.; Finn, M. G. *Curr. Protoc. Chem. Biol.* **2011**, *3*, 153–162.
237. Jin, S.; Choudhary, G.; Cheng, Y.; Dai, C.; Li, M.; Wang, B. *Chem. Commun.* **2009**, *0*, 5251–5253.
238. Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyo, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125*, 2416–2425.
239. Matochko, W. L.; Chu, K.; Jin, B.; Lee, S. W.; Whitesides, G. M.; Derda, R. *Methods.* **2012**, *58*, 47–55.
240. Matochko, W.; Li, S. C.; Tang, S. K. Y.; Derda, R. *Nuc. Acid. Res.* **2014**, *42*, 1784–1798.
241. Tjhung, K. F.; Kitov, P. I.; Ng, S.; Kitova, E. N.; Deng, L.; Klassen, J. S.; Derda, R. *J. Am. Chem. Soc.* **2016**, *138*, 32–35.
242. Hall, D. G. *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials (Volume 1 and 2), Second Edition* (Ed: Hall, D. G.), Wiley-VCH Verlag GmbH & Co. KGaA, **2011**, p. 19–20.
243. Kaiser, F.; Schwink, L.; Velder, J.; Schmalz, H–G. *J. Org. Chem.* **2002**, *67*, 9248–9256.
244. Otremba, T.; Ravoo, B. J. *Tetrahedron* **2017**, *73*, 4972–4978.
245. Wanner, J.; Romashko, D.; Werner, D. S.; May, E. W.; Peng, Y.; Schulz, R.; Foreman, K. W.; Russo, S.; Arnold, L. D.; Pingle, M.; Bergstrom, D. E.; Barany, F.; Thomson, S. *PLoS One*, **2015**, *10*, e0121793.

# Appendices

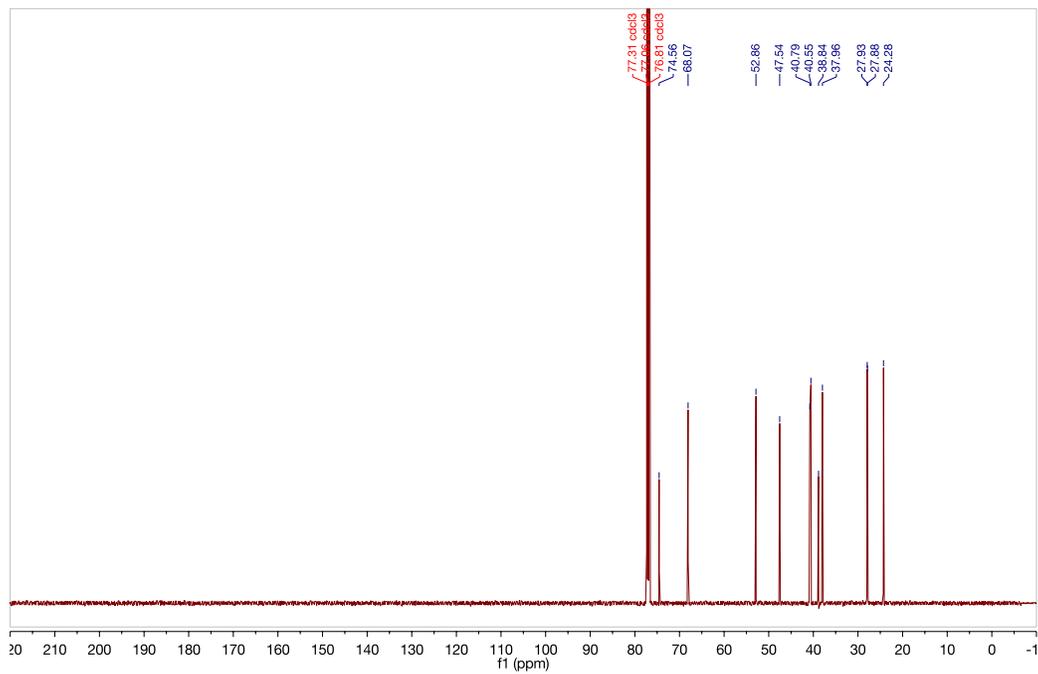
## Appendix 1: Selected copies of NMR spectra

### $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-1b in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-01-38; 499.806 MHz H1 PRESAT in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

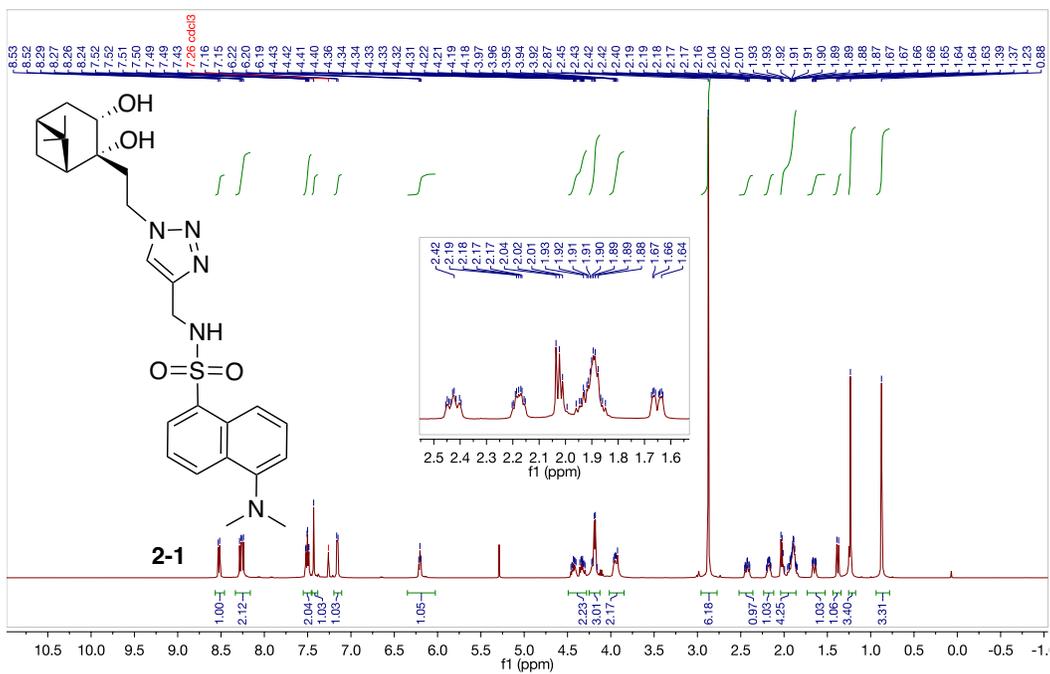


BAH-01-38; 125.691 MHz C13[H1] 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

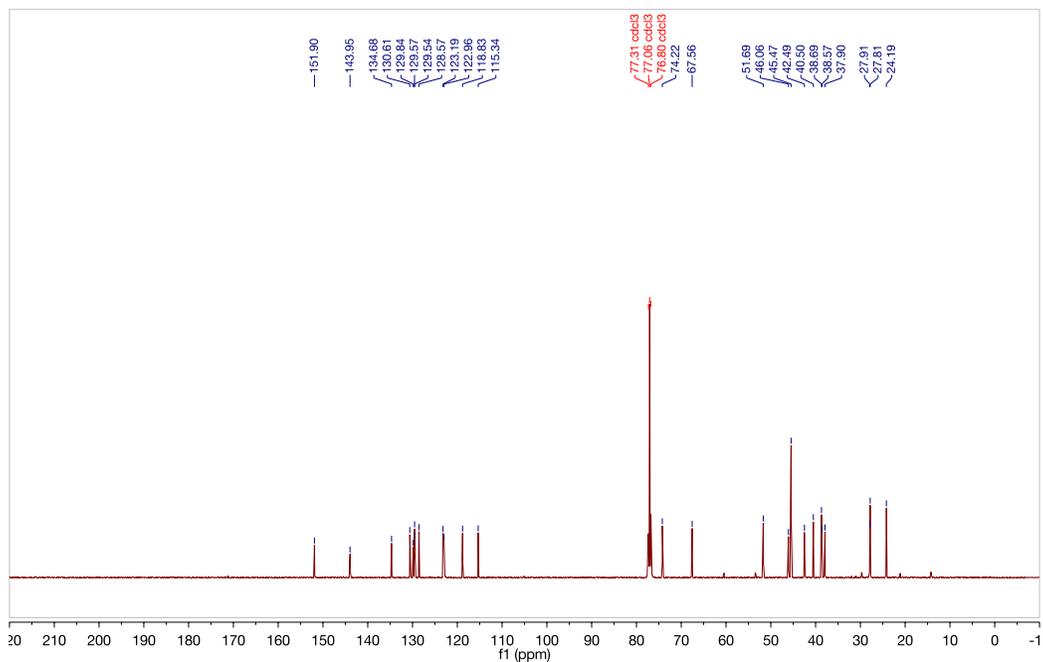


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-1 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-04-146; 499.806 MHz H1 PRESAT in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

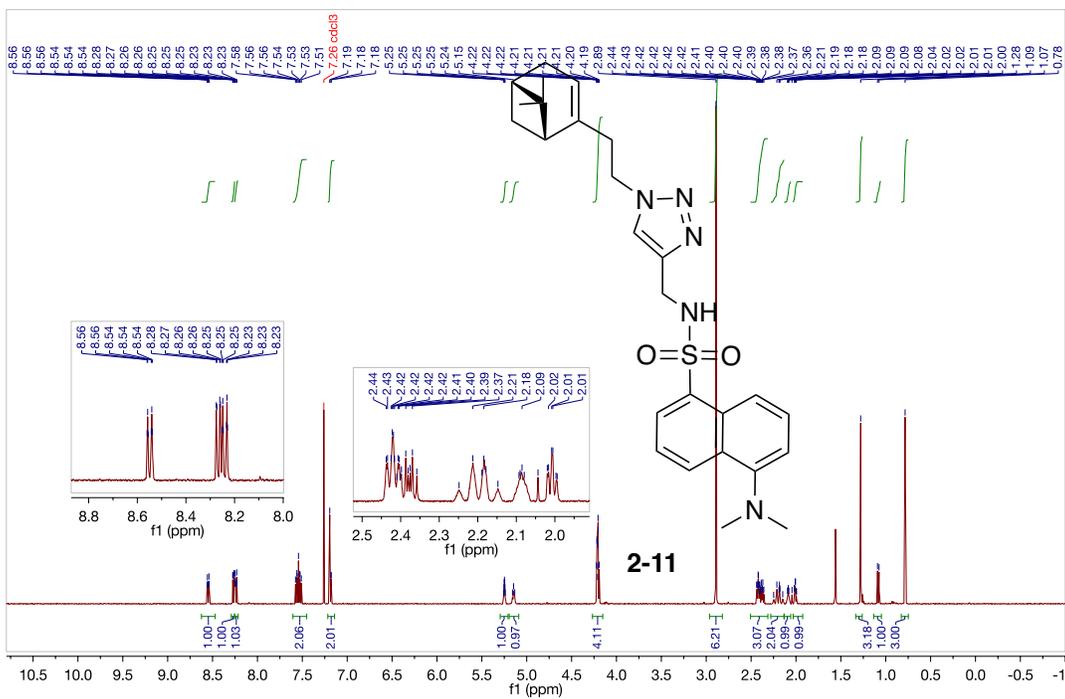


BAH-04-146; 125.691 MHz C13[H1] 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

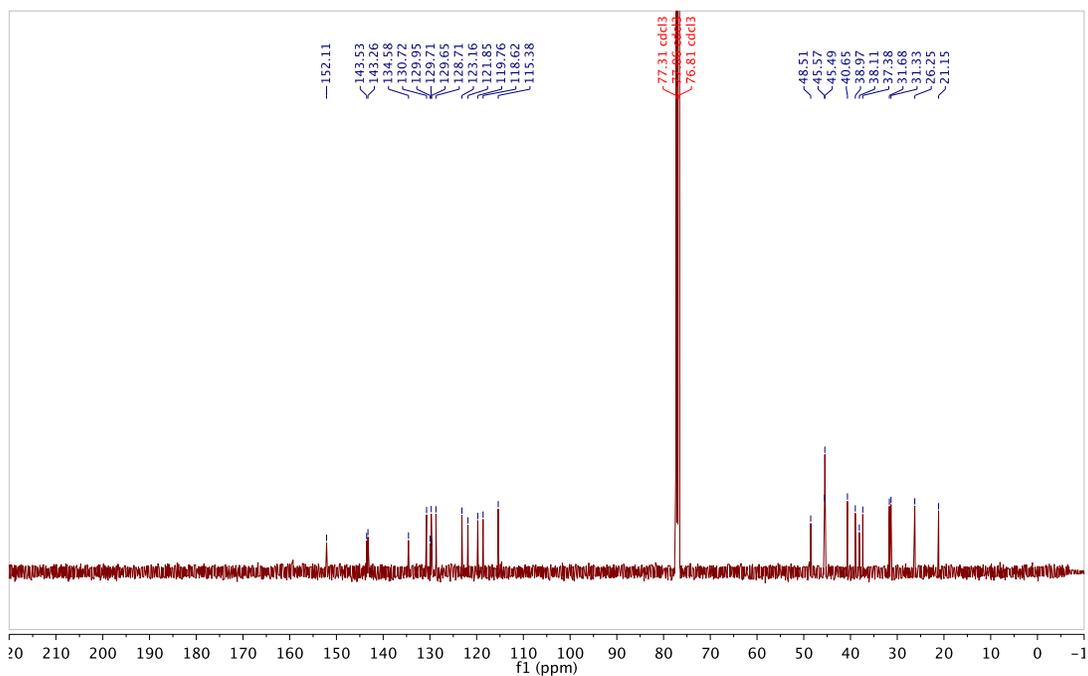


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-11 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-04-190; 498.120 MHz H1 1D in  $\text{CDCl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 26.4 C -> actual temp = 27.0 C, autotx probe



BAH-04-190; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, coldtdd probe



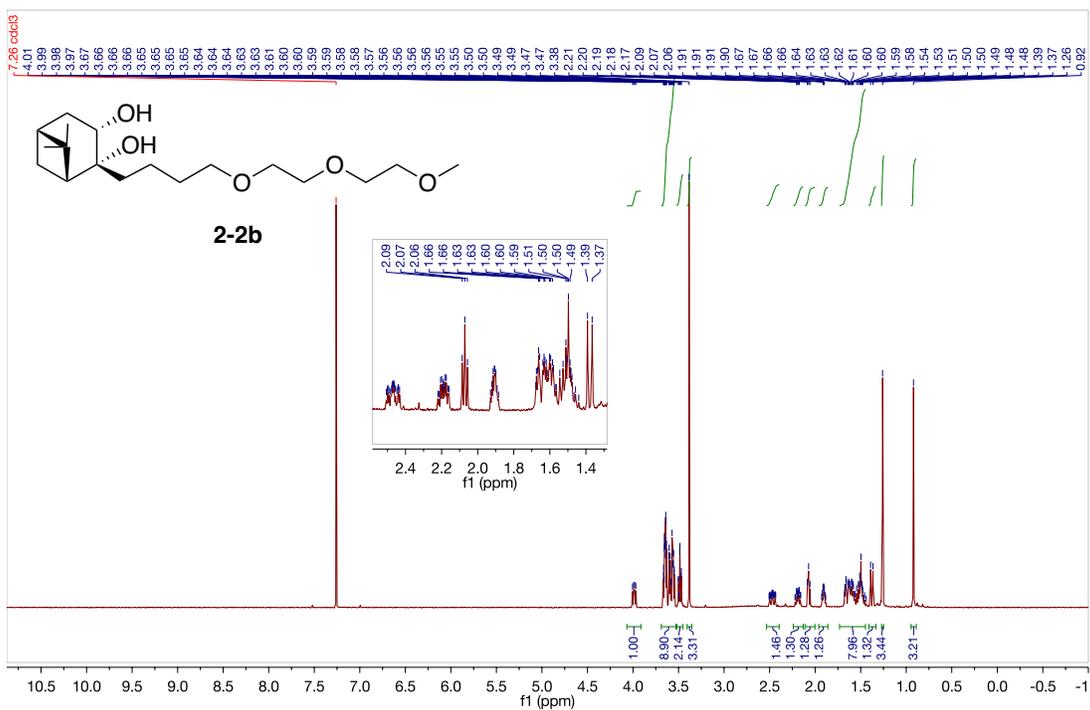




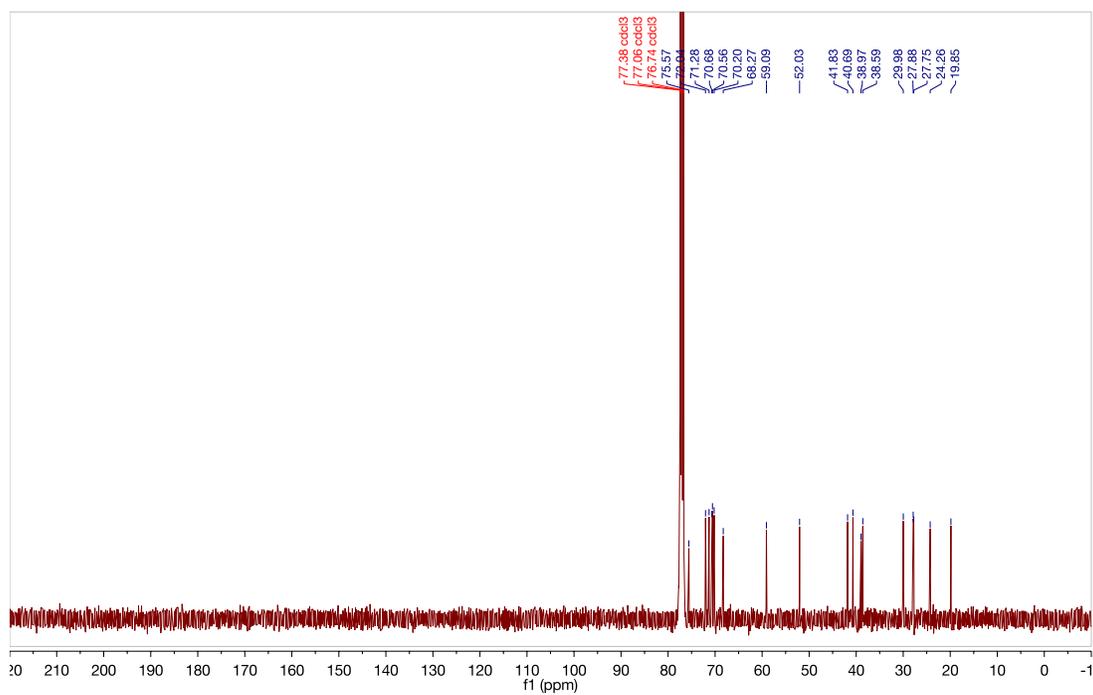


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-2b in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-03-45; 399.984 MHz H1 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 25.9 C -> actual temp = 27.0 C, onenmr probe



BAH-03-45; 100.579 MHz C13[H1] 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 25.5 C -> actual temp = 26.9 C, sw400 probe



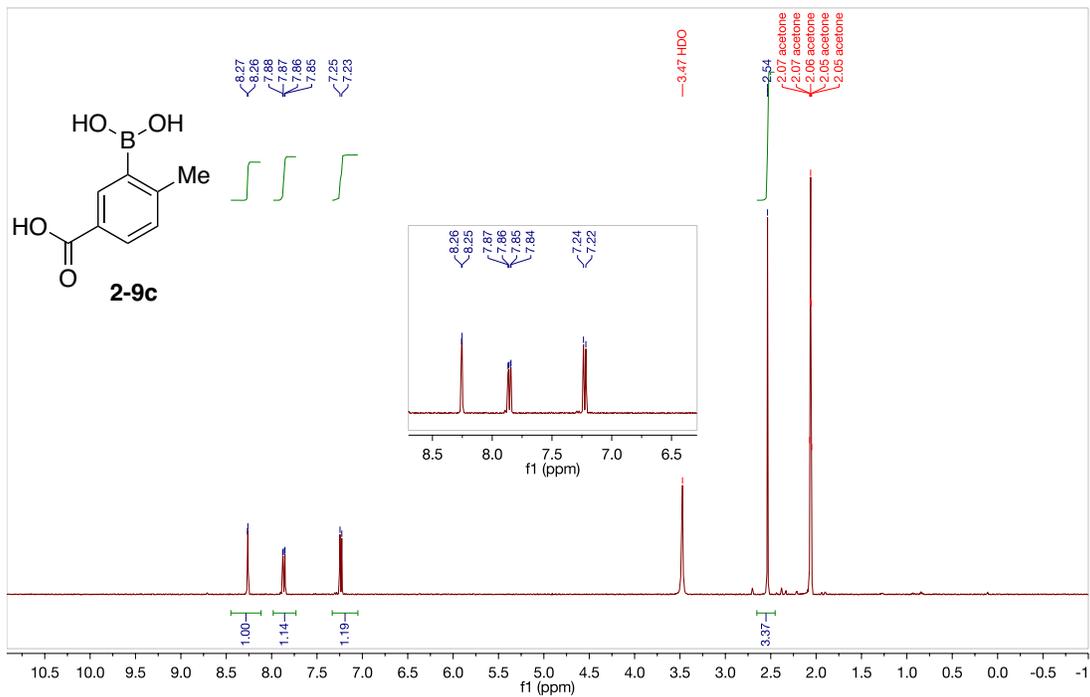




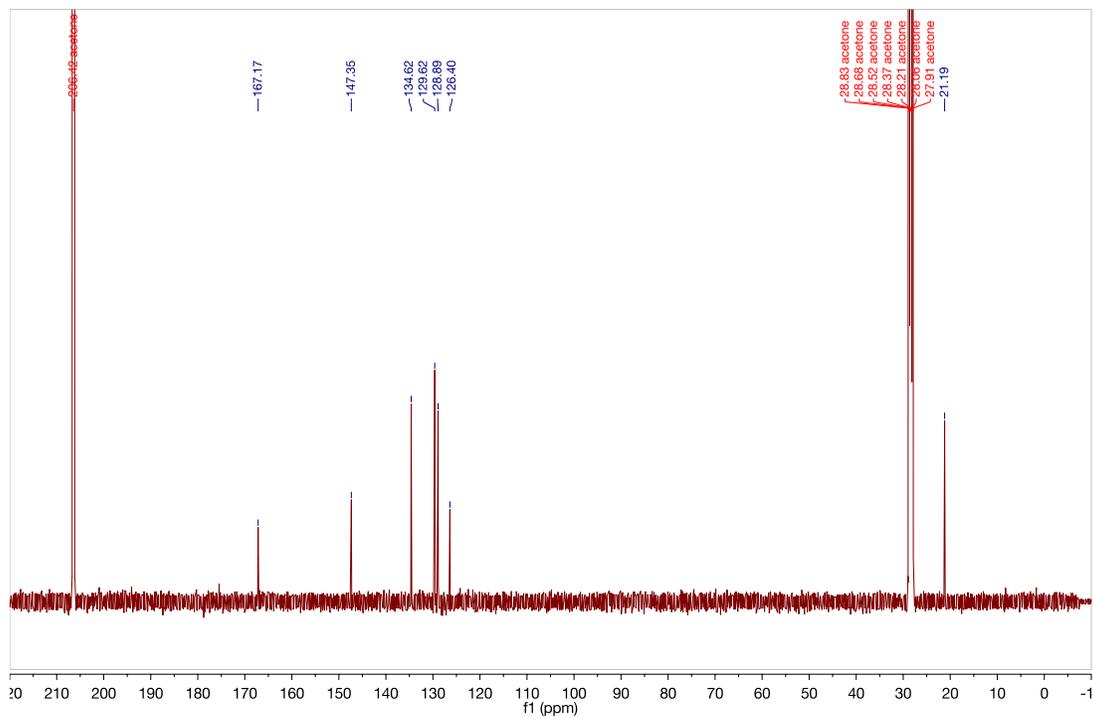


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-9c in acetone- $d_6$ at 25 °C

BAH-04-24; 399.987 MHz H1 1D in acetone (ref. to acetone @ 2.04 ppm), temp 25.9 C -> actual temp = 27.0 C, onenmr probe

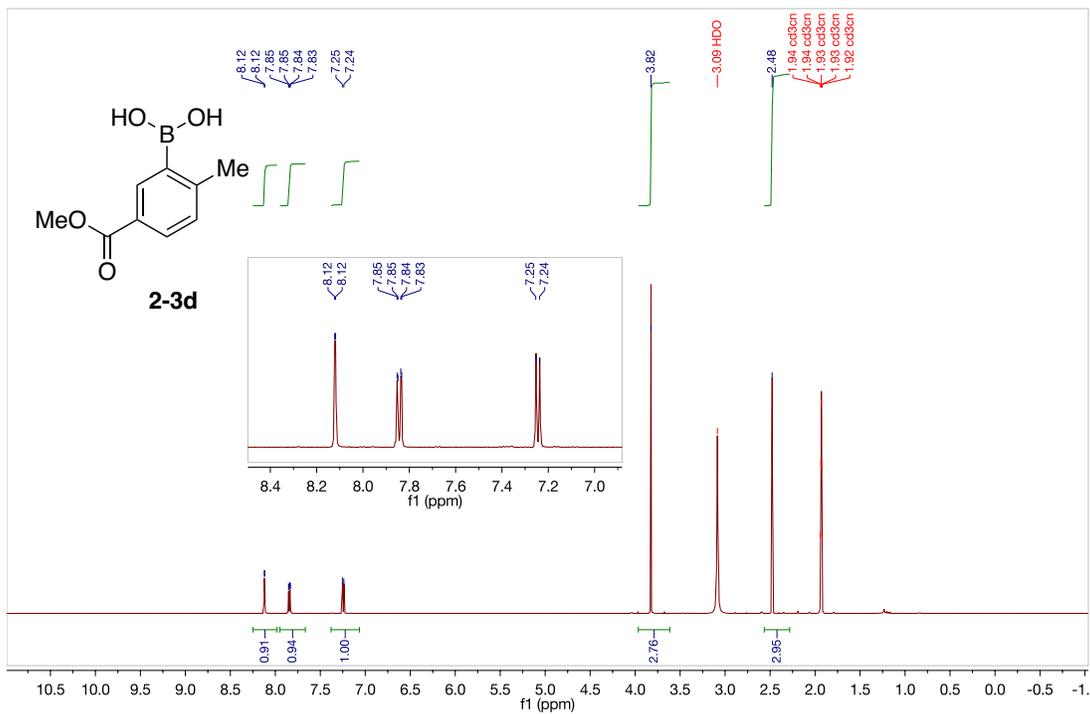


BAH-04-24; 125.692 MHz C13[H1] 1D in acetone (ref. to acetone @ 29.8 ppm), temp 27.7 C -> actual temp = 27.0 C, coldual probe

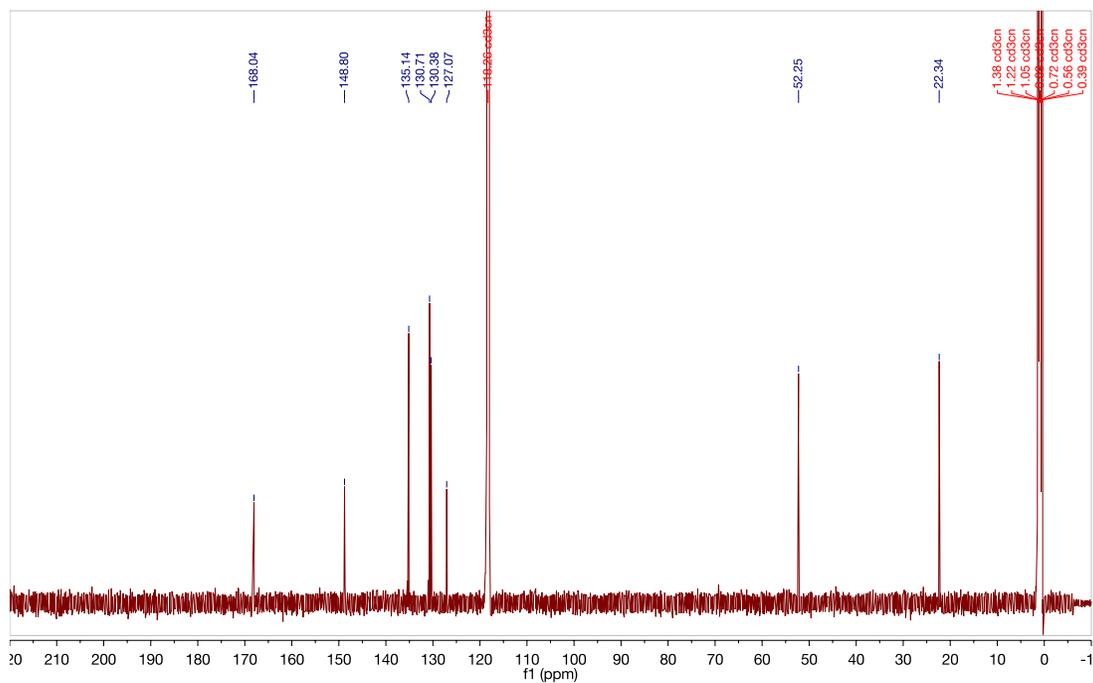


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-3d in $\text{ACN-}d_3$ at 25 °C

BAH-03-199; 499.809 MHz H1 PRESAT in cd3cn, temp 27.7 C -> actual temp = 27.0 C, cold dual probe

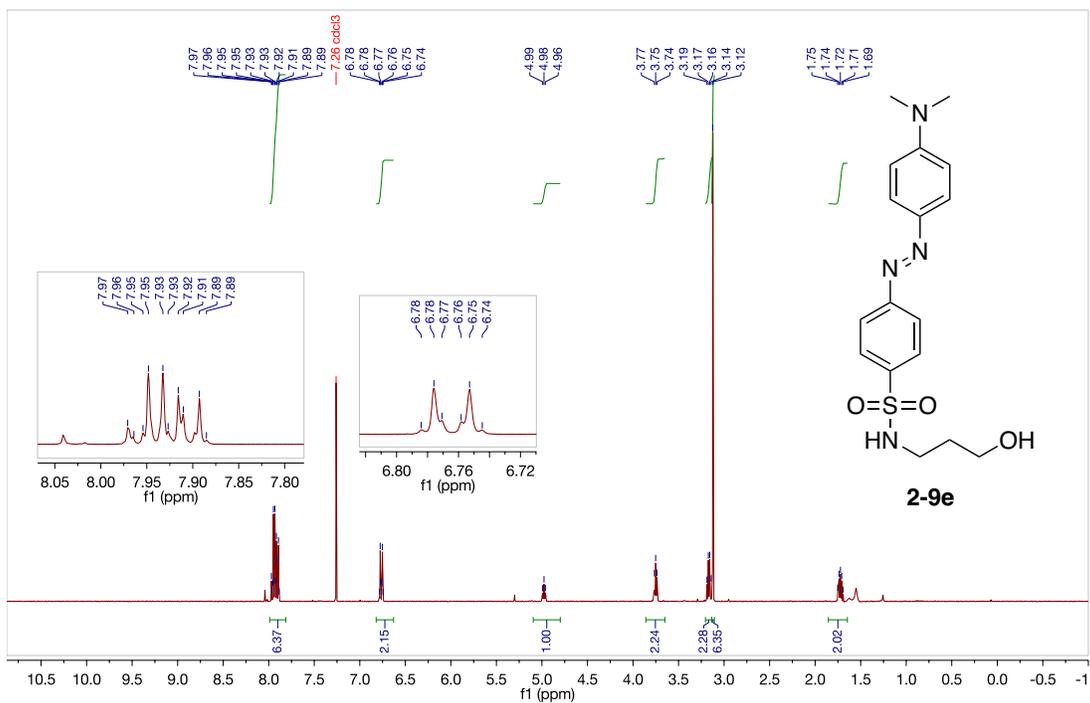


BAH-03-199; 125.692 MHz C13[H1] 1D in cd3cn, temp 27.7 C -> actual temp = 27.0 C, cold dual probe

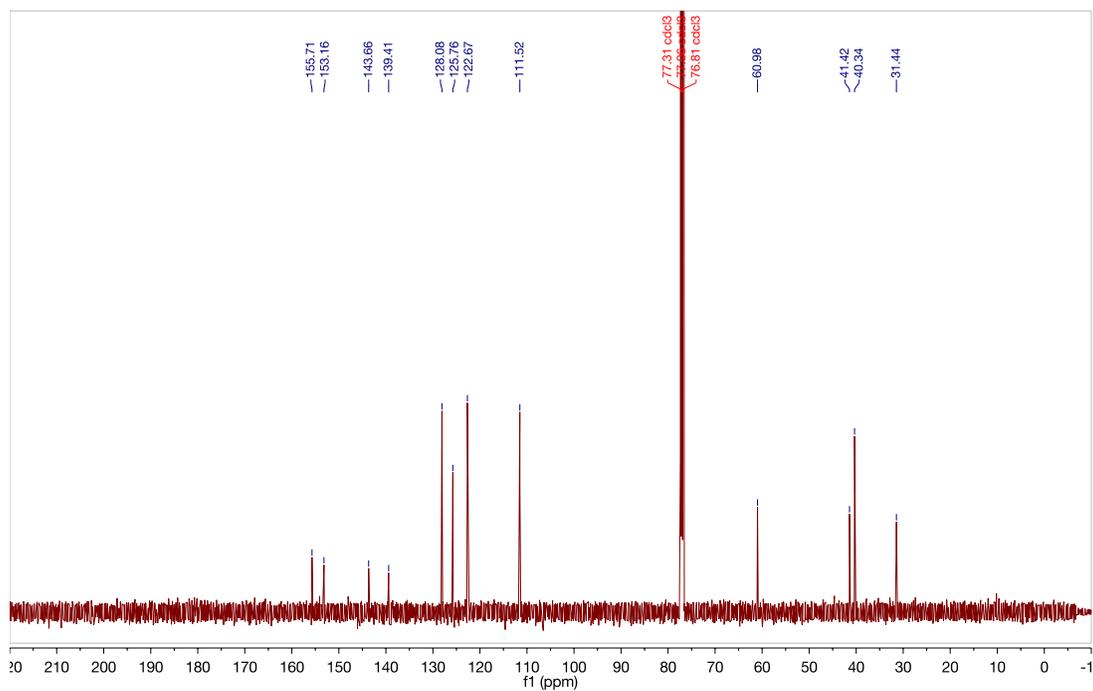


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-9e in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-04-185; 399.984 MHz H1 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 25.9 C -> actual temp = 27.0 C, onenmr probe

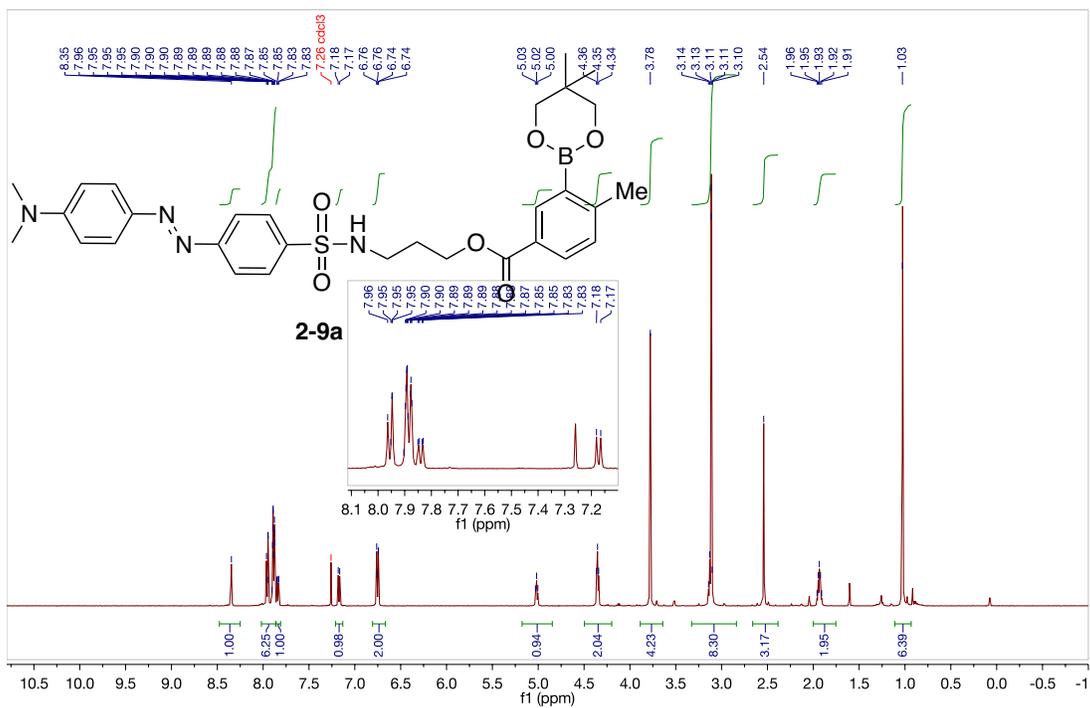


BAH-04-185; 125.691 MHz C13[H1] 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, coldual probe

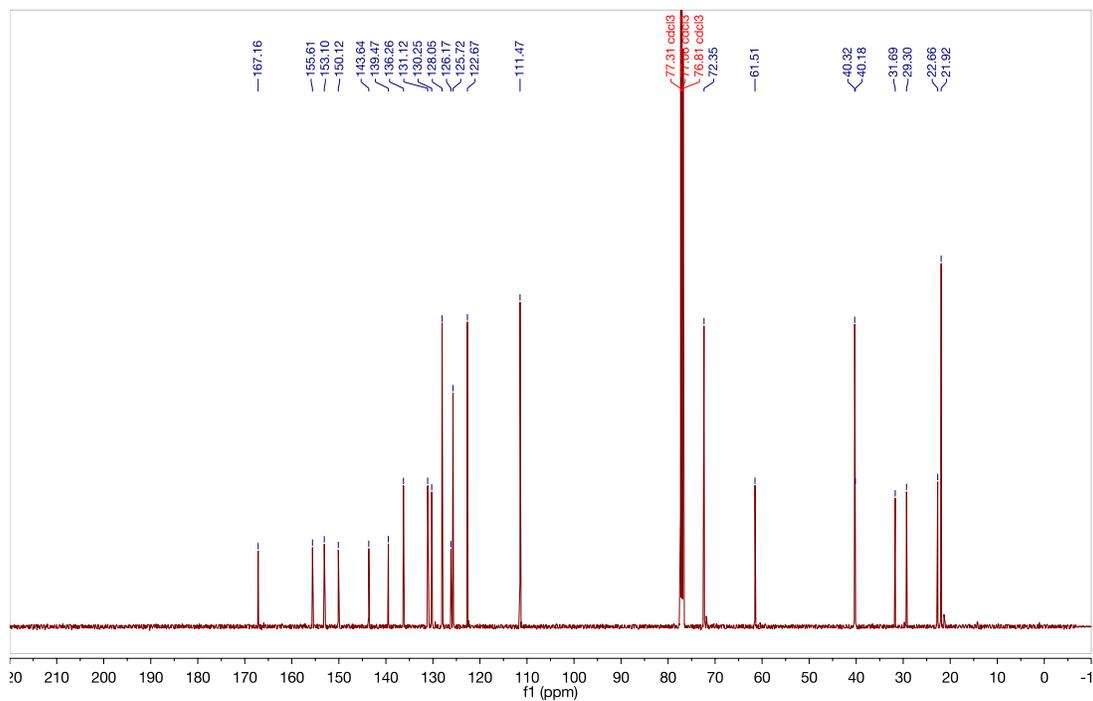


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-9a in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-05-160; 498.120 MHz H1 1D in dmsol (ref. to DMSO @ 2.49 ppm), temp 26.4 C -> actual temp = 27.0 C, autotxdb probe



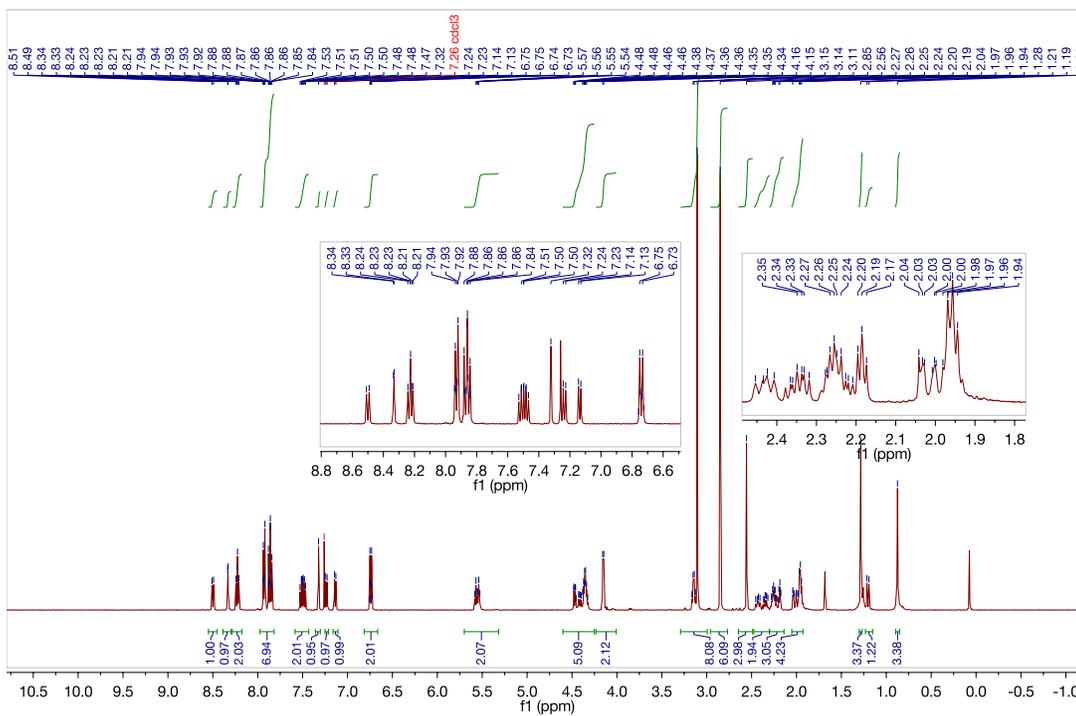
BAH-05-160; 125.691 MHz C13[H1] 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, colddual probe



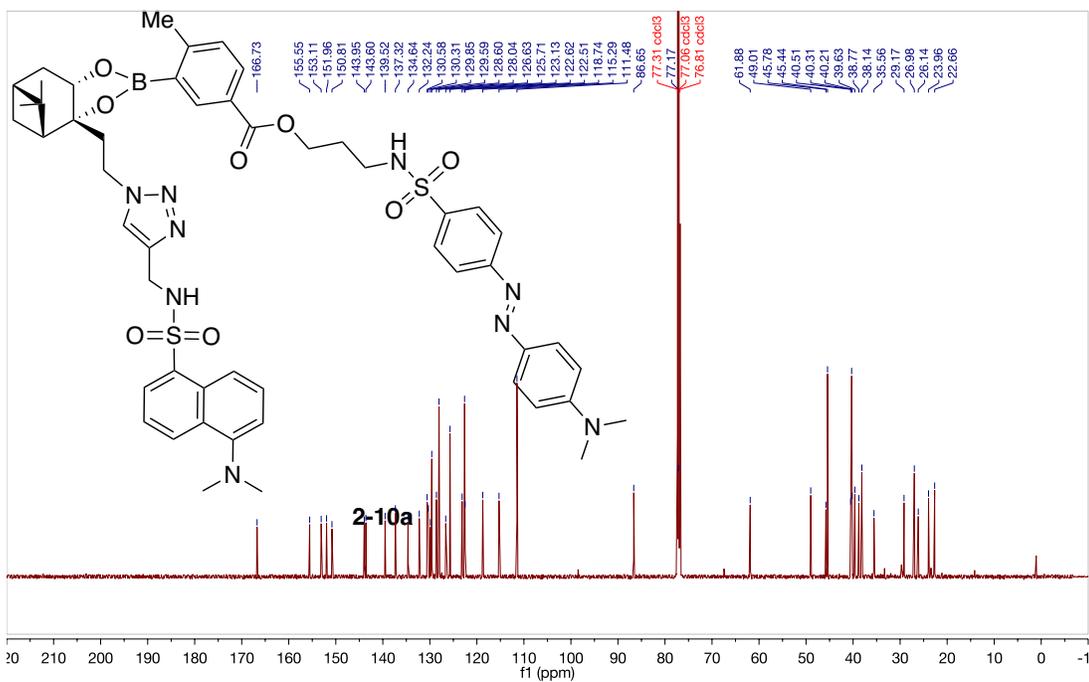


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-10a in $\text{CDCl}_3$ at 25 °C

BAH-04-188; 498.118 MHz H1 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 26.4 C -> actual temp = 27.0 C, autotx probe



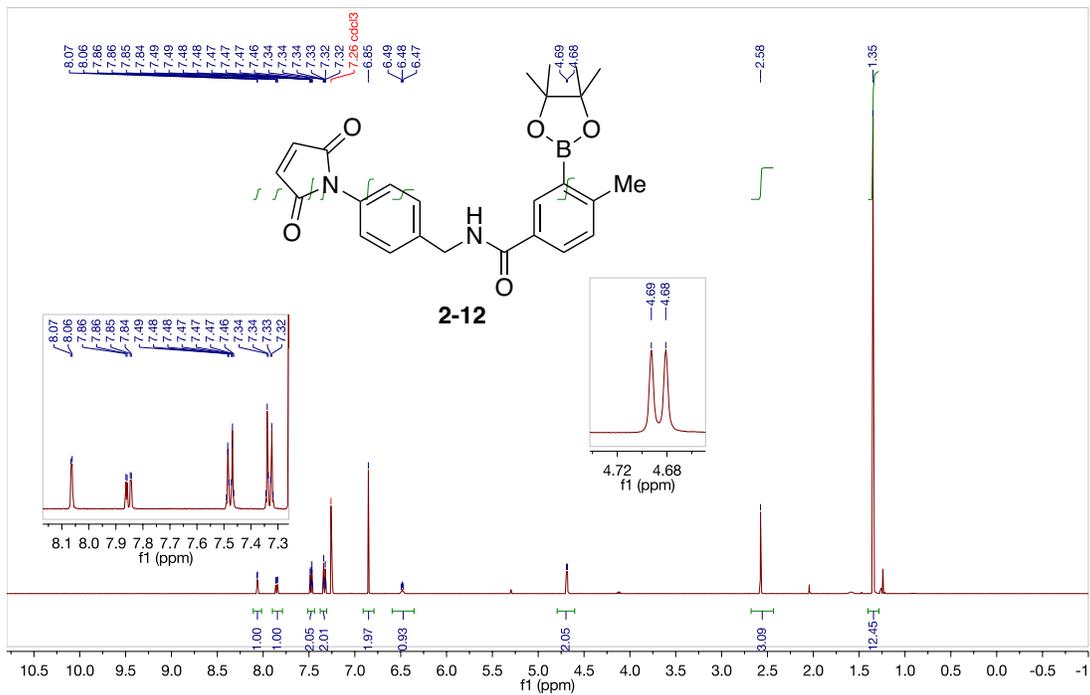
BAH-04-188; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe



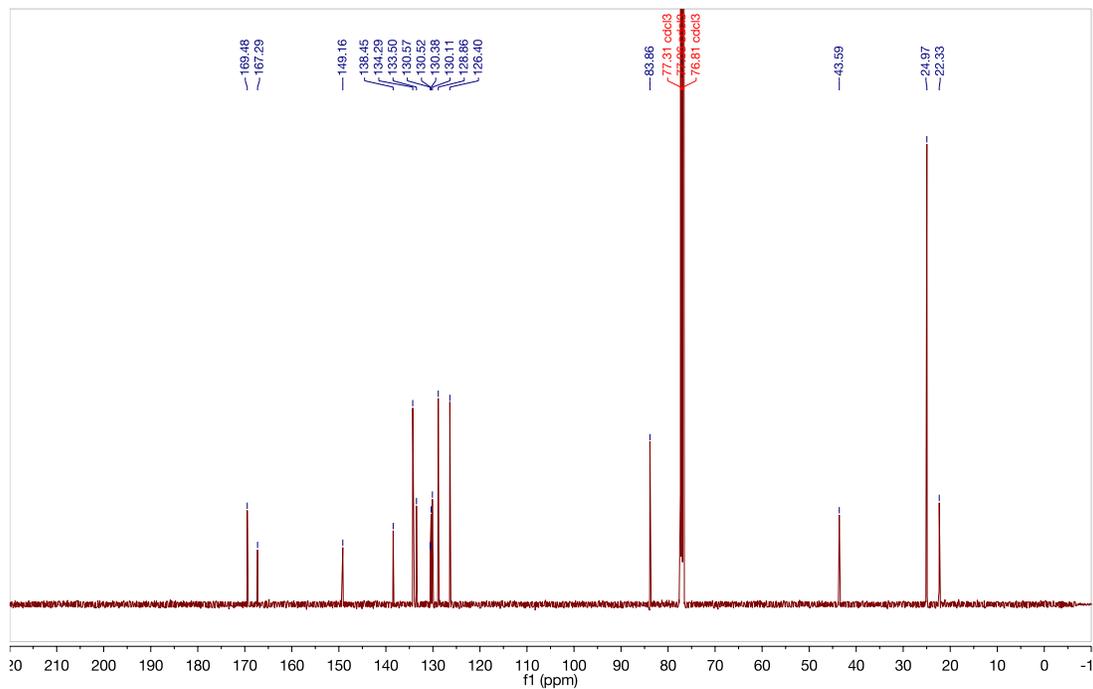


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-12 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-05-44; 498.118 MHz H1 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 26.4 C -> actual temp = 27.0 C, autoxdb probe

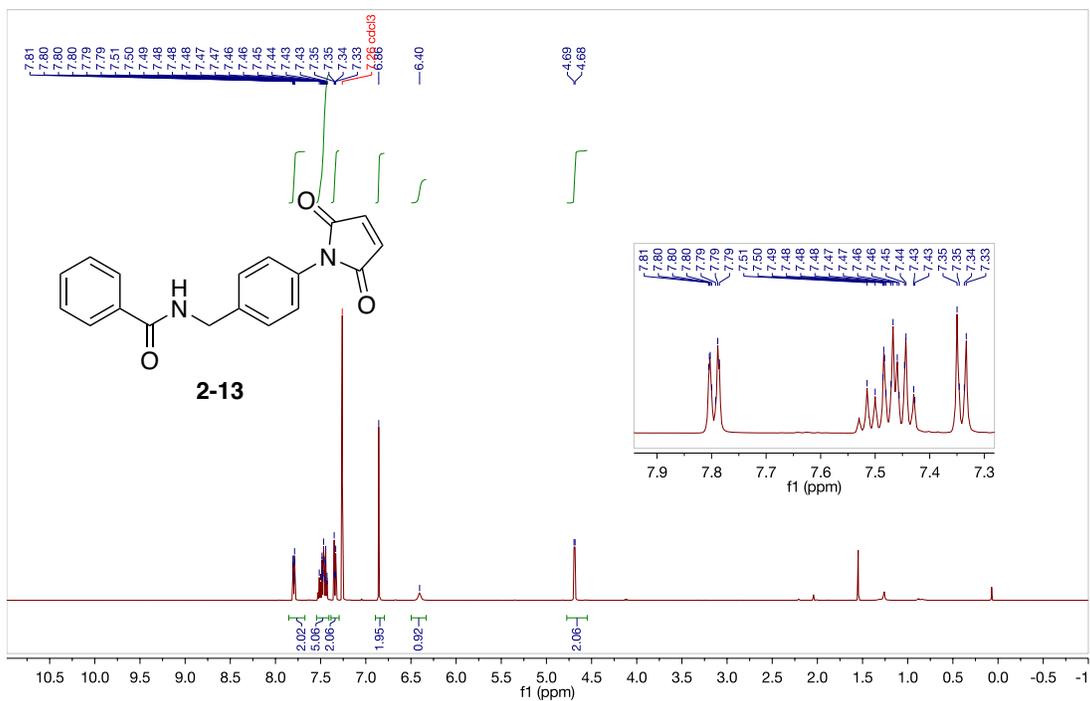


BAH-05-44; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, coldual probe

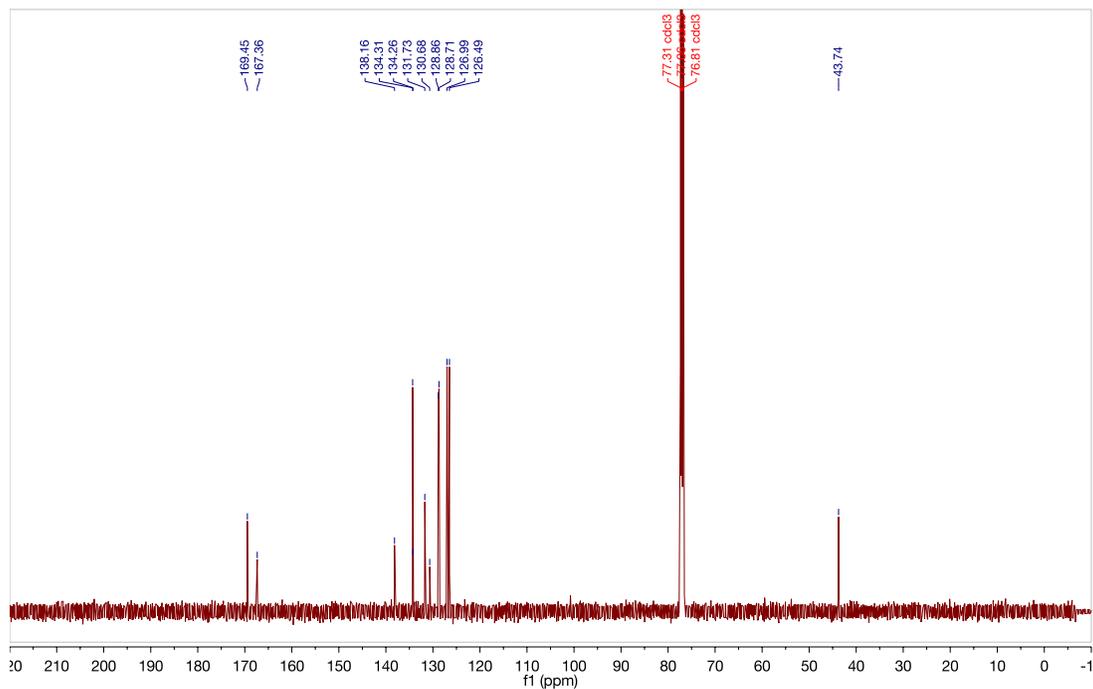


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-13 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-04-108; 499.806 MHz H1 PRESAT in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe



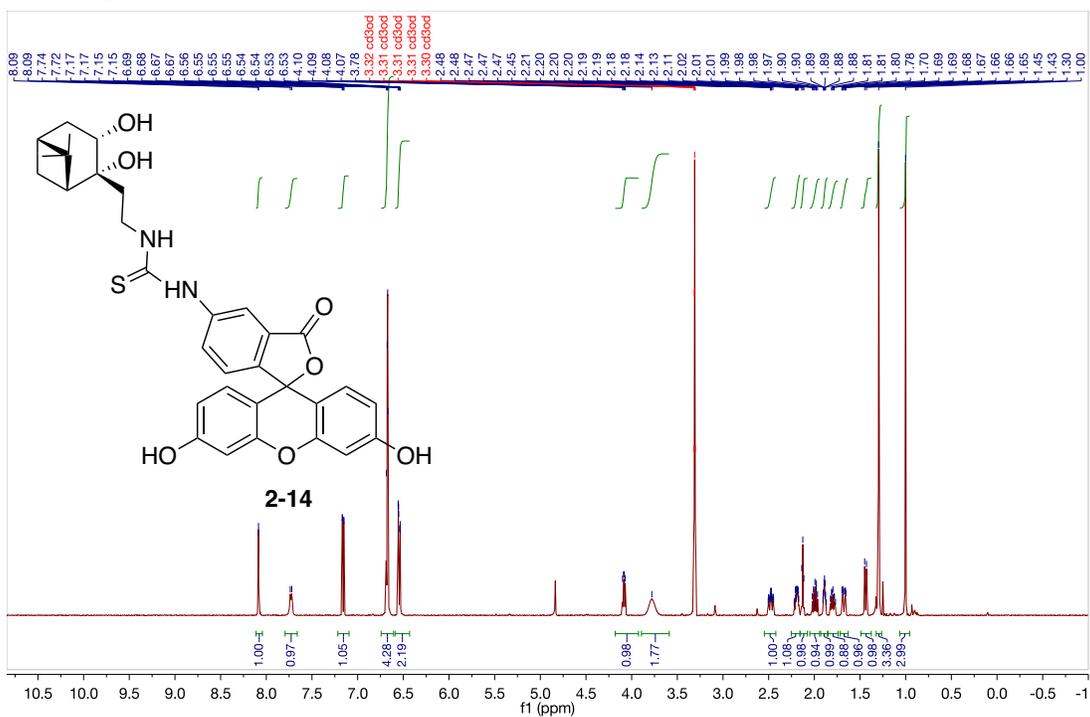
BAH-04-108; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe



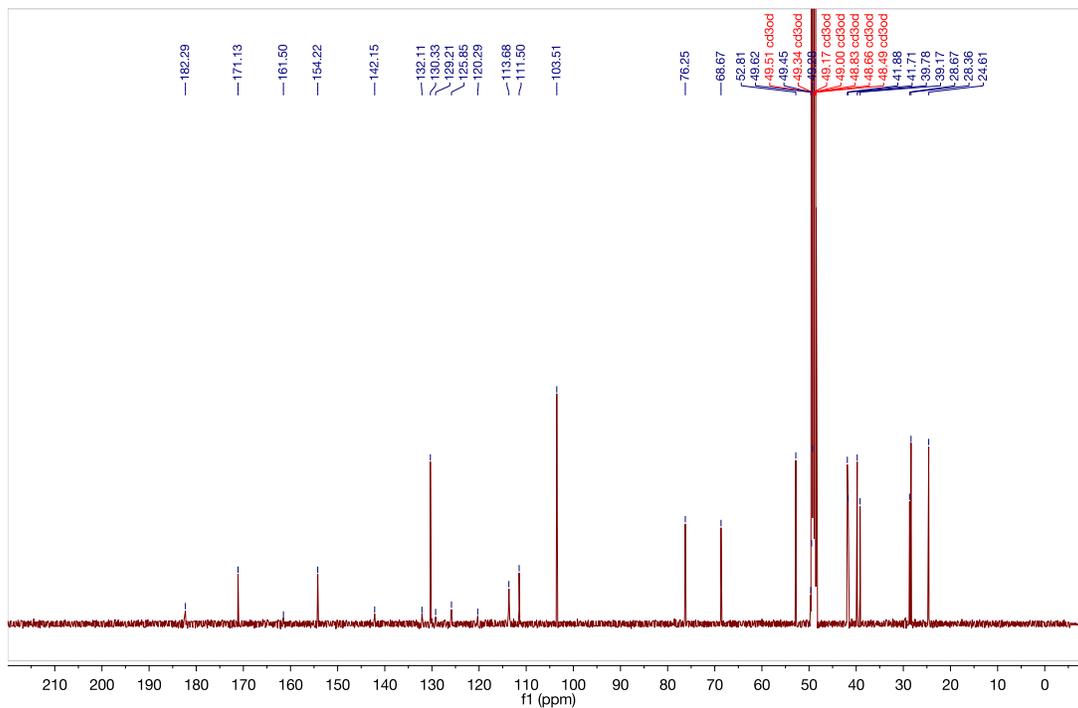




<sup>1</sup>H and <sup>13</sup>C spectrum for 2-14 in CD<sub>3</sub>OD at 25 °C

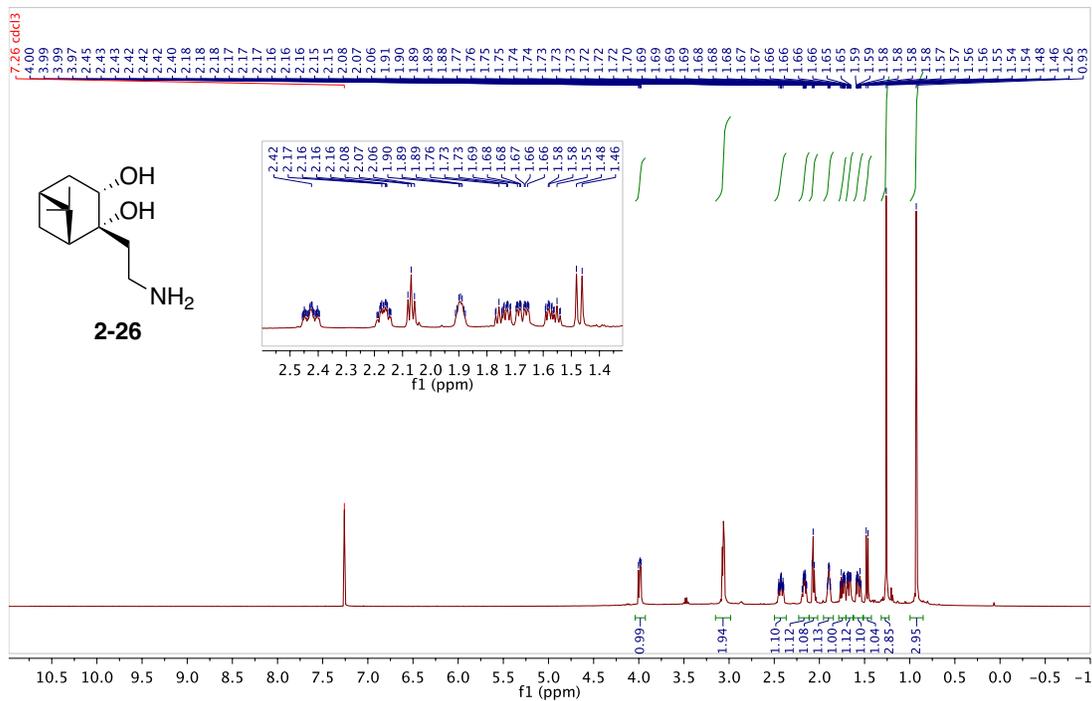


BAH-05-66; 125.691 MHz C13[H1] 1D in cd3od (ref. to CD3OD @ 49.0 ppm), temp 27.7 C -> actual temp = 27.0 C, coldual probe

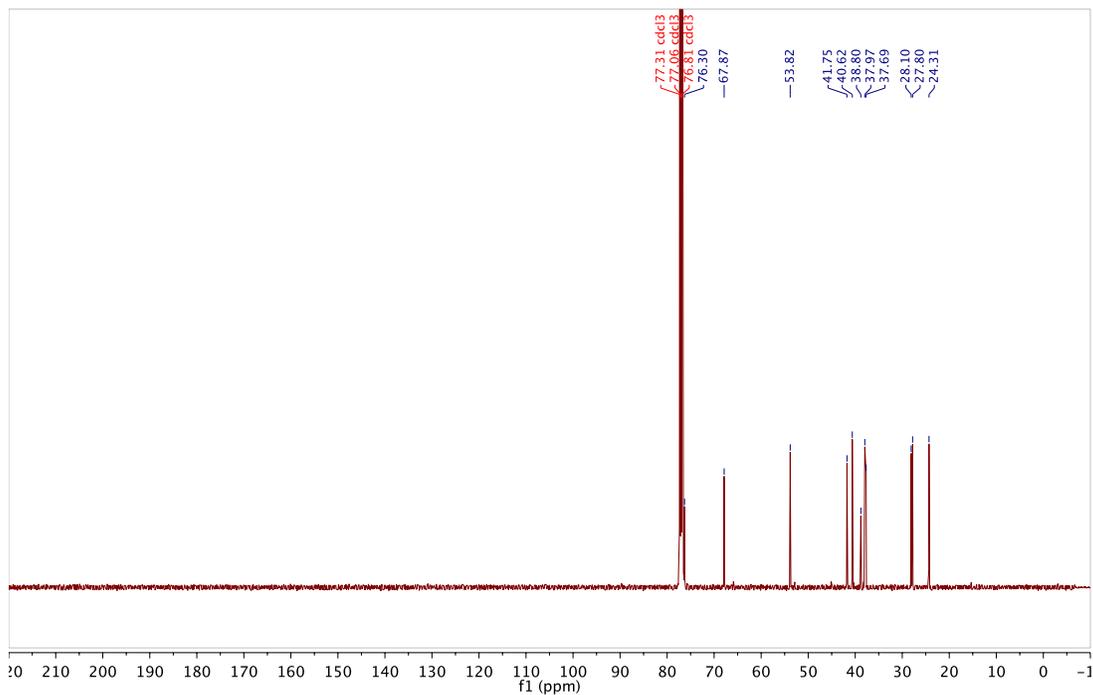


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 2-26 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-05-198; 499.806 MHz H1 PRESAT in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, coldddual probe



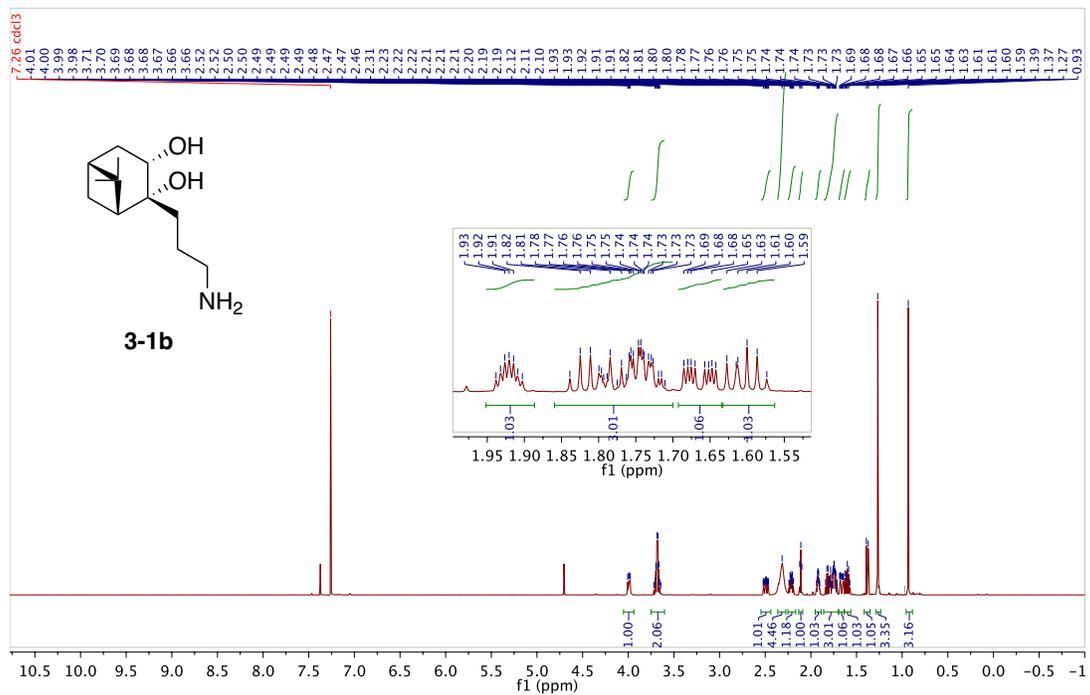
BAH-05-198; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, coldddual probe



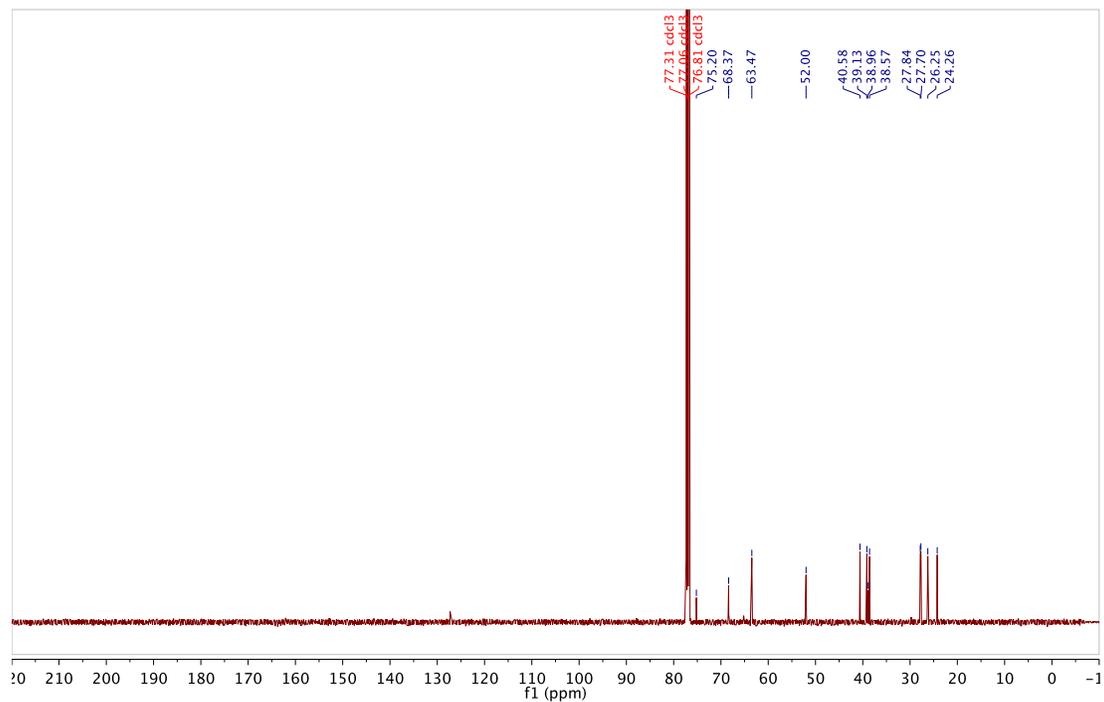


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1b in $\text{CDCl}_3$ at 25 °C

BAH-06-102; 499.806 MHz H1 PRESAT in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

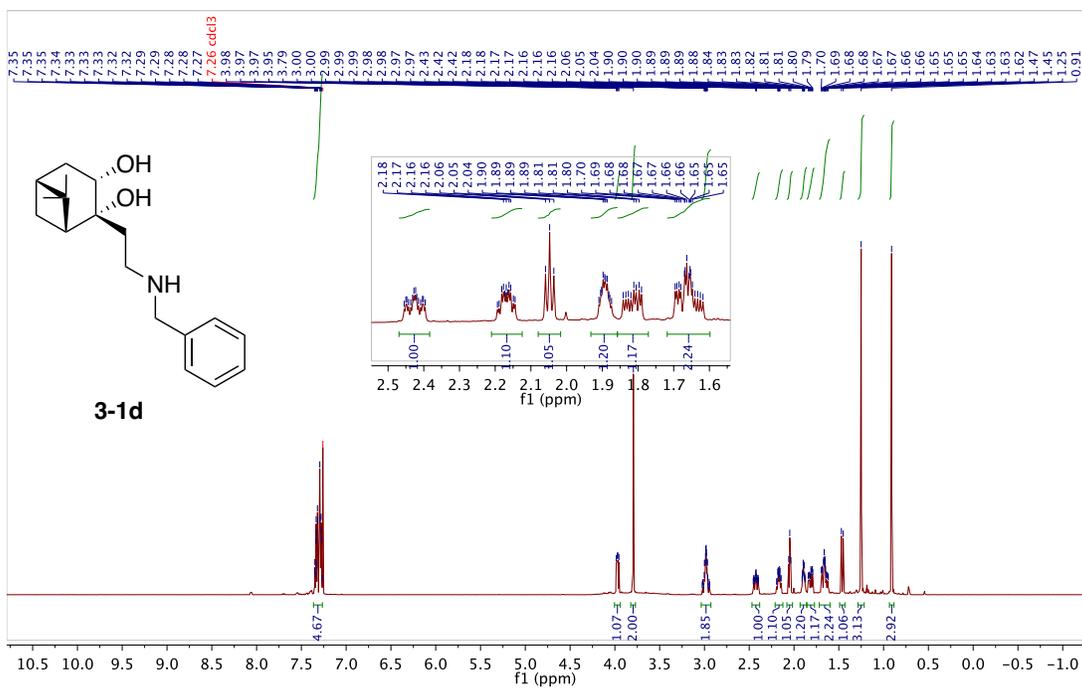


BAH-06-102; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

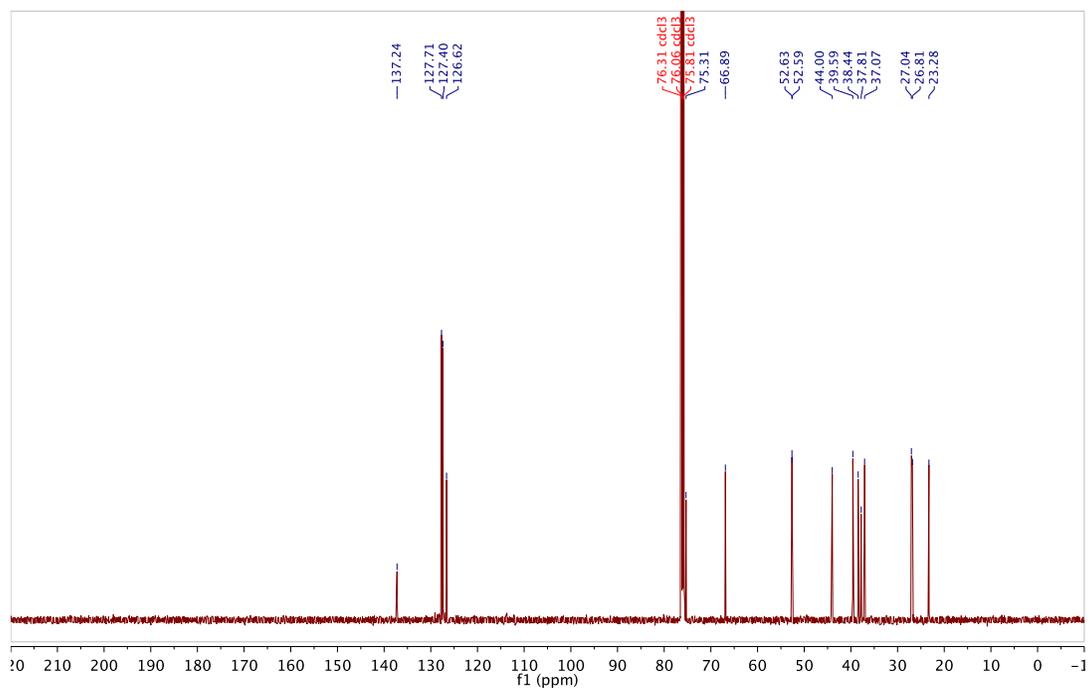


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1d in $\text{CDCl}_3$ at 25 °C

BAH-06-203; 499.797 MHz H1 PRESAT in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, coldddal probe

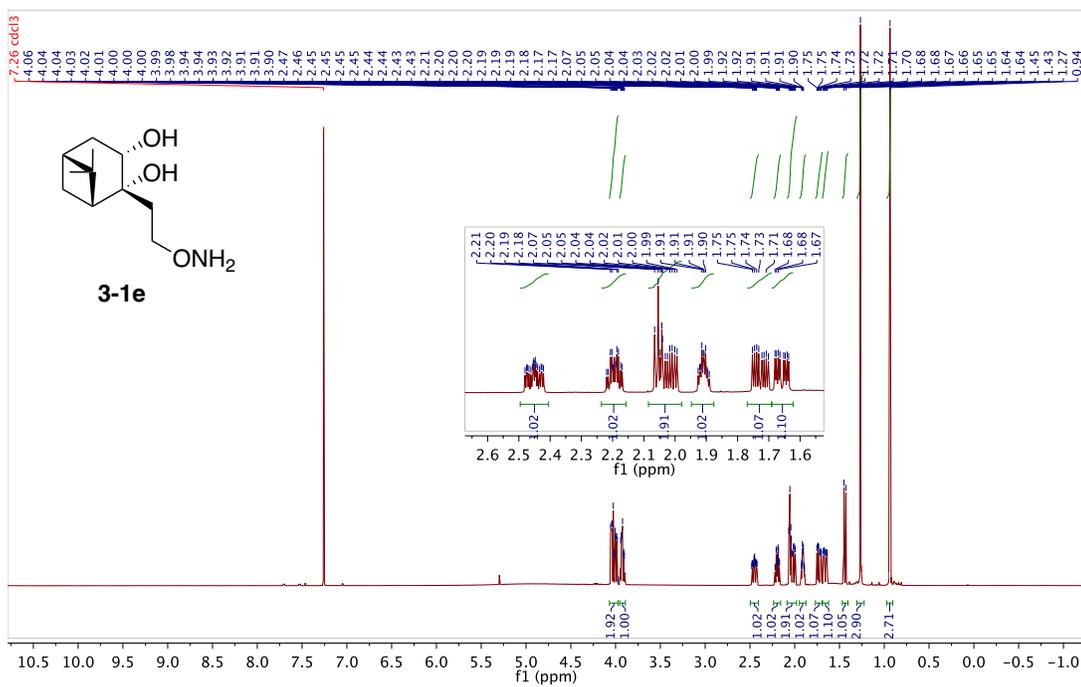


BAH-06-203; 125.688 MHz C13[H1] 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, coldddal probe

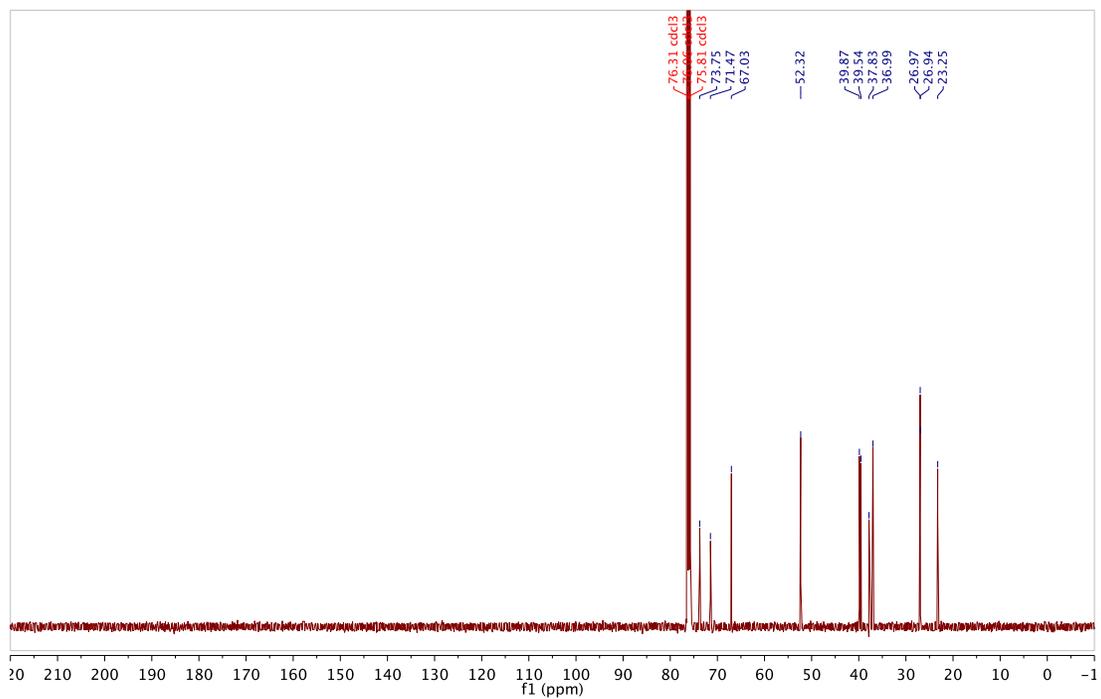


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1e in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-08-93; 499.797 MHz H1 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm) temp 27.7 C -> actual temp = 27.0 C, coldddual probe

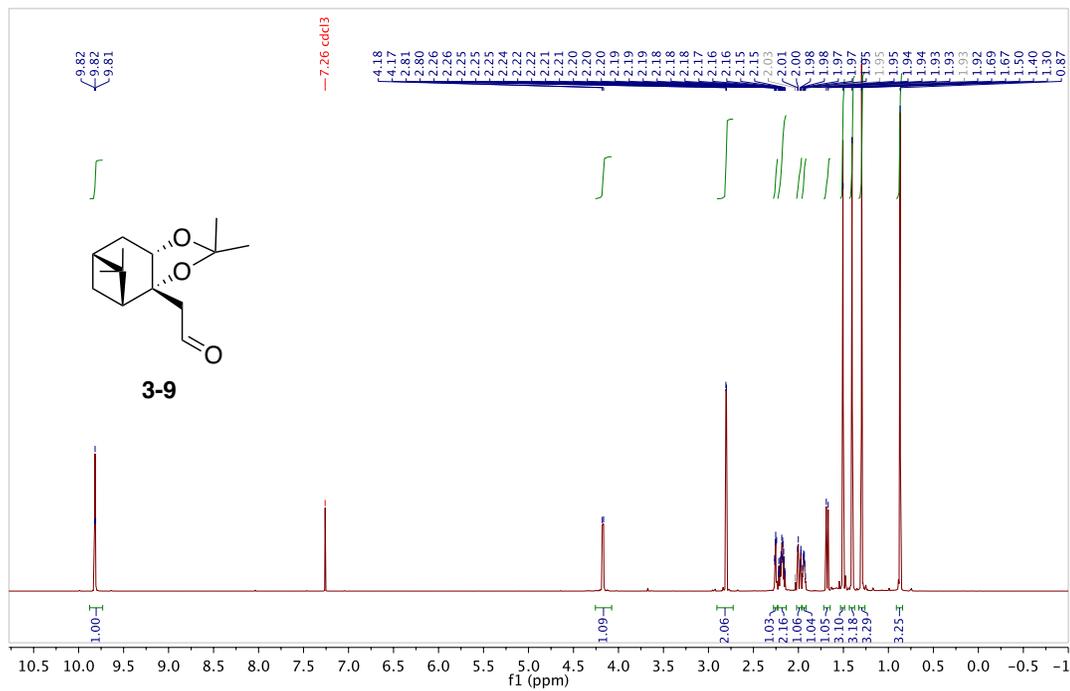


BAH-08-93; 125.688 MHz C13{H1} 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm) temp 27.7 C -> actual temp = 27.0 C, coldddual probe

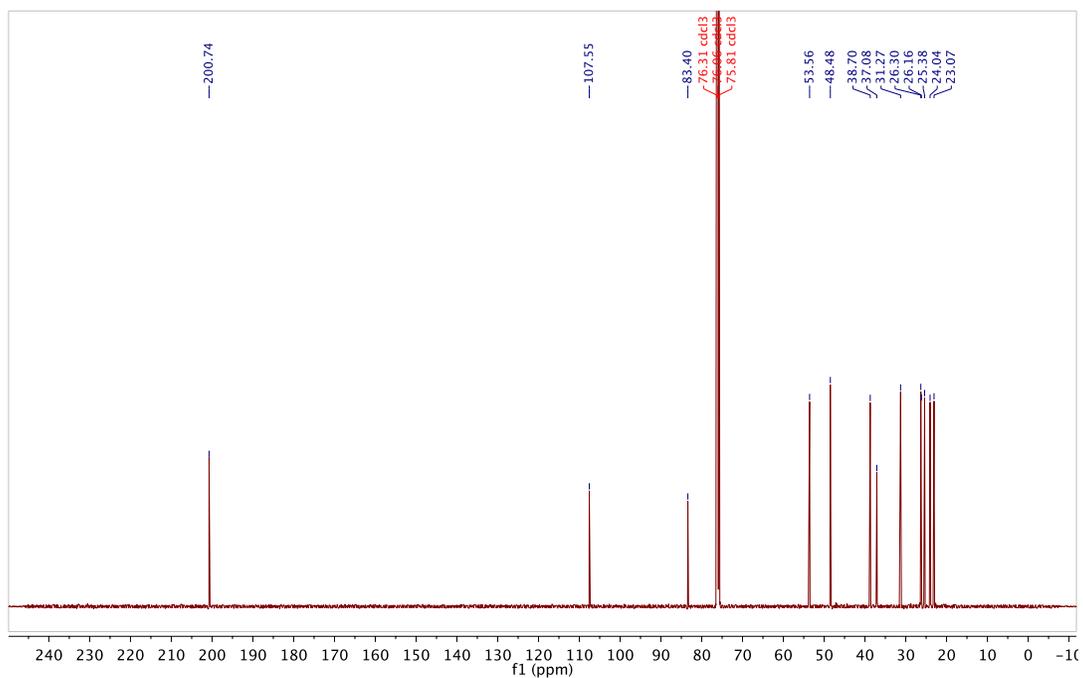


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-9 in $\text{CDCl}_3$ at 25 °C

BAH-06-78; 499.806 MHz H1 PRESAT in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

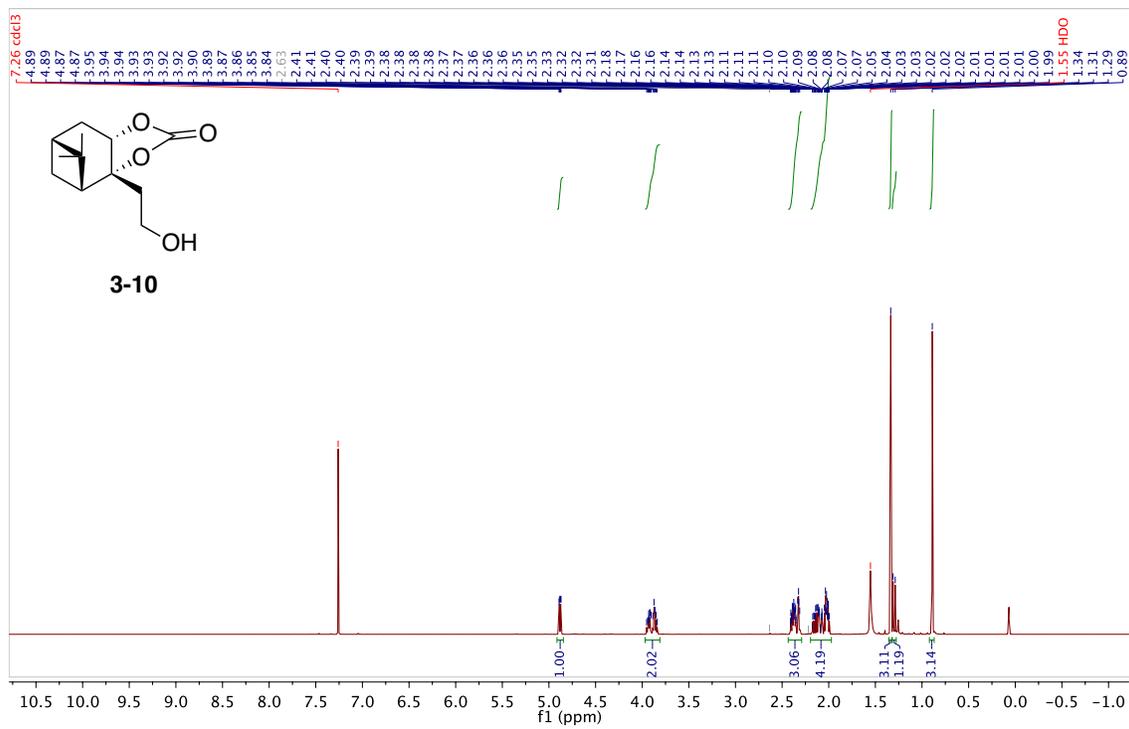


BAH-06-78; 125.691 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

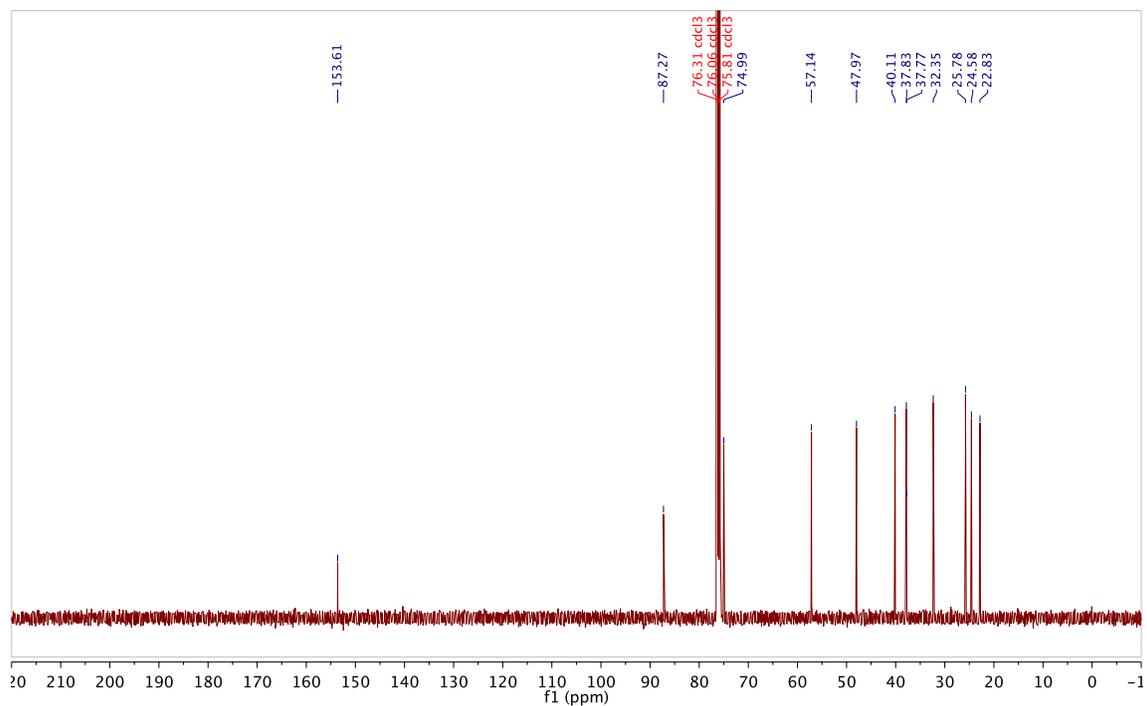


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-10 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-07-51; 499.797 MHz H1 PRESAT in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

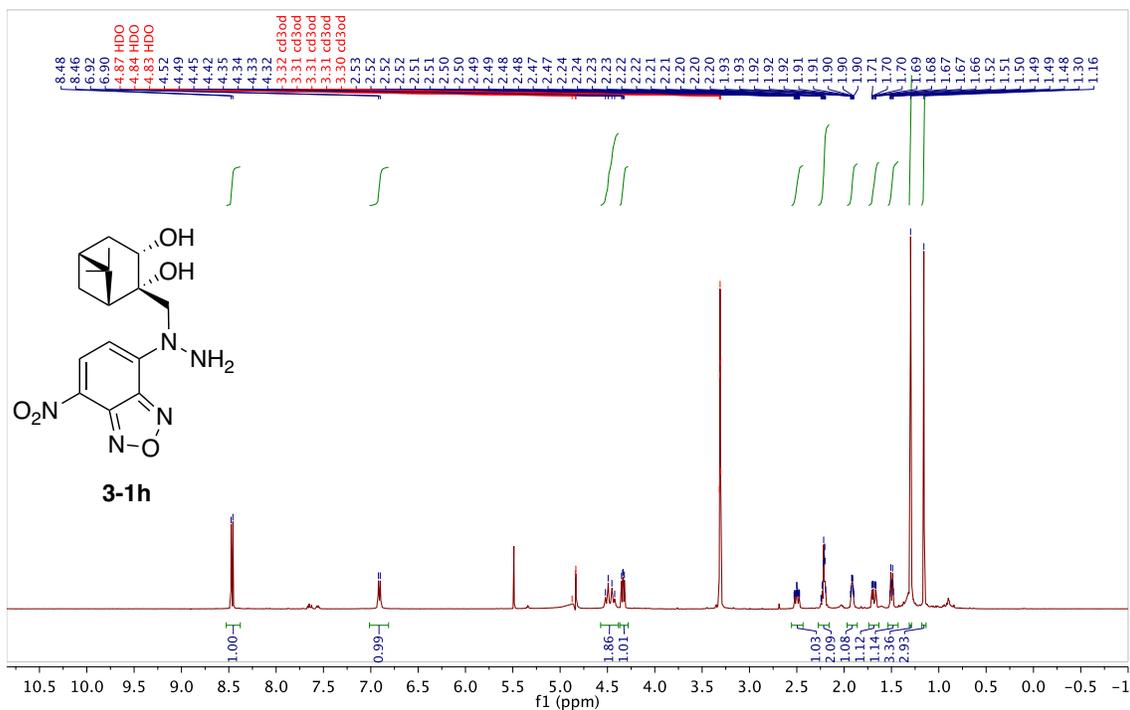


BAH-07-51; 125.688 MHz C13[H1] 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

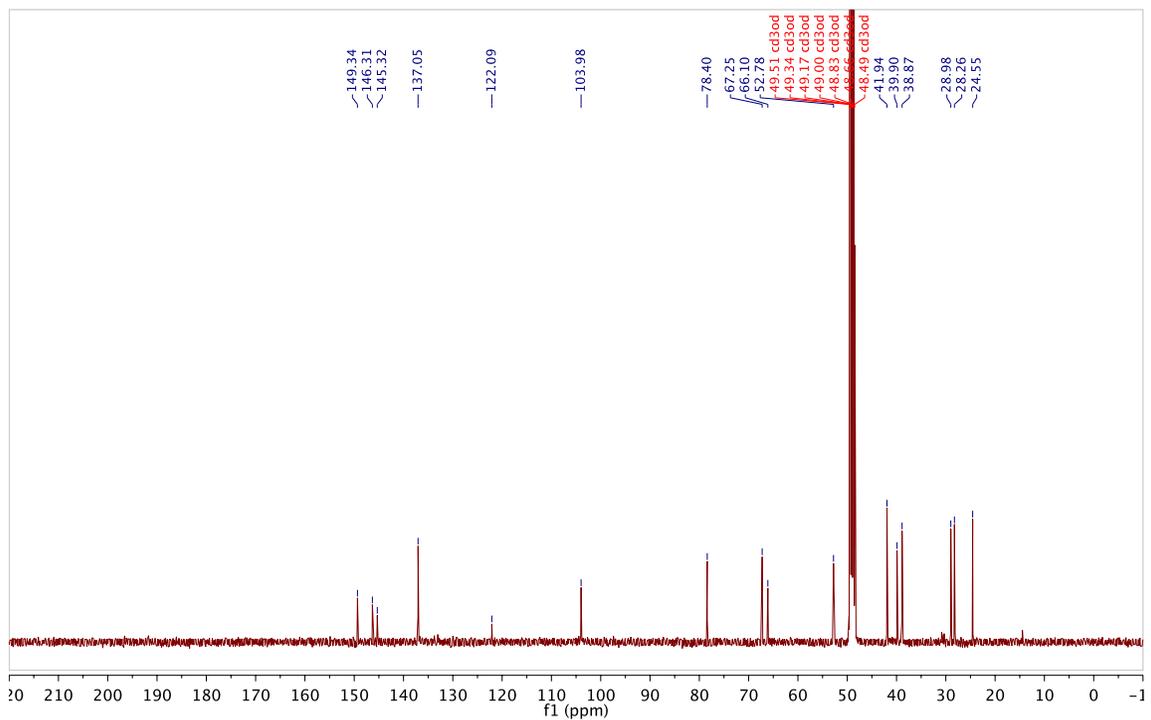


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1h in $\text{CD}_3\text{OD}$ at 25 °C

BAH-08-86; 499.799 MHz H1 1D in cd3od (ref. to CD3OD @ 3.30 ppm) temp 27.7 C -> actual temp = 27.0 C, coldlual probe

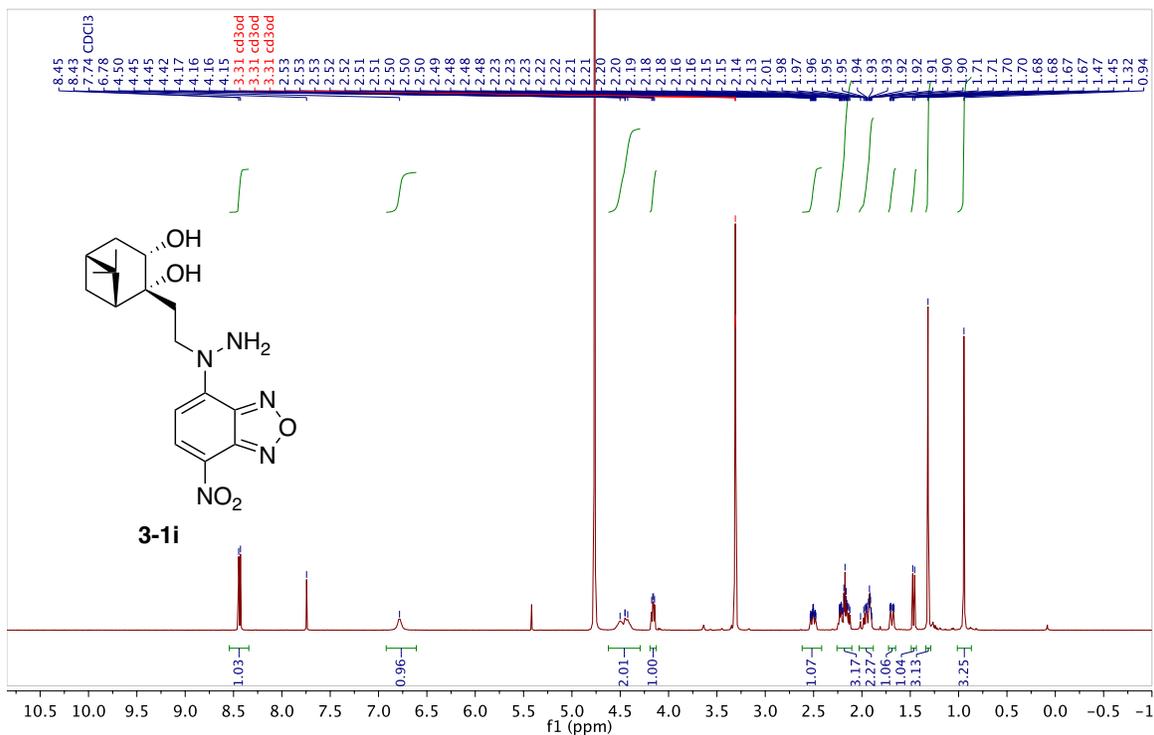


BAH-08-86 C-NMR; 125.688 MHz C13{H1} 1D in cd3od (ref. to CD3OD @ 49.0 ppm) temp 27.7 C -> actual temp = 27.0 C, coldlual probe

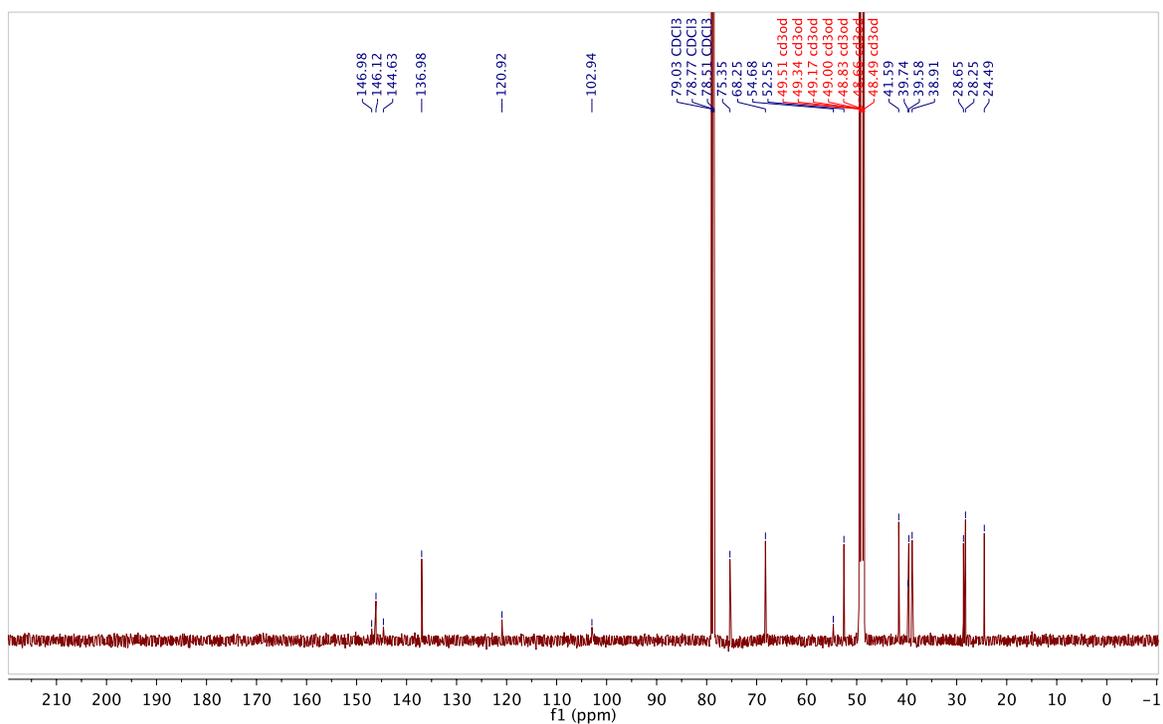


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1i in $\text{CD}_3\text{OD}$ at 25 °C

BAH-08-85; 499.799 MHz H1 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 3.30 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe



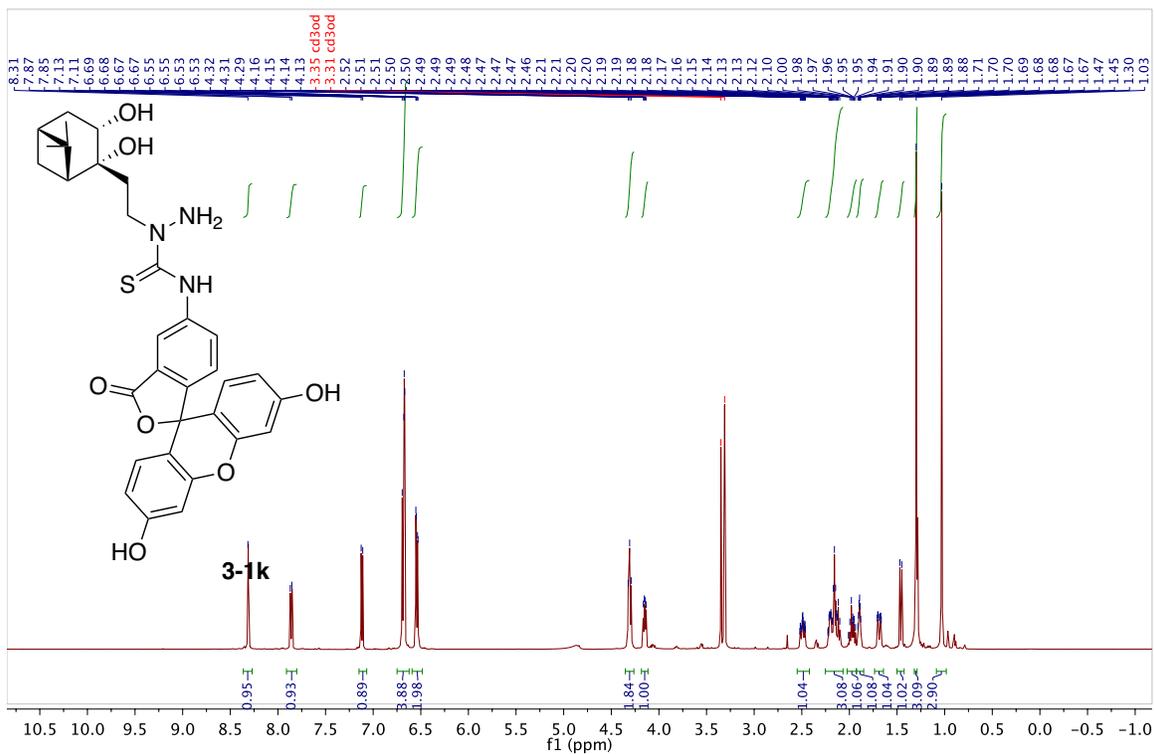
BAH-08-85; 125.688 MHz C13{H1} 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 49.0 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe



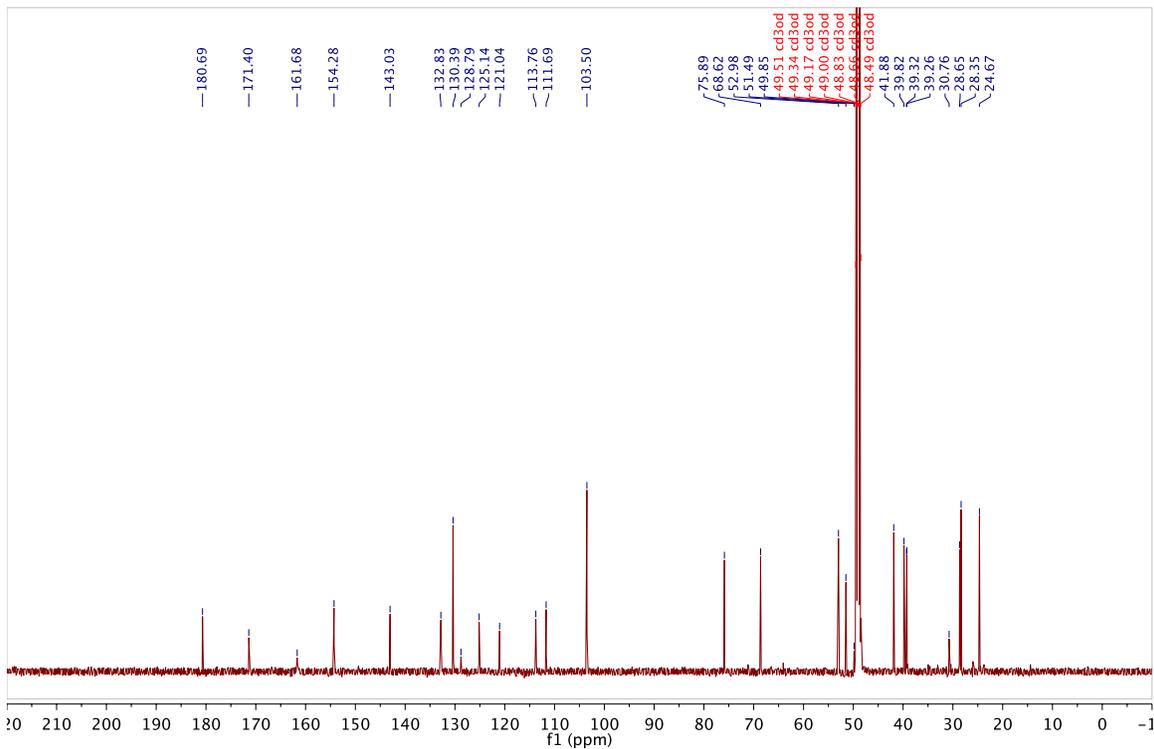


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1k in $\text{CD}_3\text{OD}$ at 25 °C

BAH-09-62; 499.799 MHz H1 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 3.30 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe

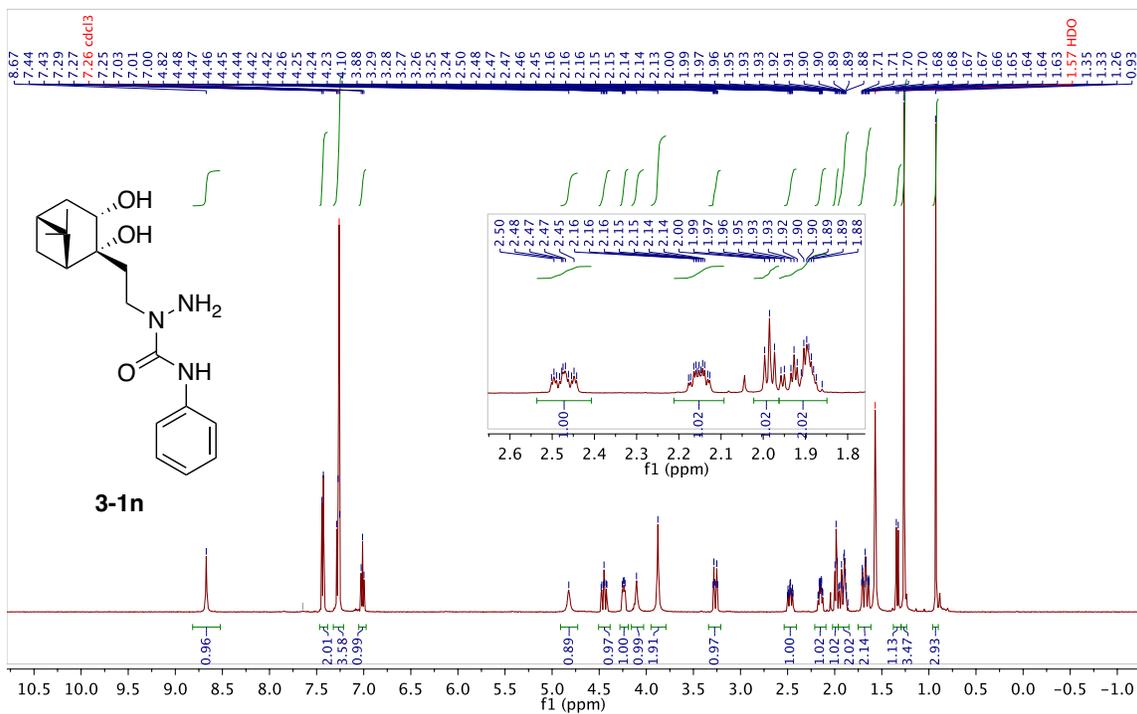


BAH-09-62; 125.688 MHz C13{H1} 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 49.0 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe

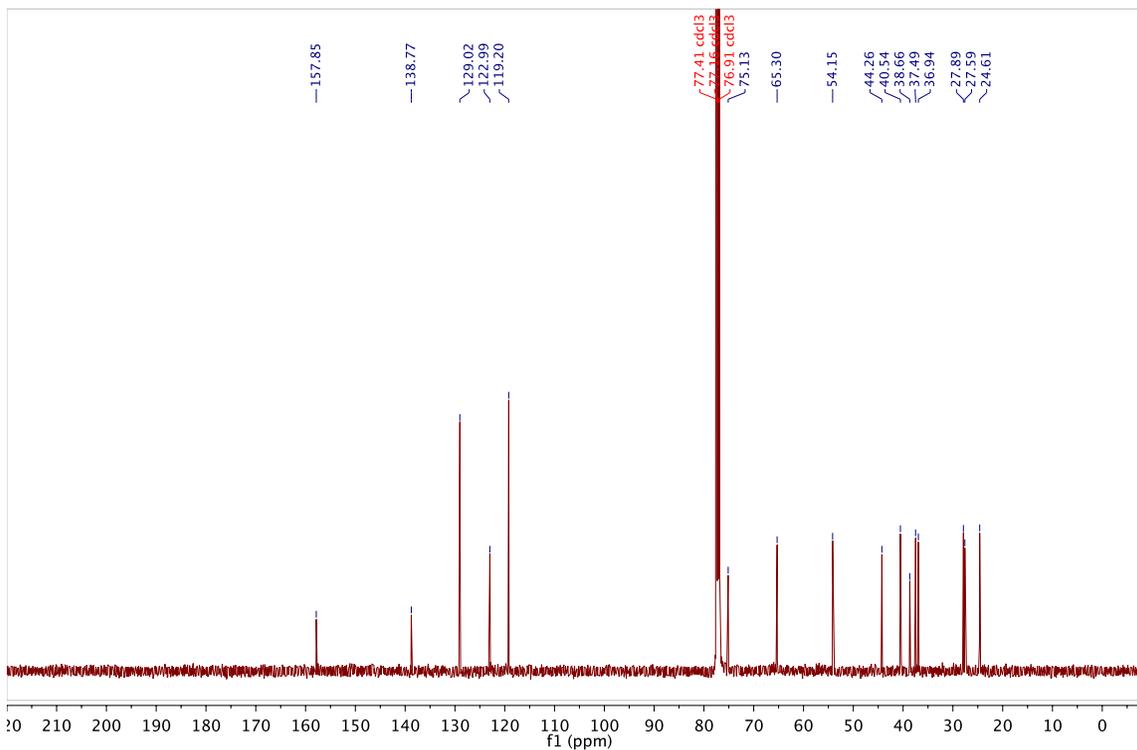


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1n in $\text{CDCl}_3$ at 25 °C

BAH-09-49\_for\_characterization; 499.797 MHz H1 1D in cdcl3 (ref. to CDCl3 @ 7.26 ppm) temp 27.7 C -> actual temp = 27.0 C, colddual probe

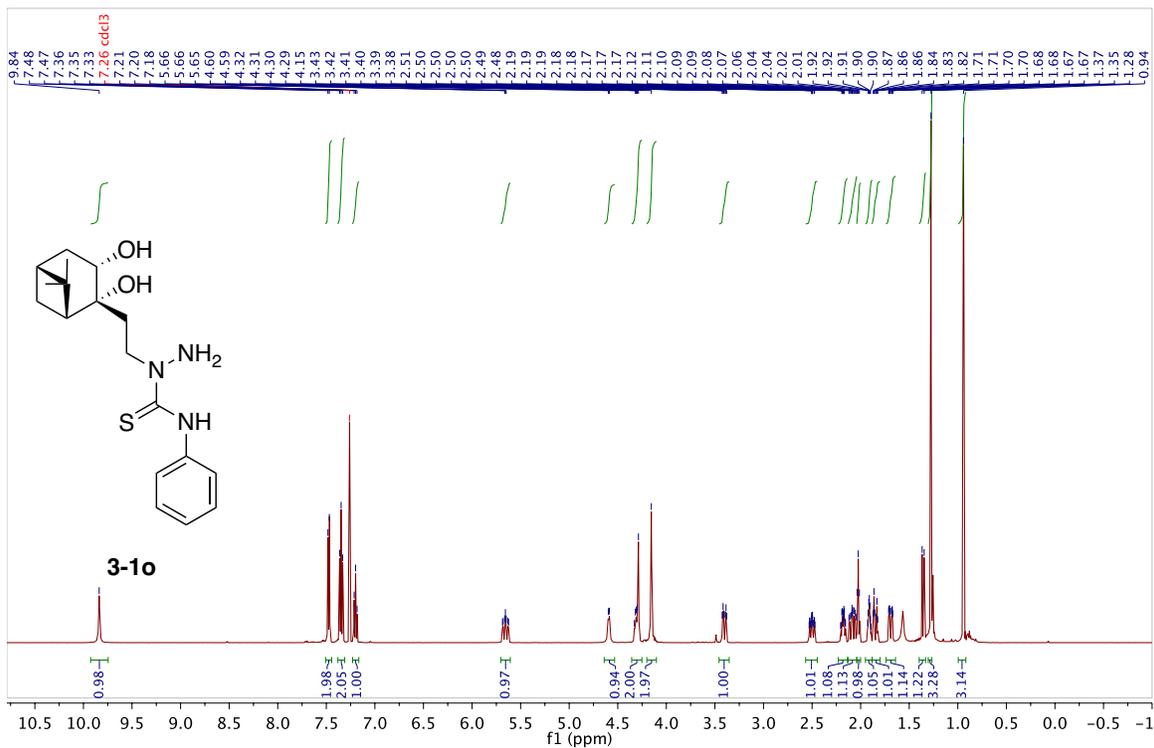


BAH-09-49; 125.688 MHz C13{H1} 1D in cdcl3 (ref. to CDCl3 @ 77.06 ppm) temp 27.7 C -> actual temp = 27.0 C, colddual probe

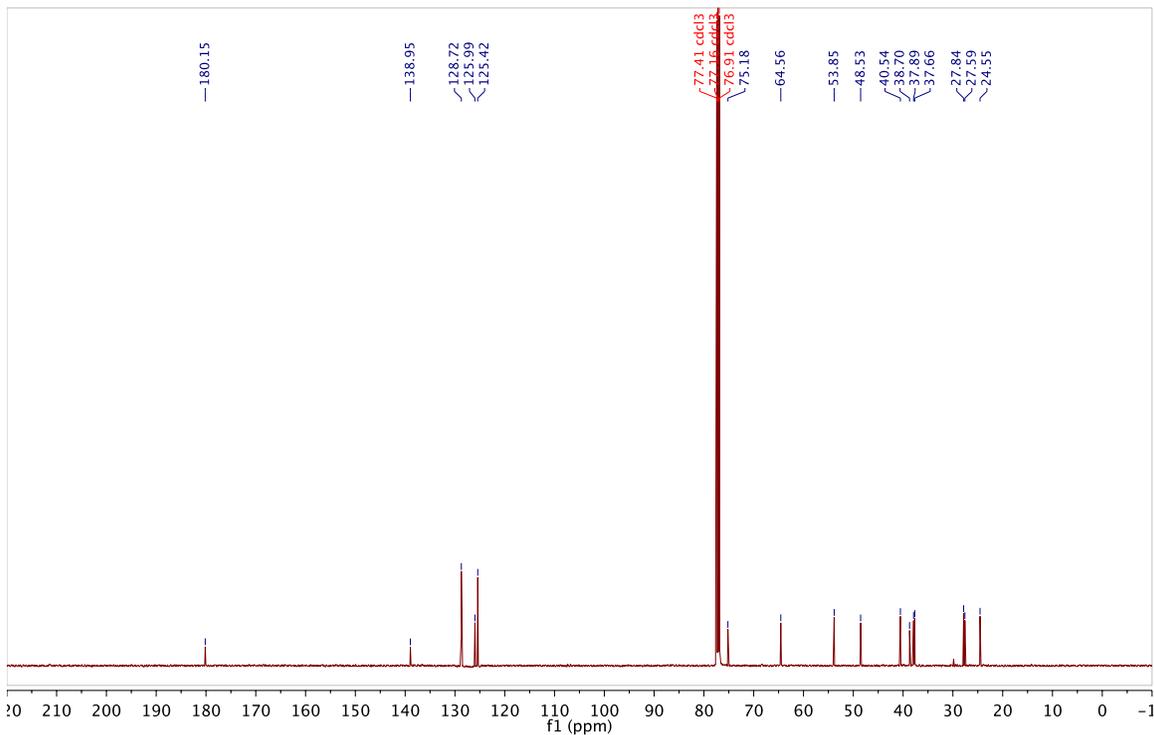


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-1o in $\text{CDCl}_3$ at 25 °C

BAH-09-69; 499.797 MHz H1 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 7.26 ppm) temp 27.7 C -> actual temp = 27.0 C, coldualdual probe

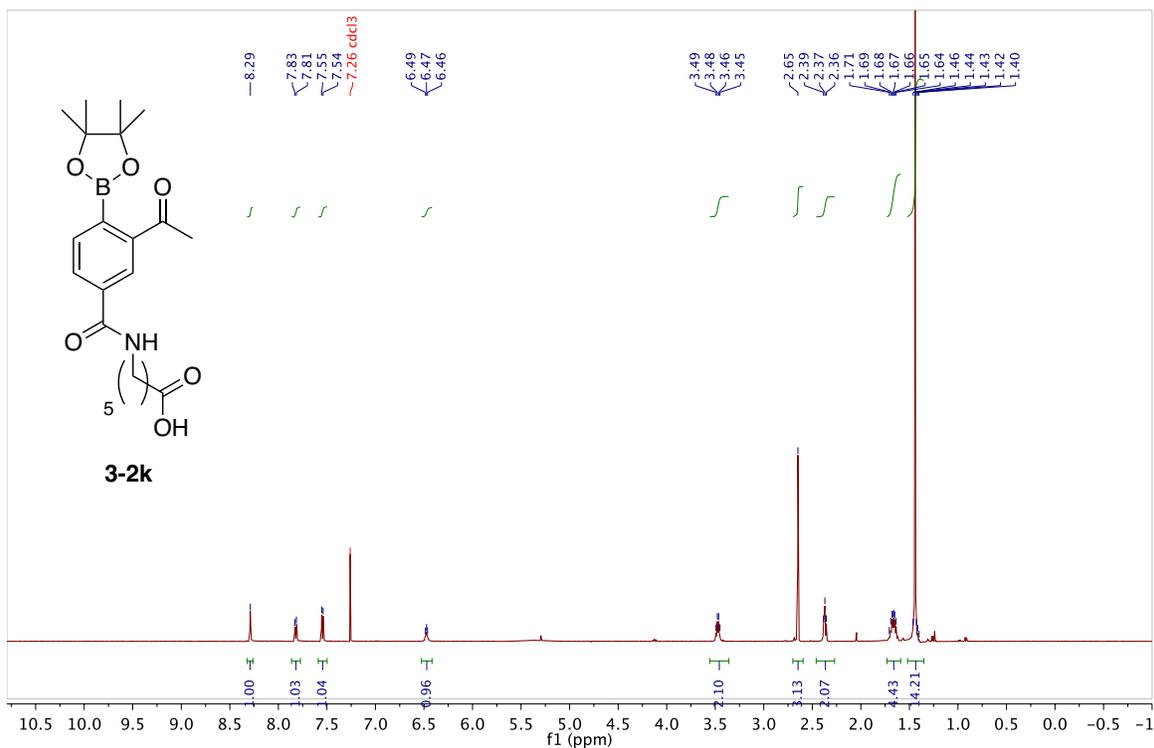


BAH-09-69; 125.688 MHz C13{H1} 1D in cdcl3 (ref. to  $\text{CDCl}_3$  @ 77.06 ppm) temp 27.7 C -> actual temp = 27.0 C, coldualdual probe

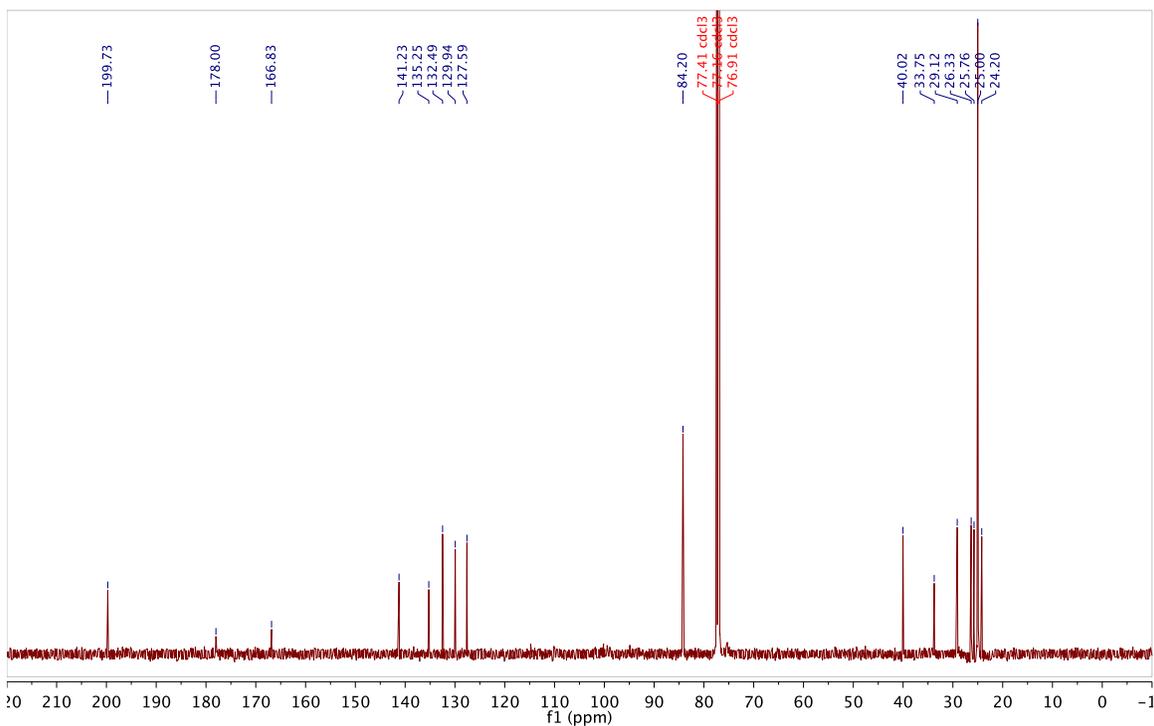


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-2k in $\text{CDCl}_3$ at 25 °C

BAH-09-102; 498.120 MHz H1 1D in  $\text{CDCl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm) temp 26.9 C -> actual temp = 27.0 C, autotx probe

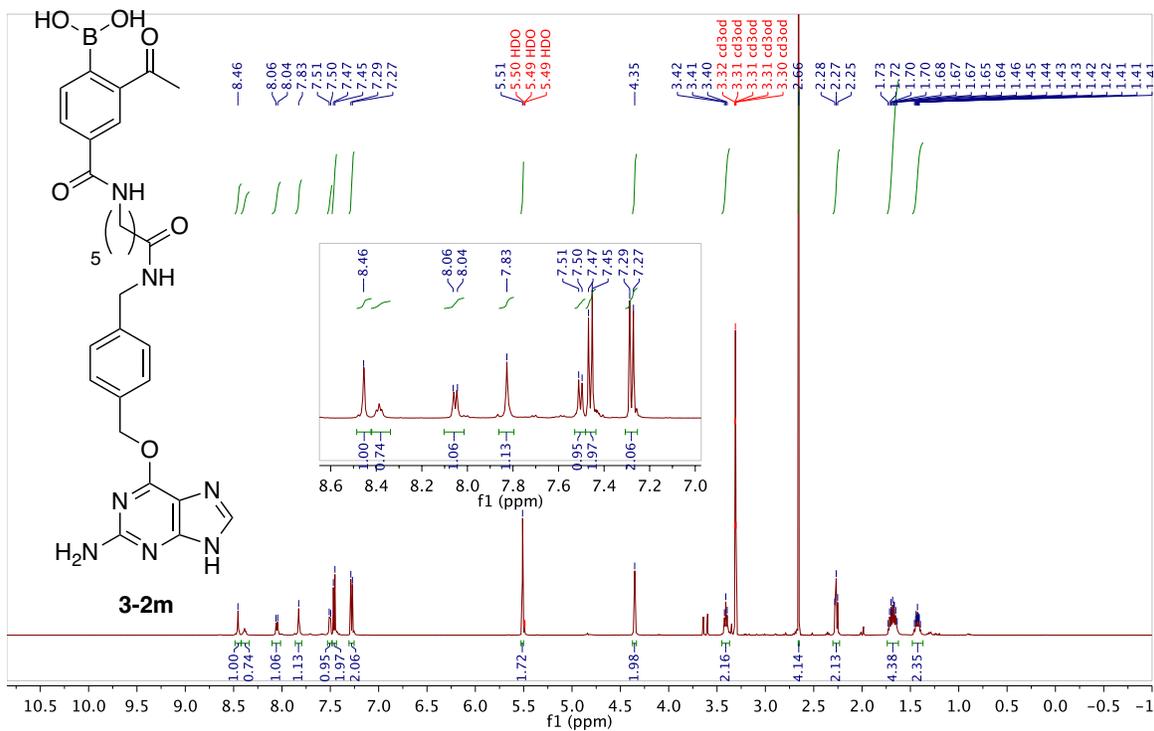


BAH-09-102; 125.688 MHz C13{H1} 1D in  $\text{cdCl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe

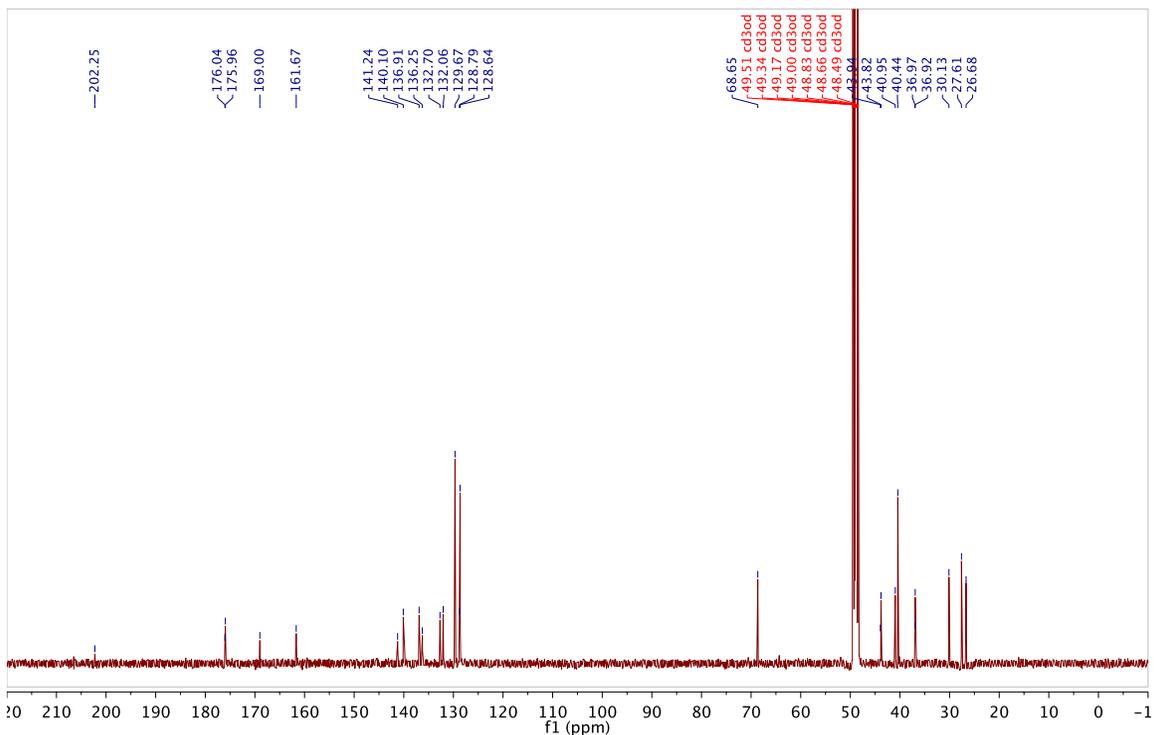


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 3-2m in $\text{CD}_3\text{OD}$ at 25 °C

BAH-09-101; 499.799 MHz H1 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 3.30 ppm) temp 27.7 C -> actual temp = 27.0 C, coldual probe

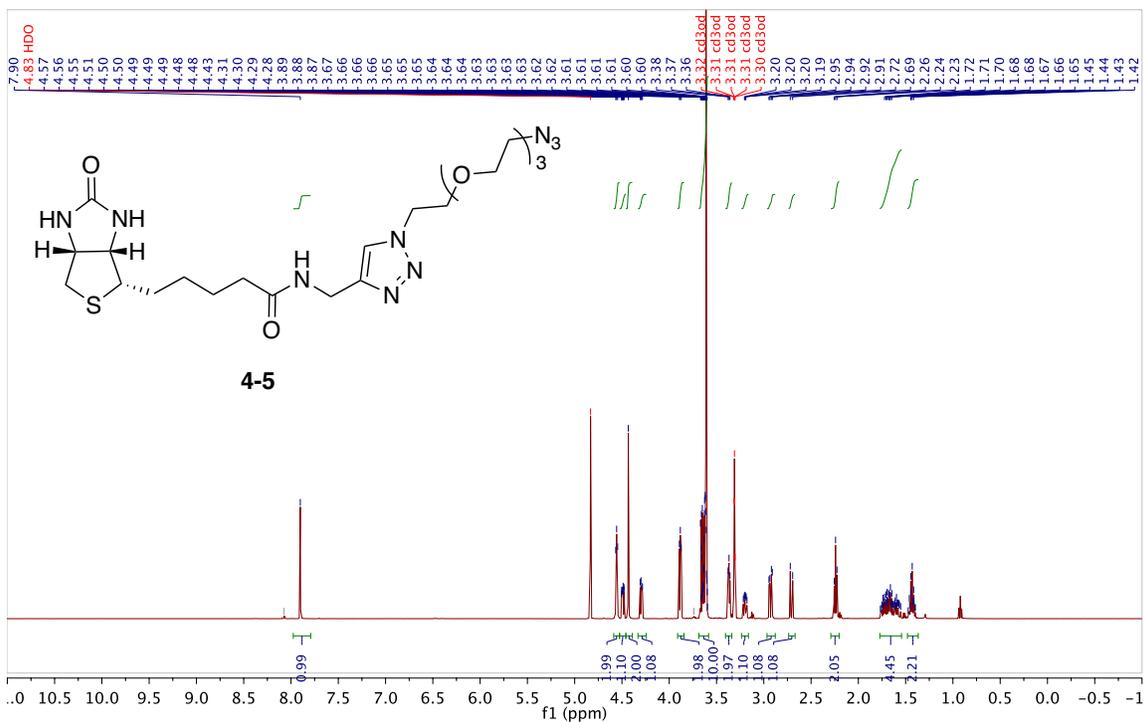


BAH-09-101\_for\_charac; 125.688 MHz C13{H1} 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 49.0 ppm) temp 27.7 C -> actual temp = 27.0 C, coldual probe

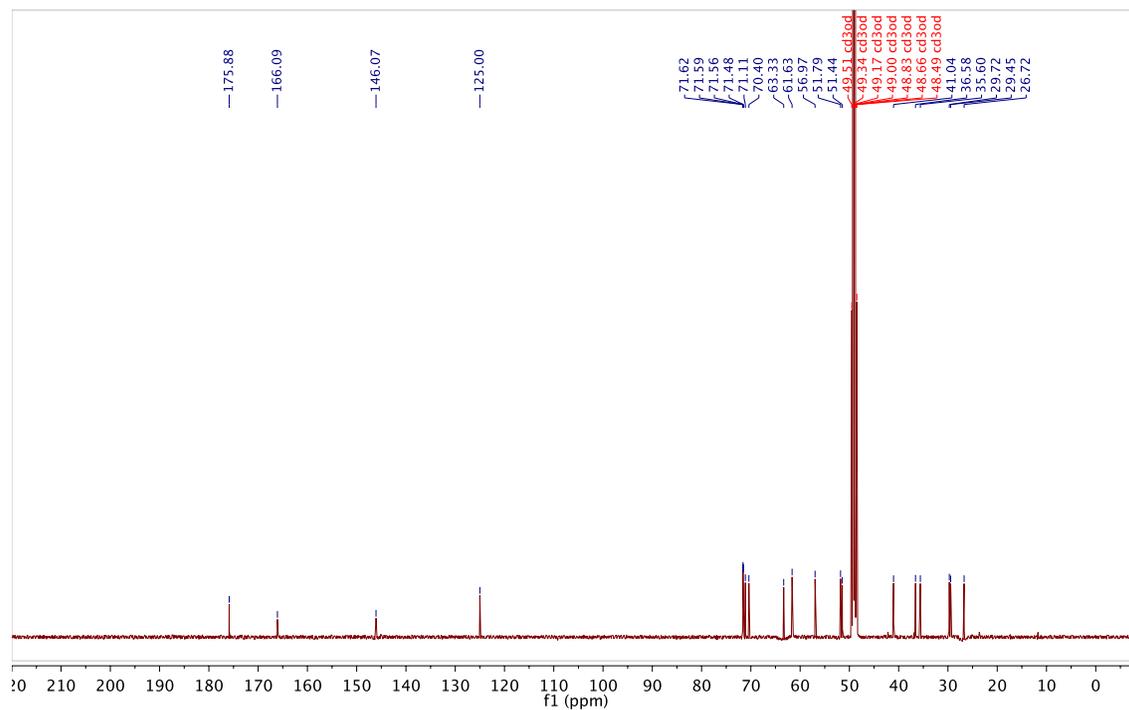


# <sup>1</sup>H and <sup>13</sup>C spectrum for 4-5 in CD<sub>3</sub>OD at 25 °C

BAH-06-48; 499.808 MHz H1 PRESAT in cd3od (ref. to CD3OD @ 3.30 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

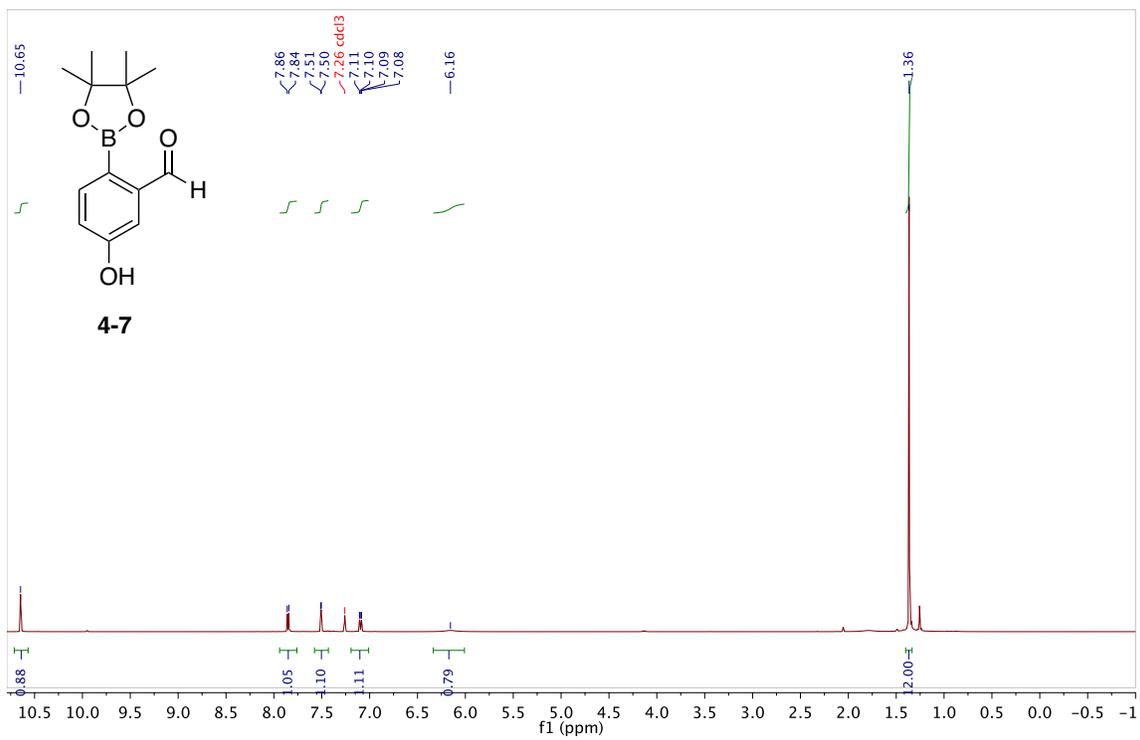


BAH-06-48; 125.691 MHz C13[H1] 1D in cd3od (ref. to CD3OD @ 49.0 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

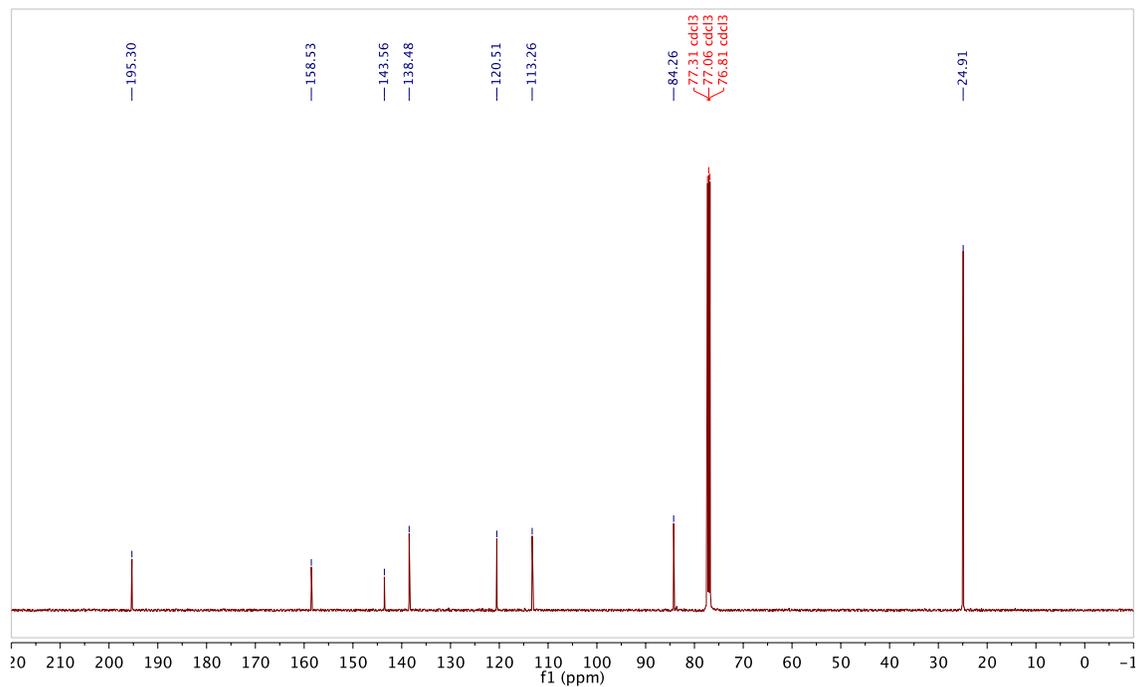


# <sup>1</sup>H and <sup>13</sup>C spectrum for 4-7 in CDCl<sub>3</sub> at 25 °C

BAH-07-24; 499.797 MHz H1 PRESAT in cdcl3 (ref. to CDCl<sub>3</sub> @ 7.26 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

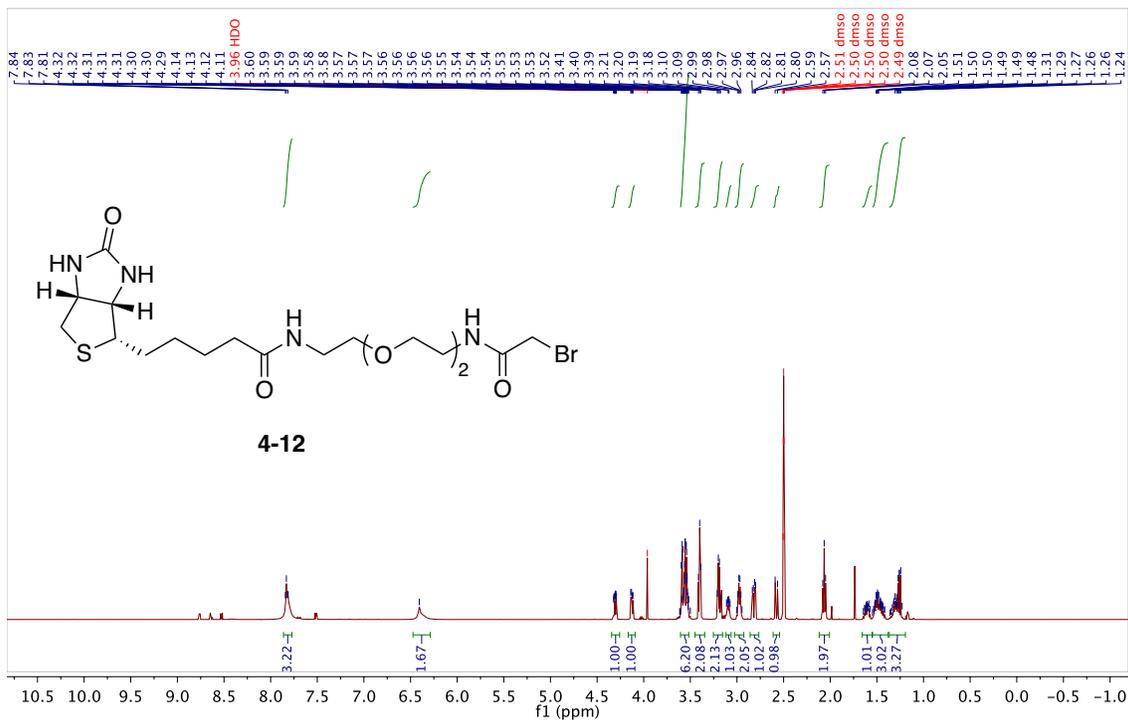


BAH-07-24; 125.688 MHz C13[H1] 1D in cdcl3 (ref. to CDCl<sub>3</sub> @ 77.06 ppm), temp 27.7 C -> actual temp = 27.0 C, cold dual probe

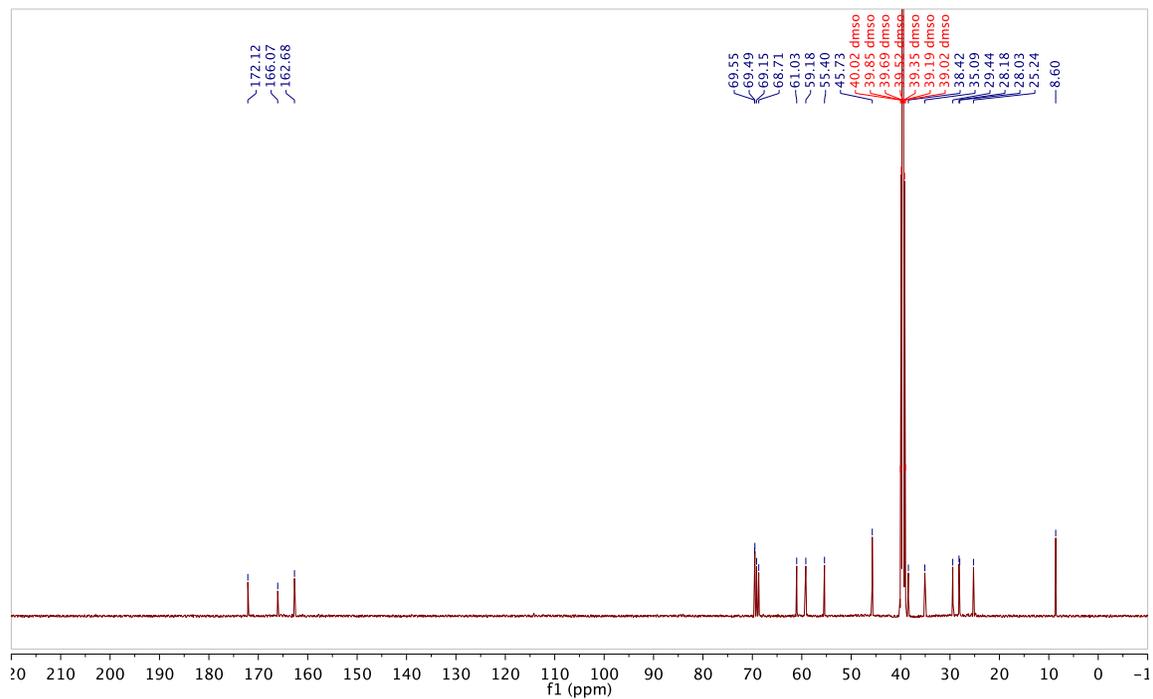


# <sup>1</sup>H and <sup>13</sup>C spectrum for 4-12 in DMSO-d<sub>6</sub> at 25 °C

BAH-07-156; 499.800 MHz H1 1D in dms0 (ref. to DMSO @ 2.49 ppm) temp 27.7 C -> actual temp = 27.0 C, coldludal probe

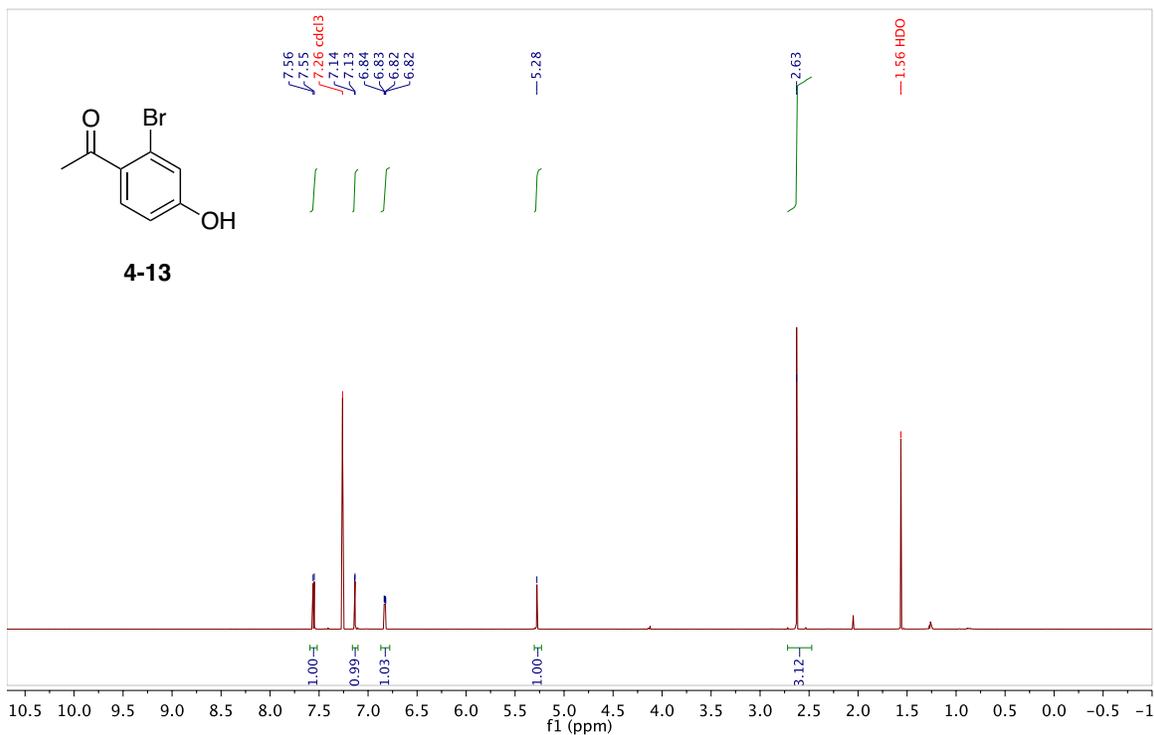


BAH-07-156; 125.689 MHz C13{H1} 1D in dms0 (ref. to DMSO @ 39.5 ppm) temp 27.7 C -> actual temp = 27.0 C, coldludal probe

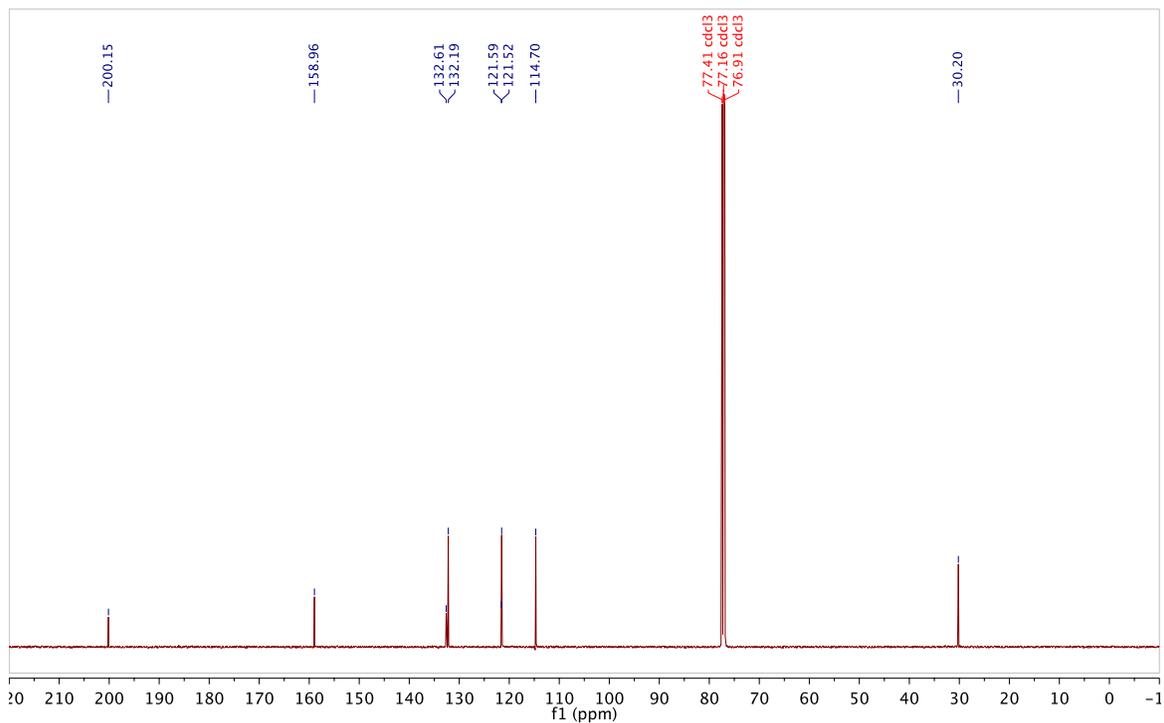


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 4-13 in $\text{CDCl}_3$ at $25^\circ\text{C}$

BAH-07-189; 699.762 MHz H1 PRESAT in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 7.26 ppm), temp 27.5 C -> actual temp = 27.0 C, coldid probe

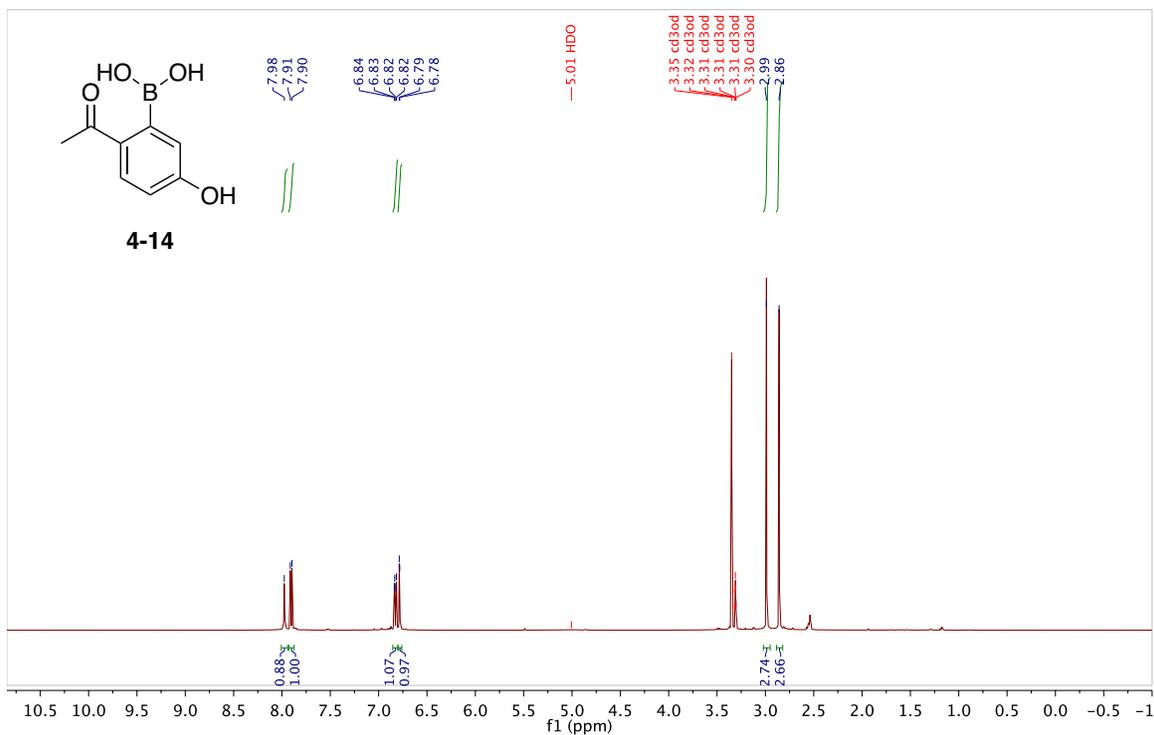


BAH-07-189-C-NMR; 125.688 MHz C13{H1} 1D in  $\text{cdcl}_3$  (ref. to  $\text{CDCl}_3$  @ 77.06 ppm) temp 27.7 C -> actual temp = 27.0 C, coldual probe

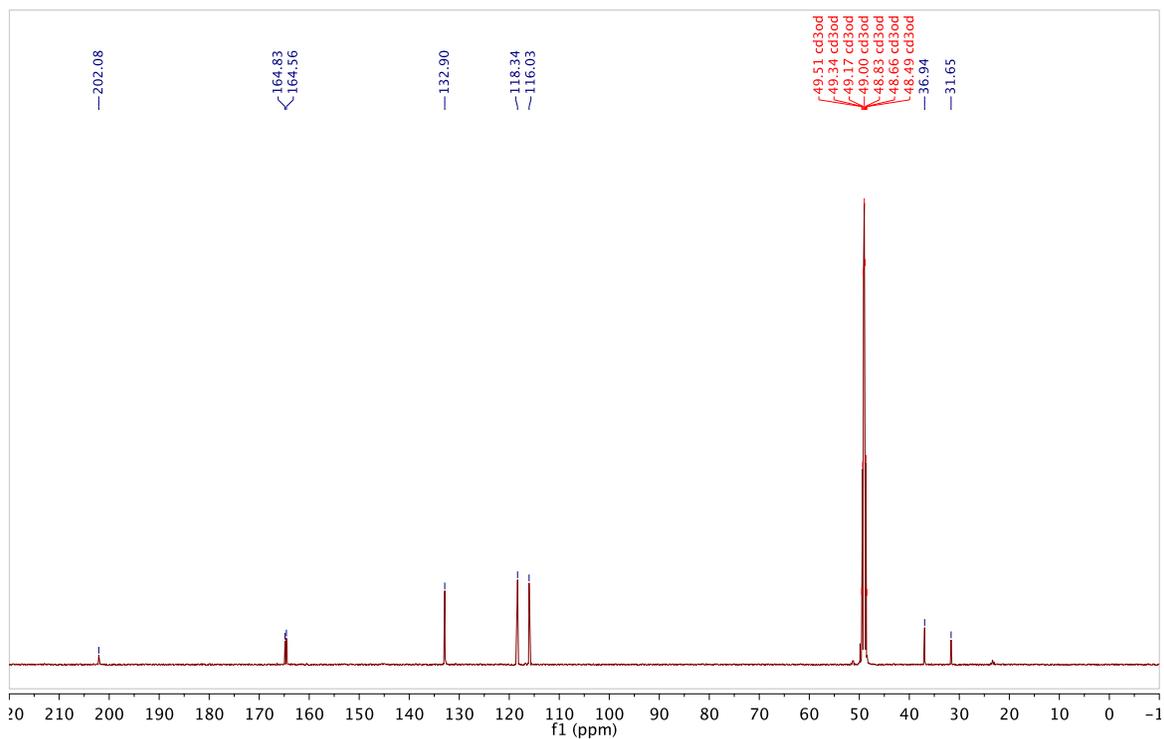


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 4-14 in $\text{CD}_3\text{OD}$ at $25^\circ\text{C}$

BAH-08-82; 499.799 MHz H1 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 3.30 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe



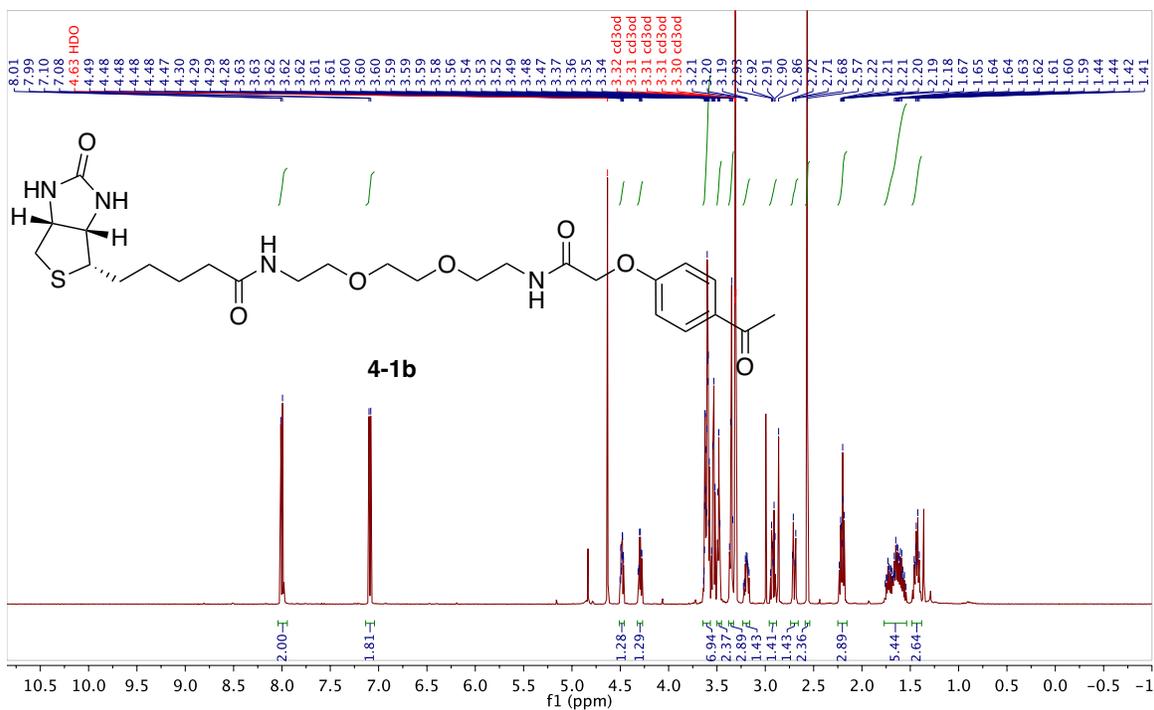
BAH-08-82; 125.688 MHz C13{H1} 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 49.0 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe



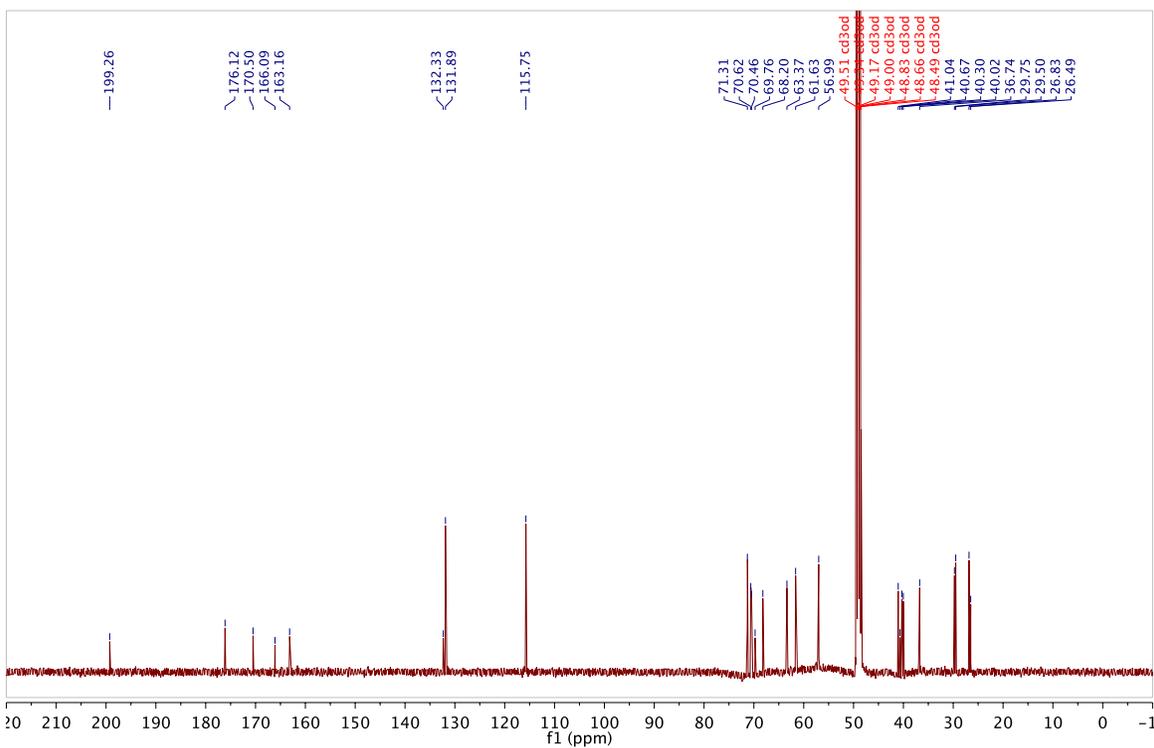


# $^1\text{H}$ and $^{13}\text{C}$ spectrum for 4-1b in $\text{CD}_3\text{OD}$ at $25^\circ\text{C}$

BAH-08-20; 499.799 MHz H1 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 3.30 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe



BAH-08-20; 125.688 MHz C13{H1} 1D in cd3od (ref. to  $\text{CD}_3\text{OD}$  @ 49.0 ppm) temp 27.7 C -> actual temp = 27.0 C, cold dual probe





## Appendix 2: MatLab script Makefigure4-8.m

```
clear all; close all;

clc

Dir='';
File = 'SequencingDataForFigure4-8.txt';
%SAVEto = [File(1:end-4) '']; % keep blank if don't want to save

SET{1} = [58 59 60]; %
SET{2} = [61 62 63]; %
SET{3} = [64 65 66]; %

TEST_SET = 1;
CONTROL_SETS = [2 3];

HITS2DISPLAY = 50; % maximum numer of hits to display
SHOWaminoACIDS = [18 19 20 21 22];

CLUSTERbyH = 1; % 1 if you want your hits to be clustered by Hamming dist.

PLOT_VOLCANO = 1; % set to 1 if you want to see the actual volcano plot

%%%%%%%%% volcano plot parameters here %%%%%%%%%%
p_cutoff = 0.1; % p-value cutoff (use 0.9 if dont care abt p)
R_cutoff = 3; % ratio cutoff
MaxX=14; % maximum on the X-scale (if plotting volcano)
vert_cutoff = 0.00001; % maximum on the Y-scale (if plotting volcano)
%%%%%%%%%

%%%%%%%%% do not change things beyond this point %%%%%%%%%%
%%%%%%%%% unless you know what you are doing %%%%%%%%%%
REREAD=1;

if REREAD

[Nuc, AA, Fr] = readMulticolumn('Dir', Dir, 'File', File, ...
                               'column', 1:max(cell2mat(SET)),...
                               'skip', 2, 'output', 'normalized+1');

end

%%%%%%%%%
%%%%%%%%%
%%%%%%%%%
```

```

% select only the aminoacids you want to see
CAA = char(AA);
AA=cellstr(CAA(:,SHOWaminoACIDS));

SQUARE=zeros(size(Fr,1),1);
IX=zeros(size(Fr,1),numel(CONTROL_SETS));

i=0;

for j=CONTROL_SETS
    i=i+1;
    ratio(:,i) = mean(Fr(:,SET{TEST_SET}), 2) ./ mean(Fr(:,SET{j}), 2);

    [~,confi(:,i)] = ttest2(Fr(:,SET{TEST_SET}),'Fr(:,SET{j})',....
        p_cutoff,'both','unequal');

    IX(:,i) = (confi(:,i) <= p_cutoff) & (ratio(:,i) >= R_cutoff);

    SQUARE = SQUARE + ratio(:,i).^2;

    if PLOT_VOLCANO
        subplot(1,numel(CONTROL_SETS),i);

        plot(log2 (ratio(:,i)),...
            -log10(confi(:,i)),'d',...
            'MarkerSize',4,...
            'MarkerFaceColor',0.5*[1 1 1],...
            'MarkerEdgeColor',0.5*[1 1 1]); hold on;

        plot( log2 (ratio(find(IX(:,i)),i)),...
            -log10(confi(find(IX(:,i)),i)),'d',...
            'MarkerSize',4,...
            'MarkerFaceColor','r',...
            'MarkerEdgeColor','r'); hold on;

        line([log2(R_cutoff) MaxX],[-log10(p_cutoff) -log10(p_cutoff)]);
        line([ log2(R_cutoff)    log2(R_cutoff)],...
            [-log10(p_cutoff) -log10(vert_cutoff)]);

        xlim([-MaxX MaxX]);

    end
end

R2 = sqrt(SQUARE);

IXall = find( (sum(IX,2)==size(IX,2)) ); % hits that satisfy all criteria
%IXall = find( (sum(IX,2)>=5) ); %hits that satisfy 5 criteria

```

```

    hits    = char(AA(IXall,:));
    Rhits   = ratio(IXall,:);
    R2hits  = R2(IXall);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% this is part where hits are clustered by H-dist %%%%%%%%%%%%%%%

% figure(2)
% if size(hits,1)>3
%     Y = pdist(hits,'hamming');
%     Z = linkage(Y,'complete');
%     [H,T,perm] = dendrogram(Z,0,'colorthreshold',20);
%     set(H,'LineWidth',2)
%
%     for i =1:size(hits,1)
%         label{i} = i;
%     end
%     set(gca,'XTick', 1:1:size(hits,1), 'XTickLabel',label);
%
%     hits    = hits(perm,:);
%     Rhits   = Rhits(perm,:);
%     R2hits  = R2(perm);
%     IXall   = IXall(perm);
% end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% display all results as heat map
figure(3)

if size(IXall,1)>=HITS2DISPLAY
    N=HITS2DISPLAY; % display only the first or defined number of hits
else
    N=size(IXall,1); %display all
end

FrPPM = round(10^6*Fr); % convert normalized fraction frequency to PPM

imagesc( log10([FrPPM(IXall(1:N),:) ratio(IXall(1:N),:) ]+1) );

set(gca,'YTick', 1:1:N, 'YTickLabel',cellstr(hits(1:N,:)),'TickDir','out',...
    'FontName','Courier New','FontSize',14);
set(gca,'XTick', 1:1:size(Fr,2)+4, 'TickDir','out');
jet1=jet;
jet1(1,:)=[0.4 0.4 0.4];
colormap(jet1);
colorbar;

% generate a plain text table for saving or copy from command line
S = char(32*ones(size(hits,1),2));
L = [ S(:,1) char(124*ones(size(hits,1),1)) S(:,1) ];
F = FrPPM(IXall,:); % display frequency in ppm

toSave = [hits    S  ];

for i=1:numel(SET)

```

```

        for j=1:numel(SET{i})
            toSave = [toSave num2str(F(:,SET{i}(j))) S];
        end
        toSave = [toSave L];
    end
    toSave = [toSave S num2str(round(Rhits)) L];

    disp(toSave);

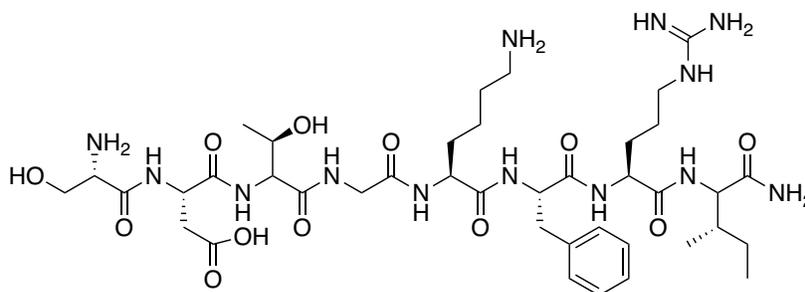
    if ~isempty(toSAVE)
        fs = fopen(fullfile(Dir,toSAVE),'w');
        RET = char(10*ones(size(toSave,1),1));
        fprintf( fs, '%s\r\n', [toSave RET]');
        fclose all;
    end

    % or you can just copy paste the results from the command line

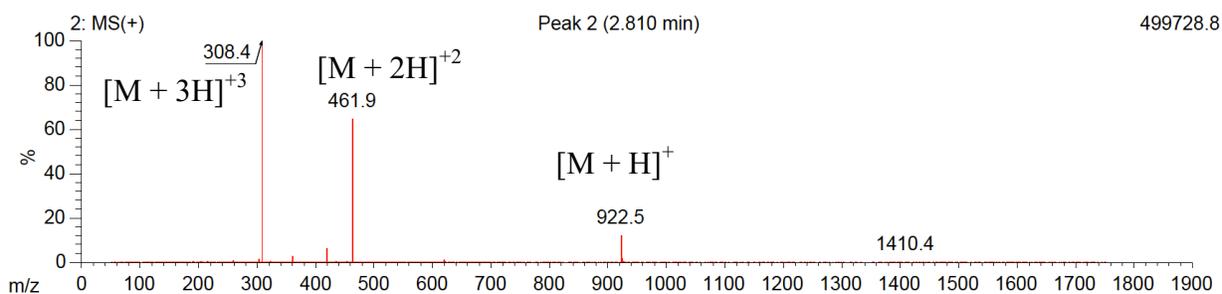
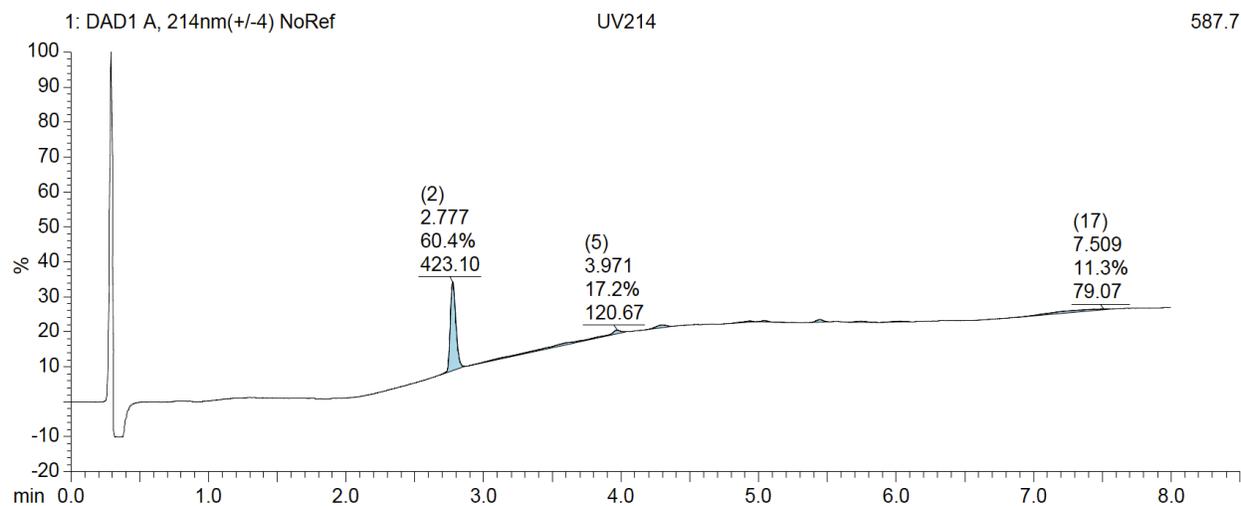
```

### Appendix 3: HRMS and HPLC-MS chromatograms of peptides 1–12

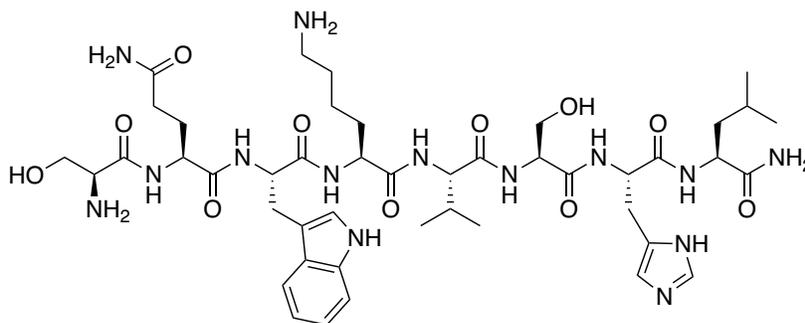
#### Peptide 1 – SDTGKFRI



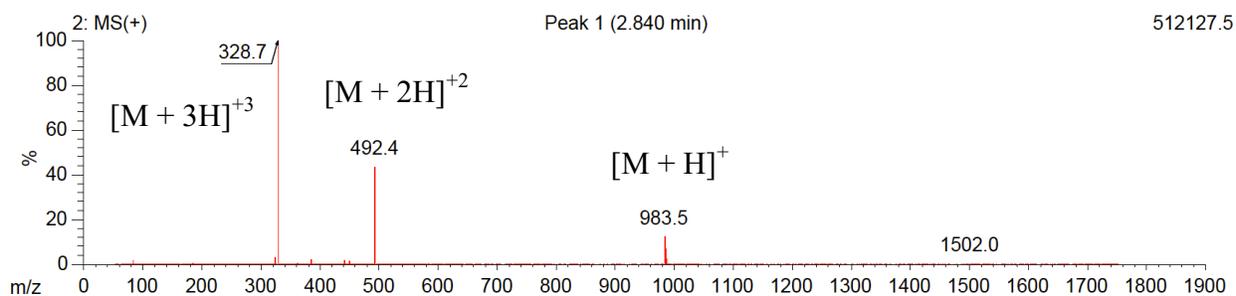
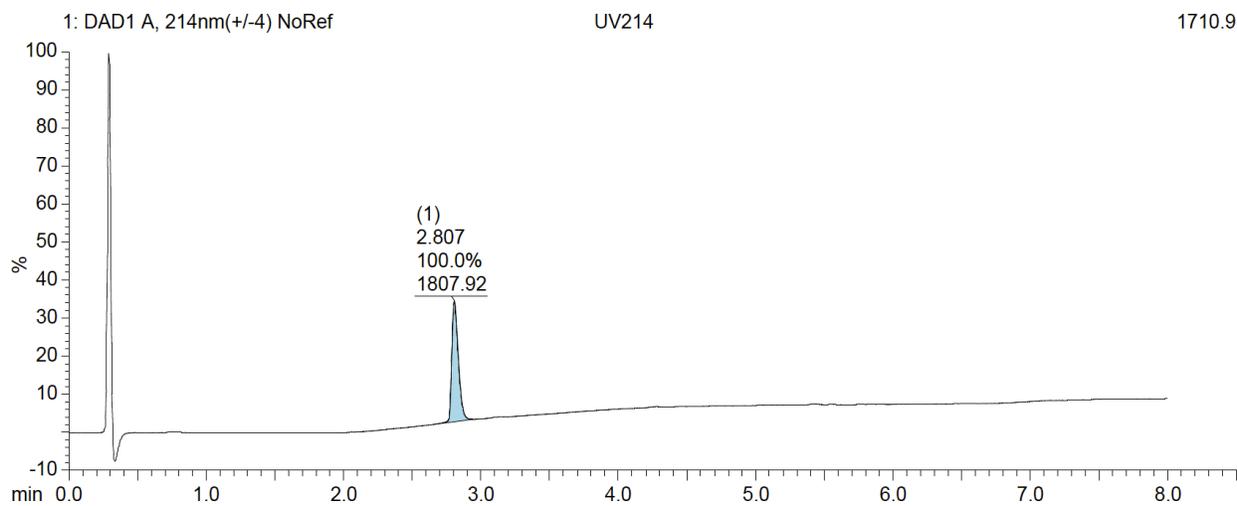
HRMS (ESI-TOF): for  $C_{40}H_{68}N_{13}O_{12}$  ( $M + H$ )<sup>+</sup>: *calcd.*: 922.5105; *found*: 922.5096.



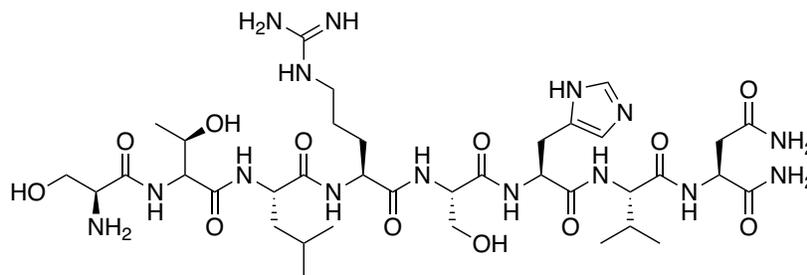
## Peptide 2 – SQWKVSHL



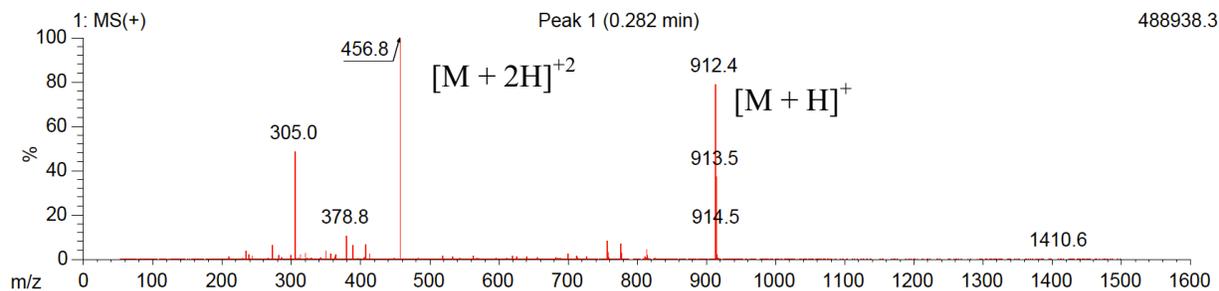
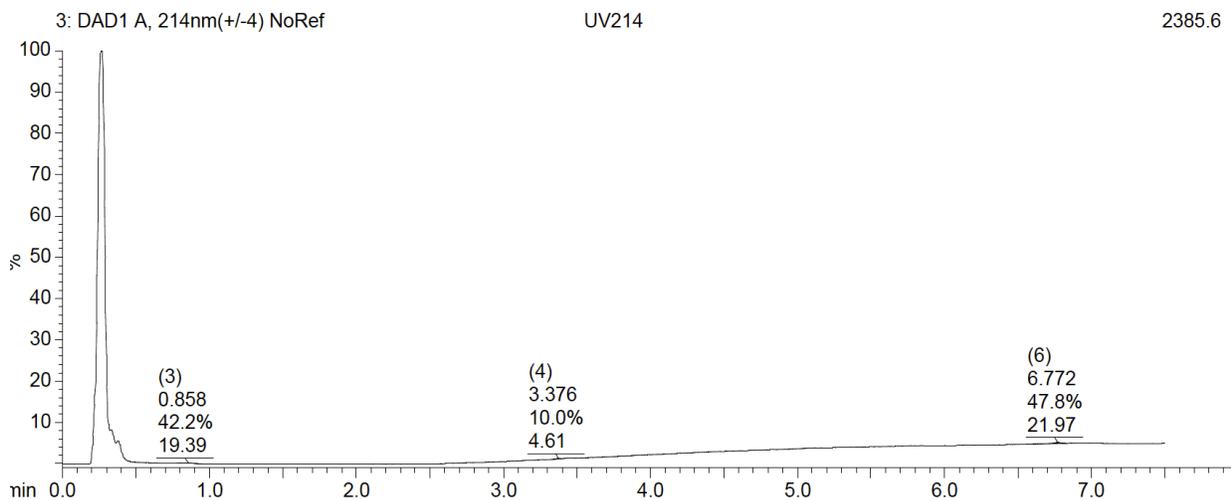
**HRMS (ESI-TOF):** for C<sub>45</sub>H<sub>71</sub>N<sub>14</sub>O<sub>11</sub> (M + H)<sup>+</sup>: *calcd.*: 983.5421; *found*: 983.5425.



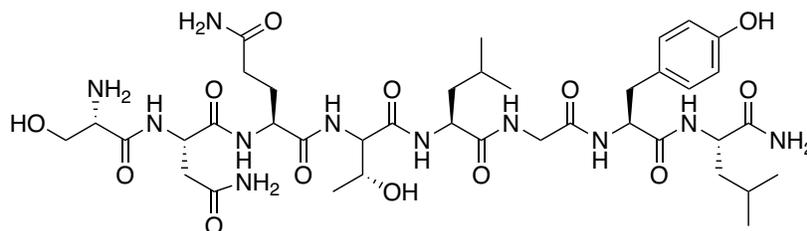
**Peptide 3 – STLRSHVN**



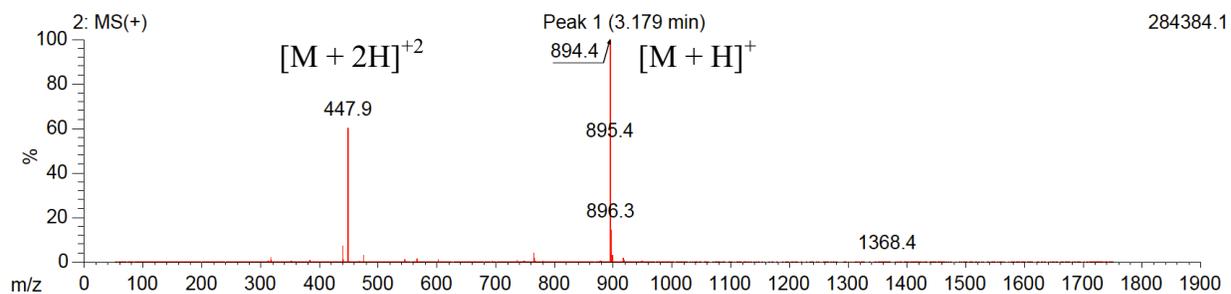
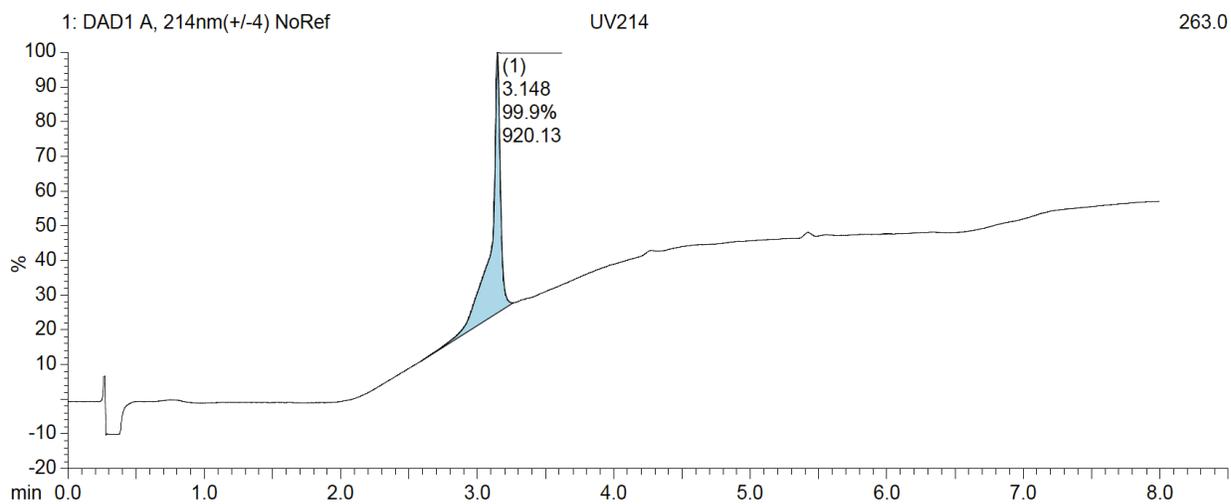
**HRMS (ESI-TOF):** for  $C_{37}H_{66}N_{15}O_{12}$  ( $M + H$ )<sup>+</sup>: *calcd.*: 912.501; *found*: 912.5005.



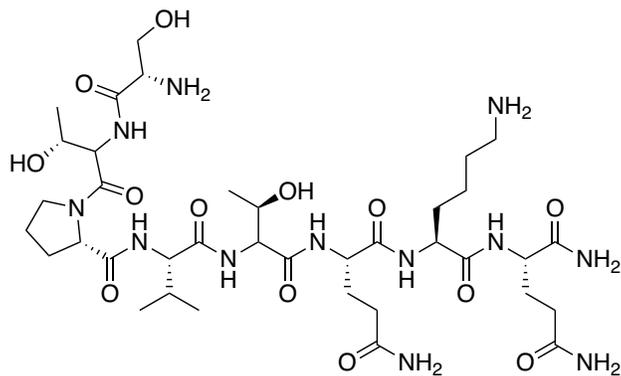
## Peptide 4 – SNQTLGYL



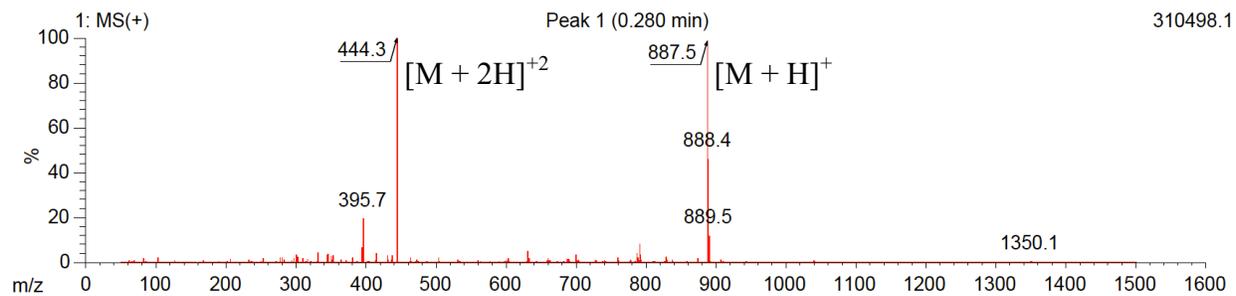
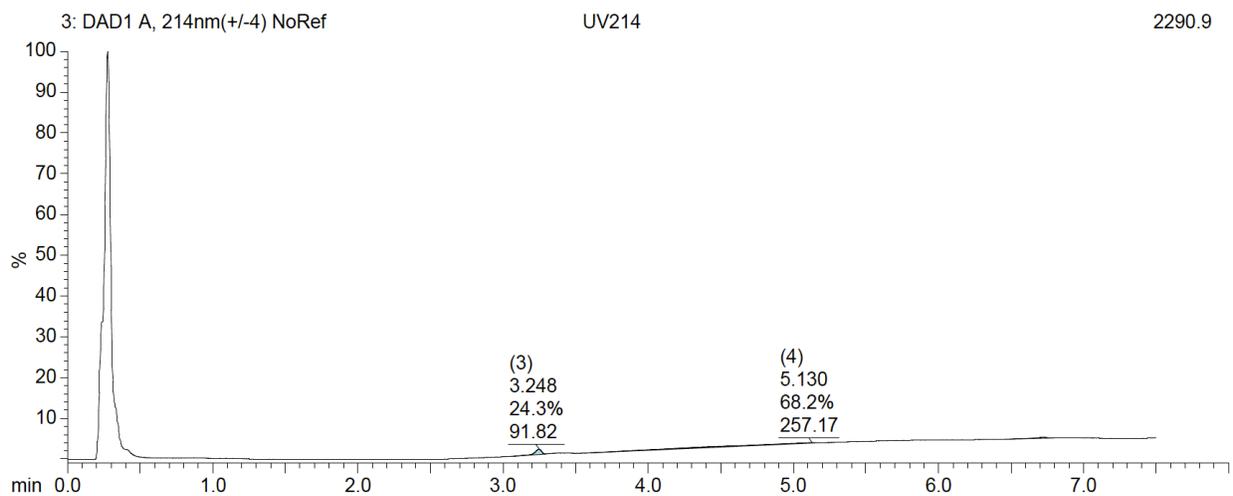
HRMS (ESI-TOF): for  $C_{39}H_{64}N_{11}O_{13}$  ( $M + H$ )<sup>+</sup>: *calcd.*: 894.468; *found*: 894.4681.



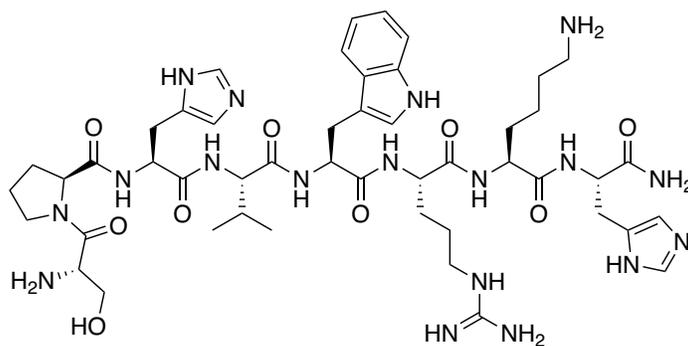
**Peptide 5 – STPVTQKQ**



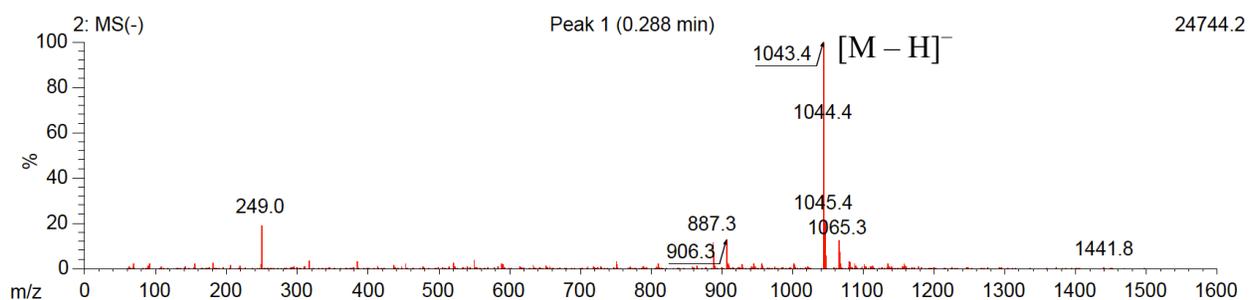
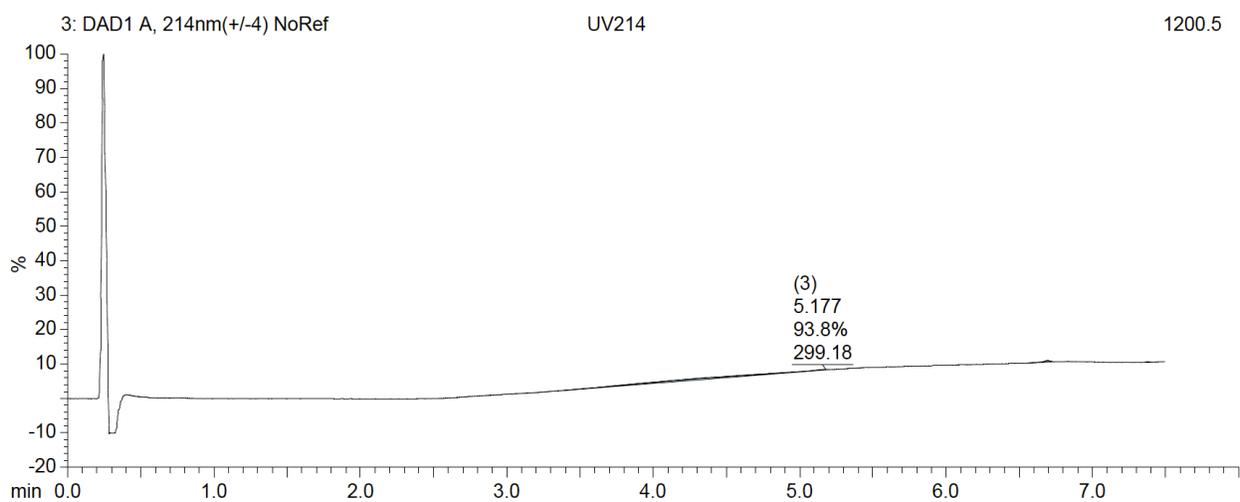
**HRMS (ESI-TOF):** for  $C_{37}H_{67}N_{12}O_{13}$  ( $M + H$ )<sup>+</sup>: *calcd.*: 887.4945; *found*: 887.4935.



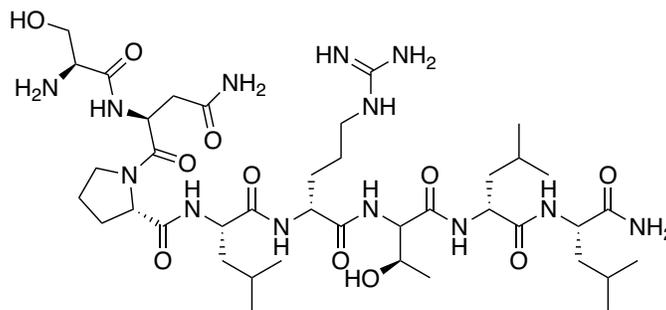
**Peptide 6 – SPHVWRKH**



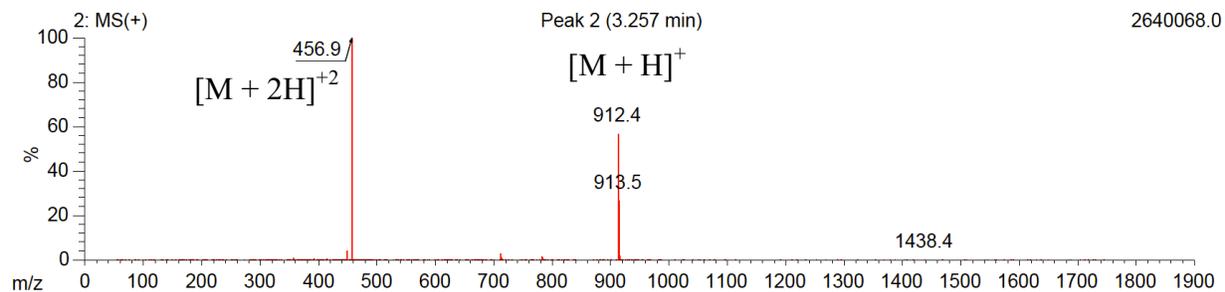
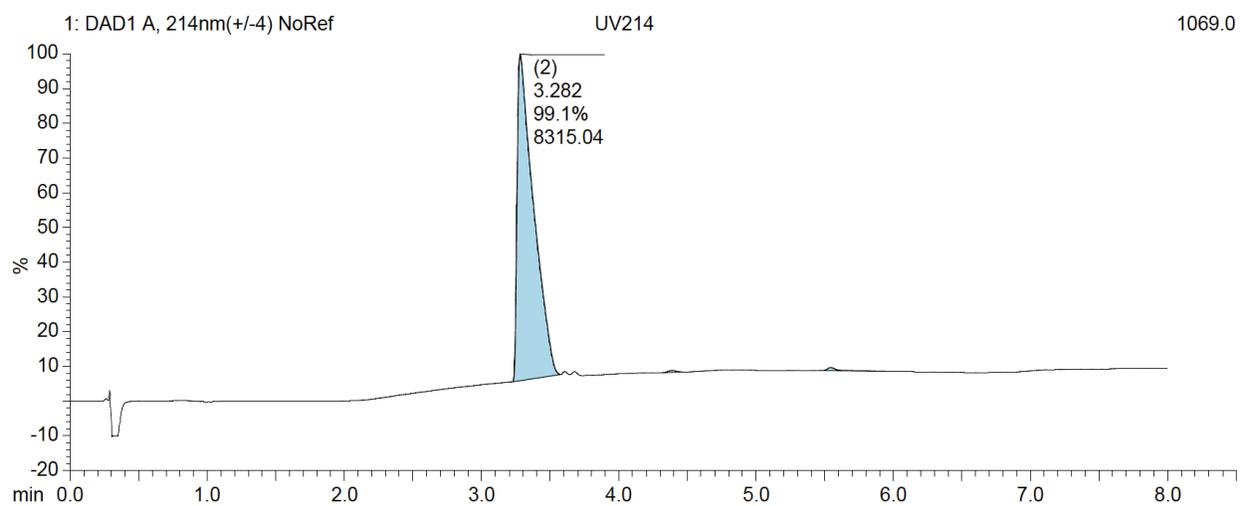
**HRMS (ESI-TOF):** for  $C_{48}H_{73}N_{18}O_9$  ( $M + H$ )<sup>+</sup>: *calcd.*: 1045.5802; *found*: 1045.5792.



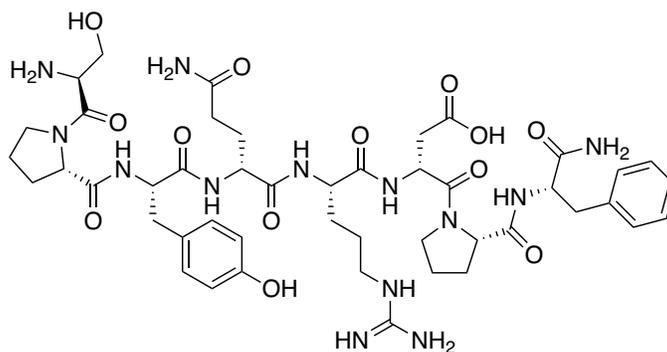
## Peptide 7 – SNPLRTLL



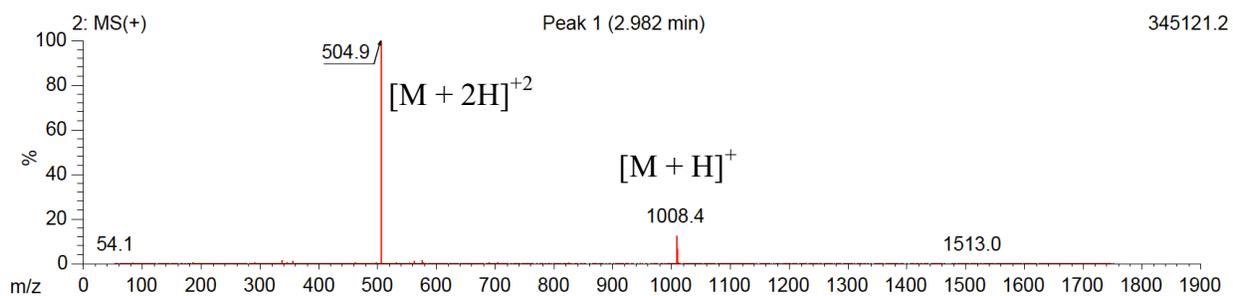
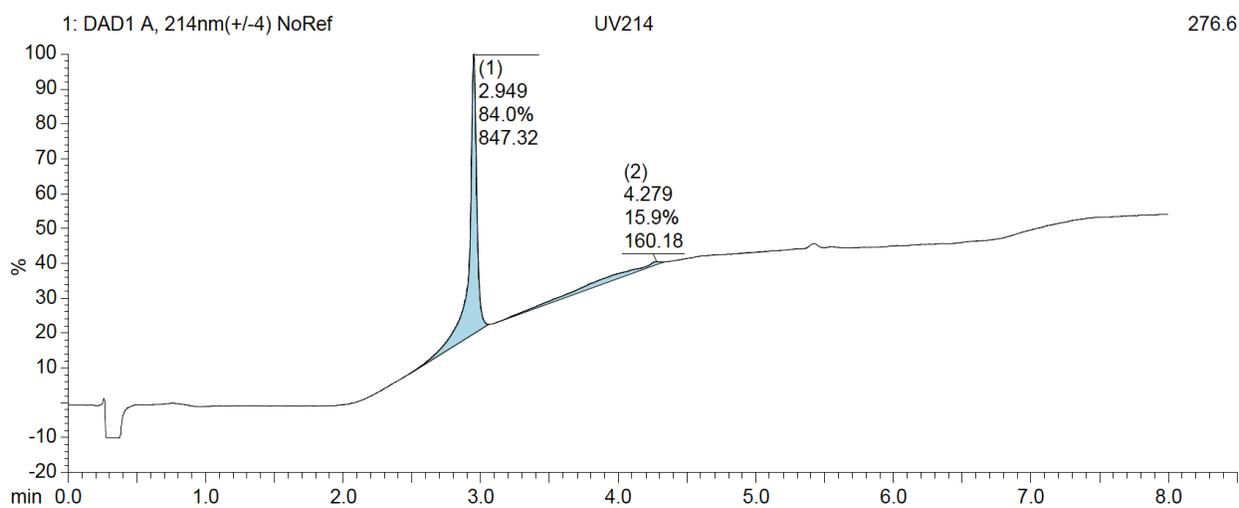
**HRMS (ESI-TOF):** for C<sub>40</sub>H<sub>74</sub>N<sub>13</sub>O<sub>11</sub> (M + H)<sup>+</sup>: *calcd.*: 912.5625; *found*: 912.5614.



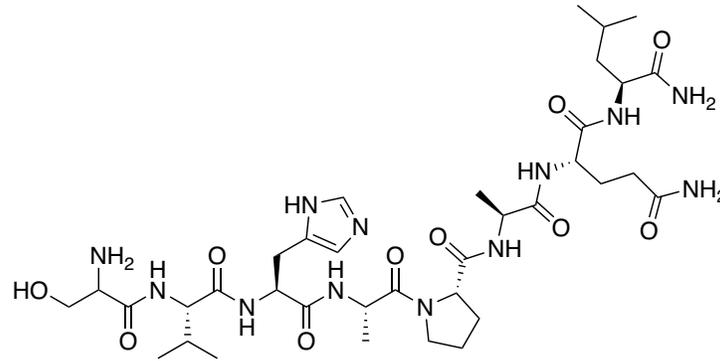
## Peptide 8 – SPYQRDPF



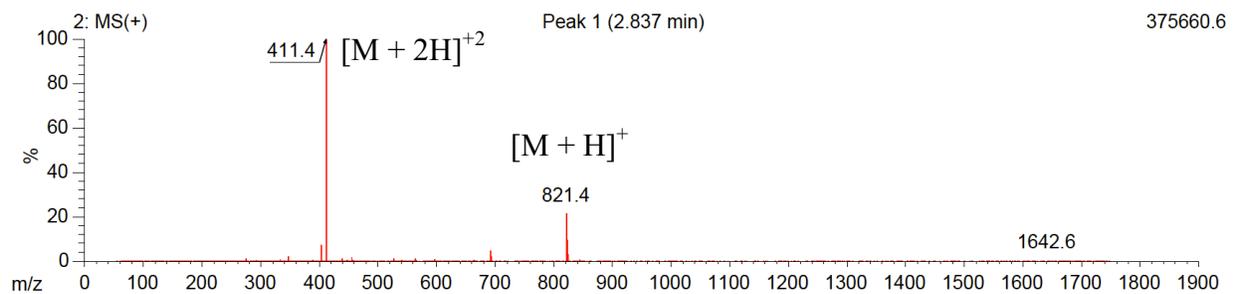
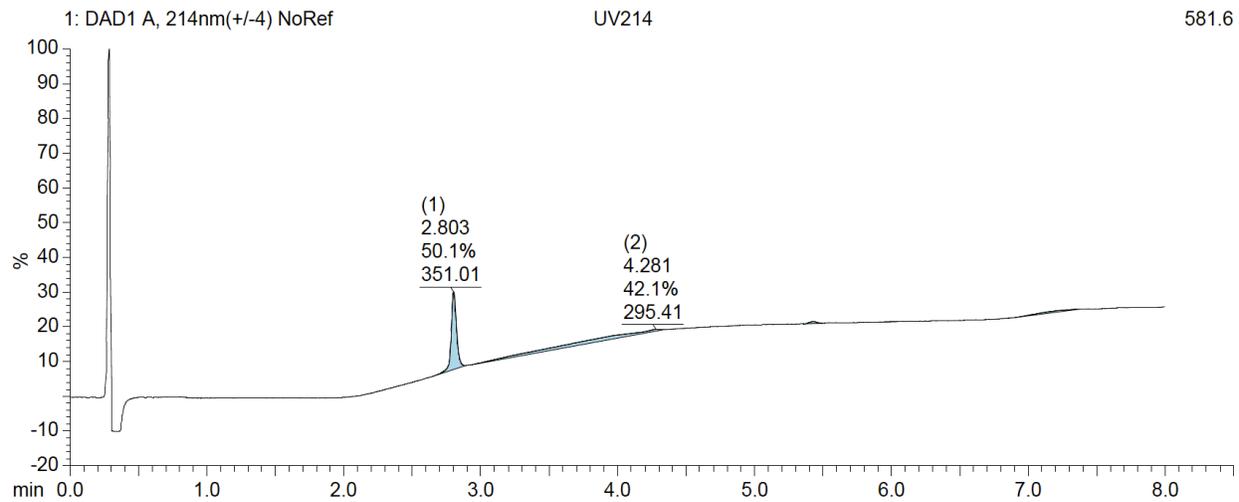
**HRMS (ESI-TOF):** for  $C_{46}H_{66}N_{13}O_{13} (M + H)^+$ : *calcd.*: 1008.4898; *found*: 1008.489.



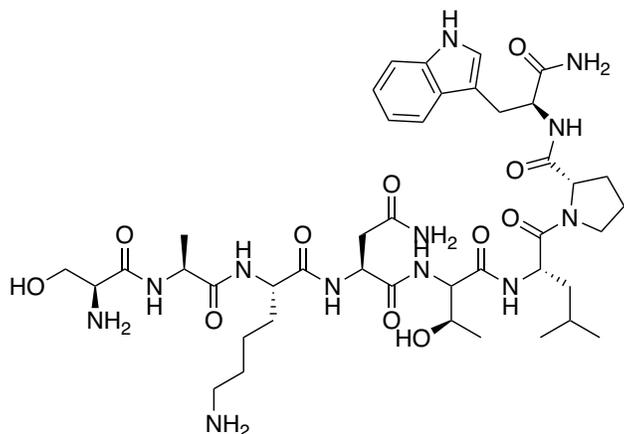
## Peptide 9 – SVHAPAQL



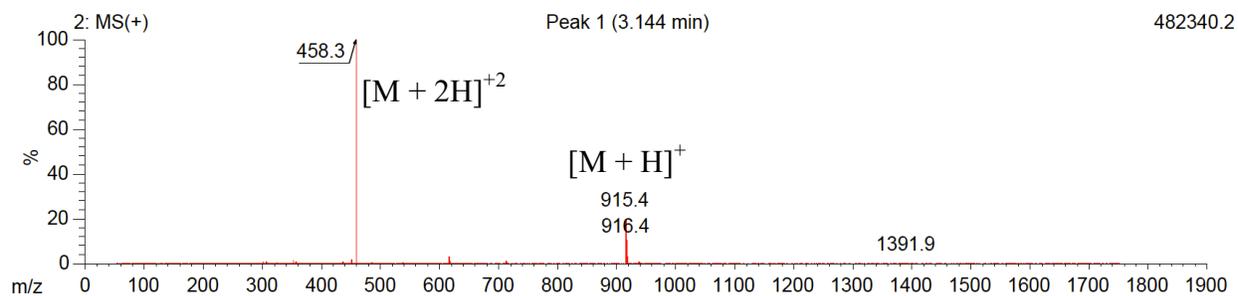
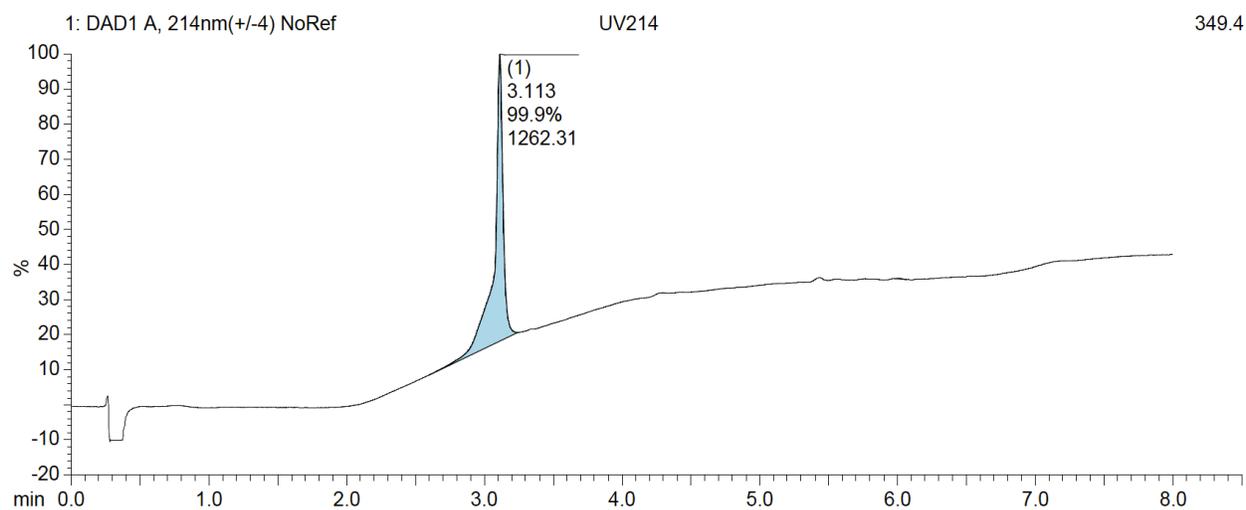
**HRMS (ESI-TOF):** for  $C_{36}H_{61}N_{12}O_{10} (M + H)^+$ : *calcd.*: 821.4628; *found*: 821.4618.



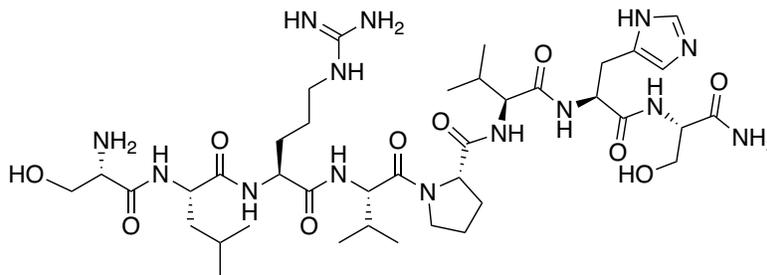
## Peptide 10 – SAKNTLPW



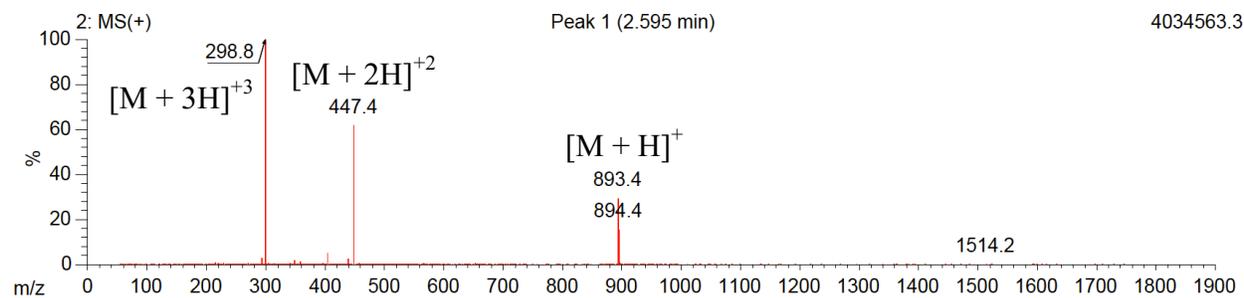
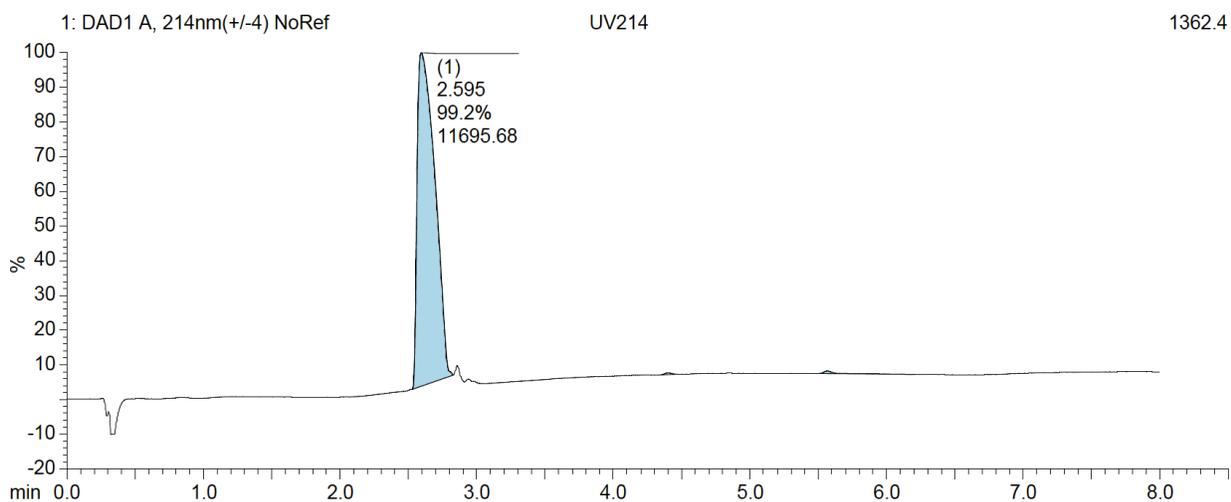
**HRMS (ESI-TOF):** for C<sub>42</sub>H<sub>67</sub>N<sub>12</sub>O<sub>11</sub> (M + H)<sup>+</sup>: *calcd.*: 915.5047; *found*: 915.5043.



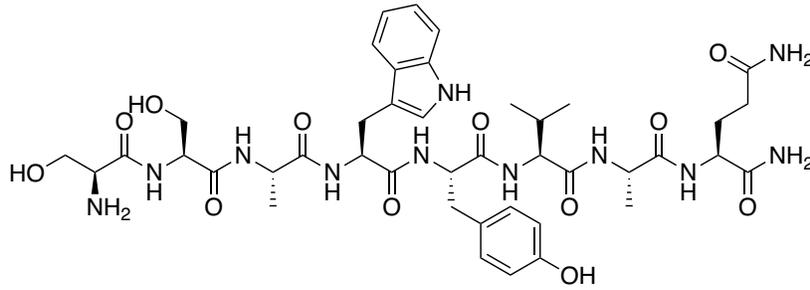
## Peptide 11 – SLRVPVHS



**HRMS (ESI-TOF):** for C<sub>39</sub>H<sub>69</sub>N<sub>14</sub>O<sub>10</sub> (M + H)<sup>+</sup>: *calcd.*: 893.5316; *found*: 893.5311.



## Peptide 12 – SSAWYVAQ



**HRMS (ESI-TOF):** for  $C_{42}H_{60}N_{11}O_{12} (M + H)^+$ : *calcd.*: 910.4417; *found*: 910.4415.

