

Integrating Computational Tools in the Design and Development of Asymmetric Catalysts and Enzyme Inhibitors

Sharon Pinus

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Department of Chemistry
McGill University
Montréal, Québec, Canada
December 2023

© Sharon Pinus, 2023

“That is the first thing I know for sure: (1.) If the questions don't make sense, neither will the answers.”

Kurt Vonnegut, *The Sirens of Titan*

Abstract

Chiral molecules are important in many biological, pharmaceutical, and industrial processes due to their asymmetric structures. One method for synthesizing chiral molecules with selectivity is by using asymmetric catalysts. Within the field of asymmetric catalysis, asymmetric organocatalysts have become more prominent in recent years, mainly due to their relatively low price and the need to avoid the use of heavy metals in some processes. As such, organocatalysis holds a central position in the landscape of modern chemistry, enabling the creation of intricate molecular structures with high enantioselectivities. However, despite their important applications, the potential of enantioselective organocatalysts remains constrained by existing limitations. The process of synthesis and optimization of a new catalyst and reaction conditions is lengthy, costly, and tedious.

To address these constraints, our research group has offered a solution in the form of VIRTUAL CHEMIST: a computational platform for the design and improvement of asymmetric catalysts. The platform is meant to be used by organic chemists and help speed up catalyst discovery and optimization. In this thesis, we aimed to demonstrate the usability of the platform by an organic chemist, validate its accuracy, and identify potential issues and limitations.

The first chapter aims to give an overview of the current computational methods for enantioselectivity predictions. This review is written from the perspective of an organic chemist and is meant to introduce the different terms and methodologies used, as well as applications. Lastly, it also discusses the differences between existing approaches and their limitations.

The second chapter introduces VIRTUAL CHEMIST and describes the process of validation and adaptation for practical use. The chapter describes the identification and handling of multiple problems, both in the software and the lack (at the time) of a chiral library for screening compounds that are commercially available or synthesizable. Validation of the software's accuracy using literature data was achieved, as predicted enantioselectivity could be compared to the published data. However, when seeking to perform experimental validation of two of the seven implemented reactions, unexpected issues arose. The next two chapters describe the investigation of these unexpected results.

Chapter 3 describes the investigation of the iminium catalyzed Diels-Alder cycloaddition. The lack of reactivity of the amines proposed in Chapter 2 became a considerable obstacle in obtaining final products and measuring the enantioselectivity. This led us to perform experimental and computational investigation of the reaction's mechanism. Our designed experiments indicate the iminium formation rate and concentration is highly dependent on the amine structure and substituents, which can then lead to very low conversion into product in the timescale of the cycloaddition reaction.

Chapter 4 describes the investigation of the Shi Epoxidation reaction, where the main hurdle was the library size for chiral ketones. Following the lessons learned in Chapters 2 and 3, we decided to first gain more insight into the reactivity of different ketones, so that when creating a virtual library, only active ones are included. Different simple ketones (achiral) were tested for reactivity in the reaction. We found that different ring sized ketones appear to have vastly different reactivities.

Lastly, when the COVID-19 pandemic hit, at the end of 2019, the need for antivirals was clear. To address this urgent need, most of our group members moved temporarily to working on this essential research. Thus, Chapter 5 describes the efforts of targeting SARS-CoV-2 by designing and synthesizing inhibitors of papain-like protease (PLpro), a viral protease which is vital for viral replication.

Overall, the work described in this thesis is a key step in improving the VIRTUAL CHEMIST platform and highlights the importance of evaluating reactivity when designing catalysts. This will improve the process of asymmetric catalysts discovery and optimization by using reliable software and will benefit all fields that require the synthesis of enantiopure molecules.

Résumé

Les molécules chirales sont essentielles en biologie, dans les médicaments et pour l'industrie tant pharmaceutique que des matériaux du fait de leur formes spécifiques. Une des méthodes pour leur préparation sélective est la catalyse asymétrique. Dans le domaine de la catalyse asymétrique, les organocatalyseurs asymétriques, apparus relativement récemment, sont d'un grand intérêt, particulièrement du fait de leur coûts attractifs et du besoin d'éviter les métaux lourds dans de nombreux procédés. De ce fait, l'organocatalyse tient une place centrale dans la chimie moderne, permettant la conception et la préparation de molécules complexes avec une bonne stéréosélectivité. Cependant, malgré leur importance, le plein potentiel des organocatalyseurs reste à explorer du fait de certaines contraintes, notamment liées aux difficultés de leur développement qui reste long et coûteux.

Afin de résoudre ces problèmes, notre groupe a conçu VIRTUAL CHEMIST, une plateforme pour la conception et l'optimisation de catalyseurs asymétriques. Cette plateforme a été développée pour être utilisée par les chimistes organiciens afin d'accélérer la découverte et l'optimisation de catalyseurs. Dans cette thèse, nous allons démontrer l'utilité de cette plateforme du point de vue du chimiste, valider sa fiabilité, identifier les limitations et problèmes potentiels et proposer des solutions à ces derniers.

Le premier chapitre présente une revue des différentes méthodes informatiques pour la prédiction de l'énantiosélectivité de réactions asymétriques, plus particulièrement du point de vue des utilisateurs (chimistes organiciens). Ainsi, nous allons introduire les différents termes propres à ces méthodes, les différentes méthodologies, et leurs applications. Les limitations de chacune de ces approches seront aussi discutées.

Le deuxième chapitre décrit le processus de validation et les modifications nécessaires pour son utilisation. Plus particulièrement, nous décrirons dans ce chapitre l'identification de problèmes spécifiques que la plateforme VIRTUAL CHEMIST peut résoudre ainsi que les problèmes rencontrés dans les différents logiciels la composant mais aussi dans le manque de bibliothèques de molécules virtuelles utilisables. L'utilisation pratique de données de la littérature (structures de catalyseurs et substrats ainsi que les énantiosélectivités mesurées) confirmera la fiabilité des différents logiciels de la plateforme. Cependant, des problèmes vont apparaître lors de cette validation utilisant des données de sept réactions connues et leur étude sera présentée dans les deux chapitres suivants.

Ainsi, le chapitre 3 présente l'étude approfondie de la cycloaddition de Diels-Alder catalysée par des iminiums. Le manque de réactivité des amines proposées dans le chapitre 2 sera un obstacle majeur à l'obtention du produit de cycloaddition et par conséquent à la mesure de l'énantiosélectivité. Cette observation nous amènera à conduire une série d'expériences et de calculs informatiques pour comprendre le mécanisme de cette réaction. Ces expériences indiqueront que la formation (vitesse de réaction et concentration finale) de l'iminium intermédiaire est dépendante de la structure de l'amine utilisée comme catalyseur ainsi que des substituants périphériques. Cette étape contrôle donc la conversion en produit et le temps de réaction. Le chapitre suivant décrira nos études de la réaction d'époxidation de Shi au cours desquelles nous allons mettre à jour le manque de bibliothèque adéquate. Forts des leçons apprises au cours des chapitres précédents, nous allons, cette fois encore, étudier la réaction de différentes cétones afin de concentrer nos études informatiques sur celles susceptibles de catalyser cette réaction. Un jeu de cétones simples sera testé et la présence et la taille de cycles de cétones cycliques sera identifié comme facteur majeur dans leur réactivité.

Finalement, lorsque la pandémie de COVID19 a frappé à la fin de 2019, le besoin d'antiviraux s'est rapidement fait sentir. Afin de contribuer à ce besoin criant, une grande partie de notre groupe s'est tournée vers cette recherche essentielle. Ainsi, le chapitre 5 présente nos efforts pour cibler le virus SARS-CoV2 par la conception et la synthèse d'inhibiteurs d'une protéase virale (PLpro) vitale pour la réplication virale.

En résumé, les travaux décrits dans cette thèse sont un pas important dans l'amélioration de la plateforme VIRTUAL CHEMIST et mettront en lumière le besoin de comprendre et d'évaluer la réactivité des catalyseurs lors de leur conception. Ces travaux permettront d'améliorer le processus de découverte et d'optimisation de catalyseurs asymétriques par l'utilisation de logiciels fiables et auront de l'impact dans tous les champs d'activité nécessitant la préparation de molécules chirales.

Acknowledgements

First and foremost, I would like to start by thanking my supervisor, **Dr. Nicolas Moitessier**, for his constant support, and invaluable guidance throughout my PhD. His expertise and encouragement played a huge role in making this thesis possible. I truly appreciate his patience, dedication, and belief in my abilities.

I would also like to thank **my family**; I am deeply grateful for your unwavering encouragement throughout this journey. Whether by letters, care packages, or phone calls, you always made sure I felt supported.

Additionally, I would like to thank **Dr. Mihai Burai-Pătrașcu**, my first friend in Canada, and one of my closest friends for the last 6 years. Thank you for all the walks, coffee breaks, and the surprise kinder eggs. I am thankful for your guidance during my first year and for your consistent readiness to offer help. It has been a privilege to work with you, and I will forever remember fondly you carefully setting up a reaction in the lab, and our many discussions over our shared projects!

I also want to thank **Julia Stille** and **Anne Labarre**. Where should I even begin? I consider myself incredibly lucky to have had you as not only lab mates, but also as my closest friends. You guys are forever my Canadian sisters. Thank you, for all the lab playlists, the tomato canning, pizza making, puzzle and Sangria evenings, craft nights, sharing on “the rocks”, and the list can go on forever. Thank you for introducing me to Canadian activities and culture and helping me appreciate the winters! Ice skating (I am just joking Julia; I know you just refer to it as “skating” in Canada!) followed by an amazing croissant got me through evening shifts during COVID. Thank you also to your wonderful families, for opening their home for me, and making me feel welcome, especially during the pandemic, when going home was not an option for me.

I am also grateful for **Jérôme Genzling**. Although you joined our group not too long ago, I sometimes feel like I have known you forever! Thank you so much, for your unfailing willingness to lend a hand and offer support. I miss early morning boxing classes with you, standing quietly in the metro because it is too early to talk. I am so glad I had the chance to work with you and learn from you and your constant curiosity.

Another person I would like to thank is **Chris (Guanyu) Wang**. For being a ray of sunshine in the lab, for bringing Ramen over whenever anyone needed cheering up, and for teaching me so much about plants. Thank you for always being so thoughtful and caring.

Although the pandemic was a rough time for us, I was so lucky to meet a lifelong friend. **Dr. Felicia Lim Phei Lin**, thank you so much for everything. Though our time working together was short, I learned so much from you during this short time, and I will always consider you a friend and mentor.

A special appreciation goes to **Dr. Itai Sharon**, for your endless patience, support, and input. Thank you for being our lab's structural biology expert, and for always being available to lend a helping hand.

I would also like to thank **Ahmed Bahaeldin**, for being such a supportive and thoughtful friend and office neighbour. Can't wait to meet you by the Red Sea!

Thank you also to **Benjamin Weiser, Steven Ma, David Hiraki**, and my wonderful lab mates in the Moitessier research group. It has been a pleasure working alongside you!

To those not explicitly named in this limited acknowledgment, your impact and support have not gone unnoticed, you are all cherished and appreciated greatly.

Table of Contents

Chapter 1	30
1.1. Introduction.....	30
1.1.1. Asymmetric Catalysis.....	30
1.1.2. Developing new catalysts.....	31
1.1.3. Computational methods for catalyst design and discovery.....	32
1.1.4. State-of-the-art approaches.....	33
1.1.5. A closer look at MM, QM, and ML.....	35
1.1.5.1. MM.....	35
1.1.5.2. QM.....	36
1.1.5.3. ML.....	37
1.2. Data.....	38
1.2.1. Developing datasets.....	38
1.2.2. Available datasets.....	41
1.3. Descriptors.....	42
1.3.1. Encoding catalysts, substrates and/or TS structures.....	42
1.3.2. Selection of descriptors and their computation.....	44
1.3.3. Number of descriptors.....	46
1.3.4. Different types of descriptors are used to describe unique atomic and molecular properties.....	46
1.3.4.1. Steric parameters.....	46
1.3.4.2. Geometric parameters.....	47
1.3.4.3. Electronic parameters.....	47
1.3.4.4. Empirical parameters.....	48
1.3.4.5. Interaction Fields. Comparative molecular field analysis (CoMFA).....	48
1.3.4.6. Comparative molecular similarity indices analysis (CoMSIA).....	48
1.4. Models.....	49
1.4.1. Model 1. Unknown mechanism/ stereoselective step (statistical and ML models).....	49
1.4.2. Model 2. Using knowledge on potential TS's (mechanism-based approach).....	51
1.4.2.1. Computational platforms for asymmetric catalyst design.....	54
1.4.2.1.1. Asymmetric Catalyst Evaluation (ACE).....	54
1.4.2.1.2. Quantum-guided molecular mechanics (Q2MM).....	55
1.4.2.1.3. An Automated Reaction Optimizer for New catalysts (AARON).....	56
1.5. Evaluation of the Models and Methods and Applications.....	56
1.5.1. Catalyst design.....	56
1.5.2. Stereoselectivity and catalytic activity optimization.....	57
1.5.3. Guiding asymmetric synthesis.....	61
1.6. Conclusions and Perspectives	64
Chapter 2	67

2.1.	Introduction.....	67
2.1.1.	Chiral molecules.	67
2.1.2.	Asymmetric catalysis.....	67
2.1.3.	Computational methods and existing software for enantioselectivity prediction..	68
2.2.	Results.....	70
2.2.1.	Running Diels-Alder with proline methyl ester hydrochloride	70
2.2.2.	Validation of VIRTUAL CHEMIST: virtual screens and experimental tests.....	71
2.2.2.1.	Virtual screening #1: ZINC database.....	74
2.2.2.2.	Virtual screening #2: ChemSpace.....	77
2.2.2.3.	Virtual screening #3: React2D library.....	79
2.2.3.	Validation by application to four real-life scenarios.....	81
2.2.3.1.	Real-life scenario I: Assessment of a small number (or just one) catalyst at a time.....	81
2.2.3.2.	Real-life scenario II: Finding a new series of catalysts for a known reaction; virtual screening #4: ZINC database	85
2.2.3.3.	Real-life scenario III: Designing and screening analogues for a hit with improved selectivity.	88
2.2.3.4.	Real-life scenario IV: Evaluation of the substrate scope for a catalyst.....	91
2.3.	Conclusions and future work.....	92
	Chapter 3	94
3.1.	Introduction.....	94
3.1.1.	Organocatalyzed Diels Alder cycloaddition.....	94
3.1.2.	Mechanism – what we know.....	95
3.1.3.	Mechanism – what we do not know.	97
3.2.	Results and Discussion.....	98
3.2.1.	Kinetic studies – iminium formation over time	98
3.2.1.1.	Different solvent systems with proline methyl ester hydrochloride salt	98
3.2.1.2.	Different amine hydrochloride salts.....	101
3.2.2.	Iminium concentration after 30 minutes.....	105
3.2.3.	Kinetic studies – Diels Alder cycloaddition	107
3.2.3.1.	Product formation with different amine salts.....	107
3.2.4.	Product formation simulation: effect of rate of each step.....	111
3.2.5.	Detailed mechanism.....	113
3.2.6.	Computational studies.....	115
3.2.6.1.	Two different hypotheses for the deprotonation of the amine and protonation of the hydroxy groups.....	115
3.2.6.2.	Initial plan	117
3.2.6.3.	Proton transfer step and re-evaluation of initial plan.....	118
3.2.6.4.	Results so far	121
3.2.6.4.1.	System 7	121
3.2.6.4.2.	Systems 6 and 8.....	123
3.3.	Conclusions and future work.....	124
	Chapter 4	127
4.1.	Introduction.....	127
4.1.1.	Alkene epoxidation with Oxone and the formation of dioxirane.	127

4.1.2.	Mechanism, proposed transition state and reactivity.....	128
4.1.3.	Asymmetric epoxidation and Shi's catalyst.....	131
4.1.4.	Reaction conditions	134
4.2.	Results	136
4.2.1.	Procedure establishment with Shi's chiral ketone catalyst.....	136
4.2.2.	Virtual screening: ChemSpace library	142
4.2.3.	Initial study of structure-activity relationship.....	144
4.3.	Conclusions and future work.....	148
	Chapter 5	151
5.1.	Introduction.....	151
5.1.1.	General information about the virus and pandemic.....	151
5.1.2.	Infection and viral replication.....	151
5.1.3.	Response to the pandemic.....	152
5.1.4.	PLpro activity and published inhibitors.....	153
5.2.	Results	158
5.2.1.	Design of a hybrid covalent inhibitor	158
5.2.2.	Synthesis and change in original design.	160
5.2.2.1.	Route I.	164
5.2.2.2.	Route II.	169
5.2.3.	Biological activity evaluation.....	175
5.3.	Conclusions and Future work.....	178
	Chapter 6 – Conclusions and Future Work.....	181
	References	189
	Supplementary information Chapter 2.....	211
	Supplementary information Chapter 3.....	239
	Supplementary information Chapter 4.....	288
	Supplementary information Chapter 5.....	292

List of Figures

Figure 1.1. The three main asymmetric catalysis fields.	31
Figure 1.2. Overall workflow for A) ML methods and B) QM/MM methods.....	35
Figure 1.3. Different methods used to compute enantioselectivity.....	37
Figure 1.4. Common descriptors used.....	49
Figure 1.5 Curtin-Hammett principle in the context of asymmetric catalysis. C: catalyst, S: substrate, P: product, either R or S enantiomer. S+C(S) and S+C(R) are the catalyst substrate complexes leading to either the S or the R product, respectively.....	52
Figure 1.6. A) Dotson and co-workers' work. B) Rinehart and co-workers' work.....	59
Figure 1.7. Betinol and co-workers' work on the application of existing models to different scenarios: synthesis of complex scaffolds, synthesis of a natural product, finding preferred catalyst for the same transformation, and finding preferred route to form the same product.....	63
Figure 2.1. Comparison of molecule screening workflows. Left- experimental chemists, right-computational chemists.....	73
Figure 2.2. Amines 2.8 and 2.11 were selected for synthesis and testing. The ZINC ID is provided under the molecules, as well as the average predicted enantioselectivity.....	75
Figure 2.3. Top 5 results of the ChemSpace virtual screening. Under each molecule is the ChemSpace ID and the ACE predicted enantioselectivity.....	78
Figure 2.4. a) Amine scaffolds used to create the combinatorial library. In the starting materials $R_1=H$. b) A library of 10 amines reacted in REACT2D with 34 aldehydes in reductive amination. The final library contained 340 chiral secondary amines.....	79
Figure 2.5. Top five amines ranked in the virtual screening of the library of analogues. Predicted selectivity is shown as % <i>ee</i> 's provided under the molecules.....	80
Figure 2.6. General reaction schemes for the 7 reaction classes that Ace was tested on, With selected examples of Ace-optimized TSs in 3D and 2D.....	82
Figure 2.7. a) Unsigned error for $\Delta\Delta G^\ddagger$ (kcal/mol) between the predicted and experimental enantioselectivity for each reaction of the 7 reaction classes. Reaction colour correlated to the reaction class on top. Red dot is the average unsigned error. Black dot is the theoretical average unsigned error we would get if the predictions made were random; For this calculation random values between -4.12 to 4.12 kcal/mol were assigned for $\Delta\Delta G^\ddagger$. This represents maximum	

selectivity margins (ratio of 1000:1). **b)** A selected example of predicted vs. observed $\Delta\Delta G^\ddagger$ for the Shi epoxidation reaction, with 51 asymmetric catalyst/substrate pairs. Positive $\Delta\Delta G^\ddagger$ represents one enantiomer whereas negative $\Delta\Delta G^\ddagger$ represents the other enantiomer.....83

Figure 2.8. Examples of substrates (chiral auxiliary Diels-Alder, **2.16**, **2.17**) and catalysts that had $\Delta\Delta G^\ddagger$ errors of 2 kcal/mol or more. In parentheses the reaction class to which the molecule was tested for.....85

Figure 2.9. a) Workflow for the virtual screening (example with Diels-Alder cycloaddition) and reducing the library size into the most diverse molecules with description of actions on the right. **b)** Second workflow created to screen molecules for enantioselectivity, with description of actions on the right. **c)** Molecules screened ranked for enantioselectivity (Diels-Alder cycloaddition and Shi epoxidation). The red bars indicate the known catalysts that were added to the screening as decoys. The graph indicates the portion of known catalysts compared to the library from the ZINC database. The graph is created by ranking order, e.g., for the Diels-Alder cycloaddition ~7% of the molecules ranked at the top from the ZINC database, compared to ~85% of the known catalysts. That is, the known enantioselective catalysts were placed correctly at the top of the list.....87

Figure 2.10. Top: Workflow for the virtual optimization of organocatalysts for the Diels-Alder cycloaddition. REACT2D was used to combine a library of primary amines (x represents a protecting group) with a library of aldehydes to create a library of imines, by imine formation reaction. Then a 1,3-dipolarcycloaddition was done using REACT2D to create a library for each isomer product in parallel. All three catalyst libraries were then evaluated for enantioselectivity in the Diels-Alder cycloaddition, *endo* and *exo* are run separately. Bottom: Results of the virtual screen plotted against experimental values (as $\Delta\Delta G^\ddagger$)- predicted on y-axes and observed on x-axes. Orange dots represent the *endo* adduct blue dots represent the *exo* adduct.....90

Figure 2.11. Results of the substrate scope study done on the Sharpless asymmetric dihydroxylation using (DHQD)2PHAL as catalyst with different alkene substrates. Each substrate result is plotted as a blue dot, with predicted enantioselectivity (as $\Delta\Delta G^\ddagger$) on the y axes and observed experimental enantioselectivity on the x axes. Positive $\Delta\Delta G^\ddagger$ represents isomers (R) and (R, R), and negative $\Delta\Delta G^\ddagger$ represents the other isomers.....91

Figure 3.1. (a) Proposed mechanism for the proline methyl ester (**3.1**)-catalyzed Diels-Alder reaction. For simplicity, only one enantiomer of the iminium and the adduct is shown (stereochemistry of products not shown). **(b)** Known catalysts used in this reaction. McMillan's imidazolinone based catalyst (**3.5**), and the pyrrolidine based catalyst used by Hayashi (**3.6**). **(c)** LUMO lowering activation with a Lewis acid (top) or with a secondary amine salt via the formation of iminium, as proposed by McMillan's group.....95

Figure 3.2. Iminium formation and equilibrium over time from proline methyl ester hydrochloride (**3.1**) and cinnamaldehyde (**3.2**) in three different solvent systems. **(a)** General reaction scheme.

Both the amine salt (**3.1**) and aldehyde are at 1 equiv. in a 1 M solution. The colors match the plot's color over time. **(b)** Total iminium -**3.1a** and **3.1b** (*cis* and *trans*, respectively) and acetal (**3.7**)– formation and equilibrium over time in CD₃OD-D₂O (19:1). Each line color corresponds to the color in **(a)**: **3.2** is blue, **3.1a** and **3.1b** are summed in the green line, **3.7** is orange. **(c)** Total iminium -**3.1a** and **3.1b** (*cis* and *trans*, respectively) and acetal (**3.7**)– formation and equilibrium over time in CD₃CN-D₂O (19:1). **(d)** Total iminium -**3.1a** and **3.1b** (*cis* and *trans*, respectively) and acetal (**3.7**)– formation and equilibrium over time in CD₃OD.....100

Figure 3.3. (a) Iminium formation equilibrium from McMillan's imidazolinone-based catalyst (**3.5**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). **(b)** Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme.....101

Figure 3.4. (a) Iminium formation equilibrium from sarcosine methyl ester HCl salt (**3.9**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). **(b)** Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme.....102

Figure 3.5. (a) Iminium formation equilibrium from piperidine HCl salt (**3.10**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). **(b)** Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme. **(c)** Iminium formation equilibrium from pyrrolidine HCl salt (**3.11**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). **(d)** Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme.....104

Figure 3.6. Structures of amines used in the study (top), and results of conversion of the amine salts into iminium after 30 minutes (bottom graph). The amine salts were used in 20 mol%, in a 1 M solution of CD₃OD-D₂O (19:1). Conversion into iminium is calculated based on integration: $(H_{\text{iminium}}/(H_{\text{iminium}}+H_{\text{amine}})) \times 100$107

Figure 3.7. (a) General reaction scheme for the iminium catalyzed Diels-Alder cycloaddition with different amine HCl salts (**3.1**, **3.11**, **3.8**), cinnamaldehyde (**3.2**) and cyclopentadiene (**3.3**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1), cyclopentadiene is used in 1 equiv. **(b)** Reaction progression over time with HCl salt of proline methyl ester (**3.1**). The plot colors correspond to the colors in the scheme, brown represents total product formed. **(c)** Reaction progression over time with HCl salt of 2-(trifluoromethyl)pyrrolidine (**3.8**). The plot colors correspond to the colors in the scheme, brown represents total product formed. **(d)** Reaction progression over time with HCl salt of pyrrolidine (**3.11**). The plot colors correspond to the colors in the scheme, brown represents total product formed.....110

Figure 3.8. Simulated time course concentration for product formation over time of the full cycloaddition reaction with proline methyl ester HCl (**3.1**), based on the data collected in the kinetic studies of iminium formation and full reaction and the best fit parameters (Chapter 3-SI **Figure S3.13**, **Table S3.4**). **(a)** Reaction scheme, first step is the iminium formation and marked in pale blue (k_1), second step is the cycloaddition reaction of the iminium formed with the cyclopentadiene and is marked in dark blue (k_2). The total product is highlighted in a grey box. **(b)** Results of the simulation plotted as total product concentration over time, plot line colors correspond to the different steps as described in **(a)**, with the original rates being marked in cyan. The lines added at ~30 minutes are meant to help visualize the difference in product formation, depending on the rate of each step. The difference should be compared to the original line. Increasing the rate of the second reaction (k_2) has a larger effect on the overall rate of the reaction, suggesting that the second step is the rate limiting step.....112

Figure 3.9. Diels Alder reaction catalyzed by proline methyl ester (**3.1**): detailed mechanism. The proposed mechanism is described while keeping the HCl and amine ratio as 1:1, based on the conditions that were used in our experiments and experiments discussed in the introduction. It is possible that locally 2 equivalents of acid are necessary for the iminium formation step (structure **3.27**).....114

Figure 3.10. (a) The two breaking bonds that were scanned. d_1 (orange) the NH--OH₂ bond (which reflects the N-H bond breaking), and d_2 (green) the C-O bond. **(b)** The model system included a network of 6 water molecules all interconnected via H-bonds to facilitate protonation/deprotonation.....115

Figure 3.11. The three systems used to test our hypothesis and study the stepwise formation of the iminium: **System 6**: the chloride counter anion located on the top face by the protonated nitrogen. **System 7**: local 2 equiv. of HCl, one chloride by the hydroxy, one by the protonated nitrogen, second proton on a water molecule (H₃O⁺) by the hydroxy. **System 8**: the same as system 6, but the chloride counter anion is placed on the bottom face, by the hydroxy group. All systems have 6 water molecules around in a H-bond network. The proton originating from the amine is highlighted in dark pink, the proton originating from the second acid equiv. (only in system 7) is highlighted in pale pink. Chlorides are highlighted in blue. All intermediates are labeled based on the relevant system.....118

Figure 3.12. Proton transfer option **III** for systems 6 & 8. **(a) III.1**: left: early steps of d_1 (N still protonated), prior to scan d_2 . Right: later steps of d_1 (N deprotonated) prior to d_2 scan. Proton transfer neutralizes the hydroxide formed (highlighted in teal). H-transfer bonds labeled in pink. **(b) III.2**: proton is removed from water network for all d_1 steps prior to d_2 scan, forming a hydroxide and a hydronium by the hydroxy (highlighted in teal).....120

Figure 3.13. Proton transfer step for system 7 prior to scan d₂, after scan d₁ for **path a**- proton placed closed to the hydroxy (left, proton highlighted in teal) and prior to scan d₁, for **path b**- proton placed on the hydroxy (right, proton highlighted in teal).....122

Figure 3.14. Geometry optimized structures of system 7 after step 3: **(a)** path a. **(b)** path b. **(c)** Both paths superimposed, path **(a)** (pink carbons), path **(b)** (grey carbons). The systems are essentially identical. They only differ on the orientation of the hydroxy, and which water molecule is protonated.....123

Figure 4.1. Two possible transition states proposed. Left: spiro TS, secondary interaction between the oxygen HOMO (n orbital) and the alkene LUMO (π^* orbital). Right: planar TS, the secondary interaction is not possible (or is minimal).....129

Figure 4.2. (a) Dioxiranes with α -trifluoromethyl groups, **4.5a**, are more reactive than their methyl counterparts, **4.6a**. The ketones from which the dioxiranes are formed likely follow the same reactivity trend in the dioxirane formation (**4.5** being more reactive than **4.6**). **(b)** 3-oxopiperidinium (**4.7**) that was tested by Denmark *et al.* and lactone formation from the Baeyer-Villiger oxidation was observed.....131

Figure 4.3. Chiral ketone catalysts published by Curci *et al.* (**4.8** and **4.9**), Shi *et al.* (**4.10**), Yang *et al.* (**4.11**), Denmark *et al.* (**4.12**), and Armstrong *et al.* (**4.13**).....132

Figure 4.4. Eight possible TSs for trisubstituted alkenes. The major competing TSs, that lead to the opposite enantiomers, are **Spiro 1** and **Planar E**. For the disfavored TS, steric clashes are highlighted with red arrows.....134

Figure 4.5. Top results of the ChemSpace virtual screening. All the highlighted ketones come from secondary alcohols that were oxidized into ketones with REACT2D. Under each molecule is the ChemSpace ID for the molecule the ketones are derived from, and the averaged %*ee* as predicted by ACE.....144

Figure 4.6. Chiral ketones with different structures tested for their reactivity **(a)** by Denmark *et al.* **(b)** by Yang *et al.*.....145

Figure 4.7. 5-membered ring ketones tested by Shi *et al.* (**4.20**, **4.21**) and Denmark *et al.* (**4.22**).....147

Figure 5.2. Antiviral drugs approved for use against COVID-19. On top- commercial name and manufacturer, in parentheses-what viral component targeted, below compounds- common name.....153

Figure 5.2. (a) Crystal structure of PLpro (PDB: 6WX4). Catalytic residues in the active sight are colored pale orange and shown as sticks, as well as a 2D illustration on the right box. **(b)** Cleavage

points of PLpro with the conserved motif LXGG. Nsp: nonstructural protein, C-term: carboxylic terminus. (c) 2D illustration of the conserved motif. Cleavage point is marked with dashed line and scissors.....154

Figure 5.3. Covalent tetrapeptide ligand **VIR251** (left). On the right, **VIR251** (pink) co-crystallized with PLpro (shown as grey surface, PDB: 6WX4). The structure is shown from two perspectives, to demonstrate the entrance to the tunnel where the active site is, as well as a side view.....155

Figure 5.4. Non-covalent inhibitor **GRL0617** (left) with IC₅₀ value underneath. On the right: co-crystal structure of **GRL0617** (pink) with SARS-CoV-2 PLpro (shown as a grey surface, PDB: 7JIR). The structure is shown from two different perspectives to demonstrate where the ligand interacts with the protein, as well as the entrance to the tunnel, where the active site is.....156

Figure 5.5. (a) The most active **GRL0617** analogues synthesized by Shen et al. with IC₅₀ values underneath. (b) Compound **XR8-83** (pink) co-crystallized with PLpro (cyan, PDB: 7LLF). Residues that form interactions with the compound are shown in stick and labeled.....157

Figure 5.6. (a) Covalent analogues of **GRL0617** synthesized by Sanders et al. with IC₅₀ values under each compound. (b) Compound **7** (pink) co-crystallized with PLpro (cyan, PDB: 8EUA). Residues that form interactions with the compound are shown in stick and labeled.....158

Figure 5.7. (a) Designed hybrid compound **5.1**. Blue dashed lines represent planned synthesis for the compound, reaction classes are written in blue. (b) Docking pose of **5.1** (pink) in PLpro's active site (teal), with potential interactions with residues (in stick, labeled). PDB: 6WX4. (c) Docked pose of **5.1** with PLpro shown as grey surface, indicating the ligand's fit in the narrow active site. (d) Overlay of 5 original co-crystallized ligand (**VIR251**, blue) with the docked pose of **5.1**, with PLpro as a grey surface, indicating similar space occupancy.....160

Figure 5.8. Example of **GRL0617** analogues designed by Ma *et al.* with inhibitory activity against PLpro, despite the amine group being replaced by a hydroxy (**Jun9-87-1**) or replaced by a hydroxy and moved from the meta position to the alkyl amine to para (**Jun9-72-2**).....169

Figure 0.9. Dose response curves for **5.1a** (left) and **GRL0617** (right) against PLpro, with Z-RLRGG-AMC as substrate.....178

Figure 5.10. Compound **XR8-83** (left), and potential analogues of compound **5.1a** (right). The region in the compounds inspired by **XR8-83** is highlighted in blue in **XR8-83**. The region changed in the analogues compared to **5.1a** is highlighted in purple.....180

List of Schemes

- Scheme 2.1.** Established procedure for carrying out the Diels-Alder cycloaddition with a secondary amine salt and obtaining the enantioselectivity from the alcohol.....71
- Scheme 2.2.** Synthesis of the two selected potential catalysts from the virtual screening. **a) i)** K₂CO₃, H₂O, 10°C to rt, 24h, then 90°C for 2h (52%). **ii)** HCl in Et₂O, THF, 0°C to rt, 24h. (quant.). **b)** PyBOP, DIPEA, DMF, 0°C to rt, 18h (89%). **ii)** HCl in Et₂O, THF, 0°C to rt, 24h. (67%).....75
- Scheme 2.3. a)** Testing of amines **2.8** and **2.11** as catalysts of the Diels-Alder cycloaddition. Each catalyst was tested both in 10 mol% and in 100 mol% (relative to the aldehyde). **b)** Cyclopentadiene (**2.1**) dimerizes into its dimer (**2.1a**) over time, at rt.....76
- Scheme 4.1.** General epoxidation conditions as published by Curci *et al.*; **a)** general structure of the dioxirane that is formed by reaction of the ketone with Oxone. **b)** acetone in water solution. **c)** biphasic solution with tetrabutylammonium hydrogensulfate as a phase transfer catalyst.....127
- Scheme 4.2.** General mechanism for the reaction of a ketone (**4.3**) with Oxone to give the epoxide product (**4.4a**). As well as 3 unproductive pathways (dashed arrows) that can inhibit product formation.....128
- Scheme 4.3. (a)** Initial epoxidation conditions as published by Shi *et al.*, with *trans* stilbene (**4.14**) as the example, using chiral ketone **4.10**. **(b)** Four proposed TSs for the epoxidation of *trans* stilbene (or any symmetrical alkene). The major enantiomer being **4.14a** provides evidence for the favored TS being **Spiro 1**.....133
- Scheme 4.4. (a)** Test reaction of Oxone's reactivity in a different oxidation reaction. **(b)** *Trans*-stilbene (**4.14**) racemic epoxidation with mCPBA.....139
- Scheme 4.5.** General scheme for the epoxidation reaction carried out with different conditions (as detailed in **Table 4.1.**). TBAHS: tetrabutylammonium hydrogensulfate, DMM: dimethoxy methane.....140
- Scheme 4.6.** General scheme for creating the virtual library of chiral cyclic ketones: cyclic ketones adjacent to a stereocenter, with ring sizes of 4-7 carbons (bottom) were filtered for, resulting in a very small library. Cyclic secondary alcohols (of inconsequential stereochemistry at the hydroxy carbon) adjacent to a stereocenter, with ring sizes of 4-7 carbons (top) were also selected in a separate library. These alcohols were then virtually oxidized into the corresponding ketones using react 2D (these ketones are highlighted to differentiate them from the ketones derived directly from the ChemSpace library. The ketone libraries were merged into a single library that was used in the virtual screening.....143

Scheme 4.7. General procedure used to test different ketones' reactivity. The ketones tested thus far are on the right. Under each ketone is the % conversion into product as measured by ¹H-NMR.....146

Scheme 5.1. (a) Synthesis of **GRL0617** analogues; PyBOP, DIPEA, DCM, rt, 24h, **GRL0617**, (74%), **5.2**, (51%), **5.3**, (73%); EDCI, HOBt·H₂O, DIPEA, DCM, rt, 24h, **5.4**, (43%). **(b)** Methylation attempts of the hydroxy group on **5.2**; **i)** (CH₃)₂SO₄, Cs₂CO₃, DMF, rt, 24h. **ii)** CH₃I, Cs₂CO₃, DMF, rt, 24h. Both conditions resulted in a mixture of methylation products. **(c)** Methylation attempt on the hydroxy group of the nitro derivative **5.4**; CH₃I, K₂CO₃, acetone, rt, 24h. Only starting material was observed by TLC. **(d)** Test alkylation on a simpler starting material, 4-nitrophenol; **i)** 3-bromopropanoic acid, KOH, H₂O, reflux for 5.5h, then rt for 19h. **ii)** 3-bromopropanoic acid, DIPEA, THF, MW, 120°C, 15 min. Both conditions resulted in a mixture of alkylated product as well as the elimination product, acrylic acid (almost 1:1).....162

Scheme 5.2. Different routes for the synthesis of **GRL0617** analogs. **(a) Route I**, amine group is kept or masked as nitro, reactivity is flipped, a halide is used instead of the hydroxy. **i)** HATU, DIPEA, acetonitrile, rt, 24h, **5.6**, (71%), **5.7**, (80%); PyBOP, DIPEA, acetonitrile, rt, 24h, **5.8**, (91%). **ii)** Boc₂O, ethanol, rt, 24h, **5.9**, (48%). **(b) Route II**, amine group replaced by different substituents. **i)** HATU, DIPEA, acetonitrile or DCM rt, 24h, **5.10**, (DCM, 54%), **5.11**, (acetonitrile, 71%), **5.12**, (acetonitrile, 48%), **5.13**, (DCM, 87%). **ii)** acetyl chloride, Et₃N, rt, DCM, 21h, 61%, **5.14**.....164

Scheme 5.3. (a) i) Na₂CO₃, THF, reflux, 24h. **ii)** DIPEA, dioxane, MW, 120°C, 1h, **5.15**. **(b) i)** CuI, K₂CO₃, DMF, 90°C, 48h, **5.7a**. **ii)** CuI, L-prolinem K₂CO₃, DMF, 90°C, 48h, **5.7b**. **iii)** Pd₂(dba)₃, BINAP, Cs₂CO₃, THF, reflux, 24h, **5.7b**. **iv)** Pd₂(dba)₃, BINAP, Cs₂CO₃, toluene, 100°C, 24h, **5.7b**. **v)** Pd(PPh₃)₄, K₂CO₃, DMF, 80°C, 20h, 27%, **5.7c**. **(c) i)** Pd₂(dba)₃, BINAP, Cs₂CO₃, toluene, 100°C, 24h, 23%, **5.15**. **ii)** Pd(PPh₃)₄, K₂CO₃, DMF, 80°C, 20h, **5.16**. **(d) i)** Pd₂(dba)₃, BINAP, KO^tBu, toluene, 100°C, 48h, **5.17**.....166

Scheme 5.4. (a) Potential oxidation products of the amination product when *t*-butyl glycinate is used. **(b)** Potential metal coordination of α-keto imine to a metal center.....168

Scheme 5.5. i) 5.18: NaH, DMF, rt, 24h; **5.19:** K₂CO₃, DMF, 60°C, 24h / NaI, acetone, rt, 5h / KOH, H₂O-THF (4:1), 80°C, 4h / K₂CO₃, DMF, MW, 120°C, 2h; **5.20:** K₂CO₃, 80°C, DMF, 24h / NaH, DMF, rt, 24h. **ii) 5.21:** DCC, DMAP, DMF, rt, 3h; **5.22:** EDCI, DMAP, DMF, rt, 24h (10%). **iii) 5.23:** K₂CO₃, DMF, rt, 24h (32%); **5.24:** K₂CO₃, DMF, rt, 24h (28%). **iv) 5.25:** TCCA, TEMPO, NaBr, acetone, rt, 24h.....171

Scheme 5.6. i) 5.26: NaOMe (25 w% in methanol), dioxane, MW, 150 °C, 30 min, then additional 1.05 equiv. of methyl-3-hydroxypropanoate and 1.5 equiv. of NaOMe were added, 170 °C for 60 more min; DBU, dioxane, MW, 150 °C, 30 min. **ii) 5.27:** DIPEA, ethanol, MW, 120 °C, 90 min.; DBU, isopropanol, MW, 150 °C, 90 min.....172

Scheme 5.7. (a) **i**) CuI, K₂CO₃, DMF-H₂O (14:1), 100°C, 20 h, **ii**) NaOH (1.6 M aq.), methanol, rt, 1 h, 36% (**5.28**) (b) CuI, Cs₂CO₃, DMF, 100°C, 3 h, 50% (**5.28**).....173

Scheme 5.8. (a) **i**) NaH, THF, 0°C to rt, 3.5 h, 60% (**5.29**); **ii**) TFA, DCM, 0°C, 2 h, quantitative yield (**5.30**). (b) DIPEA, PyBOP, DCM, 0°C to rt, 18 h, 37% (**5.1a**).....174

Scheme 5.9. Synthesis of **5.1a** analogues: **i**) HATU, DIPEA, acetonitrile, rt, 24 h, 40% (**5.31**), 78% (**5.32**). **ii**) CuI, Cs₂CO₃, DMF, 100 °C, 3 h, 58% (**5.33**), 52% (**5.34**). **iii**) DIPEA, PyBOP, DCM, 0 °C to rt, 18 h, 11% (**5.35**), 58% (**5.36**).....175

List of Tables

Table 4.1. Different conditions that were used for the epoxidation reaction, with the purpose of improving yield and reaching a result comparable to the published one.....141

Table 5.1. Biological activity evaluation of compounds synthesized against PLpro.....175

List of Abbreviations

- listed alphabetically -

3CLpro	3-chymotrypsin-like protease
AI	artificial intelligence
API	active pharmaceutical ingredient
ASO	average steric occupancy
Bn	benzyl
Boc	tert-butyloxycarbonyl
Bu	butyl
Cbz	benzyl chloroformate
COVID-19	coronavirus disease 2019
CPA	chiral phosphoric acid
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DFT	density functional theory
DIPEA	N,N-diisopropylethylamine
DMF	dimethyl formamide
DMM	dimethoxymethane
EDCI	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide chloride
EDTA	ethylenediaminetetraacetic acid
<i>ee</i>	enantiomeric excess
equiv.	equivalent
ESP	electrostatic potential
Et	ethyl
FF	force field
FMO	frontier molecular orbital
GC	gas chromatography
HATU	hexafluorophosphate azabenzotriazole tetramethyl uronium
HBD	hydrogen bond donor
HOBt	hydroxy benzotriazole

HOMO	highest occupied molecular orbital
HPLC	high performance liquid chromatography
HWE	Horner-Woodward-Emmons
IC ₅₀	half maximal inhibitory concentration
iPr	isopropyl
IR	infrared
LA	Lewis acid
LFER	linear free energy relationship
LUMO	highest unoccupied molecular orbital
MC	Monte Carlo
<i>m</i> CPBA	<i>meta</i> -chloroperoxybenzoic acid
Me	methyl
ML	machine learning
MLR	multivariate linear regression
MM	molecular mechanics
MW	microwave
NBO	natural bond orbital
NMR	nuclear magnetic resonance
nsp	nonstructural protein
OTf	triflate
<i>p</i> -	<i>para</i>
PCA	principal component analysis
Ph	phenyl
Plpro	papaine-like protease
PMB	<i>para</i> -methoxybenzyl
PyBOP	benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
QM	quantum mechanics
QSSR	quantitative structure selectivity relationship
SARS-CoV	severe acute respiratory syndrome coronavirus
SMILES	simplified molecular-input line-entry system
SVR	support vector regression

TBAHS	tetrabutylammonium hydrogensulfate
tBu	<i>tert</i> -butyl
TES	triethylsilyl ether
TFA	trifluoroacetic acid
THF	tetrahydrofurane
TLC	thin layer chromatography
TS	transition state
TSFF	transition state force field
UTS	universal training set
UV	ultraviolet

List of Author Contributions

During the course of my PhD I have co-authored 3 published manuscripts and 3 manuscripts in preparation:

Publications (chronological order). ‡ denotes first author or co-first-author

1. M. Burai-Patrascu‡, J. Pottel, **S. Pinus**, M. Bezanson, P. O. Norrby and N. Moitessier. *Nat Catal* 3, **2020**, 574–584.
2. I. Sharon‡, **S. Pinus**, M. Grogg, N. Moitessier, D. Hilvert and T. M. Schmeing. *Nat Comm* 13, **2022**, 3923.
3. J. K. Stille‡, J. Tjutrins‡, G. Wang‡, F. A. Venegas‡, C. Hennecker, A. M. Rueda, I. Sharon, N. Blaine, C. E. Miron, **S. Pinus**, A. Labarre, J. Plescia, M. Burai Patrascu, X. Zhang, A. S. Wahba, D. Vlaho, M. J. Huot, T. M. Schmeing, A. K. Mittermaier, and N. Moitessier. *Eur J Med Chem*, 229, **2022**, 114046

Manuscripts in Preparation. ‡ denotes first author or co-first-author

4. H. Y. Huang‡, **S. Pinus**, X. Zhang, G. Wang, A. M. Rueda, Y. Souaibou, S. Huck, M. Huot, D. Vlaho, J. Pottel, F. A. Venegas, Z. Lu, C. Hennecker, J. Stille, J. Tjutrins, C. E. Miron, A. Labarre, J. Plescia, M. Burai-Patrascu, A. K. Mittermaier, and N. Moitessier. Preprint on *ChemRxiv*: <https://doi.org/10.26434/chemrxiv-2023-r0v7t>
5. **S. Pinus**‡, J. Genzling, M. Burai-Patrascu and N. Moitessier, 2023, Submitted to *Nature Catalysis*
6. **S. Pinus**‡, J. Genzling‡, M. Burai-Patrascu and N. Moitessier, 2024, manuscript in preparation

This page intentionally left blank

This chapter is based on a review submitted to *Nature Catalysis*

Catalyzing Change: The Power of Computational Asymmetric Catalysis, **S. Pinus**, J. Genzling, M. Burai-Patrascu and N. Moitessier.

SP, MBP, and NM devised the structure of the review. SP collected, curated, sorted, and categorized the references. SP, JG, MBP, NM curated the datasets. SP led the writing of the review (including creating the figures), while JG, MBP and NM contributed to the writing of this manuscript.

Chapter 1

Catalyzing Change: The Power of Computational Asymmetric Catalysis

1.1. Introduction

Computational asymmetric catalysis has seen an impressive rise in the last twenty years, thanks to advancements in algorithm and method development for predicting catalyst enantioselectivity. These methods/algorithms describe reactions that can be categorized into two groups: reactions where 1) knowledge of the mechanism is not required and where leveraging experimental data to establish correlations between reaction descriptors and enantioselectivity is imperative, and 2) the mechanism (or transition state (TS) for the enantioselective step) is known and used to determine catalyst stereoselectivity by modeling the diastereomeric TSs. Although these methods have reached a level of proficiency for enantioselectivity prediction, this field remains largely obscured for experimental chemists. In this review, we aim to shed light on models, methods, and applications used in asymmetric synthesis, with accessible language suited for experimental chemists. Our hope is that these methods will ultimately be adopted by synthetic chemists for the design of novel catalysts.

1.1.1. Asymmetric Catalysis.

The field of asymmetric catalysis has revolutionized organic synthesis in the last 50 years. Catalysts have been developed to introduce stereogenic centers into molecules primarily through the formation of new C-C bonds and reduction of unsaturated bonds (e.g., carbonyls, alkenes). These chiral catalysts can take many forms, including transition metal complexes, organocatalysts, and biocatalysts (**Figure 1.1**). With transition metal complexes, the reaction is often catalyzed by the metal itself while stereochemistry is introduced in the form of metal ligands (e.g., chiral

phosphines). In organocatalysis and biocatalysis, small organic molecules and enzymes function as both catalysts and chiral directing groups. The latter two forms of asymmetric catalysts have been seen as very promising alternatives to transition metal catalysts, due to their reduced costs and toxicity. In fact, the potential impact of these alternatives was recently recognized when the Nobel Prize in Chemistry was awarded to Frances Arnold (2018, directed evolution of enzymes), then Benjamin List and David MacMillan (2020, asymmetric organocatalysis), after Barry Sharpless, Ryoji Noyori and William Knowles shared the Nobel Prize in 2001 for transition metal catalyzed reactions.

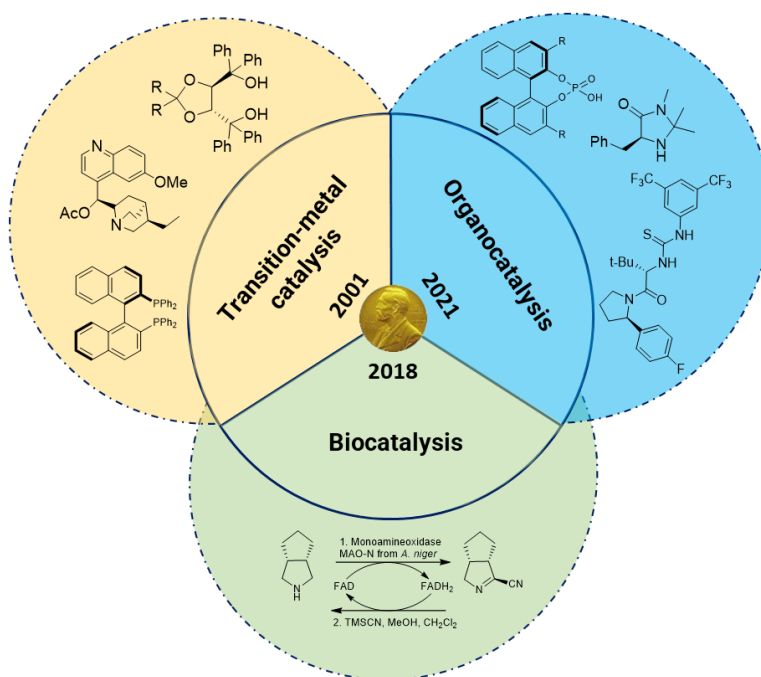


Figure 1.1. The three main asymmetric catalysis fields.

1.1.2. Developing new catalysts.

The relatively simple mechanisms of organocatalyzed reactions are a significant advantage for their development and optimization. However, the use of these catalysts has been hampered by their lower stereoselectivities, and the need for higher loading. Alternatively, development of

transition metal catalysts is facing complex mechanisms often involving multiple possible transition states (TSs), metal coordinations, additives and ligands. As a result, the very tedious “trial-and-error” approach is still commonly used. To address environmental concerns, catalysts based on greener and cheaper metals (bismuth, iron, copper) have been developed, yet the most commonly used transition metal catalysts are built around palladium, rhodium, and other toxic metals. Thus, despite tremendous progress in the field of small molecule catalysis, development of new chiral catalysts remains quite challenging and often calls for stepwise optimization. This is often a time-consuming, labour-intensive process, which requires the synthesis and evaluation of multiple novel ligands/catalysts in an iterative process that is often pursued in an empirical fashion with little guidance other than simple models and intuition. Computational methods guiding the design of new catalysts are sought after and should address this major issue.

1.1.3. Computational methods for catalyst design and discovery.

As an analogy, over the past decades, docking-based virtual screening has found extensive use and acceptance as a design tool in medicinal chemistry.¹ The low computational demands of these methods and user-friendly interfaces removed hurdles towards their widespread adoption. In contrast, computational tools that could improve the process of chemical reaction development remain largely underutilized as predictive/design methods. The power of quantum mechanics (QM) calculations, particularly density functional theory (DFT), is primarily used in a retrospective, *post hoc* fashion for understanding reaction mechanisms and for rationalizing observed selectivities, rather than in the prediction/design of new catalysts. In fact, the computational cost associated with *ab initio* QM or DFT methods, let alone the required expert knowledge, makes them unsuitable for the screening of large libraries of potential catalysts.

However, major efforts are currently ongoing to develop computational tools to assist organic chemists, and integration into organic chemistry laboratories is imminent.²

We believe that the successes of virtual screening in medicinal chemistry should be adaptable to reaction prediction, with the challenge being the availability of computationally inexpensive, rapid, and accurate methods for predicting stereoselectivities associated with complex TS structures and energies.

1.1.4. State-of-the-art approaches.

In the recent years, computer-assisted synthesis has gained significant momentum and several computational methods (most commonly machine learning (ML) methods and statistical models) applied to organic synthesis problems have been reported. For example, computer-aided synthesis planning has advanced rapidly and can propose reaction mapping³ and realistic retrosynthesis^{2, 4-6} (e.g., Chematica/Synthia^{7, 8} and AIZynthFinder⁹) can predict yields,¹⁰⁻¹² catalyst inhibition,¹³ regioselectivity,¹⁴ and chemical reactivity.¹⁵ In practice, these predictive trained methods have been successfully used to design new catalysts¹⁶ and predict the stereochemical outcome of asymmetric reactions.¹⁷ For the latter, ML techniques are advantageous over QM and molecular mechanics (MM) for two reasons: speed (orders of magnitude faster than QM) and their application to reactions with unclear or complex mechanisms (ML models developed from catalyst structures only).¹⁸ However, ML models requires a significant amount of experimental data for training¹⁹ and can hardly be applied to new reactions, as emphasized by Norrby in a viewpoint.¹⁵ Reaction/chemical mapping is more general; for example, physicochemical and QM descriptors of over 300,000 monodentate phosphine ligands have been added to a database named Kraken,²⁰ which will certainly be very useful for new reactions relying on phosphine ligands. However,

despite all these successful developments, much is left to be done in the field including the use of these methods in prospective studies.

In this chapter we will describe the development and application of computational methods for the design of asymmetric catalysts with a particular focus on organocatalysis and transition metal catalysis. For more information about biocatalysis (primarily enzymes), the readers are referred to the excellent overviews of this field by Bell *et al.*²¹ and Pyser *et al.*²² Thus, while computational methods for biocatalysis have been developed,²³ these will not be discussed herein.

In the context of computational asymmetric catalysis, an overarching goal of the computational methods is the prediction of enantioselectivity of asymmetric reactions, enabling computational reaction optimization (i.e., optimal catalysts, substrates, ligands, and conditions). Different methodologies for achieving these goals exist, and they can be divided in two main categories: **1)** methods requiring sufficient knowledge on the reaction mechanism (primarily QM and MM-based methods) **2)** data-driven methods (primarily databases and ML methods). Irrespective of category, the overall principles are similar and involve data collection, identification of meaningful patterns, parameters, or features, model building and testing, followed by model refinement (**Figure 1.2**).

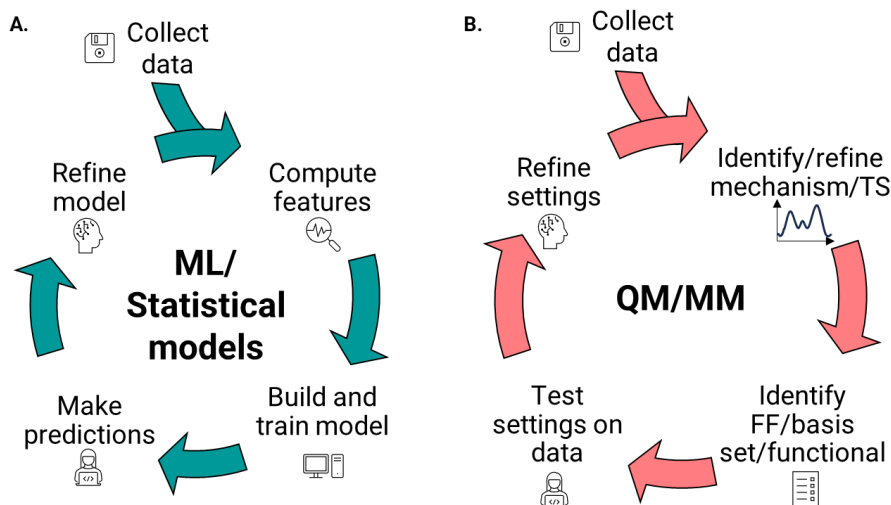


Figure 1.2. Overall workflow for A) ML methods and B) QM/MM methods.

1.1.5. A closer look at MM, QM, and ML.

1.1.5.1.MM.

In MM, the molecules are somewhat represented as charged points (atoms) connected by springs (bonds) and interacting by other means (e.g., angles, torsions, van der Waals, electrostatic) in a 3-dimensional Cartesian space. Because MM does not consider electrons or nuclei, atoms (and the interactions between them) must be parametrized *a priori*. These parameters may be obtained either experimentally (e.g., van der Waals radius of an atom) or from high level QM calculations (e.g., atomic charge, force constants) and stored in lists termed force fields (FFs). Importantly, in MM, atoms in different chemical environments are distinguished by introducing the concept of atom types (i.e., an oxygen atom in water vs. in a carbonyl group would have different atom types), each bearing specific bonding and non-bonding parameters. It is thus important to define the correct atom type for a system under scrutiny, as the total energy of the system is calculated as a sum of contributions from both bonded and non-bonded terms. Due to the size of the chemical space - $>10^{60}$ small organic molecules - the number of possible atom types is virtually extremely

large.²⁴ To make the development of FFs tractable, the number of atom types is limited yet these must be accurately assigned (together with the corresponding parameters) to each atom of any molecule. On the one hand, although the problem of accurately atom typing every molecule seems problematic, significant efforts have been made to exhaustively assign atom types to a large number of small molecules.²⁵ On the other hand, others have proposed approaches discarding atom types altogether.^{26, 27} Thus, the accuracy of MM calculations is strongly impacted by the choice of an FF and the assignment of atom types.

1.1.5.2.QM.

In contrast to MM, QM methods consider both electrons and nuclei. The orbitals required to calculate electronic terms are described by basis sets. Many different types of basis sets have been developed, and their usage depends on the system under scrutiny (i.e., metal complexes, organic molecules, etc). As a rule of thumb, the larger the basis set, the more computationally expensive the simulations. The explicit treatment of electrons significantly increases the computational cost compared to MM, due to the resource-intensive computations of electronic integrals. Depending on the desired level of accuracy, significant speed-ups are possible, either through elegant algorithms for computing electronic integrals (see the ORCA SHARK integral engine²⁸), lower-cost semi-empirical (SE) QM methods where some integrals are neglected (e.g., PM6,²⁹ GFN2-XTB³⁰), composite methods such as HF-3c/PBEh-3c, or DFT. By far the most popular method is DFT, which has seen a tremendous uptake from organic chemists. Most often employed to rationalize reaction mechanisms, DFT has more recently been used to try and explain the reactivity of different types of chemicals (through the conceptual DFT – cDFT – framework). Higher accuracy algorithms such as Møller–Plesset (MP) or coupled cluster (CC) are generally reserved for advanced theoretical work and are not commonly employed in routine organic chemistry

simulations. Overall, the accuracy of the calculations is highly dependent on the choice of method (e.g., MP2, DFT, SE) and basis sets.

1.1.5.3.ML.

ML methods differ significantly from both MM and QM. ML algorithms (and the more interpretable statistical models) aim at finding and quantifying patterns in vast amounts of data for predicting a given outcome. In chemistry, this outcome can come in the form of a single number (e.g., pKa, solubility, enantioselectivity), while the data can be in the form of structural data, reaction data, etc. As a major advantage, ML algorithms are orders of magnitude faster than both MM and QM. In contrast to both QM and MM, most ML models are “black box”, meaning that the way in which an algorithm arrives at a prediction is unknown, although explainable artificial intelligence (AI) is emerging. There are multiple flavours of ML: **supervised**, **unsupervised**, and **reinforcement**, each with their own advantages and disadvantages. For example, **supervised learning** requires labelled input and output training data (e.g., catalyst A provides an enantioselectivity of 45%ee), whereas **unsupervised learning** deals with raw data. In **reinforcement learning**, a feedback loop is employed so that the model can learn from its

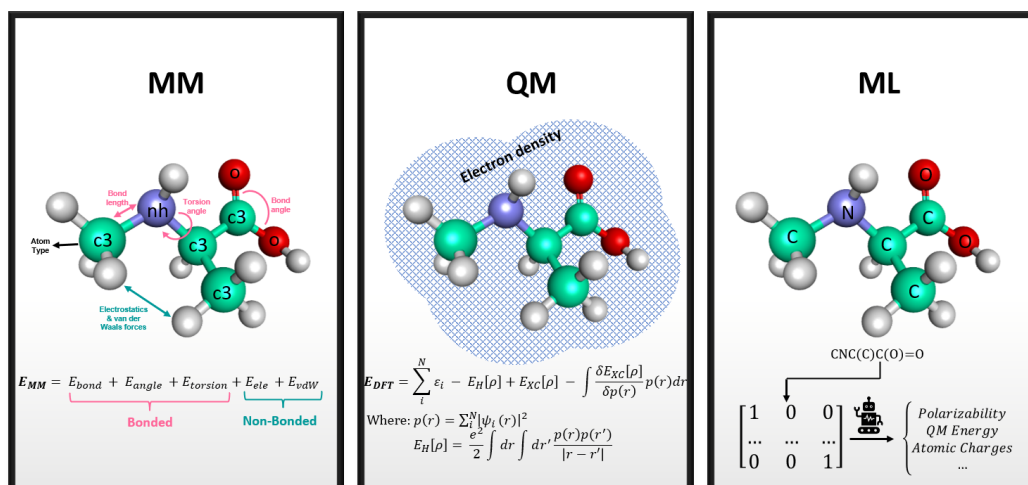


Figure 1.3. Different methods used to compute enantioselectivity.

environment and maximize its correct predictions. Irrespective of flavour or model, the key ingredients for a predictive ML model are *data quality and quantity*.

1.2. Data.

1.2.1. Developing datasets.

The first step in any modeling project is often *data collection* from the scientific literature. Since representative systems are selected and modeled based on these data (ML) and/or used to test the method (QM, MM), *data curation* is essential, as literature data may be riddled with issues (e.g., misassigned stereocenters, incomplete experimental data, reproducibility issues – same conditions used with different results). This process is a challenging task that may be prone to human error. Thus, model or method development and/or evaluation require the availability of consistent and reliable experimental enantioselectivities. The type and amount of data required to develop a predictive model depends on the approach used to model the reaction (evaluation on small sets or training on large sets).

Below, we list several aspects to developing a robust dataset.

- 1) For most methods, it is assumed that the set of substrates and catalysts collected follow a similar (if not identical) mechanism. If more than one competing mechanism is to be considered, substrates and/or catalysts favoring alternative mechanisms should be sufficiently represented. For each mechanism, one should also assume that the enantioselectivity is affected by the same factors.
- 2) In the context of catalyst design using ML methods, the model is more likely transferable to new sets of catalysts and substrates when trained on datasets with larger chemical diversity. If a component is kept constant (e.g., all the reactions in the dataset are applied to the same substrate),

then the model will not be trained to understand the impact of this component (e.g., not applicable to the search for ideal substrates for a given catalyst or prediction of the substrate cope of an asymmetric reaction).

3) Training a model or testing a method requires not only information on highly selective catalysts, but also on poorly selective catalysts. For optimal accuracy, ML models should be trained to correlate selectivity enhancing features (e.g., chemical groups) to high enantioselectivity, as well as selectivity reducing features to poor enantioselectivity, while QM and MM methods must be evaluated for their ability to distinguish good from poor catalysts.

4) TS-based methods derive the enantioselectivity from the energy difference of the diastereomeric TSs, requiring some experimental knowledge of the mechanism. More specifically, the stereoselectivity determining step (or steps) must be known, and, at the very least, a good hypothesis for the TS structure must be available. If not, a separate investigation of the mechanism must be carried out, which can be quite time-consuming.

5) In general, the larger (and more diverse) the dataset is, the more information ML models may learn. However, building a large dataset should be done with care as adding data may also result in loss of diversity and introduction of biases. For example, overrepresenting a class of catalysts may result in the model learning (or even memorizing) mostly about this chemical series (e.g., phenyl better than methyl at a given position on a ring), rather than learning general rules (e.g., steric effects). This can result in significant biases of the model and poor accuracy in the search for novel catalysts. At this point, it is important to note that information following all these criteria is rarely available in a form that can be used immediately (such as a text files or formatted tables). For QM and MM methods, the size of the set is not as relevant, but the diversity should still be a

focus of the data collection, as these methods must be able to capture various effects experienced by various catalysts (e.g., hydrogen bonds, cation- π interactions). In practice, should all this information be available (several catalysts from different classes already developed), one may question the need for a model to design novel catalysts for this reaction. Thus, the transferability of the method to other reactions may be investigated (see applications below). In the case several catalysts are already available but none providing the level of stereoselectivity targeted, a ML model may be required. However, one may question the ability of a model trained on poor to good catalysts to identify excellent catalysts (e.g., based on a different mechanism, on an interaction not experienced by other catalysts). In the case several catalysts are available but only very few are inducing reasonable stereoselectivity, weighting schemes may be used when selecting batches during model development (e.g., using `WeightedRandomSampler` and `Dataloader` in PyTorch) or weighting the loss function may be considered to consider the minority and majority classes equally. An important aspect is that the model would be as general as the data set is: *if the data set contains little variability, the model would likely fail to predict an out of set example.*

6) Depending on the method used, additional information may be needed. For example, in the case of MM, relevant force fields (FFs) are necessary (e.g., for transition metals). This will be discussed further under the descriptors section.

7) In practice, ML models are trained on a first set (referred to as training set) and tested on a second distinct set (testing set), while hyperparameters may be optimized using a third set (validation set). While simple random splitting is still often used, the similarity between these sets must be monitored and minimized. If the model memorizes input data (“ CH_3COOH : pKa 4.75”) rather than learning to predict a property (e.g., “*electronic effects make acetic acid acidic, with a pKa of 4.75*”), the model tends to be poorly predictive (e.g., “ CF_3COOH looks like CH_3COOH

and is predicted to have a pK_a of 4.75”, as opposed to “the fluorine inductive effect reduces the pK_a to 0.2”).

1.2.2. Available datasets.

When MM or QM methods are used, developers often rely on sets of a few dozens of systems to test their methods. However, as more data-intensive ML-based methods become more prevalent, and the need for datasets arises, a few curated datasets have been reported and made available. While these datasets contain information about catalysts/ligands and computed descriptors, the reaction related information still needs to be manually collected.

We have built a non-exhaustive collection of available datasets (see Supporting Information for a detailed breakdown) which may be useful for method developers. We expand on a selection of datasets below:

1. **Kraken.** This database, curated by Gensch *et al.*,²⁰ contains ~300,000 virtual monodentate organophosphorous (III) ligands for asymmetric catalysis. These ligands were combinatorially enumerated *in silico* using a set of 1558 experimental ligands (including commercially available compounds) and 576 unique, diverse substituents. For the set of 1558 ligands, physicochemical descriptors were calculated on conformer ensembles using QM methods. These descriptors were then used as input for ML models trained to predict the physicochemical profiles of the entire virtual library of 300,000 ligands.
2. **OSCAR.** A dataset of organocatalysts assembled by Gallarati *et al.*³¹ This dataset, available online (see Supporting Information), contains 4,000 catalysts collected either from literature or the Cambridge Structural Database (CSD), along with combinatorially enriched sets for carbene catalysts (over 8,000), and non-covalent dual-hydrogen-bond

donor catalysts (ca. 1.5 million). All catalysts have QM-computed stereoelectronic descriptors and DFT-optimized structures available.

- 3. VIRTUAL CHEMIST.** Upon the publication of the VIRTUAL CHEMIST platform for asymmetric catalysis, Burai-Patrascu *et al.*³² made available the data collected for the platform validation. These data include experimental conditions and %*ee* (experimental and computed) for over 350 reactions across 7 reaction classes, involving both transition metal and organocatalysis. The data are available in table format (see Supporting Information for reference 32) and on the group website (<http://www.moitessier-group.ca/>) for structures.

1.3.Descriptors.

1.3.1. Encoding catalysts, substrates and/or TS structures.

There are numerous methods through which molecular complexes and associated information (e.g., solvent, temperature, counter ions) can be represented numerically. Since enantioselectivity is affected by multiple factors, these representations usually include a combination of steric, electronic, and geometric information.^{33, 34}

In TS-based approaches molecules can be represented in several ways, depending on the approach. In MM methods, molecules are represented by Cartesian coordinates, atom types and/or FF parameters. The challenge is to optimize TS structures when most FFs have been developed for ground structures. To solve this, specific FFs have been developed (i.e., ReaxFF), while other methods rely on a combination of ground state reactants and products (i.e., empirical valence bond (EVB),³⁵ multi configurational molecular mechanics (MCMM),³⁶ and SEAM³⁷). Although these methods are available to the scientific community at large, they still require expertise in

computational chemistry, scripting, and/or computer environments to be truly usable in chemistry laboratories. In this context, another two approaches, ACE and Q2MM, were implemented into user friendly platforms (VIRTUAL CHEMIST and CatVS).

In QM methods, molecules are more accurately described than in MM through the usage of *atomic and molecular orbitals* to describe both electrons and nuclei. In terms of computational cost and time requirements, QM calculations are generally several orders of magnitude more costly than MM. However, the major advantages of QM over MM when it comes to predicting TS structures are **1)** the ability to optimize structures without the need for specific parameters (assuming all the necessary elements - in particular transition metals - are included in the basis set) and **2)** obtaining accurate energies and geometries that describe the bond breaking/forming process. While some QM methods (i.e., *ab initio*: Hartree-Fock or post-Hartree-Fock, density functional theory (DFT)) are highly accurate, they can be used effectively for catalyst design primarily in a *post-hoc* fashion and on a limited number of systems.³⁸ Alternatively, semi-empirical QM (SEQM) methods, although less accurate than the former, are significantly faster and can be envisioned as a useful tool in prospectively screening libraries of hundreds of potential catalysts due to their relatively low computational cost. However, the accuracy of SEQM in transition metal catalysis is yet to be demonstrated.³⁹

In ML-based approaches, the information is usually represented by descriptors (referred to as features). Different molecular representations exist (e.g., graph, simplified molecular-input line-entry system (SMILES)) and are linked with different molecular descriptors.⁴⁰ Among those are *system descriptors* (temperature, concentration, etc.), *steric descriptors*⁴¹ (e.g., Sterimol parameters,⁴² average steric occupancy,⁴³ %buried volume⁴⁴) *electronic descriptors* (e.g., Natural Bond Orbital (NBO)- charges, polarizability, Frontier Molecular Orbital (FMO)-gap), and

geometric descriptors (e.g., bond lengths, dihedral angles). These descriptors are often obtained from QM calculations.⁴⁵

1.3.2. Selection of descriptors and their computation.

In TS-based models, descriptors are often chosen to describe the steric and electronic effects governing the reaction. It is important to note here that these descriptors generally have **chemical meaning**. For example, in both QM and MM, electronic descriptors such as *atomic charges* and *dipole moments* may be used to understand catalyst/ligand reactivity.^{46, 47} Additionally, more advanced QM descriptors such as *local* and *global reactivity parameters* (obtainable in the cDFT framework), have often been used to rationalize the reactivity and selectivity of various chemical series in numerous reaction classes.⁴⁸

In the case of statistical or machine learning models, or unknown mechanism, thousands of descriptors may be computed and used for building the model, especially with the advent of specialized software for computing descriptors.^{49, 50} However, without carefully choosing only the most important descriptors that significantly contribute to the prediction, the model is bound to contain a large amount of noise. This affects the accuracy of the predictions. Indeed, different methods to select descriptors^{51, 52} (either supervised or unsupervised) have been developed to address this exact issue. These techniques (for a complete breakdown with examples see Supporting Information) include *filtering methods* (selection based on statistical methods like the chi-squared test), *wrapper methods* (selection based on a predictive model to generate the best descriptor combinations), *embedded methods* (selection is made by learning the importance of each feature during model training), *hybrid methods* (combination of filtering and wrapper methods), and *dimensionality reduction techniques* (selection of features after dimensionality reduction of the data).⁵³ Perhaps the most widespread method is principal component analysis

(PCA), a technique that reduces data dimensions to fewer components while retaining essential information about its diversity. This allows for simpler visualization, although interpretation might not always be straightforward.⁵⁴

If one of the goals is only to develop a predictive model, but also to have chemically interpretable descriptors that can shed light on the mechanism (statistical models), then the descriptor selection requires additional attention. First, chemical knowledge may be used when selecting these. For example, if properties truly influencing the reaction outcome have been identified experimentally, descriptors of these properties may be considered. In this case, the ML or statistical model will eventually quantify the impact of these properties. Unfortunately, irrelevant descriptors may still coincidentally correlate with the property the model aims to predict, leading to poorly predicting models. Generally, the assumption is that the different descriptors are independent of each other (as in, the change of one will not have an effect on the other). However, a counter example can be seen in the work of *Werth et al.*⁵⁵ on bifunctional hydrogen bond donor (BHD) catalysts, where the NBO charge of the catalyst was indirectly correlated with the pKa value via the LUMO energy and a separate steric parameter.

How the descriptors are computed is another fundamental aspect of both TS- and ML-based models. Many descriptors are conformation-dependent: descriptors computed only for a single conformer may not adequately represent more flexible ligands. Should a Boltzmann population average be used instead, a conformational search (hence time) must be added to the computation. A significant advantage of some available databases (e.g., Kraken and Oscar described above) is their computed descriptors, which may be used by other model developers.

1.3.3. Number of descriptors.

With the help of numerous cheminformatics tools^{49, 56, 57} thousands of descriptors can be computed for each model, although the final version of the model will ideally contain less than 10. While *more* descriptors are expected to provide a more complete representation of a reaction, they may also lead to overfitting, a common issue when developing ML or statistical models. As an indication of overfitting, the model performs well on the training set but poorly on the testing set, hence the need for significantly dissimilar training and testing sets to detect overtraining. Moreover, not all the descriptors will have a significant enough influence on the accuracy of the model and will add unnecessary noise. In general, the simpler the model (i.e., the fewer descriptors), the easier (and faster) it is to train and often the better (more generalizable) and interpretable the model will be. In practice, many descriptor combinations are evaluated for model training, and the most predictive set of descriptors is chosen for the final version of the model. However, a careful evaluation of the relevance of these descriptors should be carried out.

1.3.4. Different types of descriptors are used to describe unique atomic and molecular properties.

1.3.4.1. Steric parameters.

Steric effects play an important role in catalyst reactivity and selectivity. In the context of asymmetric catalysis, a catalyst may have different substituents, with each substituent taking up a different volume of the space around the reactive part of the catalyst, influencing the shape, stereoselection, and reactivity of the catalyst. To account for these effects, descriptors representing the steric character of the molecule and its substituents have been developed. An important aspect of some steric parameters is that they are conformation dependent, making the selection of the

conformation impactful. In practice, chemists often rely on C₂ symmetrical catalysts (e.g., BINAP) to reduce the conformational space (e.g., the number of possible conformations, the number of different faces of nucleophilic attack), which in turn simplifies their optimization. Depending on the system, these descriptors may be calculated for the lowest energy conformer, the catalytically relevant conformer, or as a Boltzmann-weighted conformation average. **Examples in text:** Sterimol parameters, % buried volume, ASO.

1.3.4.2. Geometric parameters.

In addition to sterics, one can compute descriptors providing information on the 3D shape (geometry) of the molecule. The geometry of a catalyst affects the selectivity and rate of catalyzed reactions; therefore, the selection of the conformation is critical. In addition to overall shapes, geometric parameters may also contain information such as bond lengths or bond-angle and torsion values. **Example in text:** cone angles.

1.3.4.3. Electronic parameters.

These descriptors aim to represent the electronic properties of the molecule, often of a reactive centre (e.g., the atoms participating in the reaction) or of a ligand (e.g., a phosphine ligand modulating the reactivity of metal centers). In organic chemistry, electronic properties describe information such as ability to donate or accept electrons (Lewis basicity and acidity, respectively), nucleophilicity and electrophilicity, hyperconjugation, and more. Unlike steric and geometric parameters, electronic parameters are often less influenced by conformation, and they can either be derived from a single (optimized) conformation or as a Boltzmann-weighted conformational average. Most electronic parameters describe a fragment of the molecule, such as a substituent or reactive atom, rather than the entire molecule. **Examples in text:** HOMO/LUMO gap, ESP_{MAX}.

1.3.4.4. Empirical parameters.

Descriptors do not always have to be computed theoretically. A variety of empirical descriptors can be used in model development, such as nuclear magnetic resonance (NMR) chemical shifts, NMR ^{31}P tensors, and infrared (IR) spectrum frequencies and vibrations.

1.3.4.5. Interaction Fields. Comparative molecular field analysis (CoMFA).

A method developed for ligand 3D-QSAR studies in drug discovery, which has since been implemented in asymmetric catalysis modeling.⁵⁸ The method aims to correlate reaction outcomes to molecular fields described by the steric and electronic properties of a molecule. To achieve this, molecules are first aligned and then placed in a three-dimensional energy grid. A probe atom is then added at strategic points on this grid and the interaction energy (Van der Waals and electrostatic) between the molecules and the probe atom is calculated at each grid point. These energies are the descriptors that input into regression models, most often partial least squares (PLS), which correlate catalytic activity/enantioselectivity with the computed descriptors.^{59, 60}

1.3.4.6. Comparative molecular similarity indices analysis (CoMSIA).

CoMSIA is a 3D-QSAR method developed as a natural extension of CoMFA by including molecular similarity in the computation of the molecular fields. In addition to the steric and electronic parameters captured by CoMFA, CoMSIA also includes a hydrophobicity term. The calculation of descriptors is then performed in a similar manner to CoMFA.

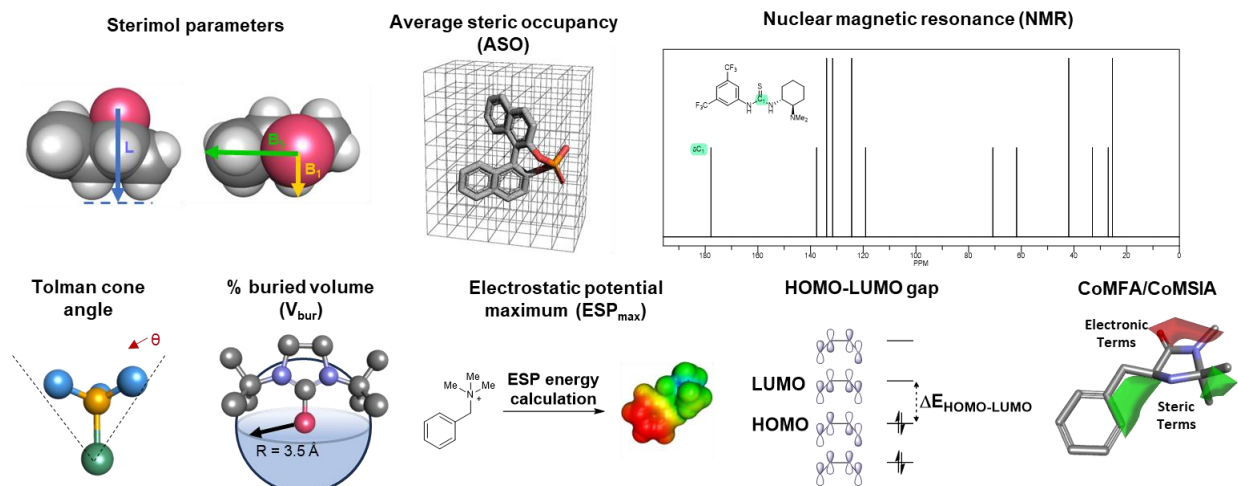


Figure 1.4. Common descriptors used.

1.4. Models.

1.4.1. Model 1. Unknown mechanism/ stereoselective step (statistical and ML models).

Modeling enantioselectivity for reactions and catalysts with unknown mechanisms relies on quantitative structure-selectivity relationships (QSSR), where statistical models can be used to correlate enantioselectivity to the structure of the catalyst and reaction components.⁶¹⁻⁶³ In addition to these, the field has evolved to also incorporate ML methods for the development of predictive models.^{64, 65} While sometimes used interchangeably in the scientific literature, statistical and ML models are in fact distinct. Statistical models are derived from the whole data (where statistical relevance is measured), while ML models are trained and tested on separate sets. Additionally, statistical models are often more interpretable than ML models.

Depending on the desired outcome, different methods can be used to relate the catalyst (and substrate, in some cases) structure to enantioselectivity. For example, Sigman and co-workers have been developing statistical models based on linear free energy relationships (LFER), where they aim to achieve both an increase in prediction accuracy as well as an intuitive understanding of the

potential mechanism.^{63, 66} This type of model aims to find a linear correlation between a variable and free energy, which in asymmetric catalysis is often the energy difference between the diastereomeric TSs ($\Delta\Delta G^\ddagger$). Since stereoselectivity is dependent on multiple variables, the model would usually be a multivariate linear regression (MLR). The input would be the numerical representations of the reaction components (descriptors), and the output would be the free energy difference.^{64, 67} These models have been successfully applied on several metal catalyzed reactions such as the Pd-catalyzed enantioselective aryl-carbonylation of sulfonimidamides,⁶⁸ Pd-catalyzed Hayashi-Heck reaction,⁶⁹ Negishi coupling,⁷⁰ and different Pd- and Ni-catalyzed cross-coupling reactions.⁷¹ Additionally, they have been applied to organocatalyzed reactions including the Mannich reaction,⁵² chiral phosphoric acid catalyzed nucleophilic addition to iminiums,^{17, 72, 73} BHD catalyzed addition of nucleophiles to nitro alkenes,⁵⁵ and others.^{66, 67}

A different type of model used successfully in asymmetric catalysis is support vector regression (SVR). With such an ML algorithm, the data may be correlated to the energy difference (enantioselectivity). In SVR the data points are correlated by a linear line, or a higher dimension curve, within a predetermined margin. The goal of the model is to identify the line that fits the majority of data points which falls within a predetermined margin.⁷⁴⁻⁷⁶ Such models have been successfully applied to organocatalyzed reactions such as the chiral phosphoric acid catalyzed thiol-nucleophiles addition to imines.⁷⁷

The major difference between MLR and SVR is that SVR does not have an underlying assumption of a linear relationship between the data points (descriptors) and the outcome (enantioselectivity), and, therefore, can be more suitable for modelling complex reactions. On the other hand, MLR can be more informative for the interpretation of the influence of each descriptor (e.g., catalyst structure, solvent, temperature), as the coefficients also contribute to their weight of

influence. An SVR model is mainly used for predicting an outcome, and the model itself, often used as a black box, is not easily interpreted. Each model has its strengths and weaknesses, and they are used based on the main goals of the modelling project, as well as the existing data. It is worth noting that other regression models such as kernel ridge and partial least squares (PLS, also referred to as projection to latent structures) have also been used.^{75, 77-79}

Once a predictive model has been developed and tested, it *should* then be adopted by organic chemists. As the field remains in its infancy, most published ML models and methods are only made available in the form of scripts and methodologies. However, there is no available user-friendly package that can be used or models that can easily be trained for novel reactions by organic chemists with minimal expertise in computer science, although this may soon change.⁸⁰

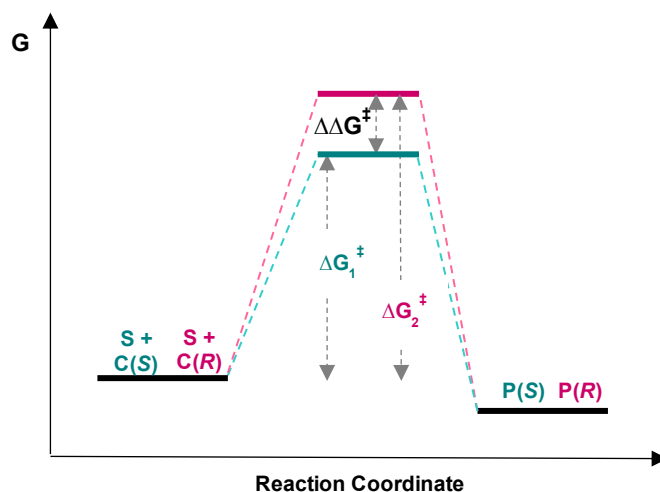


Figure 1.5 Curtin-Hammett principle in the context of asymmetric catalysis. C: catalyst, S: substrate, P: product, either R or S enantiomer. S+C(S) and S+C(R) are the catalyst substrate complexes leading to either the S or the R product, respectively.

1.4.2. Model 2. Using knowledge on potential TS's (mechanism-based approach).

When a hypothesis exists for the diastereomeric TSs of the stereoselective step, these TSs may be modeled and used directly to extract the energy difference between the diastereomeric TSs

($\Delta\Delta G^\ddagger$), and consequently to compute selectivity. Although many methods to model TSs exist, this review is not meant to be exhaustive and will focus on the most recent and advanced methods available, as well as their applications to stereoselectivity predictions. For more detailed information on TS modeling for the prediction of enantioselectivity, we refer the readers to previous reviews.⁸¹⁻⁸³

Calculating the stereoselectivity from the energy difference $\Delta\Delta G^\ddagger$ is based on the assumption that the reaction is under Curtin-Hammett control (**Figure 1.5**). Simply put, the product ratio (e.g., *R:S*) reflects the energy difference between the two competing and irreversible diastereomeric TSs (i.e., the $\Delta\Delta G^\ddagger$).

Since the difference in $\Delta\Delta G^\ddagger$ between a moderately selective catalyst (~80 %*ee*) and an excellent catalyst (>97 %*ee*) is about 1 kcal/mol, there is a requirement for the methods to be accurate enough to be able to distinguish between them. In practice, a prediction error within 1 kcal/mol is targeted. While high level calculations may fulfill this criterion, an objective is to obtain predictions quicker than experimental data and using less expensive equipment. Unfortunately, the most accurate methods for modelling TSs (e.g., MP2, DFT) are computationally demanding (time and computational resources). It follows that they can hardly compete with high throughput experimentation which provides true data (not predictions) in a time-efficient manner. As a result, these methods are primarily used to investigate reactions *post facto* rather than for designing novel catalysts.⁸⁴ Thus, faster alternative methods are necessary.

The energy difference can be calculated from single, lowest lying conformers for each diastereomeric TS identified using a conformational search algorithm.⁸⁵ Alternatively, multiple thermally accessible conformations, and their Boltzmann population distribution, can be used to obtain the energy difference. In both cases, the challenge of TS structure-based approaches is to

identify and optimize *all* potential diastereomeric TSs of the reaction under investigation. In practice, the simplicity of the Curtin-Hammet principle is overshadowed by the number of possible TS conformations, or reactions for which multiple steps could be rate-determining, or for which multiple competing mechanisms leading to opposite enantiomeric products exist.⁸⁶ For these methods to be user-friendly, an automated conformational search algorithm is needed.

Automated conformational sampling of TSs lies at the heart of tools like VIRTUAL CHEMIST/ACE, CatVS/Q2MM, QChASM/AARON (Quantum Chemistry Automation and Structure Manipulation/An Automated Reaction Optimizer for New catalysts),^{87,88} or the chemical steering wheel.⁸⁹ VIRTUAL CHEMIST/ACE, developed in our research group, is a self-contained, graphics user interface (GUI)-based asymmetric catalyst design platform.³² Designed with a chemist's needs in mind, the underlying MM methodology has been thoroughly tested on seven widespread metal and organocatalyzed reactions, with an overall accuracy of ~ 1 kcal/mol. To note, VIRTUAL CHEMIST was also applied to several scenarios that an experimental chemist might face in his project: one-by-one catalyst design, screening a library of catalysts, catalyst lead optimization through analogue search (detailed in the *Evaluation of the Models, Methods, and Applications* section), and identifying the substrate scope of a known catalyst, with demonstrated advantages over traditional asymmetric catalyst design. A similar asymmetric catalyst design platform is CatVS/Q2MM, primarily focused on organometallic catalysts. Similar to VIRTUAL CHEMIST/ACE, CatVS/Q2MM is an MM-based method and was first benchmarked on known metal-catalyzed reactions, followed by its application in a "real world" scenario, which yielded novel catalysts for the Rh-catalysed asymmetric hydrogenation of enamides (discussed in more detail in the *Evaluation of the Models, Methods, and Applications* section). Unlike VIRTUAL

CHEMIST, the free version of CatVS does not include an interface and the calculations must be run from the command line environment.

Apart from VIRTUAL CHEMIST and CatVS, another virtual platform for catalyst design is QChASM/AARON. Contrary to both VIRTUAL CHEMIST and CatVS, QChASM/AARON is an interface to various open-source tools for structural manipulation, TS search and optimization, as well as free energy calculations for %*ee* determination. While VIRTUAL CHEMIST and CatVS are primarily based on MM, the geometry optimization and energy calculations available through the QChASM/AARON interface are based on QM methods (either SEQM or DFT) accessible through software such as Gaussian,⁹⁰ Psi4,⁹¹ or ORCA.⁹² QChASM employs a GUI plugin for the Chimera visualizer,⁹³ which benefits experimental chemists with little to no expertise in command line environments.

1.4.2.1. Computational platforms for asymmetric catalyst design.

1.4.2.1.1. Asymmetric Catalyst Evaluation (ACE).

ACE is an MM-based software that predicts the stereochemical outcome of asymmetric reactions by modeling the relevant TSs of ligand/substrate/catalyst systems. The stereoinducing step for these reactions must be known *a priori*. Part of the larger VIRTUAL CHEMIST platform for asymmetric catalyst design, ACE is built on two fundamental organic chemistry principles: **1)** the Hammond-Leffler postulate and **2)** the Curtin-Hammett principle. The TSs are built in accordance with principle **1)** (i.e., the TS is most similar to the species to which it is closest in energy, either reactants or products, hence is a linear combination of reactant and product structures), while the enantiomeric excess is calculated according to principle **2)** (i.e., the %*ee* is determined according to the difference in energies between diastereomeric TSs). The preferred stereoisomers are

determined through a genetic algorithm that efficiently samples the conformational space around the ligand/substrate/catalyst system. ACE has been successfully tested on seven organo- and metal-catalyzed reactions commonly employed in asymmetric synthesis: Diels-Alder cycloaddition (with chiral auxiliaries and organocatalysts), Aldol reaction, Shi epoxidation, OsO₄-based dihydroxylation of alkenes, ZnEt₂-addition to aldehydes, and Rh-catalyzed hydrogenation of enamides, achieving accuracies of ~ 1 kcal/mol compared to experimental values.³²

1.4.2.1.2. Quantum-guided molecular mechanics (Q2MM).

Q2MM is an MM-based methodology that uses automated FF parametrization to describe TSs and predict the outcome of stereochemical reactions. Similar to ACE, Q2MM is part of a larger catalyst design platform called **CatVS**. To date, CatVS/Q2MM has been primarily employed for organometallic catalysis, with the tested reactions involving Rh-catalyzed hydrogenation of enamides, OsO₄ dihydroxylation of alkenes, ZnEt₂-addition to aldehydes, Pd-catalyzed allylation, and Ru-catalyzed hydrogenation of ketones.⁹⁴⁻⁹⁶ The FFs generated by Q2MM are reaction-specific and are known as TSFFs. Similar to ACE, the stereoinducing step must be known *a priori*. However, in contrast to ACE, Q2MM relies on reference data for a training set of model TSs that is subjected to QM calculations to determine the necessary parameters for FF parametrization. The uniqueness of Q2MM relies on the usage of the QM-derived Hessian matrix (i.e., the variations in energy with respect to geometry changes) to fit TSFF force constants for bonded parameters. Once the TSFF has been generated and validated for a reaction, Monte Carlo (MC) conformational searches are employed to find the relevant TSs and stereoisomers. The %ee's are calculated by Boltzmann-averaging the relative energies of the identified conformations. Q2MM has been tested on four metal-catalyzed reactions, achieving correlation coefficients between 0.8-0.9 between predicted and experimental data.

1.4.2.1.3. An Automated Reaction Optimizer for New catalysts (AARON).

In contrast to both ACE and Q2MM, AARON is an open-source framework that interfaces various tools for structural manipulation, TS searches, and energy calculations. However, like ACE and Q2MM, AARON is part of a larger toolkit named **QChASM**. Designed with ease-of-use in mind, AARON uses a library of TS templates to construct TSs of novel ligand/substrate/catalyst systems, followed by TS optimization at a desired level of theory (semiempirical methods or DFT). Once the TSs have been located, conformational sampling is performed using a rule-based methodology that accounts for the torsional preferences of each substituent. These conformers are then subjected to thermochemistry calculations to obtain free energies, which are Boltzmann-averaged over the populations of conformers leading to specific enantiomers to predict the %*ee*. Representative applications of AARON include Pd-catalyzed Heck allenylation, Rh-catalyzed hydrogenation of enamides, and the Lewis-base promoted propargylation of aromatic aldehydes.

1.5. Evaluation of the Models and Methods and Applications

This section is not meant to be exhaustive but, rather, to illustrate different uses of these methods.

1.5.1. Catalyst design.

The use of computers for asymmetric catalyst design has been a very promising field for two decades. For example, as early as 2003, Kozłowski and co-workers developed a model based on interaction fields for dialkylzinc addition to aldehydes catalyzed by β -amino alcohols and applied it to identify novel catalysts.⁹⁷ However, unexpectedly, twenty years later, while more validated methods are now available to the organic chemistry community, the applications to new catalysts design by groups other than the developers are still scarce.

A representative example of catalyst design is the application of CatVS to the investigation and discovery of novel catalysts.⁹⁸ TSFFs for several reactions including asymmetric dihydroxylation and rhodium-catalyzed asymmetric hydrogenation had previously been developed using Q2MM.⁹⁵ When Q2MM was integrated into CatVS, real-world case studies were carried out. Prediction of the (DHQD)2PHAL-catalyzed dihydroxylation of a dozen substrates revealed a mean unsigned error of about 0.6 kcal/mol, while the screening of rhodium ligands for Rh-catalyzed asymmetric hydrogenation of enamides was also performed. Remarkably, CatVS/Q2MM was able to distinguish between highly stereoselective and poorly stereoselective ligands. For one substrate, the four most stereoselective ligands were among the top 5 predictions. Interestingly, the use of implicit solvent was not found to improve the accuracy, in line with what was observed with ACE.⁸⁵

1.5.2. Stereoselectivity and catalytic activity optimization.

A key aspect of reaction optimization is tuning the enantioselectivity without compromising reactivity. This is especially important in the synthesis of active pharmaceutical ingredients (APIs), as the final product has strict purity requirements. For this reason, high yields of the desired product with low catalyst loading are of significance. An interesting application of this concept has been developed by Dotson *et al.*,⁶⁹ who designed a computational workflow to fine-tune enantioselectivity while simultaneously accounting for catalyst/ligand reactivity in two metal-catalyzed reactions: **1**) Pd-catalyzed Hayashi-Heck and **2**) Rh-catalyzed alkene hydroformylation. These reactions use chiral bisphosphine ligands and are pharmaceutically relevant due to their use in the synthesis of a transient receptor potential ankyrin 1 (TRPA1) antagonist (**Figure 1.6A**), **Compound 20**).⁹⁹

The Dotson workflow began with assembling a set of over 550 chiral bisphosphine ligands, for which steric, electronic, and geometric descriptors were calculated with QM. For each reaction a subset of ligands was selected for experiments to determine the regio-, enantioselectivity, and reaction yields/conversion. The latter was used to discriminate between reactive and unreactive complexes using two classification algorithms: **a)** a single-node decision tree for reaction **1**)⁶⁷ and **b)** a logistic regression classification algorithm for reaction **2**). Consequently, for each reaction, the descriptors and associated experimental data of the reactive complexes were used to train a reaction/metal-agnostic MLR model capable of correlating input data to regio- and enantioselectivity. To verify whether the workflow can be used to prospectively screen for high conversion/high enantioselectivity ligands, the last step involved a virtual screen on the database of ligands not involved in training of the classification or MLR models. Applying the developed classification and MLR models on this database led to the identification of several ligands with excellent experimental conversion and enantioselectivity (**Figure 1.6A**).

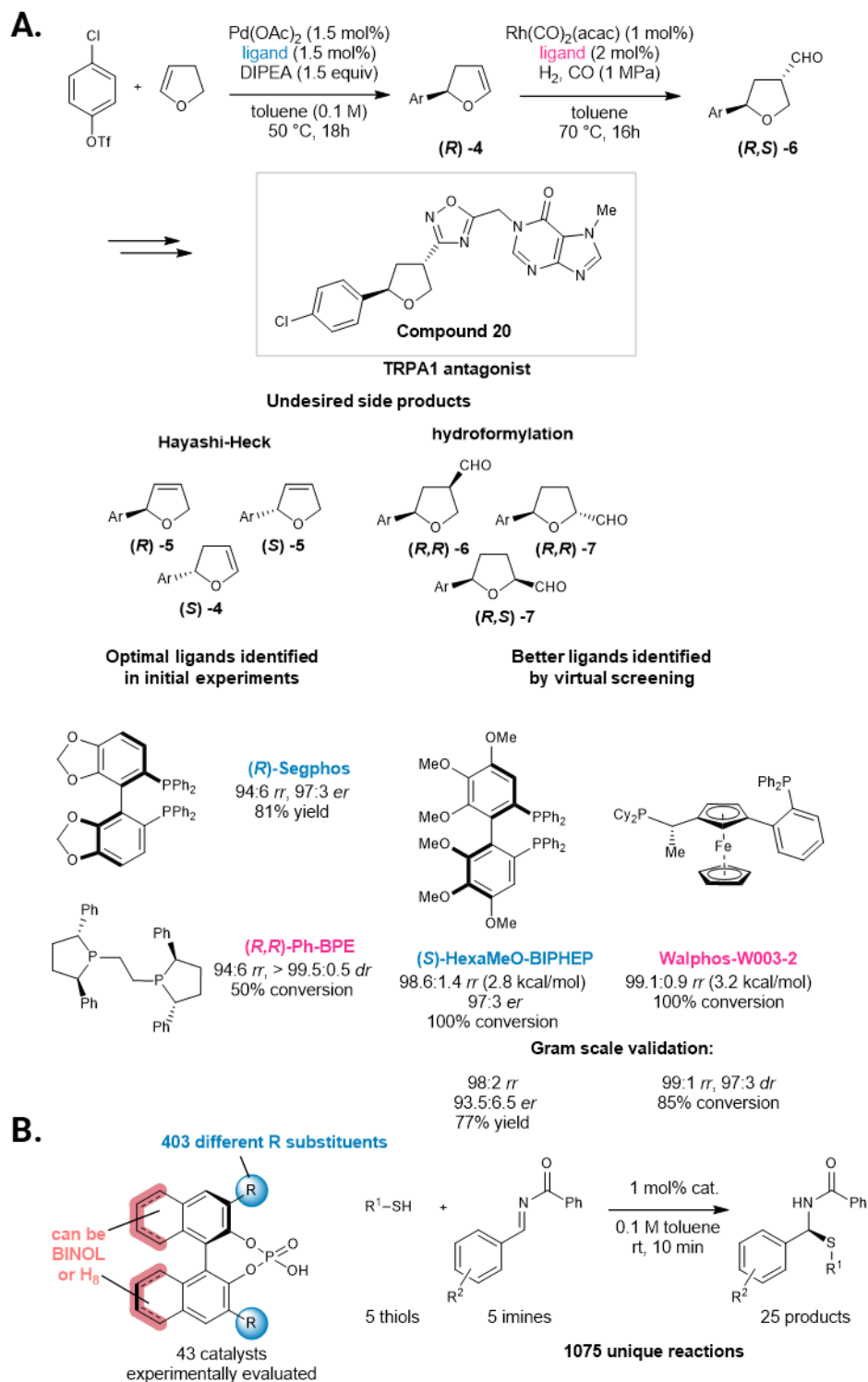


Figure 1.6. A) Dotson and co-workers' workflow. B) Rinehart and co-workers' workflow.

Similar to Dotson *et al.*, Rinehart *et al.* developed a model to predict enantioselectivity, with a focus on the chiral phosphoric acid (CPA) catalyzed thiol nucleophilic addition to imines.^{75, 77, 100} The goals of this work were to **1)** describe the components of the reactions using descriptors that are agnostic to the mechanism and **2)** develop a predictive SVR model without an assumption of a shared mechanism between the different data points. The final model related the catalyst structure to its function (enantioselectivity), in any reaction catalyzed by the input catalyst scaffold (in this case, CPA).

For this reason, the descriptors of choice were ASO for the portrayal of steric information, and ESP_{\max} for the electronic information.^{19, 43, 75} These descriptors are more abstract than the ones commonly used when building an MLR model (e.g., NBO charges, cone angles). However, the emphasis of this work is on descriptors that have been previously shown to work well,⁴³ as opposed to exploring a wide range of descriptors that relate to the reaction mechanism. Therefore, the choice of an SVR model was also appropriate. As mentioned in the section *Model*, the descriptors appearing in the final SVR model cannot generally be used to gain insights into the reaction mechanism and the factors that influence enantioselectivity. Thus, there is no underlying assumption of a shared mechanism/stereoselective step. Once the descriptor library was ready, the next step in the workflow was the use of an algorithm to divide the dataset into training and testing sets, as the more diverse the data the model is trained on, the more likely it is to be transferable to new data points. The training set, termed a universal training set (UTS), represents the variability of the chemical space of the full library.⁷⁸

With a library of 1,075 unique reactions consisting of 43 chiral phosphoric acids, 5 thiol nucleophiles, 5 imine electrophiles (25 possible products) the bulk of the work consisted of developing and testing different descriptor combinations (steric and electronic) with different

models, as well as developing the algorithm for training set selection. Eventually the best performing model and descriptors (SVR with ASO and ESP_{MAX}) were chosen. Interestingly, they demonstrated the ability of the model to predict highly enantioselective catalysts, even when the training set consists of data points of 80 %*ee* or less. This is a significant achievement, as most asymmetric catalyst developments start with only lower selectivity catalysts.

The last representative example we shall discuss in this sub-section was described by our research group during the validation of VIRTUAL CHEMIST.³² In this example, we replicated *in silico* the excellent experimental study by Gerosa *et al.*¹⁰¹ that aimed to identify selective chiral pyrrolidines as organocatalysts for the Diels-Alder cycloaddition. In this report, 22 catalysts were synthesized and tested for their ability to catalyze the Diels-Alder cycloaddition between (E)-cinnamaldehyde and cyclopentadiene. The preparation of these potential levoglucosenone-derived organocatalysts required complex synthesis and separation and characterization of stereoisomers. We developed a workflow using the modular workflow interface in VIRTUAL CHEMIST to simulate the entire process including the parallel synthesis of a small library (ca. 500+) of these organocatalysts including the ones tested by Gerosa *et al.* and evaluation of the induced stereoselectivities. We demonstrated that VIRTUAL CHEMIST not only reproduced the process successfully within just a few days, but accurately identified the most stereoselective catalysts.

1.5.3. Guiding asymmetric synthesis.

An important aspect about model development is its transferability to out-of-set reaction components, and most importantly the application of the model for a synthesis project. The work by Betinol *et al.*,⁷² exemplifies this scenario, by demonstrating how previously developed statistical models can be extrapolated to structurally diverse substrates (**Figure 1.7**). Four goals

were set at the onset of this research project: **1)** application of an existing model to the synthesis of complex scaffolds; **2)** application to the synthesis of a natural product; **3)** preferred catalyst for a given reaction; **4)** preferred route (reaction type) for the synthesis of a given product. For Goal #1 Betinol *et al.* looked into the CPA catalyzed nucleophilic addition to iminiums.^{17, 73} The model was tested on three reported reactions for the functionalization of indoles (as relevant scaffolds for biological compounds) that are catalyzed by CPAs. Importantly, these three reactions are not represented in the training set, Nonetheless, the model was able to predict the enantioselectivity with excellent accuracies (average errors between 0.29-0.54 kcal/mol). Next, for Goal #2 Betinol *et al.* tested a model designed for secondary amine catalyzed reactions,¹⁰² and demonstrated the transferability of their model to new reactions with more complicated conditions that were not represented in their training set.

With both models successfully extrapolated to more complex reactions, Betinol *et al.* moved on to goals #3 and #4. Selecting the optimal catalyst for a given reaction (Goal #3) was tested on the asymmetric epoxidation of cinnamaldehyde, while selecting the optimal catalyst to synthesize a product (Goal #4) was evaluated on the synthesis of diols via two different pathways. The results of both studies were highly encouraging, with predictions being within 1 %*ee* of experimental results.

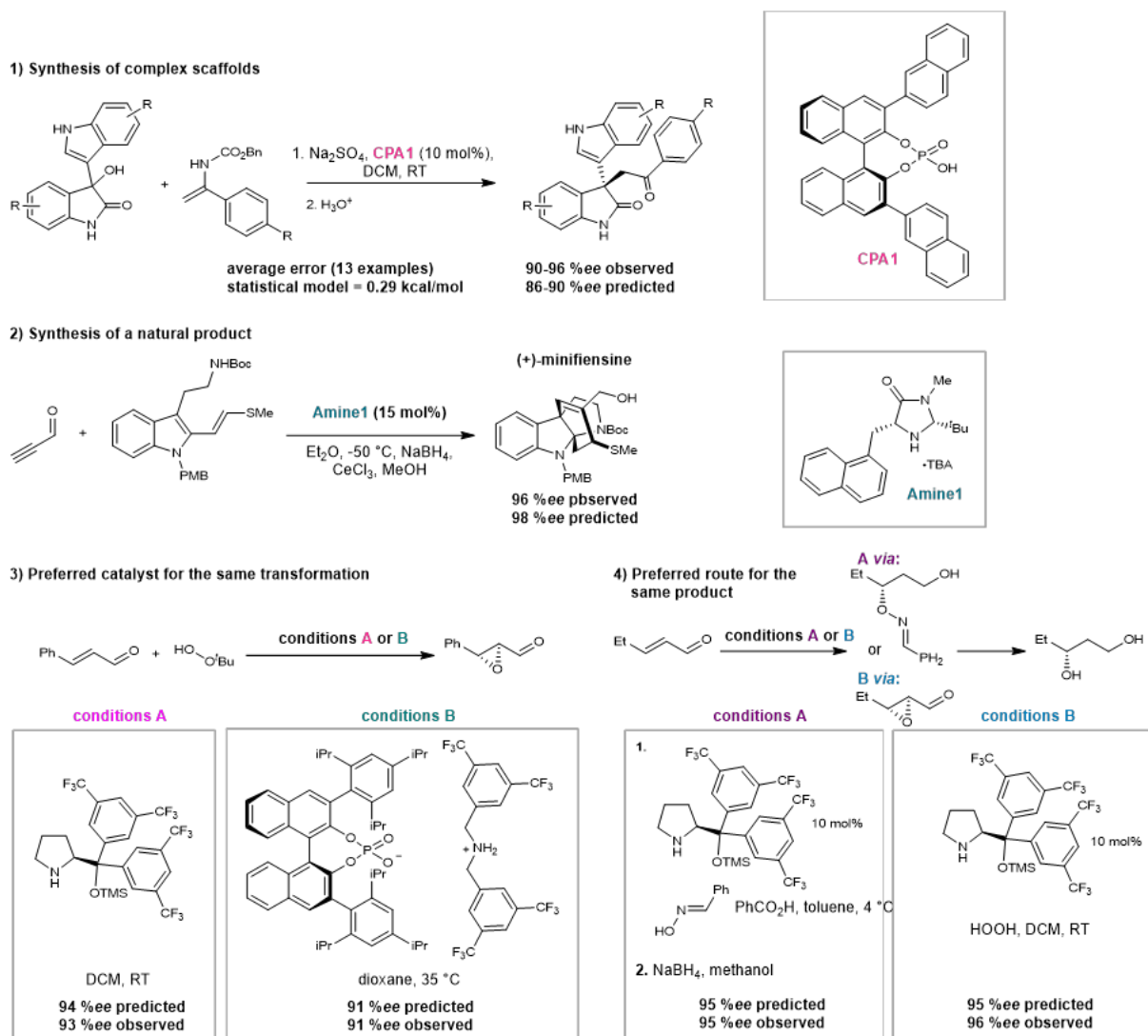


Figure 1.7. Betinol and co-workers' work on the application of existing models to different scenarios: synthesis of complex scaffolds, synthesis of a natural product, finding preferred catalyst for the same transformation, and finding preferred route to form the same product.

1.6. Conclusions and Perspectives

With the advancement of computational methods, the predictions of enantioselectivity of asymmetric transformation have gained momentum. The different approaches are broadly distinguished as ones where the mechanism is unknown (statistical models and ML), and ones where the TS for the enantioselective step is mostly known (QM, MM). Regrettably, the field still requires some understanding of the underlying computational methods and theory before being put to use by experimentalists. Consequently, practical applications of these methods are yet to come. This situation leads most organic chemistry laboratories to persist in employing the conventional, albeit laborious and time-intensive, technique of stepwise optimization. As computational methods continue to evolve and as their accessibility improves, we envision a future where these tools will be completely integrated in the toolbox of organic chemistry laboratories, and where the trained models or platforms will be able to improve the discovery rate and unveil new insights.

With the help of high-throughput experimentation (HTE) for asymmetric catalysis, reproducibility and reliability of data will increase and facilitate the development and integration of computational predictive tools. Automated HTE systems have played a crucial role in accelerating catalyst screening processes, generating vast datasets for diverse reaction conditions.¹⁰³ Integrating computational models as we described in this review alongside high-throughput screening not only help to overcome limitations related to the number of variables that can be tested (temperature, solvent etc.) but will also contribute to the generalization and robustness of the models developed on a solid dataset gathered under the same experimental conditions. This intersection of automated experimental and computation approaches can then

enhance the global efficiency of asymmetric catalysis research, leading to more accurate catalyst design strategies and predictive models.

In **Chapter 1** we introduced the latest computational approaches for the prediction of enantioselectivity. Though there have been tremendous advances in the field, the issue remains – many organic chemists are hesitant to incorporate an *in-silico* design step into their work. Additionally, many of the methods that have been developed are accessible *via* available scripts whose application requires computational expertise. This may deter some chemists from applying them to their work. We hope that by fostering a deeper understanding of computational methods and enhancing accessibility, organic chemists will incorporate predictive software into their work, thereby deriving substantial benefits from its application.

This leads us to **Chapter 2**, where we worked on the validation and tested the useability of the VIRTUAL CHEMIST platform, that has been developed by our group. The platform, which contains the enantioselectivity prediction software ACE (asymmetric catalyst evaluation), is accessible with a user interface and was developed to be used by organic chemists. In this chapter, we relay our work and the hurdles we came across, which will ultimately lead to improved versions of the platform.

This chapter is based on the work published in the paper below, with more details on the initial virtual screenings and chemistry.

Burai Patrascu, M., Pottel, J., **Pinus, S.**, Bezanson, M., Norrby, P.O., and Moitessier, N. From desktop to benchtop with automated computational workflows for computer-aided design in asymmetric catalysis. *Nat Catal* 3, 574–584 (2020).

NM developed ACE and the UI for the VIRTUAL CHEMIST platform. JP and NM developed the CONSTRUCTS software. MBP developed QUEMIST and implemented it in the VIRTUAL CHEMIST platform. MBP and SP performed the VS studies. PON developed Q2MM. All authors contributed to writing the manuscript.

Chapter 2

2.1. Introduction

2.1.1. Chiral molecules.

Small chiral molecules have a great significance in organic chemistry, in fields involving biological systems and drug discovery,¹⁰⁴ materials science (e.g., OLED),^{105, 106} agrochemicals,¹⁰⁷ and more. In the context of the pharmaceutical industry, while some drugs are sold as racemic mixtures, having access to the enantiopure compounds has a considerable benefit, since for many drugs only one enantiomer is in fact active.¹⁰⁸⁻¹¹² Organic chemists can use different approaches to synthesizing enantioenriched products: using enantiopure starting materials (e.g., amino acids, natural products such as lactic acid, chiral auxiliaries), separation of the racemic mixture (e.g., chiral HPLC, selective crystallization, chiral resolution),¹¹³⁻¹¹⁵ and using an asymmetric catalyst.¹¹⁶ The use of an asymmetric catalyst has some advantages over the others: starting materials for these catalyzed reactions are not limited to an existing library of enantiopure compounds; synthesizing an excess of the desired enantiomer is considerably less wasteful (catalytic process) than having to separate a racemic mixture, where about 50% of which might not be useable (unless dynamic resolution is used); and in many cases the catalyst can be recycled and reused.¹¹⁷ The past 25 years have seen a significant advancement in the development and use of asymmetric catalysts, many of which are also shown to be compatible with large scale synthesis.^{118, 119}

2.1.2. Asymmetric catalysis.

The development of asymmetric catalysts often relies on years of tedious efforts, screening of different molecules, and slow improvement and broadening of the substrate scope. The success

of the process partially relies on serendipity since there is not always an intuitive understanding of the structure-selectivity relationship. The considerable number of variables in the optimization of a catalyst and reaction conditions (different ligands, substituents, solvent, temperature, etc.) results in a long process that may not always lead to desirable results. This is especially relevant to any industry where large-scale synthesis requires the most efficient pathway to be taken (e.g., pharmaceutical, agricultural, food, etc.). The product should be obtained in high yield and purity, particularly when toxic heavy metals are used (for environmental and public health reasons).¹²⁰

2.1.3. Computational methods and existing software for enantioselectivity prediction.

As discussed in Chapter 1, significant advances in computational chemistry in recent years resulted in the slow integration of computational calculations into organic chemistry research. In practice, while computational methods have been improving significantly, their useability by non-expert users is still poor. Frequently, organic chemists will collaborate with computational chemists to carry out calculations. Additionally, the calculations are usually done after the experiments have been carried out and are used to rationalize observations. Although this is the standard so far, we hypothesize that computational chemistry can be made accessible to non-experts. A good example of computational chemistry being used daily by organic/medicinal chemists is in the field of drug discovery – different docking software have been developed for non-experts. Docking is usually done prior to synthesis; the ligand may be optimized based on the predictions, and only the potential best ligands would be synthesized and tested. Following our successful development of our drug discovery platform FORECASTER, now widely used by medicinal chemists all over the world, our group has developed a platform – VIRTUAL CHEMIST - with similar ideas in mind: what if an organic chemist was able to screen different potential asymmetric catalysts for a specific

reaction, optimize them *in silico*, and only synthesize and test the best options? A few key points should be considered when such a software is designed for non-experts:

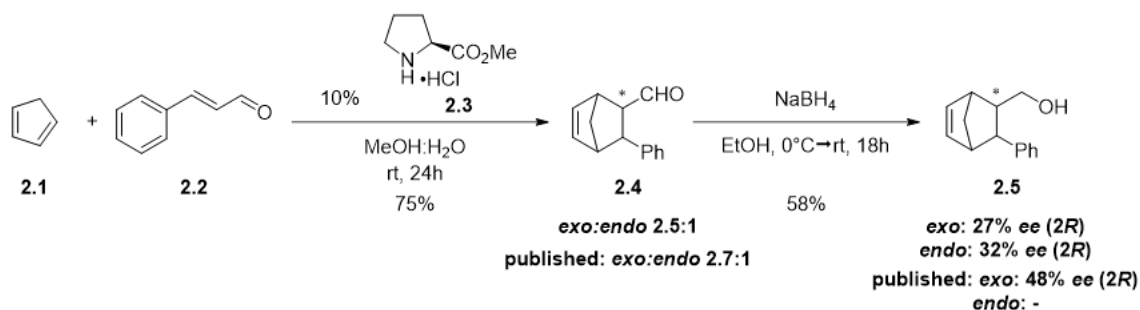
- a) **The computations should be faster than the laboratory experiment:** long calculations that take up weeks to complete are not attractive to a chemist that can run the experiment in the lab in a few days, or even in a few hours. If using the software reduces the amount of time spent on “dead ends”, investing time into optimization and testing of only a small number of potential catalysts becomes an attractive strategy.
- b) **The software should be easy to use:** aiming to be used by non-expert organic chemists, the basic concepts of the software should be clear, and the user interface should be intuitive. Similar to using Nuclear Magnetic Resonance (NMR), organic chemists usually understand the basic concepts of the experiment and are able to know how to use it and how to analyze the results. However, deeper understanding of the physics and math behind an NMR experiment is not necessary for most organic chemists.
- c) **The results need to be as accurate as possible:** this point is obvious. If the results obtained from using such a software are reliable, chemists are much more likely to incorporate it into their toolbox. However, this point poses a considerable challenge since the energy difference between different diastereomeric transition states (TS) can be as low as 1 kcal/mol (a 1 kcal/mol energy difference between diastereomeric TSs results in ~69 %*ee*). Unlike the binding energy of a ligand to a protein docking program predicts (ca. > 12 kcal/mol for nanomolar inhibitors vs. < 3 kcal/mol for inactive compounds), prediction of enantioselectivity must be considerably more accurate (with error < 1 kcal/mol) in order to distinguish between good, to moderate, to bad catalysts.

2.2. Results

2.2.1. Running Diels-Alder with proline methyl ester hydrochloride

The goal of the project was to have both an experimental and a computational validation of VIRTUAL CHEMIST. Computationally, seven reaction classes were considered (see section 2.2.3.1.), and we subsequently focused on two of them to be experimentally used for validation: the organocatalyzed Diels-Alder cycloaddition,¹²¹ and the Shi-epoxidation.¹²² Shi-epoxidation will be described in more detail in Chapter 4.

An important first step was to ensure reproducibility of published experimental results of a known catalyst. Not only is it important to establish a working experimental protocol for running the reaction (since this reaction will be used as reference for all future reactions), but it is also equally important to ensure the product can be purified and the %*ee* can be measured. For the Diels-Alder cycloaddition experiments, proline methyl ester hydrochloride (**2.3**) was used as the catalyst (**Scheme 2.1**). A procedure for running the reaction and purification of the product was established based on MacMillan's work.¹²³ For enantioselectivity measurement, many (if not most) published procedures use chiral-GC, which is unavailable in our research facility. Gratifyingly, a procedure for reducing the aldehyde product into the corresponding alcohol followed by %*ee* measurement using a chiral HPLC column was published by Gotoh *et. al.*¹²⁴ Once a full procedure (for all the steps) was established, we were well positioned to now search for potential new catalysts for the reaction and test them for their selectivity.



Scheme 2.1. Established procedure for carrying out the Diels-Alder cycloaddition with a secondary amine salt and obtaining the enantioselectivity from the alcohol.

2.2.2. Validation of VIRTUAL CHEMIST: virtual screens and experimental tests

In this section, the goal was to conduct a virtual screening of commercially available compounds, synthesize or purchase a selected few, and test their enantioselectivity experimentally. The virtual screenings were done using our in-house platform, VIRTUAL CHEMIST. Using a modular workflow, the software used will be discussed below.

VIRTUAL CHEMIST. A platform developed by our research group, that may be used during the process of asymmetric catalysts design, synthesis, and optimization. The platform streamlines the process and enables the user to create their own workflow, comprised of the different software that have been incorporated into the VIRTUAL CHEMIST platform. The user may create a virtual library of molecules, either by downloading an existing catalogue, or by using REACT2D (described below) to perform combinatorial chemistry and create a new library. The potential catalysts may then be filtered based on user selected properties (see REDUCE below) and assessed for their enantioselectivity using ACE. VIRTUAL CHEMIST allows the user to carry out a catalyst design/optimization campaign virtually, before proceeding to synthesis. Additionally, by containing a user interface, the platform may be used by organic chemists, without the need to

have any experience in computational chemistry and coding. **Figure 2.1.** illustrates the parallels of molecular screening workflows - experimentally (left) and virtually (right).

FINDERS. Filtering, Identifying, Negating Duplicates and Evaluating Reaction Substructures. A program for searching of substructures within a library of molecules, based on a reaction (or molecular structure) scheme. Also, incompatible groups can be defined and filtered out.^{125, 126}

REDUCE. Recognition and Elimination by Descriptors of Undesired Chemical Entities. Filters out user selected incompatible groups from a library (e.g., carboxylic acids, alkenes, etc.). REDUCE is also used to only keep desired substructures like amines.

SMART. Small Molecule Atom typing and Rotatable Torsions assignment. This program may be used to compute molecular descriptors (e.g., molecular weight, presence of functional groups) which will in turn be used by REDUCE to move undesired chemicals.

SELECT. Selection and Extraction of Libraries Employing Clustering Techniques. This program clusters the molecules in the library based on similarity. The clustering takes place over several iterations, until a user defined number of clusters is reached, or until the diversity of each cluster is reduced below a user defined level. Then, a library of the most diverse compounds may be created, with the most representative molecules of each cluster. This enables the user to reduce the library size to representative molecules, thus reducing computation time. SELECT may also be used to remove any molecules that are too similar to a user defined molecule (e.g., known catalyst).¹²⁶

DIVERSE. Duplicate Identification Validated by Evaluation of Regio- and Stereochemical Exactitudes. Used to ensure the library does not contain any duplicates.

REACT2D. Rapid Enumeration by an Automated Combinatorial Tool in 2D. Can be used to create a combinatorial library of products, based on two libraries of starting material and a user input general reaction scheme (e.g., a library of amines and a library of aldehydes that will undergo reductive amination).¹²⁵

CONSTRUCTS. Converting and Orienting Native Structures on Templates of Rotatable and Unoptimized Chemical Transition States. Using a template of the TS and 2D sketches of the catalysts and substrates, CONSTRUCTS assembles the reactants into TSs with reasonable geometry that will later be optimized by ACE.

ACE. Asymmetric Catalyst Evaluation. Employs a conformational search on the TSs created by CONSTRUCTS and optimizes them. Predicts enantioselectivity and favoured TSs structure.^{85, 127}

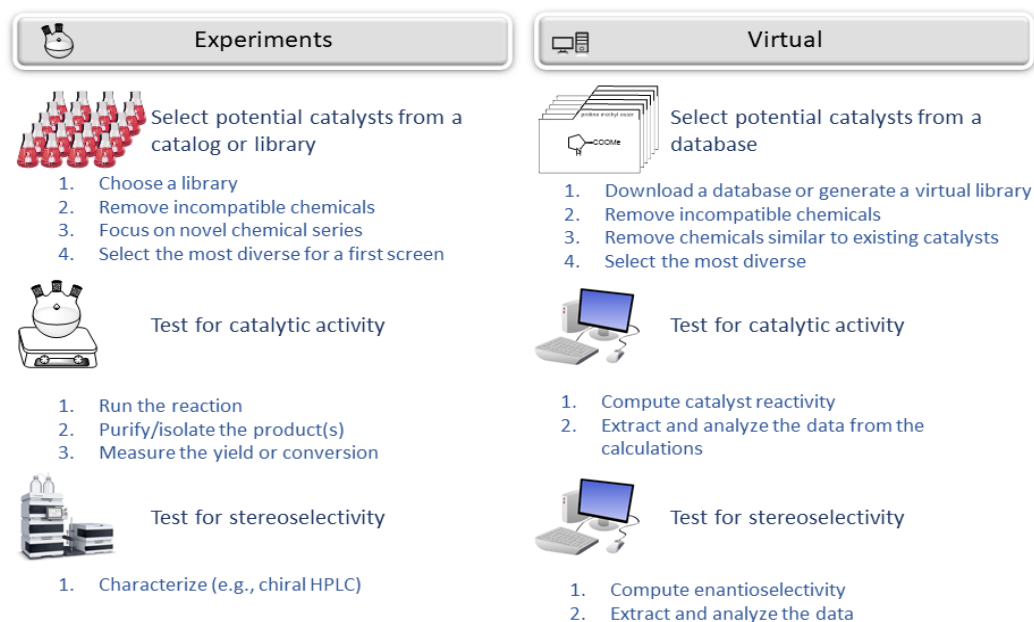


Figure 2.1. Comparison of molecule screening workflows. Left- experimental chemists, right- computational chemists.

2.2.2.1. Virtual screening #1: ZINC database

From the ZINC database,¹²⁸ 140,000 molecules containing amines were selected for screening. As this reaction is known to be catalyzed by secondary amines, we extracted those from the initial library. Thus, using FINDERS, the library was first filtered to contain only secondary amines, which represented 50,000 molecules. Next, non-compatible, reactive groups were filtered out (i.e., carboxylic acids, double bonds, aldehydes, etc.), using REDUCE, approx. 20,000 molecules remained. Out of these 20,000 molecules, the 2,000 most diverse were selected using SELECT and DIVERSE, and used as input into ACE, which predicted their enantioselectivity.

The molecules were ranked based on their predicted enantioselectivity. Initially, we were planning to closely examine the top 50 molecules. However, we soon came to the realization that we had to consider more molecules in the rank-ordered list, since challenges arose at this point: 1) Many of the top ranked molecules contained sulfonamides. These classes of functional groups are not well parameterised in MM3, the forcefield used by Ace. Thus, the predicted stereoselectivity is not expected to be reliable. 2) Although the ZINC database was supposed to only include commercially available chemicals, many of the molecules that were on our list were not commercially available at all. Many were only available as racemic mixtures, including the chiral starting material for several molecules (should we decide to synthesize them). It became apparent that the ZINC database assigns chirality to molecules, even the ones that may not be available in enantiopure form and does not provide this information. At this stage, we opted for synthesis; since the goal was to compare predicted selectivity to experimental selectivity, we prioritized secondary

amines that were synthetically feasible; ones that could be synthesized and tested quickly, over ones that were predicted to be more selective but were structurally more complex (**Figure 2.2**).

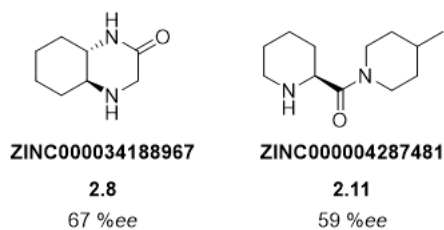
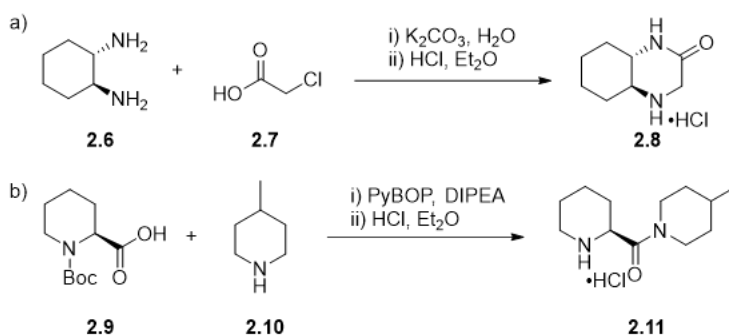


Figure 2.2. Amines **2.8** and **2.11** were selected for synthesis and testing. The ZINC ID is provided under the molecules, as well as the average predicted enantioselectivity.

The amines chosen were synthesized and tested as catalysts for the Diels-Alder cycloaddition (**Scheme 2.2**). Amine salt **2.8** was prepared by a one pot coupling and cyclization of (1*S*,2*S*)-cyclohexane-1,2-diamine (**2.6**) with 2-chloroacetic acid (**2.7**). The recrystallized free amine product was then turned into the hydrochloric acid salt using anhydrous HCl in Et₂O. Amine salt **2.11** was synthesized through a PyBOP coupling of *N*-*boc*-L-pipecolinic acid (**2.9**) with 4-methylpiperidine (**2.10**). Using anhydrous HCl in Et₂O was used both for the deprotection of the *boc*-group and to form the hydrochloric amine salt (**2.11**).

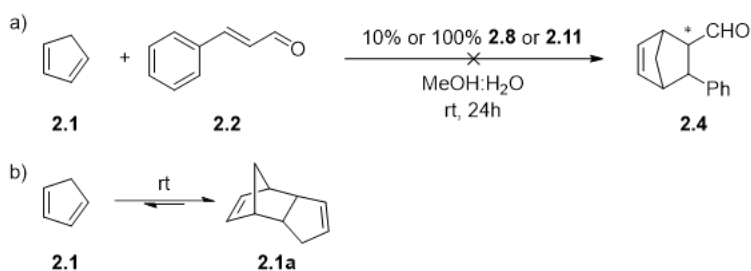


Scheme 2.2. Synthesis of the two selected potential catalysts from the virtual screening. **a) i)** K₂CO₃, H₂O, 10°C to rt, 24h, then 90°C for 2h (52%). **ii)** HCl in Et₂O, THF, 0°C to rt, 24h. (quant.). **b) PyBOP**, DIPEA, DMF, 0°C to rt, 18h (89%). **ii)** HCl in Et₂O, THF, 0°C to rt, 24h. (67%).

Once the synthesized amines were ready, the next step was testing them for their enantioselectivity in the cycloaddition, using the common substrates used in this reaction– E-

Chapter 2

cinnamaldehyde (**2.2**) and cyclopentadiene (**2.1**). Under the initial assumption that the reaction would go with enough conversion using the established protocol, we started by using 10 mol% of the amine salts. Unfortunately, no product was formed. Understanding that the amines we used may need higher loading, we decided to use 100 mol%, since at this point, we were aiming to get enough product so that the %*ee* could be measured. Disappointingly, even at 100 mol%, only trace amount of product was formed. We hypothesised that this lack of conversion could be a kinetic issue. It is possible that the product does not form in the time scale of the reaction, and perhaps the amines used are less reactive than the published ones. Since the cyclopentadiene needs to be used fresh, and will dimerize over time, the reaction could not be left running for a few days, as it will be competing with the dimerization of the cyclopentadiene. (**Scheme 2.3**) Even though the cyclopentadiene is added in excess (3 equiv., relative to the aldehyde), the reaction timescale is limited by the dimerization of the substrate.



Scheme 2.3. a) Testing of amines **2.8** and **2.11** as catalysts of the Diels-Alder cycloaddition. Each catalyst was tested both in 10 mol% and in 100 mol% (relative to the aldehyde). b) Cyclopentadiene (**2.1**) dimerizes into its dimer (**2.1a**) over time, at rt.

This set back brought to light an important drawback of the platform. VIRTUAL CHEMIST, and ACE within it, were designed to screen a library of molecules and predict their enantioselectivity. The software works under the “assumption” that the chemicals being screened will *work* as catalysts. It did not, at the time, rank or filter the molecules based on their reactivity as catalysts in the given reaction. In other words, we had no guarantee that the secondary amines that have been screened will indeed catalyze the reaction. An additional issue became apparent while analyzing

the VS results on ZINC database. The analysis became very time consuming, since a good portion of the amines were simply not synthesizable for this project. We concluded that the ZINC database is not suitable for this type of project and looked for a database that had enantiopure molecules that were commercially available, either the starting material or the molecule itself. The ZINC database was later used again for the purpose of computational validation (see section 2.2.3.1.) of VIRTUAL CHEMIST. Since the goal was only to have a large number of molecules to screen, availability and synthetic feasibility was not a concern.

2.2.2.2. Virtual screening #2: ChemSpace

Considering our conclusions from the previous virtual screening, we decided to look into a library that contained commercially available, enantiopure molecules. We ran a second virtual screening, this time using a library of chiral molecules from ChemSpace¹²⁹, which contained 100,000 molecules. The library was filtered for chiral secondary amines using FINDERS and CONVERT, leaving 50,000 molecules. Incompatible groups (including sulfonamides) were filtered out using SMART and REDUCE, and 20,000 molecules remained. The 2,000 most diverse molecules were chosen using SELECT. These 2,000 molecules were ranked for their potential reactivity using the newly implemented QUEMIST (QUantum Energy of Molecules Inducing Structural Transformations). QUEMIST can be used to compute global reactivity parameters for molecules of interest. The molecules can then be filtered using these parameters. In this case, the molecules were filtered based on the local nucleophilic Fukui function of the sp^3 nitrogen atom.^{130, 131} The top 800 molecules were then chosen (SMART) and ranked by selectivity using ACE.

While analyzing the results, we noticed that the top ranked molecules were cyclic amines (**Figure 2.3**). However, most known catalysts for this reaction are also cyclic amines,^{132, 133} and our hope was to investigate a different kind of scaffold to truly show ACE's predictive power. Interestingly, some of the molecules that came up were non-cyclic amines, that are adjacent to a ring (e.g., bottom 3 molecules in **Figure 2.3**).

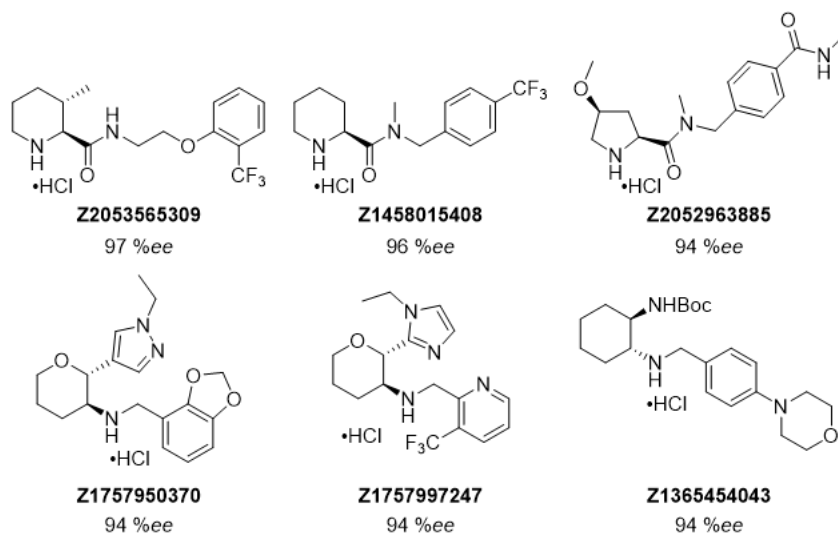


Figure 2.3. Top 5 results of the ChemSpace virtual screening. Under each molecule is the ChemSpace ID and the ACE predicted enantioselectivity.

Curious about the selectivity of such molecules, we decided to focus on these types of amines for our testing. It is important to note that ChemSpace is a library that is commonly used for screening of active molecules in drug discovery. On its own, this should not make a difference to us. However, this results in the molecules being sold in very small quantities for high prices (e.g., 1-5 mg for \$100USD or more). While a few milligrams of compound are enough for testing for bioactivity, we knew this would not be feasible for our purposes, since we need enough product to purify and measure the selectivity, in addition to the fact that we might want to run the experiment more than once and with higher loading. Our solution to this hurdle was to draw inspiration from the screening results and compile a library of molecules that we can purchase as enantiopure and

diversify by adding different groups to them. This would be done by creating a new combinatorial library, using REACT2D.

2.2.2.3. Virtual screening #3: React2D library

Based on the results of the previous screening, we chose a few scaffolds of non-cyclic, chiral, primary amines that were adjacent to a ring (**Figure 2.4**). Two of the scaffolds also contained a hydroxy (**2.13**, **2.14**) that was used for further diversification by introducing different substituents (**Figure 2.4b**). The second amine group on the diamine **2.12** was substituted with Cbz. The amine library thus contained 10 different primary amines that could be used in reductive amination. Next, we looked at different aldehydes that we had available to create the aldehyde library for REACT2D. In total, the aldehyde library contained 34 different aldehydes.

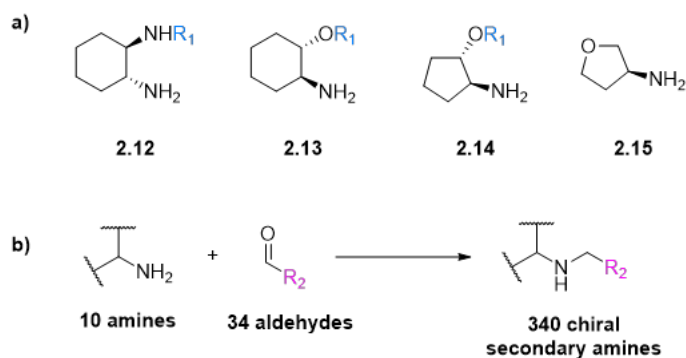


Figure 2.4. a) Amine scaffolds used to create the combinatorial library. In the starting materials $R_1=H$. b) A library of 10 amines reacted in REACT2D with 34 aldehydes in reductive amination. The final library contained 340 chiral secondary amines.

REACT2D was then used to create the library of chiral secondary amines, by using the libraries of aldehydes and of primary amines as substrates for a reductive amination (**Figure 2.4b**). The final library contained 340 molecules to be screened for enantioselectivity by ACE. Once we had the library in hand, the enantioselectivities were computed using ACE.

The top 5 molecules are shown in **Figure 2.5**. Since PMB was the largest group for R₁, it most likely had the largest effect on conformation and therefore provided better selectivity. Additionally, amine **2.15** was not ranked in the top spots. This is likely because unlike the three other scaffolds, this amine does not have an additional substituent (additional chiral center), which could result in more flexibility in the TS, leading to lower selectivity.

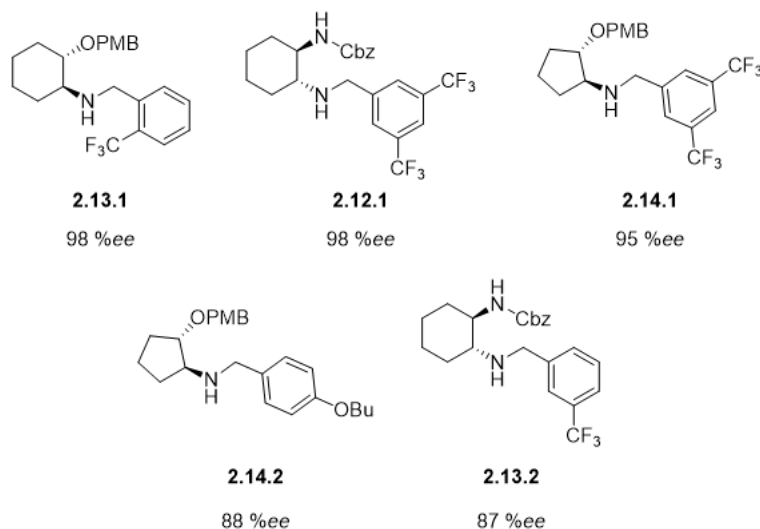


Figure 2.5. Top five amines ranked in the virtual screening of the library of analogues. Predicted selectivity is shown as %ee's provided under the molecules.

Although the screening was completed, we did not get the chance yet to synthesize and test the compounds. We decided to prioritise understanding the reactivity of the amines better, so that in the future it could reliably be computed.

At this point, it was necessary to take a step back and ask a few questions: when no product was observed - was the iminium forming? If the iminium was forming, is it possible it is not reactive enough and gets hydrolyzed too quickly? How can we filter the library, so that it will only contain amines that will give product?

2.2.3. Validation by application to four real-life scenarios

As experimental validations came with numerous challenges, we thought to exploit reported experimental data to test our platform. To do so, we considered a few “real-life” scenarios which a chemist may encounter, and apply them to VIRTUAL CHEMIST:

- I) Assessment of a small number of catalysts (or just one) at a time.
- II) Finding a new series of catalysts for a known reaction.
- III) Designing and virtually screening analogues for a hit with improved selectivity.
- IV) Evaluation of substrate scope for a catalyst.

2.2.3.1. Real-life scenario I: Assessment of a small number (or just one) catalyst at a time

The catalysts can be drawn using a 2D sketcher, and the TS can be used from an existing template for the given reaction or built based on literature information. Each structure can then be tested virtually for its potential selectivity. We applied this scenario to over 350 reactions (each reaction comprising of specific catalyst and substrates) from seven reaction classes (**Figure 2.6**):

1. Diels-Alder cycloaddition (chiral auxiliaries), **2.** Organocatalyzed Diels-Alder cycloaddition (iminium catalyzed), **3.** Organocatalyzed Aldol reaction (enamine catalyzed), **4.** Organocatalyzed Shi epoxidation, **5.** Dihydroxylation of alkenes (OsO_4), **6.** Diethylzinc addition to aldehydes, **7.** Hydrogenation of alkenes (Rh-catalyzed).

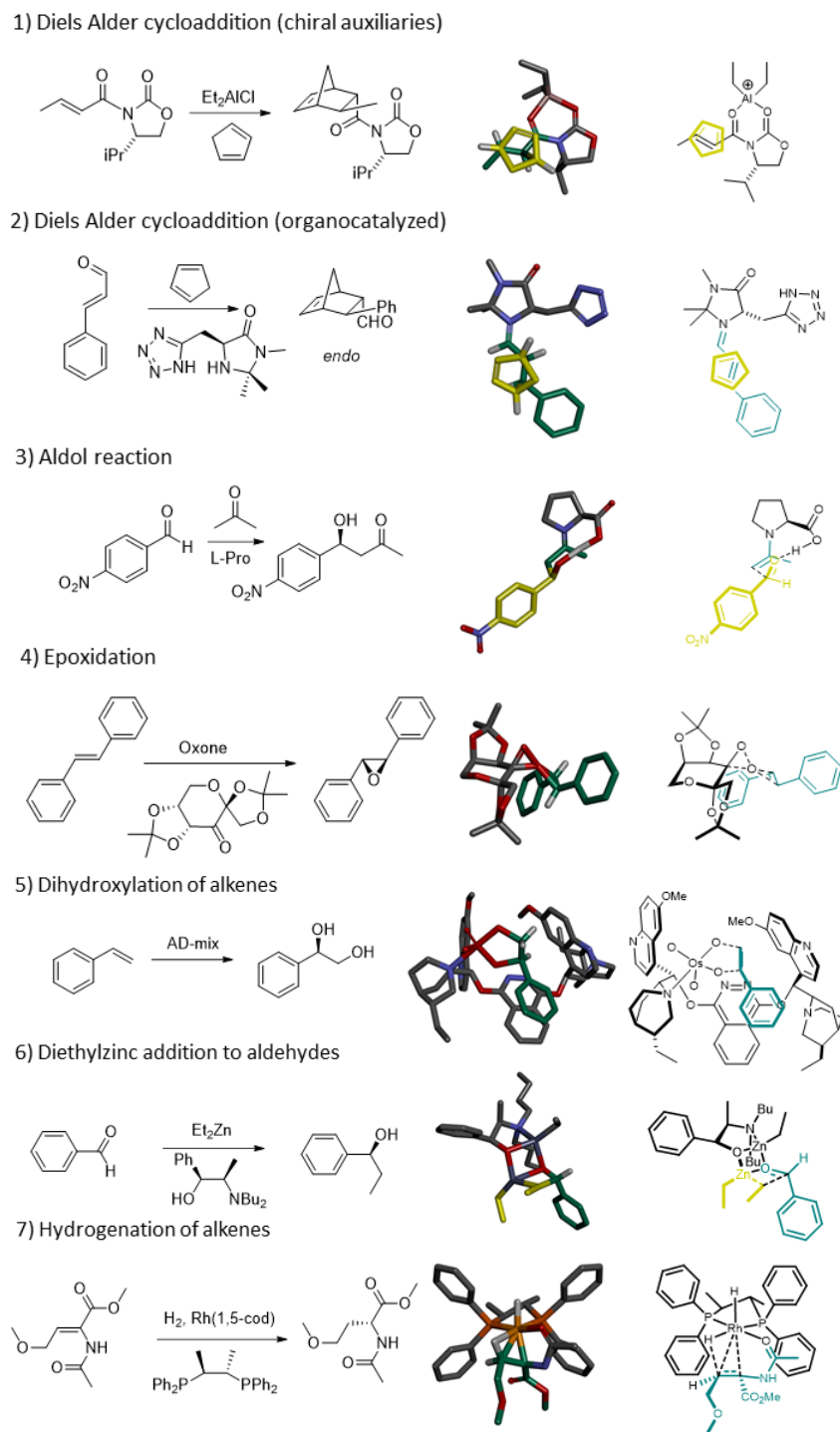


Figure 2.6. General reaction schemes for the 7 reaction classes that ACE was tested on, with selected examples of ACE-optimized TSs in 3D and 2D. Reproduced from reference Burai Patrascu *et al.*³²

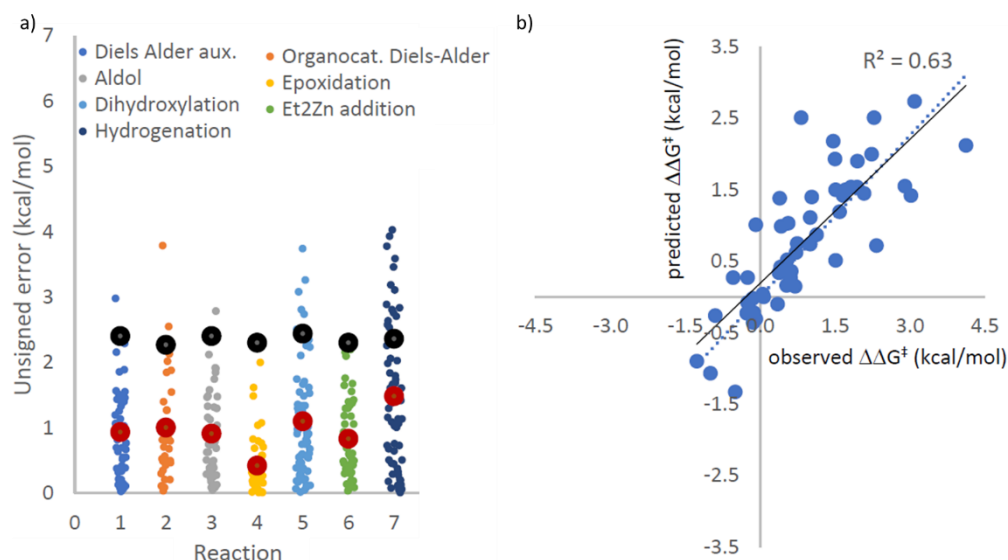


Figure 2.7. a) Unsigned error for $\Delta\Delta G^\ddagger$ (kcal/mol) between the predicted and experimental enantioselectivity for each reaction of the 7 reaction classes. Reaction colour correlated to the reaction class on top. Red dot is the average unsigned error. Black dot is the theoretical average unsigned error we would get if the predictions made were random; For this calculation random values between -4.12 to 4.12 kcal/mol were assigned for $\Delta\Delta G^\ddagger$. This represents maximum selectivity margins (ratio of 1000:1). **b)** A selected example of predicted vs. observed $\Delta\Delta G^\ddagger$ for the Shi epoxidation reaction, with 51 asymmetric catalyst/substrate pairs. Positive $\Delta\Delta G^\ddagger$ represents one enantiomer whereas negative $\Delta\Delta G^\ddagger$ represents the other enantiomer. Reproduced from reference Burai Patrascu *et al.*³²

The accuracy was evaluated by comparing the TS structure generated by ACE to previously reported TSs. Additionally, the predicted $\Delta\Delta G^\ddagger$ (energy difference between the diastereomeric TSs) computed from ACE was compared to the reported one. The unsigned error (difference between predicted and reported $\Delta\Delta G^\ddagger$) was also calculated and compared to the hypothetical unsigned error generated from random predictions to evaluate the true potential of this platform in catalyst design (Figure 2.7a).

The average unsigned error (over 5 runs) for all 7 reaction classes tested ranges between 0.94-0.97 kcal/mol. As an energy difference of ~ 1.4 kcal/mol corresponds to 90 %*er* (~ 82 %*ee*), an energy difference of 1 kcal/mol is considered the gold standard (also referred to as chemical accuracy). An unsigned average error of ~ 1 kcal/mol means that the platform can be applied to

differentiate between poor catalysts and good ones (*e.g.*, ~30 %*ee*, $\Delta\Delta G^\ddagger$ of ~0.4 kcal/mol *vs.* ~82 %*ee*, $\Delta\Delta G^\ddagger$ of ~1.4 kcal/mol) and differentiate between good catalysts to excellent ones (*e.g.*, ~82 %*ee*, $\Delta\Delta G^\ddagger$ of ~1.4 kcal/mol *vs.* ~97 %*ee*, $\Delta\Delta G^\ddagger$ of ~2.4 kcal/mol). While the predicted enantioselectivity was compared to published data, it is important to keep in mind that many of the published catalysts exhibit different enantioselectivities in different conditions (*e.g.*, solvents, temperature, additives such as acid co-catalyst, *etc.*) and are sometimes even reported with different enantioselectivities under the same conditions which are not directly accounted for in ACE. This experimental variability adds some apparent error. While ACE does have an implicit consideration of the solvent, and the temperature is considered in the Boltzmann population equation, changing the parameters did not improve the prediction accuracy.

A closer look at the false positives revealed that most of the inaccuracy in the prediction resulted from poor parametrization in the MM3 force field. A few examples for such compounds can be seen in **Figure 2.8**. For example, conjugated systems (**2.18**, **2.22**), sugar derivatives (**2.21**), sulfonamides (**2.20**), silyl ethers (**2.19**), and complex phosphines (**2.25**) are not well parametrized in MM3, leading to less accurate predictions.

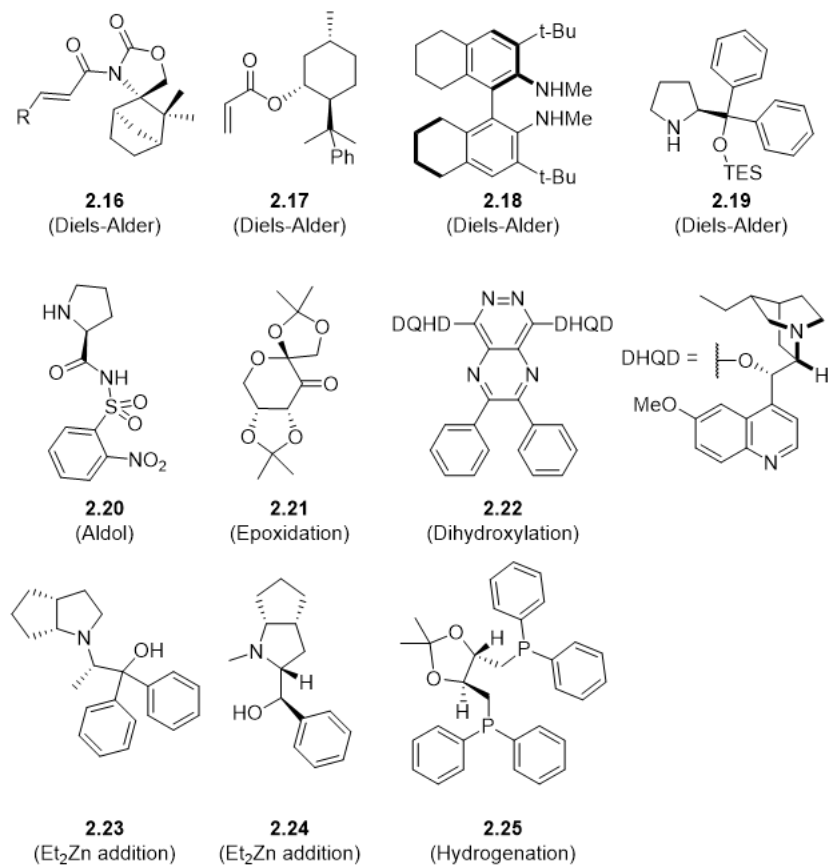


Figure 2.8. Examples of substrates (chiral auxiliary Diels-Alder, **2.16**, **2.17**) and catalysts that had $\Delta\Delta G^\ddagger$ errors of 2 kcal/mol or more. In parentheses the reaction class to which the molecule was tested for.

2.2.3.2. Real-life scenario II: Finding a new series of catalysts for a known reaction; virtual screening #4: ZINC database

This screening simulates a real-life scenario where a chemist might be interested in finding a new chemical series to catalyze a known reaction. We chose to use the Shi epoxidation and Diels-Alder cycloaddition for this scenario, since these are well-known reactions with published highly selective catalysts. The aim was to combine decoys with known catalysts into a library and evaluate whether we can recover the known catalysts within the top hit results. The purpose of this screen was to computationally verify the useability of the platform, as well as the accuracy and ability to

recognize not only selective catalysts, but also active ones. For the iminium catalysed Diels-Alder cycloaddition, ~140,000 chiral amines from the ZINC database were assembled into a library. The workflow described in **Figure 2.9** was designed; SMART was used to compute molecular descriptors for each of the molecules, such as molecular weight, charge, functional groups, etc. The molecules were filtered (REDUCE) using these descriptors, so that the library would only contain relevant molecules (*e.g.*, secondary amines, no other reactive group – aldehyde, double bond, carboxylic acid, uncharged compounds). FINDERS was then used to further specify the search, by substructure search, to only molecules containing secondary amines with a stereocenter adjacent to the amine nitrogen. With only molecules of interest in hand, SELECT was used twice. First, to remove any molecule too similar to existing, known catalysts to potentially identify new chemical series, leaving about 10,000 potential catalysts in the library. Then it was used to ensure diversity by removing analogues and keeping the most diverse structures. The purpose of this step is to optimize computing time. DIVERSE was next used to make sure no duplicates remained (*e.g.*, some libraries may include Pro·HCl and Pro·TFA), leaving 1,307 molecules for screening.

A second workflow was created to filter the 1,307 chiral secondary amines that remained in the library for their reactivity, and then selectivity. Since it was clear by experience that not all secondary amines will catalyze this reaction, QUEMIST was used to compute reactivity parameters (nucleophilicity). Using proline methyl-ester as the standard, molecules that were predicted to be less reactive (based on nucleophilicity indices) were removed using REDUCE. The 789 molecules left in the library were assembled into TSs with CONSTRUCTS, which were then input into ACE for enantioselectivity computations. In addition to the filtered molecules, the final library also

contained 6 known catalysts, which were added at this step (ACE), to assess ACE's ability to recover them. The virtual screening took place over 10 days, using a single core.

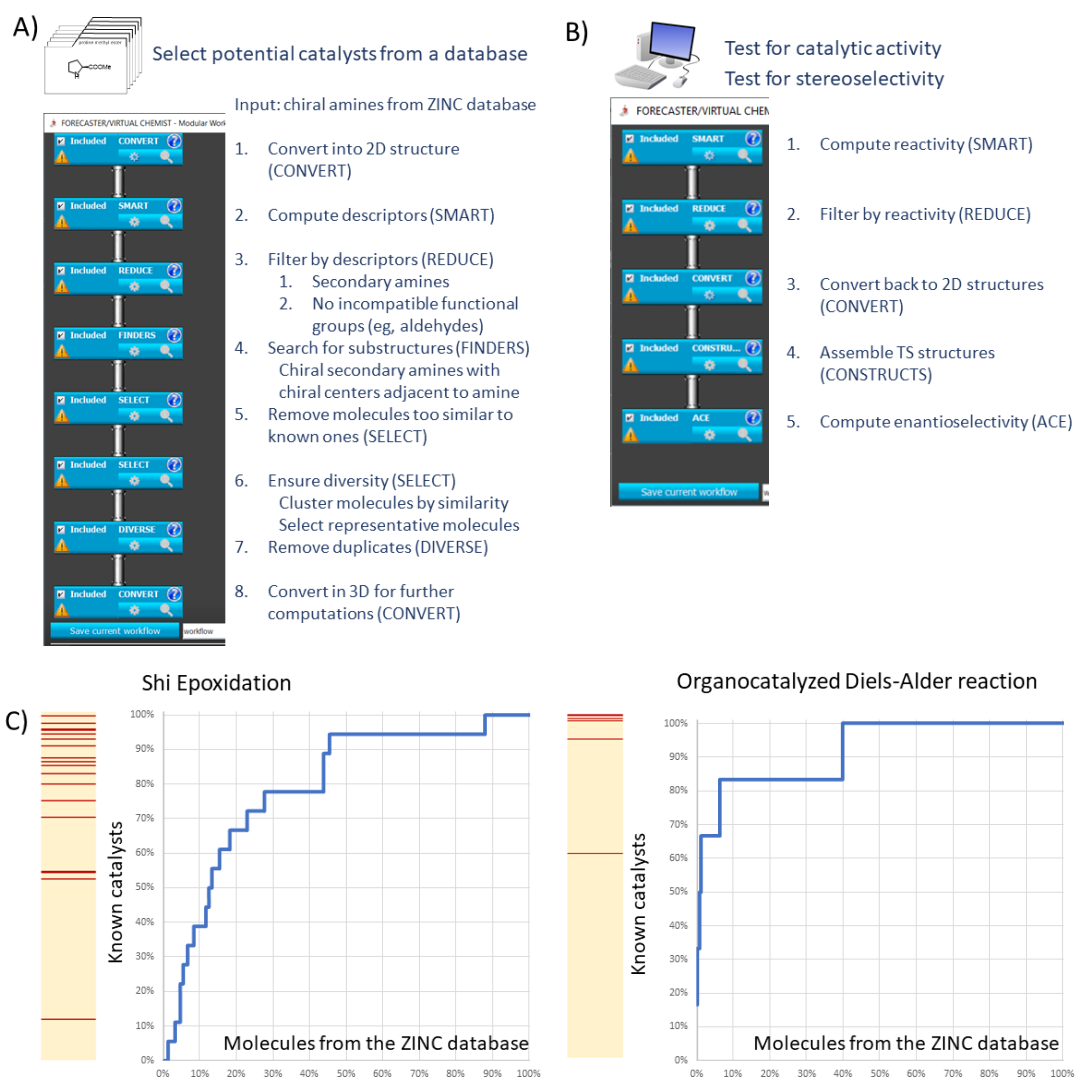


Figure 2.9. a) Workflow for the virtual screening (example with Diels-Alder cycloaddition) and reducing the library size into the most diverse molecules with description of actions on the right. b) Second workflow created to screen molecules for enantioselectivity, with description of actions on the right. c) Molecules screened ranked for enantioselectivity (Diels-Alder cycloaddition and Shi epoxidation). The red bars indicate the known catalysts that were added to the screening as decoys. The graph indicates the portion of known catalysts compared to the library from the ZINC database. The graph is created by ranking order, e.g., for the Diels-Alder cycloaddition ~7% of the molecules ranked at the top from the ZINC database, compared to ~85% of the known catalysts. That is, the known enantioselective catalysts were placed correctly at the top of the list.

For the Shi epoxidation a similar workflow was designed, this time screening a library of chiral ketones. Unlike chiral secondary amines, that are relatively abundant due to their more common use in synthesis (e.g., in medicinal chemistry), chiral ketones are less common (<300). To have a more sizable library to screen, we decided to use our in-house software REACT2D. A library of secondary alcohols that are adjacent to a chiral centre was created, and using REACT2D the alcohols were converted into chiral ketones (oxidation reaction). The converted ketones were added to the original library of chiral ketones, and the screening workflow was followed. Similar to the Diels-Alder cycloaddition, known stereoselective catalysts (18) were added.

The ability of ACE to recover the known catalysts within the top results of the ranked molecules was measured using the Area Under Receiver Operating Curve (AUROC) (**Figure 2.9c**). Most of the known selective catalysts have been ranked high, with an overall AUROC of 0.92 for the Diels-Alder cycloaddition, and 0.79 for the Shi epoxidation. This virtual screening validated the platform's capability for screening many molecules and discovering diverse chemical series as potential asymmetric catalysts. The general workflow used can be amended to guide the platform towards chemical series of interest (e.g., including a specific chemical group of interest or removing ones that will not be compatible).

2.2.3.3. Real-life scenario III: Designing and screening analogues for a hit with improved selectivity.

This scenario can be envisioned as the next step after Real-life scenario II. A chemist has a hit molecule following a screen and is interested in designing analogues and further optimizing it. To demonstrate this scenario, we used a detailed optimization study of an organocatalyst for the Diels-Alder cycloaddition by Gerosa *et al.*¹⁰¹ An initial promising core scaffold of a chiral

pyrrolidine derivative was identified, and further analogues were synthesized and tested for improved selectivity. The authors had to carry this exhaustive study out in the lab by synthesizing and testing each analogue, which clearly took a lot of effort and time. As our intention with VIRTUAL CHEMIST is to assist in these types of studies by doing a virtual screen, which is faster than the synthesis and experiment process, we simulated their workflow virtually (**Figure 2.10**).

The first step in the process was to create a virtual library. This was done by using FINDERS to create a library of amines with different R_1 groups, and a second library of aldehydes with different R_2 groups. REACT2D was then used to create the combinatorial library, producing all the products from an imine formation reaction. With the imine library, a 1,3-dipolar cycloaddition was carried out (virtually) using REACT2D again. A library of products was created for each possible diastereomer by running REACT2D as 3 parallel nodes. Next, the libraries of potential catalysts were combined, and the catalysts were assessed for their enantioselectivity in the *endo*- and *exo*-Diels-Alder cycloaddition. From the screening, the catalysts for both *endo*- and *exo*-cycloaddition were ranked and compared to the published experimental results. With a mean unsigned error of 0.33 kcal/mol, the catalyst predicted to be the most enantioselective were experimentally the best (for the *endo* product) and second best (for the *exo* product). We recognize that the experimental study probably took place over a long period of time, requiring product isolation and characterization, which requires a lot of synthetic effort. The virtual screen took place over a few days, using a standard Windows PC.

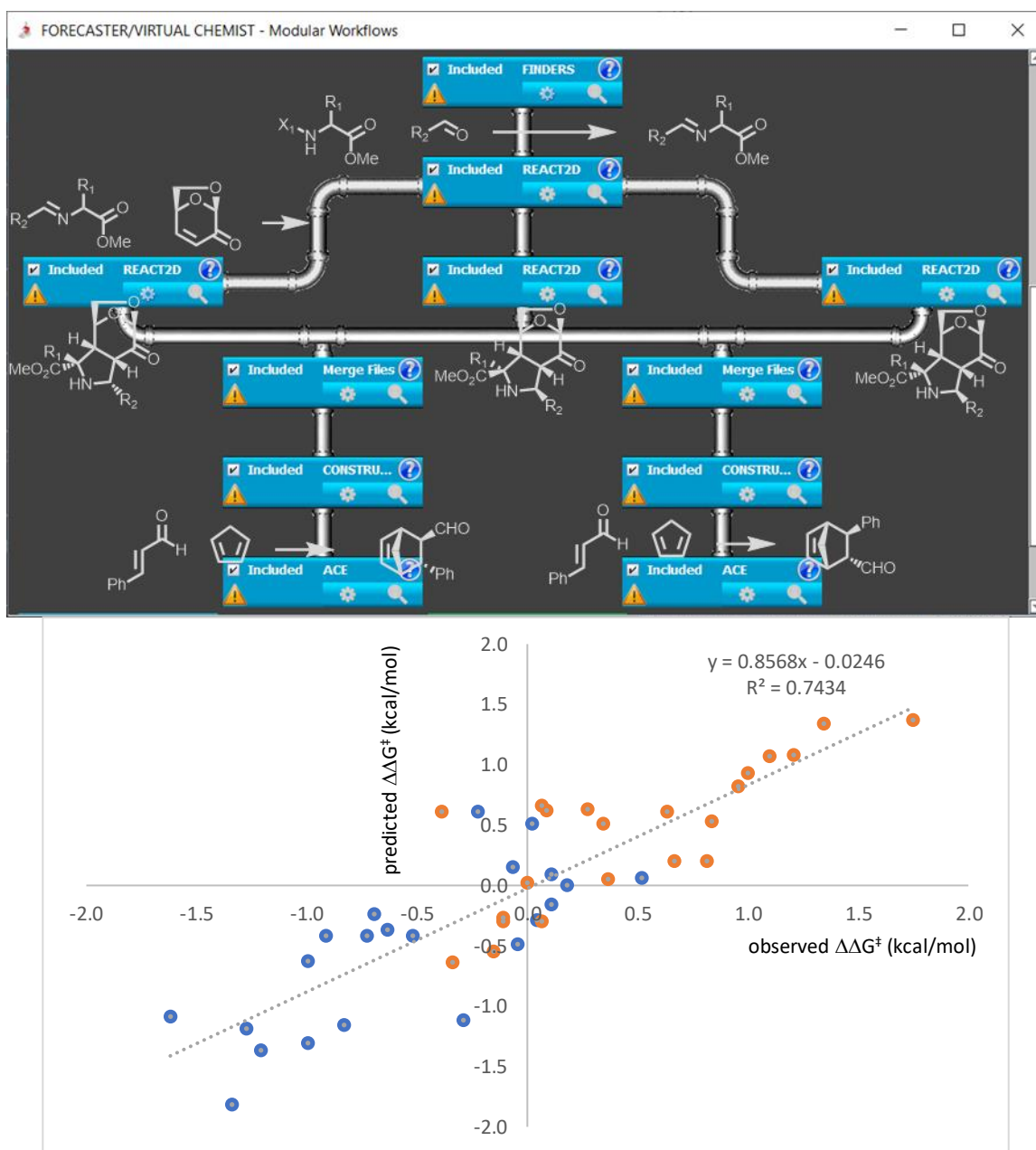


Figure 2.10. Top: Workflow for the virtual optimization of organocatalysts for the Diels-Alder cycloaddition. REACT2D was used to combine a library of primary amines (x represents a protecting group) with a library of aldehydes to create a library of imines, by imine formation reaction. Then a 1,3-dipolarcycloaddition was done using REACT2D to create a library for each isomer product in parallel. All three catalyst libraries were then evaluated for enantioselectivity in the Diels-Alder cycloaddition, *endo* and *exo* are run separately. Bottom: Results of the virtual screen plotted against experimental values (as $\Delta\Delta G^\ddagger$)- predicted on y-axes and observed on x-axes. Orange dots represent the *endo* adduct blue dots represent the *exo* adduct.

2.2.3.4. Real-life scenario IV: Evaluation of the substrate scope for a catalyst

When a new catalyst is developed, the next step is usually testing the catalyst's selectivity with a range of compounds. This measures the substrate scope of the catalyst. In this scenario, the substrate scope testing has been done virtually. For this computation we chose a well-known catalytic system for which a broad substrate scope has been published. We chose to use the OsO₄ dihydroxylation reaction using (DHQD)₂PHAL as the ligand. A virtual library of 25 substrates was created and subsequently tested. The computational results were compared to the published ones (Figure 2.11.).

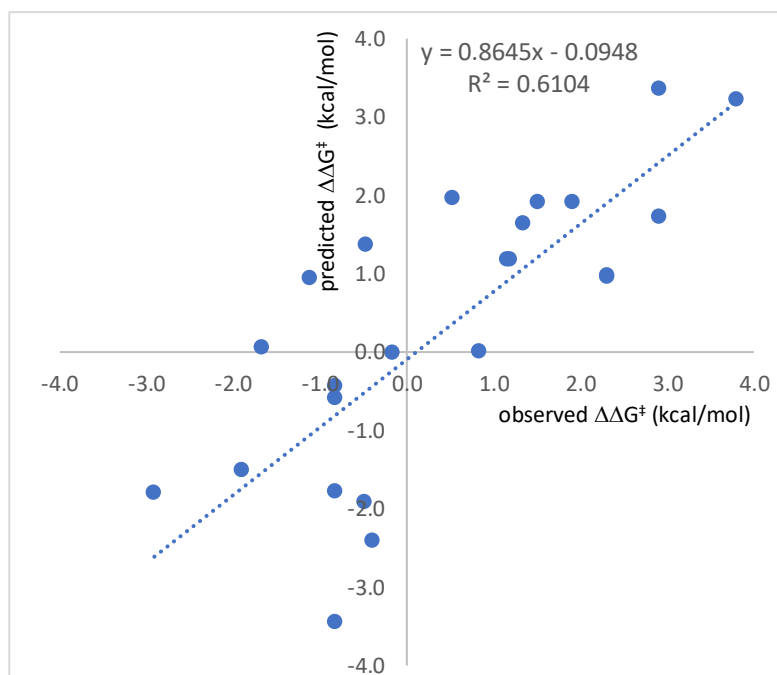


Figure 2.11. Results of the substrate scope study done on the Sharpless asymmetric dihydroxylation using (DHQD)₂PHAL as catalyst with different alkene substrates. Each substrate result is plotted as a blue dot, with predicted enantioselectivity (as $\Delta\Delta G^\ddagger$) on the y axis and observed experimental enantioselectivity on the x axis. Positive $\Delta\Delta G^\ddagger$ represents isomers (*R*) and (*R, R*), and negative $\Delta\Delta G^\ddagger$ represents the other isomers.

In agreement with experimental results, the catalyst was predicted to be highly enantioselective with ~25% of the substrates ($>97\%$ ee, > 2.5 kcal/mol), and poorly enantioselective with ~20% of the substrates ($<40\%$ ee, < 0.5 kcal/mol). 3 of the substrates were predicted to give the wrong isomer (top left quadrant), and overall, the dihydroxylation reaction had a large deviation between predicted and observed selectivity, compared to most other reactions tested, as was observed in **Figure 2.7a** as well. This is most likely due to the large size and high flexibility of the ligand, which may require more time and iterations during the conformational search. Additionally, a few of the poorer predictions result from limitations in the parametrization of the substrates (existing force fields might not be suitable for all the substrates), such as sulfur containing groups and *cis* alkenes.

2.3. Conclusions and future work

In this chapter, we worked on evaluating the useability of a software by organic chemists as well as its predictive power for different cases as examples. The initial work on the project highlighted a lot of limitations that exist. Some we cannot control, such as existing libraries of commercially available and/or synthesizable molecules, that can be used for screening as potential catalysts. Some we took on as a challenge to tackle, such as the addition of a reactivity index to the workflow, to minimize the chances of unreactive molecules in the final library. The work on developing some measurement of reactivity for amines brought to light our lack of detailed knowledge on what features are necessary for the amine to be a good catalyst (as an example here, secondary amine for the iminium catalyzed Diels-Alder cycloaddition, however the reactivity is relevant to all potential catalysts), and what step in the mechanism is hindering the formation of product – an unreactive iminium, a short lived iminium (gets hydrolyzed too quickly), or an amine that simply does not form an iminium.

As described in **Chapter 2**, we were able to validate VIRTUAL CHEMIST platform and apply it to different scenarios. However, when it came to experimental validation of the Diels-Alder cycloaddition, an issue remained that we could not be certain of what exactly influences the reactivity of the amines. Initially, we posited that it was related to poor nucleophilicity, then perhaps a combination of nucleophilicity and electron affinity. Eventually we concluded that the issue could be more complicated than it initially seemed, and we needed a deeper understanding of reactivity in this reaction. While some work has been published regarding the nuances that influence the reaction, some details of mechanistic intricacies are still missing. From here we decided to dedicate work for this specific issue and learn about the reactivity of different molecules, in this case secondary amines. This led us to the mechanistic study of the iminium catalyzed Diels-Alder cycloaddition described in **Chapter 3**.

This chapter is based on a manuscript in preparation: **S. Pinus**, J.Genzling, M. Burai-Patrascu and N. Moitessier, with additional work done by Christopher Hennecker.

All the chemistry detailed in this chapter was done by SP, the simulation was done by CH, initial computations done by MBP, follow-up computations done by JG.

Chapter 3

3.1. Introduction

3.1.1. Organocatalyzed Diels Alder cycloaddition.

Chiral secondary amines have been used as asymmetric organocatalysts in the Diels-Alder reaction for over 20 years.^{123, 132, 134-138} In their first report, the MacMillan group hypothesized that the reversible formation of the iminium with an α,β -unsaturated aldehyde could emulate Lewis acid (LA) activation by lowering the aldehyde's LUMO energy (**Figure 3.1**).¹²³ When a chiral amine is used, the iminium formed is also chiral, which leads to diastereomeric transition states. Upon completion of the cycloaddition, the iminium may be hydrolyzed to release the catalyst and the adduct. This general proposed mechanism for this iminium catalyzed cycloaddition is now widely accepted (**Figure 3.1**). However, improving the catalytic efficiency and stereoselectivity requires an in-depth understanding of this catalytic cycle.

We have recently reported VIRTUAL CHEMIST, a computational platform for asymmetric catalyst discovery which requires insights into the mechanism of a given reaction.³² Since the reactivity of the catalysts was a major issue during our previous attempts to use the platform for the design of novel catalysts, we first sought to investigate the mechanism of this transformation. Herein, we report our experimental and computational efforts to provide atomistic insight into the mechanism of this reaction, including the formation of the iminium species, in the context of previously reported data, to gain a better understanding of how the reaction mechanism proceeds and provide insights into potential optimizations.

3.1.2. Mechanism – what we know.

Following MacMillan's pioneering work, there have been a number of follow-up reports which led to a number of chiral amine catalysts with different selectivities and scope.¹³² In addition, mechanistic studies have been reported, which can aid in understanding some of the factors impacting the reaction rate and catalytic efficiency.^{139, 140}

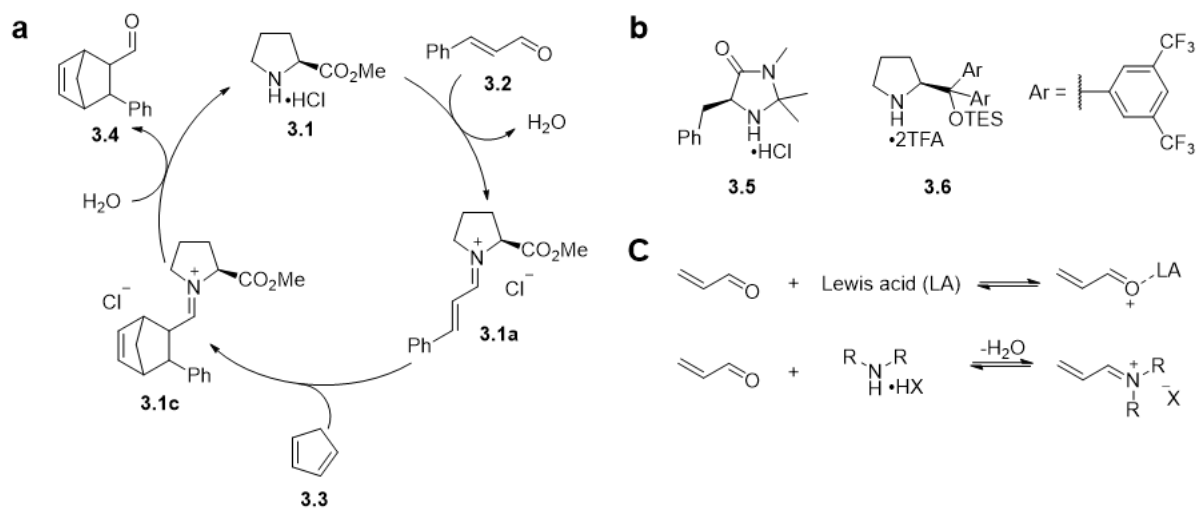


Figure 3.1. (a) Proposed mechanism for the proline methyl ester (**3.1**)-catalyzed Diels-Alder reaction. For simplicity, only one enantiomer of the iminium and the adduct is shown (stereochemistry of products not shown). (b) Known catalysts used in this reaction. MacMillan's imidazolinone based catalyst (**3.5**), and the pyrrolidine-based catalyst used by Hayashi (**3.6**). (c) LUMO lowering activation with a Lewis acid (top) or with a secondary amine salt via the formation of iminium, as proposed by MacMillan's group.

The most commonly used amine catalysts for the Diels-Alder cycloaddition reaction are based on 5-membered ring amines such as MacMillan's (**3.5**) and Jorgensen-Hayashi's (**3.6**) catalysts (**Figure 3.1b**), although primary amines have also been considered.^{132, 135} However, in theory, any secondary amine that would form the iminium could be used to catalyze the reaction. Increasing the reactivity of these amines by taking advantage of the α -effect was proposed by Tomkinson, and 6-membered, acyclic hydrazines were found to be more effective than 5-membered hydrazine and proline methyl ester.¹⁴¹ Tomkinson *et al.* rationalized these observations

by an increased acidity of the catalyst and lower LUMO of the iminium intermediate (**Figure 3.1c**).¹⁴² While working on an experimental validation of VIRTUAL CHEMIST for the asymmetric organocatalyzed Diels-Alder reaction, we performed a virtual screen on a library of chiral secondary amines following the well-documented stereochemistry control.¹⁴³ While 5-membered amines are usually preferred, two 6-membered ring amines were identified through this screening as potential stereoselective catalysts. We decided to synthesize and test them. However, no conversion was observed. Thus, we thought to go back to the mechanism to understand this established preference for 5-membered rings when amines are used, while also accounting for the work described so far in the literature. For example, in the original report by MacMillan and co-workers, it was found that, counter-intuitively, the presence of water increases the rate of iminium formation, likely because the water assist the proton transfer in the formation of the iminium.¹²³

Another report shows that less iminium is formed when water is used as a co-solvent when compared to anhydrous conditions.¹³⁹ The Tomkinson research group proposed that, in methanol, water assists in shifting the existing equilibrium towards more of the free aldehyde (vs. the dimethyl acetal), thus more aldehyde is present to form a reactive iminium, and, as mentioned above, that the proton transfer from catalyst to aldehyde is key to this process. In aprotic solvents such as acetonitrile, the water has an important role of dissolving the catalyst and activating the aldehyde through hydrogen bonding or protonation. The addition of water also led to increased enantioselectivity, compared to when no water was added, which has been attributed to fast hydrolysis of the iminium product, therefore decreasing the reversibility. This is in contrast to anhydrous conditions, where the iminium product allowed reversibility (thermodynamic control) which caused the enantioselectivity to erode over time.¹⁴⁴ From these studies, it is clear that water should be used as co-solvent when using methanol or acetonitrile as main solvent.

Work from the MacMillan group has also revealed that the iminium formed faster and to a greater extent when a more acidic co-catalyst was used (e.g., **3.5**·HCl vs. **3.5**·HClO₄). This can be explained by the first step of the proposed mechanism for iminium formation: the carbonyl needs to be activated by a Brønsted acid. Thus, a lower pK_a acid dissociates more, and more protons are available for carbonyl activation.¹³⁹

The immediate product of the cycloaddition is also an iminium. However, it has never been observed and is proposed to hydrolyze immediately into the product, releasing the catalyst. Thus, this step is not believed to be the rate-determining step (RDS). In the past, the iminium formation has been hypothesized to be the RDS and computations were carried out on simple model systems (e.g., dimethyl amine and acrolein) suggesting that the activation of the aldehyde is a key step.¹⁴⁵ However, these computations suggested that the formation of the iminium is disfavored by over 9 kcal/mol⁻¹ while we observe their formation (see below).

3.1.3. Mechanism – what we do not know.

Despite the aforementioned studies, it is still unclear what the impact of the amine reactivity on the catalytic cycle is. Although some reactivity studies were published, they were done by comparing known catalysts such as **3.5** and **3.6**, and the basic structure of the amine (e.g., ring size) was not tested.^{146, 147} It is also not clear what the relative role of the components of the salts is. For example, preliminary studies from the MacMillan group revealed that **3.5**·HClO₄ is the catalyzing species, as excess of free amine does not increase the reaction rate.¹³⁹ In contrast, when trifluoroacetic acid (TFA) is used as the co-acid, it is often used in two equivalents, although no reasoning was found for this change in conditions. In fact, excess acid may result in a racemic background reaction, although this was not discussed in these reports.^{124, 148}

3.2. Results and Discussion

3.2.1. Kinetic studies – iminium formation over time

3.2.1.1. Different solvent systems with proline methyl ester hydrochloride salt.

Aiming at providing a detailed mechanism for this reaction, we first performed kinetic studies on the formation of the iminium intermediates in different solvents and with different amines. The studies were done using $^1\text{H-NMR}$ experiments, in a deuterated solvent system (see SI for Chapter 3 for details), measurements were taken at specific time points (manually) or as an array NMR experiment (automatic measurement every set time). Measurements that were taken manually are shown as marker points with a line connecting them, whereas measurements that were taken automatically are shown as a line. Since the aim was to obtain a qualitative view of the reaction progression, the measurements were not done in duplicates. As shown in **Figure 3.2**, when cinnamaldehyde (**3.2**) is stirred in presence of 1 equivalent of methyl proline ester hydrochloride salt (**3.1**) in a mixture of methanol/water (19:1), the acetal **3.7** is quickly formed (maximum at the start of the recording), then both *cis* and *trans* iminium adducts are formed. Within approximately 50 minutes, a stable equilibrium is observed with a maximum iminium concentration of 61% (acetal concentration of 9%, cinnamaldehyde free 30%). When an acetonitrile/water (19:1) mixture was used, the reaction reached equilibrium sooner (approx. 30 minutes), however, the final concentration of the iminium was lower (35%). In acetonitrile, the acetal cannot form, and therefore the iminium formation is likely to be faster. However, since the formation of the iminium likely requires H-transfer, it may form to a lesser extent when an aprotic solvent such as acetonitrile is used (as the concentration of water is much lower).¹⁴⁸

When methanol was used as a single solvent, the initial concentration of the acetal was considerably higher (about 50%) compared to the methanol/water mix (initial acetal concentration of 20%). This observation is in agreement with the hypothesis that water is important for shifting the equilibrium from acetal back to free aldehyde. Interestingly, the final concentration of the iminium was only slightly higher (65%) than when water was added, suggesting that having about 5% water in the reaction does not reduce the final concentration of the iminium. With methanol as the only solvent, the time until equilibrium was reached was about double (110 minutes) than it was when water was added, which likely results from the higher acetal concentration and slower hydrolysis due to the lack of water.

Having that understanding of the impact different solvents have on the iminium formation with **3.1** as the catalyst, we chose to do all future experiments using the methanol/water (19:1) solvent system, as to maximize the iminium concentration we would observe, especially since many of the other amines used are likely to have much lower iminium concentration at equilibrium.

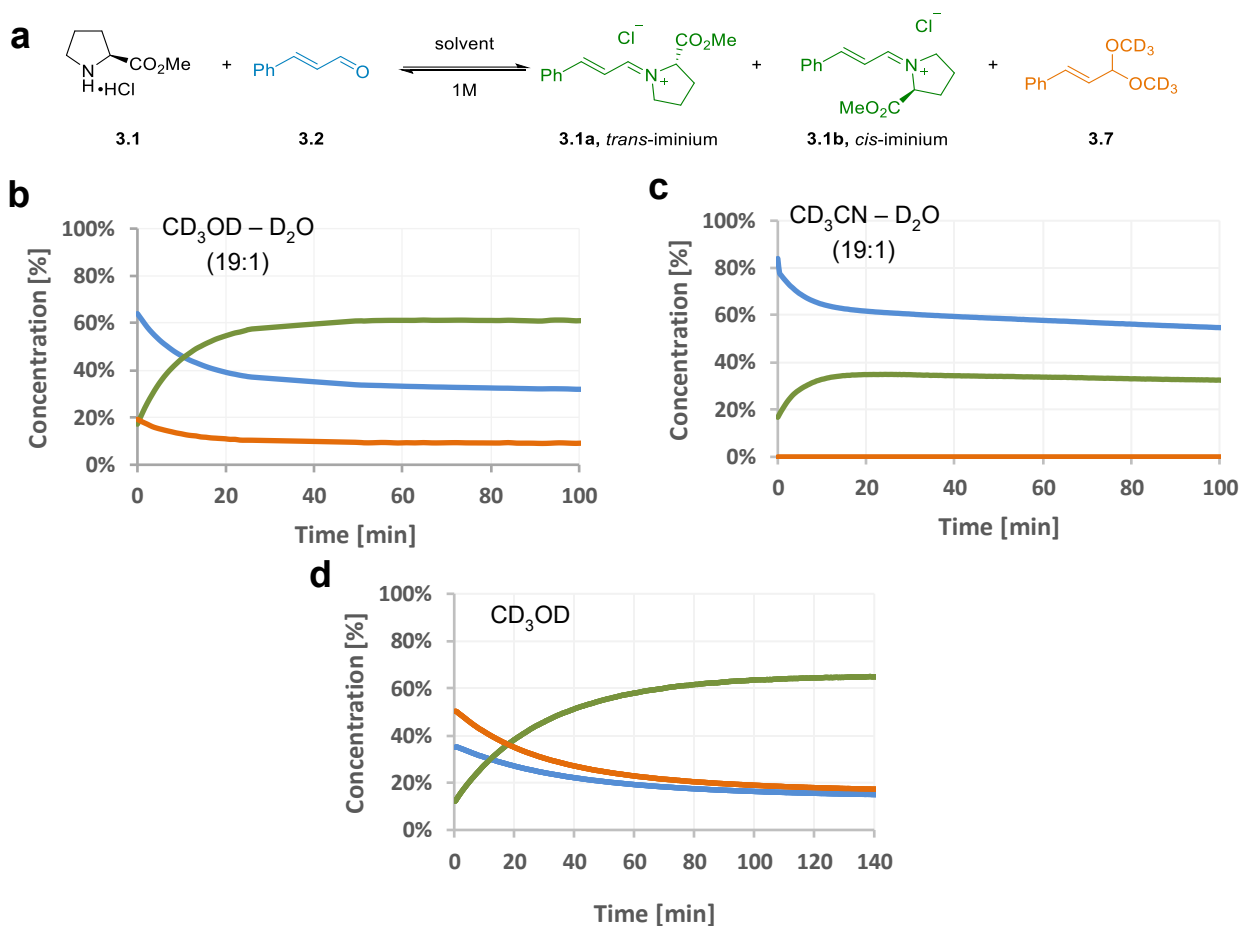


Figure 3.2. Iminium formation and equilibrium over time from proline methyl ester hydrochloride (**3.1**) and cinnamaldehyde (**3.2**) in three different solvent systems. **(a)** General reaction scheme. Both the amine salt (**3.1**) and aldehyde are at 1 equiv. in a 1 M solution. The colors match the plot's color over time. **(b)** Total iminium -**3.1a** and **3.1b** (*trans* and *cis*, respectively) and acetal (**3.7**)– formation and equilibrium over time in CD₃OD-D₂O (19:1). Each line color corresponds to the color in **(a)**: **3.2** is blue, **3.1a** and **3.1b** are summed in the green line, **3.7** is orange. **(c)** Total iminium -**3.1a** and **3.1b** (*trans* and *cis*, respectively) and acetal (**3.7**)– formation and equilibrium over time in CD₃CN-D₂O (19:1). **(d)** Total iminium -**3.1a** and **3.1b** (*trans* and *cis*, respectively) and acetal (**3.7**)– formation and equilibrium over time in CD₃OD.

3.2.1.2. Different amine hydrochloride salts.

With the optimal conditions set, we were next interested in examining the behavior of a range of different amines and getting an estimate of equilibrium concentrations and timescale to reach equilibrium. MacMillan's catalyst (**Figure 3.3, 3.5**) was used with the same conditions (methanol/water solution, 19:1). The reaction reached equilibrium within 60 minutes, with iminium concentration of 48% (acetal concentration of 12%, free cinnamaldehyde 40%).

The same experiment was done with sarcosine methyl ester hydrochloride (**Figure 3.4, 3.9**), the reaction reached equilibrium within 27 minutes, with iminium concentration of about 30% (acetal concentration of 16%, free cinnamaldehyde 56%: since the concentration is normalized based on the first NMR spectrum, an error may occur during integration since the peaks tend to broaden – leading sometimes to over slightly 100% concentration).

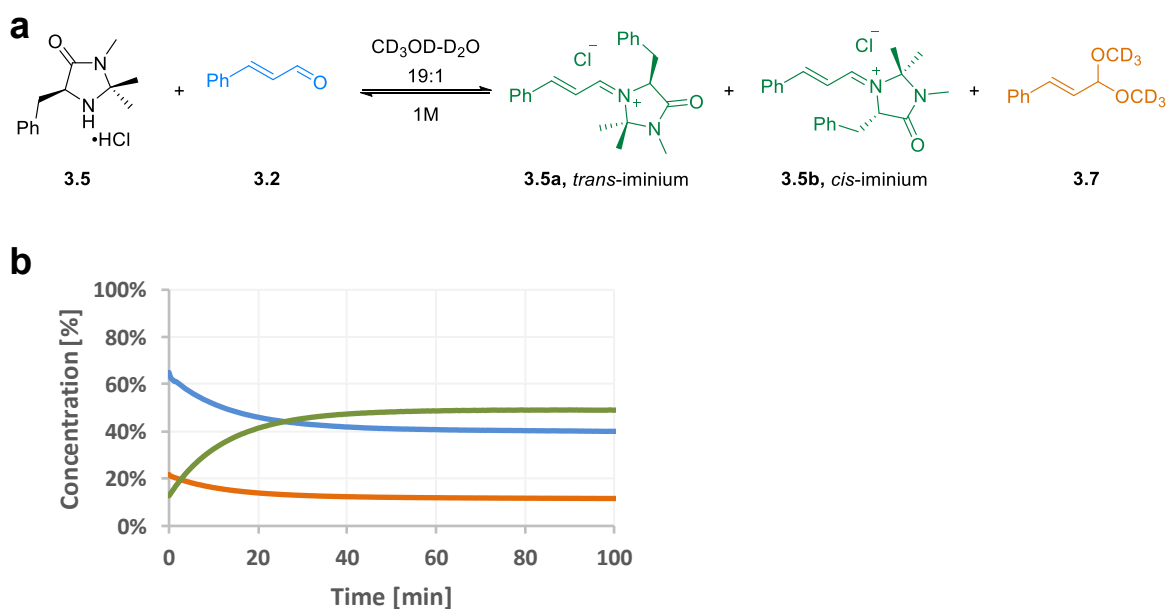


Figure 3.3. (a) Iminium formation equilibrium from MacMillan's imidazolinone-based catalyst (**3.5**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). (b) Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme.

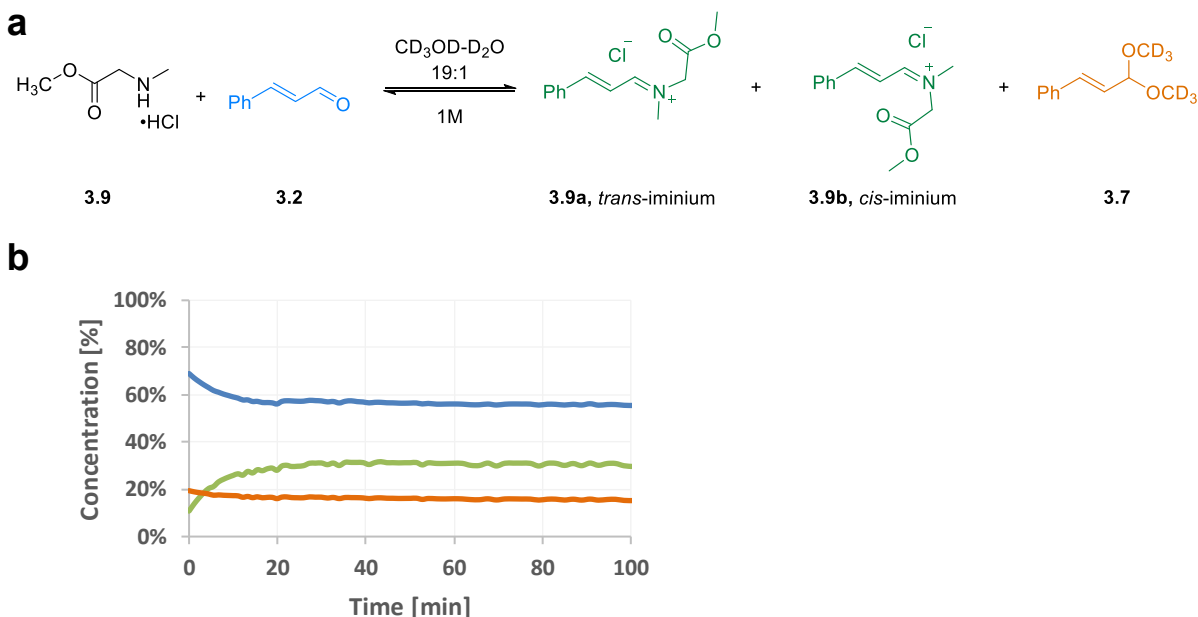


Figure 3.4. (a) Iminium formation equilibrium from sarcosine methyl ester HCl salt (**3.9**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of $\text{CD}_3\text{OD-D}_2\text{O}$ (19:1). (b) Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme.

While working on the iminium concentration at 30 minutes experiments (will be described later), we observed that the NMR tubes containing amines **3.10** and **3.11** that did not have a significant conversion into iminium after 30 minutes, turned bright yellow after 24-48 hours (for structures see **Figure 3.5**). Although the amines tested above for equilibrium have reached equilibrium within 1 hour or less, it seemed that some of the less reactive amines needed much longer time to reach equilibrium. We wondered whether the low conversions resulted from kinetic or thermodynamic effects and investigated the iminium formation over longer periods of time. To this end, we extended the duration of the experiments with these amines (**Figure 3.5**). Indeed, amine **3.10** seemed to not reach equilibrium even after 24 hours (**Figure 3.5b**, green plot), in the last measurement the iminium concentration was 39%. Pyrrolidine HCl (**3.11**) salt also appeared to react very slowly and was still increasing at the last measurement after 12 hours, with iminium concentration of 46% (**Figure 3.5d**). It should be noted that the plot for piperidine HCl (**3.10**) is

not smooth because the measurements were taken manually every few hours over 24 hours, which results in less data and therefore a “sharp” looking plot. For pyrrolidine HCl (**3.11**) an array experiment was used, where the NMR automatically runs a proton NMR every 24 seconds, which results in a smoother plot, however the trade-off is the maximum use time being 12 hours. Since we were looking at the general trend, the information for these two amines is enough to understand that the iminium formation rate may have a much bigger variability, depending on the amine structure, than we had expected before running these experiments.

We were interested in running the same experiment with amine **3.12**, however due to poor solubility at 1 equiv., we had to dilute it 10-fold (0.1 M solution). The results are described in the experimental section. Since we were hoping to have more data points for the beginning of the reaction, timepoints 0-2 hours were measured using an array experiment. Due to the time limitation on the NMR spectrometer, the rest of the measurements were done on a different spectrometer (both spectrometers are 500 MHz) with the same settings. Despite the same parameters, the graph clearly has a bump in the integration when the switch in the machines was done. Overall, the iminium formed is less than 1% after 30 hours, which confirmed this amine is not efficient in the formation of the iminium. These results suggest that the iminium formation might be a very slow process for some amine salts, rather than a thermodynamically disfavored equilibrium for these amines.

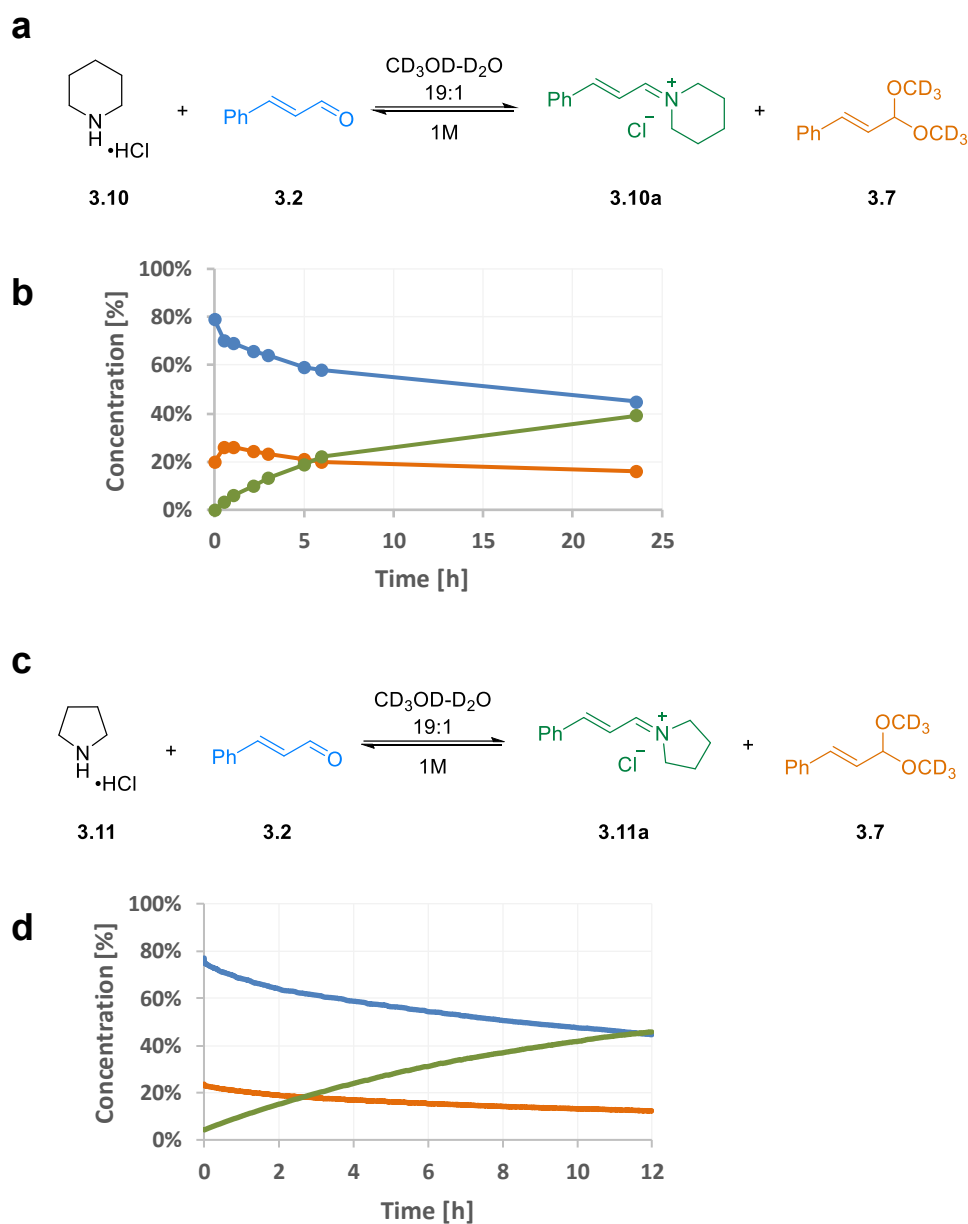


Figure 3.5. (a) Iminium formation equilibrium from piperidine HCl salt (**3.10**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). (b) Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme. (c) Iminium formation equilibrium from pyrrolidine HCl salt (**3.11**) and cinnamaldehyde (**3.2**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1). (d) Iminium formation reaction progression over time, until equilibrium is reached. The plot colors correspond to the colors in the scheme.

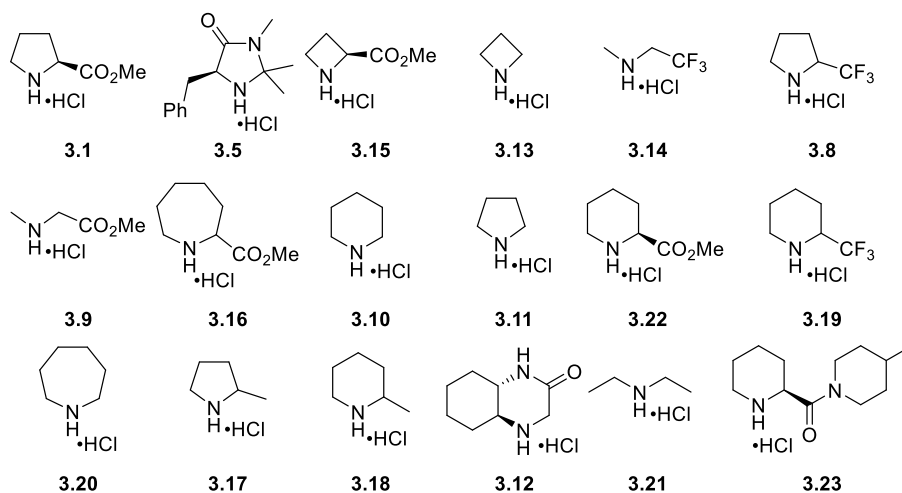
3.2.2. Iminium concentration after 30 minutes

With this information, we decided to screen more secondary amines by measuring the conversion after 30 minutes (**Figure 3.6**). We chose to monitor the conversion at conditions that more closely resemble reaction conditions, and so the amine salts were used in 0.2 equiv. to the 1 equiv. of the aldehyde, in 1 M solution. Although some amines have been observed to take longer to reach equilibrium, we were hoping to get an estimate at 30 minutes, since the cyclization reaction is expected to run over the course of a few hours to 24 hours, which means the iminium will most likely never reach equilibrium before reacting. The amount of iminium observed for some amines is drastically lower than for the reference amines investigated above. For instance, for amines such as pipercolic methyl ester (6-membered ring, **3.22**) and diethylamine (**3.21**), even when 100 mol% were used, 7% and 0% converted into iminium in this time frame, respectively.

This study (**Figure 3.6**) suggested that there is a possible correlation between the acidity of the ammonium salts used as catalysts and the formation of iminium salts, in line with the correlation observed with the pKa of the co-catalyst.¹³⁹ The data indicate that for most general scaffolds, more electron-deficient amines form iminium in larger amounts than electron-rich amines. Until now, we assumed equilibrium was reached within 45 – 60 minutes for all the amines and confirmed this behavior for the ammonium salts producing the largest concentration of iminium salts.

Overall, no amine exceeded 60% conversion into iminium, which corresponds with the kinetic studies, where we observed the equilibrium reaching a plateau with some free amine salt remaining. Additionally, it appears that overall, 5-membered ring amines are the most reactive ones at forming iminiums, although this only applies to those containing electron withdrawing

groups (**3.1**, **3.5**, **3.8**). Electron rich pyrrolidine (**3.17**), or unsubstituted (**3.11**) were considerably less reactive at the time frame of the experiment. As suspected, all the piperidine based amines performed poorly within this time frame, which suggests they would not be good scaffolds for catalyst candidates, as even the ones containing electron withdrawing substituents gave less than 5% conversion (**3.12**, **3.19**, **3.22**, **3.23**). Interestingly, azetidines (4-membered ring amines) based amine salt appeared to be equally reactive, either as the unsubstituted amine **3.13**, with 53% conversion of the salt into the iminium, or the methyl ester substituted amine **3.15**, with 56% conversion. We hypothesized that the reactivity we are observing in the formation of the iminium is likely a factor of the ability of the nitrogen to give up the proton (as when the nitrogen is protonated, it will not react as a nucleophile), which relates to the acidity and presence of electron withdrawing groups. It also relates to the structure of the amine (scaffold), that is, while azetidine **3.13** may not have an electron withdrawing group, the ring strain may result in a more active nitrogen, which would result in a faster reaction with the aldehyde.



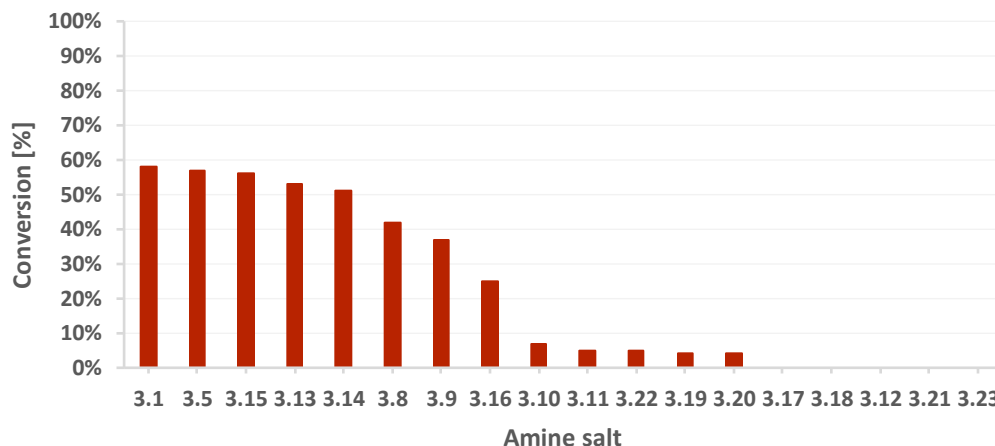


Figure 3.6. Structures of amines used in the study (top), and results of conversion of the amine salts into iminium after 30 minutes (bottom graph). The amine salts were used in 20 mol%, in a 1 M solution of CD₃OD-D₂O (19:1). Conversion into iminium is calculated based on integration: $(H_{\text{iminium}}/(H_{\text{iminium}}+H_{\text{amine}})) \times 100$.

3.2.3. Kinetic studies – Diels Alder cycloaddition

3.2.3.1. Product formation with different amine salts.

At this stage, we were interested in learning about how the formation of the iminium (rate + concentration) affected the Diels-Alder cycloaddition reaction. It had been previously proposed by Tomkinson and coworkers that the cycloaddition is the rate limiting step.¹⁴⁹ However, we were curious to see what the time dependent curve would look like for reactive and also less reactive amines. To do so, we ran the complete Diels-Alder reaction for 50-60 hours, and measured at different time points the conversion (by ¹H-NMR) of the different reaction components over time (**Figure 3.7**). Although cyclopentadiene is usually used in 3 equiv., we wanted to avoid having high excess of cyclopentadiene to not overpower all the other peaks in the NMR spectra and have therefore decided on using 1.1 equivalents. We also used 1 equiv. of the amine salts again, for the same reason explained for the iminium formation studies. Knowing some of the peaks might be very small and difficult to find in very busy spectra, we wanted to make sure the iminium would

be detectable. We then proceeded to analyze the NMR spectra by integrating the iminium formed (both *cis* and *trans*-iminium, when relevant), the cinnamaldehyde, as an indicator of the consumption of the substrate, and all the products formed integrated and combined (both *exo* and *endo*- aldehydes, as well as the corresponding acetals), since we are interested in the general progression of the reaction and wanted to simplify the results graph. We then analyzed any correlation between the rate of the iminium formation and the product formed. The results are shown in **Figure 3.7**, for three different amines (HCl salts): proline methyl ester (**3.1**), pyrrolidine (**3.11**), and 2-(trifluoromethyl)pyrrolidine (**3.8**). We expected to observe a different rate not only in the formation of the iminium (which was already monitored), but also in the cycloaddition and formation of product. Indeed, 2-(trifluoromethyl)pyrrolidine (**3.8**) appears to form product considerably faster than the other two (**Figure 3.7c**) with 51% product formed within the first 5 hours. In the same time frame, the proline methyl ester (**3.1**) and pyrrolidine (**3.11**) had 11% and 0% product formed, respectively. The pyrrolidine was considerably slower overall, as expected, and after 50 hours only 13% of the product was formed.

The final concentration measured for both (trifluoromethyl)pyrrolidine (**3.8**) and the proline methyl ester (**3.1**) at 72 hours was 88% and 78%, respectively. At this time, we had assumed the reaction reached a plateau, since the same concentration was observed at 47 hours. Interestingly, the reaction never reached completion. This may be attributed to the lower concentration of cyclopentadiene (**3.3**) that was used in these experiments, as it dimerizes over time. The concentration of cyclopentadiene was not added to the plot. Since it is highly hydrophobic and fairly volatile (bp = 40°C), it was difficult to monitor what amount was present in solution. The observations in these experiments are in agreement with the hypothesis that a more electron

deficient amine (such as **3.8**) will form an iminium with lower LUMO energy, which would therefore react faster with a nucleophile (in this case, the cyclopentadiene).

Another interesting observation is that for all three reactions, it appears that about 10% iminium is formed before the cycloaddition reaction starts taking place. This suggests that some minimal level of iminium is necessary for the reaction to occur. As even the more reactive iminiums (**3.1**, **3.8**) have this accumulation of iminium prior to reaction, it suggests that indeed the cycloaddition reaction is the rate limiting step. However, since some amount of iminium is necessary for the cycloaddition to even take place, the rate of iminium formation will also affect the overall rate and turnover of the reaction, and reactions with amines that form iminiums considerably slower (such as **3.11**, and some of the amines presented above) will also be limited by the competition of the cyclopentadiene dimerization (deactivation).

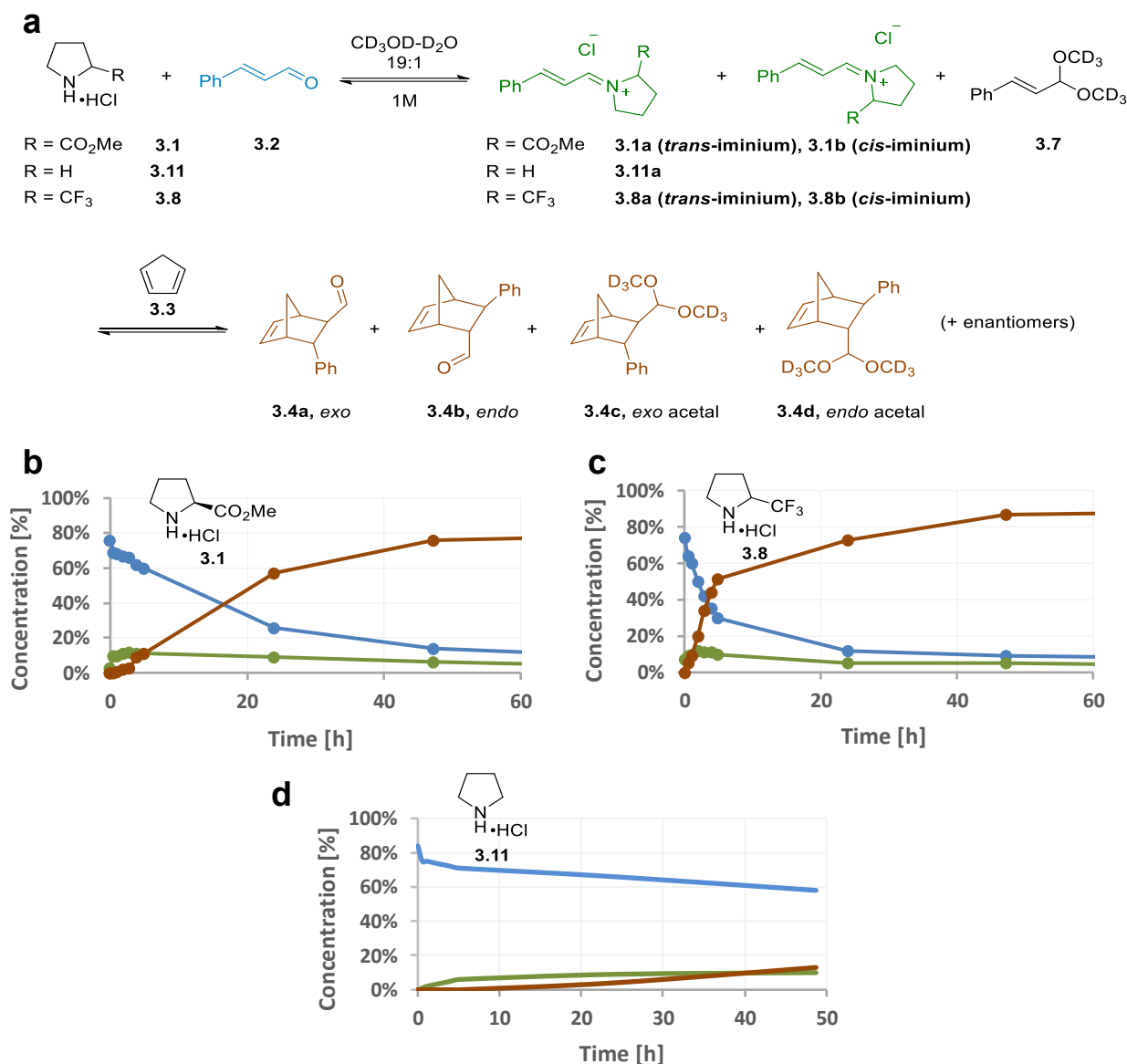


Figure 3.7. (a) General reaction scheme for the iminium catalyzed Diels-Alder cycloaddition with different amine HCl salts (**3.1**, **3.11**, **3.8**), cinnamaldehyde (**3.2**) and cyclopentadiene (**3.3**). The amine and aldehyde are both 1 equiv. in a 1 M solution of CD₃OD-D₂O (19:1), cyclopentadiene is used in 1 equiv. (b) Reaction progression over time with HCl salt of proline methyl ester (**3.1**). The plot colors correspond to the colors in the scheme, brown represents total product formed. (c) Reaction progression over time with HCl salt of 2-(trifluoromethyl)pyrrolidine (**3.8**). The plot colors correspond to the colors in the scheme, brown represents total product formed. (d) Reaction progression over time with HCl salt of pyrrolidine (**3.11**). The plot colors correspond to the colors in the scheme, brown represents total product formed.

3.2.4. Product formation simulation: effect of rate of each step.

Using the data from the iminium formation study (**Figure 3.2**) and Diels-Alder cycloaddition study with proline methyl ester HCl (**3.1**), we were able to simulate the reaction progress over time (**Figure 3.8**). Two scenarios were considered: **(i)** product formation over time with the first step (k_1 , iminium formation) being x2 faster than its observed rate in the study, and **(ii)** product formation over time with the second step (k_2 , cycloaddition) being x2 faster than its observed rate in the study. In both scenarios, only the rate of one step is changed, the other step remains unchanged to the one observed in the experimental study. As a note, the data observed is qualitative, and is used to visualize the influence of each step in the reaction on the total rate of product formation. The two scenarios are plotted in **Figure 3.8b**, showing how both steps influence the total reaction rate (product formation). The rate of the second step has a greater effect, and is therefore the rate determining step, as previously hypothesized, but the first step has a significant impact and cannot be disregarded.

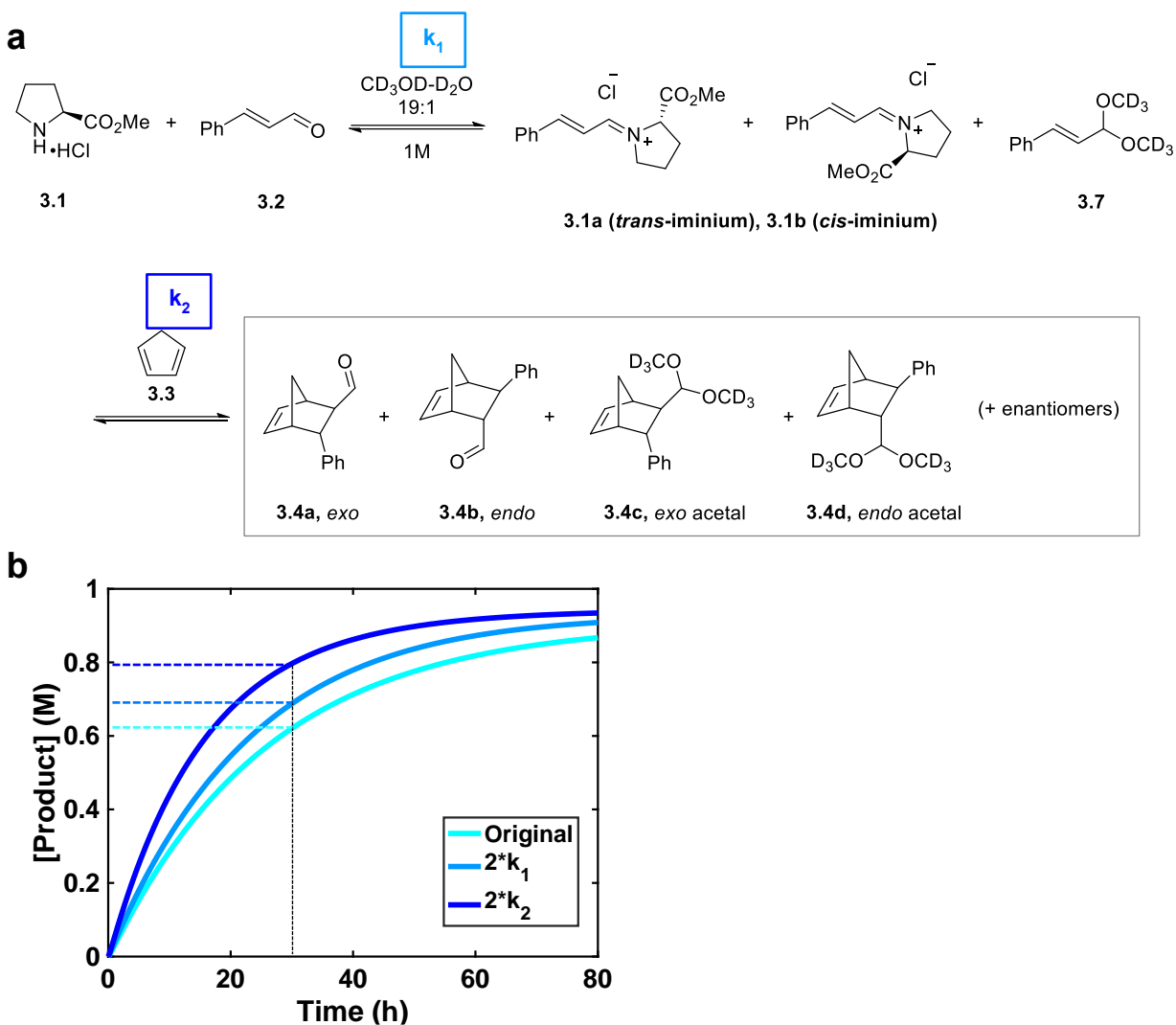


Figure 3.8. Simulated time course concentration for product formation over time of the full cycloaddition reaction with proline methyl ester HCl (**3.1**), based on the data collected in the kinetic studies of iminium formation and full reaction and the best fit parameters (Chapter 3-SI **Figure S3.13**, **Table S3.4**). **(a)** Reaction scheme, first step is the iminium formation and marked in pale blue (k_1), second step is the cycloaddition reaction of the iminium formed with the cyclopentadiene and is marked in dark blue (k_2). The total product is highlighted in a gray box. **(b)** Results of the simulation plotted as total product concentration over time, plot line colors correspond to the different steps as described in **(a)**, with the original rates being marked in cyan. The lines added at ~30 hours are meant to help visualize the difference in product formation, depending on the rate of each step. The difference should be compared to the original line. Increasing the rate of the second reaction (k_2) has a larger effect on the overall rate of the reaction, suggesting that the second step is the rate limiting step.

3.2.5. Detailed mechanism

The simulation based on our data is in agreement with previously published computations using **3.6** as the model system, in which the rate determining step is the cycloaddition step.¹⁴⁵ However, it also sheds light on the effect the iminium formation has on the reaction rate, as well as the amine structure's influence on it. Clearly, a minimal amount of iminium needs to be formed before the cycloaddition takes place. For some less reactive amines (e.g., **3.11**) this step takes place over a considerable amount of time, which does not work well for the time scale of the reaction, as the cyclopentadiene dimerizes in addition to the desire of a catalyst to increase the reaction's rate. The result is that less reactive amines will not provide product within the reaction's time frame, not because of the rate determining step being too slow or the iminium not being reactive, but because the iminium formation is kinetically too slow for the catalytic cycle to complete. This data also led us to propose a detailed mechanism (**Figure 3.9**) in which the rate determining step is the cycloaddition step. The reaction begins with the activation of the aldehyde (**3.2a**). This initial activation is directly proportional to the amount of free acid released by the ammonium salt used as organocatalyst (**3.1**→**3.1c**).

Next, the iminium forms through a hemiaminal where the amine is still protonated (**3.24**). Then, we hypothesized two options, based on conclusions from the computational study (detailed below). One of the pathways may be, where there is a localized extra equivalent of acid, leading to **3.27**, where both the nitrogen and the hydroxy are protonated. In the other pathway, first the nitrogen gets deprotonated forming **3.28**. Then, either the nitrogen of **3.27** gets deprotonated forming **3.25_7**, or the hydroxy of **3.28** gets protonated forming **3.25_6/8**. It should be noted, that both **3.25_7** and **3.25_6/8** are the same species, however, they have been labeled differently for distinction, to be able to easily follow the protons as well as the computational study. Next, the

iminium forms by the release of water (**3.1a**), and about 10% iminium is accumulated at which point the cycloaddition reaction takes place. The cycloaddition adduct (**3.26**) is likely hydrolyzed promptly, as it has not been observed, releasing the catalyst (**3.1c**, **3.1**) and product aldehyde (**3.4a** and **3.4b**), which rapidly forms the corresponding acetal (**3.4c** and **3.4d**, the acetal may also form directly from the iminium).

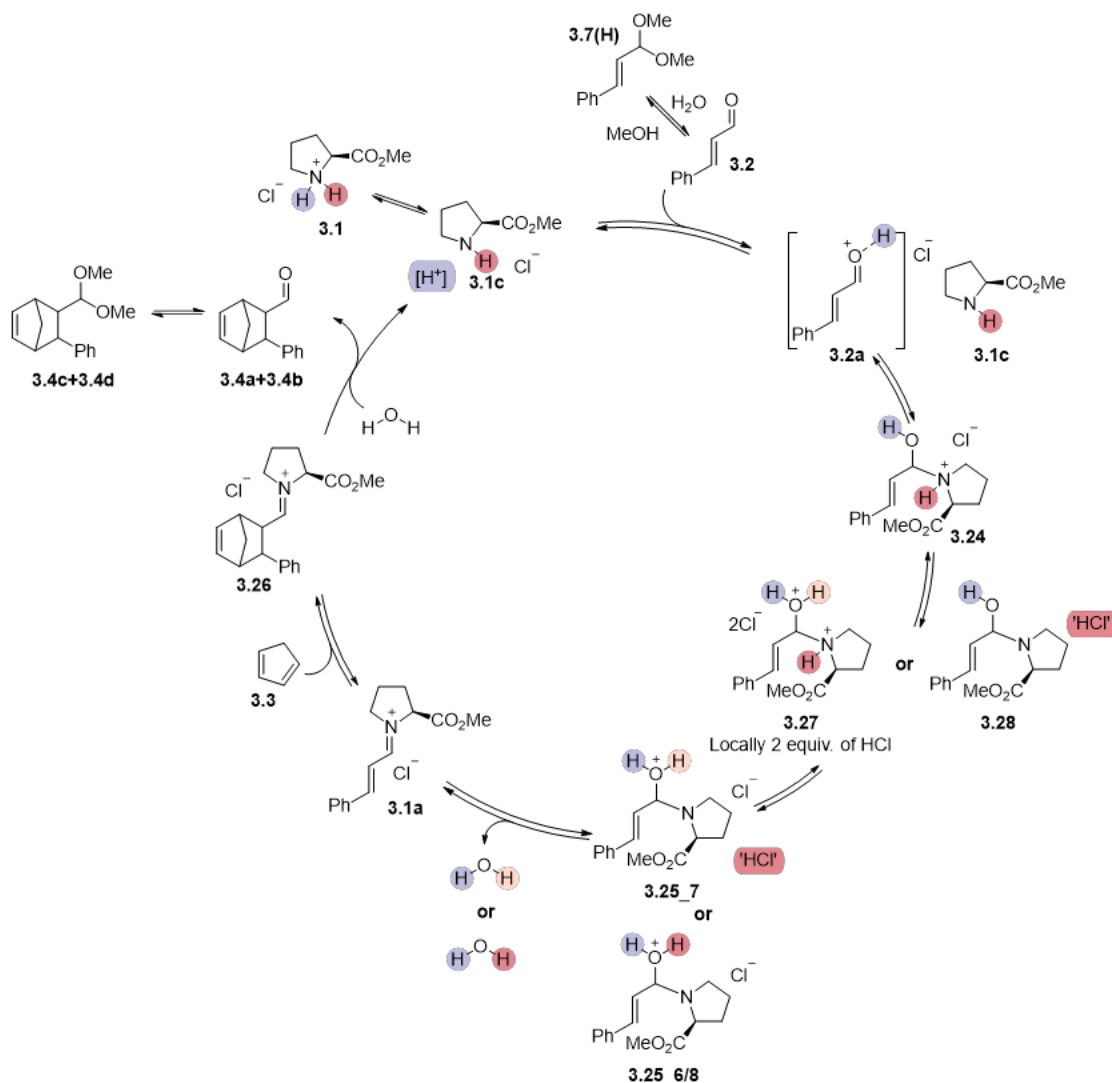


Figure 3.9. Diels Alder reaction catalyzed by proline methyl ester (**3.1**): detailed mechanism. The proposed mechanism is described while keeping the HCl and amine ratio as 1:1, based on the conditions that were used in our experiments and experiments discussed in the introduction. It is possible that locally 2 equivalents of acid are necessary for the iminium formation step (structure **3.27**).

3.2.6. Computational studies

Our goal was to determine the specific steps for the formation of the iminium species, as described in **Figure 3.9** (**3.24** \rightarrow **3.1**). We started the computations with a model system (proline methyl ester hydrochloride, **3.1**), and once we find an energetically favourable pathway, we may consider adding computations to other amine salts. We hope that computing the pathway to less reactive amines (such as piperidine-based) may shed more light on the possible reasons for reactivity differences. To analyze different mechanistic pathways, the two bonds involved in the proton transfer (**Figure 3.10a**) were scanned (geometry optimization). The bond cleavage and protonation/deprotonation were facilitated by the addition of a water molecule network to the system (**Figure 3.10b**).

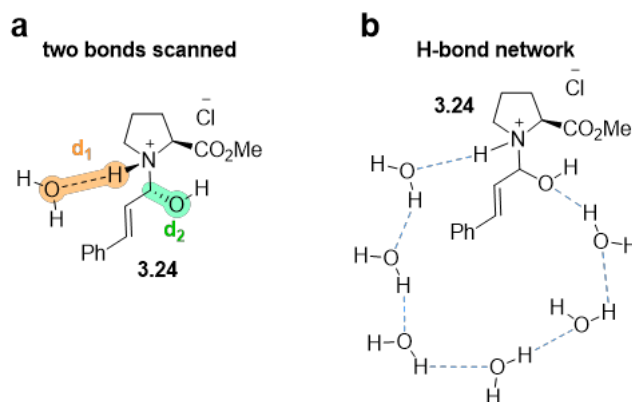


Figure 3.10. (a) The two breaking bonds that were scanned. d₁ (orange) the NH--OH₂ bond (which reflects the N-H bond breaking), and d₂ (green) the C-O bond. (b) The model system included a network of 6 water molecules all interconnected via H-bonds to facilitate protonation/deprotonation.

3.2.6.1. Two different hypotheses for the deprotonation of the amine and protonation of the hydroxy groups

Hypothesis I: the protonation of the hydroxy and deprotonation of the nitrogen may not be a concerted process (for a concerted process see **3.24** \rightarrow **3.25**, **Figure 3.9**), and may be influenced

by **1**) the placement of the counter anion (chloride in this case), which can stabilize the proton transfer, and **2**) positive charge formed on the hydroxy.

To study this idea, three systems were built and studied. For clarity, hereafter we will consider the ^+N-H to be on the top face and the $^+OH_2$ on the bottom face of the system (**Figure 3.11**). The three systems, visualized in **Figure 3.11**, are:

System 6: the chloride counter anion located on the top face by the protonated nitrogen.

System 7: local 2 equiv. of HCl, one chloride by the hydroxy, one by the protonated nitrogen, and the second proton on a water molecule (as H_3O^+) also placed initially in proximity of the hydroxy and additional chloride.

System 8: the same as system 6, only the chloride counter anion is placed on the bottom face, by the hydroxy group.

All systems have six water molecules around them in an H-bond network. The number was chosen as it was the minimal number of water molecules necessary to allow an H-bond network connecting both faces of the molecule (the ^+NH and OH of the hemiaminal).

Hypothesis II: regarding **system 7**; since the system has an additional equiv. of acid, there could be two possible pathways: **Path a**) Similarly to the two other systems, the nitrogen gets deprotonated to produce a neutral species (**3.28_7**, **Figure 3.11**), or **Path b**) where the hydroxy gets protonated by the excess acid prior to deprotonation of the nitrogen (**3.27**, **Figure 3.11**).

3.2.6.2. Initial plan

The three systems in **Figure 3.11** were studied in the following order (excluding system 7 path b, which will be discussed further below), with all the initial structures optimized with no constraints first with the same level of theory (wB97x-D3/def2-TZVP level of theory):

- 1) Breaking of the N-H (d_1) bond to form a H_3O^+ molecule – with a proximal water molecule, via the H-bond network.
- 2) Proton transfer from the top of the molecule (the nitrogen) to the bottom, for the protonation of the hydroxy.
- 3) Release of the water molecule (breaking of the C-O bond, d_2) and formation of the iminium to map the Potential Energy Surface (PES) of the transformation from the N-protonated hemiaminal to the iminium as a 3D map (each d_1 distance has a full scan of d_2)

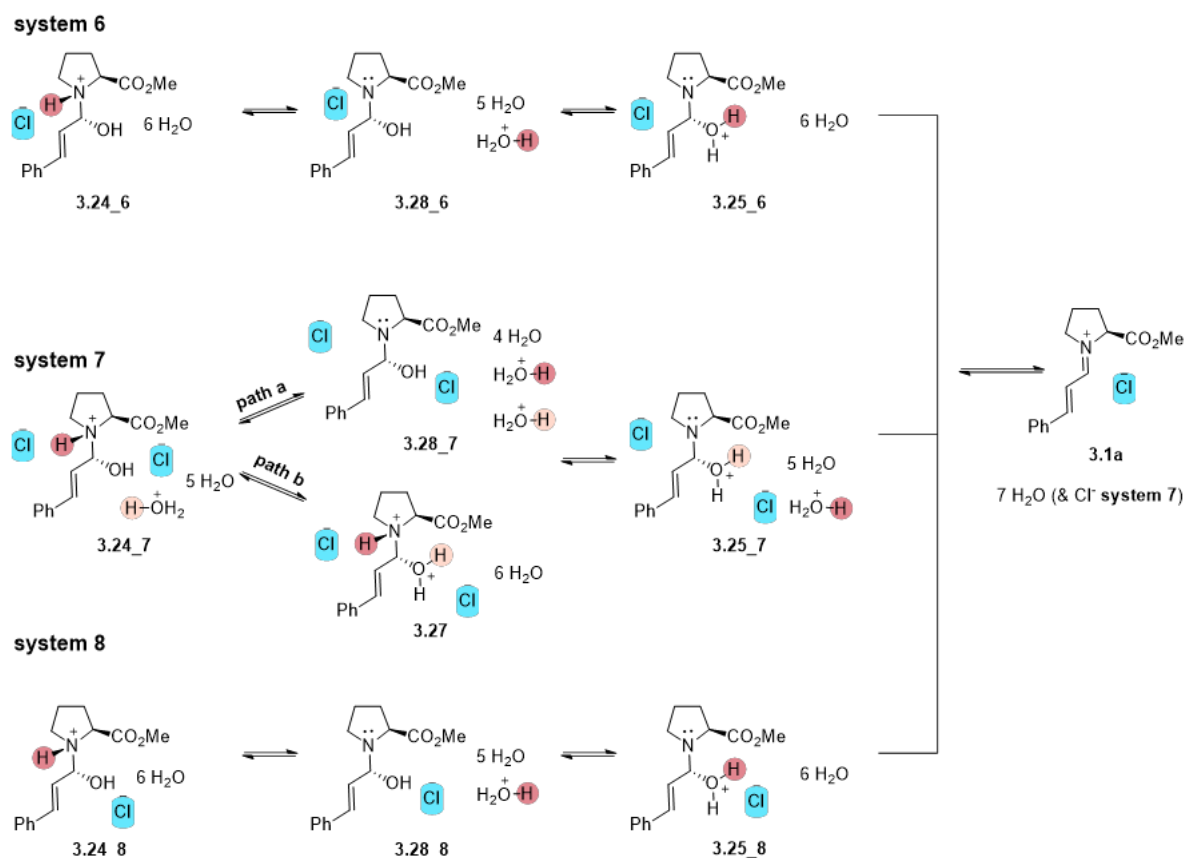


Figure 3.11. The three systems used to test our hypothesis and study the stepwise formation of the iminium: **System 6**: the chloride counter anion located on the top face by the protonated nitrogen. **System 7**: local 2 equiv. of HCl, one chloride by the hydroxy, one by the protonated nitrogen, second proton on a water molecule (H_3O^+) by the hydroxy. **System 8**: the same as system 6, but the chloride counter anion is placed on the bottom face, by the hydroxy group. All systems have 6 water molecules around in a H-bond network. The proton originating from the amine is highlighted in dark pink, the proton originating from the second acid equiv. (only in system 7) is highlighted in pale pink. Chlorides are highlighted in blue. All intermediates are labeled based on the relevant system.

3.2.6.3. Proton transfer step and re-evaluation of initial plan

Once the initial geometry optimization and d_1 scan is completed, general structure **3.28** (**Figure 3.11**) is formed. Mechanistically, our expectation is that the next step, during the later stages of d_1 scan would be the protonation of the hydroxy to form general structure **3.25**. We were optimistic that with the water molecule network around, this proton transfer from the deprotonated

nitrogen (H_3O^+) would happen independently during optimization. Unfortunately, this did not occur. One reason for this may be that DFT is a static process, consequently geometry optimization does not capture a dynamic process such as a proton transfer across an H-bond network. For this reason, we needed to find different options to model this step.

We came up with three possible options:

- I. Proceed to work on system 7, where the additional proton in the system is already placed by the hydroxy (**Figure 3.11**) and placing it in closer proximity to the hydroxy, (followed by geometry optimization prior to the second scan). The drawback of this approach is that experimental observations mentioned in section **3.1.3**, showed that the amine and acid are needed in a 1:1 ratio (though these studies were done on MacMillan's catalyst, and based on the results, potentially more experiments are needed in our case). However, local concentrations in solution are not static.
- II. Model the proton transfer between the individual water molecules with geometries scanned, to test whether the system's energy varies significantly during this process. This would be the same scan as we do for d_1 and d_2 , for each water molecule. If the energy remains essentially constant, we can conclude that the proton transfer is likely to happen spontaneously in solvent.
- III. "Forcing" the proton transfer by removing a proton and placing it on the hydroxy group. The geometry would be optimized prior to proceeding with the second scan (d_2). This possibility poses a few issues. We could not remove the proton from the H_3O^+ formed during the deprotonation of the nitrogen, since it is still on the nitrogen in the early steps of scan d_1 , which means we either had to **1**) move it for later stages of scan d_1 once the nitrogen has already been deprotonated (**Figure 3.12a**), or **2**) remove a

different proton. The issue with **1**) is that the 3D map may end up having variations due to reasons not related to the chemistry, but because the entire system's energy may be different within the scan (as the proton was moved for some and not for others). We decided to proceed with **2**) and remove a proton from one of the water molecules around (**Figure 3.12b**). The issue with this option is that it forms HO⁻ in the system, which should be neutralized throughout the scan, but is not a realistic representation of the system. Understanding none of these options is ideal, we are hopeful to review the results and asses from them how to better perform the scans.

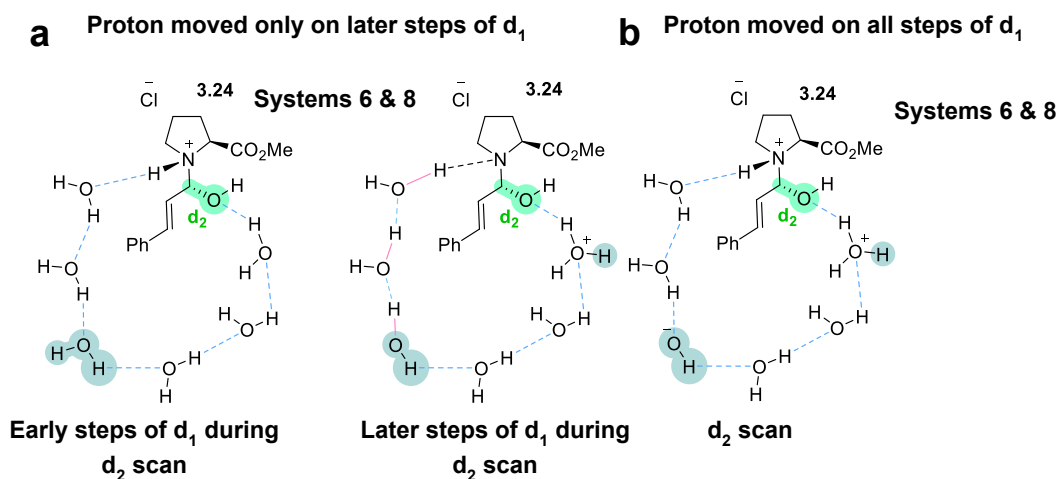


Figure 3.12. Proton transfer option **III** for systems 6 & 8. **(a) III.1:** left: early steps of d_1 (N still protonated), prior to scan d_2 . Right: later steps of d_1 (N deprotonated) prior to d_2 scan. Proton transfer neutralizes the hydroxide formed (highlighted in teal). H-transfer bonds labeled in pink. **(b) III.2:** proton is removed from water network for all d_1 steps prior to d_2 scan, forming a hydroxide and a hydronium by the hydroxy (highlighted in teal).

We proceeded with the second scan (d_2) as follows: **System 6** and **system 8** with possibility **III**, and **system 7** with possibility **I**. Since possibility **II** is the most time consuming, we decided to examine the results of the other options first.

3.2.6.4. Results so far

The computational study is still underway, herein we describe the results as we had them at the time of writing this thesis.

3.2.6.4.1. System 7

By moving the proton closer to the hydroxy in **system 7**, paths **a** and **b** in fact converge. The computation was done in the following order for each scan:

Path a. **1.** Additional HCl equiv. was placed on the bottom face, near the hydroxy, as H_3O^+ (protonated one of the existing water molecules in the network) and Cl^- , and the initial geometry was optimized. **2.** Scan d_1 was completed. **3.** The additional proton was moved closer to the hydroxy, and the geometry was optimized (for each step of d_1). (**Figure 3.13**, left) **4.** Scan d_2 .

Path b. **1.** Additional HCl equiv. was placed on the bottom face, near the hydroxy, as H_3O^+ (protonated one of the existing water molecules in the network) and Cl^- , and the initial geometry was optimized **2.** The additional proton was moved on to the hydroxy and the geometry was optimized (**Figure 3.13**, right). **3.** Scan d_1 was completed. **4.** Scan d_2 .

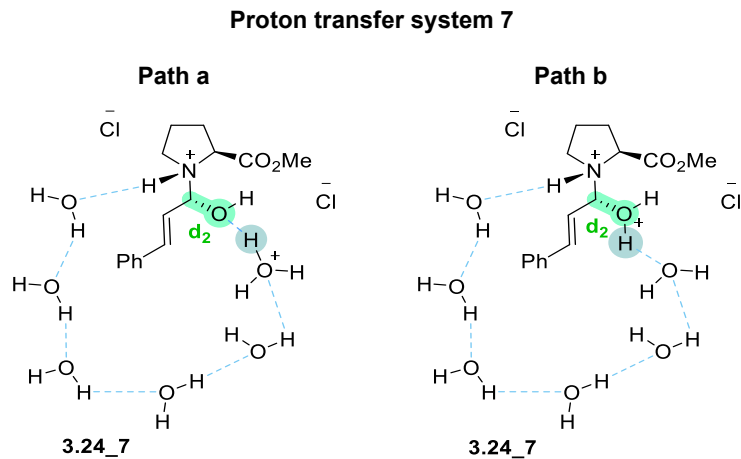


Figure 3.13. Proton transfer step for system 7 prior to scan d_2 , after scan d_1 for **path a**- proton placed closed to the hydroxy (left, proton highlighted in teal) and prior to scan d_1 , for **path b**- proton placed on the hydroxy (right, proton highlighted in teal).

Upon completion of step 3 for both paths, we noticed that the energy plots for both are very similar, as well as the geometry (**Figure 3.14**). Essentially, whether the protonation of the hydroxy/ placement of proton closer to it happens as step 3 or 2, results in the same geometry. At the end of step 3 for both paths, the additional proton is placed right in between the hydroxy and the water molecule right next to it. This positioning of the proton would theoretically allow for a quick proton transfer during scan d_2 . Since the starting point for scan d_2 is the same for both paths, we will only proceed with one of them, and once the 3D plot is complete, we should be able to see at which point of the mechanism the hydroxy is protonated (before or after the nitrogen gets deprotonated).

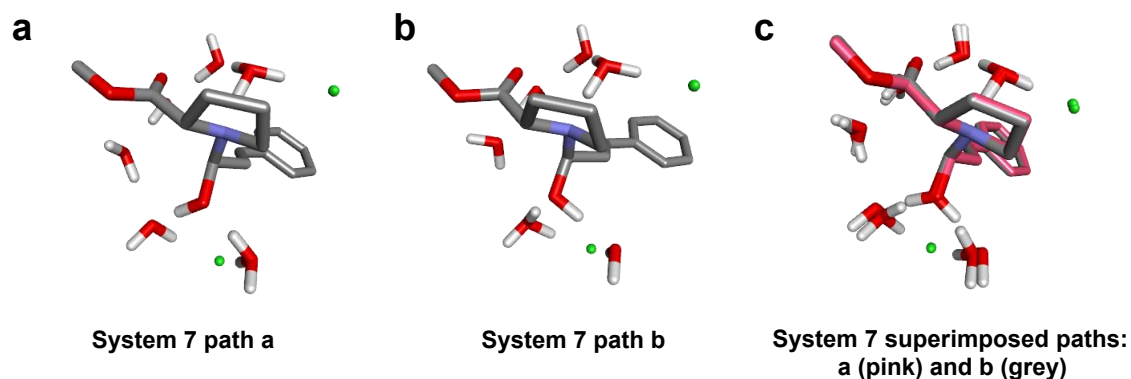


Figure 3.14. Geometry optimized structures of system 7 after step 3: **(a)** path a. **(b)** path b. **(c)** Both paths superimposed, path **(a)** (pink carbons), path **(b)** (grey carbons). The systems are essentially identical. They only differ on the orientation of the hydroxy, and which water molecule is protonated.

3.2.6.4.2. Systems 6 and 8

As discussed above, once the first scan (d_1) was completed, the proton was manually transferred (as depicted in **Figure 3.12**) on a water molecule in the proximity of the hydroxy, followed by geometry optimization. The goal of this calculation is to rearrange the hydrogens on the water network. We expect to see the proton from the ^+NH to be between the nitrogen and the closest water molecule on the top face (depending on the step of d_1 scan). For the later steps of scan d_1 we expect the artificially formed ^-OH to be neutralized, and for the hydroxy at the bottom face be “half protonated” as the proton is between the hydroxy and the closest water molecule.

The second scan (d_2) for all systems is still ongoing as of the writing of this thesis.

3.3. Conclusions and future work

One of the biggest hurdles we encountered in Chapter 2, was the lack of reactivity of the amines tested. For this reason, in this chapter we embarked on a deeper investigation of the amine structure influence on the iminium formation and the subsequent cycloaddition reaction. Our investigations revealed that while this reaction is viewed as a base-catalyzed reaction, the activated iminium intermediates formation step is acid-catalyzed. As acid-base reactions are fast, the key factor is the concentration of acid available for activation, which must reach a critical amount to enable a fast process. Amine salts that lack the reactivity to form enough iminium for the cycloaddition reaction to take place, will not produce product. The more acidic amine salts will not only form the iminium faster, but will also form a more reactive iminium, and the cycloaddition rate will also increase. 6-membered ring amine salts appear to have the lowest reactivity in the formation of iminium, regardless of their acidity. This fact might be attributed to the structure and conformation of these amines.

Although we had a general idea of reactivity based on previous publications and our observations, we did not know how big of a difference there can be between the different amines, and how much it would affect the general outcome of the reaction. It is important to keep in mind that the results are semi quantitative, since the experiments have some intrinsic errors (no replicates, a gap between mixing of the different reaction components and taking the first NMR, and little control over moisture related water in the reaction) and are meant to provide qualitative data on the different steps of the reaction progression. Since the iminium formed is conjugated and absorbs within the UV-Vis range, a different way to measure the progression of the first step could be by measuring the change in absorbance. There are a few factors that we did not yet address in

this work: the acid co-catalyst remained constant, since we were first interested only in the effect of the amine structure influence on the reaction rate. However, it has been shown that different acids influence the rate (e.g., as mentioned in the introduction, perchloric acid usually increases the rate), since the amine is added as a salt, the effect may be attributed to the counter ion present in agreement with our computational study suggesting a role for the counter anion. Additionally, it has been shown with MacMillan's imidazolinone catalyst that the acid and amine are required in 1:1 ratio, that is, adding more acid does not change the product formation rate, and adding less acid will lower the rate as if the extra free amine was not present. For this reason, we chose not to investigate the amine: acid ratio for the different amines we tested and used them in a 1:1 ratio for all the experiments. Nonetheless, it will be worth exploring whether this is true for the other amines as well. In fact, when Hayashi and co-workers used the proline-based catalyst, they used the amine and acid in a 1:2 ratio. However, this reaction was carried out under different conditions (toluene and TFA), and so the extra acid may have been necessary to assist in proton transfer, similarly to the role of the protic solvent in our studies, Furthermore, TFA is not entirely dissociated in toluene, which may necessitate the excess.

When we started the computational studies, we did not expect the small details of the reactions to complicate the investigation as much as they did. One of the major challenges was how to model the proton transfer from the protonated amine (of the hemiaminal) to the water and then to the hydroxy (of the hemiaminal), so that the iminium can be formed with water as a leaving group. This hurdle may be a limitation of the method, since in practice, we expect the proton transfer to happen spontaneously, though in a computation this might not be the case, and hence the model needs to be adjusted. Although we have yet to complete the computational study, we are optimistic of its results shedding light on the reactivity trends we observed during experiments.

Chapter 4 Preface

An important detail about carrying out published reactions for the purpose of validation is the ability to reproduce published data. In **Chapters 2** and **3** we ensured that we could reproduce the results of the Diels-Alder cycloaddition using proline methyl ester as the catalyst. In similar fashion, in **Chapter 4** we were interested in using the Shi Epoxidation as another validation tool for the VIRTUAL CHEMIST platform (as described in Chapter 2) and encountered the first problem of inability to reproduce the published procedure. **Chapter 4** is therefore dedicated to a few objectives: 1. Establishing a working, reproducible procedure. 2. Virtual screening of a library of chiral ketones. 3. Experimental tests on the reactivity of simple ketone scaffolds. This work will set the stage to virtual screening for the discovery of novel epoxidation catalysts.

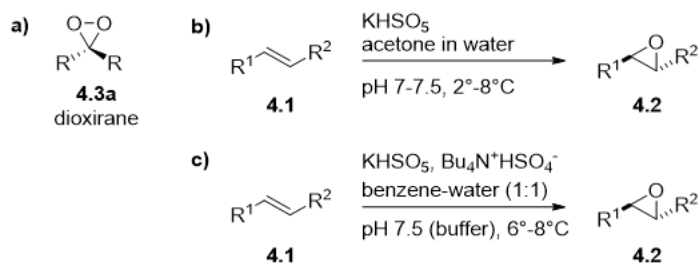
All the chemistry detailed in this chapter was done by SP, virtual screening was done in collaboration with Dr. Mihai Burai-Pătrașcu (former graduate student).

Chapter 4

4.1. Introduction

4.1.1. Alkene epoxidation with Oxone and the formation of dioxirane.

When R. E. Montgomery reported the decomposition of Oxone (potassium peroxymonosulfate, therein referred to as caroate) in the presence of ketones, it did not take long for the potential of this pathway to be further investigated.¹⁵⁰ ¹⁸O-labeling provided evidence that the Oxone likely reacts with the ketone to form a dioxirane (**4.3a**, **Scheme 4.1.a**). Curci *et al.*¹⁵¹ saw an advantage in the electrophilic nature of the dioxirane formed and hypothesized that it could be reacted with slightly nucleophilic alkenes, similarly to mCPBPA's reaction with alkenes. In their work, they used acetone as the ketone source (in catalytic amount).



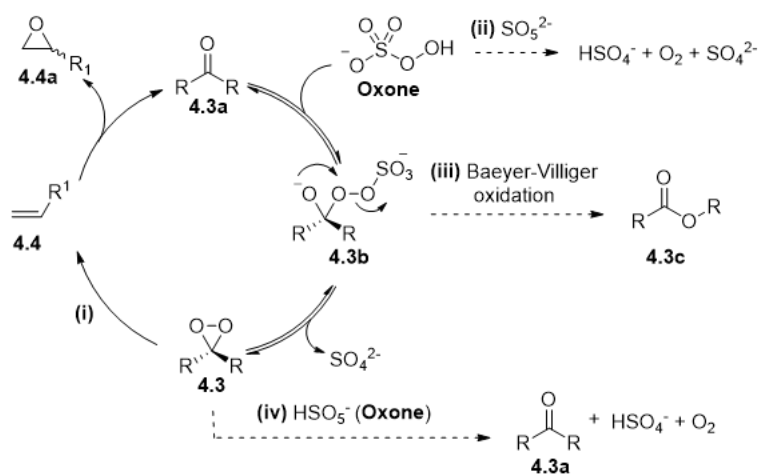
Scheme 4.1. General epoxidation conditions as published by Curci *et al.*; **a)** general structure of the dioxirane that is formed by reaction of the ketone with Oxone. **b)** acetone in water solution. **c)** biphasic solution with tetrabutylammonium hydrogensulfate as a phase transfer catalyst.

A significant benefit of this reaction was that the dioxirane formed is reactive enough to react even with unfunctionalized alkenes in mild conditions. The reaction was carried out at 2-8°C, at a neutral pH (7-7.5), either in acetone-water mixture, or in a biphasic solution with the addition of a phase transfer catalyst (such as tetrabutylammonium hydrogensulfate, TBAHS) (**Figure 4.1b** and **c**). This meant that the reaction could be used to form reactive peroxides, that may be

acid/base/temperature sensitive, as well as reaction of hydrophobic alkenes that are water insoluble.^{151, 152}

4.1.2. Mechanism, proposed transition state and reactivity.

Once there was evidence that the reactive species in the reaction is a dioxirane, a general mechanism was proposed (**Scheme 4.2**): the Oxone reacts with the ketone (**4.3**) to form a tetrahedral intermediate (**4.3b**), which cyclizes intramolecularly to form the dioxirane (**4.3a**). Next, the dioxirane can react with a nucleophile such as an alkene (**Scheme 4.2**, pathway **(i)**, **4.4**) to form the epoxide product (**4.4a**). The reaction has the potential for three different unproductive pathways (**Scheme 4.2**): **(ii)** Oxone auto decomposition; **(iii)** Baeyer-Villiger oxidation of the catalyst, which renders it unreactive; **(iv)** or reaction of the dioxirane (**4.3a**) with Oxone to reproduce the ketone (even though the catalyst is regenerated, the Oxone is consumed by this step). The unproductive pathways will be discussed further below.^{151, 153}



Scheme 4.2. General mechanism for the reaction of a ketone (**4.3**) with Oxone to give the epoxide product (**4.4a**) as well as 3 unproductive pathways (dashed arrows) that can inhibit product formation.

Another aspect of this reaction is the stereospecificity, that is, the epoxide that is formed retains the stereochemistry of the double bond. A *trans* alkene will always lead to the formation of a *trans* epoxide (as can be seen in **Scheme 4.1**), and a *cis* alkene will lead to the formation of a *cis* epoxide (for chiral molecules, both stereoisomers are formed).^{152, 154}

Two general possible transition states (TS) were proposed for the epoxidation step. The dioxirane can either approach the alkene forming a planar TS or a spiro TS (**Figure 4.1**). Observations by Baumstark and co-workers reveal that the spiro TS is likely the favoured one.¹⁵⁴ This is due to the reactivity difference between *cis* and *trans* alkenes; the *cis* hexenes were found to be more reactive than their *trans* counterpart by a factor of ~ 8 . The preference towards *cis* alkenes can be explained as a steric preference in the spiro TS, whereas in the planar TS one would expect the reactivities to be similar.¹⁵⁴ Further computational work by Houk and co-workers agreed with this hypothesis: the spiro TS was lower in energy, and this was explained by a secondary interaction between the oxygen's n orbital and the alkene's π^* (**Figure 4.1**).¹⁵⁵⁻¹⁵⁷

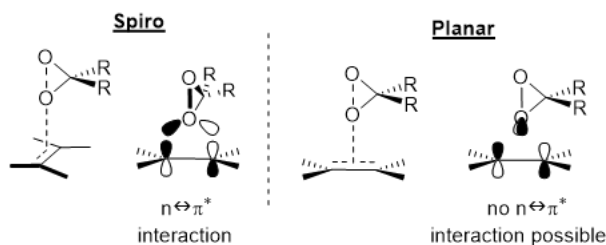


Figure 4.1. Two possible transition states proposed. Left: spiro TS, secondary interaction between the oxygen HOMO (n orbital) and the alkene LUMO (π^* orbital). Right: planar TS, the secondary interaction is not possible (or is minimal).

A few reactivity and efficiency observations are worth mentioning:

- a) α -trifluoromethyl (**Figure 4.2a**) groups on the ketone enhance the reactivity of both the ketone and the dioxirane formed (likely by increasing the electrophilicity), and reduce reaction times.¹⁵⁸
- b) The Oxone is susceptible to auto-decomposition. Trace metals in the reaction can increase the decomposition rate (**Scheme 4.2**, pathway **(ii)**). This is mitigated by the addition of a chelating agent such as Na₂EDTA. Additionally, the Oxone solution is added with Teflon needles, instead of metal ones.^{153, 159}
- c) Addition rate of the Oxone solution plays a big role in the reaction's outcome:
 - i. Once the dioxirane is formed, the presence of excess Oxone in the solution can lead to dioxirane decomposition, prior to the epoxidation reaction taking place (**Scheme 4.2**, pathway **(iv)**).
 - ii. Auto-decomposition also increases with an increase in pH as well as with high Oxone concentration (**Scheme 4.2**, pathway **(ii)**).Keeping the Oxone concentration low throughout the reaction by slowly adding a fresh solution instead of initial addition of excess Oxone can mitigate the problem of these unproductive pathways.¹⁵³
- d) The reaction has a strong dependence on pH. Specifically, the dioxirane formation has been shown to be at its maximum at pH 7.5-8.0. As pH control is necessary and formation of dioxirane is coupled with the formation of hydrogen sulfate (weakly acidic), a buffer is usually used.¹⁵³ Further discussion on the reaction's conditions, including pH, will be presented under the asymmetric epoxidation section.

- e) Baeyer-Villiger oxidation of the ketone is a probable side reaction that leads to catalyst deactivation (**Scheme 4.2**, pathway **(iii)**).^{151, 153, 159} Although the ester products have not been observed in many cases, Denmark *et al.* observed some lactone formation when testing 3-oxopiperidinium compounds as catalysts (**Figure 4.2b**).¹⁵³

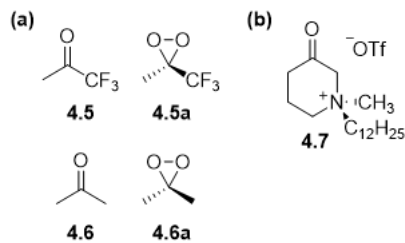


Figure 4.2. (a) Dioxiranes with α -trifluoromethyl groups, **4.5a**, are more reactive than their methyl counterparts, **4.6a**. The ketones from which the dioxiranes are formed likely follow the same reactivity trend in the dioxirane formation (**4.5** being more reactive than **4.6**). (b) 3-oxopiperidinium (**4.7**) that was tested by Denmark *et al.* and lactone formation from the Baeyer-Villiger oxidation was observed.

4.1.3. Asymmetric epoxidation and Shi's catalyst.

The first asymmetric variation was published by Curci *et al.* shortly after their initial publication, with the use of two chiral ketones, albeit producing low stereoselectivities (4.2-12.5 %*ee*) (**Figure 4.3**, structures **4.8** and **4.9**).¹⁶⁰ A big step forward came when Shi *et al.* published their fructose-derived chiral ketone catalyst for the reaction (**Figure 4.3**, **4.10**). Enantioselectivities reported between 70-95 %*ee* for multiple substrates proved the system to be promising (**Scheme 4.3a**).¹²²

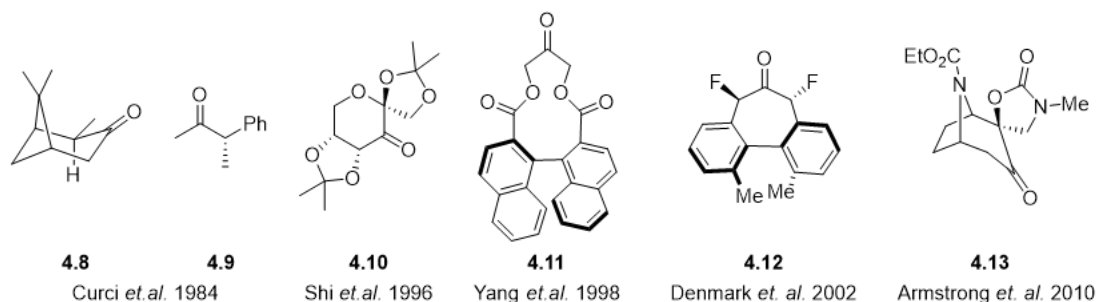
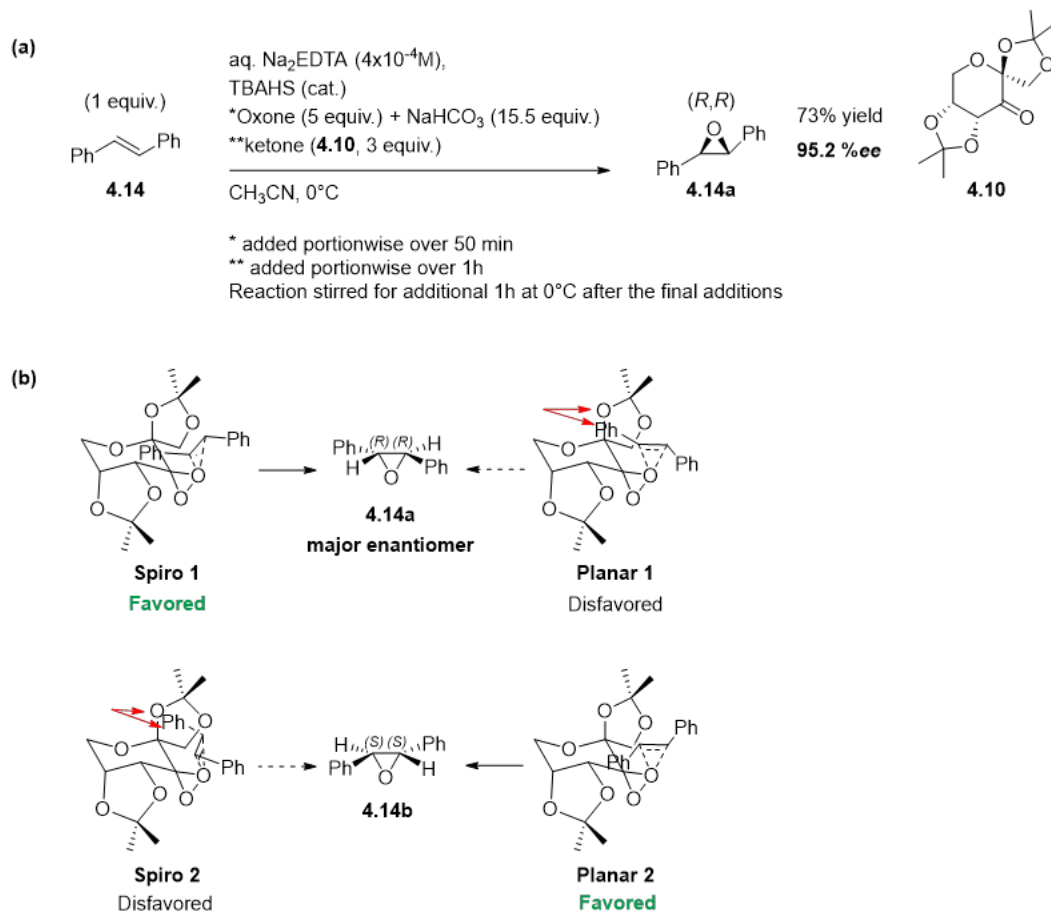


Figure 4.3. Chiral ketone catalysts published by Curci *et al.* (**4.8** and **4.9**), Shi *et al.* (**4.10**), Yang *et al.* (**4.11**), Denmark *et al.* (**4.12**), and Armstrong *et al.* (**4.13**).

A benefit of the catalyst designed by Shi is that it has been designed to offer high selectivity for *trans* alkenes. As mentioned in the previous section, the reaction likely favours the spiro-TS (**Figure 4.1**). As observed by the lower reactivity of *trans* alkenes, the steric clash in this TS means their selectivity will likely also be lower, since both faces of the alkene will clash to some degree. Using *trans*-stilbene (**4.14**) as an example for the system, Shi was able to provide further evidence that the reaction likely goes through a spiro TS (**Scheme 4.3**). The major enantiomer (*R,R*) (**4.14a**) in the reaction likely goes through a spiro TS (**Scheme 4.3b, Spiro 1**), as a planar TS (**Scheme 4.3b, Planar 1**) involves steric clashes between the phenyl of the substrate and the acetal oxygens of the ketone that render this TS less favorable, and therefore less likely. The **Planar 2** TS could also be a pathway the reaction could proceed through. However, had the reaction proceeded via this TS, the major enantiomer would have been the (*S,S*)-epoxide (**4.14b**). Theoretically, the reaction likely proceeds through both competing TS (**Spiro 1** and **Planar 2**), and the enantioselectivity depends on factors that affect the energy difference between these two TS (such as substrate structure- steric and electronic interactions). From this, one can conclude that the most favored TS (for the substrate **4.14**) is the **Spiro 1** TS.^{122, 157, 161}



Scheme 4.3. (a) Initial epoxidation conditions as published by Shi *et al.*, with *trans* stilbene (**4.14**) as the example, using chiral ketone **4.10**. (b) Four proposed TSs for the epoxidation of *trans* stilbene (or any symmetrical alkene). The major enantiomer being **4.14a** provides evidence for the favored TS being **Spiro 1**.

Shi's chiral ketone (**4.10**) also provided impressive enantioselectivities for trisubstituted alkenes (>90 %ee). However, for *cis*- and terminal- alkenes the enantioselectivities are not as good. Shi proposed eight different TSs that are possible for tri-substituted alkenes (for disubstituted some of the TSs will be equivalent) (**Figure 4.4**). Four of the TSs are "spiro" (**Spiro A-D**), the other four are planar (**Planar E-H**). As can be seen in **Figure 4.4**, TS **Spiro B-D** and **Planar F-H** all contain some steric clash with the catalyst's acetal group (highlighted with red arrows). This reduces the most likely TSs for the reaction to go through to two: **Spiro A** and **Planar E** (in dashed rectangle).

As these two TSs lead to opposite enantiomeric products, the energy difference between them will determine the percentage of enantioselectivity, as well as the major enantiomer (as was demonstrated above for the specific case of *trans*-stilbene).¹⁵⁷

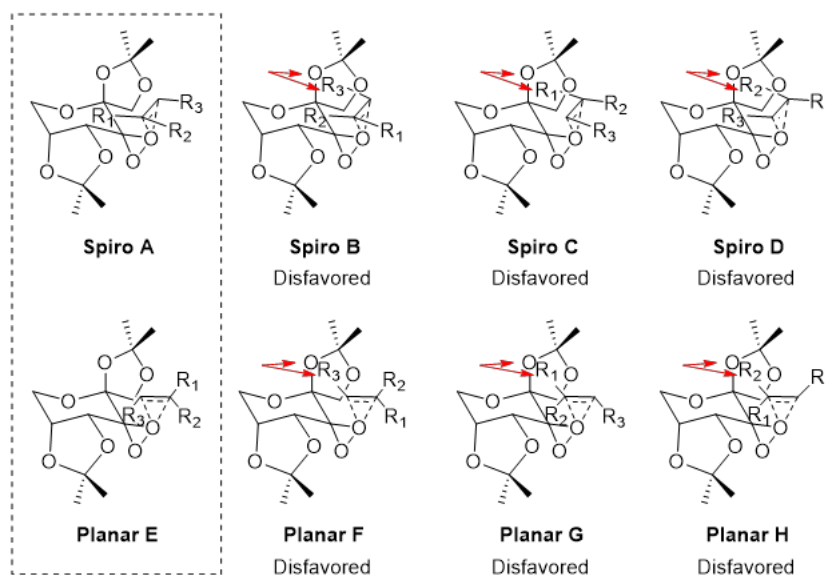


Figure 4.4. Eight possible TSs for trisubstituted alkenes. The major competing TSs, that lead to the opposite enantiomers, are **Spiro 1** and **Planar E**. For the disfavored TS, steric clashes are highlighted with red arrows.

Generally, conjugated groups, such as the *trans*-stilbene (**4.14**) favoured the product obtained by going through the Spiro TS (**Spiro 1**, **Scheme 4.3b**, or **Spiro A**, **Figure 4.4**, where $R_2=H$). This observation can be explained by the lower energy of the alkene LUMO (π^* orbital) due to conjugation, which increases the orbital interaction as shown in **Figure 4.1**. Increasing the size of the R_3 group disfavors **Planar E** (increased clash with the axial H). Similarly, decreasing the size of the R_1 group favours the **Spiro A** TS.

4.1.4. Reaction conditions

As mentioned above, the reaction's conditions have a major effect on the reaction's outcome. Interestingly, most of the work published achieved better results with a pH kept at around 7.5-8.0.

The change from running the reaction at pH 8.0 to running it at pH 8.5 in Denmark's study (with ketone **4.7**) showed a significant decrease in epoxide formation; from ~95% after 25 hours at pH 8.0, to ~12% of epoxide forming after 25 hours at pH 8.5.¹⁵³ Although they do not address this in their report, it is likely due to the Oxone auto decomposition happening faster at a higher pH, and the catalyst not being able to out-compete this reaction, leading to consumption of the oxidant.

It is important to note here that each group uses a different variation of the conditions (ketone, alkene substrate, solvent, temperature, Oxone addition rate and amount, etc.). Thus, the observations provided here of differing conditions, such as pH, are broad and meant to give an understanding and perspective, since different catalysts will not have the same reactivity at the same conditions or may even be unstable in some.

After observing that their ketone catalyst (**4.10**) was rapidly decomposing at pH 7-8, Shi and co-workers also studied the reactivity of their ketone catalyst under different pHs.¹⁶² They hypothesized that a higher pH could favor the epoxidation over the possible catalyst deactivation by Baeyer-Villiger oxidation, as well as higher nucleophilicity of the Oxone. In addition to a higher stability of their catalyst at higher pH (although they did not discuss the nature of the decomposition, one can assume that the acetals would be more labile at near neutral conditions in the presence of water). The reaction showed a significant increase in conversion at a pH of 10.5, from which they deduced that the possible racemic background reaction, as well as the auto decomposition of the Oxone, are negligible compared to the epoxidation by the ketone.¹⁶²

4.2. Results

4.2.1. Procedure establishment with Shi's chiral ketone catalyst

Our ultimate goal is to develop novel catalysts using VIRTUAL CHEMIST to identify novel chemical series. However, as discussed in Chapter 2, the first necessary step is to ensure the published results are reproducible by us, so we can compare our results to them. This step proved to take a lot more effort than anticipated, and multiple attempts and different published procedures were carried out and will be described here (**Table 4.1**). The general reaction scheme is shown in **Scheme 4.5**. Although the reaction has been published and tested on different test substrates, out of these we chose to conduct our studies using *trans*-stilbene. This is because the enantioselectivity can be measured using a chiral HPLC, whereas many of the other substrates used by Shi *et al.* (such as *trans*- β -methylstyrene) required a chiral GC, which we do not have direct access to.

Initially, the optimised conditions as published by Shi *et al.*¹⁶² were carried out (**Entry 1**). The chiral ketone (**4.10**) was added in 30 mol% (0.3 equiv.), the Oxone and the potassium carbonate were added dropwise over 90 minutes using a syringe pump, at 0°C. The only difference was that a smaller scale was used (0.6 mmol of *trans*-stilbene, instead of 1 mmol). The reaction was kept at 0°C throughout and was quenched once the addition was completed. Unfortunately, only about 5% conversion was observed (by ¹H-NMR), the rest was recovered starting material.

As the obvious difference in our conditions was the reaction's scale, we decided to run it again, but at the same scale as the published reaction (1 mmol of the *trans*-stilbene) (**Entry 2**). While this reaction afforded 18% conversion, most of the recovered material was still starting material. This conversion was significantly lower than the reported 73%.

The reaction has previously been published under different conditions, so we decided to simplify the conditions and use acetone as the ketone (**Entry 3**). The only difference here compared to the previous conditions is the use of acetone (in excess, 3 equiv.) instead of the chiral ketone, the rest remaining the same. Using 1 mmol of *trans*-stilbene, and 3 equiv. (3 mmol) of acetone, with Oxone and potassium carbonate solutions, added separately over 95 minutes. This reaction yielded only starting material, and no conversion was observed.

Since the reaction has many variables, it was hard to know which one was the culprit to the low conversion (with ketone **4.10**) in the setup. We decided to screen a wide range of conditions in the hopes of improving the yields.

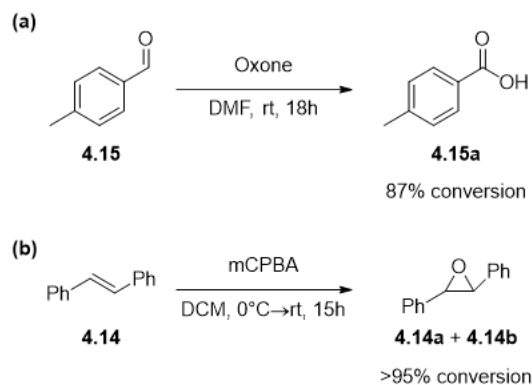
Some of the conditions were kept identical to the previous ones, 1 mmol *trans*-stilbene, 0.3 equiv. (30 mol%) of the chiral ketone catalyst, base and Oxone solutions added separately over 100 minutes using a syringe pump, all these steps being done at 0°C (**Entry 4**). This time, instead of quenching the reaction at the end of the addition, the reaction was left to stir for an additional 3.5 hours, at 0°C. Encouragingly, 38% conversion was observed.

Encouraged by this improvement, we theorized that maybe the reaction needed more stirring time. Additionally, Oxone itself can decompose overtime in storage, and the phase transfer catalyst (TBAHS) is very hygroscopic, so it was possible that the catalytic amount measured for the reaction was considerably smaller and not enough for the time scale of the reaction. To test this, two more (smaller scale) test reactions were carried out; one with twice the amount of Oxone (**Entry 5**), and the other with twice the amount of TBAHS (**Entry 6**). Both reactions were allowed to stir for longer once the addition was complete. Conversion (qualitative) was monitored by TLC, and after 20-24 hours it was clear that the reactions would not convert to product in a significant

amount. The product remained a very faint spot on TLC, and mostly starting material was observed. Although TLC is not a quantitative method, i.e., the starting material stilbene has a stronger UV absorption, and so always seems like there is more of it compared to the product, it was clear enough that the conversion was minimal. The reactions were therefore not worked-up further.

The next two reactions were carried out using other published conditions by Shi *et al.*;¹⁶³ (**Entry 7**) In this reaction the oxidant used is hydrogen peroxide instead of the Oxone, and the addition and reaction time are slightly different (see experimental section). This reaction afforded a disappointing 13% conversion.

As the amounts of Oxone and phase transfer catalyst were not the main issue, and hydrogen peroxide also did not seem to improve the results, the Oxone was tested as an oxidant in a different reaction to verify its reactivity. The reaction was an oxidation of *p*-tolualdehyde into *p*-toluic acid (**Scheme 4.4a**).¹⁶⁴ The reaction afforded 87% conversion of the starting material into product. From this, it was safe to conclude that the Oxone was still reactive and could be used in the Shi epoxidation. By this point, the variables that were ruled out as to why the reaction was not producing as much product as expected were the scale, the catalyst, stirring time after addition of the Oxone and the base, the oxidant amount, the oxidant reactivity, and phase transfer catalyst amount. It should also be mentioned that the *trans*-stilbene was tested for purity by ¹H-NMR and was also successfully used in a racemic epoxidation reaction using mCPBA as the oxidant (**Scheme 4.4b**).



Scheme 4.4. (a) Test reaction of Oxone's reactivity in a different oxidation reaction. (b) *Trans*-stilbene (4.14) racemic epoxidation with mCPBA.

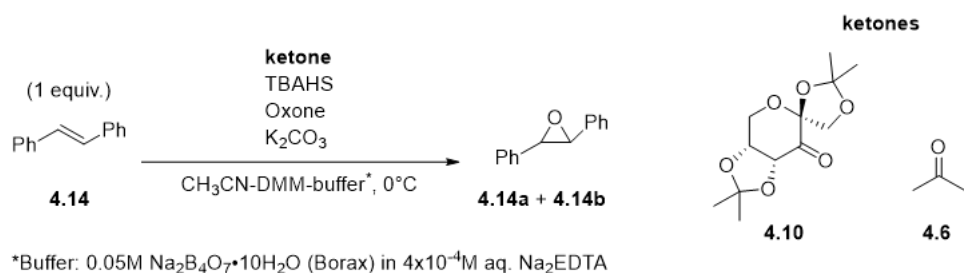
Similarly to using hydrogen peroxide, Shi's group initial published conditions were also slightly different (**Entry 8**), where the ketone catalyst, Oxone, and sodium bicarbonate were simultaneously added portion wise over 10 minutes (**Scheme 4.3a**).¹²² The reaction was stirred for an additional 12 more minutes at 0°C, and then at room temperature for 48 hours with 27 % yield.

Next, we decided to run the reaction again using the original conditions, however this time using 1 equiv. (100 mol%) of the chiral ketone catalyst instead of 0.3 equiv. (**Entry 9**). The reaction was quenched once the addition of the Oxone and base was completed (100 minutes). Additionally, more attention to small details made sure the stilbene is properly dissolved in the DMM before anything else was added, the ketone catalyst was pulverised prior to addition, and the flask was sonicated once it was added. This reaction afforded 25% conversion, which was encouraging, yet far from the reported yields. Although previously we found that additional stirring time added to the conversion, we decided to not continue testing this option as it has been mentioned that longer reaction time with the Oxone present can lead to more side products (see Introduction). We therefore hoped to achieve comparable conversion when the reaction is quenched once the addition is complete.

The next variable we changed was the addition time, using a slower addition of the Oxone and base solutions (**Entry 10**). The base solution and the Oxone solution were added using a syringe pump over 2 hours instead of over 90 minutes, and the reaction was quenched once the addition was completed, giving 35% conversion. With this promising result, the reaction was repeated (**Entry 11**), however this time the addition was done over 3.5 hours. Gratifyingly, this reaction afforded 79% conversion and 59% yield after purification of the product.

When the reaction was repeated with addition over 3 hours but with a catalytic amount of the ketone catalyst (30 mol%), only 50% conversion was observed (**Entry 12**). It is logical to assume that a lower amount of catalyst would require a longer reaction time (longer addition time). And so, the reaction was repeated, with 30 mol% of the catalyst, and Oxone and base solutions which were added over 6 hours, affording 72% conversion (**Entry 13**).

At this point, we were satisfied that we were able to modify the procedure such that we can carry out the reaction with other ketones and compare the results to our original system. We decided that using the 3.5-hour addition, with 100 mol% catalyst would be the system we would compare all other ketones (i.e., potential new catalysts) to. Once the product was in hand and purified, obtaining the %*ee* using a chiral HPLC was possible, where the results were comparable to published ones (see Supporting Information).



Scheme 4.5. General scheme for the epoxidation reaction carried out with different conditions (as detailed in **Table 4.1**). TBAHS: tetrabutylammonium hydrogensulfate, DMM: dimethoxy methane.

Table 4.1. Different conditions that were used for the epoxidation reaction, with the purpose of improving yield and reaching a result comparable to the published one.

Entry	<i>Trans</i> -stilbene ^a (4.14) [mmol]	Ketone [equiv.] (mol%)	Addition time [min] or [h]	Additional stirring time [h]	Conversion ^b [%]	Additional changes
1	0.6	0.3 (30)	90 min	-	5	-
2	1	0.3 (30)	100 min	-	18	-
3	1	3 (300)	95 min	-	-nd ^c	Acetone as ketone
4	1	0.3 (30)	100 min	3.5 h	38	-
5	0.3	0.3 (30)	30 min	20 h	-nd ^c	x 2 amount of Oxone
6	0.3	0.3 (30)	30 min	24 h	-nd ^c	x 2 amount of TBAHS
7	1	0.3 (30)	-	-	13	H ₂ O ₂ as oxidant
8	0.2	3 (300)	-	-	27	Potion wise addition
9	1	1 (100)	100 min	-	25	-

Chapter 4

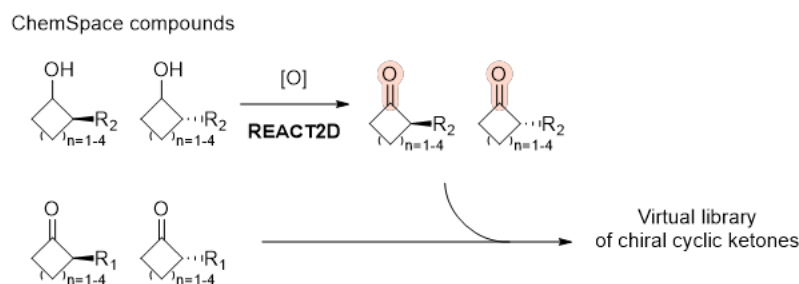
10	1	1 (100)	2 h	-	35	-
11	1	1 (100)	3.5 h	-	79	-
12	1	0.3 (30)	3 h	-	50	-
13	1	0.3 (30)	6 h	-	72	-

^a *trans*-stilbene is always 1 equiv. ^b determined by ¹H-NMR. ^c not determined.

4.2.2. Virtual screening: ChemSpace library

Once we had a working procedure in hand, we next set out to do a virtual screening on chiral ketones. As described in Chapter 2, the ZINC database was problematic, as ideally the compounds would be purchased as enantiopure or synthesizable. To address this issue, the virtual screen was done on the ChemSpace library of chiral cyclic ketones. Cyclic ketones were chosen since they have less possible conformations, and therefore more ordered TSs. The less options of different TSs there are, the better chances are for higher enantioselectivity (for this reason the reported catalysts are also cyclic). It became clear that the library size was considerably smaller when compared to the amine library, which left us with only 230 compounds to screen and filter. To increase the library size, we created a library of secondary cyclic alcohols that are next to a chiral center and used REACT2D to oxidise them into ketones (**Scheme 4.6**). While these ketones would not be directly accessible from commercial vendors, they would be synthesizable in a single step from commercial material. This new library of ketones was merged with the original cyclic ketone library, 631 ketones total, which was now ready for screening with VIRTUAL CHEMIST. The library was filtered to contain cyclic 4, 5, 6, and 7-membered chiral ketones.

The results of this screening showed very little diversity in the general scaffolds that were top ranked (**Figure 4.5**). The library had been filtered to contain 4, 5, 6, and 7-membered rings, because we wanted to have more variability in it. However, we did not have a lot of prior knowledge about the reactivity of ketones of different ring sizes, or if maybe even non-cyclic ketones could be considered. Learning from the Diels-Alder study, we decided to first try out simple ketones of various ring sizes, as well as non-cyclic ones, and test them for reactivity in this reaction. Once we had a better understanding of which ketones are reactive enough, a new library could be constructed. It is important to note that based on the proposed mechanism of the reaction, as well as some previous studies, it is known that more electron poor (i.e., more electrophilic) ketones perform better in this reaction. However, other structural influences on reactivity are less clear, and it was therefore worth studying.



Scheme 4.6. General scheme for creating the virtual library of chiral cyclic ketones: cyclic ketones adjacent to a stereocenter, with ring sizes of 4-7 carbons (bottom) were filtered for, resulting in a very small library. Cyclic secondary alcohols (of inconsequential stereochemistry at the hydroxy carbon) adjacent to a stereocenter, with ring sizes of 4-7 carbons (top) were also selected in a separate library. These alcohols were then virtually oxidized into the corresponding ketones using REACT2D (these ketones are highlighted to differentiate them from the ketones derived directly from the ChemSpace library). The ketone libraries were merged into a single library that was used in the virtual screening.

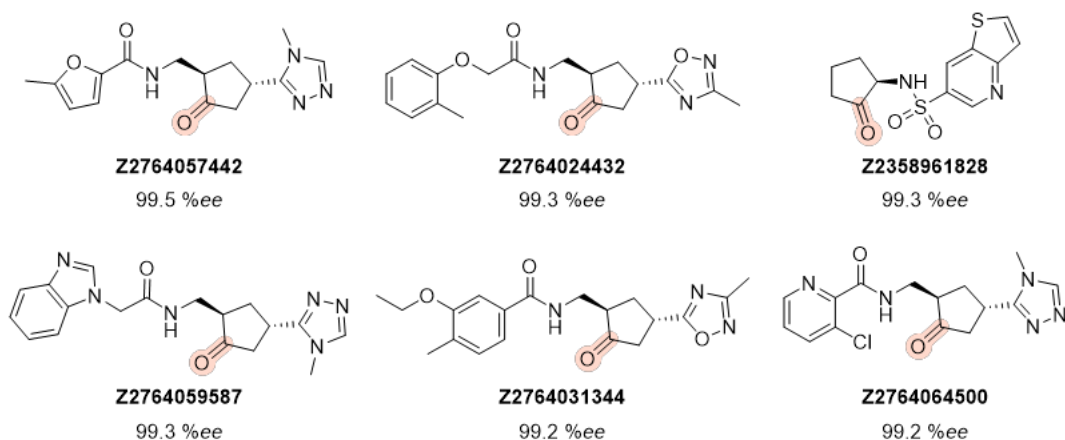


Figure 4.5. Top results of the ChemSpace virtual screening. All the highlighted ketones come from secondary alcohols that were oxidized into ketones with REACT2D. Under each molecule is the ChemSpace ID for the molecule the ketones are derived from, and the averaged %ee as predicted by ACE.

4.2.3. Initial study of structure-activity relationship

Some ketone structure probing studies have previously been reported.^{153, 165, 166} In the study done by Denmark *et al.*¹⁵³ the general conclusions were that α -substitution on acyclic ketones significantly lowered reactivity. For cyclic ketones the reactivity was ring size dependent, with cyclohexanone being the most reactive (with 67% conversion, although less reactive than acetone in their studies which had 87% conversion). The increase of electrophilicity of the ketones (trifluoroacetone or hexafluoroacetone) decreased their reactivity significantly (2-29% conversion). The last observation is in contradiction to the observation made by Yang *et al.*,¹⁵⁸ which observed higher reactivity for trifluoromethylacetone. However, in 2002 Denmark *et al.* designed a new catalyst (**Figure 4.3, 4.12**) where fluorine was added as an electron withdrawing group in order to increase reactivity under different conditions.¹⁶⁷ Denmark *et al.* also probed the ring size's influence on the reactivity of their ammonium salt catalysts (**Figure 4.6a**).¹⁵³ The ammonium group was added to serve the phase transfer function, in addition to the epoxidation

catalysis. The most reactive salt was the 6-membered ring **4.7**. Yang *et al.* developed a different series of chiral ketones (**Figure 4.6b**). Ketone **4.11a** was tested to examine the importance of the ester groups in **4.11**. It was found that the ester groups were indeed necessary for good conversion (10 % conversion was observed for **4.11a**). They rationalized this observation by comparing the electron withdrawing ability of an ester compared to the ether. The ester being a stronger electron withdrawing group, it activated the ketone better than the ether. Further work done by Shi *et al.*¹⁶⁶ will be discussed further below.

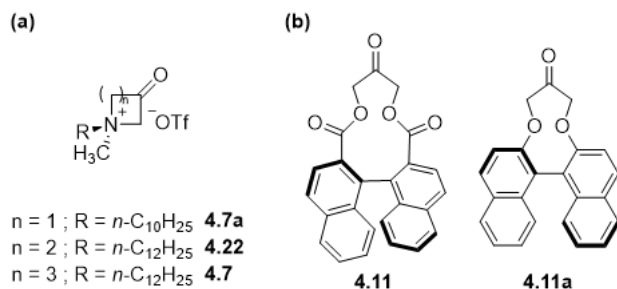


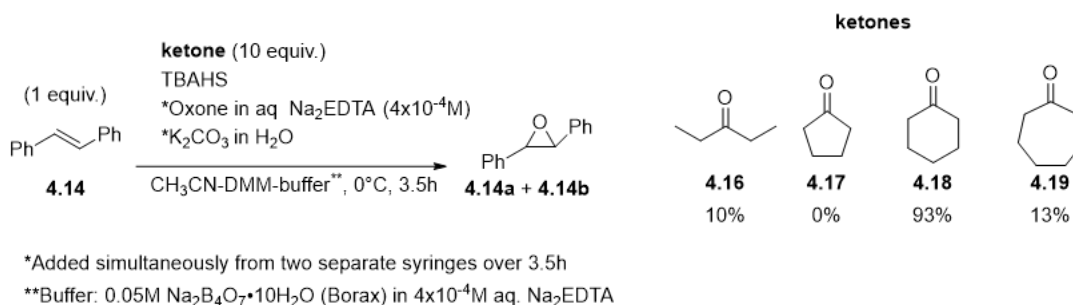
Figure 4.6. Chiral ketones with different structures tested for their reactivity (a) by Denmark *et al.* (b) by Yang *et al.*

While there appears to be a general trend, the results are likely also affected by the conditions used. Additionally, the experience of developing a procedure where the reaction works well with a known catalyst and with good reproducibility led us to conduct our own structure activity relationship of different ketones.

The goal of this study is to learn about the reactivity of different ketones under the same conditions we optimized in the previous section (**Scheme 4.7**). To get a general sense of reactivity we chose to use the ketones in excess (10 equiv.), since it is obvious some ketones will not be very reactive. We were interested in knowing if some product was formed, or none. The initial plan was: (i) find general scaffolds that show some reactivity (albeit low initially). (ii) assemble a new library of chiral ketones and secondary alcohols (oxidised into ketones using REACT2D), based on

the findings from the previous steps. **(iii)** run a virtual screen for enantioselectivity on the library. **(iv)** select a few selected ketones from the results, synthesize/purchase them, and test them under the reaction conditions for both reactivity and selectivity.

So far, four different simple ketones have been tested (**Scheme 4.7**). Cyclohexanone (**4.18**, 93% conversion) shows significantly increased reactivity compared to the other ketones, with the least reactive ketone being the cyclopentanone (**4.17**, 0% conversion). The Baeyer-Villiger oxidation product was not observed in any of the reactions. Additionally, it is safe to assume that when a low conversion is observed (**4.16** and **4.19**), the product was obtained through epoxidation with the ketone and not a direct epoxidation through the Oxone, as for **4.17**, under the same conditions, no product was observed.



Scheme 4.7. General procedure used to test different ketones' reactivity. The ketones tested thus far are on the right. Under each ketone is the % conversion into product as measured by ¹H-NMR.

As an initial analysis of the reactivity difference between the different ketones, the ketone conformation and accessibility of the carbonyl carbon may make a difference on the initial nucleophilic attack by the Oxone (**Scheme 4.2, 4.3b**).

Another difference can arise from the reactivity and stability of the dioxirane (**Scheme 4.2, 4.3a**) that is formed (if it is formed). The dioxirane can react with more equivalents of Oxone which leads to reformation of the free ketone prior to reaction with the alkene (**Scheme 4.2,**

pathway (iv)). Less stable dioxiranes will decompose faster than they will react with the alkene, which will lead to lower yields. It has been previously observed, that for noncyclic ketones such as **4.16** the lower reactivity may result from steric hindrance. The flexibility of the ethyl groups on each side of the carbonyl may impede the initial formation of the dioxirane.¹⁵³

Interestingly, cyclopentanone (**4.17**) yielded no product at all. In previous work, Shi and co-workers tested a 5-membered cyclic ketone, which resulted in very low conversion (**Figure 4.7**, **4.20** and **4.21**, 14 and 0% conversion, respectively). In their work, they did not provide a hypothesis as to the significant decrease in reactivity.¹⁶⁶ Denmark *et al.* also had no conversion when using an electron-poor 5-membered ring ketone (**Figure 4.7**, **4.22**), they concluded in their study that the dioxirane likely does not form. However, there is no hypothesis as to why this ketone would be less reactive.¹⁵³

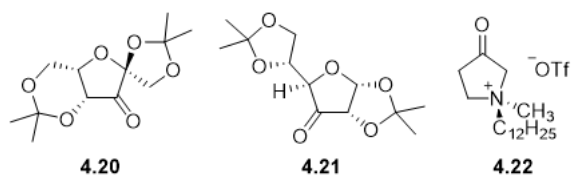


Figure 4.7. 5-membered ring ketones tested by Shi *et al.* (**4.20**, **4.21**) and Denmark *et al.* (**4.22**).

Due to time constraints, as of now only the first goal (i) has been reached, simple ketones have been tested on the reaction conditions and the conversion measured. Baeyer-Villiger oxidation product has not been observed for any of the ketones tested. There are significant reactivity differences between the different cyclic ketones, though it is yet not entirely clear where the difference is coming from.

4.3. Conclusions and future work

In this chapter a reliable procedure for running the Shi epoxidation in the lab was optimized, as well as initial work into different ketone structure reactivity. This chapter emphasizes again how little is known about different well-established reactions, and the need to further investigate the reactivity of potential catalysts.

Unfortunately, we were unable to continue through all the planned steps. However, everything is now ready to proceed with a new virtual screening. Based on our results, the library should be limited to 6-membered rings, and REACT2D should be employed to increase its size. In the chapter we used REACT2D to increase the size by oxidizing secondary alcohols into ketones. More reactions may be considered in order to increase the size, such as the hydrolysis of acetals, alkene cleavage *via* ozonolysis, and others can be employed.

This project presents a few challenges, different than the ones faced in the iminium study (Chapter 3). The reaction has many variables, and it is difficult to assess how the change of one would affect the outcome (such as the significant change in conversion when addition time was changed), including stereoselectivity. Additionally, as has been briefly mentioned in the introduction, each catalyst requires some optimization of the conditions, with the conditions used in this study being derived from Shi's work. However, it is important to remember that the conditions are not universal, and we can only infer about reactivity here under those conditions. Interestingly, the use of VS would significantly reduce this burden. If the project yields a reactive enough ketone, optimization of the conditions for ideal reactivity should be done. Practically, since the reaction necessitates addition over 3.5 hours using a syringe pump, the throughput of the procedure also limits the number of reactions that can be tested.

In terms of virtual screening, the relatively small size of the initial library for chiral ketones is another obstacle that needs to be overcome. Since most currently available libraries (such as ZINC database and ChemSpace) are directed towards drug discovery, and ketones are not a common functional group in this field, the number of available molecules in them is limited. Discussions with Liverpool Chiro Chem have been initiated.

This study highlights that VIRTUAL CHEMIST should include some more catalysts descriptors. While we found that pKa was key in the Diels Alder reaction (and a pKa predictor is being developed), the measure of reactivity (i.e., electron-poor in this case) would also be required and implementations and evaluation of reactivity descriptors have started. As a note, these descriptors must be fast enough (long DFT calculations are not an option) to maintain the throughput of the virtual screening procedures.

Chapters 1-4 describe the connection between computational chemistry and asymmetric catalysis. In **Chapter 5** we joined the efforts of targeting SARS-CoV-2 using our group's expertise in combining computational chemistry with drug discovery. Although the theme of the chapter is different, the approach remains parallel to the development of a new catalyst; published data is used to gain insight into structure-activity relationships, a compound is computationally designed, synthesized, and tested. Unlike the previous chapters, we did not conduct a virtual screen in this case due to the lack of information at the outset of this work on active compounds on PLpro, an enzyme that has been known to be difficult to target. For this reason, we chose here to use published data and proceed with rational design.

The work on this chapter is based on work published in the preprint, the chapter details the work of SP and not the entire paper.

Ying Huang, H.; **Pinus, S.**; Zhang, X. C.; Wang, G.; Rueda, A. M.; Souaibou, Y.; Huck, S.; Huot, M.; Vlaho, D.; Pottel, J.; Venegas, F.; Lu, Z.; Hennecker, C.; Stille, J.; Tjutrins, J.; Miron, C.; Labarre, A.; Plescica, J.; Burai-Patrascu, M.; Mittermaier, A.; Moitessier, N. Integration of Computational and Experimental Techniques for the Discovery of SARS-CoV-2 PLpro Covalent Inhibitors. *ChemRxiv* **2023**. This content is a preprint and has not been peer reviewed.

All the synthesis and design work described in this chapter was done by SP, except for the synthesis detailed in Scheme 5.5 which was carried out by YS, and the synthesis detailed in Scheme 5.9 which was carried out by SH. The biological assays were done by YHH.

Chapter 5

5.1. Introduction

5.1.1. General information about the virus and pandemic

SARS-CoV-2 is a single stranded positive RNA virus. It belongs to the same family (coronavirus) as other known coronaviruses, some of which are as mild as the common cold,¹⁶⁸ ¹⁶⁹ however also includes SARS-CoV (which resulted in the 2001 epidemic), and MERS (2007).¹⁷⁰ SARS-CoV-2 is responsible to the COVID-19 pandemic that started in Wuhan, China towards the end of 2019. Since the beginning of the COVID-19 pandemic there have been over 760 million confirmed cases, and 6.8 million cumulative deaths, worldwide (in Canada 4.6 million confirmed cases and 51.6 thousand deaths).¹⁷¹ In addition to the severe health effects, the pandemic also caused serious economic slump. In Canada alone, the unemployment rate spiked from 5.7% in 2019, to 9.5% in 2020, the highest it has been in almost 20 years. While the unemployment rate is encouragingly decreasing, the affect it had on individuals can take years to stabilize.¹⁷²

5.1.2. Infection and viral replication

The viral replication of SARS-CoV-2 begins with the attachment and entry into the host cell membrane followed by the translation of viral proteins into a polyprotein. Proteolysis of this polyprotein chain is achieved via two proteases – PLpro (papain-like protease) and 3CLpro (3-chymotrypsin-like protease). Following, the RNA will be replicated, and the transcription and translation of structural and accessory proteins will take place. The process concludes with the assembly, packaging, and release of SARS-CoV-2. Thus, inhibiting any of the stages of the viral replication would be a viable strategy to combat the disease.^{170, 173}

5.1.3. Response to the pandemic

As the pandemic progressed ample effort was placed into coming up with preventative measures, such as vaccines, which have been produced, tested, and put into use rapidly.^{174, 175}

Considerable effort has also been put into the development novel, and the utilization of existing drugs as anti-SARS-CoV-2 drugs.^{176, 177} The first anti-viral drug to be approved was remdesivir, approved to be repurposed for the treatment of Covid-19 patients in October 2020.¹⁷⁸ Unfortunately, despite much faith in the treatment, clinical trials revealed mixed results, and in general, it seemed to not have a significant effect on hospitalized patients as evidenced by ventilation initiation and overall mortality.¹⁷⁹

The quick response and the ability to fast track clinical trials and approval, led to the approval of three antiviral drugs (**Figure 5.1**). Paxlovid (nirmatrevir and ritonavir) developed by Pfizer, is a covalent drug targeting 3CLpro.^{180, 181} Another drug that targets 3CLpro is Xocova (Ensitrelvir), developed by Shionogi, a non-covalent anti-viral.¹⁸² Molnupiravir has been developed by Merck and targets RNA-dependent RNA polymerase (RdRp).¹⁸³ In early 2023 a study was published to show some resistance to Paxlovid in a few strains, including omicron and WA1, which demonstrates the need for continuous efforts for anti-viral drugs, as other resistant strains may appear in the future. This result also showcases the importance of having more than one target (i.e., targeting RdRp and PLpro, for instance).¹⁸⁴

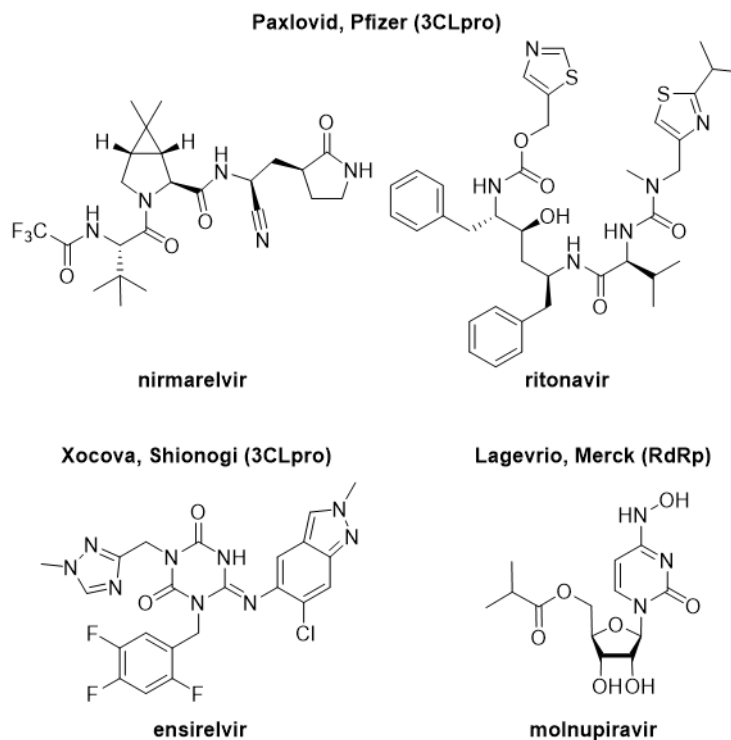


Figure 5.1. Antiviral drugs approved for use against COVID-19. On top- commercial name and manufacturer, in parentheses-what viral component targeted, below compounds- common name.

5.1.4. PLpro activity and published inhibitors

PLpro is a cysteine protease, it is one of the nonstructural proteins (nsps) encoded in the viral RNA (nsp3). It is necessary for viral replication by cleaving the viral polypeptide between nsp1-nsp2, nsp2-nsp3, and nsp3-nsp4, and is doing so by recognizing the motif LXGG (**Figure 5.2**) that is found in between them. PLpro also carries two other proteolytic activities; removal of ubiquitin (Ub) protein and of ISG15 (interferon-induced gene 15) protein, which is also referred to as Ubl (ubiquitin like protein). By doing so (deubiquitination and deISGylation), PLpro also interferes with the immune response upon viral infection.¹⁸⁵

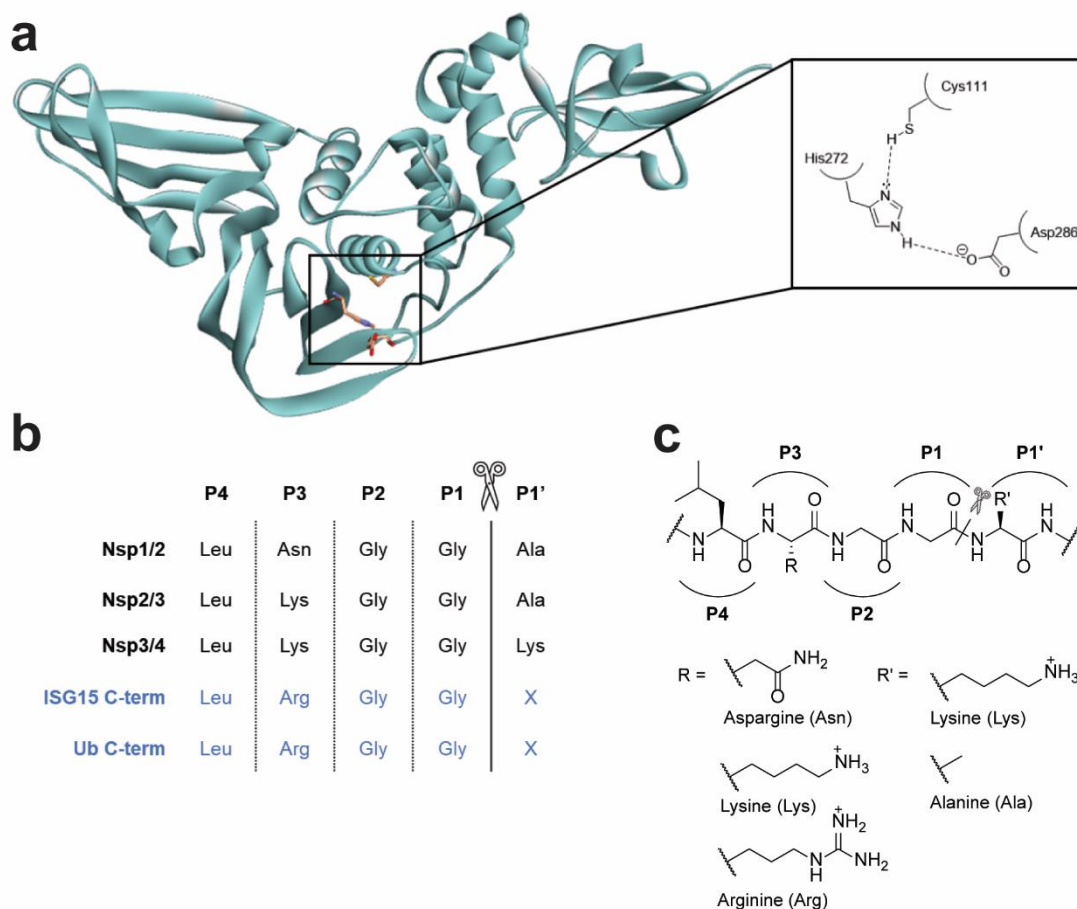


Figure 5.2. (a) Crystal structure of PLpro (PDB: 6WX4). Catalytic residues in the active site are colored pale orange and shown as sticks, as well as a 2D illustration on the right box. (b) Cleavage points of PLpro with the conserved motif LXGG. Nsp: nonstructural protein, C-term: carboxylic terminus. (c) 2D illustration of the conserved motif. Cleavage point is marked with dashed line and scissors.

As shown in **Figure 5.2a** PLpro's active site consists of a catalytic triad: cysteine, histidine, and aspartic acid. The catalytic triad's role is in activation of the serine (making it more nucleophilic) by charge relay through hydrogen bonds and enabling its reaction with a substrate amide bond.^{186, 187} A comprehensive investigation done by Rut *et al.*¹⁸⁵ on the selectivity of the active site revealed that PLpro has high specificity in the P2 position, and only glycine is accepted. Interestingly, the S3 pocket (P3 position interaction in the peptide) tolerates not only positively charged residues (such as Lys and Arg, **Figure 5.2c**), but also hydrophobic ones, thus displaying

low specificity in this position. The S4 pocket (conserved Leu position) only tolerates hydrophobic residues.

Based on these findings, Rut *et al.*¹⁸⁵ designed a covalent tetra pseudopeptide inhibitor for PLpro, **VIR251** (Figure 5.3). The pseudopeptide contains the glycine-like covalent group vinylmethyl ester (VME), which reacts with the nucleophilic cysteine at the β -position to form a covalent bond. They were able to crystallize the protein bound to the ligand, which shows PLpro's narrow tunnel-like shape of the active site. A closer look at this structure revealed that the glycine-VME (glycine-like) motif fits well into this tunnel, while also demonstrating why larger residues were not tolerated by the enzyme. **VIR251** and the detailed study carried out by Rut *et al.* provides an excellent starting point for the design of inhibitors.

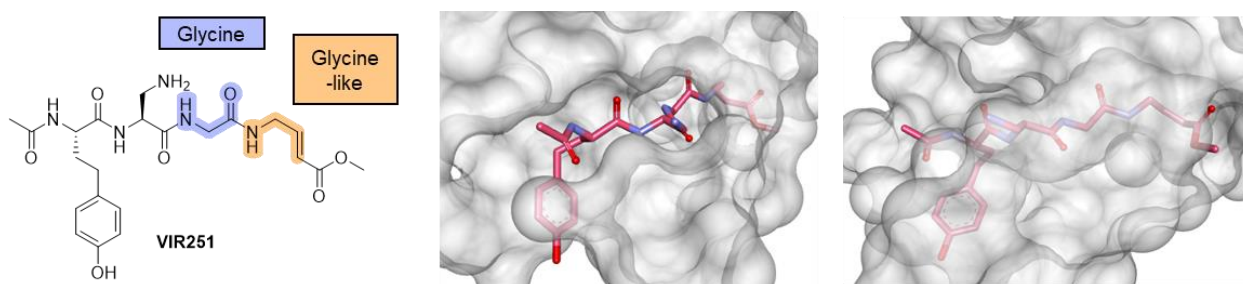


Figure 5.3. Covalent tetrapeptide ligand **VIR251** (left). On the right, **VIR251** (pink) co-crystallized with PLpro (shown as grey surface, PDB: 6WX4). The structure is shown from two perspectives, to demonstrate the entrance to the tunnel where the active site is, as well as a side view.

Based on the high sequence similarity between SARS-CoV and SARS-CoV-2 PLpro, Fu *et al.*¹⁸⁸ tested **GRL0617**, which had originally been designed for SARS-CoV PLpro in 2008,¹⁸⁹ on SARS-CoV-2 PLpro. **GRL0617** (Figure 5.4.) is a non-covalent inhibitor of PLpro, which has been shown to inhibit PLpro's deubiquitination as well as the deISGylating.¹⁸⁸ **GRL0617** has a promising IC₅₀ of 2.1 μ M, although further improvement is required to reach potency of interest for development into a therapeutic.¹⁹⁰

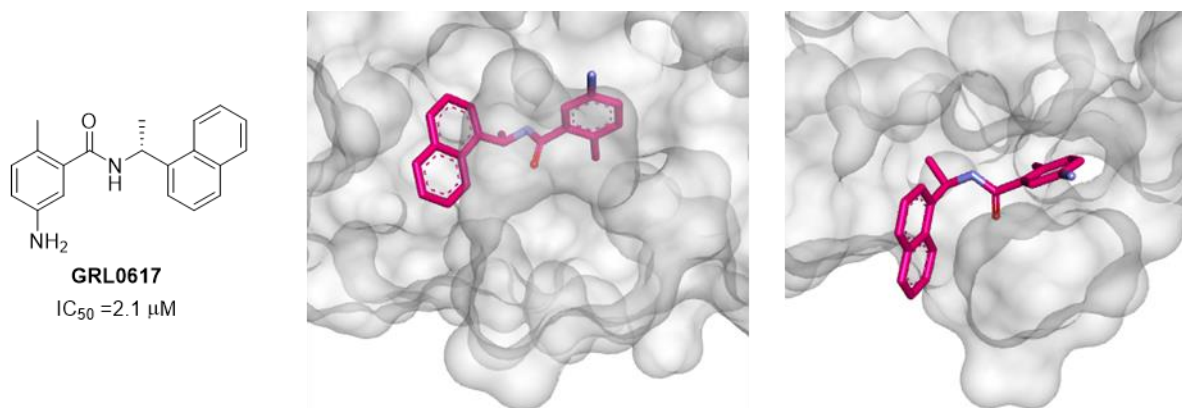


Figure 5.4. Non-covalent inhibitor **GRL0617** (left) with IC_{50} value underneath. On the right: co-crystal structure of **GRL0617** (pink) with SARS-CoV-2 PLpro (shown as a grey surface, PDB: 7JIR). The structure is shown from two different perspectives to demonstrate where the ligand interacts with the protein, as well as the entrance to the tunnel, where the active site is.

As shown in **Figure 5.4.**, **GRL0617** interacts with the enzyme at binding sites S3 and S4 (located at the entrance to the tunnel), which are wider. By doing that, the ligand blocks the entrance to the tunnel, hence blocking substrates from binding and accessing the active site. The positioning of **GRL0617** and **VIR251** in this area is very similar, and they both block the tunnel, while **VIR251** also bonds near the catalytic site, forming a covalent bond with the catalytic cysteine.

In contrast to 3CLpro, PLpro has proven to be a challenging target due to the shallow binding points, as well as the narrow tunnel where the active site resides. As mentioned above, the active site tunnel only tolerating Gly-Gly (or Gly-like) residues presents a challenge for the design of an inhibitor that is to interact with the active site residues.¹⁹¹ As a result, the success rate for finding new lead inhibitors is fairly low, and therefore **GRL0617** served as the lead for further work done by different groups.

For example, Shen *et al.*¹⁹² used **GRL0617** as the starting point, and investigated changing the different elements in the scaffold to increase interactions with the enzyme (**Figure 5.5a**). They designed the analogues with the aim of forming an interaction with Glu167 by adding a longer

chain with a H-bond donor on the aniline, as well as forming more hydrophobic interactions in the BL2 groove (the groove formed behind the BL2 loop, labeled in **Figure 5.5b**), by switching the naphthalene with different biaryl analogues. They hypothesize that adding an interaction with Glu167, which normally forms an interaction with Arg72 of ubiquitin would increase binding affinity (**Figure 5.5b**).¹⁹³ Their work led to the design of multiple inhibitors with improved inhibitory activity (IC_{50} of 0.113 μM for XR8-89 and 0.21 μM for **XR8-83**), compared to **GRL0617**. (**Figure 5.5**). The co-crystal of **XR8-83** (**Figure 5.5b**) reveals the important interactions formed between the enzyme and the ligand. As hypothesized, the longer chain with a H-bond donor indeed forms an interaction with Glu167.

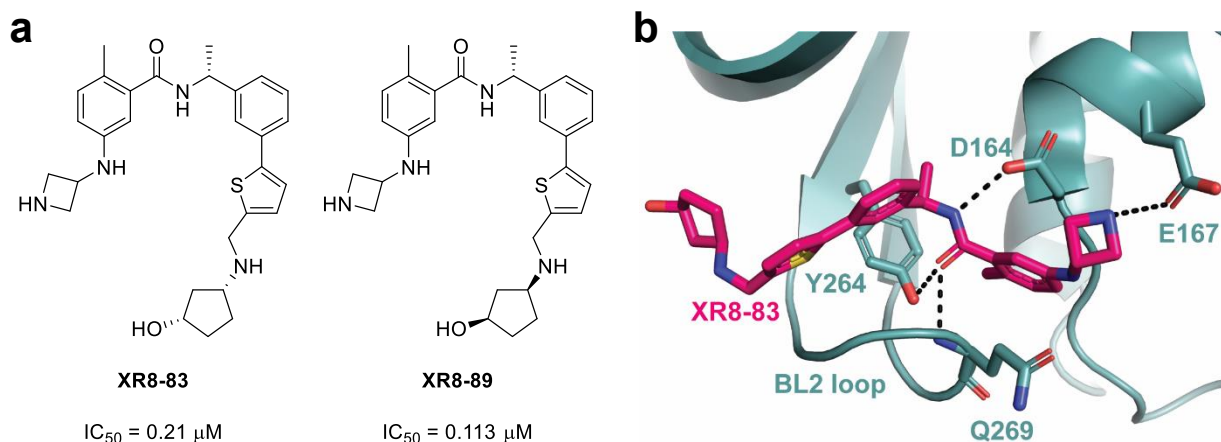


Figure 5.5. (a) The most active **GRL0617** analogues synthesized by Shen *et al.* with IC_{50} values underneath. (b) Compound **XR8-83** (pink) co-crystallized with PLpro (cyan, PDB: 7LLF). Residues that form interactions with the compound are shown in stick and labeled.

A different approach, albeit also based on **GRL0617** was taken by Sanders *et al.*¹⁹⁴ They chose to design a covalent inhibitor, which contains an electrophile, and would interact with the catalytic cysteine (**Figure 5.6a**). As a linker that would fit through the narrow tunnel into the active site *N,N'*-acetylacetohydrazine, which mimics the Gly-Gly in the natural substrate. Of the different covalent warheads tested, the fumarate ester proved to be the best one, **compound 7**, with an IC_{50}

of 0.094 μM (after a 30-minute incubation). Cell assays on Vero E6 cells infected with SARS-CoV-2 also provided promising results for the compound with EC_{50} of 1.1 μM and no cytotoxicity ($\text{CC}_{50} > 30 \mu\text{M}$). Interestingly, the analogue with NHAc *para* to the linker chain (**compound 8**) showed no significant cytoprotective effects, although having an IC_{50} of 0.230 μM , and appeared to be inactive in infected cells. **Compound 7** was co-crystallized with PLpro (PDB:7LLF), the important interactions the compound forms with the enzyme (H-bonds, and covalent bond with Cys111) can be seen in **Figure 5.6b**.

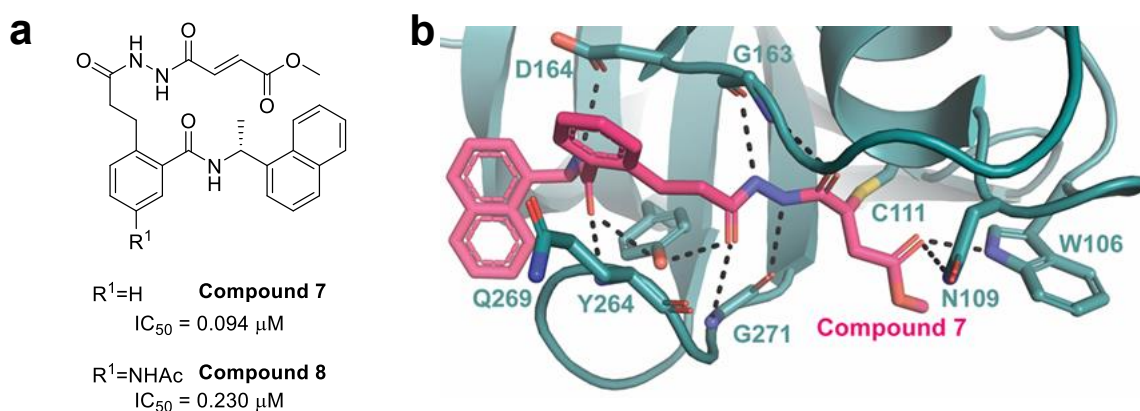


Figure 5.6. (a) Covalent analogues of **GRL0617** synthesized by Sanders *et al.* with IC_{50} values under each compound. (b) **Compound 7** (pink) co-crystallized with PLpro (cyan, PDB: 8EUA). Residues that form interactions with the compound are shown in stick and labeled.

5.2. Results

5.2.1. Design of a hybrid covalent inhibitor

Inspired by the studies on **VIR251** and **GRL0617**, we aimed to design a hybrid covalent inhibitor, in a strategy similar to that reported by Sanders *et al.*¹⁹⁴ while our work was in progress. The compound would contain a chain that resembles the Gly-Gly motif with a covalent warhead (vinyl alkyl ester, like the VME in **VIR251**) and the **GRL0617** scaffold. The crystal structure of **GRL0617** shows the methyl group pointing towards the active site, and so it could be replaced by

a longer linker that contains an electrophilic covalent warhead. The hybridinhibitor was designed to accommodate synthetic feasibility as well.

The first compound (**5.1**) was designed to retain as much of the original compounds (**GRL0617** and **VIR251**) as possible. The Gly-Gly motif was replaced with a 3-carbon chain connected to the covalent warhead via an amide bond. The group resembles glycine, it does not contain any substituents and is not bulky, which is necessary for it to reach the active site (**Figure 5.7c**).

The planned synthesis of **5.1** included synthesizing the covalent warhead and **GRL0617** part separately, and then connecting them via a Michael addition of the phenol to an acrylic ester (**Figure 5.7a**).

Before proceeding with the synthesis, we were interested in confirming the validity of the design and gaining insights into potential favourable interactions of our designed compound with PLpro. We therefore used the automated covalent docking software, FITTED, a docking program developed in our group.¹⁹⁵⁻¹⁹⁷

Docking of **5.1** to PLpro (PDB: 6WX4) revealed possible interactions of the compound within the residues around the active site (**Figure 5.7b**); residues Tyr273, Gly163, Gly271, His272, and Trp106 have the potential to form H-bonds with ligand **5.1**. Additionally, an overlay of **5.1** and **VIR251** (**Figure 5.7c, d**) in the active site, shows similarities in how both ligands occupy the space.

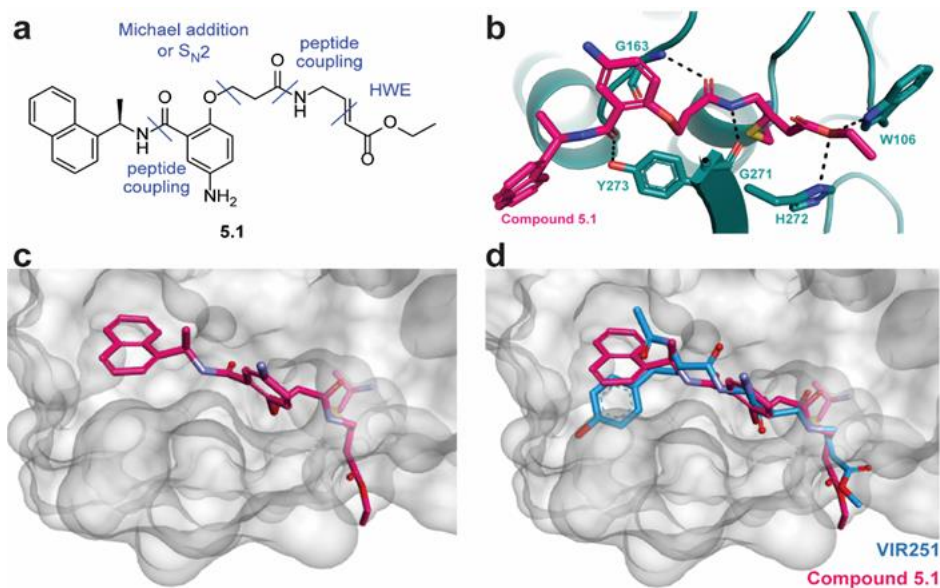
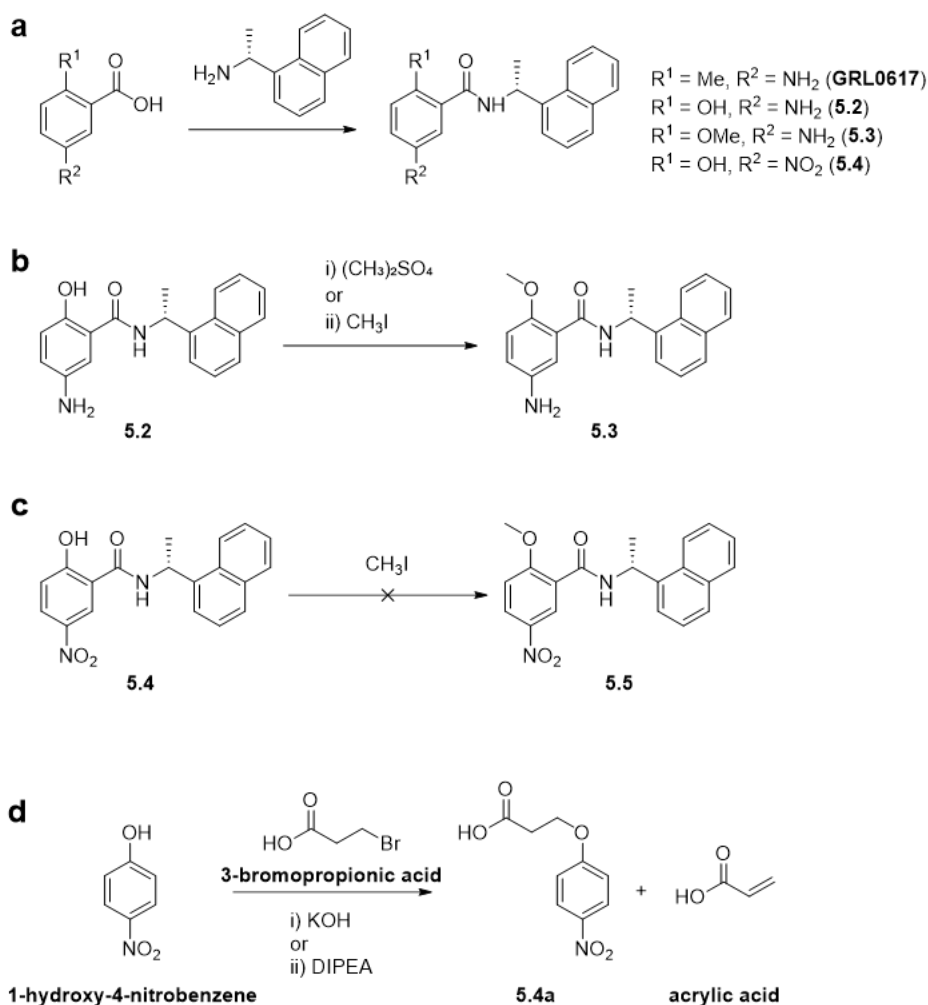


Figure 5.7. (a) Designed hybrid compound **5.1**. Blue dashed lines represent planned synthesis for the compound, reaction classes are written in blue. (b) Docking pose of **5.1** (pink) in PLpro's active site (teal), with potential interactions with residues (in stick, labeled). PDB: 6WX4. (c) Docked pose of **5.1** with PLpro shown as grey surface, indicating the ligand's fit in the narrow active site. (d) Overlay of the original co-crystallized ligand (**VIR251**, blue) with the docked pose of **5.1**, with PLpro as a grey surface, indicating similar space occupancy.

5.2.2. Synthesis and change in original design.

The first step in the synthesis was to prepare the **GRL0617**-like motif of the compound (**Scheme 5.1a**) via peptide coupling. To evaluate the impact of the substitution of the methyl carbon of **GRL0617** for an oxygen atom, the methyl ether analogue was also synthesized (**5.3**). Initially, the plan for any functionalization of the hydroxy group (including the methylation) was to use a weak base, and thus selectively alkylate the oxygen over the amine (or amide) of **5.2**. Multiple attempts to methylate the hydroxy group selectively, using methyl iodide or dimethyl sulfonate, resulted in a mixture of products, where the amide also got methylated (**Scheme 5.1b**). Although a free amine is more nucleophilic than a hydroxy group, we chose to use a weak base (Cs_2CO_3) to

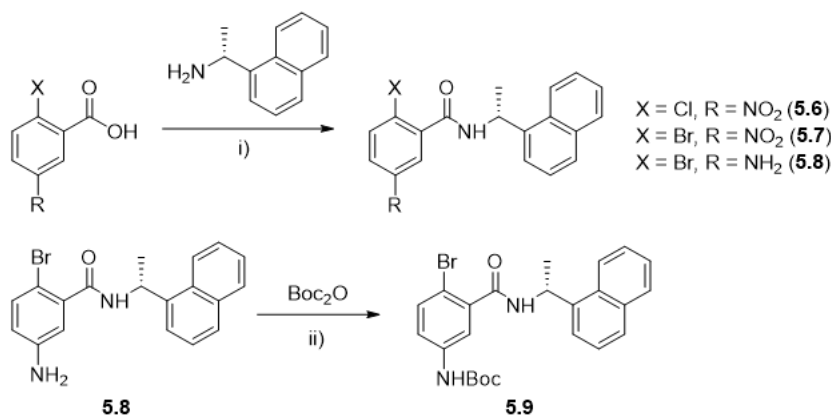
convert the latter into an alkoxide while leaving the amine unchanged. This chemoselective deprotonation would render the oxygen more nucleophilic than the amine, thus hopefully getting chemoselectivity for the methylation reaction. Unfortunately, the problem was not only the amine nitrogen, but also the amide, and we reasoned that a methylating agent might be too electrophilic. To prepare **5.3**, we used methyl ether benzoic acid as starting material and through an amide coupling obtained **5.3** (**Scheme 5.1a**). Considering the added difficulty of working with the *para*-amine group, due to both reactivity and stability (the starting material **5.2** oxidizes upon standing in room temperature over time), we decided to start with the nitro derivative of the starting material (**5.4**), with the objective to later reduce the nitro group into the amine. As we knew that the hydroxy group of this scaffold (**5.4**) would be considerably less reactive, we started by testing whether it could still be functionalized. A few test alkylation reactions on a simpler starting material (1-hydroxy-4-nitrobenzene) unfortunately yielded no product (**Scheme 5.1c, d**). Methylation with methyl-iodide resulted in only starting materials, while alkylation with 3-bromopropionic acid resulted also in a mix of starting materials, some alkylated product (**5.4a**, minor) and elimination product (acrylic acid).



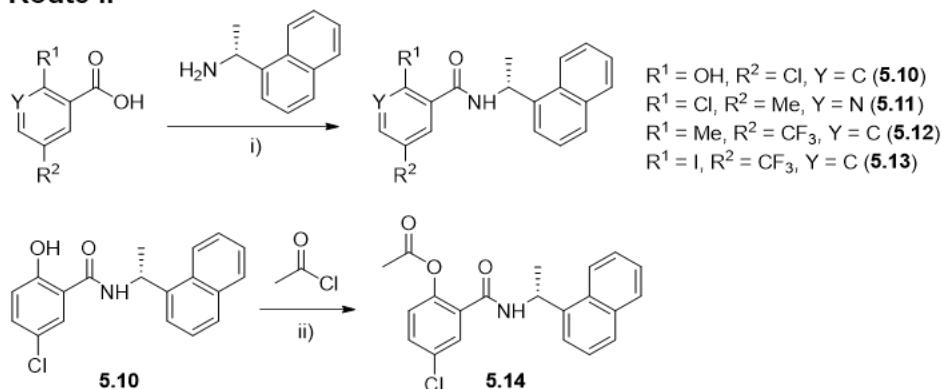
Scheme 5.1. (a) Synthesis of **GRL0617** analogues; PyBOP, DIPEA, DCM, rt, 24h, **GRL0617**, (74%), **5.2**, (51%), **5.3**, (73%); EDCI, HOBT·H₂O, DIPEA, DCM, rt, 24h, **5.4**, (43%). (b) Methylation attempts of the hydroxy group on **5.2**; i) (CH₃)₂SO₄, Cs₂CO₃, DMF, rt, 24h. ii) CH₃I, Cs₂CO₃, DMF, rt, 24h. Both conditions resulted in a mixture of methylation products. (c) Methylation attempt on the hydroxy group of the nitro derivative **5.4**; CH₃I, K₂CO₃, acetone, rt, 24h. Only starting material was observed by TLC. (d) Test alkylation on a simpler starting material, 4-nitrophenol; i) 3-bromopropanoic acid, KOH, H₂O, reflux for 5.5h, then rt for 19h. ii) 3-bromopropanoic acid, DIPEA, THF, MW, 120°C, 15 min. Both conditions resulted in a mixture of alkylated product as well as the elimination product, acrylic acid (almost 1:1).

At this point it was clear that the hydroxy is too unreactive in this scaffold (**5.4**) and we needed to consider a different alternative. We tried two routes in parallel (**Scheme 5.2**): **(I)** The reactivity of the starting material is flipped – instead of a hydroxy, the starting material would contain a halide that would later be reacted with a nucleophile in an S_NAr reaction. **(II)** Revisiting the crystal structure of **GRL0617** in PLpro, it appears that the aniline group does not make any specific interaction with the protein but is pointing away towards solvent (although it may interact with a network of water molecules in the binding site and orient the molecule in the relevant binding mode). Based on this observation, we hypothesized that the amine group could be replaced by another polar group (or a simple methyl group, as we tested with **5.11**, which led us to consider **5.10**). Additionally, while **5.1** was originally designed to have the covalent warhead group linked to the **GRL0617**-based scaffold via an ether group, other linking options were investigated (e.g., amine, C-C, ester). Referring to the docking result in **Figure 5.7b**, the ether oxygen would not necessarily form important interactions, and so we felt confident that diverting from the original plan to increase synthetic feasibility was also an option. Throughout the synthesis many challenges arose which led us to try different SMs that could lead to a slightly different product than **5.1** (e.g., changing the number carbon atoms in the Gly-like link). However, the core idea remains the same – form a hybrid of the scaffold of **GRL0617** with a covalent warhead linked with a Gly-like group, as in **VIR251**.

a. Route I



b. Route II



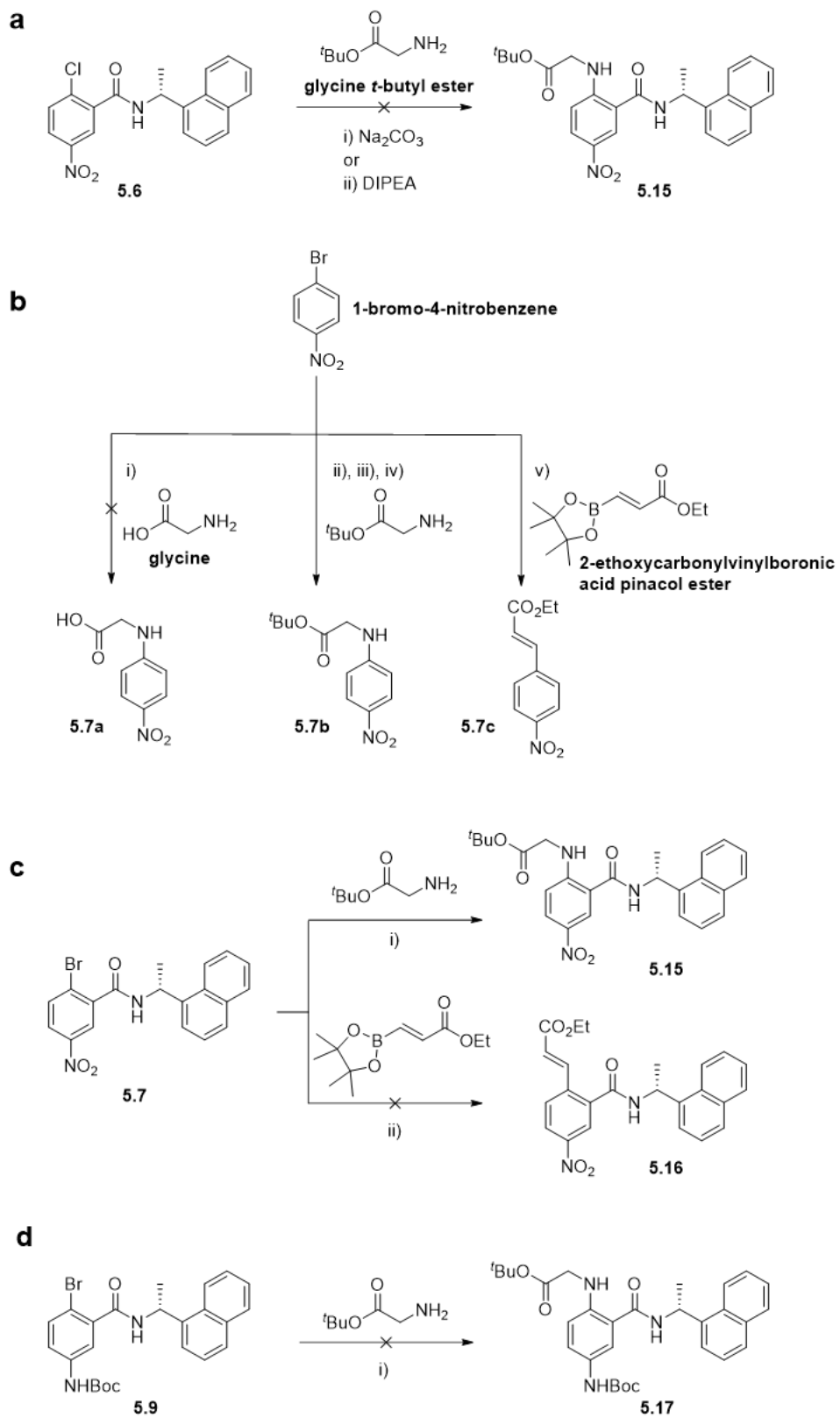
Scheme 5.2. Different routes for the synthesis of **GRL0617** analogs. **(a) Route I**, amine group is kept or masked as nitro, reactivity is flipped, a halide is used instead of the hydroxy. **i)** HATU, DIPEA, acetonitrile, rt, 24h, **5.6**, (71%), **5.7**, (80%); PyBOP, DIPEA, acetonitrile, rt, 24h, **5.8**, (91%). **ii)** Boc_2O , ethanol, rt, 24h, **5.9**, (48%). **(b) Route II**, amine group replaced by different substituents. **i)** HATU, DIPEA, acetonitrile or DCM rt, 24h, **5.10**, (DCM, 54%), **5.11**, (acetonitrile, 71%), **5.12**, (acetonitrile, 48%), **5.13**, (DCM, 87%). **ii)** acetyl chloride, Et_3N , rt, DCM, 21h, 61%, **5.14**.

5.2.2.1. Route I.

Intermediate **5.6** was first obtained by a HATU coupling reaction (**Scheme 5.2a**) then was subjected to $\text{S}_{\text{N}}\text{Ar}$ reaction conditions (**Scheme 5.3a**): using glycine *tert*-butyl ester with Na_2CO_3 in THF at reflux, as well as a microwave reaction with DIPEA in dioxane at 120°C . Unfortunately,

neither condition produced the desired product. In an attempt to increase the reactivity, we decided to work with an aryl-bromide, which is known as a better substrate for the amination reactions we were planning on conducting next.¹⁹⁸ Thus the synthesis of compound **5.7** was achieved through a HATU amide coupling reaction (**Scheme 5.2a**). The simplified substrate 1-bromo-4-nitrobenzene (**Scheme 5.3b**) was used as a model system to test different reaction conditions.

Two different Ullmann amination conditions were tested: (**Scheme 5.3b, i**) using glycine, CuI, and K₂CO₃ in DMF at 90°C. (**Scheme 5.3b, ii**) using glycine *tert*-butyl ester, L-proline as a ligand, CuI, and K₂CO₃ in DMF at 90°C.^{199, 200} Unfortunately, only starting material was recovered in both reactions with this model compound. The conditions were therefore not applied to intermediate **5.7**. A Buchwald-Hartwig amination was next attempted. Unfortunately, using glycine *tert*-butyl ester with Pd₂(dba)₃, BINAP, and Cs₂CO₃ in THF under reflux for 24h did not produce the desired product (**Scheme 5.3b, iii**). A second Buchwald-Hartwig amination with the same reagents on small scale (**Scheme 5.3b, iv**), in toluene and at 100°C for 24h was more promising as indicated by a new spot on TLC. These same reaction conditions were applied to **5.7** (**Scheme 5.3c, i**) which provided the desired product (**5.15**), albeit with low yield (13%), even after scale-up.

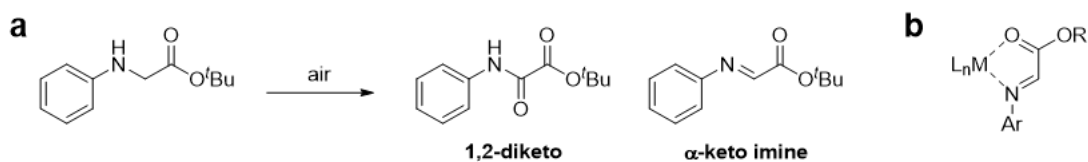


Scheme 5.3. (a) **i**) Na₂CO₃, THF, reflux, 24h. **ii**) DIPEA, dioxane, MW, 120°C, 1h, **5.15.** (b) **i**) CuI, K₂CO₃, DMF, 90°C, 48h, **5.7a.** **ii**) CuI, L-proline K₂CO₃, DMF, 90°C, 48h, **5.7b.** **iii**) Pd₂(dba)₃, BINAP, Cs₂CO₃, THF, reflux, 24h, **5.7b.** **iv**) Pd₂(dba)₃, BINAP, Cs₂CO₃, toluene, 100°C, 24h, **5.7b.** **v**) Pd(PPh₃)₄, K₂CO₃, DMF, 80°C, 20h, 27%, **5.7c.** (c) **i**) Pd₂(dba)₃, BINAP, Cs₂CO₃, toluene, 100°C, 24h, 23%, **5.15.** **ii**) Pd(PPh₃)₄, K₂CO₃, DMF, 80°C, 20h, **5.16.** (d) **i**) Pd₂(dba)₃, BINAP, KO^tBu, toluene, 100°C, 48h, **5.17.**

At this stage, we went back to the blackboard and further revised our original design considering these synthetic challenges. Another avenue we explored was using a Boc-protected aniline instead of the nitro derivatives (**5.9**). The advantage of this route is that the Boc protecting group could be removed from the aniline in the same step as the *tert*-butyl group of the glycine. Intermediate **5.9** was synthesized from aryl bromide **5.8** (**Scheme 5.2a**) which was protected with a Boc group.²⁰¹ The Boc protected SM (**5.9**) was then subjected to the same Buchwald-Hartwig amination conditions. Disappointingly, no product was obtained, and only starting materials were recovered. While various coupling conditions can be tested to optimize the reaction's yield (e.g., ligand, solvent, temperature, etc.), we decided instead to explore a different synthetic plan as we believed it was more promising. This is described below.

Exhaustive literature search into amination reactions revealed evidence about a few drawbacks in our current plan: first, it appeared that specifically glycine and glycine esters were poor substrates for the Buchwald-Hartwig amination, most likely due to the products' tendency to oxidize to α -keto imine, a π -system that can coordinate the metal and hamper the reaction (**Scheme 5.4**).^{202, 203} In addition, the coupled glycine product is also sensitive to oxidation into the 1,2-diketo derivatives, which can further decompose in the presence of oxygen and an organic solvent.²⁰²⁻²⁰⁴ This issue persists for the Ullmann amination as well since the product will be the same, and α -keto imines have been previously used as ligands in copper-catalyzed reactions, which suggests they can easily form *in situ* under reaction conditions.²⁰³ It became clear that even if we obtained

the coupled product, once we expose the *para*-aniline group (either by deprotection or by reduction), the 1,4-phenylenediamine derivative produced is very likely to be oxygen sensitive and decompose similarly to **5.2**.²⁰⁵



Scheme 5.4. (a) Potential oxidation products of the amination product when *t*-butyl glycinate is used. (b) Potential metal coordination of α -keto imine to a metal center.

With this information, we decided to make a final attempt in this route by coupling an alkyl group via the Suzuki-Miyaura cross coupling reaction. A test reaction of the boronic ester with 1-bromo-4-nitrobenzene under common conditions of Pd(PPh₃)₄ and K₃PO₄ in anhydrous DMF led to the expected product in 27% yield.²⁰⁶ Unfortunately, when the same conditions were applied to **5.7** no product was obtained, and only a mix of side products was recovered. The poor reactivity of the substrate bearing the ortho amide group compared to the simplified substrate used when testing seemed to be a recurring issue. Unfortunately, the issue could be the result of multiple parameters. It is well known that cross coupling reactions often need to be tweaked for a specific substrate. Since the reaction has many variables and optimization can be a long process, we decided to prioritize route II and potentially return to this one in the future.

5.2.2.2.Route II.

This route was worked on in parallel to **route I**, and so some of the reactions that we carried out in **route I**, and later decided to not continue with, will also be described here. In this route, the 5-NH₂ substituent was replaced by other groups. In addition to conclusions drawn by analyzing the crystal structure of PLpro with **GRL0617**, work published by Ma *et al.*²⁰⁷ provided further evidence that the amine in **GRL0617** analogues can be replaced by other groups, while maintaining activity (**Figure 5.8**).

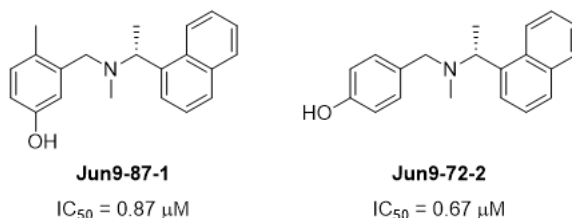
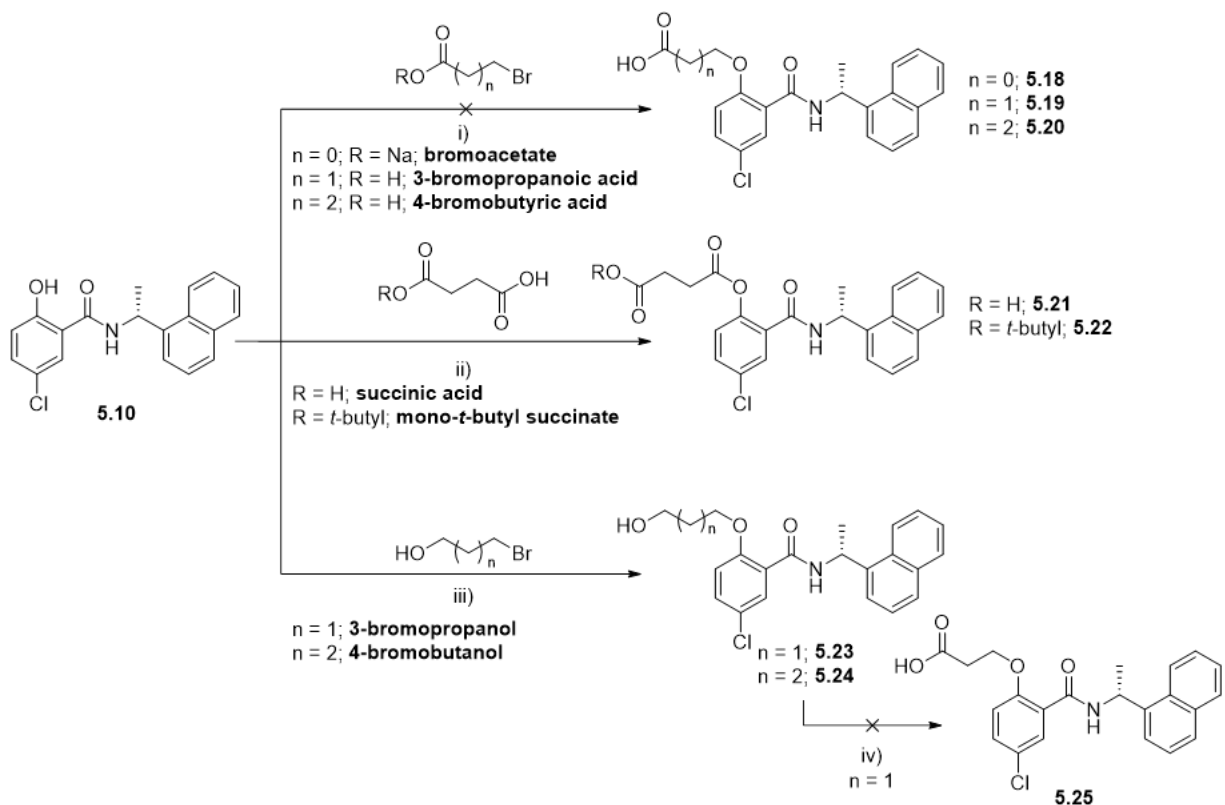


Figure 5.8. Example of **GRL0617** analogues designed by Ma *et al.* with inhibitory activity against PLpro, despite the amine group being replaced by a hydroxy (**Jun9-87-1**) or replaced by a hydroxy and moved from the meta position to the alkyl amine to para (**Jun9-72-2**).

The first scaffold, **5.10**, was used in different alkylation reactions, under the assumption that the hydroxyl group is not as deactivated as in **5.4**, but the amine group of **5.2** is not present and so would not interfere with the reaction. A simple acetylation reaction to give **5.14** worked well (**Scheme 5.2b**). However, all attempts to alkylate the hydroxy resulted in failure (**Scheme 5.5**). Initially, different alkyl bromides were used as electrophiles for an S_N2 reaction with the hydroxy of **5.10**. Using 3-bromopropionic acid (**5.19**) under different conditions resulted in recovery of starting material and some elimination product. Since the elimination product, acrylic acid, is a conjugated acid, we decided to try bromoacetate (to give product **5.18**), which would be unable to eliminate, and 4-bromobutyric acid (to give product **5.20**), which would not produce a conjugated elimination product. Unfortunately, both approaches turned out to also be unsuccessful, and only

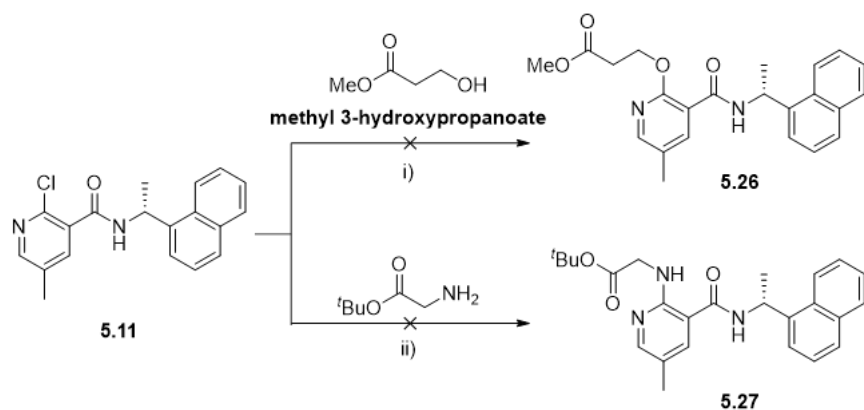
starting material was recovered (and some elimination product in the case of 4-bromobutyric acid). We hypothesized that the presence of the free carboxylic acid could be an issue, as it is more acidic than the hydroxy, so excess base is needed. The negatively charged carboxylate might then be less reactive, and in the case of 4-bromobutyric acid could also lead to cyclization forming γ -butyrolactone. The alcohol version of these groups was then considered (to form **5.23**, **5.24**).

Although the conversion was low, the product was obtained using a weak base (K_2CO_3), so selective deprotonation of the phenol was achieved. The next step was to oxidize the primary alcohol, so that the resulting carboxylic acid (**5.25**) could be used in a coupling with the covalent warhead (**5.30**). The oxidation reaction was unfortunately unsuccessful. In parallel, formation of an ester, instead of an ether, was also investigated (to form **5.21**, **5.22**). While no products were observed with succinic acid, using the mono-*tert*-butyl succinate provided some coupled product, although in low yield (10%). The ester analogue was not pursued further since the phenol, being a good leaving group, could cause the molecule to have undesired reactivity in this position. The ether scaffold was also not further pursued, as the hydroxy group proved to be unreactive. We also hypothesized that the elimination product recovered in some of the S_N2 reactions could have been formed after the nucleophilic substitution took place, since the phenol is a good leaving group. For these reasons, we decided to investigate other scaffolds, and consider a similar approach as in route I, where the reactivity between the coupling partners is switched.



Scheme 5.5. i) 5.18: NaH, DMF, rt, 24h; **5.19:** K₂CO₃, DMF, 60°C, 24h / NaI, acetone, rt, 5h / KOH, H₂O-THF (4:1), 80°C, 4h / K₂CO₃, DMF, MW, 120°C, 2h; **5.20:** K₂CO₃, 80°C, DMF, 24h / NaH, DMF, rt, 24h. **ii) 5.21:** DCC, DMAP, DMF, rt, 3h; **5.22:** EDCl, DMAP, DMF, rt, 24h (10%). **iii) 5.23:** K₂CO₃, DMF, rt, 24h (32%); **5.24:** K₂CO₃, DMF, rt, 24h (28%). **iv) 5.25:** TCCA, TEMPO, NaBr, acetone, rt, 24h.

Thus, aryl chloride **5.11** was reacted under S_NAr conditions with methyl 3-hydroxypropanoate as the nucleophile NaOMe or DBU as base and (Scheme 5.6) to form **5.26**. Unfortunately, only starting materials were recovered. At this stage, we switched to an amine nucleophile and worked initially with glycine *tert*-butyl ester (also for S_NAr reaction, Scheme 5.6, to form **5.27**). Both attempts were carried out in a microwave reactor. Using DIPEA in ethanol at 120°C or DBU in iso-propanol at 150 °C, both resulted in unreacted starting material.

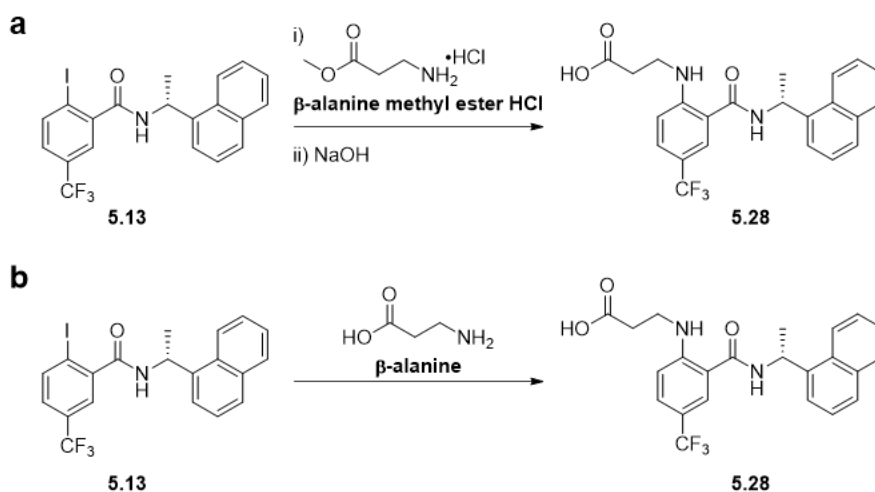


Scheme 5.6. i) 5.26: NaOMe (25 w% in methanol), dioxane, MW, 150 °C, 30 min, then additional 1.05 equiv. of methyl-3-hydroxypropanoate and 1.5 equiv. of NaOMe were added, 170 °C for 60 more min; DBU, dioxane, MW, 150 °C, 30 min. **ii) 5.27:** DIPEA, ethanol, MW, 120 °C, 90 min.; DBU, isopropanol, MW, 150 °C, 90 min.

With the previously observed problems with the aniline and the nitro group, we decided to investigate a substituent that would also make the final compound less synthetically challenging, while more drug-like. **5.13** was chosen, as iodine is known to have better reactivity in many cross-coupling reactions, as the C-I bond is weaker than other aryl halide bonds.^{198, 208} Trifluoromethyl was chosen as the *para*-substituent, as it is a well-established group in medicinal chemistry and an electron withdrawing group facilitating the reaction.²⁰⁹

We decided to revisit the Ullmann amination, using aryl iodide and β -alanine methyl ester instead of glycine (**Scheme 5.7a**). The procedure used is based on the work published by Ma *et al.*²⁰⁸ in which β -amino acid esters are used as both the ligand and the coupling partner. A small amount of water was added to the reaction to allow the hydrolysis of the ester in this one-pot reaction. The assumption is that the ester is first hydrolyzed, and the free amino carboxylate proceeds to react.²⁰⁸ However, when running the reaction, a mixture of hydrolyzed and unhydrolyzed coupled products were obtained, and an additional hydrolysis step with NaOH was needed. While the reaction did not proceed exactly as intended, this was the first time a coupled

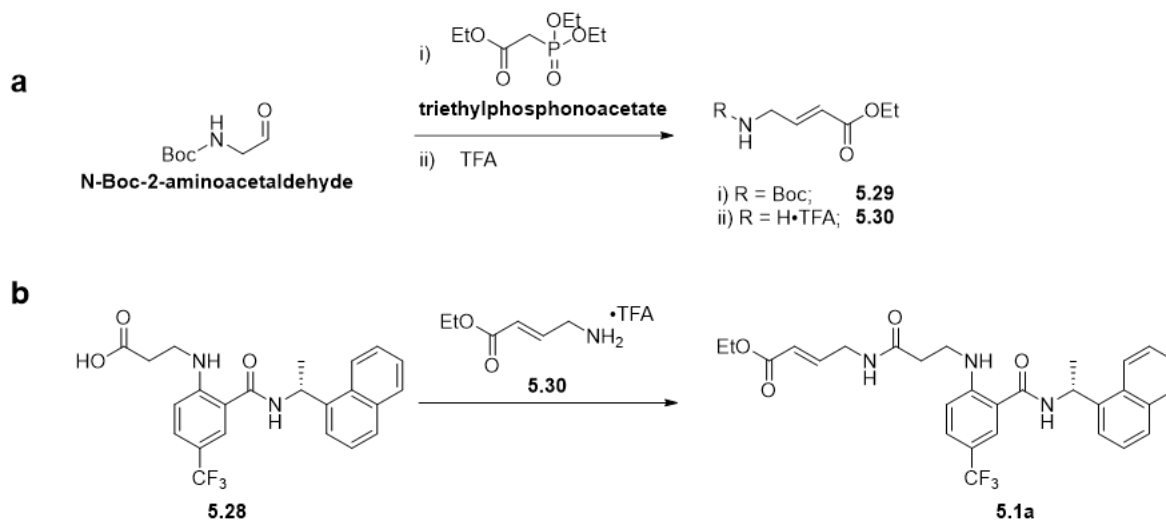
product was observed in reasonable amount. Since the ester is expected to be hydrolyzed, we decided to attempt the reaction under similar conditions with the free acid, without adding water and at a slightly lower temperature. Surprisingly, after 48h the reaction was not complete, and the major product was a side product (not characterized). The reaction was run again under modified conditions (**Scheme 5.7b**):²¹⁰ the base was switched to Cs₂CO₃ that should be more soluble in DMF and should reduce the issue of residue collecting above the solvent line, the reaction flask was purged, and the liquid mixture was degassed with argon. Running the reaction at 105°C under argon atmosphere resulted in the reaction being complete within 3h. Although the side product was still formed, the ratio was in favor of the desired product. Gratifyingly, the coupled, free acid was separated and isolated by FCC to finally give the coupled product **5.28**.



Scheme 5.7. (a) **i)** CuI, K₂CO₃, DMF-H₂O (14:1), 100°C, 20 h, **ii)** NaOH (1.6 M aq.), methanol, rt, 1 h, 36% (**5.28**) (b) CuI, Cs₂CO₃, DMF, 100°C, 3 h, 50% (**5.28**).

With the left portion of the final molecule in hand, we turned our attention to the other half. The covalent warhead was synthesized using published conditions for the Horner-Woodward-Emmons (HWE) reaction using *N*-Boc-2-aminoacetaldehyde, to form the Boc-protected product **5.29**.²¹¹ Next, Boc-deprotection was accomplished with TFA to provide the free amine salt in

quantitative yield (**Scheme 5.8a**). The two pieces of the target compound (**5.28** and **5.30**) were finally ready to be coupled to form the final product. This was achieved using PyBOP coupling, and after flash column chromatography, the final compound, **5.1a**, in 37% yield (**Scheme 5.8b**).

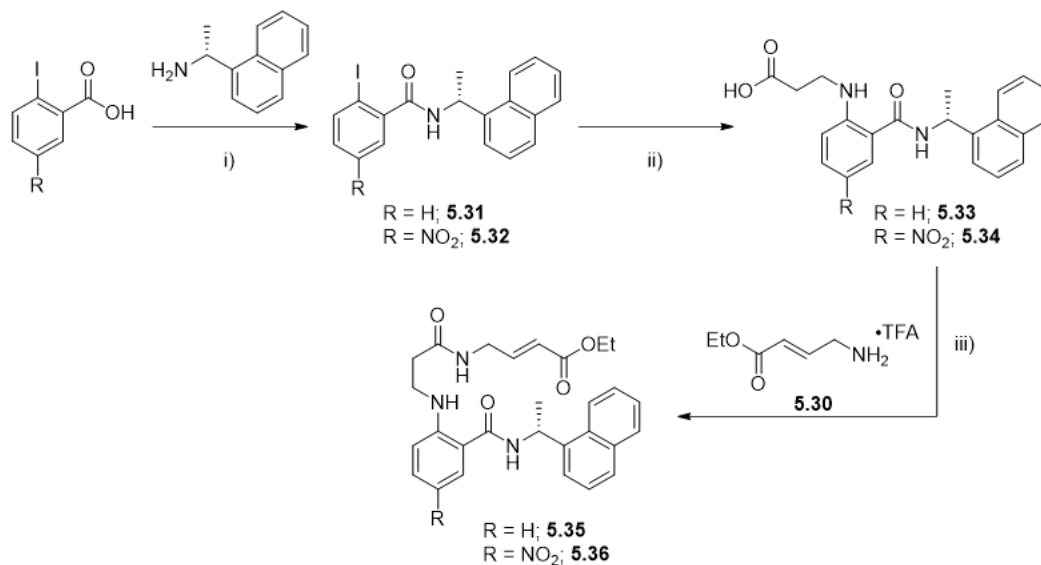


Scheme 5.8. (a) i) NaH, THF, 0°C to rt, 3.5 h, 60% (**5.29**); ii) TFA, DCM, 0°C, 2 h, quantitative yield (**5.30**). (b) DIPEA, PyBOP, DCM, 0°C to rt, 18 h, 37% (**5.1a**).

Once the chemistry was established and the synthesis of **5.1a** was completed, an undergraduate research assistant (Solène Huck) was able to synthesize analogues of the compound (**Scheme 5.9**).

The analogue with no *para* substituent on the benzene ring (**5.35**) was synthesized as a comparison to **compound 7** (**Figure 5.6**), and to learn about the influence different substituents have on the compounds' activity.

Compound **5.36** was initially synthesized with the hope to later reduce the nitro to the amine, as the amine is present in **GRL0617**. However, all reduction attempts led to decomposition or recovery of starting material, as the resulting 1,4-phenylenediamine derivative would likely be sensitive to oxidation (as was compound **5.2**). The nitro was therefore kept as it is.



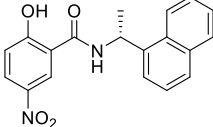
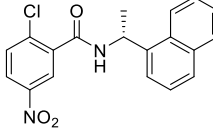
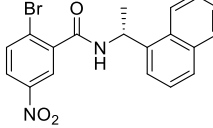
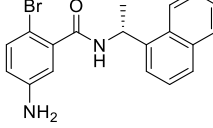
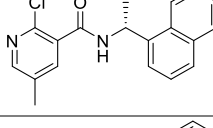
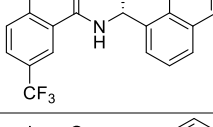
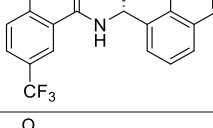
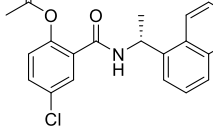
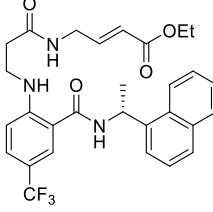
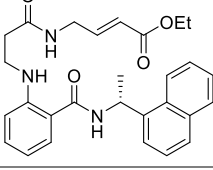
Scheme 5.9. Synthesis of **5.1a** analogues: **i)** HATU, DIPEA, acetonitrile, rt, 24 h, 40% (**5.31**), 78% (**5.32**). **ii)** CuI, Cs₂CO₃, DMF, 100 °C, 3 h, 58% (**5.33**), 52% (**5.34**). **iii)** DIPEA, PyBOP, DCM, 0 °C to rt, 18 h, 11% (**5.35**), 58% (**5.36**).

5.2.3. Biological activity evaluation

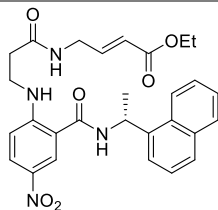
The compounds that were synthesized were assessed by Ho Ying Huang as PLpro inhibitors using a peptide (Z-RLRGG-AMC) as a fluorescent substrate (**Table 5.1**).

Table 5.1. Biological activity evaluation of compounds synthesized against PLpro.

Compound (50 μM)	Structure	Peptide (50 μM, % inhibition)
GRL0617		60%
5.2		< 5%

5.4		< 5%
5.6		12 ± 1
5.7		< 5%
5.8		60 ± 3
5.11		15 ± 3
5.12		< 5%
5.13		< 5%
5.14		12 ± 0
5.29	<chem>BocHNCC=CC(=O)OCC</chem>	< 5%
5.1a		80 ± 2
5.35		27 ± 6

5.36

 37 ± 1

When the methyl group of **GRL0617** was replaced by a hydroxy group (**5.2**) the activity was completely lost, however, when a bromine was used (**5.8**), 60% inhibition was measured. Although we do not know the exact reason as to this significant difference in reactivity, a few hypotheses can explain it: compound **5.2** is easily oxidizable, as we have observed the solution turns black over time at room temperature, which means that the concentration of **5.2** in solution over time was lower than expected.

Unlike the bromine and methyl, the hydroxy may participate better in electron donation into the aryl ring, which changed the electronic properties and may influence the conformation. Additionally, the hydroxy also makes the molecule more water soluble, therefore less likely to bind. Interestingly, all the non-covalent **GRL0617** analogues lost activity when the amine was replaced with another group, even though the role of the amine for activity is unclear from the crystal structure. Gratifyingly, despite the lack of activity of the **GRL0617** non-covalent analogue **5.12**, the hybrid compound **5.1a** with the covalent warhead displayed 80% inhibition. Unfortunately, the analogues of **5.1a** lost activity when the trifluoromethyl group was replaced by either a hydrogen (**5.35**) or a nitro (**5.36**), displaying 27 and 37% inhibition, respectively. The dose response curve of **5.1a** was consequently measured (**Figure 5.9**), and the IC_{50} was measured to be $13.3 \pm 9 \mu\text{M}$, similar to that of **GRL0617** ($5.8 \pm 1 \mu\text{M}$).

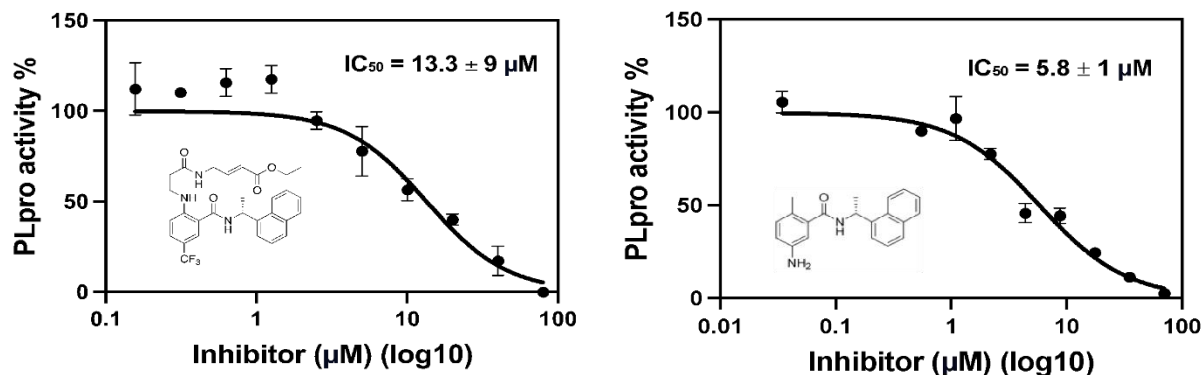


Figure 5.9. Dose response curves for **5.1a** (left) and **GRL0617** (right) against PLpro, with Z-RLRGG-AMC as substrate.

We were satisfied that the synthetic efforts led to an active compound, which is expected to be tested in cell assays. Based on the *in vitro* and cell assay results, more analogues may be synthesized and tested, potentially probing the role of the *para* substituent's inhibitory activity.

5.3. Conclusions and Future work

In this chapter a hybrid covalent inhibitor for SARS-CoV-2's PLpro was designed and synthesized. Although the original design had to be revisited and adapted as synthetic challenges arose, the original designed scaffold was achieved. The project began during the pandemic, as our group was working on targeting both 3CLpro and PLpro. As more literature was published, it became clear that targeting PLpro was more challenging than anticipated, as the active site is located within a narrow tunnel. As there were very few publications with potent compounds, it was a challenge to find a starting point for the design of a new potential inhibitor. The idea of combining the **GRL0617** scaffold with a covalent linker, similar to **VIR251**, was very appealing.

As the synthesis presented more challenges than we had anticipated, it took major efforts to make it to the last step and have the first hybrid molecule in hand. During this time, a few

publications came out that were also using **GRL0617** as a starting point, and different non-covalent analogues were made (as shown in the introduction). Recently, another group worked with a similar idea to ours and published a successful covalent inhibitor, **Compound 7**. This demonstrates the narrow starting point as well as the challenges in targeting PLpro.

Now that we have established a synthetic strategy, and two analogues were already synthesized and tested, further small changes to the scaffold may be made and tested for reactivity in a structure-activity relationship study. The NO₂ analogue was made (**5.36**), and the plan was to reduce it to the corresponding amine, with the aim of learning how the change in this group affects activity. Unfortunately, though not surprisingly, any attempt to reduce the nitro resulted in only what was likely decomposition of the unstable product. The analogue where the group was entirely removed (replaced by -H, **5.35**) was made in order to have a comparison to **compound 7** and any future analogue with different groups. For example, changes to the naphthalene side of the molecule can be carried out, and inspiration can be drawn from the work done on **XR8-89**. The “easiest” part of the molecule to introduce variance is the position of the -CF₃ group on the benzene ring. This is because different starting materials can be purchased (the same way **5.35** and **5.36** were designed and synthesized) containing different groups that can be further manipulated.

A closer look at the crystal structure of **XR8-83** may give some ideas to future analogues. The molecule has an amine 3 atoms away from the benzene ring that forms a H-bond with Glu167 (**Figure 5.5**).

Potential analogues that may be synthesized and tested may incorporate a similar element, such as the ones shown in **Figure 5.10**.

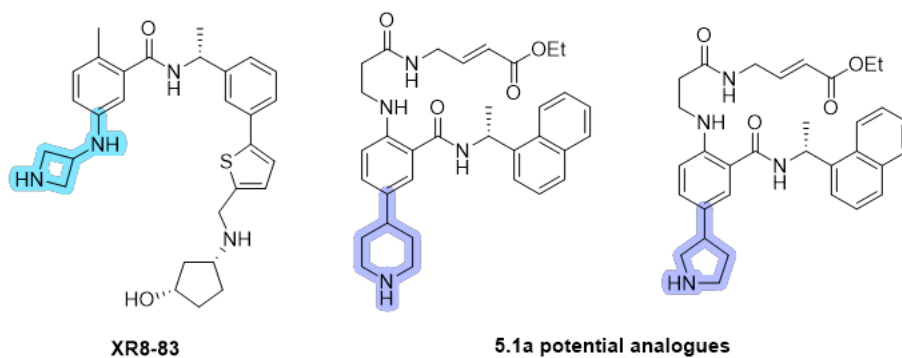


Figure 5.10. Compound **XR8-83** (left), and potential analogues of compound **5.1a** (right). The region in the compounds inspired by **XR8-83** is highlighted in blue in **XR8-83**. The region changed in the analogues compared to **5.1a** is highlighted in purple.

A limitation of this plan is the relatively lower availability of aryl iodides, which is necessary for the Ullmann coupling step, when compared to aryl bromides or chlorides, although it is not impossible to add steps to the synthesis and access further analogues.

Chapter 6 – Conclusions and Future Work

Asymmetric catalysis is an indispensable field in organic chemistry, providing an efficient route for the synthesis of chiral molecules with high enantioselectivities. The traditional way by which asymmetric catalysts are developed and optimized is unfortunately time consuming, expensive, and tedious. Advances in computational chemistry have enabled its integration into the field of medicinal chemistry and drug discovery, with docking and pharmacophore software used to screen potential ligands *in silico*, thus allowing only selected few to be synthesized. Similarly, the integration of computational methods for the prediction of enantioselectivity for the purpose of asymmetric catalysts design and optimization is becoming more prevalent. However, it still requires constant validation and improvement, as the accuracy required for the prediction of enantioselectivity is considerably higher than that required for ligand docking (~1 kcal/mol vs. ~5kcal/mol energy differences). Presently, no software exist that are available to be truly used by organic chemists, without prior computational chemistry knowledge, and very few such software are being developed. As a result, these are still seldomly applied by organic chemists, as the paradigm shift for the incorporation of such tools requires trusted validation and confidence by the chemists as well as user-friendliness.

In this thesis, we aimed to demonstrate the application of such software by an organic chemist, as well as to highlight the current challenges in the field (most importantly, the lack of a suitable library of chiral molecules for virtual screening, that is designed for the purpose of asymmetric catalysis rather than medicinal chemistry). During our work, we encountered unexpected challenges that led us to reassess what information and software development are still needed. We are optimistic that with further work demonstrating the usefulness and useability by

an organic chemist, the field will adopt computational methods more frequently, which in turn will also provide feedback to the computational chemists working on software development, as to points of improvement.

In **Chapter 1**, we provide an overview of existing computational methods for the predictions of enantioselectivity, their application, and their transferability to new reactions and/or catalysts. The information is provided from my perspective, as an organic chemist, with input from computational chemists. The goal is to demonstrate current applications of the methods to organic chemistry reactions, and to make the methods and terminology accessible and clear to organic chemists with no background in computational methods. We hope the review will inspire organic chemists to further incorporate predictive computational methods into their research.

In **Chapter 2** we sought to validate the accuracy of the VIRTUAL CHEMIST platform that has been developed in our group. We were also interested in the useability of the software by an organic chemist. Having a constant feedback loop between the computational chemists (developers) and organic chemists (users) enabled us to solve some issues in real time, while also taking note of aspects that required further attention. The initial hurdle we came across was the lack of a suitable library of chiral, enantiopure compounds that are either commercially available or synthesizable. The initial screen was done using the ZINC database, which is directed more towards medicinal chemistry. A collaboration with ChemSpace gave us access to chiral libraries that are commercially available. However, since these libraries are also aimed towards medicinal chemistry, we ran into problems of diversity (many compounds exist as a repetitive scaffold with different substituents), library size for some compounds that are less common in medicinal chemistry (chiral ketones), and price of the compounds that are sold in amounts for biological assays, and not amounts that would allow us to conduct multiple experiments, unfortunately. There is still a need for a publicly

available repository of chiral, enantiopure compounds that can be used for such endeavors. This is especially significant now. As software for virtually screening asymmetric catalysts/ligands become more powerful, they may become an inseparable part of asymmetric catalysts development and optimization projects. Another major impediment to this project was the lack of knowledge about the reactivity of potential catalysts. This issue came to light when we conducted the first virtual screening and were looking through the results for compounds that were synthesizable, and later discovered they were not active at all in catalyzing the Diels-Alder reaction. Later, when the ChemSpace library was used, we encountered diversity issues, and when looking to increase the library size we wanted to be sure only to include reactive amines. As a result of this hurdle, Dr. Mihai Burai-Patrascu from our group worked on incorporating a QM tool (QUEMIST) to the VIRTUAL CHEMIST suite, where different reactivity measures can be computed. Reactivity measures such as electron affinity and nucleophilicity can be computed and used for filtering compounds. However, an issue remained that we could not be certain of what exactly influences the reactivity of the amines in the Diels-Alder cycloaddition. Initially, we assumed it was related to poor nucleophilicity, then perhaps a combination of nucleophilicity and electron affinity, and eventually we concluded the issue is more complicated than perhaps it initially seemed, and we needed a deeper understanding of reactivity in this reaction. While some work has been published regarding the nuances that influence the reaction, some details of mechanistic intricacies were still missing. From there we decided to address this specific issue and learn about the reactivity of different molecules, in this case secondary amines.

This need for a better understanding of the reactivity observed in Chapter 2 led us to the mechanistic study of iminium catalyzed Diels-Alder cycloaddition in **Chapter 3**. In this chapter, we investigated different secondary amines as their hydrochloric acid salts and evaluated their

reactivity and rate in the formation of the iminium, as well as the reactivity and rate of the full cyclization reaction for a sample of the amines. We learned that pyrrolidine-based amines with electron withdrawing substituents (rendering the amine salt more acidic) form the iminium faster and catalyze the cycloaddition reaction faster.

While most piperidine-based amines were essentially completely unreactive in the formation of the iminium, we observed that for some (such as piperidine) the iminium formation is a kinetically slow process, which can take over 48 hours to reach equilibrium (compared to < 1h for the more reactive amines). This observation was in fact also the case for pyrrolidine, and it seems that the more electron rich pyrrolidines are most likely hindered by a very slow kinetic process. Based on the observation from product formation, we concluded that the cycloaddition is most likely the rate determining step, as about 10% iminium needs to form before the cycloaddition takes place at a reasonable rate.

Nonetheless, iminium formation rate also has an impact on the total rate of the reaction, albeit less so than the cycloaddition, as demonstrated by reaction progression simulation. In addition to the experimental studies, we also investigated the mechanistic details computationally. This turned out to be considerably more complicated than we had originally anticipated, and the results may lead to further experimental investigations.

One of the major challenges we encountered when working on the computational study was the iminium formation in a protic solution. It appeared computations suggested that the proton transfer from the amine to the oxygen in the hemiaminal in an intramolecular process was energetically disfavoured. Explicit water/methanol was necessary to facilitate this step. This led us to add a few water molecules to the system, hoping that the proton transfer would occur

independently. Despite investigating three different pathways, the proton transfer had to be done manually, which complicated the investigations and forced us to focus on a single pathway. Since this part of the project will help us understand the intricacies of our experimental results, it is still ongoing, though not completed within this thesis. Depending on the progress of the computations and the computational study's results, more experiments may need to be carried out to have the full picture of the process.

The conclusions from this project can be exploited to derive a reactivity trend that will be used in future virtual screenings. This project holds significance since our observations can certainly apply to more reactions beyond the Diels-Alder cycloaddition: iminium activation in asymmetric organocatalysis is an indispensable method to add nucleophiles to unsaturated aldehydes and ketones. Another important impact of our work, and part of the constant feedback loop in our group, is the development of a pKa predictor, to be incorporated into VIRTUAL CHEMIST.

Future work in this project is currently underway. The computational study is carried out and will reveal whether more experiments are needed. Once a 3D plot for the iminium formation is obtained, the same system for computing the mechanistic path may be used on other amines, such as pyrrolidine and trifluoromethyl pyrrolidine. Having the preferred pathway for iminium formation of each amine will enable the comparison between them and shed light on the reactivity differences. As mentioned above, a pKa predictor is being developed by our group, and once incorporated into VIRTUAL CHEMIST may be used as an indicator of the amine salt's acidity. Ideally, we would be able to define the reactivity with one or two parameters that can be used to filter un-/less reactive compounds. Once complete, a new virtual screening may take place, where not only the selectivity but also the reactivity of the amines will be computed as part of the screening and tested experimentally for selected amines.

An important factor when carrying out published reactions for the purpose of validation is the ability to reproduce published data. In Chapters 2 and 3 we had to make sure that we could reproduce the results of the Diels-Alder cycloaddition using proline methyl ester as the catalyst. Similarly, in **Chapter 4** we were interested in using the Shi Epoxidation as another validation tool for the software (as described in Chapter 2) and encountered the problems to reproduce the published procedure. Since the reaction contains many variables, it took many attempts to find a procedure that worked for us and was reproducible. This revealed the need to have more detailed procedures reported, as some apparently insignificant technical details for the chemists developing a given reaction may turn out to be critical for reproducibility. This optimized procedure was used for the rest of the studies of this reaction.

Another issue that has been mentioned in Chapter 2 was the lack of a suitable chiral library for the screening of chiral ketones. Since these scaffolds are less common in medicinal chemistry, when we were interested in doing a virtual screening, we came up with the solution of augmenting the library by also including chiral secondary alcohols that have been oxidized into ketones *in silico* using our software REACT2D. Although this should be a good solution for increasing library size in future screenings as well, the issue remained that the library lacked diversity. Similar to the conclusions that led to Chapter 3, here we also decided that we were missing experimental information about structure-activity relationship of different ketones, with the conditions used by Shi and by us. To know what general scaffolds to keep in our library (like the iminium investigation), we decided to test simple ketone scaffolds for the reactivity. The study here is of a different nature to that described in Chapter 3 since the reaction is carried out in a biphasic solution. With the oxidant and base added throughout the reaction's time frame, it is not as simple to gather information by using NMR studies. Nonetheless, the studies carried out in chapter 4 revealed that

6-membered cyclic ketones are considerably more reactive than other ketones tested and should be the only ones included in the library.

With optimized reaction conditions and a better understanding of reactivity, everything is in place for the next stage, a new virtual screening. Although this reaction can be viewed as more specific as the ketone is not used as a catalyst for other transformations, the asymmetric installation of epoxides, without the use of toxic metals, is nonetheless an important transformation, especially in medicinal chemistry and natural product synthesis.

Future work in this chapter may include a new virtual screening, including only hexanone-based scaffolds. As the issue remains, with not many chiral ketones being available, other reactions may be considered to increase the library size using REACT2D. Reactions such as ozonolysis of methylenecyclohexane derivatives and acetal hydrolysis can be used to incorporate more ketones in the library. Additionally, as a potential collaboration with Liverpool Chiral Chem is underway, their resources may be used to obtain a larger library size.

Ultimately, virtual screenings for both the iminium catalyzed Diels-Alder cycloaddition and the chiral ketone catalyzed Shi epoxidation reactions will be done with the new knowledge and inputs. With this, not only can the useability of the VIRTUAL CHEMIST platform be reevaluated, as intended, but also our investigation as to reactivity can be tested. Commercially available and/or synthesizable chiral molecules can be tested for their selectivity and reactivity.

Chapter 5 introduces a new topic, as the pandemic hit and we went into lockdown early 2020, we decided to use our lab's expertise in medicinal chemistry and join the efforts of targeting SARS-CoV-2. The chapter focuses on the design and synthesis of a potential PLpro inhibitor, the many synthetic challenges that we faced during this project, redesign of our target based on more

published information and synthetic feasibility, and finally the synthesis of a covalent inhibitor with encouraging inhibitory activity against PLpro.

Future work on this project is still ongoing in our group. With an established synthetic route, the project is ready for the synthesis of additional analogues, with hopefully improved activity against PLpro. As of the writing of this thesis, projects targeting SARS-CoV-2, and specifically PLpro are still being targeted by our group. A summer intern was working on synthesizing analogues for the molecule described in Chapter 5, and a PhD student is working on other scaffolds targeting the enzyme as well. Future students may continue work on the scaffold described in this chapter, to further improve the potency. The compounds were also sent for cell assays to establish their inhibitory activity in live cells. Based on the results of the assays, more analogues can be designed using docking. The most promising ones may be synthesized and tested, hopefully showing inhibitory activity. This would potentially lead to a potent inhibitor for PLpro.

References

- (1) Moitessier, N.; Pottel, J.; Therrien, E.; Englebienne, P.; Liu, Z.; Tomberg, A.; Corbeil, C. R. Medicinal Chemistry Projects Requiring Imaginative Structure-Based Drug Design Methods. *Accounts of Chemical Research* **2016**, *49* (9), 1646-1657. DOI: 10.1021/acs.accounts.6b00185.
- (2) Shen, Y.; Borowski, J. E.; Hardy, M. A.; Sarpong, R.; Doyle, A. G.; Cernak, T. Automation and computer-assisted planning for chemical synthesis. *Nature Reviews Methods Primers* **2021**, *1* (1). DOI: 10.1038/s43586-021-00022-5.
- (3) Schwaller, P.; Probst, D.; Vaucher, A. C.; Nair, V. H.; Kreutter, D.; Laino, T.; Reymond, J.-L. Mapping the space of chemical reactions using attention-based neural networks. *Nature machine intelligence* **2021**, *3* (2), 144-152.
- (4) Coley, C. W.; Green, W. H.; Jensen, K. F. Machine Learning in Computer-Aided Synthesis Planning. *Accounts of Chemical Research* **2018**, *51* (5), 1281-1289. DOI: 10.1021/acs.accounts.8b00087.
- (5) Molga, K.; Szymkuć, S.; Grzybowski, B. A. Chemist Ex Machina: Advanced Synthesis Planning by Computers. *Accounts of Chemical Research* **2021**, *54* (5), 1094-1106. DOI: 10.1021/acs.accounts.0c00714.
- (6) Schwaller, P.; Petraglia, R.; Zullo, V.; Nair, V. H.; Haeuselmann, R. A.; Pisoni, R.; Bekas, C.; Iuliano, A.; Laino, T. Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. *Chemical Science* **2020**, *11* (12), 3316-3325, DOI: 10.1039/C9SC05704H.
- (7) Klucznik, T.; Mikulak-Klucznik, B.; McCormack, M. P.; Lima, H.; Szymkuć, S.; Bhowmick, M.; Molga, K.; Zhou, Y.; Rickershauser, L.; Gajewska, E. P.; et al. Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory. *Chem* **2018**, *4* (3), 522-532. DOI: <https://doi.org/10.1016/j.chempr.2018.02.002>.
- (8) Mikulak-Klucznik, B.; Gołębiowska, P.; Bayly, A. A.; Popik, O.; Klucznik, T.; Szymkuć, S.; Gajewska, E. P.; Dittwald, P.; Staszewska-Krajewska, O.; Beker, W.; et al. Computational planning of the synthesis of complex natural products. *Nature* **2020**, *588* (7836), 83-88. DOI: 10.1038/s41586-020-2855-y.
- (9) Genheden, S. E., O.; Bjerrum, E. J. A Quick Policy to Filter Reactions Based on Feasibility in AI-Guided Retrosynthetic Planning. *ChemRxiv. This content is a preprint and has not been peer-reviewed.* **2020**. DOI: <https://doi.org/10.26434/chemrxiv.13280495.v1>.
- (10) Żurański, A. M.; Martinez Alvarado, J. I.; Shields, B. J.; Doyle, A. G. Predicting Reaction Yields via Supervised Learning. *Accounts of Chemical Research* **2021**, *54* (8), 1856-1865. DOI: 10.1021/acs.accounts.0c00770.
- (11) Schwaller, P.; Vaucher, A. C.; Laino, T.; Reymond, J.-L. Prediction of chemical reaction yields using deep learning. *Machine Learning: Science and Technology* **2021**, *2* (1), 015016. DOI: 10.1088/2632-2153/abc81d.

References

- (12) Anstine, D. M.; Isayev, O. Generative Models as an Emerging Paradigm in the Chemical Sciences. *Journal of the American Chemical Society* **2023**, *145* (16), 8736-8750. DOI: 10.1021/jacs.2c13467.
- (13) Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. Predicting reaction performance in C–N cross-coupling using machine learning. *Science* **2018**, *360* (6385), 186-190. DOI: 10.1126/science.aar5169.
- (14) Guan, Y.; Coley, C. W.; Wu, H.; Ranasinghe, D.; Heid, E.; Struble, T. J.; Pattanaik, L.; Green, W. H.; Jensen, K. F. Regio-selectivity prediction with a machine-learned reaction representation and on-the-fly quantum mechanical descriptors. *Chemical Science* **2021**, *12* (6), 2198-2208, DOI: 10.1039/D0SC04823B.
- (15) Jorner, K.; Tomberg, A.; Bauer, C.; Sköld, C.; Norrby, P.-O. Organic reactivity from mechanism to machine learning. *Nature Reviews Chemistry* **2021**, *5* (4), 240-255. DOI: 10.1038/s41570-021-00260-x.
- (16) Zhao, S.; Gensch, T.; Murray, B.; Niemeyer, Z. L.; Sigman, M. S.; Biscoe, M. R. Enantiodivergent Pd-catalyzed C–C bond formation enabled through ligand parameterization. *Science* **2018**, *362* (6415), 670-674. DOI: 10.1126/science.aat229.
- (17) Reid, J. P.; Sigman, M. S. Holistic prediction of enantioselectivity in asymmetric catalysis. *Nature* **2019**, *571* (7765), 343-348. DOI: 10.1038/s41586-019-1384-z.
- (18) Reid, J. P.; Sigman, M. S. Comparing quantitative prediction methods for the discovery of small-molecule chiral catalysts. *Nature Reviews Chemistry* **2018**, *2* (10), 290-305. DOI: 10.1038/s41570-018-0040-8.
- (19) Zahrt, A. F.; Henle, J. J.; Rose, B. T.; Wang, Y.; Darrow, W. T.; Denmark, S. E. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **2019**, *363* (6424), DOI: 10.1126/science.aau5631.
- (20) Gensch, T.; dos Passos Gomes, G.; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; et al. A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *Journal of the American Chemical Society* **2022**, *144* (3), 1205-1217. DOI: 10.1021/jacs.1c09718.
- (21) Bell, E. L.; Finnigan, W.; France, S. P.; Green, A. P.; Hayes, M. A.; Hepworth, L. J.; Lovelock, S. L.; Niikura, H.; Osuna, S.; Romero, E.; et al. Biocatalysis. *Nature Reviews Methods Primers* **2021**, *1* (1), 46. DOI: 10.1038/s43586-021-00044-z.
- (22) Pyser, J. B.; Chakrabarty, S.; Romero, E. O.; Narayan, A. R. H. State-of-the-Art Biocatalysis. *ACS Central Science* **2021**, *7* (7), 1105-1116. DOI: 10.1021/acscentsci.1c00273.
- (23) Finnigan, W.; Hepworth, L. J.; Flitsch, S. L.; Turner, N. J. RetroBioCat as a computer-aided synthesis planning tool for biocatalytic reactions and cascades. *Nature Catalysis* **2021**, *4* (2), 98-104. DOI: 10.1038/s41929-020-00556-z.
- (24) Kirkpatrick, P.; Ellis, C. Chemical space. *Nature* **2004**, *432* (7019), 823-823. DOI: 10.1038/432823a.
- (25) Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Von Bargen, C. D.; et al. OPLS4: Improving Force Field Accuracy on Challenging

References

- Regimes of Chemical Space. *Journal of Chemical Theory and Computation* **2021**, *17* (7), 4291-4300. DOI: 10.1021/acs.jctc.1c00302.
- (26) Qiu, Y.; Smith, D. G. A.; Boothroyd, S.; Jang, H.; Hahn, D. F.; Wagner, J.; Bannan, C. C.; Gokey, T.; Lim, V. T.; Stern, C. D.; et al. Development and Benchmarking of Open Force Field v1.0.0—the Parsley Small-Molecule Force Field. *Journal of Chemical Theory and Computation* **2021**, *17* (10), 6262-6280. DOI: 10.1021/acs.jctc.1c00571.
- (27) Wei, W.; Champion, C.; Barigye, S. J.; Liu, Z.; Labute, P.; Moitessier, N. Use of Extended-Hückel Descriptors for Rapid and Accurate Predictions of Conjugated Torsional Energy Barriers. *Journal of Chemical Information and Modeling* **2020**, *60* (7), 3534-3545. DOI: 10.1021/acs.jcim.0c00440.
- (28) Neese, F. The SHARK integral generation and digestion system. *Journal of Computational Chemistry* **2023**, *44* (3), 381-396. DOI: <https://doi.org/10.1002/jcc.26942>.
- (29) Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling* **2007**, *13* (12), 1173-1213. DOI: 10.1007/s00894-007-0233-4.
- (30) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation* **2019**, *15* (3), 1652-1671. DOI: 10.1021/acs.jctc.8b01176.
- (31) Gallarati, S.; van Gerwen, P.; Laplaza, R.; Vela, S.; Fabrizio, A.; Corminboeuf, C. OSCAR: an extensive repository of chemically and functionally diverse organocatalysts. *Chemical Science* **2022**, *13* (46), 13782-13794, DOI: 10.1039/D2SC04251G.
- (32) Burai Patrascu, M.; Pottel, J.; Pinus, S.; Bezanson, M.; Norrby, P. O.; Moitessier, N. From desktop to benchtop with automated computational workflows for computer-aided design in asymmetric catalysis. *Nat. Catal.* **2020**, *3* (7), 574-584, DOI: 10.1038/s41929-020-0468-3 Scopus.
- (33) Gallegos, L. C.; Luchini, G.; St. John, P. C.; Kim, S.; Paton, R. S. Importance of Engineered and Learned Molecular Representations in Predicting Organic Reactivity, Selectivity, and Chemical Properties. *Accounts of Chemical Research* **2021**, *54* (4), 827-836. DOI: 10.1021/acs.accounts.0c00745.
- (34) Brethomé, A. V.; Fletcher, S. P.; Paton, R. S. Conformational Effects on Physical-Organic Descriptors: The Case of Sterimol Steric Parameters. *ACS Catalysis* **2019**, *9* (3), 2313-2323. DOI: 10.1021/acscatal.8b04043.
- (35) Warshel, A.; Weiss, R. M. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *Journal of the American Chemical Society* **1980**, *102* (20), 6218-6226.
- (36) Kim, Y.; Corchado, J. C.; Villà, J.; Xing, J.; Truhlar, D. G. Multiconfiguration molecular mechanics algorithm for potential energy surfaces of chemical reactions. *The Journal of Chemical Physics* **2000**, *112* (6), 2718-2735. DOI: 10.1063/1.480846.
- (37) Jensen, F. Locating minima on seams of intersecting potential energy surfaces. An application to transition structure modeling. *Journal of the American Chemical Society* **1992**, *114* (5), 1596-1603.

References

- (38) Pairault, N.; Zhu, H.; Jansen, D.; Huber, A.; Daniliuc, C. G.; Grimme, S.; Niemeyer, J. Heterobifunctional Rotaxanes for Asymmetric Catalysis. *Angewandte Chemie International Edition* **2020**, *59* (13), 5102-5107. DOI: <https://doi.org/10.1002/anie.201913781>.
- (39) Minenkov, Y.; Sharapa, D. I.; Cavallo, L. Application of Semiempirical Methods to Transition Metal Complexes: Fast Results but Hard-to-Predict Accuracy. *Journal of Chemical Theory and Computation* **2018**, *14* (7), 3428-3439. DOI: 10.1021/acs.jctc.8b00018.
- (40) Gallarati, S.; Laplaza, R.; Corminboeuf, C. Harvesting the fragment-based nature of bifunctional organocatalysts to enhance their activity. *Organic Chemistry Frontiers* **2022**, *9* (15), 4041-4051, DOI: 10.1039/D2QO00550F.
- (41) Harper, K. C.; Bess, E. N.; Sigman, M. S. Multidimensional steric parameters in the analysis of asymmetric catalytic reactions. *Nature Chemistry* **2012**, *4* (5), 366-374. DOI: 10.1038/nchem.1297.
- (42) Verloop, A.; Tipker, J. Use of linear free energy related and other parameters in the study of fungicidal selectivity. *Pesticide Science* **1976**, *7* (4), 379-390. DOI: <https://doi.org/10.1002/ps.2780070410>.
- (43) Zahrt, A. F.; Denmark, S. E. Evaluating continuous chirality measure as a 3D descriptor in chemoinformatics applied to asymmetric catalysis. *Tetrahedron* **2019**, *75* (13), 1841-1851. DOI: <https://doi.org/10.1016/j.tet.2019.02.007>.
- (44) Falivene, L.; Credendino, R.; Poater, A.; Petta, A.; Serra, L.; Oliva, R.; Scarano, V.; Cavallo, L. SambVca 2. A Web Tool for Analyzing Catalytic Pockets with Topographic Steric Maps. *Organometallics* **2016**, *35* (13), 2286-2293. DOI: 10.1021/acs.organomet.6b00371.
- (45) Metsänen, T. T.; Lexa, K. W.; Santiago, C. B.; Chung, C. K.; Xu, Y.; Liu, Z.; Humphrey, G. R.; Ruck, R. T.; Sherer, E. C.; Sigman, M. S. Combining traditional 2D and modern physical organic-derived descriptors to predict enhanced enantioselectivity for the key aza-Michael conjugate addition in the synthesis of Prevymis™ (Ietermovir). *Chemical Science* **2018**, *9* (34), 6922-6927, DOI: 10.1039/C8SC02089B.
- (46) Maji, R.; Mallojjala, S. C.; Wheeler, S. E. Electrostatic Interactions in Asymmetric Organocatalysis. *Accounts of Chemical Research* **2023**, *56* (14), 1990-2000. DOI: 10.1021/acs.accounts.3c00198.
- (47) Wang, S.; Jiang, J. Interpretable Catalysis Models Using Machine Learning with Spectroscopic Descriptors. *ACS Catalysis* **2023**, *13* (11), 7428-7436. DOI: 10.1021/acscatal.3c00611.
- (48) Liu, S. *Conceptual density functional theory: Towards a new chemical reactivity theory*; John Wiley & Sons, 2022.
- (49) Yap, C. W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry* **2011**, *32* (7), 1466-1474. DOI: <https://doi.org/10.1002/jcc.21707>.
- (50) *Open-Source Cheminformatics Software*. <https://www.rdkit.org/>.
- (51) See, X. Y.; Wen, X.; Wheeler, T. A.; Klein, C. K.; Goodpaster, J. D.; Reiner, B. R.; Tonks, I. A. Iterative Supervised Principal Component Analysis Driven Ligand Design for Regioselective

References

- Ti-Catalyzed Pyrrole Synthesis. *ACS Catalysis* **2020**, *10* (22), 13504-13517. DOI: 10.1021/acscatal.0c03939.
- (52) Betinol, I. O.; Lai, J.; Thakur, S.; Reid, J. P. A Data-Driven Workflow for Assigning and Predicting Generality in Asymmetric Catalysis. *Journal of the American Chemical Society* **2023**, *145* (23), 12870-12883. DOI: 10.1021/jacs.3c03989.
- (53) Pudjihartono, N.; Fadason, T.; Kempa-Liehr, A. W.; O'Sullivan, J. M. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Frontiers in Bioinformatics* **2022**, *2*, Review. DOI: 10.3389/fbinf.2022.927312.
- (54) Bro, R.; Smilde, A. K. Principal component analysis. *Analytical Methods* **2014**, *6* (9), 2812-2831, DOI: 10.1039/C3AY41907J.
- (55) Werth, J.; Sigman, M. S. Connecting and Analyzing Enantioselective Bifunctional Hydrogen Bond Donor Catalysis Using Data Science Tools. *Journal of the American Chemical Society* **2020**, *142* (38), 16382-16391. DOI: 10.1021/jacs.0c06905.
- (56) Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin, V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; et al. Virtual Computational Chemistry Laboratory – Design and Description. *Journal of Computer-Aided Molecular Design* **2005**, *19* (6), 453-463. DOI: 10.1007/s10822-005-8694-y.
- (57) Moriwaki, H.; Tian, Y.-S.; Kawashita, N.; Takagi, T. Mordred: a molecular descriptor calculator. *Journal of Cheminformatics* **2018**, *10* (1), 4. DOI: 10.1186/s13321-018-0258-y.
- (58) Yamaguchi, S.; Nishimura, T.; Hibe, Y.; Nagai, M.; Sato, H.; Johnston, I. Regularized regression analysis of digitized molecular structures in organic reactions for quantification of steric effects. *Journal of Computational Chemistry* **2017**, *38* (21), 1825-1833. DOI: <https://doi.org/10.1002/jcc.24791>.
- (59) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society* **1988**, *110* (18), 5959-5967. DOI: 10.1021/ja00226a005.
- (60) Lipkowitz, K. B.; Kozlowski, M. C. Understanding Stereoinduction in Catalysis via Computer: New Tools for Asymmetric Synthesis. *Synlett* **2003**, *10*, 1547-1565.
- (61) Zahrt, A. F.; Athavale, S. V.; Denmark, S. E. Quantitative Structure-Selectivity Relationships in Enantioselective Catalysis: Past, Present, and Future. *Chem Rev* **2020**, *120* (3), 1620-1689. DOI: 10.1021/acs.chemrev.9b00425.
- (62) Hopfinger, A. J.; Wang, S.; Tokarski, J. S.; Jin, B.; Albuquerque, M.; Madhav, P. J.; Duraiswami, C. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism. *Journal of the American Chemical Society* **1997**, *119* (43), 10509-10524. DOI: 10.1021/ja9718937.
- (63) Williams, W. L.; Zeng, L.; Gensch, T.; Sigman, M. S.; Doyle, A. G.; Anslyn, E. V. The Evolution of Data-Driven Modeling in Organic Chemistry. *ACS Central Science* **2021**, *7* (10), 1622-1637. DOI: 10.1021/acscentsci.1c00535.
- (64) Sigman, M. S.; Harper, K. C.; Bess, E. N.; Milo, A. The Development of Multidimensional Analysis Tools for Asymmetric Catalysis and Beyond. *Accounts of Chemical Research* **2016**, *49* (6), 1292-1301. DOI: 10.1021/acs.accounts.6b00194.

References

- (65) Zahrt, A. F.; Henle, J. J.; Denmark, S. E. Cautionary Guidelines for Machine Learning Studies with Combinatorial Datasets. *ACS Combinatorial Science* **2020**, *22* (11), 586-591. DOI: 10.1021/acscombsci.0c00118.
- (66) Crawford, J. M.; Kingston, C.; Toste, F. D.; Sigman, M. S. Data Science Meets Physical Organic Chemistry. *Accounts of Chemical Research* **2021**, *54* (16), 3136-3148. DOI: 10.1021/acs.accounts.1c00285.
- (67) Santiago, C. B.; Guo, J.-Y.; Sigman, M. S. Predictive and mechanistic multivariate linear regression models for reaction development. *Chemical Science* **2018**, *9* (9), 2398-2412, DOI: 10.1039/C7SC04679K.
- (68) van Dijk, L.; Haas, B. C.; Lim, N.-K.; Clagg, K.; Dotson, J. J.; Treacy, S. M.; Piechowicz, K. A.; Roytman, V. A.; Zhang, H.; Toste, F. D.; et al. Data Science-Enabled Palladium-Catalyzed Enantioselective Aryl-Carbonylation of Sulfonimidamides. *Journal of the American Chemical Society* **2023**, *145* (38), 20959-20967. DOI: 10.1021/jacs.3c06674.
- (69) Dotson, J. J.; van Dijk, L.; Timmerman, J. C.; Grosslight, S.; Walroth, R. C.; Gosselin, F.; Püntener, K.; Mack, K. A.; Sigman, M. S. Data-Driven Multi-Objective Optimization Tactics for Catalytic Asymmetric Reactions Using Bisphosphine Ligands. *Journal of the American Chemical Society* **2023**, *145* (1), 110-121. DOI: 10.1021/jacs.2c08513.
- (70) Xu, J.; Grosslight, S.; Mack, K. A.; Nguyen, S. C.; Clagg, K.; Lim, N.-K.; Timmerman, J. C.; Shen, J.; White, N. A.; Sirois, L. E.; et al. Atroposelective Negishi Coupling Optimization Guided by Multivariate Linear Regression Analysis: Asymmetric Synthesis of KRAS G12C Covalent Inhibitor GDC-6036. *Journal of the American Chemical Society* **2022**. DOI: 10.1021/jacs.2c09917.
- (71) Newman-Stonebraker, S. H.; Smith, S. R.; Borowski, J. E.; Peters, E.; Gensch, T.; Johnson, H. C.; Sigman, M. S.; Doyle, A. G. Univariate classification of phosphine ligation state and reactivity in cross-coupling catalysis. *Science* **2021**, *374* (6565), 301-308. DOI: 10.1126/science.abj4213.
- (72) Betinol, I. O.; Kuang, Y.; Reid, J. P. Guiding Target Synthesis with Statistical Modeling Tools: A Case Study in Organocatalysis. *Organic Letters* **2022**, *24* (7), 1429-1433. DOI: 10.1021/acs.orglett.1c04134.
- (73) Shoja, A.; Zhai, J.; Reid, J. P. Comprehensive Stereochemical Models for Selectivity Prediction in Diverse Chiral Phosphate-Catalyzed Reaction Space. *ACS Catalysis* **2021**, *11* (19), 11897-11905. DOI: 10.1021/acscatal.1c03520.
- (74) Liu, X. H.; Song, H. Y.; Ma, X. H.; Lear, M. J.; Chen, Y. Z. Virtual screening prediction of new potential organocatalysts for direct aldol reactions. *Journal of Molecular Catalysis A: Chemical* **2010**, *319* (1), 114-118. DOI: 10.1016/j.molcata.2009.12.008.
- (75) Henle, J. J.; Zahrt, A. F.; Rose, B. T.; Darrow, W. T.; Wang, Y.; Denmark, S. E. Development of a Computer-Guided Workflow for Catalyst Optimization. Descriptor Validation, Subset Selection, and Training Set Analysis. *Journal of the American Chemical Society* **2020**, *142* (26), 11578-11592. DOI: 10.1021/jacs.0c04715.

References

- (76) Rinehart, N. I.; Zahrt, A. F.; Denmark, S. E. Leveraging Machine Learning for Enantioselective Catalysis: From Dream to Reality. *Chimia (Aarau)* **2021**, *75* (7), 592-597. DOI: 10.2533/chimia.2021.592.
- (77) Rinehart, N. I.; Zahrt, A. F.; Henle, J. J.; Denmark, S. E. Dreams, False Starts, Dead Ends, and Redemption: A Chronicle of the Evolution of a Chemoinformatic Workflow for the Optimization of Enantioselective Catalysts. *Accounts of Chemical Research* **2021**, *54* (9), 2041-2054. DOI: 10.1021/acs.accounts.0c00826.
- (78) Zahrt, A. F.; Rose, B. T.; Darrow, W. T.; Henle, J. J.; Denmark, S. E. Computational methods for training set selection and error assessment applied to catalyst design: guidelines for deciding which reactions to run first and which to run next. *Reaction Chemistry & Engineering* **2021**, *6* (4), 694-708, DOI: 10.1039/D1RE00013F.
- (79) Lipkowitz, K. B.; Pradhan, M. Computational Studies of Chiral Catalysts: A Comparative Molecular Field Analysis of an Asymmetric Diels–Alder Reaction with Catalysts Containing Bisoxazoline or Phosphinooxazoline Ligands. *The Journal of Organic Chemistry* **2003**, *68* (12), 4648-4656. DOI: 10.1021/jo0267697.
- (80) Dalmau, D. A. R., J. V. ROBERT: Bridging the Gap between Machine Learning and Chemistry. *ChemRxiv* 2023. *This content is a preprint and has not been peer-reviewed.* **2023**.
- (81) Peng, Q.; Duarte, F.; Paton, R. S. Computing organic stereoselectivity - from concepts to quantitative calculations and predictions. *Chem Soc Rev* **2016**, *45* (22), 6093-6107. DOI: 10.1039/c6cs00573j.
- (82) Pottel, J.; Moitessier, N. Efficient Transition State Modeling Using Molecular Mechanics Force Fields for the Everyday Chemist. In *Reviews in Computational Chemistry*, Reviews in Computational Chemistry, 2016; pp 152-185.
- (83) Maloney, M. P.; Stenfors, B. A.; Helquist, P.; Norrby, P.-O.; Wiest, O. Interplay of Computation and Experiment in Enantioselective Catalysis: Rationalization, Prediction, and—Correction? *ACS Catalysis* **2023**, *13* (21), 14285-14299. DOI: 10.1021/acscatal.3c03921.
- (84) Bolitho, Elizabeth M.; Coverdale, J. P. C.; Wolny, J. A.; Schünemann, V.; Sadler, P. J. Density functional theory investigation of Ru(ii) and Os(ii) asymmetric transfer hydrogenation catalysts. *Faraday Discussions* **2022**, *234* (0), 264-283, DOI: 10.1039/D1FD00075F.
- (85) Weill, N.; Corbeil, C. R.; De Schutter, J. W.; Moitessier, N. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions: Development of the molecular mechanics-based program ACE and application to asymmetric epoxidation reactions. *Journal of Computational Chemistry* **2011**, *32* (13), 2878-2889, DOI: <https://doi.org/10.1002/jcc.21869>.
- (86) Verdolino, V.; Forbes, A.; Helquist, P.; Norrby, P.-O.; Wiest, O. On the mechanism of the rhodium catalyzed acrylamide hydrogenation. *Journal of Molecular Catalysis A: Chemical* **2010**, *324* (1), 9-14. DOI: <https://doi.org/10.1016/j.molcata.2010.02.026>.
- (87) Ingman, V. M.; Schaefer, A. J.; Andreola, L. R.; Wheeler, S. E. QChASM: Quantum chemistry automation and structure manipulation. *WIREs Computational Molecular Science* **2021**, *11* (4), DOI: <https://doi.org/10.1002/wcms.1510>.

References

- (88) Guan, Y.; Ingman, V. M.; Rooks, B. J.; Wheeler, S. E. AARON: An Automated Reaction Optimizer for New Catalysts. *Journal of Chemical Theory and Computation* **2018**, *14* (10), 5249-5261. DOI: 10.1021/acs.jctc.8b00578.
- (89) Steiner, M. R., Markus. Navigating chemical reaction space with a steering wheel. *arXiv. This article is a preprint and has not been peer reviewed* **2023**. DOI: <https://doi.org/10.48550/arXiv.2308.16499>.
- (90) *Gaussian 16 Rev. C.01*; Wallingford, CT, 2016.
- (91) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Remigio, R. D.; Alenaizan, A.; et al. Psi4 1.4: Open-source software for high-throughput quantum chemistry. *The Journal of Chemical Physics* **2020**, *152* (18). DOI: 10.1063/5.0006002.
- (92) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *The Journal of Chemical Physics* **2020**, *152* (22), 224108. DOI: 10.1063/5.0004608.
- (93) Pettersen, E. F.; Goddard Td Fau - Huang, C. C.; Huang Cc Fau - Couch, G. S.; Couch Gs Fau - Greenblatt, D. M.; Greenblatt Dm Fau - Meng, E. C.; Meng Ec Fau - Ferrin, T. E.; Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. (0192-8651 (Print)). From 2004 Oct.
- (94) Rosales, A. R. Development of Force Field Parameters for Predictions in Asymmetric Catalysis and a Foray into Mechanism Elucidation. Ph.D., University of Notre Dame, Ann Arbor, 2019.
- (95) Hansen, E.; Rosales, A. R.; Tutkowski, B.; Norrby, P.-O.; Wiest, O. Prediction of Stereochemistry using Q2MM. *Accounts of Chemical Research* **2016**, *49* (5), 996-1005. DOI: 10.1021/acs.accounts.6b00037.
- (96) Rosales, A. R.; Quinn, T. R.; Wahlers, J.; Tomberg, A.; Zhang, X.; Helquist, P.; Wiest, O.; Norrby, P.-O. Application of Q2MM to predictions in stereoselective synthesis. *Chemical Communications* **2018**, *54* (60), 8294-8311, DOI: 10.1039/C8CC03695K.
- (97) Kozlowski, M. C.; Dixon, S. L.; Panda, M.; Lauri, G. Quantum Mechanical Models Correlating Structure with Selectivity: Predicting the Enantioselectivity of β -Amino Alcohol Catalysts in Aldehyde Alkylation. *Journal of the American Chemical Society* **2003**, *125* (22), 6614-6615. DOI: 10.1021/ja0293195.
- (98) Rosales, A. R.; Wahlers, J.; Limé, E.; Meadows, R. E.; Leslie, K. W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R. H.; et al. Rapid virtual screening of enantioselective catalysts using CatVS. *Nature Catalysis* **2018**, *2* (1), 41-45. DOI: 10.1038/s41929-018-0193-3.
- (99) Terrett, J. A.; Chen, H.; Shore, D. G.; Villemure, E.; Larouche-Gauthier, R.; Déry, M.; Beaumier, F.; Constantineau-Forget, L.; Grand-Maître, C.; Lépissier, L.; et al. Tetrahydrofuran-Based Transient Receptor Potential Ankyrin 1 (TRPA1) Antagonists: Ligand-Based Discovery, Activity in a Rodent Asthma Model, and Mechanism-of-Action via Cryogenic Electron Microscopy. *Journal of Medicinal Chemistry* **2021**, *64* (7), 3843-3869. DOI: 10.1021/acs.jmedchem.0c02023.

References

- (100) Ingle, G. K.; Mormino Mg Fau - Wojtas, L.; Wojtas L Fau - Antilla, J. C.; Antilla, J. C. Chiral phosphoric acid-catalyzed addition of thiols to N-acyl imines: access to chiral N,S-acetals. (1523-7052 (Electronic)). From 2011 Sep 16.
- (101) Gerosa, G. G.; Spanevello, R. A.; Suárez, A. G.; Sarotti, A. M. Joint Experimental, in Silico, and NMR Studies toward the Rational Design of Iminium-Based Organocatalyst Derived from Renewable Sources. *J Org Chem* **2015**, *80* (15), 7626-7634. DOI: 10.1021/acs.joc.5b01214.
- (102) Kuang, Y.; Lai, J.; Reid, J. P. Transferrable selectivity profiles enable prediction in synergistic catalyst space. *Chemical Science* **2023**, *14* (7), 1885-1895, 10.1039/D2SC05974F. DOI: 10.1039/D2SC05974F.
- (103) Isbrandt, E. S.; Sullivan, R. J.; Newman, S. G. High Throughput Strategies for the Discovery and Optimization of Catalytic Reactions. *Angewandte Chemie International Edition* **2019**, *58* (22), 7180-7191. DOI: <https://doi.org/10.1002/anie.201812534>.
- (104) Brandt, J. R.; Salerno, F.; Fuchter, M. J. The added value of small-molecule chirality in technological applications. *Nature Reviews Chemistry* **2017**, *1* (6), 0045. DOI: 10.1038/s41570-017-0045.
- (105) MacKenzie, L. E.; Stachelek, P. The twists and turns of chiral chemistry. *Nature Chemistry* **2021**, *13* (6), 521-522. DOI: 10.1038/s41557-021-00729-8.
- (106) Wade, J.; Hilfiker, J. N.; Brandt, J. R.; Liirò-Peluso, L.; Wan, L.; Shi, X.; Salerno, F.; Ryan, S. T. J.; Schöche, S.; Arteaga, O.; et al. Natural optical activity as the origin of the large chiroptical properties in π -conjugated polymer thin films. *Nature Communications* **2020**, *11* (1), 6137. DOI: 10.1038/s41467-020-19951-y.
- (107) Jeschke, P. Current status of chirality in agrochemicals. *Pest Management Science* **2018**, *74* (11), 2389-2404. DOI: <https://doi.org/10.1002/ps.5052>.
- (108) Evans, A. M. Comparative Pharmacology of S(+)-Ibuprofen and (RS)-Ibuprofen. *Clinical Rheumatology* **2001**, *20* (1), 9-14. DOI: 10.1007/BF03342662.
- (109) Srinivas, N. R.; Barbhaiya, R. H.; Midha, K. K. Enantiomeric drug development: Issues, considerations, and regulatory requirements. *Journal of Pharmaceutical Sciences* **2001**, *90* (9), 1205-1215. DOI: 10.1002/jps.1074.
- (110) McConathy, J.; Owens, M. J. Stereochemistry in Drug Action. *Prim Care Companion J Clin Psychiatry* **2003**, *5* (2), 70-73. DOI: 10.4088/pcc.v05n0202.
- (111) Calcaterra, A.; D'Acquarica, I. The market of chiral drugs: Chiral switches versus de novo enantiomerically pure compounds. *Journal of Pharmaceutical and Biomedical Analysis* **2018**, *147*, 323-340. DOI: <https://doi.org/10.1016/j.jpba.2017.07.008>.
- (112) Coelho, M. M.; Fernandes, C.; Remião, F.; Tiritan, M. E. Enantioselectivity in Drug Pharmacokinetics and Toxicity: Pharmacological Relevance and Analytical Methods. *Molecules* **2021**, *26* (11), 3113.
- (113) Mane, S. Racemic drug resolution: a comprehensive guide. *Analytical Methods* **2016**, *8* (42), 7567-7586, DOI: 10.1039/C6AY02015A.
- (114) Pinto, M. M. M.; Fernandes, C.; Tiritan, M. E. Chiral Separations in Preparative Scale: A Medicinal Chemistry Point of View. **2020**, *25* (8), 1931.

References

- (115) Faigl, F.; Fogassy, E.; Nógrádi, M.; Pálóvics, E.; Schindler, J. Strategies in optical resolution: a practical guide. *Tetrahedron: Asymmetry* **2008**, *19* (5), 519-536. DOI: <https://doi.org/10.1016/j.tetasy.2008.02.004>.
- (116) Kaya, C.; Birgül, K.; Bülbül, B. Fundamentals of chirality, resolution, and enantiopure molecule synthesis methods. *Chirality* **2023**, *35* (1), 4-28. DOI: <https://doi.org/10.1002/chir.23512>.
- (117) Silvestri, I. P.; Colbon, P. J. J. The Growing Importance of Chirality in 3D Chemical Space Exploration and Modern Drug Discovery Approaches for Hit-ID. *ACS Medicinal Chemistry Letters* **2021**, *12* (8), 1220-1229. DOI: 10.1021/acsmchemlett.1c00251.
- (118) García Mancheño, O.; Waser, M. Recent Developments and Trends in Asymmetric Organocatalysis. *European Journal of Organic Chemistry* **2023**, *26* (1), DOI: <https://doi.org/10.1002/ejoc.202200950>.
- (119) Bihani, M.; Zhao, J. C. G. Advances in Asymmetric Diastereodivergent Catalysis. *Advanced Synthesis & Catalysis* **2017**, *359* (4), 534-575. DOI: <https://doi.org/10.1002/adsc.201601188>.
- (120) Hawkins, J. M.; Watson, T. J. N. Asymmetric Catalysis in the Pharmaceutical Industry. **2004**, *43* (25), 3224-3228. DOI: <https://doi.org/10.1002/anie.200330072>.
- (121) Ahrendt, K. A.; Borths, C. J.; MacMillan, D. W. C. New Strategies for Organic Catalysis: The First Highly Enantioselective Organocatalytic Diels–Alder Reaction. *Journal of the American Chemical Society* **2000**, *122* (17), 4243-4244. DOI: 10.1021/ja000092s.
- (122) Tu, Y.; Wang, Z.-X.; Shi, Y. J. J. o. t. A. C. S. An efficient asymmetric epoxidation method for trans-olefins mediated by a fructose-derived ketone. **1996**, *118* (40), 9806-9807.
- (123) Ahrendt, K. A.; Borths, C. J.; MacMillan, D. W. C. New strategies for organic catalysis: The first highly enantioselective organocatalytic Diels - Alder reaction. *J. Am. Chem. Soc.* **2000**, *122* (17), 4243-4244. DOI: 10.1021/ja000092s.
- (124) Gotoh, H.; Hayashi, Y. Diarylprolinol Silyl Ether as Catalyst of an exo-Selective, Enantioselective Diels–Alder Reaction. *Org Lett* **2007**, *9* (15), 2859-2862.
- (125) Pottel, J.; Moitessier, N. Customizable Generation of Synthetically Accessible, Local Chemical Subspaces. *Journal of Chemical Information and Modeling* **2017**, *57* (3), 454-467. DOI: 10.1021/acs.jcim.6b00648.
- (126) Therrien, E.; Englebienne, P.; Arrowsmith, A. G.; Mendoza-Sanchez, R.; Corbeil, C. R.; Weill, N.; Campagna-Slater, V.; Moitessier, N. Integrating Medicinal Chemistry, Organic/Combinatorial Chemistry, and Computational Chemistry for the Discovery of Selective Estrogen Receptor Modulators with Forecaster, a Novel Platform for Drug Discovery. *Journal of Chemical Information and Modeling* **2012**, *52* (1), 210-224. DOI: 10.1021/ci2004779.
- (127) Corbeil, C. R.; Thielges, S.; Schwartzentruber, J. A.; Moitessier, N. J. A. C.-G. E. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions: development and application of a rapid and accurate program based on organic principles. **2008**, *120* (14), 2675-2678.
- (128) Irwin, J. J.; Tang, K. G.; Young, J.; Dandarchuluun, C.; Wong, B. R.; Khurelbaatar, M.; Moroz, Y. S.; Mayfield, J.; Sayle, R. A. ZINC20—A Free Ultralarge-Scale Chemical Database for

References

- Ligand Discovery. *Journal of Chemical Information and Modeling* **2020**, *60* (12), 6065-6073. DOI: 10.1021/acs.jcim.0c00675.
- (129) Zabolotna, Y.; Bonachera, F.; Horvath, D.; Lin, A.; Marcou, G.; Klimchuk, O.; Varnek, A. Chemspace Atlas: Multiscale Chemography of Ultralarge Libraries for Drug Discovery. *Journal of Chemical Information and Modeling* **2022**, *62* (18), 4537-4548. DOI: 10.1021/acs.jcim.2c00509.
- (130) Burai-Patrascu, M. From Desktop to Benchtop – Developing Computational Tools for Organic and Medicinal Chemistry. McGill University, 2020.
- (131) Fuentealba, P.; Pérez, P.; Contreras, R. On the condensed Fukui function. *The Journal of Chemical Physics* **2000**, *113* (7), 2544-2551. DOI: 10.1063/1.1305879.
- (132) Merino, P.; Marqués-López, E.; Tejero, T.; Herrera, R. P. Enantioselective Organocatalytic Diels-Alder Reactions. *Synthesis* **2010**, 1-26.
- (133) Agirre, M.; Arrieta, A.; Arrastia, I.; Cossío, F. P. Organocatalysts Derived from Unnatural α -Amino Acids: Scope and Applications. *Chemistry – An Asian Journal* **2019**, *14* (1), 44-66, DOI: <https://doi.org/10.1002/asia.201801296>.
- (134) Laina-Martín, V.; Fernández-Salas, J. A.; Alemán, J. Organocatalytic Strategies for the Development of the Enantioselective Inverse-electron-demand Hetero-Diels-Alder Reaction. *Chemistry – A European Journal* **2021**, *27* (49), 12509-12520. DOI: <https://doi.org/10.1002/chem.202101696>.
- (135) Gotoh, H.; Hayashi, Y. Diarylprolinol Silyl Ether as Catalyst of an exo-Selective, Enantioselective Diels-Alder Reaction. *Org Lett* **2007**, *9* (15), 2859-2862.
- (136) Hayashi, Y.; Gotoh, H.; Hayashi, T.; Shoji, M. Diphenylprolinol Silyl Ethers as Efficient Organocatalysts for the Asymmetric Michael Reaction of Aldehydes and Nitroalkenes. **2005**, *44* (27), 4212-4215. DOI: <https://doi.org/10.1002/anie.200500599>.
- (137) Mose, R.; Jensen, M. E.; Preegel, G.; Jorgensen, K. A. Direct Access to Multifunctionalized Norcamphor Scaffolds by Asymmetric Organocatalytic Diels-Alder Reactions. *Angew Chem Int Ed Engl* **2015**, *54* (46), 13630-13634. DOI: 10.1002/anie.201507348.
- (138) Nielsen, M.; Worgull, D.; Zweifel, T.; Gschwend, B.; Bertelsen, S.; Jorgensen, K. A. Mechanisms in aminocatalysis. *Chem Commun (Camb)* **2011**, *47* (2), 632-649. DOI: 10.1039/c0cc02417a.
- (139) Larsen, C. H.-M. Investigating Imidazolidinone Catalysts: Enantioselective Organocatalytic Diels–Alder Reactions, Conjugate Additions to Access Non-Natural α -Amino Acids, and Bimodal Catalyst Activation for the Development of Organo-Cascade Reactions. PhD thesis, California Institute of Technology, Pasadena, CA, 2005.
- (140) Evans, G.; Gibbs, T. J. K.; Jenkins, R. L.; Coles, S. J.; Hursthouse, M. B.; Platts, J. A.; Tomkinson, N. C. O. Kinetics of Iminium Ion Catalysis. *Angew. Chem. Int. Ed.* **2008**, *47* (15), 2820-2823. DOI: <https://doi.org/10.1002/anie.200705539>.
- (141) Brazier, J. B.; Cavill, J. L.; Elliott, R. L.; Evans, G.; Gibbs, T. J. K.; Jones, I. L.; Platts, J. A.; Tomkinson, N. C. O. The α -effect in cyclic secondary amines: new scaffolds for iminium ion

References

- accelerated transformations. *Tetrahedron* **2009**, *65* (48), 9961-9966. DOI: 10.1016/j.tet.2009.10.009.
- (142) An, F.; Maji, B.; Min, E.; Ofial, A. R.; Mayr, H. Basicities and Nucleophilicities of Pyrrolidines and Imidazolidinones Used as Organocatalysts. *Journal of the American Chemical Society* **2020**, *142* (3), 1526-1547. DOI: 10.1021/jacs.9b11877.
- (143) Gordillo, R.; Houk, K. N. Origins of stereoselectivity in Diels-Alder cycloadditions catalyzed by chiral imidazolidinones. *J. Am. Chem. Soc.* **2006**, *128* (11), 3543-3553.
- (144) Brazier, J. B.; Jones, K. M.; Platts, J. A.; Tomkinson, N. C. O. On the Roles of Protic Solvents in Imidazolidinone-Catalyzed Transformations. *Angew. Chem. Int. Ed.* **2011**, *50* (7), 1613-1616. DOI: <https://doi.org/10.1002/anie.201005892>.
- (145) Evans, G. J. S.; White, K.; Platts, J. A.; Tomkinson, N. C. O. Computational study of iminium ion formation: effects of amine structure. *Organic & Biomolecular Chemistry* **2006**, *4* (13), 2616-2627, DOI: 10.1039/B602645A.
- (146) Gotoh, H.; Uchamaru, T.; Hayashi, Y. Two Reaction Mechanisms via Iminium Ion Intermediates: The Different Reactivities of Diphenylprolinol Silyl Ether and Trifluoromethyl-Substituted Diarylprolinol Silyl Ether. *Chemistry* **2015**, *21* (35), 12337-12346. DOI: 10.1002/chem.201500326.
- (147) Brazier, J. B.; Hopkins, G. P.; Jirari, M.; Mutter, S.; Pommereuil, R.; Samulis, L.; Platts, J. A.; Tomkinson, N. C. O. Iminium ion catalysis: direct comparison of imidazolidinone and diarylprolinol ether reactivity. *Tetrahedron Letters* **2011**, *52* (21), 2783-2785. DOI: 10.1016/j.tetlet.2011.03.129.
- (148) Brazier, J. B.; Jones, K. M.; Platts, J. A.; Tomkinson, N. C. On the roles of protic solvents in imidazolidinone-catalyzed transformations. *Angew Chem Int Ed Engl* **2011**, *50* (7), 1613-1616. DOI: 10.1002/anie.201005892.
- (149) Evans, G.; Gibbs, T. J.; Jenkins, R. L.; Coles, S. J.; Hursthouse, M. B.; Platts, J. A.; Tomkinson, N. C. Kinetics of iminium ion catalysis. *Angew Chem Int Ed Engl* **2008**, *47* (15), 2820-2823. DOI: 10.1002/anie.200705539.
- (150) Montgomery, R. E. Catalysis of peroxymonosulfate reactions by ketones. *Journal of the American Chemical Society* **1974**, *96* (25), 7820-7821. DOI: 10.1021/ja00832a040.
- (151) Curci, R.; Fiorentino, M.; Troisi, L.; Edwards, J. O.; Pater, R. H. Epoxidation of alkenes by dioxirane intermediates generated in the reaction of potassium caroate with ketones. *The Journal of Organic Chemistry* **1980**, *45* (23), 4758-4760. DOI: 10.1021/jo01311a040.
- (152) Cicala, G.; Curci, R.; Fiorentino, M.; Laricchiuta, O. Stereo- and regioselectivities in the epoxidation of some allylic alcohols by the dioxirane intermediate generated in the reaction of potassium caroate with acetone. *The Journal of Organic Chemistry* **1982**, *47* (13), 2670-2673. DOI: 10.1021/jo00134a033.
- (153) Denmark, S. E.; Forbes, D. C.; Hays, D. S.; DePue, J. S.; Wilde, R. G. Catalytic Epoxidation of Alkenes with Oxone. *The Journal of Organic Chemistry* **1995**, *60* (5), 1391-1407. DOI: 10.1021/jo00110a049.

References

- (154) Baumstark, A. L.; McCloskey, C. J. Epoxidation of alkenes by dimethyldioxirane: Evidence for a spiro transition state. *Tetrahedron Letters* **1987**, *28* (29), 3311-3314. DOI: [https://doi.org/10.1016/S0040-4039\(00\)95499-6](https://doi.org/10.1016/S0040-4039(00)95499-6).
- (155) Jenson, C.; Liu, J.; Houk, K. N.; Jorgensen, W. L. Elucidation of Transition Structures and Solvent Effects for Epoxidation by Dimethyldioxirane. *Journal of the American Chemical Society* **1997**, *119* (52), 12982-12983. DOI: 10.1021/ja971766a.
- (156) Houk, K.; Liu, J.; DeMello, N. C.; Condroski, K. R. J. J. o. t. A. C. S. Transition states of epoxidations: diradical character, spiro geometries, transition state flexibility, and the origins of stereoselectivity. **1997**, *119* (42), 10147-10152.
- (157) Shi, Y. Organocatalytic Asymmetric Epoxidation of Olefins by Chiral Ketones. *Accounts of Chemical Research* **2004**, *37* (8), 488-496. DOI: 10.1021/ar030063x.
- (158) Yang, D.; Wong, M.-K.; Yip, Y.-C. J. T. J. o. O. C. Epoxidation of olefins using methyl (trifluoromethyl) dioxirane generated in situ. **1995**, *60* (12), 3887-3889.
- (159) Edwards, J. O.; Pater, R. H.; Curclf, R.; Furia, F. D. ON THE FORMATION AND REACTIVITY OF DIOXIRANE INTERMEDIATES IN THE REACTION OF PEROXOANIONS WITH ORGANIC SUBSTRATES*. *Photochemistry and Photobiology* **1979**, *30* (1), 63-70, DOI: <https://doi.org/10.1111/j.1751-1097.1979.tb07116.x>.
- (160) Curci, R.; Fiorentino, M.; Serio, M. R. Asymmetric Epoxidation of Unfunctionalized Alkenes by Dioxirane Intermediates Generated from Potassium Peroxomonosulphate and Chiral Ketones. *Chem Comm* **1984**, 155-156.
- (161) Wong, O. A.; Shi, Y. Organocatalytic Oxidation. Asymmetric Epoxidation of Olefins Catalyzed by Chiral Ketones and Iminium Salts. *Chemical Reviews* **2008**, *108* (9), 3958-3987. DOI: 10.1021/cr068367v.
- (162) Wang, Z.-X.; Tu, Y.; Frohn, M.; Zhang, J.-R.; Shi, Y. An Efficient Catalytic Asymmetric Epoxidation Method. *Journal of the American Chemical Society* **1997**, *119* (46), 11224-11235. DOI: 10.1021/ja972272g.
- (163) Shu, L.; Shi, Y. Asymmetric epoxidation using hydrogen peroxide (H₂O₂) as primary oxidant. *Tetrahedron Letters* **1999**, *40* (50), 8721-8724. DOI: [https://doi.org/10.1016/S0040-4039\(99\)01814-6](https://doi.org/10.1016/S0040-4039(99)01814-6).
- (164) Travis, B. R.; Sivakumar, M.; Hollist, G. O.; Borhan, B. Facile Oxidation of Aldehydes to Acids and Esters with Oxone. *Organic Letters* **2003**, *5* (7), 1031-1034. DOI: 10.1021/ol0340078.
- (165) Yang, D.; Wong, M.-K.; Yip, Y.-C.; Wang, X.-C.; Tang, M.-W.; Zheng, J.-H.; Cheung, K.-K. Design and Synthesis of Chiral Ketones for Catalytic Asymmetric Epoxidation of Unfunctionalized Olefins. *Journal of the American Chemical Society* **1998**, *120* (24), 5943-5952. DOI: 10.1021/ja980428m.
- (166) Tu, Y.; Wang, Z.-X.; Frohn, M.; He, M.; Yu, H.; Tang, Y.; Shi, Y. Structural Probing of Ketone Catalysts for Asymmetric Epoxidation. *The Journal of Organic Chemistry* **1998**, *63* (23), 8475-8485. DOI: 10.1021/jo9817218.

References

- (167) Denmark, S. E.; Matsubashi, H. Chiral Fluoro Ketones for Catalytic Asymmetric Epoxidation of Alkenes with Oxone. *The Journal of Organic Chemistry* **2002**, *67* (10), 3479-3486. DOI: 10.1021/jo020050h.
- (168) Fehr, A. R.; Perlman, S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol Biol* **2015**, *1282*, 1-23. DOI: 10.1007/978-1-4939-2438-7_1.
- (169) King, A. An uncommon cold. *New Sci* **2020**, *246* (3280), 32-35. DOI: 10.1016/s0262-4079(20)308629.
- (170) Domling, A.; Gao, L. Chemistry and Biology of SARS-CoV-2. *Chem* **2020**, *6* (6), 1283-1295. DOI: 10.1016/j.chempr.2020.04.023.
- (171) Organization, W. H. *WHO Coronavirus (COVID-19) Dashboard*. 2023. (accessed 2023 March 20).
- (172) Bank, W. *Unemployment rate from 2002 to 2021*. Statistica, 2023. (accessed 2021 March 21).
- (173) Lamers, M. M.; Haagmans, B. L. SARS-CoV-2 pathogenesis. *Nature Reviews Microbiology* **2022**, *20* (5), 270-284. DOI: 10.1038/s41579-022-00713-0.
- (174) Chaudhary, N.; Weissman, D.; Whitehead, K. A. mRNA vaccines for infectious diseases: principles, delivery and clinical translation. *Nature Reviews Drug Discovery* **2021**, *20* (11), 817-838. DOI: 10.1038/s41573-021-00283-5.
- (175) Dolgin, E. Pan-coronavirus vaccine pipeline takes form. *Nature reviews. Drug discovery* **2022**, *21* (5), 324-326. DOI: 10.1038/d41573-022-00074-6.
- (176) Li, G.; Hilgenfeld, R.; Whitley, R.; De Clercq, E. Therapeutic strategies for COVID-19: progress and lessons learned. *Nature Reviews Drug Discovery* **2023**. DOI: 10.1038/s41573-023-00672-y.
- (177) Napolitano, V.; Dabrowska, A.; Schorpp, K.; Mourão, A.; Barreto-Duran, E.; Bedyk, M.; Botwina, P.; Brandner, S.; Bostock, M.; Chykunova, Y.; et al. Acriflavine, a clinically approved drug, inhibits SARS-CoV-2 and other betacoronaviruses. *Cell Chem Biol* **2022**, *29* (5), 774-784.e778. DOI: 10.1016/j.chembiol.2021.11.006.
- (178) *FDA Approves First Treatment for COVID-19*. 2020. FDA Approves First Treatment for COVID-19 (accessed 2023/09/01).
- (179) Pan, H.; Peto, R.; Henao-Restrepo, A. M.; Preziosi, M. P.; Sathiyamoorthy, V.; Abdool Karim, Q.; Alejandria, M. M.; Hernández García, C.; Kieny, M. P.; Malekzadeh, R.; et al. Repurposed Antiviral Drugs for Covid-19 - Interim WHO Solidarity Trial Results. *N Engl J Med* **2021**, *384* (6), 497-511. DOI: 10.1056/NEJMoa2023184.
- (180) Boras, B.; Jones, R. M.; Anson, B. J.; Arenson, D.; Aschenbrenner, L.; Bakowski, M. A.; Beutler, N.; Binder, J.; Chen, E.; Eng, H.; et al. Discovery of a Novel Inhibitor of Coronavirus 3CL Protease for the Potential Treatment of COVID-19. *bioRxiv* **2021**. DOI: 10.1101/2020.09.12.293498.
- (181) Owen, D. R.; Allerton, C. M. N.; Anderson, A. S.; Aschenbrenner, L.; Avery, M.; Berritt, S.; Boras, B.; Cardin, R. D.; Carlo, A.; Coffman, K. J.; et al. An oral SARS-CoV-2 M(pro) inhibitor

References

- clinical candidate for the treatment of COVID-19. *Science* **2021**, *374* (6575), 1586-1593. DOI: 10.1126/science.abl4784.
- (182) Shionogi. Shionogi Receives U.S. FDA Fast Track Designation for Ensitrelvir Fumaric Acid, an Investigational Oral Antiviral for COVID-19. 2023.
- (183) Kabinger, F.; Stiller, C.; Schmitzová, J.; Dienemann, C.; Kokic, G.; Hillen, H. S.; Höbartner, C.; Cramer, P. Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nature Structural & Molecular Biology* **2021**, *28* (9), 740-746. DOI: 10.1038/s41594-021-00651-0.
- (184) Shuiyun, L.; Grace, N.; Ryan, L. S.; William, A. C.; Andres Emanuelli, C.; Zachary, C. L.; Nicole, L.; Huanchun, Z.; Jasper, L.; Maria, E. C.; et al. Nirmatrelvir Resistance in SARS-CoV-2 Omicron_BA.1 and WA1 Replicons and Escape Strategies. *bioRxiv* **2023**, DOI: 10.1101/2022.12.31.522389.
- (185) Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M.; et al. Activity profiling and structures of inhibitor-bound SARS-CoV-2-PLpro protease provides a framework for anti-COVID-19 drug design. *bioRxiv* **2020**. DOI: 10.1101/2020.04.29.068890.
- (186) Hedstrom, L. Serine Protease Mechanism and Specificity. *Chemical Reviews* **2002**, *102* (12), 4501-4524. DOI: 10.1021/cr000033x.
- (187) Agback, P.; Agback, T. Direct evidence of a low barrier hydrogen bond in the catalytic triad of a Serine protease. *Scientific Reports* **2018**, *8* (1), 10078. DOI: 10.1038/s41598-018-28441-7.
- (188) Fu, Z.; Huang, B.; Tang, J.; Liu, S.; Liu, M.; Ye, Y.; Liu, Z.; Xiong, Y.; Zhu, W.; Cao, D.; et al. The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery. *Nat Commun* **2021**, *12* (1), 488. DOI: 10.1038/s41467-020-20718-8.
- (189) Ratia, K.; Pegan, S.; Takayama, J.; Sleeman, K.; Coughlin, M.; Baliji, S.; Chaudhuri, R.; Fu, W.; Prabhakar, B. S.; Johnson, M. E.; et al. A noncovalent class of papain-like protease/deubiquitinase inhibitors blocks SARS virus replication. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (42), 16119-16124. DOI: 10.1073/pnas.0805240105.
- (190) Hughes, J. P.; Rees S Fau - Kalindjian, S. B.; Kalindjian Sb Fau - Philpott, K. L.; Philpott, K. L. Principles of early drug discovery. (1476-5381 (Electronic)). From 2011 Mar.
- (191) Tan, H.; Hu, Y.; Jadhav, P.; Tan, B.; Wang, J. Progress and Challenges in Targeting the SARS-CoV-2 Papain-like Protease. *Journal of Medicinal Chemistry* **2022**, *65* (11), 7561-7580. DOI: 10.1021/acs.jmedchem.2c00303.
- (192) Shen, Z.; Ratia, K.; Cooper, L.; Kong, D.; Lee, H.; Kwon, Y.; Li, Y.; Alqarni, S.; Huang, F.; Dubrovskiy, O.; et al. Design of SARS-CoV-2 PLpro Inhibitors for COVID-19 Antiviral Therapy Leveraging Binding Cooperativity. *Journal of Medicinal Chemistry* **2022**, *65* (4), 2940-2955. DOI: 10.1021/acs.jmedchem.1c01307.
- (193) Ratia, K.; Kilianski, A.; Baez-Santos, Y. M.; Baker, S. C.; Mesecar, A. Structural Basis for the Ubiquitin-Linkage Specificity and deISGylating activity of SARS-CoV papain-like protease. *PLoS pathogens* **2014**, *10* (5), DOI: 10.1371/journal.ppat.1004113.

References

- (194) Sanders, B. C.; Pokhrel, S.; Labbe, A. D.; Mathews, I. I.; Cooper, C. J.; Davidson, R. B.; Phillips, G.; Weiss, K. L.; Zhang, Q.; O'Neill, H.; et al. Potent and selective covalent inhibition of the papain-like protease from SARS-CoV-2. *Nature Communications* **2023**, *14* (1), 1733. DOI: 10.1038/s41467-023-37254-w.
- (195) Pottel, J.; Therrien, E.; Gleason, J. L.; Moitessier, N. Docking Ligands into Flexible and Solvated Macromolecules. 6. Development and Application to the Docking of HDACs and other Zinc Metalloenzymes Inhibitors. *Journal of Chemical Information and Modeling* **2014**, *54* (1), 254-265. DOI: 10.1021/ci400550m.
- (196) Therrien, E.; Weill, N.; Tomberg, A.; Corbeil, C. R.; Lee, D.; Moitessier, N. Docking Ligands into Flexible and Solvated Macromolecules. 7. Impact of Protein Flexibility and Water Molecules on Docking-Based Virtual Screening Accuracy. *Journal of Chemical Information and Modeling* **2014**, *54* (11), 3198-3210. DOI: 10.1021/ci500299h.
- (197) Labarre, A.; Stille, J. K.; Patrascu, M. B.; Martins, A.; Pottel, J.; Moitessier, N. Docking Ligands into Flexible and Solvated Macromolecules. 8. Forming New Bonds—Challenges and Opportunities. *Journal of Chemical Information and Modeling* **2022**, *62* (4), 1061-1077. DOI: 10.1021/acs.jcim.1c00701.
- (198) Stille, J. K.; Lau, K. S. Y. Mechanisms of oxidative addition of organic halides to Group 8 transition-metal complexes. *Accounts of Chemical Research* **1977**, *10* (12), 434-442. DOI: 10.1021/ar50120a002.
- (199) Ma, D.; Cai, Q.; Zhang, H. Mild Method for Ullmann Coupling Reaction of Amines and Aryl Halides. *Organic Letters* **2003**, *5* (14), 2453-2455. DOI: 10.1021/ol0346584.
- (200) Ma, D.; Cai, Q. Copper/Amino Acid Catalyzed Cross-Couplings of Aryl and Vinyl Halides with Nucleophiles. *Accounts of Chemical Research* **2008**, *41* (11), 1450-1460. DOI: 10.1021/ar8000298.
- (201) Vilaivan, T. A rate enhancement of tert-butoxycarbonylation of aromatic amines with Boc₂O in alcoholic solvents. *Tetrahedron Letters* **2006**, *47* (38), 6739-6742. DOI: <https://doi.org/10.1016/j.tetlet.2006.07.097>.
- (202) Young, H. A.; Guthrie, Q. A. E.; Proulx, C. N-Arylation of Amino Acid Esters to Expand Side Chain Diversity in Ketoxime Peptide Ligations. *The Journal of Organic Chemistry* **2020**, *85* (3), 1748-1755. DOI: 10.1021/acs.joc.9b02810.
- (203) Xie, Z.; Liu, X.; Liu, L. Copper-Catalyzed Aerobic Enantioselective Cross-Dehydrogenative Coupling of N-Aryl Glycine Esters with Terminal Alkynes. *Organic Letters* **2016**, *18* (12), 2982-2985. DOI: 10.1021/acs.orglett.6b01328.
- (204) Yamamoto, H.; Hattori, T. Chemical Ligation of Ketoxime Peptides Derived from N-Aryl Peptides and Alkoxyamines. *Synfacts* **2019**, *15* (12), 1447.
- (205) Cataldo, F.; Faucette, B.; Huang, S.; Ebenezer, W. On the early reaction stages of ozone with N,N'-substituted p-phenylenediamines (6PPD, 77PD) and N,N',N''-substituted-1,3,5-triazine "Durazone®": An electron spin resonance (ESR) and electronic absorption spectroscopy study. *Polymer Degradation and Stability* **2015**, *111*, 223-231. DOI: <https://doi.org/10.1016/j.polymdegradstab.2014.11.011>.

References

- (206) Wienhold, M.; Molloy, J. J.; Daniliuc, C. G.; Gilmour, R. Coumarins by Direct Annulation: β -Borylacrylates as Ambiphilic C3-Synthons. *Angewandte Chemie International Edition* **2021**, *60* (2), 685-689, DOI: <https://doi.org/10.1002/anie.202012099>.
- (207) Ma, C.; Sacco, M. D.; Xia, Z.; Lambrinidis, G.; Townsend, J. A.; Hu, Y.; Meng, X.; Szeto, T.; Ba, M.; Zhang, X.; et al. Discovery of SARS-CoV-2 Papain-like Protease Inhibitors through a Combination of High-Throughput Screening and a FlipGFP-Based Reporter Assay. *ACS Central Science* **2021**, *7* (7), 1245-1260. DOI: 10.1021/acscentsci.1c00519.
- (208) Ma, D.; Xia, C. CuI-Catalyzed Coupling Reaction of β -Amino Acids or Esters with Aryl Halides at Temperature Lower Than That Employed in the Normal Ullmann Reaction. Facile Synthesis of SB-214857. *Organic Letters* **2001**, *3* (16), 2583-2586. DOI: 10.1021/ol1016258r.
- (209) Nair, A. S.; Singh, A. K.; Kumar, A.; Kumar, S.; Sukumaran, S.; Koyiparambath, V. P.; Pappachen, L. K.; Rangarajan, T. M.; Kim, H.; Mathew, B. FDA-Approved Trifluoromethyl Group-Containing Drugs: A Review of 20 Years. In *Processes*, 2022; Vol. 10.
- (210) Altman, R. A.; Buchwald, S. L. Cu-catalyzed Goldberg and Ullmann reactions of aryl halides using chelating N- and O-based ligands. *Nature Protocols* **2007**, *2* (10), 2474-2479. DOI: 10.1038/nprot.2007.364.
- (211) Stress, C. J.; Sauter, B.; Schneider, L. A.; Sharpe, T.; Gillingham, D. A DNA-Encoded Chemical Library Incorporating Elements of Natural Macrocycles. *Angewandte Chemie International Edition* **2019**, *58* (28), 9570-9574, DOI: <https://doi.org/10.1002/anie.201902513>.
- (212) Gotoh, H.; Hayashi, Y. Diarylprolinol silyl ether as catalyst of an exo-selective, enantioselective diels-alder reaction. *Org. Lett.* **2007**, *9* (15), 2859-2862. DOI: 10.1021/ol071009+
- (213) Evans, D. A.; Chapman, K. T.; Bisaha, J. Asymmetric Diels-Alder cycloaddition reactions with chiral α,β -unsaturated N-acyloxazolidinones. *J. Am. Chem. Soc.* **1988**, *110* (4), 1238-1256.
- (214) Corey, E. J.; Cheng, X.-M.; Cimprich, K. A. 1-Mesityl-2,2,2-trifluoroethanol, an outstanding new chiral controller for catalyzed diels-alder reactions. *Tetrahedron Lett.* **1991**, *32* (47), 6839-6842
- (215) Oppolzer, W.; Kurth, M.; Reichlin, D.; Moffatt, F. A reinvestigation of asymmetric induction in diels-alder reactions to chiral acrylates. *Tetrahedron Lett.* **1981**, *22* (27), 2545-2548
- (216) Banks, M. R.; Blake, A. J.; Cadogan, J. I. G.; Dawson, I. M.; Gosney, I.; Grant, K. J.; Gaur, S.; Hodgson, P. K. G.; Knight, K. S.; Smith, G. W.; et al. Enantiospecific preparation of [(2,6)-endo]-5-aza-1,10,10-trimethyl-3-oxatricyclo[5.2.1.0^{2,6}]decan-4-one by a nitrene-mediated route from [(1-endo)-(-)-borneol and its utility as a chiral auxiliary in some asymmetric transformations. *Tetrahedron* **1992**, *48* (37), 7979-8006
- (217) Banks, M. R.; Blake, A. J.; Brown, A. R.; Cadogan, J. I. G.; Gaur, S.; Gosney, I.; Hodgson, P. K. G.; Thorburn, P. Exploiting steric shielding: Tuning terpenoid-derived oxazolidin-2-ones as chiral auxiliaries for the Diels-Alder reaction. *Tetrahedron Lett.* **1994**, *35* (3), 489-492
- (218) Banks, M. R.; Blake, A. J.; Cadogan, J. I. G.; Doyle, A. A.; Gosney, I.; Hodgson, P. K. G.; Thorburn, P. Asymmetric Diels-Alder reactions employing modified camphor-derived oxazolidin-2-one chiral auxiliaries. *Tetrahedron* **1996**, *52* (11), 4079-4094

References

- (219) Banks, M. R.; Cadogan, J. I. G.; Gosney, I.; Gaur, S.; Hodgson, P. K. G. Highly regio- and stereo-specific preparation of a new carbohydrate-based 1,3-oxazin-2-one by the INIR method and its applications in some asymmetric transformations. *Tetrahedron Asym.* **1994**, *5* (12), 2447-2458
- (220) Banks, M. R.; Blake, A. J.; Cadogan, J. I. G.; Dawson, I. M.; Gaur, S.; Gosney, I.; Gould, R. O.; Grant, K. J.; Hodgson, P. K. G. (5R)-7,8:9,10-Di-O-isopropylidene-2,6-dioxa-4-azaspiro[4,5]decan-3-one: A new chiral spirooxazolidin-2-one derived from D-(+)-galactose for use in asymmetric transformations. *J. Chem. Soc. Chem. Commun.* **1993**, (14), 1146-1148
- (221) Banks, M. R.; Cadogan, J. I. G.; Gosney, I.; Gould, R. O.; Hodgson, P. K. G.; McDougall, D. Preparation of enantiomerically pure fructose-derived 1,3-oxazin-2-one by INIR methodology and its application as a chiral auxiliary in some model asymmetric reactions. *Tetrahedron* **1998**, *54* (33), 9765-9784
- (222) Matsunaga, H.; Kimura, K.; Ishizuka, T.; Haratake, M.; Kunleda, T. Conformationally rigid chiral [4+2]cycloadduct-based 2-oxazolidinones as new auxiliaries. *Tetrahedron Lett.* **1991**, *32*, 7715-7718.
- (223) Armstrong, A.; Ahmed, G.; Dominguez-Fernandez, B.; Hayter, B. R.; Wailes, J. S. Enantioselective Epoxidation of Alkenes Catalyzed by 2-Fluoro-N-Carboethoxytropinone and Related Tropinone Derivatives. *J. Org. Chem.* **2002**, *67* (24), 8610-8617
- (224) Armstrong, A.; Hayter, B. R.; Moss, W. O.; Reeves, J. R.; Wailes, J. S. Alkene epoxidation catalyzed by bicyclo[3.2.1]octan-3-ones: effects of structural modifications on catalyst efficiency and epoxidation enantioselectivity. *Tetrahedron Asym.* **2000**, *11* (10), 2057-2061
- (225) Armstrong, A.; Dominguez-Fernandez, B.; Tsuchiya, T. Bicyclo[3.2.1]octanone catalysts for asymmetric alkene epoxidation: the effect of disubstitution. *Tetrahedron* **2006**, *62* (28), 6614-6620. DOI: 10.1016/j.tet.2005.12.073.
- (226) Armstrong, A.; Moss, W. O.; Reeves, J. R. Asymmetric epoxidation catalyzed by esters of [alpha]-hydroxy-8-oxabicyclo[3.2.1]octan-3-one. *Tetrahedron Asym.* **2001**, *12* (20), 2779-2781
- (227) Solladié-Cavallo, A.; Jierry, L.; Klein, A.; Schmitt, M.; Welter, R. [alpha]-Fluoro decalones as chiral epoxidation catalysts: fluorine effect. *Tetrahedron Asym.* **2004**, *15* (24), 3891-3898
- (228) Denmark, S. E.; Matsuhashi, H. Chiral Fluoro Ketones for Catalytic Asymmetric Epoxidation of Alkenes with Oxone. *J. Org. Chem.* **2002**, *67* (10), 3479-3486
- (229) Armstrong, A.; Tsuchiya, T. A new class of chiral tetrahydropyran-4-one catalyst for asymmetric epoxidation of alkenes. *Tetrahedron* **2006**, *62* (2-3), 257-263. DOI: <http://dx.doi.org/10.1016/j.tet.2005.08.112>.
- (230) Burke, C. P.; Shi, Y. Regio- and Enantioselective Epoxidation of Dienes by a Chiral Dioxirane: Synthesis of Optically Active Vinyl cis-Epoxides. *Angew. Chem. Int. Ed.* **2006**, *45* (27), 4475-4478
- (231) Wang, Z. X.; Tu, Y.; Frohn, M.; Zhang, J. R.; Shi, Y. An Efficient Catalytic Asymmetric Epoxidation Method. *J. Am. Chem. Soc.* **1997**, *119* (46), 11224-11235
- (232) Tu, Y.; Wang, Z.-X.; Shi, Y. An Efficient Asymmetric Epoxidation Method for trans-Olefins Mediated by a Fructose-Derived Ketone. *J. Am. Chem. Soc.* **1996**, *118* (40), 9806-9807. DOI: 10.1021/ja962345g.

References

- (233) Wang, Z. X.; Miller, S. M.; Anderson, O. P.; Shi, Y. Asymmetric Epoxidation by Chiral Ketones Derived from Carbocyclic Analogues of Fructose. *J. Org. Chem.* **2001**, *66* (2), 521-530
- (234) Hartikka, A.; Hojabri, L.; Bose, P. P.; Arvidsson, P. I. Synthesis and application of novel imidazole and 1H-tetrazolic acid containing catalysts in enantioselective organocatalyzed Diels-Alder reactions. *Tetrahedron Asym.* **2009**, *20* (16), 1871-1876. DOI: 10.1016/j.tetasy.2009.07.029
- (235) Ahrendt, K. A.; Borths, C. J.; MacMillan, D. W. C. New strategies for organic catalysis: The first highly enantioselective organocatalytic diels - Alder reaction [16]. *J. Am. Chem. Soc.* **2000**, *122* (17), 4243-4244. DOI: 10.1021/ja000092s.
- (236) Bonini, B. F.; Capitò, E.; Comes-Franchini, M.; Fochi, M.; Ricci, A.; Zwanenburg, B. Aziridin-2-yl methanols as organocatalysts in Diels-Alder reactions and Friedel-Crafts alkylations of N-methyl-pyrrole and N-methyl-indole. *Tetrahedron Asym.* **2006**, *17* (22), 3135-3143. DOI: 10.1016/j.tetasy.2006.11.028.
- (237) He, H.; Pei, B. J.; Chou, H. H.; Tian, T.; Chan, W. H.; Lee, A. W. M. Camphor sulfonyl hydrazines (CaSH) as organocatalysts in enantioselective Diels-Alder reactions. *Org. Lett.* **2008**, *10* (12), 2421-2424. DOI: 10.1021/ol8005826.
- (238) Lemay, M.; Ogilvie, W. W. Aqueous enantioselective organocatalytic Diels-Alder reactions employing hydrazide catalysts. A new scaffold for organic acceleration. *Org. Lett.* **2005**, *7* (19), 4141-4144. DOI: 10.1021/ol051476w.
- (239) Kano, T.; Tanaka, Y.; Maruoka, K. exo-selective asymmetric Diels-Alder reaction catalyzed by diamine salts as organocatalysts. *Org. Lett.* **2006**, *8* (13), 2687-2689. DOI: 10.1021/ol060621i.
- (240) Shi, L. X.; Sun, Q.; Ge, Z. M.; Zhu, Y. Q.; Cheng, T. M.; Li, R. T. Dipeptide-catalyzed direct asymmetric aldol reaction. *Synlett* **2004**, (12), 2215-2217
- (241) List, B.; Lerner, R. A.; Barbas III, C. F. Proline-catalyzed direct asymmetric aldol reactions [13]. *J. Am. Chem. Soc.* **2000**, *122* (10), 2395-2396. DOI: 10.1021/ja994280y.
- (242) Sakthivel, K.; Notz, W.; Bui, T.; Barbas III, C. F. Amino acid catalyzed direct asymmetric aldol reactions: A bioorganic approach to catalytic asymmetric carbon-carbon bond-forming reactions. *J. Am. Chem. Soc.* **2001**, *123* (22), 5260-5267. DOI: 10.1021/ja010037z.
- (243) Wagner, M.; Contie, Y.; Ferroud, C.; Revial, G. Enantioselective Aldol Reactions and Michael Additions Using Proline Derivatives as Organocatalysts. *Int. J. Org. Chem.* **2014**, *4*, 55-67.
- (244) Cobb, A. J. A.; Shaw, D. M.; Longbottom, D. A.; Gold, J. B.; Ley, S. V. Organocatalysis with proline derivatives: Improved catalysts for the asymmetric Mannich, nitro-Michael and aldol reactions. *Org. Biomol. Chem.* **2005**, *3* (1), 84-96. DOI: 10.1039/b414742a.
- (245) Gryko, D.; Lipiński, R. L-prolinethioamides - Efficient organocatalysts for the direct asymmetric aldol reaction. *Adv. Synth. Cat.* **2005**, *347* (15), 1948-1952. DOI: 10.1002/adsc.200505247.
- (246) Tang, Z.; Jiang, F.; Cui, X.; Gong, L. Z.; Mi, A. Q.; Jiang, Y. Z.; Wu, Y. D. Enantioselective direct aldol reactions catalyzed by L-prolinamide derivatives. *Proc. Natl Acad. Sci. USA* **2004**, *101* (16), 5755-5760. DOI: 10.1073/pnas.0307176101.

References

- (247) Gryko, D.; Chromiński, M.; Pielacińska, D. J. Prolinethioamides versus Prolinamides in Organocatalyzed Aldol Reactions—A Comparative Study. *Symmetry* **2011**, *3*, 265-282.
- (248) Tang, Z.; Jiang, F.; Yu, L. T.; Cui, X.; Gong, L. Z.; Mi, A. Q.; Jiang, Y. Z.; Wu, Y. D. Novel small organic molecules for a highly enantioselective direct aldol reaction. *J. Am. Chem. Soc.* **2003**, *125* (18), 5262-5263
- (249) Tang, Z.; Yang, Z. H.; Chen, X. H.; Cun, L. F.; Mi, A. Q.; Jiang, Y. Z.; Gong, L. Z. A highly efficient organocatalyst for direct Aldol reactions of ketones with aldehydes. *J. Am. Chem. Soc.* **2005**, *127* (25), 9285-9289
- (250) Wang, B.; Liu, X. W.; Liu, L. Y.; Chang, W. X.; Li, J. Highly efficient direct asymmetric aldol reactions catalyzed by a prolinethioamide derivative in aqueous media. *Eur. J. Org. Chem.* **2010**, (31), 5951-5954. DOI: 10.1002/ejoc.201001115.
- (251) Becker, H.; King, S. B.; Taniguchi, M.; Vanhessche, K. P. M.; Sharpless, K. B. New ligands and improved enantioselectivities for the asymmetric dihydroxylation of olefins. *J. Org. Chem.* **1995**, *60* (13), 3940-3941
- (252) Crispino, G. A.; Makita, A.; Wang, Z. M.; Sharpless, K. B. A comparison of ligands proposed for the asymmetric dihydroxylation. *Tetrahedron Lett.* **1994**, *35* (4), 543-546
- (253) Corey, E. J.; Noe, M. C.; Grogan, M. J. A mechanistically designed mono-cinchona alkaloid is an excellent catalyst for the enantioselective dihydroxylation of olefins. *Tetrahedron Lett.* **1994**, *35* (35), 6427-6430
- (254) Sharpless, K. B.; Amberg, W.; Beller, M.; Chen, H.; Hartung, J.; Kawanami, Y.; Labben, D.; Manoury, E.; Ogino, Y.; Shibata, T.; et al. New ligands double the scope of the catalytic asymmetric dihydroxylation of olefins. *J. Org. Chem.* **1991**, *56* (15), 4585-4588
- (255) Jacobsen, E. N.; Marko, I.; Mungall, W. S.; Schroder, G.; Sharpless, K. B. Asymmetric Dihydroxylation via Ligand-Accelerated Catalysis. *J. Am. Chem. Soc.* **1988**, *110*, 1968-1970.
- (256) Huang, J.; Corey, E. J. A Mechanistically Guided Design Leads to the Synthesis of an Efficient and Practical New Reagent for the Highly Enantioselective, Catalytic Dihydroxylation of Olefins. *Org. Lett.* **2003**, *5* (19), 3455-3458
- (257) Oishi, T.; Hirama, M. Synthesis of chiral 2,3-disubstituted 1,4-diazabicyclo[2.2.2]octane. New ligand for the osmium-catalyzed asymmetric dihydroxylation of olefins. *Tetrahedron Lett.* **1992**, *33* (5), 639-642.
- (258) Becker, H.; Barry Sharpless, K. A New Ligand Class for the Asymmetric Dihydroxylation of Olefins. *Angew. Chem. Int. Ed.* **1996**, *35* (4), 448-451
- (259) Sharpless, K. B.; Amberg, W.; Bennani, Y. L.; Crispino, G. A.; Hartung, J.; Jeong, K. S.; Kwong, H. L.; Morikawa, K.; Wang, Z. M.; Xu, D.; et al. The osmium-catalyzed asymmetric dihydroxylation: A new Ligand class and a process improvement. *J. Org. Chem.* **1992**, *57* (10), 2768-2771
- (260) Wang, Z. M.; Kakiuchi, K.; Sharpless, K. B. Osmium-catalyzed asymmetric dihydroxylation of cyclic cis-disubstituted olefins. *J. Org. Chem.* **1994**, *59* (23), 6895-6897

References

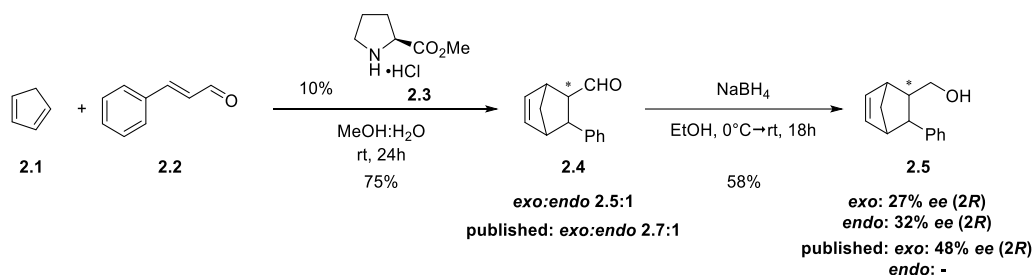
- (261) Soai, K.; Yokoyama, S.; Ebihara, K.; Hayasaka, T. A new chiral catalyst for the highly enantioselective addition of dialkylzinc reagents to aliphatic aldehydes. *J. Chem. Soc. Chem. Commun.* **1987**, (22), 1690-1691, DOI: 10.1039/C39870001690.
- (262) Ohga, T.; Umeda, S.; Kawanami, Y. Enantioselective addition of diethylzinc to aldehydes catalyzed by chiral amino alcohols. Substituent effect and nonlinear effect. *Tetrahedron* **2001**, *57* (23), 4825-4829. DOI: [http://dx.doi.org/10.1016/S0040-4020\(01\)00423-9](http://dx.doi.org/10.1016/S0040-4020(01)00423-9).
- (263) Kawanami, Y.; Mitsue, T.; Miki, M.; Sakamoto, T.; Nishitani, K. New Chiral Ligands Derived from (S)-Leucine for the Enantioselective Addition of Diethylzinc to Aldehydes. *Tetrahedron* **2000**, *56* (2), 175-178. DOI: [http://dx.doi.org/10.1016/S0040-4020\(99\)00971-0](http://dx.doi.org/10.1016/S0040-4020(99)00971-0).
- (264) Kitamura, M.; Suga, S.; Kawai, K.; Noyori, R. Catalytic asymmetric induction. Highly enantioselective addition of dialkylzinc to aldehydes. *J. Am. Chem. Soc.* **1986**, *108* (19), 6071-6072. DOI: 10.1021/ja00279a083.
- (265) Guijarro, D.; Pinho, P.; Andersson, P. G. Enantioselective Addition of Dialkylzinc Reagents to N-(Diphenylphosphinoyl) Imines Promoted by 2-Azanorbornylmethanols. *J. Org. Chem.* **1998**, *63* (8), 2530-2535. DOI: 10.1021/jo9718096.
- (266) Wassmann, S.; Wilken, J.; Martens, J. Synthesis and application of C2-symmetrical bis- β -amino alcohols based on the octahydro-cyclopenta[b]pyrrole system in the catalytic enantioselective addition of diethylzinc to benzaldehyde. *Tetrahedron Asym.* **1999**, *10* (22), 4437-4445. DOI: [http://dx.doi.org/10.1016/S0957-4166\(99\)00475-9](http://dx.doi.org/10.1016/S0957-4166(99)00475-9).
- (267) Yang, X.; Shen, J.; Da, C.; Wang, R.; Choi, M. C. K.; Yang, L.; Wong, K.-y. Chiral pyrrolidine derivatives as catalysts in the enantioselective addition of diethylzinc to aldehydes. *Tetrahedron Asym.* **1999**, *10* (1), 133-138. DOI: [http://dx.doi.org/10.1016/S0957-4166\(98\)00497-2](http://dx.doi.org/10.1016/S0957-4166(98)00497-2).
- (268) Lawrence, C. F.; Nayak, S. K.; Thijs, L.; Zwanenburg, B. N-Trityl-Aziridinyl(diphenyl)methanol as an Effective Catalyst in the Enantioselective Addition of Diethylzinc to Aldehydes. *Synlett* **1999**, *10*, 1571-1572.
- (269) Binder, C. M.; Bautista, A.; Zaidlewicz, M.; Krzemiński, M. P.; Oliver, A.; Singaram, B. Dual Stereoselectivity in the Dialkylzinc Reaction Using (-)- β -Pinene Derived Amino Alcohol Chiral Auxiliaries. *J. Org. Chem.* **2009**, *74* (6), 2337-2343. DOI: 10.1021/jo802371z.
- (270) Steiner, D.; Sethofer, S. G.; Goralski, C. T.; Singaram, B. Asymmetric addition of diethylzinc to aldehydes catalyzed by β -amino alcohols derived from limonene oxide. *Tetrahedron Asym.* **2002**, *13* (14), 1477-1483. DOI: [http://dx.doi.org/10.1016/S0957-4166\(02\)00342-7](http://dx.doi.org/10.1016/S0957-4166(02)00342-7).
- (271) Dai, W.-M.; Zhu, H.-J.; Hao, X.-J. Chiral ligands derived from abrine. Part 6: Importance of a bulky N-alkyl group in indole-containing chiral β -tertiary amino alcohols for controlling enantioselectivity in addition of diethylzinc toward aldehydes. *Tetrahedron Asym.* **2000**, *11* (11), 2315-2337. DOI: [http://dx.doi.org/10.1016/S0957-4166\(00\)00189-0](http://dx.doi.org/10.1016/S0957-4166(00)00189-0).
- (272) Joshi, S. N.; Malhotra, S. V. Enantioselective addition of diethylzinc to aldehydes catalyzed by a β -amino alcohol derived from (+)-3-carene. *Tetrahedron Asym.* **2003**, *14* (13), 1763-1766. DOI: [http://dx.doi.org/10.1016/S0957-4166\(03\)00347-1](http://dx.doi.org/10.1016/S0957-4166(03)00347-1).

References

- (273) Sibi, M. P.; Chen, J.-x.; Cook, G. R. Reversal of stereochemistry in diethylzinc addition to aldehydes by a simple change of the backbone substituent in L-serine derived ligands. *Tetrahedron Lett.* **1999**, *40* (17), 3301-3304. DOI: [http://dx.doi.org/10.1016/S0040-4039\(99\)00480-3](http://dx.doi.org/10.1016/S0040-4039(99)00480-3).
- (274) Shi, M.; Satoh, Y.; Masaki, Y. Chiral C₂-symmetric 2,5-disubstituted pyrrolidine derivatives as catalytic chiral ligands in the reactions of diethylzinc with aryl aldehydes. *J. Chem. Soc. Perkin Trans 1* **1998**, (16), 2547-2552, 10.1039/A803336F. DOI: 10.1039/A803336F.
- (275) Kossenjans, M.; Soeberdt, M.; Wallbaum, S.; Harms, K.; Martens, J.; Gunter Aurich, H. Utilization of industrial waste materials. Part 14.[dagger] Synthesis of [small beta]-amino alcohols and thiols with a 2-azabicyclo[3.3.0]octane backbone and their application in enantioselective catalysis. *J. Chem. Soc. Perkin Trans 1* **1999**, (16), 2353-2365, 10.1039/A902362C. DOI: 10.1039/A902362C.
- (276) Solà, L.; Reddy, K. S.; Vidal-Ferran, A.; Moyano, A.; Pericàs, M. A.; Riera, A.; Alvarez-Larena, A.; Piniella, J.-F. A Superior, Readily Available Enantiopure Ligand for the Catalytic Enantioselective Addition of Diethylzinc to α -Substituted Aldehydes. *J. Org. Chem.* **1998**, *63* (20), 7078-7082. DOI: 10.1021/jo981336i.
- (277) Paleo, M. R.; Cabeza, I.; Sardina, F. J. Enantioselective Addition of Diethylzinc to Aldehydes Catalyzed by N-(9-Phenylfluoren-9-yl) β -Amino Alcohols. *J. Org. Chem.* **2000**, *65* (7), 2108-2113. DOI: 10.1021/jo9917083.
- (278) Scott, J. W.; Keith, D. D.; Nix, G.; Parrish, D. R.; Remington, S.; Roth, G. R.; Townsend, J. M.; Valentine, D.; Yang, R. Catalytic asymmetric hydrogenation of methyl (E)- and (Z)-2-acetamido-3-alkylacrylates. *The Journal of Organic Chemistry* **1981**, *46* (25), 5086-5093. DOI: 10.1021/jo00338a007.
- (279) Tang, W.; Capacci, A. G.; White, A.; Ma, S.; Rodriguez, S.; Qu, B.; Savoie, J.; Patel, N. D.; Wei, X.; Haddad, N.; et al. Novel and Efficient Chiral Bisphosphorus Ligands for Rhodium-Catalyzed Asymmetric Hydrogenation. *Org. Lett.* **2010**, *12* (5), 1104-1107. DOI: 10.1021/ol1000999.
- (280) Knowles, W. S. Asymmetric hydrogenation. *Acc. Chem. Res.* **1983**, *16* (3), 106-112. DOI: 10.1021/ar00087a006.
- (281) Hopkins, J. M.; Dalrymple, S. A.; Parvez, M.; Keay, B. A. 3,3'-Disubstituted BINAP Ligands: Synthesis, Resolution, and Applications in Asymmetric Hydrogenation. *Org. Lett.* **2005**, *7* (17), 3765-3768. DOI: 10.1021/ol051413d.
- (282) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. Prediction of Enantioselectivity in Rhodium Catalyzed Hydrogenations. *J. Am. Chem. Soc.* **2009**, *131* (2), 410-411 DOI: 10.1021/ja806246h.
- (283) Hutcheon, J. A.; Chiolero A Fau - Hanley, J. A.; Hanley, J. A. Random measurement error and regression dilution bias. (1756-1833 (Electronic)). From 2010 Jun 23.
- (284) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132* (15), 154104. DOI: 10.1063/1.3382344.

Supplementary information Chapter 2

Scheme S2.4. Established procedure for carrying out the Diels-Alder cycloaddition with a secondary amine salt and obtaining the enantioselectivity from the alcohol.



Cyclopentadiene was distilled at 180°C from its dimer before using. Cinnamaldehyde was purified by silica gel chromatography (Hex: EA, 9:1) before using.

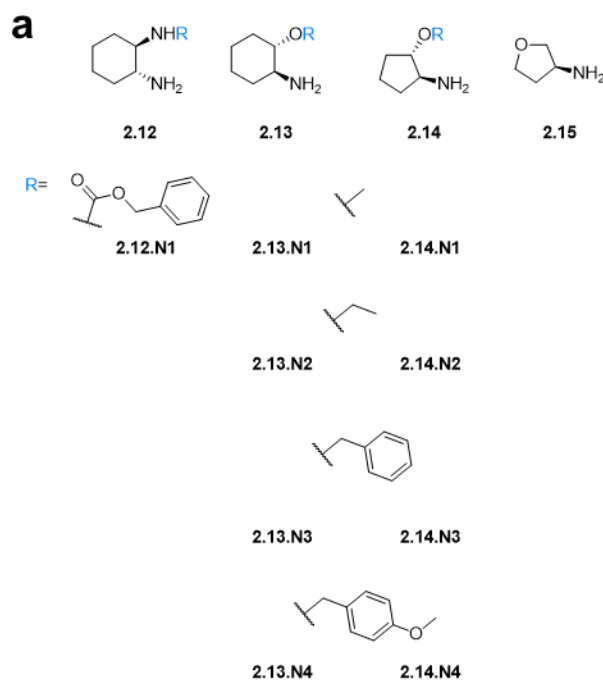
Cinnamaldehyde (3.369 mmol) was added to a solution of (L)-Pro-OMe·HCl (0.3369 mmol) in MeOH/H₂O 1M (95:5 v/v). The solution was stirred for 2 minutes before addition of the distilled diene (10.107 mmol). Reaction was monitored with TLC: toluene - iPrOH 1%, stained with KMnO₄. Cinnamaldehyde R_f=0.31, reaction adduct R_f=0.55. The reaction stirred at room temperature for 24h. After 24h, the reaction mixture was diluted with Et₂O and extracted from the aqueous layer. Organic layer was then washed successively with H₂O and brine, dried with Na₂SO₄, filtered and concentrated. Hydrolysis of the dimethyl acetal (from the MeOH) was performed by stirring the crude product mixture in TFA: H₂O: CHCl₃ (1: 1: 2, 0.5ml: 0.5ml: 1ml) for 2h at room temperature, followed by neutralization with sat. aq. NaHCO₃, addition of water, and extraction with Et₂O. Crude was then dried with Na₂SO₄, filtered and concentrated.

Purification of the Diels-Alder adduct from starting material was accomplished by silica gel chromatography (tol: iPrOH 1%). Reaction went with 82% conversion, and a d.r. of 2.5:1 *Exo:endo*.

Supplementary information Chapter 2

In order to determine the enantioselectivity of the catalyst, **2.4** was further reduced to its derived alcohol (**2.5**) using NaBH_4 .²¹²

Figure S2.5. Compounds used for virtual screening #3: **(a)** 10 different primary amine scaffolds. **(b)** 34 different aldehydes used with the primary amines to create a library of 340 chiral secondary amines, using REACT2D.



b

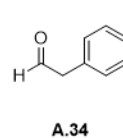
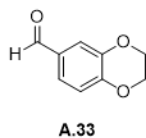
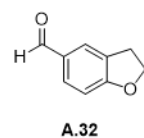
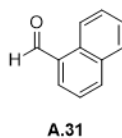
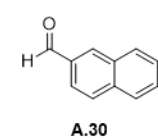
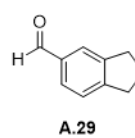
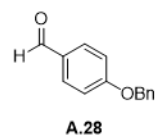
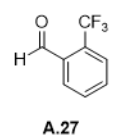
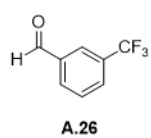
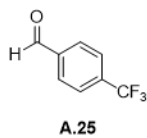
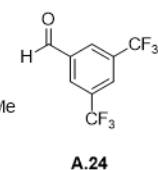
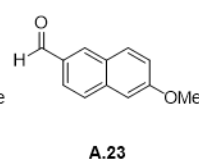
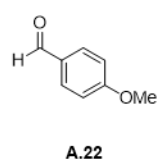
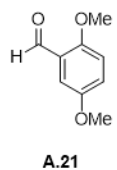
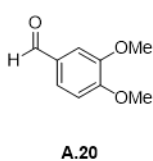
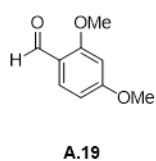
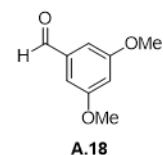
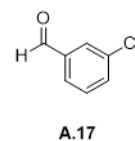
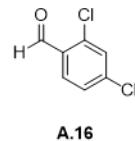
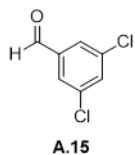
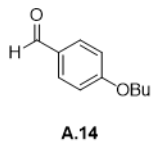
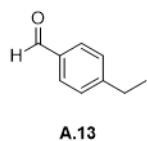
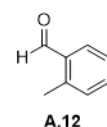
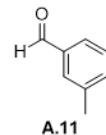
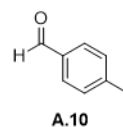
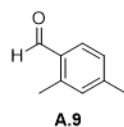
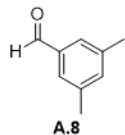
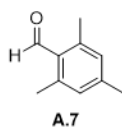
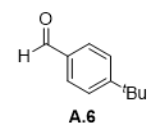
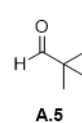
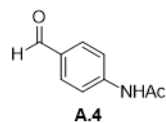
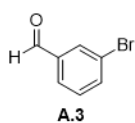
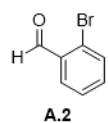
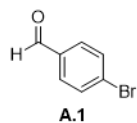


Figure S2.2. (a) Reaction scheme of chiral auxiliary asymmetric Diels-Alder cycloaddition. (b) Dienophiles. (c) Dienes.

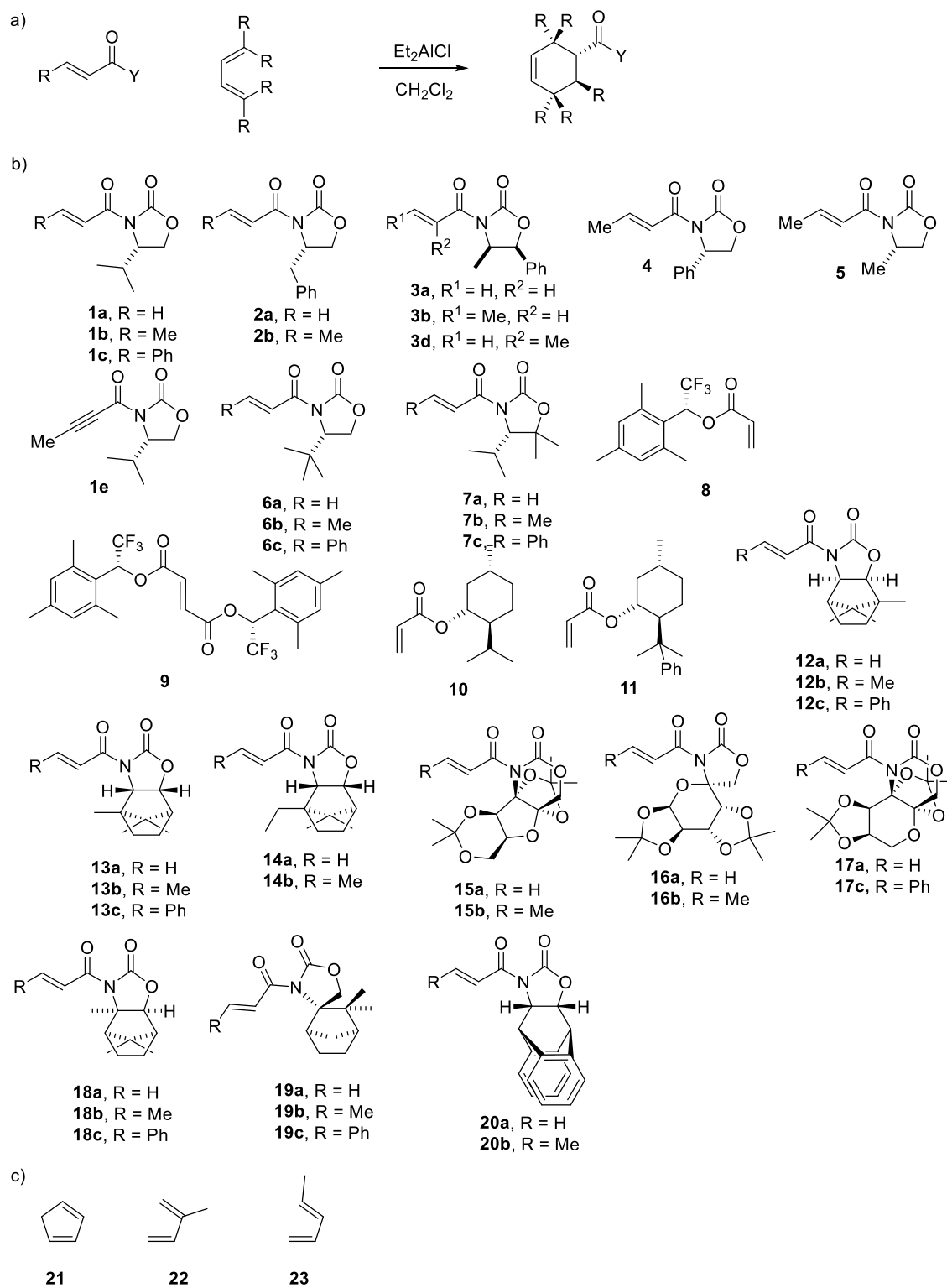


Table S2.1: Observed %*ee* (*endo* adducts only), temperature, and reference from which the data was collected.

Entry	Auxiliary	Dienophile	Diene	Temp. (°C)	Observed d.e. (%)	Ref.
1	1	a	21	-100	86 (<i>R</i>)	213
2	1	a	22	-100	66 (<i>R</i>)	213
3	1	b	21	-100	90 (<i>2R,3S</i>)	213
4	1	b	22	-100	68 (<i>S,S</i>)	213
5	1	c	21	-22	86 (<i>2S,3S</i>)	213
6	1	e	21	-22	50 (nd) ^b	213
7	2	a	21	-100	90 (<i>R</i>)	213
8	2	a	22	-100	90 (<i>R</i>)	213
9	2	a	23	-100	>99 (<i>R</i>)	213
10	2	b	21	-100	94 (<i>S,S</i>)	213
11	2	b	22	-30	88 (<i>S,S</i>)	213
12	2	b	23	-30	98 (<i>S,S</i>)	213
13	3	a	21	-100	90 (<i>S</i>)	213
14	3	b	21	-100	96 (<i>2S, 3R</i>)	213
15	3	d	21	-78	64 (nd)	213
16	4	b	22	-100	33 (<i>S,S</i>)	213
17	5	b	22	-100	58 (<i>S,S</i>)	213
18	6	a	22	-30	98 (<i>S,S</i>)	213
19	6	b	21	-100	99 (<i>2R,3S</i>)	213
19	6	b	22	-30	>98 (<i>S,S</i>)	213
21	6	c	21	-100	>99 (<i>S,S</i>)	213
22	7	a	22	-30	94 (<i>S,S</i>)	213
24	7	b	21	-100	99 (<i>2R,3S</i>)	213
23	7	b	22	-30	94 (<i>S,S</i>)	213
25	7	c	21	-100	>99 (<i>S,S</i>)	213
26 ^a	8	a	21	-78	95 (<i>S</i>)	214
27 ^a	9	-	21	-78	97 (<i>R,R</i>) ^b	214
28 ^a	10	a	21	0	47 (<i>R</i>)	215
29 ^a	11	a	21	0	64 (<i>R</i>)	215

Supplementary information Chapter 2

30	12	a	21	-78	96 (<i>S</i>)	216, 217
32	12	b	21	-78	97 (<i>2S,3R</i>)	216, 217
33	12	c	21	-22	98 (<i>R,R</i>)	216, 217
34	13	a	21	-78	81 (<i>R</i>)	218
35	13	b	21	-78	>99 (<i>2R,3S</i>)	218
36	13	c	21	-22	>99 (<i>S,S</i>)	218
37	14	a	21	-78	>95 (<i>R</i>)	218
38	14	b	21	-78	>99 (<i>2R,3S</i>)	218
39	15	a	21	-78	82 (<i>S</i>)	219
40	15	b	21	-22	94 (<i>2S,3R</i>)	219
41	16	a	21	-78	80 (<i>R</i>)	220
42	16	b	21	-78	89 (<i>2R,3S</i>)	220
43	17	a	21	-78	87 (<i>S</i>)	221
44	17	c	21	-22	63 (<i>2S,3R</i>)	221
45	18	a	21	-78	92 (<i>S</i>)	217
46	18	b	21	-78	94 (<i>2S,3R</i>)	217
47	18	c	21	-78	94 (<i>R,R</i>)	217
48	19	a	21	-78	98 (<i>R</i>)	217
49	19	b	21	-78	98 (<i>S,S</i>)	217
50	19	c	21	-78	98 (<i>2R,3S</i>)	217
51	20	a	21	-78	89 (<i>R</i>)	222
52	20	b	21	-78	96 (<i>R</i>)	222

^a Me₂AlCl is used in place of Et₂AlCl. ^b does not match with the template with CONSTRUCTS.

Figure S2.3. (a) Reaction scheme of Shi epoxidation reaction. **(b)** Chiral ketone catalysts **(c)** Alkenes.

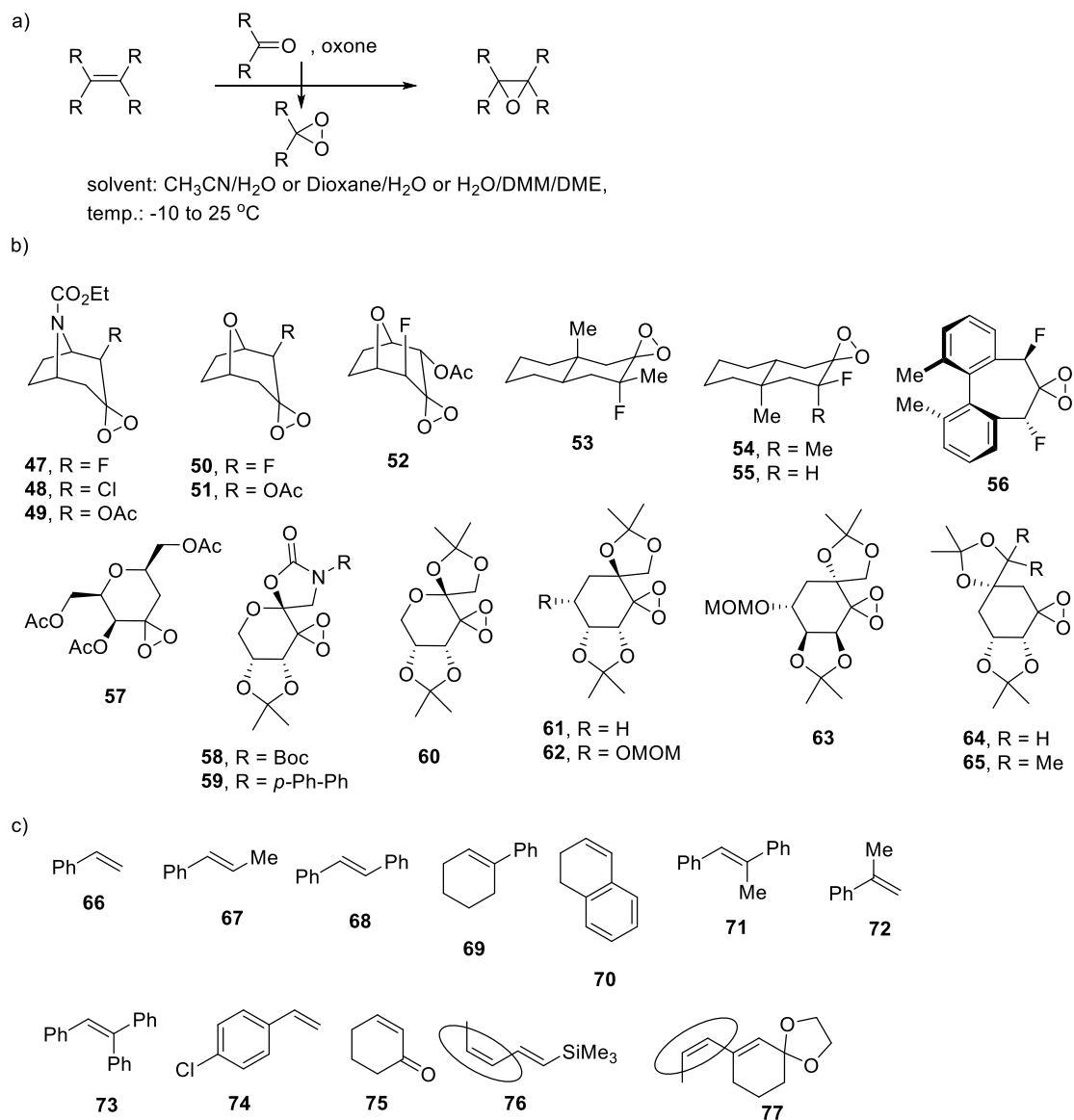


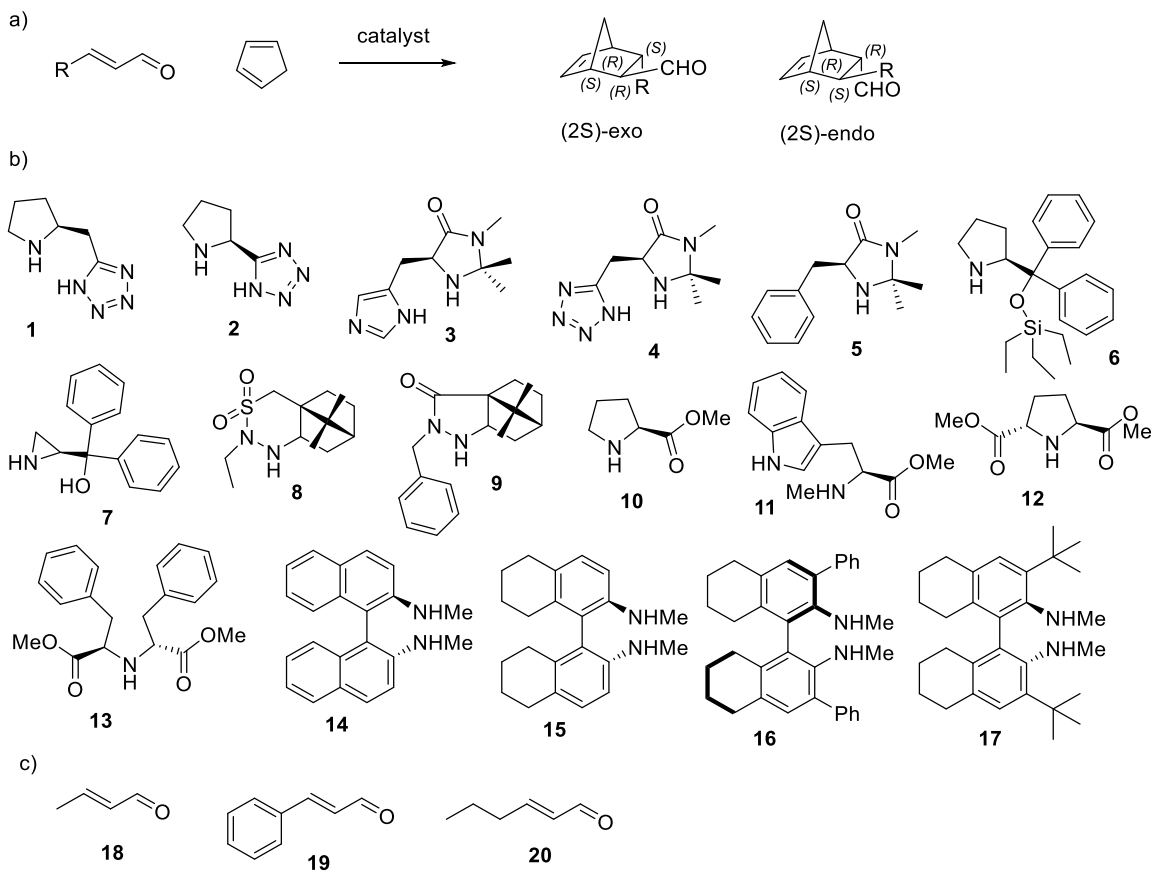
Table S2.2. Experimental information and references from which the data was collected.

Entry	Catalyst	Substrate	Temperature (°C)	Observed e.e. (%)	Ref
1	47	66	25	29 (<i>R</i>)	223
2	47	67	25	56 (<i>R,R</i>)	223
3	47	68	25	76 (<i>R,R</i>)	223
4	47	69	25	69 (<i>R,R</i>)	223
5	47	70	25	18 (<i>1R,2S</i>)	223
6	47	71	25	73 (<i>R,R</i>)	223
7	47	72	25	22 (<i>R</i>)	223
8	47	73	25	83 (<i>R</i>)	223
9	48	68	25	54 (<i>R,R</i>)	223
10	49	68	25	86 (<i>R,R</i>)	223
11	50	68	25	83 (<i>R,R</i>)	224, 225
12	51	66	25	48 (<i>R</i>)	226
13	51	68	25	93-95 (<i>R,R</i>)	224, 226
14	51	69	25	82 (<i>R,R</i>)	226
15	51	73	25	98 (<i>R,R</i>)	226
16	51	75	25	3 (<i>R,R</i>)	226
17	52	66	25	2 (<i>S</i>)	225
18	52	68	25	64 (<i>S,S</i>)	225
-	52	72	25	8 (<i>S</i>)	225
19	53	67	25	70 (<i>R,R</i>)	227
20	53	68	25	86 (<i>R,R</i>)	227
21	54	67	25	22 (<i>R,R</i>)	227
22	55	67	25	0 (<i>R,R</i>)	227
23	56	67	0	88 (<i>R,R</i>)	228
24	56	68	0	94 (<i>R,R</i>)	228
25	56	69	0	59 (<i>R,R</i>)	228
26	56	74	0	43 (<i>R</i>)	228
27	57	66	25	19 (<i>S</i>)	229
-	57	68	25	81 (<i>S,S</i>)	229
28	57	72q	25	7 (<i>S</i>)	229
29	58	77	-10	93 (<i>2S,3R</i>) ^a	230

Supplementary information Chapter 2

30	59	76	-10	92 (2 <i>S</i> ,3 <i>R</i>) ^a	230
31	60	66	-10	15-24 (<i>R</i>)	231
32	60	67	-10	88-95 (<i>R</i> , <i>R</i>)	231, 232
33	60	68	0	95-98 (<i>R</i> , <i>R</i>)	231, 232
34	60	69	-	91-98 (<i>R</i> , <i>R</i>)	231, 232
35	60	70	-	12-32 (1 <i>S</i> , <i>R</i>)	231
36	60	71	-	92-95 (<i>R</i> , <i>R</i>)	231, 232
37	60	72	-	20-28 (<i>S</i>)	231
38	60	73	-	92-96 (<i>R</i>)	231, 232
39	61	66	-10	15 (<i>R</i>)	233
40	61	67	-10	87 (<i>R</i> , <i>R</i>)	233
41	61	68	0	88 (<i>R</i> , <i>R</i>)	233
42	62	66	-10	31 (<i>R</i>)	233
43	62	67	-10	88 (<i>R</i> , <i>R</i>)	233
44	62	68	0	89 (<i>R</i> , <i>R</i>)	233
45	63	66	-10	24 (<i>S</i>)	233
46	63	67	-10	77 (<i>S</i> , <i>S</i>)	233
47	64	66	-10	14 (<i>R</i>)	233
48	64	67	-10	46 (<i>R</i> , <i>R</i>)	233
49	64	68	0	66 (<i>R</i> , <i>R</i>)	233
50	65	67	-10	38 (<i>R</i> , <i>R</i>)	233
51	65	68	0	72 (<i>R</i> , <i>R</i>)	233

^a Substrates containing 2 double bonds, CONSTRUCTS cannot differentiate double bonds by reactivity and will simply react the first one it finds. These substrates cannot be considered.

Figure S2.4. (a) Reaction scheme for the organocatalyzed Diels-Alder cycloaddition. (b) Chiral secondary amine catalysts used. (c) Dienophiles used.**Table S2.3.** Experimental information used and referenced.

Entry	Catalyst	Substrate		Observed e.e. (%)	Ref.
1	1	20	<i>exo</i>	3-17 (2 <i>S</i>)	234
2	1	20	<i>endo</i>	0-9 (2 <i>R</i>)	234
3	2	20	<i>exo</i>	3-36 (2 <i>S</i>)	234
4	2	20	<i>endo</i>	0-36 (2 <i>R</i>)	234
5	3	18	<i>exo</i>	66-89 (2 <i>R</i>)	234
6	3	18	<i>endo</i>	57-83 (2 <i>R</i>)	234
7	3	19	<i>exo</i>	68-84 (2 <i>S</i>)	234
8	3	19	<i>endo</i>	58-60 (2 <i>S</i>)	234
9	3	20	<i>exo</i>	69-72 (2 <i>S</i>)	234
10	3	20	<i>endo</i>	45-58 (2 <i>R</i>)	234

Supplementary information Chapter 2

11	4	18	<i>exo</i>	90 (2R)	234
12	4	18	<i>endo</i>	95 (2R)	234
13	4	19	<i>exo</i>	89 (2S)	234
14	4	19	<i>endo</i>	90 (2S)	234
15	4	20	<i>exo</i>	85 (2S)	234
16	4	20	<i>endo</i>	93 (2R)	234
17	5	18	<i>exo</i>	86 (2R)	234, 235
18	5	18	<i>endo</i>	90 (2R)	234, 235
19	5	19	<i>exo</i>	93 (2S)	234, 235
20	5	19	<i>endo</i>	93 (2S)	234, 235
21	5	20	<i>exo</i>	79-86 (2S)	234, 235
22	5	20	<i>endo</i>	90-96 (2R)	234, 235
23	6	19	<i>exo</i>	88 (4S)	212
24	6	19	<i>endo</i>	97 (4S)	212
25	7	18	<i>exo</i>	11-22 (nd) ^a	236
26	7	18	<i>endo</i>	10-24 (nd) ^a	236
27	7	19	<i>exo</i>	37-57 (nd) ^a	236
28	7	19	<i>endo</i>	36-66 (nd) ^a	236
29	8	18	<i>endo</i>	83 (2S)	237
30	8	19	<i>exo</i>	78 (2R)	237
31	8	19	<i>endo</i>	93 (2R)	237
32	8	20	<i>exo</i>	66 (2R)	237
33	8	20	<i>endo</i>	90 (2R)	237
34	9	19	<i>exo</i>	90 (2R)	238
35	9	19	<i>endo</i>	82-88 (2R)	238
36	9	20	<i>exo</i>	81 (2R)	238
37	10	19	<i>exo</i>	48 (2R)	235
38	11	19	<i>exo</i>	59 (2S)	235
39	12	19	<i>exo</i>	57 (2R)	235
40	13	19	<i>exo</i>	74 (2R)	235
41	14	19	<i>exo</i>	9 (R)	239
42	14	19	<i>endo</i>	15 (R)	239
43	15	19	<i>exo</i>	38 (S)	239
44	15	19	<i>endo</i>	27 (R)	239

Supplementary information Chapter 2

45	16	19	<i>exo</i>	53 (<i>R</i>)	239
46	16	19	<i>endo</i>	39 (<i>R</i>)	239
47	17	18	<i>exo</i>	88 (<i>S</i>)	239
48	17	19	<i>exo</i>	72-92 (<i>R</i>)	239
49	17	19	<i>endo</i>	68-91 (<i>R</i>)	239

^a Absolute stereochemistry not assigned.

Figure S2.5. (a) Reaction scheme for the organocatalyzed aldol reaction. (b) chiral secondary amine catalysts used. (c) aldehydes used.

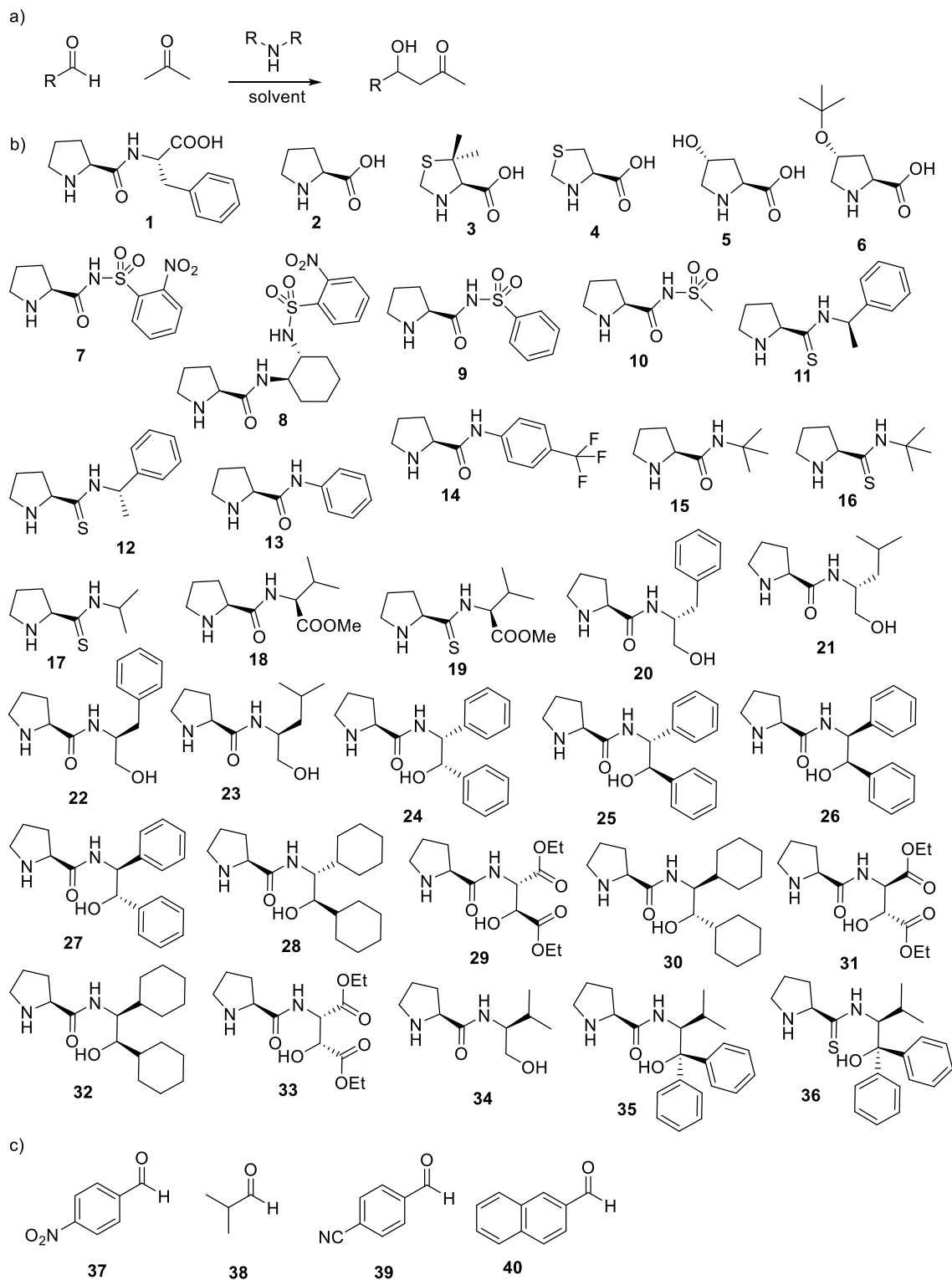
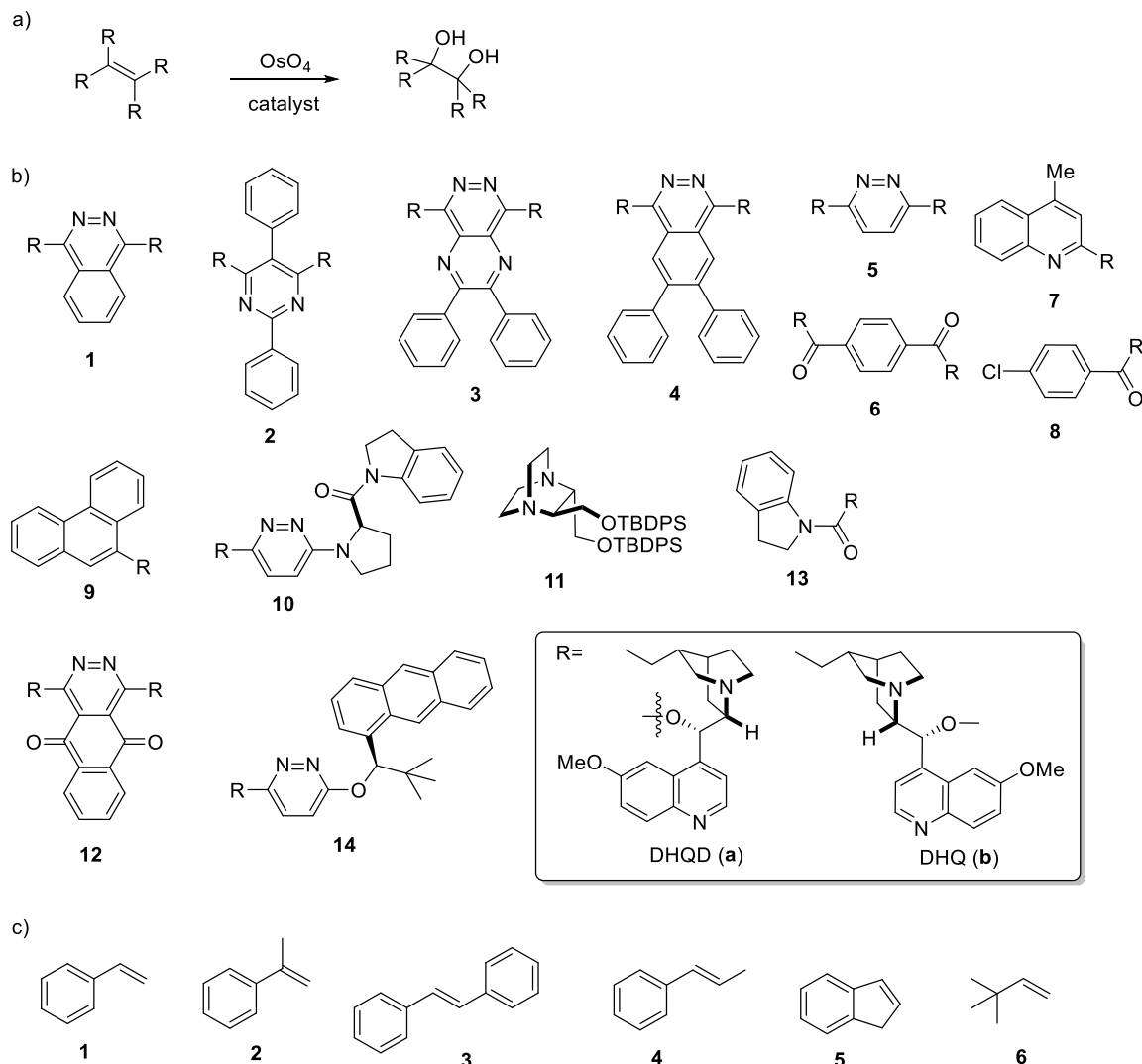


Table S2.4. Experimental information and references.

Entry	Catalyst	Substrate	Observed e.e. (%)	Ref.
1	1	37	59-73 (<i>R</i>)	240
2	1	38	77 (<i>R</i>)	240
3	2	37	76 (<i>R</i>)	241, 242
4	2	38	96 (<i>R</i>)	241, 242
5	2	40	77 (<i>R</i>)	241, 242
6	3	37	86 (<i>R</i>)	241, 242
7	3	38	94 (<i>R</i>)	241, 242
8	3	40	88 (<i>R</i>)	241, 242
9	4	37	73 (<i>R</i>)	241, 242
10	5	37	78 (<i>R</i>)	241, 242
11	6	37	62 (<i>R</i>)	241, 242
12	7	37	95 (<i>R</i>)	243
13	7	40	95 (<i>R</i>)	243
14	8	37	81 (<i>R</i>)	243
15	9	37	61-92 (<i>R</i>)	244
16	10	37	44-87 (<i>R</i>)	244
17	11	39	72-93 (<i>R</i>)	245
18	12	37	84-94 (<i>R</i>)	245
19	12	39	54-93 (<i>R</i>)	245
20	12	40	81 (<i>R</i>)	245
21	13	37	37 (<i>R</i>)	246
22	14	37	45 (<i>R</i>)	246
23	15	39	15 (<i>R</i>)	246
24	16	39	27 (<i>R</i>)	247
25	17	39	75 (<i>R</i>)	247
26	18	37	22-35 (<i>R</i>)	247
27	19	37	77-85 (<i>R</i>)	247
28	20	37	46 (<i>R</i>)	248
29	21	37	33 (<i>R</i>)	248

Supplementary information Chapter 2

30	22	37	48 (<i>R</i>)	248
31	23	37	52 (<i>R</i>)	248
32	24	37	49 (<i>R</i>)	248
33	25	37	44 (<i>R</i>)	248
34	26	37	64 (<i>R</i>)	248
35	27	37	69-93 (<i>R</i>)	248
36	27	38	98 (<i>R</i>)	248
37	27	40	84 (<i>R</i>)	248
38	28	37	15 (<i>R</i>)	249
39	29	37	72 (<i>R</i>)	249
40	30	37	15 (<i>R</i>)	249
41	31	37	87-99 (<i>R</i>)	249
42	31	38	>99 (<i>R</i>)	249
43	31	39	99 (<i>R</i>)	249
44	32	37	17 (<i>R</i>)	249
45	33	37	67 (<i>R</i>)	249
46	34	37	96 (<i>R</i>)	250
47	35	37	60 (<i>R</i>)	250
48	36	37	95 (<i>R</i>)	250
49	36	40	96 (<i>R</i>)	250

Figure S2.6. (a) Reaction scheme for the OsO₄ catalyzed dihydroxylation. (b) Chiral ligands used. (c) alkenes used.

Table S2.5. Experimental information and references.

Entry	Catalyst	Substrate	Observed e.e. (%)	Ref.
1	1a	1	97 (<i>R</i>)	251, 252
2	1b	1	97 (<i>S</i>)	251, 252
3	2a	1	80 (<i>R</i>)	251
4	3a	1	99 (<i>R</i>)	251
5	3b	1	97 (<i>S</i>)	251
6	4a	1	98 (<i>R</i>)	251
7	4b	1	96 (<i>S</i>)	251

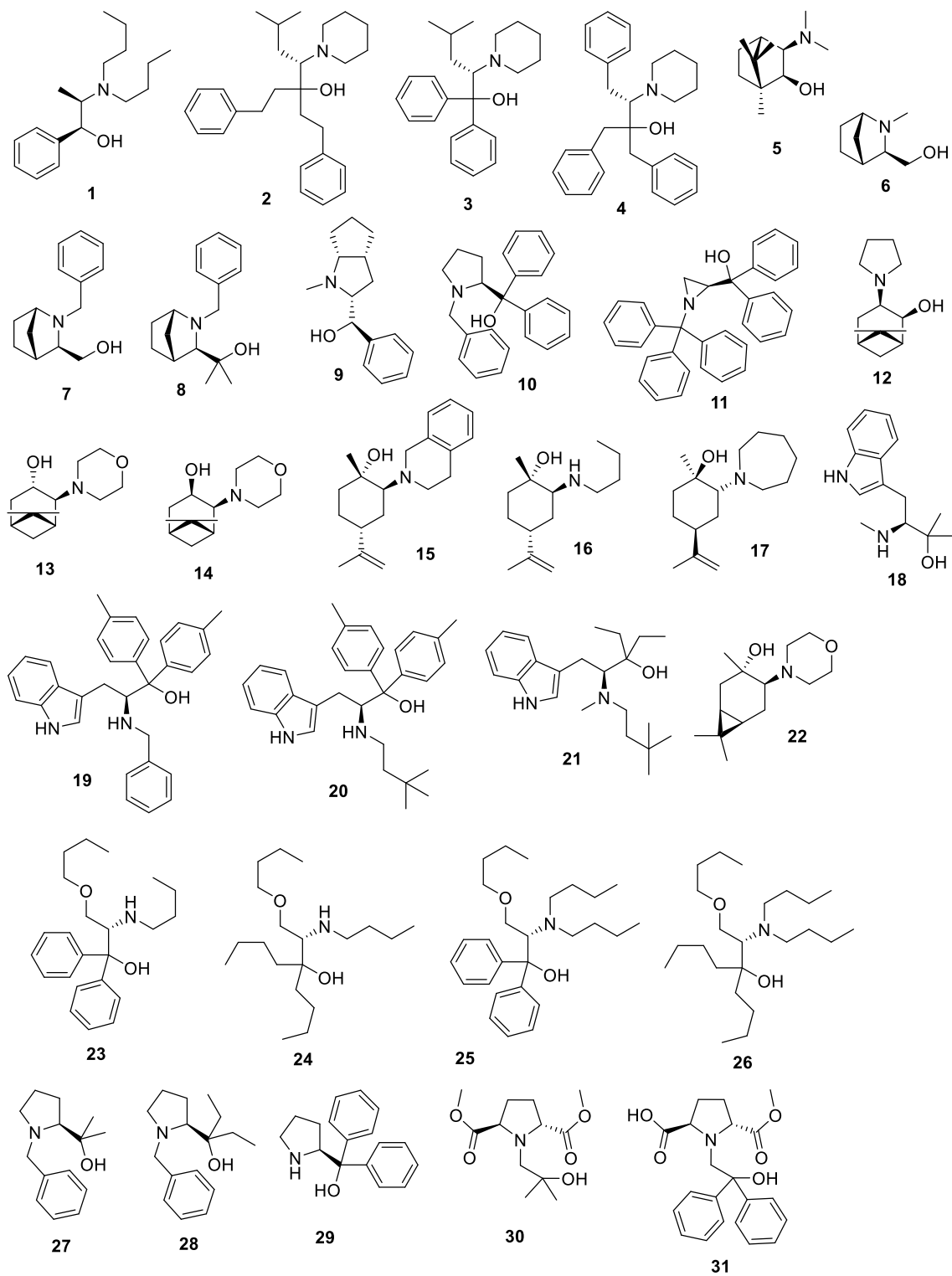
Supplementary information Chapter 2

8	5a	1	96 (<i>R</i>)	252, 253
9	5b	1	93 (<i>S</i>)	252
10	6a	1	80-92 (<i>R</i>)	252
11	6b	1	79-85 (<i>S</i>)	252
12	7a	1	87 (<i>R</i>)	254
13	8a	1	60-74 (<i>R</i>)	255 254
14	8b	1	54 (<i>S</i>)	255
15	9a	1	78 (<i>R</i>)	254
16	10	1	93 (<i>R</i>)	256
17	11	1	21 (<i>S</i>)	257
18	12a	1	89 (<i>R</i>)	258
19	12b	1	85 (<i>S</i>)	258
20	14	1	76 (<i>R</i>)	253
21	1a	2	94 (<i>R</i>)	251
22	1b	2	93 (<i>S</i>)	251
23	2a	2	69 (<i>R</i>)	251
24	3a	2	96 (<i>R</i>)	251
25	3b	2	92 (<i>S</i>)	251
26	4a	2	94 (<i>R</i>)	251
27	4b	2	94 (<i>S</i>)	251
28	5a	2	93 (<i>R</i>)	252
29	6a	2	72 (<i>R</i>)	252
30	8a	2	33 (<i>R</i>)	255
31	12a	2	82 (<i>R</i>)	258
32	1a	3	>99.5 (<i>R,R</i>)	259
33	1b	3	>99.5 (<i>S,S</i>)	259
34	5a	3	99 (<i>R,R</i>)	253
35	7a	3	98 (<i>R,R</i>)	254
36	8a	3	85-88 (<i>R,R</i>) 99 (<i>R,R</i>)	255 254
37	8b	3	78 (<i>S,S</i>)	255
38	9a	3	99 (<i>R,R</i>)	254
39	10	3	99.5 (<i>R,R</i>)	256
40	11	3	40 (<i>S,S</i>)	257

Supplementary information Chapter 2

41	14	3	92 (<i>R,R</i>)	253
42	8a	4	65 (<i>R,R</i>)	255
43	8b	4	55 (<i>S,S</i>)	255
44	10	4	97 (<i>R,R</i>)	256
45	11	4	19 (<i>S,S</i>)	257
46	12a	4	92 (<i>R,R</i>)	258
47	1a	5	42 (<i>1R,2S</i>)	251, 260
48	2a	5	35 (<i>1R,2S</i>)	251, 260
51	3a	5	20 (<i>1R,2S</i>)	251
50	4a	5	53 (<i>1R,2S</i>)	251
51	11	5	12 (<i>1S 2R</i>)	257
52	12a	5	63 (<i>1R,2S</i>)	258
53	1a	6	64	-
54	1b	6	66	-
55	2a	6	92	-
56	2b	6	87	-
57	3a	6	59	-
58	3b	6	65	-
59	4a	6	67	-
60	4b	6	73	-
61	7a	6	79 (<i>R</i>)	-
62	8a	6	44 (<i>R</i>)	-
63	9a	6	79 (<i>R</i>)	-

Figure S2.7. Chiral ligands used for the diethyl zinc addition to aldehydes.



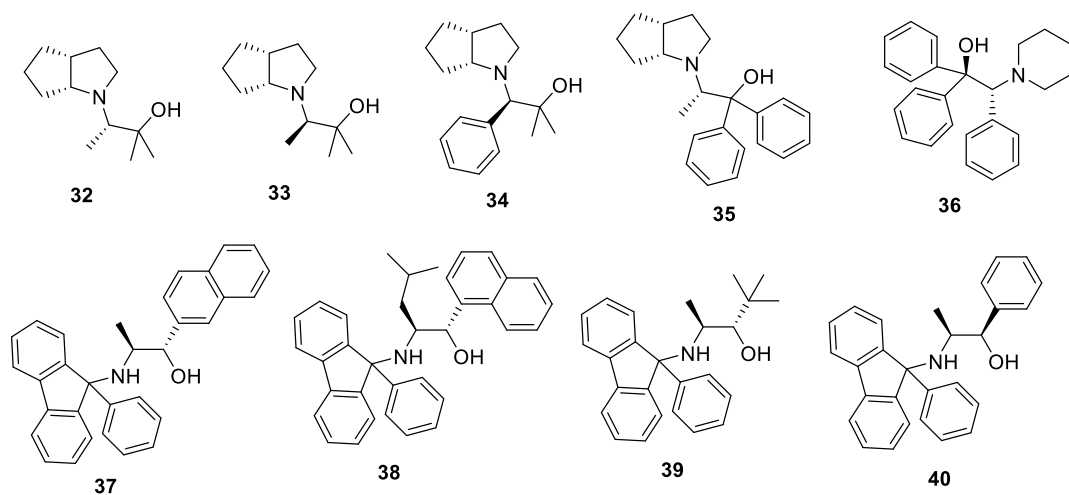


Table S2.6. Experimental information and references.

Entry	Catalyst	Substrate	Observed e.e. (%)	Ref.
1	1	41a	90 (<i>S</i>)	261
2	1	41c	78 (<i>S</i>)	261
3	1	41d	88 (<i>S</i>)	261
4	2	41a	97 (<i>R</i>)	262, 263
5	2	41c	95 (<i>R</i>)	262, 263
6	2	41d	88 (<i>R</i>)	262, 263
7	3	41a	85 (<i>R</i>)	262, 263
8	3	41d	78 (<i>R</i>)	262, 263
9	4	41a	67 (<i>R</i>)	262
10	5	41a	98 (<i>S</i>)	264
11	5	41b	90 (<i>S</i>)	264
12	5	41d	61 (<i>S</i>)	264
13	6	41a	24 (<i>R</i>)	265
14	7	41a	75 (<i>S</i>)	265
15	8	41a	24 (<i>S</i>)	265
16	9	41a	68 (<i>R</i>)	266
17	10	41a	81.8 (<i>S</i>)	267
18	11	41a	99 (<i>S</i>) ^a	268
19	11	41c	99 (<i>S</i>) ^a	268

Supplementary information Chapter 2

20	11	41d	80 (<i>S</i>) ^a	268
21	12	41a	84 (<i>S</i>)	269
22	13	41a	58 (<i>R</i>)	269
23	14	41a	99 (<i>R</i>)	269
24	15	41a	78 (<i>R</i>)	270
25	16	41a	48 (<i>R</i>)	270
26	17	41a	80 (<i>S</i>)	270
27	18	41a	31 (<i>R</i>)	271
28	19	41a	36.7 (<i>S</i>)	271
29	20	41a	57.9 (<i>R</i>)	271
30	21	41a	87.5 (<i>R</i>)	271
31	22	41a	81 (<i>R</i>)	272
32	22	41c	93 (<i>R</i>)	272
33	23	41a	66 (<i>S</i>)	273
34	23	41b	70 (<i>S</i>)	273
35	24	41a	28 (<i>R</i>)	273
36	25	41a	26 (<i>R</i>)	273
37	26	41a	79 (<i>R</i>)	273
38	26	41b	69 (<i>R</i>)	273
39	27	41a	96.7 (<i>R</i>)	267
40	28	41a	95.9 (<i>R</i>)	267
41	29	41a	36.1 (<i>R</i>)	267
42	30	41a	61 (<i>R</i>)	274
43	31	41a	91 (<i>R</i>)	274
44	32	41a	36 (<i>R</i>)	275
45	33	41a	80 (<i>S</i>)	275
46	34	41a	80 (<i>S</i>)	275
47	35	41a	86 (<i>R</i>)	275
48	36	41a	98 (<i>S</i>)	276
49	36	41b	93 (<i>S</i>)	276
50	36	41c	98 (<i>S</i>)	276
51	36	41d	92 (<i>S</i>)	276
52	37	41a	50 (<i>S</i>)	277
53	38	41a	24 (<i>S</i>)	277

Supplementary information Chapter 2

54	39	41a	97 (<i>S</i>)	277
55	40	41a	29 (<i>R</i>)	277

^a This catalyst features an aziridine which does not match with the template used in Constructs.

Figure S2.8. Ruthenium ligands and substrates used in the hydrogenation reaction.

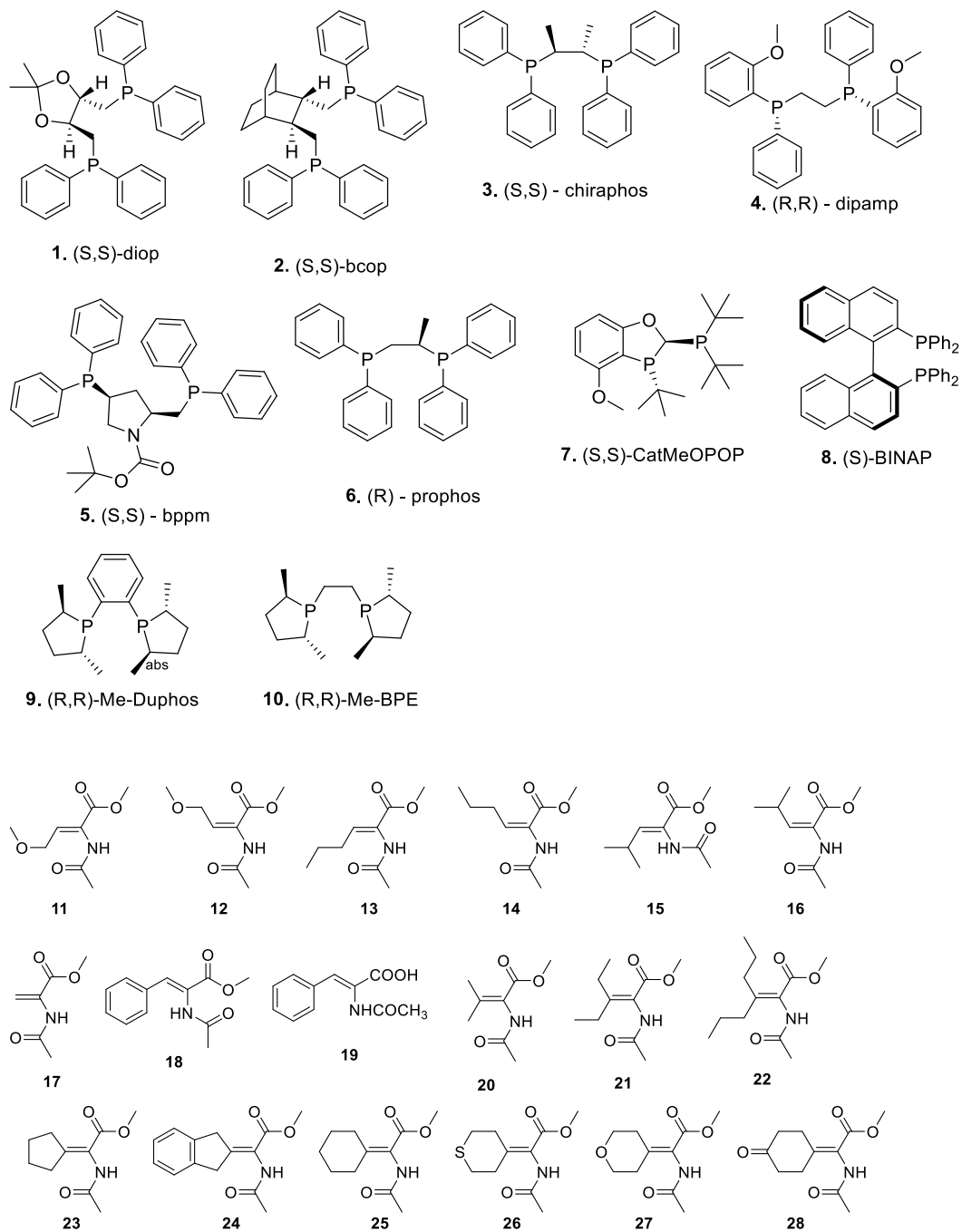


Table S2.7. Experimental information and references.

Entry	Catalyst	Substrate	Observed e.e. (%) ^a	Ref.
1	1	11	29 (<i>S</i>)	278
2	2	11	30 (<i>S</i>)	278
3	3	11	87 (<i>R</i>)	278
4	4	11	86 (<i>S</i>)	278
5	5	11	37 (<i>R</i>)	278
6	6	11	74 (<i>S</i>)	278
7	1	12	67 (<i>S</i>)	278
8	2	12	78 (<i>S</i>)	278
9	3	12	36 (<i>R</i>)	278
10	4	12	94 (<i>S</i>)	278
11	5	12	82 (<i>R</i>)	278
12	6	12	60 (<i>S</i>)	278
13	1	13	30 (<i>S</i>)	278
14	2	13	17 (<i>S</i>)	278
15	3	13	82 (<i>R</i>)	278
16	4	13	96 (<i>S</i>)	278
17	5	13	80 (<i>R</i>)	278
18	6	13	88 (<i>S</i>)	278
19	1	14	49 (<i>S</i>)	278
20	2	14	55 (<i>S</i>)	278
21	3	14	22 (<i>R</i>)	278
22	4	14	95 (<i>S</i>)	278
23	5	14	75 (<i>R</i>)	278
24	6	14	53 (<i>S</i>)	278
25	1	15	18 (<i>S</i>)	278
26	2	15	8 (<i>S</i>)	278
27	3	15	71 (<i>R</i>)	278
28	4	15	65 (<i>S</i>)	278
29	5	15	57 (<i>R</i>)	278
30	6	15	81 (<i>S</i>)	278
31	1	16	54 (<i>S</i>)	278

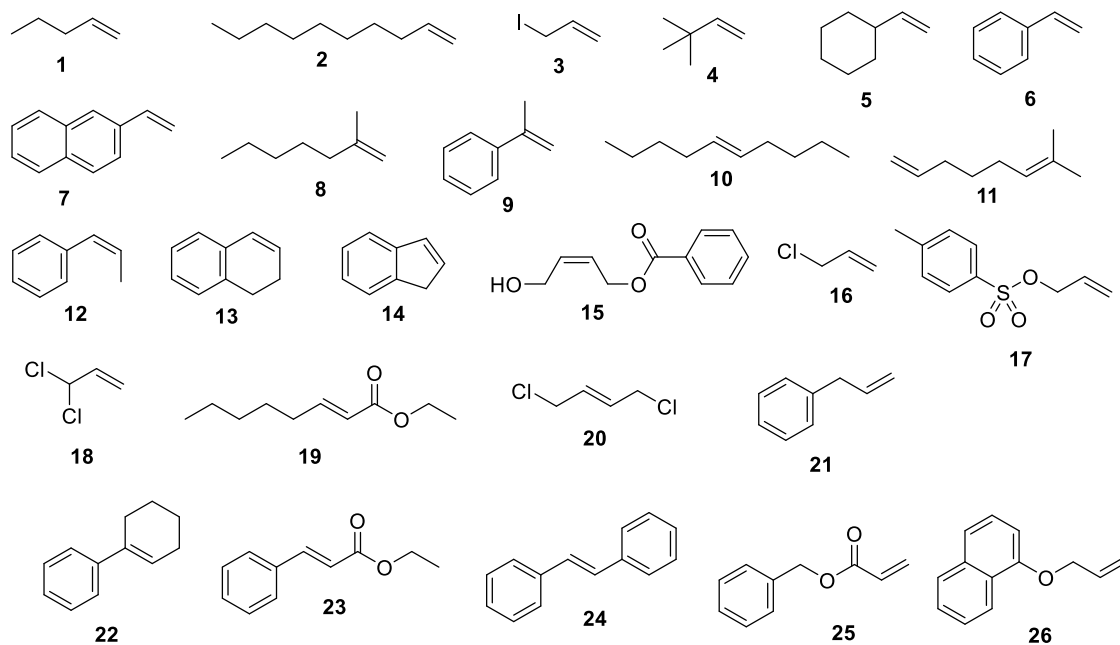
Supplementary information Chapter 2

32	2	16	24 (<i>S</i>)	278
33	3	16	24 (<i>R</i>)	278
34	4	16	78 (<i>S</i>)	278
35	5	16	72 (<i>R</i>)	278
36	6	16	67 (<i>S</i>)	278
37	1	17	56 (<i>S</i>)	278
38	2	17	45 (<i>S</i>)	278
39	3	17	79 (<i>R</i>)	278
40	4	17	95 (<i>S</i>)	278
41	5	17	82 (<i>R</i>)	278
42	6	17	80 (<i>S</i>)	278
43	7	17	98 (<i>S</i>) ^b	279
44	1	18	49 (<i>S</i>)	278
45	4	18	96 (<i>S</i>)	278, 280
46	5	18	95 (<i>R</i>)	278
47	7	18	99 (<i>S</i>) ^b	279
48	8	18	21 (<i>R</i>) ^b	281
49	1	19	83 (<i>S</i>)	280
50	3	19	95 +(<i>R</i>)	280
51	4	19	95 (<i>S</i>)	280
52	5	19	91 (<i>R</i>)	280
53	7	19	99 (<i>S</i>) ^b	279
54	8	19	15 (<i>R</i>) ^b	281
55	4	20	55 (<i>S</i>)	278
56	5	20	0	278
57	9	20	96 (<i>R</i>)	282
58	10	20	98.2 (<i>R</i>)	282
59	9	21	96.2 (<i>R</i>)	282
60	10	21	97.5 (<i>R</i>)	282
61	9	22	85.1 (<i>R</i>)	282
62	10	22	96.8 (<i>R</i>)	282
63	9	23	96.8 (<i>R</i>)	282
64	10	23	97.2 (<i>R</i>)	282
65	9	24	99.0 (<i>R</i>)	282

Supplementary information Chapter 2

66	10	24	98.6 (<i>R</i>)	282
67	9	25	96.2 (<i>R</i>)	282
68	10	25	98.6 (<i>R</i>)	282
69	9	26	95.0 (<i>R</i>)	282
70	10	26	98.4 (<i>R</i>)	282
71	9	27	98.2 (<i>R</i>)	282
72	10	27	98.6 (<i>R</i>)	282
73	9	28	93.7 (<i>R</i>)	282
74	10	28	98.0 (<i>R</i>)	282

^a Stereochemistry inverted when assigned by CONSTRUCTS. This results from the presence of Rhodium in the TS where the product will have a hydrogen. ^b These substrates do not match with the template used in CONSTRUCTS.

Figure S2.9. Substrates set used in the substrate scope study with (DHQD)₂PHAL.^a

Table S2.8. Experimental information and references.

Entry	Substrate	Observed e.e. (%)	Ref.
1	1	79 (<i>R</i>)	251
2	2	84 (<i>R</i>)	251
3	3	63 (<i>S</i>)	251
4	4	64 (<i>R</i>)	251
5	5	88 (<i>R</i>)	251
6	6	97 (<i>R</i>)	251
7	7	99 (<i>R</i>)	251
8	8	78 (<i>R</i>)	251
9	9	94 (<i>R</i>)	251
10	10	97 (<i>R,R</i>)	251
11	11^b	98 (<i>R</i>)	251
12	12	35 (1 <i>R</i> ,2 <i>S</i>)	251
13	13	15 (1 <i>R</i> ,2 <i>S</i>)	251
14	14	42 (1 <i>R</i> ,2 <i>S</i>)	251
15	15	64 (1 <i>S</i> ,2 <i>R</i>)	251
16	16	63 (<i>S</i>)	258

Supplementary information Chapter 2

17	17	40 (<i>S</i>)	258
18	18	63 (<i>S</i>)	258
19	19	99 (2 <i>S</i> ,3 <i>R</i>)	258
20	20	94 (<i>S</i> , <i>S</i>)	258
21	21	44 (<i>R</i>)	258
22	22	99 (<i>R</i> , <i>R</i>)	259
23	23	97 (2 <i>S</i> , 2 <i>R</i>)	259
24	24	>99.5 (<i>R</i> , <i>R</i>)	259
25	25	77 (<i>S</i>)	259
26	26	91 (<i>S</i>)	259

^a (DHQD)₂PHAL is catalyst **1a** in Table D5. ^b the two double bonds cannot be distinguished by CONSTRUCTS. This substrate was discarded.

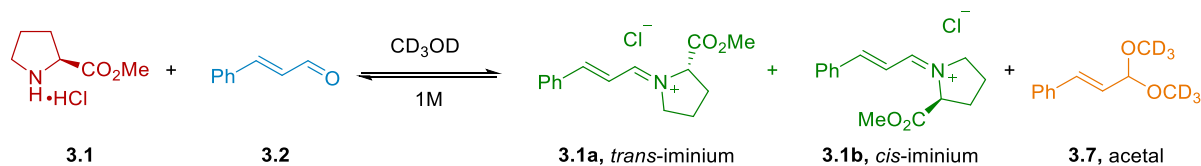
Supplementary information Chapter 3

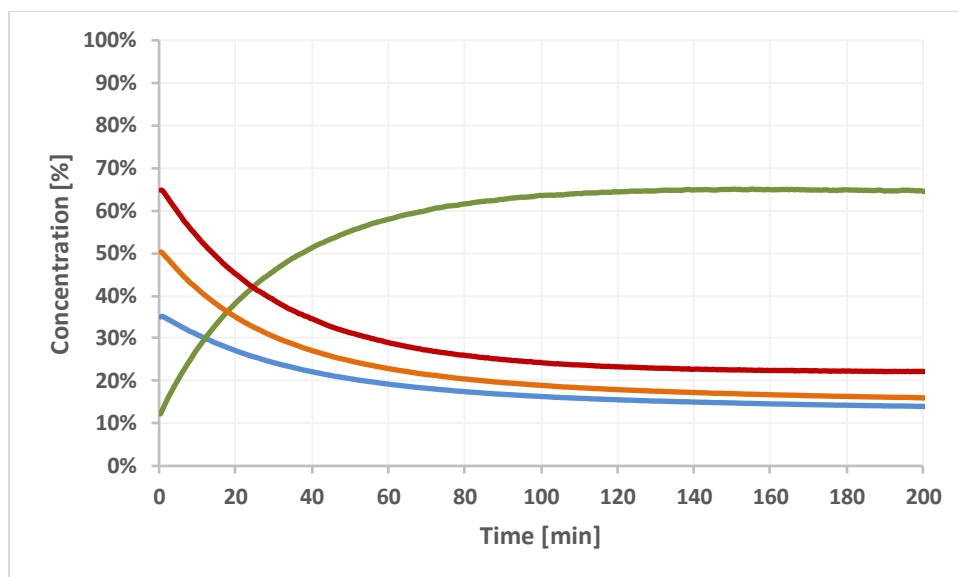
Time-dependent experiment for the formation of 3.1a and 3.1b in CD₃OD with 3.1

Methyl L-prolinate hydrochloride (**3.1**) (106.5 mg, 0.65mmol) dissolved in 0.3 μL of CD₃OD was added to a solution of cinnamaldehyde (**3.2**) (82 μL, 0.65 mmol) in 0.35 μL of CD₃OD. The time-dependent experiment was conducted on a Varian VNMRS 500 MHz NMR spectrometer where ns = 8, SW = 16ppm, RG = 20 and temp = 25°C. The initial experiment was reported as t = 0.4 min. The concentration vs time curve (**Figure S3.1**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0.4 min)

When t = 0.4 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.1a** & **3.1b**)] = 1
Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.63-9.71, acetal (**3.7**) 6.12-6.18, *trans*- and *cis*-iminium (**3.1a** & **3.1b**) 8.86-9.01 & 9.02-9.14, amine (**3.1**) 3.82-3.84.

Figure S3.1. Concentration vs time curve for the formation of 3.1a and 3.1b with 3.1





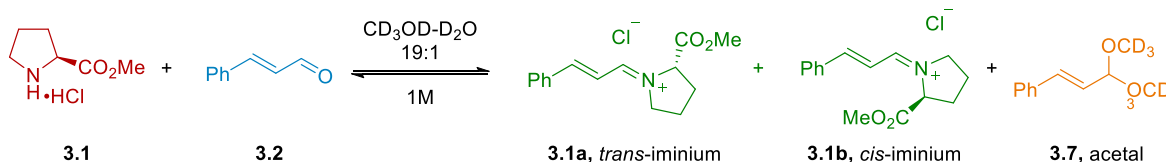
Time-dependent experiment for the formation of **3.1a** and **3.1b** in CD₃OD: D₂O (19:1) **3.1**

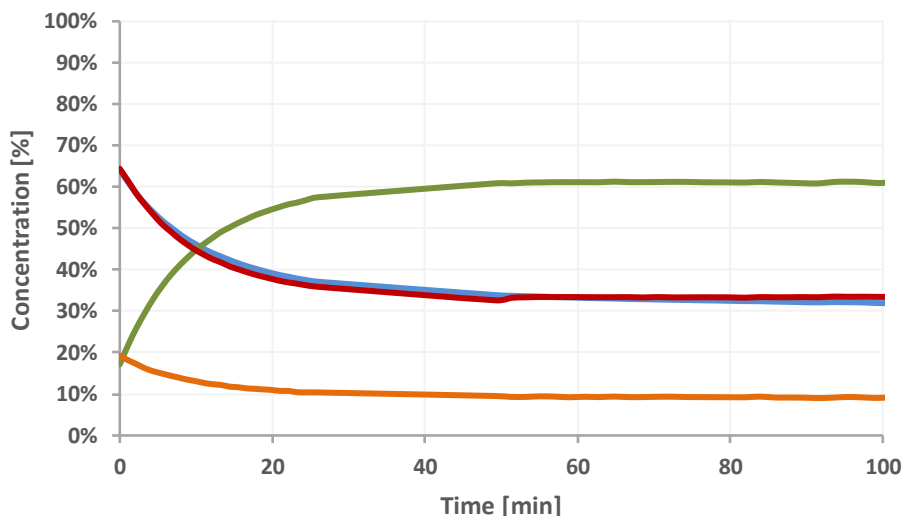
Methyl L-prolinate hydrochloride (**3.1**) (108 mg, 0.65 mmol) dissolved in 0.3 μL of CD₃OD: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (82 μL, 0.65 mmol) in 0.35 μL of CD₃OD: D₂O (19:1). The time-dependent experiment was conducted on a Varian VNMRS 500 MHz NMR spectrometer where ns = 8, SW = 16 ppm, RG = 20 and temp = 25°C. The initial experiment was reported as t = 0 min. The concentration vs time curve (**Figure S3.2**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.1a** & **3.1b**)] = 1

Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.63-9.70, acetal (**3.7**) 4.89-4.95, *trans*- and *cis*-iminium (**3.1a** & **3.1b**) 8.87-9.15, amine (**3.1**) 3.83-3.85.

Figure S3.2. Concentration vs time curve for the formation of **3.1a** and **3.1b** with **3.1**.





Time-dependent experiment for the formation of **3.1a** and **3.1b** in CD₃CN: D₂O (19:1) with **3.1**.

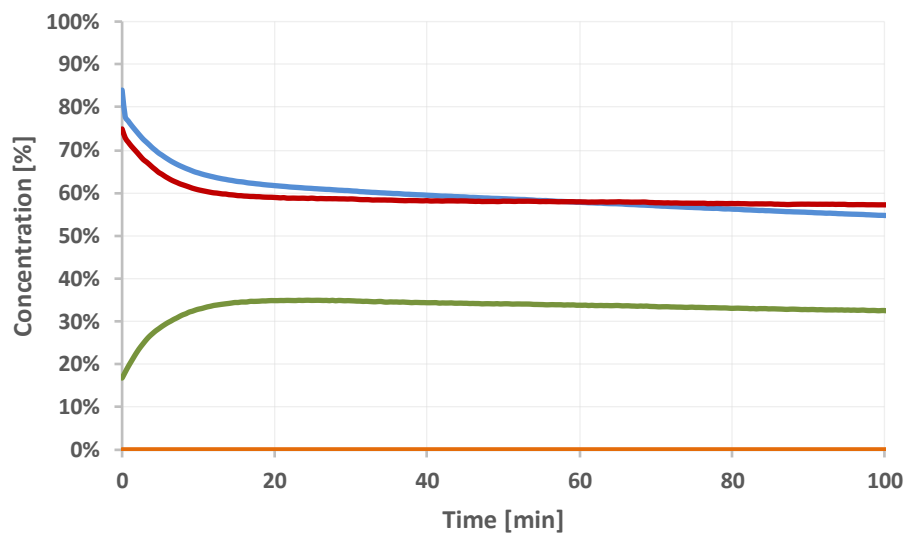
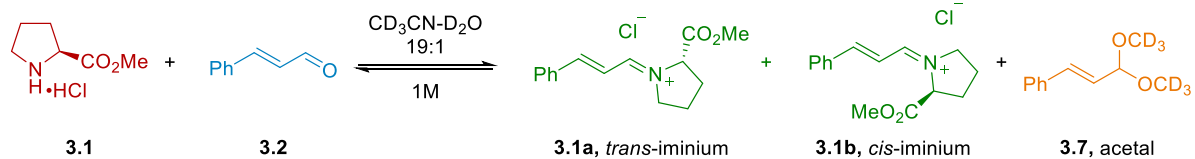
Methyl L-prolinate hydrochloride (**3.1**) (111.6 mg, 0.67mmol) dissolved in 0.3 μ L of CD₃CN: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (82 μ L, 0.65 mmol) in 0.35 μ L of CD₃CN: D₂O (19:1). The time-dependent experiment was conducted on a Varian VNMRS 500 MHz NMR spectrometer where ns = 8, SW = 16 ppm, RG = 16 and temp = 25°C. The initial experiment was reported as t = 0 min. The concentration vs time curve (**Figure S3.3**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.1a** & **3.1b**)] = 1

Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.64-9.70, *trans*- and *cis*-iminium (**3.1a** & **3.1b**) 8.93-9.23, amine (**3.1**) 3.79-3.82.

Figure S3.3. Concentration vs time curve for the formation of **3.1a** and **3.1b** with **3.1** in acetonitrile and water:

Supplementary information Chapter 3

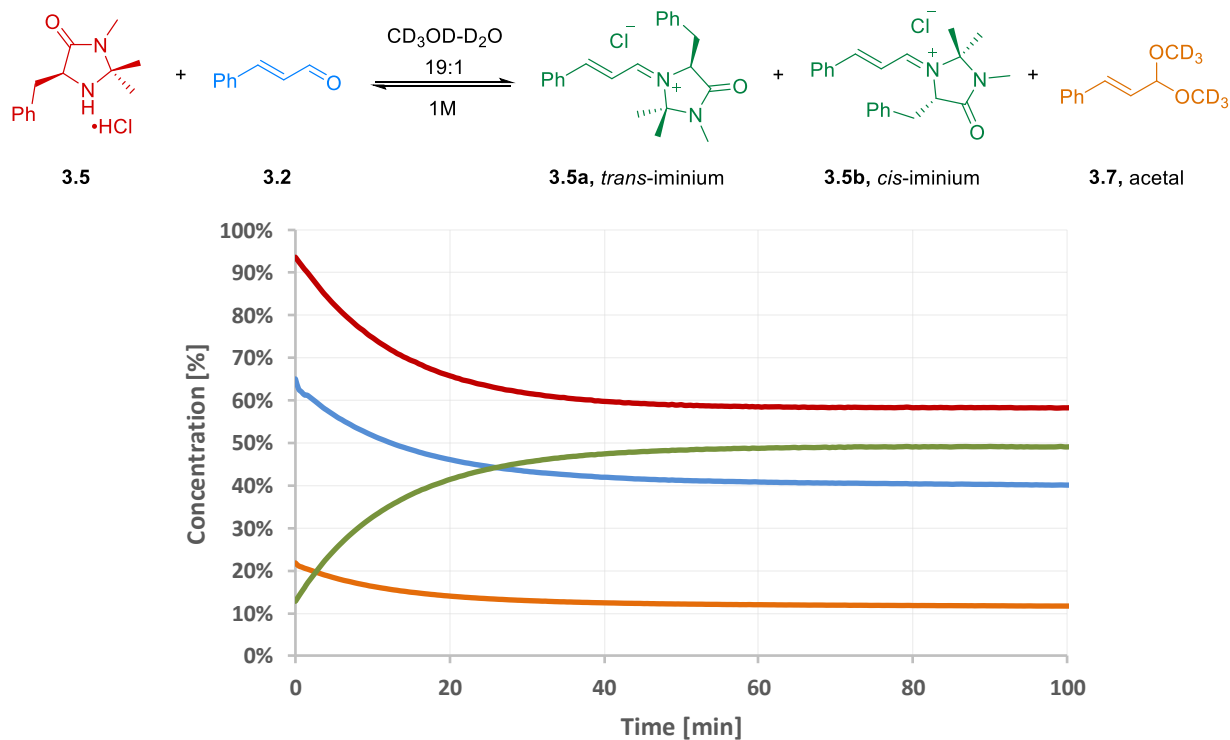


Time-dependent experiment for the formation of 3.5a and 3.5b in CD₃OD: D₂O (19:1) with MacMillan's catalyst 3.5.

MacMillan's catalyst (**3.5**, 165.5 mg, 0.65mmol) dissolved in 0.3 μL of CD₃OD: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (82 μL, 0.65 mmol) in 0.35 μL of CD₃OD: D₂O (19:1). The time-dependent experiment was conducted on a Varian VNMRs 500 MHz NMR spectrometer where ns = 8, SW = 16 ppm, RG = 16 and temp = 24°C. The initial experiment was reported as t = 0 min. The concentration vs time curve (**Figure S3.4**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.5a** & **3.5b**)] = 1
 Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.63-9.67, acetal (**3.7**) 6.11-6.20, *trans*- and *cis*-iminium (**3.5a** & **3.5b**) 0.89-0.93 & 0.95-0.97, amine (**3.5**) 1.61-1.67.

Figure S3.4. Concentration vs time curve for the formation of **3.5a** and **3.5b** with **3.5**.



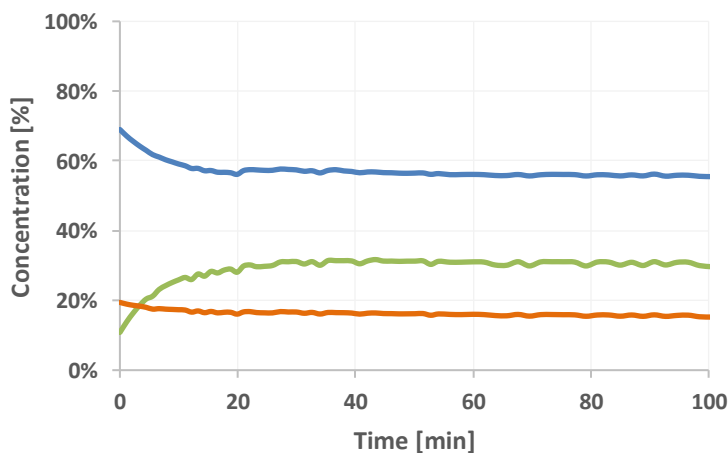
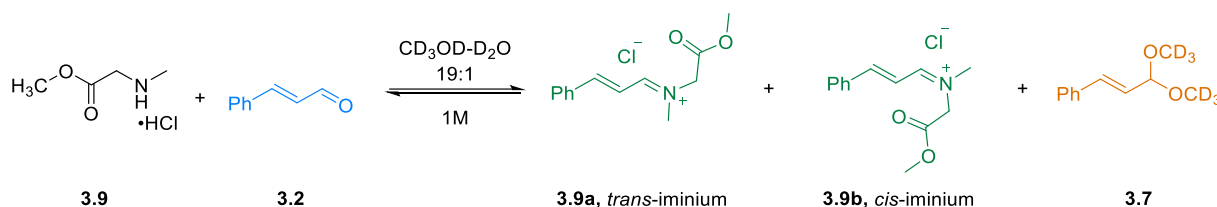
Time-dependent experiment for the formation of 3.9a and 3.9b in CD₃OD: D₂O (19:1) with 3.9.

Sarcosine methyl ester hydrochloric acid salt (**3.9**, 91.25 mg, 0.65mmol) dissolved in 30 uL of CD₃OD: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (82 uL, 0.65 mmol) in 35 uL of CD₃OD: D₂O (19:1). The time-dependent experiment was conducted on a Varian VNMRs 500 MHz NMR spectrometer from t = 0 minutes to t = 228 minutes, where ns = 8, SW = 14 ppm, RG = 20 and temp = 25°C. The concentration vs time curve (**Figure S3.5**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.9a** & **3.9b**)] = 1

Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.60-9.68; acetal (**3.7**) 4.90-4.92; *cis* and *trans*-iminium (**3.9a** & **3.9b**) 8.83-9.05.

Figure S3.5. Concentration vs time curve for the formation of **3.9a** and **3.9b** with **3.9**.

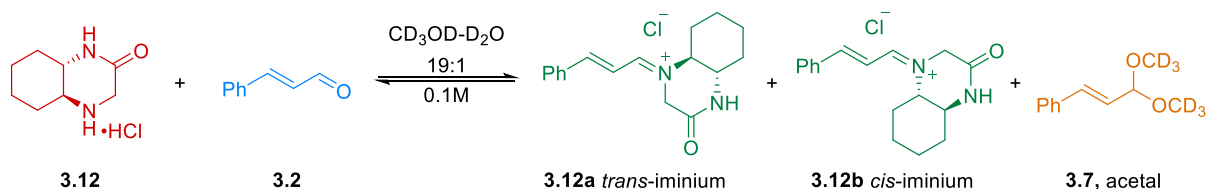


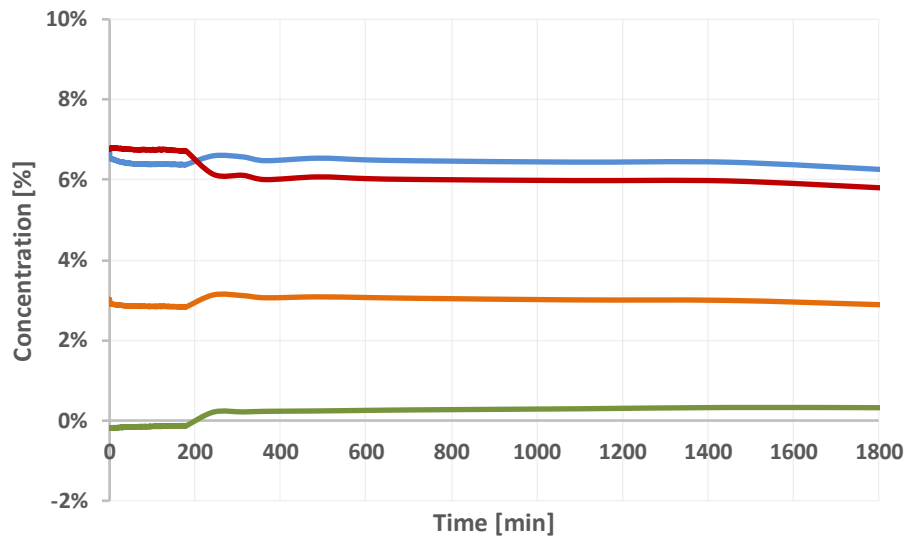
Time-dependent experiment for the formation of 3.12a and 3.12b in CD₃OD: D₂O (19:1) with 3.12.

Amine **3.12** (12.4 mg, 0.065mmol) dissolved in 30 uL of CD₃OD: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (8.2 uL, 0.065 mmol) in 35 uL of CD₃OD: D₂O (19:1). The time-dependent experiment was conducted on a Varian VNMRS 500 MHz NMR spectrometer from t = 0 minutes to t = 179.6 minutes, where ns = 8, SW = 16 ppm, RG = 34 and temp = 25°C. The Bruker 500 MHz NMR was used for t = 243.6 to t = 1803.5 minutes. where ns = 8, SW = 20 ppm, RG = 106.6 and temp = 25°C. The concentration vs time curve (**Figure S3.6**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.12a** & **3.12b**)] = 1
 Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.74-9.60, acetal (**3.7**) 6.10-6.21, trans- & cis-iminium (**3.12a**, **3.12b**)8.85-9.05-8.84, amine (**3.12**) concentration consists of the averaged integration of the following peaks: 1.63-1.34ppm (4H), 2.21-1.75ppm (4H), 3.52-3.43ppm (1H), 3.22-3.13ppm (1H). All peaks were divided by the number of protons prior to averaging.

Figure S3.6. Concentration vs time curve for the formation of **3.12a** and **3.12b** with **3.2**.





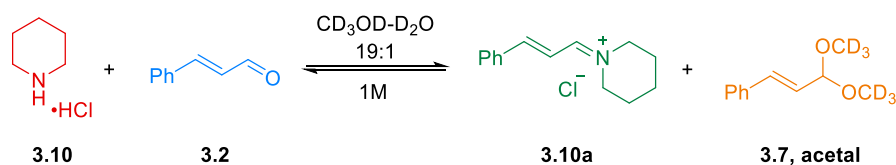
Time-dependent experiment for the formation of 3.10a in CD₃OD: D₂O (19:1) with piperidine hydrochloride (3.10).

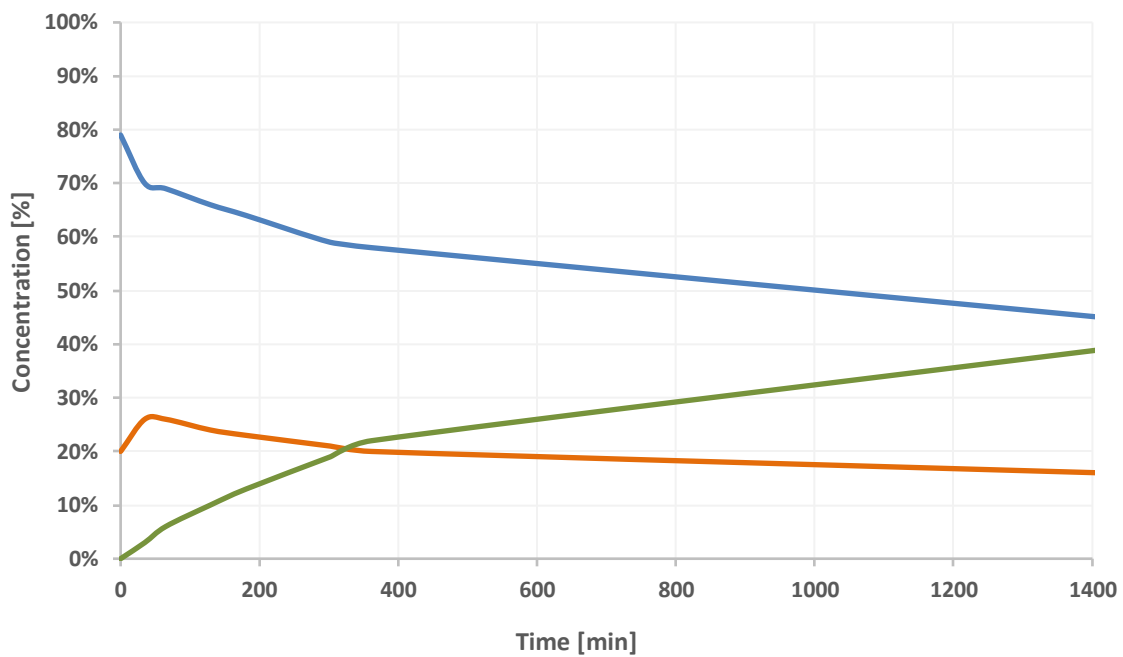
Piperidine hydrochloride (**3.10**, 79.1 mg, 0.65mmol) dissolved in 0.3 uL of CD₃OD: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (82 uL, 0.65 mmol) in 0.35 uL of CD₃OD: D₂O (19:1). The time-dependent experiment was conducted on a QANUC 500 MHz NMR spectrometer from t = 0 minutes to t = 1413.4 minutes (23.5 hours), where ns = 4, SW = 14 ppm, RG = 22 and temp = 23°C. The concentration vs time curve (**Figure S3.7**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.10a**)] = 1

Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.64-9.66, acetal (**3.7**) 4.89-4.91, iminium (**3.10a**) 8.70-8.75.

Figure S3.7. Concentration vs time curve for the formation of **3.10a** with **3.2**.





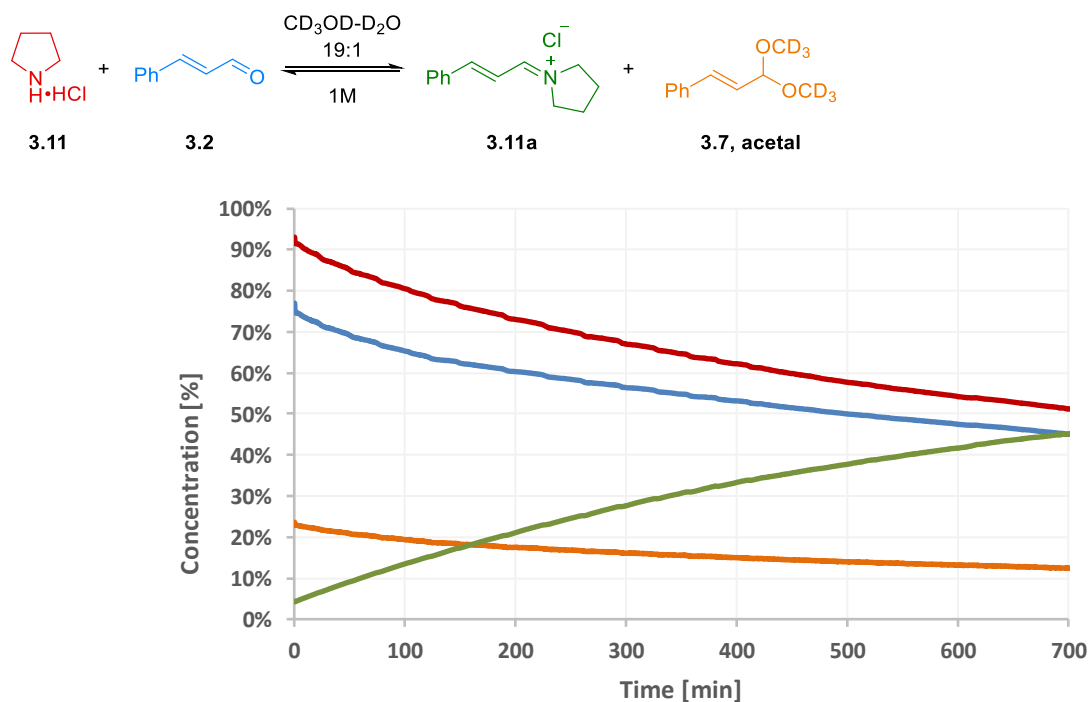
Time-dependent experiment for the formation of 3.11a in CD₃OD: D₂O (19:1) with pyrrolidine hydrochloride (3.11)

Pyrrolidine hydrochloride (**3.11**, 69.3 mg, 0.65mmol) dissolved in 0.3 uL of CD₃OD: D₂O (19:1) was added to a solution of cinnamaldehyde (**3.2**) (82 uL, 0.65 mmol) in 0.35 uL of CD₃OD: D₂O (19:1). The time-dependent experiment was conducted on a VARIAN500 MHz NMR spectrometer from t = 0 minutes to t = 720 minutes (12 hours), where ns = 8, SW = 16 ppm, RG = 34 and temp = 25°C. The concentration vs time curve (**Figure S3.8**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.11a**)] = 1

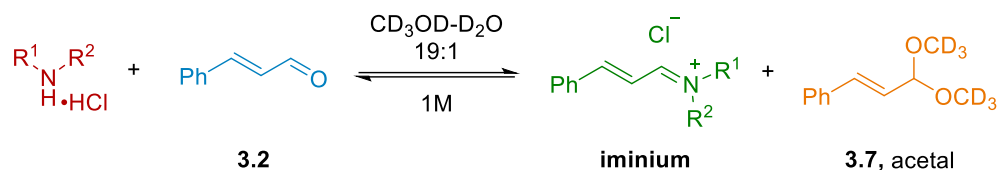
Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.62-9.67, acetal (**3.7**) 4.88-4.94, iminium (**3.11a**) 8.76-8.84, amine (**3.11**) 3.20-3.30.

Figure S3.8. Concentration vs time curve for the formation of **3.11a** with **3.2**.



Amine catalysed iminium formation at $t = 30\text{min}$

Scheme S3.1. General reaction scheme for iminium formation.



General Information (Table S3.1)

All amines were purchased from Sigma-Aldrich, and Combi Blocks unless specified. Amine hydrochlorides (3.12, 3.23) were synthesized as described in literature¹. Amine salt (3.10, 3.11, 3.18, 3.20, 3.21) were synthesized according to reported protocol.² ¹H NMR spectra were obtained on a Varian VNMRS 500 MHz NMR spectrometer or Varian Inova 500 MHz NMR Spectrometer (“QANUC 500”) at McGill University. Chemical shifts are reported in δ (ppm) referenced to the residual solvent signal. ¹H NMR spectra are tabulated in the following order: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, p = pentet, and br s = broad singlet), coupling constant(s) (J) in hertz (Hz), and number of protons. ¹H NMR time-dependent experiments were recorded on Varian VNMRS 500 MHz NMR spectrometer or Varian Inova 500 MHz NMR Spectrometer (“QANUC 500”). Chemical shifts are calibrated in δ (ppm) referenced to the residual solvent signal.

General synthesis of amine salts (3.10, 3.11, 3.18, 3.20, 3.21, 3.12)

In a flame dried flask, amine was dissolved in anhydrous THF (0.5 M) or DCM (0.5 M) and the reaction mixture was cooled to 0°C. At 0°C, HCl (2 M solution in Et₂O, 1.5 equiv.) was added dropwise. The reaction was allowed to warm to rt and stirred for 24 hours. Solvent and volatiles were removed under reduced pressure to give a white solid.

Supplementary information Chapter 3

Piperidine hydrochloride, **3.10**: 0.2 mL piperidine in 5 mL DCM, 1.5 mL HCl (2 M solution in Et₂O). 243 mg, 99% yield. ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 9.03 (s, 2H), 3.00 – 2.92 (m, 4H), 1.67 (p, *J* = 5.8 Hz, 4H), 1.58 – 1.50 (m, 2H).

Pyrrolidine hydrochloride, **3.11**: 0.3 mL pyrrolidine in 7 mL anhydrous THF, 2.7 mL HCl (2 M solution in Et₂O). 296 mg, 75% yield. ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 9.37 (s, 2H), 3.11 – 3.00 (m, 4H), 1.87 – 1.76 (m, 4H).

2-Methylpiperidine hydrochloride, **3.18**: 0.5 mL 2-methyl piperidine in 9 mL anhydrous THF, 3.2 mL HCl (2 M solution in Et₂O). 570 mg, 99% yield. ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 9.31 – 8.63 (m, 2H), 3.16 (d, *J* = 12.7 Hz, 1H), 3.03 (s, 1H), 2.77 (q, *J* = 11.9 Hz, 1H), 1.79 – 1.66 (m, 3H), 1.66 – 1.53 (m, 1H), 1.48 – 1.35 (m, 2H), 1.22 (d, *J* = 6.5 Hz, 3H).

Azepane hydrochloride, **3.20**: 0.5 mL azepane in 9 mL anhydrous THF, 3.3 mL HCl (2 M solution in Et₂O). 590 mg, 98% yield. ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 9.18 (s, 2H), 3.03 (d, *J* = 7.0 Hz, 4H), 1.76 (dtdd, *J* = 7.6, 5.8, 4.2, 2.3 Hz, 4H), 1.58 (dt, *J* = 7.4, 2.7 Hz, 4H).

Diethylamine hydrochloride, **3.21**: 0.5 mL diethyl ether in 9.5 mL anhydrous THF, 3.6 mL HCl (2 M solution in Et₂O). 497 mg, 94% yield. ¹H NMR (500 MHz, DMSO-*d*₆) δ (ppm) 9.05 (s, 2H), 2.86 (q, *J* = 7.3 Hz, 4H), 1.18 (t, *J* = 7.3 Hz, 6H).

General synthesis of amine hydrochlorides (3.12, 3.23)

(4a*S*,8a*S*)-octahydroquinoxalin-2(1H)-one hydrochloride, **3.12**: (1*S*,2*S*)-cyclohexane-1,2-diamine (500 mg, 4.379 mmol) was dissolved in water (13.27 mL, 0.33 M). The solution was cooled to 10°C and a solution of chloroacetic acid (413.74 mg, 4.379 mmol) in water (4.38 mL, 1 M) was added dropwise at 10°C over 7 minutes. A solution of potassium carbonate (968.21 mg, 7.006 mmol) in water (2.62 mL, 2.67 M) was then also added dropwise 10°C. Reaction mixture was allowed to warm up to room temperature and stirred at room temperature for 24 hours. After

24 hours, the reaction mixture was heated to 90°C (a reflux condenser was added) for 2 hours. After 2 hours the reaction was allowed back to room temperature and was concentrated under reduced pressure. The resulting solid was then dissolved in boiling ethanol, and the solution was filtered hot. The filtrate was concentrated to give a white powder that was then recrystallized from toluene to give white needle like crystals. 352 mg, 52% yield. ¹H NMR (500 MHz, Acetone) δ 6.75 (s, 1H), 3.34 (d, *J* = 3.0 Hz, 2H), 2.96 (ddd, *J* = 11.1, 9.1, 3.8 Hz, 1H), 2.34 (ddd, *J* = 11.4, 9.1, 3.7 Hz, 1H), 1.96 (s, 1H), 1.91 – 1.66 (m, 4H), 1.44 – 1.15 (m, 4H). ¹³C NMR (126 MHz, Acetone) δ 169.84, 59.39, 59.14, 51.06, 32.38, 31.69, 25.74, 24.73. HRMS (ESI+) for C₈H₁₅N₂O [M + H]⁺, calcd: 155.11789, found: 155.11792.

The hydrochloride salt of the amine was made according to the general procedure, (200 mg, 1.297 mmol) (4aS,8aS)-octahydroquinoxalin-2(1H)-one in 3 mL anhydrous THF, 0.5 mL 1,4-dioxane, 1 mL HCl (2M solution in Et₂O) quantitative yield. ¹H NMR (500 MHz, DMSO) δ 10.34 – 9.96 (m, 2H), 8.40 (s, 1H), 3.74 – 3.61 (m, 1H), 3.57 (dd, *J* = 16.7, 4.4 Hz, 1H), 3.52 – 3.37 (m, 1H), 3.00 (qd, *J* = 10.6, 3.5 Hz, 1H), 2.06 (dt, *J* = 12.2, 3.6 Hz, 1H), 1.97 – 1.88 (m, 1H), 1.75 (d, *J* = 12.2 Hz, 1H), 1.69 – 1.63 (m, 1H), 1.43 (qd, *J* = 12.3, 3.7 Hz, 1H), 1.33 – 1.11 (m, 3H). ¹³C NMR (126 MHz, DMSO) δ 162.97, 55.45, 51.91, 44.08, 30.42, 26.83, 23.55, 22.72.

(S)-(4-methylpiperidin-1-yl)(piperidin-2-yl)methanone hydrochloride, **3.23**: In a flame dried flask under argon, Boc-L-pipecolic acid (300 mg, 1.308 mmol) was dissolved in DCM (10 mL, 0.13 M). Solution was cooled to 0°C, then DIPEA (0.25 mL, 1.427 mmol), PyBOP (685 mg, 1.316 mmol), and 4-methyl piperidine (0.14 mL, 1.189 mmol) were added. Reaction was allowed slowly back to room temperature and stirred at 0°C → room temperature for 24 hours. DCM and saturated aqueous NH₄Cl were added. Phases were separated and organic phase was washed with saturated aqueous NH₄Cl (x2), then with saturated aqueous NaHCO₃ (x1), and then with brine (x1). Organic

phase was then dried (Na_2SO_4), filtered, and concentrated, to give a clear yellowish oil, that solidified upon standing. Crude was purified by flash column chromatography (5 \rightarrow 40% ethyl acetate in hexanes) to give a white solid. 274 mg, 74 % yield. ^1H NMR (500 MHz, CDCl_3) δ 5.10 – 4.77 (m, 1H), 4.61 – 4.41 (m, 1H), 4.03 – 3.74 (m, 2H), 3.45 – 3.15 (m, 1H), 3.07 – 2.83 (m, 1H), 2.63 – 2.45 (m, 1H), 1.93 – 1.30 (m, 17H), 1.18 – 1.00 (m, 2H), 0.94 (dd, $J = 6.6, 2.8$ Hz, 3H).

In a flame dried flask, tert-butyl (S)-2-(4-methylpiperidine-1-carbonyl)piperidine-1-carboxylate (189 mg, 0.609 mmol) was dissolved in anhydrous THF (2 mL, 0.3M). Solution was cooled to 0°C and HCl (2M solution in Et_2O , 2.3 mL, 4.51 mmol) was added. Reaction stirred at $0^\circ\text{C} \rightarrow$ room temperature for 24 hours. Solvent and volatiles were removed to give an off-white solid that was then recrystallized from acetonitrile to give a white solid. 70 mg, 47% yield. ^1H NMR (500 MHz, MeOD) δ 4.45 (tdt, $J = 12.3, 4.7, 2.4$ Hz, 1H), 4.31 (ddd, $J = 17.7, 12.3, 3.0$ Hz, 1H), 3.82 (ddq, $J = 11.2, 4.3, 2.1$ Hz, 1H), 3.42 – 3.35 (m, 1H), 3.19 – 2.99 (m, 2H), 2.71 (qd, $J = 12.9, 2.8$ Hz, 1H), 2.17 – 2.06 (m, 1H), 1.96 – 1.51 (m, 8H), 1.26 – 1.01 (m, 2H), 0.98 (dd, $J = 6.4, 4.1$ Hz, 3H). ^{13}C NMR (126 MHz, MeOD) δ 168.04, 57.43, 57.36, 46.89, 46.51, 44.89, 44.83, 43.86, 43.60, 35.82, 35.62, 34.89, 34.70, 32.08, 31.96, 28.19, 27.63, 22.86, 22.83, 22.81, 22.78, 22.03, 21.83. HRMS (ESI+) for $\text{C}_{12}\text{H}_{23}\text{N}_2\text{O}$ $[\text{M}+\text{H}]^+$ calcd: 211.1805, found: 211.1805.

Procedure for ^1H NMR time-dependent experiment

Cinnamaldehyde (**3.2**, 106 μL , 0.842 mmol) was added to a solution of the amine (0.2 equiv. 0.1684 mmol) dissolved in 0.842 mL of $\text{CD}_3\text{OD}:\text{D}_2\text{O}$ (19:1) unless otherwise specified.

Amine 3.1. Cinnamaldehyde (**3.2**, 126 μL , 1 mmol) was added to a solution of the Amine **3.1**: (0.2 equiv. 0.2 mmol) dissolved in 0.968 mL of $\text{CD}_3\text{OD}:\text{D}_2\text{O}$ (19:1).

Supplementary information Chapter 3

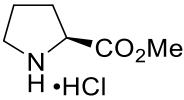
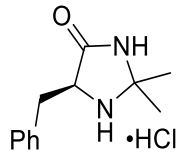
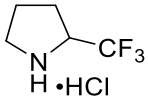
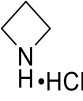
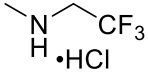
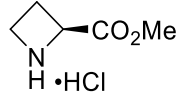
Amine 3.9. Cinnamaldehyde (**3.2**, 119 μL , 0.947 mmol) was added to a solution of the methyl methylglycinate hydrochloride (**3.9**) (0.2 equiv. 0.1894 mmol) dissolved in 0.88 mL of CD_3OD : D_2O (19:1).

After 30 minutes, the ^1H NMR of the reaction mixture was taken on a Varian VNMRS 500 MHz NMR spectrometer or Varian Inova 500 MHz NMR Spectrometer (“QANUC 500”) (Table S5). The % of iminium conversion was calculated from the concentration (integral) of iminium present in the reaction mixture during the time of the experiment ($t = 30$ min) (Table S3.2)

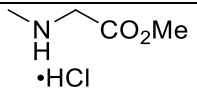
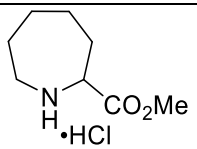
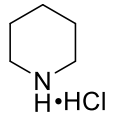
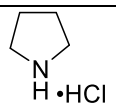
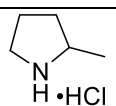
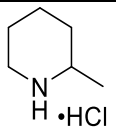
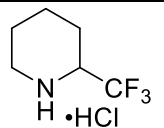
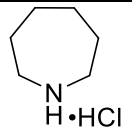
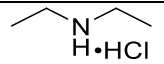
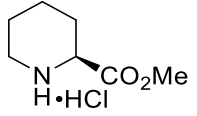
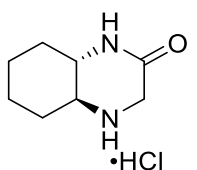
When $t = 30$ min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**Xa** & **Xb**)] = 1

Where X = relevant amine number.

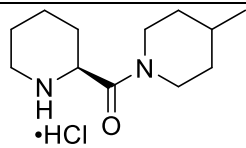
Table S3.1. List of amines and mass weighed for the experiments.

Amine	Structure	MW (g mol^{-1})	Mass (mg)
3.1		165.62	34.24
3.5		254.76	42.47
3.8		175.58	30.07
3.13		93.55	15.38
3.14		149.5	25.76
3.15		151.59	25.19

Supplementary information Chapter 3

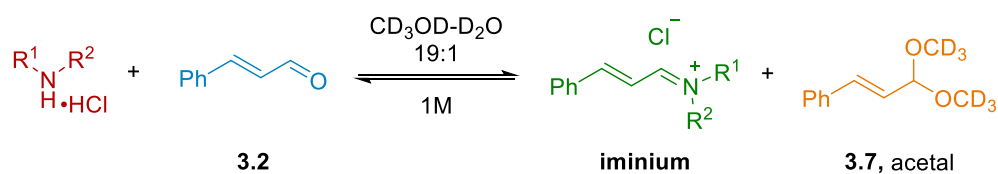
3.9		139.58	26.36
3.16		193.67	32.25
3.10		121.61	20.56
3.11		107.58	18.38
3.17		121.61	20.47
3.18		135.64	22.77
3.19		189.61	31.89
3.20		135.64	22.52
3.21		109.6	18.56
3.22		179.74	30.30
3.12		190.67	32.38

3.23



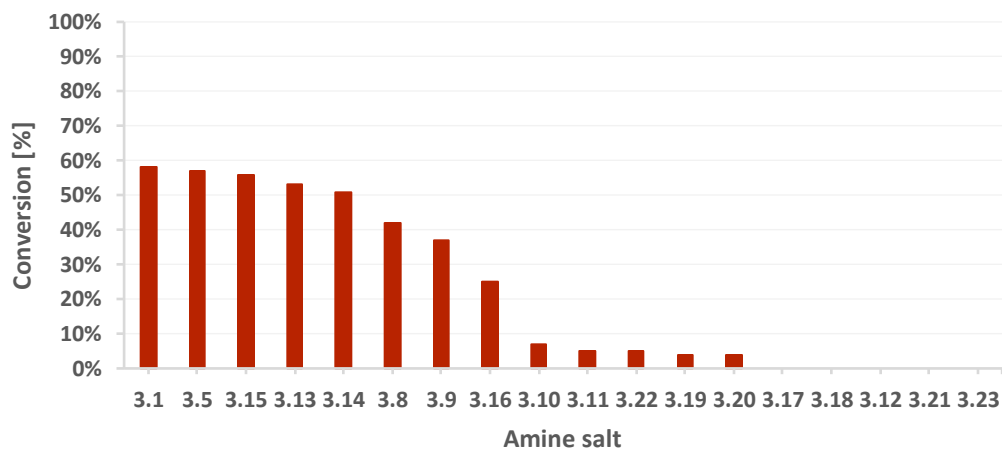
246.78

41.20

Table S3.2. Time-dependent ^1H NMR experiments on the formation of iminium after 30 min


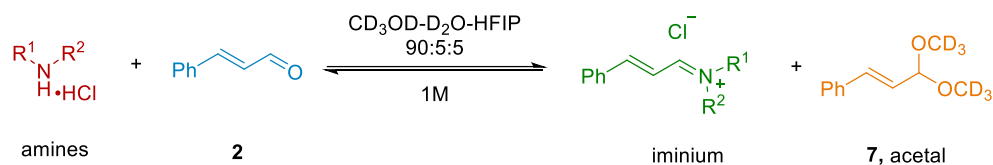
Amine	Concentration (M) at t = 30 min			
	cinnamaldehyde (3.2)	acetal (3.7)	Iminium	Amine
3.1	0.648	0.214	0.138	0.1
3.5	0.9662	0.232	0.106	0.08
3.8	0.653	0.189	0.157	0.22
3.13	0.680	0.218	0.102	0.09
3.14	0.689	0.214	0.096	0.0933
3.15	0.704	0.218	0.077	0.06
3.9	0.707	0.219	0.074	0.13
3.16	0.719	0.230	0.050	0.150
3.10	0.741	0.244	0.015	0.185
3.11	0.759	0.228	0.014	0.24
3.17	0.775	0.225	0	0.18
3.18	0.819	0.180	0	0.203
3.19	0.746	0.246	0.007	0.20
3.20	0.862	0.129	0.009	0.188
3.21	0.826	0.173	0	0.23
3.22	0.730	0.263	0.007	0.15
3.12	0.740	0.270	0.00	0.12
3.23	0.806	0.194	0	0.20

Figure S3.9. Amine(s) catalysed iminium formation at t = 30 min: percent of amine converted into iminium



Supplementary information Chapter 3

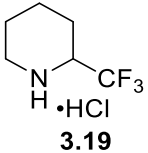
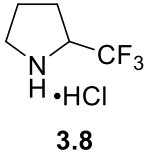
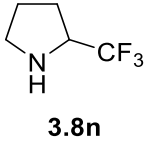
Scheme S3.2. Amine catalysed iminium formation at $t = 30$ min with HFIP as co-solvent



Procedure for 1H NMR time-dependent experiment with HIPF

Cinnamaldehyde (**3.2**, 106 μ L, 0.842 mmol) was added to a solution of the amine (0.2 equiv. 0.1684 mmol) dissolved in 0.842 mL of $CD_3OD-D_2O-HFIP$ (90:5:5).

Table S3.3. List of amines and results of time-dependent 1H NMR experiments on the formation of iminium after 30 min

Amine	Mw [gr/mol]	Amount [mg] or [μ L]	Concentration (M) at $t = 30$ min			
			Cinnamaldehyde (3.2)	Acetal (3.7)	Iminium	Amine
 3.19	189.61	29.5	0.76	0.24	0	0.21
 3.8	175.58	29.35	0.67	0.19	0.14	0.07
 3.8n	139.12	20	1	0	0	-

Time dependent Diels-Alder cycloaddition reaction

Experimental conditions for each of the amines used:

Methyl L-prolinate hydrochloride (**3.1**) (21.86 mg, 0.13mmol), cinnamaldehyde (**3.2**) (82 μ L, 0.65 mmol) and cyclopentadiene (**3.3**, 60 μ L, 0.715 mmol) were dissolved in 0.65 mL of CD₃OD: D₂O (19:1) The time-dependent experiment was conducted on a Varian Inova 500 MHz NMR spectrometer where ns = 4, SW = 16 ppm, RG = 18 and temp = 24°C. The initial experiment was reported as t = 0 min. The concentration vs time curve (Figure S5) was calibrated from the initial concentration (integral) of cinnamaldehyde (**2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.1a** & **3.1b**)] = 1

Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.57-9.67, *trans*- and *cis*-iminium (**3.1a** & **3.1b**) 8.79-9.04, *exo*- and *endo*-aldehyde (**3.4a** & **3.4b**) 9.75-9.82 & 9.41-9.45, *exo*- and *endo*-acetal (**3.4c** & **3.4d**) 4.31-4.38 & 3.85-3.94.

Pyrrolidine hydrochloride (**3.11**) (14.36 mg, 0.13mmol), cinnamaldehyde (**3.2**) (82 μ L, 0.65 mmol) and cyclopentadiene (**3.3**, 60 μ L, 0.715 mmol) was dissolved in 0.65 mL of CD₃OD: D₂O (19:1) The time-dependent experiment was conducted on a Varian Inova 500 MHz NMR Spectrometer where ns = 4, SW = 16 ppm, RG = 20 and temp = 24°C. The concentration vs time curve (**Figure S3.10**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.11a**)] = 1

Supplementary information Chapter 3

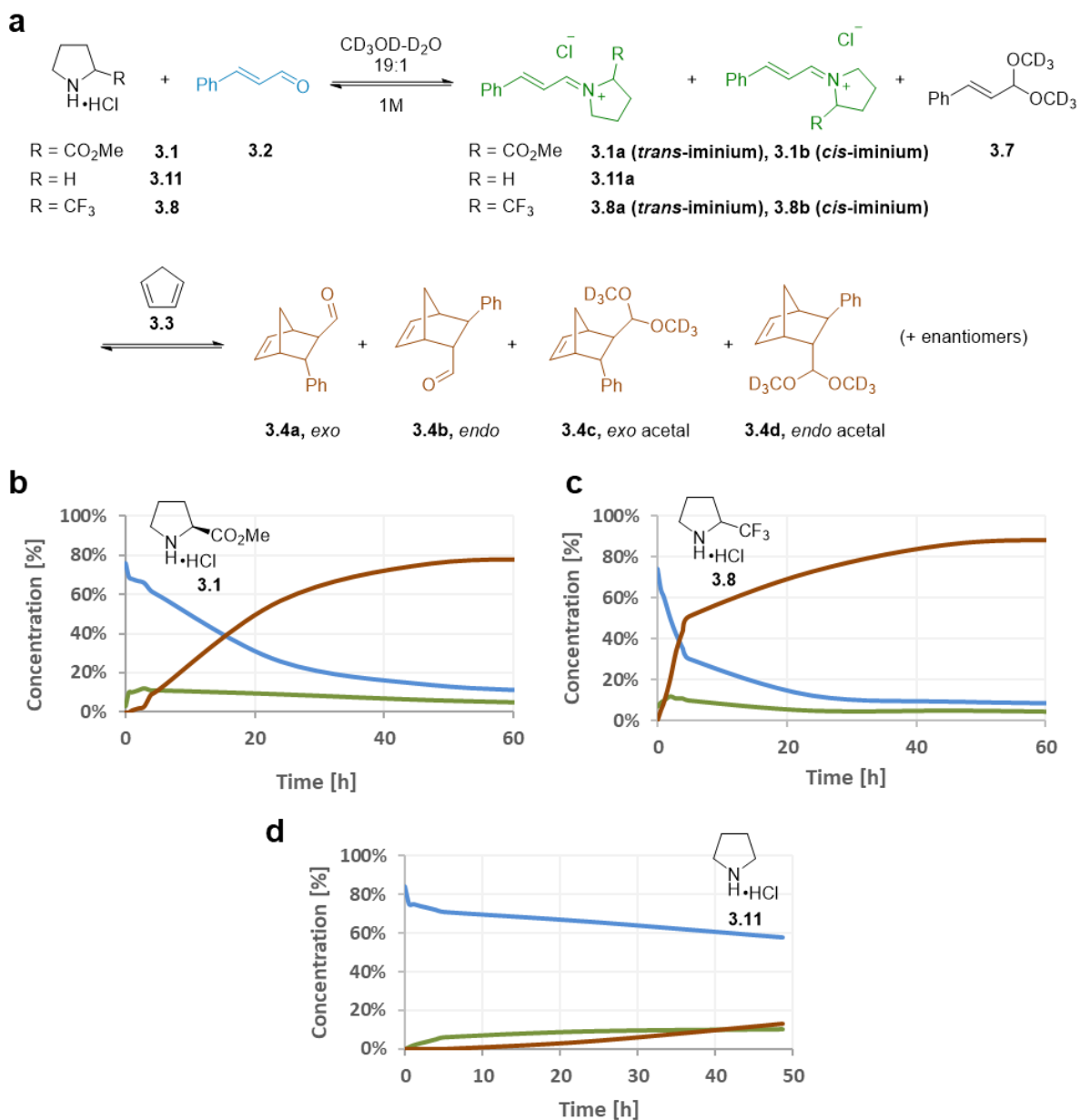
Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.57-9.66, iminium (**3.11**) 8.68-8.77, *exo*- and *endo*-aldehyde (**3.4a** & **3.4b**) 9.75-9.82 & 9.41-9.46, *exo*- and *endo*-acetal (**3.4c** & **3.4d**) 4.31-4.37 & 3.87-3.92.

2-(trifluoromethyl)pyrrolidine hydrochloride (**3.8**) (26.56 mg, 0.13mmol), cinnamaldehyde (**3.2**) (82 μ L, 0.65 mmol) and cyclopentadiene (**3.3**, 60 μ L, 0.715 mmol) was dissolved in 0.65 mL of CD₃OD: D₂O (19:1) The time-dependent experiment was conducted on a Varian Inova 500 MHz NMR Spectrometer where ns = 4, SW = 16 ppm, RG = 20 and temp = 24°C. The concentration vs time curve (**Figure S3.10**) was calibrated from the initial concentration (integral) of cinnamaldehyde (**3.2**) present in the reaction mixture during the time of the experiment (t = 0 min)

When t = 0 min, Σ integrals [cinnamaldehyde (**3.2**) + acetal (**3.7**) + iminium (**3.8a** & **3.8b**)] = 1

Chemical shifts considered for integration [ppm]: cinnamaldehyde (**3.2**) 9.59-9.64, *trans*- and *cis*-iminium (**3.8a** & **3.8b**) 9.00-9.07, *exo*- and *endo*-aldehyde (**3.4a** & **3.4b**) 9.77-9.82 & 9.49-9.54, *exo*- and *endo*-acetal (**3.4c** & **3.4d**) 4.33-4.39 & 3.88-3.94.

Figure S3.10. Concentration vs time curve for the formation of **3.4**



Time course simulation over time

Figure S3.11. Time course concentration profiles for each species measured using $^1\text{H-NMR}$. Experimental data is plotted as points and the predicted concentrations corresponding to Model S3 with best-fit parameters found in **Table S3.4** are plotted as solid lines.

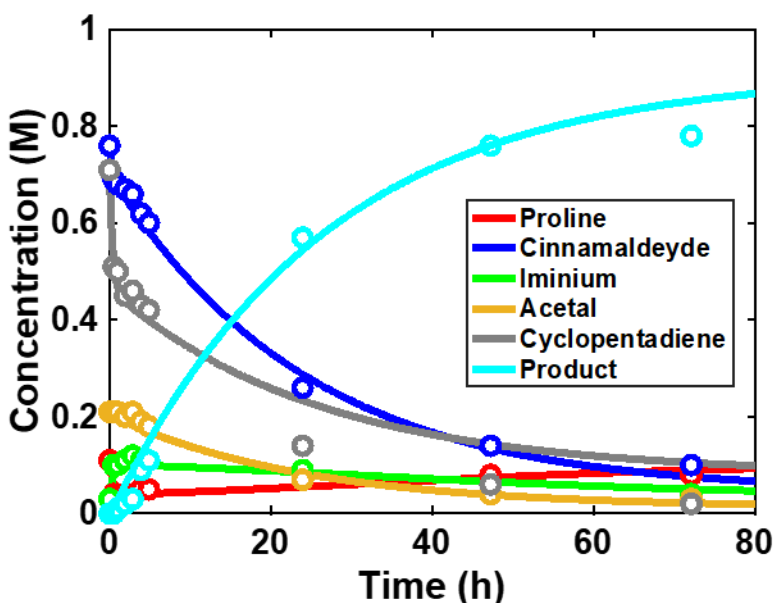


Figure S3.12. Simulated time course concentration profiles for product formation using Model S3 and the best-fit parameters found in **Table S3.4**. The profile from the original parameters is shown in cyan. The profile showing what happens when the rate of the first step of the reaction (k_1) is increased by a factor of two is shown in light blue. The profile showing what happens when the rate of the second step of the reactions (k_2) is increased by a factor of two is shown in dark blue. Increasing the rate of the second reaction (k_2) has a larger effect on the overall rate of the reaction, suggesting that the second step is the rate limiting step.

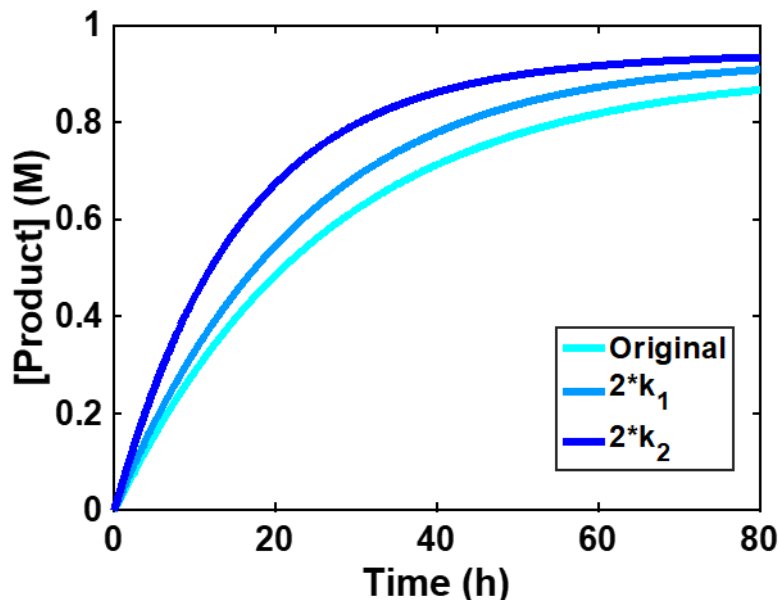


Figure S3.13. Model S3: Reaction scheme showing the four reactions that were considered during the fitting of experimental data. The first two steps show the successful conversion of Cinnamaldehyde and Cyclopentadiene into the product. The third step considers the dimerization of cyclopentadiene which has shown to occur in solution. The fourth step considers cinnamaldehyde reacting with the solvent (methanol) to produce an acetal. In the case of the fourth reaction, the concentration of methanol was assumed to be constant.

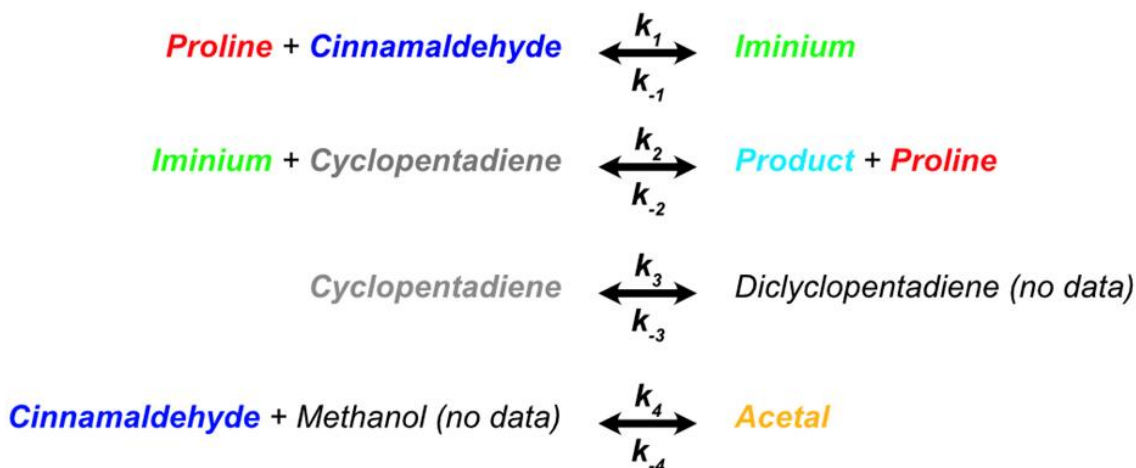


Table S3.4. Rate constants corresponding to best-fit parameters of Model S3 to the experimental data shown in **Figures S3.10b** and **S3.2**. k_1 , k_2 , k_{-2} , k_3 and k_4 are reported in $M^{-1} h^{-1}$.

k_{-1} , k_{-3} , and k_{-4} are reported in h^{-1} . Errors were calculated as reported in the data analysis section. As a note, the high errors in the parameters like k_{-2} are there because we can be reliably measured. Meaning there is a higher confidence in these parameters. Although it is important to keep in mind that the main purpose is to get qualitative understanding more than to get accurate numbers.

k_1	k_{-1}	k_2	k_{-2}
1.9 ± 0.2	0.17 ± 0.05	0.74 ± 0.06	0.03 ± 0.02
k_3	k_{-3}	k_4	k_{-4}
1.5 ± 0.4	1.2 ± 0.3	0.79 ± 0.04	2.7 ± 0.1

Data analysis

All data analysis was performed with MATLAB 2023a. The differential equations shown below (*Equations 1-6*) describe the complete reaction scheme *Model S3*, were integrated using MATLAB's built in ODE solver ode15s.

Equation (1)

$$\frac{d[\text{Proline}]_t}{dt} = -k_1 * [\text{Proline}]_t * [\text{Cinnamaldehyde}]_t + k_{-1} * [\text{Iminium}]_t - k_{-2} * [\text{Product}]_t * [\text{Proline}]_t + k_2 * [\text{Iminium}]_t * [\text{Cyclopentadiene}]_t$$

Equation (2)

$$\frac{d[\text{Cinnamaldehyde}]_t}{dt} = -k_1 * [\text{Proline}]_t * [\text{Cinnamaldehyde}]_t + k_{-1} * [\text{Iminium}]_t - k_4 * [\text{Cinnamaldehyde}]_t + k_{-4} * [\text{Acetal}]_t$$

Equation (3)

$$\frac{d[\text{Iminium}]_t}{dt} = k_1 * [\text{Proline}]_t * [\text{Cinnamaldehyde}]_t - k_{-1} * [\text{Iminium}]_t + k_{-2} * [\text{Product}]_t * [\text{Proline}]_t - k_2 * [\text{Iminium}]_t * [\text{Cyclopentadiene}]_t$$

Equation (4)

$$\frac{d[\text{Cyclopentadiene}]_t}{dt} = k_{-2} * [\text{Product}]_t * [\text{Proline}]_t - k_2 * [\text{Iminium}]_t * [\text{Cyclopentadiene}]_t - k_4 * [\text{Cyclopentadiene}]_t^2 + k_{-4} * [\text{Dicyclopentadiene}]_t$$

Equation (5)

$$\frac{d[\text{Product}]_t}{dt} = -k_{-2} * [\text{Product}]_t * [\text{Proline}]_t + k_2 * [\text{Iminium}]_t * [\text{Cyclopentadiene}]_t$$

Equation (6)

$$\frac{d[\text{Acetal}]_t}{dt} = k_4 * [\text{Cinnamaldehyde}]_t - k_{-4} * [\text{Aldehyde}]_t$$

Equation (7)

$$\frac{d[\text{Dicyclopentadiene}]_t}{dt} = k_4 * [\text{Cyclopentadiene}]_t^2 - k_{-4} * [\text{Dicyclopentadiene}]_t$$

where $k_1, k_2, k_3,$ and k_4 are the forward rate constants for each reaction, and $k_{-1}, k_{-2}, k_{-3},$ and k_{-4} are the reverse rate constants for each reaction.

$[\text{Proline}]_t, [\text{Cinnamaldehyde}]_t, [\text{Iminium}]_t, [\text{Cyclopentadiene}]_t, [\text{Product}]_t,$

$[\text{Aldehyde}]_t,$ and $[\text{Dicyclopentadiene}]_t$ are the total concentrations of each species at time t .

Kinetic parameters were globally fit to all identifiable species by minimizing the target function to the following species in solution:

[Proline]_t, [Cinnamaldehyde]_t, [Iminium]_t, [Cyclopentadiene]_t, [Product]_t, and [Aldehyde]_t,

(Equation 7)

$$RSS = \sum_{n=0}^N ([Experimental]_n - [Simulated]_n)^2$$

where the RSS is the residual sum of squared differences, [Experimental]_n is the experimental concentration for each species in solution, and [Simulate]_n is the predicted concentration of each species in solution, and *N* is the total number of time points. The minimization of this function was done in MATLAB 2023a using *fminsearch*. Errors for fitted parameters were calculated using a jackknife approach,²⁸³ in which each sample was obtained by resampling the original data with one less time point. For example, if the original dataset contained *N* points, each jackknife sample is constructed by removing a single point, and the data set would now contain *N*-1 points. There are a total of *N* data sets that can be made with *N*-1 points from the original set of data. The errors in the extracted parameters were taken as the standard deviations of these fitted sets of parameters obtained for all jackknife samples.

Computational study of iminium formation

All the computation is done at DFT level with the following keywords:

```
# avogadro generated ORCA input file
```

```
# Basic Mode
```

```
#
```

Supplementary information Chapter 3

*! wB97X-D3 Opt def2-TZVP RIJCOSX AutoAux SOSCF KDIIS defgrid2 TightSCF AnFreq
CPCM(Methanol)*

%geom

scan

B 38 44 = 1.82, 0.82, 21 #H2O....HNCR

end # B: bond; 38 45: atoms 39 and 46; 1.82 0.82: max and min distance; 21: iterations

constraints

{B 0 21 C}

end

*fullScan true # let the TS scan go to the end. If false it will stop when the max is reached to
search for TS*

MaxIter 500

end

%pal

nprocs 16

end

%scf

MaxIter 500

end

%maxcore 4900

wB97X-D3: Range-separated hybrid DFT functional with Grimme's D3 dispersion correction.²⁸⁴

Opt: Optimizing the structure as we go and don't compute only the energies

Def2-TZVP: Basis set used. Valence triple-zeta basis set (Ahlich family) with polarization

RIJCOSX: Resolution of Identity (RI) approximation to speed up calculation while introducing a very small error. Here RI-J is used for the Coulomb integrals and numerical chain-of-sphere integration for the HF Exchange integrals (COSX). This can speed up the calculation enormously. This is nowadays the default for hybrid-DFT in ORCA 5.0. The RI-J auxiliary basis set is picked automatically (AutoAux)

SOSCF + KDIIS: Second-Order Self-Consistent Field and Krylov-Subspace Direct Inversion of the Iterative Subspace algorithms. They are both used to accelerate the convergence of computations. The first one is using second-order information to improve the convergence of the SCF cycle. The second works by constructing a sequence of trial vectors from previous iterates, and then employs a direct inversion technique within a subspace (Krylov subspace in the case of

KDIIS) to generate a combination of these vectors. This combination minimizes the error in subsequent iterates, helping the algorithm converge faster to the self-consistent solution.

Defgrid2: Size of the DFT integration grid. It's the default one and has been chosen to be numerically robust and more accurate.

TightSCF: SCF Convergence tolerance. The maximum accepted is an Energy change $1.0e-08$ au. Default for geometry optimizations.

AnFreq: Performs anharmonic frequency calculations for a more detailed understanding of molecular vibrations and their interactions. These calculations are computationally more demanding than standard harmonic frequency calculations due to the inclusion of higher-order terms, but they provide a more comprehensive description of the molecular vibrational behavior.

CPCM(Methanol): Continuum solvation using Methanol, its dielectric constant and refractive index.

%geom Scan: A scan optimization is done, meaning the bond length of either bond is changed and the geometry is optimized at each point. The entire scan is done as fullScan is True

As an output of a scan job, the relaxscanscf energies are used. RELAXSCANS CF optimizes the geometry with respect to the scan coordinate considering the entire electronic structure and all molecular orbitals, unlike RELAXSCANACT.

Scan d₁: 21 iterations, the distance between the H and the O of the closest water molecule varies. From 1.82 Å to 0.82 Å with a 0.1 Å step.

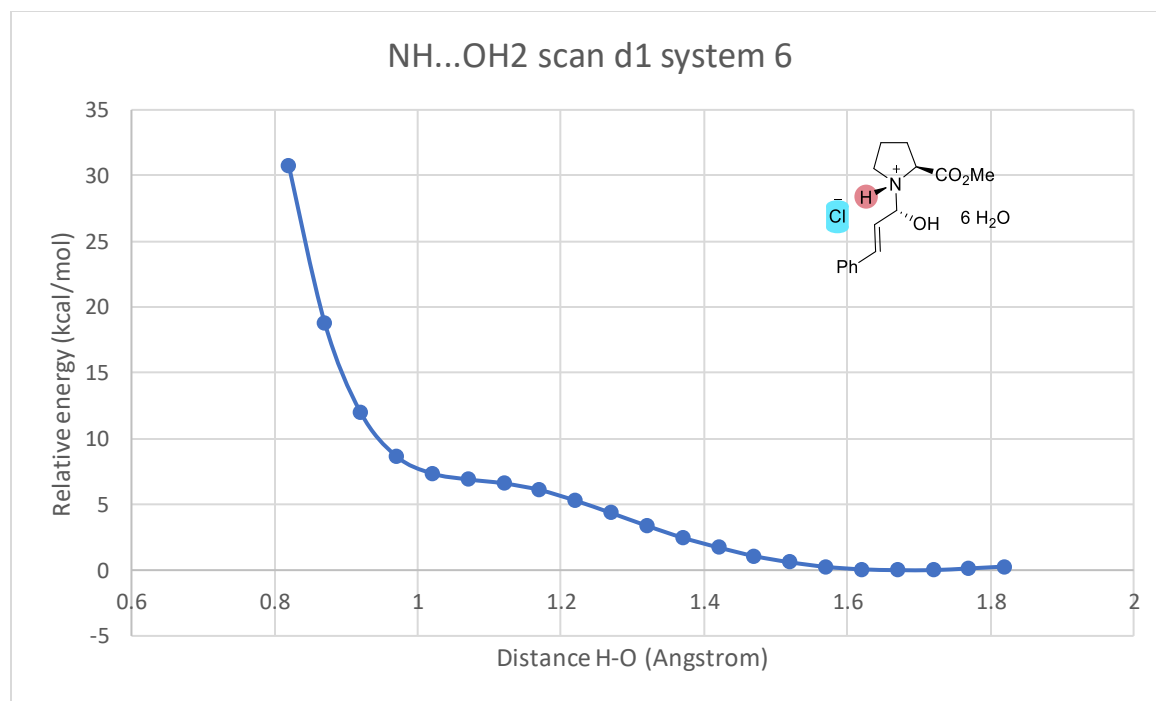
Supplementary information Chapter 3

Scan d_2 : 53 iterations, the distance between the C and the O of the protonated hydroxy varies. The distance of d_1 is kept constant (each d_1 distance has a full scan of d_2). From 1.40 Å to 4 Å, 0.1 Å step.

Results of first scan (d_1) $\text{NH}\cdots\text{OH}_2$

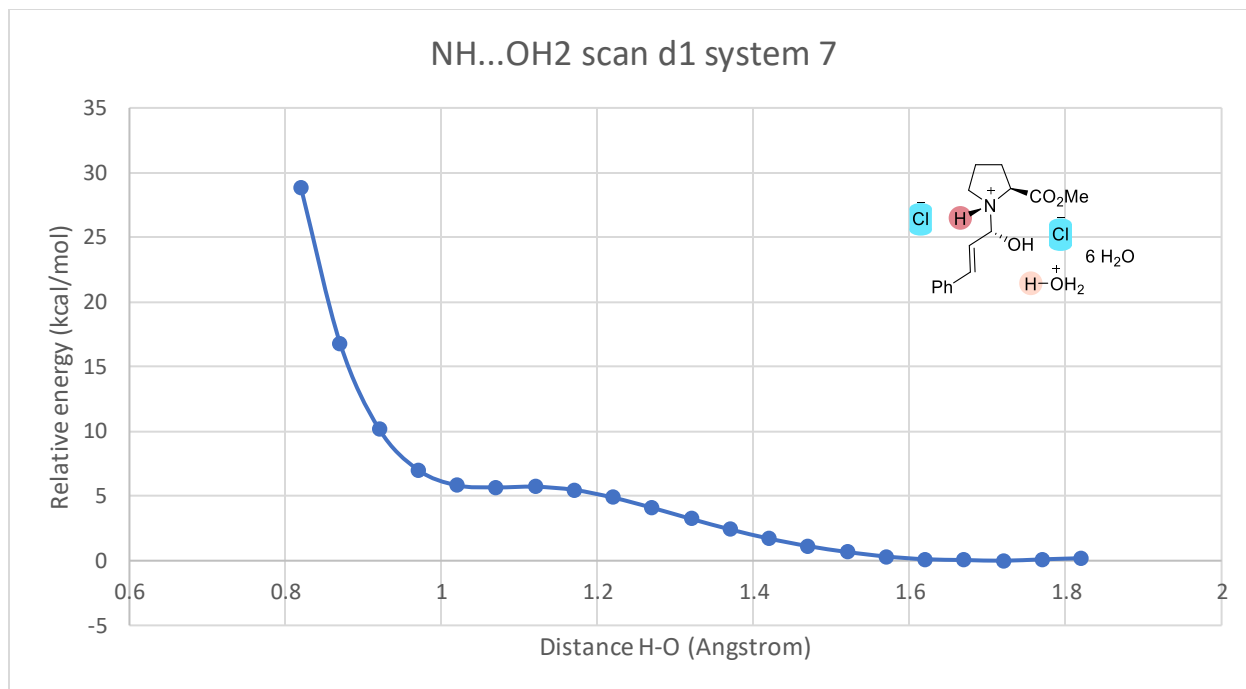
System 6: Cl^- by the N-H^+ (top face):

Figure S3.14. Plot of relative energy of the system when the bond distance d_1 decreases (and as a result the N-H bond length increases).



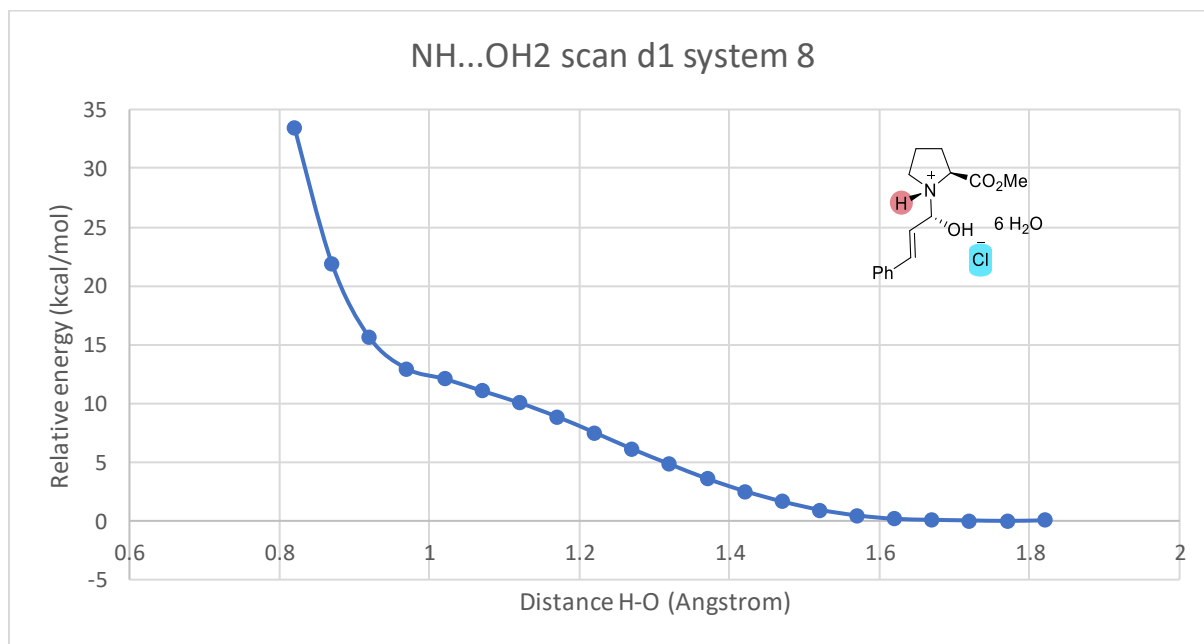
System 7: Cl^- by the N-H^+ (top face) + additional HCl by the OH (bottom face)

Figure S3.15. Plot of relative energy of the system when the bond distance d_1 decreases (and as a result the N-H bond length increases).



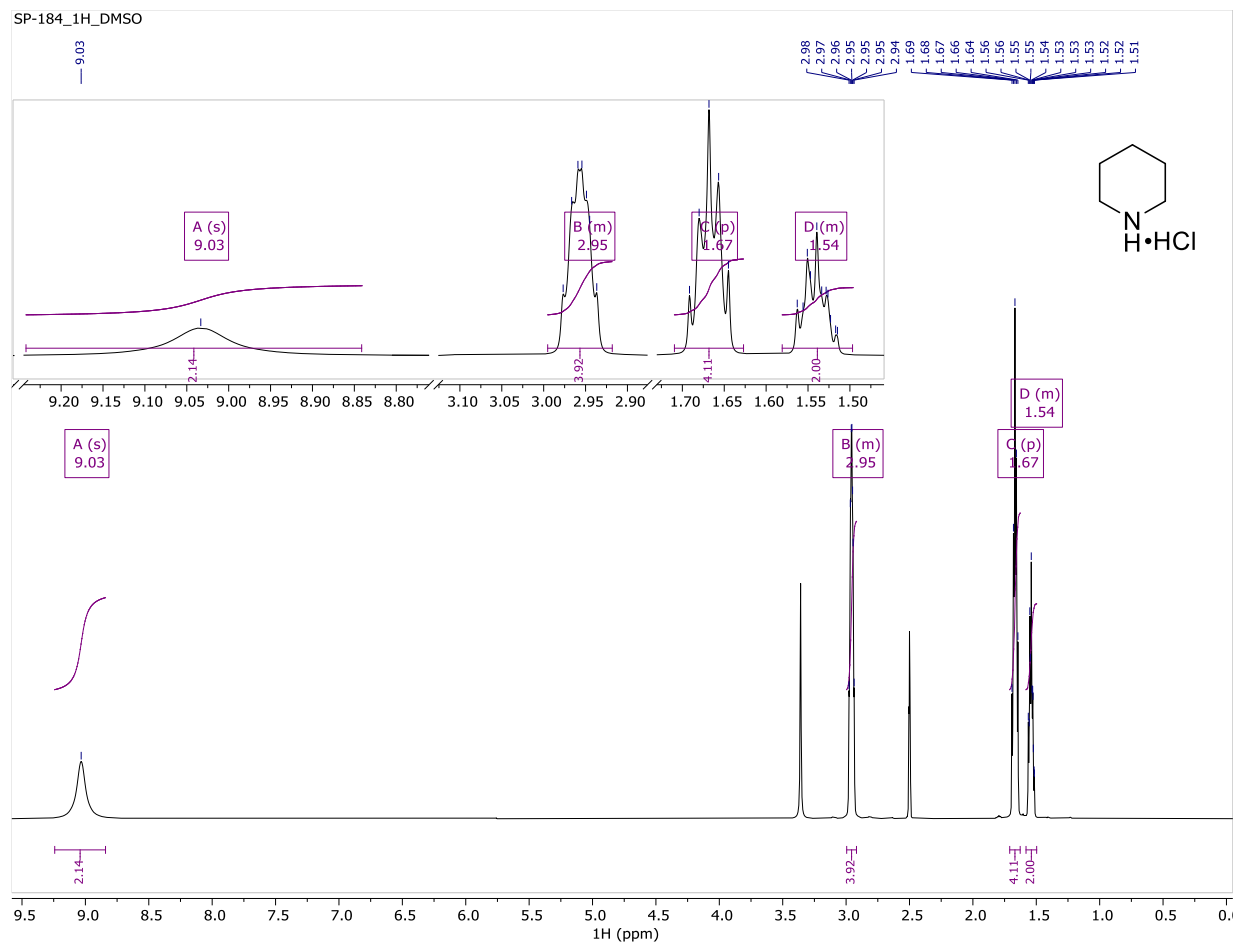
System 8: Cl- by the OH (bottom face)

Figure S3.16. Plot of relative energy of the system when the bond distance d_1 decreases (and as a result the N-H bond length increases).



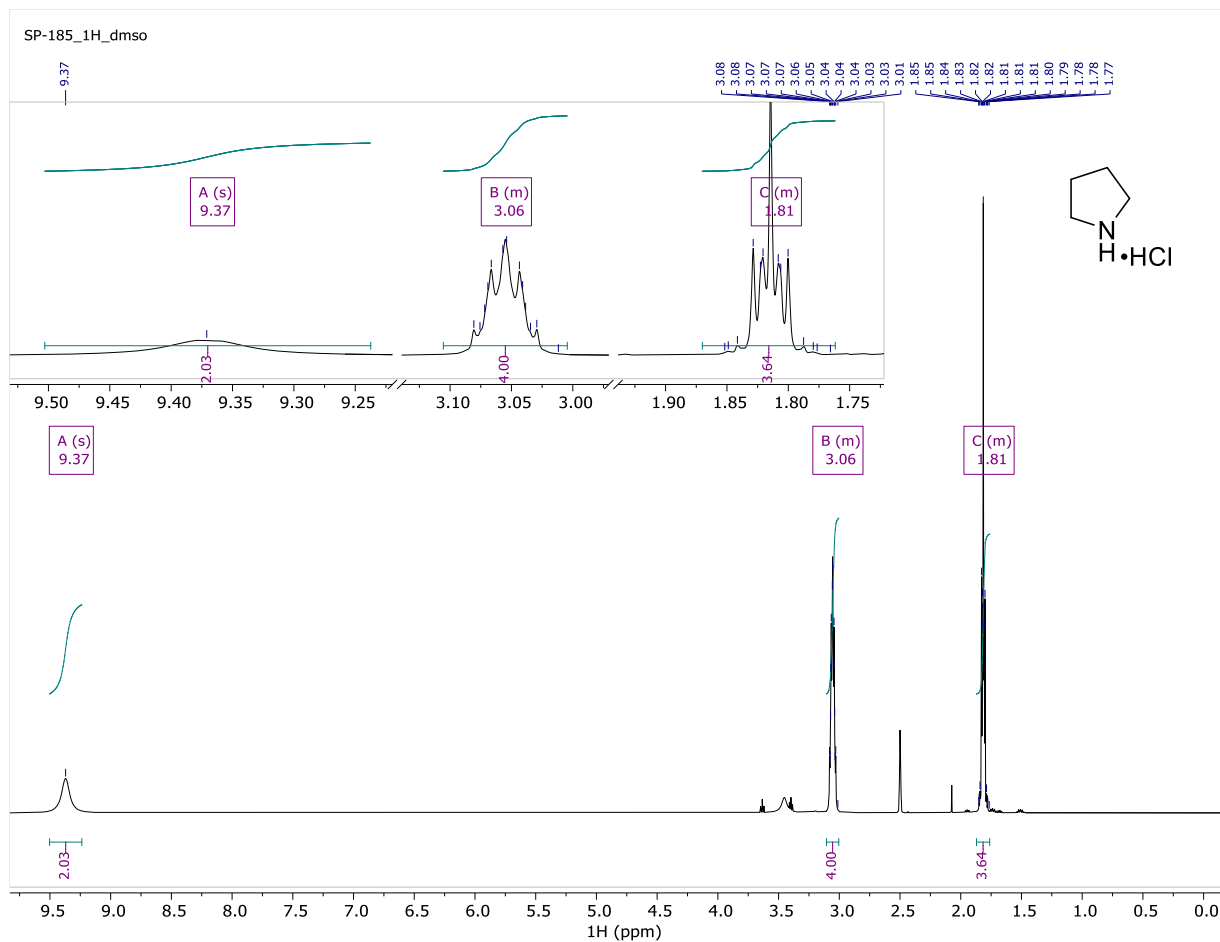
¹H NMR spectra of amine hydrochlorides

Amine **3.10** ¹H NMR (DMSO-*d*₆) Varian VNMRS 500 MHz



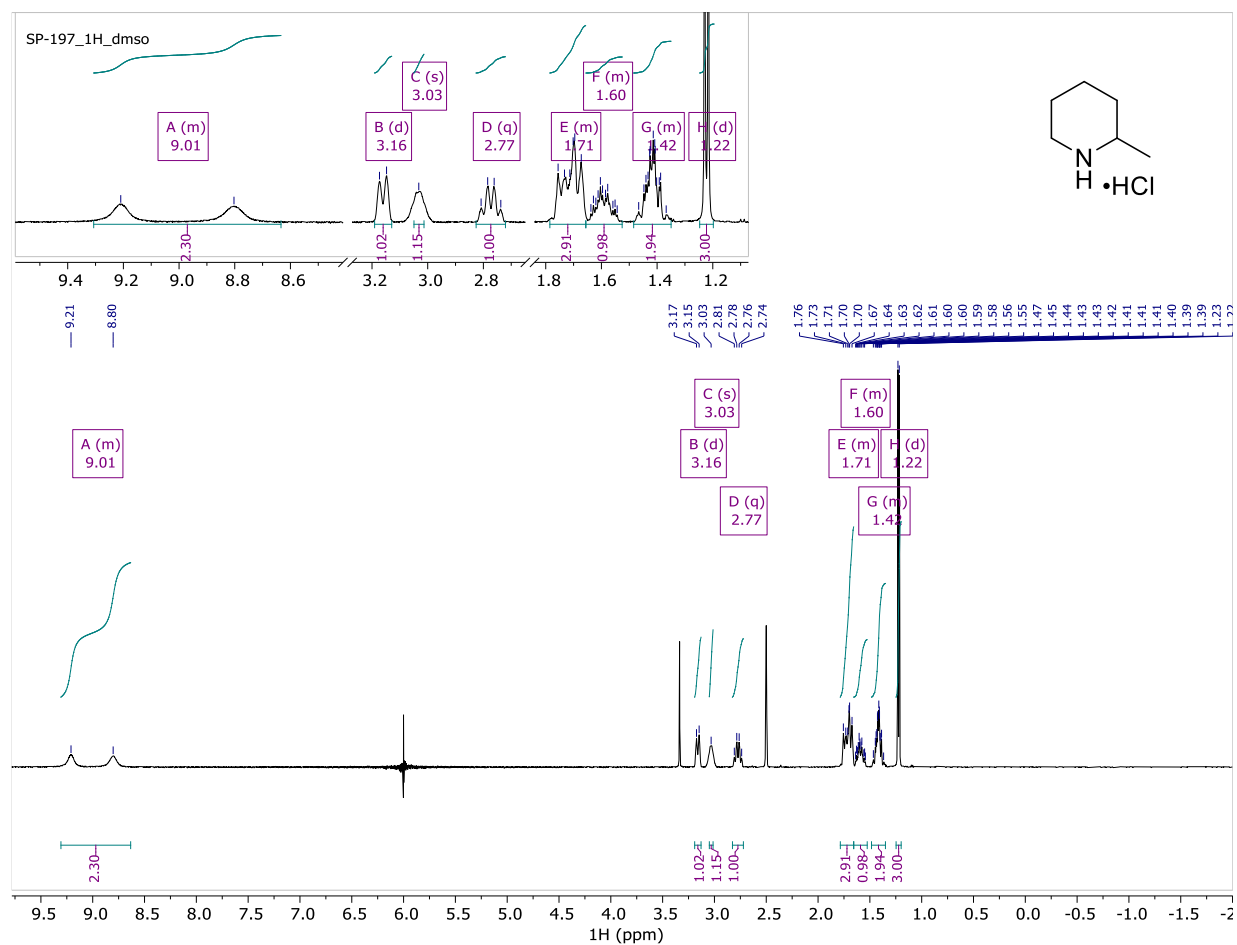
Supplementary information Chapter 3

Amine 3.11 ^1H NMR (500MHz) QANUC500 DMSO-d6



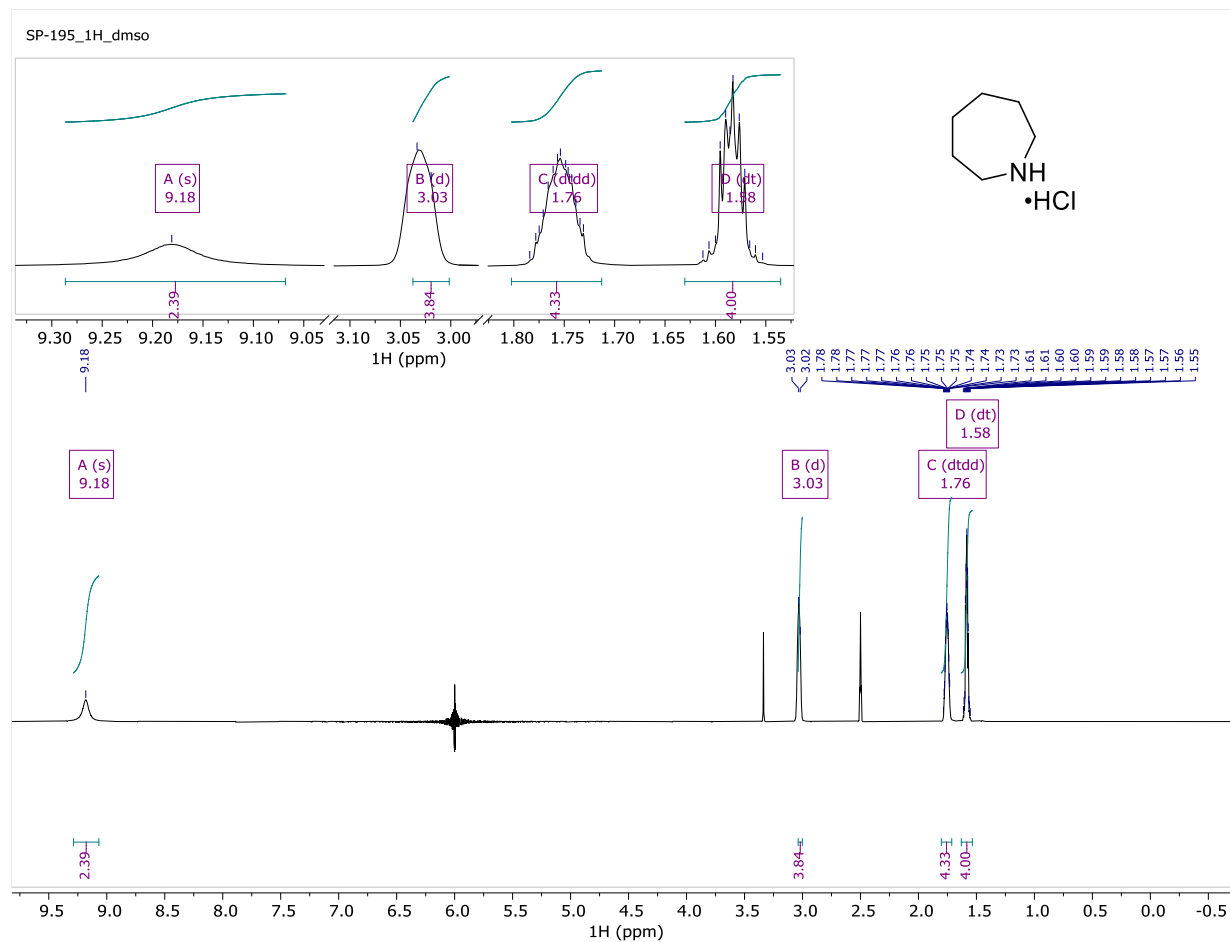
Supplementary information Chapter 3

Amine **3.18** ^1H NMR (500MHz) QANUC500 DMSO- d_6



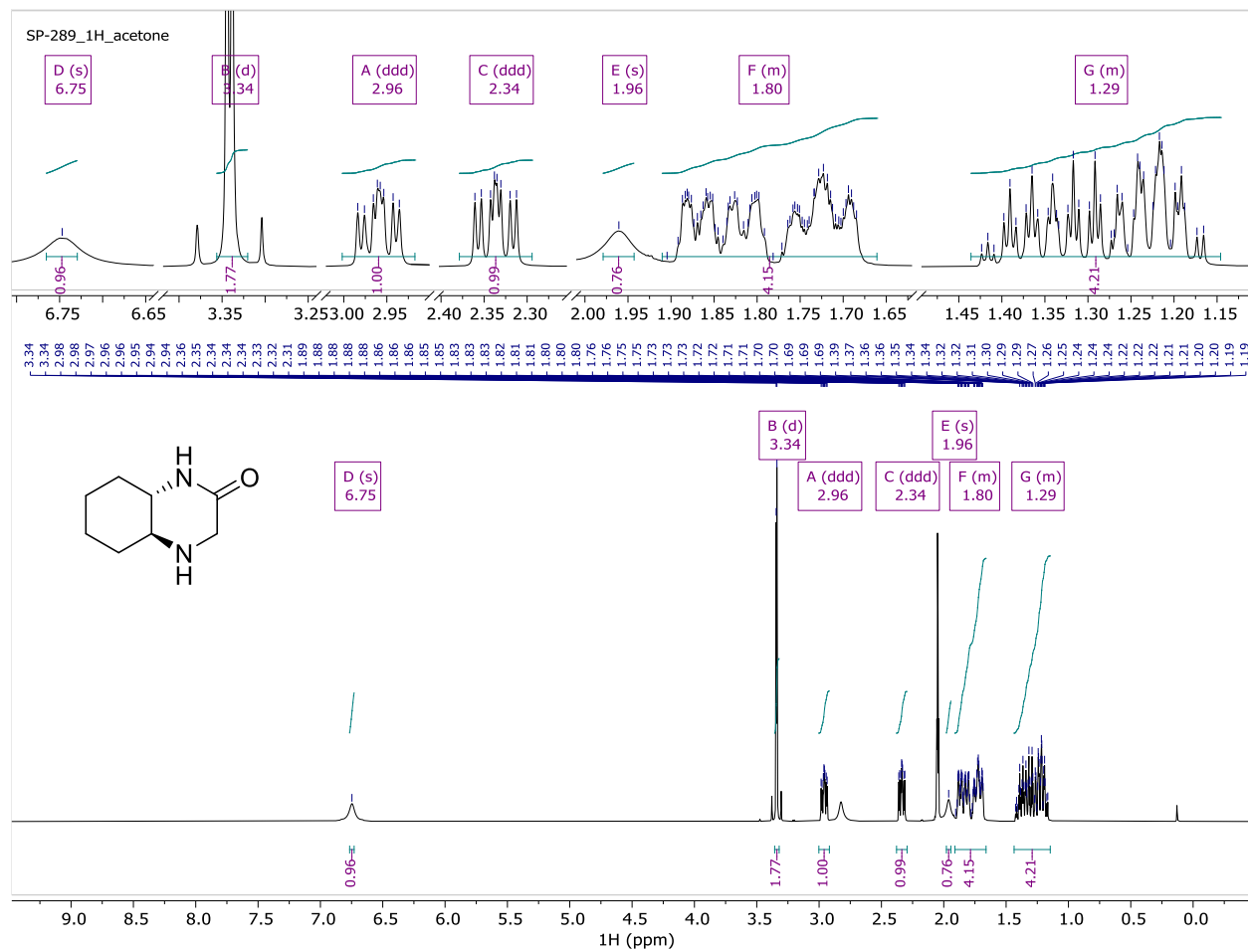
Supplementary information Chapter 3

Amine **3.20** ^1H NMR (500MHz) QANUC500 DMSO- d_6



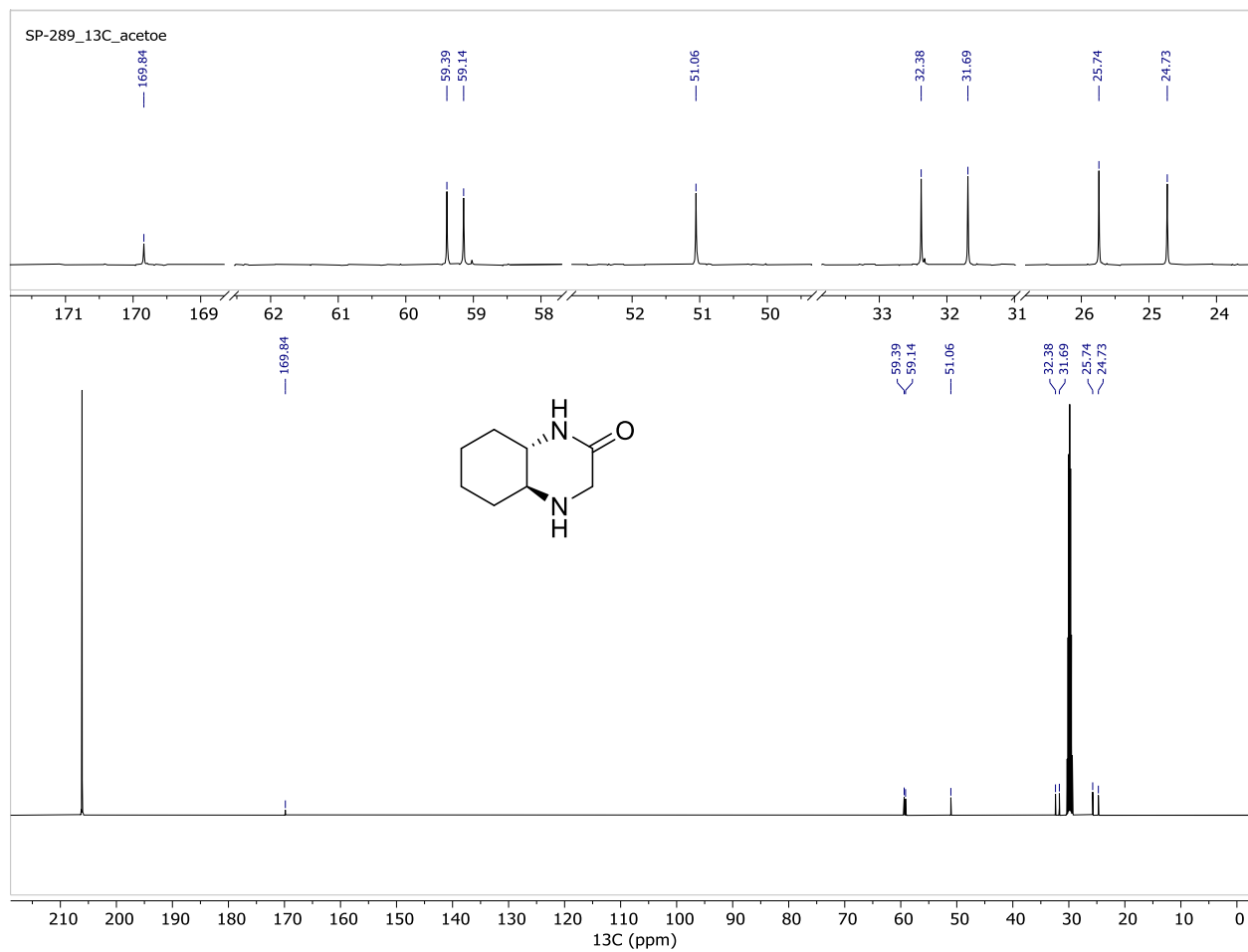
Supplementary information Chapter 3

Free amine **3.12** ^1H NMR (500MHz) Bruker500 Acetone- d_6



Supplementary information Chapter 3

Free amine **3.12** ^{13}C NMR (126MHz) Bruker500 Acetone- d_6



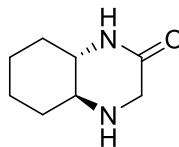
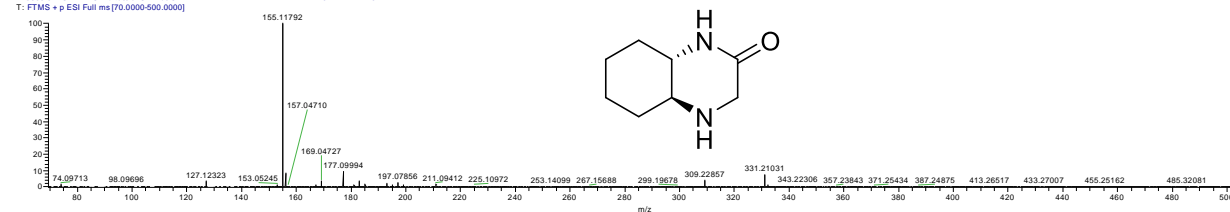
Supplementary information Chapter 3

Free amine **3.12** HRMS

220909-01ESI- HRMS-Moitessier

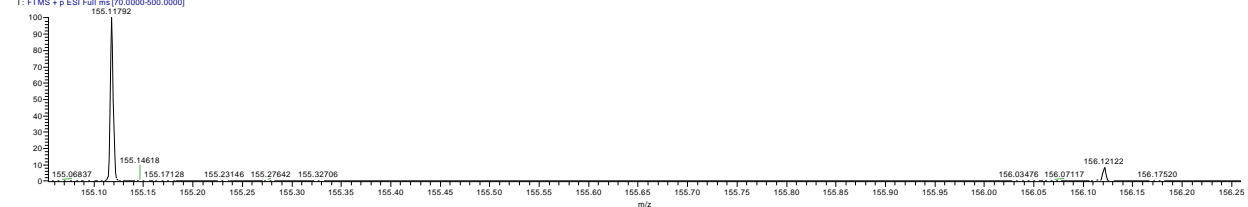
09/09/22 07:45:32

220909-01ESI- HRMS-Moitessier-S Pinus-SP-2
T: FTMS + p ESI Full ms [70.0000-500.0000]



220909-01ESI- HRMS-Moitessier-S Pinus-SP-289 #764-779 RT: 2.49-2.52 AV: 16 NL: 2.41E9

T: FTMS + p ESI Full ms [70.0000-500.0000]



220909-01ESI- HRMS-Moitessier-S Pinus-SP-289#764-779 RT: 2.49-2.52 AV: 16

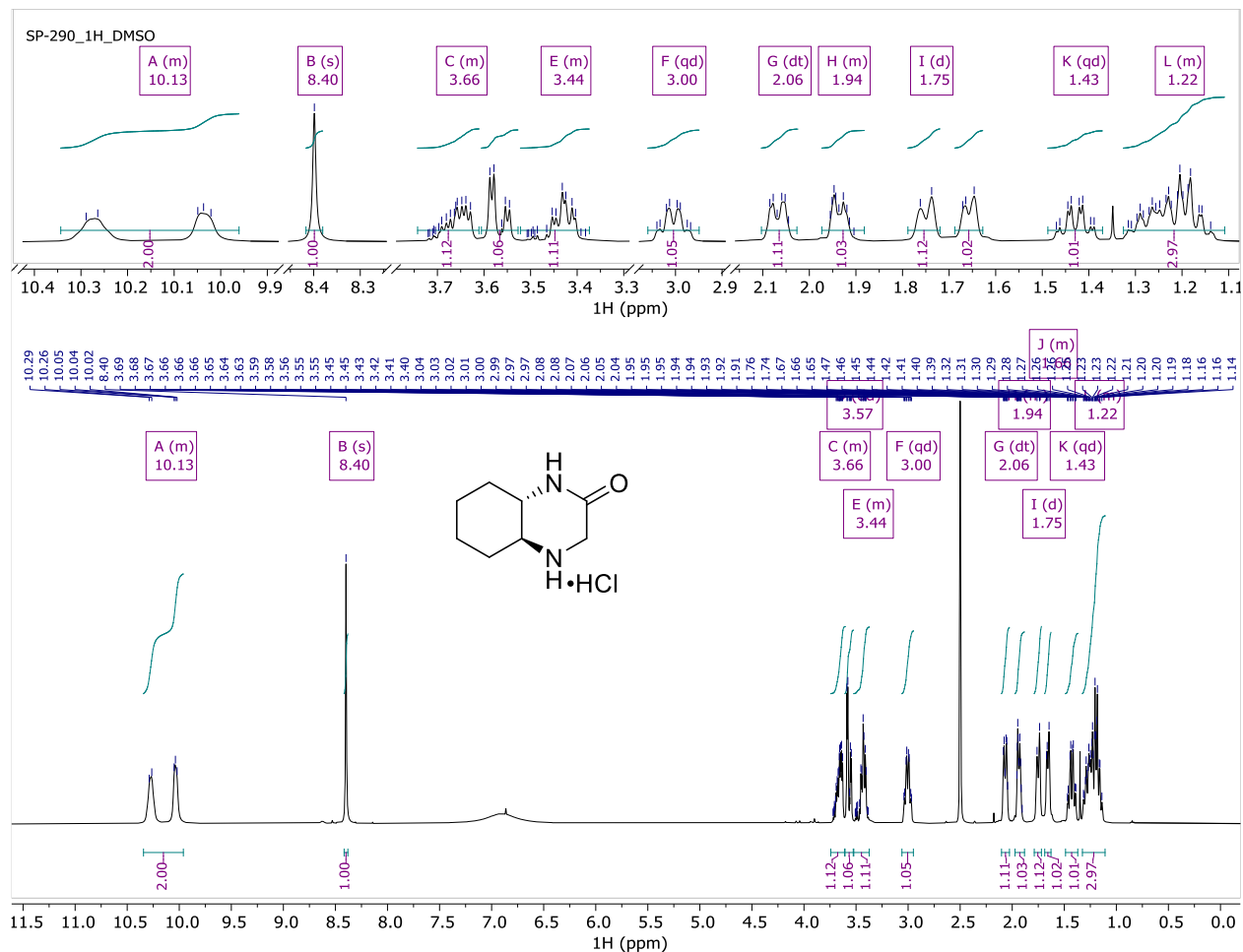
T: FTMS + p ESI Full ms [70.0000-500.0000]

m/z = 155.11337-155.12054

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
155.11792	2416333568.0	100.00	44845.52	1.00	155.11789	0.21	2.5	C ₈ H ₁₅ O N ₂

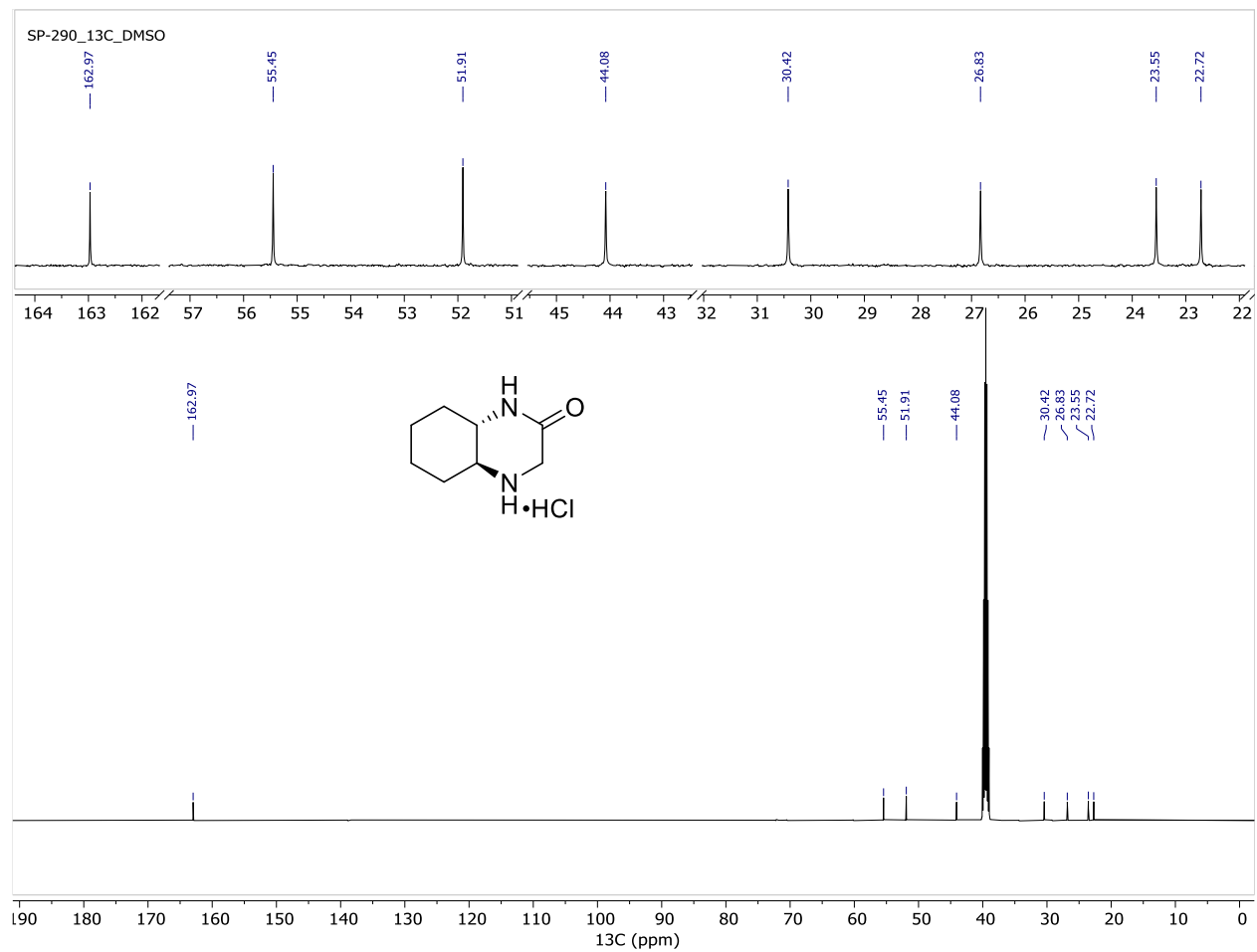
Supplementary information Chapter 3

Amine **3.12** HCl salt ^1H NMR (500MHz) Bruker500 DMSO- d_6



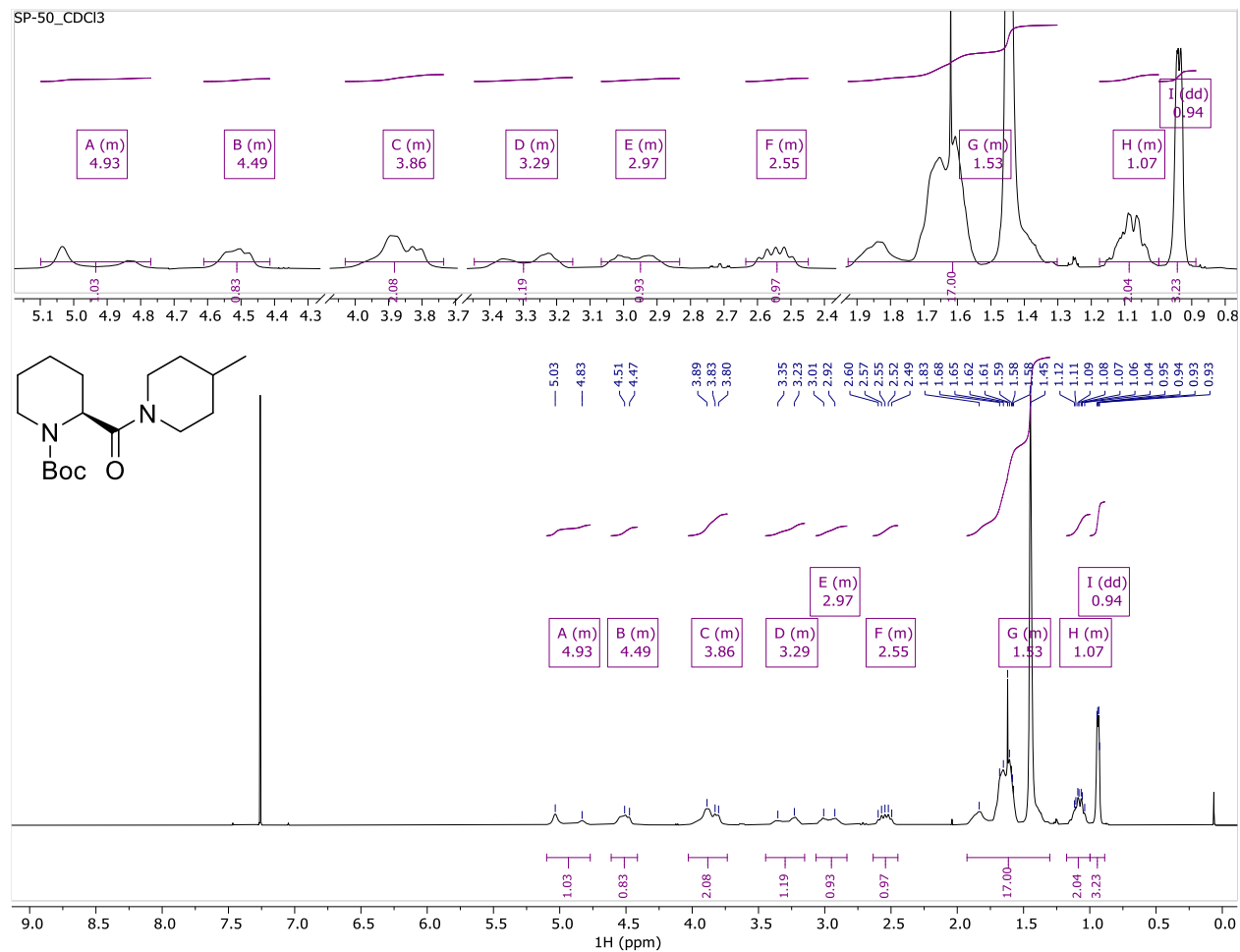
Supplementary information Chapter 3

Amine **3.12** HCl salt ^{13}C NMR (126MHz) Bruker500 DMSO- d_6



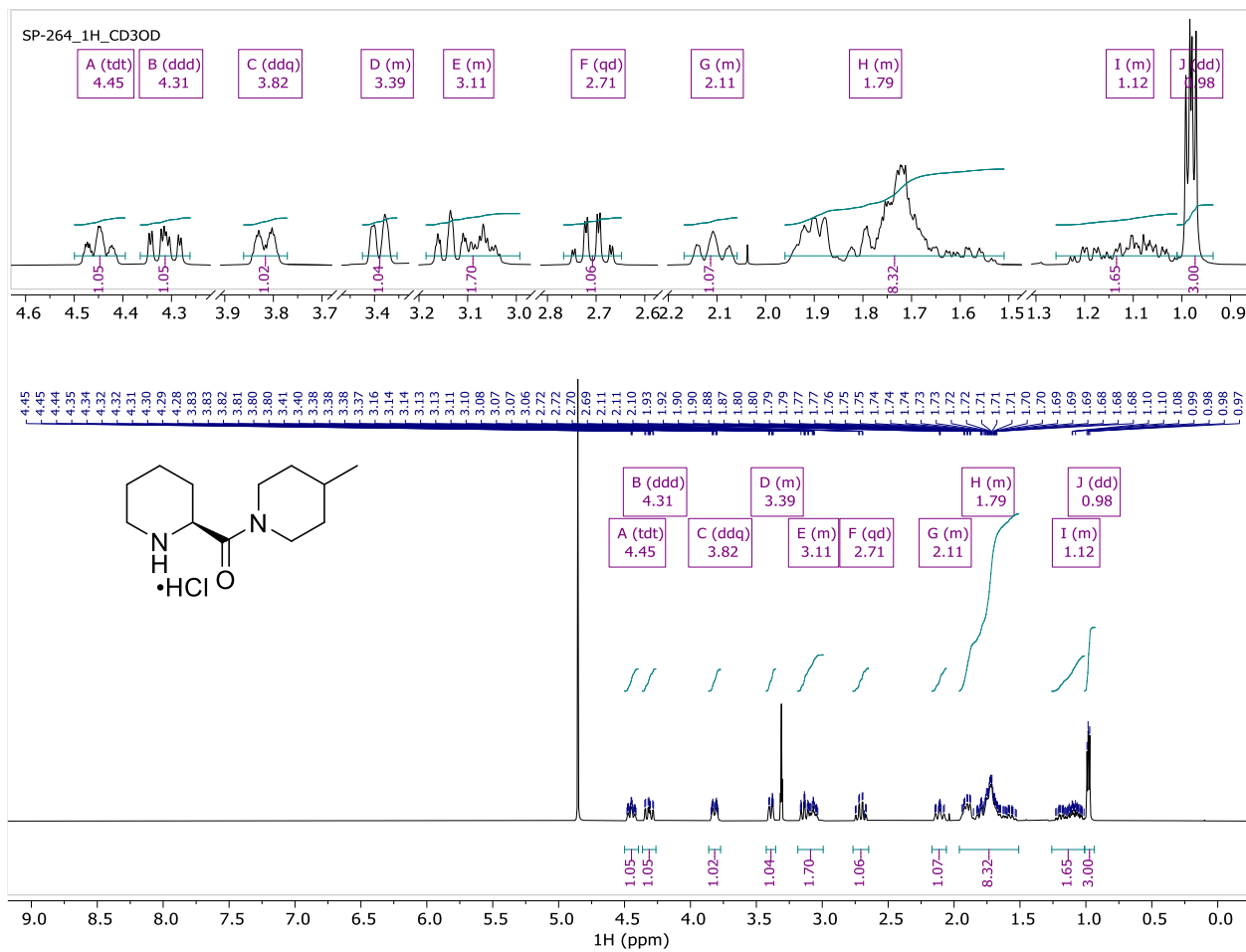
Supplementary information Chapter 3

Amine **3.23** with Boc-protecting group ^1H NMR (500MHz) VARIAN500 CDCl_3



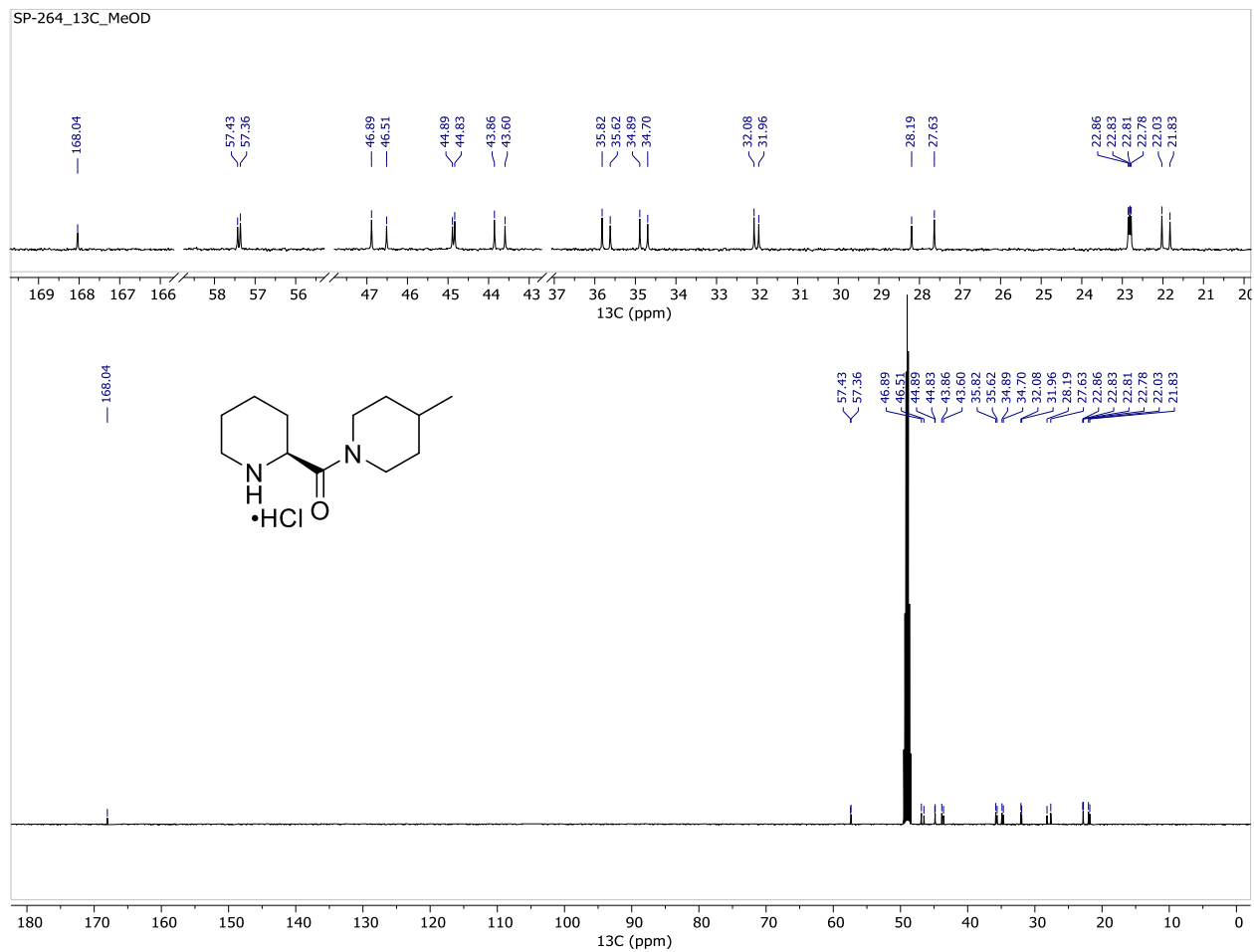
Supplementary information Chapter 3

Amine **3.23** HCl salt ^1H NMR (500MHz) Bruker500 in CD_3OD



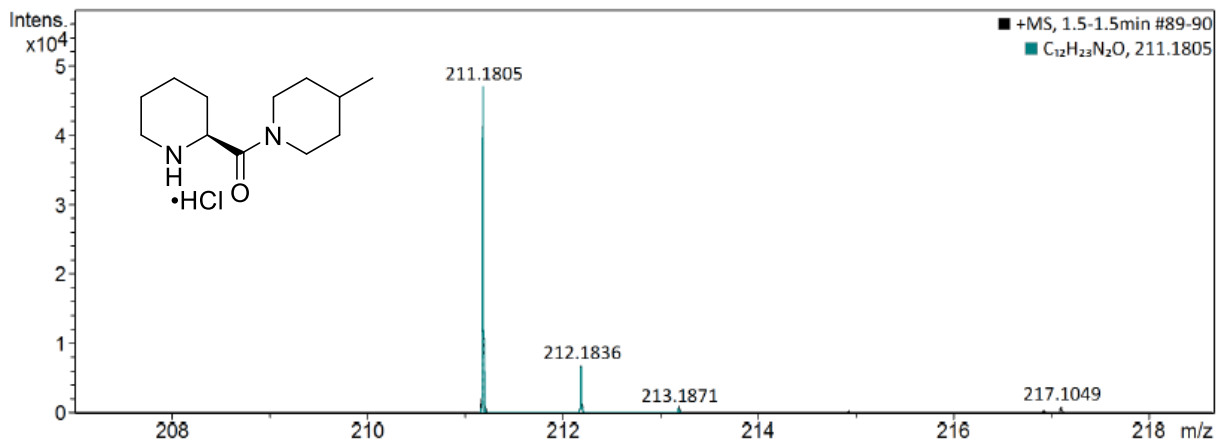
Supplementary information Chapter 3

Amine **3.23** HCl salt ^{13}C NMR (126MHz) Bruker500 DMSO- d_6



Supplementary information Chapter 3

Amine **3.23** HCl salt HRMS



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
211.1805	1	C ₁₂ H ₂₃ N ₂ O	211.1805	0.1	3.5	1	100.00	2.5	even	ok
	1	C ₁₀ H ₂₄ N ₂ NaO	211.1781	-11.3	15.9	1	100.00	-0.5	even	ok

Supplementary information Chapter 4

Table S1: Different reaction conditions used for optimization.

Reaction number	<i>Trans</i> -stilbene [mmol]	Catalyst [equiv.]	Addition time [min] or [h]	Additional stirring time [h]	Conversion [%]	Additional changes
SP-74 ^a	0.6	0.3	90 min	-	5	-
SP-77 ^a	1	0.3	100 min	-	18	-
SP-82 ^a	1	3	95 min	-	0	Acetone as ketone
SP-89 ^a	1	0.3	100 min	3.5 h	38	-
SP-94 ^a	0.3	0.3	30 min	20 h	- ^b	x 2 amount of Oxone
SP-95 ^a	0.3	0.3	30 min	24 h	- ^b	x 2 amount of TBAHS
SP-97 ^a	1	0.3	-	-	13	H ₂ O ₂ as oxidant
SP-99 ^a	0.2	3	-	-	27	Potion wise addition
SP-100	1	1	100 min	-	25	-
SP-101	1	1	2 h	-	35	-
SP-102 ^c	1	1	3.5 h	-	79	-
SP-103	1	0.3	3 h	-	50	-
SP-105	1	0.3	6 h	-	72	-

^a Unlike in the optimized conditions, in these reactions the catalyst was not pulverized, and the flask was not sonicated after addition of the solids.

^b Reaction was not worked up, based on TLC it was concluded there was no conversion.

^c Optimized conditions chosen. Using 0.3 equiv. required the reaction to take place over 6 hours (SP-105), and the conversion (72%) was comparable to using 1 equiv. for 3.5 hours. For the sake of efficiency, we chose to use 1 equiv. of catalyst, and compare other ketones tested to these conditions.

Optimized general procedure for the Shi Epoxidation:

In a three-necked flask, E-stilbene (180.25 mg, 1 mmol, 1 equiv.) was dissolved in CH₃CN-DCM (15 mL, 1:2 v/v). Mixture was sonicated and stirred for about 10 minutes (flask should be capped, to avoid loss of solvent), to ensure the stilbene dissolved. Subsequently were added: buffer (10 mL of 0.05M of Borax in 4x10⁻⁴ aq. Na₂EDTA), tetrabutylammonium hydrogensulfate (TBAHS) (15

mg, 0.04 mmol, 0.04 equiv.), Shi's ketone catalyst (258.27 mg, 1 mmol, 1 equiv.; catalyst was pulverized prior to addition). Mixture was sonicated for about 30 seconds, and then cooled down to 0°C. A syringe pump with two different syringes (equipped with Teflon needles); one containing Oxone (851.13 mg, 2.76 mmol, 2.76 equiv.) solution in 6.5 mL 4×10^{-4} aq. Na₂EDTA, the second one containing K₂CO₃ (801.589 mg, 5.8 mmol, 5.8 equiv.) solution in 6.5 mL H₂O. These two solutions were added to the reaction simultaneously over 3.5 hours at 0°C (reaction mixture should be stirring vigorously). Once the addition was completed, 30 mL of H₂O was added, and the product was extracted with Hexanes (x4, 40 mL). Combined organic layers were washed with brine (x2), dried (Na₂SO₄), filtered, and concentrated to give a yellowish pale oil which solidified upon standing. Conversion was analyzed by crude ¹H-NMR, unless specified otherwise.

Reaction with Shi's chiral ketone catalyst – optimized conditions

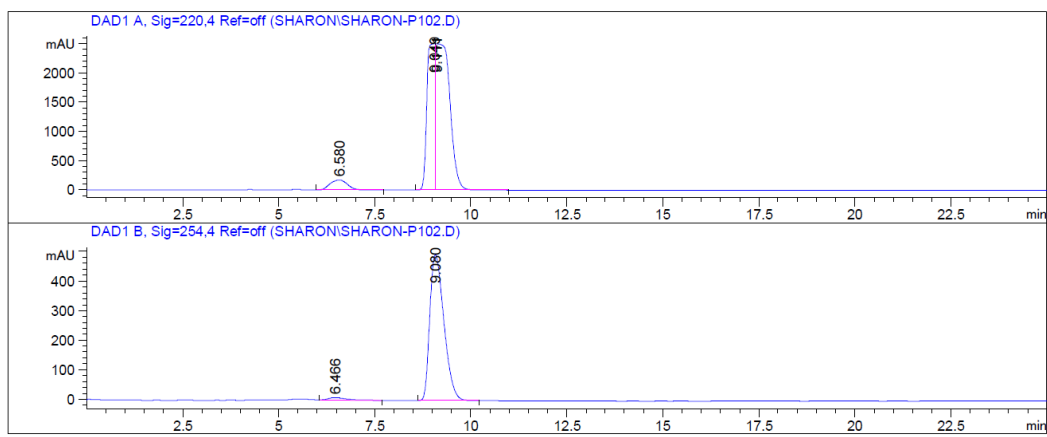
Reaction was run according to the optimized general procedure. Conversion was measured as 79%. Crude was purified by column chromatography: crude was dry loaded on silica (Et₂O and about 1% Et₃N), and run using hexanes: Et₃N 99:1 as eluent. 59% isolated yield. ¹H NMR (500 MHz, CDCl₃) δ 7.45 – 7.30 (m, 10H), 3.88 (s, 2H). Chiral separation was achieved using HPLC, Chiracel-OD column, hexanes: iso-propanol 80:20 eluent. 95.5 %*ee*.

Supplementary information Chapter 4

Sample Info : Hex:iPrOH 80:20, 1.0 mL/min, 10 uL inj., Chiracel OD, 25 min

Additional Info : Peak(s) manually integrated

=====



Signal 1: DAD1 A, Sig=220,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	6.580	VB	0.5161	5356.20508	169.76978	5.0901
2	9.049	BV	0.1773	3.62610e4	2498.39526	34.4596

Signal 2: DAD1 B, Sig=254,4 Ref=off

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	6.466	BB	0.4872	286.94458	8.10127	2.2340
2	9.080	BB	0.3999	1.25576e4	491.77200	97.7660

Totals : 1.28445e4 499.87327

Reaction with cyclohexanone

Reaction was run according to the optimized general procedure: cyclohexanone 1 mL, 10 mmol, 10 equiv.). Conversion was measured as 93%.

Reaction with cyclopentanone

Reaction was run according to the optimized general procedure: cyclohexanone (0.86 mL, 10 mmol, 10 equiv.). Conversion was measured as 0%, only starting material was observed.

Reaction with cycloheptanone

Supplementary information Chapter 4

Reaction was run according to the optimized general procedure: cycloheptanone (1.2 mL, 10 mmol, 10 equiv.). Conversion was measured as 13%.

Reaction with 3-pentanone

Reaction was run according to the optimized general procedure: 3-pentanone (1.1 mL, 10 mmol, 10 equiv.). Conversion was measured as 10%.

Supplementary information Chapter 5

(R)-5-Amino-2-methyl-N-(1-(naphthalen-1-yl)ethyl)benzamide (GRL0617). To a solution of 5-amino-2-methylbenzoic acid (200 mg, 1.32 mmol) in DCM (5.5 mL, 0.25M) under Argon at 0°C was added DIPEA (0.58 mL, 3.31 mmol, 2.5 equiv.), (R)-(+)-1-(1-naphthyl)ethylamine (0.23 mL, 1.45 mmol, 1.1 equiv.), and PyBOP (0.772 g, 1.45 mmol, 1.1 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Water and an aq. Saturated solution of NH₄Cl were added to the reaction mixture. Phases were separated and the organic phase was washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x2), and brine (x1). Combined organic layers were dried (Na₂SO₄), filtered, and concentrated to give an orange oil. Crude was purified by column chromatography on silica gel using 0-40% (10% MeOH/1% NH₄OH in DCM) in DCM as an eluent. Product was triturated with heptane and evaporated to afford a beige powder, 300 mg, 74% yield. ¹H NMR (500 MHz, CDCl₃) δ 8.23 (d, J = 8.5 Hz, 1H), 7.88 (dd, J = 8.1, 1.5 Hz, 1H), 7.81 (d, J = 8.2 Hz, 1H), 7.61 – 7.43 (m, 4H), 6.94 (d, J = 8.0 Hz, 1H), 6.62 – 6.54 (m, 2H), 6.11 (p, J = 7.0 Hz, 1H), 5.97 (d, J = 8.4 Hz, 1H), 3.53 (s, 2H), 2.29 (s, 3H), 1.77 (d, J = 6.7 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 169.18, 144.23, 138.22, 137.13, 134.09, 131.94, 131.34, 128.92, 128.60, 126.73, 126.08, 125.48, 125.32, 123.72, 122.69, 116.78, 113.50, 44.93, 20.78, 18.86. HRMS (ESI+) for C₂₀H₂₀N₂NaO (M+Na) calcd 327.1468 found 327.1465.

(R)-5-Amino-2-hydroxy-N-(1-(naphthalen-1-yl)ethyl)benzamide (5.2). To a solution of 5-aminosalicylic acid (300 mg, 1.96 mmol) in DCM (10 mL, 0.2M) under Argon at 0°C was added DIPEA (0.68 mL, 3.92 mmol, 2 equiv.), (R)-(+)-1-(1-naphthyl)ethylamine (0.35 mL, 2.15 mmol, 1.1 equiv.), and PyBOP (1.143 g, 2.15 mmol, 1.1 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. DCM and an aq. Saturated solution of NH₄Cl were added to the reaction mixture. Phases were separated and the organic phase was

washed with aq. saturated NH_4Cl (x2), aq. saturated NaHCO_3 (x2), and brine (x1). Combined organic layers were dried (Na_2SO_4), filtered, and concentrated to give an orange oil. Crude was purified by column chromatography on silica gel using 0-20% (10% MeOH/1% NH_4OH in DCM) in DCM as an eluent. Product was triturated with hexanes and evaporated to afford an off-white powder, 306 mg, 51% yield. ^1H NMR (800 MHz, CDCl_3) δ 11.64 (s, 1H), 8.11 – 8.08 (m, 1H), 7.91 – 7.87 (m, 1H), 7.84 (d, $J = 8.2$ Hz, 1H), 7.59 (dt, $J = 7.2, 1.0$ Hz, 1H), 7.57 – 7.45 (m, 3H), 6.81 (d, $J = 8.7$ Hz, 1H), 6.74 (dd, $J = 8.7, 2.7$ Hz, 1H), 6.55 (d, $J = 2.7$ Hz, 1H), 6.44 (d, $J = 7.7$ Hz, 1H), 6.07 (p, $J = 7.0$ Hz, 1H), 3.29 (s, 2H), 1.77 (d, $J = 6.8$ Hz, 3H). ^{13}C NMR (201 MHz, CDCl_3) δ 168.94, 154.77, 137.99, 137.72, 134.13, 131.20, 129.04, 128.87, 127.02, 126.18, 125.37, 123.27, 123.01, 122.87, 119.37, 114.23, 111.09, 45.17, 20.68. HRMS (ESI+) for $\text{C}_{19}\text{H}_{18}\text{N}_2\text{NaO}_2$ ($\text{M}+\text{Na}$) calcd 329.1258 found 329.1261.

(R)-5-Amino-2-methoxy-N-(1-(naphthalen-1-yl)ethyl)benzamide (5.3). To a suspension of 5-amino-2-methoxybenzoic acid (100 mg, 0.60 mmol, 1.2 equiv.) in acetonitrile (6 mL, 0.1M) under Argon at 0°C was added DIPEA (0.11 mL, 0.60 mmol, 1.2 equiv.), (R)-(+)-1-(1-naphthyl)ethylamine (0.08 mL, 0.50 mmol, 1 equiv.), and PyBOP (311.32 mg, 0.60 mmol, 1.2 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Solvent was removed under reduced pressure and ethyl acetate and an aq. Saturated solution of NH_4Cl were added. Phases were separated and the organic phase was washed with aq. saturated NH_4Cl (x2), aq. saturated NaHCO_3 (x2), and brine (x1). Combined organic layers were dried (Na_2SO_4), filtered, and concentrated to give a red oil. Crude was purified by column chromatography on silica gel using 20-100% ethyl acetate in hexanes as an eluent. Product was triturated with hexane and evaporated to afford an off-white powder, 117 mg, 73% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.39 (d, $J = 7.8$ Hz, 1H), 8.21 (d, $J = 8.4$ Hz, 1H), 7.89 – 7.84 (m, 1H), 7.79

(d, $J = 8.2$ Hz, 1H), 7.60 – 7.42 (m, 5H), 6.75 (d, $J = 2.5$ Hz, 2H), 6.11 (p, $J = 7.0$ Hz, 1H), 3.72 (s, 3H), 1.74 (d, $J = 6.8$ Hz, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 164.34, 150.87, 140.71, 139.42, 134.14, 131.17, 128.89, 128.16, 126.44, 125.83, 125.46, 123.81, 122.64, 122.45, 119.29, 118.70, 113.45, 56.78, 45.57, 21.71. HRMS (ESI+) for $\text{C}_{20}\text{H}_{21}\text{O}_2\text{N}_2$ calcd 321.15975 found 321.15966.

(R)-2-Hydroxy-N-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzamide (5.4). To a suspension of 5-nitrosalicylic acid (200 mg, 1.09 mmol) in DCM (10 mL, 0.1M) under Argon at 0°C was added EDCI (251.24 mg, 1.31 mmol, 1.2 equiv.). Reaction stirred for 30 minutes at 0°C (turned clear). HOBt· H_2O (54 mg, 0.328 mmol, 0.3 equiv.), .), (R)-(+)-1-(1-naphthyl)ethylamine (0.21 mL, 1.31 mmol, 1.2 equiv.), and DIPEA (0.19 mL, 1.09 mmol, 1 equiv.) were added. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. DCM and an aq. Saturated solution of NH_4Cl were added. Phases were separated and the organic phase was washed with aq. saturated NH_4Cl (x2), aq. saturated NaHCO_3 (x2), and brine (x1). Combined organic layers were dried (Na_2SO_4), filtered, and concentrated to give a yellow oil. Crude was purified by column chromatography on silica gel using 0-20% ethyl acetate in hexanes as an eluent. Product was triturated with heptane and evaporated to afford a yellow powder, 157 mg, 43% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.29 (d, $J = 2.6$ Hz, 1H), 8.23 (dd, $J = 9.2, 2.6$ Hz, 1H), 8.07 (d, $J = 8.4$ Hz, 1H), 7.92 – 7.82 (m, 2H), 7.62 (d, $J = 7.1$ Hz, 1H), 7.60 – 7.45 (m, 3H), 7.05 (d, $J = 9.2$ Hz, 1H), 6.77 (d, $J = 7.7$ Hz, 1H), 6.12 (p, $J = 7.0$ Hz, 1H), 1.82 (d, $J = 6.8$ Hz, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 167.61, 167.42, 139.26, 136.86, 134.16, 131.05, 129.46, 129.23, 129.18, 127.08, 126.24, 125.46, 123.07, 122.83, 122.43, 119.58, 113.57, 45.60, 20.67. HRMS (ESI+) for $\text{C}_{19}\text{H}_{16}\text{N}_2\text{NaO}_4$ (M+Na) calcd 359.1002 found 359.1005.

(R)-5-chloro-2-hydroxy-N-(1-(naphthalen-1-yl)ethyl)benzamide (5.10). To a solution of 5-chlorosalicylic acid (200 mg, 1.16 mmol, 1.2 equiv.) in DCM (10 mL, 0.1M) under Argon at 0°C

was added DIPEA (0.17 mL, 0.96 mmol, 1 equiv.), and HATU (440.67 mg, 1.16 mmol, 1.2 equiv.). DMF (0.5 mL) was added to help dissolve the starting material. Mixture stirred at 0°C for 20 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.15 mL, 0.96 mmol, 1 equiv.), and DIPEA (0.17 mL, 0.96 mmol, 1 equiv.) were added. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. DCM and water were added. Phases were separated and the aqueous phase was extracted with DCM (x3). Combined organic phases were washed with aq. saturated NH₄Cl (x1), aq. saturated NaHCO₃ (x1), and brine (x1), dried (Na₂SO₄), filtered, and concentrated to give an orange oil. Crude was purified by column chromatography on silica gel using 2-20% ethyl acetate in hexanes as an eluent to afford an off-white powder, 169 mg, 54% yield. ¹H NMR (500 MHz, CDCl₃) δ 12.29 (s, 1H), 8.11 – 8.06 (m, 1H), 7.93 – 7.83 (m, 2H), 7.64 – 7.47 (m, 4H), 7.31 (dd, J = 8.9, 2.5 Hz, 1H), 7.22 (d, J = 2.5 Hz, 1H), 6.93 (d, J = 8.8 Hz, 1H), 6.42 (d, J = 7.8 Hz, 1H), 6.10 (p, J = 7.0 Hz, 1H), 1.79 (d, J = 6.8 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 168.01, 160.41, 137.27, 134.29, 134.17, 131.12, 129.15, 129.06, 127.07, 126.23, 125.41, 125.08, 123.40, 123.04, 122.93, 120.33, 115.14, 45.37, 20.70. HRMS (ESI+) for C₁₉H₁₅O₂NCl calcd 324.07968 found 324.07927.

(R)-4-chloro-2-((1-(naphthalen-1-yl)ethyl)carbamoyl)phenyl acetate (5.14). To a solution of (R)-5-chloro-2-hydroxy-N-(1-(naphthalen-1-yl)ethyl)benzamide (1e) (73.7 mg, 0.23 mmol, 1 equiv.) in DCM (2 mL, 0.1 M) under Argon at 0°C was added triethyl amine (67 μL, 0.48 mmol, 2.1 equiv.) and acetyl chloride (18 μL, 0.25 mmol, 1.1 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Water and DCM were added. Phases were separated and the aqueous phase was extracted with DCM (x3). Combined organic layers were washed with aq. Saturated NH₄Cl (x1), aq. saturated NaHCO₃ (x1), and brine (x1), dried (Na₂SO₄), filtered, and concentrated to give a white solid, 52 mg, 61% yield ¹H NMR (500 MHz,

CDCl_3) δ 8.16 – 8.10 (m, 1H), 7.91 – 7.80 (m, 3H), 7.60 – 7.45 (m, 4H), 7.38 (ddd, $J = 8.7, 2.7, 1.5$ Hz, 1H), 6.95 (dd, $J = 8.7, 0.9$ Hz, 1H), 6.59 – 6.50 (m, 1H), 6.10 (p, $J = 7.0$ Hz, 1H), 1.77 (d, $J = 6.7$ Hz, 3H), 1.54 (s, 4H). ^{13}C NMR (126 MHz, CDCl_3) δ 168.66, 163.04, 146.24, 137.59, 134.12, 132.15, 131.81, 131.32, 130.38, 129.66, 128.95, 128.88, 127.13, 126.32, 125.38, 124.74, 123.54, 122.93, 45.13, 20.25, 20.00. HRMS (ESI+) for $\text{C}_{21}\text{H}_{18}\text{O}_3\text{NCINa}$ calcd 390.08674 found 390.08617.

(R)-2-chloro-N-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzamide (5.6). To a solution of 5-chloronitrobenzoic acid (100 mg, 0.496 mmol, 1.2 equiv.) in acetonitrile (5 mL, 0.1M) under Argon at 0°C was added DIPEA (0.08 mL, 0.455 mmol, 1.1 equiv.), and HATU (187 mg, 0.496 mmol, 1.2 equiv.). Mixture stirred at 0°C for 25 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.07 mL, 0.413 mmol, 1 equiv.) was added, a precipitate formed soon after the addition. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Solvent was removed under reduced pressure, ethyl acetate was added, and precipitate was filtered and washed with ethyl acetate (precipitate is clean product, white solid). Filtrate was then washed with aq. saturated NH_4Cl (x3), aq. saturated NaHCO_3 (x2), and brine (x2), dried (Na_2SO_4), filtered, and concentrated to give an orange solid. Washed crude was recrystallized from ethyl acetate to afford a white fluffy solid, combined product, 104 mg, 71% yield. ^1H NMR (500 MHz, DMSO) δ 9.32 (d, $J = 7.8$ Hz, 1H), 8.31 – 8.17 (m, 3H), 7.97 (dd, $J = 8.1, 1.4$ Hz, 1H), 7.87 (d, $J = 8.1$ Hz, 1H), 7.82 (d, $J = 8.8$ Hz, 1H), 7.67 – 7.50 (m, 4H), 5.91 (p, $J = 7.0$ Hz, 1H), 1.62 (d, $J = 6.9$ Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 163.40, 145.96, 139.21, 137.75, 137.15, 133.39, 131.27, 130.35, 128.70, 127.53, 126.29, 125.71, 125.48, 125.38, 123.54, 123.15, 122.69, 44.88, 21.27. HRMS (ESI+) for $\text{C}_{19}\text{H}_{15}\text{ClN}_2\text{NaO}_3$ calcd 377.0663 found 377.0657.

(R)-2-chloro-5-methyl-N-(1-(naphthalen-1-yl)ethyl)nicotinamide (5.11). To a solution of 5-chloro-5-methylnicotinic acid (150 mg, 0.874 mmol, 1.2 equiv.) in acetonitrile (8.5 mL, 0.1M) under Argon at 0°C was added DIPEA (0.15 mL, 0.874 mmol, 1.2 equiv.), and HATU (335 mg, 0.881 mmol, 1.2 equiv.). Mixture stirred at 0°C for 25 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.12 mL, 0.728 mmol, 1 equiv.) was added. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Solvent was removed under reduced pressure, ethyl acetate was added and crude was washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x2), and brine (x1), dried (Na₂SO₄), filtered, and concentrated to give a dark orange solid. Crude was recrystallized from methanol to afford a beige solid, 145 mg, 61% yield ¹H NMR (500 MHz, CDCl₃) δ 8.23 (d, J = 2.4 Hz, 1H), 8.21 – 8.17 (m, 1H), 7.91 – 7.86 (m, 2H), 7.82 (d, J = 8.1 Hz, 1H), 7.62 – 7.55 (m, 2H), 7.55 – 7.44 (m, 2H), 6.71 (d, J = 8.0 Hz, 1H), 6.13 (p, J = 7.0 Hz, 1H), 2.32 (s, 3H), 1.81 (d, J = 6.8 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 163.91, 151.31, 144.34, 140.35, 137.68, 134.12, 133.03, 131.15, 130.50, 129.02, 128.79, 126.80, 126.12, 125.40, 123.48, 122.90, 45.98, 20.82, 17.65. HRMS (ESI+) for C₁₉H₁₇ClN₂NaO calcd 347.0927 found 347.0925.

(R)-2-bromo-N-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzamide (5.7). To a solution of 2-bromo-5-nitrobenzoic acid (400 mg, 1.63 mmol, 1.2 equiv.) in acetonitrile (10 mL, 0.1M) under Argon at 0°C was added DIPEA (0.28 mL, 1.63 mmol, 1.2 equiv.), and HATU (620 mg, 1.63 mmol, 1.2 equiv.). Mixture stirred at 0°C for 25 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.22 mL, 1.36 mmol, 1 equiv.) was added. Reaction was allowed back to room temperature and stirred at room temperature for 19 hours. Solvent was removed under reduced pressure, solid was washed with ethanol and recrystallized from ethyl acetate to afford beige needle like crystals, 435 mg, 80% yield ¹H NMR (500 MHz, DMSO) δ 9.30 (d, J = 7.9 Hz, 1H), 8.22 (d, J = 8.5 Hz, 1H), 8.17 (dd, J = 8.7, 2.8 Hz, 1H), 8.13 (d, J = 2.7 Hz, 1H), 8.01 – 7.94 (m, 2H), 7.87 (d, J = 8.1 Hz, 1H), 7.68 –

7.60 (m, 2H), 7.59 – 7.51 (m, 2H), 5.90 (p, $J = 7.0$ Hz, 1H), 1.63 (d, $J = 6.9$ Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 164.39, 146.46, 139.98, 139.13, 134.47, 133.39, 130.36, 128.69, 127.54, 126.83, 126.33, 125.71, 125.48, 125.22, 123.22, 123.18, 122.77, 44.84, 21.24. HRMS (ESI+) for $\text{C}_{19}\text{H}_{15}\text{O}_3\text{N}_2\text{BrNa}$ calcd 421.01583 found 421.-1570.

(R)-5-amino-2-bromo-N-(1-(naphthalen-1-yl)ethyl)benzamide (5.8). To a suspension of 5-amino-2-bromobenzoic acid (500 mg, 2.31 mmol, 1.2 equiv.) in acetonitrile (9 mL, 0.2M) under Argon at 0°C were added DIPEA (0.4 mL, 2.31 mmol, 1.2 equiv.), PyBOP (1.24 gr, 2.38 mmol, 1.2 equiv.), and (R)-(+)-1-(1-naphthyl)ethylamine (0.31 mL, 1.93 mmol, 1 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 20 hours. Solvent was removed under reduced pressure, and ethyl acetate was added. Organic phase was washed with aq. saturated NH_4Cl (x3), then with aq. saturated NaHCO_3 (x2), aq. saturated NaCl (x2), and water (x1). Organic phase was then dried (Na_2SO_4), filtered and concentrated to give an orange solid. Crude was purified by flash column chromatography with 60 – 100% ethyl acetate in hexanes as eluent to give a pale orange solid, 496 mg, 70% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.21 (dd, $J = 8.6, 3.7$ Hz, 1H), 7.90 – 7.77 (m, 2H), 7.57 (ddd, $J = 8.6, 6.8, 1.4$ Hz, 2H), 7.53 – 7.41 (m, 2H), 7.22 (d, $J = 8.5$ Hz, 1H), 6.76 (d, $J = 2.9$ Hz, 1H), 6.50 (dd, $J = 8.6, 2.9$ Hz, 1H), 6.28 (d, $J = 8.3$ Hz, 1H), 6.14 – 6.06 (m, 1H), 3.51 (s, 3H), 1.79 (d, $J = 6.8$ Hz, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 166.87, 146.05, 137.99, 137.88, 134.02, 131.24, 128.91, 128.64, 126.73, 126.06, 125.34, 123.74, 122.89, 118.03, 115.93, 106.55, 45.65, 20.72. HRMS (ESI+) for $\text{C}_{19}\text{H}_{18}\text{ON}_2\text{Br}$ calcd 369.05970 found 369.05927.

tert-butyl (R)-(4-bromo-3-((1-(naphthalen-1-yl)ethyl)carbamoyl)phenyl)carbamate (5.9). To a solution of (R)-5-amino-2-bromo-N-(1-(naphthalen-1-yl)ethyl)benzamide (**5.8**, 250 mg, 0.68 mmol, 1 equiv.) in ethanol (1.5 mL, 0.5M) under Argon, was added Boc_2O (0.23 mL, 1.02 mmol,

1.5 equiv.). Reaction stirred at room temperature for 25 hours. Solvent was removed under reduced pressure to give a beige solid. Crude was used in the next step without further purification.

Ethyl (E)-4-((tert-butoxycarbonyl)amino)but-2-enoate (5.29). To a suspension of NaH (60% suspension in oil, 32 mg, 0.75 mmol, 1.6 equiv.) and molecular sieves in THF (3.7 mL) under Argon at 0°C was added triethyl phosphonoacetate (0.14 mL, 0.71 mmol, 1.5 equiv.). The solution stirred at 0°C for 20 minutes. N-Boc-2-aminoacetaldehyde (73.5 mg, 0.46 mmol, 1 equiv.) was dissolved in THF (5 mL, final reaction concentration is 0.17 M) and added dropwise to the reaction mixture at 0°C. The reaction stirred at 0°C for 1 hour, the ice bath was then removed, and the reaction stirred at room temperature for 2 hours. When the starting material appeared to be consumed by TLC (4:1 Hexanes: Ethyl Acetate, R_f=0.28) the reaction was quenched with ethanol. Solvent was then removed by evaporation to give the crude product as a yellow oil. Crude was purified by column chromatography on silica gel using 4-40% ethyl acetate in hexanes as an eluent to afford a colorless oil, 31 mg, 31% yield. ¹H NMR (500 MHz, CDCl₃) δ 6.90 (dt, J = 15.7, 4.9 Hz, 1H), 5.93 (dt, J = 15.7, 2.0 Hz, 1H), 4.69 (s, 1H), 4.19 (q, J = 7.1 Hz, 2H), 3.91 (d, J = 6.0 Hz, 2H), 1.45 (s, 9H), 1.28 (t, J = 7.1 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 166.10, 155.57, 144.73, 121.37, 79.92, 60.46, 41.35, 28.36, 14.24. HRMS (ESI⁺) for C₁₁H₁₉O₄NNa calcd 252.12063 found 252.12053.

Ethyl (E)-4-aminobut-2-enoate trifluoroacetic acid salt (5.30). A solution of Ethyl (E)-4-((tert-butoxycarbonyl)amino)but-2-enoate (5.29, 50 mg, 0.22 mmol, 1 equiv.) in DCM (0.5 mL, 0.4 M) was cooled to 0°C, and TFA (0.07 mL, 0.92 mmol, 4 equiv.) was added dropwise. The reaction stirred at 0°C for 10 minutes, and then allowed back to room temperature for 2 hours. Solvent was removed to give an orange oil. Product was used for the next reaction without further purification.

(R)-2-methyl-N-(1-(naphthalen-1-yl)ethyl)-5-(trifluoromethyl)benzamide (5.12). To a solution of 2-methyl-5-(trifluoromethyl)benzoic acid (200 mg, 0.980 mmol, 1.2 equiv.) in acetonitrile (6.5 mL, 0.15M) under Argon at 0°C was added DIPEA (0.17 mL, 0.980 mmol, 1.2 equiv.), and HATU (373 mg, 0.980 mmol, 1.2 equiv.). Mixture stirred at 0°C for 30 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.13 mL, 0.817 mmol, 1 equiv.) was added. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Solvent was removed under reduced pressure, ethyl acetate was added, and precipitate was filtered and washed with ethyl acetate (precipitate is clean product, beige solid). Filtrate was then washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x2), and brine (x1), dried (Na₂SO₄), filtered, and concentrated to give an orange solid. Washed crude was recrystallized from isopropanol to afford a pale yellow solid, combined product, 141 mg, 48% yield. ¹H NMR (500 MHz, CDCl₃) δ 8.22 (d, J = 8.5 Hz, 1H), 7.89 (dd, J = 8.1, 1.4 Hz, 1H), 7.83 (d, J = 8.1 Hz, 1H), 7.63 – 7.44 (m, 6H), 7.30 (d, J = 7.8 Hz, 1H), 6.15 (p, J = 6.9 Hz, 1H), 6.00 (d, J = 8.4 Hz, 1H), 2.47 (s, 3H), 1.82 (d, J = 6.7 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 167.72, 140.45, 137.68, 137.14, 134.15, 131.55, 131.28, 129.06, 128.86, 128.53, 128.26, 126.84, 126.60, 126.57, 126.54, 126.51, 126.21, 125.34, 125.01, 123.63, 123.60, 123.57, 123.54, 123.48, 122.85, 45.25, 20.68, 19.91. HRMS (ESI+) for C₂₁H₁₈F₃NNaO calcd 380.1233 found 380.1232.

(R)-2-iodo-N-(1-(naphthalen-1-yl)ethyl)-5-(trifluoromethyl)benzamide (5.13). To a solution of 2-iodo-5-(trifluoromethyl)benzoic acid (400 mg, 1.266 mmol, 1.2 equiv.) in DCM (8 mL, 0.16M) under Argon at 0°C was added DIPEA (0.22 mL, 1.266 mmol, 1.2 equiv.), and HATU (480 mg, 1.266 mmol, 1.2 equiv.). Mixture stirred at 0°C for 25 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.17 mL, 1.055 mmol, 1 equiv.) was added. Reaction was allowed back to room temperature and stirred at room temperature for 18 hours. Solvent was removed under

reduced pressure, ethyl acetate was added, and precipitate was filtered and washed with ethyl acetate (precipitate is clean product, white solid). Filtrate was then washed with aq. saturated NH_4Cl (x2), aq. saturated NaHCO_3 (x2), and brine (x1), dried (Na_2SO_4), filtered, and concentrated to give an orange solid. Washed crude was recrystallized from isopropanol to afford a pale orange solid, combined product, 431 mg, 87% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.25 (d, $J = 9.0$ Hz, 1H), 7.96 (d, $J = 8.3$ Hz, 1H), 7.89 (dd, $J = 8.2, 1.4$ Hz, 1H), 7.83 (d, $J = 8.2$ Hz, 1H), 7.64 – 7.57 (m, 2H), 7.57 – 7.51 (m, 2H), 7.48 (dd, $J = 8.2, 7.2$ Hz, 1H), 7.29 (dd, $J = 8.3, 2.2$ Hz, 1H), 6.20 – 6.10 (m, 1H), 6.03 (d, $J = 8.3$ Hz, 1H), 1.86 (d, $J = 6.7$ Hz, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 167.14, 142.99, 140.75, 137.24, 134.12, 131.26, 131.11, 130.84, 129.01, 128.95, 127.58, 127.55, 126.97, 126.24, 125.32, 125.03, 125.00, 123.70, 123.11, 122.46, 97.02, 45.76, 20.45. HRMS (ESI+) for $\text{C}_{20}\text{H}_{15}\text{F}_3\text{INNaO}$ calcd 492.0043 found 492.0055.

(R)-3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-(trifluoromethyl)phenyl)amino)propanoic acid (5.28). (R)-2-iodo-N-(1-(naphthalen-1-yl)ethyl)-5-(trifluoromethyl)benzamide (SP-307, 75 mg, 0.160 mmol, 1 equiv.), β -alanine (16 mg, 0.176 mmol, 1.1 equiv.), Cs_2CO_3 (stored in the oven, 130 mg, 0.4 mmol, 2.5 equiv.), CuI (15 mg, 0.08 mmol, 0.5 equiv.) were added to a flask. The flask was capped and purged with vacuum and argon (x3). DMF was added (1 mL, 0.16M) and solution was degassed by bubbling argon into it while stirring. Flask was kept under argon atmosphere and heated to 105°C . Reaction was monitored by TLC and was completed after 3 hours. Reaction mixture was then cooled to room temperature and filtered through celite using ethyl acetate, filtrate is a pale green solution. Water was added to the filtrate and pH was adjusted to 3 using 1M HCl . Phases were separated and aqueous phase was extracted with ethyl acetate (x5). Combined organic layers were washed with acidic water (pH 3) (x3), brine (x2), dried (Na_2SO_4), filtered and concentrated to give an orange

oil. Crude was co-evaporated with heptane to remove as much DMF as possible. Crude was purified by flash column chromatography on silica gel using 0-5% methanol in DCM as eluent. Solvent was removed from combined fractions to give a pale orange solid, 19 mg, 28% yield. ¹H NMR (500 MHz, CDCl₃) δ 8.11 (d, J = 8.4 Hz, 1H), 7.87 (dd, J = 8.0, 1.5 Hz, 1H), 7.81 (d, J = 8.2 Hz, 1H), 7.59 – 7.45 (m, 6H), 6.71 (d, J = 9.3 Hz, 1H), 6.36 (d, J = 7.6 Hz, 1H), 6.04 (p, J = 6.9 Hz, 1H), 3.51 (t, J = 6.9 Hz, 2H), 2.68 (t, J = 6.9 Hz, 2H), 1.74 (d, J = 6.8 Hz, 3H). ¹³C NMR (126 MHz, CDCl₃) δ 176.23, 167.91, 151.36, 138.02, 134.14, 131.14, q: (129.81, 129.78, 129.75, 129.72, ³J_{C-F}=3.51 Hz), 129.07, 128.71, q: (127.84, 125.69, 123.54, 121.39, ¹J_{C-F}=270 Hz), 126.81, 126.07, 125.44, q: (124.96, 124.93, 124.90, 124.87, ³J_{C-F}=3.84 Hz), 123.23, 122.75, q: (117.16, 116.90, 116.63, 116.37, ²J_{C-F}=33.07 Hz), 114.79, 111.22, 53.57, 45.39, 38.39, 21.02. . HRMS (ESI+) for C₂₃H₂₁F₃N₂NaO₃ calcd 453.1396 found 453.1396.

Ethyl (R,E)-4-(3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-(trifluoromethyl)phenyl)amino)-propanamido)but-2-enoate (5.1a). To a solution of (R)-3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-(trifluoromethyl)phenyl)amino)propanoic acid (**5.28**, 88 mg, 0.204 mmol, 1.2 equiv.) in DCM (0.6 mL) under Argon at 0°C were added DIPEA (0.04 mL, 0.229 mmol, 1.3 equiv.) and PyBOP (106 mg, 0.204 mmol, 1.2 equiv.). In a separate flask, ethyl (E)-4-aminobut-2-enoate trifluoroacetic acid salt (**5.30**, 43 mg, 0.177 mmol, 1 equiv.) was dissolved in DCM (0.6 mL) and DIPEA (0.04 mL, 0.229 mmol, 1.3 equiv.) was added to neutralize the salt. Neutralized amine mixture was added to reaction mixture (DCM final amount 1.2 mL, 0.17M, DIPEA 0.08 mL, 0.458 mmol, 2.6 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 18 hours. Solvent was removed under reduced pressure and ethyl acetate and an aq. Saturated solution of NH₄Cl were added. Phases were separated and the organic phase was washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x1), and brine

(x1). Combined organic layers were dried (Na_2SO_4), filtered, and concentrated to give an orange solid. Crude was purified by flash column chromatography on silica gel using 5-60% ethyl acetate in DCM as an eluent. The product was co-evaporated with chloroform to afford a white solid, 36 mg, 37% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.18 (d, $J = 8.4$ Hz, 1H), 7.91 – 7.83 (m, 2H), 7.80 (d, $J = 8.2$ Hz, 1H), 7.59 – 7.41 (m, 6H), 6.73 (d, $J = 8.7$ Hz, 2H), 6.62 (dt, $J = 15.7, 5.1$ Hz, 1H), 6.06 (p, $J = 7.0$ Hz, 1H), 5.70 (t, $J = 5.9$ Hz, 1H), 5.64 (dt, $J = 15.7, 1.9$ Hz, 1H), 4.11 (q, $J = 7.1$ Hz, 2H), 3.93 (dddd, $J = 17.5, 6.2, 5.1, 1.9$ Hz, 1H), 3.81 (dtd, $J = 17.5, 5.5, 1.9$ Hz, 1H), 3.53 (ddt, $J = 16.1, 13.3, 6.7$ Hz, 2H), 2.51 – 2.36 (m, 2H), 1.74 (d, $J = 6.8$ Hz, 3H), 1.23 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (126 MHz, CDCl_3) δ 170.77, 167.98, 166.08, 151.21, 143.61, 138.39, 134.12, 131.17, q: (129.63, 129.60, 129.58, 129.55, $^3J_{\text{C-F}}=3.66$ Hz), 129.11, 128.53, q: (127.85, 125.70, 123.55, 121.40, $^1J_{\text{C-F}}=270$ Hz), 126.64, 126.00, 125.51, q: (125.15, 125.13, 125.09, 125.06, $^3J_{\text{C-F}}=3.66$ Hz), 123.36, 122.79, 121.65, q: (117.16, 116.90, 116.63, 116.37, $^2J_{\text{C-F}}=32.92$ Hz), 115.18, 111.40, 60.66, 45.14, 40.13, 39.15, 36.53, 21.05, 14.27. HRMS (ESI+) for $\text{C}_{29}\text{H}_{30}\text{O}_4\text{N}_3\text{F}_3\text{Na}$ calcd 564.20806 found 564.20820.

(R)-tert-butyl (4-chloro-2-((1-(naphthalen-1-yl)ethyl)carbamoyl)phenyl) succinate (5.22). To the stirred solution of mono-tert-succinate (58.7 mg, 0.337 mmol, 1.1 equiv.) in an anhydrous DCM (10 mL) at 0°C was added DMAP (64.6 mg, 0.337 mmol, 1.1 equiv) and EDCI (41.2 mg, 0.337 mmol, 1.1 equiv.). The mixture stirred at 0°C for 5 minutes, before (R)-5-chloro-2-hydroxy-N-(1-(naphthalen-1-yl)ethyl)benzamide (**5.10**, 100 mg, 0.307 mmol, 1 equiv.) was added. The reaction then stirred at room temperature overnight. The precipitated urea was filtered and the filtrate was evaporated under vacuum. The residue was dissolved in DCM and water (15 mL) was extracted (20 mL x 3). The organic layers were combined, washed with water, dried (Na_2SO_4), filtered, and evaporated under vacuum. The crude was purified by flash chromatography on silica

gel using 20-90% ethyl acetate in hexane to afford a white solid, 15 mg, 10% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.16 (d, $J = 8.4$ Hz, 1H), 7.90 (dd, $J = 8.0, 1.6$ Hz, 1H), 7.88 – 7.82 (m, 2H), 7.64 – 7.47 (m, 5H), 7.39 (dd, $J = 8.7, 2.7$ Hz, 1H), 7.01 (d, $J = 8.7$ Hz, 1H), 6.68 (d, $J = 8.4$ Hz, 1H), 6.18 – 6.08 (m, 1H), 2.28 – 2.05 (m, 4H), 1.91 – 1.82 (m, 1H), 1.79 (d, $J = 6.8$ Hz, 3H), 1.44 (s, 9H). ^{13}C NMR (126 MHz, CDCl_3) δ 171.03, 170.38, 162.88, 146.12, 137.60, 133.90, 131.92, 131.68, 131.15, 130.32, 129.32, 128.80, 128.64, 126.92, 126.11, 125.28, 124.60, 123.46, 122.95, 81.06, 77.30, 77.05, 76.79, 45.01, 29.64, 28.52, 28.07, 20.10. HRMS (ESI+) for $\text{C}_{27}\text{H}_{28}\text{ClNNaO}_5$ ($\text{M} + \text{Na}$), calcd: 504.1548, found: 504.1558.

(R)-5-chloro-2-(3-hydroxypropoxy)-N-(1-(naphthalen-1-yl)ethyl)benzamide (5.23). **5.10** (100 mg, 0.307 mmol, 1 equiv.) and K_2CO_3 (84.8 mg, 0.614 mmol, 2 equiv.) were dissolved in DMF (10 mL) and sonicated for 5 minutes. 3-bromopropanol (51.2 mg, 0.369 mmol, 1.2 equiv) was added and the resulting solution mixture was stirred at room temperature overnight. When the reaction appeared to be complete by TLC water was added, and the mixture was extracted with EtOAc (3x 30 mL). The organic layers were combined, washed with water, dried (Na_2SO_4), filtered and evaporated under. The crude was purified by flash chromatography on silica gel using 20-90% ethyl acetate in hexane to afford a white solid, **5.23**, 37 mg, 31% yield. ^1H NMR (500 MHz, CDCl_3) δ 8.24 (d, $J = 7.6$ Hz, 1H), 8.17 (dd, $J = 10.3, 5.1$ Hz, 2H), 7.87 (d, $J = 8.1$ Hz, 1H), 7.82 (t, $J = 9.0$ Hz, 1H), 7.59 (d, $J = 7.3$ Hz, 1H), 7.56 – 7.44 (m, 4H), 7.35 – 7.27 (m, 1H), 6.79 (t, $J = 10.3$ Hz, 1H), 6.14 – 6.06 (m, 1H), 4.08 – 4.00 (m, 1H), 4.00 – 3.88 (m, 1H), 3.26 (dd, $J = 10.4, 5.3$ Hz, 2H), 1.92 (dd, $J = 12.0, 5.9$ Hz, 1H), 1.85 – 1.77 (m, 3H), 1.76 (s, 1H), 1.56 (tt, $J = 11.3, 4.6$ Hz, 2H). ^{13}C NMR (126 MHz, CDCl_3) δ 163.11, 155.40, 138.52, 133.94, 132.26, 131.78, 131.16, 128.75, 128.34, 126.63, 126.52, 125.95, 125.27, 123.62, 122.94, 122.67, 113.68, 66.44, 58.78, 45.22, 31.39, 20.51. HRMS (ESI+) for $\text{C}_{22}\text{H}_{22}\text{ClNNaO}_3$ ($\text{M} + \text{Na}$), calcd: 406.1180, found:

406.1197.

(R)-5-chloro-2-(4-hydroxybutoxy)-N-(1-(naphthalen-1-yl)ethyl)benzamide Same (5.24). The same general procedure as for **5.23** was used: **5.10** (100 mg, 0.307 mmol, 1 equiv.), K_2CO_3 (84.8 mg, 0.614 mmol, 2 equiv.) and 3-bromo-butanol (51.5 mg, 0.369 mmol, 1.2 equiv.). The crude was purified by flash chromatography on silica gel using 20-90% ethyl acetate in hexane to afford the product **5.24**, as a white solid, 36 mg, 28% yield. 1H NMR (500 MHz, $CDCl_3$) δ 8.28 – 8.20 (m, 2H), 8.18 (d, $J = 8.3$ Hz, 1H), 7.88 (dt, $J = 7.7, 1.8$ Hz, 1H), 7.85 – 7.78 (m, 1H), 7.61 (dq, $J = 7.2, 1.1$ Hz, 1H), 7.54 (ddd, $J = 8.4, 6.1, 1.5$ Hz, 1H), 7.52 – 7.44 (m, 2H), 7.32 (dt, $J = 8.7, 2.6$ Hz, 1H), 6.79 (dd, $J = 8.8, 2.3$ Hz, 1H), 6.11 (p, $J = 6.9$ Hz, 1H), 4.13 (q, $J = 7.1$ Hz, 2H), 3.97 (dtd, $J = 8.7, 6.3, 2.2$ Hz, 1H), 3.85 (dtd, $J = 8.5, 6.4, 1.8$ Hz, 1H), 3.35 (q, $J = 5.7$ Hz, 2H), 1.79 (dd, $J = 6.7, 1.3$ Hz, 3H), 1.54 – 1.38 (m, 3H), 1.31 – 1.12 (m, 5H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.20, 155.46, 138.47, 133.95, 132.27, 131.95, 131.26, 128.71, 128.35, 126.63, 126.61, 125.93, 125.23, 123.66, 122.85, 122.71, 113.69, 69.26, 61.86, 45.19, 28.82, 25.31, 21.06.

(R)-2-iodo-N-(1-(naphthalen-1-yl)ethyl)benzamide (5.31). To a suspension of 2-iodobenzoic acid (200 mg, 0.81 mmol) in acetonitrile (8 mL, 0.1M) under Argon at 0°C was added DIPEA (0.14 mL, 0.81 mmol, 1.2 equiv.), and HATU (306.5 mg, 0.81 mmol, 1.2 equiv.). Mixture stirred at 0°C for 30 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.11 mL, 0.67 mmol, 1 equiv.) was added. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Solvent was removed under reduced pressure, ethyl acetate was added, and precipitate was filtered and washed with ethyl acetate (precipitate is clean product, beige solid). Filtrate was then washed with aq. saturated NH_4Cl (x2), aq. saturated $NaHCO_3$ (x2), and brine (x1), dried (Na_2SO_4), filtered, and concentrated to give an orange solid. Washed crude was recrystallized from isopropanol to afford a white solid, combined product, 107 mg, 40% yield. 1H NMR (500

MHz, CDCl₃) δ 8.27 (d, $J = 8.5$ Hz, 1H), 7.91 – 7.85 (m, 1H), 7.82 (dd, $J = 8.2, 1.7$ Hz, 2H), 7.63 – 7.56 (m, 2H), 7.52 (ddd, $J = 8.1, 6.9, 1.2$ Hz, 1H), 7.47 (dd, $J = 8.2, 7.2$ Hz, 1H), 7.35 – 7.27 (m, 2H), 7.05 (ddd, $J = 7.9, 6.5, 2.6$ Hz, 1H), 6.19 – 6.10 (m, 1H), 5.98 (d, $J = 8.4$ Hz, 1H), 1.84 (d, $J = 6.7$ Hz, 3H).

(R)-2-iodo-N-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzamide (5.32). To a suspension of 2-iodo-5-nitrobenzoic acid (200 mg, 0.68 mmol) in acetonitrile (8 mL, 0.1M) under Argon at 0°C was added DIPEA (0.12 mL, 0.68 mmol, 1.2 equiv.), and HATU (259.5 mg, 0.68 mmol, 1.2 equiv.). Mixture stirred at 0°C for 30 minutes. (R)-(+)-1-(1-naphthyl)ethylamine (0.09 mL, 0.57 mmol, 1 equiv.) was added. Reaction was allowed back to room temperature and stirred at room temperature for 24 hours. Solvent was removed under reduced pressure, ethyl acetate was added, and precipitate was filtered and washed with ethyl acetate (precipitate is clean product, beige solid). Filtrate was then washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x2), and brine (x1), dried (Na₂SO₄), filtered, and concentrated to give an orange solid. Washed crude was recrystallized from isopropanol to afford a white solid, combined product, 199 mg, 78% yield. ¹H NMR (500 MHz, DMSO) δ 9.25 (d, $J = 7.8$ Hz, 1H), 8.22 (dd, $J = 20.1, 8.5$ Hz, 2H), 8.02 – 7.93 (m, 3H), 7.87 (d, $J = 8.1$ Hz, 1H), 7.70 – 7.60 (m, 2H), 7.59 – 7.49 (m, 2H), 5.90 (p, $J = 7.1$ Hz, 1H), 1.64 (d, $J = 6.9$ Hz, 3H).

(R)-3-((2-((1-(naphthalen-1-yl)ethyl)carbonyl)-4-nitrophenyl)amino)propanoic acid (5.33). (R)-2-iodo-N-(1-(naphthalen-1-yl)ethyl)benzamide (**5.31**, 100 mg, 0.25 mmol, 1 equiv.), β -alanine (24.4 mg, 0.27 mmol, 1.1 equiv.), Cs₂CO₃ (stored in the oven, 203 mg, 0.62 mmol, 2.5 equiv.), CuI (23.7 mg, 0.12 mmol, 0.5 equiv.) were added to a flask. The flask was capped and purged with vacuum and argon (x3). DMF was added (2.5 mL, 0.16M) and solution was degassed by bubbling argon into it while stirring. Flask was kept under argon atmosphere and heated to 105°C. Reaction

was monitored by TLC and was completed after 3 hours. Reaction mixture was then cooled to room temperature and filtered through celite using ethyl acetate, filtrate is a pale green solution. Water was added to the filtrate and pH was adjusted to 3 using 1M HCl. Phases were separated and aqueous phase was extracted with ethyl acetate (x5). Combined organic layers were washed with acidic water (pH 3) (x3), brine (x2), dried (Na₂SO₄), filtered and concentrated to give an orange oil. Crude was co-evaporated with heptane to remove as much DMF as possible. Crude was purified by flash column chromatography on silica gel using 0-5% methanol in DCM as eluent. Solvent was removed from combined fractions to give a translucent oil, 52 mg, 58% yield. ¹H NMR (500 MHz, CDCl₃) δ 8.13 (d, J = 8.4 Hz, 1H), 7.86 (dd, J = 7.7, 1.4 Hz, 1H), 7.80 (d, J = 8.1 Hz, 1H), 7.59 – 7.42 (m, 4H), 7.29 – 7.27 (m, 1H), 7.24 (d, J = 1.6 Hz, 1H), 6.70 (dd, J = 8.5, 1.1 Hz, 1H), 6.54 (ddd, J = 8.1, 7.2, 1.1 Hz, 1H), 6.34 (d, J = 7.8 Hz, 1H), 6.06 – 5.91 (m, 1H), 3.49 (t, J = 6.9 Hz, 2H), 2.67 (t, J = 6.9 Hz, 2H), 1.74 (d, J = 6.8 Hz, 3H).

(R)-3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-nitrophenyl)amino)propanoic acid (5.34).

(R)-2-iodo-N-(1-(naphthalen-1-yl)ethyl)-5-nitrobenzamide (**5.32**, 190 mg, 0.42 mmol, 1 equiv.), βalanine (41.7 mg, 0.47 mmol, 1.1 equiv.), Cs₂CO₃ (stored in the oven, 346.8 mg, 1.06 mmol, 2.5 equiv.), CuI (40.5 mg, 0.21 mmol, 0.5 equiv.) were added to a flask. The flask was capped and purged with vacuum and argon (x3). DMF was added (2.5 mL, 0.16M) and solution was degassed by bubbling argon into it while stirring. Flask was kept under argon atmosphere and heated to 105°C. Reaction was monitored by TLC and was completed after 3 hours. Reaction mixture was then cooled to room temperature and filtered through celite using ethyl acetate, filtrate is a pale green solution. Water was added to the filtrate and pH was adjusted to 3 using 1M HCl. Phases were separated and aqueous phase was extracted with ethyl acetate (x5). Combined organic layers were washed with acidic water (pH 3) (x3), brine (x2), dried (Na₂SO₄), filtered and concentrated

to give an orange oil. Crude was co-evaporated with heptane to remove as much DMF as possible. Crude was purified by flash column chromatography on silica gel using 0-5% methanol in DCM as eluent. Solvent was removed from combined fractions to give a pale orange solid, 170 mg, 52% yield. ¹H NMR (500 MHz, DMSO) δ 9.25 (d, J = 7.8 Hz, 1H), 8.22 (dd, J = 20.1, 8.5 Hz, 2H), 8.02 – 7.93 (m, 3H), 7.87 (d, J = 8.1 Hz, 1H), 7.70 – 7.60 (m, 2H), 7.59 – 7.49 (m, 2H), 5.90 (p, J = 7.1 Hz, 1H), 1.64 (d, J = 6.9 Hz, 3H).

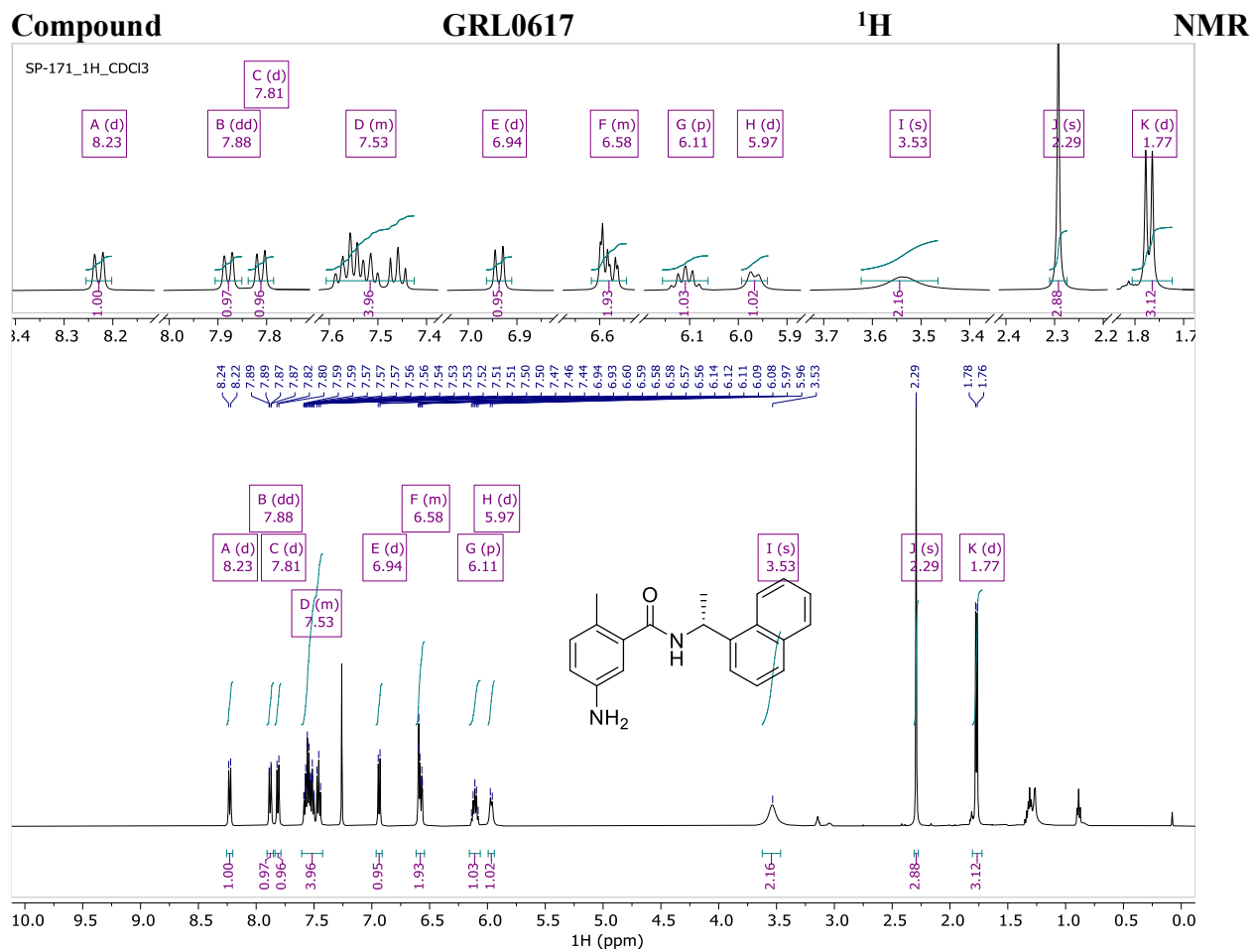
Ethyl (R,E)-4-(3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)phenyl)amino)propanamido)but-2-enoate (5.35). (R)-3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-nitrophenyl)amino)propanoic acid (**5.33**, 200 mg, 0.55 mmol, 1.2 equiv.) in DCM (3.5 mL) under Argon at 0°C were added DIPEA (0.09 mL, 0.55 mmol, 1.2 equiv.) and PyBOP (286 mg, 0.55 mmol, 1.2 equiv.). In a separate flask, ethyl (E)-4-aminobut-2-enoate trifluoroacetic acid salt (**5.30**, 104.9 mg, 0.46 mmol, 1 equiv.) was dissolved in DCM (0.6 mL) and DIPEA (0.08 mL, 0.46 mmol, 1 equiv.) was added to neutralize the salt. Neutralized amine mixture was added to reaction mixture (DCM final amount 6 mL, 0.17M, DIPEA 0.17 mL, 1.01 mmol, 2.2 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 18 hours. Solvent was removed under reduced pressure and ethyl acetate and an aq. Saturated solution of NH₄Cl were added. Phases were separated and the organic phase was washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x1), and brine (x1). Combined organic layers were dried (Na₂SO₄), filtered, and concentrated to give an orange solid. Crude was purified by flash column chromatography on silica gel using 5-60% ethyl acetate in DCM as an eluent. The product was co-evaporated with chloroform to afford a white solid, 213. mg, 52% yield. ¹H NMR (500 MHz, CDCl₃) δ 8.17 (d, J = 8.4 Hz, 1H), 7.91 – 7.85 (m, 1H), 7.81 (d, J = 8.2 Hz, 1H), 7.59 – 7.42 (m, 4H), 6.77 – 6.72 (m, 1H), 6.71 (t, J = 5.2 Hz, 0H), 6.56 (td, J = 7.5, 1.1

Hz, 1H), 6.48 (d, J = 7.8 Hz, 1H), 6.04 (p, J = 7.0 Hz, 1H), 5.92 (s, 1H), 5.75 (dt, J = 15.8, 1.9 Hz, 1H), 4.12 (q, J = 3.6 Hz, 2H), 4.02 – 3.84 (m, 2H), 3.53 (td, J = 6.6, 3.0 Hz, 2H), 2.50 (td, J = 6.6, 4.3 Hz, 2H), 2.04 (s, 0H), 1.74 (d, J = 6.8 Hz, 3H), 1.24 (t, J = 7.1 Hz, 4H). ¹³C NMR (126 MHz, CDCl₃) δ 171.26, 168.76, 166.01, 148.96, 143.70, 138.43, 134.00, 132.91, 131.10, 128.90, 128.39, 127.54, 126.58, 125.90, 125.32, 123.35, 122.56, 121.67, 116.02, 115.54, 111.99, 60.46, 44.94, 40.10, 39.49, 36.67, 20.91, 14.21. HRMS (ESI+) for C₂₈H₃₁O₄N₃ calcd 496.2207 found 496.2204.

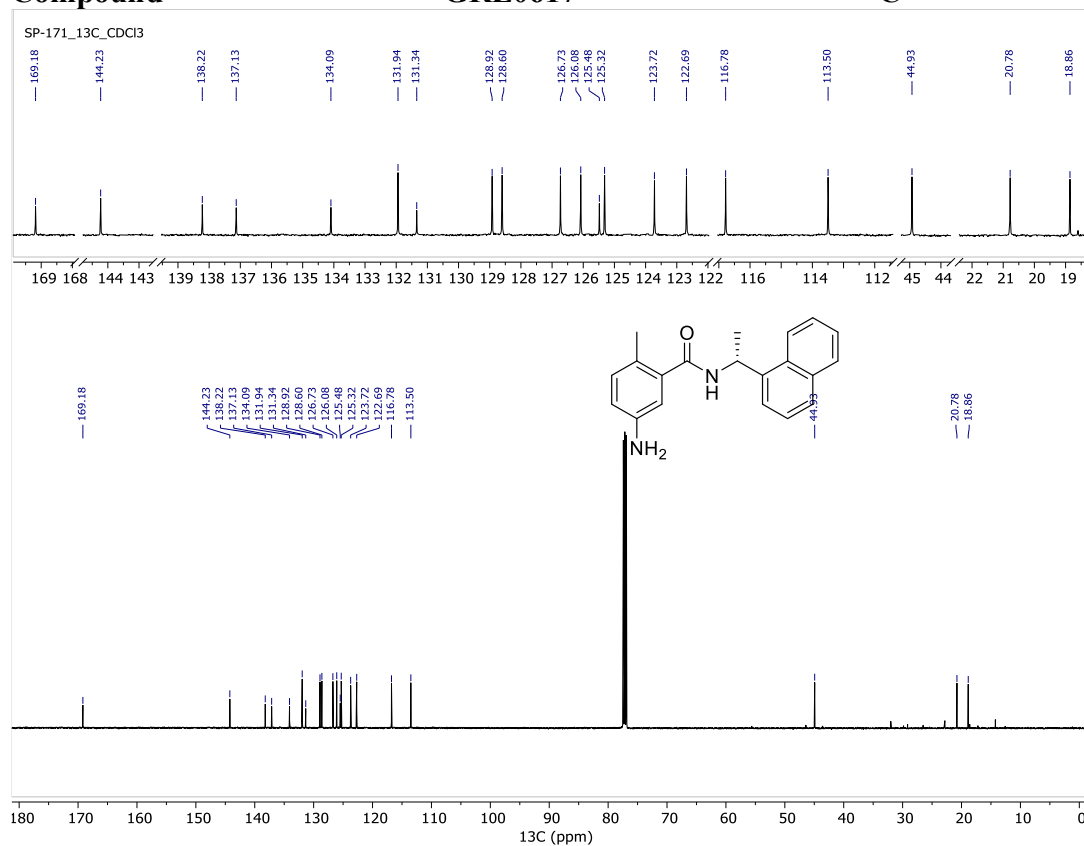
Ethyl (R,E)-4-(3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-nitrophenyl)amino)propanamido)but-2-enoate (5.36). (R)-3-((2-((1-(naphthalen-1-yl)ethyl)carbamoyl)-4-nitrophenyl)amino)propanoic (**5.34**, 140 mg, 0.34 mmol, 1.2 equiv.) in DCM (0.6 mL) under Argon at 0°C were added DIPEA (0.04 mL, 0.229 mmol, 1.2 equiv.) and PyBOP (177 mg, 0.34 mmol, 1.2 equiv.). In a separate flask, ethyl (E)-4-aminobut-2-enoate trifluoroacetic acid salt (**5.30**, 43 mg, 0.177 mmol, 1 equiv.) was dissolved in DCM (0.6 mL) and DIPEA (0.06 mL, 0.34 mmol, 1 equiv.) was added to neutralize the salt. Neutralized amine mixture was added to reaction mixture (DCM final amount 3.5 mL, 0.17M, DIPEA 0.11 mL, 0.62 mmol, 2.2 equiv.). Reaction was allowed back to room temperature and stirred at room temperature for 18 hours. Solvent was removed under reduced pressure and ethyl acetate and an aq. Saturated solution of NH₄Cl were added. Phases were separated and the organic phase was washed with aq. saturated NH₄Cl (x2), aq. saturated NaHCO₃ (x1), and brine (x1). Combined organic layers were dried (Na₂SO₄), filtered, and concentrated to give an orange solid. Crude was purified by flash column chromatography on silica gel using 5-60% ethyl acetate in DCM as an eluent. The product was co-evaporated with chloroform to afford a white solid, 148. mg, 58% yield. ¹H NMR (500 MHz, DMSO) δ 9.31 (d, J = 7.6 Hz, 1H), 8.97 (t, J = 5.7 Hz, 1H), 8.66 (d, J = 2.7 Hz, 1H), 8.26 (t, J = 5.7 Hz, 1H), 8.18 (d, J = 8.4 Hz, 1H), 8.13 (dd, J = 9.4, 2.6 Hz, 1H), 7.95 (dd, J = 8.0, 1.6

Supplementary information Chapter 5

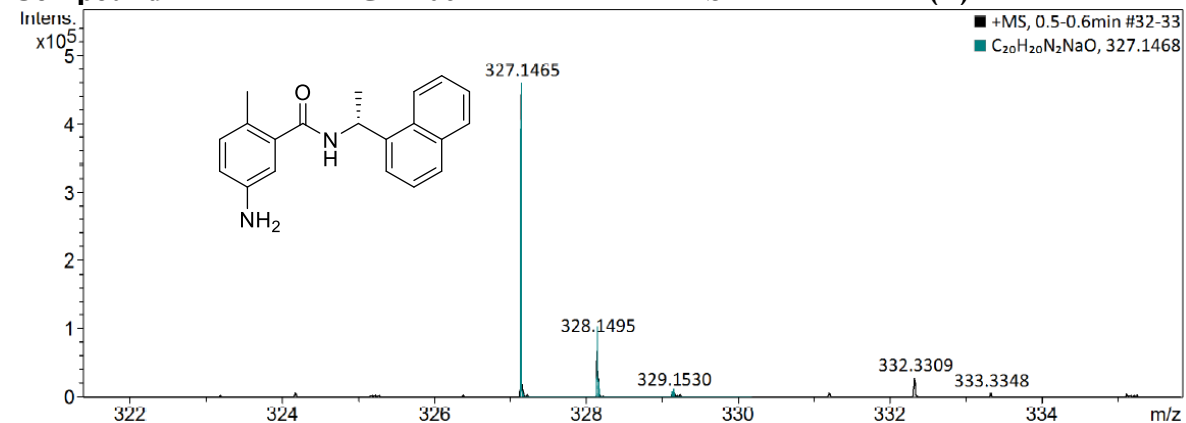
Hz, 1H), 7.84 (d, $J = 8.1$ Hz, 1H), 7.64 (dd, $J = 7.2, 1.2$ Hz, 1H), 7.60 – 7.49 (m, 3H), 6.87 (d, $J = 9.5$ Hz, 1H), 6.80 (dt, $J = 15.7, 4.7$ Hz, 1H), 5.93 (p, $J = 7.0$ Hz, 1H), 5.85 (dt, $J = 15.7, 1.9$ Hz, 1H), 4.07 (q, $J = 7.1$ Hz, 2H), 3.85 (dq, $J = 5.6, 1.9$ Hz, 2H), 3.56 – 3.43 (m, 2H), 2.47 (t, $J = 6.7$ Hz, 2H), 1.62 (d, $J = 7.0$ Hz, 3H), 1.16 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 170.15, 166.60, 165.43, 153.80, 145.74, 140.04, 134.60, 133.36, 130.39, 128.69, 128.15, 127.33, 126.22, 125.56 (d, $J = 11.0$ Hz), 123.01, 122.57, 120.31, 113.28, 110.85, 59.83, 44.66, 34.37, 21.20, 14.08. HRMS (ESI+) for $\text{C}_{28}\text{H}_{30}\text{O}_6\text{N}_4$ calcd 541.20576 found 541.20448.



Compound **GRL0617** ¹³C NMR



Compound **GRL0617** ESI (+) HRMS



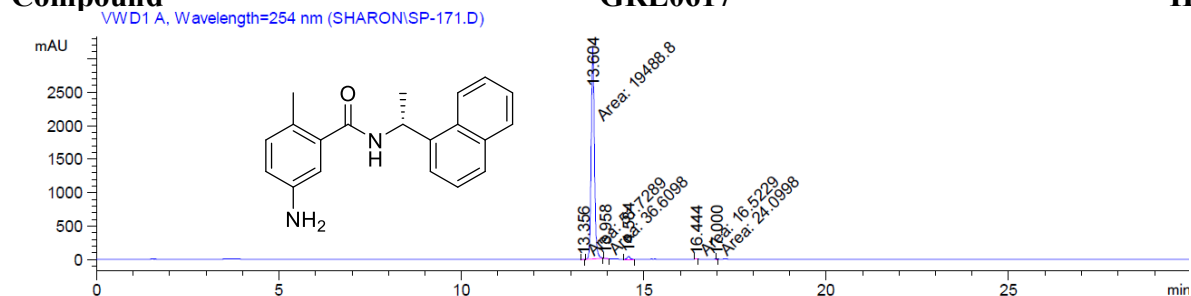
Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
327.1465	1	C ₂₀ H ₂₀ N ₂ NaO	327.1468	0.9	22.9	1	100.00	11.5	even	ok

Supplementary information Chapter 5

Compound

GRL0617

HPLC



Signal 2: VWD1 A, Wavelength=254 nm

Peak #	RetTime [min]	Type	Width [min]	Area mAU	Area *s	Height [mAU]	Area %
1	13.356	MM	0.1001	51.72893	7.51741	0.2607	
2	13.604	MM	0.1026	1.94888e4	3166.73950	98.2016	
3	13.958	MM	0.0886	36.60981	6.88862	0.1845	
4	14.584	BV	0.0846	227.94283	41.40695	1.1486	
5	16.444	MM	0.0656	16.52287	4.20057	0.0833	

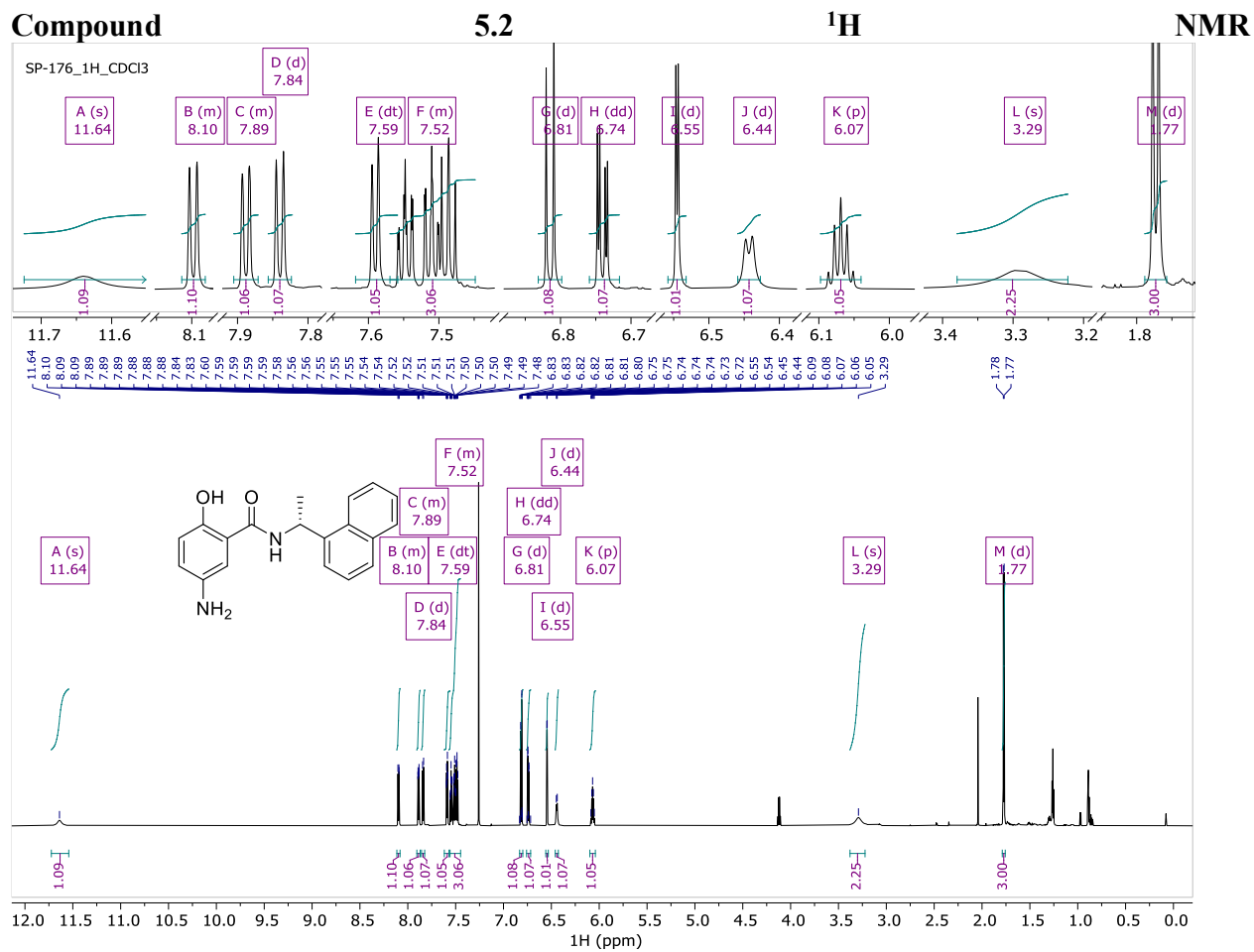
Instrument 1 2/18/2022 1:55:27 PM Anne

File C:\CHEM32\1\DATA\SHARON\SP-171.D

Sample Name: SP-171

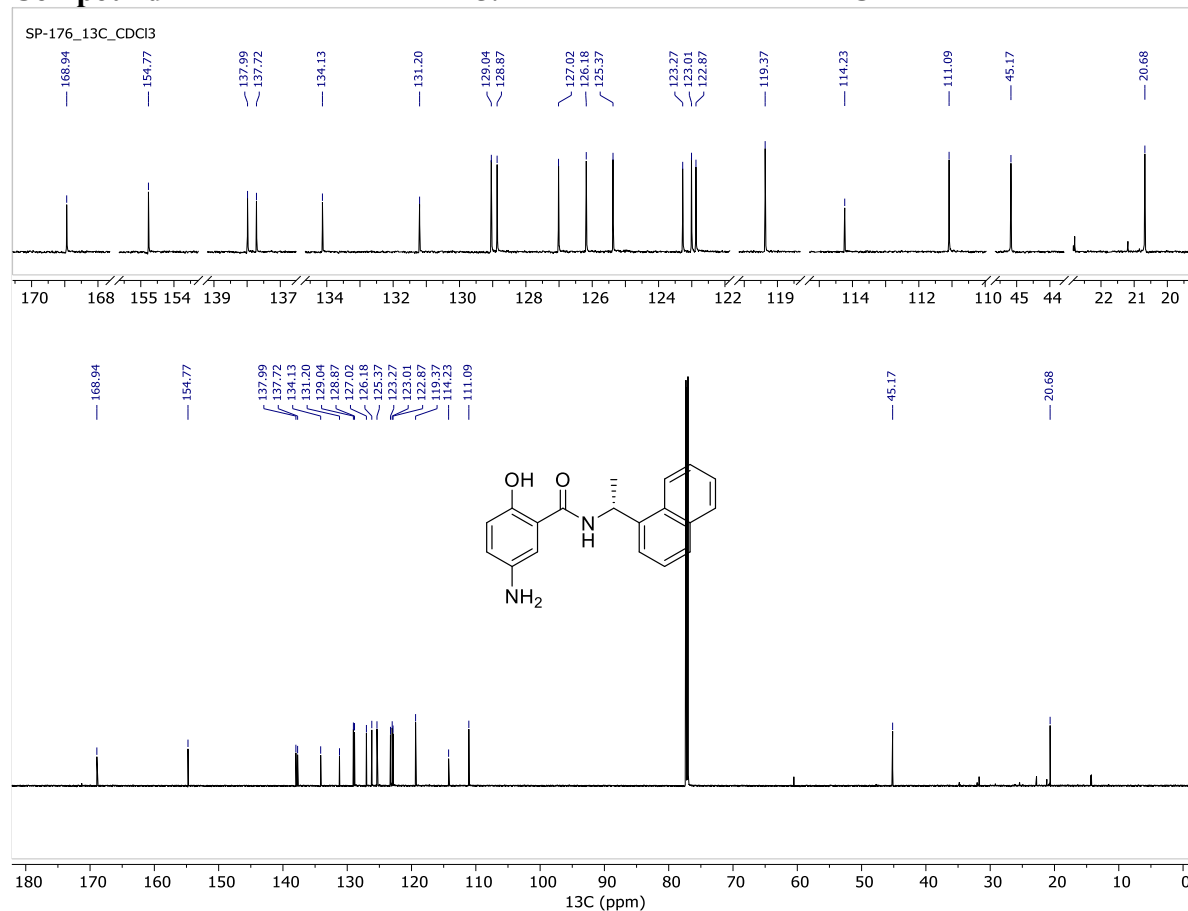
Peak #	RetTime [min]	Type	Width [min]	Area mAU	Area *s	Height [mAU]	Area %
6	17.000	MM	0.0477	24.09983	8.41466	0.1214	

Totals : 1.98457e4 3235.16771

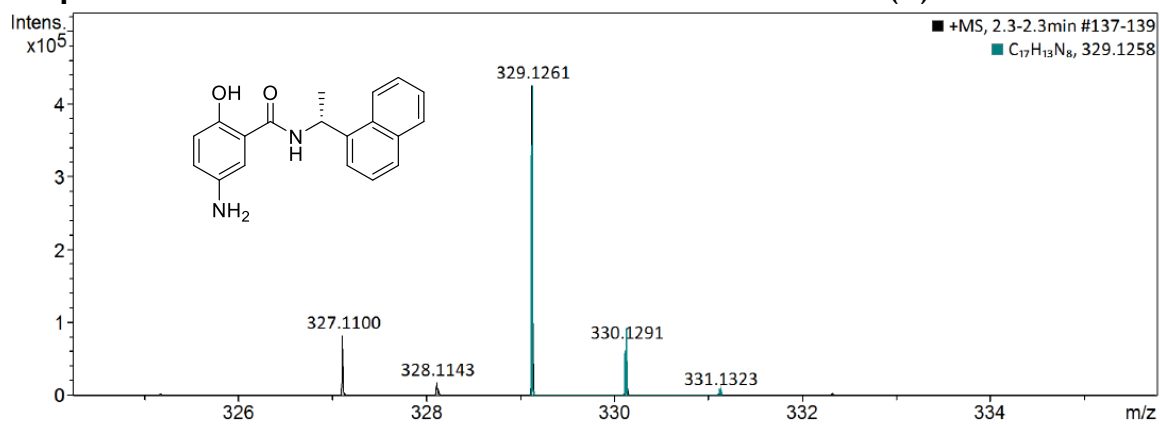


Supplementary information Chapter 5

Compound 5.2 ¹³C NMR



Compound 5.2 ESI (+) HRMS

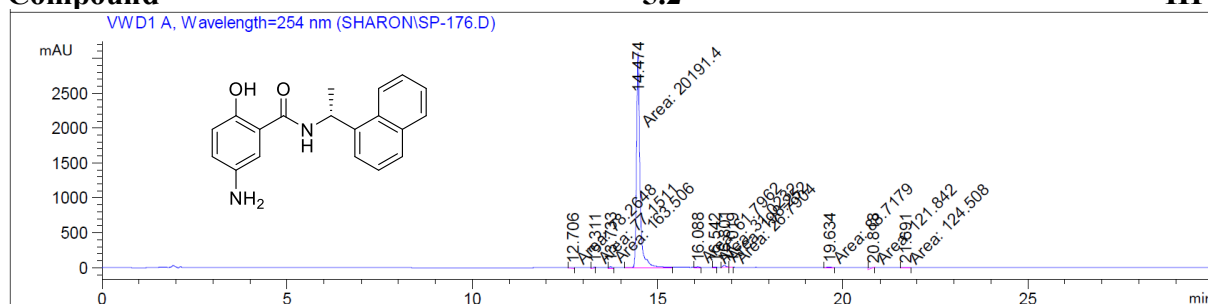


Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdB	e ⁻ Conf	N-Rule
329.1261	1	C ₁₇ H ₁₃ N ₈	329.1258	-1.1	31.0	1	100.00	15.5	even	ok
	1	C ₁₉ H ₁₈ N ₂ NaO ₂	329.1260	-0.2	32.1	1	100.00	11.5	even	ok

Compound

5.2

HPLC

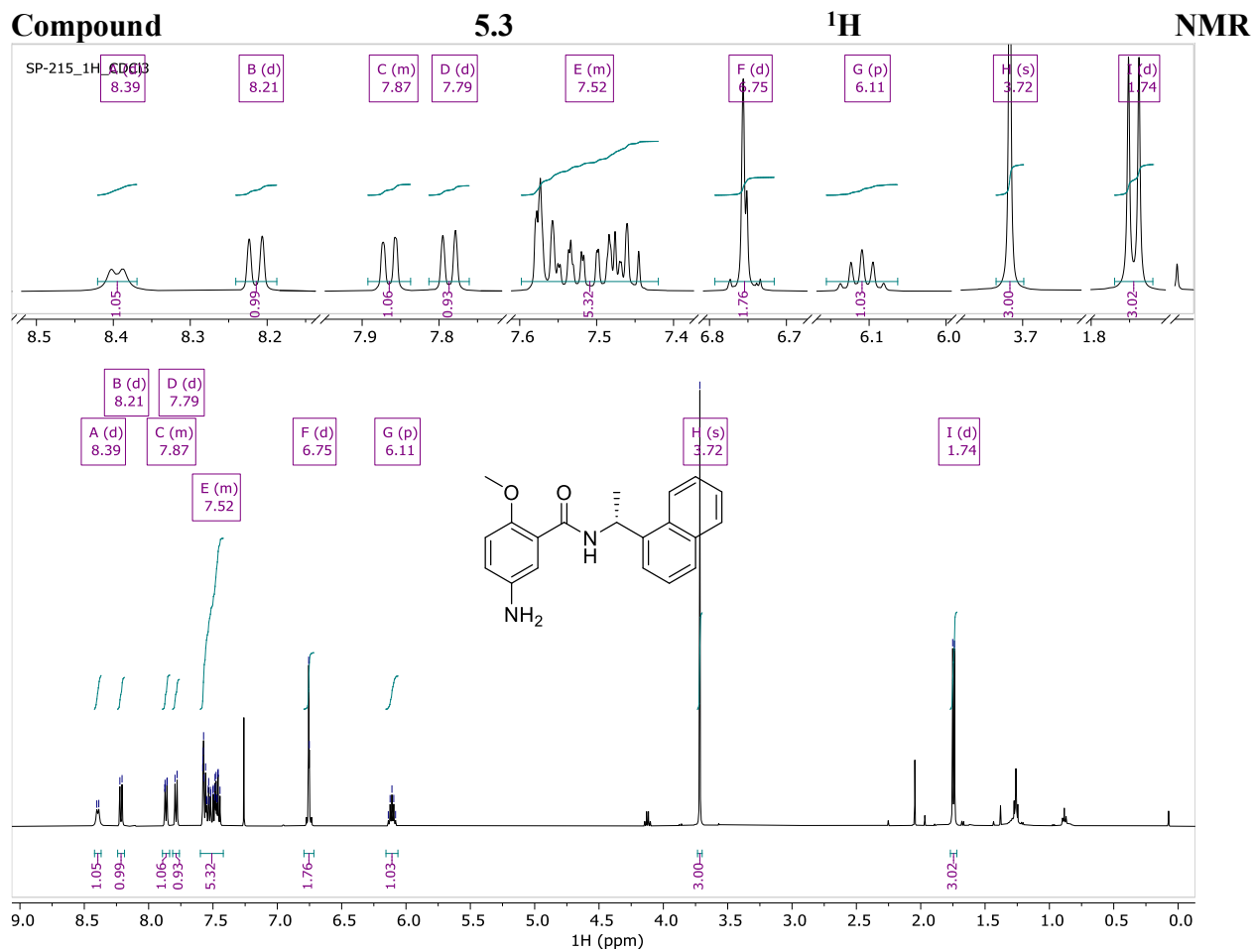


Peak #	RetTime [min]	Type	Width [min]	Area mAU	*s	Height [mAU]	Area %
1	12.706	MM	0.1368	78.26482		9.53485	0.3698
2	13.311	MM	0.1066	77.15111		12.06148	0.3646
3	13.733	MM	0.1145	163.50562		23.79793	0.7726
4	14.474	MM	0.1088	2.01914e4		3093.52954	95.4078
5	16.088	MM	0.0965	61.79617		10.67354	0.2920
6	16.542	MM	0.0875	31.02322		5.91063	0.1466
7	16.801	MM	0.1073	198.25177		30.80828	0.9368
8	17.019	MM	0.0916	26.79037		4.87599	0.1266
9	19.634	MM	0.1875	88.71795		7.88483	0.4192
10	20.818	MM	0.6437	121.84180		3.15493	0.5757
11	21.691	MM	0.2400	124.50841		8.64655	0.5883

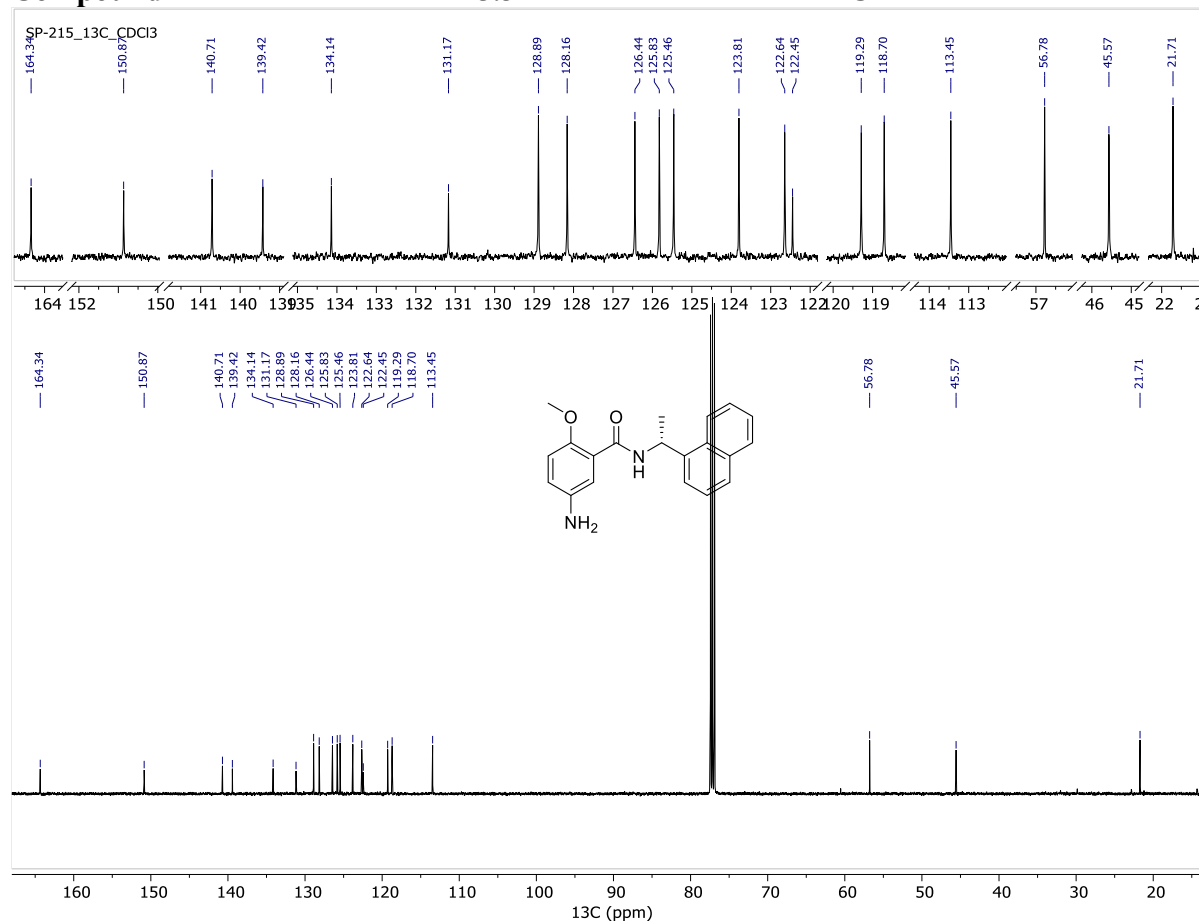
trument 1 2/18/2022 3:18:17 PM Anne

a File C:\CHEM32\1\DATA\SHARON\SP-176.D
 ple Name: SP-176

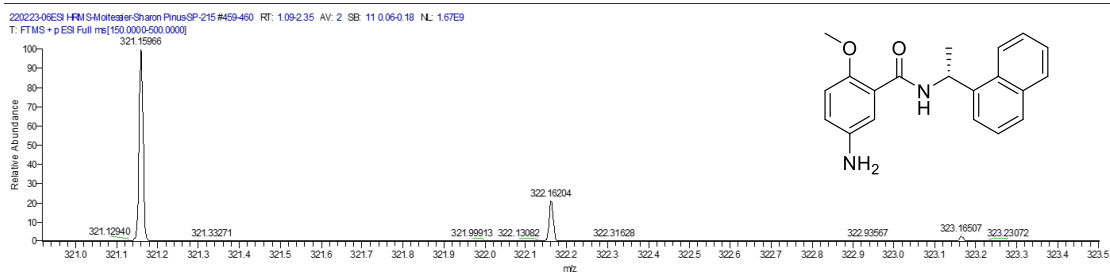
Peak #	RetTime [min]	Type	Width [min]	Area mAU	*s	Height [mAU]	Area %
Totals :				2.11633e4		3210.87855	



Compound **5.3** ¹³C NMR



Compound **5.3** ESI (+) HRMS



220223-06ESI HRMS-Moitessier-Sharon Pinus-SP-215#459-460 RT: 1.09-2.35 AV: 2

SB: 11 0.06-0.18

T: FTMS + p ESI Full ms [150.0000-500.0000]

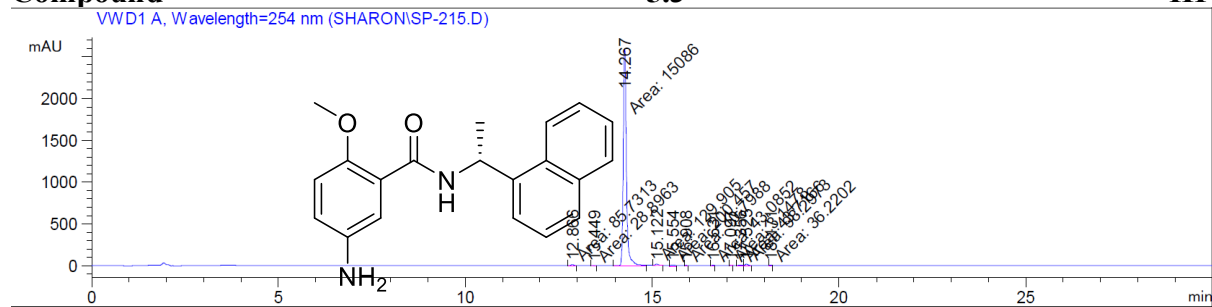
m/z = 321.15492-321.16871

m/z	Intensity	Relative Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
321.15966	1668079488.0	100.00	30597.33	1.00	321.15975	-0.30	11.5 C ₂₀ H ₂₁ O ₂ N ₂

Compound

5.3

HPLC



Signal 2: VWD1 A, Wavelength=254 nm

Peak #	RetTime [min]	Type	Width [min]	Area mAU	Area *s	Height [mAU]	Area %
1	12.866	MM	0.0985	85.73133	14.50936	0.5438	
2	13.449	MM	0.1223	28.89627	3.93894	0.1833	
3	14.267	MM	0.0965	1.50860e4	2604.19653	95.6854	
4	15.122	MM	0.1110	129.90492	19.50038	0.8239	
5	15.554	MM	0.1568	200.45659	21.30028	1.2714	

trument 1 2/18/2022 1:59:23 PM Anne

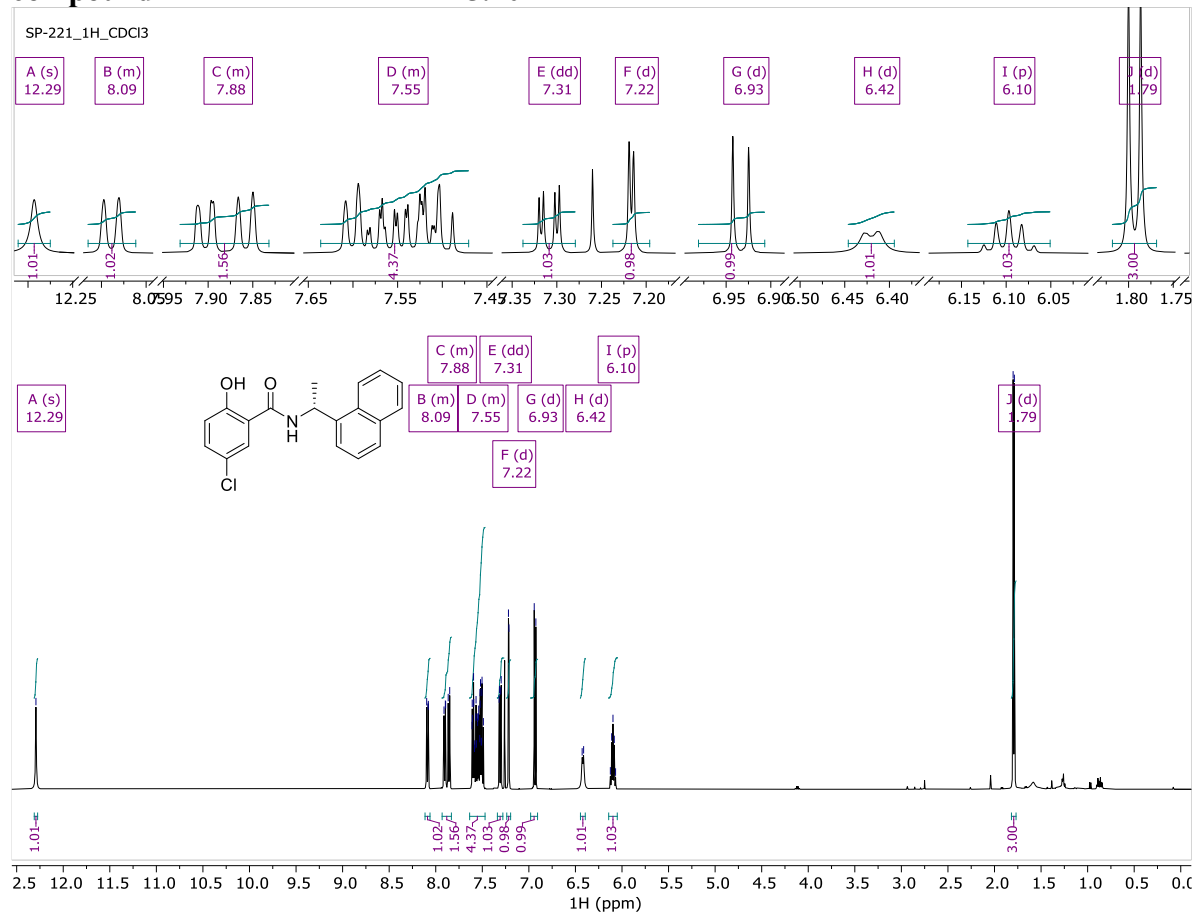
a File C:\CHEM32\1\DATA\SHARON\SP-215.D

ple Name: SP-215

Peak #	RetTime [min]	Type	Width [min]	Area mAU	Area *s	Height [mAU]	Area %
6	15.908	MM	0.0767	10.79879	2.34783	0.0685	
7	16.631	MM	0.0960	43.08517	7.48111	0.2733	
8	17.097	MM	0.0549	3.14780	1.03148	0.0200	
9	17.356	MM	0.0980	43.71657	7.43840	0.2773	
10	17.525	MM	0.1064	98.29726	15.39887	0.6235	
11	18.171	MM	0.0832	36.22017	7.25414	0.2297	

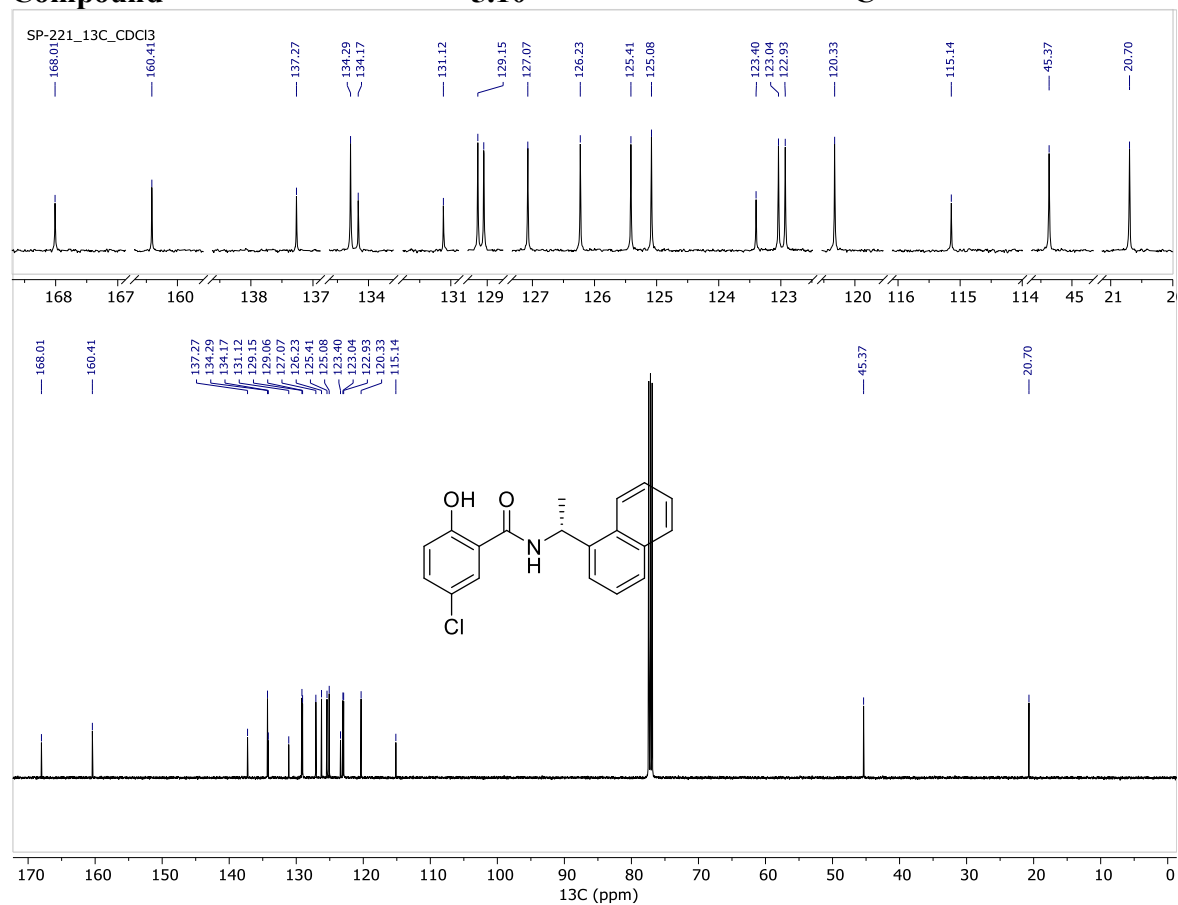
Totals : 1.57662e4 2704.39732

compound **5.10** ^1H NMR



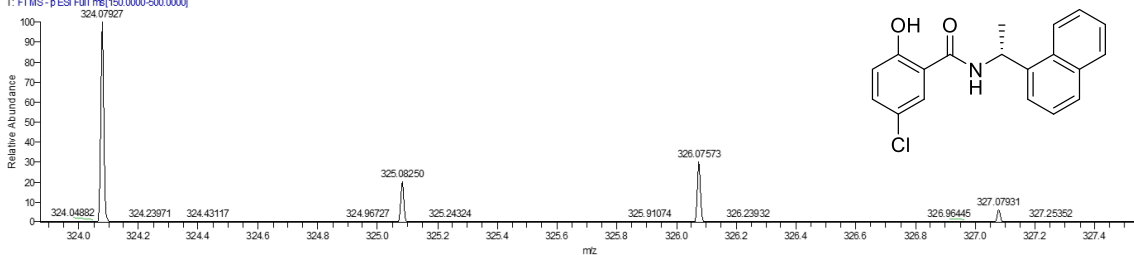
Supplementary information Chapter 5

Compound 5.10 ¹³C NMR



Compound 5.10 ESI (+) HRMS

220223-07ESI HRMS-Moitessier-Sharon Pinus-SP-221#183-191 RT: 0.43-0.45 AV: 9 NL: 3.25E8
T: FTMS - p ESI Full ms [150.0000-500.0000]



220223-07ESI HRMS-Moitessier-Sharon Pinus-SP-221#183-191 RT: 0.43-0.45 AV: 9

T: FTMS - p ESI Full ms [150.0000-500.0000]

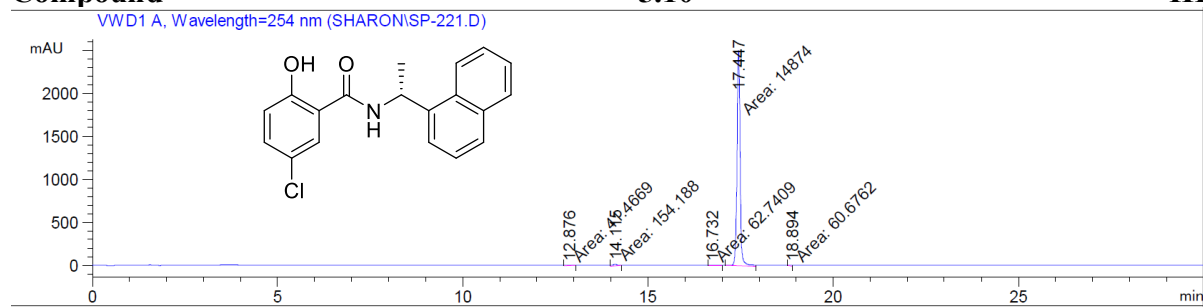
m/z = 324.04935-324.09706

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
324.07927	324588832.0	100.00	30275.50	1.00	324.07968	-1.27	12.5	C ₁₉ H ₁₅ O ₂ NCl

Compound

5.10

HPLC

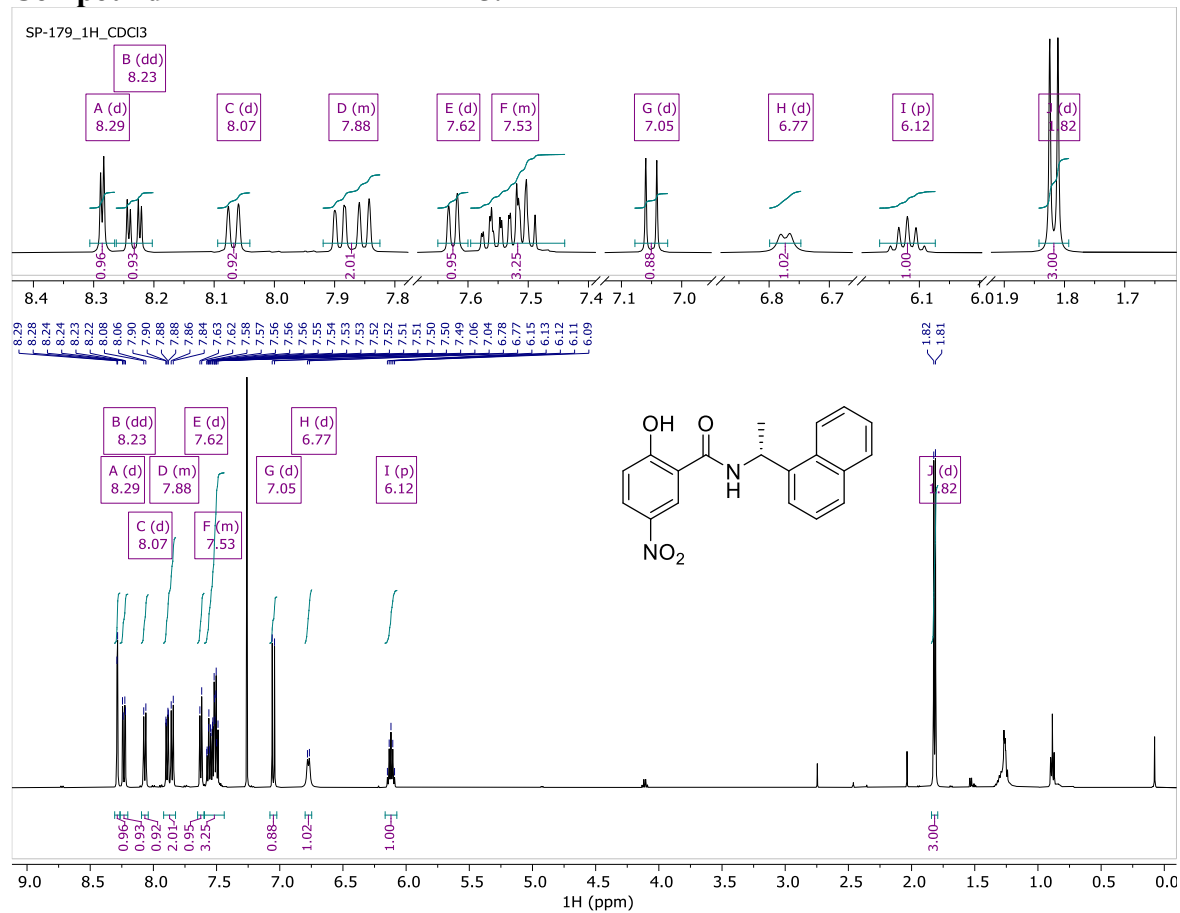


Signal 2: VWD1 A, Wavelength=254 nm

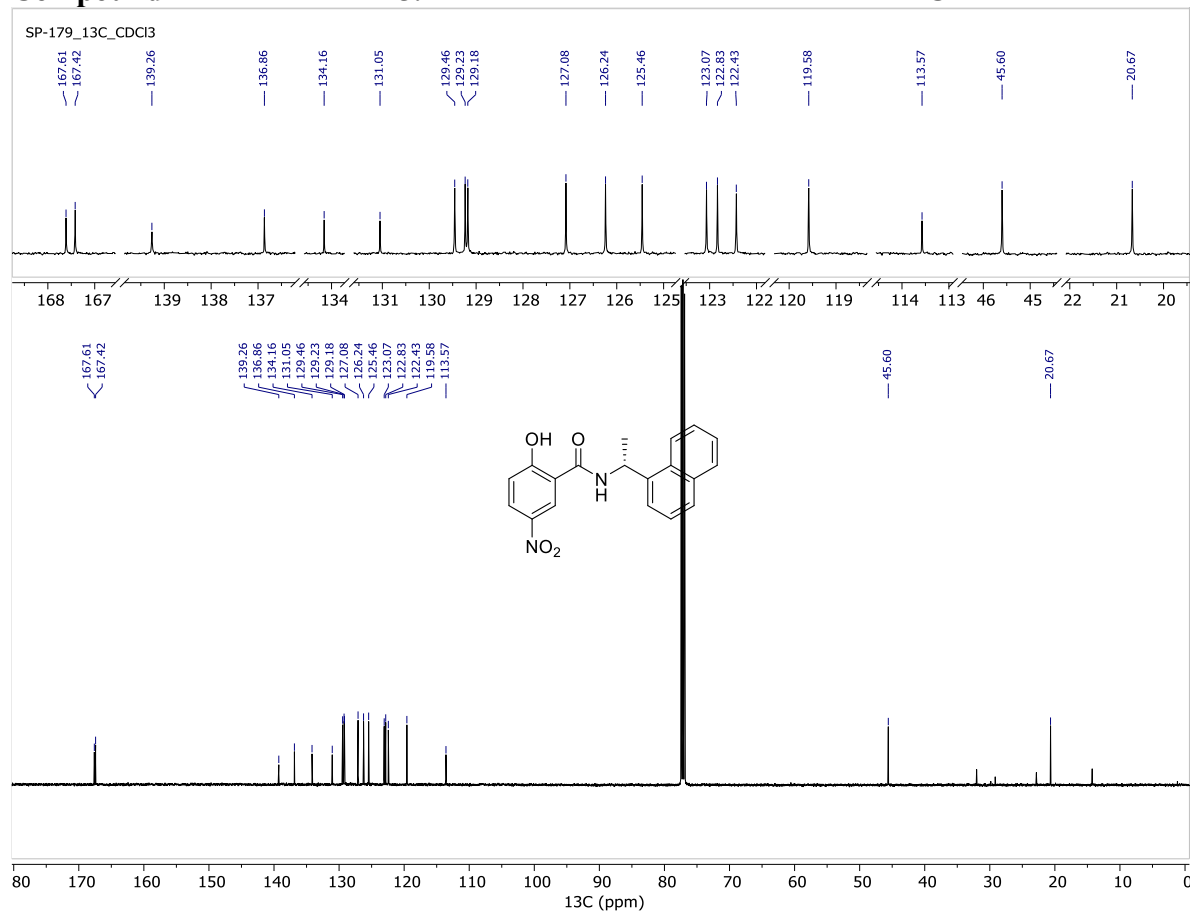
Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
1	12.876	MM	0.2562	41.46692	2.47925	0.2729
2	14.115	MM	0.1333	154.18800	19.27642	1.0149
3	16.732	MM	0.2773	62.74094	3.77148	0.4130
4	17.447	MM	0.0988	1.48740e4	2509.90356	97.8999
5	18.894	MM	0.0948	60.67622	10.66700	0.3994

Totals : 1.51931e4 2546.09772

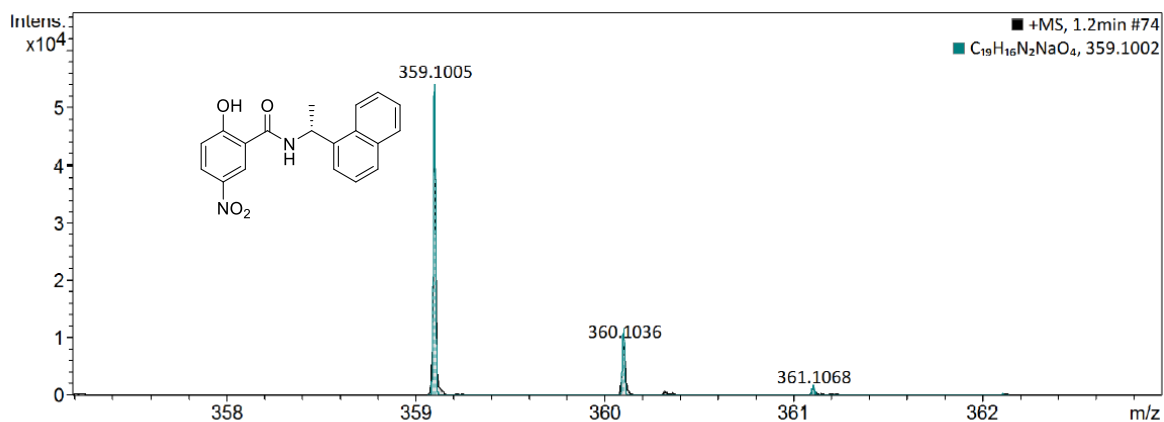
Compound 5.4 ¹H NMR



Compound **5.4** ¹³C NMR

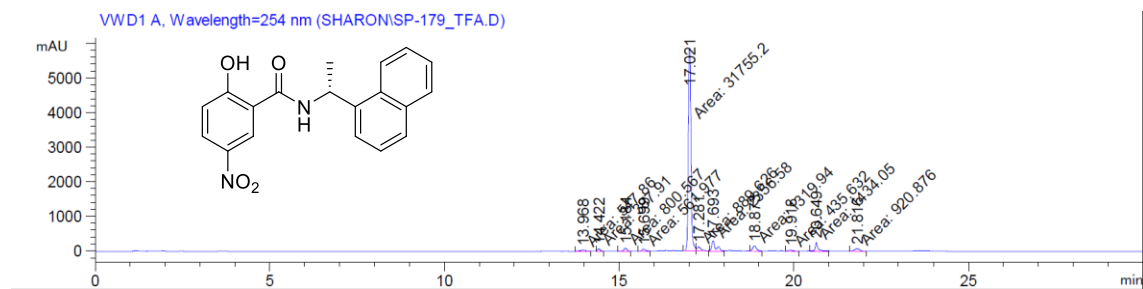


Compound **5.4** ESI (+) HRMS



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
359.1005	1	C ₁₇ H ₁₁ N ₂ O ₂	359.0999	-1.7	30.2	1	100.00	16.5	even	ok
	1	C ₁₉ H ₁₆ N ₂ NaO ₄	359.1002	-0.9	31.2	1	100.00	12.5	even	ok

Compound 5.4 HPLC



Signal 2: VWD1 A, Wavelength=254 nm

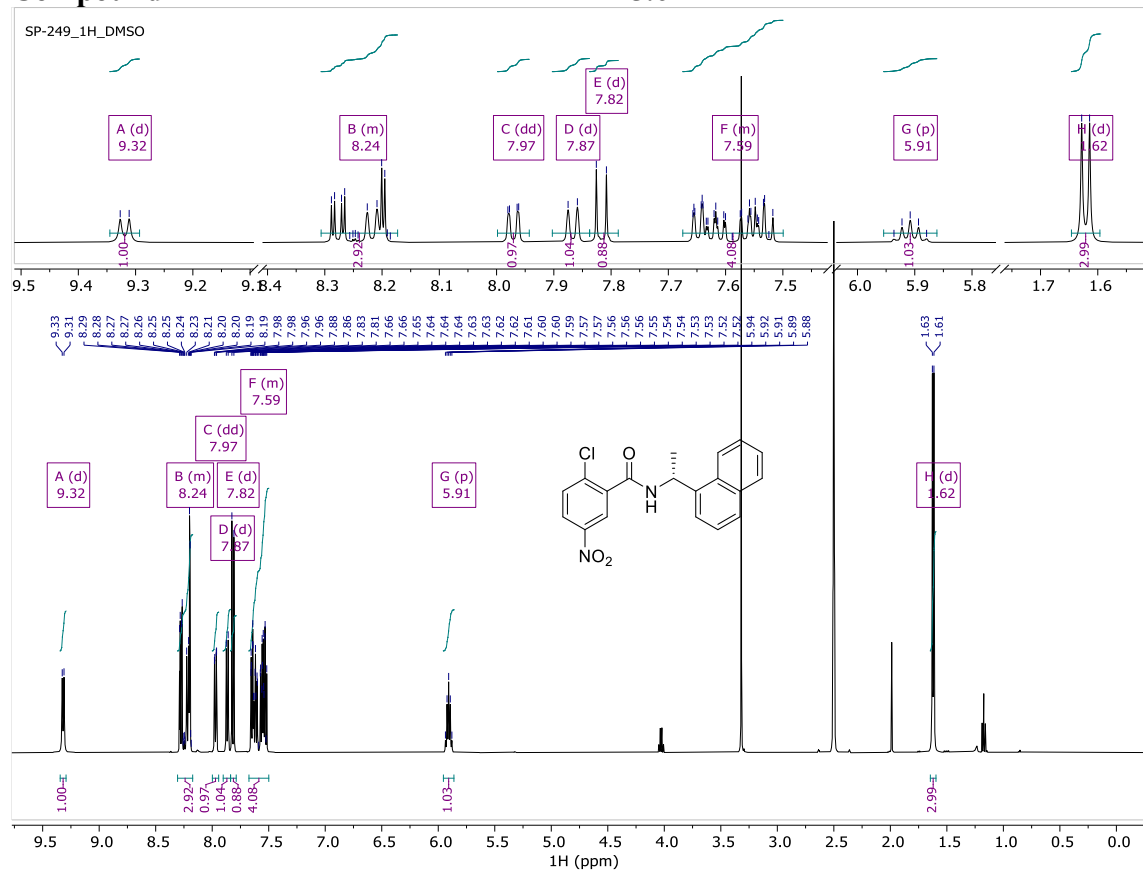
Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
1	13.968	MM	0.2057	547.86023	44.38540	1.3230
2	14.422	MM	0.0835	387.90994	77.39862	0.9368
3	15.184	MM	0.1332	800.56677	100.20741	1.9333
4	15.699	MM	0.1421	561.97729	65.90402	1.3571
5	17.021	MF	0.0902	3.17552e4	5865.05518	76.6844
6	17.281	FM	0.1335	889.62622	111.08904	2.1483
7	17.693	MM	0.1369	2356.57666	286.90073	5.6908

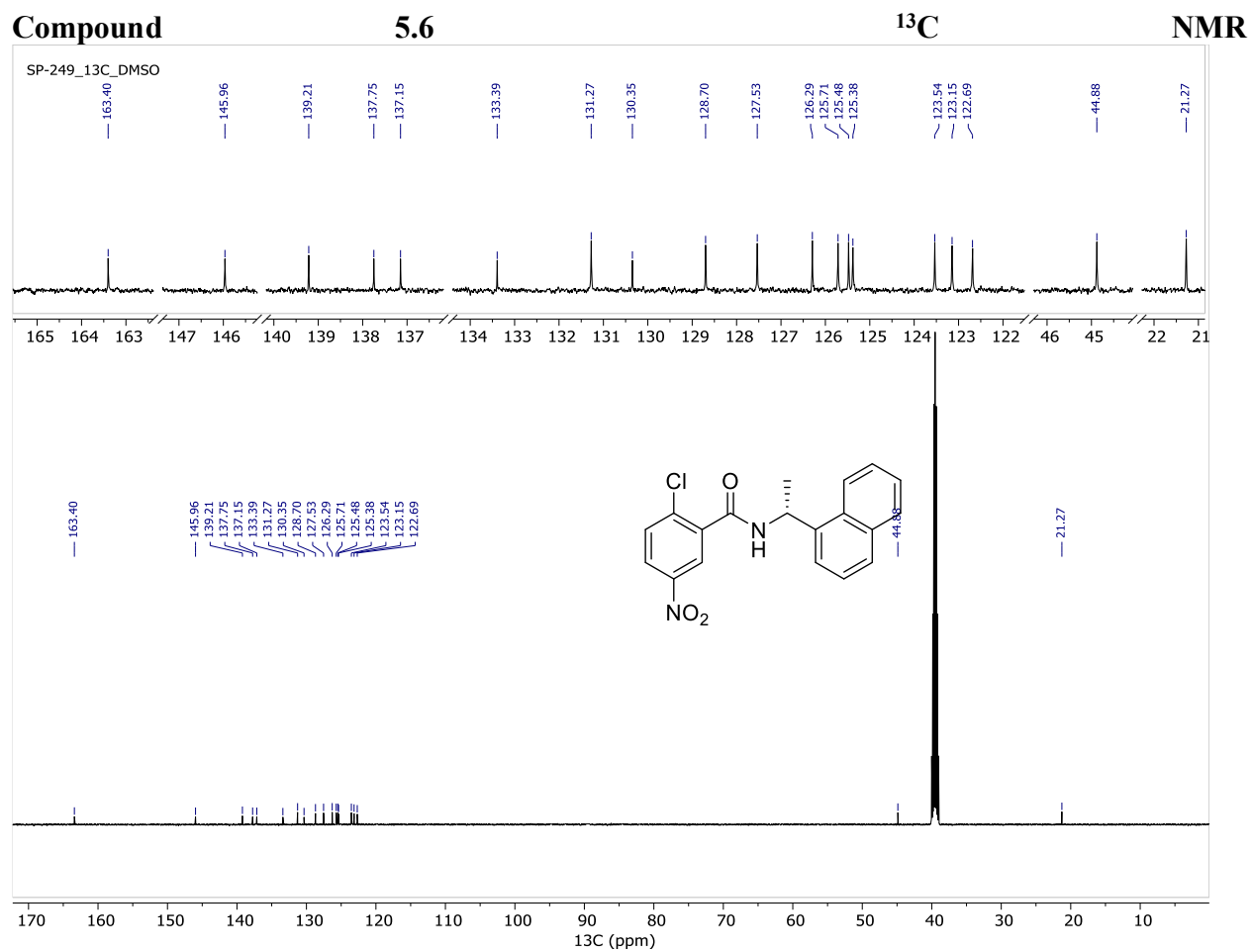
Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
8	18.873	MM	0.1367	1319.94421	160.96675	3.1875
9	19.918	MM	0.2063	435.63248	35.18786	1.0520
10	20.649	MM	0.0961	1434.05200	248.67609	3.4630
11	21.816	MM	0.2012	920.87646	76.28246	2.2238

Totals : 4.14102e4 7072.05356

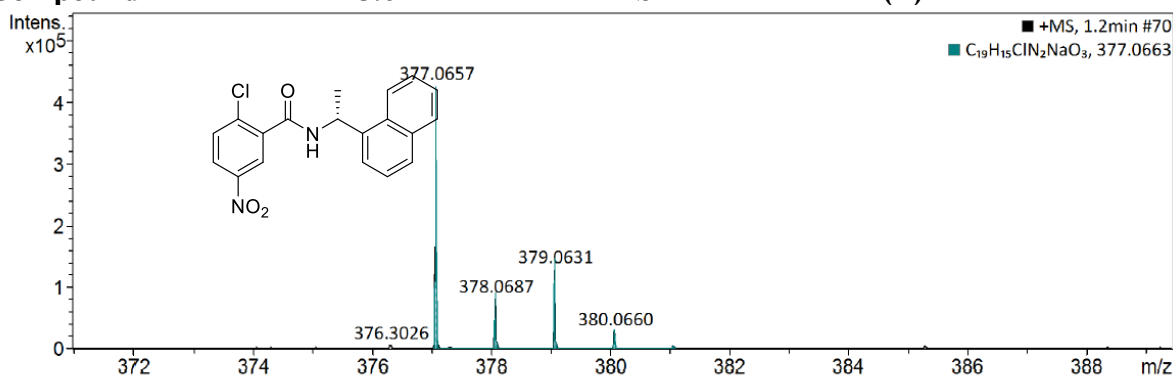
Compound

5.6

¹H-NMR

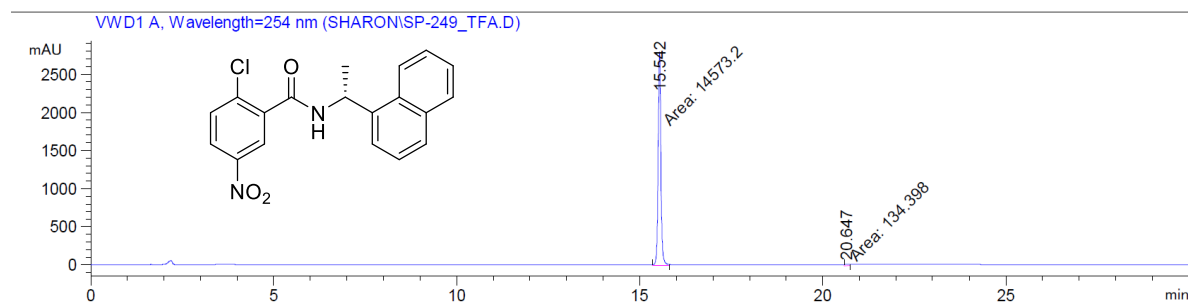


Compound 5.6 ESI (+) HRMS



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdB	e ⁻ Conf	N-Rule
377.0657	1	C17H10ClN8O	377.0661	0.9	23.8	1	100.00	16.5	even	ok
	2	C21H13O7	377.0656	-0.3	154.2	2	0.27	15.5	even	ok
393.0394	1	C19H15ClN2NaO3	377.0663	1.7	23.2	1	100.00	12.5	even	ok
	1	C19H10ClN4O4	393.0385	-2.2	23.8	1	89.69	16.5	even	ok
	2	C20H6ClN8	393.0398	1.2	28.5	2	100.00	21.5	even	ok
	3	C24H9O6	393.0394	-0.1	204.9	3	0.01	20.5	even	ok
	4	C25H5N4O2	393.0407	3.3	208.3	4	0.00	25.5	even	ok
	5	C21HN10	393.0380	-3.5	211.8	5	0.00	26.5	even	ok
	1	C22H11ClN2NaO2	393.0401	1.9	27.0	1	100.00	17.5	even	ok
2	C18H7ClN8Na	393.0374	-5.0	28.1	2	36.10	18.5	even	ok	

Compound 5.6 HPLC



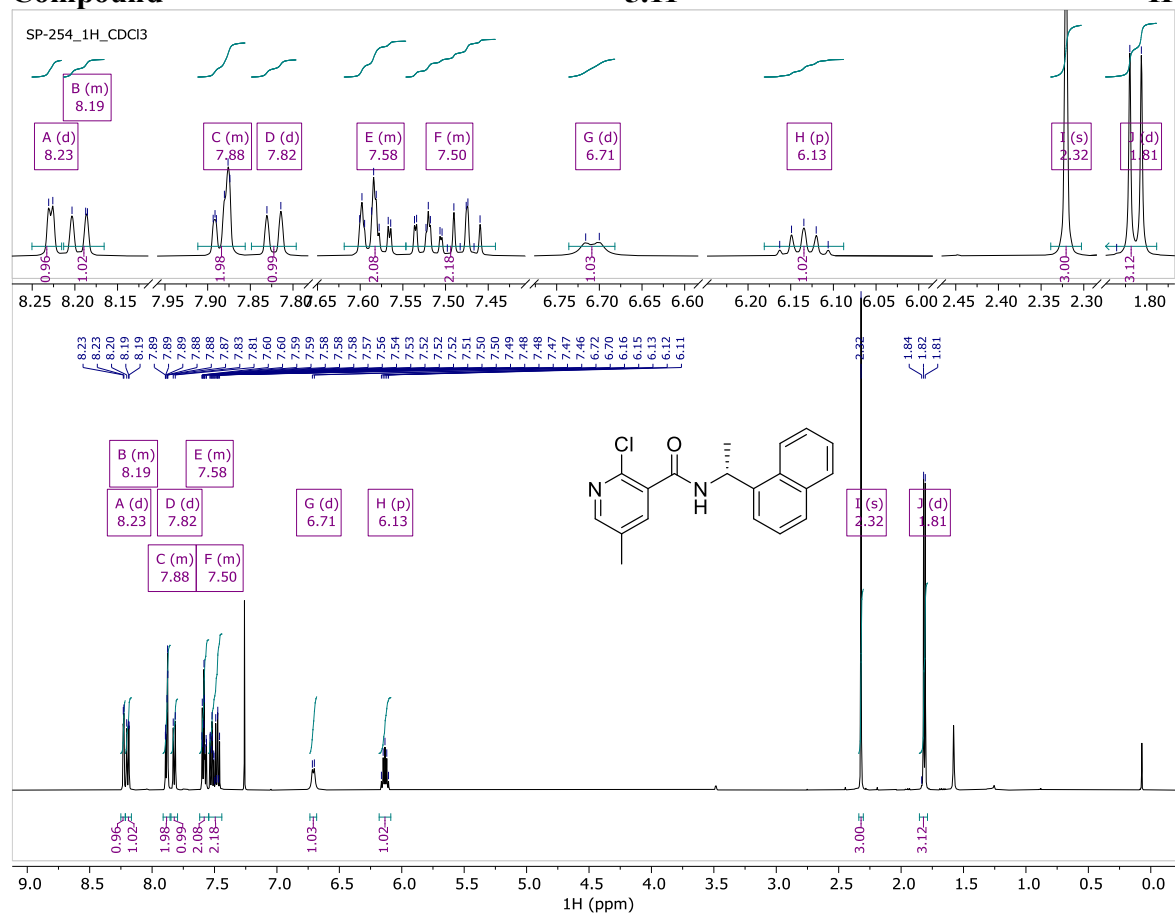
Signal 2: VWD1 A, Wavelength=254 nm

Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
1	15.542	MM	0.0865	1.45732e4	2808.60132	99.0862
2	20.647	MM	0.1314	134.39787	17.04819	0.9138

Totals : 1.47076e4 2825.64951

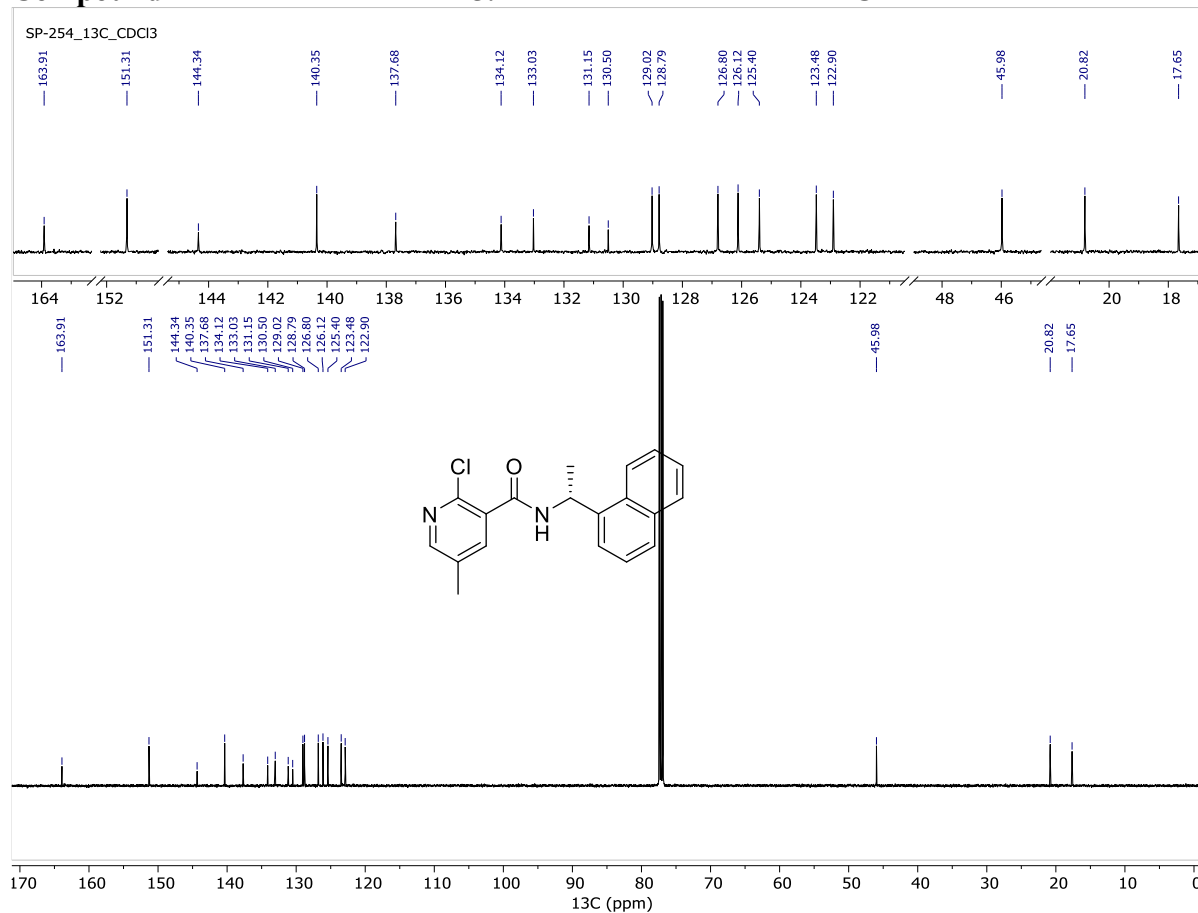
Compound

5.11

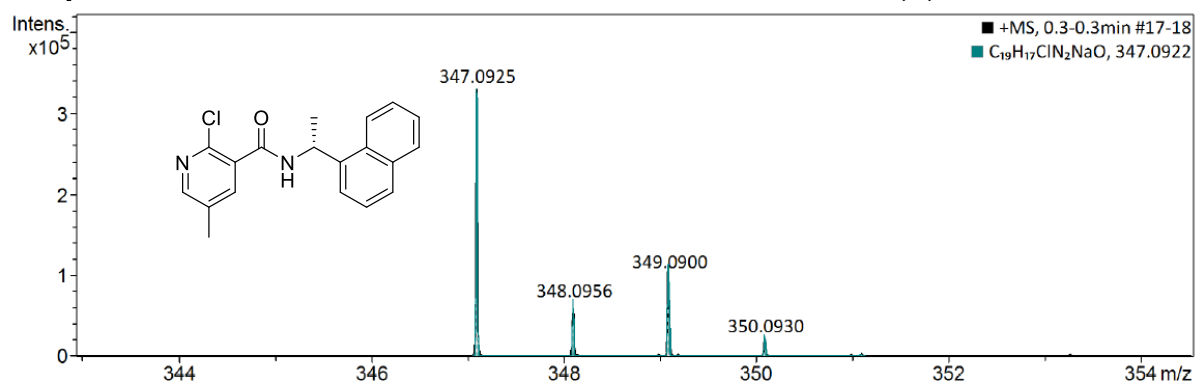
¹H-NMR

Supplementary information Chapter 5

Compound 5.11 ¹³C NMR

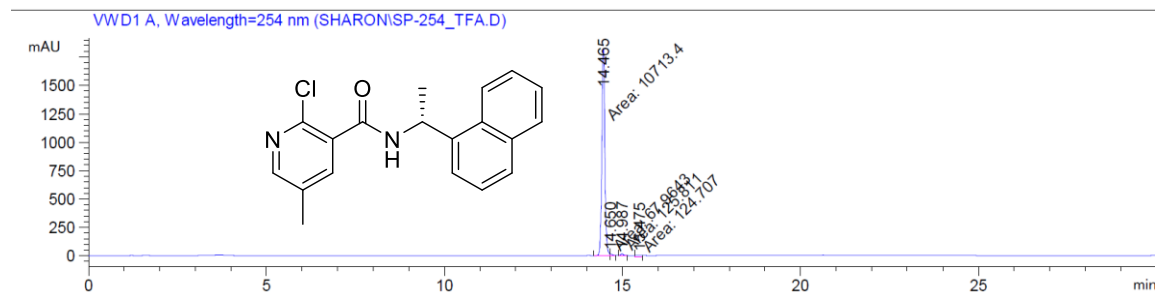


Compound 5.11 ESI (+) HRMS



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
347.0925	1	C ₂₂ H ₁₁ N ₄ O	347.0927	0.6	160.8	1	100.00	19.5	even	ok
	1	C ₁₉ H ₁₇ ClN ₂ NaO	347.0922	-1.1	21.8	1	100.00	11.5	even	ok
	1	C ₁₆ H ₂₁ ClKN ₂ O ₂	347.0923	-0.6	49.3	1	100.00	6.5	even	ok

Compound 5.11 HPLC



Signal 2: VWD1 A, Wavelength=254 nm

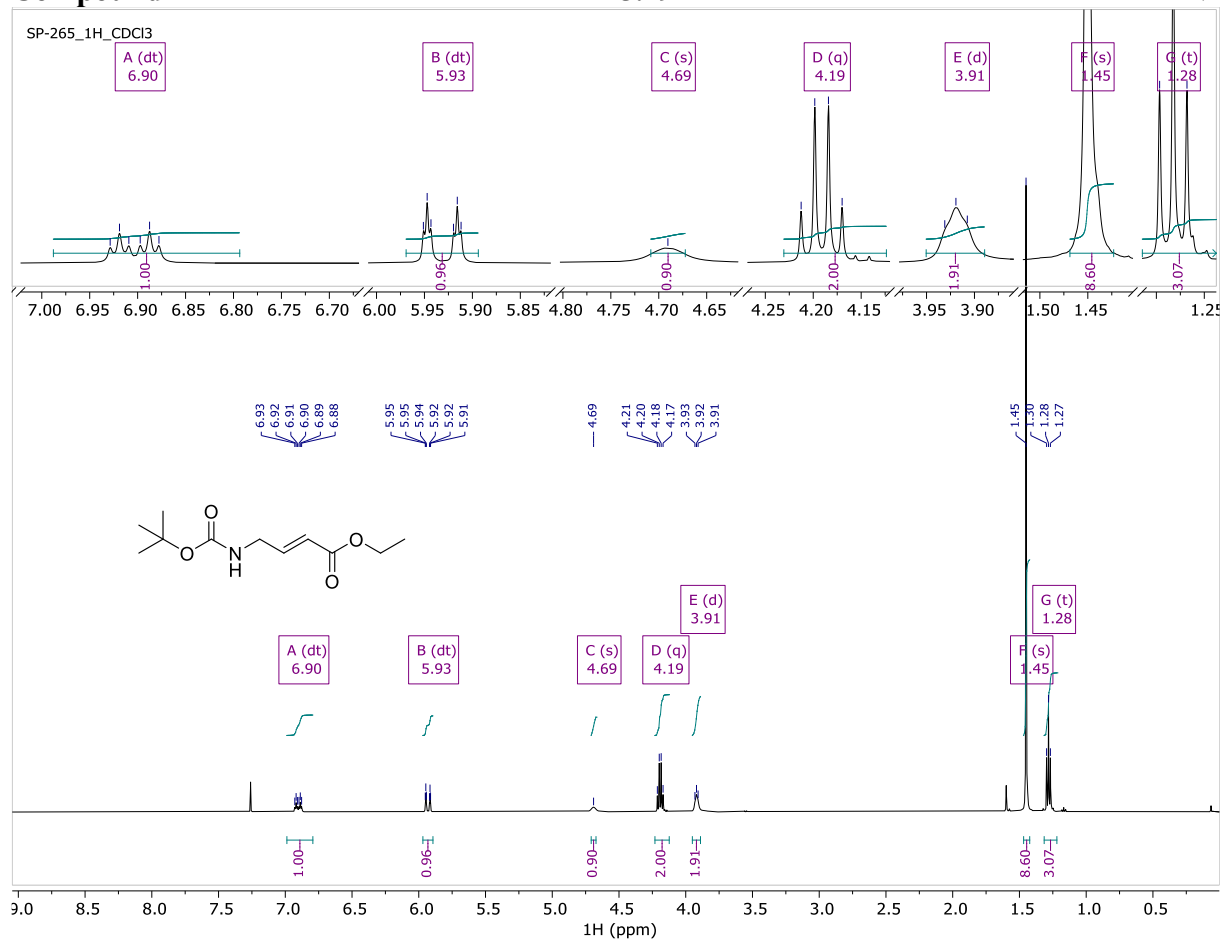
Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
1	14.465	MF	0.0976	1.07134e4	1830.32690	97.1131
2	14.650	FM	0.0820	67.96427	13.81158	0.6161
3	14.987	MM	0.1104	125.81104	18.99604	1.1404
4	15.475	MM	0.1842	124.70669	11.28097	1.1304

Totals : 1.10318e4 1874.41549

Compound

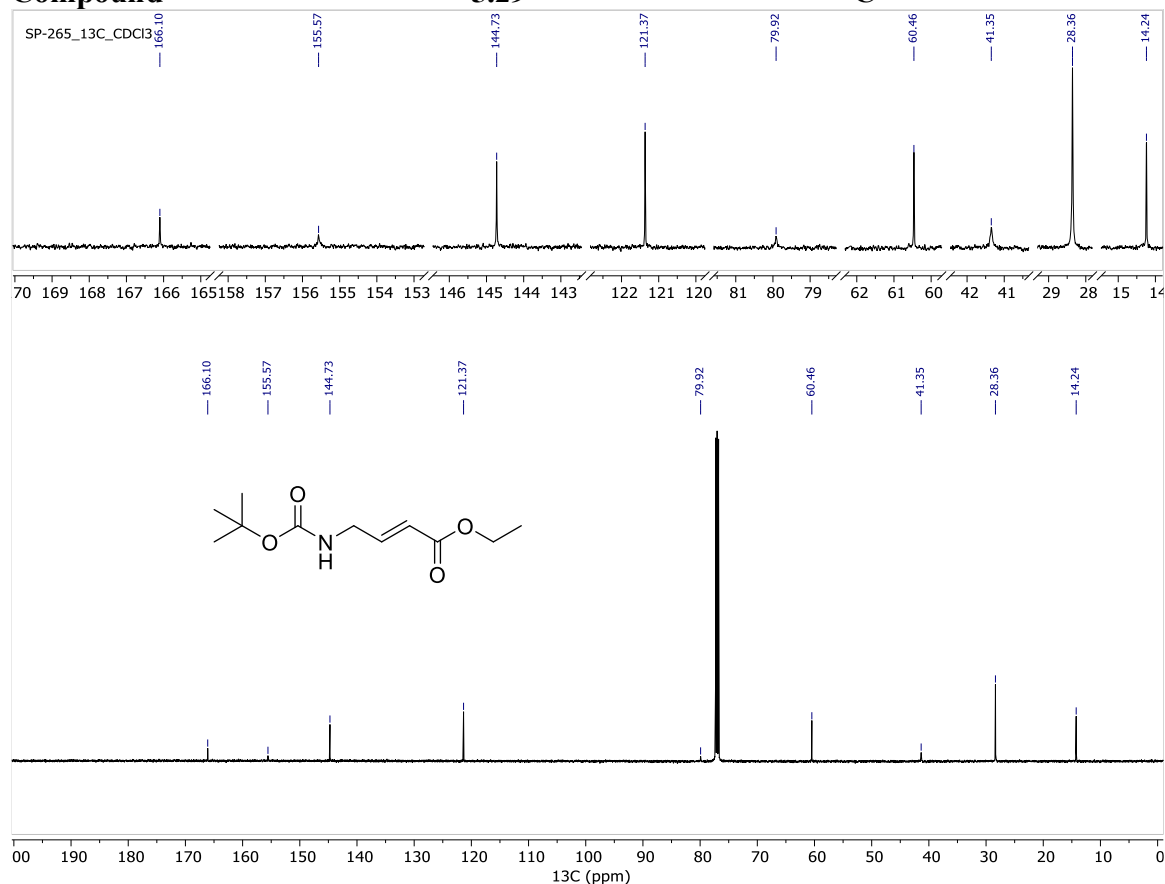
5.29

¹H-NMR

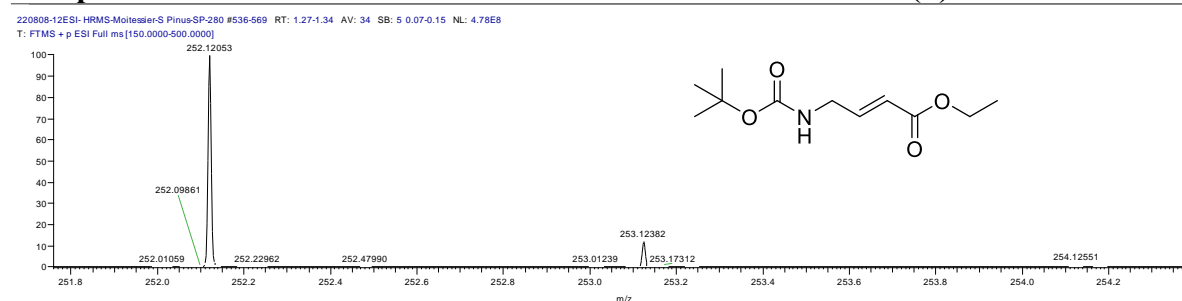


Supplementary information Chapter 5

Compound 5.29 ¹³C NMR



Compound 5.29 ESI (+) HRMS



220808-12ESI-HRMS-Moiteissier-S Pinus-SP-280#536-569 RT: 1.27-1.34 AV: 34

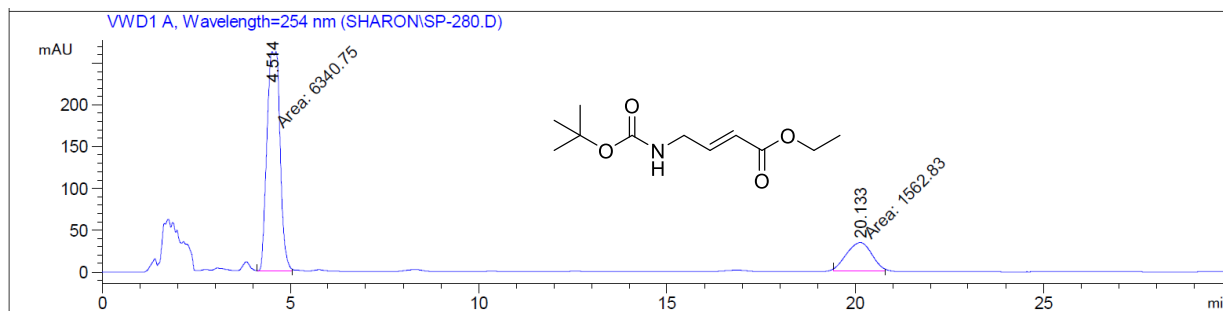
SB: 5 0.07-0.15

T: FTMS + p ESI Full ms [150.0000-500.0000]

m/z = 252.11127-252.12580

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
252.12053	478648480.0	100.00	34605.03	1.00	252.12063	-0.40	2.5	C ₁₁ H ₁₉ O ₄ N Na
					252.11956	3.83	-1.0	C ₁₀ H ₂₅ N ₂ Br

Compound 5.29 HPLC

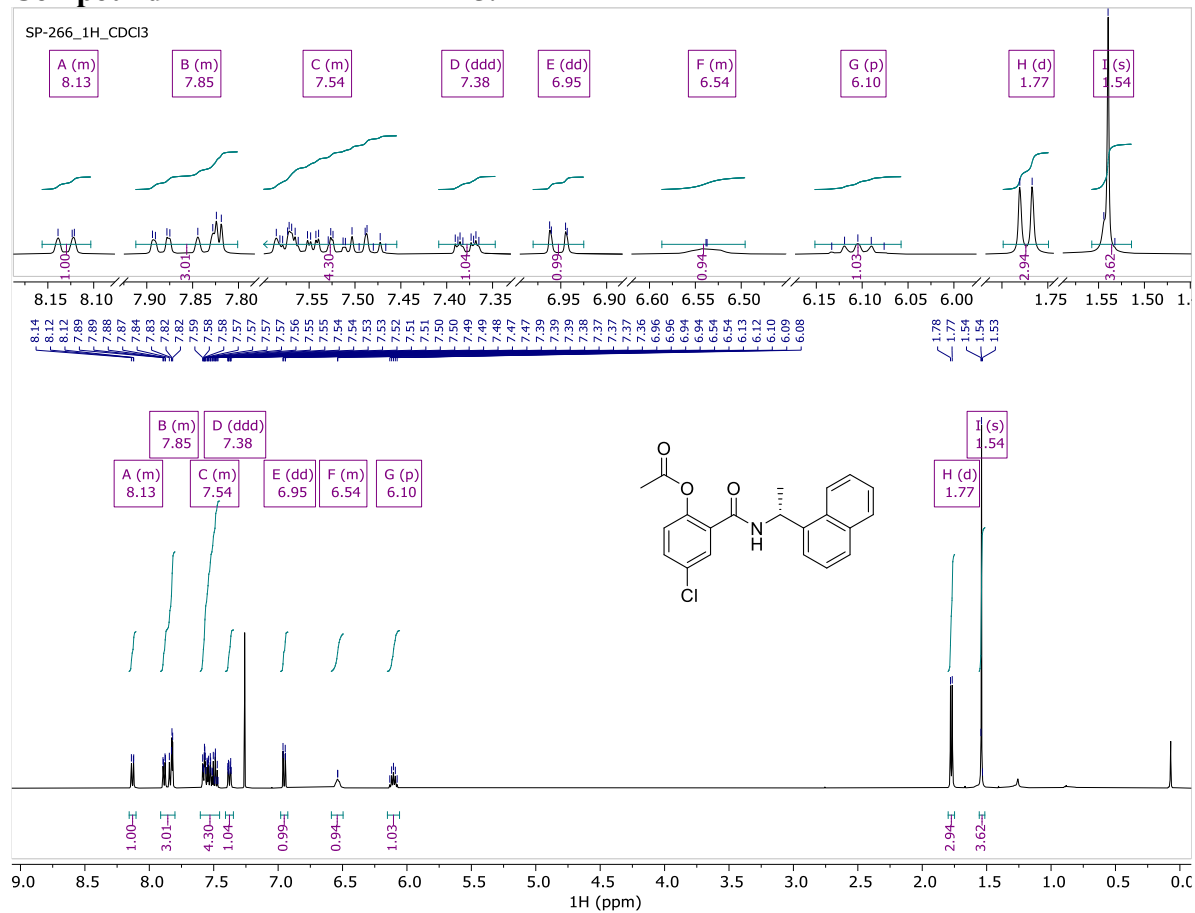


Signal 2: VWD1 A, Wavelength=254 nm

Peak #	RetTime [min]	Type	Width [min]	Area mAU	Area *s	Height [mAU]	Area %
1	4.514	MM	0.4026	6340.75146	28714.5	262.52048	80.2263
2	20.133	MM	0.7747	1562.82825	31456.6	33.62426	19.7737

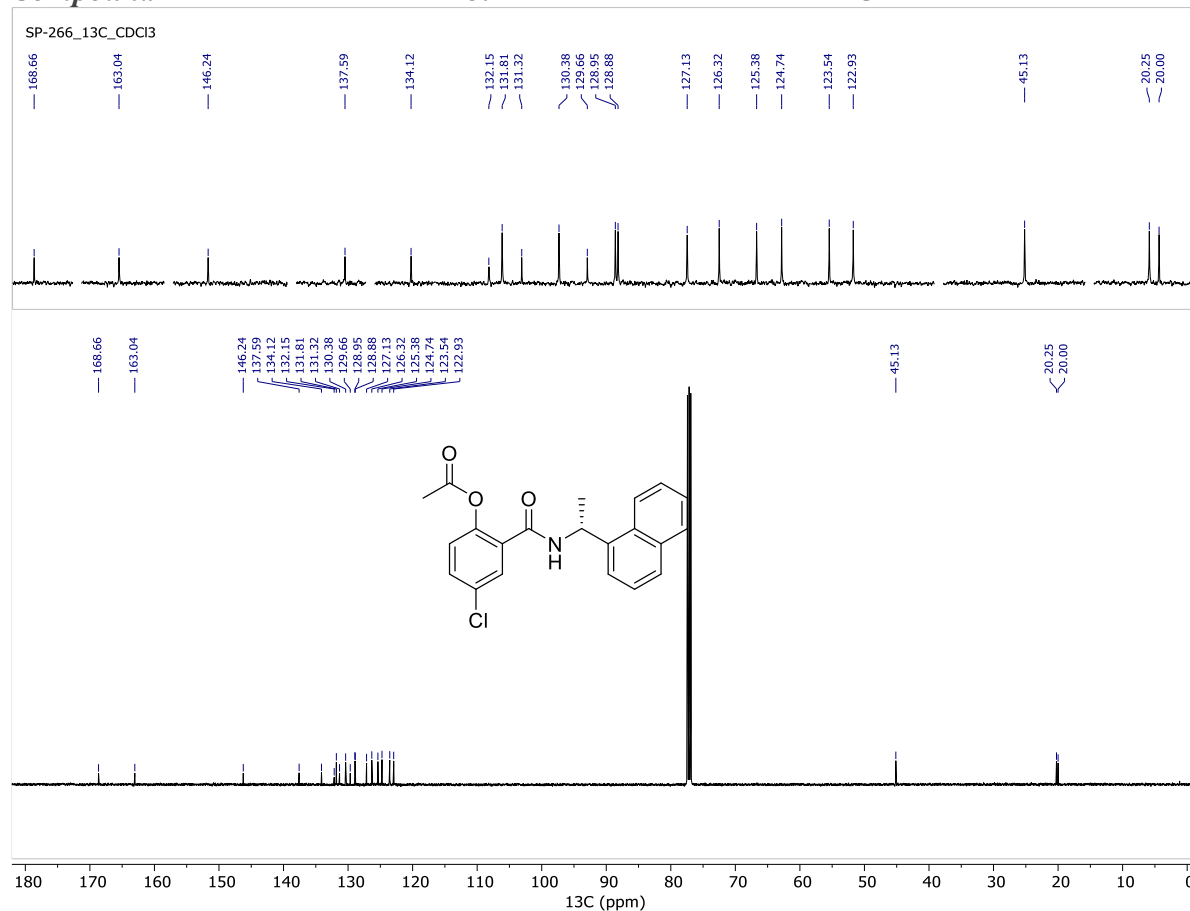
Totals : 7903.57971 296.14474

Compound 5.14 ¹H NMR

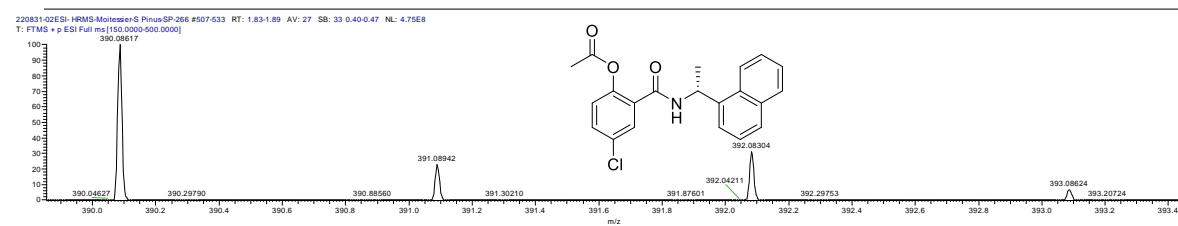


Supplementary information Chapter 5

Compound 5.14 ¹³C NMR



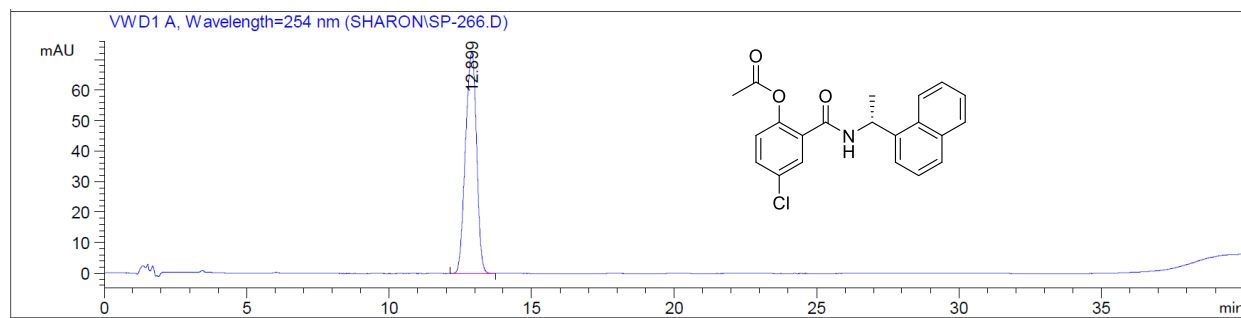
Compound 5.14 ESI (+) HRMS



220831-02ESI- HRMS-Moiteissier-S Pinus-SP-266#507-533 RT: 1.83-1.89 AV: 27
SB: 33 0.40-0.47
T: FTMS + p ESI Full ms [150.0000-500.0000]
m/z = 390.06733-390.10765

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
390.08617	474878976.0	100.00	26944.55	1.00	390.08674	-1.46	12.5	C ₂₁ H ₁₈ O ₃ NClNa

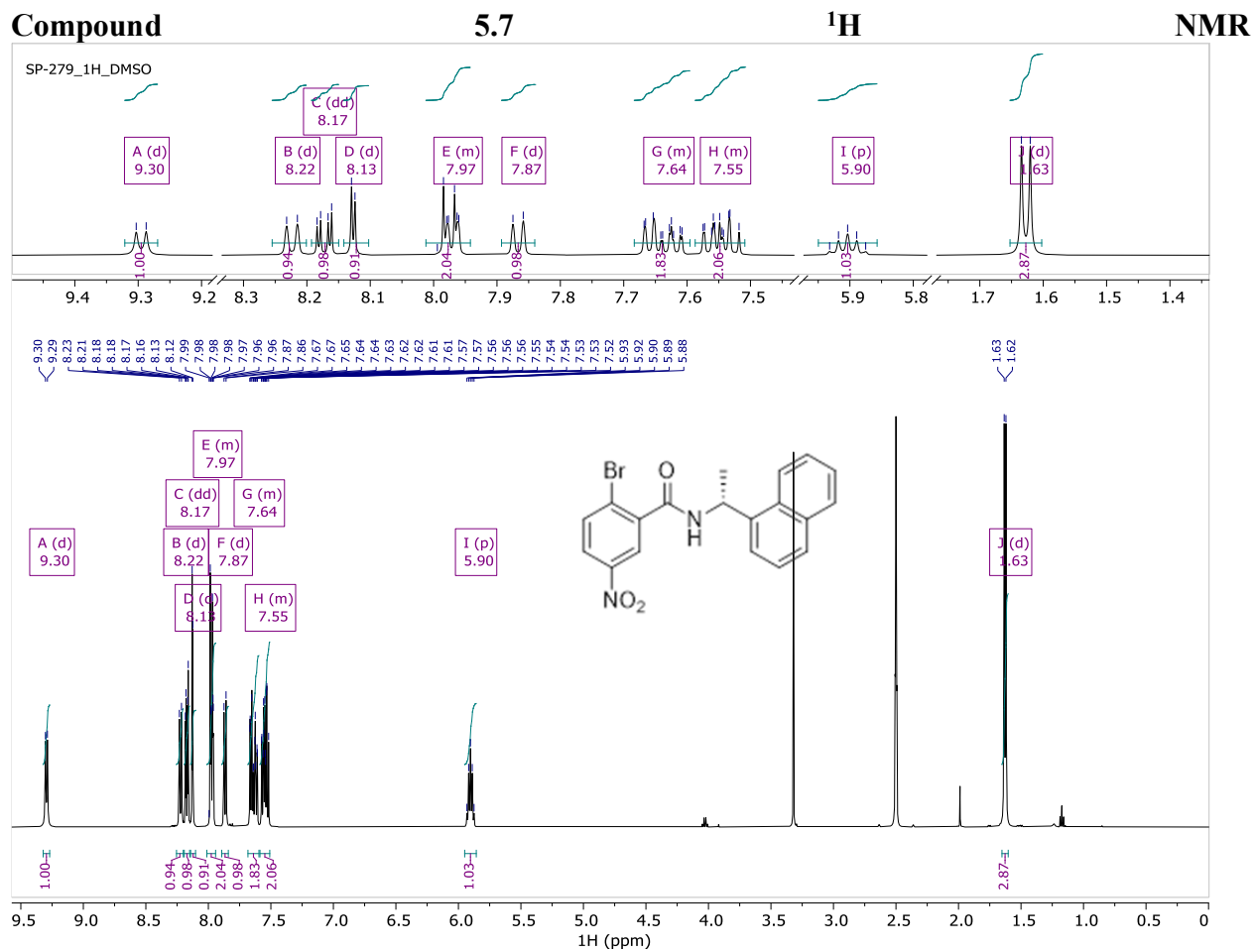
Compound 5.14 HPLC



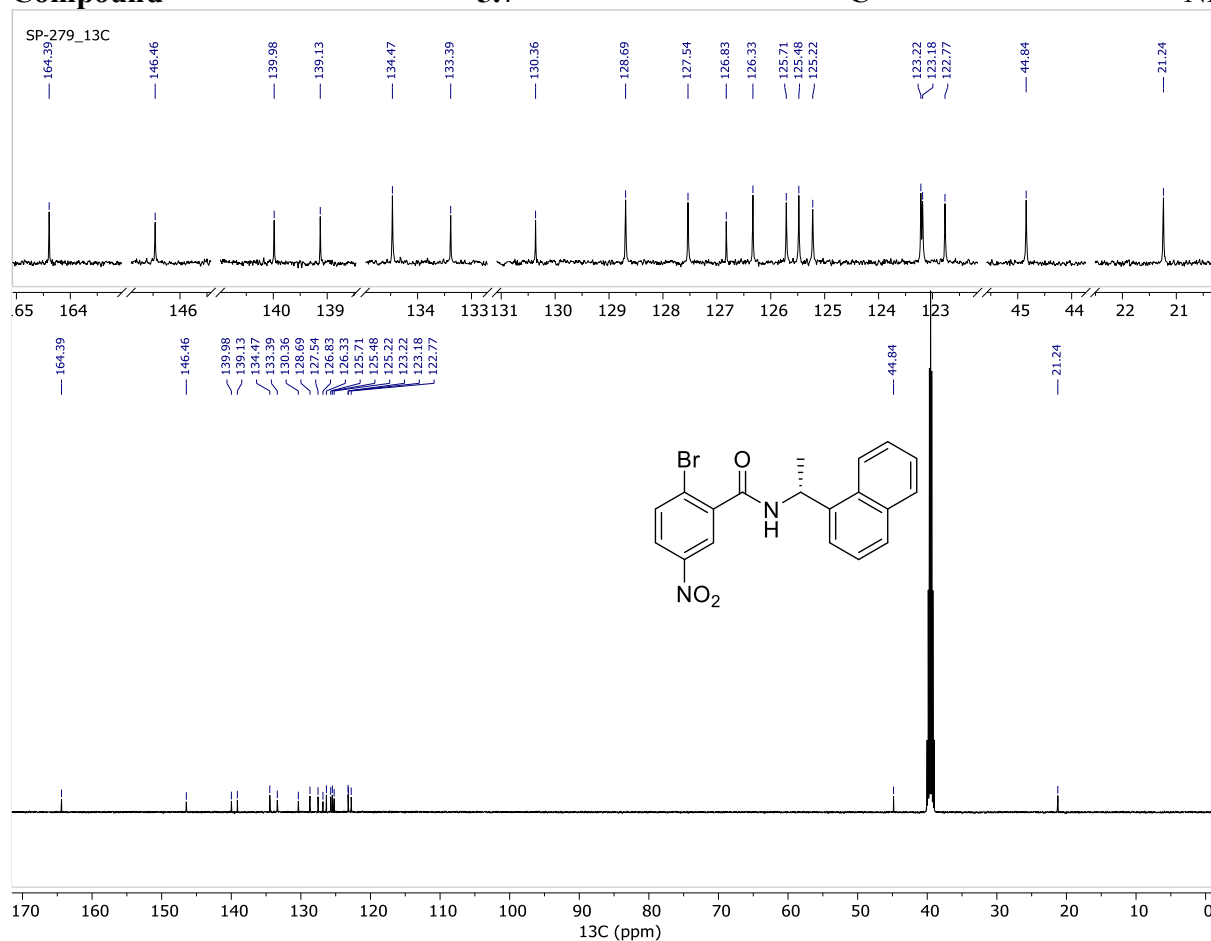
Signal 2: VWD1 A, Wavelength=254 nm

Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
1	12.899	BB	0.4423	1964.23779	72.60744	100.0000

Totals : 1964.23779 72.60744

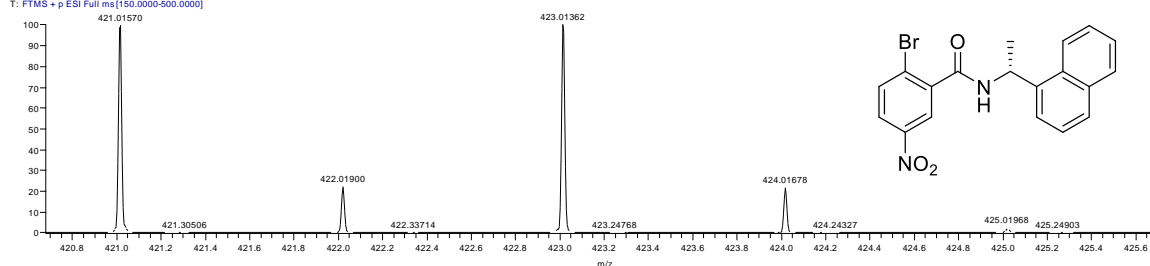


Compound 5.7 ¹³C NMR



Compound 5.7 ESI (+) HRMS

220808-13ESI- HRMS-Moitessier-S Pinus-SP-279 #218-222 RT: 0.53-0.54 AV: 5 SB: 6 0.14-0.15 NL: 5.78E7
T: FTMS + p ESI Full ms [150.0000-500.0000]



220808-13ESI- HRMS-Moitessier-S Pinus-SP-279#218-222 RT: 0.53-0.54 AV: 5

SB: 6 0.14-0.15

T: FTMS + p ESI Full ms [150.0000-500.0000]

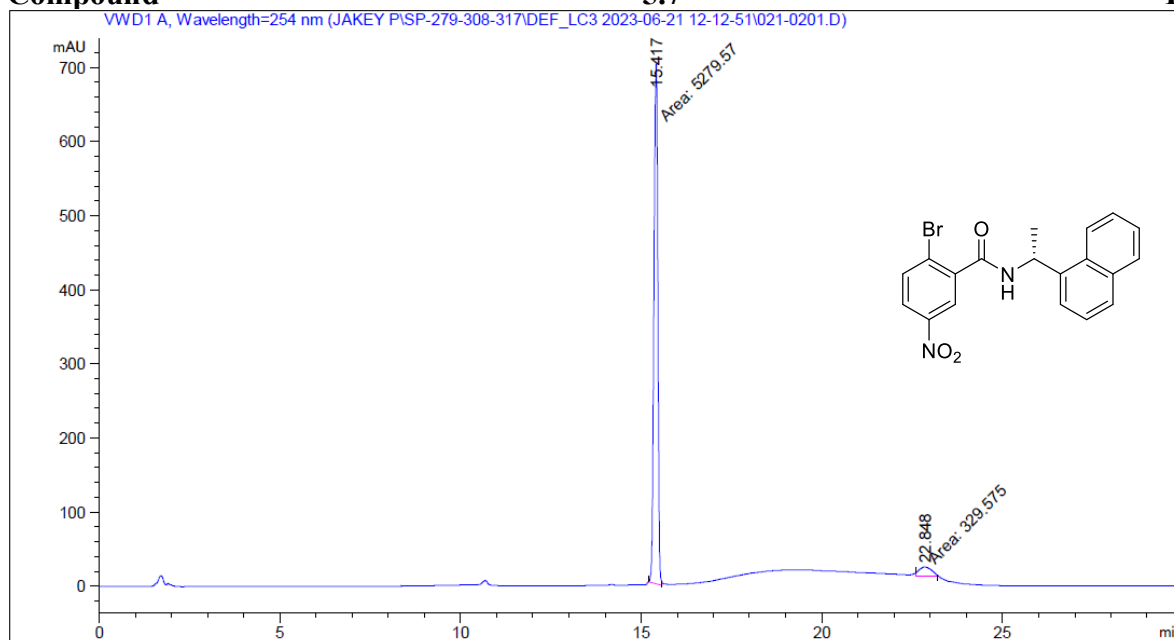
m/z= 420.99850-421.03423

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
421.01570	59765920.0	100.00	25308.31	1.00	421.01583	-0.31	12.5	C ₁₉ H ₁₅ O ₃ N ₂ Br Na

Compound

5.7

HPLC

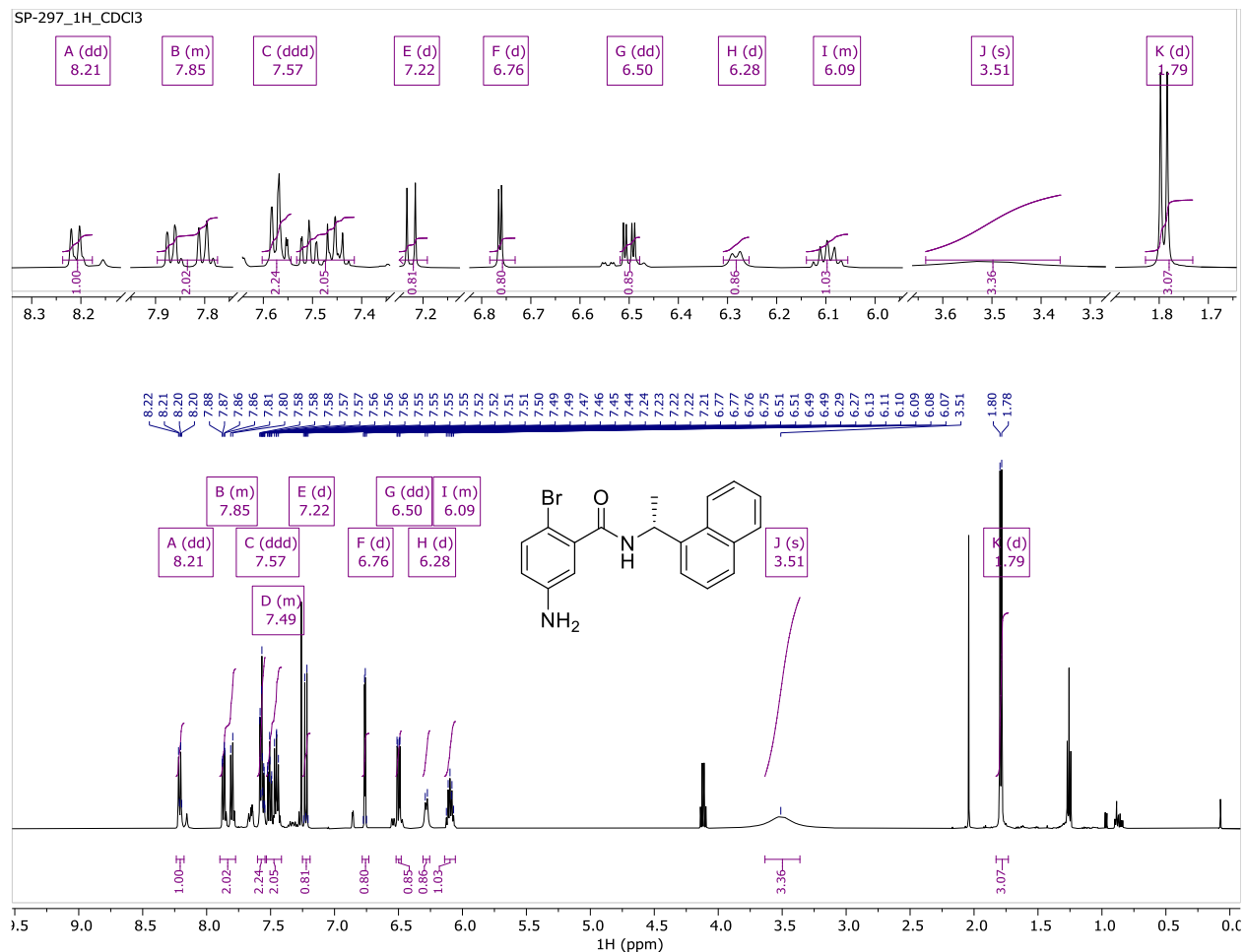


Signal 1: VWD1 A, Wavelength=254 nm

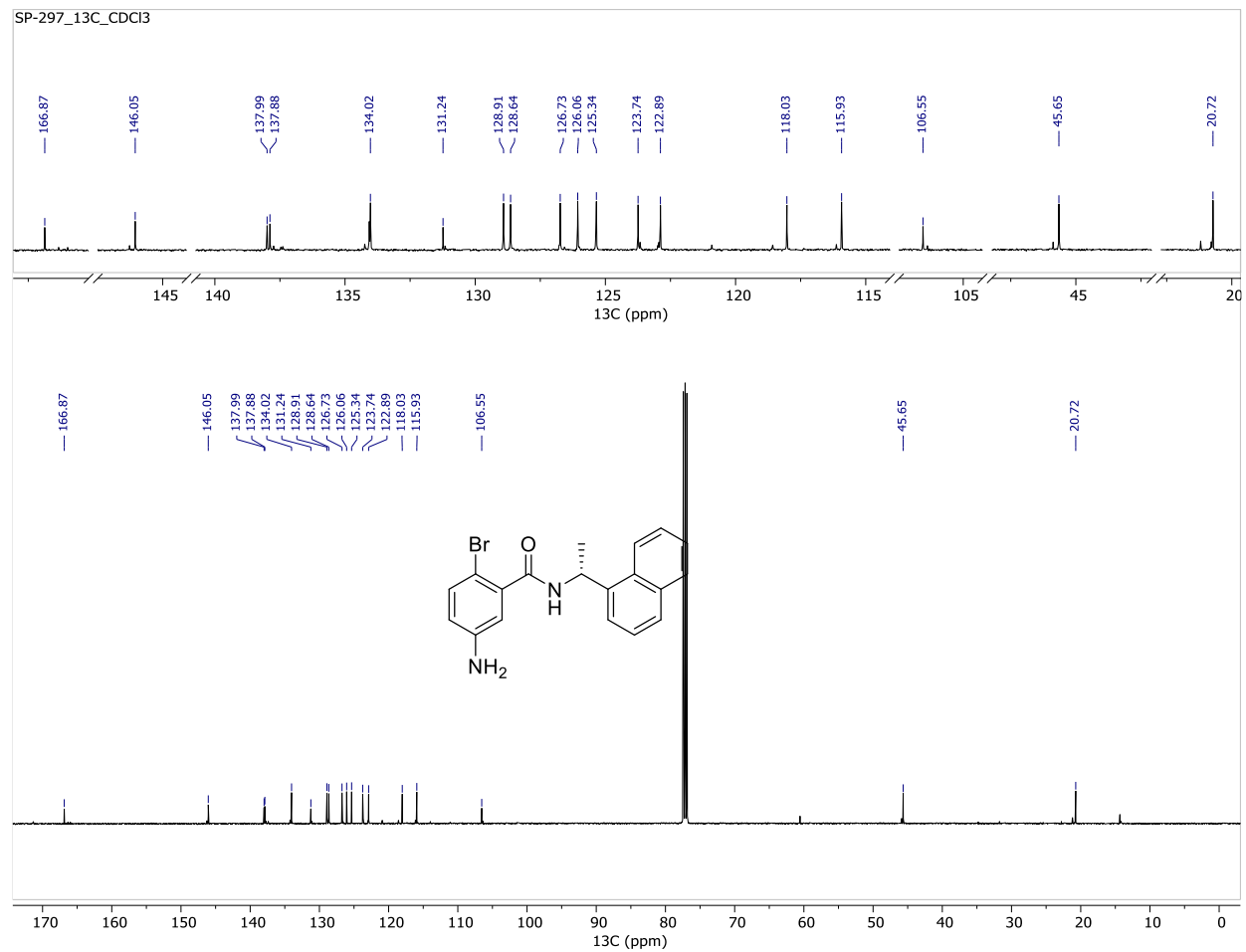
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	15.417	MM	0.1251	5279.57422	703.13385	94.1243
2	22.848	MM	0.4237	329.57489	12.96393	5.8757

Totals : 5609.14911 716.09778

Compound 5.8 ¹H-NMR



Compound 5.8 ¹³C-NMR



Supplementary information Chapter 5

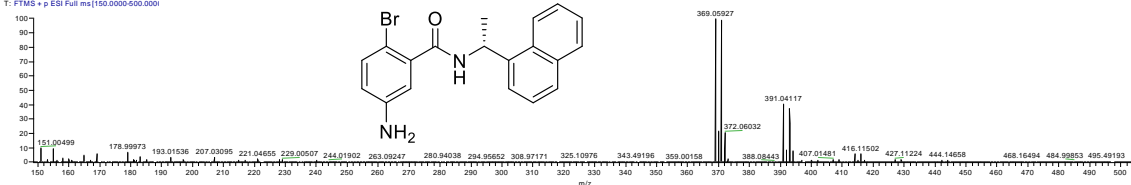
Compound 5.8 ESI (+) HRMS

230717-04-ESI-HRMS-Moitess

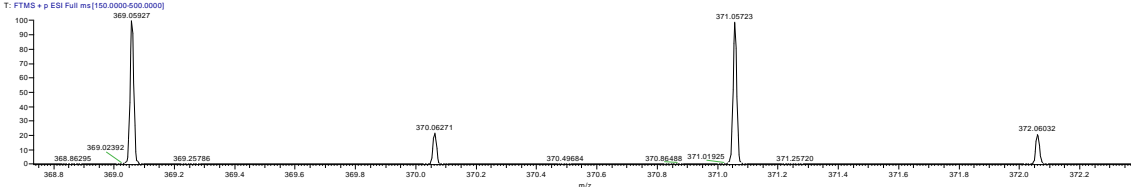
07/17/23 11:30:05

230717-04-ESI-HRMS-Moitessier-Sharon Pinus
T: FTMS + p ESI Full ms [150.0000-500.0000]

.051.07 AV: 12 SB: 23 0.31-0.36 NL: 6.07E7



230717-04-ESI-HRMS-Moitessier-Sharon Pinus-SP-297#440-451 RT: 1.05-1.07 AV: 12 SB: 23 0.31-0.36 NL: 6.07E7
T: FTMS + p ESI Full ms [150.0000-500.0000]



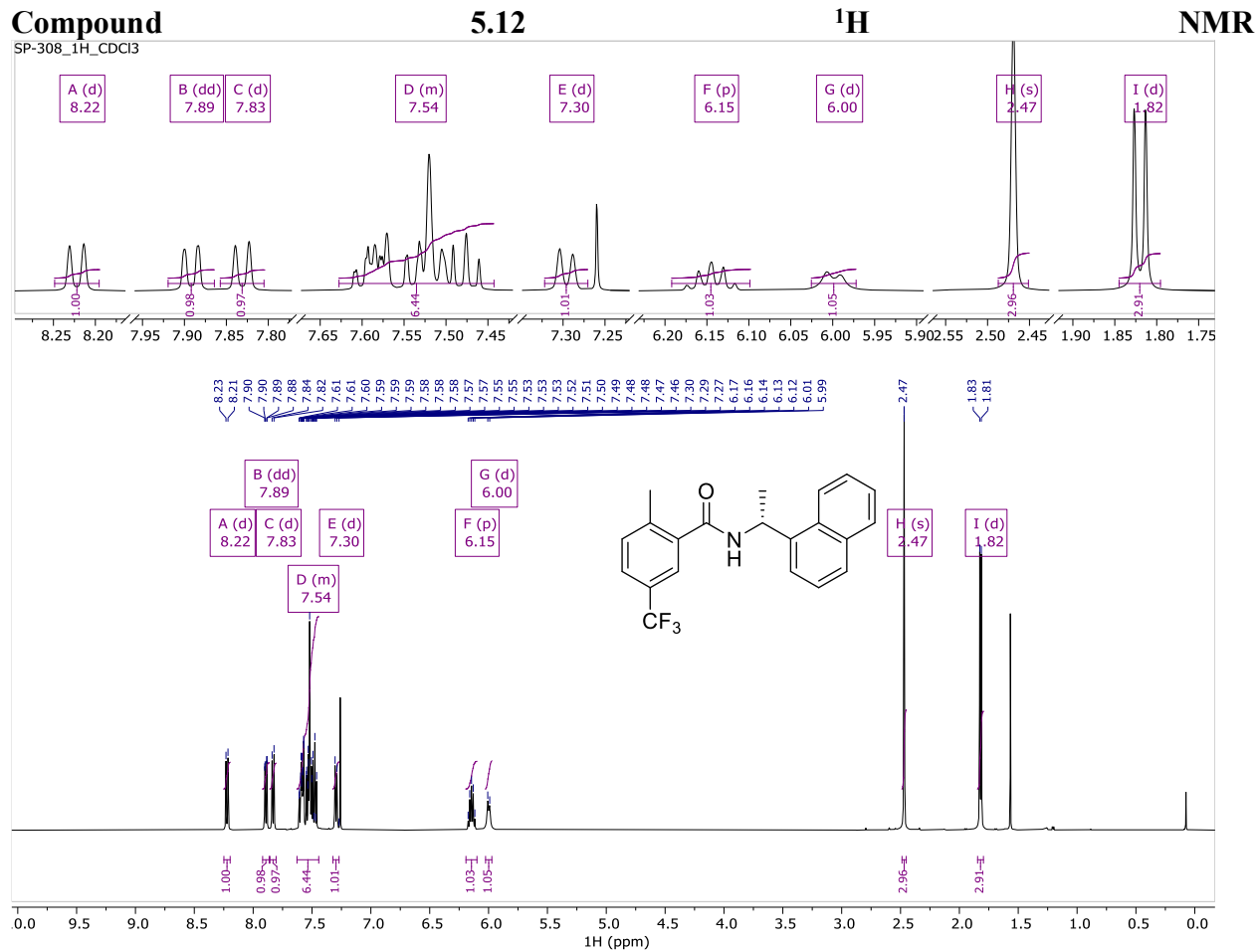
230717-04-ESI HRMS-Moitessier-Sharon Pinus-SP-297#440-451 RT: 1.05-1.07 AV: 12

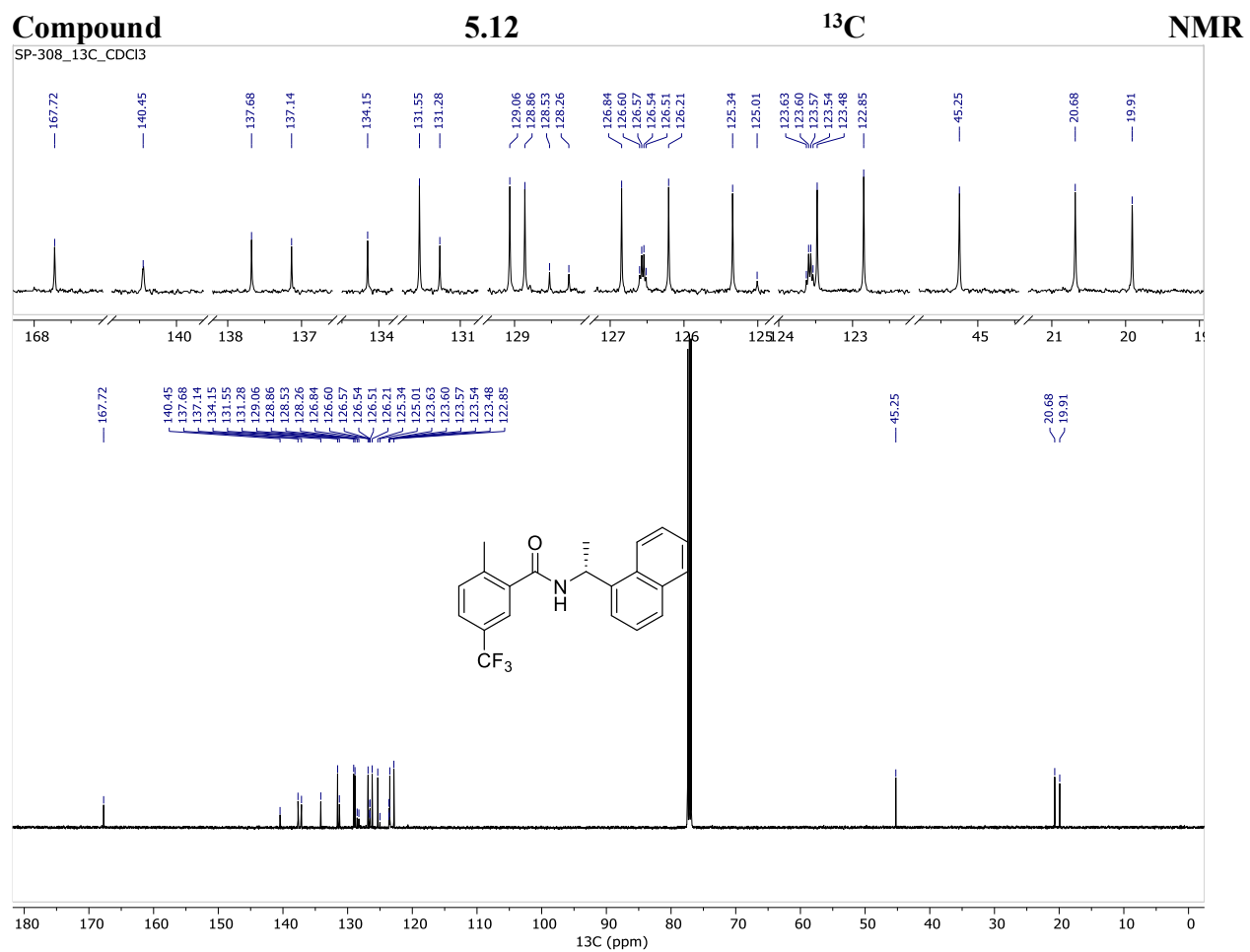
SB: 23 0.31-0.36

T: FTMS + p ESI Full ms [150.0000-500.0000]

m/z: 369.04111-369.07311

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
369.05927	61065436.0	100.00	27153.43	1.00	369.05970	-1.17	11.5	C ₁₉ H ₁₈ O ₂ Br

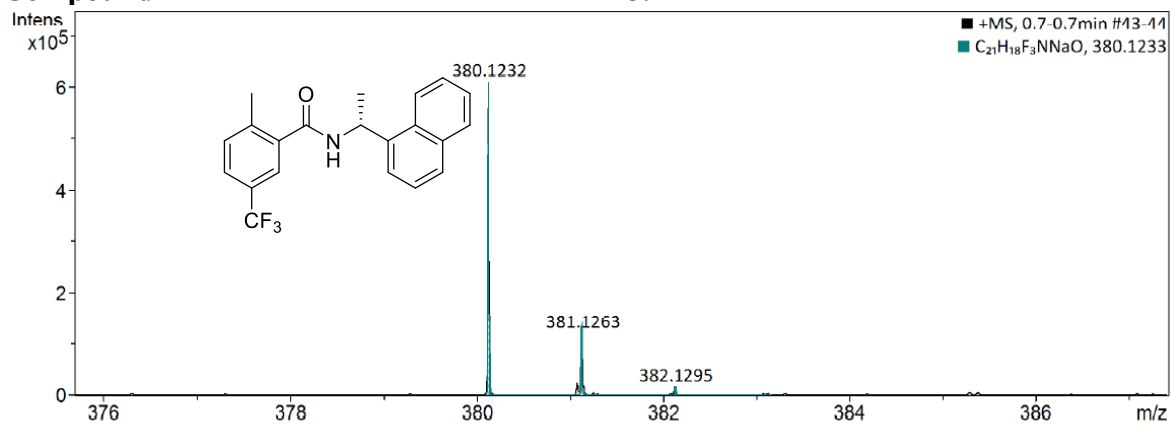




Compound

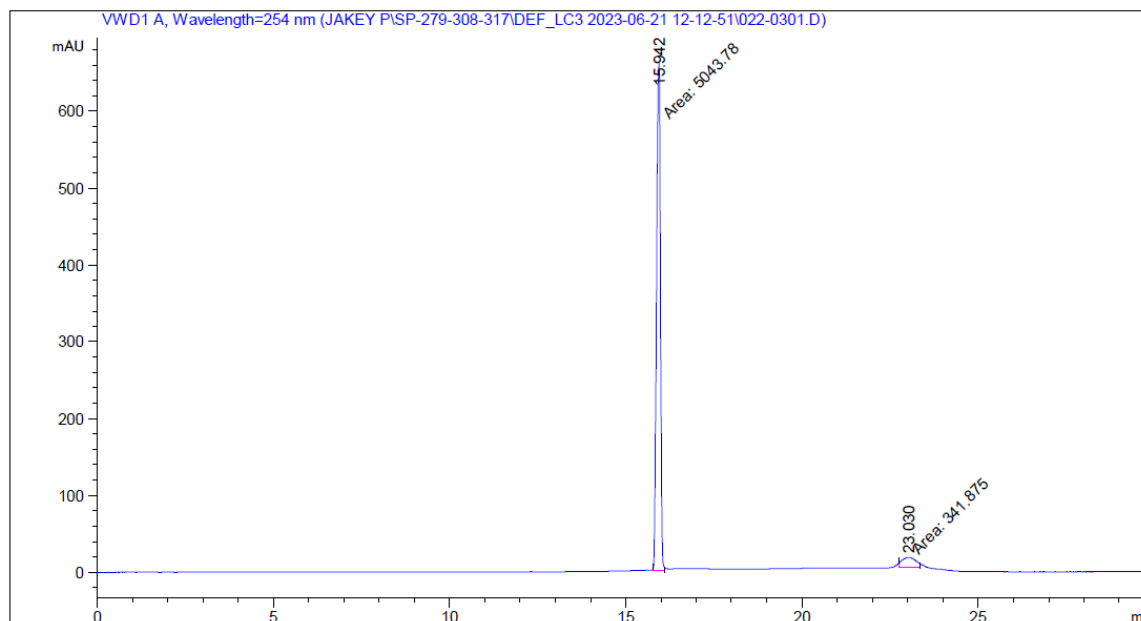
5.12

HRMS



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
380.1232	1	C ₁₈ H ₁₇ F ₃ N ₃ O ₃	380.1217	-4.2	7.3	1	100.00	10.5	even	ok
	1	C ₂₁ H ₁₈ F ₃ NNaO	380.1233	0.1	21.4	1	100.00	11.5	even	ok

Compound 5.12 HPLC

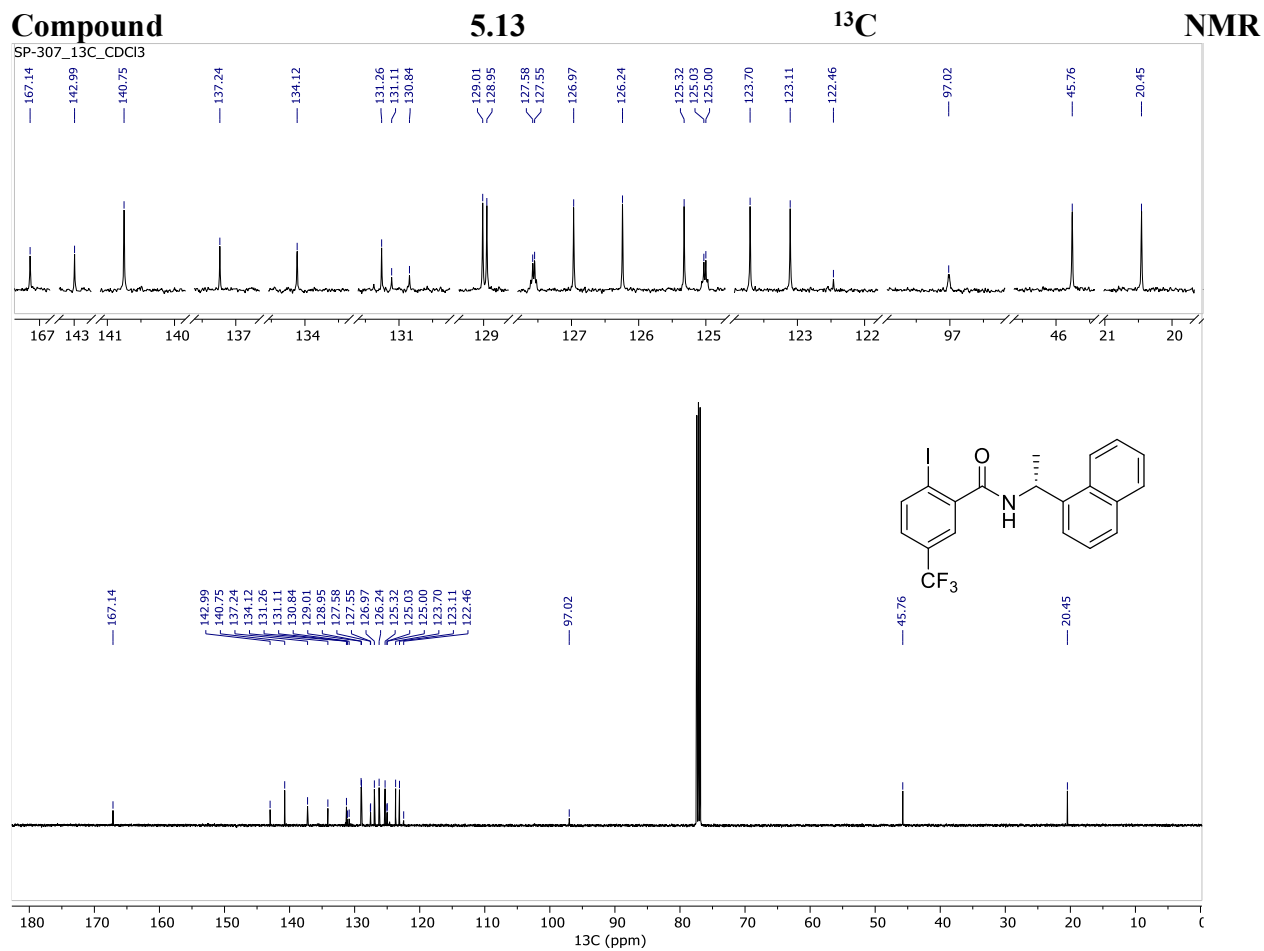


Signal 1: VWD1 A, Wavelength=254 nm

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	15.942	MM	0.1272	5043.77637	661.01917	93.6521
2	23.030	MM	0.4438	341.87506	12.83944	6.3479

Totals : 5385.65143 673.85860

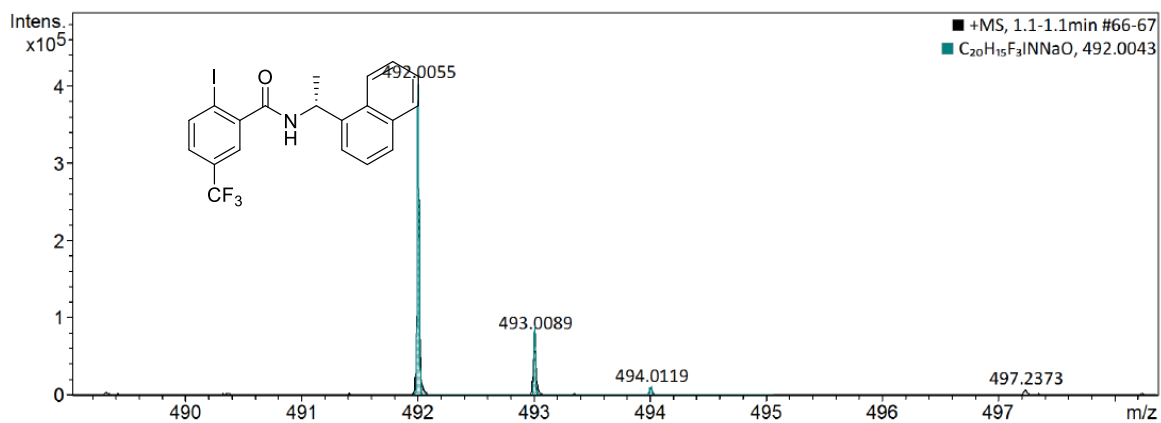




Compound

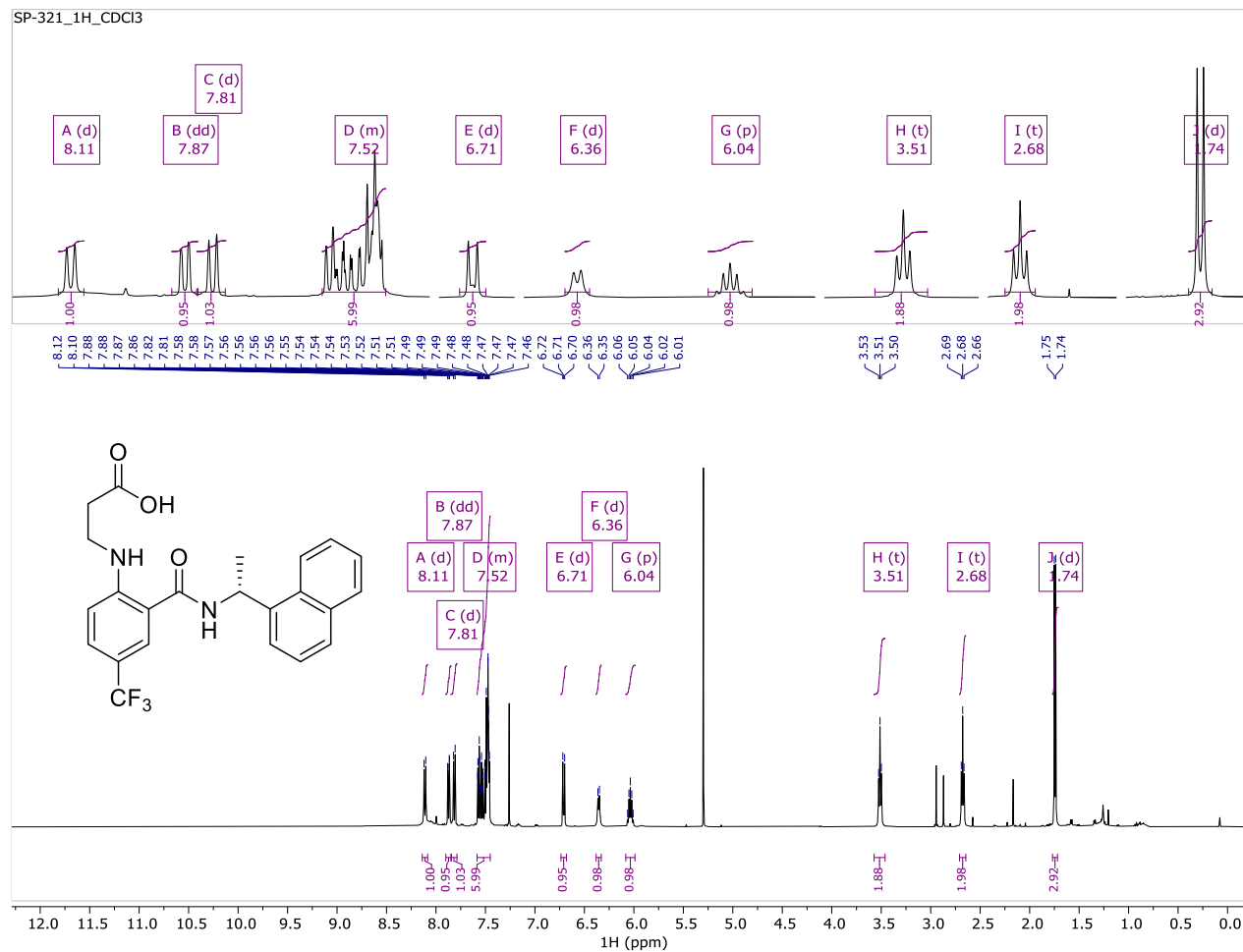
5.13

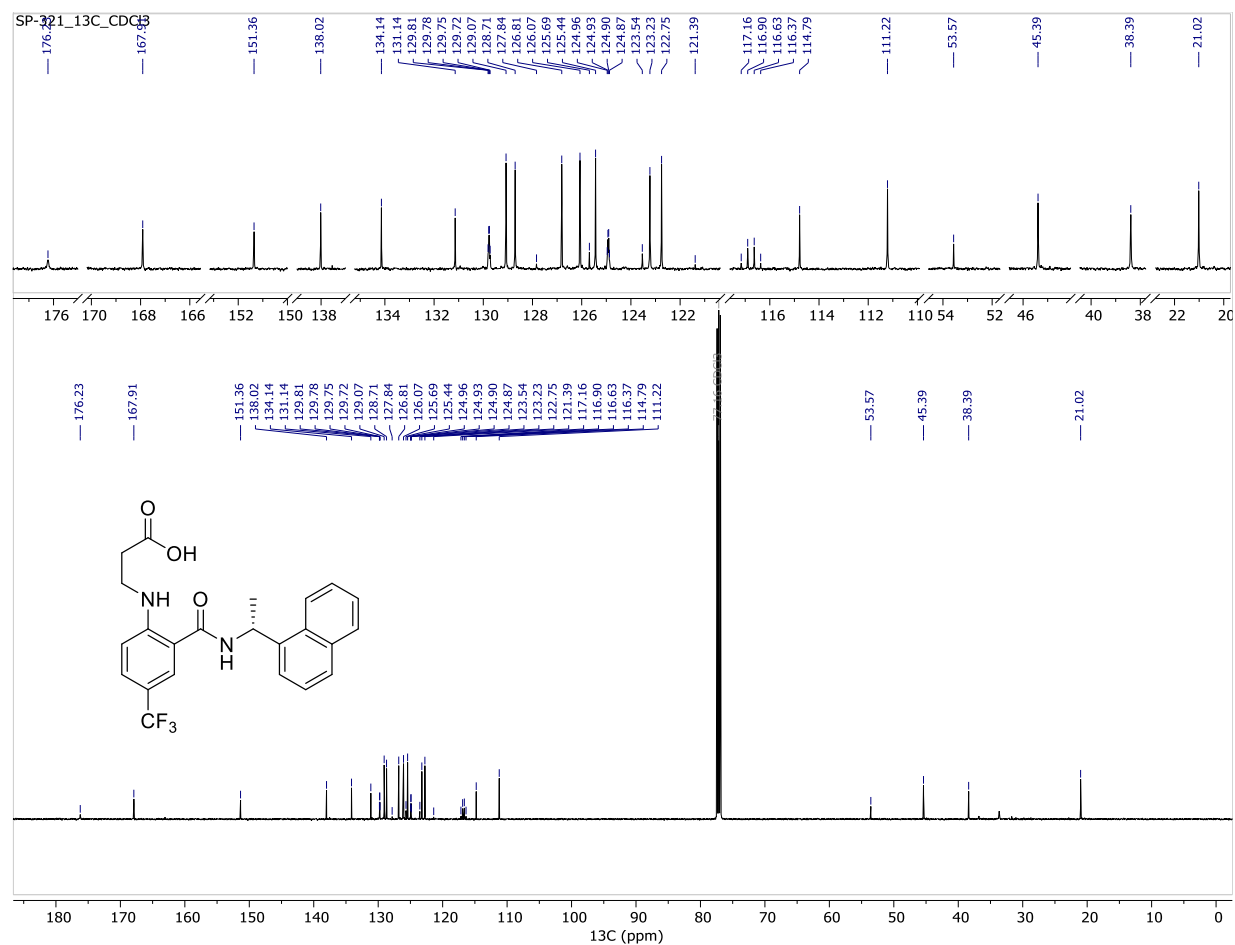
HRMS



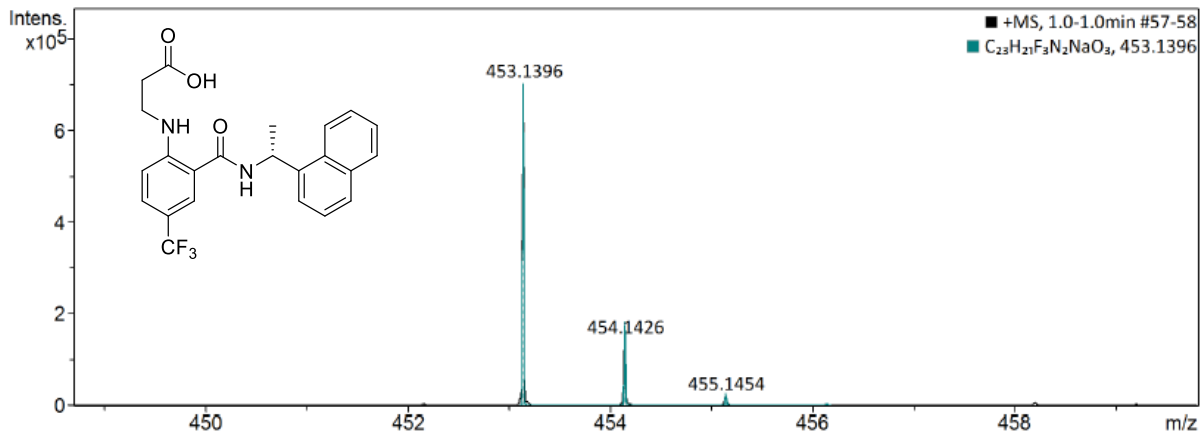
Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
492.0055	1	C22H14F3INO	492.0067	2.3	30.2	1	100.00	14.5	even	ok
	2	C6H14F3IN9O6	492.0058	0.6	52.2	2	83.96	2.5	even	ok
	3	C3H6F3IN19	492.0045	-2.1	52.5	3	56.13	8.5	even	ok
	1	C20H15F3INNaO	492.0043	-2.6	17.5	1	100.00	11.5	even	ok
	2	C5H11F3IN13NaO2	492.0048	-1.6	51.5	2	55.75	4.5	even	ok

Compound 5.28 ¹H-NMR



Compound 5.28 ¹³C-NMR

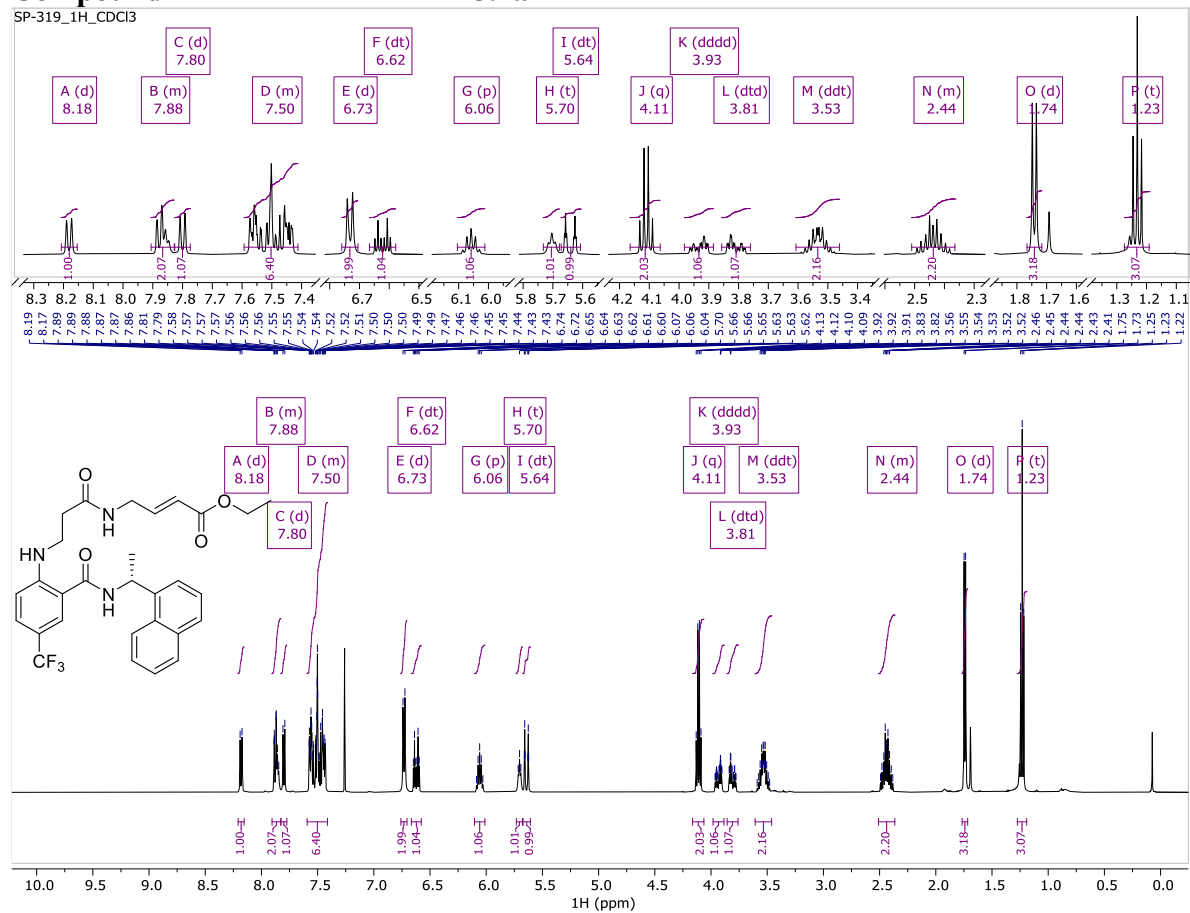
Compound 5.28 HRMS

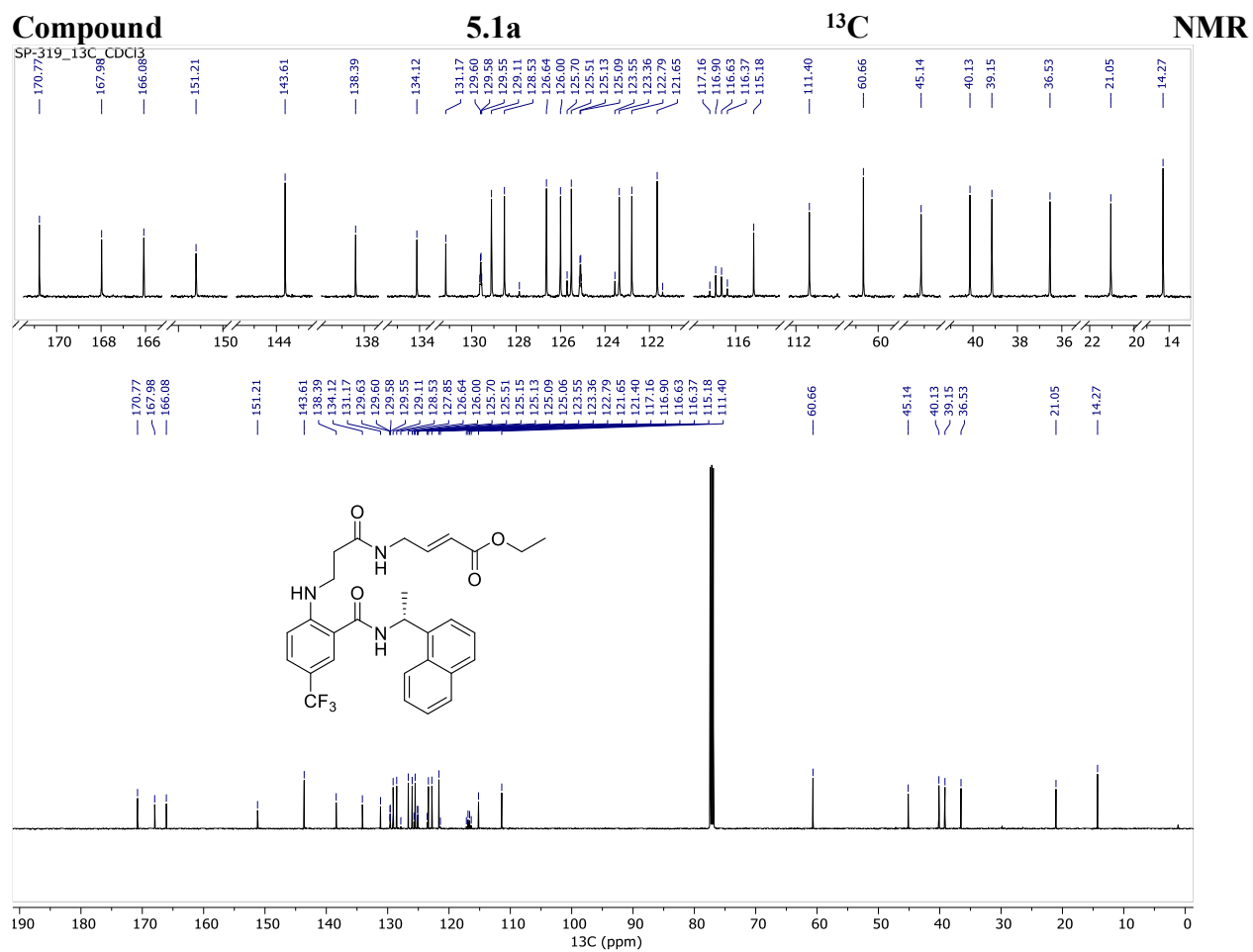


Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
431.1577	1	C23H22F3N2O3	431.1577	0.0	7.9	1	100.00	12.5	even	ok
	1	C26H23F3NaO	431.1593	3.8	8.3	1	100.00	13.5	even	ok
453.1396	1	C20H20F3N4O5	453.1380	-3.5	18.6	1	63.93	11.5	even	ok
	2	C21H16F3N8O	453.1394	-0.6	31.7	2	100.00	16.5	even	ok
	1	C23H21F3N2NaO3	453.1396	0.1	32.7	1	100.00	12.5	even	ok
469.1132	1	C23H16F3N4O4	469.1118	-3.0	36.0	1	61.57	16.5	even	ok
	2	C24H12F3N8	469.1132	-0.1	43.8	2	100.00	21.5	even	ok
	1	C19H17F3KN8O	469.1109	-4.9	4.7	1	32.23	13.5	even	ok
	2	C23H21F3KN2O3	469.1136	0.8	10.1	2	100.00	12.5	even	ok
475.1214	1	C18H14F3N10O3	475.1197	-3.5	2.3	1	49.58	15.5	even	ok

Supplementary information Chapter 5

Compound **5.1a** ^1H NMR





Supplementary information Chapter 5

Compound

230125-07ESI- HRMS-Moitessier

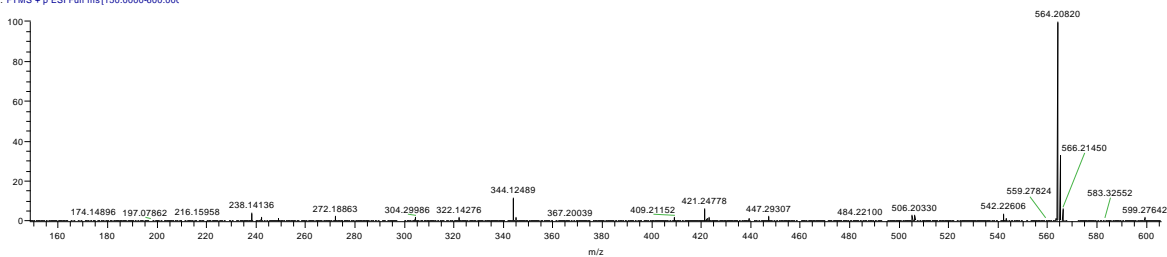
5.1a

HRMS

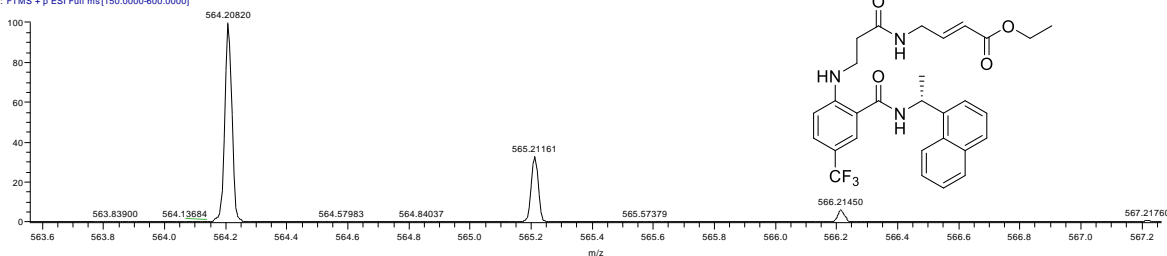
01/25/23 13:25:56

230125-07ESI- HRMS-Moitessier-Sharon Pin
T: FTMS + p ESI Full ms [150.0000-600.0000]

0.28-0.31 AV: 15 NL: 4.35E8



230125-07ESI- HRMS-Moitessier-Sharon Pinus-SP-319 #118-132 RT: 0.28-0.31 AV: 15 NL: 4.35E8
T: FTMS + p ESI Full ms [150.0000-600.0000]



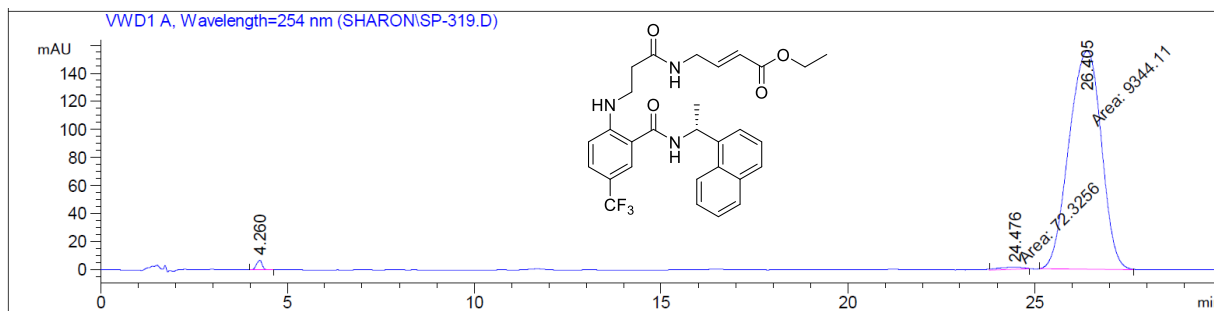
230125-07ESI- HRMS-Moitessier-Sharon Pinus-SP-319#118-132 RT: 0.28-0.31 AV: 15

T: FTMS + p ESI Full ms [150.0000-600.0000]

m/z = 564.17302-564.23129

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
564.20820	437236704.0	100.00	22745.77	1.00	564.20806	0.25	14.5	C ₂₉ H ₃₀ O ₄ N ₃ F ₃ Na

Compound 5.1a HPLC

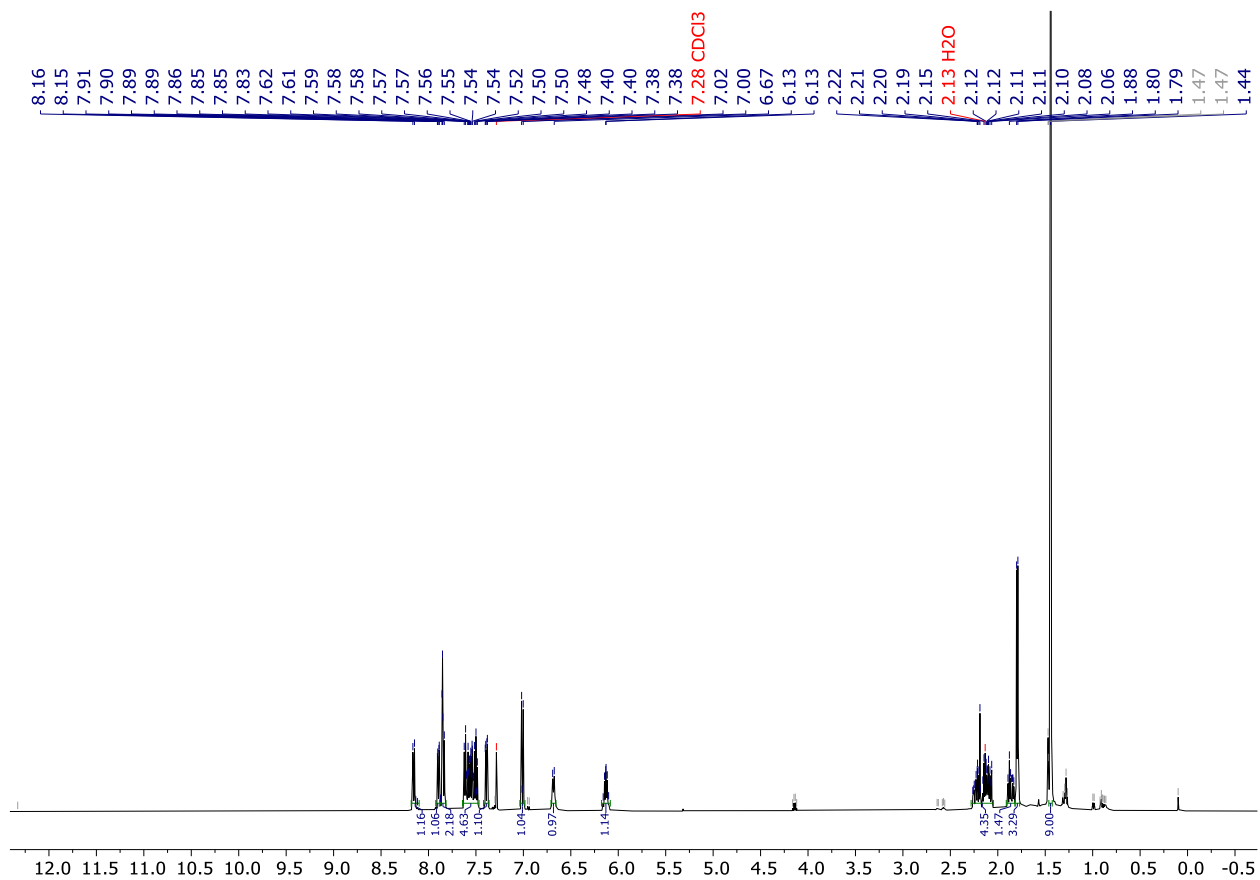
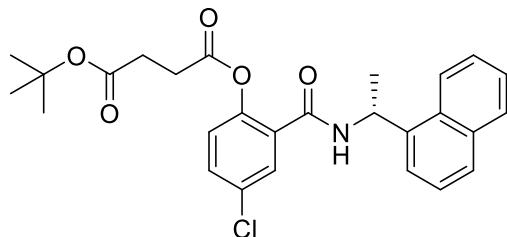


Signal 2: VWD1 A, Wavelength=254 nm

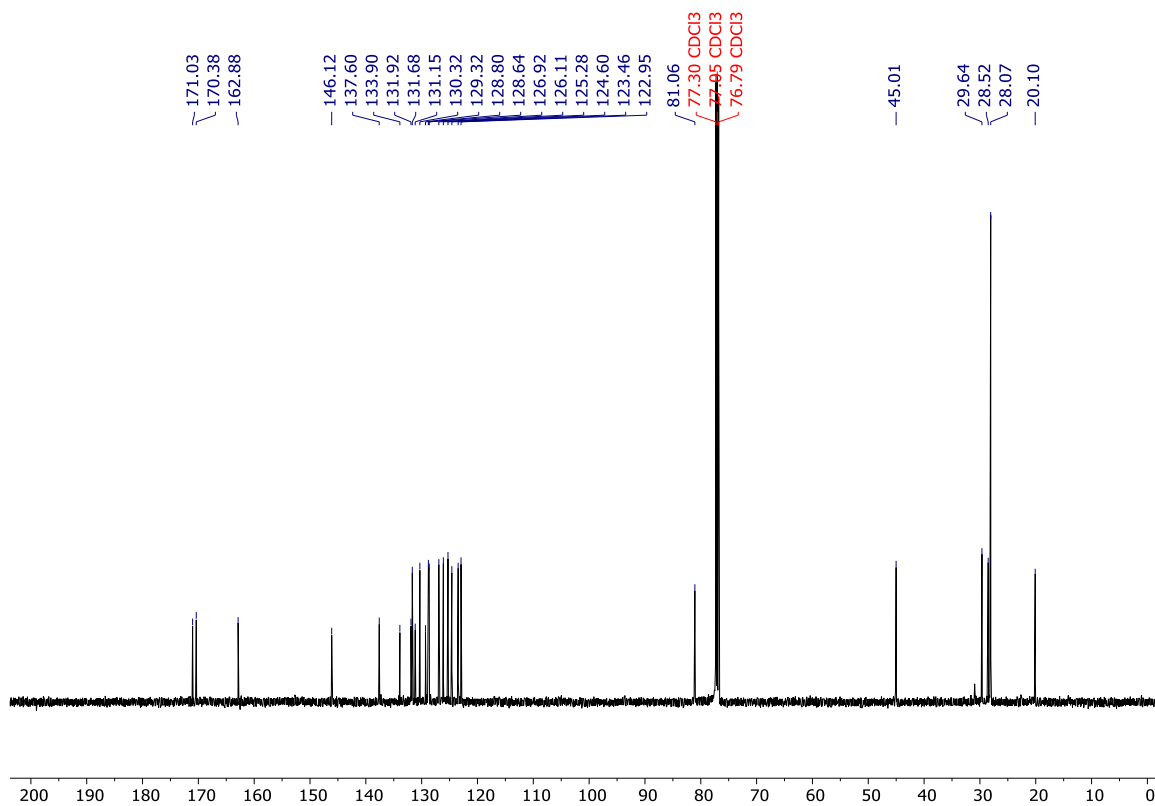
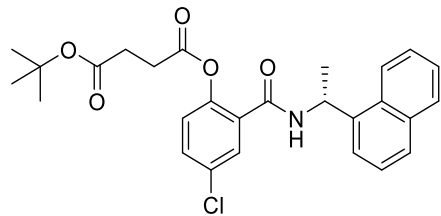
Peak #	RetTime [min]	Type	Width [min]	Area mAU *s	Height [mAU]	Area %
1	4.260	BB	0.1720	72.71426	6.92305	0.7663
2	24.476	MM	0.5737	72.32563	1.52139	0.7622
3	26.405	MM	1.0011	9344.10840	155.56308	98.4715

Supplementary information Chapter 5

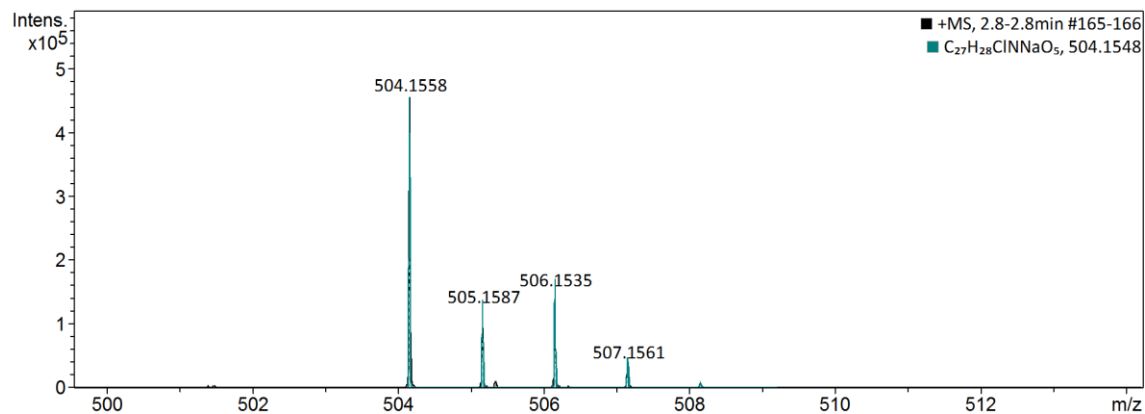
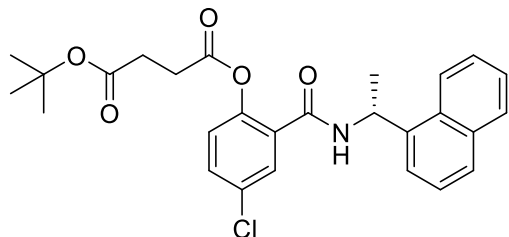
Compound 5.22: ^1H NMR (500 MHz, CDCl_3)



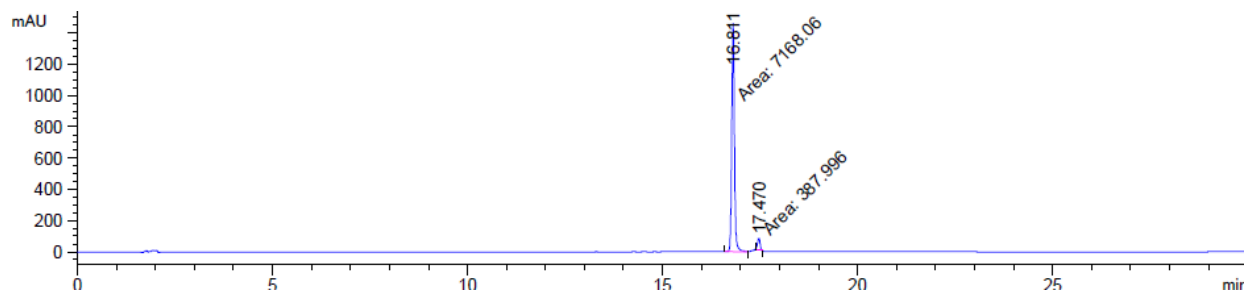
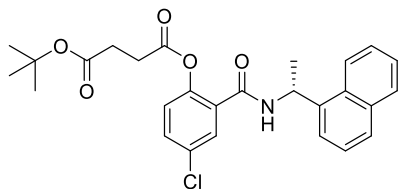
Compound 5.22: ^{13}C NMR (125 MHz, CDCl_3)



Compound 5.22: ESI(+) HRMS

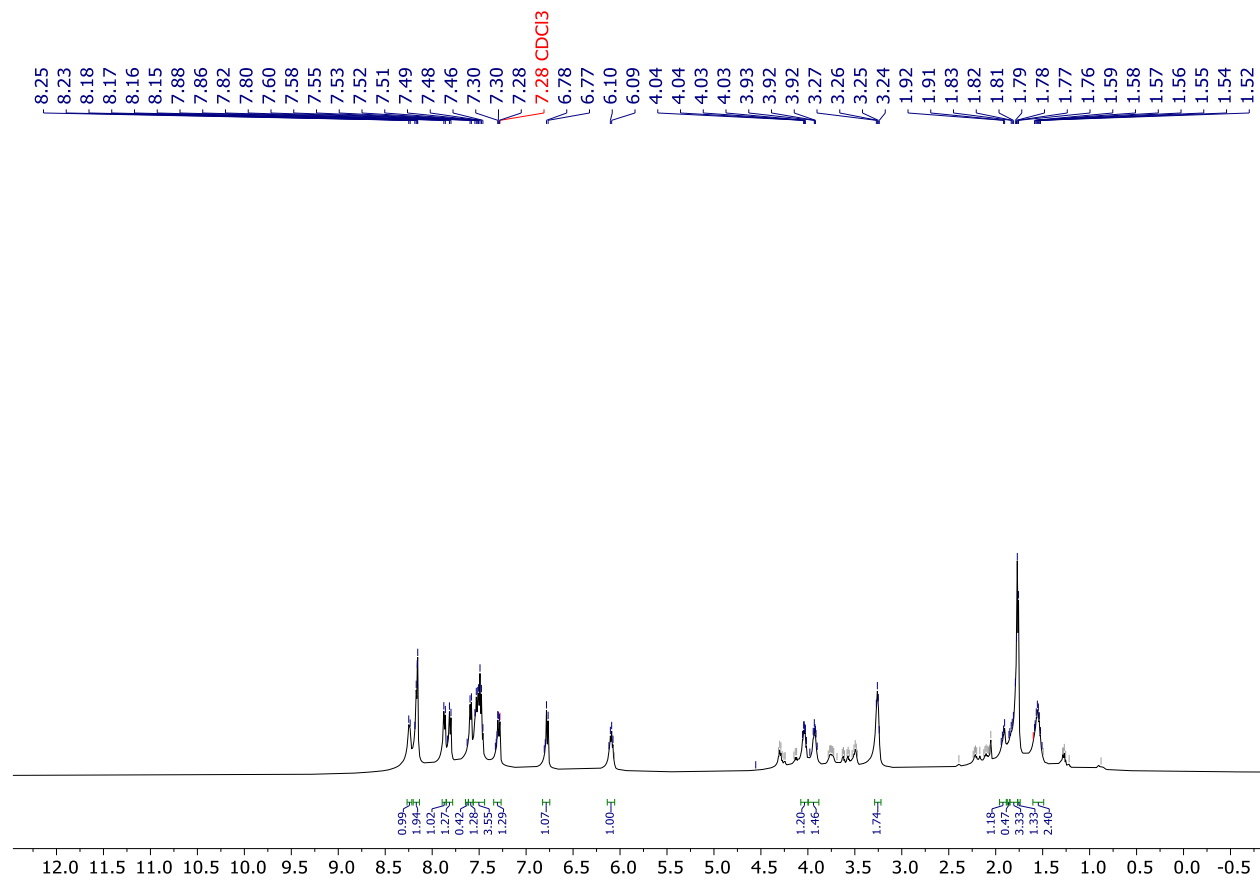
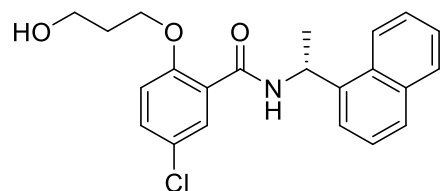


Compound 5.22: HPLC

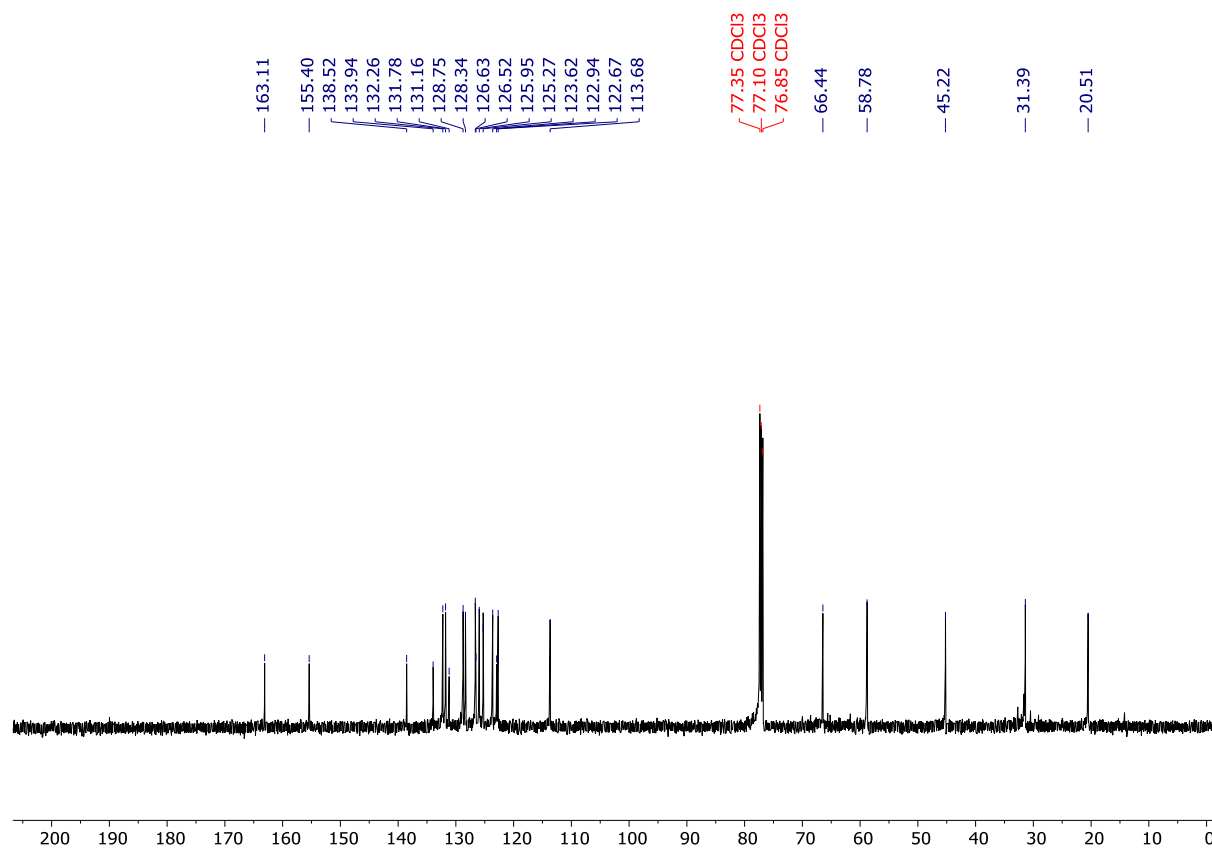
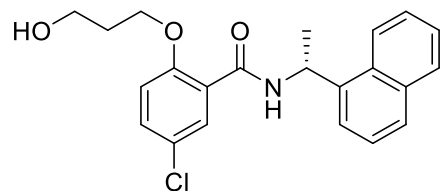


Peak #	RetTime [min]	Type	Width [min]	Area mAU*s	Height [mAU]	Area %
1	16.811	MM	0.0815	7168.05908	1466.30933	94.8651
2	17.470	MM	0.0832	387.99564	77.74462	5.1349

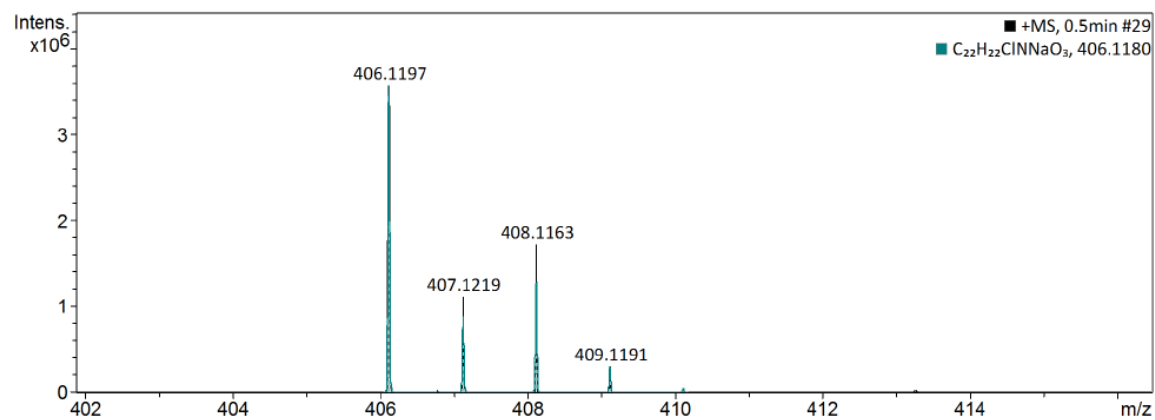
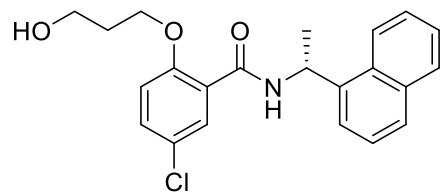
Compound 5.23: ^1H NMR (500 MHz, CDCl_3)



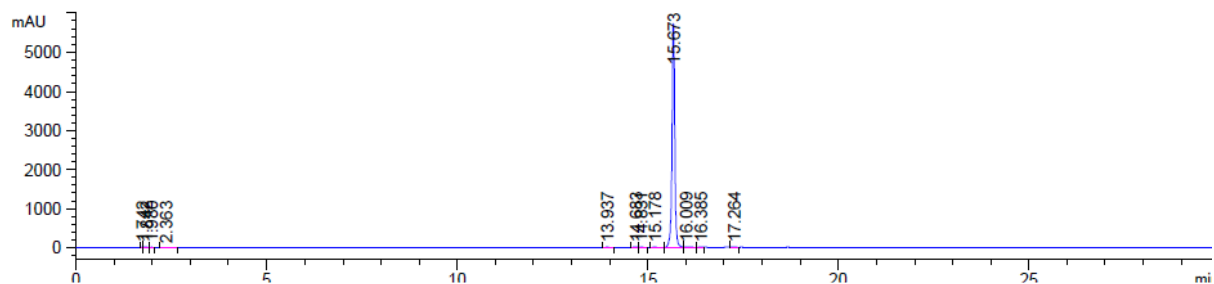
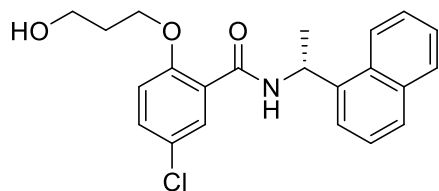
Compound 5.23: ^{13}C NMR (125 MHz, CDCl_3)



Compound 5.23: ESI(+) HRMS



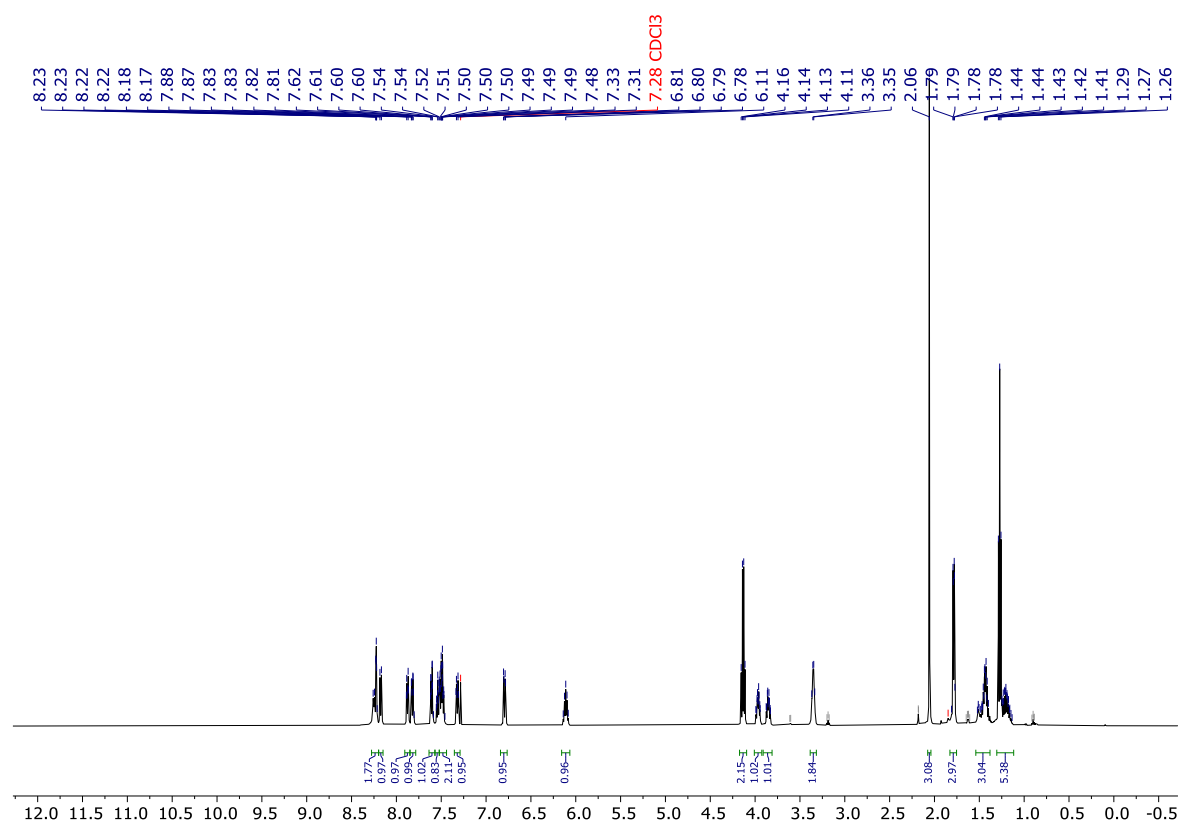
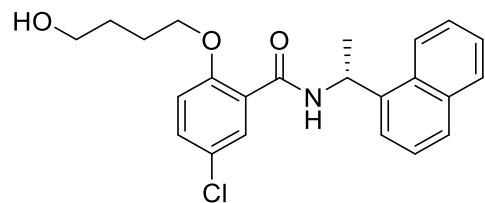
Compound 5.23: HPLC



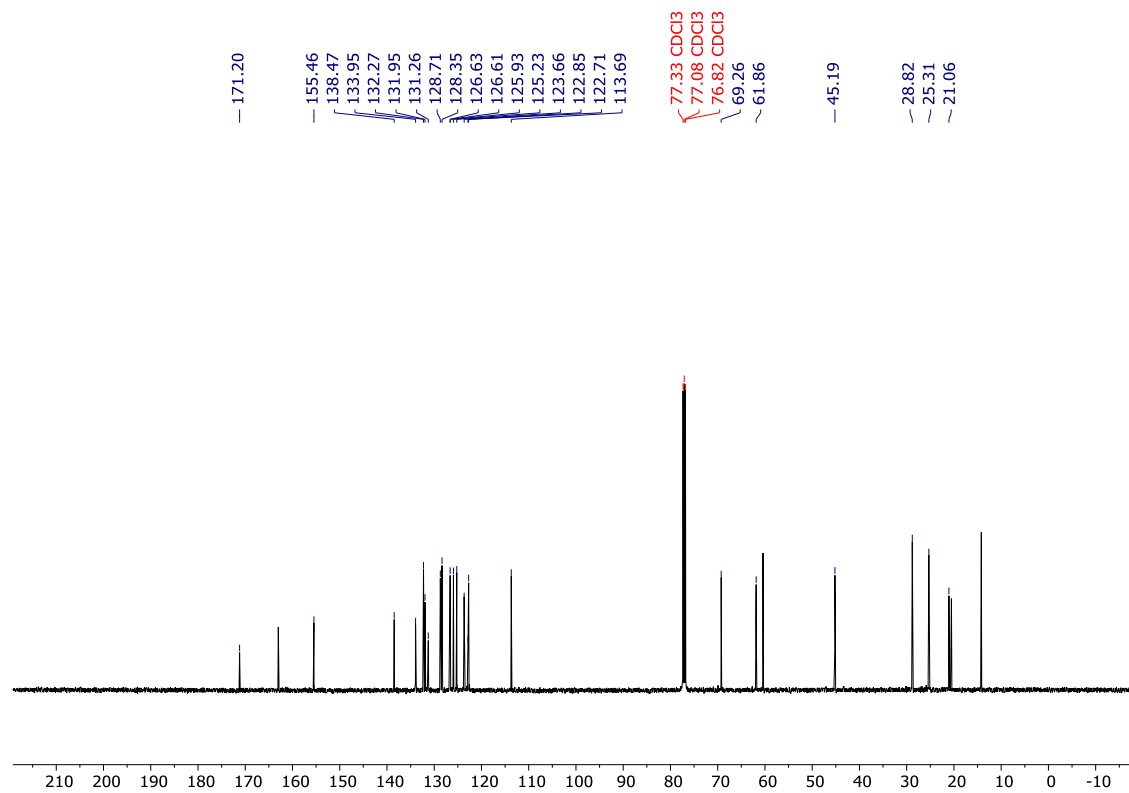
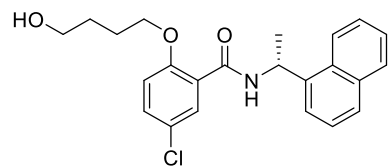
Peak #	RetTime [min]	Type	Width [min]	Area mAU	Area *s	Height [mAU]	Area %
1	1.742	BV	0.0459	21.19866		6.93027	0.0719
2	1.842	VV	0.0886	50.11782		7.74715	0.1699
3	1.980	VV	0.0733	20.22322		4.34039	0.0686
4	2.363	BB	0.0981	13.53119		2.15341	0.0459
5	13.937	BB	0.0742	14.55294		2.99071	0.0493
6	14.683	BV	0.0749	44.05370		8.95123	0.1494
7	14.831	VB	0.0721	17.37933		3.61762	0.0589
8	15.178	VB	0.0928	21.66652		3.42320	0.0735
9	15.673	BV	0.0767	2.91857e4		5744.89014	98.9648
10	16.009	VB	0.0898	60.02082		9.68371	0.2035
11	16.385	BV	0.0914	10.62652		1.74646	0.0360
12	17.264	VV	0.0857	31.91163		5.57636	0.1082

Supplementary information Chapter 5

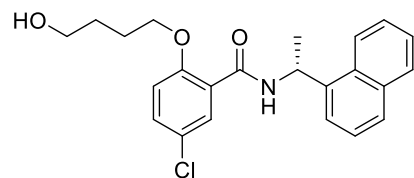
Compound 5.24: ^1H NMR (500 MHz, CDCl_3)

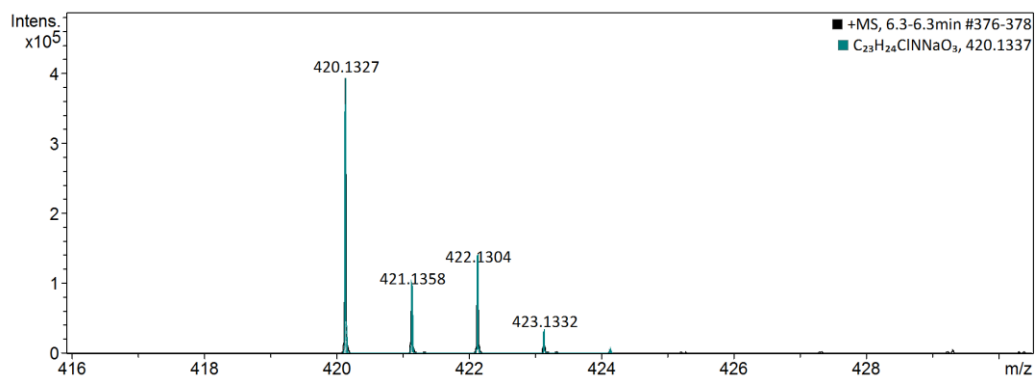


Compound 5.24: ^{13}C NMR (125 MHz, CDCl_3)

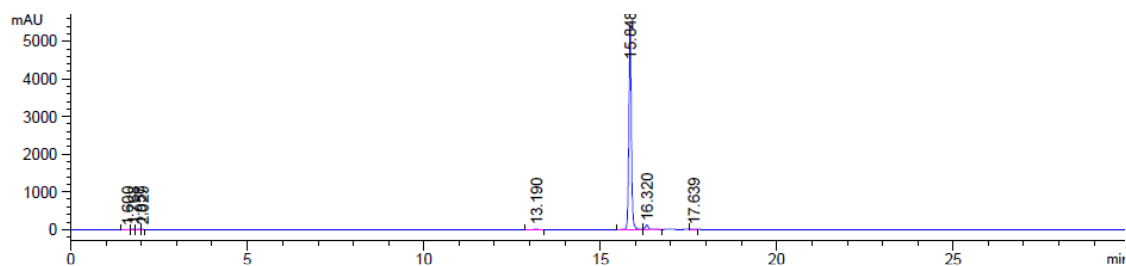
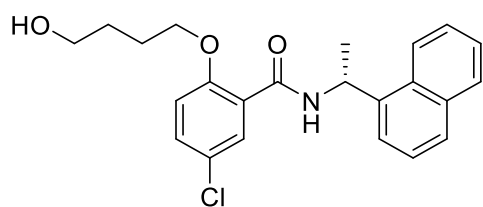


Compound 5.24: ESI(+) HRMS



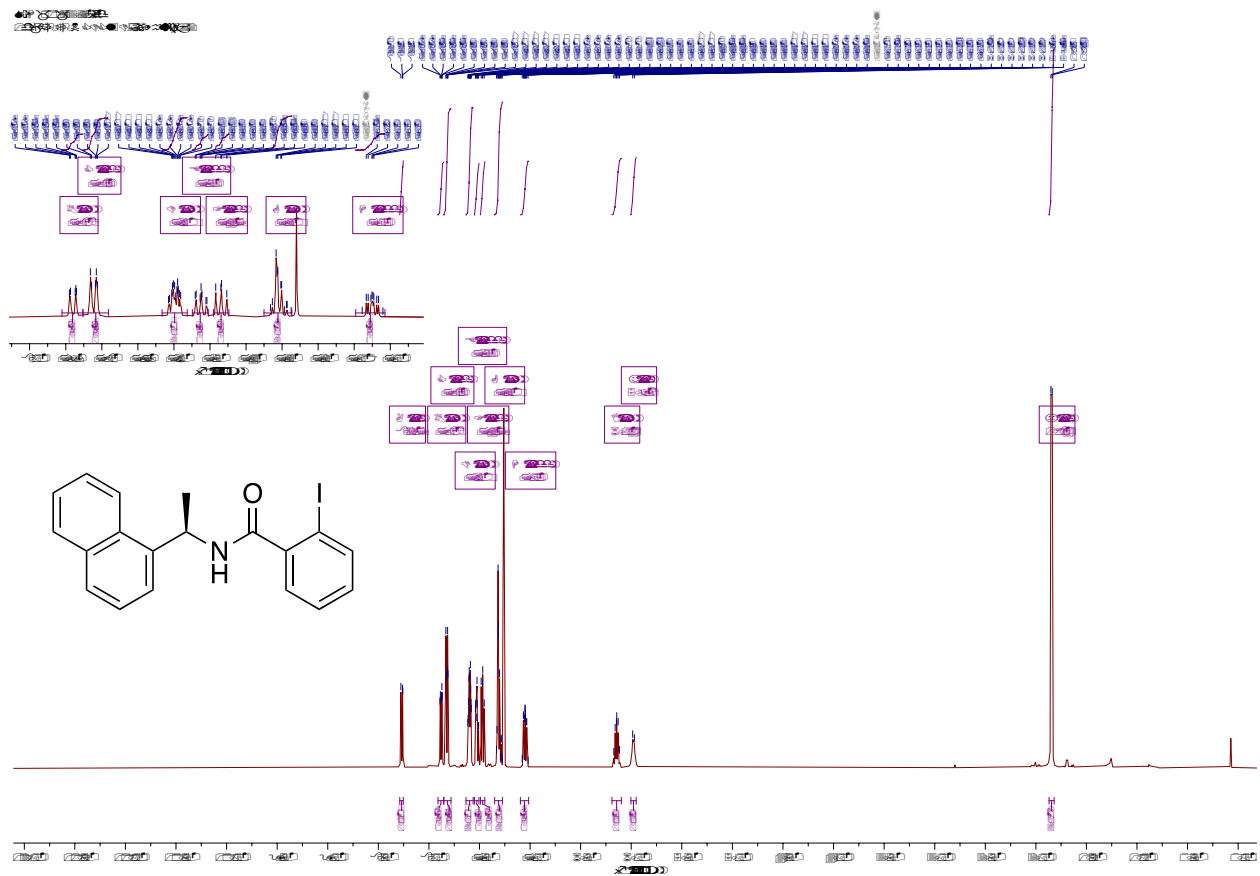


Compound 5.24: HPLC

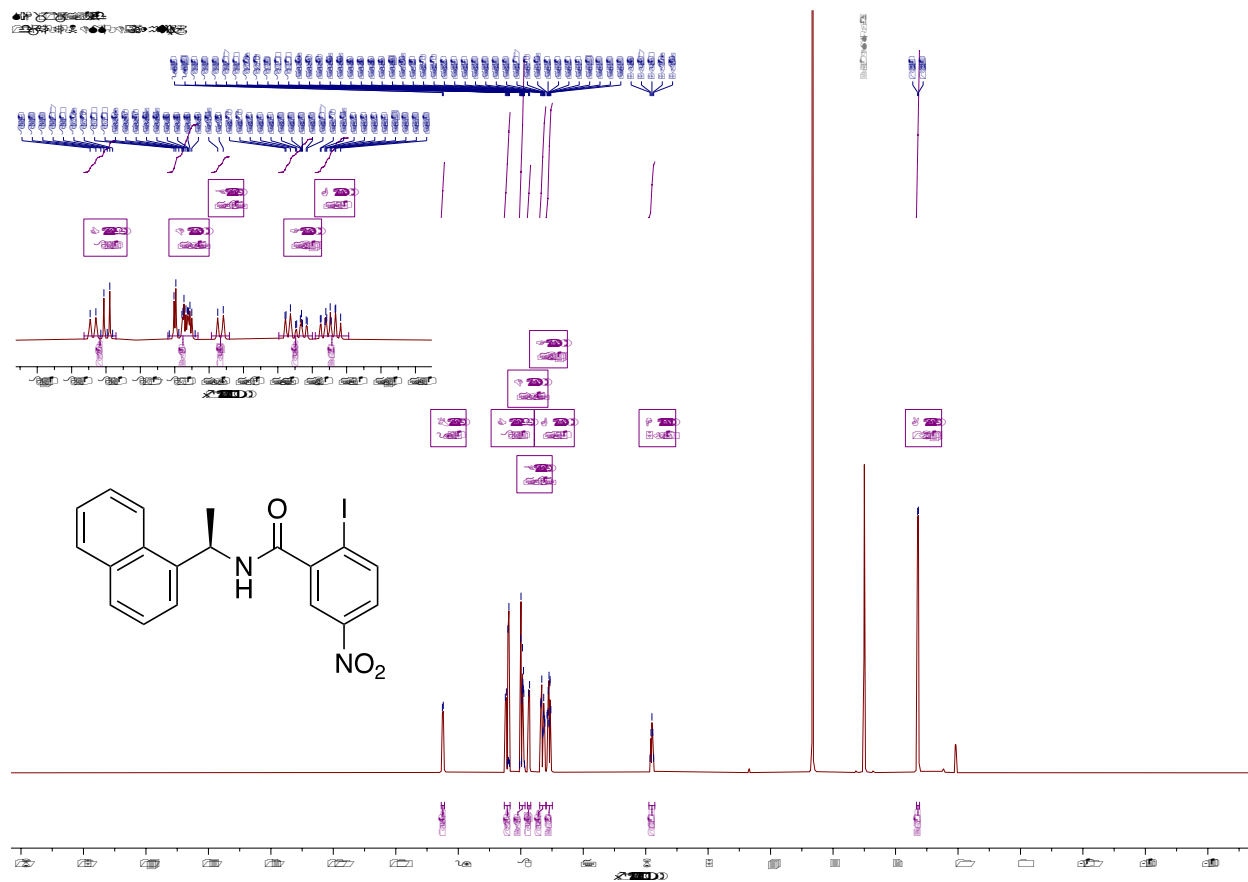


Peak #	RetTime [min]	Type	Width [min]	Area mAU	Height [mAU]	Area %
1	1.600	BB	0.1223	25.41055	3.13245	0.0924
2	1.768	BB	0.0732	36.00541	6.99340	0.1310
3	1.957	BV	0.0795	63.98516	11.75980	0.2328
4	2.029	VB	0.0575	38.19994	10.37355	0.1390
5	13.190	BV	0.0829	42.23029	7.53515	0.1536
6	15.848	BV	0.0706	2.65724e4	5534.49658	96.6755
7	16.320	VB	0.0866	683.00885	120.41161	2.4849
8	17.639	VV	0.0846	24.94331	4.53429	0.0907

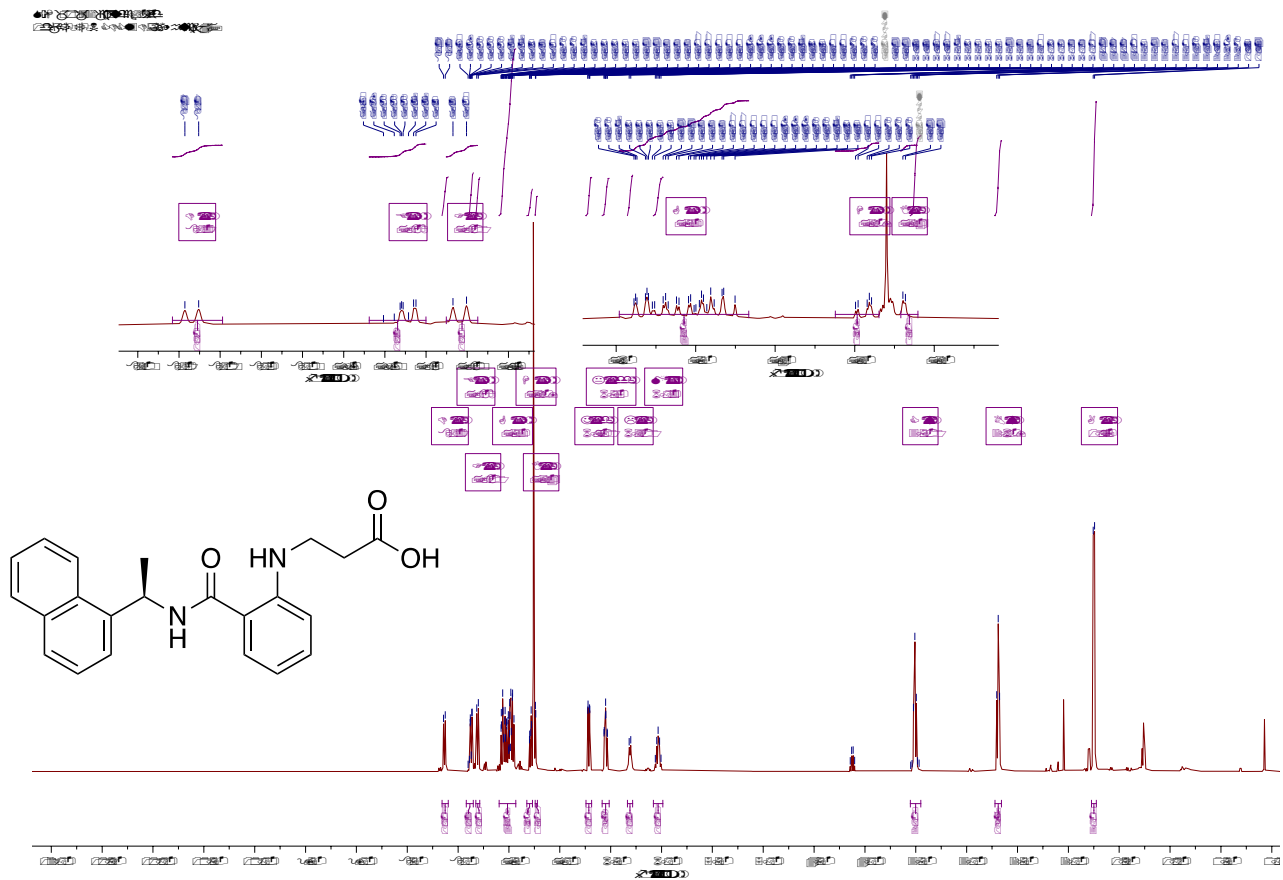
Compound 5.31 ¹H-NMR



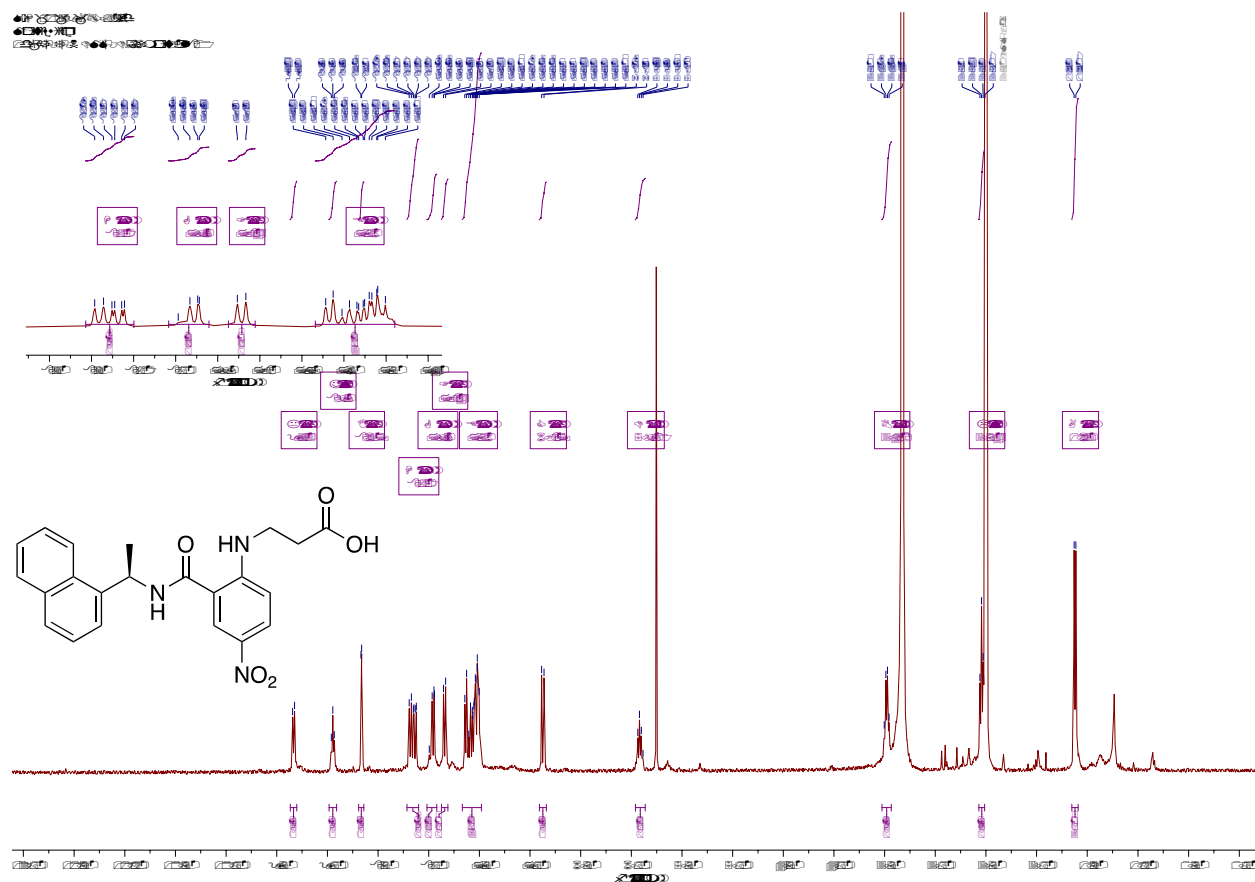
Compound 5.32 ¹H-NMR



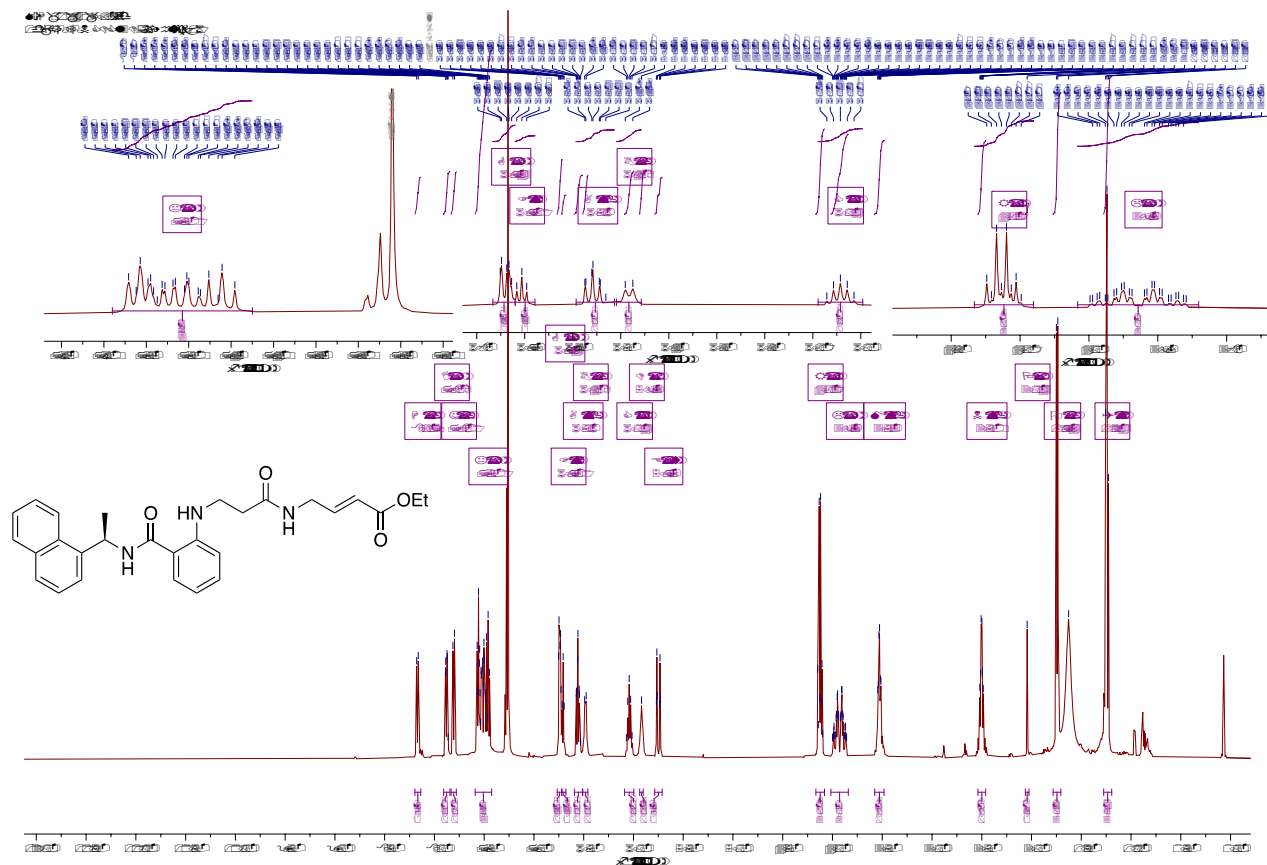
Compound 5.33 ¹H-NMR



Compound 5.34 ¹H-NMR

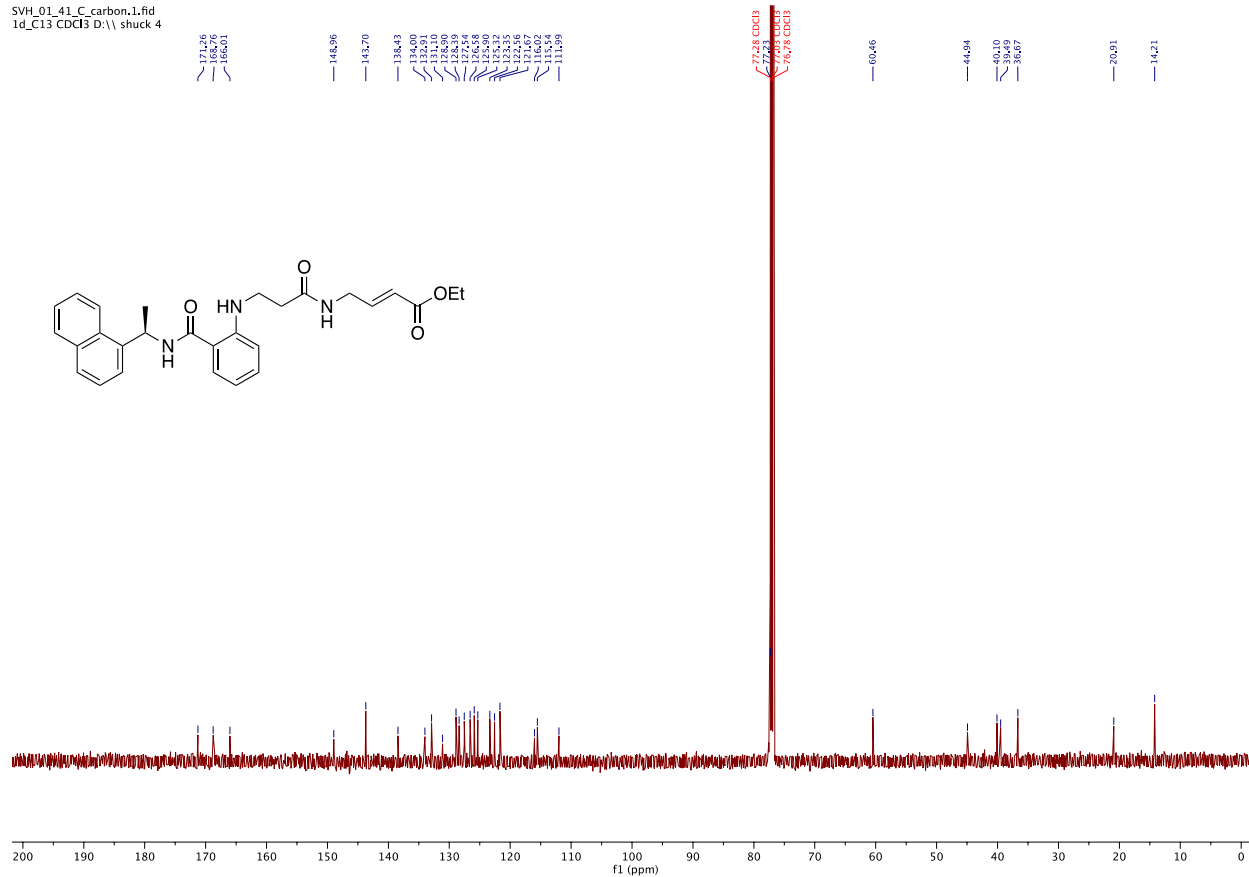


Compound 5.35 ¹H-NMR

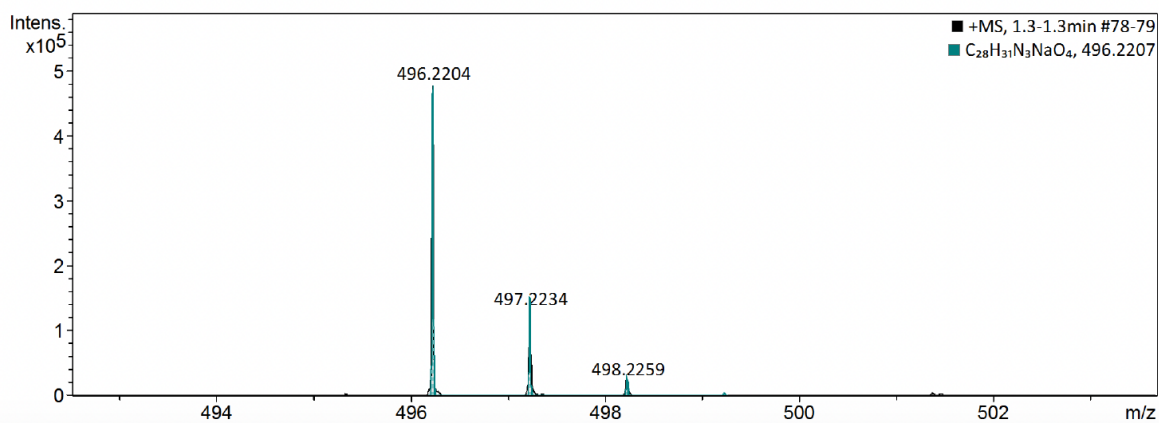
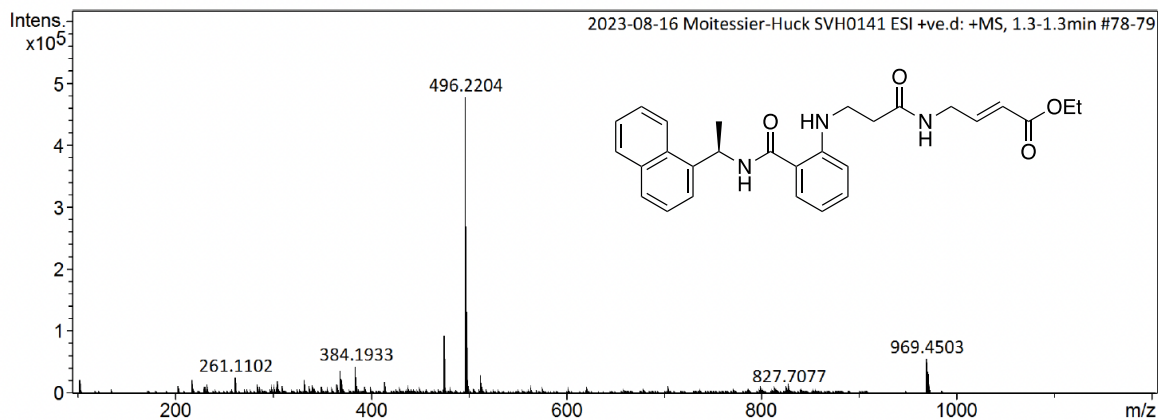


Compound 5.35 ¹³C-NMR

SVH_01_41_C_carbon.1.fid
1d_C13 CDCl3 D:\shuck 4

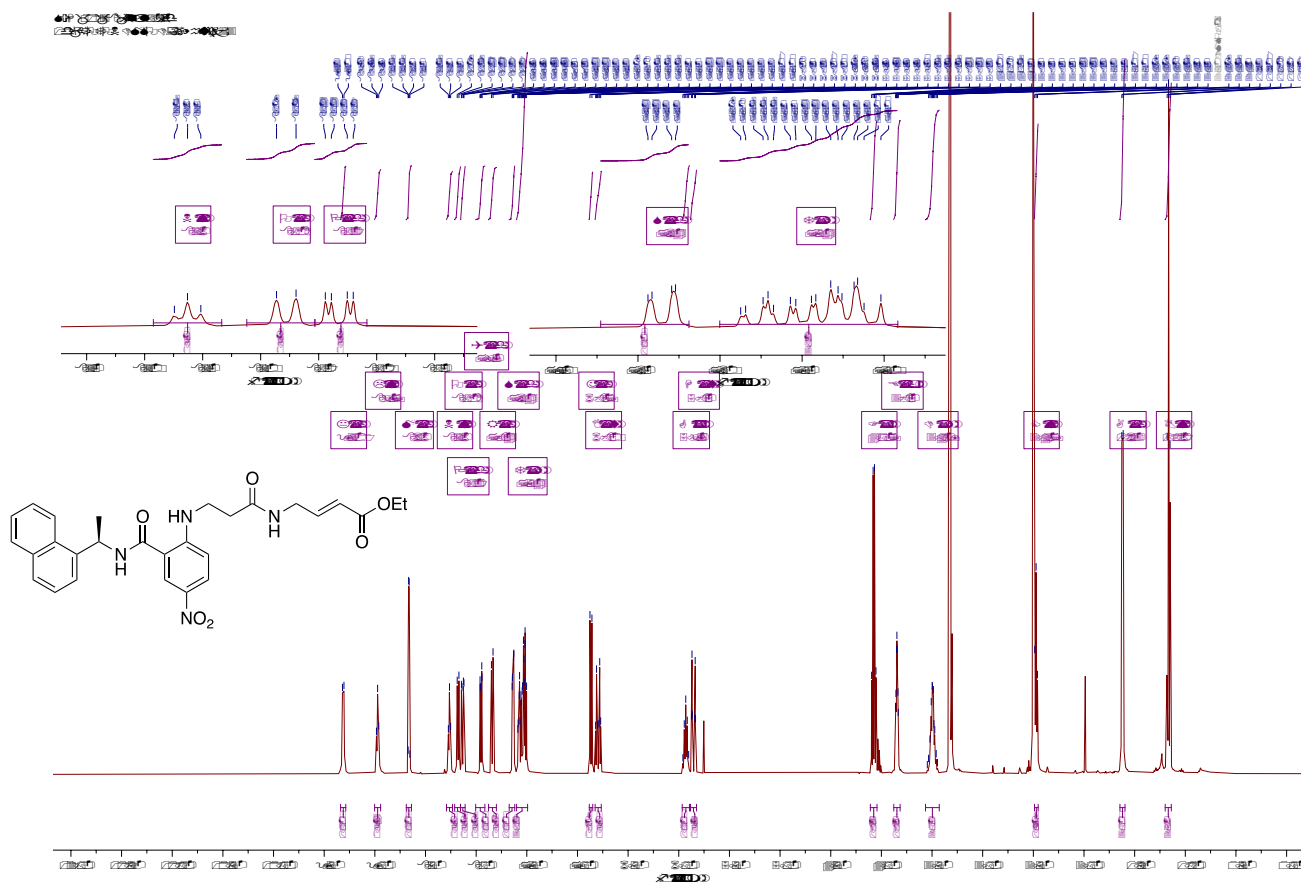


Compound 5.35 HRMS

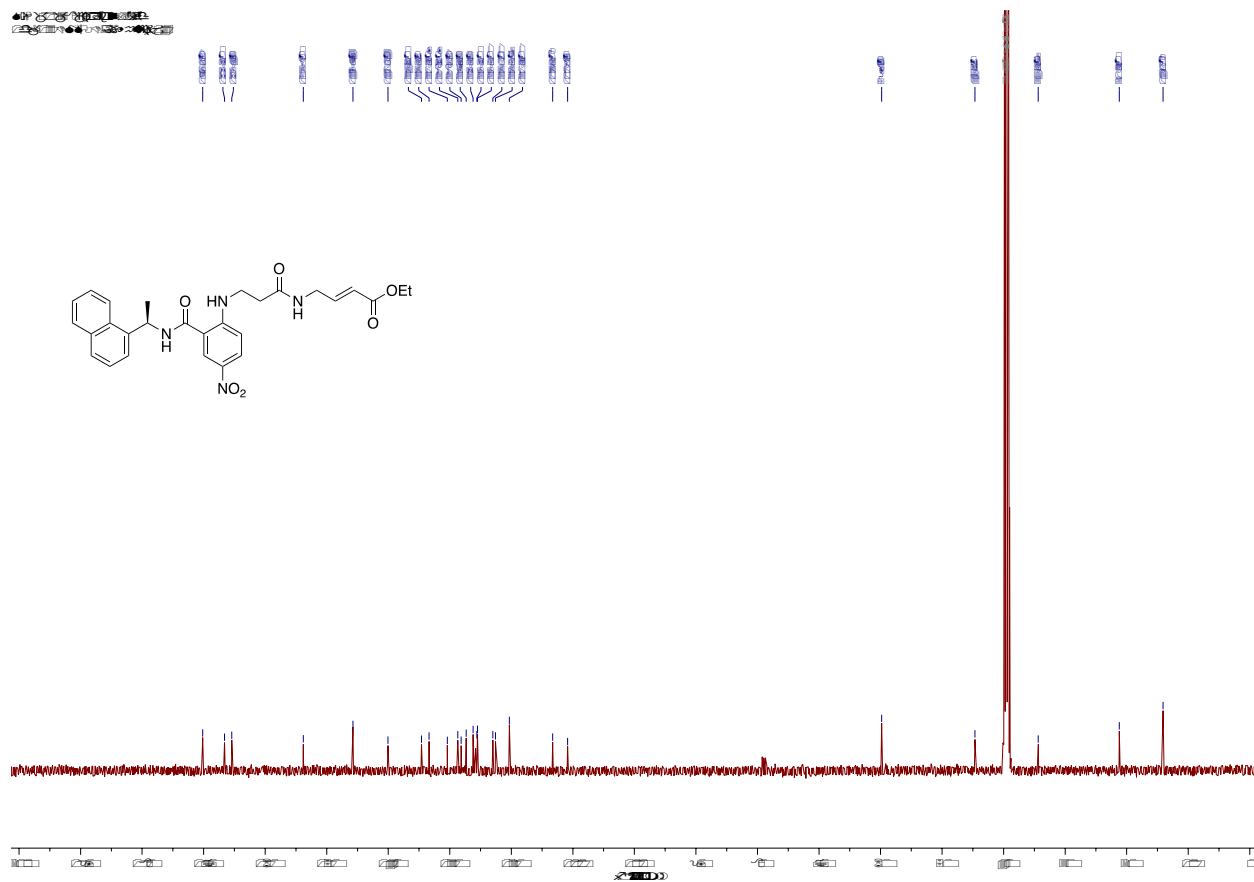


Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e ⁻ Conf	N-Rule
474.2382	1	C ₂₈ H ₃₂ N ₃ O ₄	474.2387	1.2	3.4	1	100.00	14.5	even	ok
	2	C ₂₆ H ₃₀ N ₆ O ₃	474.2374	-1.7	5.7	2	63.69	15.0	odd	ok
	1	C ₂₈ H ₃₅ NaO ₅	474.2377	-1.1	4.3	1	100.00	11.0	odd	ok
	2	C ₂₇ H ₂₉ N ₇ Na	474.2377	-1.1	6.0	2	72.71	16.5	even	ok
	3	C ₂₉ H ₃₁ N ₄ NaO	474.2390	1.7	10.0	3	57.80	16.0	odd	ok
496.2204	1	C ₂₇ H ₃₂ N ₂ O ₇	496.2204	-0.1	18.2	1	100.00	13.0	odd	ok
	2	C ₂₆ H ₂₆ N ₉ O ₂	496.2204	-0.1	26.7	2	61.98	18.5	even	ok
	1	C ₂₈ H ₃₁ N ₃ NaO ₄	496.2207	0.5	24.1	1	100.00	14.5	even	ok

Compound 5.36 ¹H-NMR



Compound 5.36 ¹³C-NMR



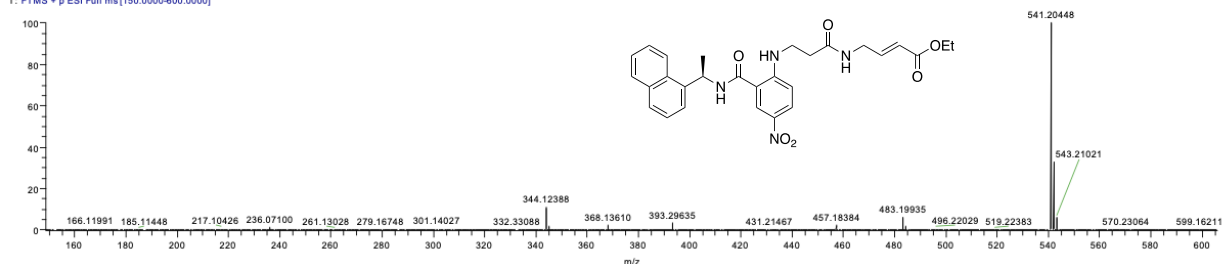
Supplementary information Chapter 5

Compound 5.36 HRMS

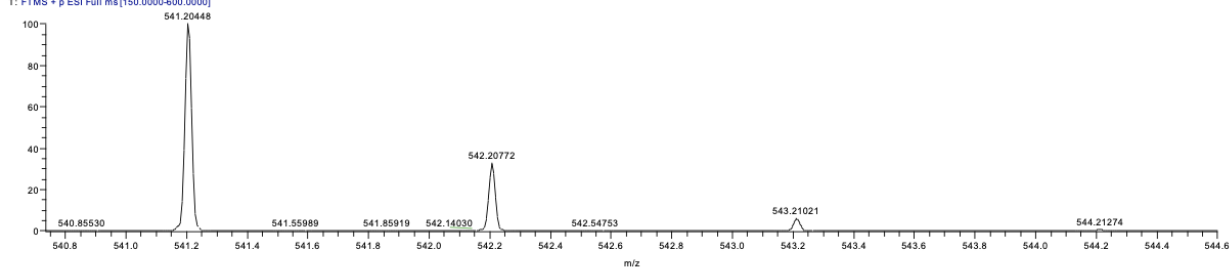
230804-07 ESI HRMS-Moiteissier-Huck So...

08/04/23 10:32:12

230804-07 ESI HRMS-Moiteissier-Huck Solene-SVH-01-38_230804103212 #217-223 RT: 0.52-0.53 AV: 4 NL: 6.15E8
T: FTMS + p ESI Full ms [150.0000-600.0000]



230804-07 ESI HRMS-Moiteissier-Huck Solene-SVH-01-38_230804103212 #217-223 RT: 0.52-0.53 AV: 4 NL: 6.15E8
T: FTMS + p ESI Full ms [150.0000-600.0000]



230804-07 ESI HRMS-Moiteissier-Huck Solene-SVH-01-38_230804103212 #217-220 RT: 0.52-0.53 AV: 4

T: FTMS + p ESI Full ms [150.0000-600.0000]

m/z = 541.17472-541.22943

m/z	Intensity	Relative	Resolution	Charge	Theo. Mass	Delta (ppm)	RDB equiv.	Composition
541.20448	623014848.0	100.00	22798.29	1.00	541.20576	-2.37	15.5	C ₂₈ H ₃₀ O ₆ N ₄ Na

